PROJECT AKARI

AKARI

知乎 @ 余命数（水无月零）

中学数学之旅

五卷合订本

Omnibus Edition

$$e^{i\pi} + 1 = 0$$

公开版 · 10 月 5 日

Project Akari

余命数

知乎@余命数

知乎@余命数

中学数学之旅

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

rikka.moe

知乎@余命数

余命数 (Midnight Forever)
知乎 @ 余命数
https://rikka.moe

知乎@余命数

中学数学之旅
An Excursion through Secondary School Mathematics
最后更新：October 7, 2026

知乎@余命数

© 2026 余命数
保留所有权利. 书稿版本，禁止修改、倒卖等任何侵犯作者著作权的行为.

知乎@余命数

知乎@余命数

书稿声明

Draft Notice

本书目前为**书稿状态**，意味着：

- 内容尚未完整，部分章节可能缺失或正在编写中
- 文中可能存在笔误、排版错误或数学推导疏漏
- 章节结构与内容安排可能在后续版本中调整
- 习题答案与索引等附录内容可能尚未完善

若您在阅读过程中发现任何错误或有改进建议，
欢迎通过以下方式联系作者：

个人网站：https://rikka.moe

电子邮箱：me@rikka.moe

书籍主页：https://rikka.moe/books

当前版本：Draft October 7, 2026

知乎@余命数

献词

Dedication

谨以此书

献给星街彗星

愿你如彗星划过星街，把梦唱到银河的尽头

每一个未曾舞蹈的日子
都是对生命的一种辜负

„Und verloren sei uns der Tag,
wo nicht ein Mal getanzt wurde!”

— 弗里德里希·尼采
Friedrich Nietzsche

— 2026 —

知乎@余命数

关于作者

余命数



Project Akari ·Midnight Forever

2025 届广西新高考二卷考生, 政史地组合, 数学 139 分.

这套书整理自我中学阶段学习数学时的笔记与讲义. 我按一条理想的阅读骨架把代数、离散与随机、几何、分析、线性结构依次铺开, 每条结论都给出推导路径, 把概念彼此生长的过程摊开在纸面上. 愿它陪你走完一段中学数学的旅程.

书籍主页

https://rikka.moe/books

更多信息、勘误表、联系方式等均可在此查看

交流与联系

在线平台

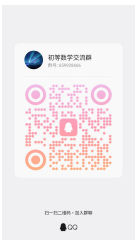
个人主页	`https://rikka.moe`
书籍主页	`https://rikka.moe/books`
知乎	@ 余命数
爱发电	`https://afdian.com/a/writeamathbook`

即时通讯

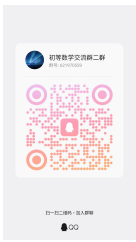
QQ	`2454167821`
交流群	`659928666`
邮箱	`me@rikka.moe`



扫码添加好友



交流群一



交流群二



爱发电赞助

知乎@余命数

欢迎通过以上方式与我交流讨论

知乎@余命数

前言

如果你基于本书制作修改版 PDF, 请保留这里的前言与声明, 谢谢. 想了解更多, 请访问我的主页https://rikka.moe

推动我写这套书的最主要动因是贫穷. 我高中就读于一所普通的省重点, 家里没钱给我补习, 能找到的老师水平也一般.

好在我有一台能联网的手机和一颗热爱学习的心, 于是便用互联网找学习资源, 听网课、看电子书.

这个过程里我攒下大量笔记, 大多记在 Obsidian 上. 高考结束后的 2025 年暑假, 我把这些笔记整理成册, 发到几个群里反响不错, 于是继续写了下去.

我也希望这本书能集百家之长. 市面上的资料和教辅很多, 我想读一本成体系、又贴合自己阅读习惯的书, 于是干脆自己动手写, 权当半个教科书.

这本书花了我不少时间. 如果你觉得有帮助, 欢迎去爱发电https://afdian.com/a/writeamathbook支持一下, 请我吃一碗拼好饭也可以. 书中的疑问或错误, 欢迎随时联系, 邮件 me@rikka.moe、QQ2454167821, 或加入 QQ 群 659928666.

知乎@余命数

吊诡的是, 我觉得教辅太贵, 所以都看电子版.

Obsidian 是一款很好用的笔记软件, 用好了可以打造属于自己的知识库.

关于联系与提问

本人经常在外打工兼职或忙别的事, QQ 私聊不一定看得到, 重要的事情请发知乎私信. 已经写清楚的问题我一般不再重复, 比如分卷本与合订本的区别、各版本之间的区别.

学习方法我也说不出什么有用的. 你观察过学得好的同学就会发现, 无论被谁请上台分享经验, 大多只能说些政治正确的话, 扎实基础、回归课本、多跟老师. 这怪不得他们, 真正掌握一项技能后, 就很难想象不懂它是什么感觉, 一项已经内化的能力, 很难再拆开讲给你听.

即便方法就是“极度的自律、极高的抗挫折能力、成百上千小时的枯燥重复”, 即便有些人同时有智商与基因、家庭托底和超强的专注力与抗挫折力, 学习也没有廉价速成的魔法. 每次有人问我, 我也只能给出和所有人一样的回答, 巩固基础、回归课本、多跟老师、总结错题.

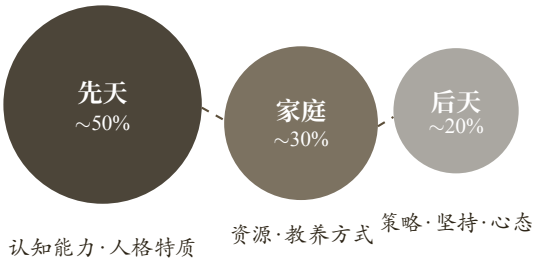
用卖课的人的话说, 就是欢迎多多支持我的项目和电子书. 所以这方面的问题, 就不必再问了.

不过, 考虑到还是有人想知道学习方法. 很好, 下面从比较科学的角度出发, 拒绝政治正确和哔哩哔哩经典学长或学姐分享学习方法的套话, 来说说好的学习方法.

学习成绩的好坏可以认为取决于三大类因素: 先天禀赋、家庭条件和后天努力. 下图用圆的大小表示各自的影响占比, 圆越大代表该因素对成绩差异的解释份额越多.

占比数据综合自 Silventoinen et al. (2020) 在 Scientific Reports 上发表的 28 国双胞胎整合分析(遗传力约 40%–70%, 共享环境约 20%–40%)与 de Zeeuw et al. (2015) 的荷兰双胞胎元分析; 本文取中位数附近的近似值.

知乎@余命数



知乎@余命数

图中三个圆分别代表：

- **先天**，约占一半：认知加工速度、工作记忆容量等人脑硬件参数，以及尽责性、抗压性等人格倾向. 这部分相对稳定，但并非锁死. 早期营养、睡眠和运动会影响其发挥. 有些人的确在某些方面的所谓的感觉更好，又或者说，假如你在婴幼儿时期，甚至还没出生的时候，就接受到了一些对神经系统不太好的物质，这都是实打实的.
- **家庭**，约占三成：家里的藏书量、父母的词汇量和教育期望、能否提供安静的学习环境与课外资源. 家庭条件好在早期尤为关键，随着年级升高其权重逐渐下降. 这里也比较显然，如果家庭是那种书香家庭，父母都是文化人，或者都是中产阶级甚至是富人阶级，请得起清华北大的私教一对一，买学区房、上重高等等，获取好的成绩的概率更大、更轻松. 如果家庭是那种一天到晚都在吵架、家庭暴力、父母没文化而且也不怎么有钱（遗憾的是编者的家庭都是如此，因此我付出了很大的努力），很大概率最后都是成绩一般甚至很差，成为混混或其他.
- **后天**，约占两成：学习方法是否科学、能不能坚持、面对难题时的心态. 虽然占比最小，却是个人最能把控的部分，也是拉开同等条件下人与人差距的主战场. 这也就是唯一可以努力的地方，除非你是家长，否则对于学生，也就是我们而言，只有此路！

虚线表示三类因素之间存在交互，即先天禀赋需要在合适的家庭环境中才能充分发挥，而后天努力的效率又受限于先天的认知基础和家庭提供的资源.

把可控因素提炼出来，成绩的近似模型可以写成：

$$\text{成绩} \approx \underbrace{(\text{认知基础} + \text{尽责性})}_{\text{较稳定}} \times \underbrace{(\text{科学策略} + \text{自我调节})}_{\text{高度可控}} + \underbrace{\text{家庭支持} + \text{教学质量}}_{\text{环境系数}} \tag{0-1}$$

说说，公式里用的是乘法和加法，含义各有侧重.“较稳定”部分与“高度可控”部分之间是乘号，即便先天基础只有六七分，只要方法和自律跟上，乘积依然可观；反过来基础再好，方法为零则乘积也为零. 加在最外面的“环境系数”是加分项，好的家庭和老师能把整体水平往上抬一截，但它们替代不了主体部分的乘积.

Poropat (2009) 的元分析，样本逾七万人，显示尽责性是大五人格中对学业成绩最强的非智力预测因子，其对大学成绩的预测效力与智力相当.

ACE 模型源自行为遗传学经典范式：A 为加性遗传，C 为共享环境，E 为非共享环境.Lee et al. (2018) 在 Nature Genetics 上发表了百万级样本的 GWAS，其多基因评分解释受教育年限约 11%–13% 的变异.

Hart & Risley (1995) 在 Meaningful Differences 中跟踪观测了 42 户家庭三年，发现专业人员家庭幼儿在三岁前累计听到约 3000 万词，领取福利金的家庭约 1000 万词，“三千万词鸿沟”由此得名；后续研究对其取样和方法有争议，但方向一致.

Duckworth & Seligman (2005) 在 Psychological Science 上对八年级生的纵向研究中发现，自控力对期末成绩的方差解释量约为 IQ 的两倍有余.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

“科学策略”指的是有实证支撑的学习技术. 合上书本逼自己回忆的主动检索, 以及在快要忘记时复习的间隔重复, 其效果远大于被动重读和划重点.“自我调节”涵盖制定计划、监控进度和在出错时及时纠正方向的能力.

先天和家庭设定了起跑线和跑道状况, 后天的高质量努力决定你能在这条道上跑多远.

所以, 动用你聪明的小脑袋瓜, 评估自身的情况, 决定从何入手, 别去问别人的学习方法. 你首先是你自己的, 如果你自己都不知道自己要做什么, 就不要偷懒问别人了.Like WTF, 如果你是个普通家庭, 你去问一个有钱人出身的竞赛背景的学霸, 问学习方法, 这本身就荒谬.

当然, 还有一个很好的学习习惯, 那就是去爱发电支持作者:), 推出更多的学习资源, 助你一臂之力.

Dunlosky et al. (2013) 在 Psychological Science in the Public Interest 上评定了十种常用学习技术, practice testing 与 distributed practice 获评高效用, rereading 与 highlighting 获评低效用.

知乎@余命数

如何获取无水印版

本书的基础版本， 含水印免费开放， 任何人都可以在我的主页`https://rikka.moe/books`下载. 如果你想获取无水印版, 可以在爱发电上支持本项目:

1. 打开爱发电页面:`https://afdian.com/a/writeamathbook`
2. 选择对应的赞助档位(通常会标注“无水印版”或“完整版”), 完成赞助
3. 赞助后在爱发电的消息里会收到无水印版 PDF 的下载链接; 如果没有收到, 请发邮件或在知乎私信告诉我你的爱发电昵称

无水印版与免费下载版的**内容完全一致**, 差别只在水印. 你不花钱也能读到一模一样的数学内容;付费只是为了去掉水印、支持项目运转.

既然可以免费下到, 为什么要花钱买?

我一直用中世纪或文艺复兴时期王公贵族资助艺术家来比喻这种捐赠的模式, 或者更恰当的, 是 18-19 世纪文人的集资预订, 又或者是数字公共品理论。

许多高中生(包括我)因为高中的时候穷, 没钱, 所以喜欢白漂电子书资料。毕竟在我那时候, 任何没有 DRM(数字版权管理技术)的电子书都能被随意复制粘贴和传播, 盗版网课也有很多。很多时候不是不想支持正版, 而是没钱。与其到处找资源, 不如加成千上万个资料群, 去点那些夸克网盘(天下没有免费的午餐, 夸克网盘对分享资料的人也是有利可图的), 下一堆资料。虽然下了或者存了以后, 可能没有时间看或者基本不看, 但至少满足了收集欲, 以及一种“我在学习”的感觉。

但是, 与其在网盘里存 500G 的名师网课和状元笔记, 或者花重金买各种名师的网课和资料, 不如把钱留给真正对你有用的东西。如果你觉得 “这本书确实对我有用”、写得好, 或者在其他方面给你带来了价值, 不妨去我的爱发电`https://afdian.com/a/writeamathbook`花 10～30 块来支持我, 并获得无水印版。

其实 10 块钱，吃一次拼好饭差不多就花完了。所以我想了一句广告语：用一次拼好饭的钱，来支持下一位创作者。即使实际上你可以白漂（毕竟这是公共品，一个人的占用不会导致另一个人占用的减少），但你的支持，能让互联网上的创作者感到生活更加美好。

更重要的是，你今天怎么对待创作者，就是你希望将来别人怎么对待你。总有一天，你也会在网上创作：可能是一篇文章，一门课。到那时你会明白，被人认可的感觉有多重要，哪怕对方只付了一次拼好饭的钱。

公共品之所以能一直存在，是因为总有一些人，在可以不付钱的时候，仍然愿意为自己认可的东西付一点点。创作者因此能继续创作，受众也因此能继续读到好东西，这才是真正的双赢。而你今天的这一点点，就是在把这个循环往下传。

知乎 知乎@余命数

试用 Akari Web 平台

知乎@余命数

除了纸质书之外，Project Akari 还在做一套在线学习工具，即 Akari Web 平台（`https://kirie.io`）.

Project Akari 是一项旨在结合人工智能教育、教育数据挖掘、学习分析，并基于认知心理学的研究成果，搭建一款旨在帮助学生高效地学习数学和计算机科学方面知识的智能辅导系统，或称为学习引擎.

余命数 知乎@余命数 知乎@余命数

因此，我们的目的就是，尽可能量化学生的知识状态，并结合算法，智能评价和判断学生目前的学习状态，辅助其真正习得和掌握一样东西.

我们的哲学是，让用户用效率最高的方式学到知识，为此我们开发了一整套配合的算法，并在课程等设计上尝试最小化认知负担，采取更互动的方式以激发主动性. 量化学习的一切，并且优化，便是一切的出发点和落脚点.

知乎@余命数

这本书写给谁

知乎@余命数

本书的目标读者有三类：

- **在校高中生**：跟着课堂教学同步使用，课上没听懂的章节回来对着书再看一遍推导. 标题带 $*$ 的选读内容可以先跳过，等基础扎实了再回头啃.
- **复读生或自学备考者**：按目录顺序系统过一遍，重点关注每节的例题和章末习题. 如果你的时间紧张，优先读函数、三角、数列和解三角形四章，即高考分值密集的区域.
- **想重温数学的成年人**：你可能忘了大半高中内容但又不想从初中补起. 挑感兴趣的章节直接读就好，遇到前置知识不清楚的地方顺着交叉引用回去翻.

不管你是哪种读者，建议的做法都一样，先看正文搞懂逻辑链，再做例题（盖上解答自己想），最后做章末习题检验. 光看不练等于白看.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

本书的结构

各卷可独立阅读，也可合订通读. 分卷本在文前附有本卷的分册前言与本卷导读；合订本在每卷分隔页之后给出本卷导读，便于按卷回看. 标题中带 * 的章节多为选读或拓展，初读时可略，需要时再补.

关于版本与系列

本书目前仍是书稿，个别地方尚未完整，可能存在笔误或疏漏，若发现错误或有建议，欢迎联系作者. 本版本是 2026 年按学习路径重排的重要更新，调整了五卷章序，把数列章尾的分析内容并入极限与导数，把平面向量提到解三角形之前.

本书属于 Project Akari 系列.Akari 取自日文“あかり”，意为光，寓意是让好的学习材料更容易被找到、更容易被用起来.

Project Akari 的另一部书稿《函数与导数》将函数、极限与导数相关内容单独成书，系统展开. 两本书可以独立阅读，也可以互相参照. 配套的 Akari Web 平台（`https://kirie.io`）提供笔记、记忆卡片和在线练习，与书稿配合使用.

本系列的基础版本免费开放，完整版本以极低价格提供. 知识不应该是奢侈品，这是我做这件事的底线.

从写作、排版到软件开发，Project Akari 的每一步都是我一个人完成的. 服务器和运营成本靠书的销售收入以及我个人的工作收入来覆盖. 这个项目既是学习方法的实践，也是工程能力的验证.

发现错误怎么办

一个人的精力和眼力终究有限，书里难免有笔误、推导疏漏乃至概念性错误. 如果你发现了，请按以下方式反馈，我能更快地定位和修正：

1. 说明在哪一卷、哪一章、哪一节，或直接给页码和小标题
2. 引用书中的原文或公式编号
3. 写出你认为正确的版本或质疑的理由

最快的渠道是知乎私信 `https://www.zhihu.com/people/yushumi-shu`，其次是邮件 `me@rikka.moe`. 有价值纠错的读者名字会列入后续版本的鸣谢名单.

致谢

这本书能走到今天，离不开很多人的帮助.

首先要感谢每一位读过草稿、提出意见的同学和朋友. 你们在群里随手指出的一道错题、一处不通顺的表达，聚沙成塔地把这本书往前推了一步.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

感谢爱发电上的赞助者们. 你们的每一笔支持都实实在在地变成了服务器费用和我碗里的拼好饭, 让我能在打工之余腾出手来继续写.

感谢家人容忍我把无数个周末和假期耗在了电脑屏幕前. 虽然你们不完全理解我在做什么, 但从没拦过我.

最后感谢你, 正在读这段字的陌生人. 一本书的价值在被翻开的那一刻才算兑现.

版权与授权

本系列的基础版本依据**知识共享署名-非商业性使用-禁止演绎 4.0 国际协议**(CC BY-NC-ND 4.0) 开放给个人学习者免费使用. 完整版本以极低价格提供, 销售收入用于覆盖服务器与运营成本. 商业使用、机构采购或翻译授权请联系作者.

防盗版与使用声明

本书为 Project Akari 系列出版物. 本 PDF 仅供合法获取者个人学习使用, 禁止以下行为:

- 将本 PDF 或其任何部分上传至任何公开网络平台、网盘、论坛或即时通讯群组;
- 对本 PDF 进行二次售卖、捆绑销售或用于任何商业用途(已获作者书面授权者除外);
- 将本书内容用作机器学习或人工智能训练语料;
- 移除、涂改或遮盖本书的水印、版权信息、读者指纹或页脚标识;
- 修改 PDF 元数据后以原书名、作者名重新分发.

如发现盗版或侵权行为, 欢迎通过以下渠道举报:

- 书籍主页:https://rikka.moe/books
- 邮箱:me@rikka.moe

版本指纹(用于溯源与正版核验):

读者标识 DRAFT-REVIEW

文档哈希 DRAFT-REVIEW

编译日期 October 7, 2026

关于修改与再分发的声明

本书由作者独自维护并发布, 是本项目唯一的**权威正版来源**. 任何对本 PDF 或其源文件的内容、水印、版权信息、读者指纹、PDF 元数据的删改、涂改或遮盖, 均视为对本声明与著作权法的违反.

修改后的版本**不得**保留以下任一标识并以任何形式分发, 否则视为侵权:

- 书名《中学数学之旅》及 Project Akari 系列名称;
- 作者署名"余命数"及相关社交账号;
- 版权页、前言中任何声明性文字;
- 页脚隐性指纹与页面水印.

欢迎读者基于本书内容做学习笔记与摘录，但请在分发时注明来源与原作者. 如需做系统性改编，如翻译、缩写、教材化使用，请先联系作者取得书面授权.

正版核验与联系方式见版权页，或访问 https://rikka.moe/books.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

目录

关于作者 i

前言 iii

本卷导读·第一卷 3

第一章 预备知识 6

1.1 因式分解 . . . 6
1.2 乘法公式与因式分解 . . . 11
1.3 恒等变形的方法 . . . 24
1.4 二次函数 . . . 35
1.5 计算与变形技巧 . . . 41
1.6 数论进阶 . . . 54

第二章 集合与映射 62

2.1 集合论初步 . . . 62
2.2 映射 . . . 78
2.3 逻辑用语初步 . . . 80
2.4 势与对角线法 * . . . 83

第三章 函数 88

3.1 函数的概念 . . . 88
3.2 定义域 . . . 90
3.3 单调性 . . . 92
3.4 奇偶性 . . . 96
3.5 周期性 . . . 111
3.6 凹凸性 . . . 119
3.7 反函数 . . . 127
3.8 函数的复合 . . . 131
3.9 函数的图像变换 . . . 135
3.10 常见多项式函数 . . . 139
3.11 指数函数 . . . 151
3.12 幂函数 . . . 160
3.13 对数函数 . . . 162
3.14 函数模型与应用 . . . 170
3.15 函数的零点(一) . . . 173

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

3.16 分段函数 . . . 179
3.17 绝对值函数 . . . 194
3.18 函数解析式的求解 . . . 200
3.19 函数值域的求解 . . . 209
3.20 多项式理论 . . . 213
3.21 三次与四次方程 . . . 220

第四章 函数方程与抽象函数 228
4.1 基本分析方法 . . . 228
4.2 从基本初等函数到函数方程 . . . 230
4.3 二次及高次多项式模型 . . . 231
4.4 三角与双曲函数模型 . . . 234
4.5 柯西方程的正则解 . . . 237
4.6 Jensen 方程与 Pexider 方程 . . . 240
4.7 迭代函数方程 . . . 243
4.8 嵌套根式与函数方程 . . . 246
4.9 解题策略小结 . . . 249

第五章 不等式 254
5.1 均值不等式 . . . 254
5.2 均值不等式的提示 . . . 258
5.3 均值不等式链 . . . 261
5.4 均值不等式的应用技巧 . . . 263
5.5 柯西不等式 . . . 267
5.6 排序不等式 . . . 278
5.7 伯努利不等式 . . . 280
5.8 糖水不等式 . . . 283
5.9 放缩思想 . . . 290
5.10 局部化与孤立思想 . . . 295
5.11 构造思想 . . . 299
5.12 齐次化与规范化 . . . 303
5.13 代数代换思想 . . . 309
5.14 排序与调整思想 . . . 313
5.15 凸性思想 . . . 317
5.16 进阶不等式 . . . 320
5.17 切比雪夫与 Schur 不等式 . . . 323
习题 . . . 327

本卷导读·第二卷 333

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第六章 三角函数 **336**
6.1 弧度制 . . . 337
6.2 正弦函数与余弦函数 . . . 339
6.3 正弦与余弦函数的图像 . . . 343
6.4 正弦函数与余弦函数的性质 . . . 345
6.5 正切函数 . . . 348
6.6 余切、正割与余割函数 . . . 352
6.7 同角三角函数的基本关系 . . . 356
6.8 诱导公式 . . . 362
6.9 和差角公式 . . . 369
6.10 倍角公式 . . . 376
6.11 万能公式与齐次式思想 . . . 387
6.12 角的变换 . . . 395
6.13 和差化积与积化和差 . . . 399
6.14 $y = A\sin(\omega x + \varphi) + B$ . . . 407
6.15 ω 的约束与求解 . . . 416
6.16 反三角函数 . . . 436

第七章 数列 **444**
7.1 数学归纳法 . . . 444
7.2 $\sum$ 和 $\prod$ 符号 . . . 448
7.3 数列的表示 . . . 460
7.4 等差数列 . . . 461
7.5 等比数列 . . . 465
7.6 差比数列 . . . 467
7.7 斐波那契数列 . . . 481
7.8 数列求和的通用策略 . . . 492
7.9 递推关系与通项公式 . . . 525
7.10 一阶递推的基本形式 . . . 530
7.11 迭代与不动点 . . . 541
7.12 数列的单调性 . . . 549
7.13 奇偶数列 * . . . 552
7.14 数列的周期性 . . . 554
7.15 子列与 Bolzano-Weierstrass 定理 * . . . 559

第八章 统计 **564**
8.1 描述数据 . . . 565
8.2 抽样方法 . . . 568
8.3 从样本到总体 . . . 575

知乎@余命数

8.4 相关与回归 . . . 578
8.5 章末习题 . . . 583
习题 . . . 583

第九章 计数原理 586
9.1 加法原理与乘法原理 . . . 587
9.2 组合与排列 . . . 597
9.3 二项式定理 . . . 602
9.4 从三项式定理到多项式定理 * . . . 607
9.5 排列组合综合应用 . . . 614
习题 . . . 644

第十章 概率 648
10.1 事件与概率 . . . 648
10.2 经典概率问题 . . . 663
10.3 条件概率 . . . 672
10.4 事件的相互独立性 . . . 677
10.5 全概率公式与贝叶斯公式 . . . 682
10.6 离散型随机变量及其分布列 . . . 689
10.7 期望与方差 . . . 692
10.8 常见分布模型 . . . 696
习题 . . . 701

本卷导读·第三卷 705

第十一章 平面向量 708
11.1 向量代数 . . . 708
11.2 定比分点与共线定理 . . . 748
11.3 向量外积与混合积 . . . 764

第十二章 解三角形 772
12.1 正弦定理和余弦定理 . . . 773
12.2 范围与综合 . . . 797
12.3 保底的建系法 . . . 832
12.4 非常规题型 . . . 838
12.5 梅涅劳斯定理与塞瓦定理 . . . 842
12.6 斯特瓦尔特定理与三角形特殊点 . . . 847
12.7 托勒密定理与西姆松定理 . . . 851

第十三章 解析几何 858

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.1 直线与圆 858
13.2 线性规划初步 * 876
13.3 二次曲线 908
13.4 设点法的技术基础 971
13.5 设点法的参量选择 983
13.6 设点法的典型应用 993
13.7 联立运算的基础准备 1003
13.8 直线方程的形式选择 1026
13.9 联立后的代数处理技巧 1036
13.10 一次式与法线式 1057
13.11 直线偶与齐次化 1064
13.12 三角形的心 1069
13.13 共圆点与共轴圆 1073
13.14 轨迹方程 1078
13.15 解析法证几何题 1089
13.16 解析几何思想与方法论 1097
13.17 隐圆 1101
13.18 直线与圆的综合应用 1111
13.19 垂直弦问题 1122
13.20 斜率和差商积 1135
13.21 圆幂与反演 1148
13.22 圆锥曲线的光学性质 1153
13.23 焦点弦 1157

第十四章 极坐标与参数方程 * 1164

第十五章 射影几何 1194

15.1 源起与观念 1194
15.2 单比与交比 1195
15.3 二次曲线的极点与极线 1200
15.4 极点与极线的几何作图 1209
15.5 笛沙格定理与帕普斯定理 1215
15.6 帕斯卡与布里昂桑 1219

本卷导读·第四卷 1227

第十六章 极限与导数 1230

16.1 实数系的完备性 1230
16.2 极限的定义 * 1232

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

16.3 数列的极限 * 1239
16.4 级数 1257
16.5 切线问题 1269
16.6 函数的连续性 1275
16.7 极限的运算 1277
16.8 无穷远处的极限与无穷小量 1282
16.9 导数的概念 1289
16.10 中值定理的应用与推广 1320
16.11 导数的综合应用 1324
16.12 高阶导数与泰勒展开初步 1340
16.13 泰勒公式与泰勒级数 1344
16.14 洛必达法则与极限计算的系统策略 1349
16.15 必要性探路与端点分析法 * 1355
16.16 极值点偏移问题 1370
16.17 牛顿迭代法 1374

第十七章 简单的积分 1382
17.1 原函数与不定积分 1382
17.2 不定积分 1386
17.3 换元积分法 1393
17.4 分部积分法 1401
17.5 定积分与微积分基本定理 1412
17.6 定积分的性质 1422
17.7 定积分的几何与物理应用 1433

第十八章 微分方程初步 1442
18.1 微分方程、阶与解 1442
18.2 可分离变量的微分方程 1444
18.3 一阶线性微分方程 1447
18.4 应用微分方程求解函数方程 1448
18.5 增长与衰减模型 1452
18.6 落体与极限速度 1455
18.7 悬链线与追逐曲线 *(选读) 1458
18.8 章末习题 1461
习题 1461

第十九章 数学思想浅谈 1464
19.1 数学对象与问题三要素 1464
19.2 自由度与变量、约束、独立参数 1466

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

19.3 自由度如何决定解的个数 1471
19.4 解题框架的四种基本路径 1473
19.5 纯构造式问题的解法 1480
19.6 直接构造困难时的反证法 1482
19.7 含满足式问题的通用策略 1483
19.8 不变量、单调量、极端原理、抽屉原理与对称 1486
19.9 谈谈二级结论 1497
19.10 二级结论的适用边界 1499
19.11 换元法为何行之有效 *(选读) 1500
19.12 换元法的典型模式 1504
19.13 本章小结与习题 1508
习题 1510

本卷导读·第五卷 **1515**

第二十章 复数 **1518**
20.1 复数的概念与几何意义 1519
20.2 复数的运算 1523
20.3 二次方程的复数解 1524
20.4 复数运算的几何意义 * 1526
20.5 代数基本定理 * 1532
20.6 曲线方程的复数表示 * 1534
20.7 本章小结与思想方法 1540

第二十一章 立体几何 **1544**
21.1 基本公理与空间元素 1544
21.2 证明线线平行的基本策略 1550
21.3 空间中的平行关系 1564
21.4 空间中的垂直关系 1570
21.5 空间向量与坐标法 1575
21.6 空间中的度量 1581
21.7 三面角与四面体 1587
21.8 球 1591

第二十二章 矩阵与行列式 **1604**
22.1 向量 1605
22.2 矩阵的定义与运算 1608
22.3 行列式 1620
22.4 线性变换与几何应用 1636

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

22.5 逆矩阵与克拉默法则 1651
22.6 秩与高斯消元 1656

第二十三章 连分数 **1664**
23.1 辗转相除法与连分数 1664
23.2 有限连分数与有理数 1667
23.3 无限连分数与无理数 1670
23.4 最佳有理逼近 1672
23.5 黄金比 1674
23.6 逼近的阶 1678

第二十四章 常数 e、π 与 γ **1686**
24.1 自然常数 e 1686
24.2 π 的无理性 1691
24.3 欧拉常数 γ 1693
24.4 巴塞尔问题 1695

第二十五章 无穷乘积 **1698**
25.1 无穷乘积的收敛性 1698
25.2 Wallis 乘积 1700
25.3 Vieta 公式 1702
25.4 欧拉乘积 1704

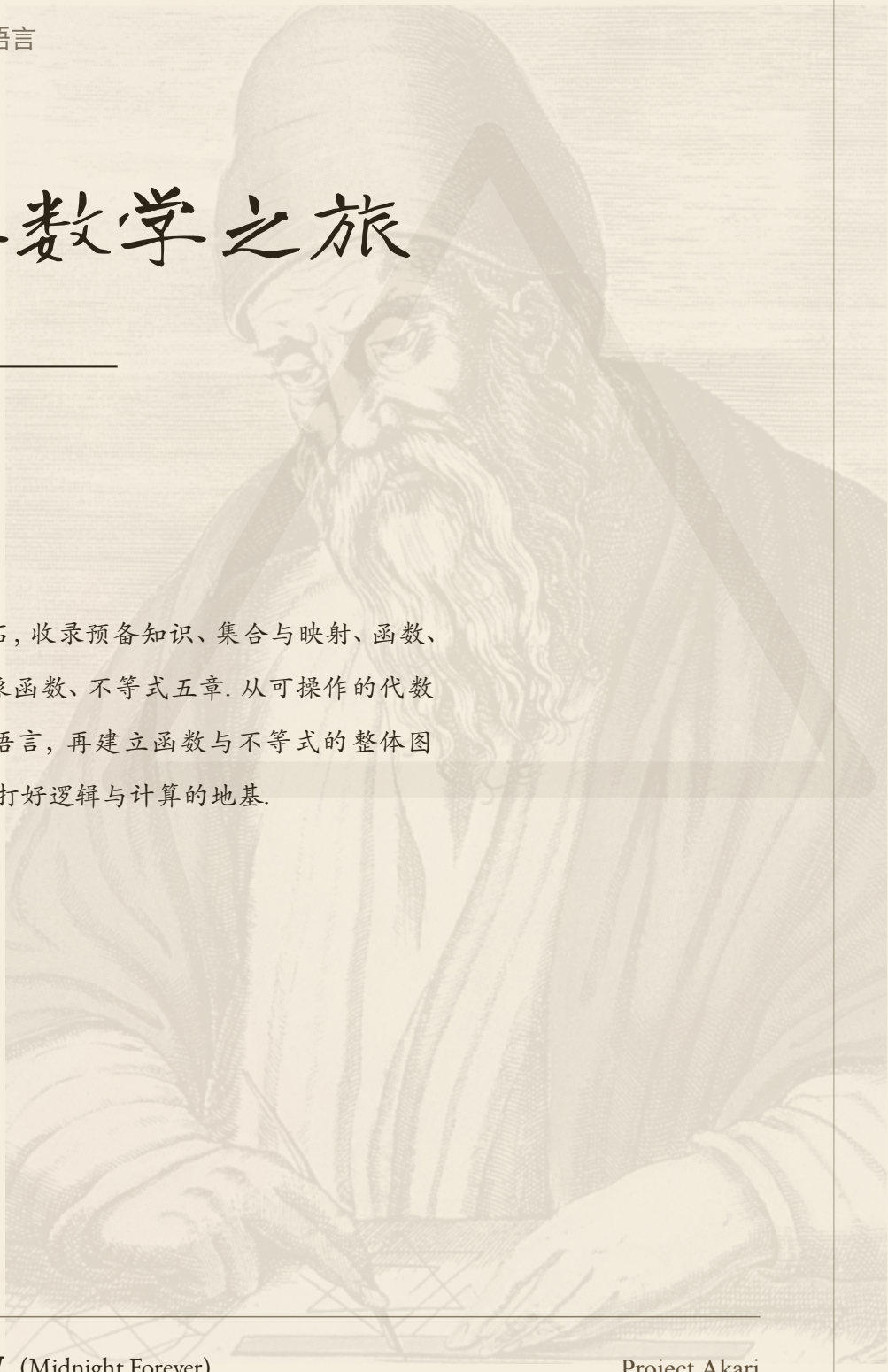
习题解答 **1707**
第 1 章预备知识 1707
第 8 章统计 1710
第 12 章解三角形 1710
第 13 章解析几何 1715
第 15 章射影几何 1716
第 16 章极限与导数 1717
第 18 章微分方程初步 1746
第 21 章立体几何 1746
第 23 章连分数 1750

术语索引 **1751**

知乎@余命数
知乎@余命数
@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第一卷 · 基础语言

中学数学之旅



本卷为系列基石，收录预备知识、集合与映射、函数、函数方程与抽象函数、不等式五章. 从可操作的代数变形走到对应语言，再建立函数与不等式的整体图景，为后续各卷打好逻辑与计算的地基.

知乎 @ 余命数 (Midnight Forever)

Project Akari

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

本卷导读

本卷线索

第一卷的主线是：**从可操作的代数变形，走到可陈述的对应语言，再落到可分析的函数与不等式.** 具体技巧服务于结构；结构又反过来解释技巧为何有效.

知乎@余命数

章节地图

▷ **预备知识** 整理因式分解、乘法公式、恒等变形、二次函数与若干数论事实，是后续各章随时调用的工具基础.

▷ **集合与映射** 用集合描述对象，用映射描述对应，并初步引入命题与量词. 函数的概念由此获得精确的表述.

▷ **函数** 本卷重心. 先建立定义域、单调性、奇偶性、周期性等性质，再讨论复合、反函数与图像变换，然后进入常见类型与综合问题. 全章论证从定义与图像出发.

▷ **函数方程与抽象函数** 在表达式未知、关系已知的设定下训练赋值与结构假设，把函数章建立的性质用作推理的支点.

▷ **不等式** 以均值不等式为主线，辅以柯西、排序等工具，并展开放缩与构造等思想. 切线法与凸性观点把函数图像的几何直觉化为估计手段.

知乎@余命数 知乎@余命数

向后展望

第二卷将把函数语言用于周期现象（三角函数）与离散对象，后者即数列初等部分，并经计数进入统计与概率. 第一卷建立的单调性、图像与不等式估计，会在那里反复出现.

知乎@余命数

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 1 章

预备知识

基础语言

「几何无王者之路。」

——欧几里得

第一章 预备知识

1.1 因式分解

因式分解把多项式写成若干整式的乘积, 我们用它分析代数结构、求解方程、简化运算. 多元多项式的因式分解常常依靠式子本身的结构, 如对称性、轮换性、齐次性.

1.1.1 特殊代数式

我们讨论三类多项式, 即对称式、轮换式与齐次式. 整理复杂多项式、降元与因式分解时经常遇到它们.

1.1.2 定义与基本性质

在给出形式定义之前, 我们先定义一个贯穿本节的概念, 即**置换**. 对称式与轮换式都要用它来描述.

1.1.2.1 置换

设有三个位置, 分别摆放着变量 x_1, x_2, x_3. 我们把它们重新摆放, 比如变成 x_2, x_3, x_1. 这一过程称为一个**置换**, 记作

$$\sigma : (x_1, x_2, x_3) \longmapsto (x_2, x_3, x_1).$$

此时第 1 个位置改放 x_2, 第 2 个位置改放 x_3, 第 3 个位置改放 x_1. 下图的箭头标明了每个变量在新排列中落到的位置.

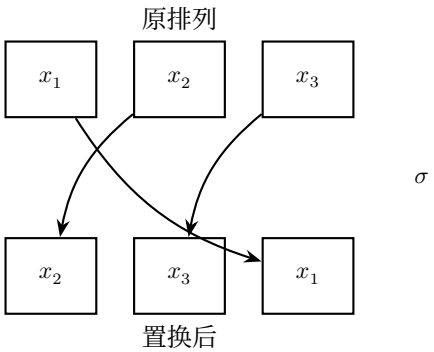


图 1.1 三元置换 σ 中变量位置的循环挪动

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

只交换两个变量而其余位置不动的置换, 称为**对换**, 记作 τ_{ij}. 每个置换都可由若干次对换拼成. 例如上图中的 σ 可由两次对换实现, 先交换 x_1, x_2, 再交换(此时的)x_1, x_3. 下图展示最简单的对换 τ_{12}.

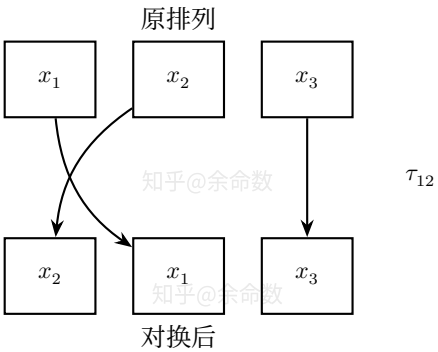


图 1.2　对换 τ_{12} 中 x_1 与 x_2 互换, x_3 不动

我们把 n 个变量的全部置换汇集起来, 称之为 n **次对称群** S_n. 它共有 $n!$ 个元素, 三个变量时有 $3! = 6$ 种不同排法. 恒等排列, 即什么都不换的那个排列, 也是其中之一.

$n!$ 读作"n 的阶乘", 即 $n! = 1 \cdot 2 \cdots n$, 见第9.2节.

有了以上直观, 我们给出对称式的定义. 对称式是在任意置换下都不改变的多项式. 由于每个置换都能由对换拼成, 检验对称性时逐一检查对换就够了.

定义 1.1 (对称式) 设 $P(x_1, \dots, x_n)$ 为 n 元多项式. 若对任意置换 $\sigma \in S_n$, 都有

$$P(x_{\sigma(1)}, \dots, x_{\sigma(n)}) \equiv P(x_1, \dots, x_n),$$

则称 P 为关于 $x_1, \dots, x_n$ 的**对称式**.

"群"是代数术语, 指一个配备了运算的集合, 运算满足四条性质: 封闭性、结合律、存在单位元、每个元素有逆元. S_n 以置换的复合为运算, 记作 $\circ$, $\tau \circ \sigma$ 表示先做 σ 再做 τ, 逐一满足: 两置换复合仍是置换, 复合满足结合律, 恒等排列是单位元, 每个置换 σ 都有逆置换 σ^{-1}, 上标 -1 标记逆置换.

例题 1.1
- $P(x, y) = x^2 + y^2$ 在交换 x, y 后不变, 为对称式.
- $P(x, y, z) = xyz + x + y + z$ 在任意交换 x, y, z 后不变, 为三元对称式.
- $P(x, y) = x - y$ 在交换 x, y 后变为 $y - x = -(x - y)$, 故非对称式.

注 1.1 (基本性质)
设 P, Q 均为关于 $x_1, \dots, x_n$ 的对称式.

- 和与积仍为对称式, 即 $P \pm Q$ 与 PQ 皆为对称式.
- 代入保持对称性, 即 f 为一元多项式时, $f(P)$ 仍为对称式.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

- 若 $P=AB$, 则对每个置换 σ 有

$$P(x_\sigma)=A(x_\sigma)B(x_\sigma)=P(x),$$

因此两个因式在置换下满足 $A(x_\sigma)/A(x)=B(x)/B(x_\sigma)$(作为有理函数相等). 例如 $P=x_1x_2x_3$ 时取 $A=x_1$、$B=x_2x_3$, 两个因式都不是对称式, 乘积却是对称式.

注 1.2 (基本对称式与表示)
对 n 元变量 $x_1,\dots,x_n$, 记

$$e_1=x_1+\cdots+x_n,\quad e_2=\sum_{1\le i<j\le n}x_ix_j,\quad \dots,\quad e_n=x_1x_2\cdots x_n,$$

($\sum$ 读作对……求和. 例如 $\sum_{1\le i<j\le n}x_ix_j$ 表示所有满足 $1\le i<j\le n$ 的乘积 x_ix_j 相加. 三元时即 $x_1x_2+x_1x_3+x_2x_3$. 系统记号见第7.2节.)称为**基本对称式**. 任一关于 $x_1,\dots,x_n$ 的对称多项式, 都可以表示为 $e_1,\dots,e_n$ 的多项式. 例如在两元情形,

$$\text{任一对称式}P(x,y)\text{ 都可写成}P(x,y)=F(x+y,xy),$$

其中 F 为某一元或二元多项式.

知乎@余命数

例题 1.2
- $x^2+y^2=(x+y)^2-2xy$, 为 e_1,e_2 的多项式.
- $x^3+y^3=(x+y)^3-3xy(x+y)$.

注 1.3 (对称性与因式)
在两元情形, 若 $P(x,y)$ 为对称式且含有因式 $(x-y)$, 则 $(x-y)$ 的次数为偶数. 原因是对换 x,y 时 $(x-y)$ 变号, 而 P 不变, 故 $(x-y)$ 的总次方必须为偶数, 例如 $(x-y)^2,(x-y)^4$ 等.

定义 1.2 (轮换式) 设 $P(x_1,\dots,x_n)$ 为 n 元多项式. 若在轮换

$$(x_1,x_2,\dots,x_n)\mapsto(x_2,x_3,\dots,x_n,x_1)$$

下恒有

$$P(x_2,x_3,\dots,x_n,x_1)\equiv P(x_1,\dots,x_n),$$

则称 P 为关于 $x_1,\dots,x_n$ 的**轮换式**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 1.3 以 a,b,c 的轮换 $a\to b\to c\to a$ 为例.

- $a^2b+b^2c+c^2a$ 在轮换下不变, 为轮换式; 在交换 a,b 时变式, 故非对称式.
- $(a-b)(b-c)(c-a)$ 在轮换下仅作因子重排, 乘积不变, 为轮换式.

注 1.4 (与对称式的关系及封闭性)

- 任一对称式在任意置换下不变, 因而在轮换下亦不变; 对称式必为轮换式, 反之不然.
- 若 P,Q 为轮换式, 则 $P\pm Q$ 与 PQ 仍为轮换式.

数
知乎@余命数

注 1.5 (因式传递性)
设 $P(a,b,c)$ 为关于 a,b,c 的轮换式.

- 若 P 含有因式 $(a-b)$, 则在轮换 $a\to b\to c\to a$ 下, P 也含 $(b-c)$ 与 $(c-a)$.
- 类似地, 若 P 含有因式 (a^2-b^2), 则含有 (b^2-c^2) 与 (c^2-a^2) 等.

这类因式成轮换族, 可一并提取.

知乎@余命数

定义 1.3 (齐次式) 设 $P(x_1,\dots,x_n)$ 为多项式. 若 P 中每一项的总次数均为同一整数 d, 则称 P 为 d **次齐次式**.

例题 1.4

- $2x^2+3xy-y^2$ 各项为 2 次, 为二次齐次式.
- $a^3+b^3+c^3-3abc$ 各项为 3 次, 为三次齐次式.
- x^2+y 各项次数不等, 非齐次式.

引理 1.1 (齐次性的等价刻画) 设 P 为非零多项式. 则 P 为 d 次齐次式, 当且仅当对任意数 t, 有

$$P(tx_1,\dots,tx_n)=t^dP(x_1,\dots,x_n).$$

证明 若 P 为 d 次齐次式, 则每一项形如 $c\,x_1^{k_1}\cdots x_n^{k_n}$, 且 $k_1+\cdots+k_n=d$. 代入 tx_i 后, 每项变为

$$c\,t^{k_1+\cdots+k_n}x_1^{k_1}\cdots x_n^{k_n}=t^d\cdot c\,x_1^{k_1}\cdots x_n^{k_n},$$

从而整体为 $t^d P(x_1, \dots, x_n)$.

反过来, 将 P 按总次数分解为若干齐次部分之和

$$P = P_m + P_{m+1} + \cdots + P_M,$$

其中非零的 P_k 是 k 次齐次部分. 代入 tx_i 后得到

$$P(tx_1, \dots, tx_n) = \sum_{k=m}^{M} t^k P_k(x_1, \dots, x_n).$$

@余命数 知乎@余命数

若它恒等于 $t^d P(x_1, \dots, x_n)$, 则比较关于 t 的各次幂系数可知, 除 $k = d$ 外所有 P_k 均为零. 因此所有非零项的总次数都等于 d, 即 P 为 d 次齐次式. □

注 1.6 (因式与降元)
若 P 为 d 次齐次式且 $P = QR$, 则 Q, R 亦必为齐次式, 且 $\deg Q + \deg R = d$. 在二元情形, 齐次性可用于降元. 设

$$F(x, y) = 0$$

为 d 次齐次方程, 若 $y \neq 0$, 令 $t = \dfrac{x}{y}$, 两边除以 y^d, 得一元方程

$$F(t, 1) = 0,$$

解出 t 后再还原 $x : y$ 的比值. 这在因式分解和解齐次方程组时常用.

注 1.7 (对称性、轮换性与齐次性的组合)
多项式往往同时具有多种结构. 例如:

- $x^2 + xy + y^2$ 是对称式、(两元)轮换式、二次齐次式.
- $a^2b + b^2c + c^2a$ 是三元轮换式、三次齐次式, 非对称式.
- $a^3 + b^3 + c^3 - 3abc$ 是三元对称式、三次齐次式.

在因式分解或恒等变形中, 通常结合多种结构使用, 如利用齐次性降元, 配合对称性或轮换性判断因式形状.

1.1.3 应用练习

知乎@余命数 知乎@余命数

练习 1.1 判断下列多项式的类型(对称式、轮换式、齐次式).

(a) $P(x,y)=x^2+xy+y^2$;

(b) $P(a,b,c)=a^2b+b^2c+c^2a$;

(c) $P(x,y,z)=x^3+y^3+z^2$;

(d) $P(a,b,c)=(a-b)(b-c)(c-a)$.

知乎@余命数

练习 1.2 已知关于 a,b,c 的轮换式 $P(a,b,c)$ 含有因式 $(a-b)$, 写出它必然含有的另外两个一次因式.

练习 1.3 因式分解

$$x(y^2-z^2)+y(z^2-x^2)+z(x^2-y^2).$$

练习 1.4 设 $F(x,y)$ 为二元 d 次齐次多项式,方程 $F(x,y)=0$ 中 $(x,y)\neq(0,0)$. 令 $t=\dfrac{x}{y}(y\neq 0)$ 或 $t=\dfrac{y}{x}(x\neq 0)$, 将方程化为一元方程, 写出一般变形过程.

1.2 乘法公式与因式分解

我们先整理常用的乘法公式及其逆运算, 再给出若干因式分解方法, 最后推广到高次多项式.

1.2.1 平方差与完全平方

性质 (平方差) 对任意数 a,b,

$$a^2-b^2=(a-b)(a+b).$$

证明 展开右端得

知乎@余命数

$$(a-b)(a+b)=a^2+ab-ba-b^2=a^2-b^2.$$

□

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 1.5 计算 $2024 \cdot 2016$.

解 这两个数关于 2020 对称. 令 $a = 2020, b = 4$, 则

$$2024 \cdot 2016 = (2020 + 4)(2020 - 4) = 2020^2 - 4^2 = 4080400 - 16 = 4080384.$$

例题 1.6 因式分解 $x^2 - 4y^2 - 4y - 1$.

解 后三项构成完全平方, 即

$$x^2 - (4y^2 + 4y + 1) = x^2 - (2y + 1)^2,$$

于是

$$x^2 - 4y^2 - 4y - 1 = (x - 2y - 1)(x + 2y + 1).$$

性质 (完全平方) 对任意数 a, b, c,

$$(a \pm b)^2 = a^2 \pm 2ab + b^2,$$

$$(a + b + c)^2 = a^2 + b^2 + c^2 + 2(ab + bc + ca).$$

证明 二项式情形直接展开即可. 对三项式, 令 $u = a + b$, 有

$$(a + b + c)^2 = (u + c)^2 = u^2 + 2uc + c^2,$$

再代回 $u^2 = (a + b)^2 = a^2 + 2ab + b^2$ 即得. □

例题 1.7 计算 101^2 与 99^2.

解

$$101^2 = (100 + 1)^2 = 100^2 + 2 \cdot 100 \cdot 1 + 1 = 10201,$$

$$99^2 = (100 - 1)^2 = 100^2 - 2 \cdot 100 \cdot 1 + 1 = 9801.$$

例题 1.8 已知 $a - b = 4, ab = 5$, 求 $(a + b)^2$.

解 由

$$(a + b)^2 = (a - b)^2 + 4ab,$$

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

得

$$(a+b)^2 = 4^2 + 4\cdot 5 = 36.$$

例题 1.9 已知 $x+\dfrac{1}{x}=5$, 求 $x^2+\dfrac{1}{x^2}$ 与 $x^4+\dfrac{1}{x^4}$.

知乎@余命数

解 平方得

$$\left(x+\frac{1}{x}\right)^2 = x^2+\frac{1}{x^2}+2 = 25,$$

知乎@余命数

从而

$$x^2+\frac{1}{x^2} = 23.$$

再平方一次,

$$\left(x^2+\frac{1}{x^2}\right)^2 = x^4+\frac{1}{x^4}+2 = 23^2,$$

故

$$x^4+\frac{1}{x^4} = 23^2-2 = 527.$$

知乎@余命数

1.2.2 立方公式与幂的和差

性质 (立方和与立方差) 对任意数 a, b,

$$a^3+b^3 = (a+b)(a^2-ab+b^2),$$

$$a^3-b^3 = (a-b)(a^2+ab+b^2).$$

知乎@余命数

证明 对 a^3+b^3 添项拆项, 得

$$a^3+b^3 = a^3+a^2b-a^2b-ab^2+ab^2+b^3 = a^2(a+b)-ab(a+b)+b^2(a+b),$$

提取公因式 $(a+b)$ 即得

$$a^3+b^3 = (a+b)(a^2-ab+b^2).$$

立方差同理. □

知乎@余命数

性质 (完全立方) 对任意数 a, b,

$$(a \pm b)^3 = a^3 \pm 3a^2b + 3ab^2 \pm b^3.$$

证明 由二项式自乘,

$$(a+b)^3 = (a+b)(a+b)^2 = (a+b)(a^2+2ab+b^2),$$

展开整理即得所述公式.$(a-b)^3$ 类似. □

例题 1.10 因式分解 $8x^3 - 27y^3$.

解

$$8x^3 - 27y^3 = (2x)^3 - (3y)^3 = (2x-3y)(4x^2+6xy+9y^2).$$

例题 1.11 化简 $\dfrac{x^3+1}{x^2-x+1}$.

解 对分子作立方和分解, 得

$$x^3 + 1 = (x+1)(x^2-x+1),$$

约去公因式, 故

$$\frac{x^3+1}{x^2-x+1} = x+1.$$

又 $x^2 - x + 1 = (x - \frac{1}{2})^2 + \frac{3}{4} > 0$, 对一切实数 x 分母不为零.

例题 1.12 已知 $x + \dfrac{1}{x} = 3$, 求 $x^3 + \dfrac{1}{x^3}$.

解

$$\left(x + \frac{1}{x}\right)^2 = x^2 + \frac{1}{x^2} + 2 = 9,$$

得 $x^2 + \dfrac{1}{x^2} = 7$. 于是

$$x^3 + \frac{1}{x^3} = \left(x + \frac{1}{x}\right)\left(x^2 - 1 + \frac{1}{x^2}\right) = 3 \cdot (7-1) = 18.$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

亦可用

$$x^3+\frac{1}{x^3}=\left(x+\frac{1}{x}\right)^3-3\cdot x\cdot\frac{1}{x}\left(x+\frac{1}{x}\right)=27-9=18.$$

定理 1.1 (幂的差) 对任意正整数 n,

$$a^n-b^n=(a-b)(a^{n-1}+a^{n-2}b+\cdots+ab^{n-2}+b^{n-1}).$$

证明 设

$$S=a^{n-1}+a^{n-2}b+\cdots+ab^{n-2}+b^{n-1},$$

则乘积

$$(a-b)S=a^n-b^n$$

其中中间各项成对抵消, 仅保留首尾两项. □

定理 1.2 (奇数次幂的和) 若 n 为正奇数, 则

$$a^n+b^n=(a+b)(a^{n-1}-a^{n-2}b+\cdots-ab^{n-2}+b^{n-1}).$$

证明 我们把幂的差公式用于 $a^n-(-b)^n$. 因 n 为奇数, $(-b)^n=-b^n$, 故

$$a^n+b^n=a^n-(-b)^n=(a+b)(a^{n-1}+a^{n-2}(-b)+\cdots+(-b)^{n-1}).$$

由于 $n-1$ 为偶数, 括号内各项符号交替, 末项为 b^{n-1}, 即得所述公式. □

例题 1.13 分解 x^5-y^5 与 x^5+y^5.

解

$$x^5-y^5=(x-y)(x^4+x^3y+x^2y^2+xy^3+y^4),$$

$$x^5+y^5=(x+y)(x^4-x^3y+x^2y^2-xy^3+y^4).$$

例题 1.14 分解 x^4+4x^2+16.

解 我们把它视为关于 x^2 的二次式. 若用十字相乘, 在整数系数范围难以直接配对. 改用结构式

$$x^4+4x^2+16=(x^2+2x+4)(x^2-2x+4),$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

展开右端验证为

$$(x^2+2x+4)(x^2-2x+4)=x^4+4x^2+16.$$

注 1.8 (索菲·热尔曼型恒等式)
对任意 a,b,

$$a^4+4b^4=(a^2+2ab+2b^2)(a^2-2ab+2b^2),$$

其中 x^4+4 为 $a=x,b=1$ 的特例.

1.2.3 基本因式分解方法

1.2.3.1 提公因式与分组分解

例题 1.15 因式分解 $3x(a-b)-2y(b-a)$.

解 $(b-a)=-(a-b)$, 故

$$3x(a-b)-2y(b-a)=3x(a-b)+2y(a-b)=(a-b)(3x+2y).$$

例题 1.16 因式分解 $ab(a-b)+bc(b-c)+ca(c-a)$.

解 有

$$(a-b)+(b-c)+(c-a)=0,\quad c-a=-(a-b)-(b-c).$$

代入, 得

$$\begin{aligned}ab(a-b)+bc(b-c)+ca(c-a)&=ab(a-b)+bc(b-c)\\&\quad+ca[-(a-b)-(b-c)]\\&=[ab(a-b)-ca(a-b)]\\&\quad+[bc(b-c)-ca(b-c)]\\&=(a-b)(ab-ca)+(b-c)(bc-ca).\end{aligned}$$

注意 $ab-ca=a(b-c)$, $bc-ca=c(b-a)$, 于是

$$\begin{aligned}&=(a-b)a(b-c)+(b-c)c(b-a)\\&=(a-b)a(b-c)-(b-c)c(a-b)\\&=(a-b)(b-c)(a-c)\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= -(a-b)(b-c)(c-a).$$

1.2.3.2 配方法与添项拆项

例题 1.17 因式分解 x^4+4.

解 视 x^4 与 4 为 $(x^2)^2$ 与 2^2, 构造中间项 $4x^2$, 得

$$x^4+4=(x^4+4x^2+4)-4x^2=(x^2+2)^2-(2x)^2.$$

平方差公式给出

$$x^4+4=(x^2-2x+2)(x^2+2x+2).$$

例题 1.18 因式分解 $x^2-6x+9-y^2$.

解

$$x^2-6x+9=(x-3)^2,$$

原式为

$$(x-3)^2-y^2=(x-3-y)(x-3+y)=(x-y-3)(x+y-3).$$

1.2.3.3 整体代换

例题 1.19 因式分解 $2(x+y)^2-5(x+y)-3$.

解 令 $A=x+y$, 则

$$2A^2-5A-3=(2A+1)(A-3),$$

代回得

$$2(x+y)^2-5(x+y)-3=[2(x+y)+1][(x+y)-3].$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1.2.4 二次多项式

1.2.4.1 十字相乘法

定义 1.4 设 $a,b,c\in\mathbb{R},a\neq 0$, 一元二次三项式

$$ax^2+bx+c$$

若可写成

$$(a_1x+c_1)(a_2x+c_2),$$

则称右式为其在实数范围内的因式分解.

展开右式, 得

$$(a_1x+c_1)(a_2x+c_2)=a_1a_2x^2+(a_1c_2+a_2c_1)x+c_1c_2,$$

比较系数, 需满足

$$a_1a_2=a,\quad c_1c_2=c,\quad a_1c_2+a_2c_1=b.$$

例题 1.20 因式分解 $3x^2+11x-4$.

解 $3x^2=3x\cdot x,-4=-1\cdot 4$, 取

$$a_1=3,\ a_2=1,\ c_1=-1,\ c_2=4,$$

得

$$(3x)\cdot 4+(x)\cdot(-1)=12x-x=11x,$$

故

$$3x^2+11x-4=(3x-1)(x+4).$$

例题 1.21 因式分解 $2x^2+7x+3$.

解 $2x^2=2x\cdot x,3=1\cdot 3$, 取

$$a_1=2,\ a_2=1,\ c_1=1,\ c_2=3,$$

得

$$(2x)\cdot 3+(x)\cdot 1=7x,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于是

$$2x^2+7x+3=(2x+1)(x+3).$$

1.2.4.2 双十字相乘法

定义 1.5 设 $a,b,c\in\mathbb{R},ac\neq 0$. 齐次二次式

$$ax^2+bxy+cy^2$$

若可写成

$$(a_1x+c_1y)(a_2x+c_2y),$$

则称右式为其因式分解.

知乎@余命数

展开右式, 得

$$(a_1x+c_1y)(a_2x+c_2y)=a_1a_2x^2+(a_1c_2+a_2c_1)xy+c_1c_2y^2,$$

比较系数, 需要

$$a_1a_2=a,\quad c_1c_2=c,\quad a_1c_2+a_2c_1=b.$$

知乎@余命数

例题 1.22 因式分解 $6x^2+5xy-6y^2$.

解 $6x^2=3x\cdot 2x,-6y^2=-2y\cdot 3y$, 取

$$a_1=3,\ a_2=2,\ c_1=-2,\ c_2=3,$$

知乎@余命数

有

$$(3x)(3y)+(2x)(-2y)=9xy-4xy=5xy,$$

故

$$6x^2+5xy-6y^2=(3x-2y)(2x+3y).$$

例题 1.23 因式分解 $x^2-3xy-10y^2$.

解 $x^2=x\cdot x,-10y^2=-5y\cdot 2y$, 取

$$a_1=1,\ a_2=1,\ c_1=-5,\ c_2=2,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

交叉项

$$(x)(2y)+(x)(-5y)=2xy-5xy=-3xy,$$

于是

$$x^2-3xy-10y^2=(x-5y)(x+2y).$$

知乎@余命数

1.2.4.3 通用公式

定理 1.3 (一元二次通用公式) 设 $a,b,c\in\mathbb{R},a\neq 0$, 判别式

$$\Delta=b^2-4ac.$$

若 $\Delta\geq 0$, 则

$$ax^2+bx+c=a(x-x_1)(x-x_2),\quad x_{1,2}=\frac{-b\pm\sqrt{\Delta}}{2a}.$$

若 $\Delta<0$, 则在实数范围内不能分解为一次因式的乘积.

证明 方程 $ax^2+bx+c=0$ 等价于

$$x^2+\frac{b}{a}x+\frac{c}{a}=0.$$

配方得到

$$\left(x+\frac{b}{2a}\right)^2=\frac{b^2-4ac}{4a^2}=\frac{\Delta}{4a^2},$$

当 $\Delta\geq 0$ 时,

$$x=\frac{-b\pm\sqrt{\Delta}}{2a}.$$

两根为 x_1,x_2 时, 有

$$ax^2+bx+c=a(x-x_1)(x-x_2),$$

展开右端可核对系数. 当 $\Delta<0$ 时, 该方程无实根; 若它能在实数范围内分解为两个一次因式的乘积, 则方程必有实根, 与 $\Delta<0$ 矛盾. □

例题 1.24 用通用公式因式分解 $3x^2-11x+10$.

解 $a=3,b=-11,c=10$, 判别式

$$\Delta=(-11)^2-4\cdot 3\cdot 10=1,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根为

$$x_1 = 2, \quad x_2 = \frac{5}{3}.$$

于是

$$3x^2 - 11x + 10 = 3(x-2)\left(x - \frac{5}{3}\right) = (3x-5)(x-2).$$

知乎@余命数

定理 1.4 (齐次二次分解) 设 $a, b, c \in \mathbb{R}, a \neq 0$, $\Delta = b^2 - 4ac$. 若 $\Delta \geq 0$, 则存在实数 λ_1, λ_2,

$$ax^2 + bxy + cy^2 = a(x - \lambda_1 y)(x - \lambda_2 y),$$

其中

$$\lambda_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a}.$$

若 $\Delta < 0$, 则在实数范围内不能分解为一次因式的乘积.

证明 令 $t = x/y (y \neq 0)$, 则

$$ax^2 + bxy + cy^2 = y^2(at^2 + bt + c).$$

若 $\Delta \geq 0$, 一元二次 $at^2 + bt + c = 0$ 有根

$$\lambda_{1,2} = \frac{-b \pm \sqrt{\Delta}}{2a},$$

且

$$at^2 + bt + c = a(t - \lambda_1)(t - \lambda_2),$$

于是当 $y \neq 0$ 时,

$$ax^2 + bxy + cy^2 = a(x - \lambda_1 y)(x - \lambda_2 y).$$

两边都是关于 x, y 的多项式, 且在 $y \neq 0$ 时相等, 故作为多项式恒等式对所有 x, y 成立. 若 $\Delta < 0$ 而仍可在实数范围内分解为两个一次齐次因式, 则令 $y = 1$ 会得到一元二次式 $at^2 + bt + c$ 的实一次因式分解, 从而有实根, 矛盾. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1.2.5 高次多项式

引理 1.2 (余式定理) 设 $P(x)$ 为多项式, 将其除以 $(x-a)$, 得

$$P(x)=(x-a)Q(x)+r,\quad r\in\mathbb{R}.$$

则 $r=P(a)$.

知乎@余命数

证明 对上式代入 $x=a$, 得 $P(a)=r$. □

定理 1.5 (因式定理) 多项式 $P(x)$ 有因式 $(x-a)$ 当且仅当 $P(a)=0$.

证明 若 $P(x)=(x-a)Q(x)$, 则 $P(a)=0$. 反之, 若 $P(a)=0$, 由余式定理,

$$P(x)=(x-a)Q(x)+P(a)=(x-a)Q(x),$$

从而 $(x-a)$ 为因式. 知乎@余命数 □

定理 1.6 (有理根定理) 设

$$P(x)=a_nx^n+\cdots+a_0$$

为整系数多项式, $a_n\neq 0$. 若最简分数 $\frac{p}{q}(\gcd(p,q)=1,q>0)$ 为 $P(x)$ 的有理根, 则 q 为 a_n 的整数因数; 若 $p\neq 0$, 则 p 为 a_0 的整数因数, 若 $p=0$, 则 $a_0=0$.

知乎@余命数

证明 由 $P\left(\frac{p}{q}\right)=0$, 两边乘以 q^n 得

$$a_np^n+a_{n-1}p^{n-1}q+\cdots+a_1pq^{n-1}+a_0q^n=0.$$

移项可得

$$a_np^n=-q(a_{n-1}p^{n-1}+\cdots+a_1pq^{n-2}+a_0q^{n-1}),$$

因此 $q\mid a_np^n$. 又 $\gcd(p,q)=1$, 故 $\gcd(p^n,q)=1$, 从而 $q\mid a_n$.

同理,

$$a_0q^n=-p(a_np^{n-1}+a_{n-1}p^{n-2}q+\cdots+a_1q^{n-1}).$$

知乎@余命数

若 $p\neq 0$, 由 $\gcd(p,q^n)=1$ 得 $p\mid a_0$; 若 $p=0$, 直接由 $P(0)=0$ 得 $a_0=0$. □

知乎@余命数

例题 1.25　因式分解 $P(x)=x^3-2x^2-5x+6$.

解　这里 $a_3=1, a_0=6$. 根据有理根定理, 可能的有理根为 $\pm1,\pm2,\pm3,\pm6$. 逐一代入, 得

$$P(1)=1-2-5+6=0,$$

故 $(x-1)$ 为因式. 综合除法(系数为 $1,-2,-5,6$ 与根 1)得商

$$x^2-x-6.$$

继续因式分解, 得

$$x^2-x-6=(x-3)(x+2),$$

于是

$$P(x)=(x-1)(x-3)(x+2).$$

例题 1.26　因式分解 $P(x)=x^3-4x^2+8$.

解　这里 $a_3=1, a_0=8$. 根据有理根定理, 可能的有理根为 $\pm1,\pm2,\pm4,\pm8$. 逐个试算, 得

$$P(2)=8-16+8=0,$$

故 $(x-2)$ 为因式. 综合除法(系数 $1,-4,0,8$ 与根 2)得商

$$x^2-2x-4.$$

判别式

$$\Delta=(-2)^2-4\cdot1\cdot(-4)=20>0,$$

一元二次公式给出根

$$x=1\pm\sqrt{5}.$$

在实数范围内可写成

$$x^2-2x-4=(x-1-\sqrt{5})(x-1+\sqrt{5}),$$

在有理数范围则保留二次因式. 于是

$$P(x)=(x-2)(x^2-2x-4)$$

知乎@余命数

或

$$P(x)=(x-2)(x-1-\sqrt{5})(x-1+\sqrt{5}).$$

1.2.6 练习

练习 1.5 因式分解 $(x+2y)^2-(x-y)^2$.

练习 1.6 因式分解 $8a^3+27b^3$.

练习 1.7 已知 $x+\dfrac{1}{x}=4$, 求 $x^2+\dfrac{1}{x^2}$ 与 $x^3+\dfrac{1}{x^3}$.

练习 1.8 因式分解 x^4-16 与 x^4+4x^2+16.

练习 1.9 因式分解 $2x^2-5x-3$ 与 $6x^2+5xy-6y^2$.

知乎@余命数

练习 1.10 对多项式 $P(x)=2x^3-x^2-8x+4$, 写出所有可能的有理根, 并完成因式分解.

1.3 恒等变形的方法

恒等变形用来证明等式与化简代数式. 我们给出五种策略, 分别对应不同的结构线索.

1.3.1 复杂化简法

我们证明恒等式 $A=B$ 时, 从结构复杂的一端入手, 经因式分解、通分、展开、配方等变形, 化到结构简单的一端.

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 1.27　在各分母均有意义时, 求证

$$\frac{b-c}{(a-b)(a-c)}+\frac{c-a}{(b-c)(b-a)}+\frac{a-b}{(c-a)(c-b)}=\frac{2}{a-b}+\frac{2}{b-c}+\frac{2}{c-a}$$

证明　左侧具有轮换对称性, 右侧为简单分式之和, 我们对左侧作裂项. 将分子表为分母因式之差, 得

$$b-c=(a-c)-(a-b)$$

知乎@余命数

于是

$$\frac{(a-c)-(a-b)}{(a-b)(a-c)}=\frac{1}{a-b}-\frac{1}{a-c}=\frac{1}{a-b}+\frac{1}{c-a}$$

由轮换对称性, 另两项同理处理, 得

知乎@余命数

$$\frac{c-a}{(b-c)(b-a)}=\frac{1}{b-c}+\frac{1}{a-b}$$

$$\frac{a-b}{(c-a)(c-b)}=\frac{1}{c-a}+\frac{1}{b-c}$$

三式相加, 每个分式恰出现两次, 得

知乎@余命数　知乎@余命数

$$\begin{aligned}\text{LHS}&=\left(\frac{1}{a-b}+\frac{1}{c-a}\right)+\left(\frac{1}{b-c}+\frac{1}{a-b}\right)+\left(\frac{1}{c-a}+\frac{1}{b-c}\right)\\&=\frac{2}{a-b}+\frac{2}{b-c}+\frac{2}{c-a}=\text{RHS}\end{aligned}$$

知乎@余命数

知乎@余命数　□

例题 1.28　因式分解 $(x^2+x-1)(x^2+x+3)-5$.

解　x^2+x 反复出现, 我们作整体换元. 令 $A=x^2+x$, 原式化为

$$(A-1)(A+3)-5=A^2+2A-3-5=A^2+2A-8$$

对 A 的二次式作因式分解, 得

$$A^2+2A-8=(A+4)(A-2)$$

代回 $A=x^2+x$, 得

$$(x^2+x+4)(x^2+x-2)$$

知乎@余命数

因 x^2+x+4 判别式 $\Delta = 1-16 < 0$, 在实数域不可约; 而 $x^2+x-2 = (x+2)(x-1)$. 故

$$\text{原式} = (x^2+x+4)(x+2)(x-1)$$

例题 1.29 化简 $\dfrac{x^4+x^2y^2+y^4}{x^2-xy+y^2}$.

解 分母形如 x^2-xy+y^2, 我们将分子分解为含此因式的乘积. 为构造平方差, 对分子配方, 得

$$\begin{aligned} x^4+x^2y^2+y^4 &= (x^4+2x^2y^2+y^4)-x^2y^2 \\ &= (x^2+y^2)^2-(xy)^2 \\ &= (x^2+y^2-xy)(x^2+y^2+xy) \end{aligned}$$

代入原式, 约去公因式, 得

$$\frac{(x^2-xy+y^2)(x^2+xy+y^2)}{x^2-xy+y^2} = x^2+xy+y^2 \quad (x^2-xy+y^2 \neq 0)$$

例题 1.30 求证 $(a+b+c)^3-(a^3+b^3+c^3) = 3(a+b)(b+c)(c+a)$.

证明 右侧为三个一次因式之积, 我们用因式定理验证左侧的根. 令 $P(a) = (a+b+c)^3-a^3-b^3-c^3$, 代入 $a=-b$ 得

$$P(-b) = c^3-(-b)^3-b^3-c^3 = 0$$

故 $(a+b) \mid P(a)$. 由轮换对称性, $(b+c)$ 与 $(c+a)$ 亦整除 P.

P 是三次齐次式, 三个一次因式之积亦为三次齐次式, 故

$$P(a) = k(a+b)(b+c)(c+a)$$

取 $a=b=c=1$ 确定系数, 得

$$\text{LHS} = 3^3-3 = 24, \quad \text{RHS} = k\cdot 2\cdot 2\cdot 2 = 8k$$

解得 $k=3$. □

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 1.31　已知 $abc=1$, 且各分母均有意义, 求证 $\dfrac{1}{1+a+ab}+\dfrac{1}{1+b+bc}+\dfrac{1}{1+c+ca}=1$.

证明　三个分母结构相似但不同, 条件 $abc=1$ 尚未使用. 我们用此条件统一分母.

第二项乘 $\dfrac{a}{a}$ 并利用 $abc=1$, 得

$$\frac{1}{1+b+bc}=\frac{a}{a+ab+abc}=\frac{a}{a+ab+1}$$

第三项乘 $\dfrac{ab}{ab}$, 得

$$\frac{1}{1+c+ca}=\frac{ab}{ab+abc+a^2bc}=\frac{ab}{ab+1+a}$$

三项分母统一为 $1+a+ab$, 分子相加, 得

$$\frac{1}{1+a+ab}+\frac{a}{1+a+ab}+\frac{ab}{1+a+ab}=\frac{1+a+ab}{1+a+ab}=1$$

□

例题 1.32　已知 $a+b+c=0$, 且各分母均有意义, 求证 $\dfrac{a^2}{2a^2+bc}+\dfrac{b^2}{2b^2+ca}+\dfrac{c^2}{2c^2+ab}=1$.

证明　分母 $2a^2+bc$ 形式特殊, 我们利用 $a=-(b+c)$ 将其因式分解, 得

$$2a^2+bc=a^2+a\cdot a+bc=a^2-a(b+c)+bc=(a-b)(a-c)$$

由轮换对称性, $2b^2+ca=(b-c)(b-a)$, $2c^2+ab=(c-a)(c-b)$.

通分后公分母为 $(a-b)(b-c)(c-a)$, 分子为

$$\begin{aligned}&-a^2(b-c)-b^2(c-a)-c^2(a-b)\\&=-a^2b+a^2c-b^2c+ab^2-c^2a+bc^2\end{aligned}$$

此式关于 a,b,c 轮换反对称(交换任意两变量变号), 且为三次齐次式, 故必有因式 $(a-b)(b-c)(c-a)$. 比较首项系数知

$$-a^2(b-c)-b^2(c-a)-c^2(a-b)=(a-b)(b-c)(c-a)$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

与分母相消, 得 1. □

例题 1.33 因式分解 $x^4+y^4+(x+y)^4$.

解 我们直接展开, 再看能否配成完全平方. 展开 $(x+y)^4$, 得

$$\begin{aligned} x^4+y^4+(x+y)^4 &= x^4+y^4+(x^4+4x^3y+6x^2y^2+4xy^3+y^4) \\ &= 2x^4+4x^3y+6x^2y^2+4xy^3+2y^4 \\ &= 2(x^4+2x^3y+3x^2y^2+2xy^3+y^4) \end{aligned}$$

括号内为四次齐次对称式. 我们猜测它是某二次齐次式的平方, 设为 $(x^2+\alpha xy+y^2)^2$. 展开比较系数, 得

$$\begin{aligned} (x^2+xy+y^2)^2 &= x^4+x^2y^2+y^4+2x^3y+2xy^3+2x^2y^2 \\ &= x^4+2x^3y+3x^2y^2+2xy^3+y^4 \end{aligned}$$

吻合. 故原式 $=2(x^2+xy+y^2)^2$.

1.3.2 双向变换法

当等式两端均较复杂时, 我们分别把它们变形到同一形式. 若 $A=C$ 且 $B=C$, 则 $A=B$.

例题 1.34 求证 $(a^2+b^2)(c^2+d^2)=(ac+bd)^2+(ad-bc)^2$.

证明 两侧均为四个变量的四次式, 我们直接展开比较. 左侧展开为

$$\mathrm{LHS}=a^2c^2+a^2d^2+b^2c^2+b^2d^2$$

右侧展开为

$$\begin{aligned} \mathrm{RHS} &= (a^2c^2+2abcd+b^2d^2)+(a^2d^2-2abcd+b^2c^2) \\ &= a^2c^2+b^2d^2+a^2d^2+b^2c^2 \end{aligned}$$

交叉项 $\pm 2abcd$ 相消, 两侧相等. □

例题 1.35 求证 $a^3+b^3+c^3-3abc=\frac{1}{2}(a+b+c)[(a-b)^2+(b-c)^2+(c-a)^2]$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 左侧为已知的因式分解公式, 右侧较复杂. 我们从右侧入手化简, 得

$$\begin{aligned}\text{RHS} &= \frac{1}{2}(a+b+c)(2a^2+2b^2+2c^2-2ab-2bc-2ca)\\ &= (a+b+c)(a^2+b^2+c^2-ab-bc-ca)\end{aligned}$$

此即 $a^3+b^3+c^3-3abc$ 的标准分解式. 展开验证, 得

$$\begin{aligned}&= a^3+ab^2+ac^2-a^2b-abc-a^2c\\ &\quad +a^2b+b^3+bc^2-ab^2-b^2c-abc\\ &\quad +a^2c+b^2c+c^3-abc-bc^2-ac^2\\ &= a^3+b^3+c^3-3abc = \text{LHS}\end{aligned}$$

□

例题 1.36 求证 $(a+b+c)^3abc-(bc+ca+ab)^3 = abc(a^3+b^3+c^3)-(b^3c^3+c^3a^3+a^3b^3)$.

证明 两侧结构相似, 左侧涉及 $(a+b+c)^3$ 与 $(bc+ca+ab)^3$, 右侧涉及 $a^3+b^3+c^3$ 与 $(bc)^3+(ca)^3+(ab)^3$. 移项使结构对称, 得

$$abc[(a+b+c)^3-(a^3+b^3+c^3)] = (bc+ca+ab)^3-[(bc)^3+(ca)^3+(ab)^3]$$

利用恒等式 $(x+y+z)^3-(x^3+y^3+z^3) = 3(x+y)(y+z)(z+x)$. 左侧展开为

$$\text{左式} = abc\cdot 3(a+b)(b+c)(c+a)$$

右侧取 $x=bc$, $y=ca$, $z=ab$, 得

$$\begin{aligned}\text{右式} &= 3(bc+ca)(ca+ab)(ab+bc)\\ &= 3\cdot c(a+b)\cdot a(b+c)\cdot b(a+c)\\ &= 3abc(a+b)(b+c)(c+a)\end{aligned}$$

两侧相等. □

例题 1.37 求证 $a^2(b+c)+b^2(c+a)+c^2(a+b)+3abc = (a+b+c)(ab+bc+ca)$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

证明 两侧复杂度相当, 我们直接展开. 左侧展开为

$$a^2b+a^2c+b^2c+b^2a+c^2a+c^2b+3abc$$

右侧展开为

$$\begin{aligned}&=a(ab+bc+ca)+b(ab+bc+ca)+c(ab+bc+ca)\\&=(a^2b+abc+a^2c)+(ab^2+b^2c+abc)+(abc+bc^2+c^2a)\\&=a^2b+a^2c+ab^2+b^2c+c^2a+bc^2+3abc\end{aligned}$$

两侧相等. □

例题 1.38 求证 $2a^2b^2+2b^2c^2+2c^2a^2-a^4-b^4-c^4=(a+b+c)(a+b-c)(a-b+c)(-a+b+c)$.

证明 先记

$$T=a^4+b^4+c^4-2a^2b^2-2b^2c^2-2c^2a^2.$$

四个因式形如 $\pm a\pm b\pm c$, 我们将 T 配方为平方差的乘积.

$$\begin{aligned}T&=(a^4+2a^2b^2+b^4)-4a^2b^2-c^4-2b^2c^2-2c^2a^2+2a^2b^2\\&=(a^2+b^2)^2-2(a^2+b^2)c^2+c^4-4a^2b^2\\&=(a^2+b^2-c^2)^2-4a^2b^2\end{aligned}$$

应用平方差公式两次, 得

$$\begin{aligned}&=(a^2+b^2-c^2-2ab)(a^2+b^2-c^2+2ab)\\&=((a-b)^2-c^2)((a+b)^2-c^2)\\&=(a-b-c)(a-b+c)(a+b-c)(a+b+c)\end{aligned}$$

注意 $-a+b+c=-(a-b-c)$, 故

$$\begin{aligned}&2a^2b^2+2b^2c^2+2c^2a^2-a^4-b^4-c^4\\&=-T=(a+b+c)(a+b-c)(a-b+c)(-a+b+c).\end{aligned}$$

□

例题 1.39 (索菲·热尔曼恒等式) 求证 $a^4+4b^4=(a^2+2b^2+2ab)(a^2+2b^2-2ab)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 右侧形如 $(M+N)(M-N)=M^2-N^2$, 其中 $M=a^2+2b^2$, $N=2ab$. 展开验证, 得

$$\begin{aligned}\text{RHS} &= (a^2+2b^2)^2-(2ab)^2\\ &= a^4+4a^2b^2+4b^4-4a^2b^2\\ &= a^4+4b^4=\text{LHS}\end{aligned}$$

□

例题 1.40 (拉格朗日恒等式) 求证 $\left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right)=\left(\sum_i a_ib_i\right)^2+\sum_{i<j}(a_ib_j-a_jb_i)^2$.

证明 我们对有限指标集合展开右侧, 得

$$\left(\sum_i a_ib_i\right)^2=\sum_i a_i^2b_i^2+2\sum_{i<j}a_ia_jb_ib_j,$$

且

$$\sum_{i<j}(a_ib_j-a_jb_i)^2=\sum_{i<j}(a_i^2b_j^2-2a_ia_jb_ib_j+a_j^2b_i^2).$$

两式相加后交叉项抵消, 得

$$\sum_i a_i^2b_i^2+\sum_{i<j}(a_i^2b_j^2+a_j^2b_i^2)=\sum_{i,j}a_i^2b_j^2.$$

而

$$\sum_{i,j}a_i^2b_j^2=\left(\sum_i a_i^2\right)\left(\sum_j b_j^2\right),$$

故恒等式成立. □

1.3.3 作差法

我们利用 $A=B\iff A-B=0$. 当等式两侧结构有局部对应关系时, 逐项作差往往比通分简洁.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 1.41 在各分母均有意义时,求证 $\dfrac{x}{ax-a^2}+\dfrac{y}{ay-a^2}+\dfrac{z}{az-a^2}=\dfrac{1}{x-a}+\dfrac{1}{y-a}+\dfrac{1}{z-a}+\dfrac{3}{a}$.

知乎@余命数

证明 左右两侧均有三个分式,结构对应. 我们逐项作差,得

$$\frac{x}{ax-a^2}-\frac{1}{x-a}=\frac{x}{a(x-a)}-\frac{a}{a(x-a)}=\frac{x-a}{a(x-a)}=\frac{1}{a}$$

知乎@余命数

另两组差亦为 $\dfrac{1}{a}$. 故

$$\text{LHS}-\left(\frac{1}{x-a}+\frac{1}{y-a}+\frac{1}{z-a}\right)=\frac{3}{a}$$

移项即得原式. □

例题 1.42 求证 $(a+b)(b+c)(c+a)=a^2(b+c)+b^2(c+a)+c^2(a+b)+2abc$.

证明 左侧为三因式之积,右侧已部分展开. 直接展开左侧比较,得

$$\begin{aligned}(a+b)(b+c)(c+a)&=(ab+ac+b^2+bc)(c+a)\\&=abc+a^2b+ac^2+a^2c+b^2c+ab^2+bc^2+abc\\&=a^2(b+c)+b^2(a+c)+c^2(a+b)+2abc\end{aligned}$$

知乎@余命数

知乎@余命数

与右侧相同. □

例题 1.43 已知 $a+b+c=0$,求证 $a^3+b^3+c^3=3abc$.

证明 我们利用标准分解式

$$a^3+b^3+c^3-3abc=(a+b+c)(a^2+b^2+c^2-ab-bc-ca)$$

代入条件 $a+b+c=0$,右边为 0. 故 $a^3+b^3+c^3=3abc$. □

例题 1.44 已知 $a+b+c=0$,且各分母均有意义,求证 $\dfrac{1}{b^2+c^2-a^2}+\dfrac{1}{c^2+a^2-b^2}+\dfrac{1}{a^2+b^2-c^2}=0$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 我们利用条件简化分母. 由 $a=-(b+c)$ 得

$$b^2+c^2-a^2=b^2+c^2-(b+c)^2=-2bc$$

由轮换对称性, 另两分母为 $-2ca$ 与 $-2ab$. 故

$$\text{LHS}=\frac{1}{-2bc}+\frac{1}{-2ca}+\frac{1}{-2ab}=-\frac{1}{2abc}(a+b+c)=0$$

知乎@余命数 □

例题 1.45 已知 $\dfrac{a}{x}=\dfrac{b}{y}=\dfrac{c}{z}$, 且 $x,y,z\neq 0$, 求证 $(a^2+b^2+c^2)(x^2+y^2+z^2)=(ax+by+cz)^2$.

证明 由等比条件, 我们设公比. 令 $\dfrac{a}{x}=\dfrac{b}{y}=\dfrac{c}{z}=k$, 则 $a=kx$, $b=ky$, $c=kz$. 代入, 得

$$\text{LHS}=k^2(x^2+y^2+z^2)\cdot(x^2+y^2+z^2)=k^2(x^2+y^2+z^2)^2$$

$$\text{RHS}=(kx^2+ky^2+kz^2)^2=k^2(x^2+y^2+z^2)^2$$

两侧相等. □

例题 1.46 在各分母均有意义时, 求证 $\dfrac{a^2-bc}{(a+b)(a+c)}+\dfrac{b^2-ca}{(b+c)(b+a)}+\dfrac{c^2-ab}{(c+a)(c+b)}=0$.

证明 左侧为轮换对称式, 我们通分验证分子为零. 公分母为 $(a+b)(b+c)(c+a)$, 分子为

$$(a^2-bc)(b+c)+(b^2-ca)(c+a)+(c^2-ab)(a+b)$$

展开, 得

$$\begin{aligned}&=(a^2b+a^2c-b^2c-bc^2)+(b^2c+ab^2-ac^2-a^2c)\\&\quad+(ac^2+bc^2-a^2b-ab^2)\\&=0\end{aligned}$$

各项两两抵消. □

知乎@余命数

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1.3.4 分析法

我们从待证结论出发, 逐步归结到已知条件. 正式书写时把推导过程整理为综合形式.

例题 1.47 已知 $a+b+c=0$, 求证 $(a^2+b^2+c^2)^2=2(a^4+b^4+c^4)$.

证明 我们展开左侧, 待证式等价于

$$a^4+b^4+c^4+2(a^2b^2+b^2c^2+c^2a^2)=2(a^4+b^4+c^4)$$

知乎@余命数

即

$$a^4+b^4+c^4=2(a^2b^2+b^2c^2+c^2a^2)(*)$$

现从条件推导 $(*)$. 由 $(a+b+c)^2=0$ 得

$$a^2+b^2+c^2=-2(ab+bc+ca)$$

知乎@余命数

两边平方, 得

$$(a^2+b^2+c^2)^2=4(ab+bc+ca)^2$$

展开两侧, 得

$$a^4+b^4+c^4+2(a^2b^2+b^2c^2+c^2a^2)=4[a^2b^2+b^2c^2+c^2a^2+2abc(a+b+c)]$$

知乎@余命数 知乎@余命数 知乎@余命数

因 $a+b+c=0$, 右边末项消失, 得

$$a^4+b^4+c^4+2(a^2b^2+b^2c^2+c^2a^2)=4(a^2b^2+b^2c^2+c^2a^2)$$

移项即得 $(*)$. □

例题 1.48 求证 $a^2+b^2+c^2-ab-bc-ca=\frac{1}{2}[(a-b)^2+(b-c)^2+(c-a)^2]$.

证明 右侧为完全平方之和, 展开后应与左侧吻合, 得

$$\begin{aligned}\text{RHS}&=\frac{1}{2}(a^2-2ab+b^2+b^2-2bc+c^2+c^2-2ca+a^2)\\&=\frac{1}{2}(2a^2+2b^2+2c^2-2ab-2bc-2ca)\\&=a^2+b^2+c^2-ab-bc-ca=\text{LHS}\end{aligned}$$

□

知乎@余命数

1.3.5 赋值法（多项式恒等定理）

定理 1.7　若两个次数不超过 n 的多项式 $P(x)$ 与 $Q(x)$ 在 $n+1$ 个不同点处取值相同，则 $P(x)\equiv Q(x)$.

证明　令 $R(x)=P(x)-Q(x)$. 则 $R(x)$ 的次数不超过 n，且它在这 $n+1$ 个不同点处均取值为 0. 若 R 不是零多项式，则由因式定理可知这些不同的零点对应 $n+1$ 个不同的一次因式，因而 R 的次数至少为 $n+1$，与 $\deg R\le n$ 矛盾. 故 $R\equiv 0$，即 $P(x)\equiv Q(x)$. □

知乎@余命数

此定理将恒等式的证明化为有限次数值验证.

例题 1.49　求证 $(x+1)(x+2)(x+3)(x+4)+1=(x^2+5x+5)^2$.

证明　设左边为 $P(x)$，右边为 $Q(x)$，均为 4 次多项式. 我们取 5 个点验证，得

- $x=-1$：$P(-1)=0\cdot 1\cdot 2\cdot 3+1=1$，$Q(-1)=(1-5+5)^2=1$.
- $x=-2$：$P(-2)=(-1)\cdot 0\cdot 1\cdot 2+1=1$，$Q(-2)=(4-10+5)^2=1$.
- $x=-3$：$P(-3)=(-2)(-1)\cdot 0\cdot 1+1=1$，$Q(-3)=(9-15+5)^2=1$.
- $x=-4$：$P(-4)=(-3)(-2)(-1)\cdot 0+1=1$，$Q(-4)=(16-20+5)^2=1$.
- $x=0$：$P(0)=24+1=25$，$Q(0)=25$.

知乎@余命数

5 个点处取值相同，由定理知 $P(x)\equiv Q(x)$. □

知乎@余命数

1.4 二次函数

设 $f(x)=ax^2+bx+c$，其中 $a\neq 0$ 为二次函数. 它的图像是抛物线.

1.4.0.0.1　配方法与顶点式　我们对一般式配方，把它改写成顶点式.

$$\begin{aligned} f(x)&=a\left(x^2+\frac{b}{a}x\right)+c\\ &=a\left(x^2+\frac{b}{a}x+\frac{b^2}{4a^2}-\frac{b^2}{4a^2}\right)+c\\ &=a\left(x+\frac{b}{2a}\right)^2+\frac{4ac-b^2}{4a} \end{aligned}$$

于是我们得到 $f(x)=a(x-h)^2+k$.

知乎@余命数

- 顶点坐标 $(h,k)=\left(-\dfrac{b}{2a},\dfrac{4ac-b^2}{4a}\right)$. 当 $a>0$ 时它是抛物线的最低点; 当 $a<0$ 时它是抛物线的最高点.
- 对称轴方程 $x=h=-\dfrac{b}{2a}$.

1.4.0.0.2 系数的几何意义 设二次项系数为 a, 一次项系数为 b, 常数项为 c. 它们的几何意义如下.

- a 决定开口方向与宽窄. 当 $a>0$ 时开口向上; 当 $a<0$ 时开口向下.$|a|$ 越大, 开口越窄.
- b 与 a 共同决定对称轴位置. 当 a,b 同号时对称轴在 y 轴左侧; 当 a,b 异号时对称轴在 y 轴右侧.
- c 是抛物线与 y 轴交点的纵坐标.

1.4.0.0.3 与 x 轴的交点 抛物线与 x 轴交点的横坐标就是方程 $ax^2+bx+c=0$ 的实根. 实根的个数由判别式 $\Delta=b^2-4ac$ 的符号决定.

- 若 $\Delta>0$, 则抛物线与 x 轴有两个不同交点.
- 若 $\Delta=0$, 则抛物线与 x 轴有唯一切点(重根).
- 若 $\Delta<0$, 则抛物线与 x 轴无交点.

1.4.1 韦达定理及其推广

设 x_1,x_2 是方程 $ax^2+bx+c=0\ (a\neq 0)$ 的两个根. 由因式分解

$$ax^2+bx+c=a(x-x_1)(x-x_2)=ax^2-a(x_1+x_2)x+ax_1x_2$$

比较系数, 我们得到

$$x_1+x_2=-\frac{b}{a},\quad x_1x_2=\frac{c}{a}$$

利用韦达定理, 我们可以在不解方程的前提下计算根的对称表达式. 一些常用变形如下.

- $x_1^2+x_2^2=(x_1+x_2)^2-2x_1x_2$
- $(x_1-x_2)^2=(x_1+x_2)^2-4x_1x_2=\dfrac{\Delta}{a^2}$
- $x_1^3+x_2^3=(x_1+x_2)[(x_1+x_2)^2-3x_1x_2]$
- $\dfrac{1}{x_1}+\dfrac{1}{x_2}=\dfrac{x_1+x_2}{x_1x_2}$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

1.4.1.1 推广至 n 次方程

设 $x_1,\dots,x_n$ 是方程 $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0\ (a_n\neq 0)$ 的 n 个根. 根与系数满足下述对称关系.

$$\sum_i x_i=-\frac{a_{n-1}}{a_n}$$

$$\sum_{i<j} x_ix_j=\frac{a_{n-2}}{a_n}$$

$$\sum_{i<j<k} x_ix_jx_k=-\frac{a_{n-3}}{a_n}$$

$$\vdots$$

$$x_1x_2\cdots x_n=(-1)^n\frac{a_0}{a_n}$$

一般地, 任取 k 个根相乘再求和, 所得结果为 $(-1)^k\dfrac{a_{n-k}}{a_n}$.

对于三次方程 $ax^3+bx^2+cx+d=0$, 设其根为 x_1,x_2,x_3, 则它们满足

$$x_1+x_2+x_3=-\frac{b}{a},\quad x_1x_2+x_2x_3+x_3x_1=\frac{c}{a},\quad x_1x_2x_3=-\frac{d}{a}$$

例题 1.50 设 α,β 是方程 $x^2-3x-5=0$ 的两根, 求以 $\dfrac{1}{\alpha-1}$ 和 $\dfrac{1}{\beta-1}$ 为根的一元二次方程.

解 我们设新根为 $y_1=\dfrac{1}{\alpha-1}, y_2=\dfrac{1}{\beta-1}$. 由原方程可知 $\alpha+\beta=3, \alpha\beta=-5$.

计算新根之和, 得

$$y_1+y_2=\frac{(\alpha-1)+(\beta-1)}{(\alpha-1)(\beta-1)}=\frac{(\alpha+\beta)-2}{\alpha\beta-(\alpha+\beta)+1}=\frac{3-2}{-5-3+1}=-\frac{1}{7}$$

计算新根之积, 得

$$y_1y_2=\frac{1}{(\alpha-1)(\beta-1)}=\frac{1}{\alpha\beta-(\alpha+\beta)+1}=-\frac{1}{7}$$

于是所求方程为 $y^2+\dfrac{1}{7}y-\dfrac{1}{7}=0$, 即 $7y^2+y-1=0$.

为验证这一结果, 我们由 $y=\dfrac{1}{x-1}$ 得 $x=\dfrac{y+1}{y}$. 代入原方程, 得

$$\left(\frac{y+1}{y}\right)^2-3\left(\frac{y+1}{y}\right)-5=0$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

两边乘 y^2 并整理, 得

$$(y+1)^2-3y(y+1)-5y^2=-7y^2-y+1=0$$

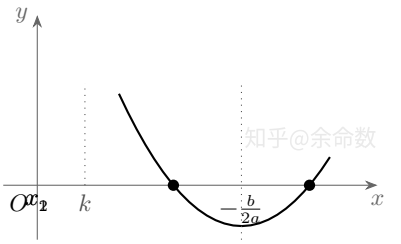
即 $7y^2+y-1=0$.

1.4.2 一元二次方程根的分布

设 $f(x)=ax^2+bx+c\ (a>0)$. 在根的分布问题中, 我们把根相对于给定值的位置关系翻译成关于系数的不等式组. 约束来自以下三点.

1. 判别式 $\Delta=b^2-4ac$ 控制根的存在性.
2. 对称轴 $x=-\dfrac{b}{2a}$ 控制根的整体位置.
3. 端点函数值 $f(k)$ 控制特定点与根的相对位置.

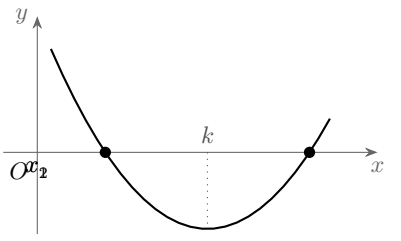
1.4.2.1 两根均大于 k



由图可知, 抛物线与 x 轴有交点, 对称轴在 k 的右侧, 且点 $(k,f(k))$ 在 x 轴上方. 因此所需条件为

$$\begin{cases}\Delta\geq 0\\ -\dfrac{b}{2a}>k\\ f(k)>0\end{cases}$$

1.4.2.2 一根大于 k, 一根小于 k

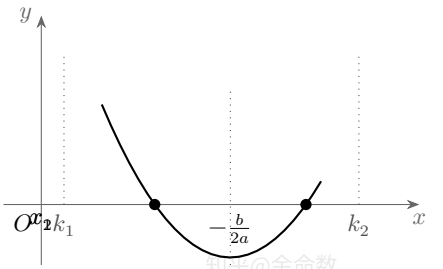


点 $(k,f(k))$ 在 x 轴下方. 由于抛物线开口向上, $f(k)<0$ 已经蕴含 $\Delta>0$. 于是条件简化为

$$f(k)<0$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1.4.2.3 两根均在区间 (k_1,k_2) 内



知乎@余命数

同理, 两根都在 (k_1,k_2) 内时, 条件可写成

$$\begin{cases}\Delta\geq 0\\ k_1<-\dfrac{b}{2a}<k_2\\ f(k_1)>0\\ f(k_2)>0\end{cases}$$

例题 1.51　若方程 $x^2-(m+1)x+m=0$ 的两根均在区间 $(0,2)$ 内, 求 m 的取值范围.

解 设 $f(x)=x^2-(m+1)x+m$.

判别式 $\Delta=(m+1)^2-4m=(m-1)^2\geq 0$ 恒成立.

对称轴满足 $0<\dfrac{m+1}{2}<2$, 解得 $-1<m<3$.

端点函数值需 $f(0)=m>0$ 与 $f(2)=4-2(m+1)+m=2-m>0$, 即 $m<2$.

取交集得 $m\in(0,2)$.

例题 1.52　若方程 $x^2-2mx+m^2-1=0$ 至少有一个根在区间 $(0,2)$ 内, 求 m 的取值范围.

解 我们考虑补集, 即两根均不在 $(0,2)$ 内.

方程化为 $(x-m)^2=1$, 解得 $x=m\pm 1$, 即两根为 $m-1$ 与 $m+1$.

两根均 ≤ 0 即 $m+1\leq 0$, 亦即 $m\leq -1$.

两根均 ≥ 2 即 $m-1\geq 2$, 亦即 $m\geq 3$.

两根分居 0 与 2 两侧, 即 $m-1\leq 0\leq 2\leq m+1$, 亦即 $m=1$.

补集为 $(-\infty,-1]\cup\{1\}\cup[3,+\infty)$, 故原问题的解为 $m\in(-1,3)\setminus\{1\}$.

例题 1.53　已知方程 $x^2+(k-3)x+k=0$ 有两个不同实根 x_1,x_2, 满足 $x_1<1<x_2<4$. 求 k 的取值范围.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解 设 $f(x)=x^2+(k-3)x+k$.

由 $x_1<1<x_2$ 可知 $f(1)<0$, 即

$$1+(k-3)+k=2k-2<0 \implies k<1$$

由 $x_2<4$ 可知 $f(4)>0$, 即

$$16+4(k-3)+k=5k+4>0 \implies k>-\frac{4}{5}$$

对称轴 $\dfrac{3-k}{2}<4$ 给出 $k>-5$, 这个条件已被 $k>-\dfrac{4}{5}$ 蕴含.

取交集得 $k\in\left(-\dfrac{4}{5},1\right)$.

知乎@余命数

例题 1.54 求使不等式 $ax^2-2x+1>0$ 对所有 $x\in(0,2]$ 恒成立的 a 的取值范围.

解 设 $f(x)=ax^2-2x+1$. 我们要求 f 在 $(0,2]$ 上的最小值大于 0.

若 $a=0$, 则 $f(x)=-2x+1$ 在 $x=1$ 处取负值, 不满足要求.

若 $a<0$, 则 $f(2)=4a-3<-3<0$, 不满足要求.

若 $a>0$, 则对称轴为 $x_0=\dfrac{1}{a}$.

知乎@余命数

- 当 $x_0\geq 2$, 即 $0<a\leq\dfrac{1}{2}$ 时, f 在 $(0,2]$ 上递减, 最小值为 $f(2)=4a-3$. 要求 $4a-3>0$, 即 $a>\dfrac{3}{4}$, 与 $a\leq\dfrac{1}{2}$ 矛盾.
- 当 $0<x_0<2$, 即 $a>\dfrac{1}{2}$ 时, 最小值为 $f\left(\dfrac{1}{a}\right)=1-\dfrac{1}{a}$. 要求 $1-\dfrac{1}{a}>0$, 即 $a>1$.

故 $a\in(1,+\infty)$.

例题 1.55 若方程 $x^2-2a|x|+a^2-4=0$ 有四个不相等的实根, 求 a 的取值范围.

解 令 $t=|x|$, 其中 $t\geq 0$, 则方程化为

$$t^2-2at+a^2-4=0$$

知乎@余命数

每个正根 t_0 对应两个 x 值 $\pm t_0$. 因此, 原方程有四个不同实根当且仅当关于 t 的方程有两个不相等的正根.

设 $g(t)=t^2-2at+a^2-4$.

判别式 $\Delta=4a^2-4(a^2-4)=16>0$, 恒有两个不同实根.

两根之和 $2a>0$, 即 $a>0$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两根之积 $a^2-4>0$, 即 $|a|>2$.

取交集得 $a\in(2,+\infty)$.

例题 1.56　已知 $f(x)=x^2+ax+1$. 若方程 $f(x)=0$ 在区间 $(0,1)$ 内有唯一解, 求 a 的取值范围.

解　唯一解只有两种可能. 一种是重根落在 $(0,1)$ 内, 另一种是两根中恰有一个落在 $(0,1)$ 内.

先考虑重根情形.$\Delta=a^2-4=0$ 给出 $a=\pm2$. 当 $a=2$ 时重根为 -1; 当 $a=-2$ 时重根为 1. 它们都不在开区间 $(0,1)$ 内.

再考虑单根情形. 我们需 $f(0)$ 与 $f(1)$ 异号. 由 $f(0)=1>0$ 可知 $f(1)=a+2<0$, 因此 $a<-2$.

再验证端点. 若 $f(1)=0$, 即 $a=-2$, 则它属于重根情形, 已被排除.

故 $a\in(-\infty,-2)$.

1.5 计算与变形技巧

1.5.1 配方法

配方法把二次式写成完全平方式与常数之和. 我们分两步做. 先提取二次项系数, 再加减一次项系数一半的平方.

练习 1.11　求函数 $y=2x^2-8x+5$ 的最小值.

例题 1.57　求代数式 $f(x,y)=x^2+4y^2-4x+8y+10$ 的最小值, 其中 $x,y\in\mathbb{R}$.

解　这个表达式含两个变量, 我们分别对 x 与 y 配方. 先将与 x 和 y 相关的项分组.

$$f(x,y)=(x^2-4x)+(4y^2+8y)+10$$

对关于 x 的部分配方, 得

$$x^2-4x=(x^2-4x+4)-4=(x-2)^2-4$$

对关于 y 的部分配方, 注意先提取二次项系数, 得

$$4y^2+8y=4(y^2+2y)=4(y^2+2y+1-1)=4(y+1)^2-4$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

将配方后的结果代回原式, 得

$$\begin{aligned} f(x,y) &= [(x-2)^2-4]+[4(y+1)^2-4]+10 \\ &= (x-2)^2+4(y+1)^2+2 \end{aligned}$$

由于平方项的非负性,有 $(x-2)^2 \geq 0$ 和 $(y+1)^2 \geq 0$. 因此, $f(x,y) \geq 0+4(0)+2=2$. 当且仅当 $(x-2)^2=0$ 且 $(y+1)^2=0$, 即 $x=2, y=-1$ 时, 等号成立. 故该代数式的最小值为 2.

知乎@余命数

例题 1.58 证明不等式 $a^2+b^2+c^2 \geq ab+bc+ca$ 对任意实数 a,b,c 恒成立.

解 移项后乘以 2 再配方, 得

$$a^2+b^2+c^2-ab-bc-ca \geq 0$$

等价于

$$2a^2+2b^2+2c^2-2ab-2bc-2ca \geq 0.$$

拆项组合得

$$\begin{aligned} &(a^2-2ab+b^2)+(b^2-2bc+c^2)+(c^2-2ca+a^2) \geq 0 \\ &(a-b)^2+(b-c)^2+(c-a)^2 \geq 0. \end{aligned}$$

知乎@余命数 知乎@余命数

实数的平方非负, 三非负数之和也非负, 故原不等式成立. 等号成立当且仅当 $a-b=0, b-c=0, c-a=0$, 即 $a=b=c$.

知乎@余命数

例题 1.59 求 $y=\dfrac{12}{x^2-8x+20}$ 的最大值.

解 这是一个分式函数, 分子为正常数. 函数取最大值, 等价于分母取最小值. 我们对分母配方.

$$\begin{aligned} \text{分母} &= x^2-8x+20 \\ &= (x^2-8x+16)-16+20 \\ &= (x-4)^2+4 \end{aligned}$$

由于 $(x-4)^2 \geq 0$ 对所有实数 x 成立, 所以分母的最小值为 $0+4=4$, 此最小值在 $x=4$ 时取得. 因此, 函数 y 的最大值为

知乎@余命数

$$y_{\max}=\frac{12}{4}=3$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

该最大值在 $x=4$ 时取得.

例题 1.60　求方程 $x^2+y^2-8x+6y+21=0$ 的所有整数解 (x,y).

解 配方, 得

$$(x^2-8x)+(y^2+6y)+21=0$$

$$(x^2-8x+16)-16+(y^2+6y+9)-9+21=0$$

$$(x-4)^2+(y+3)^2-25+21=0$$

$$(x-4)^2+(y+3)^2=4$$

由于 x,y 都是整数, 那么 $x-4$ 和 $y+3$ 也必为整数. 问题转化为求两个整数的平方和为 4. 设 $A=x-4, B=y+3$, 则 $A,B\in\mathbb{Z}$. 方程为 $A^2+B^2=4$. 整数平方的序列为 $0,1,4,9,\ldots$. 为使两个平方数的和为 4, 只有两种可能(不计顺序).

情形一为 $A^2=0$ 且 $B^2=4$. $A=0 \implies x-4=0 \implies x=4$. $B^2=4 \implies B=\pm2 \implies y+3=\pm2 \implies y=-1$ 或 $y=-5$. 得到两组解 $(4,-1)$ 和 $(4,-5)$.

情形二为 $A^2=4$ 且 $B^2=0$. $A^2=4 \implies A=\pm2 \implies x-4=\pm2 \implies x=6$ 或 $x=2$. $B=0 \implies y+3=0 \implies y=-3$. 得到两组解 $(6,-3)$ 和 $(2,-3)$.

故该方程共有四组整数解 $(4,-1),(4,-5),(6,-3),(2,-3)$.

1.5.2 换元法

表达式中某一部分反复出现时, 我们引入新变量替换, 并先确定新变量的取值范围.

练习 1.12　解方程 $4^x-3\cdot2^x+2=0$.

例题 1.61　求函数 $f(x)=\sin x+\cos x-\sin x\cos x$ 的值域.

解 令 $t=\sin x+\cos x$.

确定 t 的范围. 由辅助角公式, $t=\sqrt{2}\sin(x+\frac{\pi}{4})$, 故 $t\in[-\sqrt{2},\sqrt{2}]$.

用 t 表示原函数. 对 $t=\sin x+\cos x$ 两边平方, 得 $t^2=(\sin x+\cos x)^2=\sin^2x+\cos^2x+2\sin x\cos x=1+2\sin x\cos x$. 由此解出 $\sin x\cos x=\dfrac{t^2-1}{2}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将 t 与 $\dfrac{t^2-1}{2}$ 代入原函数表达式, 得

$$f(x)=g(t)=t-\frac{t^2-1}{2}=-\frac{1}{2}t^2+t+\frac{1}{2}$$

求二次函数 $g(t)$ 在 $[-\sqrt{2},\sqrt{2}]$ 上的值域. $g(t)=-\dfrac{1}{2}(t^2-2t)+\dfrac{1}{2}=-\dfrac{1}{2}(t-1)^2+1$. 这是一个开口向下, 对称轴为 $t=1$ 的抛物线.

由于对称轴 $t=1\in[-\sqrt{2},\sqrt{2}]$, 函数在 $t=1$ 处取得最大值. $g_{\max}=g(1)=1$.

最小值在区间的两个端点 $t=-\sqrt{2}$ 或 $t=\sqrt{2}$ 处取得. $g(\sqrt{2})=-\dfrac{1}{2}(\sqrt{2}-1)^2+1=-\dfrac{1}{2}(3-2\sqrt{2})+1=\sqrt{2}-\dfrac{1}{2}$. $g(-\sqrt{2})=-\dfrac{1}{2}(-\sqrt{2}-1)^2+1=-\dfrac{1}{2}(3+2\sqrt{2})+1=-\sqrt{2}-\dfrac{1}{2}$. 比较可知 $g(-\sqrt{2})$ 是最小值.

因此, 函数 $f(x)$ 的值域为 $\left[-\sqrt{2}-\dfrac{1}{2},1\right]$.

例题 1.62 若关于 x 的方程 $\sqrt{x-1}+\sqrt{3-x}=k$ 有实数解, 求实数 k 的取值范围.

解 根式有意义要求 $x-1\geq 0$ 且 $3-x\geq 0$, 所以 $x\in[1,3]$. 令 $t=\sqrt{x-1}+\sqrt{3-x}$, 求 $t(x)$ 在 $[1,3]$ 上的值域.

$t\geq 0$, 先求 t^2 的范围.

$$\begin{aligned}t^2&=(\sqrt{x-1}+\sqrt{3-x})^2\\&=(x-1)+(3-x)+2\sqrt{(x-1)(3-x)}\\&=2+2\sqrt{-x^2+4x-3}\end{aligned}$$

设 $u(x)=-x^2+4x-3$, 求其在 $[1,3]$ 上的范围. 配方得 $u(x)=-(x^2-4x+4-4)-3=-(x-2)^2+1$. 这是一个开口向下, 顶点为 $(2,1)$ 的抛物线. 在区间 $[1,3]$ 上, $u(x)$ 在 $x=2$ 处取得最大值 $u_{\max}=1$. 在端点 $x=1$ 或 $x=3$ 处取得最小值 $u_{\min}=0$. 故 $u(x)$ 的值域为 $[0,1]$.

因此, $\sqrt{u(x)}$ 的值域为 $[0,1]$. 从而, $t^2=2+2\sqrt{u(x)}$ 的值域为 $[2+2(0),2+2(1)]=[2,4]$.

因为 $t\geq 0$, 所以 t 的值域为 $[\sqrt{2},2]$. 故实数 k 的取值范围是 $[\sqrt{2},2]$.

例题 1.63 解方程组 $\begin{cases}x^2+y^2+x+y=8\\xy+x+y=5\end{cases}$.

解 我们观察到方程组由基本对称式 $x+y$ 与 xy 构成. 令 $u=x+y, v=xy$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们用新变量表示原方程组. 注意到 $x^2+y^2=(x+y)^2-2xy=u^2-2v$. 原方程组即转化为关于 u,v 的方程组, 得

$$\begin{cases}(u^2-2v)+u=8\\ v+u=5\end{cases}$$

解关于 u,v 的方程组. 由第二个方程得 $v=5-u$. 将其代入第一个方程, 得

$$u^2-2(5-u)+u=8$$

$$u^2-10+2u+u=8$$

$$u^2+3u-18=0$$

$$(u+6)(u-3)=0$$

解得 $u=-6$ 或 $u=3$.

回代求解 x,y 并检验范围.x,y 是关于 t 的一元二次方程 $t^2-(x+y)t+xy=0$, 即 $t^2-ut+v=0$ 的两个实数根. 该方程有实数根的充要条件是其判别式 $\Delta\geq 0$, 即 $u^2-4v\geq 0$.

当 $u=3$ 时, $v=5-3=2$. 检验范围, $u^2-4v=3^2-4(2)=9-8=1\geq 0$, 存在实数解. 此时 x,y 是方程 $t^2-3t+2=0$ 的根. 解得 $t=1$ 或 $t=2$. 所以方程组的解为 $\begin{cases}x=1\\ y=2\end{cases}$ 或 $\begin{cases}x=2\\ y=1\end{cases}$.

当 $u=-6$ 时, $v=5-(-6)=11$. 检验范围, $u^2-4v=(-6)^2-4(11)=36-44=-8<0$. 此时 $t^2+6t+11=0$ 无实数根, 故此组 u,v 值舍去.

故原方程组的解为 $\begin{cases}x=1\\ y=2\end{cases}$ 和 $\begin{cases}x=2\\ y=1\end{cases}$.

例题 1.64　在实数范围内解方程 $6x^4-25x^3+37x^2-25x+6=0$.

解　这是一个系数对称的四次方程, 也称为倒数方程. 代入 $x=0$ 得常数项 $6\neq 0$, 故 $x=0$ 不是方程的根, 可在方程两边同除以 x^2.

$$6x^2-25x+37-\frac{25}{x}+\frac{6}{x^2}=0$$

将系数相同的项结合, 得

$$6\left(x^2+\frac{1}{x^2}\right)-25\left(x+\frac{1}{x}\right)+37=0$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $t=x+\dfrac{1}{x}$. 确定 t 的范围. 若 x 为非零实数, 则 $|t|=\left|x+\dfrac{1}{x}\right|\geq 2$. 平方得 $t^2=x^2+2+\dfrac{1}{x^2}$, 故 $x^2+\dfrac{1}{x^2}=t^2-2$.

代入原方程, 得到关于 t 的二次方程, 得

$$6(t^2-2)-25t+37=0$$

$$6t^2-12-25t+37=0$$

$$6t^2-25t+25=0$$

利用十字相乘法, $(2t-5)(3t-5)=0$. 解得 $t=\dfrac{5}{2}$ 或 $t=\dfrac{5}{3}$.

检验 t 的范围并回代.$t=\dfrac{5}{2}=2.5$, 满足 $|t|\geq 2$ 的条件. $t=\dfrac{5}{3}\approx 1.67$, 不满足 $|t|\geq 2$ 的条件, 故舍去.

当 $t=\dfrac{5}{2}$ 时, 解关于 x 的方程, 得

$$x+\frac{1}{x}=\frac{5}{2}$$

$$2x^2+2=5x$$

$$2x^2-5x+2=0$$

$$(2x-1)(x-2)=0$$

解得 $x=\dfrac{1}{2}$ 或 $x=2$.

因此, 原方程的解为 $x=\dfrac{1}{2},2$.

1.5.3 待定系数法

当已知函数或多项式的结构类型时, 我们先设出含未知系数的标准形式, 再根据已知条件列方程(组)解出系数.

1.5.3.1 典例剖析

我们对下面这道题用三种方法处理.

例题 1.65 因式分解 $P(a,b,c)=a^2(b-c)+b^2(c-a)+c^2(a-b)$.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

降幂整理

解 把原式看作关于 a 的多项式, 按 a 的降幂排列.

$$\begin{aligned}&a^2(b-c)+b^2(c-a)+c^2(a-b)\\=&\,a^2b-a^2c+b^2c-ab^2+ac^2-bc^2\\=&\,(b-c)a^2-(b^2-c^2)a+(b^2c-bc^2)\\=&\,(b-c)a^2-(b-c)(b+c)a+bc(b-c)\end{aligned}$$

此时, 公因式 $(b-c)$ 显现出来.

$$\begin{aligned}=&\,(b-c)[a^2-(b+c)a+bc]\\=&\,(b-c)(a-b)(a-c)\end{aligned}$$

为美观, 常整理为轮换对称形式.

$$-(a-b)(b-c)(c-a)$$

因式定理

解 将原式看作关于 a 的多项式 $P(a)$. 若令 $a=b$ 代入原式, 得

$$P(b)=b^2(b-c)+b^2(c-b)+c^2(b-b)=b^2(b-c)-b^2(b-c)+0=0$$

据因式定理, $P(a)$ 必有因式 $(a-b)$. 由于原式是轮换对称式(a,b,c 轮换后表达式不变),同理可得,它也必有因式 $(b-c)$ 和 $(c-a)$. 原式是三次齐次式,$(a-b)(b-c)(c-a)$ 也是三次式, 故它们只相差一个常数倍. 设 $P(a,b,c)=k(a-b)(b-c)(c-a)$. 取特殊值 $a=1,b=2,c=0$ 代入检验, 得左边 $=1^2(2-0)+2^2(0-1)+0=2-4=-2$. 右边 $=k(1-2)(2-0)(0-1)=k(-1)(2)(-1)=2k$. 由 $-2=2k$ 解得 $k=-1$. 故原式 $=-(a-b)(b-c)(c-a)$.

待定系数法

解 原式是关于 a,b,c 的三次齐次式, 在轮换下不变而交换任意两个变量变号. 具有这种性质的最低次式是 $(a-b)(b-c)(c-a)$, 于是设

$$a^2(b-c)+b^2(c-a)+c^2(a-b)=k(a-b)(b-c)(c-a),$$

其中 k 为待定常数. 取 $a=0,b=1,c=2$.

左边 $=0+1\cdot(2-0)+4\cdot(0-1)=2-4=-2$.

右边 $=k(0-1)(1-2)(2-0)=k(-1)(-1)(2)=2k$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由 $-2=2k$ 得 $k=-1$. 故原式 $=-(a-b)(b-c)(c-a)$.

注 1.9
三种方法各有侧重. 降幂整理直接但繁琐, 通用性强. 因式定理依赖零点观察. 待定系数法从对称结构入手.

例题 1.66 已知二次函数 $f(x)$ 满足 $f(0)=1$, 且 $f(x+1)-f(x)=2x$. 求 $f(x)$ 的解析式.

解 题目已知 $f(x)$ 是二次函数, 故设 $f(x)=ax^2+bx+c$, 其中 $a\neq 0$.

由 $f(0)=1$ 立得 $c=1$.

接下来利用条件 $f(x+1)-f(x)=2x$. 先计算左边, 得

$$\begin{aligned} f(x+1)-f(x) &= a(x+1)^2+b(x+1)+c-(ax^2+bx+c) \\ &= a(x^2+2x+1)+bx+b-ax^2-bx \\ &= 2ax+a+b \end{aligned}$$

此式恒等于 $2x$, 即 $2ax+(a+b)=2x+0$ 对一切 x 成立.

比较两边同次项系数.x 的系数给出 $2a=2$, 即 $a=1$; 常数项给出 $a+b=0$, 即 $b=-1$.

因此 $f(x)=x^2-x+1$.

例题 1.67 将 $\dfrac{1}{n(n+1)(n+2)}$ 分解为部分分式, 其中 $n\neq 0,-1,-2$.

解 设

$$\frac{1}{n(n+1)(n+2)}=\frac{A}{n}+\frac{B}{n+1}+\frac{C}{n+2}$$

其中 A,B,C 为待定常数.

将右边通分, 得

$$\frac{A(n+1)(n+2)+Bn(n+2)+Cn(n+1)}{n(n+1)(n+2)}=\frac{1}{n(n+1)(n+2)}$$

故分子满足恒等式

$$A(n+1)(n+2)+Bn(n+2)+Cn(n+1)\equiv 1$$

由于这是关于 n 的恒等式, 可代入任意 n 值来确定系数.

令 $n=0$ 得 $A\cdot 1\cdot 2=1$, 即 $A=\dfrac{1}{2}$.

令 $n=-1$ 得 $B\cdot(-1)\cdot 1=1$, 即 $B=-1$.

令 $n=-2$ 得 $C\cdot(-2)\cdot(-1)=1$, 即 $C=\dfrac{1}{2}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此

$$\frac{1}{n(n+1)(n+2)}=\frac{1}{2n}-\frac{1}{n+1}+\frac{1}{2(n+2)}=\frac{1}{2}\left(\frac{1}{n}-\frac{2}{n+1}+\frac{1}{n+2}\right)$$

注 1.10
部分分式分解用于级数求和与积分. 上例的结果可改写为

$$\frac{1}{n(n+1)(n+2)}=\frac{1}{2}\left(\frac{1}{n(n+1)}-\frac{1}{(n+1)(n+2)}\right)$$

这一裂项形式使级数求和时相邻项大量抵消, 下一节讨论这种抵消.

知乎@余命数

1.5.4 裂项相消法

我们把通项分解为两项, 或多项之差, 求和时相邻项抵消, 仅余首尾若干项.

@余命数

1.5.4.1 常见裂项公式

以下公式均可由待定系数法或直接通分验证.

$$\frac{1}{n(n+1)}=\frac{1}{n}-\frac{1}{n+1}$$

$$\frac{1}{n(n+k)}=\frac{1}{k}\left(\frac{1}{n}-\frac{1}{n+k}\right)$$

知乎@余命数

$$\frac{1}{n(n+1)(n+2)}=\frac{1}{2}\left(\frac{1}{n(n+1)}-\frac{1}{(n+1)(n+2)}\right)$$

知乎@余命数

$$\frac{1}{\sqrt{n}+\sqrt{n+1}}=\sqrt{n+1}-\sqrt{n}$$

$$n\cdot n!=(n+1)!-n!$$

知乎@余命数

练习 1.13　求和 $S=\sum_{k=1}^{n}\frac{1}{k(k+1)}$.

例题 1.68　求和 $S=\sum_{k=1}^{n}\frac{1}{\sqrt{k}+\sqrt{k+1}}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 分母是两个根式之和, 我们做有理化. 分子分母同乘 $\sqrt{k+1}-\sqrt{k}$.

$$\frac{1}{\sqrt{k}+\sqrt{k+1}}=\frac{\sqrt{k+1}-\sqrt{k}}{(\sqrt{k+1})^2-(\sqrt{k})^2}=\sqrt{k+1}-\sqrt{k}$$

于是

$$\begin{aligned}S&=\sum_{k=1}^{n}(\sqrt{k+1}-\sqrt{k})\\&=(\sqrt{2}-\sqrt{1})+(\sqrt{3}-\sqrt{2})+\cdots+(\sqrt{n+1}-\sqrt{n})\\&=\sqrt{n+1}-1\end{aligned}$$

知乎@余命数

例题 1.69 求和 $S=\sum_{k=1}^{n}k\cdot k!$.

解 直接观察通项, 注意到 $k\cdot k!=(k+1-1)\cdot k!=(k+1)!-k!$. 这正是一个裂项结构.

$$\begin{aligned}S&=\sum_{k=1}^{n}[(k+1)!-k!]\\&=(2!-1!)+(3!-2!)+\cdots+((n+1)!-n!)\\&=(n+1)!-1\end{aligned}$$

知乎@余命数

例题 1.70 求和 $S=\sum_{k=1}^{n}\frac{1}{k(k+1)(k+2)}$.

解 由待定系数法一节的结论,

$$\frac{1}{k(k+1)(k+2)}=\frac{1}{2}\left(\frac{1}{k(k+1)}-\frac{1}{(k+1)(k+2)}\right)$$

设 $a_k=\frac{1}{k(k+1)}$, 则通项为 $\frac{1}{2}(a_k-a_{k+1})$. 于是

$$\begin{aligned}S&=\frac{1}{2}\sum_{k=1}^{n}(a_k-a_{k+1})=\frac{1}{2}(a_1-a_{n+1})\\&=\frac{1}{2}\left(\frac{1}{1\cdot 2}-\frac{1}{(n+1)(n+2)}\right)\\&=\frac{1}{4}-\frac{1}{2(n+1)(n+2)}=\frac{n(n+3)}{4(n+1)(n+2)}\end{aligned}$$

知乎@余命数

例题 1.71　求和 $S=\sum_{k=0}^{n}\dfrac{1}{(2k+1)(2k+3)}$.

解　分母是两个奇数的乘积, 其差为 2. 应用公差为 2 的裂项公式, 得

$$\frac{1}{(2k+1)(2k+3)}=\frac{1}{2}\left(\frac{1}{2k+1}-\frac{1}{2k+3}\right)$$

于是

$$\begin{aligned}S&=\frac{1}{2}\sum_{k=0}^{n}\left(\frac{1}{2k+1}-\frac{1}{2k+3}\right)\\&=\frac{1}{2}\left[\left(\frac{1}{1}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{5}\right)+\cdots+\left(\frac{1}{2n+1}-\frac{1}{2n+3}\right)\right]\\&=\frac{1}{2}\left(1-\frac{1}{2n+3}\right)=\frac{n+1}{2n+3}\end{aligned}$$

1.5.5 分离常数法

当分式分子次数不低于分母次数, 或需要分离出整数部分时, 可将其分离, 使剩余部分更易处理. 这一步把分式改写成多项式加真分式, 便于分析其结构.

练习 1.14　化简 $\dfrac{x^2+3x+5}{x+1}$.

例题 1.72　求函数 $y=\dfrac{2x+1}{x-1}(x>1)$ 的最小值.

解　把分子改写为含 $(x-1)$ 的形式, 得

$$2x+1=2(x-1)+3$$

于是

$$y=\frac{2(x-1)+3}{x-1}=2+\frac{3}{x-1}$$

由于 $x>1$, 故 $x-1>0$, 从而 $\dfrac{3}{x-1}>0$.

$\dfrac{3}{x-1}$ 随 x 增大而减小, 但始终为正, 无最小值; 然而, 当 $x\to+\infty$ 时 $y\to 2$. 故 $y>2$, 函数无最小值, 但有下确界 2.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

若题目改为 $1<x\leq 2$, 则 $0<x-1\leq 1$, 从而 $\dfrac{3}{x-1}\geq 3$. 此时 $y\geq 5$, 且等号在 $x=2$ 时取得, 故最小值为 5.

例题 1.73 设 $f(x)=\dfrac{x^2}{x-1}(x>1)$, 求 $f(x)$ 的最小值.

解 设 $t=x-1$, 则 $x=t+1$, 且 $x>1$ 等价于 $t>0$.

$$f=\frac{(t+1)^2}{t}=\frac{t^2+2t+1}{t}=t+2+\frac{1}{t}$$

对 $t+\dfrac{1}{t}(t>0)$ 应用均值不等式, 得

$$t+\frac{1}{t}\geq 2\sqrt{t\cdot\frac{1}{t}}=2$$

等号在 $t=1$ 时成立. 故 $f\geq 2+2=4$, 最小值 4 在 $t=1$, 即 $x=2$ 时取得.

例题 1.74 求 $y=\dfrac{3x^2-4x+5}{x^2+1}$ 的值域.

解 从值域中取 y 反解 x:

$$y(x^2+1)=3x^2-4x+5,$$

即

$$(y-3)x^2+4x+(y-5)=0.$$

此方程有实数解 x 的条件需分类讨论.

若 $y=3$, 方程退化为 $4x+(3-5)=0$, 即 $4x-2=0$, 解得 $x=\dfrac{1}{2}$, 故 $y=3$ 在值域中.

若 $y\neq 3$, 则方程为关于 x 的一元二次方程, 有实根的充要条件是判别式非负.

$$\Delta=16-4(y-3)(y-5)\geq 0$$

$$4-(y-3)(y-5)\geq 0$$

$$4-(y^2-8y+15)\geq 0$$

$$-y^2+8y-11\geq 0$$

$$y^2-8y+11\leq 0$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解这个一元二次不等式, 得

$$y=\frac{8\pm\sqrt{64-44}}{2}=\frac{8\pm\sqrt{20}}{2}=4\pm\sqrt{5}$$

故 $4-\sqrt{5}\leq y\leq 4+\sqrt{5}$.

结合 $y=3$ 的情形 ($3\in[4-\sqrt{5},4+\sqrt{5}]$), 函数的值域为 $[4-\sqrt{5},4+\sqrt{5}]$.

1.5.6 整体思想

把表达式中反复出现的部分看作整体, 利用已知条件化简, 避免繁琐展开. 这与换元的思路相通, 只是不强求引入新字母.

练习 1.15 已知 $a+b=3, ab=2$, 求 a^2+b^2 的值.

例题 1.75 已知 $x+\dfrac{1}{x}=3$, 求 $x^3+\dfrac{1}{x^3}$ 的值.

解 设 $t=x+\dfrac{1}{x}=3$. 逐步用 t 表示更高次的对称式.

首先, 平方得

$$t^2=x^2+2+\frac{1}{x^2}=9$$

故 $x^2+\dfrac{1}{x^2}=7$.

接下来, 利用立方和公式

$$x^3+\frac{1}{x^3}=\left(x+\frac{1}{x}\right)\left(x^2-1+\frac{1}{x^2}\right)=t\cdot\left[\left(x^2+\frac{1}{x^2}\right)-1\right]=3\times(7-1)=18$$

例题 1.76 化简 $\dfrac{(a^2+b^2)^2-(a^2-b^2)^2}{(ab)^2}$, 其中 $ab\neq 0$.

解 分子是两个平方的差, 运用平方差公式, 得

$$\begin{aligned}\text{分子}&=[(a^2+b^2)+(a^2-b^2)]\cdot[(a^2+b^2)-(a^2-b^2)]\\&=(2a^2)(2b^2)=4a^2b^2\end{aligned}$$

故原式 $=\dfrac{4a^2b^2}{a^2b^2}=4$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 1.77 已知 $(x-2)^5=a_0+a_1x+a_2x^2+a_3x^3+a_4x^4+a_5x^5$, 求 $a_1+a_2+a_3+a_4+a_5$ 的值.

解 展开 $(x-2)^5$ 可直接读取系数, 也可代入特殊值.

在恒等式中令 $x=1$, 得

$$(1-2)^5=a_0+a_1+a_2+a_3+a_4+a_5$$

$$-1=a_0+a_1+a_2+a_3+a_4+a_5$$

再令 $x=0$, 得

$$(-2)^5=a_0$$

$$a_0=-32$$

两式相减得

$$a_1+a_2+a_3+a_4+a_5=-1-(-32)=31$$

例题 1.78 设 $f(x)=x^5+ax^3+bx-8$, 已知 $f(-2)=10$, 求 $f(2)$.

解 设 $g(x)=x^5+ax^3+bx$, 则 $g(x)$ 是奇函数, $g(-x)=-g(x)$.

于是 $f(x)=g(x)-8$, 且

$$f(-x)=g(-x)-8=-g(x)-8$$

将 $f(x)+f(-x)$ 相加, 得

$$f(x)+f(-x)=[g(x)-8]+[-g(x)-8]=-16$$

故 $f(2)+f(-2)=-16$, 即 $f(2)=-16-f(-2)=-16-10=-26$.

1.6 数论进阶

?? 建立了整除、辗转相除、同余与不定方程的基本框架. 本节在此基础上引入素数分解定理、高斯函数、欧拉函数与费马小定理, 这些工具在组合计数、密码学和竞赛数论中反复出现.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

1.6.1 算术基本定理

定义 1.6 (素数) 大于 1 的整数 p 若只有 1 和 p 两个正因数, 则称 p 为**素数**(或质数). 大于 1 的非素整数称为**合数**.

定理 1.8 (算术基本定理) 每个大于 1 的整数都可以唯一地表示为有限个素数的乘积. 即对任意整数 $n>1$, 存在素数 $p_1 \le p_2 \le \cdots \le p_k$ 使得

$$n = p_1 p_2 \cdots p_k$$

且这种分解在不计顺序的意义下唯一.

证明 **存在性**. 对 n 作强归纳. $n=2$ 时显然成立. 设 $n>2$, 若 n 为素数则已毕; 若 n 为合数, 则 $n=ab, 1<a,b<n$, 由归纳假设 a 和 b 均可分解为素数乘积, 故 n 亦可.

唯一性. 设 $n = p_1 \cdots p_k = q_1 \cdots q_l$ 为两种素数分解. 由素数性质, $p_1 \mid q_1 \cdots q_l$ 蕴含 p_1 整除某个 q_j, 而 q_j 为素数, 故 $p_1 = q_j$. 两边约去 p_1, 对剩余乘积继续此过程, 最终得到 $k=l$ 且两列素数一一相等. □

将相同素数合并, n 的标准分解式可写为

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_m^{\alpha_m}, \quad \alpha_i > 0$$

由此, n 的正因数个数为

$$d(n) = (\alpha_1+1)(\alpha_2+1)\cdots(\alpha_m+1) \tag{1-1}$$

n 的所有正因数之和为

$$\sigma(n) = \prod_{i=1}^{m} \frac{p_i^{\alpha_i+1}-1}{p_i-1} \tag{1-2}$$

例题 1.79 $60 = 2^2 \times 3 \times 5$, 故 $d(60) = (2+1)(1+1)(1+1) = 12$, $\sigma(60) = \frac{2^3-1}{2-1} \cdot \frac{3^2-1}{3-1} \cdot \frac{5^2-1}{5-1} = 7 \times 4 \times 6 = 168$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

1.6.2 高斯函数

定义 1.7 (高斯函数) 对实数 x, 不超过 x 的最大整数称为 x 的**取整函数**(或高斯函数), 记作 $\lfloor x \rfloor$. 即 $\lfloor x \rfloor$ 是满足

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$$

的唯一整数.

知乎@余命数

定理 1.9 (高斯函数的基本性质) 对任意实数 x、y 和整数 n:

1. $x - 1 < \lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$;
2. $\lfloor x + n \rfloor = \lfloor x \rfloor + n$;
3. 若 $x \geq 0$ 且 n 为正整数, 则 $\left\lfloor \frac{x}{n} \right\rfloor = \left\lfloor \frac{\lfloor x \rfloor}{n} \right\rfloor$;
4. $\lfloor x \rfloor + \lfloor y \rfloor \leq \lfloor x + y \rfloor \leq \lfloor x \rfloor + \lfloor y \rfloor + 1$.

证明 (1) 由定义直接给出.(2) 设 $\lfloor x \rfloor = m$, 则 $m \leq x < m+1$, 故 $m+n \leq x+n < m+n+1$, 即 $\lfloor x+n \rfloor = m+n$.(3) 设 $\lfloor x \rfloor = m, m = qn+r, 0 \leq r < n$. 由 $m \leq x < m+1$ 得 $\frac{m}{n} \leq \frac{x}{n} < \frac{m+1}{n}$, 即 $\frac{qn+r}{n} \leq \frac{x}{n} < \frac{qn+r+1}{n}$, 故 $\left\lfloor \frac{x}{n} \right\rfloor = q = \left\lfloor \frac{m}{n} \right\rfloor$.(4) 由 $x \geq \lfloor x \rfloor$ 和 $y \geq \lfloor y \rfloor$ 得 $x + y \geq \lfloor x \rfloor + \lfloor y \rfloor$, 故 $\lfloor x+y \rfloor \geq \lfloor x \rfloor + \lfloor y \rfloor$. 另一方面 $\lfloor x \rfloor + \lfloor y \rfloor$ 和 $\lfloor x+y \rfloor$ 都是整数且差值小于 2. □

知乎@余命数 知乎@余命数 知乎@余命数

定理 1.10 (勒让德定理) 在 $n!$ 的素因数分解中, 素数 p 的指数为

$$v_p(n!) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor \tag{1-3}$$

该级数实际上是有限的, 因为当 $p^k > n$ 时各项为零.

证明 在 $1, 2, \ldots, n$ 中, p 的倍数有 $\left\lfloor \frac{n}{p} \right\rfloor$ 个, 每个贡献一个 p 因子; p^2 的倍数有 $\left\lfloor \frac{n}{p^2} \right\rfloor$ 个, 每个额外贡献一个 p 因子; 以此类推. 各层贡献求和即得. □

例题 1.80 求 100! 末尾零的个数. 末尾零的个数等于 10 的因子个数, 即 $\min(v_2(100!), v_5(100!))$. 由勒让德定理:

$$v_5(100!) = \left\lfloor \frac{100}{5} \right\rfloor + \left\lfloor \frac{100}{25} \right\rfloor + \left\lfloor \frac{100}{125} \right\rfloor = 20 + 4 + 0 = 24$$

知乎@余命数

$v_2(100!)$ 远大于 24, 故末尾有 24 个零.

1.6.3 欧拉函数

定义 1.8 (欧拉函数) 对正整数 n, 欧拉函数 $\varphi(n)$ 表示 1 到 n 中与 n 互素的整数个数.

知乎@余命数

定理 1.11 (欧拉函数的计算公式) 若 n 的标准分解式为 $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_m^{\alpha_m}$, 则

$$\varphi(n) = n \prod_{i=1}^{m} \left(1 - \frac{1}{p_i}\right) = \prod_{i=1}^{m} p_i^{\alpha_i - 1}(p_i - 1) \tag{1-4}$$

证明 先计算 $\varphi(p^k)$. 在 1 到 p^k 中, 与 p 不互素的数恰是 p 的倍数, 共 p^{k-1} 个, 故 $\varphi(p^k) = p^k - p^{k-1} = p^{k-1}(p-1)$.

再证积性. 若 $(a, b) = 1$, 则 $\varphi(ab) = \varphi(a)\varphi(b)$. 这可由中国剩余定理得出: a 的完全剩余系与 b 的完全剩余系的笛卡尔积与 ab 的完全剩余系一一对应, 互素条件保持.

知乎@余命数

将积性应用于标准分解式即得 (1-4). □

例题 1.81 $\varphi(12) = \varphi(2^2 \times 3) = 12 \times \frac{1}{2} \times \frac{2}{3} = 4$. 验证: 1 到 12 中与 12 互素的数为 $1, 5, 7, 11$, 共 4 个.

知乎@余命数

1.6.4 欧拉定理与费马小定理

定理 1.12 (欧拉定理) 若 $(a, n) = 1$, 则

$$a^{\varphi(n)} \equiv 1 \pmod{n} \tag{1-5}$$

证明 设 $r_1, r_2, \ldots, r_{\varphi(n)}$ 是模 n 的一个缩系, 即 1 到 n 中与 n 互素的 $\varphi(n)$ 个整数. 因为 $(a, n) = 1$, 所以 $ar_1, ar_2, \ldots, ar_{\varphi(n)}$ 也是模 n 的缩系, 其中任意两个模 n 互不同余, 且每个都与 n 互素. 于是

$$\prod_{i=1}^{\varphi(n)} r_i \equiv \prod_{i=1}^{\varphi(n)} ar_i = a^{\varphi(n)} \prod_{i=1}^{\varphi(n)} r_i \pmod{n}$$

知乎@余命数

两边约去 $\prod r_i$(它与 n 互素,可在模 n 下消去)即得 $a^{\varphi(n)} \equiv 1 \pmod{n}$. □

推论 1.1 (费马小定理) 若 p 为素数且 $(a,p)=1$,则

$$a^{p-1} \equiv 1 \pmod{p} \tag{1-6}$$

证明 $\varphi(p) = p-1$,代入欧拉定理即得. □

费马小定理也可写为 $a^p \equiv a \pmod{p}$(对任意整数 a,不要求 $(a,p)=1$).

例题 1.82 (RSA 公钥加密) RSA 加密利用大数分解的困难性保障安全. 选两个大素数 p 和 q,令 $n=pq$. 公开 n 和加密指数 e(满足 $(e,\varphi(n))=1$),保密解密指数 d,它满足 $ed \equiv 1 \pmod{\varphi(n)}$. 加密消息 m 为 $c \equiv m^e \pmod{n}$,解密为 $m \equiv c^d \pmod{n}$.
正确性验证:$c^d \equiv m^{ed} \pmod{n}$. 由 $ed \equiv 1 \pmod{\varphi(n)}$,设 $ed = 1 + k\varphi(n)$. 当 $(m,n)=1$ 时,由欧拉定理 $m^{\varphi(n)} \equiv 1 \pmod{n}$,故 $m^{ed} = m \cdot (m^{\varphi(n)})^k \equiv m \pmod{n}$. 当 $(m,n) \neq 1$ 时,利用中国剩余定理分别在模 p 和模 q 下验证可得同一结论.

知乎@余命数

1.6.5 中国剩余定理

定理 1.13 (中国剩余定理) 设 $m_1, m_2, \ldots, m_k$ 两两互素,则同余方程组

$$\begin{cases} x \equiv a_1 \pmod{m_1} \\ x \equiv a_2 \pmod{m_2} \\ \cdots \\ x \equiv a_k \pmod{m_k} \end{cases} \tag{1-7}$$

在模 $M = m_1 m_2 \cdots m_k$ 下有唯一解.

证明 **存在性**. 令 $M_i = M/m_i$,则 $(M_i, m_i)=1$,存在 t_i 使 $M_i t_i \equiv 1 \pmod{m_i}$. 令

$$x = \sum_{i=1}^{k} a_i M_i t_i$$

对每个 j,当 $i \neq j$ 时 $m_j \mid M_i$,故 $x \equiv a_j M_j t_j \equiv a_j \pmod{m_j}$.

唯一性. 若 x_0 和 x_1 都是解,则 $m_i \mid (x_0 - x_1)$ 对所有 i 成立. 由 m_i 两两互素,$M \mid (x_0 - x_1)$,故 $x_0 \equiv x_1 \pmod{M}$. □

例题 1.83　求解 $\begin{cases} x \equiv 2 \pmod 3 \\ x \equiv 3 \pmod 5 \\ x \equiv 2 \pmod 7 \end{cases}$.

解　$M = 3 \times 5 \times 7 = 105. M_1 = 35, M_2 = 21, M_3 = 15.$

求 t_1：$35t_1 \equiv 2t_1 \equiv 1 \pmod 3$, $t_1 = 2$. 求 t_2：$21t_2 \equiv t_2 \equiv 1 \pmod 5$, $t_2 = 1$. 求 t_3：$15t_3 \equiv t_3 \equiv 1 \pmod 7$, $t_3 = 1$.

$$x = 2 \times 35 \times 2 + 3 \times 21 \times 1 + 2 \times 15 \times 1 = 140 + 63 + 30 = 233 \equiv 23 \pmod{105}$$

验证：$23 \equiv 2 \pmod 3, 23 \equiv 3 \pmod 5, 23 \equiv 2 \pmod 7$. 解为 $x \equiv 23 \pmod{105}$.

中国剩余定理把对大模数 M 的同余问题分解为对若干小模数 m_i 的同余问题，在计算数论中用于简化大模数运算.

知乎@余命数

知乎@余命数
知乎@余命数
数
乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 2 章

集合与映射



基础语言

「数学的本质在于其自由。」

——康托尔

第二章 集合与映射

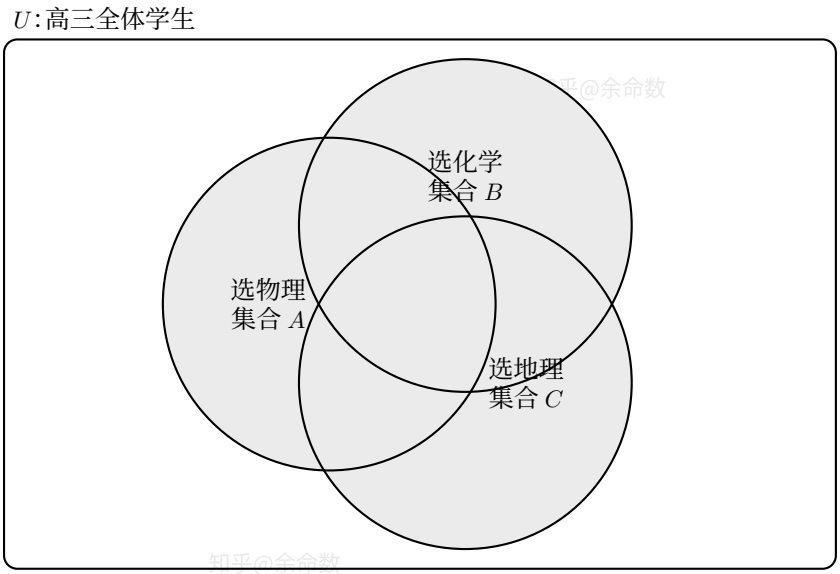


图 2.1 三集合的韦恩图:重叠区域表示交集

新高考选科后,学校要统计学生的选课情况,以便分班和安排教学资源. 设选物理、化学、地理的学生分别构成集合 A、B、C. 已知每个集合的人数,也知道两两交集与三者交集的人数. 至少选了一科的学生总数、只选了物理这一科的学生人数,都能用集合的运算不重不漏地算出.

本章引入集合与映射,用它们统一描述数集、函数定义域、方程解集等对象,并给出严谨表达推理的逻辑用语. 本章依次讨论.

- 集合的定义、表示方法、关系与运算.
- 常用数集及其扩张历程,特别是实数系的完备性.
- 逻辑用语,包括四种命题与量词否定.
- 充要条件与集合包含关系的等价转化.

章首图的选科问题,可以用本章的集合运算解决.

2.1 集合论初步

2.1.1 集合的概念与性质

我们常常把若干对象汇集起来当作一个整体来研究. 动物园里的猫科动物群、一支球队的队员、太阳系中的行星,乃至所有小于 10 的素数,都是这样的例子.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这些例子中的总体、群体或全体，称谓不同，都是指由确定对象构成的整体. 为精确描述这样的整体，数学引入了集合.

定义 2.1 (集合) 把确定的、可互相区分的事物汇集成的整体称为**集合**. 构成集合的每个事物称为该集合的**元素**.

知乎@余命数

这个定义给出集合的三条基本性质.

▷ **确定性**　任意对象是否属于给定集合，答案必须明确. 例如，所有大于 π 的实数构成集合；世界上所有著名的数学家则不能，因为著名的标准模糊.

▷ **互异性**　集合中元素互不相同. 重复出现只计一次. 例如，方程 $(x-1)^2=0$ 的解集是 $\{1\}$.

▷ **无序性**　集合中元素没有顺序. 例如，$\{1,2,3\}$ 与 $\{3,1,2\}$ 是同一个集合.

通常用大写斜体字母 $A,B,C,\ldots$ 表示集合，用小写斜体字母 $a,b,c,\ldots$ 表示元素. 若 a 属于集合 A，记作 $a\in A$，读作 a 属于 A；若 a 不属于集合 A，记作 $a\notin A$，读作 a 不属于 A.

2.1.2 集合的表示方法

知乎@余命数

2.1.2.0.1　列举法　若集合元素有限或规律明显，可把所有元素一一列出，并用花括号 {} 包起来. 例如，所有正偶数的集合写成 $\{2,4,6,8,\ldots\}$.

知乎@余命数　知乎@余命数

2.1.2.0.2　描述法　描述法给出集合中元素共同满足的条件，以此定义集合，标准形式为 $\{x\mid P(x)\}$. 例如，方程 $x^2-9=0$ 的解集写成 $\{x\in\mathbb{R}\mid x^2-9=0\}$.

知乎@余命数　知乎@余命数

2.1.2.0.3　维恩图法　英国数学家约翰·维恩用平面上的封闭曲线（通常是圆）表示集合及其关系，这种方法称为维恩图.

例题 2.1　用列举法表示集合 $A=\{d\mid d=\frac{6}{5-a}\in\mathbb{N}^*,a\in\mathbb{Z}\}$.

解　集合 A 的元素是 d，它要同时满足两条.

1. d 是正自然数，即正整数 ($d\in\mathbb{N}^*$).
2. d 由整数 a 决定，$d=\dfrac{6}{5-a}$.

因为 d 是整数，分母 $(5-a)$ 必须是 6 的整数因数；又因为 d 为正且分子为 6，分母 $(5-a)$ 也必须是正数. 于是 $(5-a)$ 是 6 的正因数，可能取值为 $1,2,3,6$.

我们将 $5-a$ 依次取 $1,2,3,6$ 代入. 当 $5-a=1$ 时 $a=4$，$d=6$；当 $5-a=2$ 时 $a=3,d=3$；当 $5-a=3$ 时 $a=2,d=2$；当 $5-a=6$ 时 $a=-1,d=1$.

因此 $A=\{1,2,3,6\}$.

知乎@余命数

注 2.1 (参数与元素的辨析)
这里常见的错误是把满足条件的参数集合 $\{-1,2,3,4\}$ 错当成集合 A. 描述法中, 竖线左侧的变量才是集合的元素; 本例中左侧变量为 d, 参数 a 只是生成元素的工具, 不属于集合.

知乎@余命数

例题 2.2 用列举法表示集合 $S=\{y \mid y=x^2,-1<x<2,y\in\mathbb{Z}\}$.

解 集合 S 的元素是满足 $y=x^2$ 且 $-1<x<2$ 的整数 y. 求解时可以先去掉整数限制, 求出 y 的连续取值范围, 再从中挑出整数.

函数 $f(x)=x^2$ 在区间 $(-1,2)$ 上的变化分三段.

- 当 x 从 -1 趋近 0 时, y 从 1 减小到 0.
- 当 $x=0$ 时, y 取最小值 0.
- 当 x 从 0 增大到 2 时, y 从 0 增大到 4.

于是当 x 遍历 $(-1,2)$ 时, y 的取值范围是 $[0,4)$, 其中 0 可取到, 4 取不到.

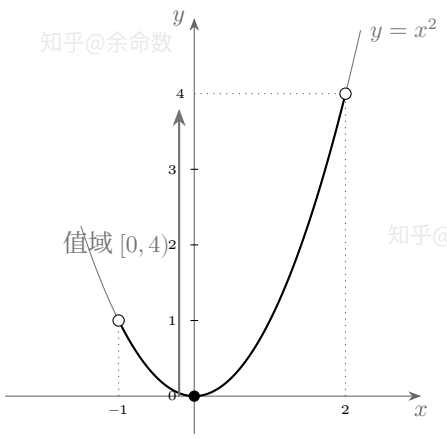


图 2.2 函数 $y=x^2$ 在定义域 $x\in(-1,2)$ 上的图像(加粗曲线). 其值域为 $[0,4)$, 其中整数点已在 y 轴上标出.

再加上 $y\in\mathbb{Z}$ 的限制, $[0,4)$ 中的整数只有 $0,1,2,3$. 因此 $S=\{0,1,2,3\}$.

2.1.3 常用数集与数系的扩张

数学中常出现几类数集, 它们的包含关系对应着数系在运算封闭性驱动下逐次扩张的过程. 所谓封闭性, 是指集合中的元素经过某种运算后, 结果仍属于该集合.

知乎@余命数
知乎@余命数

2.1.3.1 从自然数到整数与减法封闭

自然数用于计数, 即 $\mathbb{N} = \{0, 1, 2, 3, \dots\}$. 在自然数集中, 加法和乘法封闭, 减法却不一定. 例如方程

$$x + 5 = 2$$

在自然数集中无解. 为使任意减法都有意义, 我们引入**负整数**. 把正整数、0、负整数合并, 得到**整数集** $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$. 记号 $\mathbb{Z}$ 来自德语 Zahlen(意为"数"). 由定义有 $\mathbb{N} \subsetneq \mathbb{Z}$, 例如 $-1 \in \mathbb{Z}$ 但 $-1 \notin \mathbb{N}$.

知乎@余命数

2.1.3.2 从整数到有理数与除法封闭

整数集解决了减法封闭问题, 但除法又提出了新的要求. 方程

$$3x = 1$$

在整数集中无解. 为使除法封闭, 即除数不为 0 时封闭, 我们引入**分数**. 能写成两个整数之比的数构成**有理数集** $\mathbb{Q}$, 即

知乎@余命数

$$\mathbb{Q} = \left\{ x \mid x = \frac{p}{q},\ p \in \mathbb{Z},\ q \in \mathbb{Z},\ q \neq 0 \right\}.$$

记号 $\mathbb{Q}$ 来自德语 Quotient(意为"商"). 整数可看作分母为 1 的分数, 因此 $\mathbb{Z} \subsetneq \mathbb{Q}$.

2.1.3.3 从有理数到实数(补全数轴的完备性)

知乎@余命数

有理数集 $\mathbb{Q}$ 在数轴上稠密. 任意两个不同有理数之间都有无穷多个有理数. 但这不意味着有理数填满了整个数轴. 毕达哥拉斯学派发现, 边长为 1 的正方形对角线长度 $\sqrt{2}$ 不能写成两个整数之比. 这类数称为**无理数**. 著名的无理数还有圆周率 π 和自然对数的底 e.

把有理数与无理数合并, 得到能与数轴上的点一一对应的**实数集** $\mathbb{R}$. 这个描述直观但不严格, 既未说明为何不再有空隙, 也未给出实数集的构造. 我们用戴德金分割给出严格定义.

设想捕捉数轴上一个点的精确位置. 我们把有理数集分成两个非空集合 L 和 R, 使得 L 中每个数都小于 R 中每个数, 这个有序对 (L, R) 称为有理数集的一个**分割**. 我们考察分割边界所在的位置.

情况一, 边界落在有理数上. 我们取有理数 $q = 3$, 定义

- $L = \{x \in \mathbb{Q} \mid x \leq 3\}$,
- $R = \{x \in \mathbb{Q} \mid x > 3\}$.

在这个分割中, 左集合 L 有最大元素 3, 该元素标记了分割的位置.(我们也可定义 $L = \{x < 3\}, R = \{x \geq 3\}$, 此时右集合有最小元素. 边界点总会被其中一个集合包含.)

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

情况二, 边界落在空隙中. 在 $\sqrt{2}$ 所在的位置无法用有理数表示, 我们只能依有理数的性质定义分割.

- $L = \{x \in \mathbb{Q} \mid x < 0 \text{ 或} x^2 < 2\}$,
- $R = \{x \in \mathbb{Q} \mid x > 0 \text{ 且} x^2 > 2\}$.

我们分析这个分割 (L, R) 的边界.

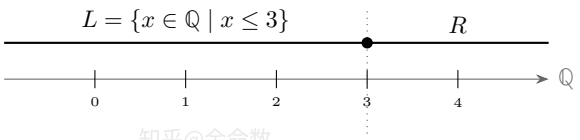
- L 没有最大元素. 任取 $x_0 \in L$ 接近 $\sqrt{2}$, 总可找到 $x_1 \in L$ 使得 $x_0 < x_1$.
- R 没有最小元素. 任取 $y_0 \in R$ 接近 $\sqrt{2}$, 总可找到 $y_1 \in R$ 使得 $y_1 < y_0$.

知乎@余命数

此时边界既不是 L 的最大元素, 也不是 R 的最小元素, 它落在有理数集的一个空隙中.

知乎@余命数

情况一:边界在有理数 3 处

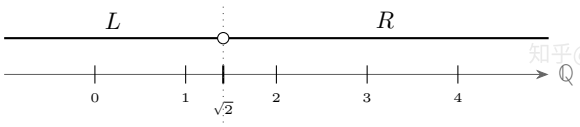


知乎@余命数

分割由 L 中的
最大元素 3 确定.

知乎@余命数

情况二:边界在空隙 $\sqrt{2}$ 处



知乎@余命数

在 $\mathbb{Q}$ 中, L 无最大元素,
R 无最小元素. 分割
定义了一个**空隙**.

图 2.3 左侧实心点表示该边界点属于 L; 右侧空心点表示边界点在有理数集 $\mathbb{Q}$ 中不被任何一个集合包含.

十九世纪德国数学家戴德金提出一种构造, 直接用**分割本身定义数**, 无需寻找填补空隙的数.

- 每一个有理数集的分割 (L, R) 就是**一个实数**.
- 若 L 有最大元素或 R 有最小元素, 这个分割定义了一个**有理数**.
- 若 L 没有最大元素**且** R 没有最小元素, 这个分割定义了一个**无理数**.

于是, 实数集 $\mathbb{R}$ 可严格定义为有理数集 $\mathbb{Q}$ 上所有分割的集合. 这个构造保证了实数轴的连续性与**完备性**, 即不存在空隙.

知乎@余命数

例题 2.3　用戴德金分割定义有理数 $q=\frac{2}{3}$, 并说明它与无理数分割的区别.

解　我们按有理数分割的特征构造

$$L=\left\{x\in\mathbb{Q}\mid x\le\frac{2}{3}\right\},\quad R=\left\{x\in\mathbb{Q}\mid x>\frac{2}{3}\right\}.$$

左集合 L 中, 由定义 $\frac{2}{3}\in L$, 且 L 中任意元素 p 都满足 $p\le\frac{2}{3}$, 故 $\frac{2}{3}$ 是 L 的**最大元素**.

右集合 R 中, 假设 R 有最小元素 $r_{\min}$, 则 $r_{\min}>\frac{2}{3}$. 我们取

$$m=\frac{1}{2}\left(r_{\min}+\frac{2}{3}\right).$$

m 是有理数, 且满足 $\frac{2}{3}<m<r_{\min}$, 因此 $m\in R$ 且 $m<r_{\min}$, 这与 $r_{\min}$ 是最小元素矛盾, 所以 R 没有最小元素.

为 $\frac{2}{3}$ 构造的分割 (L,R), 左集合有最大元素 $\frac{2}{3}$, 右集合没有最小元素; 而为 $\sqrt{2}$ 构造的分割, 左右集合都不存在边界元素. 两者的区别在此. 有理数分割的边界被其中一个集合包含, 没有产生空隙; 无理数分割的边界则悬于两个集合之间, 定义了一个空隙.

高中阶段常用的数集按包含关系排成如下链.

$$\mathbb{N}\subsetneq\mathbb{Z}\subsetneq\mathbb{Q}\subsetneq\mathbb{R}$$

2.1.4 集合间的基本关系

定义 2.2 (子集与真子集) 对于两个集合 A 和 B, 若 A 中任意元素都属于 B, 则称 A 是 B 的**子集**, 记作 $A\subseteq B$.

$$A\subseteq B\iff\forall x,\ (x\in A\implies x\in B).$$

如果 $A\subseteq B$, 且 B 中至少有一个元素不属于 A, 则称 A 是 B 的**真子集**, 记作 $A\subsetneq B$.

子集概念的两个直接推论是空集的地位与集合相等的判定.

命题 2.1 (空集与相等) 不含任何元素的集合称为**空集**, 记作 $\emptyset$.

- **空集的性质**. 空集是任何集合的子集.
- **集合相等**.$A=B\iff A\subseteq B$ 且$B\subseteq A$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

证明 按子集定义, $\emptyset \subseteq A$ 等价于命题 $\forall x,\ (x \in \emptyset \implies x \in A)$ 为真. 该蕴含命题的前提 $x \in \emptyset$ 恒为假, 因为空集不含任何元素. 在逻辑中, 前提为假的蕴含命题恒为真, 称为**空虚的真**, 故 $\emptyset \subseteq A$ 对任意集合 A 成立.

我们再证集合相等. 若 $A = B$, 则任取 $x \in A$ 都有 $x \in B$, 任取 $x \in B$ 都有 $x \in A$, 所以 $A \subseteq B$ 且 $B \subseteq A$. 反过来, 若 $A \subseteq B$ 且 $B \subseteq A$, 则任意元素 x 满足

$$x \in A \implies x \in B, \qquad x \in B \implies x \in A.$$

因而 $x \in A \iff x \in B$ 对任意 x 成立, 两个集合元素完全相同, 故 $A = B$. □

例题 2.4 已知集合 $A = \{x \mid x^2 - 3x + 2 = 0\}$, 集合 $B = \{x \mid ax = 2\}$. 若 $A \cup B = A$, 求所有使条件成立的实数 a 构成的集合.

解 $A \cup B = A$ 等价于 $B \subseteq A$, 即 B 的元素已全部在 A 中.

集合 A 由方程 $x^2 - 3x + 2 = 0$ 的解构成, 因式分解得 $(x-1)(x-2) = 0$, 解得 $x = 1$ 或 $x = 2$, 所以 $A = \{1, 2\}$.

再讨论 $B = \{x \mid ax = 2\}$ 成为 A 的子集的条件.

当 $a = 0$ 时, 方程变为 $0 \cdot x = 2$, 无解, 于是 $B = \emptyset$, 而空集是任何集合的子集, 所以 $a = 0$ 是解.

当 $a \neq 0$ 时, 方程 $ax = 2$ 有唯一解 $x = \frac{2}{a}$, 即 $B = \{\frac{2}{a}\}$. $B \subseteq A$ 要求 $\frac{2}{a} \in \{1, 2\}$.

- 由 $\frac{2}{a} = 1$ 得 $a = 2$.
- 由 $\frac{2}{a} = 2$ 得 $a = 1$.

因此实数 a 的取值集合为 $\{0, 1, 2\}$.

例题 2.5 设集合 $A = \{x \in \mathbb{R} \mid x^2 - 4x + 3 \leq 0\}$, 集合 $B = \{x \in \mathbb{R} \mid 1 - m \leq x \leq 3 + m\}$. 若 $A \subsetneq B$, 求实数 m 的取值范围.

解 由不等式 $x^2 - 4x + 3 \leq 0$ 即 $(x-1)(x-3) \leq 0$ 得 $1 \leq x \leq 3$, 于是 $A = [1, 3]$.

$A \subsetneq B$ 表示 $A \subseteq B$ 且 $A \neq B$. 为使区间 $[1, 3]$ 被 $[1 - m, 3 + m]$ 覆盖, 需要

$$\begin{cases} 1 - m \leq 1, \\ 3 + m \geq 3, \\ 1 - m \leq 3 + m, \end{cases}$$

化为 $m \geq 0$、$m \geq 0$、$m \geq -1$, 三者同时成立即 $m \geq 0$.

再排除相等的情形. 若 $A = B$, 则端点重合, 由 $1 - m = 1$ 与 $3 + m = 3$ 解得 $m = 0$, 因此为使 $A \neq B$, 必须 $m \neq 0$.

结合两个条件得 $m > 0$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 2.6　已知集合 $A=\{0,1,a\}$, $B=\{-1,a+1,a^2\}$. 若 $A=B$, 求实数 a 的值.

解 集合相等表示两个集合的元素完全相同, 与顺序无关. 我们从集合 B 中已知的 -1 入手, 确定 a.

因为 $A=B$, 所以 $-1\in A$. 而 $0\neq -1$, $1\neq -1$, 故

$$a=-1.$$

这只给出了必要条件, 还需验证充分性. 把 $a=-1$ 代入, 得

$$A=\{0,1,-1\},\qquad B=\{-1,0,1\}.$$

由集合的无序性, 两者元素相同, 所以 $A=B$ 成立.

因此 $a=-1$.

知乎@余命数

2.1.5 集合的基本运算

定义 2.3 (交集) 由所有既属于 A 又属于 B 的元素构成的集合, 称为 A 与 B 的**交集**, 记作 $A\cap B$.

$$A\cap B=\{x\mid x\in A\text{ 且}x\in B\}.$$

交集对应逻辑关系中的且, 即两个集合的公共部分.

知乎@余命数

例题 2.7　已知集合 $A=\{x\in\mathbb{Z}\mid x^2-2x-3=0\}$, 集合 $B=\{x\in\mathbb{R}\mid x>1\}$. 求 $A\cap B$.

知乎@余命数

解 先把集合化为列举法或区间等更直观的形式. 解方程 $x^2-2x-3=0$, 因式分解得 $(x-3)(x+1)=0$, 于是 $x=3$ 或 $x=-1$, 两个都是整数. 因此

知乎@余命数

$$A=\{-1,3\}.$$

集合 B 是开区间 $(1,+\infty)$.

我们按交集定义检验 A 中的元素.$-1\in A$, 但 $-1\not>1$, 故 $-1\notin B$; $3\in A$, 且 $3>1$, 故 $3\in B$.

同时属于两个集合的元素只有 3, 所以 $A\cap B=\{3\}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 2.8 设集合 $A=\{x\in\mathbb{R}\mid -2\le x\le 5\}$,集合 $B=\{x\in\mathbb{R}\mid x^2-x-6>0\}$. 求 $A\cap B$.

解 涉及不等式的集合运算, 常在数轴上分析.

集合 A 是闭区间 $[-2,5]$.

解 B 中的二次不等式 $(x-3)(x+2)>0$, 得 $x<-2$ 或 $x>3$. 因此

$$B=(-\infty,-2)\cup(3,+\infty).$$

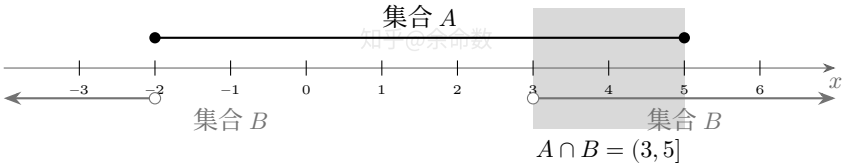


图 2.4 利用数轴求 $A\cap B$.

公共区域位于 3 到 5 之间, 端点单独判断. 左端点 $3\in A$, 但 B 要求 $x>3$, 所以 $3\notin A\cap B$; 右端点 $5\in A$ 且 $5>3$, 所以 $5\in A\cap B$.

于是 $A\cap B=(3,5]$.

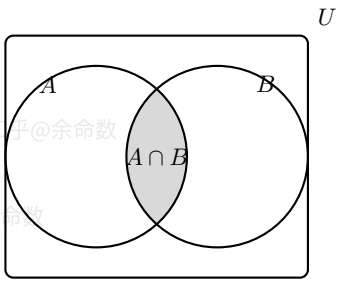


图 2.5 交集 $A\cap B$ 的维恩图示意(阴影部分)

定义 2.4 (并集) 由所有属于 A 或属于 B 的元素构成的集合, 称为 A 与 B 的**并集**, 记作 $A\cup B$.

$$A\cup B=\{x\mid x\in A\text{ 或}x\in B\}.$$

这里的或是包容性的, 包含同时属于两者的情况.

例题 2.9 已知集合 $A=\{-1,3\}$, 集合 $B=\{x\in\mathbb{R}\mid x>1\}$. 求 $A\cup B$.

解 A 含整数 -1 和 3; B 是区间 $(1,+\infty)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

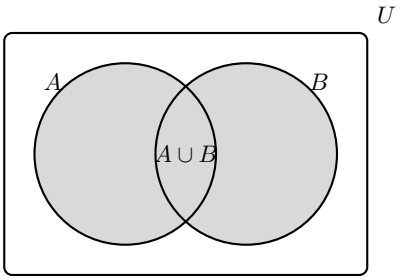


图 2.6　并集 $A \cup B$ 的维恩图示意(阴影部分)

合并时需遵循互异性.3 既属于 A, 又满足 $3 > 1$, 因此已包含在 $(1, +\infty)$ 中, 无需单独列出; -1 只属于 A, 需要单独列出.

于是

$$A \cup B = \{-1\} \cup (1, +\infty).$$

例题 2.10　设集合 $A = \{x \in \mathbb{R} \mid -2 \leq x \leq 5\}$, 集合 $B = \{x \in \mathbb{R} \mid x < -2$ 或$x > 3\}$. 求 $A \cup B$.

解　在数轴上把两个集合覆盖的区域合并. 集合 $A = [-2, 5]$, 集合 $B = (-\infty, -2) \cup (3, +\infty)$.

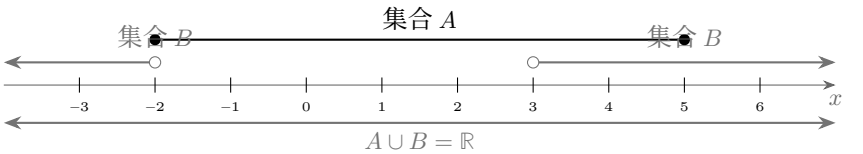


图 2.7　利用数轴求 $A \cup B$. 集合 A 填上了 B 在 $[-2, 3]$ 之间的空缺.

合并后, B 覆盖 $(-\infty, -2)$, $x = -2$ 不属于 B 但属于 A, 故 $-2 \in A \cup B$, 再并上 A 覆盖的 $[-2, 5]$ 与 B 覆盖的 $(3, +\infty)$, 得

$$A \cup B = (-\infty, -2) \cup [-2, 5] \cup (3, +\infty).$$

将 $(-\infty, -2) \cup [-2, 5]$ 合并为 $(-\infty, 5]$, 再与 $(3, +\infty)$ 取并集即得 $\mathbb{R}$.

故 $A \cup B = \mathbb{R}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 2.5 (补集) 若所研究的集合都是某个给定集合 U 的子集, 则称 U 为**全集**. 对于 U 的子集 A, 由 U 中不属于 A 的元素构成的集合, 称为 A 在 U 中的**补集**, 记作 $\complement_U A$.

$$\complement_U A = \{x \mid x \in U \text{ 且} x \notin A\}.$$

补集对应逻辑关系中的非, 即从全集中去掉 A 的元素.

知乎@余命数

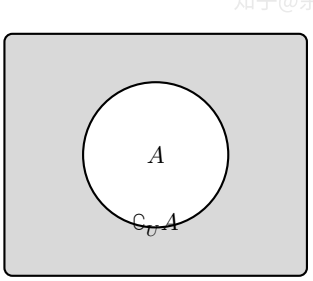


图 2.8 补集 $\complement_U A$ 的维恩图示意(阴影部分)

例题 2.11 设全集 $U = \mathbb{R}$, 集合 $A = \{x \mid -1 < x \leq 3\}$. 求 $\complement_U A$.

解 $\complement_U A$ 是所有不满足 $-1 < x \leq 3$ 的实数, 即

$$x \notin (-1, 3] \iff x \leq -1 \text{ 或} x > 3.$$

端点开闭性取反, $x > -1$ 的否定是 $x \leq -1$, $x \leq 3$ 的否定是 $x > 3$.
在数轴上表示如下.

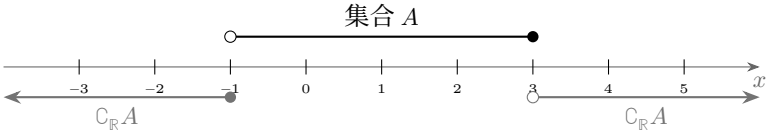


图 2.9 集合 A 与其在 $\mathbb{R}$ 中补集的数轴表示. 补集是数轴上去掉 A 后剩余的部分, 端点开闭性与 A 相反.

因此

$$\complement_U A = (-\infty, -1] \cup (3, +\infty).$$

例题 2.12 设全集 $U = \mathbb{Z}$, 集合 $A = \{x \in \mathbb{Z} \mid x^2 \leq 9\}$. 求 $\complement_U A$.

解 全集是离散的 $\mathbb{Z}$, 补集也只能由整数组成.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$x^2 \le 9$ 等价于 $-3 \le x \le 3$. 加上 $x \in \mathbb{Z}$, 得

$$A = \{-3, -2, -1, 0, 1, 2, 3\}.$$

因此 $\complement_U A$ 是所有不在 A 中的整数, 即

$$\complement_U A = \{x \in \mathbb{Z} \mid x < -3 \text{ 或} x > 3\} = \{\dots, -5, -4\} \cup \{4, 5, \dots\}.$$

知乎@余命数

注 2.2 (全集)
全集决定了补集的形式. 若在同一条件下取实数全集 $A' = \{x \in \mathbb{R} \mid x^2 \le 9\}$, 则 $\complement_{\mathbb{R}} A' = (-\infty, -3) \cup (3, +\infty)$; 而本题全集是 $\mathbb{Z}$, 补集只能是整数集合.

2.1.6 集合的运算律

集合运算满足交换律、结合律、分配律等. 下面证明德摩根定律.

定理 2.1 (德摩根定律) 对于全集 U 及其任意子集 $A, B \subseteq U$, 我们有

$$\complement_U(A \cup B) = (\complement_U A) \cap (\complement_U B)$$
$$\complement_U(A \cap B) = (\complement_U A) \cup (\complement_U B)$$

知乎@余命数　知乎@余命数

证明 我们用逻辑等价推导证明. 任取 $x \in U$.

$$\begin{aligned}
x \in \complement_U(A \cup B) &\iff x \notin (A \cup B) && \text{(补集定义)} \\
&\iff \neg(x \in A \text{ 或} x \in B) && \text{(并集定义)} \\
&\iff (x \notin A) \text{ 且} (x \notin B) && \text{(逻辑德摩根律)} \\
&\iff (x \in \complement_U A) \text{ 且} (x \in \complement_U B) && \text{(补集定义)} \\
&\iff x \in (\complement_U A) \cap (\complement_U B) && \text{(交集定义)}
\end{aligned}$$

因此第一式成立. 同理, 对任意 $x \in U$ 我们有

$$\begin{aligned}
x \in \complement_U(A \cap B) &\iff x \notin (A \cap B) && \text{(补集定义)} \\
&\iff \neg(x \in A \text{ 且} x \in B) && \text{(交集定义)} \\
&\iff (x \notin A) \text{ 或} (x \notin B) && \text{(逻辑德摩根律)} \\
&\iff (x \in \complement_U A) \text{ 或} (x \in \complement_U B) && \text{(补集定义)} \\
&\iff x \in (\complement_U A) \cup (\complement_U B) && \text{(并集定义)}
\end{aligned}$$

知乎@余命数　知乎@余命数

两个等式两边都是 U 的子集, 且对任意 $x \in U$ 给出相同的归属关系, 故两式成立. □

例题 2.13 (导出容斥原理相关恒等式) 设 A, B, C 为有限集. 证明仅属于 A 的元素个数为

$$\operatorname{card}(A) - \operatorname{card}(A \cap B) - \operatorname{card}(A \cap C) + \operatorname{card}(A \cap B \cap C).$$

解 我们取包含 $A \cup B \cup C$ 的全集 U. 仅属于 A 的元素满足 $x \in A$ 且 $x \notin B$、$x \notin C$, 即

$$A \cap (\complement_U B) \cap (\complement_U C).$$

由德摩根定律,

$$(\complement_U B) \cap (\complement_U C) = \complement_U (B \cup C),$$

所以上述集合可写成

$$A \cap \complement_U (B \cup C).$$

它的基数等于 A 中去掉 A 与 $B \cup C$ 的交集部分.

$$\operatorname{card}(A) - \operatorname{card}(A \cap (B \cup C)).$$

用分配律, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$. 我们对右端应用二元容斥原理.

$$\begin{aligned} \operatorname{card}\big((A \cap B) \cup (A \cap C)\big) &= \operatorname{card}(A \cap B) + \operatorname{card}(A \cap C) \\ &\quad - \operatorname{card}\big((A \cap B) \cap (A \cap C)\big) \\ &= \operatorname{card}(A \cap B) + \operatorname{card}(A \cap C) - \operatorname{card}(A \cap B \cap C). \end{aligned}$$

代回得

$$\begin{aligned} &\operatorname{card}(A \cap \complement_U (B \cup C)) \\ &\quad = \operatorname{card}(A) - \big(\operatorname{card}(A \cap B) + \operatorname{card}(A \cap C) - \operatorname{card}(A \cap B \cap C)\big) \\ &\quad = \operatorname{card}(A) - \operatorname{card}(A \cap B) - \operatorname{card}(A \cap C) + \operatorname{card}(A \cap B \cap C). \end{aligned}$$

证毕.

例题 2.14 设全集 $U = \mathbb{R}$, 集合 $A = \{x \mid x^2 - 2x - 8 > 0\}$, $B = \{x \mid -1 \le x \le 5\}$. 求 $\complement_U (A \cup B)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 直接算 $A\cup B$ 再求补集较繁琐,用德摩根定律更简单.

$$\complement_U(A\cup B)=(\complement_U A)\cap(\complement_U B).$$

先求 $\complement_U A$. 由 $x^2-2x-8\le 0$ 得 $(x-4)(x+2)\le 0$, 于是 $-2\le x\le 4$, 即 $\complement_U A=[-2,4]$.

再求 $\complement_U B$. 由 $B=\{x\mid -1\le x\le 5\}$ 取补集得

$$\complement_U B=\{x\mid x<-1 \text{ 或} x>5\}=(-\infty,-1)\cup(5,+\infty).$$

知乎@余命数

于是

$$(\complement_U A)\cap(\complement_U B)=[-2,4]\cap((-\infty,-1)\cup(5,+\infty))=[-2,-1).$$

例题 2.15 某班级共 50 名学生,一次考试中有两道压轴题. 答对第一题的有 30 人, 答对第二题的有 28 人, 两题都答对的有 15 人. 问两题都未答对的学生有多少人?

知乎@余命数

解 设 U 为全班学生, A 为答对第一题的学生集合, B 为答对第二题的学生集合. 题意给出

$$\text{card}(U)=50,\quad \text{card}(A)=30,\quad \text{card}(B)=28,\quad \text{card}(A\cap B)=15.$$

两题都未答对的学生属于 $(\complement_U A)\cap(\complement_U B)$. 由德摩根定律,

知乎@余命数

$$(\complement_U A)\cap(\complement_U B)=\complement_U(A\cup B).$$

因此所求人数为

$$\text{card}(\complement_U(A\cup B))=\text{card}(U)-\text{card}(A\cup B).$$

用容斥原理得

$$\begin{aligned}\text{card}(A\cup B)&=\text{card}(A)+\text{card}(B)-\text{card}(A\cap B)\\&=30+28-15=43.\end{aligned}$$

所以两题都未答对的人数为 $50-43=7$.

例题 2.16 设全集 $U=\mathbb{R}$,集合 $A=\{x\mid x^2-9>0\}$,$B=\{x\mid x^2-5x+4\le 0\}$. 求 $(\complement_U A)\cup(\complement_U B)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 一种办法是直接计算两个补集再取并集.$\complement_U A$ 满足 $x^2-9\le 0$, 即 $-3\le x\le 3$, 所以 $\complement_U A=[-3,3]$; $\complement_U B$ 满足 $(x-1)(x-4)>0$, 即 $x<1$ 或 $x>4$, 所以 $\complement_U B=(-\infty,1)\cup(4,+\infty)$. 于是

$$(\complement_U A)\cup(\complement_U B)=[-3,3]\cup(-\infty,1)\cup(4,+\infty)=(-\infty,3]\cup(4,+\infty).$$

另一种办法是先化简再借助德摩根定律. 由 $(\complement_U A)\cup(\complement_U B)=\complement_U(A\cap B)$, 先把

$$A=(-\infty,-3)\cup(3,+\infty),\quad B=[1,4]$$

取交集得 $A\cap B=(3,4]$, 再取补集得

$$\complement_U(A\cap B)=(-\infty,3]\cup(4,+\infty).$$

两种办法结果相同.

2.1.7 有限集的计数

定义 2.6 (集合的基数) 有限集合的元素个数称为该集合的**基数**, 记作 $\mathrm{card}(A)$.

基数度量有限集合的元素个数, 由此可计算子集总数.

定理 2.2 若 $\mathrm{card}(A)=n$, 则 A 的子集个数为 2^n.

证明 设 $A=\{a_1,a_2,\dots,a_n\}$. 我们构造 A 的任意子集 S, 对每个元素 a_i 独立地选择 $a_i\in S$ 或 $a_i\notin S$. 把这一组选择编码为长度为 n 的二进制序列 $(b_1,b_2,\dots,b_n)$, 其中

$$b_i=\begin{cases}1, & a_i\in S,\\ 0, & a_i\notin S.\end{cases}$$

这给出了 A 的幂集 $\mathcal{P}(A)$ 到所有长度为 n 的二进制序列之间的双射.

每次选择有 2 种, n 次独立选择共有

$$\underbrace{2\times 2\times\cdots\times 2}_{n\text{个}}=2^n$$

种结果, 所以 A 的子集个数为 2^n. □

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 2.3 (容斥原理) 对任意有限集合 A, B, C, 我们有

- $\operatorname{card}(A \cup B) = \operatorname{card}(A) + \operatorname{card}(B) - \operatorname{card}(A \cap B)$.
- $$\begin{aligned}&\operatorname{card}(A \cup B \cup C)\\ &= \operatorname{card}(A) + \operatorname{card}(B) + \operatorname{card}(C)\\ &\quad - \operatorname{card}(A \cap B) - \operatorname{card}(A \cap C) - \operatorname{card}(B \cap C)\\ &\quad + \operatorname{card}(A \cap B \cap C).\end{aligned}$$

证明 *二元情形.* 把 $A \cup B$ 分成三块互不相交的部分.

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B).$$

三集合两两不交, 故

$$\operatorname{card}(A \cup B) = \operatorname{card}(A \setminus B) + \operatorname{card}(B \setminus A) + \operatorname{card}(A \cap B).$$

又 $A = (A \setminus B) \cup (A \cap B)$, 所以

$$\operatorname{card}(A \setminus B) = \operatorname{card}(A) - \operatorname{card}(A \cap B).$$

同理, $\operatorname{card}(B \setminus A) = \operatorname{card}(B) - \operatorname{card}(A \cap B)$. 代入即得

$$\operatorname{card}(A \cup B) = \operatorname{card}(A) + \operatorname{card}(B) - \operatorname{card}(A \cap B).$$

三元情形. 令 $D = B \cup C$, 则 $A \cup B \cup C = A \cup D$. 由二元情形,

$$\operatorname{card}(A \cup B \cup C) = \operatorname{card}(A) + \operatorname{card}(D) - \operatorname{card}(A \cap D).$$

对 D 再用二元情形, 得

$$\operatorname{card}(D) = \operatorname{card}(B \cup C) = \operatorname{card}(B) + \operatorname{card}(C) - \operatorname{card}(B \cap C).$$

对 $A \cap D$, 由分配律 $A \cap D = (A \cap B) \cup (A \cap C)$, 再用二元情形,

$$\operatorname{card}(A \cap D) = \operatorname{card}(A \cap B) + \operatorname{card}(A \cap C) - \operatorname{card}(A \cap B \cap C).$$

把 $\operatorname{card}(D)$ 与 $\operatorname{card}(A \cap D)$ 代回并整理, 得到

$$\begin{aligned}&\operatorname{card}(A \cup B \cup C)\\ &= \operatorname{card}(A) + \operatorname{card}(B) + \operatorname{card}(C)\\ &\quad - \operatorname{card}(A \cap B) - \operatorname{card}(A \cap C) - \operatorname{card}(B \cap C)\end{aligned}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

$$+\operatorname{card}(A\cap B\cap C).$$

证毕. □

2.2 映射

2.2.1 映射的概念

我们先观察数集之间的对应关系. 例如, $f(x)=x^2$ 把每个实数 x 对应到 x^2; 平面上的旋转变换把每个点对应到旋转后的点. 我们把这种对应关系推广到一般集合, 就得到**映射**.

定义 2.7 (映射) 设 A, B 是两个非空集合. 如果存在法则 f, 使得对每个 $x\in A$ 都有唯一确定的 $y\in B$ 与之对应, 那么我们称 f 为从 A 到 B 的一个**映射**, 并记作

$$f: A\to B.$$

其中 A 称为映射的**定义域**, B 称为映射的**陪域**.

这一定义包含三个要素, 即定义域 A、陪域 B 和对应法则 f. 它同时对 f 提出两条要求.

- **全域性**定义域中每个元素都有像.
- **单值性**每个元素的像唯一.

若 $x\in A$ 在 B 中对应 y, 则称 y 为 x 在 f 下的**像**, 记作 $y=f(x)$; 称 x 为 y 的一个**原像**. 定义域中所有元素的像构成集合, 称为 f 的**像集**或**值域**, 记作 $f(A)$.

$$f(A)=\{f(x)\mid x\in A\}.$$

显然 $f(A)\subseteq B$; 陪域 B 包含所有可能的像, 而像集是这些像中实际取到的部分.

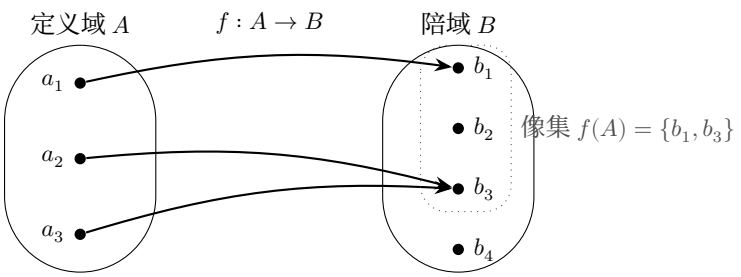


图 2.10 映射图示. 定义域 A 中每个元素都有唯一箭头指向其在陪域 B 中的像. 陪域中 b_2, b_4 没有原像; b_3 有两个原像 a_2, a_3.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 2.17　我们考察函数 $f(x)=x^2$. 将其视为映射, 分析定义域、陪域和像集对性质的影响.

解　我们先说明, 一个映射的性质由定义域、陪域和对应法则共同决定.

1. 取 $f_1:\mathbb{R}\to\mathbb{R}$, $f_1(x)=x^2$. 像集为 $f_1(\mathbb{R})=[0,+\infty)$, 是陪域的真子集.
2. 取 $f_2:\mathbb{R}\to[0,+\infty)$, $f_2(x)=x^2$. 对应法则与 f_1 相同, 但陪域缩小为非负实数集. 像集 $f_2(\mathbb{R})=[0,+\infty)$ 与陪域相等.
3. 取 $f_3:[0,+\infty)\to[0,+\infty)$, $f_3(x)=x^2$. 定义域也被限制为非负实数. 像集仍是 $[0,+\infty)$.
4. 取 $f_4:\mathbb{R}\to(-\infty,-1)$, $f_4(x)=x^2$. 这个定义不合法. 因为对任意 $x\in\mathbb{R}$ 都有 $x^2\geq 0$, 不属于陪域 $(-\infty,-1)$.

因此, 同一法则配以不同的定义域或陪域, 会得到不同的映射.

知乎@余命数

2.2.2 特殊映射

我们按像与原像的对应关系, 对映射作进一步分类.

定义 2.8 (单射、满射、双射)　设 $f:A\to B$ 为映射.

- 若对定义域中任意两个不同元素, 像也互不相同, 即

$$\forall x_1,x_2\in A,\ x_1\neq x_2\implies f(x_1)\neq f(x_2),$$

  则称 f 为**单射**或一对一映射.
- 若陪域中每个元素都至少有一个原像, 即像集等于陪域 ($f(A)=B$), 亦即

$$\forall y\in B,\ \exists x\in A,\ f(x)=y,$$

  则称 f 为**满射**或映上映射.
- 若 f 既是单射又是满射, 则称 f 为**双射**或一一对应.

单射意味着不同输入对应不同输出; 满射意味着陪域中的每个元素都被取到; 双射则给出两个集合之间一一对应, 并且是可逆的.

例题 2.18　我们回到 $f(x)=x^2$ 的例子, 判断 f_1,f_2,f_3 各为何种映射.

解　我们逐一对照单射与满射的定义来检验 f_1,f_2,f_3.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1. $f_1 : \mathbb{R} \to \mathbb{R}$ 不是单射, 因为 $f_1(-2) = 4 = f_1(2)$ 而 $-2 \neq 2$; 也不是满射, 因为像集 $[0, +\infty) \neq \mathbb{R}$, 例如 -1 没有原像.
2. $f_2 : \mathbb{R} \to [0, +\infty)$ 仍不是单射, 但它是满射. 任意 $y \in [0, +\infty)$ 都有原像 $\sqrt{y}$ 和 $-\sqrt{y}$.
3. $f_3 : [0, +\infty) \to [0, +\infty)$ 是单射. 若 $f_3(x_1) = f_3(x_2)$, 则 $x_1^2 = x_2^2$, 而 $x_1, x_2 \geq 0$, 所以 $x_1 = x_2$. 它也是满射. 任意 $y \in [0, +\infty)$ 都有唯一原像 $\sqrt{y}$.
   因此 f_3 是双射. 平方根函数 $\sqrt{x}$ 取非负实数域为定义域, 正是为了保证它是双射.

2.3 逻辑用语初步

知乎@余命数

我们先说明数学推理所用的语言. 推理要求精确、没有歧义, 这种语言就是逻辑用语.

@余命数
知乎@余命数

2.3.1 命题及其真值

定义 2.9 (命题) 我们称能判断真假的陈述句为**命题**.

我们把一个命题的判断结果称为其**真值**, 真值只有真、假两种.

例如, $3 > 2$ 是真命题, "所有素数都是奇数"是假命题(2 是素数且为偶数), 而 $x > 5$ 在 x 的值确定之前无法判断真假, 不是命题. 我们称含有变量的陈述句为谓词或开命题.

知乎@余命数

2.3.2 逻辑联结词与集合运算

我们可以通过逻辑联结词把简单命题组合成复合命题. 这些联结词与集合运算存在对应关系.

定义 2.10 (基本逻辑联结词) 设 p, q 为两个命题.

- **否定**($\neg p$, 读作"非 p"). p 真则 $\neg p$ 假, p 假则 $\neg p$ 真. 这对应集合的**补集**运算. 若 P 是使 p 为真的情形集合, 则使 $\neg p$ 为真的情形集合就是 $\complement_U P$.
- **合取**($p \wedge q$, 读作 p 且 q). 仅当 p 和 q 同时为真时, $p \wedge q$ 为真. 这对应集合的**交集** $P \cap Q$.
- **析取**($p \vee q$, 读作 p 或 q). 仅当 p 和 q 同时为假时, $p \vee q$ 为假. 这是包容性的或, 对应集合的**并集** $P \cup Q$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$\neg p \iff \complement_U P$　　$p \wedge q \iff P \cap Q$　　$p \vee q \iff P \cup Q$

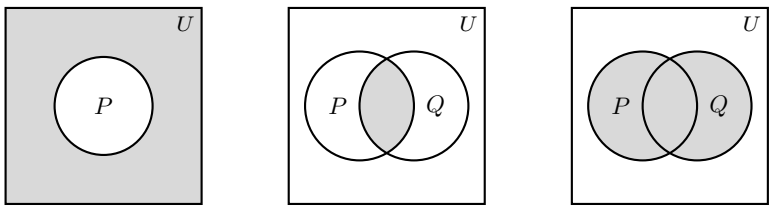


图 2.11　逻辑联结词与集合运算的对应关系

2.3.3 蕴含关系与充要条件

我们讨论条件命题, 数学中的定理大多写成“若 p 则 q”的形式.

知乎@余命数

定义 2.11 (蕴含命题) 我们称形如 “若 p 则 q” 的命题为**蕴含命题**, 并记作 $p \implies q$. p 称为**前提**或**条件**, q 称为**结论**. 我们规定, $p \implies q$ 的真值如下. **仅在 p 真且 q 假时为假**, 其余情况均为真.

我们在集合论中把蕴含关系对应到**子集关系**. 命题 $p \implies q$ 为真, 当且仅当使 p 为真的情形集合 P 是使 q 为真的情形集合 Q 的子集, 即 $P \subseteq Q$.

- 若前提 p 为假, 即 $x \notin P$, 则 x 是否在 Q 中与 $P \subseteq Q$ 的真假无关. 因此, 前提为假时蕴含命题恒为真, 我们称这种情况为**空虚的真**.

对蕴含命题 $p \implies q$, 我们构造三个关联命题.

- **逆命题** $q \implies p$.
- **否命题** $\neg p \implies \neg q$.
- **逆否命题** $\neg q \implies \neg p$.

知乎@余命数

定理 2.4　一个蕴含命题与其逆否命题**逻辑等价**, 即真值相同.

$$(p \implies q) \iff (\neg q \implies \neg p).$$

逆命题与否命题也逻辑等价.

知乎@余命数

证明 $p \implies q$ 仅在 p 真且 q 假时为假; 其逆否命题 $\neg q \implies \neg p$ 仅在“$\neg q$ 真且 $\neg p$ 假”时为假, 也就是 q 假且 p 真. 两者同时为假的情形完全相同, 所以真值相同, 即

$$(p \implies q) \iff (\neg q \implies \neg p).$$

同理, 逆命题 $q \implies p$ 仅在 q 真且 p 假时为假; 否命题 $\neg p \implies \neg q$ 仅在“$\neg p$ 真且 $\neg q$ 假”时为假, 即 p 假且 q 真. 两者同真假, 所以逆命题与否命题逻辑等价. □

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

这一定理是**反证法**的依据. 要证 $p \implies q$, 我们可以改证其逆否命题, 即由 $\neg q$ 推出 $\neg p$.

如果 $p \implies q$ 成立, 那么我们称 p 是 q 的**充分条件**, q 是 p 的**必要条件**.

- p 成立足以保证 q 成立.
- q 成立是 p 成立的必要前提.

若 $p \implies q$ 与 $q \implies p$ 同时成立, 记作 $p \iff q$, 则称 p 是 q 的**充分必要条件**, 也称充要条件. 这在集合论中对应 $P = Q$.

知乎@余命数

2.3.4 量词及其否定

为把谓词变成命题, 我们引入量词.

知乎@余命数

定义 2.12 (量词)

- **全称量词** $\forall$, 读作“对任意”或“所有”. 命题 $\forall x \in M, p(x)$ 为真, 当且仅当 M 中每个元素 x 都使 $p(x)$ 成立.
- **存在量词** $\exists$, 读作“存在”或“有些”. 命题 $\exists x \in M, p(x)$ 为真, 当且仅当 M 中至少存在一个元素 x_0 使 $p(x_0)$ 成立.

在证明与反驳中, 我们需要对量词命题作出正确的否定.

知乎@余命数

定理 2.5 (量词的否定)

$$\neg(\forall x \in M, p(x)) \iff \exists x \in M, \neg p(x)$$
$$\neg(\exists x \in M, p(x)) \iff \forall x \in M, \neg p(x)$$

证明 $\forall x \in M, p(x)$ 为假, 当且仅当 M 中存在某个 x_0 使 $p(x_0)$ 不成立, 即 $\exists x \in M, \neg p(x)$.

$\exists x \in M, p(x)$ 为假, 当且仅当 M 中没有元素使 $p(x)$ 成立, 即每个 $x \in M$ 都使 $p(x)$ 不成立, 也就是 $\forall x \in M, \neg p(x)$. □

例题 2.19 我们写出命题“所有正方形的对角线都互相垂直”的否定.

知乎@余命数

解 设 S 为正方形集合, $p(x)$ 表示“x 的对角线互相垂直”. 原命题符号化为 $\forall x \in S, p(x)$.

由量词否定规则, 其否定为 $\exists x \in S, \neg p(x)$, 即“存在一个正方形, 它的对角线不互相垂直”.

原命题为真, 因此其否定为假.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2.4 势与对角线法 *

有限集的大小用元素个数度量. 比较两个有限集合, 只需数一数谁的元素多. 无限集没有元素个数可言, 但 Cantor 发现, 无限集之间仍有大小之分, 有的无限集更大, 有的无限集较小. 区分的工具是*势*, 把双射当作配对的标准, 两个集合能一一配对就视为等势. 本节用对角线法证明实数集比自然数集更大, 并由此引出超越数的存在性.

知乎@余命数

2.4.1 可数集

知乎@余命数

定义 2.13 (可数集) 若集合 A 与自然数集 $\mathbb{N}$ 之间存在双射, 则称 A 是**可数集**. 等价地说, A 可数当且仅当它的元素可以排成一个无穷序列

$$a_1,\ a_2,\ a_3,\ \ldots$$

使得每个元素恰好出现一次.

有限集也可视为可数集(序列在某项后停止). 按此定义, $\mathbb{N}$ 本身可数. 下面看两个不那么平凡的可数集.

例题 2.20 (整数集可数) 整数集 $\mathbb{Z}$ 是可数集.

证明 把整数按如下顺序排列:

$$0,\ 1,\ -1,\ 2,\ -2,\ 3,\ -3,\ \ldots$$

第 1 项是 0, 第 $2n$ 项是 n, 第 $2n+1$ 项是 $-n (n \geq 1)$. 每个整数恰好出现一次, 这给出了 $\mathbb{N}^+$ 到 $\mathbb{Z}$ 的双射, 故 $\mathbb{Z}$ 可数. □

例题 2.21 (有理数集可数) 有理数集 $\mathbb{Q}$ 是可数集.

证明 每个正有理数可写成既约分数 p/q, 其中 $p, q \in \mathbb{N}^+$. 把所有正有理数排成一张二维表, 第 p 行第 q 列放置 p/q:

知乎@余命数

沿反对角线 $p+q = 2,\ 3,\ 4,\ \ldots$ 依次扫描, 先取 $1/1$, 再取 $1/2,\ 2/1$, 再取 $3/1,\ 2/2,\ 1/3$, 依此类推. 扫描中跳过非既约分数, 如 $2/2$, 每个正有理数恰好被取到一次. 这样得到正有理数集 $\mathbb{Q}^+$ 的一个排列.

负有理数与正有理数一一对应 $(q \mapsto -q)$, 把正、负有理数交替插入并补上 0, 得到 $\mathbb{Q}$ 的一个排列. 故 $\mathbb{Q}$ 可数. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

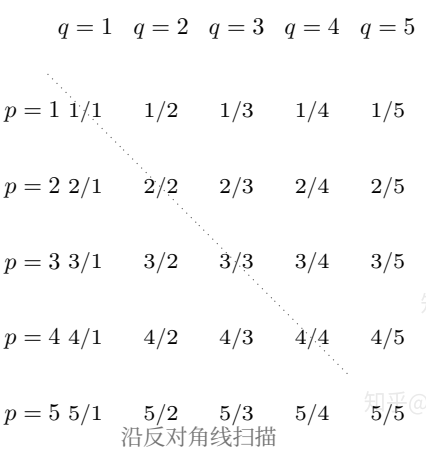


图 2.12 正有理数 p/q 排成二维表, 沿反对角线 $p+q=$ 常数 依次扫描.

可数集的并仍可数, 只要参与合并的可数集本身可枚举.

定理 2.6 (可数集的可数并) 若 $A_1, A_2, A_3, \ldots$ 是一列可数集, 则 $\bigcup_{k=1}^{\infty} A_k$ 也是可数集.

证明 把每个 A_k 的元素排成序列 $a_{k,1},\ a_{k,2},\ a_{k,3},\ \ldots$, 再像正有理数那样沿反对角线扫描二维表 $\{a_{k,j}\}$, 跳过重复元素. 每个元素都会被扫到, 合并后的集合可排成序列, 故可数. □

2.4.2 Cantor 对角线法

有理数在数轴上稠密, 任意两个实数之间都有有理数. 稠密似乎意味着元素很多, 但 $\mathbb{Q}$ 仍可数.Cantor 用对角线法证明实数集 $\mathbb{R}$ 不可数.

定理 2.7 (实数集不可数) 开区间 $(0,1)$ 不可数, 因而 $\mathbb{R}$ 不可数.

证明 反证. 假设 $(0,1)$ 可数, 则其元素可排成序列 $r_1, r_2, r_3, \ldots$. 把每个 r_n 写成十进制小数(统一用不结尾全 9 的表示, 避免 $0.4999\cdots = 0.5000\ldots$ 的歧义):

$$
\begin{aligned}
r_1 &= 0.\,d_{11}\,d_{12}\,d_{13}\,\cdots \\
r_2 &= 0.\,d_{21}\,d_{22}\,d_{23}\,\cdots \\
r_3 &= 0.\,d_{31}\,d_{32}\,d_{33}\,\cdots \\
&\ \ \vdots
\end{aligned}
$$

其中 d_{nk} 是 r_n 的第 k 位小数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

构造一个新数 $s = 0.c_1\,c_2\,c_3\,\cdots$, 其第 n 位 c_n 取与 d_{nn} 不同的值. 为避开歧义, 规定

$$c_n = \begin{cases} 1, & d_{nn} \neq 1, \\ 2, & d_{nn} = 1. \end{cases}$$

这样 $c_n \in \{1, 2\}$, 不取 9, 小数 s 不以全 9 结尾, 表示唯一.

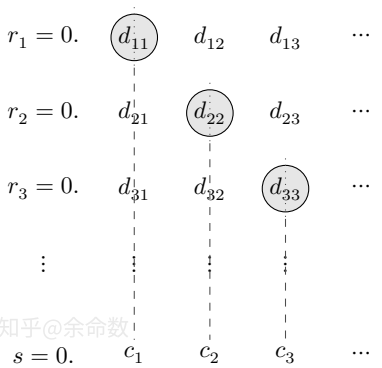


图 2.13　对角线构造: 新数 s 的第 n 位 c_n 与 r_n 的第 n 位 d_{nn} 不同(圈出), 故 s 与每个 r_n 都在某一位上相异.

对每个 n, s 的第 n 位 $c_n \neq d_{nn}$, 而 r_n 的第 n 位是 d_{nn}, 故 $s \neq r_n$. 序列 $r_1, r_2, r_3, \ldots$ 中没有任何一项等于 s.

但 $s \in (0,1)$, 因为 $s = 0.c_1c_2c_3\cdots$ 且每位取 1 或 2, 有 $0 < s \leq 0.222\cdots = 2/9 < 1$. 于是 s 属于 $(0,1)$ 却不在序列中, 与序列穷尽 $(0,1)$ 的假设矛盾.

故 $(0,1)$ 不可数. 而 $(0,1) \subset \mathbb{R}$, 若 $\mathbb{R}$ 可数则其子集 $(0,1)$ 亦可数, 矛盾, 故 $\mathbb{R}$ 不可数. □

对角线法的机制是, 给定任意一个枚举, 都能构造出一个不在其中的元素, 只要沿对角线逐位取反. 枚举者无论怎样调整顺序, 对角线数总能逃逸. 这把 $\mathbb{R}$ 比 $\mathbb{N}$ 更大这一断言从直觉提升为定理.

有理数集 $\mathbb{Q}$ 在数轴上稠密, 却可数; 实数集 $\mathbb{R}$ 不可数. 稠密与不可数是两个独立的性质, 前者描述元素在局部如何分布, 后者描述集合整体的规模.

2.4.3 代数数与超越数

定义 2.14 (代数数与超越数) 若实数 α 是某个整系数多项式方程

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \quad (a_n \neq 0,\ a_i \in \mathbb{Z})$$

的根, 则称 α 为**代数数**. 不是代数数的实数称为**超越数**.

知乎@余命数　知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

每个有理数 p/q 都是代数数, 因为它是 $qx-p=0$ 的根.$\sqrt{2}$ 也是代数数, 它是 $x^2-2=0$ 的根.

定理 2.8 (代数数可数) 所有代数数构成的集合是可数集.

证明 整系数多项式 $a_nx^n+\cdots+a_0$ 由系数序列 $(a_n,\ldots,a_0)$ 决定. 固定次数 n 与高度 $h=|a_n|+\cdots+|a_0|$, 满足 h 的多项式只有有限个. 按 n、h 从小到大排列, 所有整系数多项式可排成序列, 故多项式集合可数.

每个多项式只有有限个根, 即次数为 n 的多项式至多 n 个实根. 可数个有限集的并仍可数(由定理 2.6), 故代数数集合可数. □

定理 2.9 (超越数的存在性) 超越数存在, 且超越数集合不可数.

证明 实数集 $\mathbb{R}$ 不可数, 代数数集合可数. 若超越数也可数, 则 $\mathbb{R}=$ 代数数 $\cup$ 超越数是两个可数集的并, 仍可数, 与 $\mathbb{R}$ 不可数矛盾. 故超越数集合不可数, 超越数存在. □

这个论证是非构造性的, 它证明了超越数极多, 却没有指出任何一个具体的超越数. 具体判定一个数是否超越, 往往极其困难. 以下两个结论在此只作陈述, 不给证明.

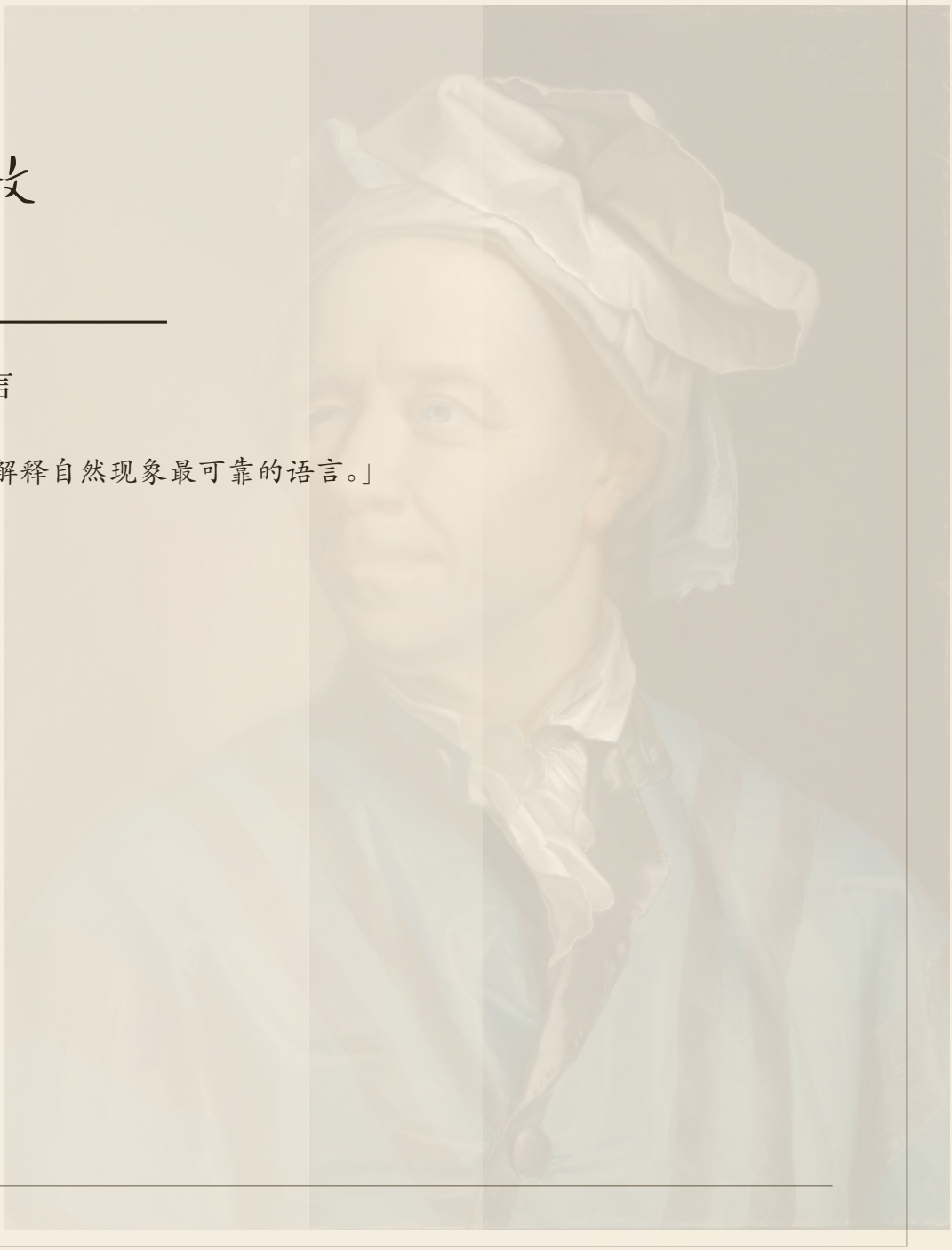
定理 2.10 (e 与 π 的超越性) 自然对数的底 e 和圆周率 π 都是超越数.

Hermite 于 1873 年证明了 e 的超越性, Lindemann 于 1882 年证明了 π 的超越性.Lindemann 的结果同时解决了古希腊三大几何难题之一的“化圆为方”问题, 即尺规作图无法作出与给定圆等积的正方形, 因为 $\sqrt{\pi}$ 是超越数, 而尺规可作的数都是代数数.

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第 3 章

函数

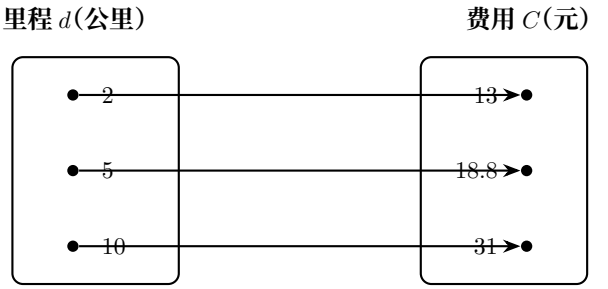


基础语言

「数学是解释自然现象最可靠的语言。」

——欧拉

第三章 函数



对应法则 $f:\ d \mapsto C$

起步价 13 元, 超出 3 公里部分每公里 2.4 元

图 3.1 从出租车计价看函数的对应关系

在物理与社会现象中, 一个量的值常常依赖于另一个量的值. 例如, 出租车的费用由行驶里程决定. 计价规则确保了对于任意给定的里程, 都存在一个**唯一确定**的费用与之对应. 数学的任务是为这类确定性的依赖关系建立一个精确的模型, 用一个普适的数学结构来刻画从一个输入集合到一个输出集合的对应关系.

本章建立函数的理论. 函数为变量之间的依赖关系提供了形式化的语言. 我们从函数作为一种特殊映射的严格定义出发, 研究单调性、奇偶性、周期性等性质, 再分析几类基本初等函数, 进而讨论函数的图像变换, 并把函数模型用于实际问题.

3.1 函数的概念

在数学中, 我们首先定义**映射** (mapping). 一个映射描述从一个集合 A 到另一个集合 B 的元素之间的对应法则. 函数是映射的特例, 其中集合 A 和 B 的元素均为**数**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.1.1 函数的定义

定义 3.1 (函数) 设 A, B 为两个非空的**实数子集**. 一个从 A 到 B 的**函数** (function) f, 是一个为 A 中**每一个**元素 x, 都在 B 中指派**唯一确定**的一个元素 y 的法则.
此函数可记作 $f : A \to B$, 其对应关系通常写作

$$y = f(x), \quad x \in A \tag{3-1}$$

我们称 x 为**自变量**, y 为**因变量**, 集合 A 为函数的**定义域**, 集合 B 为**陪域**.

函数由定义域 A、陪域 B 和对应法则 f 三部分构成. 定义要求两条来自映射的约束.

- **穷尽性**. 法则 f 必须对定义域 A 中的**所有**元素有效.
- **唯一性**.A 中任意元素 x 在 B 中的对应元素 y 是**唯**一的;一个输入不能产生多个输出.

知乎@余命数

定义允许多个不同的输入 x_1, x_2 对应同一个输出 y_0, 即 $f(x_1) = f(x_2) = y_0$.

记号 $y = f(x)$ 中, f 指整个函数(法则本身), 而 $f(x)$ 是法则 f 作用于自变量 x 后得到的具体数值.

注 3.1 (陪域与值域的辨析)
陪域和值域是两个相关但不同的概念.

- **陪域** B 是函数定义时预先指定的目标集合, 它包含了所有**可能**的输出值.
- **值域** $f(A)$ 是所有**实际**输出值的集合, 由定义域和对应法则完全确定. 其严谨定义为

$$f(A) = \{y \mid y = f(x), x \in A\} \tag{3-2}$$

值域始终是陪域的一个子集, 即 $f(A) \subseteq B$.

例如, 对于函数 $f : \mathbb{R} \to \mathbb{R}$, 其法则为 $f(x) = x^2$. 此函数的陪域是 $\mathbb{R}$, 但其值域是 $[0, +\infty)$.

我们接下来把集合 A, B 限定为数集. 于是函数表达式之间可以进行代数运算, 例如 $h(x) = f(x) + g(x)$. 我们还可以把函数与平面几何对象联系起来. 函数 f 的**图像** $\mathcal{G}_f$ 定义为点集

$$\mathcal{G}_f = \{(x, y) \in \mathbb{R}^2 \mid y = f(x), x \in A\} \tag{3-3}$$

我们借助图像在几何直观中研究函数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.1.2 函数的三要素

我们研究函数时通常关注三个对象, 即**定义域**、**对应法则**和**值域**.

定义域给出函数的合法输入集合. 任何关于函数的讨论都以定义域为前提; 我们把它称为定义域优先. 定义域记作 $\mathrm{D}(f)$ 或 $\mathrm{dom}(f)$. 若函数仅由解析式给出而未明确指定定义域, 则定义域约定为使该解析式在实数范围内有意义的所有 x 的集合, 称为**自然定义域**.

对应法则描述输入与输出之间的具体关系, 通常由**解析式**给出, 如 $f(x)=\sqrt{x^2+1}$. **图像**与**表格**也是表示对应法则的有效形式.

值域是所有函数值构成的集合, 由定义域和对应法则完全决定. 求一个函数的值域, 就是找出它在定义域上的所有输出.

3.2 定义域

知乎@余命数

我们称函数的**自然定义域** (natural domain) 为使该函数解析式在实数域 $\mathbb{R}$ 内有意义的全体自变量 x 构成的集合. 它通常由解析式中的不等式或方程确定.

3.2.1 基本代数结构的约束

我们先看函数解析式中的基本代数结构会带来哪几类典型约束.

- **分式** $\frac{f(x)}{g(x)}$. 分母不为零, 即 $g(x)\neq 0$.
- **偶次根式** $\sqrt[2n]{f(x)}$ **(**$n\in\mathbb{N}^*$**)**. 被开方数非负, 即 $f(x)\geq 0$.
- **对数** $\log_a f(x)$. 真数大于零, 即 $f(x)>0$.
- **零次幂** $[f(x)]^0$. 底数不为零, 即 $f(x)\neq 0$.

3.2.2 实际问题中的隐性定义域

当函数模型源于实际问题时, 其定义域除受上述代数结构约束外, 还必须符合问题本身的物理或逻辑意义. 例如, 代表时间、长度或数量的变量通常被限制为非负数.

例题 3.1 求函数 $f(x)=\frac{\sqrt{x-2}}{x-4}$ 的定义域.

解 该函数的定义域受两个约束. 根式 $\sqrt{x-2}$ 要求被开方数非负; 分母 $x-4$ 不得为零. 二者组成不等式组

$$\begin{cases} x-2\geq 0 \\ x-4\neq 0 \end{cases}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其解为 $x \geq 2$ 且 $x \neq 4$.

因此, 函数的定义域为 $[2,4) \cup (4,+\infty)$.

例题 3.2　求函数 $y = \sqrt{9-x^2} + \log_2(x+1)$ 的定义域.

解　函数的定义域是其所有组成部分均有意义的 x 的集合.

对于根式项 $\sqrt{9-x^2}$, 被开方数必须非负, 即 $9-x^2 \geq 0$, 解得 $-3 \leq x \leq 3$.

对于对数项 $\log_2(x+1)$, 真数必须为正, 即 $x+1>0$, 解得 $x>-1$.

函数的定义域是这两个解集的交集. 联立不等式组

$$\begin{cases} -3 \leq x \leq 3 \\ x > -1 \end{cases}$$

其解集为 $(-1,3]$. 因此, 函数的定义域为 $(-1,3]$.

例题 3.3　求函数 $f(x) = \frac{(x^2-4)^0}{\sqrt{x+1}-2}$ 的定义域.

解　此函数的定义域需同时满足三个条件.

零次幂的底数 x^2-4 不为零, 故 $x \neq 2$ 且 $x \neq -2$.

分母 $\sqrt{x+1}-2$ 不为零, 即 $\sqrt{x+1} \neq 2$, 两边平方得 $x+1 \neq 4$, 故 $x \neq 3$.

分母中的根式要求被开方数 $x+1$ 非负, 即 $x \geq -1$.

综合得 $x \in [-1,+\infty)$ 且 $x \neq 2, x \neq 3$.$x \geq -1$ 已自动排除 $x=-2$.

因此, 定义域为 $[-1,2) \cup (2,3) \cup (3,+\infty)$.

例题 3.4　求函数 $f(x) = \sqrt{\log_{0.5}\left(\frac{2x-1}{x+2}\right)}$ 的定义域.

解　这是复合函数, 其约束条件由外向内逐层分析. 最外层的根式结构要求其被开方数非负

$$\log_{0.5}\left(\frac{2x-1}{x+2}\right) \geq 0$$

由于对数底数 $0.5 \in (0,1)$, 对数函数 $\log_{0.5} t$ 严格单调递减. 又 $\log_{0.5} 1 = 0$, 于是上述不等式等价于

$$0 < \frac{2x-1}{x+2} \leq 1$$

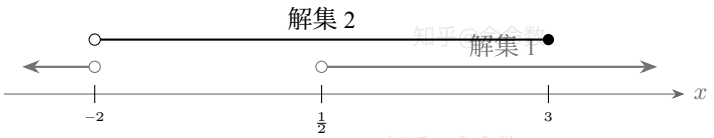
我们把它分解为两个不等式

$$\frac{2x-1}{x+2} > 0 \tag{3-4}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\frac{2x-1}{x+2} \le 1 \tag{3-5}$$

不等式 (3-4) 的解集为 $(-\infty,-2)\cup(\frac{1}{2},+\infty)$. 不等式 (3-5) 移项通分得 $\frac{x-3}{x+2}\le 0$, 其解集为 $(-2,3]$.



函数的定义域是这两个解集的交集, 即 $((-\infty,-2)\cup(\frac{1}{2},+\infty))\cap(-2,3]$.

因此, 定义域为 $(\frac{1}{2},3]$.

例题 3.5 已知函数 $f(x)$ 的定义域为 $[-1,4]$. 求函数 $g(x)=\frac{f(2x-1)}{\sqrt{x}}$ 的定义域.

解 函数 f 的定义域为 $[-1,4]$. 对于 $g(x)=\frac{f(2x-1)}{\sqrt{x}}$, f 的输入量 $2x-1$ 必须落在 $[-1,4]$ 内, 且分母 $\sqrt{x}$ 要求 $x>0$.

由 $-1\le 2x-1\le 4$ 得 $0\le x\le \frac{5}{2}$.

因此, 函数 $g(x)$ 的定义域为 $(0,\frac{5}{2}]$.

3.3 单调性

我们关心函数值怎样随自变量变化. 图像上升、下降是几何的说法, 本节把它写成一个严格的代数条件, 即单调性. 有了单调性, 我们就能比较函数值的大小, 并讨论不等式与极值问题.

3.3.1 单调性的定义与判定策略

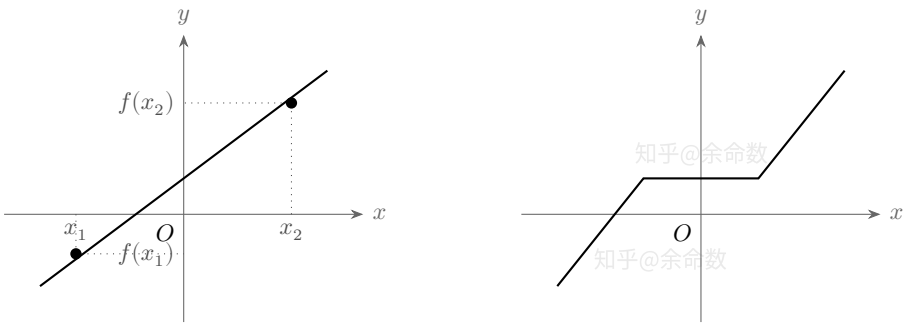
定义 3.2 (单调函数) 设函数 $y=f(x)$ 的定义域为 D, 区间 $I\subseteq D$.

- 若对于 I 内**任意** x_1,x_2, 当 $x_1<x_2$ 时, **恒有** $f(x_1)<f(x_2)$, 则称 $f(x)$ 在 I 上**严格单调递增**.
- 若对于 I 内**任意** x_1,x_2, 当 $x_1<x_2$ 时, **恒有** $f(x_1)>f(x_2)$, 则称 $f(x)$ 在 I 上**严格单调递减**.

若将上述定义中的严格不等号 $<$ 和 $>$ 分别替换为 $\le$ 和 $\ge$, 则得到**单调递增**, 或称**非减**, 和**单调递减**(或称**非增**)的定义. 严格单调函数与单调函数统称为**单调函数**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

用代数的语言说, 严格单调递增的函数保持元素的序, 我们称它**保序**; 严格单调递减的函数反转元素的序, 我们称它**逆序**.



严格单调递增　　单调递增(非严格)

图 3.2　单调性的几何直观

以严格递增为例, 判定单调性就是证明一个全称命题, 即任取 I 中的 x_1, x_2, 当 $x_1 < x_2$ 时恒有 $f(x_1) < f(x_2)$. 于是问题归结为比较两个任意函数值 $f(x_1)$ 与 $f(x_2)$ 的大小.

比较大小有两条基本途径. 一条是代数的, 直接考察差 $f(x_1) - f(x_2)$ 的符号; 当函数值恒正时, 也可以考察商 $\frac{f(x_1)}{f(x_2)}$ 与 1 的大小. 我们称它为**定义法**. 另一条是分析的, 若 f 可导, 单调性由导函数的符号决定. 我们称它为**导数法**(其理论见第16章).

3.3.2 单调性的判定方法

3.3.2.1 定义法

定义法直接应用定义. 任取 $x_1 < x_2$, 作差 $f(x_1) - f(x_2)$, 再通过恒等变形判断它的符号. 操作分四步.

1. 在指定区间 I 内, 任取 x_1, x_2, 并设 $x_1 < x_2$.
2. 构造差式 $f(x_1) - f(x_2)$.
3. 对差式进行恒等变形, 目标是将其化为若干符号易于判定的因式的乘积或商.
4. 依据区间 I 的约束条件判定差式的最终符号, 从而确定 $f(x_1)$ 与 $f(x_2)$ 的大小关系.

例题 3.6　证明函数 $g(x) = x + \frac{1}{x}$ 在区间 $(0, 1]$ 上严格单调递减, 在 $[1, +\infty)$ 上严格单调递增.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 设 x_1, x_2 是定义域中任意两点, 且 $x_1 < x_2$. 作差并通分, 得

$$g(x_1) - g(x_2) = \frac{(x_1 - x_2)(x_1x_2 - 1)}{x_1x_2}$$

由于 $x_1 < x_2$, 因子 $(x_1 - x_2)$ 恒为负. 于是差式的符号由 $(x_1x_2 - 1)$ 与分母 x_1x_2 决定.

若 $x_1, x_2 \in [1, +\infty)$ 且 $1 \leq x_1 < x_2$, 则分母 $x_1x_2 > 0$, 且 $x_1x_2 > 1$, 故因子 $(x_1x_2 - 1) > 0$. 差式的符号为 $\frac{(-)(+)}{(+)} = (-)$, 即 $g(x_1) < g(x_2)$. 因此, $g(x)$ 在 $[1, +\infty)$ 上严格单调递增.

若 $x_1, x_2 \in (0, 1]$ 且 $0 < x_1 < x_2 \leq 1$, 则分母 $x_1x_2 > 0$, 且 $0 < x_1x_2 < 1$, 故因子 $(x_1x_2 - 1) < 0$. 差式的符号为 $\frac{(-)(-)}{(+)} = (+)$, 即 $g(x_1) > g(x_2)$. 因此, $g(x)$ 在 $(0, 1]$ 上严格单调递减.

注 3.2 (单调区间的表述)
单调性是定义在**单一区间**上的性质. 例如, 函数 $f(x) = 1/x$ 在 $(-\infty, 0)$ 上严格单调递减, 在 $(0, +\infty)$ 上也严格单调递减. 但是, **不能**断言 $f(x)$ 在其定义域的并集 $(-\infty, 0) \cup (0, +\infty)$ 上是单调递减的.
取 $x_1 = -1, x_2 = 1$ 就能看出原因. 显然 $x_1 < x_2$, 但 $f(x_1) = -1 < f(x_2) = 1$, 这与单调递减的定义矛盾. 因此, 陈述函数的单调区间时, 必须分别列出, 用连词"和"或逗号连接, 不使用并集符号.

导数法

导数法(预告)

注 3.3
以下用导函数符号判定单调性, 并引用拉格朗日中值定理. **严格定义与证明见第16章**.

当 f 可导时, 我们可以用导函数 f' 的符号判定 f 的单调性.

定理 3.1 (导数与单调性的关系) 设函数 $y = f(x)$ 在区间 I 上连续, 并在 I 的内部可导.

- 若在 I 的内部恒有 $f'(x) > 0$, 则 $f(x)$ 在 I 上严格单调递增.
- 若在 I 的内部恒有 $f'(x) < 0$, 则 $f(x)$ 在 I 上严格单调递减.

证明 任取 $x_1, x_2 \in I$ 且 $x_1 < x_2$. 由于 f 在 $[x_1, x_2]$ 上连续、在 (x_1, x_2) 上可导, 由拉格朗日中值定理, 存在 $\xi \in (x_1, x_2)$, 使得

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1).$$

若 I 的内部恒有 $f'(x) > 0$, 则右端为正, 从而 $f(x_2) > f(x_1)$, 故 f 在 I 上严格单调递增. 若 I 的内部恒有 $f'(x) < 0$, 则右端为负, 从而 $f(x_2) < f(x_1)$, 故 f 在 I 上严格单调递减. □

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据这个定理，判定单调性归结为解不等式 $f'(x)>0$ 或 $f'(x)<0$.f' 的零点把区间划分为若干单调区间.

例题 3.7　求函数 $f(x)=\frac{1}{3}x^3-x^2-3x+1$ 的单调区间.

解　函数 $f(x)$ 的定义域为 $\mathbb{R}$. 其导函数为 $f'(x)=x^2-2x-3$. 令 $f'(x)=0$，即 $x^2-2x-3=0$，解得 $x=-1$ 或 $x=3$. 这两个根将实数轴划分为三个区间 $(-\infty,-1)$、$(-1,3)$、$(3,+\infty)$.

当 $x\in(-\infty,-1)\cup(3,+\infty)$ 时，$f'(x)>0$. 当 $x\in(-1,3)$ 时，$f'(x)<0$.

由于函数 $f(x)$ 在 $x=-1$ 和 $x=3$ 处连续，单调区间可扩展至闭区间. 因此，函数 $f(x)$ 的单调递增区间是 $(-\infty,-1]$ 和 $[3,+\infty)$，单调递减区间是 $[-1,3]$.

3.3.3 复合函数的单调性

复合函数 $y=f(g(x))$ 的单调性由内层函数 $u=g(x)$ 与外层函数 $y=f(u)$ 的单调性共同决定. 我们把这条规律称为同增异减.

定理 3.2 (复合函数单调性) 设复合函数 $y=f(g(x))$ 在区间 I 上有定义. 若 $u=g(x)$ 在 I 上单调，且 $y=f(u)$ 在对应的值域 $g(I)$ 上单调，则：

- 若 $f(u)$ 与 $g(x)$ 的**单调性相同**，同为非减或同为非增，则 $f(g(x))$ 在 I 上单调递增，即非减.
- 若 $f(u)$ 与 $g(x)$ 的**单调性相异**，即一非减一非增，则 $f(g(x))$ 在 I 上单调递减，即非增.

若相应的两个单调性均为严格单调，则结论也相应为严格单调.

证明　设 $x_1,x_2\in I$ 且 $x_1<x_2$. 令 $u_1=g(x_1),u_2=g(x_2)$.

若 g 和 f 均单调递增（非减），则 $x_1<x_2\implies u_1\le u_2\implies f(u_1)\le f(u_2)$. 复合函数单调递增.

若 g 和 f 均单调递减（非增），则 $x_1<x_2\implies u_1\ge u_2\implies f(u_1)\le f(u_2)$. 复合函数单调递增.

若 g 单调递增而 f 单调递减，则 $x_1<x_2\implies u_1\le u_2\implies f(u_1)\ge f(u_2)$. 复合函数单调递减.

最后一种情况（g 单调递减，f 单调递增）的证明是类似的. 若上述两个单调性均为严格单调，则相应的 $\le,\ge$ 均可替换为严格不等号，结论也相应为严格单调. □

例题 3.8　求函数 $y=\log_{0.5}(x^2-2x-3)$ 的单调区间.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

解 先求定义域. 由 $x^2-2x-3>0$ 解得 $x\in(-\infty,-1)\cup(3,+\infty)$.

我们把函数分解为外层函数 $y=f(u)=\log_{0.5}u$ 与内层函数 $u=g(x)=x^2-2x-3$.

外层函数 $f(u)$ 因底数 $0.5\in(0,1)$ 在其定义域 $(0,\infty)$ 上严格单调递减.

内层函数 $g(x)$ 是开口向上的二次函数, 对称轴为 $x=1$. 在其定义域的子区间 $(-\infty,-1)$ 上, $g(x)$ 严格单调递减; 在 $(3,+\infty)$ 上, $g(x)$ 严格单调递增.

在区间 $(-\infty,-1)$ 上, $g(x)$ 递减, $f(u)$ 递减. 根据定理, 单调性相同, 故复合函数单调递增.

在区间 $(3,+\infty)$ 上, $g(x)$ 递增, $f(u)$ 递减. 根据定理, 单调性相异, 故复合函数单调递减.

因此, 函数的单调递增区间是 $(-\infty,-1)$, 单调递减区间是 $(3,+\infty)$.

3.4 奇偶性

我们研究函数图像的对称性. 函数图像关于 y 轴对称或关于原点对称时, 函数值 $f(x)$ 与 $f(-x)$ 之间成立确定的代数恒等式, 奇偶性就是用这些恒等式描述这两种对称的工具.

知乎@余命数

3.4.1 定义

我们先关注两种对称, 即关于 y 轴的轴对称和关于原点的中心对称.

知乎@余

定义 3.3 (奇偶性) 设函数 $f(x)$ 的定义域 D 是一个**关于原点对称**的数集, 即, 若 $x\in D$, 则必有 $-x\in D$.

- 如果对于任意 $x\in D$, 恒有 $f(-x)=f(x)$, 则称 $f(x)$ 为**偶函数**. 其图像关于 y 轴对称.
- 如果对于任意 $x\in D$, 恒有 $f(-x)=-f(x)$, 则称 $f(x)$ 为**奇函数**. 其图像关于原点中心对称.

该定义有两层不可分割的内容. 定义域的对称性确保, 若 x 属于定义域, 则 $-x$ 也属于定义域, $f(-x)$ 有意义. 恒等式 $f(-x)=\pm f(x)$ 是奇偶性的判据, 且需对定义域内每一个 x 成立.

例题 3.9 (1) 下列函数在定义域上是偶函数的是(　)

A. $f(x)=\frac{x^2+x}{x+1}$

B. $f(x)=x\sin x$

C. $f(x)=\log_2(\sqrt{x^2+1}-x)$

知乎@余命数

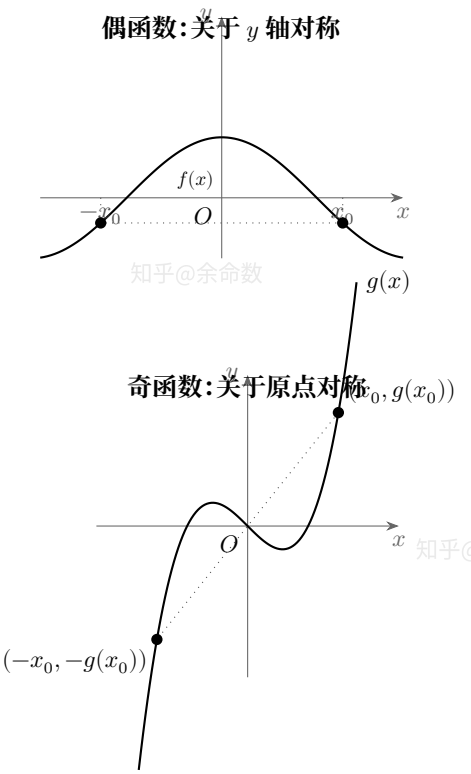


图 3.3　偶函数与奇函数的几何直观

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

D. $f(x)=\frac{\sqrt{1-x^2}}{|x+2|-2}$

(2) 已知定义在 $\mathbb{R}$ 上的函数 $f(x)$ 满足 $f(x+y)=xf(y)+yf(x)$. 定义在 $\mathbb{R}$ 上的函数 $g(x)$ 满足 $g(x+1)=(x+1)(x^2+2x)$, 则(　)

A. $f(x)$ 不是奇函数

B. $f(x)$ 既是奇函数也是偶函数

C. $g(x)$ 是奇函数

D. $g(x)$ 既不是奇函数也不是偶函数

知乎@余命数

证明 (1) 先确定每个选项的定义域是否关于原点对称, 再检验 $f(-x)$ 与 $f(x)$ 的关系.

A. $f(x)=\dfrac{x^2+x}{x+1}$ 的定义域为 $\{x\in\mathbb{R}\mid x\neq -1\}$, 不关于原点对称, 故不是偶函数.

B. $f(x)=x\sin x$ 的定义域为 $\mathbb{R}$, 关于原点对称. 又

$$f(-x)=(-x)\sin(-x)=(-x)(-\sin x)=x\sin x=f(x),$$

因此 $f(x)$ 是偶函数.

知乎@余命数

C. 观察 $f(x)=\log_2(\sqrt{x^2+1}-x)$. 注意到

$$\sqrt{x^2+1}-x=\frac{1}{\sqrt{x^2+1}+x},$$

因此

$$f(x)=\log_2\Big(\frac{1}{\sqrt{x^2+1}+x}\Big)=-\log_2(\sqrt{x^2+1}+x).$$

已知 $\log_2(\sqrt{x^2+1}+x)$ 是奇函数, 故 $f(x)$ 也是奇函数.

D. $f(x)=\dfrac{\sqrt{1-x^2}}{|x+2|-2}$ 的定义域由 $1-x^2\geq 0$ 与 $|x+2|-2\neq 0$ 决定. 前者给出 $-1\leq x\leq 1$, 后者给出 $x\neq 0$ 且 $x\neq -4$. 故定义域为 $[-1,0)\cup(0,1]$, 关于原点对称. 但

$$f(-x)=\frac{\sqrt{1-(-x)^2}}{|-x+2|-2}=\frac{\sqrt{1-x^2}}{|2-x|-2},$$

一般情况下 $f(x)\neq f(-x)$, 故不是偶函数.

只有 B 选项是偶函数.

(2) 对 $f(x)$, 由 $f(x+y)=xf(y)+yf(x)$ 对任意 $x,y\in\mathbb{R}$ 成立. 令 $x=0$ 得

知乎@余命数

$$f(y)=yf(0).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此 $f(x)=kx$, 其中 $k=f(0)$. 把它代回原方程.

$$k(x+y)=x(ky)+y(kx) \implies kx+ky=2kxy.$$

该式对任意 x,y 成立只能通过 $k=0$, 所以 $f(x)\equiv 0$, 它既是奇函数又是偶函数.

对 $g(x)$, 已知 $g(x+1)=(x+1)(x^2+2x)$. 令 $t=x+1$, 则 $x=t-1$, 于是

$$\begin{aligned} g(t)&=t((t-1)^2+2(t-1))=t(t^2-2t+1+2t-2)\\ &=t(t^2-1)=t^3-t. \end{aligned}$$

因此 $g(x)=x^3-x$. 于是 $g(-x)=(-x)^3-(-x)=-x^3+x=-(x^3-x)=-g(x)$, 故 $g(x)$ 是奇函数.

故正确选项是 B 和 C. □

知乎@余命数

3.4.2 函数的奇偶分解定理

奇偶性也适用于一般函数. 任意定义在对称区间上的函数都可以唯一地分解为一个偶函数与一个奇函数之和.

知乎@余命数

定理 3.3 (奇偶分解定理) 任何一个定义在对称区间 D 上的函数 $f(x)$, 都可以**唯**一地表示为一个偶函数 $f_e(x)$ 与一个奇函数 $f_o(x)$ 的和, 即

$$f(x)=f_e(x)+f_o(x)$$

其中,

$$f_e(x)=\frac{f(x)+f(-x)}{2} \quad 且 \quad f_o(x)=\frac{f(x)-f(-x)}{2}$$

知乎@余命数

证明 我们先证明, 若分解存在, 则其形式必如定理给出; 再验证该形式确实给出分解.

不妨设 $f(x)$ 可以分解为 $f(x)=f_e(x)+f_o(x)$, 其中 $f_e(x)$ 为偶函数, $f_o(x)$ 为奇函数. 将 x 替换为 $-x$, 得

$$f(-x)=f_e(-x)+f_o(-x)=f_e(x)-f_o(x)$$

现在得到一个关于 $f_e(x)$ 和 $f_o(x)$ 的线性方程组.

$$\begin{aligned} f_e(x)+f_o(x)&=f(x)\\ f_e(x)-f_o(x)&=f(-x) \end{aligned}$$

知乎@余命数

两式相加除以 2, 得 $f_e(x)=\frac{f(x)+f(-x)}{2}$; 两式相减除以 2, 得 $f_o(x)=\frac{f(x)-f(-x)}{2}$. 故若分解存在, 其形式必如上式. 尚需验证这样构造出的 $f_e(x)$ 与 $f_o(x)$ 分别是偶函数与奇函数. 对 $f_e(x)$, 有

$$f_e(-x)=\frac{f(-x)+f(-(-x))}{2}=\frac{f(-x)+f(x)}{2}=f_e(x),$$

故 $f_e(x)$ 是偶函数. 对 $f_o(x)$, 有

$$f_o(-x)=\frac{f(-x)-f(-(-x))}{2}=\frac{f(-x)-f(x)}{2}=-f_o(x),$$

故 $f_o(x)$ 是奇函数. 并且

$$f_e(x)+f_o(x)=\frac{f(x)+f(-x)}{2}+\frac{f(x)-f(-x)}{2}=f(x).$$

以上推导同时给出存在性与唯一性. 任何满足条件的分解必须满足该线性方程组, 而该方程组有唯一解. 证毕. □

该定理把对称区间上的任意函数分解为偶部 $f_e(x)$ 与奇部 $f_o(x)$.

例题 3.10 将函数 $f(x)=e^x$ 分解为一个偶函数与一个奇函数之和.

证明 函数 $f(x)=e^x$ 的定义域为 $\mathbb{R}$, 关于原点对称. 根据奇偶分解定理, 其偶部为

$$f_e(x)=\frac{f(x)+f(-x)}{2}=\frac{e^x+e^{-x}}{2},$$

其奇部为

$$f_o(x)=\frac{f(x)-f(-x)}{2}=\frac{e^x-e^{-x}}{2}.$$

这两个函数分别称为**双曲余弦函数**和**双曲正弦函数**, 记作 $\cosh x$ 和 $\sinh x$. 于是 $e^x=\cosh x+\sinh x$. □

3.4.3 运算封闭性与对称推广

奇函数与偶函数的奇偶性在加法和乘法下保持稳定. 我们直接验证这些运算法则.

不妨设 $f(x)$ 和 $g(x)$ 是两个定义在同一个对称区间 D 上的函数.

首先考虑加法. 令 $h(x)=f(x)+g(x)$. 考察 $h(x)$ 的奇偶性. 若 $f(x)$ 和 $g(x)$ 均为偶函数, 那么根据定义, 有 $f(-x)=f(x)$ 和 $g(-x)=g(x)$. 于是,

$$h(-x)=f(-x)+g(-x)=f(x)+g(x)=h(x)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故,两个偶函数的和仍然是一个偶函数. 若 $f(x)$ 和 $g(x)$ 均为奇函数,那么 $f(-x)=-f(x)$ 和 $g(-x)=-g(x)$. 于是,

$$h(-x)=f(-x)+g(-x)=-f(x)-g(x)=-(f(x)+g(x))=-h(x)$$

故,两个奇函数的和仍然是一个奇函数.

对乘法同理. 令 $p(x)=f(x)g(x)$. 若 $f(x)$ 和 $g(x)$ 均为偶函数,则

$$p(-x)=f(-x)g(-x)=f(x)g(x)=p(x)$$

乘积是偶函数. 若 $f(x)$ 是偶函数而 $g(x)$ 是奇函数,则

$$p(-x)=f(-x)g(-x)=f(x)(-g(x))=-f(x)g(x)=-p(x)$$

乘积是奇函数. 当 $f(x)$ 和 $g(x)$ 均为奇函数时,

$$p(-x)=f(-x)g(-x)=(-f(x))(-g(x))=f(x)g(x)=p(x)$$

故**两个奇函数的乘积是一个偶函数**.

这些运算法则与整数的符号法则结构相似. 将偶函数看作正号或 $+1$,将奇函数看作负号或 -1,函数的乘法规则对应于符号的乘法.

$$\begin{aligned}
&\text{偶}\times\text{偶}\to\text{偶} \quad &\longleftrightarrow \quad &(+1)\times(+1)=+1\\
&\text{偶}\times\text{奇}\to\text{奇} \quad &\longleftrightarrow \quad &(+1)\times(-1)=-1\\
&\text{奇}\times\text{奇}\to\text{偶} \quad &\longleftrightarrow \quad &(-1)\times(-1)=+1
\end{aligned}$$

3.4.3.1 对称性的推广

关于 y 轴和原点的对称,分别是关于直线 $x=0$ 和点 $(0,0)$ 的对称. 这一概念可直接推广到任意对称轴与对称中心.

▷ **关于直线 $x=a$ 对称** 若函数 $f(x)$ 的图像关于直线 $x=a$ 对称,则对任意使 $a\pm t$ 在定义域内的 t,都有

$$f(a-t)=f(a+t).$$

令 $x=a-t$,则 $t=a-x$,于是 $a+t=2a-x$. 代入后得等价判据

$$f(x)=f(2a-x).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

▷ **关于点 (a,b) 对称** 若函数 $f(x)$ 的图像关于点 (a,b) 中心对称，则对任意使 $a\pm t$ 在定义域内的 t，点 $(a-t,f(a-t))$ 与 $(a+t,f(a+t))$ 的中点为 (a,b)，即

$$\frac{f(a-t)+f(a+t)}{2}=b.$$

整理得 $f(a-t)+f(a+t)=2b$. 令 $x=a-t$，则 $a+t=2a-x$，代入后得等价判据

$$f(x)+f(2a-x)=2b.$$

标准的奇偶性正是这些广义对称性的特例. 当对称轴为 $x=0$ $(a=0)$ 时，$f(x)=f(-x)$，这正是偶函数的定义. 当对称中心为 $(0,0)$ $(a=0,b=0)$ 时，$f(x)+f(-x)=0$，即 $f(-x)=-f(x)$，这正是奇函数的定义.

3.4.3.2 一个更具操作性的判别方法

在处理具体函数方程以判定对称性时，常遇到 $f(x+a)=f(-x+b)$ 这类表达式. 直接套用 $f(x)=f(2a-x)$ 的形式未必方便. 其共性在于两个自变量 x 与 $2a-x$ 之和为常数 $2a$. 由此可得识别对称性的直接工具.

定理 3.4 (对称性的统一判据) 设函数 $f(x)$ 的定义域在相应对称变换下封闭，并且下列函数方程对所有使式子有意义的自变量恒成立.

1. **轴对称**. 若该方程可化为 $f(X_1)=f(X_2)$ 的形式，且自变量之和 $X_1+X_2=k$(常数)，则函数 $f(x)$ 的图像关于直线 $x=\frac{k}{2}$ 对称.
2. **点对称**. 若该方程可化为 $f(X_1)+f(X_2)=k$ 的形式，其中 k 为常数，且自变量之和 $X_1+X_2=K$(常数)，则函数 $f(x)$ 的图像关于点 $\left(\frac{K}{2},\frac{k}{2}\right)$ 中心对称.

证明 1. 对于 $f(X_1)=f(X_2)$ 且 $X_1+X_2=k$，令 $a=k/2$. 可设 $X_1=a+t$. 那么 $X_2=k-X_1=2a-(a+t)=a-t$. 于是原方程化为 $f(a+t)=f(a-t)$，这正是函数关于 $x=a$ 对称的定义式.

2. 对于 $f(X_1)+f(X_2)=k$ 且 $X_1+X_2=K$，令 $a=K/2,b=k/2$. 同样设 $X_1=a+t$，则 $X_2=K-X_1=2a-(a+t)=a-t$. 于是原方程化为 $f(a+t)+f(a-t)=2b$，这正是函数关于点 (a,b) 对称的定义式. □

该定理把函数方程转化为几何对称性的判定，关键在于检验自变量之和是否为常数.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

推论 3.1　我们现在可以迅速判定一些常见形式的对称性.

- 若 $f(x+a)=f(-x+b)$, 自变量之和为 $(x+a)+(-x+b)=a+b$. 故对称轴为 $x=\frac{a+b}{2}$. 当 $b=0$ 时, $f(x+a)=f(-x)$, 对称轴为 $x=a/2$.
- 若 $f(x+a)+f(-x+b)=c$, 自变量之和为 $a+b$, 函数值之和为 c. 故对称中心为 $\left(\frac{a+b}{2},\frac{c}{2}\right)$. 特别地, 若 $f(x+a)=-f(-x)$, 则 $f(x+a)+f(-x)=0$, 自变量之和为 a, 函数值之和为 0, 故对称中心为 $(\frac{a}{2},0)$.

知乎@余命数

证明　第一条中令 $X_1=x+a$, $X_2=-x+b$, 则 $X_1+X_2=a+b$, 由对称性的统一判据立即得到对称轴 $x=(a+b)/2$. 当 $b=0$ 时即为 $x=a/2$.

第二条中同样令 $X_1=x+a$, $X_2=-x+b$, 则 $X_1+X_2=a+b$, 且函数值之和为常数 c, 由统一判据得到对称中心 $((a+b)/2,c/2)$. 特别地, $f(x+a)=-f(-x)$ 等价于 $f(x+a)+f(-x)=0$, 此时 $b=0,c=0$, 因而对称中心为 $(a/2,0)$. □

例题 3.11　设函数 $f(x)$ 定义在 $\mathbb{R}$ 上. 已知 $f(x-1)$ 是一个偶函数, 而 $f(x+1)$ 是一个奇函数. 试证明 $f(x)$ 是一个周期函数.

证明　我们用广义对称性推导周期性. 先把题设写成 $f(x)$ 的对称性.

设 $g(x)=f(x-1)$. 由 $g(x)$ 为偶函数, 有 $g(x)=g(-x)$, 即 $f(x-1)=f(-x-1)$. 这是 $f(X_1)=f(X_2)$ 形式, 自变量之和 $(x-1)+(-x-1)=-2$. 根据上述定理, $f(x)$ 的图像关于直线 $x=-1$ 对称.

设 $h(x)=f(x+1)$. 由 $h(x)$ 为奇函数, 有 $h(x)=-h(-x)$, 即 $f(x+1)=-f(-x+1)$, 亦即 $f(x+1)+f(-x+1)=0$. 这是 $f(X_1)+f(X_2)=k$ 形式, 自变量之和 $(x+1)+(-x+1)=2$, 函数值之和为 0. 故 $f(x)$ 的图像关于点 $(1,0)$ 中心对称.

于是 $f(x)$ 满足

$$f(-1+t)=f(-1-t),\qquad \forall t\in\mathbb{R},$$

$$f(1+t)=-f(1-t),\qquad \forall t\in\mathbb{R}.$$

须寻找常数 T 使 $f(x+T)=f(x)$, 即把两条对称性复合使用.

考虑 $f(x+4)$. 为应用关于点 $(1,0)$ 的对称性, 把它写成 $f(1+(x+3))$.

$$f(x+4)=f(1+(x+3))=-f(1-(x+3))=-f(-x-2).$$

再对 $f(-x-2)$ 应用关于直线 $x=-1$ 的对称性, 把它写成 $f(-1+(-x-1))$.

$$f(-x-2)=f(-1+(-x-1))=f(-1-(-x-1))=f(x).$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

代回得

$$f(x+4) = -f(x).$$

该式说明函数具有半周期为 4 的反向重复. 再将 x 换为 $x+4$.

$$f(x+8) = -f(x+4).$$

知乎@余命数

结合两式,

$$f(x+8) = -f(x+4) = -(-f(x)) = f(x).$$

知乎@余命数

故 $f(x)$ 以 8 为周期.

轴对称与中心对称的复合由此产生平移不变性, 即周期性. □

例题 3.12 设函数 $f(x)$ 是定义在 $\mathbb{R}$ 上的增函数. 在两种不同的附加条件下讨论其性质.

情形一: 若 $f(x)$ 满足柯西函数方程 $f(x+y) = f(x) + f(y)$.

1. 求 $f(0)$ 的值.
2. 求证 $f(x)$ 是奇函数.
3. 若 $f(x^2+1) + f(3x-5) < 0$, 求实数 x 的取值范围.

情形二: 若 $f(x)$ 满足函数方程 $yf(x) - xf(y) = xy(x^2-y^2)$.

1. 求 $f(0)$ 的值.
2. 求证 $f(x)$ 是奇函数.
3. 若 $f(x^2+1) + f(3x-5) < 0$, 求实数 x 的取值范围.

证明 对函数方程赋值, 先求 $f(0)$, 再推导奇偶性, 最后用单调性解不等式.

这是柯西函数方程 $f(x+y) = f(x) + f(y)$.

(1) 令 $x = y = 0$, 得

$$f(0+0) = f(0) + f(0) \implies f(0) = 2f(0),$$

故 $f(0) = 0$.

(2) 令 $y = -x$, 得

知乎@余命数

$$f(x+(-x)) = f(x) + f(-x) \implies f(0) = f(x) + f(-x).$$

由 $f(0) = 0$ 得 $f(-x) = -f(x)$, 故 $f(x)$ 是奇函数.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(3) 不等式 $f(x^2+1)+f(3x-5)<0$ 可改写为 $f(x^2+1)<-f(3x-5)$. 由奇函数性质,

$$-f(3x-5)=f(5-3x),$$

于是原不等式等价于

$$f(x^2+1)<f(5-3x).$$

又 $f(x)$ 是增函数, 自变量的大小关系与函数值一致, 故

$$x^2+1<5-3x.$$

整理得 $x^2+3x-4<0$, 即 $(x+4)(x-1)<0$. 解得 $x\in(-4,1)$.

情形二的函数方程为

$$yf(x)-xf(y)=xy(x^2-y^2).$$

(1) 求 $f(0)$. 直接令 $y=0$ 会使方程两端皆为 0, 无法分离 $f(0)$. 取 $y=1,x=0$:

$$1\cdot f(0)-0\cdot f(1)=0\cdot 1\cdot(0^2-1^2)\implies f(0)=0.$$

(2) 证明奇偶性. 仍令 $y=-x$, 得

$$(-x)f(x)-xf(-x)=x(-x)(x^2-(-x)^2),$$

$$-x[f(x)+f(-x)]=-x^2(x^2-x^2)=0.$$

因此 $-x[f(x)+f(-x)]=0$ 对一切 $x\in\mathbb{R}$ 成立. 当 $x\neq 0$ 时, 约去 $-x$ 得 $f(x)+f(-x)=0$; 当 $x=0$ 时, $f(0)=0$ 使等式亦成立. 故 $f(-x)=-f(x)$, $f(x)$ 是奇函数.

(3) 解不等式. 情形二虽与情形一的函数方程形式不同, 但推出的奇函数与增函数性质完全相同. 因此不等式 $f(x^2+1)+f(3x-5)<0$ 的求解过程亦相同, 解集仍为 $(-4,1)$.

□

3.4.3.3 构造具有特定对称性的函数

构造具有特定奇偶性的函数, 其规律直接来自 $f(-x)=\pm f(x)$. 我们先以指数函数为例.

3.4.3.4 源于指数函数的对称构造

指数函数 $g(x)=a^x$ $(a>0,a\neq 1)$ 本身非奇非偶. 奇偶分解定理给出它的偶部与奇部, 我们由此得到两类常见函数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其偶部为 $f_e(x)=\frac{a^x+a^{-x}}{2}$, 奇部为 $f_o(x)=\frac{a^x-a^{-x}}{2}$. 验证, 任意常数 m 与它们的乘积, 即

知乎@余命数

$$f(x)=m(a^x+a^{-x}) \quad 与 \quad g(x)=m(a^x-a^{-x})$$

分别是偶函数和奇函数的典范. 它们是构造更复杂函数的基础单元.

一个看似不同的常见奇函数模型是 $f(x)=m\frac{a^x-1}{a^x+1}$. 分析其对称性的来源.

证明 其定义域为 $\{x\in\mathbb{R}\mid a^x\neq -1\}$, 即 $\mathbb{R}$. 这是一个对称的定义域. 我们考察 $f(-x)$.

知乎@余命数

$$f(-x)=m\frac{a^{-x}-1}{a^{-x}+1}$$

这里的关键技巧是分子分母同乘 a^x, 以此恢复与 $f(x)$ 的联系.

$$f(-x)=m\frac{(a^{-x}-1)a^x}{(a^{-x}+1)a^x}=m\frac{1-a^x}{1+a^x}=-m\frac{a^x-1}{a^x+1}=-f(x)$$

故 $f(x)$ 是一个奇函数. 实际上, 经简单代数变形可见, 它与先前导出的双曲函数模型有内在联系. □

知乎@余命数

3.4.3.5 源于对数函数的对称构造

对于对数函数 $f(x)=\log_a g(x)$, 其奇偶性完全取决于其宗量 $g(x)$ 的性质. 若 $f(x)$ 为奇函数, 则必须满足

知乎@余命数 知乎@余命数

$$f(-x)=\log_a g(-x)=-f(x)=-\log_a g(x)=\log_a\frac{1}{g(x)}$$

这要求其宗量必须满足关系 $g(-x)=\frac{1}{g(x)}$. 若 $f(x)$ 为偶函数, 则必须满足

$$f(-x)=\log_a g(-x)=f(x)=\log_a g(x)$$

这要求其宗量必须是偶函数, 即 $g(-x)=g(x)$.

基于此原理, 可构造出大量的对数型奇偶函数.

例题 3.13 判断函数 $f(x)=\log_a(\sqrt{x^2+1}+x)$ 的奇偶性, 其中 $a>0$ 且 $a\neq 1$.

证明 函数定义域要求 $\sqrt{x^2+1}+x>0$. 注意到 $\sqrt{x^2+1}>\sqrt{x^2}=|x|$, 因此 $\sqrt{x^2+1}+x>|x|+x$. 当 $x\geq 0$ 时, $|x|+x=2x\geq 0$. 当 $x<0$ 时, $|x|+x=0$. 故 $\sqrt{x^2+1}+x$ 恒为正, 定义域为 $\mathbb{R}$.

考察其宗量 $g(x)=\sqrt{x^2+1}+x$.

知乎@余命数

$$g(-x)=\sqrt{(-x)^2+1}-x=\sqrt{x^2+1}-x$$

为检验 $g(-x)$ 是否等于 $1/g(x)$, 计算其乘积, 由平方差公式

$$\begin{aligned} g(x)g(-x) &= (\sqrt{x^2+1}+x)(\sqrt{x^2+1}-x) \\ &= (\sqrt{x^2+1})^2 - x^2 = (x^2+1)-x^2 = 1. \end{aligned}$$

因此, $g(-x) = \frac{1}{g(x)}$. 于是

$$f(-x) = \log_a g(-x) = \log_a \frac{1}{g(x)} = -\log_a g(x) = -f(x).$$

故 $f(x) = \log_a(\sqrt{x^2+1}+x)$ 是奇函数. □

知乎@余命数

同理, 形如 $f(x) = \log_a \frac{m+x}{m-x}$ 的函数, 其宗量 $g(x) = \frac{m+x}{m-x}$ 也满足 $g(-x) = \frac{1}{g(x)}$, 故它也是奇函数.

3.4.3.6 普适的对称性构造法则

除了针对特定函数类型的改造, 还存在一些更为普适的构造方法.

- **绝对值化构造偶函数**. 对于定义域关于原点对称的任意函数 $f(x)$,函数 $h(x) = f(|x|)$ 是偶函数.

  证明 $h(-x) = f(|-x|) = f(|x|) = h(x)$. □

  这个构造的几何意义是, 舍弃 y 轴左侧的图像, 然后将 y 轴右侧的图像反射到左侧, 从而强制地创造出关于 y 轴的对称性.

知乎@余命数

- **复合函数法则**. 设函数 $f(x)$ 与 $g(x)$ 的奇偶性已知. 复合函数 $f(g(x))$ 的奇偶性遵循内偶则偶、内奇看外的原则.
  - 若 $g(x)$ 是偶函数, 则 $f(g(x))$ 必为偶函数(为什么与 $f(x)$ 的奇偶性无关?).
  - 若 $g(x)$ 是奇函数, 则 $f(g(x))$ 的奇偶性与 $f(x)$ 相同.

知乎@余命数

- **平凡模型**. 最简单的偶函数是常数函数 $f(x) = c$, 最简单(非零)的奇函数是正比例函数 $f(x) = kx$.

给定函数的奇偶性往往可归入上述某种构造原理.

例题 3.14 已知函数 $f(x) = \frac{2^x-1}{2^x+1}$, $g(x) = \cos x$. 判断函数 $h(x) = f(|x|)\cdot g(x)$ 的奇偶性.

证明 $h(x)$ 由两部分相乘而成, 须分别确定其奇偶性.

函数 $f(x)$ 契合奇函数模型 $m\frac{a^x-1}{a^x+1}$(此处 $m=1, a=2$), 故 $f(x)$ 是奇函数.

再看 $f(|x|)$, 此即绝对值化构造偶函数法则的直接应用. 令 $p(x) = f(|x|)$, 则 $p(-x) = f(|-x|) = f(|x|) = p(x)$, 故 $p(x)$ 是偶函数.

函数 $g(x)=\cos x$ 是基本的偶函数, $g(-x)=\cos(-x)=\cos x=g(x)$.

最终, $h(x)$ 是偶函数 $p(x)=f(|x|)$ 与偶函数 $g(x)=\cos x$ 的乘积. 依运算法则 (偶 $\times$ 偶 = 偶), $h(x)$ 是偶函数. □

例题 3.15 求解不等式.$\log_2(\sqrt{x^2+1}+x)+\log_2(\sqrt{4x^2+1}-2x)>0$.

证明 直接解这个不等式并不容易. 我们先把它与已知模型联系起来.

令 $f(x)=\log_2(\sqrt{x^2+1}+x)$. 已证明, 这是一个定义在 $\mathbb{R}$ 上的奇函数.

现在观察不等式中的第二项, $\log_2(\sqrt{4x^2+1}-2x)$. 它的形式与 $f(x)$ 的共轭形式 $\log_2(\sqrt{x^2+1}-x)$ 极其相似. 注意到

$$\log_2(\sqrt{4x^2+1}-2x)=\log_2(\sqrt{(2x)^2+1}-2x)$$

已知 $\sqrt{u^2+1}-u=\frac{1}{\sqrt{u^2+1}+u}$. 因此,

$$\begin{aligned}\log_2(\sqrt{(2x)^2+1}-2x)&=\log_2\left(\frac{1}{\sqrt{(2x)^2+1}+2x}\right)\\&=-\log_2(\sqrt{(2x)^2+1}+2x)=-f(2x).\end{aligned}$$

原不等式经上述函数性质化简为

$$f(x)-f(2x)>0\implies f(x)>f(2x)$$

尚需考察 $f(x)$ 的单调性. 函数 $u(x)=\sqrt{x^2+1}+x$ 在 $\mathbb{R}$ 上严格递增. 若 $x_1<x_2$, 则

$$u(x_2)-u(x_1)=(x_2-x_1)\left(1+\frac{x_1+x_2}{\sqrt{x_2^2+1}+\sqrt{x_1^2+1}}\right)>0,$$

其中分母严格大于 $|x_1+x_2|$. 又因为 $v(u)=\log_2 u$ 是增函数, 根据复合函数单调性法则, $f(x)=v(u(x))$ 是 $\mathbb{R}$ 上的增函数.

因为 $f(x)$ 是增函数, 所以由 $f(x)>f(2x)$ 可以直接得到 $x>2x$, 故 $-x>0$, 即 $x<0$.

所以, 原不等式的解集为 $(-\infty,0)$. □

例题 3.16 设 $g(x)=\ln\frac{1+x^2}{1-x^2}$, $h(x)=x^3\cos x$. 判断复合函数 $F(x)=g(h(x))$ 的奇偶性.

证明 按先内后外的次序, 先确定内外层函数的奇偶性.

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

内层函数 $h(x)=x^3\cos x$. 它是奇函数 x^3 与偶函数 $\cos x$ 的乘积. 根据运算法则(奇 $\times$ 偶 = 奇), $h(x)$ 是奇函数.

外层函数 $g(x)=\ln\frac{1+x^2}{1-x^2}$. 其定义域由 $\frac{1+x^2}{1-x^2}>0 \implies 1-x^2>0 \implies -1<x<1$ 给出, 是对称区间. 考察 $g(-x)$:

$$g(-x)=\ln\frac{1+(-x)^2}{1-(-x)^2}=\ln\frac{1+x^2}{1-x^2}=g(x)$$

故 $g(x)$ 是偶函数.

复合函数 $F(x)=g(h(x))$ 的定义域为

$$D=\{x\in\mathbb{R}\mid -1<h(x)<1\}.$$

由于 $h(x)$ 是奇函数, $x\in D$ 当且仅当 $-x\in D$, 故 D 关于原点对称. 现在判断 $F(x)$ 在其定义域 D 上的奇偶性. 对任意 $x\in D$, 考察 $F(-x)$:

$$F(-x)=g(h(-x))$$

因为 $h(x)$ 是奇函数, 所以 $h(-x)=-h(x)$. 代入得

$$F(-x)=g(-h(x))$$

又因为 $g(x)$ 是偶函数, 所以 $g(-u)=g(u)$ 对其定义域内任意的 u 成立. 于是:

$$F(-x)=g(-h(x))=g(h(x))=F(x)$$

因此, $F(x)$ 是一个偶函数. □

例题 3.17　设定义在 $\mathbb{R}\setminus\{0\}$ 上的偶函数 $f(x)$ 满足, 当 $x>0$ 时, $f(x)=x+\frac{1}{x}-2$. 求解方程 $f(x)=3$.

证明　利用偶函数性质把已知的正半轴解析式推广到负半轴, 再在整个定义域内求解方程.

分两种情况讨论.

情形一: $x>0$　当 $x>0$ 时, 方程 $f(x)=3$ 直接等价于

$$x+\frac{1}{x}-2=3 \implies x+\frac{1}{x}-5=0$$

两边同乘 x (因为 $x>0, x\neq 0$), 得到 $x^2-5x+1=0$. 利用求根公式, 解得 $x=\frac{5\pm\sqrt{25-4}}{2}=\frac{5\pm\sqrt{21}}{2}$. 由于 $\sqrt{16}<\sqrt{21}<\sqrt{25}$, 即 $4<\sqrt{21}<5$, 所以两个根 $\frac{5+\sqrt{21}}{2}$ 和 $\frac{5-\sqrt{21}}{2}$ 均为正数, 符合本情形的假设.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

情形二：$x<0$ 当 $x<0$ 时，不能直接使用给定的解析式. 但已知 $f(x)$ 是偶函数，因此

$$f(x)=f(-x)$$

因为 $x<0$，所以 $-x>0$. 故可对 $f(-x)$ 使用已知的解析式.

$$f(x)=f(-x)=(-x)+\frac{1}{-x}-2=-x-\frac{1}{x}-2$$

于是，在 $x<0$ 的区间上，方程 $f(x)=3$ 等价于

$$-x-\frac{1}{x}-2=3 \implies -x-\frac{1}{x}-5=0$$

两边同乘 $-x$（注意 $-x>0$），得到 $x^2+5x+1=0$. 解得 $x=\frac{-5\pm\sqrt{25-4}}{2}=\frac{-5\pm\sqrt{21}}{2}$. 两个根 $\frac{-5+\sqrt{21}}{2}$ 和 $\frac{-5-\sqrt{21}}{2}$ 均为负数，符合本情形的假设.

方程 $f(x)=3$ 的所有解为 $\frac{5\pm\sqrt{21}}{2}$ 和 $\frac{-5\pm\sqrt{21}}{2}$. □

例题 3.18 设 $f(x)=\log_a g(x)$ 为定义在 $\mathbb{R}$ 上的奇函数. 求证：函数 $h(x)=\frac{(g(x))^2-1}{(g(x))^2+1}$ 也是奇函数.

证明 由 $f(x)=\log_a g(x)$ 是奇函数，有 $f(-x)=-f(x)$，即

$$\log_a g(-x)=-\log_a g(x)=\log_a\left(\frac{1}{g(x)}\right).$$

对数函数是单射，两边宗量相等，故

$$g(-x)=\frac{1}{g(x)}.$$

这一关系是后续推导的出发点. 据此比较 $h(-x)$ 与 $h(x)$：

$$h(-x)=\frac{(g(-x))^2-1}{(g(-x))^2+1}$$

代入关系 $g(-x)=1/g(x)$，得

$$h(-x)=\frac{\left(\frac{1}{g(x)}\right)^2-1}{\left(\frac{1}{g(x)}\right)^2+1}=\frac{\frac{1}{(g(x))^2}-1}{\frac{1}{(g(x))^2}+1}$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

将分子分母同乘以 $(g(x))^2$ 化简繁分数:

$$h(-x)=\frac{1-(g(x))^2}{1+(g(x))^2}=-\frac{(g(x))^2-1}{(g(x))^2+1}=-h(x)$$

此式表明 $h(x)$ 是一个奇函数. 证毕.

实际上, 可观察到 $h(x)$ 的结构与典范奇函数模型 $\frac{a^u-1}{a^u+1}$ 极为相似. 若令 $u(x)=\log_a(g(x))^2$, 则 $a^{u(x)}=(g(x))^2$. 于是 $h(x)=\frac{a^{u(x)}-1}{a^{u(x)}+1}$. □

3.5 周期性

知乎@余命数

我们还关心函数的图像在水平平移后是否与自己重合. 这就是周期性, 即函数值每隔一段固定距离重复一次, 于是我们只需研究一个周期内的行为.

知乎@余命数

3.5.1 周期性的定义

函数图像沿水平方向平移后与自身重合. 下面把它写成严格的代数条件.

定义 3.4 (周期函数) 设函数 $y=f(x)$ 的定义域为 D. 若存在一个**非零常数** T, 使得对于任意 $x\in D$, 恒有 $x+T\in D$ 与 $x-T\in D$, 并且满足

$$f(x+T)=f(x) \tag{3-6}$$

则称 $f(x)$ 为**周期函数**, T 称为它的一个**周期**.

此定义蕴含三个必要要素. 其一. 周期 T 必须是非零常数, 否则它描述的不是一种重复模式. 其二. 定义域 D 必须在沿 T 的正、反两个方向平移下封闭, 这保证了周期关系可以反复迭代. 因此, 周期函数的定义域在周期方向上必然是无界的, 例如 $\mathbb{R}$ 或若干等距离散点的集合. 其三. 核心等式 $f(x+T)=f(x)$ 必须对定义域内的**所有** x 成立.

由定义可直接推得, 若 T 是一个周期, 则对任意非零整数 n, nT 也是一个周期. 若 T 是周期, 则 $f(x)=f(x-T)$, 故 $-T$ 也是周期.

定义 3.5 (最小正周期) 若一个周期函数的所有正周期中存在一个最小值, 则称此数为该函数的**最小正周期**. 通常语境下所称的周期即指最小正周期.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

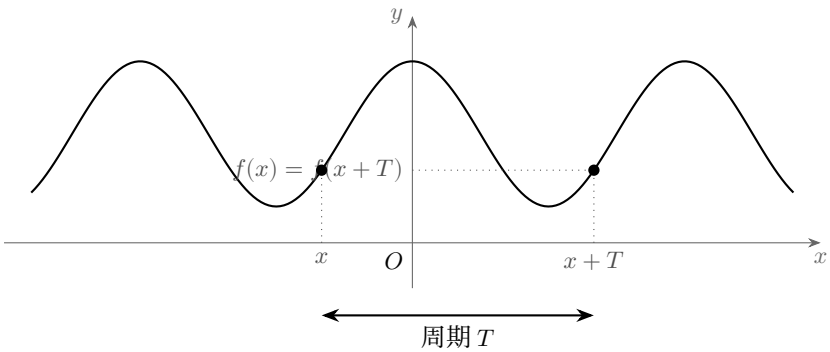


图 3.4 周期函数的几何意义. 图像的平移不变性

有些周期函数没有最小正周期. 例如, 常数函数 $f(x)=c$, 任何非零实数 T 都是它的周期, 故不存在最小的正周期. 一个更典型的例子是狄利克雷函数.

$$D(x)=\begin{cases}1, & x\in\mathbb{Q}\\ 0, & x\notin\mathbb{Q}\end{cases}$$

对于任意非零有理数 T, 当 x 是有理数时, $x+T$ 也是有理数, 故 $D(x+T)=1=D(x)$; 当 x 是无理数时, $x+T$ 也是无理数, 故 $D(x+T)=0=D(x)$. 因此, 任何非零有理数都是它的周期, 而正有理数集合没有最小值, 故狄利克雷函数不存在最小正周期.

3.5.2 周期性的四则运算

两个独立的周期性现象叠加时, 合成的现象是否仍具有周期性, 取决于两者周期的**公度性**. 例如, 将两个不同频率的纯音(对应两个正弦函数)混合, 混合音仍是周期函数当且仅当两个周期可公度.

定理 3.5 设 $f(x)$ 和 $g(x)$ 是定义在同一集合 D 上的两个周期函数, 周期分别为 T_1 和 T_2.

1. 若 T_1 与 T_2 **可公度**, 即其比值为有理数, 则存在非零整数 m,n 使 $mT_1=nT_2$. 此公共非零数 $T=mT_1=nT_2$ 是 $f(x)+g(x)$、$f(x)-g(x)$、$f(x)g(x)$ 的一个周期.
2. 若 T_1 与 T_2 **不可公度**, 则它们的和、差、积不一定是周期函数.

证明 若 T_1/T_2 为有理数, 可设 $T_1/T_2=n/m$, 其中 m,n 为非零整数. 于是 $mT_1=nT_2$, 记这个公共非零数为 T. 由于周期可作整数倍, T 同时是 f 和 g 的周期.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

对任意 $x \in D$, 有

$$f(x+T) = f(x), \qquad g(x+T) = g(x).$$

因此

$$(f+g)(x+T) = f(x+T) + g(x+T) = f(x) + g(x) = (f+g)(x),$$

差函数与积函数同理.

$$(f-g)(x+T) = f(x) - g(x), \qquad (fg)(x+T) = f(x)g(x).$$

知乎@余命数

故 T 是三者的一个周期.

若 T_1 与 T_2 不可公度, 结论只是不保证仍有周期. 例如定义在 $\mathbb{R}$ 上的函数

$$f(x) = \begin{cases} 1, & x \in \mathbb{Z}, \\ 0, & x \notin \mathbb{Z}, \end{cases} \qquad g(x) = \begin{cases} 2, & x \in \sqrt{2}\mathbb{Z}, \\ 0, & x \notin \sqrt{2}\mathbb{Z}. \end{cases}$$

其中 f 有周期 1, g 有周期 $\sqrt{2}$, 且 $1/\sqrt{2}$ 为无理数. 注意 $\mathbb{Z} \cap \sqrt{2}\mathbb{Z} = \{0\}$, 因为若 $k = \sqrt{2}m$ 且 $k, m \in \mathbb{Z}$, 当 $m \neq 0$ 时会推出 $\sqrt{2} = k/m$ 为有理数.

对 $f+g$ 而言, 函数值 3 只在 $x = 0$ 处取得; 对 $f-g$ 而言, 函数值 -1 只在 $x = 0$ 处取得; 对 fg 而言, 函数值 2 也只在 $x = 0$ 处取得. 若三者中任意一个存在非零周期 T, 将周期等式用于 $x = 0$ 都会迫使该特殊函数值也在 $x = T$ 处取得, 从而 $T = 0$, 矛盾. 因此, 在周期不可公度时, 和、差、积确实不一定仍为周期函数. □

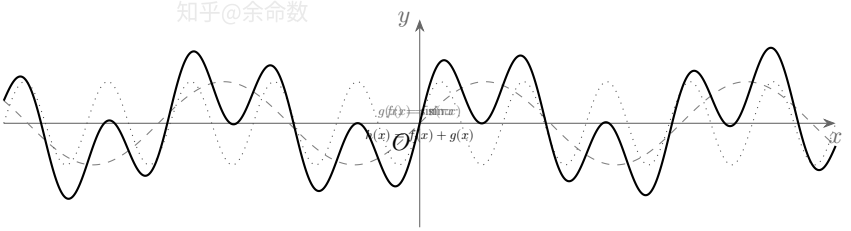


图 3.5　具有不可公度周期的两个函数(虚线与点线)之和, 以粗实线表示, 常呈现复杂的准周期形态.

例题 3.19　求函数 $f(x) = |\sin x| + \cos(2x)$ 的最小正周期.

解　函数 $f(x)$ 为两个周期函数之和: $g(x) = |\sin x|$ 与 $h(x) = \cos(2x)$.

函数 $\sin x$ 的周期为 2π, 经绝对值变换后 x 轴下方的图像翻折到上方, $[\pi, 2\pi]$ 区间的负半周与 $[0, \pi]$ 区间的正半周完全相同, 故 $g(x) = |\sin x|$ 的最小正周期为 $T_1 = \pi$. 又 $h(x) = \cos(2x)$ 的最小正周期为 $T_2 = 2\pi/|2| = \pi$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两周期之比 $T_1/T_2 = 1$ 为有理数, 故周期可公度. 因此 $f(x)$ 是周期函数, $T = \operatorname{lcm}(T_1, T_2) = \operatorname{lcm}(\pi, \pi) = \pi$ 是它的一个周期; 下面的代数变形同时说明它确实是最小正周期.

以代数变形验证此结论. 由二倍角公式,

$$\cos(2x) = 1 - 2\sin^2 x = 1 - 2|\sin x|^2$$

知乎@余命数

故 $f(x) = |\sin x| + 1 - 2|\sin x|^2$, 是关于 $|\sin x|$ 的复合函数, π 为 $f(x)$ 的周期. 尚需排除更小的正周期. 设 T 为 $f(x)$ 的周期, 令 $P(u) = u + 1 - 2u^2$, 由 $x = 0$ 得

知乎@余命数

知乎@余命数

$$P(|\sin T|) = P(0) = 1,$$

从而 $|\sin T| = 0$ 或 $|\sin T| = \frac{1}{2}$. 若 $|\sin T| = \frac{1}{2}$, 则 $|\cos T| = \frac{\sqrt{3}}{2}$, 代入 $x = \frac{\pi}{2}$ 得 $P(|\cos T|) \neq P(1)$, 矛盾. 故 $|\sin T| = 0$, 即 T 必为 π 的正整数倍, 因此最小正周期为 π.

例题 3.20 求函数 $h(x) = \sin(3x) + \cos(\frac{5}{2}x)$ 的最小正周期.

解 函数 $h(x)$ 为两个周期函数 $f(x) = \sin(3x)$ 与 $g(x) = \cos(\frac{5}{2}x)$ 的和, 分别确定这两个基函数的最小正周期, 检验其公度性, 再求最小公倍数.

$f(x) = \sin(3x)$ 的角频率为 3, 最小正周期为

知乎@余命数

$$T_1 = \frac{2\pi}{|3|} = \frac{2\pi}{3}$$

知乎@余命数

$g(x) = \cos(\frac{5}{2}x)$ 的角频率为 $\frac{5}{2}$, 最小正周期为

$$T_2 = \frac{2\pi}{|5/2|} = \frac{4\pi}{5}$$

两周期的公度性由其比值判定:

$$\frac{T_1}{T_2} = \frac{2\pi/3}{4\pi/5} = \frac{2\pi}{3} \cdot \frac{5}{4\pi} = \frac{5}{6}$$

由于比值 $5/6$ 是一个有理数, 两个周期是可公度的. 因此, 函数 $h(x)$ 必然是周期函数, 且 T_1 与 T_2 的最小公倍数是它的一个周期; 最小正周期可能比它更小, 需另行验证.

知乎@余命数

对于两个分数形式的周期 $\frac{p_1}{q_1}\pi$ 和 $\frac{p_2}{q_2}\pi$, 它们的最小公倍数可以通过以下方式计算, 得

$$T = \operatorname{lcm}\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right)\pi = \frac{\operatorname{lcm}(p_1, p_2)}{\gcd(q_1, q_2)}\pi$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在本例中,

$$T = \operatorname{lcm}\left(\frac{2\pi}{3}, \frac{4\pi}{5}\right) = \pi \cdot \operatorname{lcm}\left(\frac{2}{3}, \frac{4}{5}\right)$$

依上述法则,

知乎@余命数

$$T = \pi \cdot \frac{\operatorname{lcm}(2,4)}{\gcd(3,5)} = \pi \cdot \frac{4}{1} = 4\pi$$

因此, 函数 $h(x)$ 的最小正周期为 4π.

例题 3.21　考察函数 $h(x) = \{2x\} - \sin(\pi x)$ 的周期性, 其中 $\{u\} = u - \lfloor u \rfloor$ 表示小数部分函数.

解　此函数由两个周期函数 $f(x) = \{2x\}$ 和 $g(x) = \sin(\pi x)$ 的差构成.

基本的小数部分函数 $\{u\}$ 的最小正周期为 1, 故复合函数 $\{kx\}$ 的周期为 $1/|k|$, $f(x) = \{2x\}$ 的最小正周期为 $T_1 = 1/2$. 验证: $f(x+1/2) = \{2(x+1/2)\} = \{2x+1\}$. 加上整数不改变小数部分, 故 $\{2x+1\} = \{2x\} = f(x)$.

又 $g(x) = \sin(\pi x)$ 的角频率为 π, 最小正周期为 $T_2 = 2\pi/|\pi| = 2$.

两周期之比为

$$\frac{T_1}{T_2} = \frac{1/2}{2} = \frac{1}{4}$$

此为有理数, 故 $h(x)$ 是周期函数, T_1 与 T_2 的最小公倍数是它的一个周期; 对本例可直接验证它也是最小正周期, 计算如下.

知乎@余命数

$$T = \operatorname{lcm}\left(\frac{1}{2}, 2\right) = \operatorname{lcm}\left(\frac{1}{2}, \frac{2}{1}\right) = \frac{\operatorname{lcm}(1,2)}{\gcd(2,1)} = \frac{2}{1} = 2$$

因此, 函数 $h(x)$ 的最小正周期为 2.

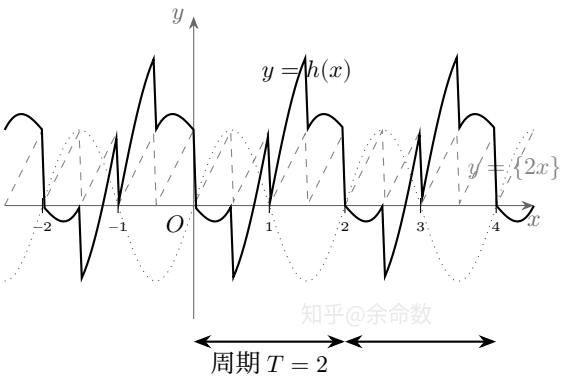


图 3.6　函数 $h(x) = \{2x\} - \sin(\pi x)$ 的图像. 尽管其构成部分周期不同 ($T_1 = 0.5, T_2 = 2$), 合成后的函数呈现出清晰的周期性, 最小正周期为 $T = 2$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.5.3 对称性与周期性的关系

函数的周期性常常由其对称性导出. 两种不同位置的对称性 (轴对称或中心对称)相复合, 可以生成平移对称, 即周期性.

定理 3.6 (对称性向周期性的转化) 设函数 $f(x)$ 的定义域为 $\mathbb{R}$.

1. **双轴对称**:若 $f(x)$ 的图像关于两条不同的直线 $x=a$ 和 $x=b$ 对称, 则 $f(x)$ 是周期函数, 且 $T=2|a-b|$ 是其一个周期.
2. **双中心对称**: 若 $f(x)$ 的图像关于两个不同的点 (a,c) 和 (b,c) 中心对称, 则 $f(x)$ 是周期函数, 且 $T=2|a-b|$ 是其一个周期.
3. **轴对称与中心对称**: 若 $f(x)$ 的图像关于直线 $x=a$ 对称, 且关于点 (b,c) 中心对称, 则当 $a\neq b$ 时, $f(x)$ 是周期函数, 且 $T=4|a-b|$ 是其一个周期; 当 $a=b$ 时, $f(x)$ 为常函数 c, 因而也是周期函数.

证明 三条结论的证明思路一致, 都是把几何对称操作写成代数恒等式, 再对恒等式复合、迭代, 最终得到 $f(x+T)=f(x)$.

函数 $f(x)$ 的图像关于两条不同的直线 $x=a$ 和 $x=b$ 对称, 故以下两个恒等式对定义域内所有 x 成立:

$$f(x)=f(2a-x) \tag{3-7}$$

$$f(x)=f(2b-x) \tag{3-8}$$

平移不变性须由这两个对称变换复合产生. 从 $f(x)$ 出发, 先应用关于 $x=a$ 的对称性 (3-7), 得 $f(x)=f(2a-x)$; 再对右端 $f(2a-x)$ 应用关于 $x=b$ 的对称性 (3-8), 即把 (3-8) 中的自变量 x 换为 $2a-x$:

$$f(2a-x)=f(2b-(2a-x))=f(x+2b-2a)$$

两步衔接即得

$$f(x)=f(x+2(b-a))$$

这正是周期函数的定义. 令 $T=2(b-a)$, 则 $f(x)=f(x+T)$. 因此, $f(x)$ 是周期函数, 且 $2|a-b|$ 是其一个周期.

更一般地, 若 $f(x)$ 的图像关于点 (a,c) 和 (b,d) 中心对称, 其代数表述为:

$$f(x)+f(2a-x)=2c \tag{3-9}$$

$$f(x)+f(2b-x)=2d \tag{3-10}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

由此解出 $f(2a-x)$ 与 $f(2b-x)$:

$$f(2a-x)=2c-f(x)$$

$$f(2b-x)=2d-f(x)$$

仍以复合变换建立远距离自变量间的联系. 自变量 $x+2(a-b)$ 可写成 $2a-(2b-x)$. 先把 (3-9) 中的自变量 x 换为 $2b-x$:

$$f(2b-x)+f(2a-(2b-x))=2c$$

整理得 $f(x+2a-2b)=2c-f(2b-x)$. 再以 (3-10) 替换右端的 $f(2b-x)$:

$$f(x+2(a-b))=2c-(2d-f(x))=f(x)+2(c-d)$$

此式说明 $f(x)$ 在平移 $2(a-b)$ 后函数值产生固定增量 $2(c-d)$, 这种结构称为**准周期性**. 为使 $f(x)$ 成为严格周期函数, 该增量须为零, 即 $2(c-d)=0$, 亦即 $c=d$. 故若 $f(x)$ 关于 (a,c) 与 (b,c)(纵坐标相同)对称, 则 $f(x+2(a-b))=f(x)$, $T=2|a-b|$ 是其一个周期.

若 $f(x)$ 关于直线 $x=a$ 对称且关于点 (b,c) 中心对称, 其代数恒等式为:

$$f(x)=f(2a-x) \tag{3-11}$$

$$f(x)+f(2b-x)=2c \tag{3-12}$$

此情形先导出反周期关系, 再迭代得到周期关系. 从 $f(x)$ 出发, 先用轴对称 (3-11), 再用中心对称 (3-12):

$$f(x)\overset{(3\text{-}11)}{=}f(2a-x)$$

由 (3-12) 解出 $f(y)=2c-f(2b-y)$, 应用于 $y=2a-x$:

$$f(2a-x)=2c-f(2b-(2a-x))=2c-f(x+2b-2a)$$

两步结合得 $f(x)=2c-f(x+2(b-a))$, 整理为反周期关系:

$$f(x+2(b-a))=2c-f(x)$$

自变量增加 $2(b-a)$ 后, 新函数值与原值关于水平线 $y=c$ 对称. 为消除此对称性并恢复恒等关系, 对该式迭代, 即把 x 换为 $x+2(b-a)$:

$$f([x+2(b-a)]+2(b-a))=2c-f(x+2(b-a))$$

$$f(x+4(b-a))=2c-(2c-f(x))$$

$$f(x+4(b-a))=f(x)$$

知乎@余命数 知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

若 $a \neq b$, 此式正是周期函数的定义, 故 $T = 4|a-b|$ 是函数的一个周期. 若 $a = b$, 则轴对称给出 $f(a+t) = f(a-t)$, 中心对称给出 $f(a+t) + f(a-t) = 2c$, 两式联立得 $f(a+t) = c$. 由于 t 任意, $f(x) \equiv c$, 因而任意非零常数都是它的周期. □

定理 3.7 (由函数方程确定周期性) 对于定义在 $\mathbb{R}$ 上的函数 $f(x)$, 若存在非零常数 a, 满足以下关系之一, 则 $f(x)$ 为周期函数.

1. 若 $f(x+a) = -f(x)$, 则 $T = 2a$ 是一个周期.
2. 若 $f(x+a) = 1/f(x)$ 且 $f(x)$ 的值域不包含零, 则 $T = 2a$ 是一个周期.
3. 若 $f(x+a) = \frac{1+f(x)}{1-f(x)}$ 且 $f(x)$ 的值域不包含 0 和 1, 则 $T = 4a$ 是一个周期.

证明 证明均通过对给定关系式进行迭代完成, 这是一种探求周期性的基本技巧.

1. $f(x+2a) = f((x+a)+a) = -f(x+a) = -(-f(x)) = f(x)$.
2. $f(x+2a) = f((x+a)+a) = 1/f(x+a) = 1/(1/f(x)) = f(x)$.
3. 这是一个更具挑战性的迭代. 值域不含 0 和 1 保证以下分式在迭代中均有意义.

$$f(x+2a) = f((x+a)+a) = \frac{1+f(x+a)}{1-f(x+a)} = \frac{1+\frac{1+f(x)}{1-f(x)}}{1-\frac{1+f(x)}{1-f(x)}}$$
$$= \frac{(1-f(x))+(1+f(x))}{(1-f(x))-(1+f(x))} = \frac{2}{-2f(x)} = -\frac{1}{f(x)}$$

此即 $f(x+2a) = -1/f(x)$. 再迭代一次, $f(x+4a) = f((x+2a)+2a) = -1/f(x+2a) = -1/(-1/f(x)) = f(x)$. 故 $T = 4a$ 是一个周期. □

例题 3.22 已知函数 $y = f(x)\,(x \in \mathbb{R})$ 不是常函数, 且满足 $f(a+b)+f(a-b) = 2f(a)f(b)$ 及 $f(1) = 1/2$. 求 $f(2024)$.

解 通过对变量 a, b 赋特殊值, 可以分析 $f(x)$ 的对称性与周期性.

首先, 在 $f(a+b)+f(a-b) = 2f(a)f(b)$ 中, 令 $a = x, b = 0$, 得 $f(x)+f(x) = 2f(x)f(0)$, 即 $2f(x) = 2f(x)f(0)$. 此式可化为 $f(x)(1-f(0)) = 0$. 由于 $f(x)$ 不是常函数 (特别地, 不是恒为零的函数), 必存在 x_0 使得 $f(x_0) \neq 0$. 因此, 必然有 $1-f(0) = 0$, 即 $f(0) = 1$.

接着, 分析函数的奇偶性. 令 $a = 0, b = x$, 得 $f(x)+f(-x) = 2f(0)f(x)$. 代入 $f(0) = 1$, 得 $f(x)+f(-x) = 2f(x)$, 这立即导出 $f(-x) = f(x)$. 故 $f(x)$ 是一个偶函数.

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

然后, 利用已知条件 $f(1)=1/2$. 令 $a=1, b=x$,

$$f(1+x)+f(1-x)=2f(1)f(x)=2\cdot\frac{1}{2}\cdot f(x)=f(x)$$

由于 $f(x)$ 是偶函数, $f(1-x)=f(-(x-1))=f(x-1)$. 于是上式变为 $f(x+1)+f(x-1)=f(x)$.

将 x 换为 $x+1$, 得 $f(x+2)+f(x)=f(x+1)$. 再将 $f(x+1)=f(x)-f(x-1)$ 代入, 得 $f(x+2)+f(x)=f(x)-f(x-1)$, 故 $f(x+2)=-f(x-1)$. 将此式中 x 再换为 $x+1$, 得 $f(x+3)=-f(x)$.

这是反周期关系 $f(x+3)=-f(x)$, 迭代得

$$f(x+6)=f((x+3)+3)=-f(x+3)=-(-f(x))=f(x)$$

因此, $f(x)$ 是周期函数, 一个周期为 $T=6$.

据此求 $f(2024)$. 由周期性, 取 2024 除以 6 的余数:

$$2024=6\times 337+2$$

故 $f(2024)=f(2)$. 为求 $f(2)$, 回到反周期关系 $f(x+3)=-f(x)$, 令 $x=-1$, 得 $f(2)=-f(-1)$. 又 $f(x)$ 是偶函数, $f(-1)=f(1)=1/2$, 故 $f(2)=-f(1)=-1/2$.

因此 $f(2024)=f(2)=-1/2$.

3.6 凹凸性

> **注 3.4**
> 本节后半使用一阶/二阶导数与中值定理判定凹凸, 并推出詹森不等式. 定义(弦在弧的上方/下方)与中点形式见下文; 导数判定与严格证明见 第16章.

函数的单调性描述了其图像的升降趋势, 这是一个关于**变化率符号**的性质. 然而, 函数图像的形态远比单纯的升降要丰富. 考虑函数 $f(x)=x^2$ 与 $g(x)=\sqrt{x}$ (在 $x>0$ 时), 它们都是单调递增的, 但它们的增长姿态截然不同. $f(x)=x^2$ 的增长越来越快, 其图像向上弯曲; 而 $g(x)=\sqrt{x}$ 的增长越来越慢, 其图像向下弯曲.

凹凸性正是为了精确刻画函数图像的这种**弯曲方向**而引入的概念. 它所关注的, 是**变化率自身的变化趋势**.

3.6.1 凹凸性的几何直观与定义

要捕捉曲线的弯曲特性, 可从两个等价的几何观点出发.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

3.6.1.0.1 割线斜率的演变 一个更具动态的观点是考察连接曲线上两点的**割线**斜率如何变化. 设想我们在一个向上弯曲的曲线上固定一点 $P_1(x_1, f(x_1))$, 然后让另一点 $P_2(x_2, f(x_2))$ 沿着曲线从左向右移动. 连接 P_1 和 P_2 的割线, 其斜率是单调递增的.

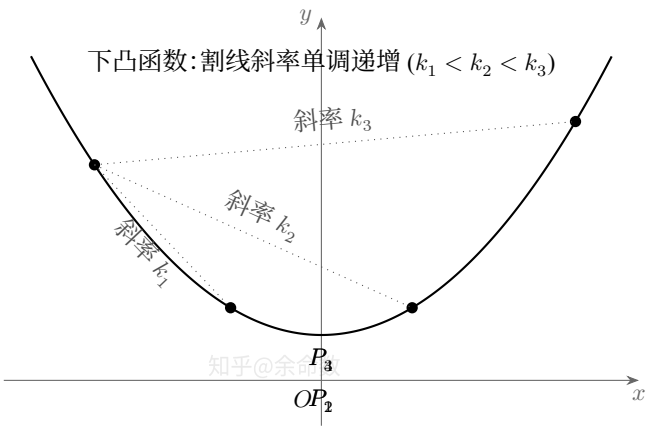


图 3.7 下凸函数割线斜率的单调性

3.6.1.0.2 函数图像与弦的相对位置 另一个等价的、也是更为经典的几何刻画, 是比较函数图像本身与连接其上任意两点的线段的相对位置, 前者称为弧, 后者称为**弦**.

- 对于一个向上弯曲的函数, 其图像总是位于连接其上任意两点的弦的**下方**(或重合).
- 对于一个向下弯曲的函数, 其图像总是位于连接其上任意两点的弦的**上方**或与之重合.

为了将这个几何观察转化为精确的代数语言, 需一种方式来表示弦上的点. 设弦的两个端点为 $A(x_1, f(x_1))$ 和 $B(x_2, f(x_2))$. 对于任意 $t \in [0,1]$, 表达式 $x_t = (1-t)x_1 + tx_2$ 给出了线段 $[x_1, x_2]$ 上的一个点. 对应地, 弦上具有相同横坐标的点的纵坐标, 是端点纵坐标的同权重线性组合, 即 $y_t = (1-t)f(x_1) + tf(x_2)$. 而函数图像上对应点的纵坐标是 $f(x_t) = f((1-t)x_1 + tx_2)$. 比较这两个纵坐标的大小, 便得到了凹凸性的严格定义.

知乎@余命数 知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

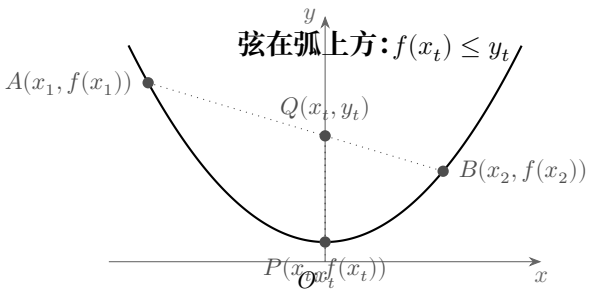


图 3.8　下凸函数的代数定义图解

定义 3.6 (凹凸函数) 设函数 $f(x)$ 定义在区间 I 上.

- 若对于任意 $x_1, x_2 \in I$ 和任意 $t \in [0,1]$, 恒有

$$f((1-t)x_1 + tx_2) \le (1-t)f(x_1) + tf(x_2)$$

则称 $f(x)$ 为 I 上的**下凸函数**.

- 若对于任意 $x_1, x_2 \in I$ 和任意 $t \in [0,1]$, 恒有

$$f((1-t)x_1 + tx_2) \ge (1-t)f(x_1) + tf(x_2)$$

则称 $f(x)$ 为 I 上的**上凸函数**.

若当 $x_1 \ne x_2$ 且 $t \in (0,1)$ 时上述不等式均取严格不等号, 则称函数为**严格**下凸或严格上凸.

这个不等式是凸性理论的基础. 特别地, 当取 $t = 1/2$ 时, 它退化为一个极其常用且直观的形式.

- **下凸函数**满足 $f\left(\dfrac{x_1+x_2}{2}\right) \le \dfrac{f(x_1)+f(x_2)}{2}$ (中点的函数值不大于函数值的中点).
- **上凸函数**满足 $f\left(\dfrac{x_1+x_2}{2}\right) \ge \dfrac{f(x_1)+f(x_2)}{2}$ (中点的函数值不小于函数值的中点).

凹凸性的导数判定

3.6.1.1 凹凸性的分析判定法

凹凸性的代数定义严谨, 但在实际操作中直接应用不等式进行判定往往十分繁琐. 需一个更具操作性的分析工具. 这一工具的建立, 源于凹凸性的几何直观. **函数图像的弯曲方向, 实质上由其切线斜率的变化趋势所决定**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们考察一个下凸函数的图像. 当我们从左至右观察其切线时, 会发现切线的倾斜程度在持续增大. 即使在函数递减的部分, 切线的斜率也是从一个较大的负数向一个较小的负数变化, 始终处在增加的过程中. 这一观察是建立分析判定法的关键.

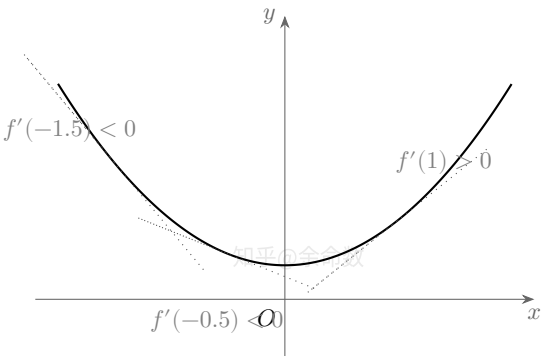


图 3.9 下凸函数: 切线斜率 $f'(x)$ 单调递增

实际上, 函数的凹凸性与其一阶导数的单调性是等价的.

定理 3.8 设函数 $f(x)$ 在区间 I 上连续, 并在 I 的内部可导.

- $f(x)$ 在 I 上是下凸函数的充分必要条件是其导函数 $f'(x)$ 在 I 的内部单调递增.
- $f(x)$ 在 I 上是上凸函数的充分必要条件是其导函数 $f'(x)$ 在 I 的内部单调递减.

证明 仅证明下凸的情形; 上凸情形可对 $-f$ 应用同样结论得到.

先设 f 为下凸函数. 由下凸定义可推出割线斜率的单调性, 即任意满足 $x_1 < x_2 < x_3$ 的三点, 有

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} \le \frac{f(x_3)-f(x_1)}{x_3-x_1} \le \frac{f(x_3)-f(x_2)}{x_3-x_2}$$

将 x_2 写成 x_1, x_3 的加权平均并代入下凸定义, 可得

$$f(x_2) \le \frac{x_3-x_2}{x_3-x_1} f(x_1) + \frac{x_2-x_1}{x_3-x_1} f(x_3),$$

整理即可得到上述两边不等式. 对任意 $a, b \in I$ 的内部且 $a < b$, 在上述不等式中令 $x_1 = a, x_3 = b$, 再分别令 $x_2 \to a^+$ 与 $x_2 \to b^-$, 得

$$f'(a) \le \frac{f(b)-f(a)}{b-a} \le f'(b).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此 $f'(a) \le f'(b)$, 即 f' 在 I 的内部单调递增.

反过来, 设 f' 在 I 的内部单调递增. 任取 $x_1 < x_2$ 与 $t \in (0,1)$, 令 $x_t = (1-t)x_1 + tx_2$. 由拉格朗日中值定理, 存在 $\xi_1 \in (x_1, x_t)$ 与 $\xi_2 \in (x_t, x_2)$, 使得

$$\frac{f(x_t)-f(x_1)}{x_t-x_1} = f'(\xi_1), \qquad \frac{f(x_2)-f(x_t)}{x_2-x_t} = f'(\xi_2).$$

因为 $\xi_1 < \xi_2$ 且 f' 单调递增, 所以前一割线斜率不大于后一割线斜率. 将 $x_t - x_1 = t(x_2-x_1)$ 与 $x_2 - x_t = (1-t)(x_2-x_1)$ 代入并整理, 得

$$f(x_t) \le (1-t)f(x_1) + tf(x_2).$$

知乎@余命数

当 $t=0$ 或 $t=1$ 时等式显然成立, 故 f 为下凸函数. □

这个定理将凹凸性的判定问题, 转化为了已熟悉的**单调性判定问题**, 只不过判定的对象是导函数 $f'(x)$. 已知, 一个函数(此处为 $f'(x)$)的单调性由其自身的导数, 即 $f''(x)$ 的符号所决定. 这便引出了最终的、也是最常用的分析判定法.

定理 3.9 (二阶导数判定法) 设函数 $f(x)$ 在区间 I 上连续, 并在 I 的内部二阶可导.

- 若在 I 的内部恒有 $f''(x) \ge 0$, 则 $f(x)$ 在 I 上是下凸的.
- 若在 I 的内部恒有 $f''(x) \le 0$, 则 $f(x)$ 在 I 上是上凸的.

证明 若 $f''(x) \ge 0$, 则对任意 $a, b \in I$ 的内部且 $a < b$, 将拉格朗日中值定理用于函数 f' 可得存在 $\xi \in (a,b)$, 使得

$$f'(b) - f'(a) = f''(\xi)(b-a) \ge 0.$$

因而 f' 在 I 的内部单调递增, 由上一条定理得 f 在 I 上下凸.

若 $f''(x) \le 0$, 同理可得 $f'(b) - f'(a) \le 0$, 即 f' 在 I 的内部单调递减, 从而 f 在 I 上上凸. □

此定理将判断凹凸性的问题最终归结为求解一个不等式, 即判断二阶导数 $f''(x)$ 的符号.

3.6.1.2 拐点

函数从一种弯曲形态过渡到另一种的临界点, 称为拐点.

定义 3.7 (拐点) 若函数 $y = f(x)$ 的图像在点 $P(x_0, f(x_0))$ 处**凹凸性发生改变**, 则称点 P 为该函数图像的一个**拐点**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据二阶导数判定法, 拐点的出现意味着二阶导数 $f''(x)$ 在该点两侧的符号发生了改变. 因此, 若函数在拐点 x_0 处二阶可导, 则必然有 $f''(x_0)=0$.$f''(x_0)=0$ 只是拐点的必要条件.(思考: 考虑函数 $f(x)=x^4$, 其在 $x=0$ 处二阶导数为零, 但该点处的凹凸性保持不变.)

例题 3.23 确定函数 $f(x)=x^3-6x^2+5x+1$ 的凹凸区间及拐点.

解 先求导, 一阶与二阶导数为

$$f'(x)=3x^2-12x+5,$$

$$f''(x)=6x-12.$$

令 $f''(x)=0$ 解得 $x=2$, 此为唯一的拐点候选. 以该点为界考察 $f''(x)$ 的符号:

- 当 $x\in(-\infty,2)$ 时, $f''(x)<0$. 根据定理, 函数 $f(x)$ 在此区间上是**上凸**的.
- 当 $x\in(2,+\infty)$ 时, $f''(x)>0$. 根据定理, 函数 $f(x)$ 在此区间上是**下凸**的.

函数在 $x=2$ 两侧的凹凸性发生改变, 因此点 $(2,f(2))$ 是拐点, 其中 $f(2)=2^3-6(2^2)+5(2)+1=8-24+10+1=-5$.

故函数 $f(x)$ 的上凸区间为 $(-\infty,2)$, 下凸区间为 $(2,+\infty)$, 拐点为 $(2,-5)$.

3.6.2 詹森不等式

凹凸性的定义本身蕴含了一个不等式结构. 当将定义中连接两点的弦与弧的关系, 从两个点自然地推广到任意多个点时, 便得到了数学分析中应用广泛的**詹森不等式**.

3.6.2.1 詹森不等式的一般形式

詹森不等式指出: **下凸函数在加权平均点处的函数值, 不大于函数值按同样权重的加权平均**.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

定理 3.10 (詹森不等式) 设 $f(x)$ 是定义在区间 I 上的函数.

1. 若 $f(x)$ 为**下凸函数**, 则对于任意正整数 n、任意 $x_1,\dots,x_n \in I$ 以及任意一组满足 $\sum_{i=1}^{n}\lambda_i = 1$ 的非负权重 $\lambda_i \geq 0$, 恒有:

$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right) \leq \sum_{i=1}^{n}\lambda_i f(x_i) \tag{3-13}$$

2. 若 $f(x)$ 为**上凸函数**, 则上述不等号方向相反.

对于严格凹凸函数, 等号成立的充分必要条件是所有正权重对应的 x_i 相等.

证明 只证明下凸情形; 上凸情形可对 $-f$ 应用下凸结论得到.

对正权重点的个数作归纳. 若只有一个正权重, 则该权重必为 1, 不等式为恒等式. 设结论对正权重点数不超过 $n-1$ 的情形成立, 现考虑有 n 个正权重的情形. 不妨这些正权重为 $\lambda_1,\dots,\lambda_n$, 且 $0<\lambda_n<1$. 令

$$\mu_i = \frac{\lambda_i}{1-\lambda_n}\quad (i=1,2,\dots,n-1), \qquad y = \sum_{i=1}^{n-1}\mu_i x_i.$$

则 $\mu_i \geq 0$ 且 $\sum_{i=1}^{n-1}\mu_i = 1$, 并且

$$\sum_{i=1}^{n}\lambda_i x_i = (1-\lambda_n)y + \lambda_n x_n.$$

由下凸定义与归纳假设,

$$\begin{aligned} f\left(\sum_{i=1}^{n}\lambda_i x_i\right) &= f((1-\lambda_n)y + \lambda_n x_n) \\ &\leq (1-\lambda_n)f(y) + \lambda_n f(x_n) \\ &\leq (1-\lambda_n)\sum_{i=1}^{n-1}\mu_i f(x_i) + \lambda_n f(x_n) \\ &= \sum_{i=1}^{n}\lambda_i f(x_i). \end{aligned}$$

因而詹森不等式成立.

若 f 严格下凸, 且至少有两个正权重对应的点不同, 则在上述归纳过程中会出现一次对两个不同点取严格加权平均的步骤, 相应不等号为严格不等号, 故整体为严格不等式. 反之, 若所有正权重对应的 x_i 相等, 两边显然相等. 严格上凸情形同理. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此定理的加权形式是其最完整的表述. 当取所有权重相等, 即 $\lambda_i = 1/n$ 时, 便得到了其更为人熟知的算术平均形式.

推论 3.2 (詹森不等式的算术平均形式) 设 $f(x)$ 是定义在区间 I 上的函数.

1. 若 $f(x)$ 为**下凸函数**, 则对于任意正整数 n 和任意 $x_1, \dots, x_n \in I$, 有:

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n} \tag{3-14}$$

   此即平均的函数值不大于函数值的平均.

2. 若 $f(x)$ 为**上凸函数**, 则不等号方向相反.

证明 (算术平均形式的证明) 在詹森不等式的一般形式中取等权重 $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1/n$, 即得

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leq \frac{f(x_1) + \cdots + f(x_n)}{n}$$

对下凸函数成立. 若 f 为上凸函数, 则应用一般形式中的上凸情形, 不等号方向相反. □

用二阶导数验证凸性并推均值

3.6.3 从琴生不等式到均值不等式

许多著名不等式是詹森不等式的特例.

例题 3.24 证明算术-几何平均值不等式: 对于正数 $x_1, \dots, x_n$, $\frac{x_1+\cdots+x_n}{n} \geq \sqrt[n]{x_1 \dots x_n}$.

解 考虑函数 $f(x) = \ln x$, 定义域为 $(0, +\infty)$. 其二阶导数为 $f''(x) = -1/x^2$. 在 $(0, +\infty)$ 内, $f''(x) < 0$ 恒成立, 故 $f(x) = \ln x$ 是严格上凸函数.

由上凸函数的詹森不等式,

$$\frac{\ln x_1 + \cdots + \ln x_n}{n} \leq \ln\left(\frac{x_1 + \cdots + x_n}{n}\right)$$

利用对数性质, 左侧化为

$$\frac{\ln(x_1 \dots x_n)}{n} = \ln\left((x_1 \dots x_n)^{1/n}\right) = \ln(\sqrt[n]{x_1 \dots x_n})$$

代入原不等式, 得 $\ln(\sqrt[n]{x_1 \dots x_n}) \leq \ln(\frac{x_1+\cdots+x_n}{n})$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由于对数函数是严格单调递增的, 上述不等关系等价于其自变量之间的不等关系:

$$\sqrt[n]{x_1 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}$$

证毕.

注 3.5 (詹森不等式的应用信号)
在处理多元函数极值问题时, 若目标表达式形如 $\sum f(x_i)$, 例如 $2^{x_1} + 2^{x_2}$ 或 $\sin A + \sin B + \sin C$, 且变量满足和为定值的约束 $\sum x_i = C$, 这便是应用詹森不等式的典型信号.

例题 3.25　在 $\triangle ABC$ 中, 求 $S = \sin A + \sin B + \sin C$ 的最大值.

解 此问题符合上述应用模式. 考虑函数 $f(x) = \sin x$, 定义域为内角范围 $(0, \pi)$, 变量满足 $A + B + C = \pi$.

函数 $f(x) = \sin x$ 的二阶导数为 $f''(x) = -\sin x$. 在区间 $(0, \pi)$ 内, $\sin x > 0$, 故 $f''(x) < 0$. 因此 $f(x)$ 是严格上凸函数.

应用上凸函数的詹森不等式,

$$\frac{\sin A + \sin B + \sin C}{3} \leq \sin\left(\frac{A + B + C}{3}\right)$$

代入约束 $A + B + C = \pi$, 得

$$\frac{S}{3} \leq \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}$$

故 $S \leq \frac{3\sqrt{3}}{2}$.

等号成立条件为 $A = B = C = \pi/3$, 此条件可以达到(当 $\triangle ABC$ 为正三角形时). 因此, S 的最大值为 $\frac{3\sqrt{3}}{2}$.

3.7 反函数

函数把定义域中的每个 x 唯一地对应到值域中的一个 y. 我们关心的问题是, 能否从 y 唯一地找回 x. 如果可以, 这个把 y 映回 x 的对应法则就称为函数 f 的**反函数**. 反函数的图像把横、纵坐标互换, 因此它与原函数的图像关于直线 $y = x$ 对称.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

3.7.1 反函数存在的条件

设给定函数 f. 一个从值域回到定义域的对应法则要成为函数, 必须让值域中的每个 y 都唯一对应定义域中的一个 x. 这正好要求 f 把不同的输入映到不同的输出, 即没有两个不同的 x 对应同一个 y.

定义 3.8 (单射) 称一个函数 $f : A \to C$ 是**单射**, 如果对于其定义域 A 内任意两个不同的元素 x_1, x_2, 其对应的函数值也必然不同, 即 $x_1 \neq x_2 \implies f(x_1) \neq f(x_2)$.

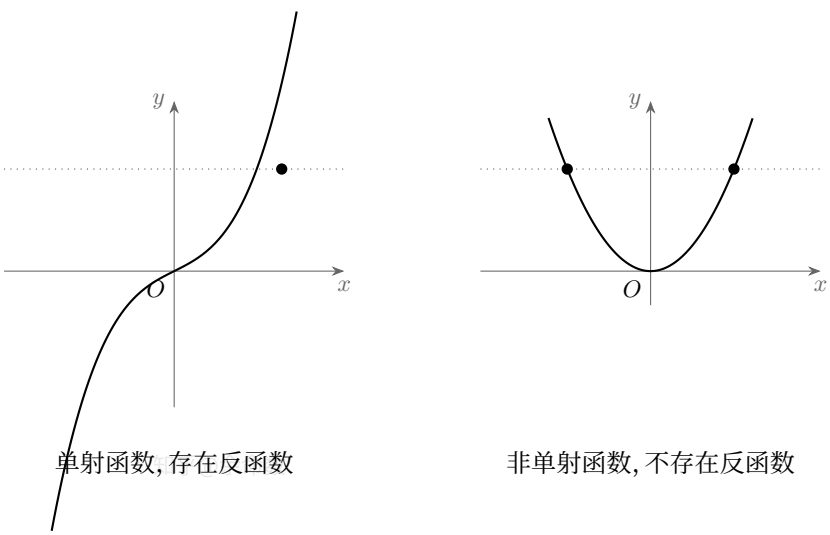
单射性给出了反函数存在的充分必要条件.

定理 3.11 设函数 f 的定义域为 A, 值域为 $C = f(A)$. 则 f 存在从 C 到 A 的反函数的充分必要条件是 f 为单射.

证明 若 f 为单射, 则对任意 $y \in C$, 由值域定义, 至少存在 $x \in A$ 使 $f(x) = y$; 又由单射性, 这样的 x 至多一个. 因此每个 $y \in C$ 都唯一对应一个 $x \in A$, 可定义 $f^{-1}(y) = x$, 这给出了从 C 到 A 的函数.

反过来, 若反函数 $f^{-1} : C \to A$ 存在, 而 f 不是单射, 则存在 $x_1 \neq x_2$ 使 $f(x_1) = f(x_2) = y$. 这会使同一个 $y \in C$ 在逆过程中同时对应 x_1 与 x_2, 与 f^{-1} 是函数矛盾. 故 f 必为单射. □

在几何上, 我们用**水平线检验法**判断一个函数是否为单射. 任取一条水平直线, 若它与函数图像至多只有一个交点, 则该函数是单射, 因而存在反函数. 严格单调函数是单射函数的一个重要子类.



单射函数, 存在反函数　　非单射函数, 不存在反函数

图 3.10 水平线检验法

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.7.2 反函数的定义与求解

定义 3.9 (反函数) 设函数 $y=f(x)$ 的定义域为 A, 值域为 C, 且 f 为单射. 对于任意 $y\in C$, 方程 $y=f(x)$ 存在唯一的解 $x\in A$. 这个解 x 依赖于 y, 因此可定义一个新函数 $g:C\to A$, 使得 $x=g(y)$. 函数 g 称为 f 的**反函数**, 记作 f^{-1}.

按照惯例, 用 x 表示自变量, y 表示因变量, 故将 $x=f^{-1}(y)$ 写为 $y=f^{-1}(x)$. 此时, 反函数的定义域是原函数的值域 C.

符号 $f^{-1}(x)$ 是反函数的标准记法, 它与 $\frac{1}{f(x)}$ 绝无关系.

确认 f 是单射后, 求解反函数可分三步.

1. **反解**. 从 $y=f(x)$ 中解出 x, 得到 $x=f^{-1}(y)$.
2. **互换**. 将 x 与 y 互换, 得到 $y=f^{-1}(x)$.
3. **注明定义域**. 反函数的定义域是原函数的值域.

例题 3.26 求函数 $f(x)=\frac{2x-1}{x+3}$ 的反函数.

解 先确定原函数 $f(x)$ 的值域. 将 $y=\frac{2x-1}{x+3}$ 变形为 $y=2-\frac{7}{x+3}$. 由于 $\frac{7}{x+3}\neq 0$, 故 $y\neq 2$. 原函数的值域为 $(-\infty,2)\cup(2,+\infty)$, 此即反函数的定义域.

接着, 从 $y=\frac{2x-1}{x+3}$ 中反解 x:

$$
\begin{aligned}
y(x+3)&=2x-1\\
xy+3y&=2x-1\\
3y+1&=x(2-y)\\
x&=\frac{3y+1}{2-y}
\end{aligned}
$$

互换 x,y 得到 $y=\frac{3x+1}{2-x}$.

因此, 所求反函数为 $f^{-1}(x)=\frac{3x+1}{2-x}$, 其定义域为 $\{x\in\mathbb{R}\mid x\neq 2\}$.

3.7.3 反函数的性质

函数 $y=f(x)$ 与它的反函数 $y=f^{-1}(x)$ 在代数与几何上互为对偶. 下面给出四条基本性质.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

定理 3.12 (反函数的性质) 设 $y = f(x)$ 是一个存在反函数 $y = f^{-1}(x)$ 的函数.

1. **图像对称性**.$y = f(x)$ 的图像与 $y = f^{-1}(x)$ 的图像关于直线 $y = x$ 对称.
2. **定义域与值域互换**.f 的定义域是 f^{-1} 的值域, f 的值域是 f^{-1} 的定义域.
3. **单调性一致性**. 若 $f(x)$ 在某个区间上单调递增(减),则其反函数 $f^{-1}(x)$ 在对应的区间上单调性相同.
4. **复合抵消性**. 对于 f^{-1} 定义域中的任意 x, $f(f^{-1}(x)) = x$. 对于 f 定义域中的任意 x, $f^{-1}(f(x)) = x$.

知乎@余命数

证明 对于图像对称性, 若点 $P(a,b)$ 在 $y = f(x)$ 的图像上, 则 $b = f(a)$. 根据反函数的定义, 这等价于 $a = f^{-1}(b)$. 故点 $Q(b,a)$ 在 $y = f^{-1}(x)$ 的图像上. 点 $P(a,b)$ 与 $Q(b,a)$ 关于直线 $y = x$ 对称, 此结论对图像上所有点均成立, 故两函数图像关于 $y = x$ 对称.

知乎@余命数 知乎@余命数

对于定义域与值域互换, 由反函数的定义, f^{-1} 的自变量取自 f 的值域 C, 而函数值回到 f 的定义域 A, 因此定义域和值域互换.

对于单调性一致性,设 f 在区间 I 上单调递增,且 $y_1 < y_2$ 为反函数定义域中对应的两个值. 令 $x_1 = f^{-1}(y_1), x_2 = f^{-1}(y_2)$. 若 $x_1 \geq x_2$, 由 f 递增得 $f(x_1) \geq f(x_2)$, 即 $y_1 \geq y_2$, 矛盾, 故 $x_1 < x_2$, 反函数递增. 若 f 单调递减, 同理由 $x_1 \leq x_2$ 会推出 $y_1 \geq y_2$ 的矛盾, 因而 $x_1 > x_2$, 反函数递减.

对于复合抵消性, 任取 x 属于 f^{-1} 的定义域, 设 $u = f^{-1}(x)$, 则按反函数定义 $f(u) = x$, 即 $f(f^{-1}(x)) = x$. 任取 x 属于 f 的定义域, 令 $y = f(x)$, 则 $f^{-1}(y) = x$, 即 $f^{-1}(f(x)) = x$. □

知乎@余命数

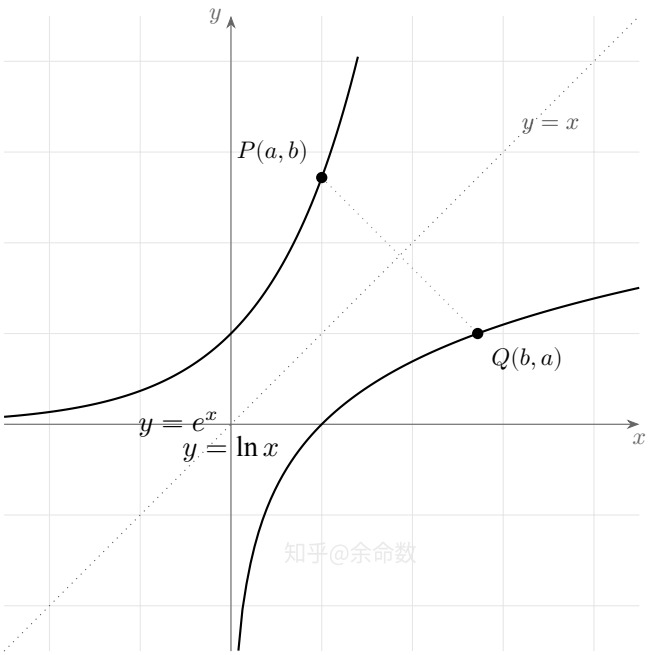
例题 3.27 已知 x_1 是方程 $3^x = \frac{4}{x}$ 的根, x_2 是方程 $\log_3 x = \frac{4}{x}$ 的根, 求 $x_1 x_2$.

解 这两个方程无法直接求解. 从函数图像分析.

方程一的根 x_1 是函数 $y_1 = 3^x$ 与 $y_g = 4/x$ 图像交点的横坐标. 方程二的根 x_2 是函数 $y_2 = \log_3 x$ 与 $y_g = 4/x$ 图像交点的横坐标. 这两个交点均唯一. 方程 $3^x = 4/x$ 等价于 $x3^x = 4$, 其定义域内必须有 $x > 0$, 而 $x3^x$ 在 $(0, +\infty)$ 上严格递增; 方程 $\log_3 x = 4/x$ 等价于 $x \log_3 x = 4$, 其解必须满足 $x > 1$, 而 $x \log_3 x$ 在 $(1, +\infty)$ 上严格递增.

分析这三个函数的对称性. $y_1 = 3^x$ 与 $y_2 = \log_3 x$ 互为反函数, 它们的图像关于直线 $y = x$ 对称. $y_g = 4/x$ 的反函数是其自身, 故其图像也关于直线 $y = x$ 对称.

知乎@余命数



知乎@余命数

图 3.11　函数与反函数图像关于 $y=x$ 对称

设 $P_1(x_1,y_1)$ 是 $y_1=3^x$ 与 $y_g=4/x$ 的一个交点. 由于 y_1 关于 $y=x$ 对称的图像是 y_2, 而 y_g 关于 $y=x$ 对称的图像仍是 y_g, y_2 与 y_g 的交点 $P_2(x_2,y_2)$ 必然是 $P_1(x_1,y_1)$ 关于直线 $y=x$ 的对称点.

因此, P_2 的坐标为 (y_1,x_1), 即 $x_2=y_1, y_2=x_1$.

由 P_1 点的定义, 有 $y_1=4/x_1$. 结合 $x_2=y_1$, 得 $x_2=4/x_1$.

故 $x_1x_2=4$.

3.8 函数的复合

给定两个函数, 我们希望依次把它们的对应法则作用到同一个输入上. 这种依次作用产生的新函数称为**复合函数**. 我们关心复合函数的定义域、值域, 以及它的单调性与奇偶性如何由内层函数和外层函数决定.

3.8.1 复合函数的定义

设有两个函数 f 和 g. 先让 g 作用于 x, 得到中间值 $u=g(x)$; 再让 f 作用于 u, 得到 $y=f(u)$. 于是从 x 到 y 的对应法则定义了一个新函数.

$$x \xrightarrow{g} u=g(x) \xrightarrow{f} y=f(u)=f(g(x))$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 3.10 (复合函数) 设有函数 $u=g(x)$ 和 $y=f(u)$,则通过代换形成的函数

$$y=f(g(x)) \tag{3-15}$$

称为由函数 g 和 f 构成的**复合函数**. 其中, x 是自变量, u 是**中间变量**. 我们称 g 为**内层函数**, f 为**外层函数**.

复合函数 $y=f(g(x))$ 有定义的前提是, 内层函数 $g(x)$ 的值域包含于外层函数 $f(u)$ 的定义域.

3.8.2 复合函数的基本性质

3.8.2.1 定义域

复合函数有定义, 当且仅当复合过程的每一步都有意义.

定理 3.13 (复合函数定义域) 设函数 f 的定义域为 D_f, 函数 g 的定义域为 D_g. 则复合函数 $y=f(g(x))$ 的定义域 $D_{f\circ g}$ 是同时满足以下两个条件的自变量 x 构成的集合.

1. $x\in D_g$(内层函数有意义);
2. $g(x)\in D_f$(外层函数有意义).

即 $D_{f\circ g}=\{x\in D_g \mid g(x)\in D_f\}$.

证明 复合函数 $f(g(x))$ 有意义, 当且仅当内层函数 $g(x)$ 先有意义且所得值可以作为 f 的输入. 前者等价于 $x\in D_g$, 后者等价于 $g(x)\in D_f$. 因此所有可取的 x 恰为

$$\{x\in D_g \mid g(x)\in D_f\}.$$

□

例题 3.28 已知函数 $f(2x-1)$ 的定义域为 $[0,1]$, 求函数 $f(\sqrt{x})$ 的定义域.

解 此问题需先确定外层函数 f 的定义域. 由题设, 当 $x\in[0,1]$ 时, $f(2x-1)$ 有意义. 即 f 的输入量, 即中间变量 $u=2x-1$ 的取值范围由 $x\in[0,1]$ 决定. 当 $x\in[0,1]$ 时, $u=2x-1\in[-1,1]$. 因此, 外层函数 $f(u)$ 的定义域为 $D_f=[-1,1]$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对于新函数 $f(\sqrt{x})$, 其定义域满足

$$\begin{cases} x \geq 0 & (\text{内层函数}\sqrt{x}\text{ 有意义}) \\ \sqrt{x} \in [-1,1] & (\text{内层函数的值在外层函数的定义域内}) \end{cases}$$

由于 $\sqrt{x} \geq 0$, 第二个条件等价于 $0 \leq \sqrt{x} \leq 1$, 平方得 $0 \leq x \leq 1$. 联立两个条件, 函数 $f(\sqrt{x})$ 的定义域为 $[0,1]$.

3.8.2.2 值域

求解复合函数 $y = f(g(x))$ 的值域, 可以由内向外进行.

1. 先在复合函数的定义域内求出内层函数 $u = g(x)$ 的值域 R_g.
2. 再将 R_g 作为外层函数 $f(u)$ 的新定义域, 在此定义域上求 $f(u)$ 的值域, 即为原复合函数的值域.

例题 3.29 求函数 $y = 4^x - 2^{x+1} + 3$ 的值域.

解 我们用换元法识别其复合结构. 令 $u = 2^x$, 则原函数可视为 $y = u^2 - 2u + 3$. 这是一个由内层函数 $u = g(x) = 2^x$ 和外层函数 $y = f(u) = u^2 - 2u + 3$ 构成的复合函数.

内层函数 $u = 2^x$ 的定义域为 $\mathbb{R}$, 其值域为 $R_g = (0, +\infty)$.

问题转化为求二次函数 $y = f(u) = (u-1)^2 + 2$ 在定义域 $u \in (0, +\infty)$ 上的值域. 此为开口向上, 对称轴为 $u = 1$ 的抛物线. 由于对称轴 $u = 1$ 位于定义域 $(0, +\infty)$ 内, 函数在 $u = 1$ 处取得最小值 $y_{\min} = f(1) = 2$.

因此, 原函数的值域为 $[2, +\infty)$.

3.8.2.3 单调性与奇偶性

下面讨论复合函数的单调性与奇偶性.

定理 3.14 (复合函数单调性) 设复合函数 $y = f(g(x))$ 在区间 I 上有定义, 且 $u = g(x)$ 与 $y = f(u)$ 在各自对应的定义域区间上单调.

- 若 $f(u)$ 与 $g(x)$ 单调性**相同**, 即同为非减或同为非增, 则 $f(g(x))$ 为**单调递增(非减)**.
- 若 $f(u)$ 与 $g(x)$ 单调性**相异**, 即一非减一非增, 则 $f(g(x))$ 为**单调递减(非增)**.

若相应的两个单调性均为严格单调, 则结论也相应为严格单调.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

证明 设 $x_1, x_2 \in I$ 且 $x_1 < x_2$，令 $u_1 = g(x_1)$，$u_2 = g(x_2)$. 若 g 与 f 均非减，则 $u_1 \le u_2$，进而 $f(u_1) \le f(u_2)$，复合函数非减. 若 g 与 f 均非增，则 $u_1 \ge u_2$，由 f 非增得 $f(u_1) \le f(u_2)$，复合函数仍非减. 若 g 非减而 f 非增，则 $u_1 \le u_2$ 推出 $f(u_1) \ge f(u_2)$，复合函数非增. 若 g 非增而 f 非减，同理有 $u_1 \ge u_2$ 推出 $f(u_1) \ge f(u_2)$，复合函数非增. 严格单调情形只需将相应不等号替换为严格不等号. □

乎@余命数

定理 3.15 (复合函数奇偶性) 设复合函数 $y = f(g(x))$ 的定义域 D 关于原点对称.

知乎@余命数

1. 若内层函数 $g(x)$ 是**偶函数**，则 $y = f(g(x))$ 必为**偶函数**.
2. 若内层函数 $g(x)$ 是**奇函数**，且外层函数 $f(u)$ 在 $g(D)$ 上为偶函数或奇函数，则复合函数的奇偶性与外层函数 $f(u)$ 相同.

证明 (奇偶性定理证明) 令 $F(x) = f(g(x))$. 1. 若 g 为偶函数：$F(-x) = f(g(-x)) = f(g(x)) = F(x)$. 故 $F(x)$ 为偶函数. 2. 若 g 为奇函数：$F(-x) = f(g(-x)) = f(-g(x))$. 若 f 为偶函数，$f(-g(x)) = f(g(x)) = F(x)$，故 $F(x)$ 为偶函数. 若 f 为奇函数，$f(-g(x)) = -f(g(x)) = -F(x)$，故 $F(x)$ 为奇函数. □

知乎@余命数

3.8.3 函数相等

两个函数相等的含义由下面的定义给出.

定义 3.11 (函数相等) 若两个函数 $f(x)$ 和 $g(x)$ 的**定义域相同**，且**对应法则也相同**，则称这两个函数相等，记为 $f(x) = g(x)$.

判断两个函数是否相等，需要同时检查这两个方面. 通常可先把解析式化到最简，再比较化简结果.

例题 3.30 判断 $f(x) = x + 2$ 与 $g(x) = \frac{x^2-4}{x-2}$ 是否为同一函数.

解 先看定义域. $f(x) = x + 2$ 的定义域为 $\mathbb{R}$. $g(x) = \frac{x^2-4}{x-2}$ 的分母不能为零，即 $x \ne 2$. 其定义域为 $\{x \in \mathbb{R} \mid x \ne 2\}$.

知乎@余命数

由于定义域不同，这两个函数不是同一函数.

知乎@余命数

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

3.9 函数的图像变换

我们从基本函数的图像出发，通过平移、伸缩、对称等操作构造更复杂函数的图像. 本节建立这些几何操作与代数表示之间的联系.

3.9.1 变换的统一原理

所有图像变换都归结为新旧坐标之间的代换关系.

设原函数图像满足方程 $y_0 = f(x_0)$，其上任意一点为 (x_0, y_0). 经过变换，该点移动至新图像上的点 (x, y). 若新旧坐标间的变换关系可表示为

$$x_0 = g(x, y)$$

$$y_0 = h(x, y)$$

则将此关系代入原方程 $y_0 = f(x_0)$，即得新图像的方程 $h(x, y) = f(g(x, y))$.

注 3.6 (坐标变换与方程变换的对偶性)
方程的变换与点的坐标变换在形式上互为逆过程. 例如，将图像上的所有点 (x_0, y_0) 向右平移 a 个单位得到新点 (x, y)，其坐标关系为 $x = x_0 + a, y = y_0$. 为建立新坐标 (x, y) 满足的方程，必须用新坐标表示旧坐标，即 $x_0 = x - a, y_0 = y$，然后代入原方程，得到 $y = f(x - a)$. 作用于坐标的变换是 $+a$，而作用于方程中变量的变换是 $-a$.

3.9.2 基本图像变换

定理 3.16 (平移与伸缩变换) 设原函数为 $y = f(x)$，a, b, ω, A 均为正实数.

- **水平平移**：$y = f(x - a)$ 的图像由 $y = f(x)$ 向右平移 a 个单位得到.
- **竖直平移**：$y = f(x) + b$ 的图像由 $y = f(x)$ 向上平移 b 个单位得到.
- **水平伸缩**：$y = f(\omega x)$ 的图像由 $y = f(x)$ 上各点的横坐标变为原来的 $1/\omega$ 倍得到(纵坐标不变).
- **竖直伸缩**：$y = Af(x)$ 的图像由 $y = f(x)$ 上各点的纵坐标变为原来的 A 倍得到(横坐标不变).

证明 设原图像上一点为 (x_0, y_0)，则 $y_0 = f(x_0)$. 水平右移 a 后新点满足 $x = x_0 + a, y = y_0$，即 $x_0 = x - a$，故新图像方程为 $y = f(x - a)$. 竖直上移 b 后新点满足 $x = x_0, y = y_0 + b$，即 $y - b = f(x)$，故 $y = f(x) + b$.

水平伸缩时新点满足 $x = x_0/\omega, y = y_0$，即 $x_0 = \omega x$，故新图像方程为 $y = f(\omega x)$. 竖直伸缩时新点满足 $x = x_0, y = Ay_0$，即 $y/A = f(x)$，故 $y = Af(x)$. □

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 3.17 (对称变换) 设原函数为 $y=f(x)$.

- 关于 **x 轴**对称的图像方程为 $y=-f(x)$.
- 关于 **y 轴**对称的图像方程为 $y=f(-x)$.
- 关于 **原点**对称的图像方程为 $y=-f(-x)$.
- 关于直线 **y=x** 对称的图像方程为 $x=f(y)$; 当 f 存在反函数时, 该图像可写为 $y=f^{-1}(x)$.

证明 关于 x 轴对称时, 新旧点满足 $x=x_0, y=-y_0$, 即 $-y=f(x)$, 故 $y=-f(x)$. 关于 y 轴对称时, 新旧点满足 $x=-x_0, y=y_0$, 即 $y=f(-x)$. 关于原点对称时, 新旧点满足 $x=-x_0, y=-y_0$, 即 $-y=f(-x)$, 故 $y=-f(-x)$.

关于直线 $y=x$ 对称时, 新旧点满足 $x=y_0, y=x_0$. 由 $y_0=f(x_0)$ 得 $x=f(y)$. 若 f 为单射并存在反函数, 则该关系等价于 $y=f^{-1}(x)$. □

上述变换都可由统一原理直接推出.

3.9.3 复合变换

对于形如 $y=Af(\omega x+\phi)+k$ 的函数, 我们须注意其图像变换的顺序. 为建立坐标关系, 先把它改写为标准形式

$$y-k=Af\left(\omega\left(x+\frac{\phi}{\omega}\right)\right)$$

从基函数 $y_0=f(x_0)$ 到目标函数的坐标映射关系为

$$x_0=\omega\left(x+\frac{\phi}{\omega}\right),\quad y_0=\frac{y-k}{A}$$

反解出新坐标 (x,y) 与旧坐标 (x_0,y_0) 的关系, 得

$$x=\frac{1}{\omega}x_0-\frac{\phi}{\omega},\quad y=Ay_0+k$$

此坐标变换过程可分解为伸缩, 或对称变换与平移变换的复合. 我们采用**先伸缩, 后平移**的作图顺序. 水平变换与竖直变换的次序无关, 但各自内部先伸缩后平移的次序不可颠倒.

例题 3.31 叙述由 $y=\sin x$ 的图像变换得到 $y=-2\sin\left(\frac{1}{2}x-\frac{\pi}{6}\right)+1$ 图像的过程.

解 将目标函数改写为标准形式 $y-1=-2\sin\left(\frac{1}{2}\left(x-\frac{\pi}{3}\right)\right)$. 变换过程如下.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1. 水平变换. (a) 将 $y=\sin x$ 图像上所有点的横坐标变为原来的 2 倍, 得到 $y=\sin(\frac{1}{2}x)$ 的图像. (b) 再将所得图像向右平移 $\frac{\pi}{3}$ 个单位, 得到 $y=\sin(\frac{1}{2}(x-\frac{\pi}{3}))$ 的图像.

2. 竖直变换. (a) 将前一步所得图像上所有点的纵坐标变为原来的 -2 倍, 即伸长到 2 倍, 再关于 x 轴对称, 得到 $y=-2\sin(\frac{1}{2}(x-\frac{\pi}{3}))$ 的图像. (b) 最后将所得图像向上平移 1 个单位, 得到最终图像.

3.9.4 绝对值变换

我们讨论含绝对值的函数, 其图像可通过对坐标轴翻折得到.

定理 3.18 (翻折变换) 设原函数为 $y=f(x)$.

1. **函数 $y=f(|x|)$ 的图像**: 由 $f(|x|)=\begin{cases}f(x), & x\geq 0\\ f(-x), & x<0\end{cases}$ 可知, 新函数为偶函数. 其图像按两步构造. 保留 $y=f(x)$ 图像在 y 轴右侧, 含 y 轴的部分, 并将此部分关于 y 轴对称翻折到左侧, 构成完整图像.
2. **函数 $y=|f(x)|$ 的图像**: 由 $|f(x)|=\begin{cases}f(x), & f(x)\geq 0\\ -f(x), & f(x)<0\end{cases}$ 可知, 新函数的值域非负. 其图像按两步构造. 保留 $y=f(x)$ 图像在 x 轴上方, 含 x 轴的部分, 并将 x 轴下方的部分关于 x 轴对称翻折到上方.

证明 对 $y=f(|x|)$, 当 $x\geq 0$ 时 $|x|=x$, 所以图像与原图像右半部分一致; 当 $x<0$ 时 $|x|=-x$, 因而左侧点 $(x,f(|x|))$ 与右侧点 $(-x,f(-x))$ 关于 y 轴对称. 这说明保留右半部分并关于 y 轴翻折即可得到全图.

对 $y=|f(x)|$, 当 $f(x)\geq 0$ 时图像点不变; 当 $f(x)<0$ 时 $|f(x)|=-f(x)$, 即原来位于 x 轴下方的点 $(x,f(x))$ 被映到关于 x 轴对称的点 $(x,-f(x))$. 因此保留 x 轴上方部分并将下方部分向上翻折即可得到全图. □

导函数对称性

3.9.5 导函数的对称性

定理 3.19 若可导函数 $y=f(x)$ 的图像关于直线 $x=a$ 轴对称, 则其导函数 $y=f'(x)$ 的图像关于点 $(a,0)$ 中心对称.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

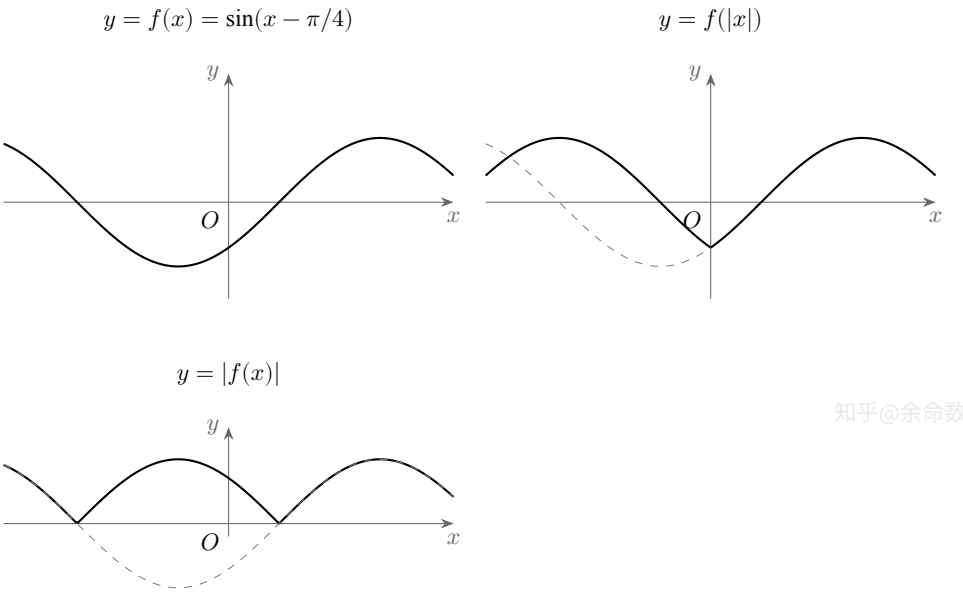


图 3.12 翻折变换示例

知乎@余命数

证明 函数 $f(x)$ 的图像关于 $x = a$ 对称, 等价于 $f(a - x) = f(a + x)$. 我们把这个恒等式两边关于 x 求导. 由链式法则得

$$[f(a-x)]' = [f(a+x)]'$$

$$f'(a-x)\cdot(a-x)' = f'(a+x)\cdot(a+x)'$$

$$-f'(a-x) = f'(a+x)$$

这正是导函数 $f'(x)$ 的图像关于点 $(a, 0)$ 中心对称的定义. □

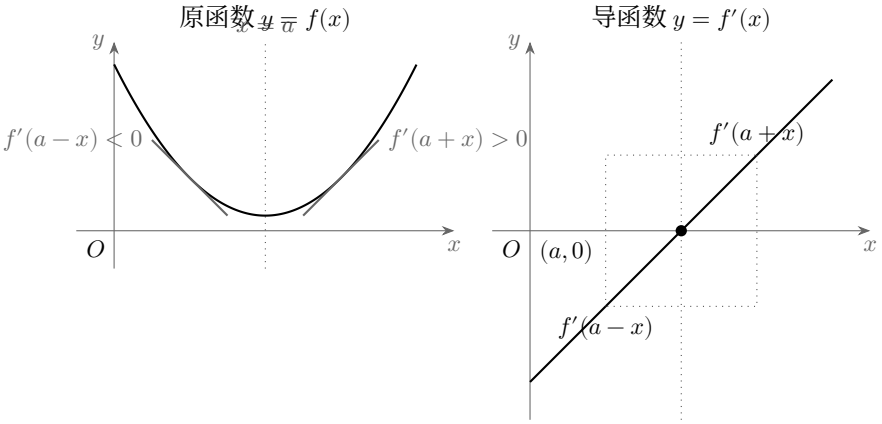


图 3.13 轴对称原函数与中心对称导函数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.10 常见多项式函数

多项式函数 $P(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 由幂函数通过有限次加法与数乘得到, 其定义域为 $\mathbb{R}$. 我们先回顾一次函数与二次函数的主要性质, 再分析三次函数的对称性、单调性与根.

注 3.7
三次函数的极值、拐点与切线条数讨论使用导数语言, 见第16章.

3.10.1 一次与二次函数回顾

一次函数 $f(x)=kx+b\ (k\neq 0)$ 的斜率 k 处处相等, 即其变化率恒定(分析语言中即 $f'(x)=k$ 为常数, 见 第16章). 在几何上, 直线是自身在每一点的切线, 并且关于其上任意一点都构成中心对称图形.

二次函数 $f(x)=ax^2+bx+c\ (a\neq 0)$ 的**变化率的变化率恒定**. 其弯曲方向在整个定义域上保持不变 (分析语言中 $f''(x)=2a$ 为非零常数, 见 第16章). 抛物线的基本性质是**轴对称性**, 对称轴为直线 $x=-\frac{b}{2a}$. 这一对称性把函数的几何与代数性质联系起来. 设方程 $ax^2+bx+c=0$ 的两根为 x_1,x_2, 根据韦达定理, $x_1+x_2=-b/a$. 因此对称轴的横坐标 $\frac{x_1+x_2}{2}=-\frac{b}{2a}$ 恰好是两根的算术平均值. 对称轴即为函数零点的几何中心.

知乎@余命数

三次函数的导数视角

知乎@余命数

3.10.2 三次函数

三次函数 $f(x)=ax^3+bx^2+cx+d\ (a\neq 0)$ 的形态与性质更丰富, 其一阶导数与二阶导数分别决定单调性、极值与凹凸性、拐点.

知乎@余命数

$$f'(x)=3ax^2+2bx+c$$

$$f''(x)=6ax+2b$$

一阶导数是二次函数, 决定原函数的单调性与极值; 二阶导数是一次函数, 决定原函数的凹凸性与拐点. 我们通过导函数研究原函数.

3.10.2.1 单调性与极值

三次函数的单调性完全取决于其导函数 $f'(x)=3ax^2+2bx+c$ 的符号. $f'(x)$ 是二次函数, 其符号由判别式 $\Delta=(2b)^2-4(3a)c=4(b^2-3ac)$ 决定.

1. **当 $\Delta>0$ 时**, 即 $b^2-3ac>0$, 方程 $f'(x)=0$ 有两个不相等的实数根 x_1,x_2. 这两个根是 $f(x)$ 的极值点. 函数 $f(x)$ 在 $(-\infty,\min\{x_1,x_2\})$ 和 $(\max\{x_1,x_2\},+\infty)$ 上的单调性相同, 在 $(\min\{x_1,x_2\},\max\{x_1,x_2\})$ 上单调性相反. 其图像呈 S 形, 有一个极大值点和一个极小值点.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. **当 $\Delta = 0$ 时**, 即 $b^2 - 3ac = 0$, 方程 $f'(x) = 0$ 有两个相等的实数根 x_0. 此时 $f'(x)$ 在 $\mathbb{R}$ 上除一点外恒为正或恒为负. 函数 $f(x)$ 在整个定义域 $\mathbb{R}$ 上是单调的, 没有极值点. 但在 x_0 处, 其切线斜率为零, 该点是一个水平拐点.
3. **当 $\Delta < 0$ 时**, 即 $b^2 - 3ac < 0$, 方程 $f'(x) = 0$ 无实数根.$f'(x)$ 在 $\mathbb{R}$ 上恒为正或恒为负. 函数 $f(x)$ 在整个定义域 $\mathbb{R}$ 上是严格单调的, 既无极值点, 也无水平切线.

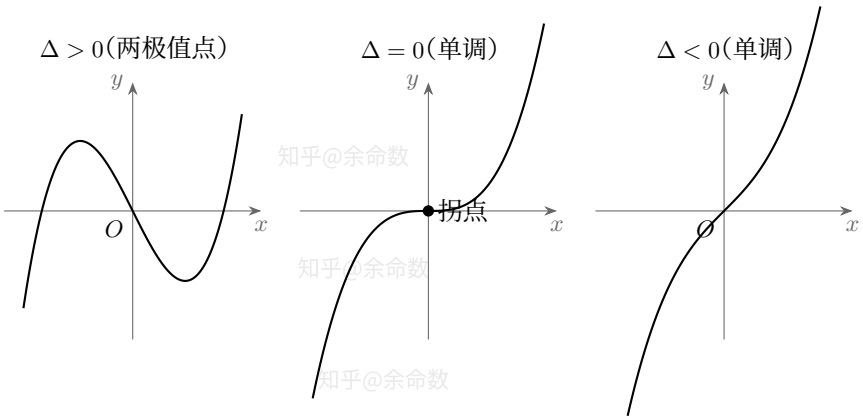


图 3.14 三次函数的三种基本形态 ($a > 0$)

3.10.2.2 对称性

二次函数的图像是轴对称的, 三次函数的图像是中心对称的.

> **定理 3.20** 任何三次函数 $f(x) = ax^3 + bx^2 + cx + d$ $(a \neq 0)$ 的图像都是中心对称图形, 其对称中心是唯一的拐点 $(x_0, f(x_0))$.

证明 函数的拐点对应凹凸性的改变, 由二阶导数 $f''(x) = 0$ 确定.

$$f''(x) = 6ax + 2b = 0 \implies x_0 = -\frac{b}{3a}$$

此即拐点的横坐标. 该点也是导函数 $f'(x) = 3ax^2 + 2bx + c$ 这条抛物线的对称轴.

为证明 $(x_0, f(x_0))$ 是对称中心, 我们验证 $f(x_0 + t) + f(x_0 - t) = 2f(x_0)$ 对任意实数 t 恒成立. 将三次函数在 x_0 处展开, 得

$$f(x_0 + t) = f(x_0) + f'(x_0)t + \frac{f''(x_0)}{2}t^2 + at^3.$$

因为 $f''(x_0) = 0$, 有

$$f(x_0 + t) = f(x_0) + f'(x_0)t + at^3, \qquad f(x_0 - t) = f(x_0) - f'(x_0)t - at^3.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两式相加即得 $f(x_0+t)+f(x_0-t)=2f(x_0)$, 故图像关于 $(x_0,f(x_0))$ 中心对称. 由于三次函数只有一个二阶导数零点, 该拐点唯一. □

由这一对称性可得一个推论. 若三次函数存在两个极值点 x_1,x_2, 则它们关于对称中心的横坐标 x_0 对称, 即 $x_0=\frac{x_1+x_2}{2}$. 同时, 两个极值 $f(x_1),f(x_2)$ 也关于对称中心的纵坐标 $f(x_0)$ 对称, 即 $f(x_0)=\frac{f(x_1)+f(x_2)}{2}$.

3.10.2.3 三次方程实根个数的判定

三次方程 $f(x)=ax^3+bx^2+cx+d=0$ 实根的个数, 等于函数图像 $y=f(x)$ 与 x 轴的交点个数. 我们由极值性质判定这个几何问题.

我们先考察导函数 $f'(x)=3ax^2+2bx+c$. 其判别式 $\Delta=4(b^2-3ac)$ 给出分类的枢纽.

先看函数不存在极值的情形. 当 $b^2-3ac\le 0$ 时, 导函数 $f'(x)$ 在实数域上至多有一个零点, 从而 $f'(x)$ 恒为非负或恒为非正. 故原函数 $f(x)$ 在整个定义域 $\mathbb{R}$ 上是单调的. 一个连续的单调函数, 其图像与任意水平直线(包括 x 轴)至多有一个交点. 又因 $\lim_{x\to\pm\infty}f(x)$ 符号相反, 故必有一个交点. 因此, 在此情形下, 方程 $f(x)=0$ 有且仅有一个实根.

当 $b^2-3ac>0$ 时, 函数 $f(x)$ 存在一个极大值与一个极小值, 不妨设它们分别出现在 x_1 与 x_2 处. 函数图像呈升-降-升或降-升-降的形态. 图像与 x 轴的交点个数, 完全取决于两个极值 $f(x_1)$ 和 $f(x_2)$ 相对于 x 轴的位置. 我们考察两个极值的乘积 $f(x_1)\cdot f(x_2)$ 的符号.

- 若 $f(x_1)\cdot f(x_2)>0$, 即两个极值点均在 x 轴的同一侧(同为正或同为负). 此时, 函数图像在达到一个极值后折返, 但在到达另一个极值前尚未触及 x 轴. 因此, 图像仅与 x 轴有一个交点, 方程仅有一个实根.
- 若 $f(x_1)\cdot f(x_2)=0$, 即其中一个极值点恰好落在 x 轴上. 此时, 函数图像在 x 轴上有一个切点(二重根)和一个穿根点. 因此, 方程有两个不相等的实根.
- 若 $f(x_1)\cdot f(x_2)<0$, 即极大值在 x 轴上方, 而极小值在 x 轴下方. 根据连续函数介值定理, 函数图像在两个极值点之间必然要穿越一次 x 轴. 加上两端趋于无穷的行为, 图像共与 x 轴有三个不同的交点. 因此, 方程有三个不相等的实根.

我们把三次方程实根个数的判定条件总结如下. 设 $f(x)=ax^3+bx^2+cx+d=0$, 且 x_1,x_2 是其导函数 $f'(x)=0$ 的根.

- **一个实根**的充要条件是 $b^2-3ac\le 0$, 或 $b^2-3ac>0$ 且 $f(x_1)f(x_2)>0$.
- **两个不等实根**的充要条件是 $b^2-3ac>0$ 且 $f(x_1)f(x_2)=0$.
- **三个不等实根**的充要条件是 $b^2-3ac>0$ 且 $f(x_1)f(x_2)<0$.

例题 3.32　确定方程 $2x^3-9x^2+12x-4=0$ 的实根个数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

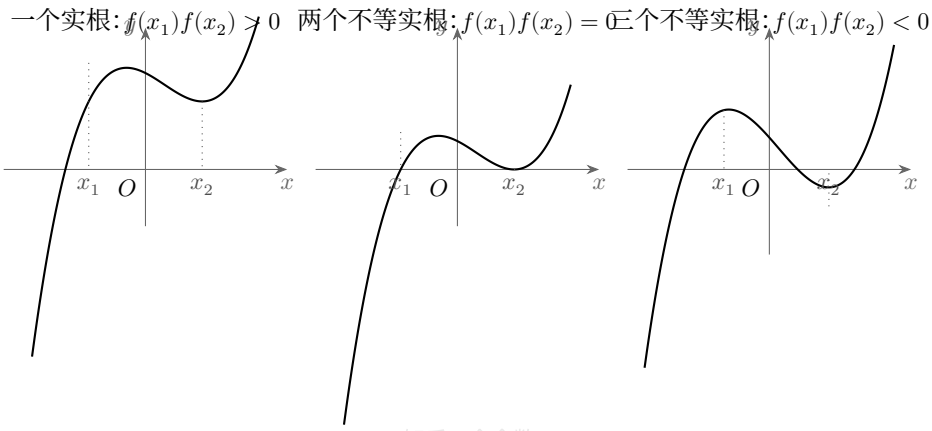


图 3.15 三次函数存在极值时, 实根个数的几何情形

解 此问题等价于考察函数 $f(x)=2x^3-9x^2+12x-4$ 的图像与 x 轴的交点个数, 策略是先由导数确定单调性与极值, 再依极值与零的位置关系判定根的数目.

求导得

$$f'(x)=6x^2-18x+12=6(x^2-3x+2)=6(x-1)(x-2)$$

令 $f'(x)=0$, 得临界点 $x_1=1$ 与 $x_2=2$, 故 $f(x)$ 有两个极值点. 两极值为

$$f(1)=2(1)^3-9(1)^2+12(1)-4=2-9+12-4=1$$

$$f(2)=2(2)^3-9(2)^2+12(2)-4=16-36+24-4=0$$

极小值 $f(2)$ 恰好为零, 即函数图像在点 $(2,0)$ 处与 x 轴相切.

极值之积为 $f(1)\cdot f(2)=1\cdot 0=0$. 依判定准则, 极值之积为零时, 函数图像与 x 轴有一个切点和一个穿根点.

因此, 原方程有两个不相等的实根, 其中一个是二重根.

例题 3.33 讨论关于 x 的方程 $x^3-3x=m$ 的实根个数随参数 m 的变化情况.

解 按参数 m 的取值对三次方程 $x^3-3x-m=0$ 的根的结构作完整分类. 令 $f(x)=x^3-3x-m$, 分析 $f(x)$ 的零点个数.

其导函数为 $f'(x)=3x^2-3=3(x^2-1)$, 令 $f'(x)=0$ 得极值点 $x_1=-1$ 与 $x_2=1$, 极值点的存在与参数 m 无关. 两极值为

$$f(-1)=(-1)^3-3(-1)-m=-1+3-m=2-m$$

$$f(1)=(1)^3-3(1)-m=1-3-m=-2-m$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

实根个数取决于两极值乘积的符号：

$$f(-1)f(1)=(2-m)(-2-m)=(m-2)(m+2)$$

据此对 m 分类：

- **当有三个不等实根时**，须 $f(-1)f(1)<0$，即 $(m-2)(m+2)<0$. 解得 $-2<m<2$.
- **当有两个不等实根时**，须 $f(-1)f(1)=0$，即 $(m-2)(m+2)=0$. 解得 $m=2$ 或 $m=-2$.
- **当仅有一个实根时**，须 $f(-1)f(1)>0$，即 $(m-2)(m+2)>0$. 解得 $m>2$ 或 $m<-2$.

知乎@余命数　知乎@余命数　知乎@余

注 3.8 (分离参数法)
一个等价且几何直观更强的视角是分离参数. 原方程 $x^3-3x=m$ 的根的个数，等价于函数 $g(x)=x^3-3x$ 的图像与水平直线 $y=m$ 的交点个数. 通过分析 $g(x)$ 的极大值 $g(-1)=2$ 与极小值 $g(1)=-2$，可直观地得出相同的结论.

例题 3.34　已知函数 $f(x)=x^3-ax^2+(a^2-2a)x$. 若 $f(x)$ 在区间 $(1,2)$ 内恰有一个极值点，求实数 a 的取值范围.

解　函数 $f(x)$ 在区间 $(1,2)$ 内恰有一个极值点，等价于导函数

$$f'(x)=3x^2-2ax+(a^2-2a)$$

知乎@余命数　知乎@余命数

在 $(1,2)$ 内恰有一个零点，且该零点处导数符号发生改变. 由于 $f'(x)$ 是开口向上的二次函数，先用端点符号确定根的位置.

计算 $f'(1)$ 与 $f'(2)$，得

$$f'(1)=3-2a+a^2-2a=a^2-4a+3=(a-1)(a-3)$$

$$f'(2)=3(4)-2a(2)+a^2-2a=a^2-6a+12$$

知乎@余命数

对于 $f'(2)=a^2-6a+12=(a-3)^2+3$，其值恒为正.

因此，条件 $f'(1)\cdot f'(2)<0$ 就等价于 $f'(1)<0$.

知乎@余命数

$$(a-1)(a-3)<0$$

解得 $1<a<3$. 在此范围内，$f'(1)<0$ 而 $f'(2)>0$，故由连续性可知 $f'(x)$ 在 $(1,2)$ 内至少有一个零点. 又因为开口向上的二次函数在 $x=1$ 处取负值，$x=2$ 处取正值，所以 $x=2$ 位于两个根的右侧之外，区间 $(1,2)$ 内恰含右侧一个根；该根为单根，导数符号由负变正，对应 $f(x)$ 的一个极小值点.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

还需检验边界情况. 若 $f'(1)=0$, 则 $a=1$ 或 $a=3$. 当 $a=1$ 时, $f'(x)=3x^2-2x-1=(3x+1)(x-1)$. 根为 $x=-1/3$ 和 $x=1$. 区间 $(1,2)$ 内无根, 不符. 当 $a=3$ 时, $f'(x)=3x^2-6x+3=3(x-1)^2$. 根为 $x=1$(二重根). 区间 $(1,2)$ 内无根, 不符.

若 $f'(2)=0$, 由于 $f'(2)=(a-3)^2+3$ 恒大于零, 此情况不会发生.

故 a 的取值范围是 $(1,3)$.

例题 3.35 设函数 $f(x)=\frac{1}{3}x^3-ax^2+1$. 若 $a>2$, 判断函数 $f(x)$ 在区间 $(0,2)$ 内的零点个数.

知乎@余命数

解 先判断 $f(x)$ 在区间 $(0,2)$ 上的单调性, 再结合端点处函数值的符号确定零点个数.

求导得

$$f'(x)=x^2-2ax=x(x-2a)$$

题设 $a>2$ 给出 $2a>4$. 对任意 $x\in(0,2)$, $x>0$ 且 $x-2a<2-4=-2<0$, 两因子异号, 故 $f'(x)=x(x-2a)<0$ 恒成立, $f(x)$ 在 $(0,2)$ 上严格单调递减.

再看区间端点: 左端点 $f(0)=1>0$; 右端点 $f(2)=\frac{1}{3}(2)^3-a(2)^2+1=\frac{8}{3}-4a+1=\frac{11}{3}-4a$. 由 $a>2$ 知 $4a>8$, 故 $f(2)=\frac{11}{3}-4a<\frac{11}{3}-8=-\frac{13}{3}<0$.

$f(x)$ 在 $[0,2]$ 上连续, 在 $(0,2)$ 上严格单调递减, 且两端点的值异号 ($f(0)>0, f(2)<0$). 由连续函数介值定理, $f(x)$ 在 $(0,2)$ 内存在唯一零点.

故函数在区间 $(0,2)$ 上恰好有 1 个零点.

例题 3.36 设 a 为实数, 函数 $f(x)=-x^3+3x+a$.

1. 求 $f(x)$ 的极值.
2. 若函数 $y=f(x)$ 恰好有两个零点, 求 a 的值.

解 (1) 极值由导数的零点决定. 求导得

$$f'(x)=-3x^2+3$$

知乎@余命数

令 $f'(x)=0$, 即 $-3x^2+3=0$, 解得 $x=\pm1$.

依 $f'(x)$ 的符号, 当 $x\in(-\infty,-1)\cup(1,\infty)$ 时, $f'(x)<0$, 函数单调递减; 当 $x\in(-1,1)$ 时, $f'(x)>0$, 函数单调递增.

因此, $f(x)$ 在 $x=-1$ 处取得极小值 $f_{极小}=f(-1)=-(-1)^3+3(-1)+a=1-3+a=a-2$, 在 $x=1$ 处取得极大值 $f_{极大}=f(1)=-(1)^3+3(1)+a=-1+3+a=a+2$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2) 函数 $y=f(x)$ 恰有两个零点, 即图像与 x 轴恰有两个交点. 对既有极大值又有极小值的三次函数, 这当且仅当其中一个极值点恰好落在 x 轴上, 即极大值或极小值之一为零. 故

$$f_{\text{极小}}=a-2=0 \quad \text{或} \quad f_{\text{极大}}=a+2=0$$

知乎@余命数

解得 $a=2$ 或 $a=-2$.

知乎@余

例题 3.37　已知函数 $f(x)=\frac{1}{3}x^3-ax^2+b$, 且其在 $x=-2$ 处取得极值.

1. 求函数 $f(x)$ 的单调区间.
2. 若函数 $f(x)$ 在闭区间 $[-3,3]$ 上有且仅有一个零点, 求实数 b 的取值范围.

解 (I) 极值点是导函数的零点. 求导得

$$f'(x)=x^2-2ax$$

题设 $f(x)$ 在 $x=-2$ 处取得极值, 故 $f'(-2)=0$, 即 $f'(-2)=(-2)^2-2a(-2)=4+4a=0$, 解得 $a=-1$.

由此 $f'(x)=x^2+2x=x(x+2)$. 令 $f'(x)>0$ 得 $x>0$ 或 $x<-2$; 令 $f'(x)<0$ 得 $-2<x<0$.

知乎@余命数

故 $f(x)$ 的严格单调递增区间是 $(-\infty,-2)$ 和 $(0,+\infty)$, 严格单调递减区间是 $(-2,0)$.

知乎@余命数

(II) 由 (I), 所研究函数为 $f(x)=\frac{1}{3}x^3+x^2+b$, 在 $(-\infty,-2)$ 上递增, 在 $(-2,0)$ 上递减, 在 $(0,\infty)$ 上递增. 取区间端点及内部极值点处的函数值:

$$f(-3)=\frac{1}{3}(-3)^3+(-3)^2+b=-9+9+b=b$$

$$f(-2)=\frac{1}{3}(-2)^3+(-2)^2+b=-\frac{8}{3}+4+b=b+\frac{4}{3} \quad \text{(区间内的极大值)}$$

$$f(0)=b \quad \text{(区间内的极小值)}$$

$$f(3)=\frac{1}{3}(3)^3+(3)^2+b=9+9+b=b+18$$

函数图像在 $[-3,3]$ 上的形态固定, 仅随 b 竖直平移. 为使函数在此区间上有且仅有一个零点, 图像的最高点与最低点须几乎完全位于 x 轴的一侧. 分两种情形:

知乎@余命数

情形一: 极大值低于 x 轴, 即 $f(-2)<0$, 亦即 $b<-4/3$. 此时 $f(-3)=b$, $f(-2)$ 与 $f(0)=b$ 均为负. 存在唯一零点须且只需区间右端点的值非负, 即 $f(3)\geq 0$, 亦即 $b\geq -18$. 合并得 $-18\leq b<-4/3$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

情形二：极小值高于 x 轴，即 $f(0)>0$，亦即 $b>0$. 此时 $f(0),f(-2),f(3)$ 均为正. 存在唯一零点须且只需 $f(-3)\le 0$，即 $b\le 0$. 此与 $b>0$ 矛盾，故此情形不可能发生.

再检验边界. 若 $f(-2)=0$ ($b=-4/3$)，则函数在 $x=-2$ 处有二重根，且 $f(3)=-4/3+18>0$，$f(0)=-4/3<0$，故 $(0,3)$ 内还有一根，不符. 若 $f(0)=0$ ($b=0$)，则 $f(x)=x^2(\frac{1}{3}x+1)$，零点为 $x=0$ 与 $x=-3$，在 $[-3,3]$ 内有两个零点，不符.

故实数 b 的取值范围是 $[-18,-4/3)$.

3.10.2.4 根与系数的关系

定理 3.21 (三次方程根与系数的关系) 设 $a\neq 0$，方程 $ax^3+bx^2+cx+d=0$ 的三个根（按重数计算，可能为复数根）为 x_1,x_2,x_3，则：

$$x_1+x_2+x_3=-b/a \tag{3-16}$$

$$x_1x_2+x_1x_3+x_2x_3=c/a \tag{3-17}$$

$$x_1x_2x_3=-d/a \tag{3-18}$$

证明 由代数基本定理和因式定理，多项式 ax^3+bx^2+cx+d 可唯一地分解为 $a(x-x_1)(x-x_2)(x-x_3)$. 展开右侧的表达式，并与左侧的同次项系数比较，即可导出上述三个关系式. □

三次函数切线与极值几何

3.10.2.5 切线问题

过平面上任意一点，可以作三次函数图像若干条切线，条数随点的位置而变. 设切点为 $(t,f(t))$，则切线方程为 $y-f(t)=f'(t)(x-t)$. 若该切线经过点 $P(x_0,y_0)$，则

$$y_0-f(t)=f'(t)(x_0-t)$$

将 $f(t)$ 和 $f'(t)$ 的表达式代入，这是一个关于切点横坐标 t 的三次方程. 此方程实根的个数即为过点 P 可作的切线条数. 对此三次方程根的个数分析，可以得到一个几何结论.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 3.22 (切线条数判定) 三次函数 $y=f(x)$ 的图像及其在拐点处的切线(称之为拐点切线), 将整个平面划分为四个区域.

- 若点 $P(x_0,y_0)$ 位于由函数图像与拐点切线所夹的两个开区域内, 则过该点可作 **3 条**不同的切线.
- 若点 P 位于函数图像上或拐点切线上(但非拐点本身), 则可作 **2 条**切线.
- 若点 P 位于其余两个开区域内, 或 P 恰为拐点, 则仅可作 **1 条**切线.

证明 经过平移坐标、减去拐点切线对应的一次函数并作非零比例伸缩, 可将一般三次函数可逆地化为标准三次曲线 $y=x^3$. 这种可逆仿射变换把直线映为直线、保持相切关系, 也保持点相对函数图像和拐点切线所夹区域的位置关系. 因此只需证明 $y=x^3$ 的情形, 其拐点为原点, 拐点切线为 $y=0$.

设切点为 (t,t^3). 该点处切线为

$$y-t^3=3t^2(x-t),$$

若切线经过 $P(u,v)$, 则

$$v=3ut^2-2t^3,\qquad \text{即}\qquad 2t^3-3ut^2+v=0. \tag{1}$$

过 P 的不同切线条数等于方程 (1) 的不同实根个数, 但重根对应重合切线需只计一次.

令 $\Phi(t)=3ut^2-2t^3$. 当 $u>0$ 时, $\Phi'(t)=6t(u-t)$, 极小值为 $\Phi(0)=0$, 极大值为 $\Phi(u)=u^3$. 故方程 $v=\Phi(t)$ 在 $0<v<u^3$ 时有三个不同实根; 在 $v=0$ 或 $v=u^3$ 且 P 不是原点时有两个不同实根; 在其他情形只有一个不同实根. 当 $u<0$ 时同理, 函数图像与 x 轴之间的开区域对应三个不同实根, 边界对应两个, 其余对应一个. 当 $u=0$ 时, 方程为 $2t^3+v=0$, 只有一个不同实根.

对 $y=x^3$ 而言, 函数图像和拐点切线 $y=0$ 所夹的两个开区域正是上述三实根情形; 图像或拐点切线上非拐点的点对应两个不同实根; 其余开区域和拐点对应一个不同实根. 由坐标变换可逆, 一般三次函数情形随之成立. □

例题 3.38 已知函数 $f(x)=-x^3+2x^2-x$, 若过点 $P(1,t)$ 可作曲线 $y=f(x)$ 的三条切线, 求实数 t 的取值范围.

解 由切线条数判定定理, 可作三条切线等价于点 $P(1,t)$ 严格位于函数图像 $y=f(x)$ 与其拐点切线之间, 故 t 须介于横坐标 $x=1$ 处函数图像的纵坐标与拐点切线的纵坐标之间.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

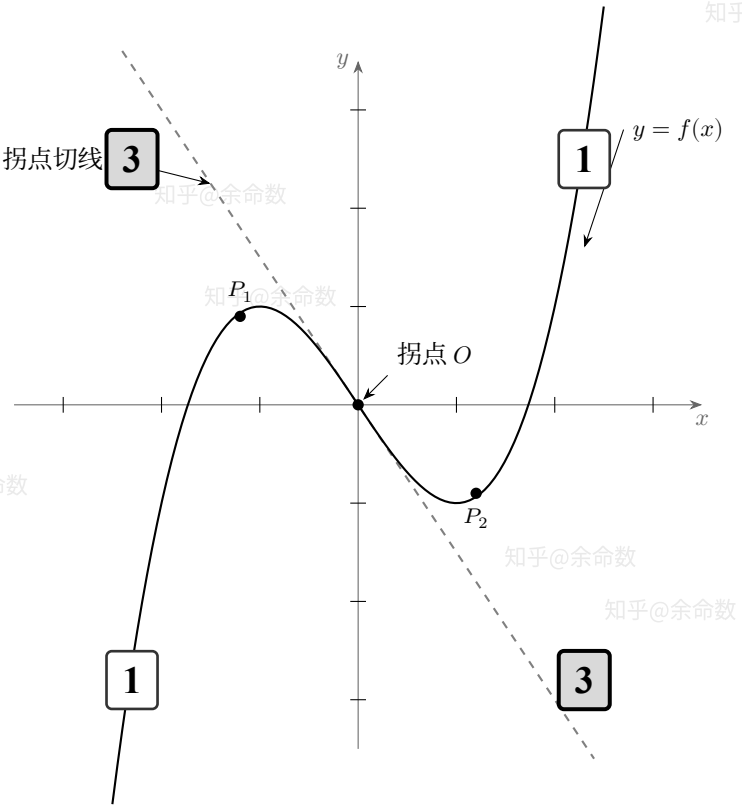


图 3.16 过平面一点作三次函数切线条数的区域划分 ($a > 0$)

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

函数在 $x=1$ 处的值为

$$f(1)=-(1)^3+2(1)^2-1=-1+2-1=0$$

再求拐点切线的方程. 由 $f'(x)=-3x^2+4x-1$ 得 $f''(x)=-6x+4$. 令 $f''(x)=0$, 解得拐点横坐标 $x_0=2/3$, 函数值 $f(2/3)=-(2/3)^3+2(2/3)^2-2/3=-8/27+8/9-2/3=-2/27$, 切线斜率 $k_0=f'(2/3)=-3(2/3)^2+4(2/3)-1=-4/3+8/3-1=1/3$.

由此拐点切线方程为 $y-(-2/27)=(1/3)(x-2/3)$, 化简得 $y=\frac{1}{3}x-\frac{8}{27}$. 在 $x=1$ 处, 切线上的纵坐标为

$$y_{\text{切线}}=\frac{1}{3}(1)-\frac{8}{27}=\frac{9-8}{27}=\frac{1}{27}$$

$P(1,t)$ 的纵坐标 t 须严格介于 $f(1)=0$ 与 $1/27$ 之间.

故实数 t 的取值范围是 $(0,1/27)$.

3.10.2.6 对称性

函数的对称性与其导函数的对称性之间存在着一种对偶关系.

定理 3.23 设可导函数 $y=f(x)$.

1. 若 $f(x)$ 的图像关于直线 $x=a$ 轴对称, 则其导函数 $y=f'(x)$ 的图像关于点 $(a,0)$ 中心对称.
2. 若 $f(x)$ 的图像关于点 (a,b) 中心对称, 则其导函数 $y=f'(x)$ 的图像关于直线 $x=a$ 轴对称.

证明 先证第一条. 函数 $f(x)$ 图像关于直线 $x=a$ 轴对称的代数表达式为 $f(a-t)=f(a+t)$. 对此恒等式两边关于 t 求导, 应用链式法则:

$$-f'(a-t)=f'(a+t),$$

即 $f'(a+t)+f'(a-t)=0$. 这正是导函数图像关于点 $(a,0)$ 中心对称的定义.

再证第二条. 函数 $f(x)$ 的图像关于点 (a,b) 对称, 等价于 $f(a+t)+f(a-t)=2b$. 这是关于变量 t 的恒等式. 我们把它两边关于 t 求导. 由链式法则得

$$\frac{d}{dt}[f(a+t)]+\frac{d}{dt}[f(a-t)]=\frac{d}{dt}[2b]$$

$$f'(a+t)\cdot(a+t)'_t+f'(a-t)\cdot(a-t)'_t=0$$

$$f'(a+t)\cdot 1+f'(a-t)\cdot(-1)=0$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由此得 $f'(a+t)=f'(a-t)$. 这正是导函数 $f'(x)$ 的图像关于直线 $x=a$ 轴对称的定义. □

我们回来看三次函数的对称性. 已经证明, 任何三次函数 $f(x)=ax^3+bx^2+cx+d$ 都关于其唯一的拐点 $(x_0,f(x_0))$ 中心对称, 其中 $x_0=-b/(3a)$. 由上述定理, 其导函数 $f'(x)=3ax^2+2bx+c$ 的图像关于直线 $x=x_0=-b/(3a)$ 轴对称. 这与二次函数的对称轴公式一致.

3.10.2.7 极值点连线的几何性质

当三次函数存在两个极值点时, 我们研究连接这两个点的割线斜率与拐点处切线斜率的关系.

定理 3.24 若三次函数 $f(x)=ax^3+bx^2+cx+d\ (a\neq 0)$ 存在两个极值点 $(x_1,f(x_1))$ 与 $(x_2,f(x_2))$, 则连接这两点的割线斜率为

$$\frac{f(x_1)-f(x_2)}{x_1-x_2}=-\frac{a}{2}(x_1-x_2)^2$$

并且该斜率与拐点处的切线斜率 $f'(x_0)$ 成正比,

$$\frac{f(x_1)-f(x_2)}{x_1-x_2}=\frac{2}{3}f'(x_0).$$

其中 $x_0=\frac{x_1+x_2}{2}$ 是函数拐点的横坐标.

证明 我们证明割线斜率与拐点处切线斜率之间的关系.

极值点 x_1,x_2 是导函数 $f'(x)=3ax^2+2bx+c=0$ 的两根, 由韦达定理 $x_1+x_2=-2b/(3a)$, $x_1x_2=c/(3a)$.

割线斜率为

$$\begin{aligned}\frac{f(x_1)-f(x_2)}{x_1-x_2}&=\frac{(ax_1^3+bx_1^2+cx_1+d)-(ax_2^3+bx_2^2+cx_2+d)}{x_1-x_2}\\&=a(x_1^2+x_1x_2+x_2^2)+b(x_1+x_2)+c\end{aligned}$$

将 $x_1^2+x_2^2=(x_1+x_2)^2-2x_1x_2=\frac{4b^2}{9a^2}-\frac{2c}{3a}$ 代入, 得

$$\begin{aligned}\text{斜率}&=a\left(\left(\frac{4b^2}{9a^2}-\frac{2c}{3a}\right)+\frac{c}{3a}\right)+b\left(-\frac{2b}{3a}\right)+c\\&=a\left(\frac{4b^2}{9a^2}-\frac{c}{3a}\right)-\frac{2b^2}{3a}+c\\&=\frac{4b^2}{9a}-\frac{c}{3}-\frac{6b^2}{9a}+c=-\frac{2b^2}{9a}+\frac{2c}{3}=\frac{-2b^2+6ac}{9a}=-\frac{2(b^2-3ac)}{9a}\end{aligned}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

同时，

$$\begin{aligned}-\frac{a}{2}(x_1-x_2)^2 &= -\frac{a}{2}((x_1+x_2)^2-4x_1x_2)\\ &= -\frac{a}{2}\left(\left(-\frac{2b}{3a}\right)^2-4\frac{c}{3a}\right)\\ &= -\frac{a}{2}\left(\frac{4b^2-12ac}{9a^2}\right)=-\frac{2(b^2-3ac)}{9a}.\end{aligned}$$

故第一个等式成立.

拐点横坐标 $x_0=-b/(3a)$，此处的导数值为

$$\begin{aligned}f'(x_0) &= 3a(x_0)^2+2bx_0+c=3a\left(-\frac{b}{3a}\right)^2+2b\left(-\frac{b}{3a}\right)+c\\ &= \frac{b^2}{3a}-\frac{2b^2}{3a}+c=-\frac{b^2}{3a}+c=\frac{-b^2+3ac}{3a}=-\frac{b^2-3ac}{3a}\end{aligned}$$

比较上述斜率与 $f'(x_0)$，得

$$\text{斜率}=-\frac{2(b^2-3ac)}{9a}=\frac{2}{3}\left(-\frac{b^2-3ac}{3a}\right)=\frac{2}{3}f'(x_0).$$

证毕. □

此定理给出如下几何事实. 连接三次函数两个极值点的直线的斜率，恰好是其拐点处切线斜率的 2/3. 拐点处的切线斜率是导函数 $f'(x)$ 的极值，对应原函数变化率最快或最慢的瞬间. 因此，这一关系把函数的宏观起伏，即由极值点决定的起伏，与其局部变化最剧烈处(拐点)联系起来.

3.11 指数函数

注 3.9
自然底 e 的极限定义与 $(e^x)'=e^x$ 见 第16章. 本节把 e 当作连续复利极限给出的常数(约 2.718)，并给出指数运算法则与图像.

在自然界与社会经济现象中，许多系统的增长或衰减过程具有一个共同特征. 其变化率与系统当前的总量成正比. 例如，在理想条件下，一个细胞群落的增殖速率与其现有的细胞数量成正比；放射性元素的衰变速率与其剩余的原子核数量成正比；一笔按复利计息的存款，其利息的增长速率与当前的本利和成正比.

为了精确地刻画这种增长模式，我们把幂运算推广到任意实指数，并由此定义指数函数.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

3.11.1 指数的扩张与指数函数的定义

我们已熟悉整数指数幂的运算, 例如 $a^3 = a \cdot a \cdot a$ 以及 $a^{-2} = 1/a^2$. 现在我们将指数的概念从整数域 $\mathbb{Z}$ 拓展到整个实数域 $\mathbb{R}$, 并确保每一步拓展都与原有运算律相容.

首先, 我们定义有理数指数幂. 对于任意正实数 a 和有理数 p/q ($p, q \in \mathbb{Z}, q > 0$), 我们定义

$$a^{p/q} = \sqrt[q]{a^p}$$

这一定义保持幂运算律 $(a^x)^y = a^{xy}$.

无理数指数幂的意义需要单独说明. 例如 2^π 不能理解为有限次乘法或开方. 其定义依赖实数系的**完备性**. 取一列有理数 $r_1, r_2, r_3, \ldots$ 逼近 π, 如 $3, 3.1, 3.14, \ldots$, 数列 $2^{r_1}, 2^{r_2}, 2^{r_3}, \ldots$ 会逼近一个确定的实数. 实数系的完备性保证该极限存在且唯一, 此极限值即定义为 2^π.

知乎@余命数

通过这种方式, 可将 a^x 的定义域从有理数集 $\mathbb{Q}$ 拓展至整个实数集 $\mathbb{R}$.

定义 3.12 (指数函数) 函数 $y = a^x$ ($a > 0$ 且 $a \neq 1$) 称为**指数函数**, 其中 x 是自变量. 其定义域为全体实数 $\mathbb{R}$.

底数 a 必须为正数, 否则函数在整个实数域上无定义, 例如, $(-2)^{1/2}$ 在实数域内无意义. 若 $a = 1$, 则 $y = 1^x = 1$ 是常数函数, 其性质退化, 通常不单独作为指数函数研究.

例题 3.39 (指数函数的定义辨析) 若函数 $y = (a^2 - 5a + 7)a^x$ 是指数函数, 求参数 a 的值.

解 要确定参数 a 的值, 我们回归指数函数的定义. 指数函数的标准形式为 $y = b^x$, 它满足两个约束.

1. 函数的系数必须为 1.
2. 底数 b 必须为正数且不等于 1.

将这些约束应用于 $y = (a^2 - 5a + 7)a^x$, 系数部分须为 1.

$$a^2 - 5a + 7 = 1$$

整理得

$$a^2 - 5a + 6 = 0$$

因式分解为 $(a - 2)(a - 3) = 0$, 解得 $a = 2$ 或 $a = 3$.

再检验这两个候选值是否满足 $a > 0$ 且 $a \neq 1$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 当 $a=2$ 时, $2>0$ 且 $2\neq 1$, 条件满足.
- 当 $a=3$ 时, $3>0$ 且 $3\neq 1$, 条件同样满足.

两个候选值均通过检验. 因此, 参数 a 的值为 2 或 3.

例题 3.40 (由点确定指数函数) 已知指数函数 $y=f(x)$ 的图像经过点 $(2,4)$, 求 $f(3)$ 的值.

解 由题设, $f(x)$ 是指数函数, 故可设 $f(x)=a^x$, 其中 $a>0$ 且 $a\neq 1$.

函数图像经过点 $(2,4)$, 即 $f(2)=4$. 于是

$$a^2=4$$

解得 $a=2$ 或 $a=-2$.

由指数函数定义, $a>0$, 故舍去 $a=-2$. 因此, 底数 a 被唯一确定为 2, 函数表达式为 $f(x)=2^x$.

于是

$$f(3)=2^3=8.$$

故 $f(3)$ 的值为 8.

3.11.2 指数函数的图像与性质

指数函数的性质由其底数 a 的取值范围决定. 我们分 $a>1$ 和 $0<a<1$ 两种情形来讨论.

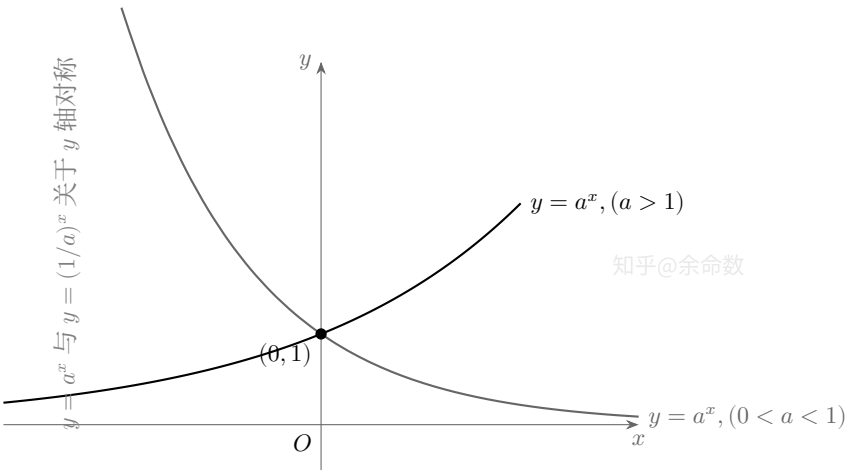


图 3.17　指数函数的图像

指数函数 $y=a^x$ $(a>0, a\neq 1)$ 的主要性质如下.

1. **定义域与值域**: 其定义域为 $\mathbb{R}$, 值域为 $(0,+\infty)$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

2. **特定点**：无论底数 a 为何值，函数图像恒过定点 $(0,1)$，因为 $a^0=1$.

3. **单调性**：

   - 当 $a>1$ 时，指数函数在 $\mathbb{R}$ 上是**严格单调递增**的.
   - 当 $0<a<1$ 时，指数函数在 $\mathbb{R}$ 上是**严格单调递减**的.

4. **渐近线**：x 轴是指数函数图像的水平渐近线.

我们以 $a>1$ 的情形为例，利用定义法严格证明其单调性.

证明 不妨设 $a>1$. 在定义域 $\mathbb{R}$ 中任取 x_1,x_2 且 $x_1<x_2$. 考察其函数值的比值

$$\frac{f(x_2)}{f(x_1)}=\frac{a^{x_2}}{a^{x_1}}=a^{x_2-x_1}.$$

知乎@余命数

由于 $x_1<x_2$，故 $x_2-x_1>0$. 当底数 $a>1$ 且指数为正数时，幂的值必然大于 1. 即 $a^{x_2-x_1}>1$. 因此，$\frac{f(x_2)}{f(x_1)}>1$. 又因为指数函数的值域为 $(0,+\infty)$，$f(x_1)$ 恒为正，故可得 $f(x_2)>f(x_1)$. 根据定义，函数 $f(x)=a^x$ 在 $a>1$ 时是严格单调递增的.$0<a<1$ 的情形可类似证明. □

例题 3.41 求解不等式 $4^{x^2+x}<(\frac{1}{2})^{x-5}$.

解 两端底数不同，不能直接比较指数. 我们先把它们化为同底数，再利用指数函数的单调性得到指数部分的不等式.

底数 4 和 $\frac{1}{2}$ 都是 2 的幂，因此统一为底数 2.

$$4=2^2,\quad \frac{1}{2}=2^{-1}$$

代入原不等式，得

$$(2^2)^{x^2+x}<(2^{-1})^{x-5}.$$

由幂的运算法则，上式等价于

$$2^{2(x^2+x)}<2^{-(x-5)}.$$

不等式两端是同一个指数函数 $f(t)=2^t$ 在两个不同点 $t_1=2(x^2+x)$ 和 $t_2=-(x-5)$ 的取值.

由于底数 $2>1$，指数函数 $f(t)=2^t$ 在 $\mathbb{R}$ 上严格单调递增. 因此，$f(t_1)<f(t_2)$ 等价于 $t_1<t_2$，即

$$2(x^2+x)<-(x-5).$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

整理得

$$2x^2+3x-5<0.$$

因式分解得 $(2x+5)(x-1)<0$.

故此不等式的解集为 $(-\frac{5}{2},1)$.

例题 3.42 求解方程 $4^x+6^x=2\cdot 9^x$.

知乎@余命数

解 方程 $4^x+6^x=2\cdot 9^x$ 含有三个不同底数. 这些底数之间存在代数联系, 它们都可由 2 和 3 的幂次组合而成, $4=2^2, 6=2\cdot 3, 9=3^2$. 因此我们用除法简化方程, 减少底数的种类.

方程两边同除以 9^x(它恒为正, 不改变解集), 得

知乎@余命数

$$\frac{4^x}{9^x}+\frac{6^x}{9^x}=2.$$

由指数运算法则 $(\frac{a}{b})^x=\frac{a^x}{b^x}$, 上式化为

$$\left(\frac{4}{9}\right)^x+\left(\frac{6}{9}\right)^x=2$$

$$\left(\left(\frac{2}{3}\right)^2\right)^x+\left(\frac{2}{3}\right)^x-2=0$$

知乎@余命数

$$\left(\frac{2}{3}\right)^{2x}+\left(\frac{2}{3}\right)^x-2=0.$$

令 $t=(\frac{2}{3})^x$. 由于指数函数的值域为正, 有 $t>0$. 上述方程化为

$$t^2+t-2=0$$

因式分解得 $(t+2)(t-1)=0$. 解得 $t=-2$ 或 $t=1$.

结合 $t>0$ 的约束, 我们舍去 $t=-2$. 因此, 唯一有效的解是 $t=1$.

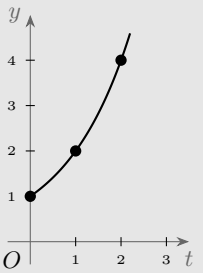
代回 $t=(\frac{2}{3})^x$, 得

$$\left(\frac{2}{3}\right)^x=1.$$

由任何非零实数的零次幂等于 1, 得 $x=0$. 因此, 原方程的解为 $x=0$.

例题 3.43 (指数增长模型的性质辨析) 某池塘中浮萍的面积 y(单位: m^2)与时间 t 的关系为指数函数模型 $y=f(t)=ka^t$ $(k>0,a>0,a\neq 1)$, 其中 t 以月为单位, 其图像如图所示. 试判断下列说法的正误.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



A. 浮萍每月增加的面积都相等.

B. 第 6 个月时, 浮萍的面积会超过 $30\,\mathrm{m}^2$.

C. 浮萍面积从 $2\,\mathrm{m}^2$ 蔓延到 $64\,\mathrm{m}^2$ 只需经过 5 个月.

D. 若浮萍面积蔓延到 $4\,\mathrm{m}^2, 6\,\mathrm{m}^2, 9\,\mathrm{m}^2$ 所经过的时间分别为 t_1, t_2, t_3, 则 $t_1 + t_3 = 2t_2$.

解 我们先从图像中读取参数. 图像经过点 $(0,1)$, 代入 $y = ka^t$ 得

$$1 = k \cdot a^0 \implies k = 1.$$

因此 $y = a^t$. 图像还经过点 $(1,2)$, 代入得 $2 = a$, 即 $a = 2$. 验证点 $(2,4)$. $f(2) = 2^2 = 4$, 与图像吻合. 因此, 浮萍面积的增长模型为 $y = f(t) = 2^t$.

我们逐一分析各个命题.

对于命题 A, 每月增加的面积为 $f(t+1) - f(t)$.

$$f(t+1) - f(t) = 2^{t+1} - 2^t = 2^t.$$

此增量为 2^t, 随 t 变化. 故 A 项错误.

对于命题 B, 第 6 个月末的面积为

$$f(6) = 2^6 = 64\,(\mathrm{m}^2).$$

由于 $64 > 30$, 故 B 项正确.

对于命题 C, 由 $2^t = 2$ 得 $t = 1$, 由 $2^t = 64 = 2^6$ 得 $t = 6$. 所需时间为 $\Delta t = 6 - 1 = 5$ 个月. 故 C 项正确.

对于命题 D, 我们检验 $t_1 + t_3 = 2t_2$ 是否成立. 由题意,

$$4 = 2^{t_1}, \quad 6 = 2^{t_2}, \quad 9 = 2^{t_3},$$

等式左侧为

$$t_1 + t_3 = \log_2 4 + \log_2 9 = \log_2(4 \cdot 9) = \log_2 36,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等式右侧为

$$2t_2 = 2\log_2 6 = \log_2(6^2) = \log_2 36.$$

二者相等, 故 D 项正确.

正确的说法为 B、C、D.

注 3.10 (等比与等差的对偶性)
选项 D 说明了一个性质. 若指数函数的因变量构成等比数列(本例中 $4, 6, 9$ 的公比为 $3/2$), 则其对应的自变量构成等差数列. 这是对数运算将乘法关系转化为加法关系的直接体现.

例题 3.44 (指数衰减模型的建立与求解) 某地区计划对总面积为 A 的老旧房屋进行平改坡工程. 经测算, 若改造模式为每年改造的面积是**当年剩余未改造面积**的一个固定百分比 p, 要在 10 年内完成工程总量的一半, 试估算这个百分比 p 的值.(参考数据: $(1/2)^{1/10} \approx 0.933$)

解 我们建立指数衰减模型描述剩余量. 设工程总面积为 A, t 年后剩余未改造的面积为 $R(t)$. 初始状态为 $R(0) = A$.

根据题意, 每年改造的面积是当年剩余面积的百分比 p. 因此第 $t+1$ 年末的剩余面积为

$$R(t+1) = R(t) - p \cdot R(t) = (1-p)R(t).$$

这是公比为 $r = 1-p$ 的递推关系, 其通项公式为 $R(t) = R(0) \cdot r^t = A(1-p)^t$.

由题意, 10 年后完成工程总量的一半, 即 $t = 10$ 时 $R(10) = A/2$. 代入模型, 得

$$A(1-p)^{10} = \frac{A}{2}.$$

消去 A, 得 $(1-p)^{10} = \frac{1}{2}$.

由此解得公比 $r = 1-p$ 为

$$1 - p = \left(\frac{1}{2}\right)^{\frac{1}{10}}.$$

利用题目给出的参考数据 $(1/2)^{1/10} \approx 0.933$, 得 $1 - p \approx 0.933$.

因此, 所求百分比 $p = 1 - \left(\frac{1}{2}\right)^{1/10} \approx 1 - 0.933 = 0.067$.

故每年约需改造当年剩余面积的 6.7%.

注 3.11 (指数模型与线性模型的对比)
这种按剩余量百分比进行改造的模式, 理论上永远无法完成 100% 的工程, 因为剩余量总是在不断减小但永不为零. 这与每年改造固定面积, 例如, 总面积的 5% 的线性模型不同. 在线性模型中, 工程将在 20 年后完成.

知乎@余命数

自然底 e 与 $(e^x)' = e^x$

知乎@余命数

3.11.3 自然底数 e 的引入

指数函数 $y = a^x$ 的底数 a 可以是任何不为 1 的正实数. 不同的底数决定函数增长或衰减的剧烈程度. 在所有可能的底数中, 确实存在一个最自然的底数, 它对应的指数函数具有独特的简洁性.

我们先看连续复利这个金融问题.

设想将一笔本金存入银行, 年利率为 r. 若每年计息一次, 一年后本利和为 $P_1 = P_0(1+r)$. 若每半年计息一次, 则每次计息的利率为 $r/2$, 一年内计息两次. 一年后本利和为 $P_2 = P_0(1+\frac{r}{2})^2$. 若每月计息一次, 则一年后本利和为 $P_{12} = P_0(1+\frac{r}{12})^{12}$.

知乎@余命数

一般地, 若一年内计息 n 次, 则每次的利率为 r/n, 一年后本利和为

$$P_n = P_0\left(1+\frac{r}{n}\right)^n.$$

计息周期无限缩短, 即计息频率 n 趋于无穷大时, 本利和会收敛到一个确定的极限值.

知乎@余命数

为了分析其结构, 我们设本金为 1 单位, 年利率为 100%, 即 $r = 1$, 以剥离无关参数. 此时, 一年后的本利和表达式为

$$\left(1+\frac{1}{n}\right)^n.$$

知乎@余命数

我们的问题转化为研究当 $n \to \infty$ 时, 此表达式的极限行为. 计算可知, 随着 n 的增大, 这个值确实在增长, 但增长的步伐越来越慢, 并稳定地趋向于一个特定的数值.

- $n = 1$∶$(1+1/1)^1 = 2$
- $n = 10$∶$(1+1/10)^{10} \approx 2.5937$
- $n = 100$∶$(1+1/100)^{100} \approx 2.7048$
- $n = 1000$∶$(1+1/1000)^{1000} \approx 2.7169$

这个极限存在且唯一. 其严格证明需要分析工具, 见 第16章. 这个极限值是一个无理数, 用字母 e 表示.

定义 3.13 (自然底数 e) 自然底数 e 定义为如下极限.

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \approx 2.71828\ldots$$

实际上, 对于任意年利率 r, 连续复利下的本利和极限为 P_0e^r. 因此, e 是单位利率下单位本金连续增长的极限倍率.

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

连续复利模型给出了 e 的一种来源. 它的另一来源在于分析性质. 我们回到最初提出的问题, 在所有的指数函数 $y=a^x$ 中找出导数与其函数值本身关系最为简洁的那一个. 在微积分中, 自然就是简洁.

可以证明, 指数函数 $f(x)=a^x$ 的导数与其自身成正比, 即 $(a^x)'=(\ln a)\cdot a^x$, 其中比例常数 $\ln a$ 完全由底数 a 决定. 我们要找一个底数 a, 使这个比例常数恰好为 1, 函数的变化率在每一点都精确地等于其函数值.

这个特殊的底数正是自然底数 e.

定理 3.25 (自然指数函数的基本性质) 在全体可导函数中, 以 e 为底的指数函数 $f(x)=e^x$, 称为**自然指数函数**是唯一满足 $f(0)=1$ 且 $f'(x)=f(x)$ 的函数. 它满足

$$(e^x)'=e^x.$$

证明 先证存在性. 由自然底数 e 的定义以及指数函数的导数公式可得 $(e^x)'=e^x$, 且 $e^0=1$, 因而 e^x 满足题设条件.

再证唯一性. 设可导函数 $g(x)$ 满足 $g'(x)=g(x)$ 且 $g(0)=1$. 构造

$$h(x)=g(x)e^{-x}.$$

由乘积求导法则,

$$h'(x)=g'(x)e^{-x}-g(x)e^{-x}=(g'(x)-g(x))e^{-x}=0.$$

因此 $h(x)$ 在 $\mathbb{R}$ 上为常数. 又 $h(0)=g(0)e^0=1$, 故 $h(x)\equiv 1$, 即 $g(x)=e^x$. 于是满足条件的可导函数唯一. □

这个性质使得 e^x 适合描述增长率正比于存量的过程. 因此, e 是最自然的底数.

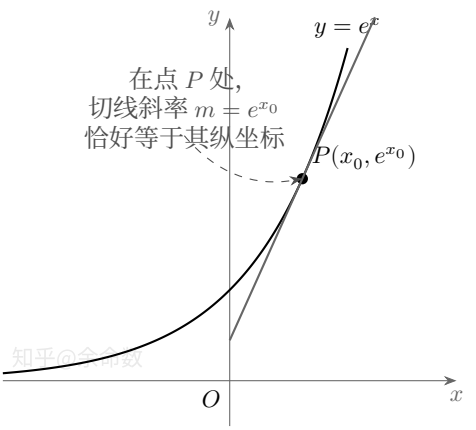


图 3.18　自然指数函数的几何意义. 图像上任意一点的切线斜率等于该点的函数值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.12 幂函数

我们已经研究了变量位于指数位置的函数 $y = a^x$, 即指数函数. 现在, 我们把注意力转向另一类基本函数, 其中变量位于底数位置.

3.12.1 定义与辨析

定义 3.14 (幂函数) 形如 $y = x^\alpha$ 的函数称为**幂函数**, 其中 x 是自变量, $\alpha \in \mathbb{R}$ 是一个常数, 称为指数.

我们先用一个例子区分幂函数与指数函数. 两族函数的表达式都出现幂运算, 但变量所处的位置不同. 幂函数 $y = x^\alpha$ 以 x 为底数、α 为常数, 指数函数 $y = a^x$ 则以 x 为指数、a 为常数. 这种差异使它们的定义域、图像形态和单调性都不一样.

- **幂函数** $y = x^\alpha$: 自变量 x 是**底数**, 指数 α 是**常数**.
- **指数函数** $y = a^x$: 自变量 x 是**指数**, 底数 a 是**常数**.

例如, $y = x^2$ 与 $y = 2^x$ 在 $x > 0$ 时都是增函数, 但增长速度完全不同. 当 x 足够大时, 2^x 会远远超过 x^2, 如 图 3.19 所示.

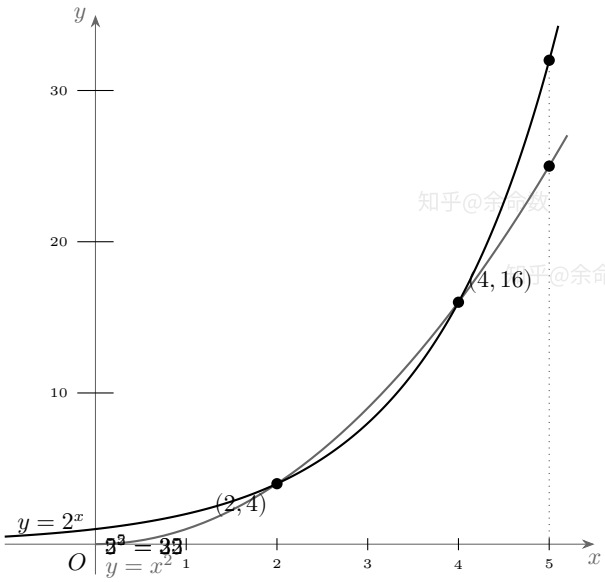


图 3.19 幂函数与指数函数的增长比较. 在 $x > 4$ 后, 指数函数的增长速度远超幂函数.

3.12.2 图像与性质分析

幂函数 $y = x^\alpha$ 的性质由指数 α 决定. 我们取几个典型值, 归纳其图像与性质的异同.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

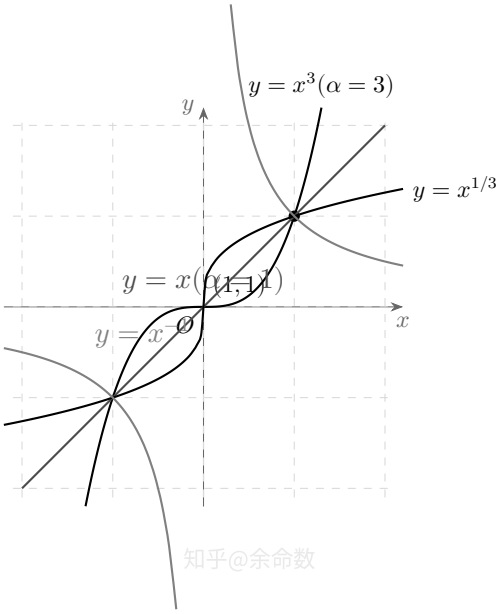


图 3.20　不同指数 α 下幂函数的图像

从图像与代数两方面看, 我们得到幂函数 $y = x^\alpha$ 的一般性质.

1. **定义域**. 定义域依赖于 α. 若 α 为正整数, 定义域为 $\mathbb{R}$. 若 α 为负整数, 定义域为 $\mathbb{R} \setminus \{0\}$. 若 α 为分数, 则需根据分母的奇偶性确定. 例如 $y = x^{1/2}$ 的定义域为 $[0, +\infty)$.
2. **奇偶性**. 奇偶性同样取决于 α. 若 $f(x) = x^\alpha$ 的定义域关于原点对称, 则当 α 是整数时, 其奇偶性与 α 的奇偶性一致. 若 $\alpha = p/q$(最简分数), 则当 q 为奇数时, 其奇偶性与分子 p 的奇偶性一致.
3. **公共点**. 无论 α 为何值 (除 $\alpha = 0$ 外), 幂函数的图像恒过定点 $(1, 1)$, 因为 $1^\alpha = 1$.
4. **单调性**(在第一象限, 即 $x > 0$ 时):
   - 当 $\alpha > 0$ 时, 函数在 $(0, +\infty)$ 上**严格单调递增**.
   - 当 $\alpha < 0$ 时, 函数在 $(0, +\infty)$ 上**严格单调递减**.
5. **凹凸性**(在第一象限, 即 $x > 0$ 时):
   - 当 $\alpha > 1$ 或 $\alpha < 0$ 时, 图像是**下凸**的(向上弯曲).
   - 当 $0 < \alpha < 1$ 时, 图像是**上凸**的(向下弯曲).

这些性质用于比较大小与求解不等式.

例题 3.45　比较 $0.7^{0.8}, 0.8^{0.7}, 0.8^{0.8}$ 的大小.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 三个数底数、指数都不同, 无法直接比较. 我们用中间量把它们拆成可以直接用单调性处理的比较.

先把 $0.8^{0.7}$ 与 $0.8^{0.8}$ 比较. 这可以视为指数函数 $f(x)=0.8^x$ 在两个不同点上的取值. 由于底数 $a=0.8\in(0,1)$, 此指数函数在 $\mathbb{R}$ 上严格单调递减. 因为 $0.7<0.8$, 由单调递减的定义得 $f(0.7)>f(0.8)$, 即 $0.8^{0.7}>0.8^{0.8}$.

再把 $0.7^{0.8}$ 与 $0.8^{0.8}$ 比较. 这可以视为幂函数 $g(x)=x^{0.8}$ 在两个不同点上的取值. 由于指数 $\alpha=0.8>0$, 此幂函数在定义域 $(0,+\infty)$ 上严格单调递增. 因为 $0.7<0.8$, 由单调递增的定义得 $g(0.7)<g(0.8)$, 即 $0.7^{0.8}<0.8^{0.8}$.

故三个数的大小关系为

$$0.7^{0.8}<0.8^{0.8}<0.8^{0.7}.$$

例题 3.46 已知函数 $f(x)=(m^2-m-1)x^{m-2}$ 是一个幂函数, 且在区间 $(0,+\infty)$ 上是减函数, 求实数 m 的值.

知乎@余命数

解 我们先从幂函数的定义提取对 m 的约束, 再用单调性筛选候选值.

根据幂函数的定义, 函数的标准形式为 $y=x^{\alpha}$, 因此自变量 x 前的系数必须为 1. 于是系数必须满足

$$m^2-m-1=1$$

整理得 $m^2-m-2=0$ 因式分解为 $(m-2)(m+1)=0$, 解得 $m=2$ 或 $m=-1$.

再用单调性筛选这两个候选值. 题设函数在区间 $(0,+\infty)$ 上是减函数. 根据幂函数的性质, 当 $x>0$ 时, $y=x^{\alpha}$ 单调递减的充要条件是其指数 $\alpha<0$. 在本题中, 指数为 $m-2$, 故

$$m-2<0 \implies m<2$$

将两个条件得到的解集求交集. 第一个条件给出 $m\in\{2,-1\}$; 第二个条件要求 $m<2$. 唯一同时满足这两个条件的 m 值是 $m=-1$.

故实数 m 的值为 -1.

3.13 对数函数

指数函数 $y=a^x$ 把自变量 x, 也就是指数, 映为函数值 y, 也就是幂. 这一对应可以逆转, 已知 y 与底数 a 就能唯一确定指数 x.

例如, $2^x=8$ 的解是 $x=3$. 但 $2^x=5$ 的解不是整数. 从 $y=2^x$ 严格单调递增的图像看, 存在唯一实数 x 使 $2^x=5$. 为表示这个数, 我们引入对数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.13.1 对数的定义与对数函数

定义 3.15 (对数) 若 $a^x = N$ $(a > 0, a \neq 1)$, 则数 x 称为以 a 为底 N 的**对数**, 记作

$$x = \log_a N$$

其中, a 称为对数的**底数**, N 称为**真数**.

这个定义表明对数与指数互为逆运算, 即 $a^x = N \iff x = \log_a N$. 于是 $\log_a N$ 就是满足 $a^x = N$ 的那个指数 x.

由此可建立对数函数.

定义 3.16 (对数函数) 函数 $y = \log_a x$ $(a > 0, a \neq 1)$ 称为**对数函数**, 其中 x 是自变量.

知乎@余命

对数函数的底数与定义域承袭自指数函数.

- **底数约束 ($a > 0, a \neq 1$)**. 与指数函数相同, 要求 $a > 0$ 且 $a \neq 1$.
- **定义域(真数约束 $x > 0$)**. 对数函数的自变量 x 是指数函数 $y = a^x$ 的因变量. 由于指数函数的值域为 $(0, +\infty)$, 对数函数的定义域也必须是 $(0, +\infty)$.

知乎@

3.13.2 对数函数的图像与性质

知乎@余命数

对数函数 $y = \log_a x$ 是指数函数 $y = a^x$ 的反函数. 因此它们的图像关于直线 $y = x$ 对称. 其余代数性质都由此导出.

我们列出对数函数 $y = \log_a x$ $(a > 0, a \neq 1)$ 的主要性质.

1. **定义域与值域**. 其定义域为 $(0, +\infty)$, 值域为 $\mathbb{R}$. 二者恰好是指数函数定义域与值域的互换.
2. **特定点**. 无论底数 a 为何值, 函数图像恒过定点 $(1, 0)$, 因为 $\log_a 1 = 0$.
3. **单调性**:
   - 当 $a > 1$ 时, 对数函数在 $(0, +\infty)$ 上是**严格单调递增**的.
   - 当 $0 < a < 1$ 时, 对数函数在 $(0, +\infty)$ 上是**严格单调递减**的.
4. **渐近线**. y 轴, 即直线 $x = 0$ 是对数函数图像的垂直渐近线.

知乎@余命数

例题 3.47 (对数与指数的复合应用) 已知对数函数 $f(x)$ 的图像与一次函数 $h(x) = \frac{1}{3}x - \frac{1}{3}$ 的图像交于 A, B 两点, 且点 B 的横坐标为 4.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

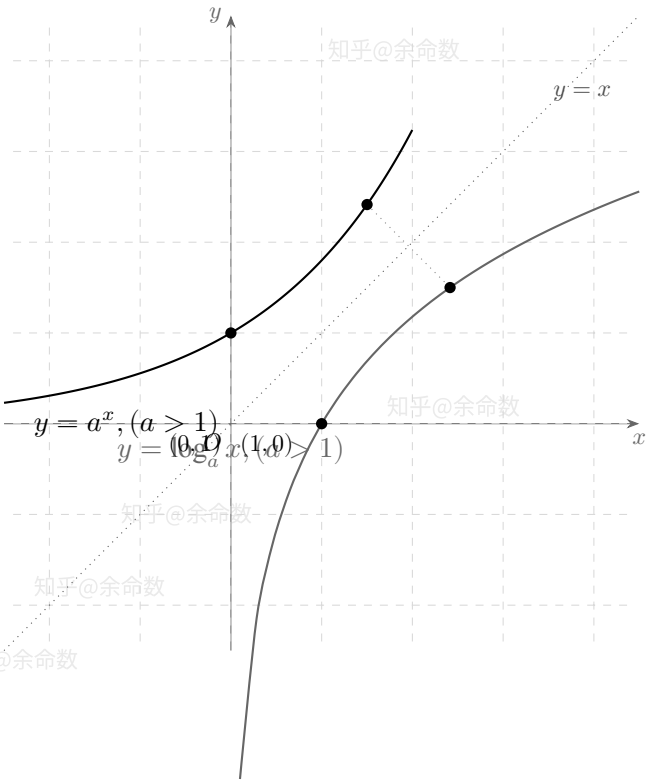
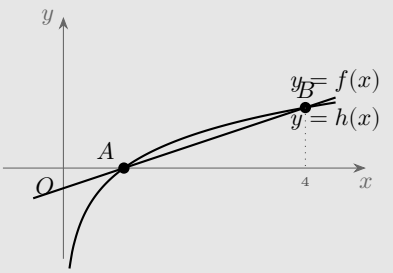


图 3.21 指数函数与对数函数图像的对称性 ($a > 1$)

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



1. 求 $f(x)$ 的解析式.
2. 若关于 x 的不等式 $4^{f(x)}<k$ 恰有 1 个整数解, 求实数 k 的取值范围.

知乎@余命数

解 先确定 $f(x)$ 的解析式. 设 $f(x)=\log_a x$, 其中 $a>0$ 且 $a\neq 1$.

点 B 是两个函数图像的公共点, 其坐标必须同时满足两个函数的解析式. 已知点 B 的横坐标为 4, 可计算其纵坐标, 得

$$h(4)=\frac{1}{3}(4)-\frac{1}{3}=1$$

因此, 点 B 的坐标为 $(4,1)$.

将点 $B(4,1)$ 的坐标代入 $f(x)=\log_a x$ 中, 得关于底数 a 的方程

$$f(4)=\log_a 4=1$$

根据对数的定义, 此式等价于 $a^1=4$, 故 $a=4$. 因此, 函数 $f(x)$ 的解析式为 $f(x)=\log_4 x$.

再分析不等式 $4^{f(x)}<k$. 把 $f(x)=\log_4 x$ 代入, 左侧化简为

$$4^{\log_4 x}<k$$

由对数与指数的互逆关系, $4^{\log_4 x}=x$, 不等式化为

$$x<k$$

讨论必须在 $f(x)$ 的定义域 $(0,+\infty)$ 内进行.

因此, 不等式的完整解集是 $0<x<k$. 我们要找参数 k 的取值, 使得开区间 $(0,k)$ 内恰好包含一个整数.

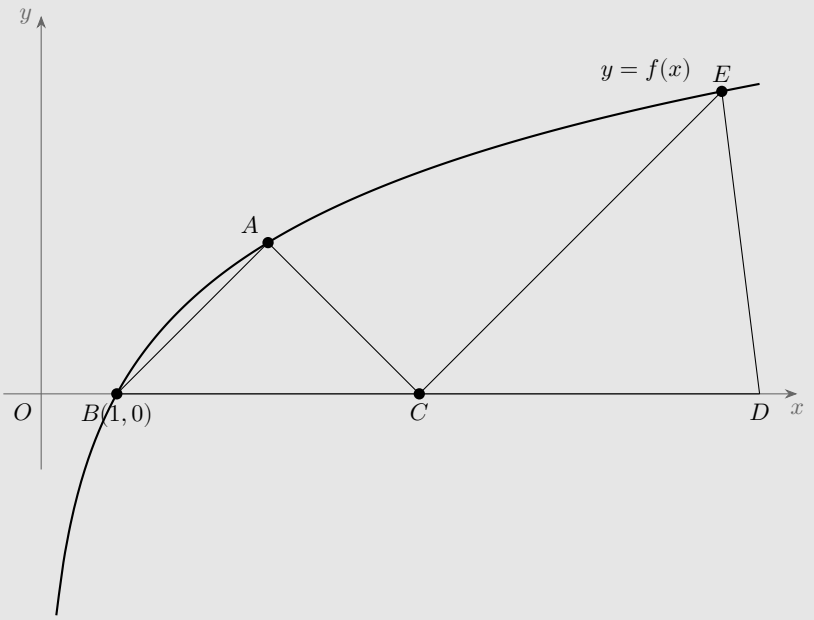
区间 $(0,k)$ 内的整数从小到大依次为 $1,2,3,\ldots$. 要使该区间内恰好包含一个整数, 这个整数只能是 1.

为了让 1 成为解集的一部分, 必须有 $1<k$. 为了让 2 不成为解集的一部分, 必须有 $k\leq 2$.

综合这两个条件, 得 $1<k\leq 2$. 故实数 k 的取值范围是 $(1,2]$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 3.48 (对数函数图像的几何应用) 如图, 对数函数 $f(x)=\log_a x\ (a>1)$ 图像上的点 A 与 x 轴上的点 $B(1,0)$ 和点 C 构成以 BC 为斜边的等腰直角三角形. 若 $\triangle ECD$ 与 $\triangle ABC$ 相似, 点 E 在函数 $f(x)$ 的图像上, 点 D 位于点 C 的右侧, 且两个三角形的相似比为 $2:1$, 求底数 a 的值.



解 此问题需将几何条件转化为点坐标之间的代数关系.

设点 A 的坐标为 (x_1,y_1). 因为 $a>1$, 函数 $f(x)=\log_a x$ 在 $(0,+\infty)$ 上严格单调递增, 所以 $y_1>0$. 又因为 $\triangle ABC$ 以 BC 为斜边且为等腰直角三角形, 点 A 在斜边上方, 故 A 到 BC 的垂线长度等于 BC 的一半, 垂足为 BC 的中点.

点 A 的纵坐标 y_1 即为三角形的高. 设 C 点坐标为 $(c,0)$, 则斜边长为 $c-1$. 得关系 $y_1=\frac{1}{2}(c-1)$. 同时, 点 A 的横坐标 x_1 必为 BC 的中点横坐标, 即 $x_1=\frac{1+c}{2}$.

从这两个关于 x_1,y_1,c 的关系式中, 可消去 c. 由 $c=2x_1-1$ 代入前式, 得 $y_1=\frac{1}{2}((2x_1-1)-1)=x_1-1$. 图像上的点 $A(x_1,y_1)$ 必须满足线性关系 $y_1=x_1-1$.

接下来, 引入 $\triangle ECD$ 的信息. 设点 E 的坐标为 (x_2,y_2). 由于 $\triangle ECD\sim\triangle ABC$ 且相似比为 2, 其对应高的比也为 2.$\triangle ECD$ 的高即为点 E 的纵坐标 y_2. 因此, $y_2=2y_1$.

因为点 $A(x_1,y_1)$ 和 $E(x_2,y_2)$ 均在函数 $f(x)=\log_a x$ 的图像上, 有

$$y_1=\log_a x_1,\quad y_2=\log_a x_2$$

结合 $y_2=2y_1$, 得 $\log_a x_2=2\log_a x_1=\log_a(x_1^2)$. 由于对数函数是单射, 此式蕴含了 $x_2=x_1^2$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

还需要一个关于 x_1, x_2 的关系. 同样利用 $\triangle ECD$ 的几何性质, 其底边 CD 的长度是 BC 的两倍, 即 $CD = 2BC = 2(c-1) = 4y_1$. 点 E 的横坐标 x_2 是 CD 的中点横坐标.

知乎@余命数

$$x_2 = c + \frac{1}{2}CD = c + 2y_1$$

现在拥有一个关于 x_1, y_1, x_2 的方程组

$$\begin{cases} y_1 = x_1 - 1 \\ x_2 = x_1^2 \\ x_2 = c + 2y_1 = (2x_1 - 1) + 2y_1 \end{cases}$$

知乎@余命数　知乎@余命数

将第一、二个式子代入第三个式子

$$x_1^2 = (2x_1 - 1) + 2(x_1 - 1) = 4x_1 - 3$$

整理得到一个关于 x_1 的一元二次方程

$$x_1^2 - 4x_1 + 3 = 0$$

因式分解得 $(x_1 - 1)(x_1 - 3) = 0$, 解得 $x_1 = 1$ 或 $x_1 = 3$.

若 $x_1 = 1$, 则 $y_1 = \log_a 1 = 0$, 这将导致 $\triangle ABC$ 退化为一个线段, 不合题意, 故舍去. 因此, $x_1 = 3$.

知乎@余命数　知乎@余命数

当 $x_1 = 3$ 时, $y_1 = x_1 - 1 = 2$. 故点 A 的坐标为 $(3, 2)$.

最后, 将点 $A(3, 2)$ 的坐标代入函数解析式以求解底数 a

$$2 = \log_a 3$$

根据对数的定义, $a^2 = 3$. 由于题设 $a > 1$, 得 $a = \sqrt{3}$.

3.13.3 对数的运算法则

对数的运算法则可直接从指数运算法则导出. 我们把指数式改写成对数语言, 即可逐条得到.

定理 3.26 (对数运算法则) 设 $a > 0, a \neq 1$, 且 $M > 0, N > 0$.

1. $\log_a(MN) = \log_a M + \log_a N$(积的对数等于对数的和).
2. $\log_a(M/N) = \log_a M - \log_a N$(商的对数等于对数的差).
3. $\log_a(M^n) = n \log_a M$ $(n \in \mathbb{R})$(幂的对数等于指数乘以底的对数).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 设 $u=\log_a M, v=\log_a N$. 根据对数定义, $M=a^u, N=a^v$.

对于积的法则

$$MN=a^u a^v=a^{u+v},$$

故 $\log_a(MN)=u+v=\log_a M+\log_a N$.

对于商的法则, 由于 $N>0$, 有

$$\frac{M}{N}=\frac{a^u}{a^v}=a^{u-v},$$

故 $\log_a(M/N)=u-v=\log_a M-\log_a N$.

对于幂的法则, 任取实数 n. 由实指数幂的运算法则

$$M^n=(a^u)^n=a^{un},$$

故 $\log_a(M^n)=un=n\log_a M$. 三条公式均得证. □

实际计算中常需要统一不同对数的底数. 换底公式正是这个工具.

定理 3.27 (换底公式) 设 $a,b>0$ 且 $a,b\neq 1, N>0$. 则

$$\log_a N=\frac{\log_b N}{\log_b a}$$

证明 设 $\log_a N=x$. 则 $a^x=N$. 对此指数式两边取以 b 为底的对数, 得

$$\log_b(a^x)=\log_b N$$

应用幂的运算法则得 $x\log_b a=\log_b N$. 由于 $a\neq 1, \log_b a\neq 0$, 故可解得 $x=\frac{\log_b N}{\log_b a}$. 将 $x=\log_a N$ 代回, 即得公式. □

换底公式有两个很有用的推论.

- $\log_a b=\frac{1}{\log_b a}$(通过令 $N=b$ 得到).
- $\log_{a^m} b^n=\frac{n}{m}\log_a b$(通过换底到以 a 为底的对数得到).

例题 3.49 求解不等式 $\log_{0.5}(x^2-2x-3)>\log_{0.5}(x-1)$.

解 由于对数的真数必须为正, 我们先列定义域约束

$$\begin{cases} x^2-2x-3>0 \\ x-1>0 \end{cases}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第一个不等式 $(x-3)(x+1)>0$ 的解集为 $(-\infty,-1)\cup(3,+\infty)$. 第二个不等式 $x>1$ 的解集为 $(1,+\infty)$. 两个解集的交集为 $(3,+\infty)$. 这是考虑原不等式时自变量 x 必须满足的范围.

在定义域 $x\in(3,+\infty)$ 内求解原不等式. 不等式两端是同一个对数函数 $f(t)=\log_{0.5}t$ 在两个不同点 $t_1=x^2-2x-3$ 和 $t_2=x-1$ 的取值.

由于底数 $a=0.5\in(0,1)$, 对数函数 $f(t)$ 在其定义域上是严格单调递减的. 因此, $f(t_1)>f(t_2)$ 等价于 $t_1<t_2$

$$x^2-2x-3<x-1$$

整理此一元二次不等式, 得

$$x^2-3x-2<0$$

方程 $x^2-3x-2=0$ 的根为 $x=\frac{3\pm\sqrt{9-4(1)(-2)}}{2}=\frac{3\pm\sqrt{17}}{2}$. 故不等式 $x^2-3x-2<0$ 的解集为 $(\frac{3-\sqrt{17}}{2},\frac{3+\sqrt{17}}{2})$.

最后, 须将此解集与先前确定的函数定义域 $(3,+\infty)$ 求交集. 注意到 $4<\sqrt{17}<5$, 因此 $\frac{3+4}{2}<\frac{3+\sqrt{17}}{2}<\frac{3+5}{2}$, 即 $3.5<\frac{3+\sqrt{17}}{2}<4$. 同时, $3=\frac{6}{2}=\frac{\sqrt{36}}{2}<\frac{3+\sqrt{17}}{2}$.

因此, 原不等式的解集为 $(3,\frac{3+\sqrt{17}}{2})$.

例题 3.50　已知函数 $f(x)=\log_a\frac{1+x}{1-x}$ $(a>0,a\neq1)$.

1. 求 $f(x)$ 的定义域.
2. 判断 $f(x)$ 的奇偶性.

解　函数的定义域要求其真数为正, 即

$$\frac{1+x}{1-x}>0$$

此分式不等式等价于 $(1+x)(1-x)>0$, 即 $(x+1)(x-1)<0$. 解得 $-1<x<1$. 故 $f(x)$ 的定义域为开区间 $(-1,1)$.

再判断奇偶性. 定义域 $(-1,1)$ 关于原点对称, 因此可以讨论奇偶性. 由

$$f(-x)=\log_a\frac{1+(-x)}{1-(-x)}=\log_a\frac{1-x}{1+x}$$

为了与 $f(x)$ 建立联系, 我们利用对数运算法则. 注意到 $\frac{1-x}{1+x}=\left(\frac{1+x}{1-x}\right)^{-1}$. 因此

$$\begin{aligned}f(-x)&=\log_a\left(\left(\frac{1+x}{1-x}\right)^{-1}\right)\\&=-1\cdot\log_a\frac{1+x}{1-x}\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= -f(x)$$

此关系对定义域内的所有 x 成立, 故 $f(x)$ 是奇函数.

注 3.12
更一般地, 若 $g(-x) = 1/g(x)$, 则函数 $f(x) = \log_a g(x)$ 必为奇函数.

3.14 函数模型与应用

数学不仅研究抽象结构, 也可用来描述自然、社会、经济中的数量关系. 把一个实际问题抽象为数学结构的过程称为**数学建模**. 建模时, 我们根据问题的内在规律与经验数据, 从已知函数类型中选择候选模型, 估算参数, 并通过比较确定与数据最吻合的表达式.

一个完整的建模过程通常包括以下几个环节.

- **数据分析与模型初选**. 审视经验数据的变化趋势是增长还是衰减、变化速率是加速、匀速还是减速等特征, 据此筛选出具有相似理论行为的函数类型作为候选模型.
- **参数估计**. 利用数据中的若干关键信息 (通常是特定的数据点), 建立关于模型待定参数的方程组, 并求解这些参数.
- **模型检验与优化**. 将拟合出的具体函数模型用于预测其他数据点, 并比较预测值与真实值的差异, 即误差, 以评估模型的优劣. 若误差过大, 则可能需要更换模型或修正参数.

知乎@余命数

以下通过实例说明.

例题 3.51 (函数模型的选择与拟合) 某养殖场随着技术的进步和规模的扩大, 肉鸡产量在不断增加. 现收集到 2020 年前 10 个月该养殖场上市的肉鸡数量 W(单位:万只)与月份 m 的数据如下表:

月份 m	1	2	3	4	5
数量 W	1.020	2.000	2.578	2.997	3.313
月份 m	6	7	8	9	10
数量 W	3.578	3.804	4.000	4.173	4.329

数量 W 和月份 m 之间可能存在以下四种函数关系:

① $W(m) = b \cdot a^m$

② $W(m) = b \cdot m^c$

③ $W(m) = b + \log_a m$

④ $W(m) = a - \frac{b}{m}$ $(a > 0, b > 0)$

1. 请从这四个函数模型中去掉一个与表格中数据不吻合的函数模型, 并说明理由.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. 请从表格中选择 2 月份和 8 月份的数据, 再从第 (1) 问剩下的三个模型中任选两个函数模型进行建模, 求出其函数表达式, 再分别求出这两个模型下 4 月份的肉鸡数量, 并说明哪个函数模型更好. (参考数据: $\sqrt[3]{16} \approx 2.519, \sqrt{2} \approx 1.414$)

解 (1) 先分析数据趋势. 由表格可知, W 随 m 增加而增加, 但增长的速率在变化. 相邻月份的增量如下

$$W(2) - W(1) \approx 0.980$$

$$W(3) - W(2) \approx 0.578$$

$$W(4) - W(3) \approx 0.419$$

$$\cdots$$

$$W(10) - W(9) \approx 0.156$$

每月增加的数量递减, 因此这是减速增长.

再分析四个候选模型的理论行为.

- 模型 ① $W(m) = b \cdot a^m$ 是一个指数增长模型. 为保证增长, 必须有 $a > 1$. 在此条件下, 其增量为 $W(m+1) - W(m) = b \cdot a^m(a-1)$, 这是一个随 m 增加而严格递增的量. 故指数模型描述的是加速增长过程.
- 模型 ② $W(m) = b \cdot m^c$ 是一个幂函数模型. 为保证增长, 需 $b > 0, c > 0$. 若 $0 < c < 1$, 其增长率是递减的; 若 $c > 1$, 其增长率是递增的.
- 模型 ③ $W(m) = b + \log_a m$ 是一个对数函数模型. 为保证增长, 需 $a > 1$. 对数函数的增长率是递减的.
- 模型 ④ $W(m) = a - b/m$ 是一个反比例函数平移后的模型. 其增长率也是递减的.

比较二者可知, 模型 ① 描述加速增长, 与数据的减速增长矛盾. 因此应去掉模型 ①.

(2) 选取 2 月份的数据 $(2, 2.000)$ 和 8 月份的数据 $(8, 4.000)$. 我们选择模型 ② 和 ③ 进行拟合.

对模型 ②. 设 $W(m) = b \cdot m^c$.

将数据点 $(2, 2)$ 和 $(8, 4)$ 代入, 得到关于参数 b, c 的方程组

$$\begin{cases} 2 = b \cdot 2^c \\ 4 = b \cdot 8^c \end{cases}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

两式相除以消去 b, 得

$$\frac{4}{2}=\frac{b\cdot 8^c}{b\cdot 2^c}\implies 2=\left(\frac{8}{2}\right)^c=4^c$$

解得 $c=1/2$. 将 $c=1/2$ 代回第一个方程得 $2=b\cdot 2^{1/2}=b\sqrt{2}$, 解得 $b=\sqrt{2}$. 因此, 幂函数模型为 $W_2(m)=\sqrt{2}\cdot m^{1/2}=\sqrt{2m}$.

利用此模型预测 4 月份的肉鸡数量

$$W_2(4)=\sqrt{2\cdot 4}=\sqrt{8}=2\sqrt{2}\approx 2\times 1.414=2.828(\text{万只})$$

对模型 ③. 设 $W(m)=b+\log_a m$.

将数据点 $(2,2)$ 和 $(8,4)$ 代入, 得到关于参数 a,b 的方程组

$$\begin{cases}2=b+\log_a 2\\ 4=b+\log_a 8\end{cases}$$

两式相减, 以消去参数 b

$$4-2=(\log_a 8)-(\log_a 2)\implies 2=\log_a(8/2)=\log_a 4$$

根据对数的定义, $a^2=4$. 由于底数 $a>0,a\neq 1$, 故 $a=2$. 将 $a=2$ 代回第一个方程得 $2=b+\log_2 2=b+1$, 解得 $b=1$. 因此, 对数函数模型为 $W_3(m)=1+\log_2 m$.

利用此模型预测 4 月份的肉鸡数量

$$W_3(4)=1+\log_2 4=1+2=3.000(\text{万只})$$

比较两模型.

把两个模型的预测值与表格中 $W(4)=2.997$ 比较.

模型 ② 的预测误差为 $|2.828-2.997|=0.169$. 模型 ③ 的预测误差为 $|3.000-2.997|=0.003$.

对数函数模型的误差远小于幂函数模型, 因此在本题中模型 ③ 更好.

例题 3.52 人耳对响度的感知与声强 I 的对数有关, 其中 I 的单位是 W/m^2. 科学家用声强级 Y(单位分贝)刻画响度, 公式为

$$Y=10\lg\left(\frac{I}{10^{-12}}\right)$$

其中 $10^{-12}\ \text{W/m}^2$ 是人耳能听到的阈值声强.

1. 平常人交谈时的声强约为 $10^{-6}\ \text{W/m}^2$, 求其声强级.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. 一般常人能听到的最低声强级是 0 分贝, 求能听到的最低声强为多少?
3. 比较理想的睡眠环境要求声强级 $Y \le 50$ 分贝. 已知熄灯后两位同学在宿舍说话的声强为 $5\times10^{-7}\,\mathrm{W/m^2}$, 问这两位同学是否会影响其他同学休息?(参考数据:$\lg 2 \approx 0.301$)

解 声强级 Y 与声强 I 由对数模型联系.

(1) 平常交谈 $I = 10^{-6}\,\mathrm{W/m^2}$

$$Y = 10\lg\left(\frac{10^{-6}}{10^{-12}}\right) = 10\lg(10^6) = 60(\text{分贝}).$$

(2) 最低可听对应 $Y = 0$. 由 $10\lg(\frac{I}{10^{-12}}) = 0$, 得 $\frac{I}{10^{-12}} = 1$, 故 $I = 10^{-12}\,\mathrm{W/m^2}$, 恰为模型中的基准声强.

(3) 宿舍谈话 $I = 5\times10^{-7}\,\mathrm{W/m^2}$

$$Y = 10\lg\left(\frac{5\times10^{-7}}{10^{-12}}\right) = 10\lg(5\times10^5) = 10(\lg 5 + 5).$$

由 $\lg 5 = \lg(10/2) = 1 - \lg 2$

$$Y = 10(6 - \lg 2) \approx 10(6 - 0.301) = 56.99(\text{分贝}),$$

高于 50 分贝的睡眠上限, 故会影响他人休息.

注 3.13 (对数标度的意义)
声学、地震学(里氏震级)、化学 (pH) 等广泛采用对数标度, 因为相关物理量的跨度极大, 而感知更依赖数量级. 对数把乘法型的数量级变化变为加法型的标度变化, 将巨大动态范围压缩到可处理的尺度.

3.15 函数的零点(一)

函数 $y = f(x)$ 的零点就是方程 $f(x) = 0$ 的实数根, 也是图像 $y = f(x)$ 与 x 轴交点的横坐标. 零点、方程的根、横截距这三种说法指同一个对象. 本节研究方程 $f(x) = 0$ 的解是否存在、有多少个、落在何处.

3.15.1 零点的存在性

我们要回答的是,如何确信一个函数存在零点,尤其是在无法直接解出 $f(x) = 0$ 时. 答案蕴含在连续性中.

设想一条连续曲线, 一端在 x 轴下方, 另一端在 x 轴上方. 当它从一端连到另一端时, 必然穿过 x 轴. 这就是零点存在性定理的直观背景.

注 3.14
零点存在性的严格证明依赖连续函数的介值定理, 见第16章. 图像连续且两端异号则必有交点, 是介值定理的几何表述.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

零点存在性 (介值定理)

定理 3.28 (零点存在性定理) 若函数 $y=f(x)$ 在闭区间 $[a,b]$ 上的图像是一条连续不断的曲线, 并且其在区间端点处的函数值异号, 即 $f(a)\cdot f(b)<0$, 则在开区间 (a,b) 内, 函数 $y=f(x)$ 至少存在一个零点.

证明 我们用介值定理. 由 $f(a)f(b)<0$ 可知 $f(a)$ 与 $f(b)$ 异号, 因此 0 严格介于 $f(a)$ 与 $f(b)$ 之间.

因为 f 在闭区间 $[a,b]$ 上连续, 由介值定理, 存在 $c\in[a,b]$ 使得 $f(c)=0$. 又 $f(a)$ 与 $f(b)$ 均不为 0, 所以 c 不可能是端点 a 或 b, 从而 $c\in(a,b)$. 因此函数在 (a,b) 内至少存在一个零点. □

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

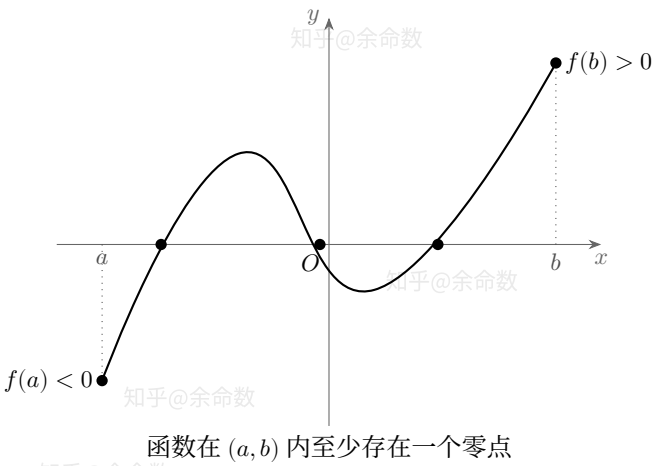


图 3.22 零点存在性定理的几何直观. 由于 $f(a)$ 与 $f(b)$ 异号, 连续曲线必然穿越 x 轴.

条件 $f(a)\cdot f(b)<0$ 只是函数在 (a,b) 内存在零点的一个**充分条件**. 函数完全可能在区间内存在零点, 但不满足该条件. 例如 $f(x)=x^2$ 在区间 $[-1,1]$ 上有零点 $x=0$, 而 $f(-1)f(1)=1>0$. 这通常发生在图像与 x 轴相切, 或穿越 x 轴偶数次时.

导数辅助判定零点

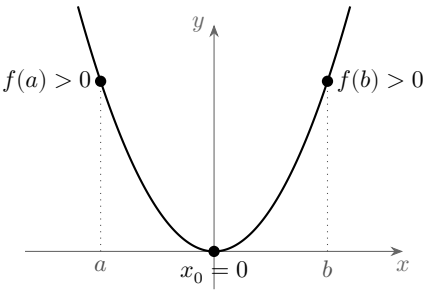
知乎@余命数

例题 3.53 函数 $f(x)=\ln x+x^2-2$ 的零点所在的区间是(　).

A.$(0,\frac{\sqrt{2}}{2})$　　B.$(\frac{\sqrt{2}}{2},1)$　　C.$(1,\sqrt{2})$　　D.$(\sqrt{2},2)$

解 函数 $f(x)$ 的定义域为 $(0,+\infty)$. 方程 $\ln x+x^2-2=0$ 没有显式根, 我们用零点存在性定理定位零点. 该定理只保证存在性, 因此先考察单调性以判定零点是否唯一.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



$f(a)f(b)>0$, 但函数仍存在零点

图 3.23　零点存在性定理条件的非必要性

导函数 $f'(x)=\frac{1}{x}+2x$. 对一切 $x>0$ 有 $\frac{1}{x}>0$ 且 $2x>0$, 故 $f'(x)>0$, f 在 $(0,+\infty)$ 上严格递增, 至多一个零点. 问题化为在四个选项中定位这唯一的零点.

先看 $x=1$

$$f(1)=\ln 1+1^2-2=-1<0.$$

由严格递增, 零点 x_0 必在 1 的右侧, 因此 $x_0>1$, 排除 A、B.

再看选项 C 的右端点 $x=\sqrt{2}$

$$f(\sqrt{2})=\ln(\sqrt{2})+(\sqrt{2})^2-2=\frac{1}{2}\ln 2>0.$$

于是 $f(1)<0$, $f(\sqrt{2})>0$. 由零点存在性定理, $(1,\sqrt{2})$ 内至少有一个零点; 结合单调性, 该零点唯一, 恰落在区间 $(1,\sqrt{2})$ 内.

故选 C.

3.15.2 零点的求解与个数判定

零点存在之后, 我们进一步确定它的位置与个数. 常用的做法是把代数分析与几何直观结合起来.

3.15.2.1 代数变换与图像分析

对于复杂方程 $f(x)=0$, 一个常用做法是把它们**等价变形**为 $g(x)=h(x)$, 其中 g,h 是我们更熟悉的函数. 于是 $f(x)=0$ 的根与 $g(x)=h(x)$ 的根相同. 这样, 求根问题就变成判断两条曲线交点个数的问题.

例题 3.54　判断函数 $f(x)=\ln(x+1)-\frac{2}{x}$ 的零点个数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 直接分析 $f(x)$ 比较复杂. 我们把它等价地写成

$$\ln(x+1)=\frac{2}{x}.$$

于是原问题等价于考察函数 $y_1=\ln(x+1)$ 与函数 $y_2=\frac{2}{x}$ 的图像的交点个数.

分别分析.$y_1=\ln(x+1)$ 由 $y=\ln x$ 向左平移 1 个单位得到，定义域为 $(-1,+\infty)$，严格递增，且经过原点. $y_2=\frac{2}{x}$ 为反比例函数，图像在第一、三象限. 在同一坐标系中画出草图, 见 图 3.24.

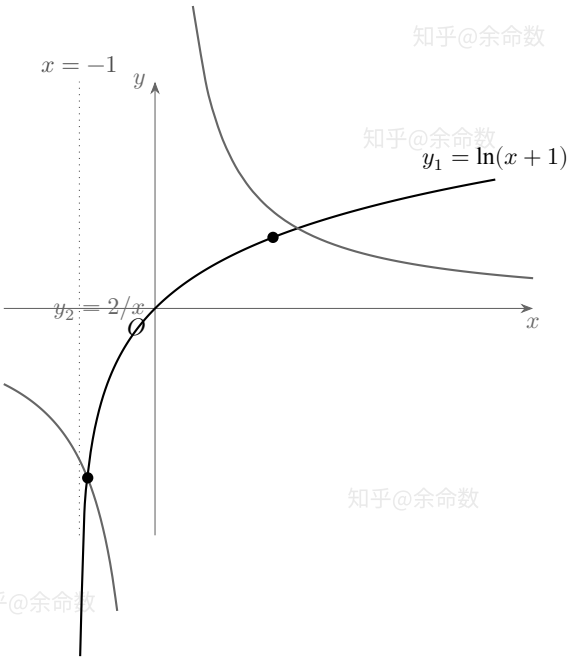


图 3.24 函数 $y_1=\ln(x+1)$ 与 $y_2=2/x$ 的图像

图像显示两条曲线在第一象限和 $(-1,0)$ 内各有一个交点. 下面严格证明这一点.

在 $(0,+\infty)$ 上, $y_1=\ln(x+1)$ 严格递增, $y_2=2/x$ 严格递减. 因此它们的图像在此区间内至多有一个交点.

只需验证存在性.$y_1(1)=\ln 2<1<2=y_2(1)$, 而 $y_1(2)=\ln 3>1=y_2(2)$, 故 $(1,2)$ 内必有交点.

在区间 $(-1,0)$ 内, $y_1=\ln(x+1)$ 的值域为 $(-\infty,0)$, $y_2=2/x$ 的值域也为 $(-\infty,0)$. 直接看图还不确定, 因此我们考虑辅助函数 $F(x)=\ln(x+1)-2/x$. 其导数 $F'(x)=\frac{1}{x+1}+\frac{2}{x^2}>0$, 故 $F(x)$ 在 $(-1,0)$ 上单调递增. 又 $\lim_{x\to-1^+}F(x)=-\infty$, $\lim_{x\to0^-}F(x)=+\infty$, 由零点存在性定理, 此区间内也恰有一个零点.

两个函数图像共有两个交点, 故原函数 $f(x)$ 有且仅有 2 个零点.

用导数判定零点个数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.15.2.2 单调性与极值分析

对于不能拆成两个简单函数的 $f(x)$, 判定零点个数通常先研究其单调性. 单调区间由 $f'(x)$ 的符号决定;极值点处的函数值则是判断零点是否存在的关键.

判定时依次做三件事.

1. 求导数 $f'(x)$, 并求解 $f'(x)=0$ 得到所有临界点.
2. 根据临界点划分函数的单调区间.
3. 计算函数在所有极值点处的函数值, 并考察函数在无穷远处的极限行为.
4. 结合单调性与上述关键点的函数值符号, 在每个单调区间内应用零点存在性定理, 最终确定零点的总个数.

知乎@余命数

例题 3.55 讨论函数 $f(x)=e^x-ax$ $(a\in\mathbb{R})$ 的零点个数.

解 我们先从 $f(x)$ 的单调性入手. 该问题等价于判断 $e^x=ax$ 的根的个数, 亦即曲线 $y=e^x$ 与直线 $y=ax$ 的交点个数.

其导函数为 $f'(x)=e^x-a$.

令 $f'(x)=0$, 得 $e^x=a$. 这个方程是否有解, 取决于 a 的值.

情形一 ($a<0$). 若 $a<0$, 则 $e^x=a$ 无解, 且 $f'(x)=e^x-a>0$ 对所有 $x\in\mathbb{R}$ 成立. 故 $f(x)$ 在 $\mathbb{R}$ 上严格单调递增. f 在 $\mathbb{R}$ 上连续且严格递增, 因此与 x 轴至多有一个交点. 又 $\lim_{x\to-\infty}f(x)=-\infty$, 且 $\lim_{x\to+\infty}f(x)=+\infty$, 由介值定理, 函数至少有一个零点. 因此, $a<0$ 时 $f(x)$ 有且仅有 1 个零点.

当 $a=0$ 时, $f(x)=e^x>0$ 对所有 $x\in\mathbb{R}$ 恒成立, 因而没有零点.

情形二 ($a>0$). 在此情形下, $f'(x)=0$ 有唯一解 $x_0=\ln a$. 当 $x<\ln a$ 时, $f'(x)<0$, $f(x)$ 单调递减. 当 $x>\ln a$ 时, $f'(x)>0$, $f(x)$ 单调递增. 因此, 函数在 $x=\ln a$ 处取得唯一的极小值, 也是最小值.

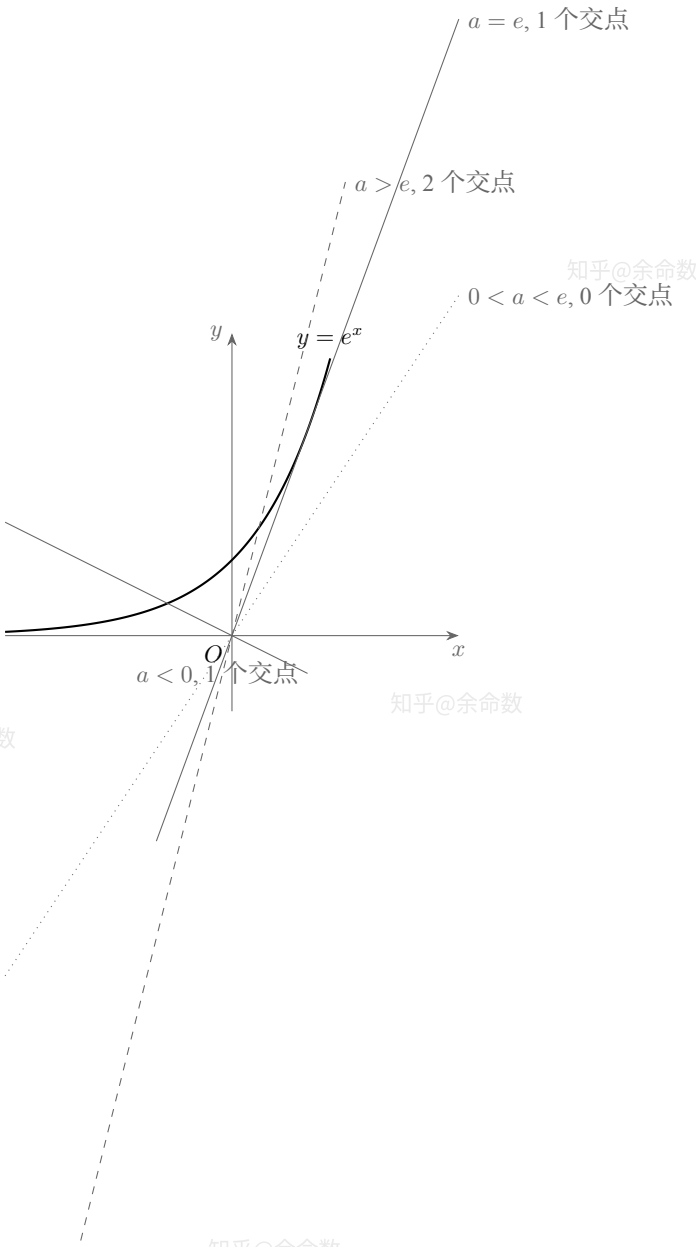
$$f_{\min}=f(\ln a)=e^{\ln a}-a\ln a=a-a\ln a=a(1-\ln a)$$

函数零点的个数, 完全取决于这个最小值与零的大小关系.

- 若 $f_{\min}>0$, 即 $a(1-\ln a)>0$. 由于 $a>0$, 这等价于 $1-\ln a>0$, 即 $\ln a<1$, 解得 $0<a<e$. 此时, 函数的最小值大于零, 图像恒在 x 轴上方, 故函数**没有零点**.
- 若 $f_{\min}=0$, 即 $a(1-\ln a)=0$. 由于 $a>0$, 这等价于 $1-\ln a=0$, 即 $\ln a=1$, 解得 $a=e$. 此时, 函数的最小值恰好为零, 图像与 x 轴仅有一个切点, 故函数有且仅有 **1 个零点**.
- 若 $f_{\min}<0$, 即 $a(1-\ln a)<0$. 由于 $a>0$, 这等价于 $1-\ln a<0$, 即 $\ln a>1$, 解得 $a>e$. 此时, 函数的最小值小于零. 考虑到 $\lim_{x\to\pm\infty}f(x)=+\infty$, 函数图像在极小值点左侧和右侧必然各与 x 轴相交一次. 故函数有 **2 个零点**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数



知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

图 3.25 方程 $e^x = ax$ 根的个数的几何解释

知乎@余命数

函数 $f(x)=e^x-ax$ 的零点个数为

- 当 $a\in(-\infty,0)\cup\{e\}$ 时, 有 1 个零点.
- 当 $a\in[0,e)$ 时, 有 0 个零点.
- 当 $a\in(e,+\infty)$ 时, 有 2 个零点.

3.16 分段函数

我们看一类函数. 它的对应法则在定义域的不同部分并不相同. 例如, 所得税的税率随收入水平分级变化, 物体在不同介质中运动遵循不同的物理定律. 单一解析表达式往往无法写出这种关系, 于是我们引入分段函数. 分析分段函数时, 我们先分别考察各片段上的局部性质, 再考察这些性质在拼接点处如何衔接, 从而判断其整体行为.

注 3.15
分段点处的连续性与可导性使用极限语言, 严格定义见第16章. 图像上可先判断分段点是否连续、是否有尖角.

3.16.1 定义与基本示例

定义 3.17 (分段函数) 一个**分段函数**是指在其定义域的不同子集上, 分别由不同的解析式所定义的函数. 其一般形式为:

$$f(x)=\begin{cases}f_1(x), & x\in D_1\\ f_2(x), & x\in D_2\\ \vdots & \\ f_n(x), & x\in D_n\end{cases}$$

其中, 各个子集 D_i 互不相交, 即 $D_i\cap D_j=\emptyset$ 对任意 $i\neq j$, 且它们的并集 $\bigcup_{i=1}^{n}D_i$ 构成了函数 $f(x)$ 的完整定义域 D.

我们以出租车计费为例. 设乘客乘坐的里程为 x, 支付的车费为 $f(x)$, 则计费规则可以写成

$$f(x)=\begin{cases}13, & 0\le x\le 3\\ 13+2.4(x-3), & x>3\end{cases}$$

其中, 第一段表示里程不超过 3 公里时车费固定为 13 元, 第二段表示里程超过 3 公里时车费为 13 元加上超出部分每公里 2.4 元. 对于任意给定的里程 x, 我们先判断它属于哪个计价区间, 再用该区间对应的公式计算费用. 上述规则是一个单一且明确的函数.

现在回头审视定义中的各个部分.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- ▷ $D_1, D_2, \ldots$ (**定义域的划分**) 这些子集对应于出租车计费的各个里程区间. 它们将整个定义域分割成若干个互不重叠的部分.
- ▷ $f_1(x), f_2(x), \ldots$ (**不同的法则**) 这些解析式对应于每个里程区间所使用的不同计价公式. 它们给出函数在各自区域的行为.
- ▷ **互不相交** ($D_i \cap D_j = \emptyset$) 这是分段函数成为**函数**的根本. 它意味着任意输入值 x 只能落入**唯一一个**子定义域 D_i 中. 这就保证了对于一个输入, 只会有一个计价公式被激活, 从而产生唯一的输出值. 一次 5 公里的行程不能同时适用 3 公里内和 3 公里外的规则.
- ▷ **并集构成完整定义域** 这保证了规则的完备性. 对于任意允许的输入值, 系统中都存在一个对应的规则来处理它, 不会出现无法计价的情况.

因此, 分段函数是一**个**定义完整的函数. 它的特殊之处在于其对应法则是**有条件的**, 需要根据自变量的取值来选择执行.

分段函数与普通函数的区别在于, 它的定义域被**划分**成若干子集. 函数在某一点 x_0 的取值, 完全取决于 x_0 属于哪个子定义域. 绘制分段函数的图像时, 我们在同一坐标系中把各个函数片段 $y = f_i(x)$ 在其各自的定义域 D_i 上的图像拼接起来. 在此过程中, 需要特别关注各个分段的**端点**. 这些点是函数性质可能发生突变的地方, 我们称之为**分段点**. 在分段点处, 我们要仔细判断该点属于前一段还是后一段, 并在图像上用实心点与空心点加以区分, 以保证函数的单值性.

知乎@余命数 知乎@余命数 知乎@余命数

例题 3.56 设函数 $f(x) = \begin{cases} x+2, & x \le 1 \\ 3, & x > 1 \end{cases}$. 求 $f(0)$ 与 $f(f(0))$ 的值.

解 为求 $f(0)$, 我们先判断自变量 0 所在的区间. 由于 $0 \le 1$, $f(0)$ 适用第一个解析式, 故 $f(0) = 0 + 2 = 2$.

为求 $f(f(0))$, 我们需计算 $f(2)$. 自变量 2 满足 $2 > 1$, 适用第二个解析式, 故 $f(2) = 3$. 因此, $f(f(0)) = 3$.

例题 3.57 设函数 $f(x) = \begin{cases} x^2+1, & x \ge 0 \\ 2x+1, & x < 0 \end{cases}$. 求 $f(3)$ 的值.

解 为求 $f(3)$, 我们先确定自变量 $x = 3$ 属于哪个子定义域.

由于 $3 \ge 0$, 自变量 $x = 3$ 满足第一个条件.

于是, 应用第一个解析式 $f(x) = x^2 + 1$ 来计算函数值.

$$f(3) = 3^2 + 1 = 9 + 1 = 10$$

知乎@余命数

故 $f(3)$ 的值为 10.

例题 3.58　设函数 $g(x)=\begin{cases}2x, & x>1\\ x-1, & x\le 1\end{cases}$. 求 $g(g(1))$ 的值.

解　求解复合函数值 $g(g(1))$ 需要从里向外逐步计算.

先计算内层函数的值 $g(1)$. 自变量 $x=1$ 满足条件 $x\le 1$. 因此, 应用第二个解析式 $g(x)=x-1$.

$$g(1)=1-1=0$$

再将此结果作为外层函数的自变量, 计算 $g(0)$. 自变量 $x=0$ 满足条件 $x\le 1$. 因此, 再次应用第二个解析式.

$$g(g(1))=g(0)=0-1=-1$$

故 $g(g(1))$ 的值为 -1.

例题 3.59　已知函数 $h(x)=\begin{cases}\sqrt{x}, & x\ge 1\\ x+1, & x<1\end{cases}$. 若 $h(x)=3$, 求 x 的值.

解　我们需寻找一个自变量 x, 使其函数值为 3. 由于函数的解析式是分段的, 我们在每个子定义域上分别讨论.

情形一: 假设 $x\ge 1$. 在此条件下, 函数的解析式为 $h(x)=\sqrt{x}$. 令 $\sqrt{x}=3$, 两边平方得 $x=9$. 我们检验此解是否满足本情形的假设. 由于 $9\ge 1$, 该解有效.

情形二: 假设 $x<1$. 在此条件下, 函数的解析式为 $h(x)=x+1$. 令 $x+1=3$, 解得 $x=2$. 我们检验此解是否满足本情形的假设. 由于 2 并不小于 1, 该解无效, 必须舍去.

综合以上两种情形, 唯一有效的解是 $x=9$.

分段点连续与可导

3.16.2　连续性与可导性

分析分段函数时, 我们关注其在分段点处的行为.

知乎@余命数　知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 3.18 (分段点的连续性) 设 x_0 是函数 $f(x)$ 的一个分段点.$f(x)$ 在 x_0 处**连续**的充要条件是其在该点的左极限、右极限均存在且等于该点的函数值, 即

$$\lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x) = f(x_0)$$

若此条件不满足, 则称函数在 x_0 处**间断**.

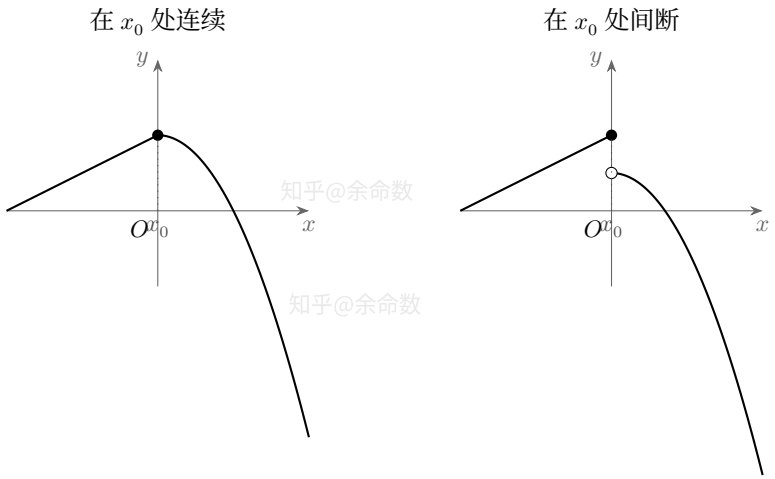


图 3.26 分段点处的连续与间断

定义 3.19 (分段点的可导性) 函数 $f(x)$ 在分段点 x_0 处**可导**的充要条件是其左导数与右导数存在且相等, 即

$$f'_-(x_0) = \lim_{h\to 0^-} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{h\to 0^+} \frac{f(x_0+h)-f(x_0)}{h} = f'_+(x_0)$$

若左右导数不相等, 则函数在 x_0 处不可导, 其图像在该点形成**尖点**或**角点**.

例题 3.60 分析函数 $f(x)=|x|$ 在 $x=0$ 处的连续性与可导性.

解 函数 $f(x)=|x|$ 可写为分段函数形式 $f(x)=\begin{cases} x, & x\geq 0 \\ -x, & x<0 \end{cases}$.

在 $x=0$ 处, 左极限 $\lim_{x\to 0^-} f(x) = \lim_{x\to 0^-}(-x) = 0$. 右极限 $\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} x = 0$. 函数值为 $f(0)=0$. 由于左极限、右极限与函数值都等于 0, 函数在 $x=0$ 处连续.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

左导数为 $f'_-(0)=\lim_{h\to0^-}\frac{|0+h|-|0|}{h}=\lim_{h\to0^-}\frac{-h}{h}=-1$. 右导数为 $f'_+(0)=\lim_{h\to0^+}\frac{|0+h|-|0|}{h}=\lim_{h\to0^+}\frac{h}{h}=1$. 由于 $f'_-(0)\neq f'_+(0)$, 函数在 $x=0$ 处不可导.

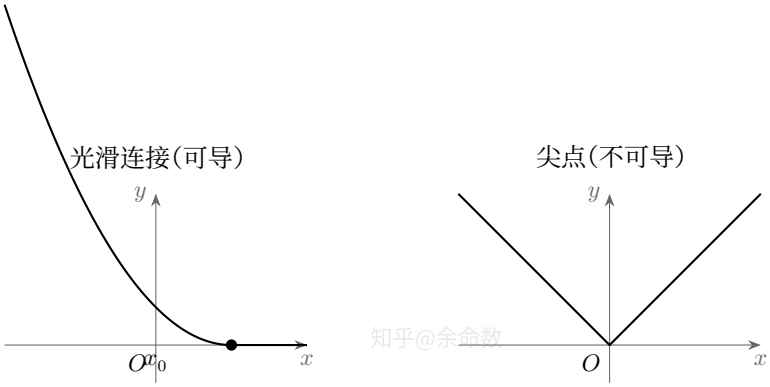


图 3.27　分段点处的光滑与尖点

知乎@余命数

3.16.3 整体性质的判定

一个函数在各分段区间上具有某种性质, 例如单调性, 并不能保证它在整个定义域上都具有此性质. 判定整体性质时, 我们还需考察函数在分段点处的衔接行为.

例如, 一个分段函数若要在整个定义域上单调递增, 必须满足以下两个条件.

1. 函数在每个子定义域上均单调递增.
2. 在任意相邻的两个子定义域 D_i, D_{i+1} (设 D_i 在 D_{i+1} 左侧) 的分界点 x_0 处, 有 $\lim_{x\to x_0^-}f(x)\leq f(x_0)$.

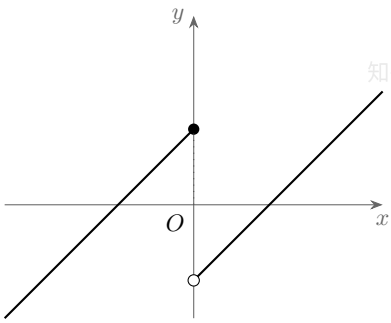


图 3.28　一个在各分段区间均递增, 但整体非单调的函数.

知乎@余命数

例题 3.61 已知函数 $f(x)=\begin{cases}(2a-1)x+7a-2, & x<1\\ a^x, & x\geq 1\end{cases}$ 在 $\mathbb{R}$ 上单调递减，求实数 a 的取值范围.

解 为使函数 $f(x)$ 在整个定义域 $\mathbb{R}$ 上单调递减，必须同时满足三个条件.

条件一是，函数在左侧子定义域 $(-\infty,1)$ 上必须单调递减. 此段为线性函数，其单调性由斜率决定.

$$2a-1<0 \implies a<\frac{1}{2}$$

条件二是，函数在右侧子定义域 $[1,+\infty)$ 上必须单调递减. 此段为指数函数，其单调性由底数决定.

$$0<a<1$$

条件三是，在分段点 $x=1$ 处，左侧区间的终点值必须不小于右侧区间的起点值，以保证单调递减趋势得以延续.

$$\lim_{x\to 1^-} f(x)\geq f(1)$$

计算得

$$(2a-1)(1)+7a-2\geq a^1$$

$$9a-3\geq a \implies 8a\geq 3 \implies a\geq\frac{3}{8}$$

将这三个条件联立，求解其交集.

$$\begin{cases}a<1/2\\ 0<a<1\\ a\geq 3/8\end{cases}$$

注意到 $3/8=0.375$ 且 $1/2=0.5$，因此三个区间的公共部分为 $[\frac{3}{8},\frac{1}{2})$.

故实数 a 的取值范围是 $[\frac{3}{8},\frac{1}{2})$.

例题 3.62 讨论方程 $f(x)=k$ 的根的个数，其中 $f(x)=\begin{cases}-x^2-2x, & x\leq 0\\ \ln(x+1), & x>0\end{cases}$.

解 此问题等价于考察函数 $y=f(x)$ 的图像与水平直线 $y=k$ 的交点个数. 我们分析函数 $f(x)$ 的图像，特别是其单调性、极值与端点行为.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

分段分析函数 $f(x)$.

当 $x \le 0$ 时, $f(x) = -x^2 - 2x = -(x+1)^2 + 1$. 这是一个开口向下、对称轴为 $x = -1$ 的抛物线. 在区间 $(-\infty, -1]$ 上, 函数单调递增. 在区间 $[-1, 0]$ 上, 函数单调递减. 函数在 $x = -1$ 处取得极大值 $f(-1) = 1$. 在分段点 $x = 0$ 处, $f(0) = 0$.

当 $x > 0$ 时, $f(x) = \ln(x+1)$. 这是一个严格单调递增的对数函数. 它在 $x \to 0^+$ 时的极限为 $\lim_{x\to 0^+} \ln(x+1) = \ln 1 = 0$.

综合两段的信息, 我们可以绘制出函数的草图. 函数在 $x = 0$ 处连续. 它从 $-\infty$ 增长到极大值 1, 然后下降到 0, 再从 0 开始无限增长.

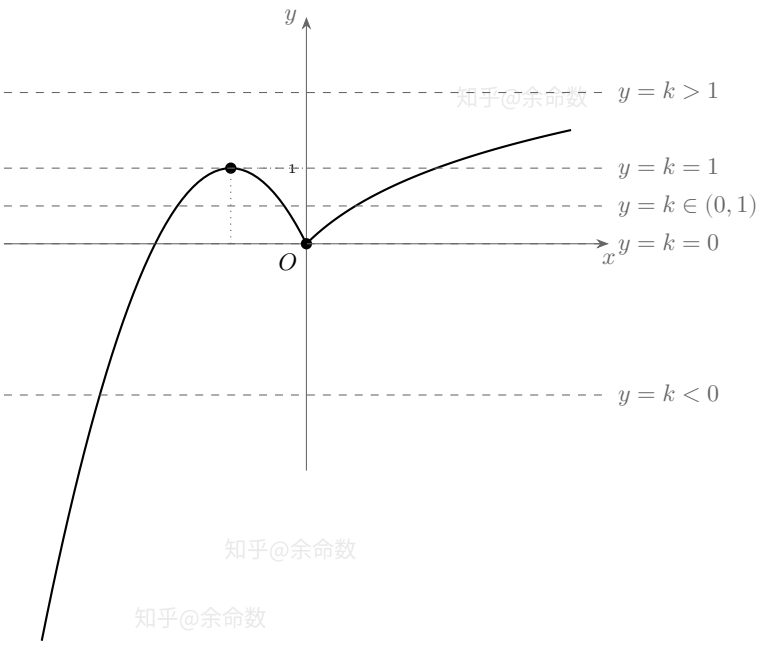


图 3.29　函数 $y = f(x)$ 与水平直线 $y = k$ 的交点情况

观察图像与水平直线 $y = k$ 的交点个数可知.

- 当 $k > 1$ 时, 直线不与第一段相交, 只与第二段有一个交点, 共 1 个交点.
- 当 $k = 1$ 时, 直线与图像在 $x = -1$ 处相切, 并在第二段有一个交点, 共 2 个交点.
- 当 $0 < k < 1$ 时, 直线与图像在第一段的两个分支以及第二段各有一个交点, 共 3 个交点.
- 当 $k = 0$ 时, 直线与图像在 $x = 0$ 和 $x = -2$ 处相交, 共 2 个交点.
- 当 $k < 0$ 时, 直线仅与图像在第一段有一个交点, 共 1 个交点.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此, 方程 $f(x)=k$ 的根的个数为

$$\text{根的个数}=\begin{cases}1, & k<0 \text{ 或} k>1\\ 2, & k=0 \text{ 或} k=1\\ 3, & 0<k<1\end{cases}$$

上述分析表明, 方程 $f(x)=k$ 的根的个数由常数 k 与函数的关键值(极大值 1 与连接点值 0)之间的关系决定.

知乎@余命数

3.16.4 函数的基本要素分析

分段函数的定义域、值域与单调性等基本性质, 由其各个组成部分的性质以及它们在分段点处的衔接方式共同决定.

3.16.4.1 定义域与值域

3.16.4.1.1 定义域 分段函数的定义域是其所有子定义域的并集.

$$D=D_1\cup D_2\cup\cdots\cup D_n$$

这通常在函数定义时已经明确给出.

3.16.4.1.2 值域 分段函数的值域是其在各个子定义域上取得的值的集合的并集. 求解过程分为两步.

知乎@余命数

1. 分别求出函数 $y=f_i(x)$ 在其对应定义域 D_i 上的值域 R_i.
2. 将所有这些局部值域合并, 即求其并集 $R=R_1\cup R_2\cup\cdots\cup R_n$.

例题 3.63 求函数 $f(x)=\begin{cases}x+2, & x\le -1\\ x^2, & -1<x<2\\ 2, & x\ge 2\end{cases}$ 的值域.

知乎@余命数

解 我们分别考察函数在每一个子定义域上的取值范围, 然后将这些范围合并.

- 当 $x\in(-\infty,-1]$ 时, $f(x)=x+2$. 这是一个单调递增的线性函数. 其在该区间上的值域为 $(-\infty,f(-1)]$, 即 $(-\infty,1]$.
- 当 $x\in(-1,2)$ 时, $f(x)=x^2$. 这是一个二次函数, 在区间 $(-1,0]$ 上递减, 在 $[0,2)$ 上递增. 其在该区间上的值域为

$$[f(0),\max(\lim_{x\to -1^+}x^2,\lim_{x\to 2^-}x^2))=[0,4).$$

知乎@余命数

- 当 $x \in [2, \infty)$ 时, $f(x) = 2$. 这是一个常数函数, 其值域为单点集 $\{2\}$.

函数 $f(x)$ 的总值域是这三个部分值域的并集.

$$(-\infty, 1] \cup [0, 4) \cup \{2\} = (-\infty, 4)$$

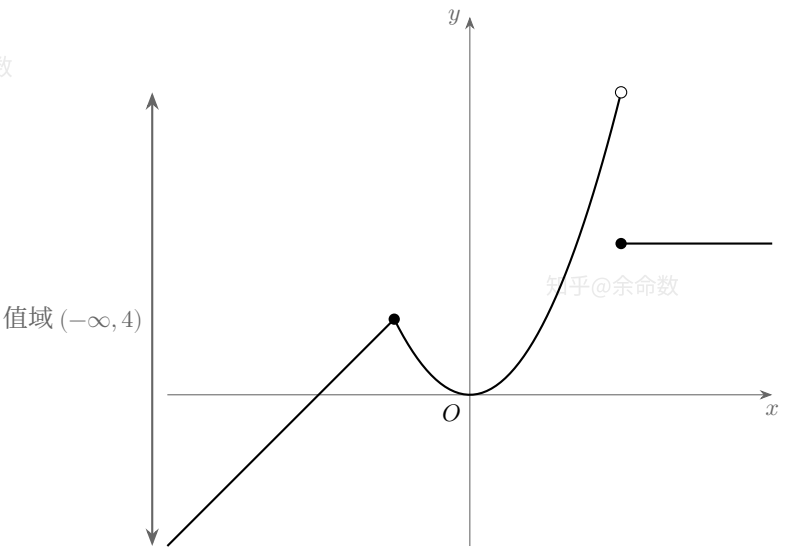


图 3.30　函数 $f(x)$ 的图像及其值域

因此, 函数的值域为 $(-\infty, 4)$.

3.16.4.2 单调性

一个分段函数若要在整个定义域上单调, 必须同时满足两个条件.

1. 函数在**每个**子定义域上均具有相同的单调性, 例如, 均为单调递增.
2. 在所有分段点处, 函数的衔接必须**保持**这种单调趋势. 对于单调递增函数, 即后一段的起始值必须不小于前一段的结束值.

若第二个条件不满足, 即使函数在每个局部区间上都单调, 其整体也不是单调函数.

3.16.5 应用示例

分段函数可以描述现实世界中的条件依赖关系. 以下示例展示如何将几何问题与生活中的计价规则转化为分段函数的数学模型.

例题 3.64　如图, 在边长为 4 的正方形 $ABCD$ 的边上有一动点 P, 沿着折线 $B \to C \to D \to A$ 运动. 设点 P 运动的路程为 x, $\triangle APB$ 的面积为 y.

1. 求 y 与 x 之间的函数关系式.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

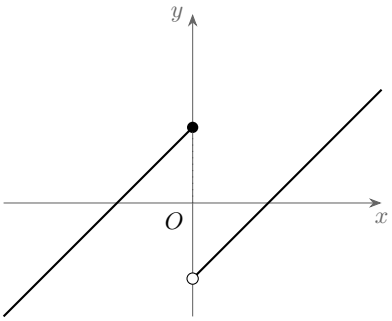
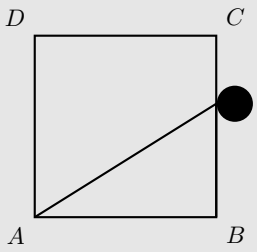


图 3.31 一个在各分段区间均递增, 但因在 $x=0$ 处发生向下跳跃而整体非单调的函数.

2. 求面积最大时, x 的取值范围.



解 (1) 三角形 APB 的底边 AB 长度固定为 4. 因此, 其面积 y 完全由高决定, 这个高是点 P 到边 AB 的垂直距离.

$$y=\frac{1}{2}\cdot 底\cdot 高=\frac{1}{2}\cdot 4\cdot h_P=2h_P$$

点 P 的运动路径由三条边构成, 其在不同边上运动时, 高 h_P 与路程 x 的关系不同. 因此 $y(x)$ 是一个分段函数.

分段分析.

分段一: 点 P 在边 BC 上运动. 此时, P 从 B 运动到 C. 路程 x 的取值范围是 $[0,4]$. 点 P 到 AB 的高 h_P 等于它从点 B 向上运动的距离, 即路程 x. 故 $h_P=x$. 面积 $y=2x$.

分段二: 点 P 在边 CD 上运动. 此时, P 从 C 运动到 D. 路程 x 从 4 增加到 $4+4=8$. 其取值范围是 $(4,8]$. 点 P 在 CD 上运动时, 它到 AB 的高 h_P 始终等于正方形的边长 4. 故 $h_P=4$. 面积 $y=2\cdot 4=8$.

分段三: 点 P 在边 DA 上运动. 此时, P 从 D 运动到 A. 路程 x 从 8 增加到 $8+4=12$. 其取值范围是 $(8,12]$. 点 P 到 AB 的高 h_P 等于它与点 A 的距离. 总路程为 12, 已走路程为 x, 剩余路程即为 PA 的长度. 故 $h_P=12-x$. 面积 $y=2(12-x)$.

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

综合以上三段, 我们得到函数关系式

$$y=f(x)=\begin{cases}2x, & 0\le x\le 4\\ 8, & 4<x\le 8\\ 2(12-x), & 8<x\le 12\end{cases}$$

(2) 为求面积的最大值, 我们分析函数 $y(x)$ 在其定义域 $[0,12]$ 上的行为.

- 在 $[0,4]$ 上, $y=2x$ 是增函数, 最大值为 $y(4)=8$.
- 在 $(4,8]$ 上, $y=8$ 是常数函数.
- 在 $(8,12]$ 上, $y=2(12-x)$ 是减函数, 其值域为 $[0,8)$.

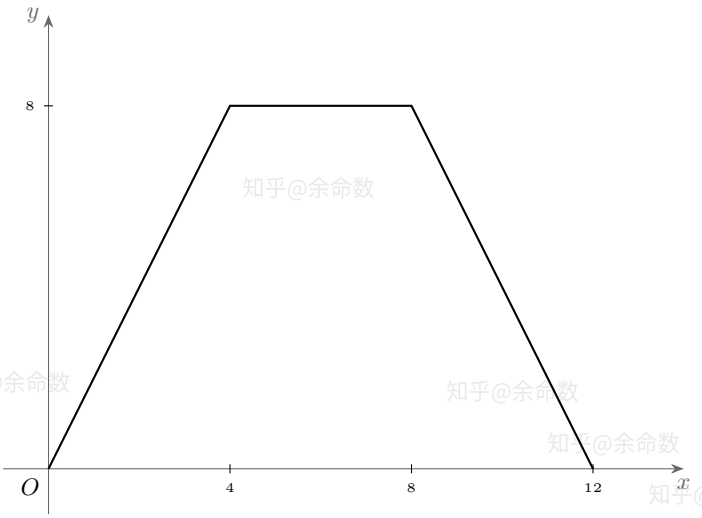


图 3.32　面积函数 $y(x)$ 的图像

通过分析各段的值域并结合图像, 函数 y 的最大值为 8. 这个最大值在 $x=4$ 时首次达到, 并一直保持到 $x=8$. 因此, 当面积最大时, x 的取值范围是 $[4,8]$.

例题 3.65　某市空调公共汽车的票价按下列规则制定:

1. 乘坐汽车 5 km 以内, 票价 2 元;
2. 5 km 以上, 每增加 5 km, 票价增加 1 元(不足 5 km 的按 5 km 计算).

已知两个相邻的公共汽车站间相距约为 1 km, 如果沿途(包括起点站和终点站)设 20 个汽车站, 请根据题意写出票价与里程之间的函数解析式.

解　首先, 我们确定函数的自变量和因变量. 设乘客乘坐的里程为 x (km), 对应的票价为 y(元).

知乎@余命数　知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

接着, 我们确定自变量 x 的取值范围. 线路共有 20 个站, 站间距为 1 km, 因此总里程为 $(20-1)\times 1=19$ km. 乘客乘坐的里程 x 的范围是 $(0,19]$.

然后, 我们根据计价规则, 将定义域 $(0,19]$ 分割成不同的计价区间. 计价规则以 5 km 为一个台阶.

区间一: $0<x\le 5$. 根据规则 (1), 在此里程范围内, 票价固定为 2 元.

区间二: $5<x\le 10$. 里程超过 5 km, 进入增价层级 A. 根据规则 (2), 票价在 2 元的基础上增加 1 元. 票价为 $2+1=3$ 元.

区间三: $10<x\le 15$. 里程超过 10 km, 进入增价层级 B. 票价在 3 元的基础上再增加 1 元. 票价为 $3+1=4$ 元.

区间四: $15<x\le 19$. 里程超过 15 km, 进入增价层级 C. 票价在 4 元的基础上再增加 1 元. 票价为 $4+1=5$ 元.

把上述各区间的结果合并, 我们得到票价 y 关于里程 x 的分段函数解析式

$$y=f(x)=\begin{cases}2, & 0<x\le 5\\ 3, & 5<x\le 10\\ 4, & 10<x\le 15\\ 5, & 15<x\le 19\end{cases}$$

这是一个阶梯函数.

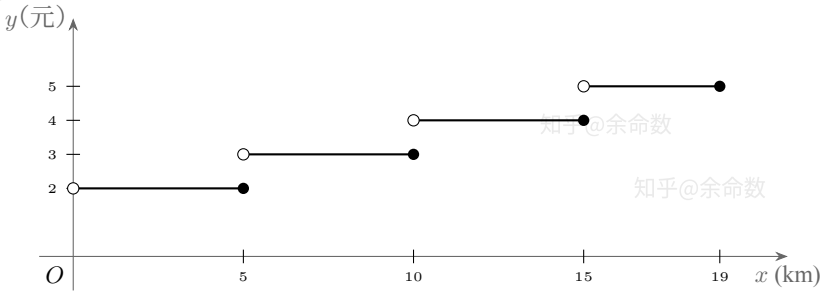


图 3.33 公交票价函数图像

例题 3.66 设函数 $f(x)=\begin{cases}-ax+1, & x<a\\ (x-2)^2, & x\ge a\end{cases}$. 若 $f(x)$ 存在最小值, 求实数 a 的最大值.

解 我们分析函数在不同参数 a 下的形态, 并确定它是否存在一个全局最小值. 函数的表达式由线性部分与二次函数部分构成, 其整体行为取决于分段点 a 的位置. 我们对 a 的取值进行分类讨论.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

情形一： $a<0$. 此时, 线性部分 $y=-ax+1$ 的斜率 $-a$ 大于零, 函数在区间 $(-\infty,a)$ 上单调递增. 当 $x\to-\infty$ 时, $f(x)\to-\infty$. 函数无下界, 因此不存在最小值.

情形二： $a=0$. 此时函数为 $f(x)=\begin{cases}1, & x<0\\ (x-2)^2, & x\geq 0\end{cases}$. 在区间 $(-\infty,0)$ 上, 函数值为常数 1. 在区间 $[0,+\infty)$ 上, 函数 $y=(x-2)^2$ 在 $x=2$ 处取得最小值 0. 因此, 函数 $f(x)$ 的最小值为 0. 故 $a=0$ 是一个可能的值.

情形三： $a>0$. 此时, 线性部分 $y=-ax+1$ 的斜率 $-a$ 小于零, 函数在区间 $(-\infty,a)$ 上单调递减.

我们根据分段点 a 与抛物线顶点 $x=2$ 的相对位置, 进一步细分.

子情形 3a: $0<a<2$. 函数在 $(-\infty,a)$ 上单调递减, 其值的下确界为 $\lim_{x\to a^-}f(x)=-a^2+1$. 在 $[a,+\infty)$ 上, 函数 $y=(x-2)^2$ 先递减后递增, 在 $x=2$ 处取得最小值 0. 要使函数存在全局最小值, 函数必须有下界, 即当 x 从左侧趋近于 a 时, 函数值不能趋于一个比全局最小值更小的值. 若 $-a^2+1<0$, 则函数在 $(-\infty,a)$ 上的下确界为负, 且无法取到, 故不存在最小值. 因此, 必须满足 $-a^2+1\geq 0$, 即 $a^2\leq 1$. 结合本情形的假设 $a>0$, 得 $0<a\leq 1$.

子情形 3b: $a\geq 2$. 函数在 $(-\infty,a)$ 上单调递减. 在 $[a,+\infty)$ 上, 由于 $a\geq 2$, 函数 $y=(x-2)^2$ 是单调递增的. 因此, 函数的图像在 $x=a$ 点处可能存在一个最低点. 函数的下确界为 $\min(\lim_{x\to a^-}f(x),f(a))=\min(-a^2+1,(a-2)^2)$. 为使最小值存在, 该下确界必须在 $x=a$ 处取到, 即 $f(a)\leq\lim_{x\to a^-}f(x)$.

$$(a-2)^2\leq -a^2+1$$

整理得 $a^2-4a+4\leq -a^2+1$, 即 $2a^2-4a+3\leq 0$. 此二次多项式的判别式 $\Delta=(-4)^2-4(2)(3)=16-24=-8<0$. 由于该二次多项式的开口向上, 故 $2a^2-4a+3$ 恒为正. 不等式 $2a^2-4a+3\leq 0$ 无解. 因此, 在 $a\geq 2$ 的情况下, 函数不存在最小值.

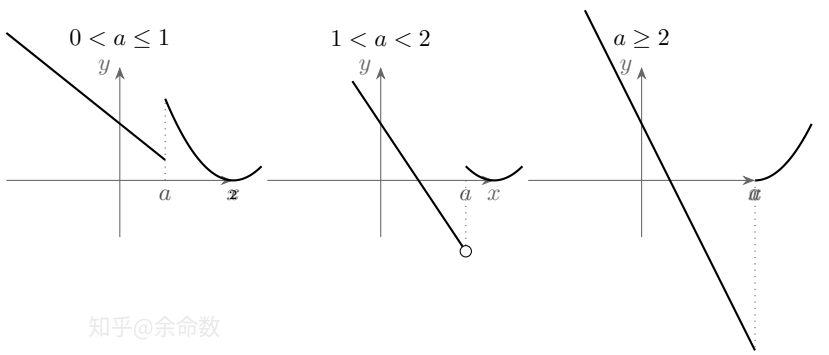


图 3.34　不同 a 值下函数图像形态示意

知乎@余命数
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

综合所有情形, 函数 $f(x)$ 存在最小值的条件是 $a \in [0,1]$. 因此, 实数 a 的最大值为 1.

例题 3.67 设函数 $f(x)=\begin{cases} x^2, & x \le a \\ x^2-2ax+a, & x>a \end{cases}$. 若存在实数 b, 使得函数 $g(x)=f(x)-b$ 有 3 个零点, 则 a 的取值范围为_____.

解 函数 $g(x)$ 有 3 个零点, 等价于方程 $f(x)=b$ 有 3 个不同的实数根. 这在几何上等价于存在一条水平直线 $y=b$, 与函数 $y=f(x)$ 的图像有 3 个不同的交点.

为产生 3 个交点, 函数 $f(x)$ 的图像必须存在峰与谷, 即非单调. 我们先分析函数 $f(x)$ 的结构.

在 $x=a$ 处,函数的左值为 $f(a)=a^2$. 其右极限为 $\lim_{x\to a^+} f(x)=\lim_{x\to a^+}(x^2-2ax+a)=a^2-2a^2+a=a-a^2$.

我们对参数 a 进行分类讨论.

情形一: $a \le 0$. 当 $x \le a$ 时, $f(x)=x^2$ 在此区间上单调递减. 当 $x>a$ 时, $f(x)=x^2-2ax+a=(x-a)^2+a-a^2$. 这是一个顶点在 $x=a$ 的开口向上的抛物线, 故在 $(a,+\infty)$ 上单调递增. 函数图像的整体形态是先下降后上升, 任何水平直线至多与其有两个交点. 因此此情形不满足条件.

情形二: $a>0$. 当 $x \le a$ 时, $f(x)=x^2$ 在 $(-\infty,0]$ 上单调递减, 在 $[0,a]$ 上单调递增. 函数在 $x=0$ 处取得一个局部最小值 $f(0)=0$. 当 $x>a$ 时, $f(x)=(x-a)^2+a-a^2$ 在 $(a,+\infty)$ 上单调递增.

函数 $f(x)$ 的整体图像形态如下. 它从 $+\infty$ 下降至 $(0,0)$, 再上升至分段点 (a,a^2), 然后在 $x=a$ 处发生一个跳跃, 从 $(a,a-a^2)$ 开始继续上升至 $+\infty$.

要使水平直线 $y=b$ 能与图像产生 3 个交点, 它必须同时穿过位于 $(0,a)$ 区间的上升段与位于 $(a,+\infty)$ 区间的上升段. 这要求在分段点 $x=a$ 处发生一个向下的跳跃, 从而形成一个局部的峰与谷.

这个向下的跳跃的代数条件是, 函数在分段点左侧的值大于其右极限.

$$f(a) > \lim_{x\to a^+} f(x)$$

$$a^2 > a-a^2$$

整理得 $2a^2-a>0$, 即 $a(2a-1)>0$.

结合本情形的假设 $a>0$, 上述不等式等价于 $2a-1>0$, 即 $a>1/2$.

当 $a>1/2$ 时, 确有 $a^2>a-a^2$. 此时, 可选取满足 $b>a-a^2$ 且 $0<b<a^2$ 的 b 值, 使水平直线 $y=b$ 与左段 $y=x^2$ 有两个交点, 与右段抛物线有一个交点, 从而产生 3 个交点.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

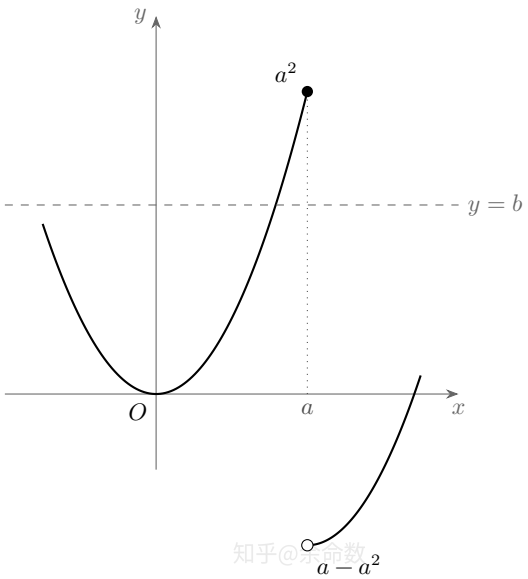


图 3.35　当 $a>1/2$ 时, 函数图像形态与水平线 $y=b$

知乎@余命数　知乎@余命数

反之, 若 $0<a\leq 1/2$, 则 $a^2\leq a-a^2$. 函数在 $x=0$ 之后是单调不减的, 任何水平直线至多有两个交点.

因此,函数 $g(x)$ 有 3 个零点存在的条件是 $a>1/2$. 故 a 的取值范围为 $(\frac{1}{2},+\infty)$.

例题 3.68　设 a 为实数, 函数 $f(x)=(x-a)^2+|x-a|-a|a-1|$.

1. 若 $f(0)>1$, 求 a 的取值范围;
2. 讨论 $f(x)$ 的单调性;
3. 当 $a>2$ 时, 讨论 $f(x)+|x|$ 在 $\mathbb{R}$ 上的零点个数.

证明 (1) 由函数表达式, $f(0)=(0-a)^2+|0-a|-a|a-1|=a^2+|a|-a|a-1|$. 条件 $f(0)>1$ 即 $a^2+|a|-a|a-1|>1$. 我们根据 a 的取值对绝对值进行分段讨论.

若 $a\geq 1$, 不等式为 $a^2+a-a(a-1)>1$, 即 $a^2+a-a^2+a>1$, 整理得 $2a>1$, 解得 $a>1/2$. 与 $a\geq 1$ 求交集, 得 $a\geq 1$.

若 $0\leq a<1$, 不等式为 $a^2+a-a(-(a-1))>1$, 即 $a^2+a+a^2-a>1$, 整理得 $2a^2>1$, 解得 $a>1/\sqrt{2}$ 或 $a<-1/\sqrt{2}$. 与 $0\leq a<1$ 求交集, 得 $1/\sqrt{2}<a<1$.

若 $a<0$, 不等式为 $a^2-a-a(-(a-1))>1$, 即 $a^2-a+a^2-a>1$, 整理得 $2a^2-2a-1>0$. 该二次不等式的解为 $a>\frac{1+\sqrt{3}}{2}$ 或 $a<\frac{1-\sqrt{3}}{2}$. 与 $a<0$ 求交集, 得 $a<\frac{1-\sqrt{3}}{2}$.

因此, a 的取值范围是 $(-\infty,\frac{1-\sqrt{3}}{2})\cup(\frac{1}{\sqrt{2}},+\infty)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2) 函数 $f(x)$ 可视为复合函数. 令 $t=|x-a|$, 则 $t\geq 0$. 原函数可表示为关于 t 的函数 $h(t)=t^2+t-a|a-1|$. 这是一个开口向上的二次函数, 对称轴为 $t=-1/2$. 由于 $t\geq 0$, 函数 $h(t)$ 在其定义域 $[0,+\infty)$ 上严格单调递增.

内层函数 $t(x)=|x-a|$ 在区间 $(-\infty,a]$ 上单调递减, 在区间 $[a,+\infty)$ 上单调递增. 根据复合函数单调性法则, 在区间 $(-\infty,a]$ 上, 内层函数递减而外层函数递增, 故 $f(x)$ 单调递减. 在区间 $[a,+\infty)$ 上, 内层函数递增而外层函数递增, 故 $f(x)$ 单调递增.

(3) 设 $g(x)=f(x)+|x|$. 当 $a>2$ 时, 我们讨论方程 $g(x)=0$ 的根的个数. 根据绝对值的零点 $x=0$ 和 $x=a$ 将函数写为分段形式.

当 $x<0$ 时, $g(x)=(x-a)^2-(x-a)-a(a-1)-x=x^2-(2a+2)x+2a$. 当 $0\leq x<a$ 时, $g(x)=(x-a)^2-(x-a)-a(a-1)+x=x^2-2ax+2a$. 当 $x\geq a$ 时, $g(x)=(x-a)^2+(x-a)-a(a-1)+x=x^2-(2a-2)x$.

我们分析函数 $g(x)$ 的单调性. 在 $(-\infty,0)$ 上, $g'(x)=2x-(2a+2)$. 由于 $x<0,a>2$, 有 $g'(x)<0$, 函数单调递减. 在 $(0,a)$ 上, $g'(x)=2x-2a=2(x-a)$. 由于 $x<a$, 有 $g'(x)<0$, 函数单调递减. 在 $(a,+\infty)$ 上, $g'(x)=2x-(2a-2)=2(x-(a-1))$. 由于 $x>a>a-1$, 有 $g'(x)>0$, 函数单调递增.

因此, 函数 $g(x)$ 在 $(-\infty,a]$ 上单调递减, 在 $[a,+\infty)$ 上单调递增. 于是 $g(x)$ 在 $x=a$ 处取得全局最小值.

我们计算几个关键点的函数值. $\lim_{x\to-\infty}g(x)=+\infty$. $g(0)=0^2-2a(0)+2a=2a$. 由于 $a>2$, 有 $g(0)>0$. $g(a)=a^2-(2a-2)a=-a^2+2a=-a(a-2)$. $\lim_{x\to+\infty}g(x)=+\infty$.

零点的个数取决于最小值 $g(a)$ 的符号.

由于 $a>2$, 有 $g(a)=-a(a-2)<0$. 又 $g(0)=2a>0$, 且 $g(x)$ 在 $[0,a]$ 上单调递减, 故函数在 $(0,a)$ 上恰有一个零点. 在 $(-\infty,0]$ 上, 函数由 $+\infty$ 单调递减到正数 $2a$, 因而没有零点. 在 $[a,+\infty)$ 上, 函数由负数 $g(a)$ 单调递增到 $+\infty$, 故在 $(a,+\infty)$ 上恰有一个零点.

因此, 当 $a>2$ 时, $g(x)=f(x)+|x|$ 在 $\mathbb{R}$ 上恒有 2 个零点. □

3.17 绝对值函数

绝对值是代数运算的基本工具, 也是距离这一几何概念的代数表达. 本节我们定义绝对值函数, 并讨论它的图像、性质与含绝对值的方程和不等式.

3.17.1 绝对值的定义与基本性质

我们先从代数定义出发.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 3.20 (绝对值) 实数 x 的**绝对值**, 记作 $|x|$, 定义为如下分段函数.

$$|x| = \begin{cases} x, & x \geq 0 \\ -x, & x < 0 \end{cases}$$

此定义的直接几何诠释是, $|x|$ 表示实数 x 在数轴上对应的点到原点的**距离**. 类似地, $|x-a|$ 表示数 x 与数 a 在数轴上对应点之间的距离. 这一几何观点常常用于化解复杂的绝对值问题.

由定义可直接导出绝对值的几个基本性质.

定理 3.29 (绝对值的基本性质) 对于任意实数 x, y, 以下性质恒成立.

1. **非负性**: $|x| \geq 0$.
2. **正定性**: $|x| = 0 \iff x = 0$.
3. **积性**: $|xy| = |x||y|$.
4. **三角不等式**: $|x+y| \leq |x| + |y|$. 等号成立的充要条件是 $xy \geq 0$.

证明 非负性由绝对值的定义立即得到. 若 $|x| = 0$, 则按定义只能有 $x = 0$; 反之 $x = 0$ 时 $|x| = 0$, 故正定性成立.

对于积性, $|xy|$ 与 $|x||y|$ 都非负, 且

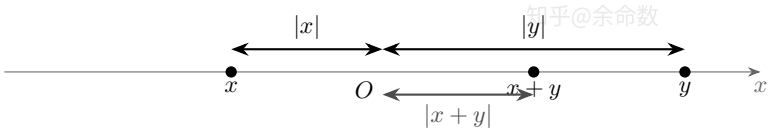
$$|xy|^2 = (xy)^2 = x^2y^2 = |x|^2|y|^2 = (|x||y|)^2.$$

两个非负数平方相等, 因而 $|xy| = |x||y|$.

对于三角不等式, 因为 $|x+y|$ 与 $|x| + |y|$ 都非负, 我们只需比较平方.

$$(|x| + |y|)^2 - |x+y|^2 = x^2 + 2|x||y| + y^2 - (x+y)^2 = 2(|x||y| - xy) \geq 0.$$

因此 $|x+y| \leq |x| + |y|$. 等号成立当且仅当 $|x||y| = xy$, 这等价于 $xy \geq 0$. □



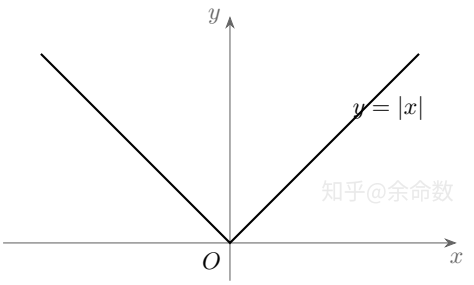
几何直观: 从原点出发, 先走 $|x|$ 再走 $|y|$ 的总路程,
不会小于直接到达终点 $x+y$ 的位移大小 $|x+y|$.

图 3.36　三角不等式的数轴诠释

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

3.17.2 绝对值函数图像与性质

函数 $f(x)=|x|$ 是最基本的分段函数之一, 其图像与性质是分析更复杂绝对值函数的基础.



图像呈 V 形, 关于 y 轴对称(偶函数).
在 $(-\infty,0]$ 上单调递减, 在 $[0,\infty)$ 上单调递增.
在原点处存在一个**尖点**, 函数在该点连续但不可导.

图 3.37 函数 $y=|x|$ 的图像

3.17.3 含绝对值的方程与不等式

求解含绝对值的方程与不等式时, 我们先化去绝对值. 常用的方法是**零点分段讨论法**, 它直接来自绝对值的定义.

该方法的步骤如下.

1. 找出所有绝对值符号内部表达式的零点.
2. 这些零点将整个实数轴划分为若干个开区间.
3. 在每个开区间内, 所有绝对值符号内部表达式的符号恒定, 从而可以去掉绝对值符号, 将原方程, 或不等式转化为一个不含绝对值的常规问题.
4. 分别求解各个区间上的问题, 并将解集与该区间的范围求交集, 得到该区间内的有效解.
5. 将所有区间上的有效解合并, 得到原问题的完整解集.

例题 3.69 求解不等式 $|2x-1|<x+2$.

解 我们先确定讨论的零点. 令绝对值内部的表达式 $2x-1=0$, 解得 $x=1/2$. 此点将实数轴划分为两个区间 $(-\infty,1/2)$ 和 $[1/2,+\infty)$.

情形一: $x\geq 1/2$. 在此区间内, $2x-1\geq 0$, 故 $|2x-1|=2x-1$. 原不等式转化为 $2x-1<x+2$, 解得 $x<3$. 将此解与本情形的约束 $x\geq 1/2$ 求交集, 得到该区间的解为 $[1/2,3)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

情形二：$x < 1/2$. 在此区间内，$2x-1<0$，故 $|2x-1| = -(2x-1) = 1-2x$. 原不等式转化为 $1-2x < x+2$，解得 $-1 < 3x$，即 $x > -1/3$. 将此解与本情形的约束 $x < 1/2$ 求交集，得到该区间的解为 $(-1/3, 1/2)$.

最后，我们将两个情形下得到的解集合并，得

$$[1/2, 3) \cup (-1/3, 1/2) = (-1/3, 3)$$

故原不等式的解集为 $(-1/3, 3)$.

对于特定形式的绝对值问题，也存在一些更简洁的等价转化.

- $|f(x)| < a$ $(a>0)$ 等价于 $-a < f(x) < a$.
- $|f(x)| > a$ $(a>0)$ 等价于 $f(x) > a$ 或 $f(x) < -a$.
- $|f(x)| = |g(x)|$ 等价于 $f(x)^2 = g(x)^2$，也等价于 $(f(x)-g(x))(f(x)+g(x)) = 0$.

导数讨论含绝对值参数

例题 3.70　已知函数 $f(x) = -x\ln x + ax$ 在 $(0,e)$ 上是增函数，函数 $g(x) = |e^x - a| + \frac{a^2}{2}$ 在 $x \in [0, \ln 3]$ 上的最大值 M 与最小值 m 的差为 $\frac{3}{2}$. 求 a 的值.

解　两个条件彼此独立，分头处理.

先看 f 的单调性. $f'(x) = a - 1 - \ln x$. f 在 $(0,e)$ 上递增要求 $f' \geq 0$ 恒成立，即 $a \geq 1 + \ln x$ 对一切 $x \in (0,e)$ 成立. $1+\ln x$ 在 $(0,e)$ 上递增，上界为 2，故必须 $a \geq 2$.

再看 $g(x) = |e^x - a| + \frac{a^2}{2}$ 在 $[0, \ln 3]$ 上的极值差. 绝对值零点为 $x = \ln a$；由 $a \geq 2$ 知 $\ln a \geq \ln 2 > 0$，且 g 在 $\ln a$ 左侧递减、右侧递增，最小值在 $x = \ln a$ 取得. 按 $\ln a$ 与 $[0, \ln 3]$ 的相对位置分类.

情形一：$a \geq 3$，即 $\ln a \geq \ln 3$. 此时 $\ln a$ 在区间右侧，g 在 $[0, \ln 3]$ 上递减，

$$M = g(0) = a - 1 + \frac{a^2}{2}, \quad m = g(\ln 3) = a - 3 + \frac{a^2}{2},$$

$M - m = 2 \neq \frac{3}{2}$，矛盾.

情形二：$2 \leq a < 3$. 此时 $\ln a \in (0, \ln 3)$，g 在 $[0, \ln a]$ 递减、在 $[\ln a, \ln 3]$ 递增，最小值 $m = \frac{a^2}{2}$. 最大值取自端点：$g(0) = a - 1 + \frac{a^2}{2}$，$g(\ln 3) = 3 - a + \frac{a^2}{2}$，且 $g(0) - g(\ln 3) = 2a - 4 \geq 0$，故 $M = g(0)$. 由 $M - m = \frac{3}{2}$ 得 $a - 1 = \frac{3}{2}$，$a = \frac{5}{2}$. 该值落在 $[2,3)$ 内，有效.

故 $a = \frac{5}{2}$.

3.17.4 距离和的最小值

绝对值函数的组合可以构造出一些优化问题中的经典模型. 其中最经典的是求解动点到若干定点距离之和的最小值问题.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 3.71 求函数 $f(x)=|x-1|+|x-3|$ 的最小值.

解 从几何的观点看, $f(x)$ 表示数轴上动点 x 到两个定点 1 和 3 的距离之和. 直观上, 当动点 x 位于两个定点之间时, 其距离之和恰好等于两定点间的距离, 此时应取得最小值. 我们用代数方法严格证明这一猜想.

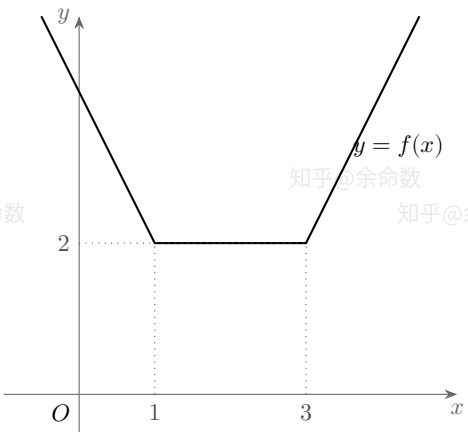
以两个零点 $x=1$ 和 $x=3$ 为分界, 我们对 x 的取值进行分段讨论.

情形一: $x\le 1$. 此时 $x-1\le 0, x-3<0$. $f(x)=-(x-1)-(x-3)=-2x+4$. 这是一个单调递减的函数.

情形二: $1<x<3$. 此时 $x-1>0, x-3<0$. $f(x)=(x-1)-(x-3)=2$. 函数在此区间上为常数.

情形三: $x\ge 3$. 此时 $x-1>0, x-3\ge 0$. $f(x)=(x-1)+(x-3)=2x-4$. 这是一个单调递增的函数.

把三段的分析合起来, 函数 $f(x)$ 的图像先下降, 然后在区间 $[1,3]$ 上保持为常数 2, 之后再上升. 因此, 函数的最小值为 2, 且在闭区间 $[1,3]$ 上的任意一点均可取得.



最小值在区间 $[1,3]$ 上取得

图 3.38 函数 $y=|x-1|+|x-3|$ 的图像

此例可以推广到更一般的情形. 对于函数 $f(x)=\sum_{i=1}^{n}|x-a_i|$, 其几何意义是动点 x 到 n 个定点 $a_1,\dots,a_n$ 的距离之和. 可以证明, 该函数的最小值在诸 a_i 的**中位数**处取得.

3.17.5 两点距离模型

函数 $f(x)=|x-1|+|x-3|$ 的图像形态具有一般性. 我们系统地考察两类与数轴上两点距离相关的函数模型.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3.17.5.1 距离和模型

考虑函数 $f(x)=|x-a|+|x-b|$, 其中 $a<b$. 此函数可视为数轴上动点 x 到两定点 a,b 的距离之和. 通过零点分段讨论, 我们得到其分段表达式

$$f(x)=\begin{cases}-2x+a+b, & x<a\\ b-a, & a\le x\le b\\ 2x-a-b, & x>b\end{cases}$$

函数图像由两条射线和一个水平线段构成, 形似平底锅. 其最小值为 $b-a$, 在区间 $[a,b]$ 上取得.

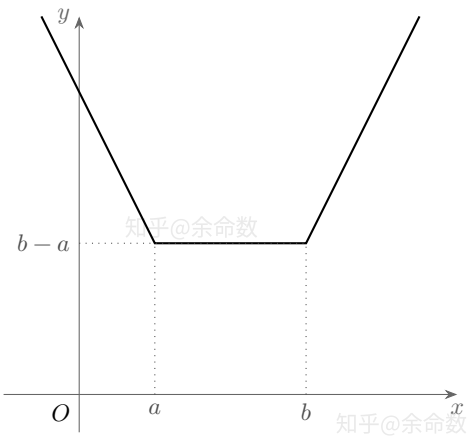


图 3.39　函数 $y=|x-a|+|x-b|$ 的图像(平底锅图像)

3.17.5.2 距离差模型

与距离和模型对应, 我们考察距离差模型 $g(x)=|x-a|-|x-b|$, 其中 $a<b$. 其分段表达式为

$$g(x)=\begin{cases}a-b, & x<a\\ 2x-a-b, & a\le x\le b\\ b-a, & x>b\end{cases}$$

函数图像由两条水平射线和一条斜率为 2 的线段构成. 函数的值域为 $[a-b,b-a]$.

3.17.6 作为范数的绝对值

对绝对值概念进行抽象和推广, 就得到近代数学中的**范数**. 在一个向量空间中, 范数是为每个向量赋予一个长度或大小的函数. 绝对值 $|x|$ 正是定义在一维实向量空间 $\mathbb{R}^1$ 上的范数. 我们之前列出的绝对值的基本性质, 是范数定义的三个公理.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

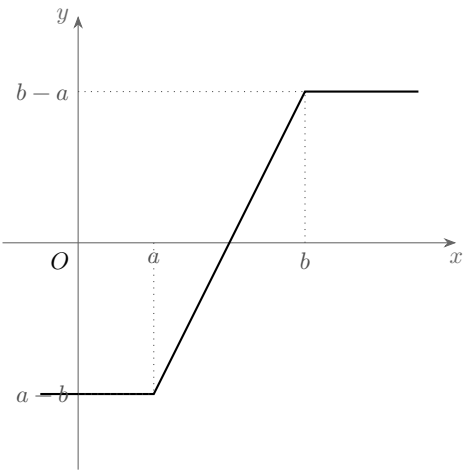


图 3.40 函数 $y=|x-a|-|x-b|$ 的图像(破锅图像)

- **正定性**:$\|v\| \geq 0$, 且 $\|v\| = 0 \iff v = 0$.
- **齐次性**:$\|\lambda v\| = |\lambda| \|v\|$(这推广了积性).
- **三角不等式**:$\|u+v\| \leq \|u\| + \|v\|$.

例如, 在二维平面 $\mathbb{R}^2$ 中, 一个向量 $\mathbf{v} = (x, y)$ 的欧几里得范数, 或称模长定义为 $\|\mathbf{v}\|_2 = \sqrt{x^2+y^2}$, 它同样满足上述三条公理. 这正是绝对值概念在更高维度上的自然推广.

3.18 函数解析式的求解

本节研究如何由已知条件确定一个函数的解析式. 已知条件可能是函数方程, 可能是图像经过的点, 也可能是奇偶性、周期性等内在性质. 我们根据这些条件的结构, 把它们转写成关于函数表达式的代数方程, 再求出这个表达式.

3.18.1 待定系数法, 当函数类型已知时

若题设已经指明函数是一次函数、二次函数或指数函数等, 那么解析式的结构便已确定, 未知的只是若干系数. 我们利用给定条件建立关于这些待定系数的方程组, 再解出它们即可.

例题 3.72 已知二次函数 $f(x)$ 满足 $f(2) = -1$, 且其图像的对称轴为直线 $x = 1$, 顶点在 x 轴下方, 与 x 轴两交点间的距离为 4. 求 $f(x)$ 的解析式.

解 我们先把几何信息转成代数约束. 对称轴 $x = 1$ 提示用顶点式, 设 $f(x) = a(x-1)^2 + k$ $(a \neq 0)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

与 x 轴两交点距离为 4, 且交点关于 $x=1$ 对称, 故两根横坐标 x_1, x_2 满足 $|x_1-x_2|=4$ 与 $\frac{x_1+x_2}{2}=1$, 解得 $x_1=-1, x_2=3$. 于是也可写成两根式

$$f(x)=a(x+1)(x-3),$$

同样内含对称轴 $x=1$. 代入过点 $(2,-1)$:

$$f(2)=a(2+1)(2-3)=-3a=-1 \implies a=\frac{1}{3}.$$

故

$$f(x)=\frac{1}{3}(x+1)(x-3)=\frac{1}{3}x^2-\frac{2}{3}x-1.$$

顶点 $(1,f(1))=(1,-\frac{4}{3})$ 在 x 轴下方, 与题设一致.

知乎@余命数

注 3.16 (选择恰当的函数形式)
二次函数的一般式、顶点式与两根式适用于不同的已知条件. 我们按已知信息选择计算量最小的形式.

例题 3.73　已知 $f(x)$ 为二次函数, 且 $f(0)=3$, $f(x+2)-f(x)=4x+2$. 求 $f(x)$.

知乎@余命数

解　函数类型已知为二次函数, 我们设其一般形式为 $f(x)=ax^2+bx+c$.

由 $f(0)=3$, 直接得 $c=3$. 故 $f(x)=ax^2+bx+3$.

现在我们处理函数方程 $f(x+2)-f(x)=4x+2$. 先计算 $f(x+2)$, 得

$$f(x+2)=a(x+2)^2+b(x+2)+3=a(x^2+4x+4)+b(x+2)+3$$

再构造差式, 得

$$\begin{aligned} f(x+2)-f(x)&=[a(x^2+4x+4)+b(x+2)+3]-[ax^2+bx+3]\\ &=(ax^2+4ax+4a+bx+2b+3)-(ax^2+bx+3)\\ &=4ax+(4a+2b) \end{aligned}$$

知乎@余命数

根据题设, 这个差式恒等于 $4x+2$:

$$4ax+(4a+2b)=4x+2.$$

这是关于 x 的恒等式. 由多项式恒等定理, 两边同次项系数相等. 比较一次项系数得 $4a=4$, 故 $a=1$. 比较常数项得 $4a+2b=2$. 将 $a=1$ 代入, 得 $4+2b=2$, 解得 $b=-1$.

至此, 所有待定系数都已确定. 故 $f(x)=x^2-x+3$.

知乎@余命数

知乎@余命数

3.18.2 利用函数性质法(奇偶性、周期性与对称性)

当函数的具体类型未知,但已知它具有某种全局性质时,这个性质就把函数在定义域一部分上的值与另一部分上的值联系起来. 我们利用这个联系写出解析式.

3.18.2.1 利用奇偶性

若已知奇函数或偶函数在部分定义域上的解析式, 我们便可用其定义把它延拓到对称的另一部分定义域上.

设 $x<0$. 为了求 $f(x)$, 我们先写 $f(-x)$, 并注意 $-x>0$. 这样 $f(-x)$ 可由已知解析式求出. 再由 $f(x)=f(-x)$, 即偶函数的情形, 或 $f(x)=-f(-x)$(奇函数)即可得到 $f(x)$.

例题 3.74 设函数 $f(x)$ 为定义在 $\mathbb{R}$ 上的奇函数,且当 $x\geq 0$ 时, $f(x)=e^x-1$. 求当 $x<0$ 时 $f(x)$ 的解析式.

解 设 $x<0$. 由奇函数定义, $f(x)=-f(-x)$.

因为 $-x>0$, 它落在已知解析式的区间 $[0,+\infty)$ 内. 将 $-x$ 代入已知表达式, 得

$$f(-x)=e^{-x}-1.$$

代回奇函数定义关系式, 得

$$f(x)=-f(-x)=-(e^{-x}-1)=1-e^{-x}.$$

此式对所有 $x<0$ 成立. 在 $x=0$ 处, 奇函数要求 $f(0)=0$, 而已知式给出 $f(0)=e^0-1=0$, 两者一致.

故函数 $f(x)$ 的完整解析式为

$$f(x)=\begin{cases}e^x-1, & x\geq 0\\ 1-e^{-x}, & x<0\end{cases}$$

3.18.2.2 利用周期性

若函数 $f(x)$ 是周期为 T 的周期函数, 那么它在任何一个长度为 T 的区间上的表达式, 就决定了它在整个定义域上的行为.

对任意自变量 x, 我们总可以加上或减去整数倍的周期 T, 把它平移到已知的基准区间内. 即

$$f(x)=f(x+nT),\quad n\in\mathbb{Z}.$$

我们要找一个整数 n, 使得 $x+nT$ 落入已知解析式的定义域.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 3.75　已知定义在 $\mathbb{R}$ 上的函数 $f(x)$ 满足 $f(x+2)=f(x)$，且当 $x\in[-1,1)$ 时，$f(x)=x^2$. 求当 $x\in[1,3)$ 时 $f(x)$ 的解析式.

解　条件 $f(x+2)=f(x)$ 表明 $f(x)$ 是一个周期为 2 的周期函数.

设 $x\in[1,3)$. 为了使用 $[-1,1)$ 上的已知解析式，我们把它平移，考虑 $x-2$. 当 $x\in[1,3)$ 时，$x-2\in[-1,1)$.

由周期性，$f(x)=f(x-2)$. 由于 $x-2$ 落在已知解析式的区间，代入 $f(t)=t^2$（取 $t=x-2$），得

$$f(x-2)=(x-2)^2.$$

因此，当 $x\in[1,3)$ 时，

$$f(x)=(x-2)^2=x^2-4x+4.$$

3.18.3　换元法与配凑法（处理复合函数）

当已知条件涉及复合函数，例如给出 $f(g(x))$ 的表达式而要求解 $f(x)$ 时，目标是消去内层函数 $g(x)$ 的影响，还原外层函数 f 对其自变量的作用. 实现这一目标的标准方法是换元法.

我们的做法是，把内层函数设为新的变量 $t=g(x)$. 变量 t 的取值范围就是 $f(t)$ 的定义域. 然后我们把右侧的表达式用 t 表示，通常是从 $t=g(x)$ 反解出 x 并代入. 最后把 $f(t)$ 中的 t 换为 x，就得到所求的解析式.

例题 3.76　已知函数 $f(\sqrt{x}+1)=x+2\sqrt{x}$，求 $f(x)$ 的解析式.

解　令 $t=\sqrt{x}+1$.

新变量 t 的取值范围由 $\sqrt{x}\geq 0$ 决定，故 $t\geq 1$.

从 $t=\sqrt{x}+1$ 反解出 $\sqrt{x}=t-1$，进而 $x=(t-1)^2$. 代入已知函数方程，得

$$f(t)=(t-1)^2+2(t-1).$$

展开化简，得

$$f(t)=(t^2-2t+1)+(2t-2)=t^2-1.$$

因此 $f(t)=t^2-1$，其中 $t\geq 1$. 换回变量 x，得

$$f(x)=x^2-1,\quad x\in[1,+\infty).$$

有时，从 $t=g(x)$ 中反解 x 是困难的或不便的. 此时我们直接看右侧表达式能否写成关于 $g(x)$ 的代数组合. 这种技巧称为配凑法.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 3.77 已知 $f(x+\frac{1}{x})=x^3+\frac{1}{x^3}$, 求 $f(x)$.

解 令 $t=x+\frac{1}{x}$. 不必反解 x, 我们直接处理右侧 $x^3+\frac{1}{x^3}$. 由恒等式 $(A+B)^3=A^3+B^3+3AB(A+B)$,

$$x^3+\frac{1}{x^3}=\left(x+\frac{1}{x}\right)^3-3\cdot x\cdot\frac{1}{x}\left(x+\frac{1}{x}\right)=t^3-3t.$$

故 $f(t)=t^3-3t$. 又 $x+\frac{1}{x}$ 的值域为 $(-\infty,-2]\cup[2,+\infty)$, 因而换元后的函数定义域为 $|x|\geq 2$, 解析式为 $f(x)=x^3-3x$.

例题 3.78 若 $f(\sin x)=2-\cos(2x)$, 则 $f(\cos x)=(\quad)$.

解 我们分两步. 先由已知关系求出 $f(x)$, 再把 $\cos x$ 代入所得解析式.

令 $t=\sin x$. 由于 $x\in\mathbb{R}$, 新变量 t 的取值范围是 $[-1,1]$.

把 $2-\cos(2x)$ 用 t 表示. 由二倍角公式, $\cos(2x)=1-2\sin^2 x$. 代入原方程, 得

$$f(\sin x)=2-(1-2\sin^2 x)=1+2\sin^2 x.$$

用 t 替换 $\sin x$, 得

$$f(t)=1+2t^2,\quad t\in[-1,1].$$

因此 $f(x)=1+2x^2$, 定义域为 $[-1,1]$.

现在计算 $f(\cos x)$. 将 x 替换为 $\cos x$, 得

$$f(\cos x)=1+2\cos^2 x.$$

再用二倍角公式 $2\cos^2 x=1+\cos(2x)$, 得

$$f(\cos x)=1+(1+\cos(2x))=2+\cos(2x).$$

故 $f(\cos x)=2+\cos(2x)$.

例题 3.79 已知 $f\left(\frac{1-x}{1+x}\right)=\frac{1-x^2}{1+x^2}$, 则 $f(x)$ 的解析式为 $(\quad)$.

解 令 $t=\frac{1-x}{1+x}$.

我们从这个关系式中反解 x. 由

$$t(1+x)=1-x$$

$$t+tx=1-x$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$tx + x = 1 - t$$

$$x(t+1) = 1 - t$$

得

$$x = \frac{1-t}{1+t}.$$

将 $x = \frac{1-t}{1+t}$ 代入 $\frac{1-x^2}{1+x^2}$, 得

$$1 - x^2 = 1 - \left(\frac{1-t}{1+t}\right)^2 = \frac{(1+t)^2 - (1-t)^2}{(1+t)^2} = \frac{4t}{(1+t)^2}$$

$$1 + x^2 = 1 + \left(\frac{1-t}{1+t}\right)^2 = \frac{(1+t)^2 + (1-t)^2}{(1+t)^2} = \frac{2(1+t^2)}{(1+t)^2}$$

两式相除, 得

$$\frac{1-x^2}{1+x^2} = \frac{4t/(1+t)^2}{2(1+t^2)/(1+t)^2} = \frac{4t}{2(1+t^2)} = \frac{2t}{1+t^2}.$$

因此 $f(t) = \frac{2t}{1+t^2}$. 由于 $t = \frac{1-x}{1+x}$ 不可能取到 -1, 该解析式在 $t \neq -1$ 的范围内确定.

将自变量换回 x, 得到 $f(x) = \frac{2x}{1+x^2}$ $(x \neq -1)$.

3.18.4 方程组法(利用对称性与对合关系)

当未知函数 $f(x)$ 出现在一个包含 $f(-x)$ 或 $f(1/x)$ 等项的方程中时, 我们可以对自变量作代换, 构造出另一个方程, 形成一个方程组. 把 $f(x)$ 等视为未知量, 解这个方程组就得到 $f(x)$.

这里的代换 $x \to -x$ 和 $x \to 1/x$ 都具有对合性质, 即连续施行两次就回到原变量. 我们常用它们来构造方程组.

例题 3.80　设函数 $f(x)$ 满足 $2f(x) + f(-x) = 3x + 4$, 求 $f(x)$ 的解析式.

解　方程中同时出现 x 与 $-x$. 我们将原方程中的 x 换为 $-x$, 得到一个新方程.

将 x 替换为 $-x$, 原方程变为

$$2f(-x) + f(-(-x)) = 3(-x) + 4.$$

化简得

$$2f(-x) + f(x) = -3x + 4.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将原方程与这个新方程联立, 得到关于 $f(x)$ 和 $f(-x)$ 的二元一次方程组

$$\begin{cases} 2f(x)+f(-x)=3x+4 & (1) \\ f(x)+2f(-x)=-3x+4 & (2) \end{cases}$$

为消去 $f(-x)$, 将方程 (1) 两边乘以 2, 再减去方程 (2):

$$(4f(x)+2f(-x))-(f(x)+2f(-x))=2(3x+4)-(-3x+4)$$

$$3f(x)=(6x+8)-(-3x+4)=9x+4$$

解得 $f(x)=3x+\frac{4}{3}$.

例题 3.81 已知 $f(x)$ 的定义域为 $\{x \mid x \neq 0\}$, 满足 $3f(x)+5f\left(\frac{1}{x}\right)=\frac{3}{x}+1$. 求函数 $f(x)$.

解 此方程涉及 $f(x)$ 与 $f(1/x)$, 且代换 $x \to 1/x$ 具有对合性.

将原方程记为 (1):

$$3f(x)+5f\left(\frac{1}{x}\right)=\frac{3}{x}+1 \quad (1)$$

在方程 (1) 中, 将所有 x 替换为 $1/x$, 得到新方程 (2):

$$3f\left(\frac{1}{x}\right)+5f(x)=3x+1 \quad (2)$$

将 $f(x)$ 和 $f(1/x)$ 视为未知量, 联立方程组

$$\begin{cases} 3f(x)+5f(1/x)=3/x+1 \\ 5f(x)+3f(1/x)=3x+1 \end{cases}$$

为消去 $f(1/x)$, 将 (1) 式乘以 3, (2) 式乘以 5:

$$9f(x)+15f(1/x)=9/x+3$$

$$25f(x)+15f(1/x)=15x+5$$

两式相减, 得到 $16f(x)=(15x+5)-(9/x+3)=15x-9/x+2$.

故 $f(x)=\frac{15}{16}x-\frac{9}{16x}+\frac{1}{8}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 3.82　已知函数 $f(x)$ 对定义域 $\{x \mid x \neq 0\}$ 内的任意实数 x 满足 $f(x) - 2f\left(\frac{2}{x}\right) = 4x$, 则 $f(x) = (\quad)$.

解　此方程涉及 $f(x)$ 与 $f(2/x)$. 若令 $g(x) = 2/x$, 则 $g(g(x)) = 2/(2/x) = x$, 因此代换 $x \to 2/x$ 同样是对合的.

原方程为

$$f(x) - 2f\left(\frac{2}{x}\right) = 4x \quad (1)$$

将 x 替换为 $2/x$:

$$f\left(\frac{2}{x}\right) - 2f\left(\frac{2}{2/x}\right) = 4\left(\frac{2}{x}\right).$$

化简得

$$f\left(\frac{2}{x}\right) - 2f(x) = \frac{8}{x} \quad (2)$$

联立方程组

$$\begin{cases} f(x) - 2f(2/x) = 4x \\ -2f(x) + f(2/x) = 8/x \end{cases}$$

为消去 $f(2/x)$, 将 (2) 式乘以 2, 再与 (1) 式相加:

$$(f(x) - 2f(2/x)) + (-4f(x) + 2f(2/x)) = 4x + 2\left(\frac{8}{x}\right)$$

$$-3f(x) = 4x + \frac{16}{x}$$

解得 $f(x) = -\frac{4}{3}x - \frac{16}{3x}$.

3.18.5 赋值法与递推(探索函数方程)

有时我们无法预知函数类型. 这时我们可以通过给变量赋特殊值来简化方程, 逐步发现函数的性质或递推关系. 这些值可以是 $0, 1, -1$, 也可以令 $y = x$ 或 $y = -x$ 等.

例题 3.83　已知函数 $f(x)$ 对任意实数 x, y 均满足 $f(x+y) = f(x)+f(y)+2xy$, 且 $f(1) = 1$. 若 $f(x)$ 为多项式函数, 求 $f(x)$ 的解析式.

解　这是一个柯西型函数方程. 我们用特殊赋值探寻它的结构.

令 $x = y = 0$, 得 $f(0) = f(0) + f(0) + 0$, 故 $f(0) = 0$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $y=-x$, 得 $f(0)=f(x)+f(-x)-2x^2$. 由于 $f(0)=0$, 故 $f(x)+f(-x)=2x^2$.

令 $y=1$, 得 $f(x+1)=f(x)+f(1)+2x$. 代入 $f(1)=1$, 得一阶差分方程

$$f(x+1)-f(x)=2x+1.$$

知乎@余命数

我们考察函数在整数点上的取值. $f(0)=0$. $f(1)=1$. $f(2)=f(1)+2(1)+1=1+3=4$. $f(3)=f(2)+2(2)+1=4+5=9$.

我们看到 $f(n)=n^2$ 对 $n=0,1,2,3$ 成立. 题设限定 $f(x)$ 为多项式函数, 且差分 $f(x+1)-f(x)=2x+1$ 是一次式, 因此 $f(x)$ 至多为二次多项式. 设 $f(x)=ax^2+bx+c$, 由 $f(x+1)-f(x)=2ax+(a+b)=2x+1$ 得 $a=1, b=0$, 又 $f(0)=0$ 得 $c=0$. 因而 $f(x)=x^2$.

现在验证这个结果. 原方程左侧为 $f(x+y)=(x+y)^2=x^2+2xy+y^2$, 右侧为 $f(x)+f(y)+2xy=x^2+y^2+2xy$. 两侧相等, 且 $f(1)=1^2=1$ 也满足.

因此, 函数的解析式为 $f(x)=x^2$.

注 3.17 (从离散到连续)
赋值法先在特殊点上给出猜想, 再用代数恒等式验证它.

例题 3.84 已知函数 $f(x)$ 的定义域为 $\mathbb{R}$, 且 $f(x+y)+f(x-y)=2f(x)f(y)$, $f(0)=1$. 请写出一个满足条件的函数 $f(x)$.

解 这是达朗贝尔函数方程. 我们用特殊赋值分析其性质.

令 $x=0$, 得 $f(y)+f(-y)=2f(0)f(y)=2f(y)$. 化简得 $f(-y)=f(y)$, 故 $f(x)$ 是一个偶函数.

令 $y=x$, 得 $f(2x)+f(0)=2(f(x))^2$. 代入 $f(0)=1$, 得 $f(2x)=2(f(x))^2-1$.

这个关系式与二倍角公式 $\cos(2x)=2\cos^2 x-1$ 完全相同.

我们猜测 $f(x)=\cos(\alpha x)$ 可能是一个解. 代入 $f(0)=1$ 的条件, 得 $f(0)=\cos(0)=1$, 满足.

现在验证它是否满足原方程. 左侧为

$$\cos(\alpha(x+y))+\cos(\alpha(x-y))=\cos(\alpha x+\alpha y)+\cos(\alpha x-\alpha y).$$

利用和差化积公式, 左侧等于 $2\cos(\alpha x)\cos(\alpha y)$, 右侧也是 $2\cos(\alpha x)\cos(\alpha y)$. 两侧恒等.

因此 $f(x)=\cos(\alpha x)$ 对任意实数 α 都是一个解. 题目只要求写出一个解, 我们取最简单的情形 $\alpha=1$.

故一个满足条件的函数是 $f(x)=\cos x$.

例题 3.85 已知函数 $f(x)$ 满足 $f(x-y)=f(x)+y(y-2x+5)$, 且 $f(-1)=3$. 求 $f(x)$ 的解析式.

解 方程右侧同时含有 $f(x), x, y$. 我们选取合适的赋值消去其中一个变量.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $x-y=-1$, 则 $y=x+1$. 代入原方程, 得

$$f(-1)=f(x)+(x+1)((x+1)-2x+5).$$

知乎@余命数

代入 $f(-1)=3$, 得

$$3=f(x)+(x+1)(6-x).$$

整理得

$$3=f(x)-x^2+5x+6.$$

由此解出

$$f(x)=x^2-5x-3.$$

知乎@余命数

验证: 对该函数,

$$f(x-y)-f(x)=y^2-2xy+5y=y(y-2x+5),$$

且 $f(-1)=1+5-3=3$, 满足题设.

这一步的做法是令 $x-y$ 为常数, 从而直接利用 $f(-1)=3$.

3.19 函数值域的求解

函数的值域是所有输出值构成的集合. 它说明函数在纵向上的范围, 也关系到方程与不等式的解集. 求值域分析的是必然输出, 而求定义域分析的是合法输入.

求值域要结合函数表达式的代数结构与几何特征, 选用合适的分析工具, 在代数变形的严谨性与几何直观的启发性之间建立联系.

注 3.18
可导函数在临界点取最值见 第16章; 本节主线是换元、配方、定义法与基本不等式.

3.19.1 观察法与基本函数

基本初等函数的值域, 是求更复杂函数值域的基础.

- 线性函数 $y=ax+b\ (a\neq 0)$: 值域为 $\mathbb{R}$.
- 二次函数 $y=ax^2+bx+c\ (a\neq 0)$: 值域为 $[\frac{4ac-b^2}{4a},+\infty)$ 当 $a>0$, 或 $(-\infty,\frac{4ac-b^2}{4a}]$(当 $a<0$).
- 反比例函数 $y=k/x\ (k\neq 0)$: 值域为 $(-\infty,0)\cup(0,+\infty)$.
- 指数函数 $y=a^x$: 值域为 $(0,+\infty)$.
- 对数函数 $y=\log_a x$: 值域为 $\mathbb{R}$.
- 正弦与余弦函数 $y=\sin x, y=\cos x$: 值域为 $[-1,1]$.

导数求值域

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

3.19.2 分析法

知乎@余命数

若函数可导, 我们可用导数求它在区间上的最大值与最小值, 从而得到值域. 连续函数在闭区间上的最值必在端点或临界点处取得.

我们先求出函数的导数 $f'(x)$, 并解出所有临界点, 即 $f'(x)=0$ 的根. 再比较函数在这些临界点以及定义域端点处的函数值, 综合得出最大值与最小值.

例题 3.86 求函数 $f(x)=x+\frac{4}{x}$ 在区间 $[1,4]$ 上的值域.

解 先考察函数在指定区间内的变化趋势. 其导函数为

$$f'(x)=1-\frac{4}{x^2}=\frac{x^2-4}{x^2}.$$

知乎@余命数

在区间 $[1,4]$ 内, 令 $f'(x)=0$, 解得 $x=2$. 此为区间内唯一的临界点.

只需比较函数在区间端点 $(x=1, x=4)$ 与临界点 $(x=2)$ 处的函数值:

$$f(1)=1+\frac{4}{1}=5$$

$$f(2)=2+\frac{4}{2}=4$$

$$f(4)=4+\frac{4}{4}=5$$

知乎@余命数

比较可知, 函数在区间 $[1,4]$ 上的最大值为 5, 最小值为 4. 因此, 函数的值域为 $[4,5]$.

知乎@余命数

3.19.3 代数法

我们用恒等变形把函数化成已知值域的结构, 或者把 $y=f(x)$ 看作关于 x 的方程, 用有解条件确定 y 的范围.

3.19.3.1 换元法与配方法

当函数可以视为某个中间变量的简单函数(特别是二次函数)时, 换元法便成为一种有力工具. 我们要准确地确定中间变量的取值范围, 这将成为外层函数的新定义域.

例题 3.87 求函数 $f(x)=\sin^2 x+2\sin x+3$ 的值域.

解 此函数可被看作一个关于 $\sin x$ 的二次函数.

进行换元, 令 $t=\sin x$. 新变量 t 的取值范围由正弦函数的值域 $[-1,1]$ 决定, 故 $t\in[-1,1]$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

原问题等价于求二次函数 $g(t)=t^2+2t+3$ 在 $t\in[-1,1]$ 上的值域.

对 $g(t)$ 配方, 以展示顶点位置:

$$g(t)=(t+1)^2+2.$$

抛物线开口向上, 对称轴为 $t=-1$.

由于对称轴 $t=-1$ 恰好是区间 $[-1,1]$ 的左端点, 函数 $g(t)$ 在该区间上严格单调递增. 因此最小值在 $t=-1$ 处取得, 最大值在 $t=1$ 处取得:

$$g_{\min}=g(-1)=(-1+1)^2+2=2$$

$$g_{\max}=g(1)=(1+1)^2+2=6$$

故原函数 $f(x)$ 的值域为 $[2,6]$.

知乎@余命数

3.19.3.2 反函数法(方程法)

我们也可以换一个视角. 方程法把 $y=f(x)$ 视为关于 x 的方程. 函数的值域, 就是使该方程有实数解的所有 y 构成的集合.

例题 3.88 求函数 $f(x)=\frac{x^2-x+1}{x^2+x+1}$ 的值域.

知乎@余命数

解 设 $y=\frac{x^2-x+1}{x^2+x+1}$. 将此式看作关于 x 的方程并反解 x.

整理方程, 消去分母, 得

$$y(x^2+x+1)=x^2-x+1$$

知乎@余命数

$$yx^2+yx+y=x^2-x+1.$$

将所有项移至一边, 得到关于 x 的标准一元二次方程形式

$$(y-1)x^2+(y+1)x+(y-1)=0.$$

函数的值域就是使得这个关于 x 的方程有实数解的所有 y 的集合.

对系数 $(y-1)$ 分类讨论.

情形一: $y-1=0$**, 即** $y=1$. 此时方程退化为 $2x=0$, 解得 $x=0$. 故 $x=0$ 是一个合法的实数解, $y=1$ 属于值域.

情形二: $y-1\neq0$**, 即** $y\neq1$. 此时方程是一元二次方程. 它有实数解的充要条件是其判别式 $\Delta\geq0$:

知乎@余命数

$$\Delta=(y+1)^2-4(y-1)(y-1)\geq0$$

$$(y+1)^2-(2(y-1))^2\geq0.$$

利用平方差公式分解:

$$((y+1)-2(y-1))((y+1)+2(y-1)) \geq 0$$

$$(-y+3)(3y-1) \geq 0$$

$$(y-3)(3y-1) \leq 0.$$

解此不等式, 得到 $\frac{1}{3} \leq y \leq 3$.

综合两种情形并注意到 $y \neq 1$, 得 y 的取值范围是 $[\frac{1}{3}, 1) \cup (1, 3] \cup \{1\}$.

合并这些集合, 最终得到函数的值域为 $[\frac{1}{3}, 3]$.

3.19.3.3 不等式法

在处理分式或根式函数时, 基本不等式, 如均值不等式可用来确定函数值的边界.

例题 3.89 求函数 $f(x) = x + \frac{4}{x}$ 的值域.

解 此函数的定义域为 $(-\infty, 0) \cup (0, +\infty)$. 我们分情况讨论.

情形一: $x > 0$. 此时 x 与 $4/x$ 均为正数, 可应用算术-几何平均值不等式:

$$f(x) = x + \frac{4}{x} \geq 2\sqrt{x \cdot \frac{4}{x}} = 2\sqrt{4} = 4.$$

等号成立的条件是 $x = \frac{4}{x}$, 即 $x^2 = 4$. 由于 $x > 0$, 解得 $x = 2$. 因此当 $x > 0$ 时, 函数的值域为 $[4, +\infty)$.

情形二: $x < 0$. 此时 x 与 $4/x$ 均为负数. 我们可对它们的相反数应用均值不等式. 令 $x = -t$, 其中 $t > 0$.

$$f(x) = -t + \frac{4}{-t} = -\left(t + \frac{4}{t}\right).$$

由于 $t > 0$, 已知 $t + \frac{4}{t} \geq 4$. 因此 $f(x) = -\left(t + \frac{4}{t}\right) \leq -4$. 等号在 $t = 2$, 即 $x = -2$ 时取得. 因此当 $x < 0$ 时, 函数的值域为 $(-\infty, -4]$.

综合两种情形, 函数 $f(x)$ 的完整值域为 $(-\infty, -4] \cup [4, +\infty)$.

3.19.4 数形结合法

函数的值域在几何上表现为图像在 y 轴上的投影. 我们也可以先画出函数的草图, 再观察纵向分布范围, 并用代数方法确认它.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

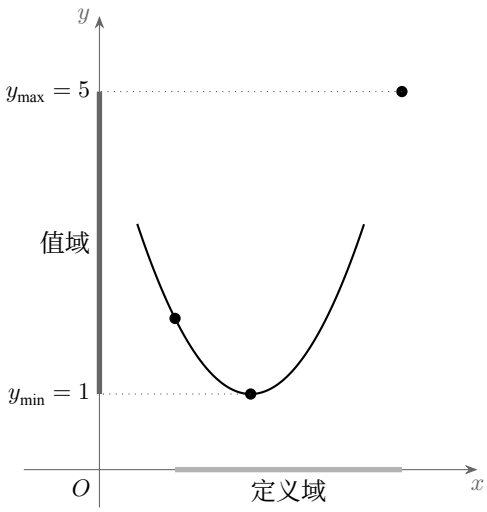


图 3.41　函数 $y=(x-2)^2+1$ 在定义域 $[1,4]$ 上的值域 $[1,5]$

3.20 多项式理论

多项式是代数运算的基本对象. 本节在数域上建立多项式的整除、因式分解与对称函数理论, 给出重因式的导数判别法, 并用 Lagrange 插值与切比雪夫多项式展示多项式的逼近作用.

3.20.1 数域与多项式

定义 3.21 (数域) 若数集 P 包含 0 与 1, 且对其中任意两数作加、减、乘、除(除数非零)所得结果仍在 P 中, 则称 P 为一个**数域**.

有理数集 $\mathbb{Q}$、实数集 $\mathbb{R}$、复数集 $\mathbb{C}$ 都是数域. 数域也能由无理数生成, 把 α 与 $\mathbb{Q}$ 中的数作四则运算得到的一切数记为 $\mathbb{Q}(\alpha)$, 它是数域. 当 α 是二次无理数时 $\mathbb{Q}(\alpha)=\{a+b\alpha \mid a,b\in\mathbb{Q}\}$, 以 $\alpha=\sqrt{2}$ 为例, $(a+b\sqrt{2})(c+d\sqrt{2})=(ac+2bd)+(ad+bc)\sqrt{2}$, 非零元素的逆 $\frac{a-b\sqrt{2}}{a^2-2b^2}$ 也在其中. $\alpha=\sqrt[3]{2}$ 时 $\alpha^2=\sqrt[3]{4}$ 落在 $\{a+b\sqrt[3]{2}\mid a,b\in\mathbb{Q}\}$ 之外, 这个二维集合对乘法不封闭, $\mathbb{Q}(\sqrt[3]{2})=\{a+b\sqrt[3]{2}+c\sqrt[3]{4}\mid a,b,c\in\mathbb{Q}\}$ 才是数域.

定义 3.22 (P 上多项式) 系数取自数域 P 的多项式 $f(x)=a_nx^n+\cdots+a_1x+a_0$ 称为 P 上的多项式. 若 $a_n\neq 0$, 称 n 为 f 的次数, 记 $\deg f=n$; 零多项式不定义次数.

本节中多项式均指某一固定数域 P 上的多项式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

整数有带余除法, 17 除以 5 得商 3 余 2. 多项式有完全平行的结构. 整除、最大公因式、因式分解这三件事都从带余除法生长出来.

3.20.2 带余除法与辗转相除

定理 3.30 (带余除法) 对 P 上任意非零多项式 $f(x)$ 与 $g(x)$, 存在唯一的 $q(x), r(x) \in P[x]$, 使得

$$f(x) = q(x)\,g(x) + r(x), \quad \deg r < \deg g \text{ 或 } r \equiv 0. \tag{3-19}$$

证明 存在性. 若 $\deg f < \deg g$, 取 $q = 0, r = f$ 即可. 设 $\deg f = n \geq m = \deg g$, f 与 g 的首项系数分别为 a_n、b_m. 令 $q_0(x) = \frac{a_n}{b_m}x^{n-m}$, 则 $f_1 = f - q_0 g$ 的次数低于 n. 对 f_1 重复此操作, 每次至少降一次, 有限步后得到 f_r 满足 $\deg f_r < \deg g$ 或 $f_r \equiv 0$. 累加各步的 q_i 即得 q, 取 $r = f_r$.

唯一性. 设 $f = q_1 g + r_1 = q_2 g + r_2$, 则 $(q_1 - q_2)g = r_1 - r_2$. 左端若非零则次数 $\geq \deg g$, 右端次数 $< \deg g$, 故两端为零, 即 $q_1 = q_2, r_1 = r_2$. □

$q(x)$ 称为商, $r(x)$ 称为余式. 当 $r \equiv 0$ 时称 g 整除 f, 记 $g \mid f$.

定义 3.23 (最大公因式) 若 $d(x)$ 同时整除 $f(x)$ 与 $g(x)$, 且 f、g 的任一公因式都整除 $d(x)$, 则称 $d(x)$ 为 f 与 g 的最大公因式. 最大公因式在相差非零常数倍的意义下唯一, 通常取其中的首一多项式, 记作 (f, g). 当 $(f, g) = 1$ 时称 f 与 g 互素.

定理 3.31 (Bézout 等式) 设 $d(x) = (f(x), g(x))$, 则存在 $u(x), v(x) \in P[x]$ 使得

$$d(x) = u(x)\,f(x) + v(x)\,g(x). \tag{3-20}$$

证明 对 f、g 反复作带余除法, 得序列

$$\begin{aligned} f &= q_1 g + r_1, \\ g &= q_2 r_1 + r_2, \\ r_1 &= q_3 r_2 + r_3, \\ &\vdots \\ r_{s-2} &= q_s r_{s-1} + r_s, \\ r_{s-1} &= q_{s+1} r_s. \end{aligned}$$

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

每步余式次数严格递减, 有限步后余式为零. 由第一式起, f 与 g 的公因式整除 r_1; 由第二式, g 与 r_1 的公因式整除 r_2; 依次传递, r_s 整除所有前项, 故 r_s 是最大公因式. 从倒数第二式倒推, 把 r_s 表为 r_{s-1}、r_{s-2} 的线性组合, 再逐次用前式代入消去 $r_{s-1}, r_{s-2}, \dots$, 最终得到 (3-20) 的形式. □

上述反复作带余除法的过程即**辗转相除法**.

例题 3.90 (求最大公因式) 求 $f(x)=x^4+x^3-x-1$ 与 $g(x)=x^3+x^2-x-1$ 的最大公因式.

解 作带余除法: $f=(x)g+(x^2-1)$, $g=(x+1)(x^2-1)+0$. 故 $(f,g)=x^2-1$(差常数因子). 倒推得 $x^2-1=f-x\,g$, 即 $u=1, v=-x$.

有了最大公因式, 下一步是因式分解. 整数的因数分解到素数为止, $12=2^2\cdot 3$, 其中 2、3 不能再拆. 多项式也有不能再分的"素多项式", 即不可约多项式.

3.20.3 因式分解唯一定理

定义 3.24 (不可约多项式) 若 P 上次数 ≥ 1 的多项式 $p(x)$ 不能分解为 P 上两个次数更低的多项式之积, 则称 $p(x)$ 在 P 上不可约.

多项式是否不可约依赖于数域, x^2-2 在 $\mathbb{Q}$ 上不可约, 在 $\mathbb{R}$ 上可分解为 $(x-\sqrt{2})(x+\sqrt{2})$.

定理 3.32 (因式分解唯一定理) 数域 P 上任一次数 ≥ 1 的多项式 $f(x)$ 可分解为 P 上有限个不可约多项式之积

$$f(x)=c\,p_1(x)^{\alpha_1}p_2(x)^{\alpha_2}\cdots p_k(x)^{\alpha_k}, \tag{3-21}$$

其中 c 为首项系数, p_i 为首一不可约多项式, $\alpha_i\geq 1$. 此分解在不计因式顺序时唯一.

证明 存在性. 对 $\deg f$ 作归纳.$\deg f=1$ 时 f 本身不可约. 设 $\deg f=n$, 若 f 不可约则已证; 否则 $f=f_1f_2$, $\deg f_1, \deg f_2<n$, 由归纳假设 f_1、f_2 均可分解, 合并即得 f 的分解.

唯一性. 先记一条要反复使用的性质, 若不可约多项式 p 整除乘积 $q_1q_2\cdots q_l$, 则 p 整除某个 q_j. 这是因为 p 不可约, 它与每个 q_j 要么互素, 要么相差常数因子; 若 p 与所有 q_j 都互素, 由 Bézout 等式可知 p 与整个乘积互素, 与 $p\mid q_1\cdots q_l$ 矛盾.

设 $f=c\,p_1^{\alpha_1}\cdots p_k^{\alpha_k}=c'\,q_1^{\beta_1}\cdots q_l^{\beta_l}$. 由 $p_1\mid f$ 及上述性质, p_1 整除某个 q_j; q_j 不可约故 p_1 与 q_j 相差常数因子, 归一为首一后相等. 消去后继续, 建立 p_i 与 q_j 的一一

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对应. 设某不可约因式 h 在两分解中幂次分别为 α、β, 由 $h^{\alpha} \mid f$ 及 h 与其余因式互素得 $h^{\alpha} \mid h^{\beta}$, 故 $\alpha \le \beta$; 对称地 $\beta \le \alpha$, 故 $\alpha = \beta$. □

分解的唯一性是这一定理的实质, 同一个多项式无论怎样分解, 最终的不可约因式(不计顺序和常数因子)完全一致. 但“不可约”依赖于数域, 同一个多项式在小域里不可分, 到大域里就能拆开.

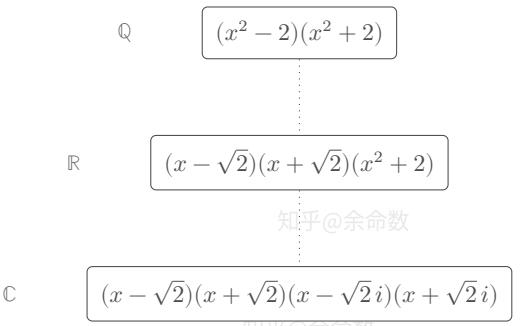


图 3.42 x^4-4 在不同数域上的因式分解, 数域越大, 不可约因式的次数越低, 直至 $\mathbb{C}$ 上完全分解为一次因式

3.20.4 重因式与导数判别

定义 3.25 (重因式) 若不可约多项式 $p(x)$ 满足 $p(x)^k \mid f(x)$ 但 $p(x)^{k+1} \nmid f(x)$, 则称 p 为 f 的 k 重因式. $k=1$ 时称单因式, $k \ge 2$ 时称重因式.

多项式的导数 $f'(x)$ 按通常求导规则定义. 导数给出重因式的判据.

定理 3.33 (重因式判别) 设 $p(x)$ 不可约, $k \ge 1$. 若 p 是 f 的 k 重因式, 则 $p^{k-1} \mid f'$ 而 $p^k \nmid f'$; 反之, 若 $p \mid f$ 且 $p^{k-1} \mid f'$, $p^k \nmid f'$, 则 p 是 f 的 k 重因式. 特别地, f 没有重因式当且仅当 $(f, f') = 1$.

证明 设 $f(x) = p(x)^k\, g(x)$, $(p, g) = 1$. 求导得

$$f'(x) = k\,p(x)^{k-1} p'(x)\, g(x) + p(x)^k\, g'(x) = p(x)^{k-1}\big[k\,p'(x)\, g(x) + p(x)\, g'(x)\big].$$

记括号内为 $h(x)$. 因 $\deg p' < \deg p$ 且 $(p, g) = 1$, $p \nmid p'g$, 故 $p \nmid h$, 即 $p^{k-1} \mid f'$ 但 $p^k \nmid f'$.

反向. 设 $p \mid f$ 且 $p^{k-1} \mid f'$, $p^k \nmid f'$. 记 p 在 f 中的重数为 $m \ge 1$. 由已证的正向部分, $p^{m-1} \mid f'$ 而 $p^m \nmid f'$. 同一不可约因式在 f' 中的重数唯一, 故 $m-1 = k-1$, 即 $m = k$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 3.91 (判别重因式) 判别 $f(x)=x^4+2x^3+2x^2+2x+1$ 是否有重因式.

解 $f'(x)=4x^3+6x^2+4x+2=2(2x^3+3x^2+2x+1)$. 对 f 与 f' 作辗转相除, 可得 $(f,f')=x+1$. 故 $x+1$ 是 f 的重因式; 它在 (f,f') 中是一重因式, 由判别定理它在 f 中是二重因式. 验证: $f=(x+1)^2(x^2+1)$, $f'=2(x+1)(2x^2+x+1)$, 两者不再有其他公共因式, 故 f 没有别的重因式.

重因式判据不必先分解, 只需算 f 与 f' 的最大公因式, 就能判断 f 是否有重根. 这一判据在数值计算和代数化简里都常用.

3.20.5 对称多项式

定义 3.26 (初等对称多项式) 对 n 个变元 $x_1,\dots,x_n$, 令

$$\begin{aligned}
\sigma_1 &= x_1+x_2+\cdots+x_n,\\
\sigma_2 &= \sum_{1\le i<j\le n} x_ix_j,\\
&\vdots\\
\sigma_k &= \sum_{1\le i_1<\cdots<i_k\le n} x_{i_1}\cdots x_{i_k},\\
&\vdots\\
\sigma_n &= x_1x_2\cdots x_n.
\end{aligned}$$

称 $\sigma_1,\dots,\sigma_n$ 为初等对称多项式. 对任意变元的置换保持不变的多项式称为对称多项式.

定理 3.34 (对称多项式基本定理) 任一对称多项式 $F(x_1,\dots,x_n)$ 可唯一地表为初等对称多项式 $\sigma_1,\dots,\sigma_n$ 的多项式.

此定理的构造性证明通过对 F 的单项式按字典序归纳, 逐步用 σ_k 替换首项, 此处从略.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 3.35 (牛顿恒等式) 记 $p_k = x_1^k + x_2^k + \cdots + x_n^k$ 为幂和, $\sigma_1, \ldots, \sigma_n$ 为初等对称多项式. 对 $k \geq 1$,

$$p_k - \sigma_1 p_{k-1} + \sigma_2 p_{k-2} - \cdots + (-1)^{k-1}\sigma_{k-1}p_1 + (-1)^k k\sigma_k = 0 \quad (k \leq n), \tag{3-22}$$

当 $k > n$ 时, 末项 $(-1)^k k\sigma_k$ 替换为 $(-1)^n \sigma_n p_{k-n}$.

知乎@

证明 对每个 x_i, 由韦达定理知 x_i 是 $t^n - \sigma_1 t^{n-1} + \sigma_2 t^{n-2} - \cdots + (-1)^n \sigma_n = 0$ 的根, 故

$$x_i^n = \sigma_1 x_i^{n-1} - \sigma_2 x_i^{n-2} + \cdots + (-1)^{n-1}\sigma_n.$$

乘以 x_i^{k-n} 并对 i 求和 $(k \geq n)$, 得 $p_k = \sigma_1 p_{k-1} - \sigma_2 p_{k-2} + \cdots + (-1)^{n-1}\sigma_n p_{k-n}$, 即 $k > n$ 的情形. 对 $k \leq n$, 对生成函数 $\prod(1 - x_i t) = \sum(-1)^k \sigma_k t^k$ 取对数后比较 t^k 系数即得 (3-22). □

牛顿恒等式把幂和 p_k 递归地表为初等对称多项式的多项式, 据此可化简一切对称式.

例题 3.92 (用牛顿恒等式求幂和) 设 x_1, x_2, x_3 满足 $\sigma_1 = 6$, $\sigma_2 = 11$, $\sigma_3 = 6$, 求 p_2 与 p_3.

知乎@余命数

解 由 (3-22), $k = 2$: $p_2 - \sigma_1 p_1 + 2\sigma_2 = 0$, $p_1 = \sigma_1 = 6$, 故 $p_2 = 6\cdot 6 - 2\cdot 11 = 36 - 22 = 14$.

知乎@余命数

$k = 3$: $p_3 - \sigma_1 p_2 + \sigma_2 p_1 - 3\sigma_3 = 0$, 故 $p_3 = 6\cdot 14 - 11\cdot 6 + 3\cdot 6 = 84 - 66 + 18 = 36$.

对称多项式从代数结构刻画根的集合. 反过来, 已知若干点上的函数值也能唯一确定一个多项式, 这由 Lagrange 插值给出.

3.20.6 Lagrange 插值

定理 3.36 (Lagrange 插值公式) 设 $x_0, x_1, \ldots, x_n$ 为数域 P 中 $n+1$ 个互异点, $y_0, y_1, \ldots, y_n$ 为给定值. 则存在唯一的次数 $\leq n$ 的多项式 $L(x)$ 满足 $L(x_i) = y_i (i = 0, \ldots, n)$, 且

$$L(x) = \sum_{i=0}^{n} y_i \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{x - x_j}{x_i - x_j}. \tag{3-23}$$

证明 存在性. 对每个 i, 令 $\ell_i(x) = \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}$. 则 $\ell_i(x_i) = 1$, $\ell_i(x_j) = 0 (j \neq i)$, 故 $L(x) = \sum y_i \ell_i(x)$ 满足插值条件, 且 $\deg L \leq n$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

唯一性. 设 L_1, L_2 均满足条件且 $\deg \le n$. 令 $H = L_1 - L_2$, 则 $H(x_i) = 0$ 对 $i = 0, \dots, n$ 成立, 即 H 有 $n+1$ 个不同根. 但 $\deg H \le n$, 故 $H \equiv 0$. □

例题 3.93 (构造插值多项式) 求次数 ≤ 2 的多项式 $L(x)$, 使 $L(0) = 1, L(1) = 2, L(2) = 0$.

解 由 (3-23),

$$\begin{aligned} L(x) &= 1 \cdot \frac{(x-1)(x-2)}{(0-1)(0-2)} + 2 \cdot \frac{x(x-2)}{(1-0)(1-2)} + 0 \cdot \frac{x(x-1)}{(2-0)(2-1)} \\ &= \frac{(x-1)(x-2)}{2} - 2x(x-2) \\ &= \frac{x^2 - 3x + 2}{2} - 2x^2 + 4x \\ &= -\frac{3}{2}x^2 + \frac{5}{2}x + 1. \end{aligned}$$

知乎@余命数

验证: $L(0) = 1$, $L(1) = -\frac{3}{2} + \frac{5}{2} + 1 = 2$, $L(2) = -6 + 5 + 1 = 0$.

Lagrange 插值用 $n+1$ 个点确定一个 n 次多项式, 但节点选取影响逼近误差. 等距节点在端点附近剧烈振荡 (Runge 现象), 而切比雪夫多项式的根作为节点可抑制这一现象.

3.20.7 切比雪夫多项式

知乎@余命数

由三倍角公式 $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$ 可见 $\cos 3\theta$ 是 $\cos\theta$ 的三次多项式. 一般地:

定理 3.37 (切比雪夫多项式) 对任意正整数 n, $\cos n\theta$ 可表为 $\cos\theta$ 的 n 次多项式 $T_n(x)$, 其首项系数为 2^{n-1}, 满足

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2x\,T_n(x) - T_{n-1}(x). \tag{3-24}$$

证明 由和差化积, $\cos(n+1)\theta + \cos(n-1)\theta = 2\cos n\theta\cos\theta$, 故

$$\cos(n+1)\theta = 2\cos\theta\cos n\theta - \cos(n-1)\theta.$$

设 $\cos n\theta = T_n(\cos\theta)$, $\cos(n-1)\theta = T_{n-1}(\cos\theta)$, 代入即得递推 (3-24). 首项系数满足 $a_{n+1} = 2a_n$, $a_1 = 1$, 故 $a_n = 2^{n-1}$. □

前几个切比雪夫多项式为

$$T_0(x) = 1, \qquad T_1(x) = x, \qquad T_2(x) = 2x^2 - 1,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$T_3(x)=4x^3-3x,\qquad T_4(x)=8x^4-8x^2+1,\qquad T_5(x)=16x^5-20x^3+5x.$$

知乎@余命数

定理 3.38 (切比雪夫多项式的根) $T_n(x)$ 的 n 个根为

$$x_k=\cos\frac{(2k-1)\pi}{2n},\quad k=1,2,\dots,n, \tag{3-25}$$

全部落在区间 $[-1,1]$ 上.

证明 $T_n(x_k)=0$ 当且仅当 $\cos n\theta_k=0$,即 $n\theta_k=\frac{(2k-1)\pi}{2}$,$\theta_k=\frac{(2k-1)\pi}{2n}$. 故 $x_k=\cos\theta_k$ 如 (3-25).$k=1,\dots,n$ 给出 n 个互异根, 且 $\theta_k\in(0,\pi)$ 保证 $x_k\in(-1,1)$. □

知乎@余命数

切比雪夫多项式在 $[-1,1]$ 上满足 $|T_n(x)|\le 1$,且其根均匀分布于 $[-1,1]$. 以 T_n 的根为插值节点可抑制 Runge 现象, 这是切比雪夫多项式用于函数逼近的根据.

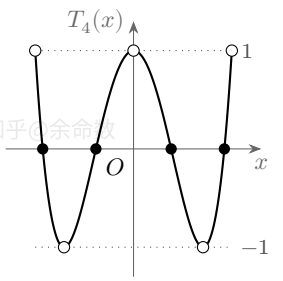


图 3.43 $T_4(x)=8x^4-8x^2+1$ 在 $[-1,1]$ 上的等波纹, 4 个根(实点)均匀分布, 5 个空点交替取 ±1

知乎@余命数 知乎@余命数 知乎@余命数

例题 3.94 (用切比雪夫多项式表示) 利用递推 (3-24) 求 $T_4(x)$, 并验证 $T_4(\cos\theta)=\cos 4\theta$.

解 $T_2=2x^2-1, T_3=2x(2x^2-1)-x=4x^3-3x, T_4=2x(4x^3-3x)-(2x^2-1)=8x^4-8x^2+1$.

验证: $\cos 4\theta=2\cos^2 2\theta-1=2(2\cos^2\theta-1)^2-1=8\cos^4\theta-8\cos^2\theta+1=T_4(\cos\theta)$.

3.21 三次与四次方程

知乎@余命数

二次方程的求根公式已为熟知. 三次与四次方程同样存在根式解, 本节给出 Cardano 公式、三次方程的三角解法, 以及四次方程的 Ferrari 降次法.

知乎@余命数 知乎@余命数

3.21.1 消去二次项

对一般三次方程 $ax^3+bx^2+cx+d=0(a\neq 0)$, 两端除以 a 化为

$$x^3+bx^2+cx+d=0.$$

作平移 $x=y-\dfrac{b}{3}$, 利用 $(y-\frac{b}{3})^3$ 展开后二次项系数为 $-b$, 与 $b(y-\frac{b}{3})^2$ 的二次项 b 相消, 方程化为不含二次项的形式

$$y^3+py+q=0. \tag{3-26}$$

此步骤称为 Vieta 平移. 以下对约化方程 (3-26) 求解.

消去二次项后, 方程只剩 y^3、y 和常数项.Cardano 的想法是把 y 拆成两个变量之和 $y=u+v$, 利用多出的自由度附加约束 $uv=-p/3$, 把三次方程降为关于 u^3 的二次方程.

3.21.2 Cardano 公式

定理 3.39 (Cardano 公式) 约化三次方程 $y^3+py+q=0$ 的根为

$$y_k=\omega^k u+\omega^{-k}v,\quad k=0,1,2, \tag{3-27}$$

其中 $\omega=e^{2\pi i/3}=-\dfrac{1}{2}+\dfrac{\sqrt{3}}{2}i$, $u=\sqrt[3]{-\dfrac{q}{2}+\sqrt{\Delta}}$, $v=\sqrt[3]{-\dfrac{q}{2}-\sqrt{\Delta}}$, $\Delta=\dfrac{q^2}{4}+\dfrac{p^3}{27}$, 且 u,v 满足 $uv=-\dfrac{p}{3}$.

证明 设 $y=u+v$, 代入 $y^3+py+q=0$ 得

$$(u^3+v^3)+(u+v)(3uv+p)+q=0.$$

u、v 为两个自由变量, 可附加约束 $3uv+p=0$, 即 $uv=-\frac{p}{3}$. 此时 $u^3+v^3=-q$. 令 $U=u^3, V=v^3$, 则

$$U+V=-q,\quad UV=(uv)^3=-\frac{p^3}{27}.$$

故 U、V 是二次方程 $t^2+qt-\frac{p^3}{27}=0$ 的两根:

$$U,V=\frac{-q\pm\sqrt{q^2+\frac{4p^3}{27}}}{2}=-\frac{q}{2}\pm\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}=-\frac{q}{2}\pm\sqrt{\Delta}.$$

知乎@余命数

取 $u=\sqrt[3]{U}, v=\sqrt[3]{V}$. 约束 $uv=-\frac{p}{3}$ 限定 u、v 的立方根搭配, 设 u_0 为 U 的一个立方根, 则 $v_0=-\frac{p}{3u_0}$ 为 V 的对应立方根.U 的三个立方根为 $u_0, \omega u_0, \omega^2 u_0$, 对应的 v 为 $v_0, \omega^2 v_0, \omega v_0$(因 $uv=-\frac{p}{3}$ 要求 v 的因子与 u 的因子乘积为 1). 故三根为

$$y_0=u_0+v_0, \quad y_1=\omega u_0+\omega^2 v_0, \quad y_2=\omega^2 u_0+\omega v_0,$$

即 (3-27). □

证明的关键一步是把 $y=u+v$ 拆成两个变量, 用 $uv=-p/3$ 消去交叉项, 剩下的 $u^3+v^3=-q$ 与 $u^3v^3=-p^3/27$ 合起来是二次方程的两根之和与积. 三次方程由此降为二次方程.

定义 3.27 (Vieta 替换) 约束 $uv=-\frac{p}{3}$ 等价于 $v=-\frac{p}{3u}$. 代换

$$y=u-\frac{p}{3u} \tag{3-28}$$

称为 Vieta 替换. 在此替换下, $y^3+py+q=0$ 化为

$$u^3-\frac{p^3}{27u^3}+q=0,$$

即关于 u^3 的二次方程 $U^2+qU-\frac{p^3}{27}=0$.

Vieta 替换把三次方程降为二次方程. 解出 u^3 的两个值后开立方得 u, 再由 (3-28) 还原 y. 注意 u^3 的两个值 U、V 满足 $V=-\frac{p^3}{27U}$, 故 u 与 $-\frac{p}{3u}$ 成对出现, 经 Vieta 替换给出相同的 y, 三对 u 值对应三个根.

例题 3.95 (用 Cardano 公式解方程) 解方程 $x^3-6x-6=0$.

解 此为约化形式, $p=-6, q=-6.\Delta=\frac{q^2}{4}+\frac{p^3}{27}=\frac{36}{4}+\frac{-216}{27}=9-8=1.$

$u=\sqrt[3]{-\frac{q}{2}+\sqrt{\Delta}}=\sqrt[3]{3+1}=\sqrt[3]{4}, v=\sqrt[3]{-\frac{q}{2}-\sqrt{\Delta}}=\sqrt[3]{3-1}=\sqrt[3]{2}.$

验证 $uv=\sqrt[3]{8}=2=-\frac{p}{3}=2$, 符合约束. 三根为

$$x_0=\sqrt[3]{4}+\sqrt[3]{2}, \quad x_1=\omega\sqrt[3]{4}+\omega^2\sqrt[3]{2}, \quad x_2=\omega^2\sqrt[3]{4}+\omega\sqrt[3]{2}.$$

因 $\Delta>0$, 方程有一个实根 x_0 与一对共轭复根 x_1, x_2.

3.21.3 三次方程的三角解法

当约化方程 $y^3+py+q=0$ 的判别式 $\Delta<0$, 即 $p<0$ 且 $\frac{3q}{2p}\sqrt{\frac{3}{-p}}$ 的绝对值 <1 时, 三个根均为实数, Cardano 公式中出现虚数的立方根(不可约情形). 此时用三角恒等变换可避开复数运算.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 3.40 (三次方程三角解法) 设约化方程 $y^3+py+q=0$ 中 $p<0$, 令 $y=2\sqrt{-\frac{p}{3}}\cos\theta$, 则方程化为

$$\cos 3\theta=\frac{3q}{2p}\sqrt{\frac{3}{-p}}.$$

若右端绝对值 ≤ 1, 则三个实根为

$$y_k=2\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2p}\sqrt{\frac{3}{-p}}\right)-\frac{2k\pi}{3}\right),\quad k=0,1,2. \tag{3-29}$$

证明 代入 $y=2\sqrt{-\frac{p}{3}}\cos\theta$, 记 $r=2\sqrt{-\frac{p}{3}}$. 则 $y^3=r^3\cos^3\theta$, $py=pr\cos\theta$. 由三倍角公式 $\cos 3\theta=4\cos^3\theta-3\cos\theta$, 即 $\cos^3\theta=\frac{\cos 3\theta+3\cos\theta}{4}$, 故

$$y^3+py=r^3\cdot\frac{\cos 3\theta+3\cos\theta}{4}+pr\cos\theta=\frac{r^3}{4}\cos 3\theta+\cos\theta\left(\frac{3r^3}{4}+pr\right).$$

由 $r=2\sqrt{-\frac{p}{3}}$, $r^2=-\frac{4p}{3}$, $\frac{3r^3}{4}=\frac{3}{4}\cdot r\cdot r^2=\frac{3}{4}\cdot r\cdot(-\frac{4p}{3})=-pr$, 故 $\frac{3r^3}{4}+pr=0$. 方程化为

$$\frac{r^3}{4}\cos 3\theta+q=0,\quad \cos 3\theta=-\frac{4q}{r^3}.$$

计算 $\frac{4q}{r^3}=\frac{4q}{8(-p/3)^{3/2}}=\frac{q}{2(-p/3)^{3/2}}=-\frac{3q}{2p}\sqrt{\frac{3}{-p}}$(因 $p<0$), 故 $\cos 3\theta=-\frac{4q}{r^3}=\frac{3q}{2p}\sqrt{\frac{3}{-p}}$. 设 $\varphi=\arccos\left(\frac{3q}{2p}\sqrt{\frac{3}{-p}}\right)$, 则 $3\theta=\pm\varphi+2k\pi$, $\theta=\frac{\varphi}{3}-\frac{2k\pi}{3}$, 取一组即可.$k=0,1,2$ 给出三个互异的 $\cos\theta$ 值, 对应三个实根 (3-29). □

例题 3.96 (三角解法求实根) 解方程 $x^3-3x+1=0$.

解 $p=-3, q=1.\Delta=\frac{1}{4}+\frac{(-3)^3}{27}=\frac{1}{4}-1=-\frac{3}{4}<0$, 属不可约情形, 用三角解法.

$r=2\sqrt{-\frac{p}{3}}=2\sqrt{1}=2.\cos 3\theta=\frac{3q}{2p}\sqrt{\frac{3}{-p}}=\frac{3}{-6}\sqrt{1}=-\frac{1}{2}$. 故 $3\theta=\pm\frac{2\pi}{3}+2k\pi$, $\theta=\frac{2\pi}{9}-\frac{2k\pi}{3}$.

三根为

$$x_k=2\cos\left(\frac{2\pi}{9}-\frac{2k\pi}{3}\right),\quad k=0,1,2.$$

即 $x_0=2\cos\frac{2\pi}{9}$, $x_1=2\cos\frac{4\pi}{9}$, $x_2=2\cos\frac{10\pi}{9}$, 均为实数.

当 $\Delta<0$ 时三个根都是实数, 但 Cardano 公式需经复数中转才能得到实根, 这是不可约情形. 三角解法利用 $\cos 3\theta$ 的三倍角公式, 直接在实数域中给出三个根.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

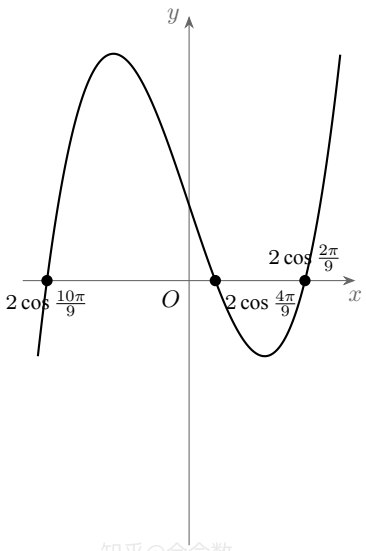


图 3.44 方程 $x^3-3x+1=0$ 的三个实根(不可约情形), Cardano 公式在此需经复数中转, 三角解法直接给出实根

3.21.4 四次方程的 Ferrari 降次法

对一般四次方程 $x^4+bx^3+cx^2+dx+e=0$, 先作平移 $x=y-\frac{b}{4}$ 消去三次项, 化为

$$y^4+py^2+qy+r=0. \tag{3-30}$$

定理 3.41 (Ferrari 降次法) 约化四次方程 (3-30) 可通过引入参数 t 分解为两个二次方程之积, t 满足一个三次预解方程. 解出 t 后, 四次方程的根由两个二次方程给出.

证明 将 (3-30) 改写为

$$y^4=-py^2-qy-r.$$

左端配方: $y^4+2ty^2+t^2=(2t-p)y^2-qy+(t^2-r)$, 即

$$(y^2+t)^2=(2t-p)y^2-qy+(t^2-r). \tag{3-31}$$

右端为 y 的二次式, 若其为完全平方, 则 (3-31) 两端开方后得到两个二次方程. 右端判别式为零的条件为

$$q^2-4(2t-p)(t^2-r)=0,$$

即

$$8t^3-4pt^2-8rt+4pr-q^2=0. \tag{3-32}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(3-32) 是关于 t 的三次方程, 称为预解方程. 任取其一实根 t_0, 则 (3-31) 右端为完全平方 $(\sqrt{2t_0-p}\,y-\frac{q}{2\sqrt{2t_0-p}})^2$(要求 $2t_0-p>0$). 开方得

$$y^2+t_0=\pm\left(\sqrt{2t_0-p}\,y-\frac{q}{2\sqrt{2t_0-p}}\right),$$

即两个二次方程

$$y^2-\sqrt{2t_0-p}\,y+t_0+\frac{q}{2\sqrt{2t_0-p}}=0,\quad y^2+\sqrt{2t_0-p}\,y+t_0-\frac{q}{2\sqrt{2t_0-p}}=0. \tag{3-33}$$

分别解之, 共得四个根. □

Ferrari 法的精髓是配方时引入参数 t, 左端 $(y^2+t)^2$ 总是完全平方, 右端是否完全平方取决于 t. 令右端判别式为零, 就把四次方程降为关于 t 的三次方程. 解出 t 后, 四次方程拆成两个二次方程.

例题 3.97 (用 Ferrari 法解方程) 解方程 $x^4-2x^2+8x-3=0$.

解 此为约化形式, $p=-2, q=8, r=-3$. 预解方程 (3-32):

$$8t^3-4(-2)t^2-8(-3)t+4(-2)(-3)-64=0,$$

即 $8t^3+8t^2+24t-40=0$, $t^3+t^2+3t-5=0$.

试根: $t=1$ 时 $1+1+3-5=0$, 故 $t_0=1$ 为实根. 此时 $2t_0-p=2-(-2)=4>0$, $\sqrt{2t_0-p}=2$.

代入 (3-33), 两方程为

$$y^2-2y+1+\frac{8}{4}=y^2-2y+3=0,\quad y^2+2y+1-\frac{8}{4}=y^2+2y-1=0.$$

第一个方程 $y^2-2y+3=0$, 判别式 $4-12=-8<0$, 根 $y=1\pm\sqrt{2}\,i$. 第二个方程 $y^2+2y-1=0$, 根 $y=-1\pm\sqrt{2}$.

故原方程四根为 $x_1=-1+\sqrt{2}, x_2=-1-\sqrt{2}, x_3=1+\sqrt{2}\,i, x_4=1-\sqrt{2}\,i$.

Ferrari 法把四次方程降为三次预解方程加两个二次方程, 做法是配方后令右端判别式为零. 三次与四次方程的根式解揭示了代数方程可解性的边界, 五次及以上的一般方程不存在根式解, 这一结论由 Abel 与 Galois 给出.

五次及以上方程没有根式解, 但数值上求根始终可行. 牛顿迭代法是其中最基本的方法, 过当前点 $(x_n, f(x_n))$ 作切线, 切线与 x 轴的交点作为下一个近似

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在根附近, 迭代以平方速度收敛, 每步有效位数大致翻倍.

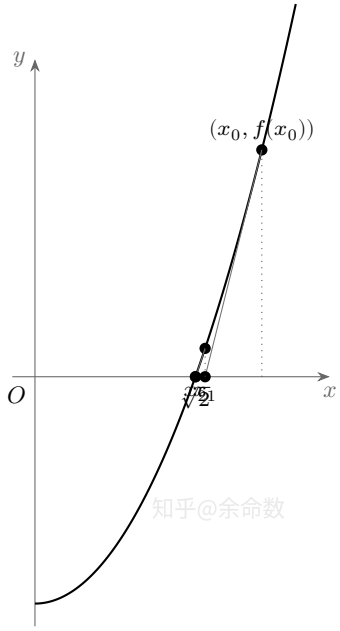


图 3.45 牛顿迭代法, 过 $(x_n, f(x_n))$ 作切线, 切线与 x 轴的交点为 x_{n+1}, 序列快速收敛到根 $\sqrt{2}$

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 4 章

函数方程与抽象函数

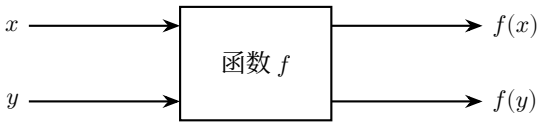
基础语言

「人追求完美，却往往只能得到近似。」

——柯西



第四章 函数方程与抽象函数



函数方程: $f(x+y)=f(x)+f(y)$

图 4.1 函数方程约束未知函数的行为

一个函数通常由其解析式定义. 然而, 在许多数学情境中, 函数的具体形态未知, 只知道它满足某些特定的代数关系式, 这就是函数方程.

抽象函数问题以一个或多个函数方程约束函数的行为. 我们从这些代数关系中提炼函数的具体性质, 再确定它的解析式.

4.1 基本分析方法

形式各异的函数方程共享若干基本分析策略. 这些策略既是模型法的依据,也在方程无法直接套用模型时承担探索的任务.

4.1.1 赋值法

赋值法就是给方程中的变量代入特殊值. 由于等式对所有变量成立, 它对特殊值也成立. 我们选取能简化结构或暴露函数性质的特殊值, 逐步推出函数的关键性质.

常见的赋值策略包括:

- 引入单位元: 在加法型方程 $f(x+y)$ 中令 $y=0$, 或在乘法型方程 $f(xy)$ 中令 $y=1$, 以探求 $f(0)$ 或 $f(1)$ 的值.
- 引入逆元: 在加法型方程中令 $y=-x$, 或在乘法型方程中令 $y=1/x$, 以探求函数的奇偶性或对称性.
- 变量取等: 令 $y=x$, 可将 $f(x+y)$ 转化为 $f(2x)$, 或将 $f(xy)$ 转化为 $f(x^2)$, 由此得到关于自变量倍数或乘方的函数关系式.

例题 4.1 (多选题) 已知函数 $f(x)$ 的定义域为 $\mathbb{R}$, 对任意 $x,y\in\mathbb{R}$, 均有 $f(x+y)=f(x)f(y)$ 成立, 且 $f(1)=2$. 若 $f(x)$ 在 $\mathbb{R}$ 上单调, 判断下列说法.

A. $f(0)=1$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

B. $f(x)$ 是偶函数

C. $f(x)$ 在 $\mathbb{R}$ 上单调递增

D. 不等式 $f(x^2-2x)<f(3)$ 的解集为 $(-1,3)$

解 我们先证明 $f(x)>0$ 对所有 x 成立. 在方程 $f(x+y)=f(x)f(y)$ 中令 $y=x$, 得到 $f(2x)=[f(x)]^2\geq 0$. 若存在 x_0 使得 $f(x_0)=0$, 则对任意 $x\in\mathbb{R}$, 有 $f(x)=f((x-x_0)+x_0)=f(x-x_0)f(x_0)=0$, 函数恒为零, 与条件 $f(1)=2$ 矛盾. 故对任意 $x\in\mathbb{R}$, $f(x)\neq 0$. 结合 $f(2x)\geq 0$ 知 $f(x)\geq 0$, 因而 $f(x)>0$.

对于 (A): 在原方程中令 $y=0$, 得 $f(x)=f(x)f(0)$. 由 $f(x)$ 恒不为零, 两边同除以 $f(x)$ 即得 $f(0)=1$. 故 (A) 正确.

对于 (B): 在原方程中令 $y=-x$ 以引入 $f(-x)$, 得 $f(x-x)=f(x)f(-x)$, 即 $f(0)=f(x)f(-x)$. 代入 $f(0)=1$, 得 $f(-x)=\dfrac{1}{f(x)}$. 由 $f(1)=2$ 知 $f(-1)=1/2$, 它既不等于 $f(1)$, 也不等于 $-f(1)$. 故 $f(x)$ 既非奇函数亦非偶函数, (B) 错误.

对于 (C): 题设已给出 $f(x)$ 在 $\mathbb{R}$ 上单调, 又 $f(1)=2>f(0)=1$, 故它只能单调递增. 进一步, 若 $h>0$, 则 $f(h)\geq f(0)=1$; 若 $f(h)=1$, 取正整数 n 使 $nh>1$, 则 $f(nh)=[f(h)]^n=1$, 与 $f(nh)\geq f(1)=2$ 矛盾. 故 $f(h)>1$, 从而 $f(x+h)=f(x)f(h)>f(x)$, 函数严格递增. 故 (C) 正确.

对于 (D): (C) 已证函数单调递增, 故不等式 $f(x^2-2x)<f(3)$ 等价于 $x^2-2x<3$. 解此一元二次不等式 $x^2-2x-3<0$, 即 $(x+1)(x-3)<0$, 解集为 $(-1,3)$. 故 (D) 正确.

正确的选项为 **(A)(C)(D)**.

可导假设下的换元与求导

4.1.2 换元构造法

当函数方程不能直接套用标准模型时, 我们识别方程中反复出现的共同结构, 把它定义为新函数, 从而把原方程化为标准方程.

例题 4.2 (2023 年 8 月浙江省 Z20 名校联盟第一次联考第 12 题) (多选)定义在 $(0,+\infty)$ 上的可导函数 $f(x)$ 满足以下条件.(1) $f(xy)=xf(y)+yf(x)$; (2) 当 $x>1$ 时, $f(x)>0$. 判断下列结论.

A. $f(1)=0$

B. $f(xy)=f(x)f(y)$

C. $f(x)$ 在 $(1,+\infty)$ 上单调递增

D. 不等式 $xf\left(x-\frac{3}{2}\right)\geq\left(\frac{3}{2}-x\right)f(x)$ 的解集为 $[2,+\infty)$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 方程 $f(xy)=xf(y)+yf(x)$ 的两边同除以 xy, 即可得到 $f(z)/z$ 的形式, 化为柯西方程.

在方程两边同时除以 xy. 定义域为 $(0,+\infty)$, x,y 均为正数, 除法可行:

$$\frac{f(xy)}{xy}=\frac{xf(y)}{xy}+\frac{yf(x)}{xy}\implies\frac{f(xy)}{xy}=\frac{f(y)}{y}+\frac{f(x)}{x}.$$

上式表明 $\frac{f(z)}{z}$ 反复出现. 令 $g(x)=\frac{f(x)}{x}$, 原方程化为标准对数型柯西方程:

$$g(xy)=g(x)+g(y).$$

题设给出 f 可导, 函数 g 也可导. 在可导正则条件下, 这个对数型柯西方程的解为 $g(x)=c\ln x$ (严格推导见 第16章). 原函数 $f(x)$ 的模型为 $f(x)=x\cdot g(x)=cx\ln x$.

由条件 (2) 确定常数 c 的符号, 当 $x>1$ 时, $f(x)>0$. 即 $cx\ln x>0$. 当 $x>1$ 时, $x>0$ 且 $\ln x>0$, 故 $c>0$.

模型 $f(x)=cx\ln x$ $(c>0)$ 既已确定, 我们逐一检验各选项.

- **(A)** $f(1)=0$: $f(1)=c\cdot 1\cdot\ln 1=0$. 正确.
- **(B)** $f(xy)=f(x)f(y)$: 即 $cxy\ln(xy)=(cx\ln x)(cy\ln y)$. 不成立. 错误.
- **(C) $f(x)$ 在 $(1,+\infty)$ 上单调递增**: 对 $f(x)$ 求导得 $f'(x)=c(\ln x+x\cdot\frac{1}{x})=c(\ln x+1)$. 因为 $c>0$, 当 $x>1$ 时, $\ln x>0$, 故 $f'(x)>0$. 函数单调递增. 正确.
- **(D) 解不等式**: 原不等式为 $xf(x-\frac{3}{2})-(\frac{3}{2}-x)f(x)\geq 0$. 移项得 $xf(x-\frac{3}{2})+(x-\frac{3}{2})f(x)\geq 0$. 左侧是 $f(x(x-\frac{3}{2}))$ 的展开式. 由条件 (1) $f(uv)=uf(v)+vf(u)$, 令 $u=x, v=x-\frac{3}{2}$, 原不等式等价于 $f(x(x-\frac{3}{2}))\geq 0$.
  自变量须为正, 即 $x>0$ 且 $x-\frac{3}{2}>0\implies x>\frac{3}{2}$. 代入模型得 $c\cdot x(x-\frac{3}{2})\ln\big(x(x-\frac{3}{2})\big)\geq 0$. 由 $c>0$ 且 $x(x-3/2)>0$, 不等式简化为 $\ln\big(x(x-\frac{3}{2})\big)\geq 0$. 对数函数递增且 $\ln 1=0$, 故真数不小于 1. $x(x-\frac{3}{2})\geq 1\implies 2x^2-3x-2\geq 0\implies(2x+1)(x-2)\geq 0$. 结合定义域 $x>\frac{3}{2}$, 解得 $x\geq 2$. 正确.

正确的选项是 **(A)(C)(D)**.

4.2 从基本初等函数到函数方程

基本初等函数的运算性质本身即构成一类函数方程.

- 线性函数 $f(x)=cx$ 保持加法结构: $f(x+y)=f(x)+f(y)$.
- 指数函数 $f(x)=a^x$ 将加法映射为乘法: $f(x+y)=f(x)f(y)$.
- 幂函数 $f(x)=x^k$ 保持乘法结构: $f(xy)=f(x)f(y)$.
- 对数函数 $f(x)=\log_a x$ 将乘法映射为加法: $f(xy)=f(x)+f(y)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这四个方程统称柯西函数方程. 左端出现和式 $x+y$ 的两类称为加法型与指数型, 左端出现乘积 xy 的两类称为乘法型与对数型; 右端把 f 拆成两项之和还是两项之积, 区分出每一对中的两个方程. 遇到一个抽象函数问题时, 先看方程的两个变量以何种方式组合, 再看右端的结合方式, 就能在四类模型中定位它.

例题 4.3　指出下列函数各满足四类方程中的哪一类, 并写出对应的函数方程:

$$\text{(1) } f(x)=5x, \qquad \text{(2) } f(x)=2^x,$$
$$\text{(3) } f(x)=x^3\ (x>0), \quad \text{(4) } f(x)=\log_2 x\ (x>0).$$

解　逐个把函数代进四类方程, 只有一种组合能使等式对定义域内任意 x,y 成立.

对于 (1): $f(x+y)=5(x+y)=5x+5y=f(x)+f(y)$, 满足加法型方程 $f(x+y)=f(x)+f(y)$.

对于 (2): 由指数律 $2^{x+y}=2^x\cdot 2^y$, 满足指数型方程 $f(x+y)=f(x)f(y)$.

对于 (3): 由幂的运算律 $(xy)^3=x^3y^3$, 满足乘法型方程 $f(xy)=f(x)f(y)$.

对于 (4): 由对数律 $\log_2(xy)=\log_2 x+\log_2 y$, 满足对数型方程 $f(xy)=f(x)+f(y)$.

方程的另一侧用法是求值. 方程本身不含待定参数, 却把一整族函数值锁在一起, 代入少数几个特殊值, 就能从一个已知的函数值推出其余各处的函数值.

例题 4.4　已知定义在 $\mathbb{R}$ 上的函数 $f(x)$ 满足 $f(x+y)=f(x)+f(y)$, 且 $f(2)=6$. 求 $f(0)$、$f(1)$、$f(-3)$ 与 $f(10)$.

解　方程对一切实数 x,y 成立, 代入特殊值即可逐步确定各处的函数值.

令 $x=y=0$, 得 $f(0)=f(0)+f(0)$, 故 $f(0)=0$. 令 $y=-x$, 得 $f(0)=f(x)+f(-x)$, 故 $f(-x)=-f(x)$, 函数为奇函数. 令 $x=y=1$, 得 $f(2)=2f(1)$, 由 $f(2)=6$ 解得 $f(1)=3$.

由 $f(1)=3$ 与 $f(2)=6$ 得 $f(3)=f(1)+f(2)=9$, 再由奇性得 $f(-3)=-f(3)=-9$. 对正整数 n, 反复使用方程得 $f(n)=nf(1)=3n$, 故 $f(10)=30$.

在连续性等正则假设下, 四类方程的连续解恰好就是本节开头列出的四类基本初等函数, 这是模型法成立的依据. 正则假设的用途与证明见 第4.5节.

4.3 二次及高次多项式模型

函数方程中出现二次交叉项或差分规律时, 解往往取二次或更高次多项式的形式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

4.3.1 二次函数型方程

定理 4.1 (二次函数型方程模型) 若二次函数 $f(x)=ax^2+bx+c$ 满足方程 $f(x+y)=f(x)+f(y)+\mu xy+\lambda$, 其中 μ,λ 为常数, 则 $a=\frac{\mu}{2}$, $c=-\lambda$, 即其解的形式为 $f(x)=\frac{\mu}{2}x^2+bx-\lambda$, 其中 b 为待定系数.

证明 将 $f(x+y)$ 展开:

$$f(x+y)=a(x+y)^2+b(x+y)+c=ax^2+ay^2+2axy+bx+by+c,$$

而

$$f(x)+f(y)+\mu xy+\lambda=ax^2+ay^2+bx+by+2c+\mu xy+\lambda.$$

等式对一切 x,y 成立, 两端逐项比较系数. 二次项系数相等是恒等式; 交叉项给出 $2a=\mu$; 常数项给出 $c=2c+\lambda$; 一次项系数 b 不受约束. 于是

$$a=\frac{\mu}{2},\qquad c=-\lambda,\qquad f(x)=\frac{\mu}{2}x^2+bx-\lambda.$$

□

注 4.1
上述推导只用到 f 是二次函数这一条件. 去掉该限制, 仅保留连续性, 解的形式仍然不变, 证明归结为加法型柯西方程, 见 第4.5节.

例题 4.5 (多选) 已知二次函数 $f(x)$ 满足对任意 $x,y\in\mathbb{R}$, 均有 $f(x+y)=f(x)+f(y)+4xy-3$, 且 $f(1)=6$. 判断下列说法.

A. $f(0)=3$

B. $f(x)$ 为偶函数

C. 函数 $y=f(x)$ 图像的对称轴为 $x=-\frac{1}{4}$

D. 函数 $f(x)$ 的最小值为 $\frac{23}{8}$

解 这是二次函数型方程. 我们直接用模型确定解析式.

据定理取参数 $\mu=4,\lambda=-3$, 函数模型为 $f(x)=\frac{4}{2}x^2+bx-(-3)=2x^2+bx+3$.

由题设 $f(1)=6$ 确定待定系数 b: $f(1)=2(1)^2+b(1)+3=5+b$, 令 $5+b=6$, 解得 $b=1$.

故 $f(x)=2x^2+x+3$.

- **对于 (A)**: $f(0)=3$.(A) 正确.
- **对于 (B)**: $f(-x)=2(-x)^2+(-x)+3=2x^2-x+3\neq f(x)$, 故 $f(x)$ 为非偶函数.(B) 错误.
- **对于 (C)**: 二次函数 $f(x)=2x^2+x+3$ 的对称轴为 $x=-\frac{1}{2(2)}=-\frac{1}{4}$.(C) 正确.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **对于 (D)**: 最小值在顶点处取得, 即 $x=-1/4$ 时. $f\left(-\frac{1}{4}\right)=2\left(-\frac{1}{4}\right)^2-\frac{1}{4}+3=\frac{1}{8}-\frac{2}{8}+\frac{24}{8}=\frac{23}{8}$.(D) 正确.

正确的选项是 **(A)(C)(D)**.

4.3.2 差分法与多项式阶数

差分算子描述的是离散函数的变化. 定义一阶前向差分 $\Delta f(x)=f(x+1)-f(x)$. 它作用在多项式上使次数降低一.

对于二次函数 $f(x)=ax^2+bx+c$, 一阶差分 $\Delta f(x)=2ax+(a+b)$ 是一次多项式, 二阶差分 $\Delta^2 f(x)=\Delta(\Delta f(x))=f(x+2)-2f(x+1)+f(x)=2a$ 是常数. 反之, 若函数在整数点上的 k 阶差分恒为常数, 则它在整数点上的取值与某个次数不超过 k 的多项式一致.

例题 4.6 (2023 年 12 月苏州市部分高中高三上联考第 12 题) (多选) 定义在 $\mathbb{R}$ 上的二次函数 $f(x)$ 同时满足以下两个条件.

I. $f(x+1)-f(x)=2x+2, \forall x\in\mathbb{R}$;

II. 当 $x\in[0,1]$ 时, $|f(x)|\leq 1$.

判断下列说法.

A. $f(0)=-1$

B. $f(x)$ 为偶函数

C. 存在 $n\in\mathbb{N}^*$, 使得 $f(n)>2023n$

D. 对任意 $x\in\mathbb{R}$, $|f(x)|<x^2+|x|+3$

解 题设已说明 $f(x)$ 为二次函数. 我们先用待定系数法确定二次项与一次项, 再用条件 II 确定常数项.

设 $f(x)=ax^2+bx+c$, 其一阶差分为

$$f(x+1)-f(x)=[a(x+1)^2+b(x+1)+c]-[ax^2+bx+c]=2ax+(a+b),$$

由条件 I 得恒等式 $2ax+(a+b)\equiv 2x+2$. 比较系数, $2a=2, a+b=2$, 解得 $a=1, b=1$. 故 $f(x)=x^2+x+c$.

由条件 II, 对任意 $x\in[0,1]$, $|f(x)|\leq 1$ 恒成立, 即 $-1\leq x^2+x+c\leq 1$. 令 $h(x)=x^2+x$.$h(x)$ 在 $[0,1]$ 上单调递增,值域为 $[0,2]$. 故 $x\in[0,1]$ 时 $f(x)=h(x)+c$ 的值域为 $[c,c+2]$. 要使该区间包含于 $[-1,1]$,须 $c\geq -1$ 且 $c+2\leq 1$,唯一解 $c=-1$.

解析式唯一确定为 $f(x)=x^2+x-1$.

- **对于 (A)**: $f(0)=-1$. 正确.
- **对于 (B)**: $f(x)$ 的对称轴为 $x=-1/2$, 为非偶函数. 错误.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **对于 (C)**: $f(n) > 2023n$ 等价于 $n^2 + n - 1 > 2023n$, 即 $n^2 - 2022n - 1 > 0$. 这是开口向上的二次函数, n 足够大时函数值必为正. 正确.
- **对于 (D)**: 由三角不等式, $|x^2 + x - 1| \le |x^2| + |x| + |-1| = x^2 + |x| + 1$. 因 $1 < 3$, $x^2 + |x| + 1 < x^2 + |x| + 3$. 故 $|f(x)| \le x^2 + |x| + 1 < x^2 + |x| + 3$ 恒成立. 正确.

正确的选项是 **(A)(C)(D)**.

4.4 三角与双曲函数模型

注 4.2
本节使用三角函数模型, 部分证明调用可微性、微分方程与积分思想, 见第6章 与 第16章. 赋值与递推求候选解不依赖可微性.

有些函数方程的结构与三角函数或双曲函数的和差角公式一致. 我们称

$$f(x+y) + f(x-y) = 2f(x)f(y)$$

为达朗贝尔方程.

三角与双曲函数模型

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

4.4.1 达朗贝尔方程

定义 4.1 (双曲函数) 双曲正弦函数 (sinh) 和双曲余弦函数 (cosh) 定义为

$$\sinh x = \frac{e^x - e^{-x}}{2}, \quad \cosh x = \frac{e^x + e^{-x}}{2} \tag{4-1}$$

$\sinh x$ 是奇函数, $\cosh x$ 是偶函数. 它们满足恒等式 $\cosh^2 x - \sinh^2 x = 1$.

定理 4.2 (达朗贝尔方程的正则解) 设函数 $f : \mathbb{R} \to \mathbb{R}$ 二次连续可微, 若满足 $f(x+y) + f(x-y) = 2f(x)f(y)$, 则其非零解的模型为 $f(x) = \cos(\omega x)$ 或 $f(x) = \cosh(\omega x)$, 其中 ω 为常数; 零函数也是方程的平凡解.

证明 令 $y = 0$, 得 $2f(x) = 2f(x)f(0)$. 由 f 不恒为零, 存在 x_0 使 $f(x_0) \neq 0$, 故 $f(0) = 1$.

将原方程对 y 求偏导两次. 一阶偏导给出 $f'(x+y) - f'(x-y) = 2f(x)f'(y)$, 令 $y = 0$ 得 $0 = 2f(x)f'(0)$, 由 f 不恒为零知 $f'(0) = 0$. 再求一次偏导得 $f''(x+y) + f''(x-y) = 2f(x)f''(y)$, 令 $y = 0$ 得 $2f''(x) = 2f(x)f''(0)$, 即

$$f''(x) = kf(x), \quad k = f''(0).$$

知乎@余命数

这是二阶常系数齐次线性微分方程. 按 k 的符号分类讨论. 若 $k < 0$, 令 $k = -\omega^2$, 通解为 $A\cos(\omega x) + B\sin(\omega x)$, 由初始条件 $f(0) = 1, f'(0) = 0$ 得 $f(x) = \cos(\omega x)$. 若 $k > 0$, 令 $k = \omega^2$, 通解为 $Ae^{\omega x} + Be^{-\omega x}$, 同样由初始条件得 $f(x) = \cosh(\omega x)$.

若 $k=0$, 解为常数函数 $f(x)=1$. 由和差角公式可反向验证零函数、$\cos(\omega x)$ 与 $\cosh(\omega x)$ 均满足原方程. □

例题 4.7 (2024 届重庆第八中学高三月考卷五) (多选)定义域为 $\mathbb{R}$ 的连续函数 $f(x)$, 对任意 $x,y\in\mathbb{R}$, 有 $f(x+y)+f(x-y)=2f(x)f(y)$, 且 $f(x)$ 不恒为 0, 则下列说法正确的是

A. $f(x)$ 为偶函数

B. $f(x)+f(0)\geq 0$

C. 若 $f(1)=0$, 则 $\sum\limits_{i=1}^{2024} f(i)=4048$

D. 若 0 为 $f(x)$ 的极小值点, 则 $f(x)$ 的最小值为 1

解 由赋值 $y=0$ 得 $2f(x)=2f(x)f(0)$, 因 f 不恒为零, $f(0)=1$; 再令 $x=0$ 得 $f(y)+f(-y)=2f(y)$, 即 $f(-y)=f(y)$.

对于 (A): 上式表明 f 为偶函数. 故 (A) 正确.

对于 (B): 在原方程中取 $x=y=\frac{t}{2}$, 得

$$f(t)+f(0)=2f\left(\frac{t}{2}\right)^2.$$

由 $f(0)=1$ 知 $f(t)+1=2f(t/2)^2\geq 0$. 故 (B) 正确.

对于 (C): 附加条件 $f(1)=0$. 在原方程中取 $y=1$ 且 $x=n$ ($n\in\mathbb{Z}$), 得

$$f(n+1)+f(n-1)=2f(n)f(1)=0.$$

又 $f(0)=1, f(1)=0$, 故 $f(2)=-1, f(3)=0, f(4)=1$, 整数点函数值以 $1,0,-1,0$ 为周期循环, 一个周期内和为 0. 2024 是 4 的倍数, 恰含整数个周期, 和为 0. 故 (C) 错误.

对于 (D): 附加条件为 0 是 $f(x)$ 的极小值点. 此时存在 $\delta>0$, 当 $|t|<\delta$ 时 $f(t)\geq f(0)=1$. 任取实数 x, 取足够大的正整数 n 使 $|x/n|<\delta$. 记 $h=x/n$, $c=f(h)$, 则 $c\geq 1$. 令 $a_k=f(kh)$, 由原方程得递推关系

$$a_{k+1}+a_{k-1}=2ca_k,\qquad a_0=1,\ a_1=c.$$

当 $c=1$ 时, 递推给出 $a_k=1$; 当 $c>1$ 时, 令 $c=\cosh u$ ($u>0$), 由同一递推及初值比较得 $a_k=\cosh(ku)\geq 1$. 故 $f(x)=a_n\geq 1$. 又 $f(0)=1$, $f(x)$ 的最小值为 1. 故 (D) 正确.

正确的选项为 **(A)(B)(D)**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

4.4.2 正切函数模型

定理 4.3 (正切函数模型) 设 I 是含 0 的开区间. 若可微函数 $f: I \to \mathbb{R}$ 对所有 $x, y, x+y \in I$ 满足

$$f(x+y) = \frac{f(x)+f(y)}{1-f(x)f(y)},$$

且分母不为 0, 则在该区间内其正则解模型为 $f(x) = \tan(\omega x)$, 其中 ω 为常数且 ωx 不跨越正切极点.

知乎@余命数

证明 令 $x = y = 0$, 得 $f(0) = 0$. 固定 x, 将原方程对 y 求导. 右端按商法则求导, 再令 $y = 0$ 并代入 $f(0) = 0$, 得

$$f'(x) = \frac{f'(0)(1-f(x)f(0)) + (f(x)+f(0))f(x)f'(0)}{(1-f(x)f(0))^2} = f'(0)(1+[f(x)]^2).$$

这是一阶可分离变量的微分方程. 令 $y = f(x)$ 及常数 $\omega = f'(0)$, 方程化为 $\frac{dy}{dx} = \omega(1+y^2)$. 分离变量积分得 $\arctan(y) = \omega x + C$. 由初始条件 $f(0) = 0$ 得 $C = 0$, 故 $f(x) = \tan(\omega x)$. □

例题 4.8 (2024 年 1 月浙江省宁波市慈溪期末第 6 题) 已知函数 $f(x)$ 的定义域为 $\mathbb{R}$, 且 $f(x+y)f(x)f(y) = f(x+y) - f(x) - f(y)$. 若 $f(1) = \sqrt{3}$, 则 $\sum_{k=1}^{2024} f(k) =$

A. 2024

B. $1012\sqrt{3}$

C. $\sqrt{3}$

D. 0

解 题目只要求整数点函数值之和. 我们直接由递推关系计算整数点即可.

令 $x = y = 0$, 得 $f(0)^3 = -f(0)$, 故 $f(0) = 0$. 在原式中令 $y = 1$ 并代入 $f(1) = \sqrt{3}$, 得

$$f(x+1)f(x)\sqrt{3} = f(x+1) - f(x) - \sqrt{3}.$$

只要 $1 - \sqrt{3}f(x) \neq 0$, 即有

$$f(x+1) = \frac{f(x)+\sqrt{3}}{1-\sqrt{3}f(x)}.$$

从 $x=1$ 起逐个计算整数点, 得

$$f(2)=\frac{\sqrt{3}+\sqrt{3}}{1-3}=-\sqrt{3},\qquad f(3)=\frac{-\sqrt{3}+\sqrt{3}}{1+3}=0.$$

再把 x 换成 $x+3$, 由 $f(3)=0$ 得

$$0=f(x+3)-f(x)-f(3)=f(x+3)-f(x),$$

故 $f(x+3)=f(x)$, 整数点函数值以 $\sqrt{3},-\sqrt{3},0$ 为周期循环, 一个周期之和为 0.

所求和式共 2024 项, $2024=3\times 674+2$.

$$\begin{aligned}\sum_{k=1}^{2024}f(k)&=674\times(f(1)+f(2)+f(3))+f(2023)+f(2024)\\&=0+f(1)+f(2)=\sqrt{3}-\sqrt{3}=0.\end{aligned}$$

答案为 0, 选 (D).

4.5 柯西方程的正则解

第4.2节 列出四类基本初等函数各自满足的函数方程. 反向的问题是, 一个未知函数满足其中某个方程时, 它能取的形式取决于对未知函数施加的正则假设. 本节处理连续性假设下的情形, 四类方程的解随之唯一确定.

4.5.1 四类柯西方程的连续解

定理 4.4 (柯西函数方程) 在连续性等正则条件下, 以下四类柯西函数方程的连续解为对应的基本初等函数.

加法型 若连续函数 $f:\mathbb{R}\to\mathbb{R}$ 满足 $f(x+y)=f(x)+f(y)$, 则 $f(x)=cx$.

指数型 若连续函数 $f:\mathbb{R}\to\mathbb{R}$ 满足 $f(x+y)=f(x)f(y)$, 则 $f(x)\equiv 0$ 或 $f(x)=a^x$ $(a>0)$.

乘法型 若连续函数 $f:(0,+\infty)\to\mathbb{R}$ 满足 $f(xy)=f(x)f(y)$, 则 $f(x)\equiv 0$ 或 $f(x)=x^k$.

对数型 若连续函数 $f:(0,+\infty)\to\mathbb{R}$ 满足 $f(xy)=f(x)+f(y)$, 则 $f(x)=c\ln x$; 当 $c\neq 0$ 时也可写作 $f(x)=\log_a x$ $(a=e^{1/c})$, 当 $c=0$ 时为零函数.

连续性的作用:解从有理数延拓到实数

证明 证明中假设 $f(x)$ 在其定义域上连续.

注 4.3 把解从有理数延拓到实数需要连续性, 其严格证明见第16章. 赋值得到的有理数结论并不依赖连续性, 连续性只是延拓到实数时的正则假设.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们先把解从整数逐步延拓到实数. 由数学归纳法, 对任意 $n \in \mathbb{Z}$ 及 $x \in \mathbb{R}$, $f(nx) = nf(x)$. 由此对任意有理数 $r = p/q$ ($p \in \mathbb{Z}, q \in \mathbb{N}^*$), 有 $f(p) = pf(1)$ 且 $f(p) = f(q \cdot \frac{p}{q}) = qf(\frac{p}{q})$. 令 $c = f(1)$, 则 $pc = qf(p/q)$, 故 $f(p/q) = (p/q)c$. 即对任意有理数 $r \in \mathbb{Q}$, $f(r) = cr$. 对任意实数 $x \in \mathbb{R}$, 取有理数序列 $\{r_n\}$ 使得 $\lim\limits_{n\to\infty} r_n = x$. 利用连续性, 有

$$f(x) = f\left(\lim_{n\to\infty} r_n\right) = \lim_{n\to\infty} f(r_n) = \lim_{n\to\infty} (cr_n) = c\left(\lim_{n\to\infty} r_n\right) = cx.$$

故加法型方程的连续解为 $f(x) = cx$.

其余三种方程经变量代换化为加法型方程.

- **指数型**: 若 $f(x)$ 不恒为零, 先令 $y = 0$ 得 $f(0) = 1$. 若存在 t 使 $f(t) = 0$, 则对任意 x 有 $f(x) = f((x-t)+t) = f(x-t)f(t) = 0$, 矛盾. 又 $f(x) = f(x/2)^2 > 0$, 故可构造 $g(x) = \ln f(x)$. $g(x+y) = \ln f(x+y) = \ln(f(x)f(y)) = \ln f(x) + \ln f(y) = g(x) + g(y)$. 故 $g(x) = cx$. 因此 $\ln f(x) = cx \implies f(x) = e^{cx} = (e^c)^x$. 令 $a = e^c$, 则 $f(x) = a^x$.
- **对数型**: 定义域为 $(0, +\infty)$. 构造 $g(t) = f(e^t)$. $g(u+v) = f(e^{u+v}) = f(e^u e^v) = f(e^u) + f(e^v) = g(u) + g(v)$. 故 $g(t) = ct$. 令 $t = \ln x$, 则 $f(x) = c\ln x$. 若 $c \neq 0$, 令 $a = e^{1/c}$, 则 $f(x) = \log_a x$; 若 $c = 0$, 则 $f \equiv 0$.
- **乘法型**: 定义域为 $(0, +\infty)$. 若 f 不恒为零, 同指数型可得 $f(x) > 0$. 构造 $g(t) = \ln f(e^t)$. $g(u+v) = \ln f(e^{u+v}) = \ln(f(e^u)f(e^v)) = \ln f(e^u) + \ln f(e^v) = g(u) + g(v)$. 故 $g(t) = ct$. 因此 $\ln f(e^t) = ct \implies f(e^t) = e^{ct} = (e^t)^c$. 令 $x = e^t$, 则 $f(x) = x^c$.

□

4.5.2 代数变式

标准的柯西方程经代数平移或调整, 可得到新的函数方程模型.

4.5.2.1 平移变式

定理 4.5 (柯西方程平移变式) 若连续函数 $f : \mathbb{R} \to \mathbb{R}$ 满足 $f(x+y) = f(x) + f(y) + \lambda$, 其中 λ 为常数, 则其函数形式为 $f(x) = cx - \lambda$, 其中 c 为待定常数.

证明 为消去常数项 λ, 令 $g(x) = f(x) + \lambda$, 则

$$\begin{aligned} g(x+y) &= f(x+y) + \lambda = (f(x) + f(y) + \lambda) + \lambda \\ &= (f(x) + \lambda) + (f(y) + \lambda) = g(x) + g(y). \end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于是 $g(x)$ 满足标准的加法型柯西方程, 故 $g(x)=cx$. 将 $f(x)=g(x)-\lambda$ 代回, 得 $f(x)+\lambda=cx$, 即 $f(x)=cx-\lambda$. □

例题 4.9 (2024 年 3 月安徽省安庆市二模第 9 题) (多选)已知定义在 $\mathbb{R}$ 上的函数 $f(x)$, 满足对任意的实数 x,y, 均有 $f(x+y)=f(x)+f(y)-1$, 且当 $x>0$ 时, $f(x)<1$. 判断下列说法.

A. $f(0)=1$

B. $f(1)+f(-1)=1$

C. $f(x)$ 为减函数

D. 函数 $y=f(x)$ 的图像关于点 $(0,1)$ 对称

解 这是 $\lambda=-1$ 的平移变式. 我们先用赋值法推出函数的性质, 再判断各选项.

对于 (A): 在方程 $f(x+y)=f(x)+f(y)-1$ 中令 $x=y=0$, 得 $f(0)=f(0)+f(0)-1$, 解得 $f(0)=1$. 故 (A) 正确.

对于 (D) 与 (B): 在方程中令 $y=-x$ 以引入 $f(-x)$, 得 $f(x-x)=f(x)+f(-x)-1$, 即 $f(0)=f(x)+f(-x)-1$. 代入 $f(0)=1$, 得 $f(x)+f(-x)=2$. 该关系式是函数图像关于点 $(0,1)$ 中心对称的代数表述, 故 (D) 正确. 在其中令 $x=1$, 得 $f(1)+f(-1)=2$. 故 (B) 错误.

对于 (C): 任取 $x_1,x_2\in\mathbb{R}$ 且 $x_2>x_1$, 令 $\Delta x=x_2-x_1>0$. 考察 $f(x_2)-f(x_1)=f(x_1+\Delta x)-f(x_1)$. 由原方程, $f(x_1+\Delta x)=f(x_1)+f(\Delta x)-1$, 故 $f(x_2)-f(x_1)=f(\Delta x)-1$. 题设给出自变量大于 0 时函数值小于 1. 由 $\Delta x>0$ 知 $f(\Delta x)<1$, 即 $f(\Delta x)-1<0$. 故 $f(x_2)<f(x_1)$, $f(x)$ 在 $\mathbb{R}$ 上递减. 故 (C) 正确.

正确的选项是 **(A)(C)(D)**.

4.5.2.2 尺度变式

定理 4.6 (带尺度因子的乘法型方程模型) 若连续函数 $f:(0,+\infty)\to\mathbb{R}$ 满足 $f(xy)=\lambda f(x)f(y)$, 其中 $\lambda\neq0$, 其非平凡解的模型为 $f(x)=\frac{1}{\lambda}x^{\alpha}$.

证明 令 $g(x)=\lambda f(x)$, 则

$$\begin{aligned}g(xy)&=\lambda f(xy)=\lambda(\lambda f(x)f(y))\\&=(\lambda f(x))(\lambda f(y))=g(x)g(y).\end{aligned}$$

于是 $g(x)$ 满足标准的柯西乘法型方程, 故其正则解为 $g(x)=x^{\alpha}$. 将 $f(x)=g(x)/\lambda$ 代回, 得 $\lambda f(x)=x^{\alpha}$, 解得 $f(x)=\frac{1}{\lambda}x^{\alpha}$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

用导数判定抽象函数性质

例题 4.10 （多选）已知定义在 $(0,+\infty)$ 上的连续函数 $f(x)$ 满足对任意 $x,y\in(0,+\infty)$，均有 $f(xy)=2f(x)f(y)$，且 $f(2)=4$. 判断下列说法.

A. $f(1)=\frac{1}{2}$

B. $f(x)$ 是奇函数

C. $f(x)$ 在 $(0,+\infty)$ 上单调递增

D. $f(\frac{x}{2})=\frac{f(x)}{8}$

解 这是 $f(xy)=\lambda f(x)f(y)$ 形式的方程，其解为带系数的幂函数. 我们套用上面的模型确定解析式，再逐一检验选项.

据定理取参数 $\lambda=2$，函数模型为 $f(x)=\frac{1}{\lambda}x^{\alpha}=\frac{1}{2}x^{\alpha}$.

由题设 $f(2)=4$ 确定指数 α：$f(2)=\frac{1}{2}(2)^{\alpha}=4$，即 $2^{\alpha}=8$，解得 $\alpha=3$.

满足所有条件的函数唯一，解析式为 $f(x)=\frac{1}{2}x^3$.

- **对于 (A)**：$f(1)=\frac{1}{2}(1)^3=\frac{1}{2}$.(A) 正确.
- **对于 (B)**：定义域 $(0,+\infty)$ 不关于原点对称，无法讨论奇偶性.(B) 错误.
- **对于 (C)**：$f'(x)=\frac{d}{dx}(\frac{1}{2}x^3)=\frac{3}{2}x^2$. 在定义域 $(0,+\infty)$ 上 $f'(x)>0$，函数严格递增.(C) 正确.
- **对于 (D)**：比较方程两端. 左端 $f(\frac{x}{2})=\frac{1}{2}\left(\frac{x}{2}\right)^3=\frac{x^3}{16}$，右端 $\frac{f(x)}{8}=\frac{\frac{1}{2}x^3}{8}=\frac{x^3}{16}$. 两端相等.(D) 正确.

正确的选项是 **(A)(C)(D)**.

4.6 Jensen 方程与 Pexider 方程

柯西方程刻画加法结构保持者. 把视角从端点移到中点，或从一个函数推广到多个函数，就得到两类自然变体.

4.6.1 Jensen 方程的中点条件

Jensen 方程把加法条件换成中点条件：

$$f\left(\frac{x+y}{2}\right)=\frac{f(x)+f(y)}{2}. \tag{4-2}$$

满足此方程的连续函数就是一次函数.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 4.7 (Jensen 方程的连续解) 若连续函数 $f:\mathbb{R}\to\mathbb{R}$ 满足 (4-2), 则 $f(x)=ax+b$, 其中 a,b 为常数.

证明 令 $g(x)=f(x)-f(0)$, 则 $g(0)=0$. 将 $f(x)=g(x)+f(0)$ 代入 (4-2):

$$g\left(\frac{x+y}{2}\right)+f(0)=\frac{g(x)+f(0)+g(y)+f(0)}{2},$$

化简得

$$g\left(\frac{x+y}{2}\right)=\frac{g(x)+g(y)}{2}. \tag{4-3}$$

在 (4-3) 中令 $y=0$, 利用 $g(0)=0$ 得 $g(x/2)=g(x)/2$.

令 $u=x/2,\ v=y/2$, 则

$$g(u+v)=g\left(\frac{2u+2v}{2}\right)=\frac{g(2u)+g(2v)}{2}.$$

由 $g(x/2)=g(x)/2$ 知 $g(2u)=2g(u), g(2v)=2g(v)$, 故

$$g(u+v)=\frac{2g(u)+2g(v)}{2}=g(u)+g(v).$$

g 满足加法型柯西方程, 由 第4.5节 的结论 $g(x)=ax$. 代回 $f(x)=g(x)+f(0)$, 得 $f(x)=ax+b$ $(b=f(0))$. □

例题 4.11 (多选)已知定义在 $\mathbb{R}$ 上的连续函数 $f(x)$ 满足对任意 $x,y\in\mathbb{R}$, 均有 $f\left(\frac{x+y}{2}\right)=\frac{f(x)+f(y)}{2}$, 且 $f(1)=3, f(2)=5$. 判断下列说法.

A. $f(0)=1$

B. $f(x)=2x+1$

C. $f(x)$ 为奇函数

D. $f(3)=7$

注 4.4
Jensen 方程与柯西加法型方程的差别在于常数项 b. 平移 $g(x)=f(x)-f(0)$ 消去常数项后, 两个方程等价.

解 由 Jensen 方程的连续解 $f(x)=ax+b$, 代入 $f(1)=3, f(2)=5$ 得

$$a+b=3,\quad 2a+b=5.$$

解得 $a=2,\ b=1$, 故 $f(x)=2x+1$.

- **对于 (A):** $f(0)=2\cdot 0+1=1$.(A) 正确.
- **对于 (B):** 已求得 $f(x)=2x+1$.(B) 正确.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **对于 (C)**：$f(-x)=-2x+1$，而 $-f(x)=-2x-1$，两者不等，f 不是奇函数.(C) 错误.
- **对于 (D)**：$f(3)=2\cdot 3+1=7$.(D) 正确.

正确的选项是 **(A)(B)(D)**.

4.6.2 Pexider 方程的多函数加法分解

Pexider 方程把柯西加法型方程从一个函数推广到三个函数：

$$f(x+y)=g(x)+h(y). \tag{4-4}$$

其中 f,g,h 是三个待定函数. 在连续性假设下，三个函数都是同一斜率的线性函数，只差常数项.

知乎@余命数

定理 4.8 (Pexider 方程的连续解) 若连续函数 $f,g,h:\mathbb{R}\to\mathbb{R}$ 满足 (4-4)，则存在常数 c,a,b 使得

$$f(x)=cx+a+b,\quad g(x)=cx+a,\quad h(y)=cy+b.$$

证明 在 (4-4) 中令 $x=0$ 得 $f(y)=g(0)+h(y)$，令 $y=0$ 得 $f(x)=g(x)+h(0)$.

记 $a=g(0)$，$b=h(0)$，则

$$h(y)=f(y)-a,\quad g(x)=f(x)-b.$$

知乎@余命数

代回 (4-4)：

$$f(x+y)=(f(x)-b)+(f(y)-a)=f(x)+f(y)-(a+b).$$

令 $F(x)=f(x)-(a+b)$，则

$$F(x+y)=f(x+y)-(a+b)=f(x)+f(y)-(a+b)-(a+b)=F(x)+F(y).$$

F 满足加法型柯西方程，故 $F(x)=cx$. 代回得

$$f(x)=cx+a+b,\quad g(x)=cx+a,\quad h(y)=cy+b.$$

□

例题 4.12 （多选）已知定义在 $\mathbb{R}$ 上的连续函数 f,g,h 满足对任意 $x,y\in\mathbb{R}$，均有 $f(x+y)=g(x)+h(y)$，且 $g(0)=1$、$h(0)=2$、$f(1)=4$. 判断下列说法.

A. $f(x)=x+3$

B. $g(x)-h(x)=-1$

C. $f(x)$ 为奇函数

D. $f(2)=5$

解 由 Pexider 方程的连续解, $f(x)=cx+a+b$、$g(x)=cx+a$、$h(y)=cy+b$.
代入 $g(0)=a=1$、$h(0)=b=2$、$f(1)=c+a+b=c+3=4$, 解得 $c=1$.
故 $f(x)=x+3$、$g(x)=x+1$、$h(x)=x+2$.

- **对于 (A)**: $f(x)=x+3$.(A) 正确.
- **对于 (B)**: $g(x)-h(x)=(x+1)-(x+2)=-1$.(B) 正确.
- **对于 (C)**: $f(0)=3\neq 0$, f 不是奇函数.(C) 错误.
- **对于 (D)**: $f(2)=2+3=5$.(D) 正确.

正确的选项是 **(A)(B)(D)**.

知乎@余命数

注 4.5
Pexider 方程也有乘法版本 $f(x+y)=g(x)\,h(y)$. 在连续性且 f 不恒为零的条件下, 令 $y=0$ 得 $f(x)=g(x)\,h(0)$, 令 $x=0$ 得 $f(y)=g(0)\,h(y)$, 代回后 f 满足指数型柯西方程, 解为 $f(x)=A\,a^x$、$g(x)=B\,a^x$、$h(y)=\frac{A}{B}\,a^y$.

4.7 迭代函数方程

知乎@余命数

前几节处理的方程都含两个自变量 x,y. 还有一类方程只含单变量, 但函数以复合方式出现: $f(f(x))$, $f(f(f(x)))$, 等等. 这类方程靠迭代处理.

4.7.1 迭代与对合

记 $f^1(x)=f(x)$, $f^{n+1}(x)=f(f^n(x))$, 称 $f^n(x)$ 为 f 的 n 次迭代.

定义 4.2 (对合) 若 $f(f(x))=x$ 对所有 x 成立, 则称 f 为一个**对合**. 对合就是恒等映射的迭代平方根.

对合的直观特征是两次操作回到原处. 常见的对合有 $f(x)=-x$、$f(x)=a-x$、$f(x)=1/x$ $(x\neq 0)$.

知乎@余命数

定理 4.9 (对合的构造) 设 g 是一一映射, 则 $f(x)=g^{-1}(-g(x))$ 是对合.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明

$$f(f(x))=g^{-1}(-g(g^{-1}(-g(x))))=g^{-1}(-(-g(x)))=g^{-1}(g(x))=x.$$

□

例题 4.13 讨论满足 $f(f(x))=x$ 的连续函数 $f:\mathbb{R}\to\mathbb{R}$ 的类型.

解 连续的对合必须严格单调. 若 f 严格递增且 $f(f(x))=x$, 假设存在 x_0 使 $f(x_0)>x_0$, 则 $f(f(x_0))>f(x_0)>x_0$, 与 $f(f(x_0))=x_0$ 矛盾; 同理 $f(x_0)<x_0$ 也矛盾. 故严格递增的对合只有 $f(x)=x$.

严格递减的连续对合有很多: $f(x)=a-x$ 是一族,更一般地 $f(x)=g^{-1}(-g(x))$ 对任意严格递增的连续 g 都给出严格递减的对合.

知乎@余命数

4.7.2 Babbage 方程的迭代根

对合是 $f^2(x)=x$ 的解. 推广到一般情形, 求 $f^n(x)=h(x)$ 的解 f, 称为 h 的**迭代 n 次根**. 当 $h(x)=x$ 时, 称为 Babbage 方程.

定理 4.10 (迭代根的单调性约束) 设 h 是一一映射, f 是 h 的连续迭代 n 次根.

1. 若 n 为偶数且 h 严格递减, 则连续迭代根不存在.
2. 若 n 为奇数, 则 h 的单调性与 f 一致.

证明 连续且一一的函数必然严格单调. 若 f 严格递增, 则 f^2 也严格递增, 进而 f^n 对偶数 n 严格递增. 若 f 严格递减, 则 f^2 严格递增, f^3 严格递减, 交替递推: f^n 在 n 为偶数时严格递增, 在 n 为奇数时与 f 同向.

当 n 为偶数时 $h=f^n$ 必须严格递增, 故 h 严格递减时无解. 当 n 为奇数时 h 的单调性就是 f 的单调性. □

例题 4.14 (1979 年 Putnam 竞赛 A2 题) 求常数 k 满足何种条件时, 存在连续函数 $f:\mathbb{R}\to\mathbb{R}$ 使得 $f(f(x))=kx^9$.

解 此即 $n=2, h(x)=kx^9$ 的迭代根问题.

当 $k<0$ 时 $h(x)=kx^9$ 严格递减. 由 定理 4.10, 偶数阶迭代根要求 h 严格递增, 故 $k<0$ 时无连续解.

当 $k\geq 0$ 时, 取 $f(x)=k^{1/4}x^3$, 则

$$f(f(x))=k^{1/4}(k^{1/4}x^3)^3=k^{1/4}\cdot k^{3/4}x^9=kx^9.$$

故 $k \geq 0$ 时存在连续解.

所求条件为 $k \geq 0$.

4.7.3 对交换性方程两边复合 f

对迭代根方程 $f^n(x) = h(x)$, 在两边复合 f:

$$f(h(x)) = f(f^n(x)) = f^n(f(x)) = h(f(x)).$$

于是 f 与 h 交换: $f \circ h = h \circ f$. 这是迭代根的必要条件, 可用于缩小候选范围.

例题 4.15 求所有整数值函数 $f : \mathbb{Z} \to \mathbb{Z}$ 满足 $f(f(n)) = n + 1$.

解 对两边复合 f: $f(n+1) = f(f(f(n))) = f(n) + 1$.

由 $f(n+1) - f(n) = 1$ 得 $f(n) = n + m$(m 为整数常数). 代回原方程:

$$f(f(n)) = f(n+m) = n + 2m.$$

要求 $n + 2m = n + 1$, 即 $2m = 1$, 在整数范围内无解.

故不存在满足条件的整数值函数.

4.7.4 共轭化方法

设 h 的迭代根 f 可以写成 $f(x) = \phi^{-1}(\beta(\phi(x)))$, 其中 ϕ 是一一映射, β 是较简单的函数. 则

$$h(x) = f^n(x) = \phi^{-1}(\beta^n(\phi(x))),$$

即 $\phi(h(x)) = \beta^n(\phi(x))$. 于是求 h 的迭代根归结为求较简单函数 β 的迭代根, 再通过 ϕ 换回.

最常用的线性化有两种, 分别取 $\beta(x) = x + c$ 与 $\beta(x) = sx$, 对应 Abel 方程与 Schröder 方程:

- 若 $\phi(h(x)) = \phi(x) + 1$, 则 $f(x) = \phi^{-1}\left(\phi(x) + \frac{1}{n}\right)$ 是 h 的迭代 n 次根.
- 若 $\phi(h(x)) = s\,\phi(x)$ $(s \neq 0, 1)$, 则 $f(x) = \phi^{-1}\left(s^{1/n}\phi(x)\right)$ 是 h 的迭代 n 次根.

例题 4.16 求 $h(x) = x + 1$ 的迭代平方根, 即求 f 使 $f(f(x)) = x + 1$.

解 $h(x) = x + 1$ 本身就是线性形式, 取 $\phi(x) = x$, $\beta(x) = x + 1$, 满足 $\phi(h(x)) = h(x) = x + 1 = \phi(x) + 1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由共轭化公式, 迭代平方根为

$$f(x)=\phi^{-1}\left(\phi(x)+\frac{1}{2}\right)=x+\frac{1}{2}.$$

验证: $f(f(x))=f\left(x+\frac{1}{2}\right)=x+1$. 成立.

例题 4.17 求 $h(x)=4x$ 的迭代平方根.

解 取 $\phi(x)=x, \beta(x)=4x$, 满足 $\phi(h(x))=4x=4\,\phi(x)$, 即 Schröder 方程中 $s=4$.

迭代平方根为

$$f(x)=\phi^{-1}\left(4^{1/2}\phi(x)\right)=2x.$$

验证: $f(f(x))=2(2x)=4x$. 成立.

注 4.6

共轭化依赖合适的 ϕ 的选取. 当 h 本身已是平移或伸缩时, ϕ 取恒等映射即可. 当 h 形式更复杂时, 求 ϕ 本身就是解一个函数方程(Abel 方程或 Schröder 方程), 难度不低于原问题. 共轭化的价值在于把一个难问题转化为另一个可能更简单的问题.

4.8 嵌套根式与函数方程

拉马努金提出过一个著名的嵌套根式:

$$\sqrt{1+2\sqrt{1+3\sqrt{1+4\sqrt{\cdots}}}}=3.$$

这个等式的验证过程展示了函数方程的一种独特用法, 把无限嵌套结构转化为函数方程, 再通过估值夹逼确定唯一解.

4.8.1 从嵌套根式到函数方程

将上述根式推广为含参数 x 的形式:

$$f(x)=\sqrt{1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{\cdots}}}}. \tag{4-5}$$

对 (4-5) 两边平方:

$$[f(x)]^2=1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{\cdots}}}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根号下的嵌套结构与 $f(x)$ 的定义 (4-5) 形式相同, 只是把参数 x 换成了 $x+1$:

$$\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{\cdots}}}=f(x+1).$$

代入即得 $f(x)$ 满足的函数方程

$$[f(x)]^2=1+x\,f(x+1),\quad f(x)\geq 0. \tag{4-6}$$

4.8.2 多项式试探的猜测与验证

知乎@余命数

(4-6) 不属于已知的柯西方程族. 一个直接策略是猜测解的形式. 设 $f(x)$ 为 n 次多项式, 则左端 $[f(x)]^2$ 为 $2n$ 次, 右端 $1+x\,f(x+1)$ 为 $n+1$ 次. 两端次数相等要求 $2n=n+1$, 即 $n=1$.

设 $f(x)=ax+b$, 代入 (4-6):

$$(ax+b)^2=1+x\big(a(x+1)+b\big)=1+x(ax+a+b).$$

展开并比较系数:

$$a^2x^2+2ab\,x+b^2=ax^2+(a+b)x+1.$$

对应系数相等:

$$a^2=a,\quad 2ab=a+b,\quad b^2=1.$$

由 $a^2=a$ 得 $a=1$(舍去 $a=0$, 否则 f 退化为常数). 代入 $2ab=a+b$: $2b=1+b$, 得 $b=1$. 验证 $b^2=1$: 成立.

故 $f(x)=x+1$ 是 (4-6) 的解. 代入 (4-5) 得

$$x+1=\sqrt{1+x\sqrt{1+(x+1)\sqrt{1+(x+2)\sqrt{\cdots}}}}.$$

取 $x=2$ 即得拉马努金等式: $3=\sqrt{1+2\sqrt{1+3\sqrt{1+4\sqrt{\cdots}}}}$.

4.8.3 估值夹逼证明唯一性

$f(x)=x+1$ 是 (4-6) 的解, 要说明它是 (4-5) 所定义的那个 f, 还需证明唯一性.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定理 4.11 (嵌套根式的唯一性) 设 $f(x)$ 满足函数方程 (4-6) 且对所有 $x \geq 1$ 满足估计

$$\frac{x+1}{2} \leq f(x) \leq 2(x+1). \tag{4-7}$$

则 $f(x) = x+1$.

证明 在 (4-7) 中将 x 替换为 $x+1$:

$$\frac{x+2}{2} \leq f(x+1) \leq 2(x+2).$$

代入函数方程 $[f(x)]^2 = 1 + x\,f(x+1)$, 得

$$1 + x \cdot \frac{x+2}{2} \leq [f(x)]^2 \leq 1 + x \cdot 2(x+2).$$

把 $x(x+2) = (x+1)^2 - 1$ 代入, 得

$$\frac{1 + (x+1)^2}{2} \leq [f(x)]^2 \leq 2(x+1)^2 - 1.$$

而 $1 + (x+1)^2 \geq (x+1)^2$, 开方得

$$\frac{x+1}{\sqrt{2}} \leq f(x) \leq \sqrt{2}\,(x+1).$$

与原估计 (4-7) 比较, 上下界的比例从 2 收缩到了 $\sqrt{2}$. 重复此过程 k 次, 得

$$\frac{x+1}{2^{1/2^k}} \leq f(x) \leq 2^{1/2^k}(x+1).$$

令 $k \to \infty$, $2^{1/2^k} \to 1$, 由夹逼定理得 $f(x) = x+1$. □

注 4.7

估计 (4-7) 本身可从嵌套结构直接推出. 下界: 每层根号内的 1 不小于 0, 放缩得 $f(x) \geq \sqrt{x\sqrt{x\sqrt{x\cdots}}} = x$. 上界: 每层根号内的 1 用更大的数替换, 得

$$f(x) \leq \sqrt{(x+1)\sqrt{(x+2)\sqrt{(x+3)\cdots}}} \leq 2(x+1).$$

由 $x \geq \frac{x+1}{2}$ $(x \geq 1)$, 这组估计蕴含 (4-7), 足以启动夹逼迭代.

例题 4.18 求嵌套根式 $g(x) = \sqrt{x + (x+1)\sqrt{(x+1) + (x+2)\sqrt{\cdots}}}$ 的值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 两边平方：$[g(x)]^2 = x + (x+1)\,g(x+1)$. 猜测 $g(x) = ax + b$, 代入：

$$(ax+b)^2 = x + (x+1)(a(x+1)+b).$$

展开比较系数：

$$a^2x^2 + 2ab\,x + b^2 = ax^2 + (2a+b+1)x + (a+b).$$

知乎@余命数

由 $a^2 = a$ 得 $a = 1$. 由 $2ab = 2a + b + 1$ 得 $2b = b + 3$, 故 $b = 3$. 验证常数项：$b^2 = 9$, $a + b = 4$, 不符.

线性猜测失败, 改设 $g(x) = x + c$, 即 $a = 1$, 直接代入：

$$(x+c)^2 = x + (x+1)(x+1+c).$$

知乎@余命数

左端 $x^2 + 2cx + c^2$, 右端 $x + (x+1)(x+1+c) = x^2 + (3+c)x + (1+c)$. 比较得 $2c = 3 + c$ 即 $c = 3$, 常数项 $c^2 = 9$ vs $1 + c = 4$, 仍不符.

此根式不存在线性解. 次数匹配条件 $2n = n + 1$ 已限定 $n = 1$, 多项式试探法在此失效, 应转向非多项式方法.

4.9 解题策略小结

函数方程没有万能解法, 但有一些反复有效的策略. 面对一个新问题时, 以下思路可以作为切入点.

知乎@余命数

4.9.1 识别标准模型

许多函数方程是基本初等函数的代数性质的推广. 遇到方程时先辨认它是否接近某个已知模型：

- $f(x+y) = f(x) + f(y)$, 加法型, 线性函数 $f(x) = cx$.
- $f(x+y) = f(x)f(y)$, 指数型, $f(x) = a^x$.
- $f(xy) = f(x)f(y)$, 乘法型, $f(x) = x^k$.
- $f(xy) = f(x) + f(y)$, 对数型, $f(x) = \log_a x$.
- $f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$, Jensen 型, $f(x) = ax + b$.
- $f(x+y) = g(x) + h(y)$, Pexider 型, 三个线性函数.

知乎@余命数

方程带有平移常数或尺度因子时, 先做变量代换消去常数项, 化回标准形式.

4.9.2 赋值探路

在方程中代入特殊值, 提取函数的关键信息：

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 令 $x=y=0$ 求 $f(0)$.
- 令 $y=-x$ 引入 $f(-x)$, 讨论奇偶性.
- 令 $y=x$ 得到 $f(2x)$ 与 $f(x)$ 的关系.
- 令 $y=1$ 或 $y=x$ 获取递推关系.

知乎@余命数

赋值本身不能给出完整解, 但能缩小候选范围, 为后续步骤提供线索.

4.9.3 变量代换消项

当方程含两个变量 x,y 时, 选择 y 关于 x 的特殊代换, 使等式两端的某些项相等或抵消:

知乎@余命数

- 令 y 等于使某项为零的值.
- 令 y 等于使左端某项等于右端某项的值, 触发抵消.
- 对称方程(x,y 可交换)与非对称方程联用, 交换 x,y 得到新方程, 与原方程联立求解.

4.9.4 模型假设与待定系数

当赋值和代换不足以确定解时, 假设解属于某个函数族(线性函数、多项式、指数函数等), 用待定系数法求解, 再验证所得解确实满足原方程.

多项式假设的次数判定, 若 f 为 n 次多项式, 比较 $f(x+y)$ 与 $f(x)+f(y)$ 两端的次数, 可以确定 n 的上限.

4.9.5 单调性与有界性

从方程本身提取函数的性质:

- 令 $x_2>x_1$, 考察 $f(x_2)-f(x_1)$ 的符号, 判定单调性.
- 若 $f(x)>0$ 对所有 x 成立, 可取对数转化为加法型方程.
- 若 f 有界且满足 $f(x+y)=f(x)+f(y)$, 则 $f\equiv 0$(有界加性函数必为零).

4.9.6 迭代与复合

当方程含 $f(f(x))$ 或 $f(\alpha(x))$ 时:

- 对两边复合 f, 将迭代根方程转化为交换性方程 $f\circ h=h\circ f$.
- 研究 α 的迭代轨道 (splinter), 若 $\alpha^m(x)=x$, 则对方程反复复合 m 次可能直接解出 f.
- 利用共轭化, 设 $f(x)=\phi^{-1}(\beta(\phi(x)))$, 将求 f 转化为求较简单的 β.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

4.9.7 嵌套结构的方程化

知乎@余命数

无限嵌套根式或无限连分式可以平方或取倒数, 转化为函数方程. 具体步骤:

- 引入参数 x, 把常数嵌套推广为 $f(x)$ 的方程.
- 猜测解的形式(通常是一次函数), 用待定系数法求解.
- 通过估值夹逼证明唯一性, 建立上下界, 反复迭代收紧, 取极限.

注 4.8

以上策略并非互斥, 实际解题中往往多策略联用. 先赋值提取信息, 再识别模型或假设函数族, 最后验证并证明唯一性, 是一条常见的路径.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

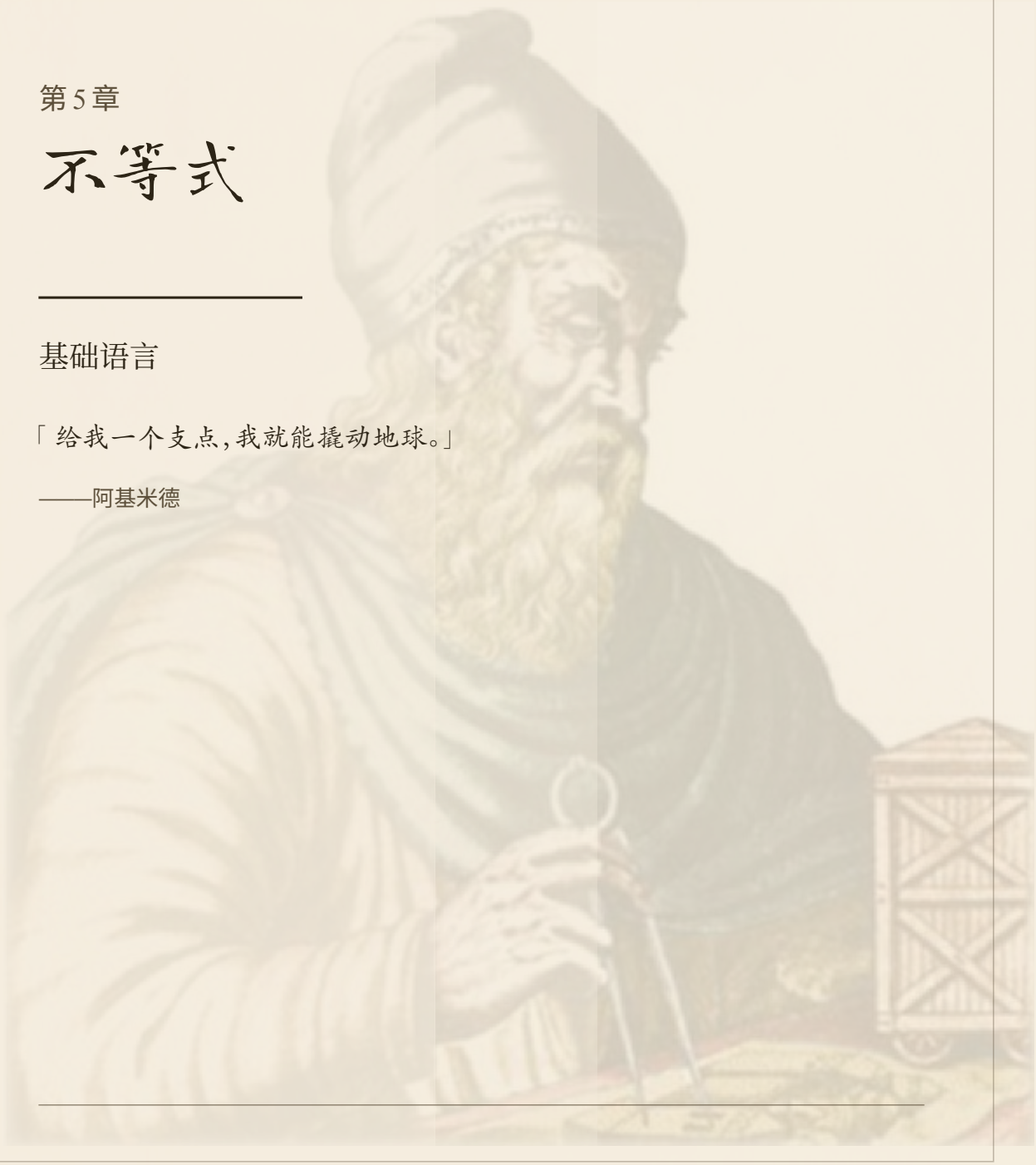
第 5 章

不等式

基础语言

「给我一个支点，我就能撬动地球。」

——阿基米德



第五章 不等式

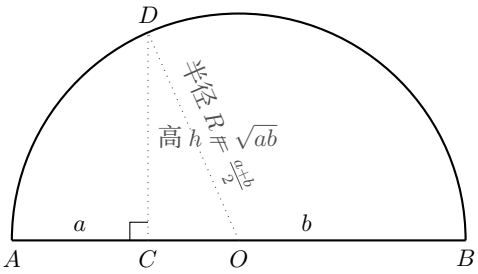


图 5.1 算术平均与几何平均的几何关系

我们研究量与量之间的大小关系. 不等式把这种关系写成严格的代数语言. 本章从最熟悉的算术-几何平均不等式出发, 逐步引入柯西不等式、排序不等式、伯努利不等式与糖水不等式, 并讨论它们在求最值、解方程和证明不等式中的应用.

本章分为两部分. 前半部分建立经典不等式本身.

- **算术-几何平均不等式**. 证明、推广及其在求最值问题中的应用.
- **柯西不等式**. 向量形式与代数形式的证明与应用.
- **排序不等式、伯努利不等式与糖水不等式**. 组合、分析与分式结构中的基本工具.

后半部分讨论证明不等式的七种思想.

- **放缩思想**. 保持方向, 剥离次要因素.
- **局部化与切线法**. 化整体为局部, 用支撑线夹住曲线.
- **构造思想**. 引入新对象, 让代数关系显形.
- **齐次化与规范化**. 次数分析, Schur 与 Muirhead.
- **代数代换**. 化约束为自由, Ravi、乘积与三角代换.
- **排序与调整**. 设序与磨平, 切比雪夫与调整法.
- **凸性思想**. 琴生不等式作为母体, 与切线法互为表里.

5.1 均值不等式

5.1.1 二元情形

我们先讨论两个变量的情形.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

定理 5.1 (二元算术-几何平均不等式) 对任意非负实数 a, b, 恒有

$$\frac{a+b}{2} \geq \sqrt{ab}$$

等号成立的充分必要条件是 $a = b$.

证明 (代数证明) 设 a, b 是两个非负实数. 实数的平方非负, $(\sqrt{a}-\sqrt{b})^2 \geq 0$. 展开此完全平方式, 得 $a - 2\sqrt{ab} + b \geq 0$. 移项整理, 即得 $a + b \geq 2\sqrt{ab}$, 进而 $\frac{a+b}{2} \geq \sqrt{ab}$.

等号成立的条件是 $(\sqrt{a}-\sqrt{b})^2 = 0$, 即 $\sqrt{a} = \sqrt{b}$, 故 $a = b$. □

证明 (几何证明) 我们给出一个几何证明. 设半圆的直径为 $a + b$.

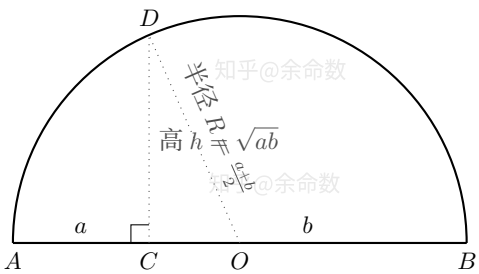


图 5.2　AM-GM 不等式的几何解释

在图中, 线段 AB 为半圆的直径, 其长度为 $a + b$. 点 C 在直径上, 满足 $AC = a, CB = b$. 半圆的半径 $R = \frac{a+b}{2}$.

过点 C 作直径的垂线, 交半圆于点 D. 在直角三角形 ADB 中(由直径所对圆周角为直角可知), 由射影定理, $CD^2 = AC \cdot CB$. 故 $CD = \sqrt{ab}$.

在直角三角形 OCD 中, 斜边 OD 是半圆的半径, 其长度为 $\frac{a+b}{2}$, 直角边 CD 的长度为 $\sqrt{ab}$. 直角三角形斜边长不小于任一直角边, 故 $OD \geq CD$. 即 $\frac{a+b}{2} \geq \sqrt{ab}$.

等号成立的条件是 $\triangle OCD$ 退化, 即点 C 与圆心 O 重合. 此时 $AC = CB$, 即 $a = b$. □

例题 5.1　求函数 $f(x) = x + \frac{9}{x}$ 在 $x > 0$ 时的最小值.

解 由 $x > 0, 9/x$ 也为正数. 对 x 和 $9/x$ 应用 AM-GM 不等式,

$$f(x) = x + \frac{9}{x} \geq 2\sqrt{x \cdot \frac{9}{x}} = 2\sqrt{9} = 6$$

函数值 $f(x)$ 的下界为 6. 为确认 6 是否为最小值, 检验等号能否成立. 等号成立的条件是 $x = \frac{9}{x}$, 即 $x^2 = 9$. 由 $x > 0$, 解得 $x = 3$.

$x = 3$ 时 $f(x) = 6$, 故函数 $f(x)$ 的最小值为 6.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

注 5.1 (应用 AM-GM 不等式求最值的三个条件)
- **各项非负**: 不等式中的各项必须为非负数.
- **和或积为定值**: 求和的最小值时, 各项的乘积须为定值; 求积的最大值时, 各项的和须为定值.
- **等号可取**: 使等号成立的条件须能在变量的允许取值范围内达到.

这三条是应用 AM-GM 不等式时的必要条件.

5.1.2 多元情形

定理 5.2 (多元算术-几何平均不等式) 对任意 n 个非负实数 $a_1, a_2, \dots, a_n$, 恒有

$$\frac{a_1 + a_2 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 a_2 \dots a_n}$$

等号成立的充分必要条件是 $a_1 = a_2 = \cdots = a_n$.

证明 (柯西归纳法) 柯西归纳法以向前推广 ($n \to 2n$) 与向后递推 ($n \to n-1$) 配合奠基 $n = 2$, 逐步覆盖一切正整数 $n \geq 2$.

先作向前推广. 设不等式对 n 个非负实数成立, 推证其对 $2n$ 个非负实数 $a_1, \dots, a_{2n}$ 亦成立. 将这 $2n$ 个数平分为两组, 每组 n 个, 对各组应用 AM-GM 不等式得

$$\frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \dots a_n},$$
$$\frac{a_{n+1} + \cdots + a_{2n}}{n} \geq \sqrt[n]{a_{n+1} \dots a_{2n}}.$$

再把这两组数的算术平均值视为两个新的非负数, 应用二元 AM-GM 不等式, 得

$$\begin{aligned}
\frac{a_1 + \cdots + a_{2n}}{2n} &= \frac{1}{2}\left(\frac{a_1 + \cdots + a_n}{n} + \frac{a_{n+1} + \cdots + a_{2n}}{n}\right) \\
&\geq \sqrt{\left(\frac{a_1 + \cdots + a_n}{n}\right)\left(\frac{a_{n+1} + \cdots + a_{2n}}{n}\right)} \\
&\geq \sqrt{(\sqrt[n]{a_1 \dots a_n})(\sqrt[n]{a_{n+1} \dots a_{2n}})} \\
&= \sqrt{\sqrt[n]{a_1 \dots a_{2n}}} = \sqrt[2n]{a_1 \dots a_{2n}}.
\end{aligned}$$

命题遂由 n 传至 $2n$; 连同 $n = 2$ 的奠基, 命题对所有 $n = 2^k$ ($k \in \mathbb{N}^*$) 成立.

再作向后递推. 设不等式对 n 个非负实数成立, 推证其对 $n-1$ 个非负实数 $a_1, \dots, a_{n-1}$ 亦成立. 为补足第 n 个数, 取这 $n-1$ 个数的算术平均值 $A_{n-1} = \frac{a_1 + \cdots + a_{n-1}}{n-1}$, 令 $a_n = A_{n-1}$. 将这 n 个数代入已知的 n 元 AM-GM 不等式, 得

$$\frac{a_1 + \cdots + a_{n-1} + a_n}{n} \geq \sqrt[n]{a_1 \dots a_{n-1} a_n}.$$

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

左端因 $a_1+\cdots+a_{n-1}=(n-1)A_{n-1}$ 化为 A_{n-1}，右端代入 $a_n=A_{n-1}$ 即 $\sqrt[n]{(a_1\dots a_{n-1})A_{n-1}}$，故

$$\begin{aligned}
&A_{n-1}\geq\sqrt[n]{(a_1\dots a_{n-1})A_{n-1}},\\
&A_{n-1}^n\geq(a_1\dots a_{n-1})A_{n-1} && \text{（两边取 } n \text{ 次方）},\\
&A_{n-1}^{n-1}\geq a_1\dots a_{n-1} && \text{（}A_{n-1}>0 \text{ 时两边同除以 } A_{n-1}\text{）},\\
&A_{n-1}\geq\sqrt[n-1]{a_1\dots a_{n-1}} && \text{（两边开 } n-1 \text{ 次方）}.
\end{aligned}$$

此即 $\frac{a_1+\cdots+a_{n-1}}{n-1}\geq\sqrt[n-1]{a_1\dots a_{n-1}}$. 若 $A_{n-1}=0$，则所有 $a_i=0$，不等式平凡成立.

向前推广使命题遍及 2 的一切幂，向后递推又把相邻两幂之间的整数尽数填满，二者合之即覆盖一切正整数 $n\geq 2$. 等号条件亦可由递推得到：$n=2$ 时两数相等；从 n 推到 $2n$ 时，等号要求两组内部均相等且两组算术平均值相等，因而 $2n$ 个数全相等；从 n 推到 $n-1$ 时，等号要求 $a_1=\cdots=a_{n-1}=a_n=A_{n-1}$，若 $A_{n-1}=0$ 则所有数均为 0. 因此等号成立当且仅当所有数相等. □

5.1.3 加权形式

定理 5.3 (加权 AM-GM 不等式) 对 n 个非负实数 $a_1,\dots,a_n$ 和 n 个满足 $\sum_{i=1}^n w_i=1$ 的正权重 $w_1,\dots,w_n$，恒有

$$w_1a_1+\cdots+w_na_n\geq a_1^{w_1}\dots a_n^{w_n}$$

等号成立的充分必要条件是 $a_1=a_2=\cdots=a_n$.

证明 若某个 $a_i=0$，则右侧 $a_1^{w_1}\dots a_n^{w_n}=0$. 若所有 a_i 均为 0，等号成立；若至少有一个 $a_i>0$，由各权重均为正，左侧 $\sum w_ia_i>0$，不等式严格成立. 下设 $a_i>0$.

该定理可由琴生不等式与对数函数的上凸性直接导出. 函数 $f(x)=\ln x$ 在 $(0,+\infty)$ 上是上凸函数，由琴生不等式有

$$\sum_{i=1}^n w_i\ln a_i\leq\ln\left(\sum_{i=1}^n w_ia_i\right).$$

利用对数性质，左端化为 $\ln(a_1^{w_1}\dots a_n^{w_n})$，故 $\ln(a_1^{w_1}\dots a_n^{w_n})\leq\ln(\sum w_ia_i)$. 对数函数单调递增，此不等式等价于 $a_1^{w_1}\dots a_n^{w_n}\leq\sum w_ia_i$. $\ln x$ 为严格上凸函数，$a_i>0$ 时等号成立当且仅当 $a_1=a_2=\cdots=a_n$. 结合零值情形，原命题的等号条件成立. □

算术平均与几何平均之间的大小关系，只是更广泛的均值序的一部分. 引入调和平均与平方平均后，可得完整的均值不等式链，详见 第5.3节.

注 5.2 (关于柯西归纳法) 柯西归纳法把向前推广与向后递推结合起来. 更多细节见https://zutomayo.org/zh-cn/posts/math_and_cauchy_induction/.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

5.2 均值不等式的提示

我们使用 AM-GM 不等式时, 必须保证各项非负、和或积为定值, 且等号条件在允许范围内可达. 否则结论可能不成立. 下面用例子说明违反这些前提会得到什么结果.

5.2.1 忽略非负前提

AM-GM 不等式仅对非负实数成立. 把它推广到全体实数会产生谬误.

例题 5.2 (变量可为负数) 设 a,b 为实数, $a+b=2$. 求 ab 的最大值.
错误分析. 不加分辨地使用 AM-GM 不等式, 会得到 $\frac{a+b}{2}\geq\sqrt{ab}$, 即 $1\geq\sqrt{ab}$, 从而 $ab\leq 1$.
这个结论不能成立. 取 $a=3,b=-1$, 满足 $a+b=2$, 但 $ab=-3$. 此时 $\sqrt{ab}$ 在实数范围内无意义, 不等式的前提已不成立.
正确分析. 此问题应视为二次函数求最值. 令 $b=2-a$, 则 $ab=a(2-a)=-a^2+2a=-(a-1)^2+1$. 这是开口向下的抛物线, 在顶点 $a=1$(此时 $b=1$)处取得最大值 1.
若题目指明 a,b 为非负实数, 两种方法的结果相同.

例题 5.3 设 x 为非零实数, 求 $f(x)=x+\frac{1}{x}$ 的最值.

解 错误分析. 忽略 x 的符号, 直接应用 AM-GM 不等式, 会得到

$$x+\frac{1}{x}\geq 2\sqrt{x\cdot\frac{1}{x}}=2$$

从而断定函数的最小值为 2. 此推理只在 $x>0$ 时有效, 因为只有 x 为正时, x 和 $1/x$ 才同时满足非负前提.

正确分析. 按 x 的符号分别讨论.

- 当 $x>0$ 时, x 和 $1/x$ 均为正数. 由 AM-GM 不等式, $f(x)=x+\frac{1}{x}\geq 2$. 等号在 $x=1$ 时成立. 在此区间, 函数的最小值为 2.
- 当 $x<0$ 时, x 和 $1/x$ 均为负数, 不满足 AM-GM 不等式的非负前提. 令 $x=-t$, 其中 $t>0$.

$$f(x)=-t+\frac{1}{-t}=-\left(t+\frac{1}{t}\right)$$

由 $t>0$, 有 $t+\frac{1}{t}\geq 2$. 因此 $f(x)=-\left(t+\frac{1}{t}\right)\leq -2$. 等号在 $t=1$, 即 $x=-1$ 时成立. 在此区间, 函数的最大值为 -2.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

该函数没有全局最小值或最大值. 它在 $x>0$ 的部分有局部最小值 2, 在 $x<0$ 的部分有局部最大值 -2.

例题 5.4 设实数 x,y 满足 $x+y=8$. 求乘积 $P=xy$ 的取值范围.

解 错误分析. 应用 AM-GM 不等式:

知乎@余命数

$$\frac{x+y}{2}\geq\sqrt{xy}\implies\frac{8}{2}\geq\sqrt{P}\implies 4\geq\sqrt{P}$$

两边平方得 $16\geq P$. 结论是 P 的最大值为 16.

这个推理有两处问题.$\sqrt{xy}$ 只在 $xy\geq 0$ 时有意义, 且 AM-GM 本身要求 x,y 非负. 若允许 x,y 为任意实数, 例如 $x=10,y=-2$, 此时 $x+y=8$, 但 $xy=-20$, $\sqrt{xy}$ 在实数域内无意义.

知乎@余命数

正确分析. 此问题不应预设 x,y 的符号, 应使用代数方法. 由 $y=8-x$, 乘积 P 表示为关于 x 的函数:

$$P(x)=x(8-x)=-x^2+8x$$

这是开口向下的二次函数. 配方:

$$P(x)=-(x^2-8x)=-((x-4)^2-16)=16-(x-4)^2$$

由 $(x-4)^2\geq 0$, 故 $P(x)\leq 16$. 当 $x=4$(此时 $y=4$)时, $P(x)$ 取得最大值 16.

知乎@余命数

x 可取任意实数, $(x-4)^2$ 可无限增大, 故 $P(x)$ 没有最小值. 因此, P 的取值范围是 $(-\infty,16]$.

最大值为 16, 且在 $x=y=4$ 处取得; 但 AM-GM 不等式没有给出 P 的下界, 因此不能代替完整的代数分析.

5.2.2 和或积非定值

应用不等式求最值时, 须构造出常数界. 若不等式的一侧仍是变量, 则无法确定最值.

例题 5.5 (界为变量) 设 $x>0$, 求 $f(x)=x^3+\frac{1}{x}$ 的最小值.
错误分析. 直接应用 AM-GM 不等式:

$$x^3+\frac{1}{x}\geq 2\sqrt{x^3\cdot\frac{1}{x}}=2\sqrt{x^2}=2x$$

不等式 $f(x)\geq 2x$ 虽成立, 但未给出常数下界.$2x$ 随 x 变化, 不能据此确定 $f(x)$ 的最小值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

正确分析. 须通过拆项构造乘积为定值的形式.

$$f(x)=x^3+\frac{1}{3x}+\frac{1}{3x}+\frac{1}{3x}\geq 4\sqrt[4]{x^3\cdot\frac{1}{3x}\cdot\frac{1}{3x}\cdot\frac{1}{3x}}=4\sqrt[4]{\frac{1}{27}}$$

取等条件为 $x^3=\frac{1}{3x}$, 即 $x^4=1/3$, $x=1/\sqrt[4]{3}$. 故最小值为 $4/\sqrt[4]{27}$.

5.2.3 取等条件无法满足

即使求出一个常数界, 也须验证使等号成立的条件能否在定义域内达到.

例题 5.6 (取等条件互相矛盾) 设 x,y 为正实数,$x\neq y$. 求 $f(x,y)=\left(x+\frac{1}{x}\right)^2+\left(y+\frac{1}{y}\right)^2$ 的最小值.

错误分析. 分别对括号内的两项应用 AM-GM 不等式: $x+\frac{1}{x}\geq 2$, 等号在 $x=1$ 时成立. $y+\frac{1}{y}\geq 2$, 等号在 $y=1$ 时成立.

因此, $f(x,y)\geq 2^2+2^2=8$.

然而使最终等号成立的条件是 $x=1$ 且 $y=1$, 这与题设 $x\neq y$ 矛盾. 因此 8 不是可取到的最小值.

正确分析. 将 $f(x,y)$ 视为关于 t 的函数 $g(t)=(t+1/t)^2$ 在 $t=x$ 和 $t=y$ 处的取值之和. $g'(t)=2(t+1/t)(1-1/t^2)$. 在 $(0,1)$ 上 $g'(t)<0$, 在 $(1,\infty)$ 上 $g'(t)>0$. 函数在 $t=1$ 处取得最小值 4. 由 $x\neq y$, $g(x)+g(y)$ 不能同时取到 $4+4=8$, 故 $f(x,y)>8$. 另一方面, 取 $x=1$, $y=1+\varepsilon$ $(\varepsilon>0)$, 则 $x\neq y$, 且当 $\varepsilon\to 0^+$ 时

$$f(1,1+\varepsilon)=4+\left(1+\varepsilon+\frac{1}{1+\varepsilon}\right)^2\to 8.$$

因此, 在约束 $x,y>0$ 且 $x\neq y$ 下, $f(x,y)$ **没有最小值**; 其下确界为 8.

用导数讨论最值

例题 5.7 (取等点不在定义域内) 设 $x\geq 2$, 求 $f(x)=x+\frac{1}{x}$ 的最小值.

错误分析. $x+\frac{1}{x}\geq 2\sqrt{x\cdot\frac{1}{x}}=2$. 由此得出最小值为 2.

这个结论错误. 等号成立的条件是 $x=1$, 但该点不在定义域 $[2,+\infty)$ 内. 函数值可以大于 2, 但永远无法达到 2.

正确分析. 须分析函数在指定定义域上的单调性. $f'(x)=1-\frac{1}{x^2}$. 当 $x\geq 2$ 时, $x^2\geq 4$, $0<\frac{1}{x^2}\leq\frac{1}{4}$, 故 $f'(x)>0$. 函数 $f(x)$ 在 $[2,+\infty)$ 上严格单调递增. 其最小值在区间的左端点 $x=2$ 处取得. 最小值为 $f(2)=2+\frac{1}{2}=2.5$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.3 均值不等式链

知乎@余命数

上一节已证明算术平均不小于几何平均. 下面我们引入调和平均与平方平均, 给出完整的不等式链.

知乎@余命数

定义 5.1 (调和平均与平方平均) 对 n 个正实数 $a_1, a_2, \dots, a_n$:

- **调和平均 (Harmonic Mean, HM)** 定义为

$$H_n = \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}}$$

- **平方平均 (Quadratic Mean, QM)** 定义为

$$Q_n = \sqrt{\frac{a_1^2 + a_2^2 + \cdots + a_n^2}{n}}$$

调和平均等于倒数的算术平均之倒数. 平方平均等于平方和的算术平均之根. 这四种平均值之间存在固定的序关系.

知乎@余命数

定理 5.4 (均值不等式链) 对任意 n 个正实数 $a_1, a_2, \dots, a_n$, 恒有

$$\frac{n}{\sum_{i=1}^{n} \frac{1}{a_i}} \le \sqrt[n]{\prod_{i=1}^{n} a_i} \le \frac{\sum_{i=1}^{n} a_i}{n} \le \sqrt{\frac{\sum_{i=1}^{n} a_i^2}{n}}$$

即 $H_n \le G_n \le A_n \le Q_n$. 在任意一步不等式中, 等号成立的充分必要条件均为 $a_1 = a_2 = \cdots = a_n$.

证明 上一节已证 $G_n \le A_n$. 这里只证其余两个不等关系.

先证 $H_n \le G_n$. 考虑 n 个正实数 $\frac{1}{a_1}, \frac{1}{a_2}, \dots, \frac{1}{a_n}$, 对它们应用算术-几何平均不等式, 得

$$\frac{\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}}{n} \ge \sqrt[n]{\frac{1}{a_1}\frac{1}{a_2}\cdots\frac{1}{a_n}}$$

整理得

$$\frac{\sum_{i=1}^{n} \frac{1}{a_i}}{n} \ge \frac{1}{\sqrt[n]{\prod_{i=1}^{n} a_i}}$$

知乎@余命数

即 $\frac{1}{H_n} \ge \frac{1}{G_n}$. 各项均为正数, 两边取倒数, 不等号方向反转, 即得 $H_n \le G_n$.

知乎@余命数

再证 $A_n \le Q_n$. 此不等式等价于 $(\frac{\sum a_i}{n})^2 \le \frac{\sum a_i^2}{n}$, 可由柯西不等式给出. 取两个 n 维向量 $\mathbf{a} = (a_1, \dots, a_n)$ 和 $\mathbf{u} = (1, \dots, 1)$, 由柯西不等式 $(\mathbf{a} \cdot \mathbf{u})^2 \le \|\mathbf{a}\|^2 \|\mathbf{u}\|^2$ 得

$$\left(\sum_{i=1}^{n} a_i \cdot 1 \right)^2 \le \left(\sum_{i=1}^{n} a_i^2 \right) \left(\sum_{i=1}^{n} 1^2 \right),$$

即

$$\left(\sum a_i \right)^2 \le \left(\sum a_i^2 \right) \cdot n.$$

不等式两边同除以 n^2, 得

$$\frac{(\sum a_i)^2}{n^2} \le \frac{\sum a_i^2}{n},$$

即 $(\frac{\sum a_i}{n})^2 \le \frac{\sum a_i^2}{n}$. 各项为正, 两边开平方根, 即得 $A_n \le Q_n$.

等号条件分别继承自所用的不等式.$H_n \le G_n$ 中等号当且仅当 $1/a_1 = \cdots = 1/a_n$, 即 $a_1 = \cdots = a_n$; $G_n \le A_n$ 的等号条件为 $a_1 = \cdots = a_n$; $A_n \le Q_n$ 中柯西等号当且仅当 $(a_1, \dots, a_n)$ 与 $(1, \dots, 1)$ 共线, 也即 $a_1 = \cdots = a_n$. 因此链中任意一步等号成立的充分必要条件均为所有数相等. □

例题 5.8 已知 x, y 为正实数, 且 $x + y = 4$. 求 $x^2 + y^2$ 的最小值.

解 我们用 $A_2 \le Q_2$:

$$\frac{x+y}{2} \le \sqrt{\frac{x^2+y^2}{2}}$$

将 $x + y = 4$ 代入, 得

$$\frac{4}{2} \le \sqrt{\frac{x^2+y^2}{2}}$$

$$2 \le \sqrt{\frac{x^2+y^2}{2}}$$

两边均为正, 两边平方, 得

$$4 \le \frac{x^2+y^2}{2}$$

由此 $x^2 + y^2 \ge 8$.

等号成立当且仅当 $x = y$. 结合 $x + y = 4$, 解得 $x = y = 2$.

$x = y = 2$ 使等号可以取到, 故 $x^2 + y^2$ 的最小值为 8.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.4 均值不等式的应用技巧

我们常用均值不等式处理和与积的最值问题. 许多问题的代数结构须先变形才满足 AM-GM 的前提. 本节讨论几种常用的变形技巧.

5.4.1 拆项与凑项创造定值

算术-几何平均不等式要求各项之和或各项之积为定值. 当条件不满足时, 可拆分或添加项, 构造新的定值表达式.

例题 5.9 求函数 $f(x)=x^2+\frac{2}{x}$ 在 $x>0$ 时的最小值.

解 表达式 $x^2+\frac{2}{x}$ 的两项之积 $x^2\cdot\frac{2}{x}=2x$ 随 x 变化, 不能直接应用二元均值不等式.

为得到定值乘积, 须消去变量 x.x^2 与两个 $1/x$ 相乘后乘积为常数, 故把 $\frac{2}{x}$ 拆成 $\frac{1}{x}+\frac{1}{x}$.

原函数变为三项之和.

$$f(x)=x^2+\frac{1}{x}+\frac{1}{x}$$

这三项均为正数, 且乘积为 $x^2\cdot\frac{1}{x}\cdot\frac{1}{x}=1$, 是定值.

应用三元 AM-GM 不等式,

$$f(x)=x^2+\frac{1}{x}+\frac{1}{x}\geq 3\sqrt[3]{x^2\cdot\frac{1}{x}\cdot\frac{1}{x}}=3\sqrt[3]{1}=3$$

下界为 3. 等号要求所有项相等, 即 $x^2=\frac{1}{x}$. 解得 $x^3=1$, 又 $x>0$, 故 $x=1$.

存在 $x=1$ 使 $f(1)=3$, 因此函数 $f(x)$ 的最小值为 3.

5.4.2 变量代换

当变量受到条件约束时, 引入新变量常可简化代数结构.

例题 5.10 设 a,b,c 为三角形的三边长. 证明, 得

$$\frac{a}{b+c-a}+\frac{b}{c+a-b}+\frac{c}{a+b-c}\geq 3$$

解 不等式左侧分母为三角形三边的线性组合. 通过变量代换可简化问题.

定义新变量 $x=b+c-a$、$y=c+a-b$、$z=a+b-c$. 由三角形不等式(任意两边之和大于第三边), x,y,z 均为正数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

反解得

$$y+z=(c+a-b)+(a+b-c)=2a \implies a=\frac{y+z}{2}$$
$$z+x=(a+b-c)+(b+c-a)=2b \implies b=\frac{z+x}{2}$$
$$x+y=(b+c-a)+(c+a-b)=2c \implies c=\frac{x+y}{2}$$

代入原不等式左侧:

$$\begin{aligned}\text{LHS}&=\frac{(y+z)/2}{x}+\frac{(z+x)/2}{y}+\frac{(x+y)/2}{z}\\&=\frac{1}{2}\left(\frac{y+z}{x}+\frac{z+x}{y}+\frac{x+y}{z}\right)\\&=\frac{1}{2}\left(\frac{y}{x}+\frac{z}{x}+\frac{z}{y}+\frac{x}{y}+\frac{x}{z}+\frac{y}{z}\right)\end{aligned}$$

整理后得到互为倒数的项对.

$$\text{LHS}=\frac{1}{2}\left[\left(\frac{x}{y}+\frac{y}{x}\right)+\left(\frac{y}{z}+\frac{z}{y}\right)+\left(\frac{z}{x}+\frac{x}{z}\right)\right]$$

对每一对应用二元均值不等式, 例如 $\frac{x}{y}+\frac{y}{x}\geq 2$. 于是

$$\text{LHS}\geq\frac{1}{2}(2+2+2)=3$$

等号成立的条件是 $\frac{x}{y}=1,\frac{y}{z}=1,\frac{z}{x}=1$, 即 $x=y=z$. 这对应 $b+c-a=c+a-b=a+b-c$, 解得 $a=b=c$. 因此当三角形为等边三角形时, 等号成立.

5.4.3 利用取等条件构造变形

对称结构的不等式, 极值点通常在所有变量相等时取到. 这一事实可反过来推断拆分方式.

例题 5.11 求函数 $f(x)=2x+\frac{1}{x^2}$ 在 $x>0$ 时的最小值.

解 为构造乘积为定值的形式, 须拆分 $2x$. 拆法由均值不等式的取等条件反推.

设函数在 x_0 处取得最小值, 应用不等式的各项必须相等. 将 $2x$ 拆为 n 个相等的项 $\frac{2x}{n}$, 与 $\frac{1}{x^2}$ 一同应用 $n+1$ 元均值不等式. 为使乘积为定值, 各项中 x 的总幂次应为零.

乘积为 $\left(\frac{2x}{n}\right)^n\cdot\frac{1}{x^2}$, 其中 x 的幂次为 $n-2$. 令 $n-2=0$, 得 $n=2$.

知乎@余命数

因此把 $2x$ 拆成 $x+x$. 对 $x, x, \frac{1}{x^2}$ 三项应用均值不等式,

$$f(x)=x+x+\frac{1}{x^2}\geq 3\sqrt[3]{x\cdot x\cdot\frac{1}{x^2}}=3\sqrt[3]{1}=3$$

等号成立条件为 $x=\frac{1}{x^2}$, 即 $x^3=1$. 由 $x>0$, 解得 $x=1$.

$x=1$ 时, $f(1)=2+1=3$. 最小值可以取到. 故函数 $f(x)$ 的最小值为 3.

5.4.4 整数拆分中的乘积最值

前面几个例子都在实数范围内寻找极值. 当变量被迫取整数时, 均值不等式的取等条件常常无法满足, 最优解要另想办法.

例题 5.12 (拆数使乘积最大) 把正整数 1976 拆成若干个正整数之和, 使这些数的乘积最大, 求最大的乘积.

解 先让题面变小. 记 $M(n)$ 为把 n 拆开所能得到的最大乘积, 从较小的数数起.

$$M(2)=2,\quad M(3)=3,\quad M(4)=2\times 2=4,$$

$$M(5)=3\times 2=6,\quad M(6)=3\times 3=9,\quad M(7)=3\times 2\times 2=12.$$

出现的因数只有 2 与 3, 由此猜想, 最优的拆法只用 2 与 3. 要断定这件事, 可以从任意一个最优拆法出发, 用改动把它整理成只用 2 与 3 的样子.

若某个因数 $a\geq 5$, 把它换成 3 与 $a-3$ 两个因数, 乘积由 a 变为

$$3(a-3)=3a-9>a,$$

因为 $a\geq 5$ 时 $2a>9$. 改动之后乘积变大, 与最优性矛盾, 所以最优拆法里的因数都不超过 4.

接着把 4 换成 2 与 2. 两个 2 的乘积仍是 4, 因数变多而乘积不变, 故可以认为最优拆法只用 2 与 3. 再看三个 2, 因为

$$2\times 2\times 2=8<9=3\times 3,$$

而两边的因数之和都是 6, 所以最优拆法里至多出现两个 2.

剩下的只是代数. 记 $n=3q+r$, 其中 $r\in\{0,1,2\}$.

- $r=0$ 时全部用 3, 最大乘积为 3^q;
- $r=1$ 时取出一个 3 与余数拼成 $2+2$, 最大乘积为 $3^{q-1}\times 2\times 2$;
- $r=2$ 时保留这个余数, 最大乘积为 $3^q\times 2$.

由 $1976=3\times 658+2$ 得, 最大乘积为 $3^{658}\times 2$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

处理连续极值靠均值不等式的取等条件, 处理整数极值靠逐项调整, 改动一步不会让结果变好就说明已经最优. 固定因数之和后, 由 第5.1节 的均值不等式, 各因数越接近越好, 连续版本的答案是每个因数都等于 $e \approx 2.718$. 整数只能取 2 与 3, 而 $2 < e < 3$, 这就是最优拆法只用 2 与 3 的来源.

5.4.5 归一化方法与齐次性

我们称一个不等式为齐次的, 如果当所有变量同乘以一个正数因子时, 不等关系保持不变. 这类不等式可在不失一般性的前提下对变量施加额外约束, 如和为 1 或积为 1, 以简化问题.

例题 5.13 (内斯比特不等式) 设 a, b, c 为正实数. 证明, 得

$$\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} \geq \frac{3}{2}$$

知乎@余命数

解 考察不等式左侧表达式的齐次性. 以 ta, tb, tc $(t > 0)$ 替换 a, b, c, 表达式的值不变.

知乎@余命数

$$\frac{ta}{tb+tc} + \frac{tb}{tc+ta} + \frac{tc}{ta+tb} = \frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b}$$

故该不等式是 0 次齐次的.

由齐次性, 可设 $a+b+c=1$. 在此约束下, $b+c = 1-a$、$c+a = 1-b$、$a+b = 1-c$. 不等式左侧变为:

$$\text{LHS} = \frac{a}{1-a} + \frac{b}{1-b} + \frac{c}{1-c} = \left(\frac{1}{1-a} - 1\right) + \left(\frac{1}{1-b} - 1\right) + \left(\frac{1}{1-c} - 1\right)$$

$$\text{LHS} = \frac{1}{b+c} + \frac{1}{c+a} + \frac{1}{a+b} - 3$$

原不等式等价于证明

知乎@余命数

$$\frac{1}{b+c} + \frac{1}{c+a} + \frac{1}{a+b} - 3 \geq \frac{3}{2} \implies \frac{1}{b+c} + \frac{1}{c+a} + \frac{1}{a+b} \geq \frac{9}{2}$$

此形式可由柯西不等式的 Engel 形式直接证明.

$$\frac{1^2}{b+c} + \frac{1^2}{c+a} + \frac{1^2}{a+b} \geq \frac{(1+1+1)^2}{(b+c)+(c+a)+(a+b)} = \frac{9}{2(a+b+c)}$$

知乎@余命数

由 $a+b+c=1$, 不等式成立.

原不等式是齐次的, 此结论对所有正实数 a, b, c 均成立.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.5 柯西不等式

注 5.3 (表述约定)
我们先写出柯西不等式的代数形式 $\left(\sum a_i b_i\right)^2 \le \left(\sum a_i^2\right)\left(\sum b_i^2\right)$；把它解释为向量点积则留到第11章.

我们关心向量内积与向量长度之间的关系. 柯西不等式给出了二者之间的一个定量联系.

5.5.1 向量形式

定理 5.5 (柯西不等式的向量形式) 对 n 维欧几里得空间中的任意两个向量 $\mathbf{a}=(a_1,\dots,a_n)$ 和 $\mathbf{b}=(b_1,\dots,b_n)$, 恒有

$$(\mathbf{a}\cdot\mathbf{b})^2 \le \|\mathbf{a}\|^2\|\mathbf{b}\|^2$$

其中 $\mathbf{a}\cdot\mathbf{b}=\sum_{i=1}^n a_i b_i$ 是向量的点积, $\|\mathbf{a}\|=\sqrt{\sum a_i^2}$ 是向量的欧几里得范数. 等价地,

$$\left(\sum_{i=1}^n a_i b_i\right)^2 \le \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right)$$

等号成立的充分必要条件是向量 $\mathbf{a}$ 与 $\mathbf{b}$ 线性相关；特别地, 任一向量为零向量时也取等号.

证明 (基于二次判别式) 我们考虑关于实数 t 的二次多项式 $P(t)=\sum_{i=1}^n (a_i t-b_i)^2$.

$P(t)$ 是若干个平方项之和, 因此对任意实数 t, 恒有 $P(t)\ge 0$.

将 $P(t)$ 展开得

$$\begin{aligned} P(t) &= \sum_{i=1}^n (a_i^2 t^2 - 2a_i b_i t + b_i^2) \\ &= \left(\sum a_i^2\right) t^2 - 2\left(\sum a_i b_i\right) t + \left(\sum b_i^2\right). \end{aligned}$$

它是关于 t 的二次函数(若 $\sum a_i^2=0$ 则为一次函数).

若 $\sum a_i^2=0$, 则所有 $a_i=0$, 不等式显然成立. 若 $\sum a_i^2>0$, 则 $P(t)$ 是开口向上的二次函数. 由它恒非负可知其图像与 t 轴至多有一个交点, 因此判别式 $\Delta\le 0$.

$$\Delta = \left(-2\sum a_i b_i\right)^2 - 4\left(\sum a_i^2\right)\left(\sum b_i^2\right) \le 0$$

$$4\left(\sum a_i b_i\right)^2 \le 4\left(\sum a_i^2\right)\left(\sum b_i^2\right)$$

知乎@余命数　知乎@余命数

此即柯西不等式.

等号成立的条件是存在实数 t_0 使 $P(t_0)=0$. 各项非负, 这等价于 $a_i t_0 - b_i = 0$ 对所有 i 成立, 即 $\mathbf{b} = t_0\mathbf{a}$, 两向量线性相关. 若 $\mathbf{a}=\mathbf{0}$ 或 $\mathbf{b}=\mathbf{0}$, 等号也成立且同样属于线性相关情形. 反过来, 若两向量线性相关, 直接代入可得等号. □

例题 5.14 已知 x, y, z 为实数, 且 $x+2y+3z=7$. 求 $x^2+y^2+z^2$ 的最小值.

解 目标是一个二次式, 约束是一个线性方程. 我们用柯西不等式建立二者之间的联系.

我们将目标表达式 $x^2+y^2+z^2$ 与约束表达式 $x+2y+3z$ 分别视为向量范数的平方与点积.

令向量 $\mathbf{a}=(x,y,z)$ 和 $\mathbf{b}=(1,2,3)$.

根据柯西不等式.

$$(x\cdot 1 + y\cdot 2 + z\cdot 3)^2 \le (x^2+y^2+z^2)(1^2+2^2+3^2)$$

代入 $x+2y+3z=7$, 得

$$7^2 \le (x^2+y^2+z^2)(1+4+9)$$

$$49 \le 14(x^2+y^2+z^2)$$

解得 $x^2+y^2+z^2 \ge \frac{49}{14} = \frac{7}{2}$.

等号成立的条件是向量 $\mathbf{a}$ 与 $\mathbf{b}$ 共线, 即存在常数 k 使得 $(x,y,z) = k(1,2,3)$. 即 $\frac{x}{1}=\frac{y}{2}=\frac{z}{3}=k$.

将 $x=k, y=2k, z=3k$ 代入 $x+2y+3z=7$, 得

$$k+2(2k)+3(3k)=7$$

$$k+4k+9k=14k=7$$

解得 $k=1/2$.

存在满足等号条件的实数 x, y, z, 即 $x=1/2, y=1, z=3/2$, 因此最小值为 $7/2$.

5.5.2 推论

我们给出柯西不等式的几个推论, 它们分别对应一些常见的代数结构.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 5.6 (柯西不等式的 Engel 形式) 设 $a_1,\dots,a_n$ 为任意实数, $b_1,\dots,b_n$ 为正实数. 则

$$\frac{a_1^2}{b_1}+\frac{a_2^2}{b_2}+\cdots+\frac{a_n^2}{b_n}\geq\frac{(a_1+a_2+\cdots+a_n)^2}{b_1+b_2+\cdots+b_n}$$

等号成立当且仅当 $\frac{a_1}{b_1}=\frac{a_2}{b_2}=\cdots=\frac{a_n}{b_n}$.

知乎@余命数

证明 它可由柯西不等式直接得到. 令向量 $\mathbf{u}=\left(\frac{a_1}{\sqrt{b_1}},\frac{a_2}{\sqrt{b_2}},\dots,\frac{a_n}{\sqrt{b_n}}\right)$ 和 $\mathbf{v}=(\sqrt{b_1},\sqrt{b_2},\dots,\sqrt{b_n})$.

由柯西不等式 $(\mathbf{u}\cdot\mathbf{v})^2\leq\|\mathbf{u}\|^2\|\mathbf{v}\|^2$, 分别计算两端, 得

$$(\mathbf{u}\cdot\mathbf{v})^2=\left(\sum_{i=1}^n\frac{a_i}{\sqrt{b_i}}\sqrt{b_i}\right)^2=\left(\sum_{i=1}^n a_i\right)^2 \qquad \text{(点积的平方)},$$

$$\|\mathbf{u}\|^2=\sum_{i=1}^n\left(\frac{a_i}{\sqrt{b_i}}\right)^2=\sum_{i=1}^n\frac{a_i^2}{b_i} \qquad (\mathbf{u}\text{ 的范数平方}),$$

$$\|\mathbf{v}\|^2=\sum_{i=1}^n(\sqrt{b_i})^2=\sum_{i=1}^n b_i \qquad (\mathbf{v}\text{ 的范数平方}).$$

代入不等式, 即得

$$\left(\sum a_i\right)^2\leq\left(\sum\frac{a_i^2}{b_i}\right)\left(\sum b_i\right),$$

整理后即为所证.

等号成立的条件是向量 $\mathbf{u}$ 与 $\mathbf{v}$ 共线, 即存在常数 k 使 $\mathbf{u}=k\mathbf{v}$. 分量形式为 $\frac{a_i}{\sqrt{b_i}}=k\sqrt{b_i}$, 即 $a_i=kb_i$, 故 $\frac{a_i}{b_i}=k$ 对所有 i 成立. 反过来, 若 $\frac{a_1}{b_1}=\cdots=\frac{a_n}{b_n}=k$, 则 $\mathbf{u}=k\mathbf{v}$, 柯西不等式取等号. □

定理 5.7 (闵可夫斯基不等式, $p=2$ 情形) 设 $a_1,\dots,a_n$ 和 $b_1,\dots,b_n$ 为任意实数. 则

$$\sqrt{a_1^2+b_1^2}+\cdots+\sqrt{a_n^2+b_n^2}\geq\sqrt{(a_1+\cdots+a_n)^2+(b_1+\cdots+b_n)^2}$$

等号成立当且仅当所有向量 (a_i,b_i) 同向共线, 即存在同一个非零向量 $\mathbf{u}$ 和非负实数 λ_i 使 $(a_i,b_i)=\lambda_i\mathbf{u}$; 零向量可视为任意方向.

证明 它在几何上即欧几里得空间中的三角不等式. 考虑 n 个二维向量 $\mathbf{v}_i=(a_i,b_i)$. 不等式左侧为 $\sum_{i=1}^n\|\mathbf{v}_i\|$, 右侧为 $\|\sum_{i=1}^n\mathbf{v}_i\|$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们对不等式两边平方, 再应用柯西不等式证明. 令 $A=\sum a_i$, $B=\sum b_i$. 先证 $n=2$ 的情形, 即

$$\sqrt{a_1^2+b_1^2}+\sqrt{a_2^2+b_2^2}\geq\sqrt{(a_1+a_2)^2+(b_1+b_2)^2}.$$

知乎@余命数

两边平方, 得

知乎@余命数

$$(a_1^2+b_1^2)+(a_2^2+b_2^2)+2\sqrt{(a_1^2+b_1^2)(a_2^2+b_2^2)}\geq(a_1^2+2a_1a_2+a_2^2)+(b_1^2+2b_1b_2+b_2^2),$$

化简得 $\sqrt{(a_1^2+b_1^2)(a_2^2+b_2^2)}\geq a_1a_2+b_1b_2$. 若右侧为负, 不等式显然成立; 若右侧非负, 两边平方得

知乎@余命数

$$(a_1^2+b_1^2)(a_2^2+b_2^2)\geq(a_1a_2+b_1b_2)^2,$$

知乎@余命数

此即二维柯西不等式.

一般情形可由数学归纳法, 重复应用 $n=2$ 的结论得到. 等号条件来自二维情形中柯西取等号且 $a_1a_2+b_1b_2\geq 0$, 这要求每次合并时参与相加的向量同向共线. 因而所有非零向量 (a_i,b_i) 必须同向共线; 反过来, 若满足此条件, 则向量长度相加等于和向量长度, 故取等号. □

注 5.4
这即权方和不等式, 1980 年代后得名.

例题 5.15 已知 x,y 为实数, 满足 $x+3y=4$. 求 $4x^2+y^2$ 的最小值.

知乎@余命数

解 目标函数 $4x^2+y^2$ 是二次和, 可写成 $(2x)^2+y^2$; 而约束 $x+3y=4$ 是线性的. 我们只需在二者之间建立联系.

知乎@余命数

为了匹配目标函数, 我们把 x 改写成 $2x$ 的一半.

$$x+3y=\frac{1}{2}(2x)+3y=4$$

令向量 $\mathbf{a}=(\frac{1}{2},3)$ 和 $\mathbf{b}=(2x,y)$. 根据柯西不等式, $(\mathbf{a}\cdot\mathbf{b})^2\leq\|\mathbf{a}\|^2\|\mathbf{b}\|^2$.

$$\left(\frac{1}{2}(2x)+3y\right)^2\leq\left(\left(\frac{1}{2}\right)^2+3^2\right)\left((2x)^2+y^2\right)$$

代入 $x+3y=4$, 得

知乎@余命数

$$4^2\leq\left(\frac{1}{4}+9\right)(4x^2+y^2)$$

知乎@余命数

$$16\leq\frac{37}{4}(4x^2+y^2)$$

由此 $4x^2+y^2\geq\frac{64}{37}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等号成立的条件是向量 **a** 与 **b** 共线, 即 $\frac{2x}{1/2}=\frac{y}{3}$. 即 $4x=y/3$, 或 $y=12x$. 代入 $x+3(12x)=4$, 得 $37x=4$, $x=4/37$. 此时 $y=12(4/37)=48/37$.

由于存在使等号成立的实数 x,y, 故 $4x^2+y^2$ 的最小值为 $\frac{64}{37}$.

例题 5.16 已知实数 a,b,c 和 x,y,z 满足 $a^2+b^2+c^2=9$ 和 $x^2+y^2+z^2=1$. 求 $ax+by+cz$ 的最大值.

解 问题涉及两个二次和的定值与一个线性组合的最值. 这是柯西不等式的直接应用.

令向量 $\mathbf{u}=(a,b,c)$ 和 $\mathbf{v}=(x,y,z)$. 根据柯西不等式.

$$(ax+by+cz)^2\le(a^2+b^2+c^2)(x^2+y^2+z^2)$$

代入已知条件:

$$(ax+by+cz)^2\le 9\cdot 1=9$$

由此 $|ax+by+cz|\le 3$, 即 $-3\le ax+by+cz\le 3$.

表达式的最大值为 3. 等号成立的条件是向量 **u** 与 **v** 共线, 即存在常数 k 使得 $(a,b,c)=k(x,y,z)$. 此时 $a^2+b^2+c^2=k^2(x^2+y^2+z^2)$, 即 $9=k^2(1)$, 解得 $k=\pm 3$. 当 $k=3$ 时, $a=3x,b=3y,c=3z$, 此时 $ax+by+cz=3x^2+3y^2+3z^2=3(x^2+y^2+z^2)=3$.

由于最大值可以取到, 故 $ax+by+cz$ 的最大值为 3.

例题 5.17 已知 a,b,c 均为正数, 且 $a^2+b^2+4c^2=3$. 证明:

1. $a+b+2c\le 3$.
2. 若 $b=2c$, 则 $\frac{1}{a}+\frac{1}{c}\ge 3$.

解

1. 待证不等式是线性的, 条件是二次的. 我们构造向量并应用柯西不等式.
   把条件 $a^2+b^2+4c^2=3$ 写成 $a^2+b^2+(2c)^2=3$. 令向量 $\mathbf{u}=(a,b,2c)$ 和 $\mathbf{v}=(1,1,1)$.
   根据柯西不等式.

$$(a\cdot 1+b\cdot 1+2c\cdot 1)^2\le(a^2+b^2+(2c)^2)(1^2+1^2+1^2)$$

$$(a+b+2c)^2\le(a^2+b^2+4c^2)(3)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入 $a^2+b^2+4c^2=3$, 得

知乎@余命数

$$(a+b+2c)^2 \le 3\cdot 3 = 9$$

由于 a,b,c 均为正数, $a+b+2c$ 为正. 两边开平方根得 $a+b+2c\le 3$. 等号成立当且仅当 $\frac{a}{1}=\frac{b}{1}=\frac{2c}{1}$, 即 $a=b=2c$. 代入条件得 $a^2+a^2+a^2=3$, $a=1$. 此时 $a=b=1, c=1/2$.

2. 当 $b=2c$ 时, 原条件 $a^2+b^2+4c^2=3$ 变为

$$a^2+(2c)^2+4c^2 = a^2+8c^2 = 3$$

我们需在此条件下证明 $\frac{1}{a}+\frac{1}{c}\ge 3$.
这个结构适合使用柯西不等式的 Engel 形式. 为了构造合适的项, 我们将 $\frac{1}{c}$ 写成 $\frac{4}{4c}$.

知乎@余命数

$$\frac{1}{a}+\frac{1}{c} = \frac{1^2}{a}+\frac{2^2}{4c} \ge \frac{(1+2)^2}{a+4c} = \frac{9}{a+4c}$$

接下来我们求 $a+4c$ 的上界. 利用条件 $a^2+8c^2=3$, 我们再次应用柯西不等式.

$$(a+4c)^2 = (1\cdot a+\sqrt{2}\cdot\sqrt{8}c)^2 \le (1^2+(\sqrt{2})^2)(a^2+(\sqrt{8}c)^2)$$

知乎@余命数 知乎@余命数

$$(a+4c)^2 \le (1+2)(a^2+8c^2) = 3\cdot 3 = 9$$

知乎@余命数

由于 a,c 为正, $a+4c\le 3$.
结合两个结果:

知乎@余命数

$$\frac{1}{a}+\frac{1}{c} \ge \frac{9}{a+4c} \ge \frac{9}{3} = 3$$

检验等号成立的条件.Engel 形式的放缩要求 $\frac{1}{a}=\frac{2}{4c}$, 即 $a=2c$; 柯西不等式的取等要求 $\frac{a}{1}=\frac{\sqrt{8}c}{\sqrt{2}}$, 即 $a=2c$. 两个条件一致. 将 $a=2c$ 代入 $a^2+8c^2=3$, 得 $(2c)^2+8c^2=12c^2=3$, $c=1/2$. 此时 $a=1$. 当 $a=1, c=1/2$ 时, $\frac{1}{a}+\frac{1}{c}=1+2=3$, 等号可以取到.

例题 5.18 已知点 $P(a,b)$ 在圆 $C: x^2+y^2=x+y$ 上, 且 $a,b\in(0,+\infty)$.

1. 求 $\frac{1}{a}+\frac{1}{b}$ 的最小值.
2. 是否存在 a,b 满足 $(a+1)(b+1)=4$? 若存在, 求出 a,b 的值; 若不存在, 说明理由.

解

1. 由题设, a,b 为正实数, 且满足 $a^2+b^2=a+b$. 所求表达式为 $\frac{1}{a}+\frac{1}{b}=\frac{a+b}{ab}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据均值不等式, $a^2+b^2 \geq 2ab$. 将条件代入, 得 $a+b \geq 2ab$. 由于 $a,b>0$, 可得 $\frac{a+b}{ab} \geq 2$.

等号成立的条件是 $a=b$. 代入圆的方程, $a^2+a^2=a+a$, 即 $2a^2=2a$. 由于 $a>0$, 解得 $a=1$. 故当 $a=b=1$ 时, $\frac{1}{a}+\frac{1}{b}$ 取得最小值 2.

2. 问题等价于判断表达式 $ab+a+b+1$ 的值能否达到 4. 先确定 $a+b$ 的范围. 由柯西不等式,

$$(a+b)^2=(1\cdot a+1\cdot b)^2 \leq (1^2+1^2)(a^2+b^2)=2(a^2+b^2)$$

将条件 $a^2+b^2=a+b$ 代入,

$$(a+b)^2 \leq 2(a+b)$$

由于 $a,b>0, a+b>0$, 两边同除以 $a+b$, 得 $a+b \leq 2$.

同时, 由均值不等式, $ab \leq \frac{a^2+b^2}{2}=\frac{a+b}{2}$.

现在考察表达式 $(a+1)(b+1)$:

$$(a+1)(b+1)=ab+a+b+1 \leq \frac{a+b}{2}+(a+b)+1=\frac{3}{2}(a+b)+1$$

由于 $a+b \leq 2$, 故

$$\frac{3}{2}(a+b)+1 \leq \frac{3}{2}(2)+1=4$$

故 $(a+1)(b+1)$ 的最大值为 4.

等号成立的条件是上述所有不等式均取等号, 即

- $a+b=2$
- $a=b$

联立解得 $a=b=1$.

将 $a=b=1$ 代入原方程 $a^2+b^2=a+b$ 检验.$1^2+1^2=1+1$, 方程成立.

因此, 存在 $a=b=1$ 使得 $(a+1)(b+1)=4$.

例题 5.19　设正实数 m,n 满足 $2m+n=2$. 求 $\frac{1}{m}+\frac{4}{n}$ 的最小值.

解　目标是一个分式和, 约束是一个线性方程, 这适合用柯西不等式.

将约束条件与目标函数相乘:

$$(2m+n)\left(\frac{1}{m}+\frac{4}{n}\right)$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

知乎@余命数

对于正数 x_1, x_2, y_1, y_2, 柯西不等式的一种形式为

$$(x_1 + x_2)(y_1 + y_2) \geq (\sqrt{x_1 y_1} + \sqrt{x_2 y_2})^2$$

令 $x_1 = 2m, x_2 = n$ 及 $y_1 = \frac{1}{m}, y_2 = \frac{4}{n}$.

$$(2m + n)\left(\frac{1}{m} + \frac{4}{n}\right) \geq \left(\sqrt{2m \cdot \frac{1}{m}} + \sqrt{n \cdot \frac{4}{n}}\right)^2$$

$$(2m + n)\left(\frac{1}{m} + \frac{4}{n}\right) \geq (\sqrt{2} + \sqrt{4})^2 = (\sqrt{2} + 2)^2 = 6 + 4\sqrt{2}$$

将 $2m + n = 2$ 代入, 得

$$2\left(\frac{1}{m} + \frac{4}{n}\right) \geq 6 + 4\sqrt{2}$$

知乎@余命数

$$\frac{1}{m} + \frac{4}{n} \geq 3 + 2\sqrt{2}$$

知乎@余命数

等号成立的条件是 $\frac{x_1}{y_1}$ 与 $\frac{x_2}{y_2}$ 的比值相等, 即 $\frac{2m}{1/m} = \frac{n}{4/n}$, 亦即 $2m^2 = \frac{n^2}{4}$, 故 $n = 2\sqrt{2}m$.

知乎@余命数

代入 $2m + n = 2$, 得

$$2m + 2\sqrt{2}m = 2 \implies m(2 + 2\sqrt{2}) = 2 \implies m = \frac{1}{1 + \sqrt{2}} = \sqrt{2} - 1$$

此时 $n = 2\sqrt{2}(\sqrt{2} - 1) = 4 - 2\sqrt{2}$.

由于 m, n 均为正数且使等号成立, 故 $\frac{1}{m} + \frac{4}{n}$ 的最小值为 $3 + 2\sqrt{2}$.

例题 5.20 已知 $a > 0, b > 0$ 且 $a + b = 5$. 求 $\sqrt{a+1} + \sqrt{b+3}$ 的最大值.

解 目标函数是根式之和, 约束是线性的. 我们用柯西不等式. 令向量 $\mathbf{u} = (\sqrt{a+1}, \sqrt{b+3})$ 和 $\mathbf{v} = (1, 1)$.

根据柯西不等式.

$$(\sqrt{a+1} \cdot 1 + \sqrt{b+3} \cdot 1)^2 \leq ((\sqrt{a+1})^2 + (\sqrt{b+3})^2)(1^2 + 1^2)$$

$$(\sqrt{a+1} + \sqrt{b+3})^2 \leq ((a+1) + (b+3))(2)$$

知乎@余命数

$$(\sqrt{a+1} + \sqrt{b+3})^2 \leq (a + b + 4)(2)$$

代入 $a+b=5$, 得

$$(\sqrt{a+1}+\sqrt{b+3})^2 \le (5+4)(2)=18$$

由于各项为正, 开平方根得 $\sqrt{a+1}+\sqrt{b+3} \le \sqrt{18}=3\sqrt{2}$.

等号成立的条件是向量 **u** 与 **v** 共线, 即 $\frac{\sqrt{a+1}}{1}=\frac{\sqrt{b+3}}{1}$. 这等价于 $a+1=b+3$, 即 $a=b+2$. 联立 $a=b+2$ 和 $a+b=5$, 解得 $b=3/2, a=7/2$.

由于存在使等号成立的正数 a,b, 故表达式的最大值为 $3\sqrt{2}$.

例题 5.21 求函数 $f(x)=\sqrt{x^2-3x+2}+\sqrt{2+3x-x^2}$ 的最大值.

解 函数是两个根式之和. 两个根号内的表达式相加后变量被消去.

$$(x^2-3x+2)+(2+3x-x^2)=4$$

这提示我们应用柯西不等式.

令向量 $\mathbf{u}=(\sqrt{x^2-3x+2},\sqrt{2+3x-x^2})$ 和 $\mathbf{v}=(1,1)$.

$$(f(x))^2=(\sqrt{x^2-3x+2}\cdot 1+\sqrt{2+3x-x^2}\cdot 1)^2$$

$$\le ((\sqrt{x^2-3x+2})^2+(\sqrt{2+3x-x^2})^2)(1^2+1^2)$$

$$\le ((x^2-3x+2)+(2+3x-x^2))(2)$$

$$\le (4)(2)=8$$

故 $f(x)\le\sqrt{8}=2\sqrt{2}$.

等号成立的条件是 $\frac{\sqrt{x^2-3x+2}}{1}=\frac{\sqrt{2+3x-x^2}}{1}$. 即 $x^2-3x+2=2+3x-x^2$, 化简得 $2x^2-6x=0$, 解得 $x=0$ 或 $x=3$.

检验这两个值是否在函数的定义域内. 定义域要求 $x^2-3x+2\ge 0$ 且 $2+3x-x^2\ge 0$. 当 $x=0$ 时, $2\ge 0$ 且 $2\ge 0$, 成立. 当 $x=3$ 时, $9-9+2=2\ge 0$ 且 $2+9-9=2\ge 0$, 成立.

由于存在使等号成立的 x 值, 函数的最大值为 $2\sqrt{2}$.

例题 5.22 已知正数 x,y 满足 $\frac{1}{2x+y}+\frac{3}{x+y}=2$. 求 $6x+5y$ 的最小值.

解 条件是一个分式和, 目标是一个线性和. 将目标表达式与条件表达式相乘, 可构造一个常数下界.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

我们先把目标 $6x+5y$ 用条件中的分母 $2x+y$ 与 $x+y$ 表示. 设 $6x+5y=A(2x+y)+B(x+y)=(2A+B)x+(A+B)y$. 比较系数得 $2A+B=6, A+B=5$. 解得 $A=1, B=4$. 故 $6x+5y=(2x+y)+4(x+y)$.

考虑乘积 $(6x+5y)\cdot 2$.

$$2(6x+5y)=\left(\frac{1}{2x+y}+\frac{3}{x+y}\right)((2x+y)+4(x+y))$$

令 $a=\frac{1}{2x+y}, b=\frac{3}{x+y}$ 和 $c=2x+y, d=4(x+y)$. 根据柯西不等式的一个推论, 或直接由 AM-GM 不等式, $(a+b)(c+d)\geq(\sqrt{ac}+\sqrt{bd})^2$.

$$\begin{aligned}\left(\frac{1}{2x+y}+\frac{3}{x+y}\right)((2x+y)+4(x+y)) &\geq \left(\sqrt{1}+\sqrt{3\cdot 4}\right)^2\\ &=(1+2\sqrt{3})^2\\ &=13+4\sqrt{3}.\end{aligned}$$

因此, $2(6x+5y)\geq 13+4\sqrt{3}$, 即 $6x+5y\geq\frac{13+4\sqrt{3}}{2}$.

等号成立的条件是 $\frac{a}{c}=\frac{b}{d}$, 即 $\frac{1/(2x+y)}{2x+y}=\frac{3/(x+y)}{4(x+y)}$. $\frac{1}{(2x+y)^2}=\frac{3}{4(x+y)^2}$, 即 $4(x+y)^2=3(2x+y)^2$. 由于各项为正, $2(x+y)=\sqrt{3}(2x+y)$, 即 $(2-\sqrt{3})y=(2\sqrt{3}-2)x$. 令 $r=\frac{y}{x}=\frac{2\sqrt{3}-2}{2-\sqrt{3}}>0$. 再由题设

$$\frac{1}{(2+r)x}+\frac{3}{(1+r)x}=2$$

可唯一确定一个正数 x, 从而得到正数 $y=rx$. 因此等号条件确实可以在题设范围内达到.

故 $6x+5y$ 的最小值为 $\frac{13+4\sqrt{3}}{2}$.

例题 5.23 已知正数 x, y, z 满足 $x+y+z=1$. 求 $\frac{x^2}{y+2z}+\frac{y^2}{z+2x}+\frac{z^2}{x+2y}$ 的最小值.

解 目标函数具有 Engel 形式的标准结构.

$$\frac{x^2}{y+2z}+\frac{y^2}{z+2x}+\frac{z^2}{x+2y}\geq\frac{(x+y+z)^2}{(y+2z)+(z+2x)+(x+2y)}$$

化简分母:

$$(y+2z)+(z+2x)+(x+2y)=3x+3y+3z=3(x+y+z)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入不等式：

$$\text{LHS} \geq \frac{(x+y+z)^2}{3(x+y+z)} = \frac{x+y+z}{3}$$

根据约束 $x+y+z=1$，得到

$$\frac{x^2}{y+2z} + \frac{y^2}{z+2x} + \frac{z^2}{x+2y} \geq \frac{1}{3}$$

等号成立的条件是 $\frac{x}{y+2z} = \frac{y}{z+2x} = \frac{z}{x+2y}$. 若 $x=y=z$，此条件显然满足. 结合 $x+y+z=1$，解得 $x=y=z=1/3$.

由于存在使等号成立的取值，故表达式的最小值为 $1/3$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

例题 5.24　已知正数 x, y 满足 $x+y=1$. 求 $\frac{1}{x^2} + \frac{8}{y^2}$ 的最小值.

解　我们用 Holder 不等式（柯西不等式的一种推广）解决这个问题. 对于三组正数 $(a_1,a_2),(b_1,b_2),(c_1,c_2)$，有

$$(a_1+a_2)(b_1+b_2)(c_1+c_2) \geq \left(\sqrt[3]{a_1b_1c_1} + \sqrt[3]{a_2b_2c_2}\right)^3$$

我们构造三组数. 令 $(a_1,a_2) = (\frac{1}{x^2}, \frac{8}{y^2})$、$(b_1,b_2) = (x,y)$、$(c_1,c_2) = (x,y)$.

代入 Holder 不等式.

$$\left(\frac{1}{x^2} + \frac{8}{y^2}\right)(x+y)(x+y) \geq \left(\sqrt[3]{\frac{1}{x^2}\cdot x\cdot x} + \sqrt[3]{\frac{8}{y^2}\cdot y\cdot y}\right)^3$$

利用 $x+y=1$ 简化左侧，并计算右侧，得

$$\left(\frac{1}{x^2} + \frac{8}{y^2}\right)\cdot 1\cdot 1 \geq \left(\sqrt[3]{1} + \sqrt[3]{8}\right)^3 = (1+2)^3 = 27$$

故 $\frac{1}{x^2} + \frac{8}{y^2} \geq 27$.

等号成立的条件是这三组数成比例，即 $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$.

$$\frac{1/x^2}{8/y^2} = \frac{x}{y}$$

$$\frac{y^2}{8x^2} = \frac{x}{y} \implies y^3 = 8x^3 \implies y = 2x$$

联立 $y=2x$ 和 $x+y=1$，解得 $x=1/3, y=2/3$.

由于存在使等号成立的正数 x, y，故表达式的最小值为 27.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

何时要用

柯西不等式联系二次型表达式, 如 $\sum a_i^2$ 与一次线性组合, 如 $\sum a_i b_i$. 只要给定其中一类的值而探求另一类的最值, 它常能直接给出结果, 例如在线性约束下求二次和的最小值, 或在二次约束下求线性组合的最值.

当目标函数为分式和, 特别是形如 $\sum \frac{a_i^2}{b_i}$ 时, 我们可考虑 Engel 形式. 分子不必是变量的平方, 也可以是常数的平方; 例如 $\frac{1}{b}$ 可写成 $\frac{1^2}{b}$, 而 $\frac{9}{c}$ 可写成 $\frac{3^2}{c}$. 若目标函数呈此结构且约束与分母之和相关, Engel 形式往往能直接导出结论.

我们还可以借助辅助向量来创造应用条件; 例如取 $(1,1,\ldots,1)$ 作为辅助向量, 可建立 $(\sum x_i)^2$ 与 $\sum x_i^2$ 之间的联系.

知乎@余命数 知乎@余命数 知乎@余命数

5.6 排序不等式

我们研究两个序列的元素如何配对, 才能使乘积之和最大或最小. 排序不等式回答这个问题.

定理 5.8 (排序不等式) 设有两组实数 $a_1 \le a_2 \le \ldots \le a_n$ 和 $b_1 \le b_2 \le \ldots \le b_n$. 设 $(p_1, p_2, \ldots, p_n)$ 是 $(1,2,\ldots,n)$ 的一个任意排列. 则有

$$\sum_{i=1}^{n} a_i b_{n-i+1} \le \sum_{i=1}^{n} a_i b_{p_i} \le \sum_{i=1}^{n} a_i b_i$$

因此, 两个有序序列的元素配对, 其乘积之和在**顺序和**时取得最大值, 在**乱序和**时介于最大与最小之间, 在**逆序和**时取得最小值.

证明 证明顺序和最大. 考虑任意非顺序的乱序和 $S = \sum a_i b_{p_i}$. 若该排列不是顺序排列 $(1,2,\ldots,n)$, 则存在两个下标 j,k 满足 $j<k$ 但 $p_j > p_k$.

S 中与 j,k 相关的项为 $a_j b_{p_j} + a_k b_{p_k}$.

交换 p_j, p_k 得到新排列, 其对应的和记为 S'. S' 中与 j,k 相关的项为 $a_j b_{p_k} + a_k b_{p_j}$.

考查两式之差.

$$\begin{aligned} S' - S &= (a_j b_{p_k} + a_k b_{p_j}) - (a_j b_{p_j} + a_k b_{p_k}) \\ &= a_k b_{p_j} - a_j b_{p_j} - a_k b_{p_k} + a_j b_{p_k} \\ &= b_{p_j}(a_k - a_j) - b_{p_k}(a_k - a_j) \\ &= (a_k - a_j)(b_{p_j} - b_{p_k}) \end{aligned}$$

由假设, $j<k \implies a_j \le a_k$, 故 $a_k - a_j \ge 0$. 又 $p_j > p_k \implies b_{p_j} \ge b_{p_k}$, 故 $b_{p_j} - b_{p_k} \ge 0$.

因此 $S' - S \ge 0$, 即 $S' \ge S$.

每当排列中出现违反顺序的两个配对, 将该配对修正为顺序后, 总和不会减小. 通过有限次这样的交换, 任何乱序排列都可调整为顺序排列, 且乘积之和单调不减. 故顺序和最大. 对序列 $b_n \le b_{n-1} \le \cdots \le b_1$ 应用同样结论, 可得逆序和最小. 所以任意乱序和介于逆序和与顺序和之间. □

知乎@余命数

例题 5.25　设 a,b,c 为正实数. 证明 $a^3+b^3+c^3 \ge a^2b+b^2c+c^2a$.

解 这个不等式为轮换式, 不能假设 a,b,c 的大小. 我们用 AM-GM 不等式证明它.

由 a,b,c 均为正实数, 分别有

$$a^3+a^3+b^3 \ge 3\sqrt[3]{a^3\cdot a^3\cdot b^3} = 3a^2b,$$

$$b^3+b^3+c^3 \ge 3b^2c, \qquad c^3+c^3+a^3 \ge 3c^2a.$$

三式相加, 得

$$3(a^3+b^3+c^3) \ge 3(a^2b+b^2c+c^2a).$$

两边同除以 3, 即得

$$a^3+b^3+c^3 \ge a^2b+b^2c+c^2a.$$

例题 5.26 (IMO) 已知 $x_i, y_i (i=1,2,\ldots,n)$ 为任意实数, 且 $x_1 \ge x_2 \ge \ldots \ge x_n$ 和 $y_1 \ge y_2 \ge \ldots \ge y_n$. 又 $z_1, z_2, \ldots, z_n$ 是 $y_1, y_2, \ldots, y_n$ 的任意一个排列. 试证:

$$\sum_{i=1}^{n}(x_i-y_i)^2 \le \sum_{i=1}^{n}(x_i-z_i)^2$$

解 展开平方项, 就可把问题化为排序不等式.

先展开右侧:

$$\begin{aligned}\sum_{i=1}^{n}(x_i-z_i)^2 &= \sum_{i=1}^{n}(x_i^2-2x_iz_i+z_i^2)\\ &= \sum_{i=1}^{n}x_i^2 - 2\sum_{i=1}^{n}x_iz_i + \sum_{i=1}^{n}z_i^2\end{aligned}$$

分析其中的项.

- $\sum x_i^2$ 由序列 (x_i) 唯一确定, 是常数.
- (z_i) 是 (y_i) 的排列, 所含元素相同, 仅顺序不同. 故 $\sum z_i^2 = \sum y_i^2$ 也是常数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 随排列 (z_i) 变化的是乱序和 $\sum x_i z_i$.

因此, 表达式 $\sum(x_i - z_i)^2$ 取最小值当且仅当 $\sum x_i z_i$ 取最大值.

由排序不等式, 两个同序排列的序列 (x_i) 与 (y_i) 的乘积之和在顺序和时最大, 即对 (y_i) 的任意排列 (z_i),

$$\sum_{i=1}^{n} x_i z_i \leq \sum_{i=1}^{n} x_i y_i$$

知乎@余命数

所以 $\sum x_i z_i$ 的最大值在 $z_i = y_i$ 时取到.

代回展开式, $\sum(x_i - z_i)^2$ 的最小值在 $z_i = y_i$ 时取到, 此时值为 $\sum(x_i - y_i)^2$.

故对任意排列 (z_i), 都有

$$\sum_{i=1}^{n}(x_i - y_i)^2 \leq \sum_{i=1}^{n}(x_i - z_i)^2$$

知乎@余命数

5.7 伯努利不等式

伯努利不等式给出 $(1+x)^n$ 的一个线性下界.

定理 5.9 (伯努利不等式, 整数形式) 对任意实数 $x > -1$ 和任意非负整数 n, 恒有

$$(1+x)^n \geq 1 + nx$$

等号成立的充分必要条件是 $n = 0$, $n = 1$ 或 $x = 0$.

证明 用数学归纳法对整数 $n \geq 0$ 证明.

奠基时, $n = 0$ 给出 $(1+x)^0 \geq 1 + 0 \cdot x$ 即 $1 \geq 1$, $n = 1$ 给出 $(1+x)^1 \geq 1 + 1 \cdot x$ 即 $1 + x \geq 1 + x$, 均成立.

设 $n = k\,(k \geq 1)$ 时命题成立, 即 $(1+x)^k \geq 1 + kx$, 推证 $n = k+1$ 时命题亦成立. 由题设 $x > -1$, 故 $1 + x > 0$, 以正数 $1 + x$ 同乘归纳假设两端, 不等号方向不变, 得

$$(1+x)^{k+1} = (1+x)^k(1+x) \geq (1+kx)(1+x).$$

展开右端得 $(1+kx)(1+x) = 1 + x + kx + kx^2 = 1 + (k+1)x + kx^2$, 而由 $k \geq 1$ 且 $x^2 \geq 0$ 有 $kx^2 \geq 0$, 故 $1 + (k+1)x + kx^2 \geq 1 + (k+1)x$. 由传递性即得

$$(1+x)^{k+1} \geq 1 + (k+1)x,$$

命题对 $n = k+1$ 成立.

由数学归纳法原理, 伯努利不等式对所有非负整数 n 成立.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等号当 $n=0$ 或 $n=1$ 时原式恒取, 当 $x=0$ 时两边也均为 1. 若 $n\geq 2$ 且 $x\neq 0$, 归纳递推给出严格不等式. 对 $n=2$ 有 $(1+x)^2-(1+2x)=x^2>0$; 若第 k 步严格, 由 $1+x>0$ 且递推中还加上 $kx^2>0$, 第 $k+1$ 步仍严格. 故等号成立当且仅当 $n=0$、$n=1$ 或 $x=0$. □

伯努利不等式可推广到实数指数. 下面的证明用到函数的凹凸性.

伯努利不等式实数形式

定理 5.10 (伯努利不等式, 实数形式) 设 $x>-1$ 且 $x\neq 0$.

- 若 $\alpha<0$ 或 $\alpha>1$, 则 $(1+x)^\alpha>1+\alpha x$.
- 若 $0<\alpha<1$, 则 $(1+x)^\alpha<1+\alpha x$.

知乎@余命数

证明 考虑辅助函数 $f(x)=(1+x)^\alpha-(1+\alpha x)$, 其定义域为 $(-1,+\infty)$. 分析导数确定其单调性与极值.

$$f'(x)=\alpha(1+x)^{\alpha-1}-\alpha=\alpha\left[(1+x)^{\alpha-1}-1\right]$$

令 $f'(x)=0$, 唯一解为 $x=0$.

当 $\alpha>1$ 时, $\alpha-1>0$. 若 $x>0$, 则 $1+x>1$、$(1+x)^{\alpha-1}>1$、$f'(x)>0$. 若 $-1<x<0$, 则 $0<1+x<1$、$(1+x)^{\alpha-1}<1$、$f'(x)<0$. 故 $f(x)$ 在 $x=0$ 处取得唯一的极小值, 也是最小值. $f_{\min}=f(0)=(1+0)^\alpha-(1+0)=0$. 因此, 对 $x\neq 0$, 恒有 $f(x)>0$, 即 $(1+x)^\alpha>1+\alpha x$.

当 $0<\alpha<1$ 时, $\alpha-1<0$. 若 $x>0$, 则 $1+x>1$、$(1+x)^{\alpha-1}<1$、$f'(x)<0$. 若 $-1<x<0$, 则 $0<1+x<1$、$(1+x)^{\alpha-1}>1$、$f'(x)>0$. 故 $f(x)$ 在 $x=0$ 处取得唯一的极大值, 也是最大值. $f_{\max}=f(0)=0$. 因此, 对 $x\neq 0$, 恒有 $f(x)<0$, 即 $(1+x)^\alpha<1+\alpha x$.

当 $\alpha<0$ 时, $\alpha-1<0$. 若 $x>0$, 则 $(1+x)^{\alpha-1}<1$, 而 $\alpha<0$, 故 $f'(x)>0$; 若 $-1<x<0$, 则 $(1+x)^{\alpha-1}>1$, 故 $f'(x)<0$. 因此 f 仍在 $x=0$ 处取得唯一最小值 0, 对 $x\neq 0$ 有 $f(x)>0$. □

例题 5.27 证明 $1.001^{1000}>2$.

知乎@余命数

解 该不等式具有 $(1+x)^n$ 的形式, 我们用伯努利不等式证明它.

令 $x=0.001$ 及 $n=1000$. 条件 $x=0.001>-1$ 和 n 为非负整数均满足. 由伯努利不等式,

$$(1+0.001)^{1000}\geq 1+1000\times 0.001=1+1=2$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等号仅在 $n=0,1$ 或 $x=0$ 时成立. 此例中 $n=1000$ 且 $x=0.001$, 均不满足等号条件. 因此不等号为严格不等号.

故 $1.001^{1000}>2$.

例题 5.28 证明序列 $a_n=\left(1+\frac{1}{n}\right)^n$ $(n\in\mathbb{N},n\geq 1)$ 严格单调递增.

解 要证对任意 $n\geq 1$, 恒有 $a_{n+1}>a_n$. 这等价于证明 $\frac{a_{n+1}}{a_n}>1$.

构造商式并化简, 得

$$\begin{aligned}\frac{a_{n+1}}{a_n}&=\frac{\left(1+\frac{1}{n+1}\right)^{n+1}}{\left(1+\frac{1}{n}\right)^n}=\frac{\left(\frac{n+2}{n+1}\right)^{n+1}}{\left(\frac{n+1}{n}\right)^n}\\&=\left(\frac{n+2}{n+1}\right)^{n+1}\cdot\left(\frac{n}{n+1}\right)^n\\&=\frac{n+2}{n+1}\cdot\left(\frac{n+2}{n+1}\cdot\frac{n}{n+1}\right)^n\\&=\frac{n+2}{n+1}\cdot\left(\frac{n^2+2n}{(n+1)^2}\right)^n\\&=\frac{n+2}{n+1}\cdot\left(\frac{(n+1)^2-1}{(n+1)^2}\right)^n\\&=\frac{n+2}{n+1}\cdot\left(1-\frac{1}{(n+1)^2}\right)^n\end{aligned}$$

对因子 $\left(1-\frac{1}{(n+1)^2}\right)^n$ 应用伯努利不等式. 令 $x=-\frac{1}{(n+1)^2}$. 由 $n\geq 1, x>-1$ 且 $x\neq 0$. 由伯努利不等式, $(1+x)^n\geq 1+nx$,

$$\left(1-\frac{1}{(n+1)^2}\right)^n\geq 1-\frac{n}{(n+1)^2}$$

代回商式并化简, 得

$$\begin{aligned}\frac{a_{n+1}}{a_n}&\geq\frac{n+2}{n+1}\cdot\left(1-\frac{n}{(n+1)^2}\right)\\&=\frac{n+2}{n+1}\cdot\frac{(n+1)^2-n}{(n+1)^2}\\&=\frac{n+2}{n+1}\cdot\frac{n^2+2n+1-n}{(n+1)^2}\\&=\frac{(n+2)(n^2+n+1)}{(n+1)^3}\end{aligned}$$

比较分子与分母.

$$(n+2)(n^2+n+1)=n^3+3n^2+3n+2$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数　$(n+1)^3 = n^3 + 3n^2 + 3n + 1$

分子比分母大 1. 故 $\frac{(n+2)(n^2+n+1)}{(n+1)^3} > 1$.

因此 $\frac{a_{n+1}}{a_n} > 1$, 即 $a_{n+1} > a_n$. 序列 a_n 严格单调递增.

5.8 糖水不等式

我们比较分式大小时, 常会遇到一类背景类似溶液浓度变化的不等式.

定理 5.11 (糖水不等式, 基本形式) 设 a, b, x 均为正实数.

- 若 $\frac{a}{b} < 1$, 则 $\frac{a}{b} < \frac{a+x}{b+x}$.
- 若 $\frac{a}{b} > 1$, 则 $\frac{a}{b} > \frac{a+x}{b+x}$.
- 若 $\frac{a}{b} = 1$, 则 $\frac{a}{b} = \frac{a+x}{b+x} = 1$.

对真分数, 分子分母同加一个正数会变大; 对假分数, 同加一个正数会变小.

证明 我们比较 $\frac{a+x}{b+x}$ 与 $\frac{a}{b}$ 的大小. 由于 a, b, x 均为正数, b 和 $b+x$ 亦为正, 于是作差, 得

知乎@余命数

$$\frac{a+x}{b+x} - \frac{a}{b} = \frac{b(a+x) - a(b+x)}{b(b+x)} = \frac{ab + bx - ab - ax}{b(b+x)} = \frac{x(b-a)}{b(b+x)}$$

知乎@余命数

分母 $b(b+x)$ 与 x 均为正, 差的符号完全由 $(b-a)$ 决定.

- 若 $\frac{a}{b} < 1$, 则 $a < b$, 即 $b - a > 0$. 差为正, 故 $\frac{a+x}{b+x} > \frac{a}{b}$.
- 若 $\frac{a}{b} > 1$, 则 $a > b$, 即 $b - a < 0$. 差为负, 故 $\frac{a+x}{b+x} < \frac{a}{b}$.
- 若 $\frac{a}{b} = 1$, 则 $a = b$, 即 $b - a = 0$. 差为零, 故 $\frac{a+x}{b+x} = \frac{a}{b}$.

□

例题 5.29 比较 $\frac{2023}{2024}$ 与 $\frac{2024}{2025}$ 的大小.

解 两个分数结构相似, 第二个分数可由第一个变化得到.

令 $a = 2023, b = 2024$, 则 $\frac{2024}{2025} = \frac{2023+1}{2024+1} = \frac{a+x}{b+x}$ $(x = 1)$.

$\frac{a}{b} = \frac{2023}{2024} < 1$ 是真分数, 由糖水不等式基本形式可得

$$\frac{2023}{2024} < \frac{2023+1}{2024+1} = \frac{2024}{2025}$$

我们可以把基本形式推广到分子、分母分别加上不同正数的情形.

定理 5.12 (糖水不等式, 一般形式) 设 a,b,x,y 均为正实数. 分数 $\frac{a+x}{b+y}$ 的值介于 $\frac{a}{b}$ 和 $\frac{x}{y}$ 之间.
若 $\frac{a}{b}<\frac{x}{y}$, 则 $\frac{a}{b}<\frac{a+x}{b+y}<\frac{x}{y}$. 若 $\frac{a}{b}>\frac{x}{y}$, 则 $\frac{a}{b}>\frac{a+x}{b+y}>\frac{x}{y}$.

证明 我们可设 $\frac{a}{b}<\frac{x}{y}$. 由于 b,y 均为正数, 该式等价于 $ay<bx$.

先证 $\frac{a}{b}<\frac{a+x}{b+y}$, 即

$$\frac{a}{b}<\frac{a+x}{b+y}\iff a(b+y)<b(a+x)\iff ay<bx,$$

这与假设一致, 故成立.

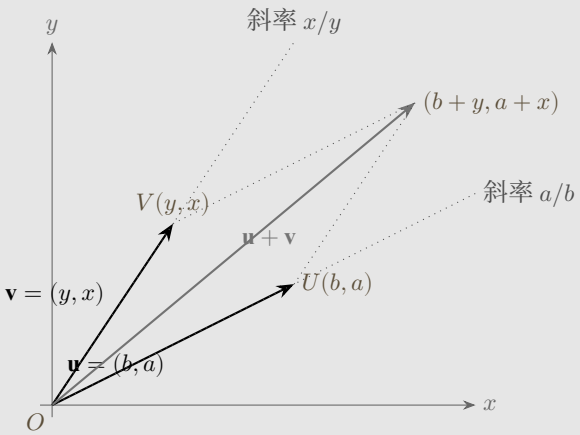
再证 $\frac{a+x}{b+y}<\frac{x}{y}$, 即

$$\frac{a+x}{b+y}<\frac{x}{y}\iff y(a+x)<x(b+y)\iff ay<bx,$$

同样成立.

若 $\frac{a}{b}>\frac{x}{y}$, 对称讨论即可. 若 $\frac{a}{b}=\frac{x}{y}$, 则三个分数值全部相等. □

注 5.5 (几何解释)
我们给一般形式的糖水不等式一个几何解释. 考虑平面直角坐标系中的两个向量 $\mathbf{u}=(b,a)$ 和 $\mathbf{v}=(y,x)$. 由 a,b,x,y 均为正数, 这两个向量位于第一象限.
向量 $\mathbf{u}$ 的斜率为 $\frac{a}{b}$, 向量 $\mathbf{v}$ 的斜率为 $\frac{x}{y}$. 两个向量的和为 $\mathbf{u}+\mathbf{v}=(b+y,a+x)$, 其斜率为 $\frac{a+x}{b+y}$.
从图中可以看出, 两个第一象限向量之和的斜率介于这两个向量各自的斜率之间.



知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 5.30　设有两个正数序列 $\{a_n\}$ 和 $\{b_n\}$. 如果序列 $c_n = \frac{a_n}{b_n}$ 严格单调递增, 证明由它们的部分和构成的序列 $S_n = \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i}$ 也严格单调递增.

解　我们需证明对任意 $n \geq 1$ 都有 $S_n < S_{n+1}$.

记 $A_n = \sum_{i=1}^{n} a_i$, $B_n = \sum_{i=1}^{n} b_i$, 则 $S_n = \frac{A_n}{B_n}$. 由 $A_{n+1} = A_n + a_{n+1}$ 且 $B_{n+1} = B_n + b_{n+1}$,

$$S_{n+1} = \frac{A_n + a_{n+1}}{B_n + b_{n+1}}.$$

由糖水不等式的一般形式, S_{n+1} 介于 $\frac{A_n}{B_n}$ 和 $\frac{a_{n+1}}{b_{n+1}}$ 之间. 因此, 只需证明 $S_n < \frac{a_{n+1}}{b_{n+1}}$, 便可得到 $S_n < S_{n+1}$.

利用 $a_i = c_i b_i$, 得

$$S_n = \frac{c_1 b_1 + c_2 b_2 + \cdots + c_n b_n}{b_1 + b_2 + \cdots + b_n}.$$

它是 $c_1, c_2, \ldots, c_n$ 关于正权重 b_i 的加权算术平均.

由 c_n 严格单调递增, $c_1 < c_2 < \cdots < c_n < c_{n+1}$. 加权平均值不超过所有元素中的最大值, 因此

$$S_n = \frac{\sum_{i=1}^{n} c_i b_i}{\sum_{i=1}^{n} b_i} \leq c_n.$$

又 $c_n < c_{n+1} = \frac{a_{n+1}}{b_{n+1}}$, 所以 $S_n < \frac{a_{n+1}}{b_{n+1}}$.

因此 $S_n < S_{n+1}$, 序列 S_n 严格单调递增.

5.8.1 函数单调性与次可加性

糖水不等式说明函数 $f(t) = \frac{t}{c+t}$ $(c > 0, t \geq 0)$ 单调递增. 结合三角不等式, 我们可用它来证明一类不等式.

例题 5.31　设 $x_1, x_2, \ldots, x_n$ 为实数, $c > 0$. 证明:

$$\frac{|\sum_{i=1}^{n} x_i|}{c + |\sum_{i=1}^{n} x_i|} \leq \sum_{i=1}^{n} \frac{|x_i|}{c + |x_i|}$$

解　令 $f(t) = \frac{t}{c+t}$ $(t \geq 0)$, 则原不等式可写作 $f(|\sum x_i|) \leq \sum f(|x_i|)$.

我们将 $f(t)$ 变形为 $1 - \frac{c}{c+t}$. 当 $t \geq 0$ 时, t 增大会导致 $c + t$ 增大, 从而使 $\frac{c}{c+t}$ 减小, 因此 $f(t)$ 单调递增.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

由三角不等式, $|\sum_{i=1}^n x_i| \le \sum_{i=1}^n |x_i|$. 由 $f(t)$ 单调递增,

$$f\left(\left|\sum_{i=1}^n x_i\right|\right) \le f\left(\sum_{i=1}^n |x_i|\right)$$

即

$$\frac{|\sum_{i=1}^n x_i|}{c+|\sum_{i=1}^n x_i|} \le \frac{\sum_{i=1}^n |x_i|}{c+\sum_{i=1}^n |x_i|}.$$

我们再验证 $f(t)$ 的次可加性. 对任意 $t_1, t_2 \ge 0$:

$$f(t_1+t_2) = \frac{t_1+t_2}{c+t_1+t_2} = \frac{t_1}{c+t_1+t_2} + \frac{t_2}{c+t_1+t_2}.$$

由 $t_1, t_2 \ge 0$, 有 $c+t_1+t_2 \ge c+t_1$ 且 $c+t_1+t_2 \ge c+t_2$. 于是

$$\frac{t_1}{c+t_1+t_2} \le \frac{t_1}{c+t_1} = f(t_1), \qquad \frac{t_2}{c+t_1+t_2} \le \frac{t_2}{c+t_2} = f(t_2)$$

相加得到 $f(t_1+t_2) \le f(t_1)+f(t_2)$. 由数学归纳法, 对任意 n 个非负变量,

$$f\left(\sum_{i=1}^n t_i\right) \le \sum_{i=1}^n f(t_i).$$

取 $t_i = |x_i|$, 即得

$$f\left(\sum_{i=1}^n |x_i|\right) \le \sum_{i=1}^n f(|x_i|) = \sum_{i=1}^n \frac{|x_i|}{c+|x_i|}.$$

故

$$\frac{|\sum_{i=1}^n x_i|}{c+|\sum_{i=1}^n x_i|} \le \frac{\sum_{i=1}^n |x_i|}{c+\sum_{i=1}^n |x_i|} \le \sum_{i=1}^n \frac{|x_i|}{c+|x_i|},$$

原不等式得证.

注 5.6

当 $c=1, n=2$ 时, 对任意实数 a, b,

$$\frac{|a+b|}{1+|a+b|} \le \frac{|a|}{1+|a|} + \frac{|b|}{1+|b|}.$$

这个不等式在度量空间理论中有用. 若 $d(x,y)=|x-y|$ 是一个度量, 则 $d'(x,y)=\frac{|x-y|}{1+|x-y|}$ 也是一个度量, 并且是有界度量.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

5.8.2 利用变量代换构造齐次形式

当约束条件为乘积形式时，我们可以通过变量代换把它齐次化.

例题 5.32　设 $a_1, a_2, \dots, a_n$ 为 n 个正实数 $(n \geq 3)$，且满足 $\prod_{i=1}^{n} a_i = 1$. 证明：

$$1 < \sum_{i=1}^{n} \frac{1}{1+a_i} < n-1$$

解　约束条件 $\prod a_i = 1$ 是非齐次的. 我们做一次变量代换，使约束自动满足，并把原不等式齐次化.

设存在正实数 $x_1, x_2, \dots, x_n$ 使得

$$a_1 = \frac{x_2}{x_1}, \quad a_2 = \frac{x_3}{x_2}, \quad \dots, \quad a_{n-1} = \frac{x_n}{x_{n-1}}, \quad a_n = \frac{x_1}{x_n}.$$

对任意满足乘积为 1 的正数 a_i，我们总可找到对应 x_i，因此这一代换是合理的.

代入后乘积条件自动满足：

$$\prod_{i=1}^{n} a_i = \frac{x_2}{x_1} \cdot \frac{x_3}{x_2} \cdots \frac{x_1}{x_n} = 1.$$

求和式中每一项变为

$$\frac{1}{1+a_i} = \frac{1}{1+x_{i+1}/x_i} = \frac{x_i}{x_i + x_{i+1}} \quad (x_{n+1} = x_1).$$

因此原不等式等价于：对任意正实数 $x_1, \dots, x_n$，

$$1 < \sum_{i=1}^{n} \frac{x_i}{x_i + x_{i+1}} < n-1.$$

先证左侧. 由 $x_j > 0$ 对所有 j 成立，$x_i + x_{i+1} < \sum_{j=1}^{n} x_j$，从而

$$\frac{x_i}{x_i + x_{i+1}} > \frac{x_i}{\sum_{j=1}^{n} x_j}.$$

对 $i = 1, \dots, n$ 求和：

$$\sum_{i=1}^{n} \frac{x_i}{x_i + x_{i+1}} > \sum_{i=1}^{n} \frac{x_i}{\sum_{j=1}^{n} x_j} = \frac{\sum_{i=1}^{n} x_i}{\sum_{j=1}^{n} x_j} = 1.$$

再证右侧. 将每一项变形为

$$\frac{x_i}{x_i + x_{i+1}} = 1 - \frac{x_{i+1}}{x_i + x_{i+1}},$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

对 $i=1,\dots,n$ 求和得

$$\sum_{i=1}^{n}\frac{x_i}{x_i+x_{i+1}}=n-\sum_{i=1}^{n}\frac{x_{i+1}}{x_i+x_{i+1}}.$$

要证右侧不等式, 只需证明 $\sum_{i=1}^{n}\frac{x_{i+1}}{x_i+x_{i+1}}>1$. 仿照左侧的放缩方法, 由 $x_i+x_{i+1}<\sum_{j=1}^{n}x_j$ 得

$$\frac{x_{i+1}}{x_i+x_{i+1}}>\frac{x_{i+1}}{\sum_{j=1}^{n}x_j}.$$

求和得

$$\sum_{i=1}^{n}\frac{x_{i+1}}{x_i+x_{i+1}}>\sum_{i=1}^{n}\frac{x_{i+1}}{\sum_{j=1}^{n}x_j}=1$$

(其中 $\sum x_{i+1}$ 只是 $x_1,\dots,x_n$ 的循环移位, 仍为 $\sum x_j$).

于是 $\sum_{i=1}^{n}\frac{x_i}{x_i+x_{i+1}}=n-$(大于 1 的数), 其值小于 $n-1$.

结合以上两方面, 原不等式得证.

5.8.3 在组合计数中的应用

糖水不等式也可用于比较组合数的大小, 从而证明某些代数不等式.

例题 5.33 已知 i,m,n 是正整数, 且 $1<i\le m<n$. 证明:

1. $n^iA_m^i<m^iA_n^i$
2. $(1+m)^n>(1+n)^m$

其中 $A_n^k=n(n-1)\dots(n-k+1)$ 表示排列数.

解 (1) 不等式 $n^iA_m^i<m^iA_n^i$ 等价于 $\frac{A_m^i}{m^i}<\frac{A_n^i}{n^i}$.

化简两边:

$$\frac{A_m^i}{m^i}=\frac{m(m-1)\dots(m-i+1)}{m\cdot m\cdots m}=\prod_{k=0}^{i-1}\frac{m-k}{m}$$

$$\frac{A_n^i}{n^i}=\frac{n(n-1)\dots(n-i+1)}{n\cdot n\cdots n}=\prod_{k=0}^{i-1}\frac{n-k}{n}$$

比较对应项 $\frac{m-k}{m}$ 与 $\frac{n-k}{n}$, $k=0,1,\dots,i-1$. 当 $k=0$ 时, 两项相等, 均为 1.

当 $1\le k\le i-1$ 时, 由于 $m<n$, $\frac{m-k}{n-k}$ 是真分数. 由糖水不等式,

$$\frac{m-k}{n-k}<\frac{(m-k)+k}{(n-k)+k}=\frac{m}{n}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因为 $m,n,n-k$ 均为正, 上式可化为 $n(m-k)<m(n-k)$, 即

$$\frac{m-k}{m}<\frac{n-k}{n}.$$

由于 $i>1$, 至少存在一个 $k\geq 1$ 使不等号取严格小于. 因此

$$\prod_{k=0}^{i-1}\frac{m-k}{m}<\prod_{k=0}^{i-1}\frac{n-k}{n}$$

即 $\frac{A_m^i}{m^i}<\frac{A_n^i}{n^i}$, 故 $n^iA_m^i<m^iA_n^i$ 成立.

(2) 用二项式定理展开两边:

$$(1+m)^n=\sum_{k=0}^{n}\binom{n}{k}m^k,\qquad (1+n)^m=\sum_{k=0}^{m}\binom{m}{k}n^k.$$

对 $k=0,1,\dots,m$, 比较 $\binom{n}{k}m^k$ 与 $\binom{m}{k}n^k$. 由 $\binom{n}{k}=\frac{A_n^k}{k!}$ 和 $\binom{m}{k}=\frac{A_m^k}{k!}$, 比较两者等价于比较 $m^kA_n^k$ 与 $n^kA_m^k$.

根据 (1) 的结论, 对 $1<k\leq m<n$, 有 $n^kA_m^k<m^kA_n^k$, 因此 $\binom{m}{k}n^k<\binom{n}{k}m^k$.

考察两个和式:

$$(1+m)^n=\sum_{k=0}^{m}\binom{n}{k}m^k+\sum_{k=m+1}^{n}\binom{n}{k}m^k$$

$$(1+n)^m=\sum_{k=0}^{m}\binom{m}{k}n^k$$

由于 $n>m$, $(1+m)^n$ 中从 $k=m+1$ 到 n 的项均为正, 故 $\sum_{k=m+1}^{n}\binom{n}{k}m^k>0$.

比较前 $m+1$ 项.

$$\sum_{k=0}^{m}\binom{n}{k}m^k>\sum_{k=0}^{m}\binom{m}{k}n^k$$

因为 $k=0,1$ 时两边相等, 而 $k=2,\dots,m$ 时左边严格大于右边.

综合以上两点,

$$(1+m)^n=\sum_{k=0}^{m}\binom{n}{k}m^k+(\text{正数})>\sum_{k=0}^{m}\binom{m}{k}n^k=(1+n)^m$$

不等式得证.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.9 放缩思想

前面各节介绍的均值、柯西、排序等不等式都是成形工具, 结构匹配时直接代入. 大量不等式问题需要自行构造中间量. 与解方程相比, 不等式推理多出一层方向性, 即它依赖实数序的传递性,

$$A \le B \quad \text{且} \quad B \le C \implies A \le C$$

要证明 $A \le C$ 而直接比较有困难时, 可寻找一个中间量 B, 分别证明 $A \le B$ 与 $B \le C$. 这个中间量就是**放缩**.

一次成功的放缩需要同时满足三个条件.

- **方向保持**: 每一步放大或缩小都必须维持不等号的指向, 并留意等号条件是否被改变.
- **目标导向**: 放缩服务于得到可处理的对象, 可以求和、可以约分、可以套用已知不等式.
- **精度适度**: 放得太多, 结论不成立; 放得太少, 计算难以继续. 在保持方向的前提下剥离次要因素, 即为精度适度.

本节介绍三种最常用的放缩技术.

5.9.1 舍项与添项

在正数范围内, 下列事实几乎不需要证明, 却十分实用.

- 分式的分母**舍去**一个正项, 分母变小, 分式变大; **添上**一个正项, 分式变小.
- 代数和中**舍去**一个非负项, 总和不变或变小; **添上**一个非负项, 总和不变或变大.

舍项放缩的作用在于统一, 即把几个略有差异的表达式化为相同结构, 从而可合并处理.

例题 5.34 设 a, b, c, d 为正实数. 证明, 得

$$1 < \frac{a}{a+b+d} + \frac{b}{a+b+c} + \frac{c}{b+c+d} + \frac{d}{a+c+d} < 2$$

解 四个分式的分母都接近 $a+b+c+d$, 但各自缺少一个不同变量, 直接相加较困难. 这时应把分母统一, 以便相加.

为证下界, 给每个分母添上所缺变量, 分母变大, 分式变小.

$$\frac{a}{a+b+d} > \frac{a}{a+b+c+d}, \qquad \frac{b}{a+b+c} > \frac{b}{a+b+c+d}, \qquad \cdots$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

四式相加, 分母统一为 $a+b+c+d$, 分子之和恰为 $a+b+c+d$.

$$\text{原式} > \frac{a+b+c+d}{a+b+c+d} = 1$$

为证上界, 需放大每个分式. 每个分式都是真分数, 由糖水不等式 (第5.8节), 分子分母同加一个正数, 分式的值增大. 给每个分式添上所缺变量, 得

$$\frac{a}{a+b+d} < \frac{a+c}{a+b+c+d}, \qquad \frac{b}{a+b+c} < \frac{b+d}{a+b+c+d},$$

$$\frac{c}{b+c+d} < \frac{c+a}{a+b+c+d}, \qquad \frac{d}{a+c+d} < \frac{d+b}{a+b+c+d}.$$

四式相加, 分母同为 $a+b+c+d$, 分子之和为 $2(a+b+c+d)$.

$$\text{原式} < \frac{2(a+b+c+d)}{a+b+c+d} = 2$$

故原式严格介于 1 与 2 之间.

注 5.7
同一个表达式, 向下放缩得到下界, 向上放缩得到上界; 两个方向都指向统一分母 $a+b+c+d$. 放缩方向完全由目标决定. 要证下界, 就把每一项缩小; 要证上界, 就把每一项放大.

5.9.2 裂项放缩

我们用放缩处理数列求和不等式. 对于难以精确求和的级数 $\sum_{k=1}^{n} a_k$, 一个基本策略是把通项 a_k 放缩成某个**可裂项相消**的差分形式 $b_{k-1}-b_k$, 求和时中间项抵消, 只留首尾.

例题 5.35　证明: 对任意正整数 n,

$$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2} < 2$$

解　通项 $\frac{1}{k^2}$ 需要改写成相邻两项之差. 当 $k \geq 2$ 时, 把分母中的 k^2 换成更小的 $k(k-1)$, 分式变大且可裂项.

$$\frac{1}{k^2} < \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$$

保留首项, 从第二项起应用此放缩.

$$\sum_{k=1}^{n} \frac{1}{k^2} = 1 + \sum_{k=2}^{n} \frac{1}{k^2} < 1 + \sum_{k=2}^{n} \left(\frac{1}{k-1} - \frac{1}{k} \right)$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

$$= 1 + \left(1 - \frac{1}{n}\right) = 2 - \frac{1}{n} < 2$$

知乎@余命数

证毕.

上述证明中, 最大的损失发生在第二项, $\frac{1}{2^2} = \frac{1}{4}$ 被放大成 $\frac{1}{2\cdot 1} = \frac{1}{2}$, 一处就放掉了 $\frac{1}{4}$ 的余量. 若要更精确的估计, 我们改用分母更接近 k^2 的裂项方式.

$$\frac{1}{k^2} < \frac{1}{k^2 - \frac{1}{4}} = \frac{1}{k - \frac{1}{2}} - \frac{1}{k + \frac{1}{2}}$$

保留前两项的精确值, 从 $k = 3$ 起放缩, 得

$$\sum_{k=1}^{n} \frac{1}{k^2} < 1 + \frac{1}{4} + \left(\frac{2}{5} - \frac{1}{n + \frac{1}{2}}\right) < \frac{33}{20} = 1.65$$

知乎@余命数

这与该级数真实极限值 $\frac{\pi^2}{6} \approx 1.6449$ 已相当接近. 同一个问题, 放缩得越贴身, 得到的界越精确; 保留越多精确首项, 损失越小.

注 5.8 (放缩的度)
放缩的力度需要与目标匹配. 若在上例中把 $\frac{1}{k^2}$ 放大为 $\frac{1}{k}$, 求和将得到发散的调和级数, 任何常数上界都不存在, 因此放缩过度会使结论失效. 一次合格的放缩必须同时满足两个条件. 放松后的对象**可处理**(能求和、能约简), 且放松的**总量落在目标允许的余量内**. 动手前应先估计各项数量级, 起主要作用的大项必须保留, 只剥离影响甚微的次要因素.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

5.9.3 基本函数不等式放缩

三角函数、指数函数与对数函数同多项式的直接代数比较往往比较困难. 处理含这些函数的不等式, 我们先用一组多项式把它们夹住, 再在多项式的范围内求解. 这些基本放缩来自函数图象与其切线的位置关系, 也是 第5.10节 切线法的素材.

定理 5.13 (正弦与正切的基本序) 对任意 $0 < x < \frac{\pi}{2}$, 恒有

$$\sin x < x < \tan x$$

证明 (面积比较)

如图, 在单位圆中取圆心角 $\angle AOP = x$ $(0 < x < \frac{\pi}{2})$, 其中 A 为单位圆与 x 轴正半轴的交点, P 为圆上对应点. 过 P 作 OA 的垂线, 垂足为 Q; 过 A 作圆的切线(垂直于半径 OA), 交射线 OP 的延长线于 T.

由三角函数定义, $PQ = \sin x$, $AT = \tan x$. 考察三个区域的面积.

- 三角形 OAP 的面积为 $\frac{1}{2} \cdot OA \cdot PQ = \frac{1}{2}\sin x$;
- 扇形 OAP 的面积为 $\frac{1}{2} \cdot 1^2 \cdot x = \frac{x}{2}$;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

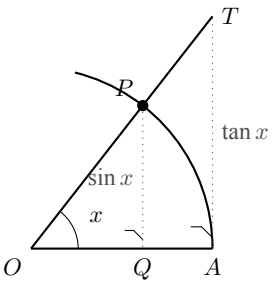


图 5.3　单位圆中的面积比较.$\triangle OAP$、扇形 OAP 与 $\triangle OAT$

- 三角形 OAT 的面积为 $\frac{1}{2}\cdot OA\cdot AT=\frac{1}{2}\tan x$.

由图形的包含关系, $\triangle OAP$ 严格含于扇形 OAP, 扇形 OAP 又严格含于 $\triangle OAT$, 故

$$\frac{1}{2}\sin x<\frac{x}{2}<\frac{1}{2}\tan x$$

两边同乘以 2, 即得 $\sin x<x<\tan x$. □

注 5.9
对 $\sin x<x<\tan x$ 各项取倒数并整理, 可得 $\cos x<\frac{\sin x}{x}<1$. 当 $x\to 0$ 时两端都趋于 1, 由夹逼原理即得重要极限 $\lim_{x\to 0}\frac{\sin x}{x}=1$. 这是微分学的基本极限, 严格极限语言见 第16章.

知乎@余命数 知乎@余命数 知乎@余命数

定理 5.14 (指数函数的基本放缩) 对任意实数 x, 恒有

$$e^x\geq 1+x$$

等号成立当且仅当 $x=0$.

指数基本放缩的导数证明

证明 (导数法) 我们构造辅助函数 $f(x)=e^x-1-x$, 只需证明 $f(x)\geq 0$. 求导得 $f'(x)=e^x-1$. 当 $x<0$ 时 $f'(x)<0$, f 单调递减; 当 $x>0$ 时 $f'(x)>0$, f 单调递增. 故 f 在 $x=0$ 处取得最小值 $f(0)=0$. 于是对任意实数 x 都有 $f(x)\geq 0$, 即 $e^x\geq 1+x$, 等号当且仅当 $x=0$. □

注 5.10 (切线放缩的几何意义)
不等式 $e^x\geq 1+x$ 的几何含义是, 曲线 $y=e^x$ 恒位于其在 $x=0$ 处的切线 $y=1+x$ 的上方 (切点除外). 这正是凸函数的特征性质, 见 第3.6节. 这种用切线夹住曲线的想法将在 第5.10节 中发展为系统性的切线法. 本节证明中使用的导数工具, 其理论基础见 第16章.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

推论 5.1 (对数函数的基本放缩) 对任意 $x > -1$, 恒有

$$\ln(1+x) \le x$$

等号成立当且仅当 $x = 0$. 等价地, 对任意 $x > 0$, 有 $\ln x \le x - 1$.

证明 设 $x > -1$, 则 $1 + x > 0$. 对不等式 $e^x \ge 1 + x$ 两边取自然对数, 由对数函数单调递增, 得 $x \ge \ln(1+x)$. 等号条件继承自指数放缩, 当且仅当 $x = 0$. 将 x 替换为 $x - 1$, 即得等价形式 $\ln x \le x - 1$ $(x > 0)$. □

例题 5.36 证明: 对任意正整数 n,

$$\left(1+\frac{1}{2}\right)\left(1+\frac{1}{4}\right)\left(1+\frac{1}{8}\right)\cdots\left(1+\frac{1}{2^n}\right) < e$$

解 左边是连乘积, 右边是 e. 由指数放缩 $e^x \ge 1 + x$, 1 加一个小量与 e 的幂联系起来, 而幂的乘积又可化为指数之和. 我们对每个因子应用严格不等式 $1 + x < e^x$ $(x > 0)$.

$$1+\frac{1}{2^k} < e^{1/2^k} \quad (k = 1, 2, \ldots, n)$$

n 个正项不等式相乘, 得

$$\prod_{k=1}^{n}\left(1+\frac{1}{2^k}\right) < \prod_{k=1}^{n} e^{1/2^k} = e^{\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^n}} = e^{1-\frac{1}{2^n}} < e$$

最后一步因为指数 $1 - \frac{1}{2^n} < 1$, 且指数函数单调递增.

例题 5.37 证明: 对任意正整数 n,

$$1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n} > \ln(n+1)$$

解 左边是调和级数的部分和, 右边是对数. 这提示我们使用对数放缩 $\ln(1+x) < x$ $(x > 0)$ 把每一项 $\frac{1}{k}$ 与对数联系起来. 取 $x = \frac{1}{k}$, 得

$$\frac{1}{k} > \ln\left(1+\frac{1}{k}\right) = \ln\frac{k+1}{k} = \ln(k+1) - \ln k$$

右边恰为相邻对数之差, 又是一个裂项结构. 对 $k = 1$ 到 n 求和, 中间项抵消:

$$\sum_{k=1}^{n}\frac{1}{k} > \sum_{k=1}^{n}(\ln(k+1) - \ln k) = \ln(n+1) - \ln 1 = \ln(n+1)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证毕.

5.10 局部化与孤立思想

注 5.11
切线法把函数的局部线性近似写成 $f(x_0)+f'(x_0)(x-x_0)$, 其严格理论见 第16章. 本节只用初等方法, 即作差定号、配方与均值不等式.

我们面对多元对称和式

$$f(a_1)+f(a_2)+\cdots+f(a_n)\geq C$$

直接处理整体时经常会遇到困难, 因为变量被约束条件互相纠缠. **局部化**思想把整体命题拆成每个变量单独成立的**局部命题**

$$f(x)\geq \ell(x)\quad (\text{对每个}x\text{ 单独成立})$$

然后把 n 个局部不等式相加. 每个变量被孤立出来, 多元问题就被降维成一元问题.

局部不等式需要精心设计. 一个合格的 $\ell(x)$ 要满足三条.

- **单项可证**. $f(x)\geq \ell(x)$ 是一元不等式, 比原问题简单.
- **求和闭合**. $\sum \ell(a_i)$ 恰好能被约束条件算尽, 得到目标常数 C.
- **等号兼容**. 局部等号条件与整体等号条件一致, 保证求和后等号可以同时成立.

满足这些条件的 $\ell(x)$ 通常取一次函数. 一次函数求和后只出现 $\sum a_i$, 最容易被约束条件吸收.

5.10.1 局部放缩

我们先看一个典型的局部放缩例子.

例题 5.38 设 a,b,c 为正实数. 证明, 得

$$\frac{a^2}{a+b}+\frac{b^2}{b+c}+\frac{c^2}{c+a}\geq \frac{a+b+c}{2}$$

解 左边是三个结构相同的分式之和, 右边只含变量的和. 这提示我们为每一项找一个线性的局部估计

$$\frac{a^2}{a+b}\geq \alpha a+\beta b$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

使得三式相加后右边恰好是 $\frac{a+b+c}{2}$.

利用齐次性, 令 $b=1$, 记 $t=\frac{a}{b}$, 把局部不等式化为单变量形式

$$g(t)=\frac{t^2}{t+1}\geq \alpha t+\beta$$

等号条件 $a=b$ 对应 $t=1$. 取 g 在 $t=1$ 处的切线. $g(1)=\frac{1}{2}$、$g'(t)=\frac{t^2+2t}{(t+1)^2}$、$g'(1)=\frac{3}{4}$. 切线为 $y=\frac{1}{2}+\frac{3}{4}(t-1)=\frac{3t-1}{4}$. 据此猜测局部不等式

$$\frac{a^2}{a+b}\geq \frac{3a-b}{4}$$

作差得

$$\frac{a^2}{a+b}-\frac{3a-b}{4}=\frac{4a^2-(3a-b)(a+b)}{4(a+b)}=\frac{a^2-2ab+b^2}{4(a+b)}=\frac{(a-b)^2}{4(a+b)}\geq 0$$

局部不等式成立, 等号当且仅当 $a=b$. 同理写出另外两式, 三式相加.

$$\sum_{\text{cyc}}\frac{a^2}{a+b}\geq \frac{(3a-b)+(3b-c)+(3c-a)}{4}=\frac{2(a+b+c)}{4}=\frac{a+b+c}{2}$$

等号成立当且仅当 $a=b=c$.

注 5.12

本例中局部不等式的线性式 $\frac{3a-b}{4}$, 是比值函数 $g(t)$ 在等号点处的切线. 这也说明局部化的作用方式, 并且引出下一小节的切线法.

切线法

5.10.2 切线法

设我们要在约束 $a_1+\cdots+a_n=S$ 下证明 $\sum f(a_i)\geq C$, 且等号在所有变量相等 ($a_i=x_0=\frac{S}{n}$) 时取得. 我们希望找到线性函数 $\ell(x)=kx+b$, 使 $f(x)\geq \ell(x)$, 并且求和后恰好得到 C. 这两个要求唯一地决定 ℓ.

- 等号兼容要求 $f(x_0)=kx_0+b$, 即直线过点 $(x_0,f(x_0))$.
- 令 $h(x)=f(x)-kx-b$, 则 $h(x)\geq 0$ 且 $h(x_0)=0$, 即 x_0 是 h 的最小值点, 故 $h'(x_0)=0$, 即 $k=f'(x_0)$.

于是 ℓ 是 f 在等号点 x_0 处的**切线**.

$$f(x)\geq f(x_0)+f'(x_0)(x-x_0)$$

这就是**切线法**. 它在等号成立点作曲线的切线, 用切线代替曲线参与求和. 线性式求和后只含 $\sum a_i$, 约束条件把它算尽.

$$\sum_{i=1}^{n}f(a_i)\geq nf(x_0)+f'(x_0)\left(\sum_{i=1}^{n}a_i-nx_0\right)=nf(x_0)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

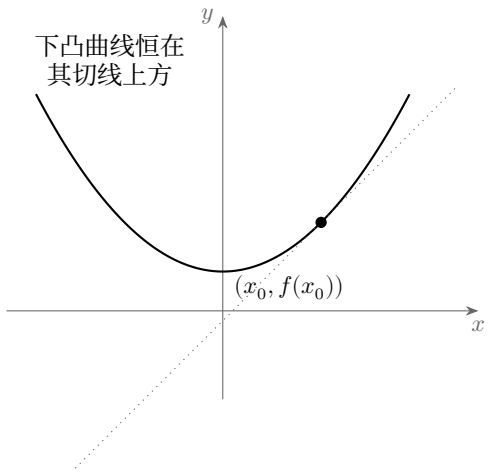


图 5.4　切线法的几何直观, $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$

几何上, 若 f 是下凸函数, 其图象恒在任一切线的上方 (第3.6节), 切线不等式自动成立. 但切线法的适用范围不止于凸函数; 只要作差后能在取值范围内定号, 即使 f 改变凸性, 方法仍然有效. 操作分四步.

- **定等号点**. 由对称性与约束条件猜测等号点 x_0(通常为变量的平均值).
- **写切线**.$\ell(x) = f(x_0) + f'(x_0)(x - x_0)$.
- **作差验证**. 计算 $f(x) - \ell(x)$, 通分后分子应分解出 $(x - x_0)^2$ 因子, 这是切线与曲线二阶接触的代数特征, 并确认其在取值范围内不变号.
- **求和消元**. 对各项求和, 线性项被约束条件算尽, 即得目标常数.

如果要证 $\sum f(a_i) \leq C$, 方向相反, 需要曲线在切线下方(上凸型).

例题 5.39　设 a, b, c 为正实数, 且 $a + b + c = 3$. 证明:

$$\frac{a}{1+a^2} + \frac{b}{1+b^2} + \frac{c}{1+c^2} \leq \frac{3}{2}$$

解　约束具有对称性, 等号在 $a = b = c = 1$ 取得. 考察单项函数 $f(x) = \frac{x}{1+x^2}$ 在 $x_0 = 1$ 处的切线. $f(1) = \frac{1}{2}$、$f'(x) = \frac{1-x^2}{(1+x^2)^2}$、$f'(1) = 0$. 切线是水平直线 $\ell(x) = \frac{1}{2}$. 作差得

$$\frac{1}{2} - \frac{x}{1+x^2} = \frac{1 + x^2 - 2x}{2(1+x^2)} = \frac{(x-1)^2}{2(1+x^2)} \geq 0$$

局部不等式 $\frac{x}{1+x^2} \leq \frac{1}{2}$ 对所有实数 x 成立. 三式相加即得

$$\sum_{\text{cyc}} \frac{a}{1+a^2} \leq \frac{3}{2}$$

等号成立当且仅当 $a = b = c = 1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 5.40 设 a,b,c 为正实数, 且 $a+b+c=3$. 证明:

$$\frac{a^2}{a+2}+\frac{b^2}{b+2}+\frac{c^2}{c+2}\geq 1$$

解 等号在 $a=b=c=1$ 取得. 考察 $f(x)=\frac{x^2}{x+2}$ 在 $x_0=1$ 处的切线. $f(1)=\frac{1}{3}$、$f'(x)=\frac{x^2+4x}{(x+2)^2}$、$f'(1)=\frac{5}{9}$. 切线为 $\ell(x)=\frac{1}{3}+\frac{5}{9}(x-1)=\frac{5x-2}{9}$. 作差得

$$\frac{x^2}{x+2}-\frac{5x-2}{9}=\frac{9x^2-(5x-2)(x+2)}{9(x+2)}=\frac{4x^2-8x+4}{9(x+2)}=\frac{4(x-1)^2}{9(x+2)}\geq 0$$

分子分解出了 $(x-1)^2$. 对 a,b,c 分别应用局部不等式并求和.

$$\sum_{\text{cyc}}\frac{a^2}{a+2}\geq\frac{5(a+b+c)-6}{9}=\frac{5\cdot 3-6}{9}=1$$

等号成立当且仅当 $a=b=c=1$.

例题 5.41 设 a,b,c 为正实数, 且 $a+b+c=1$. 证明:

$$\frac{1}{1+a^2}+\frac{1}{1+b^2}+\frac{1}{1+c^2}\leq\frac{27}{10}$$

解 等号在 $a=b=c=\frac{1}{3}$ 取得. 考察 $f(x)=\frac{1}{1+x^2}$ 在 $x_0=\frac{1}{3}$ 处的切线. $f\left(\frac{1}{3}\right)=\frac{9}{10}$、$f'(x)=\frac{-2x}{(1+x^2)^2}$、$f'\left(\frac{1}{3}\right)=-\frac{27}{50}$. 切线为 $\ell(x)=\frac{9}{10}-\frac{27}{50}\left(x-\frac{1}{3}\right)=\frac{27(2-x)}{50}$. 作差得

$$\frac{27(2-x)}{50}-\frac{1}{1+x^2}=\frac{27(2-x)(1+x^2)-50}{50(1+x^2)}=\frac{-27x^3+54x^2-27x+4}{50(1+x^2)}$$

分子按切线法的特征分解出 $(3x-1)^2$.

$$-27x^3+54x^2-27x+4=(3x-1)^2(4-3x)$$

当 $0<x<1$ 时, $(3x-1)^2\geq 0$ 且 $4-3x>0$, 故差式非负, 局部不等式成立. 三式相加.

$$\sum_{\text{cyc}}\frac{1}{1+a^2}\leq\frac{27}{50}(6-(a+b+c))=\frac{27}{50}\cdot 5=\frac{27}{10}$$

等号成立当且仅当 $a=b=c=\frac{1}{3}$.

注 5.13 (凹凸性改变时的切线法)
计算二阶导数, $f''(x)=\frac{6x^2-2}{(1+x^2)^3}$. 当 $0<x<\frac{1}{\sqrt{3}}$ 时 $f''(x)<0$, f 上凸; 当 $\frac{1}{\sqrt{3}}<x<1$ 时 $f''(x)>0$, f 下凸. 函数在区间 $(0,1)$ 内改变凹凸性, 琴生不等式要求在整个区间上凹凸一致,

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因而这里的前提不足. 切线法只要求作差后在取值范围上定号, 因此仍然适用. 它与琴生不等式的联系见 第5.15节.

注 5.14 (切线法的适用边界与对策)
切线法有以下边界, 需要相应变通.

- **作差变号**. 切线穿过了曲线, 说明在该取值范围上缺少全局的切线夹持, 需要另寻支撑线或改用他法.
- **等号落在边界**. 若最值在边界或端点取得, 例如某变量趋于 0, 对称点的切线失效. 此时可尝试过曲线**端点的割线**作为支撑线, 或改用调整法 (第5.14节).
- **约束为乘积形式**. 若约束为 $abc=1$ 这类乘积形式, 可先取对数或作代换 (第5.13节) 化为和式约束, 再行局部化.

局部化把多元问题降维成一元问题, 切线法给出了一种局部线性估计方法. 二者都是放缩.

知乎@余命数

5.11 构造思想

前面各节的方法都是对问题中**已有**的表达式施加变形. 构造思想更进一步. 我们引入一个新的数学对象, 可以是函数、方程、向量、几何图形, 甚至物理模型, 让隐藏的代数关系在新对象上显现.

构造法有规律可循. 待证式子的**结构特征**会暗示我们该构造什么. 判别式的形状对应二次函数, 内积的形状对应一对向量, 根式的形状对应一段距离, 和式加权的形状对应质点系重心. 本节展示这些对应关系.

5.11.1 构造二次函数与方程

回顾柯西不等式的经典证明 (第5.5节). 要证

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \le \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right)$$

关键在于右减左的形状酷似二次函数的判别式 $\Delta = B^2 - 4AC$. 于是构造

$$P(t) = \sum_{i=1}^{n} (a_i t - b_i)^2 = \left(\sum a_i^2\right) t^2 - 2\left(\sum a_i b_i\right) t + \sum b_i^2$$

这一步包含两个要点. 第一, $P(t)$ 是**平方和**, 平方非负, 故 $P(t) \ge 0$ 恒成立. 第二, 一个恒非负的二次函数与 t 轴至多有一个交点, 于是**判别式**把恒非负这个事实翻译成了系数之间的代数不等式 $\Delta \le 0$, 整理即得柯西不等式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

构造的对象也可以是方程. 当需要刻画某个值存在时, 方程有实根这一要求本身就会产生不等式.

例题 5.42 (判别式法求值域) 求函数 $f(x)=\dfrac{x^2+x+1}{x^2+1}$ 的值域.

解 设 y 是函数值, 则存在实数 x 使得

$$y=\frac{x^2+x+1}{x^2+1} \iff (y-1)x^2-x+(y-1)=0$$

把 y 看作参数, 构造出关于 x 的方程.y 在值域中等价于此方程有实根.

若 $y=1$, 方程化为 $-x=0$, 有解 $x=0$, 故 $y=1$ 可取.

若 $y\neq 1$, 这是关于 x 的二次方程, 有实根的充分必要条件是判别式非负.

$$\Delta=1-4(y-1)^2\geq 0 \iff (y-1)^2\leq\frac{1}{4}\iff \frac{1}{2}\leq y\leq\frac{3}{2}$$

端点可以取到: $y=\frac{1}{2}$ 时方程有重根 $x=-1$, 且 $f(-1)=\frac{1}{2}$; $y=\frac{3}{2}$ 时方程有重根 $x=1$, 且 $f(1)=\frac{3}{2}$.

故函数的值域为 $[\frac{1}{2},\frac{3}{2}]$.

注 5.15
判别式法通过构造一个以所求量为参数的方程, 用有实根替代直接求最值. 它与柯西不等式的判别式证明共同利用了一个事实. **二次式的非负性, 是不等式最基本、最直接的来源.**

5.11.2 构造向量与几何图形

许多代数结构在欧氏空间中有天然对应.

- $\sum a_ib_i$ 是向量**内积**, 由 $|\vec{m}\cdot\vec{n}|\leq|\vec{m}|\,|\vec{n}|$ 立即得到柯西不等式. 这正是 第5.5节 向量形式的观点;
- $\sqrt{a^2+b^2}$ 是平面上点 (a,b) 到原点的**距离**, 根式之和提示折线长度, 而折线不短于直线段正是闵可夫斯基不等式 (第5.5节) 的几何依据.

把代数式翻译为几何对象后, 三角不等式等直观事实可直接调用.

例题 5.43 证明: 对任意实数 x,

$$\sqrt{x^2+9}+\sqrt{x^2-6x+25}\geq\sqrt{58}$$

解 两个根式都是平方和开根的形状, 对应距离. 把第二项配方.

$$\sqrt{x^2-6x+25}=\sqrt{(x-3)^2+16}$$

在平面上构造三点 $A(0,3)$, $B(3,-4)$, 以及 x 轴上的动点 $P(x,0)$. 则

$$|PA|=\sqrt{x^2+9},\qquad |PB|=\sqrt{(x-3)^2+4^2}=\sqrt{x^2-6x+25}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由 A 在 x 轴上方而 B 在下方, 点 P 无论在 x 轴上如何移动, 折线 $A \to P \to B$ 的长度都不短于直线段 AB.

$$|PA| + |PB| \geq |AB| = \sqrt{(3-0)^2 + (-4-3)^2} = \sqrt{9+49} = \sqrt{58}$$

等号成立当且仅当 P 落在线段 AB 与 x 轴的交点处. 直线 AB 的方程为 $y = 3 - \frac{7}{3}x$, 令 $y = 0$ 得 $x = \frac{9}{7}$, 确实在线段范围内, 故等号可以取到.

原不等式成立, 且左边可取得最小值 $\sqrt{58}$.

注 5.16
构造几何图形需要恰当地摆放已知量. 本题把 $9 = 3^2$ 和 $16 = 4^2$ 摆成纵坐标, 把变量 x 摆成横坐标, 两个根式便各归其位. 若将 A, B 置于 x 轴同侧, 则需先作对称点再连线, 这是几何最值问题的标准手法.

5.11.3 构造物理与概率模型

构造的对象也可以来自物理. 在 x 轴上的点 $x_1, \dots, x_n$ 处分别放置质量为 $\lambda_1, \dots, \lambda_n$ 的质点 ($\lambda_i \geq 0, \sum \lambda_i = 1$), 则质点系的**重心**位于

$$\bar{x} = \sum_{i=1}^{n} \lambda_i x_i$$

对下凸函数 f, 曲线上各点 $(x_i, f(x_i))$ 的加权重心 $(\bar{x}, \sum \lambda_i f(x_i))$ 必落在连接各点的弦, 或其凸组合区域上, 因而位于曲线下方点 $(\bar{x}, f(\bar{x}))$ 的上方, 即

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leq \sum_{i=1}^{n} \lambda_i f(x_i)$$

这正是琴生不等式 (第3.6节). 重心模型给出琴生不等式的一个直观来源, 并且说明加权形式中 λ_i 就是权重. 我们还将在 第5.15节 中继续讨论.

注 5.17 (概率观点前瞻)
若把权重 λ_i 看作随机变量 X 取值 x_i 的概率, 则 $\bar{x} = E[X]$, 琴生不等式读作 $f(E[X]) \leq E[f(X)]$ (f 下凸). 取 $f(x) = x^2$ 即得 $E[X^2] \geq (E[X])^2$, 这正是方差非负. 柯西不等式、琴生不等式在概率论中都有对应形式, 详见 第10章.

构造辅助函数并用导数讨论

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

5.11.4 构造辅助函数

比较 A 与 B 的大小, 最直接的想法是作差 $A-B$. 构造思想把它函数化. 我们引入辅助函数 $F(x)=A(x)-B(x)$, 把比较两个数转化为研究一个函数的符号, 于是可以调用单调性这一工具(其严格判定由 第16章 给出).

例题 5.44 证明: 当 $x>0$ 时, $e^x>1+x+\dfrac{x^2}{2}$.

解 作差构造辅助函数

$$F(x)=e^x-1-x-\frac{x^2}{2}$$

则 $F(0)=0$, 且

$$F'(x)=e^x-1-x$$

由 第5.9节 的指数基本放缩, $e^x\geq 1+x$ 恒成立, 故 $F'(x)\geq 0$, F 在 $[0,+\infty)$ 上单调递增. 于是当 $x>0$ 时, $F(x)>F(0)=0$, 即 $e^x>1+x+\frac{x^2}{2}$.

注 5.18
基本放缩 $e^x\geq 1+x$ 可以递归加强. 对它求导或积分构造一次, 可得带二次项的更紧估计; 重复这一过程, 得到 e^x 的各阶多项式逼近, 这正是泰勒展开截断的雏形, 其严格理论由 第16章 给出.

例题 5.45 比较 e^{π} 与 π^e 的大小.

解 两数均为正, 取自然对数后等价于比较 π 与 $e\ln\pi$, 再同除以正数 $e\pi$, 等价于比较

$$\frac{\ln e}{e}\quad 与\quad \frac{\ln\pi}{\pi}$$

两个数具有完全相同的结构 $\frac{\ln x}{x}$, 只是代入的 x 不同, 这提示构造同一函数. 设

$$f(x)=\frac{\ln x}{x}\quad (x>0)$$

求导得 $f'(x)=\frac{1-\ln x}{x^2}$. 当 $x>e$ 时 $\ln x>1$, 故 $f'(x)<0$, f 在 $(e,+\infty)$ 上单调递减. 由 $e<\pi$ 得 $f(e)>f(\pi)$, 即 $\frac{\ln e}{e}>\frac{\ln\pi}{\pi}$. 逐步回溯等价变形, 得

$$e^{\pi}>\pi^e$$

构造法有一条统一的线索. **盯住待证式的结构**. 判别式、内积、距离、重心、差函数, 这些结构识别得越准确, 构造就越自然.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.12 齐次化与规范化

许多不等式带有**约束条件**, 例如 $a+b+c=1$ 或 $abc=1$, 或写成次数不齐的形式. 处理这类问题时, 先做次数分析往往能化繁为简. 次数分析可以把多元问题降维, 也可以把次数不齐的不等式转化为齐次形式.

5.12.1 齐次性

一个由变量 $a_1,\dots,a_n$ 构成的表达式 F 称为 d **次齐次**的, 若将所有变量同时乘以 $t>0$ 时 F 被乘以 t^d.

$$F(ta_1,\dots,ta_n)=t^d\cdot F(a_1,\dots,a_n)$$

不等式 $F\ge G$ 是**齐次的**, 当且仅当两边同次. 此时, 若对所有正实数 a_i 都有 $F\ge G$, 则对任意正数 t,

$$F(ta_i)=t^dF(a_i)\ge t^dG(a_i)=G(ta_i).$$

这个性质给出了一份自由度. **齐次不等式可以随意添加归一化约束**, 例如

$$a+b+c=1,\qquad abc=1,\qquad a^2+b^2+c^2=3$$

因为对任意原始变量, 总能找到合适的 t 使归一化条件成立.

- 若 $a+b+c=S>0$, 令 $t=\frac{1}{S}$, 则 $t(a+b+c)=1$;
- 若 $abc=P>0$, 令 $t=\frac{1}{\sqrt[3]{P}}$, 则 $(ta)(tb)(tc)=t^3P=1$.

齐次性保证了原不等式在 (a,b,c) 成立当且仅当在 (ta,tb,tc) 成立, 因而可**不妨假设**归一化条件成立. 这是 第5.4节 末内斯比特不等式中设 $a+b+c=1$ 的手法.

5.12.2 齐次化

另一些情形, 待证不等式左右次数不一致. 例如条件 $abc=1$ 出现在题设中, 而目标一侧为二次、一侧为一次. 此时**齐次化**的任务是用条件把常数 1 替换回有次数的表达式, 使等号两侧次数相同, 以便使用对称不等式操作.

例题 5.46 (齐次化) 设 a,b,c 为正实数, 且 $abc=1$. 证明:

$$a^2+b^2+c^2\ge a+b+c$$

解 左边 $a^2+b^2+c^2$ 是 2 次, 右边 $a+b+c$ 是 1 次, 不等式次数不齐. 直接用 $a^2\ge a$ 放缩只在 $a\ge 1$ 时成立, 会因遗漏 $abc=1$ 而失败.

我们用条件把右边补成与左边同次. 由 $abc=1$, 有

$$a+b+c=(a+b+c)(abc)^{1/3}.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

此式在 $abc=1$ 时与原右边相等, 故与原不等式等价. 现在两边都是 2 次, 已脱离约束, 只需证明对一切正实数

$$a^2+b^2+c^2 \geq (a+b+c)\sqrt[3]{abc}.$$

右边展开为 $a\sqrt[3]{abc}+b\sqrt[3]{abc}+c\sqrt[3]{abc}$, 即对称和 $[\frac{4}{3},\frac{1}{3},\frac{1}{3}]$; 左边是 $[2,0,0]$. 由优超关系 $(2,0,0)\succeq(\frac{4}{3},\frac{1}{3},\frac{1}{3})$(分量按不增序, $2\geq\frac{4}{3}$, $2+0\geq\frac{4}{3}+\frac{1}{3}$, 总次均为 2), Muirhead 定理, 见本节末直接给出结论. 代回 $abc=1$ 即得原不等式, 等号在 $a=b=c=1$ 取得.

此例说明, 若待证式不齐次, 可用约束补齐次数, 再考虑放缩.

例题 5.47 (规范化) 设 a,b,c 为正实数. 证明:

$$(a+b)(b+c)(c+a)\geq 8abc$$

余命数

解 不等式两端均为 3 次, 是齐次不等式. 由齐次性, 不妨规范化为 $abc=1$. 对任意正实数 a,b,c, 取 $t=(abc)^{-1/3}$, 则 $a'=ta, b'=tb, c'=tc$ 满足 $a'b'c'=1$, 而

$$\frac{(a'+b')(b'+c')(c'+a')}{a'b'c'}=\frac{(a+b)(b+c)(c+a)}{abc}$$

知乎@余命数

即原不等式在 (a,b,c) 与 (a',b',c') 处等价. 于是只需证 $abc=1$ 的情形, 此时即证 $(a+b)(b+c)(c+a)\geq 8$.

由 AM-GM,

知乎@余命数

$$a+b\geq 2\sqrt{ab},\quad b+c\geq 2\sqrt{bc},\quad c+a\geq 2\sqrt{ca}.$$

三式相乘.

$$(a+b)(b+c)(c+a)\geq 8\sqrt{a^2b^2c^2}=8abc.$$

等号当且仅当 $a=b=c$.

注 5.19 (齐次化与规范化的对称分工)
齐次化与规范化是齐次不等式处理的两面. 条件 $a+b+c=1$、$abc=1$ 等在源条件中原样出现, 便于减少一个变量, 这是**规范化**. 不等式未齐次时, 把 1 替换回条件的 1 次式 $(a+b+c)$ 或 $(abc)^{1/3}$, 这是**齐次化**. 实战常交叉使用. 先齐次化统一, 再规范化减元.

齐次化之后, 对称多项式记号 $s_1=a+b+c$、$s_2=ab+bc+ca$、$s_3=abc$ 会自然出现. 它们把三个对称多项式联系起来, 这就是下一小节的 Schur 不等式.

5.12.3 Schur 不等式

三元对称不等式中, 一个最基础的结果源于 Schur.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

定理 5.15 (Schur 不等式) 设 $a,b,c\ge 0, r\ge 0$ 为实数. 则

$$a^r(a-b)(a-c)+b^r(b-a)(b-c)+c^r(c-a)(c-b)\ge 0$$

等号成立当且仅当 $a=b=c$, 或 $r>0$ 且两数相等而第三者为 0, 例如 $a=b$, $c=0$.

证明 不妨设 $a\ge b\ge c\ge 0$, 对称性的好处见 第5.14节. 把左端按 a,b,c 重新组合.

$$a^r(a-b)(a-c)-b^r(a-b)(b-c)+c^r(c-a)(c-b)$$

注意 $(c-a)(c-b)=(a-c)(b-c)$(两个负因子之积), 故

$$\text{左端}=(a-b)\big[a^r(a-c)-b^r(b-c)\big]+c^r(a-c)(b-c).$$

知乎@余命数

逐项分析. 由 $a\ge b$, 有 $(a-b)\ge 0$. 又 $a\ge b$ 时 $a^r\ge b^r$ $(r\ge 0)$ 且 $a-c\ge b-c\ge 0$, 故

知乎@余命数

$$a^r(a-c)\ge b^r(a-c)\ge b^r(b-c),$$

知乎@余命数

即 $a^r(a-c)-b^r(b-c)\ge 0$. 因而第一项 $(a-b)[\cdots]\ge 0$. 又 $a\ge c$ 且 $b\ge c$, 故 $a-c\ge 0$, $b-c\ge 0$, 且 $c^r\ge 0$, 因而第二项 $c^r(a-c)(b-c)\ge 0$.

故左端为两个非负项之和, 故不小于 0.

至于等号条件, 若 $a=b=c$, 直接代入各项得 0. 若 $r>0$ 且 $a=b$, $c=0$, 则第一项为 0, 第二项 $c^r\cdot$ 非负 $=0$ $(c^r=0)$, 也成立. 反向, 设等号成立, 则两项皆须为零. 第一项为零要求 $a=b$, 或 $a^r(a-c)=b^r(b-c)$. 但 $r>0$ 且 $a>b$ 时 $a^r>b^r$, $a-c>b-c\ge 0$, 乘积严格更大, 故必 $a=b$. 第二项为零要求 $c=0$ ($r>0$ 时 $c^r=0$) 或 $a=b=c$. 结合得两种情形, $a=b=c$, 或 $a=b>c=0$ 及其置换. 当 $r=0$ 时 $c^r=1$, 第二项因子化为 $(a-c)(b-c)$, $a\ge b\ge c$ 且 $c>0$、$a>c$ 时该项严格正; 故 $r=0$ 时等号只在 $a=b=c$ 处成立. □

注 5.20 (Schur 的常用形式 ($r=1$))
$r=1$ 时, Schur 不等式写作

$$a(a-b)(a-c)+b(b-a)(b-c)+c(c-a)(c-b)\ge 0.$$

展开后用对称多项式 $s_1=a+b+c$、$s_2=ab+bc+ca$、$s_3=abc$ 表示为

$$s_1^3+9s_3\ge 4s_1s_2.$$

这是把三个对称多项式联系起来的不等式, 适合处理关于 $\sum a^3$、$\sum ab(a+b)$ 类目标的不等式.

例题 5.48 设 a,b,c 为正实数, 证明

$$a^3+b^3+c^3+3abc \geq ab(a+b)+bc(b+c)+ca(c+a)$$

解 右端 $\sum ab(a+b)=\sum(a^2b+ab^2)$, 故左右差为

$$\sum a^3-\sum a^2b-\sum ab^2+3abc.$$

用对称多项式表示, $\sum a^3=s_1^3-3s_1s_2+3s_3$; $\sum ab(a+b)=s_1s_2-3s_3$. 代入

$$(s_1^3-3s_1s_2+3s_3)-(s_1s_2-3s_3)+3s_3=s_1^3-4s_1s_2+9s_3.$$

这是 $r=1$ 的 Schur 不等式 $s_1^3+9s_3 \geq 4s_1s_2$ 的左端. 故成立,等号当且仅当 $a=b=c$ 或两数相等而第三者为 0. 由于 $a,b,c>0$, 等号仅在 $a=b=c$ 时取得.

5.12.4 Muirhead 定理与优超关系

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

Schur 处理的是三数对称单项式.Muirhead 定理则给出一个统一工具, 用于比较任意两个三元对称和.

定义 5.2 (对称和记号) 对非负实数指数 $\mathbf{a}=(a_1,a_2,a_3)$ 与正实数 x,y,z, 记

$$[\mathbf{a}]=\sum_{\sigma} x^{a_{\sigma(1)}}y^{a_{\sigma(2)}}z^{a_{\sigma(3)}}$$

其中 σ 取遍 $\{1,2,3\}$ 的所有使各项不同的排列, 相同单项式只计入一次. 例如 $[3,0,0]=x^3+y^3+z^3$、$[2,1,0]=\sum_{\text{sym}}x^2y=x^2y+x^2z+\cdots+z^2y$、$[1,1,1]=xyz$.

对称和记号为比较三元对称和提供统一记号, 优超关系则给出比较的判据.

定义 5.3 (优超关系) 设 $\mathbf{a}=(a_1,a_2,a_3)$ 和 $\mathbf{b}=(b_1,b_2,b_3)$ 的分量非负, 且按不增顺序排列. 称 **a 优超 b**, 记 $\mathbf{a} \succeq \mathbf{b}$, 若满足

$$a_1 \geq b_1, \quad a_1+a_2 \geq b_1+b_2, \quad a_1+a_2+a_3=b_1+b_2+b_3.$$

知乎@余命数

优超关系是 Muirhead 定理的比较判据.

定理 5.16 (Muirhead 定理) 设 $\mathbf{a} \succeq \mathbf{b}$, 则对所有正实数 x, y, z,

$$[\mathbf{a}] \geq [\mathbf{b}]$$

当 $\mathbf{a}$ 与 $\mathbf{b}$ 互为置换时等号恒成立; 否则等号成立当且仅当 $x = y = z$.

证明分两步. 第一步先证一个磨平引理, 它把一次微小的指数调整对应到对称和的单调不等式. 第二步再证优超关系总可通过有限次磨平达到.

引理 5.1 (单步磨平) 设 $\mathbf{a} = (a_1, a_2, a_3)$, $i, j \in \{1, 2, 3\}$ 且 $a_i \geq a_j$, $\epsilon \in [0, a_i - a_j]$. 记 $\mathbf{a}'$ 为把 a_i 减去 ϵ、把 a_j 加上 ϵ 所得的向量(其余分量不变). 则

$$[\mathbf{a}] \geq [\mathbf{a}']$$

证明 不妨设 $i = 1$, $j = 2$ (其余情形由对称性同理). 将对称和中的项按第三个变量取 z^{a_3} 分组, 每组只涉及前两个变量 x, y, 公因式为 $z^{a_3} \geq 0$. 配而成对, 每对差为

$$x^{a_1} y^{a_2} + x^{a_2} y^{a_1} - x^{a_1 - \epsilon} y^{a_2 + \epsilon} - x^{a_2 + \epsilon} y^{a_1 - \epsilon}.$$

提取公共因子 $x^{a_2} y^{a_2}$. 令 $R_1 = a_1 - a_2 - \epsilon \geq 0$, $R_2 = \epsilon \geq 0$, 则 $a_1 - a_2 = R_1 + R_2$, 括号内化为

$$x^{R_1 + R_2} + y^{R_1 + R_2} - x^{R_1} y^{R_2} - x^{R_2} y^{R_1} = (x^{R_1} - y^{R_1})(x^{R_2} - y^{R_2}).$$

由 $R_1, R_2 \geq 0$, 当 $x \geq y$ 时两因子同为非负; 当 $x \leq y$ 时同为非正. 故乘积非负, 每对差

$$x^{a_2} y^{a_2} \, (x^{R_1} - y^{R_1})(x^{R_2} - y^{R_2}) \geq 0.$$

乘上 $z^{a_3} \geq 0$ 后仍非负.

全部配对项相加, $[\mathbf{a}] - [\mathbf{a}'] \geq 0$. 等号当且仅当 $\epsilon = 0$ (平凡) 或 $x = y = z$. □

证明 (Muirhead 定理的证明) 设 $\mathbf{a} \succeq \mathbf{b}$, 且不妨设 $a_1 \geq a_2 \geq a_3$ 与 $b_1 \geq b_2 \geq b_3$ (重排分量不改 $[\,\cdot\,]$ 的值). 我们证明, 若 $\mathbf{a} \neq \mathbf{b}$, 则存在一次单步磨平, 把 $\mathbf{a}$ 变为某个 $\mathbf{c}$, 使 $\mathbf{c} \succeq \mathbf{b}$ 且

$$\sum_{k=1}^{3} (c_k - b_k)^2 < \sum_{k=1}^{3} (a_k - b_k)^2.$$

反复施行, 量 $\sum (a_k - b_k)^2$ 严格下降且非负, 有限步后必达到 $\mathbf{b}$. 每步由单步磨平引理使对称和不增, 故 $[\mathbf{a}] \geq \cdots \geq [\mathbf{b}]$.

知乎@余命数

设 $\mathbf{a}\neq\mathbf{b}$. 由总次数相等, 存在指标 i 使 $a_i>b_i$, 也存在指标 j 使 $a_j<b_j$. 取

$$\epsilon=\min(a_i-b_i,\ b_j-a_j)>0$$

令 $\mathbf{c}$ 为把 a_i 减去 ϵ、把 a_j 加上 ϵ 所得的向量. 则

- $c_i=a_i-\epsilon\geq b_i$, $c_j=a_j+\epsilon\leq b_j$, 其余分量不变;
- $\mathbf{c}$ 的分量之和仍等于 $\mathbf{b}$ 的分量之和;
- 对任意 $m=1,2$, 部分和 $\sum_{k=1}^{m}c_k^{\downarrow}\geq\sum_{k=1}^{m}b_k$ 仍成立, 其中 $c^{\downarrow}$ 为 $\mathbf{c}$ 按不增序重排后的向量. 这一点可由优超的标准性质得到, 即从较大分量向较小分量转移, 不破坏对 $\mathbf{b}$ 的优超关系.

因而 $\mathbf{c}\succeq\mathbf{b}$. 又

$$\sum_{k=1}^{3}(c_k-b_k)^2=\sum_{k=1}^{3}(a_k-b_k)^2-2\epsilon((a_i-b_i)+(b_j-a_j))+2\epsilon^2<\sum_{k=1}^{3}(a_k-b_k)^2$$

(由 $\epsilon\leq a_i-b_i$ 且 $\epsilon\leq b_j-a_j$ 得 $\epsilon<(a_i-b_i)+(b_j-a_j)$, 括号内为正, 且 $\epsilon>0$), 量严格下降.

反复施行上述磨平, 有限步后 $\mathbf{a}$ 化为 $\mathbf{b}$. 每次磨平由单步磨平引理使对称和不增, 故

$$[\mathbf{a}]\geq\cdots\geq[\mathbf{b}].$$

当 $\mathbf{a}$ 与 $\mathbf{b}$ 互为置换时等号恒成立; 否则至少有一次磨平 $\epsilon>0$, 引理严格等号仅在 $x=y=z$ 时发生, 故 Muirhead 等号当且仅当 $x=y=z$. □

注 5.21

Muirhead 定理把对称和的比较转化为指数向量的优超比较.AM-GM 等经典不等式都是它的特例.

- $(1,0,0)\succeq(\frac{1}{3},\frac{1}{3},\frac{1}{3})$, 给出 $x+y+z\geq3(xyz)^{1/3}$(分数指数时 Muirhead 仍适用);
- $(2,1,0)\succeq(1,1,1)$, 给出 $\sum_{\text{sym}}x^2y\geq6xyz$, 是 Schur 的近亲.

例题 5.49 设 a,b,c 为正实数. 证明:

$$a^5+b^5+c^5\geq a^3bc+b^3ca+c^3ab$$

解 左边是 $[5,0,0]$, 右边是 $[3,1,1]$. 检查优超关系.$(5,0,0)\succeq(3,1,1)$, 因为 $5\geq3$, $5+0\geq3+1$, 总次都是 5. 由 Muirhead 定理,

$$[5,0,0]\geq[3,1,1]$$

即 $a^5+b^5+c^5\geq a^3bc+b^3ca+c^3ab$. 等号当且仅当 $a=b=c$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

注 5.22 (计算量)
齐次化配合 Muirhead, 原则上能把三元对称齐次不等式化为对称和之差, 再由优超关系判定. 展开后项数往往很多, 需要先看清主导项与等号结构, 再决定是否走这条路.

5.13 代数代换思想

处理不等式时, 约束条件常常是主要困难. 三角不等式把三边锁在一个区域内, 乘积 $abc=1$ 让变量相互牵制, $x^2+y^2=1$ 将点限制在单位圆上. **代数代换**的思路是引入新变量, 把这些受限约束**化掉**, 要么变成自由的正实数, 要么变成更对称、更易处理的形式.

代换的目的通常有两个.

- **化约束为自由**. 把必须满足某条件的变量, 换成天然落在定义域内的新变量.
- **化偏倚为对称**. 把结构不对称的表达式, 换成轮换或对称的新形式, 便于放缩或套用已知不等式.

本节介绍三类最常见的代换.

5.13.1 Ravi 代换

设 a,b,c 是三角形的三边. 它们必须满足三角不等式

$$a<b+c,\quad b<c+a,\quad c<a+b.$$

直接在这个区域内讨论不等式较为繁琐.**Ravi 代换**把该受限区域转化为正实数的全体.

定义 5.4 (Ravi 代换) 设 a,b,c 为三角形的三边. 令

$$x=\frac{b+c-a}{2},\quad y=\frac{c+a-b}{2},\quad z=\frac{a+b-c}{2}$$

则 $x,y,z>0$, 且

$$a=y+z,\quad b=z+x,\quad c=x+y.$$

几何上, x,y,z 正是内切圆在三边上切点分出的切线长. 从顶点 A 出发的两段切线长相等, 记为 x; 从 B 出发记为 y; 从 C 出发记为 z. 于是

$$a=y+z,\quad b=z+x,\quad c=x+y,$$

且 $x,y,z>0$ 自动保证三角不等式成立. 反过来, 任意正实数 x,y,z 都对应一个三角形. 关于三边的不等式就转化为关于三个**自由正实数** x,y,z 的不等式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

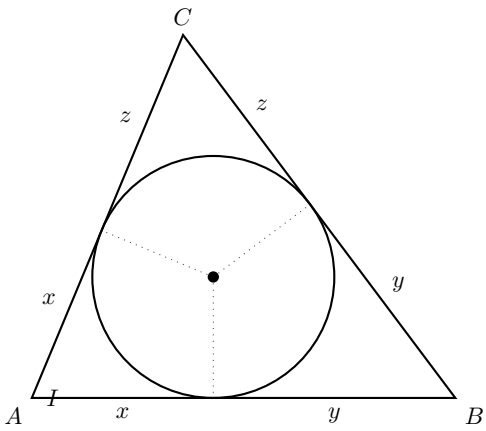


图 5.5 Ravi 代换的几何来源, 内切圆切点分边为 x, y, z

例题 5.50 设 a, b, c 为三角形的三边. 证明, 得

$$a^2 + b^2 + c^2 < 2(ab + bc + ca)$$

解 作 Ravi 代换 $a = y + z$、$b = z + x$、$c = x + y$ $(x, y, z > 0)$.

$$\begin{aligned} a^2 + b^2 + c^2 &= (y + z)^2 + (z + x)^2 + (x + y)^2 \\ &= 2(x^2 + y^2 + z^2) + 2(xy + yz + zx) \\ ab + bc + ca &= (y + z)(z + x) + (z + x)(x + y) + (x + y)(y + z) \\ &= x^2 + y^2 + z^2 + 3(xy + yz + zx) \end{aligned}$$

待证 $a^2 + b^2 + c^2 < 2(ab + bc + ca)$ 化为

$$2(x^2 + y^2 + z^2) + 2(xy + yz + zx) < 2(x^2 + y^2 + z^2) + 6(xy + yz + zx),$$

即 $0 < 4(xy + yz + zx)$. 由 $x, y, z > 0$, 右边严格为正, 结论成立.

等号仅在三角形退化时可能出现, 此时 x, y, z 中至少一个为 0, 已越出正实数范围, 故严格不等式成立.

注 5.23
Ravi 代换把三角形这个几何约束, 完全翻译成 $x, y, z > 0$ 这个代数条件. 此后 AM-GM、柯西、Schur 等工具都可在无约束的正实数上自由施展.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.13.2 乘积代换

当约束为 $abc=1$ 时，一个常用代换是

$$a=\frac{x}{y},\quad b=\frac{y}{z},\quad c=\frac{z}{x}\quad (x,y,z>0).$$

此时 $abc=1$ 自动成立，许多分式表达式中的乘积因子会两两抵消. 即便不显式写出 x,y,z，识别 $\frac{1}{a}=bc$ 这类关系，就已经是在做乘积代换.

例题 5.51　设 a,b,c 为正实数，且 $abc=1$. 证明：

$$\frac{1}{a^3(b+c)}+\frac{1}{b^3(c+a)}+\frac{1}{c^3(a+b)}\geq\frac{3}{2}$$

解　约束 $abc=1$ 的关键作用是把倒数与乘积等同.$\frac{1}{a}=bc$、$\frac{1}{b}=ca$、$\frac{1}{c}=ab$. 对左边应用柯西不等式的 Engel 形式.

$$\sum_{\text{cyc}}\frac{1}{a^3(b+c)}=\sum_{\text{cyc}}\frac{\frac{1}{a^2}}{a(b+c)}\geq\frac{\left(\sum\frac{1}{a}\right)^2}{\sum a(b+c)}.$$

分母 $\sum a(b+c)=\sum(ab+ac)=2(ab+bc+ca)$. 分子中 $\sum\frac{1}{a}=\sum bc=ab+bc+ca$（由 $\frac{1}{a}=bc$）. 令 $p=ab+bc+ca$，则

$$\sum_{\text{cyc}}\frac{1}{a^3(b+c)}\geq\frac{p^2}{2p}=\frac{p}{2}.$$

由 AM-GM，$p=ab+bc+ca\geq3\sqrt[3]{a^2b^2c^2}=3(abc)^{2/3}=3$，故 $\frac{p}{2}\geq\frac{3}{2}$. 结论成立.

等号成立当且仅当 $a=b=c$（AM-GM 等号）且 Engel 形式等号条件 $a(b+c)$ 成比例于 $\frac{1}{a}$ 同时满足，即 $a=b=c=1$.

注 5.24
此证并未显式写出 $a=\frac{x}{y}$ 等形式，但 $\frac{1}{a}=bc$ 的识别已经起到了乘积代换的作用. 若令 $a=\frac{x}{y}$、$b=\frac{y}{z}$、$c=\frac{z}{x}$，则乘积约束自动消失，分式中的 x,y,z 会大量抵消. 在更复杂的分式不等式中，显式代换往往更省事.

例题 5.52　设 a,b,c 为正实数，且 $abc=1$. 证明：

$$a^2+b^2+c^2+3\geq2(a+b+c)$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

解 由 $abc=1$ 及 AM-GM, $a+b+c\geq 3\sqrt[3]{abc}=3$. 待证等价于

$$a^2+b^2+c^2-2(a+b+c)+3\geq 0,$$

即 $(a-1)^2+(b-1)^2+(c-1)^2\geq 0$, 恒成立. 等号当且仅当 $a=b=c=1$.

我们利用 $abc=1$ 确认等号点 $a=b=c=1$ 可达. 约束 $abc=1$ 保证了等号的可达性.

5.13.3 三角代换

当变量满足 $x^2+y^2=1$ 时, 令 $x=\cos\theta, y=\sin\theta$ $(\theta\in\mathbb{R})$, 约束自动满足, 且可调用三角恒等式. 更一般地, 若 $x,y,z>0$ 且 $xy+yz+zx=1$, 可令

$$x=\tan\frac{A}{2},\quad y=\tan\frac{B}{2},\quad z=\tan\frac{C}{2},$$

其中 A,B,C 为某三角形的内角. 半角正切公式保证 $xy+yz+zx=1$ 当 $A+B+C=\pi$(三角恒等式在 第6章 中系统讲解).

例题 5.53 设 x,y 为实数, 且 $x^2+y^2=1$. 求 $x+2y$ 的最大值.

解 令 $x=\cos\theta, y=\sin\theta$. 则

$$x+2y=\cos\theta+2\sin\theta=\sqrt{5}\left(\frac{1}{\sqrt{5}}\cos\theta+\frac{2}{\sqrt{5}}\sin\theta\right)=\sqrt{5}\sin(\theta+\phi),$$

其中 $\cos\phi=\frac{2}{\sqrt{5}}$, $\sin\phi=\frac{1}{\sqrt{5}}$. 最大值为 $\sqrt{5}$, 当 $\sin(\theta+\phi)=1$ 时取得.

也可用柯西不等式.$(x+2y)^2\leq(1^2+2^2)(x^2+y^2)=5$, 最大值 $\sqrt{5}$. 三角代换给出了等号点的参数表示.

例题 5.54 设 $x,y,z>0$, 且 $xy+yz+zx=1$. 证明:

$$x+y+z\geq\sqrt{3}$$

解 由 $x^2+y^2+z^2\geq xy+yz+zx$, 得

$$(x+y+z)^2=x^2+y^2+z^2+2(xy+yz+zx)\geq 3(xy+yz+zx)=3,$$

故 $x+y+z\geq\sqrt{3}$. 等号当且仅当 $x=y=z=\frac{1}{\sqrt{3}}$.

用三角代换. 令 $x=\tan\frac{A}{2}$、$y=\tan\frac{B}{2}$、$z=\tan\frac{C}{2}$、$A+B+C=\pi$、$A,B,C>0$. 则 $xy+yz+zx=1$ 自动成立. 目标化为

$$\tan\frac{A}{2}+\tan\frac{B}{2}+\tan\frac{C}{2}\geq\sqrt{3}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在 $A=B=C=\frac{\pi}{3}$ 时, $\tan\frac{\pi}{6}=\frac{1}{\sqrt{3}}$, 三者和为 $\sqrt{3}$, 等号成立. 由 tan 在 $(0,\frac{\pi}{2})$ 上下凸, 琴生不等式给出

$$\frac{1}{3}\sum\tan\frac{A}{2}\geq\tan\frac{A+B+C}{6}=\tan\frac{\pi}{6}=\frac{1}{\sqrt{3}},$$

故 $\sum\tan\frac{A}{2}\geq\sqrt{3}$. 三角代换把代数约束 $xy+yz+zx=1$ 翻译为角度约束 $A+B+C=\pi$, 从而可调用三角函数的凸性.

> **注 5.25 (增量代换)**
> 对称二元问题中, 常令 $a=m+t, b=m-t(m=\frac{a+b}{2}$ 为平均值, $t=\frac{a-b}{2}$ 为偏差). 这把两个变量拆成平均水平与偏离程度, 便于分析对称函数对偏差的单调性, 见 第5.14节. 变量代换的更多技巧见 第5.4节.

代换法的共通线索是**识别约束的结构, 选择能自动满足该约束的参数化**. 三角形用切线长, 乘积用分式, 圆用三角函数.

5.14 排序与调整思想

我们面对对称不等式与轮换不等式时, 注意到变量位置可以互换而表达式值保持不变. 因此可以**假设变量已按大小排列**, 把无序的代数和变为有序. 排序不等式 (第5.6节) 是这一思想的起点; 本节把它推广到切比雪夫不等式, 并引入与之对偶的**调整法**.

5.14.1 不妨设序的合法性

动手之前, 我们先分清两种对称性.

- **完全对称**. 任意置换变量, 表达式不变. 例如 $a^3+b^3+c^3$ 与 $ab+bc+ca$. 此时可任意假设 $a\geq b\geq c$.
- **轮换对称**. 仅在循环置换 $(a,b,c)\mapsto(b,c,a)$ 下不变. 例如 $a^2b+b^2c+c^2a$. 此时随意假设 $a\geq b\geq c$ 会改变表达式的值, 轮换对称对反序排列一般给出不同结果. 第5.6节 的第一道例题正因如此才改用 AM-GM.

确认合法性之后, 设序才是合法的简化步骤.

5.14.2 切比雪夫不等式

排序不等式比较的是同序和与反序和, 切比雪夫不等式给出更精细的定量形式, 即两个同向单调序列的平均乘积不小于各自平均的乘积.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定理 5.17 (切比雪夫不等式) 设 $a_1 \le a_2 \le \ldots \le a_n$ 与 $b_1 \le b_2 \le \ldots \le b_n$ 为两组同序实数. 则

$$\frac{a_1b_1 + a_2b_2 + \cdots + a_nb_n}{n} \ge \left(\frac{a_1 + \cdots + a_n}{n}\right)\left(\frac{b_1 + \cdots + b_n}{n}\right)$$

即 $\sum a_ib_i \ge \frac{1}{n}(\sum a_i)(\sum b_i)$. 若两组反序 $(a_1 \le \ldots \le a_n, b_1 \ge \ldots \ge b_n)$, 则不等号反向. 等号成立当且仅当 $a_1 = \cdots = a_n$ 或 $b_1 = \cdots = b_n$.

证明 对任意 $i, j \in \{1, \ldots, n\}$, 由两组同序, $(a_i - a_j)$ 与 $(b_i - b_j)$ 同号, 或至少一个为零, 故

$$(a_i - a_j)(b_i - b_j) \ge 0.$$

展开得 $a_ib_i + a_jb_j \ge a_ib_j + a_jb_i$. 对所有 i, j 求和.

知乎@余命数

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(a_ib_i + a_jb_j) \ge \sum_{i=1}^{n}\sum_{j=1}^{n}(a_ib_j + a_jb_i).$$

左边, 对 j 求和时每个 a_ib_i 出现 n 次, 对 i 求和时每个 a_jb_j 出现 n 次, 故左边 $= 2n\sum a_ib_i$. 右边, $\sum_i\sum_j a_ib_j = (\sum a_i)(\sum b_j)$, 同理 $\sum_i\sum_j a_jb_i = (\sum a_j)(\sum b_i)$, 故右边 $= 2(\sum a_i)(\sum b_i)$. 因此

$$2n\sum a_ib_i \ge 2(\sum a_i)(\sum b_i),$$

即 $\sum a_ib_i \ge \frac{1}{n}(\sum a_i)(\sum b_i)$. 等号成立当且仅当对所有 i, j 都有 $(a_i - a_j)(b_i - b_j) = 0$, 即所有 a_i 相等或所有 b_i 相等. □

注 5.26
切比雪夫不等式对排序不等式作定量加强, 给出同序和与平均乘积之间的差距; 排序不等式只说同序和最大. 二者共享同一个代数内核, 即 $(a_i - a_j)(b_i - b_j) \ge 0$.

例题 5.55 设 $a_1 \le a_2 \le \ldots \le a_n$ 为实数, f 为递增函数. 则

$$\frac{1}{n}\sum_{i=1}^{n}a_if(a_i) \ge \left(\frac{1}{n}\sum_{i=1}^{n}a_i\right)\left(\frac{1}{n}\sum_{i=1}^{n}f(a_i)\right)$$

解 由 f 递增, 序列 $(f(a_1), \ldots, f(a_n))$ 与 $(a_1, \ldots, a_n)$ 同序. 由切比雪夫不等式即得. 这是切比雪夫最典型的应用模式, 递增序列与它在递增函数下的像, 乘积平均不小于各自平均的乘积.

知乎@余命数

例题 5.56　设 a,b,c 为正实数. 证明, 得

$$a^3+b^3+c^3 \geq \frac{1}{3}(a+b+c)(a^2+b^2+c^2)$$

解　不等式完全对称, 不妨设 $a \geq b \geq c > 0$. 则 $a^2 \geq b^2 \geq c^2$, 序列 (a,b,c) 与 (a^2,b^2,c^2) 同序. 由切比雪夫,

$$\frac{a\cdot a^2+b\cdot b^2+c\cdot c^2}{3} \geq \left(\frac{a+b+c}{3}\right)\left(\frac{a^2+b^2+c^2}{3}\right),$$

即 $\frac{a^3+b^3+c^3}{3} \geq \frac{(a+b+c)(a^2+b^2+c^2)}{9}$, 两边同乘 3 即得

$$a^3+b^3+c^3 \geq \frac{1}{3}(a+b+c)(a^2+b^2+c^2).$$

等号当且仅当 $a=b=c$.

知乎@余命数

例题 5.57　设 a,b,c 为正实数. 证明, 得

$$\frac{a^2}{b}+\frac{b^2}{c}+\frac{c^2}{a} \geq a+b+c$$

解　由 AM-GM 局部配对.

$$\frac{a^2}{b}+b \geq 2a, \quad \frac{b^2}{c}+c \geq 2b, \quad \frac{c^2}{a}+a \geq 2c.$$

知乎@余命数

三式相加.$\sum \frac{a^2}{b}+(a+b+c) \geq 2(a+b+c)$, 即 $\sum \frac{a^2}{b} \geq a+b+c$. 等号当且仅当 $a=b=c$.

此证用局部 AM-GM. 工具的选择应服务于结构, 切比雪夫适合两组同序序列的乘积和; 当结构是平方比线性时, 局部 AM-GM 往往更直接.

5.14.3 调整法

调整法的思想与排序对偶, 排序把变量按大小排好, 调整法主动**改变**变量的值, 使它们趋于相等, 并证明每一步调整都使目标函数朝我们希望的方向变化.

定理 5.18 (调整法的基本框架) 设 $f(x_1,\dots,x_n)$ 为对称函数, 在紧集 $\{x_i \geq 0,\ \sum x_i = S\}$ 上连续. 若对任意 $x_i \neq x_j$, 把它们替换为 $\frac{x_i+x_j}{2}, \frac{x_i+x_j}{2}$ (或更一般的磨平变换)后 f 的值不减, 则 f 在 $x_1=\cdots=x_n=\frac{S}{n}$ 处取得最大值; 若替换后 f 的值不增, 则在该点取得最小值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

证明 在紧集 $\{(x_1,\dots,x_n): x_i \geq 0, \sum x_i = S\}$ 上, 连续函数 f 取到极值. 设极值点为 $(x_1^*,\dots,x_n^*)$. 若某两个分量 $x_i^* \neq x_j^*$, 则磨平变换使 f 严格变化, 与极值性矛盾. 故极值点处所有分量相等. □

例题 5.58 (AM-GM 的调整法证明) 证明: 对正实数 $a_1,\dots,a_n$, 有 $\frac{a_1+\cdots+a_n}{n} \geq \sqrt[n]{a_1\dots a_n}$, 等号当且仅当 $a_1=\cdots=a_n$.

解 和 $S=a_1+\cdots+a_n$ 固定时, 只需证积 $P=a_1\dots a_n$ 在 $a_1=\cdots=a_n=\frac{S}{n}$ 时最大.

设某两个变量 $a_i \neq a_j$, 不妨 $a_i < a_j$. 令 $a_i' = a_j' = \frac{a_i+a_j}{2}$. 则 $a_i'+a_j' = a_i+a_j$, 和保持不变, 而

$$a_i'a_j' = \left(\frac{a_i+a_j}{2}\right)^2 \geq a_ia_j$$

(由 $(\sqrt{a_i}-\sqrt{a_j})^2 \geq 0$). 等号当且仅当 $a_i=a_j$. 故积严格增大.

反复将不相等的变量对磨平, 积单调递增. 在紧集 $\{a_i \geq 0 : \sum a_i = S\}$ 上, 连续函数 P 取到最大值; 若最大值点有某两个变量不等, 则可磨平使积严格增大, 矛盾. 故最大值点必使所有变量相等, 此时 $P=(\frac{S}{n})^n$. 故对任意正实数, $P \leq (\frac{S}{n})^n$, 即 AM-GM 成立.

例题 5.59 设 a,b,c 为正实数, 且 $a+b+c=3$. 证明:

$$ab+bc+ca \leq 3$$

解 由 $(a+b+c)^2 = a^2+b^2+c^2+2(ab+bc+ca) = 9$, 有 $ab+bc+ca = \frac{9-(a^2+b^2+c^2)}{2}$. 由 $a^2+b^2+c^2 \geq \frac{(a+b+c)^2}{3} = 3$(QM-AM 或柯西), 故 $ab+bc+ca \leq \frac{9-3}{2} = 3$. 等号当且仅当 $a=b=c=1$.

用调整法. 设 $a \neq b$, 令 $a'=b'=\frac{a+b}{2}$, $c'=c$. 则 $a'+b'+c'=3$ 不变, 而

$$\begin{aligned} a'b'+b'c'+c'a'-(ab+bc+ca) &= \left(\frac{a+b}{2}\right)^2 + c(a+b) - ab - c(a+b) \\ &= \frac{(a+b)^2}{4} - ab = \frac{(a-b)^2}{4} \geq 0. \end{aligned}$$

故磨平后 $ab+bc+ca$ 不减. 反复磨平至 $a=b=c=1$, 得 $ab+bc+ca \leq 3$.

注 5.27 (调整法的严密性)
有限步调整在变量取有理数时可以精确到达相等点; 对一般实数, 需在紧集, 如 $a_i \geq 0$, $\sum a_i = S$ 上利用连续函数取到最值, 再证明最值点必使所有变量相等(否则可继续调整使目标更优, 矛盾). 这一步依赖极值的存在性, 其严格基础见分析初步. 第5.12节 的 Muirhead 定理中, 磨平作用于指数; 调整法作用于变量本身.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

排序在假设中设序, 调整在过程中磨平, 两者都让对称性直接进入证明.

5.15 凸性思想

注 5.28
本节凸性思想中的二阶导数判定、Jensen 应用与 Young 不等式, 依赖 第3章 的凹凸定义与第16章. 琴生直觉与均值特例是入口; 其余性质在此基础上展开.

我们前面各节的工具, 放缩、构造、代换、排序, 大多在代数层面运作. 凸性思想换一个视角, 从函数图象的**弯曲方向**看待不等式. 若一个函数下凸, 如 $y=x^2$ 或 $y=e^x$, 连接图象上任意两点的弦必位于曲线的上方. 把这一几何事实翻译成代数语言, 就是琴生不等式.

第3.6节 已经介绍过琴生不等式的完整陈述与证明. 本节把它放在不等式思想的脉络中, 展示它如何作为均值不等式等经典结果的**母体**, 如何与切线法(第5.10节) 互为表里, 以及它的边界在哪里.

凸性、琴生与切线对照

5.15.1 琴生不等式

定理 5.19 (琴生不等式, 回顾) 设 f 为区间 I 上的下凸函数. 则对任意 $x_1,\dots,x_n\in I$ 与非负权重 λ_i ($\sum\lambda_i=1$),

$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right)\le\sum_{i=1}^{n}\lambda_i f(x_i)$$

若 f 上凸, 不等号反向. 对严格凹凸函数, 等号成立当且仅当所有正权重对应的 x_i 相等.

证明 证明见 定理 3.10. □

取等权重 $\lambda_i=\frac{1}{n}$, 得算术平均形式.

$$f\left(\frac{x_1+\cdots+x_n}{n}\right)\le\frac{f(x_1)+\cdots+f(x_n)}{n},$$

即平均的函数值不超过函数值的平均(下凸情形).

例题 5.60 (由琴生导出 AM-GM) 对正数 $x_1,\dots,x_n$, 证明 $\frac{x_1+\cdots+x_n}{n}\ge\sqrt[n]{x_1\cdots x_n}$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

解 取 $f(x)=\ln x$, 定义域 $(0,+\infty)$.$f''(x)=-\frac{1}{x^2}<0$, f 严格上凸. 由上凸的琴生不等式,

$$\frac{\ln x_1+\cdots+\ln x_n}{n}\le \ln\left(\frac{x_1+\cdots+x_n}{n}\right).$$

左边 $=\ln\sqrt[n]{x_1\dots x_n}$. 由 ln 严格递增, 得 $\sqrt[n]{x_1\dots x_n}\le\frac{x_1+\cdots+x_n}{n}$.

例题 5.61 (幂平均不等式的雏形) 设 $a_1,\dots,a_n>0, p>1$. 证明:

$$\left(\frac{a_1^p+\cdots+a_n^p}{n}\right)^{1/p}\ge\frac{a_1+\cdots+a_n}{n}$$

解 取 $f(x)=x^p$ $(x>0)$.$p>1$、$f''(x)=p(p-1)x^{p-2}>0$、f 下凸. 由琴生,

$$\left(\frac{a_1+\cdots+a_n}{n}\right)^p\le\frac{a_1^p+\cdots+a_n^p}{n}.$$

两边开 p 次方(正数, 方向不变)即得. 等号当且仅当 $a_1=\cdots=a_n$.

这是幂平均不等式 $M_p\ge M_1$ $(p>1)$ 的特例. 更一般地, 对 $p>q$, 有 $M_p\ge M_q$, 证明可通过对 $t\mapsto t^{p/q}$ 的凸性与琴生完成.

例题 5.62 (Young 不等式) 设 $a,b>0, p,q>1$ 且 $\frac{1}{p}+\frac{1}{q}=1$. 证明:

$$ab\le\frac{a^p}{p}+\frac{b^q}{q}$$

解 取 $f(x)=e^x$, 下凸. 令 $x_1=\ln(a^p), x_2=\ln(b^q)$, 权重 $\lambda_1=\frac{1}{p},\lambda_2=\frac{1}{q}$. 由琴生,

$$\exp\left(\frac{1}{p}\ln(a^p)+\frac{1}{q}\ln(b^q)\right)\le\frac{1}{p}e^{\ln(a^p)}+\frac{1}{q}e^{\ln(b^q)}.$$

左边 $=\exp(\ln a+\ln b)=ab$, 右边 $=\frac{a^p}{p}+\frac{b^q}{q}$. 等号当且仅当 $a^p=b^q$.

5.15.2 权重选择的艺术

琴生不等式的权重 λ_i 可按问题需要自由选取, 恰当选择权重可把约束条件直接编码进不等式.

例题 5.63 设 $a,b,c>0$, 且 $a+b+c=1$. 证明:

$$a\ln a+b\ln b+c\ln c\ge-\ln 3$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

解 取 $f(x)=x\ln x\ (x>0)$.$f''(x)=\frac{1}{x}>0$, f 下凸. 由琴生(等权重),

$$f\left(\frac{a+b+c}{3}\right)\le\frac{f(a)+f(b)+f(c)}{3}.$$

即 $\frac{1}{3}\ln\frac{1}{3}\le\frac{1}{3}(a\ln a+b\ln b+c\ln c)$, 故 $a\ln a+b\ln b+c\ln c\ge\ln\frac{1}{3}=-\ln 3$. 等号当且仅当 $a=b=c=\frac{1}{3}$.

这是信息论中熵的上界 $\sum a_i\ln a_i\ge-\ln n$(在 $\sum a_i=1$ 时)的三元特例.

5.15.3 与切线法的对照

切线法 (第5.10节) 与琴生不等式联系紧密, 对下凸函数曲线恒在切线上方, 即

$$f(x)\ge f(x_0)+f'(x_0)(x-x_0).$$

这正是切线法的局部不等式. 对 n 个点求和并利用 $\sum x_i=nx_0$, 得 $\sum f(x_i)\ge nf(x_0)$, 这与琴生在等权重、等号点为平均值时的结论一致.

二者的分工如下.

- **琴生**. 需函数在整个区间上凹凸性一致, 一次给出全局结论; 权重灵活, 适合平均值约束.
- **切线法**. 只需在取值范围内作差定号; 适合凹凸性改变但仍有支撑线的情形 (第5.10节 末例).

例题 5.64 (同一问题的两种视角) 设 $a,b,c>0$, 且 $a+b+c=3$. 证明:

$$a^2+b^2+c^2\ge 3$$

解 从琴生视角看, 取 $f(x)=x^2$, 下凸 ($f''=2>0$), 由琴生不等式得

$$\left(\frac{a+b+c}{3}\right)^2\le\frac{a^2+b^2+c^2}{3},$$

即 $1\le\frac{a^2+b^2+c^2}{3}$, 故 $a^2+b^2+c^2\ge 3$.

从切线法视角看, 等号点 $x_0=1$、$f(1)=1$、$f'(x)=2x$、$f'(1)=2$, 切线 $\ell(x)=1+2(x-1)=2x-1$. 作差得 $x^2-(2x-1)=(x-1)^2\ge 0$, 求和即得 $\sum a^2\ge\sum(2a-1)=2\cdot 3-3=3$.

琴生一步到位; 切线法展示了局部线性估计加求和消元的机制.

5.15.4 凹凸性改变时

若函数在区间内改变凹凸性(出现拐点), 琴生不等式要求整个区间上凹凸一致, 此处前提不足. 此时切线法、分段讨论或调整法往往更有效.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 5.65 设 $a,b,c\in(0,1)$, 且 $a+b+c=1$. 证明:

$$\frac{1}{1+a^2}+\frac{1}{1+b^2}+\frac{1}{1+c^2}\le\frac{27}{10}$$

解 此题已在 第5.10节 中用切线法证明. 函数 $f(x)=\frac{1}{1+x^2}$ 的二阶导数 $f''(x)=\frac{6x^2-2}{(1+x^2)^3}$ 在 $x=\frac{1}{\sqrt{3}}$ 处变号. 在 $(0,\frac{1}{\sqrt{3}})$ 上上凸, 在 $(\frac{1}{\sqrt{3}},1)$ 上下凸. 取值范围 $(0,1)$ 跨越拐点, 琴生要求整个区间上凹凸一致, 此处前提不足. 切线法在 $x_0=\frac{1}{3}$ 处作支撑线, 作差后分子分解出 $(3x-1)^2(4-3x)$, 在 $(0,1)$ 上定号, 顺利完成证明.

支撑线的存在只要求作差在取值范围内定号.

知乎@余命数

注 5.29 (七种思想的对应关系)

- **放缩** (第5.9节). 构造中间量, 保持方向, 把握精度;
- **局部化** (第5.10节). 化整体为局部, 切线法是最彻底的形式;
- **构造** (第5.11节). 引入新对象, 让代数关系显现;
- **齐次化与规范化** (第5.12节). 次数分析, Schur 与 Muirhead;
- **代数代换** (第5.13节). 化约束为自由, Ravi、乘积、三角;
- **排序与调整** (第5.14节). 设序与磨平, 切比雪夫与调整法;
- **凸性**(本节). 从弯曲方向看不等式, 琴生为母体, 与切线法互为表里.

5.16 进阶不等式

知乎@余命数

第5.5节 给出了柯西不等式与闵可夫斯基不等式 $p=2$ 的情形. 本节把它们推广到一般的 p 次幂, 并引入幂平均不等式与嵌入不等式.

5.16.1 赫尔德不等式

定理 5.20 (赫尔德不等式) 设 $x_i>0,y_i>0(i=1,2,\dots,n),p>1$ 且 $\frac{1}{p}+\frac{1}{q}=1$, 则

$$\sum_{i=1}^{n}x_iy_i\le\left(\sum_{i=1}^{n}x_i^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}y_i^q\right)^{\frac{1}{q}}\tag{5-1}$$

等号当且仅当 x_i^p 与 y_i^q 成比例时成立. 当 $0<p<1$ 时不等式反向.

证明 先证 Young 不等式: 对 $a,b>0,p>1,\frac{1}{p}+\frac{1}{q}=1$, 有

$$ab\le\frac{a^p}{p}+\frac{b^q}{q}\tag{5-2}$$

这由 $f(t)=\frac{t^p}{p}+\frac{1}{q}-t^{1/p}$ 在 $t>0$ 上的最小值给出, 或由凸函数 e^x 的切线不等式给出.

设 $A=(\sum x_i^p)^{1/p}$, $B=(\sum y_i^q)^{1/q}$. 若 $A=0$ 或 $B=0$ 则结论平凡. 否则令 $a_i=\frac{x_i}{A}, b_i=\frac{y_i}{B}$, 则 $\sum a_i^p=\sum b_i^q=1$. 对每个 i 应用 Young 不等式:

$$a_ib_i\le\frac{a_i^p}{p}+\frac{b_i^q}{q}$$

知乎@余命数

对 i 求和: $\sum a_ib_i\le\frac{1}{p}\sum a_i^p+\frac{1}{q}\sum b_i^q=\frac{1}{p}+\frac{1}{q}=1$. 还原即得 (5-1). □

知乎@余命数

取 $p=q=2$, 赫尔德不等式退化为柯西不等式. 赫尔德不等式是柯西不等式在一般 L^p 空间中的推广.

5.16.2 闵可夫斯基不等式

知乎@余命数

定理 5.21 (闵可夫斯基不等式) 设 $x_i>0, y_i>0(i=1,2,\dots,n), p>1$, 则

$$\left(\sum_{i=1}^n(x_i+y_i)^p\right)^{\frac{1}{p}}\le\left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^n y_i^p\right)^{\frac{1}{p}}\tag{5-3}$$

等号当且仅当 x_i 与 y_i 成比例时成立. 当 $0<p<1$ 时不等式反向.

证明 展开 $(x_i+y_i)^p=(x_i+y_i)^{p-1}\cdot(x_i+y_i)$, 求和:

知乎@余命数

$$\sum(x_i+y_i)^p=\sum x_i(x_i+y_i)^{p-1}+\sum y_i(x_i+y_i)^{p-1}$$

对右端两项分别应用赫尔德不等式(取 $\frac{1}{p}+\frac{1}{q}=1, q=\frac{p}{p-1}$):

$$\sum x_i(x_i+y_i)^{p-1}\le\left(\sum x_i^p\right)^{\frac{1}{p}}\left(\sum(x_i+y_i)^p\right)^{\frac{1}{q}}$$

同理处理第二项. 两式相加后两边除以 $(\sum(x_i+y_i)^p)^{1/q}$, 利用 $1-\frac{1}{q}=\frac{1}{p}$ 即得 (5-3). □

取 $p=2$, 闵可夫斯基不等式给出 $\sqrt{\sum(x_i+y_i)^2}\le\sqrt{\sum x_i^2}+\sqrt{\sum y_i^2}$, 即 L^2 三角不等式. 几何上, $n=2$ 情形正是三角形两边之和大于第三边的坐标表达.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.16.3 幂平均值不等式

定理 5.22 (幂平均值不等式) 对 n 个正实数 $x_1,\dots,x_n$ 和权值 $\alpha_i\ge 0(\sum\alpha_i=1)$, 定义 p 次幂平均

$$M_p=\left(\sum_{i=1}^{n}\alpha_i x_i^p\right)^{\frac{1}{p}}$$

若 $0<p<q$, 则 $M_p\le M_q$, 等号当且仅当 $x_1=\cdots=x_n$ 时成立.

证明 设 $r=\frac{q}{p}>1$. 对 $\sum\alpha_i x_i^p$ 应用 Jensen 不等式($t\mapsto t^r$ 在 $t>0$ 上为下凸函数):

$$\left(\sum_{i=1}^{n}\alpha_i x_i^p\right)^r\le\sum_{i=1}^{n}\alpha_i(x_i^p)^r=\sum_{i=1}^{n}\alpha_i x_i^q$$

知乎@余命数

两边取 $\frac{1}{q}$ 次幂即得 $M_p\le M_q$. □

取 $p=1$, $q=2$ 得算术平均不超过平方平均. 取 $p\to 0$ 的极限得几何平均 $M_0=\prod x_i^{\alpha_i}$, 于是 $M_0\le M_p$ 对一切 $p>0$ 成立, 这正是加权均值不等式.

5.16.4 嵌入不等式

知乎@余命数

定理 5.23 (嵌入不等式) 在 $\triangle ABC$ 中, 对任意实数 x、y、z, 有

$$x^2+y^2+z^2\ge 2yz\cos A+2zx\cos B+2xy\cos C \tag{5-4}$$

等号当且仅当 $x:y:z=\sin A:\sin B:\sin C$ 时成立.

证明 构造向量 $\vec{u}=(x,y,z)$, 并利用二次型非负性. 将 (5-4) 的右端移到左端:

$$x^2+y^2+z^2-2yz\cos A-2zx\cos B-2xy\cos C\ge 0$$

这是关于 x、y、z 的二次型. 由 $A+B+C=\pi$, 该二次型的系数矩阵的行列式为零(因为 $(\sin A,\sin B,\sin C)$ 是核中的向量), 且其主子式非负, 故半正定, 不等式成立.

知乎@余命数

等号条件: 二次型的核由 $\sin A:\sin B:\sin C$ 给出, 这正是取等条件. □

例题 5.66 (嵌入不等式与柯西不等式的联系) 在 (5-4) 中取 $A=B=C=\frac{\pi}{3}$ (正三角形), 得

$$x^2+y^2+z^2\ge yz+zx+xy$$

这等价于 $(x-y)^2+(y-z)^2+(z-x)^2 \geq 0$,是恒等式 $(x+y+z)^2 \geq 3(xy+yz+zx)$ 的等价形式.

知乎@余命数

例题 5.67 (用嵌入不等式证明三角形不等式) 在 $\triangle ABC$ 中证明 $a^2+b^2+c^2 \geq 4\sqrt{3}\,S$(Weitzenböck 不等式).

解 在嵌入不等式 (5-4) 中取 $x=a, y=b, z=c$:

$$a^2+b^2+c^2 \geq 2bc\cos A + 2ca\cos B + 2ab\cos C$$

由余弦定理 $2bc\cos A = b^2+c^2-a^2$ 等,代入后右边 $= a^2+b^2+c^2-(a^2+b^2+c^2)+(a^2+b^2+c^2) = a^2+b^2+c^2$,这给出恒等式. 改取 $x=y=z=1$:

知乎@余命数

$$3 \geq 2(\cos A+\cos B+\cos C)$$

即 $\cos A+\cos B+\cos C \leq \frac{3}{2}$. 再结合 $S=\frac{1}{2}ab\sin C$ 与正弦定理,可推出 Weitzenböck 不等式 $a^2+b^2+c^2 \geq 4\sqrt{3}\,S$.

5.17 切比雪夫与 Schur 不等式

知乎@余命数

第5.14节 给出了切比雪夫不等式的单侧形式并用配对求和证明, 第5.12节 引入了 Schur 不等式 $r \geq 0$ 的形式. 本节给出切比雪夫不等式的双侧完整形式, 它直接从排序不等式 (第5.6节) 推出; 再把 Schur 不等式写成一般参数 $t>0$ 的形式, 建立 $t=1$ 时与三角形不等式的联系.

5.17.1 切比雪夫不等式的双侧形式

定理 5.24 (切比雪夫不等式) 设 $a_1 \geq a_2 \geq \cdots \geq a_n$ 与 $b_1 \geq b_2 \geq \cdots \geq b_n$ 为两组同序实数, 则

$$n\sum_{i=1}^{n} a_ib_i \geq \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{i=1}^{n} b_i\right) \geq n\sum_{i=1}^{n} a_ib_{n+1-i} \tag{5-5}$$

等号成立当且仅当 $a_1=\cdots=a_n$ 或 $b_1=\cdots=b_n$.

第5.14节 证明了右半不等式 $\sum a_ib_i \geq \frac{1}{n}(\sum a_i)(\sum b_i)$. 这里用排序不等式给出双侧的统一证明.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 考虑 $b_1, b_2, \dots, b_n$ 的 n 个循环移位. 对 $k = 0, 1, \dots, n-1$, 记

$$S_k = \sum_{i=1}^{n} a_i b_{i+k},$$

其中下标 $i+k$ 模 n 取值, 即 $b_{n+j} = b_j$.$S_0 = \sum a_i b_i$ 是同序和, $S_1, \dots, S_{n-1}$ 是乱序和. 由排序不等式, $S_k \le S_0$ 对一切 k 成立. 将这 n 个不等式相加:

$$\sum_{k=0}^{n-1} S_k \le n\, S_0 = n \sum_{i=1}^{n} a_i b_i.$$

左端对每个固定的 i, b_{i+k} $(k = 0, \dots, n-1)$取遍 $b_1, \dots, b_n$ 各一次, 故

$$\sum_{k=0}^{n-1} S_k = \sum_{i=1}^{n} a_i \sum_{k=0}^{n-1} b_{i+k} = \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{j=1}^{n} b_j\right).$$

右半不等式得证.

左半不等式: 将序列 $(b_1, \dots, b_n)$ 替换为 $(-b_n, -b_{n-1}, \dots, -b_1)$, 它与 $(a_1, \dots, a_n)$ 同序. 对这两组数应用刚证的右半不等式:

$$n \sum_{i=1}^{n} a_i(-b_{n+1-i}) \ge \left(\sum a_i\right)\left(\sum(-b_{n+1-i})\right) = -\left(\sum a_i\right)\left(\sum b_i\right),$$

即 $-n\sum a_i b_{n+1-i} \ge -(\sum a_i)(\sum b_i)$, 两边乘 -1 得左半不等式.

等号条件: 右半等号要求所有 $S_k = S_0$, 即每个循环移位都达到同序和. 由排序不等式的等号条件, 这要求 $a_1 = \cdots = a_n$ 或 $b_1 = \cdots = b_n$. □

注 5.30 (Abel 求和证法)
切比雪夫不等式也可用 Abel 分部求和证明. 记 $\bar{a} = \frac{1}{n}\sum a_i, \bar{b} = \frac{1}{n}\sum b_i$, 则

$$\sum_{i=1}^{n} a_i b_i - n\bar{a}\bar{b} = \sum_{i=1}^{n} (a_i - \bar{a})(b_i - \bar{b}).$$

令 $\alpha_i = a_i - \bar{a}, \beta_i = b_i - \bar{b}$, 则 $\sum \alpha_i = 0$. 记 $A_k = \sum_{i=1}^{k} \alpha_i$. 由 a_i 递减, 前 k 个最大元素之和不小于 $k\bar{a}$, 故 $A_k \ge 0 (k < n)$.Abel 分部求和给出

$$\sum_{i=1}^{n} \alpha_i \beta_i = \sum_{k=1}^{n-1} A_k(\beta_k - \beta_{k+1}).$$

由 b_i 递减有 $\beta_k - \beta_{k+1} = b_k - b_{k+1} \ge 0$, 右端每项非负, 故 $\sum a_i b_i \ge n\bar{a}\bar{b} = \frac{1}{n}(\sum a_i)(\sum b_i)$. 这与 第5.14节 的配对求和共享同一个恒等式

$$n\sum a_i b_i - \left(\sum a_i\right)\left(\sum b_i\right) = \sum_{1 \le i < j \le n} (a_i - a_j)(b_i - b_j),$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

只是分组方式不同.

5.17.2 Schur 不等式的一般参数形式

定理 5.25 (Schur 不等式) 设 $x,y,z\geq 0,t>0$, 则

$$x^t(x-y)(x-z)+y^t(y-z)(y-x)+z^t(z-x)(z-y)\geq 0 \tag{5-6}$$

等号成立当且仅当 $x=y=z$, 或两数相等而第三者为 0.

第5.12节 用参数 $r\geq 0$ 陈述了同一不等式并给出了证明. 这里取 $t>0$, 证明思路一致, 记录如下以备引用.

证明 由对称性, 不妨设 $x\geq y\geq z\geq 0$. 将左端按 x,y 配合再分离 z 项:

$$\text{左端}=(x-y)\left[x^t(x-z)-y^t(y-z)\right]+z^t(x-z)(y-z).$$

其中 $(z-x)(z-y)=(x-z)(y-z)$(两个负因子之积). 逐项判断符号. 由 $x\geq y$ 有 $x-y\geq 0;t>0$ 时 $x^t\geq y^t$, 且 $x-z\geq y-z\geq 0$, 故 $x^t(x-z)\geq y^t(y-z)$, 第一项非负. 第二项中 $z^t\geq 0,x-z\geq 0,y-z\geq 0$, 亦非负. 两项之和 ≥ 0.

等号条件:两项同时为零. 第一项为零要求 $x=y$ 或 $x^t(x-z)=y^t(y-z);t>0$ 且 $x>y$ 时 $x^t>y^t$ 且 $x-z>y-z\geq 0$, 乘积严格更大, 故必 $x=y$. 第二项为零要求 $z=0$($t>0$ 时 $z^t=0$)或 $x=z$ 或 $y=z$. 结合得 $x=y=z$, 或 $x=y>z=0$ 及其置换. □

5.17.3 $t=1$ 与三角形不等式

取 $t=1$, Schur 不等式展开为

$$x^3+y^3+z^3+3xyz\geq xy(x+y)+yz(y+z)+zx(z+x). \tag{5-7}$$

用对称多项式 $s_1=x+y+z$, $s_2=xy+yz+zx$, $s_3=xyz$ 表示, 这就是 第5.12节 中的 $s_1^3+9s_3\geq 4s_1s_2$.

这个代数不等式与几何中的三角形不等式有直接联系. 设 a,b,c 为三角形三边, 由 Ravi 代换 $a=y+z,b=z+x,c=x+y(x,y,z>0)$, 三角形不等式 $a+b>c$ 等价于 $x,y,z>0$. 把 a,b,c 代入 (5-7) 的左端:

$$\sum a^3+3abc=\sum(y+z)^3+3(y+z)(z+x)(x+y).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

展开化简得 $2(x+y+z)^3-6xyz$. 由均值不等式 $(x+y+z)^3\geq 27xyz$, 该式非负, 于是

$$a^3+b^3+c^3+3abc\geq ab(a+b)+bc(b+c)+ca(c+a)$$

对一切三角形三边成立.

反过来, 许多关于三角形三边的对称不等式经 Ravi 代换后化为关于正实数 x,y,z 的 Schur 型不等式, Schur 不等式提供了代数保证.

例题 5.68 (Schur 与三角形不等式) 设 a,b,c 为三角形三边. 证明

$$abc\geq (b+c-a)(c+a-b)(a+b-c).$$

解 记 $s_1=a+b+c$, $s_2=ab+bc+ca$, $s_3=abc$. 展开右端:

$$(b+c-a)(c+a-b)(a+b-c)=(s_1-2a)(s_1-2b)(s_1-2c).$$

乘开得 $s_1^3-2s_1^2(a+b+c)+4s_1(ab+bc+ca)-8abc=-s_1^3+4s_1s_2-8s_3$. 于是

$$abc-(b+c-a)(c+a-b)(a+b-c)=s_3-(-s_1^3+4s_1s_2-8s_3)=s_1^3-4s_1s_2+9s_3.$$

由 Schur 不等式 $t=1$, $s_1^3+9s_3\geq 4s_1s_2$, 右端 ≥ 0. 等号当且仅当 $a=b=c$(正三角形).

另一条路线: Ravi 代换 $a=y+z$, $b=z+x$, $c=x+y$ $(x,y,z>0)$, 则 $b+c-a=2x, c+a-b=2y, a+b-c=2z$, 不等式化为 $(y+z)(z+x)(x+y)\geq 8xyz$. 由 AM-GM, $y+z\geq 2\sqrt{yz}$, $z+x\geq 2\sqrt{zx}$, $x+y\geq 2\sqrt{xy}$, 三式相乘即得. 等号当且仅当 $x=y=z$, 即 $a=b=c$.

5.17.4 应用例题

例题 5.69 (切比雪夫双侧: 反序乘积和的上界) 设 $a_1\geq a_2\geq\cdots\geq a_n>0$ 且 $\sum_{i=1}^n a_i=1$. 证明

$$\sum_{i=1}^n a_ia_{n+1-i}\leq\frac{1}{n}.$$

解 在切比雪夫双侧不等式 (5-5) 中取 $b_i=a_i$. 左半给出

$$\Big(\sum a_i\Big)^2\geq n\sum a_ia_{n+1-i},$$

即 $1\geq n\sum a_ia_{n+1-i}$, 故 $\sum a_ia_{n+1-i}\leq\frac{1}{n}$. 等号当且仅当 $a_1=\cdots=a_n=\frac{1}{n}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

右半给出 $\sum a_i^2 \geq \frac{1}{n}$，即 QM-AM. 双侧合在一起夹住反序乘积和与同序乘积和.

例题 5.70 (Schur $t=2$) 设 $x,y,z\geq 0$. 证明

$$x^4+y^4+z^4+xyz(x+y+z)\geq x^3(y+z)+y^3(z+x)+z^3(x+y).$$

解 在 Schur 不等式 (5-6) 中取 $t=2$:

$$x^2(x-y)(x-z)+y^2(y-z)(y-x)+z^2(z-x)(z-y)\geq 0.$$

展开第一项 $x^2(x-y)(x-z)=x^4-x^3(y+z)+x^2yz$, 三项求和得

$$\sum x^4-\sum x^3(y+z)+xyz\sum x\geq 0,$$

移项即得目标不等式. 等号当且仅当 $x=y=z$, 或两数相等而第三者为 0.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

习题

5.1. (放缩)证明:对任意正整数 $n\geq 2$,

$$\sum_{k=2}^{n}\frac{1}{k^2}<1-\frac{1}{n}$$

提示:我们把每一项放缩为 $\frac{1}{k^2}<\frac{1}{k(k-1)}=\frac{1}{k-1}-\frac{1}{k}$, 再裂项求和.

5.2. (放缩)证明:对任意 $x\in(0,\frac{\pi}{2})$,

$$\frac{2x}{\pi}<\sin x<x$$

知乎@余命数

提示:我们研究函数 $f(x)=\frac{\sin x}{x}$ 的单调性, 也可比较单位圆中的面积.

5.3. (局部化 / 切线法)设 $a,b,c>0$ 且 $a+b+c=3$. 证明:

$$\frac{a}{a^2+1}+\frac{b}{b^2+1}+\frac{c}{c^2+1}\leq\frac{3}{2}$$

提示:我们对 $f(x)=\frac{x}{x^2+1}$ 在 $x_0=1$ 处作切线.

5.4. (局部化)设 $a,b,c>0$ 且 $a+b+c=1$. 证明:

$$\frac{a^2}{a+1}+\frac{b^2}{b+1}+\frac{c^2}{c+1}\geq\frac{1}{4}$$

知乎@余命数

提示:我们对 $f(x)=\frac{x^2}{x+1}$ 在 $x_0=\frac{1}{3}$ 处作切线.

5.5. (构造)求函数 $f(x)=\sqrt{x^2-2x+5}+\sqrt{x^2+4x+8}$ 的最小值. 提示:我们把每个根式解释为平面上动点到定点的距离.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.6. (构造)证明:对任意实数 a,b,

$$(a^2+1)(b^2+1)\geq(ab+1)^2$$

提示:我们构造一个二次函数,或者展开后作差;也可把它看作柯西不等式的特例.

5.7. (齐次化)设 $a,b,c>0$ 且 $a+b+c=1$. 证明:

$$\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\geq 9$$

提示:我们用 AM-HM,或先齐次化再用 AM-GM.

5.8. (Schur) 设 $a,b,c\geq 0$. 证明,得

$$a^3+b^3+c^3+3abc\geq ab(a+b)+bc(b+c)+ca(c+a)$$

提示:我们把它写成 $r=1$ 的 Schur 不等式.

5.9. (Muirhead) 设 $a,b,c>0$. 证明,得

$$a^4+b^4+c^4\geq a^2bc+b^2ca+c^2ab$$

提示:我们比较对称和 $[4,0,0]$ 与 $[2,1,1]$,验证优超关系.

5.10. (Ravi 代换)设 a,b,c 为三角形的三边. 证明,得

$$\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}<2$$

提示:我们作 Ravi 代换 $a=y+z$ 等,把问题化到正实数上.

5.11. (乘积代换)设 $a,b,c>0$ 且 $abc=1$. 证明:

$$\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}\geq\frac{3}{2}$$

提示:我们把它看成内斯比特不等式;可用 $abc=1$ 简化,或直接用 Engel 形式.

5.12. (三角代换)设 x,y 为实数且 $x^2+y^2=1$. 求 $3x+4y$ 的最大值与最小值. 提示:我们令 $x=\cos\theta$, $y=\sin\theta$,或用柯西不等式.

5.13. (切比雪夫)设 $a\geq b\geq c>0$. 证明,得

$$\frac{a}{b+c}+\frac{b}{c+a}+\frac{c}{a+b}\geq\frac{3}{2}$$

提示:我们注意到 (a,b,c) 与 $(\frac{1}{b+c},\frac{1}{c+a},\frac{1}{a+b})$ 同序,用切比雪夫;也可直接用内斯比特.

5.14. (调整法)设 $a,b,c>0$ 且 $a+b+c=3$. 用调整法证明:

$$a^2+b^2+c^2\geq 3$$

提示:我们将不相等的两变量磨平为平均值,证明平方和严格减小,直至变量相等.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

5.15. (凸性 / 琴生)设 $a,b,c>0$ 且 $a+b+c=\pi$. 证明:

$$\sin a+\sin b+\sin c\le\frac{3\sqrt{3}}{2}$$

提示:我们用 $\sin x$ 在 $(0,\pi)$ 上的上凸性与琴生不等式.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
命数

第二卷 · 周期·离散·随机

中学数学之旅



本卷承接基础语言，收录三角函数、数列、统计、计数原理、概率五章. 把圆周上的周期、正整数上的递推，以及数据、计数与随机度量串成一条线；严格极限与泰勒留给分析卷.

知乎 @ 余命数 (Midnight Forever)

Project Akari

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

本卷导读

知乎@余命数

本卷线索

第二卷把周期、离散与随机串在同一条函数与度量的线索上. 三角描述圆周上的连续振动, 数列描述正整数上的递推演化, 计数为随机事件提供可数的模型空间, 统计与概率则分别从数据与公理两端逼近不确定性.

章节地图

知乎@余命数

- ▷ **三角函数** 从弧度与单位圆建立正弦、余弦等基本量, 展开图像与性质、同角关系、诱导与和差倍角. 单位圆是推导恒等式的支点, 这里的图像与变换同第一卷的函数图像、复合变换一脉相承; 三角函数既是几何量, 也是第三卷解三角形的公共语言.
- ▷ **数列** 以数学归纳法与 $\sum / \prod$ 记号为工具, 研究等差、等比、差比与斐波那契等模型, 进而讨论递推、通项、不动点、单调与周期性. 归纳法是本章论证的基石, 递推模型的每一步分析都建立在它之上.
- ▷ **统计** 从描述数据到抽样设计, 再到用样本推断总体.
- ▷ **计数原理** 加法与乘法原理是计数的起点, 由此建立排列与组合, 再以二项式定理收束.
- ▷ **概率** 在计数提供的可数模型上, 用事件与概率公理刻画不确定性, 并讨论条件概率、独立性与错排等经典问题.

统计、计数、概率三章依次沿描述、计数、度量推进, 合起来构成从处理数据到公理化不确定性的完整链条.

向后展望

第三卷进入几何主干, 平面向量先行, 再解三角形与解析几何, 参数与极坐标紧贴坐标法. 第四卷建立极限与导数的严格理论, 数列的极限行为、级数与泰勒展开以及概率中的渐近结果在那里获得完整处理.

知乎@余命数

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

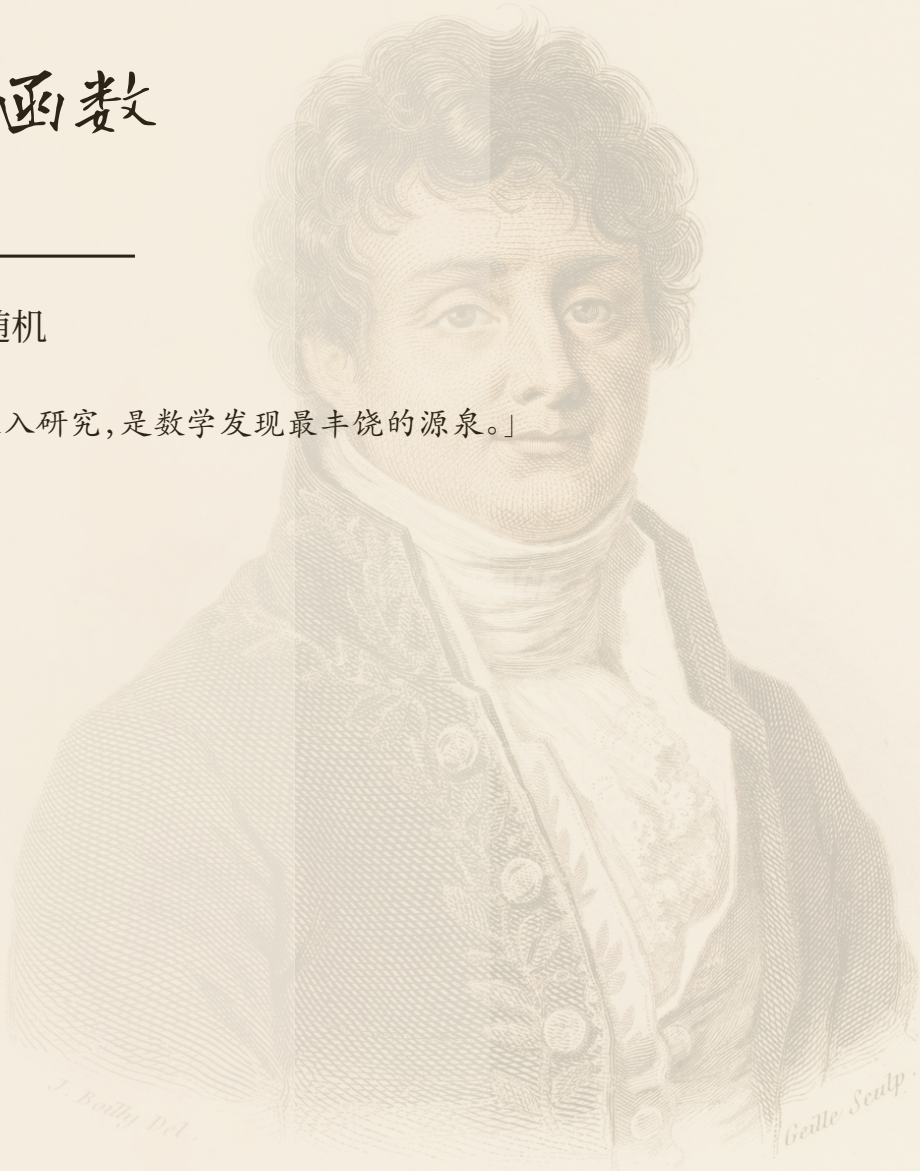
第 6 章

三角函数

周期·离散·随机

「对自然的深入研究，是数学发现最丰饶的源泉。」

——傅里叶



知乎@余命数

第六章 三角函数

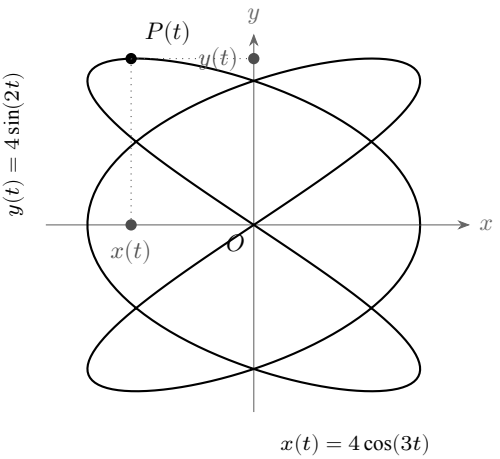


图 6.1 利萨茹曲线:两个独立三角振动的叠加

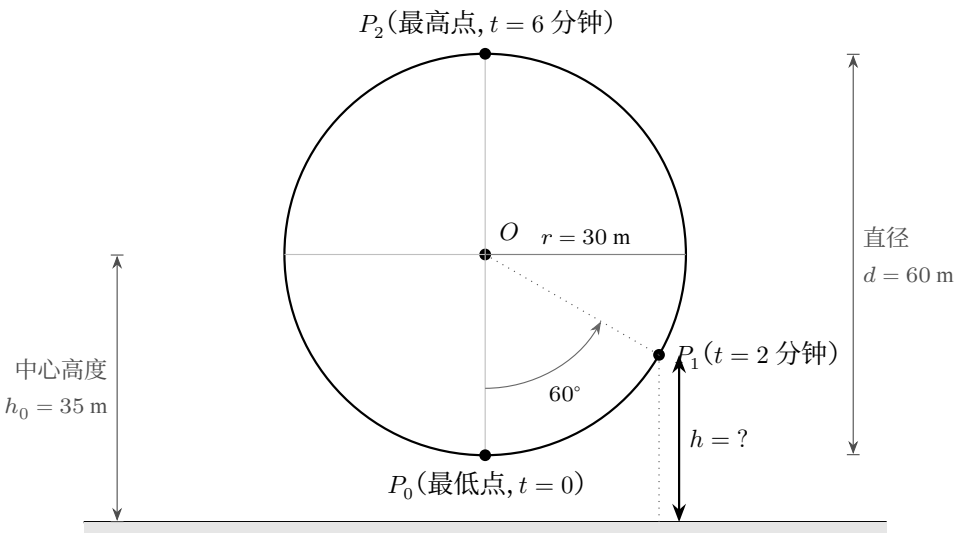
屏幕上的光点同时参与两种独立的运动. 一个余弦波控制它水平往复振动,另一个正弦波控制它垂直往复振动. 两种运动叠加后, 光点描出稳定、封闭而复杂的图案, 我们称之为利萨茹曲线. 图案的最终形态由频率比与初相位决定. 同频同相得直线段, 同频异相得椭圆(相位差为 π 时退化为直线段), 频率比为简单整数比时得封闭的多环图案. 改变水平或垂直振动的频率或起始时机, 图形随之变形. 这些现象直接体现了三角函数的周期性与相位性质.

探索课题

我们考虑一座摩天轮. 设它的半径为 30 米, 其中心点距离地面 35 米. 你从摩天轮的最低点坐上座舱, 摩天轮以恒定的角速度开始逆时针转动, 并且每 12 分钟转一圈.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数



问题

1. 启动 2 分钟后, 你距离地面的高度是多少?
2. 启动 8 分钟后, 你距离地面的高度又是多少?

我们已知最低点高度为 $35-30=5$ 米, 最高点为 $35+30=65$ 米. 从最低点到最高点需要 $12/2=6$ 分钟. 你可能会想, 既然 6 分钟升高了 60 米, 那么 2 分钟后高度就是 $5+\frac{2}{6}\times 60=25$ 米. 请仔细观察上图. 当你在 P_1 点时, 你上升的垂直高度与沿圆弧走过的路程不成正比. 你的垂直速度在不断变化 (在最低点和最高点最慢, 在与中心等高处最快). 这说明高度随时间的变化不是线性关系, 我们需要一种新的函数模型来描述这种周而复始的非线性变化.

在此之前, 我们遇到的函数, 如一次函数、二次函数、指数函数、对数函数大多描述的是增长、衰减或简单的 V 形、U 形变化. 然而现实世界中充满着**周期性现象**, 这些循环往复的运动, 无法用旧有的函数模型来刻画.

三**角函数**正是为描述周期性现象而建立的数学语言. 本章将超越直角三角形内的边角比, 从单位圆上的坐标这一更广阔的视角重新定义三角函数, 进而系统研究其**图像特征**与**性质**, 并借助解析式 $y=A\sin(\omega x+\varphi)+k$ 分析函数的行为, 解决与周期、频率、相位相关的实际问题.

6.1 弧度制

来龙去脉

我们先用度作角的单位. 把圆周分成 360 等份, 每份对应 1°. 古巴比伦人选择 360, 因为它能被许多整数整除, 便于天文、航海和日常计算.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

360 出自历史约定, 与圆的半径、周长等几何性质无关. 角度制下许多公式因此带一个转换系数 $\frac{\pi}{180}$, 表达式不够简洁.

更自然的做法是用圆自身的量来度量角, 即取弧长与半径的比值. 这个比值就是**弧度制**.

6.1.1 定义

定义 6.1 (弧度) 在圆中, 我们规定**长度等于半径长的圆弧, 它所对的圆心角的大小为 1 弧度**. 设一段弧长为 l, 它所对的圆心角为 α(单位为弧度), 所在圆的半径为 r. 那么角 α 的大小由弧长与半径的比值唯一确定.

$$\alpha = \frac{l}{r} \tag{6-1}$$

知乎@余命数

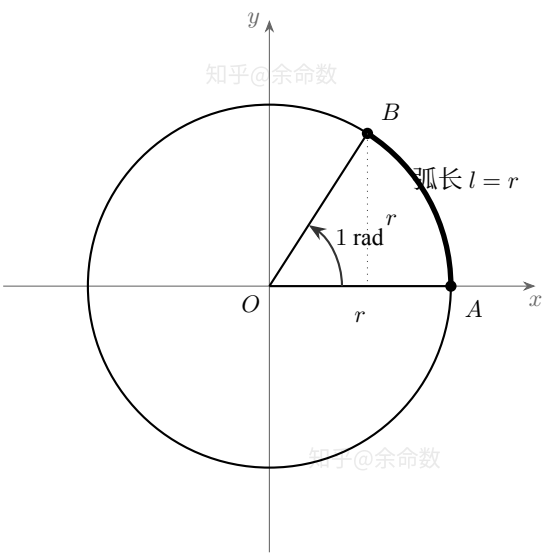


图 6.2 弧度的定义. 当弧长等于半径时, 所对应的圆心角为 1 弧度

弧度制是一个纯粹的**比值**, 没有单位(也可说单位为 1). 角的大小因此与所选圆的半径无关, 成为一个只与形状有关的纯数, 能在数学公式中与其他物理量直接、协调地参与运算.

6.1.2 角度制与弧度制的换算

我们先以周角为锚点进行换算. 一个完整圆周的弧长 $l = 2\pi r$. 由弧度制定义, 它对应的弧度角为

$$\alpha = \frac{l}{r} = \frac{2\pi r}{r} = 2\pi(\text{弧度}).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

一个周角等于 360°, 所以

$$360^\circ = 2\pi \text{ rad} \quad \Longleftrightarrow \quad 180^\circ = \pi \text{ rad} \tag{6-2}$$

由此得到任意角度的换算公式.

- **角度化弧度**.$1^\circ = \frac{\pi}{180}$ rad, 那么 $n^\circ = n \cdot \frac{\pi}{180}$ rad.
- **弧度化角度**.$1 \text{ rad} = \left(\frac{180}{\pi}\right)^\circ \approx 57.3^\circ$.

常用特殊角的换算

我们把这些特殊角的弧度值列在下表中.

角度制	0°	30°	45°	60°	90°	180°	270°	360°
弧度制	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π

从现在起, 在讨论函数、方程等数学问题时, 凡未指明单位的角度, 一**律默认为弧度制**. 例如, $\sin 2$ 指的是 $\sin(2 \text{ rad})$, 不是 $\sin(2^\circ)$. 这是高中和高等数学的标准约定.

6.2 正弦函数与余弦函数

在初等几何中, 我们最初在**直角三角形**内定义正弦、余弦.

$$\sin A = \frac{\text{对边}}{\text{斜边}}, \quad \cos A = \frac{\text{邻边}}{\text{斜边}}.$$

这个定义只适用于锐角 $A \in (0, \frac{\pi}{2})$. 它能处理直角三角形的边角计算, 却不能说明 $\sin\frac{2\pi}{3}$ 或 $\cos(-\frac{\pi}{6})$ 的意义, 更无法把正弦、余弦看成定义在 $\mathbb{R}$ 上的函数.

要描述任意大小的旋转, 例如摩天轮或行星轨道, 就要突破直角三角形的限制. **单位圆**把角 α 的终边与圆上点的坐标直接联系起来.$\sin\alpha$ 是该点的纵坐标, $\cos\alpha$ 是该点的横坐标. 这样, 定义域从 $(0, \frac{\pi}{2})$ 扩展到整个实数集 $\mathbb{R}$, 周期性也直接来自圆周运动.

知乎@余命数

6.2.1 单位圆上的参数化

定义 6.2 (任意角的三角函数) 在平面直角坐标系中, 我们构造一个以原点 O 为圆心, 半径 $r = 1$ 的圆, 称之为**单位圆**, 其方程为 $x^2 + y^2 = 1$. 设角 α 的顶点与原点重合, 始边与 x 轴的非负半轴重合. 其终边与单位圆交于唯一点 $P(x, y)$. 那么点 P 的坐标与角 α 之间的关系如下.

- 点 P 的**纵坐标** y 称为角 α 的**正弦** (Sine), 记作 $\sin\alpha = y$.
- 点 P 的**横坐标** x 称为角 α 的**余弦** (Cosine), 记作 $\cos\alpha = x$.

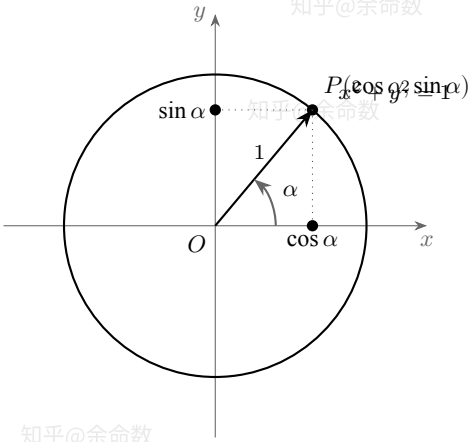


图 6.3 单位圆中三角函数的定义

由单位圆方程 $x^2 + y^2 = 1$ 可得, 对任意角 α 都有

$$\sin^2\alpha + \cos^2\alpha = 1.$$

这个等式就是同角三角函数的平方关系, 它的几何意义是单位圆上任意一点到原点的距离为 1.

欧拉公式预告

注 6.1 (欧拉公式的启示)
在复分析领域, 欧拉公式说明了一个更为精细的联系. 它给出 $e^{i\alpha} = \cos\alpha + i\sin\alpha$. 因此, 复数 $e^{i\alpha}$ 是单位圆上辐角为 α 的点. 正弦与余弦就是这个基本复指数函数在实轴与虚轴上的投影, 它统一了指数函数与三角函数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

6.2.2 图像的生成

$\sin x$ 和 $\cos x$ 已经是单位圆上动点的纵、横坐标. 把它们对应的坐标值沿角度 x 展开, 就得到一维的函数图像.

我们让质点 P 从 $(1,0)$ 出发沿单位圆逆时针运动.

- 把 P 的**纵坐标** $y=\sin x$ 投影到水平数轴上, 得到**正弦曲线**;
- 把 P 的**横坐标** $x'=\cos x$ 投影到同一条水平数轴上, 得到**余弦曲线**.

这两条曲线是最基本的波动形态.

6.2.3 性质的几何根源

正弦与余弦函数的性质都可以从单位圆直接读出.

6.2.3.1 定义域与值域

- **定义域**. 对于任意实数角 $x\in\mathbb{R}$, 其终边总能在单位圆上唯一确定一个点 P. 因此点 P 的横纵坐标 $\cos x$ 和 $\sin x$ 总是有定义的, 故 $y=\sin x$ 和 $y=\cos x$ 的定义域均为 $\mathbb{R}$.
- **值域**. 单位圆上的点, 其横坐标 x 和纵坐标 y 的取值范围都被限制在 $[-1,1]$ 之内, 故 $y=\sin x$ 和 $y=\cos x$ 的值域均为 $[-1,1]$.

6.2.3.2 周期性

当角 x 增加 2π(一个完整的圆周)时, 其终边回到原来的位置, 单位圆上的点 P 也重合. 于是点的坐标值将重复出现.

$$\sin(x+2\pi)=\sin x,\quad \cos(x+2\pi)=\cos x$$

因此, 2π 是它们的一个周期. 我们还可以验证, 这也是它们的最小正周期. 周期性是圆周运动最根本的属性.

6.2.3.3 奇偶性与对称性

我们考察角 x 与角 $-x$ 的终边. 它们在单位圆上所对应的点 $P(\cos x,\sin x)$ 与 $P'(\cos(-x),\sin(-x))$ 关于 x 轴对称.

由对称性可知, 点 P 与 P' 的横坐标相同, 纵坐标互为相反数. 据此得到

$$\cos(-x)=\cos x\quad(\text{横坐标不变})$$

$$\sin(-x)=-\sin x\quad(\text{纵坐标反号})$$

因此, $y=\cos x$ 是**偶函数**, 其图像关于 y 轴对称; $y=\sin x$ 是**奇函数**, 其图像关于原点对称.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

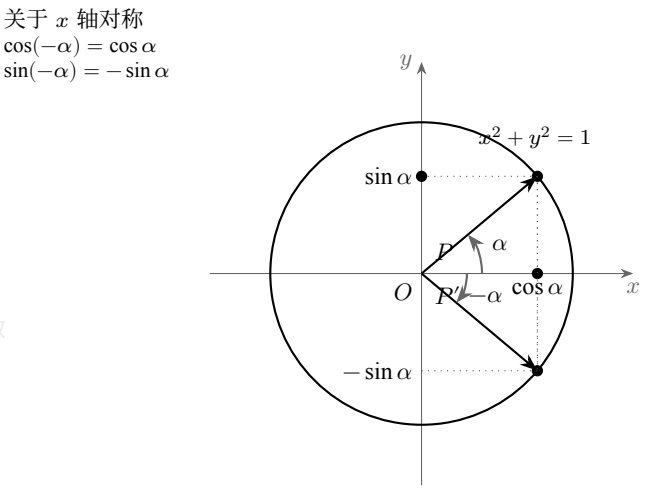
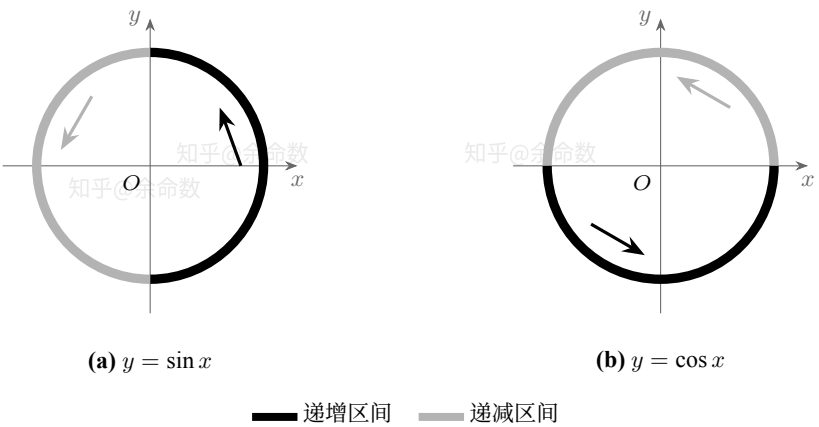


图 6.4 三角函数的奇偶性

6.2.3.4 单调性

函数的单调性完全取决于动点 P 在单位圆上运动时,其横纵坐标的变化趋势.



(a) $y=\sin x$ **(b)** $y=\cos x$

图 6.5 三角函数在单位圆上的单调性

- **对于** $y=\sin x$**(考察纵坐标)**. 当角 x 从 $-\frac{\pi}{2}$ 运动至 $\frac{\pi}{2}$ 时, 动点 P 在单位圆的右半圆上从最低点运动至最高点, 其纵坐标 y 从 -1 单调地**增加**至 1. 当角 x 从 $\frac{\pi}{2}$ 运动至 $\frac{3\pi}{2}$ 时, 动点 P 在单位圆的左半圆上从最高点运动至最低点, 其纵坐标 y 从 1 单调地**减少**至 -1.
- **对于** $y=\cos x$**(考察横坐标)**. 当角 x 从 0 运动至 π 时, 动点 P 在单位圆的上半圆上从最右点运动至最左点, 其横坐标 x' 从 1 单调地**减少**至 -1. 当角 x 从 π 运动至 2π 时, 动点 P 在单位圆的下半圆上从最左点运动至最右点, 其横坐标 x' 从 -1 单调地**增加**至 1.

我们将这些观察与周期性结合, 就能得到它们在整个定义域上的所有单调区间.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

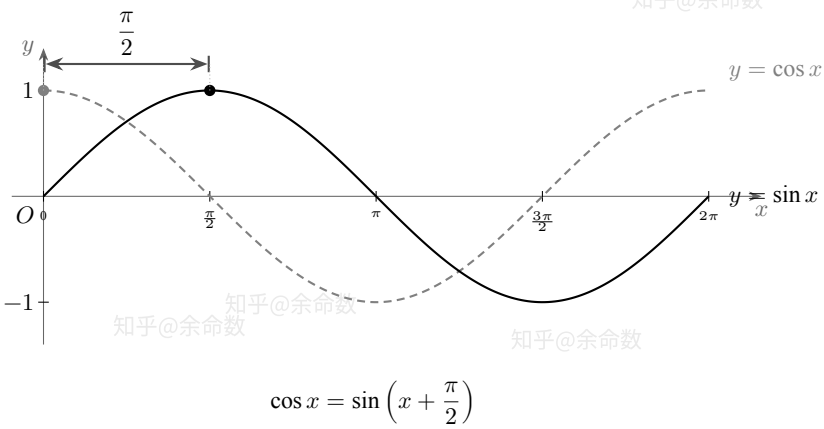
6.2.4 正弦与余弦的内在联系

正弦与余弦的图像形状完全相同，只是位置错开. 下面我们用一个诱导公式说明这种关系.

定理 6.1 (诱导公式) 对于任意角 x, 恒有

$$\cos x = \sin\left(x + \frac{\pi}{2}\right)$$

证明 我们用几何方法证明. 在单位圆上，将角 x 的终边逆时针旋转 $\frac{\pi}{2}$，得到角 $x + \frac{\pi}{2}$ 的终边. 设角 x 终边上的点为 $P(\cos x, \sin x)$，则角 $x + \frac{\pi}{2}$ 终边上的点为 $P'(-\sin x, \cos x)$. 根据定义, $\sin(x + \frac{\pi}{2})$ 就是点 P' 的纵坐标, 即 $\cos x$. □



$$\cos x = \sin\left(x + \frac{\pi}{2}\right)$$

图 6.6 正弦函数与余弦函数的相位关系.$\cos x$ 是 $\sin x$ 向左平移 $\frac{\pi}{2}$ 个单位

余弦函数可以由正弦函数向左平移 $\frac{\pi}{2}$ 得到, 两者描述的是同一类波动, 只是相位不同. 余弦比正弦超前 $\frac{\pi}{2}$, 正弦比余弦滞后 $\frac{\pi}{2}$. 它们都是简谐运动的基础.

6.3 正弦与余弦函数的图像

从单位圆到函数图像

当角 x 从 0 变到 2π 时, 单位圆上的点 $P(\cos x, \sin x)$ 刚好转一周. 在这个过程中, P 的纵坐标 $y = \sin x$ 和横坐标 $x' = \cos x$ 都在周期性变化.

我们把这种圆周运动画成函数图像. 让单位圆沿 x 轴正方向滚动, 同时把 P 的纵坐标投影到对应的 x 轴位置上. 投影点描出的轨迹就是正弦函数的图像.

这条曲线称为**正弦曲线**, 是最基本的波动形态.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

6.3.1 五点作图法

画一条曲线需要很多点. 先标出一个周期内的最大、最小和零点, 就能快速勾出轮廓. 对 $y=\sin x$ 在 $[0,2\pi]$ 上, 这五个关键点是

- **起点(零值点)**.$(0,0)$
- **最高点**.$(\frac{\pi}{2},1)$
- **中间点(零值点)**.$(\pi,0)$
- **最低点**.$(\frac{3\pi}{2},-1)$
- **终点(零值点)**.$(2\pi,0)$

用光滑曲线连接这五点, 就得到一个周期的正弦图像. 由于函数以 2π 为周期, 把这个波形向左右重复延伸, 即得完整图像.

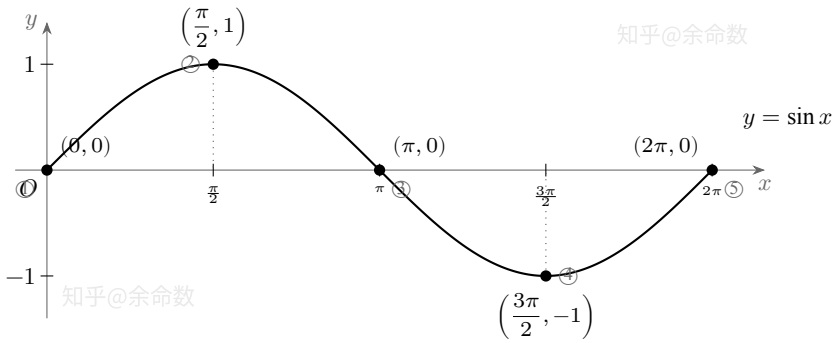


图 6.7 正弦函数 $y=\sin x$ 在 $[0,2\pi]$ 上的图象及五个关键点

6.3.2 余弦函数的图像

同样地, 把单位圆上点 $P(\cos x,\sin x)$ 的**横坐标** $\cos x$ 沿角度展开, 就得到**余弦曲线** $y=\cos x$. 它在 $[0,2\pi]$ 上的五个关键点是 $(0,1),(\frac{\pi}{2},0),(\pi,-1),(\frac{3\pi}{2},0),(2\pi,1)$.

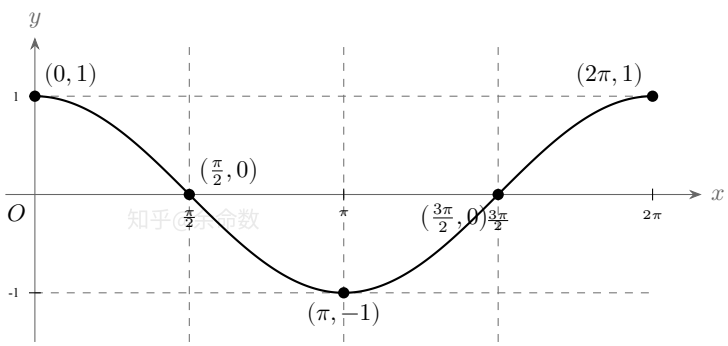


图 6.8 函数 $y=\cos x$ 在 $[0,2\pi]$ 上的五点法作图

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

6.4 正弦函数与余弦函数的性质

有了正弦、余弦函数的图像，我们便可以研究它们的性质. 这些性质包括定义域、值域、周期性、奇偶性、单调性和对称性. 我们还将用它们来理解 $y = A\sin(\omega x + \varphi) + B$.

6.4.1 定义域、值域与周期性

从单位圆的定义看，角 x 旋转一周（2π 弧度），终边上的点 P 回到原位，其坐标 $(\cos x, \sin x)$ 也重复出现. 这正是三角函数周期性的来源. 我们有 $\sin(x+2\pi) = \sin x$，$\cos(x + 2\pi) = \cos x$. 因此 2π 是它们的一个周期，也是最小正周期.

- **定义域**. 从单位圆定义可知，任意实数角 α 都有唯一确定的正弦值和余弦值与之对应. 因此，函数 $y = \sin x$ 和 $y = \cos x$ 的定义域均为实数集 $\mathbb{R}$.
- **值域**. 在单位圆 $x^2 + y^2 = 1$ 上，点的横、纵坐标的取值范围都是 $[-1, 1]$. 因此，函数 $y = \sin x$ 和 $y = \cos x$ 的值域均为 $[-1, 1]$.
  - 当 $x = \frac{\pi}{2} + 2k\pi, k \in \mathbb{Z}$ 时，$y = \sin x$ 取得最大值 1.
  - 当 $x = \frac{3\pi}{2} + 2k\pi, k \in \mathbb{Z}$ 时，$y = \sin x$ 取得最小值 -1.
  - 当 $x = 2k\pi, k \in \mathbb{Z}$ 时，$y = \cos x$ 取得最大值 1.
  - 当 $x = \pi + 2k\pi, k \in \mathbb{Z}$ 时，$y = \cos x$ 取得最小值 -1.
- **周期性**. 正弦函数和余弦函数都是周期函数，它们的最小正周期都是 2π. 我们有 $\sin(x + 2k\pi) = \sin x$，$\cos(x + 2k\pi) = \cos x$，其中 $k \in \mathbb{Z}$.

6.4.2 奇偶性与单调性

- **奇偶性**. 在单位圆上，角 x 与 $-x$ 的终边关于 x 轴对称. 设角 x 终边上点为 $P(\cos x, \sin x)$，则角 $-x$ 终边上点为 $P'(\cos x, -\sin x)$. 根据定义，我们得到

$$\cos(-x) = \cos x$$

$$\sin(-x) = -\sin x$$

  因此
  - $y = \cos x$ 是**偶函数**，其图像关于 y 轴对称.
  - $y = \sin x$ 是**奇函数**，其图像关于原点对称.
- **单调性**. 我们先在一个周期内观察单调性，再由周期性推广到整个定义域.
  - 对于函数 $y = \sin x$：
    - **单调递增区间**为 $\left[-\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + 2k\pi\right], k \in \mathbb{Z}$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

- **单调递减区间**为 $\left[\frac{\pi}{2}+2k\pi,\frac{3\pi}{2}+2k\pi\right],k\in\mathbb{Z}$.

- 对于函数 $y=\cos x$：
  - **单调递增区间**为 $\left[-\pi+2k\pi,2k\pi\right],k\in\mathbb{Z}$.
  - **单调递减区间**为 $\left[2k\pi,\pi+2k\pi\right],k\in\mathbb{Z}$.

记单调区间时，我们可以在单位圆上观察点的运动. 对于 $\sin x$，我们看纵坐标 y. 从 -90° $(-\pi/2)$ 转到 90° $(\pi/2)$，点在圆的右半部分，纵坐标 y 从 -1 **增**至 1；从 90° 转到 270° $(3\pi/2)$，点在圆的左半部分，纵坐标 y 从 1 **减**至 -1. 对于 $\cos x$，我们看横坐标 x. 从 180° (π) 转到 360°（2π 或 0），点在圆的下半部分移动到右侧，横坐标 x 从 -1 **增**至 1；从 0° 转到 180°，点在圆的上半部分，横坐标 x 从 1 **减**至 -1.

6.4.3 对称性

我们讨论对称轴和对称中心.

- **函数** $y=\sin x$.
  - **对称轴**. 图像的最高点和最低点所在的竖直线，即直线 $x=k\pi+\frac{\pi}{2},k\in\mathbb{Z}$.
  - **对称中心**. 图像与 x 轴的交点，即点 $(k\pi,0),k\in\mathbb{Z}$.
- **函数** $y=\cos x$.
  - **对称轴**. 图像的最高点和最低点所在的竖直线，即直线 $x=k\pi,k\in\mathbb{Z}$.（这包含了作为偶函数对称轴的 y 轴 $x=0$）
  - **对称中心**. 图像与 x 轴的交点，即点 $\left(k\pi+\frac{\pi}{2},0\right),k\in\mathbb{Z}$.

对称轴经过函数取到**最值**的位置；对称中心是函数图像与 x 轴的交点，也是图像的拐点. 若函数满足 $f(a+x)=f(a-x)$，则其图像关于直线 $x=a$ 对称；若满足 $f(a+x)=-f(a-x)$（或 $f(x)+f(2a-x)=0$），则图像关于点 $(a,0)$ 中心对称.

6.4.4 性质汇总

我们将正弦函数与余弦函数的性质汇总成表.

知乎@余命数
知乎@余命数
知乎@余命数

知乎@余命数

表 6.1　正弦函数与余弦函数性质对比

性质	$y=\sin x$	$y=\cos x$
定义域	$\mathbb{R}$	$\mathbb{R}$
值域	$[-1,1]$	$[-1,1]$
最小正周期	2π	2π
奇偶性	奇函数	偶函数
单调递增区间	$[2k\pi-\frac{\pi}{2},2k\pi+\frac{\pi}{2}]$	$[(2k-1)\pi,2k\pi]$
单调递减区间	$[2k\pi+\frac{\pi}{2},2k\pi+\frac{3\pi}{2}]$	$[2k\pi,(2k+1)\pi]$
对称轴	直线 $x=k\pi+\frac{\pi}{2}$	直线 $x=k\pi$
对称中心	点 $(k\pi,0)$	点 $(k\pi+\frac{\pi}{2},0)$

表中所有 $k\in\mathbb{Z}$

知乎@余命数　知乎@余命数

6.4.5 高考真题选讲

例题 6.1 (2012 年福建卷文科数学第 8 题) 函数 $f(x)=\sin\left(x-\frac{\pi}{4}\right)$ 的图象的一条对称轴是

A. $x=\frac{\pi}{4}$　　C. $x=-\frac{\pi}{4}$

B. $x=\frac{\pi}{2}$　　D. $x=-\frac{\pi}{2}$

解 答案：C.

令 $x-\frac{\pi}{4}=k\pi+\frac{\pi}{2},k\in\mathbb{Z}$，得对称轴方程 $x=k\pi+\frac{3\pi}{4}$.

当 $k=-1$ 时，$x=-\frac{\pi}{4}$，对应选项 C.

例题 6.2 (2019 年北京卷理科数学第 9 题) 函数 $f(x)=\sin^2 2x$ 的最小正周期是_____.

解 答案：$\frac{\pi}{2}$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

用降幂公式 $\sin^2 2x=\dfrac{1-\cos 4x}{2}$, 故

$$f(x)=\frac{1}{2}-\frac{1}{2}\cos 4x,$$

周期 $T=\dfrac{2\pi}{4}=\dfrac{\pi}{2}$.

例题 6.3 (2017 年全国 □ 卷理科数学第 6 题) 设函数 $f(x)=\cos\left(x+\dfrac{\pi}{3}\right)$, 则下列结论错误的是

A. $f(x)$ 的一个周期为 -2π

B. $y=f(x)$ 的图象关于直线 $x=\dfrac{8\pi}{3}$ 对称

C. $f(x+\pi)$ 的一个零点为 $x=\dfrac{\pi}{6}$

D. $f(x)$ 在 $\left(\dfrac{\pi}{2},\pi\right)$ 单调递减

知乎@余命数

解 答案: D.

A: cos 的周期为 $2k\pi$, $k=-1$ 时 $T=-2\pi$, 正确.

B: $f\left(\dfrac{8\pi}{3}\right)=\cos\left(\dfrac{8\pi}{3}+\dfrac{\pi}{3}\right)=\cos 3\pi=-1$, 取最小值, 故 $x=\dfrac{8\pi}{3}$ 是对称轴, 正确.

C: $f\left(\dfrac{\pi}{6}+\pi\right)=\cos\left(\dfrac{\pi}{6}+\pi+\dfrac{\pi}{3}\right)=\cos\dfrac{3\pi}{2}=0$, 正确.

D: 当 $x\in\left(\dfrac{\pi}{2},\pi\right)$ 时, $x+\dfrac{\pi}{3}\in\left(\dfrac{5\pi}{6},\dfrac{4\pi}{3}\right)$, cos 在该区间内先减后增(在 $x+\dfrac{\pi}{3}=\pi$ 处取最小值), 并非单调递减, 故 D 错误.

6.5 正切函数

正弦、 余弦对应单位圆上点的坐标 (x,y). 角 α 确定了终边上一点 $P(\cos\alpha,\sin\alpha)$. 这条终边 OP 是一条射线, 除点的坐标外还有一个几何量, 即**斜率**. 根据斜率的定义, $k_{OP}=\dfrac{\Delta y}{\Delta x}=\dfrac{\sin\alpha-0}{\cos\alpha-0}=\dfrac{\sin\alpha}{\cos\alpha}$ (当 $\cos\alpha\neq 0$ 时). 这个比值就是角 α 的**正切**.

正切的几何意义可以直接画出来. 过单位圆右侧点 $(1,0)$ 作 x 轴的垂线, 这条直线与圆**相切**. 延长角 α 的终边, 与这条切线交于点 T. 由相似三角形可知, 点 T 的纵坐标是 $\dfrac{\sin\alpha}{\cos\alpha}$. 所以正切值就是**单位圆切线上的一段有向线段的长度**.

知乎@余命数

6.5.1 定义与图像

定义 6.3 (正切) 我们定义角 α 的**正切**为

$$\tan\alpha = \frac{\sin\alpha}{\cos\alpha} \tag{6-3}$$

它的定义域是所有使得分母 $\cos\alpha \neq 0$ 的角 α 的集合,也就是 $\alpha \neq k\pi + \frac{\pi}{2}, k \in \mathbb{Z}$.

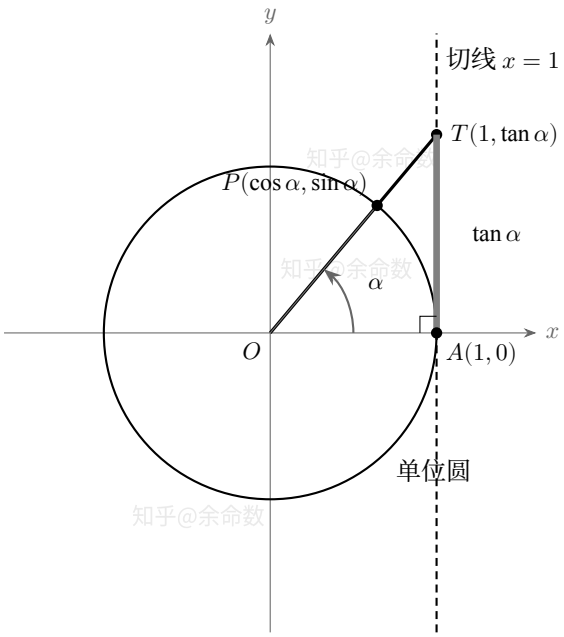


图 6.9　正切函数的几何意义

正切函数的图像与正弦、余弦曲线有明显的不同. 由于定义域中排除了使 $\cos x = 0$ 的点, 函数图像在这些点上是**不连续**的. 这些不连续点 $x = k\pi + \frac{\pi}{2}$ 形成了函数图像的一条条**垂直渐近线**.

在一个周期 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 内,取几个关键点即可描绘其形态. $(-\frac{\pi}{4}, -1)$、$(0,0)$、$(\frac{\pi}{4}, 1)$. 图像从左侧贴近渐近线 $x = -\frac{\pi}{2}$ 的负无穷处开始, 穿过这三点, 最终向上贴近渐近线 $x = \frac{\pi}{2}$ 的正无穷处.

6.5.2 正切函数的性质

观察正切函数的定义与图像, 我们可总结出它的主要性质.

- **定义域**.$\{x \in \mathbb{R} \mid x \neq k\pi + \frac{\pi}{2}, k \in \mathbb{Z}\}$.
- **值域**. 实数集 $\mathbb{R}$. 正切函数可以取到任意实数值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

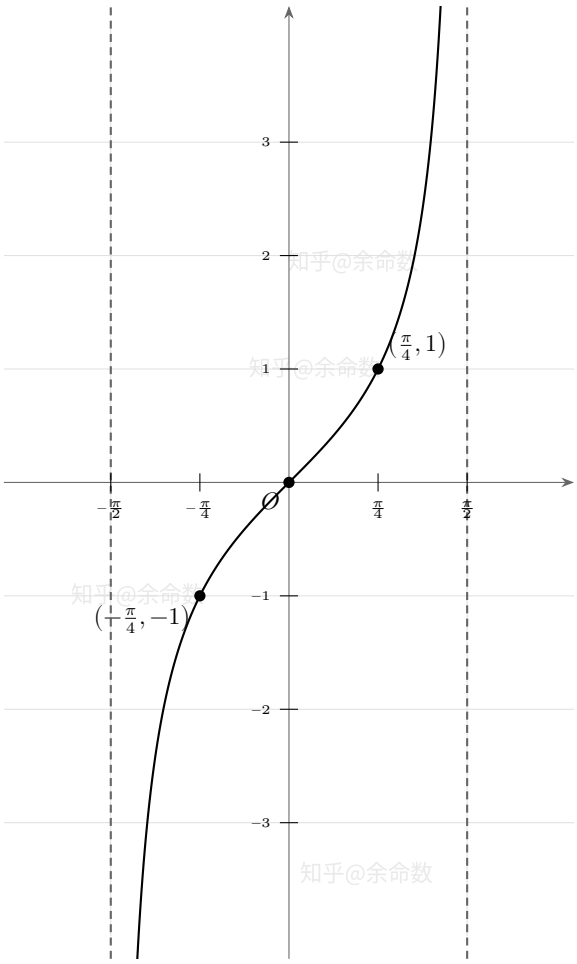


图 6.10 函数 $y = \tan x$ 的图像

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **周期性**. 观察图像, 其形态每隔 π 的长度就重复一次. 可以验证

$$\tan(x+\pi)=\frac{\sin(x+\pi)}{\cos(x+\pi)}=\frac{-\sin x}{-\cos x}=\tan x$$

所以正切函数是周期函数, 其**最小正周期为** π.

- **奇偶性**. 根据定义

$$\tan(-x)=\frac{\sin(-x)}{\cos(-x)}=\frac{-\sin x}{\cos x}=-\tan x$$

所以正切函数是**奇函数**, 其图像关于原点 $(0,0)$ 对称. 更一般地, 图像关于所有零点 $(k\pi,0),k\in\mathbb{Z}$ 中心对称.

- **单调性**. 正切函数在它的每一个连续区间上都是**单调递增**的. 其单调递增区间为

$$\left(k\pi-\frac{\pi}{2},k\pi+\frac{\pi}{2}\right),k\in\mathbb{Z}$$

常见的误解是以为 $y=\tan x$ 在整个定义域上单调递增. 其实它在每个连续区间 $(-\frac{\pi}{2},\frac{\pi}{2})$、$(\frac{\pi}{2},\frac{3\pi}{2})$ 等内部都递增, 但不能跨区间说整体递增. 例如取 $x_1=\frac{\pi}{4}$ 和 $x_2=\frac{3\pi}{4}$, 有 $x_1<x_2$, 而 $\tan x_1=1,\tan x_2=-1$, 所以 $\tan x_1>\tan x_2$, 这与整体单调递增矛盾.

6.5.3 高考真题选讲

例题 6.4 (2020 年全国 □ 卷理科数学第 9 题) 已知 $2\tan\theta-\tan\left(\theta+\frac{\pi}{4}\right)=7$, 则 $\tan\theta=$

A. -2　　C. 1

B. -1　　D. 2

解 答案: D.

用正切和角公式 $\tan\left(\theta+\frac{\pi}{4}\right)=\frac{\tan\theta+1}{1-\tan\theta}$ 代入, 得

$$2\tan\theta-\frac{\tan\theta+1}{1-\tan\theta}=7.$$

两边乘以 $1-\tan\theta$ 并整理, 得

$$2\tan\theta-2\tan^2\theta-\tan\theta-1=7-7\tan\theta,$$

即 $2\tan^2\theta-8\tan\theta+8=0$, 亦即 $(\tan\theta-2)^2=0$, 故 $\tan\theta=2$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 6.5 (2020 年浙江卷第 13 题) 已知 $\tan\theta = 2$, 则 $\cos 2\theta =$_____, $\tan\left(\theta - \frac{\pi}{4}\right) =$_____.

解 答案: $-\frac{3}{5}, \frac{1}{3}$.

用万能公式求 $\cos 2\theta$:

$$\cos 2\theta = \frac{1-\tan^2\theta}{1+\tan^2\theta} = \frac{1-4}{1+4} = -\frac{3}{5}.$$

用正切差角公式求 $\tan\left(\theta - \frac{\pi}{4}\right)$:

$$\tan\left(\theta - \frac{\pi}{4}\right) = \frac{\tan\theta - 1}{1+\tan\theta} = \frac{2-1}{1+2} = \frac{1}{3}.$$

6.6 余切、正割与余割函数

我们已经在单位圆上用点 (x, y) 的坐标定义了 $\sin\alpha, \cos\alpha$, 用终边斜率定义了 $\tan\alpha$. 现在我们把这三个函数分别取倒数, 得到另外三个函数.

- **余切 (Cotangent)** $\cot\alpha = \dfrac{1}{\tan\alpha} = \dfrac{x}{y}$.
- **正割 (Secant)** $\sec\alpha = \dfrac{1}{\cos\alpha} = \dfrac{1}{x}$.
- **余割 (Cosecant)** $\csc\alpha = \dfrac{1}{\sin\alpha} = \dfrac{1}{y}$.

这三个倒数函数在历史上曾被用来简化除法运算; 在现代数学中, 它们在积分、微分方程和某些物理问题里仍然出现.

6.6.1 定义与几何意义

定义 6.4 (余切、正割、余割) 对于任意角 α, 我们定义.

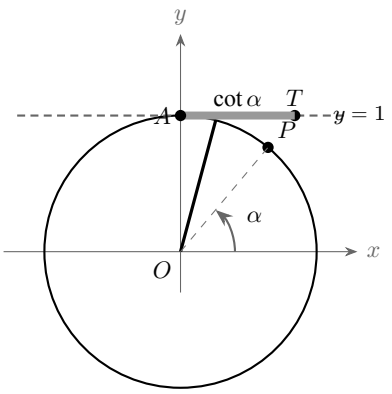
- **余切** $\cot\alpha = \dfrac{\cos\alpha}{\sin\alpha}$, 定义域为 $\{\alpha \in \mathbb{R} \mid \sin\alpha \neq 0\}$.
- **正割** $\sec\alpha = \dfrac{1}{\cos\alpha}$, 定义域为 $\{\alpha \in \mathbb{R} \mid \cos\alpha \neq 0\}$.
- **余割** $\csc\alpha = \dfrac{1}{\sin\alpha}$, 定义域为 $\{\alpha \in \mathbb{R} \mid \sin\alpha \neq 0\}$.

这三个函数同样可以在单位圆上用切线段的长度来解释.

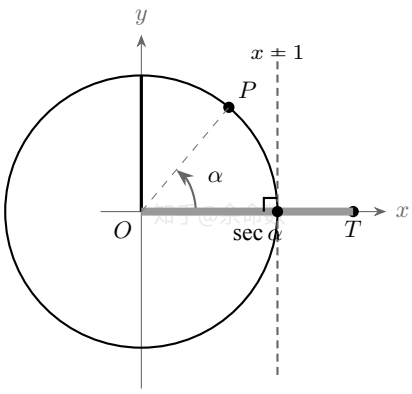
- 余切 $\cot\alpha$ 的几何意义如下. 作单位圆过点 $(0, 1)$ 的切线 $y = 1$. 角 α 的终边或其反向延长线与该切线交于点 T_c. 点 T_c 的横坐标即为 $\cot\alpha$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

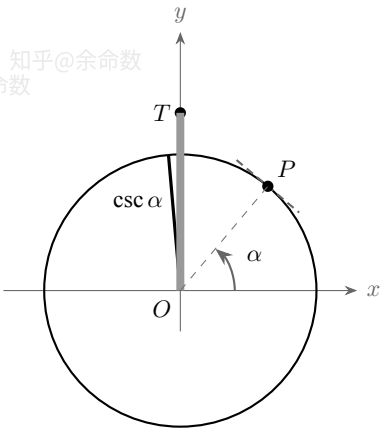
知乎@余命数



(a) 余切.$\cot\alpha = \overline{AT}$



(b) 正割.$\sec\alpha = \overline{OT}$



(c) 余割.$\csc\alpha = \overline{OT}$

图 6.11　余切、正割、余割的几何意义

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 正割 $\sec\alpha$ 与余割 $\csc\alpha$ 的几何意义如下. 作单位圆过终边上点 $P(\cos\alpha,\sin\alpha)$ 的切线. 该切线与 x 轴的交点为 T_s, 与 y 轴的交点为 T_{cs}. 原点到这两个交点的有向距离 OT_s 和 OT_{cs} 分别为 $\sec\alpha$ 和 $\csc\alpha$.

知乎@余命数

6.6.2 函数图像与性质

这三个函数的图像与性质由倒数定义直接决定. 凡是使分母为零的点, 都是图像的垂直渐近线.

6.6.2.1 余切函数 $y=\cot x$

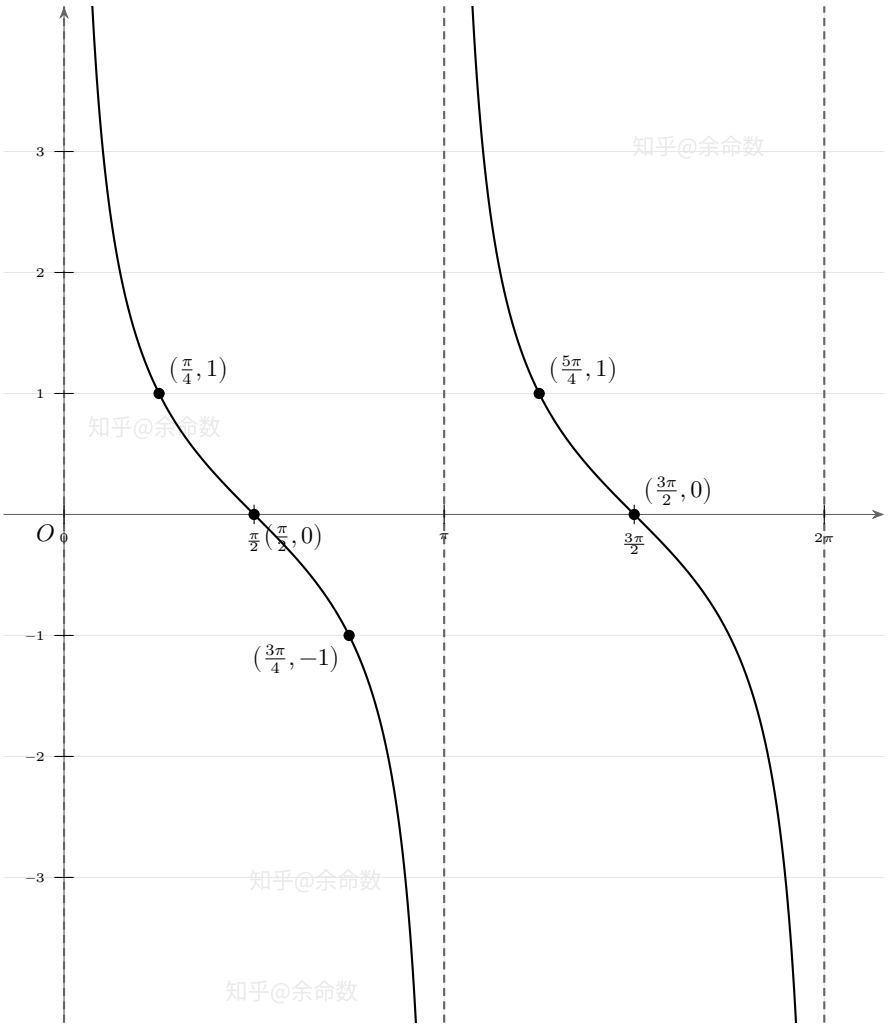


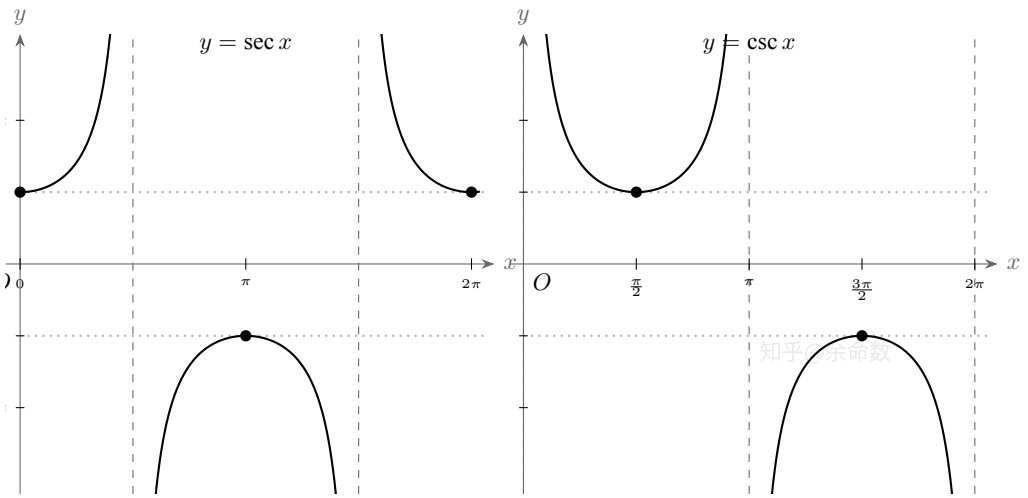
图 6.12 函数 $y=\cot x$ 的图像

- **定义域** $\{x\in\mathbb{R}\mid x\neq k\pi,k\in\mathbb{Z}\}$.
- **值域** $\mathbb{R}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **周期性** 最小正周期为 π.
- **奇偶性** $\cot(-x) = -\cot x$, 是奇函数, 图像关于原点对称.
- **单调性** 在每个定义区间 $(k\pi, (k+1)\pi)$ 内单调递减.

6.6.2.2 正割函数 $y = \sec x$ 与余割函数 $y = \csc x$



(a) 正割函数　　(b) 余割函数

图 6.13　正割函数和余割函数的图像

- **定义域** $\sec x$ 为 $\{x \mid x \neq k\pi + \frac{\pi}{2}\}$; $\csc x$ 为 $\{x \mid x \neq k\pi\}$.
- **值域**均为 $(-\infty, -1] \cup [1, \infty)$.
- **周期性**最小正周期均为 2π.
- **奇偶性** $y = \sec x$ 是偶函数; $y = \csc x$ 是奇函数.

6.6.3 性质汇总

我们将六个三角函数的性质汇总成表.

知乎@余命数

知乎@余命数

表 6.2 六种三角函数重要性质对比

函数	定义域	值域	最小正周期	奇偶性
$y=\sin x$	$\mathbb{R}$	$[-1,1]$	2π	奇函数
$y=\cos x$	$\mathbb{R}$	$[-1,1]$	2π	偶函数
$y=\tan x$	$\{x \mid x \neq k\pi+\frac{\pi}{2}\}$	$\mathbb{R}$	π	奇函数
$y=\cot x$	$\{x \mid x \neq k\pi\}$	$\mathbb{R}$	π	奇函数
$y=\sec x$	$\{x \mid x \neq k\pi+\frac{\pi}{2}\}$	$(-\infty,-1]\cup[1,\infty)$	2π	偶函数
$y=\csc x$	$\{x \mid x \neq k\pi\}$	$(-\infty,-1]\cup[1,\infty)$	2π	奇函数

注. 表中所有 $k\in\mathbb{Z}$

知乎@余命数

6.7 同角三角函数的基本关系

我们已经用单位圆上点的坐标定义了 $\sin\alpha,\cos\alpha$, 用终边斜率定义了 $\tan\alpha$. 它们来自同一个角, 因此有固定的联系. 三角计算与恒等变形从这些关系出发.

6.7.1 两个基本关系式

这些关系可以直接从定义得到.

- **第一条, 单位圆方程**.
  点 $P(x,y)$ 在单位圆 $x^2+y^2=1$ 上, 而 $x=\cos\alpha, y=\sin\alpha$. 将坐标代入圆的方程, 得到对任意角 α 都成立的平方关系.
- **第二条, 正切的定义**.
  我们定义了 $\tan\alpha=\dfrac{\sin\alpha}{\cos\alpha}$. 这本身就是一个关系式, 我们称之为商数关系.

定理 6.2 (同角三角函数基本关系) 对于任意角 α, 以下关系式恒成立.

1. **平方关系**

$$\sin^2\alpha+\cos^2\alpha=1 \tag{6-4}$$

2. **商数关系**

$$\tan\alpha=\frac{\sin\alpha}{\cos\alpha}\quad\left(\alpha\neq k\pi+\frac{\pi}{2},k\in\mathbb{Z}\right) \tag{6-5}$$

证明 我们用定义直接证明.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

设角 α 的终边上有一点 $P(x,y)$(非原点),它到原点 O 的距离为 $r=\sqrt{x^2+y^2}$. 则

$$\sin\alpha=\frac{y}{r},\quad \cos\alpha=\frac{x}{r},\quad \tan\alpha=\frac{y}{x}\quad (x\neq 0)$$

关于平方关系,直接将定义代入等式左侧计算,得

$$\sin^2\alpha+\cos^2\alpha=\left(\frac{y}{r}\right)^2+\left(\frac{x}{r}\right)^2=\frac{y^2}{r^2}+\frac{x^2}{r^2}=\frac{x^2+y^2}{r^2}$$

因为 $x^2+y^2=r^2$,所以上式可以继续化简,得

$$\frac{x^2+y^2}{r^2}=\frac{r^2}{r^2}=1$$

这就证明了 $\sin^2\alpha+\cos^2\alpha=1$ 恒成立.

关于商数关系,同样从定义出发. 在定义域要求 $\cos\alpha\neq 0$,即 $x\neq 0$ 的前提下,

$$\frac{\sin\alpha}{\cos\alpha}=\frac{y/r}{x/r}=\frac{y}{x}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

而 $\frac{y}{x}$ 是 $\tan\alpha$ 的定义. 故商数关系 $\tan\alpha=\frac{\sin\alpha}{\cos\alpha}$ 成立. 该关系式的成立条件 $\cos\alpha\neq 0$,意味着角 α 的终边不能落在 y 轴上,即 $\alpha\neq k\pi+\frac{\pi}{2},k\in\mathbb{Z}$. □

平方关系和商数关系是所有三角恒等变形的代数起点. 更复杂的三角恒等式都可以通过对它们作代数运算得到.

6.7.2 公式的变形与应用

知乎@余命数

这两个公式形式简单,但变形很多.

知乎@余命数

- **知一求二**. 由平方关系开方,得

$$\sin\alpha=\pm\sqrt{1-\cos^2\alpha},\qquad \cos\alpha=\pm\sqrt{1-\sin^2\alpha}.$$

知乎@余命数

正负号由角 α 所在象限唯一确定. 开方符号是这一步最容易出错的地方.

- **和差与积的转化**. 我们可以用下面的恒等式把 $\sin\alpha\pm\cos\alpha$ 与 $\sin\alpha\cos\alpha$ 互相转化.

$$(\sin\alpha\pm\cos\alpha)^2=\sin^2\alpha+\cos^2\alpha\pm 2\sin\alpha\cos\alpha=1\pm 2\sin\alpha\cos\alpha$$

知乎@余命数

知道和或差,就能求出积;知道积,也能求出和或差的平方.

- **切化弦与弦化切**. 商数关系把正切与正、余弦联系在一起.
  - 遇到正切,可以统一化为正余弦计算,即切化弦.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

- 遇到关于 $\sin\alpha$ 和 $\cos\alpha$ 的齐次式(分子分母各项的幂次相同), 可以分子分母同除以 $\cos^k\alpha$, 其中 k 为幂次, 从而转化为关于 $\tan\alpha$ 的式子, 即弦化切.

例题 6.6 (知一求二) 已知角 α 的终边在第三象限, 且 $\cos\alpha=-\frac{4}{5}$, 求 $\sin\alpha$ 和 $\tan\alpha$ 的值.

解 我们先用平方关系求出 $\sin\alpha$ 的绝对值, 再由象限确定符号, 最后用商数关系求 $\tan\alpha$.

将 $\cos\alpha=-\frac{4}{5}$ 代入平方关系, 得

$$\sin^2\alpha=1-\cos^2\alpha=1-\left(-\frac{4}{5}\right)^2=1-\frac{16}{25}=\frac{9}{25},$$

故 $\sin\alpha=\pm\frac{3}{5}$. 角 α 的终边在第三象限, 其正弦值为负, 因而 $\sin\alpha=-\frac{3}{5}$. 再由商数关系得

$$\tan\alpha=\frac{\sin\alpha}{\cos\alpha}=\frac{-3/5}{-4/5}=\frac{3}{4}.$$

开方符号由象限决定. 忽略第三象限这个条件, $\sin\alpha$ 的符号就无法确定, 答案也会出错. 任何涉及开方的三角运算, 都应先检查象限.

例题 6.7 (和差积关系的运用) 已知角 $\alpha\in(\frac{\pi}{2},\pi)$, 且 $\sin\alpha+\cos\alpha=\frac{1}{5}$, 求下列各式的值.

(1) $\sin\alpha\cos\alpha$ (2) $\sin\alpha-\cos\alpha$ (3) $\tan\alpha$

解 我们可以对已知等式两边平方, 利用完全平方公式 $(a+b)^2=a^2+b^2+2ab$ 把和转化为积.

(1) 将已知等式 $\sin\alpha+\cos\alpha=\frac{1}{5}$ 两边平方, 并以平方关系 $\sin^2\alpha+\cos^2\alpha=1$ 代入, 得

$$\begin{aligned}\frac{1}{25}&=(\sin\alpha+\cos\alpha)^2=\sin^2\alpha+\cos^2\alpha+2\sin\alpha\cos\alpha\\&=1+2\sin\alpha\cos\alpha,\end{aligned}$$

移项解出

$$\sin\alpha\cos\alpha=-\frac{12}{25}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2) 直接求 $\sin\alpha-\cos\alpha$ 不便, 我们改求它的平方. 由上一步结果,

$$\begin{aligned}(\sin\alpha-\cos\alpha)^2&=\sin^2\alpha+\cos^2\alpha-2\sin\alpha\cos\alpha\\&=1-2\left(-\frac{12}{25}\right)=1+\frac{24}{25}=\frac{49}{25},\end{aligned}$$

开方得 $\sin\alpha-\cos\alpha=\pm\frac{7}{5}$. 由 $\alpha\in(\frac{\pi}{2},\pi)$ 知 $\sin\alpha>0$ 且 $\cos\alpha<0$, 故 $\sin\alpha-\cos\alpha$ 为正, 即 $\sin\alpha-\cos\alpha=\frac{7}{5}$.

(3) 联立 $\sin\alpha+\cos\alpha=\frac{1}{5}$ 与 $\sin\alpha-\cos\alpha=\frac{7}{5}$, 两式相加得 $\sin\alpha=\frac{4}{5}$, 相减得 $\cos\alpha=-\frac{3}{5}$, 这与 α 在第二象限的符号一致. 由商数关系,

$$\tan\alpha=\frac{\sin\alpha}{\cos\alpha}=\frac{4/5}{-3/5}=-\frac{4}{3}.$$

平方把和差变成积;开方从积得到和差;象限条件决定开方符号.

知乎@余命数

例题 6.8 (弦化切齐次式处理) 已知 $\tan\alpha=-3$, 求 $\dfrac{2\sin\alpha-3\cos\alpha}{4\sin\alpha+5\cos\alpha}$.

解 所求表达式的分子、分母都是关于 $\sin\alpha$ 和 $\cos\alpha$ 的一次齐次式(每一项的幂次均为 1), 因此可以分子分母同除以 $\cos\alpha$, 把整个式子转化为关于 $\tan\alpha$ 的表达式.

当 $\tan\alpha=-3$ 有意义时, 必有 $\cos\alpha\neq0$, 因此可在分子分母同除以 $\cos\alpha$.

$$\begin{aligned}\text{原式}&=\frac{\frac{2\sin\alpha-3\cos\alpha}{\cos\alpha}}{\frac{4\sin\alpha+5\cos\alpha}{\cos\alpha}}\\&=\frac{2\frac{\sin\alpha}{\cos\alpha}-3\frac{\cos\alpha}{\cos\alpha}}{4\frac{\sin\alpha}{\cos\alpha}+5\frac{\cos\alpha}{\cos\alpha}}\\&=\frac{2\tan\alpha-3}{4\tan\alpha+5}\end{aligned}$$

知乎@余命数

将 $\tan\alpha=-3$ 代入上式, 得

$$\text{原式}=\frac{2(-3)-3}{4(-3)+5}=\frac{-6-3}{-12+5}=\frac{-9}{-7}=\frac{9}{7}.$$

知乎@余命数

弦化切用于关于 $\sin\alpha,\cos\alpha$ 的齐次分式. 通过分子分母同除以 $\cos\alpha$ 的适当幂次, 把两个变量化为一个变量 $\tan\alpha$. 二次齐次式则同除以 $\cos^2\alpha$.

知乎@余命数

例题 6.9 (切化弦由正切求正余弦) 已知 $\tan\alpha=2$, 求 $\dfrac{1}{\sin^2\alpha-\sin\alpha\cos\alpha}$ 的值.

解 我们用两种方法求解.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

法一数形结合. 由 $\tan\alpha=2>0$ 知角 α 在第一或第三象限. 构造一个直角三角形, 使其一锐角 α' 满足 $\tan\alpha'=2=\frac{2}{1}$, 即对边为 2、邻边为 1, 由勾股定理斜边为 $\sqrt{2^2+1^2}=\sqrt{5}$. 于是 $|\sin\alpha|=\frac{2}{\sqrt{5}}$, $|\cos\alpha|=\frac{1}{\sqrt{5}}$, 再按象限分类.

- 若 α 在第一象限, 则 $\sin\alpha=\frac{2}{\sqrt{5}}$, $\cos\alpha=\frac{1}{\sqrt{5}}$,

$$\text{所求表达式}=\frac{1}{(\frac{2}{\sqrt{5}})^2-(\frac{2}{\sqrt{5}})(\frac{1}{\sqrt{5}})}=\frac{1}{\frac{4}{5}-\frac{2}{5}}=\frac{1}{\frac{2}{5}}=\frac{5}{2}.$$

- 若 α 在第三象限, 则 $\sin\alpha=-\frac{2}{\sqrt{5}}$, $\cos\alpha=-\frac{1}{\sqrt{5}}$,

$$\text{所求表达式}=\frac{1}{(-\frac{2}{\sqrt{5}})^2-(-\frac{2}{\sqrt{5}})(-\frac{1}{\sqrt{5}})}=\frac{1}{\frac{4}{5}-\frac{2}{5}}=\frac{1}{\frac{2}{5}}=\frac{5}{2}.$$

两种情形下结果相同, 故所求表达式的值均为 $\frac{5}{2}$.

法二代数变形. 所求表达式并非齐次式, 我们先把它改造成齐次式. 分子 1 可以写成 $\sin^2\alpha+\cos^2\alpha$; 代入后分子分母皆为二次齐次式, 再同除以 $\cos^2\alpha$, 得

$$\begin{aligned}\text{原式}&=\frac{\sin^2\alpha+\cos^2\alpha}{\sin^2\alpha-\sin\alpha\cos\alpha}\quad\text{(分子以平方关系代换)}\\&=\frac{\tan^2\alpha+1}{\tan^2\alpha-\tan\alpha}\qquad\text{(同除以 }\cos^2\alpha\text{)}.\end{aligned}$$

代入 $\tan\alpha=2$, 得

$$\text{原式}=\frac{2^2+1}{2^2-2}=\frac{5}{2}.$$

构造三角形法先求出绝对值, 最后需要由象限补符号. 代数变形法把 1 写成 $\sin^2\alpha+\cos^2\alpha$, 把非齐次式化为齐次式, 适用范围更广.

例题 6.10 (2023 年 6 月全国甲卷) 设甲为 $\sin^2\alpha+\sin^2\beta=1$, 乙为 $\sin\alpha+\cos\beta=0$, 则(　)

A. 甲是乙的充分条件但不是必要条件

B. 甲是乙的必要条件但不是充分条件

C. 甲是乙的充要条件

D. 甲既不是乙的充分条件也不是乙的必要条件

解 判断甲、乙的充分必要关系, 需分别检验"乙 $\Longrightarrow$ 甲"与"甲 $\Longrightarrow$ 乙"两个方向.

先证"乙 $\Longrightarrow$ 甲". 由条件乙 $\sin\alpha+\cos\beta=0$ 得 $\sin\alpha=-\cos\beta$, 两边平方并以平方关系 $\cos^2\beta=1-\sin^2\beta$ 代入,

$$\sin^2\alpha=\cos^2\beta=1-\sin^2\beta,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

移项即得 $\sin^2\alpha+\sin^2\beta=1$, 此即条件甲. 故“乙 $\Longrightarrow$ 甲”成立, 甲是乙的必要条件.

再考察“甲 $\Longrightarrow$ 乙”. 由条件甲 $\sin^2\alpha+\sin^2\beta=1$ 及平方关系得 $\sin^2\alpha=1-\sin^2\beta=\cos^2\beta$, 开方有 $\sin\alpha=\pm\cos\beta$, 此式含 $\sin\alpha=\cos\beta$ 与 $\sin\alpha=-\cos\beta$ 两种可能, 未必推出条件乙, 取 $\alpha=\beta=\frac{\pi}{4}$ 即可说明. 此时

$$\sin^2\frac{\pi}{4}+\sin^2\frac{\pi}{4}=\frac{1}{2}+\frac{1}{2}=1$$

满足甲, 而 $\sin\frac{\pi}{4}+\cos\frac{\pi}{4}=\sqrt{2}\neq 0$ 不满足乙. 故“甲 $\Longrightarrow$ 乙”不成立, 甲不是乙的充分条件.

由于“乙 $\Longrightarrow$ 甲”成立而“甲 $\Longrightarrow$ 乙”不成立, 甲是乙的必要条件, 但不是充分条件, 答案为 B.

知乎@余命数

6.7.3 高考真题选讲

例题 6.11 (2020 年全国 □ 卷理科数学第 9 题) 已知 $\alpha\in(0,\pi)$, 且 $3\cos 2\alpha-8\cos\alpha=5$, 则 $\sin\alpha=$

A. $\frac{\sqrt{5}}{3}$　　C. $\frac{1}{3}$

B. $\frac{2}{3}$　　D. $\frac{\sqrt{5}}{9}$

解 答案: A.

用二倍角公式 $\cos 2\alpha=2\cos^2\alpha-1$ 代入, 得

$$3(2\cos^2\alpha-1)-8\cos\alpha=5,$$

整理得 $3\cos^2\alpha-4\cos\alpha-4=0$, 解得 $\cos\alpha=2$(舍去)或 $\cos\alpha=-\frac{2}{3}$.

由 $\alpha\in(0,\pi)$ 且 $\cos\alpha<0$ 知 $\alpha\in\left(\frac{\pi}{2},\pi\right)$, 故

$$\sin\alpha=\sqrt{1-\cos^2\alpha}=\sqrt{1-\frac{4}{9}}=\frac{\sqrt{5}}{3}.$$

例题 6.12 (2019 年全国 □ 卷理科数学第 10 题) 已知 $\alpha\in\left(0,\frac{\pi}{2}\right)$, $2\sin 2\alpha=\cos 2\alpha+1$, 则 $\sin\alpha=$

A. $\frac{1}{5}$　　C. $\frac{\sqrt{3}}{3}$

B. $\frac{\sqrt{5}}{5}$　　D. $\frac{2\sqrt{5}}{5}$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解 答案:B.

用二倍角公式展开:$4\sin\alpha\cos\alpha = 2\cos^2\alpha$. 由 $\alpha \in \left(0, \dfrac{\pi}{2}\right)$ 知 $\cos\alpha > 0$, 两边除以 $2\cos\alpha$ 得 $\cos\alpha = 2\sin\alpha$.

代入 $\sin^2\alpha + \cos^2\alpha = 1$ 得 $5\sin^2\alpha = 1$, 故 $\sin\alpha = \dfrac{\sqrt{5}}{5}$.

6.8 诱导公式

我们已经会求 $\frac{\pi}{6}$、$\frac{\pi}{4}$、$\frac{\pi}{3}$ 这些锐角的三角函数值. 但实际旋转的角度是任意的, 如 $\sin\frac{2\pi}{3}$ 或 $\cos(-\frac{\pi}{6})$ 这类值就不能直接读出.

$120°$、$210°$、$-30°$ 等角的终边与 $30°$ 的终边在单位圆上具有对称性, 或关于 y 轴对称, 或关于原点对称, 或关于 x 轴对称.

诱导公式把这些对称性写成代数等式, 从而把任意角的三角函数计算化为锐角的三角函数计算.

我们以 α 为基准, 在单位圆上考察 $2k\pi+\alpha$、$\pi+\alpha$、$-\alpha$、$\pi-\alpha$ 等角的对称点, 建立它们与 α 的三角函数关系. 为看清象限, 先设 α 为锐角;所得的公式对任意角 α 都成立.

设角 α 的终边与单位圆交于点 $P(\cos\alpha, \sin\alpha)$.

6.8.0.1 关于原点与 x 轴的对称

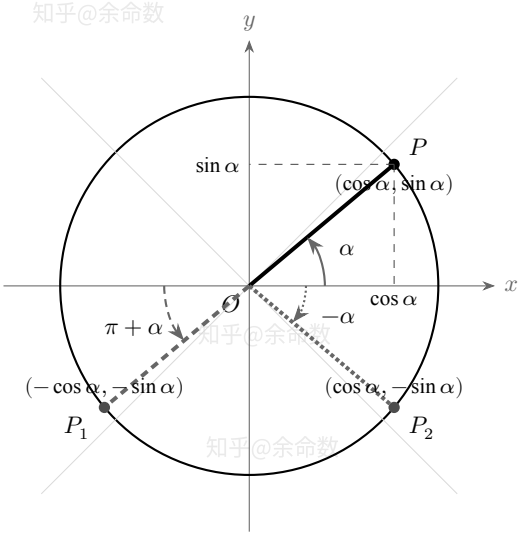


图 6.14 诱导公式的几何意义. 角 α、$\pi+\alpha$ 和 $-\alpha$ 的终边关系

- **周期性(公式一)** 角 $2k\pi+\alpha$ 与角 α 的终边完全相同, 所以它们的三角函数值也完全相同. 这是周期性的体现.

$$\sin(2k\pi+\alpha) = \sin\alpha \quad \cos(2k\pi+\alpha) = \cos\alpha \quad \tan(2k\pi+\alpha) = \tan\alpha$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

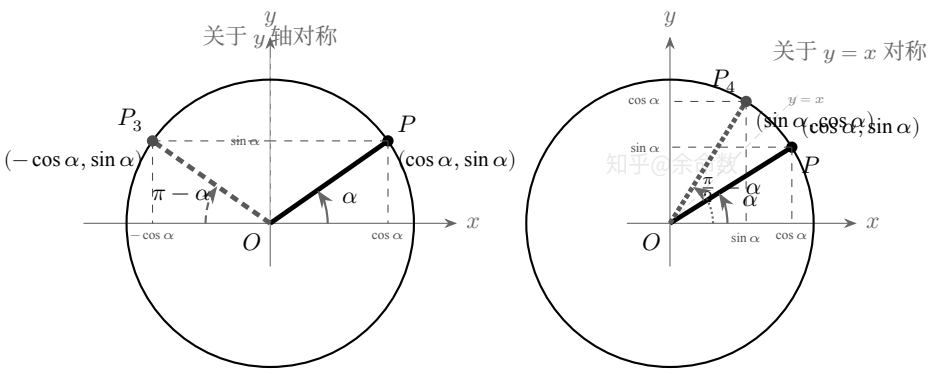
- **关于原点对称(公式二)**角 $\pi+\alpha$ 的终边是角 α 终边的反向延长线, 点 P_1 与点 P 关于原点对称. 因此 P_1 的坐标为 $(-\cos\alpha,-\sin\alpha)$.

$$\sin(\pi+\alpha)=-\sin\alpha\quad\cos(\pi+\alpha)=-\cos\alpha\quad\tan(\pi+\alpha)=\tan\alpha$$

- **关于 x 轴对称(公式三)**角 $-\alpha$ 的终边与角 α 的终边关于 x 轴对称. 点 P_2 与点 P 的横坐标相同, 纵坐标互为相反数. 因此 P_2 的坐标为 $(\cos\alpha,-\sin\alpha)$.

$$\sin(-\alpha)=-\sin\alpha\quad\cos(-\alpha)=\cos\alpha\quad\tan(-\alpha)=-\tan\alpha$$

6.8.0.2 关于 y 轴与直线 $y=x$ 的对称



(a) 角 α 和 $\pi-\alpha$　　**(b)** 角 α 和 $\frac{\pi}{2}-\alpha$

图 6.15　互余角和互补角的终边关系

- **关于 y 轴对称(公式四)**角 $\pi-\alpha$ 的终边与角 α 的终边关于 y 轴对称. 点 P_3 与点 P 的纵坐标相同, 横坐标互为相反数. 因此 P_3 的坐标为 $(-\cos\alpha,\sin\alpha)$.

$$\sin(\pi-\alpha)=\sin\alpha\quad\cos(\pi-\alpha)=-\cos\alpha\quad\tan(\pi-\alpha)=-\tan\alpha$$

- **关于直线 $y=x$ 对称(公式五)**角 $\frac{\pi}{2}-\alpha$ 的终边与角 α 的终边关于直线 $y=x$ 对称. 点 P_4 的坐标是点 P 坐标的横纵互换. 因此 P_4 的坐标为 $(\sin\alpha,\cos\alpha)$.

$$\sin(\frac{\pi}{2}-\alpha)=\cos\alpha\quad\cos(\frac{\pi}{2}-\alpha)=\sin\alpha\quad\tan(\frac{\pi}{2}-\alpha)=\cot\alpha$$

6.8.0.3 名称变换的诱导公式

上面的公式中, 函数名称没有改变. 而涉及 $\frac{\pi}{2}$ 和 $\frac{3\pi}{2}$ 的变换, 则会引起函数名称的改变, 正弦变余弦, 余弦变正弦.

- **角 $\frac{\pi}{2}+\alpha$(公式六)**可由已有公式推导.

$$\sin(\frac{\pi}{2}+\alpha)=\sin\Big(\pi-(\frac{\pi}{2}-\alpha)\Big)=\sin(\frac{\pi}{2}-\alpha)=\cos\alpha$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\cos(\frac{\pi}{2}+\alpha)=\cos\left(\pi-(\frac{\pi}{2}-\alpha)\right)=-\cos(\frac{\pi}{2}-\alpha)=-\sin\alpha$$

几何上, 角 $\frac{\pi}{2}+\alpha$ 的终边是将角 α 的终边逆时针旋转 90° 得到的, 点 (x,y) 旋转 90° 变为 $(-y,x)$, 同样可得.

$$\sin(\frac{\pi}{2}+\alpha)=\cos\alpha\quad\cos(\frac{\pi}{2}+\alpha)=-\sin\alpha\quad\tan(\frac{\pi}{2}+\alpha)=-\cot\alpha$$

6.8.1 奇变偶不变, 符号看象限

这些公式的内在规律可以概括为一句口诀, **奇变偶不变, 符号看象限**, 适用于所有 $k\cdot\frac{\pi}{2}\pm\alpha\ (k\in\mathbb{Z})$ 形式的诱导公式.

- **奇变偶不变**. 指角 $k\cdot\frac{\pi}{2}\pm\alpha$ 中 k 的奇偶性.
  - 当 k 为偶数时, 如 $\pi\pm\alpha=2\cdot\frac{\pi}{2}\pm\alpha$、$2\pi\pm\alpha=4\cdot\frac{\pi}{2}\pm\alpha$, 函数名称不变 ($\sin\to\sin,\cos\to\cos$).
  - 当 k 为奇数时, 如 $\frac{\pi}{2}\pm\alpha$、$\frac{3\pi}{2}\pm\alpha$, 函数名称改变 ($\sin\leftrightarrow\cos$).
- **符号看象限**. 判断结果正负时, 先设 α 为锐角, 确定原角 $k\cdot\frac{\pi}{2}\pm\alpha$ 所在象限, 再取原函数在该象限的符号.

符号看象限时, 我们看的是原函数在原角所在象限的符号. 例如, 判断 cos 在第四象限的符号, 结果为正. 设 α 为锐角只是为了方便判断象限; 这条口诀对任意角 α 都成立, 依据是三角函数的周期性与对称性.

6.8.2 典例解析

例题 6.13 计算 $\sin(-\frac{13\pi}{6})+\cos(\frac{19\pi}{3})-\tan(-\frac{29\pi}{4})$ 的值.

解 我们先把负角化成正角, 再把角化到 $[0,2\pi)$ 内, 最后用诱导公式化为锐角. 先用公式三去掉负号,

$$\begin{aligned}\text{原式}&=\sin(-\frac{13\pi}{6})+\cos(\frac{19\pi}{3})-\tan(-\frac{29\pi}{4})\\&=-\sin(\frac{13\pi}{6})+\cos(\frac{19\pi}{3})+\tan(\frac{29\pi}{4})\quad\text{(奇偶性)}.\end{aligned}$$

再由周期性把每一项化为 $[0,2\pi)$ 内的熟悉角,

$$\begin{aligned}&\sin(\frac{13\pi}{6})=\sin(2\pi+\frac{\pi}{6})=\sin(\frac{\pi}{6}),\\&\cos(\frac{19\pi}{3})=\cos(6\pi+\frac{\pi}{3})=\cos(\frac{\pi}{3}),\\&\tan(\frac{29\pi}{4})=\tan(7\pi+\frac{\pi}{4})=\tan(\pi+\frac{\pi}{4})=\tan(\frac{\pi}{4})\quad\text{(周期 }\pi\text{)}.\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入特殊角值 $\sin(\frac{\pi}{6})=\frac{1}{2}$、$\cos(\frac{\pi}{3})=\frac{1}{2}$、$\tan(\frac{\pi}{4})=1$，得

$$\text{原式}=-\left(\frac{1}{2}\right)+\frac{1}{2}+1=1.$$

奇偶性、周期性与诱导公式的组合，把任意角的三角函数化为锐角的三角函数，再落到一组熟记的特殊值上.

注 6.2 (方法总结)
任意角的三角函数求值可以分三步.

- **化正角**对负角用 $\sin(-\alpha)=-\sin\alpha$、$\cos(-\alpha)=\cos\alpha$、$\tan(-\alpha)=-\tan\alpha$.
- **化小角**利用周期性，把角加上或减去 2π（正切是 π）的整数倍，化到 $[0,2\pi)$ 之类的熟悉区间.
- **化锐角**再用诱导公式把大于 $\frac{\pi}{2}$ 的角化为锐角.

例题 6.14　化简分式 $\dfrac{\sin(\pi+\alpha)\cos(2\pi-\alpha)\tan(3\pi-\alpha)}{\cos(\frac{\pi}{2}+\alpha)\sin(-\alpha-\pi)}$.

解　我们用口诀把表达式中的每一项化为只含 α 的三角函数. 遇到 $\sin(-\alpha-\pi)$ 这类带多重符号的角，先处理括号内的负号.

分子的三项依次化简，得

$$\begin{aligned}
&\sin(\pi+\alpha)=-\sin\alpha && \text{（偶不变，III 象限 sin 为负），}\\
&\cos(2\pi-\alpha)=\cos(-\alpha)=\cos\alpha && \text{（周期性，IV 象限 cos 为正），}\\
&\tan(3\pi-\alpha)=\tan(-\alpha)=-\tan\alpha && \text{（周期 }\pi\text{，II 象限 tan 为负）.}
\end{aligned}$$

分母的两项为

$$\cos\left(\frac{\pi}{2}+\alpha\right)=-\sin\alpha \qquad \text{（奇变，II 象限 cos 为负），}$$

$$\sin(-\alpha-\pi)=\sin(-(\pi+\alpha))=-\sin(\pi+\alpha)=-(-\sin\alpha)=\sin\alpha \qquad \text{（先提内层负号）.}$$

将上述结果代回原分式，

$$\text{原式}=\frac{(-\sin\alpha)(\cos\alpha)(-\tan\alpha)}{(-\sin\alpha)(\sin\alpha)}=\frac{\sin\alpha\cdot\cos\alpha\cdot\tan\alpha}{-\sin^2\alpha},$$

再以商数关系 $\tan\alpha=\frac{\sin\alpha}{\cos\alpha}$ 切化弦，得

$$\text{原式}=\frac{\sin\alpha\cdot\cos\alpha\cdot(\frac{\sin\alpha}{\cos\alpha})}{-\sin^2\alpha}=\frac{\sin^2\alpha}{-\sin^2\alpha}=-1.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

化简结果为 -1.

注 6.3 (经验总结)
处理复杂的三角表达式时, 我们把每一项单独化简; 对 $\sin(-\alpha-\pi)$ 这类复合角, 先提取括号内的负号; 最后若同时出现 sin、cos、tan, 可以把 tan 写成 $\frac{\sin}{\cos}$, 方便约分.

例题 6.15 在 $\triangle ABC$ 中, 求证.
(1) $\sin(A+B)=\sin C$ (2) $\cos A=-\cos(B+C)$ (3) $\sin\frac{A+B}{2}=\cos\frac{C}{2}$

解 在三角形中, $A+B+C=\pi$. 我们用它把所证式中的角改写为诱导公式的标准形式, 即可得证.

(1) 由 $A+B+C=\pi$ 得 $A+B=\pi-C$, 代入等式左边并用诱导公式 $\sin(\pi-C)=\sin C$,

$$\sin(A+B)=\sin(\pi-C)=\sin C,$$

故 $\sin(A+B)=\sin C$.

(2) 由 $A+B+C=\pi$ 得 $B+C=\pi-A$, 考察等式右边并用诱导公式 $\cos(\pi-A)=-\cos A$,

$$-\cos(B+C)=-\cos(\pi-A)=-(-\cos A)=\cos A,$$

右边即等于左边的 $\cos A$, 故 $\cos A=-\cos(B+C)$.

(3) 由 $A+B+C=\pi$, 两边同除以 2 得 $\frac{A+B}{2}+\frac{C}{2}=\frac{\pi}{2}$, 即 $\frac{A+B}{2}=\frac{\pi}{2}-\frac{C}{2}$, 代入等式左边并用诱导公式 $\sin(\frac{\pi}{2}-x)=\cos x$,

$$\sin\frac{A+B}{2}=\sin\left(\frac{\pi}{2}-\frac{C}{2}\right)=\cos\frac{C}{2},$$

故 $\sin\frac{A+B}{2}=\cos\frac{C}{2}$.

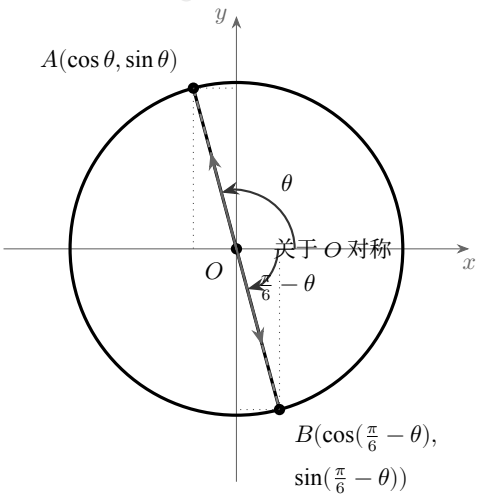
注 6.4 (经验总结)
在三角形中, $A+B+C=\pi$ 可以把两角和写成 π 减第三角, 也可以把半角写成 $\frac{\pi}{2}$ 减第三角的一半. 这些是诱导公式的标准形式.

例题 6.16 (2023 年 12 月广东佛山高三测试)若点 $A(\cos\theta,\sin\theta)$ 关于原点对称点为 $B\left(\cos(\frac{\pi}{6}-\theta),\sin(\frac{\pi}{6}-\theta)\right)$, 写出 θ 的一个取值为 _____.

解 点 A 在单位圆上, 关于原点对称的点 B 对应的角与 A 对应的角相差 π 的奇数倍.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数



点 A 与点 B 关于原点对称, 意味着其终边互为反向延长线, 对应的角相差 π 的奇数倍. 设 $\theta_A=\theta, \theta_B=\dfrac{\pi}{6}-\theta$, 由对称性有

$$\theta_B-\theta_A=(2k+1)\pi,\quad k\in\mathbb{Z}.$$

代入 θ_A, θ_B 并整理,

$$\begin{aligned}\left(\frac{\pi}{6}-\theta\right)-\theta&=(2k+1)\pi,\\ \frac{\pi}{6}-2\theta&=(2k+1)\pi,\\ -2\theta=(2k+1)\pi-\frac{\pi}{6}&=\frac{12k\pi+6\pi-\pi}{6}=\frac{(12k+5)\pi}{6},\end{aligned}$$

解得

$$\theta=-\frac{(12k+5)\pi}{12},\quad k\in\mathbb{Z}.$$

取 $k=0$ 得 $\theta=-\dfrac{5\pi}{12}$, 取 $k=-1$ 得 $\theta=-\dfrac{(-12+5)\pi}{12}=\dfrac{7\pi}{12}$. 因此 θ 的一个可能取值为 $\dfrac{7\pi}{12}$ 或 $-\dfrac{5\pi}{12}$ 等.

注 6.5 (经验总结)
本题把几何对称转化为代数关系. 几何上, 关于原点对称说明两个角的终边在同一直线上; 代数上, 这等价于两个角相差 π 的奇数倍, 即 $\theta_2-\theta_1=(2k+1)\pi$. 建立方程后, 再用代数运算求解. 处理单位圆问题时, 这种几何直观常常能直接给出突破口.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

例题 6.17 （2023 年 11 月广东金太阳联考）设 $\alpha \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right], \beta \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right]$，且 $\sin\alpha + \cos\alpha = \sqrt{2}\cos\beta$，则（　）

A. $\alpha + \beta = \frac{\pi}{4}$

B. $\alpha - \beta = \frac{\pi}{4}$

C. $\alpha + \beta = \frac{\pi}{2}$

D. $\alpha - \beta = -\frac{\pi}{4}$

解 等式左边可用辅助角公式合并，右边是 $\sqrt{2}\cos\beta$，结合两边在给定区间上的值域即可锁定 α 和 β.

对等式左边用辅助角公式变形，

$$\sin\alpha + \cos\alpha = \sqrt{2}\left(\frac{\sqrt{2}}{2}\sin\alpha + \frac{\sqrt{2}}{2}\cos\alpha\right) = \sqrt{2}\sin\left(\alpha + \frac{\pi}{4}\right),$$

代回原等式并约去 $\sqrt{2}$，得

$$\sin\left(\alpha + \frac{\pi}{4}\right) = \cos\beta.$$

此方程在给定区间上不能直接求解，我们改先考察两边的值域. 由 $\alpha \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right]$ 知 $\alpha + \frac{\pi}{4} \in \left[\frac{\pi}{2}, \frac{3\pi}{4}\right]$，在该区间上 sin 的值域为 $\left[\sin\frac{3\pi}{4}, \sin\frac{\pi}{2}\right] = \left[\frac{\sqrt{2}}{2}, 1\right]$；而 $\beta \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right]$，cos 的值域为 $\left[\cos\frac{\pi}{2}, \cos\frac{\pi}{4}\right] = \left[0, \frac{\sqrt{2}}{2}\right]$. 两边值域的唯一公共值为 $\frac{\sqrt{2}}{2}$，故必须同时满足

$$\sin\left(\alpha + \frac{\pi}{4}\right) = \frac{\sqrt{2}}{2} \quad \text{且} \quad \cos\beta = \frac{\sqrt{2}}{2}.$$

在 $\alpha + \frac{\pi}{4} \in \left[\frac{\pi}{2}, \frac{3\pi}{4}\right]$ 中 $\sin(\alpha + \frac{\pi}{4}) = \frac{\sqrt{2}}{2}$ 给出 $\alpha + \frac{\pi}{4} = \frac{3\pi}{4}$，即 $\alpha = \frac{\pi}{2}$；在 $\beta \in \left[\frac{\pi}{4}, \frac{\pi}{2}\right]$ 中 $\cos\beta = \frac{\sqrt{2}}{2}$ 给出 $\beta = \frac{\pi}{4}$. 因此 $\alpha - \beta = \frac{\pi}{2} - \frac{\pi}{4} = \frac{\pi}{4}$，答案为 B.

注 6.6 (经验总结)
当三角方程难以直接求解，且自变量范围已知时，可以分别求出等式两边的值域，取交集后往往能得到唯一可能的函数值，从而确定变量的取值.

6.8.3 高考真题选讲

例题 6.18 (2010 年全国 □ 卷文科数学第 3 题) 已知 $\sin\alpha = \frac{2}{3}$，则 $\cos(\pi - 2\alpha) =$

A. $-\frac{\sqrt{5}}{3}$　　C. $\frac{1}{9}$

B. $-\frac{1}{9}$　　D. $\frac{\sqrt{5}}{3}$

解 答案：B. 先用诱导公式 $\cos(\pi - 2\alpha) = -\cos 2\alpha$，再用倍角公式 $\cos 2\alpha = 1 - 2\sin^2\alpha$，

$$\cos(\pi - 2\alpha) = -(1 - 2\sin^2\alpha) = -\left(1 - 2\cdot\frac{4}{9}\right) = -\frac{1}{9}.$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

6.9 和差角公式

我们已经会求 30°、45°、60° 等特殊角的三角函数值. 遇到 75° 或 15° 这类角时, 我们想用已知角表示它们. 例如

$$75^\circ = 45^\circ + 30^\circ, \qquad 15^\circ = 45^\circ - 30^\circ.$$

一个自然的猜测是 $\cos(\alpha-\beta)$ 等于 $\cos\alpha-\cos\beta$. 但这个等式不成立. 取 $\alpha = 60^\circ, \beta = 30^\circ$, 则 $\cos(60^\circ - 30^\circ) = \cos 30^\circ = \frac{\sqrt{3}}{2}$, 而 $\cos 60^\circ - \cos 30^\circ = \frac{1}{2} - \frac{\sqrt{3}}{2}$, 两者不相等.

我们关心两个角的和差如何进入三角函数. 下面的推导借助单位圆上的向量点积.

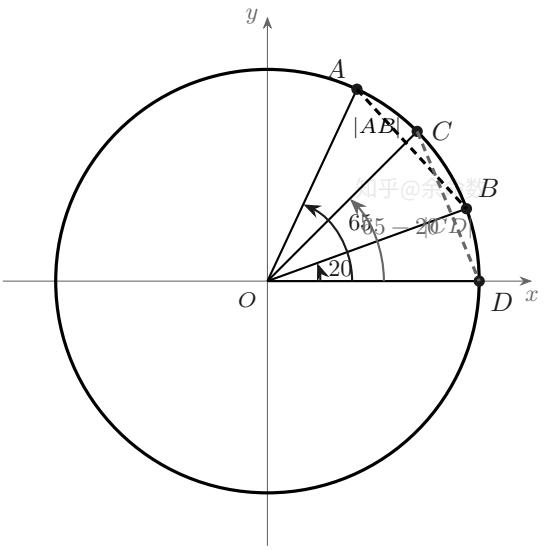
6.9.1 两角差的余弦

定理 6.3 (两角差的余弦) 对于任意角 α, β, 恒有

$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta \tag{6-6}$$

知乎@余命数

证明 我们用单位圆上等弧对等弦的性质来证明. 如图所示, 在单位圆上取四点 A, B, C, D, 对应的角分别为 $\alpha, \beta, \alpha-\beta, 0$.



知乎@余命数
知乎@余命数

证明依赖等弧对等弦. 圆心角 $\angle AOB = \angle COD = \alpha - \beta$, 故弦长相等, 即 $|AB| = |CD|$. 两边平方得 $|AB|^2 = |CD|^2$, 用两点间距离公式把两端分别展开,

$$(\cos\alpha - \cos\beta)^2 + (\sin\alpha - \sin\beta)^2 = (\cos(\alpha-\beta) - 1)^2 + \sin^2(\alpha-\beta).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

左端展开并利用 $\sin^2\theta+\cos^2\theta=1$ 化简,

$$\begin{aligned}&\cos^2\alpha-2\cos\alpha\cos\beta+\cos^2\beta+\sin^2\alpha-2\sin\alpha\sin\beta+\sin^2\beta\\&=2-2(\cos\alpha\cos\beta+\sin\alpha\sin\beta),\end{aligned}$$

右端同理化简,

$$\begin{aligned}&\cos^2(\alpha-\beta)-2\cos(\alpha-\beta)+1+\sin^2(\alpha-\beta)\\&=2-2\cos(\alpha-\beta).\end{aligned}$$

两端相等, 即

$$2-2(\cos\alpha\cos\beta+\sin\alpha\sin\beta)=2-2\cos(\alpha-\beta),$$

整理即得

$$\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta.$$

□

这个公式是和差角公式的起点, 其余公式都由它通过代数变形得到. 它的结构可记为余余正正, 符号反.

6.9.2 公式体系的构建

我们由 $\cos(\alpha-\beta)$ 推出其余和差角公式.

- **推导** $\cos(\alpha+\beta)$
  将 $\alpha+\beta$ 看作 $\alpha-(-\beta)$.

$$\begin{aligned}\cos(\alpha+\beta)&=\cos(\alpha-(-\beta))\\&=\cos\alpha\cos(-\beta)+\sin\alpha\sin(-\beta)\\&=\cos\alpha\cos\beta-\sin\alpha\sin\beta\end{aligned}$$

- **推导** $\sin(\alpha+\beta)$
  利用诱导公式 $\sin\theta=\cos(\frac{\pi}{2}-\theta)$.

$$\begin{aligned}\sin(\alpha+\beta)&=\cos\left(\frac{\pi}{2}-(\alpha+\beta)\right)=\cos\left(\left(\frac{\pi}{2}-\alpha\right)-\beta\right)\\&=\cos(\frac{\pi}{2}-\alpha)\cos\beta+\sin(\frac{\pi}{2}-\alpha)\sin\beta\\&=\sin\alpha\cos\beta+\cos\alpha\sin\beta\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 6.4 (和差角公式)

$$\sin(\alpha\pm\beta)=\sin\alpha\cos\beta\pm\cos\alpha\sin\beta \qquad \text{(正余余正, 符号同)}$$

$$\cos(\alpha\pm\beta)=\cos\alpha\cos\beta\mp\sin\alpha\sin\beta \qquad \text{(余余正正, 符号反)}$$

$$\tan(\alpha\pm\beta)=\frac{\tan\alpha\pm\tan\beta}{1\mp\tan\alpha\tan\beta} \qquad \text{(分子符号同, 分母符号反)}$$

其中正切公式要求 $\tan\alpha,\tan\beta,\tan(\alpha\pm\beta)$ 均有意义, 且分母 $1\mp\tan\alpha\tan\beta\neq0$.

证明 我们用 $\cos(\alpha-\beta)$ 的公式、奇偶性与诱导公式直接推出正弦、余弦的和差公式. 例如

知乎@余命数

$$\cos(\alpha+\beta)=\cos(\alpha-(-\beta))=\cos\alpha\cos\beta-\sin\alpha\sin\beta,$$

而

$$\sin(\alpha+\beta)=\cos\left(\frac{\pi}{2}-(\alpha+\beta)\right)=\sin\alpha\cos\beta+\cos\alpha\sin\beta.$$

知乎@余命数

将 β 替换为 $-\beta$ 即得相应的差角公式.

当相关正切均有意义时, 由商数关系可得

$$\begin{aligned}\tan(\alpha+\beta)&=\frac{\sin(\alpha+\beta)}{\cos(\alpha+\beta)}\\&=\frac{\sin\alpha\cos\beta+\cos\alpha\sin\beta}{\cos\alpha\cos\beta-\sin\alpha\sin\beta}.\end{aligned}$$

在 $\cos\alpha\cos\beta\neq0$ 时, 分子分母同除以 $\cos\alpha\cos\beta$, 得

$$\tan(\alpha+\beta)=\frac{\tan\alpha+\tan\beta}{1-\tan\alpha\tan\beta}.$$

差角正切公式同理, 或将 β 替换为 $-\beta$ 得到. 分母非零正对应 $\cos(\alpha\pm\beta)\neq0$, 即 $\tan(\alpha\pm\beta)$ 有意义. □

知乎@余命数

6.9.3 辅助角公式

我们经常遇到函数 $f(x)=a\sin x+b\cos x$; 它表示两个同频率但不同相位的简谐波叠加. 和差角公式把这个和式合并成一个单一的正弦或余弦函数, 从而可以直接分析最值、周期和单调性.

我们称这种变形为**辅助角公式**.

$$a\sin x+b\cos x=\sqrt{a^2+b^2}\left(\frac{a}{\sqrt{a^2+b^2}}\sin x+\frac{b}{\sqrt{a^2+b^2}}\cos x\right)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

关键在于构造一个辅助角. 令 $\cos\varphi=\frac{a}{\sqrt{a^2+b^2}}$, $\sin\varphi=\frac{b}{\sqrt{a^2+b^2}}$, 则 $\tan\varphi=\frac{b}{a}$. 于是

$$\begin{aligned}\text{原式}&=\sqrt{a^2+b^2}(\cos\varphi\sin x+\sin\varphi\cos x)\\&=\sqrt{a^2+b^2}\sin(x+\varphi)\end{aligned}$$

注 6.7 (辅助角变形)
要把 $a\sin x+b\cos x$ 写成 $\sqrt{a^2+b^2}\sin(x+\varphi)$ 的形式, 我们先提取系数 $\sqrt{a^2+b^2}$, 再把括号内配成 $\sin x\cos\varphi+\cos x\sin\varphi$ 或 $\cos x\cos\varphi+\sin x\sin\varphi$ 的形式, 最后由 $\tan\varphi=\dfrac{b}{a}$ 与 a,b 的符号确定 φ 所在象限.

6.9.4 典例解析

例题 6.19 (给值求值) 已知 $\sin\alpha=\frac{5}{13}$、$\alpha\in(\frac{\pi}{2},\pi)$、$\cos\beta=-\frac{3}{5}$、$\beta$ 是第三象限角. 求 $\sin(\alpha+\beta)$ 的值.

解 我们先用象限条件补全各角的正弦、余弦, 再代入和差角公式. 由 $\alpha\in(\frac{\pi}{2},\pi)$ 知 $\cos\alpha<0$, 故

$$\cos\alpha=-\sqrt{1-\sin^2\alpha}=-\sqrt{1-(\frac{5}{13})^2}=-\sqrt{\frac{144}{169}}=-\frac{12}{13};$$

又 β 是第三象限角, $\sin\beta<0$, 故

$$\sin\beta=-\sqrt{1-\cos^2\beta}=-\sqrt{1-(-\frac{3}{5})^2}=-\sqrt{\frac{16}{25}}=-\frac{4}{5}.$$

代入正弦和角公式 $\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta$, 得

$$\begin{aligned}\sin(\alpha+\beta)&=\left(\frac{5}{13}\right)\left(-\frac{3}{5}\right)+\left(-\frac{12}{13}\right)\left(-\frac{4}{5}\right)\\&=-\frac{15}{65}+\frac{48}{65}=\frac{33}{65}.\end{aligned}$$

例题 6.20 (函数性质分析) 求函数 $f(x)=\sqrt{3}\sin x-\cos x$ 的最小正周期和单调递增区间.

解 我们先用辅助角公式把函数化为 $y=A\sin(\omega x+\varphi)+k$ 的形式, 再读出周期与单调性. 此处 $a=\sqrt{3},b=-1$, 故 $\sqrt{a^2+b^2}=\sqrt{3+1}=2$, 从而

$$f(x)=2\left(\frac{\sqrt{3}}{2}\sin x-\frac{1}{2}\cos x\right)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
&= 2\left(\sin x\cos\frac{\pi}{6} - \cos x\sin\frac{\pi}{6}\right)\\
&= 2\sin\left(x-\frac{\pi}{6}\right).
\end{aligned}
$$

由标准形式 $2\sin(x-\frac{\pi}{6})$ 知最小正周期 $T=\frac{2\pi}{|\omega|}=\frac{2\pi}{1}=2\pi$. 求单调递增区间, 须令内层 $u=x-\frac{\pi}{6}$ 落在正弦的标准递增区间 $2k\pi-\frac{\pi}{2}\le u\le 2k\pi+\frac{\pi}{2}$ $(k\in\mathbb{Z})$ 上, 即

$$
\begin{aligned}
&2k\pi-\frac{\pi}{2}\le x-\frac{\pi}{6}\le 2k\pi+\frac{\pi}{2},\\
&2k\pi-\frac{\pi}{3}\le x\le 2k\pi+\frac{2\pi}{3}\quad(\text{两端加 }\frac{\pi}{6}),
\end{aligned}
$$

故单调递增区间为 $\left[2k\pi-\frac{\pi}{3},2k\pi+\frac{2\pi}{3}\right],k\in\mathbb{Z}$.

例题 6.21 (2024 年 1 月金丽衢十二校第一次联考) 已知 α 是第二象限角, $\beta\in(0,\frac{\pi}{2})$, $\tan(\alpha+\frac{\pi}{4})=-\frac{1}{4}$. 现将角 α 的终边逆时针旋转 β 后得到角 γ, 若 $\tan\gamma=\frac{1}{7}$, 则 $\tan\beta=$ _____.

解 我们把逆时针旋转 β 理解为 $\gamma=\alpha+\beta$, 于是 $\beta=\gamma-\alpha$, 再用差角公式求 $\tan\beta$. 这要求我们先求出 $\tan\gamma$ 和 $\tan\alpha$.

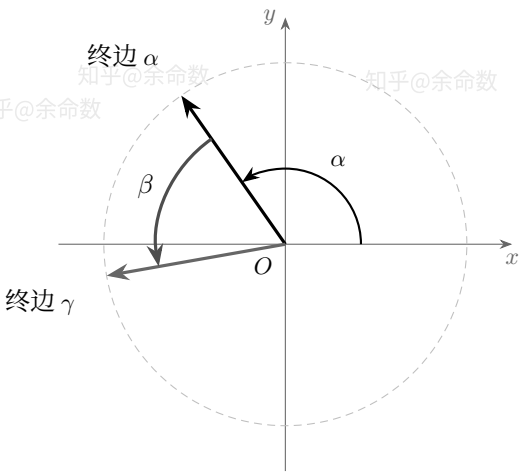


图 6.16　角 α 的终边逆时针旋转角 β 得到角 γ 的终边

先求 $\tan\alpha$. 由正切和角公式

$$
\tan\left(\alpha+\frac{\pi}{4}\right)=\frac{\tan\alpha+\tan\frac{\pi}{4}}{1-\tan\alpha\tan\frac{\pi}{4}}=\frac{\tan\alpha+1}{1-\tan\alpha},
$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入已知值 $\tan(\alpha+\frac{\pi}{4})=-\frac{1}{4}$ 得方程 $\frac{1+\tan\alpha}{1-\tan\alpha}=-\frac{1}{4}$, 交叉相乘解得

$$4(1+\tan\alpha)=-(1-\tan\alpha)\implies 4+4\tan\alpha=-1+\tan\alpha\implies 3\tan\alpha=-5,$$

即 $\tan\alpha=-\frac{5}{3}$ (α 是第二象限角, $\tan\alpha$ 为负, 与题意相符).

再由 $\tan\gamma=\frac{1}{7}$ 与差角公式求 $\tan\beta=\tan(\gamma-\alpha)$,

$$\begin{aligned}\tan\beta=\tan(\gamma-\alpha)&=\frac{\tan\gamma-\tan\alpha}{1+\tan\gamma\tan\alpha}\\&=\frac{\frac{1}{7}-\left(-\frac{5}{3}\right)}{1+\frac{1}{7}\cdot\left(-\frac{5}{3}\right)}\\&=\frac{\frac{3+35}{21}}{1-\frac{5}{21}}=\frac{\frac{38}{21}}{\frac{16}{21}}=\frac{38}{16}=\frac{19}{8}.\end{aligned}$$

故 $\tan\beta=\dfrac{19}{8}$.

注 6.8 (旋转与差角)
几何描述给出终边旋转关系 $\gamma=\alpha+\beta$, 于是 $\beta=\gamma-\alpha$. 要求 $\tan\beta$, 只需求出 $\tan\gamma$ 和 $\tan\alpha$. 而已知 $\tan(\alpha+\frac{\pi}{4})$ 可通过正切和角公式化为关于 $\tan\alpha$ 的方程.

例题 6.22 (2023 年 12 月江苏省决胜新高考三大联考) 已知实数 m,n 满足 $(m+1)(n+1)=2$, 则 m,n 可能是(　)

A. $m=\tan\frac{\pi}{16},n=\tan\frac{3\pi}{16}$

B. $m=\tan\frac{\pi}{8},n=\tan\frac{3\pi}{8}$

C. $m=\cos\frac{\pi}{16},n=\tan\frac{3\pi}{16}$

D. $m=\cos\frac{\pi}{8},n=\tan\frac{3\pi}{8}$

解 我们先把代数式变形, 看出它与正切和角公式的结构相同, 再用三角函数检验选项. 将 $(m+1)(n+1)=2$ 展开并移项, 得

$$mn+m+n+1=2\implies m+n=1-mn.$$

当 $1-mn\neq 0$ 时, 上式可化为 $\frac{m+n}{1-mn}=1$, 此形式与正切和角公式 $\tan(A+B)=\frac{\tan A+\tan B}{1-\tan A\tan B}$ 一致, 故令 $m=\tan A,n=\tan B$, 条件即转化为 $\tan(A+B)=1$, 即 $A+B=k\pi+\frac{\pi}{4}$ ($k\in\mathbb{Z}$).

逐一检验哪个选项满足 $A+B=k\pi+\frac{\pi}{4}$.

- **A**: $A=\frac{\pi}{16},B=\frac{3\pi}{16}$, $A+B=\frac{4\pi}{16}=\frac{\pi}{4}$, 满足 ($k=0$), 故 A 可能.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **B**: $A=\frac{\pi}{8}, B=\frac{3\pi}{8}, A+B=\frac{\pi}{2}$, 此时 $\tan(A+B)$ 无定义, 对应分母 $1-mn=0$, 即 $mn=1$; 代入原式 $mn+m+n+1=2$ 得 $m+n=0$, 但 $m=\tan\frac{\pi}{8}>0, n=\tan\frac{3\pi}{8}>0$, 和不能为 0, 故 B 错误.
- **C**: $m=\cos\frac{\pi}{16}>\frac{1}{2}$, 且 $n=\tan\frac{3\pi}{16}>\tan\frac{\pi}{8}=\sqrt{2}-1$, 因而 $(m+1)(n+1)>\frac{3}{2}\cdot\sqrt{2}>2$, 不满足条件.
- **D**: $m=\cos\frac{\pi}{8}>0$, 且 $n=\tan\frac{3\pi}{8}>1$, 因而 $(m+1)(n+1)>1\cdot 2=2$, 不满足条件.

正确答案为 A.

注 6.9 (结构识别)
把 $(m+1)(n+1)=2$ 化简为 $m+n=1-mn$, 于是 $\dfrac{m+n}{1-mn}=1$. 这个式子与正切和角公式 $\tan(A+B)=\dfrac{\tan A+\tan B}{1-\tan A\tan B}$ 结构相同, 因此可令 $m=\tan A, n=\tan B$, 问题化为检验 $A+B=k\pi+\frac{\pi}{4}$.

例题 6.23 (2023 年 12 月山东中学联盟高三联考) 已知 $\cos(\beta-\alpha)=\dfrac{3}{5}$, 且 $\tan\alpha\tan\beta=\dfrac{1}{2}$, 求 $\cos 2(\alpha+\beta)$ 的值(　)

A. $-\dfrac{2}{25}$
B. $\dfrac{23}{25}$
C. $\dfrac{2}{25}$
D. $-\dfrac{23}{25}$

解 我们把 $\cos\alpha\cos\beta$ 与 $\sin\alpha\sin\beta$ 看作两个整体未知数, 解方程组求出它们, 再求目标值.

由已知条件得到两个关系式, 即 $\cos(\beta-\alpha)=\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta=\frac{3}{5}$ 与 $\tan\alpha\tan\beta=\frac{\sin\alpha\sin\beta}{\cos\alpha\cos\beta}=\frac{1}{2}$. 令 $X=\cos\alpha\cos\beta, Y=\sin\alpha\sin\beta$, 上述两式化为关于 X,Y 的方程组

$$\begin{cases} X+Y=\frac{3}{5} \\ \frac{Y}{X}=\frac{1}{2} \end{cases}$$

由第二式得 $X=2Y$, 代入第一式有 $3Y=\frac{3}{5}$, 即 $Y=\frac{1}{5}, X=\frac{2}{5}$, 故

$$\sin\alpha\sin\beta=\frac{1}{5},\quad \cos\alpha\cos\beta=\frac{2}{5}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

要计算 $\cos 2(\alpha+\beta)$, 我们先求 $\cos(\alpha+\beta)$, 再用二倍角公式 $\cos 2\theta = 2\cos^2\theta - 1$. 由余弦和角公式,

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta = X - Y = \frac{2}{5} - \frac{1}{5} = \frac{1}{5},$$

代入二倍角公式,

$$\begin{aligned}\cos 2(\alpha+\beta) &= 2\cos^2(\alpha+\beta) - 1 = 2\left(\frac{1}{5}\right)^2 - 1\\ &= \frac{2}{25} - \frac{25}{25} = -\frac{23}{25}.\end{aligned}$$

正确答案为 D.

6.9.5 高考真题选讲

例题 6.24 (2011 年浙江卷理科数学第 6 题) 若 $0 < \alpha < \frac{\pi}{2}$、$-\frac{\pi}{2} < \beta < 0$、$\cos\left(\frac{\pi}{4}+\alpha\right) = \frac{1}{3}$、$\cos\left(\frac{\pi}{4}-\frac{\beta}{2}\right) = \frac{\sqrt{3}}{3}$, 则 $\cos\left(\alpha+\frac{\beta}{2}\right) =$

A. $\frac{\sqrt{3}}{3}$　　C. $\frac{5\sqrt{3}}{9}$

B. $-\frac{\sqrt{3}}{3}$　　D. $-\frac{\sqrt{6}}{9}$

解 答案:C. 注意到 $\alpha+\frac{\beta}{2} = \left(\frac{\pi}{4}+\alpha\right) - \left(\frac{\pi}{4}-\frac{\beta}{2}\right)$, 可直接用余弦差角公式. 先求两个角的正弦值. 由 $0 < \alpha < \frac{\pi}{2}$ 知 $\frac{\pi}{4} < \frac{\pi}{4}+\alpha < \frac{3\pi}{4}$, 故 $\sin\left(\frac{\pi}{4}+\alpha\right) = \sqrt{1-\frac{1}{9}} = \frac{2\sqrt{2}}{3}$; 由 $-\frac{\pi}{2} < \beta < 0$ 知 $\frac{\pi}{4} < \frac{\pi}{4}-\frac{\beta}{2} < \frac{\pi}{2}$, 故 $\sin\left(\frac{\pi}{4}-\frac{\beta}{2}\right) = \sqrt{1-\frac{1}{3}} = \frac{\sqrt{6}}{3}$.

代入余弦差角公式,

$$\cos\left(\alpha+\frac{\beta}{2}\right) = \frac{1}{3}\cdot\frac{\sqrt{3}}{3} + \frac{2\sqrt{2}}{3}\cdot\frac{\sqrt{6}}{3} = \frac{\sqrt{3}}{9} + \frac{4\sqrt{3}}{9} = \frac{5\sqrt{3}}{9}.$$

6.10 倍角公式

来龙去脉

我们在和角公式中令两个角相同, 即设 $\alpha = \beta$, 便得到二**倍角公式**.

二倍角公式把角和它的二倍角联系起来. 由它移项可得**降幂公式**; 把倍角换成半角又可得**半角公式**. 这些公式在化简与求值中都会用到.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

6.10.1 公式体系

定理 6.5 (倍角、半角公式)　1. **二倍角公式**

$$\begin{aligned}\sin(2\alpha) &= 2\sin\alpha\cos\alpha\\ \cos(2\alpha) &= \cos^2\alpha-\sin^2\alpha\\ &= 2\cos^2\alpha-1\\ &= 1-2\sin^2\alpha\\ \tan(2\alpha) &= \frac{2\tan\alpha}{1-\tan^2\alpha}\quad(\tan\alpha\text{ 有意义且 }\tan\alpha\neq\pm1)\end{aligned}$$

2. **降幂公式**

$$\begin{aligned}\sin^2\alpha &= \frac{1-\cos(2\alpha)}{2}\\ \cos^2\alpha &= \frac{1+\cos(2\alpha)}{2}\end{aligned}$$

3. **半角公式**

$$\begin{aligned}\sin\frac{\alpha}{2} &= \pm\sqrt{\frac{1-\cos\alpha}{2}}\\ \cos\frac{\alpha}{2} &= \pm\sqrt{\frac{1+\cos\alpha}{2}}\\ \tan\frac{\alpha}{2} &= \pm\sqrt{\frac{1-\cos\alpha}{1+\cos\alpha}}\end{aligned}$$

在分母有意义时, 还可写成

$$\tan\frac{\alpha}{2}=\frac{\sin\alpha}{1+\cos\alpha}\quad(\cos\alpha\neq-1),\qquad \tan\frac{\alpha}{2}=\frac{1-\cos\alpha}{\sin\alpha}\quad(\sin\alpha\neq0).$$

半角公式根号前的正负号由角 $\frac{\alpha}{2}$ 所在的象限决定. 正切的后两个无根号形式在化简中更为常用, 但使用前必须检查分母是否为零.

证明　我们用两角和的三角函数公式推导这些公式.

先推二倍角公式. 在两角和公式中令两个角相等, 即设 $\beta=\alpha$. 对正弦函数, 由 $\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta$ 得

$$\sin(2\alpha)=\sin(\alpha+\alpha)=\sin\alpha\cos\alpha+\cos\alpha\sin\alpha=2\sin\alpha\cos\alpha;$$

对余弦函数, 由 $\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta$ 得

$$\cos(2\alpha)=\cos(\alpha+\alpha)=\cos^2\alpha-\sin^2\alpha,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

再结合平方关系 $\sin^2\alpha+\cos^2\alpha=1$, 得 $\cos(2\alpha)$ 的另外两种常用形式

$$\cos(2\alpha)=\cos^2\alpha-(1-\cos^2\alpha)=2\cos^2\alpha-1,$$

$$\cos(2\alpha)=(1-\sin^2\alpha)-\sin^2\alpha=1-2\sin^2\alpha;$$

对正切函数, 由商数关系及已证二倍角公式,

$$\tan(2\alpha)=\frac{\sin(2\alpha)}{\cos(2\alpha)}=\frac{2\sin\alpha\cos\alpha}{\cos^2\alpha-\sin^2\alpha},$$

为得到用 $\tan\alpha$ 表示的形式, 分子分母同除以 $\cos^2\alpha$(要求 $\cos\alpha\neq0$), 得

$$\tan(2\alpha)=\frac{2\frac{\sin\alpha}{\cos\alpha}}{1-\frac{\sin^2\alpha}{\cos^2\alpha}}=\frac{2\tan\alpha}{1-\tan^2\alpha}.$$

再推降幂公式. 它由余弦二倍角公式移项得到. 由 $\cos(2\alpha)=1-2\sin^2\alpha$ 移项得 $\sin^2\alpha=\frac{1-\cos(2\alpha)}{2}$; 同理由 $\cos(2\alpha)=2\cos^2\alpha-1$ 得 $\cos^2\alpha=\frac{1+\cos(2\alpha)}{2}$.

最后推半角公式. 在降幂公式中令角为 $\frac{\alpha}{2}$ 即可. 在 $\sin^2\theta=\frac{1-\cos(2\theta)}{2}$ 中令 $\theta=\frac{\alpha}{2}$, 则 $2\theta=\alpha$, 得

$$\sin^2\frac{\alpha}{2}=\frac{1-\cos\alpha}{2}\implies\sin\frac{\alpha}{2}=\pm\sqrt{\frac{1-\cos\alpha}{2}};$$

同理在 $\cos^2\theta=\frac{1+\cos(2\theta)}{2}$ 中作相同代换, 得

$$\cos^2\frac{\alpha}{2}=\frac{1+\cos\alpha}{2}\implies\cos\frac{\alpha}{2}=\pm\sqrt{\frac{1+\cos\alpha}{2}}.$$

正切的半角公式还有无根号形式. 利用商数关系及倍角、降幂公式,

$$\tan\frac{\alpha}{2}=\frac{\sin\frac{\alpha}{2}}{\cos\frac{\alpha}{2}}=\frac{2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}}{2\cos^2\frac{\alpha}{2}}=\frac{\sin\alpha}{1+\cos\alpha},$$

或分子分母同乘 $2\sin\frac{\alpha}{2}$, 得另一形式

$$\tan\frac{\alpha}{2}=\frac{\sin\frac{\alpha}{2}}{\cos\frac{\alpha}{2}}=\frac{2\sin^2\frac{\alpha}{2}}{2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}}=\frac{1-\cos\alpha}{\sin\alpha}.$$

这两个有理形式避免了开方和符号判断, 常在化简中使用. □

知乎@余命数

6.10.2 内在联系与应用策略

- **$\cos(2\alpha)$ 的三种形式**
  $\cos(2\alpha)$ 既能连起倍角与单角, 移项后又给出升幂、降幂公式, 是处理三角幂次的工具.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

- **角的相对性**
  倍角是相对的.α 是 $\frac{\alpha}{2}$ 的倍角，4α 是 2α 的倍角. 公式可以跨层次使用，例如 $\sin\alpha = 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}$，$\cos(4x) = 2\cos^2(2x) - 1$.
- **1 的代换与弦切互化**
  当式子中出现 1 时，可写成 $1 = \sin^2\alpha + \cos^2\alpha$ 来齐次化. 例如 $\dfrac{1-\cos\alpha}{\sin\alpha} = \tan\dfrac{\alpha}{2}$.

6.10.3 典例解析

例题 6.25　已知 $\tan\alpha = -\frac{1}{7}$，$\sin\beta = \frac{\sqrt{10}}{10}$ 且 β 为锐角，求 $\tan(2\alpha+\beta)$ 的值.

解　我们要求 $\tan(2\alpha+\beta)$，先求 $\tan 2\alpha$ 与 $\tan\beta$，再用正切和角公式.

已知 $\tan\alpha = -\frac{1}{7}$，由正切二倍角公式

$$\tan(2\alpha) = \frac{2\tan\alpha}{1-\tan^2\alpha} = \frac{2(-\frac{1}{7})}{1-(-\frac{1}{7})^2} = \frac{-\frac{2}{7}}{1-\frac{1}{49}} = \frac{-\frac{2}{7}}{\frac{48}{49}} = -\frac{7}{24}.$$

知乎@余命数

再求 $\tan\beta$. 由 β 为锐角知 $\cos\beta > 0$，故

$$\cos\beta = \sqrt{1-\sin^2\beta} = \sqrt{1-\left(\frac{\sqrt{10}}{10}\right)^2} = \sqrt{\frac{90}{100}} = \frac{3\sqrt{10}}{10},$$

从而 $\tan\beta = \frac{\sin\beta}{\cos\beta} = \frac{\sqrt{10}/10}{3\sqrt{10}/10} = \frac{1}{3}$. 最后由正切和角公式，

$$\tan(2\alpha+\beta) = \frac{\tan(2\alpha)+\tan\beta}{1-\tan(2\alpha)\tan\beta} = \frac{-\frac{7}{24}+\frac{1}{3}}{1-\left(-\frac{7}{24}\right)\left(\frac{1}{3}\right)}$$

$$= \frac{\frac{-7+8}{24}}{1+\frac{7}{72}} = \frac{\frac{1}{24}}{\frac{79}{72}} = \frac{1}{24}\cdot\frac{72}{79} = \frac{3}{79}.$$

因此 $\tan(2\alpha+\beta) = \frac{3}{79}$.

知乎@余命数

注 6.10 (变量准备) 本题需要先求 $\tan 2\alpha$ 与 $\tan\beta$，再把 $2\alpha+\beta$ 看作整体和角代入正切和角公式.

例题 6.26 (倍角公式的连乘技巧)（2024 年 1 月湖南邵阳高三第一次联考）已知 $\cos\theta\cos 2\theta\cos 4\theta = -\frac{1}{8}$，$\theta \in (\frac{\pi}{7}, \frac{\pi}{4})$，则 $2\cos^2 4\theta - \cos\theta =$ _____.

解 左侧是余弦的倍角连乘. 我们在等式两边同乘一个合适的正弦因子, 连续使用二倍角公式把它折叠成简洁方程. 设 $P=\cos\theta\cos 2\theta\cos 4\theta=-\frac{1}{8}$, 在等式两边同乘 $8\sin\theta$(由 $\theta\in(\frac{\pi}{7},\frac{\pi}{4})$ 知 $\sin\theta\neq 0$), 反复用 $2\sin x\cos x=\sin 2x$ 折叠,

$$\begin{aligned}
8\sin\theta\cos\theta\cos 2\theta\cos 4\theta &= -\sin\theta,\\
4(2\sin\theta\cos\theta)\cos 2\theta\cos 4\theta &= -\sin\theta,\\
4\sin 2\theta\cos 2\theta\cos 4\theta &= -\sin\theta,\\
2(2\sin 2\theta\cos 2\theta)\cos 4\theta &= -\sin\theta,\\
2\sin 4\theta\cos 4\theta &= -\sin\theta,\\
\sin 8\theta &= -\sin\theta \quad \text{(折叠完成)}.
\end{aligned}$$

由 $\sin 8\theta=-\sin\theta=\sin(-\theta)$ 及正弦函数的性质, 得

$$8\theta=-\theta+2k\pi \quad \text{或} \quad 8\theta=\pi-(-\theta)+2k\pi,\quad k\in\mathbb{Z},$$

即 $9\theta=2k\pi\implies\theta=\frac{2k\pi}{9}$, 或 $7\theta=(2k+1)\pi\implies\theta=\frac{(2k+1)\pi}{7}$. 将 $\theta\in(\frac{\pi}{7},\frac{\pi}{4})$ 即 $\theta\in(\frac{4\pi}{28},\frac{7\pi}{28})$ 代入检验: 对 $\theta=\frac{2k\pi}{9}$, 取 $k=1$ 得 $\theta=\frac{2\pi}{9}=\frac{56\pi}{252}$, 而 $\frac{\pi}{7}=\frac{36\pi}{252}$, $\frac{\pi}{4}=\frac{63\pi}{252}$, 由 $36<56<63$ 知 $\theta=\frac{2\pi}{9}$ 是一个解; 对 $\theta=\frac{(2k+1)\pi}{7}$, $k=0$ 给出 $\theta=\frac{\pi}{7}$ 不在区间内, $k\geq 1$ 给出 $\theta\geq\frac{3\pi}{7}>\frac{\pi}{4}$. 故唯一解为 $\theta=\frac{2\pi}{9}$.

所求表达式 $2\cos^2 4\theta-\cos\theta$ 用降幂公式 $2\cos^2 x=\cos(2x)+1$ 化为 $\cos(8\theta)+1-\cos\theta$. 由 $9\theta=2\pi$ 知 $8\theta=2\pi-\theta$, 故 $\cos(8\theta)=\cos(2\pi-\theta)=\cos\theta$, 代入得

$$\cos\theta+1-\cos\theta=1.$$

注 6.11 (经验总结)
形如 $\cos x\cos(2x)\cos(4x)\cdots$ 的连乘式, 常用方法是两边同乘最简角的正弦(这里为 $\sin\theta$), 连续使用 $2\sin\alpha\cos\alpha=\sin(2\alpha)$, 把式子逐步折叠成一个简洁的方程.

例题 6.27 (2024 年广州广雅中学二调) 已知锐角 θ 满足 $2\cos 2\theta=1+\sin 2\theta$, 则 $\tan\theta=$(　)

A. $\frac{1}{3}$

B. $\frac{1}{2}$

C. 2

D. 3

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

解 方程两边都含倍角 2θ, 先把它化为只含单角 θ 的式子. 左边用余弦二倍角公式,

$$2\cos 2\theta = 2(\cos^2\theta - \sin^2\theta) = 2(\cos\theta - \sin\theta)(\cos\theta + \sin\theta);$$

右边把 1 写成 $\sin^2\theta + \cos^2\theta$ 后与 $\sin 2\theta$ 配成完全平方,

$$1 + \sin 2\theta = (\sin^2\theta + \cos^2\theta) + 2\sin\theta\cos\theta = (\sin\theta + \cos\theta)^2.$$

于是原方程化为

$$2(\cos\theta - \sin\theta)(\cos\theta + \sin\theta) = (\sin\theta + \cos\theta)^2.$$

由 θ 是锐角知 $\sin\theta + \cos\theta \neq 0$, 两边可同除以 $\sin\theta + \cos\theta$, 得

$$2(\cos\theta - \sin\theta) = \sin\theta + \cos\theta,$$

$$2\cos\theta - 2\sin\theta = \sin\theta + \cos\theta,$$

$$\cos\theta = 3\sin\theta.$$

再在 $\cos\theta = 3\sin\theta$ 两边同除以 $\cos\theta$(θ 为锐角, $\cos\theta \neq 0$), 得

$$1 = 3\frac{\sin\theta}{\cos\theta} \implies \tan\theta = \frac{1}{3}.$$

答案为 A.

注 6.12 (经验总结)
解这道题, 要活用 $1 \pm \sin 2\theta = (\sin\theta \pm \cos\theta)^2$.

- 常数 1 可写成 $\sin^2\theta + \cos^2\theta$, 与 $\sin 2\theta$ 搭配后得到完全平方.
- 把 2θ 统一化为单角 θ, 便于与另一边比较.
- 两边约去 $\sin\theta + \cos\theta$ 前, 先说明它不为零; 这里锐角条件正好保证这一点.

例题 6.28 (2023 年 11 月湖北省部分高中联考协作体期中考) 若 $\theta \in (0, \frac{\pi}{2})$, $\dfrac{\sin\theta}{2\cos^2\theta} = \dfrac{\sin\theta - \cos\theta}{1 - \sin 2\theta}$, 则 $\tan\theta =$ _____.

解 等式两边的分母都可化为单角 θ 的形式, 先统一角, 再化简方程. 左边分母 $2\cos^2\theta$ 已是单角形式; 右边分母可配方,

$$1 - \sin 2\theta = \sin^2\theta + \cos^2\theta - 2\sin\theta\cos\theta = (\sin\theta - \cos\theta)^2,$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

代入原方程得

$$\frac{\sin\theta}{2\cos^2\theta}=\frac{\sin\theta-\cos\theta}{(\sin\theta-\cos\theta)^2}.$$

化简右式前须考虑分母 $\sin\theta-\cos\theta$ 是否为零，若 $\sin\theta-\cos\theta=0$，则 $\tan\theta=1$，$\theta=\frac{\pi}{4}$，此时 $1-\sin 2\theta=0$，原方程右边分母为零而无意义，故 $\theta=\frac{\pi}{4}$ 不在取值范围内，$\sin\theta-\cos\theta\neq 0$，可约分得

$$\frac{\sin\theta}{2\cos^2\theta}=\frac{1}{\sin\theta-\cos\theta}.$$

交叉相乘，

$$\sin\theta(\sin\theta-\cos\theta)=2\cos^2\theta,$$
$$\sin^2\theta-\sin\theta\cos\theta-2\cos^2\theta=0,$$

这是关于 $\sin\theta,\cos\theta$ 的二次齐次方程. 由 $\theta\in(0,\frac{\pi}{2})$ 知 $\cos\theta\neq 0$，两边同除以 $\cos^2\theta$，把方程化为关于 $\tan\theta$ 的方程，

$$\tan^2\theta-\tan\theta-2=0,$$

因式分解得 $(\tan\theta-2)(\tan\theta+1)=0$，解得 $\tan\theta=2$ 或 $\tan\theta=-1$. 由 $\theta\in(0,\frac{\pi}{2})$ 知 $\tan\theta>0$，舍去 $\tan\theta=-1$.

答案为 2.

6.10.4 三倍角与 n 倍角公式(拓展)

来龙去脉

和角公式与二倍角公式能处理 $\alpha\pm\beta$ 与 2α. 把 3α 拆成 $2\alpha+\alpha$，就能用同样的办法推出三倍角公式.

三倍角公式无需死记硬背，但它的推导展示出如何把未知角拆成已知角的和，再把结果化归到单一三角函数. 这种递推思路还可以继续推广到任意整数倍角 $n\alpha$.

定理 6.6 (三倍角公式)

$$\sin(3\alpha)=3\sin\alpha-4\sin^3\alpha$$
$$\cos(3\alpha)=4\cos^3\alpha-3\cos\alpha$$

证明 把 3α 写成 $2\alpha+\alpha$，再用和角公式与二倍角公式展开. 先推 $\sin(3\alpha)$，

$$\sin(3\alpha)=\sin(2\alpha+\alpha)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
&= \sin(2\alpha)\cos\alpha + \cos(2\alpha)\sin\alpha && \text{(应用和角公式)}\\
&= (2\sin\alpha\cos\alpha)\cos\alpha + (1-2\sin^2\alpha)\sin\alpha \\
&\qquad\qquad \text{(展开 } 2\alpha\text{, 注意 } \cos(2\alpha) \text{ 选用含 sin 的形式)}\\
&= 2\sin\alpha\cos^2\alpha + \sin\alpha - 2\sin^3\alpha && \text{(化简)}\\
&= 2\sin\alpha(1-\sin^2\alpha) + \sin\alpha - 2\sin^3\alpha && \text{(将 } \cos^2\alpha \text{ 换掉, 实现“归一”)}\\
&= 2\sin\alpha - 2\sin^3\alpha + \sin\alpha - 2\sin^3\alpha \\
&= 3\sin\alpha - 4\sin^3\alpha.
\end{aligned}
$$

类似地推 $\cos(3\alpha)$, 把结果只保留 $\cos\alpha$,

$$
\begin{aligned}
\cos(3\alpha) &= \cos(2\alpha+\alpha) \\
&= \cos(2\alpha)\cos\alpha - \sin(2\alpha)\sin\alpha && \text{(应用和角公式)}\\
&= (2\cos^2\alpha - 1)\cos\alpha - (2\sin\alpha\cos\alpha)\sin\alpha \\
&\qquad\qquad \text{(展开 } 2\alpha\text{, 注意 } \cos(2\alpha) \text{ 选用含 cos 的形式)}\\
&= 2\cos^3\alpha - \cos\alpha - 2\cos\alpha\sin^2\alpha && \text{(化简)}\\
&= 2\cos^3\alpha - \cos\alpha - 2\cos\alpha(1-\cos^2\alpha) && \text{(将 } \sin^2\alpha \text{ 换掉, 实现“归一”)}\\
&= 2\cos^3\alpha - \cos\alpha - 2\cos\alpha + 2\cos^3\alpha \\
&= 4\cos^3\alpha - 3\cos\alpha.
\end{aligned}
$$

□

6.10.4.1 推广至 n 倍角

三倍角公式启发我们, $\cos(n\alpha)$ 可以表示为关于 $\cos\alpha$ 的多项式, 并且存在递推关系. 考虑和差角公式.

$$\cos(n\alpha+\alpha) = \cos(n\alpha)\cos\alpha - \sin(n\alpha)\sin\alpha$$

$$\cos(n\alpha-\alpha) = \cos(n\alpha)\cos\alpha + \sin(n\alpha)\sin\alpha$$

将以上两式相加, 可得

$$\cos((n+1)\alpha) + \cos((n-1)\alpha) = 2\cos(n\alpha)\cos\alpha$$

整理后得到递推公式.

$$\cos((n+1)\alpha) = 2\cos\alpha\cdot\cos(n\alpha) - \cos((n-1)\alpha) \tag{6-7}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

这个公式表明，只要知道相邻两项 $\cos((n-1)\alpha)$ 和 $\cos(n\alpha)$，就可以推出下一项 $\cos((n+1)\alpha)$.

验证一下.

- 令 $n=1$，递推式变为 $\cos(2\alpha)=2\cos\alpha\cdot\cos\alpha-\cos(0\alpha)=2\cos^2\alpha-1$. 与二倍角公式完全吻合.
- 令 $n=2$，递推式变为 $\cos(3\alpha)=2\cos\alpha\cdot\cos(2\alpha)-\cos\alpha$. 将 $\cos(2\alpha)=2\cos^2\alpha-1$ 代入，得

$$\begin{aligned}\cos(3\alpha)&=2\cos\alpha(2\cos^2\alpha-1)-\cos\alpha\\&=4\cos^3\alpha-2\cos\alpha-\cos\alpha\\&=4\cos^3\alpha-3\cos\alpha\end{aligned}$$

与我们刚刚推导的三倍角公式完全吻合.

例题 6.29 求值 $\dfrac{3\sin 15^\circ-4\sin^3 15^\circ}{4\cos^3 15^\circ-3\cos 15^\circ}$.

解 分子、分母的结构分别对应正弦、余弦三倍角公式，直接套用即可. 由正弦三倍角公式 $\sin(3\alpha)=3\sin\alpha-4\sin^3\alpha$ 取 $\alpha=15^\circ$，得

$$3\sin 15^\circ-4\sin^3 15^\circ=\sin(3\times 15^\circ)=\sin 45^\circ=\frac{\sqrt{2}}{2};$$

由余弦三倍角公式 $\cos(3\alpha)=4\cos^3\alpha-3\cos\alpha$ 取 $\alpha=15^\circ$，得

$$4\cos^3 15^\circ-3\cos 15^\circ=\cos(3\times 15^\circ)=\cos 45^\circ=\frac{\sqrt{2}}{2}.$$

将分子、分母代回原分式，

$$\text{原式}=\frac{\sin 45^\circ}{\cos 45^\circ}=\tan 45^\circ=1.$$

定理 6.7 (正切的三倍角公式)

$$\tan(3\alpha)=\frac{3\tan\alpha-\tan^3\alpha}{1-3\tan^2\alpha}$$

该公式要求 $\tan\alpha$ 与 $\tan(3\alpha)$ 均有意义，且 $1-3\tan^2\alpha\neq 0$.

证明 推导正切的三倍角公式有两种思路.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

其一, 利用商数关系, 得

$$\tan(3\alpha)=\frac{\sin(3\alpha)}{\cos(3\alpha)}=\frac{3\sin\alpha-4\sin^3\alpha}{4\cos^3\alpha-3\cos\alpha},$$

知乎@余命数

分子分母同除以 $\cos^3\alpha$(要求 $\cos\alpha\neq0$),

$$\begin{aligned}\tan(3\alpha)&=\frac{3\dfrac{\sin\alpha}{\cos\alpha}\dfrac{1}{\cos^2\alpha}-4\dfrac{\sin^3\alpha}{\cos^3\alpha}}{4-3\dfrac{1}{\cos^2\alpha}}\\&=\frac{3\tan\alpha(1+\tan^2\alpha)-4\tan^3\alpha}{4-3(1+\tan^2\alpha)}\\&=\frac{3\tan\alpha-\tan^3\alpha}{1-3\tan^2\alpha},\end{aligned}$$

上式分母 $1-3\tan^2\alpha$ 非零, 正对应 $\cos(3\alpha)\neq0$, 即 $\tan(3\alpha)$ 有意义.

知乎@余命数

其二, 利用和角公式,

$$\begin{aligned}\tan(3\alpha)&=\tan(2\alpha+\alpha)\\&=\frac{\tan(2\alpha)+\tan\alpha}{1-\tan(2\alpha)\tan\alpha}&&\text{(应用和角公式)}\\&=\frac{\frac{2\tan\alpha}{1-\tan^2\alpha}+\tan\alpha}{1-\frac{2\tan\alpha}{1-\tan^2\alpha}\cdot\tan\alpha}&&\text{(展开 }\tan(2\alpha)\text{)}\\&=\frac{\frac{2\tan\alpha+\tan\alpha(1-\tan^2\alpha)}{1-\tan^2\alpha}}{\frac{(1-\tan^2\alpha)-2\tan^2\alpha}{1-\tan^2\alpha}}&&\text{(通分)}\\&=\frac{2\tan\alpha+\tan\alpha-\tan^3\alpha}{1-3\tan^2\alpha}\\&=\frac{3\tan\alpha-\tan^3\alpha}{1-3\tan^2\alpha}.\end{aligned}$$

知乎@余命数 知乎@余命数

□

例题 6.30 (2024 年 2 月青桐鸣联考) 已知 $\tan\alpha=\dfrac{\sqrt{2}}{2}$, 求 $\tan3\alpha$ 的值, $\tan3\alpha=$ _____.

知乎@余命数

解 可直接代入正切三倍角公式, 也可分步用 $3\alpha=2\alpha+\alpha$ 计算. 现用分步法.

法一采用分步计算. 先由正切二倍角公式,

$$\tan(2\alpha)=\frac{2\tan\alpha}{1-\tan^2\alpha}=\frac{2\cdot\frac{\sqrt{2}}{2}}{1-\left(\frac{\sqrt{2}}{2}\right)^2}=\frac{\sqrt{2}}{1-\frac{2}{4}}=\frac{\sqrt{2}}{\frac{1}{2}}=2\sqrt{2};$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再由正切和角公式求 $\tan(3\alpha)=\tan(2\alpha+\alpha)$,

$$\tan(3\alpha)=\frac{\tan(2\alpha)+\tan\alpha}{1-\tan(2\alpha)\tan\alpha}=\frac{2\sqrt{2}+\frac{\sqrt{2}}{2}}{1-(2\sqrt{2})\left(\frac{\sqrt{2}}{2}\right)}=\frac{\frac{5\sqrt{2}}{2}}{1-2}=-\frac{5\sqrt{2}}{2}.$$

法二直接代入三倍角公式. 将 $\tan\alpha=\frac{\sqrt{2}}{2}$ 代入 $\tan(3\alpha)=\dfrac{3\tan\alpha-\tan^3\alpha}{1-3\tan^2\alpha}$,

$$\tan(3\alpha)=\frac{3\left(\frac{\sqrt{2}}{2}\right)-\left(\frac{\sqrt{2}}{2}\right)^3}{1-3\left(\frac{\sqrt{2}}{2}\right)^2}=\frac{\frac{3\sqrt{2}}{2}-\frac{\sqrt{2}}{4}}{1-\frac{3}{2}}=\frac{\frac{5\sqrt{2}}{4}}{-\frac{1}{2}}=-\frac{5\sqrt{2}}{2}.$$

两种方法结果一致, 答案为 $-\dfrac{5\sqrt{2}}{2}$.

6.10.5 高考真题选讲

以下三题均用倍角公式将 $\cos 2x$ 或 $\sin^2 x$ 降幂, 再以换元 $t=\cos x$ 或 $t=\sin x$ 转化为闭区间上二次函数的最值问题.

例题 6.31 (2021 年北京卷第 7 题) 函数 $f(x)=\cos x-\cos 2x$ 的奇偶性及最大值为

A. 奇函数, 最大值为 2　　C. 奇函数, 最大值为 $\frac{9}{8}$

B. 偶函数, 最大值为 2　　D. 偶函数, 最大值为 $\frac{9}{8}$

解 答案:D.

由 $f(-x)=\cos(-x)-\cos(-2x)=\cos x-\cos 2x=f(x)$ 知 $f(x)$ 为偶函数.

用二倍角公式 $\cos 2x=2\cos^2 x-1$ 降幂,

$$f(x)=\cos x-(2\cos^2 x-1)=-2\cos^2 x+\cos x+1.$$

令 $t=\cos x\in[-1,1]$, 则

$$g(t)=-2t^2+t+1=-2\left(t-\frac{1}{4}\right)^2+\frac{9}{8}.$$

对称轴 $t=\dfrac{1}{4}\in[-1,1]$, 故 $f(x)$ 的最大值为 $\dfrac{9}{8}$.

例题 6.32 (2017 年全国 □ 卷理科数学第 14 题) 函数 $f(x)=\sin^2 x+\sqrt{3}\cos x-\frac{3}{4}\left(x\in\left[0,\frac{\pi}{2}\right]\right)$ 的最大值是_____.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 答案:1.

用 $\sin^2 x = 1 - \cos^2 x$ 降幂,

$$f(x) = 1 - \cos^2 x + \sqrt{3}\cos x - \frac{3}{4} = -\cos^2 x + \sqrt{3}\cos x + \frac{1}{4}.$$

令 $t = \cos x$. 由 $x \in \left[0, \frac{\pi}{2}\right]$ 得 $t \in [0, 1]$.

$$g(t) = -t^2 + \sqrt{3}t + \frac{1}{4} = -\left(t - \frac{\sqrt{3}}{2}\right)^2 + 1.$$

对称轴 $t = \frac{\sqrt{3}}{2} \in [0, 1]$, 故最大值为 1.

例题 6.33 (2016 年全国 □ 卷文科数学第 11 题) 函数 $f(x) = \cos 2x + 6\cos\left(\frac{\pi}{2} - x\right)$ 的最大值为

A. 4　　C. 6

B. 5　　D. 7

解 答案:B.

由诱导公式 $\cos\left(\frac{\pi}{2} - x\right) = \sin x$ 及二倍角公式 $\cos 2x = 1 - 2\sin^2 x$,

$$f(x) = 1 - 2\sin^2 x + 6\sin x.$$

令 $t = \sin x \in [-1, 1]$, 则

$$g(t) = -2t^2 + 6t + 1 = -2\left(t - \frac{3}{2}\right)^2 + \frac{11}{2}.$$

对称轴 $t = \frac{3}{2} \notin [-1, 1]$, 抛物线开口向下, 故 $g(t)$ 在 $[-1, 1]$ 上单调递增, 最大值在 $t = 1$ 处取得:$g(1) = -2 + 6 + 1 = 5$.

6.11 万能公式与齐次式思想

来龙去脉

在处理三角函数问题时, 同一个方程或表达式里常同时出现 $\sin\alpha, \cos\alpha, \tan\alpha$ 等不同三角函数. 化简这类问题的首要步骤, 是把它们统一成同类函数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

已用弦化切处理过**齐次式**. 把这个思路再推进一步. 若能用一个单一的代数变量把所有三角函数都表示出来, 那么任何三角方程都能化为纯代数方程, 从而避开三角函数特有的变形.

万能公式说明这样的表示是可行的. 它以二倍角公式和 $1=\sin^2\frac{\alpha}{2}+\cos^2\frac{\alpha}{2}$ 为根据, 把 $\sin\alpha,\cos\alpha,\tan\alpha$ 都写成半角正切 $t=\tan\frac{\alpha}{2}$ 的有理分式, 完成了三角函数名之间的统一.

之所以冠以“万能”之名, 在于它的适用范围不限于某一类三角方程. 只要一个表达式由 $\sin\alpha,\cos\alpha,\tan\alpha$ 经四则运算复合而成, 代入万能公式后所得的必然是关于 t 的有理分式. 有理分式的化简与求值只依赖代数运算, 不再涉及任何三角恒等变形. 一个代换把整类问题归约为代数问题, “万能”二字正由此而来.

这一代换最早服务于积分学. 对于形如 $\int R(\sin x,\cos x)\,dx$ 的有理三角积分, 令 $t=\tan\frac{x}{2}$ 可将被积函数连同 dx 一并化为 t 的有理函数, 从而把超越的积分问题转化为有理函数的积分. 这一技巧以德国数学家魏尔斯特拉斯的名字命名, 称为**魏尔斯特拉斯代换**, 其积分应用见 第17.3节.

6.11.1 公式体系

定理 6.8 (万能公式) 当 $\alpha\neq(2k+1)\pi$ $(k\in\mathbb{Z})$ 时, 令 $t=\tan\frac{\alpha}{2}$, 则有

$$\sin\alpha=\frac{2t}{1+t^2} \tag{6-8}$$

$$\cos\alpha=\frac{1-t^2}{1+t^2} \tag{6-9}$$

$$\tan\alpha=\frac{2t}{1-t^2}\quad(\text{当}t\neq\pm1\text{ 时}) \tag{6-10}$$

6.11.2 齐次化构造

证明 (推导) 先推 $\sin\alpha$. 从二倍角公式出发, 利用 $1=\sin^2\frac{\alpha}{2}+\cos^2\frac{\alpha}{2}$ 把分子改造成齐次分式,

$$\begin{aligned}\sin\alpha&=\sin\left(2\cdot\frac{\alpha}{2}\right)=2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}\\&=\frac{2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}}{1}&&\text{(除以 1)}\\&=\frac{2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}}{\sin^2\frac{\alpha}{2}+\cos^2\frac{\alpha}{2}}.&&\text{(1 的代换, 构造齐次分式)}\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

此时分子分母都是关于 $\sin\frac{\alpha}{2},\cos\frac{\alpha}{2}$ 的二次齐次式, 同除以 $\cos^2\frac{\alpha}{2}$(设 $\cos\frac{\alpha}{2}\neq 0$),

$$\sin\alpha=\frac{\frac{2\sin(\alpha/2)\cos(\alpha/2)}{\cos^2(\alpha/2)}}{\frac{\sin^2(\alpha/2)+\cos^2(\alpha/2)}{\cos^2(\alpha/2)}}=\frac{2\frac{\sin(\alpha/2)}{\cos(\alpha/2)}}{\frac{\sin^2(\alpha/2)}{\cos^2(\alpha/2)}+1}=\frac{2\tan\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}=\frac{2t}{1+t^2}.$$

再推 $\cos\alpha$, 采用相同策略. 由二倍角公式及 1 的代换,

$$\begin{aligned}\cos\alpha&=\cos\left(2\cdot\frac{\alpha}{2}\right)=\cos^2\frac{\alpha}{2}-\sin^2\frac{\alpha}{2}\\&=\frac{\cos^2\frac{\alpha}{2}-\sin^2\frac{\alpha}{2}}{1}=\frac{\cos^2\frac{\alpha}{2}-\sin^2\frac{\alpha}{2}}{\sin^2\frac{\alpha}{2}+\cos^2\frac{\alpha}{2}},\end{aligned}$$

分子分母同除以 $\cos^2\frac{\alpha}{2}$, 得

$$\cos\alpha=\frac{1-\frac{\sin^2(\alpha/2)}{\cos^2(\alpha/2)}}{1+\frac{\sin^2(\alpha/2)}{\cos^2(\alpha/2)}}=\frac{1-\tan^2\frac{\alpha}{2}}{1+\tan^2\frac{\alpha}{2}}=\frac{1-t^2}{1+t^2}.$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

最后推 $\tan\alpha$. 当 $\cos\alpha\neq 0$ 时, 由商数关系及刚得的两式,

$$\tan\alpha=\frac{\sin\alpha}{\cos\alpha}=\frac{\frac{2t}{1+t^2}}{\frac{1-t^2}{1+t^2}}=\frac{2t}{1-t^2},$$

此处要求 $1-t^2\neq 0$, 这正对应 $\cos\alpha\neq 0$. □

万能公式提供了一条纯代数的解题路径. 它的计算量通常较大, 且要注意 $\alpha=(2k+1)\pi$ 时 $\tan\frac{\alpha}{2}$ 无定义. 因此它在其他方法, 如辅助角公式不易见效时, 是一种可靠但繁琐的备选.

6.11.3 典例解析

知乎@余命数

例题 6.34 (解三角方程) 求解方程 $\sin\theta+\cos\theta=1$.

解 此题可用多种方法求解, 下用万能公式法演示. 令 $t=\tan\frac{\theta}{2}$, 将万能公式代入原方程,

$$\frac{2t}{1+t^2}+\frac{1-t^2}{1+t^2}=1.$$

知乎@余命数

因分母 $1+t^2>0$ 恒成立, 可直接去分母,

知乎@余命数

$$2t+1-t^2=1+t^2\implies 2t^2-2t=0\implies 2t(t-1)=0,$$

解得 $t=0$ 或 $t=1$. 由 $t=\tan\frac{\theta}{2}$ 反解: $t=0$ 给出 $\frac{\theta}{2}=k\pi$, 即 $\theta=2k\pi$; $t=1$ 给出 $\frac{\theta}{2}=k\pi+\frac{\pi}{4}$, 即 $\theta=2k\pi+\frac{\pi}{2}$ ($k\in\mathbb{Z}$).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

万能公式的使用前提是 $\cos\frac{\theta}{2}\neq 0$,即 $\theta\neq(2k+1)\pi$,故须单独检验 $\theta=(2k+1)\pi$,当 $\theta=\pi$ 时 $\sin\pi+\cos\pi=0-1=-1\neq 1$,这些点不是解.

原方程的解集为 $\{\theta\mid\theta=2k\pi$ 或 $\theta=2k\pi+\frac{\pi}{2},k\in\mathbb{Z}\}$.

例题 6.35 (2024 年 1 月浙江省宁波九校联考) 已知 $\tan A=\dfrac{1}{2}$,求

$$\frac{\cos 4A-4\cos 2A+3}{\cos 4A+4\cos 2A+3}$$

的值(　)

A. $\dfrac{1}{16}$

B. $\dfrac{1}{8}$

C. $\dfrac{1}{4}$

D. $\dfrac{1}{2}$

解 题目给出单角的正切值,要求一个含其二倍角和四倍角的复杂分式的值,宜用由切求弦的思路,把所有三角函数都用 $\tan A$ 表示.

由万能公式的余弦表达式直接求 $\cos 2A$,

$$\cos 2A=\frac{1-\tan^2 A}{1+\tan^2 A}=\frac{1-(\frac{1}{2})^2}{1+(\frac{1}{2})^2}=\frac{\frac{3}{4}}{\frac{5}{4}}=\frac{3}{5}.$$

再把 $2A$ 视为整体,用余弦二倍角公式 $\cos(2\theta)=2\cos^2\theta-1$ 求 $\cos 4A$,

$$\cos 4A=\cos(2\cdot 2A)=2\cos^2(2A)-1=2\left(\frac{3}{5}\right)^2-1=2\left(\frac{9}{25}\right)-1=\frac{18}{25}-1=-\frac{7}{25}.$$

将 $\cos 2A=\frac{3}{5}$ 与 $\cos 4A=-\frac{7}{25}$ 代入原分式,

$$\begin{aligned}
\text{原式}&=\frac{(-\frac{7}{25})-4(\frac{3}{5})+3}{(-\frac{7}{25})+4(\frac{3}{5})+3}=\frac{-\frac{7}{25}-\frac{12}{5}+3}{-\frac{7}{25}+\frac{12}{5}+3},\\
\text{分子}&=\frac{-7-60+75}{25}=\frac{8}{25},\\
\text{分母}&=\frac{-7+60+75}{25}=\frac{128}{25},\\
\text{原式}&=\frac{8/25}{128/25}=\frac{8}{128}=\frac{1}{16}.
\end{aligned}$$

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

另法：令 $x=\cos 2A$，将 $\cos 4A=2x^2-1$ 代入，分子为 $2x^2-1-4x+3=2(x-1)^2$，分母为 $2x^2+4x+2=2(x+1)^2$，故原式 $=\frac{2(\cos 2A-1)^2}{2(\cos 2A+1)^2}=\left(\frac{\cos 2A-1}{\cos 2A+1}\right)^2$；代入 $\cos 2A=3/5$ 得 $\left(\frac{-2/5}{8/5}\right)^2=\frac{1}{16}$.

两种方法都得 $\frac{1}{16}$，故选 A.

例题 6.36 (2023 年 12 月福建百校 & 河北金科联考) 若 $\tan\left(\alpha+\dfrac{\pi}{4}\right)=-3$，$\tan\beta=3$，求

$$\frac{\cos(\alpha-\beta)}{\sin(\alpha+\beta)}$$

的值(　)

A. -1

B. $\dfrac{7}{5}$

C. $\dfrac{3}{5}$

D. $\dfrac{4}{5}$

解 题目给出两个角的正切相关值，要求解一个复杂的三角分式. 直接展开会很繁琐，该分式的结构正适合弦化切.

先求 $\tan\alpha$. 由 $\tan(\alpha+\frac{\pi}{4})=-3$ 及和角公式，

$$\frac{\tan\alpha+1}{1-\tan\alpha}=-3,$$

解得 $\tan\alpha+1=-3(1-\tan\alpha)=-3+3\tan\alpha$，即 $2\tan\alpha=4$，故 $\tan\alpha=2$.

再对目标分式弦化切. 将其分子分母分别用和差角公式展开，

$$\text{原式}=\frac{\cos\alpha\cos\beta+\sin\alpha\sin\beta}{\sin\alpha\cos\beta+\cos\alpha\sin\beta},$$

这是关于 $\sin,\cos$ 的伪齐次式，分子分母同除以 $\cos\alpha\cos\beta$（设其不为 0），

$$\begin{aligned}\text{原式}&=\frac{\frac{\cos\alpha\cos\beta+\sin\alpha\sin\beta}{\cos\alpha\cos\beta}}{\frac{\sin\alpha\cos\beta+\cos\alpha\sin\beta}{\cos\alpha\cos\beta}}\\&=\frac{1+\frac{\sin\alpha}{\cos\alpha}\frac{\sin\beta}{\cos\beta}}{\frac{\sin\alpha}{\cos\alpha}+\frac{\sin\beta}{\cos\beta}}=\frac{1+\tan\alpha\tan\beta}{\tan\alpha+\tan\beta}.\end{aligned}$$

代入 $\tan\alpha=2$ 与 $\tan\beta=3$，

$$\text{原式}=\frac{1+(2)(3)}{2+3}=\frac{1+6}{5}=\frac{7}{5}.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

正确答案为 B.

例题 6.37 (2023 年 11 月江淮十校第二次联考) 已知角 θ 为第二象限角, 且满足 $\sin(\theta-\frac{\pi}{3})\sin(\pi+\theta)=\cos 2\theta$, 则 $\tan\theta=$(　)

A. $\dfrac{\sqrt{3}-\sqrt{11}}{2}$

B. $\dfrac{\sqrt{3}-\sqrt{7}}{2}$

C. $\dfrac{-\sqrt{3}-\sqrt{11}}{2}$

D. $\dfrac{-\sqrt{3}-\sqrt{7}}{2}$

解 这个恒等式把所有三角函数都化为单角 θ 的正弦和余弦后, 即可整理成齐次方程, 再弦化切求解.

用和差角公式、诱导公式与倍角公式把方程各项展开,

$$\sin\left(\theta-\frac{\pi}{3}\right)=\sin\theta\cos\frac{\pi}{3}-\cos\theta\sin\frac{\pi}{3}=\frac{1}{2}\sin\theta-\frac{\sqrt{3}}{2}\cos\theta,$$

$$\sin(\pi+\theta)=-\sin\theta,$$

$$\cos(2\theta)=\cos^2\theta-\sin^2\theta,$$

代入原方程,

$$\left(\frac{1}{2}\sin\theta-\frac{\sqrt{3}}{2}\cos\theta\right)(-\sin\theta)=\cos^2\theta-\sin^2\theta.$$

展开并化简, 两边同乘 2 后移项至同一侧, 得到关于 $\sin\theta,\cos\theta$ 的二次齐次方程,

$$-\frac{1}{2}\sin^2\theta+\frac{\sqrt{3}}{2}\sin\theta\cos\theta=\cos^2\theta-\sin^2\theta,$$

$$-\sin^2\theta+\sqrt{3}\sin\theta\cos\theta=2\cos^2\theta-2\sin^2\theta\quad(\text{两边同乘 }2),$$

$$\sin^2\theta+\sqrt{3}\sin\theta\cos\theta-2\cos^2\theta=0\quad(\text{移项合并}).$$

因为 θ 是第二象限角, $\cos\theta\neq 0$, 两边同除以 $\cos^2\theta$ 弦化切, 化为关于 $\tan\theta$ 的一元二次方程, 再用求根公式,

$$\tan^2\theta+\sqrt{3}\tan\theta-2=0,$$

$$\tan\theta=\frac{-\sqrt{3}\pm\sqrt{(\sqrt{3})^2-4(1)(-2)}}{2(1)}=\frac{-\sqrt{3}\pm\sqrt{3+8}}{2}=\frac{-\sqrt{3}\pm\sqrt{11}}{2}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

θ 为第二象限角, 其正切必为负: $\frac{-\sqrt{3}+\sqrt{11}}{2}$ 因 $\sqrt{11}>\sqrt{3}$ 而为正, 舍去; $\frac{-\sqrt{3}-\sqrt{11}}{2}$ 为负, 符合题意.

故正确答案为 C.

例题 6.38 (2024 届长郡中学月考) 已知 $\alpha\in(0,\frac{\pi}{2})$, 且 $3\sqrt{2}\sin(\alpha-\frac{\pi}{4})=\cos\alpha$, 则 $\sin\alpha+2\cos\alpha=$(　)

A. -2

B. $\dfrac{25}{16}$

C. $-\dfrac{25}{16}$

D. 2

解 先由方程解出 α 的某个三角函数值,再代入目标式. 展开 $3\sqrt{2}\sin(\alpha-\frac{\pi}{4})=\cos\alpha$:

$$3\sqrt{2}\Big(\sin\alpha\cdot\frac{\sqrt{2}}{2}-\cos\alpha\cdot\frac{\sqrt{2}}{2}\Big)=\cos\alpha,$$

$$3(\sin\alpha-\cos\alpha)=\cos\alpha,$$

$$3\sin\alpha=4\cos\alpha.$$

因 $\alpha\in(0,\frac{\pi}{2})$, $\cos\alpha\neq0$, 故 $\tan\alpha=\frac{4}{3}$.

法一采用求弦代入. 由 $\tan\alpha=\frac{4}{3}$ 与 α 锐角, 构造边长 $3,4,5$ 的直角三角形, 得 $\sin\alpha=\frac{4}{5}$, $\cos\alpha=\frac{3}{5}$, 于是

$$\sin\alpha+2\cos\alpha=\frac{4}{5}+\frac{6}{5}=2.$$

法二采用 tan 表达弦. 在 $\alpha\in(0,\frac{\pi}{2})$ 下,

$$\sin\alpha=\frac{\tan\alpha}{\sqrt{1+\tan^2\alpha}}=\frac{4/3}{5/3}=\frac{4}{5},\qquad\cos\alpha=\frac{1}{\sqrt{1+\tan^2\alpha}}=\frac{3}{5},$$

正确答案为 D.

6.11.4 基本思路

上面的例题反复用到以下四条思路.

- **思想一**(配凑与拆解)
  已知角与目标角不一致时, 需要构造它们之间的联系.
  - **拼凑角**, 如已知 $\tan(\alpha-\beta)$ 与 $\tan\beta$ 求 $\tan\alpha$, 可用 $\alpha=(\alpha-\beta)+\beta$.
  - **拆分角**, 如推导三倍角公式时, 把 3α 写成 $2\alpha+\alpha$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **互补与互余**, 在三角形中, 利用 $A+B+C=\pi$ 构造 $\pi-C$ 或 $\frac{\pi}{2}-\frac{C}{2}$ 等形式.

角的变换最终都是把未知角用已知角表示, 从而便于套用公式.

- **思想**二(识别公式结构)
观察代数式的结构, 看它接近哪个公式.
  - **公式识别**, 由 $(m+1)(n+1)=2$ 得 $\frac{m+n}{1-mn}=1$, 可联想到 $\tan(A+B)$ 的结构; $1\pm\sin 2\theta$ 可联想到 $(\sin\theta\pm\cos\theta)^2$.
  - **齐次式构造**, 分式可考虑弦化切; $\sin\alpha+2\cos\alpha$ 等非齐次式也可除以 $1=\sin^2\alpha+\cos^2\alpha$ 化为齐次式.

这种思路把复杂式子变形成熟悉的公式模型.

- **思想**三(整体代换)
当局部变量, 如 α,β 难以直接求解时, 可把某些组合视为整体.
  - **整体代换**, 例如已知 $\cos(\beta-\alpha)$ 与 $\tan\alpha\tan\beta$, 可把 $\cos\alpha\cos\beta$ 与 $\sin\alpha\sin\beta$ 设为整体未知数 X,Y 处理.

关注关系本身, 能简化复杂问题.

- **思想四**(归一)
式子里出现多种三角函数时, 常把它们化为同一类函数, 例如正切.
  - **齐次方程化正切方程**, 二次齐次方程 $\sin^2\theta+A\sin\theta\cos\theta+B\cos^2\theta=0$ 可化为关于 $\tan\theta$ 的代数方程.
  - **万能公式化有理分式**, 万能公式把各三角函数写成 $t=\tan(\frac{\alpha}{2})$ 的有理分式.

变量归一后, 代数运算和化简更容易进行.

注 6.13 (三角恒等变换的做法)

1. **分析条件**
拿到题目后, 先分析已知条件和求解目标, 重点关注两个方面.
   - **分析角**, 已知角与目标角之间是否存在和、差、倍、半、互补、互余等关系?
   - **分析结构**, 表达式是齐次式吗? 是否可以因式分解? 是否与某个公式的结构相似? 函数名是否统一? 幂次是否一致?
2. **变形与转化**
根据对条件与结构的分析, 选择最合适的思想方法变形.
   - 若**角不统一**, 优先使用**角的变换**思想, 进行拆角或凑角.
   - 若**名不统一**, 优先使用**统一思想**, 进行弦化切或应用同角关系.
   - 若**幂不统一**, 优先使用**降幂公式**或**升幂公式**(二倍角余弦).
   - 若**结构特殊**, 应用**结构变形**思想, 直接套用模型或进行因式分解.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 若问题**极其复杂**, 考虑**整体思想**或启用**万能公式**作为最后手段.

3. **求解与检验**
   经过变形后, 问题通常会化简为一个可以直接求解的代数问题或简单的三角函数求值问题.
   - **求解**. 进行最后的代数运算, 得出结果.
   - **检验**. 将结果带回思考, 是否符合题目给出的角的范围, 如象限、锐角等? 在变形过程中, 是否有除以一个可能为零的式子? 逻辑上是否存在漏洞?

6.11.5 高考真题选讲

例题 6.39 (2021 年全国 □ 卷数学第 6 题) 若 $\tan\theta=-2$, 则 $\dfrac{\sin\theta(1+\sin 2\theta)}{\sin\theta+\cos\theta}=$

A. $-\frac{6}{5}$　　C. $\frac{2}{5}$

B. $-\frac{2}{5}$　　D. $\frac{6}{5}$

解 答案:C. 将 $1+\sin 2\theta=\sin^2\theta+\cos^2\theta+2\sin\theta\cos\theta=(\sin\theta+\cos\theta)^2$ 代入,

$$\frac{\sin\theta(1+\sin 2\theta)}{\sin\theta+\cos\theta}=\frac{\sin\theta(\sin\theta+\cos\theta)^2}{\sin\theta+\cos\theta}=\sin\theta(\sin\theta+\cos\theta)=\sin^2\theta+\sin\theta\cos\theta.$$

再除以 $\sin^2\theta+\cos^2\theta=1$ 弦化切,

$$\sin^2\theta+\sin\theta\cos\theta=\frac{\tan^2\theta+\tan\theta}{1+\tan^2\theta}=\frac{4-2}{1+4}=\frac{2}{5}.$$

6.12 角的变换

来龙去脉

我们在解三角变换题时,常把已知角与目标角写成不同形式. 例如已知 $\sin(\theta-\frac{\pi}{3})$, 却要求 $\cos(2\theta+\frac{\pi}{3})$, 它们乍看之下没有明显联系.

我们的目标是把目标角用已知角表示出来. 通过加减、倍半等代数运算, 把它向已知角的方向改写, 就能直接套用和、差、倍、半角公式.

换元法把这种关系写得更清楚. 令 $x=\theta-\frac{\pi}{3}$, 则复杂的角表达式改写为关于 x 的简单式子, 倍角、半角的关系也变得更明显.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

6.12.1 典例剖析

例题 6.40 (2024 年 1 月浙江省嘉兴市高三上期末) 已知 $\sin(\theta-\frac{\pi}{3})=\frac{\sqrt{5}}{5}$, 则 $\cos(2\theta+\frac{\pi}{3})=$ _____.

解 我们先识别已知角与目标角. 已知角是 $\theta-\frac{\pi}{3}$, 目标角是 $2\theta+\frac{\pi}{3}$. 令 $x=\theta-\frac{\pi}{3}$, 就能把两者关系显化.

换元简化令 $x=\theta-\frac{\pi}{3}$, 则已知条件变为 $\sin x=\frac{\sqrt{5}}{5}$. 再把目标角用 x 表示. 由 $x=\theta-\frac{\pi}{3}$ 得 $\theta=x+\frac{\pi}{3}$.

$$\begin{aligned}2\theta+\frac{\pi}{3}&=2\left(x+\frac{\pi}{3}\right)+\frac{\pi}{3}\\&=2x+\frac{2\pi}{3}+\frac{\pi}{3}\\&=2x+\pi\end{aligned}$$

转化问题换元后, 原问题等价于 已知 $\sin x=\frac{\sqrt{5}}{5}$, 求 $\cos(2x+\pi)$ 的值.

现在求解新问题. 先用诱导公式, 再用二倍角公式.

$$\cos(2x+\pi)=-\cos(2x)$$

已知 $\sin x$, 我们选用只含 $\sin x$ 的余弦二倍角公式.

$$-\cos(2x)=-(1-2\sin^2 x)=2\sin^2 x-1$$

代入已知值 $\sin x=\frac{\sqrt{5}}{5}$:

$$2\left(\frac{\sqrt{5}}{5}\right)^2-1=2\left(\frac{5}{25}\right)-1=\frac{10}{25}-1=\frac{2}{5}-1=-\frac{3}{5}$$

故最终答案为 $-\frac{3}{5}$.

例题 6.41 (2023 年 11 月 THUSSAT 诊断性测试) 若 $\frac{9\pi}{16}<\alpha<\frac{17\pi}{16}$, 且满足 $\cos\left(2\alpha-\frac{\pi}{8}\right)=-\frac{1}{4}$, 则 $\sin\left(\frac{9\pi}{16}-\alpha\right)=$(　　)

A. $-\frac{\sqrt{6}}{4}$

B. $\frac{\sqrt{6}}{4}$

C. $-\frac{\sqrt{10}}{4}$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

D. $\dfrac{\sqrt{10}}{4}$

解 这道题的角关系更隐蔽. 我们仍用换元法揭示它, 再根据角的范围判断三角函数值的符号.

换元简化令 $x=2\alpha-\frac{\pi}{8}$. 已知条件变为 $\cos x=-\frac{1}{4}$. 再把目标角 $\frac{9\pi}{16}-\alpha$ 用 x 表示. 由 $x=2\alpha-\frac{\pi}{8}$, 可得 $2\alpha=x+\frac{\pi}{8}\implies\alpha=\frac{x}{2}+\frac{\pi}{16}$. 把它代入目标角, 得

$$\frac{9\pi}{16}-\alpha=\frac{9\pi}{16}-\left(\frac{x}{2}+\frac{\pi}{16}\right)=\frac{8\pi}{16}-\frac{x}{2}=\frac{\pi}{2}-\frac{x}{2}$$

转化问题于是原问题等价于下面这个问题. 已知 $\cos x=-\frac{1}{4}$, 求 $\sin(\frac{\pi}{2}-\frac{x}{2})$ 的值.

下面用诱导公式与半角关系求解. 根据诱导公式, $\sin(\frac{\pi}{2}-\frac{x}{2})=\cos(\frac{x}{2})$. 问题进一步化为已知 $\cos x$ 求 $\cos\frac{x}{2}$, 即半角问题. 由二倍角公式 $\cos x=2\cos^2\frac{x}{2}-1$, 解得 $\cos^2\frac{x}{2}$.

$$2\cos^2\frac{x}{2}=1+\cos x=1+\left(-\frac{1}{4}\right)=\frac{3}{4}$$

$$\cos^2\frac{x}{2}=\frac{3}{8}\implies\cos\frac{x}{2}=\pm\sqrt{\frac{3}{8}}=\pm\frac{\sqrt{6}}{4}$$

接下来由范围确定符号. 符号由角 $\frac{x}{2}$ 所在象限决定. 先根据 α 的范围求出 x 的范围, 再得到 $\frac{x}{2}$ 的范围. 已知 $\frac{9\pi}{16}<\alpha<\frac{17\pi}{16}$. 两边同乘以 2, 得 $\frac{9\pi}{8}<2\alpha<\frac{17\pi}{8}$. 再同减去 $\frac{\pi}{8}$, 得 $\frac{9\pi}{8}-\frac{\pi}{8}<2\alpha-\frac{\pi}{8}<\frac{17\pi}{8}-\frac{\pi}{8}$. 即 $\pi<x<2\pi$. 两边同除以 2, 得 $\frac{\pi}{2}<\frac{x}{2}<\pi$. 所以, 角 $\frac{x}{2}$ 在第二象限. 在第二象限, 余弦值为负, 因此 $\cos\frac{x}{2}$ 取负值.

$$\cos\frac{x}{2}=-\frac{\sqrt{6}}{4}$$

所以 $\sin(\frac{9\pi}{16}-\alpha)=-\frac{\sqrt{6}}{4}$. 正确答案为 A.

6.12.2 四大思想

这四条思路在上面的例题中反复出现.

- **思想一配凑与拆解**
  当已知角与目标角不一致时, 我们要构造它们之间的联系.
  - **拼凑角**. 已知 $\tan(\alpha-\beta)$ 与 $\tan\beta$ 求 $\tan\alpha$, 可用 $\alpha=(\alpha-\beta)+\beta$.
  - **拆分角**. 推导三倍角公式时, 把 3α 写成 $2\alpha+\alpha$.
  - **换元凑角**. 令 $x=\theta-\frac{\pi}{3}$, 则目标角 $2\theta+\frac{\pi}{3}=2x+\pi$.

  角的变换最终都是把未知角用已知角表示, 从而直接应用公式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

- **思想二识别公式结构**
  观察代数式的结构, 看它接近哪个公式.
  - **公式识别**. 由 $(m+1)(n+1)=2$ 得 $\frac{m+n}{1-mn}=1$, 可联想到 $\tan(A+B)$ 的结构; $1\pm\sin 2\theta$ 可联想到 $(\sin\theta\pm\cos\theta)^2$.
  - **齐次式构造**. 分式可考虑弦化切; $\sin\alpha+2\cos\alpha$ 等非齐次式也可除以 $1=\sin^2\alpha+\cos^2\alpha$ 化为齐次式.

  它把复杂式子变形成常见公式模型.
- **思想三整体代换**
  当局部变量(如 α,β)难以直接求解时, 可把某些组合视为整体.
  - **整体代换**. 例如已知 $\cos(\beta-\alpha)$ 与 $\tan\alpha\tan\beta$, 可把 $\cos\alpha\cos\beta$ 与 $\sin\alpha\sin\beta$ 设为整体未知数 X,Y 处理.

  这样注意力放在关系本身, 从而简化复杂问题.
- **思想四归一**
  式子里出现多种三角函数时, 常把它们化为同一类函数, 例如正切.
  - **齐次方程化正切方程**. 二次齐次方程 $\sin^2\theta+A\sin\theta\cos\theta+B\cos^2\theta=0$ 可化为关于 $\tan\theta$ 的代数方程.
  - **万能公式化有理分式**. 万能公式把各三角函数写成 $t=\tan(\frac{\alpha}{2})$ 的有理分式.

  变量归一后, 代数运算和化简变得更直接.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

6.12.3 高考真题选讲

知乎@余命数

例题 6.42 (2012 年江苏卷第 11 题) 设 α 为锐角, 若 $\cos\left(\alpha+\frac{\pi}{6}\right)=\frac{4}{5}$, 则 $\sin\left(2\alpha+\frac{\pi}{12}\right)=$ _____.

解 令 $\beta=\alpha+\frac{\pi}{6}$, 则 $2\alpha+\frac{\pi}{12}=2\left(\alpha+\frac{\pi}{6}\right)-\frac{\pi}{4}=2\beta-\frac{\pi}{4}$, 于是

$$\sin\left(2\alpha+\frac{\pi}{12}\right)=\sin\left(2\beta-\frac{\pi}{4}\right)=\frac{\sqrt{2}}{2}(\sin 2\beta-\cos 2\beta).$$

知乎@余命数

由 α 为锐角知 $\beta=\alpha+\frac{\pi}{6}\in\left(\frac{\pi}{6},\frac{2\pi}{3}\right)$, 又 $\cos\beta=\frac{4}{5}$, 故 $\sin\beta=\frac{3}{5}$. 由此

$$\sin 2\beta=2\cdot\frac{3}{5}\cdot\frac{4}{5}=\frac{24}{25},\quad \cos 2\beta=2\cdot\frac{16}{25}-1=\frac{7}{25}.$$

代回原式,

$$\sin\left(2\alpha+\frac{\pi}{12}\right)=\frac{\sqrt{2}}{2}\left(\frac{24}{25}-\frac{7}{25}\right)=\frac{\sqrt{2}}{2}\cdot\frac{17}{25}=\frac{17\sqrt{2}}{50}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

6.13 和差化积与积化和差

至此,我们已会处理角的和差,例如 $\sin(\alpha+\beta)$. 更复杂的问题中还会出现三角函数的和差,例如 $\sin\alpha+\sin\beta$. 这两者形式不同,但可以通过公式互相转化.

积化和差与和差化积是和差角公式的变形与组合. 它们互为逆运算.

- **积化和差**. 将两个三角函数的乘积化为和差. 它常用于降幂和化简,把乘积项变成可以逐项处理的和差项.
- **和差化积**. 将两个三角函数的和差化为乘积. 它常用于分式约分和解三角方程,把 $f(x)+g(x)=0$ 化为 $A(x)B(x)=0$ 的形式.

这两套公式是三角恒等变换中处理复杂变形的常用工具.

6.13.1 公式

6.13.1.1 积化和差

我们回顾正弦与余弦的和差角公式.

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta \quad \cdots ①$$

$$\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta \quad \cdots ②$$

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta \quad \cdots ③$$

$$\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta \quad \cdots ④$$

通过将以上四式两两相加或相减,即可得到积化和差公式. 例如,$(①+②)\div 2$ 可得

$$\sin\alpha\cos\beta=\frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)]$$

定理 6.9 (积化和差公式)

$$\sin\alpha\cos\beta=\frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)]$$

$$\cos\alpha\sin\beta=\frac{1}{2}[\sin(\alpha+\beta)-\sin(\alpha-\beta)]$$

$$\cos\alpha\cos\beta=\frac{1}{2}[\cos(\alpha+\beta)+\cos(\alpha-\beta)]$$

$$\sin\alpha\sin\beta=-\frac{1}{2}[\cos(\alpha+\beta)-\cos(\alpha-\beta)]$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

证明 由和差角公式

$$\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta,$$
$$\sin(\alpha-\beta)=\sin\alpha\cos\beta-\cos\alpha\sin\beta,$$
$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta,$$
$$\cos(\alpha-\beta)=\cos\alpha\cos\beta+\sin\alpha\sin\beta.$$

将前两式相加、相减, 分别得到前两个积化和差公式; 将后两式相加、相减, 分别得到后两个公式. □

6.13.1.2 和差化积公式

在积化和差公式的基础上, 我们进行一次巧妙的换元. 令 $\theta=\alpha+\beta, \phi=\alpha-\beta$, 则 $\alpha=\frac{\theta+\phi}{2}$, $\beta=\frac{\theta-\phi}{2}$. 将此代入第一个积化和差公式 $\sin\alpha\cos\beta=\frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)]$ 中, 反解可得

$$\sin\theta+\sin\phi=2\sin\frac{\theta+\phi}{2}\cos\frac{\theta-\phi}{2}$$

知乎@余命数

定理 6.10 (和差化积公式)

$$\sin\theta+\sin\phi=2\sin\frac{\theta+\phi}{2}\cos\frac{\theta-\phi}{2}$$
$$\sin\theta-\sin\phi=2\cos\frac{\theta+\phi}{2}\sin\frac{\theta-\phi}{2}$$
$$\cos\theta+\cos\phi=2\cos\frac{\theta+\phi}{2}\cos\frac{\theta-\phi}{2}$$
$$\cos\theta-\cos\phi=-2\sin\frac{\theta+\phi}{2}\sin\frac{\theta-\phi}{2}$$

证明 在积化和差公式中令

$$\alpha=\frac{\theta+\phi}{2},\qquad \beta=\frac{\theta-\phi}{2}.$$

则 $\alpha+\beta=\theta, \alpha-\beta=\phi$. 将这一换元分别代入四个积化和差公式并整理, 即得四个和差化积公式. □

这四个公式的结构可统一如下. 两角正弦之和, 右端正弦在前; 两角正弦之差, 右端余弦在前; 两角余弦之和, 右端余弦在前; 两角余弦之差, 右端正弦在前且带负号. 各式右端的角度永远是和的一半、差的一半.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

6.13.2 典例剖析

例题 6.43 (和差化积的应用) 化简分式 $\dfrac{\sin 5x+\sin 3x}{\cos 5x+\cos 3x}$.

解 本题的分子、分母都是三角函数的和. 我们先用和差化积公式化简, 再约分.

对分子应用和差化积

$$\sin 5x+\sin 3x=2\sin\frac{5x+3x}{2}\cos\frac{5x-3x}{2}=2\sin 4x\cos x$$

对分母应用和差化积

$$\cos 5x+\cos 3x=2\cos\frac{5x+3x}{2}\cos\frac{5x-3x}{2}=2\cos 4x\cos x$$

知乎@余命数

整合与约分将化简后的结果代回原分式.

$$\text{原式}=\frac{2\sin 4x\cos x}{2\cos 4x\cos x}$$

知乎@余命数

假设 $\cos 4x\neq 0$ 且 $\cos x\neq 0$, 可约去公因式 $2\cos x$.

$$\text{原式}=\frac{\sin 4x}{\cos 4x}=\tan 4x$$

知乎@余命数

例题 6.44 (积化和差的应用) 求函数 $f(x)=\sin(x+\frac{\pi}{6})\cos(x-\frac{\pi}{3})$ 的最大值.

解 本题是两个三角函数的乘积, 直接求最值较繁琐. 我们把它用积化和差化为标准正弦函数.

知乎@余命数

应用积化和差公式令 $\alpha=x+\frac{\pi}{6}$, $\beta=x-\frac{\pi}{3}$. 应用公式 $\sin\alpha\cos\beta=\frac{1}{2}[\sin(\alpha+\beta)+\sin(\alpha-\beta)]$. 先计算 $\alpha+\beta$ 和 $\alpha-\beta$.

知乎@余命数

$$\alpha+\beta=\left(x+\frac{\pi}{6}\right)+\left(x-\frac{\pi}{3}\right)=2x-\frac{\pi}{6}$$

知乎@余命数

$$\alpha-\beta=\left(x+\frac{\pi}{6}\right)-\left(x-\frac{\pi}{3}\right)=\frac{\pi}{6}+\frac{\pi}{3}=\frac{\pi}{2}$$

代入公式后.

$$\begin{aligned}f(x)&=\frac{1}{2}\left[\sin\left(2x-\frac{\pi}{6}\right)+\sin\left(\frac{\pi}{2}\right)\right]\\&=\frac{1}{2}\left[\sin\left(2x-\frac{\pi}{6}\right)+1\right]\\&=\frac{1}{2}\sin\left(2x-\frac{\pi}{6}\right)+\frac{1}{2}\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

分析最值经过积化和差, 我们把 $f(x)$ 化为 $y=A\sin(\omega x+\varphi)+k$ 的形式. 已知 $\sin(2x-\frac{\pi}{6})$ 的取值范围是 $[-1,1]$. 因此, $f(x)$ 的最大值在 $\sin(2x-\frac{\pi}{6})=1$ 时取得.

$$f(x)_{\max}=\frac{1}{2}(1)+\frac{1}{2}=1$$

函数的最大值为 1.

例题 6.45 (2022 年 6 月南京大学强基测试复试) 已知 $\alpha,\beta\in(0,\pi)$, 且 $\cos\alpha+\cos\beta-\cos(\alpha+\beta)=\frac{3}{2}$, 试求 α,β 的值.

解 这个方程混合了单角与和角, 直接求解 α,β 困难. 我们先把各项化为相同的角, 再统一处理.

统一视角并应用公式观察到方程中有 $\cos\alpha+\cos\beta$ 和 $\cos(\alpha+\beta)$. 我们先把它们都与 $\frac{\alpha+\beta}{2}$ 和 $\frac{\alpha-\beta}{2}$ 联系起来. 对 $\cos\alpha+\cos\beta$ 应用**和差化积**.

$$\cos\alpha+\cos\beta=2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}$$

对 $\cos(\alpha+\beta)$ 应用二**倍角公式**, 将其视为 $\cos(2\cdot\frac{\alpha+\beta}{2})$.

$$\cos(\alpha+\beta)=2\cos^2\frac{\alpha+\beta}{2}-1$$

换元与代数化将上述两式代入原方程, 得

$$2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}-\left(2\cos^2\frac{\alpha+\beta}{2}-1\right)=\frac{3}{2}$$

我们再做换元以简化结构. 令 $a=\cos\frac{\alpha+\beta}{2}$, $b=\cos\frac{\alpha-\beta}{2}$. 则方程变为

$$2ab-(2a^2-1)=\frac{3}{2}$$

整理得到一个关于 a,b 的代数方程

$$2ab-2a^2+1=\frac{3}{2}\implies 4ab-4a^2+2=3\implies 4a^2-4ab+1=0$$

配方与求解接下来对关于 a,b 的式子配方.

$$(4a^2-4ab+b^2)-b^2+1=0$$

$$(2a-b)^2+(1-b^2)=0$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由范围定值. 因为 $\alpha,\beta\in(0,\pi)$, 所以 $\alpha-\beta\in(-\pi,\pi)$, 则 $\frac{\alpha-\beta}{2}\in(-\frac{\pi}{2},\frac{\pi}{2})$. 在此区间内, $b=\cos\frac{\alpha-\beta}{2}>0$. 同时 $b\leq 1$. 因此 $b^2\leq 1\implies 1-b^2\geq 0$. 接着, 方程 $(2a-b)^2+(1-b^2)=0$ 是两个非负数之和等于零, 那么唯一的可能是两个非负数同时为零.

$$\begin{cases}(2a-b)^2=0\implies 2a=b\\ 1-b^2=0\end{cases}$$

由 $1-b^2=0$ 及 $b>0$ 可得 $b=1$. 代入 $2a=b$ 可得 $a=\frac{1}{2}$.

反解定角将 a,b 的值代回去.

$$b=\cos\frac{\alpha-\beta}{2}=1\implies\frac{\alpha-\beta}{2}=0\implies\alpha=\beta$$

$$a=\cos\frac{\alpha+\beta}{2}=\frac{1}{2}$$

因为 $\alpha=\beta$, 所以 $\cos\frac{2\alpha}{2}=\cos\alpha=\frac{1}{2}$. 又因为 $\alpha\in(0,\pi)$, 所以 $\alpha=\frac{\pi}{3}$.

因此, $\alpha=\beta=\dfrac{\pi}{3}$.

例题 6.46 (2010 年清华大学自主招生测试) 求 $\sin^4 10^\circ+\sin^4 50^\circ+\sin^4 70^\circ$ 的值.

解 本题角度特殊、幂次较高. 直接计算较繁琐, 我们先用降幂公式降次, 再结合和差化简化简求和.

降幂处理处理高次幂通常先用降幂公式降次.

$$\sin^2 x=\frac{1-\cos 2x}{2}\implies\sin^4 x=\left(\frac{1-\cos 2x}{2}\right)^2=\frac{1-2\cos 2x+\cos^2 2x}{4}$$

将 $10^\circ,50^\circ,70^\circ$ 分别代入, 得

$$\begin{aligned}\text{原式}&=\frac{1-2\cos 20^\circ+\cos^2 20^\circ}{4}+\frac{1-2\cos 100^\circ+\cos^2 100^\circ}{4}\\&\quad+\frac{1-2\cos 140^\circ+\cos^2 140^\circ}{4}\\&=\frac{1}{4}\Big[3-2(\cos 20^\circ+\cos 100^\circ+\cos 140^\circ)\\&\quad+(\cos^2 20^\circ+\cos^2 100^\circ+\cos^2 140^\circ)\Big]\end{aligned}$$

问题转化为计算两个括号内的和.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

计算一次项之和令 $S_1 = \cos 20^\circ + \cos 100^\circ + \cos 140^\circ$. 先对后两项使用和差化积.

$$\begin{aligned}\cos 100^\circ + \cos 140^\circ &= 2\cos\frac{100^\circ + 140^\circ}{2}\cos\frac{100^\circ - 140^\circ}{2}\\ &= 2\cos 120^\circ \cos(-20^\circ) = 2\left(-\frac{1}{2}\right)\cos 20^\circ = -\cos 20^\circ\end{aligned}$$

知乎@余命数

所以, $S_1 = \cos 20^\circ + (-\cos 20^\circ) = 0$.

计算二次项之和令 $S_2 = \cos^2 20^\circ + \cos^2 100^\circ + \cos^2 140^\circ$. 再次进行降幂处理.

$$\begin{aligned}S_2 &= \frac{1+\cos 40^\circ}{2} + \frac{1+\cos 200^\circ}{2} + \frac{1+\cos 280^\circ}{2}\\ &= \frac{3}{2} + \frac{1}{2}(\cos 40^\circ + \cos 200^\circ + \cos 280^\circ)\end{aligned}$$

计算括号内的和 $S_3 = \cos 40^\circ + \cos 200^\circ + \cos 280^\circ$. 根据诱导公式, $\cos 200^\circ = \cos(180^\circ + 20^\circ) = -\cos 20^\circ$, $\cos 280^\circ = \cos(360^\circ - 80^\circ) = \cos 80^\circ$.

$$S_3 = \cos 40^\circ - \cos 20^\circ + \cos 80^\circ$$

对 $\cos 40^\circ + \cos 80^\circ$ 应用和差化积.

$$\begin{aligned}\cos 40^\circ + \cos 80^\circ &= 2\cos\frac{40^\circ + 80^\circ}{2}\cos\frac{40^\circ - 80^\circ}{2}\\ &= 2\cos 60^\circ \cos(-20^\circ) = 2\left(\frac{1}{2}\right)\cos 20^\circ = \cos 20^\circ\end{aligned}$$

所以, $S_3 = \cos 20^\circ - \cos 20^\circ = 0$. 因此, $S_2 = \frac{3}{2} + \frac{1}{2}(0) = \frac{3}{2}$.

整合求值将 $S_1 = 0$ 和 $S_2 = \frac{3}{2}$ 代回原式.

$$\text{原式} = \frac{1}{4}\left[3 - 2(0) + \frac{3}{2}\right] = \frac{1}{4}\left(3 + \frac{3}{2}\right) = \frac{1}{4}\left(\frac{9}{2}\right) = \frac{9}{8}$$

最终结果为 $\dfrac{9}{8}$.

知乎@余命数

来龙去脉

我们已研究了最基本的正弦函数 $y = \sin x$; 它的振幅为 1, 周期为 2π, 从原点开始振动. 现实中的周期现象更复杂.

- 声音有高低**音量**之分;
- 琴弦有不同**音调**之别;
- 振动有不同**起始时刻**之差;
- 潮汐有不同**基准海平面**之高低.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

描述这些现象需要为基础波形 $y=\sin x$ 引入参数，写成 $y=A\sin(\omega x+\varphi)+B$ $(A>0,\omega>0)$. 参数 A,ω,φ,B 分别对应振幅、周期、相位和垂直平移四种几何变换. 我们在本节中研究每个参数对函数图像的单独作用，以及它们共同作用时的效果.

$y=A\sin(\omega x+\varphi)+B$ 的图像可看成由 $y=\sin x$ 经过一系列几何变换得到. 参数 A,B,ω,φ 分别控制一种特定形变. 我们逐一考察每个参数的效果.

6.13.2.1 参数 B 与垂直平移

参数 B 控制图像的**整体上下平移**. 它不改变图像的形状(振幅、周期)，只改变其垂直位置.

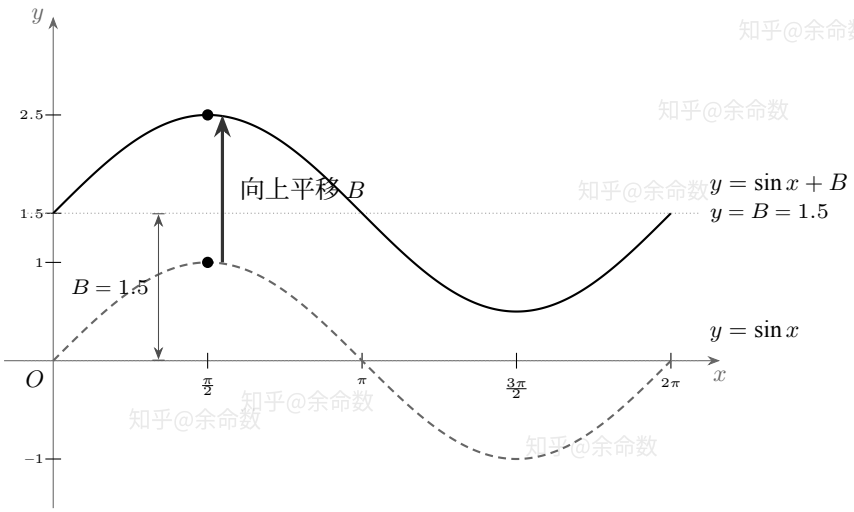


图 6.17　参数 B 导致图像垂直平移，得到 $y=\sin x+B$

- **变换规则**. 将 $y=f(x)$ 的图像向上 $(B>0)$ 或向下 $(B<0)$ 平移 $|B|$ 个单位长度，得到 $y=f(x)+B$ 的图像.
- **性质影响**.
  - **对称中心线**. 图像的平衡位置由 x 轴 $(y=0)$ 平移至直线 $y=B$.
  - **最值**. 最大值由 A 变为 $A+B$，最小值由 $-A$ 变为 $-A+B$.

6.13.2.2 参数 A 与垂直伸缩(振幅变换)

参数 A $(A>0)$ 控制图像在垂直方向上的**拉伸或压缩**，即改变图像的**振幅**.

- **变换规则**. 将 $y=f(x)$ 图像上所有点的纵坐标变为原来的 A 倍(横坐标不变)，得到 $y=Af(x)$ 的图像.
- **性质影响**.
  - **值域**. 值域由 $[-1,1]$ 变为 $[-A,A]$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

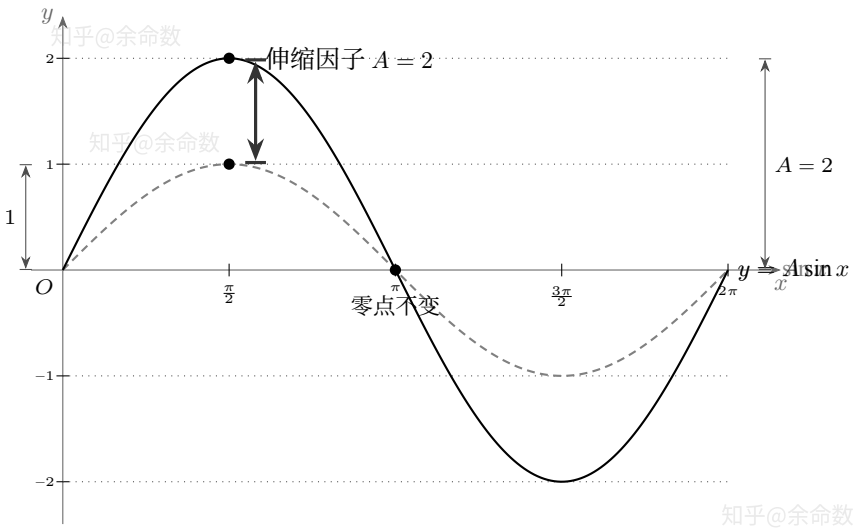


图 6.18 参数 A 使图像在垂直方向伸缩 A 倍(振幅变为 A)

- **最值**. 最大值为 A, 最小值为 $-A$.
- **不变性**. 图像的周期、对称轴位置和零点位置均**不改变**.

6.13.2.3 参数 ω 与水平伸缩(周期变换)

参数 ω ($\omega>0$) 控制图像在水平方向上的**压缩或拉伸**, 直接改变图像的**周期**.

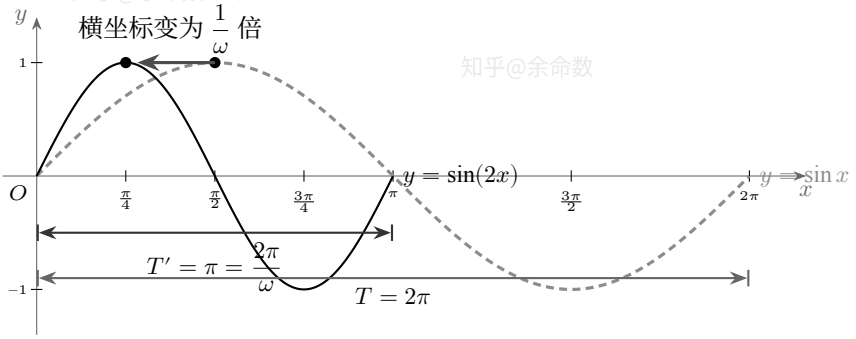


图 6.19 参数 ω 使图像在水平方向压缩 ω 倍(周期变为 $\frac{2\pi}{\omega}$)

- **变换规则**. 将 $y=f(x)$ 图像上所有点的横坐标变为原来的 $\frac{1}{\omega}$ 倍(纵坐标不变), 得到 $y=f(\omega x)$ 的图像.
  - 当 $\omega>1$ 时, 图像被**压缩**.
  - 当 $0<\omega<1$ 时, 图像被**拉伸**.
- **性质影响**.
  - **周期**. 最小正周期由 T 变为 $T'=\dfrac{T}{\omega}$. 对于正弦函数, 由 2π 变为 $\dfrac{2\pi}{\omega}$.
  - **不变性**. 图像的振幅、值域和平衡位置均**不改变**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

6.13.2.4 参数 φ 与水平平移(相位变换)

参数 φ 控制图像的**整体左右平移**. 它不改变图像的形状, 只改变其水平位置, 即**相位**.

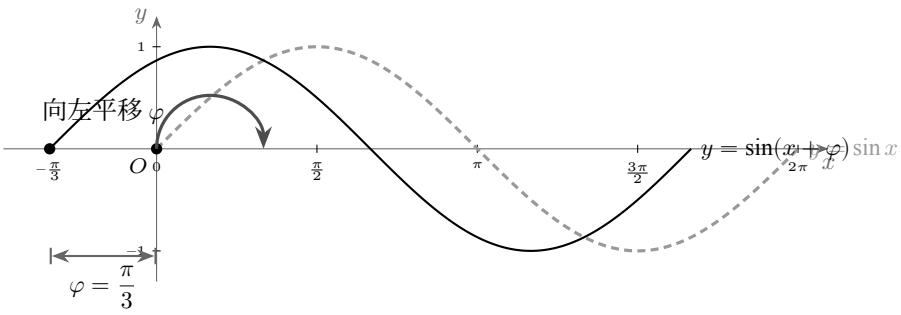


图 6.20　参数 φ 使图像沿 x 轴平移; $\varphi > 0$ 时向左平移, $\varphi < 0$ 时向右平移

- **变换规则**. 将 $y = f(x)$ 的图像进行平移.
  - **向左**平移 φ 个单位 ($\varphi > 0$), 得到 $y = f(x + \varphi)$ 的图像.
  - **向右**平移 φ 个单位 ($\varphi > 0$), 得到 $y = f(x - \varphi)$ 的图像.

  即向左平移对应函数中的加号, 向右平移对应函数中的减号.
- **性质影响**.
  - **位置**. 图像上所有点的位置发生水平移动, 因此对称轴、对称中心、最值点、零点的位置都发生改变.
  - **不变性**. 图像的形状完全不变, 即振幅、周期和值域都**不改变**.

6.14 $y = A\sin(\omega x + \varphi) + B$

6.14.1 由 $y = \sin x$ 到 $y = A\sin(\omega x + \varphi) + B$ 的图像变换

我们想把 $y = \sin x$ 的图像变成 $y = A\sin(\omega x + \varphi) + B$. 伸缩与平移的顺序不同, 得到的结果也会不同. 我们先伸缩、后平移, 这样最不容易混淆.

路径一

$$
\begin{aligned}
y = \sin x &\xrightarrow{\text{周期变换}} y = \sin(\omega x) \\
&\xrightarrow{\text{相位变换}} y = \sin(\omega x + \varphi) \\
&\xrightarrow{\text{振幅变换}} y = A\sin(\omega x + \varphi) \\
&\xrightarrow{\text{垂直平移}} y = A\sin(\omega x + \varphi) + B
\end{aligned}
$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

我们一步一步地看.

1. **周期变换**. 将 $y=\sin x$ 的图像上所有点的横坐标变为原来的 $\frac{1}{\omega}$ 倍(纵坐标不变),得到 $y=\sin(\omega x)$.
2. **相位变换**. 将 $y=\sin(\omega x)$ 的图像**向左**平移 $\frac{\varphi}{\omega}$ 个单位长度,得到 $y=\sin\left(\omega\left(x+\frac{\varphi}{\omega}\right)\right)=\sin(\omega x+\varphi)$.
3. **振幅变换**. 将 $y=\sin(\omega x+\varphi)$ 的图像上所有点的纵坐标变为原来的 A 倍(横坐标不变),得到 $y=A\sin(\omega x+\varphi)$.
4. **垂直平移**. 将 $y=A\sin(\omega x+\varphi)$ 的图像向上 $(B>0)$ 或向下 $(B<0)$ 平移 $|B|$ 个单位长度,得到最终图像.

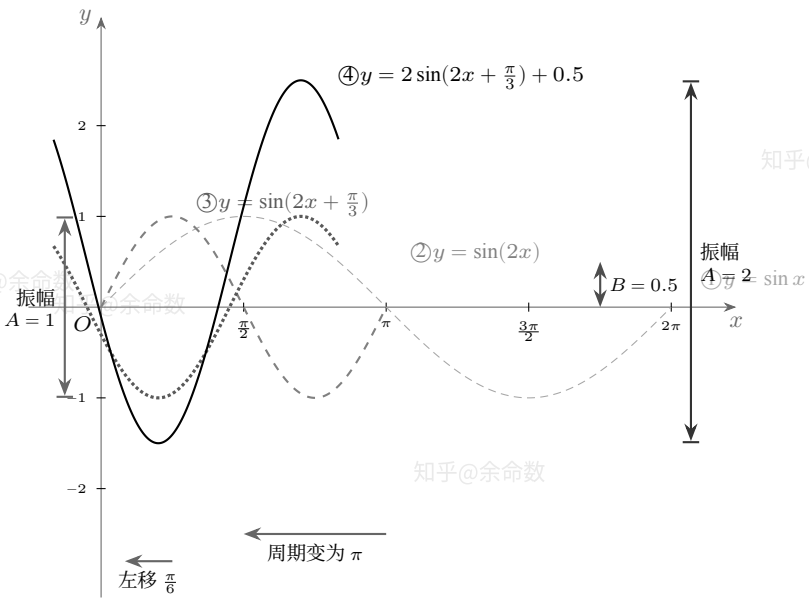


图 6.21 由 $y=\sin x$ 到 $y=2\sin(2x+\frac{\pi}{3})+0.5$ 的图像变换过程

注 6.14 (平移量的确定)
我们进行水平平移时,先把 x 的系数 ω 提出,得到 $y=\sin\left(\omega\left(x+\frac{\varphi}{\omega}\right)\right)$.

- 从 $y=\sin(\omega x)$ 到 $y=\sin(\omega x+\varphi)$,图像向左平移了 $\frac{\varphi}{\omega}$ 个单位.
- 另一条路径先把 $y=\sin x$ 向左平移 φ 个单位,得到 $y=\sin(x+\varphi)$,再把横坐标缩为原来的 $\frac{1}{\omega}$ 倍,于是平移量也缩为 $\frac{\varphi}{\omega}$,最终结果仍是 $y=\sin(\omega x+\varphi)$.

因此,我们采用先伸缩、后平移的路径;平移量是 $\frac{\varphi}{\omega}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

6.14.2 整体换元思想

我们分析函数 $f(x) = A\sin(\omega x + \varphi) + B$ 的性质时，把内层函数 $u = \omega x + \varphi$ 看成一个整体. 先求出基础函数 $y = A\sin u + B$ 的对应性质区间，再反解出 x 的取值范围.

- **值域与最值**. 当 $\sin(\omega x + \varphi) = 1$ 时，$y_{\max} = |A| + B$. 当 $\sin(\omega x + \varphi) = -1$ 时，$y_{\min} = -|A| + B$.
- **单调区间**.（以 $A > 0, \omega > 0$ 为例）要求单调递增区间，即求满足 $2k\pi - \frac{\pi}{2} \leq \omega x + \varphi \leq 2k\pi + \frac{\pi}{2}$ 的 x 的范围. 要求单调递减区间，即求满足 $2k\pi + \frac{\pi}{2} \leq \omega x + \varphi \leq 2k\pi + \frac{3\pi}{2}$ 的 x 的范围.
- **对称性**.（以 $A > 0, \omega > 0$ 为例）对称轴由 $\omega x + \varphi = k\pi + \frac{\pi}{2}$ 解出. 对称中心由 $\omega x + \varphi = k\pi$ 解出其横坐标.

6.14.3 典例剖析

例题 6.47 (由图求参) 已知函数 $f(x) = A\sin(\omega x + \varphi) + B$ $(A > 0, \omega > 0, |\varphi| < \frac{\pi}{2})$ 的部分图像如图所示，求其解析式.

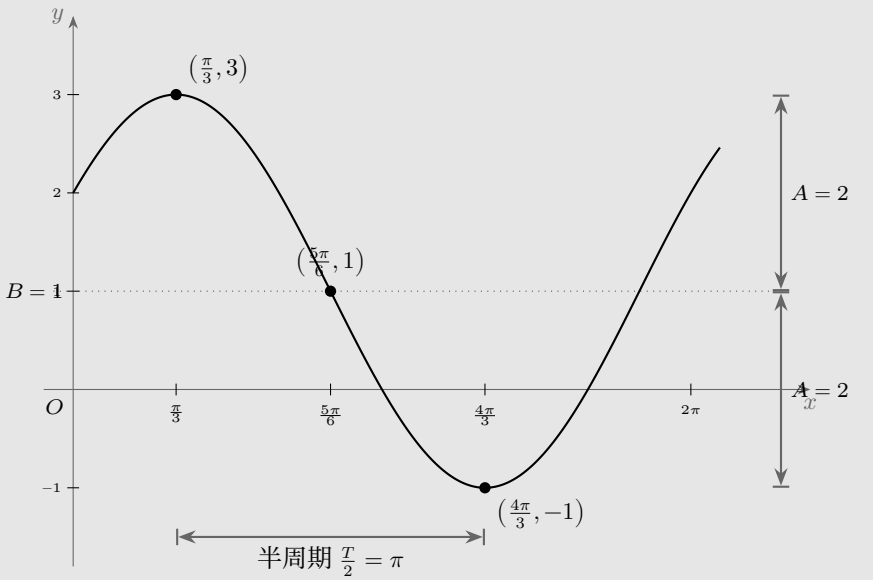


图 6.22　函数 $f(x) = A\sin(\omega x + \varphi) + B$ 的部分图像

解 我们按 **定** B、**定** A、**定** ω、**定** φ 的顺序，从图像中逐一求出参数.

先确定平衡位置 B

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

由图可知, 函数的最大值 $y_{\max}=3$, 最小值为 $y_{\min}=-1$.

$$B=\frac{y_{\max}+y_{\min}}{2}=\frac{3+(-1)}{2}=1$$

再确定振幅 A

$$A=\frac{y_{\max}-y_{\min}}{2}=\frac{3-(-1)}{2}=2$$

接着确定角频率 ω

图中给出了一个最高点 $\left(\frac{\pi}{3},3\right)$ 和一个相邻的最低点 $\left(\frac{4\pi}{3},-1\right)$. 它们之间的水平距离恰好是半个周期.

$$\frac{T}{2}=\frac{4\pi}{3}-\frac{\pi}{3}=\pi \implies T=2\pi$$

由周期公式 $T=\frac{2\pi}{\omega}$ 可得

$$\omega=\frac{2\pi}{T}=\frac{2\pi}{2\pi}=1$$

最后确定初相 φ

将 A,B,ω 的值代入, 函数解析式为 $f(x)=2\sin(x+\varphi)+1$.

我们取最高点 $\left(\frac{\pi}{3},3\right)$ 代入来求 φ.

$$3=2\sin\left(\frac{\pi}{3}+\varphi\right)+1 \implies 2=2\sin\left(\frac{\pi}{3}+\varphi\right) \implies \sin\left(\frac{\pi}{3}+\varphi\right)=1$$

于是内层函数满足 $\frac{\pi}{3}+\varphi=\frac{\pi}{2}+2k\pi, k\in\mathbb{Z}$.

$$\frac{\pi}{3}+\varphi=\frac{\pi}{2}+2k\pi \implies \varphi=\frac{\pi}{6}+2k\pi$$

由条件 $|\varphi|<\frac{\pi}{2}$, 取 $k=0$, 得 $\varphi=\frac{\pi}{6}$.

答案. 所求函数的解析式为 $f(x)=2\sin\left(x+\frac{\pi}{6}\right)+1$

注 6.15 (五点法反求 φ)

我们确定 φ 时, 把五点法中的特殊点代入, 这些点对应最高点、最低点或与平衡位置的交点. 这样得到的方程容易求解.

- **代入最高点**. 此时 $\sin(\omega x+\varphi)=1$, 即 $\omega x+\varphi=\frac{\pi}{2}+2k\pi$.
- **代入最低点**. 此时 $\sin(\omega x+\varphi)=-1$, 即 $\omega x+\varphi=-\frac{\pi}{2}+2k\pi$.
- **代入中心线上升点**. 此时 $\sin(\omega x+\varphi)=0$ 且函数递增, 即 $\omega x+\varphi=2k\pi$.

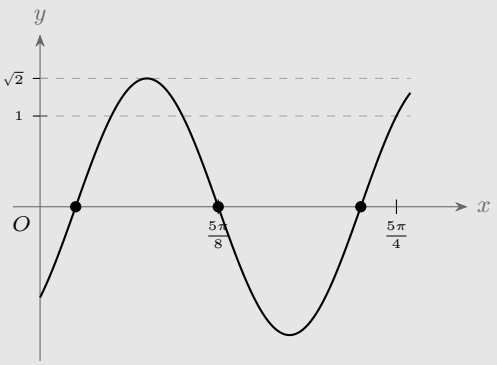
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **代入中心线下降点**. 此时 $\sin(\omega x+\varphi)=0$ 且函数递减, 即 $\omega x+\varphi=\pi+2k\pi$.

这样得到的方法称为“五点法求 φ”.

综合题中的导数判别

例题 6.48 (多选, 2023 山东师大附中 10 月月考) 函数 $f(x)=\sqrt{2}\sin(\omega x+\varphi)$ $(\omega>0,|\varphi|<\frac{\pi}{2})$ 的部分图象如图所示, 则下列说法中正确的有(　)



A. $f(x)$ 的最小正周期 T 为 π

B. $f(x)$ 向右平移 $\dfrac{3\pi}{8}$ 个单位后得到的新函数是偶函数

C. 若方程 $f(x)=1$ 在 $(0,m)$ 上共有 4 个根, 则这 4 个根的和为 $\dfrac{7\pi}{2}$

D. $f(x)(x\in\left[0,\dfrac{5\pi}{4}\right])$ 图象上的动点 M 到直线 $2x-y+4=0$ 的距离最小时, M 的横坐标为 $\dfrac{\pi}{4}$

知乎@余命数

解　我们先用图像确定函数解析式, 再逐一判断四个选项. (D 选项用到切线斜率, 见 第16章.)

假设选项 A 正确, 即 $T=\pi$. 则 $\omega=\frac{2\pi}{T}=2$. 函数为 $f(x)=\sqrt{2}\sin(2x+\varphi)$. 将图中的零点 $(\frac{5\pi}{8},0)$ 代入, 得

知乎@余命数

$$\sqrt{2}\sin\left(2\cdot\frac{5\pi}{8}+\varphi\right)=0\implies\sin\left(\frac{5\pi}{4}+\varphi\right)=0$$

即 $\frac{5\pi}{4}+\varphi=k\pi, k\in\mathbb{Z}$.

$$\varphi=k\pi-\frac{5\pi}{4}$$

由条件 $|\varphi|<\frac{\pi}{2}$, 取 $k=1$, 得 $\varphi=\pi-\frac{5\pi}{4}=-\frac{\pi}{4}$. 于是函数解析式为 $f(x)=\sqrt{2}\sin(2x-\frac{\pi}{4})$. 代入 $x=\frac{\pi}{8}$ 和 $x=\frac{5\pi}{8}$ 均得 $f(x)=0$, 与图象吻合.

逐一判断选项

- **对于 A**. 它基于 $T=\pi$ 推导出的解析式, 故 A 正确.
- **对于 B**. 将 $f(x)$ 向右平移 $\frac{3\pi}{8}$ 个单位得到 $g(x)=f(x-\frac{3\pi}{8})$.

$$\begin{aligned} g(x) &= \sqrt{2}\sin\left(2\left(x-\frac{3\pi}{8}\right)-\frac{\pi}{4}\right) \\ &= \sqrt{2}\sin\left(2x-\frac{3\pi}{4}-\frac{\pi}{4}\right)=\sqrt{2}\sin(2x-\pi)=-\sqrt{2}\sin(2x) \end{aligned}$$

该函数是奇函数, 故 B 错误.

- **对于 C**. 解方程 $f(x)=1 \implies \sqrt{2}\sin(2x-\frac{\pi}{4})=1 \implies \sin(2x-\frac{\pi}{4})=\frac{\sqrt{2}}{2}$.

$$2x-\frac{\pi}{4}=2k\pi+\frac{\pi}{4} \quad \text{或} \quad 2x-\frac{\pi}{4}=2k\pi+\frac{3\pi}{4}$$

$$2x=2k\pi+\frac{\pi}{2} \quad \text{或} \quad 2x=2k\pi+\pi$$

$$x=k\pi+\frac{\pi}{4} \quad \text{或} \quad x=k\pi+\frac{\pi}{2}$$

在 $(0,m)$ 内的 4 个正根, 从小到大依次为. $x_1=\frac{\pi}{4}$, $x_2=\frac{\pi}{2}$（当 $k=0$）$x_3=\pi+\frac{\pi}{4}=\frac{5\pi}{4}$, $x_4=\pi+\frac{\pi}{2}=\frac{3\pi}{2}$, 它们对应 $k=1$. 这 4 个根的和为 $\frac{\pi}{4}+\frac{2\pi}{4}+\frac{5\pi}{4}+\frac{6\pi}{4}=\frac{14\pi}{4}=\frac{7\pi}{2}$. 故 C 正确.

- **对于 D**. 求动点 $M(x,f(x))$ 到直线 $2x-y+4=0$ 的距离最小值, 即求

$$d(x)=\frac{|2x-f(x)+4|}{\sqrt{5}}$$

在 $x\in[0,\frac{5\pi}{4}]$ 上的最小值. 由于 $2x-f(x)+4\geq 4-\sqrt{2}>0$, 可直接比较 $h(x)=2x-f(x)+4$. 先找内部驻点.

$$h'(x)=2-f'(x)=2-2\sqrt{2}\cos\left(2x-\frac{\pi}{4}\right)$$

令 $h'(x)=0$, 得 $\cos(2x-\frac{\pi}{4})=\frac{\sqrt{2}}{2}$, 所以区间内的候选点为 $x=0,\frac{\pi}{4},\pi,\frac{5\pi}{4}$. 再比较端点与候选点.

$$h(0)=5, \quad h\left(\frac{\pi}{4}\right)=\frac{\pi}{2}+3, \quad h(\pi)=2\pi+5, \quad h\left(\frac{5\pi}{4}\right)=\frac{5\pi}{2}+3.$$

其中 $\frac{\pi}{2}+3<5<2\pi+5$ 且 $\frac{\pi}{2}+3<\frac{5\pi}{2}+3$, 故距离在 $x=\frac{\pi}{4}$ 处最小.D 正确.

正确的选项为 ACD.

切线条件用导数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 6.49 (多选，2023 湖南长郡中学月考) 已知 $x_0=\dfrac{2\pi}{3}$ 是函数 $f(x)=\sin(2x+\varphi)(0<\varphi<\pi)$ 的一个零点，则(　)

A. $f(x)$ 在区间 $\left[0,\dfrac{5\pi}{12}\right]$ 上单调递减

B. $f(x)$ 在区间 $\left(-\dfrac{\pi}{12},\dfrac{11\pi}{12}\right)$ 只有一个极值点

C. 直线 $x=\dfrac{7\pi}{6}$ 是曲线 $y=f(x)$ 的对称轴

D. 直线 $y=\dfrac{\sqrt{3}}{2}-x$ 是曲线 $y=f(x)$ 的切线

解 我们先用给定的零点求出参数，再判断各选项.

我们先确定函数解析式. 将零点 $x_0=\frac{2\pi}{3}$ 代入 $f(x)=\sin(2x+\varphi)$.

$$\sin\left(2\cdot\frac{2\pi}{3}+\varphi\right)=\sin\left(\frac{4\pi}{3}+\varphi\right)=0$$

即 $\frac{4\pi}{3}+\varphi=k\pi,k\in\mathbb{Z}$.

$$\varphi=k\pi-\frac{4\pi}{3}$$

由条件 $0<\varphi<\pi$，取 $k=2$，得 $\varphi=2\pi-\frac{4\pi}{3}=\frac{2\pi}{3}$. 因此 $f(x)=\sin(2x+\frac{2\pi}{3})$.

逐一判断选项

- **对于 A**. 求单调递减区间. 令 $u=2x+\frac{2\pi}{3}$，$\sin u$ 的递减区间为 $[\frac{\pi}{2}+2k\pi,\frac{3\pi}{2}+2k\pi]$.

$$\frac{\pi}{2}+2k\pi\le 2x+\frac{2\pi}{3}\le\frac{3\pi}{2}+2k\pi$$

$$-\frac{\pi}{6}+2k\pi\le 2x\le\frac{5\pi}{6}+2k\pi\implies -\frac{\pi}{12}+k\pi\le x\le\frac{5\pi}{12}+k\pi$$

当 $k=0$ 时，一个单调递减区间为 $[-\frac{\pi}{12},\frac{5\pi}{12}]$. 题目给出的区间 $[0,\frac{5\pi}{12}]$ 是上述区间的子集，因此 $f(x)$ 在此区间上单调递减. 故 A 正确.

- **对于 B**. 求极值点. 极值点满足 $u=2x+\frac{2\pi}{3}=\frac{\pi}{2}+k\pi$.

$$2x=\frac{\pi}{2}+k\pi-\frac{2\pi}{3}=-\frac{\pi}{6}+k\pi\implies x=\frac{k\pi}{2}-\frac{\pi}{12}$$

在区间 $(-\frac{\pi}{12},\frac{11\pi}{12})$ 内寻找满足条件的 x：若 $k=1,x=\frac{\pi}{2}-\frac{\pi}{12}=\frac{5\pi}{12}$，在区间内. 若 $k=2,x=\pi-\frac{\pi}{12}=\frac{11\pi}{12}$，不在区间内. 若 $k=0,x=-\frac{\pi}{12}$，不在区间内. 因此该区间内只有一个极值点. 故 B 正确.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

- **对于 C**. 判断对称轴. 对称轴方程为 $x=\frac{k\pi}{2}-\frac{\pi}{12}$. 将 $x=\frac{7\pi}{6}$ 代入该式看能否取得.

$$\frac{7\pi}{6}=\frac{k\pi}{2}-\frac{\pi}{12}\implies\frac{14\pi}{12}+\frac{\pi}{12}=\frac{k\pi}{2}$$
$$\implies\frac{15\pi}{12}=\frac{k\pi}{2}\implies\frac{5\pi}{4}=\frac{k\pi}{2}\implies k=\frac{5}{2}$$

  k 不是整数, $x=\frac{7\pi}{6}$ 不是对称轴. 故 C 错误.

- **对于 D**. 判断切线. 需找一个点 x_t, 同时满足 $f'(x_t)=-1$ 和 $f(x_t)=\frac{\sqrt{3}}{2}-x_t$. $f'(x)=2\cos(2x+\frac{2\pi}{3})$. 令 $f'(x_t)=-1\implies\cos(2x_t+\frac{2\pi}{3})=-\frac{1}{2}$. 一个简单的解是 $2x_t+\frac{2\pi}{3}=\frac{2\pi}{3}\implies x_t=0$. 验证 $x_t=0$. 切点坐标为 $x=0$, $y=f(0)=\sin(\frac{2\pi}{3})=\frac{\sqrt{3}}{2}$. 切点为 $(0,\frac{\sqrt{3}}{2})$. 将该点代入切线方程 $y=\frac{\sqrt{3}}{2}-x$, 左边为 $\frac{\sqrt{3}}{2}$, 右边为 $\frac{\sqrt{3}}{2}-0=\frac{\sqrt{3}}{2}$. 左右相等, 说明点在直线上. 斜率和点均满足, 该直线是曲线在 $x=0$ 处的切线. 故 D 正确.

正确的选项为 ABD.

6.14.4 高考真题选讲

以下选编近年高考中与正弦型函数 $A\sin(\omega x+\varphi)+b$ 相关的试题, 涵盖化简、性质与图像变换.

例题 6.50 (2016 年浙江卷理科数学第 10 题) 已知 $2\cos^2 x+\sin 2x=A\sin(\omega x+\varphi)+b\,(A>0)$, 则 $A=$_____, $b=$_____.

解 用降幂公式 $\cos^2 x=\frac{1+\cos 2x}{2}$ 把 $2\cos^2 x$ 降为一次,

$$2\cos^2 x+\sin 2x=(1+\cos 2x)+\sin 2x=\sin 2x+\cos 2x+1.$$

再用辅助角公式合并 $\sin 2x+\cos 2x$,

$$\sin 2x+\cos 2x=\sqrt{2}\sin\left(2x+\frac{\pi}{4}\right),$$

所以 $2\cos^2 x+\sin 2x=\sqrt{2}\sin(2x+\frac{\pi}{4})+1$, 故 $A=\sqrt{2}, b=1$.

例题 6.51 (2021 年全国乙卷文科数学第 4 题) 函数 $f(x)=\sin\frac{x}{3}+\cos\frac{x}{3}$ 的最小正周期和最大值分别是

A. 3π 和 $\sqrt{2}$　　C. 6π 和 $\sqrt{2}$

B. 3π 和 2　　D. 6π 和 2

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 答案:C. 用辅助角公式合并,

$$f(x)=\sin\frac{x}{3}+\cos\frac{x}{3}=\sqrt{2}\sin\left(\frac{x}{3}+\frac{\pi}{4}\right),$$

最小正周期 $T=\frac{2\pi}{1/3}=6\pi$, 最大值为 $\sqrt{2}$.

例题 6.52 (2021 年全国 □ 卷数学第 4 题) 下列区间中, 函数 $f(x)=7\sin\left(x-\frac{\pi}{6}\right)$ 单调递增的区间是

A. $\left(0,\frac{\pi}{2}\right)$　　C. $\left(\pi,\frac{3\pi}{2}\right)$

B. $\left(\frac{\pi}{2},\pi\right)$　　D. $\left(\frac{3\pi}{2},2\pi\right)$

解 答案:A. 令 $-\frac{\pi}{2}+2k\pi\leqslant x-\frac{\pi}{6}\leqslant\frac{\pi}{2}+2k\pi\,(k\in\mathbb{Z})$, 解得 $-\frac{\pi}{3}+2k\pi\leqslant x\leqslant\frac{2\pi}{3}+2k\pi$. 取 $k=0$, 单调递增区间为 $[-\frac{\pi}{3},\frac{2\pi}{3}]$, 而 $(0,\frac{\pi}{2})\subset[-\frac{\pi}{3},\frac{2\pi}{3}]$.

例题 6.53 (2017 年全国 □ 卷理科数学第 9 题) 已知曲线 $C_1:y=\cos x$, $C_2:y=\sin\left(2x+\frac{2\pi}{3}\right)$, 则

A. 把 C_1 上各点的横坐标伸长到原来的 2 倍, 纵坐标不变, 再把得到的曲线向右平移 $\frac{\pi}{6}$ 个单位长度, 得到曲线 C_2

B. 把 C_1 上各点的横坐标伸长到原来的 2 倍, 纵坐标不变, 再把得到的曲线向左平移 $\frac{\pi}{12}$ 个单位长度, 得到曲线 C_2

C. 把 C_1 上各点的横坐标缩短到原来的 $\frac{1}{2}$ 倍, 纵坐标不变, 再把得到的曲线向右平移 $\frac{\pi}{6}$ 个单位长度, 得到曲线 C_2

D. 把 C_1 上各点的横坐标缩短到原来的 $\frac{1}{2}$ 倍, 纵坐标不变, 再把得到的曲线向左平移 $\frac{\pi}{12}$ 个单位长度, 得到曲线 C_2

解 答案:D. 先统一函数名, 把 C_2 化为余弦形式,

$$C_2:y=\sin\left(2x+\frac{2\pi}{3}\right)=\cos\left(2x+\frac{2\pi}{3}-\frac{\pi}{2}\right)=\cos\left(2x+\frac{\pi}{6}\right).$$

从 $C_1:y=\cos x$ 出发, 横坐标缩短到原来的 $\frac{1}{2}$ 倍, 即 ω 从 1 变为 2, 得到 $y=\cos 2x$; 再向左平移 $\frac{\pi}{12}$ 个单位, 得到 $y=\cos\left[2\left(x+\frac{\pi}{12}\right)\right]=\cos\left(2x+\frac{\pi}{6}\right)$, 即 C_2.

例题 6.54 (2016 年全国 □ 卷理科数学第 14 题) 函数 $y=\sin x-\sqrt{3}\cos x$ 的图象可由函数 $y=\sin x+\sqrt{3}\cos x$ 的图象至少向右平移 _____ 个单位长度得到.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 分别化简两个函数,

$$y=\sin x+\sqrt{3}\cos x=2\sin\left(x+\frac{\pi}{3}\right),\quad y=\sin x-\sqrt{3}\cos x=2\sin\left(x-\frac{\pi}{3}\right).$$

设向右平移 φ $(\varphi>0)$ 个单位,则 $2\sin\left(x+\frac{\pi}{3}-\varphi\right)=2\sin\left(x-\frac{\pi}{3}\right)$,需要

$$\frac{\pi}{3}-\varphi=-\frac{\pi}{3}+2k\pi\quad(k\in\mathbb{Z}),$$

解得 $\varphi=\frac{2\pi}{3}-2k\pi$. 取 $k=0$ 得最小正值 $\varphi=\frac{2\pi}{3}$.

6.15 ω 的约束与求解

注 6.16
个别变式使用导数表示切线斜率,见第16章.

我们常需要根据函数在区间上的性质,反向推断解析式中的参数,尤其是决定周期的参数 ω. 这些性质包括单调性、最值点的数量、零点的个数等. 我们把这类问题统称为参数 ω 的约束与求解.

这类问题可归结为建立周期与区间长度之间的不等关系. 函数在给定区间内能容纳多少个周期,决定了该区间内最值点和零点的数量. 下面给出两条常用的分析路径.

- **整体换元法**. 将 $u=\omega x+\varphi$ 视为一个整体,分析当 x 在给定区间内变化时 u 所在的区间. 然后考察标准正弦函数 $y=\sin u$ 在这个有效区间内必须包含哪些特征点(最高点、最低点、零点),以及必须排除哪些特征点,从而建立关于区间端点的不等式来求解 ω.
- **特征点定位法**,亦称 ω 卡根法. 直接写出函数 $f(x)$ 的所有特征点横坐标的通式(该通式含有 ω). 然后根据题意,确定第 n 个特征点必须落在给定区间内,而第 $n+1$ 个特征点必须落在区间外或恰好在端点上. 通过这种卡位的方式,建立关于 ω 的不等式组.

这两条路径最终都归结为建立并求解不等式.

用导数找极值/周期约束

例题 6.55 (2022 年 6 月全国甲卷 (理) 第 11 题) 设函数 $f(x)=\sin(\omega x+\frac{\pi}{3})$ 在区间 $(0,\pi)$ 恰有三个极值点,两个零点,则 ω 的取值范围是(　)

A. $\left(\frac{5}{3},\frac{13}{6}\right]$　　C. $\left(\frac{13}{6},\frac{8}{3}\right]$

B. $\left(\frac{5}{3},\frac{19}{6}\right)$　　D. $\left(\frac{13}{6},\frac{19}{6}\right)$

解 我们从整体换元与特征点定位两个角度处理.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

法一采用整体换元法. 令 $t=\omega x+\frac{\pi}{3}$, 则 $f(x)$ 的性质等价于 $y=\sin t$ 的性质. 当 $x\in(0,\pi)$ 时,

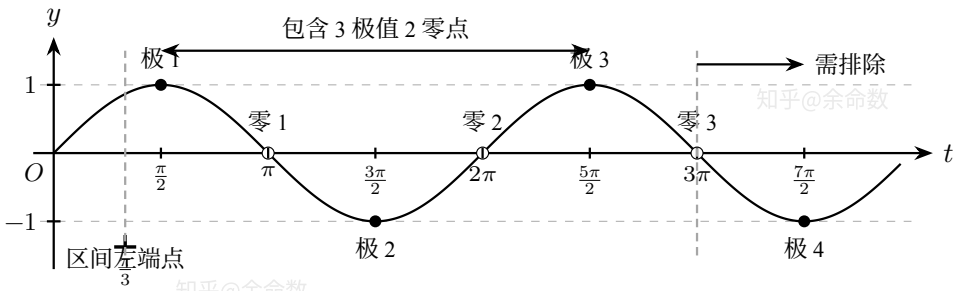
$$t\in\left(\frac{\pi}{3},\omega\pi+\frac{\pi}{3}\right).$$

要求 f 在 $(0,\pi)$ 内恰有三个极值点、两个零点, 等价于 $\sin t$ 在上述区间内恰有三个极值点、两个零点.

$t>0$ 时 $\sin t$ 的极值点与零点分别如下.

- 极值点 ($\sin t=\pm1$).$\frac{\pi}{2}$, $\frac{3\pi}{2}$, $\frac{5\pi}{2}$, $\frac{7\pi}{2}$, ...
- 零点 ($\sin t=0$).π, 2π, 3π, 4π, ...

区间左端为 $t=\frac{\pi}{3}$. 恰含三个极值点时, 这三点只能是 $\frac{\pi}{2}$, $\frac{3\pi}{2}$, $\frac{5\pi}{2}$; 恰含两个零点时, 这两点只能是 π, 2π.



区间 $\left(\frac{\pi}{3},\omega\pi+\frac{\pi}{3}\right)$ 必须包含的最后一个点是第三个极值点 $\frac{5\pi}{2}$. 同时, 它必须排除第三个零点 3π 和第四个极值点 $\frac{7\pi}{2}$. 其中, 3π 是更小的需要被排除的点. 因此, 区间的右端点 $\omega\pi+\frac{\pi}{3}$ 必须满足.

$$\frac{5\pi}{2}<\omega\pi+\frac{\pi}{3}\leq3\pi$$

由于 $x\in(0,\pi)$, 即使 $\omega\pi+\frac{\pi}{3}=3\pi$ 成立, 对应的 x 值也不在原区间内, 因此右边界可以取等号. 解此不等式组.

$$\frac{5\pi}{2}-\frac{\pi}{3}<\omega\pi\leq3\pi-\frac{\pi}{3}$$

$$\frac{13\pi}{6}<\omega\pi\leq\frac{8\pi}{3}$$

$$\frac{13}{6}<\omega\leq\frac{8}{3}$$

法二采用特征点定位法(ω 卡根法). 直接写出极值点与零点的横坐标通式. 由 $f(x)=0$ 得 $\omega x+\frac{\pi}{3}=k\pi$, 零点 $x_k=\frac{(3k-1)\pi}{3\omega}$, $k\in\mathbb{Z}$. 由 $f'(x)=0$ 得 $\omega x+\frac{\pi}{3}=k\pi+\frac{\pi}{2}$, 极值点 $x'_k=\frac{(6k+1)\pi}{6\omega}$, $k\in\mathbb{Z}$.

需在 $(0,\pi)$ 内恰有三个正极值点、两个正零点.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 零点 $(x_k>0).k=1,2,3,\cdots \implies x_1=\frac{2\pi}{3\omega},x_2=\frac{5\pi}{3\omega},x_3=\frac{8\pi}{3\omega},\ldots$
- 极值点 $(x'_k>0).k=0,1,2,3,\cdots \implies x'_0=\frac{\pi}{6\omega},x'_1=\frac{7\pi}{6\omega},x'_2=\frac{13\pi}{6\omega},x'_3=\frac{19\pi}{6\omega},\ldots$

知乎@余命数

于是第三个极值点 x'_2 与第二个零点 x_2 须在 $(0,\pi)$ 内, 而第四个极值点 x'_3 与第三个零点 x_3 须在区间外, 即 $\geq\pi$. 等价地说, 区间内诸点中最大者小于 π, 区间外诸点中最小者 $\geq\pi$.

比较得 $x'_2=\frac{13\pi}{6\omega}$ 大于 $x_2=\frac{5\pi}{3\omega}$, 故区间内最大点为 x'_2; $x_3=\frac{8\pi}{3\omega}$ 小于 $x'_3=\frac{19\pi}{6\omega}$, 故区间外最小点为 x_3. 于是

$$\begin{cases}x'_2=\dfrac{13\pi}{6\omega}<\pi\\ x_3=\dfrac{8\pi}{3\omega}\geq\pi\end{cases}$$

分别解得 $\omega>\frac{13}{6}$ 与 $\omega\leq\frac{8}{3}$. 合并解得

$$\frac{13}{6}<\omega\leq\frac{8}{3}$$

知乎@余命数

知乎@余命数

两种方法结果一致, 故选 C.

求解此类关于 ω 的范围问题, 最终都是建立不等式.

知乎@余命数

- **整体换元法**侧重于图像理解. 它将问题转化为在标准正弦曲线上框选一个区间, 使得该区间恰好包含指定数量的特征点. 关键在于确定区间的左右端点必须卡在哪两个特征点之间.
- **特征点定位法**侧重于代数计算. 它通过求解通式, 直接定位最后一个应在区间内的点和第一个应在区间外的点, 然后利用区间端点将它们卡住. 这种方法需要细心处理多个点的通式和大小比较.

知乎@余命数

无论哪种方法, 处理不等式的边界条件 (等号能否取到) 都决定结果是否正确. 这通常取决于原区间的开闭性以及函数在端点处的取值情况.

我们根据函数在固定长度的区间内展现出的性质, 如零点、极值点的数量, 反向约束决定周期性伸缩的参数 ω.

问题等价于判断函数图像在区间 (a,b) 内能容纳多少个波形.ω 越大, 周期 $T=\frac{2\pi}{\omega}$ 越小, 图像被压缩得越厉害, 同样长度的区间内就能容纳更多的零点和极值点; ω 越小则反之.

知乎@余命数

我们的任务就是将“恰有 N 个极值点, M 个零点”这一性质描述, 翻译成一个关于 ω 的数学不等式.

于是, 卡位思想是解决此类问题的关键. 它指的是, 我们必须找到最后一个应该在区间内的特征点, 以及第一个不应该在区间内的特征点, 然后用区间的端点将它们卡住.

具体地, 假设区间为 (a,b), 得

知乎@余命数

- **内卡条件**. 最后一个需要被包含的特征点(设其横坐标为 $x_{\text{尾内}}$), 必须严格落在区间 (a,b) 内部. 这给出关于 ω 的一个下界约束.

$$x_{\text{尾内}} < b$$

- **外卡条件**. 第一个不应被包含的特征点(设其横坐标为 $x_{\text{首外}}$), 必须落在区间右侧或恰好在右端点上. 这给出关于 ω 的一个上界约束.

$$x_{\text{首外}} \geq b$$

联立这两个不等式, 即可解出 ω 的范围.

下面给出两条路径, 都用 $x_{\text{尾内}}$ 和 $x_{\text{首外}}$ 来建立不等式, 二者等价, 操作不同.

6.15.0.1 整体换元法(图像视角)

固定标准图像, 移动考察区间.

操作步骤.

1. **换元**. 令 $t = \omega x + \varphi$, 将问题转化为分析标准函数 $y = \sin t$.
2. **定区间**. 根据 $x \in (a,b)$, 确定新变量 t 的范围 $(\omega a + \varphi, \omega b + \varphi)$.
3. **数点**. 在 $y = \sin t$ 的标准图像上, 从 t 的左端点开始, 数出题目要求的 N 个极值点和 M 个零点.
4. **卡位**. 确定 t 的右端点 $\omega b + \varphi$ 必须满足.

$$t_{\text{尾内}} < \omega b + \varphi \leq t_{\text{首外}}$$

其中 $t_{\text{尾内}}$ 是最后一个应包含点的横坐标, $t_{\text{首外}}$ 是第一个应排除点的横坐标.

5. **求解**. 解此不等式, 得到 ω 的范围.

图像视角的思路直观, 与函数图像的联系紧密, 能清晰地看到特征点被框选的过程.

6.15.0.2 特征点定位法(代数视角)

固定考察区间, 计算伸缩后的特征点.

操作步骤.

1. **求通式**. 直接写出原函数 $f(x)$ 的所有极值点横坐标通式 x_k' 和零点横坐标通式 x_k(均含有 ω).
2. **定序号**. 根据 $x > a$ 的条件, 确定满足题意的第 1、2、……、N 个极值点和第 1、2、……、M 个零点的具体表达式.
3. **找关键点**. 在所有应包含的点中, 找出横坐标最大的那个 $x_{\text{尾内}}$. 在所有不应包含的点中, 找出横坐标最小的那个 $x_{\text{首外}}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

4. **卡位**. 根据卡位思想, 建立不等式组.

$$\begin{cases} x_{\text{尾内}} < b \\ x_{\text{首外}} \geq b \end{cases}$$

5. **求解**. 解此不等式组, 得到 ω 的范围.

代数视角的方法程序化, 在特征点类型单一, 如只问零点或初相 φ 简单时, 计算可能更直接.

在建立不等式时, 等号能否取到, 决定最终结果是否正确. 判断依据如下.

- **看原区间开闭**. 若原区间是开区间 (a,b), 特征点恰好落在端点 b 上时, 它并不属于区间内. 因此, 外卡条件中的 $x_{\text{首外}}$ 可以取到 b, 即 $x_{\text{首外}} \geq b$.
- **看函数在端点取值**. 如果题目对函数在端点处的行为有特殊要求, 例如, 在闭区间 $[a,b]$ 内恰有 N 个极值点, 则需要额外考虑端点是否可能成为极值点, 这会影响不等式的取舍.

处理边界时, 必须仔细审题, 弄清题中恰好一词的确切含义.

例题 6.56 (单调性约束) 已知函数 $f(x)=\sin(\omega x+\frac{\pi}{4})$ 在区间 $(0,\frac{\pi}{2})$ 上单调递增, 求正实数 ω 的取值范围.

解 这里的条件是对函数在特定区间上单调性的要求. 要使函数在指定区间上单调, 该区间必须是函数某个标准单调区间的子集.

对于正弦函数 $y=\sin t$, 单调递增区间为 $\left[2k\pi-\frac{\pi}{2},2k\pi+\frac{\pi}{2}\right],k\in\mathbb{Z}$.

令 $t=\omega x+\frac{\pi}{4}$. 当 $x\in(0,\frac{\pi}{2})$ 时, t 的取值范围为

$$t\in\left(\omega\cdot 0+\frac{\pi}{4},\omega\cdot\frac{\pi}{2}+\frac{\pi}{4}\right)=\left(\frac{\pi}{4},\frac{\omega\pi}{2}+\frac{\pi}{4}\right)$$

要使 $f(x)$ 在 $(0,\frac{\pi}{2})$ 上单调递增, 就必须使区间 $(\frac{\pi}{4},\frac{\omega\pi}{2}+\frac{\pi}{4})$ 是 $y=\sin t$ 某个标准单调递增区间的子集.

由于该区间的左端点为 $\frac{\pi}{4}$, 它自然落在 $k=0$ 时所对应的标准递增区间 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 内部. 因此, 只需保证该区间的右端点不超过标准递增区间的右端点即可.

$$\frac{\omega\pi}{2}+\frac{\pi}{4}\leq\frac{\pi}{2}$$

解此不等式.

$$\frac{\omega\pi}{2}\leq\frac{\pi}{2}-\frac{\pi}{4}=\frac{\pi}{4}$$

$$\omega\leq\frac{1}{2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

又因为题目要求 ω 为正实数, 即 $\omega > 0$.

ω 的取值范围是 $(0, \frac{1}{2}]$.

例题 6.57 (零点个数与闭区间约束) 已知函数 $f(x) = \cos(\omega x + \frac{\pi}{6})$ 在闭区间 $[0, \pi]$ 上恰有三个零点, 求 ω 的取值范围.

解 这里把考察对象换成零点, 区间也换成闭区间. 端点处若恰有零点, 同样计入总数, 因此边界条件的判断更为严格.

令 $t = \omega x + \frac{\pi}{6}$. 当 $x \in [0, \pi]$ 时, t 的取值范围为

$$t \in \left[\omega \cdot 0 + \frac{\pi}{6}, \omega\pi + \frac{\pi}{6}\right] = \left[\frac{\pi}{6}, \omega\pi + \frac{\pi}{6}\right]$$

问题转化为, 标准余弦函数 $y = \cos t$ 在闭区间 $[\frac{\pi}{6}, \omega\pi + \frac{\pi}{6}]$ 上恰有三个零点.

标准余弦函数 $y = \cos t$ 的零点位于 $t = k\pi + \frac{\pi}{2}, k \in \mathbb{Z}$. 其正值零点依次如下: $\frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi}{2}, \ldots$ 考察区间从 $t = \frac{\pi}{6}$ 开始. 要使其恰好包含三个零点, 这三个零点必须是 $\frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}$.

根据卡位思想, 区间的右端点 $\omega\pi + \frac{\pi}{6}$ 必须大于等于第三个零点 $\frac{5\pi}{2}$, 同时必须严格小于第四个零点 $\frac{7\pi}{2}$.

$$\frac{5\pi}{2} \leq \omega\pi + \frac{\pi}{6} < \frac{7\pi}{2}$$

- 左边界. 当 $\omega\pi + \frac{\pi}{6} = \frac{5\pi}{2}$ 时, 第三个零点恰好落在区间的右端点上. 由于原区间是闭区间, 该点被计入, 满足恰有三个零点的要求, 故可以取等号.
- 右边界. 当 $\omega\pi + \frac{\pi}{6} = \frac{7\pi}{2}$ 时, 第四个零点被包含, 变为四个零点, 不满足题意, 故不能取等号.

解此不等式组.

$$\frac{5\pi}{2} - \frac{\pi}{6} \leq \omega\pi < \frac{7\pi}{2} - \frac{\pi}{6}$$

$$\frac{14\pi}{6} \leq \omega\pi < \frac{20\pi}{6}$$

$$\frac{7\pi}{3} \leq \omega\pi < \frac{10\pi}{3}$$

$$\frac{7}{3} \leq \omega < \frac{10}{3}$$

ω 的取值范围是 $[\frac{7}{3}, \frac{10}{3})$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 6.58 (对称性与极值唯一性约束) 已知函数 $f(x)=\sin(\omega x+\frac{\pi}{4})$ $(\omega>0)$ 的图像关于直线 $x=\frac{\pi}{2}$ 对称, 且在区间 $(0,\pi)$ 上有唯一最大值, 求 ω 的值.

解 这里的条件是对称性与极值唯一性的组合, 需要从两个性质出发, 建立关于 ω 的方程和不等式, 并求出它们的交集.

函数图像关于直线 $x=a$ 对称, 意味着 $x=a$ 必为函数的一条对称轴, 即函数在该点取得极值. 因此, $x=\frac{\pi}{2}$ 必须是函数的一个极值点. 令 $t=\omega x+\frac{\pi}{4}$, 则当 $x=\frac{\pi}{2}$ 时, t 必须满足 $t=k\pi+\frac{\pi}{2}, k\in\mathbb{Z}$.

$$\omega\cdot\frac{\pi}{2}+\frac{\pi}{4}=k\pi+\frac{\pi}{2}$$

$$\frac{\omega\pi}{2}=k\pi+\frac{\pi}{4}$$

$$\omega=2k+\frac{1}{2},\quad k\in\mathbb{Z}$$

再利用唯一最大值条件筛选. 当 $x\in(0,\pi)$ 时, 相位

$$t=\omega x+\frac{\pi}{4}$$

的取值区间为

$$\left(\frac{\pi}{4},\omega\pi+\frac{\pi}{4}\right).$$

若 $k\geq 1$, 则 $\omega=2k+\frac{1}{2}\geq\frac{5}{2}$, 于是相位区间至少为

$$\left(\frac{\pi}{4},\frac{11\pi}{4}\right),$$

其中同时包含正弦函数的两个最大值相位 $\frac{\pi}{2}$ 与 $\frac{5\pi}{2}$, 这会使 $f(x)$ 在 $(0,\pi)$ 上至少有两个最大值点, 与题意矛盾. 因此只能有 $k=0$.

结合 $\omega>0, k<0$ 不合题意; 而 $k=0$ 时 $\omega=\frac{1}{2}$, 此时相位区间为 $(\frac{\pi}{4},\frac{3\pi}{4})$, 其中恰含一个最大值相位 $\frac{\pi}{2}$, 符合题意. 因此, 唯一符合所有条件的 ω 值是 $\frac{1}{2}$.

ω 的值为 $\dfrac{1}{2}$.

上面三道变式的切入点不同, 但都遵循同一逻辑. **将函数的性质语言, 翻译为关于参数 ω 的代数约束语言**.

6.15.0.3 思想一. 单调性条件的区间包含

当题目给出的条件是单调性这类贯穿整个区间的性质时, 我们的策略从数点转变为区间包含.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1. **确定标准区间**. 写出基础函数, 如 $y=\sin t$ 具备该性质的标准区间通式. 例如, 单调递增区间为 $[2k\pi-\frac{\pi}{2},2k\pi+\frac{\pi}{2}]$.
2. **建立包含关系**. 利用整体换元法, 得到 t 的实际考察区间 $(\omega a+\varphi,\omega b+\varphi)$. 要使原函数在 (a,b) 上具备该性质, 就必须保证 t 的实际区间是某个标准性质区间的**子集**.
3. **转化为不等式**. 将子集关系转化为关于区间端点的不等式. 例如, 要使函数在 $(0,\frac{\pi}{2})$ 上单调递增, 就必须满足.

$$\left(\frac{\pi}{4},\frac{\omega\pi}{2}+\frac{\pi}{4}\right)\subseteq\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$$

知乎@余命数

这直接引出了关于 ω 的决定性不等式.

知乎@余命数

6.15.0.4 思想二. 开闭区间对等号的影响

原问题定义域的开闭性, 决定卡位不等式中等号能否取到.
在建立卡位不等式 $x_{\text{尾内}}\le b<x_{\text{首外}}$ 时, 必须仔细审视边界.

- **开区间** (a,b). 特征点若恰好落在端点 b 上, 则不计入区间内. 因此, 第一个区间外的点可以等于 b, 即 $x_{\text{首外}}\ge b$.
- **闭区间** $[a,b]$. 特征点若恰好落在端点 b 上, 则计入区间内. 例如变式二中, 第三个零点可以恰好等于区间的右端点, 从而不等式变为 $\frac{5\pi}{2}\le\omega\pi+\frac{\pi}{6}$.

对边界条件的精确判断, 依赖严格推理.

知乎@余命数

6.15.0.5 思想三. 多约束条件的分解与交集

当题目给出多个性质条件时, 应将它们分解, 逐一翻译成独立的数学约束, 最后求解这些约束的**交集**.

例如变式三中的问题可分解为两个独立任务.

1. **处理对称性**. 对称性通常对应函数在某点取得极值, 这是一个很强的条件, 往往能导出一个关于 ω 的**方程**, 将 ω 的取值范围从连续区间量化为离散点集, 如 $\omega=4m+\frac{1}{2}$.
2. **处理唯一性**. 极值点的唯一性, 是一个关于函数周期与区间长度关系的几何约束, 它通常导出一个关于 ω 的**不等式**, 为 ω 划定一个连续的取值范围, 如 $\omega\le 2$.
3. **求解交集**. 最后, 寻找在不等式划定的范围内, 满足方程结构的所有解. 这通常意味着在一个有限的离散点集中进行筛选, 从而得到唯一确定的答案.

化整为零, 分别约束, 再求交集, 是处理多条件问题的常用策略.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 6.59 (2022 年 6 月全国乙卷 (理) 第 15 题) 记函数 $f(x)=\cos(\omega x+\varphi)$ $(\omega>0,0<\varphi<\pi)$ 的最小正周期为 T. 若 $f(T)=\frac{\sqrt{3}}{2}$, $x=\frac{\pi}{9}$ 为 $f(x)$ 的零点, 则 ω 的最小值为 _____.

解 我们利用函数的两个特定点的函数值, 反向求解参数 φ 和 ω. 先用周期性确定 φ, 再用零点条件建立关于 ω 的方程, 最后根据 ω 的正性求其最小值.

根据最小正周期的定义, 有 $T=\frac{2\pi}{\omega}$, 故 $\omega T=2\pi$. 将 $x=T$ 代入第一个已知条件 $f(T)=\frac{\sqrt{3}}{2}$ 中.

$$f(T)=\cos(\omega T+\varphi)=\cos(2\pi+\varphi)=\frac{\sqrt{3}}{2}$$

利用余弦函数的周期性, 上式简化为.

$$\cos\varphi=\frac{\sqrt{3}}{2}$$

结合题目给出的范围 $0<\varphi<\pi$, 可唯一确定 φ 的值.

$$\varphi=\frac{\pi}{6}$$

已知 $x=\frac{\pi}{9}$ 是函数 $f(x)$ 的一个零点, 即 $f(\frac{\pi}{9})=0$.

$$f\left(\frac{\pi}{9}\right)=\cos\left(\omega\cdot\frac{\pi}{9}+\varphi\right)=0$$

余弦函数值为零的点, 其相位是 $\frac{\pi}{2}$ 的奇数倍. 因此,

$$\omega\frac{\pi}{9}+\varphi=k\pi+\frac{\pi}{2},\quad k\in\mathbb{Z}$$

将已求得的 $\varphi=\frac{\pi}{6}$ 代入, 得

$$\omega\frac{\pi}{9}+\frac{\pi}{6}=k\pi+\frac{\pi}{2}$$

从上述方程中解出 ω.

$$\omega\frac{\pi}{9}=k\pi+\frac{\pi}{2}-\frac{\pi}{6}=k\pi+\frac{\pi}{3}$$

两边同乘以 $\frac{9}{\pi}$, 得到 ω 关于整数 k 的表达式.

$$\omega=9k+3$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

题目要求 $\omega>0$, 因此.

$$9k+3>0 \implies 9k>-3 \implies k>-\frac{1}{3}$$

由于 k 必须是整数, 满足此条件的最小整数 k 为 0. 将 $k=0$ 代入, 即可得到 ω 的最小正值.

$$\omega_{\min}=9(0)+3=3$$

知乎@余命数

因此, ω 的最小值为 3.

例题 6.60 (2024 年 1 月浙江省宁波市宁波九校第 15 题) 将函数 $f(x)=\sin(\omega x+\frac{\pi}{6})$ $(\omega>0)$ 的图像向左平移 $\frac{\pi}{3}$ 个单位长度后得到函数 $g(x)$ 的图像, 且 $g(x)$ 在区间 $[\frac{\pi}{6},\frac{2\pi}{3}]$ 上单调递增, 则 ω 的取值范围是 _____.

解 我们先进行图像平移, 得到 $g(x)$ 的解析式, 再利用单调性条件建立不等式.

将函数 $f(x)$ 的图像向左平移 $\frac{\pi}{3}$ 个单位, 意味着用 $x+\frac{\pi}{3}$ 替换原解析式中的 x.

$$g(x)=f\left(x+\frac{\pi}{3}\right)=\sin\left[\omega\left(x+\frac{\pi}{3}\right)+\frac{\pi}{6}\right]=\sin\left(\omega x+\frac{\omega\pi}{3}+\frac{\pi}{6}\right)$$

问题转化为, 函数 $g(x)=\sin\left(\omega x+\frac{\omega\pi}{3}+\frac{\pi}{6}\right)$ 在区间 $[\frac{\pi}{6},\frac{2\pi}{3}]$ 上单调递增. 令 $t=\omega x+\frac{\omega\pi}{3}+\frac{\pi}{6}$. 要使 $g(x)$ 单调递增, 其内层函数 t 的取值区间必须是标准正弦函数 $y=\sin t$ 某个单调递增区间的子集.

先确定当 $x\in[\frac{\pi}{6},\frac{2\pi}{3}]$ 时, t 的取值范围.

- 当 $x=\frac{\pi}{6}$ 时(左端点), $t_{\min}=\omega\left(\frac{\pi}{6}\right)+\frac{\omega\pi}{3}+\frac{\pi}{6}=\frac{\omega\pi}{2}+\frac{\pi}{6}$.
- 当 $x=\frac{2\pi}{3}$ 时(右端点), $t_{\max}=\omega\left(\frac{2\pi}{3}\right)+\frac{\omega\pi}{3}+\frac{\pi}{6}=\omega\pi+\frac{\pi}{6}$.

因此, t 的取值区间为 $[\frac{\omega\pi}{2}+\frac{\pi}{6},\omega\pi+\frac{\pi}{6}]$.

标准正弦函数 $y=\sin t$ 的单调递增区间为 $[2k\pi-\frac{\pi}{2},2k\pi+\frac{\pi}{2}]$, $k\in\mathbb{Z}$. 由于 $\omega>0$, t 的区间下限 $\frac{\omega\pi}{2}+\frac{\pi}{6}$ 恒为正. 因此, 只需考虑 $k\geq 0$ 的情况.

当 $k=0$ 时, 标准递增区间为 $[-\frac{\pi}{2},\frac{\pi}{2}]$. t 的区间必须是该标准区间的子集, 得.

$$\begin{cases}\frac{\omega\pi}{2}+\frac{\pi}{6}\geq-\frac{\pi}{2} & (\text{此式对}\omega>0\text{ 恒成立})\\ \omega\pi+\frac{\pi}{6}\leq\frac{\pi}{2}\end{cases}$$

解第二个不等式.

$$\omega\pi\leq\frac{\pi}{2}-\frac{\pi}{6}=\frac{\pi}{3} \implies \omega\leq\frac{1}{3}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

还需检验是否存在 $k \geq 1$ 的解. 当 $k=1$ 时, 标准递增区间为 $[\frac{3\pi}{2}, \frac{5\pi}{2}]$. 此时需满足.

$$\begin{cases} \frac{\omega\pi}{2} + \frac{\pi}{6} \geq \frac{3\pi}{2} \implies \frac{\omega\pi}{2} \geq \frac{4\pi}{3} \implies \omega \geq \frac{8}{3} \\ \omega\pi + \frac{\pi}{6} \leq \frac{5\pi}{2} \implies \omega\pi \leq \frac{7\pi}{3} \implies \omega \leq \frac{7}{3} \end{cases}$$

$\omega \geq \frac{8}{3}$ 与 $\omega \leq \frac{7}{3}$ 矛盾, 故无解. 对于更大的 k 值, 同样无解.

结合 $\omega > 0$ 的前提条件, 最终的取值范围为 $0 < \omega \leq \frac{1}{3}$.

因此, ω 的取值范围是 $(0, \frac{1}{3}]$.

例题 6.61 (变式四) 已知函数 $f(x) = \cos(\omega x + \frac{\pi}{3})$ $(\omega > 0)$ 在区间 $[0, \frac{\pi}{2}]$ 上的值域为 $[-\frac{1}{2}, \frac{1}{2}]$, 求 ω 的取值范围.

解 这里的约束条件是函数在特定区间上的值域. 我们先分析内层函数 $t = \omega x + \frac{\pi}{3}$ 的取值区间, 再考察标准余弦函数 $y = \cos t$ 在该区间上的最值行为.

令 $t = \omega x + \frac{\pi}{3}$. 当 $x \in [0, \frac{\pi}{2}]$ 时, t 的取值范围为

$$t \in \left[\omega \cdot 0 + \frac{\pi}{3}, \omega \cdot \frac{\pi}{2} + \frac{\pi}{3}\right] = \left[\frac{\pi}{3}, \frac{\omega\pi}{2} + \frac{\pi}{3}\right]$$

问题转化为, 函数 $y = \cos t$ 在区间 $[\frac{\pi}{3}, \frac{\omega\pi}{2} + \frac{\pi}{3}]$ 上的值域为 $[-\frac{1}{2}, \frac{1}{2}]$.

由于 $\omega > 0$, 该区间的右端点大于左端点. 在 $[\frac{\pi}{3}, \pi]$ 上, $\cos t$ 单调递减. 左端点处函数值已经等于 $\frac{1}{2}$, 因而值域的最大值必为 $\frac{1}{2}$. 要使值域的最小值恰好为 $-\frac{1}{2}$, 区间右端点必须恰好到达第一个满足 $\cos t = -\frac{1}{2}$ 的点 $t = \frac{2\pi}{3}$; 若未到达, 最小值大于 $-\frac{1}{2}$, 若越过, 最小值小于 $-\frac{1}{2}$. 因此

$$\frac{\omega\pi}{2} + \frac{\pi}{3} = \frac{2\pi}{3}.$$

解得

$$\omega = \frac{2}{3}.$$

ω 的取值范围为 $\{\frac{2}{3}\}$.

例题 6.62 (变式五) 已知函数 $f(x) = \sin(\omega x)$ $(\omega > 0)$ 在区间 $[0, \pi]$ 上至少有一个最大值点, 但不存在最小值点, 求 ω 的取值范围.

解 这里的约束条件是至少有一个最大值点、同时不存在最小值点. 我们仍然用卡位思想处理.

对于标准正弦函数 $y = \sin t$,

- 其正的最大值点位于 $t = 2k\pi + \frac{\pi}{2}, k \in \mathbb{Z}$. 第一个正的最大值点是 $\frac{\pi}{2}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 其正的最小值点位于 $t=2k\pi+\frac{3\pi}{2},k\in\mathbb{Z}$. 第一个正的最小值点是 $\frac{3\pi}{2}$.

令 $t=\omega x$. 当 $x\in[0,\pi]$ 时, t 的取值范围是 $[0,\omega\pi]$. 问题转化为, 区间 $[0,\omega\pi]$ 需满足.

知乎@余命数

1. **至少包含一个最大值点**. 即区间的右端点必须大于等于第一个最大值点的位置.

$$\omega\pi\geq\frac{\pi}{2}$$

2. **不包含任何最小值点**. 即区间的右端点必须严格小于第一个最小值点的位置.

$$\omega\pi<\frac{3\pi}{2}$$

解第一个不等式.

$$\omega\geq\frac{1}{2}$$

知乎@余命数

解第二个不等式.

$$\omega<\frac{3}{2}$$

将两个条件合并, ω 的取值范围是 $[\frac{1}{2},\frac{3}{2})$.

知乎@余命数

切线斜率为负的导数表述

例题 6.63 (变式六) 已知函数 $f(x)=\sin(\omega x-\frac{\pi}{3})$ $(\omega>0)$ 满足对任意 $x\in(0,\pi)$, 都有 $f(x)+f(\pi-x)=0$. 若 $f(x)$ 在点 $(\frac{\pi}{2},f(\frac{\pi}{2}))$ 处的切线斜率为负, 求 ω 的最小值.

解 第一个条件给出隐式的对称性, 第二个条件约束函数在该对称中心处的变化趋势. 我们先把几何语言翻译为代数条件, 再结合导数信息求解.

条件 $f(x)+f(\pi-x)=0$ 对任意 $x\in(0,\pi)$ 成立, 描述的是函数图像关于点 $(\frac{\pi}{2},0)$ 中心对称. 要使点 $(\frac{\pi}{2},0)$ 成为 $y=\sin t$ 型函数的对称中心, 该点对应的相位 t 必须是 π 的整数倍. 令 $t=\omega x-\frac{\pi}{3}$. 将 $x=\frac{\pi}{2}$ 代入, 得

知乎@余命数

$$\omega\cdot\frac{\pi}{2}-\frac{\pi}{3}=k\pi,\quad k\in\mathbb{Z}$$

解出 ω 关于 k 的表达式.

$$\frac{\omega\pi}{2}=k\pi+\frac{\pi}{3}\implies\omega=2k+\frac{2}{3}$$

函数在点 $x=\frac{\pi}{2}$ 处的切线斜率为负, 即 $f'(\frac{\pi}{2})<0$. 首先求导数.

$$f'(x)=\omega\cos\left(\omega x-\frac{\pi}{3}\right)$$

将 $x=\frac{\pi}{2}$ 代入, 得

$$f'\left(\frac{\pi}{2}\right)=\omega\cos\left(\omega\frac{\pi}{2}-\frac{\pi}{3}\right)<0$$

由前面对对称中心的分析, $\omega\frac{\pi}{2}-\frac{\pi}{3}=k\pi$. 代入上式, 得

$$\omega\cos(k\pi)<0$$

因为已知 $\omega>0$, 所以必须满足 $\cos(k\pi)<0$. 这当且仅当 k 为**奇数**时成立.

现在有两个关于 ω 的条件.

1. $\omega=2k+\frac{2}{3}$
2. k 是奇数
3. $\omega>0$

要使 $\omega>0$, 则 $2k+\frac{2}{3}>0\implies k>-\frac{1}{3}$. 结合 k 为奇数, 满足条件的 k 的最小值为 1. 当 $k=1$ 时, 得 ω 的最小值.

$$\omega_{\min}=2(1)+\frac{2}{3}=\frac{8}{3}$$

因此, ω 的最小值为 $\dfrac{8}{3}$.

处理此类参数约束题时, 应将题设条件逐一转化为相位方程或不等式, 并充分利用所有约束信息.

例题 6.64 (2023 年 12 月重庆南开中学校高三第四次期中质检) 已知函数 $f(x)=\sin(\omega x+\varphi)$ $(\omega>0, 0<\varphi<\frac{\pi}{2})$, 曲线 $y=f(x)$ 的一个对称中心为 $(\frac{\pi}{12},0)$, 一条对称轴为 $x=\frac{\pi}{4}$, 则 ω 的最小值为 _____.

解 我们用一个对称中心与一条对称轴约束 ω 与 φ, 可以从代数方程与几何关系两条路推进.

法一采用代数法, 把几何性质写成相位方程后消元. 对称中心 $(\frac{\pi}{12},0)$ 要求相位为 π 的整数倍.

$$\omega\cdot\frac{\pi}{12}+\varphi=k_1\pi,\quad k_1\in\mathbb{Z}. \tag{6-11}$$

对称轴 $x=\frac{\pi}{4}$ 要求相位为 $\frac{\pi}{2}$ 的奇数倍:

$$\omega\cdot\frac{\pi}{4}+\varphi=k_2\pi+\frac{\pi}{2},\quad k_2\in\mathbb{Z}. \tag{6-12}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(6-12) 减去 (6-11) 消去 φ.

$$\omega\left(\frac{\pi}{4}-\frac{\pi}{12}\right)=(k_2-k_1)\pi+\frac{\pi}{2}\Longrightarrow\ \omega=6(k_2-k_1)+3.$$

再消 ω 以求 φ: 将 (6-11) 乘以 3 得 $\omega\frac{\pi}{4}+3\varphi=3k_1\pi$, 减去 (6-12),

$$2\varphi=\left(3k_1-k_2-\frac{1}{2}\right)\pi\ \Longrightarrow\ \varphi=\frac{3k_1-k_2-1/2}{2}\pi.$$

由 $0<\varphi<\frac{\pi}{2}$ 得 $\frac{1}{2}<3k_1-k_2<\frac{3}{2}$. k_1,k_2 为整数, 故 $3k_1-k_2$ 为整数, 唯一可能是 1, 即 $k_2=3k_1-1$. 代回 ω 的式子: $\omega=12k_1-3$. 由 $\omega>0$ 得 $k_1>\frac{1}{4}$, 最小整数 $k_1=1$, 此时

$$\omega_{\min}=9.$$

法二采用几何法. 对称中心与对称轴的水平距离必为四分之一周期的奇数倍. 取 $d=\left|\frac{\pi}{4}-\frac{\pi}{12}\right|=\frac{\pi}{6}$, 周期 $T=\frac{2\pi}{\omega}$, 则

$$\frac{\pi}{6}=\frac{2n+1}{4}T=\frac{(2n+1)\pi}{2\omega},\ n\in\{0,1,2,\dots\},$$

解得 $\omega=3(2n+1)$, 候选值 $\omega\in\{3,9,15,\dots\}$.

对每个候选值, 由对称中心条件 $\varphi=k\pi-\omega\frac{\pi}{12}$ 检验是否存在 $\varphi\in(0,\frac{\pi}{2})$. 当 $\omega=3$ 时 $\varphi=k\pi-\frac{\pi}{4}$, 对一切整数 k 均不在 $(0,\frac{\pi}{2})$ 内, 舍去. 当 $\omega=9$ 时, $k=1$ 给出 $\varphi=\frac{\pi}{4}\in(0,\frac{\pi}{2})$; 再验对称轴:

$$9\cdot\frac{\pi}{4}+\frac{\pi}{4}=\frac{5\pi}{2}$$

恰为 $\frac{\pi}{2}$ 的奇数倍, 故 $\omega=9$ 有效. 候选值随 n 递增, 第一个有效值即最小值.

两种方法均得 ω 的最小值为 9.

例题 6.65 (变式七) 已知函数 $f(x)=\cos(\omega x+\varphi)$ $(\omega>0,0<\varphi<\frac{\pi}{2})$ 的图像关于直线 $x=\frac{\pi}{3}$ 和 $x=\frac{5\pi}{6}$ 均对称, 求 ω 的最小值.

解 这里给出两条对称轴. 由对称轴可得到函数在特定点的相位信息; 两条对称轴之间的距离又与周期有确定的几何关系. 函数图像的两条对称轴之间的距离, 必定是半周期 $T/2$ 的整数倍. 两条对称轴的距离为.

$$d=\left|\frac{5\pi}{6}-\frac{\pi}{3}\right|=\frac{\pi}{2}$$

周期 $T=\frac{2\pi}{\omega}$. 因此有.

$$d=n\cdot\frac{T}{2},\quad n\in\{1,2,3,\dots\}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\frac{\pi}{2} = n \cdot \frac{1}{2} \cdot \frac{2\pi}{\omega} = \frac{n\pi}{\omega}$$

解出 ω 关于正整数 n 的表达式.

$$\omega = 2n, \quad n \in \{1, 2, 3, \dots\}$$

这给出了 ω 的一个离散候选值集 $\{2, 4, 6, 8, \dots\}$. 还需要利用 φ 的范围来筛选出符合条件的最小 n 值.

因为 $x = \frac{\pi}{3}$ 是一条对称轴, 所以函数在该点的相位必须是 π 的整数倍 (对于余弦函数).

$$\omega \cdot \frac{\pi}{3} + \varphi = k\pi, \quad k \in \mathbb{Z}$$

$$\varphi = k\pi - \omega\frac{\pi}{3}$$

将 $\omega = 2n$ 代入, 得

$$\varphi = k\pi - \frac{2n\pi}{3} = \left(k - \frac{2n}{3}\right)\pi$$

根据题设 $0 < \varphi < \frac{\pi}{2}$, 有

$$0 < \left(k - \frac{2n}{3}\right)\pi < \frac{\pi}{2} \implies 0 < k - \frac{2n}{3} < \frac{1}{2}$$

需寻找最小的正整数 n, 使得存在整数 k 满足此不等式.

- 当 $n = 1$ 时, $0 < k - \frac{2}{3} < \frac{1}{2} \implies \frac{2}{3} < k < \frac{7}{6}$. 此时整数 $k = 1$ 满足条件.

因此, 满足条件的最小正整数 n 是 1. 对应的 ω 最小值为 $\omega_{\min} = 2(1) = 2$.

ω 的最小值为 2.

例题 6.66 (变式八) 已知函数 $f(x) = \cos(\omega x + \varphi)$ $(\omega > 0, 0 < \varphi < \pi)$. 若将 $y = f(x)$ 的图像向右平移 $\frac{\pi}{12}$ 个单位后得到的函数 $g(x)$ 为偶函数, 且 $f(\frac{\pi}{6}) = 0$, 求 ω 的最小值.

解 这里把对称性条件施加在一个变换后的函数上. 我们先进行图像变换的代数运算, 再将函数的性质转化为关于参数的方程组. 函数 $g(x) = f(x - \frac{\pi}{12}) = \cos\left(\omega(x - \frac{\pi}{12}) + \varphi\right) = \cos\left(\omega x - \frac{\omega\pi}{12} + \varphi\right)$. $g(x)$ 为偶函数, 意味着其图像关于 y 轴对称. 对于余弦型函数, 这要求其相位中的常数项 (不含 x 的部分) 为 π 的整数倍.

$$-\frac{\omega\pi}{12} + \varphi = k\pi, \quad k \in \mathbb{Z} \tag{6-13}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$f(\frac{\pi}{6})=0$ 意味着函数在 $x=\frac{\pi}{6}$ 处的相位是 $\frac{\pi}{2}$ 的奇数倍.

$$\omega\frac{\pi}{6}+\varphi=m\pi+\frac{\pi}{2},\quad m\in\mathbb{Z} \tag{6-14}$$

由方程 (6-13) 可得 $\varphi=k\pi+\frac{\omega\pi}{12}$. 将其代入方程 (6-14):

$$\omega\frac{\pi}{6}+\left(k\pi+\frac{\omega\pi}{12}\right)=m\pi+\frac{\pi}{2}$$

$$\omega\left(\frac{\pi}{6}+\frac{\pi}{12}\right)=(m-k)\pi+\frac{\pi}{2}$$

$$\omega\frac{\pi}{4}=(m-k)\pi+\frac{\pi}{2}$$

令 $n=m-k$, n 仍为整数.

$$\omega=4n+2,\quad n\in\mathbb{Z}$$

将 $\omega=4n+2$ 代回 φ 的表达式.

$$\varphi=k\pi+\frac{(4n+2)\pi}{12}=k\pi+\frac{(2n+1)\pi}{6}$$

根据题设 $0<\varphi<\pi$, 有, 得

$$0<k\pi+\frac{(2n+1)\pi}{6}<\pi\implies 0<k+\frac{2n+1}{6}<1$$

$$-\frac{2n+1}{6}<k<1-\frac{2n+1}{6}=\frac{5-2n}{6}$$

需寻找最小的整数 n(使得 $\omega>0$), 使得上述关于 k 的不等式区间内存在整数解.

- 因为 $\omega=4n+2>0$, 所以 $n>-1/2$. 故 n 的最小整数取值为 0.
- 当 $n=0$ 时, 不等式变为 $-\frac{1}{6}<k<\frac{5}{6}$. 此时存在整数解 $k=0$.

因此, 使得 ω 取最小值的整数 n 为 0.

$$\omega_{\min}=4(0)+2=2$$

ω 的最小值为 2.

例题 6.67 (变式九) 已知函数 $f(x)=\sin(\omega x+\varphi)$ $(\omega>0,0<\varphi<\frac{\pi}{2})$ 满足对任意 $x\in\mathbb{R}$, 都有 $f(\frac{\pi}{3}+x)=f(\frac{\pi}{3}-x)$, 且函数在区间 $(0,\frac{\pi}{2})$ 内恰有两个零点, 求 ω 的取值范围.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 这里将一个明确的对称轴条件与一个零点数量的卡位条件相结合, 需要建立一个关于 ω 和整数参数 k 的不等式系统来求解.

$f(\frac{\pi}{3}+x)=f(\frac{\pi}{3}-x)$ 表明, 直线 $x=\frac{\pi}{3}$ 是函数的一条对称轴. 因此, 函数在 $x=\frac{\pi}{3}$ 处的相位必须是 $\frac{\pi}{2}$ 的奇数倍.

$$\omega\frac{\pi}{3}+\varphi=k\pi+\frac{\pi}{2},\quad k\in\mathbb{Z} \tag{6-15}$$

函数在 $(0,\frac{\pi}{2})$ 内恰有两个零点. 令 $t=\omega x+\varphi$. 当 $x\in(0,\frac{\pi}{2})$ 时, t 的取值范围是 $(\varphi,\frac{\omega\pi}{2}+\varphi)$. 函数的零点对应 $t=m\pi, m\in\mathbb{Z}$. 由于 $0<\varphi<\frac{\pi}{2}$, t 的区间左端点 φ 在 $(0,\frac{\pi}{2})$ 内. 因此, 区间内可能包含的第一个零点是 $t=\pi$. 要恰好包含两个零点, 这两个零点必须是 π 和 2π. 根据卡位思想, 区间 $(\varphi,\frac{\omega\pi}{2}+\varphi)$ 必须包含 2π, 但不包含 3π.

$$2\pi<\frac{\omega\pi}{2}+\varphi\le 3\pi\quad\cdots(*)$$

现在有一个方程和一个不等式, 它们共同约束着 ω 和 φ. 从方程 (6-15) 中解出 φ:

$$\varphi=k\pi+\frac{\pi}{2}-\frac{\omega\pi}{3}$$

利用 φ 的范围 $0<\varphi<\frac{\pi}{2}$:

$$0<k\pi+\frac{\pi}{2}-\frac{\omega\pi}{3}<\frac{\pi}{2}\implies -\frac{\pi}{2}<k\pi-\frac{\omega\pi}{3}<0\implies -\frac{1}{2}<k-\frac{\omega}{3}<0$$

$$k<\frac{\omega}{3}<k+\frac{1}{2} \tag{6-16}$$

这个关系表明, 对于一个给定的 ω, 整数 k 的取值是唯一确定的.

接着, 将 φ 的表达式代入不等式 $(*)$:

$$2\pi<\frac{\omega\pi}{2}+\left(k\pi+\frac{\pi}{2}-\frac{\omega\pi}{3}\right)\le 3\pi$$

$$2\pi<\frac{\omega\pi}{6}+k\pi+\frac{\pi}{2}\le 3\pi$$

$$\frac{3\pi}{2}<\frac{\omega\pi}{6}+k\pi\le\frac{5\pi}{2}$$

$$\frac{3}{2}-k<\frac{\omega}{6}\le\frac{5}{2}-k\implies 9-6k<\omega\le 15-6k$$

现在需要在 $\omega>0$ 的前提下, 寻找存在整数 k 使得不等式 $9-6k<\omega\le 15-6k$ 和不等式 (6-16) $(3k<\omega<3k+1.5)$ 同时成立.

为使两个关于 ω 的区间有交集, 区间的端点必须满足.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- $3k < 15 - 6k \implies 9k < 15 \implies k < \frac{5}{3}$
- $9 - 6k < 3k + 1.5 \implies 7.5 < 9k \implies k > \frac{5}{6}$

结合 k 为整数, 唯一可能的取值是 $k = 1$.

当 $k = 1$ 时, 将两个区间具体化.

- 由 (6-16) 得: $3(1) < \omega < 3(1) + 1.5 \implies 3 < \omega < 4.5$.
- 由主不等式得: $9 - 6(1) < \omega \leq 15 - 6(1) \implies 3 < \omega \leq 9$.

取这两个区间的交集, 得到 ω 的最终取值范围.

因此, ω 的取值范围是 $(3, \frac{9}{2})$.

通过对例题及其变式题的剖析, 我们总结出处理复杂 ω 约束问题的几点原则.

知乎@余命数

6.15.0.6 原则一. 对称性的两种处理

对称性是三角函数最重要的性质之一, 也是 ω 约束问题的高频考点.

6.15.0.6.1　1. 代数法　将几何性质直接转化为代数方程.

- **对称轴** $x = x_0$ 当且仅当相位 $\omega x_0 + \varphi = k\pi + \frac{\pi}{2}$ 对 sin 成立, 或 $k\pi$ 对 cos 成立.
- **对称中心** $(x_0, 0)$ 当且仅当相位 $\omega x_0 + \varphi = k\pi$(对 sin)或 $k\pi + \frac{\pi}{2}$ (对 cos.
- **偶函数**当且仅当初相 $\varphi = k\pi$(对 cos)或 $k\pi + \frac{\pi}{2}$ (对 sin.
- **隐式对称** $f(a + x) = f(a - x)$ 当且仅当对称轴为 $x = a$.
- **隐式对称** $f(a + x) = -f(a - x)$ 当且仅当对称中心为 $(a, 0)$.

知乎@余命数

6.15.0.6.1.1　说明　代数翻译是最直接的方法. 它将几何性质转化为关于 ω, φ 和整数 k, m 的不定方程组. 当涉及多个整数参数时, 需要细致地讨论整数解的存在性.

知乎@余命数

6.15.0.6.2　2. 几何法　利用对称点之间的几何距离关系.

- 两条**同类**对称点(轴与轴, 中心与中心)的距离是**半周期**的整数倍.
- 两条**异类**对称点(轴与中心)的距离是**四分之一周期**的奇数倍.

6.15.0.6.2.1　说明　几何法在题目给出两个对称点时, 能迅速建立一个仅关于 ω 和一个整数参数 n 的方程, 简化计算. 但候选集必须代回原条件, 结合 φ 的范围进行检验, 以排除不符合初相条件的伪解.

6.15.0.7 原则二. 多约束条件的分解

当题目给出多个不同类型的约束条件时, 如例 20 及其变式, 分而治之是有效的策略.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1. **分解**. 将复杂的题目描述分解为若干个独立的性质陈述. 例如, 平移后为偶函数且某点为零点.
2. **翻译**. 对每一个性质陈述, 选择最恰当的方法(代数或几何), 将其翻译为关于参数的方程或不等式约束.
3. **整合**. 联立求解所有约束. 这个过程通常演变为.
   - 方程与方程联立, 通过消元法求解.
   - 方程与不等式联立, 将由方程解出的参数表达式代入不等式, 从而约束整数参数的取值.
   - 不等式与不等式联立, 求解不等式组的交集.

知乎@余命数

6.15.0.7.0.1 说明 这种分治-整合的策略将复杂问题拆解为若干子问题, 降低出错风险. 关键在于准确地翻译每一个条件, 并在最后进行严谨的逻辑整合.

6.15.0.8 原则三. 存在性与唯一性

知乎@余命数

从至少或至多有一个, 到恰好有一个, 问题的逻辑复杂度在提升.

知乎@余命数

- **存在性**(至少/至多)通常对应一个单边的不等式. 例如, 至少一个最大值点意味着考察区间的右端点必须越过第一个最大值点, 即 $\omega\pi \geq \frac{\pi}{2}$.
- **唯一性**(恰好一个)通常对应一个双边不等式, 即卡位思想. 例如, 唯一最大值意味着区间必须包含第一个最大值点, 同时排除第二个最大值点和第一个最小值点.

6.15.0.8.0.1 说明 处理这类逻辑词汇时, 要把它精确地转化为区间端点的开闭与大小关系. 这既考验计算, 也考验对数学语言精确性的把握.

知乎@

- 对于**单一性质**(单调性、值域、零点个数), **整体换元法**通常是最稳健的选择.
- 对于**两个或以上明确的对称点**, **几何距离法**是首选的捷径, 但务必进行后续检验.

知乎@余命数

- 对于**复合的、混合的性质约束**, **建立方程组(代数法)**, 并采用分治-整合的策略, 是最可靠的方法.

ω 约束问题的每种性质都有明确的几何意义, 需要在代数表示与几何意义之间切换.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

6.15.1 高考真题选讲

例题 6.68 (2008 年辽宁卷理科数学第 16 题) 已知 $f(x)=\sin\left(\omega x+\dfrac{\pi}{3}\right)$ $(\omega>0)$, $f\left(\dfrac{\pi}{6}\right)=f\left(\dfrac{\pi}{3}\right)$, 且 $f(x)$ 在区间 $\left(\dfrac{\pi}{6},\dfrac{\pi}{3}\right)$ 上有最小值无最大值, 则 $\omega=$_____.

解 答案: $\dfrac{14}{3}$.

由 $f\left(\dfrac{\pi}{6}\right)=f\left(\dfrac{\pi}{3}\right)$ 知, $x=\dfrac{\pi}{4}$, 即 $\dfrac{\pi}{6}$ 与 $\dfrac{\pi}{3}$ 的中点是一条对称轴. 又因 $f(x)$ 在 $\left(\dfrac{\pi}{6},\dfrac{\pi}{3}\right)$ 上有最小值无最大值, 该最小值在 $x=\dfrac{\pi}{4}$ 处取得, 故

$$\omega\cdot\frac{\pi}{4}+\frac{\pi}{3}=2k\pi-\frac{\pi}{2},\quad k\in\mathbb{Z},$$

解得 $\omega=\dfrac{24k-10}{3}$.

由 $\omega>0$ 知 $k\geqslant 1$. 当 $k=1$ 时, $\omega=\dfrac{14}{3}$, 半周期 $\dfrac{\pi}{\omega}=\dfrac{3\pi}{14}$ 大于区间长度 $\dfrac{\pi}{6}$, 故区间内恰含一个最小值点而不含最大值点, 符合题意. 当 $k=2$ 时, $\omega=\dfrac{38}{3}$, 半周期 $\dfrac{3\pi}{38}$ 小于区间长度 $\dfrac{\pi}{6}$, 区间内已出现最大值, 不合题意.

故 $\omega=\dfrac{14}{3}$.

例题 6.69 (2018 年北京卷理科数学第 11 题) 设函数 $f(x)=\cos\left(\omega x-\dfrac{\pi}{6}\right)$ $(\omega>0)$, 若 $f(x)\leqslant f\left(\dfrac{\pi}{4}\right)$ 对任意实数 x 都成立, 则 ω 的最小值为_____.

解 答案: $\dfrac{2}{3}$.

$f(x)\leqslant f\left(\dfrac{\pi}{4}\right)$ 对任意实数 x 成立, 说明 $x=\dfrac{\pi}{4}$ 是最大值点. 对余弦型函数, 最大值点处的相位等于 $2k\pi$:

$$\omega\cdot\frac{\pi}{4}-\frac{\pi}{6}=2k\pi,\quad k\in\mathbb{Z},$$

解得 $\omega=8k+\dfrac{2}{3}$.

由 $\omega>0$ 知 $k\geqslant 0$. 当 $k=0$ 时, $\omega=\dfrac{2}{3}$, 即为所求最小值.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 6.70 (2018 年江苏卷第 7 题) 已知函数 $y=\sin(2x+\varphi)\left(-\dfrac{\pi}{2}<\varphi<\dfrac{\pi}{2}\right)$ 的图象关于直线 $x=\dfrac{\pi}{3}$ 对称, 则 φ 的值为_____.

知乎@余命数

解 答案: $-\dfrac{\pi}{6}$.

正弦型函数的对称轴处, 相位等于 $k\pi+\dfrac{\pi}{2}$:

$$2\cdot\frac{\pi}{3}+\varphi=k\pi+\frac{\pi}{2},\quad k\in\mathbb{Z},$$

解得 $\varphi=k\pi-\dfrac{\pi}{6}$.

由 $-\dfrac{\pi}{2}<\varphi<\dfrac{\pi}{2}$ 知 $k=0$, 故 $\varphi=-\dfrac{\pi}{6}$.

知乎@余命数

例题 6.71 (2013 年山东卷理科数学第 5 题) 函数 $y=\sin(2x+\varphi)$ 的图象沿 x 轴向左平移 $\dfrac{\pi}{8}$ 个单位后, 得到一个偶函数的图象, 则 φ 的一个可能的值为

A. $\dfrac{3\pi}{4}$　　C. 0

B. $\dfrac{\pi}{4}$　　D. $-\dfrac{\pi}{4}$

知乎@余命数 知乎@余命数 知乎@余命数

解 答案: B.

向左平移 $\dfrac{\pi}{8}$ 个单位后, 函数变为

$$y=\sin\left[2\left(x+\frac{\pi}{8}\right)+\varphi\right]=\sin\left(2x+\frac{\pi}{4}+\varphi\right).$$

该函数为偶函数, 故初相 $\dfrac{\pi}{4}+\varphi=k\pi+\dfrac{\pi}{2}$, $k\in\mathbb{Z}$, 即 $\varphi=k\pi+\dfrac{\pi}{4}$.

当 $k=0$ 时, $\varphi=\dfrac{\pi}{4}$, 对应选项 B.

知乎@余命数 知乎

6.16 反三角函数

函数 $y=f(x)$ 存在反函数 $x=f^{-1}(y)$ 的充要条件是 f 为一一映射. 三角函数由于周期性都是多对一, 例如 $\frac{\pi}{6}, \frac{5\pi}{6}, \frac{13\pi}{6}, \ldots$ 都对应同一个正弦值 $\frac{1}{2}$.

因此不能直接把整个三角函数反转, 否则一个输入值会对应无穷多个输出角, 不符合函数定义.

知乎@余命数

解决方法是限制定义域, 截取一段单调且值域完整的分支. 这段分支需满足两个条件.

1. **完备性**. 函数值恰好覆盖原函数的全部值域. 例如 $y=\sin x$ 在选定区间上的值域必须是 $[-1,1]$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. **单调性**. 函数在该区间上严格单调,从而是一一映射,保证反函数存在且唯一.

满足上述条件的区间称为**主值区间**. 在其上构造的反函数称为**反三角函数**. 它要回答的问题是, 在主值区间内哪个角的三角函数值等于给定值.

6.16.1 反正弦函数

6.16.1.1 定义与主值区间的选取

考察函数 $y=\sin x$. 为了定义其反函数, 需选取一个单调区间, 使得该区间上的函数值域覆盖 $[-1,1]$.

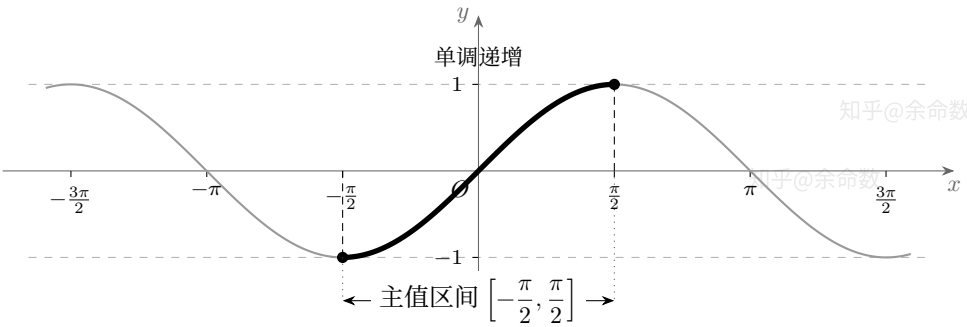


图 6.23　正弦函数主值区间 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 的选取

观察图像可知, 区间 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 是一个理想的选择. 它同时满足满射性(值域恰好覆盖 $[-1,1]$)、严格单调递增, 以及关于原点对称并包含锐角范围的对称性, 因而便于计算与记忆. 虽然 $[\frac{\pi}{2},\frac{3\pi}{2}]$ 等区间也满足单调性和值域条件, 但数学界统一规定选取 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 作为反正弦函数的主值区间.

定义 6.5 (反正弦) 函数 $y=\sin x$ 在其主值区间 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 上的反函数, 称为**反正弦函数**, 记作 $y=\arcsin x$.

- **定义域**: $[-1,1]$
- **值域**: $[-\frac{\pi}{2},\frac{\pi}{2}]$

符号 $\arcsin x$ 表达的意义是. 一个在 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 内的角, 其正弦值为 x.

6.16.1.2 性质与图像

- **单调性**: 正弦函数在主值区间的单调递增性, 决定了反正弦函数在其整个定义域 $[-1,1]$ 上也是**单调递增**的.
- **奇偶性**: 正弦函数主值支的图像关于原点对称, 这一几何特性被反函数继承. 故 $y=\arcsin x$ 是**奇函数**, 满足 $\arcsin(-x)=-\arcsin x$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

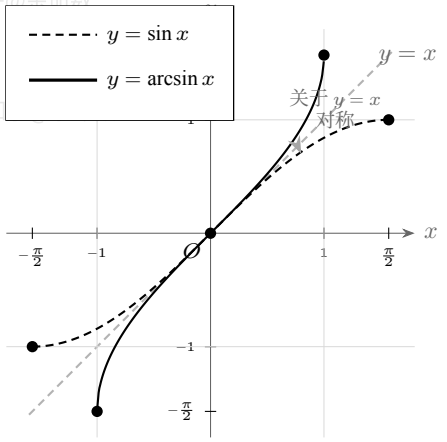
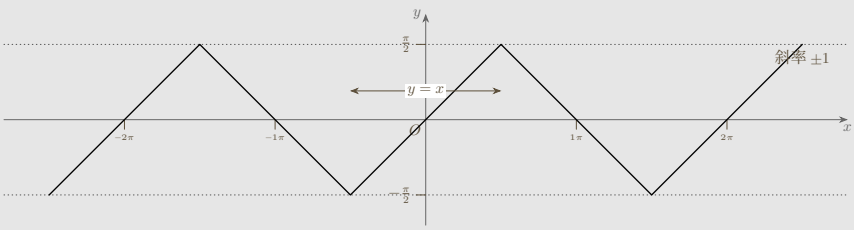


图 6.24 反正弦函数与正弦函数主值支的对称关系

注 6.17 ($\sin(\arcsin x)$ 与 $\arcsin(\sin x)$)
根据反函数的定义, $\sin(\arcsin x)=x$ 与 $\arcsin(\sin x)=x$ 两个式子看似相同, 但成立范围不同.
此关系式对于 $\arcsin x$ 的整个定义域 $x\in[-1,1]$ 是**恒成立**的.
此关系式**仅当** $x\in[-\frac{\pi}{2},\frac{\pi}{2}]$(主值区间)时成立.

- 当 x 超出主值区间时, 运算结果会将 x **折叠**回主值区间.
- 函数 $y=\arcsin(\sin x)$ 的图像是周期为 2π 的锯齿波.



6.16.2 反余弦函数

6.16.2.1 定义与主值区间的选取

考察 $y=\cos x$. 区间 $[-\frac{\pi}{2},\frac{\pi}{2}]$ 上 $\cos x$ 先增后减, 不是单射, 不能作为反余弦的主值区间.

为了定义反函数, 需选取一个既能覆盖值域 $[-1,1]$ 又严格单调的区间. 数学界选定 $[0,\pi]$ 作为标准.

- 在 $[0,\pi]$ 上, $\cos x$ 从 1 严格单调递减至 -1.
- 该区间是非负的, 且包含了 0 点, 形式最简.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

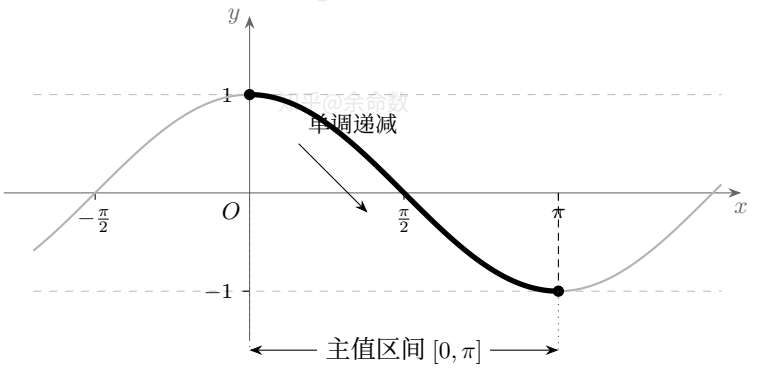


图 6.25　余弦函数主值区间 $[0,\pi]$ 的选取示意图

定义 6.6 (反余弦) 函数 $y=\cos x$ 在其主值区间 $[0,\pi]$ 上的反函数, 称为**反余弦函数**, 记作 $y=\arccos x$.

- **定义域**: $[-1,1]$
- **值域**: $[0,\pi]$

符号 $\arccos x$ 表达的意义是. 一个在 $[0,\pi]$ 内的角, 其余弦值为 x.

6.16.2.2 性质与图像

- **单调性**: 在其整个定义域 $[-1,1]$ 上**单调递减**.
- **奇偶性**: 既非奇函数也非偶函数. 但它具有一个重要的对称关系.

$$\arccos(-x)=\pi-\arccos x.$$

证明 (关于 $\arccos(-x)=\pi-\arccos x$) 令 $\alpha=\arccos x$. 由定义知 $\cos\alpha=x$ 且 $\alpha\in[0,\pi]$. 现证 $\pi-\alpha=\arccos(-x)$. 根据反余弦函数的定义, 只需验证两点.

1. $\cos(\pi-\alpha)=-x$.
2. $\pi-\alpha$ 落在反余弦的值域 $[0,\pi]$ 内.

第一点, 由诱导公式可得 $\cos(\pi-\alpha)=-\cos\alpha=-x$, 成立. 第二点, 因为 $0\le\alpha\le\pi$, 所以 $-\pi\le-\alpha\le0$, 故 $0\le\pi-\alpha\le\pi$, 也成立. 证毕. □

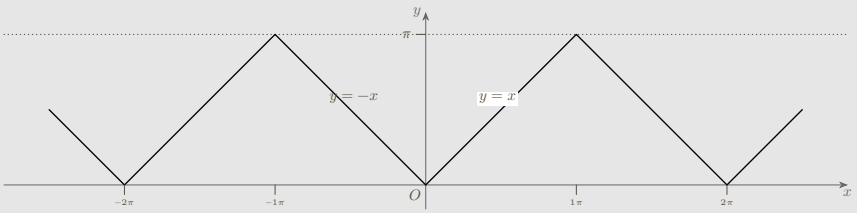
注 6.18 ($\cos(\arccos x)$ 与 $\arccos(\cos x)$)
1. 恒等式 $\cos(\arccos x)=x$
此关系式对于 $\arccos x$ 的定义域 $x\in[-1,1]$ 是**恒成立**的.
2. 关系式 $\arccos(\cos x)=x$ **的成立范围**
此关系式**仅当** $x\in[0,\pi]$(主值区间)时成立.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

- 当 x 超出此范围时, $\arccos(\cos x)$ 的结果是**返回到主值区间** $[0,\pi]$ **内**.
- 几何意义. 结果是与 x 终边位置具有相同余弦值, 且位于 0 到 π 之间的那个角.
- 例如.$\arccos(\cos(2\pi))=0$; $\arccos(\cos(-\frac{\pi}{3}))=\frac{\pi}{3}$.

函数 $y=\arccos(\cos x)$ 的图像是一条在 x 轴上方、周期为 2π 的三角波.



6.16.3 重要的恒等关系

定理 6.11 (余角关系) 对于任意 $x\in[-1,1]$, 恒有.

$$\arcsin x+\arccos x=\frac{\pi}{2} \tag{6-17}$$

证明 令 $\alpha=\arcsin x$. 由定义知 $\sin\alpha=x$ 且 $\alpha\in[-\frac{\pi}{2},\frac{\pi}{2}]$. 考察角 $\frac{\pi}{2}-\alpha$. 首先,其值域为 $0\le\frac{\pi}{2}-\alpha\le\pi$,这恰好是反余弦函数的值域. 其次,其值为 $\cos(\frac{\pi}{2}-\alpha)=\sin\alpha=x$. 根据反余弦函数的唯一定义,一个值域在 $[0,\pi]$ 且余弦值为 x 的角,必定是 $\arccos x$. 故 $\arccos x=\frac{\pi}{2}-\alpha=\frac{\pi}{2}-\arcsin x$. 移项即得证. □

定理 6.12 (反正切的加法公式) 对于任意实数 a,b,

$$\arctan a+\arctan b=\begin{cases}\arctan\left(\frac{a+b}{1-ab}\right), & \text{if } ab<1\\ \frac{\pi}{2}, & \text{if } ab=1 \text{ and } a>0\\ -\frac{\pi}{2}, & \text{if } ab=1 \text{ and } a<0\\ \pi+\arctan\left(\frac{a+b}{1-ab}\right), & \text{if } ab>1 \text{ and } a>0\\ -\pi+\arctan\left(\frac{a+b}{1-ab}\right), & \text{if } ab>1 \text{ and } a<0\end{cases} \tag{6-18}$$

证明 令 $A=\arctan a, B=\arctan b$. 则 $\tan A=a,\tan B=b$, 且 $A,B\in(-\frac{\pi}{2},\frac{\pi}{2})$. 考察其和 $A+B$ 的正切值.

$$\tan(A+B)=\frac{\tan A+\tan B}{1-\tan A\tan B}=\frac{a+b}{1-ab}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这说明 $A+B$ 与 $\arctan\left(\frac{a+b}{1-ab}\right)$ 的正切值相同,它们之间相差 π 的整数倍. 即 $A+B=k\pi+\arctan\left(\frac{a+b}{1-ab}\right)$. 要点在于确定整数 k 的值, 这取决于 $A+B$ 所在的区间.

情形一. $ab<1$ 因为 $\cos A>0$ 且 $\cos B>0$, 有

$$\cos(A+B)=\cos A\cos B(1-\tan A\tan B)=\cos A\cos B(1-ab)>0.$$

又 $A+B\in(-\pi,\pi)$, 所以 $A+B\in(-\frac{\pi}{2},\frac{\pi}{2})$. 这正是 arctan 的主值区间, 因而 $k=0$.

情形二. $ab=1$ 此时 a 与 b 同号. 若 $a>0$, 则 $A,B\in(0,\frac{\pi}{2})$ 且 $\tan A\tan B=1$, 故 $A+B=\frac{\pi}{2}$. 若 $a<0$, 则 $A,B\in(-\frac{\pi}{2},0)$, 同理 $A+B=-\frac{\pi}{2}$.

情形三. $ab>1$ **且** $a>0$ 此时必有 $b>0$.$A,B\in(0,\frac{\pi}{2})$, 故 $A+B\in(0,\pi)$. $ab>1\implies 1-ab<0\implies\frac{a+b}{1-ab}<0\implies\tan(A+B)<0$. 这说明 $A+B$ 位于第二或第四象限. 结合 $A+B\in(0,\pi)$, 可知 $A+B\in(\frac{\pi}{2},\pi)$. 而 $\arctan(\frac{a+b}{1-ab})$ 位于 $(-\frac{\pi}{2},0)$. 要使等式成立, 必须有 $k=1$.

情形四. $ab>1$ **且** $a<0$ 此时必有 $b<0$.$A,B\in(-\frac{\pi}{2},0)$, 故 $A+B\in(-\pi,0)$. $ab>1\implies 1-ab<0\implies\frac{a+b}{1-ab}>0\implies\tan(A+B)>0$. 这说明 $A+B$ 位于第一或第三象限. 结合 $A+B\in(-\pi,0)$, 可知 $A+B\in(-\pi,-\frac{\pi}{2})$. 而 $\arctan(\frac{a+b}{1-ab})$ 位于 $(0,\frac{\pi}{2})$. 要使等式成立, 必须有 $k=-1$. □

知乎@余命数　知乎@余命数　知乎@余命数

6.16.4 典例剖析

例题 6.72 (复合函数求值) 求 $\cos\left(\arcsin\left(-\frac{4}{5}\right)\right)$ 的值.

解 令 $\alpha=\arcsin\left(-\frac{4}{5}\right)$. 现求 $\cos\alpha$. 根据反正弦函数的定义, 可知.

$$\sin\alpha=-\frac{4}{5}\quad\text{且}\quad\alpha\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$$

利用同角三角函数平方关系 $\sin^2\alpha+\cos^2\alpha=1$,

$$\cos^2\alpha=1-\sin^2\alpha=1-\left(-\frac{4}{5}\right)^2=1-\frac{16}{25}=\frac{9}{25}$$

所以 $\cos\alpha=\pm\frac{3}{5}$.

接着, 利用角的范围确定符号. 因为 $\sin\alpha=-\frac{4}{5}$ 是负值, 且 $\alpha\in[-\frac{\pi}{2},\frac{\pi}{2}]$, 所以角 α 必定落在第四象限, 或其边界, 即 $\alpha\in[-\frac{\pi}{2},0)$. 在第四象限, 余弦值为正.

$$\therefore\cos\alpha=\frac{3}{5}$$

例题 6.73 (反正切的和) 求值.$\arctan 2+\arctan 3$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 直接应用反正切的加法公式. 令 $a=2, b=3$. 注意到 $a>0, b>0$ 且 $ab=2\times 3=6>1$. 因此, 应使用第二种情形的公式.

$$\begin{aligned}\arctan 2+\arctan 3 &= \pi+\arctan\left(\frac{2+3}{1-2\cdot 3}\right)\\ &= \pi+\arctan\left(\frac{5}{-5}\right)\\ &= \pi+\arctan(-1)\end{aligned}$$

根据反正切函数的定义, $\arctan(-1)$ 是在 $(-\frac{\pi}{2}, \frac{\pi}{2})$ 内正切值为 -1 的角, 即 $-\frac{\pi}{4}$.

$$\therefore \arctan 2+\arctan 3=\pi+\left(-\frac{\pi}{4}\right)=\frac{3\pi}{4}$$

例题 6.74 (透彻理解恒等式) 求值.$\arccos\left(\cos\left(\frac{8\pi}{3}\right)\right)$.

解 角 $\frac{8\pi}{3}$ 不在反余弦函数的主值区间 $[0,\pi]$ 内, 所以结果不是它本身. 先需计算内层函数的值.

$$\cos\left(\frac{8\pi}{3}\right)=\cos\left(2\pi+\frac{2\pi}{3}\right)=\cos\left(\frac{2\pi}{3}\right)=-\frac{1}{2}$$

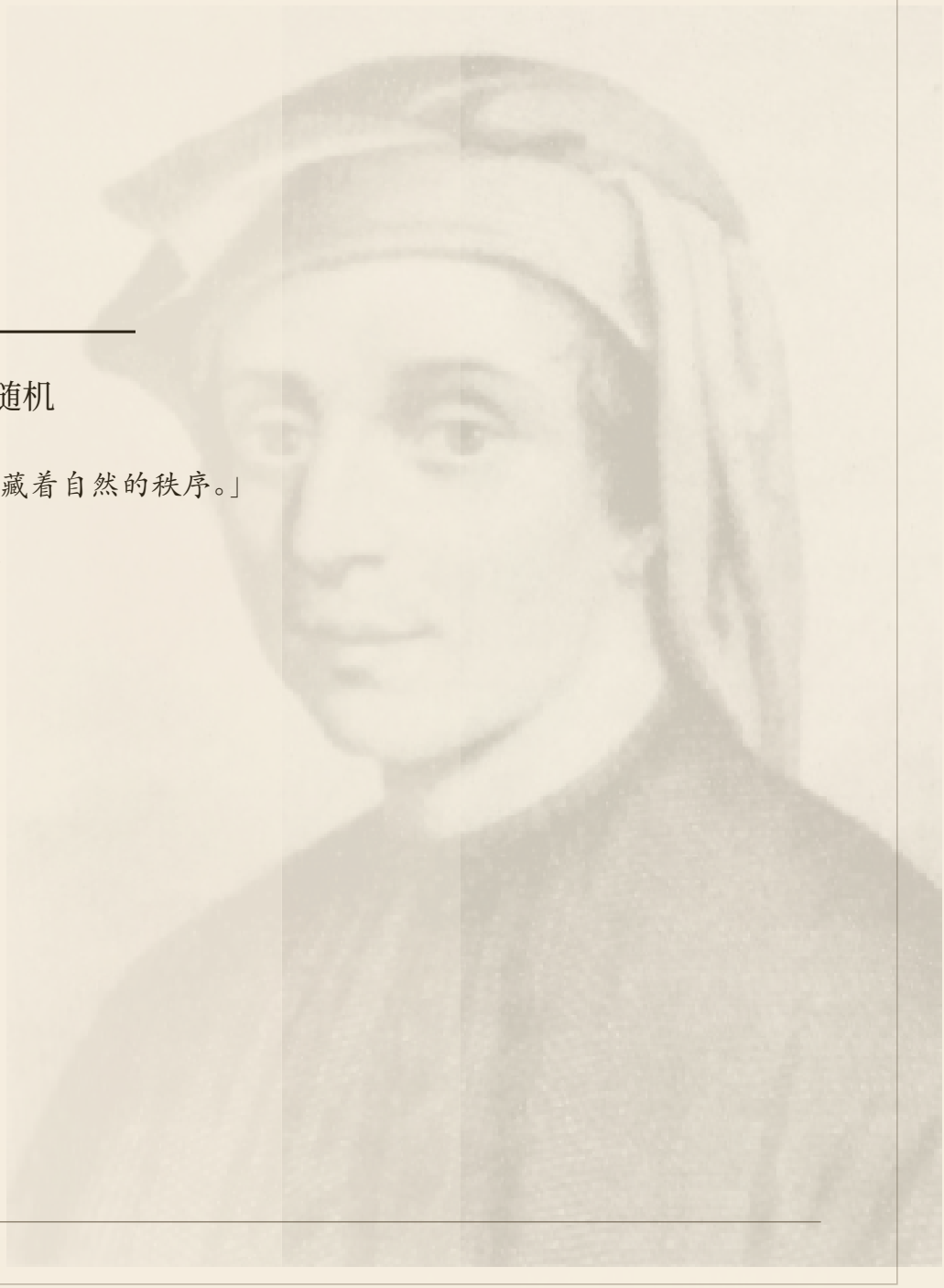
接着, 原问题转化为求 $\arccos(-\frac{1}{2})$. 需寻找在 $[0,\pi]$ 区间内, 余弦值为 $-\frac{1}{2}$ 的角. 该角为 $\frac{2\pi}{3}$.

$$\therefore \arccos\left(\cos\left(\frac{8\pi}{3}\right)\right)=\frac{2\pi}{3}$$

知乎@余命数

第 7 章

数列

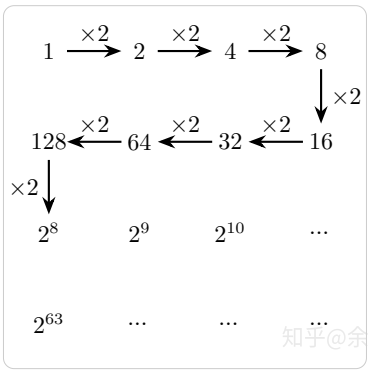


周期·离散·随机

「序列之中，藏着自然的秩序。」

——斐波那契

第七章 数列



知乎@余命数

棋盘上的米粒

古老的传说中, 国际象棋的发明者向国王提出一个请求. 他在棋盘第 1 格放 1 粒米, 第 2 格放 2 粒, 第 3 格放 4 粒, 之后每一格的米粒数都是前一格的两倍, 直到填满全部 64 格. 国王欣然应允, 却不知自己许下了一个无法兑现的承诺.

这个序列 $1, 2, 4, 8, \ldots$ 增长速度惊人. 到第 20 格, 米粒数已超过一百万; 到第 41 格, 已超过当时全世界的米粒总产量. 我们要算清 64 格总共有多少米粒, 就要刻画这串数字的生成规律, 并找到不必逐一相加便能求出总数的办法.

这个故事引出一个基本数学对象, 即**数列**, 特别是**等比数列**. 复利增长、细胞分裂等利滚利现象, 都可以用等比数列描述. 本章我们建立描述数列每一项的记法, 并学习高效计算其和的方法.

7.1 数学归纳法

在数学的探索中, 我们常常从观察具体的例子开始.

$$1 = 1^2$$

$$1 + 3 = 4 = 2^2$$

$$1 + 3 + 5 = 9 = 3^2$$

$$1 + 3 + 5 + 7 = 16 = 4^2$$

观察这一组等式. 前 n 个正奇数的和似乎总是等于 n^2. 要确定这个规律对所有正整数 n 成立, 无论是 $n = 100$ 还是 $n = 10^{10}$, 验证再多例子也无法构成严格的**证明**. 需要一种从有限验证出发、得到关于无限情况的确定性结论的方法.

这个逻辑工具就是**数学归纳法**, 它把归纳猜想与演绎证明联系起来, 是离散数学[1]的重要证明方法.

[1] 离散数学是数学的几个分支的总称, 研究基于离散空间的数学结构. 与连续变化的实数不同, 离散数学的研究对象, 例如整数、图和数学逻辑中的命题, 拥有不等、分立的值. 因此离散数学不包含微积分和分析等连续数学的内容.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

公理 7.1 (数学归纳法) 设 $P(n)$ 是一个关于正整数 n 的命题, 且起始值 n_0 为正整数. 如果以下两个条件都被满足,

- **奠基步骤:** 命题 $P(n)$ 对起始值 $n = n_0$ 成立(通常 $n_0 = 1$).
- **归纳步骤:** **假设**命题对任意整数 $k \geq n_0$ 成立, 即 $P(k)$ 为真, 能够**推导出**命题对下一个整数 $k+1$ 也成立, 即 $P(k+1)$ 为真.

那么, 可断定, 命题 $P(n)$ 对所有大于或等于 n_0 的整数 n 都成立.

数学归纳法的原理可以用一排无穷无尽的多米诺骨牌来比喻. 要让所有的骨牌都倒下, 需要满足两个条件,

1. 首先要推倒第一块骨牌, 这就是**奠基步骤** $P(n_0)$. 第一块不倒, 连锁反应就无从谈起.
2. 骨牌的排列必须保证, 任意一块(第 k 块)倒下时, 都必定会撞倒第 $k+1$ 块, 这就是**归纳步骤** $P(k) \implies P(k+1)$. 它保证了倒下这个性质可以无休止地传递下去.

只要这两个条件同时满足, 整排骨牌就将依次倒下.

例题 7.1 使用数学归纳法证明,对所有正整数 n,等式 $1+3+5+\cdots+(2n-1) = n^2$ 成立.

证明 设命题 $P(n)$ 为 $1+3+5+\cdots+(2n-1) = n^2$.

当 $n = 1$ 时, 等式左边为 1, 右边为 $1^2 = 1$, 两端相等, 故 $P(1)$ 成立, 奠基步骤完成.

假设当 $n = k$($k \geq 1$ 为任意整数)时命题成立, 即

$$1+3+5+\cdots+(2k-1) = k^2.$$

由此只需证明当 $n = k+1$ 时命题也成立, 即

$$1+3+5+\cdots+(2(k+1)-1) = (k+1)^2.$$

对目标等式的左边作恒等变形, 将其分离为前 k 个奇数之和与新增的末项, 得

$$\begin{aligned} \text{左边} &= 1+3+5+\cdots+(2k-1)+(2(k+1)-1) \\ &= \underbrace{[1+3+5+\cdots+(2k-1)]}_{\text{归纳假设部分}}+(2k+2-1) \\ &= k^2+(2k+1) \quad \text{(代入归纳假设)} \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= k^2 + 2k + 1$$

$$= (k+1)^2$$

$$= \text{右边}.$$

知乎@余命数

上式表明, 若第 k 块骨牌倒下, 它必然会推倒第 $k+1$ 块, 归纳步骤成立.

由数学归纳法原理可知, 命题 $P(n)$ 对所有正整数 n 均成立. □

知乎@余命数

归纳假设不是循环论证. 我们在归纳步骤中未直接断言结论. 我们证明的是一个**条件命题**, "**如果** $P(k)$ 成立, **那么** $P(k+1)$ 成立". 我们把 $P(k)$ 作为一个前提条件, 来证明这个传递的过程是有效的. 两步缺一不可. 只有奠基步骤, 我们只能证明命题对起始值成立, 无法推广; 只有归纳步骤, 我们只证明了"如果一块倒下, 下一块也会倒", 但如果第一块就没倒, 那么什么都不会发生. 数学归纳法同样能处理整除与不等式问题.

知乎@余命数

例题 7.2 证明当实数 $x > -1$ 时, 对所有整数 $n \geq 1$, 不等式 $(1+x)^n \geq 1+nx$ 成立.

解 设命题 $P(n)$ 为 $(1+x)^n \geq 1+nx$.

当 $n=1$ 时, 不等式左边为 $(1+x)^1 = 1+x$, 右边为 $1+1\cdot x = 1+x$, 两端相等, 故 $P(1)$ 成立.

知乎@余命数

假设当 $n=k$ $(k \geq 1)$ 时命题成立, 即

$$(1+x)^k \geq 1+kx.$$

由此只需证明当 $n=k+1$ 时命题也成立, 即 $(1+x)^{k+1} \geq 1+(k+1)x$.

由条件 $x > -1$ 知 $1+x > 0$, 故不等式两边可同乘正数 $1+x$ 而不改变方向. 将目标不等式左边写为 $(1+x)^{k+1} = (1+x)^k \cdot (1+x)$, 代入归纳假设得

$$(1+x)^{k+1} \geq (1+kx)(1+x) \quad \text{(归纳假设, 两边同乘 } 1+x>0\text{)}$$

知乎@余命数

$$= 1 + x + kx + kx^2$$

$$= 1 + (k+1)x + kx^2.$$

知乎@余命数

由于 $k \geq 1$ 且 $x^2 \geq 0$, 故 $kx^2 \geq 0$, 从而

$$1 + (k+1)x + kx^2 \geq 1 + (k+1)x.$$

由不等式的传递性将两式连接得

$$(1+x)^{k+1} \geq 1 + (k+1)x + kx^2 \geq 1 + (k+1)x,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即 $(1+x)^{k+1} \geq 1+(k+1)x$, 归纳步骤成立.

由数学归纳法原理, 命题对所有整数 $n \geq 1$ 成立.

此即**伯努利不等式**.

例题 7.3　证明对所有正整数 n, 表达式 7^n-1 能被 6 整除.

解　设命题 $P(n)$ 为"7^n-1 能被 6 整除".

当 $n=1$ 时, $7^1-1=6$ 是 6 的倍数, 故 $P(1)$ 成立.

假设当 $n=k\,(k\geq 1)$ 时命题成立, 即 7^k-1 是 6 的倍数, 可记作 $7^k-1=6m$, 其中 m 为整数, 于是 $7^k=6m+1$. 由此只需证明当 $n=k+1$ 时 $7^{k+1}-1$ 能被 6 整除.

对目标表达式变形以引入归纳假设中的 7^k, 得

$$\begin{aligned} 7^{k+1}-1 &= 7\cdot 7^k-1 \\ &= 7(6m+1)-1 \quad (\text{代入 } 7^k=6m+1) \\ &= 42m+6 \\ &= 6(7m+1). \end{aligned}$$

因 m 为整数, 故 $7m+1$ 亦为整数, 从而 $7^{k+1}-1$ 等于 6 与一整数之积, 符合能被 6 整除的定义, 归纳步骤成立.

由数学归纳法原理, 对所有正整数 n, 7^n-1 都能被 6 整除.

例题 7.4 (缺一格的棋盘)　设 n 为正整数, 从 $2^n\times 2^n$ 的棋盘中挖去任意一格, 证明剩下的部分都能用 L 形骨牌(2 × 2 的方块去掉一格)恰好铺满.

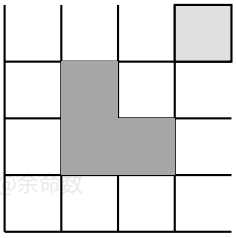


图 7.1　4×4 棋盘挖去一格(右上角的浅色格). 先在中心放一块 L 形骨牌, 使不含挖去格的三个象限各少一格, 四个 2×2 小棋盘于是都成为"挖去一格"的形状.

证明　对 n 归纳.$n=1$ 时, 2×2 的棋盘挖去一格, 剩下一格恰好是一块 L 形骨牌, 结论成立.

设结论对 n 成立, 考虑 $2^{n+1}\times 2^{n+1}$ 的棋盘挖去一格. 把它等分成四个 $2^n\times 2^n$ 的象限, 挖去的格落在其中一个象限里. 在棋盘正中心放一块 L 形骨牌, 让它的三

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

格分别落在另外三个象限里紧贴中心的那一格. 这样一来, 四个象限各自恰好少了一格, 含挖去格的象限少的是原来那格, 另外三个象限少的是被骨牌占去的那格. 四个象限都成了“$2^n \times 2^n$ 挖去一格”的形状, 由归纳假设各自都能铺满, 合起来就铺满了整个棋盘.

归纳假设能用在四个象限上, 是因为每个象限的形状与假设中的形状完全相同. 换一个尺寸或换一种骨牌, 这一步就未必走得通. □

例题 7.5 (直线分平面的区域数) 平面上 n 条直线两两不平行, 且任意三条不共点, 证明这 n 条直线把平面分成的区域数为 $\dfrac{n(n+1)}{2}+1$.

证明 记区域数为 R_n.$n=1$ 时一条直线把平面分成 2 个区域, 而 $\dfrac{1\times 2}{2}+1=2$, 公式成立.

设 n 条直线已把平面分成 R_n 个区域, 添入第 $n+1$ 条直线. 它与原有的 n 条直线各有一个交点, 由题设这些交点互不重合, 于是这条直线上共有 n 个交点, 它们把直线分成 $n+1$ 段. 每一段都落在某个已有的区域里, 并把该区域一分为二, 所以区域数增加 $n+1$, 即

$$R_{n+1}=R_n+(n+1).$$

由归纳假设 $R_n=\dfrac{n(n+1)}{2}+1$, 得

$$R_{n+1}=\frac{n(n+1)}{2}+1+(n+1)=\frac{(n+1)(n+2)}{2}+1,$$

公式对 $n+1$ 也成立. □

涉及新定义时还会再用到数学归纳法. 现先认识两个简化运算的符号.

7.2 $\sum$ 和 $\prod$ 符号

进入数列前, 有一类操作会反复出现, 即计算数列中许多项的总和或总乘积. 例如, 要表达前 100 个正整数的和, 写出 $1+2+3+\cdots+100$ 尚可接受, 但若要表达一个更抽象或更长的和, 比如 $a_1+a_2+\cdots+a_n$, 这种写法就显得冗长且不够精确. 为简洁、清晰、无歧义地表达这种运算, 数学家引入了两个实用符号, 即求和符号 ($\sum$) 和连乘符号 ($\prod$).

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 7.1 (求和符号) 一个数列 $\{a_n\}$ 从第 m 项到第 n 项的和 $a_m + a_{m+1} + \cdots + a_n$, 其中 $m \le n$ 可以用求和符号 $\sum$ (Sigma) 紧凑地写成

$$\sum_{k=m}^{n} a_k$$

其中各部分的名称如下.

- k 被称为**求和指标**. 它是一个哑变量, 仅起占位作用, 可用任何其他不引起混淆的字母代替, 如 i, j 等.
- m 是求和的**下界**, 即求和指标的起始值.
- n 是求和的**上界**, 即求和指标的终止值.
- a_k 是被求和的**通项**.

知乎@余命数

例题 7.6
- 前 100 个正整数的和可写成 $1 + 2 + \cdots + 100 = \sum_{k=1}^{100} k$.
- 前 n 个正奇数的和可写成 $1 + 3 + 5 + \cdots + (2n-1) = \sum_{k=1}^{n} (2k-1)$.
- 等差数列前 n 项和可写成 $S_n = a_1 + \cdots + a_n = \sum_{k=1}^{n} a_k$.

定义 7.2 (连乘符号) 与求和符号类似, 一个数列 $\{a_n\}$ 从第 m 项到第 n 项的积 $a_m \cdot a_{m+1} \cdot \cdots \cdot a_n$ 可以用连乘符号 $\prod$ (Pi) 表示如下.

$$\prod_{k=m}^{n} a_k$$

其各部分的名称与求和符号类似, 分别为**连乘指标**、**下界**和**上界**.

例题 7.7
- 阶乘可写成 $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n = \prod_{k=1}^{n} k$.
- 等比数列前 n 项的积可写成 $P_n = a_1 \cdot a_2 \cdot \cdots \cdot a_n = \prod_{k=1}^{n} a_k = \prod_{k=1}^{n} (a_1 q^{k-1})$.

知乎@余命数

它们继承了加法和乘法的基本定律, 如分配律、结合律和交换律.
设 $\{a_k\}, \{b_k\}$ 为数列, c 为常数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

- **常数倍数法则**.

$$\sum_{k=m}^{n} ca_k = c\sum_{k=m}^{n} a_k \quad \text{而} \quad \prod_{k=m}^{n} (ca_k) = c^{n-m+1}\prod_{k=m}^{n} a_k$$

注意连乘时, 常数 c 被乘了 $n-m+1$ 次.

- **加法/乘法法则**.

$$\sum_{k=m}^{n} (a_k \pm b_k) = \sum_{k=m}^{n} a_k \pm \sum_{k=m}^{n} b_k$$

$$\prod_{k=m}^{n} (a_k b_k) = \Big(\prod_{k=m}^{n} a_k\Big)\Big(\prod_{k=m}^{n} b_k\Big)$$

- **拆分法则**. 对于任意整数 p 满足 $m \le p < n$,

$$\sum_{k=m}^{n} a_k = \sum_{k=m}^{p} a_k + \sum_{k=p+1}^{n} a_k$$

知乎@余命数

$$\prod_{k=m}^{n} a_k = \left(\prod_{k=m}^{p} a_k\right)\left(\prod_{k=p+1}^{n} a_k\right)$$

知乎@余命数

- **指标平移**.

$$\sum_{k=m}^{n} a_k = \sum_{j=m+p}^{n+p} a_{j-p}$$

知乎@余命数

这相当于做了一个变量代换 $j = k + p$.

逐条说明这些法则.

7.2.1 运算法则

7.2.1.1 线性性质

这条性质允许把复杂的求和分解为若干个简单的部分. 该性质仅适用于**求和符号**.

知乎@余命数

$$\sum_{k=m}^{n} (c \cdot a_k \pm d \cdot b_k) = c\sum_{k=m}^{n} a_k \pm d\sum_{k=m}^{n} b_k$$

其中 c, d 是与求和指标 k 无关的常数. 这个性质可以看作两条规则的结合. 常数可提出求和号外, 各项可分别求和.

例题 7.8 利用等差数列求和公式计算 $\sum_{k=1}^{n}(4k-3)$.

知乎@余命数

解 直接将通项 $4k-3$ 看作首项为 1、末项为 $4n-3$ 的等差数列求和固然可行，此处则借助线性性质将其分解为更基本的求和.

$$\begin{aligned}\sum_{k=1}^{n}(4k-3) &= \sum_{k=1}^{n}(4k)-\sum_{k=1}^{n}3 && \text{（加减法拆分）}\\ &= 4\sum_{k=1}^{n}k-3\sum_{k=1}^{n}1 && \text{（常数因子提出）}\end{aligned}$$

代入两个基本求和公式 $\sum_{k=1}^{n}k=\frac{n(n+1)}{2}$ 与 $\sum_{k=1}^{n}1=n$，得

$$\begin{aligned}\sum_{k=1}^{n}(4k-3) &= 4\cdot\frac{n(n+1)}{2}-3\cdot n\\ &= 2n(n+1)-3n\\ &= 2n^2-n.\end{aligned}$$

这与直接用等差数列求和公式 $S_n=\frac{n(a_1+a_n)}{2}=\frac{n(1+4n-3)}{2}=n(2n-1)=2n^2-n$ 所得结果一致.

知乎@余命数

注 7.1 (线性性质不适用于乘积)
求和的线性性质**不适用于**乘积. 一般地，

$$\sum_{k=1}^{n}(a_kb_k)\neq\left(\sum_{k=1}^{n}a_k\right)\left(\sum_{k=1}^{n}b_k\right)$$

例如，$\sum_{k=1}^{2}k^2=1^2+2^2=5$，但 $(\sum_{k=1}^{2}k)^2=(1+2)^2=9$.
连乘符号没有线性性质. 一般地，

$$\prod_{k=1}^{n}(a_k+b_k)\neq\prod_{k=1}^{n}a_k+\prod_{k=1}^{n}b_k$$

@余命数

7.2.1.2 指标平移

这项技巧通过改变求和的起点和终点，来简化求和式内部的表达式，使其变为我们熟悉的形式. 这好比解方程时换元.

具体做法是，令新的指标 j 与旧的指标 k 满足一个简单线性关系，如 $j=k+c$ 或 $j=k-c$.

$$\sum_{k=m}^{n}a_k=\sum_{j=m+c}^{n+c}a_{j-c}$$

在替换时，要同时更新**下界**、**上界**和**通项**中的变量.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 7.9 计算和式 $S=\sum_{k=3}^{n+2}(k-2)^2$.

解 这个和式的通项是 $(k-2)^2$, 上下界也比较奇特. 把它变成 $\sum j^2$ 的标准形式即可直接套用平方和公式.

作换元, 令新指标 $j=k-2$, 并同时更新求和的三要素, 得

- **下界**. 当旧指标 $k=3$ 时, 新指标 $j=3-2=1$.
- **上界**. 当旧指标 $k=n+2$ 时, 新指标 $j=(n+2)-2=n$.
- **通项**.$a_k=(k-2)^2$ 变为 $a_{j+2}=j^2$.

于是原来的和式转化为

$$S=\sum_{j=1}^{n}j^2.$$

套用平方和公式(将在后续章节证明)

$$\sum_{j=1}^{n}j^2=\frac{n(n+1)(2n+1)}{6},$$

即得 $S=\frac{n(n+1)(2n+1)}{6}$.

7.2.1.3 嵌套求和

当求和的通项本身又是一个求和时, 就构成了嵌套求和.

$$\sum_{i=m}^{n}\left(\sum_{j=p}^{q}a_{i,j}\right)$$

当内层求和的上下界 (p,q) 与外层指标 (i) 无关时, 可自由地**交换求和次序**.

$$\sum_{i=m}^{n}\sum_{j=p}^{q}a_{i,j}=\sum_{j=p}^{q}\sum_{i=m}^{n}a_{i,j}$$

理解这个性质的最佳方式是想象一个 $m\times n$ 的数字矩阵[2]. 上述等式无非是说, 先按行求和再把各行之和加起来, 与先按列求和再把各列之和加起来, 结果是相等的.

[2] 矩阵即按行列排成的数表.

例题 7.10 证明 $\left(\sum_{i=1}^{m}x_i\right)\left(\sum_{j=1}^{n}y_j\right)=\sum_{i=1}^{m}\sum_{j=1}^{n}x_iy_j$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 从左边入手, 将第一个和式视作一个整体 $C=\sum_{i=1}^{m} x_i$, 则

$$\left(\sum_{i=1}^{m} x_i\right)\left(\sum_{j=1}^{n} y_j\right)=C\cdot\sum_{j=1}^{n} y_j.$$

由求和的线性性质, 常数 C 可乘入第二个和式内部, 得 $\sum_{j=1}^{n}(C\cdot y_j)$; 再把 $C=\sum_{i=1}^{m} x_i$ 代回, 有

$$\sum_{j=1}^{n}\left(\left(\sum_{i=1}^{m} x_i\right)y_j\right).$$

对于内层的和 $\sum_{i=1}^{m} x_i, y_j$ 关于指标 i 是常数, 故同样可乘入, 得到

$$\sum_{j=1}^{n}\left(\sum_{i=1}^{m} x_i y_j\right).$$

有限求和可交换次序, 因此

$$\sum_{j=1}^{n}\sum_{i=1}^{m} x_i y_j=\sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j,$$

这正是右侧的嵌套求和形式. 此恒等式在概率统计中计算期望与方差时常被引用. □

7.2.1.4 裂项求和与求积

我们后续会用专门的章节来分析裂项求和与求积. 这里的要点是用 ∑ 和 ∏ 符号写出它的结构.

- **裂项求和**. 如果通项可以表示为 $a_k=f(k)-f(k+1)$, 那么

$$\begin{aligned}\sum_{k=m}^{n} a_k &= \sum_{k=m}^{n}(f(k)-f(k+1))\\ &=(f(m)-f(m+1))+(f(m+1)-f(m+2))\\ &\quad+\cdots+(f(n)-f(n+1))\\ &=f(m)-f(n+1)\end{aligned}$$

  中间的项 $-f(m+1)$ 与 $+f(m+1)$ 等前后抵消, 如同收起的望远镜.

- **裂项求积**. 如果通项可以表示为 $a_k=\frac{f(k)}{f(k+1)}$, 那么

$$\prod_{k=m}^{n} a_k=\prod_{k=m}^{n}\frac{f(k)}{f(k+1)}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{f(m)}{f(m+1)} \cdot \frac{f(m+1)}{f(m+2)} \cdot \cdots \cdot \frac{f(n)}{f(n+1)}$$
$$= \frac{f(m)}{f(n+1)}$$

7.2.2 有趣的应用 *

例题 7.11 计算和式 $S = \sum_{k=0}^{n-1}(k+1)^3$.

知乎@余命数

解 这个和式的通项是 $(k+1)^3$, 求和指标从 0 到 $n-1$, 与熟知的立方和公式 $\sum_{j=1}^{n} j^3$ 形式接近, 故作指标平移化为标准形式.

令新指标 $j = k+1$, 求和的三要素随之更新, 得

- **下界**. 当旧指标 $k=0$ 时, 新指标 $j = 0+1 = 1$.
- **上界**. 当旧指标 $k = n-1$ 时, 新指标 $j = (n-1)+1 = n$.
- **通项**.$(k+1)^3$ 变为 j^3.

原和式转化为 $S = \sum_{j=1}^{n} j^3$, 套用立方和公式(将在后续章节证明)即得

$$S = \left(\frac{n(n+1)}{2}\right)^2 = \frac{n^2(n+1)^2}{4}.$$

例题 7.12 计算和式 $S_n = \sum_{k=1}^{n} \frac{1}{k(k+2)}$.

知乎@余命数

解 通项 $\frac{1}{k(k+2)}$ 的形式暗示了裂项求和的可能, 以待定系数法将其分解为部分分式, 即

$$\frac{1}{k(k+2)} = \frac{A}{k} + \frac{B}{k+2}.$$

通分得 $1 = A(k+2) + Bk$. 令 $k=0$ 得 $A = 1/2$, 令 $k=-2$ 得 $B = -1/2$, 故

$$\frac{1}{k(k+2)} = \frac{1}{2}\left(\frac{1}{k} - \frac{1}{k+2}\right).$$

代回原和式,

$$S_n = \sum_{k=1}^{n} \frac{1}{2}\left(\frac{1}{k} - \frac{1}{k+2}\right) = \frac{1}{2}\sum_{k=1}^{n}\left(\frac{1}{k} - \frac{1}{k+2}\right).$$

这是间隔两项的裂项求和. 写出前后各项以观察抵消规律.

$$2S_n = \left(\frac{1}{1} - \frac{1}{3}\right) + \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{5}\right) + \left(\frac{1}{4} - \frac{1}{6}\right) + \cdots$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$+\left(\frac{1}{n-1}-\frac{1}{n+1}\right)+\left(\frac{1}{n}-\frac{1}{n+2}\right).$$

从第三项起, $-\frac{1}{3}$ 与 $+\frac{1}{3}$ 抵消, $-\frac{1}{4}$ 与 $+\frac{1}{4}$ 抵消, 以此类推, 最终只留下开头的两项正数与结尾的两项负数.

$$2S_n=\frac{1}{1}+\frac{1}{2}-\frac{1}{n+1}-\frac{1}{n+2}=\frac{3}{2}-\frac{n+2+n+1}{(n+1)(n+2)}=\frac{3}{2}-\frac{2n+3}{(n+1)(n+2)}.$$

故

$$S_n=\frac{3}{4}-\frac{2n+3}{2(n+1)(n+2)}=\frac{n(3n+5)}{4(n+1)(n+2)}.$$

例题 7.13　计算二重和 $S=\sum_{i=1}^{n}\sum_{j=i}^{n}\frac{i}{j(j+1)}$.

解　内层 $\sum_{j=i}^{n}$ 的求和范围依赖外层指标 i, 直接计算内层和有困难, 故交换求和次序.

指标 (i,j) 需满足 $1\le i\le j\le n$. 当前的求和次序是固定 i、令 j 从 i 扫到 n(按行求和); 改为固定 j 再看 i 的取值范围, 由条件 $1\le i\le j\le n$ 可知 j 从 1 到 n, 固定 j 时 i 从 1 到 j. 于是得到等价形式

$$S=\sum_{j=1}^{n}\sum_{i=1}^{j}\frac{i}{j(j+1)}.$$

在内层和 $\sum_{i=1}^{j}$ 中, $\frac{1}{j(j+1)}$ 与指标 i 无关, 可提出.

$$S=\sum_{j=1}^{n}\left(\frac{1}{j(j+1)}\sum_{i=1}^{j}i\right).$$

内层为标准等差数列求和 $\sum_{i=1}^{j}i=\frac{j(j+1)}{2}$, 代回得

$$S=\sum_{j=1}^{n}\left(\frac{1}{j(j+1)}\cdot\frac{j(j+1)}{2}\right)=\sum_{j=1}^{n}\frac{1}{2}=n\cdot\frac{1}{2}=\frac{n}{2}.$$

交换求和次序后, 复杂的二重和归约为平凡的 n 个 $\frac{1}{2}$ 相加.

例题 7.14　计算连乘积 $P_n=\prod_{k=2}^{n}\frac{k^3-1}{k^3+1}$, 其中 $n\ge 2$.

解　连乘积的目标通常是裂项相消, 要点在于因式分解通项. 用立方差与立方和公式

$$a^3-b^3=(a-b)(a^2+ab+b^2),\quad a^3+b^3=(a+b)(a^2-ab+b^2)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

分解分子分母, 得

$$k^3-1=(k-1)(k^2+k+1),\quad k^3+1=(k+1)(k^2-k+1),$$

故每一项为

$$\frac{k^3-1}{k^3+1}=\frac{(k-1)(k^2+k+1)}{(k+1)(k^2-k+1)}.$$

此式直接看不出相消关系, 将其拆成两个部分的乘积.

$$P_n=\prod_{k=2}^{n}\left(\frac{k-1}{k+1}\right)\cdot\prod_{k=2}^{n}\left(\frac{k^2+k+1}{k^2-k+1}\right).$$

第一个乘积是标准的裂项求积.

$$\prod_{k=2}^{n}\frac{k-1}{k+1}=\frac{1}{3}\cdot\frac{2}{4}\cdot\frac{3}{5}\cdot\cdots\cdot\frac{n-1}{n+1}=\frac{1\cdot 2}{n(n+1)}=\frac{2}{n(n+1)}.$$

知乎@余命数

对于第二个乘积, 令 $f(k)=k^2-k+1$, 则 $f(k+1)=(k+1)^2-(k+1)+1=k^2+k+1$, 于是通项的第二部分可写为 $\frac{f(k+1)}{f(k)}$.

$$\prod_{k=2}^{n}\frac{f(k+1)}{f(k)}=\frac{f(3)}{f(2)}\cdot\frac{f(4)}{f(3)}\cdot\cdots\cdot\frac{f(n+1)}{f(n)}=\frac{f(n+1)}{f(2)}=\frac{(n+1)^2-(n+1)+1}{2^2-2+1}=\frac{n^2+n+1}{3}.$$

知乎@余命数

两部分相乘即得

$$P_n=\frac{2}{n(n+1)}\cdot\frac{n^2+n+1}{3}=\frac{2(n^2+n+1)}{3n(n+1)}.$$

例题 7.15 计算和式 $S_n=\sum_{k=1}^{n}\frac{k}{k^4+k^2+1}$.

知乎@余命

解 通项分母 k^4+k^2+1 形式特殊, 是突破口. 通过配方法和平方差公式对其因式分解, 即所谓苏菲·热尔曼恒等式[3].

知乎@余命数

$$k^4+k^2+1=(k^4+2k^2+1)-k^2=(k^2+1)^2-k^2=(k^2-k+1)(k^2+k+1).$$

知乎@余命数

注意到分母两因子之差 $(k^2+k+1)-(k^2-k+1)=2k$ 恰为分子的常数倍, 据此作裂项, 得

$$\begin{aligned}\frac{k}{k^4+k^2+1}&=\frac{k}{(k^2-k+1)(k^2+k+1)}\\&=\frac{1}{2}\cdot\frac{2k}{(k^2-k+1)(k^2+k+1)}\end{aligned}$$

[3] 索菲·热尔曼恒等式是一个常用的因式分解技巧, 其经典形式为 $a^4+4b^4=(a^2+2b^2-2ab)(a^2+2b^2+2ab)$. 其做法是通过增项再减项配成一个平方差, 例如将 a^4+4b^4 变为 $(a^4+4a^2b^2+4b^4)-4a^2b^2$. 本例中对 k^4+k^2+1 的分解正是该做法的直接应用, $(k^4+2k^2+1)-k^2$.

$$= \frac{1}{2} \cdot \frac{(k^2+k+1)-(k^2-k+1)}{(k^2-k+1)(k^2+k+1)}$$
$$= \frac{1}{2}\left(\frac{1}{k^2-k+1} - \frac{1}{k^2+k+1}\right).$$

令 $f(k) = \frac{1}{k^2-k+1}$. 因 $k^2+k+1 = (k+1)^2-(k+1)+1$, 故 $\frac{1}{k^2+k+1} = f(k+1)$, 原和式化为

$$S_n = \sum_{k=1}^{n} \frac{1}{2}\left(f(k) - f(k+1)\right) = \frac{1}{2}\sum_{k=1}^{n}(f(k) - f(k+1))$$
$$= \frac{1}{2}\left[(f(1)-f(2)) + (f(2)-f(3)) + \cdots + (f(n)-f(n+1))\right]$$
$$= \frac{1}{2}(f(1) - f(n+1)).$$

计算首末两项, $f(1) = \frac{1}{1^2-1+1} = 1$, $f(n+1) = \frac{1}{(n+1)^2-(n+1)+1} = \frac{1}{n^2+n+1}$. 代入得

$$S_n = \frac{1}{2}\left(1 - \frac{1}{n^2+n+1}\right) = \frac{1}{2} \cdot \frac{n^2+n}{n^2+n+1} = \frac{n(n+1)}{2(n^2+n+1)}.$$

例题 7.16　计算连乘积 $P_n = \prod_{k=2}^{n}\left(1 - \frac{1}{k^2}\right)$.

解　直接逐项相乘会非常复杂. 连乘计算的关键第一步通常是**因式分解**通项, 以期发现可裂项相消的结构. 通项为

$$a_k = 1 - \frac{1}{k^2} = \frac{k^2-1}{k^2} = \frac{(k-1)(k+1)}{k \cdot k},$$

代入连乘积得

$$P_n = \prod_{k=2}^{n} \frac{(k-1)(k+1)}{k \cdot k}.$$

展开乘积以观察相消规律.

$$P_n = \frac{(2-1)(2+1)}{2 \cdot 2} \cdot \frac{(3-1)(3+1)}{3 \cdot 3} \cdot \frac{(4-1)(4+1)}{4 \cdot 4} \cdot \cdots \cdot \frac{(n-1)(n+1)}{n \cdot n}$$
$$= \frac{1 \cdot 3}{2 \cdot 2} \cdot \frac{2 \cdot 4}{3 \cdot 3} \cdot \frac{3 \cdot 5}{4 \cdot 4} \cdot \cdots \cdot \frac{(n-1)(n+1)}{n \cdot n}.$$

将分子与分母分别归类.

$$P_n = \frac{1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-1)}{2 \cdot 3 \cdot 4 \cdot \cdots \cdot n} \cdot \frac{3 \cdot 4 \cdot 5 \cdot \cdots \cdot (n+1)}{2 \cdot 3 \cdot 4 \cdot \cdots \cdot n}.$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

两式都呈大量对消. 第一式

$$\frac{1\cdot 2\cdot 3\cdot\cdots\cdot(n-1)}{2\cdot 3\cdot\cdots\cdot(n-1)\cdot n}=\frac{1}{n},$$

第二式

$$\frac{3\cdot 4\cdot 5\cdot\cdots\cdot n\cdot(n+1)}{2\cdot 3\cdot 4\cdot\cdots\cdot n}=\frac{n+1}{2}.$$

故

$$P_n=\frac{1}{n}\cdot\frac{n+1}{2}=\frac{n+1}{2n}.$$

例题 7.17 计算 $P=\prod_{k=2}^{63}\log_k(k+1)$.

解 通项是对数, 底数与真数都在变化, 无法直接计算. 处理不同底数对数的乘积, 标准思路是用**对数换底公式** $\log_a b=\frac{\ln b}{\ln a}$ 把各项统一到同一底(自然对数 ln). 于是 $\log_k(k+1)=\frac{\ln(k+1)}{\ln k}$, 原乘积变为

$$P=\prod_{k=2}^{63}\frac{\ln(k+1)}{\ln k}.$$

展开后得裂项求积形式.

$$P=\frac{\ln(3)}{\ln(2)}\cdot\frac{\ln(4)}{\ln(3)}\cdot\frac{\ln(5)}{\ln(4)}\cdot\cdots\cdot\frac{\ln(63)}{\ln(62)}\cdot\frac{\ln(64)}{\ln(63)}.$$

前一项的分子恰为后一项的分母, 依次抵消.

$$P=\frac{\cancel{\ln(3)}}{\ln(2)}\cdot\frac{\cancel{\ln(4)}}{\cancel{\ln(3)}}\cdot\frac{\cancel{\ln(5)}}{\cancel{\ln(4)}}\cdot\cdots\cdot\frac{\cancel{\ln(63)}}{\cancel{\ln(62)}}\cdot\frac{\ln(64)}{\cancel{\ln(63)}}=\frac{\ln(64)}{\ln(2)}.$$

反向使用换底公式, 或直接计算得

$$P=\log_2(64)=\log_2(2^6)=6.$$

例题 7.18 计算 $P_n=\prod_{k=1}^{n}\cos\left(\frac{x}{2^k}\right)$, 其中 $\sin(x/2^n)\neq 0$.

解 这个连乘积涉及三角函数, 常规因式分解不再适用, 需寻找三角恒等式来创造相消结构. 核心恒等式是正弦二倍角公式 $\sin(2\theta)=2\sin\theta\cos\theta$, 由此得 $\cos\theta=\frac{\sin(2\theta)}{2\sin\theta}$. 此式把一个 cos 项与两个不同角度的 sin 项联系起来, 结构上正适合裂项. 乘积中并无 sin 项, 故引入一个**辅助乘子**来启动连锁反应, 最合适的乘子是与乘积中角度最小的项 $\cos(x/2^n)$ 配套的 $\sin(x/2^n)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $P_n = \cos\left(\frac{x}{2}\right)\cos\left(\frac{x}{4}\right)\ldots\cos\left(\frac{x}{2^n}\right)$, 两边同乘 $2^n\sin(x/2^n)$, 得

$$\begin{aligned}2^n\sin\left(\frac{x}{2^n}\right)P_n &= 2^n\cos\left(\frac{x}{2}\right)\ldots\cos\left(\frac{x}{2^{n-1}}\right)\left[\sin\left(\frac{x}{2^n}\right)\cos\left(\frac{x}{2^n}\right)\right]\\ &= 2^{n-1}\cdot 2\cos\left(\frac{x}{2}\right)\ldots\left[\frac{1}{2}\sin\left(2\cdot\frac{x}{2^n}\right)\right]\\ &= 2^{n-1}\cos\left(\frac{x}{2}\right)\ldots\cos\left(\frac{x}{2^{n-1}}\right)\sin\left(\frac{x}{2^{n-1}}\right).\end{aligned}$$

最小角度的 sin, cos 项合并为角度大一倍的 sin 项, 此过程持续下去.

$$\begin{aligned}\ldots &= 2^{n-2}\cos\left(\frac{x}{2}\right)\ldots\sin\left(\frac{x}{2^{n-2}}\right)\\ &\vdots\\ &= 2\cos\left(\frac{x}{2}\right)\sin\left(\frac{x}{2}\right)\\ &= \sin(x).\end{aligned}$$

故 $2^n\sin\left(\frac{x}{2^n}\right)P_n = \sin(x)$. 由 $\sin(x/2^n)\neq 0$, 解出

$$P_n = \frac{\sin x}{2^n\sin(x/2^n)}.$$

例题 7.19　计算 $P_n = \prod_{k=2}^{n}\left(1-\frac{2}{k(k+1)}\right)$, 其中 $n\geq 2$.

解　先对通项作代数变形, 得

$$a_k = 1-\frac{2}{k(k+1)} = \frac{k(k+1)-2}{k(k+1)} = \frac{k^2+k-2}{k(k+1)},$$

分解分子得 $a_k = \frac{(k+2)(k-1)}{k(k+1)}$. 将此分式**分组**, 使每一组都构成简单的裂项结构.

$$a_k = \frac{k-1}{k}\cdot\frac{k+2}{k+1}.$$

原连乘积随之拆分为两个连乘积之积.

$$P_n = \left(\prod_{k=2}^{n}\frac{k-1}{k}\right)\cdot\left(\prod_{k=2}^{n}\frac{k+2}{k+1}\right).$$

分别计算. 第一个

$$\prod_{k=2}^{n}\frac{k-1}{k} = \frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdot\ldots\cdot\frac{n-1}{n} = \frac{1}{n},$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第二个

$$\prod_{k=2}^{n}\frac{k+2}{k+1}=\frac{4}{3}\cdot\frac{5}{4}\cdot\frac{6}{5}\cdot\cdots\cdot\frac{n+2}{n+1}=\frac{n+2}{3}.$$

两部分相乘即得

$$P_n=\frac{1}{n}\cdot\frac{n+2}{3}=\frac{n+2}{3n}.$$

这些例题显示, 复杂的和式与积式经过适当变形, 往往会暴露出隐藏的结构. 其中用到的技巧按作用方式可归为几类.

最基础的操作是**变形与化归**, 它把陌生或复杂的形式通过等价变换变回熟悉的形式. **指标平移**就是一例, 令 $j=k+1$, 可把 $\sum(k+1)^3$ 写成标准立方和 $\sum j^3$. 多数例题的第一步是对通项做代数变形, 例如把 $\prod(1-1/k^2)$ 改写成 $\frac{(k-1)(k+1)}{k^2}$. 对数换底公式、三角二倍角公式、立方和/差公式等恒等式, 都能把通项改写成便于下一步处理的形式.

裂项是和式与积式求值中最常用的技巧. 它的做法是把 a_k 拆成 $f(k)-f(k+1)$ 或 $f(k)/f(k+1)$, 让中间项在累加或连乘中相互抵消.

构造裂项需要观察通项结构. 有理分式常用**部分分式分解**, 如 $\sum\frac{1}{k(k+2)}$; 复杂代数式则要利用恒等式, 例如用**索菲·热尔曼恒等式**分解 k^4+k^2+1. 有时把通项拆成多个因子的乘积, 每组各自构成简单裂项结构, 这就是**分组裂项**, 如 $\prod(1-\frac{2}{k(k+1)})$ 中的处理.

对二重和或多重和, 交换求和次序能化简内层计算. 例如 $\sum_{i=1}^{n}\sum_{j=i}^{n}$ 若按 i 先、j 后求和, 内层依赖外层; 改为先固定 j、再对 i 求和, 内层变量可能变成常数, 计算随之简化.

主动引入辅助因子可创造裂项或相约条件. 计算 $\prod\cos(x/2^k)$ 时, 乘上 $\sin(x/2^n)$ 后用二倍角公式 $\sin(2\theta)=2\sin\theta\cos\theta$ 逐级合并, 便得到简洁结果.

7.3 数列的表示

来龙去脉

在数学中, 当我们想研究一串按照特定规律排列的数时, 需一套精确的记法来描述它们. 数列, 就是这套记法的基础.

定义 7.3 (数列) 按一定次序排列的一列数称为**数列**. 数列中的每一个数都称为这个数列的**项**.
排在第一位的数称为第一项(或首项), 排在第二位的数称为第二项, ……, 排在第 n 位的数称为第 n 项.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

一个数列通常记作 $\{a_n\}$. 例如，奇数数列可以写成 $1,3,5,7,\ldots$. 其中，首项 $a_1=1$, 第二项 $a_2=3$, 等等.

要完全确定一个数列, 有两种主要的方式.

定义 7.4 (通项公式) 如果数列 $\{a_n\}$ 的第 n 项 a_n 与项数 n 之间的关系可以用一个公式 $a_n=f(n)$ 来表示, 那么这个公式就叫做这个数列的**通项公式**.

通项公式的优点是直截了当, 给定 n 就可直接算出对应的项.

定义 7.5 (递推公式) 如果已知数列 $\{a_n\}$ 的第一项, 或前几项, 且从第二项, 或某一项开始的任意一项 a_n 都可由它前面的一项, 或几项来确定, 那么这个关系式就叫做这个数列的**递推公式**.

递推公式描述相邻项之间的关系, 即从已知项推出下一项, 层层传递.

例题 7.20 对于奇数数列 $1,3,5,7,\ldots$

- **通项公式**是 $a_n=2n-1$. 例如, 第 100 项是 $a_{100}=2(100)-1=199$.
- **递推公式**是 $a_1=1, a_{n+1}=a_n+2\ (n\geq 1)$.

7.4 等差数列

来龙去脉

在所有数列中, 最简单的一种规律是每次固定增加或减少一个量. 比如, 楼房的层数、自然数的序列、储蓄罐里每天固定存入 10 元后的总额, 都是等差数列.

定义 7.6 (等差数列) 如果一个数列从第二项起, 每一项与它的前一项的差等于同一个常数, 那么这个数列就叫做**等差数列**. 这个常数叫做等差数列的**公差**, 通常用字母 d 表示.

根据定义, 公差 $d=a_{n+1}-a_n$.

- 当 $d>0$ 时, 后一项总比前一项大 $(a_{n+1}>a_n)$.
- 当 $d<0$ 时, 后一项总比前一项小 $(a_{n+1}<a_n)$.
- 当 $d=0$ 时, 数列是一个常数列, 所有项都相等 $(a_{n+1}=a_n)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

7.4.1 通项公式与前 n 项和

7.4.1.1 通项公式

需建立一个从项数 n 到项的值 a_n 的直接联系, 从而不必一步步地从 a_1 开始计算.

我们从定义 $a_{k+1}=a_k+d$ 出发.

从首项 a_1 到达第 n 项 a_n 要走 $n-1$ 步, 每步相差一个公差 d.

$$
\begin{aligned}
a_2 &= a_1+d \\
a_3 &= a_2+d=(a_1+d)+d=a_1+2d \\
a_4 &= a_3+d=(a_1+2d)+d=a_1+3d \\
&\vdots
\end{aligned}
$$

注意到, 第 n 项 a_n 是在 a_1 的基础上, 累加了 $n-1$ 个公差 d 得到的.

一种更严谨的推导方法是**累加法**. 将定义式改写为 $a_{k+1}-a_k=d$. 然后, 将从 $k=1$ 到 $k=n-1$ 的所有等式写出来.

$$
\begin{aligned}
a_2-a_1 &= d \\
a_3-a_2 &= d \\
a_4-a_3 &= d \\
&\vdots \\
a_n-a_{n-1} &= d
\end{aligned}
$$

将这 $n-1$ 个等式全部相加, 左侧的中间项会两两抵消, 例如 $-a_2$ 与 $+a_2$, 这正是**裂项相消**.

$$
(a_2-a_1)+(a_3-a_2)+\cdots+(a_n-a_{n-1})=\underbrace{d+d+\cdots+d}_{n-1\text{ 个}}
$$

化简后, 左边只剩下首末两项的差, 右边是 $n-1$ 个 d 的和.

$$
a_n-a_1=(n-1)d
$$

移项后便得到等差数列的通项公式.

$$
a_n=a_1+(n-1)d \tag{7-1}
$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

直观的归纳和累加法给出了公式的图像. 要严格证明公式 $a_n = a_1 + (n-1)d$ 对**所有**正整数 n 成立, 可以使用**数学归纳法**. 先验证 $n = 1$ 时命题成立, 再证明若 $n = k$ 时成立则 $n = k + 1$ 也成立.

设命题 $P(n)$ 为 $a_n = a_1 + (n-1)d$.

当 $n = 1$ 时, 公式左边为 a_1, 右边为 $a_1 + (1-1)d = a_1$, 两端相等, 故 $P(1)$ 成立, 奠基步骤完成.

假设当 $n = k(k \geq 1$ 且 k 为整数)时命题成立, 即 $a_k = a_1 + (k-1)d$. 由此只需证明当 $n = k + 1$ 时 $a_{k+1} = a_1 + ((k+1)-1)d = a_1 + kd$ 也成立.

由等差数列的定义 $a_{k+1} = a_k + d$ 出发, 代入归纳假设得

$$\begin{aligned} a_{k+1} &= (a_1 + (k-1)d) + d \quad \text{（代入归纳假设）} \\ &= a_1 + kd - d + d \\ &= a_1 + kd, \end{aligned}$$

与归纳目标 $a_{k+1} = a_1 + kd$ 一致, 归纳步骤成立. 故若命题对 $n = k$ 成立, 则对 $n = k + 1$ 亦成立.

由数学归纳法原理, 命题 $P(n)$ 对所有正整数 n 均成立, 通项公式 $a_n = a_1 + (n-1)d$ 得证.

7.4.1.2 前 n 项和公式的推导与视角

求 $S_n = a_1 + a_2 + \cdots + a_n$ 有多种推导途径, 下面给出三种方法.

法一采用倒序相加.

倒序相加相传由年幼的高斯发现, 要点是创造对称性.

将和 S_n 写两遍, 第二遍顺序颠倒.

$$\begin{aligned} S_n &= a_1 + a_2 + \cdots + a_{n-1} + a_n \\ S_n &= a_n + a_{n-1} + \cdots + a_2 + a_1 \end{aligned}$$

将这两式按位相加, 得到.

$$2S_n = (a_1 + a_n) + (a_2 + a_{n-1}) + \cdots + (a_{n-1} + a_2) + (a_n + a_1)$$

考察每一对的和. 对于第 k 对 $(a_k + a_{n-k+1})$:

$$\begin{aligned} a_k + a_{n-k+1} &= [a_1 + (k-1)d] + [a_1 + (n-k)d] \\ &= 2a_1 + (k - 1 + n - k)d \\ &= 2a_1 + (n-1)d \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个结果是一个与 k 无关的常数, 恰好等于首项与末项之和 a_1+a_n. 由于一共有 n 对这样的和, 得

$$2S_n = n(a_1+a_n)$$

于是得第一个求和公式, 它连接了和、项数、首项与末项.

$$S_n = \frac{n(a_1+a_n)}{2} \tag{7-2}$$

法二采用代入与分组.

此法说明求和公式中两个组成部分的来源. 将通项公式 $a_k = a_1+(k-1)d$ 直接代入和式 S_n 的每一项, 得

知乎@余命数

$$\begin{aligned} S_n &= a_1+a_2+a_3+\cdots+a_n \\ &= a_1+[a_1+d]+[a_1+2d]+\cdots+[a_1+(n-1)d] \end{aligned}$$

接着, 将式子中的 a_1 和包含 d 的部分分开重新组合, 得

$$S_n = \underbrace{(a_1+a_1+\cdots+a_1)}_{n\text{个}} + \underbrace{[0\cdot d+1\cdot d+2\cdot d+\cdots+(n-1)d]}_{\text{公差部分}}$$

第一部分显然是 na_1. 第二部分可以提出公因子 d:

$$d[1+2+\cdots+(n-1)]$$

括号内是前 $n-1$ 个正整数的和, 这是一个我们熟知的简单等差数列, 其和为 $\frac{(n-1)n}{2}$. 将两部分加起来, 我们便得到求和公式的第二种形式.

$$S_n = na_1+\frac{n(n-1)}{2}d \tag{7-3}$$

法三从 S_n 的函数结构入手.

我们在提示框中提到, S_n 是关于 n 的一个没有常数项的二次函数. 可利用这一点, 反过来推导出公式本身. 设 $S_n = An^2+Bn$. 我们的任务是确定系数 A 和 B. 我们利用数列的一个基本关系. 当 $n\geq 2$ 时, $a_n = S_n - S_{n-1}$.

$$\begin{aligned} a_n &= (An^2+Bn)-[A(n-1)^2+B(n-1)] \\ &= (An^2+Bn)-[A(n^2-2n+1)+B(n-1)] \\ &= An^2+Bn-An^2+2An-A-Bn+B \\ &= (2A)n+(B-A) \end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个结果表明，a_n 是关于 n 的一次式，这正是等差数列的特征. 比较通项公式 $a_n = dn + (a_1 - d)$，得系数的对应关系.

$$公差 d = 2A \implies A = \frac{d}{2}$$

$$a_1 = (2A)\cdot 1 + (B - A) = A + B \implies B = a_1 - A = a_1 - \frac{d}{2}$$

将求出的 A 和 B 代回 $S_n = An^2 + Bn$：

$$S_n = \frac{d}{2}n^2 + \left(a_1 - \frac{d}{2}\right)n$$

知乎@余命数

整理后即得

$$S_n = na_1 + \frac{n(n-1)}{2}d$$

注 7.2 (等差数列的函数视角)

- **通项公式的本质**.$a_n = dn + (a_1 - d)$. 故 a_n 是关于项数 n 的**一次函数**.
- **求和公式的本质**.$S_n = \frac{d}{2}n^2 + (a_1 - \frac{d}{2})n$. 故 S_n 是关于项数 n 的**二次函数**，且其**常数项为零**.
- 这一视角在判断题、选择题中能提供更宏观的判断依据.

7.5 等比数列

来龙去脉

知乎@余命数

等差数列刻画求和的规律，等比数列刻画求积的规律. 现实中翻倍、按比例增长或衰减的模型，在米粒与棋盘问题里已有体现.

定义 7.7 (等比数列) 如果一个数列从第二项起，每一项与它的前一项的比等于同一个常数，那么这个数列就叫做**等比数列**. 这个常数叫做等比数列的**公比**，通常用字母 q 表示 $(q \neq 0)$.

根据定义，公比 $q = \dfrac{a_{n+1}}{a_n}$.

7.5.1 通项公式与前 n 项和

7.5.1.1 通项公式

等差数列累加公差得到通项, 等比数列则连乘公比. 目标同样是建立项数 n 到项值 a_n 的直接关系.

法一采用注意力法.

这个方法很直观. 从首项 a_1 到第 n 项 a_n 需要连乘 $n-1$ 次公比 q.

知乎@余命数

$$
\begin{aligned}
a_2 &= a_1 \cdot q = a_1 q^1 \\
a_3 &= a_2 \cdot q = (a_1 q^1) \cdot q = a_1 q^2 \\
a_4 &= a_3 \cdot q = (a_1 q^2) \cdot q = a_1 q^3 \\
&\vdots
\end{aligned}
$$

规律是, 第 n 项 a_n 中, 公比 q 的指数比项数 n 小 1. 由此, 我们归纳出等比数列的通项公式.

知乎@余命数

$$
a_n = a_1 q^{n-1} \tag{7-4}
$$

法二采用累乘法.

这种方法与等差数列的累加法完全对称. 等比数列的定义式可写作 $\dfrac{a_{k+1}}{a_k} = q$. 将从 $k=1$ 到 $k=n-1$ 的所有等式写出来.

$$
\begin{aligned}
\frac{a_2}{a_1} &= q \\
\frac{a_3}{a_2} &= q \\
\frac{a_4}{a_3} &= q \\
&\vdots \\
\frac{a_n}{a_{n-1}} &= q
\end{aligned}
$$

知乎@余命数

将这 $n-1$ 个等式全部相乘, 左边的分子和分母依次约去, 这正是乘法中的**裂项相消**.

$$
\frac{a_2}{a_1} \cdot \frac{a_3}{a_2} \cdot \frac{a_4}{a_3} \cdot \cdots \cdot \frac{a_n}{a_{n-1}} = \underbrace{q \cdot q \cdot \cdots \cdot q}_{n-1\text{个}}
$$

化简后, 左边只剩下首末两项的比, 右边是 $n-1$ 个 q 的乘积.

$$
\frac{a_n}{a_1} = q^{n-1}
$$

两边同乘 a_1, 再次得到通项公式.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

法三采用数学归纳法.

要严格证明该公式对所有正整数 n 成立, 可以使用数学归纳法.

设命题 $P(n)$ 为 $a_n = a_1 q^{n-1}$.

当 $n = 1$ 时, 公式左边为 a_1, 右边为 $a_1 q^{1-1} = a_1 q^0 = a_1$ (约定 $q \neq 0$), 两端相等, 故 $P(1)$ 成立, 奠基步骤完成.

归纳假设 $a_k = a_1 q^{k-1}$ ($k \geq 1$ 为整数), 归纳目标是证明 $a_{k+1} = a_1 q^{(k+1)-1} = a_1 q^k$.

知乎@余命数

由等比数列定义 $a_{k+1} = a_k \cdot q$ 出发, 代入归纳假设得

$$\begin{aligned} a_{k+1} &= (a_1 q^{k-1}) \cdot q \quad \text{（代入归纳假设）} \\ &= a_1 q^{(k-1)+1} \\ &= a_1 q^k, \end{aligned}$$

与归纳目标一致, 归纳步骤成立. 故若命题对 $n = k$ 成立, 则对 $n = k + 1$ 亦成立.

由数学归纳法原理, 命题 $P(n)$ 对所有正整数 n 均成立.

知乎@余命数

7.6 差比数列

我们已经分别讨论了等差数列与等比数列. 现在把两者对应项相乘, 得到一类新的数列.

知乎@余命数

若一个数列 $\{c_n\}$ 的每一项都是某个等差数列的对应项与某个等比数列的对应项的乘积, 我们就称它为**差比数列**. 其通项同时带有等差与等比的特征, 而求和通常使用错位相减法.

定义 7.8 (差比数列) 一个数列 $\{c_n\}$ 若能表示为 $c_n = a_n \cdot b_n$ 的形式, 其中 $\{a_n\}$ 是一个首项为 a_1、公差为 d 的等差数列, $\{b_n\}$ 是一个首项为 b_1、公比为 q 的等比数列, 则称 $\{c_n\}$ 为**差比数列**. 其通项公式为

$$c_n = (a_1 + (n-1)d) \cdot b_1 q^{n-1}$$

知乎@余命数

例如, 我们在前面求和过的数列 $\{n \cdot 2^n\}$, 就是由等差数列 $\{n\}$, 其中 $a_1 = 1, d = 1$ 和等比数列 $\{2^n\}$, 其中 $b_1 = 2, q = 2$ 相乘得到的.

7.6.1 差比求和

我们来求差比数列 $\{(a_1 + (n-1)d) \cdot b_1 q^{n-1}\}$ 的前 n 项和. 标准工具是错位相减法. 其做法是**用等比数列部分的公比 q 去乘 S_n, 再与原式相减, 使相减后的新数列化简为一个等比数列的和**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们先推导一般情形. 为简化书写, 设差比数列为 $c_n=(a+(n-1)d)r^{n-1}$. 令 $S_n=\sum_{k=1}^{n}(a+(k-1)d)r^{k-1}$.

$$S_n=a+(a+d)r+(a+2d)r^2+\cdots+(a+(n-1)d)r^{n-1}$$

两边同乘以公比 r $(r\neq 1)$, 得

$$rS_n=\qquad ar+(a+d)r^2+\cdots+(a+(n-2)d)r^{n-1}+(a+(n-1)d)r^n$$

两式相减, 得

$$\begin{aligned}(1-r)S_n&=a+[(a+d)r-ar]+[(a+2d)r^2-(a+d)r^2]+\cdots\\&\quad+[(a+(n-1)d)r^{n-1}-(a+(n-2)d)r^{n-1}]-(a+(n-1)d)r^n\\(1-r)S_n&=a+[dr+dr^2+\cdots+dr^{n-1}]-(a+(n-1)d)r^n\end{aligned}$$

观察上式, 中括号内的部分是一个首项为 dr, 公比为 r 的等比数列, 共有 $n-1$ 项.

$$[dr+dr^2+\cdots+dr^{n-1}]=\frac{dr(1-r^{n-1})}{1-r}$$

代入后得到关于 S_n 的方程, 解出 S_n 即可. 这个一般公式较复杂, 我们重点理解推导过程.

注 7.3 (差比数列求和要领)
求差比数列的和时, 我们先把通项拆成等差部分与等比部分, 再用公比 q 乘和式并作**错位相减**. 相减后的结果通常呈 **首项 + 等比数列 + 尾项**结构, 从而化归成熟悉的形式.

例题 7.21 求数列 $\{(2n+1)\cdot 3^n\}$ 的前 n 项和 S_n.

解 该数列是等差数列 $\{2n+1\}$(首项 3, 公差 2)与等比数列 $\{3^n\}$ 的乘积, 后者首项为 3, 公比为 3.

写出和式 S_n.

$$S_n=3\cdot 3^1+5\cdot 3^2+7\cdot 3^3+\cdots+(2n+1)3^n$$

两边同乘以公比 3, 得

$$3S_n=\qquad 3\cdot 3^2+5\cdot 3^3+\cdots+(2n-1)3^n+(2n+1)3^{n+1}$$

上式减下式, 得

$$S_n-3S_n=3\cdot 3^1+(5-3)3^2+(7-5)3^3+\cdots$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$+((2n+1)-(2n-1))3^n-(2n+1)3^{n+1}$$

$$-2S_n=9+2\cdot3^2+2\cdot3^3+\cdots+2\cdot3^n$$
$$-(2n+1)3^{n+1}$$

$$-2S_n=9+2(3^2+3^3+\cdots+3^n)-(2n+1)3^{n+1}$$

括号内是一个首项为 $3^2=9$, 公比为 3, 项数为 $n-1$ 的等比数列.

$$3^2+3^3+\cdots+3^n=\frac{9(3^{n-1}-1)}{3-1}=\frac{9}{2}(3^{n-1}-1)=\frac{1}{2}(3^{n+1}-9)$$

代回原式, 得

$$-2S_n=9+2\cdot\frac{1}{2}(3^{n+1}-9)-(2n+1)3^{n+1}$$
$$-2S_n=9+3^{n+1}-9-(2n+1)3^{n+1}$$
$$-2S_n=(1-(2n+1))3^{n+1}$$
$$-2S_n=-2n\cdot3^{n+1}$$

两边同除以 -2, 最终得到

$$S_n=n\cdot3^{n+1}$$

7.6.2 苹果公式

错位相减法需要构造辅助等式. 另一种思路是用待定系数法把差比数列的通项 c_k 拆成 $F(k)-F(k-1)$, 通过裂项相消求和.

把这种做法推广, 可以得到适用于一般差比数列的求和公式. 该公式直接给出待定系数法的结果, 代入参数即可计算.

定理 7.1 (差比数列求和公式) 对于一个通项为 $c_k=(ak+b)q^{k-1}$ 的差比数列, 其中 a,b,q 为常数, $a\neq0,q\neq1$, 其前 n 项和为

$$S_n=\sum_{k=1}^{n}(ak+b)q^{k-1}=(An+B)q^n-B$$

其中系数 A 与 B 由下式确定

$$A=\frac{a}{q-1},\quad B=\frac{b-A}{q-1}=\frac{b(q-1)-a}{(q-1)^2}=\frac{bq-b-a}{(q-1)^2}$$

证明 我们用待定系数裂项法来证明. 寻找一个函数 $F(k)$, 使差比数列的通项 $c_k=(ak+b)q^{k-1}$ 等于其差分 $F(k)-F(k-1)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

通项 c_k 是线性多项式 $(ak+b)$ 与指数项 q^{k-1} 的乘积, 因此反差分函数 $F(k)$ 也应为多项式乘以指数项的形式. 设

$$F(k)=(Pk+Q)q^k,$$

其中 P,Q 为待定常数. 计算其差分并令其与 c_k 相等, 得

$$\begin{aligned}F(k)-F(k-1)&=(Pk+Q)q^k-[P(k-1)+Q]q^{k-1}\\&=q(Pk+Q)q^{k-1}-[P(k-1)+Q]q^{k-1}\\&=[q(Pk+Q)-(Pk-P+Q)]q^{k-1}\\&=[(Pq-P)k+(P+Qq-Q)]q^{k-1}\\&=[P(q-1)k+P+Q(q-1)]q^{k-1}.\end{aligned}$$

要求此式恒等于 $c_k=(ak+b)q^{k-1}$.

$$[P(q-1)k+P+Q(q-1)]q^{k-1}=(ak+b)q^{k-1}.$$

比较两端 q^{k-1} 的系数多项式中 k 的同次幂系数, 得到关于 P,Q 的方程组.

- k 的一次项系数为 $P(q-1)=a \implies P=\dfrac{a}{q-1}$.
- 常数项为 $P+Q(q-1)=b \implies Q(q-1)=b-P \implies Q=\dfrac{b-P}{q-1}$.

这正是定理中系数 A 和 B 的定义, 即 $A=P,B=Q$.

通项 c_k 已分解为 $F(k)-F(k-1)$ 的形式, 求和随之成为标准的裂项相消.

$$\begin{aligned}S_n=\sum_{k=1}^{n}c_k&=\sum_{k=1}^{n}[F(k)-F(k-1)]\\&=(F(1)-F(0))+(F(2)-F(1))+\cdots+(F(n)-F(n-1))\\&=F(n)-F(0).\end{aligned}$$

将 $F(k)=(Ak+B)q^k$ 代入首末两项. 于是 $F(n)=(An+B)q^n$, $F(0)=(A\cdot0+B)q^0=B$. 故

$$S_n=(An+B)q^n-B.$$

证毕. □

证明 (Abel 变换证法) 考虑和式 $S_n=\sum_{k=1}^{n}(ak+b)q^{k-1}$. 阿贝尔变换把求和项 c_k 写成 u_kv_k, 让其中一部分易于求和, 另一部分易于求差.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

阿贝尔变换公式为

$$\sum_{k=1}^{n} u_k v_k = U_n v_n - \sum_{k=1}^{n-1} U_k (v_{k+1} - v_k), \text{ 其中} U_k = \sum_{i=1}^{k} u_i$$

对通项 $c_k = (ak+b)q^{k-1}$, 可取

- 令 $v_k = ak + b$, 这是等差数列项, 其差分 $v_{k+1} - v_k$ 为常数, 能简化新的求和.
- 令 $u_k = q^{k-1}$, 这是等比数列项, 其前 k 项和 U_k 有简洁的封闭形式.

接着, 计算阿贝尔变换所需的各个构件, 得

- 差分项 $v_{k+1} - v_k = [a(k+1)+b] - [ak+b] = a$.
- 部分和 $U_k = \sum_{i=1}^{k} u_i = \sum_{i=1}^{k} q^{i-1} = \frac{q^k-1}{q-1}$.
- 最终项 $v_n = an + b$ 以及 $U_n = \frac{q^n-1}{q-1}$.

将这些构件代入阿贝尔变换公式, 得

$$\begin{aligned} S_n &= U_n v_n - \sum_{k=1}^{n-1} U_k (v_{k+1} - v_k) \\ &= \left(\frac{q^n-1}{q-1}\right)(an+b) - \sum_{k=1}^{n-1} \left(\frac{q^k-1}{q-1}\right) a \\ &= \frac{(an+b)(q^n-1)}{q-1} - \frac{a}{q-1} \sum_{k=1}^{n-1} (q^k - 1) \end{aligned}$$

接下来计算新的求和部分, 得

$$\begin{aligned} \sum_{k=1}^{n-1} (q^k - 1) &= \left(\sum_{k=1}^{n-1} q^k\right) - \left(\sum_{k=1}^{n-1} 1\right) \\ &= \frac{q(q^{n-1}-1)}{q-1} - (n-1) \end{aligned}$$

将此结果代回 S_n 的表达式并整理, 目标是分离出含 q^n 的项与常数项.

$$\begin{aligned} S_n &= \frac{(an+b)q^n - (an+b)}{q-1} - \frac{a}{q-1}\left[\frac{q^n-q}{q-1} - (n-1)\right] \\ &= \frac{(an+b)q^n}{q-1} - \frac{an+b}{q-1} - \frac{a(q^n-q)}{(q-1)^2} + \frac{a(n-1)}{q-1} \end{aligned}$$

首先, 合并所有包含 q^n 的项, 得

$$\begin{aligned} q^n \left[\frac{an+b}{q-1} - \frac{a}{(q-1)^2}\right] &= q^n \left[\frac{(an+b)(q-1)-a}{(q-1)^2}\right] \\ &= q^n \left[\frac{an(q-1)+b(q-1)-a}{(q-1)^2}\right] \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= q^n \left[\frac{a(q-1)}{(q-1)^2} n + \frac{bq-b-a}{(q-1)^2} \right]$$
$$= \left(\frac{a}{q-1} n + \frac{bq-b-a}{(q-1)^2} \right) q^n$$

这恰好是 $(An+B)q^n$ 的形式, 其中 $A = \dfrac{a}{q-1}$ 且 $B = \dfrac{bq-b-a}{(q-1)^2}$.

接着, 合并所有不含 q^n 的项, 即常数部分.

$$\begin{aligned}
&-\frac{an+b}{q-1} - \frac{-aq}{(q-1)^2} + \frac{a(n-1)}{q-1} \\
&= \frac{-(an+b)+a(n-1)}{q-1} + \frac{aq}{(q-1)^2} \\
&= \frac{-(an+b)(q-1)+a(n-1)(q-1)+aq}{(q-1)^2} \\
&= \frac{-an(q-1)-b(q-1)+an(q-1)-a(q-1)+aq}{(q-1)^2} \\
&= \frac{-b(q-1)-a(q-1)+aq}{(q-1)^2} \\
&= \frac{-bq+b-aq+a+aq}{(q-1)^2} = \frac{-bq+b+a}{(q-1)^2} \\
&= -\left(\frac{bq-b-a}{(q-1)^2} \right) = -B
\end{aligned}$$

将两部分结果组合起来, 最终得到

$$S_n = (An+B)q^n - B$$

□

证明 (Gosper 裂项法证法) Gosper 算法系统地寻找给定通项 c_k 的反差分函数 $F(k)$, 使得 $c_k = F(k) - F(k-1)$. 一旦找到 $F(k)$, 求和便可通过裂项相消 $S_n = F(n) - F(0)$ 直接得到.

通项 $c_k = (ak+b)q^{k-1}$ 是典型的超几何项(其相邻项之比为有理函数).Gosper 算法断言, 其反差分函数 $F(k)$ 若存在, 则必具相似结构, 即一个多项式乘以一个指数项.

求和公式的结果含 $(An+B)q^n$, 因此我们猜测反差分函数 $F(k)$ 具 $F(k) = (Pk+Q)q^k$ 的形式, 其中 P, Q 为待定常数. 再求解差分方程定出 P 与 Q.

设 $F(k) = (Pk+Q)q^k$ 并计算其差分, 得

$$\begin{aligned}
F(k) - F(k-1) &= (Pk+Q)q^k - [P(k-1)+Q]q^{k-1} \\
&= q \cdot (Pk+Q)q^{k-1} - [Pk-P+Q]q^{k-1}
\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
&= [q(Pk+Q) - (Pk - P + Q)]q^{k-1} \\
&= [(Pq - P)k + (P + Qq - Q)]q^{k-1} \\
&= [P(q-1)k + P + Q(q-1)]q^{k-1}
\end{aligned}
$$

要求这个差分结果恒等于原数列的通项 $c_k = (ak+b)q^{k-1}$.

$$[P(q-1)k + P + Q(q-1)]q^{k-1} = (ak+b)q^{k-1}$$

这是一个关于 k 的多项式恒等式. 通过比较等式两边 k 的同次幂的系数, 可建立关于 P, Q 的线性方程组.

- 比较 k^1 的系数, 得

$$P(q-1) = a \implies P = \frac{a}{q-1}$$

- 比较 k^0(常数项)的系数, 得

$$P + Q(q-1) = b \implies Q(q-1) = b - P \implies Q = \frac{b-P}{q-1}$$

将求得的 P 代入 Q 的表达式中, 得

$$Q = \frac{b - \frac{a}{q-1}}{q-1} = \frac{\frac{b(q-1)-a}{q-1}}{q-1} = \frac{b(q-1)-a}{(q-1)^2} = \frac{bq-b-a}{(q-1)^2}$$

因此, 用这个系统的方法求出的常数 P, Q 与定理中给出的系数 A, B 完全一致, 即 $A = P, B = Q$.

我们已找到反差分函数 $F(k) = (Ak+B)q^k$. 根据离散微积分基本定理, 和式 S_n 等于 $F(n)$ 与 $F(0)$ 之差.

$$S_n = \sum_{k=1}^{n} c_k = F(n) - F(0)$$

计算 $F(n)$ 和 $F(0)$, 得

- $F(n) = (An+B)q^n$
- $F(0) = (A \cdot 0 + B)q^0 = B$

因此, 得

$$S_n = (An+B)q^n - B$$

□

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.22 求数列 $\{(2k-1)\cdot 3^{k-1}\}$ 的前 n 项和 S_n.

解 先识别数列的类型. 其通项 $c_k=(2k-1)\cdot 3^{k-1}$ 是等差数列项 $(2k-1)$ 与等比数列项 3^{k-1} 的乘积, 因此是一个标准的差比数列.

我们可以直接套用差比数列求和公式. 先将通项与标准形式 $(ak+b)q^{k-1}$ 比对, 确定参数 a,b,q.

- 线性部分 $2k-1$ 对应 $ak+b$, 可得 $a=2,b=-1$.
- 指数部分 3^{k-1} 对应 q^{k-1}, 可得 $q=3$.

接着, 计算公式所需的中间系数 A 和 B.

知乎@余命数

$$A=\frac{a}{q-1}=\frac{2}{3-1}=1$$

知乎@余命数

$$B=\frac{bq-b-a}{(q-1)^2}=\frac{(-1)(3)-(-1)-2}{(3-1)^2}=\frac{-3+1-2}{4}=\frac{-4}{4}=-1$$

知乎@余命数

最后, 将求得的 A,B 以及参数 q 代入求和公式 $S_n=(An+B)q^n-B$, 得

$$\begin{aligned}S_n&=(1\cdot n+(-1))\cdot 3^n-(-1)\\&=(n-1)3^n+1\end{aligned}$$

因此, 该数列的前 n 项和为 $S_n=(n-1)3^n+1$.

知乎@余命数

验证. 为检验结果, 我们计算前两项的和. $S_2=c_1+c_2=(2\cdot 1-1)3^0+(2\cdot 2-1)3^1=1+3\cdot 3=10$. 将 $n=2$ 代入推导的公式, 得 $S_2=(2-1)3^2+1=1\cdot 9+1=10$. 结果吻合.

例题 7.23 求数列 $\{k\cdot 2^k\}$ 的前 n 项和 S_n.

知乎@余命数

解 该数列的通项 $c_k=k\cdot 2^k$ 也是差比数列, 但指数 k 与标准公式 $(ak+b)q^{k-1}$ 不一致. 套用公式前要先做代数变形.

目标是把 2^k 化为含 q^{k-1} 的形式.

$$2^k=2\cdot 2^{k-1}$$

将此代回通项表达式, 得

$$c_k=k\cdot(2\cdot 2^{k-1})=(2k)\cdot 2^{k-1}$$

经过变形, 通项现在完全符合 $(ak+b)q^{k-1}$ 的结构. 进行参数比对.

- 线性部分 $2k$ 对应 $ak+b$, 可得 $a=2,b=0$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 指数部分 2^{k-1} 对应 q^{k-1}, 可得 $q=2$.

  接着, 计算系数 A, B.

$$A=\frac{a}{q-1}=\frac{2}{2-1}=2$$

$$B=\frac{bq-b-a}{(q-1)^2}=\frac{(0)(2)-0-2}{(2-1)^2}=\frac{-2}{1}=-2$$

  将 A, B, q 代入求和公式 $S_n=(An+B)q^n-B$, 得

$$\begin{aligned}S_n&=(2n-2)\cdot 2^n-(-2)\\&=2(n-1)\cdot 2^n+2\\&=(n-1)\cdot 2^{n+1}+2\end{aligned}$$

  因此, 该数列的前 n 项和为 $S_n=(n-1)2^{n+1}+2$.

例题 7.24　求和式 $S_n=\sum_{k=1}^{n}\frac{k+1}{3^k}$ 的值.

解　该和式的通项 $c_k=\dfrac{k+1}{3^k}$ 要先化成差比数列形式. 利用求和的线性性质, 把通项拆成熟悉的部分之和.

将通项 c_k 拆分.

$$c_k=\frac{k}{3^k}+\frac{1}{3^k}$$

因此, 原求和式可分解为两个独立和式之和.

$$S_n=\sum_{k=1}^{n}\frac{k}{3^k}+\sum_{k=1}^{n}\frac{1}{3^k}$$

分别计算这两个和式.

令 $T_n=\sum_{k=1}^{n}\frac{k}{3^k}$. 其通项为 $d_k=k\cdot\left(\frac{1}{3}\right)^k$. 这是一个差比数列. 为了应用公式, 将其变形为标准形式 $(ak+b)q^{k-1}$.

$$d_k=k\cdot\frac{1}{3}\cdot\left(\frac{1}{3}\right)^{k-1}=\left(\frac{1}{3}k\right)\cdot\left(\frac{1}{3}\right)^{k-1}$$

与标准形式比对, 得参数 $a=\frac{1}{3}, b=0, q=\frac{1}{3}$. 计算系数 A, B, 得

$$A=\frac{a}{q-1}=\frac{1/3}{1/3-1}=\frac{1/3}{-2/3}=-\frac{1}{2}$$

知乎@余命数　知乎@余命数

$$B=\frac{bq-b-a}{(q-1)^2}=\frac{0-0-1/3}{(1/3-1)^2}=\frac{-1/3}{(-2/3)^2}=\frac{-1/3}{4/9}=-\frac{3}{4}$$

代入公式 $T_n=(An+B)q^n-B$, 得

$$T_n=\left(-\frac{1}{2}n-\frac{3}{4}\right)\left(\frac{1}{3}\right)^n-\left(-\frac{3}{4}\right)=\frac{3}{4}-\left(\frac{n}{2}+\frac{3}{4}\right)\frac{1}{3^n}$$

知乎@余命数

令 $G_n=\sum_{k=1}^{n}\frac{1}{3^k}=\sum_{k=1}^{n}\left(\frac{1}{3}\right)^k$. 这是一个首项为 $1/3$, 公比为 $1/3$ 的等比数列.

知乎@余命数

$$G_n=\frac{\frac{1}{3}\left(1-(\frac{1}{3})^n\right)}{1-\frac{1}{3}}=\frac{\frac{1}{3}\left(1-\frac{1}{3^n}\right)}{2/3}=\frac{1}{2}\left(1-\frac{1}{3^n}\right)=\frac{1}{2}-\frac{1}{2\cdot 3^n}$$

知乎@余命数

将两部分相加.

$$\begin{aligned}S_n&=T_n+G_n\\&=\left[\frac{3}{4}-\left(\frac{n}{2}+\frac{3}{4}\right)\frac{1}{3^n}\right]+\left[\frac{1}{2}-\frac{1}{2\cdot 3^n}\right]\\&=\frac{3}{4}+\frac{1}{2}-\left(\frac{n}{2}+\frac{3}{4}+\frac{1}{2}\right)\frac{1}{3^n}\\&=\frac{5}{4}-\left(\frac{n}{2}+\frac{5}{4}\right)\frac{1}{3^n}\\&=\frac{5}{4}-\frac{2n+5}{4\cdot 3^n}\end{aligned}$$

知乎@余命数

例题 7.25 已知数列 $\{c_n\}$ 的前 n 项和 $S_n=(n-1)2^n+1$.

1. 求数列 $\{c_n\}$ 的通项公式.
2. 求数列 $\{(-1)^{k+1}c_k\}$ 的前 n 项和 T_n.

知乎@余命数

解 (1) 这是一个由前 n 项和 S_n 反求通项 c_n 的问题. 利用基本关系式 $c_n=S_n-S_{n-1}$, 当 $n\geq 2$ 时.

知乎@余命数

$$S_{n-1}=((n-1)-1)2^{n-1}+1=(n-2)2^{n-1}+1$$

对于 $n\geq 2$:

知乎@余命数

$$\begin{aligned}c_n&=S_n-S_{n-1}\\&=[(n-1)2^n+1]-[(n-2)2^{n-1}+1]\\&=(n-1)\cdot 2\cdot 2^{n-1}-(n-2)2^{n-1}\\&=[2(n-1)-(n-2)]2^{n-1}\\&=(2n-2-n+2)2^{n-1}\end{aligned}$$

$$= n \cdot 2^{n-1}$$

需检验此公式在 $n=1$ 时是否成立. 当 $n=1$ 时, $c_1 = S_1 = (1-1)2^1 + 1 = 1$. 将 $n=1$ 代入推导的公式, $c_1 = 1 \cdot 2^{1-1} = 1 \cdot 2^0 = 1$. 公式对 $n=1$ 也成立. 因此, 数列 $\{c_n\}$ 的通项公式为 $c_n = n \cdot 2^{n-1}$.

(2) 需求和的数列是 $\{d_k\}$, 其通项为 $d_k = (-1)^{k+1}c_k = (-1)^{k+1}k \cdot 2^{k-1}$. 这个通项看似复杂, 仍是差比数列. 重组后即可匹配标准形式.

$$\begin{aligned} d_k &= k \cdot [(-1)^{k+1}2^{k-1}] \\ &= k \cdot [(-1)^2(-1)^{k-1}2^{k-1}] \\ &= k \cdot [(-2)^{k-1}] \end{aligned}$$

现在通项 $d_k = k \cdot (-2)^{k-1}$ 是一个标准形式的差比数列. 与 $(ak+b)q^{k-1}$ 比对, 得参数 $a=1, b=0, q=-2$.

计算系数 A, B, 得

$$A = \frac{a}{q-1} = \frac{1}{-2-1} = -\frac{1}{3}$$

$$B = \frac{bq-b-a}{(q-1)^2} = \frac{0-0-1}{(-2-1)^2} = -\frac{1}{9}$$

代入求和公式 $T_n = (An+B)q^n - B$:

$$\begin{aligned} T_n &= \left(-\frac{1}{3}n - \frac{1}{9}\right)(-2)^n - \left(-\frac{1}{9}\right) \\ &= -\frac{3n+1}{9}(-2)^n + \frac{1}{9} \\ &= \frac{1-(3n+1)(-2)^n}{9} \end{aligned}$$

此即为所求的前 n 项和.

例题 7.26 (当公比为 1 时) 一个数列的通项为 $c_k = (4k-3) \cdot 1^k$. 求其前 n 项和 S_n.

解 初看此题, 通项 $c_k = (4k-3) \cdot 1^k$ 像是线性部分 $(4k-3)$ 与指数部分 1^k 的乘积, 容易让人套用差比数列求和公式.

作参数比对, 先将通项变形为标准形式 $(ak+b)q^{k-1}$. 于是

$$c_k = (4k-3) \cdot 1^k = (4k-3) \cdot 1 \cdot 1^{k-1},$$

可得 $a=4, b=-3, q=1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

随之计算系数 A, 得

$$A=\frac{a}{q-1}=\frac{4}{1-1}=\frac{4}{0},$$

知乎@余命数

此时出现除以零的错误. 差比数列求和公式在此失效, 因为公式推导(无论是错位相减法还是待定系数法)都隐含 $q\neq 1$ 的前提. 当 $q=1$ 时, 分母 $(q-1)$ 为零, 公式不再成立.

知乎@余命数

须回到数列本身分析. 当公比 $q=1$ 时, 指数部分 $q^{k-1}=1^{k-1}=1$ 对所有 k 恒成立, 等比部分退化为常数数列 $\{1,1,1,\dots\}$, 因此原数列的通项 c_k 发生**退化**.

$$c_k=(4k-3)\cdot 1^{k-1}=4k-3,$$

实质上是一个早已熟悉的**等差数列**.

问题随之化归为求解首项 $c_1=4(1)-3=1$、公差 $d=4$ 的等差数列前 n 项和. 直接使用等差数列求和公式

知乎@余命数

$$S_n=\frac{n(c_1+c_n)}{2},$$

知乎@余命数

其中末项 $c_n=4n-3$, 得

$$S_n=\frac{n(1+(4n-3))}{2}=\frac{n(4n-2)}{2}=n(2n-1)=2n^2-n.$$

知乎@余命数 知乎@余命数

注 7.4 (启示)
公式都有适用范围. 差比数列求和公式要求 $q\neq 1$. 当 $q=1$ 时, 差比数列退化为等差数列. 解题时应先确定数列类型, 再选择对应公式.

例题 7.27 (综合应用) 已知数列 $\{c_n\}$ 满足关系式 $\sum_{k=1}^{n}\frac{c_k}{3^k}=2n^2+n$.

1. 求数列 $\{c_n\}$ 的通项公式.
2. 设 $d_n=\dfrac{(-1)^n c_n}{3^{n-1}}$, 求数列 $\{d_n\}$ 的前 n 项和 S_n.

知乎@余命数 知乎@余命数

解 (1) 题目给出变换后的数列 $u_n=c_n/3^n$ 的前 n 项和, 要求反求原数列 c_n 的通项.

令新数列 $\{u_n\}$ 的通项为 $u_n=\dfrac{c_n}{3^n}$. 那么, 题目给出的条件即为数列 $\{u_n\}$ 的前 n 项和 $T_n=\sum_{k=1}^{n}u_k=2n^2+n$.

知乎@余命数

可利用基本关系式 $u_n=T_n-T_{n-1}$, 当 $n\geq 2$ 时来求解 u_n.

$$T_{n-1}=2(n-1)^2+(n-1)=2(n^2-2n+1)+n-1=2n^2-3n+1$$

对于 $n \geq 2$:

$$\begin{aligned} u_n &= T_n - T_{n-1} \\ &= (2n^2 + n) - (2n^2 - 3n + 1) \\ &= 4n - 1 \end{aligned}$$

知乎@余命数

接下来,检验此公式在 $n = 1$ 时是否成立. 当 $n = 1$ 时,$u_1 = T_1 = 2(1)^2 + 1 = 3$. 将 $n = 1$ 代入已得的公式, $u_1 = 4(1) - 1 = 3$. 公式对 $n = 1$ 也成立. 因此, 数列 $\{u_n\}$ 的通项公式为 $u_n = 4n - 1$.

最后, 从 $u_n = \dfrac{c_n}{3^n}$ 反解出 c_n. 于是

$$c_n = u_n \cdot 3^n = (4n - 1)3^n$$

(2) 需求和的数列是 $\{d_n\}$, 其通项 $d_k = \dfrac{(-1)^k c_k}{3^{k-1}}$. 首先, 将第一问求得的 $c_k = (4k - 1)3^k$ 代入, 得

知乎@余命数

$$d_k = \frac{(-1)^k (4k - 1)3^k}{3^{k-1}} = (-1)^k (4k - 1) \cdot 3$$

这个通项实质上是差比数列, 但形式需变形后才与标准公式 $(ak + b)q^{k-1}$ 匹配. 关键是处理交错项 $(-1)^k$.

$$d_k = 3 \cdot (-1) \cdot (-1)^{k-1} \cdot (4k - 1) = (-3)(4k - 1) \cdot (-1)^{k-1}$$

知乎@余命数

$$d_k = (-12k + 3) \cdot (-1)^{k-1}$$

现在通项完全符合标准形式. 进行参数比对.

- 线性部分 $-12k + 3$ 对应 $ak + b$, 可得 $a = -12, b = 3$.
- 指数部分 $(-1)^{k-1}$ 对应 q^{k-1}, 可得 $q = -1$.

接着, 计算系数 A, B, 得

$$A = \frac{a}{q - 1} = \frac{-12}{-1 - 1} = \frac{-12}{-2} = 6$$

$$B = \frac{bq - b - a}{(q - 1)^2} = \frac{(3)(-1) - 3 - (-12)}{(-1 - 1)^2} = \frac{-3 - 3 + 12}{(-2)^2} = \frac{6}{4} = \frac{3}{2}$$

知乎@余命数

将 A, B, q 代入求和公式 $S_n = (An + B)q^n - B$, 得

$$\begin{aligned} S_n &= \left(6n + \frac{3}{2}\right)(-1)^n - \frac{3}{2} \\ &= \frac{12n + 3}{2}(-1)^n - \frac{3}{2} \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{(12n+3)(-1)^n - 3}{2}$$

注 7.5 (启示)
本题综合了数列的多个知识点.

1. 第一问引入辅助数列 $\{u_n\}$, 将复杂关系式**化归**为由和求通项的问题.
2. 第二问不能直接把 d_k 的原始形式代入公式, 而要先做代数恒等变形, 严格化为 $(ak+b)q^{k-1}$ 的标准结构. 处理交错项 $(-1)^k = -(-1)^{k-1}$, 是识别公比 $q=-1$ 的关键.
3. 推导完成后, 一旦问题化为标准形式, 代入公式即可得到结果, 比错位相减法更直接.

7.6.3 裂项相消法

错位相减法通过构造把差比数列求和化归为等比数列求和. 另一种思路是把通项 c_n 直接拆成 $F(n)-F(n-1)$, 像处理 $\frac{1}{n(n+1)}$ 那样裂项相消.

这需要从观察规律转向主动构造. 此法计算上可能更繁琐, 但它说明求和与差分互为逆运算, 也是差分方程的基础[4].

[4] 后面讨论概率相关题目时, 还会回到这里.

7.6.3.1 待定系数法

我们寻找一个函数 $F(n)$, 使得 $c_n = F(n) - F(n-1)$. 差比数列的通项 $c_n = (a_1+(n-1)d)b_1q^{n-1}$ 是 n 的一次多项式与指数函数的乘积. 由此猜测, 函数 $F(n)$ 也应为多项式乘以指数函数的形式.

设差比数列通项为 $c_n = (\alpha n+\beta)q^n$, 当 $q \neq 0$ 时, 常见的差比数列通项都可以化为此形式. 设待定函数 $F(n)$ 具有 $F(n) = (Pn+Q)q^n$ 的形式, 其中 P 和 Q 是待定常数.

接着, 计算 $F(n)-F(n-1)$, 得

$$\begin{aligned} F(n)-F(n-1) &= (Pn+Q)q^n - [P(n-1)+Q]q^{n-1} \\ &= q(Pn+Q)q^{n-1} - [Pn-P+Q]q^{n-1} \\ &= [q(Pn+Q) - (Pn-P+Q)]q^{n-1} \\ &= [Pqn+Qq-Pn+P-Q]q^{n-1} \\ &= [n(Pq-P) + (P+Qq-Q)]q^{n-1} \end{aligned}$$

要使 $c_n = F(n)-F(n-1)$, 只需将上式与 $c_n = (\alpha n+\beta)q\cdot q^{n-1} = [(\alpha q)n+\beta q]q^{n-1}$ 作比较, 令对应项系数相等, 即可解出 P 和 Q.

例题 7.28 再次求解数列 $\{(2n+1)\cdot 3^n\}$ 的前 n 项和 S_n.

解

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

设 $F(n)=(Pn+Q)3^n$.

希望 $c_n=F(n)-F(n-1)$.

$$\begin{aligned}F(n)-F(n-1)&=(Pn+Q)3^n-(P(n-1)+Q)3^{n-1}\\&=3(Pn+Q)3^{n-1}-(Pn-P+Q)3^{n-1}\\&=[3Pn+3Q-Pn+P-Q]3^{n-1}\\&=[2Pn+(P+2Q)]3^{n-1}\end{aligned}$$

为了与 $c_n=(2n+1)3^n$ 比较, 将 c_n 也变形.

$$c_n=(2n+1)3^n=3(2n+1)3^{n-1}=(6n+3)3^{n-1}$$

比较 $[2Pn+(P+2Q)]3^{n-1}$ 与 $(6n+3)3^{n-1}$ 的系数, 得

- n 的系数为 $2P=6\implies P=3$
- 常数项系数为 $P+2Q=3\implies 3+2Q=3\implies Q=0$

解得 $Q=0$, 所以待定函数是 $F(n)=3n\cdot 3^n=n\cdot 3^{n+1}$.

$$S_n=\sum_{k=1}^{n}c_k=F(n)-F(0)$$

计算 $F(n)$ 和 $F(0)$, 得

$$F(n)=n\cdot 3^{n+1}$$

$$F(0)=0\cdot 3^{0+1}=0$$

因此,

$$S_n=n\cdot 3^{n+1}-0=n\cdot 3^{n+1}$$

我们得到与错位相减法相同的结果, 但路径不同. 这一方法把依赖技巧的问题变成按固定步骤做代数运算的问题.

7.7 斐波那契数列

在 1202 年, 意大利数学家斐波那契在他的著作《算盘书》中提出了一个看似简单的问题.

假设一对兔子每个月可以生一对新的兔子, 而新生的兔子又需要一个月的时间才能成熟并开始生育. 如果从一对新生的兔子开始, 那么一年后会有多少对兔子?

这个问题的答案, $1,1,2,3,5,8,13,\ldots$, 催生了数学中最著名、最迷人的数列.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

定义 7.9 (斐波那契数列) 斐波那契数列 $\{F_n\}$ 由以下线性递推关系定义

$$F_n = F_{n-1} + F_{n-2} \quad (n \geq 2)$$

其初始项通常定义为 $F_0 = 0, F_1 = 1$. 数列的前几项为

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

7.7.1 比内公式

递推公式计算效率低. 要从 F_0, F_1 一路递推到 F_{100} 很繁琐. 与等差、等比数列一样, 目标是建立直接从 n 计算 F_n 的**通项公式**.

这是一种二阶线性递推关系, 可以使用**特征方程法**求解. 该方法先猜测解的形式. 我们大胆假设数列的通项具有等比数列的形式, 即 $F_n = x^n$, 其中 $x \neq 0$. 将此形式代入递推关系 $F_n - F_{n-1} - F_{n-2} = 0$ 中, 得

知乎@余命数

知乎@余命数

$$x^n - x^{n-1} - x^{n-2} = 0$$

知乎@余命数

两边同除以 x^{n-2}, 得到这个数列的**特征方程**

$$x^2 - x - 1 = 0$$

利用二次方程求根公式, 解得两个根

知乎@余命数

$$\varphi = \frac{1+\sqrt{5}}{2}, \quad \psi = \frac{1-\sqrt{5}}{2}$$

φ 是著名的**黄金分割比**. 由于 $F_n = \varphi^n$ 和 $F_n = \psi^n$ 都是满足递推关系的解, 那么它们的任意线性组合 $F_n = A\varphi^n + B\psi^n$ 也同样满足. 接着, 利用初始条件 $F_0 = 0, F_1 = 1$ 来确定常数 A 和 B.

知乎@余命数

$$\begin{aligned} n = 0: &\quad F_0 = A\varphi^0 + B\psi^0 = A + B = 0 \implies B = -A \\ n = 1: &\quad F_1 = A\varphi^1 + B\psi^1 = A\varphi - A\psi = A(\varphi - \psi) = 1 \end{aligned}$$

我们计算 $\varphi - \psi = \frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2} = \sqrt{5}$. 于是 $A\sqrt{5} = 1$, 解得 $A = \frac{1}{\sqrt{5}}$, 从而 $B = -\frac{1}{\sqrt{5}}$. 将 A, B 代回, 得斐波那契数列的通项公式, 即**比内公式**.

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right] \tag{7-5}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个充满了无理数 $\sqrt{5}$ 的公式, 对于任意非负整数 n, 其计算结果永远是一个整数. 在代数运算中, 所有的 $\sqrt{5}$ 项都相互抵消了.

7.7.2 前 n 项和

斐波那契数列的求和公式也很简洁. 现推导 $S_n = \sum_{k=1}^{n} F_k$ 的封闭表达式.

由定义式 $F_{k+2} = F_{k+1} + F_k$ 移项, 得到裂项所需的形式

$$F_k = F_{k+2} - F_{k+1}$$

它表明, 每一个斐波那契数都可以表示为后两个斐波那契数之差. 接着将和式 S_n 中的每一项都用这种形式替换:

$$\begin{aligned} S_n &= F_1 + F_2 + F_3 + \cdots + F_n \\ &= (F_3 - F_2) + (F_4 - F_3) + (F_5 - F_4) + \cdots + (F_{n+2} - F_{n+1}) \end{aligned}$$

这是**裂项相消**求和. 中间项 $-F_3$ 与 $+F_3$, $-F_4$ 与 $+F_4$ 等依次抵消, 最后只剩下 $-F_2$ 和 $+F_{n+2}$.

$$S_n = F_{n+2} - F_2$$

根据我们的定义 $F_2 = 1$, 得最终的求和公式

$$\sum_{k=1}^{n} F_k = F_{n+2} - 1 \tag{7-6}$$

这个结果就是斐波那契恒等式. 前 n 个斐波那契数的和等于 $F_{n+2} - 1$.

例题 7.29　我们验证 $n = 5$.

$$S_5 = F_1 + F_2 + F_3 + F_4 + F_5 = 1 + 1 + 2 + 3 + 5 = 12$$

根据公式, 结果应为

$$F_{5+2} - 1 = F_7 - 1 = 13 - 1 = 12$$

7.7.3 平方和

求出前 n 项和后, 自然可以研究**前 n 项的平方和** $\sum_{k=1}^{n} F_k^2$. 这个和也有简洁的封闭形式.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定理 7.2 对于斐波那契数列 $\{F_n\}$, 其前 n 项平方和为

$$\sum_{k=1}^{n} F_k^2 = F_n F_{n+1}$$

证明 需找一个关于 F_k^2 的恒等式来实现裂项. 从恒等式出发, 对任意 $k \geq 1$:

$$F_{k+1} - F_{k-1} = (F_k + F_{k-1}) - F_{k-1} = F_k.$$

据此把 F_k^2 中的一个 F_k 替换掉.

$$F_k^2 = F_k \cdot (F_{k+1} - F_{k-1}) = F_k F_{k+1} - F_k F_{k-1}.$$

令 $A_k = F_k F_{k+1}$, 则 $F_k F_{k-1}$ 即 A_{k-1}, 上式写成裂项形式

$$F_k^2 = A_k - A_{k-1}.$$

对整个数列求和.

$$\begin{aligned}\sum_{k=1}^{n} F_k^2 &= \sum_{k=1}^{n} (A_k - A_{k-1}) \\ &= (A_1 - A_0) + (A_2 - A_1) + (A_3 - A_2) + \cdots + (A_n - A_{n-1}).\end{aligned}$$

裂项求和时, 中间项 $-A_1$ 与 $+A_1$, $-A_2$ 与 $+A_2$ 等依次抵消, 最后只剩下 $-A_0$ 和 $+A_n$.

$$\sum_{k=1}^{n} F_k^2 = A_n - A_0.$$

计算 $A_n = F_n F_{n+1}$, $A_0 = F_0 F_{0+1} = F_0 F_1 = 0 \cdot 1 = 0$, 故

$$\sum_{k=1}^{n} F_k^2 = F_n F_{n+1} - 0 = F_n F_{n+1}.$$

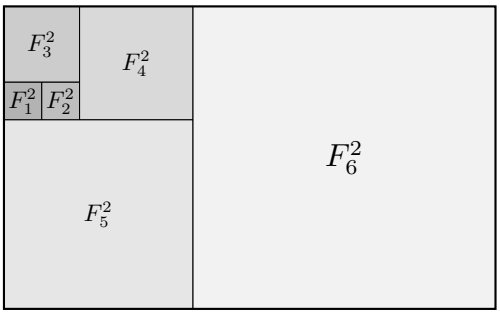
证毕. □

这个恒等式还有一个几何证明. 用边长为 F_k 的正方形面积表示 F_k^2, 再把它们拼成一个矩形.

观察这个拼接过程.

- 从 $F_1^2 (1 \times 1)$ 开始, 这是一个 $F_1 \times F_2$ 的矩形.
- 加上 $F_2^2 (1 \times 1)$, 得一个 2×1 的矩形, 其面积为 $F_1^2 + F_2^2 = F_2 F_3$.
- 加上 $F_3^2 (2 \times 2)$, 得一个 2×3 的矩形, 其面积为 $\sum_{k=1}^{3} F_k^2 = F_3 F_4$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



- 加上 $F_4^2(3\times 3)$, 得一个 5×3 的矩形, 其面积为 $\sum_{k=1}^{4} F_k^2 = F_4F_5$.

每一步, 把边长为 F_n 的正方形拼到长宽为 F_{n-1} 和 F_n 的矩形上, 得到长为 F_n、宽为 $F_{n-1}+F_n = F_{n+1}$ 的新矩形. 这个过程不断进行, 说明所有正方形面积之和等于最后两个斐波那契数所构成矩形的面积.

7.7.4 相邻项乘积之和

在分析了平方和之后, 自然要问相邻两项乘积之和是否也有简洁闭式.

$$\sum_{k=1}^{n} F_kF_{k+1}.$$

答案是肯定的, 现给出按奇偶情形区分的精确公式.

定理 7.3　设 $S_n = \sum_{k=1}^{n} F_kF_{k+1}$. 则

$$S_n = \begin{cases} F_{n+1}^2 & \text{如果} n \text{是奇数} \\ F_nF_{n+2} & \text{如果} n \text{是偶数} \end{cases}$$

证明　这个结论按奇偶区分, 可用交错数学归纳法证明. 我们同时证明对任意正整数 m, 命题在 $n=2m-1$ 和 $n=2m$ 时都成立.

设 $S_n = \sum_{k=1}^{n} F_kF_{k+1}$. 于是有两个递推关系.

$$S_{2m} = S_{2m-1} + F_{2m}F_{2m+1} \tag{1}$$

$$S_{2m+1} = S_{2m} + F_{2m+1}F_{2m+2} \tag{2}$$

对奇数 $n=1, S_1 = F_1F_2 = 1\cdot 1 = 1$, 公式结果为 $F_{1+1}^2 = F_2^2 = 1^2 = 1$, 命题成立. 对偶数 $n=2, S_2 = S_1 + F_2F_3 = 1+1\cdot 2 = 3$, 公式结果为 $F_2F_{2+2} = F_2F_4 = 1\cdot 3 = 3$, 命题成立.

假设当 $m=k$ 时, 即对于 $n=2k-1$ 和 $n=2k$, 命题均成立.

$$S_{2k-1} = F_{2k}^2 \qquad \text{(归纳假设 1)}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

$$S_{2k} = F_{2k}F_{2k+2} \qquad \text{(归纳假设 2)}$$

需证当 $m = k+1$ 时, 即对于 $n = 2k+1$ 和 $n = 2k+2$, 命题也成立.

先证 $S_{2k+1} = F_{(2k+1)+1}^2 = F_{2k+2}^2$. 由关系式 (2) 与归纳假设 2 出发, 得

$$\begin{aligned} S_{2k+1} &= S_{2k} + F_{2k+1}F_{2k+2} \\ &= (F_{2k}F_{2k+2}) + F_{2k+1}F_{2k+2} \\ &= (F_{2k} + F_{2k+1})F_{2k+2} \\ &= F_{2k+2} \cdot F_{2k+2} \quad (\text{斐波那契定义 } F_{2k} + F_{2k+1} = F_{2k+2}) \\ &= F_{2k+2}^2, \end{aligned}$$

即奇数 $n = 2k+1$ 的目标.

再证 $S_{2k+2} = F_{2k+2}F_{(2k+2)+2} = F_{2k+2}F_{2k+4}$. 由关系式 (1)(将 m 换为 $k+1$) 与刚证得的 $S_{2k+1} = F_{2k+2}^2$ 出发, 得

$$\begin{aligned} S_{2k+2} &= S_{2k+1} + F_{2k+2}F_{2k+3} \\ &= F_{2k+2}^2 + F_{2k+2}F_{2k+3} \\ &= F_{2k+2}(F_{2k+2} + F_{2k+3}) \\ &= F_{2k+2}F_{2k+4} \quad (\text{斐波那契定义 } F_{2k+2} + F_{2k+3} = F_{2k+4}). \end{aligned}$$

即偶数 $n = 2k+2$ 的目标.

奠基步骤成立, 归纳步骤证明了从 k 到 $k+1$ 的传递性, 故该定理对所有正整数 n 均成立. □

7.7.5 对称项之和

斐波那契数列中还有许多对称恒等式. 下列关系式说明任意一项与它前后等距两项之间的联系.

定理 7.4 对于斐波那契数列 $\{F_n\}$, 任意一项的三倍, 等于它前面第二项与后面第二项之和.

$$3F_n = F_{n+2} + F_{n-2} \quad (n \geq 2)$$

证明 从右侧入手, 反复用斐波那契数列的定义化简 $F_{n+2} + F_{n-2}$. 将 F_{n+2} 向下展开, 用 F_n 及其相邻项表示.

$$F_{n+2} = F_{n+1} + F_n$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= (F_n + F_{n-1}) + F_n$$
$$= 2F_n + F_{n-1}.$$

对 F_{n-2}, 由定义式 $F_n = F_{n-1} + F_{n-2}$ 移项得

$$F_{n-2} = F_n - F_{n-1},$$

故 F_{n-2} 也以 F_n 和 F_{n-1} 表示. 将两式代回原表达式, 得

$$\begin{aligned} F_{n+2} + F_{n-2} &= (2F_n + F_{n-1}) + (F_n - F_{n-1}) \\ &= 2F_n + F_n + F_{n-1} - F_{n-1} \\ &= 3F_n. \end{aligned}$$

等式右侧化简为左侧, 证毕. □

黄金分割极限

注 7.6
数值上 F_{n+1}/F_n 很快靠近 $\varphi \approx 1.618$; 现用比内公式作严格极限论证, 依赖 $|q| < 1$ 时 $q^n \to 0$ 等结论, 见 第16章.

例题 7.30 (黄金分割极限) 证明斐波那契数列相邻两项的比值, 当 n 趋于无穷时, 收敛于黄金分割比 φ.

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \frac{1+\sqrt{5}}{2}$$

解 这是数列极限问题, 描述斐波那契数列的渐近行为. 递推公式无法直接处理, 需用通项公式, 即比内公式. 记 $\varphi = \frac{1+\sqrt{5}}{2}$ 和 $\psi = \frac{1-\sqrt{5}}{2}$, 比内公式为 $F_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}$. 构造比值

$$\frac{F_{n+1}}{F_n} = \frac{\frac{\varphi^{n+1} - \psi^{n+1}}{\sqrt{5}}}{\frac{\varphi^n - \psi^n}{\sqrt{5}}} = \frac{\varphi^{n+1} - \psi^{n+1}}{\varphi^n - \psi^n}.$$

求极限需识别主导项. 注意到 $|\varphi| = \frac{1+\sqrt{5}}{2} \approx 1.618 > 1$, $|\psi| = \left|\frac{1-\sqrt{5}}{2}\right| \approx 0.618 < 1$, 当 $n \to \infty$ 时 $\psi^n \to 0$ 而 $\varphi^n \to \infty$, 故 φ^n 为主导项. 将分子分母同除以 φ^n:

$$\frac{F_{n+1}}{F_n} = \frac{\varphi - \psi \cdot (\frac{\psi}{\varphi})^n}{1 - (\frac{\psi}{\varphi})^n}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因 $|\frac{\psi}{\varphi}| = |\frac{1-\sqrt{5}}{1+\sqrt{5}}| < 1$, 故 $\lim_{n\to\infty}\left(\frac{\psi}{\varphi}\right)^n = 0$, 从而

$$\lim_{n\to\infty}\frac{F_{n+1}}{F_n} = \frac{\varphi - \psi\cdot 0}{1-0} = \varphi.$$

斐波那契数列的相邻项之比确收敛于黄金分割比 $\varphi = \frac{1+\sqrt{5}}{2}$.

知乎@余命数

例题 7.31 计算前 10 项斐波那契数的平方和 $\sum_{k=1}^{10} F_k^2$.

解 逐项计算平方再相加很繁琐. 利用已证的平方和公式, 问题可化简.

已证明的平方和公式为

$$\sum_{k=1}^{n} F_k^2 = F_n F_{n+1}$$

知乎@余命数

在本题中, $n = 10$. 因此, 只需要计算出 F_{10} 和 F_{11} 即可. 回忆斐波那契数列.

知乎@余命数

$$1, 1, 2, 3, 5, 8, 13, 21, 34, \underbrace{55}_{F_{10}}, \underbrace{89}_{F_{11}}, ...$$

所以 $F_{10} = 55$, $F_{11} = 89$.

代入公式, 得

知乎@余命数

$$\sum_{k=1}^{10} F_k^2 = F_{10}F_{11} = 55 \times 89$$

计算可得

$$55 \times 89 = 55 \times (90 - 1) = 4950 - 55 = 4895$$

因此, 前 10 项斐波那契数的平方和为 4895.

例题 7.32 求 $\sum_{k=1}^{8}(F_k + F_{k+1})^2$ 的值.

知乎@余命数

解 通项是完全平方式, 直接计算复杂, 先展开再拆成已知形式, 得

$$(F_k + F_{k+1})^2 = F_k^2 + 2F_kF_{k+1} + F_{k+1}^2,$$

故原求和式拆分为三部分.

$$\sum_{k=1}^{8}(F_k + F_{k+1})^2 = \sum_{k=1}^{8} F_k^2 + 2\sum_{k=1}^{8} F_kF_{k+1} + \sum_{k=1}^{8} F_{k+1}^2.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

逐个处理. 部分一为平方和, 由公式 $\sum_{k=1}^{n} F_k^2 = F_n F_{n+1}$ 得

$$\sum_{k=1}^{8} F_k^2 = F_8 F_9 = 21 \times 34 = 714.$$

部分二为相邻项乘积之和. 因 $n = 8$ 是偶数, 用公式 $\sum_{k=1}^{n} F_k F_{k+1} = F_n F_{n+2}$, 得

$$2\sum_{k=1}^{8} F_k F_{k+1} = 2(F_8 F_{10}) = 2 \times (21 \times 55) = 2310.$$

部分三 $\sum_{k=1}^{8} F_{k+1}^2$ 是变形的平方和. 换元 $j = k + 1$ 得

$$\sum_{k=1}^{8} F_{k+1}^2 = F_2^2 + F_3^2 + \cdots + F_9^2 = \left(\sum_{k=1}^{9} F_k^2\right) - F_1^2,$$

由平方和公式

$$\left(\sum_{k=1}^{9} F_k^2\right) - F_1^2 = (F_9 F_{10}) - 1^2 = (34 \times 55) - 1 = 1869.$$

三部分相加得

$$714 + 2310 + 1869 = 4893.$$

最终结果为 4893.

例题 7.33 证明 $\sum_{k=1}^{n} F_{2k-1} = F_{2n}$.

证明 这个和式只对奇数项求和, 没有现成公式可用, 但通项形式适合裂项, 需把 F_{2k-1} 写成两项之差. 由斐波那契数列的定义, 偶数项与其前两项的关系为

$$F_{2k} = F_{2k-1} + F_{2k-2},$$

移项得

$$F_{2k-1} = F_{2k} - F_{2k-2},$$

此式把一个奇数项表示成两个相邻偶数项之差. 对整个和式裂项.

$$\begin{aligned}\sum_{k=1}^{n} F_{2k-1} &= \sum_{k=1}^{n} (F_{2k} - F_{2k-2}) \\ &= (F_2 - F_0) + (F_4 - F_2) + (F_6 - F_4) + \cdots + (F_{2n} - F_{2n-2}).\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

裂项求和时,中间项 $-F_2$ 与 $+F_2$,$-F_4$ 与 $+F_4$ 等全部抵消,最终只剩第一项的 $-F_0$ 和最后一项的 $+F_{2n}$.

$$\sum_{k=1}^{n} F_{2k-1} = F_{2n} - F_0.$$

由定义 $F_0 = 0$,得

$$\sum_{k=1}^{n} F_{2k-1} = F_{2n}.$$

证毕. □

例题 7.34 (23-24 河南漯河期末) 分析并判断以下关于斐波那契数列 $\{F_n\}$ (定义为 $F_0 = 0, F_1 = 1, \ldots$)的四个命题是否正确,并给出严格的证明或反例.
命题 A 数列 $\{F_n\}$ 是一个递增数列.
命题 B 89 是数列 $\{F_n\}$ 中的一项.
命题 C 对所有 $n \geq 1$,恒等式 $\sum_{k=1}^{n} F_k^2 = F_n F_{n+1}$ 成立.(原题是要求从前 10 项注意到,稍微有些不严谨.)
命题 D 比值 $\dfrac{F_n}{F_{n+1}}$ 收敛,且其极限为 $\dfrac{\sqrt{5}-1}{2}$.

解 逐一判断四个命题的真伪.

命题 A 错误. 严格递增数列要求对任意 $n \geq 0$ 都有 $F_{n+1} > F_n$. 但在本数列中,有 $F_1 = 1$ 且 $F_2 = 1$,不满足 $F_2 > F_1$. 因此,斐波那契数列是一个**非减数列**,但不是严格递增数列.

命题 B 正确. 通过逐项计算,得数列的前几项,得

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ \underline{89},\ \ldots$$

可见 89 是数列的第 11 项 ($F_{11} = 89$).

命题 C 正确. 这正是在本节平方和部分,通过代数裂项和几何拼接两种方法证明过的恒等式. 它对所有正整数 n 都成立.

命题 D 正确. 在前面的例题中,已证明斐波那契数列相邻两项的比值收敛于黄金分割比 φ.

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \varphi = \frac{1+\sqrt{5}}{2}$$

本题所求的极限是其倒数. 在已接受 $\lim F_{n+1}/F_n = \varphi$ 的前提下,由极限运算法则(第16章),得

$$\lim_{n\to\infty} \frac{F_n}{F_{n+1}} = \frac{1}{\lim_{n\to\infty} \frac{F_{n+1}}{F_n}} = \frac{1}{\varphi}$$

知乎@余命数
知乎@余命数
知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对 $\frac{1}{\varphi}$ 作分子有理化.

$$\frac{1}{\varphi}=\frac{2}{1+\sqrt{5}}=\frac{2(\sqrt{5}-1)}{(\sqrt{5}+1)(\sqrt{5}-1)}=\frac{2(\sqrt{5}-1)}{5-1}=\frac{2(\sqrt{5}-1)}{4}=\frac{\sqrt{5}-1}{2}$$

因此, 该命题正确. 这个值有时被称为黄金分割比的共轭, 记作 $\frac{1}{\varphi}$ 或 Φ.

命题 A 错误, 命题 B、C、D 均正确.

例题 7.35 (2024 年新高考一卷第 8 题) 已知定义在 $\mathbb{R}$ 上的函数 $f(x)$ 满足 $f(x)>f(x-1)+f(x-2)$, 且当 $x<3$ 时, $f(x)=x$. 判断下列结论中哪一个必定正确?

(A) $f(10)>100$

(B) $f(20)>1000$

(C) $f(10)<1000$

(D) $f(20)<10000$

解 问题要求理解不等式 $f(x)>f(x-1)+f(x-2)$ 的含义. 该函数的增长速度**超过**斐波那契数列, 因此可用一个标准斐波那契数列作为 $f(x)$ 的**下界**, 再据此估值.

不等式只给出下限, $f(x)$ 的值可以比下限大很多, 例如 $f(3)$ 可以是 3.1, 也可以是 10^9. 因此, 任何给出 $f(x)$ **上界**的结论, 如 (C) 和 (D), 都无法从已知条件保证, 不可能是必定正确的. 只需关心与下界相关的 (A) 和 (B).

构造一个基准数列 $\{g_n\}$, 它满足斐波那契的**等式**关系.

$$g_n=g_{n-1}+g_{n-2}\quad(n\geq 3).$$

用 $f(x)$ 的已知值设定初始项. 对于整数 n, 得

$$f(1)=1\quad(\text{因为}1<3)$$

$$f(2)=2\quad(\text{因为}2<3),$$

故设 $g_1=1, g_2=2$. 计算 $\{g_n\}$ 的前几项, 得

$$g_1=1,\quad g_2=2,\quad g_3=g_2+g_1=3,$$

$$g_4=g_3+g_2=5,\quad g_5=g_4+g_3=8.$$

数列 $\{g_n\}$ 正是斐波那契数列 $\{F_n\}$ 的平移. 于是 $g_n=F_{n+1}$.

用数学归纳法证明 $f(n)>g_n=F_{n+1}$ 对所有整数 $n\geq 3$ 成立.

当 $n=3$ 时, 由题设 $f(3)>f(2)+f(1)=2+1=3$, 而 $g_3=3$, 故 $f(3)>g_3$ 成立, 奠基步骤完成.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

假设当 $3 \le k < n$ 时不等式 $f(k) > g_k$ 成立, 需证 $f(n) > g_n$. 由题设

$$f(n) > f(n-1) + f(n-2).$$

根据归纳假设, $f(n-1) > g_{n-1}$ 且 $f(n-2) > g_{n-2}$(只要 $n-2 \ge 3$). 当 $n-2=2$ 即 $n=4$ 时, $f(2)=g_2$, 需单独考虑.

$$\text{若} n=4, \quad f(4) > f(3) + f(2) > g_3 + g_2 = g_4;$$

$$\text{若} n \ge 5, \quad f(n) > f(n-1) + f(n-2) > g_{n-1} + g_{n-2} = g_n.$$

综合 $n=4$ 与 $n \ge 5$ 的讨论, 对所有整数 $n \ge 3$ 都有 $f(n) > g_n = F_{n+1}$.

用此下界检验选项 (A) 和 (B).

选项 (A) $f(10) > 100$. 由结论 $f(10) > g_{10} = F_{11}$. 计算斐波那契数列 $F_1 = 1, \dots, F_8 = 21, F_9 = 34, F_{10} = 55, F_{11} = 89$, 只能确定 $f(10) > 89$, 并不能保证 $f(10)$ 一定大于 100. 可构造满足条件但 $f(10) < 100$ 的函数. 令 $x < 3$ 时 $f(x) = x$, $x \ge 3$ 时递推定义 $f(x) = f(x-1) + f(x-2) + 0.1$. 它满足题设不等式, 而整数点上依次有 $f(3) = 3.1, f(4) = 5.2, \dots, f(10) = 94.4 < 100$. 故 (A) 不一定正确.

选项 (B) $f(20) > 1000$. 由结论 $f(20) > g_{20} = F_{21}$. 继续计算斐波那契数列, 得

$$F_{11} = 89, \quad F_{12} = 144, \quad F_{13} = 233, \quad F_{14} = 377,$$

$$F_{15} = 610, \quad F_{16} = 987, \quad F_{17} = 1597, \quad F_{18} = 2584,$$

$$F_{19} = 4181, \quad F_{20} = 6765, \quad F_{21} = 10946.$$

得 $f(20) > F_{21} = 10946$. 因 $10946 > 1000$, 故 $f(20) > 1000$ 一定成立, (B) 正确.

本题把函数与数列结合在一起. 条件 $f(x) > f(x-1) + f(x-2)$ 与斐波那契递推 $F_n = F_{n-1} + F_{n-2}$ 只差一点, 即不等号. 这说明 f 的增长速度不低于斐波那契数列.

处理这类不等式时, 可先构造一个满足等式的标准斐波那契数列 g_n 作为下界. 令 $g_1 = f(1) = 1$, $g_2 = f(2) = 2$, 则 $g_n = g_{n-1} + g_{n-2}$, 也就是 $g_n = F_{n+1}$. 再用数学归纳法证明 $f(n) > g_n$ 对所有 $n \ge 3$ 成立, 下界就落到实处.

7.8 数列求和的通用策略

我们已经讨论了等差数列与等比数列的求和公式. 实际遇到的数列常常比这两种模型更复杂, 既非等差, 也非等比. 我们要计算这些数列的和, 就需要一套通用的代数技巧.

这些技巧的目标一致. 我们总是把陌生、复杂的求和问题**化归**为熟悉或简单的形式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

7.8.1 基本公式法

当数列的通项 a_n 是关于 n 的多项式时，我们利用求和符号 $\sum$ 的线性性质，把原求和式拆成若干基本幂次和，再用已知公式计算.

我们先写下几个基本的自然数幂次和公式.

- 常数的和是 $\sum_{k=1}^{n} c = cn$.
- 一次方的和是 $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.
- 二次方的和是 $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

下面补充平方和公式的推导.

7.8.1.1 扰动法

我们从寻找与目标和式 $\sum k^2$ 相关的裂项求和开始.$(k+1)^3 - k^3$ 是关于 k 的二次多项式，正好把 k^2 与裂项联系起来.

考虑恒等式

$$(k+1)^3 - k^3 = k^3 + 3k^2 + 3k + 1 - k^3 = 3k^2 + 3k + 1$$

这个恒等式对所有 k 均成立. 现在，对这个等式的两边同时从 $k=1$ 到 n 求和，得

$$\sum_{k=1}^{n} \left[(k+1)^3 - k^3\right] = \sum_{k=1}^{n} (3k^2 + 3k + 1)$$

分别计算等式左右两边.

左侧 (LHS) 是一个典型的裂项求和，计算如下.

$$\begin{aligned}\sum_{k=1}^{n} \left[(k+1)^3 - k^3\right] &= (2^3 - 1^3) + (3^3 - 2^3) + (4^3 - 3^3) \\ &\quad + \cdots + ((n+1)^3 - n^3) \\ &= -1^3 + (2^3 - 2^3) + (3^3 - 3^3) + \ldots \\ &\quad + (n^3 - n^3) + (n+1)^3 \\ &= (n+1)^3 - 1\end{aligned}$$

右侧 (RHS) 利用求和的线性性质拆分，得

$$\sum_{k=1}^{n} (3k^2 + 3k + 1) = 3\sum_{k=1}^{n} k^2 + 3\sum_{k=1}^{n} k + \sum_{k=1}^{n} 1$$

令 $S_2 = \sum_{k=1}^{n} k^2$，并代入已知的一次和与常数和公式，得

$$\text{RHS} = 3S_2 + 3 \cdot \frac{n(n+1)}{2} + n$$

联立左右两侧的结果, 有

$$(n+1)^3-1=3S_2+\frac{3n(n+1)}{2}+n$$

任务现转化为解出 S_2.

$$\begin{aligned}3S_2&=(n+1)^3-1-\frac{3n(n+1)}{2}-n\\&=(n+1)^3-(n+1)-\frac{3n(n+1)}{2}\\&=(n+1)\left[(n+1)^2-1-\frac{3n}{2}\right]\\&=(n+1)\left[n^2+2n+1-1-\frac{3n}{2}\right]\\&=(n+1)\left[n^2+\frac{n}{2}\right]\\&=(n+1)\cdot n\left(n+\frac{1}{2}\right)\\&=\frac{n(n+1)(2n+1)}{2}\end{aligned}$$

两边同除以 3, 即得平方和公式

$$S_2=\sum_{k=1}^{n}k^2=\frac{n(n+1)(2n+1)}{6}$$

7.8.1.2 数学归纳法

已通过扰动法得到公式, 现用归纳法严格证明它.

设命题 $P(n)$ 为 $\sum_{k=1}^{n}k^2=\frac{n(n+1)(2n+1)}{6}$.

当 $n=1$ 时, 左侧 $=\sum_{k=1}^{1}k^2=1$, 右侧 $=\frac{1(1+1)(2\cdot1+1)}{6}=\frac{1\cdot2\cdot3}{6}=1$, 两端相等, 故 $P(1)$ 成立, 奠基步骤完成.

假设当 $n=m(m\geq1$ 为任意整数)时命题成立, 即

$$\sum_{k=1}^{m}k^2=\frac{m(m+1)(2m+1)}{6}.$$

由此只需证明当 $n=m+1$ 时

$$\sum_{k=1}^{m+1}k^2=\frac{(m+1)((m+1)+1)(2(m+1)+1)}{6}=\frac{(m+1)(m+2)(2m+3)}{6}$$

也成立. 我们从左侧开始推演.

$$\text{左侧}=\sum_{k=1}^{m+1}k^2=\left(\sum_{k=1}^{m}k^2\right)+(m+1)^2$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
&= \frac{m(m+1)(2m+1)}{6} + (m+1)^2 \quad \text{（代入归纳假设）} \\
&= (m+1)\left[\frac{m(2m+1)}{6} + (m+1)\right] \quad \text{（提取公因子 } (m+1)\text{）} \\
&= (m+1)\left[\frac{2m^2+m+6(m+1)}{6}\right] \\
&= \frac{(m+1)(2m^2+7m+6)}{6} \\
&= \frac{(m+1)(m+2)(2m+3)}{6} \quad \text{（因式分解 } 2m^2+7m+6\text{）} \\
&= \text{右侧}.
\end{aligned}
$$

故若命题对 $n=m$ 成立, 则对 $n=m+1$ 亦成立, 归纳步骤完成.

由数学归纳法原理, 命题 $P(n)$ 对所有正整数 n 均成立.

知乎@余命数

7.8.1.3 待定系数法

我们注意到, d 次多项式的求和结果是一个 $d+1$ 次多项式. 由于 k^2 是关于 k 的二次多项式, 其前 n 项和 $S_n=\sum_{k=1}^{n} k^2$ 必然是关于 n 的三次多项式.

设 $S_n = An^3+Bn^2+Cn+D$. 我们的任务是确定系数 A, B, C, D. 我们利用关系 $a_n = S_n - S_{n-1}$, 得到

$$
\begin{aligned}
n^2 &= S_n - S_{n-1} \\
&= (An^3+Bn^2+Cn+D) - [A(n-1)^3+B(n-1)^2+C(n-1)+D] \\
&= A(n^3-(n^3-3n^2+3n-1)) + B(n^2-(n^2-2n+1)) + C(n-(n-1)) \\
&= A(3n^2-3n+1) + B(2n-1) + C \\
&= 3An^2 + (2B-3A)n + (A-B+C)
\end{aligned}
$$

比较等式两边 n 的同次幂系数, 得到

- n^2 的系数: $3A=1 \implies A=\frac{1}{3}$
- n 的系数: $2B-3A=0 \implies 2B=3(\frac{1}{3})=1 \implies B=\frac{1}{2}$
- 常数项: $A-B+C=0 \implies C=B-A=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$

因此, $S_n=\frac{1}{3}n^3+\frac{1}{2}n^2+\frac{1}{6}n+D$. 为了确定 D, 我们利用初始条件 $S_1=1$.

$$
S_1 = \frac{1}{3}(1)^3+\frac{1}{2}(1)^2+\frac{1}{6}(1)+D = \frac{1}{3}+\frac{1}{2}+\frac{1}{6}+D = 1+D
$$

由于 $S_1=1$, 故 $1+D=1 \implies D=0$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

将系数代回, 得到

$$\begin{aligned} S_n &= \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n = \frac{2n^3+3n^2+n}{6} \\ &= \frac{n(2n^2+3n+1)}{6} = \frac{n(n+1)(2n+1)}{6}. \end{aligned}$$

例题 7.36 求和 $S_n = \sum_{k=1}^{n}(k^2-3k+2)$.

解 我们利用求和符号的线性性质将和式分解, 得

知乎@余命数

$$S_n = \sum_{k=1}^{n} k^2 - \sum_{k=1}^{n} 3k + \sum_{k=1}^{n} 2$$

再将常数因子提出, 得

知乎@余命数

$$S_n = \sum_{k=1}^{n} k^2 - 3\sum_{k=1}^{n} k + 2\sum_{k=1}^{n} 1$$

知乎@余命数

最后代入基本幂次和公式, 得

$$\begin{aligned} S_n &= \frac{n(n+1)(2n+1)}{6} - 3\cdot\frac{n(n+1)}{2} + 2n \\ &= \frac{n}{6}[(n+1)(2n+1) - 9(n+1) + 12] \\ &= \frac{n}{6}[2n^2+3n+1-9n-9+12] \\ &= \frac{n}{6}[2n^2-6n+4] \\ &= \frac{n}{3}(n^2-3n+2) \\ &= \frac{n(n-1)(n-2)}{3} \end{aligned}$$

知乎@余命数 知乎@余命数 知乎@余命数

7.8.2 裂项相消法

若数列的通项 a_k 能分解为两项之差, 即 $a_k = f(k) - f(k+c)$ 或 $a_k = f(k+c) - f(k)$, 我们就可以把求和化简.

把数列的每一项都拆成两部分, 求和时中间的项前后两两抵消, 最终只剩下有限的几项(通常是首尾几项), 从而大幅简化计算.

这一方法的难点在于构造裂项. 以下给出常见的几类代数结构.

- **部分分式型 (I) - 基本型.**

$$\frac{1}{k(k+d)} = \frac{1}{d}\left(\frac{1}{k} - \frac{1}{k+d}\right)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这是最基本的分式裂项. 其构造的要点是在分子上创造出与分母因子相匹配的差, 即 $(k+d)-k=d$. 通过在原式分子分母同乘以 d, 再进行拆分即可得到.

- **部分分式型 (II) - 奇偶项.**

$$\frac{1}{(2k-1)(2k+1)}=\frac{1}{2}\left(\frac{1}{2k-1}-\frac{1}{2k+1}\right)$$

这是上述公式在 $d=2$ 时的特例, 适用于分母是由公差为 2 的等差数列相邻项构成的情况.

- **根式有理化型.**

$$\frac{1}{\sqrt{k+d}+\sqrt{k}}=\frac{1}{d}(\sqrt{k+d}-\sqrt{k})$$

构造时对分母进行有理化. 分子分母同乘以其共轭表达式 $\sqrt{k+d}-\sqrt{k}$, 利用平方差公式 $(\sqrt{a}+\sqrt{b})(\sqrt{a}-\sqrt{b})=a-b$ 消去分母的根号, 使其变为常数 d.

- **平方差裂项型.**

$$\frac{2k+1}{k^2(k+1)^2}=\frac{1}{k^2}-\frac{1}{(k+1)^2}$$

从右侧反向推导更为直观. 通分 $\frac{1}{k^2}-\frac{1}{(k+1)^2}$, 其分子为 $(k+1)^2-k^2=(k^2+2k+1)-k^2=2k+1$, 恰好匹配左侧. 此形式在处理分母为相邻项的完全平方之积时极为有效.

- **指数型.**

$$\frac{2^k}{(2^k-1)(2^{k+1}-1)}=\frac{1}{2^k-1}-\frac{1}{2^{k+1}-1}$$

同样从右侧通分验证. 其分子为 $(2^{k+1}-1)-(2^k-1)=2^{k+1}-2^k=2\cdot 2^k-2^k=2^k$, 与左侧分子完全吻合. 本质是利用了指数函数 $f(k)=2^k-1$ 的差分性质.

- **混合型(指数与多项式).**

$$\frac{2^k(k-1)}{k(k+1)}=\frac{2^{k+1}}{k+1}-\frac{2^k}{k}$$

发现此类混合裂项的通常途径, 是定义函数 $f(k)=\frac{2^k}{k}$ 并考察其差分 $f(k+1)-f(k)$. 计算 $f(k+1)-f(k)=\frac{2^{k+1}}{k+1}-\frac{2^k}{k}$ 并通分, 即可得到左侧的表达式.

- **交错分式型.**

$$\frac{(-1)^k(k+1)}{(2k+1)(2k+3)}=\frac{1}{4}\left(\frac{(-1)^k}{2k+1}-\frac{(-1)^{k+1}}{2k+3}\right)$$

从右侧方括号内通分, 分子为 $(-1)^k(2k+3)-(-1)^{k+1}(2k+1)$. 利用 $(-1)^{k+1}=-(-1)^k$, 分子化为 $(-1)^k[(2k+3)+(2k+1)]=(-1)^k(4k+4)=4(-1)^k(k+1)$, 再与系数 $\frac{1}{4}$ 相乘, 恰好得到左侧.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **交错根式型.**

$$\frac{(-1)^k}{\sqrt{k}+\sqrt{k-1}}=(-1)^k\sqrt{k}+(-1)^{k-1}\sqrt{k-1}$$

- **高阶部分分式型.**

$$\frac{1}{k(k+1)(k+2)}=\frac{1}{2}\left(\frac{1}{k(k+1)}-\frac{1}{(k+1)(k+2)}\right)$$

知乎@余命数

这是我们在待定系数法一节中推导过的例子, 以 $f(k)=\frac{A}{k(k+1)}$ 为差分原函数进行构造, 是基本部分分式思想的延伸.

- **阶乘添项型 (I).**

$$k\cdot k!=(k+1)!-k!$$

这是阶乘裂项中最经典的形式. 关键技巧在于将因子 k 改写为 $(k+1)-1$, 然后利用乘法分配律展开, 即

$$k\cdot k!=((k+1)-1)k!=(k+1)k!-k!$$

知乎@余命数

- **阶乘添项型 (II).**

$$\frac{k}{(k+1)!}=\frac{1}{k!}-\frac{1}{(k+1)!}$$

知乎@余命数

与上一条技巧完全相同, 我们把分子 k 改写为 $(k+1)-1$.

$$\frac{k}{(k+1)!}=\frac{(k+1)-1}{(k+1)!}=\frac{k+1}{(k+1)!}-\frac{1}{(k+1)!}=\frac{1}{k!}-\frac{1}{(k+1)!}$$

例题 7.37 求和 $S_n=\sum_{k=1}^{n}\frac{1}{k(k+2)}$.

解 通项 $a_k=\frac{1}{k(k+2)}$ 符合部分分式分解的结构, 因此

知乎@余命数

知乎@余命数

$$a_k=\frac{1}{2}\left(\frac{1}{k}-\frac{1}{k+2}\right).$$

这是一个间隔为 2 的裂项. 我们写出和式以观察相消规律.

$$\begin{aligned}2S_n&=\sum_{k=1}^{n}\left(\frac{1}{k}-\frac{1}{k+2}\right)\\&=\left(1-\frac{1}{3}\right)+\left(\frac{1}{2}-\frac{1}{4}\right)+\left(\frac{1}{3}-\frac{1}{5}\right)+\cdots\\&\quad+\left(\frac{1}{n-1}-\frac{1}{n+1}\right)+\left(\frac{1}{n}-\frac{1}{n+2}\right)\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

抵消发生在前一项的负项与后一项的正项之间, 例如 $-\frac{1}{3}$ 与 $+\frac{1}{3}$. 最终, 只有开头的两项正数 1 和 $\frac{1}{2}$, 以及结尾的两项负数 $-\frac{1}{n+1}$ 和 $-\frac{1}{n+2}$ 被保留下来.

$$2S_n = 1 + \frac{1}{2} - \frac{1}{n+1} - \frac{1}{n+2} = \frac{3}{2} - \frac{2n+3}{(n+1)(n+2)}$$

因此,

$$S_n = \frac{3}{4} - \frac{2n+3}{2(n+1)(n+2)} = \frac{n(3n+5)}{4(n+1)(n+2)}$$

例题 7.38　已知正项数列 $\{a_n\}$ 中, $a_1 = 1, a_{n+1}^2 - a_n^2 = 1$, 则数列 $\left\{\dfrac{1}{a_{n+1}+a_n}\right\}$ 的前 99 项和为(　)

(A) 4950

(B) 10

(C) 9

(D) $\dfrac{1}{4950}$

解　已知的是 $\{a_n\}$ 的递推, 所求却是 $b_n = \frac{1}{a_{n+1}+a_n}$ 的前 99 项和. 沟通已知与目标的关键, 在于把递推式等价变形为 b_n 的简单表达式.

递推左侧 $a_{n+1}^2 - a_n^2 = 1$ 是平方差, 因式分解得

$$(a_{n+1} - a_n)(a_{n+1} + a_n) = 1.$$

因子 $a_{n+1} + a_n$ 恰为 b_n 的分母. 因 $\{a_n\}$ 为正项数列, $a_{n+1} + a_n \neq 0$, 两边同除得

$$\frac{1}{a_{n+1} + a_n} = a_{n+1} - a_n.$$

分式被化归为差分, 原求和化为裂项相消:

$$\begin{aligned} S_{99} &= \sum_{n=1}^{99} \frac{1}{a_{n+1}+a_n} = \sum_{n=1}^{99} (a_{n+1} - a_n) \\ &= (a_2 - a_1) + (a_3 - a_2) + \cdots + (a_{100} - a_{99}) = a_{100} - a_1. \end{aligned}$$

余下只需求 a_{100}. 令 $c_n = a_n^2$, 则 $c_{n+1} - c_n = 1$, 故 $\{a_n^2\}$ 是首项 $c_1 = a_1^2 = 1$、公差 1 的等差数列, 通项 $c_n = n$, 即 $a_n^2 = n$. 由正项知 $a_n = \sqrt{n}$, 从而 $a_{100} = 10$.

$$S_{99} = a_{100} - a_1 = 10 - 1 = 9.$$

故选 (C).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.39 数列 $\{a_n\}$ 的通项公式为 $a_n=\dfrac{2n+1}{n^2(n+1)^2}(n\in\mathbb{N}^*)$, 该数列的前 8 项和为 _____.

解 通项较复杂, 宜先裂项. 分母 $n^2(n+1)^2$ 恰为 $\frac{1}{n^2}$ 与 $\frac{1}{(n+1)^2}$ 通分后的分母, 提示 a_n 可能是二者之差. 直接验证:

$$\frac{1}{n^2}-\frac{1}{(n+1)^2}=\frac{(n+1)^2-n^2}{n^2(n+1)^2}=\frac{2n+1}{n^2(n+1)^2}=a_n.$$

故 $a_n=\frac{1}{n^2}-\frac{1}{(n+1)^2}$, 属标准裂项. 前 8 项和

$$S_8=\sum_{n=1}^{8}\Big(\frac{1}{n^2}-\frac{1}{(n+1)^2}\Big)=1-\frac{1}{9^2}=\frac{80}{81}.$$

例题 7.40 (2013 全国课标 I 卷) 已知等差数列 $\{a_n\}$ 的前 n 项和为 S_n, 满足 $S_3=0, S_5=-5$.

1. 求 $\{a_n\}$ 的通项公式;
2. 求数列 $\left\{\dfrac{1}{a_{2n-1}a_{2n+1}}\right\}$ 的前 n 项和.

解 设首项 a_1、公差 d. 由 $S_n=na_1+\frac{n(n-1)}{2}d$ 与已知条件,

$$\begin{cases}S_3=3a_1+3d=0\\S_5=5a_1+10d=-5\end{cases}\Longrightarrow\begin{cases}a_1+d=0\\a_1+2d=-1\end{cases},$$

解得 $d=-1, a_1=1$, 通项 $a_n=2-n$.

第二问令 $b_n=\frac{1}{a_{2n-1}a_{2n+1}}$. 由通项 $a_{2n-1}=3-2n, a_{2n+1}=1-2n$, 故

$$b_n=\frac{1}{(3-2n)(1-2n)}.$$

分母两因子之差为 2, 裂项得

$$b_n=\frac{1}{2}\Big(\frac{1}{1-2n}-\frac{1}{3-2n}\Big).$$

令 $f(k)=\frac{1}{3-2k}$, 则 $f(k+1)=\frac{1}{1-2k}$, 故 $b_k=\frac{1}{2}(f(k+1)-f(k))$.

令所求前 n 项和为 T_n, 则

$$T_n=\sum_{k=1}^{n}b_k=\frac{1}{2}\sum_{k=1}^{n}(f(k+1)-f(k))$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{1}{2}\left[(f(2)-f(1))+(f(3)-f(2))+\cdots+(f(n+1)-f(n))\right]$$

这是一个首尾相连的裂项求和, 中间项相互抵消, 最终结果为

$$\begin{aligned} T_n &= \frac{1}{2}(f(n+1)-f(1)) \\ &= \frac{1}{2}\left(\frac{1}{3-2(n+1)}-\frac{1}{3-2(1)}\right) \\ &= \frac{1}{2}\left(\frac{1}{1-2n}-1\right) \\ &= \frac{1}{2}\left(\frac{1-(1-2n)}{1-2n}\right) \\ &= \frac{n}{1-2n} \end{aligned}$$

例题 7.41 (2014 大纲卷) 已知等差数列 $\{a_n\}$ 的前 n 项和为 S_n、$a_1=10$、a_2 为整数, 且对任意 $n\in\mathbb{N}^*$ 都有 $S_n\le S_4$.

1. 求 $\{a_n\}$ 的通项公式;
2. 设 $b_n=\dfrac{1}{a_n a_{n+1}}$, 求数列 $\{b_n\}$ 的前 n 项和 T_n.

解 条件 $S_n\le S_4$ 对所有正整数 n 成立, 意味着数列 $\{S_n\}$ 在 $n=4$ 处取得最大值. 对于等差数列, 其前 n 项和 S_n 是关于 n 的二次函数, $S_n=\frac{d}{2}n^2+(a_1-\frac{d}{2})n$. 欲使 S_n 存在最大值, 其二次项系数必须为负, 即公差 $d<0$.

S_n 在 $n=4$ 取得最大值, 这也意味着从第 5 项开始, 数列的项 a_n 开始对总和产生负贡献, 或零贡献. 这等价于 $S_4\ge S_3$ 且 $S_4\ge S_5$. 由 $S_n-S_{n-1}=a_n$, 上述条件转化为:

$$\begin{cases} a_4\ge 0 \\ a_5\le 0 \end{cases}$$

利用通项公式 $a_n=a_1+(n-1)d$ 及已知 $a_1=10$, 得关于公差 d 的不等式组:

$$\begin{cases} 10+3d\ge 0 \\ 10+4d\le 0 \end{cases} \Longrightarrow \begin{cases} d\ge -\frac{10}{3} \\ d\le -\frac{10}{4} \end{cases}$$

即 $-\frac{10}{3}\le d\le -\frac{5}{2}$.

又因为 $a_2=a_1+d=10+d$ 是整数, 且 $a_1=10$ 是整数, 所以公差 d 必须是整数. 在区间 $[-\frac{10}{3},-\frac{5}{2}]$, 即 $[-3.33\ldots,-2.5]$ 内唯一的整数是 $d=-3$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由此唯一确定该等差数列. 其通项公式为:

$$a_n = 10 + (n-1)(-3) = 13 - 3n$$

第二问求数列 $\{b_n\}$ 的和, 其通项为 $b_n = \dfrac{1}{a_n a_{n+1}}$. 代入 $a_n = 13 - 3n$, 有 $a_{n+1} = 13 - 3(n+1) = 10 - 3n$.

$$b_n = \frac{1}{(13-3n)(10-3n)}$$

分母两因子之差为 $(13-3n)-(10-3n)=3$. 可进行裂项分解:

$$b_n = \frac{1}{3}\left(\frac{1}{10-3n} - \frac{1}{13-3n}\right)$$

令 $f(k) = \dfrac{1}{13-3k}$, 那么 $f(k+1) = \dfrac{1}{13-3(k+1)} = \dfrac{1}{10-3k}$. 于是, $b_k = \dfrac{1}{3}(f(k+1) - f(k))$.

所求前 n 项和 T_n 为:

$$\begin{aligned} T_n &= \sum_{k=1}^{n} b_k = \frac{1}{3}\sum_{k=1}^{n}(f(k+1) - f(k)) \\ &= \frac{1}{3}\left[(f(2)-f(1)) + (f(3)-f(2)) + \cdots + (f(n+1)-f(n))\right] \end{aligned}$$

此为裂项求和, 结果为:

$$\begin{aligned} T_n &= \frac{1}{3}(f(n+1) - f(1)) \\ &= \frac{1}{3}\left(\frac{1}{13-3(n+1)} - \frac{1}{13-3(1)}\right) \\ &= \frac{1}{3}\left(\frac{1}{10-3n} - \frac{1}{10}\right) \\ &= \frac{1}{3}\left(\frac{10-(10-3n)}{10(10-3n)}\right) \\ &= \frac{n}{10(10-3n)} \end{aligned}$$

例题 7.42 (2017 全国 III 卷) 设数列 $\{a_n\}$ 满足 $a_1 + 3a_2 + \cdots + (2n-1)a_n = 2n$.

1. 求 $\{a_n\}$ 的通项公式;
2. 求数列 $\left\{\dfrac{a_n}{2n+1}\right\}$ 的前 n 项和.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 (1) 设 $T_n = a_1 + 3a_2 + \cdots + (2n-1)a_n = 2n$. 这个结构提示使用作差法来分离出通项 a_n. 当 $n \geq 2$ 时, 有 $T_{n-1} = a_1 + 3a_2 + \cdots + (2(n-1)-1)a_{n-1} = 2(n-1)$.

考察 T_n 与 T_{n-1} 的差:

$$\begin{aligned} T_n - T_{n-1} &= (a_1 + \cdots + (2n-1)a_n) - (a_1 + \cdots + (2n-3)a_{n-1}) \\ &= (2n-1)a_n \end{aligned}$$

另一方面,

$$T_n - T_{n-1} = 2n - 2(n-1) = 2$$

因此, 对于 $n \geq 2$, 有 $(2n-1)a_n = 2$, 即 $a_n = \dfrac{2}{2n-1}$.

接着, 需检验 $n = 1$ 的情况. 在原条件中令 $n = 1$, 可得 $1 \cdot a_1 = 2(1)$, 解得 $a_1 = 2$. 将 $n = 1$ 代入已得的公式, $a_1 = \dfrac{2}{2(1)-1} = 2$. 公式对 $n = 1$ 也成立. 故数列 $\{a_n\}$ 的通项公式为 $a_n = \dfrac{2}{2n-1}$.

(2) 设新数列为 $\{c_n\}$, 其通项为 $c_n = \dfrac{a_n}{2n+1}$. 将 a_n 的表达式代入, 得

$$c_n = \frac{2/(2n-1)}{2n+1} = \frac{2}{(2n-1)(2n+1)}$$

分母两因子之差为 $(2n+1) - (2n-1) = 2$, 这是一个典型的裂项求和结构.

$$c_n = \frac{1}{2n-1} - \frac{1}{2n+1}$$

令所求前 n 项和为 S_n, 得

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \left(\frac{1}{2k-1} - \frac{1}{2k+1} \right) \\ &= \left(\frac{1}{1} - \frac{1}{3} \right) + \left(\frac{1}{3} - \frac{1}{5} \right) + \cdots + \left(\frac{1}{2n-1} - \frac{1}{2n+1} \right) \end{aligned}$$

中间项全部抵消, 得到:

$$S_n = 1 - \frac{1}{2n+1} = \frac{2n}{2n+1}$$

例题 7.43 (2011 大纲卷) 设数列 $\{a_n\}$ 满足 $a_1 = 0$, 且 $\dfrac{1}{1-a_{n+1}} - \dfrac{1}{1-a_n} = 1$.

1. 求 $\{a_n\}$ 的通项公式;
2. 设 $b_n = \dfrac{1-\sqrt{a_{n+1}}}{\sqrt{n}}$, 记 $S_n = \sum_{k=1}^{n} b_k$, 求证: $S_n < 1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 (1) 观察给定的递推关系, 其形式暗示了换元法. 令新数列 $\{c_n\}$ 的通项为 $c_n = \dfrac{1}{1-a_n}$. 则原递推关系化为 $c_{n+1} - c_n = 1$. 故数列 $\{c_n\}$ 是一个公差为 1 的等差数列.

其首项为 $c_1 = \dfrac{1}{1-a_1} = \dfrac{1}{1-0} = 1$. 根据等差数列通项公式, $c_n = c_1 + (n-1)d = 1 + (n-1)\cdot 1 = n$.

由 $c_n = n$ 反解 a_n:

$$\frac{1}{1-a_n} = n \implies 1 - a_n = \frac{1}{n} \implies a_n = 1 - \frac{1}{n} = \frac{n-1}{n}$$

该公式即为 $\{a_n\}$ 的通项公式.

(2) 为了对 $\{b_n\}$ 求和, 先需化简其通项. 将 $a_{n+1} = \dfrac{(n+1)-1}{n+1} = \dfrac{n}{n+1}$ 代入 b_n 的表达式, 得

$$\sqrt{a_{n+1}} = \sqrt{\frac{n}{n+1}} = \frac{\sqrt{n}}{\sqrt{n+1}}$$

于是,

$$\begin{aligned} b_n &= \frac{1 - \frac{\sqrt{n}}{\sqrt{n+1}}}{\sqrt{n}} \\ &= \frac{(\sqrt{n+1} - \sqrt{n})/\sqrt{n+1}}{\sqrt{n}} \\ &= \frac{\sqrt{n+1} - \sqrt{n}}{\sqrt{n}\sqrt{n+1}} \\ &= \frac{\sqrt{n+1}}{\sqrt{n}\sqrt{n+1}} - \frac{\sqrt{n}}{\sqrt{n}\sqrt{n+1}} \\ &= \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} \end{aligned}$$

通项 b_n 就此分解为可以裂项相消的形式.

接着, 计算前 n 项和 S_n, 得

$$\begin{aligned} S_n &= \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} \left(\frac{1}{\sqrt{k}} - \frac{1}{\sqrt{k+1}} \right) \\ &= \left(\frac{1}{\sqrt{1}} - \frac{1}{\sqrt{2}} \right) + \left(\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{3}} \right) + \cdots + \left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} \right) \end{aligned}$$

中间项全部抵消, 得到:

$$S_n = \frac{1}{\sqrt{1}} - \frac{1}{\sqrt{n+1}} = 1 - \frac{1}{\sqrt{n+1}}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因为 n 为正整数, $n+1 \geq 2$, 所以 $\sqrt{n+1} > 0$, 从而 $\dfrac{1}{\sqrt{n+1}} > 0$. 因此, $S_n = 1 - \dfrac{1}{\sqrt{n+1}} < 1$. 证毕.

例题 7.44 (2006 全国 I 卷理) 设数列 $\{a_n\}$ 的前 n 项和为 S_n, 且满足 $S_n = \dfrac{4}{3}a_n - \dfrac{1}{3}\cdot 2^{n+1} + \dfrac{2}{3}$ $(n \in \mathbb{N}^*)$.

1. 求首项 a_1 与通项公式 a_n;
2. 设 $T_n = \dfrac{2^n}{S_n}$, 求证: $\sum_{i=1}^{n} T_i < \dfrac{3}{2}$.

解 (1) 首先求首项 a_1. 当 $n=1$ 时, 数列的前 1 项和即为其本身, 故有 $S_1 = a_1$. 将此条件代入已知关系式, 得

$$a_1 = \frac{4}{3}a_1 - \frac{1}{3}\cdot 2^{1+1} + \frac{2}{3}$$

$$a_1 = \frac{4}{3}a_1 - \frac{4}{3} + \frac{2}{3} \implies -\frac{1}{3}a_1 = -\frac{2}{3}$$

解得 $a_1 = 2$.

接下来求通项公式. 已知 S_n 与 a_n 的关系, 需建立一个只含数列项的递推关系. 通项与前 n 项和的最基本关系式 $a_n = S_n - S_{n-1}$, 当 $n \geq 2$ 时是实现这一目标的关键, 因为它能把关于前 n 项和的条件转化为相邻两项之间的递推关系.

对于 $n \geq 2$, 有

$$S_n = \frac{4}{3}a_n - \frac{1}{3}\cdot 2^{n+1} + \frac{2}{3}$$

$$S_{n-1} = \frac{4}{3}a_{n-1} - \frac{1}{3}\cdot 2^{n} + \frac{2}{3}$$

两式相减, 左侧为 $S_n - S_{n-1} = a_n$.

$$a_n = \left(\frac{4}{3}a_n - \frac{2}{3}\cdot 2^n\right) - \left(\frac{4}{3}a_{n-1} - \frac{1}{3}\cdot 2^n\right) + \frac{2}{3} - \frac{2}{3}$$

$$a_n = \frac{4}{3}a_n - \frac{4}{3}a_{n-1} - \frac{1}{3}(2^{n+1} - 2^n) = \frac{4}{3}a_n - \frac{4}{3}a_{n-1} - \frac{1}{3}\cdot 2^n$$

整理上式以得到 a_n 与 a_{n-1} 的关系:

$$-\frac{1}{3}a_n = -\frac{4}{3}a_{n-1} - \frac{1}{3}\cdot 2^n \implies a_n = 4a_{n-1} + 2^n \quad (n \geq 2)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这是一个一阶线性递推关系. 为求解此式, 可两边同除以 4^n, 将其转化为一个更简单的形式.

$$\frac{a_n}{4^n}=\frac{4a_{n-1}}{4^n}+\frac{2^n}{4^n}\implies\frac{a_n}{4^n}=\frac{a_{n-1}}{4^{n-1}}+\left(\frac{1}{2}\right)^n$$

令 $c_n=\dfrac{a_n}{4^n}$, 则 $c_1=\dfrac{a_1}{4}=\dfrac{1}{2}$, 且 $c_n-c_{n-1}=(\frac{1}{2})^n$. 通过累加法,

$$c_n=c_1+\sum_{k=2}^{n}\left(\frac{1}{2}\right)^k=\frac{1}{2}+\frac{(\frac{1}{2})^2(1-(\frac{1}{2})^{n-1})}{1-\frac{1}{2}}=\frac{1}{2}+\frac{1}{2}\left(1-\frac{1}{2^{n-1}}\right)=1-\frac{1}{2^n}$$

因此 $a_n=4^nc_n=4^n\left(1-\frac{1}{2^n}\right)=4^n-2^n$. 检验 $n=1$ 时, $a_1=4^1-2^1=2$, 与所求一致. 故通项公式为 $a_n=4^n-2^n$.

(2) 先需得到 S_n 的显式表达式.

$$S_n=\sum_{k=1}^{n}a_k=\sum_{k=1}^{n}(4^k-2^k)=\sum_{k=1}^{n}4^k-\sum_{k=1}^{n}2^k$$

利用等比数列求和公式:

$$S_n=\frac{4(4^n-1)}{4-1}-\frac{2(2^n-1)}{2-1}=\frac{4^{n+1}-4}{3}-(2^{n+1}-2)=\frac{4^{n+1}-3\cdot2^{n+1}+2}{3}$$

S_n 的表达式看似复杂, 但若将其中的指数项看作一个整体, 它的结构便清晰起来. 将 2^{n+1} 视为整体变量后, 分子可直接因式分解, 从而为后续裂项提供结构. 令 $x=2^{n+1}$, 则 $4^{n+1}=(2^2)^{n+1}=(2^{n+1})^2=x^2$. 于是分子变为 $x^2-3x+2=(x-1)(x-2)$.

$$S_n=\frac{(2^{n+1}-1)(2^{n+1}-2)}{3}=\frac{2(2^{n+1}-1)(2^n-1)}{3}$$

现在可计算 T_n 的通项, 得

$$T_n=\frac{2^n}{S_n}=\frac{3\cdot2^n}{2(2^{n+1}-1)(2^n-1)}=\frac{3\cdot2^{n-1}}{(2^{n+1}-1)(2^n-1)}$$

分母是结构相似的项的乘积, 这强烈暗示了裂项相消的可能性. 尝试将其分解为 $\dfrac{A}{2^n-1}-\dfrac{B}{2^{n+1}-1}$ 的形式.

$$\frac{A(2^{n+1}-1)-B(2^n-1)}{(2^n-1)(2^{n+1}-1)}=\frac{(2A-B)2^n-(A-B)}{(2^n-1)(2^{n+1}-1)}$$

比较分子 $3\cdot2^{n-1}=\frac{3}{2}\cdot2^n$, 得到 $2A-B=\frac{3}{2}$ 且 $A-B=0$, 解得 $A=B=\frac{3}{2}$. 故

$$T_n=\frac{3}{2}\left(\frac{1}{2^n-1}-\frac{1}{2^{n+1}-1}\right)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这是一个标准的裂项形式. 求其前 n 项和:

$$\begin{aligned}\sum_{i=1}^{n} T_i &= \frac{3}{2}\sum_{i=1}^{n}\left(\frac{1}{2^i-1}-\frac{1}{2^{i+1}-1}\right)\\ &= \frac{3}{2}\left[\left(\frac{1}{2^1-1}-\frac{1}{2^2-1}\right)+\cdots+\left(\frac{1}{2^n-1}-\frac{1}{2^{n+1}-1}\right)\right]\\ &= \frac{3}{2}\left(1-\frac{1}{2^{n+1}-1}\right)\end{aligned}$$

由于 $n\geq 1$, 则 $n+1\geq 2$, $2^{n+1}-1\geq 3>0$, 所以 $\dfrac{1}{2^{n+1}-1}>0$. 因此,

知乎@余命数

$$\sum_{i=1}^{n} T_i = \frac{3}{2}-\frac{3}{2(2^{n+1}-1)}<\frac{3}{2}$$

证毕.

例题 7.45 (2023 年 4 月杭州、长沙、郑州、西安部分名校联考) 已知正项数列 $\{a_n\}$ 的前 n 项和为 S_n, 且 $a_1=1$, $a_n=\sqrt{S_n}+\sqrt{S_{n-1}}(n\in\mathbb{N}^*$ 且 $n\geq 2)$.

1. 求数列 $\{a_n\}$ 的通项公式;
2. 设数列 $\left\{\dfrac{a_{n+2}}{2^n a_n a_{n+1}}\right\}$ 的前 n 项和为 T_n, 求证:$T_n<1$.

解 (1) 给定的递推关系式 $a_n=\sqrt{S_n}+\sqrt{S_{n-1}}$ 结构特殊, 直接处理较为困难. 观察此式的结构, 它与平方差公式 $(\sqrt{x}-\sqrt{y})(\sqrt{x}+\sqrt{y})=x-y$ 有着紧密的联系. 已知, 数列通项与其前 n 项和之间存在基本关系 $a_n=S_n-S_{n-1}$, 当 $n\geq 2$ 时. 这启发将 a_n 替换为 $(\sqrt{S_n}-\sqrt{S_{n-1}})(\sqrt{S_n}+\sqrt{S_{n-1}})$.

对于 $n\geq 2$, 有

知乎@余命数

$$a_n=S_n-S_{n-1}=(\sqrt{S_n}-\sqrt{S_{n-1}})(\sqrt{S_n}+\sqrt{S_{n-1}})$$

将已知条件 $a_n=\sqrt{S_n}+\sqrt{S_{n-1}}$ 代入上式, 得到

$$a_n=(\sqrt{S_n}-\sqrt{S_{n-1}})a_n$$

由于 $\{a_n\}$ 为正项数列, 故 $a_n>0$, 等式两边可以约去 a_n.

$$\sqrt{S_n}-\sqrt{S_{n-1}}=1\quad(n\geq 2)$$

故数列 $\{\sqrt{S_n}\}$ 从第二项起是一个公差为 1 的等差数列. 这一步把含根号的递推关系转化为等差数列问题, 是求出 S_n 与 a_n 的关键转折.

计算其首项:$\sqrt{S_1}=\sqrt{a_1}=\sqrt{1}=1$. 因此, 数列 $\{\sqrt{S_n}\}$ 是一个首项为 1, 公差为 1 的等差数列. 其通项为 $\sqrt{S_n}=1+(n-1)\cdot 1=n$. 故 $S_n=n^2$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当 $n \geq 2$ 时,

$$a_n = S_n - S_{n-1} = n^2 - (n-1)^2 = 2n - 1$$

当 $n = 1$ 时, $a_1 = S_1 = 1^2 = 1$. 将 $n = 1$ 代入 $2n-1$ 得 1, 公式同样成立. 所以, $\{a_n\}$ 的通项公式为 $a_n = 2n - 1$.

(2) 设所求数列为 $\{b_n\}$, 其通项为 $b_n = \dfrac{a_{n+2}}{2^n a_n a_{n+1}}$. 将 $a_n = 2n-1$ 代入, 得

$$a_{n+1} = 2(n+1) - 1 = 2n + 1$$

$$a_{n+2} = 2(n+2) - 1 = 2n + 3$$

于是,

$$b_n = \frac{2n+3}{2^n(2n-1)(2n+1)}$$

此通项结构复杂, 直接求和十分困难. 现将其分解为可以裂项相消的形式. 经观察, 有恒等式

$$b_n = \frac{1}{2^{n-1}(2n-1)} - \frac{1}{2^n(2n+1)}$$

(请自行通分验证此等式为何成立.) 这个形式是 $f(k) - f(k+1)$ 的理想裂项结构, 其中 $f(k) = \dfrac{1}{2^{k-1}(2k-1)}$.

计算前 n 项和 T_n, 得

$$\begin{aligned} T_n &= \sum_{k=1}^{n} b_k = \sum_{k=1}^{n} \left(\frac{1}{2^{k-1}(2k-1)} - \frac{1}{2^k(2k+1)} \right) \\ &= \left(\frac{1}{2^0 \cdot 1} - \frac{1}{2^1 \cdot 3} \right) + \left(\frac{1}{2^1 \cdot 3} - \frac{1}{2^2 \cdot 5} \right) \\ &\quad + \cdots + \left(\frac{1}{2^{n-1}(2n-1)} - \frac{1}{2^n(2n+1)} \right) \end{aligned}$$

中间项逐一抵消, 最终只剩下首项与末项:

$$T_n = \frac{1}{1} - \frac{1}{2^n(2n+1)} = 1 - \frac{1}{2^n(2n+1)}$$

因为 n 是正整数, 所以 $2^n > 0$ 且 $2n+1 > 0$, 即 $\dfrac{1}{2^n(2n+1)}$ 是一个正数. 因此, $T_n = 1 - (\text{一个正数}) < 1$. 证毕.

构造裂项还有更系统的做法.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.8.2.1 待定系数法

这个方法根据通项 a_k 的结构**猜测**出差分原函数 $f(k)$ 的代数形式, 再用待定系数法精确求解.

目标是把 a_k 表示为 $f(k)-f(k+1)$ 或其变体. 考虑一个比 $\frac{1}{k(k+1)}$ 更复杂的例子.

例题 7.46　求和 $S_n=\sum_{k=1}^{n}\frac{1}{k(k+1)(k+2)}$.

解 通项 $a_k=\frac{1}{k(k+1)(k+2)}$, 需找一个函数 $f(k)$ 使得 $a_k=f(k)-f(k+1)$. a_k 的分母是三个连续项的乘积, 故猜测原函数 $f(k)$ 的分母少一项, 即只含两个连续因子.

设 $f(k)=\frac{A}{k(k+1)}$, 其中 A 是待定常数.

计算 $f(k)-f(k+1)$ 并令其等于 a_k, 得

$$\begin{aligned}f(k)-f(k+1)&=\frac{A}{k(k+1)}-\frac{A}{(k+1)(k+2)}\\&=A\left(\frac{1}{k(k+1)}-\frac{1}{(k+1)(k+2)}\right)\\&=A\cdot\frac{(k+2)-k}{k(k+1)(k+2)}\\&=\frac{2A}{k(k+1)(k+2)}\end{aligned}$$

欲使该结果等于 $a_k=\frac{1}{k(k+1)(k+2)}$. 通过比较分子, 得 $2A=1$, 解得 $A=\frac{1}{2}$.

由此得到通项分解

$$a_k=\frac{1}{2}\left(\frac{1}{k(k+1)}-\frac{1}{(k+1)(k+2)}\right)$$

这是一个标准的裂项形式, 其中 $f(k)=\frac{1}{2k(k+1)}$.

现在进行求和.

$$\begin{aligned}S_n&=\sum_{k=1}^{n}\frac{1}{2}\left(\frac{1}{k(k+1)}-\frac{1}{(k+1)(k+2)}\right)\\&=\frac{1}{2}\left[\left(\frac{1}{1\cdot2}-\frac{1}{2\cdot3}\right)+\left(\frac{1}{2\cdot3}-\frac{1}{3\cdot4}\right)\right.\\&\qquad\left.+\cdots+\left(\frac{1}{n(n+1)}-\frac{1}{(n+1)(n+2)}\right)\right]\\&=\frac{1}{2}\left(\frac{1}{2}-\frac{1}{(n+1)(n+2)}\right)\\&=\frac{1}{4}-\frac{1}{2(n+1)(n+2)}=\frac{n(n+3)}{4(n+1)(n+2)}\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.8.2.2 差分与和分 *

裂项相消法的背后是离散化的微积分. 如果微分与积分互为逆运算, 那么数列中也有对应的运算, 即差分与和分.

定义 7.10 (差分算子) 设 $f(k)$ 是任意函数. 我们定义其(前向)**差分** $\Delta f(k)$ 为

$$\Delta f(k) = f(k+1) - f(k)$$

有了差分算子, 裂项求和就纳入了差分与和分的框架. 求和 $\sum_{k=m}^{n}(f(k+1)-f(k))$ 正是 $\sum_{k=m}^{n}\Delta f(k)$.

由此, **任何求和问题 $\sum_{k=1}^{n} a_k$ 都等价于寻找一个函数 $F(k)$, 称为 a_k 的反差分, 使得 $\Delta F(k) = F(k+1)-F(k) = a_k$.** 一旦找到反差分 $F(k)$, 求和只需计算端点, 得

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n} \Delta F(k) = F(n+1) - F(1)$$

寻找反差分 $F(k)$ 的过程, 也就是构造裂项的过程, 这一思路可以推广到更广的通项.

寻找 $F(k)$ 时, 对多项式型的通项可再次运用待定系数法.

例题 7.47 运用反差分的思想, 推导平方和公式 $\sum_{k=1}^{n} k^2$.

解 现求 $S_n = \sum_{k=1}^{n} k^2$. 通项 $a_k = k^2$. 专注于寻找一个函数 $F(k)$ 使得 $\Delta F(k) = F(k+1) - F(k) = k^2$.

差分运算会使多项式的次数降低 1. 因此, 一个二次多项式 k^2 的反差分 $F(k)$ 必然是一个三次多项式. 设 $F(k) = Ak^3 + Bk^2 + Ck$.(常数项在差分中会消失, 故可省略).

计算其差分 $\Delta F(k)$, 得

$$\begin{aligned}\Delta F(k) &= [A(k+1)^3 + B(k+1)^2 + C(k+1)] - [Ak^3 + Bk^2 + Ck] \\ &= A((k+1)^3 - k^3) + B((k+1)^2 - k^2) + C((k+1) - k) \\ &= A(3k^2 + 3k + 1) + B(2k+1) + C \\ &= (3A)k^2 + (3A + 2B)k + (A + B + C)\end{aligned}$$

要求 $\Delta F(k) = k^2 = 1 \cdot k^2 + 0 \cdot k + 0$. 通过比较等式两边 k 的同次幂系数, 得到一个线性方程组.

$$k^2\text{的系数}: 3A = 1$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$k\text{的系数:}3A+2B=0$$

$$\text{常数项:}A+B+C=0$$

知乎@余命数

解这个方程组. 由方程一得 $A=\frac{1}{3}$. 代入方程二得 $3(\frac{1}{3})+2B=0 \implies 1+2B=0 \implies B=-\frac{1}{2}$. 代入方程三得 $\frac{1}{3}-\frac{1}{2}+C=0 \implies C=\frac{1}{6}$.

已找到反差分函数.

$$F(k)=\frac{1}{3}k^3-\frac{1}{2}k^2+\frac{1}{6}k$$

根据离散微积分基本定理, 平方和为

$$S_n=\sum_{k=1}^{n}k^2=F(n+1)-F(1)$$

计算 $F(1)=\frac{1}{3}-\frac{1}{2}+\frac{1}{6}=\frac{2-3+1}{6}=0$. 计算 $F(n+1)$, 得

$$\begin{aligned}F(n+1)&=\frac{1}{3}(n+1)^3-\frac{1}{2}(n+1)^2+\frac{1}{6}(n+1)\\&=\frac{n+1}{6}[2(n+1)^2-3(n+1)+1]\\&=\frac{n+1}{6}[2(n^2+2n+1)-3n-3+1]\\&=\frac{n+1}{6}[2n^2+4n+2-3n-2]\\&=\frac{n+1}{6}[2n^2+n]\\&=\frac{n(n+1)(2n+1)}{6}\end{aligned}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

因此, $S_n=F(n+1)-0=\frac{n(n+1)(2n+1)}{6}$.

7.8.3 Gosper 裂项法 *

裂项相消法在特定结构的数列求和问题上成效显著. 无论部分分式分解还是阶乘恒等式, 都是把通项 a_k 改写为 $F(k)-F(k-1)$ 的形式. 然而, 这些技巧在很大程度上依赖于观察、经验与代数变形.

我们希望有一种系统性的方法, 也就是一种**算法**, 能够机械地判断一个给定的通项 a_k 是否存在一个简单的反差分 $F(k)$, 并在存在时直接把它算出来.

对于数学中最为常见和重要的一类序列, **超几何序列**, 美国数学家 R. William Gosper 在 20 世纪 70 年代给出了一个完整的解决方案.Gosper 算法将寻找裂项公式这一创造性活动, 转化为了一套可以由计算机执行的确定程序.

如今, 在 Mathematica 或 Maple 等计算机代数系统中输入复杂求和指令时, 其背后往往就是 Gosper 算法或其后继算法在运行.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

7.8.3.1 超几何项

Gosper 算法的适用范围, 是被称为超几何项的序列.

定义 7.11 (超几何项) 一个序列 $\{a_k\}$ 如果其相邻项之比 $\dfrac{a_{k+1}}{a_k}$ 是一个关于 k 的**有理函数**, 即两个关于 k 的多项式之比, 则称 $\{a_k\}$ 是一个**超几何项**或**超几何序列**.

这个定义涵盖了我们常遇到的大部分基本构件.

- 常数与幂函数.$a_k=c$ 时 $\dfrac{a_{k+1}}{a_k}=1$;$a_k=k^m$ 时 $\dfrac{a_{k+1}}{a_k}=\left(\dfrac{k+1}{k}\right)^m$.
- 指数函数.$a_k=q^k$ 时 $\dfrac{a_{k+1}}{a_k}=q$.
- 阶乘.$a_k=k!$ 时 $\dfrac{a_{k+1}}{a_k}=k+1$.
- 二项式系数.$a_k=\binom{n}{k}$ 时 $\dfrac{a_{k+1}}{a_k}=\dfrac{n-k}{k+1}$.

更重要的是, 超几何项的集合在乘法、除法、平移等运算下是封闭的. 即, 由以上基本构件通过乘除组合而成的复杂项, 例如 $a_k=\dfrac{(k^2+1)k!}{2^k\binom{2k}{k}}$, 仍然是超几何项.

7.8.3.2 Gosper 的洞察与方法

面对一个超几何求和 $\sum a_k$, 现寻找其反差分 $F(k)$, 使得 $a_k=F(k)-F(k-1)$. Gosper 的第一个关键洞察如下.

定理 7.5 (Gosper 判别方程) 设非零超几何项 a_k 的相邻项比值为 $r(k)=\dfrac{a_k}{a_{k-1}}$. 若存在有理函数 $y(k)$ 满足

$$r(k)y(k)-y(k-1)=r(k),$$

则 $F(k)=y(k)a_k$ 是 a_k 的一个超几何反差分, 即 $F(k)-F(k-1)=a_k$.

证明 由 $F(k)=y(k)a_k$ 得

$$F(k)-F(k-1)=y(k)a_k-y(k-1)a_{k-1}=\left(y(k)-\frac{y(k-1)}{r(k)}\right)a_k.$$

判别方程 $r(k)y(k)-y(k-1)=r(k)$ 等价于 $y(k)-\frac{y(k-1)}{r(k)}=1$, 因而 $F(k)-F(k-1)=a_k$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

又因为 a_k 是超几何项, $r(k)$ 为有理函数; $y(k)$ 为有理函数时,

$$\frac{F(k)}{F(k-1)}=\frac{y(k)}{y(k-1)}\,r(k)$$

也是关于 k 的有理函数(在相应表达式有意义处). 因此 $F(k)$ 也是超几何项. □

Gosper 完整算法进一步说明, 对超几何项求超几何反差分时, 搜索这样的有理函数 $y(k)$ 不会漏掉可求解的情形. 因而这个判别方程大幅缩小了对未知函数 $F(k)$ 的搜寻范围.

将 $F(k)=y(k)a_k$ 代入裂项关系式 $a_k=F(k)-F(k-1)$, 得

$$a_k=y(k)a_k-y(k-1)a_{k-1}$$

在 $a_k\neq 0$ 的情况下, 两边同除以 a_k, 得

$$1=y(k)-y(k-1)\frac{a_{k-1}}{a_k}$$

令 $r(k)=\dfrac{a_k}{a_{k-1}}$ 为相邻项的比值(这是一个已知的有理函数), 则上式变为

$$r(k)=r(k)y(k)-y(k-1)$$

这是一个关于未知有理函数 $y(k)$ 的函数方程.Gosper 的完整算法给出了求解这个方程的系统步骤, 要点是证明 $y(k)$ 必须具有一种特殊的多项式结构, 从而将问题转化为一个求解多项式系数的线性方程组问题.

用一个例子体会这种待定系数求解的思想.

例题 7.48　求和 $S_n=\sum_{k=1}^{n}(k^2+1)k!$.

解　通项 $a_k=(k^2+1)k!$ 是一个超几何项. 遵循 Gosper 的思想, 尝试寻找其反差分 $F(k)$, 并猜测 $F(k)$ 的结构与 a_k 相似, 设 $F(k)=P(k)\cdot k!$, 其中 $P(k)$ 是待定多项式.

计算 $F(k)-F(k-1)$, 得

$$\begin{aligned}F(k)-F(k-1)&=P(k)k!-P(k-1)(k-1)!\\&=P(k)\cdot k\cdot(k-1)!-P(k-1)(k-1)!\\&=[kP(k)-P(k-1)](k-1)!\end{aligned}$$

欲使该结果等于 $a_k=(k^2+1)k!=k(k^2+1)(k-1)!$.

由此得到关于未知多项式 $P(k)$ 的方程

$$kP(k)-P(k-1)=k(k^2+1)=k^3+k$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

接下来，比较等式两边多项式的次数以确定 $P(k)$ 的次数. 设 $\deg(P) = d$，则 $P(k) = c_d k^d + \ldots$ ($c_d \neq 0$). 方程左侧的最高次项来自 $kP(k)$, 其次数为 $d+1$. 方程右侧的最高次项是 k^3, 次数为 3. 因此, 必然有 $d+1=3$, 解得 $d=2$. 故 $P(k)$ 是一个二次多项式.

设 $P(k) = Ak^2 + Bk + C$. 代入方程 $kP(k) - P(k-1) = k^3 + k$:

$$\begin{aligned}\text{LHS} &= k(Ak^2+Bk+C) - [A(k-1)^2 + B(k-1) + C] \\ &= (Ak^3 + Bk^2 + Ck) - [A(k^2-2k+1) + B(k-1) + C] \\ &= Ak^3 + (B-A)k^2 + (C+2A-B)k + (-A+B-C)\end{aligned}$$

$\text{RHS} = 1 \cdot k^3 + 0 \cdot k^2 + 1 \cdot k + 0$.

比较等式两边各项系数, 得到一个线性方程组.

k^3系数: $A = 1$

k^2系数: $B - A = 0 \implies B = A = 1$

k^1系数: $C + 2A - B = 1 \implies C + 2(1) - 1 = 1 \implies C = 0$

k^0系数: $-A + B - C = 0 \implies -1 + 1 - 0 = 0$.(验证成立)

解得 $P(k) = 1 \cdot k^2 + 1 \cdot k + 0 = k^2 + k = k(k+1)$.

因此, 找到的反差分为

$$F(k) = P(k)k! = k(k+1)k! = k \cdot (k+1)!$$

于是求和只剩端点值的计算.

$$\begin{aligned}S_n = \sum_{k=1}^{n} a_k &= \sum_{k=1}^{n} [F(k) - F(k-1)] \\ &= F(n) - F(0) \\ &= n(n+1)! - 0 \cdot (0+1)! \\ &= n(n+1)!\end{aligned}$$

例题 7.49 求和 $S_n = \sum_{k=1}^{n} k \cdot k!$.

解 通项 $a_k = k \cdot k!$ 是超几何项. 寻找其反差分 $F(k)$, 使得 $a_k = F(k) - F(k-1)$.

一个自然的猜测是 $F(k)$ 与 $k!$ 具有相同的结构, 形如 $F(k) = P(k) \cdot k!$, 其中 $P(k)$ 是一个待定的多项式. 计算差分, 得

$$F(k) - F(k-1) = P(k)k! - P(k-1)(k-1)!$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= [kP(k) - P(k-1)](k-1)!$$

知乎@余命数

欲使该结果等于 $a_k = k \cdot k! = k^2(k-1)!$.

由此得到关于 $P(k)$ 的方程

$$kP(k) - P(k-1) = k^2$$

比较等式两边多项式的次数. 若 $\deg(P) = d$, 则左侧最高次项来自 $kP(k)$, 次数为 $d+1$. 右侧次数为 2. 故 $d+1=2 \implies d=1$.$P(k)$ 是一个一次多项式.

设 $P(k) = Ak + B$. 代入方程, 得

$$k(Ak+B) - [A(k-1)+B] = k^2$$

$$Ak^2 + Bk - Ak + A - B = k^2$$

$$Ak^2 + (B-A)k + (A-B) = 1 \cdot k^2 + 0 \cdot k + 0$$

知乎@余命数

比较系数. $A=1$. $B-A=0 \implies B=1$. $A-B=0 \implies 1-1=0$, 验证成立.

解得 $P(k) = k+1$. 因此, 反差分是 $F(k) = (k+1)k! = (k+1)!$.

求和.

$$S_n = \sum_{k=1}^{n} a_k = F(n) - F(0)$$

$$= (n+1)! - (0+1)!$$

知乎@余命数

$$= (n+1)! - 1$$

例题 7.50　求和 $S_n = \sum_{k=1}^{n} k \cdot 2^k$.

知乎@余命数

解　通项 $a_k = k \cdot 2^k$ 是超几何项. 猜测其反差分具有形式 $F(k) = P(k) \cdot 2^k$. 计算差分, 得

知乎@余命数

$$F(k) - F(k-1) = P(k)2^k - P(k-1)2^{k-1}$$

$$= [2P(k) - P(k-1)]2^{k-1}$$

欲使该结果等于 $a_k = k \cdot 2^k = 2k \cdot 2^{k-1}$.

由此得到关于 $P(k)$ 的方程

$$2P(k) - P(k-1) = 2k$$

比较次数可知 $\deg(P) = 1$. 设 $P(k) = Ak + B$. 代入方程, 得

$$2(Ak+B) - [A(k-1)+B] = 2k$$

$$2Ak + 2B - Ak + A - B = 2k$$

$$Ak + (A + B) = 2k + 0$$

比较系数. $A = 2$. $A + B = 0 \implies 2 + B = 0 \implies B = -2$.

解得 $P(k) = 2k - 2$. 因此, 反差分是 $F(k) = (2k - 2)2^k = (k - 1)2^{k+1}$.

求和.

$$\begin{aligned} S_n = \sum_{k=1}^{n} a_k &= F(n) - F(0) \\ &= (n-1)2^{n+1} - (0-1)2^{0+1} \\ &= (n-1)2^{n+1} - (-1) \cdot 2 \\ &= (n-1)2^{n+1} + 2 \end{aligned}$$

这个结果与使用错位相减法得到的结果完全一致.

例题 7.51 求和 $S_n = \sum_{k=1}^{n} k(k+1)(k+2)$.

解 通项 $a_k = k(k+1)(k+2)$ 是一个三次多项式. 寻找其反差分 $F(k)$ 使得 $F(k) - F(k-1) = a_k$. 差分运算会使多项式的次数减一, 故 $F(k)$ 必然是一个四次多项式.

这种上升阶乘幂的形式 $k^{\overline{m}} = k(k+1)\ldots(k+m-1)$ 具有简单的差分性质,故猜测反差分 $F(k)$ 具有更高一阶的上升阶乘幂形式,即 $F(k) = C \cdot k(k+1)(k+2)(k+3)$.

计算其差分, 得

$$\begin{aligned} F(k) - F(k-1) &= C \cdot k(k+1)(k+2)(k+3) - C \cdot (k-1)k(k+1)(k+2) \\ &= C \cdot k(k+1)(k+2)\left[(k+3) - (k-1)\right] \quad \text{（提取公因子 } C \cdot k(k+1)(k+2)\text{）} \\ &= C \cdot k(k+1)(k+2) \cdot 4. \end{aligned}$$

令此式等于 $a_k = k(k+1)(k+2)$, 比较可知 $4C = 1 \implies C = \dfrac{1}{4}$.

因此, 反差分是 $F(k) = \dfrac{1}{4}k(k+1)(k+2)(k+3)$.

求和.

$$\begin{aligned} S_n = \sum_{k=1}^{n} a_k &= F(n) - F(0) \\ &= \frac{1}{4}n(n+1)(n+2)(n+3) - \frac{1}{4} \cdot 0 \cdot (1) \cdot (2) \cdot (3) \\ &= \frac{n(n+1)(n+2)(n+3)}{4} \end{aligned}$$

这种方法比将 a_k 展开为 $k^3 + 3k^2 + 2k$ 再利用幂和公式计算要简洁得多.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.52　求和 $S_n = \sum_{k=1}^{n} k^2$.

解 通项 $a_k = k^2$ 是一个二次多项式. 寻找其反差分 $F(k)$ 使得 $a_k = F(k) - F(k-1)$. 由于差分运算会使多项式的次数降低 1, 一个二次多项式的反差分必然是三次多项式. 设 $F(k) = Ak^3 + Bk^2 + Ck$(常数项在差分中会消失, 故可省略).

计算差分 $F(k) - F(k-1)$, 得

$$\begin{aligned} F(k) - F(k-1) &= (Ak^3 + Bk^2 + Ck) - [A(k-1)^3 + B(k-1)^2 + C(k-1)] \\ &= A(3k^2 - 3k + 1) + B(2k - 1) + C \\ &= 3Ak^2 + (2B - 3A)k + (A - B + C) \end{aligned}$$

欲使该结果等于 $a_k = 1 \cdot k^2 + 0 \cdot k + 0$.

比较系数可得

$$3A = 1 \implies A = 1/3$$

$$2B - 3A = 0 \implies 2B = 3(1/3) = 1 \implies B = 1/2$$

$$A - B + C = 0 \implies C = B - A = 1/2 - 1/3 = 1/6$$

因此, 反差分是 $F(k) = \dfrac{1}{3}k^3 + \dfrac{1}{2}k^2 + \dfrac{1}{6}k = \dfrac{k(k+1)(2k+1)}{6}$.

求和.

$$\begin{aligned} S_n = \sum_{k=1}^{n} a_k &= F(n) - F(0) \\ &= \frac{n(n+1)(2n+1)}{6} - 0 \\ &= \frac{n(n+1)(2n+1)}{6} \end{aligned}$$

例题 7.53　求和 $S_n = \sum_{k=1}^{n} \dfrac{k}{(k+1)!}$.

解 设 $F(k) = \dfrac{C}{(k+1)!}$. 计算差分, 得

$$\begin{aligned} F(k) - F(k-1) &= \frac{C}{(k+1)!} - \frac{C}{k!} \\ &= \frac{C - C(k+1)}{(k+1)!} \\ &= \frac{-Ck}{(k+1)!} \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

欲使该结果等于 $a_k=\dfrac{k}{(k+1)!}$. 比较可知 $-C=1 \implies C=-1$.

因此, 反差分是 $F(k)=-\dfrac{1}{(k+1)!}$.

求和.

$$\begin{aligned} S_n=\sum_{k=1}^{n} a_k &= F(n)-F(0) \\ &= -\frac{1}{(n+1)!}-\left(-\frac{1}{(0+1)!}\right) \\ &= 1-\frac{1}{(n+1)!} \end{aligned}$$

知乎@余命数

例题 7.54 求和 $S_n=\sum_{k=1}^{n}\frac{1}{(2k-1)(2k+1)}$.

解 通项 $a_k=\dfrac{1}{(2k-1)(2k+1)}$, 寻找其反差分 $F(k)$.

观察通项分母的结构, 猜测反差分 $F(k)$ 的分母可能只包含其中一个因子, 设 $F(k)=\dfrac{C}{2k+1}$.

计算差分 $F(k)-F(k-1)$, 得

$$\begin{aligned} F(k)-F(k-1) &= \frac{C}{2k+1}-\frac{C}{2(k-1)+1} \\ &= \frac{C}{2k+1}-\frac{C}{2k-1} \\ &= C\left(\frac{(2k-1)-(2k+1)}{(2k+1)(2k-1)}\right) \\ &= \frac{-2C}{(2k-1)(2k+1)} \end{aligned}$$

欲使该结果等于 $a_k=\dfrac{1}{(2k-1)(2k+1)}$. 比较分子可知 $-2C=1 \implies C=-1/2$.

因此, 反差分是 $F(k)=-\dfrac{1}{2(2k+1)}$.

求和.

$$\begin{aligned} S_n=\sum_{k=1}^{n} a_k &= F(n)-F(0) \\ &= -\frac{1}{2(2n+1)}-\left(-\frac{1}{2(2\cdot 0+1)}\right) \\ &= -\frac{1}{2(2n+1)}+\frac{1}{2} \\ &= \frac{1}{2}\left(1-\frac{1}{2n+1}\right) \end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{1}{2} \cdot \frac{2n}{2n+1} = \frac{n}{2n+1}$$

此方法提供了一种绕过部分分式分解, 直接获得裂项公式的系统途径.

待定系数裂项法可按下列次序推进. 先观察通项 a_k 的代数结构, 猜测其反差分 $F(k)$ 可能的形式. 关键原则是, $F(k)$ 中非多项式的部分应与 a_k 保持一致, 而多项式部分的次数或系数待定.

设 $P(k), Q(k)$ 为关于 k 的多项式. 常见猜想结构如下.

- 若 $a_k = P(k)$(多项式), 则猜想 $F(k) = Q(k)$, 其中 $\deg(Q) = \deg(P) + 1$.
- 若 $a_k = P(k)q^k$(差比数列), 则猜想 $F(k) = Q(k)q^k$.
- 若 $a_k = P(k)k!$(阶乘相关), 则猜想 $F(k) = Q(k)k!$ 或 $Q(k)(k-1)!$.
- 若 a_k 为分式, 则猜想 $F(k)$ 为一个结构更简单的分式.

写出核心关系式 $F(k) - F(k-1) = a_k$, 并将带有待定系数的猜想形式代入此方程. 随后进入纯代数计算.

- 对代入后的方程化简, 通常需要通分或提取公因子, 如 q^{k-1} 或 $(k-1)!$.
- 化简后, 方程两边会呈现出关于 k 的多项式相等的形式.
- 通过比较等式两边多项式的最高次项, 确定待定多项式 $Q(k)$ 的次数.
- 设出 $Q(k)$ 的一般形式, 如 $Ak^2 + Bk + C$, 然后通过比较等式两边 k 的所有同次幂系数, 建立一个关于待定系数的线性方程组.
- 解这个方程组, 求出所有待定系数的值.

将已求出的系数代回, 得到确定的反差分函数 $F(k)$. 利用裂项求和的最终公式计算总和. 如果求和从 $k = 1$ 开始, 则

$$S_n = \sum_{k=1}^{n} a_k = F(n) - F(0)$$

注意仔细计算 $F(n)$ 和 $F(0)$ 的值.

7.8.4 错位相减法

我们用错位相减法处理**差比数列**求和. 差比数列的通项形如 $c_n = a_n b_n$, 其中 $\{a_n\}$ 是等差数列, $\{b_n\}$ 是等比数列.

求和时模仿等比数列求和公式的推导过程. 设等比部分的公比为 q.

1. 写出和式 $S_n = \sum_{k=1}^{n} a_k b_k$.
2. 将整个等式两边同乘以公比 q, 得到 qS_n.
3. 将 S_n 与 qS_n 的表达式对齐项错位相减, 得到 $(1-q)S_n$.
4. 相减后的结果, 通常会变成一个简单的等比数列之和, 外加首末两个独立项.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个过程通过乘公比再作差, 把差比数列化简为一个纯粹的等比数列.

例题 7.55 求和 $S_n=\sum_{k=1}^{n} k\cdot x^k\ (x\neq 1)$.

解 这是由等差数列 $\{k\}$ 与等比数列 $\{x^k\}$ 构成的差比数列. 先写出和式

知乎@余命数

$$S_n=1\cdot x+2\cdot x^2+3\cdot x^3+\cdots+n\cdot x^n$$

两边同乘以公比 x, 得

$$xS_n=\qquad 1\cdot x^2+2\cdot x^3+\cdots+(n-1)\cdot x^n+n\cdot x^{n+1}$$

两式相减.

$$\begin{aligned}(1-x)S_n&=x+(2-1)x^2+(3-2)x^3+\cdots+(n-(n-1))x^n-n\cdot x^{n+1}\\&=(x+x^2+x^3+\cdots+x^n)-n\cdot x^{n+1}\end{aligned}$$

知乎@余命数

括号内是一个首项为 x, 公比为 x 的等比数列之和. 知乎@余命数

$$(1-x)S_n=\frac{x(1-x^n)}{1-x}-nx^{n+1}$$

知乎@余命数

两边同除以 $(1-x)$, 得

$$S_n=\frac{x(1-x^n)}{(1-x)^2}-\frac{nx^{n+1}}{1-x}$$

7.8.5 倒序相加法

当数列的和式具有某种对称性 (典型特征是首末项之和、第二项与倒数第二项之和……遵循统一的规律) 时, 可把和式 S_n 的各项顺序颠倒, 得到另一个表达式, 再将两个表达式按位相加. 如果每一对应项的和都是常数或简单形式, 问题就得以简化. 高斯求和等差数列是此方法的经典应用.

阿贝尔变换与离散分部

知乎@余命数

知乎@余命数

注 7.7
裂项求和与分部积分在形式上对应. 前者对差分求和, 后者对导数积分.

7.8.6 阿贝尔 (Abel) 变换 *

知乎@余命数

在数列求和中常会遇到形如 $\sum_{k=1}^{n} a_kb_k$ 的乘积和. 错位相减法处理等差乘以等比这一特殊情况. 更一般的乘积求和可用 Abel 变换, 也称分部求和法.

这个方法把求和项 a_kb_k 的求和难度在 $\{a_k\}$ 与 $\{b_k\}$ 之间转换. 可能 $\{a_k\}$ 易于求和, 例如等比数列, 其前 k 项和 A_k 有简洁公式, 而 $\{b_k\}$ 的**相邻项差分** $\Delta b_k=b_{k+1}-b_k$ 很简单, 例如多项式, 差分后降次.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

分部求和法把原求和式中对 a_k 的逐项累加，转化为对 A_k 与 Δb_k 乘积的求和，其中 A_k 是 a_k 的和，Δb_k 是 b_k 的差.

这样用一个易于求和的部分换取一个易于求差的部分，把原问题转化为阶次更低、结构更简单的求和问题.

熟悉微积分的场合中，这与分部积分法 $\int u\,dv = uv - \int v\,du$ 完全对偶.

定理 7.6 (阿贝尔变换) 设 $\{a_k\}$ 与 $\{b_k\}$ 为两个数列，并记 $A_n = \sum_{i=1}^{n} a_i$ 为 $\{a_k\}$ 的前 n 项和(规定 $A_0 = 0$). 则

$$\sum_{k=1}^{n} a_k b_k = A_n b_n - \sum_{k=1}^{n-1} A_k (b_{k+1} - b_k)$$

证明 现将求和式中的 a_k 替换掉，引入其部分和 A_k. 数列的通项 a_k 与其部分和 A_k 的最基本关系是差分关系. 当 $k \geq 1$ 时，a_k 等于其第 k 项部分和与第 $k-1$ 项部分和之差.

$$a_k = A_k - A_{k-1}$$

这是整个推导的逻辑起点.

将上述关系代入原求和式，并利用求和的线性性质将其拆分为两个部分，以便进行后续的代数变形.

$$\begin{aligned}\sum_{k=1}^{n} a_k b_k &= \sum_{k=1}^{n} (A_k - A_{k-1}) b_k \\ &= \sum_{k=1}^{n} A_k b_k - \sum_{k=1}^{n} A_{k-1} b_k\end{aligned}$$

为了将两个求和式合并，需让其中求和指标 k 所对应的项具有相同的结构，特别是让 A_k 的下标对齐. 对第二个求和式 $\sum_{k=1}^{n} A_{k-1} b_k$ 进行指标平移.

令新指标 $j = k - 1$. 当旧指标 k 从 1 变化到 n 时，新指标 j 从 0 变化到 $n-1$. 同时，旧通项中的 k 必须用 $j+1$ 替换. 于是

$$\sum_{k=1}^{n} A_{k-1} b_k = \sum_{j=0}^{n-1} A_j b_{j+1}$$

由于求和指标只是一个哑变量，可将 j 重新写为 k:

$$\sum_{k=0}^{n-1} A_k b_{k+1}$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

现在的表达式变为 $\sum_{k=1}^{n} A_k b_k - \sum_{k=0}^{n-1} A_k b_{k+1}$. 这两个和式的求和范围不完全重合. 为了合并它们, 将各自范围中不重合的项分离出来.

知乎@余命数

$$\begin{aligned}\sum_{k=1}^{n} a_k b_k &= \left(\sum_{k=1}^{n-1} A_k b_k + A_n b_n\right) - \left(A_0 b_1 + \sum_{k=1}^{n-1} A_k b_{k+1}\right) \\ &= A_n b_n - A_0 b_1 + \sum_{k=1}^{n-1} A_k b_k - \sum_{k=1}^{n-1} A_k b_{k+1}\end{aligned}$$

根据规定 $A_0 = 0$, 所以 $A_0 b_1 = 0$. 接着合并两个求和范围完全相同的和式, 得

$$\begin{aligned}\sum_{k=1}^{n} a_k b_k &= A_n b_n + \sum_{k=1}^{n-1} (A_k b_k - A_k b_{k+1}) \\ &= A_n b_n - \sum_{k=1}^{n-1} A_k (b_{k+1} - b_k)\end{aligned}$$

证毕. □

例题 7.56 再次使用阿贝尔变换求解差比数列和 $S_n = \sum_{k=1}^{n} k \cdot x^k$ $(x \neq 1)$.

解 需将通项 $k \cdot x^k$ 分解为 $a_k b_k$ 两部分. 阿贝尔变换的作用在于, 如果 $b_{k+1} - b_k$ 比 b_k 简单, 且 $A_k = \sum a_i$ 易于计算, 那么变换后的新求和式会变得更简单.

- 考虑 k: 它是一个一次多项式. 其差分 $(k+1) - k = 1$ 是一个常数, 形式急剧简化. 这使其成为 b_k 的理想选择.
- 考虑 x^k: 它是一个等比数列. 其部分和 $A_k = \sum_{i=1}^{k} x^i$ 有一个简洁的求和公式. 这使其成为 a_k 的理想选择.

因此令 $a_k = x^k$ 且 $b_k = k$.

根据所选, 计算阿贝尔变换公式所需的各个部分, 得

知乎@余命数

- $b_k = k \Longrightarrow b_{k+1} - b_k = 1$.
- $A_k = \sum_{i=1}^{k} x^i = \frac{x(1-x^k)}{1-x}$.
- $A_n = \frac{x(1-x^n)}{1-x}$.
- $A_n b_n = \frac{nx(1-x^n)}{1-x}$.

将上述构件代入阿贝尔变换公式, 得

知乎@余命数

$$\sum_{k=1}^{n} k x^k = A_n b_n - \sum_{k=1}^{n-1} A_k (b_{k+1} - b_k)$$

$$S_n = \frac{nx(1-x^n)}{1-x} - \sum_{k=1}^{n-1} \frac{x(1-x^k)}{1-x} \cdot 1$$

问题转化为计算新的、但更简单的求和式 $\sum_{k=1}^{n-1} \frac{x(1-x^k)}{1-x}$.

$$\begin{aligned}\sum_{k=1}^{n-1} \frac{x(1-x^k)}{1-x} &= \frac{x}{1-x} \sum_{k=1}^{n-1} (1-x^k) \\ &= \frac{x}{1-x} \left(\sum_{k=1}^{n-1} 1 - \sum_{k=1}^{n-1} x^k \right) \\ &= \frac{x}{1-x} \left((n-1) - \frac{x(1-x^{n-1})}{1-x} \right)\end{aligned}$$

代回, 得

$$S_n = \frac{nx(1-x^n)}{1-x} - \frac{x}{1-x} \left(n-1 - \frac{x(1-x^{n-1})}{1-x} \right)$$

继续通分整理, 得

$$\begin{aligned}S_n &= \frac{nx(1-x^n)(1-x) - x(n-1)(1-x) + x^2(1-x^{n-1})}{(1-x)^2} \\ &= \frac{x-(n+1)x^{n+1}+nx^{n+2}}{(1-x)^2}.\end{aligned}$$

这与错位相减法得到的结果一致. 阿贝尔变换将一个需要巧妙构造的问题, 转化为了一个遵循固定范式的程序化计算.

7.8.6.1 部分分式型

我们来推导 $\frac{1}{k(k+d)}$ 的裂项形式.

希望找到 $f(k)$ 使得 $f(k) - f(k+1)$ 与 $\frac{1}{k(k+d)}$ 成比例. 差分运算通常会使分式的分母变得更复杂 (通分导致阶次升高) . 因此, 一个合理的猜想是, 反差分函数 $f(k)$ 的分母应该比原通项**更简单**. 原通项分母是两个因子的乘积, 我们猜想 $f(k)$ 的分母只包含其中一个因子.

我们不妨猜想 $f(k) = \frac{1}{k}$.

计算这个猜想的差分, 得

$$f(k) - f(k+1) = \frac{1}{k} - \frac{1}{k+1} = \frac{(k+1)-k}{k(k+1)} = \frac{1}{k(k+1)}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个结果对于 $d=1$ 的情况直接成功. 对于更一般的情况, 我们猜想 $f(k)=\frac{1}{k}$.

$$f(k)-f(k+d)=\frac{1}{k}-\frac{1}{k+d}=\frac{(k+d)-k}{k(k+d)}=\frac{d}{k(k+d)}$$

这个结果与待求的 $\frac{1}{k(k+d)}$ 只差一个常数因子 d. 因此, 两边同除以 d 即得

$$\frac{1}{k(k+d)}=\frac{1}{d}\left(\frac{1}{k}-\frac{1}{k+d}\right)$$

这个裂项公式便由系统的猜想与验证推得.

7.8.6.2 阶乘添项型

我们来推导 $k\cdot k!$ 的裂项形式.

我们寻找一个函数 $f(k)$, 其差分 $f(k+1)-f(k)$ 等于 $k\cdot k!$. 阶乘世界的基本运算单元是 $k!$ 本身. 一个自然的猜想是, 反差分函数 $f(k)$ 应该与 $k!$ 结构相似.

最大胆的猜想是 $f(k)=k!$.

计算该猜想的差分, 得

$$f(k+1)-f(k)=(k+1)!-k!$$

利用阶乘的性质 $(k+1)!=(k+1)\cdot k!$:

$$(k+1)k!-k!=((k+1)-1)k!=k\cdot k!$$

猜想一次成功. 我们系统地推导出了恒等式 $k\cdot k!=(k+1)!-k!$.

7.8.6.3 平方差裂项型

我们寻找 $f(k)$ 使得 $f(k)-f(k+1)$ 等于目标通项. 观察分母 $k^2(k+1)^2$, 它提示我们反差分函数 $f(k)$ 的分母可能更简单, 例如只包含 k^2.

我们猜想 $f(k)=\dfrac{1}{k^2}$.

计算这个猜想的差分, 得

$$f(k)-f(k+1)=\frac{1}{k^2}-\frac{1}{(k+1)^2}=\frac{(k+1)^2-k^2}{k^2(k+1)^2}$$

计算分子, 得

$$(k+1)^2-k^2=(k^2+2k+1)-k^2=2k+1$$

因此,

$$f(k)-f(k+1)=\frac{2k+1}{k^2(k+1)^2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.9 递推关系与通项公式

知乎@余命数

我们希望通过数列的局部规则得到它的整体表达式. 递推关系描述相邻项之间的联系, 通项公式则给出每一项的具体值. 本节建立由递推关系推导通项公式的代数框架.

7.9.1 和与项的互化关系

在数列研究中, 前 n 项和 S_n 与通项 a_n 包含的信息是等价的.S_n 描述了数列的累积性质, 而 a_n 描述了数列的局部特征. 两者通过差分运算相互转化.

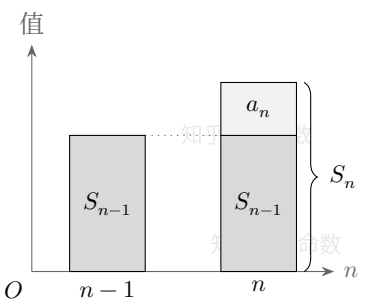
7.9.1.1 基本原理

定义 7.12　设数列 $\{a_n\}$ 的前 n 项和为 $S_n=\sum_{k=1}^{n}a_k$. 则通项 a_n 与和 S_n 的关系如下.

$$a_n=\begin{cases}S_1 & n=1\\ S_n-S_{n-1} & n\geq 2\end{cases}$$

几何上, 这对应于从总长度中减去前一部分的长度, 剩余即为当前项.

知乎@余命数　知乎@余命数



知乎@余命数

7.9.1.2 由 S_n 的解析式求 a_n

知乎@余命数

这是最直接的应用. 计算时需严格区分 $n=1$ 与 $n\geq 2$ 的情况, 最后检验是否可以合并.

例题 7.57　已知数列 $\{a_n\}$ 的前 n 项和为 $S_n=n^2+1$. 求通项公式.

解　先计算首项, 得

$$n=1:\quad a_1=S_1=1^2+1=2$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再计算一般项 $(n \geq 2)$, 得

$$\begin{aligned} a_n &= S_n - S_{n-1} \\ &= (n^2+1) - [(n-1)^2+1] \\ &= n^2 - (n^2 - 2n + 1) \\ &= 2n - 1 \end{aligned}$$

最后检验一致性. 令 $n = 1$ 代入 $2n-1$, 得 $2(1)-1 = 1$. 由于 $1 \neq a_1$ (实际 $a_1 = 2$), 故公式不能合并.

故通项公式为

$$a_n = \begin{cases} 2 & n = 1 \\ 2n-1 & n \geq 2 \end{cases}$$

注 7.8

若 $S_n = An^2 + Bn + C$, 则 $\{a_n\}$ 为等差数列的充要条件是常数项 $C = 0$. 上例中 $C = 1 \neq 0$, 故 a_1 游离于等差规则之外.

7.9.1.3 由含 S_n 的递推关系求 a_n

当已知条件给出 S_n 与 a_n 的混合关系时, 我们用移位作差消去 S_n, 把它化为 a_n 与 a_{n-1} 的纯递推关系.

例题 7.58 已知数列 $\{a_n\}$ 满足 $a_1 = 1$, 且 $2S_n = 3a_n - 1$. 求通项公式.

解 构造相邻项的等式即可消去 S_n.

当 $n \geq 2$ 时, 写出原式与降阶后的式子.

$$2S_n = 3a_n - 1 \tag{7-7}$$

$$2S_{n-1} = 3a_{n-1} - 1 \tag{7-8}$$

两式相减 (7-7) − (7-8), 得

$$2(S_n - S_{n-1}) = 3a_n - 3a_{n-1}$$

代入 $S_n - S_{n-1} = a_n$, 得

$$2a_n = 3a_n - 3a_{n-1}$$

整理得

$$a_n = 3a_{n-1}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $\{a_n\}$ 从第 2 项开始构成公比 $q=3$ 的等比数列.

下面检验首项. 由已知 $a_1=1$. 当 $n=2$ 时, 由递推式得 $a_2=3a_1=3$.

若通项为 $a_n=1\cdot 3^{n-1}$, 则 $a_1=3^0=1$, 符合.

因此, 通项公式为 $a_n=3^{n-1}$.

7.9.1.4 连乘积与项的转化

类似于和与项的关系, 若定义前 n 项积 $T_n=\prod_{k=1}^{n}a_k(a_k\neq 0)$, 则有乘法形式的互化关系.

定义 7.13

$$a_n=\begin{cases}T_1 & n=1\\ \frac{T_n}{T_{n-1}} & n\geq 2\end{cases}$$

例题 7.59 已知数列 $\{a_n\}$ 的前 n 项积 $T_n=n^2$. 求通项公式.

解 当 $n=1$ 时, $a_1=T_1=1^2=1$.

当 $n\geq 2$ 时,

$$a_n=\frac{T_n}{T_{n-1}}=\frac{n^2}{(n-1)^2}=\left(\frac{n}{n-1}\right)^2$$

检验 $n=1$: $\left(\frac{1}{0}\right)^2$ 无意义, 或趋于无穷, 且原式计算 $n=1$ 时代入公式分母为零. 故公式不可合并, 通项公式为

$$a_n=\begin{cases}1 & n=1\\ \left(\frac{n}{n-1}\right)^2 & n\geq 2\end{cases}$$

常见类型可归纳如下.

7.9.1.5 S_n 为显式函数

当 S_n 以 n 的函数形式给出时, 直接利用定义求解. 注意常数项往往导致首项不符合后续规律.

例题 7.60 (2022 全国卷) 已知数列 $\{a_n\}$ 的前 n 项和为 $S_n=n^2+4n-3$. 求通项公式.

解 先考察 $n=1$ 的情形.

$$a_1=S_1=1^2+4(1)-3=2$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再考察 $n \geq 2$ 的情形.

$$\begin{aligned} a_n &= S_n - S_{n-1} \\ &= (n^2+4n-3) - [(n-1)^2 + 4(n-1) - 3] \\ &= n^2 + 4n - 3 - (n^2 - 2n + 1 + 4n - 4 - 3) \\ &= n^2 + 4n - 3 - (n^2 + 2n - 6) \\ &= 2n + 3 \end{aligned}$$

知乎@余命数

再检验一致性. 令 $n=1$, 得 $2(1)+3=5$. 由于 $5 \neq a_1$(实际 $a_1=2$), 故通项公式不统一, 即

$$a_n = \begin{cases} 2 & n=1 \\ 2n+3 & n \geq 2 \end{cases}$$

7.9.1.6 线性混合递推

知乎@余命数

当条件给出 a_n 与 S_n 的线性关系时, 利用移位作差法消去 S_n, 转化为 a_n 与 a_{n-1} 的纯递推关系.

知乎@余命数

例题 7.61 (2013 全国卷) 若数列 $\{a_n\}$ 的前 n 项和为 $S_n = \frac{2}{3}a_n + \frac{1}{3}$, 求数列 $\{a_n\}$ 的通项公式.

解 先确定首项 a_1. 令 $n=1$, 得

$$S_1 = \frac{2}{3}a_1 + \frac{1}{3} \implies a_1 = \frac{2}{3}a_1 + \frac{1}{3} \implies \frac{1}{3}a_1 = \frac{1}{3} \implies a_1 = 1$$

知乎@余命数

对于 $n \geq 2$, 利用移位作差法.

$$\begin{aligned} S_n &= \frac{2}{3}a_n + \frac{1}{3} \\ S_{n-1} &= \frac{2}{3}a_{n-1} + \frac{1}{3} \end{aligned}$$

知乎@余命数

两式相减得 $a_n = \frac{2}{3}a_n - \frac{2}{3}a_{n-1}$. 移项整理, 得

$$\frac{1}{3}a_n = -\frac{2}{3}a_{n-1} \implies a_n = -2a_{n-1}$$

由此可知, 数列 $\{a_n\}$ 从第 2 项开始满足公比 $q=-2$ 的等比关系. 下面验证首项. 若公式 $a_n = a_1 \cdot q^{n-1}$ 对 $n=1$ 成立, 则需 $1 \cdot (-2)^0 = 1$, 这与实际 a_1 一致.

因此, 通项公式为 $a_n = (-2)^{n-1}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.9.1.7 隐式与二次关系

当关系式中包含 a_n 的高次项时, 通常需结合代数变形, 如配方, 挖掘其内在的结构.

例题 7.62 (2015 全国 I 卷) 已知 $a_n>0$, 且 $a_n^2+2a_n=4S_n+3$. 求 $\{a_n\}$ 的通项公式.

解 先求 a_1. 令 $n=1$, 得

$$a_1^2+2a_1=4a_1+3 \implies a_1^2-2a_1-3=0 \implies (a_1-3)(a_1+1)=0$$

因 $a_n>0$, 解得 $a_1=3$.

对于 $n\geq 2$, 观察原式结构, 等号左边配方为 $(a_n+1)^2-1$.

$$(a_n+1)^2=4S_n+4$$

同理可写出

$$(a_{n-1}+1)^2=4S_{n-1}+4$$

两式相减, 得

$$(a_n+1)^2-(a_{n-1}+1)^2=4(S_n-S_{n-1})=4a_n$$

利用平方差公式展开左边, 或者直接展开计算, 得

$$(a_n^2+2a_n+1)-(a_{n-1}+1)^2=4a_n$$

$$a_n^2-2a_n+1=(a_{n-1}+1)^2$$

$$(a_n-1)^2=(a_{n-1}+1)^2$$

两边开方. 由于 $a_n>0$, 只能取正号 $a_n-1=a_{n-1}+1$(若取负号则导致 $a_n+a_{n-1}=0$, 矛盾).

$$a_n-a_{n-1}=2$$

故 $\{a_n\}$ 是首项为 3, 公差为 2 的等差数列.

通项公式为 $a_n=3+(n-1)2=2n+1$.

7.9.1.8 结构化消元

当递推式呈现特定的代数结构, 如平方差形式、根式形式时, 应优先考虑整体代换.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.63 (2020 辽宁) 已知数列 $\{a_n\}$ 各项均为正数, 其前 n 项和为 S_n. 若 $a_1=1$, 且 $\sqrt{S_{n+1}}+\sqrt{S_n}=a_{n+1}$. 求通项公式.

解 利用 $a_{n+1}=S_{n+1}-S_n$, 将原式变形为

$$\sqrt{S_{n+1}}+\sqrt{S_n}=S_{n+1}-S_n$$

利用平方差公式分解右侧, 得

$$\sqrt{S_{n+1}}+\sqrt{S_n}=(\sqrt{S_{n+1}}-\sqrt{S_n})(\sqrt{S_{n+1}}+\sqrt{S_n})$$

由于各项为正, $\sqrt{S_{n+1}}+\sqrt{S_n}\neq 0$, 两边约去该公因子, 得

$$1=\sqrt{S_{n+1}}-\sqrt{S_n}$$

故数列 $\{\sqrt{S_n}\}$ 是一个等差数列, 公差 $d=1$.

首项为 $\sqrt{S_1}=\sqrt{a_1}=1$. 所以 $\sqrt{S_n}=1+(n-1)\cdot 1=n$. 两边平方得 $S_n=n^2$.

最后还原 a_n: 当 $n=1$ 时, $a_1=1$. 当 $n\geq 2$ 时, $a_n=S_n-S_{n-1}=n^2-(n-1)^2=2n-1$. 经检验, $n=1$ 时 $2(1)-1=1$, 符合.

故通项公式为 $a_n=2n-1$.

知乎@余命数

7.10 一阶递推的基本形式

当递推关系仅涉及相邻两项的差或比时, 数列的通项可以通过遍历所有项的积累效应(求和或求积)直接还原.

7.10.1 累加法

形如 $a_{n+1}-a_n=f(n)$ 的递推关系. 该式表明, 数列的增长率是关于 n 的函数.

知乎@余命数

定理 7.7 若 $a_{n+1}-a_n=f(n)$, 则

$$a_n=a_1+\sum_{k=1}^{n-1}f(k)\quad(n\geq 2)$$

知乎@余命数

证明 对递推式从 $k=1$ 到 $k=n-1$ 求和, 得

$$\sum_{k=1}^{n-1}(a_{k+1}-a_k)=\sum_{k=1}^{n-1}f(k).$$

左侧裂项相消为 $a_n - a_1$, 故

$$a_n = a_1 + \sum_{k=1}^{n-1} f(k).$$

□

知乎@余命数

裂项相消求解的要点在于计算 $\sum f(k)$. 若 $f(k)$ 能被拆解为 $g(k+1) - g(k)$ 的形式, 求和过程将极度简化.

例题 7.64 (根式裂项) 已知数列 $\{a_n\}$ 满足 $a_1 = 1$, 且 $a_{n+1} = a_n + \frac{1}{\sqrt{n}+\sqrt{n+1}}$. 求 a_n.

解 将递推项变形, 用分母有理化构造差分形式, 得

$$f(n) = \frac{1}{\sqrt{n}+\sqrt{n+1}} = \frac{\sqrt{n+1}-\sqrt{n}}{(n+1)-n} = \sqrt{n+1}-\sqrt{n}$$

原递推式化为 $a_{n+1} - a_n = \sqrt{n+1} - \sqrt{n}$.

知乎@余命数

对 $n \geq 2$ 累加, 得

$$\begin{aligned} a_n &= a_1 + \sum_{k=1}^{n-1} (\sqrt{k+1} - \sqrt{k}) \\ &= 1 + [(\sqrt{2}-\sqrt{1}) + (\sqrt{3}-\sqrt{2}) + \cdots + (\sqrt{n}-\sqrt{n-1})] \end{aligned}$$

中间项全部抵消, 只留首尾.

$$a_n = 1 + (\sqrt{n} - \sqrt{1}) = \sqrt{n}$$

检验 $n = 1$, $\sqrt{1} = 1$, 符合. 故 $a_n = \sqrt{n}$.

例题 7.65 (分式裂项) 已知 $a_1 = 2$, 且 $a_{n+1} = a_n + \frac{1}{n^2+2n}$. 求 a_n.

知乎@余命数

解 对 $f(n) = \frac{1}{n(n+2)}$ 作部分分式分解, 得

$$\frac{1}{n(n+2)} = \frac{1}{2}\left(\frac{1}{n} - \frac{1}{n+2}\right)$$

知乎@余命数

累加求和(注意跨度为 2, 需保留两项), 得

$$\begin{aligned} a_n - a_1 &= \frac{1}{2}\sum_{k=1}^{n-1}\left(\frac{1}{k} - \frac{1}{k+2}\right) \\ &= \frac{1}{2}\left[\left(1-\frac{1}{3}\right) + \left(\frac{1}{2}-\frac{1}{4}\right) + \left(\frac{1}{3}-\frac{1}{5}\right) + \cdots + \left(\frac{1}{n-1}-\frac{1}{n+1}\right)\right] \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

剩余项为起始的正项 $1,\frac{1}{2}$ 与末尾的负项 $-\frac{1}{n},-\frac{1}{n+1}$:

$$a_n-2=\frac{1}{2}\left(1+\frac{1}{2}-\frac{1}{n}-\frac{1}{n+1}\right)$$

整理得

$$a_n=\frac{11}{4}-\frac{2n+1}{2n(n+1)}$$

例题 7.66 (2020 内蒙古) 在数列 $\{a_n\}$ 中, $a_1=3, a_{n+1}=a_n+\frac{3}{n(n+1)}$. 求通项公式 a_n.

解 差分项为 $f(n)=\frac{3}{n(n+1)}$. 作部分分式分解, 把 $f(n)$ 裂项.

$$f(n)=3\left(\frac{1}{n}-\frac{1}{n+1}\right)$$

当 $n\geq 2$ 时累加, 得

$$\begin{aligned}a_n&=a_1+\sum_{k=1}^{n-1}f(k)\\&=3+3\sum_{k=1}^{n-1}\left(\frac{1}{k}-\frac{1}{k+1}\right)\\&=3+3\left[\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\cdots+\left(\frac{1}{n-1}-\frac{1}{n}\right)\right]\end{aligned}$$

中间项相互抵消, 只留首尾.

$$a_n=3+3\left(1-\frac{1}{n}\right)=6-\frac{3}{n}$$

检验 $n=1$: $6-3=3$, 与已知一致. 故 $a_n=6-\frac{3}{n}$.

7.10.1.1 构造性累加

当递推关系不直接为差分形式, 但两边结构相似时, 可通过构造新数列转化为累加问题.

例题 7.67 (2021 全国乙卷) 在数列 $\{a_n\}$ 中, $a_1=2$, 且 $\frac{a_{n+1}}{n+1}=\frac{a_n}{n}+\ln(1+\frac{1}{n})$. 求 a_n.

解 等式两边形式均为项除以项数. 令 $b_n=\frac{a_n}{n}$, 则原式化为

$$b_{n+1}=b_n+\ln\left(\frac{n+1}{n}\right)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即 $b_{n+1}-b_n=\ln(n+1)-\ln n$.

这是一个标准的累加形式. 对于 $n\geq 2$:

$$\begin{aligned} b_n &= b_1+\sum_{k=1}^{n-1}[\ln(k+1)-\ln k] \\ &= \frac{a_1}{1}+(\ln n-\ln 1) \\ &= 2+\ln n \end{aligned}$$

所以 $b_n=2+\ln n$.

还原 a_n:

$$a_n=nb_n=n(2+\ln n)$$

经检验 $n=1$ 时成立.

7.10.2 累乘法

形如 $a_{n+1}=f(n)\cdot a_n$ 的递推关系. 该式表明, 数列的增长比例是关于 n 的函数.

定理 7.8 若数列各项非零, 且 $\frac{a_{n+1}}{a_n}=f(n)$、$f(n)\neq 0$, 则

$$a_n=a_1\cdot\prod_{k=1}^{n-1}f(k)\quad(n\geq 2)$$

证明 对等式 $\frac{a_{k+1}}{a_k}=f(k)$ 从 $k=1$ 到 $k=n-1$ 连乘, 得

$$\prod_{k=1}^{n-1}\frac{a_{k+1}}{a_k}=\prod_{k=1}^{n-1}f(k).$$

左侧相邻分子分母逐项约去, 只剩 $\frac{a_n}{a_1}$. 由于 $a_1\neq 0$, 两边同乘 a_1, 即得

$$a_n=a_1\prod_{k=1}^{n-1}f(k).$$

□

因子缩并求解要点在于将 $f(k)$ 写成分子与分母错位对应的形式, 使连乘积发生大规模约分.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.68 (Wallis 型约分) 已知数列 $\{a_n\}$ 满足 $a_1=1$, $a_2=1$, 且 $a_{n+1}=\frac{n^2}{n^2-1}a_n$ $(n\geq 2)$. 求 a_n.

解 注意 $n^2-1=(n-1)(n+1)$. 递推式从 $n=2$ 开始, 故需由 a_2 起累乘.

将因子结构化.

$$\frac{a_{k+1}}{a_k}=\frac{k\cdot k}{(k-1)(k+1)}$$

累乘展开, 得

$$a_n=a_2\cdot\prod_{k=2}^{n-1}\frac{k^2}{k^2-1}=a_2\cdot\left(\frac{2\cdot 2}{1\cdot 3}\right)\cdot\left(\frac{3\cdot 3}{2\cdot 4}\right)\cdot\left(\frac{4\cdot 4}{3\cdot 5}\right)\cdots\left(\frac{(n-1)(n-1)}{(n-2)n}\right)$$

观察交叉约分规律.

$$a_n=a_2\cdot\frac{2}{1}\cdot\underbrace{\left(\frac{2}{3}\cdot\frac{3}{2}\right)}_{1}\cdot\underbrace{\left(\frac{3}{4}\cdot\frac{4}{3}\right)}_{1}\cdots\cdot\frac{n-1}{n}$$

剩余项为

$$a_n=a_2\cdot 2\cdot\frac{n-1}{n}=\frac{2(n-1)}{n}a_2$$

由 $a_2=1$, 得

$$a_1=1,\qquad a_n=\frac{2(n-1)}{n}\quad(n\geq 2).$$

例题 7.69 (指数型累乘) 已知 $a_1=1$, 且 $a_{n+1}=3^n a_n$. 求 a_n.

解 由递推式可知

$$\frac{a_{n+1}}{a_n}=3^n$$

累乘得

$$a_n=a_1\cdot\prod_{k=1}^{n-1}3^k=1\cdot 3^1\cdot 3^2\cdot 3^3\cdots 3^{n-1}$$

于是转化为指数求和(等差数列).

$$a_n=3^{1+2+\cdots+(n-1)}=3^{\frac{(n-1)n}{2}}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.70 (2023 全国乙卷) 数列 $\{a_n\}$ 满足 $a_1=\frac{1}{2}$, 且 $S_n=n^2a_n$. 求 $\{a_n\}$ 的通项公式.

解 我们先寻找相邻项的比值关系. 当 $n\geq 2$ 时, 由 $a_n=S_n-S_{n-1}$ 出发,

$$a_n=n^2a_n-(n-1)^2a_{n-1}$$

移项整理, 得

$$(n-1)^2a_{n-1}=(n^2-1)a_n$$

$$(n-1)^2a_{n-1}=(n-1)(n+1)a_n$$

因 $n\geq 2$, 可约去 $n-1$:

$$(n-1)a_{n-1}=(n+1)a_n \implies \frac{a_n}{a_{n-1}}=\frac{n-1}{n+1}$$

下面用累乘法求解 a_n $(n\geq 2)$.

$$\begin{aligned} a_n&=a_1\cdot\prod_{k=2}^{n}\frac{k-1}{k+1}\\ &=\frac{1}{2}\cdot\left(\frac{1}{3}\cdot\frac{2}{4}\cdot\frac{3}{5}\cdot\frac{4}{6}\cdots\frac{n-2}{n}\cdot\frac{n-1}{n+1}\right)\end{aligned}$$

观察分子分母的对消模式(跨度为 2). 分子保留 $1,2$ 两项. 分母保留 $n,n+1$ 两项.

$$a_n=\frac{1}{2}\cdot\frac{1\cdot 2}{n(n+1)}=\frac{1}{n(n+1)}$$

检验 $n=1$: $\frac{1}{1(2)}=\frac{1}{2}$, 与 a_1 一致. 故 $a_n=\frac{1}{n(n+1)}$.

7.10.2.1 加权和的剥离

形如 $\sum c_ka_k=f(n)$ 的关系式, 实质上是给出了一个新数列 $\{c_na_n\}$ 的前 n 项和. 利用差分即可求出 c_na_n, 进而求出 a_n.

例题 7.71 (2019 江苏) 已知 $\sum_{k=1}^{n}2^{k-1}a_k=9-6n$. 求数列 $\{a_n\}$ 的通项公式.

解 令 $T_n=\sum_{k=1}^{n}2^{k-1}a_k=9-6n$. 则数列 $\{2^{n-1}a_n\}$ 的第 n 项可表示为 T_n-T_{n-1}.

当 $n=1$ 时,

$$2^0a_1=T_1=9-6(1)=3 \implies a_1=3$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当 $n \geq 2$ 时,

$$\begin{aligned} 2^{n-1}a_n &= T_n - T_{n-1} \\ &= (9-6n) - [9-6(n-1)] \\ &= 9-6n-(9-6n+6) \\ &= -6 \end{aligned}$$

解出 a_n:

$$a_n = \frac{-6}{2^{n-1}} = -3 \cdot \frac{2}{2^{n-1}} = -3 \cdot \left(\frac{1}{2}\right)^{n-2}$$

检验 $n=1$: 代入公式得 $-3 \cdot (\frac{1}{2})^{-1} = -6$, 而实际 $a_1 = 3$. 故公式不统一, 通项须分段给出.

$$a_n = \begin{cases} 3 & n = 1 \\ -3 \cdot (\frac{1}{2})^{n-2} & n \geq 2 \end{cases}$$

7.10.3 结构匹配与变形

有时递推关系并不直接呈现为差或比的形式, 但通过移项或同除, 可以构造出具备上述特征的新数列.

同型构造若递推式形如 $g(n+1)a_{n+1} = g(n)a_n$, 则可定义 $b_n = g(n)a_n$, 此时 $b_{n+1} = b_n$, 即 $\{b_n\}$ 为常数列.

例题 7.72 已知 $a_1 = 1$, 且 $na_{n+1} = (n+1)a_n + 2n(n+1)$. 求 a_n.

解 观察系数 n 与 $n+1$, 将等式两边同除以 $n(n+1)$, 得

$$\frac{a_{n+1}}{n+1} = \frac{a_n}{n} + 2$$

令 $b_n = \frac{a_n}{n}$. 则 $b_{n+1} - b_n = 2$. 数列 $\{b_n\}$ 是公差为 2 的等差数列.
首项为 $b_1 = \frac{a_1}{1} = 1$.

$$b_n = 1 + (n-1)2 = 2n-1$$

还原 a_n:

$$a_n = nb_n = n(2n-1) = 2n^2 - n$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.10.4 线性递推与代数构造

处理 $a_{n+1}=pa_n+f(n)$ $(p\neq 1)$ 形式的递推关系, 一般通过待定系数法或同除变形, 构造新的等比数列或等差数列.

7.10.4.1 常数项构造

当 $f(n)=q$ 为常数时, 即 $a_{n+1}=pa_n+q$. 设 $a_{n+1}+\lambda=p(a_n+\lambda)$, 展开比对得 $\lambda(p-1)=q$, 即 $\lambda=\frac{q}{p-1}$. 此时数列 $\{a_n+\lambda\}$ 为公比为 p 的等比数列.

7.10.4.2 同除构造

当 $f(n)$ 为指数函数 $c\cdot p^n$ 或多项式时, 常在等式两边同除以 p^{n+1}, 转化为累加形式.

例题 7.73 已知 $a_1=1$, 且 $a_{n+1}=2a_n+3^n$. 求 a_n.

证明 (解析) 观察非齐次项 3^n, 等式两边同除以 3^{n+1}, 得

$$\frac{a_{n+1}}{3^{n+1}}=\frac{2a_n}{3\cdot 3^n}+\frac{3^n}{3^{n+1}}$$

整理得

$$\frac{a_{n+1}}{3^{n+1}}=\frac{2}{3}\cdot\frac{a_n}{3^n}+\frac{1}{3}$$

令 $b_n=\frac{a_n}{3^n}$, 则 $b_{n+1}=\frac{2}{3}b_n+\frac{1}{3}$. 设 $b_{n+1}+\lambda=\frac{2}{3}(b_n+\lambda)$, 解得 $\lambda=-1$. 数列 $\{b_n-1\}$ 是首项为 $b_1-1=\frac{1}{3}-1=-\frac{2}{3}$, 公比为 $\frac{2}{3}$ 的等比数列.

$$b_n-1=\left(-\frac{2}{3}\right)\cdot\left(\frac{2}{3}\right)^{n-1}=-\left(\frac{2}{3}\right)^n$$

$$b_n=1-\left(\frac{2}{3}\right)^n$$

还原 a_n:

$$a_n=3^n\cdot b_n=3^n\left[1-\left(\frac{2}{3}\right)^n\right]=3^n-2^n$$

□

7.10.5 非线性递推的变换技巧

通过代数变换, 可将部分非线性递推转化为线性递推.

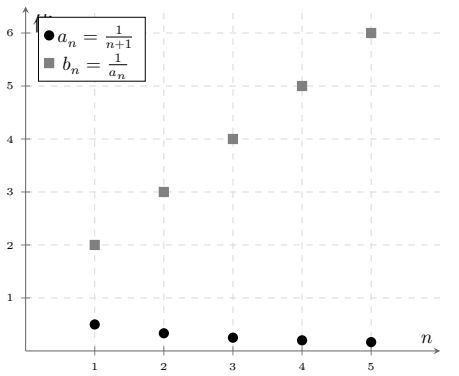
知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

7.10.5.1 倒数变换

针对分式线性形式 $a_{n+1}=\frac{Aa_n}{Ba_n+C}$. 若 $C\neq 0$, 取倒数得

$$\frac{1}{a_{n+1}}=\frac{Ba_n+C}{Aa_n}=\frac{C}{A}\cdot\frac{1}{a_n}+\frac{B}{A}$$

令 $b_n=\frac{1}{a_n}$, 即转化为 $b_{n+1}=pb_n+q$ 形式.



知乎@余命数

倒数变换将双曲线型衰减序列转化为线性增长序列.

7.10.5.2 对数变换

针对幂指形式 $a_{n+1}=Ca_n^k$ $(a_n>0)$. 两边取对数, 得

$$\ln a_{n+1}=k\ln a_n+\ln C$$

令 $b_n=\ln a_n$, 即转化为线性递推.

知乎@余命数

7.10.6 二阶常系数线性递推

形如 $a_{n+2}=pa_{n+1}+qa_n$ 的递推关系, 其通项结构由特征方程 $x^2-px-q=0$ 的根决定.

知乎@余命数

定理 7.9 设 $q\neq 0$, 特征方程 $x^2-px-q=0$ 的两根为 λ_1,λ_2.

1. 若 $\lambda_1\neq\lambda_2$, 则 $a_n=c_1\lambda_1^n+c_2\lambda_2^n$.
2. 若 $\lambda_1=\lambda_2=\lambda$, 则 $a_n=(c_1+c_2n)\lambda^n$.

其中 c_1,c_2 由初始值 a_1,a_2 确定.

证明 若 x 是特征方程的根, 则序列 $u_n=x^n$ 满足

$$u_{n+2}=x^{n+2}=px^{n+1}+qx^n=pu_{n+1}+qu_n.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当 $\lambda_1 \neq \lambda_2$ 时, λ_1^n 与 λ_2^n 都是递推方程的解, 因而其线性组合 $c_1\lambda_1^n + c_2\lambda_2^n$ 也是解. 因 $q \neq 0$, 两根均非零, 由初值方程

$$c_1\lambda_1 + c_2\lambda_2 = a_1, \qquad c_1\lambda_1^2 + c_2\lambda_2^2 = a_2$$

的系数行列式为 $\lambda_1\lambda_2(\lambda_2 - \lambda_1) \neq 0$, 可唯一确定 c_1, c_2.

当 $\lambda_1 = \lambda_2 = \lambda$ 时, 有 $p = 2\lambda, q = -\lambda^2$, 且 $\lambda \neq 0$. 除 λ^n 外, 序列 $n\lambda^n$ 也满足递推式:

$$(n+2)\lambda^{n+2} = 2\lambda(n+1)\lambda^{n+1} - \lambda^2 n\lambda^n.$$

知乎@余命数

因而 $(c_1 + c_2 n)\lambda^n$ 是解. 由 a_1, a_2 得到的两元一次方程组行列式为 $\lambda^3 \neq 0$, 故 c_1, c_2 也唯一确定.

最后, 任意两个满足同一二阶递推且前两项相同的数列, 其差数列前两项为零, 由递推式可归纳推出以后各项也全为零. 因此上述形式给出了满足初值的唯一通项. □

例题 7.74 求斐波那契数列变体 $a_1 = 1, a_2 = 5$, 且 $a_{n+2} = 5a_{n+1} - 6a_n$ 的通项.

证明 (解析) 先写出特征方程.

$$x^2 - 5x + 6 = 0 \implies (x-2)(x-3) = 0$$

特征根为 $\lambda_1 = 2, \lambda_2 = 3$. 设通项为 $a_n = c_1 \cdot 2^n + c_2 \cdot 3^n$. 再代入初值.

$$\begin{cases} 2c_1 + 3c_2 = 1 \\ 4c_1 + 9c_2 = 5 \end{cases}$$

解得 $c_1 = -1, c_2 = 1$. 故 $a_n = 3^n - 2^n$. □

知乎@余命数

知乎@余命数

7.10.7 特征方程法的由来

二阶常系数线性递推关系形如 $a_{n+2} = pa_{n+1} + qa_n$. 特征方程 $x^2 - px - q = 0$ 是基于**试探解**的自然产物, 其背后的数学逻辑与微分方程求解一脉相承.

回到一阶递推 $a_{n+1} = pa_n$. 一阶递推 $a_{n+1} = pa_n$ 的通项公式为等比数列形式

$$a_n = a_1 \cdot p^{n-1} = C \cdot p^n$$

这提示我们, 指数形式 x^n 是线性递推关系最自然的解的结构. 因此, 对于二阶关系, 我们假设通项公式也具有 $a_n = x^n$ 的形式, 其中 $x \neq 0$ 为待定常数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将 $a_n = x^n$ 代入原递推关系式,

$$a_{n+2} - pa_{n+1} - qa_n = 0$$

知乎@余命数

得到

$$x^{n+2} - px^{n+1} - qx^n = 0$$

知乎@余命数

由于 $x \neq 0$, 等式两边可以同时除以 x^n.

$$x^2 - px - q = 0$$

这便得到了该递推关系的**特征方程**.

这个方程表明, 只有当 x 取特征方程的根 (λ_1, λ_2) 时, 函数 x^n 才是原递推关系的解.

- **基本解.** λ_1^n 和 λ_2^n 都是满足条件的特解.
- **线性叠加.** 由于递推关系是线性的(不含 a_n^2 或 $a_n a_{n+1}$ 等项), 特解的线性组合依然是解.

知乎@余命数

由此构造出通解, 得

$$a_n = c_1\lambda_1^n + c_2\lambda_2^n$$

知乎@余命数 知乎@余命数

特征方程这一名称借用于线性代数, 见 第22章. 将递推关系改写为矩阵形式, 结构会更清楚.

定义向量 $\mathbf{v}_n = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$. 递推关系可以表示为线性变换.

$$\begin{pmatrix} a_{n+2} \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} p & q \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$$

知乎@余命数

即 $\mathbf{v}_{n+1} = A\mathbf{v}_n$. 由此推导 $\mathbf{v}_n = A^n\mathbf{v}_0$.

求解 A^n 的要点在于求矩阵 A 的特征值. 计算 A 的特征多项式 $|\lambda I - A| = 0$:

知乎@余命数 知乎@余命数

$$\det \begin{pmatrix} \lambda - p & -q \\ -1 & \lambda \end{pmatrix} = (\lambda - p)\lambda - (-q)(-1) = \lambda^2 - p\lambda - q = 0$$

知乎@余命数

这与代数法推导出的方程完全一致. 数列的通项公式, 实质上就是矩阵 A 在特征向量基下的幂运算结果.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.10.8 隔项与周期

部分数列的规律需依据项数的奇偶性或周期性分别讨论.

7.10.8.1 奇偶项讨论

若递推关系跨度为 2(如 a_{n+2} 与 a_n 相关), 或系数含 $(-1)^n$, 通常将 $\{a_{2k}\}$ 与 $\{a_{2k-1}\}$ 视为两个子数列分别求解.

7.10.8.2 周期性

若计算发现 $a_{n+k}=a_n$ 对某定值 k 成立, 则数列为周期数列. 利用归纳法或模运算确定通项. 常见于分式迭代, 如 $a_{n+1}=1-\frac{1}{a_n}$(周期为 3).

例题 7.75 已知 $a_1=2$, 且 $a_{n+1}=-a_n+2$. 求 a_n.

证明 (解析) 计算前几项, 得 $a_1=2$. $a_2=-2+2=0$. $a_3=-0+2=2$. 观察知数列呈 $2,0,2,0,\ldots$ 交替.

下面分奇偶讨论.

$$a_n=\begin{cases}2 & n\text{ 为奇数}\\ 0 & n\text{ 为偶数}\end{cases}$$

也可合写为 $a_n=1+(-1)^{n+1}$. □

7.11 迭代与不动点

在前文中, 我们讨论了递推公式, 即由前一项或几项计算后一项的规则. 若把这一规则看成一个函数 f, 则数列的生成过程就是该函数的反复作用. 我们称这种重复应用函数的过程为**迭代**.

除了生成数列, 我们还用迭代求解方程和优化模型. 从古巴比伦人计算平方根到现代计算机训练神经网络, 都建立在迭代之上.

注 7.9
压缩映射、导数判别与牛顿迭代依赖极限与导数, 见第16章.

7.11.1 不动点与方程求解

我们研究的不动点, 是既在几何上有直观意义, 又能把方程求解与迭代过程联系起来的一个对象.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 7.14 (不动点) 设 $f : D \to D$ 是一个函数. 如果存在 $x^* \in D$ 使得

$$f(x^*) = x^*$$

则称 x^* 为函数 f 的**不动点**.

从几何上看, 函数 $y = f(x)$ 的图像与直线 $y = x$ 的交点的横坐标, 就是该函数的不动点.

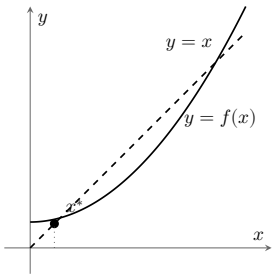


图 7.2 不动点的几何意义是 $y = f(x)$ 与 $y = x$ 的交点 (x^*, x^*), 其横坐标 x^* 满足 $f(x^*) = x^*$.

求解代数方程 $g(x) = 0$ 是数学中古老的问题. 对于五次及以上的多项式方程或超越方程, 如 $x = \cos x$, 通常不存在代数求根公式. 此时可以把求根问题改写为迭代问题.

考虑方程 $g(x) = 0$. 可经代数变形改写为等价形式 $x = f(x)$. 例如

- 简单移项给出 $x = x + \lambda g(x)$, 其中 $\lambda \neq 0$.
- 牛顿法构造给出 $x = x - \frac{g(x)}{g'(x)}$.

此时, 方程 $g(x) = 0$ 的**根**, 就是辅助函数 $f(x)$ 的**不动点**. 一旦将方程转化为 $x = f(x)$ 的形式, 我们便可以构造迭代序列, 得

$$x_{n+1} = f(x_n), \quad n = 0, 1, 2, \ldots \tag{7-9}$$

给定初始猜测 x_0, 得序列 $x_0, x_1, x_2, \ldots$. 若该序列收敛于 x^*, 且 f 是连续函数, 则

$$x^* = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} f(x_n) = f(x^*)$$

即极限值 x^* 必为 f 的不动点, 也就是原方程 $g(x) = 0$ 的根.

7.11.2 不动点法求通项的原理

不动点法的工作方式是坐标变换. 它把原点平移到不动点, 或者以不动点为基准建立参照系, 从而把复杂的非齐次或非线性递推关系化为简单的等比或等差数列.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

7.11.2.1 线性递推中的平移变换

考虑一阶线性递推关系

$$a_{n+1} = pa_n + q \quad (p \neq 1, q \neq 0)$$

这是一个非齐次线性方程, 含有常数项 q. 需消去常数项 q, 将其转化为齐次形式, 即等比数列.

设该迭代函数 $f(x) = px + q$ 的不动点为 c. 则 c 满足

$$c = pc + q$$

将原递推式与不动点方程相减, 得

$$\begin{aligned} a_{n+1} &= pa_n + q \\ c &= pc + q \\ \Longrightarrow\ a_{n+1} - c &= p(a_n - c) \end{aligned}$$

令 $b_n = a_n - c$. 上式即为 $b_{n+1} = pb_n$. 故, 虽然数列 $\{a_n\}$ 本身不是等比数列, 但它**相对于不动点 c 的距离** $\{a_n - c\}$ 构成了一个公比为 p 的等比数列.

原理. 找不动点就是找一个参照位置. 通过变换 $b_n = a_n - c$(坐标平移), 非齐次的仿射变换 $x \mapsto px + q$ 化为齐次的线性变换 $y \mapsto py$.

7.11.2.2 分式线性递推中的比值变换

考虑分式线性递推关系(莫比乌斯变换)

$$a_{n+1} = \frac{pa_n + q}{ra_n + s} \quad (r \neq 0, ps - qr \neq 0)$$

设其迭代函数 $f(x) = \frac{px+q}{rx+s}$. 若直接迭代, 通项公式很复杂. 不动点法通过构造交比把它线性化.

7.11.2.3 有两个相异不动点 α, β

不动点满足 $f(\alpha) = \alpha, f(\beta) = \beta$. 我们考察 a_{n+1} 与不动点的距离之比.

$$\begin{aligned} a_{n+1} - \alpha &= \frac{pa_n + q}{ra_n + s} - \frac{p\alpha + q}{r\alpha + s} \\ &= \frac{(pa_n + q)(r\alpha + s) - (p\alpha + q)(ra_n + s)}{(ra_n + s)(r\alpha + s)} \\ &= \frac{(ps - qr)(a_n - \alpha)}{(ra_n + s)(r\alpha + s)} \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

同理可得

$$a_{n+1}-\beta=\frac{(ps-qr)(a_n-\beta)}{(ra_n+s)(r\beta+s)}$$

两式相除, 消去复杂的公因子, 得

$$\frac{a_{n+1}-\alpha}{a_{n+1}-\beta}=\left(\frac{r\beta+s}{r\alpha+s}\right)\cdot\frac{a_n-\alpha}{a_n-\beta}$$

令 $b_n=\frac{a_n-\alpha}{a_n-\beta}$. 上式即为 $b_{n+1}=kb_n$, 其中 $k=\frac{r\beta+s}{r\alpha+s}$ 为常数.

原理. 这种变换是射影几何中的坐标变换. 把两个不动点 α,β 映射为 0 和 ∞, 非线性迭代就化为简单的乘法迭代(等比数列).

7.11.2.4 有唯一重根不动点 α

此时 $\alpha=\beta$, 上述比值变换失效 (分母为零) . 此时不动点 α 满足 $r\alpha^2+(s-p)\alpha-q=0$, 且判别式为零, 即 $\alpha=\frac{p-s}{2r}$. 我们考察倒数变换

$$\frac{1}{a_{n+1}-\alpha}=\frac{ra_n+s}{(pa_n+q)-\alpha(ra_n+s)}$$

利用 α 的代数性质化简分母, 最终可得

$$\frac{1}{a_{n+1}-\alpha}=\frac{1}{a_n-\alpha}+\frac{r}{p-r\alpha}$$

令 $b_n=\frac{1}{a_n-\alpha}$. 上式即为 $b_{n+1}=b_n+d$.

原理. 当两个不动点重合时, 几何上的旋转/缩放退化为平移. 倒数变换把原数列化为等差数列. 这对应矩阵的若尔当标准型理论. 特征值相等时无法对角化, 只能化为若尔当块, 对应线性增长(等差数列).

7.11.2.5 线性化的一般框架

利用不动点求通项公式, 是动力系统线性化的一个例子.

1. **寻找平衡位置.** 解方程 $f(x)=x$ 找到系统的不动点.
2. **建立新坐标系.** 以不动点为参照, 定义新数列 b_n, 如 a_n-c 或 $\frac{a_n-\alpha}{a_n-\beta}$.
3. **线性化迭代.** 在新坐标系下, 复杂的迭代关系退化为最基本的线性关系 ($b_{n+1}=kb_n$ 或 $b_{n+1}=b_n+d$).
4. **还原.** 利用等比/等差数列公式求出 b_n, 再逆变换回 a_n.

压缩映射原理

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.11.3 压缩映射原理

知乎@余命数

有些迭代形式 $x_{n+1}=f(x_n)$ 会发散. 迭代是否收敛取决于函数 f 是否压缩空间. 波兰数学家巴拿赫 (Stefan Banach) 将其形式化为**压缩映射原理**(巴拿赫不动点定理), 是数值分析中迭代法收敛性的基础.

定义 7.15 (压缩映射) 设 f 定义在区间 $[a,b]$ 上, 且其值域仍在 $[a,b]$ 内. 若存在常数 $0\le k<1$, 使得对于区间内任意 x,y, 都有

$$|f(x)-f(y)|\le k|x-y|$$

则称 f 为该区间上的一个**压缩映射**, 常数 k 称为压缩系数.

直观上, 压缩映射使得任意两点在经过函数作用后, 其距离变得更近.

定理 7.10 (压缩映射原理) 若 f 是闭区间 $[a,b]$ 上的压缩映射, 则

1. f 在 $[a,b]$ 内存在且仅存在一个不动点 x^*.
2. 对于任意初始值 $x_0\in[a,b]$, 迭代序列 $x_{n+1}=f(x_n)$ 均收敛于 x^*.
3. 误差满足估计式 $|x_n-x^*|\le\frac{k^n}{1-k}|x_1-x_0|$.

知乎@余命数

证明 任取 $x_0\in[a,b]$, 定义迭代 $x_{n+1}=f(x_n)$. 因为 $f([a,b])\subseteq[a,b]$, 所以所有 x_n 都属于 $[a,b]$.

知乎@余命数

由压缩条件,

$$|x_{n+1}-x_n|=|f(x_n)-f(x_{n-1})|\le k|x_n-x_{n-1}|\le k^n|x_1-x_0|.$$

因而当 $m>n$ 时,

知乎@余命数

$$|x_m-x_n|\le\sum_{j=n}^{m-1}|x_{j+1}-x_j|\le\sum_{j=n}^{m-1}k^j|x_1-x_0|\le\frac{k^n}{1-k}|x_1-x_0|.$$

知乎@余命数

这说明 $\{x_n\}$ 是柯西数列. 闭区间 $[a,b]$ 完备, 故存在 $x^*\in[a,b]$ 使 $x_n\to x^*$.

压缩映射满足 Lipschitz 条件, 因而连续. 对 $x_{n+1}=f(x_n)$ 两边取极限, 得 $x^*=f(x^*)$, 所以不动点存在.

若 u,v 都是不动点, 则

$$|u-v|=|f(u)-f(v)|\le k|u-v|.$$

由于 $0\le k<1$, 只能有 $|u-v|=0$, 即 $u=v$. 不动点唯一.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在上面的估计中令 $m \to \infty$, 即得

$$|x_n - x^*| \le \frac{k^n}{1-k}|x_1 - x_0|.$$

因此任意初始值产生的迭代序列都收敛到唯一不动点, 且满足所给误差估计. □

若函数 f 可导, 根据拉格朗日中值定理, $|f(x)-f(y)| = |f'(\xi)||x-y|$. 因此, f 在区间内恒有 $|f'(x)| \le k < 1$ 是它成为压缩映射的一个充分条件.

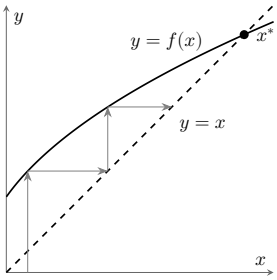


图 7.3 迭代过程的几何表示(蛛网图). 数列 $\{x_n\}$ 在 $y = x$ 与 $y = f(x)$ 之间折返, 逼近交点, 即不动点.

反之, 若 $f(x)$ 的图像与 $y = x$ 不相交, 则不动点不存在, 迭代序列无从收敛. 取 $f(x) = x + \frac{1}{2}$, 其图像是 $y = x$ 的一条平行线, 二者永不相交; 迭代 $x_{n+1} = x_n + \frac{1}{2}$ 的蛛网呈阶梯状持续上移, 远离任何可能的极限.

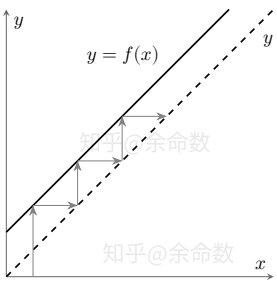
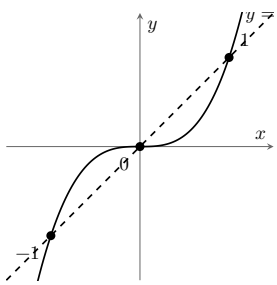


图 7.4 不动点不存在的情形, 如 $f(x) = x + \frac{1}{2}$ 与 $y = x$ 平行, 无交点, 迭代序列发散.

不动点还可以不唯一. 当 $f(x)$ 的图像与 $y = x$ 多次相交时, 每个交点都是不动点. 取 $f(x) = x^3$, 它与 $y = x$ 在 $x = -1, 0, 1$ 三处相交, 故有三个不动点. 此时迭代序列收敛到哪一个, 取决于初始值 x_0 的选取, 从 $x_0 \in (0,1)$ 出发迭代趋于 0, 从 $x_0 > 1$ 出发迭代趋于 $+\infty$, 不动点 $x^* = 1$ 虽存在却非吸引点.



知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

图 7.5　不动点不唯一的情形, 如 $f(x)=x^3$ 与 $y=x$ 有三个交点 $x^*=-1,0,1$, 迭代收敛与否取决于初始值.

即便不动点存在且迭代收敛, 收敛速度也未必快. 收敛的快慢取决于 f 在不动点处的斜率.$|f'(x^*)|$ 越接近 1, 每步缩减的误差越小, 所需迭代步数越多. 取 $f(x)=-0.9x+1$, 不动点 $x^*=\frac{10}{19}\approx 0.526$, 此处 $|f'(x^*)|=0.9$. 从 $x_0=0$ 出发, 蛛网在 $y=x$ 与 $y=f(x)$ 之间反复折返, 锯齿幅度每步仅缩减一成, 数十步后方才逼近交点.

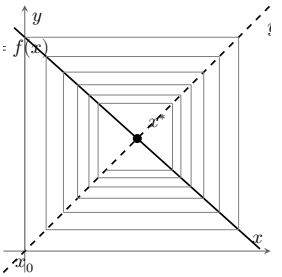


图 7.6　收敛缓慢的情形, 如 $f(x)=-0.9x+1$, $|f'(x^*)|=0.9$ 接近 1, 蛛网锯齿幅度每步仅缩减一成, 需数十步方能逼近不动点.

7.11.4 巴比伦算法

在计算器出现之前, 古巴比伦人就用一种基于迭代的算法计算 $\sqrt{S}$ 的值.

考虑方程 $x^2=S$ $(S>0)$. 将其改写为 $x=\frac{S}{x}$. 若 x_n 小于 $\sqrt{S}$, 则 $\frac{S}{x_n}$ 必然大于 $\sqrt{S}$; 反之亦然. $\sqrt{S}$ 位于 x_n 与 $\frac{S}{x_n}$ 之间. 取二者的算术平均值作为下一次的近似, 应当比 x_n 更接近真实值.

由此得到**巴比伦迭代公式**.

$$x_{n+1}=\frac{1}{2}\left(x_n+\frac{S}{x_n}\right)$$

巴比伦算法收敛的导数讨论

例题 7.76　利用巴比伦算法计算 $\sqrt{2}$ 的近似值, 取初始值 $x_0=1$.

解　代入公式 $S=2$:

$$x_1=\frac{1}{2}\left(1+\frac{2}{1}\right)=1.5$$

$$x_2=\frac{1}{2}\left(1.5+\frac{2}{1.5}\right)=\frac{1}{2}\left(\frac{3}{2}+\frac{4}{3}\right)=\frac{17}{12}\approx 1.41666\ldots$$

$$x_3=\frac{1}{2}\left(\frac{17}{12}+\frac{2}{17/12}\right)=\frac{1}{2}\left(\frac{17}{12}+\frac{24}{17}\right)=\frac{577}{408}\approx 1.4142156\ldots$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

已知 $\sqrt{2}\approx 1.41421356$. 仅迭代三次, 误差已小于 10^{-5}.

考察该迭代的收敛性. 设函数 $f(x)=\frac{1}{2}(x+\frac{S}{x})$. 不动点方程为 $x=\frac{1}{2}(x+\frac{S}{x}) \implies 2x=x+\frac{S}{x} \implies x=\frac{S}{x} \implies x^2=S$. 不动点确为 $\sqrt{S}$.

计算导数, 得

$$f'(x)=\frac{1}{2}\left(1-\frac{S}{x^2}\right)$$

在不动点 $x^*=\sqrt{S}$ 处,

$$f'(\sqrt{S})=\frac{1}{2}\left(1-\frac{S}{(\sqrt{S})^2}\right)=0$$

由于 $|f'(\sqrt{S})|=0<1$, 根据收敛原理, 该迭代在 $\sqrt{S}$ 附近收敛很快(二阶收敛).

牛顿迭代法

7.11.5 牛顿迭代法

函数一章已经介绍过这个迭代. 我们指出, 巴比伦算法是**牛顿迭代法**的一个特例.

计算机求解方程 $f(x)=0$ 时通常使用迭代法.

牛顿法的几何直观是用切线逼近曲线. 在点 $(x_n,f(x_n))$ 处作曲线 $y=f(x)$ 的切线. 切线方程为

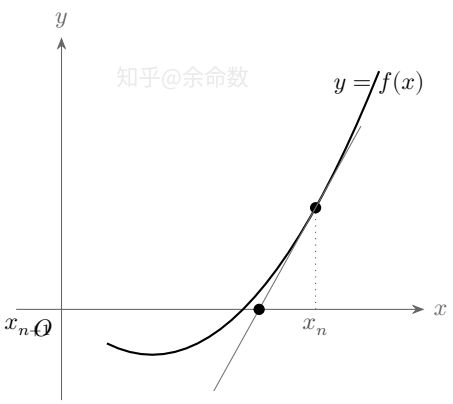
$$y-f(x_n)=f'(x_n)(x-x_n)$$

令 $y=0$, 求出切线与 x 轴的交点, 作为下一次的近似值 x_{n+1}:

$$-f(x_n)=f'(x_n)(x_{n+1}-x_n)$$

整理得牛顿迭代公式.

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \tag{7-10}$$



知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

图 7.7　牛顿迭代法的几何意义(利用切线的根作为曲线根的近似).

例题 7.77　利用牛顿法推导求 k 次方根 $\sqrt[k]{S}$ 的迭代公式.

解　求 $\sqrt[k]{S}$ 等价于求解方程 $x^k - S = 0$. 设 $f(x) = x^k - S$, 则 $f'(x) = kx^{k-1}$. 代入牛顿迭代公式,

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^k - S}{kx_n^{k-1}} \\ &= x_n - \frac{1}{k}\left(\frac{x_n^k}{x_n^{k-1}} - \frac{S}{x_n^{k-1}}\right) \\ &= x_n - \frac{1}{k}x_n + \frac{1}{k}\frac{S}{x_n^{k-1}} \\ &= \frac{k-1}{k}x_n + \frac{1}{k}\frac{S}{x_n^{k-1}} \end{aligned}$$

当 $k = 2$ 时, 公式变为 $x_{n+1} = \frac{1}{2}x_n + \frac{1}{2}\frac{S}{x_n}$, 这正是巴比伦算法.

7.12 数列的单调性

观察一个数列, 除了关心各项的数值, 我们还要看它的整体趋势. 随着 n 增大, 数列有的持续增大, 有的持续减小, 有的上下波动, 描述这种增减趋势的数学语言就是**单调性**.

这个概念与函数的单调性定义相同, 都刻画取值随变量的增减变化. 函数的单调性描述因变量如何随自变量连续变化, 数列的单调性则描述数列的项如何随离散的整数变量 n 变化. 判断并证明数列的单调性, 是分析其收敛性、估计其界限、解决相关不等式问题的前提.

注 7.10
用连续函数导数判别数列单调性属分析工具, 见 第16章.

7.12.1 定义

定义 7.16 (单调数列) 设有一个数列 $\{a_n\}$.

- 如果对于任意正整数 n, 恒有 $a_{n+1} > a_n$, 那么称数列 $\{a_n\}$ 为**递增数列**.
- 如果对于任意正整数 n, 恒有 $a_{n+1} < a_n$, 那么称数列 $\{a_n\}$ 为**递减数列**.
- 如果对于任意正整数 n, 恒有 $a_{n+1} \geq a_n$, 那么称数列 $\{a_n\}$ 为**非减数列**.
- 如果对于任意正整数 n, 恒有 $a_{n+1} \leq a_n$, 那么称数列 $\{a_n\}$ 为**非增数列**.

递增数列和递减数列统称为**严格单调数列**. 非减数列和非增数列统称为**单调数列**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

7.12.2 判断方法

判断一个数列的单调性,实质上是比较相邻两项 a_{n+1} 与 a_n 的大小关系. 根据通项公式的结构特征,常用以下几种方法.

7.12.2.1 作差法

作差法直接套用定义. 比较相邻两项的差 $d_n = a_{n+1} - a_n$,判断差值的符号.

- 若对于一切 $n \in \mathbb{N}^*$,恒有 $a_{n+1} - a_n > 0$,则 $\{a_n\}$ 是递增数列.
- 若对于一切 $n \in \mathbb{N}^*$,恒有 $a_{n+1} - a_n < 0$,则 $\{a_n\}$ 是递减数列.

对于非严格单调的情况,只需将不等号换为 $\geq$ 或 $\leq$.

例题 7.78 判断数列 $\left\{\dfrac{n}{2n+1}\right\}$ 的单调性.

解 解析: 设 $a_n = \dfrac{n}{2n+1}$. 考察相邻两项之差.

$$\begin{aligned} a_{n+1} - a_n &= \frac{n+1}{2(n+1)+1} - \frac{n}{2n+1} \\ &= \frac{n+1}{2n+3} - \frac{n}{2n+1} \\ &= \frac{(n+1)(2n+1) - n(2n+3)}{(2n+3)(2n+1)} \\ &= \frac{(2n^2+3n+1) - (2n^2+3n)}{(2n+3)(2n+1)} \\ &= \frac{1}{(2n+3)(2n+1)} \end{aligned}$$

由于对于任意正整数 n,分母 $(2n+3)(2n+1)$ 恒为正数. 因此,$a_{n+1} - a_n = \dfrac{1}{(2n+3)(2n+1)} > 0$ 恒成立. 故数列 $\left\{\dfrac{n}{2n+1}\right\}$ 是一个递增数列.

7.12.2.2 作商法

当数列的各项均为正数时,比较 a_{n+1} 与 a_n 的大小等价于比较它们的比值 $\dfrac{a_{n+1}}{a_n}$ 与 1 的大小. 这种方法尤其适用于通项中含有指数、阶乘或连乘积形式的数列. 设数列 $\{a_n\}$ 满足 $a_n > 0$.

- 若对于一切 $n \in \mathbb{N}^*$,恒有 $\dfrac{a_{n+1}}{a_n} > 1$,则 $\{a_n\}$ 是递增数列.
- 若对于一切 $n \in \mathbb{N}^*$,恒有 $\dfrac{a_{n+1}}{a_n} < 1$,则 $\{a_n\}$ 是递减数列.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 7.79　判断数列 $\left\{\dfrac{3^n}{n!}\right\}$ 的单调性.

知乎@余命数

解　解析：设 $a_n=\dfrac{3^n}{n!}$. 显然对于所有 $n\in\mathbb{N}^*$, $a_n>0$, 故可以使用作商法. 考察相邻两项之比.

知乎@余命数

$$\begin{aligned}\frac{a_{n+1}}{a_n}&=\frac{3^{n+1}/(n+1)!}{3^n/n!}\\&=\frac{3^{n+1}}{(n+1)!}\cdot\frac{n!}{3^n}\\&=\frac{3^{n+1}}{3^n}\cdot\frac{n!}{(n+1)n!}\\&=3\cdot\frac{1}{n+1}\\&=\frac{3}{n+1}\end{aligned}$$

知乎@余命数

需比较这个比值与 1 的大小关系. 当 $n=1$ 时, $\dfrac{a_2}{a_1}=\dfrac{3}{2}>1$. 当 $n=2$ 时, $\dfrac{a_3}{a_2}=\dfrac{3}{3}=1$. 当 $n\geq 3$ 时, $n+1\geq 4$, 此时 $\dfrac{3}{n+1}<1$.

知乎@余命数

因此, $a_2>a_1$ 且 $a_3=a_2$, 而对于 $n\geq 3$, 有 $a_{n+1}<a_n$. 该数列不是一个单调数列. 它先增后不变, 然后开始递减. 具体地, $a_1=3, a_2=4.5, a_3=4.5, a_4=3.375,\ldots$.

7.12.2.3　函数法

如果数列的通项 a_n 可以看作是某个函数 $f(x)$ 在正整数点上的取值, 即 $a_n=f(n)$, 那么函数 $f(x)$ 的单调性便可以为我们提供关于数列 $\{a_n\}$ 单调性的相关信息. 若函数 $f(x)$ 在区间 $[1,+\infty)$ 上是单调递增, 或递减的, 那么数列 $\{a_n\}$ 也必然是单调递增, 或递减的.

函数法/导数判别数列单调

知乎@余命数

在实践中, 我们通常利用导数来判断函数 $f(x)$ 的单调性.

知乎@余命数

注 7.11 (注意)
我们只关心函数在整数点上的行为. 因此, 即使函数 $f(x)$ 不是在整个 $[1,+\infty)$ 区间上单调, 只要它在某个区间 $[N,+\infty)$(N 为正整数)上单调, 我们就可以断定数列 $\{a_n\}$ 从第 N 项开始是单调的.

知乎@余命数

例题 7.80　判断数列 $\left\{\dfrac{\ln n}{n}\right\}$ 的单调性.

解 解析: 设 $a_n = \dfrac{\ln n}{n}$. 考虑构造函数 $f(x) = \dfrac{\ln x}{x}$ for $x \geq 1$. 利用导数来研究其单调性.

$$f'(x) = \frac{(\ln x)' \cdot x - \ln x \cdot (x)'}{x^2} = \frac{\frac{1}{x} \cdot x - \ln x}{x^2} = \frac{1 - \ln x}{x^2}$$

令 $f'(x) = 0$, 得 $1 - \ln x = 0$, 解得 $x = e$. 当 $1 \leq x < e$ 时, $\ln x < 1$, 故 $1 - \ln x > 0$, $f'(x) > 0$. 当 $x > e$ 时, $\ln x > 1$, 故 $1 - \ln x < 0$, $f'(x) < 0$.

故函数 $f(x)$ 在 $(0, e)$ 上单调递增, 在 $(e, +\infty)$ 上单调递减.

由于 $e \approx 2.718$, 需考察整数点 $n = 1, 2, 3$ 附近的行为. $a_1 = \dfrac{\ln 1}{1} = 0$. $a_2 = \dfrac{\ln 2}{2} \approx \dfrac{0.693}{2} = 0.3465$. $a_3 = \dfrac{\ln 3}{3} \approx \dfrac{1.098}{3} = 0.366$. $a_4 = \dfrac{\ln 4}{4} = \dfrac{2\ln 2}{4} = \dfrac{\ln 2}{2} \approx 0.3465$.

有 $a_1 < a_2 < a_3$, 而 $a_3 > a_4$. 由于函数 $f(x)$ 在 $x > e$ 后单调递减, 那么对于所有满足 $n \geq 3$ 的整数, 必然有 $a_{n+1} = f(n+1) < f(n) = a_n$. 因此, 该数列从第 3 项开始是严格递减的, 但它不是一个整体上的单调数列.

注 7.12 (单调性判断方法总结)

- **作差法:** 通用性最强, 是定义的直接体现. 适用于大多数多项式或分式形式的通项.
- **作商法:** 适用于各项为正, 且通项包含指数、阶乘等乘法结构的情况.
- **函数法:** 通项可延拓为连续函数时可用; **求导判别** 见 第16章.

选择哪种方法, 取决于通项 a_n 的结构.

7.13 奇偶数列 *

整数的奇偶性是其最基本的数论属性. 当将研究对象从单个整数扩展到按规律排列的整数序列时, 一个自然的问题是, 数列中各项的奇偶性是否也遵循某种规律.

研究数列的奇偶性, 相当于在模 2 的算术体系下看数列. 这种视角转换常能把复杂的递推关系化简为简单、甚至周期性的新关系. 奇偶性分析除了用于判断项的基本性质, 还用于解决数论问题, 例如证明某些项不可能是完全平方数. 数列在二元结构中常表现出周期性.

7.13.1 定义与判断

对于一个各项均为整数的数列 $\{a_n\}$, 可根据其项的奇偶性对其进行分类, 并引出其伴随的奇偶性序列.

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 7.17 (奇数列与偶数列) 设 $\{a_n\}$ 是一个整数数列.

- 如果对于任意正整数 n, 数列的项 a_n 均为奇数, 则称 $\{a_n\}$ 为**奇数列**.
- 如果对于任意正整数 n, 数列的项 a_n 均为偶数, 则称 $\{a_n\}$ 为**偶数列**.

判断一个数列的奇偶性, 主要有两种途径. 一种是从通项公式入手, 另一种是从递推关系入手.

7.13.1.1 基于通项公式的分析

如果数列的通项公式 $a_n = f(n)$ 已知, 可直接分析函数值 $f(n)$ 在 n 取遍正整数时的奇偶性. 这通常需要讨论 n 的奇偶性.

例题 7.81 判断数列 $a_n = n^2 + 3n + 5$ 的奇偶性.

解 解析: 将通项进行因式分解和重组, 以说明其结构:

$$a_n = n(n+3) + 5$$

知乎@余命数　知乎@余命数

对项数 n 的奇偶性讨论.

知乎@余命数

- **若 n 为偶数**, 设 $n = 2k$ ($k \in \mathbb{N}^*$). 则 n 是偶数, 故 $n(n+3)$ 是偶数. $a_n =$ (偶数) + 5 = (偶数) + (奇数) = (奇数).
- **若 n 为奇数**, 设 $n = 2k - 1$ ($k \in \mathbb{N}^*$). 则 $n + 3$ 是偶数, 故 $n(n+3)$ 是偶数. $a_n =$ (偶数) + 5 = (偶数) + (奇数) = (奇数).

综合两种情况, 对于所有正整数 n, a_n 均为奇数. 因此, $\{a_n\}$ 是一个奇数列.

7.13.1.2 基于递推关系的分析

当数列由递推关系定义时, 直接读出每一项的性质并不容易. 但可以分析奇偶性如何通过递推关系**遗传**下去. 这需要模 2 的同余算术. 基本运算法则如下:

知乎@余命数

- 奇 $\pm$ 奇 = 偶, $1 + 1 \equiv 0 \pmod 2$
- 奇 $\pm$ 偶 = 奇, $1 + 0 \equiv 1 \pmod 2$
- 偶 $\pm$ 偶 = 偶, $0 + 0 \equiv 0 \pmod 2$
- 奇 $\times$ 奇 = 奇, $1 \times 1 \equiv 1 \pmod 2$
- 奇 $\times$ 偶 = 偶, $1 \times 0 \equiv 0 \pmod 2$
- 偶 $\times$ 偶 = 偶, $0 \times 0 \equiv 0 \pmod 2$

知乎@余命数

通过将原递推关系中的每一项都替换为其模 2 的余数(奇数 $\to 1$, 偶数 $\to 0$), 可得到一个关于奇偶性的新递推关系.

例题 7.82 分析斐波那契数列 $\{F_n\}$ $(F_1=1, F_2=1)$ 的各项奇偶性规律.

知乎@余命数

解 解析: 斐波那契数列的递推关系为 $F_{n+2}=F_{n+1}+F_n$. 考察其各项模 2 的余数, 设 $b_n \equiv F_n \pmod 2$. 则 $\{b_n\}$ 满足递推关系 $b_{n+2} \equiv b_{n+1}+b_n \pmod 2$.

从初始项开始计算这个新的奇偶性序列, 得 $F_1=1 \implies b_1=1$, 为奇 $F_2=1 \implies b_2=1$, 为奇 $b_3 \equiv b_2+b_1=1+1 \equiv 0$, 为偶 $b_4 \equiv b_3+b_2=0+1 \equiv 1$, 为奇 $b_5 \equiv b_4+b_3=1+0 \equiv 1$, 为奇 $b_6 \equiv b_5+b_4=1+1 \equiv 0$, 为偶

序列 $\{b_n\}$ 的前几项为 $1,1,0,1,1,0,\ldots$. 观察到, 序列以奇、奇、偶的模式循环出现, 其周期为 3. 因此, 斐波那契数列中 F_n 是偶数, 当且仅当项数 n 是 3 的倍数.

例题 7.83 已知数列 $\{a_n\}$ 满足 $a_1=1, a_{n+1}=5a_n+8$. 证明 $\{a_n\}$ 是一个奇数列.

解 证明: 使用数学归纳法, 并结合奇偶性分析. 设命题 $P(n)$ 为"a_n 是一个奇数".

奠基步骤: 当 $n=1$ 时, $a_1=1$, 是一个奇数. 故 $P(1)$ 成立.

归纳步骤: 假设当 $n=k$ $(k \geq 1)$ 时命题成立, 即 a_k 是一个奇数. 需证明当 $n=k+1$ 时, a_{k+1} 也是一个奇数.

知乎@余命数

考察递推关系 $a_{k+1}=5a_k+8$. 根据归纳假设, a_k 是奇数. 因此:

$$\begin{aligned} a_{k+1} &= (\text{奇数}) \times a_k + (\text{偶数}) \\ &= (\text{奇数}) \times (\text{奇数}) + (\text{偶数}) \\ &= (\text{奇数}) + (\text{偶数}) \\ &= (\text{奇数}) \end{aligned}$$

这证明了 $P(k+1)$ 成立.

或者, 可使用更为形式化的同余语言. $a_{n+1} \equiv 5a_n+8 \pmod 2$. 由于 $5 \equiv 1 \pmod 2$ 且 $8 \equiv 0 \pmod 2$, 上式化为:

知乎@余命数

$$a_{n+1} \equiv 1 \cdot a_n + 0 \pmod 2 \implies a_{n+1} \equiv a_n \pmod 2$$

故后一项的奇偶性与前一项完全相同. 由于首项 $a_1=1$ 是奇数,即 $a_1 \equiv 1 \pmod 2$. 因此, 对于所有正整数 n, 都有 $a_n \equiv a_1 \equiv 1 \pmod 2$. 即 $\{a_n\}$ 的每一项都是奇数.

根据数学归纳法原理, 命题得证.

7.14 数列的周期性

知乎@余命数

来龙去脉

数列的项在经过固定步长后开始重复取值, 这种现象称为数列的周期性.

研究周期性直接通向数列的代数结构. 当递推关系涉及模运算或有限集合上的运算时, 周期性几乎必然出现. 判断周期性的方法常用于解决数论与组合问题, 也用于处理项数过大、无法逐项计算的数列问题.

7.14.1 定义与性质

先对数列的周期性给出严格的数学定义.

定义 7.18 (周期数列) 对于数列 $\{a_n\}$, 如果存在一个正整数 T, 使得对于所有大于或等于某个正整数 n_0 的 n, 恒有等式 $a_{n+T}=a_n$ 成立, 则称数列 $\{a_n\}$ 为**周期数列**.

周期数列的定义中满足条件的正整数 T 不止一个, 需要区分周期与最小正周期.

知乎@余命数

定义 7.19 (周期与最小正周期) 在上述定义中, 满足条件的任意一个正整数 T 都被称为数列 $\{a_n\}$ 的一个**周期**.
在所有的周期中, 最小的那个正整数被称为该数列的**最小正周期** (或主周期). 没有特别说明时, 数列的周期指的就是最小正周期.

一个直接的推论是, 若 T 是数列的最小正周期, 则任何形如 kT ($k\in\mathbb{N}^*$) 的整数都是该数列的周期. 周期数列最重要的性质是, 其所有项的取值落在一个有限集合内; 长度为 T 的一段取值重复出现, 即可确定整个数列.

反过来, 任一周期都是最小正周期的整数倍.

命题 7.1 (最小正周期的整除性) 若 T 是数列 $\{a_n\}$ 的最小正周期, T' 是它的任一周期, 则 $T\mid T'$.

证明 由带余除法, $T'=qT+r$, 其中 $0\le r<T$. 因 T 是周期, qT 也是周期, 故 $a_{n+qT}=a_n$; 又 T' 是周期, $a_{n+T'}=a_n$. 于是

$$a_{n+r}=a_{n+T'-qT}=a_{n+T'}=a_n,$$

即 r 也是周期. 但 $0\le r<T$ 且 T 是最小正周期, 故 $r=0$, 即 $T\mid T'$. □

周期数列的取值被一个长度为 T 的循环块完全决定, 由此得到两条常用性质.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

命题 7.2 (周期求和) 若数列 $\{a_n\}$ 以 T 为周期, 则任意连续 T 项之和为定值, 即对任意正整数 n,

$$\sum_{k=0}^{T-1} a_{n+k} = \sum_{k=1}^{T} a_k.$$

证明 记 $S_n = \sum_{k=0}^{T-1} a_{n+k}$. 把求和起点从 n 平移到 $n+1$, 得

$$S_{n+1} = \sum_{k=0}^{T-1} a_{n+1+k} = \sum_{k=1}^{T} a_{n+k} = S_n - a_n + a_{n+T}.$$

由 $a_{n+T} = a_n$ 知 $S_{n+1} = S_n$, 故 S_n 与 n 无关. □

知乎@余命数

命题 7.3 (和与积的周期) 若 $\{a_n\}$ 以 T_1 为周期, $\{b_n\}$ 以 T_2 为周期, 则 $\{a_n+b_n\}$ 与 $\{a_n b_n\}$ 均以 $L = \operatorname{lcm}(T_1, T_2)$ 为周期; 两者的最小正周期都整除 L.

知乎@余命数

证明 L 同时是 T_1 与 T_2 的倍数, 故

$$a_{n+L} + b_{n+L} = a_n + b_n, \qquad a_{n+L}\, b_{n+L} = a_n\, b_n.$$

即 L 是和数列与积数列的周期. 由 命题 7.1, 最小正周期整除任一周期, 故整除 L. □

知乎@余命数 知乎@余命数

还有一种更广义的最终周期性, 高中阶段较少涉及.

定义 7.20 (最终周期数列) 如果一个数列从某一项之后才开始呈现周期性, 即存在正整数 N 和 T, 使得对于所有 $n > N$, 恒有 $a_{n+T} = a_n$ 成立, 则称该数列为**最终周期数列**.
数列的前 N 项 $\{a_1, a_2, \ldots, a_N\}$ 构成了数列的**预周期**部分.

知乎@余命数 知乎@余命数 余命数

7.14.2 周期性的判定

判定由递推关系定义的数列是否具有周期性, 关键在于考察系统状态的演化. 对于一个 k 阶递推关系 $a_{n+k} = f(a_{n+k-1}, \ldots, a_n)$, 第 n 步的状态可以用 k 维向量 $\mathbf{v}_n = (a_n, a_{n+1}, \ldots, a_{n+k-1})$ 完全描述. 若某一步的状态 $\mathbf{v}_m$ 与之前某一步的状态 $\mathbf{v}_l$ $(l < m)$ 完全相同, 由递推关系的确定性可知, 此后各项将重复之前的演化路径, 数列进入周期.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这一观察与同余理论结合后尤为有效. 如果一个整数数列 $\{a_n\}$ 在模 m 的意义下被考察, 那么每一项 $a_n \pmod m$ 的取值只有 m 种可能. 相应地, k 维状态向量 $\mathbf{v}_n$ 的每个分量也只有 m 种可能, 总的状态数不超过 m^k 种. 根据**鸽巢原理**, 在至多 m^k+1 次状态转移后, 必然有两个状态完全相同. 我们由此得到一个有效的结论.

定理 7.11　设整数数列 $\{a_n\}$ 由 k 阶整系数线性递推关系

$$a_{n+k} = c_{k-1}a_{n+k-1} + \cdots + c_1 a_{n+1} + c_0 a_n$$

定义. 对任意正整数 m, 其模 m 的余数数列必然是最终周期数列; 等价地, 存在正整数 l 与 T, 使得从第 l 个状态开始按周期 T 重复.

证明　在模 m 的意义下, 状态向量

$$\mathbf{v}_n = (a_n, a_{n+1}, \dots, a_{n+k-1}) \pmod m$$

的每个分量只有 m 种可能, 因此状态总数不超过 m^k 个. 递推关系的系数为整数, 所以在模 m 下, 下一状态

$$\mathbf{v}_{n+1} = (a_{n+1}, \dots, a_{n+k})$$

由当前状态 $\mathbf{v}_n$ 唯一确定.

考察 m^k+1 个状态 $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{m^k+1}$. 根据鸽巢原理, 存在 $1 \le l < r \le m^k+1$, 使得 $\mathbf{v}_l = \mathbf{v}_r$. 由于后续演化由状态唯一决定, 归纳可得对所有 $j \ge 0$, 都有 $\mathbf{v}_{l+j} = \mathbf{v}_{r+j}$. 因此从第 l 个状态开始, 状态序列以 $T = r-l$ 为周期, 余数数列 $\{a_n \pmod m\}$ 也从某一项开始以 T 为周期. 它是最终周期数列. □

定理保证了最终周期性的存在. 具体求周期时, 逐项计算状态 $\mathbf{v}_n$ 直至首次出现重复 $\mathbf{v}_r = \mathbf{v}_l$ $(l<r)$, 则周期为 $r-l$; 对模 m 的 k 阶递推, 这样的重复在 m^k+1 个状态内必然出现. 最小正周期通常远小于上界 m^k, 须由计算确认.

例题 7.84　已知数列 $\{a_n\}$ 满足 $a_1 = 1$, $a_{n+1} = 2a_n \pmod 7$. 求 $\{a_n\}$ 的最小正周期, 并计算 $a_1 + a_2 + \cdots + a_{2024}$.

解　解析: 这是一阶递推, 状态即 a_n 本身, 取值只有 7 种. 从 $a_1 = 1$ 起逐项计算:

$$a_1 = 1, \quad a_2 = 2, \quad a_3 = 4, \quad a_4 = 8 \equiv 1 \pmod 7.$$

$a_4 = a_1$, 此后序列重复, 故最小正周期为 3, 循环块为 $(1, 2, 4)$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

一个周期内三项之和为 $1+2+4=7$. 由 命题 7.2, 任意连续 3 项之和都等于 7. 把 2024 项按每 3 项一组划分:

$$2024 = 3 \times 674 + 2.$$

前 2022 项恰为 674 个完整周期, 和为 $674 \times 7 = 4718$; 余下两项为 $a_1 = 1, a_2 = 2$. 故

$$a_1 + a_2 + \cdots + a_{2024} = 4718 + 1 + 2 = 4721.$$

例题 7.85 已知数列 $\{a_n\}$ 满足 $a_1 = 1, a_2 = 1$ 且 $a_{n+2} = a_{n+1} - a_n$. 分析其周期性.

解 解析: 从初始项逐项计算, 观察规律.

$$a_1 = 1$$

$$a_2 = 1$$

$$a_3 = a_2 - a_1 = 1 - 1 = 0$$

$$a_4 = a_3 - a_2 = 0 - 1 = -1$$

$$a_5 = a_4 - a_3 = -1 - 0 = -1$$

$$a_6 = a_5 - a_4 = -1 - (-1) = 0$$

$$a_7 = a_6 - a_5 = 0 - (-1) = 1$$

$$a_8 = a_7 - a_6 = 1 - 0 = 1$$

可见 $a_7 = a_1$ 且 $a_8 = a_2$. 由于这是二阶递推关系, 当连续两项重复时, 后续所有项也将随之重复.

该数列为 $1, 1, 0, -1, -1, 0, 1, 1, \ldots$. 其重复的周期块为 $(1, 1, 0, -1, -1, 0)$. 因此, 该数列是一个周期数列, 其最小正周期为 6.

注 7.13 (拓展旁注)
复单位根/复指数写法见 第20章.

注 7.14 (延伸思考)
该递推关系的特征方程为 $x^2 - x + 1 = 0$, 其根为 $x = \frac{1 \pm \sqrt{-3}}{2} = e^{\pm i\pi/3}$, 这是单位圆上的 6 次单位根. 这揭示了该数列周期为 6 的代数原因.

例题 7.86 设 $\{F_n\}$ 为斐波那契数列 ($F_1 = 1, F_2 = 1$). 求 F_{2024} 的个位数字.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 解析: 整数的个位数字等于它模 10 的余数. 考察斐波那契数列模 10 后的余数数列 $\{a_n\}$, 其中 $a_n \equiv F_n \pmod{10}$.

该数列满足递推关系 $a_{n+2} \equiv a_{n+1} + a_n \pmod{10}$, 初始项为 $a_1 = 1, a_2 = 1$. 状态由相邻两项构成的序对 (a_n, a_{n+1}) 决定, 这样的序对最多有 $10 \times 10 = 100$ 种. 由鸽巢原理, 该数列必在有限步后进入周期. 现通过计算确定周期.

数列 $\{a_n\}$ 的项为:

$$1,1,2,3,5,8,3,1,4,5,9,4,3,7,0,7,7,4,1,5,$$

$$6,1,7,8,5,3,8,1,9,0,9,9,8,7,5,2,7,9,6,5,$$

$$1,6,7,3,0,3,3,6,9,5,4,9,3,2,5,7,2,9,1,0,\ldots$$

当计算到第 61 项和第 62 项时, 得 $a_{61} \equiv a_{60} + a_{59} \equiv 0 + 1 = 1$ $a_{62} \equiv a_{61} + a_{60} \equiv 1 + 0 = 1$

此时状态序对 $(a_{61}, a_{62}) = (1, 1)$, 与初始状态 (a_1, a_2) 相同. 因此, 该数列的最小正周期为 60.

为了求 F_{2024} 的个位数字, 计算 2024 模 60 的余数.

$$2024 = 60 \times 33 + 44$$

因此, $a_{2024} = a_{44}$.

查阅已计算的序列, 第 44 项是 3. 故 F_{2024} 的个位数字是 3.

注 7.15 (周期性的分析路径)

- **存在性**. 整系数线性递推在模 m 下必为最终周期数列, 周期不超过 m^k(k 为阶数), 依据是鸽巢原理.
- **求周期**. 逐项计算状态直至首次重复 $\mathbf{v}_r = \mathbf{v}_l$, 周期即为 $r - l$; 最小正周期须实际算出, 通常远小于上界.
- **用性质**. 周期求和把长求和化为一个周期内的求和; 和与积的周期取 lcm, 用于由已知周期合成新数列的周期.

7.15 子列与 Bolzano-Weierstrass 定理 *

7.15.1 子列

从数列 $\{a_n\}$ 中按原有顺序选取一部分项, 得到的新数列称为子列.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定义 7.21 (子列) 设 $\{a_n\}$ 是一个数列. 若 $n_1 < n_2 < n_3 < \cdots$ 是一列严格递增的正整数, 则

$$a_{n_1}, a_{n_2}, a_{n_3}, \ldots$$

称为 $\{a_n\}$ 的一个**子列**, 记作 $\{a_{n_k}\}$.

下标 n_k 严格递增, 保证子列中各项的先后顺序与原数列一致. 取 $n_k = 2k-1$, 得到奇数项子列 $a_1, a_3, a_5, \ldots$; 取 $n_k = 2k$, 得到偶数项子列 $a_2, a_4, a_6, \ldots$.

子列继承了原数列的收敛性.

定理 7.12 (子列的极限) 若数列 $\{a_n\}$ 收敛于 L, 则它的任意子列 $\{a_{n_k}\}$ 也收敛于 L.

证明 设 $\lim_{n\to\infty} a_n = L$. 对任意 $\epsilon > 0$, 存在正整数 N, 使得当 $n > N$ 时, $|a_n - L| < \epsilon$.

严格递增的正整数序列满足 $n_k \geq k$. 取 $K = N$, 当 $k > K$ 时, $n_k \geq k > N$, 故 $|a_{n_k} - L| < \epsilon$.

因此 $\lim_{k\to\infty} a_{n_k} = L$. □

这个定理的逆否命题提供了一个判断数列发散的工具.

定理 7.13 (发散判据) 若数列 $\{a_n\}$ 有两个子列分别收敛于不同的极限, 则 $\{a_n\}$ 发散.

证明 设子列 $\{a_{n_k}\}$ 收敛于 L_1, $\{a_{m_k}\}$ 收敛于 L_2, 且 $L_1 \neq L_2$.

用反证法. 假设 $\{a_n\}$ 收敛于某个 L. 根据子列的极限定理, $\{a_{n_k}\}$ 和 $\{a_{m_k}\}$ 都应收敛于 L. 但它们分别收敛于 L_1 和 L_2, 故 $L = L_1$ 且 $L = L_2$, 与 $L_1 \neq L_2$ 矛盾.

因此 $\{a_n\}$ 发散. □

数列 $a_n = (-1)^n$ 是最直接的例子. 它的奇数项子列 $a_1, a_3, a_5, \cdots = -1, -1, -1, \ldots$ 收敛于 -1; 偶数项子列 $a_2, a_4, a_6, \cdots = 1, 1, 1, \ldots$ 收敛于 1. 两个子列极限不同, 故 $\{a_n\}$ 发散.

这个判据同样适用于函数极限. 在 第16.2节 中, 函数 $g(x) = \sin(1/x)$ 在 $x \to 0$ 时没有极限, 证明方法正是构造两条趋近于 0 的路径 $x_n = \dfrac{1}{2n\pi + \pi/2}$ 和 $y_n = \dfrac{1}{n\pi}$, 对应的函数值子列分别趋于 1 和 0. 两条路径给出不同的极限, 函数极限便不存在. 数列的子列判据与函数的路径判据在逻辑上完全一致.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

7.15.2 Bolzano-Weierstrass 定理

收敛数列的子列仍收敛, 但发散数列的子列未必发散. 一个有界的发散数列, 虽然整体没有极限, 却可能存在收敛的子列.

定理 7.14 (Bolzano–Weierstrass 定理) 任何有界数列都存在收敛的子列.

定理的证明采用二分法. 反复对分区间, 每次保留包含无穷多项的那一半, 由此构造一列嵌套收缩的区间, 最终锁定一个收敛子列的极限点.

证明 设 $\{a_n\}$ 是有界数列. 存在实数 a, b, 使得对所有 n, $a \le a_n \le b$.

将区间 $[a, b]$ 等分为两个子区间. 由于数列有无穷多项, 至少有一个子区间包含无穷多项. 记该子区间为 $I_1 = [a_1, b_1]$.

对 I_1 重复此操作, 等分并选取包含无穷多项的那一半, 记为 $I_2 = [a_2, b_2]$.

如此继续, 得到一列嵌套区间

$$[a, b] \supseteq I_1 \supseteq I_2 \supseteq \cdots$$

知乎@余命数

每个区间的长度是前一个的一半:

$$|I_k| = \frac{b-a}{2^k}$$

知乎@余命数

当 $k \to \infty$ 时, $|I_k| \to 0$.

嵌套区间的左端点序列 $\{a_k\}$ 非减且有上界 b, 右端点序列 $\{b_k\}$ 非增且有下界 a. 根据 第16.1节 中的完备性公理, $\{a_k\}$ 有上确界, $\{b_k\}$ 有下确界. 令

$$L = \sup\{a_k : k \ge 1\} = \inf\{b_k : k \ge 1\}$$

则 $a_k \le L \le b_k$ 对所有 k 成立. 由于 $|I_k| \to 0$, L 是嵌套区间的唯一公共点.

从每个区间 I_k 中选取一项 a_{n_k}, 要求下标严格递增. 这是可行的, 因为每个 I_k 包含无穷多项, 总能在其中找到下标大于 n_{k-1} 的项.

知乎@余命数

由于 $a_{n_k} \in I_k = [a_k, b_k]$, 有 $a_k \le a_{n_k} \le b_k$. 当 $k \to \infty$ 时, $a_k \to L$ 且 $b_k \to L$, 由夹逼准则得 $a_{n_k} \to L$.

知乎@余命数 知乎@余命数

因此 $\{a_{n_k}\}$ 是 $\{a_n\}$ 的一个收敛子列. □

Bolzano-Weierstrass 定理与完备性公理 (第16.1节) 等价. 在有理数系 $\mathbb{Q}$ 中, 有界数列未必有收敛子列, 例如 $\sqrt{2}$ 的十进制逼近序列 $1, 1.4, 1.41, \ldots$ 在 $\mathbb{Q}$ 中没有极限. 定理成立的前提正是实数轴没有空隙.

单调收敛定理要求数列单调且有界, Bolzano-Weierstrass 定理去掉了单调性条件, 代价是仅保证子列收敛, 整体未必收敛. 对于有界但振荡的数列 (如 $a_n = (-1)^n$), 整体极限不存在, 但奇数项子列和偶数项子列分别收敛.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

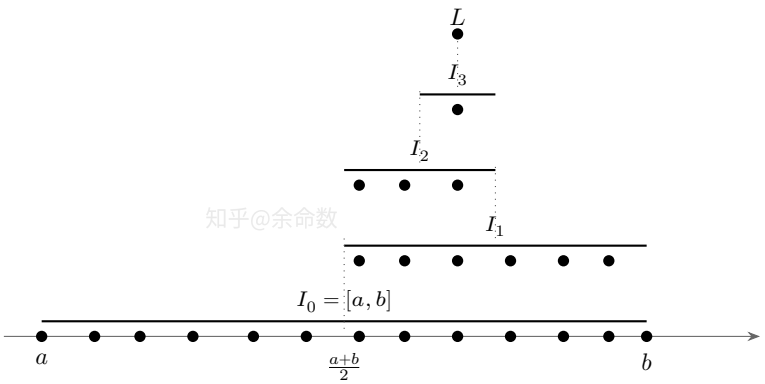


图 7.8 二分法构造收敛子列, 每一步保留包含无穷多项的半区间, 区间长度趋于零, 嵌套区间收缩到极限点 L.

注 7.16 (分析续篇的位置)
本章止于初等数列工具. 数列极限的 ε-N 语言、无穷级数, 以及泰勒公式与泰勒级数, 已迁入第16章(在极限定义与高阶导数之后). 等比公比 $|q| < 1$ 时各项越来越接近零、以及等比无穷和的图像, 可先由有限和的极限来理解. 严格证明见 第16章.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 8 章

统计

周期·离散·随机

「概率论不过是把常识化成计算。」

——拉普拉斯



知乎@余命数

第八章 统计

知乎@余命数

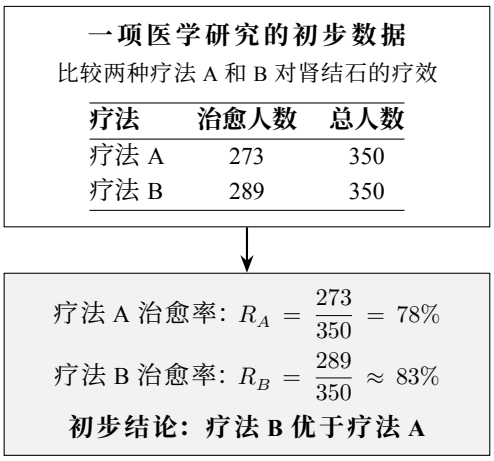


图 8.1 辛普森悖论示例(一):总体数据与初步结论

知乎@余命数

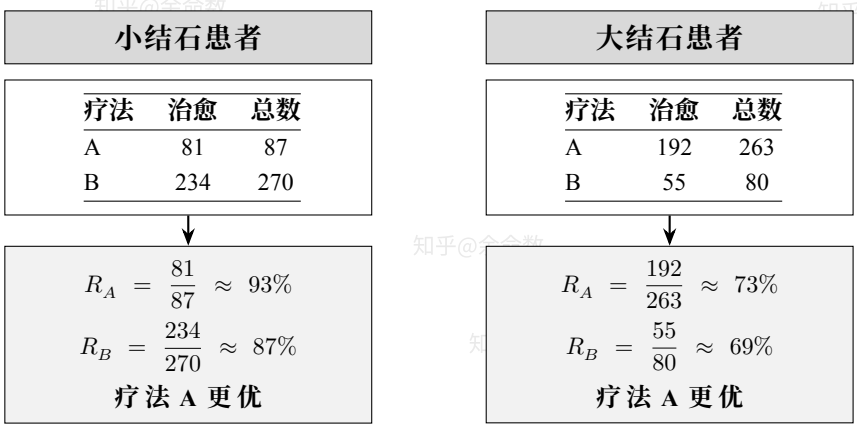


图 8.2 辛普森悖论示例(二):分组数据展现真相

知乎@余命数

辛普森悖论

我们来看一个思想实验. 某医疗中心汇总了疗法 A 与疗法 B 对肾结石的临床数据. 从整体来看, 疗法 B 的治愈率为 83%, 疗法 A 的治愈率为 78%. 只看这两个数字, 我们大概会推广疗法 B.

但是, 如果把病人按结石大小分成小结石与大结石两组, 结果就变了. 在小结石患者群中, 疗法 A 的治愈率高于疗法 B; 在大结石患者群中, 疗法 A 的治愈率也高于疗法 B. 结石大小是一个重要的潜在变量, 它改变了我们从汇总数据得出的结论.

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

一个疗法在每个分组都表现更好，合并后反而显得更差，这种现象称为**辛普森悖论**. 它说明，简单汇总可能掩盖数据内部的结构，从而扭曲事实. 数据本身不会说谎，但如果我们呈现和解读数据的方式不当，就会得到错误的结论.

为了避免这类陷阱，我们不能只满足于表面的数字. 我们需要一套系统的方法，来科学地收集数据，严谨地分析数据内部的结构，并最终做出可靠的推断. 这套方法就是**统计学**.

8.1 描述数据

我们来讨论描述性统计. 它的起点是一组原始数据，例如一个班级所有同学的身高记录. 这些数据在整理之前只是一些无序的数字. 我们的任务是把它们整理、提炼，使其蕴含的规律显现出来.

8.1.1 总体与样本

我们先界定研究对象的范围.

定义 8.1 (总体与样本)
- **总体：**我们所研究的对象的全体构成的集合. 例如，中国所有高三学生的身高.
- **个体：**构成总体的每一个对象. 例如，每一位高三学生的身高数据.
- **样本：**从总体中实际抽取并观测的一部分个体的集合. 例如，随机抽取的 1000 名高三学生的身高数据.
- **样本容量：**样本中所包含的个体数量，通常记为 n.

在大多数情况下，我们受成本、时间或可行性的限制，无法观测整个总体. 于是我们研究一个有代表性的**样本**，再由它推断**总体**的特征. 样本是我们手中可测的数据，总体是我们要推断的对象.

8.1.2 频率分布

拿到样本数据后，我们先了解数据的整体分布形态.

定义 8.2 (频率分布) 我们将样本数据分组，统计每组内数据出现的次数，称为**频数**，并计算频数占样本总量的比例，称为**频率**，然后用表格或图形把这种分布情况呈现出来，这就得到**频率分布**.

常用的图形有条形图（适用于离散或分类数据），以及适用于连续数据的**频率分布直方图**.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 8.1 某校随机抽取了 50 名学生, 测量其身高(单位 cm), 然后制作其频率分布表和直方图.

168, 175, 159, 171, 166, 182, 163, 170, 176, 169, 161, 173, 178, 165, 172, 157, 180, 167, 174, 169, 164, 171, 177, 162, 168, 170, 175, 166, 179, 158, 163, 172, 168, 174, 160, 173, 167, 171, 165, 181, 169, 176, 164, 170, 166, 172, 161, 175, 168, 173.

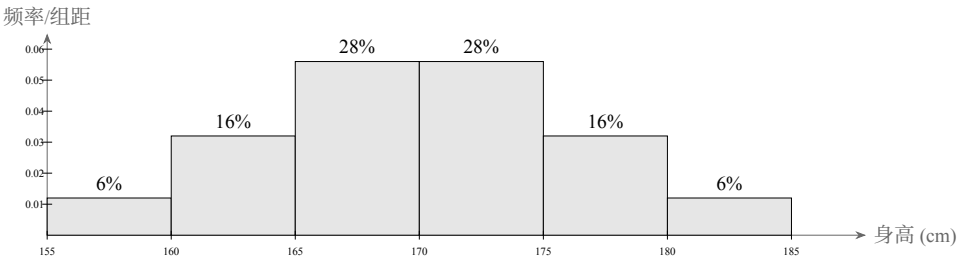


图 8.3 50 名学生身高频率分布直方图

知乎@余命数

解 这些数据的最大值为 182, 最小值为 157. 取组距为 5, 把数据分成 6 组, 得到频率分布表.

分组	频数	频率	频率/组距
$[155, 160)$	3	0.06	0.012
$[160, 165)$	8	0.16	0.032
$[165, 170)$	14	0.28	0.056
$[170, 175)$	14	0.28	0.056
$[175, 180)$	8	0.16	0.032
$[180, 185)$	3	0.06	0.012
合计	50	1.00	—

在直方图中, 每个矩形的高度正比于该组的频率, 纵轴通常标为频率/组距. 这样所有矩形的**面积之和**恰好为 1.

8.1.3 数据的数字特征

图形给出直观印象, 但若要精确比较和分析, 还需用具体数字概括样本的主要特征.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

8.1.3.1 衡量中心趋势

定义 8.3 (样本均值) 对于一个容量为 n 的样本 $\{x_1, x_2, \dots, x_n\}$, 其**样本均值**, 记为 $\bar{x}$, 定义为.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

样本均值即数据集中所有数值的算术平均, 它描述数据围绕的中心位置.

8.1.3.2 衡量离散程度

但仅知道中心位置还不够. 两组数据可以有相同的均值, 但围绕中心的分散程度可能不同. 为此我们需要一个度量离散程度的指标. 一个直接的办法是计算每个数据点到均值的平均距离 $\frac{1}{n}\sum|x_i - \bar{x}|$, 称为平均绝对离差. 绝对值在数学处理上不够便利, 因此统计学改用离差平方和来度量变异.

定义 8.4 (样本方差与样本标准差 (高中课本定义)) 对于样本 $\{x_1, \dots, x_n\}$ 及其均值 $\bar{x}$, 其**样本方差**, 记为 s^2, 定义为.

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

样本方差的算术平方根 $s = \sqrt{s^2}$ 称为**样本标准差**.

高中课本以 n 为分母定义样本方差, 统计学则采用 $n-1$ 为分母的定义.

定义 8.5 (样本方差与样本标准差) 对于样本 $\{x_1, \dots, x_n\}$ 及其均值 $\bar{x}$, 其**样本方差**, 记为 s^2, 定义为.

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

样本方差的算术平方根 $s = \sqrt{s^2}$ 称为**样本标准差**.

样本方差的分母是 $n-1$, 这常令人困惑. 既然样本均值是对 n 个数据求平均, 计算平均离差平方时却要除以 $n-1$, 这个调整称为**贝塞尔校正**.

我们计算样本方差 s^2, 是为了**估计**未知的**总体方差** σ^2. 总体方差由数据围绕**总体真实均值** μ 的离散程度决定, 但 μ 往往未知, 只能用**样本均值** $\bar{x}$ 替代, 这就会引入偏差.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

样本均值 $\bar{x}$ 是由这 n 个数据点决定的中心. 算术平均值有一个最优性质, 它使该组数据的离差平方和 $\sum(x_i-c)^2$ 达到**最小值**. 因此用 $\bar{x}$ 计算的离差平方和 $\sum(x_i-\bar{x})^2$, 比用任何其他中心点(包括未知的 μ)计算的都小.

知乎@余命数

$$\sum_{i=1}^{n}(x_i-\bar{x})^2 \le \sum_{i=1}^{n}(x_i-\mu)^2$$

于是, 若用 $\bar{x}$ 并除以 n 计算方差, 所得值会系统性地**偏小**. 在统计学中, 这种系统性估计偏差称为**偏误**.

为了修正这一偏差, 我们需把偏小的结果适当放大. 数学推导表明, 分母取 $n-1$ 恰可校正偏差, 使得 s^2 在平均意义上等于所估的 σ^2. 这样的估计量称为**无偏估计量**.

知乎@余命数

$n-1$ 在统计学中又称**自由度**.n 个独立样本点算出样本均值 $\bar{x}$ 后, 满足约束 $\sum(x_i-\bar{x})=0$. 在这一约束下, 知道其中 $n-1$ 个离差, 第 n 个就被完全确定. 因此估计总体离散程度时, 实际可独立变化的信息只有 $n-1$ 个.

知乎@余命数

标准差 s 的单位与原始数据相同, 它度量数据点偏离其均值的典型距离.s 越大, 数据越分散; s 越小, 数据越集中.

知乎@余命数

8.2 抽样方法

8.2.1 简单随机抽样

我们先讨论简单随机抽样. 它是所有抽样方法中最基础的一种, 其目的在于消除抽样过程中的主观偏见与系统性误差, 并确保样本具有**代表性**. 要理解它的定义, 我们回到一个根本问题, 即如何保证公平性.

知乎@余命数

假设我们调查全校学生的平均身高, 但只测量了校篮球队的成员. 这样得到的样本均值会系统性地高于真实值, 样本不具备代表性. 这种因抽样方法缺陷导致的系统性偏差称为**抽样偏误**. 为了克服偏误, 我们需要一种机制, 使总体中每个个体都有平等的机会被选中.

这一思想源自抽签. 公平的抽签要求每一支签被抽中的可能性完全相同. 我们把抽签的思想数学化, 便得到简单随机抽样的定义.

定义 8.6 (简单随机抽样) 设一个总体含有 N 个个体, 从中逐个不放回地抽取 n 个个体作为样本 ($n \le N$). 如果在整个抽样过程中, **每一次抽取时, 总体中余下的每一个个体被抽到的机会都相等**, 则称这种抽样为**简单随机抽样**.

这个定义刻画了一个动态的、程序化的**过程**, 并为公平性提供可操作的保证.

- **逐个不放回**, 明确抽样的基本机制. 一个个体一旦被抽中, 便移出总体, 不再参与后续抽取. 这是社会调查与质量检验的标准做法.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **每一次抽取时机会都相等**, 这一定义的关键, 也是公平性的数学形式. 它要求抽第一步时每个个体被抽中的概率为 $1/N$; 抽出一个后, 剩下的 $N-1$ 个个体每一个在第二步被抽中的概率为 $1/(N-1)$, 以此类推. 该条件确保任何阶段抽样过程对剩余个体均无偏好.

这种程序公平性的来源是**随机化**, 它也是现代统计学的基础. 引入纯粹随机机制, 可打乱总体中可能存在的内在结构, 例如学生按身高、成绩或班级排序. 随机化下某一次抽得的样本未必恰好代表总体, 但**抽样方法**本身无偏, 长期来看由此方法产生的样本能公正反映总体.

上述过程化的定义可导出一个在组合数学上极为重要的等价性质, 即在简单随机抽样下, 总体中每一个容量为 n 的可能子集被抽为最终样本的概率完全相等.

从 N 个个体中抽取 n 个, 总共存在 $\binom{N}{n}$ 个不同的样本组合. 考虑其中任意一个特定的样本组合 $\{a_1, a_2, \dots, a_n\}$, 按照定义的抽样过程, 以特定顺序, 例如 $a_1, a_2, \dots, a_n$ 被抽出的概率为

$$\frac{1}{N} \times \frac{1}{N-1} \times \frac{1}{N-2} \times \cdots \times \frac{1}{N-n+1} = \frac{1}{P_N^n}.$$

该样本中的 n 个个体可按 $n!$ 种不同顺序被抽出, 且每种顺序的发生概率相同, 故此特定样本组合被抽中的总概率为

$$\begin{aligned} n! \times \frac{1}{P_N^n} &= n! \times \frac{(N-n)!}{N!} && (\text{由 } P_N^n = N!/(N-n)!) \\ &= \frac{1}{N!/(n!(N-n)!)} = \frac{1}{\binom{N}{n}}. \end{aligned}$$

因此, 任何一个容量为 n 的样本组合, 其被选中的概率都是 $1/\binom{N}{n}$. 这一性质是所有基于古典概型的统计推断计算的理论基础, 它把一个动态的抽样过程与一个静态的、所有可能结果等概率的样本空间联系起来.

例题 8.2 设总体容量为 N, 从中抽取一个容量为 n 的简单随机样本. 证明总体中任意一个个体 a 被抽入样本的概率为 $\frac{n}{N}$.

证明 我们用等价性质来证明. 由简单随机抽样的等价性质, 任何容量为 n 的子集被抽为样本的概率均等, 故所有可能样本组合总数为 $\binom{N}{n}$. 若样本须包含特定个体 a, 只需从其余 $N-1$ 个个体中再抽取 $n-1$ 个与之组合, 故包含 a 的样本数为 $\binom{N-1}{n-1}$. 于是 a 被抽入样本的概率为

$$P(a \text{ 于样本里}) = \frac{\text{包含 } a \text{ 的样本数}}{\text{总样本数}} = \frac{\binom{N-1}{n-1}}{\binom{N}{n}}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将组合数按定义展开并约分, 得

$$\begin{aligned}\frac{\binom{N-1}{n-1}}{\binom{N}{n}} &= \frac{(N-1)!}{(n-1)!(N-n)!}\cdot\frac{n!(N-n)!}{N!} && \text{(展开组合数)}\\ &= \frac{(N-1)!}{(n-1)!}\cdot\frac{n(n-1)!}{N(N-1)!} && \text{(消去 } (N-n)!\text{)}\\ &= \frac{n}{N}.\end{aligned}$$

同一结论亦可由顺序视角导出. 将事件“个体 a 被抽入样本”分解为 n 个互斥事件 E_k(“a 在第 k 次抽取时被抽中”, $k=1,2,\dots,n$)的并集, 则 $P(a\text{ 于样本里})=\sum_{k=1}^{n}P(E_k)$. 显然 $P(E_1)=P(a\text{ 在第 1 次被抽中})=\frac{1}{N}$. 对于 E_2, 须第 1 次未抽中 a 且第 2 次抽中, 由乘法公式

$$P(E_2)=\frac{N-1}{N}\times\frac{1}{N-1}=\frac{1}{N}.$$

一般地, E_k 要求前 $k-1$ 次均未抽中 a 且第 k 次抽中, 故

$$\begin{aligned}P(E_k) &= \frac{N-1}{N}\times\frac{N-2}{N-1}\times\cdots\times\frac{N-k+1}{N-k+2}\times\frac{1}{N-k+1}\\ &= \frac{1}{N} && \text{(中间项依次约分)}.\end{aligned}$$

即在简单随机抽样过程中, 任一个体在任一次被抽中的概率都相同. 因此 a 被抽入容量为 n 的样本的总概率为

$$P(a\text{ 于样本里})=\sum_{k=1}^{n}P(E_k)=\sum_{k=1}^{n}\frac{1}{N}=\frac{n}{N}.$$

□

这个结论是简单随机抽样公平性的另一种体现. 它与每个样本组合被抽中的概率相等这一性质等价, 在理论上极为重要. 它保证了使用样本均值 $\bar{x}$ 估计总体均值 μ 时, 这种估计在方法论上无偏, 因为总体中每一个成员都有完全相同的机会对这个估计值做出贡献.

例题 8.3 在第二次世界大战期间, 盟军分析缴获的德军坦克的序列号, 以估计其总产量. 设坦克的序列号从 1 到 N 连续编号, 其中 N 是未知的总产量. 盟军缴获了 n 辆坦克, 其序列号构成从总体 $\{1,2,\dots,N\}$ 中抽取的容量为 n 的简单随机样本. 若观测到的最大序列号为 m, 试给出 N 的一个合理估计.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 令观测到的 n 个序列号为 $x_1, x_2, \dots, x_n$, 样本最大值 $m = \max\{x_1, \dots, x_n\}$. 由于 $N \geq m$, 而 m 几乎必然是低估(恰好抽到最大序列号坦克的概率极小), 需在 m 之上追加一个修正项.

这 n 个被抽中的序列号将从 1 到 N 的整个序列分割开来. 在真正的随机样本中, 成员在总体中应大致均匀散布, 故可用样本内部的平均间距估计总体中未观测到的末尾间距.n 个样本点大致将区间 $[1, m]$ 分割为 n 个小段, 平均每段长度约为 m/n, 以此估计从 m 到 N 这段末尾间距的长度, 便得到初步模型 $\hat{N} \approx m + \frac{m}{n}$.

此估计尚有改进空间. 注意 n 个样本点将总体 N 个数字分成 $n+1$ 个间隔(从 0 到最小样本、样本间的间隔、以及从最大样本到 $N+1$), 在随机抽样下这些间隔的期望长度应相等, 由此消除系统性偏差, 得到更精确的估计量

$$\hat{N} = m + \frac{m}{n} - 1.$$

例如, 缴获 $n = 5$ 辆坦克且最大序列号 $m = 120$ 时, 总产量的估计为

$$\hat{N} = 120 + \frac{120}{5} - 1 = 120 + 24 - 1 = 143.$$

知乎@余命数

德国坦克问题是统计推断中参数估计的一个经典范例. 它展示了统计学如何基于一个不完整的随机样本信息, 对一个未知的总体参数(这里是总体最大值 N)做出推断. 这里给出的 $\hat{N} = \frac{n+1}{n}m - 1$ 是基于样本最大值的常用无偏估计量.

8.2.2 分层抽样

定义 8.7 (分层抽样) 当总体由若干个有显著差异、互不重叠的子群体, 称为**层**构成时, 我们将总体划分为若干层, 然后**在每一层内部分别进行简单随机抽样**, 最后把各层抽取的样本合并起来. 这种方法称为**分层抽样**.

辛普森悖论的例子告诉我们, 当总体存在显著的内部结构时, 对这种结构视而不见可能会导致错误的结论. 简单随机抽样凭纯粹的随机化可保证长期公平性, 但无法保证**每一次**抽样都能准确反映总体的内部结构. 例如, 在一个男女比例为 $1:4$ 的总体中进行容量为 10 的简单随机抽样, 可能因随机性的波动抽到一个男女比例为 $1:1$ 的样本, 从而对依赖于性别的某些特征做出有偏的估计.

知乎@余命数

分层抽样是对纯粹随机性的一种约束和优化. 它承认并利用了总体的内在异质性, 做法是先分层后抽样. 这样, 一个异质性高的复杂总体被分解为若干个同质性高的子总体(层), 我们在每个子总体内部实施随机化.

这种方法有几条优势. 一条是**提升估计精度**. 确保所有子群体都按其在总体中的重要性(通常是规模)被恰当地代表, 分层抽样可降低抽样误差, 得到比同等规模简单随机抽样更精确的总体参数估计值, 即方差更小的估计量. 另一条是**保证子群体代表性**. 它确保样本中包含来自所有层的足够数量的个体, 我们既能对

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

总体进行推断, 又能对各子群体自身的特征进行分析和比较. 第三条是实施方便. 在组织层面, 把大规模的抽样任务分解到各个层级, 往往更易于管理.

在实际操作中, 最常见的分层抽样方法是**按比例分配**, 即每个层所抽取的样本容量占总样本容量的比例, 等于该层所含个体数占总体个体数的比例. 设总体容量为 N, 被划分为 L 个层, 每层的容量分别为 $N_1, N_2, \dots, N_L$. 若要抽取容量为 n 的样本, 则在第 h 层 ($h = 1, \dots, L$) 中抽取的样本容量 n_h 为

$$n_h = n \cdot \frac{N_h}{N}$$

例题 8.4 某大学共有学生 $N = 5000$ 人, 其中低年级学生 $N_1 = 3000$ 人, 高年级学生 $N_2 = 2000$ 人. 学校希望估计全体学生平均每周的自习时间. 根据以往经验, 低年级学生自习时间的波动性较大, 而高年级学生则相对稳定. 现采用分层抽样方法, 抽取一个容量为 $n = 100$ 的样本.

1. 采用按比例分配的原则, 应在两个年级各抽取多少名学生?
2. 假设抽样结果为低年级样本的平均自习时间 $\bar{x}_1 = 15$ 小时, 高年级样本的平均自习时间 $\bar{x}_2 = 25$ 小时. 试估计全校学生平均每周的自习时间.
3. 试讨论为何在这种情况下, 分层抽样的估计结果比简单随机抽样可能更为可靠.

解 总体根据年级这一关键变量自然地分为两层, 这是分层抽样的典型应用场景.

(1) 按比例分配, 第 h 层抽取样本量 $n_h = n \cdot \frac{N_h}{N}$. 代入数据, 低年级层应抽取 $n_1 = 100 \cdot \frac{3000}{5000} = 60$ 人, 高年级层应抽取 $n_2 = 100 \cdot \frac{2000}{5000} = 40$ 人.

(2) 由于样本结构与总体结构成比例, 以各层在总体中的比例 $W_h = N_h/N$ 为权重对各层样本均值加权平均, 即得总体均值 μ 的估计

$$\begin{aligned}\bar{x}_{st} &= W_1\bar{x}_1 + W_2\bar{x}_2 = \frac{N_1}{N}\bar{x}_1 + \frac{N_2}{N}\bar{x}_2 \\ &= \frac{3000}{5000} \cdot 15 + \frac{2000}{5000} \cdot 25 = 0.6 \cdot 15 + 0.4 \cdot 25 = 9 + 10 = 19 \text{ 小时}.\end{aligned}$$

故全校学生平均每周自习时间估计为 19 小时.

(3) 简单随机抽样依赖纯粹的随机机会获得代表性. 本例中容量为 100 的简单随机样本, 其年级比例可能因随机性偏离真实的 $3:2$, 若偶然抽到 40 名低年级与 60 名高年级学生, 且样本均值仍为 15 与 25, 则整体均值为 $0.4 \cdot 15 + 0.6 \cdot 25 = 21$ 小时, 高估真实总体均值. 分层抽样强制样本结构与总体结构一致, 从根本上消除了样本中各子群体比例不当引入的抽样误差, 将随机性限制在各同质层内部, 使最终的加权估计量更为稳定和精确.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此例导出了分层抽样下总体均值的估计方法, 即 $\bar{x}_{st}=\sum_{h=1}^{L}W_h\bar{x}_h$. 这是一个统计性质优良的无偏估计量.

例题 8.5　某市教育部门计划对一项新的教学改革方案进行民意调查, 以了解全体高中生的支持率. 该市共有高中生 $N=10000$ 人, 其中高一年级 $N_1=4000$ 人, 高二年级 $N_2=3500$ 人, 高三年级 $N_3=2500$ 人. 由于不同年级的学生面临的学业压力和课程内容不同, 他们对改革方案的态度可能存在显著差异. 因此, 决定采用分层抽样的方法, 抽取一个容量为 $n=400$ 的样本.

1. 采用按比例分配的原则, 计算应从每个年级分别抽取多少名学生?
2. 调查结束后, 样本数据显示高一、高二、高三学生的支持率分别为 85%、70%、60%. 请估计该市全体高中生对该方案的支持率.

解　各年级是天然的层, 层内学生同质性较高而层间可能存在显著差异, 这是分层抽样的典型应用场景.

1. 按比例分配要求样本中各层的比例与总体中各层的比例保持一致. 各年级学生占总体的比例分别为 $W_1=\frac{N_1}{N}=\frac{4000}{10000}=0.4$、$W_2=\frac{N_2}{N}=\frac{3500}{10000}=0.35$、$W_3=\frac{N_3}{N}=\frac{2500}{10000}=0.25$. 据此分配样本容量 $n=400$, 高一抽取 $n_1=n\times W_1=400\times0.4=160$ 人, 高二抽取 $n_2=n\times W_2=400\times0.35=140$ 人, 高三抽取 $n_3=n\times W_3=400\times0.25=100$ 人.
2. 已知各层内样本支持率 $\hat{p}_1=0.85,\hat{p}_2=0.70,\hat{p}_3=0.60$. 各层在总体中的权重不同, 不能将三者取算术平均, 须以各层比例 W_h 为权重加权平均, 得

$$\begin{aligned}\hat{p}_{st}&=W_1\hat{p}_1+W_2\hat{p}_2+W_3\hat{p}_3\\&=0.4\times0.85+0.35\times0.70+0.25\times0.60\\&=0.34+0.245+0.15=0.735.\end{aligned}$$

故该市全体高中生对该方案的支持率约为 73.5%.

本例说明, 当层间均值差异显著时, 简单随机抽样的总体方差会同时包含层内波动与层间差异; 按比例分层抽样先在各层内抽样再加权汇总, 可消除由层间结构混杂带来的额外方差. 因此在相同样本量下, 分层抽样通常能给出更稳定的总体均值估计.

例题 8.6　一家连锁企业拥有两种类型的门店, 即旗舰店和社区店. 共有旗舰店 $N_1=50$ 家, 社区店 $N_2=450$ 家. 根据历史数据, 旗舰店的月均销售额为 $\mu_1=200$ 万元, 销售额的标准差为 $\sigma_1=50$ 万元; 社区店的月均销售额为

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

$\mu_2 = 80$ 万元, 标准差为 $\sigma_2 = 15$ 万元. 现计划抽取 $n = 50$ 家门店来估计整个连锁企业的平均月销售额. 试比较简单随机抽样和按比例分层抽样两种方法的效率.

解 抽样方法的效率由其估计量的方差衡量, 方差越小则估计波动性越小、精确度越高, 效率也越高. 以下比较两种抽样方法下样本均值 $\bar{x}$ 的方差.

先计算总体相关参数. 总体规模 $N = N_1 + N_2 = 500$, 各层权重 $W_1 = \frac{N_1}{N} = 0.1$, $W_2 = \frac{N_2}{N} = 0.9$, 总体均值

$$\mu = W_1\mu_1 + W_2\mu_2 = 0.1 \times 200 + 0.9 \times 80 = 20 + 72 = 92 \text{ 万元}.$$

总体方差 σ^2 由层内方差的加权平均与层间方差两部分构成, 代入数据得

$$\begin{aligned}\sigma^2 &= \sum_{h=1}^{2} W_h\sigma_h^2 + \sum_{h=1}^{2} W_h(\mu_h - \mu)^2 \\ &= (0.1 \times 50^2 + 0.9 \times 15^2) + [0.1(200-92)^2 + 0.9(80-92)^2] \\ &= (250 + 202.5) + (1166.4 + 129.6) = 452.5 + 1296 = 1748.5.\end{aligned}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

1. 简单随机抽样. 在此方法下样本均值 $\bar{x}_{srs}$ 的方差为 (忽略有限总体校正系数)

$$\mathrm{Var}(\bar{x}_{srs}) = \frac{\sigma^2}{n} = \frac{1748.5}{50} = 34.97.$$

知乎@余命数

2. 按比例分层抽样按比例分配样本量 $n_1 = nW_1 = 50 \times 0.1 = 5$ 家, $n_2 = nW_2 = 50 \times 0.9 = 45$ 家. 分层抽样下总体均值估计量 $\bar{x}_{st} = W_1\bar{x}_1 + W_2\bar{x}_2$ 的方差为

$$\begin{aligned}\mathrm{Var}(\bar{x}_{st}) &= \sum_{h=1}^{2} W_h^2 \frac{\sigma_h^2}{n_h} = \frac{W_1^2\sigma_1^2}{nW_1} + \frac{W_2^2\sigma_2^2}{nW_2} = \frac{1}{n}\sum_{h=1}^{2} W_h\sigma_h^2 \\ &= \frac{1}{50}(0.1 \times 50^2 + 0.9 \times 15^2) = \frac{1}{50}(250 + 202.5) = \frac{452.5}{50} = 9.05.\end{aligned}$$

知乎@余命数

比较二者, $\mathrm{Var}(\bar{x}_{st}) = 9.05 < \mathrm{Var}(\bar{x}_{srs}) = 34.97$, 分层抽样的估计量方差远小于简单随机抽样. 故在相同样本量下, 分层抽样提供更为精确的估计.

本例用到分层抽样的两步, 即按比例分配样本量和对结果进行加权汇总. 按比例分配确保样本在结构上是总体的缩影, 从而保证了样本的代表性. 加权平均是从样本数据推断总体参数的正确逻辑, 它承认并利用了总体的内在分层结构, 避免了因某一层样本比例过高或过低而导致的估计偏差. 如果研究者错误地对 $85\%, 70\%, 60\%$ 求简单算术平均, 将得到 $(0.85 + 0.7 + 0.6)/3 \approx 71.7\%$, 这会低估总体支持率, 因为占比较大的高一年级的较高支持率没有得到足够的体现.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

8.2.3 系统抽样

定义 8.8 (系统抽样) 将总体的 N 个个体按序编号. 确定抽样间隔 $k = \lfloor N/n \rfloor$. 在第 1 到第 k 个个体中随机抽取一个作为起始个体 i. 之后, 依次抽取编号为 $i+k, i+2k, \ldots$ 的个体. 这种方法称为**系统抽样**.

8.3 从样本到总体

第8.2节 解决了怎么抽的问题. 样本到手之后, 真正的问题才开始, 那就是手里的数字离总体的真值有多远. 答案可以算得很准, 不过得从一段历史讲起.

1936 年美国总统选举,《文学文摘》(Literary Digest) 杂志预测共和党候选人兰登将以压倒性优势击败寻求连任的罗斯福, 依据是约 240 万份问卷回复, 规模空前, 而且这家杂志从 1916 年起届届测准. 开票结果是罗斯福以 62% 对 38% 大胜, 预测误差高达十九个百分点, 杂志不久便倒闭了. 同一年, 成立不久的盖洛普用 3000 人的样本正确预测了罗斯福获胜, 甚至提前一个月预测出了《文学文摘》会给出什么数字, 误差不到一个百分点.240 万人输给 3000 人, 原因不在人数, 在样本之外.

知乎@余命数　知乎@余命数　知乎@余命数

《文学文摘》的名单取自电话簿和俱乐部会员名册. 大萧条尚未过去, 四户人家才有一部电话, 穷人几乎进不了名单, 而穷人恰恰支持罗斯福.1000 万份寄出的问卷只收回 240 万份, 愿意回信的人又与不肯回信的人不同. 两道筛子朝着同一个方向偏, 样本再大, 也只是在更大的规模上重复同一个错误. 这种系统性的偏移叫**抽样偏差**, 它与抽签式的随机波动是两种不同性质的误差. 偏差要靠抽样设计消除, 随机误差则可以靠数学度量, 接着我们就来看看数学怎么把随机误差算清楚.

8.3.1 盒子模型与标准误

把总体想成一只盒子. 要估计总体中具有某种特征的个体占比 p, 就为每个个体放一张票, 具有该特征的记 1, 其余记 0, 于是盒中 1 的比例恰为 p. 从总体中做样本量为 n 的简单随机抽样, 相当于从盒中无放回地抽 n 张票. 第10.8节 讨论过, 当总体远大于样本时, 无放回与有放回几乎无异, 样本中 1 的张数 X 近似服从二项分布 $B(n,p)$, 期望为 np, 方差为 $np(1-p)$.

实际报告里用的是比例 $\hat{p} = X/n$, 即样本中 1 的占比. 由期望的线性性与方差的性质 (第10.7节),

$$E(\hat{p}) = p, \qquad \sqrt{D(\hat{p})} = \sqrt{\frac{p(1-p)}{n}}. \tag{8-1}$$

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

期望恰好等于真值, 说明简单随机抽样没有系统偏差, $\hat{p}$ 的波动大小由方差的平方根刻画. 这个平方根称为**标准误**, 它度量机会误差的典型大小, 样本比例偏离真值的幅度通常在一两倍标准误之内.

例题 8.7 (一千人够不够) 某健康调查的 6672 名受访者中男性占 46%. 一位研究者从中简单随机抽取 100 人复核, 样本中男性占比的波动有多大? 若改抽 400 人呢?

解 把男性记 1、女性记 0, 则 $p = 0.46$, 样本占比 $\hat{p}$ 的标准误为

$$\sqrt{\frac{0.46 \times 0.54}{100}} = \sqrt{0.002484} \approx 0.050,$$

约五个百分点. 换成 400 人, 标准误为

$$\sqrt{\frac{0.46 \times 0.54}{400}} \approx 0.025,$$

恰好减半. 样本量翻两番, 精度只翻一番, 这就是标准误的**平方根律**, 它由 式 (8-1) 中分母的 n 直接决定. 精度提高一倍要付出四倍的调查成本, 统计的代价在这里明码标价.

100 人样本的标准误约 5%, 400 人约 2.5%, 精度随 $\sqrt{n}$ 增长.

式 (8-1) 里没有总体规模的位置. 只要总体远大于样本, 从千万人口的国家和从十万人口的城市各抽一千人, 对占比估计的精度几乎相同. 这与直觉相悖, 却正是抽样的经济性所在. 新闻里民调常标注“误差范围 ± 3 个百分点”, 那正是约一千人的样本、按最保守的 $p = 0.5$ 算出的两倍标准误

$$2\sqrt{\frac{0.5 \times 0.5}{1000}} \approx 3.2\%.$$

8.3.2 正态近似

标准误给出波动的尺度, 要算出概率还需要分布形状. 第10.8节 介绍的正态分布在此接手. 当样本量不太小、p 不太靠近 0 或 1 时, $\hat{p}$ 的分布近似正态, 中心在 p, 标准差就是标准误. 样本比例落在真值左右一个标准误内的概率约为 68%, 落在两个标准误内约为 95%.

例题 8.8 (电话公司的市场调查) 某电话公司有 10 万用户, 其中 20% 年收入超过五万元. 公司简单随机抽取 400 人做市场研究, 抽到的用户中高收入者占比介于 18% 与 22% 之间的概率约是多少?

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 高收入人数 $X \sim B(400,\ 0.2)$ 近似，样本占比 $\hat{p}$ 的标准误为

$$\sqrt{\frac{0.2 \times 0.8}{400}} = \sqrt{0.0004} = 0.02.$$

以 20% 为中心、2% 为标准差，区间 $[18\%, 22\%]$ 恰好是两个标准差之内的对称区间，由正态近似，概率约为 95%.

约为 95%.

8.3.3 置信区间

实际调查的方向与上面的例子相反. 真值 p 未知，手里只有一个样本占比 $\hat{p}$，要反过来给 p 圈定范围. 思路是把 $\hat{p}$ 的波动折算回去. 标准误 式 (8-1) 里的 p 用观测值 $\hat{p}$ 代替，得到标准误的估计

$$\widehat{\mathrm{SE}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}},$$

然后以 $\hat{p}$ 为中心向两边各取两倍标准误，得到区间

$$\hat{p} \pm 2\,\widehat{\mathrm{SE}} = \left[\hat{p} - 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]. \tag{8-2}$$

定义 8.9 (置信区间) 按 式 (8-2) 构造的区间称为总体比例 p 的**置信区间**，置信水平约为 95%. 把两端的标准误倍数换成 1 倍或 3 倍，置信水平相应约为 68% 或 99.7%.

例题 8.9 (选前民调) 某镇有 25000 名选民，随机抽取 1600 人，其中 917 人支持某候选人. 给出全体选民支持率的 95% 置信区间.

解 样本支持率

$$\hat{p} = \frac{917}{1600} \approx 0.573,$$

标准误的估计

$$\widehat{\mathrm{SE}} = \sqrt{\frac{0.573 \times 0.427}{1600}} \approx \sqrt{0.000153} \approx 0.012,$$

置信区间为 $57.3\% \pm 2.4\%$，即 $[54.9\%, 59.8\%]$. 以 95% 的把握，该候选人在全体选民中的支持率落在这个区间内，过半在望，但优势大小仍有不确定的成分.

95% 置信区间约为 $[54.9\%, 59.8\%]$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

“95% 的把握”的确切含义值得多说一句. 总体支持率是一个固定的数, 不随机, 随机的是区间, 每抽一次样就得到一个新区间.95% 说的是方法的长期能力, 反复抽样反复构造区间, 约有 95% 的区间套住真值, 5% 的错过. 单次调查得到的区间要么套住了, 要么没套住, 无法自知, 但采用这个方法的赌注是划算的.

置信区间也有边界.p 极端靠近 0 或 1 而样本又不大时, 二项分布明显偏斜, 正态近似失真, 式 (8-2) 便不再可靠. 判断的办法是想象总体与样本同构成, 此时预期 1 的张数 $n\hat{p}$ 与 0 的张数 $n(1-\hat{p})$ 都该不少于五左右, 正态近似才稳妥.

现在可以完整地复盘 1936 年了.《文学文摘》的名单偏向富人, 回信又自我筛选, 抽样偏差把 19 个百分点系统地搬走了, 任何标准误都救不回来. 盖洛普从同样的名单里随机抽 3000 人, 偏差消失, 标准误只有约一个百分点, 所以他连别人的错误都能提前测准. 好的推断是两件事的合成, 无偏的设计挡住系统误差, 标准误诚实地报告剩下的随机误差, 缺一不可.

练习 8.1 某市为估计成年居民中经常锻炼者的比例, 简单随机调查了 900 人, 其中 432 人经常锻炼. 求样本比例、标准误的估计, 并给出全市经常锻炼者比例的 95% 置信区间.

练习 8.2 若要把上题的标准误减半, 样本量需增大到多少? 由此说明误差范围 ± 1.5 个百分点大致需要多大的样本.

练习 8.3 某网络平台发起在线问卷, 收到 50 万份回复, 其中 83% 反对某项收费调整, 平台据此宣布“绝大多数用户反对”. 用抽样偏差与标准误的语言评论这一结论.

8.4 相关与回归

前两节的统计对象一次只有一个变量. 更多的问题里, 每个个体带着两个数字出场, 父亲的身高与儿子的身高, 学习年限与收入, 温度与饮料销量. 单看两组平均数会丢掉最重要的信息, 两个变量如何相互联动, 这要从画图开始.

8.4.1 散点图与相关系数

把每一对数据 (x_i, y_i) 画成一个点, 得到的图叫**散点图**. 下图是计算机模拟的两组父子身高数据, 左图散点紧贴一条上升趋势的直线, 右图散点松散, 直线趋势

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

微弱. 两个图的平均数与标准差几乎相同, 区别全在点围绕直线的紧密程度, 需要一个数来度量它.

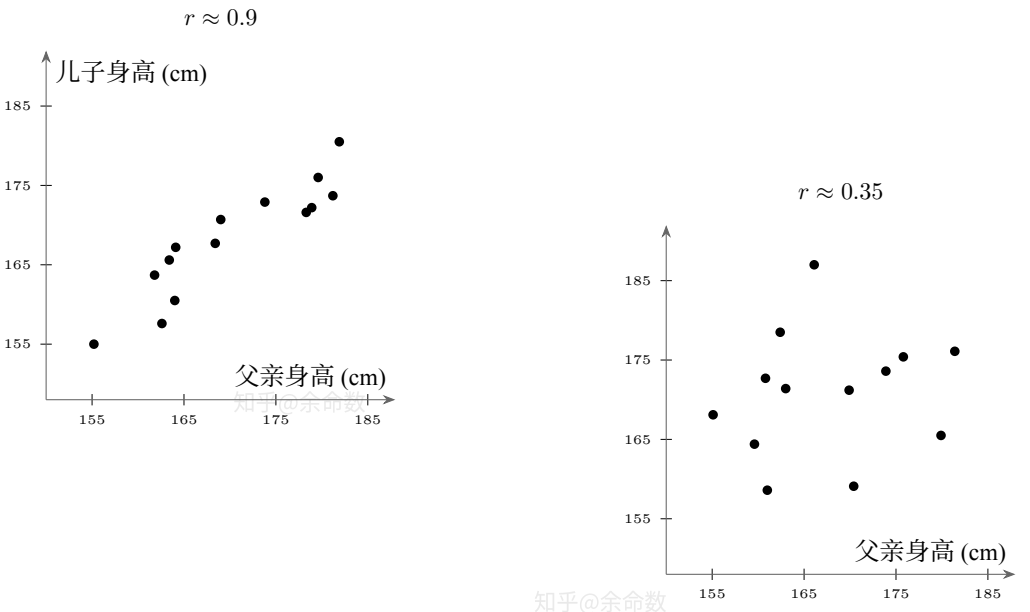


图 8.4　模拟的两组父子身高数据, 平均数与标准差几乎相同, 相关强弱迥异.

度量的想法是把两个变量都标准化, 再看它们同起同落还是各行其是. 记 $\bar{x}, \bar{y}$ 为两组平均数, s_x, s_y 为两组标准差 (第8.1节), 偏差 $d_{x,i} = x_i - \bar{x}$, $d_{y,i} = y_i - \bar{y}$. 衡量一对偏差是否同步, 最自然的乘积 $d_{x,i}\, d_{y,i}$ 同号时增大、异号时缩小, 于是把全部乘积加起来除以恰当的尺度, 得到**相关系数**

$$r = \frac{\sum_i d_{x,i}\, d_{y,i}}{\sqrt{\left(\sum_i d_{x,i}^2\right)\left(\sum_i d_{y,i}^2\right)}}. \tag{8-3}$$

写成标准化的样子即 $r = \dfrac{1}{n}\sum_i \dfrac{d_{x,i}}{s_x} \cdot \dfrac{d_{y,i}}{s_y}$, 分子是标准化变量的平均乘积, 分母的折算把 r 的取值压进 $[-1, 1]$.

> **注 8.1 ($|r| \le 1$ 的两行证明)**
> 对任意实数 t, $\sum_i (d_{x,i} - t\, d_{y,i})^2 \ge 0$. 展开得 $t^2 \sum d_y^2 - 2t \sum d_x d_y + \sum d_x^2 \ge 0$ 对一切 t 成立, 故关于 t 的二次函数判别式非正, $\left(\sum d_x d_y\right)^2 \le \sum d_x^2 \sum d_y^2$, 代入 式 (8-3) 即 $|r| \le 1$. 等号成立当且仅当存在 t 使所有 $d_{x,i} = t\, d_{y,i}$, 即全部点恰好落在一条直线上, 此时 $r = \pm 1$, 称完全相关.

r 的符号读作趋势方向.$r > 0$ 表示一个变量增大时另一个倾向于增大, 散点整体向上爬; $r < 0$ 表示一增一减, 散点下坡. 绝对值大小读作围绕直线的紧密度, $|r|$ 越接近 1 点带越窄, 越接近 0 云团越散. 有两点常被误读.$r = 0.8$ 并非表示 80% 的点在直线上, r 也不衡量两倍于 $r = 0.4$ 的线性强度, 它只是单调的紧密度指标. 另

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

外 r 只度量**线性**关联, 点落在一条弯曲的光滑曲线上时 r 可以等于 0, 后文的习题会给出这样的例子.

例题 8.10 (手算相关系数) 五名学生的数学成绩 x 与物理成绩 y 为 $(1,3),(2,1),(3,4),(4,2),(5,5)$. 计算 r.

解 $\bar{x}=3,\bar{y}=3$, 偏差

$$d_x=(-2,-1,0,1,2),\qquad d_y=(0,-2,1,-1,2).$$

逐对相乘求和

$$\sum d_x d_y=0+2+0-1+4=5,\qquad \sum d_x^2=\sum d_y^2=10.$$

代入 式 (8-3),

$$r=\frac{5}{\sqrt{10\times 10}}=\frac{1}{2}.$$

$r=0.5$, 中等程度的正相关.

8.4.2 回归直线与最小二乘法

散点图上点带既已成形, 自然想画一条直线概括它, 并用它由一个变量预测另一个. 设直线为 $\hat{y}=a+bx$, 每个点离直线的竖直距离 y_i-a-bx_i 称为**残差**. 让直线贴住全部点, 标准做法是让残差平方和

$$Q(a,b)=\sum_i (y_i-a-bx_i)^2$$

尽可能小, 这就是**最小二乘法**, 相应的直线称为**回归直线**. 最小值可以用配方法干净地求出, 不需要任何新工具.

定理 8.1 (回归系数公式) 使 $Q(a,b)$ 最小的直线过平均点 $(\bar{x},\bar{y})$, 斜率与截距为

$$b=\frac{\sum_i d_{x,i}\,d_{y,i}}{\sum_i d_{x,i}^2}=r\,\frac{s_y}{s_x},\qquad a=\bar{y}-b\,\bar{x}.$$

证明 改写残差. 记 $c=a-\bar{y}+b\bar{x}$, 则 $y_i-a-bx_i=d_{y,i}-b\,d_{x,i}-c$, 于是

$$Q=\sum_i \big(d_{y,i}-b\,d_{x,i}-c\big)^2=\sum_i \big(d_{y,i}-b\,d_{x,i}\big)^2-2c\sum_i\big(d_{y,i}-b\,d_{x,i}\big)+n\,c^2.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由偏差和为零, $\sum d_y = \sum d_x = 0$, 中项消失, 剩下

$$Q = \sum_i \left(d_{y,i} - b\, d_{x,i}\right)^2 + n\, c^2 \geq \sum_i \left(d_{y,i} - b\, d_{x,i}\right)^2,$$

等号当且仅当 $c = 0$, 即 $a = \bar{y} - b\bar{x}$, 直线必过平均点. 对剩下的二次式再展开,

$$\sum d_y^2 - 2b \sum d_x d_y + b^2 \sum d_x^2,$$

这是关于 b 的开口向上的二次函数, 顶点在

$$b = \frac{\sum_i d_{x,i}\, d_{y,i}}{\sum_i d_{x,i}^2}.$$

最后, 由 $s_x^2 = \frac{1}{n}\sum d_x^2$ 与 $s_y^2 = \frac{1}{n}\sum d_y^2$,

$$r\,\frac{s_y}{s_x} = \frac{\sum d_x d_y}{n\, s_x s_y} \cdot \frac{s_y}{s_x} = \frac{\sum d_x d_y}{\sum d_x^2} = b.$$

□

两处细节值得停一停. 回归线必过平均点, 且最优时残差和 $\sum_i (y_i - \hat{y}_i) = \sum (d_{y,i} - b d_{x,i}) = 0$, 高估与低偿恰好抵消. 斜率 $b = r\, s_y / s_x$ 把两种信息合在一起, r 管方向与紧密程度, 标准差之比管量纲换算, 父亲身高每增加一个标准差, 儿子身高平均增加 r 个标准差.

例题 8.11 (手算回归线) 求上例五名学生成绩数据的回归直线, 并据此预测 $x = 6$ 时的 y.

解 斜率

$$b = \frac{\sum d_x d_y}{\sum d_x^2} = \frac{5}{10} = \frac{1}{2}, \qquad a = \bar{y} - b\,\bar{x} = 3 - \frac{1}{2} \times 3 = \frac{3}{2},$$

回归直线 $\hat{y} = \dfrac{1}{2}x + \dfrac{3}{2}$. 预测 $\hat{y}(6) = \dfrac{15}{4}$. 五个点的残差为 $(1, -\frac{3}{2}, 1, -\frac{3}{2}, 1)$, 残差和恰为零, 残差平方和为 7.5, 已是所有直线能达到的最小值.

$\hat{y} = \frac{1}{2}x + \frac{3}{2}$, 预测值 $\frac{15}{4}$.

例题 8.12 (教育与收入) 美国 2005 年一项抽样调查记录了 562 名 25～29 岁男性的受教育年限与年收入, 摘要统计为平均受教育 12.5 年、标准差 3 年, 平均收入约 30000 美元、标准差约 24000 美元, 相关系数约 0.25. 求由受教育年限预测收入的回归直线, 并解读斜率.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 斜率

$$b = r\frac{s_y}{s_x} \approx 0.25 \times \frac{24000}{3} = 2000, \qquad a \approx 30000 - 2000 \times 12.5 = 5000,$$

回归直线为 $\hat{y} \approx 2000x + 5000$(美元),其中 x 为教育年限,$\hat{y}$ 为预测年收入. 斜率读作在该调查人群中,教育年限每多一年,平均收入约高 2000 美元. 截距 5000 美元对应零年教育的预测,已经处在数据边缘,外推到远离数据的地方并不可靠,同一调查中女性样本的回归截距甚至是负数,这并非说不上学反而倒欠钱,只是直线在数据范围外的延伸没有意义.

斜率约 2000 美元/年,教育每多一年平均收入约高 2000 美元.

8.4.3 均值回归

把斜率换算成标准差语言,会看到一件有趣的事. 预测值的偏差为 $b\,d_x$,除以 s_y 后等于 r 乘以 d_x/s_x,用回归线预测时,一个高出平均 2 个标准差的 x,对应的 y 平均只高出 $2r$ 个标准差. 只要 $|r| < 1$,预测就比观测值更靠近平均.

高尔顿在十九世纪末研究父子身高时正遇到了这个现象. 极高的父亲,儿子平均略矮于父亲;极矮的父亲,儿子平均略高于父亲,身高像被一股力量往平均处拉. 他把这条直线称为“回归线”,“回归”一词从此进入统计学,本意正是**向平均回归**. 数值上,设父亲身高平均 170 cm、标准差 6 cm,儿子身高平均 172 cm、标准差 6 cm,$r = 0.5$,回归线为 $\hat{y} = 0.5x + 87$. 父亲 182 cm(高出平均两个标准差),儿子预测 178 cm,只高出平均一个标准差,高个家庭的后代平均上向中心退了半步,但仍高于人群平均.

8.4.4 相关不等于因果

回归线的斜率常被读成干预效果,教育每多一年收入高 2000 美元,让所有人都多读一年书,收入并不会因此普遍上涨. 上述数据来自观察性调查,研究者只是记录了自然形成的差异,受教育多的人与受教育少的人在能力、家境、机遇上都不同,这些因素与教育纠缠在一起,统计学称为**混杂**. 斜率诚实地度量了人群中的关联,却分不清关联里有多少来自教育本身. 要回答因果问题,靠的是第8.2节 里随机化分组的对照实验,把其他因素的分布搅平,斜率才能读成干预效果.

这正是章首辛普森悖论的回声. 潜在变量既能让汇总数据颠倒分组结论,也能让两个本无因果的变量呈现高度相关. 夏季气温升高同时推高冰淇淋销量与溺水人数,两者的相关系数不低,冰淇淋却救不了溺水者. 遇到相关先问三个问题,数据如何产生,有无潜在的第三变量,换成对照实验结论还成立吗. 相关系数与回归线是描述数据的利器,因果的门槛它们自己跨不过去.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

练习 8.4　六次观测的数据为 $(1,2),(2,1),(3,4),(4,3),(5,5),(6,4)$. 计算相关系数, 求回归直线, 并预测 $x=8$ 时的 y.

练习 8.5　五个点 $(1,8),(2,3),(3,0),(4,3),(5,8)$ 全部落在抛物线 $y=(x-3)^2$ 上. 验证 $r=0$, 并说明 y 由 x 完全决定.

练习 8.6　某地统计发现家庭藏书量与孩子考试成绩的相关系数约为 0.5. 有人提议“多买书就能提高成绩”. 给出至少两条理由说明这个推断不成立.

8.5 章末习题

习题

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

8.1. (线性变换)一组数据 $x_1,\dots,x_n$ 的平均数为 $\bar{x}$, 标准差为 s. 数据 $y_i=3x_i+2$ 的平均数与标准差各是多少? 提示: 平均数线性, $\bar{y}=3\bar{x}+2$; 方差按 $D(aX+b)=a^2D(X)$ 缩放, 标准差乘 $|3|$.

8.2. (样本均值的标准误) 某地区成年男性身高的标准差约为 7 cm. 从该总体简单随机抽取 49 人, 样本平均身高偏离总体平均的典型幅度是多少? 提示: 样本均值的方差为 $\dfrac{\sigma^2}{n}$, 标准误为 $\dfrac{7}{\sqrt{49}}=1$ cm.

8.3. (过半的概率) 某修正案在选民中的真实支持率为 52%. 随机访问 2500 人, 样本支持率超过一半的概率约是多少? 提示: 标准误 $\sqrt{0.52\times0.48/2500}\approx1\%$, 50% 在真值下方两个标准误处, 用正态近似.

8.4. (样本量设计) 要使支持率的 95% 置信区间半宽不超过 2 个百分点, 按最保守的 $p=0.5$ 估计, 至少需要多大样本? 提示: 解 $2\sqrt{0.25/n}\le0.02$.

8.5. (置信区间的读法) 某次调查得到支持率的 95% 置信区间为 $[51\%,55\%]$, 于是有人断言“真值有 95% 的概率落在这个区间里”. 这句话错在哪里? 提示: 真值是固定常数, 不随机; 随机的是区间, 95% 说的是反复抽样时约 95% 的区间套住真值.

8.6. (父子身高回归) 设父亲身高平均 172 cm、标准差 7 cm, 儿子身高平均 174 cm、标准差 7 cm, 相关系数 0.6. 求由父亲身高预测儿子身高的回归直线, 并预测身高 190 cm 的父亲之子的身高. 提示: $b=r\,s_y/s_x=0.6$, $a=174-0.6\times172$, 代入 $x=190$.

8.7. (向平均回归) 某体能训练营按首日测试把学员分组, 教练发现首日成绩最差的一组次日平均进步明显, 首日成绩最好的一组次日平均反而退步, 于是认为“分组训练法

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

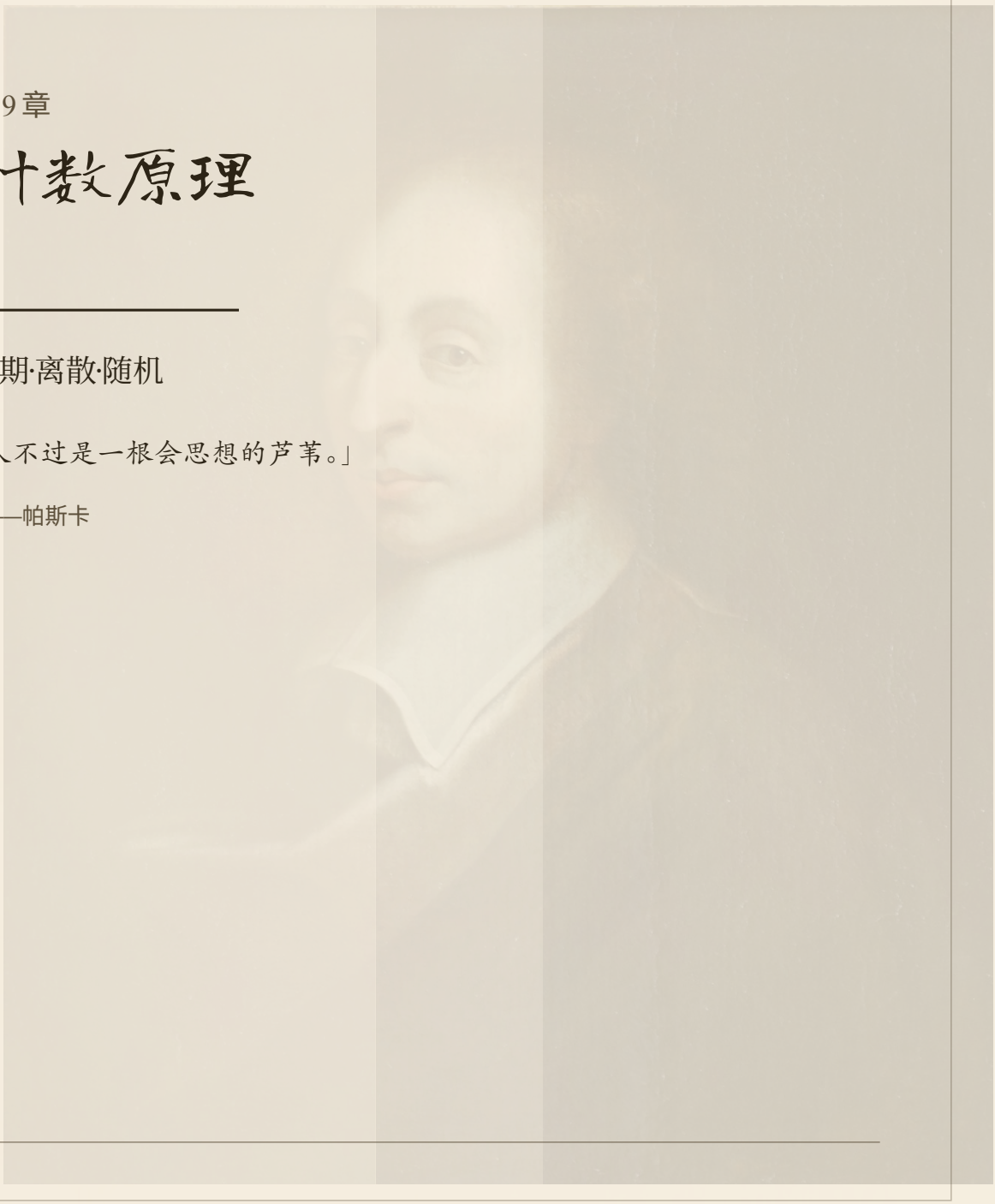
见效”. 给出另一种解释. 提示: 首日极端成绩含有极端的随机波动, 次日波动回归常态, 与训练法无关, 这正是高尔顿的均值回归.

8.8. (相关的符号)判断下列相关系数的符号并简述理由, 二手车的车龄与价格, 身高与体重, 海拔与水的沸点温度. 提示: 依次为负、正、负, 分别来自折旧、体格的同向增长与气压降低沸点下降.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 9 章

计数原理



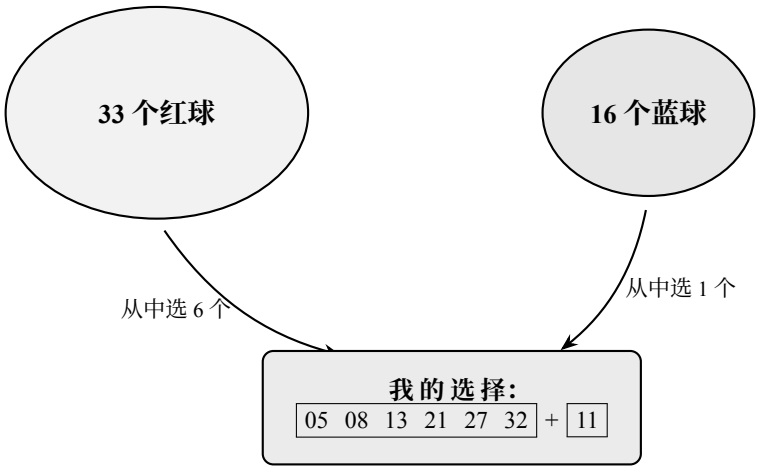
周期·离散·随机

「人不过是一根会思想的芦苇。」

——帕斯卡

第九章 计数原理

知乎@余命数



知乎@余命数

中头奖究竟有多难?

图 9.1 彩票抽奖示意图: 从 33 个红球中选 6 个, 从 16 个蓝球中选 1 个

知乎@余命数 知乎@余命数

中头奖究竟有多难

知乎上有这么一个有趣的问题“给你选的话, 你想选几百万资产, 还是上清华北大?”[1] 我相信每个人的答案不同, 但是许多人也为这百万现金有过追求, 正如每周开奖的彩票吸引着无数人, 每个人都梦想着能选中那组独一无二的幸运号码. 但这种幸运究竟有多么渺茫, 要精确计算中得头奖的概率, 必须先回答一个根本问题, 即从所有可能的号码组合中到底有多少种不同的组合.

[1] 免责声明. 给一些没有幽默感的人说句话, 如果你不知道的话, 这是一个笑话, 可能并不好笑, 如果你不经常上知乎这个软件平台, 自然是不知道的.

知乎@余命数

以一种常见的彩票游戏为例. 需从 33 个红色球中选出 6 个, 再从 16 个蓝色球中选出 1 个. 这项任务可以分为两步. 第一步挑选红球, 第二步挑选蓝球. 挑选蓝球很简单, 有 16 种选择. 真正的挑战是第一步.

从 33 个红球中不重复地选出 6 个, 一共有多少种选法?

不妨试着列举, $(1,2,3,4,5,6),(1,2,3,4,5,7),\ldots$ 很快我们就会发现, 这是一个不可能完成的任务, 其总数将是一个天文数字. 更重要的是, 我们只关心选出了**哪些**号码, 而不在乎它们被摇出的**先后顺序**. 选出 ‘05、08、13、21、27、32‘ 和选出 ‘32、21、13、08、27、05‘ 是完全一样的结果.

这种只选不排的问题, 正是我们要研究的对象. 我们需要一套方法, 用来系统地计算在各种限制下有多少种方法. 这些方法既包括需要考虑顺序的**排列**, 也包括不需要考虑顺序的**组合**. 这就是本章将要探索的计数原理.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

9.1 加法原理与乘法原理

知乎@余命数

我们要处理形形色色的计数问题. 无论问题表面多么不同, 归根到底都用到两条原理, 即**加法原理**与**乘法原理**. 它们本身很朴素, 组合起来却能产生各种计数方法.

9.1.1 加法原理

从北京到上海, 可以乘坐高铁(有 3 个不同的车次), 也可以乘坐飞机(有 5 个不同的航班). 我们来看看一共有多少种选择.

乘坐高铁与乘坐飞机是两类**互斥**的选择. 一旦选定高铁的某个车次, 就不可能再选任何一班飞机. 因此总选择数是两类选择数之和, 这是加法原理.

定理 9.1 (加法原理) 如果完成一件事有 k 类方法, 这些类别两两互斥, 且任何一类方法中的任何一种方法都能独立地完成任务. 在第一类方法中有 m_1 种不同的途径, 在第二类方法中有 m_2 种不同的途径, …, 在第 k 类方法中有 m_k 种不同的途径, 那么完成这件事共有

$$N = m_1 + m_2 + \cdots + m_k$$

种不同的方法.

@余命数

证明 设第 i 类方法构成有限集合 A_i, 则 $|A_i| = m_i$. 类别两两互斥意味着当 $i \neq j$ 时 $A_i \cap A_j = \emptyset$. 完成这件事的全部方法是并集

余命数

$$A_1 \cup A_2 \cup \cdots \cup A_k.$$

知乎@余命数

因为这些集合两两不交, 对有限集合反复使用两个集合的并集计数公式 $|A \cup B| = |A| + |B|$, 可得

$$|A_1 \cup \cdots \cup A_k| = |A_1| + \cdots + |A_k| = m_1 + \cdots + m_k.$$

因此完成任务的方法总数为 $N = m_1 + \cdots + m_k$. □

加法原理的要点在于分类恰当. 当一个问题可以拆成几个互不相干、彼此独立的情形时, 总数就是各情形之和. 在上面的例子中, 乘坐高铁是一类, 乘坐飞机是另一类; 总选择数就是 $3 + 5 = 8$ 种.

从集合论的观点看, 如果 A_i 代表第 i 类方法的集合, 那么这些集合两两不交, 即 $A_i \cap A_j = \emptyset$ for $i \neq j$. 完成任务的总方法数等于这些集合并集的大小.

$$|A_1 \cup A_2 \cup \cdots \cup A_k| = |A_1| + |A_2| + \cdots + |A_k|$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

下面看几个例子.

例题 9.1 某大学的课程目录中, 数学系开设了 5 门不同的选修课, 计算机系开设了 3 门不同的选修课. 一名学生希望从中选择一门课来学习, 他有多少种不同的选择?

解 这是加法原理最直接的应用, 学生的目标是选择一门课. 可选课程分为互斥的两类. 选数学系课程有 5 种, 选计算机系课程有 3 种. 因一名学生只能选其中一门课, 二者互斥, 由加法原理总的选择数为

$$N = 5 + 3 = 8,$$

即该学生共有 8 种不同的选择.

例题 9.2 使用字母 $\{A, B, C\}$ 和数字 $\{1, 2\}$, 可以构成多少个长度为 1 或长度为 2 的密码? 密码中的字符不能重复.

解 构成一个密码这件事, 可按密码长度分为两种互斥情况.

当密码长度为 1 时, 从 5 个可用字符 $\{A, B, C, 1, 2\}$ 中任选一个, 有 5 种方法.

当密码长度为 2 且字符不重复时, 这是一个分步过程. 先为第一位选一个字符, 有 5 种选择; 因不能重复, 第二位只剩 4 种选择. 由乘法原理, 长度为 2 的密码共有 $5 \times 4 = 20$ 个.(分步相乘的依据, 是下一小节将严格表述的**乘法原理**.)

一个密码的长度不可能既是 1 又是 2, 二者互斥, 由加法原理, 总的密码个数为

$$N = 5 + 20 = 25,$$

即总共可构成 25 个这样的密码.

例题 9.3 (斐波那契数列) 要登上一个有 10 级台阶的楼梯, 如果每次只能上 1 级或 2 级, 总共有多少种不同的登法?

解 直接列举所有情况会非常困难. 以 $f(n)$ 记登上 n 级台阶的方法数, 所求是 $f(10)$.

考虑踏上第 n 级台阶的最后一步, 它只有两种互斥可能. 要么从第 $n-1$ 级一步跨上, 要么从第 $n-2$ 级一步跨两阶上. 前者须先到达第 $n-1$ 级, 有 $f(n-1)$ 种方法; 后者须先到达第 $n-2$ 级, 有 $f(n-2)$ 种方法. 由于最后一步的出发点唯一, 二者互斥, 由加法原理得到递推关系

$$f(n) = f(n-1) + f(n-2).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为使用此关系, 需先确定初始条件. 对 1 级台阶只有一种登法, $f(1)=1$; 对 2 级台阶有连续两步 1 级或一步 2 级两种登法, $f(2)=2$.

由此依次推算

$$\begin{aligned} f(3) &= f(2)+f(1)=2+1=3,\\ f(4) &= f(3)+f(2)=3+2=5,\\ f(5) &= 5+3=8,\\ f(6) &= 8+5=13,\\ f(7) &= 13+8=21,\\ f(8) &= 21+13=34,\\ f(9) &= 34+21=55,\\ f(10) &= 55+34=89, \end{aligned}$$

故总共有 89 种不同的登法.

这个例子说明, 加法原理可以用来支撑递归构造. 把一个大问题分解为几个互斥的、结构相同的子问题, 再把它们的解相加, 就得到原问题的解. 序列 $1,2,3,5,8,\ldots$ 是斐波那契数列的一个变体.

例题 9.4 在一个平面直角坐标系中, 一个质点从原点 $(0,0)$ 出发, 每步只能使 y 坐标或 x 坐标加 1, 即向上或向右移动一个单位. 问该质点到达点 $(3,2)$ 有多少条不同的路径?

解 令 $P(m,n)$ 表示从 $(0,0)$ 到达点 (m,n) 的路径数, 所求是 $P(3,2)$. 与登台阶一例同理, 考虑质点是如何到达终点 (m,n) 的. 最后一步必然来自左邻点 $(m-1,n)$ 或下邻点 $(m,n-1)$, 前者对应向右一步到达, 后者对应向上一步到达, 二者互斥. 前者须先到达 $(m-1,n)$, 有 $P(m-1,n)$ 种方法; 后者须先到达 $(m,n-1)$, 有 $P(m,n-1)$ 种方法. 由加法原理得到递推关系

$$P(m,n)=P(m-1,n)+P(m,n-1).$$

边界条件是明显的. 从 $(0,0)$ 到达 x 轴上任意点 $(m,0)$ 只有一条路径(一直向右), 即 $P(m,0)=1$; 同理 $P(0,n)=1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

据此由近及远依次推算

$$
\begin{aligned}
&P(1,0)=1, && P(0,1)=1,\\
&P(1,1)=P(0,1)+P(1,0)=2,\\
&P(2,0)=1, && P(0,2)=1,\\
&P(2,1)=P(1,1)+P(2,0)=3,\\
&P(1,2)=P(0,2)+P(1,1)=3,\\
&P(2,2)=P(1,2)+P(2,1)=6,\\
&P(3,1)=P(2,1)+P(3,0)=4,\\
&P(3,2)=P(2,2)+P(3,1)=10,
\end{aligned}
$$

故共有 10 条不同的路径.

例题 9.5 正整数 $N=180$ 的所有正因数中, 有多少个是 6 的倍数或者是 10 的倍数?

解 对 N 作素数分解, $N=180=2^2\cdot 3^2\cdot 5^1$, 故其任一正因数 d 都可表示为 $d=2^a\cdot 3^b\cdot 5^c$, 其中 $0\le a\le 2$、$0\le b\le 2$、$0\le c\le 1$.

此题看似须用容斥原理, 因为一个数可能既是 6 的倍数又是 10 的倍数; 然而通过更精细的分类, 即可纯粹地套用加法原理. 将目标因数分为两类.

第一类是 6 的倍数. 因数是 $6=2\cdot 3$ 的倍数, 意味着 $a\ge 1$ 且 $b\ge 1$. 此时 a 可取 $\{1,2\}$ 共 2 种, b 可取 $\{1,2\}$ 共 2 种, c 可取 $\{0,1\}$ 共 2 种, 由乘法原理这类因数共有 $2\times 2\times 2=8$ 个.

第二类是 10 的倍数但不是 6 的倍数. 因数是 $10=2\cdot 5$ 的倍数, 意味着 $a\ge 1$ 且 $c\ge 1$; 同时它不是 6 的倍数, 即条件 $a\ge 1$ 且 $b\ge 1$ 不成立. 既然已要求 $a\ge 1$, 欲使其非 6 的倍数, 必须 $b<1$, 即 $b=0$. 于是 a 仍有 2 种选择, b 仅 1 种, c 仅 1 种, 这类因数共有 $2\times 1\times 1=2$ 个.

这两类因数的定义互斥 (一类是 6 的倍数, 另一类不是 6 的倍数), 由加法原理, 满足条件的因数总数为

$$N=8+2=10,$$

即共有 10 个这样的因数.

精确定义分类标准 (A 和 $B\setminus A$), 把一个看似需要容斥原理的问题转化为纯粹的加法原理问题. 这显示了恰当分类在计数中的作用.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

9.1.2 乘法原理

现在考虑一个不同的场景. 假设你正在为一次晚宴搭配服装, 有 3 件不同的衬衫和 4 条不同的裤子. 搭配方式的总数按下面的步骤算出.

完成搭配一套服装这个任务, 须**依次**完成两个步骤. 先选一件衬衫, 再选一条裤子. 第一步的选择, 例如选白色衬衫并不影响第二步的选择数量(仍有 4 条裤子可选). 对每一件衬衫都有 4 条裤子与之搭配, 有 3 件衬衫, 故总的搭配方式就是把每一步的选择数相乘. 这就是乘法原理.

定理 9.2 (乘法原理) 如果完成一件事需要依次分成 k 个步骤, 完成第一个步骤有 m_1 种不同的方法, 并且在任意已经完成前 $i-1$ 个步骤的选择之后, 第 i 个步骤总有 m_i 种不同的方法 $(i=2,\dots,k)$, 那么完成这件事共有

$$N = m_1 \times m_2 \times \cdots \times m_k$$

种不同的方法.

证明 按分步过程构造选择树. 第 1 步有 m_1 个分支; 对每一条已成前缀, 第 2 步都有 m_2 个分支, 故前两步共 $m_1 m_2$ 种前缀. 一般地, 若前 $i-1$ 步共有 $m_1 m_2 \cdots m_{i-1}$ 种前缀, 由题设每条前缀均可接上 m_i 种第 i 步选择, 故前 i 步共有

$$m_1 m_2 \cdots m_{i-1} m_i$$

种前缀. 对 $i = 1, 2, \dots, k$ 归纳, 完成全部 k 步的方法数为 $m_1 m_2 \cdots m_k$. □

从集合论的观点看, 乘法原理是**笛卡尔积**基数的直接体现. 为此先引入有序对与笛卡尔积.

一个**有序对** (a,b) 是由两个元素组成的集合, 但与普通集合 $\{a,b\}$ 不同, 它规定了元素的次序. 除非 $a=b$, 否则 $(a,b) \neq (b,a)$. 这是分步思想的数学抽象. 第一步的选择是第一个元素, 第二步的选择是第二个元素, 顺序是二者的决定性特征.

知乎@余命数

定义 9.1 (笛卡尔积) 设 A 和 B 为两个集合, 它们的笛卡尔积, 记作 $A \times B$(读作“A cross B”), 是所有可能的有序对 (a,b) 的集合, 其中 $a \in A$ 且 $b \in B$. 用集合符号表示为

$$A \times B = \{(a,b) \mid a \in A, b \in B\}$$

这个概念可以推广到任意 k 个集合. 设 $A_1, A_2, \ldots, A_k$ 为 k 个集合, 它们的笛卡尔积是所有可能的有序 k-元组 $(a_1, a_2, \ldots, a_k)$ 的集合, 其中对任意 $i = 1, 2, \ldots, k$, 都有 $a_i \in A_i$.

$$A_1 \times A_2 \times \cdots \times A_k = \{(a_1, a_2, \ldots, a_k) \mid a_i \in A_i \text{ 对于} i = 1, \ldots, k\}$$

笛卡尔积的大小, 即其元素的个数, 恰好就是乘法原理给出的结果. 若这些集合都是有限集, 则

$$|A_1 \times A_2 \times \cdots \times A_k| = |A_1| \times |A_2| \times \cdots \times |A_k| = \prod_{i=1}^{k} |A_i|$$

知乎@余命数

这个定义把分步完成任务这一直观操作数学化. 每一个步骤对应于从一个集合中选择一次, 而整个任务的所有可能结果构成一个笛卡尔积.

回到服装搭配的问题. 如果衬衫的集合是 $S = \{s_1, s_2, s_3\}$, 裤子的集合是 $P = \{p_1, p_2, p_3, p_4\}$, 那么一套搭配就是一个有序对 (s,p), 其中 $s \in S, p \in P$. 所有可能的搭配构成了集合 $S \times P$, 其大小为

$$|S \times P| = |S| \times |P| = 3 \times 4 = 12$$

下面用笛卡尔积写几个搭配计数的例子.

例题 9.6 一家餐厅的午市套餐允许顾客从 2 种汤中选择一种, 并从 3 种主菜中选择一种, 汤的集合为 $T = \{$罗宋汤, 奶油蘑菇汤$\}$, 主菜的集合为 $M = \{$牛排, 意面, 披萨$\}$. 请问共有多少种不同的套餐组合? 并写出这个组合集合.

知乎@余命数

解 每一种套餐组合都可以被视为一个有序对 (汤, 主菜). 所有可能的套餐组合构成的集合, 是汤的集合 T 与主菜的集合 M 的笛卡尔积 $T \times M$.

其大小为

$$|T \times M| = |T| \times |M| = 2 \times 3 = 6$$

这个集合的具体内容是

$$T \times M = \{(\text{罗宋汤}, \text{牛排}), (\text{罗宋汤}, \text{意面}), (\text{罗宋汤}, \text{披萨}),$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(奶油蘑菇汤, 牛排), (奶油蘑菇汤, 意面), (奶油蘑菇汤, 披萨)}

共有 6 种不同的套餐组合.

例题 9.7　组装一台电脑, CPU 有 2 种选择 ($C = \{\text{Intel}, \text{AMD}\}$), 内存有 3 种选择 ($R = \{8\text{GB}, 16\text{GB}, 32\text{GB}\}$), 硬盘有 2 种选择 ($H = \{1\text{TB SSD}, 2\text{TB SSD}\}$), 并且显卡也有 3 种选择 ($G = \{\text{GPU-A}, \text{GPU-B}, \text{GPU-C}\}$). 一种完整的配置方案由 CPU、内存、硬盘、显卡各选一种构成. 请问有多少种不同的配置方案?

解　每一种配置方案都可以被视为一个有序**四元组** (cpu, ram, hdd, gpu). 所有可能的配置方案集合是这四个选择集合的笛卡尔积 $C \times R \times H \times G$.

根据笛卡尔积的基数公式, 总的方案数是

$$|C \times R \times H \times G| = |C| \times |R| \times |H| \times |G| = 2 \times 3 \times 2 \times 3 = 36$$

所以, 共有 36 种不同的电脑配置方案.

9.1.3 原理的辨析与应用

区分加法原理和乘法原理是计数推理的常见起点. 一个简单的判断方法如下.

- **加法原理**关注分类. 完成任务的各种方式是平行的, 选择其中**任何一种**方式都能完成任务, 对应关键词**或者**.
- **乘法原理**关注分步. 完成任务需要经历**所有**步骤, 一步都不能少, 对应关键词**并且**.

在更复杂的问题中, 这两个原理往往需要结合使用.

例题 9.8　求在所有三位正整数中, 至少含有一个数字 7 的数有多少个?

解　直接计算至少含有一个 7 的数字个数会比较复杂. 须分类讨论恰好含有一个 7、恰好含有两个 7、恰好含有三个 7 三种情形, 思路可行却繁琐易错. 换一个角度往往能化繁为简.

不妨先计算它的**反面**, 即一个 7 都没有的三位正整数有多少个.

先看三位正整数的总数. 这是一个分步过程. 依次确定百位、十位、个位. 百位不能为 0, 有 9 种选择 (1 到 9); 十位和个位可取 0 到 9 中任意数字, 各有 10 种选择. 由乘法原理, 共有

$$9 \times 10 \times 10 = 900$$

个三位正整数.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

表 9.1 加法原理与乘法原理的比较

比较维度	加法原理	乘法原理
要领	**分类讨论**. 完成任务有若干类彼此独立的方法.	**分步完成**. 完成任务需要经过若干个连续进行的步骤.
关键词	或者、任选其一、情况一、情况二……、分为几类	并且、依次、接着、每一个……都……
任务关系	各类方法**相互排斥**. 选择任何一类中的一种方法, 任务即告完成.	各个步骤**相互依存**. 必须完成所有步骤, 任务才算完成.
判断准则	完成任务的方式是**择一即可**.	完成任务的步骤是**缺一不可**.
集合模型	**不交并集** $\lvert A_1 \cup \cdots \cup A_k\rvert = \sum_{i=1}^{k} \lvert A_i\rvert$, 其中 $A_i \cap A_j = \emptyset$ 对所有 $i \neq j$.	**笛卡尔积** $\lvert A_1 \times \cdots \times A_k\rvert = \prod_{i=1}^{k} \lvert A_i\rvert$.
经典比喻	从北京到上海, 可以坐火车**或**飞机.	搭配一套服装, 需要选上衣**和**裤子.

再算**不含数字 7** 的三位数. 同样分步. 百位不能是 0 且不能是 7, 有 8 种选择 (1,2,3,4,5,6,8,9); 十位和个位都不能是 7, 各有 9 种选择 (0,1,2,3,4,5,6,8,9). 由乘法原理, 不含数字 7 的三位数共有

$$8 \times 9 \times 9 = 648$$

个.

所有三位数可分成互斥的两类, 即至少含有一个 7 的和一个 7 也没有的, 二者之和即全部三位数. 由加法原理的变形, 也称**补集原理**, 是**容斥原理**的最简形式得

$$(\text{至少含有一个 7 的个数}) = (\text{总个数}) - (\text{一个 7 也没有的个数}),$$

$$N = 900 - 648 = 252,$$

即至少含有一个数字 7 的三位正整数共有 252 个.

从问题的反面入手,可以避开复杂的分类讨论. 上面的例子把一次加法与乘法的混合应用, 转化为两次乘法应用和一次减法. 这种策略在组合计数中相当实用.

接着看几个关于乘法原理的应用.

例题 9.9 令集合 $A = \{a_1, a_2, a_3\}$ 且 $B = \{1, 2, 3, 4, 5\}$. 从集合 A 到集合 B 可以定义多少个不同的函数 $f : A \to B$?

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 定义一个函数 $f: A \to B$, 就是为 A 中的**每一个**元素在 B 中指定**唯一一个**与之对应的元素, 称为它的像.A 有三个元素 a_1, a_2, a_3, 因此这项任务可分解为三个连续的决策. 先定 $f(a_1)$, 再定 $f(a_2)$, 最后定 $f(a_3)$.

这三步彼此独立.$f(a_1)$ 可取 B 中任意一个元素, 有 5 种选择; 同样 $f(a_2)$ 与 $f(a_3)$ 也各有 5 种选择. 由乘法原理, 方法数是各步选择数之积

$$N = 5 \times 5 \times 5 = 5^3 = 125,$$

故总共可定义 125 个不同的函数.

此例说明了一个普遍结论. 从一个包含 m 个元素的集合到一个包含 n 个元素的集合, 可以定义 n^m 个不同的函数.

例题 9.10 在一场田径比赛的决赛中, 共有 8 名选手. 假设没有并列名次, 那么总共有多少种可能的最终排名?

解 确定最终排名, 等价于为第 1 名到第 8 名这 8 个位置依次安排一位选手, 这是一个典型的分步过程.

第 1 名可从 8 位选手中产生, 有 8 种选择; 第 1 名一经确定, 第 2 名只能从剩下 7 位中产生, 有 7 种选择; 以此类推, 第 i 名有 $9-i$ 种选择, 直至第 8 名只剩最后一位选手, 仅 1 种选择. 由乘法原理, 所有可能的排名总数为

$$N = 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40320.$$

这个连乘积用专门的符号**阶乘**表示, 记作 $8!$.

将 n 个不同的物体进行全排列, 其方法数恰好是 $n! = n \times (n-1) \times \cdots \times 1$. 阶乘是乘法原理最直接的推论之一, 也是组合数与排列数计算中反复出现的工具.

例题 9.11 求正整数 $N = 720$ 的所有正因数的个数.

解 逐一列举 720 的因数既繁琐又易遗漏, 宜从其素数分解的结构入手. 由算术基本定理先作素数分解

$$720 = 72 \times 10 = (8 \times 9) \times (2 \times 5) = (2^3 \times 3^2) \times (2 \times 5) = 2^4 \cdot 3^2 \cdot 5^1,$$

故 720 的任一正因数 d 必具 $2^a \cdot 3^b \cdot 5^c$ 的形式, 其中 $0 \le a \le 4$、$0 \le b \le 2$、$0 \le c \le 1$.

构造一个因数, 等价于依次确定三个指数 a, b, c 的取值.a 可取 $0, 1, 2, 3, 4$ 中任一值, 共 $4+1=5$ 种选择;b 可取 $0, 1, 2$ 中任一值, 共 $2+1=3$ 种选择;c 可取 $0, 1$ 中任一值, 共 $1+1=2$ 种选择. 三次决策相互独立, 由乘法原理得正因数的总个数为

$$\tau(720) = (4+1) \times (2+1) \times (1+1) = 5 \times 3 \times 2 = 30,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即 720 共有 30 个正因数.

这个方法具有普适性. 对于任意正整数 $N = p_1^{a_1} p_2^{a_2} \dots p_k^{a_k}$, 其正因数的个数为 $\tau(N) = (a_1+1)(a_2+1)\dots(a_k+1)$, 这也是乘法原理在数论中的一项应用.

例题 9.12 (路径着色) 有 4 个排成一排的方格, 现在用 5 种不同的颜色对其进行染色, 要求相邻的方格颜色不能相同. 问有多少种不同的染色方法?

解 将 4 个方格从左到右依次编号为 1, 2、3、4, 染色是一个分步决策过程, 每一步只需避开左邻一格的颜色.

第 1 格无任何限制, 可从 5 种颜色中任选一种, 有 5 种方法. 此后第 i 格 ($i = 2,3,4$) 只需与第 $i-1$ 格不同色, 无论第 $i-1$ 格具体选了哪种颜色, 都恰好排除一种可能, 故各有 $5-1=4$ 种选择. 由乘法原理, 总的染色方法数为

$$N = 5 \times 4 \times 4 \times 4 = 5 \times 4^3 = 320,$$

即共有 320 种不同的染色方法.

这个问题的要点在于, 每一步选择的**数量**是固定的, 不受前一步**具体选择**的影响, 例如, 无论第 1 格是红色还是蓝色, 第 2 格都剩下 4 种选择. 如果问题的限制条件更为复杂, 例如首尾方格颜色不同, 则可能需要更精细的分类讨论, 即加法与乘法原理的结合.

例题 9.13 (集合的子集个数) 一个包含 n 个不同元素的有限集合 S, 共有多少个不同的子集?(子集包括空集 $\emptyset$ 和集合 S 本身)

解 此题有一个著名答案, 其证法之一恰巧运用乘法原理, 要点在于转变思考角度.

传统想法是挑选元素组成子集. 挑选 0 个元素的子集, 挑选 1 个元素的子集, ..., 再用加法原理求和. 这条路可行但更繁, 第9.2节 会用组合数给出另一种算法.

换一个角度. 对集合 S 中的**每一个**元素, 逐一决定它的归属. 设 $S = \{s_1, s_2, \dots, s_n\}$, 为构造一个子集 $A \subseteq S$, 依次对每个元素决策其是否属于 A. 对 s_1 有属于或不属于两种选择, 对 s_2 亦然, 且这一决策与对 s_1 的决策无关; 如此继续, 直至对 s_n 仍有两种选择. 当 n 个决策全部完成, 一个确定的子集就被唯一地构造出来. 若对所有元素都判定不属于, 得到空集 $\emptyset$; 若都判定属于, 则得到 S 本身.

这 n 个决策相互独立, 各有 2 种选择, 由乘法原理得总方法数

$$N = \underbrace{2 \times 2 \times \dots \times 2}_{n\text{次}} = 2^n,$$

故一个包含 n 个元素的集合, 其子集的个数为 2^n.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个证明展示了一种常用的组合技巧. 把从整体中挑选一部分的观点, 改为对每个个体进行独立决策, 从而把一个分类求和问题转化为一个连续决策问题. 这种视角转换在解决高级计数问题时相当常用.

知乎@余命数

9.2 组合与排列

在上一节, 我们用加法原理和乘法原理处理了计数问题的两类基本情形. 本节用这两条原理定义并计算两个对象, 即**组合**与**排列**.

回到本章开头提出的彩票问题.

从 33 个红球中不重复地选出 6 个, 一共有多少种选法?

这里我们只关心**选出了哪些**号码,不在乎它们的**先后顺序**. 选出集合 $\{5,8,13,21,27,32\}$ 与选出集合 $\{32,27,21,13,8,5\}$ 是完全等价的. 这种只选不排的问题, 正是我们要研究的对象.

知乎@余命数

9.2.1 组合

知乎@余命数　知乎@余命数

定义 9.2 (组合) 从 n 个不同的元素中, **不计顺序**地选取 k 个元素组成一个子集, 称为一个从 n 个元素中取出 k 个元素的**组合**. 所有不同组合的个数, 称为组合数, 记作 $C(n,k)$ 或更常用的 $\binom{n}{k}$(读作"从 n 中选 k 个").

知乎@余命数　知乎@余命数

关于组合数的记号

组合数在不同的教材和领域中有多种常见的记法, 例如 $C(n,k)$, C_n^k, 甚至 C_k^n. 它们都表示同一个数学概念.

在本书中, 我们主要采用国际通用的**二项式系数**记法

$$\binom{n}{k}$$

这个符号读作"从 n 中选 k 个", 它与常见记法满足如下恒等关系

$$\binom{n}{k} \equiv C(n,k) \equiv C_n^k$$

知乎@余命数

直接计数 $\binom{n}{k}$ 并不容易, 但可以借助乘法原理, 用一个思想实验推导出来. 以彩票问题为例, 计算 $\binom{33}{6}$.

思想实验从顺序重要的情形入手. 设想从袋子里一个接一个地、**按顺序**取出 6 个球. 根据乘法原理逐步计数.

- 第一个球有 33 种选择;

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 第二个球有 32 种选择;
- …
- 第六个球有 $33-6+1=28$ 种选择.

所以, 有序的选法总数为 $33\times32\times31\times30\times29\times28$.

现在回到顺序不重要的现实. 考虑一个具体的组合, 例如集合 $\{1,2,3,4,5,6\}$. 在刚才的有序计数中, 这个组合对应于这 6 个数字的全部全排列. $(1,2,3,4,5,6)$、$(1,2,3,4,6,5)$、…、$(6,5,4,3,2,1)$. 这 6 个数字的全排列数, 根据乘法原理, 是 $6\times5\times4\times3\times2\times1=6!$(6 的阶乘).

算出的有序选法总数, 把每一个真正想要的无序组合都重复计数了 $6!$ 次. 因此, 组合数等于有序选法总数除以这个重复倍数.

$$\binom{33}{6}=\frac{33\times32\times31\times30\times29\times28}{6\times5\times4\times3\times2\times1}$$

这个逻辑可以推广到一般情况. 从 n 个元素中按顺序选出 k 个, 有 $n(n-1)\cdots(n-k+1)$ 种方法. 而这 k 个被选出的元素自身的全排列数为 $k!$. 每一个组合都对应着 $k!$ 个有序的选法. 因此, 得组合数的计算公式.

定理 9.3 (组合数公式) 从 n 个不同元素中取出 k 个元素的组合数 $(0\le k\le n)$ 为

$$\binom{n}{k}=\frac{n(n-1)\cdots(n-k+1)}{k!}=\frac{n!}{k!(n-k)!}$$

其中 $n!$ 表示 n 的阶乘, $n!=n\times(n-1)\times\cdots\times1$. 特别地, 我们定义 $0!=1$.

证明 用双重计数. 对同一计数问题以两种方式计算, 所得必等, 由此解出未知量.

所计对象是从含 n 个不同元素的集合 S 中选出 k 个元素并按顺序排列的方法总数, 暂记为 $A(n,k)$, 即下一小节将定义的**排列数**.

法一直接用乘法原理. 构造一个 k-排列即填充 k 个有序位置. 第一位有 n 种选法, 第二位有 $n-1$ 种, …, 第 k 位有 $n-k+1$ 种. 故

$$A(n,k)=n(n-1)\cdots(n-k+1).$$

分子分母同乘 $(n-k)!$, 写成阶乘形式

$$A(n,k)=\frac{n!}{(n-k)!}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

法二经组合数间接计算. 构造 k-排列可分两步. 先从 n 个元素中不计顺序选出 k 个, 有 $\binom{n}{k}$ 种;再对这 k 个作全排列, 有 $k!$ 种. 由乘法原理,

$$A(n,k)=\binom{n}{k}\times k!.$$

两式相等, 得

$$\binom{n}{k}\times k!=\frac{n!}{(n-k)!}.$$

两边同除以 $k!$($k\geq 0$ 时 $k!\geq 1$), 即

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}.$$

若改用 $A(n,k)=n(n-1)\cdots(n-k+1)$, 则得常用形式

$$\binom{n}{k}=\frac{n(n-1)\cdots(n-k+1)}{k!}.$$

特别地, $k=0$ 时选出空集仅一种方法, 公式给出 $\binom{n}{0}=\frac{n!}{0!\,n!}=1$, 与事实相符, 亦说明定义 $0!=1$ 的合理性. □

现在可以回答彩票问题了.

$$\binom{33}{6}=\frac{33\times 32\times 31\times 30\times 29\times 28}{6\times 5\times 4\times 3\times 2\times 1}=1,107,568$$

挑选红球的方法数超过了一百万种, 而这只是中头奖的第一步. 完成整张彩票, 还要从 16 个蓝球中选出 1 个, 有 $\binom{16}{1}=16$ 种选法. 根据乘法原理, 全部可能的号码组合共有

$$\binom{33}{6}\times\binom{16}{1}=1,107,568\times 16=17,721,088$$

种. 全部组合超过一千七百万种, 而其中只有一种是当期的头奖号码. 至于中头奖的概率应当如何严格定义与计算, 我们还将在 第10章 中回到这个问题.

在结束组合的讨论之前, 再补充组合数的一个重要性质, 它能简化计算.

定理 9.4 (组合数的对称性) 对于所有满足 $0\leq k\leq n$ 的整数 n,k, 有

$$\binom{n}{k}=\binom{n}{n-k}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 代数上, 由组合数公式直接可得

$$\binom{n}{n-k} = \frac{n!}{(n-k)!\,(n-(n-k))!} = \frac{n!}{(n-k)!\,k!} = \binom{n}{k}$$

这一等式的组合意义同样清楚. 从 n 个元素中选出 k 个带走, 与选出 $n-k$ 个留下, 是同一个决定的两种说法, 两种计数必然相等. □

这个性质的实际价值在于化大为小. 当 k 接近 n 时, 直接计算 $\binom{n}{k}$ 的分子分母都很长, 转而计算 $\binom{n}{n-k}$ 就轻快得多. 例如

$$\binom{100}{98} = \binom{100}{2} = \frac{100\times 99}{2\times 1} = 4950$$

知乎@余命数

9.2.2 排列

理解了组合之后, 排列便很自然. 排列关心选出了哪些元素, 并关心这些元素被排成的**顺序**.

定义 9.3 (排列) 从 n 个不同的元素中, 选取 k 个元素并按照**一定的顺序**排成一列, 称为一个从 n 个元素中取出 k 个元素的**排列**. 所有不同排列的个数, 称为排列数. 排列数的常见记号有 $A(n,k)$、A_n^k、$P(n,k)$、P_n^k 等; 与组合数的情形类似, 本书统一使用 $A(n,k)$.

计算排列数 $A(n,k)$, 把完成一次排列看作两个步骤.

1. 先从 n 个元素中, 选出要参与排列的 k 个元素. 这一步有 $\binom{n}{k}$ 种方法.
2. 将选出的 k 个元素进行全排列. 这一步有 $k!$ 种方法.

必须完成**所有**步骤, 任务才算完成, 由乘法原理得到排列数与组合数之间的基本关系

$$A(n,k) = \binom{n}{k} \times k!$$

这个关系说明了排列与组合的联系. 一个排列, 实质上是一个组合再加上对所选元素的排序.

将组合数的公式代入, 我们就能得到排列数的计算公式.

定理 9.5 (排列数公式) 从 n 个不同元素中取出 k 个元素的排列数 $(0 \le k \le n)$ 为

$$A(n,k) = \frac{n!}{k!(n-k)!} \times k! = \frac{n!}{(n-k)!} = n(n-1)\cdots(n-k+1)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 目标是计算从 n 个不同元素中，按顺序取出 k 个元素的所有方法的总数，即 $A(n,k)$. 把构造这样一个排列的过程视为一个分步完成的任务，直接应用乘法原理.

设想有 k 个空置的位置，需从一个含 n 个元素的集合 S 中选元素填充，且每个元素最多用一次.

填充位置 1 时，可从集合 S 中任选一个元素，有 n 种选择. 填充位置 2 时，集合 S 中只剩下 $n-1$ 个元素可用，故有 $n-1$ 种选择. 填充位置 3 时，已用去两个元素，剩下 $n-2$ 个元素可用，故有 $n-2$ 种选择. 依此类推，填充第 k 个（也是最后一个）位置时，已用去 $k-1$ 个元素，剩下 $n-(k-1)=n-k+1$ 个元素可用，故有 $n-k+1$ 种选择.

根据乘法原理，完成这 k 个步骤的总方法数是每一步方法数的乘积. 因此，

$$A(n,k)=n\times(n-1)\times(n-2)\times\cdots\times(n-k+1)$$

这证明了公式的第一个等式.

为得到阶乘形式的表达式，对上式做代数变换，在分子和分母上同时乘以 $(n-k)!$.

$$A(n,k)=n(n-1)\cdots(n-k+1)\times\frac{(n-k)(n-k-1)\cdots1}{(n-k)(n-k-1)\cdots1}$$

分子部分现在是一个从 n 开始到 1 的完整连乘积，即 $n!$. 分母部分则是 $(n-k)!$. 所以，

$$A(n,k)=\frac{n!}{(n-k)!}$$

这证明了公式的第二个等式.

最后，由上一节推导出的排列与组合的关系 $A(n,k)=\binom{n}{k}\times k!$ 验证一致性.

$$\binom{n}{k}\times k!=\frac{n!}{k!(n-k)!}\times k!=\frac{n!}{(n-k)!}=A(n,k)$$

这既验证了公式，也再次说明一个 k-排列可以看作先进行一次 k-组合，再对选出的元素进行一次全排列. □

例题 9.14 一个班级有 10 名学生，需要从中选出 3 人，分别担任班长、学习委员和体育委员. 有多少种不同的选举结果?

解 此题是典型的排列问题，因为职位不同，人选的顺序至关重要. 例如，张三任班长、李四任学委，与李四任班长、张三任学委，是两种不同的结果.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

直接套用排列数公式, 即从 10 人中选出 3 人排列, 得

$$A(10,3)=\frac{10!}{(10-3)!}=\frac{10!}{7!}=10\times 9\times 8=720.$$

也可走先组合、后排序的路径. 先从 10 人中不计顺序地选出 3 人, 有 $\binom{10}{3}=\dfrac{10\times 9\times 8}{3\times 2\times 1}=120$ 种方法; 再为这 3 人分配 3 个不同职位, 有 $3!=6$ 种方法. 由乘法原理, 总的选举结果为

$$N=\binom{10}{3}\times 3!=120\times 6=720,$$

与直接使用排列数公式的结果一致, 正说明排列与组合的联系.

9.3 二项式定理

我们在上一节定义了组合数 $\binom{n}{k}$, 即 n 个不同元素中选取 k 个的方案数. 这个符号也被称为**二项式系数**, 与形如 $(x+y)$ 的二项式有内在联系. 本节说明这一联系, 它集中体现在二项式定理上.

从简单的例子开始, 观察展开式的系数规律.

$$\begin{aligned}
(x+y)^0&=1\\
(x+y)^1&=1x+1y\\
(x+y)^2&=1x^2+2xy+1y^2\\
(x+y)^3&=1x^3+3x^2y+3xy^2+1y^3\\
(x+y)^4&=1x^4+4x^3y+6x^2y^2+4xy^3+1y^4
\end{aligned}$$

观察这些展开式, 系数 $1,4,6,4,1$ 恰好是 $\binom{4}{0},\binom{4}{1},\binom{4}{2},\binom{4}{3},\binom{4}{4}$ 的值. 这并非巧合.

定理 9.6 (二项式定理) 对于任意可交换的变量或数 x,y 和任意非负整数 n, 下述展开式成立.

$$(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k$$

展开来看, 即

$$(x+y)^n=\binom{n}{0}x^n+\binom{n}{1}x^{n-1}y+\binom{n}{2}x^{n-2}y^2+\cdots+\binom{n}{n-1}xy^{n-1}+\binom{n}{n}y^n$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 考虑乘积的完全展开形式

$$(x+y)^n = \underbrace{(x+y)(x+y)\cdots(x+y)}_{n\text{个因子}}$$

根据乘法分配律, 展开后的每一项都是通过从 n 个 $(x+y)$ 因子中, 各自选择 x 或 y 相乘得到的. 例如, 要得到 x^n 这一项, 须从每一个因子中都选择 x.

现在考虑一个更普遍的项, 形如 $x^{n-k}y^k$ 是如何产生的. 为构成这一项, 必须从 n 个 $(x+y)$ 因子中.

- 恰好从 k 个因子中选择 y.
- 从其余的 $n-k$ 个因子中选择 x.

每一种这样的选择方案都贡献一个 $x^{n-k}y^k$ 项. 因此, 展开式中 $x^{n-k}y^k$ 的系数, 等于构成该项的不同选择方案的总数.

这个选择方案的总数是一个熟悉的问题. 它等价于从 n 个不同的因子中, 选出 k 个来提供 y 的方案数.

根据组合的定义, 这个方案数恰好是 $\binom{n}{k}$.

因此, $x^{n-k}y^k$ 项的系数必然是 $\binom{n}{k}$.

k 可从 0(一个 y 都不选, 得到 x^n)取到 n, 全部选 y 时则得到 y^n, 把所有这些项加起来, 就得到二项式定理的完整形式

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k}y^k$$

□

这个证明解释了组合数为何会作为系数出现在二项式的展开式中.

9.3.1 杨辉三角

将二项式系数 $\binom{n}{k}$ 排列成一个三角形的阵列, 我们就得到了著名的杨辉三角.

杨辉三角最引人注目的性质是, **每个数都等于它肩上两个数之和**. 例如, 在 $n=4$ 行, 6 是由它上一行 ($n=3$) 中的 3 和 3 相加得到的. 这个性质在数学上称为**帕斯卡恒等式**.

定理 9.7 (帕斯卡恒等式) 对于所有满足 $n \geq 1$ 且 $1 \leq k \leq n-1$ 的整数 n, k, 有

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

代数代入组合数公式再通分即可验证这个恒等式. 下面给出一个组合学证明.

证明 考虑一个组合问题. 从 n 个人中选出 k 人组成班委会, 方法数是 $\binom{n}{k}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

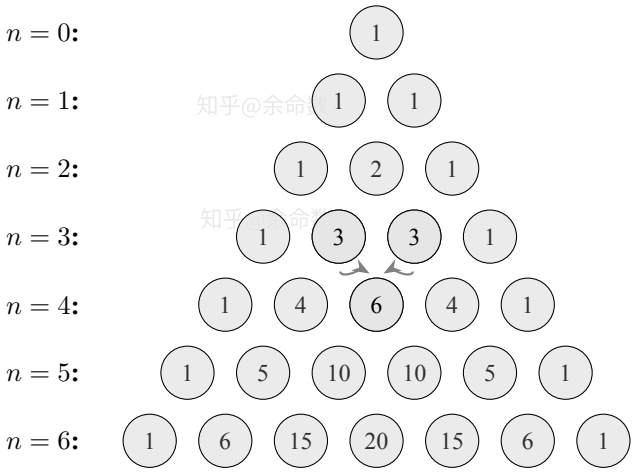


图 9.2 注意恒等式 $\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k}$ 的一个例子, $\binom{4}{2}=\binom{3}{1}+\binom{3}{2}$

换一种方式计数. 指定一个特殊的人李华, 最终选出的委员会只有两种互斥情况.

- 李华在班委中. 此时只需从剩下 $n-1$ 个人中再选出 $k-1$ 个人填补剩余名额, 方法数是 $\binom{n-1}{k-1}$.
- 李华不在班委中. 此时必须从剩下 $n-1$ 个人中选出全部 k 个班委成员, 方法数是 $\binom{n-1}{k}$.

由加法原理, 选班委的总方法数等于上述两种互斥情况的方法数之和. 故

$$\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k}$$

□

9.3.2 二项式系数的性质

再来观察杨辉三角, 可从中猜测出二项式系数的一系列规律. 每一行从左读到右与从右读到左完全相同; 每一行的数先增大、后减小, 在正中间达到最大; 每一行所有数的和恰好是上一行的两倍. 这些从数值表中归纳出的猜想, 都可以借助组合数公式和二项式定理严格证明.

首先是**对称性**, $\binom{n}{k}=\binom{n}{n-k}$, 这是杨辉三角每一行左右镜像的原因. 我们已在第9.2节 中证明这一性质; 在二项式定理中交换 x 与 y 的位置, 展开式也保持不变.

接下来看每一行的和所对应的规律. 证明它的关键技巧在于, 二项式定理是一个关于 x,y 的**恒等式**, 对其中的变量赋予特殊的值, 就能得到各种各样的组合恒等式. 这种方法称为**赋值法**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

定理 9.8 (二项式系数之和) 对于任意非负整数 n, 有

$$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\cdots+\binom{n}{n}=2^n$$

知乎@余命数

证明 在二项式定理 $(x+y)^n=\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k$ 中令 $x=y=1$,左边变为 $(1+1)^n=2^n$, 右边恰为全部二项式系数之和. □

这个等式其实是已知事实的另一种表述. 在 第9.1节 中, 我们曾用乘法原理证明, 一个 n 元集合的子集共有 2^n 个. 另一方面, 按子集的大小分类计数, 大小为 k 的子集有 $\binom{n}{k}$ 个, 对 k 从 0 到 n 求和同样数出了全部子集. 两种方法数的是同一个对象, 结果自然相等. 这是推导组合数公式时用过的双重计数方法.

定理 9.9 (二项式系数的交错和) 对于任意正整数 n, 有

$$\binom{n}{0}-\binom{n}{1}+\binom{n}{2}-\cdots+(-1)^n\binom{n}{n}=0$$

从而奇数项二项式系数之和等于偶数项二项式系数之和, 二者都等于 2^{n-1}.

@余命数

知乎@余命数

证明 在二项式定理中令 $x=1, y=-1$. 由于 $n\geq 1$, 左边为 $(1+(-1))^n=0$, 右边恰为交错和 $\sum_{k=0}^{n}(-1)^k\binom{n}{k}$, 故交错和为 0. 把所有偶数项系数留在一边、奇数项系数移到另一边, 即得奇数项之和等于偶数项之和; 又由二项式系数之和为 2^n, 所以每一方都等于 $2^n/2=2^{n-1}$. □

知乎@余命数

最后是先增后减、中间最大的规律, 这在数学上称为**单峰性**.

定理 9.10 (二项式系数的单峰性) 对于固定的正整数 n, 二项式系数序列

$$\binom{n}{0},\binom{n}{1},\ldots,\binom{n}{n}$$

先严格递增, 再严格递减. 当 n 为偶数时, 最大值在正中间一项 $\binom{n}{n/2}$ 处取得; 当 n 为奇数时, 中间两项 $\binom{n}{(n-1)/2}$ 与 $\binom{n}{(n+1)/2}$ 相等且同为最大值.

证明 考察相邻两项的比值. 对于 $1\leq k\leq n$,

$$\frac{\binom{n}{k}}{\binom{n}{k-1}}=\frac{n!}{k!(n-k)!}\cdot\frac{(k-1)!\,(n-k+1)!}{n!}=\frac{n-k+1}{k}$$

知乎@余命数

这个比值大于 1 当且仅当 $n-k+1>k$, 即 $k<\dfrac{n+1}{2}$. 因此

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 当 $k < \dfrac{n+1}{2}$ 时, $\binom{n}{k} > \binom{n}{k-1}$, 序列严格递增;
- 当 $k > \dfrac{n+1}{2}$ 时, $\binom{n}{k} < \binom{n}{k-1}$, 序列严格递减.

若 n 为偶数, 则 $\dfrac{n+1}{2}$ 不是整数, 序列在 $k = \dfrac{n}{2}$ 处取得唯一的最大值 $\binom{n}{n/2}$. 若 n 为奇数, 则当 $k = \dfrac{n+1}{2}$ 时比值恰为 1, 即中间两项相等, 它们同时成为最大值 $\binom{n}{(n-1)/2} = \binom{n}{(n+1)/2}$. 这也再次印证了对称性. □

单峰性说明, $(x+y)^n$ 展开式中最中间的项系数最大. 例如 $(x+y)^{10}$ 的展开式中, 系数最大的项是 $\binom{10}{5}x^5y^5 = 252x^5y^5$. 这一性质在二项分布等后续内容中还会反复出现.

9.3.3 杨辉三角的奇偶图案 *(选读)

把杨辉三角中每个二项式系数按奇偶涂色, 奇数涂黑, 偶数留白, 得到的图形会向自己相似, 每往下一层, 图案都由三个上一层的图案拼成, 中间留出一块空白. 本节数清每一行里究竟有多少个奇数.

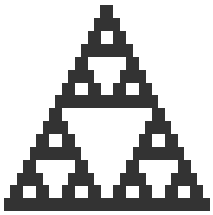


图 9.3 杨辉三角第 0 到第 15 行的奇偶图案, 黑格表示奇数. 图案自相似, 每一层由三个上一层图案拼成, 中间留空.

先看指数是 2 的幂的那些行. 在模 2 的意义下计算 $(1+x)^{2^k}$, 用平方代替乘法.

$$(1+x)^{2^k} = \left((1+x)^{2^{k-1}}\right)^2 \equiv \left(1+x^{2^{k-1}}\right)^2 = 1 + 2x^{2^{k-1}} + x^{2^k} \equiv 1 + x^{2^k} \pmod 2.$$

最右边只有两项, 中间各项的系数都是偶数. 反复使用这一步, 从 $k=0$ 推起, 可知第 2^k 行里只有两端的两个 1 是奇数.

再看一般的行. 记 n 的二进制写成 $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_s}$, 其中 $k_1 > k_2 > \cdots > k_s \geq 0$, s 是 n 的二进制中 1 的个数. 由上一段的结论,

$$(1+x)^n = \prod_{i=1}^{s} (1+x)^{2^{k_i}} \equiv \prod_{i=1}^{s} \left(1 + x^{2^{k_i}}\right) \pmod 2.$$

右边每个括号贡献一次选择, 要么取 1, 要么取 $x^{2^{k_i}}$. 由于各个 2^{k_i} 互不相同, 不同的选择得到的指数互不相同, 展开后不会有两项合并. 于是右边展开后恰好有 2^s 项系数为 1, 其余项系数为 0.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 9.11 (一行的奇数个数) 设 n 的二进制表示中 1 的个数为 s, 则杨辉三角第 n 行中奇数的个数等于 2^s. 特别地, 第 2^k-1 行的 2^k 个数全是奇数.

第 2^k-1 行的二进制的 k 位全是 1, 故 $s=k$, 奇数个数为 2^k, 恰好等于该行的项数, 整行都是奇数. 第 2^k 行则恰好转过一轮, 只剩两端两个奇数, 中间全为偶数.

例题 9.15 求杨辉三角第 100 行中奇数的个数.

解 把 100 写成二进制

$$100=64+32+4=1100100_{(2)},$$

其中 1 的个数 $s=3$, 故第 100 行中奇数有 $2^3=8$ 个.

奇数的占比还可以整体估算. 在第 0 到第 2^k-1 行中, 二进制含 j 个 1 的行共有 $\binom{k}{j}$ 行, 每行的奇数个数为 2^j, 于是这些行里的奇数总数为

$$\sum_{j=0}^{k}\binom{k}{j}2^j=3^k.$$

而这 2^k 行的项数共为 $1+2+\cdots+2^k=2^k\left(2^k+1\right)/2$. 两者相除得

$$\frac{3^k}{2^k\left(2^k+1\right)/2}=\frac{2\cdot 3^k}{2^k\left(2^k+1\right)},$$

分母中的 2^{2k} 增长更快, 比值的极限为 0. 行数越多, 奇数所占的比例越接近于零, 杨辉三角里几乎所有的数都是偶数.

9.4 从三项式定理到多项式定理 *

一旦得出二项式定理, 自然会考虑是否存在类似的三项式定理、四项式定理…. 沿用证明二项式定理时的组合思想来分析 $(x+y+z)^n$ 的展开式.

$$(x+y+z)^n=\underbrace{(x+y+z)(x+y+z)\cdots(x+y+z)}_{n\text{个因子}}$$

展开后的任意一项必然具有 $x^{k_1}y^{k_2}z^{k_3}$ 的形式. 由于从 n 个因子中各取一个变量相乘, 这些变量的指数总和必须等于 n, 即

$$k_1+k_2+k_3=n\quad(\text{其中}k_1,k_2,k_3\text{为非负整数})$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

接下来确定项 $x^{k_1}y^{k_2}z^{k_3}$ 的系数. 这个系数等于从 n 个因子中, 选出 k_1 个提供 x、选出 k_2 个提供 y、再选出 k_3 个提供 z 的总方案数. 可以分步计数来解决.

1. 为 x 选择因子. 从 n 个可用的因子中, 选出 k_1 个. 这有 $\binom{n}{k_1}$ 种方法.
2. 为 y 选择因子. 从剩下的 $n-k_1$ 个因子中, 选出 k_2 个. 这有 $\binom{n-k_1}{k_2}$ 种方法.
3. 为 z 选择因子. 最后剩下的 $n-k_1-k_2=k_3$ 个因子, 全部用于提供 z. 这只有 $\binom{k_3}{k_3}=1$ 种方法.

根据乘法原理, 总方案数为

$$\begin{aligned}\binom{n}{k_1}\binom{n-k_1}{k_2} &= \frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\\ &= \frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!k_3!}\\ &= \frac{n!}{k_1!k_2!k_3!}\end{aligned}$$

这个结果引出了三项式系数的定义.

定义 9.4 (三项式系数) 对于满足 $k_1+k_2+k_3=n$ 的非负整数 k_1,k_2,k_3, 三项式系数被定义为

$$\binom{n}{k_1,k_2,k_3}=\frac{n!}{k_1!k_2!k_3!}$$

分组问题

三项式系数 $\binom{n}{k_1,k_2,k_3}$ 有一个非常直观的组合解释. 它表示将 n 个不同的物体分成三组, 各组物体的数量分别为 k_1,k_2,k_3 的方法数. 这与上面的推导过程(将 n 个因子分成三组)完全一致. 它也等价于对包含 k_1 个 x、k_2 个 y 和 k_3 个 z 的多重集进行全排列的数目.

9.4.1 多项式定理

有了三项式系数的定义, 可将二项式定理正式推广到三项式.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 9.12 (三项式定理) 对于任意可交换的变量或数 x, y, z 和任意非负整数 n, 下述展开式成立.

$$(x+y+z)^n = \sum_{k_1+k_2+k_3=n} \binom{n}{k_1, k_2, k_3} x^{k_1} y^{k_2} z^{k_3}$$

其中求和遍历所有满足条件的非负整数解 (k_1, k_2, k_3).

在证明之前, 先弄清三项式定理中求和符号 $\sum_{k_1+k_2+k_3=n}$ 的含义.

这个记号表示, 需要对所有满足以下两个条件的有序整数三元组 (k_1, k_2, k_3) 进行求和.

1. k_1, k_2, k_3 均为**非负整数**, 即 $k_1 \geq 0, k_2 \geq 0, k_3 \geq 0$.
2. 它们的和恰好等于 n, 即 $k_1 + k_2 + k_3 = n$.

这个过程也被称为"遍历 n 的所有整数分拆到三个部分".

以具体的例子 $n = 2$ 来阐明. 需要找出所有满足 $k_1 + k_2 + k_3 = 2$ 的非负整数解 (k_1, k_2, k_3), 系统地列出它们.

- 如果 $k_1 = 2$, 那么 $k_2 = 0, k_3 = 0$. 得到 $(2, 0, 0)$.
- 如果 $k_1 = 1$, 那么 $k_2 + k_3 = 1$. 这有两种可能, 即 $(1, 1, 0)$ 和 $(1, 0, 1)$.
- 如果 $k_1 = 0$, 那么 $k_2 + k_3 = 2$. 这有三种可能, 即 $(0, 2, 0), (0, 1, 1), (0, 0, 2)$.

所以, 对于 $n = 2$, 求和符号意味着将对应于以下 6 个三元组的项全部加起来.

$$(2,0,0), (0,2,0), (0,0,2), (1,1,0), (1,0,1), (0,1,1)$$

因此, $(x+y+z)^2$ 的展开式实际上是

$$\begin{aligned}\sum_{k_1+k_2+k_3=2} \binom{2}{k_1, k_2, k_3} x^{k_1} y^{k_2} z^{k_3} = \\ \binom{2}{2,0,0} x^2 y^0 z^0 + \binom{2}{0,2,0} x^0 y^2 z^0 + \binom{2}{0,0,2} x^0 y^0 z^2 \\ + \binom{2}{1,1,0} x^1 y^1 z^0 + \binom{2}{1,0,1} x^1 y^0 z^1 + \binom{2}{0,1,1} x^0 y^1 z^1 \\ = 1x^2 + 1y^2 + 1z^2 + 2xy + 2xz + 2yz\end{aligned}$$

这与直接用分配律展开得到的结果一致.

证明 用组合论证确定展开式中任意一个通项 $x^{k_1} y^{k_2} z^{k_3}$ 的系数, 其中 k_1, k_2, k_3 是满足 $k_1 + k_2 + k_3 = n$ 的非负整数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

考虑乘积的展开形式

$$(x+y+z)^n = \underbrace{(x+y+z)(x+y+z)\cdots(x+y+z)}_{n\text{个因子}}$$

根据乘法分配律, 展开后的每一项都是从这 n 个 $(x+y+z)$ 因子中的每一个里选择一个变量(x, y 或 z)相乘得到的.

为构成一个具体的项 $x^{k_1}y^{k_2}z^{k_3}$, 必须

- 恰好从 n 个因子中的 k_1 个因子里选择变量 x.
- 恰好从剩下的 $n-k_1$ 个因子中的 k_2 个因子里选择变量 y.
- 恰好从最后剩下的 $n-k_1-k_2=k_3$ 个因子里选择变量 z.

因此, 项 $x^{k_1}y^{k_2}z^{k_3}$ 在最终展开式中出现的次数, 即它的系数, 等于完成上述三步选择的总方案数.

用乘法原理计算这个方案数, 得

1. **为 x 选择因子.** 从 n 个不同的因子中, 选出 k_1 个来提供变量 x. 这是一个组合问题, 方法数是 $\binom{n}{k_1}$.
2. **为 y 选择因子.** 为 x 选好因子后, 还剩下 $n-k_1$ 个因子. 从这些剩余因子中, 选出 k_2 个来提供变量 y. 方法数是 $\binom{n-k_1}{k_2}$.
3. **为 z 选择因子.** 最后剩下 $n-k_1-k_2=k_3$ 个因子. 从这 k_3 个因子中选出 k_3 个来提供变量 z. 方法数是 $\binom{k_3}{k_3}=1$.

根据乘法原理, 总方案数是以上三步方法数的乘积.

$$\begin{aligned}
\text{系数} &= \binom{n}{k_1}\binom{n-k_1}{k_2}\binom{k_3}{k_3} \\
&= \frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdot\frac{k_3!}{k_3!0!} \\
&= \frac{n!}{k_1!\cancel{(n-k_1)!}}\cdot\frac{\cancel{(n-k_1)!}}{k_2!(k_3)!}\cdot 1 \quad (\text{因为} n-k_1-k_2=k_3) \\
&= \frac{n!}{k_1!k_2!k_3!} \\
&= \binom{n}{k_1,k_2,k_3}
\end{aligned}$$

已经证明, 对于任意满足 $k_1+k_2+k_3=n$ 的非负整数组 (k_1,k_2,k_3), 其对应项 $x^{k_1}y^{k_2}z^{k_3}$ 的系数恰好是三项式系数 $\binom{n}{k_1,k_2,k_3}$.

把所有这些可能的项相加, 就构成 $(x+y+z)^n$ 的完整展开式. 定理得证. □

例题 9.16 求在 $(2x-y+3z^2)^5$ 的展开式中, x^2yz^4 项的系数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 将表达式视为 $(A+B+C)^5$，其中 $A=2x, B=-y, C=3z^2$，所求项 x^2yz^4 应具 $A^{k_1}B^{k_2}C^{k_3}$ 的形式，$k_1+k_2+k_3=5$. 由 $(2x)^{k_1}$ 贡献 x^2 得 $k_1=2$；由 $(-y)^{k_2}$ 贡献 y^1 得 $k_2=1$；由 $(3z^2)^{k_3}$ 贡献 z^4，即 $z^{2k_3}=z^4$，得 $k_3=2$. 指数和 $k_1+k_2+k_3=2+1+2=5$，与 $n=5$ 相符.

据三项式定理，含 $A^2B^1C^2$ 的完整项为 $\binom{5}{2,1,2}A^2B^1C^2$，其中三项式系数

$$\binom{5}{2,1,2}=\frac{5!}{2!\,1!\,2!}=\frac{120}{2\cdot 1\cdot 2}=30.$$

代回 A,B,C 并整理得

$$\begin{aligned}30\cdot(2x)^2\cdot(-y)\cdot(3z^2)^2&=30\cdot(4x^2)\cdot(-y)\cdot(9z^4)\\&=(30\cdot 4\cdot(-1)\cdot 9)\,x^2yz^4\\&=-1080\,x^2yz^4,\end{aligned}$$

故 x^2yz^4 项的系数是 -1080.

这个思想可以自然地推广到任意数量的项，得到更一般的**多项式定理**.

考虑展开式 $(x_1+x_2+\cdots+x_m)^n$. 其通项必然具有 $x_1^{k_1}x_2^{k_2}\cdots x_m^{k_m}$ 的形式，其中指数和为 n，即 $k_1+k_2+\cdots+k_m=n$.

遵循之前的组合论证，该项的系数等于将 n 个不同的因子分成 m 组，并指定第 i 组提供变量 x_i 的方案数，其中第 i 组的大小恰好为 k_i. 这个数称为**多项式系数**. 其组合意义是

将 n 个不同的物体分成 m 个有区别的组，使得第 i 组恰好有 k_i 个物体的总方法数.

与二项式系数的关系

二项式系数是多项式系数的一个特例. 从 n 个元素中选出 k 个，等价于将 n 个元素分成两组. 一组是被选中的 k 个元素，另一组是未被选中的 $n-k$ 个元素. 因此，

$$\binom{n}{k}=\binom{n}{k,n-k}=\frac{n!}{k!(n-k)!}$$

这个视角统一了这些计数工具，它们数的都是把 n 个元素分成若干组的方法数.

有了对这个系数组合意义的理解，我们正式陈述多项式定理. 该定理给出展开式的完整形式，同时给出计算多项式系数的代数公式.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 9.13 (多项式定理) 对于任意非负整数 n、整数 $m \geq 2$ 以及两两可交换的变量或数 $x_1, x_2, \ldots, x_m$, 有

$$(x_1 + x_2 + \cdots + x_m)^n = \sum_{k_1+k_2+\cdots+k_m=n} \binom{n}{k_1, k_2, \ldots, k_m} x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}$$

其中**多项式系数**定义为

$$\binom{n}{k_1, k_2, \ldots, k_m} = \frac{n!}{k_1! k_2! \cdots k_m!}$$

证明 我们用组合论证. 基本思路是分析代数展开的组合结构, 从而确定各项的系数.

首先, 将 $(x_1 + x_2 + \cdots + x_m)^n$ 写作 n 个相同因子的乘积

$$\underbrace{(x_1 + x_2 + \cdots + x_m)(x_1 + x_2 + \cdots + x_m) \cdots (x_1 + x_2 + \cdots + x_m)}_{n\text{个因子}}$$

根据乘法分配律, 展开后的每一项都是从这 n 个因子中的**每一个**里选择一个变量 $(x_1, x_2, \ldots, \text{或} x_m)$相乘得到的.

考虑展开式中一个**通项**的形式, 它必然是 $x_1^{k_1} x_2^{k_2} \cdots x_m^{k_m}$, 其中 $k_1, k_2, \ldots, k_m$ 是非负整数. 由于从 n 个因子中进行选择, 指数的总和必须等于 n, 即

$$k_1 + k_2 + \cdots + k_m = n$$

现在的任务是计算这个通项的系数. 这个系数等于能够产生该项的不同选择方案的总数, 即从 n 个因子中, 选出 k_1 个来提供 x_1, 选出 k_2 个来提供 x_2, …, 并选出 k_m 个来提供 x_m 的总方法数.

这实质上是将 n 个不同的因子(不妨将它们编号为 $1, 2, \ldots, n$)分成 m 个不同组的问题, 第 i 组的大小为 k_i. 用乘法原理, 按变量次序依次选取因子.

1. 从 n 个可用因子中为 x_1 选 k_1 个, 有 $\binom{n}{k_1}$ 种方法;
2. 再从剩下的 $n - k_1$ 个中为 x_2 选 k_2 个, 有 $\binom{n-k_1}{k_2}$ 种方法;
3. 再从剩下的 $n - k_1 - k_2$ 个中为 x_3 选 k_3 个, 有 $\binom{n-k_1-k_2}{k_3}$ 种方法;
4. 依此类推, 直至从最后剩下的 $n - k_1 - \cdots - k_{m-1} = k_m$ 个因子中为 x_m 选 k_m 个, 有 $\binom{k_m}{k_m} = 1$ 种方法.

由**乘法原理**, 总方案数是以上各步方法数的乘积.

$$\text{系数} = \binom{n}{k_1}\binom{n-k_1}{k_2}\binom{n-k_1-k_2}{k_3} \cdots \binom{k_m}{k_m}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

现在将这些二项式系数用阶乘形式展开, 观察化简过程.

$$\begin{aligned}\text{系数} &= \frac{n!}{k_1!(n-k_1)!}\cdot\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\cdot\frac{(n-k_1-k_2)!}{k_3!(n-k_1-k_2-k_3)!}\cdots\frac{k_m!}{k_m!0!}\\ &= \frac{n!}{k_1!\cancel{(n-k_1)!}}\cdot\frac{\cancel{(n-k_1)!}}{k_2!\cancel{(n-k_1-k_2)!}}\cdot\frac{\cancel{(n-k_1-k_2)!}}{k_3!\cdots}\cdots 1\\ &= \frac{n!}{k_1!k_2!k_3!\cdots k_m!}\end{aligned}$$

这是多项式系数 $\binom{n}{k_1,k_2,\ldots,k_m}$ 的定义.

已经证明, 对于任意满足 $k_1+\cdots+k_m=n$ 的非负整数组 $(k_1,\ldots,k_m)$, 其对应项 $x_1^{k_1}\cdots x_m^{k_m}$ 的系数恰好是多项式系数 $\binom{n}{k_1,\ldots,k_m}$.

把所有这些可能的项 (遍历所有满足条件的指数集) 加总, 便构成 $(x_1+\cdots+x_m)^n$ 的完整展开式. 定理得证. □

9.4.2 骰子与多项式 *(选读)

多项式乘法还能给计数问题提供另一套语言. 这一节用它解决一个关于骰子的经典问题, 答案由 Sicherman 给出.

例题 9.17 (点数重新标注的两枚骰子) 普通骰子的六个面分别标着 $1,2,3,4,5,6$. 能否把两枚骰子的点数重新标注成正整数, 允许同一枚骰子上出现相同的点数, 使得两枚骰子都不是普通的骰子, 而掷出的点数之和的各种情况仍与两枚普通骰子逐一相同?

解 把一枚骰子翻译成一个多项式. 若六个面的点数是 $a_1,a_2,\ldots,a_6$, 就记

$$A(x)=x^{a_1}+x^{a_2}+\cdots+x^{a_6}.$$

这个多项式的系数和 $A(1)$ 等于面数 6. 掷两枚骰子得到和为 s 的方案数, 恰好是乘积 $A(x)B(x)$ 中 x^s 的系数, 因为每一项 $x^{a_i}x^{b_j}=x^{a_i+b_j}$ 的指数是两个点数之和.

普通骰子对应

$$D(x)=x+x^2+x^3+x^4+x^5+x^6=x(1+x)(1+x+x^2)(1-x+x^2),$$

末一个等号把右端逐项乘开即可验证. 普通的两枚骰子对应 $D(x)^2$, 于是问题变成, 找出两个系数为 0 或 1、最低次项为 x、系数和都等于 6 的多项式 $A(x)$ 与 $B(x)$, 使

$$A(x)B(x)=D(x)^2.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这是一次多项式的分解. 由上面的分解式

$$D(x)^2 = x^2(1+x)^2\left(1+x+x^2\right)^2\left(1-x+x^2\right)^2,$$

把右端的八个因子分给 A 与 B 两方. 令 $x=1$ 看系数和, 得 $A(1)=B(1)=6$. 因子 $1+x$ 在 $x=1$ 处取 2, 而 6 只含一个因子 2, 所以两枚骰子各得一个 $1+x$. 因子 $1+x+x^2$ 在 $x=1$ 处取 3, 同理两枚骰子各得一个. 又因为两枚骰子的点数都是正整数, 各自的多项式都要被 x 整除, 于是 x^2 也是一方一个.

剩下的因子 $(1-x+x^2)^2$ 有两种分法. 一方一个, 得到 $A=B=x(1+x)(1+x+x^2)(1-x+x^2)=D$, 两枚都是普通骰子, 不合要求. 把两个都分给同一方, 则

$$A(x)=x(1+x)(1+x+x^2)=x+2x^2+2x^3+x^4,$$

$$B(x)=x(1+x)(1+x+x^2)(1-x+x^2)^2=x+x^3+x^4+x^5+x^6+x^8.$$

两个多项式各有六个系数为 1 的项(第二式中 x^2 与 x^7 的系数为零), 面数都是 6, 点数都是正整数. 对应的两枚骰子分别是

$$1,\,2,\,2,\,3,\,3,\,4 \quad 与 \quad 1,\,3,\,4,\,5,\,6,\,8.$$

由构造过程 $A(x)B(x)=D(x)^2$, 两枚骰子掷出的和从 2 到 12, 各种和的方案数依次是 $1,2,3,4,5,6,5,4,3,2,1$, 与两枚普通骰子完全一致.

这样的一对骰子称为 Sicherman 骰子. 全过程没有枚举任何一种掷法, 只做了多项式分解, 原因是乘法把点数的和与指数的和对应了起来. 这类充当计数载体的多项式称为**生成函数**, 第3.20节 处理牛顿和时用过同一个工具.

9.5 排列组合综合应用

排列与组合给出了计数的代数工具. 要把复杂的问题对应到这些基本模型上, 我们先对问题进行剖析与转化.

9.5.1 相邻与不相邻问题

在排列问题中, 元素之间的相对位置关系是最常出现的约束条件之一. 相邻与不相邻构成一对典型的对立约束, 分别对应**捆绑法**与**插空法**两种经典的处理技巧.

9.5.1.1 相邻问题与捆绑法

若若干特定元素必须排在一起, 即相邻, 我们就把它们在逻辑上视为一个不可分割的整体, 即一个复合元素, 再与其他元素一起排列.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

捆绑法

解决元素相邻的排列问题，通常遵循以下步骤.

1. **捆绑**. 将所有必须相邻的 m 个元素视为一个单一的、不可分割的大元素.
2. **外排**. 将这个大元素与其余的 $n-m$ 个元素进行全排列. 如果共有 k 个元素（包括大元素），则此步的方案数为 $k!$.
3. **内排**. 考虑被捆绑的 m 个元素内部自身的排列方式. 此步的方案数通常为 $m!$.
4. **相乘**. 根据乘法原理，将外排和内排的方案数相乘，得到最终结果.

例题 9.18　有 5 名男生和 3 名女生站成一排，如果要求 3 名女生必须站在一起，问有多少种不同的排法？

解　这是典型的相邻问题，我们用捆绑法处理. 令 3 名女生组成的整体为 G，则待排列的对象变为 5 名男生 $B_1,\dots,B_5$ 与 G，共 $5+1=6$ 个元素. 对这 6 个元素作全排列，由 $A(6,6)=6!=720$ 种方法.

在整体 G 内部，3 名女生自身尚可重排，有 $A(3,3)=3!=6$ 种方法. 外排与内排两步相互独立，由乘法原理得

$$N=6!\times 3!=720\times 6=4320,$$

即共有 4320 种不同的排法.

例题 9.19　用数字 0、1、2、3、4、5 组成没有重复数字的六位数，要求偶数数字必须相邻，奇数数字也必须相邻，问能组成多少个这样的六位数？

解　先按奇偶性分集合. 令偶数集合 $E=\{0,2,4\}$，奇数集合 $O=\{1,3,5\}$，各有 3 个元素. 题目要求偶数相邻且奇数也相邻，故需两次捆绑. 将 3 个偶数捆绑为整体 B_E，3 个奇数捆绑为整体 B_O. 外层只需排列 B_E,B_O 这两个大元素，全排列数为 $2!$. 但所成者为六位数，首位不能为 0，因此按哪个大元素居首分类讨论.

情况一.B_O 居首. 此时排列为 (B_O,B_E)，首位必为奇数（1, 3 或 5），不可能是 0，所有排列皆有效.B_O 内部全排列 $3!=6$ 种，B_E 内部全排列 $3!=6$ 种，此情况共 $3!\times 3!=36$ 种.

情况二.B_E 居首. 此时排列为 (B_E,B_O)，首位须取 B_E 中一数字且不能为 0，故 B_E 内部排列的首位只能选 2 或 4（2 种），余下 2 个偶数，含 0 排在后两位有 $2!$ 种，B_E 内部有效排列数为 $2\times 2!=4$ 种；B_O 内部仍有 $3!=6$ 种. 此情况共 $4\times 6=24$ 种.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

两情况互斥, 由加法原理得

$$N = 36 + 24 = 60,$$

知乎@余命数

即共能组成 60 个这样的六位数.

9.5.1.2 不相邻问题与插空法

若若干特定元素必须互不相邻, 我们就先排列那些没有位置限制的元素. 这些元素会自然形成一些空隙, 再将要求不相邻的元素插入这些空隙.

插空法

解决元素互不相邻的排列问题, 通常遵循以下步骤.

1. **排不受限元素**. 首先排列那些没有不相邻限制的元素. 假设有 n 个这样的元素, 它们的排列数为 $n!$.
2. **造空**. 这 n 个元素排好后, 将产生 $n+1$ 个可供插入的空隙(包括队伍的最前端和最后端).
3. **插入**. 将 m 个要求不相邻的元素, 插入到这 $n+1$ 个空隙中. 由于每个空隙最多只能放一个元素, 这相当于从 $n+1$ 个空隙中选出 m 个进行排列. 方案数为 $A(n+1, m)$.
4. **相乘**. 根据乘法原理, 将上述步骤的方案数相乘, 得到最终结果.

例题 9.20 将 3 本不同的小说和 4 本不同的诗集排成一排, 要求 3 本小说互不相邻, 问有多少种不同的排法?

解 这是典型的互不相邻问题, 我们用插空法处理. 小说是受限元素, 诗集是不受限元素. 先排列 4 本不同的诗集, 有 $A(4,4) = 4! = 24$ 种方法; 4 本诗集排好后形成 5 个可供插入的空隙

$$_P_1_P_2_P_3_P_4_,$$

知乎@余命数

将 3 本不同的小说插入这 5 个空隙, 每空至多一本, 等价于从 5 个空隙中选出 3 个并考虑小说顺序, 有 $A(5,3) = 5 \times 4 \times 3 = 60$ 种方法. 由乘法原理, 总的排法数为

$$N = 4! \times A(5,3) = 24 \times 60 = 1440,$$

即共有 1440 种不同的排法.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

不相邻与不全相邻的辨析

在处理计数问题时,我们必须先弄清约束条件的含义. 互不相邻与不全相邻是两个容易混淆的概念.

- **互不相邻**. 指任意两个指定的元素都不能相邻. 这是**插空法**的适用范围.
- **不全相邻**. 指指定的元素不全部连在一起即可, 允许其中一部分相邻.

对于不全相邻问题, 直接计数通常很困难, 因为它包含多种复杂情况. 更简洁的策略是使用**补集思想**(正难则反),即用总的排列数减去其对立面,也就是全部相邻的排列数.

例题 9.21　有 5 名男生和 3 名女生站成一排,回答下列问题.

1. 3 名女生互不相邻的排法有多少种?
2. 3 名女生不全相邻的排法有多少种?

解　本题意在澄清互不相邻与不全相邻的区别.

(1) 3 名女生互不相邻,适用插空法. 先排 5 名男生,有 $5!$ 种方法;5 名男生形成 6 个空隙;将 3 名女生插入这 6 个空隙,有 $A(6,3)$ 种方法. 由乘法原理得

$$N_1 = 5! \times A(6,3) = 120 \times (6 \times 5 \times 4) = 120 \times 120 = 14400.$$

(2) 3 名女生不全相邻, 宜用补集思想, 其对立面是 3 名女生全部相邻. 先算 8 人全排列总数 $8! = 40320$;对立事件即 3 名女生全部相邻,是前例用捆绑法算过的情形, 将 3 名女生捆绑与 5 名男生共 6 个元素全排列, 再乘以女生内部全排列, 得 $N_{\text{全相邻}} = 6! \times 3! = 720 \times 6 = 4320$. 由总数减去对立事件数

$$N_2 = 8! - (6! \times 3!) = 40320 - 4320 = 36000.$$

比较 $N_1 < N_2$, 不全相邻的情况包含了恰有两名女生相邻以及三名女生互不相邻两种情况, 而 N_1 只是后者.

9.5.2　捆绑法与插空法的原理

上一节引入了处理相邻与不相邻问题的两种流程, 即捆绑法与插空法. 本节分析它们为何有效,即把它们还原为乘法原理的构造性解释.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

9.5.2.1 捆绑法

相邻问题的约束是某些元素必须保持相对位置不变, 这破坏了全排列中每个元素独立性的前提. 捆绑法的做法, 是把一个带有复杂**约束**的计数问题, 通过逻辑上的等价变换, 转化为一个无约束的、分步的**构造**问题.

原理

现在证明一个事实. 一个满足 m 个元素相邻的排列, 可以由以下两个独立的步骤唯一地构造出来.

1. **内部构造**. 将被指定的 m 个元素进行全排列, 形成一个有序的、不可分割的复合元素.
2. **外部构造**. 将此复合元素与其余 $n-m$ 个普通元素进行全排列.

我们来分析这个构造过程的完备性与唯一性.

- **完备性**. 任何一个满足条件的最终排列, 都能通过这个两步构造法得到. 在任意一个满足条件的排列中, 那 m 个相邻的元素必然以某种顺序排列, 对应内部构造, 并且这个由它们组成的元素块必然在整个序列中占据一个位置 (对应外部构造). 因此, 不存在任何一种合法的排列会被构造过程遗漏.
- **唯一性**. 不存在两个不同的构造步骤序列会产生同一个最终排列. 改变内部构造, 例如, 将捆绑的 $\{A,B\}$ 顺序从 (A,B) 改为 (B,A), 最终排列中这部分元素的顺序也必然改变. 同样, 改变外部构造, 例如, 移动复合元素的位置, 最终排列也必然不同. 因此, 每一个独特的内部构造与外部构造组合都唯一对应一个最终的合法排列.

由于构造步骤与合法排列之间建立了一一对应, 合法排列的总数就精确等于构造步骤的总数. 根据乘法原理, 构造步骤的总数是每一步方法数的乘积.

- 内部构造的方法数是对 m 个元素的全排列, 即 $m!$.
- 外部构造的方法数是将 $n-m+1$ 个元素($n-m$ 个普通元素和 1 个复合元素)进行全排列, 即 $(n-m+1)!$.

因此, 总方法数为 $(n-m+1)!\times m!$. 这从根本上证明了捆绑法计算逻辑的正确性. 它把一个静态的满足约束的排列集合有多大的问题, 转变为一个动态的生成一个这样的排列有多少种方法的问题.

9.5.2.2 插空法

不相邻问题的约束是某些元素之间必须被其他元素隔开. 直接处理这种排斥性约束是困难的. 插空法通过逆向思维, 把一个复杂的**定位**问题转化为一个有序的**填充**问题.

插空法同样是一个构造性过程, 其正确性依赖于以下分步构造的唯一性与完备性.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1. **构建框架**. 首先排列那些没有位置限制的 n 个元素. 这一步的目的是创建一个稳定的骨架或框架.
2. **填充框架**. 将要求不相邻的 m 个元素, 放置到由上述框架产生的空隙中.

我们审视这个过程的数学基础.

- **空隙的定义**.n 个已经排好的元素 $U_1, U_2, \ldots, U_n$, 会在它们之间以及两端, 严格地定义出 $n+1$ 个不同的、可识别的空隙位置.

$$_U_1_U_2_\cdots_U_n_$$

这些空隙是放置不相邻元素的理想容器, 因为任何被放入不同空隙的两个元素, 都必然被至少一个框架元素分隔, 从而天然满足不相邻的约束.

- **填充即排列**. 将 m 个**不同**的受限元素放入 $n+1$ 个空隙中, 并且每个空隙至多放一个元素时, 这个过程本身是一个排列问题. 它等价于下面两步.
  1. 从 $n+1$ 个空隙中**选出** m 个空隙.
  2. 将 m 个不同的元素**排入**这 m 个选定的空隙中.

  它是排列 $A(n+1, m)$ 的组合意义.

与捆绑法类似, 这个两步构造过程(构建框架 $\to$ 填充框架)与所有满足互不相邻条件的最终排列之间也构成一一对应关系.

- **完备性**. 任何一个合法的排列, 其不受限元素必然构成一个有序子序列, 对应框架构建, 而受限元素必然分布在由这些不受限元素隔开的位置上 (对应填充框架).
- **唯一性**. 如果框架的排列不同, 或者填充空隙的方案不同, 最终得到的总排列也必然不同.

因此, 满足条件的总排列数等于构建框架的方法数与填充框架的方法数的乘积.

- 构建框架的方法数是对 n 个不受限元素的全排列, 即 $n!$.
- 填充框架的方法数是从 $n+1$ 个空隙中选 m 个进行排列, 即 $A(n+1, m)$.

最终结果为 $n! \times A(n+1, m)$. 这就证明了插空法是乘法原理的一个构造性应用.

例题 9.22 (2020.4 学军中学高二下期中) 一条街道上有 10 盏路灯, 将路灯依次排列并编号 1-10. 有关部门要求晚上这 10 盏路灯中相邻的两盏灯不能全开, 且这 10 盏路灯中至少打开两盏路灯. 则符合要求的开法总数是 _____.

解 此问题的约束是开着的灯互不相邻; 又因只关心哪些位置的灯是开的, 灯本身无差别, 所以这是一个组合问题. 直接挑选 k 个不相邻的位置并不方便, 我们反过来思考. 利用关着的灯为开着的灯制造空隙, 这就是插空思想.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

按打开的灯数 $k(k\geq 2)$ 分类. 设打开 k 盏灯, 则有 $10-k$ 盏灯关闭. 先将这 $10-k$ 盏关着的灯视作相同元素排成一排, 形成 $(10-k)+1=11-k$ 个空隙, 含两端; 为保证 k 盏灯彼此不相邻, 将它们放入这 $11-k$ 个空隙, 每空至多一盏. 因灯无差别, 此即从 $11-k$ 个空隙中选 k 个, 方法数为 $\binom{11-k}{k}$. 由 $k\geq 2$ 及 $11-k\geq k$, 即 $2k\leq 11$ 得 $k\in\{2,3,4,5\}$. 各情况相加

$$\begin{aligned}N&=\binom{9}{2}+\binom{8}{3}+\binom{7}{4}+\binom{6}{5}\\&=\tfrac{9\cdot 8}{2}+\tfrac{8\cdot 7\cdot 6}{3!}+\binom{7}{3}+6\\&=36+56+35+6=133,\end{aligned}$$

故符合要求的开法总数是 133.

同一结果亦可由变换法导出. 问题即从 10 个位置中选出若干个互不相邻的位置, 是纯组合问题. 设打开 k 盏灯, 等价于从 $\{1,\dots,10\}$ 中取 k 个数且任意两数不相邻. 令选出的数为 $1\leq x_1<x_2<\cdots<x_k\leq 10$, 约束可写为 $x_{i+1}-x_i\geq 2$.

为处理此不等式, 构造双射 $y_i=x_i-(i-1)$. 则 $y_1=x_1\geq 1, y_k=x_k-(k-1)\leq 10-(k-1)=11-k$, 且

$$y_{i+1}-y_i=(x_{i+1}-i)-(x_i-(i-1))=(x_{i+1}-x_i)-1\geq 1,$$

即 $y_1<y_2<\cdots<y_k$. 于是从 $\{1,\dots,10\}$ 中选 k 个互不相邻的数 x_i, 等价于从 $\{1,\dots,11-k\}$ 中选 k 个不同的数 y_i, 方法数同为 $\binom{11-k}{k}$.k 的取值范围仍由 $k\geq 2$ 与 $11-k\geq k$ 决定, 得 $k\in\{2,3,4,5\}$, 相加得

$$N=\binom{9}{2}+\binom{8}{3}+\binom{7}{4}+\binom{6}{5}=36+56+35+6=133.$$

例题 9.23 (2014 重庆理) 某次联欢会要安排 3 个歌舞类节目、2 个小品类节目和 1 个相声类节目的演出顺序, 则同类节目不相邻的排法种数是(　).

A. 72　　B. 120　　C. 144　　D. 168

解 本题涉及对多个元素集合的不相邻约束, 正面处理很复杂, 使用**容斥原理**(正难则反)正合适.

首先假定所有 6 个节目都是独一无二的. 总的排列数(没有任何限制)为

$$N_{总}=6!=720$$

需要排除所有不符合同类节目不相邻要求的排列. 同类节目不相邻的对立面是至少有一类同类节目相邻. 具体地, 就是 2 个小品相邻或者 3 个歌舞中有节目相邻.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设 U 为所有排列的全集,$|U|=720$. 令 A 为 2 个小品节目相邻的排列构成的集合. 令 B 为至少有 2 个歌舞节目相邻的排列构成的集合. 要求解的是 $|U|-|A\cup B|$. 根据容斥原理, $|A\cup B|=|A|+|B|-|A\cap B|$.

将 2 个小品节目捆绑成一个复合元素. 现在排列 {歌舞 1, 歌舞 2, 歌舞 3, 相声, (小品块)} 这 5 个对象, 有 5! 种方法. 小品块内部有 2! 种排列.

$$|A|=5!\times 2!=120\times 2=240$$

知乎@余命数

直接计算至少有 2 个歌舞相邻比较复杂, 计算其补集, 即 3 个歌舞节目互不相邻(记作 B^c). 为此使用插空法. 先排列非歌舞节目 {小品 1, 小品 2, 相声}, 有 3! 种方法. 这 3 个节目形成了 4 个空隙. 将 3 个不同的歌舞节目插入这 4 个空隙中, 有 $A(4,3)$ 种方法.

知乎@余命数

$$|B^c|=3!\times A(4,3)=6\times 24=144$$

因此, $|B|=|U|-|B^c|=720-144=576$.

最后计算 $|A\cap B|$, 这个集合表示小品相邻 **且**歌舞至少有 2 个相邻. 同样可利用补集思想计算这个交集的大小, 即 $|A\cap B|=|A|-|A\cap B^c|$. $|A\cap B^c|$ 表示小品相邻 **且**歌舞互不相邻. 再次使用插空法. 将被捆绑的小品块和相声节目作为框架. 排列 {(小品块), 相声} 有 2! 种方法. 小品块内部有 2! 种排列. 故框架的排列数为 $2!\times 2!=4$. 这个框架形成了 3 个空隙. 将 3 个不同的歌舞节目插入这 3 个空隙中, 有 $A(3,3)=3!$ 种方法.

$$|A\cap B^c|=(2!\times 2!)\times 3!=4\times 6=24$$

余命数

所以, $|A\cap B|=|A|-|A\cap B^c|=240-24=216$.

不符合要求的排列总数为

知乎@余命数

$$|A\cup B|=|A|+|B|-|A\cap B|=240+576-216=600$$

知乎@余命数

知乎@余命数

因此, 符合要求的排列种数为

知乎@余命数

$$N=|U|-|A\cup B|=720-600=120$$

故选 B.

9.5.3 定序问题与倍缩策略

在排列问题中, 除了相邻与不相邻的约束外, 另一类常见的限制是**定序问题**, 即要求某几个元素必须保持特定的相对顺序. 例如, 在排列 A、B、C、D 时, 要求 A 必须始终在 B 的左边.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

面对这类问题, 一种有效的思想是**倍缩策略**, 做法是先放后缩. 先忽略定序的约束, 对所有相关元素进行全排列, 得到一个总数. 然后分析在这个总数中, 所关注的、被定了序的元素有多少种排列方式. 由于题目只允许其中的一种特定顺序, 原来的计数夸大或重复了, 其倍数恰好是被定序元素的全排列数. 因此, 用总排列数除以这个重复倍数, 即可得到最终答案.

这个思想来自组合数公式, 而组合数公式本身就是倍缩策略最经典的体现.

$$\binom{n}{k}=\frac{A(n,k)}{k!}$$

此公式的组合意义是, 要从 n 个元素中选出 k 个而不计顺序, 可以先选出 k 个并排列它们, 共 $A(n,k)$ 种方法. 但我们不关心这 k 个元素的顺序, 而它们自身有 $k!$ 种排列, 因此将 $A(n,k)$ 除以 $k!$ 以消除顺序的影响.

倍缩策略是用总排列数除以被定序元素的全排列数. 这种除法是乘法原理的逆向应用, 背后是集合划分与一一对应的思想. 我们来说明它为何有效.

设集合 S 含 n 个元素, 我们要计算 S 的所有排列中, 满足特定 m 个元素 $M=\{e_1,e_2,\dots,e_m\}$ 保持固定相对顺序, 例如 e_1 在 e_2 之前, e_2 在 e_3 之前, 以此类推的排列总数. 令这个符合条件的排列集合为 $\mathcal{A}$, 其大小为 $N=|\mathcal{A}|$.

直接计算 N 较困难, 因此引入一个更大的集合作为参照. 令 $\mathcal{U}$ 为 S 中所有 n 个元素的全排列集合, 则 $|\mathcal{U}|=n!$.

关键在于建立集合 $\mathcal{U}$ 与 $\mathcal{A}$ 之间的关系. 我们用构造性论证与乘法原理表达 $|\mathcal{U}|$.

考虑如何构造出 $\mathcal{U}$ 中的任意一个排列. 每一个排列都可由一个两步过程唯一地生成.

1. **选择一个骨架排列.**
   从符合定序条件的集合 $\mathcal{A}$ 中任意选择一个排列. 在这个排列中, 那 m 个特殊元素的相对顺序是正确的. 这一步有 N 种选择.
2. **对特殊元素进行内部重排.**
   在上一步选出的骨架排列中, 那 m 个特殊元素占据了 m 个确定的位置. 保持其他 $n-m$ 个元素不动, 只将这 m 个特殊元素在这 m 个位置上进行任意的全排列. 这一步有 $m!$ 种方法.

根据**乘法原理**, 通过这两个步骤, 总共可以构造出 $N\times m!$ 个不同的排列.

现在证明这个构造过程与集合 $\mathcal{U}$ 等价, 即它不重不漏地生成了 $\mathcal{U}$ 中的所有 $n!$ 个排列.

- **无遗漏:** $\mathcal{U}$ 中的任何一个排列都能被上述方法构造出来. 任取一个 $\mathcal{U}$ 中的排列 p, 其中 m 个特殊元素的位置和内部顺序是任意的. 可以通过一个标准化操作, 将这 m 个元素按照题目要求的固定顺序重新排列, 从而得到一个唯一对应的骨架排列(它必然属于集合 $\mathcal{A}$). 这证明了任何排列 p 都有其构造来源.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **无重复：**两个不同的构造步骤组合，不会产生同一个最终排列. 假设通过 $(\text{骨架}_1, \text{重排}_1)$ 和 $(\text{骨架}_2, \text{重排}_2)$ 得到了同一个最终排列 p. 一个排列中，所有非特殊元素的位置，以及所有特殊元素所占据的位置集合，都是唯一的. 因此，骨架_1 和 骨架_2 必须具有相同的结构，即相同的非特殊元素排列和相同的特殊元素占位. 又因为在骨架排列中特殊元素的相对顺序固定，所以 $\text{骨架}_1 = \text{骨架}_2$. 既然骨架相同，且最终结果 p 相同，那么施加在骨架上的内部重排操作也必须相同，即 $\text{重排}_1 = \text{重排}_2$. 因此每一种构造组合都唯一对应一个最终排列.

上述论证建立了从构造步骤对到全排列集合 $\mathcal{U}$ 的一一对应关系. 因此它们的大小必然相等.

$$|\mathcal{U}| = N \times m!$$

已知 $|\mathcal{U}| = n!$，所以

$$n! = N \times m!$$

由此解出最初的目标 N.

$$N = \frac{n!}{m!}$$

这个推导把倍缩策略归结为对乘法原理方程 $n! = N \times m!$ 的求解.

从另一个角度看，这个过程等价于将 $n!$ 个全排列进行**分组**或**划分**. 我们将所有排列按照除 m 个特殊元素的内部顺序外、其他一切相同的规则分组. 例如，排列 (e_2, x, e_1, y) 和 (e_1, x, e_2, y) 就属于同一组，因为它们的骨架（非 e 元素的位置和 e 元素所占的位置）是相同的. 每一组内的排列个数，恰好等于这 m 个特殊元素自身的全排列数，即 $m!$ 个. 而题目要求的定序排列，恰好是每一组中那**唯一一个**满足特定顺序的排列. 因此，符合条件的排列总数 N 等于这个分组的数量. 根据除法原理，分组的数量为

$$N = \frac{\text{总元素数}}{\text{每组的元素数}} = \frac{n!}{m!}$$

这两种视角都证明了倍缩策略的正确性. 它把一个带有复杂相对顺序约束的问题，转化为关于等价类的计数问题.

倍缩策略

若在一组共 n 个元素的排列中，有 m 个特定元素被要求保持固定的相对顺序，则符合条件的排列数可按以下步骤计算.

1. 计算这 n 个元素的全排列数，即 $n!$.
2. 这 m 个被定序的元素在上述全排列中，自身有 $m!$ 种不同的排列方式.
3. 由于约束只允许这 $m!$ 种排列中的一种，将总数除以 $m!$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

最终的排列数为 $\dfrac{n!}{m!}$. 如果有多组元素被定序, 则需除以各组元素内部全排列数的乘积.

例题 9.24 (2024 年 1 月浙江省宁波九校联考) 体育课上, 罗老师让 8 名身高各不相同的同学排队, 要求排成前后两排, 每排 4 人, 且每排同学从左到右身高依次递增, 则不同排法的种数为(　).

A. 60　　B. 70　　C. 80　　D. 90

解 此问题的突破口在于, 一旦确定了哪些同学在前排、哪些同学在后排, 那么由于身高依次递增的严格约束, 每一排的站位顺序就**唯一确定**了. 因此, 问题从一个排列问题, 退化为一个分组或组合问题.

任务可以简化为把 8 名同学分成两组, 一组是前排组, 另一组是后排组, 每组 4 人.

确定前排组的人选即可完成这个任务. 从 8 名同学中选出 4 名到前排, 这是一个组合问题, 其方法数为

$$\binom{8}{4}=\frac{8\times7\times6\times5}{4\times3\times2\times1}=70$$

一旦这 4 名同学被选定, 他们只能以身高从低到高的一种方式站立.

剩下的 $8-4=4$ 名同学自动成为后排组, 同样, 他们的站立方式也只有一种.

由乘法原理, 总的排法数就是分组的方法数, 即 $70\times1\times1=70$.

也可以从**倍缩策略**的角度理解这个问题, 这种视角更能说明重复计数来自哪里.

首先不考虑任何身高顺序的限制, 将 8 名同学安排到 8 个具体的位置上(4 个前排位置, 4 个后排位置). 这相当于 8 名同学的全排列, 总方法数为 $8!$.

现在引入约束条件. 对于最终被安排在前排的 4 名同学, 无论他们是谁, 在 $8!$ 种总排列中, 他们在这 4 个前排位置上可以有 $4!$ 种不同的排列方式. 然而, 题目要求他们必须按身高递增排列, 这只是 $4!$ 种排列中的一**种**. 因此, 对前排的排列多算了 $4!$ 倍.

同理, 对于最终被安排在后排的 4 名同学, 他们的顺序也被唯一确定, 所以对后排的排列也多算了 $4!$ 倍.

由倍缩策略, 将总排列数除以这两个重复的倍数, 得到符合条件的排法数.

$$N=\frac{8!}{4!\times4!}=\frac{40320}{24\times24}=\frac{40320}{576}=70$$

这与从组合角度得到的结果一致. 故选 B.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.25 (2008 安徽理) 12 名同学合影, 站成前排 4 人后排 8 人, 现摄影师要从后排 8 人中抽 2 人调整到前排, 若其他人的相对顺序不变, 则不同调整方式的种数是(　).

A. $C_8^2A_3^2$　　C. $C_8^2A_6^2$

B. $C_8^2A_6^6$　　D. $C_8^2A_5^2$

解 此问题分两步. 先**选人**, 从后排 8 名同学中选 2 人调到前排, 只关心选哪两人、不涉及顺序, 是组合问题, 方法数为

$$\binom{8}{2}$$

再**排位**, 此时前排 $4+2=6$ 名同学, 后排剩 $8-2=6$ 名.

问题的约束是其他人的相对顺序不变, 即

- 原前排 4 人的相对顺序保持不变.
- 原后排剩下的 6 人的相对顺序也保持不变.

知乎@余命数

由于后排 6 人的位置和相对顺序都已固定, 我们只需考虑前排这 6 人的新排法.

在前排的 6 个位置上, 需要安排 4 名旧同学和 2 名新同学. 其中, 4 名旧同学的相对顺序是固定的. 这是一个典型的**定序问题**.

我们用插空法解决排位问题. 把前排的 6 个位置看作 6 个空位, 为新选出的 2 名同学选择位置. 为第一位新同学选择位置, 有 6 种选择. 为第二位新同学选择位置, 有 5 种选择. 这两位新同学是不同的个体, 顺序是重要的. 因此为这两位新同学安排位置的方法数是一个排列问题, 共有

$$A(6,2)=6\times5=30$$

一旦这两位新同学的位置确定, 剩下的 4 个位置将由 4 名旧同学按原有相对顺序填入, 方法只有 1 种.

或者, 也可以运用**倍缩策略**思考排位问题. 如果不考虑任何顺序限制, 将这 6 名同学(4 旧 2 新)排在前排 6 个位置上, 共有 6! 种方法. 但题目要求 4 名旧同学的相对顺序不变. 在这 6! 种排列中, 这 4 名旧同学自身有 4! 种排列方式, 而只接受其中的 1 种. 因此, 符合条件的排法数为

$$\frac{6!}{4!}=\frac{720}{24}=30=A(6,2)$$

知乎@余命数
知乎@余命数

最后由乘法原理, 将选人和排位两个步骤的方法数相乘, 得到总的调整方式种数.

$$N = \binom{8}{2} \times A(6,2)$$

知乎@余命数

在题目选项的记法中, 即为 $C_8^2 A_6^2$. 故选 C.

知乎@余命数

例题 9.26 (重复元素的排列) 计算由单词 MISSISSIPPI 中所有字母构成的不同排列(字符串)的数量.

解 这个问题的新复杂性在于元素重复. 单词 MISSISSIPPI 共含 11 个字母, 其中 1 个 M, 4 个 I, 4 个 S, 2 个 P.

直接应用排列公式行不通, 因为它要求所有元素可区分. 但借助倍缩策略与一个思想实验, 此问题即可解决.

首先假想所有字母都可区分. 给相同字母加上下标来区分它们.

$$M_1, I_1, S_1, S_2, I_2, S_3, S_4, I_3, P_1, P_2, I_4$$

现在我们有 11 个完全不同的对象, 将它们全排列, 总方法数为 11!.

知乎@余命数 知乎@余命数

接下来去除下标, 回到原问题. 考虑任意一个具体的排列, 例如 $S_1 I_1 P_1 M_1 I_2 S_2 S_3 I_3 P_2 S_4 I_4$. 现在关注其中的 4 个 I (I_1, I_2, I_3, I_4). 在这个排列的特定位置上, 这 4 个 I 自身可以有 4! 种排列方式, 例如 I_2 和 I_1 交换位置. 在假想的可区分世界里, 这些都是不同的排列. 但在现实世界中, 所有 I 都相同, 因此这 4! 种排列都对应同一个字符串 SIPMISSIPSI.

最初计算的 11! 对每一个最终字符串, 都因 I 的存在而重复计数了 4! 倍. 因此必须用总数除以 4! 来消除由 I 引起的重复.

同理, 对于 4 个 S, 也重复计数了 4! 倍. 对于 2 个 P, 重复计数了 2! 倍.

由倍缩策略, 将总的假想排列数除以各重复元素内部排列所产生的重复倍数.

知乎@余命数

$$N = \frac{11!}{4! \times 4! \times 2! \times 1!} = \frac{39916800}{(24) \times (24) \times (2) \times (1)} = \frac{39916800}{1152} = 34650$$

因此共有 34650 个不同的排列.

此例导出一个普遍结论. 对于包含 n 个对象的多重集, 其中有 k_1 个第一类相同对象,k_2 个第二类相同对象,…,k_m 个第 m 类相同对象,其中 $k_1+k_2+\cdots+k_m = n$, 其不同的排列数为

$$\binom{n}{k_1, k_2, \ldots, k_m} = \frac{n!}{k_1! k_2! \cdots k_m!}$$

这是我们之前遇到的**多项式系数**. 倍缩策略为其提供了一个组合解释.

@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.27 (圆排列) 将 n 位不同的宾客安排在一个圆形的餐桌旁就座, 有多少种不同的就座方式?(如果两种就座方式中, 每位宾客的左邻和右邻都完全相同, 则认为这两种方式是相同的.)

解 这是典型的圆排列问题. 圆桌没有起点或终点的概念, 判断标准从绝对位置变为相对位置.

我们再次用倍缩策略, 从一个更熟悉的线性排列开始.

首先暂时忽略圆桌的特性, 设想将这 n 位宾客排成一条直线. 那么总的线性排列数为 $n!$.

现在将任意一个线性排列, 例如 $(P_1, P_2, \dots, P_n)$, 弯曲成一个圆圈, 使其首尾相连. 旋转这个线性排列, 依次得到下列各排列.

$$\begin{aligned}(P_1, P_2, P_3, \dots, P_n)\\ (P_n, P_1, P_2, \dots, P_{n-1})\\ (P_{n-1}, P_n, P_1, \dots, P_{n-2})\\ \vdots\\ (P_2, P_3, P_4, \dots, P_1)\end{aligned}$$

知乎@余命数

通过 n 次旋转, 得到 n 个不同的线性排列. 然而, 当将它们都布置到圆桌上时, 它们描述的是**同一种**就座方式, 因为在每一种情况中, P_1 的左边总是 P_n、右边总是 P_2; P_2 的左边总是 P_1、右边总是 P_3, 以此类推. 相对位置关系完全一样.

这说明了一个关键事实. 最初计算的 $n!$ 种线性排列, 可以将它们 n 个一组进行划分. 每一组中的 n 个线性排列都对应着同一个圆排列. 因此初始计数 $n!$ 将每一种真正的圆排列都重复计算了 n 倍.

由倍缩策略, 将线性排列总数除以这个重复的倍数 n, 即得到圆排列的数目.

知乎@余命数

$$N = \frac{n!}{n} = (n-1)!$$

因此有 $(n-1)!$ 种不同的就座方式.

圆排列的一个自然延伸是**项链问题**. 将 n 颗不同的珠子串成一条项链, 求不同的串法数. 项链既能旋转, 也能从空间中拿起来**翻转**. 对于 $n > 2$ 且珠子两两不同的情况, 每一次翻转都会将一个圆排列变成另一个镜像圆排列, 但二者对应同一条项链. 因此之前的计数 $(n-1)!$ 将每条项链重复计算了 2 倍. 项链的不同串法数为 $\dfrac{(n-1)!}{2}$.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.28 (棋盘问题 *) 在一个 $m \times n$ 的棋盘格上, 一个棋子从左下角 $(0,0)$ 出发, 要到达右上角 (m,n). 如果每次只能向上或向右移动一个单位, 总共有多少条不同的最短路径?

解 这个经典的格路计数问题解法众多, 倍缩策略为它提供了一个新视角.

首先分析路径的构成. 任何一条从 $(0,0)$ 到 (m,n) 的最短路径, 都必须包含 m 次向右移动与 n 次向上移动, 分别记作 R 与 U. 因此, 每一条不同的路径都唯一对应一个由 m 个 R 和 n 个 U 构成的序列, 该序列总长度为 $m+n$.

例如, 在 2×1 的棋盘上, 从 $(0,0)$ 到 $(2,1)$ 的路径如下.

- 路径 $(0,0) \to (1,0) \to (2,0) \to (2,1)$ 对应序列 RRU.
- 路径 $(0,0) \to (1,0) \to (1,1) \to (2,1)$ 对应序列 RUR.
- 路径 $(0,0) \to (0,1) \to (1,1) \to (2,1)$ 对应序列 URR.

知乎@余命数

于是原问题转化为计算由 m 个 R 和 n 个 U 构成的不同序列的数量.

这是一个带有重复元素的排列问题, 是倍缩策略的用武之地.

设想 $m+n$ 次移动, 它们分别是 $R_1, R_2, \dots, R_m$ 和 $U_1, U_2, \dots, U_n$. 将这 $m+n$ 个不同的操作进行全排列, 总方法数为 $(m+n)!$.

在现实问题中, 所有的向右移动 R 没有区别, 所有的向上移动 U 也没有区别. 在假想的 $(m+n)!$ 种排列中, 那 m 个不同的向右操作 $R_1, \dots, R_m$ 自身可以有 $m!$ 种排列方式. 但这些都对应同一个由 m 个 R 组成的移动模式. 因此计数因 R 的存在而重复了 $m!$ 倍.

同理, 计数也因 U 的存在而重复了 $n!$ 倍.

将总的假想排列数除以由重复元素产生的倍数, 得到最终的路径数.

$$N = \frac{(m+n)!}{m!n!}$$

知乎@余命数

这个结果是组合数 $\binom{m+n}{m}$, 也等于 $\binom{m+n}{n}$. 这为组合数提供了一个新的组合解释.

$$\binom{m+n}{m} = \text{从 (0,0) 到 (m,n) 的格路数}$$

知乎@余命数

也可以直接用组合的定义解决. 一条路径总共需要 $m+n$ 步. 只需在这 $m+n$ 步中决定哪几步向右移动 (R), 剩下的自动是向上移动 (U). 从 $m+n$ 个总步数中选出 m 步来执行向右移动, 方法数为

$$\binom{m+n}{m}$$

这与倍缩策略得到的结果一致. 这种将排列问题转化为序列编码、再用倍缩策略求解的方法, 是组合数学中的常用技巧.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.29 (2024 年 1 月湖南省邵阳市高三第一次联考) 城步苗族自治县六月六山歌节是湖南省四大节庆品牌之一, 至今已举办 25 届. 假设在即将举办的第 26 届六月六山歌节中, 组委会要在原定排好的 10 个本土歌舞节目中增加 2 个歌王对唱节目. 若保持原来 10 个节目的相对顺序不变, 则不同的排法种数为(　).

A. 110　　B. 144　　C. 132　　D. 156

解　问题的约束是保持原来 10 个节目的相对顺序不变. 记原来的 10 个节目为 $O_1, O_2, \ldots, O_{10}$, 则在最终的节目单中 O_1 必须出现在 O_2 之前, O_2 必须出现在 O_3 之前, 以此类推. 这是一个典型定序问题, 适合用倍缩策略.

增加 2 个新的歌王对唱节目后(视这两个新节目为不同个体 N_1, N_2), 总节目数量为 $10+2=12$ 个.

方法一采用倍缩策略.

首先忽略相对顺序不变的约束, 考虑将这 12 个不同节目全排列. 总的排列方法数为

$$12!$$

现在引入约束条件. 在上述 12! 种排列中, 那 10 个原来的节目自身可以有 10! 种不同的内部排列方式. 然而, 题目要求只接受它们按原有顺序排列的**那一种**.

对于任何一种满足条件的最终排法, 最初的 12! 计数都将其重复计算了 10! 次(因为这 10! 次内部重排都被算作不同情况). 因此必须用总排列数除以这个重复的倍数.

符合条件的排法种数为

$$N=\frac{12!}{10!}=\frac{12\times 11\times (10!)}{10!}=12\times 11=132$$

方法二采用选位法(插空思想的变体).

也可从构造的角度思考. 总共有 12 个节目位置需要安排, 问题可转化为先为 2 个新节目选择位置, 剩下的位置自然由 10 个旧节目按其固有顺序填充. 从 12 个时间槽中为 2 个不同的新节目选定位置, 等价于从 12 个位置中选出 2 个进行排列, 方法数为 $A(12,2)=12\times 11=132$. 新节目位置一旦确定, 只剩 $12-2=10$ 个空位, 而 10 个旧节目的相对顺序固定, 填充这 10 个空位的方式**唯一确定**, 方法数为 1. 由乘法原理, 总的排法种数为 $A(12,2)\times 1=132$.

两种方法得到相同的结果. 故选 C.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.30 (待插入元素相同) 沿用上题背景, 假设增加的 2 个歌王对唱节目是**完全相同**的, 例如, 都是同一首歌的两个不同时段重播, 若仍要保持原来 10 个节目的相对顺序不变, 则不同的排法种数为(　).

A. 55　　B. 66　　C. 78　　D. 132

解 这个问题与原题的区别在于新增的 2 个节目从可区分变为不可区分, 这直接影响计数方式.

方法一采用选位法(组合思想).

总共有 12 个节目位置. 任务是为 2 个**相同**的新节目选择位置. 由于新节目相同, 选择第 3 位和第 8 位与选择第 8 位和第 3 位最终产生的节目单相同. 因此这是一个纯粹的组合问题.

从 12 个可用位置中选出 2 个位置安放这两个相同的新节目. 方法数为

$$\binom{12}{2}=\frac{12\times 11}{2\times 1}=66$$

一旦新节目的位置被确定, 剩下的 10 个位置将由 10 个旧节目按其固有顺序唯一填充. 由乘法原理, 总排法种数为 $\binom{12}{2}\times 1=66$.

方法二采用倍缩策略的叠加.

从 12 个节目全排列出发, 总数为 12!.

首先应用第一个约束, 即 10 个旧节目的相对顺序不变. 因此将总数除以 10! 来消除这 10 个节目内部顺序的重复计数.

$$N_1=\frac{12!}{10!}$$

接着应用第二个约束.2 个新节目是完全相同的. 即在上述结果中, 仍然将这两个新节目视为可区分的, 从而对它们自身的 2! 种排列进行了重复计数. 因此, 必须**再次**应用倍缩策略, 除以 2!.

最终的排法种数为

$$N=\frac{12!}{10!\times 2!}=\frac{A(12,2)}{2!}=\frac{132}{2}=66$$

这个结果是组合数 $\binom{12}{2}$ 的定义. 它把组合数 $\binom{n}{k}$ 解释为在 n 个对象的排列中同时对选出的 k 个元素和未选的 $n-k$ 个元素施加定序, 或视其为相同约束的结果. 故选 B.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.31 (增加不相邻约束) 沿用原题背景, 即增加的 2 个歌王对唱节目是**不同**的. 若既要保持原来 10 个节目的相对顺序不变, 又要求这两个新增的歌王对唱节目**不能相邻**, 则不同的排法种数为(　).

A. 110　　B. 120　　C. 121　　D. 132

知乎@余命数

解 这个问题在原题的基础上增加了一个不相邻约束. 这种定序与不相邻的复合约束, 是插空法与倍缩思想结合的典型场景.

方法一采用插空法.

处理不相邻问题最直接的方法是用插空法. 在这里, 10 个旧节目是骨架, 2 个新节目是等待插入的元素.

先排列骨架. 先安排 10 个旧节目. 由于它们的相对顺序是固定的, 排列它们的方法只有一**种**.

$$O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{10}$$

再创造空隙. 这 10 个排好的节目形成了 11 个可供插入的空隙(包括最前和最后).

$$_O_1_O_2_\cdots_O_{10}_$$

接着插入元素. 将 2 个**不同**的新节目插入到这 11 个空隙中, 每个空隙至多放一个节目, 以保证它们不相邻. 这相当于从 11 个空隙中选出 2 个, 并将 2 个不同的节目排列进去. 这是一个排列问题. 方法数为

$$A(11,2) = 11 \times 10 = 110$$

由乘法原理, 总的排法种数为 $1 \times A(11,2) = 110$.

方法二采用补集思想(正难则反).

也可以从原问题的总数中减去不满足不相邻约束的情况, 即减去 2 个新节目必须相邻的情况.

原问题(只要求旧节目定序)的总排法数, 前面已计算过是 132.

现在计算旧节目定序且新节目相邻的排法数. 使用**捆绑法**. 将 2 个不同的新节目捆绑成一个大节目块. 这个大节目块内部, 2 个新节目可以有 $2! = 2$ 种排列方式.

现在的问题转化为将这个大节目块和 10 个旧节目进行排列, 同时保持 10 个旧节目的相对顺序不变. 这等价于将 1 个大节目块插入到 10 个旧节目形成的 11 个空隙(包括首尾)中. 选择插入位置的方法数有 $\binom{11}{1} = 11$ 种.

因此新节目相邻的排法数为

$$N_{\text{相邻}} = (\text{内部排列}) \times (\text{外部插入}) = 2! \times \binom{11}{1} = 2 \times 11 = 22$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

最后用总数减去相邻数.

$$N_{\text{不相邻}} = N_{\text{总}} - N_{\text{相邻}} = 132 - 22 = 110$$

两种方法再次得到一致的结果. 故选 A.

9.5.4 分组问题

在组合计数中, 经常需要将一个包含 n 个不同元素的集合, 划分成若干个互不相交的子集, 即组. 解决这类问题的要点在于精确辨析一个要素, 即这些最终形成的组本身是**可区分的**还是**不可区分的**.

这两种模型的根本差异可以从一个场景看出. 设有 6 本完全不同的书, 需要将它们分成三份, 每份 2 本.

- **情景一.** 将这些书分给甲、乙、丙三个人, 每人 2 本.
- **情景二.** 将这些书打包成三份, 每份 2 本, 堆在桌子上.

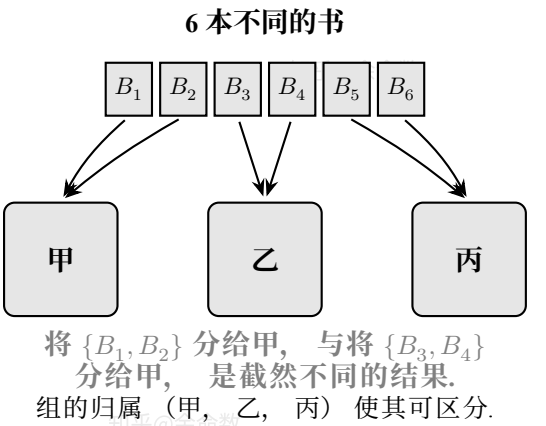


图 9.4 组可区分的有序分组

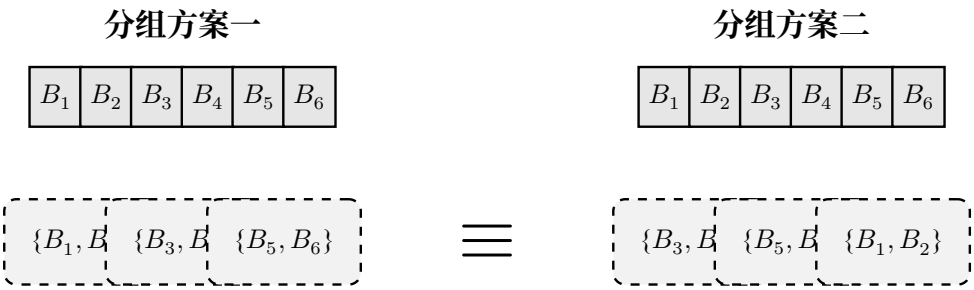
9.5.4.1 有序分组

在情景一中, 最终分到书的三个人甲、乙、丙是明确可区分的. 甲拿到 $\{B_1, B_2\}$ 和乙拿到 $\{B_1, B_2\}$ 是两种完全不同的分配结果. 因此, 这些组因为其归属(甲、乙、丙)而被赋予了身份, 是可区分的. 这称为**有序分组**或**定序分组**.

计算这类问题的过程是一个连续的组合决策.

- 首先, 从 6 本书中为甲选出 2 本, 方法数为 $\binom{6}{2}$.
- 接着, 从剩下的 4 本书中为乙选出 2 本, 方法数为 $\binom{4}{2}$.
- 最后, 剩下的 2 本书自动归丙所有, 方法数为 $\binom{2}{2}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



方案一和方案二只是三堆书的摆放顺序不同，实质上是同一种分组结果.
这些堆没有身份标签，是不可区分的.

图 9.5　组不可区分的无序分组

由乘法原理, 总的分配方法数为

$$N_1 = \binom{6}{2}\binom{4}{2}\binom{2}{2} = \frac{6\times 5}{2}\times\frac{4\times 3}{2}\times 1 = 15\times 6\times 1 = 90$$

将这个结果用阶乘形式展开, 可以得到

$$\frac{6!}{2!4!}\times\frac{4!}{2!2!}\times\frac{2!}{2!0!} = \frac{6!}{2!2!2!}$$

这是将 6 个不同元素分成大小为 2, 2, 2 的三个**有序**组的方法数, 其形式与多项式系数一致.

有序分组模型

将 n 个不同元素分成 m 个**可区分的**组, 例如, 分配给 m 个不同的人, 或放入 m 个有标签的盒子, 各组的元素个数分别为 $k_1, k_2, \ldots, k_m$, 其中 $\sum k_i = n$, 其总方法数为

$$N = \binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{k_m}{k_m} = \frac{n!}{k_1!k_2!\cdots k_m!}$$

证明　我们用构造性论证证明此公式. 思路是将完成一次有序分组这一任务, 分解为一系列连续且独立的组合选择, 然后应用乘法原理.

设原始集合为 S, 包含 n 个不同元素. 现将其划分为 m 个有序子集 $G_1, G_2, \ldots, G_m$, 使得对每个 $i = 1, \ldots, m$, 子集 G_i 的大小为 $|G_i| = k_i$.

可按顺序为这 m 个可区分的组选择它们的成员.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

首先为第一个组 G_1 选择成员. 需从 S 中的 n 个元素里不计顺序地选出 k_1 个. 根据组合数的定义, 完成这一步的方法数为

$$\binom{n}{k_1}$$

在确定 G_1 的成员后, 集合 S 中还剩下 $n-k_1$ 个元素可供分配. 接下来为第二个组 G_2 选择成员. 需从这 $n-k_1$ 个剩余元素中选出 k_2 个. 方法数为

$$\binom{n-k_1}{k_2}$$

这个过程依次进行. 当为第 i 个组 G_i 选择成员时, 已有 $k_1+k_2+\cdots+k_{i-1}$ 个元素被分配. 因此, 需从剩下的 $n-\sum_{j=1}^{i-1}k_j$ 个元素中, 选出 k_i 个.

直至最后一个组 G_m, 此时剩下的元素数量为 $n-\sum_{j=1}^{m-1}k_j=k_m$. 需从这 k_m 个元素中选出 k_m 个, 方法数为 $\binom{k_m}{k_m}=1$.

由于整个分组任务必须完成所有这些步骤, 根据乘法原理总的方法数 N 是每一步方法数的乘积.

$$N=\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{n-\sum_{j=1}^{m-2}k_j}{k_{m-1}}\binom{k_m}{k_m}$$

这便证明了公式的第一个等式.

为了推导其更简洁的阶乘形式, 将每一个组合数用其阶乘定义展开.

$$\begin{aligned}N&=\frac{n!}{k_1!(n-k_1)!}\times\frac{(n-k_1)!}{k_2!(n-k_1-k_2)!}\times\cdots\times\frac{(k_m)!}{k_m!(k_m-k_m)!}\\&=\frac{n!}{k_1!\cancel{(n-k_1)!}}\times\frac{\cancel{(n-k_1)!}}{k_2!\cancel{(n-k_1-k_2)!}}\times\cdots\times 1\end{aligned}$$

观察到前一项分母中的阶乘项, 恰好被后一项的分子抵消. 这种级联约分会一直持续到最后.

最终所有中间的阶乘项都被消去, 只剩下最初的分子 $n!$ 和每一项分母中代表各组大小的阶乘 $k_i!$.

$$N=\frac{n!}{k_1!k_2!\cdots k_m!}$$

这与多项式系数的形式完全吻合. □

例题 9.32 某学校的志愿者社团共有 9 名成员, 需要将他们分配到 A、B、C 三个不同的社区服务中心.A 中心需要 2 名成员, B 中心需要 3 名成员, C 中心需要 4 名成员. 问共有多少种不同的分配方案?

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 这个问题要求将 9 名不同成员分配到三个可区分的中心, 且每个中心所需人数已经确定, 因此是一个有序分组问题. 把分配过程视为一个连续的决策步骤.

首先为 A 中心从 9 名成员中选派 2 人. 这是一个组合问题, 因为成员在此步骤中仅被选中, 其顺序无关紧要. 方法数为

$$\binom{9}{2}$$

接着从余下的 $9-2=7$ 名成员中为 B 中心选派 3 人. 方法数为

$$\binom{7}{3}$$

最后剩下的 $7-3=4$ 名成员将全部被派往 C 中心. 方法数为

$$\binom{4}{4}$$

根据乘法原理, 完成整个分配任务的总方案数是上述步骤方法数的乘积.

$$N=\binom{9}{2}\binom{7}{3}\binom{4}{4}=\frac{9\times 8}{2\times 1}\times\frac{7\times 6\times 5}{3\times 2\times 1}\times 1=36\times 35\times 1=1260$$

此结果也可直接通过有序分组的通用公式得到. 将 $n=9$ 个元素分成大小为 $k_1=2,k_2=3,k_3=4$ 的可区分三组, 其方法数由多项式系数给出.

$$\binom{9}{2,3,4}=\frac{9!}{2!3!4!}=\frac{362880}{2\times 6\times 24}=\frac{362880}{288}=1260$$

两种计算路径的结果一致. 故共有 1260 种不同的分配方案.

例题 9.33 将 5 名新入职的员工分配到三个不同的项目中, 其中项目 A 需要 1 名员工, 项目 B 需要 2 名员工, 项目 C 需要 2 名员工. 问有多少种不同的分配方案?

解 此任务的要点是将 5 名可区分的员工, 分配到三个可区分的项目 (A、B、C) 中. 这是一个有序分组问题, 即使其中两个项目所需的人数相同, 都是 2 人, B 与 C 这两个项目本身的差异性使分组是**有序**的.

我们采用连续组合选择的方式来解决. 首先, 为项目 A 从 5 名员工中选定 1 人. 方法数为

$$\binom{5}{1}=5$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

接着, 为项目 B 从剩下的 4 名员工中选定 2 人. 方法数为

$$\binom{4}{2}=\frac{4\times 3}{2}=6$$

最后剩下的 2 名员工自动被分配到项目 C. 方法数为

$$\binom{2}{2}=1$$

依据乘法原理, 总的分配方案数为

$$N=\binom{5}{1}\binom{4}{2}\binom{2}{2}=5\times 6\times 1=30$$

也可直接应用有序分组公式验证. 将 $n=5$ 个元素分成大小为 $k_1=1,k_2=2,k_3=2$ 的三组, 其方法数为

$$\binom{5}{1,2,2}=\frac{5!}{1!2!2!}=\frac{120}{1\times 2\times 2}=30$$

两种路径再次得到相同的结果. 故共有 30 种不同的分配方案.

例题 9.34 在一副标准的 52 张扑克牌(所有牌均不相同)中, 将所有牌平均发给东、南、西、北四位玩家, 每位玩家获得 13 张牌. 问总共有多少种不同的发牌结果?

解 这个问题要求将 52 张不同的牌分配给四个可区分的玩家. 这是将 $n=52$ 个不同元素分成四组、每组大小均为 $k=13$ 的有序分组问题.

我们模拟发牌的连续过程来构建解决方案. 首先为东这位玩家从 52 张牌中发出 13 张. 方法数为

$$\binom{52}{13}$$

接着为南这位玩家从剩下的 $52-13=39$ 张牌中发出 13 张. 方法数为

$$\binom{39}{13}$$

然后为西这位玩家从剩下的 $39-13=26$ 张牌中发出 13 张. 方法数为

$$\binom{26}{13}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

最后剩下的 13 张牌全部发给北这位玩家. 方法数为

$$\binom{13}{13}$$

根据乘法原理, 总的发牌结果数为

$$N=\binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13}$$

现在利用有序分组的通用公式从更宏观的视角验证这个结果. 该问题是把 $n=52$ 个元素分成大小分别为 $k_1=13,k_2=13,k_3=13,k_4=13$ 的四个可区分的组. 其总方法数由以下公式直接给出.

$$N=\frac{52!}{13!13!13!13!}=\frac{52!}{(13!)^4}$$

我们验证这两种表达方式的等价性, 以弄清公式的来龙去脉.

$$\begin{aligned}
&\binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13}\\
&=\frac{52!}{13!(52-13)!}\cdot\frac{39!}{13!(39-13)!}\cdot\frac{26!}{13!(26-13)!}\cdot\frac{13!}{13!(13-13)!}\\
&=\frac{52!}{13!\cdot 39!}\cdot\frac{39!}{13!\cdot 26!}\cdot\frac{26!}{13!\cdot 13!}\cdot\frac{13!}{13!\cdot 0!}\\
&=\frac{52!}{13!\cdot \cancel{39!}}\cdot\frac{\cancel{39!}}{13!\cdot \cancel{26!}}\cdot\frac{\cancel{26!}}{13!\cdot 13!}\cdot 1\quad(\text{注意}0!=1)\\
&=\frac{52!}{13!\cdot 13!\cdot 13!\cdot 13!}
\end{aligned}$$

两种方法的结果完全吻合. 最终结果十分庞大, 通常无需计算其具体值, 以阶乘或组合数形式表示即可.

9.5.4.2 无序分组

现在转向情景二. 在这里只是将书打包成三份, 这些份或堆之间没有任何身份标签, 因此不可区分. 例如, 分组结果 $\{\{B_1,B_2\},\{B_3,B_4\},\{B_5,B_6\}\}$ 与 $\{\{B_3,B_4\},\{B_1,B_2\},\{B_5,B_6\}\}$ 是完全**相同**的一种分组, 因为它们都只是把这六本书分成了同样的三堆.

直接计算无序分组较困难, 但可借助有序分组的结果和倍缩策略. 我们先把这些组视为有区别的, 比如给它们贴上 A、B、C 的标签. 正如有序分组计算所得, 这样做有 90 种方法.

现在分析这种贴标签的操作导致了多少倍的重复计数. 考虑一个具体的无序分组结果, 例如 $G=\{\{B_1,B_2\},\{B_3,B_4\},\{B_5,B_6\}\}$. 在我们的有序分组计数中, 这个单一结果对应于将 A、B、C 三个标签分配给这三个小组的全部分配方式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- (A: $\{B_1, B_2\}$, B: $\{B_3, B_4\}$, C: $\{B_5, B_6\}$)
- (A: $\{B_1, B_2\}$, C: $\{B_3, B_4\}$, B: $\{B_5, B_6\}$)
- ……

知乎@余命数

将 A、B、C 三个标签全排列, 共有 $3! = 6$ 种方式. 即每一个我们真正想要的无序分组, 在有序分组的计算中都被重复计数了 3! 次.

因此根据倍缩策略, 我们必须用有序分组的结果除以这个重复的倍数.

$$N_2 = \frac{N_1}{3!} = \frac{\binom{6}{2}\binom{4}{2}\binom{2}{2}}{3!} = \frac{90}{6} = 15$$

因此将 6 本不同的书分成无法区分的三堆, 每堆 2 本, 只有 15 种方法.

这个思想可以推广. 如果将 n 个元素分组, 其中有 m 个组的大小完全相同, 那么在计算有序分组后, 需要除以 $m!$ 来消除这些相同大小的组之间的顺序带来的重复.

知乎@余命数

无序分组模型

将 n 个不同元素分组.

1. **各组大小均不相同**:若各组大小为 $k_1, k_2, \ldots, k_m$ 且彼此不等, 则无序分组与有序分组的结果相同, 为 $\dfrac{n!}{k_1!k_2!\cdots k_m!}$.(因为组的大小本身就成为了它们的天然标签)
2. **部分组大小相同**:若其中有 m_1 个组的大小为 k_1, m_2 个组的大小为 k_2, …, 则总方法数为

$$N = \frac{\binom{n}{\cdots}}{m_1!m_2!\cdots} = \frac{n!}{(k_1!)^{m_1}(k_2!)^{m_2}\cdots(m_1!)(m_2!)\cdots}$$

先按有序分组计算, 然后对每一类大小相同的组, 除以该类组数的阶乘.

知乎@余命数

知乎@余命数

例题 9.35 将 6 个颜色互不相同的球全部放入三个完全相同的盒子里, 每个盒子里至少放一个球且数目不同, 则不同的放球方法有(　)种.

A. 10　　B. 20　　C. 360　　D. 60

解 相异球放入相同盒子, 且附加每盒至少一个、各盒数目互异的约束. 先确定各组球数, 再计算分配方案.

设三盒球数为 k_1, k_2, k_3, 需 $k_1 + k_2 + k_3 = 6$, $k_i \geq 1$ 且两两不同. 即把 6 分拆为三个不同正整数之和. 设 $1 \leq k_1 < k_2 < k_3$. 若 $k_1 = 1$, 则 $k_2 + k_3 = 5$, 唯一解

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

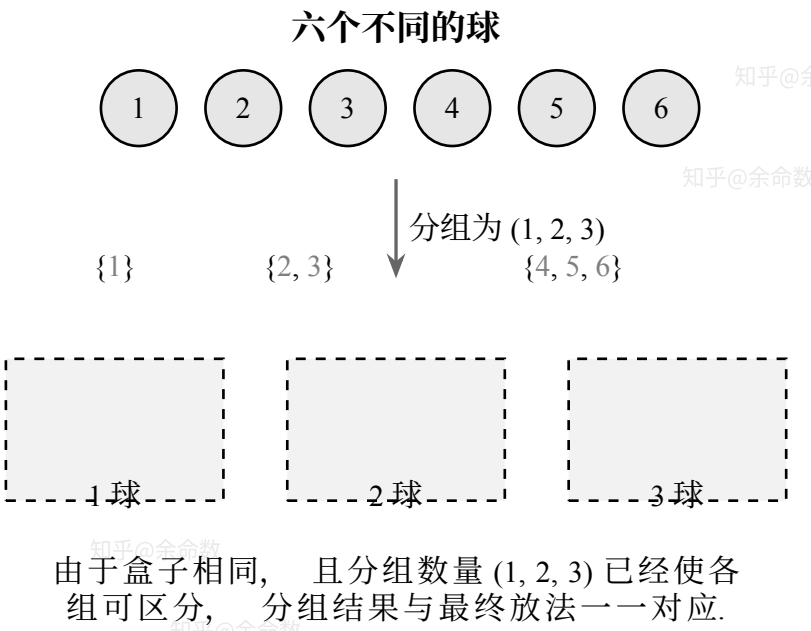


图 9.6　分组过程图解

$k_2=2, k_3=3$, 得方案 $(1,2,3)$. 若 $k_1\geq 2$, 最小和为 $2+3+4=9>6$, 不可能. 故唯一数量方案是分成 $1,2,3$ 三组.

再把 6 个不同球按大小 $1,2,3$ 分组, 先选 1 个有 $\binom{6}{1}$ 种, 再选 2 个有 $\binom{5}{2}$ 种, 余下 3 个自成一组, 故

$$N=\binom{6}{1}\binom{5}{2}\binom{3}{3}=6\times 10=60.$$

盒子相同, 分组本应无序; 但三组大小分别为 $1,2,3$, 大小本身即可区分各组, 无需再除以 3!. 因而放法总数即 60. 故选 D.

例题 9.36　将 10 名实习生分配到三个不同的部门 A、B、C 进行实习, 其中 A 部门 3 人, B 部门 3 人, C 部门 4 人. 问有多少种分配方案?

解　这是一个典型有序分组问题, 因为部门 A、B、C 明确可区分.

分步选择. 首先为 A 部门从 10 人中选出 3 人, 有 $\binom{10}{3}$ 种方法. 然后为 B 部门从剩下的 7 人中选出 3 人, 有 $\binom{7}{3}$ 种方法. 最后剩下的 4 人全部到 C 部门, 有 $\binom{4}{4}$ 种方法.

根据乘法原理, 总方案数为

$$N=\binom{10}{3}\binom{7}{3}\binom{4}{4}=\frac{10\cdot 9\cdot 8}{3\cdot 2\cdot 1}\times\frac{7\cdot 6\cdot 5}{3\cdot 2\cdot 1}\times 1=120\times 35=4200$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此共有 4200 种不同的分配方案.

例题 9.37 将 10 名实习生分成三组, 两组 3 人, 一组 4 人, 再将这三组派往三个不同的城市进行社会调研. 问有多少种派遣方案?

解 问题包含分组与分配两阶段.

方法一采用分步处理, 先分组, 后分配. 先将 10 人分成两组 3 人、一组 4 人, 此时组本身**没有身份**, 只关心成员构成; 两个 3 人组大小相同, 不可区分. 若先按有序分组计算, 方法数为 $\binom{10}{3}\binom{7}{3}\binom{4}{4}=4200$; 因两个 3 人组不可区分, 需再除以 2!, 得

$$N_{\text{分组}}=\frac{\binom{10}{3}\binom{7}{3}\binom{4}{4}}{2!}=2100.$$

再将这三组(人员构成已定, 因而可区分)派往三个不同城市, 有 $3!=6$ 种分配. 总方案数 $N=2100\times 6=12600$.

方法二采用直接分配. 将 10 人直接分到三个城市岗位, 其中两城各 3 人、一城 4 人, 按哪个城市得 4 人分类. 城市 1 得 4 人、其余各 3 人, 方法数为 $\binom{10}{4}\binom{6}{3}\binom{3}{3}=4200$; 城市 2 或城市 3 得 4 人时同样各为 4200. 由加法原理, $N=3\times 4200=12600$.

两种方法结果一致, 并体现计数路径的等价关系

$$\left(\frac{\binom{10}{3}\binom{7}{3}\binom{4}{4}}{2!}\right)\times 3!=3\times\binom{10}{4}\binom{6}{3}.$$

9.5.5 相同元素的分配与隔板法

到目前为止, 我们处理的元素都是**可区分**的, 例如不同的人、不同的书、不同的节目. 但在另一大类问题中, 待分配的元素是**完全相同**的, 例如把 10 个完全相同的名额分配给 3 个班级. 此时哪几个元素分在一起不再有意义, 唯一重要的是每个接收者**分到几个**. 处理这类问题的标准工具是**隔板法**.

我们从一个具体问题开始分析.

例题 9.38 把 10 个完全相同的小球放入 3 个不同的盒子中, 每个盒子至少放 1 个球, 有多少种不同的放法?

解 元素完全相同意味着无法像处理分组问题那样选出某些元素组成一组. 但换一个视角, 把 10 个小球在桌上排成一行.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

只要在这一行小球中插入 2 块隔板, 小球就被分成三段; 把三段依次放入三个盒子, 就得到一种符合要求的放法. 例如

$$\bullet\bullet \mid \bullet\bullet\bullet \mid \bullet\bullet\bullet\bullet\bullet$$

对应“盒子一放 2 个、盒子二放 3 个、盒子三放 5 个”.

10 个小球排成一行, 球与球之间共有 $10-1=9$ 个内部空隙. 要保证每盒至少 1 球, 隔板不能放在两端, 同一个空隙里也不能放两块隔板, 因此只需从这 9 个空隙中选出 2 个放置隔板. 这是一个组合问题, 方法数为

$$\binom{9}{2}=\frac{9\times 8}{2\times 1}=36$$

因此共有 36 种不同的放法.

这个构造有效, 是因为放隔板的方案与放球的方案之间存在一一对应. 每一种放球方案唯一决定 2 块隔板的位置, 反过来, 每一种隔板放法也唯一决定一种放球方案. 于是我们得到一般性结论.

定理 9.14 (隔板法) 把 n 个完全相同的元素放入 k 个不同的盒子 ($n\geq k\geq 1$), 每个盒子至少放 1 个, 则不同的放法数为

$$\binom{n-1}{k-1}$$

等价地, 方程 $x_1+x_2+\cdots+x_k=n$ 的**正整数解** $(x_1,x_2,\ldots,x_k)$ 的个数也是 $\binom{n-1}{k-1}$.

证明 将 n 个相同元素排成一行, 它们之间形成 $n-1$ 个内部空隙. 从中选出 $k-1$ 个空隙各插入一块隔板, 元素就被分成 k 段, 且每段非空; 按顺序将各段放入 k 个盒子, 即得一种合法放法.

这个映射是双射. 任意一种合法放法, 把各盒中的元素依次排开, 在盒与盒的交界处放回隔板, 就唯一还原出一种隔板放法; 反之, 任意一种隔板放法显然唯一决定一种放球方案. 因此放法总数等于隔板放法数, 即从 $n-1$ 个空隙中选出 $k-1$ 个的组合数 $\binom{n-1}{k-1}$.

令 x_i 表示第 i 个盒子中元素的个数, 则每个盒子至少 1 个等价于 $x_i\geq 1$, 放法与方程 $x_1+\cdots+x_k=n$ 的正整数解一一对应, 故正整数解的个数同为 $\binom{n-1}{k-1}$. □

如果允许盒子为空, 一个借球技巧可以把它化归为已经解决的情形.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 9.15 (允许空盒的情形) 把 n 个完全相同的元素放入 k 个不同的盒子, 允许某些盒子为空, 则不同的放法数为

$$\binom{n+k-1}{k-1}$$

等价地, 方程 $x_1+x_2+\cdots+x_k=n$ 的**非负整数解**的个数也是 $\binom{n+k-1}{k-1}$.

证明 向每个盒子中借入 1 个相同元素, 于是盒中共有 $n+k$ 个元素. 对每一种允许为空的原始放法, 加上借入的元素后每个盒子都至少有 1 个元素, 对应一种把 $n+k$ 个元素放入 k 个盒子且每盒非空的放法; 反过来, 对 $n+k$ 个元素的任意一种每盒非空的放法, 从每个盒子中各拿走 1 个元素, 就唯一还原出一种允许为空的原始放法.

二者一一对应, 故所求放法数等于"$n+k$ 个相同元素放入 k 个不同盒子、每盒非空"的放法数, 由隔板法即 $\binom{n+k-1}{k-1}$. 非负整数解的情形同样由 $x_i\geq 0$ 与放法的一一对应得到. □

隔板法使用要点

- **适用条件**. 待分配的元素**完全相同**, 接收的盒子**彼此不同**, 且只关心每个盒子分到几个.
- **每盒非空**.n 个元素、k 个盒子, 放法数为 $\binom{n-1}{k-1}$, 对应**正整数解**计数.
- **允许空盒**. 放法数为 $\binom{n+k-1}{k-1}$, 对应**非负整数解**计数; 可通过每盒先借 1 个化为非空情形.
- **每盒至少 m 个**. 先给每盒各预放 m 个, 或等价地换元, 剩余元素再按非空情形分配.

例题 9.39 某校把 12 个三好学生名额分配给高三年级的 4 个班, 每班至少 1 个名额, 有多少种分配方案? 若允许某些班没有名额, 又有多少种?

解 名额是完全相同的元素, 班级是彼此不同的盒子, 这是隔板法的标准场景.

第一问. 每班至少 1 个名额. 由隔板法, 分配方案数为

$$\binom{12-1}{4-1}=\binom{11}{3}=\frac{11\times 10\times 9}{3\times 2\times 1}=165$$

第二问. 允许有班级没有名额. 由允许空盒的公式, 分配方案数为

$$\binom{12+4-1}{4-1}=\binom{15}{3}=\frac{15\times 14\times 13}{3\times 2\times 1}=455$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 9.40　求方程 $x_1+x_2+x_3+x_4=20$ 满足 $x_1\geq 1$、$x_2\geq 2$、$x_3\geq 3$、$x_4\geq 4$ 的整数解的个数.

解　各未知数的下界不同, 先通过预付把它们统一化为正整数解问题. 这等价于先往第 2、3、4 个盒子里分别预放 $1,2,3$ 个元素, 再把剩下的元素按每盒至少 1 个分配.

令 $y_1=x_1$、$y_2=x_2-1$、$y_3=x_3-2$、$y_4=x_4-3$, 则原约束等价于 $y_i\geq 1$, 且

知乎@余命数

$$y_1+y_2+y_3+y_4=20-(0+1+2+3)=14$$

由隔板法, 这个方程的正整数解个数为

$$\binom{14-1}{4-1}=\binom{13}{3}=\frac{13\times 12\times 11}{3\times 2\times 1}=286$$

所以原方程满足条件的整数解共有 286 个.

9.5.6 涂色问题 *

知乎@余命数

在加法原理与乘法原理一节中, 我们解决过 4 个方格排成一行、相邻不同色的问题. 从左到右逐格染色, 每格只需避开左边一格的颜色, 乘法原理直接给出答案. 但如果这些方格首尾相接围成一个**环**, 或者染色的对象换成地图、多面体的顶点, 单纯的分步就不再够用, 因为后面某一步的选择数会依赖于前面某两步的颜色是否相同. 这时必须把加法原理与乘法原理结合起来分类讨论. 下面用两个经典的例子说明其中的思想.

例题 9.41 (环形方格)　把 4 个方格围成一圈, 即第 1 格与第 4 格也相邻, 用 5 种不同的颜色染色, 要求相邻方格颜色不同, 有多少种染色方法?

知乎@余命数

解　与链状情形的差别在于, 染第 4 格时既要避开第 3 格的颜色, 又要避开第 1 格的颜色, 而第 1 格和第 3 格是否同色直接影响第 4 格的选择数. 因此按第 1 格与第 3 格是否同色分类.

先染第 1 格, 有 5 种选择; 第 2 格与第 1 格不同色, 有 4 种选择.

情况一. 第 3 格与第 1 格同色第 3 格只需避开第 2 格的颜色, 而第 1 格的颜色恰好满足条件, 因此第 3 格的颜色被唯一确定(1 种选择). 此时第 4 格只需避开第 3 格, 亦即第 1 格这一种颜色, 有 4 种选择. 此类共有 $5\times 4\times 1\times 4=80$ 种.

情况二. 第 3 格与第 1 格不同色第 3 格既要避开第 2 格、又要避开第 1 格的颜色, 有 $5-2=3$ 种选择. 此时第 4 格要避开第 3 格与第 1 格这两种不同的颜色, 有 $5-2=3$ 种选择. 此类共有 $5\times 4\times 3\times 3=180$ 种.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两类互斥, 由加法原理, 总的染色方法数为

$$N = 80 + 180 = 260$$

这个按首尾是否呼应分类的思想可以推广到任意个方格的环, 并得到一般公式. 用 k 种颜色染 n 个环形方格 ($n \geq 2$), 相邻不同色的方法数为

$$a_n = (k-1)^n + (-1)^n(k-1)$$

它的推导需要一点递推技巧, 我们把它留作本章习题. 可以验证 $n = 4$, $k = 5$ 时 $a_4 = 4^4 + 4 = 260$, 与上面的结果一致.

例题 9.42 (2003 全国理) 用 5 种不同的颜色给一个四棱锥 $S-ABCD$ 的 5 个顶点染色, 要求同一条棱的两个端点颜色不同, 有多少种染色方法?

解 顶点 S 与底面四个顶点都相邻, 底面 $ABCD$ 本身又是一个四元环. 仿照上例, 按 A 与 C 是否同色分类处理底面.

先染 S, 有 5 种选择; 再染 A, 与 S 不同, 有 4 种选择; 染 B, 与 S, A 都不同, 有 3 种选择.

情况一. C 与 A 同色 C 只需与 S, B 不同色, 而 A 的颜色恰满足此条件, 因此 C 的颜色被唯一确定(1 种选择). 此时 D 与 S、A、C 相邻, 而 A, C 同色, 所以 D 只需避开 2 种颜色, 有 3 种选择. 此类共有 $5 \times 4 \times 3 \times 1 \times 3 = 180$ 种.

情况二. C 与 A 不同色 C 既要与 S, B 不同色, 又不与 A 同色, 故 C 有 $5-3=2$ 种选择. 此时 D 要避开 S、A、C 这 3 种互不相同的颜色, 有 2 种选择. 此类共有 $5 \times 4 \times 3 \times 2 \times 2 = 240$ 种.

由加法原理, 总的染色方法数为

$$N = 180 + 240 = 420$$

涂色问题的难点在于找到影响后续选择数的关键位置, 如本节的第 1 格与第 3 格、顶点 A 与 C, 并围绕它展开分类.

习题

9.1. (加法原理与乘法原理) 用数字 $0, 1, 2, \ldots, 9$ 组成没有重复数字的三位偶数, 共能组成多少个? 按个位是否为 0 分类, 注意百位不能为 0. 答案 328.

9.2. (组合数的对称性) 计算 $\binom{15}{12} + \binom{20}{18}$. 先利用 $\binom{n}{k} = \binom{n}{n-k}$ 化大为小. 答案 645.

9.3. (组合与补集) 从 5 名男生和 4 名女生中选出 3 人参加活动, 要求至少有 1 名女生, 有多少种不同的选法? 用总数减去全是男生的情形. 答案 74.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

9.4. (捆绑法)7 名同学站成一排, 其中甲、乙两人必须相邻, 有多少种不同的排法? 把甲、乙视为一个复合元素, 注意内部排列. 答案 1440.

9.5. (插空法)7 名同学站成一排, 其中甲、乙两人互不相邻, 有多少种不同的排法? 先排其余 5 人, 再在空隙中安排甲、乙. 答案 3600.

9.6. (倍缩策略)7 名同学站成一排, 要求甲、乙、丙三人的相对顺序保持不变(甲在乙前, 乙在丙前, 三人不必相邻), 有多少种不同的排法? 在全排列中, 三人的 3! 种内部顺序只有 1 种被接受. 答案 840.

9.7. (无序分组)将 6 名不同的同学分成两组, 每组 3 人, 有多少种不同的分法? 若改为分给甲、乙两个小组, 每组 3 人, 又有多少种? 前一问中两个组不可区分, 需除以 2!; 后一问中组有归属, 是有序分组. 答案 10 与 20.

9.8. (隔板法)把 12 个相同的小球放入 4 个不同的盒子, 每个盒子至少 1 个, 有多少种放法? 若允许盒子为空, 又有多少种? 分别对应正整数解与非负整数解的计数. 答案 165 与 455.

9.9. (二项式定理)求 $\left(2x-\dfrac{1}{x}\right)^6$ 展开式中的常数项. 通项为 $\binom{6}{k}(2x)^{6-k}\left(-\frac{1}{x}\right)^k$, 令 x 的指数为 0. 答案 -160.

9.10. (赋值法)求 $(x+2y)^7$ 展开式中所有项的系数之和. 令 $x=y=1$. 答案 2187.

9.11. (交错和) 求 $\binom{10}{0}+\binom{10}{2}+\binom{10}{4}+\cdots+\binom{10}{10}$ 的值. 偶数项二项式系数之和等于 2^{n-1}. 答案 512.

9.12. (环形涂色)用 k 种颜色给 n 个围成一圈的方格染色 ($n\geq 2$), 相邻方格颜色不同, 记方法数为 a_n. 证明:

$$a_n=(k-1)^n+(-1)^n(k-1)$$

先染第 1 格, 再沿环依次染色, 可得递推关系 $a_n+a_{n-1}=k(k-1)^{n-1}$ ($n\geq 3$); 剪开环后, 首尾同色与首尾不同色的情形分别对应 a_{n-1} 与 a_n. 再由递推式用数学归纳法或构造等比数列求出通项.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 10 章

概率

周期·离散·随机

「偶然性，也可以成为严格科学的对象。」

——柯尔莫哥洛夫



第十章 概率

知乎@余命数

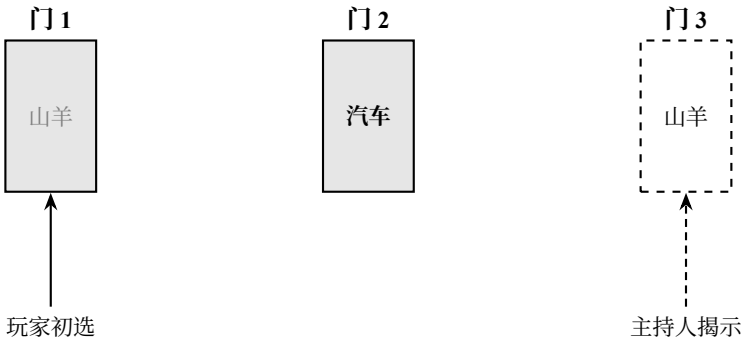


图 10.1 蒙提霍尔问题:玩家初选门 1 后, 主持人揭示门 3 为山羊

蒙提霍尔问题

我们先来看一个源自美国电视游戏的著名谜题. 舞台上有三扇关闭的门, 其中一扇门后面是一辆汽车, 另外两扇门后面各是一只山羊. 我们选择一扇门, 例如门 1, 但暂时不能打开. 此时, 知道门后真相的主持人会打开剩下两扇门中的一扇, 我们便会看到其后面是一只山羊, 例如门 3.

接着他问,“你是坚持最初选择的门 1, 还是换成剩下那扇未打开的门 2?”第一感觉往往是, 既然门 3 已经被排除, 剩下门 1 和门 2, 汽车在其中一扇门后的机会似乎均等, 各有一半. 那么换与不换, 结果应该没有区别.

但这个直觉是错误的; **交换选择会使赢得汽车的机会翻倍**. 要解释这种看似合理的直觉为何失效, 我们需要一套能够精确量化可能性大小的数学工具, 即**概率论**. 用严谨的逻辑分析不确定性, 许多直觉性的判断便会现出原形.

待到 第10.5节 建立条件概率与贝叶斯公式之后, 我们再来回看蒙提霍尔问题, 给出严格解答.

知乎@余命数

10.1 事件与概率

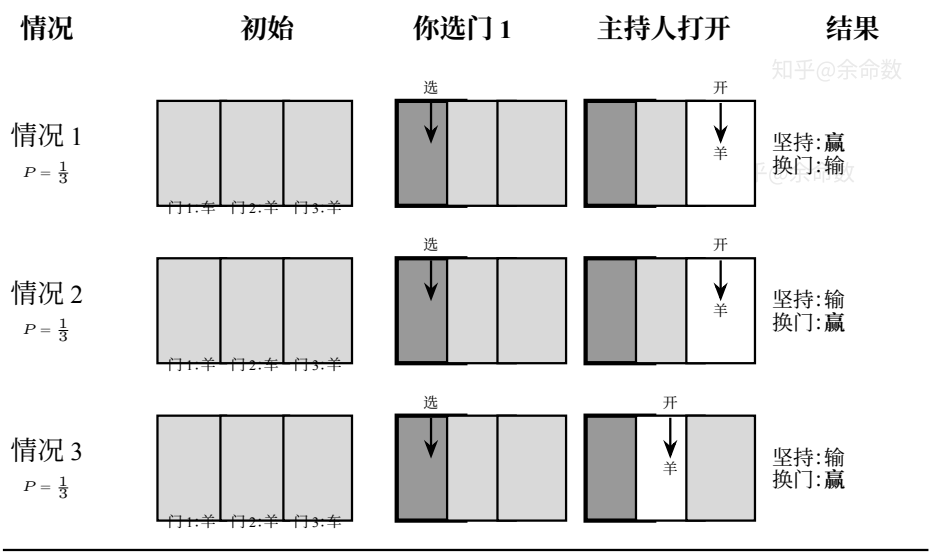
抛掷一枚硬币会出现正面还是反面, 投掷一颗骰子会得到几点, 我们每天都面对诸如此类的不确定性. 概率论提供了一套描述和量化这种不确定性的数学语言. 本章从最基本的概念(事件)出发, 搭建起初步的概率论体系.

10.1.1 随机试验与事件

要讨论可能性的大小, 我们先要精确界定所讨论的对象. 概率论的起点是**随机试验**.

一个过程或试验, 若它满足以下三个基本特征, 我们便称之为随机试验.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



在 3 种等可能情况中，　坚持策略仅在情况 1 获胜 ($P = \frac{1}{3}$), 换门策略在情况 2 和 3 获胜 ($P = \frac{2}{3}$)

图 10.2　蒙提霍尔问题:穷举所有可能情况(假设玩家初选门 1)

1. **可重复性**试验可以在相同的条件下重复进行.
2. **可知性**试验的所有可能结果是明确已知的.
3. **不确定性**每次试验前无法预知具体结果.

知乎@余命数　知乎@余命数

例题 10.1 (随机试验的例子)

- **抛掷一枚硬币**可以在相同条件下重复抛掷, 所有可能结果是正面或反面, 但每次抛掷前不知道具体是哪一面朝上.
- **掷一颗骰子**所有可能结果是点数 $\{1,2,3,4,5,6\}$.
- **记录某路口一小时内通过的汽车数量**所有可能结果是 $\{0,1,2,\dots\}$.
- **从区间 $[0,1]$ 上随机取一个数**所有可能结果是该区间内的任何一个实数.

对于随机试验, 我们把它所有可能的基本结果构成的集合称为**样本空间**, 通常用大写希腊字母 Ω 表示. 样本空间中的每一个元素, 即每一个不能再分解的基本结果, 称为一个**样本点**.

定义 10.1 (样本空间与样本点) 随机试验 E 的所有可能结果构成的集合称为 E 的**样本空间**, 记为 Ω. 样本空间 Ω 中的元素, 即试验的每一个基本结果, 称为**样本点**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 10.2 (样本空间的确定)

- 抛掷一枚硬币一次, 记正面为 H, 反面为 T. 样本空间为 $\Omega = \{H, T\}$.
- 掷一颗骰子一次, 样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- 抛掷一枚硬币两次, 记录每次的结果. 样本空间为

$$\Omega = \{(H, H),\ (H, T),\ (T, H),\ (T, T)\}.$$

- 从装有 2 个红球和 3 个白球的袋中, 不放回地抽取 2 个球. 若关心球的颜色组合, 则样本空间可设为

$$\Omega = \{\{\text{红}, \text{红}\},\ \{\text{红}, \text{白}\},\ \{\text{白}, \text{白}\}\}.$$

若对每个球加以区分, 例如 R_1, R_2, W_1, W_2, W_3, 则样本空间会复杂得多. **样本空间的定义方式取决于我们所关心的问题**.

知乎@余命数

上面最后一个例子表明, 定义样本空间时, 我们既要罗列结果, 也要围绕研究目标建模. 样本空间是为解决特定问题而构建的数学模型.

10.1.1.1 模型的选择

10.1.1.2 等可能性

在应用古典概型时, 一个根本性的要求是样本空间中的所有样本点必须**等可能**. 然而, 等可能性往往只是建模时做出的一种**理想化假设**. 通过恰当地定义样本点, 我们可以构造出满足等可能性的样本空间.

知乎@余命数

例题 10.3 (区分与不区分的抉择) 一个袋中装有 2 个红球和 2 个白球, 除颜色外完全相同. 从中不放回地随机抽取 2 个球, 求抽到 2 个红球的概率是多少?

解 对此试验, 我们至少可以构想出两种不同的样本空间.

第一种只关心最终抽出的颜色组合, 不区分同色球. 可能的结果有三种, 即两红、一红一白、两白. 样本空间取为

$$\Omega_1 = \{\text{两红}, \text{一红一白}, \text{两白}\}$$

如果草率地认定这三个结果等可能, 我们就会得出 $P(\text{两红}) = \frac{1}{3}$ 的错误结论. 这个模型虽然直观, 但样本点并不等可能; 它忽略了不同球之间的差异, 压缩了信息.

第二种改从抽取过程入手, 区分同色球以构造等可能的样本点. 即便这些球物理上无法分辨, 数学上仍可给它们贴上标签 R_1, R_2, W_1, W_2. 于是抽取两球等

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

同于从这 4 个可区分的物体中取出 2 个, 不考虑顺序. 所有可能的组合构成新样本空间

$$\Omega_2 = \{\{R_1, R_2\}, \{W_1, W_2\}, \{R_1, W_1\}, \{R_1, W_2\}, \{R_2, W_1\}, \{R_2, W_2\}\}$$

知乎@余命数

这个样本空间共有 $\binom{4}{2} = 6$ 个样本点. 由于每个球被抽到的机会均等, 我们可以合理地假设这 6 个组合是**等可能的**.

现在考察事件 A, 它表示抽到 2 个红球. 在 Ω_2 中, 这个事件对应于子集 $A = \{\{R_1, R_2\}\}$. 该事件只包含了 1 个样本点. 根据古典概型, 正确的概率为

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega_2)} = \frac{1}{6}$$

这个对比清楚地展示了建模思路. 在数学上将物理上不可区分的物体视为可区分的, 便能构造出满足等可能性的样本空间, 从而套用古典概型. **好的样本空间能把一个复杂的、非等可能的问题化为一个简单的、等可能的组合计数问题**.

样本空间的构建还依赖于所关心问题的**细节层次**. 一个过于粗糙的样本空间可能丢失解决问题所需的信息, 而一个过于精细的样本空间则可能徒增不必要的计算复杂性. 例如, 在连续抛掷三次硬币的试验中, 若关心恰好出现两次正面的概率, 那么记录每次抛掷结果的**有序**样本空间 $\Omega = \{HHH, HHT, \ldots, TTT\}$ 是必要的. 若只关心正面的总数, 则样本空间可取为 $\Omega = \{0, 1, 2, 3\}$, 它虽然合法, 但样本点不等可能, 反而增加了计算难度.

10.1.1.3 事件的定义与表示

定义了样本空间之后, 我们便可定义事件. 在日常语言中, 事件指发生的事情.

知乎@余命数

定义 10.2 (事件) 样本空间 Ω 的任意一个**子集** A ($A \subseteq \Omega$), 称为一个**随机事件**, 简称**事件**. 当且仅当一次试验的结果, 即样本点属于子集 A 时, 我们称**事件 A 发生**.

这个定义是现代概率论的基础. 它把随机、不确定的概念(事件的发生)与确定、严谨的数学对象(集合)联系起来.

例题 10.4 (事件的集合表示) 考虑掷一颗骰子的试验, 其样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$. 看几个事件的例子.

- 事件 A 表示掷出的点数为偶数. 用集合表示, $A = \{2, 4, 6\}$.
- 事件 B 表示掷出的点数大于 4. 用集合表示, $B = \{5, 6\}$.
- 事件 C 表示掷出的点数为 7. 这是一个不可能发生的情况, 对应的集合是空集 $C = \emptyset$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 事件 D 表示掷出的点数小于 10. 这是一个必然发生的情况, 对应的集合是整个样本空间 $D=\Omega$.

从这个例子中可见两种特殊的事件.

- **必然事件**每次试验中必定发生的事件. 它对应于整个样本空间 Ω.
- **不可能事件**每次试验中都绝不会发生的事件. 它对应于空集 $\emptyset$.

知乎@余命数

10.1.2 事件的运算与关系

既然事件就是集合, 集合论中的各种关系和运算便对应事件之间的逻辑关系, 这套语言足以精确描述复杂的概率问题. 我们以掷一颗骰子的试验为例, 令样本空间 $\Omega=\{1,2,3,4,5,6\}$.

定义 10.3 (包含关系) 若事件 A 的发生必然导致事件 B 的发生, 则称事件 B **包含**事件 A, 记作 $A\subseteq B$ 或 $B\supseteq A$.

例如, 把点数为 2 或 4 的事件记作 $C=\{2,4\}$, 把点数为偶数的事件记作 $B=\{2,4,6\}$. 那么 $C\subseteq B$, 因为掷出 2 点或 4 点必然意味着掷出了偶数点.

定义 10.4 (和事件) 事件"A 与 B **至少有一个**发生"称为 A 与 B 的**和事件**, 记作 $A\cup B$.

例如, 把点数不大于 4 的事件记作 $A=\{1,2,3,4\}$, 把点数为偶数的事件记作 $B=\{2,4,6\}$. 那么和事件 $A\cup B=\{1,2,3,4,6\}$, 表示点数不大于 4 或者点数为偶数.

定义 10.5 (积事件) 事件 "A 与 B **同时**发生" 称为 A 与 B 的**积事件**, 记作 $A\cap B$ 或 AB.

例如, 对于上述事件 A,B, 积事件 $A\cap B=\{2,4\}$, 表示点数既不大于 4 又是偶数.

定义 10.6 (互斥事件) 若事件 A 与 B **不能同时**发生, 则称它们为**互斥事件**或**互不相容事件**. 在集合上表现为它们的交集是空集, 即 $A\cap B=\emptyset$.

例如, 事件 $B=\{2,4,6\}$ 与事件 $D=\{1,3,5\}$ 互斥(前者为偶数, 后者为奇数), 因为 $B\cap D=\emptyset$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 10.7 (差事件) 事件“A 发生而事件 B **不发生**”称为 A 与 B 的**差事件**, 记作 $A-B$. 它等价于 $A\cap\bar{B}$.

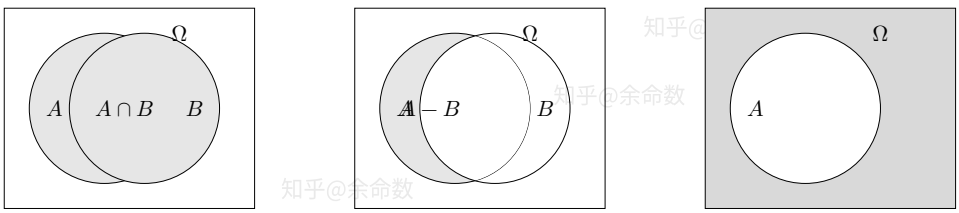
例如, 令 $A=\{1,2,3,4\}, B=\{2,4,6\}$. 那么差事件 $A-B=\{1,3\}$, 表示点数不大于 4 却不是偶数.

定义 10.8 (对立事件) 事件“A **不发生**”称为 A 的**对立事件**或**逆事件**, 记作 $\bar{A}$. 它由样本空间 Ω 中所有不属于 A 的元素构成.

事件 A 与其对立事件 $\bar{A}$ 满足 $A\cup\bar{A}=\Omega$ 且 $A\cap\bar{A}=\emptyset$.A 和 $\bar{A}$ 既互斥又完备, 它们构成对样本空间的一个划分.

例如, 若事件 $A=\{1,2,3,4\}$ 表示点数不大于 4, 则其对立事件为 $\bar{A}=\{5,6\}$ (点数大于 4).

这些事件间的关系和运算, 可用韦恩图直观地表示.



$A\cup B$(整个涂色区域)　　$A-B=A\cap\bar{B}$　　对立事件 $\bar{A}$

图 10.3　事件运算的韦恩图表示

事件的运算满足与集合运算完全相同的定律, 例如交换律、结合律、分配律以及德·摩根定律.

定理 10.1 (德·摩根定律) 对于任意两个事件 A 和 B:

1. $\overline{A\cup B}=\bar{A}\cap\bar{B}$　(A 或 B 都不发生等价于 A 不发生且 B 不发生)
2. $\overline{A\cap B}=\bar{A}\cup\bar{B}$　即 A 和 B 不都发生等价于 A 不发生或 B 不发生

证明 对任意样本点 $\omega\in\Omega$, 由补集与并集的定义依次推得

$$\omega\in\overline{A\cup B}\iff\omega\notin A\cup B\iff(\omega\notin A)\text{ 且}(\omega\notin B)\iff\omega\in\bar{A}\cap\bar{B}.$$

该等价对一切 ω 成立, 故 $\overline{A\cup B}=\bar{A}\cap\bar{B}$.

交换 $\cup$ 与 $\cap$ 的角色, 同理有

$$\omega\in\overline{A\cap B}\iff\omega\notin A\cap B\iff(\omega\notin A)\text{ 或}(\omega\notin B)\iff\omega\in\bar{A}\cup\bar{B},$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

从而 $\overline{A \cap B} = \bar{A} \cup \bar{B}$. □

例题 10.5 高中生李华决定报名参加学校的社团活动, 有三种选择, 即 ACG 社、辩论社与虚拟联合国, 依次记作 A、B、C. 设 A, B, C 分别表示李华报名参加了 ACG 社、报名参加了辩论社、报名参加了虚拟联合国这三个事件. 请使用 A, B, C 的运算关系表示下列事件. (1) 李华至少报名参加了一个社团. (2) 李华同时报名参加了全部三个社团. (3) 李华报名参加了 ACG 社, 但没有报名参加辩论社. (4) 李华三个社团都没有报名参加. (5) 李华恰好只报名参加了一个社团.

解 本题要点在于理解或、且、非等逻辑词与事件运算的对应关系.

(1) 至少报名参加一个社团意味着参加了 A 或参加了 B 或参加了 C. 这是三个事件的和事件.

$$A \cup B \cup C$$

(2) 同时报名参加了全部三个社团意味着参加了 A 且参加了 B 且参加了 C. 这是三个事件的积事件.

$$A \cap B \cap C \quad \text{或简记为} \quad ABC$$

(3) "报名参加了 ACG 社, 但没有报名参加辩论社"意味着事件 A 发生且事件 B 不发生, 即 $\bar{B}$ 发生.

$$A \cap \bar{B} \quad \text{或记为} \quad A - B$$

(4) 三个社团都没有报名参加是至少报名参加一个社团的对立事件.

$$\overline{A \cup B \cup C}$$

根据德·摩根定律, 这也可以表示为没有参加 A 且没有参加 B 且没有参加 C.

$$\bar{A} \cap \bar{B} \cap \bar{C}$$

(5) 恰好只报名参加一个社团是一个更复杂的逻辑, 它由三个互斥的部分组成, 即只参加了 A、只参加了 B、只参加了 C. 只参加了 A 意味着"参加了 A, 且没参加 B, 且没参加 C", 即 $A\bar{B}\bar{C}$. 因此, 该事件为这三个互斥事件的和事件.

$$(A\bar{B}\bar{C}) \cup (\bar{A}B\bar{C}) \cup (\bar{A}\bar{B}C)$$

这个练习把日常语言的逻辑结构分解, 再映射到集合运算.

知乎关注@余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 10.6　某电子设备由两个部件构成. 令事件 A 表示部件 1 在一年内失效, 事件 B 表示部件 2 在一年内失效. 已知 $P(A)=0.1$、$P(B)=0.08$、$P(AB)=0.02$. 求该设备在一年内能正常工作的概率.

解　设备能正常工作, 当且仅当部件 1 没有失效且部件 2 没有失效, 即事件 $\bar{A}\cap\bar{B}$.

直接计算 $P(\bar{A}\cap\bar{B})$ 较为困难, 转而利用其与和事件的关系. 由德·摩根定律

$$\bar{A}\cap\bar{B}=\overline{A\cup B},$$

其直观意义是两个部件都没坏恰好为至少有一个部件坏了的对立事件, 于是

$$P(\bar{A}\cap\bar{B})=P(\overline{A\cup B})=1-P(A\cup B).$$

再由广义加法公式 $P(A\cup B)=P(A)+P(B)-P(AB)$ 代入已知数据,

$$P(A\cup B)=0.1+0.08-0.02=0.16,$$

故设备正常工作的概率为

$$P(\text{正常工作})=1-P(A\cup B)=1-0.16=0.84.$$

例题 10.7　某班级有 50 名学生, 参加了数学、物理、化学三科的兴趣小组. 已知参加数学的有 30 人, 物理 25 人, 化学 20 人. 同时参加数学和物理的有 15 人, 数学和化学 10 人, 物理和化学 8 人. 三科都参加的有 5 人. 问题如下. (1) 至少参加一科的有多少人? (2) 恰好只参加一科的有多少人?

解　设 M,P,C 分别表示参加数学、物理、化学兴趣小组的学生构成的集合. 这是一个典型的容斥原理问题.

(1) 至少参加一科的人数即为 $\text{card}(M\cup P\cup C)$. 根据三事件的容斥原理公式, 我们先把单个事件的基数相加, 再减去两两相交事件的基数, 最后加上三者相交事件的基数.

$$\begin{aligned}\text{card}(M\cup P\cup C)=&(\text{card}(M)+\text{card}(P)+\text{card}(C))\\&-(\text{card}(MP)+\text{card}(MC)+\text{card}(PC))\\&+\text{card}(MPC)\end{aligned}$$

代入数据.

$$\begin{aligned}\text{card}(M\cup P\cup C)&=(30+25+20)-(15+10+8)+5\\&=75-33+5=47\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

所以至少参加一科的有 47 人;另外 3 人什么都没参加.

(2) 恰好只参加一科的人数需要从总数中剔除参加两科及三科的人,利用韦恩图辅助思考会非常清晰. 设 N_1, N_2, N_3 分别为恰好参加一科、两科、三科的人数.

三科都参加的人数 $N_3 = \text{card}(MPC) = 5$ 人. 在此基础上求恰好参加两科的人数. 恰好参加数学和物理两科,即参加这两科但不参加化学的人数为

$$\text{card}(MP) - \text{card}(MPC) = 15 - 5 = 10$$

同理,恰好参加物理和化学两科的人数为 $8 - 5 = 3$,恰好参加数学和化学两科的人数为 $10 - 5 = 5$,故恰好参加两科的总人数 $N_2 = 10 + 3 + 5 = 18$ 人.

再由 N_3 与 N_2 反求恰好只参加一科的人数. 恰好只参加数学一科的人数,等于参加数学的总人数减去所有与数学有交集的恰好两科和三科的人数.

$$\begin{aligned} &30 - (\text{恰好数物}) - (\text{恰好数化}) - (\text{三科都参加}) \\ &= 30 - 10 - 5 - 5 = 10 \end{aligned}$$

同理,恰好只参加物理一科的人数为 $25 - 10 - 3 - 5 = 7$,恰好只参加化学一科的人数为 $20 - 5 - 3 - 5 = 7$,故恰好只参加一科的总人数 $N_1 = 10 + 7 + 7 = 24$ 人.

检验.$N_1 + N_2 + N_3 = 24 + 18 + 5 = 47$,与 (1) 的结果吻合.

例题 10.8 证明对于任意两个事件 A, B,以下等式恒成立.

$$(A \cup B) \cap (A \cup \bar{B}) = A$$

解 我们用分配律来证明.

把事件的并集与交集分别类比为数的加法与乘法,分配律 $x(y+z) = xy + xz$ 在此对应于 $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$,因而公共因子 A 可提取. 整个化简只需三步.

$$\begin{aligned} (A \cup B) \cap (A \cup \bar{B}) &= A \cup (B \cap \bar{B}) && \text{(分配律)} \\ &= A \cup \emptyset && \text{(对立事件} B \cap \bar{B} = \emptyset\text{)} \\ &= A && \text{(空集与任意事件的并集为自身).} \end{aligned}$$

我们用逻辑分析来证明.

分析事件 $(A \cup B) \cap (A \cup \bar{B})$ 发生的条件. 该事件发生意味着 $(A \cup B)$ 与 $(A \cup \bar{B})$ 同时发生. 分两种情况.

- **若事件 A 发生**,则 $A \cup B$ 与 $A \cup \bar{B}$ 均必然发生,整个积事件发生.
- **若事件 A 不发生,即 $\bar{A}$ 发生**,则要使 $A \cup B$ 发生须 B 发生,要使 $A \cup \bar{B}$ 发生须 $\bar{B}$ 发生,二者不能同时成立.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

该复合事件发生的充要条件即事件 A 发生，故它等价于事件 A.

熟练运用事件的运算法则，可以像代数一样简洁地化简和证明复杂的事件关系.

知乎@余命数

例题 10.9　从一副标准的 52 张扑克牌中随机抽取一张. 定义下列事件. 设 A 表示抽到红桃. 设 B 表示抽到黑桃. 设 C 表示抽到红色牌，即红桃或方片. 设 D 表示抽到 J、Q、K 中的一张（人头牌）.
请判断下列说法的正误. (1) 事件 A 与 B 是互斥的吗? (2) 事件 A 与 C 是互斥的吗? (3) 事件 A 与 B 是对立事件吗? (4) 事件 C 与 $\bar{C}$ 是对立事件吗? (5) 事件 A 与 D 是互斥的吗?

解　此题旨在厘清互斥与对立这两个容易混淆的概念的区别.

(1) **正确**. 一张牌不能既是红桃又是黑桃，所以它们的交集为空，即 $A \cap B = \emptyset$.

(2) **错误**. 红桃本身就是红色的牌. 事件 A 的发生必然导致事件 C 的发生，即 $A \subseteq C$. 因此它们的交集 $A \cap C = A \neq \emptyset$.

(3) **错误**. 对立事件必须满足两个条件. 第一，它们互斥 ($A \cap B = \emptyset$)；第二，它们的并集是整个样本空间 ($A \cup B = \Omega$). 从 (1) 可知，事件 A 和 B 确实满足互斥的条件. 但是，它们的并集 $A \cup B$ 代表抽到红桃或黑桃，这并未包含所有可能性，例如方片和梅花. 因此 $A \cup B \neq \Omega$. 由于第二个条件不满足，它们不是对立事件.

(4) **正确**. 我们验证对立事件的两个条件. 事件 C 是抽到红色牌，其对立事件 $\bar{C}$ 就是没有抽到红色牌，即抽到黑色牌. (1) 一张牌不能既红又黑，所以 $C \cap \bar{C} = \emptyset$，即两事件互斥. (2) 任何一张牌非红即黑，所以 $C \cup \bar{C} = \Omega$（并集为全集）. 两个条件都满足，因此它们是对立事件.

(5) **错误**. 存在一张牌既是红桃又是人头牌，即红桃 J、红桃 Q、红桃 K. 因此 $A \cap D \neq \emptyset$.

互斥只要求事件不能同时发生 ($A \cap B = \emptyset$). 对立则更为严格，它在互斥的基础上，还要求两事件的并集穷尽所有可能性 ($A \cup B = \Omega$). 任何对立事件必然互斥，但互斥事件不一定对立.

至此，事件的形式化描述已经建立. 我们把原本不确定的概念**事件**转化为严谨的数学对象**集合**. 接下来为每一个事件 A 赋予一个数值 $P(A)$，用以量化它发生的可能性. 这就是概率.

10.1.3 概率的定义

在详细分析概率的严格数学体系之前，我们先回到一个更符合直觉的问题. 抛掷均匀硬币、正面朝上的概率是 $\frac{1}{2}$，这句话究竟意味着什么，需要仔细说明.

知乎@余命数

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

10.1.3.1 概率的统计定义

最朴素、也最符合科学实验精神的一种想法是, 概率是一个事件在大量重复试验中发生的**频率的稳定值**.

知乎@余命数

具体地, 在相同条件下对一个随机试验重复进行 n 次, 其中事件 A 发生了 n_A 次. 我们称比值 $f_n(A)=\frac{n_A}{n}$ 为事件 A 在这 n 次试验中发生的**频率**. 大量的实践经验表明, 当试验次数 n 越来越大时, 这个频率 $f_n(A)$ 会表现出稳定性, 它会围绕着某个固定的数值摆动, 且摆动的幅度越来越小. 这个现象被称为**大数定律**. 于是, 一个自然的定义便产生了.

定义 10.9 (概率的统计定义) 在不变的条件下, 重复进行 n 次试验, 事件 A 发生的频率 $f_n(A)$ 随着 n 的增大稳定于某一常数 p. 我们就称这个常数 p 为事件 A 的概率, 记作 $P(A)=p$.

这个定义将概率与可观测的物理世界联系起来, 非常直观. 但要充当数学定义, 它存在一些基本的困难.

首先, 它依赖于频率会稳定于一个常数这个事实; 但这个事实本身需要用概率论的理论 (大数定律) 来证明. 用一个需要被证明的结论来作定义, 这在逻辑上不够严谨.

知乎@余命数

其次, 我们永远无法做无穷多次试验以求那个极限值. 任何实际的频率都只是概率的一个**近似值**.

最后, 也是对于实践最不利的, 对于很多现实问题, 试验根本不具备重复进行的可能性. 例如, 明天会下雨的概率或某支球队赢得下届世界杯的概率, 这些事件都是一次性的, 无法重复一万个明天或一万个下届世界杯来统计频率.

由于这些缺陷, 统计定义不能作为概率论的逻辑基础. 于是数学家们尝试从另一个角度出发.

知乎@余命数

10.1.3.2 古典概型及其不足

历史上最早取得成功的数学定义是**古典概型**, 它避开了重复试验, 转而分析试验结果的内在对称性. 它基于一个非常直观的思想. 如果一个试验有 n 个**等可能**的基本结果, 而事件 A 包含了其中的 m 个结果, 那么事件 A 发生的概率就是 $P(A)=\frac{m}{n}$. 这种定义简洁、易于理解, 并且在我们处理像掷骰子、抽扑克牌这类理想化问题时非常有效.

然而, 古典概型的应用范围同样有限. 它要求样本空间是有限的, 并且依赖于等可能性这个无法被定义和验证的前提. 对于一枚质地不均匀的硬币, 或一个包含无穷结果的试验, 如从区间 $[0,1]$ 内随机取一个数, 古典概型便无能为力.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

10.1.3.3 概率的公理化定义

为了克服上述定义的局限, 我们需要一个更具普适性的概率定义. 这个定义从一些最基本、不证自明的公理出发, 用逻辑演绎的方式构建起整个概率论的体系. 这就是苏联数学家安德雷·柯尔莫戈洛夫在 1933 年提出的**概率公理化定义**.

它不关心概率的数值从何而来(可以来自频率统计, 也可以来自对称性假设), 只关心这些数值必须遵守哪些基本的规则.

定义 10.10 (概率的公理化定义) 设 E 是一个随机试验, Ω 是其样本空间. 对于 Ω 中的每一个事件 A, 赋予一个实数, 记为 $P(A)$, 称为事件 A 的**概率**. 如果这个赋值函数 $P(\cdot)$ 满足以下三条公理, 则称之为一个**概率测度**.

1. **非负性公理**对于每一个事件 A, 有 $P(A) \geq 0$.
2. **规范性公理**必然事件的概率为 1, 即 $P(\Omega) = 1$.
3. **可加性公理**若 $A_1, A_2, \ldots$ 是一列**两两互斥**的事件, 即对于任意 $i \neq j$, 都有 $A_i \cap A_j = \emptyset$, 则有

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots \tag{10-1}$$

对于有限个互斥事件, 此公理同样成立; $P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n)$.

柯尔莫戈洛夫的定义规定了任何一种合理的概率度量**必须遵守的根本法则**.

首先, "对于…每一个事件 A, 赋予一个实数…记为 $P(A)$"这句话, 实质上定义了一个**函数**. 我们在高中阶段所熟悉的函数, 其定义域和值域通常都是数的集合, 例如 $f(x) = x^2$. 而这里的概率 $P(\cdot)$ 是一个更广义的函数. 它的**定义域**是样本空间 Ω 的所有事件构成的集合, 函数的自变量是一个**集合**; 其**值域**则是实数集 $\mathbb{R}$(根据公理, 实际上是区间 $[0, 1]$). 因此, $P(A)$ 的写法与 $f(x)$ 完全对应, 它表示将事件 A 这个输入, 映射到它的概率值这个输出.

而概率测度这一术语, 意为测量. 就像长度测量一维线段、面积测量二维图形一样, 概率测度就是一套用来**测量**一个事件发生可能性大小的规则. 任何满足这三条公理的函数 $P(\cdot)$, 都被认为是一种合法的、自洽的测量方式.

10.1.3.4 三条公理的内涵

10.1.3.4.1 非负性 这是最基本的要求. 一个事件发生的可能性, 最低对应概率为 0, 即事件不发生. 此公理为概率设定了下界.

10.1.3.4.2 规范性 这是为测量设定一个**单位尺度**. 规范性的意思就是标准化. 将必然发生这件事的确定性程度规定为 1(或者说 100%). 于是, 所有其他事件的概

知乎@余命数

率都成为了相对于这个完全确定的基准的一个比例, 自然地落在了 0 和 1 之间. 就像将 1 米定义为一个标准长度一样, 这里将完全的确定性标定为 1.

知乎@余命数

10.1.3.4.3 可加性 对于一系列两两互斥的事件 A_i, 其并集的概率等于它们各自概率的和. 这是三条公理中最实用的一条, 它将整体的概率与局部的概率联系起来.

概率的公理化定义, 实质上是为可能性这个模糊概念建立了一个数学模型. 它声称, 任何一个函数, 只要它取非负值, 把总的可能性校准为 1, 并且允许将互斥可能性的概率进行相加, 也就是满足三条公理, 那么它就是一个有效的概率模型.

这三条公理形式简单, 却是概率论的基础; 所有我们熟知的概率性质都可以从它们推导出来.

性质 不可能事件的概率为零.

$$P(\emptyset) = 0$$

证明 考虑必然事件 Ω 与不可能事件 $\emptyset$. 二者互斥 ($\Omega \cap \emptyset = \emptyset$), 且并集为 $\Omega \cup \emptyset = \Omega$. 由公理 3(可加性),

$$P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset).$$

代入 $\Omega \cup \emptyset = \Omega$, 上式化为 $P(\Omega) = P(\Omega) + P(\emptyset)$; 再由公理 2, $P(\Omega) = 1$, 得

$$1 = 1 + P(\emptyset),$$

知乎@余命数 知乎@余命数 知乎@余命数

故 $P(\emptyset) = 0$. □

知乎@余命数 知乎@余命数 乎@余命数

反之则不一定成立. 概率为 0 的事件不一定是不可能事件. 例如, 在从区间 $[0,1]$ 上随机取一个数的试验中, 取到数字 0.5 的概率是 0, 但它显然是一个可能发生的结果.

性质 对于任意事件 A, 其对立事件 $\bar{A}$ 的概率为

$$P(\bar{A}) = 1 - P(A)$$

这个性质十分常用, 因为直接计算事件 A 的概率往往比较复杂, 而计算其对立事件 $\bar{A}$ 的概率则比较简单.

证明 对任意事件 A 与其对立事件 $\bar{A}$, 二者互斥 ($A \cap \bar{A} = \emptyset$) 且完备 ($A \cup \bar{A} = \Omega$). 由公理 3(可加性),

$$P(A \cup \bar{A}) = P(A) + P(\bar{A}).$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

又 $A \cup \bar{A} = \Omega$, 由公理 2, $P(A \cup \bar{A}) = P(\Omega) = 1$, 故

$$1 = P(A) + P(\bar{A}),$$

移项即得 $P(\bar{A}) = 1 - P(A)$. □

性质 对于任意事件 A, 其概率介于 0 和 1 之间.

$$0 \leq P(A) \leq 1$$

证明 由公理 1(非负性)直接得 $P(A) \geq 0$.

为证 $P(A) \leq 1$, 套用刚证得的对立事件概率性质

$$P(A) = 1 - P(\bar{A}).$$

公理 1 保证 $P(\bar{A}) \geq 0$, 故

$$P(A) = 1 - P(\bar{A}) \leq 1 - 0 = 1,$$

即 $0 \leq P(A) \leq 1$. □

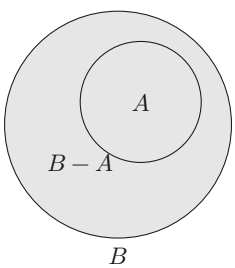
性质 若事件 A 包含于事件 B ($A \subseteq B$), 则事件 A 的概率不大于事件 B 的概率.

$$P(A) \leq P(B)$$

并且, 事件差的概率为

$$P(B - A) = P(B) - P(A)$$

证明 当 $A \subseteq B$ 时, 事件 B 可分解为两个互斥部分的并集. 这两个部分分别是事件 A 与事件 $B - A$, 即属于 B 但不属于 A 的部分.



知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$B = A \cup (B - A)$, 且 $A \cap (B - A) = \emptyset$. 由公理 3 的可加性对此不交并分解,

$$P(B) = P(A) + P(B - A),$$

移项得事件差的概率公式 $P(B - A) = P(B) - P(A)$. 又由公理 1（非负性）有 $P(B - A) \geq 0$, 代入上式即 $P(B) - P(A) \geq 0$, 故

$$P(A) \leq P(B).$$

□

性质 对于任意两个事件 A 和 B, 它们并集的概率为

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

证明 为套用公理 3, 将事件 $A \cup B$ 分解为互斥事件的并. 由韦恩图可见 $A \cup B$ 可写成三个互斥区域 $A - B$、$B - A$、$A \cap B$ 的并集; 更简洁的分解是

$$A \cup B = A \cup (B - A),$$

其中 A 与 $B - A$ 互斥. 由公理 3,

$$P(A \cup B) = P(A) + P(B - A).$$

另一方面, $B = (A \cap B) \cup (B - A)$ 且 $A \cap B$ 与 $B - A$ 互斥, 故 $P(B) = P(A \cap B) + P(B - A)$, 即 $P(B - A) = P(B) - P(A \cap B)$. 代回上式即得

$$P(A \cup B) = P(A) + (P(B) - P(A \cap B)) = P(A) + P(B) - P(A \cap B).$$

□

这个公式表明, 当简单地将两个事件的概率相加时, 我们把它们**共同发生**的部分 $(A \cap B)$ 计算了两遍, 因此必须减去一次. 这个思想可以推广到任意多个事件, 并引出著名的**容斥原理**. 例如, 对于三个事件 A, B, C, 有

$$P(A \cup B \cup C) = \sum P(A) - \sum P(AB) + P(ABC)$$

其中 $\sum P(A) = P(A) + P(B) + P(C)$, $\sum P(AB) = P(AB) + P(AC) + P(BC)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

10.2 经典概率问题

我们这一节来看两个著名的古典概率问题, 即生日问题与错排问题. 为求解它们, 我们会用到上一节建立的全部工具, 包括样本空间的建模、对立事件与容斥原理.

注 10.1 (集合符号提示)
处理多个事件的关系时, 逐一写出 $A_1 \cup A_2 \cup \cdots \cup A_n$ 会非常繁琐. 为此我们引入一套简洁的符号, 用来表示多个事件的并集与交集.
已知 $A \cup B$ 表示 A 或 B 发生, $A \cap B$ 表示 A 与 B 同时发生. 我们把它推广到 n 个事件 $A_1, A_2, \ldots, A_n$, 于是得到以下记号.

- 大并集符号 $\bigcup$

$$\bigcup_{i=1}^{n} A_i \quad \text{读作从 i 等于 1 到 n, } A_i \text{ 的并集}$$

它等价于 $A_1 \cup A_2 \cup \cdots \cup A_n$, 所代表的事件是 $A_1, A_2, \ldots, A_n$ 中**至少有一个**发生.

- 大交集符号 $\bigcap$

$$\bigcap_{i=1}^{n} A_i \quad \text{读作从 i 等于 1 到 n, } A_i \text{ 的交集}$$

它等价于 $A_1 \cap A_2 \cap \cdots \cap A_n$, 所代表的事件是 $A_1, A_2, \ldots, A_n$ **全部都**发生.

10.2.1 生日问题 *

我们先看一个经典的问题.

例题 10.10 (生日问题) 在一个有 n 个人的房间里 ($n \leq 365$), 假设每个人的生日等可能地分布在一年 365 天中的任何一天, 且不考虑闰年. 求至少有两个人生日相同的概率是多少? 计算当 $n = 23$ 时, 这个概率大约是多少.

解 直接计算至少有两人生日相同的概率颇为困难, 因为它混合了恰好两人相同、恰好三人相同、两对不同的人生日相同等多种情形. 于是我们转求它的对立事件 A, 即"所有人的生日都不同".

样本空间 Ω 是 n 个人所有可能的生日组合, 每人都有 365 种可能, 故 $\text{card}(\Omega) = 365^n$. 事件 A 的样本点数相当于从 365 天中选出 n 个不同生日并分配给 n 个人, 即

$$\text{card}(A) = P_{365}^{n} = 365 \times 364 \times \cdots \times (365 - n + 1),$$

因而

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{P_{365}^{n}}{365^n} = \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{365 - n + 1}{365}.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

所求的至少有两人生日相同的概率为

$$P(\bar{A}) = 1 - P(A) = 1 - \prod_{k=0}^{n-1} \left(1 - \frac{k}{365}\right).$$

代入 $n = 23$, 由 $P(A) = \frac{P_{365}^{23}}{365^{23}} \approx 0.4927$ 得

$$P(\bar{A}) = 1 - P(A) \approx 1 - 0.4927 = 0.5073.$$

仅 23 人的房间里, 至少两人生日相同的概率已超过 50%. 这是一个反直觉的结果, 其根源在于配对数量随 n 二次增长.

我们常感到意外, 因为直觉倾向于**线性思考**. 它容易把问题理解为我与另外 22 个人生日相同的概率, 或者简单地比较人数 23 与天数 365, 从而认为概率应当很小.

问题的真正要点是**任意两个人**生日相同的概率. 任意二字把组合学引入问题. 对 n 个人, 总共有 $\binom{n}{2} = \frac{n(n-1)}{2}$ 个可能的配对. 当 $n = 23$ 时, 总共有

$$\binom{23}{2} = \frac{23 \times 22}{2} = 253 \text{ 个配对.}$$

知乎@余命数

于是, 一个只有 23 人的房间里, 存在 **253** 个可能出现生日相同的配对. 单个配对生日相同的概率很低, 但配对数如此之多, 至少出现一对相同的概率便相当高. 直觉没有捕捉到由人数增加导致的**配对数量的二次方增长**. 从上面的解法可见, 直接处理至少一词非常复杂; 转向其对立事件(所有人生日都不同), 则使问题简化.

事件 A(所有人生日都不同)的概率可以写成一个连乘积.

$$P(A) = 1 \times \left(1 - \frac{1}{365}\right) \times \left(1 - \frac{2}{365}\right) \times \ldots$$

知乎@余命数

每个乘上的因子都小于 1. 多个小于 1 的数相乘, 结果会**迅速减小**. 这种衰减是指数性的, 比直觉中的线性减少快得多. 因此 $P(A)$ 迅速从 1 降到 0.5 以下, 其对立事件 $P(\bar{A})$ 便迅速攀升超过 0.5.

知乎@余命数

生日问题的模型在计算机科学中有一个重要应用, 即**哈希碰撞**.

哈希函数把任意长度的数据, 如一个文件、一张图片映射为固定长度的、看似随机的字符串, 我们称这个字符串为**哈希值**, 它相当于数据的数字指纹. **碰撞**则是指两个**不同**的原始数据经过哈希函数计算后得到**完全相同**的哈希值.

知乎@余命数

哈希值的可能空间 (对应生日问题中的 365 天) 通常非常巨大, 例如 2^{128} 或 2^{256}. 然而生日问题表明, 我们并不需要 2^{128} 个不同文件才会看到明显的碰撞风险. 按生日问题的近似, 当文件数量达到 $\sqrt{2^{128}} = 2^{64}$ 量级时, 碰撞概率已经不可忽视.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这一原理是著名的**生日攻击**的基础. 攻击者可以利用这一概率原理, 以远低于暴力破解的计算量伪造数字签名或寻找加密算法的弱点. 因此, 设计安全的密码系统时必须考虑生日问题所揭示的碰撞风险.

10.2.2 错排问题

知乎@余命数

例题 10.11 (德·蒙莫尔问题) n 位客人参加晚宴, 他们都把自己的帽子交给了衣帽间的服务员. 晚宴结束后, 服务员由于匆忙, 把 n 顶帽子随机地还给了每位客人. 求没有任何一位客人拿到自己帽子的概率 p_n. 当 $n \to \infty$ 时, 这个概率收敛到何值?

这是一个经典的概率与组合问题. 我们从古典概型出发求解它.

先确定样本空间的总大小. 服务员把 n 顶不同的帽子随机还给 n 位客人, 总的排列方式有 $n!$ 种.

$$\text{card}(\Omega) = n!$$

接下来, 我们需要计算事件 A 所包含的样本点数, 其中事件 A 表示没有任何一位客人拿到自己的帽子. 这是一个复杂的计数问题, 因此我们先用严格的方式定义这一类排列.

定义 10.11 (错排与错排数) 设一个 n 个元素的排列中没有任何元素出现在它原来的位置上, 我们就称它为一个 n 元**错排**. 这类排列的总数称为**错排数**, 记为 D_n 或 $!n$.
例如, 当 $n = 3$ 时, 集合 $\{1, 2, 3\}$ 的全排列共有 $3! = 6$ 种, 其中错排是 $(2, 3, 1)$ 与 $(3, 1, 2)$, 因此 $D_3 = 2$.

根据上面的定义, 我们要求的样本点数就是 D_n. 于是原概率问题化为求 D_n 的通项公式.

$$p_n = P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{D_n}{n!}$$

我们用两种方法求 D_n.

解

知乎@余命数

直接计算所有人都错位的排列数 D_n 很困难. 我们先算其对立面, 即至少有一人拿对自己帽子的排列数, 再用总排列数 $n!$ 减去它.

令 A_i 为"第 i 位客人拿到自己帽子"的排列所构成的集合 $(i = 1, \dots, n)$, 我们需要计算 $|A_1 \cup A_2 \cup \cdots \cup A_n|$. 由容斥原理,

$$|A_1 \cup \cdots \cup A_n| = \sum |A_i| - \sum_{i<j} |A_i \cap A_j| + \cdots + (-1)^{n-1} |A_1 \cap \cdots \cap A_n|.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其中 $|A_i|$ 表示第 i 位客人拿对帽子而其余 $n-1$ 人任意排, 有 $(n-1)!$ 种, 这样的集合有 $\binom{n}{1}$ 个; $|A_i \cap A_j|$ 表示第 i,j 位客人拿对帽子而其余 $n-2$ 人任意排, 有 $(n-2)!$ 种, 这样的集合有 $\binom{n}{2}$ 个. 以此类推, k 位指定客人拿对帽子的排列数为 $(n-k)!$, 共有 $\binom{n}{k}$ 组.

由此至少一人拿对的总数为

$$\binom{n}{1}(n-1)!-\binom{n}{2}(n-2)!+\cdots+(-1)^{n-1}\binom{n}{n}(n-n)!$$

$$=\frac{n!}{1!(n-1)!}(n-1)!-\frac{n!}{2!(n-2)!}(n-2)!+\cdots=\frac{n!}{1!}-\frac{n!}{2!}+\frac{n!}{3!}-\cdots+\frac{(-1)^{n-1}n!}{n!},$$

故错排数 D_n 为

$$\begin{aligned}D_n&=n!-\left(\frac{n!}{1!}-\frac{n!}{2!}+\cdots+\frac{(-1)^{n-1}n!}{n!}\right)\\&=\frac{n!}{0!}-\frac{n!}{1!}+\frac{n!}{2!}-\cdots+\frac{(-1)^n n!}{n!}=n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}.\end{aligned}$$

分析第 n 个元素的位置还可建立 D_n 的递推关系. 考虑元素 n, 在一个错排中它不能在第 n 位, 只能在其余 $n-1$ 个位置之一, 设其在第 k 位 $(k\neq n)$, 这有 $n-1$ 种选择. 对此选择再就元素 k 的位置分两种情形. 若元素 k 恰好放在第 n 位, 相当于 k 与 n 互换位置, 剩下 $n-2$ 个元素需完全错排, 方法数为 D_{n-2}; 若元素 k 不在第 n 位, 则 n 已占第 k 位, 对剩下 $n-1$ 个元素与 $n-1$ 个位置, 其中元素不含 n、位置不含 k, 问题化为将元素 $j\in\{1,\ldots,n-1\}, j\neq k$ 放在非 j 号位且将元素 k 放在非 n 号位, 等价于一个 $n-1$ 元错排, 方法数为 D_{n-1}. 由乘法与加法原理得二阶递推关系

$$D_n=(n-1)(D_{n-1}+D_{n-2})\quad(n\geq 3),$$

初始值 $D_1=0, D_2=1$. 由此递推亦可推出通项公式.

回到最初的概率问题, 代入 D_n 的通项公式,

$$p_n=\frac{D_n}{n!}=\frac{n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}}{n!}=\sum_{k=0}^{n}\frac{(-1)^k}{k!}.$$

注 10.2

由 $p_n=\sum_{k=0}^{n}\frac{(-1)^k}{k!}$ 可直接算出前几项, $p_2=0.5$、$p_3\approx 0.333$、$p_4=0.375$、$p_5=0.3666\ldots$, 数值上很快接近 0.37. 通项公式与递推关系解决了有限 n 的错排计数; 极限常数见下.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

错排概率极限 $1/e$

当 $n\to\infty$ 时, 部分和 $\sum_{k=0}^{n}\frac{(-1)^k}{k!}$ 收敛到 e^x 在 $x=-1$ 处的泰勒级数.

$$\lim_{n\to\infty}p_n=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}=e^{-1}=\frac{1}{e}\approx 0.36788.$$

知乎@余命数

因而大规模宴会中无人拿对帽子的概率趋于常数 $1/e$. 级数与余项的严格语言见第16章.

10.2.2.1 错排数的性质与公式

上述例题给出了关于错排数的一些重要结论, 我们把它汇总如下.

定理 10.2 (错排数的通项公式) n 元错排数 D_n 的显式通项公式为

$$D_n=n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}=\frac{n!}{0!}-\frac{n!}{1!}+\frac{n!}{2!}-\cdots+(-1)^n\frac{n!}{n!}\tag{10-2}$$

证明 设全集为 n 个元素的全排列集合, 其大小为 $n!$. 对每个 $i=1,2,\dots,n$, 令 A_i 表示“第 i 个元素仍在第 i 个位置”的排列集合. 错排就是不属于任何 A_i 的排列, 因而

$$D_n=n!-\left|A_1\cup A_2\cup\cdots\cup A_n\right|.$$

若指定 r 个元素固定在原位, 其余 $n-r$ 个元素可任意排列, 因此对应排列数为 $(n-r)!$. 这样的 r 个固定元素集合有 $\binom{n}{r}$ 种. 由容斥原理,

$$\left|A_1\cup\cdots\cup A_n\right|=\sum_{r=1}^{n}(-1)^{r-1}\binom{n}{r}(n-r)!.$$

所以

知乎@余命数

$$D_n=n!+\sum_{r=1}^{n}(-1)^r\binom{n}{r}(n-r)!=\sum_{r=0}^{n}(-1)^r\frac{n!}{r!}=n!\sum_{r=0}^{n}\frac{(-1)^r}{r!}.$$

□

定理 10.3 (错排数的递推关系) 错排数 D_n 满足以下两个等价的递推关系, 初始值为 $D_1=0, D_2=1$.

1. **(二阶递推)** $D_n=(n-1)(D_{n-1}+D_{n-2})$, 　对于 $n\geq 3$.
2. **(一阶递推)** $D_n=nD_{n-1}+(-1)^n$, 　对于 $n\geq 2$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 对二阶递推, 在一个 n 元错排中元素 n 不能在第 n 位, 设它被放到第 k 位 $(k \neq n)$, 这有 $n-1$ 种选择. 若元素 k 放在第 n 位, 则 k 与 n 形成互换, 剩下 $n-2$ 个元素必须错排, 有 D_{n-2} 种方法; 若元素 k 不放在第 n 位, 则删去第 k 位和元素 n 后, 剩余问题等价于一个 $n-1$ 元错排, 有 D_{n-1} 种方法. 因此

$$D_n = (n-1)(D_{n-1} + D_{n-2}) \quad (n \geq 3).$$

一阶递推由通项公式直接导出:

$$D_n = n! \sum_{r=0}^{n} \frac{(-1)^r}{r!} = n\,(n-1)! \sum_{r=0}^{n-1} \frac{(-1)^r}{r!} + (-1)^n = nD_{n-1} + (-1)^n.$$

初始值 $D_1 = 0, D_2 = 1$ 可直接列举得到. □

当 n 较小时, 使用递推关系通常比展开通项公式更快捷. 例如 $D_3 = 2(D_2 + D_1) = 2(1+0) = 2, D_4 = 3(D_3 + D_2) = 3(2+1) = 9$.

极限概率 $D_n/n! \to 1/e$

性质 (极限概率) n 个元素的全排列中, 是一个错排的概率 $p_n = D_n/n!$ 收敛于 $1/e$.

$$\lim_{n\to\infty} \frac{D_n}{n!} = \frac{1}{e} \approx 0.36788 \tag{10-3}$$

证明 由错排数通项公式,

$$\frac{D_n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

右端是指数函数泰勒级数 $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ 在 $x = -1$ 处的前 n 项部分和. 因此当 $n \to \infty$ 时,

$$\frac{D_n}{n!} \to \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} = e^{-1} = \frac{1}{e}.$$

□

最佳整数近似 $D_n = \lfloor n!/e + 1/2 \rfloor$

注 10.3
小 n 时用递推或通项有限和;现用 e^{-1} 的交错级数余项说明四舍五入公式.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

性质 (最佳整数近似) 对于所有正整数 $n \geq 1$, D_n 是最接近 $\frac{n!}{e}$ 的整数. 因此 D_n 可以通过四舍五入计算得到.

$$D_n = \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor.$$

这个性质为快速估算甚至精确计算 D_n 提供了一个极为便捷的方法.

证明 由 e^{-1} 的交错级数展开与错排数通项公式,

$$\frac{n!}{e} - D_n = n! \sum_{k=n+1}^{\infty} \frac{(-1)^k}{k!}.$$

交错级数余项的绝对值小于首个被舍去项, 所以

$$\left| \frac{n!}{e} - D_n \right| < \frac{n!}{(n+1)!} = \frac{1}{n+1} \leq \frac{1}{2}.$$

因此整数 D_n 与 $\frac{n!}{e}$ 的距离小于 $1/2$, 它就是距离 $\frac{n!}{e}$ 最近的整数, 等价地

$$D_n = \left\lfloor \frac{n!}{e} + \frac{1}{2} \right\rfloor.$$

□

现在看一些错排理论在具体问题中的应用.

例题 10.12 (信件与信封问题) 将 4 封写好的不同信件, 随机装入写好对应地址的 4 个信封中, 每个信封恰好装一封信. 求 4 封信全部装错信封的装法共有多少种?

解 这是最直接的错排问题. 元素是 4 封信, 位置是 4 个信封, 所求即 4 个元素的错排数 D_4. 代入错排数通项公式, 当 $n = 4$ 时

$$\begin{aligned} D_4 &= 4! \left(\frac{1}{0!} - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \right) \\ &= 24 \left(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} \right) \\ &= 24 \left(\frac{1}{2} - \frac{1}{6} + \frac{1}{24} \right) \\ &= 12 - 4 + 1 = 9. \end{aligned}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

或由递推公式, 已知 $D_2=1, D_3=2$,

$$D_4=(4-1)(D_3+D_2)=3\times(2+1)=9.$$

故共有 9 种全部装错的方法.

公式给出答案, 而公式背后正是容斥原理的一次直接应用. 为看清来路, 我们不妨手算 D_4.

总排列数,即样本空间大小为 $4!=24$. 记 A_i 为第 i 封信装对信封 $(i=1,2,3,4)$, 由容斥原理

$$D_4=24-|A_1\cup A_2\cup A_3\cup A_4|,$$

右端 $|A_1\cup\cdots\cup A_4|$ 即至少一封装对的情况数, 按容斥原理逐项展开, 得

- 任选 1 封装对, 其余 3 封任意排, 共 $\binom{4}{1}\times 3!=4\times 6=24$ 种.
- 指定 2 封装对时在上一项中被重复计入, 减去 $\binom{4}{2}\times 2!=6\times 2=12$ 种.
- 指定 3 封装对时前两层分别被加、减各 3 次而净计入 0 次, 补回 $\binom{4}{3}\times 1!=4\times 1=4$ 种.
- 四封全装对累计被计入 $4-6+4=2$ 次, 应保留 1 次, 再减 $\binom{4}{4}\times 0!=1$ 种.

于是至少一封装对的情况数为

$$\binom{4}{1}3!-\binom{4}{2}2!+\binom{4}{3}1!-\binom{4}{4}0!=24-12+4-1=15,$$

从而 $D_4=24-15=9$, 与公式计算吻合.

例题 10.13 (枚举法) 三位同学甲、乙、丙, 分别坐在 1 号、2 号、3 号座位上. 现在让他们重新就座, 要求每个人都不能坐在自己原来的位置上. 请问有多少种不同的坐法?

解 此为 $n=3$ 的错排问题, 总数很少, 用枚举法直接列出即可. 对甲、乙、丙三人作全排列, 括号内按顺序表示坐在 1、2、3 号座位上的人.

1. (甲, 乙, 丙)$\to$ 所有人都坐对了.
2. (甲, 丙, 乙)$\to$ 甲坐对了.
3. (乙, 甲, 丙)$\to$ 丙坐对了.
4. (乙, 丙, 甲)$\to$ **没有任何人坐对, 这是一个错排**.
5. (丙, 甲, 乙)$\to$ **没有任何人坐对, 这是另一个错排**.
6. (丙, 乙, 甲)$\to$ 乙坐对了.

逐项核对, 仅第 4、第 5 种坐法满足“每个人都不能坐在自己原来的位置上”, 故共有 2 种不同的坐法, 验证了 $D_3=2$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 10.14 (四球四盒问题) 有 4 个写着号码 1, 2、3, 4 的球, 和 4 个对应号码的盒子. 现将 4 个球随机放入 4 个盒子中, 每盒一个. 求至少有一个球放进与自己号码相同的盒子的情况有多少种?

解 所求为至少一个球放对的情况数. 直接分类恰好一个对、恰好两个对等会很繁琐, 转而用容斥原理直接计算并事件. 令 A_i 为"第 i 号球放进第 i 号盒"的事件, 则所求即 $|A_1 \cup A_2 \cup A_3 \cup A_4|$. 由容斥原理,

$$\begin{aligned}|A_1 \cup A_2 \cup A_3 \cup A_4| &= \binom{4}{1}3! - \binom{4}{2}2! + \binom{4}{3}1! - \binom{4}{4}0! \\ &= 4 \times 6 - 6 \times 2 + 4 \times 1 - 1 \times 1 \\ &= 24 - 12 + 4 - 1 = 15,\end{aligned}$$

其中每一项 $\binom{4}{k}(4-k)!$ 表示选定 k 个球放对、其余 $4-k$ 个球任意放的取法数. 故至少一个球放对有 15 种.

所有球都放错的情况数 D_4 则是 $24 - 15 = 9$ 种.

例题 10.15 一位粗心的药剂师手边有为 4 位病人准备的 4 种不同药品. 他随机地将这 4 种药品分发给了这 4 位病人, 每人一种. 请问, 恰好只有两位病人拿到了正确的药品的情况有多少种?

解 此为部分错排问题, 可分解为两个独立步骤, 再用乘法原理相乘.

先从 4 位病人中选出拿到正确药品的 2 位, 这是组合问题, 选法数为 $\binom{4}{2} = \frac{4\times3}{2\times1} = 6$ 种. 剩下 2 位病人则需完全错排. 设他们为甲、乙, 对应药品为药 A、药 B, 要让两人都拿错, 只能甲拿药 B 且乙拿药 A, 故 $D_2 = 1$ 种.

由乘法原理将两步方法数相乘,

$$\begin{aligned}\text{总数} &= (\text{选出 2 个正确病人的方法数}) \times (\text{剩下 2 个病人错排的方法数}) \\ &= \binom{4}{2} \times D_2 = 6 \times 1 = 6.\end{aligned}$$

故恰好只有两位病人拿到正确药品的情况有 6 种.

例题 10.16 书架上有《红楼梦》《西游记》《三国演义》《水浒传》四本书, 按从左到右的顺序排好. 现将它们取下后随机放回, 求恰好有一本书放回原位的概率.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解 此题结合古典概型与部分错排. 将 4 本不同的书随机放回, 样本空间总数为

$$\text{card}(\Omega) = 4! = 24.$$

事件恰好有一本书放回原位可分解为两步. 先从 4 本书中选出 1 本放回原位, 有 $\binom{4}{1} = 4$ 种选法, 再让其余 3 本完全错排($D_3 = 2$ 种). 由乘法原理,

$$\text{card}(A) = \binom{4}{1} \times D_3 = 4 \times 2 = 8.$$

代入古典概型公式 $P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$,

$$P(\text{恰好一本对}) = \frac{8}{24} = \frac{1}{3}.$$

故恰好有一本书放回原位的概率是 $\frac{1}{3}$.

10.3 条件概率

到目前为止, 我们计算概率时总是站在对整个试验结果一无所知的立场上. 现实中, 我们常常已经掌握关于试验结果的部分信息. 例如掷完骰子后有人瞥见点数为偶数, 体检后化验单上写着某项指标异常. 新的信息会改变我们对事件可能性的判断. 这一节研究的问题是, 在已知事件 A 发生的前提下, 事件 B 发生的概率该如何计算.

10.3.1 条件概率的定义

我们先看一个引例. 掷一颗骰子, 样本空间为 $\Omega = \{1, 2, 3, 4, 5, 6\}$. 记事件 B 为点数大于 3, 显然

$$P(B) = \frac{\text{card}(B)}{\text{card}(\Omega)} = \frac{3}{6} = \frac{1}{2}.$$

现在增加一条信息. 有人提前表明点数为偶数. 记这个事件为 $A = \{2, 4, 6\}$. 在这条信息之下, 点数大于 3 的概率需要重新计算.

信息改变了问题的格局. 既然 A 已经发生, 试验结果只可能落在 $\{2, 4, 6\}$ 之中, 原来的样本空间 Ω 实际上已经**收缩**为 A. 在这三个等可能的结果里, 满足点数大于 3 的只有 4 和 6 两个. 所以我们直观地认为, 所求概率应该是 $\frac{2}{3}$, 它大于原来的 $\frac{1}{2}$.

把这个过程用符号重写一遍. 在古典概型下, “在 A 发生的条件下 B 发生”的样本点, 恰好构成交集 $A \cap B$; 而新的样本空间就是 A 本身. 于是

$$\text{所求概率} = \frac{\text{card}(A \cap B)}{\text{card}(A)} = \frac{\text{card}(A \cap B)/\text{card}(\Omega)}{\text{card}(A)/\text{card}(\Omega)} = \frac{P(A \cap B)}{P(A)}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

最右端的表达式只用到了概率, 不再依赖古典概型的计数结构. 这提示我们把这个比值直接当作条件概率的定义.

定义 10.12 (条件概率) 设 A, B 是样本空间 Ω 中的两个事件, 且 $P(A) > 0$. 称

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

为**在事件 A 发生的条件下事件 B 发生的条件概率**.

定义中要求 $P(A) > 0$ 是自然的. 如果 A 根本不可能发生, 或者它以概率零发生, 那么已知 A 发生这个前提便失去意义. 读法上, $P(B \mid A)$ 读作在 A 发生的条件下 B 的概率, 竖线之后的部分是条件.

我们可以从两个等价的角度理解条件概率, 它们随时切换会带来便利.

- **收缩的样本空间**. 已知 A 发生后, 样本空间从 Ω 缩小为 A, 各样本点的相对比例保持不变, 再重新归一化.
- **概率的比值**. 直接用公式 $P(B \mid A) = P(AB)/P(A)$ 计算, 其中 $P(AB)$ 与 $P(A)$ 都在原样本空间中求得.

例题 10.17 掷一颗骰子, 记 A 为点数为偶数, B 为点数大于 3. 用定义验证引例中的直观结果.

解 有 $A = \{2, 4, 6\}$, $B = \{4, 5, 6\}$, 于是 $A \cap B = \{4, 6\}$.

$$P(A) = \frac{3}{6} = \frac{1}{2}, \qquad P(A \cap B) = \frac{2}{6} = \frac{1}{3}.$$

由条件概率的定义,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3},$$

与引例中收缩样本空间的算法一致.

例题 10.18 (两个孩子的家庭) 某家庭有两个孩子. 假设男孩与女孩的出生机会均等, 四个结果(依长幼顺序记为男男、男女、女男、女女)等可能. (1) 已知这个家庭至少有一个男孩, 求两个孩子都是男孩的概率. (2) 已知老大是男孩, 求两个孩子都是男孩的概率.

解 记 $B = \{(\text{男}, \text{男})\}$ 为两个孩子都是男孩这一事件.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

(1) 记 A_1 为至少有一个男孩, 则

$$A_1 = \{(男, 男), (男, 女), (女, 男)\}, \qquad P(A_1) = \frac{3}{4}.$$

由于 $B \subseteq A_1$, 有 $B \cap A_1 = B$, 故 $P(B \cap A_1) = \frac{1}{4}$. 所以

$$P(B \mid A_1) = \frac{P(B \cap A_1)}{P(A_1)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

(2) 记 A_2 为老大是男孩, 则 $A_2 = \{(男, 男), (男, 女)\}$, $P(A_2) = \frac{1}{2}$. 所以

$$P(B \mid A_2) = \frac{P(B \cap A_2)}{P(A_2)} = \frac{1/4}{1/2} = \frac{1}{2}.$$

两问的答案不同, 因为两条信息的内容不同. 至少有一个男孩只排除了女女这一种结果, 而老大是男孩还额外排除了女男. 信息越多, 样本空间收缩得越厉害, 概率的改变也越大. 使用条件概率时必须把已知的信息精确地写成事件; 含糊其词会得出错误的答案.

条件概率 $P(B \mid A)$ 中的 A 一旦固定, 它作为 B 的函数, 自身就是一个合法的概率测度.

知乎@余命数

性质 (条件概率是概率测度) 设 $P(A) > 0$. 令 $Q(B) = P(B \mid A)$, 则 $Q(\cdot)$ 满足概率的三条公理.

1. 对任意事件 B, $Q(B) \geq 0$;
2. $Q(\Omega) = 1$;
3. 若 $B_1, B_2, \ldots$ 两两互斥, 则 $Q(B_1 \cup B_2 \cup \ldots) = Q(B_1) + Q(B_2) + \ldots$.

知乎@余命数

证明 (1) 由 $P(A \cap B) \geq 0$ 且 $P(A) > 0$, 得

$$Q(B) = \frac{P(A \cap B)}{P(A)} \geq 0.$$

知乎@余命数

(2) 由于 $\Omega \cap A = A$,

$$Q(\Omega) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1.$$

(3) 若 $B_1, B_2, \ldots$ 两两互斥, 则 $B_1 \cap A, B_2 \cap A, \ldots$ 也两两互斥. 对 P 使用可加性公理,

$$Q\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \frac{P((\bigcup_{i=1}^{\infty} B_i) \cap A)}{P(A)} = \frac{P(\bigcup_{i=1}^{\infty}(B_i \cap A))}{P(A)}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{\sum_{i=1}^{\infty} P(B_i \cap A)}{P(A)} = \sum_{i=1}^{\infty} Q(B_i).$$

□

这条性质带来一批推论. 在条件 A 之下, 对立事件公式、加法公式等全部照旧成立. 例如

$$P(\bar{B} \mid A) = 1 - P(B \mid A),$$

$$P(B \cup C \mid A) = P(B \mid A) + P(C \mid A) - P(B \cap C \mid A).$$

它们与普通概率的用法相同, 只是所有概率都挂在同一个条件 A 上.

10.3.2 乘法公式

把条件概率的定义式交叉相乘, 立即得到一个计算积事件概率的公式.

定理 10.4 (乘法公式) 设 $P(A) > 0$, 则

$$P(A \cap B) = P(A)\, P(B \mid A).$$

对称地, 若 $P(B) > 0$, 则 $P(A \cap B) = P(B)\, P(A \mid B)$.

证明 由条件概率的定义 $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$, 两边乘以 $P(A)$ 即得 $P(A \cap B) = P(A)\, P(B \mid A)$. 对称地交换 A, B 得另一形式. □

要求 A 与 B 同时发生的概率, 可以先让 A 发生(概率 $P(A)$), 再在 A 已经发生的前提下让 B 发生, 后者概率为 $P(B \mid A)$. 这与计数中的乘法原理一脉相承. 一个分两步完成的过程, 总方法数等于各步方法数的乘积; 概率的乘法公式正是这个思想在概率上的投影.

例题 10.19 (不放回抽球) 袋中装有 3 个红球和 2 个白球, 除颜色外完全相同. 不放回地依次取出 2 个球, 求两次都取到红球的概率.

解 记 A_1 为第一次取到红球, A_2 为第二次取到红球.

第一次取球时, 袋中有 3 个红球共 5 个球, 故

$$P(A_1) = \frac{3}{5}.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

在 A_1 已经发生的条件下, 袋中剩下 2 个红球共 4 个球, 故

$$P(A_2 \mid A_1) = \frac{2}{4} = \frac{1}{2}.$$

由乘法公式,

$$P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1) = \frac{3}{5} \times \frac{1}{2} = \frac{3}{10}.$$

用古典概型验证. 从 5 个球中取 2 个, 共有 $\binom{5}{2} = 10$ 种等可能的组合, 其中全是红球的有 $\binom{3}{2} = 3$ 种, 概率为 $\frac{3}{10}$. 两种方法结果一致.

乘法公式可以推广到任意多个事件的积. 要求 n 个事件依次全部发生的概率, 只需沿着时间顺序, 每一步都在此前各步均已发生的条件下计算下一步的概率, 再全部乘起来.

定理 10.5 (乘法公式的链式形式) 设 $A_1, A_2, \ldots, A_n$ 是 n 个事件 ($n \geq 2$), 且 $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, 则

$$\begin{aligned} &P(A_1 \cap A_2 \cap \cdots \cap A_n) \\ &\quad = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1}). \end{aligned}$$

证明 对 n 作数学归纳法.

$n = 2$ 时, 结论就是乘法公式本身.

假设结论对 $n-1$ 个事件成立 ($n \geq 3$). 把 $A_1 \cap \cdots \cap A_{n-1}$ 视为一个事件 C. 由归纳假设成立的前提 $P(A_1 \cap \cdots \cap A_{n-1}) > 0$, 即 $P(C) > 0$, 可以对 C 与 A_n 使用乘法公式.

$$P(C \cap A_n) = P(C)\,P(A_n \mid C).$$

再由归纳假设展开 $P(C) = P(A_1 \cap \cdots \cap A_{n-1})$, 即得

$$\begin{aligned} P(A_1 \cap \cdots \cap A_n) = P(A_1)\,P(A_2 \mid A_1) \cdots P(A_{n-1} \mid A_1 \cap \cdots \cap A_{n-2}) \\ \times P(A_n \mid A_1 \cap \cdots \cap A_{n-1}). \end{aligned}$$

归纳成立, 结论对一切 $n \geq 2$ 成立.

条件 $P(A_1 \cap \cdots \cap A_{n-1}) > 0$ 的用意在于保证链上每一项都有意义. 由于 $A_1 \cap \cdots \cap A_{n-1} \subseteq A_1 \cap \cdots \cap A_{k-1}$ 对任意 $k \leq n$ 成立, 由概率的单调性, $P(A_1 \cap \cdots \cap A_{n-1}) > 0$ 蕴含 $P(A_1 \cap \cdots \cap A_{k-1}) > 0$, 故每个条件概率 $P(A_k \mid A_1 \cap \cdots \cap A_{k-1})$ 都有定义. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 10.20　一批产品共 10 件, 其中有 3 件次品. 从中不放回地依次抽取 3 件, 求“前两次抽到正品、第三次抽到次品”的概率.

解 记 A_i 为第 i 次抽到正品 $(i=1,2)$, A_3 为第三次抽到次品. 要求 $P(A_1\cap A_2\cap A_3)$.

沿着抽取顺序逐个写出条件概率.

$$P(A_1)=\frac{7}{10},\qquad P(A_2\mid A_1)=\frac{6}{9},\qquad P(A_3\mid A_1\cap A_2)=\frac{3}{8}.$$

第一个等号是因为 10 件产品中有 7 件正品; 第二个等号是因为抽走 1 件正品后, 剩下 9 件中有 6 件正品; 第三个等号是因为抽走 2 件正品后, 剩下 8 件中有 3 件次品.

由链式乘法公式,

$$P(A_1\cap A_2\cap A_3)=\frac{7}{10}\times\frac{6}{9}\times\frac{3}{8}=\frac{126}{720}=\frac{7}{40}.$$

链式乘法公式把一个整体上难以直接计数的事件, 拆解为一系列每一步都容易判断的条件概率. 抽样问题中带先后顺序的事件, 用它处理格外顺手.

回到本节的出发点. 条件概率刻画了信息对判断的影响. 一般情况下 $P(B\mid A)\neq P(B)$, 知道 A 发生会改变对 B 的估计. 但也存在一类特殊的事件对, A 的发生与否完全不影响 B 的概率, 即 $P(B\mid A)=P(B)$. 这种互不影响的关系在概率论中极为重要, 下一节专门研究它.

10.4 事件的相互独立性

上一节末尾留下一个问题, 即是否存在 $P(B\mid A)=P(B)$ 的情形, 也就是事件 A 的发生完全不改变事件 B 发生的概率. 抛两枚硬币时, 第一枚的结果不会影响第二枚; 一台设备的两个独立部件, 一个的寿命也不会影响另一个. 这种互不影响的关系就是本节要研究的**相互独立**.

10.4.1 两个事件的独立性

从条件概率出发, $P(B\mid A)=P(B)$ 等价于

$$\frac{P(A\cap B)}{P(A)}=P(B)\quad\Longleftrightarrow\quad P(A\cap B)=P(A)\,P(B).$$

最右端的等式有很多好处. 它对 A,B 完全对称; 不要求 $P(A)>0$; 只涉及乘积, 验证起来十分方便. 我们就用它作为独立性的定义.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 10.13 (两个事件的相互独立) 设 A, B 是样本空间 Ω 中的两个事件. 若

$$P(A \cap B) = P(A)\,P(B),$$

则称事件 A 与事件 B **相互独立**, 简称**独立**; 否则称它们**不独立**或**相依**.

当 $P(A) > 0$ 时, 独立性等价于 $P(B \mid A) = P(B)$; 当 $P(B) > 0$ 时, 又等价于 $P(A \mid B) = P(A)$. **一个事件的发生, 不改变另一个事件发生的概率**.

例题 10.21 抛两枚均匀的硬币. 记 A 为第一枚正面朝上, B 为第二枚正面朝上. 判断 A 与 B 是否相互独立.

解 样本空间为 $\{(H,H),(H,T),(T,H),(T,T)\}$, 四个样本点等可能.

$$P(A) = \frac{2}{4} = \frac{1}{2}, \qquad P(B) = \frac{2}{4} = \frac{1}{2}, \qquad P(A \cap B) = P(\{(H,H)\}) = \frac{1}{4}.$$

由于 $P(A \cap B) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(A)\,P(B)$, 故事件 A 与 B 相互独立.

性质 (独立性与对立事件) 若事件 A 与 B 相互独立, 则 A 与 $\bar{B}$, $\bar{A}$ 与 B, $\bar{A}$ 与 $\bar{B}$ 也都相互独立.

证明 只证 A 与 $\bar{B}$ 独立, 其余两组的证法完全相同.

由差事件概率公式, $P(A \cap \bar{B}) = P(A - B) = P(A) - P(A \cap B)$. 代入独立性条件 $P(A \cap B) = P(A)P(B)$:

$$P(A \cap \bar{B}) = P(A) - P(A)\,P(B) = P(A)\,(1 - P(B)) = P(A)\,P(\bar{B}).$$

这正是 A 与 $\bar{B}$ 独立的定义式. □

这条性质符合直觉. 如果 A 的发生影响不了 B, 那么它也影响不了 B 不发生. 在计算中它非常实用, 只要四个事件中有一对独立, 其余三对的乘积公式全部免费成立.

注 10.4 (独立与互斥)
互斥与独立描述的是两个完全不同的维度, 初学者极易混淆.

- **互斥**是事件层面的关系, 它意味着 $A \cap B = \emptyset$, 两个事件不能同时发生. 一个发生了, 另一个就**绝不发生**, 这其实是极强的相互影响.
- **独立**是概率层面的关系, 它要求 $P(A \cap B) = P(A)P(B)$, 一个事件的发生不改变另一个事件的概率.

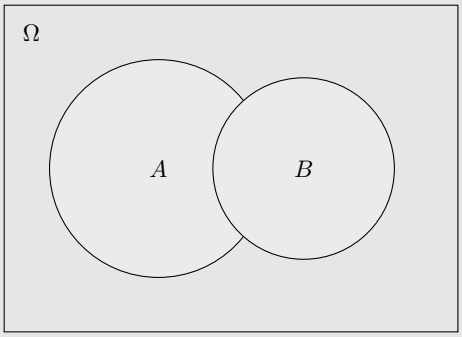
知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

二者在正概率下甚至互相排斥. 若 $P(A)>0, P(B)>0$ 且 A, B 互斥, 则

$$P(A\cap B)=P(\emptyset)=0<P(A)\,P(B),$$

故事件 A, B 必不独立. 反过来, 若 A, B 独立且概率都为正, 则 $P(A\cap B)=P(A)P(B)>0$, 两个事件必定相容.
先看两个极端情形. 不可能事件 $\emptyset$ 与任何事件独立, 因为 $P(\emptyset\cap B)=0=P(\emptyset)P(B)$ 恒成立; 必然事件 Ω 也与任何事件独立, 因为 $P(\Omega\cap B)=P(B)=P(\Omega)P(B)$ 恒成立.

例题 10.22 (人教 A 版必修二 226 页复习参考题 10) 下图是一个古典概型的样本空间 Ω 和事件 A, B 的 Venn 图. 已知样本空间的总样本点数 $\text{card}(\Omega)=24$, 事件 A 包含的样本点数 $\text{card}(A)=12$, 事件 B 包含的样本点数 $\text{card}(B)=8$, 事件 $A\cup B$ 包含的样本点数 $\text{card}(A\cup B)=16$.



(1) 求 $\text{card}(AB)$、$P(AB)$、$P(A\cup B)$ 以及 $P(\bar{A}\bar{B})$.
(2) 判断事件 A 与 B 是否互斥? 是否相互独立?

解 本题把事件的关系与概率的基本性质放进同一个模型中考查.

(1) 计算基本事件的样本点数与概率. 由集合的容斥原理计算事件 A 和 B 交集(积事件)的样本点数,

$$\text{card}(A\cup B)=\text{card}(A)+\text{card}(B)-\text{card}(A\cap B).$$

代入已知数值 $16=12+8-\text{card}(A\cap B)$, 解得积事件 AB, 即 $A\cap B$ 的样本点数 $\text{card}(AB)=4$. 由古典概型定义算出相关概率:

$$P(AB)=\frac{\text{card}(AB)}{\text{card}(\Omega)}=\frac{4}{24}=\frac{1}{6},$$
$$P(A\cup B)=\frac{\text{card}(A\cup B)}{\text{card}(\Omega)}=\frac{16}{24}=\frac{2}{3}.$$

知乎@余命数

至于 $P(\bar{A}\bar{B})$, 由德·摩根定律 $\bar{A}\cap\bar{B}=\overline{A\cup B}$, 它即事件 $A\cup B$ 的对立事件概率,

$$P(\bar{A}\bar{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-\frac{2}{3}=\frac{1}{3}.$$

(2) 判断事件的关系.

互斥性. 互斥要求交集为空, 即 $\text{card}(A\cap B)=0$; 而 (1) 已得 $\text{card}(AB)=4\neq 0$, 故事件 A 与 B **不互斥**.

独立性. 相互独立要求 $P(AB)=P(A)P(B)$.(1) 已给左边 $P(AB)=\frac{1}{6}$; 由古典概型算右边

$$P(A)=\frac{\text{card}(A)}{\text{card}(\Omega)}=\frac{12}{24}=\frac{1}{2},\qquad P(B)=\frac{\text{card}(B)}{\text{card}(\Omega)}=\frac{8}{24}=\frac{1}{3},$$

故 $P(A)P(B)=\frac{1}{2}\times\frac{1}{3}=\frac{1}{6}=P(AB)$ 成立, 事件 A 与 B **相互独立**.

这道题还透露出一个信息. 独立与互斥的判定途径完全不同. 互斥与否看集合是否相交, 独立与否看概率的乘积关系, 两者各算各的账.

独立性在工程中的一个直接应用是评估系统的可靠性.

例题 10.23 (串联与并联系统) 某设备的两个部件各自独立工作, 在一个使用周期内, 部件 1 正常工作的概率为 0.9, 部件 2 正常工作的概率为 0.8. (1) 若两个部件**串联**, 即任何一个失效都会导致系统失效, 求系统正常工作的概率. (2) 若两个部件**并联**, 即只要还有一个正常系统就能运行, 求系统正常工作的概率.

解 记 A_i 为部件 i 正常工作 $(i=1,2)$, 则 $P(A_1)=0.9$, $P(A_2)=0.8$, 且 A_1 与 A_2 相互独立.

(1) 串联系统正常, 当且仅当两个部件都正常, 即事件 $A_1\cap A_2$. 由独立性,

$$P(\text{系统正常})=P(A_1\cap A_2)=P(A_1)\,P(A_2)=0.9\times 0.8=0.72.$$

(2) 并联系统失效, 当且仅当两个部件都失效, 即事件 $\bar{A}_1\cap\bar{A}_2$. 由上面的性质, $\bar{A}_1$ 与 $\bar{A}_2$ 也相互独立, 故

$$P(\text{系统失效})=P(\bar{A}_1)\,P(\bar{A}_2)=(1-0.9)(1-0.8)=0.1\times 0.2=0.02.$$

系统正常的概率为

$$P(\text{系统正常})=1-0.02=0.98.$$

对比两个结果. 单个部件的可靠性最高不过 0.9, 并联之后系统可靠性提升到 0.98. 用多个部件互为备份来对冲单点失效的风险, 这就是工程上**冗余设计**的概率原理.

知乎@余命数

10.4.2 多个事件的独立性

三个或更多事件的独立性, 不能简单地用两两独立来概括.

知乎@余命数

定义 10.14 (n 个事件的相互独立) 设 $A_1, A_2, \dots, A_n$ 是 n 个事件. 若对任意满足 $2 \le k \le n$ 的整数 k, 以及任意 k 个不同的下标 $1 \le i_1 < i_2 < \cdots < i_k \le n$, 都有

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})\,P(A_{i_2}) \cdots P(A_{i_k}),$$

则称这 n 个事件**相互独立**.

以 $n = 3$ 为例, 相互独立要求以下四条同时成立.

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C),$$

$$P(BC) = P(B)P(C), \quad P(ABC) = P(A)P(B)P(C).$$

知乎@余命数

前三条只涉及两个事件, 合起来称为**两两独立**. 两两独立推不出相互独立, 最后一条三者联合的等式可能单独失效.

例题 10.24 (两两独立但不相互独立) 抛两枚均匀的硬币. 记 A 为第一枚正面朝上, B 为第二枚正面朝上, C 为两枚结果相同. 验证 A, B, C 两两独立, 但不相互独立.

知乎@余命数

解 样本空间仍是 $\{(H,H),(H,T),(T,H),(T,T)\}$. 逐一计算, 得

知乎@余命数

$$P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2}.$$

先算两两积事件.

$$P(A \cap B) = P(\{(H,H)\}) = \frac{1}{4} = P(A)P(B),$$

$$P(A \cap C) = P(\{(H,H)\}) = \frac{1}{4} = P(A)P(C),$$

$$P(B \cap C) = P(\{(H,H)\}) = \frac{1}{4} = P(B)P(C).$$

三个事件两两独立. 再算三者同时发生.

$$P(A \cap B \cap C) = P(\{(H,H)\}) = \frac{1}{4} \neq \frac{1}{8} = P(A)P(B)P(C).$$

所以 A, B, C 不相互独立.

直观的解释如下. 单独看 A 或 B, 都改变不了 C 的概率; 但 A 与 B **合起来**完全决定了 C, 两枚的结果都知道了, 是否相同也就知道了. 两两独立只排除了单个事件的影响, 管不住多个事件的联合影响, 这正是定义中要求任意 k 个的原因.

多个事件的相互独立最常见的出场方式, 是**独立重复试验**. 同一个试验在相同条件下重复进行, 各次的结果互不影响.

知乎@余命数

例题 10.25 (重复独立试验) 连续抛一枚均匀的硬币 5 次, 求恰好出现 3 次正面的概率.

解 把 5 次抛掷的结果按顺序记录下来, 每一种完整的结果序列都是一个样本点, 例如

知乎@余命数

$$(\text{正}, \text{正}, \text{反}, \text{正}, \text{反}).$$

各次抛掷相互独立, 由乘法公式, 这个特定序列的概率为

$$\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\left(\frac{1}{2}\right)^5=\frac{1}{32}.$$

同理, 任何一个指定了每次正反的序列, 概率都是 $\frac{1}{32}$.

事件恰好 3 次正面包含的序列, 相当于在 5 个位置中选出 3 个放正, 共有 $\binom{5}{3}=10$ 个. 这些序列两两互斥, 由可加性公理,

$$P(\text{恰好 3 次正面})=\binom{5}{3}\times\frac{1}{32}=10\times\frac{1}{32}=\frac{5}{16}.$$

知乎@余命数 知乎@余命数 余命数 知乎@余命数

每种指定序列的概率用独立性算, 序列的条数用组合数算, 这个两步走是处理独立重复试验的标准动作, 将来研究二项分布时还会反复用到它.

本节从互不影响的直觉出发, 把独立性写成了乘积公式, 并用它拆解了系统可靠性、重复试验等问题. 下一节换一个方向. 当一个事件可以由若干种原因引发时, 我们如何把各种原因的概率拼出该事件的总概率; 反过来, 看到结果之后, 又如何反推每种原因的可能性.

10.5 全概率公式与贝叶斯公式

很多事件的概率难以直接求得, 但它总是由若干种可以一一枚举的原因引发的. 每种原因之下, 事件的条件概率往往好算得多. 把各种原因的情形拼起来, 就得到了事件的总概率, 这就是**全概率公式**. 反过来, 试验结束后看到结果, 回头追问哪种原因的可能性最大, 这就是**贝叶斯公式**. 这两个公式是条件概率的重要应用, 本章开头悬而未决的蒙提霍尔问题, 也将在本节得到正式的解答.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

10.5.1 全概率公式

先引入一个术语. 把样本空间切成若干互不重叠的块, 每块代表一种原因或情形.

定义 10.15 (样本空间的划分) 设 $B_1, B_2, \dots, B_n$ 是样本空间 Ω 中的一组事件. 若它们满足

1. **两两互斥**. 对任意 $i \neq j$, $B_i \cap B_j = \emptyset$;
2. **完备**. $B_1 \cup B_2 \cup \cdots \cup B_n = \Omega$;
3. **正概率**. $P(B_i) > 0$ $(i = 1, 2, \dots, n)$,

则称 $B_1, B_2, \dots, B_n$ 为样本空间 Ω 的一个**划分**, 或称**完备事件组**.

知乎@余命数

最熟悉的划分莫过于 $\{A, \bar{A}\}$. 任何事件与它的对立事件, 恰好把样本空间一分为二.

定理 10.6 (全概率公式) 设 $B_1, B_2, \dots, B_n$ 是样本空间 Ω 的一个划分, 则对任意事件 A,

$$P(A) = \sum_{i=1}^{n} P(B_i)\, P(A \mid B_i).$$

余命数

证明 由于 $B_1, \dots, B_n$ 的并集是整个 Ω,

$$A = A \cap \Omega = A \cap (B_1 \cup B_2 \cup \cdots \cup B_n) = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n).$$

知乎@余命数

由 B_i 两两互斥, 可知 $A \cap B_i$ 也两两互斥. 应用可加性公理,

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i).$$

再由 $P(B_i) > 0$, 对每一项使用乘法公式 $P(A \cap B_i) = P(B_i)\, P(A \mid B_i)$, 即得

$$P(A) = \sum_{i=1}^{n} P(B_i)\, P(A \mid B_i).$$

□

全概率公式按互斥的原因划分事件. 直接求 $P(A)$ 困难时, 找一组把问题切成若干简单情形的划分 $\{B_i\}$, 如图 10.4, 在每种情形下分别计算条件概率, 再按各情形自身的概率加权求和.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

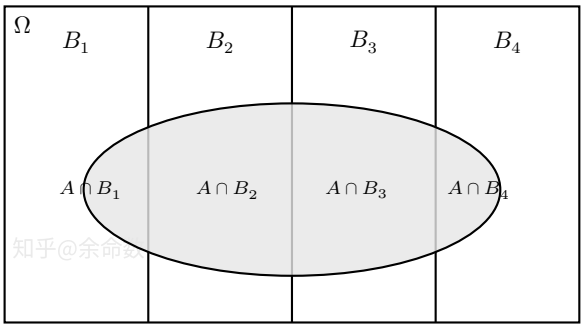


图 10.4 全概率公式的几何图像. 划分 B_1, B_2, B_3, B_4 把事件 A 切成互斥的四块, $P(A)$ 是各块概率按 $P(B_i)$ 加权之和

例题 10.26 (抽签的公平性) 5 张签中有 1 张有奖. 甲、乙两人依次不放回地各抽一张签. 求乙抽到奖签的概率, 并与甲比较.

解 乙能否中奖, 取决于甲抽走了什么. 记 B 为甲抽到奖签, 则 $\{B, \bar{B}\}$ 构成一个划分, 且

$$P(B) = \frac{1}{5}, \qquad P(\bar{B}) = \frac{4}{5}.$$

记 A 为乙抽到奖签. 我们分两种情形写出条件概率.

- 若甲已经抽走奖签, 则乙绝不可能中奖, 即 $P(A \mid B) = 0$.
- 若甲没中奖, 则剩下 4 张签中有 1 张奖签, 即 $P(A \mid \bar{B}) = \frac{1}{4}$.

由全概率公式,

$$P(A) = P(B)\,P(A \mid B) + P(\bar{B})\,P(A \mid \bar{B}) = \frac{1}{5} \times 0 + \frac{4}{5} \times \frac{1}{4} = \frac{1}{5}.$$

整个推理过程可以画成一棵概率树 (图 10.5). 第一阶段按甲的结果分枝, 第二阶段按乙的结果分枝; 每条边上标注相应的 (条件) 概率, 沿路径相乘得到各结局的联合概率, 再把属于乙中奖的叶子加起来, 就是全概率公式在树上读出的结果.

乙中奖的概率与甲完全相同, 都是 $\frac{1}{5}$. 同样的计算可以推广到更多人. 无论先抽后抽, 每个人中奖的概率都相等. 抽签作为一种分配手段是公平的, 先到者并不占便宜, 直觉里后抽的人机会被先抽者抢走的担心, 被全概率公式证伪了.

例题 10.27 (两个供应商) 某工厂从甲、乙两家供应商处采购同一种零件, 甲厂供货占 60%, 次品率为 2%; 乙厂供货占 40%, 次品率为 5%. 从仓库中随机取出一个零件, 求它是次品的概率.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

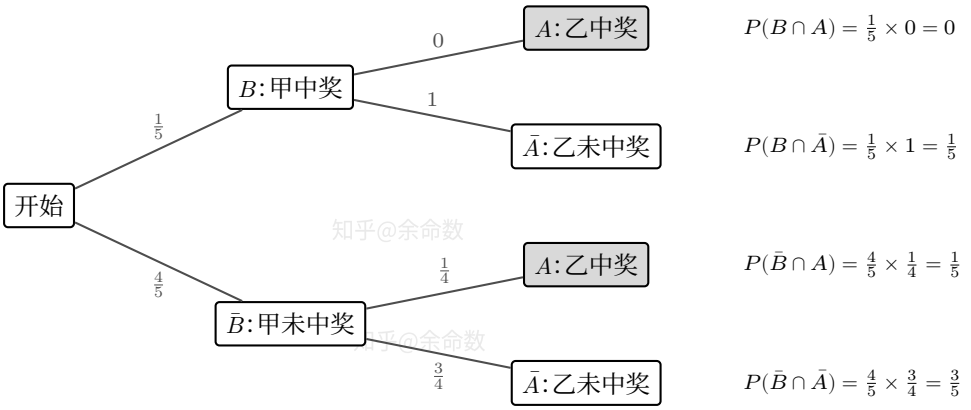


图 10.5　抽签问题的概率树. 沿路径相乘、按叶求和. 两个灰色叶子对应乙中奖, 概率之和为 $0+\frac{1}{5}=\frac{1}{5}$

解　取出的零件来自哪家供应商, 构成一个划分. 记 B_1 为零件来自甲厂, B_2 为零件来自乙厂, A 为零件是次品. 已知

$$P(B_1)=0.6,\quad P(B_2)=0.4,\quad P(A\mid B_1)=0.02,\quad P(A\mid B_2)=0.05.$$

由全概率公式,

$$\begin{aligned}P(A)&=P(B_1)\,P(A\mid B_1)+P(B_2)\,P(A\mid B_2)\\&=0.6\times 0.02+0.4\times 0.05=0.012+0.020=0.032.\end{aligned}$$

零件为次品的概率是 3.2%.

10.5.2　贝叶斯公式

上一个例题还可以追问一步. 如果抽到的零件**确实是次品**, 它更可能来自哪家供应商, 这样的问题颠倒了条件与结果的顺序. 全概率公式由原因推出结果, 现在要由结果反推原因.

定理 10.7 (贝叶斯公式)　设 $B_1,B_2,\dots,B_n$ 是样本空间 Ω 的一个划分, A 是满足 $P(A)>0$ 的事件. 则对每个 $i=1,2,\dots,n$,

$$P(B_i\mid A)=\frac{P(B_i)\,P(A\mid B_i)}{\sum_{j=1}^{n}P(B_j)\,P(A\mid B_j)}.$$

证明　由条件概率的定义,

$$P(B_i\mid A)=\frac{P(B_i\cap A)}{P(A)}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对分子使用乘法公式 $P(B_i \cap A) = P(B_i)\,P(A \mid B_i)$, 对分母使用全概率公式, 即得所证等式. □

公式中各项有约定俗成的名字.$P(B_i)$ 称为**先验概率**, 它表示在看到结果 A 之前, 我们对各种原因可能性的原有判断; $P(B_i \mid A)$ 称为**后验概率**, 它表示结果 A 发生后, 我们对各种原因可能性的修正判断. 贝叶斯公式回答的正是如何用证据更新信念. 后验概率与先验概率以及该原因产生证据的能力 $P(A \mid B_i)$ 成正比.

例题 10.28 (次品溯源) 继续上面的供应商例题. 已知取出的零件是次品, 求它来自甲厂的概率.

解 由贝叶斯公式,

$$P(B_1 \mid A) = \frac{P(B_1)\,P(A \mid B_1)}{P(B_1)\,P(A \mid B_1) + P(B_2)\,P(A \mid B_2)} = \frac{0.012}{0.032} = \frac{3}{8} = 0.375.$$

甲厂供货占 60%, 但次品中只有 37.5% 来自甲厂. 两个数字不矛盾. 甲厂的次品率低, 摊到每件次品上的嫌疑自然小于它的供货份额. 后验概率把供货份额与质量两方面一并计入.

知乎@余命数

例题 10.29 (医学筛查) 某种疾病在人群中的患病率为 0.1%. 现有一种检测方法. 患病者检测呈阳性的概率为 99%(灵敏度), 健康者检测呈阳性的概率为 1%, 即假阳性率. 某人检测结果为阳性, 求他真正患病的概率.

解 记 B 为此人患病, A 为检测呈阳性. 已知

$$P(B) = 0.001, \quad P(\bar{B}) = 0.999, \quad P(A \mid B) = 0.99, \quad P(A \mid \bar{B}) = 0.01.$$

由贝叶斯公式,

$$P(B \mid A) = \frac{P(B)\,P(A \mid B)}{P(B)\,P(A \mid B) + P(\bar{B})\,P(A \mid \bar{B})} = \frac{0.001 \times 0.99}{0.001 \times 0.99 + 0.999 \times 0.01}.$$

计算分子分母, 得

$$P(B \mid A) = \frac{0.00099}{0.00099 + 0.00999} = \frac{0.00099}{0.01098} \approx 0.0902.$$

检测呈阳性的人, 真正患病的概率只有约 9%.

直觉的猜测往往是“检测有 99% 的准确率, 阳性结果应该有九成九的把握”, 计算给出的数字却不到一成. 反差来自基础概率. 人群中的健康者数量是患病者的近千倍, 哪怕只有 1% 的健康者被误判, 假阳性的人数也会淹没真阳性的人数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这正是先验概率的作用, 面对罕见事件, 单一证据很难彻底翻转极低的先验, 需要追加独立的复查来逐步更新判断.

这个结论用自然频率数一遍人会看得更清楚 (图 10.6).

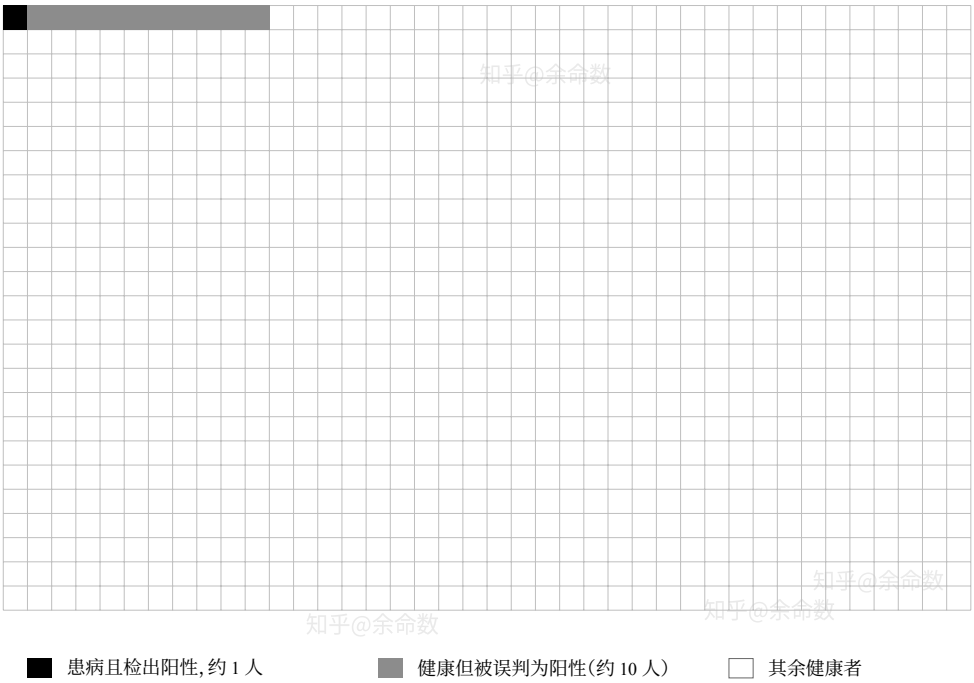


图 10.6　医学筛查的自然频率图像. 每格代表 1 人, 共 1000 人. 阳性者约 11 人(1 黑格加 10 灰格), 其中真正患病的只有约 1 人, 占 $\frac{1}{11} \approx 9\%$

10.5.3 回到蒙提霍尔问题

现在补齐本章开头的欠账. 回顾蒙提霍尔问题的规则. 三扇门后有一辆汽车和两只山羊, 车的位置等可能, 见图 10.1; 玩家初选门 1; 主持人知道车的位置, 并且遵守固定的规则, 即他总是打开剩下一扇有山羊的门, 如果两扇可选的门后都是山羊, 即车恰好在门 1 后, 他随机打开其中一扇. 主持人打开门 3 之后, 需要比较坚持门 1 与换到门 2 哪个胜算更大.

记 C_i 为车在门 i 后 $(i = 1, 2, 3)$, H_3 为主持人打开门 3. 由对称性,

$$P(C_1) = P(C_2) = P(C_3) = \frac{1}{3}.$$

按照主持人的规则, 我们逐个写出条件概率.

- 若车在门 1 后, 门 2、门 3 后都是山羊, 主持人随机开一扇, 于是 $P(H_3 \mid C_1) = \frac{1}{2}$.
- 若车在门 2 后, 主持人只能打开门 3, 于是 $P(H_3 \mid C_2) = 1$.
- 若车在门 3 后, 主持人不能打开门 3, 于是 $P(H_3 \mid C_3) = 0$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

由全概率公式,

$$P(H_3)=\frac{1}{3}\times\frac{1}{2}+\frac{1}{3}\times 1+\frac{1}{3}\times 0=\frac{1}{2}.$$

再由贝叶斯公式, 分别计算两种策略的胜率, 得

$$P(C_1\mid H_3)=\frac{P(C_1)\,P(H_3\mid C_1)}{P(H_3)}=\frac{\frac{1}{3}\times\frac{1}{2}}{\frac{1}{2}}=\frac{1}{3},$$

$$P(C_2\mid H_3)=\frac{P(C_2)\,P(H_3\mid C_2)}{P(H_3)}=\frac{\frac{1}{3}\times 1}{\frac{1}{2}}=\frac{2}{3}.$$

坚持门 1 的胜率仍是 $\frac{1}{3}$, 换到门 2 的胜率是 $\frac{2}{3}$. 换门确实使胜率翻倍, 章首宣称的结论得到严格的证明.

其实还有一个更干脆的论证, 一条式子都不用写. 注意主持人永远不会打开有车的门, 因此有下面两条.

- 若玩家初选正确(概率 $\frac{1}{3}$), 换门必然扑空;
- 若玩家初选错误(概率 $\frac{2}{3}$), 车在两扇未选的门之一后面, 主持人被迫打开了另一扇山羊门, 剩下那扇未开的门后必为汽车, 换门必胜.

换门获胜, 当且仅当初选错误, 故换门胜率为 $P(\text{初选错误})=\frac{2}{3}$. 这层关系用策略树直接表示 (图 10.7).

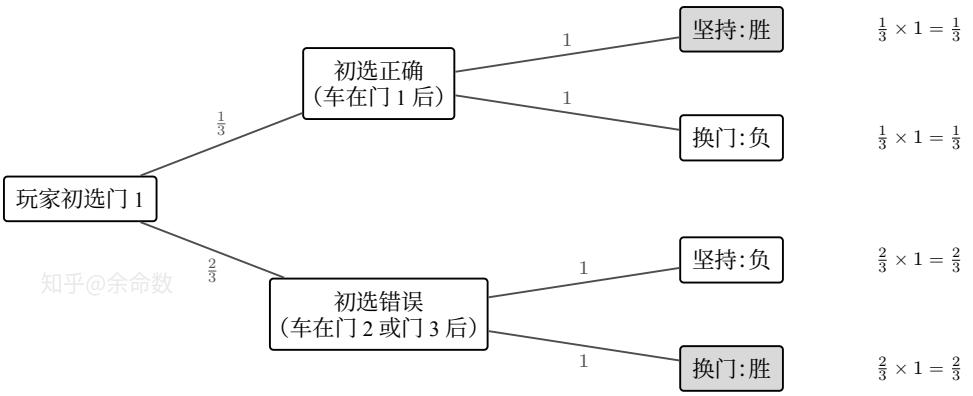


图 10.7 蒙提霍尔问题的策略树. 主持人总会打开一扇山羊门, 胜负在初选时就已注定. 坚持胜与换门胜各对应一个灰色叶子, 概率分别为 $\frac{1}{3}$ 与 $\frac{2}{3}$

直觉在这里失灵,是因为直觉把主持人开门当成了一次随机的排除. 三扇门随便排除一扇, 剩下两扇自然五五开. 但主持人是**带着知识**开门的, 他刻意避开汽车. 这个有偏向的动作, 把初选错误的那 $\frac{2}{3}$ 概率全部集中到了剩下那扇未开的门上. 章首的三情况穷举图 (图 10.2) 数出了同样的答案, 而条件概率与贝叶斯公式把这条信息的来龙去脉记成了明账. 先验是三个 $\frac{1}{3}$, 主持人开门这条证据按 $P(H_3\mid C_i)$ 加权, 后验就变成 $\frac{1}{3}$ 与 $\frac{2}{3}$ 的分化.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

不过到目前为止，我们回答的问题都是某事件发生的概率是多少. 更多的场合里，我们关心的其实是试验产生的数值. 掷骰子的点数、抽检中的次品数、一名射手的命中环数. 把这些数值作为研究对象，需要一个新的概念，即随机变量.

知乎@余命数

10.6 离散型随机变量及其分布列

知乎@余命数

前面几节里，我们关心的都是某事件发生的概率. 但一次试验的结果本身也可以变成数. 掷一颗骰子时，我们自然关心点数；抽检一批产品时，我们关心次品的个数；连抛硬币时，我们也可把正面记为 1、反面记为 0. 给每个样本点配上一个数，这个数就成了新的研究对象.

10.6.1 随机变量的概念

定义 10.16 (随机变量) 设随机试验的样本空间为 Ω. 若对每个样本点 $\omega \in \Omega$，都有唯一确定的实数 $X(\omega)$ 与之对应，则称这个对应关系 X 为一个**随机变量**，通常用大写字母 X, Y, Z 等表示.

知乎@余命数

随机变量并不神秘. 用 第2章 中已经熟悉的话来说，它就是定义在样本空间 Ω 上、取值于实数集 $\mathbb{R}$ 的一个**函数**. 它与普通函数的唯一区别，是定义域里的每个样本点都带着一个概率；于是 X 取某个值这件事也有了概率.

例题 10.30 (随机变量的例子)

- 掷一颗骰子，令 X 为掷出的点数. 样本空间 $\Omega = \{1, 2, \dots, 6\}$，$X(\omega) = \omega$.
- 抛两枚硬币，令 Y 为出现正面的次数. 样本空间为

$$\{(H,H), (H,T), (T,H), (T,T)\},$$

  Y 在四个样本点处的取值分别为 $2, 1, 1, 0$.
- 抛一枚硬币，令

$$Z = \begin{cases} 1, & \text{正面朝上}, \\ 0, & \text{反面朝上}. \end{cases}$$

  试验结果本身不是数，人为赋值之后，Z 也是一个随机变量.

若一个随机变量的所有可能取值可以一一列举出来（有限个，或像 $0, 1, 2, \dots$ 那样排成一列），就称它为**离散型随机变量**. 上面三个例子都是离散型的；本章只讨论离散型. 取值可数无限时，分布列与期望中的求和改为相应的级数，其收敛性

知乎@余命数

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

需要单独判断. 像灯泡寿命这种取值充满一个区间的随机变量, 称为连续型; 严格处理需要积分, 本书把它放在 第17章.

引入随机变量之后, 事件获得了新的表达方式.$\{X=2\}$ 表示 X 取值为 2 这一事件, 它对应样本点的集合 $\{\omega\in\Omega: X(\omega)=2\}$; 同样, $\{X\le 3\}$、$\{1<X\le 5\}$ 也都是事件. 例如抛两枚硬币时, $\{Y=1\}=\{(H,T),(T,H)\}$, 所以

$$P(Y=1)=\frac{2}{4}=\frac{1}{2}.$$

随机变量把事件翻译成了关于数的等式与不等式, 许多实际问题因此变得易于书写.

10.6.2 分布列

光知道随机变量能取哪些值还不够, 还要知道它取每个值的概率有多大. 把这两样信息列成一张表, 就得到了分布列.

定义 10.17 (分布列) 设离散型随机变量 X 的所有可能取值为 $x_1,x_2,\dots,x_n$, 取每个值的概率为

$$P(X=x_i)=p_i\quad(i=1,2,\dots,n).$$

则称该表为随机变量 X 的**概率分布列**, 简称**分布列**.

X	x_1	x_2	$\cdots$	x_n
P	p_1	p_2	$\cdots$	p_n

知乎@余命数

性质 (分布列的性质) 离散型随机变量 X 的分布列满足下面两条.

1. $p_i\ge 0\ (i=1,2,\dots,n)$;
2. $p_1+p_2+\cdots+p_n=1$.

证明 (1) 每个 p_i 都是事件 $\{X=x_i\}$ 的概率, 由非负性公理即得.

(2) 事件 $\{X=x_1\},\{X=x_2\},\dots,\{X=x_n\}$ 两两互斥 (X 一次只能取一个值), 且它们的并集是整个样本空间, 因为 X 总要取某个值. 由可加性公理与规范性公理,

$$\sum_{i=1}^{n}p_i=\sum_{i=1}^{n}P(X=x_i)=P\Big(\bigcup_{i=1}^{n}\{X=x_i\}\Big)=P(\Omega)=1.$$

□

知乎@余命数

这两条性质是检验一张分布列是否合法的标准, 也常用来反求分布列中待定的参数.

例题 10.31 (两颗骰子点数之和的分布列) 掷两颗均匀的骰子, 记 X 为两颗骰子的点数之和. 求 X 的分布列.

解 X 的可能取值为 $2,3,\dots,12$. 两颗骰子的结果共有 36 个等可能样本点, 和为 k 的样本点数等于方程 $x_1+x_2=k$ 满足 $1\le x_1,x_2\le 6$ 的整数解个数, 例如和为 4 的有 $(1,3),(2,2),(3,1)$ 共 3 个. 逐一枚举, 得下表. 为书写紧凑, 第二行记录的是样本点个数, 概率为样本点数除以 36.

X	2	3	4	5	6	7	8	9	10	11	12
样本点数	1	2	3	4	5	6	5	4	3	2	1

故 X 的分布列为

知乎@余命数

$$P(X=k)=\frac{6-|k-7|}{36},\quad k=2,3,\dots,12.$$

直接检验, $1+2+3+4+5+6+5+4+3+2+1=36$, 各概率之和 $\frac{36}{36}=1$, 符合分布列的两条性质.

例题 10.32 (抽检中的次品数) 一批产品共 5 件, 其中有 2 件次品. 从中不放回地取出 3 件, 记 X 为取到的次品件数. 求 X 的分布列.

解 X 的可能取值为 $0,1,2$. 从 5 件中取 3 件, 共有 $\binom{5}{3}=10$ 种等可能取法. 取到 k 件次品等价于从 2 件次品中选 k 件、从 3 件正品中选 $3-k$ 件, 故

$$P(X=k)=\frac{\binom{2}{k}\binom{3}{3-k}}{\binom{5}{3}},\quad k=0,1,2.$$

分别代入得

$$P(X=0)=\frac{\binom{2}{0}\binom{3}{3}}{\binom{5}{3}}=\frac{1}{10},\quad P(X=1)=\frac{\binom{2}{1}\binom{3}{2}}{\binom{5}{3}}=\frac{6}{10},\quad P(X=2)=\frac{\binom{2}{2}\binom{3}{1}}{\binom{5}{3}}=\frac{3}{10}.$$

故分布列为

X	0	1	2
P	$\frac{1}{10}$	$\frac{6}{10}$	$\frac{3}{10}$

直接检验, $\frac{1}{10}+\frac{6}{10}+\frac{3}{10}=1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

求分布列, 即列出 X 的所有可能取值, 逐一计算 $P(X=x_i)$, 并用两条性质检验. 计算每个取值的概率, 是把概率问题归结为计数问题, 这一步依赖 第9章 与前面各节建立的工具.

分布列完整地刻画了随机变量, 但一张表往往信息过载. 比较两个随机变量时, 我们常希望各用一个数字概括它们. 一个数字代表平均水平, 另一个代表波动程度. 这就是下一节要定义的期望与方差.

10.7 期望与方差

分布列给出了离散型随机变量的全部信息, 但一整张表读起来并不轻松. 比较两名射手的水平、估计一家保险公司的盈亏, 我们想要的往往是几个有代表性的数字. 本节引入最常用的两个数字, **期望**刻画平均水平, **方差**刻画波动程度.

10.7.1 数学期望

先从一个思想实验说起. 掷一颗骰子 n 次, 把每次的点数都记下来, 计算这 n 个数的算术平均值. 当 n 很大时, 根据频率的稳定性, 点数 k 出现的频率会稳定在 $\frac{1}{6}$ 附近, 于是平均值稳定在

$$1\times\frac{1}{6}+2\times\frac{1}{6}+\cdots+6\times\frac{1}{6}=3.5$$

附近. 这个按概率加权的平均值, 就是期望的雏形.

定义 10.18 (数学期望) 设离散型随机变量 X 的分布列为 $P(X=x_i)=p_i$ $(i=1,2,\dots,n)$. 称

$$E(X)=\sum_{i=1}^{n}x_ip_i$$

为随机变量 X 的**数学期望**, 简称**期望**, 也称**均值**.

期望是以概率为权重的加权平均. 取值越容易出现, 权重越大, 对平均水平的贡献也越大. 期望未必是 X 能取到的值, 骰子的期望 3.5 就不在 $\{1,2,\dots,6\}$ 之中, 它描述的是长期平均意义下的中心位置.

例题 10.33 求上一节抽检次品数一例中, 次品数 X 的期望.

解 由分布列

X	0	1	2
P	$\frac{1}{10}$	$\frac{6}{10}$	$\frac{3}{10}$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

按期望定义直接计算,

$$E(X)=0\times\frac{1}{10}+1\times\frac{6}{10}+2\times\frac{3}{10}=\frac{12}{10}=1.2.$$

平均每次抽检会抽到 1.2 件次品.

期望有一组用起来极为顺手的运算性质.

性质 (期望的线性性质) 设 X,Y 是离散型随机变量, a,b 为常数, 则

1. $E(aX+b)=aE(X)+b$;
2. $E(X+Y)=E(X)+E(Y)$.

证明 (1) 设 X 的分布列为 $P(X=x_i)=p_i$. 随机变量 $aX+b$ 取值 ax_i+b 的概率也是 p_i, 故

$$E(aX+b)=\sum_{i=1}^{n}(ax_i+b)\,p_i=a\sum_{i=1}^{n}x_ip_i+b\sum_{i=1}^{n}p_i=aE(X)+b.$$

(2) 在有限样本空间 Ω 上, 期望可以改写成对样本点求和的形式. 把 $\sum_{\omega}X(\omega)\,P(\{\omega\})$ 中取值相同的项合并, 恰好得到 $\sum_i x_iP(X=x_i)$. 于是

$$\begin{aligned}E(X+Y)&=\sum_{\omega\in\Omega}(X(\omega)+Y(\omega))\,P(\{\omega\})\\&=\sum_{\omega\in\Omega}X(\omega)\,P(\{\omega\})+\sum_{\omega\in\Omega}Y(\omega)\,P(\{\omega\})=E(X)+E(Y).\end{aligned}$$

□

第二条性质最省力的地方在于, 它不要求 X 与 Y 独立. 无论两个随机变量如何纠缠, 和的期望永远等于期望的和.

例题 10.34 (两颗骰子点数之和的期望) 掷两颗均匀的骰子, 求点数之和 X 的期望.

解 法一采用分布列. 由上一节求得的分布列,

$$\begin{aligned}E(X)=\frac{1}{36}\big(&2\times1+3\times2+4\times3+5\times4+6\times5+7\times6\\&+8\times5+9\times4+10\times3+11\times2+12\times1\big)=\frac{252}{36}=7.\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

法二采用线性性质. 设 X_1, X_2 分别为两颗骰子的点数, 则 $X = X_1 + X_2$. 单颗骰子的期望为

$$E(X_1) = E(X_2) = \frac{1+2+3+4+5+6}{6} = 3.5,$$

故

$$E(X) = E(X_1) + E(X_2) = 3.5 + 3.5 = 7.$$

两种方法结果一致. 把一个复杂的随机变量拆成若干简单随机变量之和, 再用线性性质求期望, 是概率论中反复使用的技巧.

10.7.2 方差

期望相同, 水平未必一样. 看两名射手.

例题 10.35 (两名射手的成绩) 甲、乙两名射手射击一次, 命中环数 X_1, X_2 的分布列分别为

X_1	8	9	10
P	0.2	0.6	0.2

X_2	8	9	10
P	0.4	0.2	0.4

比较两人的射击水平.

解 按期望定义计算两人平均环数,

$$E(X_1) = 8 \times 0.2 + 9 \times 0.6 + 10 \times 0.2 = 9,$$

$$E(X_2) = 8 \times 0.4 + 9 \times 0.2 + 10 \times 0.4 = 9.$$

两人平均环数相同, 但直观上甲的成绩集中在 9 环附近, 乙则经常打出 8 环或 10 环, 甲更稳定. 期望看不出来这个差别, 还需要一个专门度量偏离平均水平程度的数字.

度量偏离, 自然的想法是看各取值与期望之差 $x_i - E(X)$. 直接加权平均会被正负抵消(结果恒为 0), 把差平方再平均就解决了这个问题.

定义 10.19 (方差与标准差) 设离散型随机变量 X 的分布列为 $P(X = x_i) = p_i$ $(i = 1, 2, \ldots, n)$. 称

$$D(X) = \sum_{i=1}^{n} \big(x_i - E(X)\big)^2 p_i$$

为随机变量 X 的**方差**; 称 $\sqrt{D(X)}$ 为 X 的**标准差**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

方差是与期望偏差的平方的期望, 可以记作 $D(X)=E[(X-E(X))^2]$. 方差越小, 取值越集中于期望附近, 随机变量越稳定; 方差越大, 波动越剧烈. 标准差与 X 同量纲, 解释起来更直观.

按定义算方差, 每算一个都要先减期望再平方, 相当繁琐. 下列公式把计算化为两个期望之差.

性质 (方差的计算公式)

$$D(X)=E(X^2)-\big[E(X)\big]^2.$$

证明 记 $\mu=E(X)$. 展开平方并套用期望的线性性质,

$$\begin{aligned}D(X)&=\sum_{i=1}^{n}(x_i-\mu)^2p_i=\sum_{i=1}^{n}(x_i^2-2\mu x_i+\mu^2)\,p_i\\&=\sum_{i=1}^{n}x_i^2p_i-2\mu\sum_{i=1}^{n}x_ip_i+\mu^2\sum_{i=1}^{n}p_i\\&=E(X^2)-2\mu\cdot\mu+\mu^2=E(X^2)-\mu^2.\end{aligned}$$

知乎@余命数

□

性质 设 a,b 为常数, 则

$$D(aX+b)=a^2D(X).$$

证明 由期望的线性性质, $E(aX+b)=aE(X)+b$, 于是

$$(aX+b)-E(aX+b)=aX+b-aE(X)-b=a(X-E(X)).$$

两边平方取期望即

$$D(aX+b)=E[a^2(X-E(X))^2]=a^2E[(X-E(X))^2]=a^2D(X).$$

常数平移 b 不影响方差, 因为整批数据一起移动, 离散程度不变; 缩放 a 会把偏差放大 a 倍, 方差放大 a^2 倍. □

例题 10.36 完成射手比较. 计算 X_1,X_2 的方差, 再求骰子点数的方差.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 套用计算公式 $D(X)=E(X^2)-[E(X)]^2$. 对甲射手, $E(X_1^2)=64\times0.2+81\times0.6+100\times0.2=81.4$, 故

$$D(X_1)=81.4-9^2=0.4.$$

对乙射手, $E(X_2^2)=64\times0.4+81\times0.2+100\times0.4=81.8$, 故

$$D(X_2)=81.8-9^2=0.8.$$

$D(X_1)<D(X_2)$, 甲的成绩更稳定, 与直观判断一致, 期望比平均水平, 方差比稳定程度, 两个数字合起来才构成完整的评价.

骰子点数 X 同法处理, $E(X^2)=\frac{1^2+2^2+\cdots+6^2}{6}=\frac{91}{6}$, 故

$$D(X)=\frac{91}{6}-\left(\frac{7}{2}\right)^2=\frac{91}{6}-\frac{49}{4}=\frac{182-147}{12}=\frac{35}{12}\approx2.92.$$

期望与方差是分布列的摘要. 前者回答平均是多少, 后者回答围绕平均晃多远. 下一节要认识的几个分布模型, 都配有现成的期望与方差公式, 届时可以直接套用.

10.8 常见分布模型

现实中有几类随机试验反复出现. 一次成败的尝试、重复多次的成败尝试、不放回抽检、受大量微小因素叠加影响的测量, 都对应着现成的分布公式与数字特征. 认出模型, 就可以直接套用. 本节认识四个出场率最高的分布模型.

10.8.1 两点分布

最简单的随机试验只有两种结果, 成功与失败、正面与反面、合格与不合格.

定义 10.20 (两点分布) 设随机变量 X 只取 0 和 1 两个值, 分布列为

X	0	1
P	$1-p$	p

其中 $0<p<1$. 则称 X 服从**两点分布**, 也称**伯努利分布**.

按定义直接算其数字特征,

$$E(X)=0\times(1-p)+1\times p=p,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$E(X^2) = 0^2 \times (1-p) + 1^2 \times p = p,$$

$$D(X) = E(X^2) - [E(X)]^2 = p - p^2 = p(1-p).$$

10.8.2 二项分布

把同一个两点分布的试验在相同条件下独立地重复 n 次, 记录成功的总次数, 这就是第四节讨论过的独立重复试验, 也称 n **重伯努利试验**.

定义 10.21 (二项分布) 在 n 重伯努利试验中, 每次试验成功的概率为 p $(0 < p < 1)$, 记 X 为成功的总次数. 则

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \ldots, n,$$

称 X 服从参数为 n, p 的**二项分布**, 记作 $X \sim B(n, p)$.

这个公式的来路, 第四节已经走过一遍. 任何一个指定的、恰好含 k 次成功的结果序列, 例如成成败成…, 由独立性其概率为 $p^k(1-p)^{n-k}$; 这样的序列相当于在 n 个位置中选出 k 个放成功, 共 $\binom{n}{k}$ 个, 且两两互斥, 相加即得.

各概率之和为 1 可由我们已在 第9.3节 中了解过的二项式定理直接验证,

$$\sum_{k=0}^{n} P(X = k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = \big(p + (1-p)\big)^n = 1.$$

二项分布的名字正来自这个二项展开.

定理 10.8 (二项分布的期望) 若 $X \sim B(n, p)$, 则

$$E(X) = np.$$

证明 由期望定义,

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}.$$

$k = 0$ 的项为 0. 对 $k \geq 1$, 利用恒等式 $k\binom{n}{k} = n\binom{n-1}{k-1}$ (两边都是“从 n 人中选 k 人并指定一名组长”的计数), 得

$$\begin{aligned} E(X) &= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k (1-p)^{n-k} = np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\ &= np\big(p + (1-p)\big)^{n-1} = np. \end{aligned}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

□

直觉上这个结论理所当然. 每次试验平均贡献 p 次成功, n 次独立重复自然平均成功 np 次. 用期望的线性性质也能看出这一点. 记 X_i 为第 i 次试验成功的两点分布变量, 则 $X = X_1 + \cdots + X_n$, 故 $E(X) = nE(X_i) = np$.

定理 10.9 (二项分布的方差) 若 $X \sim B(n,p)$, 则

$$D(X) = np(1-p).$$

证明 先算 $E[X(X-1)]$. 对 $k \geq 2$, 利用 $k(k-1)\binom{n}{k} = n(n-1)\binom{n-2}{k-2}$:

$$\begin{aligned} E[X(X-1)] &= \sum_{k=2}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k} \\ &= n(n-1)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2}(1-p)^{(n-2)-(k-2)} \\ &= n(n-1)p^2 (p+(1-p))^{n-2} = n(n-1)p^2. \end{aligned}$$

于是

$$E(X^2) = E[X(X-1)] + E(X) = n(n-1)p^2 + np,$$

$$D(X) = E(X^2) - [E(X)]^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 = np(1-p).$$

□

例题 10.37 (投篮问题) 某篮球运动员的罚球命中率为 0.6. 他独立罚球 10 次, 记 X 为命中次数. (1) 求恰好命中 6 次的概率; (2) 求 X 的期望与方差.

解 每次罚球是一次伯努利试验, $X \sim B(10, 0.6)$.

(1) 由分布公式,

$$P(X=6) = \binom{10}{6} \times 0.6^6 \times 0.4^4 = 210 \times 0.046656 \times 0.0256 \approx 0.2508.$$

(2) 直接套公式,

$$E(X) = np = 10 \times 0.6 = 6, \qquad D(X) = np(1-p) = 10 \times 0.6 \times 0.4 = 2.4.$$

他平均命中 6 次, 命中次数围绕 6 上下波动, 标准差为 $\sqrt{2.4} \approx 1.55$ 次.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

10.8.3 超几何分布

二项分布要求各次试验相互独立, 放回抽样满足这个要求; 不放回抽样时, 每次抽取都改变了剩下的成分, 独立性被破坏, 分布也随之改变.

定义 10.22 (超几何分布) 一批产品共 N 件, 其中次品 M 件. 从中不放回地取出 n 件 ($n \le N$), 记 X 为取到的次品件数. 则

$$P(X=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},\quad k=0,1,2,\dots,\min(M,n),$$

称 X 服从**超几何分布**.

这个公式我们在第六节的抽检次品数一例中已经用过. 分母是全部取法数,分子是"k 件次品从 M 件次品中选、$n-k$ 件正品从 $N-M$ 件正品中选"的取法数.

性质 (超几何分布的期望) 若 X 服从参数为 N, M, n 的超几何分布, 则

$$E(X)=\frac{nM}{N}.$$

证明 记 $C=\binom{N}{n}$. 由期望定义与恒等式 $k\binom{M}{k}=M\binom{M-1}{k-1}$ ($k\ge 1$),

知乎@余命数

$$E(X)=\frac{1}{C}\sum_{k=1}^{n}k\binom{M}{k}\binom{N-M}{n-k}=\frac{M}{C}\sum_{k=1}^{n}\binom{M-1}{k-1}\binom{N-M}{n-k}.$$

右端求和是范德蒙恒等式. 从 $M-1$ 个一类元素与 $N-M$ 个二类元素中共取出 $n-1$ 个, 按一类元素取 $k-1$ 个分类求和, 总数为 $\binom{N-1}{n-1}$. 故

知乎@余命数

$$E(X)=\frac{M\binom{N-1}{n-1}}{\binom{N}{n}}=M\cdot\frac{n}{N}=\frac{nM}{N}.$$

□

结果的读法很直白. 次品率为 $\frac{M}{N}$, 抽 n 件平均抽到 $\frac{nM}{N}$ 件次品. 第六节的例子中 $N=5$、$M=2$、$n=3$, 公式给出 $E(X)=\frac{3\times 2}{5}=1.2$, 与当时按分布列硬算的结果一致.

注 10.5 (放回与不放回)
同样的次品率 $p=\frac{M}{N}$, 放回抽检 n 次, 次品数服从 $B(n,p)$; 不放回抽检 n 件, 次品数服从超几何分布. 二者的期望都是 np, 但方差不同, 因为不放回时各次抽取互相牵制, 波动更小. 当总体很大而抽样量相对很小, 例如从十万件中抽十件时, 抽走几件几乎不改变总体成分, 不放回近似于放回, 超几何分布可用二项分布近似计算, 这是统计抽检中常用的简化.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

正态分布与 3σ 直观

注 10.6
离散分布(两点、二项、超几何)是本章主干. 正态分布给出钟形曲线与 3σ 规则的图像直觉;密度曲线面积的严格含义见 第17章.

10.8.4 正态分布

最后一个模型属于连续型随机变量, 严格处理需要积分; 它的应用极广, 本节给出直观认识.

测量误差、身高体重、考试成绩、工厂零件的尺寸偏差, 即由无数微小因素叠加而成的随机量, 其取值分布往往呈现同一种形态. 中间高、两边低、左右对称的钟形曲线. 这背后有一条严格的数学定理, 称为中心极限定理. 大量独立微小随机因素的叠加, 会趋向同一类分布, 即**正态分布**.

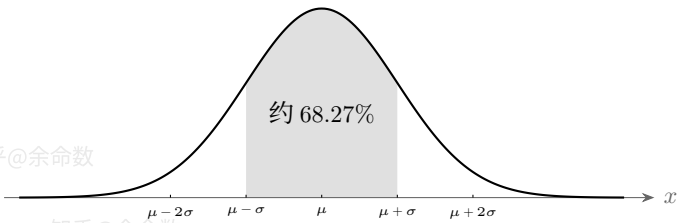


图 10.8 正态分布的密度曲线. 以 μ 为中心、σ 为尺度的钟形曲线, 阴影部分面积约为 68.27%

正态分布由两个参数完全确定. 期望 μ 决定曲线的中心位置, 标准差 σ 决定曲线的胖瘦(σ 越小曲线越瘦高, 取值越集中). 记作 $X\sim N(\mu,\sigma^2)$. 它的概率用曲线下方的面积来度量.$P(a\le X\le b)$ 等于曲线在区间 $[a,b]$ 上与 x 轴围成的面积, 总面积为 1.

性质 (3σ 规则) 若 $X\sim N(\mu,\sigma^2)$, 则

$$P(\mu-\sigma\le X\le\mu+\sigma)\approx 68.27\%,$$

$$P(\mu-2\sigma\le X\le\mu+2\sigma)\approx 95.45\%,$$

$$P(\mu-3\sigma\le X\le\mu+3\sigma)\approx 99.73\%.$$

由曲线的对称性, 中心两侧的概率可以互相换算, 例如

$$P(X\ge\mu+a)=P(X\le\mu-a)=\frac{1-P(\mu-a\le X\le\mu+a)}{2}.$$

几乎全部的取值都落在期望的 3σ 之内, 这是正态分布最实用的画像.

例题 10.38 某年级的一次数学考试成绩 X 近似服从 $N(80,10^2)$. (1) 估计成绩在 70 分到 90 分之间的学生比例; (2) 估计成绩不低于 100 分的学生比例.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 (1) $70 = 80 - 10, 90 = 80 + 10$, 即区间 $[\mu - \sigma, \mu + \sigma]$. 由 3σ 规则,

$$P(70 \le X \le 90) \approx 68.27\%.$$

知乎@余命数

约有六成八的学生成绩落在 70 到 90 分之间.

(2) $100 = 80 + 2 \times 10 = \mu + 2\sigma$. 由对称性,

$$P(X \ge 100) = \frac{1 - P(60 \le X \le 100)}{2} \approx \frac{1 - 95.45\%}{2} \approx 2.28\%.$$

知乎@余命数

百分以上的学生约占 2.3%.

本章以样本空间与概率公理为起点, 用条件概率、独立性、全概率公式与贝叶斯公式刻画信息, 并用随机变量、分布列、期望与方差描述分布, 最后给出几个常用分布模型. 正态分布钟形曲线与 3σ 规则的严格推导以积分为基础 (第17章).

知乎@余命数

知乎@余命数

习题

10.1. (事件运算)设 A, B 为两个事件. 化简: $(A \cup B) \cap (\bar{A} \cup B)$. 提示: 对 B 用分配律提取公因子.

10.2. (古典概型)从 $1, 2, \ldots, 100$ 中任取一个整数, 求它能被 6 或 8 整除的概率. 提示: 用容斥原理, 注意 6 与 8 的最小公倍数是 24.

知乎@余命数

10.3. (错排)5 位同学每人准备一份礼物, 随机交换后每人收到一份. 求没有任何一人拿到自己礼物的方案数. 提示: 即错排数 D_5, 用递推 $D_n = (n-1)(D_{n-1} + D_{n-2})$ 及 $D_3 = 2$, $D_4 = 9$.

知乎@余命数

10.4. (条件概率) 掷两颗均匀的骰子, 已知两颗的点数不同, 求至少有一颗为 6 点的概率. 提示: 在收缩后的样本空间中直接计数.

知乎@余命数

10.5. (独立性)甲、乙、丙三人各自独立破译同一份密码, 能破译的概率分别为 $\frac{1}{5}$、$\frac{1}{4}$、$\frac{1}{3}$. 求密码被破译的概率. 提示: 考虑对立事件三人都未能破译.

10.6. (全概率公式)三个箱子中分别装有 2 红 1 白、1 红 2 白、3 红共 3 个球. 随机选一个箱子, 从中取出一个球, 求取到红球的概率. 提示: 以选中哪个箱子作样本空间的划分.

10.7. (贝叶斯公式) 承上题. 已知取到的是红球, 求它来自装有 3 个红球的箱子的概率. 提示: 分母用第 6 题的结果.

10.8. (期望与方差) 设随机变量 X 的分布列为 $P(X = 0) = 0.2$、$P(X = 1) = 0.5$、$P(X = 2) = 0.3$. 求 $E(X)$ 与 $D(X)$. 提示: 方差用计算公式 $D(X) = E(X^2) - [E(X)]^2$.

10.9. (二项分布) 某选手投篮命中率为 0.7, 独立投篮 4 次. 求恰好命中 3 次的概率. 提示: 命中次数服从 $B(4, 0.7)$.

10.10. (正态分布)设 $X \sim N(2, 9)$, 用 3σ 规则估计 $P(X > 5)$. 提示: $\sigma = 3$, 而 $5 = \mu + \sigma$.

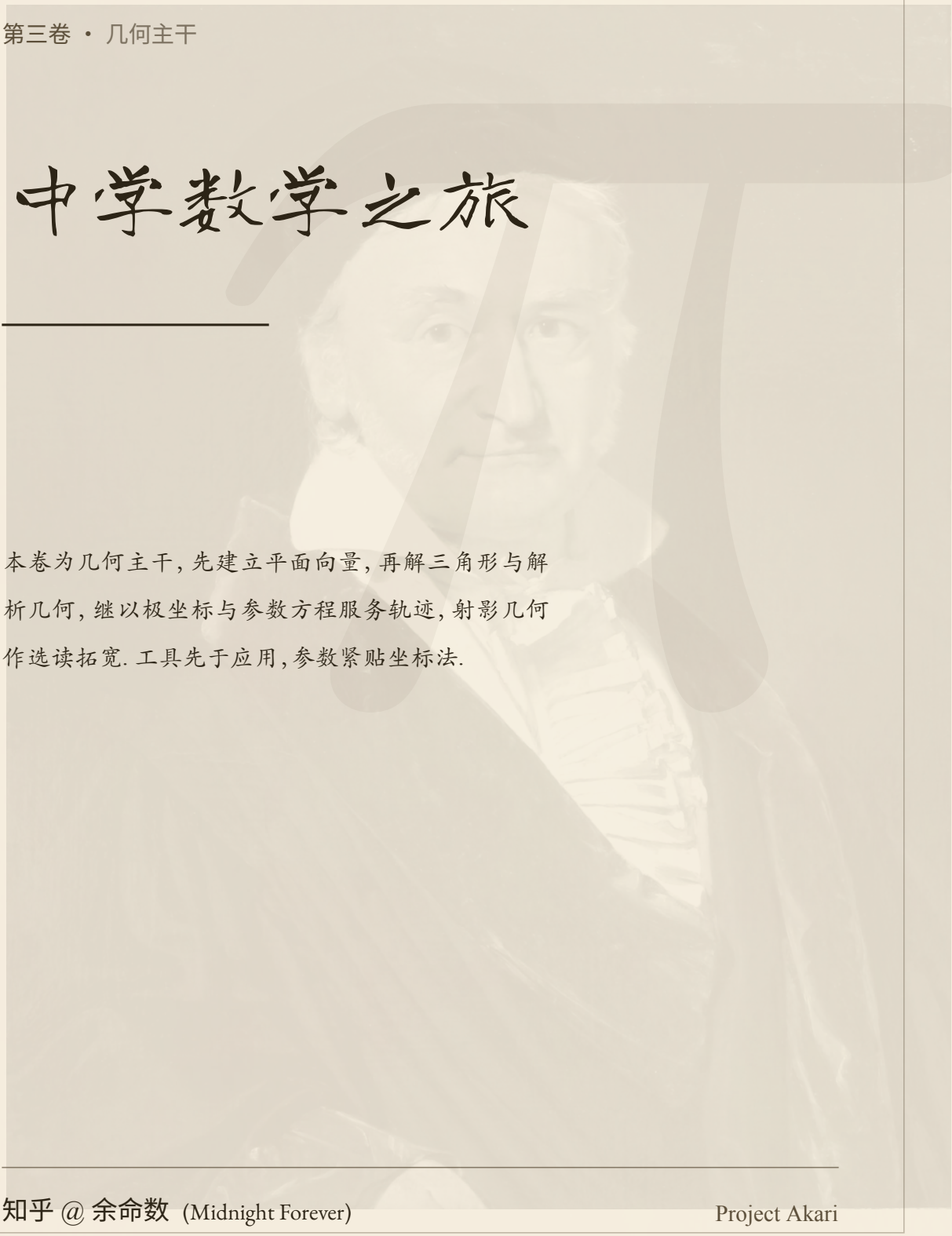
知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
@余命数
知乎@余命数

第三卷 · 几何主干

中学数学之旅



本卷为几何主干，先建立平面向量，再解三角形与解析几何，继以极坐标与参数方程服务轨迹，射影几何作选读拓宽. 工具先于应用，参数紧贴坐标法.

知乎 @ 余命数 (Midnight Forever)

Project Akari

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
命数
知乎@余命数
知乎@余命数
知乎@余命数

本卷导读

本卷线索

第三卷先建立向量语言，再用代数框架翻译几何. 坐标、方程与不变式让位置关系变得可计算、可证明，几何判断由此落到可以操作的推演上.

知乎@余命数

章节地图

- ▷ **平面向量**　运算、数量积、坐标表示与定比分点. 这套语言贯穿解三角形的四心表示与解析几何的全程，是本卷两章的公共底座.
- ▷ **解三角形**　以正弦、余弦定理处理边角关系，配合最值与建系法，几何量由此转成可解的方程.
- ▷ **解析几何**　在向量底座上统一直线、圆与二次曲线；设点法、设线法贯穿题型.
- ▷ **极坐标与参数方程**　承接解析几何，处理离心角、轨迹与旋转对称，为曲线提供另一套描述方式.
- ▷ **射影几何**　交比、调和与极点极线，从观念与结构上回看二次曲线的内在联系.

向后展望

第四卷将拧成分析主干：极限与导数（含数列极限、级数、泰勒）、积分与微分方程，并以思想方法章收束. 第三卷建立的坐标语言，是这些扩展的共同底座.

知乎@余命数

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

知乎@余命数

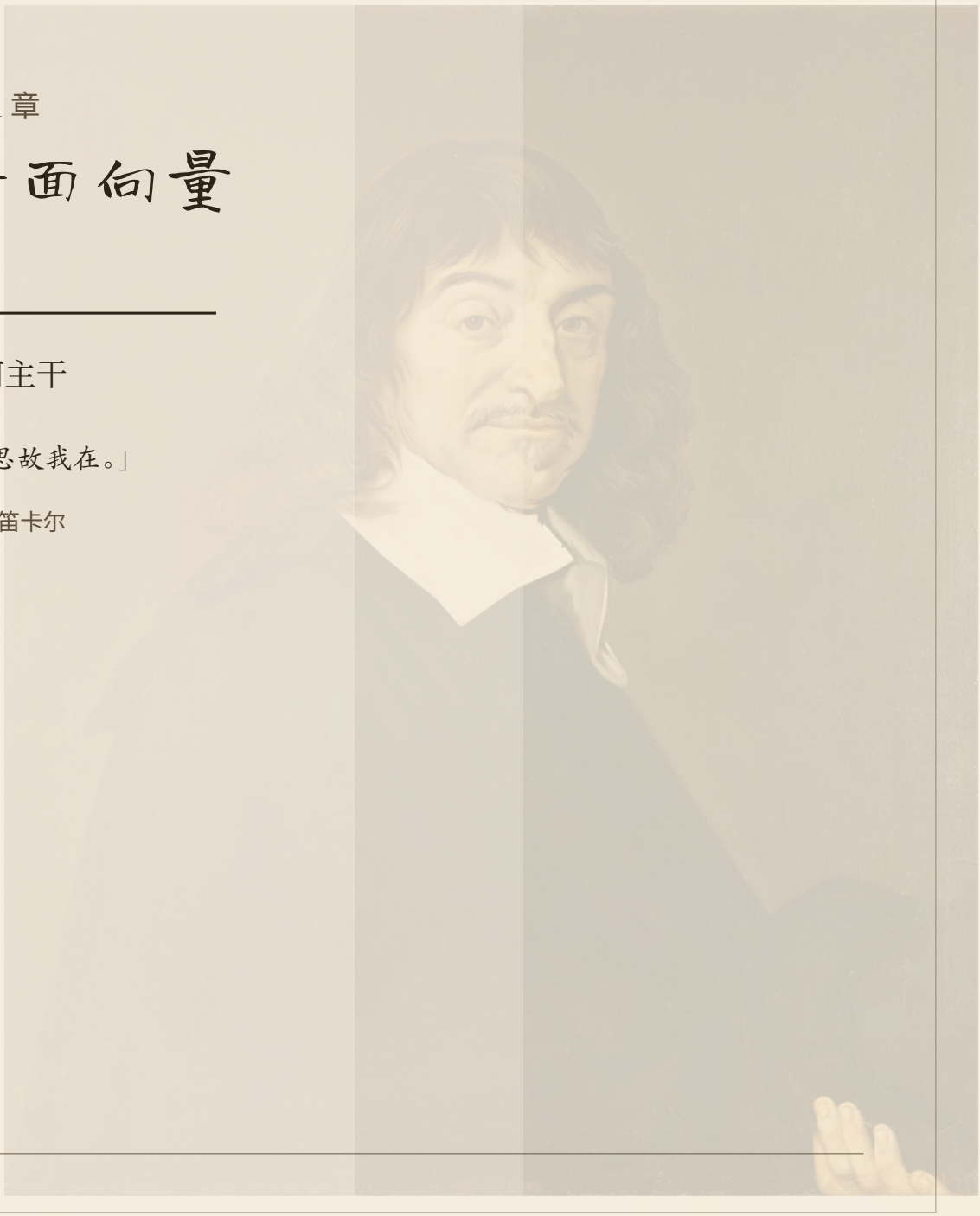
知乎@余命数

知乎@余命数

知乎@余命数

第 11 章

平面向量



几何主干

「我思故我在。」

——笛卡尔

第十一章 平面向量

知乎@余命数

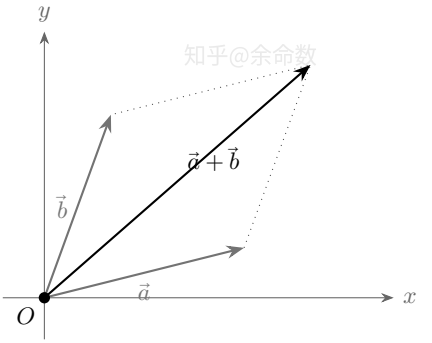
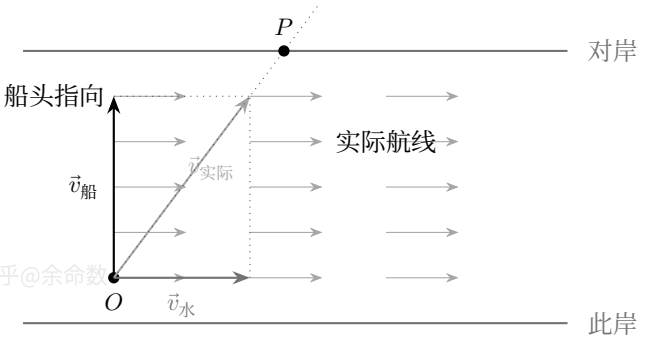


图 11.1 向量将方向与长度从具体线段中抽象出来, 使平移、投影与夹角可计算

本章把几何量改写成可运算的对象. 向量既服务三角形中的边角关系, 也是解析几何中点、直线与轨迹的默认语言.

11.1 向量代数



经典几何学的研究对象是点、线、面等不动的形体. 然而物理世界充满了运动与力, 这些量兼具大小与方向. 要将这些动态的、有方向的量纳入数学体系, 需要引入一种新的代数结构.

想象一艘船试图渡过一条湍急的河流. 船长始终将船头垂直指向对岸, 此时船速为 $\vec{v}_{船}$, 但河水以 $\vec{v}_{水}$ 将船向下游推. 最终, 船的实际航线沿着一条斜线(速度记为 $\vec{v}_{实际}$)到达对岸的下游某处.

在物理中, 力的合成与相对运动都涉及一些既带 **大小**, 如速率与力的大小, 又带 **方向**的物理量. 如何用数学语言精确描述这些带方向的量并加以计算, 是新的课题. 船长想精确地在正对岸靠岸,就必须调整船头方向以抵消水流影响. 此时,代数中的普通数字, 即标量, 已然无能为力. 需要引入一种能把大小与方向融为一体

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

的数学工具, 这就是 **向量**. 这个结构的起点, 是一个比向量更原始、更具体的几何实体, 即有向线段.

11.1.1 从有向线段到向量

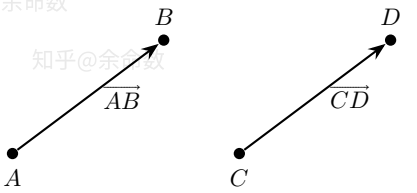
11.1.1.1 有向线段

先定义一个局限于特定位置的几何对象.

定义 11.1 (有向线段) 在空间中给定有序的一对点 A 和 B, 从点 A 指向点 B 的线段称为一条**有向线段**, 记作 $\overrightarrow{AB}$.

- 点 A 称为该有向线段的**起点**.
- 点 B 称为该有向线段的**终点**.

一条有向线段 $\overrightarrow{AB}$ 包含三个基本要素, 即**起点**、**长度**与从 A 指向 B 的**方向**, 三者分别给出位置、大小与方向. 如图所示, 有向线段 $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$ 是两个不同的几何



有向线段 $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$

对象, 因为它们的起点不同. 然而, 物理学的位移或力并不关心起点, 只要长度相等且方向相同, $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$ 就代表同一个物理实体. 例如, 向东北方向移动 5 公里这一位移, 无论从何处开始, 其本身是不变的.

为了捕捉这种独立于位置的有向的大小, 需定义一种等价关系.

定义 11.2 (有向线段的等价) 两条有向线段 $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$ 被认为是**等价**的, 如果其中一条可以通过平移变换得到另一条.

从几何上看, 这等价于要求四边形 $ABDC$ 是一个平行四边形.(注意顶点的顺序). 这同时保证了 $\overrightarrow{AB}$ 和 $\overrightarrow{CD}$ 的长度相等且方向相同.

这种等价关系具有数学上的完备性:

- **自反性**: $\overrightarrow{AB}$ 与自身等价.
- **对称性**: 若 $\overrightarrow{AB}$ 与 $\overrightarrow{CD}$ 等价, 则 $\overrightarrow{CD}$ 与 $\overrightarrow{AB}$ 等价.
- **传递性**: 若 $\overrightarrow{AB}$ 与 $\overrightarrow{CD}$ 等价, 且 $\overrightarrow{CD}$ 与 $\overrightarrow{EF}$ 等价, 则 $\overrightarrow{AB}$ 与 $\overrightarrow{EF}$ 等价.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这种关系将所有向线段划分成一个个互不相交的集合, 每个集合中的所有成员都相互等价. 这个集合, 这个抽象的整体, 就是向量.

定义 11.3 (向量) 一个**向量**是所有相互等价的有向线段所构成的**等价类**.

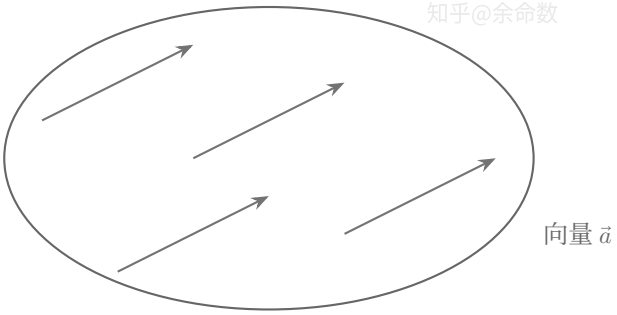


图 11.2 向量 $\vec{a}$ 是一个抽象概念, 它是所有这些长度和方向相同的有向线段的集合. 其中任意一条都是该向量的一个**几何表示**或**代表**.

这个从具体到抽象的定义,将向量从对特定起点和终点的关注中解放出来,使之成为**自由向量**. 可任意平移它的几何表示, 而不改变向量本身. 这为后续的代数运算提供了理论基础.

11.1.2 向量的基本属性与线性运算

已将向量定义为一个抽象的等价类, 它封装了大小与方向这两个信息, 并摆脱了具体位置的束缚. 接着的任务, 是为这些数学对象建立一套自洽的代数运算法则.

11.1.2.1 向量的基本属性

在定义运算之前, 先要明确向量自身的几个基本属性.

- **模**: 向量 $\vec{a}$ 的模, 记作 $|\vec{a}|$, 是其等价类中任意一个有向线段代表的长度. 模是一个非负实数, 它唯一地刻画了向量的大小.
- **相等**: 两个向量 $\vec{a}$ 和 $\vec{b}$ 相等, 记作 $\vec{a}=\vec{b}$, 当且仅当它们的模相等且方向相同. 从等价类的角度看, 它们属于同一个集合.
- **零向量**: 模为 0 的向量称为零向量, 记作 $\vec{0}$. 它的起点与终点重合, 其几何表示为一个点. 由于零向量没有长度, 方向视为任意. 在向量加法中, 零向量是类似于数字 0 的(加法单位元): $\vec{a}+\vec{0}=\vec{a}$.
- **单位向量**: 模为 1 的向量称为单位向量. 单位向量是纯粹的方向载体. 对于任意一个非零向量 $\vec{a}$, 与其方向相同的单位向量可以通过将其自身除以其模得到:

$$\vec{e}_{\vec{a}}=\frac{\vec{a}}{|\vec{a}|}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此, 任何向量都可以表示为其模与它的单位向量的乘积: $\vec{a} = |\vec{a}|\vec{e}_{\vec{a}}$.

- **相反向量**: 与向量 $\vec{a}$ 模相等但方向相反的向量, 称为 $\vec{a}$ 的相反向量, 记作 $-\vec{a}$. 它是向量加法中的逆元: $\vec{a} + (-\vec{a}) = \vec{0}$. 若 $\vec{a}$ 的一个代表是 $\overrightarrow{AB}$, 则 $-\vec{a}$ 的一个代表是 $\overrightarrow{BA}$.

逐项说明如下.

11.1.2.1.1 模 向量 $\vec{a}$ 的模, 记作 $|\vec{a}|$, 在几何上直观地表现为其任意代表有向线段的长度. 然而, 其更基本意义在于, 模是在向量空间上定义的一个**范数**. 范数是一个将向量映射到非负实数的函数, 它满足三个基本公理:

- **正定性**: $|\vec{a}| \geq 0$, 且 $|\vec{a}| = 0$ 当且仅当 $\vec{a} = \vec{0}$.
- **齐次性**: 对于任意实数 λ, 有 $|\lambda\vec{a}| = |\lambda||\vec{a}|$.
- **三角不等式**: $|\vec{a} + \vec{b}| \leq |\vec{a}| + |\vec{b}|$.

模给出长度的度量, 也是赋予一个纯代数空间以几何拓扑结构的根本工具. 有了模, 可定义向量之间的**距离** ($d(\vec{a}, \vec{b}) = |\vec{a} - \vec{b}|$), 进而分析收敛、极限与连续性等分析学概念. 因此, 模把代数与分析联系起来.

11.1.2.1.2 相等 两个向量 $\vec{a}$ 和 $\vec{b}$ 相等, 记作 $\vec{a} = \vec{b}$, 当且仅当它们的模相等且方向相同. 这一看似平凡的定义, 实则已经排除了位置的影响. 从对向量的构造性定义出发, $\vec{a} = \vec{b}$ 的真正含义是, 向量 $\vec{a}$ 和 $\vec{b}$ 属于**同一个等价类**. 即, 代表 $\vec{a}$ 的任意有向线段 $\overrightarrow{AB}$, 与代表 $\vec{b}$ 的任意有向线段 $\overrightarrow{CD}$, 都是相互等价的. 这种等价关系的确立, 使得向量摆脱了其几何代表的具体位置, 成为一个**自由向量**. 正是这种自由, 才使得可将不同位置的向量进行平移、首尾相接或共起点放置, 从而为向量的代数运算, 如加法提供了理论上的合法性与可能性.

11.1.2.1.3 零向量 模为 0 的向量称为零向量, 记作 $\vec{0}$. 它在几何上对应一个点, 其方向被规定为任意. 这一定义的代数意义在于, 零向量是向量加法运算的**单位元**. 在任何一个代数群中, 单位元是那个与其他元素运算后不改变该元素的特殊元素. 对于向量空间, 这表现为对于任意向量 $\vec{a}$, 恒有

$$\vec{a} + \vec{0} = \vec{a}$$

11.1.2.1.4 单位向量 模为 1 的向量称为单位向量. 如果说模是向量大小信息的载体, 那么单位向量就是方向信息的纯粹载体. 将一个非零向量 $\vec{a}$ 除以其模, 即 $\vec{e}_{\vec{a}} = \frac{\vec{a}}{|\vec{a}|}$, 这一过程被称为**规范化**. 规范化操作说明了向量的一个基本结构分解, 即任何非零向量 $\vec{a}$ 都可以唯一地表示为其模与它的单位向量的乘积:

$$\vec{a} = |\vec{a}|\vec{e}_{\vec{a}}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这是一种向量的**极分解**, 类似于复数中的极坐标表示 $z = re^{i\theta}$. 它将一个向量蕴含的两种信息, 即大小与方向, 进行了彻底的分离. 在诸多物理与工程应用中, 常关心过程的方向, 此时单位向量便成为描述该过程最简洁有力的工具. 所有单位向量的终点集合, 构成了向量空间中的单位球面(在二维平面上即单位圆), 这是研究向量旋转等变换的一个基本几何舞台.

11.1.2.1.5 相反向量 与向量 $\vec{a}$ 模相等但方向相反的向量, 称为 $\vec{a}$ 的相反向量, 记作 $-\vec{a}$. 在代数结构中, 相反向量是向量加法运算的**逆元**. 对于群中的每一个元素, 都必须存在一个逆元, 使得该元素与逆元运算的结果是单位元. 对于向量空间, 这表现为对于任意向量 $\vec{a}$, 恒有

$$\vec{a} + (-\vec{a}) = \vec{0}$$

相反向量的存在, 保证了向量加法群的完整性, 并由此自然地导出了向量减法的定义, 即减去一个向量等价于加上它的相反向量. 没有相反向量, 就没有封闭的代数系统, 向量的线性运算也将无从谈起. 从几何上看, 若 $\vec{a}$ 的一个代表是 $\overrightarrow{AB}$, 则 $-\vec{a}$ 的一个代表是 $\overrightarrow{BA}$, 这正给出了相反向量的几何直观.

例题 11.1 判断命题甲"$\vec{a} = \vec{b}$"是命题乙"$|\vec{a}| = |\vec{b}|$"的什么条件?

解 此题考察对向量概念的把握, 要点在于区分向量相等与向量的模相等.

若 $\vec{a} = \vec{b}$, 即向量 $\vec{a}$ 和 $\vec{b}$ 的大小和方向完全相同. 既然大小相同, 那么它们的模必然相等, 即 $|\vec{a}| = |\vec{b}|$. 因此, 甲能够推出乙, 甲是乙的充分条件.

反之, 若 $|\vec{a}| = |\vec{b}|$, 这只说明两个向量的长度相等. 它们的方向可能不同. 例如, 令 $\vec{a}$ 是一个非零向量, 而 $\vec{b} = -\vec{a}$, 则 $|\vec{a}| = |\vec{b}|$, 但 $\vec{a} \neq \vec{b}$. 因此, 乙无法推出甲, 甲不是乙的必要条件.

"$\vec{a} = \vec{b}$"是"$|\vec{a}| = |\vec{b}|$"的**充分不必要条件**.

11.1.2.2 向量加法

向量加法的几何动机源于对连续位移的描述. 想象一个质点从点 A 移动到点 B, 产生位移 $\overrightarrow{AB}$; 接着又从点 B 移动到点 C, 产生位移 $\overrightarrow{BC}$. 整个过程的净效果是从起点 A 直接到达终点 C 的位移 $\overrightarrow{AC}$. 这个首尾相接、连接始末的直观过程, 被抽象为向量加法的基本法则.

定理 11.1 (向量加法的三角形法则) 设 $\vec{a}$ 和 $\vec{b}$ 为任意两个向量. 取 $\vec{a}$ 的一个代表 $\overrightarrow{AB}$, 并取 $\vec{b}$ 的一个以 B 为起点的代表 $\overrightarrow{BC}$. 则向量和 $\vec{a} + \vec{b}$ 由有向线段 $\overrightarrow{AC}$ 所代表.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明　此定理**定义**了向量加法, 也给出了具体的几何作图方法. 定义必须唯一且自洽, 向量和的最终结果不应依赖于最初为向量 $\vec{a}$ 和 $\vec{b}$ 选取的几何代表的位置. 任务是证明这一点, 即证明向量加法是**良定义的**.

设 $\overrightarrow{AB}$ 是向量 $\vec{a}$ 的一个代表, $\overrightarrow{BC}$ 是向量 $\vec{b}$ 的一个代表. 根据定义, 它们的和 $\vec{a}+\vec{b}$ 是由 $\overrightarrow{AC}$ 所代表的向量.

现在, 在平面上另选一点 A', 并作 $\overrightarrow{A'B'}$ 作为 $\vec{a}$ 的另一个代表. 同样, 作 $\overrightarrow{B'C'}$ 作为 $\vec{b}$ 的另一个代表. 现证由这条新路径得到的和向量代表 $\overrightarrow{A'C'}$ 与原始的代表 $\overrightarrow{AC}$ 是等价的.

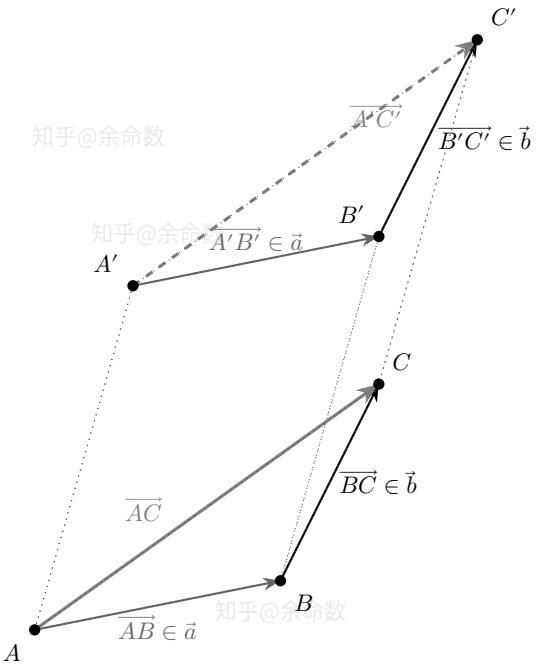


图 11.3　向量加法定义的良定义性

根据向量相等的定义, 有向线段等价于可通过平移重合. 因为 $\overrightarrow{AB}$ 与 $\overrightarrow{A'B'}$ 都是向量 $\vec{a}$ 的代表, 所以它们等价. 即四边形 $ABB'A'$ 是一个平行四边形. 由平行四边形的性质, 得 $\overrightarrow{AA'}$ 与 $\overrightarrow{BB'}$ 等价.

同理, 因为 $\overrightarrow{BC}$ 与 $\overrightarrow{B'C'}$ 都是向量 $\vec{b}$ 的代表, 所以它们等价. 即四边形 $BCC'B'$ 是一个平行四边形. 由此得 $\overrightarrow{BB'}$ 与 $\overrightarrow{CC'}$ 等价.

接着, 利用向量等价关系的传递性: 既然 $\overrightarrow{AA'}$ 等价于 $\overrightarrow{BB'}$, 且 $\overrightarrow{BB'}$ 等价于 $\overrightarrow{CC'}$, 那么 $\overrightarrow{AA'}$ 必然等价于 $\overrightarrow{CC'}$.

最后, $\overrightarrow{AA'}$ 与 $\overrightarrow{CC'}$ 等价, 这一结论的几何意义是四边形 $ACC'A'$ 是一个平行四边形. 根据平行四边形的性质, 其另一对边也必然平行且等长, 即 $\overrightarrow{AC}$ 与 $\overrightarrow{A'C'}$ 等价.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这就证明了, 无论如何选取向量 $\vec{a}$ 和 $\vec{b}$ 的代表, 最终得到的和向量的代表总是彼此等价的, 它们都属于同一个唯一的等价类. 因此, 向量加法的三角形法则是一个良定义的运算. □

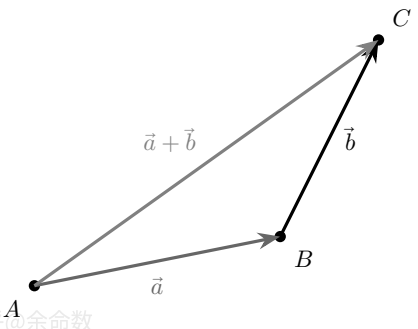


图 11.4 三角形法则: 向量 $\vec{a}$ 和 $\vec{b}$ 首尾相接, 其和为从 $\vec{a}$ 的起点到 $\vec{b}$ 的终点的向量.

当两个向量的作用是同时发生时, 例如, 共同作用于一点的两个力, 平行四边形法则提供了更直观的几何图像.

定理 11.2 (向量加法的平行四边形法则) 设两个向量 $\vec{a}$ 和 $\vec{b}$ 具有共同的起点 O. 以它们的代表 $\overrightarrow{OA}$ 和 $\overrightarrow{OB}$ 为邻边作平行四边形 $OACB$, 则从公共起点 O 出发的对角线 $\overrightarrow{OC}$ 代表它们的和 $\vec{a}+\vec{b}$.

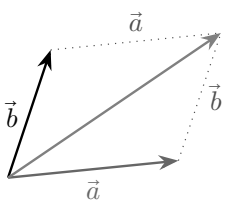
证明 平行四边形法则与三角形法则等价. 在平行四边形 $OACB$ 中, 对边平行且等长, 因而 $\overrightarrow{AC}$ 是向量 $\overrightarrow{OB}$ 的一个等价代表, 即 $\overrightarrow{AC}=\vec{b}$. 又 $\overrightarrow{OA}=\vec{a}$. 在 $\triangle OAC$ 中应用三角形法得

$$\overrightarrow{OC}=\overrightarrow{OA}+\overrightarrow{AC}=\vec{a}+\vec{b}.$$

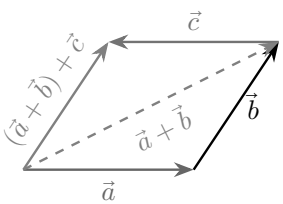
因此对角线 $\overrightarrow{OC}$ 代表向量和 $\vec{a}+\vec{b}$. □

向量加法如同实数加法一样, 满足以下运算律:

- **交换律**: $\vec{a}+\vec{b}=\vec{b}+\vec{a}$
- **结合律**: $(\vec{a}+\vec{b})+\vec{c}=\vec{a}+(\vec{b}+\vec{c})$



交换律: $\vec{a}+\vec{b}=\vec{b}+\vec{a}$



结合律

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

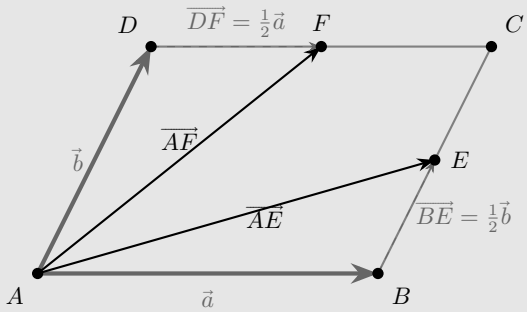
例题 11.2　化简向量表达式 $\overrightarrow{AD}+\overrightarrow{DE}+\overrightarrow{EB}$.

解　前一个向量的终点恰为后一个向量的起点, 形成首尾相接的链条. 由向量加法结合律,

$$\begin{aligned}(\overrightarrow{AD}+\overrightarrow{DE})+\overrightarrow{EB}&=\overrightarrow{AE}+\overrightarrow{EB}\\&=\overrightarrow{AB},\end{aligned}$$

故原表达式化简为 $\overrightarrow{AB}$. 一串首尾相接的向量之和, 即从第一个向量的起点直接指向最后一个向量终点的向量.

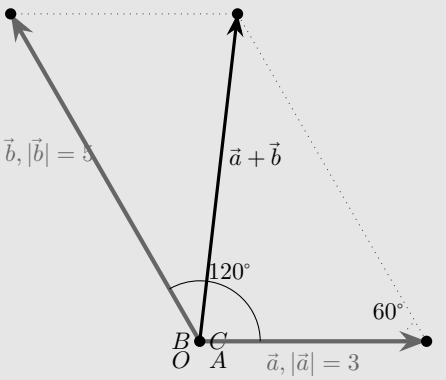
例题 11.3　在平行四边形 $ABCD$ 中, 点 E,F 分别是边 BC,DC 的中点. 若 $\overrightarrow{AB}=\vec{a},\overrightarrow{AD}=\vec{b}$, 试用 $\vec{a},\vec{b}$ 表示向量 $\overrightarrow{AE}$ 和 $\overrightarrow{AF}$.



解　此问题要求将目标向量沿一条由已知基向量组成的路径表示. 表示 $\overrightarrow{AE}$, 取路径 $A\to B\to E$. 由三角形法则 $\overrightarrow{AE}=\overrightarrow{AB}+\overrightarrow{BE}$, 已知 $\overrightarrow{AB}=\vec{a}$. 因 E 为 BC 中点, $\overrightarrow{BE}=\dfrac{1}{2}\overrightarrow{BC}$; 又平行四边形中 $\overrightarrow{BC}=\overrightarrow{AD}=\vec{b}$, 故 $\overrightarrow{BE}=\dfrac{1}{2}\vec{b}$, 代入得 $\overrightarrow{AE}=\vec{a}+\dfrac{1}{2}\vec{b}$. 表示 $\overrightarrow{AF}$, 取路径 $A\to D\to F$. 由三角形法则 $\overrightarrow{AF}=\overrightarrow{AD}+\overrightarrow{DF}$, 已知 $\overrightarrow{AD}=\vec{b}$. 因 F 为 DC 中点, $\overrightarrow{DF}=\dfrac{1}{2}\overrightarrow{DC}$; 又 $\overrightarrow{DC}=\overrightarrow{AB}=\vec{a}$, 故 $\overrightarrow{DF}=\dfrac{1}{2}\vec{a}$, 代入得 $\overrightarrow{AF}=\vec{b}+\dfrac{1}{2}\vec{a}$.

例题 11.4　已知两个非零向量 $\vec{a},\vec{b}$ 满足 $|\vec{a}|=3,|\vec{b}|=5$, 且它们的夹角为 120°. 求 $|\vec{a}+\vec{b}|$ 的值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



在 $\triangle OAC$ 中应用余弦定理

解 求向量和的模长, 是平行四边形法则的直接数量化应用. 设 $\overrightarrow{OA} = \vec{a}, \overrightarrow{OB} = \vec{b}$. 以 OA, OB 为邻边作平行四边形 $OACB$. 根据平行四边形法则, $\vec{a} + \vec{b} = \overrightarrow{OC}$. 所求 $|\vec{a} + \vec{b}|$ 就是对角线 OC 的长度.

在 $\triangle OAC$ 中, $OA = |\vec{a}| = 3$, $AC = OB = |\vec{b}| = 5$. 由 $OB \parallel AC$, $\angle OAC$ 与 $\angle AOB$ 互补, 即 $\angle OAC = 180^\circ - 120^\circ = 60^\circ$. 在 $\triangle OAC$ 中应用余弦定理求第三边 OC:

$$|\overrightarrow{OC}|^2 = |\overrightarrow{OA}|^2 + |\overrightarrow{AC}|^2 - 2|\overrightarrow{OA}||\overrightarrow{AC}|\cos(\angle OAC)$$

$$|\vec{a} + \vec{b}|^2 = 3^2 + 5^2 - 2 \cdot 3 \cdot 5 \cdot \cos(60^\circ)$$

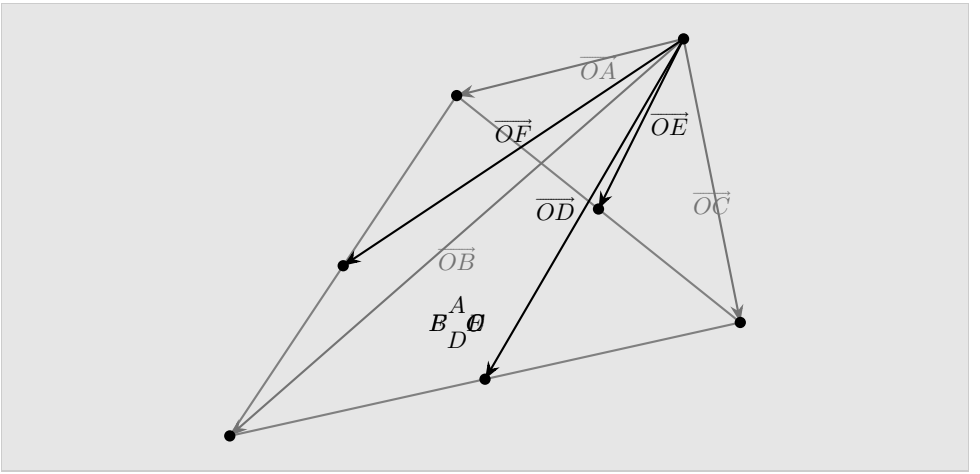
$$= 9 + 25 - 30 \cdot \tfrac{1}{2} = 19,$$

故 $|\vec{a} + \vec{b}| = \sqrt{19}$.

例题 11.5 设 O 为 $\triangle ABC$ 所在平面内任意一点, 点 D, E, F 分别是边 BC、CA、AB 的中点. 求证

$$\overrightarrow{OD} + \overrightarrow{OE} + \overrightarrow{OF} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}.$$

知乎@余命数



解 这是关于中点与任意基点的普适命题, 证明要点是把等式两边的向量都用以 O 为起点的向量表示. 由中线模型,

$$\overrightarrow{OD}=\tfrac{1}{2}(\overrightarrow{OB}+\overrightarrow{OC}),\quad \overrightarrow{OE}=\tfrac{1}{2}(\overrightarrow{OA}+\overrightarrow{OC}),\quad \overrightarrow{OF}=\tfrac{1}{2}(\overrightarrow{OA}+\overrightarrow{OB}).$$

三式相加得

$$\begin{aligned}\overrightarrow{OD}+\overrightarrow{OE}+\overrightarrow{OF}&=\tfrac{1}{2}(\overrightarrow{OB}+\overrightarrow{OC})+\tfrac{1}{2}(\overrightarrow{OA}+\overrightarrow{OC})+\tfrac{1}{2}(\overrightarrow{OA}+\overrightarrow{OB})\\&=\tfrac{1}{2}(2\overrightarrow{OA}+2\overrightarrow{OB}+2\overrightarrow{OC})\\&=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC},\end{aligned}$$

与右边相等, 命题得证.

11.1.2.3 向量减法

如果说向量加法回答了下一步去哪的问题, 那么向量减法则回答了如何从 B 到 A 的问题. 它描述了两个由同一起点出发的向量终点之间的相对位移.

定义 11.4 (向量减法) 向量 $\vec{a}$ 与 $\vec{b}$ 的差定义为 $\vec{a}$ 加上 $\vec{b}$ 的相反向量, 即

$$\vec{a}-\vec{b}=\vec{a}+(-\vec{b})$$

为了说明其基本的几何意义, 考虑共起点的向量 $\overrightarrow{OA}=\vec{a}$ 和 $\overrightarrow{OB}=\vec{b}$. 根据三角形法则, 有

$$\overrightarrow{OB}+\overrightarrow{BA}=\overrightarrow{OA}$$

移项可得

$$\overrightarrow{BA}=\overrightarrow{OA}-\overrightarrow{OB}=\vec{a}-\vec{b}$$

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

知乎@余命数

这一结论极为重要. **两个共起点向量之差, 是连接它们终点的向量, 方向由减向量**

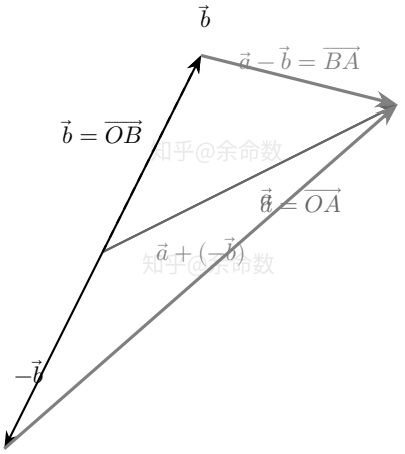
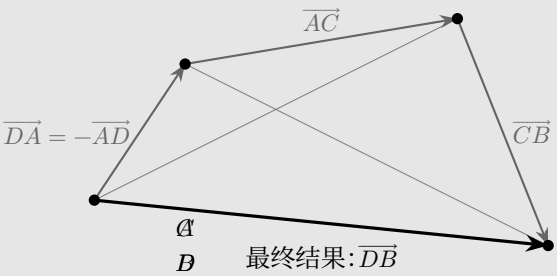


图 11.5 向量减法的几何意义. 左图: 共起点时, 差向量由减向量终点指向被减向量终点. 右图:其等价于加上相反向量.

的终点指向被减向量的终点.

例题 11.6 在四边形 $ABCD$ 中, 化简表达式 $\overrightarrow{AC}-\overrightarrow{AD}+\overrightarrow{CB}$.



解 处理加减混合运算, 先把减法转为加法, 减一向量即加其相反向量,

$$\overrightarrow{AC}-\overrightarrow{AD}+\overrightarrow{CB}=\overrightarrow{AC}+\overrightarrow{DA}+\overrightarrow{CB}.$$

由向量加法交换律重排各项使之首尾相接,

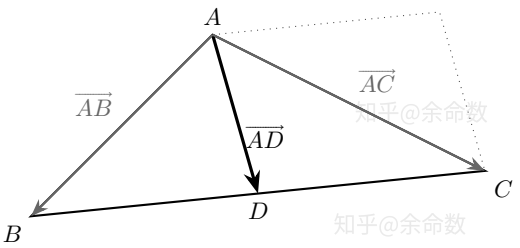
$$\begin{aligned}\overrightarrow{DA}+\overrightarrow{AC}+\overrightarrow{CB}&=(\overrightarrow{DA}+\overrightarrow{AC})+\overrightarrow{CB}\\&=\overrightarrow{DC}+\overrightarrow{CB}=\overrightarrow{DB},\end{aligned}$$

故原式化简为 $\overrightarrow{DB}$.

在 $\triangle ABC$ 中, 若 D 是边 BC 的中点, 则

$$\overrightarrow{AB}+\overrightarrow{AC}=2\overrightarrow{AD} \tag{11-1}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

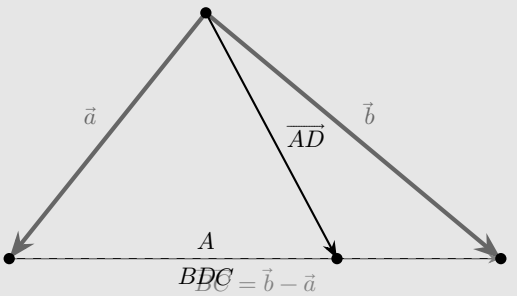


证明 在 $\triangle ABD$ 中, 有 $\overrightarrow{AB}=\overrightarrow{AD}+\overrightarrow{DB}$. 在 $\triangle ACD$ 中, 有 $\overrightarrow{AC}=\overrightarrow{AD}+\overrightarrow{DC}$. 将上述两式相加:

$$\begin{aligned}\overrightarrow{AB}+\overrightarrow{AC}&=(\overrightarrow{AD}+\overrightarrow{DB})+(\overrightarrow{AD}+\overrightarrow{DC})\\&=2\overrightarrow{AD}+(\overrightarrow{DB}+\overrightarrow{DC})\end{aligned}$$

因为 D 是线段 BC 的中点, 向量 $\overrightarrow{DB}$ 与 $\overrightarrow{DC}$ 是相反向量. 因此, $\overrightarrow{DB}+\overrightarrow{DC}=\vec{0}$. 代入即得 $\overrightarrow{AB}+\overrightarrow{AC}=2\overrightarrow{AD}$. □

例题 11.7 在 $\triangle ABC$ 中, 点 D 在边 BC 上, 且满足 $BD=2DC$. 设 $\overrightarrow{AB}=\vec{a},\overrightarrow{AC}=\vec{b}$, 试用 $\vec{a},\vec{b}$ 表示向量 $\overrightarrow{AD}$.



解 取路径 $A\to B\to D$, 由向量加法法则 $\overrightarrow{AD}=\overrightarrow{AB}+\overrightarrow{BD}$, 其中 $\overrightarrow{AB}=\vec{a}$ 已知, 只需用基向量表示 $\overrightarrow{BD}$. 点 D 在线段 BC 上, 条件 $BD=2DC$ 给出 $BD=\dfrac{2}{3}BC$, 即 $\overrightarrow{BD}=\dfrac{2}{3}\overrightarrow{BC}$. 由向量减法的几何意义 $\overrightarrow{BC}=\overrightarrow{AC}-\overrightarrow{AB}=\vec{b}-\vec{a}$,代入得 $\overrightarrow{BD}=\dfrac{2}{3}(\vec{b}-\vec{a})$. 回代到 $\overrightarrow{AD}$ 中,

$$\begin{aligned}\overrightarrow{AD}&=\overrightarrow{AB}+\overrightarrow{BD}=\vec{a}+\tfrac{2}{3}(\vec{b}-\vec{a})\\&=\vec{a}+\tfrac{2}{3}\vec{b}-\tfrac{2}{3}\vec{a}=\tfrac{1}{3}\vec{a}+\tfrac{2}{3}\vec{b}.\end{aligned}$$

11.1.2.4 标量乘法

标量乘法定义了一个实数 λ 与一个向量 $\vec{a}$ 之间的运算. 其结果 $\lambda\vec{a}$ 仍是一个向量, 几何上表示对原向量 $\vec{a}$ 进行沿其自身方向的伸缩或反向.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 11.5 (标量乘法) 实数 λ 与向量 $\vec{a}$ 的乘积 $\lambda\vec{a}$ 是一个向量, 满足:

1. **模**: $|\lambda\vec{a}| = |\lambda||\vec{a}|$.
2. **方向**:
   - 若 $\lambda > 0$, $\lambda\vec{a}$ 的方向与 $\vec{a}$ 相同.
   - 若 $\lambda < 0$, $\lambda\vec{a}$ 的方向与 $\vec{a}$ 相反.
   - 若 $\lambda = 0$, $\lambda\vec{a} = \vec{0}$.



图 11.6 标量乘法对向量 $\vec{a}$ 的作用.

标量乘法与向量加法共同满足分配律, 这建立起向量代数与实数代数之间的联系:

- $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$
- $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$

这些线性运算共同构成了向量空间的基本结构, 是后续所有解析几何理论的代数基础.

11.1.2.5 分析向量的除法

在建立了向量的加、减、数乘以及两种乘积(数量积与向量积)之后, 能否定义向量的除法成为自然的问题. 要回答它, 须先回到除法运算的根本意义.

在实数域中, 除法是乘法的逆运算.a 除以 b (a/b), 就是寻找一个唯一的数 x, 使得

$$b \cdot x = a$$

若对于任意的 a 和非零的 b, 都能找到这样一个唯一的 x, 则此代数系统中的除法是**良定义的**.

将这个思想模板套用到向量上, 定义"向量 $\vec{a}$ 除以向量 $\vec{b}$", 就意味着我们要寻找一个**唯**一的向量 $\vec{x}$, 使得

$$\vec{b} \text{ 某种符号 } \vec{x} = \vec{a}$$

其中某种符号代表一种向量乘法. 然而, 向量的乘法有两种截然不同的形式, 即数量积 ($\cdot$) 与向量积 ($\times$), 问题因此分为两路, 分别考察这两种乘法是否允许良定义的逆运算.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

先尝试基于数量积来定义除法. 为此需要解方程:

$$\vec{b}\cdot\vec{x}=\vec{a}$$

立即遇到一个无法逾越的障碍, **类型不匹配**. 方程的左侧, $\vec{b}\cdot\vec{x}$, 根据定义是一个**标量**(一个实数). 而方程的右侧, $\vec{a}$, 是一个**向量**. 一个标量永远不可能等于一个非零向量, 这在根本上是矛盾的.

为了让探索继续, 不妨调整问题, 改用一个标量 k 除以一个向量 $\vec{b}$, 这对应于求解关于向量 $\vec{x}$ 的方程.

$$\vec{b}\cdot\vec{x}=k$$

这个方程在代数类型上是自洽的. 根据数量积的几何解释, 它要求向量 $\vec{x}$ 在向量 $\vec{b}$ 方向上的投影是一个固定的值. 设 $\vec{b}$ 的方向上的单位向量为 $\vec{e}_{\vec{b}}$, 则投影长度为

$$|\vec{x}|\cos\theta=\frac{k}{|\vec{b}|}$$

其中 θ 是 $\vec{b}$ 与 $\vec{x}$ 的夹角. 如图所示, 所有满足条件的向量 $\vec{x}$ 的终点构成了一个垂

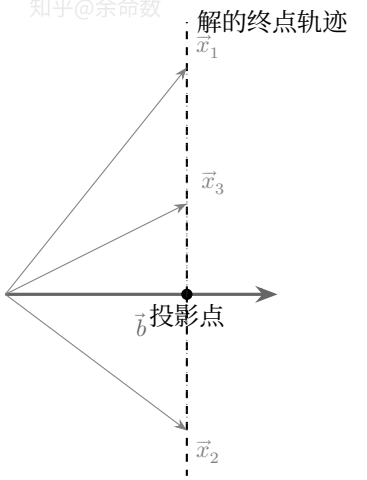


图 11.7　方程 $\vec{b}\cdot\vec{x}=k$ 的解. 任何以 O 为起点, 终点落在垂直于 $\vec{b}$ 的特定直线, 或平面上的向量 $\vec{x}$ 都是该方程的解.

直于 $\vec{b}$ 的直线或平面, 分别对应二维空间与三维空间. 即方程的解有**无穷多个**. 由于除法要求解是唯一的, **数量积不存在逆运算**. 它在运算中丢失了关于 $\vec{x}$ 垂直于 $\vec{b}$ 的分量信息, 这种信息损失是不可逆的.

接着, 尝试基于向量积来定义除法. 为此需要解方程:

$$\vec{b}\times\vec{x}=\vec{a}$$

这一次, 方程两边的类型是匹配的, 向量积的结果是一个向量. 这看起来更有希望.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

然而, 先面临一个**可解性的限制**. 根据向量积的定义, 向量 $\vec{b}\times\vec{x}$ 必然同时垂直于 $\vec{b}$ 和 $\vec{x}$. 即, 等式右边的向量 $\vec{a}$ 必须也垂直于 $\vec{b}$, 即 $\vec{a}\cdot\vec{b}=0$. 如果 $\vec{a}$ 不垂直于 $\vec{b}$, 那么这个方程**根本没有解**. 一个良定义的除法不应该只对极少数特殊情况有解.

即使满足可解性条件 $\vec{a}\cdot\vec{b}=0$, 解的唯一性问题依然存在. 假设我们找到了一个解 $\vec{x}_0$, 使得 $\vec{b}\times\vec{x}_0=\vec{a}$. 接着构造一个新的向量 $\vec{x}=\vec{x}_0+\lambda\vec{b}$, 其中 λ 是任意实数. 计算 $\vec{b}\times\vec{x}$:

$$\begin{aligned}\vec{b}\times\vec{x}&=\vec{b}\times(\vec{x}_0+\lambda\vec{b})\\&=(\vec{b}\times\vec{x}_0)+(\vec{b}\times\lambda\vec{b})\quad\text{(向量积对加法满足分配律)}\\&=\vec{a}+\lambda(\vec{b}\times\vec{b})\end{aligned}$$

由于任何向量与自身的向量积都是零向量 ($\vec{b}\times\vec{b}=\vec{0}$), 上式变为:

$$\vec{b}\times\vec{x}=\vec{a}+\vec{0}=\vec{a}$$

故, 如果 $\vec{x}_0$ 是一个解, 那么所有形如 $\vec{x}_0+\lambda\vec{b}$ 的向量都是解! 几何上, 所有的解向

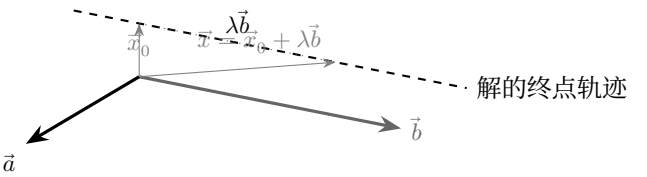


图 11.8 方程 $\vec{b}\times\vec{x}=\vec{a}$ 的解. 若 $\vec{x}_0$ 是一个解, 则所有终点在过 $\vec{x}_0$ 终点且平行于 $\vec{b}$ 的直线上的向量都是解.

量的终点构成了一条平行于向量 $\vec{b}$ 的直线. 再次得到**无穷多个解**. 其根源在于, 向量积运算存在**零因子**, 即两个非零向量的向量积可以是零向量, 当它们共线时. 这与实数乘法中"$ab=0\implies a=0\text{ or }b=0$"的性质截然不同. 拥有零因子的代数结构通常都不支持除法.

于是, 无论基于数量积还是向量积, 都无法建立起一个良定义的向量除法. 其根本原因在于, 两种乘法都是多对一的映射, 它们在运算过程中都不可避免地丢失了输入向量的某些信息, 使得逆向求解无法得到唯一的结果. 且向量代数系统, 尤其是向量积, 包含零因子, 这从根本上破坏了除法所依赖的唯一性.

11.1.3 共线定理与平面向量基本定理

上述运算共同构成了向量的线性结构, 要点是共线与线性组合的概念.

11.1.3.1 向量共线定理

向量 $\vec{a}$ ($\vec{a}\neq\vec{0}$) 与向量 $\vec{b}$ 共线的充分必要条件是, 存在**唯**一的实数 λ, 使得

$$\vec{b}=\lambda\vec{a}\tag{11-2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 **必要性（由共线 $\Longrightarrow$ 存在 λ）**假设向量 $\vec{a}$ 与 $\vec{b}$ 共线，且 $\vec{a}\neq\vec{0}$. 若 $\vec{b}=\vec{0}$，取 $\lambda=0$ 即可. 若 $\vec{b}\neq\vec{0}$，且与 $\vec{a}$ 方向相同，则其单位向量相同，$\frac{\vec{a}}{|\vec{a}|}=\frac{\vec{b}}{|\vec{b}|}$. 整理得 $\vec{b}=\frac{|\vec{b}|}{|\vec{a}|}\vec{a}$. 令 $\lambda=\frac{|\vec{b}|}{|\vec{a}|}>0$. 若与 $\vec{a}$ 方向相反，则单位向量互为相反，$\frac{\vec{b}}{|\vec{b}|}=-\frac{\vec{a}}{|\vec{a}|}$. 整理得 $\vec{b}=-\frac{|\vec{b}|}{|\vec{a}|}\vec{a}$. 令 $\lambda=-\frac{|\vec{b}|}{|\vec{a}|}<0$.

充分性（由存在 λ $\Longrightarrow$ 共线）假设存在实数 λ，使得 $\vec{b}=\lambda\vec{a}$. 根据向量数乘的定义，向量 $\vec{b}$ 的方向或者与 $\vec{a}$ 相同 ($\lambda>0$)，或者相反 ($\lambda<0$)，或者是零向量 ($\lambda=0$). 任何情况下，$\vec{a}$ 与 $\vec{b}$ 均共线.

唯一性假设存在 $\lambda_1\neq\lambda_2$ 使得 $\vec{b}=\lambda_1\vec{a}$ 且 $\vec{b}=\lambda_2\vec{a}$. 则 $\lambda_1\vec{a}=\lambda_2\vec{a}$，故 $(\lambda_1-\lambda_2)\vec{a}=\vec{0}$. 由于 $\vec{a}\neq\vec{0}$，必有 $\lambda_1-\lambda_2=0$，即 $\lambda_1=\lambda_2$. 这与假设矛盾. □

例题 11.8 判断命题 p“存在实数 k，使得 $\vec{a}=k\vec{b}$”是命题 q“向量 $\vec{a}$ 与 $\vec{b}$ 共线”的什么条件？

解 本题旨在辨析向量共线的代数表示与其几何定义之间的细微差别，其陷阱在于对零向量的考虑.

若存在实数 k 使得 $\vec{a}=k\vec{b}$，根据数乘的几何意义，$\vec{a}$ 与 $\vec{b}$ 必然共线. 所以，p 是 q 的充分条件.

反之，若向量 $\vec{a}$ 与 $\vec{b}$ 共线，可取反例. 当 $\vec{b}=\vec{0}$ 且 $\vec{a}\neq\vec{0}$ 时，根据定义，任何非零向量都与零向量共线，所以此时命题 q 为真. 然而，不存在任何实数 k 能够使得一个非零向量 $\vec{a}$ 等于 $k\vec{0}=\vec{0}$. 即此时命题 p 为假.

由于找到一个 q 为真而 p 为假的情况，所以 q 不能推出 p. p 不是 q 的必要条件. 命题 p 是命题 q 的**充分不必要条件**.

此定理将几何上的共线关系，转化为代数上的倍数关系，是进行向量论证和计算的基础.

11.1.3.2 平面向量基本定理

当将视野扩展到整个平面，能否用少数几个基本向量的线性运算表示出平面内的所有向量，成为自然的问题.

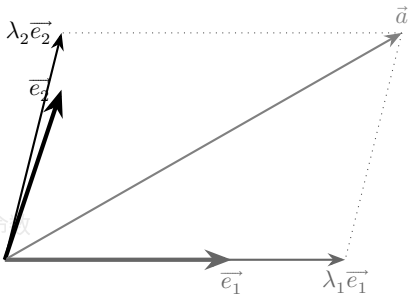
定理 11.3 (平面向量基本定理) 如果 $\overrightarrow{e_1}$ 和 $\overrightarrow{e_2}$ 是同一平面内的两个**不共线**向量，那么对于该平面内的**任意**一个向量 $\vec{a}$，存在**唯**一的一对实数 λ_1,λ_2，使得

$$\vec{a}=\lambda_1\overrightarrow{e_1}+\lambda_2\overrightarrow{e_2} \tag{11-3}$$

我们称不共线的向量 $\overrightarrow{e_1},\overrightarrow{e_2}$ 为表示该平面内所有向量的一组**基底**.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

证明 **存在性**: 取诸向量共起点的代表 $\overrightarrow{OE_1}=\overrightarrow{e_1},\overrightarrow{OE_2}=\overrightarrow{e_2},\overrightarrow{OA}=\vec{a}$. 过点 A 作直线平行于 OE_2, 交直线 OE_1 于 A_1; 再过点 A 作直线平行于 OE_1, 交直线 OE_2 于 A_2.

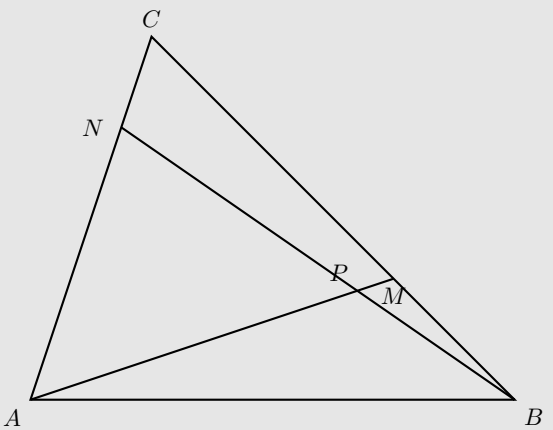


由平行四边形法则,$\vec{a}=\overrightarrow{OA_1}+\overrightarrow{OA_2}$. 由于 $\overrightarrow{OA_1}$ 与 $\overrightarrow{e_1}$ 共线,故存在 λ_1 使 $\overrightarrow{OA_1}=\lambda_1\overrightarrow{e_1}$. 同理存在 λ_2 使 $\overrightarrow{OA_2}=\lambda_2\overrightarrow{e_2}$. 故 $\vec{a}=\lambda_1\overrightarrow{e_1}+\lambda_2\overrightarrow{e_2}$.

唯一性: 假设另有一对实数 (μ_1,μ_2) 满足要求. 则 $\lambda_1\overrightarrow{e_1}+\lambda_2\overrightarrow{e_2}=\mu_1\overrightarrow{e_1}+\mu_2\overrightarrow{e_2}$, 整理得 $(\lambda_1-\mu_1)\overrightarrow{e_1}=(\mu_2-\lambda_2)\overrightarrow{e_2}$. 由于 $\overrightarrow{e_1},\overrightarrow{e_2}$ 不共线, 此等式成立的唯一可能是系数均为零, 即 $\lambda_1=\mu_1$ 且 $\lambda_2=\mu_2$. □

这一定理的重大意义在于, 它为平面内的所有向量建立了一个坐标系. 任何向量都可以在这个坐标系下被唯一地分解.

例题 11.9 在 $\triangle ABC$ 中, 点 M 在边 BC 上, 满足 $BM:MC=1:2$. 点 N 在边 AC 上, 满足 $AN:NC=3:1$. 线段 AM 与 BN 交于点 P. 求 $AP:PM$ 的值.



解 此题是关于共线和交点问题的经典模型, 思路如下: **对于同一个向量, 其在同一组基底下的表示是唯一的.**

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

以点 A 为原点, $\overrightarrow{AB}=\vec{b}$ 和 $\overrightarrow{AC}=\vec{c}$ 为基向量. 由于 A,P,M 三点共线, 向量 $\overrightarrow{AP}$ 与 $\overrightarrow{AM}$ 共线. 因此存在实数 λ, 使得 $\overrightarrow{AP}=\lambda\overrightarrow{AM}$. 由定比分点公式 $\overrightarrow{AM}=\dfrac{2\overrightarrow{AB}+1\overrightarrow{AC}}{1+2}=\dfrac{2}{3}\vec{b}+\dfrac{1}{3}\vec{c}$, 代入得 $\overrightarrow{AP}$ 的第一个表达式

$$\overrightarrow{AP}=\lambda\left(\frac{2}{3}\vec{b}+\frac{1}{3}\vec{c}\right)=\frac{2\lambda}{3}\vec{b}+\frac{\lambda}{3}\vec{c}\quad\cdots\cdots(1).$$

又 B,P,N 三点共线, 向量 $\overrightarrow{BP}$ 与 $\overrightarrow{BN}$ 共线, 故存在实数 μ 使 $\overrightarrow{BP}=\mu\overrightarrow{BN}$. 由向量减法 $\overrightarrow{AP}-\overrightarrow{AB}=\mu(\overrightarrow{AN}-\overrightarrow{AB})$, 整理得 $\overrightarrow{AP}=(1-\mu)\overrightarrow{AB}+\mu\overrightarrow{AN}$. 已知 $\overrightarrow{AB}=\vec{b}$, $\overrightarrow{AN}=\dfrac{3}{4}\overrightarrow{AC}=\dfrac{3}{4}\vec{c}$, 代入得第二个表达式

$$\overrightarrow{AP}=(1-\mu)\vec{b}+\frac{3\mu}{4}\vec{c}\quad\cdots\cdots(2).$$

由平面向量基本定理, $\overrightarrow{AP}$ 在基底 $\{\vec{b},\vec{c}\}$ 下的表示唯一, 故 (1) 与 (2) 中 $\vec{b},\vec{c}$ 的系数分别相等,

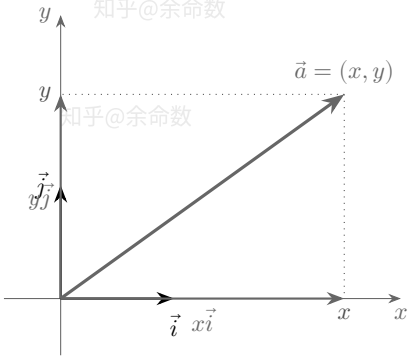
$$\begin{cases}\frac{2\lambda}{3}=1-\mu\\ \frac{\lambda}{3}=\frac{3\mu}{4}\end{cases}$$

这是一个关于 λ,μ 的二元一次方程组. 解得 $\lambda=\frac{9}{10}$.

$\overrightarrow{AP}=\frac{9}{10}\overrightarrow{AM}$ 意味着 $AP=\frac{9}{10}AM$. 因此 $PM=\frac{1}{10}AM$, 故 $AP:PM=9:1$.

11.1.4 向量的坐标表示

为了计算的方便, 通常选一组最特殊基底. 在直角坐标系中, 分别与 x 轴和 y 轴正方向相同, 且长度为 1 的单位向量, 记作 $\vec{i}$ 和 $\vec{j}$.



根据平面向量基本定理, 任何向量 $\vec{a}$ 都可以唯一地表示为 $\vec{a}=x\vec{i}+y\vec{j}$. 我们便将这对唯一的实数 (x,y) 称为向量 $\vec{a}$ 的**坐标**, 记作

$$\vec{a}=(x,y)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

向量的运算也就相应地转化为其坐标的运算.

向量的坐标运算

设 $\vec{a}=(x_1,y_1),\vec{b}=(x_2,y_2)$, 则

- **加法** $\vec{a}+\vec{b}=(x_1+x_2,y_1+y_2)$
- **减法** $\vec{a}-\vec{b}=(x_1-x_2,y_1-y_2)$
- **数乘** $k\vec{a}=(kx_1,ky_1)$
- **模** $|\vec{a}|=\sqrt{x_1^2+y_1^2}$

证明 这些运算规则把向量代数与解析几何联系起来. 其证明的基础在于平面向量基本定理, 即任何一个向量都可以唯一地表示为一组基向量的线性组合. 在所选取的标准笛卡尔坐标系中, 这组基底是相互正交的单位向量 $\vec{i}$ 和 $\vec{j}$.

设 $\vec{a}=(x_1,y_1)$ 和 $\vec{b}=(x_2,y_2)$. 根据坐标的定义, 这等价于

$$\vec{a}=x_1\vec{i}+y_1\vec{j}\quad\text{以及}\quad\vec{b}=x_2\vec{i}+y_2\vec{j}$$

知乎@余命数

从向量加法的抽象定义出发, 利用向量运算的代数性质推导.

$$\begin{aligned}\vec{a}+\vec{b}&=(x_1\vec{i}+y_1\vec{j})+(x_2\vec{i}+y_2\vec{j})\\&=(x_1\vec{i}+x_2\vec{i})+(y_1\vec{j}+y_2\vec{j})&&\text{(向量加法的交换律与结合律)}\\&=(x_1+x_2)\vec{i}+(y_1+y_2)\vec{j}&&\text{(标量乘法对加法的分配律)}\end{aligned}$$

根据平面向量基本定理的唯一性, 向量和 $\vec{a}+\vec{b}$ 在基底 $\{\vec{i},\vec{j}\}$ 下的坐标是唯一的, 即 (x_1+x_2,y_1+y_2).

同理, 向量减法可以看作是加上一个相反向量.

$$\begin{aligned}\vec{a}-\vec{b}&=(x_1\vec{i}+y_1\vec{j})-(x_2\vec{i}+y_2\vec{j})\\&=(x_1\vec{i}-x_2\vec{i})+(y_1\vec{j}-y_2\vec{j})\\&=(x_1-x_2)\vec{i}+(y_1-y_2)\vec{j}\end{aligned}$$

知乎@余命数
知乎@余命数
知乎@余命数

因此, 差向量 $\vec{a}-\vec{b}$ 的坐标为 (x_1-x_2,y_1-y_2).

设 $k\in\mathbb{R}$ 为一个标量.

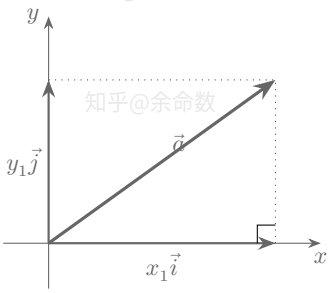
$$\begin{aligned}k\vec{a}&=k(x_1\vec{i}+y_1\vec{j})\\&=k(x_1\vec{i})+k(y_1\vec{j})&&\text{(标量乘法对向量加法的分配律)}\\&=(kx_1)\vec{i}+(ky_1)\vec{j}&&\text{(标量乘法的结合律)}\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此, 向量 $k\vec{a}$ 的坐标为 (kx_1, ky_1).

向量 $\vec{a}$ 的模是其几何长度. 向量 $\vec{a} = x_1\vec{i} + y_1\vec{j}$ 可以看作是两个向量 $x_1\vec{i}$ 和 $y_1\vec{j}$ 的和. 向量 $x_1\vec{i}$ 位于 x 轴上, 其长度为 $|x_1\vec{i}| = |x_1||\vec{i}| = |x_1|$. 向量 $y_1\vec{j}$ 位于 y 轴上, 其长度为 $|y_1\vec{j}| = |y_1||\vec{j}| = |y_1|$. 由于基向量 $\vec{i}$ 和 $\vec{j}$ 相互垂直, 这两个分量向量也相互垂直.

知乎@余命数



根据向量加法的平行四边形法则(在此处退化为矩形), 向量 $\vec{a}$ 是以 $x_1\vec{i}$ 和 $y_1\vec{j}$ 为邻边的矩形的对角线. 根据勾股定理, 该对角线的长度, 即 $\vec{a}$ 的模为:

$$|\vec{a}|^2 = |x_1\vec{i}|^2 + |y_1\vec{j}|^2 = x_1^2 + y_1^2$$

知乎@余命数

因此, $|\vec{a}| = \sqrt{x_1^2 + y_1^2}$. □

知乎@余命数

一个基本的结论是, 如果两点坐标为 $A(x_1, y_1)$ 和 $B(x_2, y_2)$, 那么由 A 指向 B 的向量 $\overrightarrow{AB}$ 的坐标为

$$\overrightarrow{AB} = (x_2 - x_1, y_2 - y_1)$$

即"**终点坐标减去起点坐标**".

例题 11.10　已知点 $A(-1,2), B(2,1), C(3,4)$, 求向量 $\vec{u} = 2\overrightarrow{AB} - \overrightarrow{BC}$ 的坐标及其模.

解　首先, 根据终点减起点法则求出基准向量的坐标.

$$\overrightarrow{AB} = (2-(-1), 1-2) = (3,-1)$$

知乎@余命数

$$\overrightarrow{BC} = (3-2, 4-1) = (1,3)$$

知乎@余命数

接下来, 进行向量的线性运算.

$$\vec{u} = 2\overrightarrow{AB} - \overrightarrow{BC} = 2(3,-1) - (1,3)$$
$$= (6,-2) - (1,3) = (5,-5)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

最后, 求出向量 $\vec{u}$ 的模.

$$|\vec{u}| = \sqrt{5^2 + (-5)^2} = \sqrt{50} = 5\sqrt{2}$$

例题 11.11 平面内向量 $\overrightarrow{OA} = (1,7), \overrightarrow{OB} = (5,1), \overrightarrow{OP} = (2,1)$. 点 X 为直线 OP 上的一个动点.

(1) 当 $\overrightarrow{XA} \cdot \overrightarrow{XB}$ 取最小值时, 求 $\overrightarrow{OX}$ 的坐标;

(2) 当点 X 满足 (1) 的条件和结论时, 求 $\cos\angle AXB$ 的值.

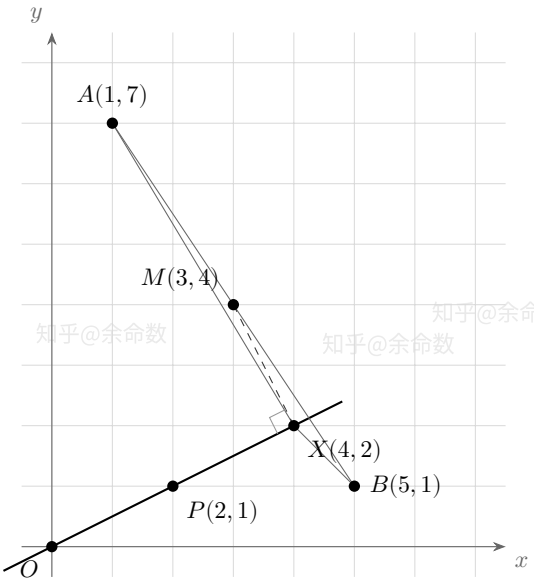


图 11.9 向量关系几何示意图

解 (1) 先对目标数量积 $\overrightarrow{XA} \cdot \overrightarrow{XB}$ 进行几何转化. 引入线段 AB 的中点 M. 则 $\overrightarrow{XA} = \overrightarrow{XM} + \overrightarrow{MA}$, 且 $\overrightarrow{XB} = \overrightarrow{XM} + \overrightarrow{MB}$. 由于 M 是中点, $\overrightarrow{MB} = -\overrightarrow{MA}$. 代入数量积表达式:

$$\overrightarrow{XA} \cdot \overrightarrow{XB} = (\overrightarrow{XM} + \overrightarrow{MA}) \cdot (\overrightarrow{XM} - \overrightarrow{MA}) = |\overrightarrow{XM}|^2 - |\overrightarrow{MA}|^2$$

注意到, 点 A, B 是定点, 故中点 M 和线段长 $|\overrightarrow{MA}|$ 均为定值. 因此, 欲使 $\overrightarrow{XA} \cdot \overrightarrow{XB}$ 取得最小值, 等价于使 $|\overrightarrow{XM}|^2$ 取得最小值, 这也等价于使点 X 到定点 M 的距离 $|\overrightarrow{XM}|$ 最小.

点 X 是直线 OP 上的动点, 故此问题转化为寻找直线上一点 X, 使其到直线外一点 M 的距离最短. 根据平面几何的基本定理, 这个最短距离在 X 为 M 在直线 OP 上的正交投影点时取得. 此时, 向量 $\overrightarrow{MX}$ 必须与直线 OP 的方向向量 $\overrightarrow{OP}$ 垂直.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

首先, 计算中点 M 的坐标, 得

$$\overrightarrow{OM}=\frac{1}{2}(\overrightarrow{OA}+\overrightarrow{OB})=\frac{1}{2}((1,7)+(5,1))=\frac{1}{2}(6,8)=(3,4)$$

由于点 X 在直线 OP 上, 存在唯一的实数 t, 使得 $\overrightarrow{OX}=t\overrightarrow{OP}=t(2,1)=(2t,t)$. 向量 $\overrightarrow{MX}$ 可以表示为:

$$\overrightarrow{MX}=\overrightarrow{OX}-\overrightarrow{OM}=(2t,t)-(3,4)=(2t-3,t-4)$$

施加垂直条件 $\overrightarrow{MX}\cdot\overrightarrow{OP}=0$:

$$(2t-3,t-4)\cdot(2,1)=0$$

$$2(2t-3)+1(t-4)=0\implies 4t-6+t-4=0\implies 5t=10\implies t=2$$

当 $t=2$ 时, $\overrightarrow{XA}\cdot\overrightarrow{XB}$ 取得最小值. 此时, $\overrightarrow{OX}$ 的坐标为:

$$\overrightarrow{OX}=2\overrightarrow{OP}=2(2,1)=(4,2)$$

(2) 根据 (1) 的结论, 取点 X 的坐标为 $(4,2)$. 任务是计算 $\cos\angle AXB$. 根据数量积的夹角公式:

$$\cos\angle AXB=\frac{\overrightarrow{XA}\cdot\overrightarrow{XB}}{|\overrightarrow{XA}||\overrightarrow{XB}|}$$

先计算分子和分母所需的各个向量与模长.

$$\overrightarrow{XA}=\overrightarrow{OA}-\overrightarrow{OX}=(1,7)-(4,2)=(-3,5)$$

$$\overrightarrow{XB}=\overrightarrow{OB}-\overrightarrow{OX}=(5,1)-(4,2)=(1,-1)$$

接着计算数量积, 得

$$\overrightarrow{XA}\cdot\overrightarrow{XB}=(-3)(1)+(5)(-1)=-3-5=-8$$

计算模长, 得

$$|\overrightarrow{XA}|=\sqrt{(-3)^2+5^2}=\sqrt{9+25}=\sqrt{34}$$

$$|\overrightarrow{XB}|=\sqrt{1^2+(-1)^2}=\sqrt{1+1}=\sqrt{2}$$

最后, 将这些值代入夹角公式:

$$\cos\angle AXB=\frac{-8}{\sqrt{34}\sqrt{2}}=\frac{-8}{\sqrt{68}}=\frac{-8}{2\sqrt{17}}=-\frac{4}{\sqrt{17}}=-\frac{4\sqrt{17}}{17}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

11.1.5 平面向量的数量积

已定义向量的加、减、数乘运算, 其结果仍然是向量. 接下来要考虑两个向量能否相乘, 若能, 乘积是向量还是标量. 物理学给出基本启示. 当一个力 F 作用于一个物体, 使其产生位移 s 时, 力所做的功 W 是一个标量, 其大小同时取决于力的大小、位移的大小及它们之间的夹角. 具体地, $W=|\vec{F}||\vec{s}|\cos\theta$. 这个例子启发我们定义一种新的向量乘法, 其结果是一个标量, 故称其为**数量积**, 又因其符号形如点, 也称为**点积**.

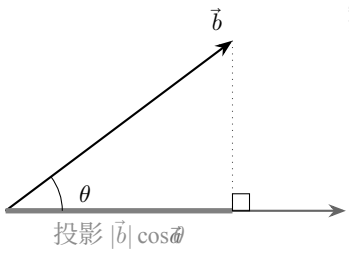
11.1.5.1 数量积的定义与几何意义

定义 11.6 (数量积) 已知两个非零向量 $\vec{a}$ 和 $\vec{b}$, 它们的夹角为 θ. 我们定义它们的数量积为一个**标量**.

$$\vec{a}\cdot\vec{b}=|\vec{a}||\vec{b}|\cos\theta \tag{11-4}$$

如果其中一个向量为零向量, 则数量积为 0.

数量积 $\vec{a}\cdot\vec{b}$ 的几何意义是, 向量 $\vec{a}$ 的模 $|\vec{a}|$ 与向量 $\vec{b}$ 在 $\vec{a}$ 方向上的**投影** $|\vec{b}|\cos\theta$ 的乘积.



11.1.5.2 数量积的性质与坐标表示

- $\vec{a}\cdot\vec{a}=|\vec{a}|^2$ 或 $|\vec{a}|=\sqrt{\vec{a}\cdot\vec{a}}$.
- $\vec{a}\cdot\vec{b}=\vec{b}\cdot\vec{a}$(交换律).
- $(\lambda\vec{a})\cdot\vec{b}=\lambda(\vec{a}\cdot\vec{b})$.
- $(\vec{a}+\vec{b})\cdot\vec{c}=\vec{a}\cdot\vec{c}+\vec{b}\cdot\vec{c}$(分配律).

注 11.1 (垂直的判定)
判断两个非零向量是否垂直, 是数量积最重要的应用之一.

$$\vec{a}\perp\vec{b}\iff\theta=90^\circ\iff\cos\theta=0\iff\vec{a}\cdot\vec{b}=0$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设基向量 $\vec{i}=(1,0),\vec{j}=(0,1)$. 显然 $|\vec{i}|=|\vec{j}|=1,\vec{i}\cdot\vec{j}=0$. 对于 $\vec{a}=(x_1,y_1)=x_1\vec{i}+y_1\vec{j}$ 和 $\vec{b}=(x_2,y_2)=x_2\vec{i}+y_2\vec{j}$, 有

$$\begin{aligned}\vec{a}\cdot\vec{b}&=(x_1\vec{i}+y_1\vec{j})\cdot(x_2\vec{i}+y_2\vec{j})\\&=x_1x_2(\vec{i}\cdot\vec{i})+x_1y_2(\vec{i}\cdot\vec{j})+y_1x_2(\vec{j}\cdot\vec{i})+y_1y_2(\vec{j}\cdot\vec{j})\\&=x_1x_2(1)+x_1y_2(0)+y_1x_2(0)+y_1y_2(1)\\&=x_1x_2+y_1y_2\end{aligned}$$

知乎@余命数

得数量积的坐标计算公式

$$\vec{a}\cdot\vec{b}=x_1x_2+y_1y_2 \tag{11-5}$$

由此, 可推导出坐标形式下的模长和夹角公式.

- **模长** $|\vec{a}|=\sqrt{x_1^2+y_1^2}$
- **夹角** $\cos\theta=\dfrac{\vec{a}\cdot\vec{b}}{|\vec{a}||\vec{b}|}=\dfrac{x_1x_2+y_1y_2}{\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2}}$

例题 11.12　已知向量 $\vec{a},\vec{b}$ 满足 $|\vec{a}|=2,|\vec{b}|=3$, 且它们的数量积 $\vec{a}\cdot\vec{b}=-3$. 求向量差 $\vec{a}-2\vec{b}$ 的模 $|\vec{a}-2\vec{b}|$.

解　处理向量模的问题, 一个有力的代数工具是利用公式 $|\vec{v}|^2=\vec{v}\cdot\vec{v}$. 将此方法应用于向量 $\vec{a}-2\vec{b}$:

$$\begin{aligned}|\vec{a}-2\vec{b}|^2&=(\vec{a}-2\vec{b})\cdot(\vec{a}-2\vec{b})\\&=\vec{a}\cdot\vec{a}-\vec{a}\cdot(2\vec{b})-(2\vec{b})\cdot\vec{a}+(2\vec{b})\cdot(2\vec{b})\\&=|\vec{a}|^2-2(\vec{a}\cdot\vec{b})-2(\vec{b}\cdot\vec{a})+4(\vec{b}\cdot\vec{b})\\&=|\vec{a}|^2-4(\vec{a}\cdot\vec{b})+4|\vec{b}|^2.\end{aligned}$$

知乎@余命数

代入已知值 $|\vec{a}|=2,\ |\vec{b}|=3,\ \vec{a}\cdot\vec{b}=-3$,

$$\begin{aligned}|\vec{a}-2\vec{b}|^2&=(2)^2-4(-3)+4(3)^2\\&=4+12+4(9)=52,\end{aligned}$$

故所求模为 $|\vec{a}-2\vec{b}|=\sqrt{52}=2\sqrt{13}$.

11.1.6 走路法

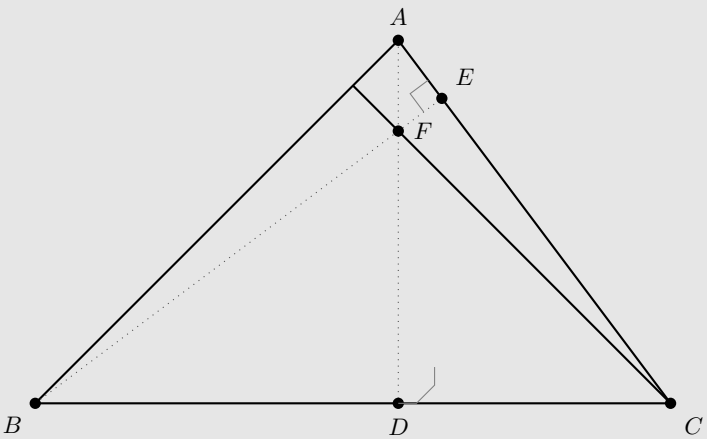
向量加法的三角形法则 $\overrightarrow{AB}+\overrightarrow{BC}=\overrightarrow{AC}$ 是向量代数中最基本的运算之一, 它在几何上直观地表现为一个连续的位移过程. 这种首尾相接、连接始末的构造, 我

知乎@余命数

们称之为**走路法**. 它的基本意义在于, 它为将复杂的几何路径问题转化为一系列有序的代数运算提供了最直接的途径. 任何一个从点 P_1 到点 P_n 的路径, 无论多么曲折, 都可以被分解为向量之和 $\overrightarrow{P_1P_2}+\overrightarrow{P_2P_3}+\cdots+\overrightarrow{P_{n-1}P_n}=\overrightarrow{P_1P_n}$.

这种方法的优势在于其系统性与普适性. 它将对图形空间位置关系的直观依赖, 转化为对向量代数法则的严格运用. 相较于综合几何学中常常需要依赖辅助线构造和灵感的解题模式, 向量走路法提供了一种更为程序化的路径, 建立向量关系, 进行代数演算, 直至获得最终的几何结论. 诚然, 任何有效的方法都包含其独特的技巧, 例如参考点的巧妙选取, 其优势在于, 它用极少的公理和运算规则, 便能构建起解决大量几何问题的统一框架.

例题 11.13 在 $\triangle ABC$ 中, 高线 AD 与 BE 相交于点 F. 求证: $CF\perp AB$.



解 三角形两条高线交于 F, 需证第三条也过 F, 即垂心. 将垂直关系代数化: $\vec{u}\perp\vec{v}$ 当且仅当 $\vec{u}\cdot\vec{v}=0$.

取两条已知高线的交点 F 为原点, 记顶点 A,B,C 的位置向量为 $\vec{a},\vec{b},\vec{c}$. 高线 AD 过原点, 故 $\vec{a}$ 与 AD 共线, $AD\perp BC$ 等价于 $\vec{a}\cdot\overrightarrow{BC}=0$. 由走路法 $\overrightarrow{BC}=\vec{c}-\vec{b}$, 得

$$\vec{a}\cdot(\vec{c}-\vec{b})=0.$$

同理, $BE\perp AC$ 给出

$$\vec{b}\cdot(\vec{c}-\vec{a})=0.$$

欲证 $CF\perp AB$, 即 $\vec{c}\cdot(\vec{b}-\vec{a})=0$.

展开前两式: $\vec{a}\cdot\vec{c}=\vec{a}\cdot\vec{b}$, $\vec{b}\cdot\vec{c}=\vec{b}\cdot\vec{a}$. 由数量积交换律 $\vec{a}\cdot\vec{b}=\vec{b}\cdot\vec{a}$, 得

$$\vec{a}\cdot\vec{c}=\vec{b}\cdot\vec{c}.$$

知乎@余命数

于是

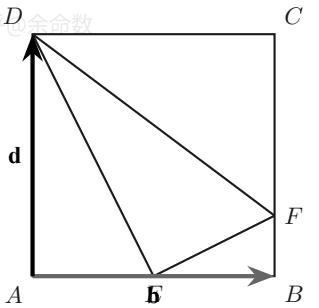
$$\vec{c}\cdot(\vec{b}-\vec{a})=\vec{c}\cdot\vec{b}-\vec{c}\cdot\vec{a}=\vec{b}\cdot\vec{c}-\vec{a}\cdot\vec{c}=0,$$

故 $CF\perp AB$. 证明全程只用数量积代数, 未建坐标, 也未添辅助线.

例题 11.14　如图, 在正方形 $ABCD$ 中, E 是 AB 中点, $3BF=FC$, 求证: $DE\perp EF$.

证明　此题涉及中点与定比分点, 这类条件在向量代数中具有简洁的表述, 并且正方形的垂直与等长边为向量运算提供了简洁的代数性质. 因此, 采用走路法证明.

走路法的要点在于, 选取一个合适的基点为原点, 并以从该点出发的向量为基本元素来表示图中的其他所有向量, 从而将几何问题转化为代数问题. 以点 A 为原点, 令基向量为 $\overrightarrow{AB}=\mathbf{b}$ 和 $\overrightarrow{AD}=\mathbf{d}$.



由于 $ABCD$ 是正方形, 这两个基向量的长度相等, 且相互垂直. 这一几何性质反映在代数上, 即

$$|\mathbf{b}|=|\mathbf{d}|,\quad \text{且}\quad \mathbf{b}\cdot\mathbf{d}=0.$$

接着, 用基向量表示图中相关点的向量. 这好比从原点 A 走到目标点. 点 D 的位置向量即为 $\overrightarrow{AD}=\mathbf{d}$. 点 E 是 AB 的中点, 故其位置向量为 $\overrightarrow{AE}=\frac{1}{2}\overrightarrow{AB}=\frac{1}{2}\mathbf{b}$. 点 F 在边 BC 上, 由 $3BF=FC$ 知, F 是 BC 的一个四等分点, 且 $BF=\frac{1}{4}BC$. 注意到 $\overrightarrow{BC}=\overrightarrow{AC}-\overrightarrow{AB}=(\overrightarrow{AB}+\overrightarrow{AD})-\overrightarrow{AB}=\overrightarrow{AD}=\mathbf{d}$. 于是, $\overrightarrow{BF}=\frac{1}{4}\overrightarrow{BC}=\frac{1}{4}\mathbf{d}$. 从 A 到 F 的路径可以分解为先到 B 再到 F, 故

$$\overrightarrow{AF}=\overrightarrow{AB}+\overrightarrow{BF}=\mathbf{b}+\frac{1}{4}\mathbf{d}.$$

现在构造待证垂直的两个线段所对应的向量 $\overrightarrow{DE}$ 与 $\overrightarrow{EF}$.

$$\overrightarrow{DE}=\overrightarrow{AE}-\overrightarrow{AD}=\frac{1}{2}\mathbf{b}-\mathbf{d},$$
$$\overrightarrow{EF}=\overrightarrow{AF}-\overrightarrow{AE}=\left(\mathbf{b}+\frac{1}{4}\mathbf{d}\right)-\frac{1}{2}\mathbf{b}=\frac{1}{2}\mathbf{b}+\frac{1}{4}\mathbf{d}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为了证明 $DE \perp EF$, 只需验证其方向向量的数量积为零. 计算数量积, 得

$$\begin{aligned}\overrightarrow{DE}\cdot\overrightarrow{EF} &= \left(\frac{1}{2}\mathbf{b}-\mathbf{d}\right)\cdot\left(\frac{1}{2}\mathbf{b}+\frac{1}{4}\mathbf{d}\right)\\ &= \frac{1}{4}(\mathbf{b}\cdot\mathbf{b})+\frac{1}{8}(\mathbf{b}\cdot\mathbf{d})-\frac{1}{2}(\mathbf{d}\cdot\mathbf{b})-\frac{1}{4}(\mathbf{d}\cdot\mathbf{d})\\ &= \frac{1}{4}|\mathbf{b}|^2+\frac{1}{8}(0)-\frac{1}{2}(0)-\frac{1}{4}|\mathbf{d}|^2.\end{aligned}$$

利用 $|\mathbf{b}|^2=|\mathbf{d}|^2$ 这一性质, 上式的结果为

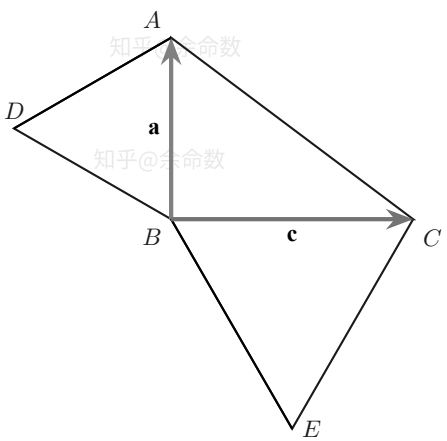
$$\overrightarrow{DE}\cdot\overrightarrow{EF}=\frac{1}{4}|\mathbf{b}|^2-\frac{1}{4}|\mathbf{b}|^2=0.$$

数量积为零, 故向量 $\overrightarrow{DE}$ 与 $\overrightarrow{EF}$ 相互正交, 故线段 DE 与 EF 垂直. □

例题 11.15 如图, 以 $\triangle ABC$ 的 AB 和 BC 边分别向外作等边 $\triangle ABD$ 和 $\triangle BCE$, 且 $AB \perp BC$. 求证: $BE \perp AD$.

证明 此题的条件涉及在三角形边上构建新的几何图形, 且包含旋转与垂直关系, 这正是向量法展现其优势的领域. 向量的代数运算能简洁地处理旋转与投影.

要点在于选择合适的原点与基向量. 点 B 是两个等边三角形的公共顶点, 且是垂直关系 $AB \perp BC$ 的交点, 因此以点 B 为原点进行向量分析最为便捷. 令基向量为 $\overrightarrow{BA}=\mathbf{a}$ 和 $\overrightarrow{BC}=\mathbf{c}$.



根据题设, $AB \perp BC$, 即基向量 $\mathbf{a}$ 与 $\mathbf{c}$ 相互垂直. 其代数体现为它们的数量积为零:

$$\mathbf{a}\cdot\mathbf{c}=0.$$

欲证 $BE \perp AD$, 这等价于证明 $\overrightarrow{BE}\cdot\overrightarrow{AD}=0$. 为此, 需用基向量 $\mathbf{a}$ 和 $\mathbf{c}$ 表示出 $\overrightarrow{BE}$ 和 $\overrightarrow{AD}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

向量 $\overrightarrow{BE}$ 是由向量 $\overrightarrow{BC}=\mathbf{c}$ 绕点 B 向外顺时针旋转 60° 得到的. 向量 $\overrightarrow{BD}$ 是由向量 $\overrightarrow{BA}=\mathbf{a}$ 绕点 B 向外(逆时针)旋转 60° 得到的. 接着, 表示目标向量: 向量 $\overrightarrow{BE}$ 已由上述旋转定义. 向量 $\overrightarrow{AD}$ 可以通过走路得到, 从 A 到 D 等价于从 A 先回到 B, 再从 B 走到 D. 故

$$\overrightarrow{AD}=\overrightarrow{AB}+\overrightarrow{BD}=-\overrightarrow{BA}+\overrightarrow{BD}=-\mathbf{a}+\overrightarrow{BD}.$$

现在, 计算数量积 $\overrightarrow{BE}\cdot\overrightarrow{AD}$:

$$\begin{aligned}\overrightarrow{BE}\cdot\overrightarrow{AD}&=\overrightarrow{BE}\cdot(-\mathbf{a}+\overrightarrow{BD})\\&=\overrightarrow{BE}\cdot\overrightarrow{BD}-\overrightarrow{BE}\cdot\mathbf{a}.\end{aligned}$$

利用数量积的几何定义 $\mathbf{x}\cdot\mathbf{y}=|\mathbf{x}||\mathbf{y}|\cos\theta$ 来计算右侧的两项. 对于 $\overrightarrow{BE}\cdot\overrightarrow{BD}$: $|\overrightarrow{BE}|=|\overrightarrow{BC}|=|\mathbf{c}|$, 且 $|\overrightarrow{BD}|=|\overrightarrow{BA}|=|\mathbf{a}|$. 向量 $\overrightarrow{BD}$ 由 $\mathbf{a}$ 逆时针旋转 60° 得到, 向量 $\overrightarrow{BE}$ 由 $\mathbf{c}$ 顺时针旋转 60° 得到. 由于 $\mathbf{a}$ 与 $\mathbf{c}$ 的夹角为 90°, 那么 $\overrightarrow{BD}$ 与 $\overrightarrow{BE}$ 之间的夹角为 $60^\circ+90^\circ+60^\circ=210^\circ$. 不过, 在计算数量积时, 通常取不超过 180° 的夹角, 即 $360^\circ-210^\circ=150^\circ$. 因此

$$\overrightarrow{BE}\cdot\overrightarrow{BD}=|\overrightarrow{BE}||\overrightarrow{BD}|\cos(150^\circ)=|\mathbf{c}||\mathbf{a}|\cos(150^\circ).$$

对于 $\overrightarrow{BE}\cdot\mathbf{a}$: $|\overrightarrow{BE}|=|\mathbf{c}|$, $|\mathbf{a}|$ 为已知. 向量 $\mathbf{a}$ 与 $\mathbf{c}$ 的夹角为 90°. 向量 $\overrightarrow{BE}$ 由 $\mathbf{c}$ 顺时针旋转 60° 得到, 因此 $\overrightarrow{BE}$ 与 $\mathbf{a}$ 的夹角为 $90^\circ+60^\circ=150^\circ$. 所以

$$\overrightarrow{BE}\cdot\mathbf{a}=|\overrightarrow{BE}||\mathbf{a}|\cos(150^\circ)=|\mathbf{c}||\mathbf{a}|\cos(150^\circ).$$

将以上两式代入数量积的表达式中, 得

$$\overrightarrow{BE}\cdot\overrightarrow{AD}=|\mathbf{c}||\mathbf{a}|\cos(150^\circ)-|\mathbf{c}||\mathbf{a}|\cos(150^\circ)=0.$$

数量积为零, 结论得以证明. 故 $BE\perp AD$. □

前面两个范例展示了平面几何问题的一种代数化处理, 即向量法, 或称走路法. 它把几何图形的性质与关系, 转化为向量空间的代数运算.

向量法先要**建立分析的框架**. 这包括:

- **选取基点**: 在图形中选定一个合适的点为原点. 通常, 这个点是图形的特殊点, 例如顶点, 中点, 或是多条重要线段的交点. 一个好的基点可以大幅简化后续的向量表示.
- **定义基向量**: 从基点出发, 选取一至两个线性无关的向量为基向量. 在平面问题中, 通常是两个不共线的向量. 在正方形或矩形中, 沿两条邻边方向的向量是自然的选择; 在一般三角形中, 沿两条边的向量亦是如此.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个框架一旦建立, 图形中的任何一个点的位置, 都可以由从基点出发的一个唯一的位置向量来描述; 任何一个有向线段, 都可以表示为基向量的线性组合.

框架建立之后, 就要**翻译几何条件**. 将题设中所有的几何关系, 如长度、中点、定比分点、平行、垂直等, 转化为向量的代数语言. 这一步决定整个解法能否走通.

- 长度关系与角度关系, 通常与向量的模和数量积相关. 例如, $AB \perp BC$ 意味着 $\overrightarrow{BA} \cdot \overrightarrow{BC} = 0$.
- 点 M 是线段 AB 的中点, 则 $\overrightarrow{OM} = \frac{1}{2}(\overrightarrow{OA} + \overrightarrow{OB})$.
- 点 P 在直线 AB 上, 且 $AP : PB = \lambda : \mu$, 则 $\overrightarrow{OP} = \frac{\mu\overrightarrow{OA} + \lambda\overrightarrow{OB}}{\lambda + \mu}$.

通过这种方式, 整个几何问题被翻译成了一个关于基向量的代数问题.

接着, **构造并计算目标向量**. 将待证的几何关系, 例如 $DE \perp EF$ 同样用向量语言表达出来, 即 $\overrightarrow{DE} \cdot \overrightarrow{EF} = 0$. 然后, 利用向量的加减法法则, 即走路法则, 如 $\overrightarrow{DE} = \overrightarrow{OE} - \overrightarrow{OD}$ 将这些目标向量用基向量表示出来.

最后, 只需通过纯粹的**代数运算**, 结合基向量自身的性质, 例如它们的模长关系和数量积, 对目标表达式化简并验证是否符合要证明的即可.

11.1.7 极化恒等式

在向量的代数体系中, 数量积是极为重要的运算, 它将向量的几何关系, 如夹角与代数运算联系起来. 数量积的几何定义 $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}|| \cdot ||\mathbf{v}|| \cos\theta$ 依赖于夹角 θ. 但在许多几何问题中, 角度是未知的, 我们更易获知的是线段的长度, 即向量的模.

自然要问, 数量积能否仅由模长表达. 极化恒等式给出肯定回答, 它在数量积与向量模长之间建立起直接的联系.

考虑向量和的模长平方 $||\mathbf{u} + \mathbf{v}||^2$. 根据模的定义, 它等于向量自身的数量积:

$$||\mathbf{u} + \mathbf{v}||^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v})$$

利用数量积的分配律展开上式, 得

$$(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \mathbf{v} \cdot \mathbf{v}$$

注意到 $\mathbf{u} \cdot \mathbf{u} = ||\mathbf{u}||^2$ 且数量积满足交换律 $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, 上式可化为

$$||\mathbf{u} + \mathbf{v}||^2 = ||\mathbf{u}||^2 + ||\mathbf{v}||^2 + 2(\mathbf{u} \cdot \mathbf{v})$$

移项整理, 得数量积的一种表达形式. 这个思想可以推广并总结为如下的命题.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 11.4 (极化恒等式) 对于任意两个向量 $\mathbf{u},\mathbf{v}$, 其数量积可以通过向量的模长表示:

$$\mathbf{u}\cdot\mathbf{v}=\frac{1}{2}\left(||\mathbf{u}+\mathbf{v}||^2-||\mathbf{u}||^2-||\mathbf{v}||^2\right)$$
$$\mathbf{u}\cdot\mathbf{v}=\frac{1}{4}\left(||\mathbf{u}+\mathbf{v}||^2-||\mathbf{u}-\mathbf{v}||^2\right)$$

证明 第一种形式由展开 $||\mathbf{u}+\mathbf{v}||^2$ 得出.

$$||\mathbf{u}+\mathbf{v}||^2=(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v})=||\mathbf{u}||^2+2\mathbf{u}\cdot\mathbf{v}+||\mathbf{v}||^2.$$

移项并除以 2, 得

$$\mathbf{u}\cdot\mathbf{v}=\frac{1}{2}\left(||\mathbf{u}+\mathbf{v}||^2-||\mathbf{u}||^2-||\mathbf{v}||^2\right).$$

第二种形式更具对称性, 其推导过程也颇具启发性. 已有

$$||\mathbf{u}+\mathbf{v}||^2=||\mathbf{u}||^2+||\mathbf{v}||^2+2(\mathbf{u}\cdot\mathbf{v})\quad(1)$$

同理, 计算 $||\mathbf{u}-\mathbf{v}||^2$:

$$\begin{aligned}||\mathbf{u}-\mathbf{v}||^2&=(\mathbf{u}-\mathbf{v})\cdot(\mathbf{u}-\mathbf{v})\\&=||\mathbf{u}||^2-2(\mathbf{u}\cdot\mathbf{v})+||\mathbf{v}||^2\quad(2)\end{aligned}$$

将 (1) 式与 (2) 式相减, 即可消去模长平方项, 得到

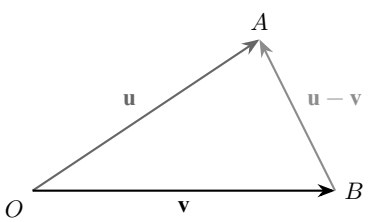
$$||\mathbf{u}+\mathbf{v}||^2-||\mathbf{u}-\mathbf{v}||^2=4(\mathbf{u}\cdot\mathbf{v})$$

整理后即得恒等式的第二种形式. □

此恒等式具有基本的几何意义. 我们重写式 (2) 可得

$$||\mathbf{u}-\mathbf{v}||^2=||\mathbf{u}||^2+||\mathbf{v}||^2-2(\mathbf{u}\cdot\mathbf{v})$$

若以 $\mathbf{u}=\overrightarrow{OA}$ 和 $\mathbf{v}=\overrightarrow{OB}$ 为邻边构成三角形, 则第三边即为向量 $\overrightarrow{BA}=\overrightarrow{OA}-\overrightarrow{OB}=\mathbf{u}-\mathbf{v}$.



知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 θ 为 $\mathbf{u}$ 和 $\mathbf{v}$ 的夹角, 代入数量积的几何定义 $\mathbf{u}\cdot\mathbf{v}=||\mathbf{u}||||\mathbf{v}||\cos\theta$, 上式即变为

$$||\mathbf{u}-\mathbf{v}||^2=||\mathbf{u}||^2+||\mathbf{v}||^2-2||\mathbf{u}||||\mathbf{v}||\cos\theta$$

这正是三角形中的余弦定理. 因此, **极化恒等式可以被视为余弦定理的向量形式**, 它将几何中的长度与角度关系融入了向量代数的框架之中. 当一个问题只涉及边长而不涉及角度时, 此恒等式是计算数量积, 进而判断垂直关系的不二之选.

例题 11.16 设有两个向量 $\mathbf{a},\mathbf{b}$, 满足 $||\mathbf{a}+\mathbf{b}||=\sqrt{10}$ 及 $||\mathbf{a}-\mathbf{b}||=\sqrt{6}$, 求数量积 $\mathbf{a}\cdot\mathbf{b}$ 的值.

解 此问题直接给出了向量和与差的模长, 而要求它们的数量积, 这正是极化恒等式应用的典型情景. 采用极化恒等式的第二种形式:

$$\mathbf{a}\cdot\mathbf{b}=\frac{1}{4}\left(||\mathbf{a}+\mathbf{b}||^2-||\mathbf{a}-\mathbf{b}||^2\right)$$

将题设中的已知值代入上式, 得

$$\mathbf{a}\cdot\mathbf{b}=\frac{1}{4}\left((\sqrt{10})^2-(\sqrt{6})^2\right)$$

知乎@余命数

进行简单的代数计算, 得

$$\mathbf{a}\cdot\mathbf{b}=\frac{1}{4}(10-6)=\frac{1}{4}(4)=1.$$

因此, 这两个向量的数量积为 1.

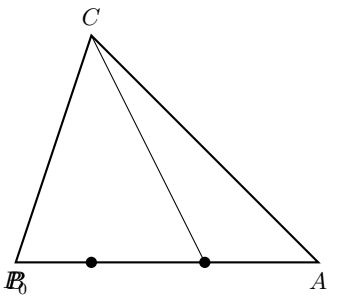
例题 11.17 设 $\triangle ABC$, P_0 是边 AB 上一定点, 满足 $P_0B=\frac{1}{4}AB$. 若对于边 AB 上任意一点 P, 恒有不等式 $\overrightarrow{PB}\cdot\overrightarrow{PC}\geq\overrightarrow{P_0B}\cdot\overrightarrow{P_0C}$ 成立, 试确定 $\triangle ABC$ 必须满足的几何条件.

解 该命题的条件是一个关于动点 P 的向量不等式. 要点是表达式 $\overrightarrow{PB}\cdot\overrightarrow{PC}$ 在点 $P=P_0$ 处取得最小值. 这种涉及最值的问题, 提示可将该表达式构造成一个函数, 并利用函数的性质来求解.

知乎@余命数

以点 B 为基点, 设 $\overrightarrow{BA}=\mathbf{a}$ 和 $\overrightarrow{BC}=\mathbf{c}$.

知乎@余命数



知乎@余命数

点 P 在线段 AB 上，它的位置可以用参数 λ 描述. 令 $\overrightarrow{BP}=\lambda\overrightarrow{BA}=\lambda\mathbf{a}$，其中 $\lambda\in[0,1]$. 于是, 向量 $\overrightarrow{PB}$ 和 $\overrightarrow{PC}$ 可以用基向量表示:

$$\overrightarrow{PB}=-\overrightarrow{BP}=-\lambda\mathbf{a},$$
$$\overrightarrow{PC}=\overrightarrow{BC}-\overrightarrow{BP}=\mathbf{c}-\lambda\mathbf{a}.$$

接着, 计算它们的数量积, 得

$$\begin{aligned}f(\lambda)=\overrightarrow{PB}\cdot\overrightarrow{PC}&=(-\lambda\mathbf{a})\cdot(\mathbf{c}-\lambda\mathbf{a})\\&=-\lambda(\mathbf{a}\cdot\mathbf{c})+\lambda^2(\mathbf{a}\cdot\mathbf{a})\\&=||\mathbf{a}||^2\lambda^2-(\mathbf{a}\cdot\mathbf{c})\lambda.\end{aligned}$$

知乎@余命数

这是一个关于变量 λ 的二次函数, 其二次项系数 $||\mathbf{a}||^2$ 为正, 故其图像是开口向上的抛物线，存在唯一的最小值. 根据二次函数的性质，其对称轴位置在 $\lambda=-\frac{-(\mathbf{a}\cdot\mathbf{c})}{2||\mathbf{a}||^2}=\frac{\mathbf{a}\cdot\mathbf{c}}{2||\mathbf{a}||^2}$. 此即函数 $f(\lambda)$ 取最小值的点.

根据题设,最小值在 $P=P_0$ 处取得. 需确定 P_0 对应的参数 λ_0. 由 $P_0B=\frac{1}{4}AB$ 可知, $||\overrightarrow{BP_0}||=\frac{1}{4}||\overrightarrow{BA}||$. 由于 P_0 在 A,B 之间, 故 $\overrightarrow{BP_0}$ 与 $\overrightarrow{BA}$ 同向, 所以 $\overrightarrow{BP_0}=\frac{1}{4}\overrightarrow{BA}=\frac{1}{4}\mathbf{a}$. 因此, $\lambda_0=\frac{1}{4}$.

令函数取得最小值的 λ 等于 λ_0:

$$\frac{\mathbf{a}\cdot\mathbf{c}}{2||\mathbf{a}||^2}=\frac{1}{4}.$$

由此得一个关于基向量的代数关系:

$$2(\mathbf{a}\cdot\mathbf{c})=||\mathbf{a}||^2.$$

最后一步, 需将这个代数关系翻译回几何语言. 考察三角形的边长关系. $||\overrightarrow{AC}||^2=||\overrightarrow{BC}-\overrightarrow{BA}||^2=||\mathbf{c}-\mathbf{a}||^2$. 利用推导极化恒等式时的展开方式, 有

$$||\mathbf{c}-\mathbf{a}||^2=(\mathbf{c}-\mathbf{a})\cdot(\mathbf{c}-\mathbf{a})=||\mathbf{c}||^2-2(\mathbf{a}\cdot\mathbf{c})+||\mathbf{a}||^2.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将已得的条件 $2(\mathbf{a}\cdot\mathbf{c})=||\mathbf{a}||^2$ 代入上式, 得

$$||\overrightarrow{AC}||^2=||\mathbf{c}||^2-||\mathbf{a}||^2+||\mathbf{a}||^2=||\mathbf{c}||^2.$$

故 $||\overrightarrow{AC}||=||\mathbf{c}||=||\overrightarrow{BC}||$. 因此, $\triangle ABC$ 必须满足的几何条件是 $AC=BC$, 即它是一个以 C 为顶角的等腰三角形. 反过来, 若 $AC=BC$, 则上述等价变形给出 $2(\mathbf{a}\cdot\mathbf{c})=||\mathbf{a}||^2$, 从而二次函数的对称轴正是 $\lambda=\frac{1}{4}$, 不等式也确实对边 AB 上任意点 P 成立.

例题 11.18 在 $\triangle ABC$ 中, D 是 BC 的中点, E,F 是 AD 上的两个三等分点, 且沿 A 到 D 的方向依次为 E,F. 已知 $\overrightarrow{BA}\cdot\overrightarrow{CA}=4$, $\overrightarrow{BF}\cdot\overrightarrow{CF}=-1$, 求 $\overrightarrow{BE}\cdot\overrightarrow{CE}$ 的值.

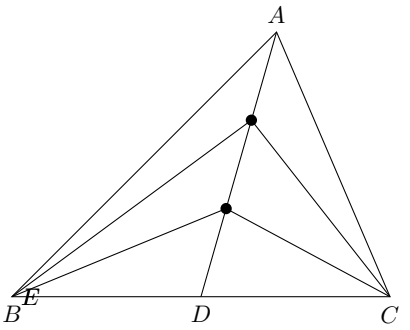
解 此题的条件与结论均涉及以不同顶点为起点的向量数量积. 这种结构提示寻找一个统一的代数形式. 注意到所有数量积都涉及线段 BC 的两个端点 B 和 C, 这启发将 BC 的中点 D 作为参照点.

考察形如 $\overrightarrow{PX}\cdot\overrightarrow{PY}$ 的数量积, 当一个参照点 M 存在时, 可进行分解. 对于 $\overrightarrow{BA}\cdot\overrightarrow{CA}$, 将向量起点统一到中点 D. $\overrightarrow{BA}=\overrightarrow{DA}-\overrightarrow{DB}$ $\overrightarrow{CA}=\overrightarrow{DA}-\overrightarrow{DC}$ 由于 D 是 BC 的中点, $\overrightarrow{DC}=-\overrightarrow{DB}$. 代入上式, 得 $\overrightarrow{CA}=\overrightarrow{DA}+\overrightarrow{DB}$ 于是, 数量积可以展开为:

$$\overrightarrow{BA}\cdot\overrightarrow{CA}=(\overrightarrow{DA}-\overrightarrow{DB})\cdot(\overrightarrow{DA}+\overrightarrow{DB})=||\overrightarrow{DA}||^2-||\overrightarrow{DB}||^2=AD^2-BD^2.$$

这实际上是极化恒等式思想的一个直接应用, 它将数量积与中线长和半弦长联系起来. 根据题设, 有

$$AD^2-BD^2=4\quad(1)$$



完全同理, 对于点 F 和 E, 可得到相同的代数结构:

$$\overrightarrow{BF}\cdot\overrightarrow{CF}=DF^2-BD^2$$

$$\overrightarrow{BE}\cdot\overrightarrow{CE}=DE^2-BD^2$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据题设 $\overrightarrow{BF}\cdot\overrightarrow{CF}=-1$, 有

$$DF^2-BD^2=-1 \quad (2)$$

现在问题转化为一个关于长度的代数问题. 由题设约定, F 是靠近 D 的分点, 则 $DF=\frac{1}{3}AD$, 且 $DE=\frac{2}{3}AD$. 将 $DF=\frac{1}{3}AD$ 代入 (2) 式, 得

$$\left(\frac{1}{3}AD\right)^2-BD^2=-1 \implies \frac{1}{9}AD^2-BD^2=-1 \quad (3)$$

现有关于 AD^2 和 BD^2 的二元一次方程组, 由 (1) 和 (3) 构成:

$$AD^2-BD^2=4$$
$$\frac{1}{9}AD^2-BD^2=-1$$

两式相减, 消去 BD^2:

$$\left(1-\frac{1}{9}\right)AD^2=4-(-1) \implies \frac{8}{9}AD^2=5 \implies AD^2=\frac{45}{8}.$$

将 AD^2 的值代回 (1) 式:

$$BD^2=AD^2-4=\frac{45}{8}-4=\frac{45-32}{8}=\frac{13}{8}.$$

现计算 $\overrightarrow{BE}\cdot\overrightarrow{CE}=DE^2-BD^2$. 利用 $DE=\frac{2}{3}AD$, 有 $DE^2=\frac{4}{9}AD^2$. 代入已求出的值:

$$\overrightarrow{BE}\cdot\overrightarrow{CE}=\frac{4}{9}AD^2-BD^2=\frac{4}{9}\left(\frac{45}{8}\right)-\frac{13}{8}=\frac{5}{2}-\frac{13}{8}=\frac{20-13}{8}=\frac{7}{8}.$$

故所求值为 $\frac{7}{8}$.

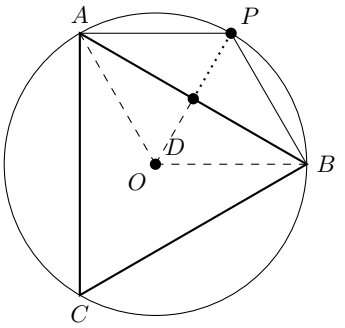
例题 11.19 已知等边 $\triangle ABC$ 内接于半径为 2 的 $\odot O$, 点 P 是 $\odot O$ 上的一个动点, 则 $\overrightarrow{PA}\cdot\overrightarrow{PB}$ 的取值范围是?

解 这个问题的要点是处理一个动点相关的数量积. 动点 P 的轨迹是圆, 其几何特征是到圆心 O 的距离恒定. 下一步是将数量积 $\overrightarrow{PA}\cdot\overrightarrow{PB}$ 与点 P 的位置建立联系.

同样地, 极化恒等式的思想启示引入 AB 的中点. 设 D 为 AB 的中点. 将 $\overrightarrow{PA}$ 和 $\overrightarrow{PB}$ 用过 D 的路径来表示: $\overrightarrow{PA}=\overrightarrow{PD}+\overrightarrow{DA}$ $\overrightarrow{PB}=\overrightarrow{PD}+\overrightarrow{DB}$ 由于 D 是 AB 的中点, $\overrightarrow{DB}=-\overrightarrow{DA}$. 代入得 $\overrightarrow{PB}=\overrightarrow{PD}-\overrightarrow{DA}$ 计算数量积, 得

$$\overrightarrow{PA}\cdot\overrightarrow{PB}=(\overrightarrow{PD}+\overrightarrow{DA})\cdot(\overrightarrow{PD}-\overrightarrow{DA})=||\overrightarrow{PD}||^2-||\overrightarrow{DA}||^2=PD^2-AD^2.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



通过这个变换, 数量积的范围问题转化为了几何长度 PD 的范围问题. 问题分解为两个子任务, 计算定长 AD, 以及确定动长 PD 的范围.

首先, 计算 AD.$\triangle ABC$ 是内接于 $\odot O$ 的等边三角形, 半径 $r = OP = OA = OB = 2$. 在 $\triangle OAB$ 中, $OA = OB = 2$, $\angle AOB = \frac{360^\circ}{3} = 120^\circ$.$D$ 是 AB 的中点, 因此 $OD \perp AB$. 在 Rt$\triangle ODA$ 中, $\angle AOD = \frac{1}{2}\angle AOB = 60^\circ$. 于是 $AD = OA\sin(60^\circ) = 2 \cdot \frac{\sqrt{3}}{2} = \sqrt{3}$. 因此, 常数项 $AD^2 = 3$.

接着, 确定 PD 的范围.P 是圆周上的动点, D 是圆内一定点.P 到 D 的距离范围取决于 D 到圆心 O 的距离. 在 Rt$\triangle ODA$ 中, $OD = OA\cos(60^\circ) = 2 \cdot \frac{1}{2} = 1$. 点 P 在半径为 2 的圆上, 点 D 与圆心 O 的距离为 1. 根据几何知识, PD 的最大值发生在 P, O, D 三点共线且 O 在 P, D 之间时:

$$PD_{\max} = PO + OD = r + OD = 2 + 1 = 3.$$

PD 的最小值发生在 P, O, D 三点共线且 D 在 P, O 之间时:

$$PD_{\min} = PO - OD = r - OD = 2 - 1 = 1.$$

所以, PD 的取值范围是 $[1, 3]$. 从而, PD^2 的取值范围是 $[1^2, 3^2] = [1, 9]$.

最后, 回到原表达式:

$$\overrightarrow{PA} \cdot \overrightarrow{PB} = PD^2 - AD^2 = PD^2 - 3.$$

由于 $PD^2 \in [1, 9]$, 那么

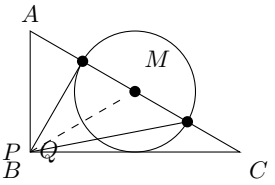
$$\overrightarrow{PA} \cdot \overrightarrow{PB} \in [1 - 3, 9 - 3] = [-2, 6].$$

因此, 所求的取值范围是 $[-2, 6]$.

例题 11.20 如图, Rt$\triangle ABC$ 中, $\angle ABC = 90^\circ$、$AB = 2$、$BC = 2\sqrt{3}$.M 是线段 AC 上一动点, 若以 M 为圆心, 半径为 1 的圆与线段 AC 交于 P, Q 两点, 则 $\overrightarrow{BP} \cdot \overrightarrow{BQ}$ 的最小值为(　)

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 此题要求一个数量积的最小值, 其中涉及的向量起点固定, 终点则由一个移动的圆心确定. 这看似复杂, 但其结构恰好契合了极化恒等式的思想.



注意到点 M 是圆心, 而 P,Q 是圆与直径 AC 的交点, 故 M 必然是线段 PQ 的中点. 根据极化恒等式的思想, 可将 $\overrightarrow{BP}\cdot\overrightarrow{BQ}$ 以中点 M 分解:

$$\overrightarrow{BP}=\overrightarrow{BM}+\overrightarrow{MP}$$

$$\overrightarrow{BQ}=\overrightarrow{BM}+\overrightarrow{MQ}$$

由于 M 是 PQ 中点, $\overrightarrow{MQ}=-\overrightarrow{MP}$. 于是

$$\overrightarrow{BP}\cdot\overrightarrow{BQ}=(\overrightarrow{BM}+\overrightarrow{MP})\cdot(\overrightarrow{BM}-\overrightarrow{MP})=||\overrightarrow{BM}||^2-||\overrightarrow{MP}||^2.$$

$||\overrightarrow{MP}||$ 是圆的半径, 题设为 1, 故 $||\overrightarrow{MP}||^2=1$. 表达式化为:

$$\overrightarrow{BP}\cdot\overrightarrow{BQ}=||\overrightarrow{BM}||^2-1.$$

欲求 $\overrightarrow{BP}\cdot\overrightarrow{BQ}$ 的最小值, 只需寻求 $||\overrightarrow{BM}||^2$ 的最小值. 这等价于寻找动点 M 在线段 AC 上运动时, 点 B 到 M 的最短距离. 几何直观表明, 点到线段的最短距离是该点到直线距离(垂线段长). 令 h_B 为 Rt$\triangle ABC$ 斜边 AC 上的高. 斜边 AC 的长度由勾股定理给出: $AC=\sqrt{AB^2+BC^2}=\sqrt{2^2+(2\sqrt{3})^2}=\sqrt{4+12}=4$. 根据三角形面积公式, $\frac{1}{2}AB\cdot BC=\frac{1}{2}AC\cdot h_B$.

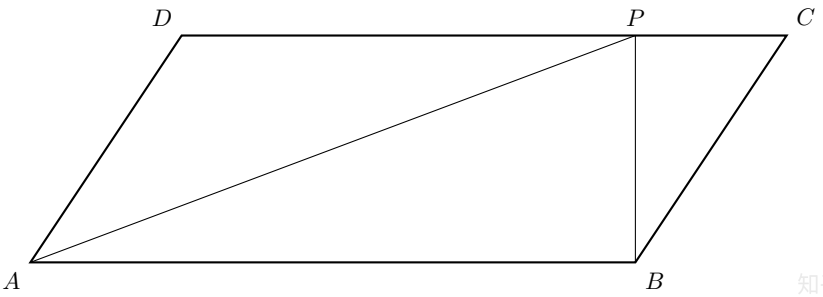
$$\frac{1}{2}\cdot 2\cdot 2\sqrt{3}=\frac{1}{2}\cdot 4\cdot h_B\implies h_B=\sqrt{3}.$$

此垂足 H 位于线段 AC 上, 且 $AH=\frac{AB^2}{AC}=1$, 因而当 $M=H$ 时, 以 M 为圆心、半径为 1 的圆与线段 AC 仍有两个交点, 符合题设限制. 此即 $||\overrightarrow{BM}||$ 的最小值. 于是 $||\overrightarrow{BM}||^2$ 的最小值为 $(\sqrt{3})^2=3$. 因此, $\overrightarrow{BP}\cdot\overrightarrow{BQ}$ 的最小值为 $3-1=2$.

例题 11.21 在平行四边形 $ABCD$ 中, 已知 $AB=8, AD=5$. 点 P 在边 CD 上, 满足 $\overrightarrow{CP}=3\overrightarrow{PD}$, 且 $\overrightarrow{AP}\cdot\overrightarrow{BP}=2$. 求 $\overrightarrow{AB}\cdot\overrightarrow{AD}$ 的值.

解 此题在一个平行四边形中给出了边长与一个数量积的特定值, 要求另一个数量积. 这是典型的借助向量代数运算建立联系的问题. 以点 A 为基点, 令基向量为 $\overrightarrow{AB}$ 和 $\overrightarrow{AD}$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数



现利用已知条件 $\overrightarrow{AP}\cdot\overrightarrow{BP}=2$ 来构造一个包含目标 $\overrightarrow{AB}\cdot\overrightarrow{AD}$ 的方程. 首先, 用基向量表示 $\overrightarrow{AP}$ 和 $\overrightarrow{BP}$. 点 P 在 CD 上, 可沿路径 $A\to D\to P$ 表示 $\overrightarrow{AP}$.

$$\overrightarrow{AP}=\overrightarrow{AD}+\overrightarrow{DP}.$$

由 $\overrightarrow{CP}=3\overrightarrow{PD}$ 可知 P 是 CD 的四等分点, 且 $\overrightarrow{DP}=\frac{1}{4}\overrightarrow{DC}$. 在平行四边形中, $\overrightarrow{DC}=\overrightarrow{AB}$. 故 $\overrightarrow{DP}=\frac{1}{4}\overrightarrow{AB}$. 所以, $\overrightarrow{AP}=\overrightarrow{AD}+\frac{1}{4}\overrightarrow{AB}$.

接着表示 $\overrightarrow{BP}$. 路径可以是 $B\to A\to P$.

$$\overrightarrow{BP}=\overrightarrow{BA}+\overrightarrow{AP}=-\overrightarrow{AB}+\left(\overrightarrow{AD}+\frac{1}{4}\overrightarrow{AB}\right)=\overrightarrow{AD}-\frac{3}{4}\overrightarrow{AB}.$$

现在将这两个表达式代入已知条件:

$$\left(\overrightarrow{AD}+\frac{1}{4}\overrightarrow{AB}\right)\cdot\left(\overrightarrow{AD}-\frac{3}{4}\overrightarrow{AB}\right)=2.$$

展开此数量积, 这与极化恒等式的推导过程一致:

$$||\overrightarrow{AD}||^2-\frac{3}{4}(\overrightarrow{AD}\cdot\overrightarrow{AB})+\frac{1}{4}(\overrightarrow{AB}\cdot\overrightarrow{AD})-\frac{3}{16}||\overrightarrow{AB}||^2=2.$$

利用数量积的交换律 $\overrightarrow{AB}\cdot\overrightarrow{AD}=\overrightarrow{AD}\cdot\overrightarrow{AB}$, 合并同类项, 得

$$||\overrightarrow{AD}||^2-\frac{1}{2}(\overrightarrow{AB}\cdot\overrightarrow{AD})-\frac{3}{16}||\overrightarrow{AB}||^2=2.$$

代入已知边长 $AD=5, AB=8$:

$$5^2-\frac{1}{2}(\overrightarrow{AB}\cdot\overrightarrow{AD})-\frac{3}{16}(8^2)=2.$$

$$25-\frac{1}{2}(\overrightarrow{AB}\cdot\overrightarrow{AD})-\frac{3}{16}(64)=2.$$

$$25-\frac{1}{2}(\overrightarrow{AB}\cdot\overrightarrow{AD})-12=2.$$

$$13-\frac{1}{2}(\overrightarrow{AB}\cdot\overrightarrow{AD})=2.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解得 $\frac{1}{2}(\overrightarrow{AB}\cdot\overrightarrow{AD})=11$, 故 $\overrightarrow{AB}\cdot\overrightarrow{AD}=22$.

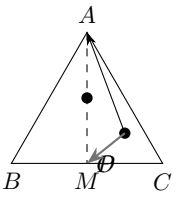
例题 11.22　已知 $\triangle ABC$ 是边长为 2 的等边三角形, P 为平面 ABC 内一点, 则 $\overrightarrow{PA}\cdot(\overrightarrow{PB}+\overrightarrow{PC})$ 的最小值是(　)

解　此表达式包含三个与动点 P 相关的向量, 直接处理较为困难. 可利用向量加法的几何意义将其简化. 令 M 为边 BC 的中点. 根据向量加法的平行四边形法则, 向量和 $\overrightarrow{PB}+\overrightarrow{PC}$ 可以被合成为

$$\overrightarrow{PB}+\overrightarrow{PC}=2\overrightarrow{PM}.$$

于是, 原表达式化为

$$\overrightarrow{PA}\cdot(2\overrightarrow{PM})=2(\overrightarrow{PA}\cdot\overrightarrow{PM}).$$



现在问题转化为求 $2(\overrightarrow{PA}\cdot\overrightarrow{PM})$ 的最小值. 表达式中仍有两个与 P 相关的向量. 为了进一步化简, 再次利用中点. 线段 AM 是定长线段, 取其中心 O 作为新的参照点.

$$\overrightarrow{PA}=\overrightarrow{PO}+\overrightarrow{OA}$$

$$\overrightarrow{PM}=\overrightarrow{PO}+\overrightarrow{OM}$$

由于 O 是 AM 的中点, $\overrightarrow{OM}=-\overrightarrow{OA}$. 代入后, 数量积变为

$$\overrightarrow{PA}\cdot\overrightarrow{PM}=(\overrightarrow{PO}+\overrightarrow{OA})\cdot(\overrightarrow{PO}-\overrightarrow{OA})=||\overrightarrow{PO}||^2-||\overrightarrow{OA}||^2.$$

所以, 原表达式为

$$2(||\overrightarrow{PO}||^2-||\overrightarrow{OA}||^2).$$

这是一个关于动点 P 到定点 O 距离的函数. O 与 A 均为定点, 故 $||\overrightarrow{OA}||^2$ 是一个常数. 为使整个表达式最小, 需最小化 $||\overrightarrow{PO}||^2$, 其最小值在 $P=O$ 时取得, 为 0. 此时, 整个表达式的最小值为 $2(0-||\overrightarrow{OA}||^2)=-2||\overrightarrow{OA}||^2$.

最后, 计算 $||\overrightarrow{OA}||$. $\triangle ABC$ 是边长为 2 的等边三角形, 其高(也是中线) AM 的长度为 $2\cdot\sin(60^\circ)=2\cdot\frac{\sqrt{3}}{2}=\sqrt{3}$. O 是 AM 的中点, 所以 $||\overrightarrow{OA}||=\frac{1}{2}||\overrightarrow{AM}||=\frac{\sqrt{3}}{2}$. 因此, $||\overrightarrow{OA}||^2=\left(\frac{\sqrt{3}}{2}\right)^2=\frac{3}{4}$. 所求的最小值为 $-2\cdot\frac{3}{4}=-\frac{3}{2}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 11.23 在 $\triangle ABC$ 中, AD 为 BC 边上的中线, E 为 AD 延长线上一点, 且 $AD = 2DE$. 已知 $\overrightarrow{AB} \cdot \overrightarrow{AC} = 10$, $\overrightarrow{EB} \cdot \overrightarrow{EC} = -2$, 求 $|BC|$ 的长度.

解 此题与前面关于中点分解的例题结构高度相似, 提供了关于中线端点以及延长线上一点的向量数量积信息, 这是应用中线分解的明确信号. 设 D 为 BC 的中点. 对于平面上任意一点 X, 有恒等式

$$\overrightarrow{XB} \cdot \overrightarrow{XC} = XD^2 - BD^2.$$

将此恒等式分别应用于点 A 和点 E. 当 $X = A$ 时:

$$\overrightarrow{AB} \cdot \overrightarrow{AC} = AD^2 - BD^2.$$

由题设 $\overrightarrow{AB} \cdot \overrightarrow{AC} = 10$, 得

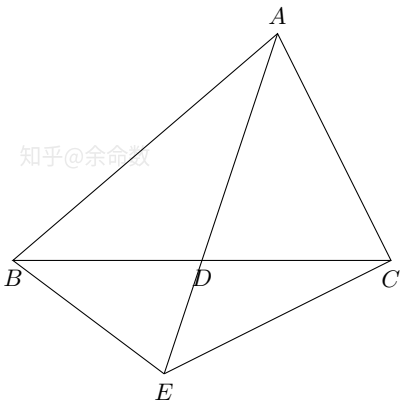
$$AD^2 - BD^2 = 10 \quad (1).$$

当 $X = E$ 时:

$$\overrightarrow{EB} \cdot \overrightarrow{EC} = ED^2 - BD^2.$$

由题设 $\overrightarrow{EB} \cdot \overrightarrow{EC} = -2$, 得

$$ED^2 - BD^2 = -2 \quad (2).$$



同时, 几何条件 $AD = 2DE$ 意味着长度的平方关系为 $AD^2 = 4DE^2$. 将此关系代入方程 (1):

$$4DE^2 - BD^2 = 10 \quad (3).$$

现在得一个关于 DE^2 和 BD^2 的二元一次方程组:

$$4DE^2 - BD^2 = 10$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$DE^2 - BD^2 = -2$$

将两方程相减, 消去 BD^2:

$$3DE^2 = 12 \implies DE^2 = 4.$$

将 $DE^2 = 4$ 代回方程 (2):

$$4 - BD^2 = -2 \implies BD^2 = 6.$$

所求是 $|BC|$. 由于 D 是 BC 的中点, $|BC| = 2|BD|$. 因此, $|BC| = 2\sqrt{BD^2} = 2\sqrt{6}$.

极化恒等式的适用场合如下.

知乎@余命数

- **中点结构**

  这是最重要且最常见的结构. 当题目中出现 **线段的中点** M, 并且需要处理与该线段两端点相关的向量数量积时, 几乎总能受益于此思想. 具体表现为计算形如 $\overrightarrow{PA}\cdot\overrightarrow{PB}$ 的数量积, 其中 M 是 AB 的中点. 通过分解 $\overrightarrow{PA} = \overrightarrow{PM} + \overrightarrow{MA}$ 和 $\overrightarrow{PB} = \overrightarrow{PM} + \overrightarrow{MB} = \overrightarrow{PM} - \overrightarrow{MA}$, 数量积立即转化为长度的平方差:

  $$\overrightarrow{PA}\cdot\overrightarrow{PB} = ||\overrightarrow{PM}||^2 - ||\overrightarrow{MA}||^2.$$

  这种转换是极化恒等式思想的直接体现, 它将一个依赖于角度和动点的复杂数量积, 转化为一个仅与距离相关的简单代数式.

- **最值问题**

  当问题要求某个数量积的最小值或最大值, 且其中涉及动点时, 极化恒等式是实现问题转化的关键工具. 比如之前的例题, 通过中点分解, 将 $\overrightarrow{BP}\cdot\overrightarrow{BQ}$ 的最值问题转化为求 $||\overrightarrow{BM}||^2$ 的最值, 将 $\overrightarrow{PA}\cdot(\overrightarrow{PB}+\overrightarrow{PC})$ 的最值问题转化为求 $||\overrightarrow{PO}||^2$ 的最值. 这实质上是将一个向量代数的最值问题, 降维成一个更直观的几何距离的最值问题, 即“求一点到某轨迹(直线、线段、圆)的最短或最长距离”.

- **向量和的出现**

  当表达式中出现形如 $\overrightarrow{PB} + \overrightarrow{PC}$ 的向量和时, 这是一个强烈的简化信号. 利用中点 M 可将此和式合并为 $2\overrightarrow{PM}$. 这一步往往是应用极化恒等式的前奏, 它将三个与动点 P 相关的向量 $(\overrightarrow{PA}, \overrightarrow{PB}, \overrightarrow{PC})$ 减少到两个 $(\overrightarrow{PA}, \overrightarrow{PM})$, 从而创造出可以应用中点分解的结构.

- **纯长度信息的代数运算**

  当题目给出的条件主要为边长, 却要求计算依赖于角度的数量积时, 这暗示需绕开角度, 直接建立长度与数量积的关系. 之前的几道例题就是一个典型的例子, 通过向量分解与展开, 建立了一个包含已知长度和未知数量积的代

数方程, 最终解出目标值. 这整个代数展开的过程, 与推导极化恒等式时展开和与差再相减的步骤一致.

极化恒等式既是一个公式, 也是一种**对称化**和**代数化**的策略. 其应用的关键, 是识别出几何图形中由中点带来的对称性, 并利用这种对称性将复杂的向量关系转化为简单的标量(长度)关系. 当你面对一个向量问题, 尤其是涉及中点、最值和向量和时, 应当首先思考能否通过分解与变换, 构造出 $(X+Y)\cdot(X-Y)$ 这一标志性结构.

知乎@余命数

11.2 定比分点与共线定理

向量既表示大小与方向, 其代数运算也直接描述几何关系. 点的共线、共面等位置关系, 在向量语言下有着极为简洁的代数刻画.

11.2.1 系数和为常数的轨迹

先考察平面上任意三点 P, A, B 共线的充要条件. 从几何直观上看, P 在直线 AB 上, 意味着向量 $\overrightarrow{AP}$ 与向量 $\overrightarrow{AB}$ 是共线的. 用代数语言描述, 即存在唯一的实数 λ, 使得

知乎@余命数 知乎@余命数 知乎@余命数

$$\overrightarrow{AP}=\lambda\overrightarrow{AB}.$$

这是一个重要的基础. 然而在很多问题中, 我们更习惯于从一个公共的基点(原点) O 出发来描述所有点的位置. 可将上式用位置向量 $\overrightarrow{OP}, \overrightarrow{OA}, \overrightarrow{OB}$ 来改写:

$$\overrightarrow{OP}-\overrightarrow{OA}=\lambda(\overrightarrow{OB}-\overrightarrow{OA}).$$

知乎@余命数

整理后可得

$$\overrightarrow{OP}=(1-\lambda)\overrightarrow{OA}+\lambda\overrightarrow{OB}.$$

知乎@余命数

这是一个简洁的形式. 若令 $x=1-\lambda$ 及 $y=\lambda$, 则上式变为

知乎@余命数

$$\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB},\quad \text{其中}\quad x+y=1.$$

知乎@余命数

这便得到了平面上三点共线的向量表示定理. 它的逆命题同样成立. 这个结论是后续讨论的基础.

知乎@余命数

定理 11.5 (三点共线定理) 设 O 为平面内任意一点, 且 $A\neq B$, 则 P, A, B 三点共线的充要条件是, 存在唯一的实数对 (x, y) 使得

$$\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB},\quad \text{且}\quad x+y=1.$$

余命数

证明 为证明该充要条件, 需分别证明其充分性与必要性.

先证明, **充分性.** 假设存在实数对 (x,y) 满足 $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$ 且 $x+y=1$. 需证明 P,A,B 三点共线.

为证明三点共线, 只需证明连接其中两点的向量, 例如 $\overrightarrow{AP}$, 与连接另外两点的向量, 例如 $\overrightarrow{AB}$, 是共线向量即可. 构造向量 $\overrightarrow{AP}$:

$$\overrightarrow{AP}=\overrightarrow{OP}-\overrightarrow{OA}.$$

将假设条件 $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$ 代入上式, 得

$$\overrightarrow{AP}=(x\overrightarrow{OA}+y\overrightarrow{OB})-\overrightarrow{OA}=(x-1)\overrightarrow{OA}+y\overrightarrow{OB}.$$

利用系数和条件 $x+y=1$, 有 $x-1=-y$. 将此关系代入, 得

$$\overrightarrow{AP}=-y\overrightarrow{OA}+y\overrightarrow{OB}=y(\overrightarrow{OB}-\overrightarrow{OA}).$$

注意到 $\overrightarrow{OB}-\overrightarrow{OA}$ 正是向量 $\overrightarrow{AB}$. 于是, 得

$$\overrightarrow{AP}=y\overrightarrow{AB}.$$

此式表明向量 $\overrightarrow{AP}$ 是向量 $\overrightarrow{AB}$ 的一个标量倍. 因此, $\overrightarrow{AP}$ 与 $\overrightarrow{AB}$ 共线. 又因为这两个向量拥有共同的起点 A, 故 P,A,B 三点必然位于同一条直线上. 充分性得证.

接下来证明, 必要性与唯一性. 假设 P,A,B 三点共线. 需证明存在唯一的实数对 (x,y) 满足 $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$ 且 $x+y=1$.

因为 $A\neq B$ 且 P,A,B 三点共线, 向量 $\overrightarrow{AP}$ 与非零向量 $\overrightarrow{AB}$ 共线. 故存在唯一的实数 λ, 使得

$$\overrightarrow{AP}=\lambda\overrightarrow{AB}.$$

将此式用以 O 为基点的位置向量表示:

$$\overrightarrow{OP}-\overrightarrow{OA}=\lambda(\overrightarrow{OB}-\overrightarrow{OA}).$$

整理可得 $\overrightarrow{OP}$ 的表达式:

$$\overrightarrow{OP}=\overrightarrow{OA}+\lambda\overrightarrow{OB}-\lambda\overrightarrow{OA}=(1-\lambda)\overrightarrow{OA}+\lambda\overrightarrow{OB}.$$

令 $x=1-\lambda$ 且 $y=\lambda$. 此时, 找到满足 $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$ 的系数. 验证其和: $x+y=(1-\lambda)+\lambda=1$. 由于对于给定的共线三点, λ 的值, 即 $\overrightarrow{AP}$ 在 $\overrightarrow{AB}$ 上的投影系数是唯一确定的, 故 x 和 y 的值也是唯一确定的. 必要性与唯一性得证.

故该定理成立. □

接着要考虑的是, 当系数和为任意常数 k, 即 $x+y=k$ 时, 点 P 的轨迹.

考察这个条件: $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$, 且 $x+y=k$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

首先考虑特殊情况. 当 $k=0$ 时, $y=-x$. 于是

$$\overrightarrow{OP}=x\overrightarrow{OA}-x\overrightarrow{OB}=x(\overrightarrow{OA}-\overrightarrow{OB})=x\overrightarrow{BA}.$$

故向量 $\overrightarrow{OP}$ 与向量 $\overrightarrow{BA}$ 共线. 由于 $\overrightarrow{OP}$ 的起点是原点 O, 故点 P 的轨迹是一条经过原点 O 且平行于直线 AB 的直线.

接着, 我们考察 $k\neq 0$ 的一般情况. 我们的策略是通过代数变形, 将表达式整理成能够说明其几何意义的形式. 从条件 $x+y=k$ 中, 可用 x 表示 y, 即 $y=k-x$. 将其代入向量表达式:

$$\overrightarrow{OP}=x\overrightarrow{OA}+(k-x)\overrightarrow{OB}.$$

为了分离变量 x 的影响, 将上式按 x 和常数项重新组合:

$$\overrightarrow{OP}=x\overrightarrow{OA}+k\overrightarrow{OB}-x\overrightarrow{OB}=k\overrightarrow{OB}+x(\overrightarrow{OA}-\overrightarrow{OB}).$$

注意到括号中的项正是向量 $\overrightarrow{BA}$, 于是

$$\overrightarrow{OP}=k\overrightarrow{OB}+x\overrightarrow{BA}.$$

这个结构是向量几何中直线的参数方程. 为了更清晰地理解其几何轨迹, 可将常数项移到等式左边:

$$\overrightarrow{OP}-k\overrightarrow{OB}=x\overrightarrow{BA}.$$

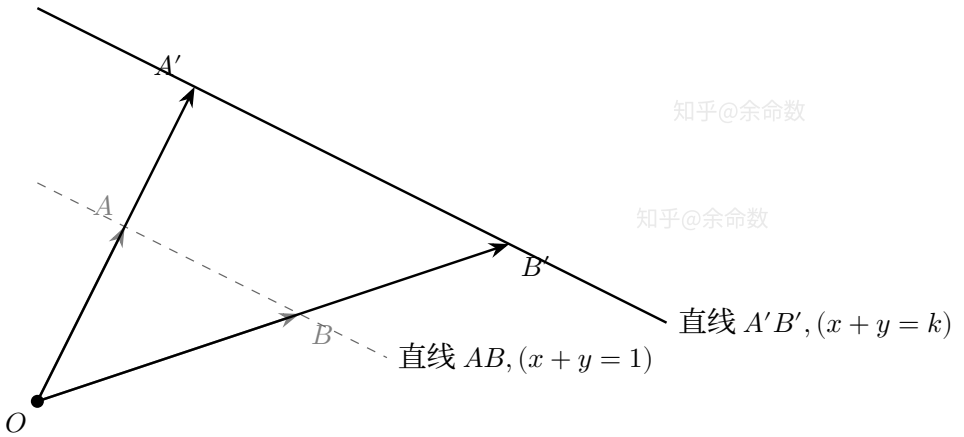
令 B' 是这样一个点, 其位置向量为 $\overrightarrow{OB'}=k\overrightarrow{OB}$. 由于 k 和点 B 都是固定的, 所以 B' 是一个定点. 上述方程可以写为

$$\overrightarrow{B'P}=x\overrightarrow{BA}.$$

此式的几何意义非常明确, 从定点 B' 到动点 P 的向量 $\overrightarrow{B'P}$, 总是与一个固定的向量 $\overrightarrow{BA}$ 共线. 因此, 动点 P 的轨迹必然是一条经过定点 B' 且方向与 $\overrightarrow{BA}$, 即直线 AB 相同的直线.

同理, 我们也可以定义一个定点 A' 满足 $\overrightarrow{OA'}=k\overrightarrow{OA}$, 通过消去 x, 将会证明点 P 的轨迹也经过 A'. 因此, 当 $x+y=k$ $(k\neq 0)$ 时, 点 P 的轨迹是一条经过点 $A'(k\overrightarrow{OA})$ 和 $B'(k\overrightarrow{OB})$ 的直线, 并且该直线平行于直线 AB.

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



知乎@余命数

将这一重要结论总结如下.

定理 11.6 (等和线定理) 给定平面上不共线的三点 O, A, B, 对于动点 P, 若其位置向量满足

$$\overrightarrow{OP} = x\overrightarrow{OA} + y\overrightarrow{OB},$$

则点 P 的轨迹由系数和 $x + y$ 的值唯一确定:

- 若 $x + y = 1$, 则 P 的轨迹是直线 AB.
- 若 $x + y = k$ (k 为非零常数), 则 P 的轨迹是一条经过点 $A'(k\overrightarrow{OA})$ 和 $B'(k\overrightarrow{OB})$ 的直线, 且该直线平行于 AB.
- 若 $x + y = 0$, 则 P 的轨迹是一条经过原点 O 且平行于 AB 的直线.

这条由系数和为常数所确定的直线, 我们称之为等和线.

知乎@余命数

证明 分三种情形证明.

- 情况一 $x + y = 1$
  此情况即为前述的三点共线定理, 其充分性与必要性已得到严格证明. 它是整个等和线理论的基础.
- 情况二 $x + y = 0$
  充分性: 假设 $\overrightarrow{OP} = x\overrightarrow{OA} + y\overrightarrow{OB}$ 且 $x + y = 0$. 由 $x + y = 0$ 可得 $y = -x$. 将其代入向量表达式:

  $$\overrightarrow{OP} = x\overrightarrow{OA} - x\overrightarrow{OB} = x(\overrightarrow{OA} - \overrightarrow{OB}) = x\overrightarrow{BA}.$$

  知乎@余命数

  此式表明, 向量 $\overrightarrow{OP}$ 与向量 $\overrightarrow{BA}$ 共线. 由于 $\overrightarrow{OP}$ 的起点为 O, 故点 P 必然位于那条经过原点 O 且与直线 AB 平行的直线上.

知乎@余命数

必要性：假设点 P 位于经过原点 O 且平行于直线 AB 的直线上. 即向量 $\overrightarrow{OP}$ 与向量 $\overrightarrow{AB}$(或 $\overrightarrow{BA}$)共线. 因此, 存在唯一的实数 t 使得

$$\overrightarrow{OP} = t\overrightarrow{BA}.$$

将 $\overrightarrow{BA}$ 写成 $\overrightarrow{OA} - \overrightarrow{OB}$, 得:

$$\overrightarrow{OP} = t(\overrightarrow{OA} - \overrightarrow{OB}) = t\overrightarrow{OA} - t\overrightarrow{OB}.$$

若令 $x = t$ 及 $y = -t$, 则上式符合 $\overrightarrow{OP} = x\overrightarrow{OA} + y\overrightarrow{OB}$ 的形式. 此时系数和为 $x + y = t + (-t) = 0$. 证毕.

- 情况三 $x + y = k$(k 为非零常数)

充分性：假设 $\overrightarrow{OP} = x\overrightarrow{OA} + y\overrightarrow{OB}$ 且 $x + y = k$. 定义点 A' 为满足 $\overrightarrow{OA'} = k\overrightarrow{OA}$ 的定点. 考察向量 $\overrightarrow{A'P}$:

$$\overrightarrow{A'P} = \overrightarrow{OP} - \overrightarrow{OA'} = (x\overrightarrow{OA} + y\overrightarrow{OB}) - k\overrightarrow{OA} = (x-k)\overrightarrow{OA} + y\overrightarrow{OB}.$$

利用系数和条件 $x + y = k$, 有 $x - k = -y$. 将此关系代入, 得

$$\overrightarrow{A'P} = -y\overrightarrow{OA} + y\overrightarrow{OB} = y(\overrightarrow{OB} - \overrightarrow{OA}) = y\overrightarrow{AB}.$$

此式表明, 向量 $\overrightarrow{A'P}$ 与向量 $\overrightarrow{AB}$ 共线. 因为 A' 是一个定点, 所以动点 P 的轨迹是一条经过 A' 且平行于直线 AB 的直线. 类似地, 可证明该直线也经过点 $B'(k\overrightarrow{OB})$.

必要性：假设点 P 位于经过点 $A'(k\overrightarrow{OA})$ 且平行于直线 AB 的直线上. 即向量 $\overrightarrow{A'P}$ 与向量 $\overrightarrow{AB}$ 共线. 因此, 存在唯一的实数 t 使得

$$\overrightarrow{A'P} = t\overrightarrow{AB}.$$

用位置向量展开上式:

$$\overrightarrow{OP} - \overrightarrow{OA'} = t(\overrightarrow{OB} - \overrightarrow{OA}).$$

解出 $\overrightarrow{OP}$:

$$\overrightarrow{OP} = \overrightarrow{OA'} + t\overrightarrow{OB} - t\overrightarrow{OA}.$$

将 $\overrightarrow{OA'} = k\overrightarrow{OA}$ 代入, 得

$$\overrightarrow{OP} = k\overrightarrow{OA} + t\overrightarrow{OB} - t\overrightarrow{OA} = (k-t)\overrightarrow{OA} + t\overrightarrow{OB}.$$

若令 $x = k - t$ 及 $y = t$, 则上式符合 $\overrightarrow{OP} = x\overrightarrow{OA} + y\overrightarrow{OB}$ 的形式. 此时系数和为 $x + y = (k-t) + t = k$. 证毕.

知乎@余命数

综合以上三种情况, 定理得证. □

这一结论极为有用, 它将一个代数条件 ($x+y=k$) 和一个清晰的几何轨迹(一条特定的直线)对应起来, 是解决向量与几何综合问题的有力工具.

为进一步理解等和线定理的含义, 可从三种观点分析这一结论.

11.2.1.1 仿射坐标系

知乎@余命数

我们熟悉的笛卡尔坐标系, 是由一个原点和两个相互垂直的单位向量基底构成的. 而等和线的概念实质上是在一个更广义的坐标系, 即**仿射坐标系**中讨论.

给定一个基点 O 和两个线性无关的向量 $\overrightarrow{OA}$ 和 $\overrightarrow{OB}$, 它们构成了一个仿射坐标系, 记为 $(O;\overrightarrow{OA},\overrightarrow{OB})$. 在此坐标系中, 平面上任意一点 P 的位置向量 $\overrightarrow{OP}$ 都可以被唯一地表示为

$$\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}.$$

这里的数对 (x,y) 就是点 P 在该仿射坐标系下的 **仿射坐标**.

从这个视角来看, 等和线定理的条件 $x+y=k$ 就十分清晰. 它是在仿射坐标系下的一条**线性方程**. 已知, 在笛卡尔坐标系中, 任何形如 $ax+by=c$ 的线性方程都对应一条直线. 仿射坐标系保持了这种最基本的线性性质.

- $x+y=1$: 这条特殊的等和线, 即直线 AB, 在仿射坐标系 $(O;\overrightarrow{OA},\overrightarrow{OB})$ 中是如同笛卡尔坐标系中 $x+y=1$ 的, 是连接两个单位点 $A(1,0)$ 和 $B(0,1)$ 的基准线.
- $x+y=k$: 这族方程对应的是一族互相平行的直线, 正如在笛卡尔坐标系中 $x+y=k$ 对应一族斜率为 -1 的平行直线系一样.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

因此, 等和线定理实质上是陈述了在向量定义的仿射坐标系中, 点的共线关系等价于其仿射坐标满足一个线性方程.

11.2.1.2 几何变换

等和线定理也可以被理解为一种几何变换作用的结果. 具体的说, 是**位似变换**.

以原点 O 为位似中心, 位似比为 k 的位似变换, 将点 P' 映射到点 P, 其定义就是 $\overrightarrow{OP}=k\overrightarrow{OP'}$.

现再看等和线的推导过程. 我们证明了, 如果点 P' 的位置向量满足 $\overrightarrow{OP'}=x'\overrightarrow{OA}+y'\overrightarrow{OB}$ 且 $x'+y'=1$, 即 P' 在直线 AB 上, 那么经过位似变换后的点 P, 其位置向量 $\overrightarrow{OP}=k\overrightarrow{OP'}=kx'\overrightarrow{OA}+ky'\overrightarrow{OB}$. 令 $x=kx'$ 及 $y=ky'$, 则点 P 的仿射坐标为 (x,y), 且其系数和为

$$x+y=kx'+ky'=k(x'+y')=k(1)=k.$$

知乎@余命数

这精确地说明了: **系数和为 k 的等和线, 正是系数和为 1 的基准线(直线 AB)在以 O 为中心、比值为 k 的位似变换下的像**. 位似变换的一个基本性质就是将直线映射为与之平行的直线. 这就从几何变换的角度解释了为何所有的等和线都相互平行.

11.2.1.3 重心坐标

对于 $\triangle OAB$, 平面内任意一点 P 都可以被唯一表示为三个顶点的线性组合:

$$\overrightarrow{OP}=u\overrightarrow{OA}+v\overrightarrow{OB}+w\overrightarrow{OO},\quad \text{其中}\quad u+v+w=1.$$

由于 $\overrightarrow{OO}$ 是零向量, 上式简化为 $\overrightarrow{OP}=u\overrightarrow{OA}+v\overrightarrow{OB}$, 且 $u+v+w=1$. 这里的 (u,v,w) 称为点 P 关于 $\triangle OAB$ 的**重心坐标**.

现再看条件 $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$. 这可以看作是重心坐标表示中第三个系数 w (对应于顶点 O)的信息. 比较两种表示, 有 $u=x, v=y$. 由重心坐标的系数和为 1, 即 $u+v+w=1$, 得

$$x+y+w=1\implies w=1-(x+y).$$

于是, 等和线的条件 $x+y=k$ 等价于一个关于重心坐标的条件:

$$w=1-k.$$

等和线 $x+y=k$ 上所有点的重心坐标中, 对应于顶点 O 的分量 w 是一个常数 $1-k$. 在重心坐标系中, 一个顶点对应的坐标分量为常数的点的轨迹, 是一条平行于该顶点对边的直线. 在这里, w 是对应顶点 O 的坐标, 其对边是线段 AB. 因此, $w=1-k$ 所描述的点的轨迹是一条平行于直线 AB 的直线.

11.2.1.4 推广至高维

以上观点, 特别是仿射坐标与重心坐标的观点, 很容易推广到三维乃至更高维空间. 例如在三维空间中, 给定基点 O 和三个不共面的点 A,B,C, 那么满足

$$\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}+z\overrightarrow{OC},\quad \text{且}\quad x+y+z=k$$

的点 P 的轨迹, 就是一个平行于平面 ABC 的平面. 这正是等和线定理在高维空间中的自然推广.

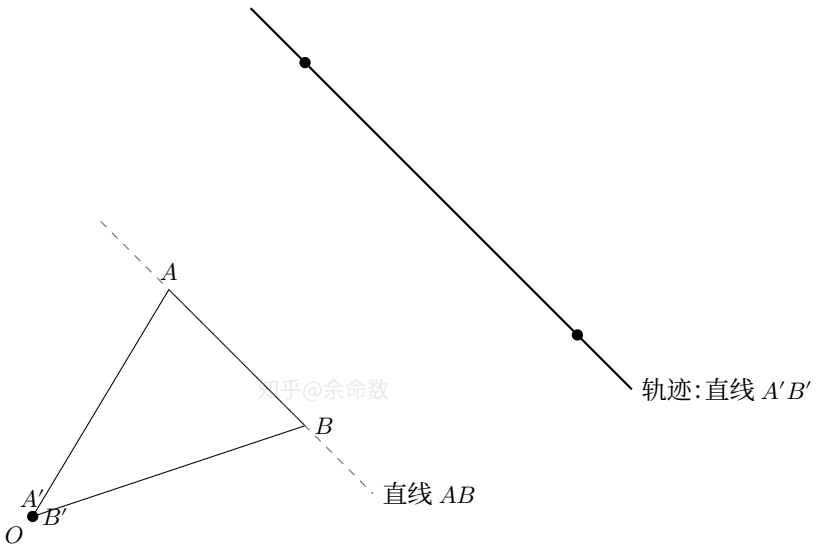
例题 11.24 在 $\triangle OAB$ 中, 点 P 位于平面内, 且满足 $\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}$. 若 $\triangle OAB$ 的面积为 1, 求在下列条件下点 P 的轨迹.

1. $x+y=2$;
2. $x+2y=2$ 且 $x\geq 0, y\geq 0$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 此题旨在将向量系数的代数约束, 翻译为点 P 的几何轨迹.

(1) 条件为 $x+y=2$. 根据等和线定理, 当系数和为常数 k 时, 点 P 的轨迹是一条平行于 AB 的直线. 此处 $k=2$. 这条直线经过点 A' 和 B', 其中 $\overrightarrow{OA'}=2\overrightarrow{OA}$ 且 $\overrightarrow{OB'}=2\overrightarrow{OB}$. 因此, 点 P 的轨迹是直线 $A'B'$.



(2) 条件为 $x+2y=2$, 且 $x\geq 0, y\geq 0$. 这个条件 $x+2y=2$ 需要先化归为等和线形式. 条件两边同除以 2, 得到

$$\frac{x}{2}+y=1.$$

这提示应该构造一个新基向量, 使得其系数恰为 $\frac{x}{2}$ 和 y.

$$\overrightarrow{OP}=x\overrightarrow{OA}+y\overrightarrow{OB}=\frac{x}{2}(2\overrightarrow{OA})+y\overrightarrow{OB}.$$

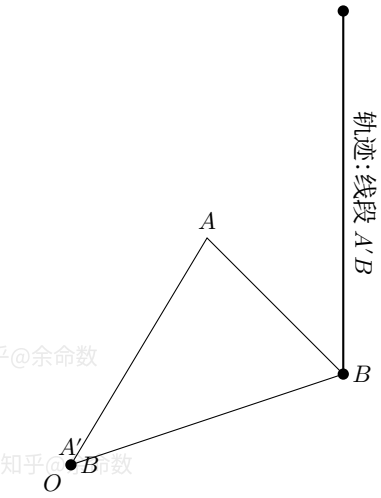
令 $\overrightarrow{OA'}=2\overrightarrow{OA}$. 那么

$$\overrightarrow{OP}=\frac{x}{2}\overrightarrow{OA'}+y\overrightarrow{OB}.$$

令 $x'=\frac{x}{2}$ 和 $y'=y$. 则 $x'+y'=1$. 根据三点共线定理, 点 P 的轨迹在直线 $A'B$ 上.

然而, 题目中还有附加条件 $x\geq 0, y\geq 0$. $x\geq 0 \implies x'\geq 0$. $y\geq 0 \implies y'\geq 0$. 因此, 点 P 的轨迹是系数 x', y' 均非负的线段, 即线段 $A'B$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



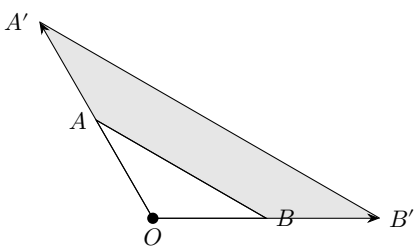
这个例子说明, 任何关于系数的线性约束 $ax+by=c$, 都可以通过对基向量的伸缩变换, 转化为系数和为 1 的标准形式, 从而确定其轨迹为一条直线.

例题 11.25 给定 $\triangle OAB$, 其中 $OA=OB=1$, $\angle AOB=120^\circ$. 点 P 在平面内运动, 其位置向量 $\overrightarrow{OP}=s\overrightarrow{OA}+t\overrightarrow{OB}$. 若实数 s,t 满足 $s\geq 0, t\geq 0$ 且 $1\leq s+t\leq 2$, 求点 P 所形成区域的面积.

解 此题要求由向量系数的不等式约束所确定的点的轨迹区域的面积. 这是一个典型的由等和线界定区域的问题. 分析点 P 的轨迹边界. 约束条件为 $1\leq s+t\leq 2$, 同时 $s\geq 0, t\geq 0$.

- 考虑边界 $s+t=1$. 根据等和线定理, 这对应一条直线. 又因 $s\geq 0, t\geq 0$, 点 P 的轨迹是连接点 $A(s=1,t=0)$ 和点 $B(s=0,t=1)$ 的线段 AB.
- 考虑边界 $s+t=2$. 根据等和线定理, 这对应另一条与 AB 平行的直线. 这条直线经过点 $A'(s=2,t=0)$ 和点 $B'(s=0,t=2)$. 其中, $\overrightarrow{OA'}=2\overrightarrow{OA}$ 且 $\overrightarrow{OB'}=2\overrightarrow{OB}$. 同样因为 $s\geq 0, t\geq 0$, 点 P 的轨迹是连接 A' 和 B' 的线段 $A'B'$.

当 $s+t$ 的值在 $[1,2]$ 区间内连续变化时, 点 P 的轨迹会扫过线段 AB 与线段 $A'B'$ 之间的所有区域. 因此, 点 P 所形成的区域是一个以 A,B,B',A' 为顶点的梯形.



知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为求该梯形的面积, 可用大三角形 $\triangle OA'B'$ 的面积减去小三角形 $\triangle OAB$ 的面积. 首先, 计算 $\triangle OAB$ 的面积, 得

$$S_{\triangle OAB}=\frac{1}{2}|\overrightarrow{OA}||\overrightarrow{OB}|\sin(\angle AOB)=\frac{1}{2}\cdot 1\cdot 1\cdot\sin(120^\circ)=\frac{1}{2}\cdot\frac{\sqrt{3}}{2}=\frac{\sqrt{3}}{4}.$$

$\triangle OA'B'$ 是 $\triangle OAB$ 以 O 为中心, 位似比为 2 的放大图形. 它的边长 $OA'=2, OB'=2$, 夹角不变. 其面积为:

$$S_{\triangle OA'B'}=\frac{1}{2}|\overrightarrow{OA'}||\overrightarrow{OB'}|\sin(\angle A'OB')=\frac{1}{2}\cdot 2\cdot 2\cdot\sin(120^\circ)=\sqrt{3}.$$

或者, 利用面积的位似性质:

$$S_{\triangle OA'B'}=2^2\cdot S_{\triangle OAB}=4\cdot\frac{\sqrt{3}}{4}=\sqrt{3}.$$

因此, 梯形 $ABB'A'$ 的面积为

$$S_{\text{梯形}}=S_{\triangle OA'B'}-S_{\triangle OAB}=\sqrt{3}-\frac{\sqrt{3}}{4}=\frac{3\sqrt{3}}{4}.$$

11.2.2 奔驰定理

已分析了由两个基向量线性组合所定义的点的轨迹. 对于三角形的三个顶点, 它们的向量之和具有确定的几何意义, 这正是三角形一个基本中心, 即重心的内在向量性质.

定理 11.7 (重心向量定理) 对于平面上任意 $\triangle ABC$ 及其重心 G, 以下两个向量恒等式成立:

1. 对于任意选定的基点 O, 顶点的位置向量之和等于重心位置向量的三倍:
$$\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}=3\overrightarrow{OG}.$$
2. 对于平面上任意一点 P, 从该点出发到三个顶点的向量之和等于到重心向量的三倍:
$$\overrightarrow{PA}+\overrightarrow{PB}+\overrightarrow{PC}=3\overrightarrow{PG}.$$

证明 将从重心的基本几何性质出发, 即重心是中线的交点, 且按 $2:1$ 的比例分割中线, 来推导这两个恒等式.

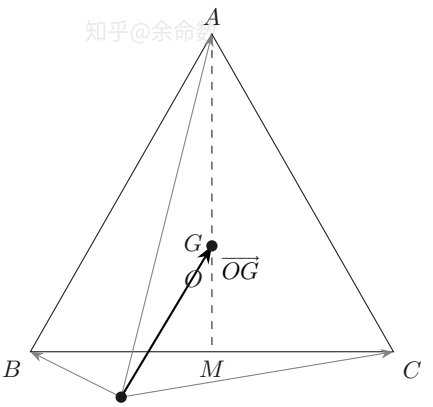
1. 证明 $\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}=3\overrightarrow{OG}$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设 M 为边 BC 的中点. 根据中点向量公式, 对于任意选定的基点 O, 点 M 的位置向量 $\overrightarrow{OM}$ 可以表示为:

$$\overrightarrow{OM}=\frac{1}{2}(\overrightarrow{OB}+\overrightarrow{OC}).$$

重心 G 是 $\triangle ABC$ 的三条中线的交点. 特别地, G 位于中线 AM 上. 由重心的性质可知, 点 G 将中线 AM 分成两段, 且满足比例关系 $AG:GM=2:1$.



可利用定比分点公式来表示点 G 的位置向量 $\overrightarrow{OG}$. 因为 G 在线段 AM 上, 且 $AG:GM=2:1$, 可将 $\overrightarrow{OG}$ 表示为 $\overrightarrow{OA}$ 和 $\overrightarrow{OM}$ 的线性组合:

$$\overrightarrow{OG}=\frac{1\cdot\overrightarrow{OA}+2\cdot\overrightarrow{OM}}{1+2}=\frac{1}{3}(\overrightarrow{OA}+2\overrightarrow{OM}).$$

现在, 将 $\overrightarrow{OM}$ 的表达式 $\frac{1}{2}(\overrightarrow{OB}+\overrightarrow{OC})$ 代入上式, 得

$$\overrightarrow{OG}=\frac{1}{3}\left(\overrightarrow{OA}+2\cdot\left(\frac{1}{2}(\overrightarrow{OB}+\overrightarrow{OC})\right)\right).$$

化简括号内的表达式, 得:

$$\overrightarrow{OG}=\frac{1}{3}(\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}).$$

将等式两边同乘以 3, 便得到了第一个恒等式, 得

$$3\overrightarrow{OG}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}.$$

此式的成立不依赖于基点 O 的选择, 故对任意点 O 均成立.

2. 证明 $\overrightarrow{PA}+\overrightarrow{PB}+\overrightarrow{PC}=3\overrightarrow{PG}$

此恒等式可通过更换基点由第一个恒等式导出. 在第一个恒等式中, 基点 O 是任意的. 可将这个任意点 O 替换为点 P. 于是, 恒等式变为:

$$\overrightarrow{PA}+\overrightarrow{PB}+\overrightarrow{PC}=3\overrightarrow{PG}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这种方法虽然简洁, 但为了更清晰地展示推导过程, 也可从向量的减法出发证明. 对于任意点 P, 可将从 P 出发的向量分解为经过基点 O 的路径:

$$\overrightarrow{PA} = \overrightarrow{OA} - \overrightarrow{OP}$$

$$\overrightarrow{PB} = \overrightarrow{OB} - \overrightarrow{OP}$$

$$\overrightarrow{PC} = \overrightarrow{OC} - \overrightarrow{OP}$$

将这三个向量表达式相加:

$$\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC} = (\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}) - 3\overrightarrow{OP}.$$

利用已证明的第一个恒等式 $\overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC} = 3\overrightarrow{OG}$, 将其代入上式, 得

$$\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC} = 3\overrightarrow{OG} - 3\overrightarrow{OP}.$$

提取公因子 3:

$$\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC} = 3(\overrightarrow{OG} - \overrightarrow{OP}).$$

根据向量减法的定义, $\overrightarrow{OG} - \overrightarrow{OP}$ 正是向量 $\overrightarrow{PG}$. 因此,

$$\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC} = 3\overrightarrow{PG}.$$

至此, 定理的两个形式均已得证. □

这个定理, 尤其是它的一个重要推论, 在几何学中被赋予了一个形象的名称, 即奔驰定理.

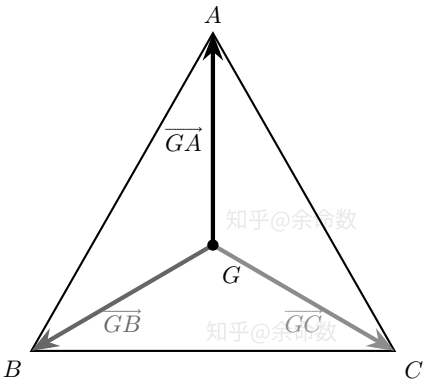
11.2.2.1 定理的几何形象与名称由来

若将基点 O 选为重心 G 本身, 那么定理的第一个恒等式就变为

$$\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = 3\overrightarrow{GG} = \vec{0}.$$

这个简洁而基本的结论 $\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = \vec{0}$ 就是奔驰定理在基点取为重心时的特例. 它表明, 如果将三个从重心出发指向顶点的向量首尾相接, 它们将构成一个封闭的三角形. 而如果将这三个向量的起点都置于重心 G, 它们会形成一个静态的力平衡, 指向三个方向. 这三个向量构成的 Y 形结构, 与梅赛德斯-奔驰的标志惊人地相似, 定理因此得名. 这也从物理学的角度解释了重心为什么被称为质量中心.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵



奔驰定理是处理涉及重心问题时的一个基础出发点. 它将三个顶点的信息集中到重心这一点上, 大幅简化了许多复杂的向量关系.

11.2.3 二次和的最小值与斯坦纳定理

奔驰定理说明了从任意点到三角形三顶点向量之和的简洁性质. 一个自然的延伸是寻找从平面上哪一点出发到三顶点的**距离平方和**最小. 这个问题在几何、物理与统计中都有着重要的应用, 物理上它对应力矩, 统计上它对应方差, 其答案由如下的斯坦纳定理或称莱布尼茨公式给出.

定理 11.8 (向量的斯坦纳定理) 对于平面上任意 $\triangle ABC$, 其重心为 G. 对于平面上任意一点 P, 以下恒等式成立:

$$PA^2 + PB^2 + PC^2 = GA^2 + GB^2 + GC^2 + 3PG^2.$$

或者用向量模长的形式写作:

$$||\overrightarrow{PA}||^2 + ||\overrightarrow{PB}||^2 + ||\overrightarrow{PC}||^2 = ||\overrightarrow{GA}||^2 + ||\overrightarrow{GB}||^2 + ||\overrightarrow{GC}||^2 + 3||\overrightarrow{PG}||^2.$$

这个恒等式说明了一个基本结构, 从任意点 P 出发的距离平方和等于一个不依赖于 P 的常数, 即从重心 G 出发的距离平方和, 加上一个正比于 P 到重心 G 距离平方的变量. 这在物理学中, 与质点系的转动惯量中的平行轴定理具有完全相同的数学形式.

证明 此定理的证明是奔驰定理的一个直接且精彩的应用. 出发点是将所有与动点 P 相关的向量, 都通过重心 G 这个不动点进行分解. 对于向量 $\overrightarrow{PA}$, 可插入点 G, 得到路径 $P \to G \to A$:

$$\overrightarrow{PA} = \overrightarrow{PG} + \overrightarrow{GA}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

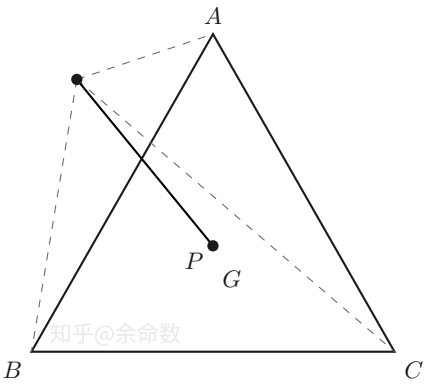
接着, 计算其模长的平方, 得

$$\begin{aligned}||\overrightarrow{PA}||^2 &= (\overrightarrow{PG}+\overrightarrow{GA})\cdot(\overrightarrow{PG}+\overrightarrow{GA})\\ &= ||\overrightarrow{PG}||^2+2(\overrightarrow{PG}\cdot\overrightarrow{GA})+||\overrightarrow{GA}||^2.\end{aligned}$$

同理, 可得到关于点 B 和 C 的表达式:

$$||\overrightarrow{PB}||^2 = ||\overrightarrow{PG}||^2+2(\overrightarrow{PG}\cdot\overrightarrow{GB})+||\overrightarrow{GB}||^2,$$

$$||\overrightarrow{PC}||^2 = ||\overrightarrow{PG}||^2+2(\overrightarrow{PG}\cdot\overrightarrow{GC})+||\overrightarrow{GC}||^2.$$



将这三个等式相加, 得到所关心的距离平方和:

$$\begin{aligned}||\overrightarrow{PA}||^2+||\overrightarrow{PB}||^2+||\overrightarrow{PC}||^2 &= \left(3||\overrightarrow{PG}||^2\right)\\ &\quad+\left(||\overrightarrow{GA}||^2+||\overrightarrow{GB}||^2+||\overrightarrow{GC}||^2\right)\\ &\quad+2(\overrightarrow{PG}\cdot\overrightarrow{GA}+\overrightarrow{PG}\cdot\overrightarrow{GB}+\overrightarrow{PG}\cdot\overrightarrow{GC}).\end{aligned}$$

利用数量积的分配律, 可将最后一项中的公因子 $\overrightarrow{PG}$ 提取出来:

$$2\overrightarrow{PG}\cdot(\overrightarrow{GA}+\overrightarrow{GB}+\overrightarrow{GC}).$$

此时, 奔驰定理的结论 $\overrightarrow{GA}+\overrightarrow{GB}+\overrightarrow{GC}=\vec{0}$ 发挥了决定性作用. 上述交叉项因此变为

$$2\overrightarrow{PG}\cdot\vec{0}=0.$$

于是, 总和的表达式中交叉项消失, 化简为:

$$||\overrightarrow{PA}||^2+||\overrightarrow{PB}||^2+||\overrightarrow{PC}||^2 = 3||\overrightarrow{PG}||^2+||\overrightarrow{GA}||^2+||\overrightarrow{GB}||^2+||\overrightarrow{GC}||^2.$$

这便完成了恒等式的证明. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

推论 11.1 在平面上, 使得与 $\triangle ABC$ 三个顶点的距离平方和 $PA^2+PB^2+PC^2$ 取得最小值的点 P, 唯一地是该三角形的重心 G.

证明 根据斯坦纳定理的恒等式, 表达式 $PA^2+PB^2+PC^2$ 由两部分构成. 第一部分, $GA^2+GB^2+GC^2$, 对于一个给定的三角形, 是一个确定的常数. 第二部分, $3PG^2$, 是唯一与动点 P 位置相关的变量. 由于模长的平方 PG^2 总是非负的, 即 $PG^2\geq 0$, 那么 $3PG^2$ 的最小值显然为 0. 这个最小值在且仅在 P 与 G 重合时 $(P=G)$ 取得. 因此,总和在点 $P=G$ 时达到其最小值,最小值为 $GA^2+GB^2+GC^2$. □

11.2.4 对角线数量积定理

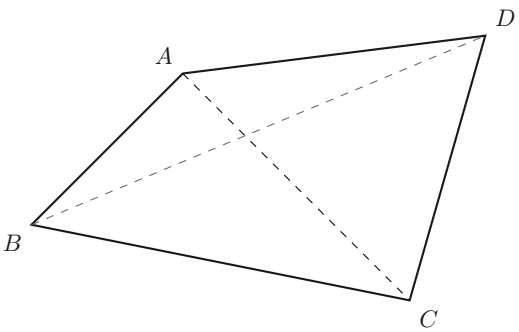
在详细研究了三角形的向量性质后, 我们自然地将目光转向更为复杂的四边形. 四边形的四条边长与其两条对角线的长度和方向之间存在确定的代数关系, 这个关系由一个关于对角线向量数量积的简洁恒等式所刻画.

定理 11.9 对于平面上任意一个四边形 $ABCD$, 其边长与对角线向量的数量积满足如下关系:

$$2(\overrightarrow{AC}\cdot\overrightarrow{BD})=(AD^2+BC^2)-(AB^2+CD^2).$$

该恒等式表明, 对角线向量的数量积等于两组对边长度平方和之差的一半.

证明 此定理的证明是向量代数处理几何问题的典范. 其做法是引入一个任意的基点 O 作为原点, 将所有几何量(边长和对角线)用相对于 O 的位置向量 $\vec{a},\vec{b},\vec{c},\vec{d}$ 来表示, 从而将几何问题彻底转化为代数恒等式的验证.



知乎@余命数

令 $\overrightarrow{OA}=\vec{a},\overrightarrow{OB}=\vec{b},\overrightarrow{OC}=\vec{c},\overrightarrow{OD}=\vec{d}$. 从恒等式右侧的边长平方和之差入手计算.

$$\begin{aligned}(AD^2+BC^2)-(AB^2+CD^2)&=||\vec{d}-\vec{a}||^2+||\vec{c}-\vec{b}||^2\\&\quad-||\vec{b}-\vec{a}||^2-||\vec{d}-\vec{c}||^2\\&=(\vec{d}-\vec{a})\cdot(\vec{d}-\vec{a})\\&\quad+(\vec{c}-\vec{b})\cdot(\vec{c}-\vec{b})\\&\quad-(\vec{b}-\vec{a})\cdot(\vec{b}-\vec{a})\\&\quad-(\vec{d}-\vec{c})\cdot(\vec{d}-\vec{c}).\end{aligned}$$

知乎@余命数

将所有模长平方展开, 得

$$\begin{aligned}&=||\vec{d}||^2-2\vec{a}\cdot\vec{d}+||\vec{a}||^2\\&\quad+||\vec{c}||^2-2\vec{b}\cdot\vec{c}+||\vec{b}||^2\\&\quad-(||\vec{b}||^2-2\vec{a}\cdot\vec{b}+||\vec{a}||^2)\\&\quad-(||\vec{d}||^2-2\vec{c}\cdot\vec{d}+||\vec{c}||^2).\end{aligned}$$

知乎@余命数

展开并移除括号后, 所有的模长平方项 $||\vec{a}||^2,||\vec{b}||^2,||\vec{c}||^2,||\vec{d}||^2$ 都相互抵消. 剩下的项为:

$$=-2\vec{a}\cdot\vec{d}-2\vec{b}\cdot\vec{c}+2\vec{a}\cdot\vec{b}+2\vec{c}\cdot\vec{d}.$$

知乎@余命数

提取公因子 2, 并重新组合这些项, 以便进行因式分解:

知乎@余命数

$$\begin{aligned}&=2(\vec{a}\cdot\vec{b}-\vec{a}\cdot\vec{d}-\vec{b}\cdot\vec{c}+\vec{c}\cdot\vec{d})\\&=2(\vec{a}\cdot(\vec{b}-\vec{d})-\vec{c}\cdot(\vec{b}-\vec{d}))\\&=2((\vec{a}-\vec{c})\cdot(\vec{b}-\vec{d})).\end{aligned}$$

最后, 将位置向量的差还原为几何向量: $\vec{a}-\vec{c}=\overrightarrow{OA}-\overrightarrow{OC}=\overrightarrow{CA}=-\overrightarrow{AC}$. $\vec{b}-\vec{d}=\overrightarrow{OB}-\overrightarrow{OD}=\overrightarrow{DB}=-\overrightarrow{BD}$. 于是, 最终的表达式为:

$$2((-\overrightarrow{AC})\cdot(-\overrightarrow{BD}))=2(\overrightarrow{AC}\cdot\overrightarrow{BD}).$$

这就证明了 $(AD^2+BC^2)-(AB^2+CD^2)=2(\overrightarrow{AC}\cdot\overrightarrow{BD})$. □ □

这个定理形式简洁, 还引出了一条关于对角线垂直的几何性质.

知乎@余命数

推论 11.2 (正交对角线定理) 一个四边形的两条对角线相互垂直的充要条件是, 其两组对边的长度平方和相等. 即,

$$AC \perp BD \iff AB^2 + CD^2 = BC^2 + AD^2.$$

证明 根据向量的性质, 两条对角线 AC 与 BD 相互垂直, 等价于它们的向量数量积为零, 即 $\overrightarrow{AC} \cdot \overrightarrow{BD} = 0$. 将此条件代入刚刚证明的对角线数量积定理, 得

$$2(0) = (AD^2 + BC^2) - (AB^2 + CD^2).$$

$$0 = (AD^2 + BC^2) - (AB^2 + CD^2).$$

移项即得

$$AB^2 + CD^2 = BC^2 + AD^2.$$

由于推导过程中的每一步都是可逆的, 因此该条件的充分性与必要性均成立. □

这个推论在解决有关等形、菱形以及其他对角线垂直的四边形问题时, 提供了一个非常有效的边长关系判据.

11.3 向量外积与混合积

数量积把两个向量压缩成一个标量, 刻画夹角与投影. 外积走另一条路, 两个向量相乘仍得向量, 方向垂直于原二向量所在平面, 大小等于它们张成的平行四边形面积. 三个向量的混合积把外积与内积串联起来, 给出平行六面体的体积, 并提供共面的代数判据.

11.3.1 向量外积的定义

定义 11.7 (向量外积) 设 $\mathbf{a}$、$\mathbf{b}$ 是两个向量, 夹角为 $\theta(0 \le \theta \le \pi)$. 它们的外积 $\mathbf{a} \times \mathbf{b}$ 是一个向量, 其大小为

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}|\,|\mathbf{b}| \sin\theta \tag{11-6}$$

方向规定为: 当 $\mathbf{a}$ 与 $\mathbf{b}$ 不共线时, $\mathbf{a} \times \mathbf{b}$ 同时垂直于 $\mathbf{a}$ 与 $\mathbf{b}$, 且 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{a} \times \mathbf{b}$ 构成右手系, 即右手四指从 $\mathbf{a}$ 转向 $\mathbf{b}$, 拇指所指即 $\mathbf{a} \times \mathbf{b}$ 的方向. 当 $\mathbf{a}$ 与 $\mathbf{b}$ 共线时, $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

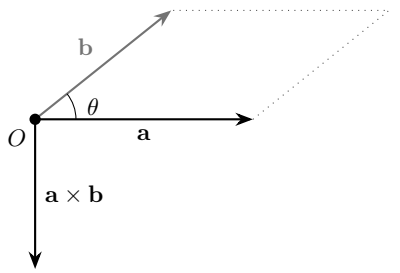


图 11.10　外积的方向由右手定则确定, 大小为平行四边形面积

由 (11-6), $|\mathbf{a}\times\mathbf{b}|$ 等于以 $\mathbf{a}$、$\mathbf{b}$ 为邻边的平行四边形面积. 当 $\mathbf{a}\times\mathbf{b}=\mathbf{0}$ 时 $\sin\theta=0$, 即 $\theta=0$ 或 π, 两个向量共线. 这给出平行的判据:

$$\mathbf{a}\times\mathbf{b}=\mathbf{0}\iff \mathbf{a}\parallel\mathbf{b}$$

11.3.2 外积的运算律

定理 11.10 (外积的运算律) 对任意向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 和实数 λ、μ:

1. 反交换律: $\mathbf{b}\times\mathbf{a}=-\mathbf{a}\times\mathbf{b}$.
2. 数乘结合: $(\lambda\mathbf{a})\times(\mu\mathbf{b})=\lambda\mu(\mathbf{a}\times\mathbf{b})$.
3. 左分配律: $(\mathbf{a}+\mathbf{b})\times\mathbf{c}=\mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$.
4. 右分配律: $\mathbf{a}\times(\mathbf{b}+\mathbf{c})=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}$.

证明 (反交换律) 交换 $\mathbf{a}$ 与 $\mathbf{b}$ 的位置, 夹角从 θ 变为 $\pi-\theta$, $\sin\theta$ 不变, 故大小相同. 右手定则中四指转向相反, 拇指方向翻转, 故方向相反. □

证明 (左分配律) 将各向量投影到垂直于 $\mathbf{c}$ 的平面 π 上. 对 π 内任一向量 $\mathbf{v}'$, $\mathbf{v}'\times\mathbf{c}$ 等于将 $\mathbf{v}'$ 在 π 内旋转 90° 后乘以 $|\mathbf{c}|$. 旋转是线性变换, 故投影的加法保持, 分配律成立. □

右分配律由左分配律与反交换律推出:

$$\mathbf{a}\times(\mathbf{b}+\mathbf{c})=-(\mathbf{b}+\mathbf{c})\times\mathbf{a}=-\mathbf{b}\times\mathbf{a}-\mathbf{c}\times\mathbf{a}=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}$$

11.3.3 外积的坐标公式

取空间直角坐标系的基向量 $\mathbf{i}$、$\mathbf{j}$、$\mathbf{k}$(构成右手系). 由定义直接得到

$$\mathbf{i}\times\mathbf{i}=\mathbf{j}\times\mathbf{j}=\mathbf{k}\times\mathbf{k}=\mathbf{0},\quad \mathbf{i}\times\mathbf{j}=\mathbf{k},\quad \mathbf{j}\times\mathbf{k}=\mathbf{i},\quad \mathbf{k}\times\mathbf{i}=\mathbf{j}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设 $\mathbf{a}=(x_1,y_1,z_1)$, $\mathbf{b}=(x_2,y_2,z_2)$, 按分配律展开:

$$\mathbf{a}\times\mathbf{b}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\x_1&y_1&z_1\\x_2&y_2&z_2\end{vmatrix}=(y_1z_2-y_2z_1)\mathbf{i}+(z_1x_2-z_2x_1)\mathbf{j}+(x_1y_2-x_2y_1)\mathbf{k}\tag{11-7}$$

定理 11.11 (拉格朗日恒等式) 对任意两个向量 $\mathbf{a}$、$\mathbf{b}$,

$$|\mathbf{a}\times\mathbf{b}|^2=|\mathbf{a}|^2|\mathbf{b}|^2-(\mathbf{a}\cdot\mathbf{b})^2\tag{11-8}$$

证明 由 $|\mathbf{a}\times\mathbf{b}|=|\mathbf{a}||\mathbf{b}|\sin\theta$ 与 $\mathbf{a}\cdot\mathbf{b}=|\mathbf{a}||\mathbf{b}|\cos\theta$,

$$|\mathbf{a}\times\mathbf{b}|^2=|\mathbf{a}|^2|\mathbf{b}|^2\sin^2\theta=|\mathbf{a}|^2|\mathbf{b}|^2(1-\cos^2\theta)=|\mathbf{a}|^2|\mathbf{b}|^2-(\mathbf{a}\cdot\mathbf{b})^2$$

□

例题 11.26 (三角形面积) 已知 $\triangle ABC$ 的顶点 $A(1,0,0)$、$B(2,1,0)$、$C(0,1,1)$. 用外积求 $\triangle ABC$ 的面积.

解 $\overrightarrow{AB}=(1,1,0)$, $\overrightarrow{AC}=(-1,1,1)$. 外积

$$\overrightarrow{AB}\times\overrightarrow{AC}=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\1&1&0\\-1&1&1\end{vmatrix}=(1,-1,2)$$

模长 $|\overrightarrow{AB}\times\overrightarrow{AC}|=\sqrt{1+1+4}=\sqrt{6}$. 三角形面积等于平行四边形面积之半:

$$S_{\triangle ABC}=\frac{1}{2}|\overrightarrow{AB}\times\overrightarrow{AC}|=\frac{\sqrt{6}}{2}$$

11.3.4 二重外积与雅可比恒等式

外积的结果仍是向量, 可继续参与外积运算.$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ 垂直于 $\mathbf{a}\times\mathbf{b}$, 而 $\mathbf{a}\times\mathbf{b}$ 垂直于 $\mathbf{a}$、$\mathbf{b}$ 所在平面, 故 $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ 落回该平面, 可被 $\mathbf{a}$、$\mathbf{b}$ 线性表示.

定理 11.12 (二重外积公式) 对任意三个向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$,

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}\tag{11-9}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 取坐标系使 $\mathbf{a}=(a,0,0)$, $\mathbf{b}=(b_1,b_2,0)$, $\mathbf{c}=(c_1,c_2,c_3)$. 则 $\mathbf{a}\times\mathbf{b}=(0,0,ab_2)$, 故

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=(-ab_2c_2,\ ab_2c_1,\ 0)$$

另一方面,

$$(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}=ac_1(b_1,b_2,0)-(b_1c_1+b_2c_2)(a,0,0)=(-ab_2c_2,\ ab_2c_1,\ 0)$$

两端相等. □

知乎@余命数

定理 11.13 (雅可比恒等式) 对任意三个向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$,

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}+(\mathbf{c}\times\mathbf{a})\times\mathbf{b}=\mathbf{0}\tag{11-10}$$

证明 对三项分别套用 (11-9):

$$(\mathbf{a}\times\mathbf{b})\times\mathbf{c}=(\mathbf{a}\cdot\mathbf{c})\mathbf{b}-(\mathbf{b}\cdot\mathbf{c})\mathbf{a}$$

$$(\mathbf{b}\times\mathbf{c})\times\mathbf{a}=(\mathbf{b}\cdot\mathbf{a})\mathbf{c}-(\mathbf{c}\cdot\mathbf{a})\mathbf{b}$$

$$(\mathbf{c}\times\mathbf{a})\times\mathbf{b}=(\mathbf{c}\cdot\mathbf{b})\mathbf{a}-(\mathbf{a}\cdot\mathbf{b})\mathbf{c}$$

三式相加, 含 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 的项各成对抵消, 得 $\mathbf{0}$. □

二重外积公式表明外积不满足结合律, $(\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ 与 $\mathbf{a}$、$\mathbf{b}$ 共面, $\mathbf{a}\times(\mathbf{b}\times\mathbf{c})$ 与 $\mathbf{b}$、$\mathbf{c}$ 共面, 两者一般不同.

11.3.5 混合积

知乎@余命数

定义 11.8 (混合积) 三个向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 的混合积定义为

$$[\mathbf{a},\mathbf{b},\mathbf{c}]=(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}\tag{11-11}$$

定理 11.14 (混合积的几何意义) $|[\mathbf{a},\mathbf{b},\mathbf{c}]|$ 等于以 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 为邻边的平行六面体的体积. 当 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 构成右手系时混合积为正, 左手系时为负.

证明 $\mathbf{a}\times\mathbf{b}$ 垂直于 $\mathbf{a}$、$\mathbf{b}$ 所在平面, $|\mathbf{a}\times\mathbf{b}|$ 为底面平行四边形面积.$\mathbf{c}$ 在 $\mathbf{a}\times\mathbf{b}$ 方向上的投影 $h=|\mathbf{c}|\cos\varphi$($\varphi$ 为 $\mathbf{c}$ 与 $\mathbf{a}\times\mathbf{b}$ 的夹角)为平行六面体的高. 故

$$(\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}=|\mathbf{a}\times\mathbf{b}|\,|\mathbf{c}|\cos\varphi=\text{底面积}\times\text{高}=\text{体积}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

右手系下 $\cos\varphi>0$, 混合积为正; 左手系下 $\cos\varphi<0$, 混合积为负. □

11.3.6 混合积的坐标公式与轮换对称性

设 $\mathbf{a}=(x_1,y_1,z_1)$, $\mathbf{b}=(x_2,y_2,z_2)$, $\mathbf{c}=(x_3,y_3,z_3)$. 由 (11-7) 算出 $\mathbf{a}\times\mathbf{b}$ 再与 $\mathbf{c}$ 作内积, 得

$$[\mathbf{a},\mathbf{b},\mathbf{c}]=\begin{vmatrix} x_1 & y_1 & z_1\\ x_2 & y_2 & z_2\\ x_3 & y_3 & z_3\end{vmatrix} \tag{11-12}$$

知乎@余命数

定理 11.15 (混合积的轮换对称性) 混合积在轮换下不变, 对换下变号:

$$[\mathbf{a},\mathbf{b},\mathbf{c}]=[\mathbf{b},\mathbf{c},\mathbf{a}]=[\mathbf{c},\mathbf{a},\mathbf{b}]=-[\mathbf{b},\mathbf{a},\mathbf{c}]$$

证明 轮换是行列式 (11-12) 的行轮换, 值不变; 对换是行列式的行对换, 值变号. □

定理 11.16 (共面判定) 三个向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$ 共面的充分必要条件是

$$[\mathbf{a},\mathbf{b},\mathbf{c}]=0 \tag{11-13}$$

证明 共面时平行六面体退化为平面图形, 体积为零, 故 $[\mathbf{a},\mathbf{b},\mathbf{c}]=0$. 反之, $[\mathbf{a},\mathbf{b},\mathbf{c}]=0$ 意味着 $\mathbf{c}$ 垂直于 $\mathbf{a}\times\mathbf{b}$; 若 $\mathbf{a}\times\mathbf{b}\neq\mathbf{0}$, $\mathbf{c}$ 落在 $\mathbf{a}$、$\mathbf{b}$ 所在平面内; 若 $\mathbf{a}\times\mathbf{b}=\mathbf{0}$, $\mathbf{a}$ 与 $\mathbf{b}$ 共线, 三向量自然共面. □

例题 11.27 (四面体体积) 四面体 $OABC$ 的顶点为 $O(0,0,0)$、$A(2,0,0)$、$B(1,2,0)$、$C(0,1,2)$. 求其体积.

解 $\overrightarrow{OA}=(2,0,0)$, $\overrightarrow{OB}=(1,2,0)$, $\overrightarrow{OC}=(0,1,2)$. 混合积

$$[\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}]=\begin{vmatrix} 2 & 0 & 0\\ 1 & 2 & 0\\ 0 & 1 & 2\end{vmatrix}=2\times(2\times2-0\times1)=8$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

四面体体积为平行六面体体积的 $\frac{1}{6}$（底面三角形面积为平行四边形之半，锥体体积再取 $\frac{1}{3}$），故

$$V=\frac{1}{6}|[\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}]|=\frac{8}{6}=\frac{4}{3}$$

例题 11.28 (四点共面判定) 判断 $A(1,0,1)$、$B(2,1,0)$、$C(0,1,1)$、$D(1,1,2)$ 是否共面.

解 取 $\overrightarrow{AB}=(1,1,-1)$，$\overrightarrow{AC}=(-1,1,0)$，$\overrightarrow{AD}=(0,1,1)$. 混合积

$$[\overrightarrow{AB},\overrightarrow{AC},\overrightarrow{AD}]=\begin{vmatrix}1&1&-1\\-1&1&0\\0&1&1\end{vmatrix}=1\times(1-0)-1\times(-1-0)+(-1)\times(-1-0)=1+1+1=3\neq 0$$

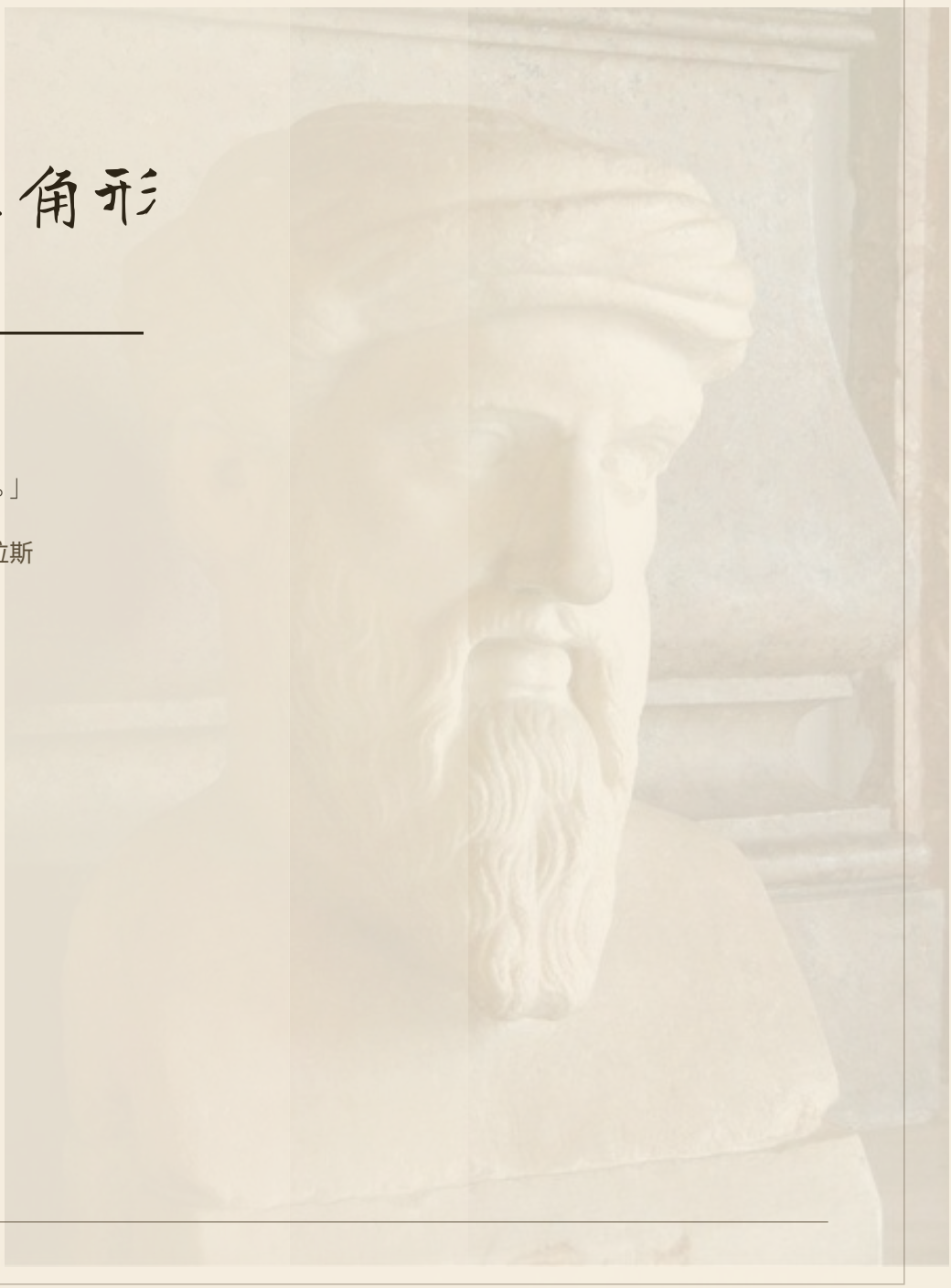
混合积非零，四点不共面.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

知乎@余命数

第 12 章

解三角形



几何主干

「万物皆数。」

——毕达哥拉斯

第十二章 解三角形

注 12.1
本章若干证明(尤其是四心的向量表示)使用 第11章 的数量积与定比分点. 边角关系由正弦定理、余弦定理与面积公式处理;建系法给出坐标保底路线.

大问题

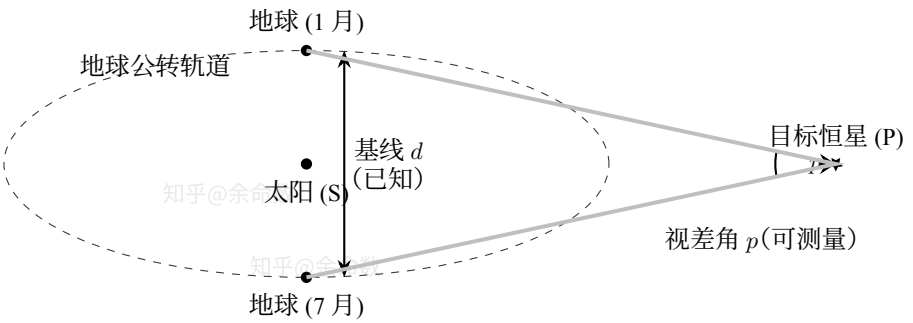


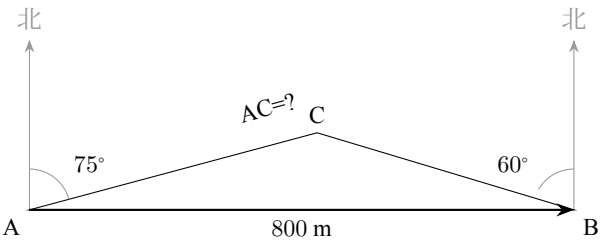
图 12.1 用三角形丈量宇宙

地球到恒星的距离可以用三角测量求出. 我们永远无法用一把尺子去测量地球到恒星的距离. 然而, 天文学家利用地球的公转, 构建了一个横跨数亿公里的巨型三角尺. 他们在 1 月和 7 月, 即地球位于轨道两侧时, 分别观测同一颗邻近恒星 P. 由于观测点不同, P 在遥远的背景星空中的位置会产生微小的位移, 这个位移的角度 p(视差角)是可以测量的. 在由 1 月地球、7 月地球与目标恒星 P 构成的这个非常狭长的三角形中, 我们已知底边长度, 即地球轨道直径 d, 以及可由视差角 p 推出的两个底角. **仅凭这些数据就能算出那条遥不可及的腰长, 即地星距离.** 这需要一套普适的、能够连接任意三角形边与角的完备的法则.

探索课题

在一次定向越野中, 你从点 A 出发, 向正东方向走了 800 米到达点 B. 此时你发现目标 C 在你的北偏西 60° 方向. 你的队友留在 A 点, 他测得 C 点位于 A 点的北偏东 75° 方向.

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



请问目标点 C 距离出发点 A 有多远?(即求 AC 的长度)

在 $\triangle ABC$ 中, 我们已知一条边 $AB = 800\text{m}$. 角 $\angle CAB = 90^\circ - 75^\circ = 15^\circ$. 在 B 点, 正东方向为射线 AB, 正北方向与其垂直, 正西方向为射线 BA. 北偏西 60° 是指从正北方向向西转 60°. 因此, 射线 BC 与正西方向 (射线 BA) 的夹角为 $90^\circ - 60^\circ = 30^\circ$. 所以 $\angle ABC = 30^\circ$. 于是, $\angle C = 180^\circ - 15^\circ - 30^\circ = 135^\circ$.

此时已拥有三角形的一边和所有角 (AAS), 但它不是直角三角形, 勾股定理和初中锐角三角函数定义都无法直接使用. 虽然可通过作高 $CD \perp AB$, 将问题转化为解两个直角三角形, 但这需要解一个关于边长的复杂方程, 过程相当繁琐. 于是需要一种更普适、更直接的工具, 能够建立任意三角形中边与其对角之间的关系.

本章围绕**正弦定理**与**余弦定理**展开. 已知三角形的部分边或角, 我们要解出其余元素. 正弦定理处理边角相对的关系, 余弦定理处理两边一夹角或三边的情形; 再配合面积公式 $S = \frac{1}{2}ab\sin C$、$S = \frac{abc}{4R}$ 与 $S = pr$, 边、角、面积、外接圆半径与内切圆半径就连成一个整体. 于是三角函数的应用从直角三角形推广到任意三角形, 几何问题被翻译成代数运算. 范围、最值、四心、四线, 都可以用这两个定理处理.

12.1 正弦定理和余弦定理

12.1.1 正弦定理

12.1.1.1 来龙去脉

在直角三角形中, 边角关系由 $\sin A = \frac{\text{对边}}{\text{斜边}}$ 给出. 已知一个锐角和一条边, 就可以求出其余边.

我们也想在任意三角形中把边和它的对角联系起来. 正弦定理指出, 在任意三角形中, 对边与其对角正弦值的比值是常数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 12.1 (正弦定理) 在一个非退化三角形中, 各边和它所对角的正弦值的比相等, 且等于其外接圆的直径. 即

$$\frac{a}{\sin A}=\frac{b}{\sin B}=\frac{c}{\sin C}=2R \tag{12-1}$$

其中 R 为该三角形外接圆的半径.

12.1.1.2 定理的证明

正弦定理的几何证明把任意三角形转化为已熟悉的直角三角形.

证明 (几何法证明) 把任意三角形化归为直角三角形, 是证明的入口. 在任意非退化 $\triangle ABC$ 中, 从顶点 C 向对边 AB 作高 h_c. 在 Rt$\triangle ADC$ 中, 有 $h_c=b\sin A$. 在

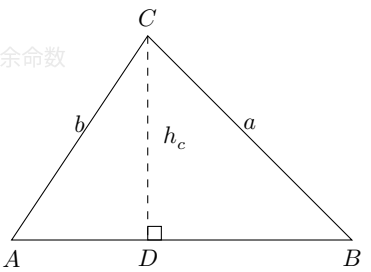
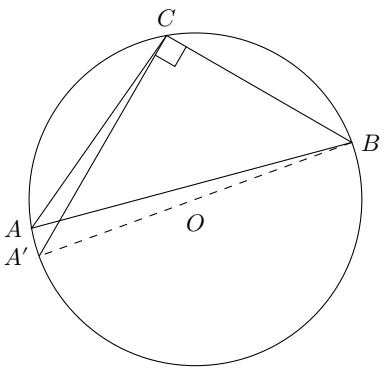


图 12.2 作高法证明

Rt$\triangle BDC$ 中, 有 $h_c=a\sin B$. 因此, $b\sin A=a\sin B$. 由于非退化三角形的内角均在 $(0,\pi)$ 内, 所以 $\sin A>0,\sin B>0$. 两边同除以 $\sin A\sin B$, 得 $\frac{b}{\sin B}=\frac{a}{\sin A}$. 同理, 作另一条高, 可以证明 $\frac{b}{\sin B}=\frac{c}{\sin C}$. 故 $\frac{a}{\sin A}=\frac{b}{\sin B}=\frac{c}{\sin C}$.

尚需证明这个比值等于 $2R$. 为此, 构造一个以直径为斜边的直角三角形. 作 $\triangle ABC$ 的外接圆, 圆心为 O. 连接 BO 并延长, 交圆于点 A'. 连接 $A'C$.



在 $\triangle A'BC$ 中, 因为 BA' 是直径, 所以 $\angle A'CB=90^\circ$. 又因为 A 和 A' 对应同一条弦 BC, 它们所成圆周角相等或互补, 因而 $\sin A=\sin A'$. 在 Rt$\triangle A'BC$ 中,

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$\sin A' = \frac{\text{对边}}{\text{斜边}} = \frac{BC}{BA'} = \frac{a}{2R}$. 所以 $\sin A = \frac{a}{2R}$, 即 $\frac{a}{\sin A} = 2R$. 与前面的比值等式合起来, 正弦定理得证. □

12.1.1.3 正弦定理的应用

正弦定理的常用变形

1. **边化角**: $a = 2R\sin A, b = 2R\sin B, c = 2R\sin C$.(用于消去边, 得到纯角的三角函数关系)
2. **角化边**: $\sin A = \frac{a}{2R}, \sin B = \frac{b}{2R}, \sin C = \frac{c}{2R}$.(用于消去角, 得到纯边的代数关系)
3. **边角比例**: $a : b : c = \sin A : \sin B : \sin C$.(用于边角关系的整体代换)

适用场景

正弦定理联系的是一组对边与对角. 它主要解决以下两类问题.

1. 已知两角和任意一边, 求其他两边和另一角.(AAS 或 ASA)
2. 已知两边和其中一边的对角, 求另一边的对角.(SSA) 此种情形可能出现一解、两解或无解, 需结合边长与角的大小逐一讨论.

12.1.2 余弦定理

12.1.2.1 勾股定理的推广

勾股定理 $a^2 + b^2 = c^2$ 只在直角三角形中成立. 当 $\angle C \neq 90^\circ$ 时, $a^2 + b^2$ 与 c^2 不再相等. 余弦定理给出这个差值的表达式 $c^2 = a^2 + b^2 - 2ab\cos C$.

- 当 $\angle C$ 是锐角时, $\cos C > 0$, 故 $c^2 < a^2 + b^2$.
- 当 $\angle C$ 是钝角时, $\cos C < 0$, 故 $c^2 > a^2 + b^2$.
- 当 $\angle C = 90^\circ$ 时, $\cos C = 0$, 余弦定理化为 $c^2 = a^2 + b^2$, 即勾股定理.

余弦定理可视为勾股定理在任意三角形中的推广.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 12.2 (余弦定理) 非退化三角形任何一边的平方, 等于其他两边平方的和减去这两边与它们夹角的积的两倍. 即

$$a^2 = b^2 + c^2 - 2bc\cos A \tag{12-2}$$

$$b^2 = a^2 + c^2 - 2ac\cos B \tag{12-3}$$

$$c^2 = a^2 + b^2 - 2ab\cos C \tag{12-4}$$

其推论形式(用于求角)如下.

$$\cos A = \frac{b^2 + c^2 - a^2}{2bc} \tag{12-5}$$

12.1.2.2 两种代数证明

余弦定理可以用坐标法和向量法两种代数方式证明.

证明 (坐标法证明) 将 $\triangle ABC$ 放置在平面直角坐标系中, 使顶点 C 与原点重合, 边 a, 即 CB 落在 x 轴正半轴上. 由三角函数定义, 各顶点坐标为 $C(0,0)$、$B(a,0)$、

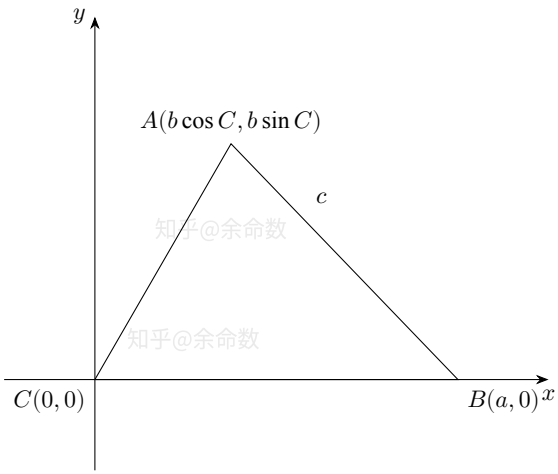


图 12.3 坐标法证明

$A(b\cos C, b\sin C)$. 用两点间距离公式计算边 c, 即线段 AB 的平方, 得

$$\begin{aligned}
c^2 = |AB|^2 &= (b\cos C - a)^2 + (b\sin C - 0)^2 && \text{(距离公式)}\\
&= (b^2\cos^2 C - 2ab\cos C + a^2) + b^2\sin^2 C && \text{(展开)}\\
&= a^2 + b^2(\cos^2 C + \sin^2 C) - 2ab\cos C && \text{(合并同类项)}\\
&= a^2 + b^2 - 2ab\cos C. && (\sin^2 C + \cos^2 C = 1)
\end{aligned}$$

□

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 在 $\triangle ABC$ 中, 向量关系 $\overrightarrow{AB}=\overrightarrow{CB}-\overrightarrow{CA}$ 两边取数量积, 得

$$|\overrightarrow{AB}|^2=(\overrightarrow{CB}-\overrightarrow{CA})\cdot(\overrightarrow{CB}-\overrightarrow{CA}) \qquad \text{(数量积展开)}$$

$$c^2=|\overrightarrow{CB}|^2-2\overrightarrow{CB}\cdot\overrightarrow{CA}+|\overrightarrow{CA}|^2 \qquad \text{(差平方公式)}$$

由数量积定义, $\overrightarrow{CB}\cdot\overrightarrow{CA}=|\overrightarrow{CB}||\overrightarrow{CA}|\cos C=ab\cos C$. 代入得 $c^2=a^2-2ab\cos C+b^2$. □

12.1.2.3 余弦定理的应用

适用场景

余弦定理联系的是三边与一角. 它主要解决以下两类问题.

1. 已知三边, 求三个角.(SSS)
2. 已知两边和它们的夹角, 求第三边和其他两个角.(SAS)

同时, 余弦定理的推论可用于判断三角形形状.$b^2+c^2-a^2>0$ 时 A 为锐角; 等于 0 时为直角; 小于 0 时为钝角.

12.1.3 三角形面积公式

中学阶段常用的三角形面积公式有五个.

- **底高公式**: $S=\frac{1}{2}ah_a$, 是面积定义本身.
- **三角公式**: $S=\frac{1}{2}ab\sin C$, 用角代替高, 是解三角形中最常用的面积公式.
- **外接圆公式**: $S=\frac{abc}{4R}$, 由正弦定理把面积与外接圆半径联系起来.
- **内切圆公式**: $S=pr$, 由内心分割三角形, 把面积与内切圆半径联系起来.
- **海伦公式**: $S=\sqrt{p(p-a)(p-b)(p-c)}$, 面积完全由三边决定, 适用于已知三边 (SSS) 的情形.

12.1.3.1 $S=\frac{1}{2}ab\sin C$

这是解三角形中最常用的面积公式. 它用已知角来计算通常不易直接得到的高.

证明 从最原始的底高公式 $S=\frac{1}{2}b\cdot h_b$ 出发. 这里的 h_b 是边 b 上的高. 如图, 从顶点 B 向边 AC 作高 BD, 其长度为 h_b. 在 Rt$\triangle BDC$ 中, 由正弦函数的定义有 $\sin C=\frac{BD}{BC}=\frac{h_b}{a}$, 因此 $h_b=a\sin C$. 把它代入以 b 为底的面积公式, 得 $S=\frac{1}{2}b\cdot h_b=\frac{1}{2}b(a\sin C)=\frac{1}{2}ab\sin C$. 同理 $S=\frac{1}{2}bc\sin A=\frac{1}{2}ac\sin B$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

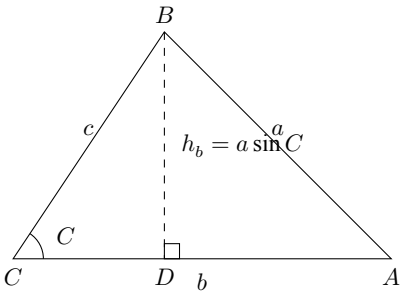


图 12.4 作高法推导三角公式

结构与应用

此公式的结构特征是**两边及夹角** (SAS). 已知或可求出两边及其夹角时, 用它求面积最直接. 反过来, 已知面积和两边, 也可用它求夹角的正弦值.

12.1.3.2 外接圆公式

面积与外接圆半径之间可由三角公式和正弦定理直接建立联系.

证明 由三角公式 $S = \frac{1}{2}ab\sin C$ 和正弦定理 $\frac{c}{\sin C} = 2R$, 得 $\sin C = \frac{c}{2R}$. 代入后 $S = \frac{1}{2}ab\left(\frac{c}{2R}\right) = \frac{abc}{4R}$. □

结构与应用

此公式联系的是三边与外接圆半径. 它常用于以下情形.

1. 已知三边和面积, 求外接圆半径 R.
2. 已知三边和外接圆半径 R, 求面积 S.
3. 把面积 S 与三边之积 abc 相互转化.

12.1.3.3 内切圆公式

内切圆公式把面积与内切圆半径联系起来, 证明方法是把三角形按内心分割成三个小三角形.

证明 设 $\triangle ABC$ 的内心为 I, 内切圆半径为 r. 连接 IA、IB、IC, 则 $\triangle ABC$ 被分成 $\triangle IBC$, $\triangle ICA$ 和 $\triangle IAB$ 三个小三角形. 由内心性质, 内心 I 到三边的距离都等于内切圆半径 r, 这也就是三个小三角形的高. 因此 $S_{\triangle ABC} = S_{\triangle IBC} + S_{\triangle ICA} + S_{\triangle IAB} = \frac{1}{2}ar + \frac{1}{2}br + \frac{1}{2}cr = \frac{1}{2}(a+b+c)r$. 记半周长 $p = \frac{a+b+c}{2}$, 则 $S = pr$. □

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

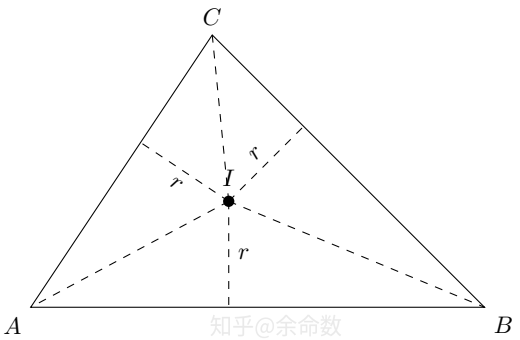


图 12.5　面积分割法推导内切圆公式

结构与应用

该公式联系面积、周长与内切圆半径. 求内切圆半径时, 先用海伦公式求出面积 S, 再用 $S=pr$ 求 r.

12.1.3.4 海伦公式

海伦公式不用角度, 只凭三边长度即可求面积. 其证明综合了三角恒等变换与代数变形.

证明 利用余弦定理把与角相关的面积公式转化为只与边长相关的表达式.

从三角形面积公式 $S=\frac{1}{2}ab\sin C$ 出发, 两边平方, 并利用 $\sin^2 C=1-\cos^2 C$ 和余弦定理 $\cos C=\frac{a^2+b^2-c^2}{2ab}$, 可得

$$\begin{aligned}
S^2&=\left(\frac{1}{2}ab\sin C\right)^2=\frac{1}{4}a^2b^2\sin^2 C\\
&=\frac{1}{4}a^2b^2(1-\cos^2 C)\\
&=\frac{1}{4}a^2b^2(1+\cos C)(1-\cos C)\\
&=\frac{1}{4}a^2b^2\left(1+\frac{a^2+b^2-c^2}{2ab}\right)\left(1-\frac{a^2+b^2-c^2}{2ab}\right)\\
&=\frac{a^2b^2}{4}\left(\frac{2ab+a^2+b^2-c^2}{2ab}\right)\left(\frac{2ab-(a^2+b^2-c^2)}{2ab}\right)\\
&=\frac{1}{16}\left((a+b)^2-c^2\right)\left(c^2-(a-b)^2\right)\\
&=\frac{1}{16}(a+b+c)(a+b-c)(c-a+b)(c+a-b)
\end{aligned}$$

令半周长 $p=\frac{a+b+c}{2}$, 则 $a+b-c=2(p-c)$、$a+c-b=2(p-b)$、$b+c-a=2(p-a)$. 代入上式, 得

$$S^2=\frac{1}{16}\cdot(2p)\cdot 2(p-c)\cdot 2(p-b)\cdot 2(p-a)=p(p-a)(p-b)(p-c)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两边开方, 得到

$$S=\sqrt{p(p-a)(p-b)(p-c)}.$$

□

结构与应用

该公式只与三边长度有关. 已知三边 (SSS) 求面积时，用海伦公式不必先求角.

12.1.4 三角形的四心

正弦定理与余弦定理既用于边角的计算, 也联系着三角形的代数信息与几何形态. 三角形四心, 即**外心**、**内心**、**重心**、**垂心**, 分别是中垂线、角平分线、中线、高线这四组特殊直线的交点, 本节给出它们的定义、性质与向量公式.

12.1.4.1 外心

给定不在同一直线上的三点 A、B、C, 到它们距离相等的点 P 就是外接圆的圆心. 这个圆称为三角形的**外接圆**, 圆心称为**外心**.

定义 12.1 (外心)

- **几何定义**三角形三条边的**垂直平分线**（中垂线）的交点.
- **性质定义**到三角形三**个顶点距离相等**的点. 这个距离就是外接圆的半径 R.

性质 (外心的位置特征) 外心的位置由三角形的形状决定.

- **锐角三角形**外心在三角形**内部**.
- **直角三角形**外心在**斜边的中点**.
- **钝角三角形**外心在三角形**外部**.

证明 由后文外心的向量性质可得外心 O 的重心坐标为

$$O=\frac{(\sin 2A)A+(\sin 2B)B+(\sin 2C)C}{\sin 2A+\sin 2B+\sin 2C}.$$

又

$$\sin 2A+\sin 2B+\sin 2C=4\sin A\sin B\sin C>0.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

一个点的重心坐标全为正时在三角形内部; 有一个坐标为 0 且其余为正时在对应边上; 有一个坐标为负时在三角形外部.

若三角形为锐角三角形, 则 A, B, C 均小于 90°, 故 $\sin 2A, \sin 2B, \sin 2C$ 全为正, 外心在内部. 若三角形为直角三角形, 例如 $C = 90^\circ$, 则 $\sin 2C = 0$, 且 $A + B = 90^\circ$ 给出 $\sin 2A = \sin 2B$, 所以外心为 AB 的中点, 即斜边中点. 若三角形为钝角三角形, 则钝角对应的 $\sin 2A$、$\sin 2B$ 或 $\sin 2C$ 为负, 外心在三角形外部. □

引理 12.1 (外心与中心角) 若 O 为 $\triangle ABC$ 的外心, 则对应中心角与圆周角满足

$$\cos\angle BOC = \cos 2A, \quad \cos\angle COA = \cos 2B, \quad \cos\angle AOB = \cos 2C.$$

特别地, 在锐角三角形中, 有 $\angle BOC = 2A$、$\angle COA = 2B$、$\angle AOB = 2C$.

证明 以 BC 所在弧为例. 圆心角 $\angle BOC$ 与圆周角 $\angle BAC$ 截同一条弧 BC. 由圆周角定理, 同弧所对圆心角等于圆周角的两倍, 因而在锐角情形下

$$\angle BOC = 2A.$$

若 A 为钝角, 则向量 $\overrightarrow{OB}$ 与 $\overrightarrow{OC}$ 之间的夹角取较小中心角, 为 $2\pi - 2A$, 其余情形同理. 因为

$$\cos(2\pi - 2A) = \cos 2A,$$

所以总有 $\cos\angle BOC = \cos 2A$. 对另外两对顶点同理可得其余两个等式. □

阅读提示

使用外心向量时, 真正需要的常常是点积中的余弦关系. 对于钝角三角形, 例如若 A 为钝角, 则较小的 $\angle BOC = 2(\pi - A)$, 但其余弦仍等于 $\cos 2A$.

定理 12.3 (外心与正弦定理的绑定关系) 外心与正弦定理的联系是它的主要性质.

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C} = 2R$$

证明 见正弦定理的证明. □

知乎@余命数

推论 12.1 (外心的向量性质) 若 O 为 $\triangle ABC$ 的外心, 则

$$(\sin 2A)\overrightarrow{OA}+(\sin 2B)\overrightarrow{OB}+(\sin 2C)\overrightarrow{OC}=\vec{0}$$

证明 设 $\vec{S}=(\sin 2A)\overrightarrow{OA}+(\sin 2B)\overrightarrow{OB}+(\sin 2C)\overrightarrow{OC}$. 证明 $\vec{S}=\vec{0}$ 只需证明它同时与两个不共线向量垂直, 这里取 $\overrightarrow{OA}$ 和 $\overrightarrow{OB}$.

$\vec{S}$ 与 $\overrightarrow{OA}$ 的点积为

$$\vec{S}\cdot\overrightarrow{OA}=(\sin 2A)(\overrightarrow{OA}\cdot\overrightarrow{OA})+(\sin 2B)(\overrightarrow{OB}\cdot\overrightarrow{OA})+(\sin 2C)(\overrightarrow{OC}\cdot\overrightarrow{OA}).$$

O 是外心, 设外接圆半径为 R, 则 $|\overrightarrow{OA}|=|\overrightarrow{OB}|=|\overrightarrow{OC}|=R$. 由点积定义与引理 12.1, 我们写出.

- $\overrightarrow{OA}\cdot\overrightarrow{OA}=R^2$;
- $\overrightarrow{OB}\cdot\overrightarrow{OA}=R^2\cos(2C)$;
- $\overrightarrow{OC}\cdot\overrightarrow{OA}=R^2\cos(2B)$.

即使钝角三角形, 例如 C 为钝角, 仍有 $\cos(\angle AOB)=\cos(2(\pi-C))=\cos(2C)$, 上式同样成立.

代入得

$$\vec{S}\cdot\overrightarrow{OA}=R^2[\sin(2A)+\sin(2B)\cos(2C)+\sin(2C)\cos(2B)]$$

由和角公式, $\sin(2B)\cos(2C)+\sin(2C)\cos(2B)=\sin(2B+2C)$, 故

$$\vec{S}\cdot\overrightarrow{OA}=R^2[\sin(2A)+\sin(2B+2C)]$$

又 $A+B+C=\pi$, 所以 $2B+2C=2\pi-2A$, 于是

$$\sin(2B+2C)=\sin(2\pi-2A)=-\sin(2A)$$

因此

$$\vec{S}\cdot\overrightarrow{OA}=R^2[\sin(2A)-\sin(2A)]=0.$$

同理可得 $\vec{S}\cdot\overrightarrow{OB}=0$.

若 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 不共线, 则同时与二者垂直的平面向量只能是零向量. 若 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 共线, 则 AB 为外接圆直径, $C=90^\circ$, 此时同样的对称计算给出 $\vec{S}\cdot\overrightarrow{OC}=0$; 而 $\overrightarrow{OA}$ 与 $\overrightarrow{OC}$ 不共线, 仍得到 $\vec{S}=\vec{0}$.

因此

$$(\sin 2A)\overrightarrow{OA}+(\sin 2B)\overrightarrow{OB}+(\sin 2C)\overrightarrow{OC}=\vec{0}.$$

□

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

12.1.4.2 内心

来龙去脉

在三角形内部到三边距离相等的点是内切圆的圆心, 称为**内心**.

定义 12.2 (内心)
- **几何定义**三角形三条**内角平分线**的交点.
- **性质定义**到三角形三**边距离相等**的点. 这个距离就是内切圆的半径 r.

性质 (内心的位置特征) 内心永远在三角形**内部**.

证明 此结论可通过内心的重心坐标定义直接证明. 一个点 P 位于 $\triangle ABC$ 内部的充分必要条件是, 存在三个正实数 $\alpha,\beta,\gamma>0$ 且 $\alpha+\beta+\gamma=1$, 使得对于任意参考点 O, 都有 $\overrightarrow{OP}=\alpha\overrightarrow{OA}+\beta\overrightarrow{OB}+\gamma\overrightarrow{OC}$.

内心的向量公式为 $\overrightarrow{OI}=\frac{a\overrightarrow{OA}+b\overrightarrow{OB}+c\overrightarrow{OC}}{a+b+c}$. 它可以重写为

$$\overrightarrow{OI}=\left(\frac{a}{a+b+c}\right)\overrightarrow{OA}+\left(\frac{b}{a+b+c}\right)\overrightarrow{OB}+\left(\frac{c}{a+b+c}\right)\overrightarrow{OC}$$

令 $\alpha=\frac{a}{a+b+c}$, $\beta=\frac{b}{a+b+c}$ 及 $\gamma=\frac{c}{a+b+c}$. 由于 a,b,c 是三角形的边长, 它们必然是正数, 即 $a>0,b>0,c>0$. 因此, 分母 $a+b+c$ 也是正数, 这确保了系数 α,β,γ 均为正数.

检验系数之和.

$$\alpha+\beta+\gamma=\frac{a}{a+b+c}+\frac{b}{a+b+c}+\frac{c}{a+b+c}=\frac{a+b+c}{a+b+c}=1$$

由于内心 I 的重心坐标系数满足 $\alpha,\beta,\gamma>0$ 且 $\alpha+\beta+\gamma=1$ 的条件, 根据内部点的定义, 内心 I 必定永远位于 $\triangle ABC$ 的内部.

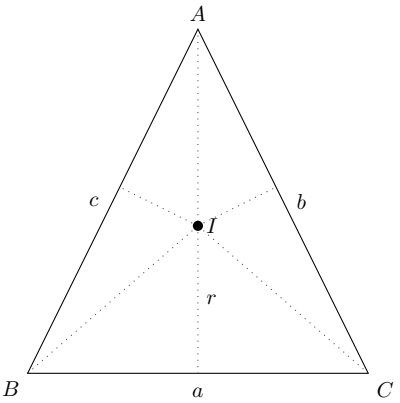
证毕. □

定理 12.4 (内心与面积公式) 内心与面积公式 $S=pr$ 把内切圆半径 r、半周长 p 和面积 S 联系起来.

$$S=\frac{1}{2}(a+b+c)r=pr$$

证明 设 $\triangle ABC$ 的内心为 I, 内切圆半径为 r. 连接 IA,IB,IC, 则 $\triangle ABC$ 被分成三个小三角形 $\triangle BIC$, $\triangle CIA$ 和 $\triangle AIB$.

知乎@余命数



知乎@余命数

于是

$$S = S_{\triangle BIC} + S_{\triangle CIA} + S_{\triangle AIB}$$

内心 I 到三边的距离都等于 r, 这是每个小三角形以原边为底时的高. 因此

知乎@余命数

$$S = \frac{1}{2}ar + \frac{1}{2}br + \frac{1}{2}cr = \frac{1}{2}(a+b+c)r = pr,$$

知乎@余命数

其中 $p = \frac{a+b+c}{2}$. □

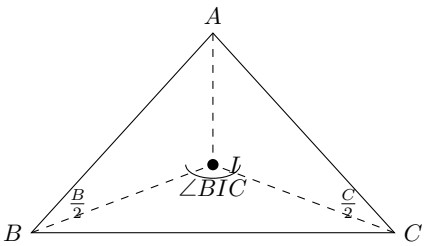
引理 12.2 (内心与顶点角) 以 $\angle BIC$ 为例, 内心与两顶点连线所成角和第三个顶点角之间有固定关系.

$$\angle BIC = 90^\circ + \frac{A}{2}$$

知乎@
知乎@余命数

证明 如图所示, 考虑 $\triangle BIC$. 根据三角形内角和定理, 有.

$$\angle BIC + \angle IBC + \angle ICB = 180^\circ$$



I 是 $\triangle ABC$ 的内心, 所以 BI, CI 分别是 $\angle B, \angle C$ 的平分线.

$$\angle IBC = \frac{1}{2}\angle B, \qquad \angle ICB = \frac{1}{2}\angle C.$$

知乎@余命数

代入 $\triangle BIC$ 的内角和公式, 得

$$\angle BIC = 180^\circ - \frac{1}{2}(\angle B + \angle C)$$

又 $\angle B + \angle C = 180^\circ - \angle A$, 故

$$\angle BIC = 180^\circ - \frac{1}{2}(180^\circ - \angle A) = 90^\circ + \frac{A}{2}.$$

知乎@余命数

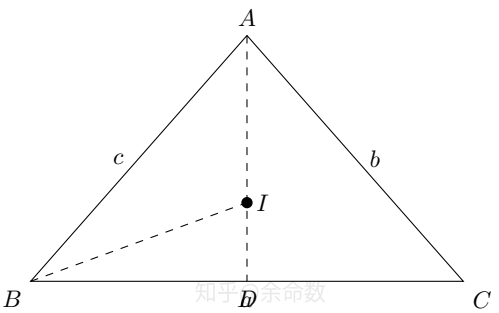
□

推论 12.2 (内心的向量表示) 以任意点 O 为参考点, 内心的向量表达式用对边边长作为权重.

$$\overrightarrow{OI} = \frac{a\overrightarrow{OA} + b\overrightarrow{OB} + c\overrightarrow{OC}}{a+b+c}$$

知乎@余命数

证明 现用角平分线定理和定比分点公式推导.



知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

设 $\angle A$ 的平分线交边 BC 于点 D, 由角平分线定理

$$\frac{BD}{DC} = \frac{AB}{AC} = \frac{c}{b},$$

于是

$$\overrightarrow{OD} = \frac{b\overrightarrow{OB} + c\overrightarrow{OC}}{b+c}. \tag{1}$$

内心 I 在角平分线 AD 上. 在 $\triangle ABD$ 中, BI 也是 $\angle B$ 的平分线, 再用角平分线定理得

$$\frac{AI}{ID} = \frac{AB}{BD} = \frac{c}{\frac{ac}{b+c}} = \frac{b+c}{a}.$$

因此

$$\overrightarrow{OI} = \frac{a\overrightarrow{OA} + (b+c)\overrightarrow{OD}}{a+b+c}.$$

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

把 (1) 代入并约去 $b+c$, 即得

$$\overrightarrow{OI}=\frac{a\overrightarrow{OA}+b\overrightarrow{OB}+c\overrightarrow{OC}}{a+b+c}.$$

□

阅读提示

该公式常用于判定内心的位置或用内心求参数.

知乎@余命数

@余命数

12.1.4.3 重心

来龙去脉

三角形三条**中线**的交点称为**重心**, 它也是均匀薄板的质量中心.

定义 12.3 (重心) 三角形三条**中线**(顶点与对边中点的连线)的交点.

性质 (重心的基本性质)

- 重心总在三角形**内部**.
- 重心把每条中线分成 **2:1** 的两段, 靠近顶点的段长是靠近中点段长的两倍. 设 D 为 BC 中点, 则 $AG:GD=2:1$.
- 三条中线把三角形分成**六个面积相等**的小三角形.

知乎@余命数

证明 设 D,E,F 分别为 BC,CA,AB 的中点. 两条中线 AD 与 BE 交于 G. 由中点向量公式,

$$\overrightarrow{OD}=\frac{\overrightarrow{OB}+\overrightarrow{OC}}{2},\qquad \overrightarrow{OE}=\frac{\overrightarrow{OC}+\overrightarrow{OA}}{2}.$$

解直线 AD 与 BE 的交点可得

$$\overrightarrow{OG}=\frac{\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}}{3}.$$

因此 G 同时也在第三条中线 CF 上, 三条中线共点. 该式又可写成

$$\overrightarrow{OG}=\frac{1}{3}\overrightarrow{OA}+\frac{1}{3}\overrightarrow{OB}+\frac{1}{3}\overrightarrow{OC},$$

三个正系数之和为 1, 所以 G 在三角形内部.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由 D 为 BC 中点,

$$\overrightarrow{OG}=\frac{\overrightarrow{OA}+2\overrightarrow{OD}}{3},$$

说明 G 位于中线 AD 上且 $AG:GD=2:1$. 对另外两条中线同理. 三条中线把原三角形先分成三个等底等高的小三角形, 再由 $2:1$ 分点关系把每个小三角形二等分, 因而得到六个面积相等的小三角形. □

定理 12.5 (重心的向量表示) 以任意点 O 为参考点, 重心 G 的向量表达式是三顶点向量的算术平均.

$$\overrightarrow{OG}=\frac{\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}}{3}$$

证明 设 G 为 $\triangle ABC$ 的重心, D 为边 BC 的中点, O 为参考点.

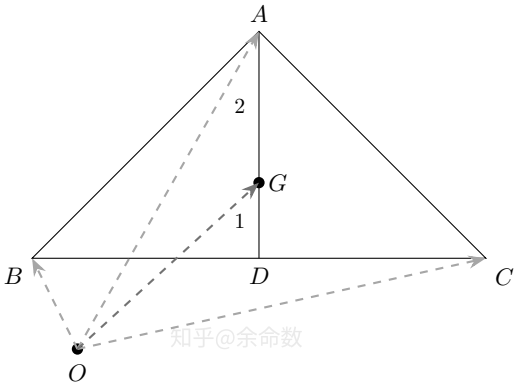


图 12.6　重心 G 在中线 AD 上且 $AG:GD=2:1$

D 是 BC 的中点, 所以

$$\overrightarrow{OD}=\frac{\overrightarrow{OB}+\overrightarrow{OC}}{2}. \tag{1}$$

又 $AG:GD=2:1$, 由定比分点公式

$$\overrightarrow{OG}=\frac{\overrightarrow{OA}+2\overrightarrow{OD}}{3}.$$

把 (1) 代入并化简, 即得

$$\overrightarrow{OG}=\frac{\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}}{3}.$$

□

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

推论 12.3 重心满足

$$\overrightarrow{GA}+\overrightarrow{GB}+\overrightarrow{GC}=\vec{0}$$

证明 由定理 12.5, 有

$$3\overrightarrow{OG}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}$$

由定理 12.5, $3\overrightarrow{OG}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}$. 又

$$\overrightarrow{GA}=\overrightarrow{OA}-\overrightarrow{OG},\quad \overrightarrow{GB}=\overrightarrow{OB}-\overrightarrow{OG},\quad \overrightarrow{GC}=\overrightarrow{OC}-\overrightarrow{OG},$$

三式相加得

$$\overrightarrow{GA}+\overrightarrow{GB}+\overrightarrow{GC}=(\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC})-3\overrightarrow{OG}=\vec{0}.$$

□

阅读提示

上述推论可用于判定重心, 也常用于向量运算.

定理 12.6 (边与中线的关系) 边长的平方和与中线长的平方和满足.

$$a^2+b^2+c^2=\frac{4}{3}(m_a^2+m_b^2+m_c^2)$$

证明 由中线长公式

$$4m_a^2=2b^2+2c^2-a^2,\quad 4m_b^2=2a^2+2c^2-b^2,\quad 4m_c^2=2a^2+2b^2-c^2.$$

三式相加得

$$4(m_a^2+m_b^2+m_c^2)=3(a^2+b^2+c^2),$$

整理即得

$$a^2+b^2+c^2=\frac{4}{3}(m_a^2+m_b^2+m_c^2).$$

□

12.1.4.4 垂心

来龙去脉

三角形的三条高线, 或其延长线交于一点, 这个交点称为**垂心**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 12.4 (垂心) 三角形三条**高线**(或其延长线)的交点.

性质 (垂心的位置特征) 与外心类似, 垂心的位置由三角形形状决定.

- **锐角三角形**垂心在三角形**内部**.
- **直角三角形**垂心就是**直角顶点**.
- **钝角三角形**垂心在三角形**外部**.

证明 高线从顶点出发并垂直于对边所在直线. 若三角形为锐角三角形, 每条高都落在三角形内部, 三条高的交点也在内部. 若三角形为直角三角形, 两条直角边本身就是高, 它们交于直角顶点, 故垂心为直角顶点.

若三角形为钝角三角形, 设 $A > 90^\circ$. 此时从 B、C 引出的两条高分别落在边 AC、AB 的延长线上, 它们的交点位于三角形外部; 第三条高也经过该点. 因此垂心在三角形外部. □

性质 (垂足圆性质) 垂心与顶点、垂足之间常构成共圆关系. 例如, 若 D, E 分别是从 A, B 向 BC, AC 作高得到的垂足, 则 A, B, D, E 四点共圆且直径为 AB, C, D, H, E 四点共圆且直径为 CH.

证明 因为 D 是从 A 向 BC 作高的垂足,所以 $AD \perp BC$. 又 $D \in BC$,故 $AD \perp BD$, 即 $\angle ADB = 90^\circ$. 因为 E 是从 B 向 AC 作高的垂足, 所以 $BE \perp AC$. 又 $E \in AC$, 故 $BE \perp AE$, 即 $\angle AEB = 90^\circ$. 所以 D, E 都在以 AB 为直径的圆上, 即 A, B, D, E 四点共圆.

设 H 为垂心. 由于 H 在高 AD 上, 有 $HD \perp CD$, 故 $\angle CDH = 90^\circ$. 同理 H 在高 BE 上, 有 $HE \perp CE$, 故 $\angle CEH = 90^\circ$. 因此 D, E 都在以 CH 为直径的圆上, 即 C, D, H, E 四点共圆. □

定理 12.7 (垂心与外心的向量关系) 若 O 为外心, H 为垂心, 则以外心 O 为原点时有

$$\overrightarrow{OH} = \overrightarrow{OA} + \overrightarrow{OB} + \overrightarrow{OC}$$

证明 以外心 O 为原点, 记

$$\vec{a} = \overrightarrow{OA}, \quad \vec{b} = \overrightarrow{OB}, \quad \vec{c} = \overrightarrow{OC}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因为 O 是外心, 所以 $|\vec{a}|=|\vec{b}|=|\vec{c}|=R$. 令点 H 满足

$$\overrightarrow{OH}=\vec{a}+\vec{b}+\vec{c}.$$

则

$$\overrightarrow{AH}=\overrightarrow{OH}-\overrightarrow{OA}=\vec{b}+\vec{c}.$$

而

$$\overrightarrow{BC}=\vec{c}-\vec{b}.$$

于是

$$\overrightarrow{AH}\cdot\overrightarrow{BC}=(\vec{b}+\vec{c})\cdot(\vec{c}-\vec{b})=|\vec{c}|^2-|\vec{b}|^2=0.$$

因此 $AH\perp BC$. 同理 $BH\perp AC$, $CH\perp AB$. 所以 H 为垂心, 于是

$$\overrightarrow{OH}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}.$$

□

阅读提示

由这个公式可直接得到 $\overrightarrow{AH}=\overrightarrow{OB}+\overrightarrow{OC}$, 并由此证明 $AH\perp BC$.

定理 12.8 (欧拉线定理) 外心 O、重心 G、垂心 H 三点共线, 这条线称为欧拉线; 且 G 在 O 与 H 之间, 满足

$$\overrightarrow{OH}=3\overrightarrow{OG}\quad(\text{即}OG:GH=1:2)$$

证明 用向量法, 取**外心 O** 作为原点.

以 O 为原点时,

$$\overrightarrow{OG}=\frac{\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}}{3},\qquad \overrightarrow{OH}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}.$$

两式比较得

$$\overrightarrow{OH}=3\overrightarrow{OG}.$$

因此 $\overrightarrow{OH}$ 与 $\overrightarrow{OG}$ 共线, O、G、H 三点在同一直线上. 又因为标量 3>0,G 在射线 OH 上, 且

$$|\overrightarrow{OH}|=3|\overrightarrow{OG}|,\quad |\overrightarrow{GH}|=2|\overrightarrow{OG}|,$$

所以 $OG:GH=1:2$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

阅读提示

欧拉线说明 O、G、H 的位置关系由比值固定. 已知其中两点, 即可确定第三点.

四心向量公式

设 O 为参考点, 则

- 重心 G 满足 $\overrightarrow{OG}=\dfrac{\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}}{3}$;
- 内心 I 满足 $\overrightarrow{OI}=\dfrac{a\overrightarrow{OA}+b\overrightarrow{OB}+c\overrightarrow{OC}}{a+b+c}$;
- 垂心 H 满足 $\overrightarrow{OH}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}$(O 为外心).

这些公式常分别与外接圆半径、内切圆半径、中线、高线等各组条件配合使用.

12.1.5 圆幂定理与托勒密定理

正弦定理和余弦定理处理的是三角形的边角关系; 当问题中出现圆时, 还需要圆内线段关系的工具. **圆幂定理**与**托勒密定理**用乘积等式把圆与线段长度联系起来.

在平面上给定一个圆和一个点 P, 过 P 作任意直线与圆相交, 则该直线被圆截得的两条线段长度之积是常数. 这个常数称为点 P 关于该圆的**幂**.

按点 P 与圆的位置关系, 圆幂定理分为三种情形. 点在圆内、圆上或圆外.

12.1.5.1 相交弦定理

定理. 圆内的两条相交弦, 被交点分成的两条线段长的积相等. 如图, 若圆内两弦 AB、CD 交于点 P, 可得

$$PA\cdot PB=PC\cdot PD \tag{12-6}$$

证明 (证明) 连接 AC 与 BD. 在 $\triangle PAC$ 和 $\triangle PDB$ 中,

- $\angle APC=\angle DPB$(对顶角相等);
- $\angle PAC=\angle PDB$(同弧 $\overset{\frown}{CB}$ 所对的圆周角相等).

因此 $\triangle PAC\sim\triangle PDB$(AA 相似). 对应边成比例, 有 $\frac{PA}{PD}=\frac{PC}{PB}$. 于是 $PA\cdot PB=PC\cdot PD$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

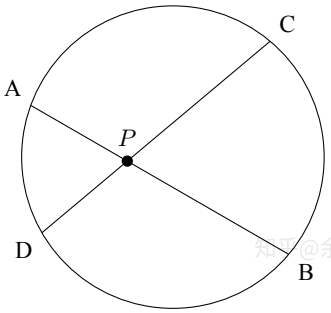


图 12.7 相交弦定理

12.1.5.2 切割线定理

定理. 从圆外一点引圆的切线和割线, 切线长的平方等于该点到割线与圆两交点距离的乘积. 如图, 若 PT 是圆的切线(T 为切点), PAB 是割线, 可得

$$PT^2 = PA \cdot PB \tag{12-7}$$

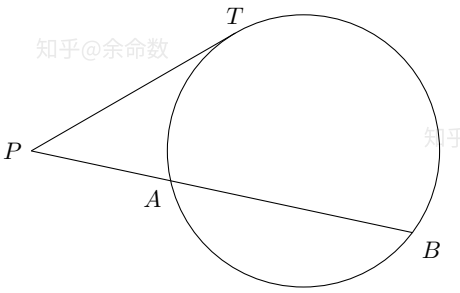


图 12.8 切割线定理

证明 (证明) 连接 AT 与 BT. 在 $\triangle PAT$ 和 $\triangle PTB$ 中,

- $\angle APT = \angle TPB$(公共角);
- $\angle PTA = \angle PBT$(弦切角定理).

因此 $\triangle PAT \sim \triangle PTB$ (AA 相似). 对应边成比例, 有 $\frac{PA}{PT} = \frac{PT}{PB}$. 于是 $PT^2 = PA \cdot PB$. □

三种形态的圆幂定理统一描述同一事实. 点 P 关于圆 O（半径为 R）的幂 $p(P) = |PO^2 - R^2|$ 等于上述各定理中那个不变的乘积.

12.1.5.3 托勒密定理

圆内接四边形的对角线关系

圆内接四边形的两组对边乘积之和等于两条对角线的乘积. 这就是托勒密定理.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

托勒密定理. 若四边形 $ABCD$ 内接于圆, 则

$$AB \cdot CD + BC \cdot DA = AC \cdot BD \tag{12-8}$$

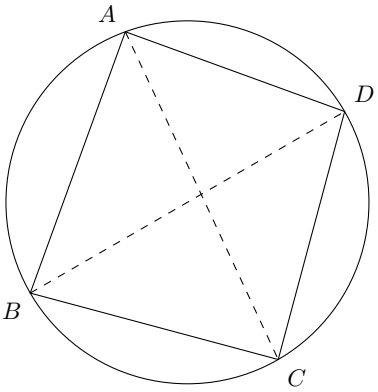


图 12.9　托勒密定理

知乎@余命数

证明 在对角线 AC 上取一点 K, 使 $\angle ABK = \angle DBC$, 再构造两对相似三角形.

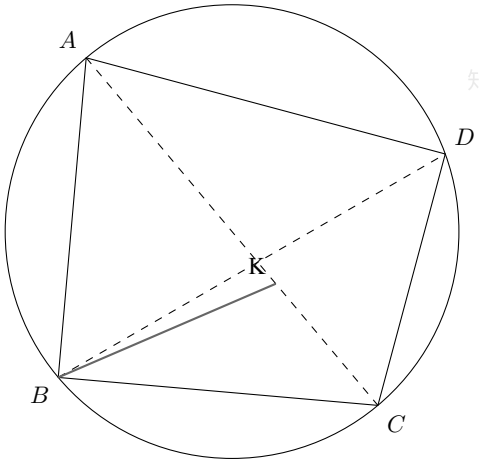


图 12.10　证明辅助线

如图, 对角线 AC 上取点 K 使 $\angle ABK = \angle DBC$.

先看 $\triangle ABK$ 与 $\triangle DBC$. 在两三角形中,

- $\angle ABK = \angle DBC$(作图);
- $\angle BAK = \angle BDC$(同弧 $\overset{\frown}{BC}$ 所对的圆周角相等).

故 $\triangle ABK \sim \triangle DBC$, 于是

$$\frac{AB}{DB} = \frac{AK}{DC}, \quad AB \cdot DC = AK \cdot DB. \tag{1}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再看 $\triangle ABD$ 与 $\triangle KBC$. 在两三角形中,

- $\angle ABD = \angle KBC$;
- $\angle ADB = \angle KCB$(同弧 $\overset{\frown}{AB}$ 所对的圆周角相等).

故 $\triangle ABD \sim \triangle KBC$, 于是

$$\frac{AD}{KC} = \frac{BD}{BC}, \quad AD \cdot BC = KC \cdot BD. \tag{2}$$

知乎@余命数

将 (1) 与 (2) 相加,

$$AB \cdot DC + AD \cdot BC = (AK + KC) \cdot BD = AC \cdot BD.$$

□

应用与拓展

- **托勒密不等式**任意凸四边形 $ABCD$ 满足 $AB \cdot CD + BC \cdot DA \geq AC \cdot BD$, 等号成立当且仅当四边形内接于圆. 该不等式可用于证明四点共圆, 也可处理某些几何不等式.
- **三角恒等式**托勒密定理应用于直径为 1 的圆内接四边形, 可导出两角和与差的正弦、余弦公式.

12.1.6 重要线段与关联定理

知乎@余命数

中线、角平分线、高线除了围成三角形的重心、内心、垂心外, 它们各自的长度也是解三角形中的常见量. 本节给出它们的长度公式, 这些公式是余弦定理与面积公式的直接应用.

12.1.6.1 中线

已知三边求中线长度时, 可在被中线分成的两个小三角形中分别使用余弦定理, 再利用互补角消去余弦项.

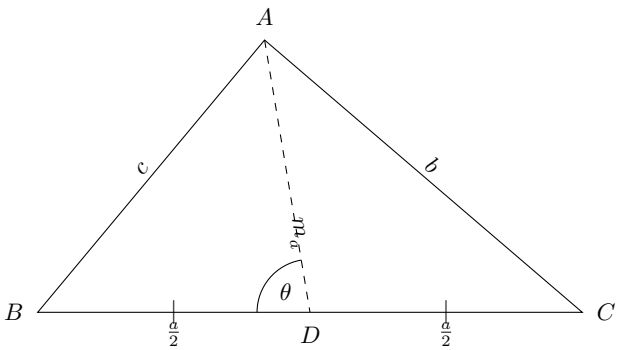
定理 12.9 (中线长公式) 设 m_a 为 $\triangle ABC$ 中边 a 上的中线, 则其长度为.

$$m_a = \frac{1}{2}\sqrt{2(b^2 + c^2) - a^2} \tag{12-9}$$

等价地, 公式也常写为 $4m_a^2 = 2b^2 + 2c^2 - a^2$.

证明 设 D 为 BC 的中点, 则 AD 为中线 m_a, 且 $BD = DC = \frac{a}{2}$.

知乎@余命数



知乎@余命数

在 $\triangle ABD$ 中, 设 $\angle ADB=\theta$. 由余弦定理,

$$c^2=m_a^2+\frac{a^2}{4}-am_a\cos\theta. \tag{1}$$

在 $\triangle ACD$ 中, $\angle ADC=\pi-\theta$, 故 $\cos\angle ADC=-\cos\theta$. 由余弦定理,

知乎@余命数

$$b^2=m_a^2+\frac{a^2}{4}+am_a\cos\theta. \tag{2}$$

(1)+(2) 消去 $\cos\theta$:

$$b^2+c^2=2m_a^2+\frac{a^2}{2},$$

整理即得 $4m_a^2=2(b^2+c^2)-a^2$. □

阿波罗尼斯定理

中线长公式等价于 $b^2+c^2=2\left(m_a^2+\left(\frac{a}{2}\right)^2\right)$, 即阿波罗尼斯定理.

12.1.6.2 角平分线

角平分线把角与边联系起来. 我们可用面积法求它的长度, 先把大三角形分成两个小三角形, 使它们的面积之和等于原三角形面积. 此外, 角平分线还满足角平分线定理.

定理 12.10 (角平分线长公式) 设 t_a 为 $\triangle ABC$ 中角 A 的内角平分线, 则其长度为.

$$t_a=\frac{2bc}{b+c}\cos\frac{A}{2} \tag{12-10}$$

证明 设 AD 为角 A 的平分线 t_a, 则 $\angle BAD=\angle CAD=\frac{A}{2}$. 由面积关系

$$S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle ACD}$$

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

得

$$\frac{1}{2}bc\sin A=\frac{1}{2}ct_a\sin\frac{A}{2}+\frac{1}{2}bt_a\sin\frac{A}{2}.$$

用 $\sin A=2\sin\frac{A}{2}\cos\frac{A}{2}$ 并约去 $\frac{1}{2}\sin\frac{A}{2}$（$A\in(0,\pi)$ 时不为零），

$$2bc\cos\frac{A}{2}=(b+c)t_a,$$

于是 $t_a=\frac{2bc}{b+c}\cos\frac{A}{2}$. □

另一个常用形式

角平分线公式还有纯边长形式

$$t_a^2=bc\left[1-\left(\frac{a}{b+c}\right)^2\right],$$

可由 $\cos^2\frac{A}{2}=\frac{1+\cos A}{2}$ 与余弦定理导出.

定理 12.11 (角平分线定理) 三角形的内角平分线把对边分成两段，这两段的比等于夹此角两边的比. 若 AD 是角 A 的平分线，则

$$\frac{AB}{AC}=\frac{BD}{CD},\qquad \text{即}\qquad \frac{c}{b}=\frac{BD}{CD}. \tag{12-11}$$

证明 在 $\triangle ABD$ 和 $\triangle ACD$ 中分别使用正弦定理，

$$\frac{BD}{\sin\frac{A}{2}}=\frac{c}{\sin\angle ADB},\qquad \frac{CD}{\sin\frac{A}{2}}=\frac{b}{\sin\angle ADC}.$$

由于 $\angle ADB+\angle ADC=\pi$，所以 $\sin\angle ADB=\sin\angle ADC$. 约去 $\sin\frac{A}{2}$ 即得 $\frac{c}{b}=\frac{BD}{CD}$. □

12.1.6.3 高线

高线的长度既可用三角函数定义直接求，也可通过面积公式间接求.

定理 12.12 (高线长) 设 h_a 为 $\triangle ABC$ 中边 a 上的高，其长度可以通过以下方式求得

1. **利用三角函数定义** $h_a=c\sin B=b\sin C$.
2. **利用面积公式** $S=\frac{1}{2}ah_a\implies h_a=\frac{2S}{a}$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

证明 设从 A 向 BC 作垂线，垂足为 D，则 $AD = h_a$. 在直角三角形 ABD 中，$h_a = AB\sin B = c\sin B$；在直角三角形 ACD 中，$h_a = AC\sin C = b\sin C$. 因此

$$h_a = c\sin B = b\sin C.$$

另一方面，以 $BC = a$ 为底、h_a 为高计算面积，有

$$S = \frac{1}{2}ah_a.$$

因为 $a > 0$，整理得

$$h_a = \frac{2S}{a}.$$

□

解题思路

- 已知两边及夹角，如 b, c, A 时，先用余弦定理求出 a，再由 $S = \frac{1}{2}bc\sin A$ 和 $h_a = \frac{2S}{a}$ 求高.
- 已知两角及一边，如 A, B, c 时，可先用正弦定理求出 b，再由 $h_a = b\sin C (C = \pi - A - B)$ 求高.

12.2 范围与综合

12.2.1 取值范围

给定一个三角形，我们常问它的某条边、某个角或某个面积表达式的取值范围. 这类问题可分三步处理. 先用正弦定理或余弦定理把多变量几何关系化归为单变量函数. 再依据三角形存在的约束确定这个函数的定义域. 最后用辅助角公式、二次函数或导数等工具求它的最值或范围.

12.2.1.0.1 构建目标函数 我们把问题中的多个边、角几何量经正弦定理与余弦定理改写，最终化归为一个以单个角或单个边长比值为自变量的函数.

- **正弦定理**是处理边的齐次式或比例式的首选工具. 对于形如 $\frac{a^2+b^2}{c^2}$ 或 $\frac{a+b}{c}$ 的表达式，其代数结构的对称性表明其值仅与三角形的形状，即角有关，与具体尺寸无关. 利用 $a = 2R\sin A, b = 2R\sin B, c = 2R\sin C$ 边化角，可消去外接圆半径 R 这一尺寸因素，得到纯粹的三角函数关系式.
- **余弦定理**在处理含边长平方和或已知两边一夹角的情形时直接决定结果，它直接建立起边与角之间的非线性方程.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

随后, 必须利用三角形内角和定理 $A+B+C=\pi$ 消去多余的角变量, 从而完成单变量目标函数的构建.

12.2.1.0.2 确定定义域 一个代数表达式能描述几何对象, 当且仅当该对象真实存在. 因此, 我们先用三角形存在的几何约束确定函数的定义域, 再分析函数.

- **基本约束**. 三角形的任意内角都必须在 $(0,\pi)$ 范围内.
- **结构约束**. 在变量化归后, 由自变量表示的其他角也必须满足基本约束. 例如, 若已将其他角用变量角 B 表示为 $A=\frac{\pi}{2}-2B, C=\frac{\pi}{2}+B$, 则必须同时满足 $B>0, A>0, C>0$ 所构成的不等式组, 从而解出 B 的最终有效范围.
- **附加约束**. 若题目明确限定了三角形的形状, 如锐角三角形, 则必须增加所有角均小于 $\frac{\pi}{2}$ 的条件, 这通常需要通过余弦定理转化为关于边或角的不等式.

忽略对定义域的精确界定, 是此类问题中最根本的错误来源.

12.2.1.0.3 分析与求解 于是原问题化为一个定义在精确区间上的函数分析问题. 我们根据目标函数的具体形式, 选用相应的代数或分析工具.

- 对于形如 $y=a\sin x+b\cos x$ 的函数, **辅助角公式**是标准处理方法.
- 对于形如 $y=a\sin^2 x+b\sin x+c$ 的函数, 可通过换元法转化为**二次函数**在闭区间上的最值问题.
- 对于更复杂的函数, **求导法**是判断其单调性、寻找极值点的最通用的工具.

例题 12.1 (2022 年 6 月新高考 I 卷第 18 题) 记 $\triangle ABC$ 的内角 A,B,C 的对边分别为 a,b,c, 已知 $\frac{\cos A}{1+\sin A}=\frac{\sin 2B}{1+\cos 2B}$. (1) 若 $C=\frac{2\pi}{3}$, 求 B; (2) 求 $\frac{a^2+b^2}{c^2}$ 的最小值.

解 先把给定的三角恒等式化简, 以得到角 A 与 B 之间的关系. 右侧用二倍角

$$\frac{\sin 2B}{1+\cos 2B}=\frac{2\sin B\cos B}{1+(2\cos^2 B-1)}=\frac{2\sin B\cos B}{2\cos^2 B}=\tan B.$$

左侧用诱导公式与半角公式

$$\frac{\cos A}{1+\sin A}=\frac{\sin(\frac{\pi}{2}-A)}{1+\cos(\frac{\pi}{2}-A)}=\tan\left(\frac{\pi}{4}-\frac{A}{2}\right).$$

原式化为 $\tan(\frac{\pi}{4}-\frac{A}{2})=\tan B$. 令 $x=\frac{\pi}{4}-\frac{A}{2}\in(-\frac{\pi}{4},\frac{\pi}{4})$, 则 $B=x+k\pi$. 由 $B\in(0,\pi)$, 只可能 $k=0$ 或 1; 若 $k=1$ 则 $A+B>\pi$, 与内角和矛盾. 故 $k=0$, $\frac{\pi}{4}-\frac{A}{2}=B$, 即 $A+2B=\frac{\pi}{2}$.

(1) 已知 $C=\frac{2\pi}{3}$, 则 $A+B=\pi-C=\frac{\pi}{3}$. 联立 $\begin{cases}A+2B=\frac{\pi}{2}\\ A+B=\frac{\pi}{3}\end{cases}$, 相减得 $B=\frac{\pi}{6}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2) 求 $\frac{a^2+b^2}{c^2}$ 的最小值. 该式关于边长二次齐次, 其值只依赖形状. 由正弦定理 $a=2R\sin A$ 等,

$$\frac{a^2+b^2}{c^2}=\frac{\sin^2 A+\sin^2 B}{\sin^2 C}.$$

代入 $A=\frac{\pi}{2}-2B, C=\frac{\pi}{2}+B$:

$$\frac{\sin^2 A+\sin^2 B}{\sin^2 C}=\frac{\cos^2(2B)+\sin^2 B}{\cos^2 B}.$$

内角为正给出 $B\in(0,\frac{\pi}{4})$. 令 $t=\sin^2 B\in(0,\frac{1}{2})$, 问题化为求 $g(t)=\frac{(1-2t)^2+t}{1-t}=\frac{4t^2-3t+1}{1-t}$ 在 $(0,\frac{1}{2})$ 上的最小值. 变形得 $g(t)=-4t-1+\frac{2}{1-t}$, 求导 $g'(t)=-4+\frac{2}{(1-t)^2}$. 令 $g'=0$ 得 $t=1-\frac{\sqrt{2}}{2}\in(0,\frac{1}{2})$; 在该点左侧 $g'<0$、右侧 $g'>0$, 故为最小值点,

$$g\Big(1-\frac{\sqrt{2}}{2}\Big)=-4\Big(1-\frac{\sqrt{2}}{2}\Big)-1+\frac{2}{\sqrt{2}/2}=4\sqrt{2}-5.$$

综上, (1) $B=\frac{\pi}{6}$; (2) 最小值为 $4\sqrt{2}-5$.

例题 12.2 (2023 年 8 月浙江省 A9 协作体返校联考第 17 题) 在 $\triangle ABC$ 中, 角 A,B,C 所对的边分别为 a,b,c, 且满足 $a\sin C\cos B+b\sin A\cos C=\frac{\sqrt{3}}{2}a$. (1) 求角 A; (2) 若 $\triangle ABC$ 为锐角三角形, 求 $4\sin^2 B-4\sin B\sin C$ 的取值范围.

解 对第 (1) 问, 由已知条件 $a\sin C\cos B+b\sin A\cos C=\frac{\sqrt{3}}{2}a$ 出发. 根据正弦定理, 有 $b\sin A=a\sin B$. 将此关系代入原式中, 可以消去变量 b, 使得 a 成为一个可约去的公因子.

$$a\sin C\cos B+(a\sin B)\cos C=\frac{\sqrt{3}}{2}a$$

由于 $a\neq 0$, 提取公因子 a 并利用和角公式, 可得

$$\sin(B+C)=\frac{\sqrt{3}}{2}$$

在 $\triangle ABC$ 中, $B+C=\pi-A$, 故 $\sin(\pi-A)=\sin A=\frac{\sqrt{3}}{2}$. 由于 $A\in(0,\pi)$, 满足此条件的角 A 为 $\frac{\pi}{3}$ 或 $\frac{2\pi}{3}$.

对第 (2) 问, 欲求 $4\sin^2 B-4\sin B\sin C$ 的取值范围. 首先, 将目标表达式化归为关于单个角 B 的函数. 由于第 (2) 问已限定 $\triangle ABC$ 为锐角三角形, A 必为锐角; 结合 (1) 可得 $A=\frac{\pi}{3}$, 从而 $C=\frac{2\pi}{3}-B$. 代入表达式中.

$$f(B)=4\sin^2 B-4\sin B\sin\Big(\frac{2\pi}{3}-B\Big)$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

$$= 4\sin^2 B - 4\sin B\left(\frac{\sqrt{3}}{2}\cos B + \frac{1}{2}\sin B\right)$$
$$= 2\sin^2 B - \sqrt{3}\sin(2B)$$

为了便于分析，将所有项都用 $2B$ 的三角函数来表示. 利用降幂公式 $2\sin^2 B = 1-\cos(2B)$, 上式可化为.

$$f(B) = 1 - (\sqrt{3}\sin(2B) + \cos(2B))$$

使用辅助角公式进行合并, 得到 $f(B) = 1 - 2\sin\left(2B + \frac{\pi}{6}\right)$.

接下来确定自变量 B 的精确定义域. 锐角三角形的几何约束给出如下代数不等式组:

$$\begin{cases} A = \frac{\pi}{3} < \frac{\pi}{2} & \text{(已满足)} \\ B \in (0, \frac{\pi}{2}) \\ C = \frac{2\pi}{3} - B \in (0, \frac{\pi}{2}) \implies \frac{\pi}{6} < B < \frac{2\pi}{3} \end{cases}$$

将 B 的所有约束条件取交集, 得到其最终的取值范围为 $B \in (\frac{\pi}{6}, \frac{\pi}{2})$.

最后，基于 B 的定义域来求解函数的值域. 由 $B \in (\frac{\pi}{6}, \frac{\pi}{2})$，可得 $2B + \frac{\pi}{6} \in (\frac{\pi}{2}, \frac{7\pi}{6})$. 在此开区间内, $\sin(2B + \frac{\pi}{6})$ 的取值从 $\sin(\frac{\pi}{2}) = 1$ 单调递减至 $\sin(\frac{7\pi}{6}) = -\frac{1}{2}$. 由于区间为开区间, 端点值均无法取到, 故 $\sin(2B + \frac{\pi}{6}) \in (-\frac{1}{2}, 1)$. 对该不等式进行代数变换, 即可得到 $f(B)$ 的范围.

$$-2\sin\left(2B + \frac{\pi}{6}\right) \in (-2, 1)$$

$$f(B) = 1 - 2\sin\left(2B + \frac{\pi}{6}\right) \in (-1, 2)$$

(1) $A = \frac{\pi}{3}$ 或 $A = \frac{2\pi}{3}$.(2) $4\sin^2 B - 4\sin B\sin C$ 的取值范围为 $(-1, 2)$.

方法论要点

上面的例子有一个共同模式. 先把三角形的几何约束化为单变量函数的定义域, 再用三角恒等式、辅助角公式或基本不等式求它的最值或范围.

1. **翻译与化归**. 利用正弦定理与余弦定理, 将问题中的边、角关系统一为纯粹的三角函数关系式. 随后, 利用内角和定理消去多余变量, 构建单变量目标函数.
2. **定义域的确定**. 这是解题的逻辑基础. 必须从三角形存在的固有几何约束, 如所有内角为正以及题目附加的条件约束, 如锐角三角形出发,

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将它们全部转化为关于自变量的不等式组, 从而精确界定函数的定义域.

3. **函数分析**. 在确定的定义域上, 根据函数的形式, 灵活选用辅助角公式、二次函数理论或求导法等分析工具, 求解其最值或范围.

求范围时容易忽略一点. 由自变量表示的其他角也必须为正. 边化角与角化边的选择若不明显, 可先观察表达式是否关于边齐次, 或是否能通过恒等变形消去角.

例题 12.3 (2023 年 1 月重庆康德卷一诊第 17 题) 在 $\triangle ABC$ 中, 角 A, B, C 的对边分别为 a, b, c, 且 $b = c(\cos A + \sin A)$. (1) 求角 C; (2) 求 $\frac{a+\sqrt{2}b}{c}$ 的最大值.

解 对第 (1) 问, 已知条件 $b = c(\cos A + \sin A)$ 是一个典型的边角混合关系式. 解决此类问题的要点在于统一语言体系. 利用正弦定理将边长 b, c 替换为对应角的正弦, 是消除边长、得到纯角关系式的最直接路径.

根据正弦定理, 有 $b = 2R\sin B$ 和 $c = 2R\sin C$. 代入已知等式中,

$$2R\sin B = 2R\sin C(\cos A + \sin A)$$

知乎@余命数

约去非零公因式 $2R$, 得到一个纯粹的角的关系式.

$$\sin B = \sin C\cos A + \sin C\sin A$$

在 $\triangle ABC$ 中, 由 $B = \pi - (A + C)$, 可得 $\sin B = \sin(\pi - (A + C)) = \sin(A + C)$. 利用和角公式展开, 得 $\sin B = \sin A\cos C + \cos A\sin C$. 联立上述两个关于 $\sin B$ 的表达式, 可得

知乎@余命数

$$\sin A\cos C + \cos A\sin C = \sin C\cos A + \sin C\sin A$$

整理上式, 消去 $\cos A\sin C$ 项, 得到 $\sin A\cos C = \sin C\sin A$. 由于 $A \in (0, \pi)$, 故 $\sin A \neq 0$. 两边同除以 $\sin A$ 可得 $\cos C = \sin C$. 又因 $C \in (0, \pi)$, 满足此条件的角 C 只能是 $C = \frac{\pi}{4}$.

对第 (2) 问, 求 $\frac{a+\sqrt{2}b}{c}$ 的最大值. 该表达式是关于边长的一次齐次式, 同样应用正弦定理进行边化角.

知乎@余命数

$$\frac{a+\sqrt{2}b}{c} = \frac{2R\sin A + \sqrt{2}(2R\sin B)}{2R\sin C} = \frac{\sin A + \sqrt{2}\sin B}{\sin C}$$

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

接着, 将表达式中的所有角用角 A 来表示, 以构建单变量函数. 由 (1) 知 $C=\frac{\pi}{4}$, 故 $B=\pi-(A+C)=\frac{3\pi}{4}-A$. 代入上式, 得

$$f(A)=\frac{\sin A+\sqrt{2}\sin(\frac{3\pi}{4}-A)}{\sin(\pi/4)}=\sqrt{2}\sin A+2\sin\left(\frac{3\pi}{4}-A\right)$$

利用和角公式展开, 得

$$\begin{aligned}f(A)&=\sqrt{2}\sin A+2\left(\sin\frac{3\pi}{4}\cos A-\cos\frac{3\pi}{4}\sin A\right)\\&=\sqrt{2}\sin A+2\left(\frac{\sqrt{2}}{2}\cos A+\frac{\sqrt{2}}{2}\sin A\right)\\&=2\sqrt{2}\sin A+\sqrt{2}\cos A\end{aligned}$$

使用辅助角公式将其合并, 得到 $f(A)=\sqrt{10}\sin(A+\phi)$, 其中角 ϕ 满足 $\tan\phi=\frac{1}{2}$.

须确定自变量 A 的定义域. 三角形的几何存在性要求其内角为正.

$$\begin{cases}A>0\\B=\frac{3\pi}{4}-A>0\end{cases}\implies A<\frac{3\pi}{4}$$

故 $A\in(0,\frac{3\pi}{4})$. 函数 $f(A)=\sqrt{10}\sin(A+\phi)$ 的最大值为 $\sqrt{10}$, 这当且仅当 $\sin(A+\phi)=1$ 时取得, 即 $A+\phi=\frac{\pi}{2}$. 由于 $\tan\phi=\frac{1}{2}$, 可知 ϕ 是一个锐角. 因此, 使得函数取最大值时的角 $A=\frac{\pi}{2}-\phi$ 位于区间 $(0,\frac{\pi}{2})$ 内, 此值满足定义域 $A\in(0,\frac{3\pi}{4})$, 故最大值 $\sqrt{10}$ 可以取到.

(1) $C=\frac{\pi}{4}$.(2) $\frac{a+\sqrt{2}b}{c}$ 的最大值为 $\sqrt{10}$.

例题 12.4 (巴蜀中学 2023 届高三适应性月考 (七) 第 17 题) 在 $\triangle ABC$ 中, 角 A,B,C 的对边分别为 a,b,c, 已知 $3\sin C+4\cos C=5$. (1) 求证 $\tan C=\frac{3}{4}$; (2) 若 $a^2+b^2=1$, 求边 c 的最小值.

解 对第 (1) 问,已知等式 $3\sin C+4\cos C=5$ 是一个标准的 $p\sin x+q\cos x$ 形式. 利用辅助角公式,等式左侧可化为 $5\left(\frac{3}{5}\sin C+\frac{4}{5}\cos C\right)=5$. 令 $\cos\phi=\frac{3}{5}$, $\sin\phi=\frac{4}{5}$,则 $\tan\phi=\frac{4}{3}$. 原式变为 $\sin(C+\phi)=1$. 由于 $C\in(0,\pi)$ 且 ϕ 为锐角,故 $C+\phi\in(\phi,\pi+\phi)$. 在此区间内, 满足 $\sin(C+\phi)=1$ 的唯一解是 $C+\phi=\frac{\pi}{2}$. 因此 $C=\frac{\pi}{2}-\phi$. 故 $\tan C=\tan(\frac{\pi}{2}-\phi)=\cot\phi=\frac{\cos\phi}{\sin\phi}=\frac{3/5}{4/5}=\frac{3}{4}$.

对第 (2) 问, 由 (1) 知 $\tan C=\frac{3}{4}$, 且 C 为三角形内角, 故 C 为锐角. 由此可得 $\cos C=\frac{4}{5}$. 根据余弦定理 $c^2=a^2+b^2-2ab\cos C$, 代入已知条件 $a^2+b^2=1$ 和 $\cos C=\frac{4}{5}$:

$$c^2=1-2ab\left(\frac{4}{5}\right)=1-\frac{8}{5}ab$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

欲求 c 的最小值, 即求 c^2 的最小值, 这等价于寻求 ab 的最大值. 根据算术平均值-几何平均值不等式, 对于正数 a, b, 有

$$1 = a^2 + b^2 \geq 2\sqrt{a^2b^2} = 2ab \implies ab \leq \frac{1}{2}$$

当且仅当 $a = b = \frac{\sqrt{2}}{2}$ 时, 等号成立, 此时 ab 取得最大值 $\frac{1}{2}$. 将 ab 的最大值代入 c^2 的表达式中.

$$c^2 = 1 - \frac{8}{5}ab \geq 1 - \frac{8}{5}\left(\frac{1}{2}\right) = 1 - \frac{4}{5} = \frac{1}{5}$$

故 c^2 的最小值为 $\frac{1}{5}$. 因为 c 是边长, 必为正数, 所以 c 的最小值为 $\sqrt{\frac{1}{5}} = \frac{\sqrt{5}}{5}$.

(1) 证明见上.(2) 边 c 的最小值为 $\frac{\sqrt{5}}{5}$.

例题 12.5 (2023 年 3 月温州市高三二模第 20 题) 已知 $\triangle ABC$ 满足

$$2\sin C\sin(B-A) = 2\sin A\sin C - \sin^2 B.$$

(1) 试问角 B 是否可能为直角? 请说明理由; (2) 若 $\triangle ABC$ 为锐角三角形, 求 $\dfrac{\sin C}{\sin A}$ 的取值范围.

解 原方程为 $2\sin C\sin(B-A) = 2\sin A\sin C - \sin^2 B$. 将 $\sin(B-A)$ 展开.$2\sin C(\sin B\cos A - \cos B\sin A) = 2\sin A\sin C - \sin^2 B$. 利用正弦定理与余弦定理进行角化边, 把三角函数关系转化为边的代数关系.

$$2\frac{c}{2R}\left(\frac{b}{2R}\frac{b^2+c^2-a^2}{2bc} - \frac{a}{2R}\frac{a^2+c^2-b^2}{2ac}\right) = 2\frac{a}{2R}\frac{c}{2R} - \left(\frac{b}{2R}\right)^2$$

$$2c\left(\frac{b^2+c^2-a^2}{2c} - \frac{a^2+c^2-b^2}{2c}\right) = 2ac - b^2$$

$$(b^2+c^2-a^2) - (a^2+c^2-b^2) = 2ac - b^2$$

$$2b^2 - 2a^2 = 2ac - b^2$$

整理得到三边之间满足的核心关系式.$3b^2 - 2a^2 = 2ac$.

对第 (1) 问, 采用反证法. 假设角 B 可能为直角, 即 $B = \frac{\pi}{2}$. 根据勾股定理, 此时三边满足 $b^2 = a^2 + c^2$. 将此关系代入已得的核心关系式中, 得 $3(a^2+c^2) - 2a^2 = 2ac \implies a^2 - 2ac + 3c^2 = 0$. 这是一个关于 a 的齐次二次方程. 将其视为关于 a 的一元二次方程, 考察其判别式. $\Delta = (-2c)^2 - 4(1)(3c^2) = 4c^2 - 12c^2 = -8c^2$. 因为 c 是边长, $c > 0$, 所以 $\Delta < 0$. 该方程关于 a 没有实数解. 因此, 最初的假设不成立, 角 B 不可能为直角.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对第 (2) 问, 所求为 $\frac{\sin C}{\sin A}$, 根据正弦定理, 这等价于求 $\frac{c}{a}$ 的取值范围. 设 $k=\frac{c}{a}$, 则 $c=ka$. 将 $c=ka$ 代入核心关系式 $3b^2-2a^2=2ac$ 中, 用 a 和 k 表示 b^2: $3b^2-2a^2=2a(ka)=2ka^2 \implies b^2=\frac{2k+2}{3}a^2$.

$\triangle ABC$ 为锐角三角形, 意味着 $\cos A>0, \cos B>0, \cos C>0$.

- 由 $\cos A>0 \implies b^2+c^2-a^2>0$: $\frac{2k+2}{3}a^2+(ka)^2-a^2>0 \implies 3k^2+2k-1>0$. 分解得 $(3k-1)(k+1)>0$. 因为 $k=\frac{c}{a}>0$, 所以 $k>\frac{1}{3}$.
- 由 $\cos C>0 \implies a^2+b^2-c^2>0$: $a^2+\frac{2k+2}{3}a^2-(ka)^2>0 \implies 3k^2-2k-5<0$. 分解得 $(3k-5)(k+1)<0$. 因为 $k>0$, 所以 $k<\frac{5}{3}$.
- 由 $\cos B>0 \implies a^2+c^2-b^2>0$: $a^2+(ka)^2-\frac{2k+2}{3}a^2>0 \implies 3k^2-2k+1>0$. 该二次函数判别式 $\Delta=(-2)^2-4(3)(1)=-8<0$, 开口向上, 故恒成立.

综合以上三个条件, 得 k 的取值范围是 $\frac{1}{3}<k<\frac{5}{3}$.

(1) 角 B 不可能为直角, 理由见上.(2) $\frac{\sin C}{\sin A}$ 的取值范围为 $\left(\frac{1}{3},\frac{5}{3}\right)$.

12.2.1.0.4 解题流程 解三角形的最值与范围问题遵循同一条由几何到代数、由多维到一维的路径. 把与三角形形态相关的几何问题, 转化为求单变量函数在特定区间上最值的代数问题.

1. **建立代数模型**. 利用正弦定理与余弦定理, 将问题中的所有边、角关系, 统一转化为纯粹的三角函数关系式或纯粹的边长关系式.
2. **变量化归**. 利用三角形内角和定理 $A+B+C=\pi$, 消去多余的角, 将一个含有多个变量的表达式, 转化为一个只含单个变量的单变量函数.
3. **确定定义域**. 这是解题的逻辑基础. 必须从三角形存在的固有几何约束, 如所有内角为正以及题目附加的条件约束, 如锐角三角形出发, 将它们全部转化为关于自变量的不等式组, 从而精确界定函数的定义域.
4. **函数分析**. 在确定的定义域上, 根据函数的形式, 灵活选用辅助角公式、二次函数理论或求导法等分析工具, 求解其最值或范围.

常见逻辑疏漏

1. **定义域挖掘不全**. 最常见的错误是, 在变量化归后, 只考虑了自变量本身大于 0, 而忘记了由它表示的另外两个角也必须大于 0 的约束.
2. **工具选择失当**. 面对边化角还是角化边的抉择时犹豫不决. 经验法则是, 当表达式是关于边的齐次式时, 几乎总是边化角更优; 当三角函数关系复杂, 但能通过变形消去角, 得到纯粹的边长关系时, 角化边则能化繁为简.
3. **代数变形出错**. 在换元、配方、求导等纯代数运算步骤中出现计算失误. 解三角形问题往往计算量较大, 须步步为营、细心验算.

知乎@余命数

12.2.2 等分线

我们从三角形的一个顶点向对边引一条线段, 把原三角形分成两个子三角形. 整体与子三角形的约束通过公共边、公共角或互补角联系起来, 于是可用两次余弦定理建立方程.

12.2.2.1 解题蓝图

12.2.2.1.1 全局与局部关系的构建 先在整个 $\triangle ABC$ 中用余弦定理建立三边 a,b,c 与某个已知角之间的关系式, 得到全局约束. 再在被分割线段所在的子三角形中再次应用余弦定理, 引入分割线段的长度作为局部约束.

12.2.2.1.2 利用几何联系消元 两个子三角形共享分割线, 且相邻角互补, 如 $\angle ADB+\angle ADC=\pi$, 故 $\cos(\angle ADB)=-\cos(\angle ADC)$. 这一等式常用来消去未知角, 得到只含边长的方程.

12.2.2.1.3 联立求解 把全局方程与局部方程联立, 解这个关于未知边长的代数方程组, 即可得到答案.

斯图尔特定理的向量形式

当分割线段属于一般线段时, 直接在子三角形中应用余弦定理可能会引入新的未知角, 从而使问题复杂化. 此时, 向量方法提供了一条更具普适性且更为直接的路径. 若点 D 在边 BC 上, 且满足 $BD:DC=\mu:\lambda$, 则由向量的定比分点公式可知.

$$\overrightarrow{AD}=\frac{\lambda}{\lambda+\mu}\overrightarrow{AB}+\frac{\mu}{\lambda+\mu}\overrightarrow{AC}$$

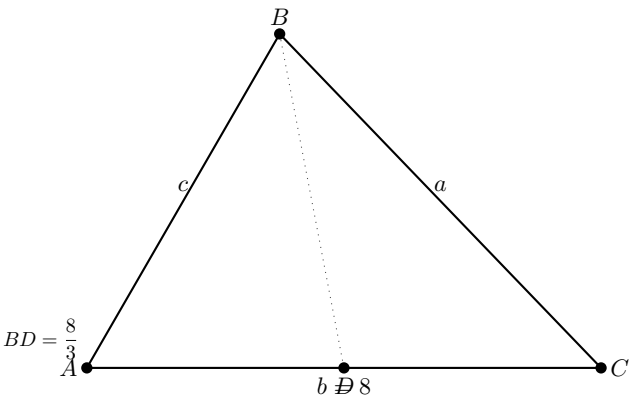
对此向量等式两边取自身的数量积, 即平方, 即可获得一个仅含三边与角 A 的、关于分割线长 AD 的表达式, 从而避免引入新的未知角. 这一关系实质上是斯图尔特定理的向量表述.

例题 12.6 (中线分割模型) 在 $\triangle ABC$ 中, 角 A,B,C 的对边分别为 a,b,c. 已知 $\frac{2\cos C}{a}=\frac{2}{b}+\frac{\sin C}{b\sin A}$. 若 $b=8$, D 为边 AC 的中点, 且 $BD=\frac{8}{3}$, 求 $\triangle ABC$ 的面积.

解 本题的线段是中线. 我们联立全局三角形与局部中线给出的两个方程.

首先, 分析初始条件. 将等式 $\frac{2\cos C}{a}=\frac{2}{b}+\frac{\sin C}{b\sin A}$ 两边同乘以 ab 得 $2b\cos C=2a+\frac{a\sin C}{\sin A}$. 由正弦定理, $\frac{a}{\sin A}=\frac{c}{\sin C}$, 代入上式得 $2b\cos C=2a+c$. 再次利用余弦定理 $\cos C=\frac{a^2+b^2-c^2}{2ab}$ 代入, 得 $2b\cdot\frac{a^2+b^2-c^2}{2ab}=2a+c$, 化简得 $a^2+b^2-c^2=2a^2+ac$, 即

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵



$b^2=a^2+c^2+ac$. 将其与余弦定理 $b^2=a^2+c^2-2ac\cos B$ 对比, 可得 $\cos B=-\frac{1}{2}$, 故 $B=\frac{2\pi}{3}$.

接着, 构建关于边长 a,c 的方程组. 其一, 是对全局 $\triangle ABC$ 应用余弦定理, 这构成了第一个方程.

$$b^2=a^2+c^2-2ac\cos B \implies 8^2=a^2+c^2-2ac\left(-\frac{1}{2}\right)$$

$$a^2+c^2+ac=64 \quad \cdots\cdots(1)$$

其二, 是利用中线 BD 的信息. 中线长公式是连接中线与三边关系最直接的工具.

$$BD^2=\frac{2(a^2+c^2)-b^2}{4}$$

代入已知值

$$\left(\frac{8}{3}\right)^2=\frac{2(a^2+c^2)-8^2}{4} \implies \frac{64}{9}=\frac{2(a^2+c^2)-64}{4}$$

整理得 $2(a^2+c^2)=\frac{256}{9}+64=\frac{832}{9}$, 从而得到第二个方程.

$$a^2+c^2=\frac{416}{9} \quad \cdots\cdots(2)$$

联立方程 (1) 与 (2), 将 (2) 代入 (1) 中, 即可求得 ac 的值.

$$\frac{416}{9}+ac=64 \implies ac=64-\frac{416}{9}=\frac{160}{9}$$

最终, $\triangle ABC$ 的面积为

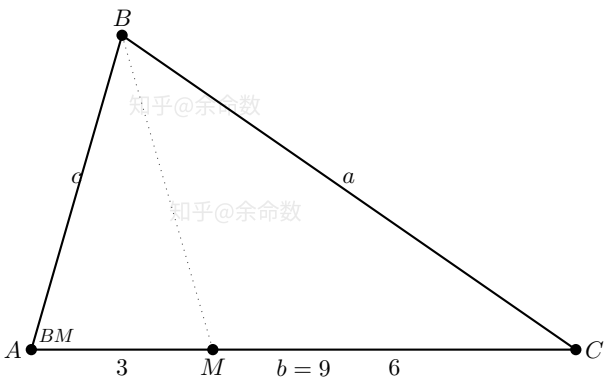
$$S=\frac{1}{2}ac\sin B=\frac{1}{2}\cdot\frac{160}{9}\cdot\sin\frac{2\pi}{3}=\frac{80}{9}\cdot\frac{\sqrt{3}}{2}=\frac{40\sqrt{3}}{9}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(1) $B=\frac{2\pi}{3}$.(2) $\triangle ABC$ 的面积为 $\frac{40\sqrt{3}}{9}$.

例题 12.7 (普通分割线模型) 在 $\triangle ABC$ 中, 角 A,B,C 的对边分别为 a,b,c, 且 $2b\cos C=2a+c$. 设 $b=9$, 点 M 是边 AC 上一点, $AM:MC=1:2$, 且 $\angle MAB=\angle MBA$. 求 $\triangle BMC$ 的面积.

解 本题的线段 BM 是一般线段, 直接在子三角形中用余弦定理会引入新的未知角. 我们用向量方法避免这个未知角.



首先, 由 $2b\cos C=2a+c$ 可推得 $B=\frac{2\pi}{3}$. 由 $AM:MC=1:2$ 及 $AC=b=9$, 可得 $AM=3,MC=6$. 条件 $\angle MAB=\angle MBA$ 表明 $\triangle ABM$ 为等腰三角形, 故 $BM=AM=3$.

接着, 构建关于边长 a,c 的方程组. 全局方程由 $\triangle ABC$ 的余弦定理给出.

$$9^2=a^2+c^2-2ac\cos\frac{2\pi}{3}\implies a^2+c^2+ac=81\quad\cdots\cdots(1)$$

局部方程的建立是此题的要点. 利用向量的定比分点公式.

$$\overrightarrow{BM}=\frac{2}{3}\overrightarrow{BA}+\frac{1}{3}\overrightarrow{BC}$$

对此向量等式两边取自身的数量积, 得

$$|\overrightarrow{BM}|^2=\frac{4}{9}|\overrightarrow{BA}|^2+\frac{1}{9}|\overrightarrow{BC}|^2+\frac{4}{9}(\overrightarrow{BA}\cdot\overrightarrow{BC})$$

将边长与角度代入, 得

$$3^2=\frac{4}{9}c^2+\frac{1}{9}a^2+\frac{4}{9}(c\cdot a\cdot\cos B)=\frac{1}{9}(4c^2+a^2-2ac)$$

整理得到第二个方程.

$$a^2+4c^2-2ac=81\quad\cdots\cdots(2)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

联立方程组 $\begin{cases} a^2+c^2+ac=81 \\ a^2+4c^2-2ac=81 \end{cases}$，解得 $a=c=3\sqrt{3}$.

在 $\triangle ABC$ 中，$a=c$ 意味着这是一个顶角为 120° 的等腰三角形，故 $A=C=30^\circ$. $\triangle BMC$ 的面积为

$$S_{\triangle BMC}=\frac{1}{2}\cdot BC\cdot MC\cdot\sin C=\frac{1}{2}\cdot(3\sqrt{3})\cdot 6\cdot\sin 30^\circ=\frac{9\sqrt{3}}{2}$$

(1) $B=\frac{2\pi}{3}$.(2) $\triangle BMC$ 的面积为 $\frac{9\sqrt{3}}{2}$.

12.2.2.2 内角平分线

角平分线把三角形的角与边联系起来. 处理角平分线时，我们常用面积法或角平分线定理.

1. **若涉及角平分线长度，首选面积法**. 面积法依据的事实是，大三角形的面积等于两个小三角形的面积之和，即 $S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle ACD}$. 我们把这个等式改写为含角平分线长度、两边及夹角的方程.
2. **若涉及对边线段比例，首选角平分线定理**. 角平分线定理断言，角平分线分割对边所成的两条线段之比，等于夹此角的两邻边之比，即 $\frac{BD}{CD}=\frac{c}{b}$.

例题 12.8 已知 $\triangle ABC$ 的内角 A,B,C 的对边分别为 a,b,c，且 $2b\cos A-a=2c$. 设 $\angle ABC$ 的角平分线 BD 交 AC 于点 D，若 $BD=2$，求 $\triangle ABC$ 的面积的最小值.

解 本题已知角平分线长度求面积最值. 我们用面积分割法建立关系，再用基本不等式求最值.

首先，分析条件 $2b\cos A-a=2c$. 根据射影定理 $c=a\cos B+b\cos A$，代入可得 $2b\cos A-a=2(a\cos B+b\cos A)$，消去 $2b\cos A$ 项得 $-a=2a\cos B$. 由于 $a\neq 0$，故 $\cos B=-\frac{1}{2}$，即 $B=\frac{2\pi}{3}$.

角平分线 BD 将 $\triangle ABC$ 分割为 $\triangle ABD$ 和 $\triangle CBD$. 利用面积关系.

$$S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle CBD}$$

角平分线将角 B 分为两个 $\frac{B}{2}=\frac{\pi}{3}$ 的角. 利用三角面积公式展开，得

$$\frac{1}{2}ac\sin B=\frac{1}{2}c\cdot BD\cdot\sin\frac{B}{2}+\frac{1}{2}a\cdot BD\cdot\sin\frac{B}{2}$$

代入已知值 $B=\frac{2\pi}{3}, BD=2$：

$$\frac{1}{2}ac\sin\frac{2\pi}{3}=\frac{1}{2}c(2)\sin\frac{\pi}{3}+\frac{1}{2}a(2)\sin\frac{\pi}{3}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两边同乘以 $\frac{2}{\sin(\pi/3)}$, 即 $\frac{4}{\sqrt{3}}$, 并利用 $\sin(2\pi/3)=\sin(\pi/3)$:

$$ac=2(a+c)$$

欲求 $\triangle ABC$ 面积 $S=\frac{1}{2}ac\sin B=\frac{\sqrt{3}}{4}ac$ 的最小值, 即求 ac 的最小值. 根据基本不等式, $a+c\geq 2\sqrt{ac}$. 将其代入已得的关系式中.

$$ac=2(a+c)\geq 2(2\sqrt{ac})=4\sqrt{ac}$$

因为 $ac>0$, 两边同除以 $\sqrt{ac}$ 得 $\sqrt{ac}\geq 4 \implies ac\geq 16$. 当且仅当 $a=c$ 时, 等号成立. 此时 $a^2=4a \implies a=4$. 故 ac 的最小值为 16.

因此, $\triangle ABC$ 面积的最小值为 $S_{\min}=\frac{\sqrt{3}}{4}\times 16=4\sqrt{3}$.

(1) $B=\frac{2\pi}{3}$.(2) $\triangle ABC$ 的面积的最小值为 $4\sqrt{3}$.

例题 12.9　在 $\triangle ABC$ 中, $\sin A+\sqrt{3}\cos A=0$. 给出以下三个条件.(1) $a=4\sqrt{3},b=4$;(2) $b^2-a^2+c^2+10b=0$;(3) $S_{\triangle ABC}=15\sqrt{3}$. 若这三个条件中有两个正确, 请选出正确的条件并求 $\angle BAC$ 的角平分线 AD 的长.

解　首先, 由 $\sin A+\sqrt{3}\cos A=0$ 可得 $\tan A=-\sqrt{3}$. 因 $A\in(0,\pi)$, 故 $A=\frac{2\pi}{3}$.

接着, 对三个条件进行检验, 以确定哪两个是相容的. 若条件 (1) 正确, $a=4\sqrt{3},b=4$. 由余弦定理 $a^2=b^2+c^2-2bc\cos A$, $48=16+c^2-2(4)c(-\frac{1}{2}) \implies c^2+4c-32=0$, 解得 $c=4$. 若条件 (2) 正确, $a^2=b^2+c^2+10b$. 对比余弦定理 $a^2=b^2+c^2+bc$, 得 $10b=bc$, 故 $c=10$. 若条件 (3) 正确, $S=\frac{1}{2}bc\sin A=15\sqrt{3} \implies \frac{1}{2}bc(\frac{\sqrt{3}}{2})=15\sqrt{3} \implies bc=60$.

组合判断. 若 (1)(2) 正确, 则 $c=4$ 与 $c=10$ 矛盾. 若 (1)(3) 正确, 由 (1) 知 $b=4,c=4$, 则 $bc=16$, 与 (3) 的 $bc=60$ 矛盾. 若 (2)(3) 正确, 由 (2) 知 $c=10$, 代入 (3) 得 $10b=60 \implies b=6$, 此组条件相容.

故正确的条件是 (2) 和 (3), 此时 $\triangle ABC$ 的三要素为 $A=\frac{2\pi}{3},b=6,c=10$.

欲求角平分线 AD 的长, 利用面积分割法 $S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle ACD}$

$$\frac{1}{2}bc\sin A=\frac{1}{2}c\cdot AD\cdot\sin\frac{A}{2}+\frac{1}{2}b\cdot AD\cdot\sin\frac{A}{2}$$

代入数值

$$15\sqrt{3}=\frac{1}{2}(10)\cdot AD\cdot\sin\frac{\pi}{3}+\frac{1}{2}(6)\cdot AD\cdot\sin\frac{\pi}{3}$$

$$15\sqrt{3}=(5\cdot AD+3\cdot AD)\cdot\frac{\sqrt{3}}{2}=8\cdot AD\cdot\frac{\sqrt{3}}{2}=4\sqrt{3}\cdot AD$$

解得 $AD=\frac{15}{4}$.

(1) $A=\frac{2\pi}{3}$.(2) 正确的条件为 (2) 和 (3), 角平分线 AD 的长为 $\frac{15}{4}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 12.10 在 $\triangle ABC$ 中, 面积 $S=\frac{\sqrt{3}}{4}(b^2+c^2-a^2)$.$AD$ 为 $\angle A$ 的平分线. (1) 证明 $AD=\frac{\sqrt{3}bc}{b+c}$;(2) 若 $BD=2DC$, $AD=4\sqrt{3}$, 求 $\triangle ABC$ 的周长.

解 本题是角平分线两大性质的深度综合应用.

对第 (1) 问, 由面积公式 $S=\frac{1}{2}bc\sin A$ 和余弦定理 $b^2+c^2-a^2=2bc\cos A$, 代入已知条件得 $\frac{1}{2}bc\sin A=\frac{\sqrt{3}}{4}(2bc\cos A)\implies \tan A=\sqrt{3}$, 故 $A=\frac{\pi}{3}$.

由 AD 是角平分线, 利用面积分割法 $S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle ACD}$

$$\frac{1}{2}bc\sin A=\frac{1}{2}c\cdot AD\cdot\sin\frac{A}{2}+\frac{1}{2}b\cdot AD\cdot\sin\frac{A}{2}$$

$$bc\sin A=(b+c)AD\sin\frac{A}{2}$$

利用二倍角公式 $\sin A=2\sin\frac{A}{2}\cos\frac{A}{2}$, 整理得 $AD=\frac{2bc}{b+c}\cos\frac{A}{2}$. 代入 $A=\frac{\pi}{3}$, 得 $\cos\frac{A}{2}=\cos\frac{\pi}{6}=\frac{\sqrt{3}}{2}$.

$$AD=\frac{2bc}{b+c}\left(\frac{\sqrt{3}}{2}\right)=\frac{\sqrt{3}bc}{b+c}$$

对第 (2) 问, 联立角平分线的两大性质建立方程组. 由角平分线定理及 $BD=2DC$, 有 $\frac{c}{b}=\frac{AB}{AC}=\frac{BD}{DC}=2$, 即 $c=2b$. 由第 (1) 问的结论及 $AD=4\sqrt{3}$, 有 $4\sqrt{3}=\frac{\sqrt{3}bc}{b+c}$, 即 $bc=4(b+c)$.

将 $c=2b$ 代入第二个方程: $b(2b)=4(b+2b)\implies 2b^2=12b$. 因 $b\neq 0$, 解得 $b=6$, 故 $c=12$.

利用余弦定理求边长 a: $a^2=b^2+c^2-2bc\cos A=6^2+12^2-2(6)(12)\cos\frac{\pi}{3}=36+144-72=108$. 故 $a=\sqrt{108}=6\sqrt{3}$.

$\triangle ABC$ 的周长为 $L=a+b+c=6\sqrt{3}+6+12=18+6\sqrt{3}$.

(1) 证明见上.(2) $\triangle ABC$ 的周长为 $18+6\sqrt{3}$.

12.2.2.3 外角平分线与调和分割

内角平分线定理给出了边长比例; 外角平分线定理把这个比例关系推广到对边的延长线上.

定理 12.13 (外角平分线定理) 三角形一个顶点处的外角平分线, 若与对边的延长线相交, 则交点到该对边两端点的距离之比, 等于夹此角的两邻边之比. 如图, 若 AD 为 $\triangle ABC$ 顶点 A 处的外角平分线, 交 BC 延长线于点 D(假设 $AB\neq AC$), 则

$$\frac{AB}{AC}=\frac{BD}{CD} \tag{12-12}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明

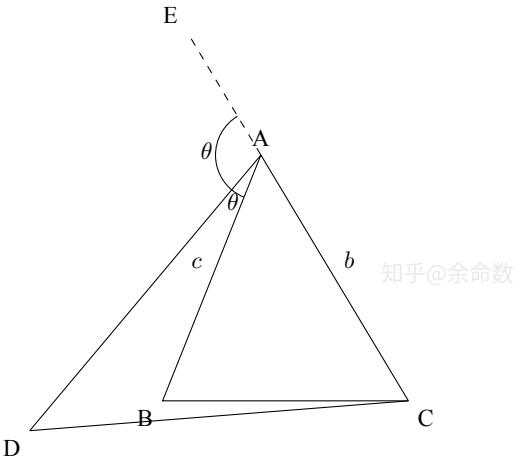


图 12.11　外角平分线定理证明示意

本定理的证明思路与内角平分线定理的证明高度一致, 均巧妙地运用了正弦定理. 考察 $\triangle ABD$ 与 $\triangle ACD$. 设直线 EAC 为一条直线, AD 平分外角 $\angle EAB$. 令 $\angle EAD = \angle DAB = \theta$.

在 $\triangle ABD$ 中, 应用正弦定理可得

$$\frac{BD}{\sin(\angle DAB)} = \frac{AB}{\sin(\angle BDA)} \quad \text{即} \quad \frac{BD}{\sin\theta} = \frac{c}{\sin(\angle BDA)}$$

在 $\triangle ACD$ 中, 其顶点 A 处的内角为 $\angle CAD$. 注意到 $\angle EAD$ 与 $\angle CAD$ 互为补角, 因为它们构成平角 $\angle EAC$. 故 $\angle CAD = \pi - \theta$. 因此 $\sin(\angle CAD) = \sin(\pi - \theta) = \sin\theta$. 应用正弦定理于 $\triangle ACD$:

$$\frac{CD}{\sin(\angle CAD)} = \frac{AC}{\sin(\angle CDA)} \quad \text{即} \quad \frac{CD}{\sin\theta} = \frac{b}{\sin(\angle CDA)}$$

接着, 需把这两个方程联系起来. 点 B, C, D 共线, 故 $\angle BDA$ 与 $\angle CDA$ 互为补角, 它们的正弦值相等.

$$\sin(\angle BDA) = \sin(\pi - \angle CDA) = \sin(\angle CDA)$$

从 $\triangle ABD$ 的关系式中, 有 $\sin(\angle BDA) = \frac{c\sin\theta}{BD}$. 从 $\triangle ACD$ 的关系式中, 有 $\sin(\angle CDA) = \frac{b\sin\theta}{CD}$.

令二者相等, 得

$$\frac{c\sin\theta}{BD} = \frac{b\sin\theta}{CD}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

由于 AD 是非退化三角形的外角平分线, θ 不为 0 或 π, 故 $\sin\theta \neq 0$. 两边可约去 $\sin\theta$, 得到

$$\frac{c}{BD} = \frac{b}{CD} \quad \text{即} \quad \frac{AB}{AC} = \frac{BD}{CD}$$

定理得证. □

调和分割

内角与外角平分线定理共同说明了一种基本的几何对称性. 对于 $\triangle ABC$ 顶点 A 的内、外角平分线, 它们分别交对边 BC (或其延长线) 于点 D_{in} 和 D_{ext}. 这两个点将线段 BC 以相同的比例进行**内分割**与**外分割**. 满足这种关系的点 D_{in}、D_{ext} 与线段端点 B、C 共同构成了一组**调和点列**. 同时, 一个顶点的内角平分线与其外角平分线总是相互垂直.

12.2.2.4 中线与阿波罗尼斯定理

中线连接三角形的一个顶点与其对边中点. 中线的长度与三边长度满足阿波罗尼斯定理.

定理 12.14 (中线长公式) 设 m_a 为 $\triangle ABC$ 中边 a 上的中线, 则其长度满足

$$4m_a^2 = 2(b^2 + c^2) - a^2 \tag{12-13}$$

等价地, 该公式亦可表示为 $m_a = \frac{1}{2}\sqrt{2(b^2 + c^2) - a^2}$.

证明 在两个由中线分割形成的相邻三角形中, 对互补角应用余弦定理, 即可消去未知角. 如图, 设 D 为 BC 的中点, 则 AD 为中线 m_a, 且 $BD = DC = \frac{a}{2}$.

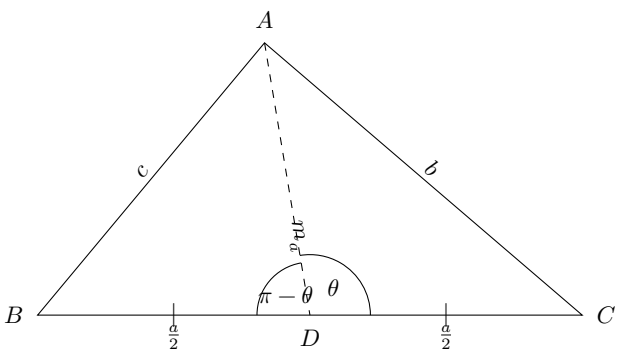


图 12.12 中线长公式的余弦定理证明

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在 $\triangle ACD$ 中, 设 $\angle ADC=\theta$. 由余弦定理得

$$b^2=m_a^2+\left(\frac{a}{2}\right)^2-2m_a\left(\frac{a}{2}\right)\cos\theta=m_a^2+\frac{a^2}{4}-am_a\cos\theta \quad \cdots\cdots(1)$$

在 $\triangle ABD$ 中, 角 $\angle ADB=\pi-\theta$, 故 $\cos(\angle ADB)=\cos(\pi-\theta)=-\cos\theta$. 再次应用余弦定理.

$$c^2=m_a^2+\left(\frac{a}{2}\right)^2-2m_a\left(\frac{a}{2}\right)\cos(\pi-\theta)=m_a^2+\frac{a^2}{4}+am_a\cos\theta \quad \cdots\cdots(2)$$

式 (1) 与 (2) 的结构仅在含 $\cos\theta$ 的项上符号相反, 因此两式相加即可消去该未知角.

$$b^2+c^2=\left(m_a^2+\frac{a^2}{4}-am_a\cos\theta\right)+\left(m_a^2+\frac{a^2}{4}+am_a\cos\theta\right)$$

$$b^2+c^2=2m_a^2+\frac{a^2}{2}$$

将上式两边同乘以 2, 即得定理的最终形式

$$2b^2+2c^2=4m_a^2+a^2 \Longrightarrow\ 4m_a^2=2(b^2+c^2)-a^2$$

定理得证. □

阿波罗尼斯定理

中线长公式是阿波罗尼斯定理的代数重述. 该定理说, 三角形任意两边长的平方和, 等于第三边上中线长的平方与该边一半长度的平方和的两倍, 即

$$b^2+c^2=2\left(m_a^2+\left(\frac{a}{2}\right)^2\right)$$

它建立了两边平方和与第三边上的中线长及该边一半长度的平方和之间的精确关系, 可直接用于处理中线与三边长度的问题.

12.2.3 面积问题

面积公式 $S=\frac{1}{2}ab\sin C$ 把边、角与面积联系起来. 已知其中两个量, 就可以求第三个量. 我们利用这个公式在面积、边长乘积与夹角正弦之间转换.

12.2.3.1 面积公式的两种用法

处理面积问题时, 我们常把面积公式当作边角互化的媒介.

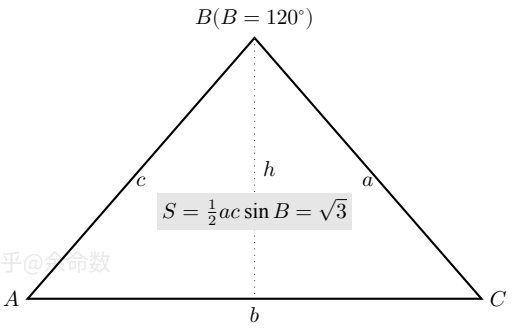
- 已知面积时, 可由 $ab=\frac{2S}{\sin C}$ 求两边之积, 或由 $\sin C=\frac{2S}{ab}$ 求夹角正弦.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 未知面积时, 可利用 $S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle ACD}$ 建立方程; 若条件中出现 S 与 $b^2+c^2-a^2$ 的关系, 注意到 $\frac{S}{2bc\cos A}=\frac{1}{4}\tan A$, 这常常是求角的信号.

例题 12.11 记 $\triangle ABC$ 的内角 A,B,C 的对边分别为 a,b,c, 已知 $\frac{a^2+c^2-b^2}{\cos B}=8$. (1) 求 ac; (2) 若 $b=2a\cos A\cos 2C+2c\cos C\cos 2A$, 求 $\triangle ABC$ 的面积.

解 第 (1) 问的表达式直接给出余弦定理的结构. 第 (2) 问的条件对称, 提示我们用正弦定理边化角, 再用三角恒等式找出角之间的联系.



知乎@余命数

对第 (1) 问, 注意到已知等式的分子 $a^2+c^2-b^2$ 是余弦定理中用以表达 $\cos B$ 的关键部分. 根据余弦定理, 有 $a^2+c^2-b^2=2ac\cos B$. 将此关系代入已知条件 $\frac{a^2+c^2-b^2}{\cos B}=8$ 中.

知乎@余命数 知乎@余命数

$$\frac{2ac\cos B}{\cos B}=8$$

由于 B 是三角形内角, 若 $\cos B=0$, 则 $B=\frac{\pi}{2}$, 但此时原式分母为零, 故 $B\neq\frac{\pi}{2}$, 因此 $\cos B\neq 0$, 可以约去.

知乎@余命数 知乎@余命数

$$2ac=8\implies ac=4$$

对第 (2) 问, 条件 $b=2a\cos A\cos 2C+2c\cos C\cos 2A$ 的结构对称, 是边化角的典型信号. 利用正弦定理 $a=2R\sin A, b=2R\sin B, c=2R\sin C$, 代入并约去 $2R$ 得

$$\sin B=2\sin A\cos A\cos 2C+2\sin C\cos C\cos 2A$$

注意到 $2\sin x\cos x=\sin(2x)$, 上式可化为.

$$\sin B=\sin(2A)\cos 2C+\sin(2C)\cos(2A)$$

右侧恰好是正弦的和角公式 $\sin(X+Y)=\sin X\cos Y+\cos X\sin Y$ 的形式.

知乎@余命数

$$\sin B=\sin(2A+2C)=\sin(2(A+C))$$

知乎@余命数

在 $\triangle ABC$ 中, $A+C=\pi-B$, 故 $2(A+C)=2\pi-2B$.

$$\sin B=\sin(2\pi-2B)=-\sin(2B)$$

再次使用二倍角公式 $\sin(2B)=2\sin B\cos B$:

$$\sin B=-2\sin B\cos B$$

因为 $B\in(0,\pi)$, $\sin B\neq 0$, 两边同除以 $\sin B$, 得

$$1=-2\cos B\implies \cos B=-\frac{1}{2}$$

故 $B=\frac{2\pi}{3}$.

接着, 计算三角形的面积.

$$S_{\triangle ABC}=\frac{1}{2}ac\sin B=\frac{1}{2}(4)\sin\frac{2\pi}{3}=2\left(\frac{\sqrt{3}}{2}\right)=\sqrt{3}$$

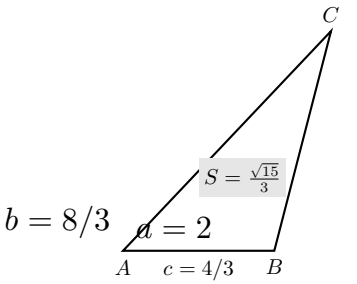
(1) $ac=4$.(2) $\triangle ABC$ 的面积为 $\sqrt{3}$.

例题 12.12　在 $\triangle ABC$ 中, 角 A,B,C 所对的边分别为 a,b,c. 从下列三个条件中选择两个, 使得 $\triangle ABC$ 存在, 并回答下列问题: (1) $b+c=2a$ (2) $A=\frac{2\pi}{3}$ (3) $3c\sin B=4a\sin C$.

(i) 求 $\cos B$ 的值;

(ii) 当 $a=2$ 时, 求 $\triangle ABC$ 的面积.

解　本题要求从三个条件中选出相容的两个. 我们先把每个条件化为最简洁的代数关系, 再通过计算检验组合.



首先, 对三个备选条件进行代数化简. 条件 (1) $b+c=2a$ 是关于三边和的线性关系. 条件 (2) $A=\frac{2\pi}{3}$ 确定了一个钝角. 条件 (3) $3c\sin B=4a\sin C$, 利用正弦定理进行角化边, 由 $\frac{b}{\sin B}=\frac{c}{\sin C}\implies \sin B=\frac{b}{c}\sin C$, 代入得 $3c\left(\frac{b}{c}\sin C\right)=4a\sin C$. 由于 $\sin C\neq 0$, 约去得 $3b=4a$, 这是一个关于两边的比例关系.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

接着组合判断. 若选择 (1) 和 (2), 即 $b+c=2a$ 且 $A=\frac{2\pi}{3}$. 由余弦定理 $a^2=b^2+c^2-2bc\cos(\frac{2\pi}{3})=b^2+c^2+bc$. 又 $(b+c)^2=(2a)^2 \implies b^2+c^2+2bc=4a^2 \implies b^2+c^2=4a^2-2bc$. 代入余弦定理式.$a^2=(4a^2-2bc)+bc \implies bc=3a^2$. 即 b,c 是关于 x 的方程 $x^2-(2a)x+3a^2=0$ 的两根. 其判别式 $\Delta=(-2a)^2-4(1)(3a^2)=-8a^2<0$, 方程无实数解, 故此组合不成立.

若选择 (2) 和 (3), 即 $A=\frac{2\pi}{3}$ 且 $3b=4a$. 由正弦定理 $\frac{a}{\sin A}=\frac{b}{\sin B}$, 可得 $\frac{a}{\sin(2\pi/3)}=\frac{4a/3}{\sin B}$, 约去 a 得 $\frac{2}{\sqrt{3}}=\frac{4}{3\sin B}$, 解得 $\sin B=\frac{2\sqrt{3}}{3}>1$, 此为不可能事件, 故此组合不成立.

因此, 唯一成立的组合是 (1) 和 (3), 即 $b+c=2a$ 且 $3b=4a$.

对 (i), 求 $\cos B$. 由 $3b=4a$ 得 $b=\frac{4}{3}a$. 代入 $b+c=2a$ 得 $\frac{4}{3}a+c=2a \implies c=\frac{2}{3}a$. 三边之比为 $a:b:c=a:\frac{4}{3}a:\frac{2}{3}a=3:4:2$. 应用余弦定理.

$$\cos B=\frac{a^2+c^2-b^2}{2ac}=\frac{(3k)^2+(2k)^2-(4k)^2}{2(3k)(2k)}=\frac{9-16+4}{12}=-\frac{3}{12}=-\frac{1}{4}$$

对 (ii), 求面积. 当 $a=2$ 时, 由三边比例关系可知 $c=\frac{2}{3}a=\frac{4}{3}$. 要求面积, 需要 $\sin B$. 由 $\cos B=-\frac{1}{4}$, 且 $B\in(0,\pi)$, 可得

$$\sin B=\sqrt{1-\cos^2 B}=\sqrt{1-\left(-\frac{1}{4}\right)^2}=\frac{\sqrt{15}}{4}$$

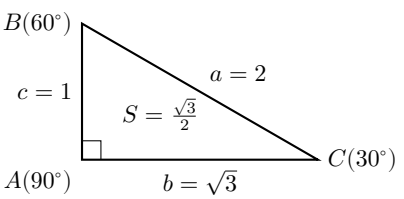
计算面积, 得

$$S_{\triangle ABC}=\frac{1}{2}ac\sin B=\frac{1}{2}(2)\left(\frac{4}{3}\right)\left(\frac{\sqrt{15}}{4}\right)=\frac{\sqrt{15}}{3}$$

选择条件 (1) 和 (3).(i) $\cos B=-\frac{1}{4}$.(ii) $\triangle ABC$ 的面积为 $\frac{\sqrt{15}}{3}$.

例题 12.13 (2023 年 3 月金丽衢十二校联考第 18 题) 在 $\triangle ABC$ 中, 角 A,B,C 的对边分别为 a,b,c. 已知 $(1-\sin C)(1-\cos 2B)=\sin 2B\cos C$, 且 $a=2,c=1$. (1) 证明 $C=2B-\frac{\pi}{2}$; (2) 求 $\triangle ABC$ 的面积.

解 本题需要化简一个复杂的三角恒等式, 找出角 C 与角 B 的关系, 再结合边长条件确定三角形的全部元素.



知乎@余命数

对第 (1) 问, 由已知等式 $(1-\sin C)(1-\cos 2B)=\sin 2B\cos C$ 出发. 将等式展开, 并将含有 $\sin C$ 和 $\cos C$ 的项移到同一边, 这通常是使用和角公式的前兆.

$$\begin{aligned}1-\cos 2B&=\sin C(1-\cos 2B)+\sin 2B\cos C\\&=\sin C-\sin C\cos 2B+\sin 2B\cos C\\&=\sin C+(\sin 2B\cos C-\cos 2B\sin C)\\&=\sin C+\sin(2B-C).\end{aligned}$$

左侧利用二倍角公式 $1-\cos 2B=2\sin^2 B$ 进行转化. 右侧是两个正弦之和, 使用和差化积公式

$$\sin x+\sin y=2\sin\frac{x+y}{2}\cos\frac{x-y}{2}$$

来化简, 得

$$\begin{aligned}\sin C+\sin(2B-C)&=2\sin\frac{C+2B-C}{2}\cos\frac{C-(2B-C)}{2}\\&=2\sin B\cos(C-B).\end{aligned}$$

于是, 原式等价于 $2\sin^2 B=2\sin B\cos(C-B)$. 由于 $B\in(0,\pi)$, $\sin B\neq 0$, 可约去 $2\sin B$, 得到 $\sin B=\cos(C-B)$. 利用诱导公式 $\sin B=\cos(\frac{\pi}{2}-B)$, 方程变为 $\cos(\frac{\pi}{2}-B)=\cos(C-B)$. 令 $u=\frac{\pi}{2}-B, v=C-B$. 因 $u\in(-\frac{\pi}{2},\frac{\pi}{2}), v\in(-\pi,\pi)$, 由 $\cos u=\cos v$ 的通解 $v=\pm u+2k\pi$ 可知此处只能有 $k=0$, 故 $\frac{\pi}{2}-B=\pm(C-B)$. 若 $\frac{\pi}{2}-B=C-B$, 则 $C=\frac{\pi}{2}$. 由正弦定理 $\frac{a}{\sin A}=\frac{c}{\sin C}$ 得 $\frac{2}{\sin A}=\frac{1}{\sin(\pi/2)}=1$, 故 $\sin A=2$, 此为不可能事件. 因此, 必然是 $\frac{\pi}{2}-B=-(C-B)=B-C$, 即 $C=2B-\frac{\pi}{2}$.

对第 (2) 问, 求三角形面积. 由 (1) 知 $C=2B-\frac{\pi}{2}$, 故 $\sin C=\sin(2B-\frac{\pi}{2})=-\cos(2B)$. 同时, $A=\pi-B-C=\pi-B-(2B-\frac{\pi}{2})=\frac{3\pi}{2}-3B$, 故 $\sin A=\sin(\frac{3\pi}{2}-3B)=-\cos(3B)$. 根据正弦定理 $\frac{a}{\sin A}=\frac{c}{\sin C}$, 有 $\frac{2}{-\cos(3B)}=\frac{1}{-\cos(2B)}$, 即 $2\cos(2B)=\cos(3B)$. 利用二倍角和三倍角公式 $2(2\cos^2 B-1)=4\cos^3 B-3\cos B$, 整理得 $4\cos^3 B-4\cos^2 B-3\cos B+2=0$. 这是一个关于 $\cos B$ 的三次方程,通过因式分解可得 $(2\cos B-1)(2\cos^2 B-\cos B-2)=0$. 解得 $\cos B=\frac{1}{2}$ 或 $\cos B=\frac{1\pm\sqrt{17}}{4}$.

此时, 必须检验解的有效性. 三角形的几何存在性要求所有内角为正. $A>0\implies\frac{3\pi}{2}-3B>0\implies B<\frac{\pi}{2}$. $C>0\implies 2B-\frac{\pi}{2}>0\implies B>\frac{\pi}{4}$. 故角 B 的取值范围是 $(\frac{\pi}{4},\frac{\pi}{2})$, 那么 $\cos B\in(0,\frac{\sqrt{2}}{2})$. $\cos B=\frac{1}{2}$ 满足此范围, 故 $B=\frac{\pi}{3}$ 是有效解. $\cos B=\frac{1+\sqrt{17}}{4}>1$, 舍去.$\cos B=\frac{1-\sqrt{17}}{4}<0$, 舍去.

因此, 得唯一的解 $B=\frac{\pi}{3}$. 最后计算面积, 得

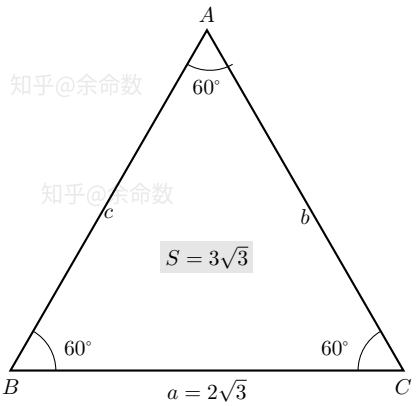
$$S_{\triangle ABC}=\frac{1}{2}ac\sin B=\frac{1}{2}(2)(1)\sin\frac{\pi}{3}=\frac{\sqrt{3}}{2}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(1) 证明见上.(2) $\triangle ABC$ 的面积为 $\frac{\sqrt{3}}{2}$.

例题 12.14 (2023 年 1 月桐庐中学高三第一学期期末考试第 17 题) 在 $\triangle ABC$ 中, 角 A, B, C 的对边分别为 a, b, c, 且 $b^2 - \frac{2\sqrt{3}}{3}bc\sin A + c^2 = a^2$. (I) 求角 A; (II) 若 $a = 2\sqrt{3}$, 且 $\frac{2a}{\tan A} = \frac{b}{\tan B} + \frac{c}{\tan C}$, 求 $\triangle ABC$ 的面积.

解 第 (I) 问把面积公式与余弦定理结合起来. 第 (II) 问把正切化为正弦、余弦, 再用正弦、余弦定理化简, 确定三角形的形状.



对于第 (I) 问, 将已知条件 $b^2 - \frac{2\sqrt{3}}{3}bc\sin A + c^2 = a^2$ 变形为 $b^2 + c^2 - a^2 = \frac{2\sqrt{3}}{3}bc\sin A$. 注意到左侧是余弦定理的关键部分, 即 $b^2 + c^2 - a^2 = 2bc\cos A$. 代入后得到 $2bc\cos A = \frac{2\sqrt{3}}{3}bc\sin A$. 由于 b, c 是边长, 不为 0, 可约去 $2bc$, 得到 $\cos A = \frac{\sqrt{3}}{3}\sin A$. 在 $\triangle ABC$ 中, $A \in (0, \pi)$, $\cos A$ 与 $\sin A$ 不可能同时为零, 故 $\tan A = \sqrt{3}$, 解得 $A = \frac{\pi}{3}$.

对于第 (II) 问, 处理第二个条件 $\frac{2a}{\tan A} = \frac{b}{\tan B} + \frac{c}{\tan C}$. 将正切化为正弦与余弦的比值.

$$\frac{2a\cos A}{\sin A} = \frac{b\cos B}{\sin B} + \frac{c\cos C}{\sin C}$$

此式结构强烈暗示使用正弦定理 $\frac{x}{\sin X} = 2R$. 代入得 $2\cdot(2R\cos A) = 2R\cos B + 2R\cos C$. 约去公因式 $2R$, 得到一个纯粹的角的关系.$2\cos A = \cos B + \cos C$. 代入已求出的 $A = \frac{\pi}{3}$, 即 $\cos A = 1/2$, 得 $1 = \cos B + \cos C$. 利用内角和定理 $B + C = \frac{2\pi}{3}$, 将 C 用 B 表示.

$$\cos B + \cos\left(\frac{2\pi}{3} - B\right) = 1$$

展开得 $\cos B + \cos\frac{2\pi}{3}\cos B + \sin\frac{2\pi}{3}\sin B = 1$, 即 $\frac{1}{2}\cos B + \frac{\sqrt{3}}{2}\sin B = 1$. 利用辅助角公式 $\sin(B + \frac{\pi}{6}) = 1$. 由于 $B \in (0, 2\pi/3)$, $B + \pi/6 \in (\pi/6, 5\pi/6)$, 故 $B + \frac{\pi}{6} = \frac{\pi}{2}$, 解得 $B = \frac{\pi}{3}$. 因此, $C = \frac{2\pi}{3} - B = \frac{\pi}{3}$, $\triangle ABC$ 是等边三角形.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

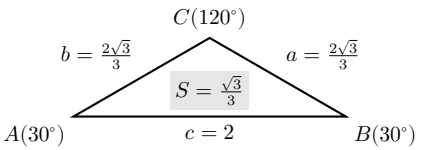
最后, 计算面积. 已知 $a=2\sqrt{3}$, 根据等边三角形面积公式 $S=\frac{\sqrt{3}}{4}\times(\text{边长})^2$:

$$S_{\triangle ABC}=\frac{\sqrt{3}}{4}a^2=\frac{\sqrt{3}}{4}(2\sqrt{3})^2=3\sqrt{3}$$

(I) $A=\frac{\pi}{3}$.(II) $\triangle ABC$ 的面积为 $3\sqrt{3}$.

例题 12.15 (2023 年 3 月苏锡常镇一模试卷第 18 题) 在 $\triangle ABC$ 中,角 A,B,C 的对边分别为 a,b,c. 已知 $1+\sin 2A=(3\tan B+2)\cos 2A$. (1) 若 $C=\frac{3\pi}{4}$, 求 $\tan B$ 的值;(2) 若 $A=B$, $c=2$, 求 $\triangle ABC$ 的面积.

解 本题的三角恒等式混合了 $\sin 2A$、$\cos 2A$、$\tan B$. 我们把它化为关于 $\tan A$ 和 $\tan B$ 的代数关系.



首先, 化简题设给出的关系式 $1+\sin 2A=(3\tan B+2)\cos 2A$. 在本题两问的条件下, 均可排除 $\cos 2A=0$ 的情形. 若 $\cos 2A=0$, 则 $A=\frac{\pi}{4}$ 或 $A=\frac{3\pi}{4}$; 其中 $A=\frac{\pi}{4}$ 代回原式不成立, 而 $A=\frac{3\pi}{4}$ 又与第 (1) 问的 $A+B=\frac{\pi}{4}$ 及第 (2) 问的 $A=B$ 均不相容. 因而两边同除以 $\cos 2A$ 得 $\sec 2A+\tan 2A=3\tan B+2$. 利用恒等式 $\sec x+\tan x=\frac{1+\sin x}{\cos x}=\frac{(\sin(x/2)+\cos(x/2))^2}{\cos^2(x/2)-\sin^2(x/2)}=\frac{\cos(x/2)+\sin(x/2)}{\cos(x/2)-\sin(x/2)}=\tan(x/2+\pi/4)$, 可得 $\tan(A+\pi/4)=3\tan B+2$. 这是一个更简洁的关系.

对第 (1) 问, 当 $C=\frac{3\pi}{4}$ 时, $A+B=\frac{\pi}{4}$. $\tan(A+\pi/4)=\tan(\pi/2-B)=\cot B=1/\tan B$. 代入关系式得 $1/\tan B=3\tan B+2$. 令 $t=\tan B$, 则 $1/t=3t+2$, 即 $3t^2+2t-1=0$. 解得 $t=\frac{1}{3}$ 或 $t=-1$. 因为 $B\in(0,\pi/4)$, $\tan B>0$, 故 $\tan B=\frac{1}{3}$.

对第 (2) 问, 当 $A=B$ 时, $\tan A=\tan B$. 代入 $\tan(A+\pi/4)=3\tan B+2$. 令 $t=\tan A=\tan B$,则 $\frac{\tan A+\tan(\pi/4)}{1-\tan A\tan(\pi/4)}=\frac{t+1}{1-t}=3t+2$. $t+1=(3t+2)(1-t)=-3t^2+t+2$, 整理得 $3t^2=1$, 故 $t^2=1/3$. 由于 A,B 是三角形内角且 $A+B<\pi$, A,B 必为锐角, 故 $\tan A>0$. 所以 $\tan A=\tan B=\frac{1}{\sqrt{3}}$, 解得 $A=B=\frac{\pi}{6}$. 角 $C=\pi-A-B=\frac{2\pi}{3}$.

为求面积, 利用正弦定理求边长 a,b. $\frac{a}{\sin A}=\frac{c}{\sin C}\implies a=\frac{c\sin A}{\sin C}=\frac{2\sin(\pi/6)}{\sin(2\pi/3)}=\frac{2(1/2)}{\sqrt{3}/2}=\frac{2}{\sqrt{3}}$. 由于 $A=B$, 所以 $a=b=\frac{2\sqrt{3}}{3}$.

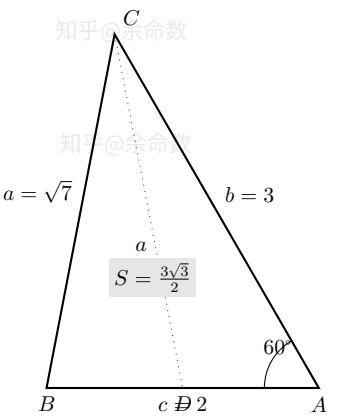
最后计算面积: $S_{\triangle ABC}=\frac{1}{2}ab\sin C=\frac{1}{2}\left(\frac{2\sqrt{3}}{3}\right)^2\sin\frac{2\pi}{3}=\frac{1}{2}\cdot\frac{4}{3}\cdot\frac{\sqrt{3}}{2}=\frac{\sqrt{3}}{3}$.

(1) $\tan B=\frac{1}{3}$.(2) $\triangle ABC$ 的面积为 $\frac{\sqrt{3}}{3}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 12.16 (2023 年 2 月深圳市高三一调第 18 题) 记 $\triangle ABC$ 的内角 A, B, C 的对边分别为 a, b, c, 已知 $b + c = 2a \sin\left(C + \frac{\pi}{6}\right)$. (1) 求 A; (2) 设 AB 的中点为 D, 若 $CD = a$, 且 $b - c = 1$, 求 $\triangle ABC$ 的面积.

解 第 (1) 问是边角互化. 第 (2) 问中 CD 是中线, 我们对 $\triangle ABC$ 与 $\triangle ACD$ 分别使用余弦定理建立方程.



对第 (1) 问, 利用和角公式展开已知条件并进行边化角:

$$b + c = 2a\left(\sin C \cos\frac{\pi}{6} + \cos C \sin\frac{\pi}{6}\right) = \sqrt{3}\, a \sin C + a \cos C.$$

两边同除以 $2R$, 得

$$\sin B + \sin C = \sqrt{3} \sin A \sin C + \sin A \cos C.$$

利用 $\sin B = \sin(A + C) = \sin A \cos C + \cos A \sin C$ 代入, 化简得

$$\cos A \sin C + \sin C = \sqrt{3} \sin A \sin C.$$

约去 $\sin C$ 得 $\cos A + 1 = \sqrt{3} \sin A$, 即

$$\sqrt{3} \sin A - \cos A = 1.$$

利用辅助角公式得 $2\sin(A - \frac{\pi}{6}) = 1$. 由于 $A \in (0, \pi)$, 故 $A - \frac{\pi}{6} \in (-\frac{\pi}{6}, \frac{5\pi}{6})$, 从而 $A - \frac{\pi}{6} = \frac{\pi}{6}$ 或 $\frac{5\pi}{6}$. 后者给出 $A = \pi$, 不符合三角形内角条件, 故 $A = \frac{\pi}{3}$.

对第 (2) 问, 采用一外一内、双余弦定理的思路. 对 $\triangle ABC$ 应用余弦定理.

$$a^2 = b^2 + c^2 - 2bc \cos A = b^2 + c^2 - bc.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对 $\triangle ACD$ 应用余弦定理, 其中 $AD=c/2$、$CD=a$、$AC=b$.

$$CD^2=AC^2+AD^2-2(AC)(AD)\cos A$$

$$\Longrightarrow\ a^2=b^2+\left(\frac{c}{2}\right)^2-2\cdot b\cdot\frac{c}{2}\cdot\cos\frac{\pi}{3}=b^2+\frac{c^2}{4}-\frac{bc}{2}.$$

联立两个关于 a^2 的表达式.$b^2+c^2-bc=b^2+\frac{c^2}{4}-\frac{bc}{2}$, 化简得 $\frac{3}{4}c^2=\frac{1}{2}bc$, 即 $3c=2b$. 结合题目给出的 $b-c=1$, 解方程组得 $c=2, b=3$.

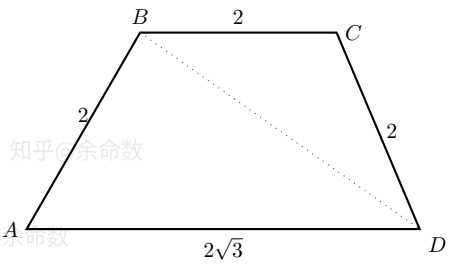
最后计算面积, 得

$$S_{\triangle ABC}=\frac{1}{2}bc\sin A=\frac{1}{2}(3)(2)\sin\frac{\pi}{3}=\frac{3\sqrt{3}}{2}$$

(1) $A=\frac{\pi}{3}$.(2) $\triangle ABC$ 的面积为 $\frac{3\sqrt{3}}{2}$.

例题 12.17 (2023 年湖南长沙中学高三月考 (六) 第 18 题) 如图, 在平面四边形 $ABCD$ 中, $AB=BC=CD=2, AD=2\sqrt{3}$. (1) 若 DB 平分 $\angle ADC$, 证明 $A+C=\pi$; (2) 记 $\triangle ABD$ 与 $\triangle BCD$ 的面积分别为 S_1 和 S_2, 求 $S_1^2+S_2^2$ 的最大值.

解 我们借助公共边 BD, 在 $\triangle ABD$ 与 $\triangle BCD$ 中用余弦定理建立角 A 与角 C 的关系.



首先, 在 $\triangle ABD$ 和 $\triangle BCD$ 中对公共边 BD 应用余弦定理. $BD^2=2^2+(2\sqrt{3})^2-2(2)(2\sqrt{3})\cos A=16-8\sqrt{3}\cos A$. $BD^2=2^2+2^2-2(2)(2)\cos C=8-8\cos C$. 联立两式,得到关于角 A 和角 C 的恒等关系.$16-8\sqrt{3}\cos A=8-8\cos C$, 即 $\sqrt{3}\cos A-\cos C=1\quad(*)$.

对第 (1) 问, 若 DB 平分 $\angle ADC$, 则 $\angle ADB=\angle CDB$. 在 $\triangle ABD$ 和 $\triangle BCD$ 中应用正弦定理, $\frac{AB}{\sin\angle ADB}=\frac{BD}{\sin A}$ 和 $\frac{BC}{\sin\angle CDB}=\frac{BD}{\sin C}$. 由于 $AB=BC=2$ 且 $\angle ADB=\angle CDB$, 可得 $\sin A=\sin C$. 故 $A=C$ 或 $A+C=\pi$. 若 $A=C$, 代入 $(*)$ 式得 $(\sqrt{3}-1)\cos A=1\ \Longrightarrow\ \cos A=\frac{\sqrt{3}+1}{2}>1$, 不可能. 因此, 唯一可能的情况是 $A+C=\pi$.

对第 (2) 问,目标函数为 $S_1^2+S_2^2=(2\sqrt{3}\sin A)^2+(2\sin C)^2=12\sin^2 A+4\sin^2 C$. 为利用关系式 $(*)$,将函数用余弦表示.$S_1^2+S_2^2=12(1-\cos^2 A)+4(1-\cos^2 C)=16-$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$12\cos^2 A-4\cos^2 C$. 由 $(*)$ 得 $\cos C=\sqrt{3}\cos A-1$. 令 $x=\cos A$,代入上式,将目标函数化为关于 x 的一元二次函数. $f(x)=16-12x^2-4(\sqrt{3}x-1)^2=-24x^2+8\sqrt{3}x+12$.

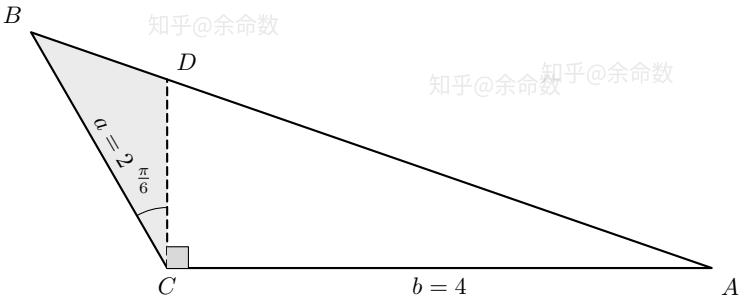
接下来确定自变量 $x=\cos A$ 的取值范围.A,C 是三角形内角,必须满足 $-1<\cos A<1$ 和 $-1<\cos C<1$. $-1<x<1$ 且 $-1<\sqrt{3}x-1<1 \implies 0<\sqrt{3}x<2 \implies 0<x<\frac{2}{\sqrt{3}}$. 取交集得 x 的定义域为 $(0,1)$.

最后,求二次函数 $f(x)=-24x^2+8\sqrt{3}x+12$ 在区间 $(0,1)$ 上的最大值. 该函数是开口向下的抛物线,对称轴为 $x=-\frac{8\sqrt{3}}{2(-24)}=\frac{\sqrt{3}}{6}$. 由于 $0<\frac{\sqrt{3}}{6}<1$,函数在对称轴处取得最大值. $f_{\max}=f(\frac{\sqrt{3}}{6})=-24(\frac{3}{36})+8\sqrt{3}(\frac{\sqrt{3}}{6})+12=-2+4+12=14$.

(1) 证明见上.(2) $S_1^2+S_2^2$ 的最大值为 14.

例题 12.18 (2024 湖南邵阳模拟) 在 $\triangle ABC$ 中,角 A,B,C 的对边分别为 a,b,c,且 $\triangle ABC$ 的周长为 $\frac{a\sin B}{\sin A+\sin B-\sin C}$. (1) 求 C;(2) 若 $a=2,b=4$,D 为边 AB 上一点,$\angle BCD=\frac{\pi}{6}$,求 $\triangle BCD$ 的面积.

解 第 (1) 问用正弦定理把边角关系式化为边长关系. 第 (2) 问用面积分割法,并注意到 $\angle ACD$ 是直角.



对第 (1) 问,已知周长 $a+b+c=\frac{a\sin B}{\sin A+\sin B-\sin C}$. 利用正弦定理,$\sin X=\frac{x}{2R}$,等式右边可化为.

$$\frac{a\cdot\frac{b}{2R}}{\frac{a}{2R}+\frac{b}{2R}-\frac{c}{2R}}=\frac{ab}{a+b-c}$$

于是得纯边的关系式 $a+b+c=\frac{ab}{a+b-c}$. 交叉相乘得 $(a+b+c)(a+b-c)=ab$,即 $(a+b)^2-c^2=ab$. 展开为 $a^2+2ab+b^2-c^2=ab$,整理得 $a^2+b^2-c^2=-ab$. 根据余弦定理,$\cos C=\frac{a^2+b^2-c^2}{2ab}=\frac{-ab}{2ab}=-\frac{1}{2}$. 因为 $C\in(0,\pi)$,所以 $C=\frac{2\pi}{3}$.

对第 (2) 问,现求 $\triangle BCD$ 的面积. 已知 $BC=a=2$ 和 $\angle BCD=\pi/6$,任务转化为求线段 CD 的长度. 注意到一个隐藏条件.

$$\angle ACD=\angle C-\angle BCD=\frac{2\pi}{3}-\frac{\pi}{6}=\frac{\pi}{2}$$

这说明 $\triangle ACD$ 是一个直角三角形.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

用面积分割法.

$$S_{\triangle ABC}=S_{\triangle ACD}+S_{\triangle BCD}.$$

分别计算三块面积, 得

$$\begin{aligned}
S_{\triangle ABC}&=\frac{1}{2}ab\sin C=\frac{1}{2}\cdot 2\cdot 4\cdot\sin\frac{2\pi}{3}=2\sqrt{3},\\
S_{\triangle ACD}&=\frac{1}{2}AC\cdot CD\cdot\sin(\angle ACD)=\frac{1}{2}\cdot 4\cdot CD\cdot\sin\frac{\pi}{2}=2\,CD,\\
S_{\triangle BCD}&=\frac{1}{2}BC\cdot CD\cdot\sin(\angle BCD)=\frac{1}{2}\cdot 2\cdot CD\cdot\sin\frac{\pi}{6}=\frac{1}{2}CD.
\end{aligned}$$

根据面积关系建立方程.

$$2\sqrt{3}=2\,CD+\frac{1}{2}CD=\frac{5}{2}CD,$$

解得 $CD=\frac{4\sqrt{3}}{5}$.

最终目标面积:

$$S_{\triangle BCD}=\frac{1}{2}CD=\frac{1}{2}\cdot\frac{4\sqrt{3}}{5}=\frac{2\sqrt{3}}{5}$$

(1) $C=\frac{2\pi}{3}$.(2) $\triangle BCD$ 的面积为 $\frac{2\sqrt{3}}{5}$.

12.2.4 内切圆与外接圆综合问题

外接圆半径 R 和内切圆半径 r 把三角形的角、边、面积与周长联系起来. 我们借助它们在边、角、面积、周长之间切换, 建立所需的方程或函数.

12.2.4.1 基本工具

12.2.4.1.1 外接圆 外接圆的半径 R 是一个全局性的度量, 规定了整个三角形的尺寸上限.

- **正弦定理**, 把边角与 R 联系起来的法则.

$$\frac{a}{\sin A}=\frac{b}{\sin B}=\frac{c}{\sin C}=2R$$

- **面积公式**.$S=\frac{abc}{4R}$ 实质上是正弦定理与 $S=\frac{1}{2}ab\sin C$ 结合的产物.
- **应用**. 见到 R 或外接圆, 即用正弦定理, 可:
  1. 由一边一角求 R.
  2. 由 R 和一个角求其对边.
  3. 由 R 和一条边求其对角的正弦.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

12.2.4.1.2 内切圆 内切圆半径 r 与三角形的周长和面积紧密相连.

- **面积公式**, 与 r 密切相关.

$$S = pr = \frac{1}{2}(a+b+c)r$$

- **切线长性质**. 从一个顶点到内切圆两个切点的距离相等, 且等于**半周长减去该顶点的对边**. 例如, 从顶点 A 出发的切线长为 $p-a$.
- **衍生公式**. 由此可推导出连接角、半径和边长的实用公式

$$\tan\frac{A}{2} = \frac{r}{p-a}$$

- **应用**. 见到 r 或内切圆, 即用 $S=pr$ 和切线长性质, 可:
  1. 在面积、周长、半径三者中知二求一.
  2. 利用切线长性质建立关于边长的方程.

例题 12.19 (2023 年 2 月七彩联盟返校联考, 18) 如图, 平面四边形 $ABCD$ 中, $AD=5, CD=3, \angle ADC=120^\circ$. $\triangle ABC$ 的内角 A, B, C 的对边分别为 a, b, c, 且满足 $\frac{a+b}{c} = \frac{\sin A-\sin C}{\sin A-\sin B}$.
(1) 求四边形 $ABCD$ 的外接圆半径 R;
(2) 求 $\triangle ABC$ 内切圆半径 r 的取值范围.

解 第 (1) 问化简比例式求出关键角, 再用正弦定理求外接圆半径 R. 第 (2) 问把 r 表示为某角的函数, 在三角形内角约束下求其定义域, 从而得到 r 的范围.

对于第 (1) 问.

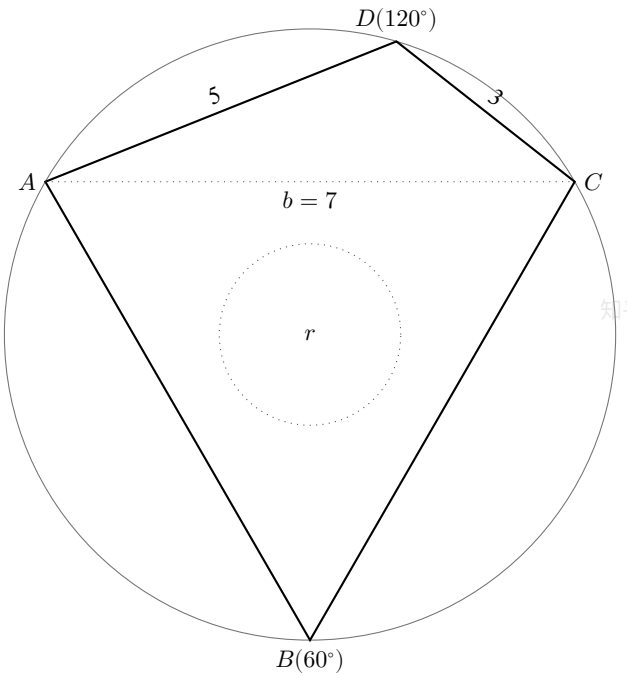
目标是求外接圆半径 R, 由正弦定理 $\frac{x}{\sin X} = 2R$, 须先求出 $\triangle ABC$ 中某条边及其对角. 题目的已知条件是一个复杂的比例式, 这是突破口, 先化简它以求出某个角.

这是一个边角混合的比例式, 利用正弦定理边化角是标准操作. 令 $a = k\sin A, b = k\sin B, c = k\sin C$ 代入 $\frac{a+b}{c} = \frac{\sin A-\sin C}{\sin A-\sin B}$:

$$\frac{k(\sin A+\sin B)}{k\sin C} = \frac{\sin A-\sin C}{\sin A-\sin B}$$

交叉相乘得 $(\sin A+\sin B)(\sin A-\sin B) = \sin C(\sin A-\sin C)$, 即 $\sin^2 A-\sin^2 B = \sin C\sin A-\sin^2 C$.

知乎@余命数



知乎@余命数

得到纯角的方程后仍嫌复杂, 故再次角化边, 转为可用余弦定理解的边长关系. 再次利用正弦定理的变形 $k^2\sin^2 X = x^2$:

$$a^2 - b^2 = ca - c^2 \implies a^2 + c^2 - b^2 = ac$$

观察左侧, 它与余弦定理的分子形式完全一致.

$$\cos B = \frac{a^2 + c^2 - b^2}{2ac} = \frac{ac}{2ac} = \frac{1}{2}$$

故 $B = 60^\circ$.

角 B 已求出, 由正弦定理只需再求其对边 b, 即 AC 即可得 R. 边 AC 恰好是 $\triangle ADC$ 的一条边, 而 $\triangle ADC$ 信息完整(两边一夹角), 可用余弦定理求出 AC.

既然四边形 $ABCD$ 有外接圆, 说明 A, B, C, D 四点共圆. 即 $\triangle ABC$ 和 $\triangle ADC$ 共享同一个外接圆. 可在 $\triangle ADC$ 中利用余弦定理求出公共边 AC 的长度.

$$\begin{aligned} AC^2 &= AD^2 + DC^2 - 2(AD)(DC)\cos(\angle ADC) \\ &= 5^2 + 3^2 - 2(5)(3)\cos 120^\circ \\ &= 25 + 9 - 30(-\frac{1}{2}) = 34 + 15 = 49 \end{aligned}$$

所以 $AC = 7$. 在 $\triangle ABC$ 中, 边 AC 就是边 b.

知乎@余命数

接着, 在 $\triangle ABC$ 中, 有边 $b=7$ 和其对角 $B=60^\circ$. 应用正弦定理求外接圆半径 R:

$$2R=\frac{b}{\sin B}=\frac{7}{\sin 60^\circ}=\frac{7}{\sqrt{3}/2}=\frac{14}{\sqrt{3}}\implies R=\frac{7\sqrt{3}}{3}$$

对于第 (2) 问.

目标是求内切圆半径 r 的取值范围. 由两个公式 $S=pr$ 和 $r=(p-b)\tan(B/2)$ 出发,后者直接把 r 与边角联系起来. 已知 $B=60^\circ$ 和 $b=7$,该公式化为 $r=\frac{a+c-7}{2\sqrt{3}}$, 问题转化为求 $a+c$ 的取值范围.

利用公式 $r=(p-b)\tan\frac{B}{2}$:

$$r=\left(\frac{a+c-b}{2}\right)\tan 30^\circ=\frac{a+c-7}{2\sqrt{3}}$$

问题转化为求 $a+c$ 的取值范围.

求 $a+c$ 的范围.a,c 是变量, 但其所在的 $\triangle ABC$ 已确定 (外接圆半径 R 和角 B 已定). 用正弦定理把 a,c 用角 A,C 表示, 再利用 $A+C$ 为定值, 转化为关于单角 A 的三角函数值域问题.

在 $\triangle ABC$ 中, 利用正弦定理 $a=2R\sin A, c=2R\sin C$:

$$a+c=2R(\sin A+\sin C)=\frac{14\sqrt{3}}{3}(\sin A+\sin C)$$

利用内角和定理 $A+C=\pi-B=\frac{2\pi}{3}$, 即 $C=\frac{2\pi}{3}-A$.

$$\begin{aligned}\sin A+\sin C&=\sin A+\sin\left(\frac{2\pi}{3}-A\right)\\&=\sin A+\sin\frac{2\pi}{3}\cos A-\cos\frac{2\pi}{3}\sin A\\&=\sin A+\frac{\sqrt{3}}{2}\cos A+\frac{1}{2}\sin A\\&=\frac{3}{2}\sin A+\frac{\sqrt{3}}{2}\cos A\\&=\sqrt{3}\left(\frac{\sqrt{3}}{2}\sin A+\frac{1}{2}\cos A\right)\\&=\sqrt{3}\sin\left(A+\frac{\pi}{6}\right)\end{aligned}$$

所以 $a+c=\frac{14\sqrt{3}}{3}\cdot\sqrt{3}\sin\left(A+\frac{\pi}{6}\right)=14\sin\left(A+\frac{\pi}{6}\right)$.

接着求 $\sin(A+\pi/6)$ 的范围,须精确确定角 A 的定义域. 三角形内角必须为正, 这是最基本的约束.

在 $\triangle ABC$ 中, $A>0$ 且 $C>0$. $C=\frac{2\pi}{3}-A>0\implies A<\frac{2\pi}{3}$. 所以 $A\in(0,\frac{2\pi}{3})$. 因此 $A+\frac{\pi}{6}\in(\frac{\pi}{6},\frac{5\pi}{6})$. 在此区间内, $\sin(A+\frac{\pi}{6})$ 的取值范围是 $(\sin\frac{\pi}{6},1]$, 即 $(\frac{1}{2},1]$.

知乎@余命数

于是 $a+c=14\sin(A+\frac{\pi}{6})$ 的取值范围是 $(14\cdot\frac{1}{2},14\cdot 1]$, 即 $(7,14]$. 代入 r 的表达式 $r=\frac{a+c-7}{2\sqrt{3}}$: r 的取值范围是 $\left(\frac{7-7}{2\sqrt{3}},\frac{14-7}{2\sqrt{3}}\right]$, 即 $\left(0,\frac{7}{2\sqrt{3}}\right]=\left(0,\frac{7\sqrt{3}}{6}\right]$.

(1) $R=\frac{7\sqrt{3}}{3}$.(2) r 的取值范围是 $\left(0,\frac{7\sqrt{3}}{6}\right]$.

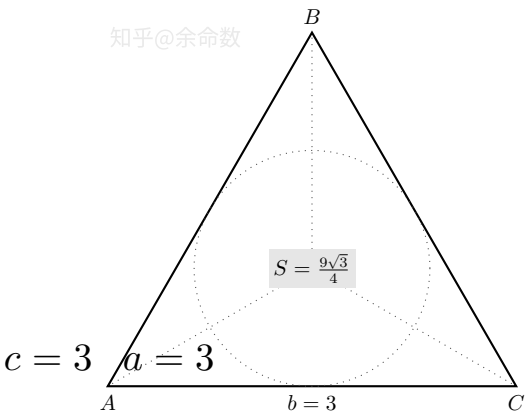
例题 12.20 (2023 年 8 月河北名校高三开学考第 18 题) 已知 $\triangle ABC$ 的内角 A,B,C 的对边分别为 a,b,c, 且 $b\cos C+c\cos B=2\sqrt{3}\sin A$.

(1) 求 $\triangle ABC$ 外接圆的面积;

(2) 记 $\triangle ABC$ 内切圆的半径为 r, 若 $B=\frac{\pi}{3},r=\frac{\sqrt{3}}{2}$, 求 $\triangle ABC$ 的面积.

知乎@余命数

解 第 (1) 问用射影定理化简边角关系, 再用正弦定理求 R. 第 (2) 问由 r 与 B 反推边长关系, 再用 $S=pr$ 与 $S=\frac{1}{2}ac\sin B$ 求面积.



(1) 由射影定理, $b\cos C+c\cos B=a$, 故已知式化为 $a=2\sqrt{3}\sin A$, 即

$$\frac{a}{\sin A}=2\sqrt{3}.$$

再由正弦定理 $\frac{a}{\sin A}=2R$, 得 $2R=2\sqrt{3}$, 故 $R=\sqrt{3}$. 外接圆面积为

$$S_{\text{外接圆}}=\pi R^2=3\pi.$$

(2) 已知 $B=\frac{\pi}{3},r=\frac{\sqrt{3}}{2}$ 以及 (1) 中的 $R=\sqrt{3}$. 先求 b:

$$b=2R\sin B=2\sqrt{3}\sin\frac{\pi}{3}=3.$$

再由 $r=(p-b)\tan\frac{B}{2}$,

$$\frac{\sqrt{3}}{2}=\left(\frac{a+c-b}{2}\right)\tan\frac{\pi}{6}=\frac{a+c-3}{2\sqrt{3}},$$

知乎@余命数

化简得 $a+c=6$. 于是半周长 $p=\frac{(a+c)+b}{2}=\frac{9}{2}$, 面积

$$S_{\triangle ABC}=pr=\frac{9}{2}\cdot\frac{\sqrt{3}}{2}=\frac{9\sqrt{3}}{4}.$$

法二采用 $S=\frac{1}{2}ac\sin B$ 作验证. 余弦定理改写为

$$b^2=(a+c)^2-2ac-2ac\cos B.$$

代入 $b=3$、$a+c=6$、$B=\frac{\pi}{3}$,

$$9=36-2ac-2ac\cdot\frac{1}{2}=36-3ac,$$

得 $ac=9$, 于是

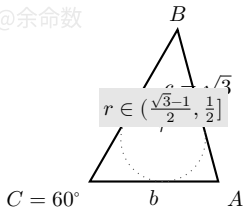
$$S_{\triangle ABC}=\frac{1}{2}\cdot 9\cdot\sin\frac{\pi}{3}=\frac{9\sqrt{3}}{4},$$

与前法一致.

(1) 外接圆面积为 3π.(2) $\triangle ABC$ 的面积为 $\frac{9\sqrt{3}}{4}$.

例题 12.21 (2023 杭州二中高三年级 4 月月考第 17 题) 在锐角 $\triangle ABC$ 中, 角 A,B,C 所对的边分别为 a,b,c, 已知 $\sqrt{3}\tan A\tan B=\tan A+\tan B+\tan C$.
(1) 求角 C 的大小;
(2) 若 $c=\sqrt{3}$, 求 $\triangle ABC$ 内切圆半径的取值范围.

解 我们先把 r 化为单角函数. 由正切恒等式求出角 C, 再用 $r=4R\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2}$ 与内角和消去变量, 最后在锐角约束的定义域上求值域.



(1) 任意非直角 $\triangle ABC$ 中有恒等式

$$\tan A+\tan B+\tan C=\tan A\tan B\tan C.$$

证明 (恒等式证明) 由 $A+B+C=\pi$ 得 $A+B=\pi-C$, 从而

$$\tan(A+B)=\tan(\pi-C)\implies\frac{\tan A+\tan B}{1-\tan A\tan B}=-\tan C.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

展开得

$$\tan A+\tan B=-\tan C(1-\tan A\tan B)=-\tan C+\tan A\tan B\tan C.$$

整理即得 $\tan A+\tan B+\tan C=\tan A\tan B\tan C$. □

将恒等式代入 $\sqrt{3}\tan A\tan B=\tan A+\tan B+\tan C$, 得

$$\sqrt{3}\tan A\tan B=\tan A\tan B\tan C.$$

知乎@余命数

锐角三角形中 $A,B\in(0,\pi/2)$, 故 $\tan A\tan B\neq 0$, 两边同除后 $\tan C=\sqrt{3}$. 又 C 为锐角, 得 $C=\frac{\pi}{3}$.

(2) 采用公式 $r=4R\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2}$. 由正弦定理与 $c=\sqrt{3}$、$C=\frac{\pi}{3}$,

$$2R=\frac{c}{\sin C}=\frac{\sqrt{3}}{\sin(\pi/3)}=\frac{\sqrt{3}}{\sqrt{3}/2}=2,$$

故 $R=1$. 再代入 $C=\frac{\pi}{3}$,

$$r=4\sin\frac{A}{2}\sin\frac{B}{2}\cdot\frac{1}{2}=2\sin\frac{A}{2}\sin\frac{B}{2}.$$

知乎@余命数 知乎@余命数

由 $A+B=\frac{2\pi}{3}$ 得 $B=\frac{2\pi}{3}-A$, 于是

$$r(A)=2\sin\frac{A}{2}\sin\left(\frac{\pi}{3}-\frac{A}{2}\right).$$

知乎@余命数

锐角约束给出

$$\begin{cases}A\in(0,\pi/2),\\ B=\frac{2\pi}{3}-A\in(0,\pi/2)\implies \frac{\pi}{6}<A<\frac{2\pi}{3},\\ C=\frac{\pi}{3}\in(0,\pi/2),\end{cases}$$

知乎@余命数

取交集得 $A\in(\frac{\pi}{6},\frac{\pi}{2})$.

令 $x=\frac{A}{2}$, 则 $x\in(\frac{\pi}{12},\frac{\pi}{4})$, 且

$$r(x)=2\sin x\sin(\frac{\pi}{3}-x).$$

积化和差 $2\sin\alpha\sin\beta=\cos(\alpha-\beta)-\cos(\alpha+\beta)$ 给出

$$r(x)=\cos(2x-\frac{\pi}{3})-\cos\frac{\pi}{3}=\cos(2x-\frac{\pi}{3})-\frac{1}{2}.$$

知乎@余命数

由 $x \in (\frac{\pi}{12}, \frac{\pi}{4})$ 得 $2x - \frac{\pi}{3} \in (-\frac{\pi}{6}, \frac{\pi}{6})$, 在此区间上

$$\cos(2x - \frac{\pi}{3}) \in (\frac{\sqrt{3}}{2}, 1],$$

故

$$r \in \Big(\frac{\sqrt{3}}{2} - \frac{1}{2}, 1 - \frac{1}{2}\Big] = \Big(\frac{\sqrt{3}-1}{2}, \frac{1}{2}\Big].$$

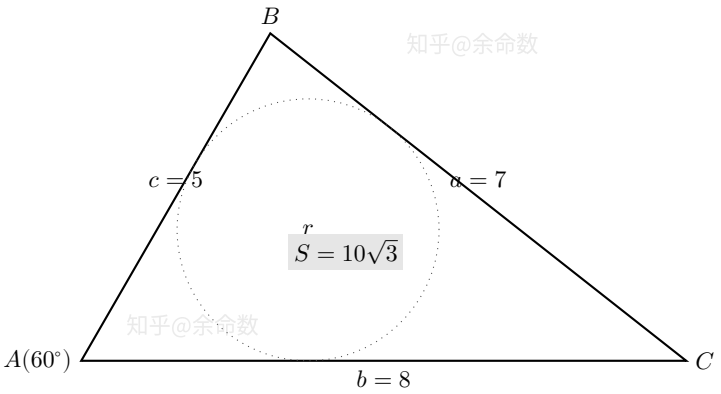
(1) $C = \frac{\pi}{3}$.(2) $\triangle ABC$ 内切圆半径的取值范围为 $\Big(\frac{\sqrt{3}-1}{2}, \frac{1}{2}\Big]$.

例题 12.22 (2023-2024 年武汉九月调研) 设 $\triangle ABC$ 的内角 A, B, C 所对的边分别为 a, b, c, 且 $2a\cos B = 2c - b$.

(1.) 求角 A;

(2.) 若 $a = 7$, 且 $\triangle ABC$ 的内切圆半径 $r = \sqrt{3}$, 求 $\triangle ABC$ 的面积 S.

解 第 (1) 问用余弦定理把 $\cos B$ 化为边长关系, 再反过来求角 A. 第 (2) 问已知 r 与角 A, 由 $r = (p-a)\tan(A/2)$ 求出半周长 p, 再用 $S = pr$ 求面积.



对于第 (1) 问.

条件 $2a\cos B = 2c - b$ 混合了边与角, 先由余弦定理把 $\cos B$ 用边长表示, 化为纯边关系.

根据余弦定理, $\cos B = \frac{a^2+c^2-b^2}{2ac}$. 代入已知等式:

$$2a \cdot \frac{a^2 + c^2 - b^2}{2ac} = 2c - b$$

化简得

$$\frac{a^2 + c^2 - b^2}{c} = 2c - b$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两边同乘 c, 得 $a^2+c^2-b^2=c(2c-b)=2c^2-bc$.

由此得到一个纯边长关系式. 把它整理出 $b^2+c^2-a^2$ 这一项, 即可套用余弦定理求角 A.

整理上述等式:

$$a^2+c^2-b^2=2c^2-bc$$

$$b^2+c^2-a^2=bc$$

由余弦定理, $\cos A=\frac{b^2+c^2-a^2}{2bc}$. 代入已得的关系:

$$\cos A=\frac{bc}{2bc}=\frac{1}{2}$$

因为 $A\in(0,\pi)$, 所以 $A=\frac{\pi}{3}$.

对于第 (2) 问.

已知 $a=7$, $r=\sqrt{3}$, 且由第 (1) 问得 $A=\pi/3$. 涉及内切圆半径与已知角时, 公式 $r=(p-a)\tan(A/2)$ 直接给出半周长 p, 再由 $S=pr$ 得面积.

利用公式 $r=(p-a)\tan\frac{A}{2}$, 其中 p 是半周长. 将已知值代入, 得

$$\sqrt{3}=(p-7)\tan\left(\frac{\pi/3}{2}\right)=(p-7)\tan\frac{\pi}{6}$$

即:

$$\sqrt{3}=(p-7)\cdot\frac{1}{\sqrt{3}}$$

交叉相乘得 $\sqrt{3}\cdot\sqrt{3}=p-7$, 即 $3=p-7$. 解得半周长 $p=10$.

最后, 利用三角形面积与内切圆半径的关系式 $S=pr$ 求面积:

$$S=10\cdot\sqrt{3}=10\sqrt{3}$$

(1) $A=\frac{\pi}{3}$.(2) $\triangle ABC$ 的面积为 $10\sqrt{3}$.

12.2.5 多三角形问题

若一个三角形被内部线段分成多个小三角形, 我们就得到**多三角形**问题. 这些小三角形之间的公共边、公共角、互补角以及线段和差, 就是联系它们的关键.

- **公共边**: 两个三角形共享一条边, 这条边的长度把二者联系起来.
- **公共角**: 两个三角形共享一个角, 这个角的三角函数值是等价的.
- **互补角**: 分割线在一条边上产生的两个邻角, 如 $\angle ADB$ 和 $\angle ADC$ 互为补角, 由此 $\sin\angle ADB=\sin\angle ADC$ 且 $\cos\angle ADB=-\cos\angle ADC$.
- **线段和差**: 被分割的边长等于各分段之和, 如 $AB=AD+DB$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

下列的例题展示如何利用这些公共量在多个三角形之间建立联系并求解.

练习 12.1 2023 武汉二调 在 $\triangle ABC$ 中, $AB=2$, D 为 AB 中点, $CD=\sqrt{2}$.
(1) 若 $BC=\sqrt{2}$, 求 AC 的长;
(2) 若 $\angle BAC=2\angle BCD$, 求 AC 的长.

练习 12.2 2023 福建质检 在 $\triangle ABC$ 中, D 为边 BC 上一点, 且 $\angle CAD=\angle B$.
(1) 证明 $AC^2=BC\cdot CD$;
(2) 若 $\cos C=\frac{3}{4}$、$AD=2\sqrt{2}$、$CD=2$, 求 $\triangle ABD$ 的面积.

12.3 保底的建系法

在处理几何构型时, 正弦与余弦定理提供连接边角关系的代数工具. 然而, 当问题的几何条件非常规, 或图形本身缺乏明显特征时, 纯粹的几何方法或三角变换可能变得很复杂. 这时我们就把点、线放进平面直角坐标系, 给每个关键点配上坐标. 长度、角度、位置关系都变成代数表达式与方程, 几何推理就变成了代数运算. 这条路有时更长, 但按部就班, 逻辑上是确定的. 我们称这种方法为**建系法**.

建系法的第一步, 是为图形选一个尽量简化后续计算的坐标系. 我们通常按下面三条原则选取.

- **原点特殊化**: 尽可能将图形的特殊点, 如顶点、中点、垂足置于坐标原点.
- **轴线特殊化**: 尽可能将图形的特殊直线, 如对称轴、某条边所在的直线与坐标轴重合.
- **利用对称性**: 若图形具有对称性, 应使坐标系与之相适应, 从而使点的坐标呈现简洁的对称形式.

坐标系建立之后, 我们就把几何条件写成代数方程, 再解方程(组)即得解答.

例题 12.23 用坐标法证明阿波罗尼斯定理(中线长公式). 在 $\triangle ABC$ 中, 设 m_a 为边 a 上的中线, 则 $b^2+c^2=2\left(m_a^2+\left(\frac{a}{2}\right)^2\right)$.

解 为使坐标表示尽可能简洁, 将中线 m_a 所对的边 BC 置于 x 轴上, 并取该边的中点 D 作为坐标原点.

设 $D(0,0)$, 则 $B(-\frac{a}{2},0)$, $C(\frac{a}{2},0)$. 设顶点 A 的坐标为 (x_0,y_0).

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

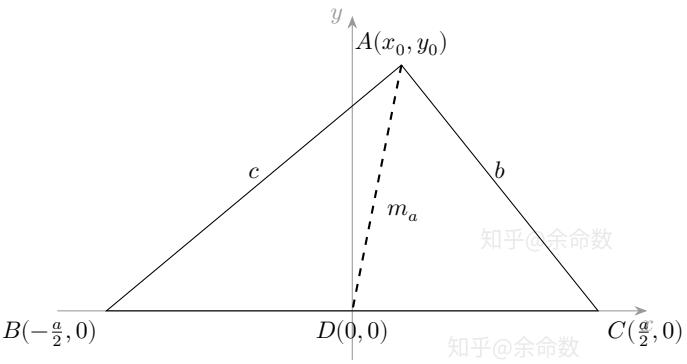


图 12.13　坐标法证明阿波罗尼斯定理

接着, 由两点间距离公式将三条边长 b,c 以及中线长 m_a 用坐标表示出来.

$$b^2=|AC|^2=\left(x_0-\frac{a}{2}\right)^2+y_0^2=x_0^2-ax_0+\frac{a^2}{4}+y_0^2 \qquad \text{(距离公式)}$$

$$c^2=|AB|^2=\left(x_0+\frac{a}{2}\right)^2+y_0^2=x_0^2+ax_0+\frac{a^2}{4}+y_0^2 \qquad \text{(距离公式)}$$

$$m_a^2=|AD|^2=x_0^2+y_0^2 \qquad (D \text{ 为原点})$$

需建立 b^2,c^2,m_a^2 与 a^2 之间的关系. 观察 b^2 与 c^2 的表达式, 它们仅有一次项符号不同. 将两式相加, 可以消去 x_0 的一次项.

$$\begin{aligned} b^2+c^2&=\left(x_0^2-ax_0+\frac{a^2}{4}+y_0^2\right)+\left(x_0^2+ax_0+\frac{a^2}{4}+y_0^2\right)\\ &=2x_0^2+2y_0^2+2\left(\frac{a^2}{4}\right)\\ &=2(x_0^2+y_0^2)+\frac{a^2}{2} \end{aligned}$$

而 $x_0^2+y_0^2=m_a^2$, 且 $\frac{a^2}{2}=2\left(\frac{a}{2}\right)^2$. 代入上式, 得

$$b^2+c^2=2m_a^2+2\left(\frac{a}{2}\right)^2=2\left(m_a^2+\left(\frac{a}{2}\right)^2\right).$$

定理得证.

例题 12.24 (2023 届湖北高三圆创三月联合测评第 18 题) 在 $\triangle ABC$ 中, D 是边 BC 上的一点, $\angle CAD=\frac{\pi}{4}$, $\frac{AB}{AC}=\frac{\sqrt{2}BD}{CD}$.
(1) 求 $\angle BAD$;
(2) 若 $AB=AD=2$, 求 $\triangle ABC$ 的面积.

解 题目给出一个含四条线段的比例式. 若直接建立坐标并用距离公式计算, 会很繁琐. 我们借助坐标系重新整理几何关系, 特别是面积关系, 把问题代数化.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对第 (1) 问, 先建立一个坐标系; 同一个三角形的面积既可用坐标 (底乘高) 计算, 也可用两边夹角的三角公式计算. 把两种表达式相等, 就得到关于未知角的方程.

令点 D 为坐标原点 $(0,0)$, 并使边 BC 所在的直线与 x 轴重合. 设顶点 A 的坐标为 (x_A,h), 其中 $h>0$ 为 $\triangle ABC$ 中 BC 边上的高.

知乎@余命数

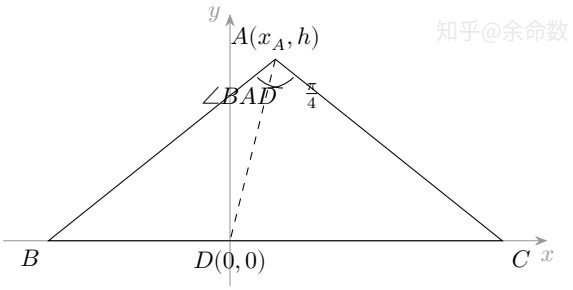


图 12.14 利用坐标系分析面积关系

在 $\triangle ACD$ 中, 面积有两种算法. 一方面, 基于坐标, 底为 CD, 高为 h, 故 $S_{\triangle ACD}=\frac{1}{2}\cdot CD\cdot h$. 另一方面,利用三角面积公式,$S_{\triangle ACD}=\frac{1}{2}AC\cdot AD\cdot\sin(\angle CAD)$. 联立两式, 得 $CD\cdot h=AC\cdot AD\cdot\sin(\angle CAD)$. 整理后得到一个关于边长比例的表达式

$$\frac{AC}{CD}=\frac{h}{AD\cdot\sin(\angle CAD)}$$

同理, 在 $\triangle ABD$ 中, 面积也有两种算法. $S_{\triangle ABD}=\frac{1}{2}\cdot BD\cdot h=\frac{1}{2}AB\cdot AD\cdot\sin(\angle BAD)$. 整理后得到另一个比例表达式

$$\frac{AB}{BD}=\frac{h}{AD\cdot\sin(\angle BAD)}$$

接着, 将题目给出的比例式 $\frac{AB}{AC}=\frac{\sqrt{2}BD}{CD}$ 变形, 使其结构与已得的比例式相匹配

$$\frac{AB}{BD}\cdot\frac{CD}{AC}=\sqrt{2}$$

将由面积关系得到的两个比例表达式代入上式, 得

$$\left(\frac{h}{AD\cdot\sin(\angle BAD)}\right)\cdot\left(\frac{AD\cdot\sin(\angle CAD)}{h}\right)=\sqrt{2}$$

变量 h 和 AD 被完全消去,这正是坐标法带来的简化. 于是得到一个纯粹的角的关系式

$$\frac{\sin(\angle CAD)}{\sin(\angle BAD)}=\sqrt{2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入已知条件 $\angle CAD=\frac{\pi}{4}$, 即 $\sin(\angle CAD)=\frac{\sqrt{2}}{2}$

$$\frac{\sqrt{2}/2}{\sin(\angle BAD)}=\sqrt{2}$$

解得 $\sin(\angle BAD)=\frac{1}{2}$. 又 $\angle BAD+\angle CAD=\angle BAC<\pi$, 且 $\angle CAD=\frac{\pi}{4}$, 故 $\angle BAD<\frac{3\pi}{4}$, 排除 $\frac{5\pi}{6}$. 因此 $\angle BAD=\frac{\pi}{6}$.

对第 (2) 问, 由第 (1) 问可知 $\angle BAD=\frac{\pi}{6}$. 在 $\triangle ABD$ 中, 已知 $AB=AD=2$, 故该三角形为等腰三角形. 因此, $\angle B=\angle ADB=\frac{1}{2}(\pi-\angle BAD)=\frac{1}{2}(\pi-\frac{\pi}{6})=\frac{5\pi}{12}$.

于是可确定 $\triangle ABC$ 的所有内角.

$$A=\angle BAD+\angle CAD=\frac{\pi}{6}+\frac{\pi}{4}=\frac{5\pi}{12}$$

$$C=\pi-A-B=\pi-\frac{5\pi}{12}-\frac{5\pi}{12}=\frac{2\pi}{12}=\frac{\pi}{6}$$

至此, 已知了 $\triangle ABC$ 的一条边 $c=AB=2$ 以及所有角. 为求面积 $S=\frac{1}{2}bc\sin A$, 还需计算边长 $b=AC$. 在 $\triangle ABC$ 中应用正弦定理, 得

$$\frac{AC}{\sin B}=\frac{AB}{\sin C}$$

$$\frac{b}{\sin(5\pi/12)}=\frac{2}{\sin(\pi/6)}$$

已知 $\sin(\pi/6)=1/2$, 以及 $\sin(5\pi/12)=\sin(75^\circ)=\sin(45^\circ+30^\circ)=\sin 45^\circ\cos 30^\circ+\cos 45^\circ\sin 30^\circ=\frac{\sqrt{2}}{2}\frac{\sqrt{3}}{2}+\frac{\sqrt{2}}{2}\frac{1}{2}=\frac{\sqrt{6}+\sqrt{2}}{4}$.

代入求解 b

$$b=\frac{2\cdot\sin(5\pi/12)}{1/2}=4\sin(5\pi/12)=4\cdot\frac{\sqrt{6}+\sqrt{2}}{4}=\sqrt{6}+\sqrt{2}$$

最后, 计算 $\triangle ABC$ 的面积, 得

$$\begin{aligned}S_{\triangle ABC}&=\frac{1}{2}\cdot AB\cdot AC\cdot\sin A\\&=\frac{1}{2}\cdot c\cdot b\cdot\sin A\\&=\frac{1}{2}\cdot 2\cdot(\sqrt{6}+\sqrt{2})\cdot\sin\frac{5\pi}{12}\\&=(\sqrt{6}+\sqrt{2})\cdot\frac{\sqrt{6}+\sqrt{2}}{4}\\&=\frac{(\sqrt{6}+\sqrt{2})^2}{4}=\frac{6+2+2\sqrt{12}}{4}=\frac{8+4\sqrt{3}}{4}=2+\sqrt{3}\end{aligned}$$

(1) $\angle BAD=\frac{\pi}{6}$.(2) $\triangle ABC$ 的面积为 $2+\sqrt{3}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 12.25 (浙江省十校联盟 2023 届高三第三次联考第 18 题) 如图, 在 $\triangle ABC$ 中, D 为边 BC 上一点, $DC=3, AD=5, AC=7, \angle DAC=\angle ABC$.
(1) 求 $\angle ADC$ 的大小;
(2) 求 $\triangle ABC$ 的面积.

解 对第 (1) 问, 所求角位于 $\triangle ADC$ 中, 而该三角形的三边长均已知, 故可直接应用余弦定理. 在 $\triangle ADC$ 中,

$$\begin{aligned}\cos(\angle ADC) &= \frac{AD^2+DC^2-AC^2}{2(AD)(DC)} \\ &= \frac{5^2+3^2-7^2}{2(5)(3)} = \frac{25+9-49}{30} = \frac{-15}{30} = -\frac{1}{2}\end{aligned}$$

由于 $\angle ADC$ 为三角形内角, 其范围为 $(0,\pi)$, 故 $\angle ADC=\frac{2\pi}{3}$.

对第 (2) 问, 我们运用坐标法, 把几何关系转化为代数方程求解.

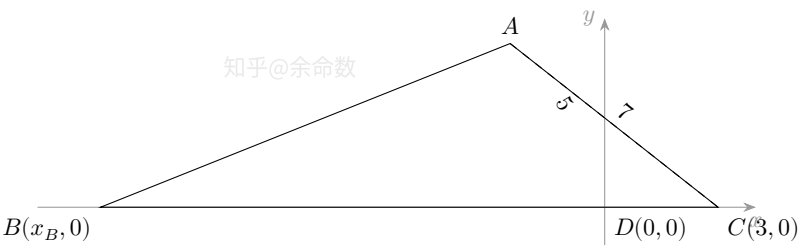


图 12.15 坐标法求解 $\triangle ABC$

建立坐标系. 以点 D 为坐标原点 $(0,0)$, 并使边 BC 所在的直线与 x 轴重合. 根据已知条件 $DC=3$, 可得点 C 的坐标为 $(3,0)$.

点 A 的坐标可以由其与 D,C 的距离关系确定. 设 $A(x_A,y_A)$, 则有方程组

$$\begin{cases}|AD|^2 = x_A^2+y_A^2 = 5^2 = 25 \\ |AC|^2 = (x_A-3)^2+y_A^2 = 7^2 = 49\end{cases}$$

将第一式代入第二式展开后得 $x_A^2-6x_A+9+y_A^2=49 \implies 25-6x_A+9=49$, 解得 $x_A=-5/2$. 进而求得 $y_A^2=25-(-5/2)^2=75/4$. 不妨取 $y_A>0$, 则 A 的坐标为 $\left(-\frac{5}{2},\frac{5\sqrt{3}}{2}\right)$.

点 B 在 x 轴上, 其位置未知, 设其坐标为 $(x_B,0)$. 由于 D 在 B,C 之间, 可知 $x_B<0$.

接着, 将题设的几何条件 $\angle DAC=\angle ABC$ 转化为代数等式. 由于两角均为三角形内角, 且 $\angle DAC$ 为锐角, 故此条件等价于 $\cos(\angle DAC)=\cos(\angle ABC)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在 $\triangle ADC$ 中，利用余弦定理可求得

$$\cos(\angle DAC)=\frac{AD^2+AC^2-DC^2}{2(AD)(AC)}=\frac{5^2+7^2-3^2}{2(5)(7)}=\frac{65}{70}=\frac{13}{14}$$

对于 $\angle ABC$，利用坐标与向量点积求其大小. $\overrightarrow{BA}=(A_x-B_x,A_y-B_y)=(-5/2-x_B,5\sqrt{3}/2)$. $\overrightarrow{BC}=(C_x-B_x,C_y-B_y)=(3-x_B,0)$.

$\cos(\angle ABC)=\frac{\overrightarrow{BA}\cdot\overrightarrow{BC}}{|\overrightarrow{BA}||\overrightarrow{BC}|}=\frac{(-5/2-x_B)(3-x_B)}{|\overrightarrow{BA}|\cdot(3-x_B)}=\frac{-5/2-x_B}{|\overrightarrow{BA}|}$.

令 $u=-5/2-x_B$. 由于 $x_B<-5/2$（否则 B 角为钝角），故 $u>0$. $|\overrightarrow{BA}|^2=u^2+(5\sqrt{3}/2)^2=u^2+75/4$.

于是，$\cos(\angle ABC)=\frac{u}{\sqrt{u^2+75/4}}$.

由等角条件建立关于 u 的方程

$$\frac{13}{14}=\frac{u}{\sqrt{u^2+75/4}}$$

两边平方得 $\frac{169}{196}=\frac{u^2}{u^2+75/4}$，整理得 $169(u^2+75/4)=196u^2$. $27u^2=169\cdot 75/4\implies u^2=\frac{169\cdot 25}{36}$，解得 $u=\frac{13\cdot 5}{6}=\frac{65}{6}$.

由 $u=-5/2-x_B=65/6$，解得 $x_B=-5/2-65/6=-15/6-65/6=-80/6=-40/3$. 点 B 的坐标为 $(-40/3,0)$.

$\triangle ABC$ 的底边 BC 长度为 $|BC|=|x_C-x_B|=|3-(-40/3)|=49/3$. 其对应的高为点 A 的 y 坐标，即 $h=\frac{5\sqrt{3}}{2}$.

因此，$\triangle ABC$ 的面积为

$$S=\frac{1}{2}\cdot|BC|\cdot h=\frac{1}{2}\cdot\frac{49}{3}\cdot\frac{5\sqrt{3}}{2}=\frac{245\sqrt{3}}{12}$$

(1) $\angle ADC=\frac{2\pi}{3}$.(2) $\triangle ABC$ 的面积为 $\frac{245\sqrt{3}}{12}$.

阅读提示

建系法把几何的直观与代数的严谨结合起来.

- **优点**：思路清晰，程序化，降低了对几何直觉和构造性技巧的依赖. 对于几何关系复杂、难以直接识别突破口的题目，建系法往往能通过系统性的计算得出结论.
- **缺点**：若坐标系选择不当，或几何图形本身不含特殊角度、特殊位置关系，可能会导致计算量急剧增大，从而引入计算风险.

当常规的几何方法或三角变换难以推进时，建系法通过为图形赋予代数结构，把几何思辨转入系统的代数运算轨道.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

12.4 非常规题型

前面几节用正弦、余弦定理和面积公式处理了许多常规三角形问题. 本节处理一些不那么直接的问题. 我们依旧依赖同样的定理, 只是需要更灵活地组合它们, 并引入一些代数技巧.

例题 12.26 已知 $\triangle ABC$ 中, $c^2=\sqrt{2}ab$.
(1) 求 $\cos C$ 的最小值;
(2) 证明: $C-A\leq\frac{\pi}{6}$.

解 (1) 求解 $\cos C$ 的最小值, 连接边角关系最直接的是余弦定理. 根据余弦定理,

$$\cos C=\frac{a^2+b^2-c^2}{2ab}$$

代入已知条件 $c^2=\sqrt{2}ab$ 消去变量 c

$$\cos C=\frac{a^2+b^2-\sqrt{2}ab}{2ab}$$

为分析其最小值, 分离表达式.

$$\cos C=\frac{a^2+b^2}{2ab}-\frac{\sqrt{2}ab}{2ab}=\frac{a^2+b^2}{2ab}-\frac{\sqrt{2}}{2}$$

$\frac{\sqrt{2}}{2}$ 是常数, 因此 $\cos C$ 的最小值取决于 $\frac{a^2+b^2}{2ab}$ 的最小值.

根据 AM-GM 不等式, 对于任意正数 a,b(三角形边长),

$$a^2+b^2\geq 2\sqrt{a^2b^2}=2ab$$

因此,

$$\frac{a^2+b^2}{2ab}\geq\frac{2ab}{2ab}=1$$

当且仅当 $a^2=b^2$, 即 $a=b$ 时, 等号成立. 此时三角形为等腰三角形, 是有效的几何构型.

代回 $\cos C$ 的表达式, 其最小值为

$$(\cos C)_{\min}=1-\frac{\sqrt{2}}{2}$$

对第 (2) 问, 证明不等式 $C-A\leq\frac{\pi}{6}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由第 (1) 问的结论, $\cos C \geq 1-\frac{\sqrt{2}}{2} > 0$, 故角 C 必为锐角, 即 $C \in (0, \frac{\pi}{2})$. 若 $C \leq A$, 则 $C - A \leq 0 < \frac{\pi}{6}$, 结论成立. 故只需考虑 $C > A$ 的情形. 此时 $0 < C - A < C < \frac{\pi}{2}$. 在区间 $(0, \frac{\pi}{2})$ 内, 正弦函数单调递增, 故只需证明 $\sin(C-A) \leq \sin\frac{\pi}{6} = \frac{1}{2}$.

利用正弦定理与余弦定理导出一个关于 $\sin(C-A)$ 的恒等式.

$$\begin{aligned}\sin(C-A) &= \sin C\cos A - \cos C\sin A \\ &= \frac{c}{2R}\frac{b^2+c^2-a^2}{2bc} - \frac{a^2+b^2-c^2}{2ab}\frac{a}{2R} \\ &= \frac{1}{4R}\left(\frac{b^2+c^2-a^2}{b} - \frac{a^2+b^2-c^2}{b}\right) \\ &= \frac{2c^2-2a^2}{4Rb} = \frac{c^2-a^2}{2Rb}\end{aligned}$$

将 $b = 2R\sin B$ 代入分母, 可得一个更为简洁的形式

$$\sin(C-A) = \frac{c^2-a^2}{b^2}\sin B$$

接着, 将题目给出的条件 $c^2 = \sqrt{2}ab$ 代入上式, 得

$$\sin(C-A) = \frac{\sqrt{2}ab-a^2}{b^2}\sin B = \left(\sqrt{2}\frac{a}{b} - \left(\frac{a}{b}\right)^2\right)\sin B$$

分析这个表达式的取值. 令 $x = \frac{a}{b}$, 考察函数 $f(x) = \sqrt{2}x - x^2$. 这是一个开口向下的二次函数, 其最大值在顶点处取得. 该函数的对称轴为 $x = -\frac{\sqrt{2}}{2(-1)} = \frac{\sqrt{2}}{2}$. 其最大值为 $f\left(\frac{\sqrt{2}}{2}\right) = \sqrt{2}\left(\frac{\sqrt{2}}{2}\right) - \left(\frac{\sqrt{2}}{2}\right)^2 = 1 - \frac{1}{2} = \frac{1}{2}$. 因此对任意边长比 $\frac{a}{b}$, 有 $\sqrt{2}\frac{a}{b} - \left(\frac{a}{b}\right)^2 \leq \frac{1}{2}$. 又 $B \in (0, \pi)$, 故 $0 < \sin B \leq 1$. 综合得

$$\sin(C-A) = \left(\sqrt{2}\frac{a}{b} - \left(\frac{a}{b}\right)^2\right)\sin B \leq \frac{1}{2}\cdot 1 = \frac{1}{2}.$$

在当前 $C > A$ 的情形下, $C - A \in (0, \frac{\pi}{2})$, 由正弦函数的单调性可得 $C - A \leq \frac{\pi}{6}$. 结合 $C \leq A$ 时结论显然成立, 原命题得证.

(1) $\cos C$ 的最小值为 $1-\frac{\sqrt{2}}{2}$.(2) 证明见上.

例题 12.27　已知 $\triangle ABC$ 中, $\sin A = \cos B = \tan C$.
(1) 求 $2A + C$ 的值;
(2) 证明: $c > b > \frac{2}{5}a$.

解 (1) 令 $\sin A = \cos B = \tan C = k$. 因 A, B, C 为三角形内角且 $k > 0$, 有 $A \in (0, \pi)$, 而 $\cos B > 0$、$\tan C > 0$ 迫使 B, C 均为锐角.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

由 $\sin A=\cos B$ 与 B 锐, $\cos B=\sin(\frac{\pi}{2}-B)$, 故 $\sin A=\sin(\frac{\pi}{2}-B)$, 通解为 $A=\frac{\pi}{2}-B$ 或 $A=\frac{\pi}{2}+B$. 若 $A=\frac{\pi}{2}-B$, 则 $A+B=\frac{\pi}{2}$, $C=\frac{\pi}{2}$, 但此时 $\tan C$ 无定义, 矛盾. 故必有 $A=\frac{\pi}{2}+B$, 即 $A-B=\frac{\pi}{2}$. 于是 $C=\pi-A-B=\frac{3\pi}{2}-2A$, 代入得

$$2A+C=2A+\left(\frac{3\pi}{2}-2A\right)=\frac{3\pi}{2}.$$

(2) 由正弦定理, $c>b>\frac{2}{5}a$ 等价于 $\sin C>\sin B>\frac{2}{5}\sin A$. 由 (1), $A=\frac{\pi}{2}+B$, $C=\frac{\pi}{2}-2B$; $C>0$ 要求 $B<\frac{\pi}{4}$.

先证 $\sin C>\sin B$, 即 $\cos(2B)>\sin B$. 由题设 $\cos B=\tan C=\cot(2B)$, 得 $\cos(2B)=2\sin B\cos^2 B$, 故不等式化为 $2\sin B\cos^2 B>\sin B$. 因 $B\in(0,\frac{\pi}{4})$, $\sin B>0$, 约去后得 $\cos^2 B>\frac{1}{2}$, 即 $\cos B>\frac{1}{\sqrt{2}}$, 亦即 $B<\frac{\pi}{4}$, 恰为 $C>0$ 的必要条件, 故成立.

再证 $\sin B>\frac{2}{5}\sin A$, 即 $\sin B>\frac{2}{5}\cos B$, 或 $\tan B>\frac{2}{5}$. 仍用 $\cos B=\cot(2B)$. 令 $t=\tan B$, 则 $\cos B=\frac{1}{\sqrt{1+t^2}}$, $\cot(2B)=\frac{1-t^2}{2t}$, 得

$$\frac{1}{\sqrt{1+t^2}}=\frac{1-t^2}{2t}.$$

由于 $B\in(0,\pi/4)$, $t\in(0,1)$, 方程两边均为正. 两边平方得

$$\frac{1}{1+t^2}=\frac{(1-t^2)^2}{4t^2}\implies 4t^2=(1+t^2)(1-t^2)^2$$

令 $u=t^2$, 得关于 u 的方程 $4u=(1+u)(1-u)^2$. 展开得 $4u=(1+u)(1-2u+u^2)=1-2u+u^2+u-2u^2+u^3=u^3-u^2-u+1$. 整理得一个关于 u 的三次方程

$$g(u)=u^3-u^2-5u+1=0$$

需证明 $t>\frac{2}{5}$, 这等价于证明 $u=t^2>\left(\frac{2}{5}\right)^2=\frac{4}{25}$. 分析函数 $g(u)$. $g'(u)=3u^2-2u-5=(3u-5)(u+1)$. 因为 $B\in(0,\pi/4)$, 所以 $t=\tan B\in(0,1)$, 故 $u=t^2\in(0,1)$. 在此区间内, $u+1>0$ 且 $3u-5<0$, 故 $g'(u)<0$. 这说明函数 $g(u)$ 在其定义域 $(0,1)$ 上是严格单调递减的.

先计算 $u=\frac{4}{25}$ 处的函数值.

$$g\left(\frac{4}{25}\right)=\left(\frac{4}{25}\right)^3-\left(\frac{4}{25}\right)^2-5\left(\frac{4}{25}\right)+1=\frac{64}{15625}-\frac{16}{625}-\frac{4}{5}+1$$

$$=\frac{64-16\cdot 25-4\cdot 3125+15625}{15625}=\frac{64-400-12500+15625}{15625}=\frac{2789}{15625}>0$$

因为 $g(u)$ 单调递减, 且 $g(\frac{4}{25})>0$, 而 u_0 是使 $g(u_0)=0$ 的根, 故必有 $u_0>\frac{4}{25}$. 因此 $\tan^2 B>\frac{4}{25}$, 从而 $\tan B>\frac{2}{5}$.$\sin B>\frac{2}{5}\sin A$ 得证.

与前一部分 $\sin C>\sin B$ 合并, 不等式 $c>b>\frac{2}{5}a$ 得证.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

12.4.1 什么时候不宜使用建系法

前一节把建系法确立为处理复杂几何问题的一条保底路径. 然而普适性不等于最优性. 获得答案之后, 我们常常追问有没有更短、更直接的路. 过度依赖建系法, 有时会把结构清晰的问题变成冗长计算, 掩盖其内在几何联系.

因此需要敏锐识别那些不宜使用建系法的情形. 遇到下面几类情形, 我们优先尝试其他方法.

- **几何结构特征极其明显时**. 当题目条件或图形本身强烈暗示了某种特定的几何关系时, 回归纯粹的几何或三角方法往往更为高效.
  - **相似关系**: 若题目中出现平行线、角相等(尤其是跨越不同三角形的角相等)等条件, 应首先审视是否存在相似三角形. 相似关系能直接建立起边长的比例式, 其效率远高于通过坐标计算多条边长再求比值.
  - **圆的性质**: 若问题涉及共圆点、切线、相交弦等, 应优先考虑圆幂定理、托勒密定理、弦切角定理等. 这些定理是处理圆相关问题的专用工具, 能把复杂的几何构型直接转化为简洁的代数乘积式, 而坐标法需要联立直线与圆的方程, 计算量远大于此.
- **问题只涉及比例关系时**. 当问题的条件与结论主要涉及边长的比例时, 正弦定理、角平分线定理等工具能直接在比例的层面上操作. 坐标法引入了绝对的坐标值, 反而可能把一个简洁的比例问题复杂化.
- **向量方法更为直接时**. 对于涉及中点、重心、定比分点以及证明共线、垂直等问题, 向量运算往往比坐标运算更为简洁和普适. 向量法不依赖于坐标系的选择, 其运算直接反映几何变换本身.

为更精细地理解这种选择, 再看上一节的例题, 对比其几何解法与坐标法的思想差异.

例题 12.28 (方法对比分析) 在 $\triangle ABC$ 中, D 为边 BC 上一点, $DC=3$, $AD=5$, $AC=7$, $\angle DAC=\angle ABC$. 求 $\triangle ABC$ 的面积.

解

解此题要洞察到条件 $\angle DAC=\angle ABC$ 的几何内涵. 考察 $\triangle DAC$ 与 $\triangle ABC$. 这两个三角形中

- $\angle C$ 是公共角, 即 $\angle DCA=\angle BCA$.
- $\angle DAC=\angle ABC$(题目已知).

根据 AA 相似判定, 得 $\triangle BAC\sim\triangle ADC$.

这一发现是解题的要点. 相似关系直接给出所有对应边的比例关系

$$\frac{BC}{AC}=\frac{AC}{DC}=\frac{AB}{AD}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入已知值 $AC=7, DC=3, AD=5$, 可直接解出所有未知边长

$$BC=\frac{AC^2}{DC}=\frac{7^2}{3}=\frac{49}{3}$$

$$AB=\frac{AC\cdot AD}{DC}=\frac{7\cdot 5}{3}=\frac{35}{3}$$

接下来, 为求面积 $S=\frac{1}{2}AB\cdot BC\sin B$, 只需 $\sin B$. 由 $\angle B=\angle DAC$, 在 $\triangle ADC$ 中利用余弦定理求出 $\cos(\angle DAC)$, 进而得到 $\sin(\angle DAC)$, 即 $\sin B$.

$$\cos(\angle DAC)=\frac{AD^2+AC^2-DC^2}{2(AD)(AC)}=\frac{5^2+7^2-3^2}{2(5)(7)}=\frac{65}{70}=\frac{13}{14}$$

$$\sin B=\sin(\angle DAC)=\sqrt{1-(13/14)^2}=\frac{3\sqrt{3}}{14}$$

最终面积为

$$S_{\triangle ABC}=\frac{1}{2}\cdot\frac{35}{3}\cdot\frac{49}{3}\cdot\frac{3\sqrt{3}}{14}=\frac{245\sqrt{3}}{12}$$

此解法逻辑清晰, 计算直接, 利用了问题的内在几何结构.

如前所示, 坐标法同样能解此题. 其路径如下. 建立坐标系 $\to$ 确定 A, C, D 坐标 $\to$ 设 B 的坐标为 $(x_B, 0)$ $\to$ 利用向量点积将 $\cos(\angle DAC)=\cos(\angle ABC)$ 转化为关于 x_B 的复杂方程 $\to$ 解方程求出 x_B $\to$ 计算底和高求得面积.

这个过程每一步都遵循固定的代数规则,完全可行. 然而它把一个简洁的相似关系拆解为一系列繁琐的坐标运算和方程求解. 在这一过程中, $\angle DAC=\angle ABC$ 这一条件被当作代数方程中的一个等量关系使用, 相似三角形之间的简洁联系便在繁琐计算中被忽略.

方法选择的准则

建系法普适但失于灵巧. 结构精巧的问题中, 几何关系往往更简洁; 当几何线索难以引导时, 建系法仍是稳健的路径. 以最能说明问题结构的方法求解, 本身就是一种基本的数学理解.

12.5 梅涅劳斯定理与塞瓦定理

正弦定理与余弦定理处理三角形的边角计算. 当问题涉及三点共线或三线共点时, 两个经典定理提供了直接的代数判据, 梅涅劳斯定理判定共线, 塞瓦定理判定共点.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

12.5.1 梅涅劳斯定理

定理 12.15 (梅涅劳斯定理) 在 $\triangle ABC$ 中, 点 X、Y、Z 分别在边 AB、BC、CA 或其延长线上, 其中有奇数个点在边的延长线上. 三点 X、Y、Z 共线的充分必要条件是

$$\frac{AX}{XB}\cdot\frac{BY}{YC}\cdot\frac{CZ}{ZA}=-1 \tag{12-14}$$

其中各线段取有向长度(点在边上时比值为正, 在延长线上时为负).

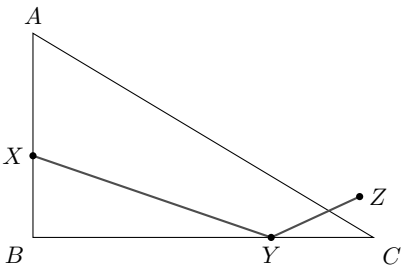


图 12.16　梅涅劳斯定理中截线 XYZ 与三角形三边, 或延长线的交点

证明 (面积法证明) 设直线 XYZ 截 $\triangle ABC$ 的三边, 或延长线于 X、Y、Z. 利用面积比与共边三角形的关系:

$$\frac{AX}{XB}=\frac{S_{\triangle AXY}}{S_{\triangle BXY}},\quad \frac{BY}{YC}=\frac{S_{\triangle BXY}}{S_{\triangle CXY}},\quad \frac{CZ}{ZA}=\frac{S_{\triangle CXY}}{S_{\triangle AXY}}$$

三式相乘, 分子分母轮换相消, 得

$$\frac{AX}{XB}\cdot\frac{BY}{YC}\cdot\frac{CZ}{ZA}=\frac{S_{\triangle AXY}}{S_{\triangle BXY}}\cdot\frac{S_{\triangle BXY}}{S_{\triangle CXY}}\cdot\frac{S_{\triangle CXY}}{S_{\triangle AXY}}=1$$

此处面积取有向面积, 当截线穿过三角形时恰有一个或三个交点在延长线上, 有向面积的符号使得乘积为 -1. □

证明 (向量法证明) 记 $\overrightarrow{AX}=\lambda_1\overrightarrow{XB},\overrightarrow{BY}=\lambda_2\overrightarrow{YC},\overrightarrow{CZ}=\lambda_3\overrightarrow{ZA}$. 选取 $\overrightarrow{AB}$ 与 $\overrightarrow{AC}$ 为基向量, 则

$$\overrightarrow{XY}=\overrightarrow{AY}-\overrightarrow{AX}=\left(\frac{1}{1+\lambda_2}-\frac{\lambda_1}{1+\lambda_1}\right)\overrightarrow{AB}+\frac{\lambda_2}{1+\lambda_2}\overrightarrow{AC}$$

$$\overrightarrow{XZ}=\overrightarrow{AZ}-\overrightarrow{AX}=\frac{1}{1+\lambda_3}\overrightarrow{AC}-\frac{\lambda_1}{1+\lambda_1}\overrightarrow{AB}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

X、Y、Z 三点共线当且仅当 $\overrightarrow{XY}$ 与 $\overrightarrow{XZ}$ 共线, 即两组系数对应成比例. 经计算, 这等价于 $\lambda_1\lambda_2\lambda_3 = -1$, 也就是 (12-14). □

向量形式给出了梅涅劳斯定理的简洁表达: $\lambda_1\lambda_2\lambda_3 = -1$, 且充分性与必要性在同一个计算中完成.

例题 12.29 (三个圆的外公切线交点共线) 考虑三个半径两两不同的圆, 每两个圆有一组外公切线, 三组外公切线的交点记为 P、Q、R. 这三个点共线.

解 设三个圆心为 A、B、C, 半径分别为 r_A、r_B、r_C. 圆 A 与圆 B 的外公切线交点 P 在 AB 的延长线上, 满足 $\frac{PA}{PB} = \frac{r_A}{r_B}$. 同理 $\frac{QB}{QC} = \frac{r_B}{r_C}$, $\frac{RC}{RA} = \frac{r_C}{r_A}$. 三式相乘:

$$\frac{PA}{PB}\cdot\frac{QB}{QC}\cdot\frac{RC}{RA} = \frac{r_A}{r_B}\cdot\frac{r_B}{r_C}\cdot\frac{r_C}{r_A} = 1$$

三个点分别在三角形三边的延长线上(偶数个在延长线上时乘积为 1, 对应梅涅劳斯定理的有向长度约定), 由梅涅劳斯定理知 P、Q、R 共线.

12.5.2 塞瓦定理

定理 12.16 (塞瓦定理) 在 $\triangle ABC$ 中, 点 X、Y、Z 分别在边 AB、BC、CA 或其延长线上, 其中有偶数个点在边的延长线上. 三条线段 AY、BZ、CX 交于一点或相互平行的充分必要条件是

$$\frac{AX}{XB}\cdot\frac{BY}{YC}\cdot\frac{CZ}{ZA} = 1 \tag{12-15}$$

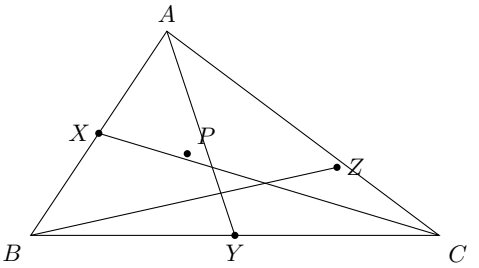


图 12.17 塞瓦定理, 三线 AY、BZ、CX 共点于 P

证明 (用梅涅劳斯定理证明) 设 AY、BZ、CX 交于点 P. 对 $\triangle ABY$ 和截线 XPC 使用梅涅劳斯定理:

$$\frac{AX}{XB}\cdot\frac{BC}{CY}\cdot\frac{YP}{PA} = -1$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再对 $\triangle ACY$ 和截线 BPZ 使用梅涅劳斯定理:

$$\frac{YB}{BC}\cdot\frac{CZ}{ZA}\cdot\frac{AP}{PY}=-1$$

两式相乘, $\frac{YP}{PA}\cdot\frac{AP}{PY}=1$, $\frac{BC}{CY}\cdot\frac{YB}{BC}=\frac{YB}{CY}$, 得

$$\frac{AX}{XB}\cdot\frac{BY}{YC}\cdot\frac{CZ}{ZA}=1$$

□

证明 (面积法证明) 设三条直线 AY、BZ、CX 交于 P. 利用共顶点三角形的面积比:

$$\frac{AX}{XB}=\frac{S_{\triangle CPA}}{S_{\triangle CPB}},\quad \frac{BY}{YC}=\frac{S_{\triangle APB}}{S_{\triangle APC}},\quad \frac{CZ}{ZA}=\frac{S_{\triangle BPC}}{S_{\triangle BPA}}$$

三式相乘, 分子分母轮换相消, 得 1. □

证明 (向量法证明) 记 $\overrightarrow{AX}=\lambda_1\overrightarrow{XB}$, $\overrightarrow{BY}=\lambda_2\overrightarrow{YC}$, $\overrightarrow{CZ}=\lambda_3\overrightarrow{ZA}$. 点 P 在 BZ 上, 可设 $\overrightarrow{AP}=(1-s)\overrightarrow{AB}+\frac{s}{1+\lambda_3}\overrightarrow{AC}$; 点 P 也在 CX 上, 可设 $\overrightarrow{AP}=\frac{t\lambda_1}{1+\lambda_1}\overrightarrow{AB}+(1-t)\overrightarrow{AC}$. 比较两组系数, 解出 P 的向量表达式, 再要求 P 在 AY 上, 得到 $\lambda_1\lambda_2\lambda_3=1$. □

塞瓦定理的向量形式与梅涅劳斯定理统一为: $\lambda_1\lambda_2\lambda_3=-1$ 判定共线, $\lambda_1\lambda_2\lambda_3=1$ 判定共点.

12.5.2.1 塞瓦定理与三角形五心

塞瓦定理给出了三线共点的统一判据. 三角形的重心、内心、垂心都可以用塞瓦定理验证:

- **重心**: 三条中线的分比 $\lambda_1=\lambda_2=\lambda_3=1$, 乘积为 1.
- **内心**: 三条角平分线的分比由角平分线定理给出 $\lambda_1=\frac{b}{a}$, $\lambda_2=\frac{c}{b}$, $\lambda_3=\frac{a}{c}$, 乘积为 1.
- **垂心**: 三条高线的分比 $\lambda_1=\frac{b\cos A}{a\cos B}$ 等, 乘积中余弦项轮换相消为 1.

12.5.3 张角定理

定理 12.17 (张角定理) 在 $\triangle ABC$ 中, D 是 BC 边上任意一点, 则

$$\frac{\sin\angle BAC}{AD}=\frac{\sin\angle CAD}{AB}+\frac{\sin\angle BAD}{AC}\tag{12-16}$$

当射线 AD 落在 $\angle BAC$ 内部时, 逆定理成立, 即若四点满足上式, 则 B、C、D 三点共线.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

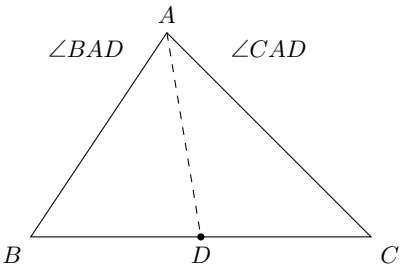


图 12.18 张角定理

知乎@余命数

证明 由 $S_{\triangle ABC}=S_{\triangle ABD}+S_{\triangle ACD}$, 展开面积公式:

$$\frac{1}{2}AB\cdot AC\cdot\sin\angle BAC=\frac{1}{2}AB\cdot AD\cdot\sin\angle BAD+\frac{1}{2}AC\cdot AD\cdot\sin\angle CAD$$

两边同除以 $\frac{1}{2}AB\cdot AC\cdot AD$ 即得 (12-16). □

张角定理建立了顶角被分成的两个角与对边分点之间的正弦关系, 常用于处理共线问题和角平分线长度公式.

例题 12.30 (角平分线长度公式) 利用张角定理推导三角形角平分线长度公式.

解 设 AD 是 $\angle A$ 的平分线, 则 $\angle BAD=\angle CAD=\frac{A}{2}$. 由张角定理:

$$\frac{\sin A}{AD}=\frac{\sin\frac{A}{2}}{AB}+\frac{\sin\frac{A}{2}}{AC}=\sin\frac{A}{2}\left(\frac{1}{AB}+\frac{1}{AC}\right)$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

利用 $\sin A=2\sin\frac{A}{2}\cos\frac{A}{2}$, 两边约去 $\sin\frac{A}{2}$:

$$\frac{2\cos\frac{A}{2}}{AD}=\frac{1}{AB}+\frac{1}{AC}=\frac{AB+AC}{AB\cdot AC}$$

所以角平分线长度为

$$AD=\frac{2\,AB\cdot AC\cdot\cos\frac{A}{2}}{AB+AC}$$

结合余弦半角公式 $\cos\frac{A}{2}=\sqrt{\frac{s(s-a)}{bc}}$ (s 为半周长), 可得

$$AD=\frac{2\sqrt{bc\,s(s-a)}}{b+c}\tag{12-17}$$

其中 $b=AC, c=AB$.

知乎@余命数

例题 12.31 (共线判定) 在 $\triangle ABC$ 中, D 是 BC 边上的点, P 是平面上一点. 若 $\frac{\sin\angle BAP}{AP}=\frac{\sin\angle CAP}{AB}+\frac{\sin\angle BAP}{AC}$ 且 D 在 BC 上满足 $\angle BAP=\angle BAD$, 则 B、D、C 共线.

解 由张角定理的逆定理, 满足 (12-16) 的四点 A、B、C、D 中, B、C、D 必共线. 这给出了一个用正弦关系判定共线的方法.

知乎@余命数

12.6 斯特瓦尔特定理与三角形特殊点

12.6.1 斯特瓦尔特定理

斯特瓦尔特定理给出了三角形中从顶点到对边任意一点的距离公式, 它是中线长公式、角平分线长公式和余弦定理的推广.

知乎@余命数

定理 12.18 (斯特瓦尔特定理) 在 $\triangle ABC$ 中, 点 D 是 BC 边上任意一点, 记 $BD=m, CD=n, BC=a=m+n$, 则

$$m\,b^2+n\,c^2=a\,(AD^2+mn) \tag{12-18}$$

其中 $b=AC, c=AB$.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

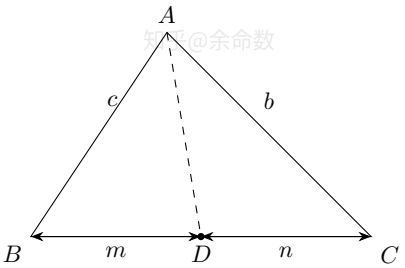


图 12.19　斯特瓦尔特定理

证明 (余弦定理证明) $\angle ADB$ 与 $\angle ADC$ 互补, 故 $\cos\angle ADB+\cos\angle ADC=0$. 由余弦定理:

$$\cos\angle ADB=\frac{AD^2+m^2-c^2}{2\,AD\cdot m},\quad \cos\angle ADC=\frac{AD^2+n^2-b^2}{2\,AD\cdot n}$$

两式相加并乘以 $2\,AD\cdot mn$:

$$n(AD^2+m^2-c^2)+m(AD^2+n^2-b^2)=0$$

知乎@余命数

整理得 $(m+n)\,AD^2 = m\,b^2 + n\,c^2 - mn(m+n)$, 即 $a\,AD^2 = m\,b^2 + n\,c^2 - a\,mn$. □

证明 (向量法证明) 设 $\overrightarrow{BD} = \lambda\overrightarrow{DC}$, 则 $\lambda = \frac{m}{n}$, 且

$$\overrightarrow{AD} = \frac{n}{a}\overrightarrow{AB} + \frac{m}{a}\overrightarrow{AC}$$

两边取模平方:

$$AD^2 = \frac{n^2}{a^2}\,c^2 + \frac{m^2}{a^2}\,b^2 + \frac{2mn}{a^2}\,bc\cos A$$

用余弦定理 $2bc\cos A = b^2 + c^2 - a^2$ 代入并整理, 即得 (12-18). □

12.6.1.1 中线长公式

当 D 为 BC 中点时, $m = n = \frac{a}{2}$, AD 为中线长 l_a. 代入 (12-18):

$$\frac{a}{2}(b^2 + c^2) = a\left(l_a^2 + \frac{a^2}{4}\right)$$

解出

$$l_a^2 = \frac{2b^2 + 2c^2 - a^2}{4} \tag{12-19}$$

12.6.1.2 角平分线长公式

当 AD 为 $\angle A$ 的平分线时, 由角平分线定理 $\frac{BD}{DC} = \frac{c}{b}$, 故 $m = \frac{ac}{b+c}$, $n = \frac{ab}{b+c}$. 代入 (12-18) 并整理:

$$\omega_a^2 = bc\left(1 - \frac{a^2}{(b+c)^2}\right) = \frac{4bc\,s(s-a)}{(b+c)^2} \tag{12-20}$$

其中 $s = \frac{a+b+c}{2}$ 为半周长.

12.6.1.3 高线长公式与海伦公式

高线长 h_a 可直接由面积公式 $S = \frac{1}{2}ah_a$ 与余弦定理得到:

$$h_a = b\sin C = b\sqrt{1 - \cos^2 C} = \frac{1}{2a}\sqrt{(a+b+c)(a+b-c)(b+c-a)(c+a-b)}$$

结合 $S = \frac{1}{2}ah_a$ 与半周长 $s = \frac{a+b+c}{2}$, 得到海伦公式:

$$S = \sqrt{s(s-a)(s-b)(s-c)} \tag{12-21}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

12.6.2 三角形的旁心

定义 12.5 (旁心) 三角形一个内角的平分线与另外两个角的外角平分线交于一点, 称为该角对应的**旁心**. 旁心是三角形旁切圆的圆心, 旁切圆与一条边和另外两条边的延长线相切.

三角形有三个旁心, 分别记为 I_A、I_B、I_C, 对应顶点 A、B、C 的对角.

定理 12.19 (旁心的向量表示) 设 I_A 为 $\triangle ABC$ 对应顶点 A 的旁心, 以平面上任一点 O 为参考点, 则

$$\overrightarrow{OI_A}=\frac{-a\overrightarrow{OA}+b\overrightarrow{OB}+c\overrightarrow{OC}}{-a+b+c} \tag{12-22}$$

知乎@余命数

证明 I_A 在 $\angle A$ 的内角平分线上, 同时在 $\angle B$ 和 $\angle C$ 的外角平分线上. 由角平分线定理, I_A 到 BC 的距离等于到 AB、AC 延长线的距离. 利用塞瓦定理的向量形式, I_A 对应的分比为 $\lambda_1=-\frac{b}{a}$, $\lambda_2=\frac{c}{b}$, $\lambda_3=-\frac{a}{c}$, 乘积 $\lambda_1\lambda_2\lambda_3=1$, 满足塞瓦条件. 代入塞瓦定理的向量表达式即得. □

知乎@余命数

旁心 I_A 的重心坐标为 $(-a:b:c)$, 其中 a 前的负号反映了 I_A 在 $\angle A$ 的对侧. 当三角形为锐角三角形时, I_A 在 $\angle A$ 对侧的三角形外部.

12.6.3 费马点

知乎@余命数

定义 12.6 (费马点) 在 $\triangle ABC$ 中, 到三个顶点距离之和 $PA+PB+PC$ 最小的点 P 称为费马点.

知乎@余命数

定理 12.20 (费马点的构造) 若 $\triangle ABC$ 的所有内角都小于 120°, 费马点 F 满足 $\angle AFB=\angle BFC=\angle CFA=120^\circ$. 若某个内角 $\geq 120^\circ$, 则该角顶点就是费马点.

证明 以 BC 为边向 $\triangle ABC$ 外侧作正三角形 BCD, 连接 AD. 对三角形内任一点 P, 把 $\triangle BPC$ 绕 B 旋转 60°, 使 C 落在 D 上, 记 P 的像为 P'. 旋转保持距离, 故 $P'D=PC$; 又 $BP'=BP$ 且 $\angle PBP'=60^\circ$, 故 $\triangle PBP'$ 是正三角形, $PP'=PB$. 于是

$$PA+PB+PC=PA+PP'+P'D\geq AD,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

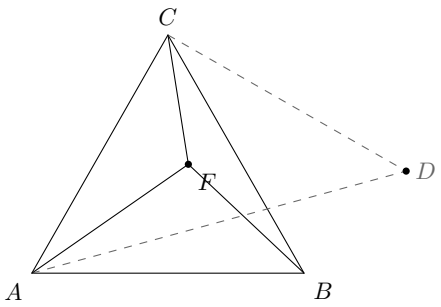


图 12.20 费马点 F, 以 BC 为边向外作正三角形 BCD, AD 与外接圆弧的交点即为 F

等号当且仅当 A、P、P'、D 依次共线, 此时 P 落在 AD 上. 故 $\triangle ABC$ 内使 $PA+PB+PC$ 最小的点 F 在 AD 上, 最小值为 AD.

再定 F 处的三个角. 取等号时 A、F、F'、D 依次共线, 而 $\triangle FBF'$ 是正三角形, 故 $\angle AFB = 180^\circ - \angle BFF' = 120^\circ$. 旋转把 C 映到 D、把 F 映到 F', 故 $\angle BFC = \angle BF'D = 180^\circ - \angle BF'F = 120^\circ$. 于是 $\angle CFA = 360^\circ - 120^\circ - 120^\circ = 120^\circ$. □

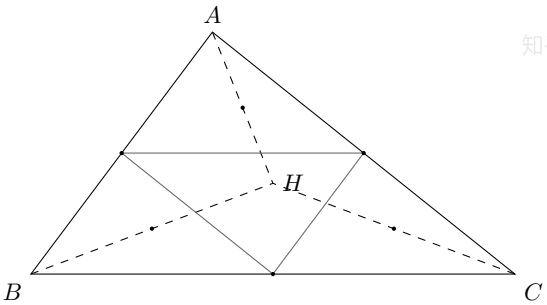
余命数 知乎@余命数

12.6.4 九点圆

定理 12.21 (九点圆定理) 在 $\triangle ABC$ 中, 以下九个点共圆:

1. 三边的中点(共 3 个);
2. 三条高线的垂足(共 3 个);
3. 三个顶点到垂心的连线的中点(共 3 个).

这个圆称为三角形的九点圆, 其半径等于外接圆半径的一半.



知乎@余命数

图 12.21 九点圆, 三边中点、三高线垂足、三顶点与垂心连线中点共圆

证明 设 H 为垂心, O 为外心, R 为外接圆半径. 记三边中点为 M_a、M_b、M_c, 三高线垂足为 D_a、D_b、D_c, 三顶点与垂心连线中点为 E_a、E_b、E_c, 并记 N 为 OH 的中点.

三边中点 M_a、M_b、M_c 构成中位三角形, 其外接圆半径为 $\frac{R}{2}$.

以 O 为原点记顶点向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$，则 $|\mathbf{a}|=|\mathbf{b}|=|\mathbf{c}|=R$，垂心向量 $\mathbf{h}=\mathbf{a}+\mathbf{b}+\mathbf{c}$，故 $\overrightarrow{ON}=\frac{\mathbf{h}}{2}$ 且

$$\overrightarrow{NM_a}=\frac{\mathbf{b}+\mathbf{c}}{2}-\frac{\mathbf{h}}{2}=-\frac{\mathbf{a}}{2},\qquad \overrightarrow{NE_a}=\frac{\mathbf{a}+\mathbf{h}}{2}-\frac{\mathbf{h}}{2}=\frac{\mathbf{a}}{2}.$$

轮换即得 $|NM_a|=|NM_b|=|NM_c|=|NE_a|=|NE_b|=|NE_c|=\frac{R}{2}$. 这六点共圆，圆心为 N，半径 $\frac{R}{2}$，其中 $M_aM_bM_c$ 的外接圆就是这个圆.

高线垂足 D_a 也在这个圆上. 因 A、D_a、H 共线，而 E_a 是 AH 的中点，故直线 E_aD_a 就是高线所在直线，垂直于 BC；又 M_a、D_a 都在直线 BC 上，故 $\angle M_aD_aE_a=90^\circ$. 由 $\overrightarrow{M_aE_a}=\mathbf{a}$，即 $|M_aE_a|=R$，故 M_aE_a 是圆的直径，由直径所对的圆周角为直角知 D_a 在圆上. 轮换得 D_b、D_c 也在圆上.

九个点同在半径为 $\frac{R}{2}$ 的圆上，圆心在 OH 的中点处. □

定理 12.22 (欧拉公式) 设 O 为外心，I 为内心，R 为外接圆半径，r 为内切圆半径，则

$$OI^2=R^2-2Rr \tag{12-23}$$

证明 点 I 对以 O 为圆心、R 为半径的圆的幂为 OI^2-R^2，故 $OI^2=R^2-2Rr$ 等价于该幂等于 $-2Rr$.

利用内心的向量表示 $\overrightarrow{OI}=\frac{a\overrightarrow{OA}+b\overrightarrow{OB}+c\overrightarrow{OC}}{a+b+c}$，计算 $|\overrightarrow{OI}|^2$ 并利用 $|\overrightarrow{OA}|=|\overrightarrow{OB}|=|\overrightarrow{OC}|=R$ 以及 $\overrightarrow{OA}\cdot\overrightarrow{OB}=R^2\cos 2C$ 等关系，再代入 $a=2R\sin A$ 与 $r=4R\sin\frac{A}{2}\sin\frac{B}{2}\sin\frac{C}{2}$，化简得 $OI^2=R^2-2Rr$. □

欧拉公式给出了外心与内心距离的精确表达式. 由 $OI^2\geq 0$ 立得 $R\geq 2r$，即**欧拉不等式**，等号当且仅当三角形为正三角形时成立.

12.7 托勒密定理与西姆松定理

第12.5节 与 第12.6节 处理了三角形内部的线段关系. 本节转向圆内接四边形与三角形外接圆上的点，给出托勒密定理与西姆松定理，它们是圆与三角形结合的两个经典结论.

12.7.1 托勒密定理

圆内接四边形有对角互补的角度关系，但角度关系定不下对角线的长度. 托勒密定理把四条边和两条对角线压进一个等式，量出五条便锁定第六条. 取一条对角线为直径、圆直径为 1，这个等式给出两角和的正弦公式 $\sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 12.23 (托勒密定理) 圆内接四边形 $ABCD$ 中, 两条对角线的乘积等于两组对边乘积之和

$$AC \cdot BD = AB \cdot CD + AD \cdot BC \tag{12-24}$$

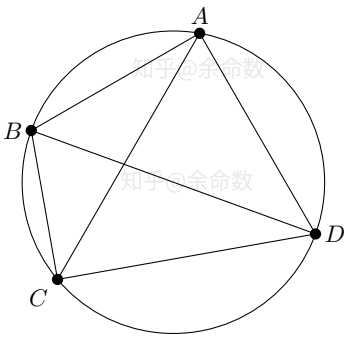


图 12.22 圆内接四边形 $ABCD$

证明 在 BD 上取点 E 使 $\angle BAE = \angle CAD$. 由 $\angle ABE = \angle ACD$, 同弧 AD 所对圆周角相等, $\triangle ABE \sim \triangle ACD$, 故

$$\frac{AB}{AC} = \frac{BE}{CD} \implies AB \cdot CD = AC \cdot BE$$

又 $\angle BAC = \angle EAD$, 由 $\angle BAE = \angle CAD$ 两端加 $\angle EAC$ 即得, 且 $\angle ACB = \angle ADE$ (同弧 AB 所对), 故 $\triangle ABC \sim \triangle AED$, 于是

$$\frac{BC}{ED} = \frac{AC}{AD} \implies AD \cdot BC = AC \cdot ED$$

两式相加, 得 $AB \cdot CD + AD \cdot BC = AC \cdot BE + AC \cdot ED = AC \cdot (BE + ED) = AC \cdot BD$. □

等号来自 E 恰好落在 BD 上, 使 $BE + ED = BD$. 构造一点拆分对角线、再用相似把乘积折算到同一段上, 是圆内接四边形度量问题的常用手法.

例题 12.32 (正五边形与黄金比) 正五边形 $ABCDE$ 的边长为 s, 对角线长为 d. 用托勒密定理求 d/s.

解 在圆内接四边形 $ABCD$ 中, $AB = BC = CD = s$, $AC = BD = d$. 由托勒密定理 (12-24),

$$AC \cdot BD = AB \cdot CD + AD \cdot BC$$

即 $d^2 = s^2 + d \cdot s$, 亦即 $d^2 - sd - s^2 = 0$. 解得 $d/s = (1 + \sqrt{5})/2 = \varphi$(黄金比).

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 12.33 (托勒密定理求对角线) 圆内接四边形 $ABCD$ 中 $AB=3, BC=4$, $CD=5, DA=6$. 求对角线 AC 与 BD 的比值.

解 由托勒密定理 $AC\cdot BD = AB\cdot CD + AD\cdot BC = 3\cdot 5 + 6\cdot 4 = 39$. 在 $\triangle ABC$ 与 $\triangle ADC$ 中分别对 AC 用余弦定理, 由圆内接四边形对角互补 $B+D=\pi$ 得 $\cos D = -\cos B$.

$$AC^2 = 25 - 24\cos B = 61 + 60\cos B$$

解得 $\cos B = -3/7$, $AC^2 = 247/7$. 同理对 BD 用 $\cos C = -\cos A$.

$$BD^2 = 45 - 36\cos A = 41 + 40\cos A$$

解得 $\cos A = 1/19$, $BD^2 = 819/19$. 故

$$\frac{AC}{BD} = \sqrt{\frac{247/7}{819/19}} = \sqrt{\frac{247\cdot 19}{7\cdot 819}} = \sqrt{\frac{13\cdot 19^2}{7\cdot 9\cdot 7\cdot 13}} = \frac{19}{21}.$$

12.7.2 托勒密不等式

去掉四点共圆的条件, 托勒密等式不再成立. 对任意四点, 乘积 $AC\cdot BD$ 不超过两组对边乘积之和, 等号恰好把共圆情形挑出来. 这样托勒密定理就从计算工具升级为共圆判定.

定理 12.24 (托勒密不等式) 对平面上任意四点 A、B、C、D,

$$AC\cdot BD \le AB\cdot CD + AD\cdot BC \tag{12-25}$$

等号成立当且仅当 A、B、C、D 四点共圆; 四点共线的退化情形需另加次序条件, 例如四点共线且按 A、B、C、D 的次序排列时取等号.

证明 取点 E 使 $\triangle ABE \sim \triangle ACD$(对应顶点 $A\to A, B\to C, E\to D$), 即 $\angle BAE = \angle CAD$ 且 $\frac{AB}{AC} = \frac{AE}{AD}$. 则 $\frac{AB}{AC} = \frac{BE}{CD}$, 故 $AB\cdot CD = AC\cdot BE$. 又 $\angle BAC = \angle EAD$, $\frac{AB}{AC} = \frac{AE}{AD}$, 故 $\triangle BAC \sim \triangle EAD$, $\frac{BC}{ED} = \frac{AC}{AD}$, $AD\cdot BC = AC\cdot ED$. 于是

$$AB\cdot CD + AD\cdot BC = AC\cdot BE + AC\cdot ED = AC\cdot(BE+ED) \ge AC\cdot BD$$

最后一步用三角不等式 $BE+ED\ge BD$, 等号当 E 在线段 BD 上时成立. 此时 $\angle ABD$ 与 $\angle ABE$ 相等, 而 $\triangle ABE\sim\triangle ACD$ 给出 $\angle ABE = \angle ACD$, 故 $\angle ABD = \angle ACD$; 点 B 与点 C 对线段 AD 张出相等的圆周角, 于是 A、B、C、D 四点共圆. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明复用托勒密定理的构造, 区别在于点 E 一般不在 BD 上, 最后一步由三角不等式 $BE+ED\geq BD$ 产生不等号.E 落在 BD 上时 $\angle ABD=\angle ACD$, 这是四点共圆的刻画, 等号条件由此浮现.

12.7.3 西姆松定理

托勒密定理处理圆内接四边形的度量关系. 西姆松定理转向另一个问题, 即三角形外接圆上一点到三边作垂线时, 三个垂足何时共线. 对圆内任意一点, 三个垂足构成垂足三角形; 当点落在外接圆上, 垂足三角形退化为一条线段, 三个垂足共线. 这条线把点在圆上这一条件翻译成垂足共线的几何事实.

定理 12.25 (西姆松定理) 设 P 是 $\triangle ABC$ 外接圆上一点, 异于 A、B、C. 过 P 分别向三边 BC、CA、AB(或其延长线)作垂线, 垂足为 X、Y、Z. 则 X、Y、Z 三点共线.

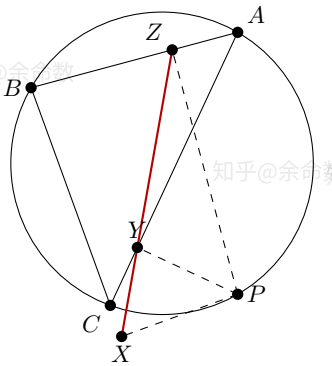


图 12.23 西姆松线, 外接圆上一点到三边的垂足所共的线

证明 以下用两直线的有向夹角 $\angle(l_1,l_2)$ 叙述. 它只依赖两条直线的方向, 以 180° 为周期, 于是三个垂足落在线段上还是落在线段延长线上不必分开讨论.

连接 PX、PY、PZ. 因 $\angle PXC=\angle PYC=90^\circ$, P、X、C、Y 四点共圆, PC 为直径, 同一弦 PX 在 Y 与 C 处张出相等的圆周角, 故

$$\angle(YP,\,YX)=\angle(CP,\,CX).$$

又 $\angle PZA=\angle PYA=90^\circ$, P、Z、A、Y 四点共圆, PA 为直径, 同一弦 PZ 张出

$$\angle(YP,\,YZ)=\angle(AP,\,AZ).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由 P、A、B、C 四点共圆, 弦 BP 在圆周上张出 $\angle(CB,\,CP)=\angle(AB,\,AP)$. 直线 CX 即直线 CB, 直线 AZ 即直线 AB, 于是

$$\angle(YX,\,YZ)=\angle(YX,\,YP)+\angle(YP,\,YZ)=-\angle(CP,\,CX)+\angle(AP,\,AZ)=\angle(CB,\,CP)+\angle(AP,\,AB)=$$

两条直线的有向夹角为零即二者方向相同, 而它们都过点 Y, 故 X、Y、Z 三点共线. □

证明靠两组四点共圆传递等角, $PXCY$ 与 $PZAY$ 分别以 PC、PA 为直径. P 在外接圆上给出 $\angle(CB,CP)=\angle(AB,AP)$, 经圆周角传递化为 $\angle(YX,YZ)=0$, 迫使 X、Y、Z 共线. 逆定理把论证反过来, 从垂足共线推出 $\angle(CB,CP)=\angle(AB,AP)$, 即 P 落在外接圆上. 正逆合在一起, 西姆松定理给出外接圆的垂足刻画.

定理 12.26 (西姆松定理的逆) 设 P 是 $\triangle ABC$ 所在平面上一点. 若 P 到三边 BC、CA、AB 的垂足 X、Y、Z 共线, 则 P 在 $\triangle ABC$ 的外接圆上.

证明 设 X、Y、Z 共线, 仍用两直线的有向夹角. 由 $\angle PXC=\angle PYC=90^\circ$ 得 P、X、C、Y 四点共圆, 故 $\angle(YP,\,YX)=\angle(CP,\,CX)$; 由 X、Y、Z 共线得 $\angle(YP,\,YX)=\angle(YP,\,YZ)$; 又 P、Z、A、Y 四点共圆 (以 PA 为直径), 故 $\angle(YP,\,YZ)=\angle(AP,\,AZ)$. 三式合并, 并对两边同取反向角, 得

$$\angle(CB,\,CP)=\angle(CX,\,CP)=\angle(AZ,\,AP)=\angle(AB,\,AP),$$

其中 CX 即直线 CB, AZ 即直线 AB. 点 A 与点 C 对线段 PB 张出相等的有向角, 故 A、B、C、P 四点共圆, 即 P 在 $\triangle ABC$ 的外接圆上. □

定义 12.7 (西姆松线) 外接圆上一点 P 到三边的垂足所在的直线称为 P 对应的**西姆松线**.

西姆松定理与托勒密定理常在同一题中配合, 前者管外接圆上点的垂足共线, 后者管圆内接四边形的度量等式. 庞培定理是托勒密定理的一个干净推论.

例题 12.34 (庞培定理) 正 $\triangle ABC$ 外接圆上一点 P 到三顶点的距离 PA、PB、PC 中, 最大者等于其余两者之和.

解 设 P 在弧 BC 上且 $P\neq A$, 则 PA 最大. 对圆内接四边形 $PABC$ 用托勒密定理 (12-24),

$$PA\cdot BC=PB\cdot AC+PC\cdot AB$$

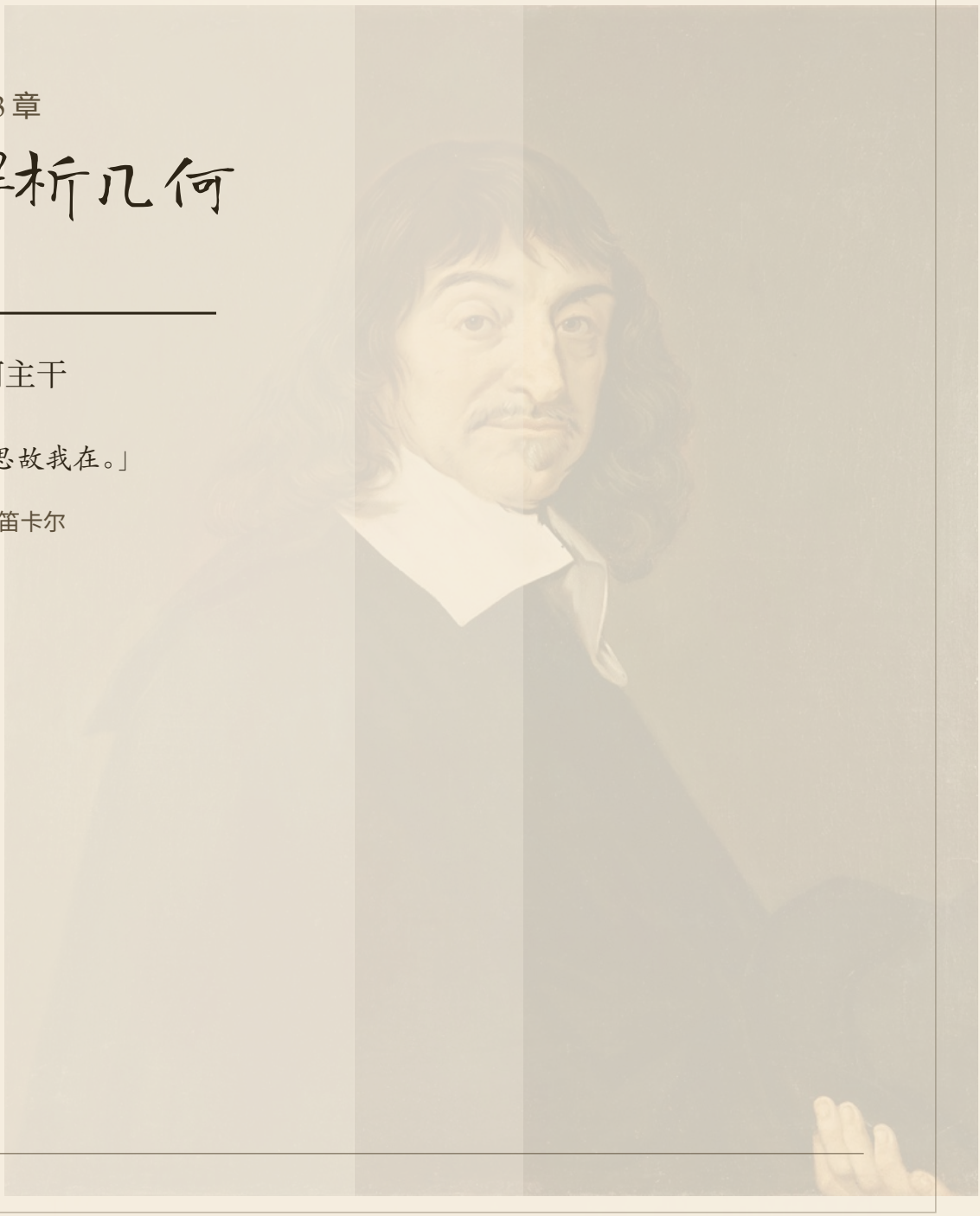
由 $AB=BC=CA$(正三角形), 两端约去边长得 $PA=PB+PC$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 13 章

解析几何



几何主干

「我思故我在。」

——笛卡尔

第十三章 解析几何

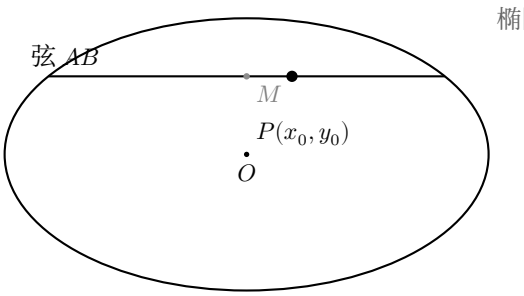


图 13.1 解析几何把几何对象翻译为坐标, 把几何关系翻译为代数方程

我们求解解析几何问题, 基本做法是把几何对象译为坐标, 把几何关系译为代数方程. 在建立方程之前, 需先选择参量. 一条路径用曲线上动点的坐标作参量, 另一条路径用直线方程的斜率与截距作参量. 前者对应本章后半部分讨论的**设点法**, 亦称点参法, 后者对应**设线法**, 亦称线参法、联立法.

注 13.1
平面向量的运算、数量积、定比分点与共线判定见第11章; 本章默认使用这些语言.

本章在向量底座之上展开. 先建立直线与圆、线性规划(选学)以及椭圆、双曲线、抛物线等二次曲线的一般理论. 随后三节讨论设点法, 依次处理其技术基础、参量选择与典型应用. 再三节讨论设线法, 依次处理联立准备、直线方程形式的选择以及联立后的代数处理. 参数方程与极坐标将另设专章, 离心角等概念仅在设点法中作为工具引用. 在此基础上, 六节技巧专题分别处理一次式与法线式、直线偶与齐次化、三角形的心与共轴圆、轨迹方程与解析法证题. 最后一节回到方法论本身, 讨论何时设点、何时设线, 并指出与射影几何等后续章节的衔接.

13.1 直线与圆

在建立了向量的代数体系之后, 我们便拥有了描述几何对象的有力工具. 此前, 几何对象的描述依赖于公理和逻辑演绎; 解析几何则把几何对象表示为代数方程, 从而将对形的研究转化为对数的运算. 将从最基本的几何图形直线与圆开始这一进程.

13.1.1 直线的方程

一条直线, 在欧几里得的观念中是没有宽度、无限延伸的长度. 用代数语言精确地捕捉这一概念, 可从不同的几何约束出发, 得到形式各异而彼此等价的方程.

描述直线方向最直观的量是其倾斜程度.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 13.1 (倾斜角与斜率) 在平面直角坐标系中, 对于一条直线 l, 我们把 x 轴正向绕 l 上任意一点逆时针旋转到与直线 l 重合时所转过的最小正角 α 称为该直线的**倾斜角**; 当 l 与 x 轴平行或重合时, 规定 $\alpha=0$. 倾斜角的取值范围是 $[0,\pi)$. 当 $\alpha\neq\frac{\pi}{2}$ 时, 倾斜角 α 的正切值 $\tan\alpha$ 称为该直线的**斜率**, 记作 m.

注意到, 当直线与 x 轴平行时, $\alpha=0, m=0$. 当直线与 x 轴垂直时, $\alpha=\frac{\pi}{2}$, 其斜率不存在.

斜率刻画了直线变化的方向, 它等于直线在竖直方向上的变化量与水平方向上的变化量之比, 这个比值与直线上点的选取无关. 设直线 l 的斜率为 m, 且经过一个定点 $P_0(x_0,y_0)$. 对于直线上异于 P_0 的任意一点 $P(x,y)$, 有

$$m=\frac{y-y_0}{x-x_0}$$

整理此式, 即得直线的**点斜式方程**:

$$y-y_0=m(x-x_0) \tag{13-1}$$

此方程表明, 一个定点 (x_0,y_0) 与一个斜率 m 即可确定一条直线.

若已知的定点是直线与 y 轴的交点 $(0,b)$, 我们称 b 为直线在 y 轴上的**截距**. 将该点坐标代入点斜式方程, 得到 $y-b=m(x-0)$, 化简即得直线的**斜截式方程**:

$$y=mx+b \tag{13-2}$$

这是将函数思想与几何图形结合最紧密的方程形式.

过两点有且只有一条直线是基本的几何公理. 设已知直线上两点 $P_1(x_1,y_1)$ 和 $P_2(x_2,y_2)$, 其中 $x_1\neq x_2$ 且 $y_1\neq y_2$. 首先, 可确定该直线的斜率:

$$m=\frac{y_2-y_1}{x_2-x_1}$$

接着, 以 $P_1(x_1,y_1)$ 为定点, 利用点斜式方程, 可得

$$y-y_1=\frac{y_2-y_1}{x_2-x_1}(x-x_1)$$

将此式变形, 得到更具对称性的**两点式方程**:

$$\frac{y-y_1}{y_2-y_1}=\frac{x-x_1}{x_2-x_1} \tag{13-3}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

若已知的两点恰好是直线与坐标轴的两个交点，即 x 截距点 $(a,0)$ 和 y 截距点 $(0,b)$，其中 a,b 均非零. 将 $x_1=a,y_1=0$ 和 $x_2=0,y_2=b$ 代入两点式方程：

$$\frac{y-0}{b-0}=\frac{x-a}{0-a}$$

$$\frac{y}{b}=\frac{x-a}{-a}=-\frac{x}{a}+1$$

整理即得直线的**截距式方程**：

$$\frac{x}{a}+\frac{y}{b}=1 \tag{13-4}$$

上述各种方程形式直观地反映了直线的几何要素，但它们都存在局限性，例如点斜式和斜截式不能表示垂直于 x 轴的直线，而截距式不能表示过原点或与坐标轴平行的直线. 需一种能够涵盖所有情况的普适形式.

将任意一种形式的方程移项整理，都可以化为 $Ax+By+C=0$ 的形式. 例如，点斜式 $y-y_0=m(x-x_0)$ 可化为 $mx-y+(y_0-mx_0)=0$. 由此定义直线的一**般式方程**：

$$Ax+By+C=0 \quad (A,B\text{不全为零}) \tag{13-5}$$

当 $B\neq 0$，方程可化为 $y=-\frac{A}{B}x-\frac{C}{B}$，即斜截式. 当 $B=0$，由于 A,B 不全为零，必有 $A\neq 0$，方程化为 $x=-\frac{C}{A}$，这是一条垂直于 x 轴的直线. 因此，一般式方程囊括了平面内所有直线.

为了更为精细地理解直线方程的内在结构，我们回归到向量的视角. 一条直线 l 可以由其上一点 P_0 和一个方向向量 $\vec{d}$ 唯一确定. 对于直线上任意一点 P，向量 $\overrightarrow{P_0P}$ 与 $\vec{d}$ 共线. 设 O 为原点，P_0,P 的位置向量分别为 $\overrightarrow{p_0},\vec{p}$，则 $\overrightarrow{P_0P}=\vec{p}-\overrightarrow{p_0}$. 于是存在实数 t 使得 $\vec{p}-\overrightarrow{p_0}=t\vec{d}$. 此即直线的**向量参数方程**：

$$\vec{p}=\overrightarrow{p_0}+t\vec{d},\quad t\in\mathbb{R}$$

在笛卡尔坐标系下，设 $\vec{p}=(x,y),\overrightarrow{p_0}=(x_0,y_0),\vec{d}=(a,b)$. 向量方程等价于坐标表示的**参数方程**：

$$\begin{cases}x=x_0+at\\ y=y_0+bt\end{cases}\qquad(t\in\mathbb{R})$$

若 a,b 均不为零，消去参数 t 可得 $\frac{x-x_0}{a}=\frac{y-y_0}{b}$，这被称为**对称式方程**，整理即得 $b(x-x_0)-a(y-y_0)=0$，也就是 $bx-ay+(ay_0-bx_0)=0$. 这是 $Ax+By+C=0$ 的形式，其中 $A=b,B=-a,C=ay_0-bx_0$.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

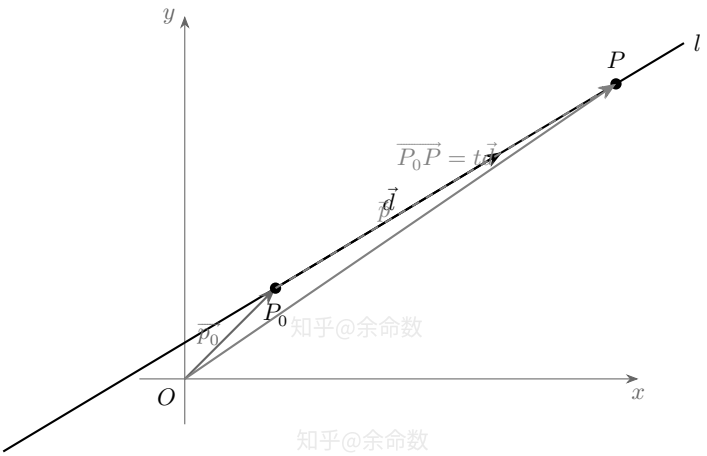


图 13.2　直线的向量参数表示

同时, 直线的方向亦可由一个与之垂直的**法向量** $\vec{n}$ 确定. 设直线过点 P_0, 法向量为 $\vec{n}$. 对线上任意点 P, 向量 $\overrightarrow{P_0P}$ 恒与 $\vec{n}$ 正交, 故其数量积为零.

$$\vec{n} \cdot \overrightarrow{P_0P} = 0 \quad \text{或} \quad \vec{n} \cdot (\vec{p} - \overrightarrow{p_0}) = 0$$

设 $\vec{n} = (A, B), \vec{p} = (x, y), \overrightarrow{p_0} = (x_0, y_0)$, 上述**点法式方程**展开为 $A(x - x_0) + B(y - y_0) = 0$, 即 $Ax + By - (Ax_0 + By_0) = 0$. 令 $C = -(Ax_0 + By_0)$, 这又一次导出了直线的一般形式 $Ax + By + C = 0$. 这一推导说明了一个基本事实, 即**在直线的一般式方程 $Ax + By + C = 0$ 中, 系数 A 和 B 恰好构成该直线的一个法向量 $\vec{n} = (A, B)$.**

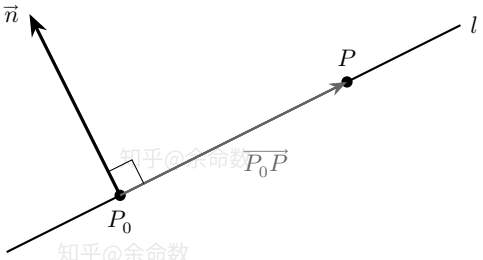


图 13.3　直线的法向量表示法

13.1.2 圆的方程

圆是继直线之后最简单的曲线. 其定义蕴含着一种纯粹的对称性.

定义 13.2 (圆) 平面内到一个定点 C 的距离等于一个定长 R 的所有点 P 的集合, 称为**圆**. 定点 C 称为**圆心**, 定长 R 称为**半径**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.1.2.1 圆的标准方程

这一定义可以直接翻译为代数语言. 引入笛卡尔坐标系, 设圆心 C 的坐标为 (h,k), 圆上任意一点 P 的坐标为 (x,y), 半径为 $R>0$. 把点 P 到点 C 的距离等于 R 用距离公式写出, 即得

$$\sqrt{(x-h)^2+(y-k)^2}=R$$

由于距离 R 是非负的, 可对等式两边进行平方, 得到一个等价的、形式更简洁的方程, 此即圆的**标准方程**:

$$(x-h)^2+(y-k)^2=R^2 \tag{13-6}$$

此方程直观地展现了圆的三个几何要素: 圆心横坐标 h, 纵坐标 k 和半径 R. 特别地, 当圆心在原点 $(0,0)$ 时, 方程简化为 $x^2+y^2=R^2$.

从向量的角度看, 设圆心 C 和圆上一点 P 的位置向量分别为 $\vec{c}$ 和 $\vec{p}$. 定义 $|\vec{p}-\vec{c}|=R$ 即为圆的向量方程. 这也等价于 $(\vec{p}-\vec{c})\cdot(\vec{p}-\vec{c})=R^2$, 它与坐标形式的方程是完全一致的.

13.1.2.2 圆的一般方程

将圆的标准方程 $(x-h)^2+(y-k)^2=R^2$ 展开, 得

$$x^2-2hx+h^2+y^2-2ky+k^2-R^2=0$$

整理后, 方程具有以下形式:

$$x^2+y^2+Dx+Ey+F=0$$

其中 $D=-2h, E=-2k, F=h^2+k^2-R^2$. 这被称为圆的**一般方程**. 其特点是:

- x^2 和 y^2 的系数相同, 且通常化为 1.
- 不含 xy 这样的交叉项.

自然要问, 是否任何形如 $x^2+y^2+Dx+Ey+F=0$ 的方程都表示一个圆; 即给定一组系数 D,E,F, 能否找到对应的圆心 (h,k) 和半径 R, 使得该方程与某个圆的标准方程等价.

为了回答这个问题, 我们采用配方法, 试图将一般方程还原为标准形式. 配方法的做法是, 把二次项和一次项组合成一个完全平方项, 使方程的形式更清晰.

$$x^2+y^2+Dx+Ey+F=0$$

$$(x^2+Dx)+(y^2+Ey)=-F$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为了配方, 需在 (x^2+Dx) 中加上 $(\frac{D}{2})^2=\frac{D^2}{4}$, 在 (y^2+Ey) 中加上 $(\frac{E}{2})^2=\frac{E^2}{4}$. 为了保证等式依然成立, 需在方程的右侧也加上相应的项:

$$\left(x^2+Dx+\frac{D^2}{4}\right)+\left(y^2+Ey+\frac{E^2}{4}\right)=-F+\frac{D^2}{4}+\frac{E^2}{4}$$

$$\left(x+\frac{D}{2}\right)^2+\left(y+\frac{E}{2}\right)^2=\frac{D^2+E^2-4F}{4}$$

方程的左侧是两个完全平方项的和, 形式上已经非常接近圆的标准方程. 为了使形式完全一致, 将方程右侧的表达式记作 $\frac{\Delta_c}{4}$, 其中 $\Delta_c=D^2+E^2-4F$.

$$\left(x+\frac{D}{2}\right)^2+\left(y+\frac{E}{2}\right)^2=\frac{\Delta_c}{4}$$

进一步变形, 得到

$$\left(x-\left(-\frac{D}{2}\right)\right)^2+\left(y-\left(-\frac{E}{2}\right)\right)^2=\left(\frac{\sqrt{\Delta_c}}{2}\right)^2$$

方程的几何意义完全取决于 Δ_c 的符号:

- **若** $\Delta_c>0$: 此时, $\frac{\sqrt{\Delta_c}}{2}$ 是一个正实数, 可以作为圆的半径 R. 方程表示一个圆心为 $(-\frac{D}{2},-\frac{E}{2})$, 半径为 $R=\frac{\sqrt{\Delta_c}}{2}$ 的圆. 即, 当且仅当 $D^2+E^2>4F$ 时, 一般方程 $x^2+y^2+Dx+Ey+F=0$ 才表示一个真正的圆.
- **若** $\Delta_c=0$: 此时, 方程变为

  $$\left(x+\frac{D}{2}\right)^2+\left(y+\frac{E}{2}\right)^2=0$$

  由于任何实数的平方都是非负的, 两个非负数的和为零, 意味着它们必须同时为零. 因此, $x=-\frac{D}{2}$ 且 $y=-\frac{E}{2}$. 方程只有一个解, 表示一个点 $(-\frac{D}{2},-\frac{E}{2})$. 可将这个点视为半径为 0 的退化圆, 即圆心与圆周重合.
- **若** $\Delta_c<0$: 此时, 方程变为

  $$\left(x+\frac{D}{2}\right)^2+\left(y+\frac{E}{2}\right)^2=\frac{\Delta_c}{4}<0$$

  由于任何实数的平方都是非负的, 两个非负数的和也必然是非负的. 因此, 方程的左侧恒为非负, 而右侧为负数. 即, 不存在任何实数 x 和 y 能够满足该方程. 因此, 方程没有实数解, 不表示任何几何图形.

一般方程 $x^2+y^2+Dx+Ey+F=0$ 只有在满足条件 $D^2+E^2>4F$ 时, 才表示一个圆. 此时, 圆心为 $(-\frac{D}{2},-\frac{E}{2})$, 半径为 $\frac{\sqrt{D^2+E^2-4F}}{2}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

13.1.3 直线与圆的位置关系

研究直线与圆的位置关系, 实质上是研究它们的方程组解的情况. 设直线 $l: Ax+By+C=0$ 与圆 $C:(x-h)^2+(y-k)^2=R^2$. 从代数角度看, 可联立这两个方程, 消去一个变量, 例如 y, 得到一个关于另一个变量, 例如 x 的一元二次方程. 该二次方程的判别式 Δ 的符号决定了交点的个数: $\Delta>0$ 对应两个交点, 即相交; $\Delta=0$ 对应一个交点, 即相切; $\Delta<0$ 对应无交点, 即相离.

然而, 这种纯代数的方法往往计算繁琐, 且掩盖了问题内在的几何结构. 一种更直接、更具几何意义的判别方法是比较圆心到直线的距离与半径的大小.

13.1.3.1 点到直线的距离

需一个公式来计算点 $P_0(x_0,y_0)$ 到直线 $l: Ax+By+C=0$ 的距离 d. 设 $P_1(x_1,y_1)$ 是直线 l 上任意一点. 我们要求的是向量 $\overrightarrow{P_1P_0}$ 在直线法向量 $\vec{n}=(A,B)$ 方向上的投影的长度.

这个投影的长度 d 就是 P_0 到直线 l 的距离. 根据向量投影的公式:

$$d=\frac{|\overrightarrow{P_1P_0}\cdot\vec{n}|}{|\vec{n}|}$$

将坐标代入: $\overrightarrow{P_1P_0}=(x_0-x_1,y_0-y_1)$, $\vec{n}=(A,B)$.

$$d=\frac{|A(x_0-x_1)+B(y_0-y_1)|}{\sqrt{A^2+B^2}}=\frac{|Ax_0+By_0-(Ax_1+By_1)|}{\sqrt{A^2+B^2}}$$

由于点 $P_1(x_1,y_1)$ 在直线 l 上, 它的坐标满足直线方程, 即 $Ax_1+By_1+C=0$, 从而 $Ax_1+By_1=-C$. 代入上式, 即得点到直线的距离公式.

定理 13.1 (点到直线的距离) 点 $P_0(x_0,y_0)$ 到直线 $l: Ax+By+C=0$, 其中 A,B 不同时为 0 的距离为

$$d=\frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}$$

证明 因为 A,B 不同时为 0, 所以法向量 $\vec{n}=(A,B)$ 非零, 且直线 l 上可取一点 $P_1(x_1,y_1)$. 点 P_0 到直线的距离等于向量 $\overrightarrow{P_1P_0}$ 在法向量方向上的投影长度:

$$d=\frac{|\overrightarrow{P_1P_0}\cdot\vec{n}|}{|\vec{n}|}=\frac{|A(x_0-x_1)+B(y_0-y_1)|}{\sqrt{A^2+B^2}}.$$

又 P_1 在直线 l 上, 故 $Ax_1+By_1+C=0$, 即 $Ax_1+By_1=-C$. 代入上式即得

$$d=\frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

□

13.1.3.2 几何判别法

设圆 C 的圆心为 C, 半径为 R. 设直线为 l. 我们考察圆心 C 到直线 l 的距离, 记为 d. 设 H 是从 C 到 l 的垂足, 那么根据定义, $d = |CH|$. 接着考虑直线上任意一点 P. 点 P 是直线与圆的公共点, 当且仅当点 P 到圆心 C 的距离为 R, 即 $|CP| = R$.

在直角三角形 $\triangle CHP$ 中(若 P 与 H 不重合), 根据勾股定理, 有

$$|CP|^2 = |CH|^2 + |HP|^2$$

代入 d 和 R 的定义, 我们寻求满足下式的点 P:

$$R^2 = d^2 + |HP|^2$$

这个方程连接着代数关系与几何直观. 分析此方程解的存在性, 即可判断公共点的个数. 将方程变形为:

$$|HP|^2 = R^2 - d^2$$

方程的解, 即点 P 的位置, 取决于线段 HP 的长度 $|HP|$, 而 $|HP|$ 的值则完全由 $R^2 - d^2$ 的符号决定.

- **若** $d < R$, 则 $d^2 < R^2$, 从而 $R^2 - d^2 > 0$. 此时, 方程 $|HP|^2 = R^2 - d^2$ 有唯一正实数解 $|HP| = \sqrt{R^2 - d^2}$. 即在直线 l 上, 存在两个关于垂足 H 对称的点 P_1 和 P_2, 使得它们到 H 的距离等于这个正值. 这两个点 P_1, P_2 均满足 $|CP_1| = |CP_2| = R$, 故它们是直线上仅有的两个在圆上的点. 因此, 直线与圆**相交**, 有两个公共点. 此时直线称为圆的**割线**.
- **若** $d = R$, 则 $d^2 = R^2$, 从而 $R^2 - d^2 = 0$. 此时, 方程变为 $|HP|^2 = 0$, 其唯一解为 $|HP| = 0$. 即点 P 必须与垂足 H 重合. 因此, 直线与圆有且仅有一个公共点, 即垂足 H. 此时直线与圆**相切**, 该公共点称为**切点**.
- **若** $d > R$, 则 $d^2 > R^2$, 从而 $R^2 - d^2 < 0$. 此时, 方程 $|HP|^2 = R^2 - d^2$ 的右侧为负数. 由于任何实数的平方都不可能为负, 该方程没有实数解. 即在直线 l 上不存在任何点 P 能够满足 $|CP| = R$. 因此, 直线与圆**相离**, 没有公共点.

13.1.3.3 圆的切线

切线是位置关系中的临界状态, 具有简洁的几何性质. 注意到当直线与圆相切时, 圆心到直线的距离等于半径, 且这段距离本身就是由圆心到切点的半径. 即, **过切点的半径所在的直线与切线垂直**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

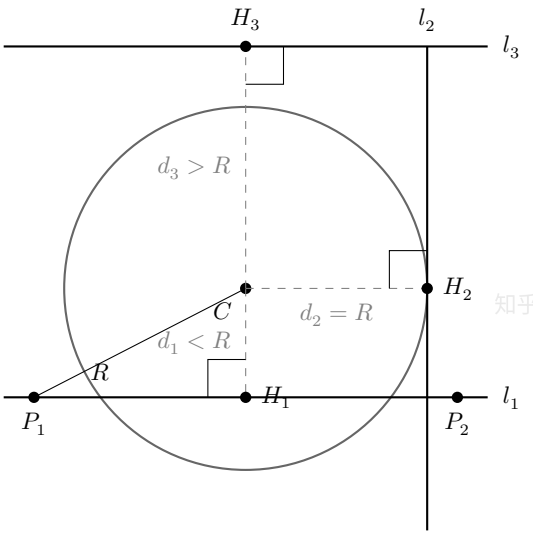


图 13.4 直线与圆的三种位置关系及其几何判定

这个性质提供一种求切线方程的实用方法. 设圆的标准方程为 $(x-h)^2+(y-k)^2=R^2$, 切点为 $P_0(x_0,y_0)$. 向量 $\overrightarrow{CP_0}=(x_0-h,y_0-k)$ 就是切线的一个法向量. 根据直线的点法式方程, 切线方程为

$$(x_0-h)(x-x_0)+(y_0-k)(y-y_0)=0$$

为了得到一个更具对称性的形式, 对上式进行代数变换:

$$(x_0-h)(x-h-(x_0-h))+(y_0-k)(y-k-(y_0-k))=0$$

$$(x_0-h)(x-h)-(x_0-h)^2+(y_0-k)(y-k)-(y_0-k)^2=0$$

移项得到

$$(x_0-h)(x-h)+(y_0-k)(y-k)=(x_0-h)^2+(y_0-k)^2$$

由于点 $P_0(x_0,y_0)$ 在圆上, 它满足圆的方程, 即 $(x_0-h)^2+(y_0-k)^2=R^2$. 于是, 得圆的切线方程的**标准形式**:

$$(x_0-h)(x-h)+(y_0-k)(y-k)=R^2 \tag{13-7}$$

这个形式的特点在于, 它是将圆的标准方程 $(x-h)^2+(y-k)^2=R^2$ 中的一个 $(x-h)$ 和一个 $(y-k)$ 分别替换为 (x_0-h) 和 (y_0-k) 得到的, 这种代换模式在解析几何的其他二次曲线的切线问题中也普遍存在.

知乎@余命数

13.1.4 弦长、圆幂与两圆位置

前面给出了位置关系的几何判别. 再补充几件在解题中常用的工具: 弦长公式、圆幂、两圆位置与公切线, 以及直径所对圆周角. 它们既是直线与圆位置关系的自然延伸, 也是 第13.3节 中二次曲线弦长/点差的原型.

13.1.4.1 弦长公式

设圆 C 圆心为 C, 半径 R, 直线 l 与圆交于 P_1, P_2, 弦心距 $d = |CH| < R$(H 为垂足). 由直角三角形 CHP_1:

知乎@余命数

$$|P_1P_2| = 2\sqrt{R^2 - d^2}. \tag{13-8}$$

若采用设线法, 将 $y = kx + m$ 与圆方程联立得 $ax^2 + bx + c = 0$, 则弦长亦可写为

$$|P_1P_2| = \sqrt{1 + k^2}\,\frac{\sqrt{\Delta}}{|a|},$$

其中 $\Delta = b^2 - 4ac$. 两式等价, 几何式直接用 d, 代数式直接用韦达, 选型见 第13.7节.

知乎@余命数

例题 13.1 (弦长的几何与代数对照) 圆 $x^2 + y^2 = 25$, 直线 $x - y + 1 = 0$. 求弦长.

解 圆心 $O(0,0)$, $R = 5$. 弦心距

知乎@余命数

知乎@余命数

$$d = \frac{|0 - 0 + 1|}{\sqrt{1^2 + (-1)^2}} = \frac{1}{\sqrt{2}}.$$

故 $|P_1P_2| = 2\sqrt{25 - \dfrac{1}{2}} = 2\sqrt{\dfrac{49}{2}} = 7\sqrt{2}$.

代数验算: 联立 $y = x + 1$ 与 $x^2 + (x+1)^2 = 25$, 得 $2x^2 + 2x - 24 = 0$, 即 $x^2 + x - 12 = 0$、$\Delta = 49$、$a = 1$、$k = 1$,

$$|P_1P_2| = \sqrt{2} \cdot \frac{7}{1} = 7\sqrt{2}.$$

13.1.4.2 圆幂与相交弦

定义 13.3 (点对圆的幂) 定点 P 对圆 C 的**幂**由过 P 的任一条与圆相交的直线的有向乘积 $\overrightarrow{PA} \cdot \overrightarrow{PB}$ 定义, 它与直线如何选取无关. 若圆方程为 $x^2 + y^2 + Dx + Ey + F = 0$, 点 $P(x_0, y_0)$ 的幂等于把点代入圆方程左边所得值(对标准圆 $(x-h)^2 + (y-k)^2 - R^2$ 则为 $|PC|^2 - R^2$).

经典几何形式:

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **相交弦定理**：圆内两弦 AB, CD 交于 P，则 $|PA|\cdot|PB| = |PC|\cdot|PD|$；
- **切割线定理**：自圆外一点 P 作割线交圆于 A, B，切线切于 T，则 $|PA|\cdot|PB| = |PT|^2$.

例题 13.2 (切割线) 圆 $x^2+y^2=4$，点 $P(4,0)$. 过 P 作圆的切线，求切线长；再过 P 作割线交圆于 A, B，验证 $|PA|\cdot|PB|$ 等于切线长的平方.

解 幂 $= |PO|^2 - R^2 = 16-4 = 12$，故切线长 $\sqrt{12} = 2\sqrt{3}$. 取水平割线 $y=0$ 与圆交于 $(\pm 2, 0)$，则 $A=(-2,0)$，$B=(2,0)$（有向时一端为负），乘积的绝对值 $|4-(-2)|\cdot|4-2|$ 在无向长度下为 $6\cdot 2 = 12$，与幂一致.

13.1.4.3 两圆的位置关系与公切线

设两圆圆心距为 d，半径 R_1, R_2 $(R_1 \ge R_2 > 0)$. 则

- $d > R_1 + R_2$：外离（4 条公切线：2 外切，2 内切）；
- $d = R_1 + R_2$：外切（3 条公切线）；
- $|R_1 - R_2| < d < R_1 + R_2$：相交于两点（2 条外公切线）；
- $d = |R_1 - R_2|$：内切（1 条外公切线）；
- $d < |R_1 - R_2|$：内含（无公切线，同心时 $d=0$ 另论）.

求公切线可用代数方法，设切线 $y = kx + m$（或 $x = my + n$），令圆心到直线距离等于相应半径，解 (k, m)；竖直切线单独讨论.

例题 13.3 (两圆公切线条数) 圆 $C_1: x^2+y^2=1$ 与 $C_2: (x-3)^2+y^2=1$. 判断位置关系并说明公切线条数.

解 圆心距 $d=3$、$R_1=R_2=1$、$R_1+R_2=2$、$|R_1-R_2|=0$. 因 $d>2$，两圆外离，有 4 条公切线.

13.1.4.4 直径所对的圆周角

定理 13.2 (Thales) 若 AB 为圆的直径，P 为圆上异于 A, B 的点，则 $\angle APB = \dfrac{\pi}{2}$，即 $\overrightarrow{PA}\cdot\overrightarrow{PB} = 0$.

证明 取坐标系使圆心为原点、直径 AB 沿 x 轴：$A=(-R,0)$，$B=(R,0)$，圆上点 $P=(R\cos\theta, R\sin\theta)$ $(\sin\theta \ne 0)$. 则

$$\overrightarrow{PA} = R(-1-\cos\theta, -\sin\theta), \qquad \overrightarrow{PB} = R(1-\cos\theta, -\sin\theta).$$

知乎@余命数

点积

$$\overrightarrow{PA}\cdot\overrightarrow{PB}=R^2[(1+\cos\theta)(-1+\cos\theta)+\sin^2\theta]=R^2[\cos^2\theta-1+\sin^2\theta]=0.$$

故 $\angle APB=\dfrac{\pi}{2}$. □

这一事实在 第13.3节 与设点法中反复出现:“以 PQ 为直径的圆过定点 F” $\Longleftrightarrow$ $\overrightarrow{FP}\cdot\overrightarrow{FQ}=0$.

例题 13.4 (弦心距与中点) 圆 $x^2+y^2-4x-2y-4=0$ 被直线 $x+y-1=0$ 所截得的弦的中点为 M. 求 M 的坐标与弦长.

解 配方得圆心 $C(2,1)$, $R=3$. 弦的中点 M 是圆心到直线的垂足, 故 $CM\perp$ 直线. 直线法向量 $(1,1)$, 故

$$M=(2,1)+t(1,1)=(2+t,1+t),$$

且 M 在直线上: $(2+t)+(1+t)-1=0 \implies 2+2t=0 \implies t=-1$. 于是 $M=(1,0)$. 弦心距 $d=|CM|=\sqrt{2}$, 弦长 $2\sqrt{9-2}=2\sqrt{7}$.

13.1.5 圆中范围与最值例题

圆上动点引发的 range 与最值问题, 通常要把目标式翻译为斜率、截距或距离, 再用相切条件求极值.

例题 13.5 (斜率型) 已知实数 x,y 满足方程 $y=\sqrt{-x^2+4x-1}$, 则 $\frac{y}{x}$ 的最大值为(　)

A. 0　　C. $\sqrt{3}$

B. 1　　D. 2

解 答案:C. 方程 $y=\sqrt{-x^2+4x-1}$ 化为 $(x-2)^2+y^2=3(y\geqslant 0)$, 表示的图形是一个以 $(2,0)$ 为圆心, $\sqrt{3}$ 为半径的半圆; 令 $\frac{y}{x}=k$, 即 $y=kx$, 如图所示, 当直线与半圆相切时, 圆心到直线的距离 $d=\frac{|2k|}{\sqrt{1+k^2}}=\sqrt{3}$, 解得 $k=\sqrt{3}$ 或 $k=-\sqrt{3}$(负值不满足条件, 舍去), 所以 $\frac{y}{x}$ 的最大值为 $\sqrt{3}$, 故选:C.

例题 13.6 (斜率型) 如果实数 x,y 满足 $(x-2)^2+y^2=2$, 则 $\frac{y}{x}$ 的范围是(　)

A. $(-1,1)$　　C. $(-\infty,-1)\cup(1,+\infty)$

B. $[-1,1]$　　D. $(-\infty,-1]\cup[1,+\infty)$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 答案: B. 设 $\frac{y}{x}=k$, 则 $y=kx$ 表示经过原点的直线, k 为直线的斜率. 如果实数 x,y 满足 $(x-2)^2+y^2=2$ 和 $\frac{y}{x}=k$, 即直线 $y=kx$ 同时经过原点和圆上的点 (x,y). 其中圆心 $C(2,0)$, 半径 $r=\sqrt{2}$ 从图中可知, 斜率取最大值时对应的直线斜率为正且刚好与圆相切, 设此时切点为 E

例题 13.7 (直线型) (2024·江西吉安·宁冈中学校考一模) 已知点 $P(x,y)$ 是圆 $x^2+y^2-6x-4y+12=0$ 上的动点, 则 $x+y$ 的最大值为(　)

A. $5+\sqrt{2}$　　C. 6

B. $5-\sqrt{2}$　　D. 5

解 答案: A. 由 $(x-3)^2+(y-2)^2=1$, 令 $\begin{cases}x=3+\cos\theta\\ y=2+\sin\theta\end{cases}$, 则 $x+y=5+\sqrt{2}\sin\left(\theta+\frac{\pi}{4}\right)$, 所以当 $\sin\left(\theta+\frac{\pi}{4}\right)=1$ 时, $x+y$ 的最大值为 $5+\sqrt{2}$. 故选: A

例题 13.8 (直线型) 已知点 $P(x,y)$ 是圆 $C:(x-a)^2+y^2=3(a>0)$ 上的一动点, 若圆 C 经过点 $A(1,\sqrt{2})$, 则 $y-x$ 的最大值与最小值之和为(　)

A. 4　　C. -4

B. $2\sqrt{6}$　　D. $-2\sqrt{6}$

解 答案: C. 因为圆 $C:(x-a)^2+y^2=3(a>0)$ 经过点 $A(1,\sqrt{2})$, $(1-a)^2+2=3$, 又 $a>0$, 所以 $a=2$, $y-x$ 可看成是直线 $y=x+b$ 在 y 轴上的截距. 如图所示, 当直线 $y=x+b$ 与圆相切时, 纵截距 b 取得最大值或最小值, 此时 $\frac{|2-0+b|}{\sqrt{2}}=\sqrt{3}$, 解得 $b=-2\pm\sqrt{6}$, 所以 $y-x$ 的最大值为 $-2+\sqrt{6}$, 最小值为 $-2-\sqrt{6}$, 故 $y-x$ 的最大值与最小值之和为 -4. 故选: C.

例题 13.9 (距离型) 已知点 $P(m,n)$ 在圆 $C:(x-2)^2+(y-2)^2=9$ 上运动, 则 $(m+2)^2+(n+1)^2$ 的最大值为__, 最小值为__, $\sqrt{m^2+n^2}$ 的范围为__.

解 答案: $64;4;[3-2\sqrt{2},3+2\sqrt{2}]$. 由圆 C 的圆心为 $(2,2)$, 半径为 3, 且 P 在圆 C 上, 则 $(m+2)^2+(n+1)^2$ 表示在圆 C 上点到 $(-2,-1)$ 距离的平方. 而圆心到 $(-2,-1)$ 的距离为 $\sqrt{[2-(-2)]^2+[2-(-1)]^2}=5>3$, 所以在圆 C 上点到 $(-2,-1)$ 距离的最大值为 8, 最小值为 2, 故 $(m+2)^2+(n+1)^2$ 的最大值为 64, 最小值为 4; 又 $\sqrt{m^2+n^2}$ 表示在圆 C 上点到原点的距离, 而圆心到原点距离为 $2\sqrt{2}<3$,

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.10 (周长面积型) (2024·高三·河南·开学考试)若直线 $l:kx-y+2-k=0$ 与圆 $C:x^2+y^2-4x-2y-4=0$ 交于 A,B 两点, 则当 $\triangle ABC$ 周长最小时, $k=$(　)

A. $\frac{1}{2}$　　C. 1

B. $-\frac{1}{2}$　　D. -1

解 答案: C. 直线 $l:kx-y+2-k=0$ 的方程可化为 $y-2=k(x-1)$ 所以直线 l 恒过定点 $D(1,2)$, 因为 $1^2+2^2-4\times1-2\times2-4=-7<0$

例题 13.11 (数量积型) 已知 PQ,MN 是半径为 5 的圆 O 上的两条动, $|\overrightarrow{PQ}|=6,|\overrightarrow{MN}|=8$, 则 $|\overrightarrow{PM}+\overrightarrow{QN}|$ 最大值是(　)

A. 7　　C. 14

B. 12　　D. 16

余命数
知乎@余命数

解 答案: C. 如图, 连接 MO,OQ,OP,ON, 作 $PQ\perp OE,MN\perp OD$, 易知 E 是 QP 的中点, D 是 MN 的中点, 由股定理得 $OE=4,OD=3$, 故 $\overrightarrow{PM}+\overrightarrow{QN}=\overrightarrow{OM}-\overrightarrow{OP}+\overrightarrow{ON}-\overrightarrow{OQ}=(\overrightarrow{OM}+\overrightarrow{ON})-(\overrightarrow{OP}+\overrightarrow{OQ})=2(\overrightarrow{OD}-\overrightarrow{OE})$, 故 $|\overrightarrow{PM}+\overrightarrow{QN}|=2|\overrightarrow{OD}-\overrightarrow{OE}|\leqslant2(|\overrightarrow{OD}|+|\overrightarrow{OE}|)=14$, 当 $\overrightarrow{OE},\overrightarrow{OD}$ 反向时等号成立, 故 C 正确. 故选: C

例题 13.12 (坐标与角度型) 已知圆 $C:(x-1)^2+y^2=1$, 点 $P(x_0,y_0)$ 在直线 $x-y+1=0$ 上运动. 若 C 上存在点 Q, 使 $\angle CPQ=30^\circ$, 则 x_0 的取值范围是__.

解 答案: $[-1,1]$. 圆 $C(1,0)$, P 在直线 $x-y+1=0$ 上, 若圆存在点 Q, 使得 $\angle CPQ=30^\circ$, 当 P 在直线 $x-y+1=0$ 上运动, 极端情况, PQ 与圆 C 相切, $\angle CPQ=30^\circ$. 在 $RT\triangle CPQ$ 中, $|CQ|=1$, 所以 $|CP|=2$. 所以以 $(1,0)$ 为圆心, 2 为半径的圆与直线交于 P,P_1 两点. 符合条件的点在线段 PP_1 之间. 所以 $\begin{cases}x-y+1=0\\(x-1)^2+y^2=4\end{cases}\Rightarrow\begin{cases}x=1\\y=2\end{cases}$ 或 $\begin{cases}x=-1\\y=0\end{cases}$ 故 x_0 的取值范围为 $[-1,1]$. 故答案为: $[-1,1]$

例题 13.13 (坐标与角度型) 已知 x,y 满足 $x^2+y^2=4y-3$, 则 $\dfrac{\sqrt{3}x+y}{\sqrt{x^2+y^2}}$ 的最大值为(　)

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

A. 1　　C. $\sqrt{3}$

B. 2　　D. $\sqrt{5}$

解 答案:C.

例题 13.14 (长度和差型) 已知复数 $z_1=a+bi, z_2=c+di, a,b,c,d\in R$, 若 $|z_1|=|z_2|=2$, 且 $ac+bd=2$, 则 $|a+b-4|+2|c+d|$

例题 13.15 (方程中的参数型) (2024·山东泰安·二模) 已知在矩形 $ABCD$ 中, $AB=1, AD=\sqrt{3}$,动点 P 在以点 C 为圆心且与 BD 相切的圆上,则 $\vec{AP}\cdot\vec{AD}$ 的最大值为__;若 $\vec{AP}=m\vec{AB}+n\vec{AD}(m,n\in R)$, 则 $m+n$ 的最大值为__.

解 答案: $\frac{9}{2}, 3$. 如图, 以点为原点, 以 BA, BC 所在的直线为 y, x 轴建立如图所示的坐标系, 则 $B(0,0)$、$A(0,1)$、$D(\sqrt{3},1)$、$C(\sqrt{3},0)$、$\vec{AD}=(\sqrt{3},0)\therefore$ 动点 P 在以点 C 为圆心且与 BD 相切的圆上, 设圆的半径为 r, $\therefore BC=\sqrt{3}, CD=1, \therefore BD=\sqrt{(\sqrt{3})^2+1^2}=2\therefore\frac{1}{2}BC\cdot CD=\frac{1}{2}BD\cdot r, \therefore r=\frac{\sqrt{3}}{2}\therefore$ 圆的方程为 $(x-\sqrt{3})^2+y^2=\frac{3}{4}$,设点 P 的坐标为 $(\frac{\sqrt{3}}{2}\cos\theta+\sqrt{3}, \frac{\sqrt{3}}{2}\sin\theta), \theta\in[0,2\pi]$,则 $\vec{AP}=(\frac{\sqrt{3}}{2}\cos\theta+\sqrt{3}, \frac{\sqrt{3}}{2}\sin\theta-1)$, $\vec{AP}\cdot\vec{AD}=\sqrt{3}(\frac{\sqrt{3}}{2}\cos\theta+\sqrt{3})=\frac{3}{2}\cos\theta+3\in[-\frac{3}{2}, \frac{9}{2}]$, 故 $\vec{AP}\cdot\vec{AD}$ 的最大值为 $\frac{9}{2}$; $\therefore\vec{AP}=m\vec{AB}+n\vec{AD}(m,n\in R)$, $\vec{AB}=(0,-1)$,

例题 13.16 (多选题) (多选题) 设点 $P(x,y)$ 为圆 $C: x^2+y^2=1$ 上一点, 已知点 $A(4,0), B(5,0)$, 则下列结论正确的有(　)

A. $x+y$ 的最大值为 $\sqrt{2}$

B. $x^2+y^2-4x-4y$ 的最小值为 -8

C. 存在点 P 使 $|PB|=\sqrt{2}|PA|$

D. 过 A 点作圆 C 的切线, 则切线长为 $\sqrt{15}$

解 答案: AD. 对于 A, 设 $x+y=k$, 则 $C(0,0)$ 点到直线 $x+y=k$ 的距离 $\frac{|0+0-k|}{\sqrt{2}}\leqslant 1$, 解得 $-\sqrt{2}\leqslant k\leqslant\sqrt{2}$, 得 k 的最大值为 $\sqrt{2}$, 故 A 正确; 对于 B, 令 $x^2+y^2-4x-4y=1-4x-4y=k$, 则 $C(0,0)$ 点到直线 $1-4x-4y=k$ 的距离 $\frac{|1-k|}{\sqrt{32}}\leqslant 1$, 解得 $-\sqrt{32}+1\leqslant k\leqslant\sqrt{32}+1$, 得 k 的最小值为 $-4\sqrt{2}+1$, 故 B 错误; 对于 C, 假设存在点 P 使 $|PB|=\sqrt{2}|PA|$, 设 $P(x,y)$, 则 $\sqrt{(x-5)^2+y^2}=\sqrt{2}\sqrt{(x-4)^2+y^2}$,化简得 $(x-3)^2+y^2=2$,因此满足 $|PB|=\sqrt{2}|PA|$ 的点 P 在圆 $(x-3)^2+y^2=2$ 上,此圆圆心为 $(3,0)$,半径为 $\sqrt{2}$,而 $\sqrt{2}+1<3$,因此与圆 C 外离, 所以不存在点 P 使 $|PB|=\sqrt{2}|PA|$, 故 C 错误; 对于 D, 圆 C 的圆心为 $C(0,0)$, 半

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

径为 $r=1$, 则过 A 点作圆 C 的切线, 则切线长为 $\sqrt{|AC|^2-r^2}=\sqrt{4^2-1}=\sqrt{15}$, 故 D 正确. 故选: AD.

例题 13.17 (多选题) (2024·高三·辽宁鞍山·开学考试)已知直线 $kx-y+k=0$, 圆 $C: x^2+y^2-6x+5=0$, $P(x_0,y_0)$ 为圆 C 上任意一点, 则下列说法正确的是(　)

A. $x_0^2+y_0^2$ 的最大值为 5
B. $\frac{y_0}{x_0}$ 的最大值为 $\frac{2\sqrt{5}}{5}$
C. 直线 l 与圆 C 相切时, $k=\pm\frac{\sqrt{3}}{3}$
D. 圆心 C 到直线 l 的距离最大为 4

解 答案: BC. 圆 C 的方程可化为 $(x-3)^2+y^2=2^2$, 所以圆 C 的圆心为 $C(3,0)$, 半径 $r=2$、$|OC|=3$、$P(x_0,y_0)$ 是圆上的点, 所以 $x_0^2+y_0^2$ 的最大值为 $(3+2)^2=25$, A 选项错误.

例题 13.18 若实数 x,y 满足条件 $x^2+y^2=1$, 则 $\frac{x+y-1}{x+1}$ 的范围是(　)

A. $[\frac{1}{4},+\infty)$
B. $(-\infty,\frac{1}{4}]$
C. $(-\infty,\frac{1}{2}]$
D. $[\frac{1}{2},+\infty)$

解 答案: B. 令 $\frac{x+y-1}{x+1}=k$, 可得 $(k-1)x-y+k+1=0$, 则直线 $(k-1)x-y+k+1=0$ 与圆 $x^2+y^2=1$ 有公共点, 所以, $\frac{|k+1|}{\sqrt{(k-1)^2+1}}\leq 1$, 解得 $k\leq\frac{1}{4}$, 即 $\frac{x+y-1}{x+1}$ 的取值范围是 $(-\infty,\frac{1}{4}]$, 故选: B.

例题 13.19 (2024·山东日照·二模)若实数 x,y 满足条件 $x^2+y^2=1$, 则 $\frac{y-2}{x+1}$ 的范围是(　)

A. $[0,\sqrt{2}]$
B. $[-3,5]$
C. $(-\infty,-1]$
D. $(-\infty,-\frac{3}{4}]$

解 答案: D. $\frac{y-2}{x+1}$ 的几何意义即圆上的点 (x,y) 到定点 $(-1,2)$ 的斜率, 由图知, 斜率的范围处在圆的两条切线率之间, 其中 AC 斜率不存在, 设 AB 的斜率为 k, 则 AB 的方程为 $y=k(x+1)+2=kx+k+2$, 由切线性质有, $\frac{|k+2|}{\sqrt{1+k^2}}=1$, 解得 $k=-\frac{3}{4}$, 故 $\frac{y-2}{x+1}$ 的取值范围为 $(-\infty,-\frac{3}{4}]$, 故选: D.

例题 13.20 已知 $P(m,n)$ 为圆 $C:(x-1)^2+(y-1)^2=1$ 上任意一点, 则 $\frac{m+n}{m+1}$ 的最大值为(　)

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

A. $\frac{\sqrt{3}}{3}$　　C. $1+\frac{\sqrt{3}}{3}$

B. $-\frac{\sqrt{3}}{3}$　　D. $1-\frac{\sqrt{3}}{3}$

解 答案:C. $\frac{m+n}{m+1}=\frac{m+1+n-1}{m+1}=1+\frac{n-1}{m+1}$,由于 $P(m,n)$ 为圆 $C:(x-1)^2+(y-1)^2=1$ 上任意一点, 故 $\frac{n-1}{m+1}$ 可看作圆上任意一点 $P(m,n)$ 到定点 $A(-1,1)$ 的斜率, 当直线 PA 与圆相切时, 此时斜率最大, 由于相切时, $|AC|=2,|CP|=1$ 故 $|PA|=\sqrt{3}$, 此时斜率 $k=\frac{|PC|}{|AP|}=\frac{1}{\sqrt{3}}$, 故 $\frac{m+n}{m+1}$ 的最大值为 $1+\frac{\sqrt{3}}{3}$, 故选:C

@余命数 知乎@余命数

例题 13.21 点 $P(x,y)$ 在圆 $(x-2)^2+(y+3)^2=1$ 上, 则 $x+y$ 的范围是__.

解 答案:$[-\sqrt{2}-1,\sqrt{2}-1]$. 设 $x=2+\cos\theta,y=-3+\cos\theta$,即 $P(2+\cos\theta,-3+\sin\theta)$, 所以 $x+y=\sin\theta+\cos\theta-1=\sqrt{2}\sin(\theta+\frac{\pi}{4})-1$, 因为 $-1\le\sin(\theta+\frac{\pi}{4})\le1$, 所以 $-\sqrt{2}-1\le x+y\le\sqrt{2}-1$. 故答案为:$[-\sqrt{2}-1,\sqrt{2}-1]$

例题 13.22 已知 x,y 满足 $x^2+y^2+2x-4y=0$, 则 $2x+y$ 的范围是__.

解 答案:$[-5,5]$. 因为 $x^2+y^2+2x-4y=0$, 所以 $(x+1)^2+(y-2)^2=5$, 表示以 $(-1,2)$ 为圆心, $\sqrt{5}$ 为半径的圆, 即点 (x,y) 为圆 $(x+1)^2+(y-2)^2=5$ 上的点. 令 $2x+y=z$, 即 $2x+y-z=0$, 当直线与圆 $(x+1)^2+(y-2)^2=5$ 相切时 z 取得最值, 所以 $d=\frac{|2\times(-1)+2-z|}{\sqrt{2^2+1^2}}=\sqrt{5}$, 即 $|z|=5$, 解得 $z=\pm5$, 所以 $-5\le2x+y\le5$. 故答案为:$[-5,5]$

例题 13.23 如果实数 x,y 满足等式 $x^2+y^2+4x-2y-4=0$, 那么 x^2+y^2 的最大值是__;$2x-y$ 的最大值是__.

解 答案:$14+6\sqrt{5};3\sqrt{5}-5$. 由 $x^2+y^2+4x-2y-4=0$, 得 $(x+2)^2+(y-1)^2=9$, x^2+y^2 的几何意义为圆 $(x+2)^2+(y-1)^2=9$ 上的动点到原点距离的平方. 因为圆心 $(-2,1)$ 到原点的距离为 $\sqrt{5}$, 所以圆上的动点到原点距离的最大值为 $\sqrt{5}+3$, 则 x^2+y^2 的最大值是 $(\sqrt{5}+3)^2=14+6\sqrt{5}$. 令 $2x-y=t$, 则 $-t$ 是直线 $2x-y=t$ 在 y 轴上的截距, 当直线与圆相切时, 直线 $2x-y=t$ 在 y 轴上的截距, 一个是最大值, 一个是最小值, 此时, 圆心 $(-2,1)$ 到直线 $2x-y=t$ 的距离 $d=\frac{|-4-1-t|}{\sqrt{5}}=3$, 解得 $t=-5\pm3\sqrt{5}$, 所以 $2x-y$ 的最大值为 $3\sqrt{5}-5$. 故答案为:$14+6\sqrt{5};3\sqrt{5}-5$.

例题 13.24 直线 $l:kx-y-2k+2=0(k\in R)$ 过定点 Q, 若 P 为圆 $C:(x-2)^2+(y-3)^2=4$ 上任意一点, 则 $|PQ|$ 的最大值为(　)

A. 1　　C. 4

B. 3　　D. 2

解 答案: B. 由 $kx-y-2k+2=0(k\in R)$, 得 $y-2=k(x-2)$, 所以直线过定点 $Q(2,2)$, 由 $C:(x-2)^2+(y-3)^2=4$, 知圆心坐标 $(2,3)$, 半径为 2, 所以 Q 到圆心的距离为 $d=\sqrt{(2-2)^2+(2-3)^2}=1<2$, 则 Q 在圆内, 则 $|PQ|$ 的最大值为 $d+2=3$, 故选: B

知乎@余命数

知乎@余命数

13.1.6 本节小结

- **直线方程**: 点斜 / 斜截 / 两点 / 截距 / 一般式 $Ax+By+C=0$; 向量参数式 $\vec{p}=\overrightarrow{p_0}+t\vec{d}$.
- **圆**: 标准式 $(x-h)^2+(y-k)^2=R^2$; 一般式 $x^2+y^2+Dx+Ey+F=0$(配方得圆心半径).
- **位置关系**: 比较弦心距 d 与 R; 切线替换式 $(x_0-h)(x-h)+(y_0-k)(y-k)=R^2$.
- **弦长**: 几何式 $2\sqrt{R^2-d^2}$; 代数式 $\sqrt{1+k^2}\,\sqrt{\Delta}/|a|$.
- **圆幂**: 相交弦 / 切割线; 两圆位置由圆心距与半径比较; 直径所对圆周角为直角.

知乎@余命数

13.1.7 本节习题

13.1.　求点 $(1,-2)$ 到直线 $3x-4y+5=0$ 的距离, 并判断该点在直线哪一侧对距离公式中绝对值的影响.

13.2.　圆 $x^2+y^2=16$, 直线 $y=x+m$. 求使直线与圆相切的 m, 并写出切线方程.

13.3.　圆 $(x-1)^2+(y+2)^2=9$ 被直线 $x-2y+1=0$ 所截, 求弦长与弦中点坐标.

13.4.　自点 $P(6,0)$ 向圆 $x^2+y^2=4$ 作切线, 求切线长; 若割线过 P 交圆于 A,B, 说明 $|PA|\cdot|PB|$ 的定值.

13.5.　两圆 $x^2+y^2=1$ 与 $(x-2)^2+y^2=4$, 判断它们相交、相切还是相离, 并求公共弦所在直线方程(两圆方程相减).

13.6.　AB 为圆 $x^2+y^2=r^2$ 的直径, P 在圆上. 用坐标证明 $\overrightarrow{PA}\cdot\overrightarrow{PB}=0$, 并说明这与直径所对圆周角为直角这一事实的等价性.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.2 线性规划初步 *

在前面的章节中, 我们运用解析几何的工具来描述和分析确定的几何图形. 接着将视野转向一个更具动态和决策性的领域, 即如何在给定的限制条件下做出最优的选择. 这是运筹学要解决的问题, 而线性规划是处理此类问题最基本且最实用的数学工具之一.

从实质上讲, 线性规划是一种数学建模技术, 用于解决资源分配的最优化问题. 几乎所有组织, 无论是商业公司, 政府机构, 还是非营利组织, 都面临着一个共同的挑战, 即在资源, 如资金, 原材料, 人力, 时间有限的情况下, 如何实现某个目标的最大化, 如利润, 效率, 社会福利或最小化, 如成本, 风险, 资源消耗.

一个问题能够被构建为线性规划模型, 必须具备三个数学要素:

- **目标函数**: 一个需要被最优化的量, 并且这个量可以表示为决策变量的线性函数. 线性意味着变量之间的关系是成比例的, 例如, 生产两倍的产品会获得两倍的利润.
- **决策变量**: 一组可控制的变量, 它们的取值共同构成一个决策方案. 例如, 生产线上各种产品的产量, 或投资组合中各项资产的金额.
- **约束条件**: 一系列描述决策变量必须满足的限制条件的线性等式或不等式. 它们反映了现实世界中资源的稀缺性或规则的限制.

因此, 线性规划的宏观定义是**在一组线性约束条件下, 寻找一组决策变量的值, 使得一个线性目标函数达到最优(最大或最小)的数学方法.**

13.2.0.1 经济学中的资源优化问题

线性规划的思想与经济学的基本原则, 即研究稀缺资源如何有效配置, 不谋而合. 因此, 它在经济学领域得到了极其广泛和基本的应用. 一个经典的**生产计划问题**可以具体说明这一点.

设想一个企业, 它利用有限的资源, 如劳动力, 机器工时, 原材料来生产多种产品. 企业的目标是制定一个生产计划, 也就是决定每种产品的产量, 以实现总利润的最大化.

- **决策变量**: 设企业生产 n 种产品, 对应的产量分别为 $x_1, x_2, \ldots, x_n$. 这就是企业需要决定的方案.
- **目标函数**: 每种产品 i 都有一个单位利润 c_i. 企业的总利润 Z 就是所有产品利润的总和, 即 $Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n$. 这是一个典型的线性函数, 企业的目标是 $\max Z$.
- **约束条件**: 生产过程受到多种资源的限制. 假设有 m 种资源.
  - 对于第 j 种资源, 例如, 钢材, 其总供应量是有限的, 设为 b_j.
  - 生产每单位的第 i 种产品, 需要消耗 a_{ji} 个单位的第 j 种资源.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 因此, 对第 j 种资源的总消耗量不能超过其总供应量, 这就构成了一条线性不等式约束:

$$a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n \le b_j$$

  企业有多少种资源限制, 就会有多少条这样的约束.

- 同时, 产量不能为负, 即 $x_i \ge 0$ 对所有 i 成立.

通过这样的建模过程, 一个复杂的企业生产决策问题, 被清晰地转化为一个标准的线性规划数学模型.

除了生产计划, 线性规划在经济学中还有诸多其他应用, 例如:

- **运输问题**: 如何安排从多个产地到多个销地的货物运输方案, 以最低的总运输成本满足所有销地的需求.
- **投资组合选择**: 在预算和风险承受能力的约束下, 如何分配资金到不同的金融资产(股票, 债券等), 以最大化预期回报.
- **营养学问题**, 又称饮食问题: 如何从多种食物中进行搭配, 以最低的成本满足人体每日对各种营养素(维生素, 蛋白质等)的最低需求. 这是最早被研究的线性规划问题之一.

在所有这些场景中, 线性规划都提供了一个系统化的框架, 将现实世界的优化问题转化为一个结构清晰, 并且存在有效求解算法的数学问题. 接下来, 将分析如何利用解析几何的知识, 来求解这类问题.

13.2.1 线性规划问题的数学模型

下面用一个具体问题引入线性规划的数学结构.

例题 13.25　一个工坊生产两种产品, 产品 A 和产品 B.

- 每生产一件产品 A, 需要 2 个单位的原料甲和 1 个单位的原料乙, 可获利 3 万元.
- 每生产一件产品 B, 需要 1 个单位的原料甲和 2 个单位的原料乙, 可获利 4 万元.

工坊每日可用的原料甲总量为 8 个单位, 原料乙总量为 7 个单位. 问每日应如何安排两种产品的生产件数, 才能获得最大利润?

要将此问题数学化, 需确定三个要素:

1. **决策变量**: 这是需决定的未知量. 在本例中, 设每日生产产品 A 为 x 件, 产品 B 为 y 件.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. **约束条件**: 这是决策变量必须满足的限制, 通常由一系列线性等式或不等式表示.
   - 生产件数不能为负: $x \geq 0, y \geq 0$.
   - 原料甲的限制: $2x + y \leq 8$.
   - 原料乙的限制: $x + 2y \leq 7$.
3. **目标函数**: 这是一个关于决策变量的线性函数, 希望将其最大化或最小化. 在本例中, 目标是最大化总利润 z, 即 $z = 3x + 4y$.

故该问题被抽象为一个标准的二维线性规划模型:

$$\text{最大化} \quad z = 3x + 4y$$

$$\text{满足约束} \quad \begin{cases} 2x + y \leq 8 \\ x + 2y \leq 7 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

13.2.2 可行域与图解法

对于二维线性规划问题, 可借助解析几何的工具, 将整个优化过程转化为一个直观的几何搜寻任务. 这种方法被称为**图解法**, 其原理建立在对约束条件和目标函数的几何解释之上.

约束条件组中的每一个线性不等式, 如 $ax + by \leq c$, 都将笛卡尔平面分割为两部分, 即满足条件的点构成的**闭半平面**和不满足条件的点构成的开半平面. 所有约束条件所对应的闭半平面的公共交集, 构成了所有可行解 (x, y) 的集合. 我们称这个几何区域为**可行域**. 在二维平面上, 只要可行域有界, 它必然是一个**凸多边形**. 凸性意味着连接区域内任意两点的线段完全位于该区域之内.

对于我们的生产计划问题, 可行域由以下四个半平面交织而成: $x \geq 0$, 即 y 轴右侧; $y \geq 0$, 即 x 轴上方; $2x + y \leq 8$, 即直线 $2x + y = 8$ 左下方; $x + 2y \leq 7$, 即直线 $x + 2y = 7$ 左下方.

接着, 我们考察目标函数 $z = 3x + 4y$. 这个方程可以改写为斜截式:

$$y = -\frac{3}{4}x + \frac{z}{4}$$

对于任意一个特定的利润值 z_0, 方程 $y = -\frac{3}{4}x + \frac{z_0}{4}$ 都代表一条斜率为 $-\frac{3}{4}$ 的直线. 当 z 的值连续变化时, 得的是一族斜率相同但 y 轴截距 $\frac{z}{4}$ 不同的平行直线. 我们称这族直线为目标函数的**等值线**, 因为在同一条直线上的所有点 (x, y) 都会产生相同的目标函数值 z.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

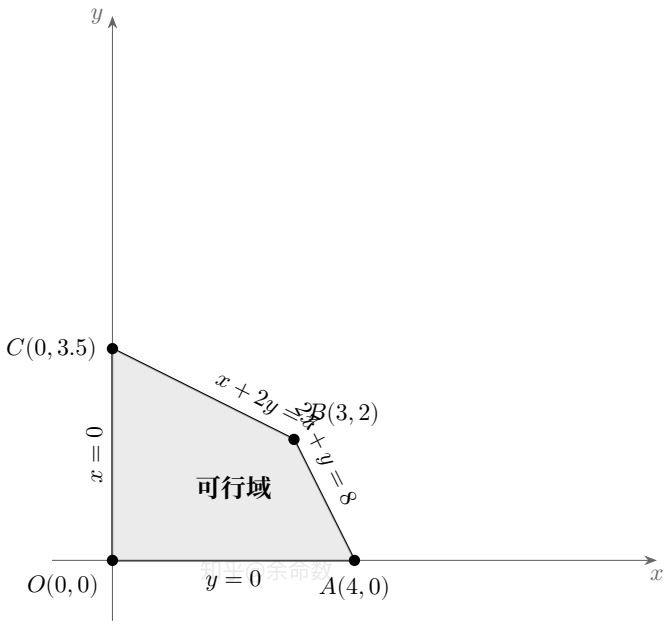


图 13.5　由约束条件构成的凸多边形可行域

我们的任务是最大化 z. 由上式可见, z 与 y 轴截距 $\frac{z}{4}$ 是成正比的. 因此, **最大化目标函数 z 在几何上等价于寻找一条与可行域有公共点, 且具有最大 y 轴截距的等值线**.

接下来确定使 z 增大的平移方向. 回忆直线的一般式方程 $Ax + By - z = 0$, 已知向量 $\vec{n} = (A, B)$ 是该直线的一个法向量. 在本例中, 目标函数 $z = 3x + 4y$ 对应的法向量为 $\vec{n} = (3, 4)$. 几何上, 法向量指向的方向是直线抬升最快的方向, 也就是使得 $3x + 4y$ 的值增大的方向. 因此, 我们的求解过程可以想象为, 将等值线 $3x + 4y = z$ 沿着其法向量 $\vec{n} = (3, 4)$ 的方向平移, 直到它即将离开可行域.

如图所示, 等值线族 $3x + 4y = z$ 是一系列斜率为 $-\frac{3}{4}$ 的平行线. 当我们沿着法向量 $\vec{n} = (3, 4)$ 的方向推动这条直线时, 即向右上方平移, z 的值不断增大. 这个平移过程是有限制的, 因为该直线必须始终与灰色表示的可行域保持接触. 当等值线移动到恰好经过顶点 $B(3, 2)$ 时, 它达到了在满足约束条件下的最远位置. 如果再向上平移丝毫, 它将与可行域完全分离. 因此, 可行域中的顶点 B 就是使目标函数值达到最大的点.

这个几何过程说明了线性规划问题的一个根本性质, 并提供下一节将要介绍的更普适的代数解法的理论基础.

13.2.3 线性规划基本定理

图解法提供了实用的几何直观, 但对于高维度问题或精度要求高的场景, 需一种更普适的代数方法. 该方法的理论基础是下列的定理.

知乎@余命数
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

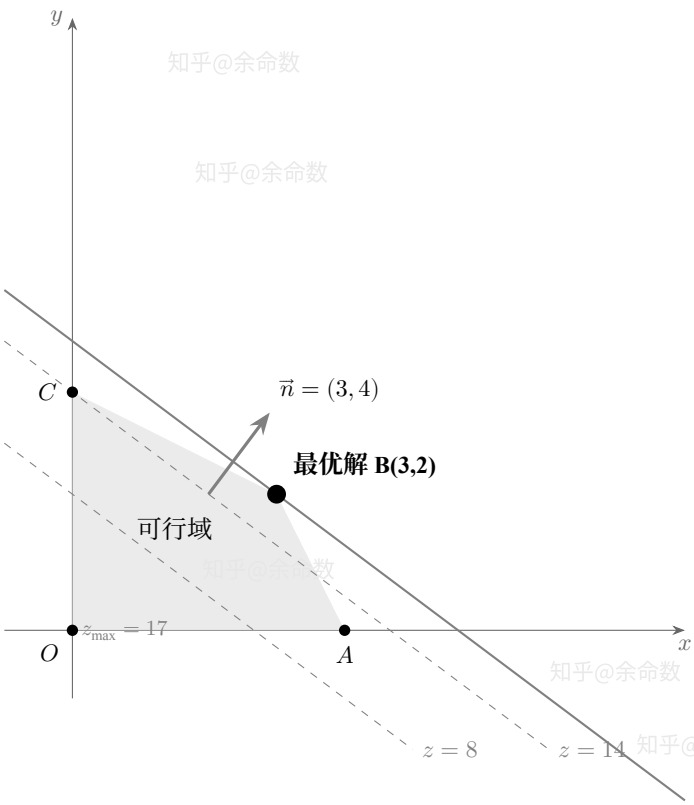


图 13.6 图解法原理, 沿法向量方向平移等值线

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 13.3 (线性规划基本定理) 对于本节图解法所处理的二维线性规划问题, 若其可行域是非空有界凸多边形, 那么目标函数的最优解一定可以在可行域的某个顶点处取得. 若存在多个最优解, 则它们至少覆盖可行域的一条边, 因而该边的端点(顶点)同样是最优解.

证明 考察目标函数 $z = ax + by$ 所代表的平行直线族 l_z. 设可行域为 S. 因为 S 是非空闭且有界的凸多边形, 所以连续函数 z 在 S 上必然存在最大值和最小值. 以最大值为例证明, 最小值情形同理. 设取得最大值的等值线为 l_{z^*}. l_{z^*} 与 S 至少有一个公共点. 若此公共点位于 S 的内部, 则总可将 l_{z^*} 再微小平移一点而不离开 S, 从而得到一个更大的 z 值, 这与 z^* 是最大值矛盾. 因此, 最优解必须出现在 S 的边界上. 若 l_{z^*} 与 S 的交集是边界上的一点, 则该点必为顶点. 若 l_{z^*} 与 S 的交集是边界上的一条边, 那么这条边上的所有点都是最优解, 这自然也包括了该边的两个顶点. 故必有顶点是最优解. □

基本定理将一个可能包含无穷多个解的搜索问题, 简化为对有限个顶点的考察. 这引出了**顶点检验法**:

1. **找出可行域的所有顶点**. 每个顶点都是两条或多条边界直线方程的交点, 可通过解方程组得到.
2. **计算目标函数在每个顶点的值**.
3. **比较这些值**, 其中最大者即为问题的最大值, 最小者即为最小值.

对于引例, 我们已通过图解法确定了四个顶点:

$$O(0,0), \quad A(4,0), \quad C(0,3.5), \quad B(3,2).$$

代入目标函数 $z = 3x + 4y$ 进行检验:

- $z(O) = 3(0) + 4(0) = 0$
- $z(A) = 3(4) + 4(0) = 12$
- $z(C) = 3(0) + 4(3.5) = 14$
- $z(B) = 3(3) + 4(2) = 17$

比较可知, 最大值为 17, 在顶点 $B(3,2)$ 处取得. 这与图解法的结果一致.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.26 求目标函数 $z = 5x + 3y$ 在下列约束条件下

$$\begin{cases} 2x + y \leq 8 \\ x + y \leq 6 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

的最大值.

解 先需确定由该线性不等式组定义的可行域. 该区域是一个位于第一象限的闭合凸多边形. 接着, 求解边界直线的交点, 找出其所有顶点. 可行域的顶点由边界

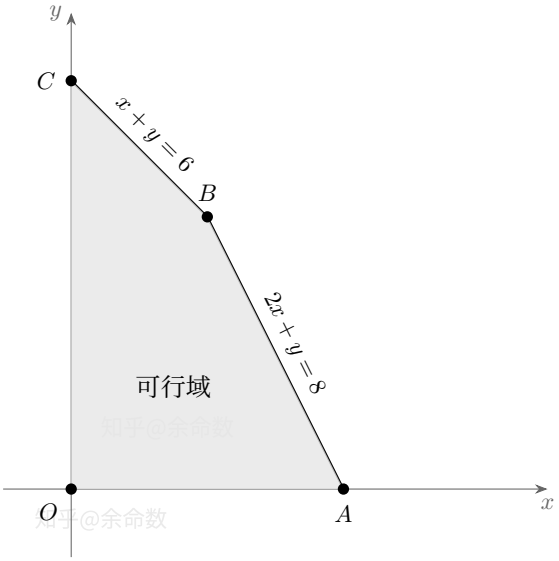


图 13.7 例 1 的可行域

线的交点确定:

- 顶点 O:$x = 0$ 与 $y = 0$ 的交点, $O(0, 0)$.
- 顶点 A:$2x + y = 8$ 与 $y = 0$ 的交点, 解得 $A(4, 0)$.
- 顶点 C:$x + y = 6$ 与 $x = 0$ 的交点, 解得 $C(0, 6)$.
- 顶点 B:$2x + y = 8$ 与 $x + y = 6$ 的交点. 联立方程组

$$\begin{cases} 2x + y = 8 \\ x + y = 6 \end{cases} \Longrightarrow \begin{cases} x = 2 \\ y = 4 \end{cases}$$

故交点为 $B(2, 4)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据线性规划基本定理, 最大值必在这些顶点之一处取得. 将各顶点坐标代入目标函数 $z=5x+3y$ 进行检验: 比较上表中的目标函数值, 可见最大值为 22, 它在

顶点坐标 (x,y)	目标函数值 $z=5x+3y$
$O(0,0)$	$5(0)+3(0)=0$
$A(4,0)$	$5(4)+3(0)=20$
$B(2,4)$	$5(2)+3(4)=10+12=22$
$C(0,6)$	$5(0)+3(6)=18$

顶点 $B(2,4)$ 处取得.

例题 13.27　求目标函数 $z=2x+2y$ 在约束条件

$$\begin{cases} x+y\le 7 \\ x\ge 1 \\ y\ge 2 \end{cases}$$

下的最大值.

解　同样地, 先确定可行域的顶点. 该可行域是一个三角形. 顶点由以下边界线的

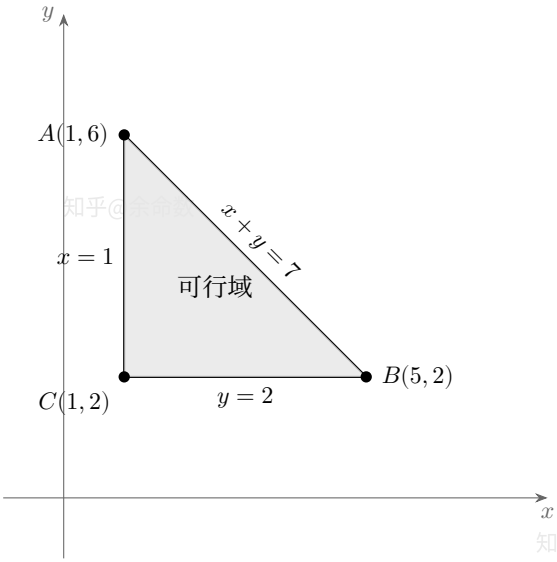


图 13.8　例 2 的可行域

交点构成:

- 顶点 C: $x=1$ 与 $y=2$ 的交点, $C(1,2)$.
- 顶点 A: $x=1$ 与 $x+y=7$ 的交点, 解得 $A(1,6)$.
- 顶点 B: $y=2$ 与 $x+y=7$ 的交点, 解得 $B(5,2)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将这些顶点代入目标函数 $z=2x+2y$ 中：观察结果，可见最大值 14 同时在两个顶

顶点坐标 (x,y)	目标函数值 $z=2x+2y$
$A(1,6)$	$2(1)+2(6)=14$
$B(5,2)$	$2(5)+2(2)=14$
$C(1,2)$	$2(1)+2(2)=6$

知乎@余命数

点 $A(1,6)$ 和 $B(5,2)$ 处取得.

这种情况源于目标函数等值线与可行域边界的平行关系. 注意到目标函数 $z=2x+2y$ 可以改写为 $y=-x+z/2$. 其等值线的斜率为 -1. 可行域的一条边界线 $x+y=7$ 的斜率同样为 -1. 即目标函数的等值线与可行域的边界 AB 平行. 当等值线平移至与线段 AB 重合时，整个线段上的所有点都将成为最优解. 因此，该线性规划问题的最大值为 14，且最优解为线段 AB 上的任意一点.

在某些问题中，决策变量必须取整数，这类问题称为**整数线性规划**. 虽然其通用解法远比标准线性规划复杂，但在二维平面上，若可行域有界，有时可以结合几何方法求解. 其中一类特殊问题是计算可行域内的整点（格点）总数.

例题 13.28 在笛卡尔坐标平面上，求满足不等式组

$$\begin{cases} y \le 3x \\ y \ge \frac{1}{3}x \\ x+y \le 100 \end{cases}$$

的整点 (x,y) 的个数.

在解决此问题之前，先介绍一个引理.

引理 13.1 (皮克定理) 对于一个顶点均为整点的简单多边形，其面积 S 与其内部的整点数 I、边界上的整点数 B 之间满足如下关系：

$$S=I+\frac{B}{2}-1$$

证明 证明 $\Phi(P)=I_P+\frac{B_P}{2}-1$ 与面积 $S(P)$ 具有相同的**可加性**；一旦在基本图形上验证 $S=\Phi$，即可推广到任意简单多边形.

取两个顶点为整点的简单多边形 P_1,P_2，共享公共边 e，且并 $P=P_1\cup P_2$ 仍为简单多边形. 设 e 上共有 k 个整点，含两端点. 面积显然可加：$S(P)=S(P_1)+S(P_2)$. 内部与边界整点满足

- 内部：$I_P=I_{P_1}+I_{P_2}+(k-2)$（$e$ 上不含端点的 $k-2$ 个点由边界转入内部）.

- 边界：$B_P = B_{P_1} + B_{P_2} - 2k + 2$（$e$ 上的点在 P_1, P_2 中各计一次，合并后中间 $k-2$ 点不再属于边界，两端点仍属边界）.

于是

$$\begin{aligned}\Phi(P) &= I_P + \frac{B_P}{2} - 1 \\ &= (I_{P_1} + I_{P_2} + k - 2) + \frac{B_{P_1} + B_{P_2} - 2k + 2}{2} - 1 \\ &= \Phi(P_1) + \Phi(P_2).\end{aligned}$$

故 Φ 与 S 同为可加. 只需在基本图形上验证 $S = \Phi$. 取边平行于坐标轴的整点矩形 R，边长为整数 a, b：

$$S_R = ab, \quad B_R = 2a + 2b, \quad I_R = (a-1)(b-1) = ab - a - b + 1.$$

$$\Phi(R) = I_R + \frac{B_R}{2} - 1 = ab = S_R.$$

定理对一切轴对齐整点矩形成立.

进一步，任一两条直角边分别平行于坐标轴的整点直角三角形，都可以看作某个上述整点矩形沿对角线分成的两个全等整点三角形之一. 这两个三角形的面积相等，且由平移或中心对称可知其内部整点数与边界整点数也相同，因而它们的 Φ 值相等. 由矩形的公式和可加性可知，每个这样的整点直角三角形也满足 $S = \Phi$.

任何一个顶点为整点的简单多边形，都可以被三**角剖分**，即分解为有限个顶点同为整点的三角形的并. 同时，任何一个顶点为整点的三角形，都可以通过一个外接的、边与坐标轴平行的矩形，减去其角上至多三个顶点为整点的直角三角形得到. 由于已证明 Φ 的可加性，并且在矩形和上述直角三角形上验证了 $S = \Phi$，那么由它们通过加减运算构成的任意三角形也必然满足. 最后，既然任意三角形都满足皮克定理，那么由这些三角形拼接而成的任意简单多边形也必然满足此定理.

□

解 此问题要求对一个由线性不等式定义的凸多边形区域内的整点进行计数. 直接枚举的方法虽然可行，但过程繁琐. 一个更简洁的途径是运用几何与数论的联系.

先需精确地刻画该不等式组所定义的可行域. 该区域是由三条直线 $l_1 : y = 3x$，$l_2 : y = \frac{1}{3}x$ 以及 $l_3 : x + y = 100$ 所围成的封闭区域. 通过联立方程组，可确定这个区域的顶点.

- 直线 $y = 3x$ 与 $y = \frac{1}{3}x$ 的交点为 $O(0, 0)$.
- 直线 $y = 3x$ 与 $x + y = 100$ 的交点为 $A(25, 75)$.
- 直线 $y = \frac{1}{3}x$ 与 $x + y = 100$ 的交点为 $B(75, 25)$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

该可行域是一个以 O, A, B 为顶点的三角形. 要点在于, **所有顶点均为整点**, 这一特性为应用**皮克定理**提供了理想的条件.

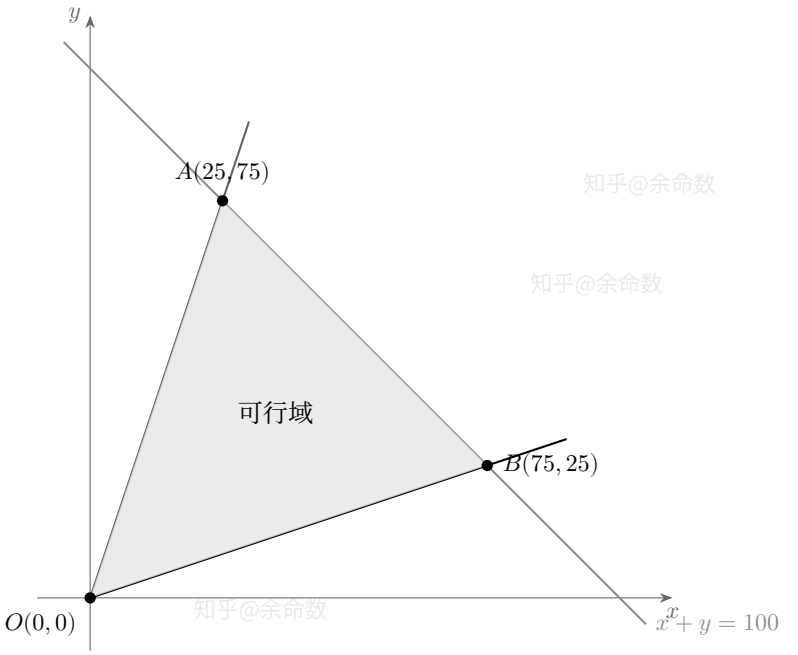


图 13.9 由不等式组定义的可行域 $\triangle OAB$

皮克定理指出, 对于一个顶点均为整点的简单多边形, 其面积 S 与其内部的整点数 I、边界上的整点数 B 之间满足如下关系:

$$S = I + \frac{B}{2} - 1$$

现求总整点数 $N = I + B$. 借助皮克定理, 可将 I 表示为 $I = S - \frac{B}{2} + 1$, 从而得到总整点数的计算公式

$$N = \left(S - \frac{B}{2} + 1\right) + B = S + \frac{B}{2} + 1$$

接下来的任务是分别计算该三角形的面积 S 与其边界上的整点数 B.

对于顶点为 $(x_O, y_O), (x_A, y_A), (x_B, y_B)$ 的三角形, 其面积可由鞋带公式给出:

$$\begin{aligned} S &= \frac{1}{2}|(x_O y_A + x_A y_B + x_B y_O) - (y_O x_A + y_A x_B + y_B x_O)| \\ &= \frac{1}{2}|(0 \cdot 75 + 25 \cdot 25 + 75 \cdot 0) - (0 \cdot 25 + 75 \cdot 75 + 25 \cdot 0)| \\ &= \frac{1}{2}|625 - 5625| = 2500 \end{aligned}$$

知乎@余命数 知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

面积算毕后, 转向边界整点数 B 的计算. 连接两个整点 (x_1, y_1) 和 (x_2, y_2) 的线段上, 其包含的整点个数, 含端点为 $\gcd(|x_2 - x_1|, |y_2 - y_1|) + 1$, 其中 gcd 表示最大公约数.

- **边 OA**(从 $O(0,0)$ 到 $A(25,75)$):整点数为 $\gcd(25,75)+1 = 25+1 = 26$.
- **边 OB**(从 $O(0,0)$ 到 $B(75,25)$):整点数为 $\gcd(75,25)+1 = 25+1 = 26$.
- **边 AB** (从 $A(25,75)$ 到 $B(75,25)$): 整点数为 $\gcd(|75-25|, |25-75|)+1 = \gcd(50,50)+1 = 50+1 = 51$.

知乎@余命数

知乎@余命数

边界上的总整点数 B 是三条边上的整点数之和, 并减去被重复计算的三个顶点, 得

$$B = 26 + 26 + 51 - 3 = 100$$

现在, 已具备所有要素, 可以得出最终的整点总数. 将 $S = 2500$ 和 $B = 100$ 代入公式:

$$N = S + \frac{B}{2} + 1 = 2500 + \frac{100}{2} + 1 = 2500 + 50 + 1 = 2551$$

知乎@余命数

因此, 满足给定不等式组的整点共有 2551 个.

解 [补集计数法] 亦可补集计数, 先数一个包含目标区域且易计的大区域中的整点, 再减去不满足原约束的点.

取 $T: x \geq 0, y \geq 0, x + y \leq 100$. T 包含全部目标整点. 固定 $x = k \in \{0, \ldots, 100\}$ 时 $0 \leq y \leq 100 - k$, 有 $101 - k$ 个整数 y, 故

知乎@余命数

知乎@余命数

知乎@余命数

$$N(T) = \sum_{k=0}^{100} (101 - k) = \frac{101 \times 102}{2} = 5151.$$

令 T_A 为 T 中 $y < \frac{1}{3}x$ 的整点集, T_B 为 $y > 3x$ 的整点集, 则所求 $N = N(T) - N(T_A) - N(T_B)$. 变换 $\phi: (x, y) \mapsto (y, x)$(关于 $y = x$ 对称)将 T 映到自身, 并将 T_A 一一映到 T_B, 故 $N(T_A) = N(T_B)$, 从而 $N = N(T) - 2N(T_B)$.

T_B 由 $x \geq 0, y \geq 3x + 1, x + y \leq 100$ 刻画. 固定 $x = k$ 时需 $3k + 1 \leq 100 - k$, 即 $k \leq 24$, 且 y 有 $100 - 4k$ 个取值, 故

$$N(T_B) = \sum_{k=0}^{24} (100 - 4k) = \frac{(100 + 4) \times 25}{2} = 1300.$$

于是 $N = 5151 - 2 \times 1300 = 2551$.

例题 13.29 点 $P(1,1), Q(2,2), R(0,-1)$ 在由方程 $|x-1| + |y-1| = 1$ 确定的曲线所围成的图形中的个数有

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

A. 3　　B. 2　　C. 1　　D. 0

解 此问题需要先识别方程 $|x-1|+|y-1|=1$ 定义的几何图形, 再判断给定点与该图形的位置关系. 该方程描述的是一个区域的边界, 而区域内部及边界上的点 (x,y) 则满足不等式

$$|x-1|+|y-1|\leq 1$$

为了清晰地说明其几何形态, 可引入一个坐标平移变换. 令 $x'=x-1$ 及 $y'=y-1$. 在新的 $x'y'$ 坐标系中, 不等式转化为

$$|x'|+|y'|\leq 1$$

这是一个标准形式, 它表示一个以新坐标系原点为中心, 顶点分别位于坐标轴上 $(\pm 1,0)$ 和 $(0,\pm 1)$ 的正方形区域. 将这个几何图形变换回原始的 xy 坐标系, 可知该区域是一个以点 $(1,1)$ 为对称中心, 顶点为 $(2,1),(1,2),(0,1),(1,0)$ 的旋转正方形.

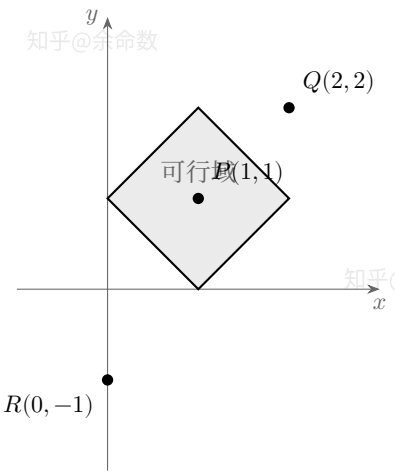


图 13.10　由 $|x-1|+|y-1|\leq 1$ 定义的区域及相关点

接着, 通过代数方法精确检验每个点的位置:

- 对于点 $P(1,1)$:

$$|1-1|+|1-1|=0+0=0$$

由于 $0\leq 1$, 点 P 位于该区域内, 且为其对称中心.

- 对于点 $Q(2,2)$:

$$|2-1|+|2-1|=1+1=2$$

由于 $2>1$, 点 Q 位于该区域之外.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 对于点 $R(0,-1)$:

$$|0-1|+|-1-1|=|-1|+|-2|=1+2=3$$

由于 $3>1$, 点 R 同样位于该区域之外.

仅有点 P 满足条件, 因此在图形中的点的个数为 1. 正确的选项为 C.

例题 13.30　坐标平面内区域 $M\left\{(x,y)\middle|\begin{cases}x-y+1\geq 0\\ x+y-1\leq 0\\ kx-y-1\leq 0\end{cases},(0\leq k\leq 1)\right\}$ 的面积可用函数 $f(k)$ 表示, 若 $f(k)=8$, 则 $k=$ 是多少.

解　区域 M 由含参数 k 的线性不等式组确定, 已知面积 $f(k)=8$, 求 k. 先把不等式写成边界直线:

$$\begin{cases}y\leq x+1 & (l_1)\\ y\leq -x+1 & (l_2)\\ y\geq kx-1 & (l_3)\end{cases}$$

当 $0\leq k<1$ 时, 可行域为三直线围成的三角形; 当 $k=1$ 时区域无界, 不可能有有限面积 8. 故只需讨论 $0\leq k<1$. 顶点由边界两两相交得到:

- l_1 与 l_2: $\begin{cases}y=x+1\\ y=-x+1\end{cases}$ 得 $A=(0,1)$, 与 k 无关.
- l_1 与 l_3: $x+1=kx-1$ 得 $x=\frac{2}{k-1}$, $y=\frac{k+1}{k-1}$, 故 $B=\left(\frac{2}{k-1},\frac{k+1}{k-1}\right)$.
- l_2 与 l_3: $-x+1=kx-1$ 得 $x=\frac{2}{k+1}$, $y=\frac{k-1}{k+1}$, 故 $C=\left(\frac{2}{k+1},\frac{k-1}{k+1}\right)$.

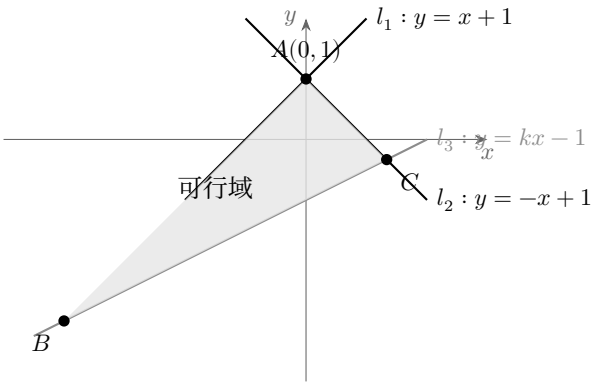


图 13.11　可行域示意图(以 $k=0.5$ 为例)

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

以 BC 为底(落在 $l_3: kx-y-1=0$ 上)算面积. 底长平方为

$$\begin{aligned}|BC|^2 &= \left(\frac{2}{k-1}-\frac{2}{k+1}\right)^2+\left(\frac{k+1}{k-1}-\frac{k-1}{k+1}\right)^2\\ &= \left(\frac{4}{k^2-1}\right)^2+\left(\frac{4k}{k^2-1}\right)^2=\frac{16(1+k^2)}{(k^2-1)^2}.\end{aligned}$$

因 $0\le k<1, k^2-1<0$, 故 $|BC|=\frac{4\sqrt{1+k^2}}{1-k^2}$. 顶点 $A(0,1)$ 到 l_3 的距离为高:

$$h=\frac{|k\cdot 0-1-1|}{\sqrt{k^2+1}}=\frac{2}{\sqrt{k^2+1}}.$$

于是

$$f(k)=\frac{1}{2}|BC|\cdot h=\frac{4}{1-k^2}.$$

令 $f(k)=8$: $\frac{4}{1-k^2}=8$, 即 $2k^2=1, k=\pm\frac{\sqrt{2}}{2}$. 由 $0\le k\le 1$ 舍去负根, 得 $k=\frac{\sqrt{2}}{2}$.

例题 13.31 对于每个自然数 n, 抛物线 $y=(n^2+n)x^2-(2n+1)x+1$ 与 x 轴交于 A_n, B_n 两点, 以 $|A_nB_n|$ 表示该两点的距离, 则 $|A_1B_1|+|A_2B_2|+\cdots+|A_{2000}B_{2000}|$ 的值是

A. $\frac{1999}{2000}$　　C. $\frac{1999}{2001}$

B. $\frac{2000}{2001}$　　D. $\frac{2001}{2001}$

解 此问题考察了二次函数与 x 轴交点距离的计算, 并将其与级数求和, 特别是裂项相消法, 巧妙地结合在一起.

对于给定的自然数 n, 抛物线 $y=(n^2+n)x^2-(2n+1)x+1$ 与 x 轴的交点 A_n, B_n 的横坐标是二次方程

$$(n^2+n)x^2-(2n+1)x+1=0$$

的两个实数根. 设这两个根为 x_1 和 x_2. 根据韦达定理, 有

$$x_1+x_2=\frac{2n+1}{n^2+n},\quad x_1x_2=\frac{1}{n^2+n}$$

两交点之间的距离 $|A_nB_n|$ 即为两根之差的绝对值 $|x_1-x_2|$. 可利用根与系数的关系来计算这个距离, 得

$$|A_nB_n|=|x_1-x_2|=\sqrt{(x_1-x_2)^2}=\sqrt{(x_1+x_2)^2-4x_1x_2}$$

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将韦达定理的结果代入上式, 得

$$\begin{aligned}|A_nB_n| &= \sqrt{\left(\frac{2n+1}{n^2+n}\right)^2 - 4\left(\frac{1}{n^2+n}\right)} \\ &= \sqrt{\frac{(2n+1)^2-4(n^2+n)}{(n^2+n)^2}} \\ &= \sqrt{\frac{4n^2+4n+1-4n^2-4n}{(n^2+n)^2}} \\ &= \sqrt{\frac{1}{(n^2+n)^2}} = \frac{1}{n^2+n}\end{aligned}$$

由于 n 是自然数, $n \geq 1$, 故 $n^2+n>0$, 分母开方后无需加绝对值.

注意到 $n^2+n=n(n+1)$, 于是得距离的简洁表达式:

$$|A_nB_n| = \frac{1}{n(n+1)}$$

这是一个可以进行裂项分解的结构. 将其分解为部分分式:

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$$

现在的任务是计算这个级数的和. 这是一个典型的裂项相消.

$$\begin{aligned}S &= \sum_{n=1}^{2000} |A_nB_n| = \sum_{n=1}^{2000} \frac{1}{n(n+1)} \\ &= \sum_{n=1}^{2000} \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{2000} - \frac{1}{2001}\right)\end{aligned}$$

观察到, 除首项和末项外, 所有中间项都相互抵消. 因此, 和为

$$S = \frac{1}{1} - \frac{1}{2001} = \frac{2001-1}{2001} = \frac{2000}{2001}$$

故正确选项为 B.

例题 13.32　若当点 $P(m,n)$ 为圆 $x^2+(y-1)^2=1$ 上任一点时, 不等式 $m+n+c \geq 0$ 恒成立, 则 c 的取值范围是?

解　需寻找参数 c 的范围, 使得一个线性不等式在一个圆域上恒成立. 这是一个典型的线性规划思想在解析几何中的应用.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

题目所给的条件是,对于满足方程 $m^2+(n-1)^2=1$ 的任意实数对 (m,n),不等式 $m+n+c\geq 0$ 均成立. 可将此不等式变形为 $c\geq -(m+n)$. 此不等式恒成立,意味着 c 的值必须大于或等于表达式 $-(m+n)$ 在圆上的所有可能取值.c 必须大于或等于 $-(m+n)$ 的最大值.

$$c\geq \max_{m^2+(n-1)^2=1}\{-(m+n)\}$$

求 $-(m+n)$ 的最大值,等价于求 $z=m+n$ 的最小值.

将此问题置于几何框架中考察. 目标函数 $z=m+n$ 在 mn-坐标平面上代表了一族斜率为 -1 的平行直线. 需寻找当这条直线与圆 $m^2+(n-1)^2=1$ 有公共点时,z 所能取到的最小值. 该圆的圆心为 $C(0,1)$,半径为 $R=1$. 当直线 $m+n-z=0$ 与圆相切时,目标函数 z 会取得其最大值或最小值. 相切的条件是圆心到直线的距离等于半径.

$$d=\frac{|1\cdot 0+1\cdot 1-z|}{\sqrt{1^2+1^2}}=R=1$$

$$\frac{|1-z|}{\sqrt{2}}=1$$

解此方程,得 $|1-z|=\sqrt{2}$,从而 $1-z=\sqrt{2}$ 或 $1-z=-\sqrt{2}$. 这给出了 z 的两个临界值:$z=1-\sqrt{2}$ 和 $z=1+\sqrt{2}$. 显然,$z=m+n$ 在圆上的最小值为 $z_{\min}=1-\sqrt{2}$.

已确定 $m+n$ 的最小值为 $1-\sqrt{2}$. 因此,$-(m+n)$ 的最大值为 $-(1-\sqrt{2})=\sqrt{2}-1$. 要使不等式 $c\geq -(m+n)$ 恒成立,c 必须大于或等于 $-(m+n)$ 的最大值. 所以,$c\geq\sqrt{2}-1$. c 的取值范围是 $[\sqrt{2}-1,+\infty)$.

例题 13.33 已知坐标平面内有曲线 C 的方程为 $x^2\cos^2\theta+y^2\sin^2\theta=4$.

(1) 当 θ 在什么范围内时,由曲线 C 上的所有点都位于不等式 $x^2+y^2\leq 16$ 所表示的图形之中(这里 $0\leq\theta\leq\pi$),并求出此范围;

(2) 当 $\dfrac{\pi}{6}<\theta<\dfrac{\pi}{3}$ 时,求满足曲线 C 的方程的点 (x,y) 的取值范围,并用图形表示.

解 (1) 先分析曲线 C 的方程. 该方程可改写为

$$\frac{x^2}{4/\cos^2\theta}+\frac{y^2}{4/\sin^2\theta}=1$$

当 $\cos\theta\neq 0$ 且 $\sin\theta\neq 0$ 时,此方程表示一个中心在原点的椭圆,其在 x 轴和 y 轴上的半轴长分别为 $a=\frac{2}{|\cos\theta|}$ 和 $b=\frac{2}{|\sin\theta|}$.

不等式 $x^2+y^2\leq 16$ 表示一个以原点为中心、半径为 4 的闭圆盘. 要使椭圆上的所有点都位于该圆盘内,椭圆的半长轴必须不大于圆的半径. 椭圆上任意点到

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

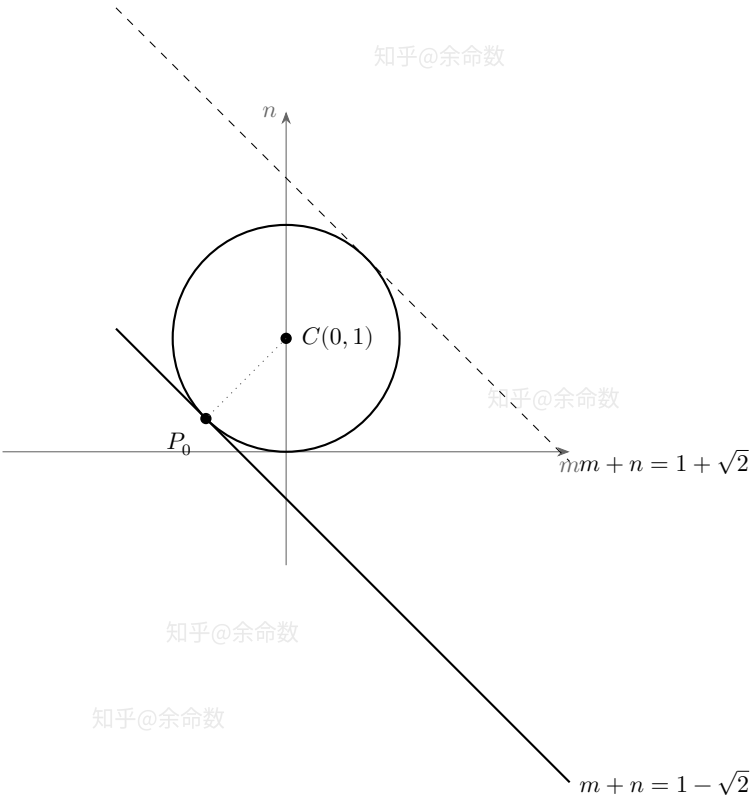


图 13.12　目标函数 $z = m + n$ 的几何解释

知乎@余命数
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

原点的距离的最大值不能超过 4. 这个最大距离就是椭圆的半长轴长度. 因此, 须满足条件:

$$\max\{a,b\}=\max\left\{\frac{2}{|\cos\theta|},\frac{2}{|\sin\theta|}\right\}\leq 4$$

这等价于同时满足以下两个不等式:

$$\frac{2}{|\cos\theta|}\leq 4 \quad 和 \quad \frac{2}{|\sin\theta|}\leq 4$$

化简得到

$$|\cos\theta|\geq\frac{1}{2} \quad 和 \quad |\sin\theta|\geq\frac{1}{2}$$

现在在给定区间 $0\leq\theta\leq\pi$ 内求解这两个不等式.

- 对于 $|\sin\theta|\geq\frac{1}{2}$: 在 $[0,\pi]$ 上, $\sin\theta\geq 0$, 故不等式为 $\sin\theta\geq\frac{1}{2}$. 解得 $\theta\in\left[\frac{\pi}{6},\frac{5\pi}{6}\right]$.
- 对于 $|\cos\theta|\geq\frac{1}{2}$: 即 $\cos\theta\geq\frac{1}{2}$ 或 $\cos\theta\leq-\frac{1}{2}$.
  - 在 $[0,\pi]$ 上, $\cos\theta\geq\frac{1}{2}$ 的解为 $\theta\in\left[0,\frac{\pi}{3}\right]$.
  - 在 $[0,\pi]$ 上, $\cos\theta\leq-\frac{1}{2}$ 的解为 $\theta\in\left[\frac{2\pi}{3},\pi\right]$.

  因此, $|\cos\theta|\geq\frac{1}{2}$ 的解集为 $\left[0,\frac{\pi}{3}\right]\cup\left[\frac{2\pi}{3},\pi\right]$.

为了使两个条件同时成立, 需取上述两个解集的交集:

$$\begin{aligned}&\left([0,\tfrac{\pi}{3}]\cup[\tfrac{2\pi}{3},\pi]\right)\cap[\tfrac{\pi}{6},\tfrac{5\pi}{6}]\\&=[\tfrac{\pi}{6},\tfrac{\pi}{3}]\cup[\tfrac{2\pi}{3},\tfrac{5\pi}{6}].\end{aligned}$$

所以, θ 的取值范围是

$$\left[\frac{\pi}{6},\frac{\pi}{3}\right]\cup\left[\frac{2\pi}{3},\frac{5\pi}{6}\right].$$

(2) 当 θ 在开区间 $\left(\frac{\pi}{6},\frac{\pi}{3}\right)$ 内变化时, 需确定满足下列方程的点 (x,y) 所构成的区域:

$$x^2\cos^2\theta+y^2\sin^2\theta=4.$$

当 θ 取区间的两个端点时, 得两个边界椭圆:

- 当 $\theta\to\frac{\pi}{6}$, $\cos^2\theta\to\frac{3}{4}$, $\sin^2\theta\to\frac{1}{4}$. 曲线趋近于 $x^2\left(\frac{3}{4}\right)+y^2\left(\frac{1}{4}\right)=4$, 即 $3x^2+y^2=16$.
- 当 $\theta\to\frac{\pi}{3}$, $\cos^2\theta\to\frac{1}{4}$, $\sin^2\theta\to\frac{3}{4}$. 曲线趋近于 $x^2\left(\frac{1}{4}\right)+y^2\left(\frac{3}{4}\right)=4$, 即 $x^2+3y^2=16$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

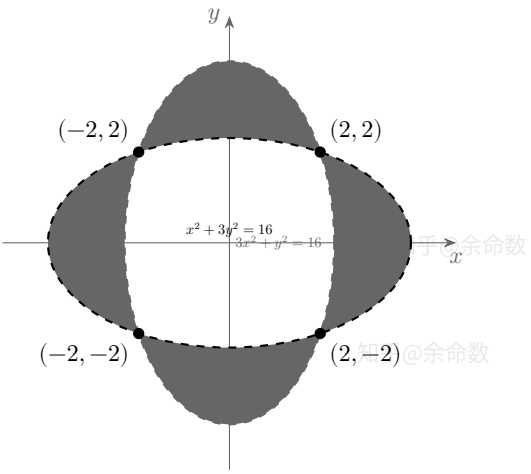


图 13.13　点 (x,y) 的取值范围(阴影部分不含边界, 但包含四个交点)

对于任意一个固定的点 (x,y), 方程 $x^2\cos^2\theta+y^2\sin^2\theta=4$ 可以看作是关于变量 $u=\cos^2\theta$ 的线性方程: $x^2u+y^2(1-u)=4$, 即 $u(x^2-y^2)=4-y^2$. 当 $\theta\in(\frac{\pi}{6},\frac{\pi}{3})$ 时, $u=\cos^2\theta\in(\frac14,\frac34)$. 因此, 点 (x,y) 属于所求区域的充要条件是, 存在 $u\in(\frac14,\frac34)$ 使得 $u(x^2-y^2)=4-y^2$.

当 $x^2\neq y^2$ 时, 这等价于数值 4 严格地位于 $x^2u+y^2(1-u)$ 当 $u=\frac14$ 和 $u=\frac34$ 时的两个值之间. 这两个值分别是 $\frac14x^2+\frac34y^2$ 和 $\frac34x^2+\frac14y^2$. 所以, 这一部分点 (x,y) 所在的区域由以下复合不等式定义:

$$\left(\tfrac14x^2+\tfrac34y^2<4\quad\text{且}\quad\tfrac34x^2+\tfrac14y^2>4\right)$$

或

$$\left(\tfrac14x^2+\tfrac34y^2>4\quad\text{且}\quad\tfrac34x^2+\tfrac14y^2<4\right).$$

整理后即

$$(x^2+3y^2<16\quad\text{且}\quad3x^2+y^2>16)$$

或

$$(x^2+3y^2>16\quad\text{且}\quad3x^2+y^2<16).$$

当 $x^2=y^2$ 时, 原式左端恒为 x^2, 因而还需满足 $x^2=4$, 即得到四个点 $(\pm2,\pm2)$. 这四个点是两条边界椭圆的交点, 且对开区间内任意 θ 都满足原方程.

因此, 该区域是位于椭圆 $3x^2+y^2=16$ 和 $x^2+3y^2=16$ 之间的开区域, 并包含两椭圆的四个交点 $(\pm2,\pm2)$.

图形表示在下一页:

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.2.4 饮食问题与计划经济应用

线性规划的作用，在于把现实世界的优化问题提炼为结构清晰的数学模型. 这一从具体问题到抽象模型的转化过程，被称为**数学建模**，它是应用数学的一项基本技能. 其要点在于准确地识别出决策变量，目标函数和约束条件.

下面用一个经典的**饮食问题**完整地展示这一过程. 这是一个典型的成本最小化问题.

知乎@余命数

例题 13.34 (营养膳食的成本优化) 一位营养师需要为一位运动员设计一份最低成本的早餐食谱，该食谱必须满足特定的营养需求. 早餐由燕麦和牛奶两种食物构成.

- **营养信息与成本**：
  - 每 100 克燕麦：含蛋白质 10 克，维生素 C 5 毫克. 成本为 3 元.
  - 每 100 毫升牛奶：含蛋白质 4 克，维生素 C 20 毫克. 成本为 5 元.
- **营养需求**：早餐必须至少提供 44 克的蛋白质和 60 毫克的维生素 C.

问：应该如何搭配燕麦和牛奶的份量，才能在满足营养需求的前提下，使得总成本最低？

解 设购买燕麦 x 个单位，每单位 100 克，牛奶 y 个单位，每单位 100 毫升，则 $x \geq 0$，$y \geq 0$. 按单价，目标是最小化总成本

$$z = 3x + 5y.$$

营养约束为蛋白质不少于 44 克、维生素 C 不少于 60 毫克：

$$10x + 4y \geq 44, \qquad 5x + 20y \geq 60 \ (\text{即} x + 4y \geq 12).$$

完整模型为

$$\begin{cases} 10x + 4y \geq 44 \\ x + 4y \geq 12 \\ x \geq 0 \\ y \geq 0 \end{cases}$$

用顶点检验法. 可行域向右上方无限延伸，是**无界区域**. 可行域的顶点由边界线的交点构成：

- 顶点 A：$10x + 4y = 44$ 与 $x = 0$ 的交点，解得 $A(0, 11)$.
- 顶点 C：$x + 4y = 12$ 与 $y = 0$ 的交点，解得 $C(12, 0)$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

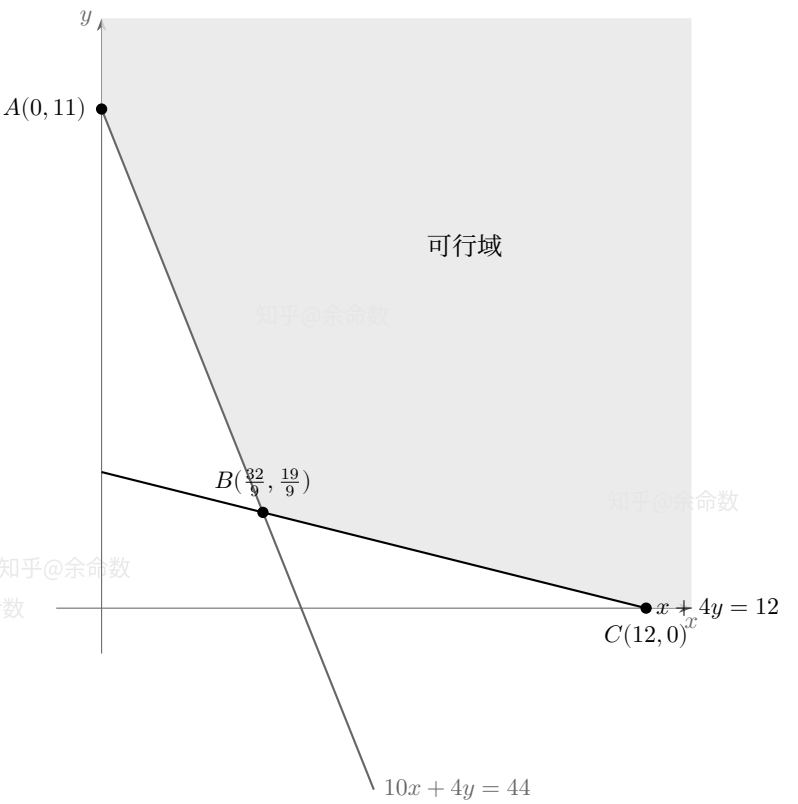


图 13.14　饮食问题的无界可行域

知乎@余命数

顶点	坐标 (x,y)	成本 $z=3x+5y$
A	$(0,11)$	$3(0)+5(11)=55$
B	$(32/9,19/9)$	$3(32/9)+5(19/9)=(96+95)/9=191/9\approx 21.22$
C	$(12,0)$	$3(12)+5(0)=36$

- 顶点 B: $10x+4y=44$ 与 $x+4y=12$ 的交点. 两式相减得 $9x=32 \implies x=32/9$. 代入第二式, $32/9+4y=12 \implies 4y=12-32/9=76/9 \implies y=19/9$. 故 $B(32/9,19/9)$.

即使可行域无界, 线性规划基本定理的一个推论指出, 如果最优解存在, 它也必然在顶点处取得. 可检验目标函数 $z=3x+5y$ 的等值线 $y=-\frac{3}{5}x+\frac{z}{5}$. 当希望最小化 z 时, 相当于在图中寻找截距最小, 即最靠下方的、且与可行域有交点的等值线. 几何直观显示, 这条线必定会首先触碰到可行域的某个角落, 即顶点.

将顶点坐标代入目标函数: 比较可知, 最小成本为 $191/9$ 元, 在顶点 $B(32/9,19/9)$ 处取得.

因此, 最优的膳食搭配方案是购买 $32/9\approx 3.56$ 个单位, 即 356 克的燕麦和 $19/9\approx 2.11$ 个单位, 即 211 毫升的牛奶.

分析计划经济中的应用. 先理解不同经济体系如何影响优化目标. 在市场经济中, 企业通常以**利润最大化**或**成本最小化**为目标. 而在**计划经济**体系中, 经济决策由中央计划机构, 如国家计划委员会做出, 资源依据国家层面的宏观战略目标进行指令性分配.

因此, 计划经济下的优化目标函数可以是一个衡量国家战略优先级的**综合指数**或**计划产值**. 例如, 在优先发展重工业的时期,

知乎@余命数

钢材的价值权重可能远高于消费品; 在专注于农业机械化时, 拖拉机的权重则会最高. 线性规划作为一种数学工具具有这种普适性, 不管目标是利润还是战略价值, 其底层的优化逻辑都是相通的.

例题 13.35 (重工业与农业的平衡发展) 一个地区的计划委员会需要为下一个五年计划周期制定主要产品的生产指标. 该地区主要利用电力、劳动力和铁矿石这三种资源, 生产钢材与拖拉机两种主要产品, 前者是重工业的基础, 后者支援农业.

- **可用资源总量**:
  - 电力: 1800 万千瓦时
  - 熟练劳动力: 2100 万工时
- **生产消耗**(以万吨为单位的钢材和万台为单位的拖拉机):
  - 生产 1 万吨钢材 (x) 需要: 5 万千瓦时电力, 10 万工时劳动力.
  - 生产 1 万台拖拉机 (y) 需要: 20 万千瓦时电力, 15 万工时劳动力.

知乎@余命数

- **内部结构与最低任务约束**:
  - **内部消耗**:每生产 1 万台拖拉机, 需要消耗 2 万吨钢材作为原材料.
  - **最低保障**:为满足基本建设和农业需求, 必须保证钢材产量至少为 60 万吨, 拖拉机产量至少为 30 万台.
- **国家优先指数**: 根据国家以农促工, 工农并举的战略方针, 每单位钢材的优先指数计为 3 分, 每单位拖拉机的优先指数计为 10 分.

问: 应如何规划钢材和拖拉机的产量, 才能使该地区的总国家优先指数最高?

知乎@余命数

解 这是一个典型的、具有多重及内部关联约束的线性规划问题. 将严格按照数学建模的步骤来构建和求解.

设钢材的计划产量为 x 万吨, 拖拉机的计划产量为 y 万台. 目标是最大化国家优先指数 z, 其目标函数为:

$$\text{最大化}\quad z = 3x + 10y$$

决策变量 (x, y) 必须满足一系列约束条件.

- **资源约束**:

$$5x + 20y \le 1800 \quad (\text{电力约束}) \implies x + 4y \le 360$$

$$10x + 15y \le 2100 \quad (\text{劳动力约束}) \implies 2x + 3y \le 420$$

- **内部消耗约束**:用于生产拖拉机的钢材 $(2y)$ 不能超过钢材的总产量 (x).

$$2y \le x \implies x - 2y \ge 0$$

- **最低任务约束**:

$$x \ge 60, \quad y \ge 30$$

获得了完整的线性规划模型. 接着, 通过顶点检验法求解. 首先确定可行域的顶点. 可行域的顶点由边界线的交点构成:

- 顶点 A:$x = 60$ 与 $y = 30$ 的交点. 检验:$60 - 2(30) = 0 \ge 0$. 故 $A(60, 30)$ 是一个顶点.
- 顶点 B:$y = 30$ 与 $2x + 3y = 420$ 的交点.$2x + 3(30) = 420 \implies 2x = 330 \implies x = 165$. 故 $B(165, 30)$.
- 顶点 C:$2x + 3y = 420$ 与 $x + 4y = 360$ 的交点. 联立求解得 $x = 120, y = 60$. 故 $C(120, 60)$.

知乎@余命数

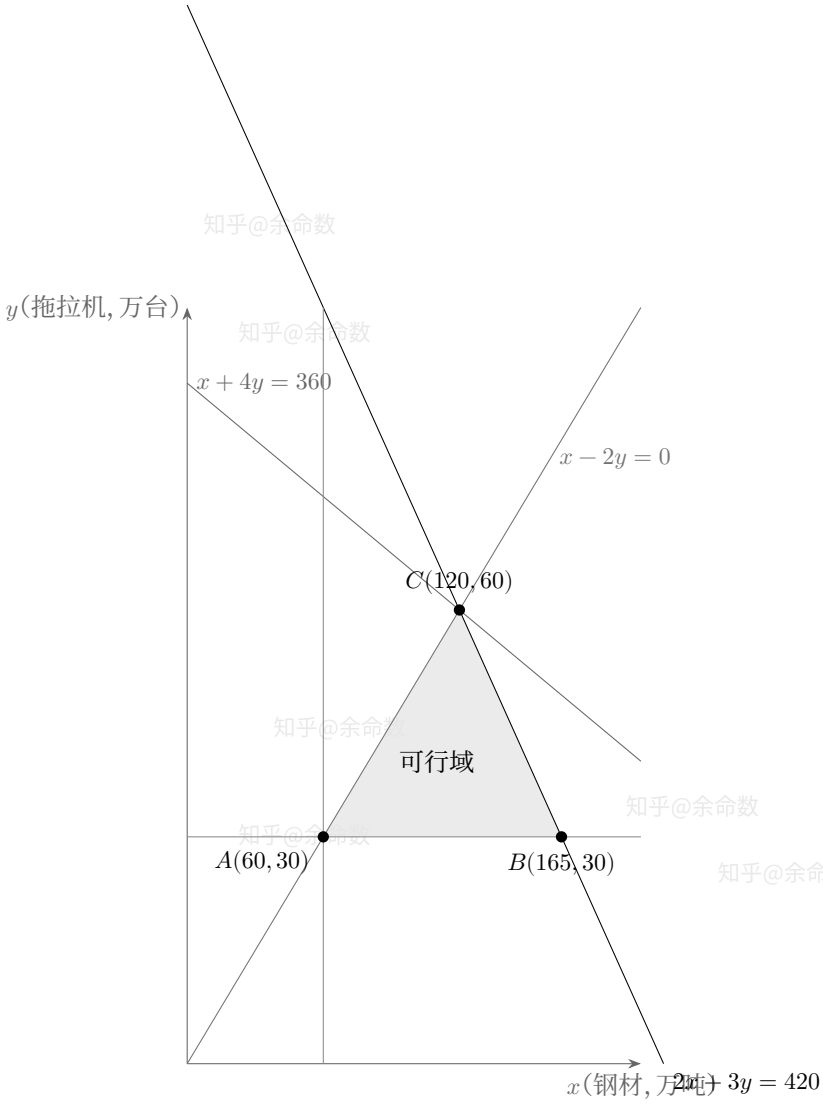


图 13.15 计划经济问题的可行域

知乎@余命数

(需验证其他交点是否构成可行域. 例如 $x = 2y$ 与 $x + 4y = 360$ 的交点也是 $(120, 60)$, 表明三条约束线在此交于一点.$x = 60$ 与 $x = 2y$ 的交点为 $(60, 30)$, 即顶点 A. 故可行域确实为三角形 ABC.)

最后, 将顶点代入目标函数 $z = 3x + 10y$ 进行检验: 比较可知, 最大的国家优

顶点坐标 (x, y)	优先指数 $z = 3x + 10y$
$A(60, 30)$	$3(60) + 10(30) = 180 + 300 = 480$
$B(165, 30)$	$3(165) + 10(30) = 495 + 300 = 795$
$C(120, 60)$	$3(120) + 10(60) = 360 + 600 = 960$

先指数为 960, 在顶点 $C(120, 60)$ 处取得.

因此, 计划委员会应下达生产 120 万吨钢材和 60 万台拖拉机的指标. 在此计划下, 电力消耗为 $5(120) + 20(60) = 600 + 1200 = 1800$, 达到上限. 劳动力消耗为 $10(120) + 15(60) = 1200 + 900 = 2100$, 达到上限. 内部钢材消耗为 $2(60) = 120$, 与钢材产量相等. 此解表明, 最优策略是将所有可用资源全部投入生产, 直到电力和劳动力同时耗尽, 并且生产的钢材恰好全部用于制造拖拉机, 没有剩余.

例题 13.36　假设你是国家计划委员会的规划专家, 需要为某一年的生产制定主要指标. 你的决策涉及三个主要部门: 钢铁、机械制造与农业, 后两者分别以机床和粮食为代表.

- **决策变量**:
  - 钢材年产量: x(万吨)
  - 机床年产量: y(万台)
  - 粮食年产量: z(百万吨)
- **可用资源约束**:
  - **国家投资总额**: 60 亿元. 生产 1 万吨钢材需投资 0.3 亿元; 1 万台机床需投资 0.8 亿元; 1 百万吨粮食需投资 0.1 亿元.
  - **技术工人总工时**: 3000 万工时. 生产 1 万吨钢材需 5 万工时; 1 万台机床需 20 万工时; 粮食生产主要依赖普通劳力, 对技术工人需求极小, 忽略不计.
- **内部结构与社会需求约束**:
  - **钢材内部消耗**: 每制造 1 万台机床, 需消耗 2 万吨钢材. 剩余的钢材才能用于其他建设.
  - **粮食基本需求**: 工业部门 (钢铁, 机床) 的工人需要粮食来维持生产. 假设每生产 1 万吨钢材或 1 万台机床, 需要消耗 0.2 百万吨粮食. 同时, 国家必须保证至少 80 百万吨的粮食储备与民用.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **最低工业指标**：为奠定工业化基础，钢材产量至少为 40 万吨，机床产量至少为 5 万台.
- **国家战略价值函数**：根据优先发展重工业，特别是装备制造业的国策，经测算，每单位产出的战略价值为：钢材 5 分，机床 12 分，粮食 2 分.

问：应如何制定 x, y, z 的年度生产计划，使得国家战略价值总分最高？

解 这是一个三维线性规划问题，其复杂性与真实性远超二维平面. 先系统地构建其数学模型.

目标函数是最大化国家战略价值 W：

$$\text{最大化} \quad W = 5x + 12y + 2z$$

约束条件如下：

1. 投资约束：$0.3x + 0.8y + 0.1z \le 60 \implies 3x + 8y + z \le 600$
2. 技术工人约束：$5x + 20y \le 3000 \implies x + 4y \le 600$
3. 钢材内部平衡：（此为隐性约束，净钢材为 $x - 2y$，必须非负）$x - 2y \ge 0$
4. 粮食需求约束：$z \ge 0.2x + 0.2y + 80$
5. 最低指标约束：$x \ge 40, y \ge 5$

三个变量下无法直接使用二维图解法，但可以利用投资约束消去 z. 目标函数中 z 的系数为正，因而在最优解处必会用尽投资预算；否则可以在不减少 x, y 的前提下增加 z，使 W 继续增大.

因此最优解满足

$$3x + 8y + z = 600, \quad z = 600 - 3x - 8y.$$

将其代入目标函数：

$$W = 5x + 12y + 2(600 - 3x - 8y) = 1200 - x - 4y.$$

所以最大化 W 等价于在可行条件下最小化 $x + 4y$.

同时，粮食需求约束给出

$$600 - 3x - 8y \ge 0.2x + 0.2y + 80$$

即

$$16x + 41y \le 2600.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

于是只需考察关于 x, y 的约束

$$\begin{cases} x + 4y \le 600 \\ 16x + 41y \le 2600 \\ x - 2y \ge 0 \\ x \ge 40 \\ y \ge 5 \end{cases}$$

在这些约束中,x 与 y 的下界分别为 $x \ge 40, y \ge 5$,且目标 $x+4y$ 的两个系数均为正. 因此其最小值在 $x = 40, y = 5$ 处取得. 该点满足 $x-2y = 30 \ge 0$、$x+4y = 60 \le 600$、$16x + 41y = 845 \le 2600$, 确为可行点.

代回投资约束求得

$$z = 600 - 3(40) - 8(5) = 440.$$

最终的战略价值总分为

$$W = 5(40) + 12(5) + 2(440) = 200 + 60 + 880 = 1140.$$

因此, 按照题设给出的目标函数与资源数据, 最优的年度生产计划是生产钢材 40 万吨, 机床 5 万台, 粮食 440 百万吨. 可见, 在线性规划建模中, 目标函数系数与资源消耗系数共同决定最优方向; 若希望模型体现“粮食只满足最低需求、资源优先投向工业”的政策目标, 则需要相应调整目标函数或增加额外约束.

13.2.5 金融领域的投资组合优化

金融学的主要议题之一是在不确定性中做出最优决策. 对任何投资者而言, 这都归结于一场在**回报**与**风险**之间的永恒博弈. 通常, 潜在回报越高的资产, 其价格波动的风险也越大. 一个纯粹追求最高回报的策略可能会导致无法承受的亏损, 而一个完全回避风险的策略则可能使资产价值停滞不前.

现代投资组合理论, 由诺贝尔奖得主哈里·马科维茨奠基, 为解决这一难题提供了科学的框架. 它主张通过**分散化**投资于多种不完全相关的资产, 在不牺牲预期回报的前提下降低整个投资组合的风险. 理论上, 存在一系列被称为*有效前沿*的投资组合, 它们在给定风险水平下提供了最高的回报.

虽然完整的马科维茨模型因其对风险 (通常用方差衡量) 的处理而属于二次规划的范畴, 但线性规划为解决这类问题提供了一个极其实用且易于理解的入门工具和近似方法. 在许多现实场景中, 风险可以用一系列线性的规则和约束有效控制, 从而将问题转化为一个标准的线性规划模型.

在线性规划的视角下, 投资组合优化的任务被清晰地解构为:

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **决策变量**: 决定投入到每一种金融资产, 如股票, 债券, 房地产等的资金数额 $x_1, x_2, \dots, x_n$.
- **目标函数**: 通常是最大化整个投资组合的总预期回报, 即 Maximize$Z = r_1x_1 + r_2x_2 + \cdots + r_nx_n$, 其中 r_i 是资产 i 的预期回报率.
- **约束条件**: 这部分把所有现实世界的限制和投资者的偏好转化为数学语言.
  - **预算约束**: $\sum x_i \le$ 总投资额.
  - **风险控制约束**: 对高风险资产类别的投资上限, 如 $\sum_{k\in\text{高风险}} x_k \le$ 风险敞口上限.
  - **分散化约束**: 对单一资产或行业的最低或最高投资比例, 如 $x_j \ge$ 最低要求, 或 $x_j \le$ 单一资产上限.
  - **监管与合规约束**: 法律或基金章程规定的其他任何线性限制.

通过这种方式, 线性规划将投资者的主观目标与客观限制, 即高回报、预算与风险, 转化为一个精确的数学问题, 并提供了一个保证找到最优解的系统性方法; 最优解未必唯一, 此时可行域的整条边, 或整个面上的点都同样最优.

例题 13.37 (风险与回报的权衡) 一位基金经理负责管理一笔总额为 100 万元的资金. 她的目标是在遵守基金的投资章程的前提下, 将这笔资金分配给两种类型的股票: 高增长科技股和蓝筹公用事业股, 以期获得最大的年度预期回报.

- **投资标的与预期回报**:
  - 高增长科技股 (T): 风险较高, 预期年回报率为 10%.
  - 蓝筹公用事业股 (U): 风险较低, 预期年回报率为 4%.
- **基金投资章程约束**:
  - **总预算**: 投入两种股票的总金额不得超过 100 万元.
  - **风险敞口限制**: 对高风险的科技股, 投资额不得超过 60 万元.
  - **稳健性要求**: 为保证投资组合的稳定性, 对公用事业股的投资额不得低于 10 万元.

问: 基金经理应如何分配资金, 才能使总的预期年回报最大化?

解 这是一个典型的投资组合优化问题, 可通过线性规划进行建模求解.

设投入高增长科技股的资金为 x 万元, 投入蓝筹公用事业股的资金为 y 万元. 目标是最大化总预期回报 Z(单位: 万元), 其目标函数为:

$$\text{最大化} \quad Z = 0.10x + 0.04y$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此决策必须满足基金章程规定的所有约束条件：

1. 预算约束：$x + y \le 100$
2. 风险敞口约束：$x \le 60$
3. 稳健性约束：$y \ge 10$
4. 非负约束：$x \ge 0$.

通过顶点检验法求解. 首先绘制由上述不等式组定义的可行域. 可行域是一

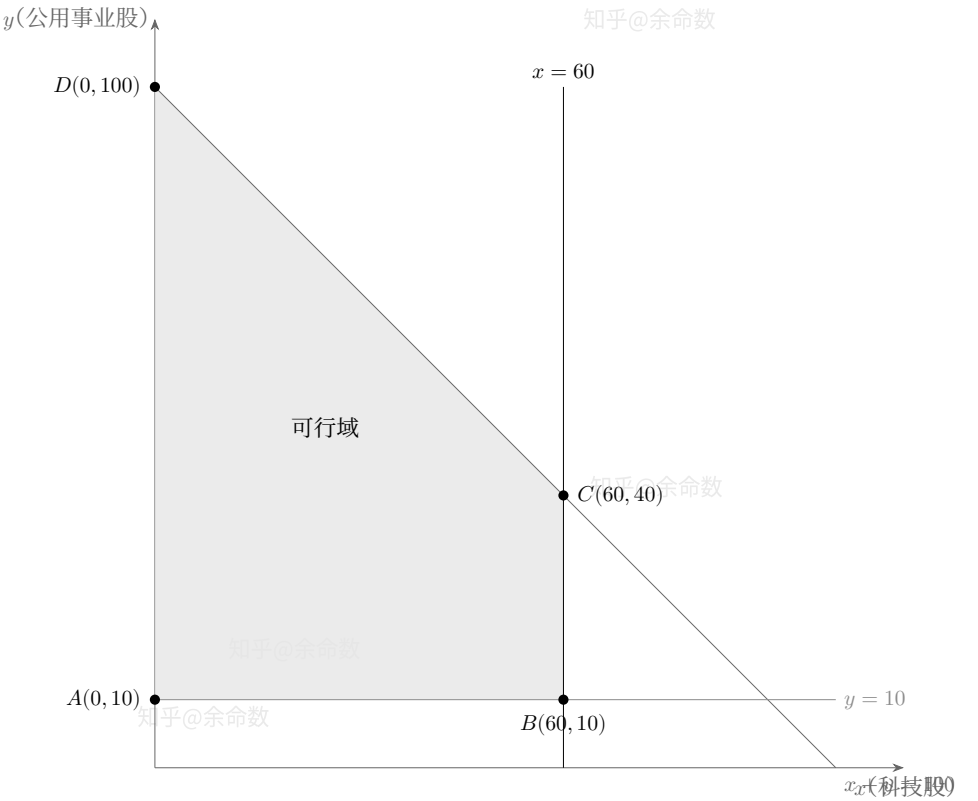


图 13.16　投资组合问题的可行域

个四边形，其顶点由边界线的交点构成：

- 顶点 A：$x = 0$ 与 $y = 10$ 的交点. 故 $A(0, 10)$.
- 顶点 B：$x = 60$ 与 $y = 10$ 的交点. 故 $B(60, 10)$.
- 顶点 C：$x = 60$ 与 $x + y = 100$ 的交点，解得 $y = 40$. 故 $C(60, 40)$.
- 顶点 D：$x = 0$ 与 $x + y = 100$ 的交点，解得 $y = 100$. 故 $D(0, 100)$.

根据线性规划基本定理，最大回报必在这些顶点之一处取得. 将各顶点坐标代入目标函数 $Z = 0.10x + 0.04y$ 进行检验：

比较上表中的回报值，可见最大值为 7.6 万元，它在顶点 $C(60, 40)$ 处取得.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

顶点坐标 (x,y)	预期回报 $Z=0.1x+0.04y$
$A(0,10)$	$0.1(0)+0.04(10)=0.4$
$B(60,10)$	$0.1(60)+0.04(10)=6+0.4=6.4$
$C(60,40)$	$0.1(60)+0.04(40)=6+1.6=7.6$
$D(0,100)$	$0.1(0)+0.04(100)=4$

因此, 最优的投资策略是将 60 万元投入高增长科技股, 将 40 万元投入蓝筹公用事业股. 这个解的经济学含义是, 为了追求最高回报, 基金经理应该将高风险资产的头寸加到规则允许的上限 ($x=60$), 然后将剩余的资金全部投入到稳健资产中, 直到用尽全部预算.

知乎@余命数

例题 13.38 (构建固定收益投资组合) 一家保险公司的资产管理部门需要将一笔 500 万元的资金投资于一个由国债和企业债构成的固定收益投资组合. 目标是在满足严格的风险和流动性要求下, 最大化年度利息总收入.

- **投资标的**:
  - A 类资产(政府主权债券 - 国债): 预期年收益率 3%, 风险评级为 1(极低风险).
  - B 类资产, 即高收益企业债券: 预期年收益率 7%, 风险评级为 5(中等风险).
- **投资规定与风控要求**:
  - **总预算**: 总投资额不超过 500 万元.
  - **平均风险控制**: 整个投资组合的加权平均风险评级不得超过 2.
  - **流动性要求**: 出于流动性考虑, 投资于国债的资金必须至少占总投资额的 40%.

问: 应如何配置国债和企业债的投资金额, 才能使年度利息收入最大化?

解 需在多重财务约束下, 寻找最优的资产配置方案. 将此问题构建为一个线性规划模型.

设投资于国债的金额为 x 万元, 投资于企业债的金额为 y 万元. 目标是最大化年度利息总收入 Z(单位: 万元), 其目标函数为:

$$\text{最大化}\quad Z=0.03x+0.07y$$

投资决策必须满足下列约束条件:

1. **预算约束**: $x+y\le 500$

知乎@余命数

2. **平均风险约束**：当实际投资额 $x+y>0$ 时，投资组合的加权平均风险评级由 $\frac{1\cdot x+5\cdot y}{x+y}$ 给出. 对非零投资组合有

$$\begin{aligned}\frac{x+5y}{x+y}\le 2 &\implies x+5y\le 2(x+y)\\ &\implies x+5y\le 2x+2y\\ &\implies 3y\le x\\ &\implies x-3y\ge 0.\end{aligned}$$

零投资点可视为退化可行点；它显然不会给最大收益，不影响最优解.

3. **流动性约束**：$x\ge 0.40(x+y)$. 化简得

$$\begin{aligned}x\ge 0.40(x+y) &\implies x\ge 0.4x+0.4y\\ &\implies 0.6x\ge 0.4y\\ &\implies 3x\ge 2y\\ &\implies 3x-2y\ge 0.\end{aligned}$$

4. **非负约束**：$x\ge 0, y\ge 0$.

通过顶点检验法求解. 首先确定可行域. 可行域是一个由原点和另外两个顶

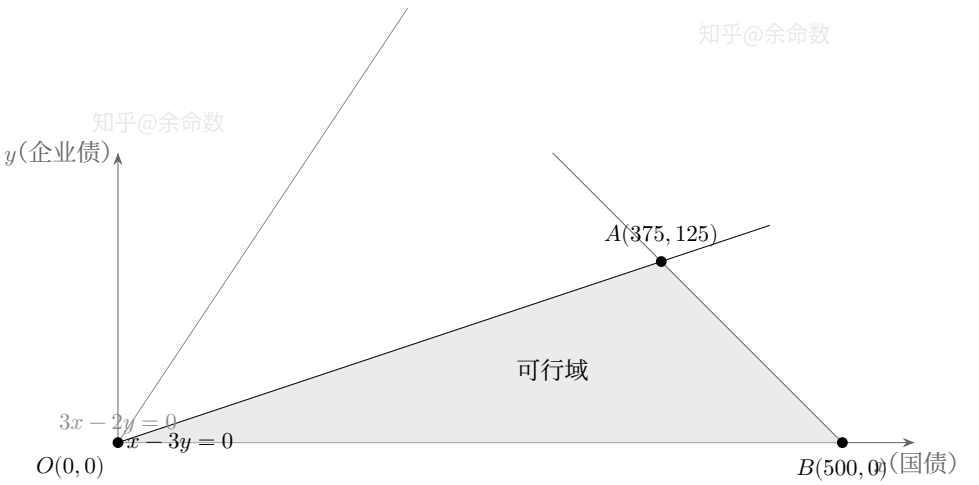


图 13.17　债券投资组合问题的可行域

点构成的三角形. 它的顶点由以下边界线的交点确定：(注意到 $3x-2y\ge 0 \implies y\le 1.5x$ 是一个比 $x-3y\ge 0 \implies y\le x/3$ 更宽松的约束，因此可行域的边界由 $y=x/3, x+y=500$ 和 $y=0$ 决定).

- 顶点 O：原点 $(0,0)$.
- 顶点 B：$x+y=500$ 与 $y=0$ 的交点. 故 $B(500,0)$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

- 顶点 A: $x+y=500$ 与 $x-3y=0$ 的交点. 联立求解, $3y+y=500 \implies 4y=500 \implies y=125$, 则 $x=375$. 故 $A(375,125)$.

(需检验顶点 A 是否满足所有约束.$3(375)-2(125)=1125-250>0$, 满足流动性约束. 故 A 是有效顶点).

将顶点代入目标函数 $Z=0.03x+0.07y$ 进行检验:

顶点坐标 (x,y)	利息收入 $Z=0.03x+0.07y$
$O(0,0)$	$0.03(0)+0.07(0)=0$
$B(500,0)$	$0.03(500)+0.07(0)=15$
$A(375,125)$	$0.03(375)+0.07(125)=11.25+8.75=20$

比较结果, 最大利息收入为 20 万元, 在顶点 $A(375,125)$ 处取得.

因此, 最优的资产配置方案是投资 375 万元于国债, 投资 125 万元于企业债. 此解的金融学含义是, 在满足预算的前提下, 风险控制主导决策. 投资者应将尽可能多的资金投入到高收益的企业债中, 直到组合的平均风险达到规定的上限 ($x-3y=0$). 此时, 任何进一步增加企业债投资的行为都将违反风险规定.

知乎@余命数

13.3 二次曲线

直线由一次方程 $Ax+By+C=0$ 给出, 圆是特殊的二次方程 $x^2+y^2+Dx+Ey+F=0$. 一个自然的下一步是把二次推到最一般的情形, 考察

$$Ax^2+Bxy+Cy^2+Dx+Ey+F=0$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

所定义的点集. 这一族曲线统称二**次曲线**, 也叫**圆锥曲线**, 包括椭圆(含圆)、抛物线、双曲线, 以及退化成一点、一条或两条直线乃至空集的若干情形. 行星轨道、光学反射面、炮弹弹道的简化模型, 都落在这一代数框架里.

13.3.1 二次曲线的一般理论

13.3.1.1 从几何定义到代数方程

三种经典曲线各有几何定义, 后文各小节会给出. **椭圆**是到两定点(焦点)距离之和为常数的点的轨迹. **双曲线**是到两定点距离之差的绝对值为常数的点的轨迹. **抛物线**是到定点与到定直线距离相等的点的轨迹, 定点称为焦点, 定直线称为准线.

取焦点在 x 轴、中心或顶点在原点、对称轴与坐标轴重合的坐标系, 把距离公式代入定义, 再两边平方消去根式, 得到**只含** x^2,y^2 **与常数**的方程, 例如

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1, \qquad \frac{x^2}{a^2}-\frac{y^2}{b^2}=1, \qquad y^2=2px.$$

知乎@余命数 知乎@余命数

这是标准方程. 它们都是二次方程, 但不含交叉项 xy, 也不含一次项 x, y（抛物线中 x 的一次项是顶点在原点、轴沿坐标轴时的特殊情形）.

现实中的曲线往往中心偏离原点、轴相对坐标轴倾斜. 要把标准形放到任意位置, 需要平面上的**刚体运动**. 刚体运动只改变图形在坐标系中的位置与朝向, 而**不改变**任何两点之间的距离, 因而也不改变长度、角度与形状. 在欧氏平面上, 任意刚体运动都可以分解为两步: **旋转**绕某点转过一个角 θ, 改变朝向; **平移**把图形整体沿某个向量挪动, 改变位置. 若还允许镜面反射, 就得到更广的等距变换; 但二次曲线的标准分类只需旋转加平移. 把曲线先绕原点逆时针旋转角 θ, 再平移到 (h, k), 则旧坐标 (x, y) 与贴着曲线的新坐标 (x', y') 之间满足

$$x = x'\cos\theta - y'\sin\theta + h,$$

$$y = x'\sin\theta + y'\cos\theta + k.$$

以椭圆 $\dfrac{x'^2}{a^2} + \dfrac{y'^2}{b^2} = 1$ 为例, 把上式代入并展开. 因为 x, y 都是 x', y' 的一次式, 而方程关于 x', y' 是二次的, 代回 (x, y) 之后总次数仍是 2, 并且一定会出现四类项: x^2, y^2 项来自二次型本身; xy 项来自旋转, 它使两个坐标混在一起; x, y 项来自平移, 它把原点挪开; 常数项则把原点处的方程值搬到新位置. 整理后恰是

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \tag{13-9}$$

双曲线、抛物线经同样的旋转加平移, 也落入同一形式. 平面上任意一个**总次数恰好为** 2 的实系数多项式方程, 都可以写成 (13-9), 其中 A, B, C 不全为零（否则退化为直线方程）. 这就是二次曲线的代数定义. 六个系数正是标准圆锥曲线经刚体运动后留下的全部自由度.

定义 13.4 (二次曲线的代数定义) 平面上一条**二次曲线**是所有满足方程

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{13-10}$$

的点 (x, y) 的集合, 其中 $A, B, C, D, E, F \in \mathbb{R}$, 且二次项系数 A, B, C 不全为零.

同一几何曲线的方程可以整体乘一个非零常数, 六个系数只有五个自由度. 系数不同时, 方程可以给出椭圆、抛物线、双曲线, 也可以给出退化情形（一点、一对相交直线、一对平行直线、一条二重直线, 或空集）. 圆锥曲线一称来自正圆锥面的截口.

13.3.1.2 六个系数各自做什么

把 (13-9) 拆成三块来读.

知乎@余命数

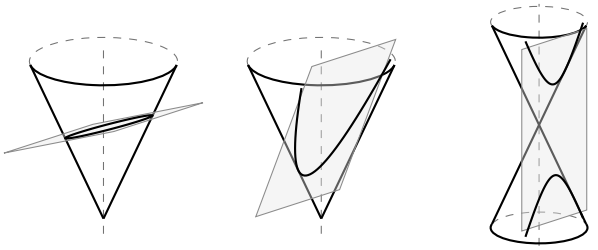


图 13.18 平面截正圆锥面的立体示意: (1) 截面与所有母线相交得椭圆, (2) 截面与一条母线平行得抛物线, (3) 截面切到对顶两半得双曲线. 截线是锥面与截面方程联立解出的精确曲线.

块	项	几何角色
二次型	$Ax^2+Bxy+Cy^2$	形状与朝向(椭圆 / 抛物 / 双曲类型, 以及轴相对坐标轴的倾斜)
一次项	$Dx+Ey$	中心或顶点相对原点的平移
常数项	F	与原点相对位置、整体尺度的定标(与 A 至 E 联立后决定大小)

下面分项说明.

1. A **与** C, 即对角二次项. 当 $B=0$ 时, 方程可写成 $Ax^2+Cy^2+\cdots=0$, 此时对称轴已与坐标轴平行. 若 A,C 同号且非零, 曲线属椭圆类型, 再经平移与缩放可到 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, 也可能退化为空集或一点. 若 A,C 异号, 曲线属双曲线类型. 若 A,C 恰有一个为零, 曲线属抛物线类型, 例如 $y^2=2px$ 对应 $A=0,\ C=1$. $|A|$ 与 $|C|$ 的相对大小决定沿 x 方向挤得多还是沿 y 方向挤得多. 椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 写成 $b^2x^2+a^2y^2-a^2b^2=0$ 时, $A=b^2,\ C=a^2$.
2. B(交叉项 xy). $B\neq 0$ 表示对称轴**相对坐标轴发生了旋转**. 例如把 $x^2-y^2=1$ 旋转 45°, 得到 $xy=\frac{1}{2}$ 这类方程, 此时 $A=C=0,\ B\neq 0$. 消掉 B 的办法是绕原点旋转坐标系; 旋转角由

$$\cot 2\theta=\frac{A-C}{B}\quad(B\neq 0)$$

决定. 若 $B=0$, 轴已与坐标轴平行, 无需旋转.

3. D **与** E, 即一次项. 它们编码**平移**. 把原点挪到椭圆与双曲线的中心, 或抛物线的顶点后, 适当选取时 $D=E=0$. 例如圆 $(x-h)^2+(y-k)^2=r^2$ 展开后 $D=-2h,\ E=-2k$. 对椭圆与双曲线这类有中心的曲线, 令 $x=x'+x_0,\ y=y'+y_0$ 代入 $Q=Ax^2+Bxy+Cy^2+Dx+Ey+F$, 并要求新方程中 x',y' 的一次项系数同时为零, 即得关于中心 (x_0,y_0) 的线性组

$$2Ax_0+By_0+D=0,\qquad Bx_0+2Cy_0+E=0.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当系数行列式 $4AC-B^2=-\Delta\neq 0$(此时必为椭圆型或双曲线型)时,中心唯一确定.

4. F(常数项). 固定二次型与一次项后, F 决定曲线相对于原点的尺度与是否实点. 对方程 $x^2+y^2+F=0$, 当 $F<0$ 时为圆, 当 $F=0$ 时为原点, 当 $F>0$ 时为空集. 整体乘非零常数会同时改写 F 与其他系数, 几何图形不变.

解析几何的常用操作是**坐标变换**,先旋转消掉 B,再平移消掉(中心型的)D,E,得到**标准方程**, 再从标准方程读出顶点、焦点、离心率等几何量. 曲线类型其实不必做完变换也能读出来, 关键在于一个旋转不变量.

13.3.1.3 判别式 $\Delta=B^2-4AC$ 的推导

我们只讨论绕原点的旋转, 平移不改变二次项系数 A,B,C, 也不改变 B^2-4AC. 令

$$x=x'\cos\theta-y'\sin\theta,$$

$$y=x'\sin\theta+y'\cos\theta,$$

并简记 $c=\cos\theta$, $s=\sin\theta$. 将

$$Q=Ax^2+Bxy+Cy^2$$

代入, 收集 x'^2, $x'y'$, y'^2 的系数, 得到

$$A'=Ac^2+Bcs+Cs^2,$$

$$B'=2(C-A)sc+B(c^2-s^2),$$

$$C'=As^2-Bcs+Cc^2.$$

利用倍角公式 $c^2-s^2=\cos 2\theta$, $2sc=\sin 2\theta$, 交叉项可写成

$$B'=B\cos 2\theta-(A-C)\sin 2\theta. \tag{13-11}$$

我们的目标是选 θ 使 $B'=0$. 若 $B=0$ 且 $A=C$, 任意 θ 都行, 此时曲线本就是圆型, 各向同性;若 $B=0$ 且 $A\neq C$, 取 $\theta=0$ 即可. 若 $B\neq 0$, 令 $B'=0$ 得

$$\cot 2\theta=\frac{A-C}{B},$$

在 $2\theta\in(0,\pi)$ 上总有解, 故**总能**消去交叉项, 使方程化为

$$A'x'^2+C'y'^2+D'x'+E'y'+F'=0.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

消交叉项之后, 类型由 A',C' 的符号决定. 若 $A'C'>0$, 曲线属椭圆类型, 含圆, 此时旋转后 $A'=C'$, 且可化为 $x'^2+y'^2=r^2$. 若 $A'C'=0$, 曲线属抛物线类型, 此时恰有一个二次项消失. 若 $A'C'<0$, 曲线属双曲线类型. 希望**不先算出** θ **与** A',C', 直接从原始 A,B,C 读出 $A'C'$ 的符号. 关键是证明

$$B'^2-4A'C'=B^2-4AC. \tag{13-12}$$

把 A',B',C' 的表达式代入左边, 先算 $A'+C'$ 与 $A'-C'$.

$$A'+C'=(Ac^2+Bcs+Cs^2)+(As^2-Bcs+Cc^2)=A+C,$$

$$A'-C'=A(c^2-s^2)+2Bcs+C(s^2-c^2)=(A-C)\cos 2\theta+B\sin 2\theta.$$

知乎@余命数

再由 (13-11) 知

$$B'=B\cos 2\theta-(A-C)\sin 2\theta.$$

知乎@余命数

于是

$$\begin{aligned}B'^2+(A'-C')^2&=\left[B\cos 2\theta-(A-C)\sin 2\theta\right]^2+\left[(A-C)\cos 2\theta+B\sin 2\theta\right]^2\\&=B^2(\cos^2 2\theta+\sin^2 2\theta)+(A-C)^2(\sin^2 2\theta+\cos^2 2\theta)\\&=B^2+(A-C)^2.\end{aligned}$$

余命数 知乎@余命数

另一方面, 恒等式

$$(A'-C')^2+B'^2=(A'+C')^2-4A'C'$$

知乎@余命数

给出

$$B'^2-4A'C'=B'^2+(A'-C')^2-(A'+C')^2=B^2+(A-C)^2-(A+C)^2.$$

右边化简为

知乎@余命数

$$B^2+(A-C)^2-(A+C)^2=B^2+(A^2-2AC+C^2)-(A^2+2AC+C^2)=B^2-4AC.$$

于是 (13-12) 成立, 即 B^2-4AC 在旋转下不变.

取定使 $B'=0$ 的 θ 后,

$$-4A'C'=B^2-4AC.$$

于是 $A'C'$ 的符号与 $-(B^2-4AC)$ 一致. 记

$$\Delta=B^2-4AC,$$

称为二次曲线的**判别式**(二次型判别式). 分类表如下.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$\Delta = B^2 - 4AC$	$A'C'$	类型	典型标准形(非退化)
$\Delta < 0$	> 0	椭圆型	$\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$(圆对应 $a = b$)
$\Delta = 0$	$= 0$	抛物线型	$y^2 = 2px$
$\Delta > 0$	< 0	双曲线型	$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$

我们补充三点, 避免把类型与实图形混为一谈.

1. Δ 只由二次项决定, 平移系数 D, E, F 不进入 Δ. 同一 Δ 下, 改 D, E, F 可把椭圆变成空集或一点, 把双曲线变成一对相交直线, 把抛物线变成一对平行直线等**退化**情形. 完整区分退化与否需要更大的不变量, 即 3×3 系数矩阵的行列式. 本书先用 Δ 判断椭圆、抛物与双曲大类, 再在标准形中用参数是否满足几何约束, 如 $a > c > 0$ 排除退化.
2. 当 $B = 0$ 且 $A = C \neq 0$ 时 $\Delta = -4A^2 < 0$, 属于椭圆型; 再配合一次项与常数, 可化为圆.
3. Δ 与一元二次方程 $at^2 + bt + c = 0$ 的判别式 $b^2 - 4ac$ 同形. 二次型 $Ax^2 + Bxy + Cy^2$ 在 $x \neq 0$ 的射线上可看作关于方向斜率 $t = y/x$ 的二次式 $A + Bt + Ct^2$, 其判别式正是 $B^2 - 4AC$. 它刻画各方向上二次型取值是否变号. 变号对应双曲型, 恒同号对应椭圆型, 恰有一方向为零对应抛物型.

下面从椭圆开始, 逐一给出几何定义、标准方程与性质.

13.3.2 椭圆

13.3.2.1 椭圆的几何定义

我们从一个不依赖坐标系的几何构造出发, 用两焦点和距离之和刻画椭圆.

定义 13.5 (椭圆) 平面内到两个定点 F_1, F_2 的距离之和等于一个常数(该常数必须大于两定点之间的距离)的所有点 P 的轨迹, 称为**椭圆**. 这两个定点 F_1, F_2 称为椭圆的**焦点**.

拉绳法给出这个定义的直观图像. 把一根长度固定的绳子的两端固定在 F_1, F_2 两图钉上, 用笔尖 P 拉紧绳子并移动, 笔尖经过的轨迹就是椭圆.

13.3.2.2 椭圆的标准方程

为了把上述几何定义写成代数方程, 我们建立一个坐标系. 最能简化计算的坐标系是把 x 轴放在两焦点所在直线上, 原点放在线段 F_1F_2 的中点. 设两焦点之间的距离为 $2c$ ($c > 0$), 则 $F_1(-c, 0), F_2(c, 0)$. 设点 P 到两焦点的距离之和为常数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

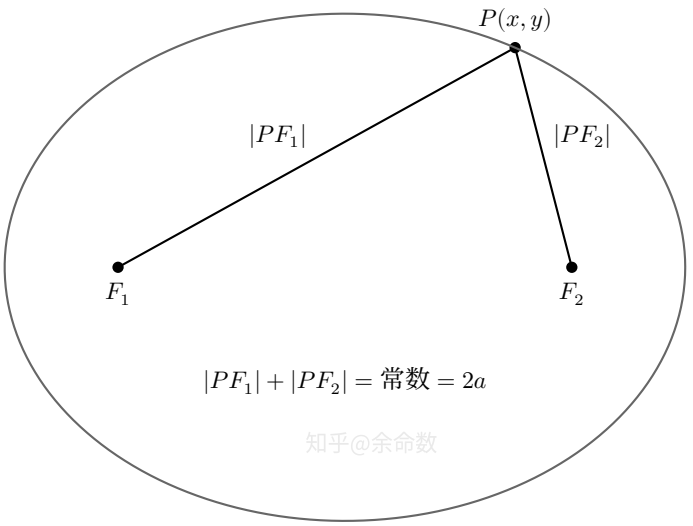


图 13.19 椭圆的焦点定义

$2a$. 根据三角形两边之和大于第三边, 必须有 $2a > |F_1F_2| = 2c$, 即 $a > c$. 设 P 的坐标为 (x, y). 根据定义, 有

$$|PF_1| + |PF_2| = 2a$$

利用两点间距离公式, 将其写为坐标形式:

$$\sqrt{(x+c)^2+y^2} + \sqrt{(x-c)^2+y^2} = 2a$$

这是含两个根式的方程. 我们通过代数运算消去根式. 首先, 将其中一个根式移到等式右侧:

$$\sqrt{(x+c)^2+y^2} = 2a - \sqrt{(x-c)^2+y^2}$$

对等式两边同时平方:

$$(x+c)^2+y^2 = 4a^2 - 4a\sqrt{(x-c)^2+y^2} + (x-c)^2 + y^2$$

展开并化简. 注意到 x^2, y^2, c^2 项在两侧被消去:

$$x^2+2xc+c^2+y^2 = 4a^2 - 4a\sqrt{(x-c)^2+y^2} + x^2 - 2xc + c^2 + y^2$$

$$4xc - 4a^2 = -4a\sqrt{(x-c)^2+y^2}$$

两边同除以 4, 整理得到

$$a^2 - cx = a\sqrt{(x-c)^2+y^2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为了消去最后的根式, 再对等式两边平方:

$$(a^2-cx)^2=a^2\left((x-c)^2+y^2\right)$$

$$a^4-2a^2cx+c^2x^2=a^2(x^2-2xc+c^2+y^2)$$

$$a^4-2a^2cx+c^2x^2=a^2x^2-2a^2cx+a^2c^2+a^2y^2$$

注意到 $-2a^2cx$ 项在两侧也被消去. 将含有变量的项移到一边, 常数项移到另一边:

$$a^2x^2-c^2x^2+a^2y^2=a^4-a^2c^2$$

知乎@余命数

$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2)$$

知乎@余命数

由于已假设 $a>c$, 故 a^2-c^2 为正数. 我们把这个量定义为 b:

$$b^2=a^2-c^2\quad(b>0)$$

将此式代入, 方程变为

$$b^2x^2+a^2y^2=a^2b^2$$

最后, 用 a^2b^2 除以方程两边(由 $a>c>0$ 知 $b^2>0$, 故 $a^2b^2\neq0$), 得到椭圆的标准方程:

知乎@余命数

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\quad(a>b>0)\tag{13-13}$$

若焦点位于 y 轴上, 即 $F_1(0,-c),F_2(0,c)$, 遵循同样的推导, 可得标准方程为 $\frac{x^2}{b^2}+\frac{y^2}{a^2}=1$. 通常我们约定 a 代表半长轴, 故 $a>b$ 总是成立.

知乎@余命数

13.3.2.3 椭圆的几何性质

椭圆的标准方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 编码了它的全部几何信息.

- **对称性**: 由于方程中 x,y 均以平方形式出现, 将 (x,y) 替换为 $(-x,y)$、$(x,-y)$ 或 $(-x,-y)$ 时方程均不变. 故椭圆关于 y 轴、x 轴以及原点中心对称.
- **范围**: 从方程可知

  $$\frac{x^2}{a^2}\leq1,\qquad\frac{y^2}{b^2}\leq1,$$

  知乎@余命数

  即 $|x|\leq a$ 且 $|y|\leq b$. 椭圆是有界图形, 完全落在直线 $x=\pm a$ 与 $y=\pm b$ 围成的矩形内部.
- **顶点**: 椭圆与对称轴的交点称为顶点, 即 $(\pm a,0)$ 与 $(0,\pm b)$.
- **轴**: 连接顶点 $(\pm a,0)$ 的线段称为**长轴**, 长度为 $2a$. 连接顶点 $(0,\pm b)$ 的线段称为**短轴**, 长度为 $2b$. 其中 a 为**半长轴长**, b 为**半短轴长**.

- **焦点与轴长的关系**：定义 $b^2=a^2-c^2$ 可改写为

$$a^2=b^2+c^2.$$

几何上，这对应以半长轴长 a 为斜边、半短轴长 b 与焦距 c 为直角边的直角三角形.

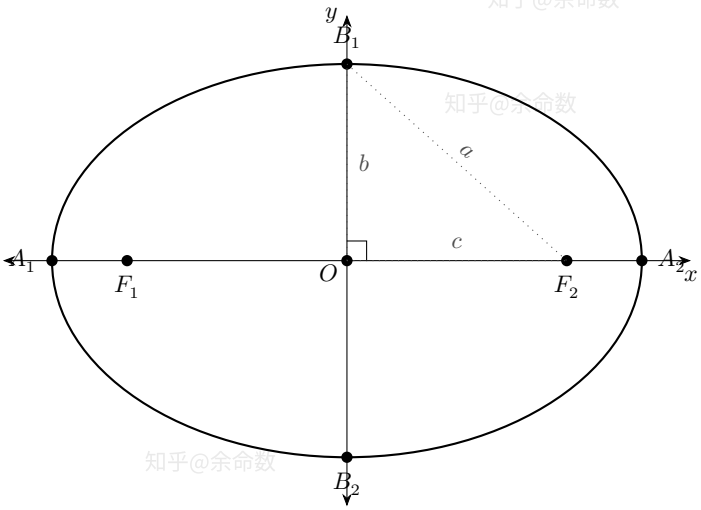


图 13.20 椭圆的标准方程几何要素，其中 $a^2=b^2+c^2$

例题 13.39 设 F_1 为椭圆 $\dfrac{x^2}{25}+\dfrac{y^2}{9}=1$ 的右焦点，AB 为过原点的弦，则 $\triangle ABF_1$ 面积的最大值为

A. 20　　C. $13\dfrac{1}{2}$

B. 15　　D. 12

解 我们需求以焦点为顶点、以中心弦为底边的三角形的最大面积.

先从椭圆方程 $\dfrac{x^2}{25}+\dfrac{y^2}{9}=1$ 中提取几何参数：半长轴 $a=5$，半短轴 $b=3$. 半焦距 $c=\sqrt{a^2-b^2}=\sqrt{25-9}=\sqrt{16}=4$. 由于 F_1 是右焦点，它的坐标为 $(4,0)$.

设弦 AB 是过原点的一条直线. 设其与椭圆的交点为 $A(x_A,y_A)$ 和 $B(x_B,y_B)$. 由于直线过原点且椭圆关于原点中心对称，A,B 也关于原点对称，即 $x_B=-x_A, y_B=-y_A$.

计算 $\triangle ABF_1$ 的面积. 把 $\triangle ABF_1$ 看成 $\triangle OAF_1$ 与 $\triangle OBF_1$ 拼接而成，它们共享底边 OF_1. 底边长为 $|OF_1|=c=4$. $\triangle OAF_1$ 的高是点 A 到 x 轴的距离，即 $|y_A|$. $\triangle OBF_1$ 的高是点 B 到 x 轴的距离，即 $|y_B|=|-y_A|=|y_A|$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

于是 $\triangle ABF_1$ 的总面积 S 为

$$\begin{aligned} S &= S_{\triangle OAF_1} + S_{\triangle OBF_1} \\ &= \frac{1}{2}|OF_1| \cdot |y_A| + \frac{1}{2}|OF_1| \cdot |y_B| \\ &= \frac{1}{2}c|y_A| + \frac{1}{2}c|y_A| \\ &= c|y_A|. \end{aligned}$$

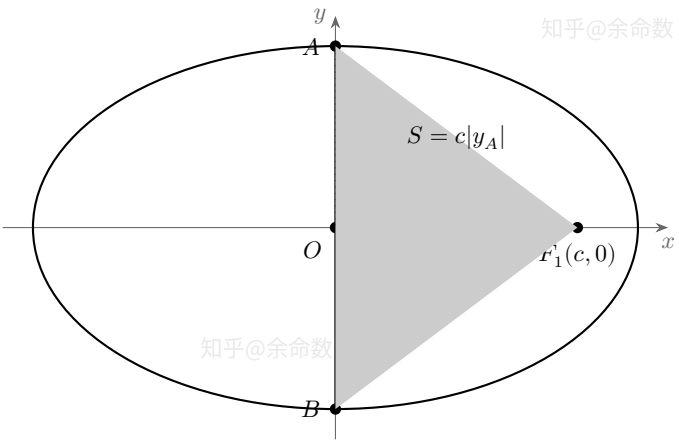


图 13.21　焦点-中心弦三角形面积的构成

要使面积 $S = c|y_A|$ 最大化，需最大化点 A 的纵坐标的绝对值 $|y_A|$. 椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ 上任意一点满足 $|y| \le b$, 最大值在短轴顶点 $x = 0$ 处取得.

因此, 当弦 AB 为椭圆的短轴时, $|y_A|$ 取得最大值 b, 此时 $\triangle ABF_1$ 的面积也达到最大值.

$$S_{\max} = c \cdot b$$

代入本题的数值, 得:

$$S_{\max} = 4 \times 3 = 12$$

故选 D.

这个例题的解法说明了一个普遍结论, 我们把它写成定理.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

定理 13.4 (焦点-中心弦三角形面积) 对于中心在原点的椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ $(a>b>0, c^2=a^2-b^2)$, 由一个焦点 $F(c,0)$ 和一条非长轴的过原点弦 AB 构成的非退化三角形 $\triangle ABF$, 其面积 S 的取值范围为:

$$0<S\le cb$$

- 当弦 AB 为短轴时, 面积取得**最大值** cb.
- 当弦 AB 趋近于长轴时(此时三角形退化为线段), 面积趋近于**最小值** 0.

证明 设椭圆的方程为 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$, 其右焦点为 $F(c,0)$. 设过原点的弦 AB 由直线 l 确定, A,B 为 l 与椭圆的两个交点. 由于椭圆关于原点中心对称, 若 $A(x_A,y_A)$, 则 $B(-x_A,-y_A)$.

计算 $\triangle ABF$ 的面积 S. 把 $\triangle ABF$ 的面积看成两个共底边的小三角形 $\triangle OAF$ 与 $\triangle OBF$ 的面积之和.

以 OF 所在的 x 轴为基准.

- $\triangle OAF$ 的底边为 $|OF|=c$, 其高为点 A 到 x 轴的距离, 即 $|y_A|$. 因此, $S_{\triangle OAF}=\frac{1}{2}c|y_A|$.
- $\triangle OBF$ 的底边也为 $|OF|=c$, 其高为点 B 到 x 轴的距离, 即 $|y_B|$. 由于 $y_B=-y_A$, 有 $|y_B|=|y_A|$. 因此, $S_{\triangle OBF}=\frac{1}{2}c|y_A|$.

将两者相加, 得到 $\triangle ABF$ 的总面积:

$$S=S_{\triangle OAF}+S_{\triangle OBF}=\frac{1}{2}c|y_A|+\frac{1}{2}c|y_A|=c|y_A|$$

因此, S 与 $|y_A|$ 成正比; 求 S 的最大值等价于求椭圆上 $|y_A|$ 的最大值.

椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 上任意一点 (x,y) 满足方程, 可推断 y 的取值范围:

$$\frac{y^2}{b^2}=1-\frac{x^2}{a^2}\le 1$$

$$y^2\le b^2 \implies |y|\le b$$

因此, 椭圆上任意一点的纵坐标的绝对值 $|y|$ 的最大值为 b, 且在 $x=0$ 处取得, 对应短轴顶点 $(0,\pm b)$.

把 $|y_A|$ 的最大值代入面积公式, 得到面积最大值:

$$S_{\max}=c\cdot b=cb$$

这种情况发生在 A 为 $(0,b)$ 或 $(0,-b)$ 时, 此时弦 AB 正是椭圆的短轴.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再考察面积的最小值.$|y_A|$ 的最小值为 0, 发生在 A 为长轴顶点 $(\pm a, 0)$ 时. 此时弦 AB 为椭圆的长轴. 在这种情况下, 三个点 A, B, F 共线, 三角形退化为一条线段, 其面积为 0. 对于任意非退化的三角形, 弦 AB 不能是长轴, 故 $|y_A| > 0$, 从而 $S > 0$. 面积可以随着弦 AB 任意地接近长轴而趋近于 0.

因此 S 的取值范围为 $0 < S \leq cb$. □

例题 13.40　在椭圆 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1 (a > b > 0)$ 上取三点, 其横坐标满足 $x_1 + x_3 = 2x_2$. 三点与某一个焦点连结的线段长分别为 r_1, r_2, r_3, 则 r_1, r_2, r_3 满足

A. $2r_2 = r_1 + r_3$

B. $\dfrac{1}{r_1} + \dfrac{1}{r_3} = \dfrac{2}{r_2}$

C. $r_2^2 = r_1 \cdot r_3$

D. 以上结论全不对

解　我们用椭圆的焦半径公式, 把 r_i 的关系转化为 x_i 的关系.

设这三点为 $P_1(x_1, y_1)$、$P_2(x_2, y_2)$、$P_3(x_3, y_3)$. 设它们与右焦点 $F_2(c, 0)$ 相连. (选择左焦点 $F_1(-c, 0)$ 会得到相同的结论). 根据前面证明的焦半径公式, 椭圆上一点 (x, y) 到右焦点 $F_2(c, 0)$ 的距离为

$$r = a - ex$$

其中 $e = \frac{c}{a}$ 是椭圆的离心率.

将此公式应用于给定的三点:

$$r_1 = a - ex_1$$

$$r_2 = a - ex_2$$

$$r_3 = a - ex_3$$

这是 r_i 与 x_i 之间的一阶线性关系.

题目条件是三点的横坐标构成等差数列, 即 $x_1 + x_3 = 2x_2$. 我们看 r_1, r_2, r_3 是否也成等差数列. 计算 $r_1 + r_3$, 得

$$\begin{aligned} r_1 + r_3 &= (a - ex_1) + (a - ex_3) \\ &= 2a - e(x_1 + x_3) \end{aligned}$$

将条件 $x_1 + x_3 = 2x_2$ 代入上式, 得

$$r_1 + r_3 = 2a - e(2x_2) = 2(a - ex_2)$$

知乎@余命数

注意到 $a-ex_2$ 正是 r_2 的表达式. 因此

$$r_1+r_3=2r_2$$

于是 r_1,r_2,r_3 也成等差数列.

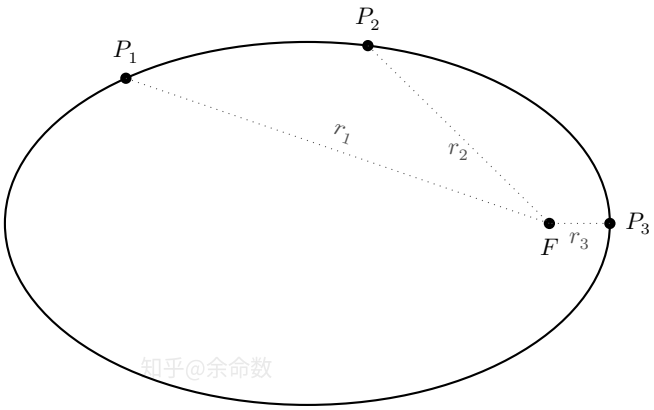


图 13.22 横坐标成等差数列的点的焦半径

知乎@余命数

(若选择左焦点 $F_1(-c,0)$, 焦半径公式为 $r=a+ex$. 同样地, $r_1+r_3=(a+ex_1)+(a+ex_3)=2a+e(x_1+x_3)=2a+e(2x_2)=2(a+ex_2)=2r_2$. 结论不变.)

故选 A. 这个结论说明了椭圆的一个内在性质, 即**焦半径的长度是其端点横坐标的线性函数**, 等差数列的结构因此可以在横坐标与焦半径之间传递.

13.3.2.4 椭圆的参数方程

标准方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 描述了坐标 (x,y) 满足的代数关系, 但也可把曲线看成质点运动的轨迹, 用一个独立变量 (通常代表时间或角度) 表示每个坐标. 这就是参数方程.

为得到椭圆的参数方程, 观察标准方程:

$$\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2=1$$

这个结构与三角恒等式 $\cos^2 t+\sin^2 t=1$ 形式相同. 令

$$\frac{x}{a}=\cos t,\quad \frac{y}{b}=\sin t$$

由此, 直接得到椭圆的**参数方程**:

$$\begin{cases}x=a\cos t\\ y=b\sin t\end{cases}\quad (t\text{为参数})\tag{13-14}$$

当 t 从 0 变到 2π 时, 点 (x,y) 恰好沿椭圆逆时针运行一周.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

参数 t 的几何意义需要辨析. 一个常见的误解是认为 t 是点 $P(x,y)$ 的位置向量与 x 轴正向的夹角, 即极角. 实际上, t 是**离心角**, 需借助一个辅助几何构造说明.

引入椭圆的**长轴辅助圆**, 它与椭圆同中心, 半径为 a, 方程为 $x^2+y^2=a^2$. 如

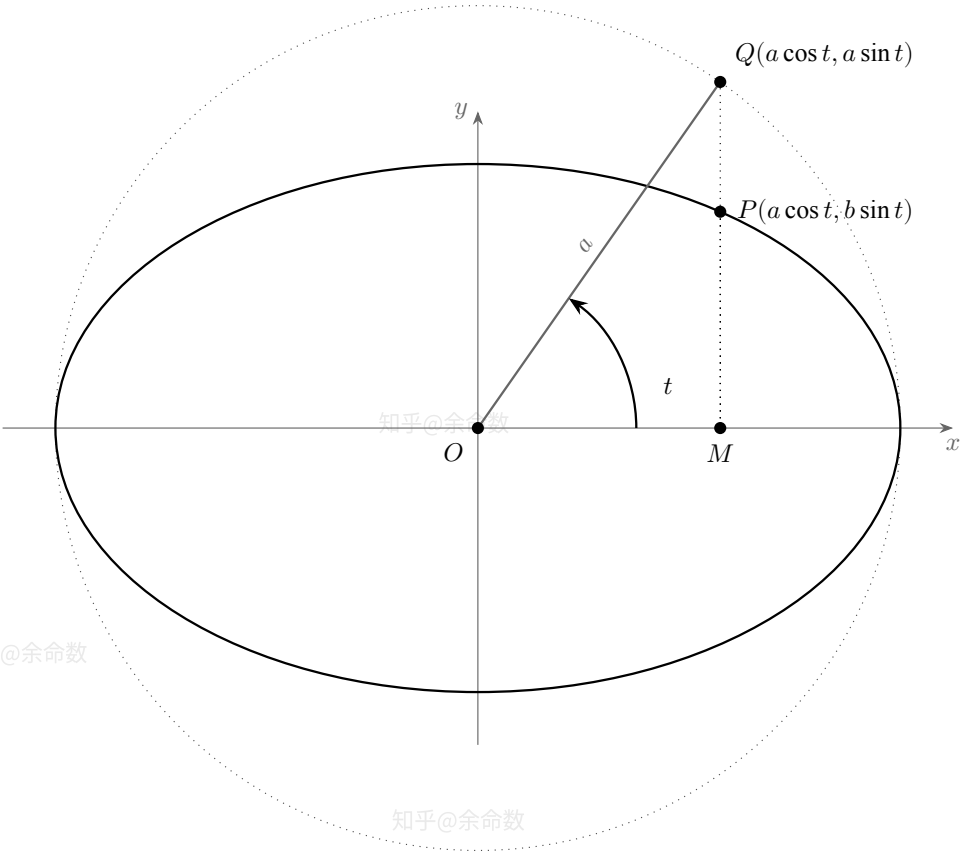


图 13.23　参数 t(离心角)的几何意义

图所示, 离心角 t 的构造如下.

1. 从椭圆上任意一点 $P(x,y)$ 向长轴即 x 轴作垂线, 垂足为 M.
2. 延长或缩短线段 MP 交长轴辅助圆于点 Q.
3. 则有向角 $\angle MOQ$, 其中 O 为原点的大小就是参数 t.

下面验证参数方程. 点 Q 位于半径为 a 的圆上, 且其极角为 t, 因此其坐标为 $(a\cos t, a\sin t)$. 由于点 P 与点 Q 在同一条垂线上, 它们的横坐标相同, 故 $x=a\cos t$. 将此 x 代入椭圆的标准方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$:

$$\frac{(a\cos t)^2}{a^2}+\frac{y^2}{b^2}=1 \implies \cos^2 t+\frac{y^2}{b^2}=1$$

$$\frac{y^2}{b^2}=1-\cos^2 t=\sin^2 t \implies y^2=b^2\sin^2 t$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由此可得 $y=\pm b\sin t$. 结合点 P 与 Q 的位置关系 (同在 x 轴上方或下方), 可知 $y=b\sin t$. 这一构造解释了离心角 t 的含义, 也给出了从圆到椭圆的视角, 把长轴辅助圆沿短轴方向均匀压缩, 或拉伸 $\frac{b}{a}$ 倍, 即得椭圆.

参数方程在切线、面积、弧长与轨道等问题中都很方便.

例题 13.41 设椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1(a>b>0)$ 与 y 轴正向的交点为 B. 求以 B 为直角顶点的椭圆内接等腰直角三角形的个数.

解 我们要求以短轴顶点 B 为直角顶点的椭圆内接等腰直角三角形个数. 个数由离心率决定, 需对方程根的存在性分类讨论.

椭圆与 y 轴正向交于 $B(0,b)$. 设等腰直角三角形为 $\triangle BPQ$, 其中 P,Q 为椭圆上另两点. 取腰 BP 的方向向量为 $(\cos\theta,-\sin\theta)$, 与之垂直且朝向椭圆内部的腰 BQ 取方向 $(-\sin\theta,-\cos\theta)$.

知乎@余命数

射线 BP 的参数方程为 $\begin{cases}x=t\cos\theta\\ y=b-t\sin\theta\end{cases}$ $(t>0)$. 把它代入椭圆方程 $b^2x^2+a^2y^2=a^2b^2$:

$$b^2(t\cos\theta)^2+a^2(b-t\sin\theta)^2=a^2b^2$$

展开整理, 得到关于 $t=|BP|$ 的二次方程:

$$(b^2\cos^2\theta+a^2\sin^2\theta)t^2-(2a^2b\sin\theta)t=0$$

此方程的非零解即为 $|BP|$:

$$|BP|=\frac{2a^2b\sin\theta}{a^2\sin^2\theta+b^2\cos^2\theta}$$

同理, 对于射线 BQ, 其参数方程为 $\begin{cases}x=-t\sin\theta\\ y=b-t\cos\theta\end{cases}$ $(t>0)$. 代入椭圆方程得 $|BQ|$:

$$|BQ|=\frac{2a^2b\cos\theta}{a^2\cos^2\theta+b^2\sin^2\theta}$$

由等腰条件 $|BP|=|BQ|$:

$$\frac{\sin\theta}{a^2\sin^2\theta+b^2\cos^2\theta}=\frac{\cos\theta}{a^2\cos^2\theta+b^2\sin^2\theta}$$

交叉相乘并整理, 得

$$\sin\theta(a^2\cos^2\theta+b^2\sin^2\theta)=\cos\theta(a^2\sin^2\theta+b^2\cos^2\theta)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

假设 $\cos\theta \neq 0$, 两边同除以 $\cos^3\theta$, 得

$$\tan\theta(a^2 + b^2\tan^2\theta) = a^2\tan^2\theta + b^2$$

令 $T = \tan\theta$, 得到关于 T 的三次方程:

$$b^2T^3 - a^2T^2 + a^2T - b^2 = 0$$

注意到 $T = 1$ 是该方程的一个根, 因为 $b^2(1)^3 - a^2(1)^2 + a^2(1) - b^2 = 0$. 因此, 可分解因式:

$$(T-1)[b^2T^2 + (b^2 - a^2)T + b^2] = 0$$

三角形个数取决于该方程实数根的个数. 其中一个根为 $T = \tan\theta = 1$, 对应 $\theta = 45^\circ$, 此时两腰关于 y 轴对称, 底边 PQ 水平. 这种三角形总存在.

其余可能解来自二次方程 $b^2T^2 + (b^2 - a^2)T + b^2 = 0$. 判别式

$$\Delta = (b^2 - a^2)^2 - 4b^4.$$

知乎@余命数

由 $c^2 = a^2 - b^2$ 即 $b^2 - a^2 = -c^2$, 得 $\Delta = c^4 - 4b^4 = (c^2 - 2b^2)(c^2 + 2b^2)$. 因 $c^2 + 2b^2 > 0$, Δ 的符号由 $c^2 - 2b^2$ 决定:

$$c^2 - 2b^2 > 0 \iff 3c^2 > 2a^2 \iff e > \frac{\sqrt{6}}{3}.$$

分类讨论如下:

1. **当** $e > \frac{\sqrt{6}}{3}$, 即 $c^2 > 2b^2$, $\Delta > 0$. 此时二次方程有两个不相等的实数根 T_1, T_2. 由于二次方程的两根乘积为 1、两根之和为 $\frac{c^2}{b^2} > 0$, 它们均为正数, 因而都对应某个 $\theta \in (0, \frac{\pi}{2})$. 注意 $T = 1$ 不是该二次方程的根(代入得 $3b^2 - a^2 \neq 0$, 除非 $c^2 = 2b^2$). 于是共有三个不同的可用实根 $1, T_1, T_2$. 每个根对应一个 θ 值, 从而对应一个等腰直角三角形. 故此时共有 **3** 个这样的三角形.
2. **当** $e = \frac{\sqrt{6}}{3}$, 即 $c^2 = 2b^2$, $\Delta = 0$. 此时二次方程有一个重根 $T = -\frac{b^2 - a^2}{2b^2} = \frac{c^2}{2b^2} = \frac{2b^2}{2b^2} = 1$. 即三次方程有三重根 $T = 1$. 因此只存在一个 $\tan\theta = 1$ 的解. 故此时只有 **1** 个这样的三角形.
3. **当** $e < \frac{\sqrt{6}}{3}$, 即 $c^2 < 2b^2$, $\Delta < 0$. 此时二次方程没有实数根. 三次方程唯一的实数根是 $T = 1$. 故此时也只有 **1** 个这样的三角形.

椭圆内接等腰直角三角形的个数取决于离心率 e:

- 若 $e > \frac{\sqrt{6}}{3}$, 存在 3 个.
- 若 $0 < e \leq \frac{\sqrt{6}}{3}$, 存在 1 个.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.3.2.5 椭圆的切线方程

推导经过椭圆上一点的切线方程. 下面把隐式微分与代数变形结合起来. 考虑标准椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. 为求任意一点 $P(x_0,y_0)$ 处的切线斜率, 对方程两边关于 x 隐式微分:

$$\frac{d}{dx}\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}\right)=\frac{d}{dx}(1)$$

$$\frac{2x}{a^2}+\frac{2y}{b^2}\frac{dy}{dx}=0$$

由此解出切线斜率 $\frac{dy}{dx}$:

$$\frac{dy}{dx}=-\frac{b^2x}{a^2y}$$

在点 $P(x_0,y_0)$ 处, 切线斜率为 $m=-\frac{b^2x_0}{a^2y_0}$. 由点斜式 $y-y_0=m(x-x_0)$, 得:

$$y-y_0=-\frac{b^2x_0}{a^2y_0}(x-x_0)$$

为把它化为更简洁对称的形式, 两边同乘 a^2y_0:

$$a^2y_0(y-y_0)=-b^2x_0(x-x_0)$$

$$a^2yy_0-a^2y_0^2=-b^2xx_0+b^2x_0^2$$

将含变量的项移到一边:

$$b^2xx_0+a^2yy_0=b^2x_0^2+a^2y_0^2$$

注意到点 $P(x_0,y_0)$ 在椭圆上, 故 $\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}=1$, 即 $b^2x_0^2+a^2y_0^2=a^2b^2$. 代入切线方程, 得:

$$b^2xx_0+a^2yy_0=a^2b^2$$

最后, 两边同除以 a^2b^2, 得到椭圆切线方程的标准形式:

$$\frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1 \tag{13-15}$$

这个形式具有替换对称性, 把标准方程中的一个 x 与一个 y 分别换成切点坐标 x_0,y_0 就得到切线方程.

若切点由参数方程给出, $P(a\cos t,b\sin t)$, 将 $x_0=a\cos t,y_0=b\sin t$ 代入上式, 得切线方程的参数形式:

$$\frac{x(a\cos t)}{a^2}+\frac{y(b\sin t)}{b^2}=1$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\frac{x\cos t}{a}+\frac{y\sin t}{b}=1 \tag{13-16}$$

例题 13.42　证明自椭圆的两焦点向其任意切线所作垂线段长的乘积为定值.

证明　下面用参数方程证明这个椭圆的几何性质. 考虑中心在原点, 焦点在 x 轴上的椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. 其焦点为 $F_1(-c,0)$ 和 $F_2(c,0)$. 利用参数方程, 椭圆上任意一点 P 可表示为 $(a\cos\theta, b\sin\theta)$. 据此, 过此点的切线方程为:

$$\frac{x\cos\theta}{a}+\frac{y\sin\theta}{b}=1$$

为使用点到直线的距离公式, 把它写成一般形式:

$$\left(\frac{\cos\theta}{a}\right)x+\left(\frac{\sin\theta}{b}\right)y-1=0$$

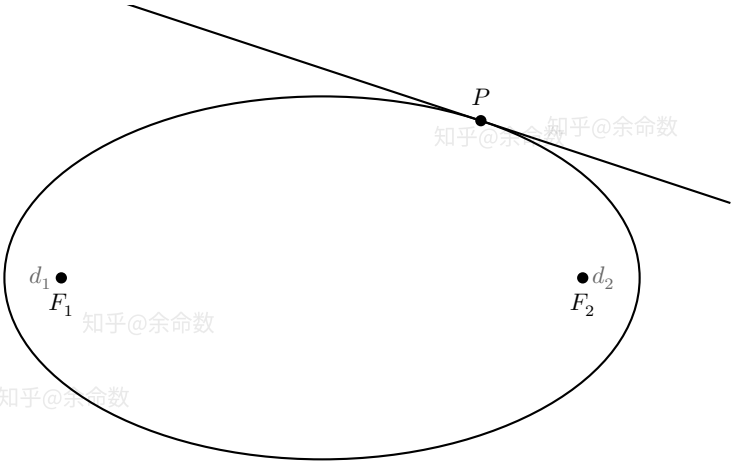


图 13.24　焦点到切线的垂线段

设两焦点到该切线的距离分别为 d_1 和 d_2. 由点到直线的距离公式:

$$d_1=\frac{\left|\frac{\cos\theta}{a}(-c)+\frac{\sin\theta}{b}(0)-1\right|}{\sqrt{\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}}}=\frac{\left|-\frac{c\cos\theta}{a}-1\right|}{\sqrt{\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}}}$$

$$d_2=\frac{\left|\frac{\cos\theta}{a}(c)+\frac{\sin\theta}{b}(0)-1\right|}{\sqrt{\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}}}=\frac{\left|\frac{c\cos\theta}{a}-1\right|}{\sqrt{\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}}}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

下面计算乘积 d_1d_2, 分子分母分别相乘:

$$d_1d_2=\frac{\left|\left(-\frac{c\cos\theta}{a}-1\right)\left(\frac{c\cos\theta}{a}-1\right)\right|}{\frac{\cos^2\theta}{a^2}+\frac{\sin^2\theta}{b^2}}=\frac{\left|1-\frac{c^2\cos^2\theta}{a^2}\right|}{\frac{b^2\cos^2\theta+a^2\sin^2\theta}{a^2b^2}}$$

分别化简分子和分母. 对于分子:

$$\left|1-\frac{c^2\cos^2\theta}{a^2}\right|=\frac{|a^2-c^2\cos^2\theta|}{a^2}$$

对于分母:

$$\begin{aligned}b^2\cos^2\theta+a^2\sin^2\theta&=b^2\cos^2\theta+a^2(1-\cos^2\theta)\\&=b^2\cos^2\theta+a^2-a^2\cos^2\theta\\&=a^2-(a^2-b^2)\cos^2\theta\\&=a^2-c^2\cos^2\theta\end{aligned}$$

把化简后的结果代回原式:

$$d_1d_2=\frac{\frac{|a^2-c^2\cos^2\theta|}{a^2}}{\frac{a^2-c^2\cos^2\theta}{a^2b^2}}=\frac{|a^2-c^2\cos^2\theta|}{a^2}\cdot\frac{a^2b^2}{a^2-c^2\cos^2\theta}$$

由于 $a>c$ 且 $|\cos\theta|\leq 1$, 有 $a^2>c^2\geq c^2\cos^2\theta$, 因此 $a^2-c^2\cos^2\theta$ 恒正, 绝对值可以去掉.

$$d_1d_2=\frac{a^2-c^2\cos^2\theta}{a^2}\cdot\frac{a^2b^2}{a^2-c^2\cos^2\theta}=b^2$$

这个常数 b^2 不依赖切点位置, 即不依赖 θ. □

13.3.2.6 椭圆的光学与声学性质

我们来看椭圆在光学与声学中的一个重要反射性质.

定理 13.5 (椭圆的反射性质) 从椭圆一个焦点发出的射线(光线或声波), 经椭圆内壁一次反射后必汇聚于另一焦点.

证明 根据物理学中的反射定律, 入射角等于反射角. 在几何上, 这等价于证明: 在椭圆上任意一点 P 处的切线, 是两条焦半径 F_1P 和 F_2P 所构成的角的外角平分线. 一个更便于向量分析的等价命题是: **该点处的法线是 $\angle F_1PF_2$ 的内角平分线**.

设椭圆方程为 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, 焦点为 $F_1(-c,0),F_2(c,0)$. 取椭圆上任意一点 $P(x_0,y_0)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

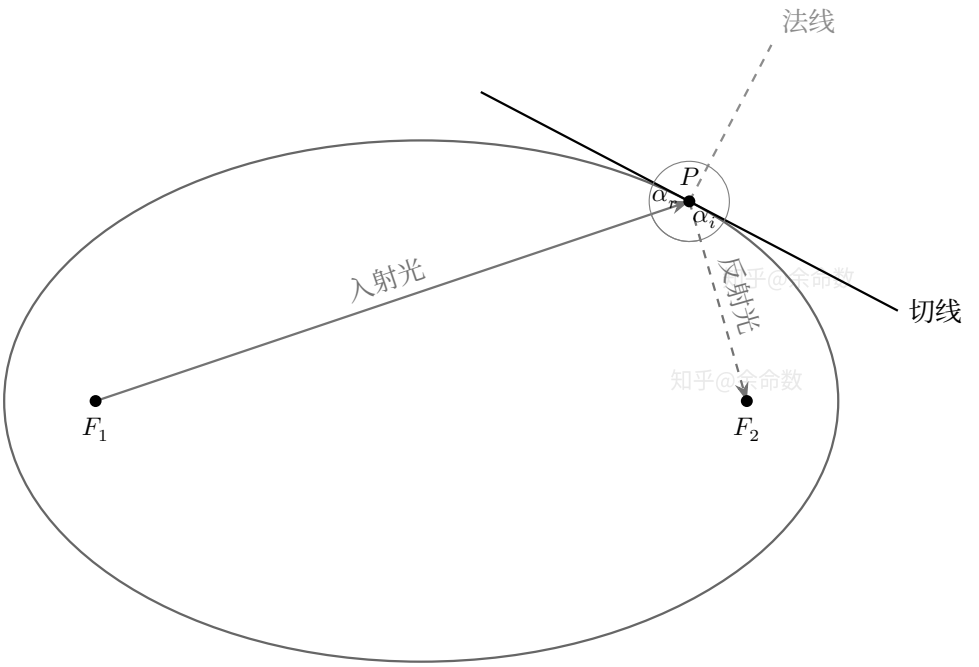


图 13.25　椭圆的反射性质：入射角 α_i 等于反射角 α_r.

要证明法线平分 $\angle F_1PF_2$，我们用向量法证明法向量与 $\overrightarrow{PF_1}, \overrightarrow{PF_2}$ 的夹角相等. 过点 $P(x_0, y_0)$ 的切线方程为 $\frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1$，故其一个法向量为 $\vec{n} = (\frac{x_0}{a^2}, \frac{y_0}{b^2})$.

向量 $\overrightarrow{PF_1} = (-c - x_0, -y_0)$ 和 $\overrightarrow{PF_2} = (c - x_0, -y_0)$. 在计算夹角之前，先证明一个引理.

引理 13.2 (焦半径公式) 对于椭圆上一点 $P(x_0, y_0)$，其到两焦点的距离为

$$|PF_1| = a + ex_0, \quad |PF_2| = a - ex_0$$

其中 $e = \frac{c}{a}$ 是离心率.

证明 直接计算 $|PF_2|^2$，得

$$\begin{aligned}
|PF_2|^2 &= (x_0 - c)^2 + y_0^2 = x_0^2 - 2cx_0 + c^2 + b^2\left(1 - \frac{x_0^2}{a^2}\right) \\
&= x_0^2 - 2cx_0 + c^2 + b^2 - \frac{b^2x_0^2}{a^2} \\
&= x_0^2 - 2cx_0 + c^2 + (a^2 - c^2) - \frac{(a^2 - c^2)x_0^2}{a^2} \quad (\text{代入} b^2 = a^2 - c^2) \\
&= a^2 - 2cx_0 + x_0^2\frac{c^2}{a^2} = a^2 - 2a\left(\frac{c}{a}\right)x_0 + \left(\frac{c}{a}x_0\right)^2 \\
&= (a - ex_0)^2.
\end{aligned}$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

由于 P 在椭圆上, $|x_0| \le a$ 且 $0 < e < 1$, 故 $a - ex_0 > 0$, 因此 $|PF_2| = a - ex_0$. 由椭圆定义 $|PF_1| + |PF_2| = 2a$, 得 $|PF_1| = 2a - (a - ex_0) = a + ex_0$. □

下面计算法向量 $\vec{n}$ 与焦半径向量 $\overrightarrow{PF_1}, \overrightarrow{PF_2}$ 的夹角余弦.

$$\begin{aligned}\cos\alpha_1 = \cos(\vec{n}, \overrightarrow{PF_1}) &= \frac{\vec{n}\cdot\overrightarrow{PF_1}}{|\vec{n}||\overrightarrow{PF_1}|} = \frac{(\frac{x_0}{a^2}, \frac{y_0}{b^2})\cdot(-c-x_0, -y_0)}{|\vec{n}||PF_1|} \\ &= \frac{-\frac{cx_0}{a^2} - \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}}{|\vec{n}|(a+ex_0)} = \frac{-\frac{c}{a}\frac{x_0}{a} - \left(\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2}\right)}{|\vec{n}|(a+ex_0)} \\ &= \frac{-e\frac{x_0}{a} - 1}{|\vec{n}|(a+ex_0)} = -\frac{1}{a|\vec{n}|}\frac{ex_0+a}{a+ex_0} = -\frac{1}{a|\vec{n}|}\end{aligned}$$

$$\begin{aligned}\cos\alpha_2 = \cos(\vec{n}, \overrightarrow{PF_2}) &= \frac{\vec{n}\cdot\overrightarrow{PF_2}}{|\vec{n}||\overrightarrow{PF_2}|} = \frac{(\frac{x_0}{a^2}, \frac{y_0}{b^2})\cdot(c-x_0, -y_0)}{|\vec{n}||PF_2|} \\ &= \frac{\frac{cx_0}{a^2} - \frac{x_0^2}{a^2} - \frac{y_0^2}{b^2}}{|\vec{n}|(a-ex_0)} = \frac{e\frac{x_0}{a} - 1}{|\vec{n}|(a-ex_0)} \\ &= \frac{1}{a|\vec{n}|}\frac{ex_0-a}{a-ex_0} = -\frac{1}{a|\vec{n}|}\end{aligned}$$

由 $\cos\alpha_1 = \cos\alpha_2$ 且两向量夹角都在 $[0, \pi]$, 得 $\alpha_1 = \alpha_2$. 故法线平分 $\angle F_1PF_2$. 由反射定律, 从 F_1 发出并经 P 反射的光线与 $\overrightarrow{PF_2}$ 共线, 即朝向另一焦点 F_2. □

证明 (平面几何法) 我们对椭圆上任意点 P, 证明两条焦半径 F_1P, F_2P 与切线所成角相等.

设 l 为椭圆在 P 处的切线. 由凸性, l 与椭圆只交于 P, 且整个椭圆位于 l 一侧. 延长 F_1P 至 G, 使 $|PG| = |PF_2|$.

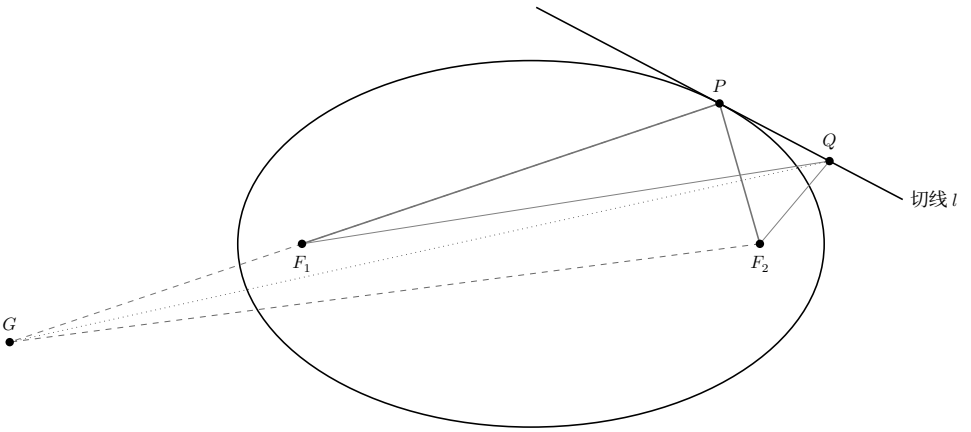


图 13.26 椭圆反射性质的几何证明

由椭圆定义, $|F_1P| + |PF_2| = 2a$. 结合构造得 $|F_1G| = |F_1P| + |PG| = 2a$, 故 G 落在以 F_1 为圆心、半径 $2a$ 的圆上.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

考察 $\triangle F_2PG$. 由于 $|PG|=|PF_2|$, 这是以 P 为顶点的等腰三角形.

下面证明切线 l 平分外角 $\angle F_2PG$. 在等腰三角形 $\triangle F_2PG$ 中, 顶点角 $\angle F_2PG$ 的平分线必是底边 F_2G 的垂直平分线. 因此, 任务转化为证明: **切线 l 是线段 F_2G 的垂直平分线**.

用反证法. 假设切线 l 不是 F_2G 的垂直平分线. 设 l' 是 F_2G 的垂直平分线. 由于 $\triangle F_2PG$ 是等腰三角形, l' 必过顶点 P.

取 l' 上异于 P 的任意一点 Q. 由于 l' 是 F_2G 的中垂线, 有 $|QF_2|=|QG|$. 计算点 Q 的焦半径之和, 得

$$|QF_1|+|QF_2|=|QF_1|+|QG|$$

知乎@余命数

在 $\triangle QF_1G$ 中, 由三角形两边之和大于第三边, 有

$$|QF_1|+|QG|>|F_1G|$$

而已知 $|F_1G|=2a$, 故

$$|QF_1|+|QF_2|>2a$$

知乎@余命数

这个不等式表明, 直线 l' 上除 P 外的任一点 Q 都在椭圆外部. 根据切线定义, l' 就是椭圆在点 P 处的切线. 因此, $l'=l$.

已证明椭圆在点 P 处的切线 l 就是 F_2G 的垂直平分线. l 是 F_2G 的中垂线, 其上任意一点到 F_2,G 的距离相等, 故 $\angle F_2Pl=\angle GPl$. 由于 $\angle GPl$ 与 $\angle F_1Pl'$, 其中 l' 是 l 的反向延长线是对顶角, 它们相等. 因此 $\angle F_2Pl=\angle F_1Pl'$, 即入射光线 F_1P 的延长线与反射光线 PF_2 关于法线对称. 故从 F_1 出发的光线经反射后必经过 F_2. □

知乎@余命数　知乎@余命数

例题 13.43　已知椭圆 $\dfrac{x^2}{25}+\dfrac{y^2}{16}=1$ 中的两焦点 $F_1(3,0),F_2(-3,0)$, 由点 F_1 射出一光线至椭圆上一点 A_1, 经反射后经过另一焦点 F_2 后至椭圆上另一点 A_2, 再反射回到焦点 F_1, 则所经过的路径长度为何?

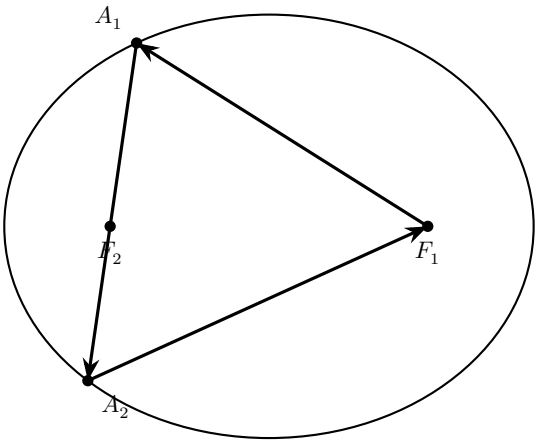
解　由椭圆方程 $\frac{x^2}{25}+\frac{y^2}{16}=1$ 读出 $a^2=25,b^2=16$, 故 $a=5$(焦距 $c=\sqrt{a^2-b^2}=3$, 与题设焦点一致). 光线路径分为三段: $F_1\to A_1$, 反射后经 F_2 到 A_2, 再反射回 F_1. 椭圆反射性质保证 A_1 处反射线过 F_2, 因而 A_1,F_2,A_2 共线, $|A_1A_2|=|A_1F_2|+|F_2A_2|$.

总路径长

$$L=|F_1A_1|+|A_1A_2|+|A_2F_1|=|F_1A_1|+(|A_1F_2|+|F_2A_2|)+|A_2F_1|=(|F_1A_1|+|A_1F_2|)+(|F_2A_2|+|A_2F_1|).$$

知乎@余命数

知乎@余命数



知乎@余命数

图 13.27 光线在椭圆内的反射路径

由椭圆定义, 椭圆上任一点到两焦点距离之和为 $2a$, 故

$$|F_1A_1|+|A_1F_2|=2a,\qquad |F_2A_2|+|A_2F_1|=2a,$$

从而 $L=4a$. 代入 $a=5$ 得 $L=20$, 与反射点具体位置无关.

13.3.2.7 椭圆的导演圆

椭圆反射性质说明了焦点与单一切线的联系. 从同一点出发的两条切线具有确定的几何关系, 它由蒙日研究后得到的一个轨迹刻画, 称为**导演圆**(或**蒙日圆**).

定理 13.6 (椭圆的导演圆) 与椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 相切的两条互相垂直的切线交点的轨迹是一个与椭圆同中心的圆, 方程为

$$x^2+y^2=a^2+b^2$$

证明 下面用关于切线斜率的二次方程和韦达定理来证明.

设 $P(x_0,y_0)$ 是轨迹上任意一点. 从 P 可引两条椭圆的切线, 其中一条设为

$$y-y_0=m(x-x_0)\quad 即\quad y=mx+(y_0-mx_0)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

先建立直线 $y=mx+k$ 与椭圆相切的条件. 将直线方程代入椭圆方程并整理成关于 x 的二次方程:

$$\frac{x^2}{a^2}+\frac{(mx+k)^2}{b^2}=1 \implies (b^2+a^2m^2)x^2+2a^2mkx+a^2(k^2-b^2)=0$$

直线与椭圆相切, 即此二次方程有唯一解, 亦即其判别式为零:

$$(2a^2mk)^2-4(b^2+a^2m^2)a^2(k^2-b^2)=0$$

化简得 $a^2m^2k^2-(b^2+a^2m^2)(k^2-b^2)=0$, 最终得到切线条件:

$$k^2=a^2m^2+b^2$$

知乎@余命数

对过点 $P(x_0,y_0)$ 的切线 $y=mx+(y_0-mx_0)$, 截距 $k=y_0-mx_0$, 代入切线条件:

$$(y_0-mx_0)^2=a^2m^2+b^2$$

展开并整理为关于斜率 m 的二次方程:

知乎@余命数

$$y_0^2-2mx_0y_0+m^2x_0^2=a^2m^2+b^2$$

知乎@余命数

$$(x_0^2-a^2)m^2-(2x_0y_0)m+(y_0^2-b^2)=0$$

当 $x_0^2\neq a^2$ 时, 该方程的两个根 m_1,m_2 正是从 P 引出的两条非竖直切线的斜率.

若两条切线互相垂直, 则 $m_1m_2=-1$. 由韦达定理,

$$m_1m_2=\frac{y_0^2-b^2}{x_0^2-a^2}.$$

知乎@余命数

因此

$$\frac{y_0^2-b^2}{x_0^2-a^2}=-1$$

$$y_0^2-b^2=-(x_0^2-a^2)=-x_0^2+a^2$$

整理得到

$$x_0^2+y_0^2=a^2+b^2.$$

反过来, 若 $P(x_0,y_0)$ 满足 $x_0^2+y_0^2=a^2+b^2$ 且 $x_0^2\neq a^2$, 则该二次方程的两根满足

$$m_1m_2=\frac{y_0^2-b^2}{x_0^2-a^2}=-1,$$

知乎@余命数

故对应的两条切线互相垂直. 同时, 该点在椭圆外部, 从一点到椭圆可作两条实切线, 因而这两个斜率给出所需切线.

最后考虑竖直切线的特殊情形. 椭圆的竖直切线为 $x = \pm a$, 水平切线为 $y = \pm b$, 它们互相垂直, 交点为 $(\pm a, \pm b)$, 也满足 $x^2 + y^2 = a^2 + b^2$. 反之, 圆 $x^2 + y^2 = a^2 + b^2$ 上若 $x_0^2 = a^2$, 则必有 $y_0^2 = b^2$, 正是这些特殊交点.

因此所求交点的轨迹恰为以原点为中心、半径 $\sqrt{a^2 + b^2}$ 的圆. □

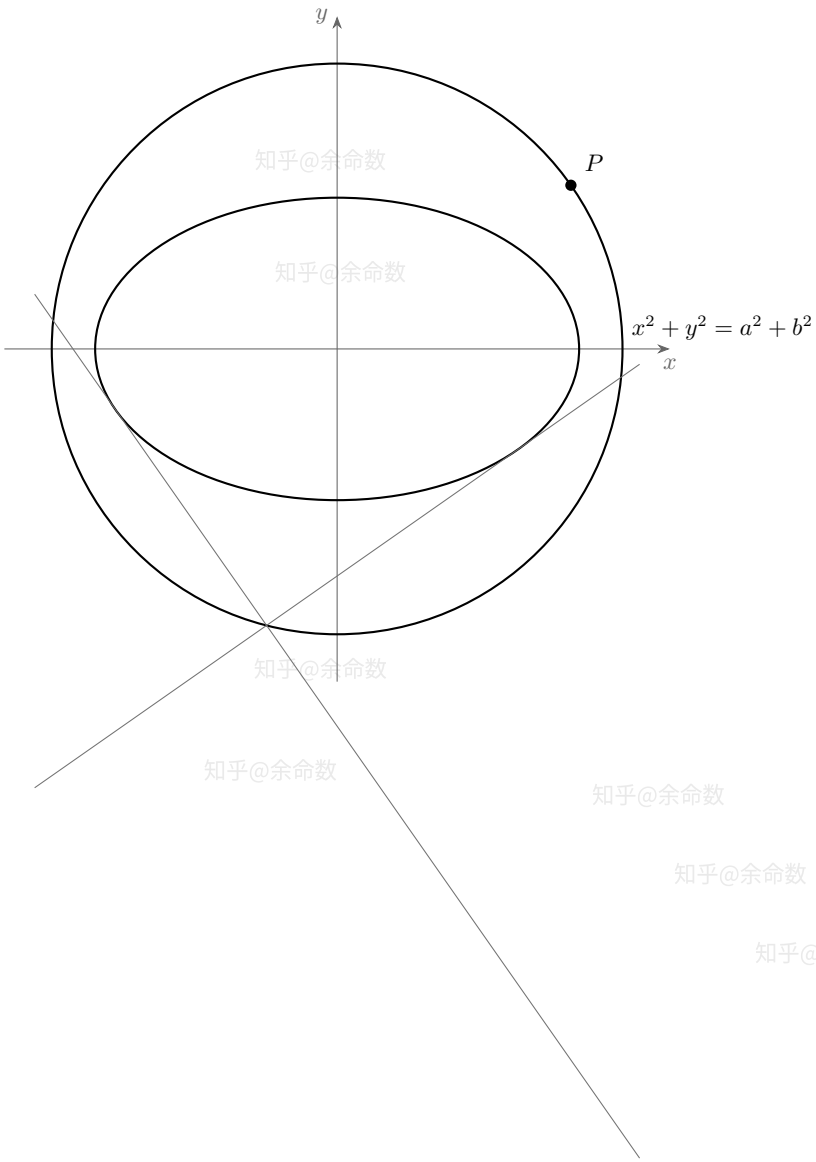


图 13.28 椭圆及其导演圆

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.3.2.8 离心率

椭圆的三个基本几何量为 a, b, c. 椭圆可以接近圆形, 也可以极为扁长, 这提示有一个量在决定它的形状. 我们需要一个与尺寸无关、只描述形状的参数.

椭圆由焦点间距 $2c$ 与距离之和 $2a$ 决定; b 只是由 $b^2 = a^2 - c^2$ 派生. 因此, 椭圆形状蕴含在 a 与 c 的相对关系中. 一个消除尺寸影响的自然度量就是比值 c/a, 这就是离心率.

定义 13.6 (离心率) 椭圆的半焦距 c 与半长轴长 a 的比值称为椭圆的**离心率**, 记作 e.

$$e = \frac{c}{a}$$

离心率字面含义是偏离中心. 这个比值量化了焦点偏离中心的程度, 且相对于椭圆整体尺寸 a.

下面用极限情况说明其几何含义:

- **当 $e \to 0$ 时**: 即 $c \to 0$. 从几何上看, 两个焦点 F_1, F_2 不断靠近, 最终在中心 O 处重合. 此时, 椭圆的定义 $|PF_1| + |PF_2| = 2a$ 就退化为 $|PO| + |PO| = 2a$, 即 $|PO| = a$. 这正是半径为 a 的圆的定义. 在代数上, $c \to 0$ 意味着 $b^2 = a^2 - c^2 \to a^2$, 即 $b \to a$, 椭圆的长短轴相等. 因此, **圆可视为离心率为 0 的特殊椭圆**.
- **当 $e \to 1$ 时**: 即 $c \to a$. 此时 $b^2 = a^2 - c^2 \to 0$, 即 $b \to 0$. 椭圆的短轴长度趋近于零, 整个图形在竖直方向上被压扁, 极限情况下退化为连接两个焦点(此时也与顶点重合)的**线段**. 离心率越接近 1, 椭圆就越扁长.

因此, 离心率 e ($0 \le e < 1$) 度量了椭圆从圆 ($e = 0$) 到线段 ($e \to 1$) 的形态变化.

离心率还有更统一的表述.

定理 13.7 (圆锥曲线的统一焦点-准线定义) 平面内到一个定点 F 的距离与到一条定直线 l 的距离之比为正常数 e 的点的轨迹是一条圆锥曲线, F 称为焦点, l 称为准线(F 不在 l 上), 这个常数 e 就是离心率.

- 当 $0 < e < 1$ 时, 轨迹是**椭圆**.
- 当 $e = 1$ 时, 轨迹是**抛物线**.
- 当 $e > 1$ 时, 轨迹是**双曲线**.

证明 通过平移和旋转坐标轴, 不妨设焦点为 $F(0, 0)$, 准线为 $x = d$ ($d > 0$). 对任意点 $P(x, y)$, 到焦点距离为 $\sqrt{x^2 + y^2}$, 到准线距离为 $|x - d|$. 条件 $|PF| = e\,\mathrm{dist}(P, l)$ 等价于

$$x^2 + y^2 = e^2(x - d)^2.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

展开得

$$(1-e^2)x^2+y^2+2e^2dx-e^2d^2=0. \tag{*}$$

当 $0<e<1$ 时, 配方得

$$\frac{\left(x+\frac{e^2d}{1-e^2}\right)^2}{\frac{e^2d^2}{(1-e^2)^2}}+\frac{y^2}{\frac{e^2d^2}{1-e^2}}=1,$$

这是椭圆的标准形式. 当 $e=1$ 时, 式 (*) 化为

$$y^2=-2d\left(x-\frac{d}{2}\right),$$

这是抛物线的标准形式. 当 $e>1$ 时, 配方得

$$\frac{\left(x-\frac{e^2d}{e^2-1}\right)^2}{\frac{e^2d^2}{(e^2-1)^2}}-\frac{y^2}{\frac{e^2d^2}{e^2-1}}=1,$$

这是双曲线的标准形式. 因此, 焦点-准线距离比为正常数 e 的轨迹按 e 的取值分别为椭圆、抛物线和双曲线. □

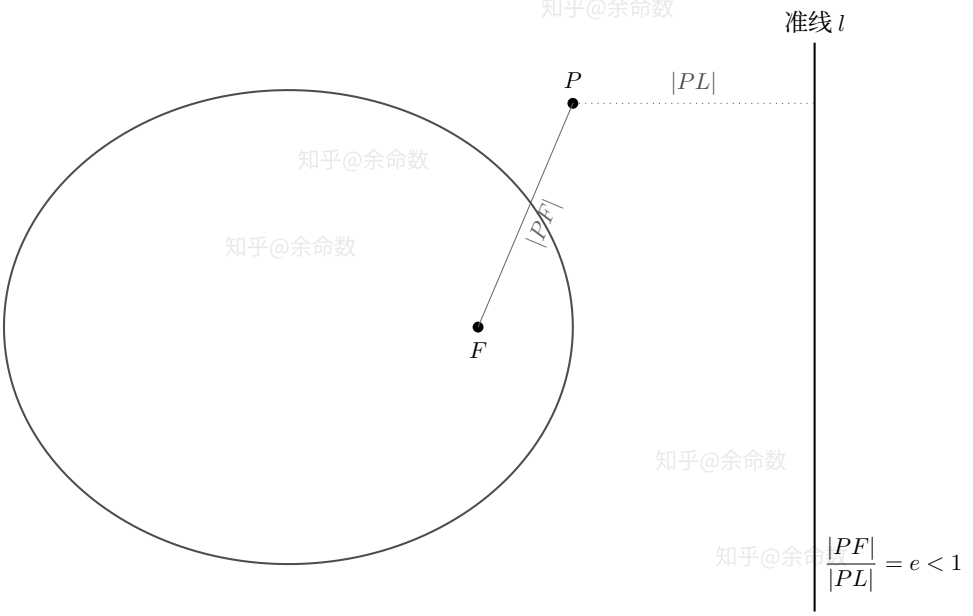


图 13.29 椭圆的焦点-准线定义

可以证明, 对椭圆, 统一定义中的常数 e 与前面定义的 $\frac{c}{a}$ 相同. 准线位置为 $x=\pm\frac{a^2}{c}=\pm\frac{a}{e}$. 统一定义表明, 椭圆、抛物线和双曲线是同一个几何法则 $|PF|=e|PL|$ 在参数 e 取不同值时的表现.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.44　求椭圆 $16x^2+25y^2=400$ 的离心率.

解 先给定方程化为标准形式. 将方程 $16x^2+25y^2=400$ 两边同除以 400, 得到

$$\frac{16x^2}{400}+\frac{25y^2}{400}=1$$

化简得

$$\frac{x^2}{25}+\frac{y^2}{16}=1$$

这是焦点在 x 轴上的椭圆标准方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. 比较系数得 $a^2=25, b^2=16$, 故 $a=5, b=4$.

离心率 $e=\frac{c}{a}$, 其中 c 为半焦距.a, b, c 满足 $a^2=b^2+c^2$. 因此

$$c^2=a^2-b^2=25-16=9 \implies c=3$$

于是离心率为

$$e=\frac{c}{a}=\frac{3}{5}$$

例题 13.45　已知椭圆的两个焦点与短轴的一个顶点构成一个顶角为 120° 的等腰三角形, 求该椭圆的离心率.

解 我们把角度条件转化为关于 a, b, c 的代数关系. 设椭圆的标准方程为 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$. 其焦点为 $F_1(-c,0), F_2(c,0)$, 短轴顶点为 $B(0,b)$.

由题意, $\triangle F_1BF_2$ 是以 B 为顶点的等腰三角形, 顶角 $\angle F_1BF_2=120^\circ$. y 轴, 即直线 OB 是该等腰三角形的对称轴, 故平分顶角.

$$\angle F_2BO=\frac{1}{2}\angle F_1BF_2=\frac{1}{2}\times 120^\circ=60^\circ$$

考察直角三角形 $\triangle F_2OB$, 有

$$\tan(\angle F_2BO)=\frac{|OF_2|}{|OB|}=\frac{c}{b}$$

代入角度值, 得 b, c 的关系:

$$\tan(60^\circ)=\frac{c}{b} \implies \sqrt{3}=\frac{c}{b} \implies c=\sqrt{3}b$$

为求离心率 $e=\frac{c}{a}$, 引入 a. 由 $a^2=b^2+c^2$ 得

$$a^2=b^2+(\sqrt{3}b)^2=b^2+3b^2=4b^2$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由此可得 $a=2b$. 又把 $b=c/\sqrt{3}$ 代入 $a=2b$, 得 $a=2c/\sqrt{3}$. 整理得 c 与 a 的比值:

$$\frac{c}{a}=\frac{\sqrt{3}}{2}$$

因此, 该椭圆的离心率为 $e=\frac{\sqrt{3}}{2}$.

例题 13.46 设 P 为椭圆上任意一点, F_1,F_2 为其两焦点. 若角 $\angle F_1PF_2$ 的最大值为 90°, 求该椭圆的离心率.

知乎@余命数

解 我们需先确定 $\angle F_1PF_2$ 取最大值的条件, 再由此建立方程.

对椭圆上任意点 P, 焦半径 $|PF_1|+|PF_2|=2a$ 固定, 底边 $|F_1F_2|=2c$ 也固定. 设 $r_1=|PF_1|, r_2=|PF_2|$, $\angle F_1PF_2=\theta$. 由余弦定理,

$$(2c)^2=r_1^2+r_2^2-2r_1r_2\cos\theta=(r_1+r_2)^2-2r_1r_2(1+\cos\theta).$$

又 $r_1+r_2=2a$, 故

$$\cos\theta=\frac{2b^2}{r_1r_2}-1.$$

要使 θ 最大, 即使 $\cos\theta$ 最小, 需使 r_1r_2 最大. 在 $r_1+r_2=2a$ 固定时, 由均值不等式 $r_1r_2\le a^2$, 等号当且仅当 $r_1=r_2=a$, 即 P 为短轴顶点.

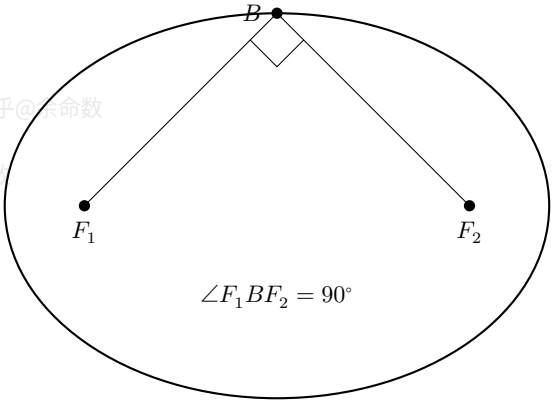


图 13.30 角 $\angle F_1PF_2$ 在短轴顶点处取得最大值

当 P 位于短轴顶点, 例如 $B(0,b)$ 时, $\triangle F_1BF_2$ 是等腰三角形. 由题意, 该角最大值为 90°, 即 $P=B(0,b)$ 时 $\triangle F_1BF_2$ 为等腰直角三角形.

该三角形腰长 $|BF_1|=|BF_2|=a$ (由对称性及 $a^2=b^2+c^2$ 可得). 斜边 $|F_1F_2|=2c$.

由勾股定理,

$$|BF_1|^2+|BF_2|^2=|F_1F_2|^2$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将 a 和 c 代入, 得

$$a^2 + a^2 = (2c)^2$$

$$2a^2 = 4c^2$$

整理得离心率 e:

$$\frac{c^2}{a^2} = \frac{2}{4} = \frac{1}{2}$$

因此, 离心率 $e = \sqrt{\frac{1}{2}} = \frac{\sqrt{2}}{2}$.

例题 13.47　一根长度为 L 的刚性细杆 AB, 其两端点 A 和 B 分别在 x 轴和 y 轴上滑动. 点 P 是杆上的一点, 满足 $|PA| = l_2$ 且 $|PB| = l_1$, 其中 $l_1 + l_2 = L$. 求点 P 的轨迹方程, 并确定该轨迹的离心率.

解　我们用参数方程描述这个轨迹的生成过程.

建立坐标系. 设杆 AB 与 x 轴正向夹角为 θ, 则端点 A 在 x 轴上, B 在 y 轴上.

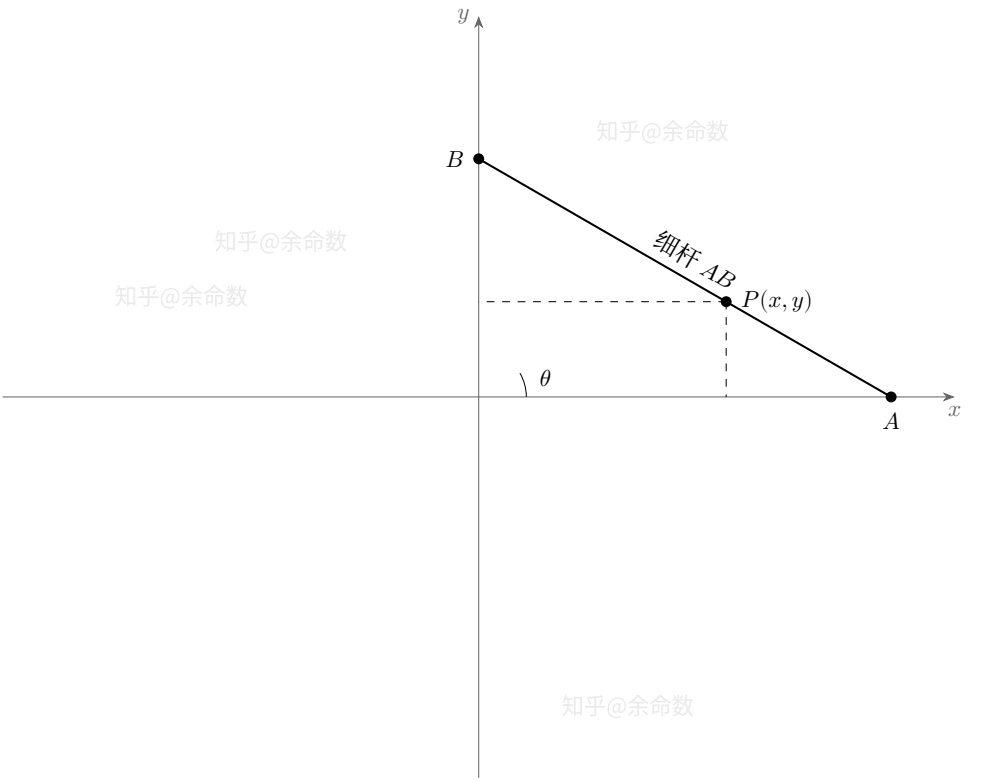


图 13.31　细杆滑动构造椭圆

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

过 P 作 x 轴的垂线,垂足为 M;作 y 轴的垂线,垂足为 N. 在直角三角形 APM 中,$\angle PAM=\theta$,故

$$y=|PM|=|PA|\sin\theta=l_2\sin\theta$$

知乎@余命数在直角三角形 BPN 中,$\angle PBN=\angle OAB=\theta$,故

$$x=|PN|=|PB|\cos\theta=l_1\cos\theta$$

得点 P 轨迹的参数方程:

$$\begin{cases}x=l_1\cos\theta\\ y=l_2\sin\theta\end{cases}\quad(\theta\text{为参数})$$

知乎@余命数

这是椭圆参数方程. 为得到笛卡尔方程,消去 θ.

$$\frac{x}{l_1}=\cos\theta,\quad \frac{y}{l_2}=\sin\theta$$

由 $\cos^2\theta+\sin^2\theta=1$,得

$$\left(\frac{x}{l_1}\right)^2+\left(\frac{y}{l_2}\right)^2=1$$

$$\frac{x^2}{l_1^2}+\frac{y^2}{l_2^2}=1$$

知乎@余命数
知乎@余命数

此即点 P 的轨迹方程,是中心在原点的椭圆.

知乎@余命数

其离心率取决于 l_1 和 l_2 的相对大小,分两种情况.

- **若** $l_1>l_2$:此时半长轴 $a=l_1$,半短轴 $b=l_2$. 半焦距 $c=\sqrt{a^2-b^2}=\sqrt{l_1^2-l_2^2}$. 离心率为 $e=\frac{c}{a}=\frac{\sqrt{l_1^2-l_2^2}}{l_1}$.
- **若** $l_2>l_1$:此时半长轴 $a=l_2$(在 y 轴上),半短轴 $b=l_1$. 半焦距 $c=\sqrt{a^2-b^2}=\sqrt{l_2^2-l_1^2}$. 离心率为 $e=\frac{c}{a}=\frac{\sqrt{l_2^2-l_1^2}}{l_2}$.
- **若** $l_1=l_2$:轨迹退化为圆 $\frac{x^2+y^2}{l_1^2}=1$,可视作离心率 $e=0$ 的特殊椭圆.

因此,离心率取决于点 P 在杆上的相对位置.

例题 13.48 设椭圆 $E:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的长轴辅助圆为 $C:x^2+y^2=a^2$. 过椭圆上满足横坐标非零的上半部点 P 作其切线 l_P. 过 P 点的竖直线与辅助圆 C 交于点 Q(设 Q 在 x 轴上方). 过 Q 作圆 C 的切线 l_Q.

1. 证明:两条切线 l_P 和 l_Q 相交于 x 轴上同一点.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. 若由两条切线 l_P, l_Q 以及 y 轴所围成的三角形面积的**最小值**等于 ab, 求椭圆 E 的离心率 e.

解 (1) 我们用参数方程描述点的位置, 以简化切线方程. 设椭圆上点 P 的离心角为 $t \in (0, \pi)$, 则 $P(a\cos t, b\sin t)$. 当 $\cos t = 0$ 时, 两条切线均平行于 x 轴, 不与 x 轴相交; 下面先讨论 $\cos t \neq 0$ 的一般情形. 由离心角的几何意义, 辅助圆上对应点 Q 的坐标为 $Q(a\cos t, a\sin t)$.

过点 P 的椭圆切线 l_P 方程为

$$\frac{x\cos t}{a} + \frac{y\sin t}{b} = 1$$

过点 Q 的圆 C 的切线 l_Q 方程为(用圆的切线公式 $xx_0 + yy_0 = R^2$):

$$x(a\cos t) + y(a\sin t) = a^2 \implies x\cos t + y\sin t = a$$

为找两切线的交点, 令 $y = 0$, 分别求它们在 x 轴上的截距. 对于 l_P: $\frac{x\cos t}{a} = 1 \implies x = \frac{a}{\cos t}$. 对于 l_Q: $x\cos t = a \implies x = \frac{a}{\cos t}$. 由于两条切线在 x 轴上的截距完全相同, 它们必然相交于 x 轴上的同一点, 记为 $T(\frac{a}{\cos t}, 0)$. 若 $\cos t = 0$, 则两条切线分别为 $y = b$ 与 $y = a$, 它们互相平行, 不构成有限三角形; 因此第 (2) 问的面积最值只需在 $\cos t \neq 0$ 时讨论.

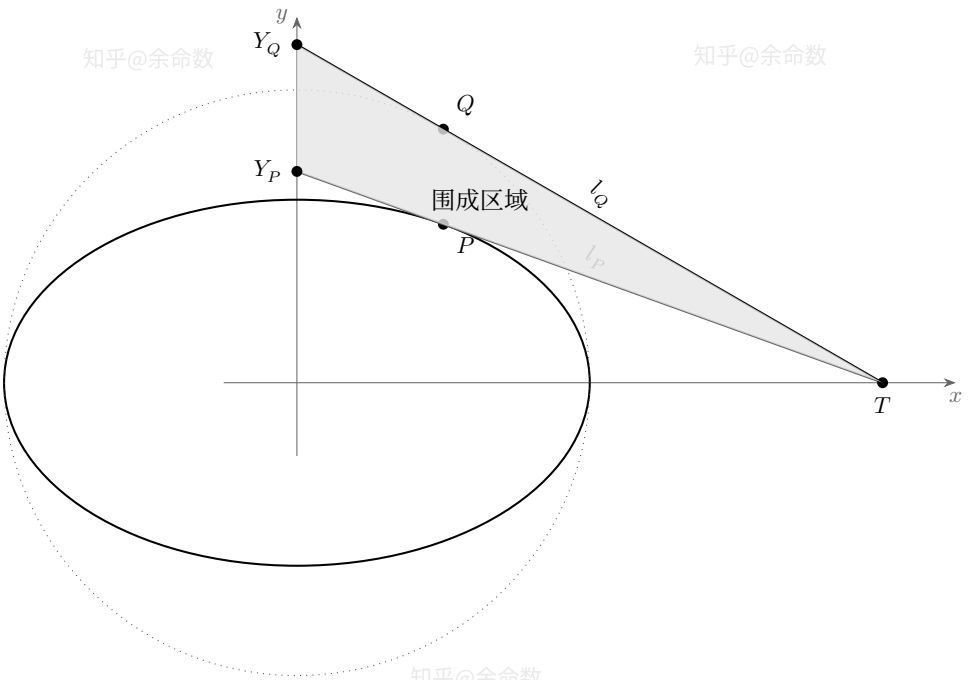


图 13.32　两切线与 y 轴围成的三角形

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2) 所述三角形由两条切线与 y 轴围成, 其三个顶点为两条切线在 y 轴上的截距点及 x 轴上的公共交点 T. 先求两条切线在 y 轴上的截距. 对于 l_P, 令 $x=0$, 得 $\frac{y\sin t}{b}=1 \implies y=\frac{b}{\sin t}$. 对于 l_Q, 令 $x=0$, 得 $y\sin t=a \implies y=\frac{a}{\sin t}$.

该区域是以 y 轴上一段线段为底、顶点为 T 的三角形. 其底边长为 $\left|\frac{a}{\sin t}-\frac{b}{\sin t}\right|=\frac{a-b}{\sin t}$ (由 $t\in(0,\pi)$ 知 $\sin t>0$, 且 $a>b$). 其高为点 T 到 y 轴的距离, 即 T 点横坐标的绝对值 $|\frac{a}{\cos t}|$.

三角形面积 $S(t)$ 为

$$\begin{aligned} S(t) &= \frac{1}{2}\cdot \text{底}\cdot \text{高} \\ &= \frac{1}{2}\cdot\frac{a-b}{\sin t}\cdot\left|\frac{a}{\cos t}\right| \\ &= \frac{a(a-b)}{|2\sin t\cos t|} \\ &= \frac{a(a-b)}{|\sin(2t)|}. \end{aligned}$$

下面求 $S(t)$ 的最小值. 分母 $|\sin(2t)|$ 的最大值为 1, 在 $2t=\pi/2$ 或 $3\pi/2$ 即 $t=\pi/4$ 或 $3\pi/4$ 时取得, 此时 $S(t)$ 取最小值.

$$S_{\min}=\frac{a(a-b)}{1}=a(a-b)$$

由题设, 该最小值等于 ab.

$$a(a-b)=ab$$

由于 $a\neq 0$, 两边同除以 a, 得

$$a-b=b \implies a=2b$$

这个关系确定了椭圆的形状. 接着计算离心率:

$$e^2=1-\frac{b^2}{a^2}=1-\frac{b^2}{(2b)^2}=1-\frac{1}{4}=\frac{3}{4}$$

故椭圆 E 的离心率 $e=\frac{\sqrt{3}}{2}$.

例题 13.49 以正方形 $ABCD$ 的相对顶点 A,C 为焦点的椭圆, 恰过正方形四边的中点, 则椭圆的离心率为(　).

解 我们用正方形的几何对称性与椭圆焦点定义来建立坐标系.

以正方形中心为原点, 焦点所在对角线 AC 为 x 轴建立平面直角坐标系. 设椭圆的半焦距为 c, 则焦点坐标为 $F_1(-c,0)$ 和 $F_2(c,0)$. 这两个焦点也即是正方形

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

的顶点 A 和 C. AC 是正方形的对角线, 长度为 $2c$, 中心在原点, 故另两顶点 B, D 的坐标为 $(0, c), (0, -c)$.

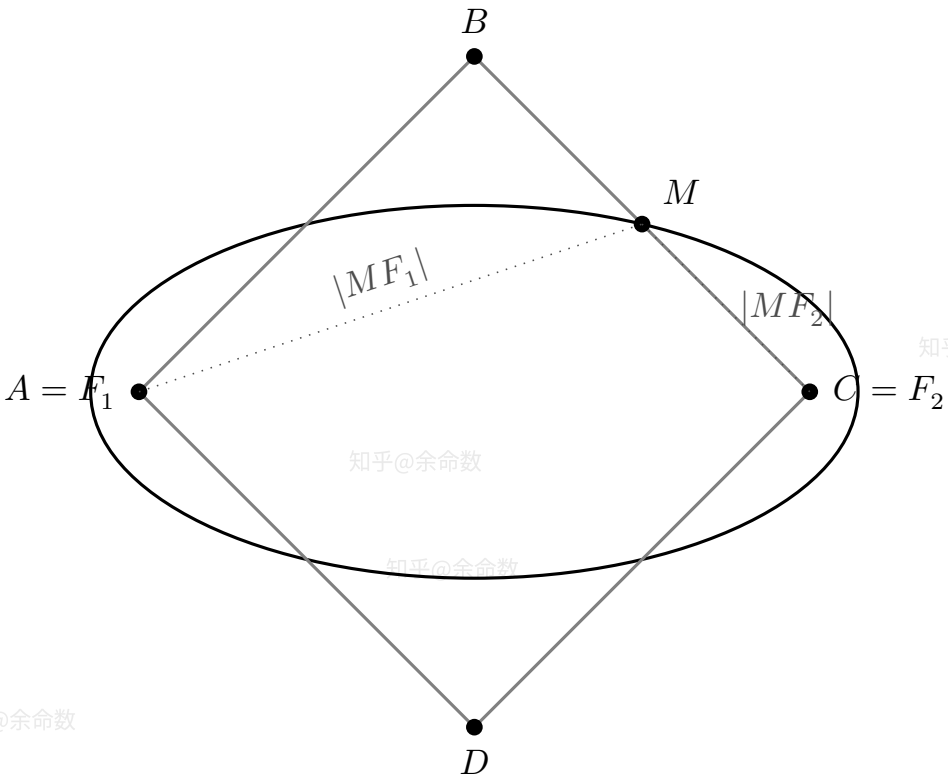


图 13.33　椭圆与内切正方形

椭圆经过正方形四边的中点. 取边 BC 的中点 M 来分析. 其坐标为 $M\left(\frac{c+0}{2}, \frac{0+c}{2}\right) = \left(\frac{c}{2}, \frac{c}{2}\right)$.

由椭圆定义, 点 M 到两焦点的距离之和等于 $2a$:

$$|MF_1| + |MF_2| = 2a$$

分别计算这两个焦半径, 得

$$\begin{aligned} |MF_1| &= \sqrt{\left(\frac{c}{2} - (-c)\right)^2 + \left(\frac{c}{2} - 0\right)^2} \\ &= \sqrt{\left(\frac{3c}{2}\right)^2 + \left(\frac{c}{2}\right)^2} \\ &= \sqrt{\frac{9c^2}{4} + \frac{c^2}{4}} \\ &= \frac{\sqrt{10}c}{2}. \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\begin{aligned}|MF_2| &= \sqrt{\left(\frac{c}{2}-c\right)^2+\left(\frac{c}{2}-0\right)^2}\\ &= \sqrt{\left(-\frac{c}{2}\right)^2+\left(\frac{c}{2}\right)^2}\\ &= \sqrt{\frac{c^2}{4}+\frac{c^2}{4}}\\ &= \frac{\sqrt{2}c}{2}.\end{aligned}$$

将这些距离代入定义式:

$$\begin{aligned}2a &= \frac{\sqrt{10}c}{2}+\frac{\sqrt{2}c}{2}\\ &= \frac{c(\sqrt{10}+\sqrt{2})}{2}.\end{aligned}$$

由此得离心率 e:

$$e=\frac{c}{a}=\frac{c}{c(\sqrt{10}+\sqrt{2})/4}=\frac{4}{\sqrt{10}+\sqrt{2}}.$$

为化简,对分母有理化:

$$\begin{aligned}e &= \frac{4(\sqrt{10}-\sqrt{2})}{(\sqrt{10}+\sqrt{2})(\sqrt{10}-\sqrt{2})}\\ &= \frac{4(\sqrt{10}-\sqrt{2})}{10-2}\\ &= \frac{4(\sqrt{10}-\sqrt{2})}{8}\\ &= \frac{\sqrt{10}-\sqrt{2}}{2}.\end{aligned}$$

13.3.3 双曲线

13.3.3.1 双曲线的几何定义

定义 13.7 (双曲线) 平面内到两个定点 F_1, F_2 的距离之差的绝对值等于一个常数(该常数必须小于 $|F_1F_2|$ 且为正)的所有点 P 的轨迹,称为**双曲线**. 定点 F_1, F_2 称为双曲线的**焦点**.

记 $|F_1F_2|=2c\,(c>0)$,距离差的绝对值为 $2a\,(a>0)$,且 $2a<2c$ 即 $a<c$. 令 $b^2=c^2-a^2\,(b>0)$.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

13.3.3.2 双曲线的标准方程

取焦点 $F_1(-c,0)$, $F_2(c,0)$, 中心为原点. 由定义 $||PF_1|-|PF_2||=2a$, 经两次平方消根(与椭圆推导平行, 符号在差上), 得

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{13-17}$$

此为**焦点在 x 轴**的标准方程; 焦点在 y 轴时为 $\dfrac{y^2}{a^2}-\dfrac{x^2}{b^2}=1$.

- **顶点**: $(\pm a,0)$; 实轴长 $2a$, 虚轴长 $2b$.
- **渐近线**: $y=\pm\dfrac{b}{a}x$(由 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=0$ 因式分解).
- **离心率**: $e=\dfrac{c}{a}>1$.
- **准线**: $x=\pm\dfrac{a}{e}=\pm\dfrac{a^2}{c}$; 统一定义 $|PF|=e\cdot$(到准线距离) 对双曲线同样成立.

13.3.3.3 双曲线的参数与切线

右支常用参数: $x=a\sec\theta$, $y=b\tan\theta$; 或双曲函数 $x=a\cosh t$, $y=b\sinh t$. 过点 $P_0(x_0,y_0)$ 在双曲线 (13-17) 上的切线, 由替换法则:

$$\frac{xx_0}{a^2}-\frac{yy_0}{b^2}=1.$$

(与椭圆切线仅差一个符号.)

例题 13.50 (双曲线基本量) 双曲线 $\dfrac{x^2}{9}-\dfrac{y^2}{16}=1$. 求 a,b,c,e 与渐近线方程.

解 $a=3$、$b=4$、$c=5$、$e=\dfrac{5}{3}$; 渐近线 $y=\pm\dfrac{4}{3}x$.

例题 13.51 (双曲线弦的中点) 直线 $y=x+m$ 与双曲线 $\dfrac{x^2}{4}-\dfrac{y^2}{2}=1$ 交于两点, 若弦中点横坐标为 3, 求 m.

解 联立 $\dfrac{x^2}{4}-\dfrac{(x+m)^2}{2}=1$, 整理得

$$x^2+4mx+2m^2+4=0.$$

由韦达 $x_1+x_2=-4m$, 故 $x_M=-2m$. 令 $x_M=3$ 得 $m=-\dfrac{3}{2}$. 此时 $\Delta=16m^2-4(2m^2+4)=8m^2-16=8\cdot\dfrac{9}{4}-16=2>0$, 实交点存在.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

注 13.2 (与椭圆的对照)
椭圆 $a^2=b^2+c^2, e<1$; 双曲线 $c^2=a^2+b^2, e>1$. 点差法中椭圆为 $\frac{x_M}{a^2}+k\frac{y_M}{b^2}=0$, 双曲线为 $\frac{x_M}{a^2}-k\frac{y_M}{b^2}=0$. 设点/设线的操作框架对二者通用.

13.3.4 抛物线

13.3.4.1 抛物线的几何定义

定义 13.8 (抛物线) 平面内到一个定点 F 的距离等于到一条定直线 l 的距离的所有点 P 的轨迹, 称为**抛物线**, F 称为焦点, l 称为准线(F 不在 l 上).

这一定义恰是圆锥曲线统一定义中 $e=1$ 的情形.

知乎@余命数

13.3.4.2 抛物线的标准方程

取焦点 $F(\frac{p}{2},0)$ $(p>0)$, 准线 $x=-\frac{p}{2}$, 顶点在原点, 得

$$y^2=2px \quad (\text{亦常写 } y^2=4ax, \text{此时焦参数 } 2a=p). \tag{13-18}$$

开口向上为 $x^2=2py$; 开口向左/下则 p 取负或改写坐标.

- **焦点**: $(\frac{p}{2},0)$; **准线**: $x=-\frac{p}{2}$.
- **通径**: 过焦点且垂直于轴的弦, 长为 $2p$.
- **焦半径**: 对 $P(x_0,y_0)$ 在抛物线上, $|PF|=x_0+\frac{p}{2}$.
- **离心率**: $e=1$.

知乎@余命数

13.3.4.3 参数方程与切线

常用参数: $y^2=2px$ 上点 $(\frac{p}{2}t^2,\,pt)$; 或 $y^2=4ax$ 上 $(at^2,2at)$. 过 $P_0(x_0,y_0)$ 的切线: $yy_0=p(x+x_0)$(对 $y^2=2px$).

例题 13.52 (抛物线焦点弦) 抛物线 $y^2=4x$, 即 $p=2$ 过焦点的弦 AB 斜率为 k $(k\neq 0)$. 求弦长 $|AB|$ 关于 k 的表达式.

解 焦点 $F(1,0)$, 直线 $y=k(x-1)$. 联立 $y^2=4x$:

$$k^2(x-1)^2=4x.$$

展开 $(k^2)x^2-(2k^2+4)x+k^2=0$. 设根 x_1,x_2,

$$x_1+x_2=\frac{2k^2+4}{k^2},\quad x_1x_2=1.$$

弦长

$$|AB| = \sqrt{1+k^2}\,\frac{\sqrt{\Delta}}{|k^2|},$$

其中 $\Delta = (2k^2+4)^2 - 4k^2\cdot k^2 = 16k^2+16 = 16(k^2+1)$, 故

$$|AB| = \sqrt{1+k^2}\cdot\frac{4\sqrt{k^2+1}}{k^2} = \frac{4(1+k^2)}{k^2}.$$

例题 13.53 (抛物线平行弦中点) 求抛物线 $y^2 = 2px$ 上一族斜率为 k ($k \neq 0$) 的平行弦的中点轨迹.

知乎@余命数

解 由点差法, 中点满足 $y_M = \dfrac{p}{k}$ 且 $x_M > \dfrac{p}{2k^2}$, 即水平射线. 椭圆与双曲线的平行弦中点落在共轭直径方向上; 抛物线无中心, 其平行弦中点轨迹是平行于轴的射线.

注 13.3 (三类曲线对照)

	椭圆	抛物线	双曲线
e	$0 \le e < 1$	$e = 1$	$e > 1$
定义	$\lvert PF_1\rvert + \lvert PF_2\rvert = 2a$	$\lvert PF\rvert = \operatorname{dist}(P, l)$	$\big\lvert\lvert PF_1\rvert - \lvert PF_2\rvert\big\rvert = 2a$
标准形	$\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$	$y^2 = 2px$	$\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$

统一定义 $|PF| = e\cdot$(到准线距离) 把三者串在同一参数 e 上(椭圆段已述).

13.3.5 圆锥曲线的代数方法

13.3.5.1 从直线与圆锥曲线的位置关系谈起

前面几节已经建立了椭圆、双曲线和抛物线的标准方程与基本性质. 下面用代数工具研究直线与圆锥曲线的关系, 把相交、相切、弦长、中点等几何概念翻译成代数语言.

研究直线与圆锥曲线的位置关系, 第一步是求解由它们方程组成的方程组.

设直线 l 的方程为 $Ax+By+C=0$, 圆锥曲线 $\mathcal{C}$ 的方程为 $f(x,y)=0$. 联立方程组.

$$\begin{cases} Ax+By+C=0 \\ f(x,y)=0 \end{cases}$$

通过消元, 例如把 $y = kx+m$ 代入 $f(x,y)=0$, 通常得到一个关于 x 或 y 的一元二次方程.

$$ax^2+bx+c=0 \quad (a\neq 0)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个一元二次方程联系了几何与代数, 其性质直接反映直线与曲线的几何关系.

13.3.5.1.1 判别式与交点个数 一元二次方程实数根的个数决定直线与圆锥曲线的交点个数. 判别式 $\Delta = b^2 - 4ac$ 的符号直接给出位置关系.

- $\Delta > 0 \iff$ 方程有两个不等实根 $\iff$ 直线与曲线**相交**于两点.
- $\Delta = 0 \iff$ 方程有两个相等实根 $\iff$ 直线与曲线**相切**于一点.
- $\Delta < 0 \iff$ 方程无实根 $\iff$ 直线与曲线**相离**, 无公共点.

当直线平行于双曲线的渐近线或抛物线的对称轴时, 联立消元后得到一个一次方程, 此时直线与曲线只有一个交点, 属相交于一点的特殊情形. 因此 $\Delta = 0$ 是相切的充分条件, 但运用时要保证二次项系数不为零.

这个结论展示了解析几何的要旨, 把几何概念转化为可计算的代数结构. 几何图形的相交与代数方程判别式之间的一一对应, 依赖如下逻辑链条:

在笛卡尔坐标系中, 任何几何图形, 如直线、椭圆都是满足某一代数方程的点的集合. 一个**交点**同时位于直线上和圆锥曲线上, 因此其坐标 (x, y) 同时满足直线方程和圆锥曲线方程.

在代数中, 寻找同时满足多个方程的解称为**解方程组**. 因此, 寻找几何交点的问题转化为求解代数方程组的问题; 方程组实数解的组数等于几何图形的交点个数.

为解这个方程组, 通常用**消元法**, 从线性方程表出 $y = kx + m$ 或 $x = my + n$, 代入二次曲线方程. 这个代入过程把二维平面上寻找点 (x, y) 的问题降维为一元二次方程 $ax^2 + bx + c = 0$ 的问题, 即寻找满足条件的横坐标 x.

一元二次方程的每个实数根对应一个交点的横坐标. 若方程有两个不等实根 x_1, x_2, 则对应两个不同交点; 若有两个相等实根 x_0, 则对应一个交点; 若没有实根, 则没有交点. 因此, 一元二次方程实数根的个数与几何图形的交点个数一一对应.

于是问题归结为判断一元二次方程 $ax^2 + bx + c = 0$ 实数根的个数, 这正是**判别式** $\Delta = b^2 - 4ac$ 的作用. 由求根公式 $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ 知, Δ 的符号决定 $\sqrt{\Delta}$ 是实数还是虚数, 从而决定实数根的个数.

以上建立了从几何到代数的逻辑通路.

$$\text{几何交点个数} \xrightarrow{\text{坐标化}} \text{方程组实数解的组数}$$

$$\xrightarrow{\text{消元}} \text{一元二次方程实数根的个数}$$

$$\xrightarrow{\text{判别式}} \Delta \text{ 的符号}$$

13.3.5.1.2 韦达定理 当 $\Delta \geq 0$ 时, 交点存在. 自然想到直接解出 x_1, x_2 再进行后续计算, 但这条路通常非常繁琐.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

回顾求根公式:

$$x_{1,2}=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

除非 Δ 为完全平方数, 否则 x_1,x_2 一般是含无理数的复杂表达式, 直接代入后续计算会很繁重.

但这些计算中蕴含规律. 例如计算弦中点横坐标 $\frac{x_1+x_2}{2}$ 时:

$$\frac{x_1+x_2}{2}=\frac{1}{2}\left(\frac{-b+\sqrt{\Delta}}{2a}+\frac{-b-\sqrt{\Delta}}{2a}\right)=\frac{1}{2}\left(\frac{-2b}{2a}\right)=-\frac{b}{2a}$$

知乎@余命数

注意到最棘手的无理部分 $\sqrt{\Delta}$ 在求和过程中抵消了, 结果 $-\frac{b}{2a}$ 只与方程系数 a,b 有关. 这提示我们, 可能不必解出根的具体值, 只需根的整体性质, 如和或积.

韦达定理给出了这个工具,一元二次方程两根的和与积可直接由系数确定,无需开方.

定理 13.8 (韦达定理) 若一元二次方程 $ax^2+bx+c=0$ $(a\neq 0)$ 的两根为 x_1,x_2, 则它们满足以下关系.

$$x_1+x_2=-\frac{b}{a},\quad x_1x_2=\frac{c}{a}$$

知乎@余命数

证明 因为 x_1,x_2 是方程 $ax^2+bx+c=0$ 的两根, 所以二次多项式可分解为

$$ax^2+bx+c=a(x-x_1)(x-x_2).$$

知乎@余命数

展开右式得

$$a(x-x_1)(x-x_2)=a(x^2-(x_1+x_2)x+x_1x_2).$$

比较 x 的系数与常数项, 得

$$b=-a(x_1+x_2),\qquad c=ax_1x_2.$$

由于 $a\neq 0$, 两式同除以 a, 即得

$$x_1+x_2=-\frac{b}{a},\qquad x_1x_2=\frac{c}{a}.$$

知乎@余命数

□

韦达定理催生了解析几何中“设而不求”的思想,即设交点 $P_1(x_1,y_1),P_2(x_2,y_2)$, 但不求出它们, 而把所有需要的几何量用对称式 x_1+x_2,x_1x_2 表示.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在解析几何中, 一元二次方程的系数 a,b,c 由直线参数 k 与曲线参数 a^2,b^2 组合而成, 韦达定理串起如下逻辑链:

$$\text{几何要素参数} \xrightarrow{\text{联立}} \text{二次方程系数} \xrightarrow{\text{韦达定理}} \text{交点坐标的整体性质}$$

知乎@余命数

13.3.5.2 基本几何量的计算

当直线与圆锥曲线交于两点 P_1,P_2 时, 线段 P_1P_2 称为弦. 其长度 $|P_1P_2|$ 可用韦达定理计算.

设直线方程为 $y=kx+m$. 由两点间距离公式,

$$|P_1P_2|=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}=\sqrt{1+k^2}|x_1-x_2|$$

而 $|x_1-x_2|$ 可由韦达定理表示, 不必求解根:

$$(x_1-x_2)^2=(x_1+x_2)^2-4x_1x_2=\left(-\frac{b}{a}\right)^2-4\frac{c}{a}=\frac{b^2-4ac}{a^2}=\frac{\Delta}{a^2}$$

因此, $|x_1-x_2|=\frac{\sqrt{\Delta}}{|a|}$.

知乎@余命数

定理 13.9 (弦长公式) 若直线 $y=kx+m$ 与圆锥曲线交于两点, 且联立消元后得到的关于 x 的一元二次方程为 $ax^2+bx+c=0$ $(a\neq 0,\Delta>0)$, 则弦长为

$$|P_1P_2|=\sqrt{1+k^2}\frac{\sqrt{\Delta}}{|a|} \tag{13-19}$$

若直线方程为 $x=my+n$, 联立消元后得到关于 y 的方程为 $ay^2+by+c=0$ $(a\neq 0,\Delta>0)$, 则弦长为

$$|P_1P_2|=\sqrt{1+m^2}\frac{\sqrt{\Delta}}{|a|} \tag{13-20}$$

证明 先证 $y=kx+m$ 的情形. 设两个交点为 $P_1(x_1,y_1),P_2(x_2,y_2)$. 因为两点都在直线 $y=kx+m$ 上, 所以

$$y_1-y_2=k(x_1-x_2).$$

由两点间距离公式,

$$|P_1P_2|=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}=\sqrt{1+k^2}\,|x_1-x_2|.$$

知乎@余命数
知乎@余命数
知乎@余命数

另一方面, x_1, x_2 是方程 $ax^2+bx+c=0$ 的两个实根. 由韦达定理,

$$(x_1-x_2)^2=(x_1+x_2)^2-4x_1x_2=\left(-\frac{b}{a}\right)^2-4\frac{c}{a}=\frac{b^2-4ac}{a^2}=\frac{\Delta}{a^2}.$$

因 $\Delta>0$, 得 $|x_1-x_2|=\frac{\sqrt{\Delta}}{|a|}$, 从而

$$|P_1P_2|=\sqrt{1+k^2}\frac{\sqrt{\Delta}}{|a|}.$$

对直线 $x=my+n$ 的情形, 改以 y_1, y_2 为两个实根并重复同样推导, 即得

$$|P_1P_2|=\sqrt{1+m^2}\frac{\sqrt{\Delta}}{|a|}.$$

□

此公式把弦长用联立方程的代数系数表达, 全程不必解出交点坐标.

例题 13.54　求直线 $y=x-1$ 被抛物线 $y^2=4x$ 截得的弦长.

解　直线 $y=x-1$ 代入抛物线方程 $y^2=4x$, 消去 y 后整理为关于 x 的一元二次方程

$$(x-1)^2=4x,$$

$$x^2-6x+1=0,$$

其系数为 $a=1, b=-6, c=1$, 直线斜率 $k=1$. 判别式 $\Delta=b^2-4ac=(-6)^2-4=32>0$ 表明直线与抛物线有两个不同的交点. 代入弦长公式得

$$|P_1P_2|=\sqrt{1+k^2}\frac{\sqrt{\Delta}}{|a|}=\sqrt{2}\cdot\frac{\sqrt{32}}{1}=\sqrt{2}\cdot 4\sqrt{2}=8.$$

例题 13.55　求椭圆 $\frac{x^2}{4}+y^2=1$ 的过右焦点 F 且斜率为 k 的弦的长度.

解　对于椭圆 $\frac{x^2}{4}+y^2=1$, 有 $a^2=4, b^2=1$, 故半焦距 $c=\sqrt{a^2-b^2}=\sqrt{3}$, 右焦点为 $F(\sqrt{3},0)$. 过 F 且斜率为 k 的直线 $y=k(x-\sqrt{3})$ 代入椭圆方程, 整理得

$$\frac{x^2}{4}+k^2(x-\sqrt{3})^2=1,$$

$$(1+4k^2)x^2-8\sqrt{3}k^2x+(12k^2-4)=0.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

其判别式

$$\begin{aligned}\Delta &= (-8\sqrt{3}k^2)^2 - 4(1+4k^2)(12k^2-4)\\ &= 192k^4 - 16(12k^4 - k^2 - 1) = 16(k^2+1),\end{aligned}$$

代入弦长公式得

$$|P_1P_2| = \sqrt{1+k^2}\frac{\sqrt{16(k^2+1)}}{1+4k^2} = \frac{4(1+k^2)}{1+4k^2}.$$

知乎@余命数

13.3.5.3 弦中点问题

知乎@余命数

13.3.5.3.1 韦达定理 设弦 P_1P_2 的中点为 $M(x_M, y_M)$, 联立直线与曲线方程得 $ax^2+bx+c=0$. 由中点坐标公式和韦达定理, 中点横坐标为

$$x_M = \frac{x_1+x_2}{2} = -\frac{b}{2a}$$

知乎@余命数

中点纵坐标 y_M 代入直线方程 $y_M = kx_M + m$ 即得. 这是处理中点问题的常用做法.

知乎@余命数

13.3.5.3.2 点差法 韦达定理法普适, 但在某些情况下运算较繁. 点差法为弦中点与弦斜率问题提供了一条更简洁的代数路径. 设弦端点为 $P_1(x_1,y_1), P_2(x_2,y_2)$, 中点为 $M(x_M, y_M)$. 以椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 为例. P_1, P_2 都在椭圆上, 故其坐标满足方程.

知乎@余命数 知乎@余命数

$$\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}=1 \quad ① \qquad 和 \qquad \frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}=1 \quad ②$$

知乎@余命数 知乎@余命数

两式相减(此即点差), 并用平方差公式:

$$\frac{(x_1-x_2)(x_1+x_2)}{a^2}+\frac{(y_1-y_2)(y_1+y_2)}{b^2}=0$$

代入中点坐标 $x_M=\frac{x_1+x_2}{2}, y_M=\frac{y_1+y_2}{2}$ 与斜率 $k=\frac{y_1-y_2}{x_1-x_2}$:

$$\frac{(x_1-x_2)(2x_M)}{a^2}+\frac{k(x_1-x_2)(2y_M)}{b^2}=0$$

当 $x_1\neq x_2$ 时, 约去 $2(x_1-x_2)$, 整理得

$$\frac{x_M}{a^2}+k\frac{y_M}{b^2}=0$$

此式建立了弦中点 (x_M, y_M) 与弦斜率 k 之间的关系, 不必涉及直线截距.

知乎@余命数 知乎@余命数

定理 13.10 (弦中点斜率关系) 对于中心在原点的椭圆与双曲线, 其一条非竖直弦的斜率为 k, 中点坐标为 (x_M, y_M), 则满足

- 椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$: $\frac{x_M}{a^2}+k\frac{y_M}{b^2}=0$. 当 $x_M\neq 0$ 时, 可写为 $k\cdot\frac{y_M}{x_M}=-\frac{b^2}{a^2}$, 即 $k\cdot k_{OM}=-\frac{b^2}{a^2}$.
- 双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$: $\frac{x_M}{a^2}-k\frac{y_M}{b^2}=0$. 当 $x_M\neq 0$ 时, 可写为 $k\cdot\frac{y_M}{x_M}=\frac{b^2}{a^2}$, 即 $k\cdot k_{OM}=\frac{b^2}{a^2}$.

证明 设弦的两个端点为 $P_1(x_1,y_1), P_2(x_2,y_2)$. 因弦非竖直, $x_1\neq x_2$, 且

$$k=\frac{y_1-y_2}{x_1-x_2},\qquad x_M=\frac{x_1+x_2}{2},\qquad y_M=\frac{y_1+y_2}{2}.$$

对椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, 将 P_1, P_2 的方程相减, 得

$$\frac{x_1^2-x_2^2}{a^2}+\frac{y_1^2-y_2^2}{b^2}=0.$$

用平方差公式并代入中点和斜率, 有

$$\frac{(x_1-x_2)(2x_M)}{a^2}+\frac{(y_1-y_2)(2y_M)}{b^2}=0\quad\Longrightarrow\quad\frac{x_M}{a^2}+k\frac{y_M}{b^2}=0.$$

若 $x_M\neq 0$, 两边同乘 $\frac{b^2}{x_M}$, 得 $k\frac{y_M}{x_M}=-\frac{b^2}{a^2}$.

对双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, 同样将两端点的方程相减, 得

$$\frac{x_1^2-x_2^2}{a^2}-\frac{y_1^2-y_2^2}{b^2}=0.$$

代入中点和斜率后得到

$$\frac{x_M}{a^2}-k\frac{y_M}{b^2}=0.$$

若 $x_M\neq 0$, 即得 $k\frac{y_M}{x_M}=\frac{b^2}{a^2}$. □

例题 13.56 已知直线 $y=x+m$ 与双曲线 $\frac{x^2}{4}-\frac{y^2}{2}=1$ 交于两点, 若两交点所构成弦的中点的横坐标为 3, 求实数 m 的值.

解 将 $y=x+m$ 代入双曲线方程, 整理得

$$\frac{x^2}{4}-\frac{(x+m)^2}{2}=1,$$

$$x^2+4mx+2m^2+4=0.$$

知乎@余命数

设两交点为 $P_1(x_1,y_1),P_2(x_2,y_2)$，弦中点为 $M(x_M,y_M)$. 由韦达定理 $x_1+x_2=-4m$，故中点横坐标 $x_M=\dfrac{x_1+x_2}{2}=-2m$. 题设 $x_M=3$，即 $-2m=3$，解得 $m=-\dfrac{3}{2}$.

例题 13.57 求椭圆 $\frac{x^2}{9}+\frac{y^2}{4}=1$ 中，被点 $M(1,1)$ 平分的弦所在的直线方程.

解 本题直接应用点差法得到的弦中点斜率关系. 对于椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$，弦的斜率 k 与弦中点 $M(x_M,y_M)$ 满足 $k\cdot\dfrac{y_M}{x_M}=-\dfrac{b^2}{a^2}$. 此处 $a^2=9,b^2=4$，中点 $M(1,1)$，代入得 $k=-\dfrac{4}{9}$. 所求直线过 $M(1,1)$ 且斜率为 $-\dfrac{4}{9}$，由点斜式 $y-1=-\dfrac{4}{9}(x-1)$，整理得 $4x+9y-13=0$.

例题 13.58 求抛物线 $y^2=2px\,(p>0)$ 的一组斜率为 $k\,(k\neq0)$ 的平行弦的中点轨迹方程.

解 设一条斜率为 k 的弦的端点为 $P_1(x_1,y_1),P_2(x_2,y_2)$. 两点在抛物线上，故 $y_1^2=2px_1,y_2^2=2px_2$. 将两式相减并利用平方差公式，得

$$(y_1-y_2)(y_1+y_2)=2p(x_1-x_2).$$

弦的斜率存在且不为零，故 $x_1\neq x_2$，上式可写为 $\dfrac{y_1-y_2}{x_1-x_2}(y_1+y_2)=2p$. 设中点为 $M(x_M,y_M)$，则 $k=\dfrac{y_1-y_2}{x_1-x_2},y_1+y_2=2y_M$，代入得 $k\cdot2y_M=2p$，即 $y_M=\dfrac{p}{k}$. 此关系对所有斜率为 k 的弦的中点成立，故中点轨迹位于平行于 x 轴的直线 $y=\dfrac{p}{k}$ 上. 还需保证过中点且斜率为 k 的直线能截得实弦，设中点为 $(x_M,\dfrac{p}{k})$，弦端点可写作 $(x_M\pm u,\dfrac{p}{k}\pm ku)$，代入抛物线方程得 $k^2u^2=2p(x_M-\dfrac{p}{2k^2})$，故须 $x_M>\dfrac{p}{2k^2}$. 因而完整轨迹为射线

$$y=\frac{p}{k},\quad x>\frac{p}{2k^2}.$$

13.3.5.3.3 点差法的原理与思想 点差法是处理圆锥曲线弦中点问题的一种方法，把两个满足曲线方程的点的坐标代入方程，再相减.

以椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 为例，设弦 P_1P_2 的端点为 $P_1(x_1,y_1),P_2(x_2,y_2)$. 由于这两点都在椭圆上，坐标满足椭圆方程：

$$\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}=1\quad\text{(I)}$$

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}=1 \quad \text{(II)}$$

单独看每个方程只得到一个点的信息; 但解析几何研究点与点之间的关系, 点差法作 I-II 正是为了建立这种关系.

$$\left(\frac{x_1^2}{a^2}-\frac{x_2^2}{a^2}\right)+\left(\frac{y_1^2}{b^2}-\frac{y_2^2}{b^2}\right)=0$$

作差消去了等式右侧的常数项 1, 使等式变为齐次形式, 便于因式分解. 利用平方差公式, 上式可分解为:

$$\frac{(x_1-x_2)(x_1+x_2)}{a^2}+\frac{(y_1-y_2)(y_1+y_2)}{b^2}=0$$

代数变形完成后, 下面赋予这些表达式以几何意义. 观察式中四个关键组合项:

- x_1-x_2 与 y_1-y_2: 这两项是坐标的差值, 它们的比值 $\frac{y_1-y_2}{x_1-x_2}$ 正是弦 P_1P_2 的斜率 k, 这里要求 $x_1\neq x_2$.
- x_1+x_2 与 y_1+y_2: 这两项是坐标的和, 它们的一半 $\frac{x_1+x_2}{2}$ 和 $\frac{y_1+y_2}{2}$ 恰好是弦 P_1P_2 的中点 M 的坐标 (x_M,y_M).

引入弦斜率 k 与中点坐标 (x_M,y_M), 可把上述代数等式转化为几何方程:

$$\frac{(x_1-x_2)\cdot(2x_M)}{a^2}+\frac{k(x_1-x_2)\cdot(2y_M)}{b^2}=0$$

约去公因子 $2(x_1-x_2)$(弦不垂直于 x 轴), 得到仅联系弦中点与斜率的关系式:

$$\frac{x_M}{a^2}+k\frac{y_M}{b^2}=0$$

点差法分代数与几何两个层面.

1. **代数层面**. 通过作差消去常数项, 利用因式分解将二次关系转化为关于坐标和与坐标差的乘积形式.
2. **几何层面**. 将代数中的和与差分别诠释为几何中的中点与斜率.

它绕开了韦达定理法中联立方程的繁琐, 直接给出圆锥曲线自身关于平行弦中点轨迹的几何属性. 这种从点的方程到弦的性质的推演, 正是解析几何处理问题的方式.

联立、韦达、弦长、点差等工具已经齐备. 面对具体问题, 还须在两类参量中作出选择. 一类是**以曲线上动点的坐标为参量**, 另一类是**以直线方程中的参数为参量**. 前者称为**设点法**, 也叫点参法; 后者称为**设线法**, 也叫线参法或联立法.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

13.3.6 设点法与设线法预告

- **设点法**. 设 $P(x_0,y_0)$ 或弦端 A,B 在曲线上, 把距离、垂直、定比等条件写成关于坐标的等式, 再用曲线方程整体代换. 基本技术、参量选择与典型应用见 第13.4节 至 第13.6节.
- **设线法**. 设直线 $y=kx+m$(或截距式、横截式), 与曲线联立得一元二次方程, 由判别式与韦达读出交点个数、中点、弦长等对称量. 联立步骤、直线形式选择与后处理见 第13.7节 至 第13.9节.

二者可以并用, 点对点约束, 如 $\overrightarrow{F_1P}\cdot\overrightarrow{F_1Q}=0$ 常先设点再借韦达; 弦长、中点轨迹等则多以设线开局. 何时取舍见 第13.16节.

13.3.7 曲线系方程 *

处理一族具有共同性质的曲线, 例如过共同交点时, 逐一分析每条曲线方程不经济. 曲线系方程的思想是用一个含参数的方程统一表示这一族曲线, 把寻找特定曲线的问题转化为确定参数值的问题.

其理论基础如下:

定理 13.11 (曲线系原理) 设曲线 C_1,C_2 的方程分别为 $f(x,y)=0$ 和 $g(x,y)=0$. 对任意实数 λ, 方程

$$f(x,y)+\lambda g(x,y)=0 \tag{13-21}$$

的解集, 当左端不恒为零时可看作一条曲线包含 C_1 和 C_2 的所有公共点.

证明 设 $P(x_0,y_0)$ 是 C_1 和 C_2 的任意公共点. 根据定义, P 的坐标必然同时满足 C_1 和 C_2 的方程, 即

$$f(x_0,y_0)=0 \quad 且 \quad g(x_0,y_0)=0$$

现在, 将点 $P(x_0,y_0)$ 的坐标代入曲线系方程 $f(x,y)+\lambda g(x,y)=0$ 的左侧:

$$f(x_0,y_0)+\lambda g(x_0,y_0)=0+\lambda\cdot 0=0$$

故等式恒成立. 因此点 $P(x_0,y_0)$ 位于方程 $f(x,y)+\lambda g(x,y)=0$ 表示的曲线上. 此结论对所有公共点成立, 故该曲线经过 C_1 和 C_2 的所有公共点. □

这个原理的优势在于, 处理过交点问题时可以避免求解交点坐标, 直接在代数结构层面解决.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.3.7.1 共点直线系

最简单且最基础的曲线系是共点直线系. 若两直线 $l_1: A_1x+B_1y+C_1=0$ 和 $l_2: A_2x+B_2y+C_2=0$ 相交, 则过其交点的所有直线(除去 l_2)可由直线系方程表示:

$$(A_1x+B_1y+C_1)+\lambda(A_2x+B_2y+C_2)=0 \quad (\lambda\in\mathbb{R}) \tag{13-22}$$

此方程是曲线系原理在 f,g 均为一次多项式时的直接应用. 对任意 λ, 该方程都是关于 x,y 的一次方程, 表示一条过 l_1,l_2 交点的直线.

下面看几个例子.

例题 13.59 求经过两直线 $l_1: 2x-y-1=0$ 和 $l_2: x+y-5=0$ 的交点, 且与点 $P(1,-2)$ 的距离为 $\sqrt{5}$ 的直线方程.

知乎@余命数

解 所求直线属于过 l_1,l_2 交点的直线系, 其方程设为 $(2x-y-1)+\lambda(x+y-5)=0$, 整理为一般式

$$(2+\lambda)x+(\lambda-1)y-(1+5\lambda)=0.$$

点 $P(1,-2)$ 到该直线的距离为 $\sqrt{5}$, 由距离公式

$$d=\frac{|(2+\lambda)(1)+(\lambda-1)(-2)-(1+5\lambda)|}{\sqrt{(2+\lambda)^2+(\lambda-1)^2}}=\sqrt{5}.$$

分子 $|2+\lambda-2\lambda+2-1-5\lambda|=|-6\lambda+3|$, 分母根号内 $(2+\lambda)^2+(\lambda-1)^2=2\lambda^2+2\lambda+5$. 两边平方得

知乎@余命数

$$\frac{(-6\lambda+3)^2}{2\lambda^2+2\lambda+5}=5,$$

$$26\lambda^2-46\lambda-16=0 \quad\Longrightarrow\quad 13\lambda^2-23\lambda-8=0,$$

解得 $\lambda=\dfrac{23\pm3\sqrt{105}}{26}$. 代回直线系方程, 得所求两条直线为

知乎@余命数

$$\left(2+\frac{23+3\sqrt{105}}{26}\right)x+\left(\frac{23+3\sqrt{105}}{26}-1\right)y-\left(1+5\cdot\frac{23+3\sqrt{105}}{26}\right)=0,$$

$$\left(2+\frac{23-3\sqrt{105}}{26}\right)x+\left(\frac{23-3\sqrt{105}}{26}-1\right)y-\left(1+5\cdot\frac{23-3\sqrt{105}}{26}\right)=0.$$

等价地, 两边同乘以 26 后可写为

$$(75+3\sqrt{105})x+(-3+3\sqrt{105})y-(141+15\sqrt{105})=0,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$(75-3\sqrt{105})x+(-3-3\sqrt{105})y-(141-15\sqrt{105})=0.$$

13.3.7.2 过圆锥曲线交点的曲线系

设 $C_1:f(x,y)=0$ 和 $C_2:g(x,y)=0$ 是两条圆锥曲线, 则 $C_1+\lambda C_2=0$ 表示经过它们所有交点的圆锥曲线系.

例题 13.60 已知两圆 $C_1:x^2+y^2-2x-4y-4=0$ 和 $C_2:x^2+y^2+2x-6y=0$.

1. 求经过 C_1,C_2 交点的直线的方程.
2. 求经过 C_1,C_2 交点, 且过点 $P(3,1)$ 的圆的方程.

余命数 知乎@余命数 知乎@余命数

解 (1) 设经过两圆交点的曲线系方程为

$$(x^2+y^2-2x-4y-4)+\lambda(x^2+y^2+2x-6y)=0.$$

该方程表示一条直线, 当且仅当 x^2,y^2 的系数 $1+\lambda$ 同时为零, 即 $\lambda=-1$. 此时

$$(x^2+y^2-2x-4y-4)-(x^2+y^2+2x-6y)=-4x+2y-4=0,$$

化简得 $2x-y+2=0$, 此直线即两圆的**根轴**.

(2) 所求圆是过 C_1,C_2 交点的曲线系中的一员. 将点 $P(3,1)$ 代入以确定 λ, 其中 C_1 部分为 $3^2+1^2-2\cdot3-4\cdot1-4=-4$, C_2 部分为 $3^2+1^2+2\cdot3-6\cdot1=10$, 故 $-4+10\lambda=0$, 解得 $\lambda=\dfrac{2}{5}$. 代回曲线系方程并两边乘以 5,

$$5(x^2+y^2-2x-4y-4)+2(x^2+y^2+2x-6y)=0,$$

$$7x^2+7y^2-6x-32y-20=0,$$

知乎@余命数 知乎@余命数

即所求圆的方程.

例题 13.61 求经过椭圆 $C:x^2+4y^2=4$ 与两条直线 $l_1:x-y=0$ 及 $l_2:x+y=0$ 的所有四个交点, 并且经过点 $P(2,1)$ 的圆锥曲线的方程.

解 此题若直接求解四个交点坐标再用待定系数法求曲线方程, 计算极为繁琐, 曲线系方法提供了简捷的解法. 将两条直线 l_1,l_2 视为一条退化的二次曲线, 其方程为 $(x-y)(x+y)=x^2-y^2=0$, 图形即直线 $y=x$ 与 $y=-x$ 的并集. 问题转化为求经过椭圆 $C:x^2+4y^2-4=0$ 与退化二次曲线 $L:x^2-y^2=0$ 交点的曲线. 依曲线系原理设其方程为

$$(x^2+4y^2-4)+\lambda(x^2-y^2)=0,$$

知乎@余命数

此曲线族自动经过所有四个交点. 将点 $P(2,1)$ 代入以确定 λ:$(4+4-4)+\lambda(4-1)=0$, 即 $4+3\lambda=0$, 解得 $\lambda=-\dfrac{4}{3}$. 代回曲线系方程并两边乘以 3,

$$3(x^2+4y^2-4)-4(x^2-y^2)=0,$$
$$-x^2+16y^2-12=0 \quad\Longrightarrow\quad x^2-16y^2+12=0,$$

这是一条双曲线.

13.3.7.3 直线系与圆系

13.3.7.3.1 直线系方程 直线系是具有某种共同几何属性的一族直线的集合. 根据属性的不同, 其方程有几种标准形式.

- **共点直线系**
  经过两已知相交直线 $l_1:A_1x+B_1y+C_1=0$ 和 $l_2:A_2x+B_2y+C_2=0$ 交点的直线系方程为

  $$(A_1x+B_1y+C_1)+\lambda(A_2x+B_2y+C_2)=0 \quad (\lambda\in\mathbb{R})$$

  此方程族不包含 l_2 本身; 若要包含 l_2, 可用对称形式 $\mu(A_1x+B_1y+C_1)+\lambda(A_2x+B_2y+C_2)=0$.
  更基本的形式是过定点 $P_0(x_0,y_0)$ 的直线系, 方程可设为

  $$y-y_0=k(x-x_0)$$

  其中斜率 k 为参数. 此形式不能表示垂直于 x 轴的直线 $x=x_0$. 解题时通常需单独讨论斜率不存在的情况.
- **平行直线系**
  与已知直线 $Ax+By+C=0$ 平行的一族直线方程为

  $$Ax+By+\lambda=0 \quad (\lambda\in\mathbb{R})$$

  从几何上看, 该族直线具有相同法向量 (A,B), 因此斜率相等. 参数 λ 控制直线与原点的距离.

例题 13.62 求与直线 $3x-4y+1=0$ 平行, 且与圆 $x^2+y^2=4$ 相切的直线方程.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 所求直线属于与 $3x-4y+1=0$ 平行的直线系, 其方程设为 $3x-4y+\lambda=0$. 直线与圆 $x^2+y^2=4$ 相切, 等价于圆心 $O(0,0)$ 到该直线的距离等于半径 $R=2$. 圆心到直线的距离

$$d=\frac{|3(0)-4(0)+\lambda|}{\sqrt{3^2+(-4)^2}}=\frac{|\lambda|}{5}.$$

令 $d=R$, 即 $\dfrac{|\lambda|}{5}=2$, 解得 $|\lambda|=10$, $\lambda=\pm10$. 故所求切线有两条: $3x-4y+10=0$ 与 $3x-4y-10=0$.

13.3.7.3.2 圆系方程 与直线系类似, 圆系方程是解决两圆交点、切点等问题的实用工具.

- **过两圆交点的圆系**
  设两圆 $C_1: x^2+y^2+D_1x+E_1y+F_1=0$ 和 $C_2: x^2+y^2+D_2x+E_2y+F_2=0$ 相交, 则经过其交点的圆系方程为

  $$(x^2+y^2+D_1x+E_1y+F_1)+\lambda(x^2+y^2+D_2x+E_2y+F_2)=0 \quad (\lambda\neq-1)$$

  特别地, 当 $\lambda=-1$ 时, 方程中的二次项 x^2, y^2 被消去, 得到一个一次方程, 它表示过两圆公共弦所在的直线, 即两圆的**根轴**.
- **过圆与直线交点的圆系**
  这是上述情形的一个重要特例. 可把直线 $l: Ax+By+C=0$ 视为半径无限大的退化圆. 经过圆 $C: x^2+y^2+Dx+Ey+F=0$ 与直线 l 交点的圆系方程为

  $$(x^2+y^2+Dx+Ey+F)+\lambda(Ax+By+C)=0$$

- **相切圆系**
  与已知直线 $l: Ax+By+C=0$ 相切于点 $P_0(x_0,y_0)$ 的圆系方程, 可看成过直线 l 与半径为零的点圆 $(x-x_0)^2+(y-y_0)^2=0$ 交点的圆系. 其方程为

  $$(x-x_0)^2+(y-y_0)^2+\lambda(Ax+By+C)=0$$

例题 13.63 求经过圆 $C: x^2+y^2-4x=0$ 与直线 $l: x-2y-2=0$ 的交点, 且面积最小的圆的方程.

解 所求圆属于过圆 C 与直线 l 交点的圆系, 其方程设为

$$(x^2+y^2-4x)+\lambda(x-2y-2)=0,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

整理为 $x^2+y^2+(\lambda-4)x-2\lambda y-2\lambda=0$, 配方得

$$\left(x+\frac{\lambda-4}{2}\right)^2+(y-\lambda)^2=\left(\frac{\lambda-4}{2}\right)^2+\lambda^2+2\lambda.$$

知乎@余命数

面积 $S=\pi R^2$ 最小等价于半径平方 R^2 最小. 令 $f(\lambda)=R^2=\dfrac{\lambda^2-8\lambda+16}{4}+\lambda^2+2\lambda=\dfrac{5}{4}\lambda^2+4$, 在 $\lambda=0$ 时取最小值 4. 此时圆即原始圆 $C:x^2+y^2-4x=0$.

此结果的几何意义是, 过圆与直线两交点的所有圆中, 以公共弦为直径的圆面积最小. 圆心 $C(2,0)$, 直线 l 斜率 $\dfrac{1}{2}$. 以公共弦为直径的圆, 其圆心必在直线 l 上. 圆系方程的圆心坐标为 $(-\dfrac{\lambda-4}{2},\ \lambda)$, 代入 l 的方程 $x-2y-2=0$ 得

$$-\frac{\lambda-4}{2}-2\lambda-2=0\implies -5\lambda=0\implies \lambda=0,$$

与求最值的结果一致. 故面积最小的圆是 $x^2+y^2-4x=0$.

例题 13.64　求与直线 $y=2x$ 相切于点 $P(1,2)$, 且过点 $Q(3,4)$ 的圆的方程.

解　这是相切圆系的经典应用. 切线方程 $2x-y=0$, 切点 $P(1,2)$. 所求圆属于与直线 $2x-y=0$ 相切于点 $(1,2)$ 的圆系, 其方程设为

$$(x-1)^2+(y-2)^2+\lambda(2x-y)=0.$$

知乎@余命数　知乎@余命数

该圆过点 $Q(3,4)$, 将 Q 代入确定 λ: $(3-1)^2+(4-2)^2+\lambda(2\cdot3-4)=0$, 即 $8+2\lambda=0$, 解得 $\lambda=-4$. 代回圆系方程, 展开整理

$$(x^2-2x+1)+(y^2-4y+4)-8x+4y=0,$$
$$x^2+y^2-10x+5=0,$$

即所求圆的方程.

13.3.7.4　二次曲线系

下面把曲线系原理从直线和圆推广到一般二次曲线. 基本形式 $C_1+\lambda C_2=0$ 描述经过 C_1,C_2 所有(至多四个)交点的一族二次曲线.

其最有力的应用策略是将其中一条二次曲线退化为两条直线.

13.3.7.4.1　退化的二次曲线　二次方程 $f(x,y)=0$ 所代表的图形并不总是椭圆、双曲线或抛物线. 当 $f(x,y)$ 可分解为两个一次多项式乘积时, 其图形是两条直线

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

的并集. 例如, 两条直线 $l_1: A_1x+B_1y+C_1=0$ 和 $l_2: A_2x+B_2y+C_2=0$ 的并集可由一个二次方程描述:

知乎@余命数

$$(A_1x+B_1y+C_1)(A_2x+B_2y+C_2)=0$$

这个方程的解集正是满足 $l_1=0$ 或 $l_2=0$ 的点的集合. 这样的图形称为**退化的二次曲线**. 它是理解和构造二次曲线系的基础.

13.3.7.4.2 过四点的二次曲线系 现在考虑一个常用场景, 求经过一条已知二次曲线 C 与两条已知直线 l_1, l_2 的四个交点的曲线. 设 $C=0, l_1=0, l_2=0$ 分别是它们的方程. 可构造退化二次曲线 $l_1l_2=0$. 根据曲线系原理, 方程

$$C+\lambda(l_1l_2)=0 \tag{13-23}$$

表示经过这四个交点的一族二次曲线.

例题 13.65 求经过椭圆 $C_1: x^2+2y^2=3$ 与圆 $C_2: x^2+y^2=2$ 的四个交点, 且过原点的二次曲线方程.

解 所求二次曲线属于过 C_1, C_2 交点的曲线系, 其方程设为

$$(x^2+2y^2-3)+\lambda(x^2+y^2-2)=0,$$

知乎@余命数

此曲线族自动经过 C_1, C_2 的所有交点. 该曲线经过原点, 将 $(0,0)$ 代入得 $-3-2\lambda=0$, 解得 $\lambda=-\dfrac{3}{2}$. 代回曲线系方程, 两边乘以 2,

$$\begin{aligned} 2(x^2+2y^2-3)-3(x^2+y^2-2)&=0,\\ -x^2+y^2=0 \quad\Longrightarrow\quad y^2&=x^2, \end{aligned}$$

分解为 $(y-x)(y+x)=0$, 图形是直线 $y=x$ 与 $y=-x$ 的并集, 即退化的双曲线.

13.3.7.4.3 从相交到相切 曲线系除处理离散交点外, 也能处理相切这种重合交点的极限情形. 几何上, 切线是割线当两个交点无限逼近并重合时的极限. 代数上, 解方程组所得二次方程的判别式 $\Delta\to 0$, 两个不等实根合并为一个重根.

基于此思想可推广曲线系处理相切问题. 当 l_1, l_2 重合为 l 时, 退化二次曲线 $l_1l_2=0$ 演变成 $l^2=0$. $l^2=0$ 的几何意义是一条双重直线. 由此得到处理相切问题的实用曲线系:

$$C+\lambda l^2=0 \tag{13-24}$$

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此方程描述一族与二次曲线 C 在它与直线 l 的两个交点处有二重接触的曲线. 若 l 与 C 交于两点, 则该曲线系中所有曲线都在这两点与 C 相切. 若 l 与 C 相切, 则该曲线系中所有曲线都在该切点处与 C 有更高阶的接触.

13.3.7.4.4　单点切触的曲线系　构造一族与某条直线 l 在某点 P_0 相切的曲线时, 可用一个特例. 与已知直线 $l: Ax+By+C=0$ 相切于点 $P_0(x_0,y_0)$ 的圆系, 可由 P_0 处的半径为零的点圆与直线 l 线性组合而成. 点圆方程为 $(x-x_0)^2+(y-y_0)^2=0$. 圆系方程为

$$(x-x_0)^2+(y-y_0)^2+\lambda(Ax+By+C)=0 \tag{13-25}$$

例题 13.66　求与抛物线 $C: y^2=4x$ 在顶点 $(0,0)$ 处相切, 且经过点 $P(1,2)$ 的圆的方程.

解　把曲线与曲线相切转化为代数构造, 即两条曲线在某点相切, 当且仅当在该点有公共切线. 抛物线 $y^2=4x$ 在顶点 $(0,0)$ 处的切线为 y 轴 $x=0$, 故原问题等价于求一个过点 $P(1,2)$ 且与直线 $l: x=0$ 相切于点 $O(0,0)$ 的圆. 依相切圆系构造, 切点 $O(0,0)$ 对应点圆 $x^2+y^2=0$, 切线 $l: x=0$, 故所有满足相切条件的圆的方程为

$$(x^2+y^2)+\lambda x=0.$$

该圆又须过 $P(1,2)$, 代入得 $5+\lambda=0$, 解得 $\lambda=-5$. 代回圆系方程, 即得所求圆 $x^2+y^2-5x=0$. 这一构造把相切的几何约束提炼为相切圆系的代数结构, 再以一次代入定参, 计算量随之减少.

13.3.8 仿射变换 *

在处理椭圆几何问题时, 常见的困难是基本原理(联立方程、韦达定理)清晰, 但直接计算往往十分繁重.

根本原因在于椭圆方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 中两个坐标变量的不对等. 相比之下, 圆 $x^2+y^2=R^2$ 具有高度对称性, 其几何性质, 如半径处处相等、垂径定理也更简洁.

由此提出一个根本问题, 即能否找到一种几何变换, 把椭圆化归为圆, 从而将复杂的椭圆问题转化到熟悉的圆的领域中去解决; 若存在这样的变换, 它需要满足什么条件, 又该如何利用.

13.3.8.1 构造伸缩变换

寻找一个坐标映射 $T:(x,y)\mapsto(x',y')$, 把椭圆 $C: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上的点 (x,y) 映成圆 C' 上的点 (x',y').

知乎@余命数

为使变换尽可能简单, 我们保持坐标系的基本结构不变. 一个自然的想法是对坐标轴进行独立伸缩. 观察椭圆方程, 可将其改写为

$$\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2=1$$

我们希望新坐标 (x',y') 满足圆的方程, 例如 $x'^2+y'^2=R^2$. 比较上述两个方程的形式, 一个非常自然的对应关系是

$$\frac{x'}{R}=\frac{x}{a}\quad\text{且}\quad\frac{y'}{R}=\frac{y}{b}$$

为进一步简化, 选择一个特定的目标圆. 例如, 令变换不改变 x 轴尺度, 即 $x'=x$. 此时, 为使方程匹配, 取目标圆半径 $R=a$. 由 $\frac{x'}{a}=\frac{x}{a}$ 得 $x'=x$. 由 $\frac{y'}{a}=\frac{y}{b}$ 得 $y'=\frac{a}{b}y$.

至此, 我们构造出满足需求的坐标变换

$$T:\begin{cases}x'=x\\ y'=\frac{a}{b}y\end{cases}\tag{13-26}$$

这个变换在几何上相当于把平面沿 y 轴方向伸缩 $\frac{a}{b}$ 倍, $a>b$ 时伸长, $a<b$ 时压缩, 而 x 轴方向不变. 通过这个变换, 椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 被映到圆 $x'^2+y'^2=a^2$.

13.3.8.2 伸缩变换后对几何性质的影响

我们已找到把椭圆化为圆的变换, 但要真正应用它, 还须考察其他几何对象和几何关系在变换下的表现.

定理 13.12 (共线性保持) 在伸缩变换 $T:(x,y)\mapsto(x,\frac{a}{b}y)$ 下, 直线映为直线.

证明 设原平面内直线方程为 $Ax+By+C=0$. 变换的逆关系为 $x=x',y=\frac{b}{a}y'$. 将其代入直线方程, 得到新坐标系下的方程,

$$A(x')+B\left(\frac{b}{a}y'\right)+C=0\implies Ax'+\frac{Bb}{a}y'+C=0$$

由于 A,B 不全为零, 故 $A,\frac{Bb}{a}$ 也不全为零. 此方程是关于 x',y' 的一次方程, 故其像仍为一条直线. □

定理 13.13 (平行性保持) 在伸缩变换 T 下, 两条互不重合的平行直线映为两条仍互不重合且平行的直线.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

证明 设原平面内两条平行直线为 $l_1: Ax+By+C_1=0$ 和 $l_2: Ax+By+C_2=0$. 由定理 1 的证明, 它们的像分别为 $l_1': Ax'+\frac{Bb}{a}y'+C_1=0$ 和 $l_2': Ax'+\frac{Bb}{a}y'+C_2=0$. 由于 x', y' 的系数对应相等, 仅常数项不同, 故 $l_1' \parallel l_2'$. □

定理 13.14 (定比分点保持) 在伸缩变换 T 下, 非退化线段的有限定比分点(包括中点)映为像线段的对应定比分点.

证明 设非退化线段 P_1P_2 的端点为 $P_1(x_1,y_1), P_2(x_2,y_2)$. 设 M 为 P_1P_2 的有限定比分点, 满足 $\overrightarrow{P_1M}=\lambda\overrightarrow{MP_2}$ $(\lambda\neq -1)$.

其坐标为 $M\left(\frac{x_1+\lambda x_2}{1+\lambda}, \frac{y_1+\lambda y_2}{1+\lambda}\right)$.

各点经变换后的像为 $P_1'(x_1, \frac{a}{b}y_1), P_2'(x_2, \frac{a}{b}y_2), M'\left(\frac{x_1+\lambda x_2}{1+\lambda}, \frac{a}{b}\frac{y_1+\lambda y_2}{1+\lambda}\right)$. 计算像线段 $P_1'P_2'$ 上满足 $\overrightarrow{P_1'M''}=\lambda\overrightarrow{M''P_2'}$ 的点 M'':

$$x_{M''}=\frac{x_1+\lambda x_2}{1+\lambda},\quad y_{M''}=\frac{\frac{a}{b}y_1+\lambda\frac{a}{b}y_2}{1+\lambda}=\frac{a}{b}\frac{y_1+\lambda y_2}{1+\lambda}$$

可见 M'' 的坐标与 M' 完全相同. 当 $\lambda=1$ 时, 此结论对中点成立. □

定理 13.15 (相切性保持) 在伸缩变换 T 下, 若一条直线与一条二次曲线相切, 则它们的像也相切.

证明 相切的代数定义是, 直线与二次曲线的方程联立后, 在切点处得到二重交点. 伸缩变换 T 是可逆仿射变换, 其逆变换也存在且唯一. 将像直线与像二次曲线用逆变换代回原坐标, 等价于原直线与原二次曲线的方程组, 因而交点及其重数一一对应. 因此, 原来的二重交点在变换后仍对应二重交点, 相切关系得以保持. □

该变换**不保持**长度和角度. 例如, 点 $(0,b)$ 变换为 $(0,a)$, 其到原点的距离从 b 变为 a. 两条垂直的直线 $y=x$ 和 $y=-x$ 变换为 $y'=\frac{a}{b}x'$ 和 $y'=-\frac{a}{b}x'$, 它们不再垂直(除非 $a=b$). 因此, 所有依赖长度、距离、垂直关系的计算, 必须在逆变换后回到原坐标系中进行.

定理 13.16 (多边形面积变换) 在伸缩变换 T 下, 面积为 S 的简单多边形映为面积为 S' 的简单多边形, 且 $S'=\frac{a}{b}S$.

证明 先证明此结论对任意三角形成立. 设原平面内三角形顶点为 $P_1(x_1,y_1), P_2(x_2,y_2), P_3(x_3,y_3)$. 其面积 $S_\triangle$ 可由坐标面积公式(鞋带公式)给出,

$$S_\triangle=\frac{1}{2}|x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)|$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

经变换 T 后, 像三角形顶点为 $P_1'(x_1, \frac{a}{b}y_1), P_2'(x_2, \frac{a}{b}y_2), P_3'(x_3, \frac{a}{b}y_3)$. 其面积 $S_{\triangle}'$ 为

$$\begin{aligned} S_{\triangle}' &= \frac{1}{2}\left|x_1\left(\frac{a}{b}y_2 - \frac{a}{b}y_3\right) + x_2\left(\frac{a}{b}y_3 - \frac{a}{b}y_1\right) + x_3\left(\frac{a}{b}y_1 - \frac{a}{b}y_2\right)\right| \\ &= \frac{1}{2}\left|\frac{a}{b}\left(x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)\right)\right| \end{aligned}$$

由于 a, b 均为正数, 可将因子 $\frac{a}{b}$ 提出绝对值符号,

$$S_{\triangle}' = \frac{a}{b} \cdot \frac{1}{2}|x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)| = \frac{a}{b}S_{\triangle}$$

对于任意一个多边形, 如四边形, 总可将其剖分为有限个不重叠的三角形. 设一个多边形被剖分为 n 个三角形 $\triangle_1, \triangle_2, \ldots, \triangle_n$. 其总面积 $S = \sum_{i=1}^{n} S_{\triangle_i}$. 经变换后, 该多边形的像由 n 个像三角形 $\triangle_1', \triangle_2', \ldots, \triangle_n'$ 组成. 其总面积 $S' = \sum_{i=1}^{n} S_{\triangle_i}'$. 根据已证的三角形面积变换关系, $S_{\triangle_i}' = \frac{a}{b}S_{\triangle_i}$ 对所有 i 成立. 因此

$$S' = \sum_{i=1}^{n}\left(\frac{a}{b}S_{\triangle_i}\right) = \frac{a}{b}\sum_{i=1}^{n} S_{\triangle_i} = \frac{a}{b}S$$

此结论对任意多边形均成立. □

基于上述理论, 我们形成清楚的解题流程:

1. **变换**. 识别问题中的椭圆, 施加伸缩变换, 将其转化为圆. 同时, 把问题中相关的点与线一起做同步变换.
2. **求解**. 在变换后的圆中, 利用圆的简洁几何性质, 如对称性与垂径定理, 来解决问题. 这一步通常在代数上或几何上都更为简单.
3. **逆变**. 将圆中得到的结论, 依据变换的几何不变性, 如平行、中点、相切, 或按确定的规律换算面积, 翻译回原始的椭圆.

例题 13.67 求内接于椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ 的三角形的最大面积.

解 设椭圆 $C: \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ 内接 $\triangle PQR$. 施行仿射变换 $T: (x, y) \mapsto (x', y') = (x, \frac{a}{b}y)$, 椭圆 C 变为圆 $C': x'^2 + y'^2 = a^2$, 内接 $\triangle PQR$ 变为圆 C' 的内接 $\triangle P'Q'R'$. 问题转化为求圆 $x'^2 + y'^2 = a^2$ 内接三角形的最大面积. 圆内接三角形以等边三角形面积为最大, 半径 $R = a$ 的圆内接等边三角形面积为 $S'_{\max} = \frac{3\sqrt{3}}{4}a^2$. 依面积变换规律 $S' = \frac{a}{b}S$, 即 $S = \frac{b}{a}S'$, 故椭圆内接三角形的最大面积为

$$S_{\max} = \frac{b}{a}S'_{\max} = \frac{b}{a} \cdot \frac{3\sqrt{3}}{4}a^2 = \frac{3\sqrt{3}}{4}ab.$$

此方法将椭圆最值问题经一个线性变换化归为圆的初等最值问题.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.68　设直线 l 与椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 交于 A,B 两点，M 为弦 AB 的中点. 若直线 OM（O 为坐标原点）与椭圆交于点 P. 证明，椭圆在点 P 处的切线与直线 l 平行.

解　这是椭圆共轭直径的基本性质，直接代数证明较为繁琐，仿射变换则直观且基本. 对整个图形施行伸缩变换 $T:(x,y)\mapsto(x',y')=(x,\frac{a}{b}y)$. 在变换后的平面中，椭圆变为圆 $C':x'^2+y'^2=a^2$，直线 l 变为 l'，交点 A,B 变为 A',B'. 由变换的中点保持性质，弦 AB 的中点 M 变为弦 $A'B'$ 的中点 M'；原点 O 不变，直线 OM 变为 OM'，交点 P 变为 P'. 由相切保持性质，椭圆在 P 处的切线 l_P 变为圆在 P' 处的切线 $l'_{P'}$.

转入圆的平面分析：$A'B'$ 为弦，M' 为其中点，O 为圆心. 由垂径定理，$OM'\perp l'$. 又 P' 在 OM' 上，故 OP' 为半径，圆在 P' 处的切线 $l'_{P'}\perp OP'$. 于是 $l'\perp OM'$ 且 $l'_{P'}\perp OM'$，同垂直于 OM' 的两直线平行，故 $l'_{P'}\parallel l'$.

仿射变换保持平行性，故原像直线 l_P 与 l 亦平行，即椭圆在点 P 处的切线与直线 l 平行.

13.3.9　离心率例题

离心率是圆锥曲线的一个主要参数，求离心率的途径十分多样，可以从定义出发建立 a,c 的方程，也可以利用焦点三角形，借助正余弦定理，数形结合，基本不等式等.

例题 13.69 (建立关于 a 和 c 的一次或二次方程与不等式)　如图，设椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左焦点为 F，上顶点为 A，右顶点为 B，且 $\overrightarrow{FA}\cdot\overrightarrow{AB}=0$，则 C 的离心率为 __.

解　答案：$\frac{\sqrt{5}-1}{2}$. 由椭圆方程可知，$F(-c,0)$、$A(0,b)$、$B(a,0)$，则 $\overrightarrow{FA}=(c,b)$，$\overrightarrow{AB}=(a,-b)$. 由 $\overrightarrow{FA}\cdot\overrightarrow{AB}=0$ 得 $ac-b^2=0$，即 $ac=b^2=a^2-c^2$，整理得 $c^2+ac-a^2=0$，即 $e^2+e-1=0$，解得 $e=\frac{\sqrt{5}-1}{2}$（负值舍去）.

例题 13.70 (圆锥曲线第一定义)　已知 P 为椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 上一点，F_1、F_2 分别为其左、右焦点，O 为坐标原点，$|PO|=\frac{\sqrt{3}}{2}a$，且 $\overrightarrow{PF_1}\cdot\overrightarrow{PF_2}=\frac{3}{4}a^2$，则 C 的离心率为（　）

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 13.71 (圆锥曲线第二定义) 古希腊数学家欧几里得在《几何原本》中描述了圆锥曲线的共性，并给出了圆锥曲线的统一定义，他指出，平面内到定点的距离与到定直线的距离的比是常数 e 的点的轨迹叫做圆锥曲线；当 $0<e<1$ 时，轨迹为椭圆；当 $e=1$ 时，轨迹为抛物线；当 $e>1$ 时，轨迹为双曲线. 则方程 $\frac{\sqrt{(x-4)^2+y^2}}{|25-4x|}=\frac{1}{5}$ 表示的圆锥曲线的离心率 e 等于(　)

A. $\frac{1}{5}$　　C. $\frac{5}{4}$

B. $\frac{4}{5}$　　D. 5

解 答案：B. 因为 $\frac{\sqrt{(x-4)^2+y^2}}{|25-4x|}=\frac{\sqrt{(x-4)^2+y^2}}{4|x-\frac{25}{4}|}=\frac{1}{5}$，

例题 13.72 (圆锥曲线第三定义 (斜率之积)) 已知双曲线 $E:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 与直线 $y=kx$ 相交于 A，B 两点，点 P 为双曲线 E 上的一个动点，记直线 PA，PB 的斜率分别为 k_1，k_2，若 $k_1k_2=\frac{1}{4}$，且双曲线 E 的右焦点到其一条渐近线的距离为 1，则双曲线 E 的离心率为__.

解 答案：$\frac{\sqrt{5}}{2}$. 设点 $A(x_1,y_1)$、$B(-x_1,-y_1)$、$P(x_0,y_0)$，则 $\frac{x_1^2}{a^2}-\frac{y_1^2}{b^2}=1$ 且 $\frac{x_0^2}{a^2}-\frac{y_0^2}{b^2}=1$，两式相减，得 $\frac{x_1^2-x_0^2}{a^2}=\frac{y_1^2-y_0^2}{b^2}$，所以 $\frac{y_1^2-y_0^2}{x_1^2-x_0^2}=\frac{b^2}{a^2}$，因为 $k_{PA}\cdot k_{PB}=k_1k_2=\frac{(y_0-y_1)(y_0+y_1)}{(x_0-x_1)(x_0+x_1)}=\frac{1}{4}$，所以 $\frac{b^2}{a^2}=\frac{1}{4}$，所以 $\frac{b}{a}=\frac{1}{2}$，所以双曲线的渐近线方程为 $y=\pm\frac{1}{2}x$，即 $x\pm2y=0$，因为焦点 $F_2(c,0)$ 到渐近线 $x-2y=0$ 的距离为 1，所以 $\frac{c}{\sqrt{5}}=1$，可得 $c=\sqrt{5}$，又因为 $c^2=a^2+b^2$，所以 $a=2,b=1$，所以双曲线的离心率 $e=\frac{\sqrt{5}}{2}$. 故答案为：$\frac{\sqrt{5}}{2}$

例题 13.73 (利用数形结合求解) 如图 1 所示，双曲线具有光学性质，从双曲线右焦点发出的光线经过双曲线镜面反射，其反射光线的反向延长线经过双曲线的左焦点. 若双曲线 $E:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的左、右焦点分别为 F_1,F_2，从 F_2 发出的光线经过图 2 中的 A,B 两点反射后，分别经过点 C 和 D，且 $\tan\angle CAB=-\frac{12}{5}$，$|\overrightarrow{BD}|^2=\overrightarrow{AD}\cdot\overrightarrow{BD}$，则双曲线 E 的离心率为(　)

A. $\frac{6}{5}$　　C. $\frac{2\sqrt{10}}{5}$

B. $\frac{\sqrt{37}}{5}$　　D. $\frac{\sqrt{14}}{3}$

解 答案：B. 如图，由 $|\overrightarrow{BD}|^2=\overrightarrow{AD}\cdot\overrightarrow{BD}$，有 $\overrightarrow{BD}^2+\overrightarrow{DA}\cdot\overrightarrow{BD}=0$，可得 $\overrightarrow{BD}\cdot(\overrightarrow{BD}+\overrightarrow{DA})=0$，可得 $\overrightarrow{BD}\cdot\overrightarrow{BA}=0$，有 $BD\perp AB$. 在 $Rt\triangle ABD$ 中，由 $\tan\angle F_1AB=\frac{12}{5}$，不妨设 $|\overrightarrow{BF_1}|=12m(m>0)$，则 $|\overrightarrow{AB}|=5m$，由勾股定理得 $|\overrightarrow{AF_1}|=13m$，又由双曲线的定义可得 $|\overrightarrow{AF_2}|=13m-2a$，$|\overrightarrow{BF_2}|=12m-2a$，根据 $|\overrightarrow{BF_1}|+|\overrightarrow{BF_2}|=|\overrightarrow{AB}|$ 可得 $(13m-2a)+(12m-2a)=5m$，解得 $a=5m$，所以 $|\overrightarrow{BF_2}|=2m$，在 $Rt\triangle F_1BF_2$ 中，

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

$2c=|\overrightarrow{F_1F_2}|=\sqrt{144m^2+4m^2}=2\sqrt{37}m$, 可得 $c=\sqrt{37}m$, 故双曲线 E 的离心率为 $e=\frac{c}{a}=\frac{\sqrt{37}m}{5m}=\frac{\sqrt{37}}{5}$. 故选:B.

例题 13.74 (利用正弦定理) 已知 F_1,F_2 为双曲线 $C:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的左、右焦点, M 为 C 左支上一点. 设 $\angle MF_1F_2=\alpha$, $\angle MF_2F_1=\beta$, 且 $\sin\frac{\alpha+\beta}{2}=2\sin\frac{\alpha-\beta}{2}$, 则 C 的离心率为(　)

乎@余命数

知乎@余命数

例题 13.75 (利用余弦定理) (2024·河北衡水·模拟预测) 已知双曲线 $E:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的右焦点为 F,过点 F 作直线 l 与渐近线 $bx-ay=0$ 直, 垂足为点 P, 延长 PF 交 E 于点 Q. 若 $\overrightarrow{FQ}=3\overrightarrow{PF}$, 则 E 的离心率为(　)

A. $\frac{6}{5}$　　C. $\frac{4}{3}$

B. $\frac{5}{4}$　　D. $\sqrt{2}$

解 答案:B. 设 $F(c,0)$,O 为坐标原点,则 $|PF|=\frac{|bc|}{\sqrt{a^2+b^2}}=\frac{bc}{c}=b$,从而 $\cos\angle OFP=\frac{b}{c}$. 设 E 的左焦点为 F', $|QF|=t$, 连接 QF', 由双曲线的定义, 得 $|QF'|=t+2a$. 在 $\triangle QF'F$ 中, 由余弦定理, 得 $(t+2a)^2=t^2+(2c)^2-2t\times 2c\times\left(-\frac{b}{c}\right)$, 解得 $t=\frac{b^2}{a-b}$. 由 $\overrightarrow{FQ}=3\overrightarrow{PF}$, 得 $\frac{b^2}{a-b}=3b$, 解得 $\frac{b}{a}=\frac{3}{4}$,

知乎@余命数

例题 13.76 (内切圆问题) (2024·高三·广东广州·期中) 已知点 P 是双曲线 $C:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 右支上一点, F_1、F_2 分别为双曲线 C 的左、右焦点, $\triangle PF_1F_2$ 的内切圆与 x 轴相切于点 N, 若 $\overrightarrow{PN}=\frac{1}{4}\overrightarrow{PF_1}+\frac{3}{4}\overrightarrow{PF_2}$, 则双曲线 C 的离心率为__.

解 答案: 2. 直线 PF_1,PF_2 分别与内切圆的切点为 M,Q, 如图所示. 由切线的性质可得 $\begin{cases}|PM|=|PQ|\\|F_1M|=|F_1N|\\|F_2Q|=|F_2N|\end{cases}$ 由双曲线的定义可得 $|PF_1|-|PF_2|=2a$, 即 $|PM|+|F_1M|-(|PQ|+|F_2Q|)=2a$, 所以 $|F_1M|-|F_2Q|=2a$, 即 $|F_1N|-|F_2N|=2a$, 又 $|F_1N|+|F_2N|=2c$, 因此 $|F_1N|=c+a$. 设 $\overrightarrow{PN}=t\overrightarrow{PF_1}$, 则 $\overrightarrow{PN}=\overrightarrow{PF_1}+\overrightarrow{F_1N}=\overrightarrow{PF_1}+t\overrightarrow{F_1F_2}=\overrightarrow{PF_1}+t(\overrightarrow{PF_2}-\overrightarrow{PF_1})=(1-t)\overrightarrow{PF_1}+t\overrightarrow{PF_2}$, 又 $\overrightarrow{PN}=\frac{1}{4}\overrightarrow{PF_1}+\frac{3}{4}\overrightarrow{PF_2}$, 因此 $t=\frac{3}{4}$, 于是 $\overrightarrow{F_1N}=\frac{3}{4}\overrightarrow{F_1F_2}$, 即 $|F_1N|=\frac{3}{2}c$,

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.77 (圆与双曲线共焦点) (2024·四川成都·模拟预测) 已知椭圆 $C_1:\frac{x^2}{a_1^2}+\frac{y^2}{b_1^2}=1(a_1>b_1>0)$ 与双曲线 $C_2:\frac{x^2}{a_2^2}-\frac{y^2}{b_2^2}=$

知乎@余命数

例题 13.78 (利用最大顶角) 已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b\geqslant 0)$, 点 A,B 是长轴的两个端点, 若椭圆上存在点 P, 使得 $\angle APB=120^\circ$, 则该圆的离心率的取值范围是(　)

知乎@余命数

A. $\left[\frac{\sqrt{6}}{3},1\right)$　　C. $\left(0,\frac{\sqrt{2}}{2}\right]$

B. $\left[\frac{\sqrt{3}}{2},1\right)$　　D. $\left(0,\frac{3}{4}\right]$

解 答案: A. 如图. 当 P 在上顶点时, $\angle APB$ 最大, 此时 $\angle APB\geqslant 120^\circ$, 则 $\angle APO\geqslant 60^\circ$, 所以 $\tan\angle APO\geqslant\tan 60^\circ=\sqrt{3}$, 即 $\frac{a}{b}\geqslant\sqrt{3}$、$a^2\geqslant 3b^2$、$a^2\geqslant 3(a^2-c^2)$ 所以 $2a^2\leqslant 3c^2$, 则 $\frac{c}{a}\geqslant\frac{\sqrt{6}}{3}$ 所以椭圆的离心率的取值范围是 $\left[\frac{\sqrt{6}}{3},1\right)$, 故选: A

例题 13.79 (基本不等式) 设椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 的右焦点为 F, 椭圆 C 上的两点 A,B 关于原点对称, 且满足 $\overrightarrow{FA}\cdot\overrightarrow{FB}=0$, $|FB|\leqslant|FA|\leqslant\sqrt{3}|FB|$, 则椭圆 C 的离心率的取值范围为(　)

A. $\left[\frac{\sqrt{2}}{2},1\right)$　　C. $\left[\sqrt{3}-1,1\right)$

B. $\left[\frac{\sqrt{2}}{2},\sqrt{3}-1\right]$　　D. $\left[\frac{\sqrt{2}}{2},\frac{\sqrt{3}}{2}\right]$

知乎@余命数

解 答案: B. 如图所示. 设椭圆的左焦点 F', 由椭圆的对称性可知, 四边形 $AFBF'$ 为平行四边形, 又 $\overrightarrow{FA}\cdot\overrightarrow{FB}=0$, 即 $FA\perp FB$, 所以四边形 $AFBF'$ 为矩形, $\therefore|AB|=|FF'|=2c$, 设 $|AF'|=n$, $|AF|=m$, 在直角 $\triangle ABF'$ 中, $m+n=2a$, $m^2+n^2=4c^2$, 得 $mn=2b^2$, 所以 $\frac{m}{n}+\frac{n}{m}=\frac{2c^2}{b^2}$, 令 $\frac{m}{n}=t$, 得 $t+\frac{1}{t}=\frac{2c^2}{b^2}$, 又 $|FB|\leqslant|FA|\leqslant\sqrt{3}|FB|$, 得 $\frac{m}{n}=t\in[1,\sqrt{3}]$, 所以 $t+\frac{1}{t}=\frac{2c^2}{b^2}\in\left[2,\frac{4\sqrt{3}}{3}\right]$, 所以 $\frac{c^2}{b^2}\in\left[1,\frac{2\sqrt{3}}{3}\right]$, 即 $\frac{b^2}{a^2}\in\left[2\sqrt{3}-3,\frac{1}{2}\right]$, 所以 $e^2\in\left[\frac{1}{2},4-2\sqrt{3}\right]$ 所以椭圆 C 的离心率的取值范围为 $e\in\left[\frac{c}{a}\in\left[\frac{\sqrt{2}}{2},\sqrt{3}-1\right]\right]$, 故选: B

例题 13.80 (十二·已知 $\overrightarrow{PF_1},\overrightarrow{PF_2}$ 范围) (2024·四川省南充市白塔中学高三开学考试) 已知 F_1、F_2 分别为椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左、右焦点, A 为右顶点, B 为上顶点, 若在线段 AB 的内部存在不同的两点 $P_i(i=1,2)$, 使得 $\overrightarrow{P_iF_1}\cdot\overrightarrow{P_iF_2}$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.81 ($|PF_1|=\lambda|PF_2|$) 已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 的左、右焦点分别为 $F_1(-c,0),F_2(c,0)$，若椭圆 C 上存在一点 P，使得 $\frac{\sin\angle PF_2F_1}{\sin\angle PF_1F_2}=\frac{c}{a}$，则椭圆 C 的离心率的取值范围为(　)

A. $(0,\frac{\sqrt{2}}{2})$　　C. $(\sqrt{2}-1,1)$

B. $(0,\sqrt{2}-1)$　　D. $(\frac{\sqrt{2}}{2},1)$

解 答案：C. 在 $\triangle PF_1F_2$ 中，由正弦定理可得 $\frac{|PF_2|}{\sin\angle PF_1F_2}=\frac{|PF_1|}{\sin\angle PF_2F_1}$，又由 $\frac{\sin\angle PF_2F_1}{\sin\angle PF_1F_2}=\frac{c}{a}$，得 $\frac{|PF_1|}{|PF_2|}=\frac{c}{a}$. 设点 $P(x_0,y_0)$，可得 $|PF_1|=a+ex_0,|PF_2|=a-ex_0$，则 $a(a+ex_0)=c(a-ex_0)$，解得 $x_0=\frac{a(e-1)}{e(e+1)}$. 由椭圆的几何性质可得 $x_0>-a$，即 $\frac{a(e-1)}{e(e+1)}>-a$，整理得 $e^2+2e-1>0$，解得 $e<-\sqrt{2}-1$ 或 $e>\sqrt{2}-1$. 又由 $e\in(0,1)$，所以椭圆的离心率的取值范围是 $(\sqrt{2}-1,1)$，故选：C.

例题 13.82 (中点弦问题)（2024·陕西西安·西安市大明宫中学校考模拟预测）已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的焦距为 $2c$，左焦点为 F，直线 l 与 C 相交于 A,B 两点，点 P 是线段 AB 的中点，P 的横坐标为 $\frac{1}{3}c$. 若直线 l 与直线 PF 的斜率之积等于 $-\frac{3}{16}$，则 C 的离心率为__.

解 答案：$\frac{1}{2}$. $F(-c,0)$，设 $A(x_1,y_1),B(x_2,y_2)$. 因为点 P 是线段 AB 的中点，P 的横坐标为 $\frac{1}{3}c$，所以 $x_1+x_2=\frac{2c}{3}$，$P\left(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2}\right)=\left(\frac{c}{3},\frac{y_1+y_2}{2}\right)$，则 $k_{PF}=\frac{\frac{y_1+y_2}{2}}{\frac{x_1+x_2}{2}+c}=\frac{\frac{y_1+y_2}{2}}{\frac{4c}{3}}=\frac{3(y_1+y_2)}{8c}$.

由直线 l 与 C 相交于 A,B 两点，得 $\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}=1$，$\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}=1$，两式相减得 $\frac{x_1^2-x_2^2}{a^2}+\frac{y_1^2-y_2^2}{b^2}=0$，即 $\frac{(x_1-x_2)(x_1+x_2)}{a^2}+\frac{(y_1-y_2)(y_1+y_2)}{b^2}=0$，所以 $\frac{(y_1-y_2)(y_1+y_2)}{(x_1-x_2)(x_1+x_2)}=-\frac{b^2}{a^2}$，即 $k_l\cdot\frac{y_1+y_2}{x_1+x_2}=-\frac{b^2}{a^2}$.

所以 $k_l=-\frac{b^2}{a^2}\frac{x_1+x_2}{y_1+y_2}=-\frac{b^2}{a^2}\frac{2c}{3(y_1+y_2)}$.

则 $k_l\cdot k_{PF}=-\frac{b^2}{a^2}\frac{2c}{3(y_1+y_2)}\cdot\frac{3(y_1+y_2)}{8c}=-\frac{b^2}{4a^2}=-\frac{3}{16}$，所以 $\frac{b^2}{a^2}=\frac{3}{4}$，故 $e^2=1-\frac{b^2}{a^2}=\frac{1}{4}$，得 $e=\frac{1}{2}$.

例题 13.83 (已知焦点三角形两底角) 已知 F_1,F_2 分别是圆 $D:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左右两个焦点，若在 D 上存在点 P 使 $\angle F_1PF_2=90^\circ$，且满足 $2\angle PF_1F_2=\angle PF_2F_1$，则椭圆的离心率为(　)

A. $\sqrt{3}$　　C. $\frac{\sqrt{3}}{2}$

B. $\sqrt{3}-1$　　D. $\frac{\sqrt{3}}{3}$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

解 答案：B. 在 $\triangle PF_1F_2$ 中 $\angle F_1PF_2=90^\circ$，且满足 $2\angle PF_1F_2=\angle PF_2F_1$，所以 $\angle PF_1F_1=60^\circ$，$\angle PF_1F_2=30^\circ$，所以 $|PF_1|=|F_1F_2|\sin 30^\circ=c,|PF_1|=|F_1F_2|\cos 30^\circ=\sqrt{3}c$，所以 $|PF_1|+|PF_1|=\sqrt{3}c+c=2a$，所以 $e=\frac{c}{a}=\frac{2}{\sqrt{3}+1}=\sqrt{3}-1$；故选：$B$

例题 13.84 (利用渐近线的斜率) （2024·山东淄博二模）若双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的一条渐近线方程为 $y=\frac{1}{2}x$，则离心率 e 为（　）

A. $\frac{\sqrt{3}}{2}$　　C. $\sqrt{3}$

B. $\frac{\sqrt{5}}{2}$　　D. $\sqrt{5}$

解 答案：B. $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 渐近线方程为 $y=\pm\frac{b}{a}x$，则 $\frac{b}{a}=\frac{1}{2}$. 离心率 $e=\frac{c}{a}=\sqrt{\left(\frac{c}{a}\right)^2}=\sqrt{\frac{c^2}{a^2}}=\sqrt{\frac{a^2+b^2}{a^2}}=\sqrt{1+\left(\frac{b}{a}\right)^2}=\sqrt{1+\left(\frac{1}{2}\right)^2}=\frac{\sqrt{5}}{2}$. 故选：B.

例题 13.85 (坐标法) （2024·全国·模拟预测）已知 F_1,F_2 分别是双曲线 $C:\frac{y^2}{a^2}-\frac{x^2}{b^2}=1(a>0,b>0)$ 的上、下焦点，过点 F_2 且与 y 轴垂直的直线与 C 的一条渐近线相交于点 P，且 P 在第四象限，四边形 PF_1QF_2 为平行四边形. 若直线 QF_2 的倾斜角 $\alpha\in[\frac{2\pi}{3},\frac{3\pi}{4}]$，则 C 的离心率的取值范围是__.

解 答案：$[\frac{\sqrt{21}}{3},\sqrt{5}]$. 如图所示，由双曲线的对称性可知 Q 也在双曲线的渐近线上，且 Q 在第二象限，由 $PF_2\perp y$ 轴可知 $QF_1\perp y$ 轴，设 $Q(x_0,c)$，又 Q 在渐近线 $y=-\frac{a}{b}x$ 上，所以 $Q\left(-\frac{bc}{a},c\right)$，则 $\tan\alpha=k_{QF_2}=-\frac{2a}{b}$，因为 $\alpha\in[\frac{2\pi}{3},\frac{3\pi}{4}]$，所以 $\tan\alpha\in[-\sqrt{3},-1]$

例题 13.86 (利用焦半径的取值范围) 已知双曲线 $M:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的左、右焦点分别为 $F_1,F_2,|F_1F_2|=2c$. 若双曲线 M 的右支上存在点 P，使 $\frac{a}{\sin\angle PF_1F_2}=\frac{3c}{\sin\angle PF_2F_1}$，则双曲线 M 的离心率的取值范围为__.

解 答案：$(1,\frac{2+\sqrt{7}}{3})$. 依题意，点 P 在双曲线的右支，P 不与双曲线顶点重合，在 $\triangle PF_1F_2$ 中，由正弦定理得：$\frac{|PF_2|}{\sin\angle PF_1F_2}=\frac{|PF_1|}{\sin\angle PF_2F_1}$，因 $\frac{a}{\sin\angle PF_1F_2}=\frac{3c}{\sin\angle PF_2F_1}$，于是得 $\frac{|PF_2|}{a}=\frac{|PF_1|}{3c}$，而点 P 在双曲线 M 的右支上，即 $|PF_1|-|PF_2|=2a$，从而有 $|PF_2|=\frac{2a^2}{3c-a}$，点 P 在双曲线 M 的右支上运动，并且异于顶点，于是有 $|PF_2|>c-a$，因此 $\frac{2a^2}{3c-a}>c-a$，而 $c>a>0$，整理得 $3c^2-4ac-a^2<0$，即 $3e^2-4e-1<0$，解得 $\frac{2-\sqrt{7}}{3}<e<\frac{2+\sqrt{7}}{3}$，又 $e>1$，故有 $1<e<\frac{2+\sqrt{7}}{3}$. 所以双曲线 M 的离心率的取值范围为 $(1,\frac{2+\sqrt{7}}{3})$. 故答案为：$(1,\frac{2+\sqrt{7}}{3})$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 13.87 (十九·四心问题) (2024·湖北·模拟预测) 斜率为 1 的直线与双曲线 $E:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 交于两点 A,B, 点 C 是 E 上的一点, 满足 $AC\perp BC,\triangle OAC,\triangle OBC$ 的重心分别为 $P,Q,\triangle ABC$ 的外心为 R. 记直线 OP,OQ,OR 的斜率为 k_1,k_2,k_3. 若 $k_1k_2k_3=-27$, 则双曲线 E 的离心率为__.

解 答案: 2. 不妨取 AC,BC 的中点 M,N. 因为 $\triangle OAC$ 的重心为 P, 且 P 在中线 OM 上, 所以 $k_1=k_{OP}=k_{OM},k_2=k_{OQ}=k_{ON}$. 由中点弦结论知, $k_{OM}k_{AC}=k_{OM}k_{BC}=\frac{b^2}{a^2}$、$\therefore k_1k_{AC}=k_2k_{BC}=\frac{b^2}{a^2}$、$\therefore k_1k_{AC}k_2k_{BC}=\left(\frac{b^2}{a^2}\right)^2$, 因为 $AC\perp BC$, 所以 $k_{AC}\cdot k_{BC}=-1$, $\therefore k_1k_2=-\left(\frac{b^2}{a^2}\right)^2$, 又由 $AC\perp BC$, 可得 $\triangle ABC$ 的外心 R 为 AB 的中点, 于是由中点弦结论知 $k_{OR}\cdot k_{AB}=\frac{b^2}{a^2}$, 又 $k_{AB}=1$, 所以 $k_{OR}=\frac{b^2}{a^2}$, 即 $k_3=\frac{b^2}{a^2}$. 由 $k_1k_2k_3=-27$ 得, $-\left(\frac{b^2}{a^2}\right)^3=-27$, 解得 $\frac{b^2}{a^2}=3$, 所以双曲线 E 的离心率 $e=\sqrt{1+\frac{b^2}{a^2}}=2$.

例题 13.88 (平面截圆锥 (丹林球) 问题) "用一个不垂直于圆锥的轴的平面截圆锥, 当圆锥的轴与截面所成的角不同时, 可以得到不同的截口曲线", 利用这个原理, 小强在家里用两个射灯在墙上投影出两个相同的椭圆 (图 1), 射灯射出的光锥视为圆锥且一条母线恰好与墙面垂直. 图 2 是一个射灯投影的直观图, 圆锥 PO 的轴截面 APB 是等边三角形, 椭圆 O_1 所在平面为 α, $PB\perp\alpha$, 则椭圆 O_1 的离心率为__.

解 答案: $\frac{\sqrt{3}}{3}$. 设 $AB=2r$, 由于 $PB\perp\alpha$, 所以 $PB\perp AM$, 在等边三角形 PAB 中, 点 M 为 PB 的中点, 于是 $AM=\sqrt{3}r$, 在平面 α 中, 由椭圆的对称性可知, $AO_1=MO_1=\frac{\sqrt{3}}{2}r$, 连接 OO_1,PO_1, 延长 PO_1 与 AB 交于点 Q, ……

13.4 设点法的技术基础

已在 第13.3节 中建立过联立、韦达定理与切线公式. 从本节起, 我们把注意力转向另一种参量选择, 即设点法.

设点法把曲线上动点的坐标直接设为参数, 以曲线方程为约束, 把几何条件写成关于这些坐标的等式. 它的动机在于减少变量, 让动点始终骑在曲线上, 并把曲线方程当作全局代换式反复使用. 这一视角对应 第19章 中的满足式条件, 即把题设条件理解为对自由度的命名, 参见 第19.4节 的条件分类与 第19.7节 的参数翻译策略.

设线法先引一条截线 $y=kx+m$, 通过联立与韦达定理读出端点对称量; 设点法直接给动点命名, 以曲线方程作整体代换, 把计算留在坐标层面. 选择标

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

准是题面先给点还是先给线: 条件用端点坐标对称式表达, 如中点、定比、垂直 $x_1x_2+y_1y_2=0$ 或斜率积, 选设点法; 条件用直线参数表达, 如截距、斜率、弦长, 或需要整体韦达和积, 选设线法. 这一区分在 第13.5节 中进一步展开.

做法可作如下概括. 设曲线上动点 (x_0,y_0), 以曲线方程为约束, 将几何条件写成关于 x_0,y_0 的代数式并作整体代换. 本节整理其中常用的技术, 供 第13.5节 与 第13.6节 选用.

13.4.1 两点式直线方程的直写与变形

已在 第13.1节 中了解过了直线方程的各形式. 这里换一个角度, 即见点写方程, 再看这些形式的关系, 这一视角在设点法中尤为常用.

设有两点 $P_1(x_1,y_1),P_2(x_2,y_2)$. 当 $x_1\neq x_2$ 且 $y_1\neq y_2$ 时, 两点式

$$\frac{y-y_1}{y_2-y_1}=\frac{x-x_1}{x_2-x_1}$$

可对称地理解为, 直线上任一点 $P(x,y)$ 的坐标, 都可用 P_1,P_2 的坐标线性插值而得. 引入参数 λ 令

$$\frac{x-x_1}{x_2-x_1}=\lambda,\qquad \frac{y-y_1}{y_2-y_1}=\lambda,$$

即得直线的参数式

$$\begin{cases}x=x_1+\lambda(x_2-x_1)\\ y=y_1+\lambda(y_2-y_1)\end{cases}\qquad(\lambda\in\mathbb{R}).$$

这里 $\lambda=0$ 对应 P_1, $\lambda=1$ 对应 P_2, $\lambda=\frac{1}{2}$ 对应中点.

> **注 13.4 (线性插值)**
> 线性插值是由两个已知端点的值按比例还原中间值的方法. 当 λ 从 0 变到 1 时, $x_1+\lambda(x_2-x_1)$ 从 x_1 匀速变到 x_2, y 坐标同理; 横, 纵坐标共用同一个 λ 同步过渡, 点 $P(x,y)$ 便从 P_1 沿线段匀速移动到 P_2.$\lambda<0$ 或 $\lambda>1$ 时, P 落在 P_1P_2 的延长线上. 两点式中两个比值相等, 说的正是横, 纵坐标同步插值这一事实.

13.4.1.0.1 分母为零的退化情形 当 $x_1=x_2$ 时, 直线竖直, 方程直写为 $x=x_1$; 当 $y_1=y_2$ 时, 直线水平, 方程直写为 $y=y_1$. 实战中先判断两端点是否共竖/共横, 再决定是否套用标准两点式, 可避免分母为零的形式错误.

13.4.1.0.2 行列式形式与共线判定 两点式可改写为齐次形式

$$(y-y_1)(x_2-x_1)-(x-x_1)(y_2-y_1)=0,$$

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即

$$\begin{vmatrix} x-x_1 & y-y_1 \\ x_2-x_1 & y_2-y_1 \end{vmatrix} = 0.$$

它与“P, P_1, P_2 三点共线”的行列式判据, 见 第13.4.5小节是同一代数条件, 只是把第三点取为动点 P.

13.4.1.0.3　与点斜式, 斜截式, 截距式的互化　当 $x_2 \neq x_1$ 时, 斜率 $m = \frac{y_2-y_1}{x_2-x_1}$, 两点式化为点斜式 $y-y_1 = m(x-x_1)$; 进一步若直线过 y 轴上点 $(0,b)$, 则化为斜截式 $y = mx+b$. 与截距式 $\frac{x}{a}+\frac{y}{b}=1$ 的互化, 只需取 $P_1=(a,0), P_2=(0,b)$ 代入两点式即可.

三式之间的灵活切换, 在设点法中常常决定计算量, 凡涉及弦端点坐标, 优先用两点式或其参数式; 凡涉及斜率, 退回点斜式; 凡涉及截距, 直接截距式.

例题 13.89　已知直线过 $P_1(2,1)$ 与 $P_2(4,5)$, 分别用两点式, 点斜式, 截距式写出它的方程, 并求其在 x 轴, y 轴上的截距.

解　两点式: $\frac{y-1}{5-1}=\frac{x-2}{4-2}$, 即 $\frac{y-1}{4}=\frac{x-2}{2}$.

斜率 $m=\frac{5-1}{4-2}=2$, 故点斜式 $y-1=2(x-2)$, 化为斜截式 $y=2x-3$.

令 $y=0$ 得 $x=\frac{3}{2}$; 令 $x=0$ 得 $y=-3$. 故 x 截距 $a=\frac{3}{2}$, y 截距 $b=-3$, 截距式 $\frac{x}{3/2}+\frac{y}{-3}=1$, 即 $\frac{2x}{3}-\frac{y}{3}=1$.

三式互化验证: 两点式 $\frac{y-1}{4}=\frac{x-2}{2} \Rightarrow y-1=2(x-2) \Rightarrow y=2x-3$, 与点斜式, 斜截式一致.

13.4.2　圆锥曲线方程的结构认识

标准方程(椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, 双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, 抛物线 $y^2=2px$)的推导见第13.3节. 这里把它们当作曲线上点必须满足的约束, 在设点时作整体代换, 而不再逐项展开.

三类标准方程的代数骨架可对照如下:

- **椭圆**: $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, 和为一, 两项同号, 界定有界闭曲线, $|x|\le a, |y|\le b$;
- **双曲线**: $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, 差为一, 两项异号, 界定两支无界曲线, $|x|\ge a$(以 x 轴焦点型为例);
- **抛物线**: $y^2=2px$, 一平一次, 二次项与一次项平衡, 界定开口无界曲线.

统一二次型观点下, 它们都是 $Ax^2+Bxy+Cy^2+Dx+Ey+F=0$ 经旋转平移消交叉项与部分线性项后的标准形(判别式 $\Delta=B^2-4AC$ 的符号分类见 第13.3节).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当动点 $P(x_0,y_0)$ 在曲线上时，曲线方程成为可整体代换的恒等式. 例如椭圆上：

$$\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}=1 \quad\Longrightarrow\quad b^2x_0^2+a^2y_0^2=a^2b^2,\quad \frac{y_0^2}{b^2}=1-\frac{x_0^2}{a^2}.$$

这种方程作约束，整体代换的意识，是设点法区别于设线法的根本. 设线法联立直线 $y=kx+m$ 与曲线，通过韦达定理读出 x_1+x_2,x_1x_2 等对称量；而设点法让动点坐标 (x_0,y_0) 自始至终骑在曲线上，每当出现 x_0,y_0 的二次组合，就以曲线方程作一次整体约简.

例题 13.90 设 $P(x_0,y_0)$ 在椭圆 $\frac{x^2}{9}+\frac{y^2}{4}=1$ 上，求 $b^2x_0^2+a^2y_0^2$ 之值；又问 $4x_0^2-9y_0^2$ 是否为定值.

解 此处 $a^2=9,b^2=4$. 由 P 在椭圆上，$\frac{x_0^2}{9}+\frac{y_0^2}{4}=1$，两边乘 36 得 $4x_0^2+9y_0^2=36$.

第一式：$b^2x_0^2+a^2y_0^2=4x_0^2+9y_0^2=36$（整体代换，不必求 x_0,y_0）.

第二式：由 $\frac{y_0^2}{4}=1-\frac{x_0^2}{9}$ 得 $y_0^2=\frac{4}{9}(9-x_0^2)$. 故

$$4x_0^2-9y_0^2=4x_0^2-9\cdot\frac{4}{9}(9-x_0^2)=4x_0^2-4(9-x_0^2)=8x_0^2-36.$$

此值随 x_0 变化，不是定值；但若附加条件，如 x_0 给定便可定值. 此例说明，整体代换遇到差型组合时，往往需配合消元.

13.4.3 点差法的原理与公式推导

点差法是设点法处理弦中点问题的主要工具. 相关结论已经在 第13.3节 的二次曲线语境中见过；这里给出完整的两式相减、因式分解、代换中点推导，并明确适用条件与推广.

定理 13.17 (椭圆的弦中点斜率关系) 设椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上两点 $P_1(x_1,y_1)$，$P_2(x_2,y_2)$ $(x_1\neq x_2)$，弦 P_1P_2 的中点为 $M(x_0,y_0)$，斜率

$$k=\frac{y_2-y_1}{x_2-x_1}.$$

则

$$k=-\frac{b^2x_0}{a^2y_0}\quad(y_0\neq0),$$

或等价地 $k\cdot k_{OM}=-\frac{b^2}{a^2}$，其中 $k_{OM}=\frac{y_0}{x_0}$ $(x_0\neq0)$ 为中心 O 到中点 M 的斜率.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

证明 因 P_1, P_2 在椭圆上,

$$\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}=1, \qquad \frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}=1.$$

两式相减:

$$\frac{x_1^2-x_2^2}{a^2}+\frac{y_1^2-y_2^2}{b^2}=0,$$

即

$$\frac{(x_1-x_2)(x_1+x_2)}{a^2}+\frac{(y_1-y_2)(y_1+y_2)}{b^2}=0.$$

由中点公式 $x_1+x_2=2x_0, y_1+y_2=2y_0$, 且 $x_1\neq x_2$, 两边除以 x_1-x_2 得

$$\frac{2x_0}{a^2}+\frac{y_1-y_2}{x_1-x_2}\cdot\frac{2y_0}{b^2}=0.$$

注意 $\frac{y_1-y_2}{x_1-x_2}=\frac{y_2-y_1}{x_2-x_1}=k$, 故

$$\frac{x_0}{a^2}+k\cdot\frac{y_0}{b^2}=0 \quad\Longrightarrow\quad k=-\frac{b^2x_0}{a^2y_0} \quad (y_0\neq 0).$$

当 $x_0\neq 0$ 时, $k_{OM}=\frac{y_0}{x_0}$, 于是 $k\cdot k_{OM}=-\frac{b^2}{a^2}$. □

13.4.3.0.1 适用条件

- 推导中约去了 x_1-x_2, 故要求弦非竖直 ($x_1\neq x_2$). 当弦竖直时, k 不存在, 中点 $M=(x_1,0)$ (两端关于 x 轴对称), 应直接用几何结论, 不必强行套用斜率公式.
- 当 $y_0=0$ 时, 中点在 x 轴上, 由 $\frac{x_0}{a^2}+k\cdot\frac{y_0}{b^2}=0$ 得 $x_0=0$, 即弦的中点为原点; 实战中 $y_0=0$ 且 $x_0\neq 0$ 时弦必竖直 ($x_1=x_2$, 斜率不存在).
- 中点 M 必须落在椭圆内部 (否则过 M 给定斜率的直线截不到实弦), 即 $\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}<1$.

13.4.3.0.2 对称差分写法 只关心中点时, 可对称地设 $M(x_0,y_0)$ 为中点, 两端点为 $P_1(x_0+u,y_0+v), P_2(x_0-u,y_0-v)$. 代入椭圆方程并相减, 平方差直接给出

$$\frac{x_0u}{a^2}+\frac{y_0v}{b^2}=0, \qquad \text{即} \qquad k=\frac{v}{u}=-\frac{b^2x_0}{a^2y_0} \quad (u\neq 0).$$

这一写法省去四个端点坐标, 中点显式可见, 在 第13.5节 的双动点设而不求中尤为顺手.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.4.3.0.3 双曲线与抛物线的推广 对双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, 同样的相减推导给出

$$k=\frac{b^2x_0}{a^2y_0}\quad(y_0\neq 0),\qquad \text{即}\qquad k\cdot k_{OM}=\frac{b^2}{a^2}.$$

对抛物线 $y^2=2px$, 设弦两端 $P_1(x_1,y_1),P_2(x_2,y_2)$, 中点 $M(x_0,y_0)$, 两式相减得

$$y_1^2-y_2^2=2p(x_1-x_2)\quad\Longrightarrow\quad(y_1-y_2)(y_1+y_2)=2p(x_1-x_2).$$

当 $x_1\neq x_2$ 时, $k=\frac{y_2-y_1}{x_2-x_1}$ 满足 $k\cdot 2y_0=2p$, 即 $k=\frac{p}{y_0}$ $(y_0\neq 0)$. 这正是抛物线弦中点斜率公式.

例题 13.91 求椭圆 $\frac{x^2}{16}+\frac{y^2}{9}=1$ 中, 过点 $M(2,1)$ 且以 M 为中点的弦所在直线方程.

解 先检验 M 在椭圆内部: $\frac{4}{16}+\frac{1}{9}=\frac{1}{4}+\frac{1}{9}=\frac{13}{36}<1$, 成立.

由弦中点斜率关系 ($a^2=16,b^2=9,x_0=2,y_0=1$):

$$k=-\frac{b^2x_0}{a^2y_0}=-\frac{9\cdot 2}{16\cdot 1}=-\frac{9}{8}.$$

故所求弦的方程为

$$y-1=-\frac{9}{8}(x-2),\quad\text{即}\quad 9x+8y-26=0.$$

复核: 联立 $y=-\frac{9}{8}(x-2)+1$ 与椭圆, 所得二次方程的两根之和为 4, 即 $2x_0$, 与中点横坐标一致, 从略.

注 13.5
点差法给出的是**中点** ($\lambda=\frac{1}{2}$) 情形. 若弦上一点分弦成定比 λ, 中点公式不再适用, 但两式相减、因式分解的代数骨架仍然成立. 这一推广(定比点差法)在 第13.5节 中系统展开.

13.4.4 面积的行列式表达与坐标运算

三角形的面积, 已在 第13.3节 中以鞋带公式简短见过. 这里用二维向量叉乘重新推导并推广, 因为叉乘观点是后续多边形面积与 第15.3节 中极点/极线讨论的共同入口.

设有向向量 $\vec{u}=(u_1,u_2),\vec{v}=(v_1,v_2)$. 定义二阶行列式

$$\det(\vec{u},\vec{v})=\begin{vmatrix}u_1 & v_1\\ u_2 & v_2\end{vmatrix}=u_1v_2-u_2v_1.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

它是平面叉乘 $u_1v_2-u_2v_1$, 几何意义是 $|\det(\vec{u},\vec{v})|$ 等于 $\vec{u},\vec{v}$ 张成的平行四边形面积, 符号反映绕向(逆时针为正).

定理 13.18 (三角形面积的行列式公式) 设平面上三点 P_1,P_2,P_3 的坐标分别为 (x_1,y_1)、(x_2,y_2)、(x_3,y_3). 则 $\triangle P_1P_2P_3$ 的有向面积为

$$\begin{aligned} S_{\to} &= \frac{1}{2}\det(\overrightarrow{P_1P_2},\overrightarrow{P_1P_3}) \\ &= \frac{1}{2}[(x_2-x_1)(y_3-y_1)-(y_2-y_1)(x_3-x_1)], \end{aligned}$$

其通常(无向)面积为 $|S_{\to}|$. 等价地, 展开后得对称的**鞋带公式**

$$S=\frac{1}{2}\left|x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)\right|.$$

知乎@余命数

证明 以 P_1 为简化原点记 $\vec{p}=\overrightarrow{P_1P_2}=(x_2-x_1,y_2-y_1)$,$\vec{q}=\overrightarrow{P_1P_3}=(x_3-x_1,y_3-y_1)$. 两向量张成的平行四边形面积为 $|\vec{p}|\,|\vec{q}|\sin\theta$. 由平面叉乘的几何意义

$$\det(\vec{p},\vec{q})=|\vec{p}|\,|\vec{q}|\sin\theta,$$

故平行四边形面积为 $|\det(\vec{p},\vec{q})|$, 三角形为其一半:

$$S=\frac{1}{2}|\det(\vec{p},\vec{q})|=\frac{1}{2}|(x_2-x_1)(y_3-y_1)-(y_2-y_1)(x_3-x_1)|.$$

知乎@余命数

去绝对值的代数符号取决于 P_1,P_2,P_3 的绕向(逆时针为正), 这正是 $S_{\to}$ 的定义.

为得鞋带公式, 展开行列式:

$$\begin{aligned} &(x_2-x_1)(y_3-y_1)-(y_2-y_1)(x_3-x_1) \\ &=x_2y_3-x_2y_1-x_1y_3+x_1y_1-(y_2x_3-y_2x_1-y_1x_3+y_1x_1) \\ &=x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2). \end{aligned}$$

取绝对值即得鞋带公式. □

知乎@余命数

13.4.4.0.1　原点特款　若三角形有一顶点在原点 O, 另两顶点为 $A(x_1,y_1)$ 与 $B(x_2,y_2)$, 则

知乎@余命数

$$S_{\triangle OAB}=\frac{1}{2}|x_1y_2-x_2y_1|.$$

这是设点法中原点与两动点三角形问题的标准入口, 见 第13.6节.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.4.4.0.2 四边形分割 对四边形 $ABCD$，可沿对角线 AC 分割为 $\triangle ABC$ 与 $\triangle ADC$，面积相加；或直接用鞋带公式推广到 n 边形（顶点逆时针排列 $P_1,\dots,P_n$，$P_{n+1}=P_1$）：

$$S=\frac{1}{2}\left|\sum_{i=1}^{n}(x_iy_{i+1}-y_ix_{i+1})\right|.$$

鞋带之名正来自交替相乘后求和，如鞋带交叉穿引.

知乎@余命数

例题 13.92 设 $A(2,0)$、$B(0,1)$、$C(-2,0)$ 为椭圆 $\frac{x^2}{4}+y^2=1$ 上的三点，用行列式公式求 $\triangle ABC$ 的面积.

知乎@余命数

解 由鞋带公式

$$S=\frac{1}{2}|x_A(y_B-y_C)+x_B(y_C-y_A)+x_C(y_A-y_B)|.$$

代入 $A=(2,0),B=(0,1),C=(-2,0)$，得

$$S=\frac{1}{2}|2\cdot(1-0)+0\cdot(0-0)+(-2)\cdot(0-1)|=\frac{1}{2}|2+0+2|=2.$$

有向行列式为正，说明顺序 $A\to B\to C$ 为逆时针. 几何上视 AC 为底（长 4，落在 x 轴上），B 到 x 轴高为 1，直观面积 $\frac{1}{2}\cdot 4\cdot 1=2$，与行列式法一致.

知乎@余命数

13.4.5 三点共线的坐标刻画

已在 第11.2节 中了解过三点共线的向量刻画与定比分点公式. 在坐标下它们有四种等价写法，证明后可按条件选用.

定理 13.19 (三点共线的四种坐标刻画) 设 $P_1(x_1,y_1),P_2(x_2,y_2),P_3(x_3,y_3)$ 互异，则下列四条件等价：

1. **斜率相等**：当 $x_2\neq x_1,x_3\neq x_1$ 时，$\frac{y_2-y_1}{x_2-x_1}=\frac{y_3-y_1}{x_3-x_1}$；
2. **向量共线**：$\det(\overrightarrow{P_1P_2},\overrightarrow{P_1P_3})=0$，即 $(x_2-x_1)(y_3-y_1)-(y_2-y_1)(x_3-x_1)=0$；
3. **行列式为零**：$\begin{vmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ x_3 & y_3 & 1\end{vmatrix}=0$；
4. **定比分点**：存在 $\lambda\neq -1$ 使 P_2 分 P_1P_3 成定比 λ（即 $\overrightarrow{P_1P_2}=\lambda\overrightarrow{P_2P_3}$），坐标满足 $x_2=\frac{x_1+\lambda x_3}{1+\lambda}$，$y_2=\frac{y_1+\lambda y_3}{1+\lambda}$.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

证明 (1)⇔(2)：当 $x_2\neq x_1, x_3\neq x_1$ 时，斜率相等即 $\frac{y_2-y_1}{x_2-x_1}=\frac{y_3-y_1}{x_3-x_1}$，交叉相乘恰为向量共线条件. 若某分母为零（竖直弦），则三横坐标相等，向量共线条件自动成立.

(2)⇔(3)：展开向量共线条件，与面积定理证明中的展开相同，得

$$x_1(y_2-y_3)+x_2(y_3-y_1)+x_3(y_1-y_2)=0,$$

这正是三阶行列式为零.

(2)⇔(4)：向量共线意味着存在实数 λ 使 $\overrightarrow{P_1P_2}=\lambda\overrightarrow{P_2P_3}$（当 P_2 不在 P_1P_3 的反向延长线上时 $\lambda\neq -1$ 可保证分点在直线上且公式有意义）. 写成分量即定比分点坐标公式. 反之，定比分点公式直接给出 $\overrightarrow{P_1P_2}=\lambda\overrightarrow{P_2P_3}$，故向量共线. □

已知斜率或易求斜率时用 (1)；涉及面积或叉乘时用 (2)，同时 $S=0$ 即共线；需要齐次或对称书写时用 (3)；题中显含分比 λ 或中点时用 (4).

定比分点公式的两个常用变形：$\lambda=1$ 即中点；给定 P 与 P_1 反解 P_2：$x_2=\frac{(1+\lambda)x_P-x_1}{\lambda}$. 后者是 第13.5节 定比点差法的入口.

例题 13.93　抛物线 $C: y^2=4x$，焦点 $F(1,0)$. 直线 l 过 F 与 C 交于 A, B，设 $|AB|=8$ 且 l 不垂直于 x 轴.

1. 求 l 的方程；
2. 设 P 在 AB 上满足 $\overrightarrow{AP}=3\overrightarrow{PB}$，求 $\triangle OAP$ 的面积，其中 O 为原点.

解 (1) 设 $l: y=k(x-1)$，与 $y^2=4x$ 联立消 y：

$$k^2(x-1)^2=4x\Longrightarrow k^2x^2-(2k^2+4)x+k^2=0.$$

根 x_1, x_2 之和 $\frac{2k^2+4}{k^2}=2+\frac{4}{k^2}$，积 1. 由 第13.3节 弦长公式

$$|AB|=\sqrt{1+k^2}\,|x_1-x_2|=\sqrt{1+k^2}\sqrt{\left(2+\frac{4}{k^2}\right)^2-4}=\frac{4(1+k^2)}{k^2}.$$

令 $|AB|=8$：$\frac{4(1+k^2)}{k^2}=8\Rightarrow k^2=1$. 取 $k=1$，即 $l: y=x-1$（$k=-1$ 同理，见 (2) 末）.

(2) 此时 A, B 的横坐标满足 $x^2-6x+1=0$，解得 $x=3\pm2\sqrt{2}$，对应 $y=2\pm2\sqrt{2}$. 设

$$A(3-2\sqrt{2},\,2-2\sqrt{2}),\qquad B(3+2\sqrt{2},\,2+2\sqrt{2}).$$

由 $\overrightarrow{AP}=3\overrightarrow{PB}$（定比 $\lambda=3$），

$$x_P=\frac{x_A+3x_B}{4},\qquad y_P=\frac{y_A+3y_B}{4}.$$

代入得 $x_P=3+\sqrt{2}$，$y_P=2+\sqrt{2}$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

对 $\triangle OAP$ 用原点特款 $S=\frac{1}{2}|x_Ay_P-x_Py_A|$:

$$x_Ay_P=(3-2\sqrt{2})(2+\sqrt{2})=2-\sqrt{2},$$
$$x_Py_A=(3+\sqrt{2})(2-2\sqrt{2})=2-4\sqrt{2},$$

故

$$S=\frac{1}{2}|(2-\sqrt{2})-(2-4\sqrt{2})|=\frac{1}{2}\cdot 3\sqrt{2}=\frac{3\sqrt{2}}{2}.$$

若取 $k=-1$, 由关于 x 轴的对称性, 同样得 $S=\frac{3\sqrt{2}}{2}$.

13.4.6 切线方程与切点弦方程

椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 在其上一点 $P_0(x_0,y_0)$ 处的切线方程已在 第13.3节 中推导过. 把椭圆方程中的一个 x 和一个 y 替换为切点的 x_0,y_0. 我们称这一做法为**替换法则**.

$$\text{椭圆:}\quad \frac{xx_0}{a^2}+\frac{yy_0}{b^2}=1,$$
$$\text{双曲线:}\quad \frac{xx_0}{a^2}-\frac{yy_0}{b^2}=1,$$
$$\text{抛物线}y^2=2px:\quad y\,y_0=p(x+x_0).$$

抛物线情形将 y^2 替换为 $y\,y_0$, $2x$ 替换为 $x+x_0$(平方项变相乘, 一次项变平均), 与前两类曲线遵守同一法则.

定理 13.20 (切点弦方程) 设 $T(x_T,y_T)$ 为圆锥曲线 $\mathcal{C}$ 外一点, 自 T 作 $\mathcal{C}$ 的两条切线分别切于 P_1,P_2, 则这两切点的连线 P_1P_2, 称为 T 关于 $\mathcal{C}$ 的**切点弦**的方程仍取替换法则形式. 具体地, 对于

$$\text{椭圆}\frac{x^2}{a^2}+\frac{y^2}{b^2}=1:\quad \frac{xx_T}{a^2}+\frac{yy_T}{b^2}=1;$$

$$\text{双曲线}\frac{x^2}{a^2}-\frac{y^2}{b^2}=1:\quad \frac{xx_T}{a^2}-\frac{yy_T}{b^2}=1;$$

$$\text{抛物线}y^2=2px:\quad y\,y_T=p(x+x_T).$$

证明 以椭圆为例. 设 $P_1=(x_1,y_1),P_2=(x_2,y_2)$ 均在椭圆上. 由替换法则, 过 P_1 的切线为 $\frac{x_1x}{a^2}+\frac{y_1y}{b^2}=1$. 该切线过外点 T, 故

$$\frac{x_1x_T}{a^2}+\frac{y_1y_T}{b^2}=1.$$

知乎@余命数

同理, 过 P_2 的切线亦过 T:

$$\frac{x_2x_T}{a^2}+\frac{y_2y_T}{b^2}=1.$$

这说明 P_1,P_2 都满足同一个一次方程

$$\frac{x_Tx}{a^2}+\frac{y_Ty}{b^2}=1.$$

此方程表示一条直线, 且经过 P_1 和 P_2, 故必为弦 P_1P_2 所在直线. 双曲线与抛物线情形的证明完全平行, 从略. □

注 13.6 (同构萌芽)
切点弦方程与切线方程*形式上完全相同*, 把曲线方程替换一遍代入 (x_T,y_T) 即可. 当 T 在曲线上时, 该方程退化为切线(单点弦); 当 T 在曲线外时, 表示切点弦. 这一同构性源于二次曲线方程的代数对称性. 当 T 在曲线内部时, 同一形式仍定义一条由 T 与曲线纯代数规定的直线, 称为 T 的**极线**. 极点与极线以及它们同切点弦、切线的统一, 见 第15.3节.

例题 13.94　椭圆 $C:\frac{x^2}{4}+y^2=1$ 外一点 $T(0,2)$.

1. 用替换法则求 T 关于 C 的切点弦方程;
2. 验证切点弦两端点处的两条切线确实过 T.

解　(1) 由切点弦方程(椭圆), $a^2=4,b^2=1,T=(0,2)$:

$$\frac{0\cdot x}{4}+\frac{2\cdot y}{1}=1,\quad 即\quad y=\frac{1}{2}.$$

故切点弦 $\ell:y=\frac{1}{2}$.

(2) 联立 ℓ 与 C: $\frac{x^2}{4}+\frac{1}{4}=1\Rightarrow x^2=3\Rightarrow x=\pm\sqrt{3}$. 切点为

$$P_1(\sqrt{3},\tfrac{1}{2}),\qquad P_2(-\sqrt{3},\tfrac{1}{2}).$$

过 P_1 的切线: $\frac{\sqrt{3}x}{4}+\frac{y}{2}=1$, 代入 $T(0,2)$: $0+1=1$, 通过. 过 P_2 的切线: $\frac{-\sqrt{3}x}{4}+\frac{y}{2}=1$, 代入 T: 亦过 T. 两切线均过 T, 与切点弦由 T 经替换法则唯一确定这一结论一致.

13.4.7 比例性质在坐标关系中的运用

合分比与等比是初等数学最基本的代数工具. 若 $\frac{a}{b}=\frac{c}{d}$, 则 $\frac{a+b}{b}=\frac{c+d}{d}$ 称为合比, $\frac{a-b}{b}=\frac{c-d}{d}$ 称为分比, $\frac{a}{a+b}=\frac{c}{c+d}$ 称为反合比. 在坐标关系处理中, 这些性质常把韦达式的和差积比重组为新的可用形式.

13.4.7.0.1　合分比的小型化简　设直线 $y=kx+m$ 与圆锥曲线交于 A,B, 韦达定理 (第13.3节) 给出 x_1+x_2,x_1x_2. 但许多题型需要 x_1 与 x_2 的**比值**类信息, 如 $\frac{AF}{FB}$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

或 $\overrightarrow{AP}=\lambda\overrightarrow{PB}$ 中的 λ. 此时可设 $t=\frac{x_1}{x_2}$, 则 $x_1=t\,x_2$. 由和 $x_2(1+t)=-\frac{\beta}{\alpha}$, 由积 $t\,x_2^2=\frac{\gamma}{\alpha}$. 两式消去 x_2, 得

$$\frac{(1+t)^2}{t}=\frac{\beta^2}{\alpha\gamma}.$$

这种把 x_1,x_2 比值约束嵌入韦达和积的写法, 是后韦达, 见 第13.9节的常用工艺.

13.4.7.0.2 等比的应用 等比性质 $\frac{a_1+\cdots+a_k}{b_1+\cdots+b_k}=r$（当各 $\frac{a_i}{b_i}=r$ 同号成立时）在涉及多个分式同结构时立刻化简. 例: 若 k 个椭圆上的点 $P_i=(x_i,y_i)$ 均满足 $\frac{x_i}{a}+\frac{y_i}{b}=\frac{1}{2}$, 全部相加得 $\frac{\sum x_i}{a}+\frac{\sum y_i}{b}=\frac{k}{2}$, 一次即得整体坐标和的关系.

例题 13.95 直线 $y=x-1$ 截抛物线 $y^2=4x$ 于 A,B, F 为焦点 $F(1,0)$. 求 $\frac{AF}{FB}$(带符号的有向分比), 并用韦达 + 合分比复核对结果.

解 联立 $y=x-1$ 与 $y^2=4x$: $(x-1)^2=4x$, 即 $x^2-6x+1=0$.A,B 的横坐标 x_1,x_2 满足 $x_1x_2=1$ 且 $x_1+x_2=6$. 取 A 对应较小根 $x_1=3-2\sqrt{2}$, B 对应较大根 $x_2=3+2\sqrt{2}$. 因 $x_F=1\in(x_1,x_2)$, 有向分比

$$\frac{AF}{FB}=\frac{1-x_1}{x_2-1}=\frac{1-(3-2\sqrt{2})}{(3+2\sqrt{2})-1}=\frac{-2+2\sqrt{2}}{2+2\sqrt{2}}=\frac{\sqrt{2}-1}{\sqrt{2}+1}=3-2\sqrt{2}$$

(最后一步分母有理化: $\frac{(\sqrt{2}-1)^2}{2-1}=3-2\sqrt{2}$).

$x_1x_2=1\Rightarrow x_2=\frac{1}{x_1}$. 故

$$\frac{1-x_1}{x_2-1}=\frac{1-x_1}{\frac{1}{x_1}-1}=\frac{(1-x_1)\,x_1}{1-x_1}=x_1=3-2\sqrt{2},$$

与直接代入法一致. 此处 $x_1x_2=1$ 这一韦达积关系, 把比值问题瞬间归于单根 x_1 的代换, 正是合分比性质嵌入韦达的结果.

13.4.8 本节习题

13.7. 椭圆 $\frac{x^2}{9}+\frac{y^2}{4}=1$ 上一点 P 满足 $x_P+y_P=2$, 求整体代换量 $4x_P^2+9y_P^2$ 之值.(提示: 由曲线方程先得 $4x_P^2+9y_P^2=36$.)

13.8. 写出过 $P_1(-1,2)$ 与 $P_2(3,-1)$ 的直线的两点式, 点斜式与截距式三种方程, 并写出其参数式, 约定 $\lambda=0$ 在 P_1, $\lambda=1$ 在 P_2.

13.9. 求椭圆 $\frac{x^2}{4}+y^2=1$ 中, 斜率为 $\frac{1}{2}$ 的弦的中点轨迹, 给出其方程与取值范围.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.10.　用行列式公式（鞋带公式）求三点 $(2,3),(-1,5),(4,-1)$ 所围三角形的面积，并据有向面积符号判定三点绕向.

13.11.　椭圆 $C:\frac{x^2}{4}+y^2=1$ 外一点 $T(4,0)$，用替换法则求 T 关于 C 的切点弦方程，并求切点弦与 C 的两公共点坐标.

13.12.　直线 $y=k(x-1)$ 截抛物线 $y^2=4x$ 于 A,B，$F(1,0)$ 为焦点. 已知 $x_1x_2=1$（韦达积），用合分比性质证明：$\frac{AF}{FB}=\frac{1-x_1}{x_2-1}$ 可仅由 x_1 表出，即与 x_2 通过 $x_1x_2=1$ 联系，并对 $k=1$ 给出具体数值.

13.13.　椭圆 $\frac{x^2}{25}+\frac{y^2}{16}=1$ 上过内点 $M(1,2)$ 且以 M 为中点的弦，求该弦所在直线的斜率与方程.

13.14.　设 $A(x_1,y_1),B(x_2,y_2)$ 为抛物线 $y^2=4x$ 上两点，O 为原点. 用原点特款面积公式写出 $S_{\triangle OAB}$ 关于 x_1,y_1,x_2,y_2 的表达式；若 AB 中点为 $(3,1)$，求 $S_{\triangle OAB}$ 是否可由此唯一确定，并说明理由.

13.5 设点法的参量选择

参量选取直接决定后续计算繁简. 第13.4节 已介绍设点法的基本技术，以曲线上动点坐标 (x_0,y_0) 为参数，以曲线方程作整体约束，把几何条件直译为关于这些坐标的代数等式. 本节讨论四种常见策略：单动点参数法、双动点参数法、定比点差法、旋转设点法. 这四种选择都对应 第19章 中的**满足式条件**，即把题设条件理解为对自由度的命名，用参数把约束直译成方程，参见 第19.4节 与 第19.7节. 这些选择与具体题型的对应在 第13.6节 中展开.

判断标准只看题面条件. 只有一个自由动点时，用单动点参数法；两个独立动点，如弦的两端时，用双动点参数法，即设而不求，或改用设线法；弦上一点分弦成定比 λ（含中点 $\lambda=1$）时，用定比点差法；显含旋转、等长垂直弦或定斜率积时，用旋转设点法.

13.5.1 单动点参数法

凡题中只有一个自由动点，其余对象由它与曲线的几何关系唯一决定，一律设该点坐标为参量. 曲线方程消掉一个自由度，所余问题往往是单变量函数的最值或定值.

13.5.1.0.1　动点在曲线上　设动点 $P(x_0,y_0)$ 在曲线上并写出约束 $F(x_0,y_0)=0$；把切线、距离、面积、夹角等条件写成关于 x_0,y_0 的代数式；用曲线方程整体代换，把二元目标化为一元函数 $f(t)$，再求导、配方或用基本不等式. 若目标含斜率 $k=\frac{y_0}{x_0}$，可令 $t=\frac{y_0}{x_0}$ 并配合齐次结构；定义域须避开 $x_0=0$ 或 $y_0=0$ 等退化点.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 13.96 (切线截距) 椭圆 $\frac{x^2}{4}+y^2=1$ 上一点 $P(x_0,y_0)$ $(x_0y_0\neq 0)$, 过 P 作椭圆的切线, 分别交 x 轴、y 轴于 A,B. 求 $|AB|$ 的最小值.

解

由 第13.3节 切线公式(替换法则), 切线为

$$\frac{x\,x_0}{4}+y\,y_0=1.$$

令 $y=0$ 得 $A=(\frac{4}{x_0},0)$; 令 $x=0$ 得 $B=(0,\frac{1}{y_0})$. 故

$$|AB|^2=\frac{16}{x_0^2}+\frac{1}{y_0^2}.$$

由 P 在椭圆上, $y_0^2=1-\frac{x_0^2}{4}$. 令 $t=x_0^2\in(0,4)$, 则

$$f(t)=\frac{16}{t}+\frac{4}{4-t},\qquad t\in(0,4).$$

$f'(t)=-\frac{16}{t^2}+\frac{4}{(4-t)^2}$. 令 $f'(t)=0$ 得 $\frac{t}{4-t}=2$ (取正根), 即 $t=\frac{8}{3}$. 又 $f''(t)=\frac{32}{t^3}+\frac{8}{(4-t)^3}>0$, 故 $t=\frac{8}{3}$ 为唯一极小点, 且

$$f\left(\tfrac{8}{3}\right)=6+3=9.$$

因此 $|AB|_{\min}=3$. 取最小值时 $x_0^2=\frac{8}{3}$, $y_0^2=\frac{1}{3}$, 切点为 $(\pm\frac{2\sqrt{6}}{3},\pm\frac{\sqrt{3}}{3})$ 四点之一.

13.5.1.0.2 动点不在曲线上 有时唯一的自由动点并不在目标曲线上, 而是相对于曲线的位置已定, 如定点在圆内或椭圆外. 此时仍把真正变化的曲线上点 P 设为参量, 把定点到 P 的距离、夹角、面积等写成 P 的函数, 再用曲线方程消元. 几何上, 这相当于以定点为观测点, 考察曲线上点的参数化像.

例题 13.97 (圆内定点对动点的张角) 圆 $x^2+y^2=r^2$ 上动点 $P(x_0,y_0)$, 定点 $A(a,0)$ $(0<a<r)$. 设 $\angle OAP=\theta$, 其中 O 为原点. 用单动点参数表示 $\cos\theta$, 并指出 θ 的取值范围思路.

解 向量 $\overrightarrow{AO}=(-a,0)$, $\overrightarrow{AP}=(x_0-a,y_0)$. 由夹角公式

$$\cos\theta=\frac{\overrightarrow{AO}\cdot\overrightarrow{AP}}{|\overrightarrow{AO}|\,|\overrightarrow{AP}|}=\frac{a-x_0}{\sqrt{(x_0-a)^2+y_0^2}}.$$

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

又 $x_0^2+y_0^2=r^2$, 故 $(x_0-a)^2+y_0^2=r^2-2ax_0+a^2$. 令 $t=x_0\in[-r,r]$, 则

$$\cos\theta=\frac{a-t}{\sqrt{r^2-2at+a^2}}.$$

这是 t 的单变量函数. 对 θ 的范围, 只需研究上式在 $[-r,r]$ 的最小值; θ 最大时 $\cos\theta$ 最小, 发生在 $t=r$ 或 $t=-r$ 处的端点. 关键一步是把夹角条件写成内积, 再用圆方程把角度问题化为一元函数.

13.5.1.0.3　单动点策略的边界　最佳适用面是一自由动点加切线、距离、面积、夹角条件. 若出现两个独立动点, 强行只设一个会丢失自由度, 应改用双动点或设线法; 若条件显含旋转对称, 应改用旋转设点法.

13.5.2 双动点参数法

知乎@余命数

题中出现两个独立动点(典型如弦 AB 的两端)时, 单点参量已不足. 双动点参数法同时给两个端点命名, 再**设而不求**, 即不分别求 x_1,y_1,x_2,y_2, 而是用相减、相加提取对称量 $x_1\pm x_2$、$y_1\pm y_2$、x_1x_2、y_1y_2. 其基本骨架在 第13.4节 中系统推导过; 本节强调对称差分的写法.

13.5.2.0.1　操作骨架

1. 设 $A(x_1,y_1),B(x_2,y_2)$ 在曲线上, 写出 $F(x_1,y_1)=0, F(x_2,y_2)=0$;
2. **相减**, 得 $(x_1-x_2),(y_1-y_2)$ 的因式, 常引出斜率 $k=\frac{y_2-y_1}{x_2-x_1}$ 与中点关系;
3. **相加**, 得 x_1+x_2、y_1+y_2 等和式, 便于写中点或垂直条件 $x_1x_2+y_1y_2=0$;
4. 把几何条件写成对称式, 整体消元.

当 $M(x_0,y_0)$ 是弦 AB 的中点时, 最省的对称差分写法是

$$A=(x_0+u,\,y_0+v),\qquad B=(x_0-u,\,y_0-v).$$

弦长 $|AB|=2\sqrt{u^2+v^2}$, 斜率 $k=\frac{v}{u}$, 中点显式可见. 第13.4节 的点差法正是这种对称差分在 $\lambda=1$ 时的特例.

例题 13.98 (中点与斜率) 椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上弦 AB 的中点为 $M(x_0,y_0)$. 用对称差分求弦 AB 的斜率 ($y_0\neq0$).

解 设

$$A=(x_0+u,\,y_0+v),\qquad B=(x_0-u,\,y_0-v),$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其中 $u \neq 0$. 两端点都在椭圆上.

$$\frac{(x_0+u)^2}{a^2}+\frac{(y_0+v)^2}{b^2}=1, \qquad \frac{(x_0-u)^2}{a^2}+\frac{(y_0-v)^2}{b^2}=1.$$

两式相减,

$$\frac{4x_0u}{a^2}+\frac{4y_0v}{b^2}=0 \quad \Longrightarrow \quad \frac{x_0u}{a^2}+\frac{y_0v}{b^2}=0.$$

由 $k=\frac{v}{u}$ 得

$$k=-\frac{b^2x_0}{a^2y_0} \qquad (y_0\neq 0).$$

这就是 第13.4节 中的弦中点斜率关系; 对称差分把四个端点坐标压缩成 (x_0,y_0) 与 (u,v) 两组量.

知乎@余命数

例题 13.99 (垂直弦) 椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 上两点 $A(x_1,y_1),B(x_2,y_2)$,满足 $OA\perp OB$(O 为原点). 用双动点参数法推出关于 x_1,x_2,y_1,y_2 的对称关系, 并说明如何过渡到定值 $\frac{1}{|OA|^2}+\frac{1}{|OB|^2}$.

解 约束为

$$\frac{x_1^2}{4}+\frac{y_1^2}{3}=1, \qquad \frac{x_2^2}{4}+\frac{y_2^2}{3}=1, \qquad x_1x_2+y_1y_2=0.$$

由 $OA\perp OB$ 得 $y_1y_2=-x_1x_2$. 相加视角给出

$$\frac{1}{|OA|^2}+\frac{1}{|OB|^2}=\frac{x_1^2+y_1^2+x_2^2+y_2^2}{(x_1^2+y_1^2)(x_2^2+y_2^2)}.$$

这只是双点参写出的对称式; 若还要具体求值, 可引入弦 AB 的直线方程 $y=kx+m$ 联立椭圆读韦达(设线法), 或改用参数角直接吃掉垂直条件, 见旋转设点法中的例. 双动点参数法在此类点式对称条件中指明该算什么, 而定值计算常需与设线法或参数法配合.

知乎@余命数

13.5.2.0.2 双点参 vs 设线法 双动点参数法适合用端点坐标对称式表达的条件, 如中点、垂直 $x_1x_2+y_1y_2=0$、定比. 设线法适合用直线参数表达的条件, 如截距、斜率、弦长. 中点 (x_0,y_0) 与斜率 k 由点差联系; 韦达的 x_1+x_2,x_1x_2 即对称量的线参版本. 选择标准只看题面先给出的是点还是线.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.5.3 定比点差法

第13.4节 的点差法处理中点（$\lambda=1$ 的内分）. 若弦 AB 上一点 M 分 $\overrightarrow{AB}$ 成定比 $\lambda=\frac{\overrightarrow{AM}}{\overrightarrow{MB}}(\lambda\neq-1)$, 中点公式失效, 但两式相减、因式分解的代数骨架仍成立. 第15.3节 的调和分割把定比写成

$$P=\frac{A+\lambda B}{1+\lambda},\qquad Q=\frac{A-\lambda B}{1-\lambda},$$

其中 P,Q 是线段 AB 的一对调和共轭点. 该节的极点与极线定理进一步说明, 当割线绕定点 P 旋转时, 所有调和对点 Q 落在同一条直线上, 即 P 关于二次曲线的极线. 本节把这一推广写成独立的代数方法, 称为**定比点差法**.

知乎@余命数

知乎@余命数

定理 13.21 (定比点差关系) 设椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上一点 A, 点 $M(x_0,y_0)$ 分 $\overrightarrow{AB}$ 成定比 $\lambda(\lambda\neq-1)$, 由此确定点 B, 斜率 $k=\frac{y_B-y_A}{x_B-x_A}(x_A\neq x_B)$. 引入定比差分坐标

$$A=(x_0-\lambda u,\,y_0-\lambda v),\qquad B=(x_0+u,\,y_0+v),$$

其中 $k=\frac{v}{u}(u\neq0)$. 则 B 也在椭圆上当且仅当

$$\frac{-2x_0u+(\lambda-1)u^2}{a^2}+\frac{-2y_0v+(\lambda-1)v^2}{b^2}=0\quad\text{（已约去因子 }\lambda+1\text{）.}$$

特别地, 取中点 $\lambda=1$ 时退化为

$$k=-\frac{b^2x_0}{a^2y_0}\qquad(y_0\neq0).$$

证明 由定比定义 $\overrightarrow{AM}=\lambda\overrightarrow{MB}$, 上述 A,B 的坐标写法保证 M 恰为分点. 设 A 在椭圆上, 则 B 在椭圆上等价于

$$\frac{(x_0-\lambda u)^2}{a^2}+\frac{(y_0-\lambda v)^2}{b^2}=1,\qquad\frac{(x_0+u)^2}{a^2}+\frac{(y_0+v)^2}{b^2}=1.$$

知乎@余命数

两式相减, 第一项为

$$\frac{-2x_0u(\lambda+1)+u^2(\lambda^2-1)}{a^2}.$$

第二项同理. 因 $\lambda^2-1=(\lambda-1)(\lambda+1)$ 且 $\lambda\neq-1$, 约去 $\lambda+1$ 即得结论. 当 $\lambda=1$ 时 u^2,v^2 项消失, 留下 $\frac{x_0u}{a^2}+\frac{y_0v}{b^2}=0$, 即 $k=\frac{v}{u}=-\frac{b^2x_0}{a^2y_0}$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.5.3.0.1 实用写法与相关点代入 考场上可固定 $M(x_0,y_0)$ 与斜率 k, 令

$$A=(x_0-\lambda t,\ y_0-\lambda kt),\qquad B=(x_0+t,\ y_0+kt),$$

代入椭圆两式相减, 得到关于 t 的方程; 再由弦真实存在($\Delta\ge 0$ 或 $t\neq 0$)定 k. 中点 $\lambda=1$ 时一步回收点差公式.

若已知分点 M 而求一个端点的轨迹, 更直接的操作是**相关点代入**. 把 M 与已知端点的关系反解, 把未知端点用 M 的坐标表出, 再代回曲线方程. 它也适用于外分、中心对称等反射型映射.

知乎@余命数 知乎@余命数 知乎@余命数

例题 13.100 (中点 ($\lambda=1$)) 求椭圆 $\frac{x^2}{16}+\frac{y^2}{9}=1$ 中, 以 $M(2,1)$ 为中点的弦所在直线方程.

解 先检验 M 在椭圆内部. $\frac{4}{16}+\frac{1}{9}=\frac{13}{36}<1$. 由定比点差在 $\lambda=1$ 时的结论,

$$k=-\frac{9\cdot 2}{16\cdot 1}=-\frac{9}{8}.$$

故弦方程 $y-1=-\frac{9}{8}(x-2)$, 即 $9x+8y-26=0$.

例题 13.101 (一般定比 ($\lambda=2$), 轨迹) 抛物线 $C:y^2=4x$ 上动点 $P_0(x_0,y_0)$, 定点 $A(4,0)$. 点 M 分 $\overrightarrow{AP_0}$ 成定比 $\lambda=2$, 即 $\overrightarrow{AM}=\frac{2}{3}\overrightarrow{AP_0}$. 求 M 的轨迹.

解 由定比分点公式

知乎@余命数 知乎@余命数

$$M=\frac{A+2P_0}{3}=\Big(\frac{4+2x_0}{3},\ \frac{2y_0}{3}\Big).$$

反解得 $x_0=\frac{3x_M-4}{2}$, $y_0=\frac{3y_M}{2}$. 代回 $y_0^2=4x_0$.

$$\Big(\frac{3y_M}{2}\Big)^2=4\cdot\frac{3x_M-4}{2}\Longrightarrow y_M^2=\frac{8}{3}\Big(x_M-\frac{4}{3}\Big).$$

轨迹是开口向右的抛物线, 顶点 $(\frac{4}{3},0)$, 焦参数 $p'=\frac{4}{3}$.

一般 λ 时, 同法得

$$v^2=\frac{4\lambda}{1+\lambda}\Big(u-\frac{4}{1+\lambda}\Big),$$

知乎@余命数

顶点 $u_0=\frac{4}{1+\lambda}$, 焦参数 $p'=\frac{2\lambda}{1+\lambda}$. $\lambda=2$ 时 $p'=\frac{4}{3}$, 与上一致; $\lambda\to+\infty$ 时 $p'\to 2$, 轨迹趋近原抛物线.

知乎@余命数

例题 13.102 (外分与相关点代入) 椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 上动点 $M(x_M,y_M)$, N 为 M 关于定点 $(1,0)$ 的对称点. 求 N 的轨迹方程.

解 $(1,0)$ 为 MN 中点, 故 $N=(2-x_M,-y_M)$; 反解得 $M=(2-x_N,-y_N)$. 代回椭圆.

$$\frac{(2-x_N)^2}{4}+\frac{y_N^2}{3}=1\Longrightarrow\frac{(x_N-2)^2}{4}+\frac{y_N^2}{3}=1.$$

轨迹是原椭圆向右平移 2 个单位所得的合同椭圆. 这是相关点代入在 $\lambda=-1$ 反射情形的标准用法, 即反解 Φ^{-1} 代回曲线, 不必逐项展开联立.

13.5.3.0.2　定比点差法选用清单

- 已知中点求弦 $\Rightarrow\lambda=1$ 点差公式, 一步出斜率;
- 已知定比分点求轨迹 $\Rightarrow$ 反解曲线上点, 代回方程;
- 外分、中心对称 $\Rightarrow$ 反射型 $\Phi=\Phi^{-1}$, 轨迹为合同或平移曲线;
- $\lambda=-1$ 时分母 $1+\lambda=0$, 公式改写为中点公式, 不可硬套内分式.

13.5.4　旋转设点法

当几何条件涉及两线段垂直且等长, 或更一般地涉及向量的定角旋转时, 把一个动点写成另一个动点绕定点的旋转像, 通常比设两个独立点再配两条曲线约束更省参量.

13.5.4.0.1　等腰直角三角形的等长垂直条件(旋转 90°)

设 $\overrightarrow{OP_2}$ 由 $\overrightarrow{OP_1}$ 绕原点逆时针旋转角 θ 得到, 则

$$\begin{pmatrix}x_2\\y_2\end{pmatrix}=\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}\begin{pmatrix}x_1\\y_1\end{pmatrix},\tag{13-27}$$

即

$$(x_2,y_2)=(x_1\cos\theta-y_1\sin\theta,\ x_1\sin\theta+y_1\cos\theta).$$

最常用的特例是 $\theta=\frac{\pi}{2}$, 此时

$$(x_2,y_2)=(-y_1,x_1).$$

它把 $OP_1\perp OP_2$ 与 $|OP_1|=|OP_2|$ 两个条件压缩为一个代数关系. 若旋转中心为任意定点 $C(x_c,y_c)$, 先平移到 C 再旋转, 最后加回 C.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.103 (圆上旋转 90°) 圆 $x^2+y^2=r^2$ 上点 $A(x_1,y_1)$, 设 B 为 A 绕原点逆时针旋转 90° 的像. 证明 B 仍在圆上, 并求弦 AB 的中点轨迹.

知乎@余命数

解 由旋转公式 $B=(-y_1,x_1)$, 则

$$x_B^2+y_B^2=y_1^2+x_1^2=r^2,$$

故 B 在圆上. 中点

$$M=\Big(\frac{x_1-y_1}{2},\ \frac{y_1+x_1}{2}\Big),$$

因此

$$|OM|^2=\frac{(x_1-y_1)^2+(x_1+y_1)^2}{4}=\frac{2x_1^2+2y_1^2}{4}=\frac{r^2}{2}.$$

中点轨迹为 $x^2+y^2=\frac{r^2}{2}$(去掉使 A,B 重合的退化点).

13.5.4.0.2 直角三角形的仅垂直条件(斜率积 -1 或点乘为零) 若条件只要求垂直而不要求等长, 旋转矩阵会引入多余等长约束, 此时应回到点乘

$$OA\perp OB\iff x_1x_2+y_1y_2=0.$$

在抛物线这类可用参数 t 表示点的曲线上, 垂直条件常与参数积直接挂钩.

例题 13.104 (抛物线过定点的直角弦) 抛物线 $y^2=2px\ (p>0)$ 上两点 A,B 满足 $\overrightarrow{OA}\perp\overrightarrow{OB}$. 证明直线 AB 恒过定点 $(2p,0)$.

证明 用参数设点. 令 $A(\frac{p}{2}t_1^2,\,pt_1)$, $B(\frac{p}{2}t_2^2,\,pt_2)$, 其中 $t_1\neq t_2$. 条件 $\overrightarrow{OA}\cdot\overrightarrow{OB}=0$ 即

$$\frac{p}{2}t_1^2\cdot\frac{p}{2}t_2^2+(pt_1)(pt_2)=0\implies\frac{p^2}{4}t_1^2t_2^2+p^2t_1t_2=0.$$

若 $t_1t_2=0$, 则 A 或 B 为顶点 O, 与弦 AB 的通常设定冲突, 故 $t_1t_2\neq0$. 两边除以 $p^2t_1t_2$ 得

$$\frac{1}{4}t_1t_2+1=0\implies t_1t_2=-4.$$

抛物线上两点弦的标准方程为

$$y(t_1+t_2)=2x+p\,t_1t_2.$$

将 $t_1t_2=-4$ 代入, 得

$$y(t_1+t_2)=2x-4p.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

把点 $(2p,0)$ 代入. 右端为 $2\cdot(2p)-4p=0$, 左端为 $0\cdot(t_1+t_2)=0$, 恒成立. 当 $t_1+t_2=0$ 时弦方程退化为 $x=2p$, 仍过 $(2p,0)$.

因此所有满足条件的弦 AB 所在直线恒过定点 $(2p,0)$. □

13.5.4.0.3　一般斜率积 $k_{OA}\cdot k_{OB}=c$　垂直只是斜率积 $c=-1$ 的特例. 更一般地, 设 A,B 是从原点出发的两条射线与曲线的交点, 若斜率积

$$k_{OA}\cdot k_{OB}=c$$

为常数 ($c\neq 0$), 则点乘条件变为 $y_1y_2=c\,x_1x_2$. 在抛物线 $y^2=2px$ 上, 用参数 t 可把这个条件直接翻译成 t_1t_2 的常数.

例题 13.105 (定斜率积的抛物线弦) 抛物线 $y^2=2px$ ($p>0$) 上两点 A,B 满足 $k_{OA}\cdot k_{OB}=c$ ($c\neq 0$). 求弦 AB 过的定点.

解 设 $A(\frac{p}{2}t_1^2,\,pt_1)$, $B(\frac{p}{2}t_2^2,\,pt_2)$. 由 $k_{OA}=\frac{2}{t_1}$, $k_{OB}=\frac{2}{t_2}$, 条件 $k_{OA}k_{OB}=c$ 给出

$$\frac{4}{t_1t_2}=c\implies t_1t_2=\frac{4}{c}.$$

抛物线弦方程

$$y(t_1+t_2)=2x+p\,t_1t_2$$

代入 $t_1t_2=\frac{4}{c}$ 得

$$y(t_1+t_2)=2x+\frac{4p}{c}.$$

令 $y=0$, 则 $2x+\frac{4p}{c}=0$, 即 $x=-\frac{2p}{c}$. 所以弦 AB 恒过定点 $(-\frac{2p}{c},0)$. 当 $c=-1$ 时回到垂直情形的定点 $(2p,0)$.

13.5.4.0.4　三类旋转设点法的分工

- 等长且垂直 $\Rightarrow$ 旋转矩阵 (13-27), $\theta=\frac{\pi}{2}$ 时 $(x,y)\mapsto(-y,x)$;
- 仅垂直 $\Rightarrow$ 点乘 $x_1x_2+y_1y_2=0$, 或抛物线参数下的 $t_1t_2=-4$;
- 定斜率积 $k_{OA}k_{OB}=c\Rightarrow$ 用参数把条件化为 $t_1t_2=\frac{4}{c}$ 等常数关系, 再读弦方程.

13.5.5 本节小结

四种策略的对照如下. 单动点参数法设一动点 (x_0,y_0), 分曲线上与曲线外两种情形, 经设点、直译、消元、求解处理切线截距、距离、夹角等问题. 双动点参数法

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

用双点参把两端点同时命名, 强调设而不求. 相减得斜率或中点, 相加得和式; 中点情形可用对称差分 $A=(x_0+u,y_0+v)$, $B=(x_0-u,y_0-v)$, 与 第13.4节 的点差法一致. 定比点差法用 λ 推广中点点差, 覆盖中点、一般定比轨迹与外分/反射; 第15.3节 的极点极线定理说明调和对点的共线性. 旋转设点法处理等长垂直、仅垂直、一般斜率积三类旋转条件, 用旋转矩阵、点乘或参数积把几何约束一次写成代数关系.

大题常把几种手法组合. 例如 A,B 在椭圆上且 $\angle AOB=\frac{\pi}{2}$, P 为 AB 中点求轨迹时, 可用点乘处理垂直, 再以 $\lambda=1$ 写中点并消去参量. 第13.6节 把这些选择与具体题型对应起来.

13.5.6 本节习题

知乎@余命数

13.15. 椭圆 $\frac{x^2}{4}+y^2=1$ 上一点 $P(x_0,y_0)$ $(x_0y_0\neq 0)$, 过 P 的切线交两坐标轴于 A,B. 求 $|AB|$ 的最小值. 提示:切线 $\frac{xx_0}{4}+yy_0=1$, $|AB|^2=\frac{16}{x_0^2}+\frac{1}{y_0^2}$, 令 $t=x_0^2$ 消元.

13.16. 圆 $x^2+y^2=r^2$ 上动点 $P(x_0,y_0)$, 定点 $A(a,0)$ $(0<a<r)$. 用单动点参数法表示 $\cos\angle OAP$, 并说明 $\angle OAP$ 何时最大.

13.17. 椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上弦 AB 的中点为 $M(x_0,y_0)$. 设 $A=(x_0+u,y_0+v)$, $B=(x_0-u,y_0-v)$, 用双点参推导 $k_{AB}=-\frac{b^2x_0}{a^2y_0}$ $(y_0\neq 0)$.

知乎@余命数 知乎@余命数 知乎@余命数

13.18. 椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 上 A,B 满足 $OA\perp OB$. 用双动点参数法写出关于 x_1,x_2,y_1,y_2 的对称关系, 并说明如何求 $\frac{1}{|OA|^2}+\frac{1}{|OB|^2}$.

知乎@余命数

13.19. 求椭圆 $\frac{x^2}{25}+\frac{y^2}{16}=1$ 中, 以 $M(1,2)$ 为中点的弦所在直线方程. 先检验 M 在椭圆内部, 再套 $\lambda=1$ 定比点差.

知乎@余命数

13.20. 抛物线 $y^2=2px$ 上动点 P_0, 定点 $A(p,0)$, M 分 $\overrightarrow{AP_0}$ 成定比 $\lambda=1$(中点). 求 M 的轨迹;再对一般 $\lambda>0$ 写出轨迹焦参数随 λ 的变化.

13.21. 圆 $x^2+y^2=r^2$ 上点 $A(x_1,y_1)$, B 为 A 绕原点逆时针旋转 90° 的像. 证明 B 在圆上, 并求弦 AB 中点轨迹.

13.22. 抛物线 $y^2=2px$ 上 A,B 满足 $k_{OA}\cdot k_{OB}=c$ $(c\neq 0)$. 用参数 t_1,t_2 证明 $t_1t_2=\frac{4}{c}$, 并由此证明弦 AB 恒过定点 $\left(-\frac{2p}{c},0\right)$. 取 $c=-1$ 验证即为直角情形.

13.23. 抛物线 $y^2=4x$ 上动点 $P(x_0,y_0)$ 到直线 $y=x+4$ 的距离为 d. 用单动点参数法把 d^2 表示为 x_0 的函数, 并求最小值.

知乎@余命数

13.24. 椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 上两点 A,B 关于原点 O 对称, P 为椭圆上异于 A,B 的任一点. 用双动点参数法证明 $k_{PA}\cdot k_{PB}=-\frac{3}{4}$(提示:用椭圆方程消去 $y_0^2-y_1^2$).

知乎@余命数

13.25.　椭圆 $\frac{x^2}{4}+y^2=1$ 上动点 $A(x_0,y_0)$, 定点 $M(1,0)$. 点 B 满足 $\overrightarrow{AM}=2\overrightarrow{MB}$. 用相关点代入求 B 的轨迹方程.

13.26.　圆 $x^2+(y-1)^2=1$ 上异于原点的一点 $A(x_1,y_1)$. 将 OA 绕原点逆时针旋转 90° 得到 OB, 用旋转设点法求 B 的轨迹方程(提示: 用旋转关系把 A 的坐标用 B 表示).

13.6 设点法的典型应用

我们已经看到设点法的基本技术与参量选择, 见 第13.4节 与 第13.5节. 本节用它们处理五类题型, 即向量轨迹、线段比例、定截距弦、斜率关系与原点三角形. 每类附代表性例题. 一般顺序是先选定参量, 再把几何条件写成坐标关系, 最后用曲线方程整体代换. 联立通法见 第13.7节.

知乎@余命数

13.6.1 向量条件求轨迹

我们先设轨迹点 $M=(X,Y)$ 与曲线上动点 $P=(x_0,y_0)$. 如果题面以向量给出定点、动点与 M 的关系, 我们把这些向量等式改写成 X,Y 与 x_0,y_0 的线性关系. 反解出 $x_0=x_0(X,Y)$, $y_0=y_0(X,Y)$, 再代入曲线方程 $F(x_0,y_0)=0$, 整理即得 M 的轨迹方程. 最后讨论使 P 仍在曲线上、或使分母非零的定义域.

例题 13.106 (中点轨迹) 椭圆 $C:\frac{x^2}{9}+\frac{y^2}{4}=1$ 上动点 P, 定点 $A(3,0)$, M 为线段 AP 的中点. 求 M 的轨迹方程.

解 设 $M=(X,Y)$, $P=(x_0,y_0)$. 由中点公式 $\overrightarrow{OM}=\frac{1}{2}(\overrightarrow{OA}+\overrightarrow{OP})$, 得

$$X=\frac{3+x_0}{2},\quad Y=\frac{y_0}{2},$$

即 $x_0=2X-3$, $y_0=2Y$. 因 $P\in C$,

$$\frac{(2X-3)^2}{9}+\frac{(2Y)^2}{4}=1\ \Longrightarrow\ \frac{(2X-3)^2}{9}+Y^2=1.$$

配方后得到 $\left(X-\frac{3}{2}\right)^2\Big/\frac{9}{4}+Y^2=1$. 故 M 的轨迹是以 $(\frac{3}{2},0)$ 为中心, 半轴长 $\frac{3}{2}$ 与 1 的椭圆(去掉使 P 越出 C 的边界点即可).

例题 13.107 (向量伸缩轨迹) 抛物线 $C:y^2=4x$ 上动点 P, 定点 $F(1,0)$(焦点). 设点 M 满足 $\overrightarrow{OM}=\overrightarrow{OF}+\frac{1}{2}\overrightarrow{OP}$. 求 M 的轨迹方程, 并说明轨迹类型.

知乎@余命数

解 设 $P=(x_0,y_0)$, $M=(X,Y)$. 由 $y_0^2=4x_0$ 及

$$(X,Y)=(1,0)+\frac{1}{2}(x_0,y_0)=\Big(1+\frac{x_0}{2},\ \frac{y_0}{2}\Big),$$

得 $x_0=2(X-1)$, $y_0=2Y$. 代入抛物线约束:

$$(2Y)^2=4\cdot 2(X-1)\implies 4Y^2=8(X-1)\implies Y^2=2(X-1).$$

故 M 的轨迹是抛物线 $y^2=2(x-1)$, 顶点 $(1,0)$, 开口向右, 焦参数为 1(由原抛物线 $y^2=4x$ 经半幅伸缩与平移合成).

13.6.1.0.1 位置关系的点式判断 轨迹题常附带点在曲线内外的讨论. 对 $F(x,y)=\frac{x^2}{a^2}+\frac{y^2}{b^2}-1=0$ 型椭圆, 我们直接把点坐标代入 F 看符号, 不必先求交点. 标准椭圆上 $F=0$, 内部 $F<0$, 外部 $F>0$. 例如 $F(3,1)=\frac{9}{9}+\frac{1}{4}-1=\frac{1}{4}>0$ 给出 $(3,1)$ 在椭圆 $\frac{x^2}{9}+\frac{y^2}{4}=1$ 外; $F(1,1)=\frac{1}{9}+\frac{1}{4}-1<0$ 给出该点在内. 这是设点法不求交点即判位置的典型用法, 也常用于斜率/中点题中先验中点是否在曲线内部.

知乎@余命数

例题 13.108 (位置关系检验) 设 $F(x,y)=\frac{x^2}{9}+\frac{y^2}{4}-1$. 判断点 $M(3,1)$ 与 $N(1,1)$ 相对椭圆 $F=0$ 的位置, 并说明在轨迹题中的用途.

解 $F(3,1)=\frac{1}{4}>0$, 故 M 在椭圆外; $F(1,1)=\frac{13}{36}-1<0$, 故 N 在椭圆内. 这可用于求中点弦与轨迹范围. 求中点弦时, 先算 $F(x_M,y_M)<0$ 确认中点在内部, 弦才存在; 求轨迹范围时, 用 F 的符号截取参数区间.

13.6.2 线段比例问题

知乎@余命数

定比分点把比例条件翻译为线性关系. 设 P 分 $\overrightarrow{AB}$ 成定比 λ, 即 $\overrightarrow{AP}=\lambda\overrightarrow{PB}$, 则

$$P=\frac{A+\lambda B}{1+\lambda}\quad(\lambda\neq -1),$$

坐标形式

$$x_P=\frac{x_A+\lambda x_B}{1+\lambda},\qquad y_P=\frac{y_A+\lambda y_B}{1+\lambda}.$$

在曲线约束下, 常可反解端点, 或配合 第13.5节 的定比点差法写中点/定比弦方程.

13.6.2.0.1 操作骨架

知乎@余命数

1. 读出 λ(中点 $\lambda=1$; 外分 $\lambda<0$ 且 $\lambda\neq -1$);
2. 用定比公式表达端点之一, 或写差分 A,B 相对分点的坐标;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3. 把两端点代入曲线, 相减得 λ 型点差, 或直接消元求坐标;
4. 若需求弦方程, 已知两端点时优先用不除零的两点式整式形式.

例题 13.109 (定比 $\lambda=2$ 求端点) 抛物线 $y^2=8x$ 上两点 A,B, 点 $P(3,0)$ 满足 $\overrightarrow{AP}=2\overrightarrow{PB}$. 求 A,B 的坐标.

解 由 $\overrightarrow{AP}=2\overrightarrow{PB}$ 得 $P=\frac{A+2B}{3}$, 故 $A=3P-2B=(9-2x_B,-2y_B)$. A 在抛物线上: $(-2y_B)^2=8(9-2x_B)$, 即 $4y_B^2=72-16x_B$. 又 $y_B^2=8x_B$, 故 $32x_B=72-16x_B$、$48x_B=72$、$x_B=\frac{3}{2}$. 从而 $y_B^2=12$, $y_B=\pm2\sqrt{3}$; $x_A=9-3=6$, $y_A=\mp4\sqrt{3}$. 两组解: $A=(6,-4\sqrt{3})$, $B=(\frac{3}{2},2\sqrt{3})$ 或 $A=(6,4\sqrt{3})$, $B=(\frac{3}{2},-2\sqrt{3})$. 二者对应同一类弦关于 x 轴的对称.

例题 13.110 (中点弦) 椭圆 $\frac{x^2}{25}+\frac{y^2}{16}=1$ 中, 以 $M(1,2)$ 为中点的弦所在直线方程(先检验 M 在椭圆内部).

知乎@余命数

解 先算 $F(1,2)=\frac{1}{25}+\frac{4}{16}-1=\frac{1}{25}+\frac{1}{4}-1=\frac{29}{100}-1<0$, 故 M 在内部, 中点弦存在. 由 第13.4节 弦中点斜率关系 ($a^2=25,b^2=16,\lambda=1$):

$$k=-\frac{b^2x_M}{a^2y_M}=-\frac{16\cdot1}{25\cdot2}=-\frac{8}{25}.$$

点斜式: $y-2=-\frac{8}{25}(x-1)$, 即 $8x+25y-58=0$. 此即所求弦方程.

13.6.2.0.2　由两端点直写弦方程　已知两端点时, 优先用不除零的两点式

$$(y-y_1)(x_2-x_1)=(y_2-y_1)(x-x_1),$$

避开斜率分类. 当 $x_1=x_2$ 时直写 $x=x_1$; 当 $y_1=y_2$ 时直写 $y=y_1$. 整式形式在竖直弦时仍可用, 而斜率式 $m=\frac{y_2-y_1}{x_2-x_1}$ 会失效.

例题 13.111 (两端点直写弦方程) 椭圆 $C:\frac{x^2}{9}+\frac{y^2}{4}=1$ 上两点 $A(0,2),B(3,0)$. 直接写出弦 AB 的方程.

解 由两点式整式形式得到

$$(y-2)(3-0)=(0-2)(x-0)\implies3(y-2)=-2x\implies2x+3y-6=0.$$

知乎@余命数

我们验证如下. 过 A 时 $0+6-6=0$; 过 B 时 $6+0-6=0$. 若误用斜率式 $m=-\frac{2}{3}$ 再写 $y-2=-\frac{2}{3}x$, 结果相同, 但当弦竖直时斜率式失效, 整式形式始终可用.

13.6.3 定截距弦问题

截距式 $\frac{x}{a_0}+\frac{y}{b_0}=1$ 适合弦与两轴交点已知或参与条件的题型, 可避开斜率分类. 截距也可用端点坐标表达, 若弦交 x 轴于 $(a_0,0)$, 交 y 轴于 $(0,b_0)$, 则 a_0,b_0 即为截距.

13.6.3.0.1 操作骨架

1. 若截距已知, 直接写 $\frac{x}{a_0}+\frac{y}{b_0}=1$, 解出 y(或 x)代入曲线;
2. 整理为一元二次, 用韦达读 x_A+x_B, x_Ax_B(不必求根);
3. 若中点已知, 可用点参差分 $A=(x_M+u,y_M+v)$, $B=(x_M-u,y_M-v)$ 求弦长与面积;
4. 弦长 $|AB|=2\sqrt{u^2+v^2}$; 面积 $S_{\triangle OAB}=\frac{1}{2}|x_Ay_B-x_By_A|$.

例题 13.112 (截距弦的韦达) 椭圆 $C:\frac{x^2}{16}+\frac{y^2}{9}=1$, 直线 $\ell:\frac{x}{6}+\frac{y}{4}=1$ 与 C 交于 A,B. 求 x_A+x_B 与 x_Ax_B.

解 由 ℓ 得 $y=4-\frac{2}{3}x$. 代入椭圆:

$$\frac{x^2}{16}+\frac{\left(4-\frac{2}{3}x\right)^2}{9}=1.$$

整理(两边乘 144):$9x^2+16\left(4-\frac{2}{3}x\right)^2=144$, 即

$$9x^2+16\left(16-\frac{16}{3}x+\frac{4}{9}x^2\right)=144,$$

$$9x^2+256-\frac{256}{3}x+\frac{64}{9}x^2=144.$$

再乘 9:$81x^2+2304-768x+64x^2=1296$, 即 $145x^2-768x+1008=0$. 由韦达定理(第13.3节):$x_A+x_B=\frac{768}{145}$, $x_Ax_B=\frac{1008}{145}$. 不必求根即可进入后续弦长或中点计算.

13.6.3.0.2 弦长与面积的坐标视角

中点已知时, 点参差分是联立法的对照路径, 设 $A=(x_M+u,y_M+v)$, $B=(x_M-u,y_M-v)$, 则斜率 $k=\frac{v}{u}$, 弦长 $|AB|=2\sqrt{u^2+v^2}$, 面积可由行列式直接展开. 设线、联立、韦达与弦长的系统讨论见 第13.7节.

例题 13.113 (弦长与 $\triangle OAB$ 面积) 椭圆 $C:\frac{x^2}{16}+\frac{y^2}{9}=1$ 的弦 AB 中点为 $M(1,\frac{1}{2})$. 求弦长 $|AB|$ 与 $\triangle OAB$ 的面积.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 由弦中点斜率关系 $(a^2=16, b^2=9)$:

$$k=-\frac{b^2x_M}{a^2y_M}=-\frac{9\cdot 1}{16\cdot\frac{1}{2}}=-\frac{9}{8}.$$

设

$$A=(1+u,\tfrac{1}{2}+v),\quad B=(1-u,\tfrac{1}{2}-v),$$

则 $k=\frac{v}{u}=-\frac{9}{8}$, 即 $v=-\frac{9}{8}u$. 代入 $A\in C$:

$$\frac{(1+u)^2}{16}+\frac{\left(\frac{1}{2}-\frac{9}{8}u\right)^2}{9}=1.$$

整理 $\left(\frac{1}{2}-\frac{9}{8}u\right)^2=\frac{(4-9u)^2}{64}$, 故

$$\frac{(1+u)^2}{16}+\frac{(4-9u)^2}{576}=1.$$

乘 576:

$$36(1+2u+u^2)+(4-9u)^2=576,$$

$$36+72u+36u^2+16-72u+81u^2=576,$$

$$117u^2+52=576\implies 117u^2=524\implies u^2=\frac{524}{117}.$$

$\gcd(524,117)=1$, 不可再约. 又 $F(1,\frac{1}{2})=\frac{1}{16}+\frac{1}{36}-1=-\frac{131}{144}<0$, 确认 M 在椭圆内, 弦确实存在.

弦长: 由 $u^2+v^2=u^2(1+\frac{81}{64})=u^2\cdot\frac{145}{64}$,

$$|AB|=2\sqrt{u^2+v^2}=2|u|\cdot\frac{\sqrt{145}}{8}=\frac{\sqrt{145}}{4}\sqrt{\frac{524}{117}}.$$

面积: 直接展开 $x_Ay_B-x_By_A$,

$$x_Ay_B-x_By_A=(1+u)(\tfrac{1}{2}-v)-(1-u)(\tfrac{1}{2}+v)=u-2v=u-2\Big(-\frac{9}{8}u\Big)=\frac{13}{4}u,$$

故

$$S_{\triangle OAB}=\frac{1}{2}|x_Ay_B-x_By_A|=\frac{1}{2}\cdot\frac{13}{4}|u|=\frac{13}{8}\sqrt{\frac{524}{117}}.$$

13.6.4 斜率关系问题

斜率条件 $k=\frac{y_1-y_2}{x_1-x_2}$、$k_1k_2=c$、$k_1+k_2=s$ 可与曲线方程, 点差公式联立, 转化为对称式. 椭圆与双曲线的点差符号不同, 需分别处理.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定理 13.22 (原点弦的斜率积与垂直) 设椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 上两点 $A(x_1,y_1),B(x_2,y_2)$, 且 $x_1x_2\neq 0$. 记 $k_1=\frac{y_1}{x_1},k_2=\frac{y_2}{x_2}$.

1. 若 $OA\perp OB$, 则 $k_1k_2=-1$;
2. 若弦 AB 的中点为 $M(x_0,y_0)$ $(y_0\neq 0)$, 则弦斜率

$$k_{AB}=-\frac{b^2x_0}{a^2y_0}$$

(点差公式, 见 第13.4节).

证明 (1) 由 $OA\perp OB$ 得 $x_1x_2+y_1y_2=0$, 即 $y_1y_2=-x_1x_2$. 故

$$k_1k_2=\frac{y_1y_2}{x_1x_2}=-1.$$

(2) 由 A,B 在椭圆上, 两式相减并因式分解(点差法):

$$\frac{(x_1-x_2)(x_1+x_2)}{a^2}+\frac{(y_1-y_2)(y_1+y_2)}{b^2}=0.$$

除以 x_1-x_2 $(x_1\neq x_2)$, 并以 $x_0=\frac{x_1+x_2}{2},y_0=\frac{y_1+y_2}{2}$ 代入, 即得

$$k_{AB}=\frac{y_1-y_2}{x_1-x_2}=-\frac{b^2x_0}{a^2y_0}.$$

□

13.6.4.0.1 操作骨架

1. 把斜率条件写成过原点的 $y_i=k_ix_i$, 或一般弦的 $k=\frac{y_2-y_1}{x_2-x_1}$;
2. 用点差得中点斜率关系, 或用 k_1k_2,k_1+k_2 的对称式;
3. 椭圆上 $k=-\frac{b^2x_0}{a^2y_0}$, 双曲线上 $k=\frac{b^2x_0}{a^2y_0}$ (符号相反, 见 第13.4节);
4. 消元得中点轨迹或参数关系, 并截取内部范围.

例题 13.114 (椭圆的定斜率弦中点轨迹) 椭圆 $C:\frac{x^2}{9}+\frac{y^2}{5}=1$ 的弦 AB 斜率为 1, 中点为 $M(x_0,y_0)$. 求 M 满足的关系式及轨迹范围.

解 由 第13.3节 弦中点斜率关系 $(a^2=9,b^2=5)$:

$$k=-\frac{b^2x_0}{a^2y_0}=-\frac{5x_0}{9y_0}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $k=1$:$9y_0=-5x_0$, 即 $y_0=-\frac{5}{9}x_0$. 中点须在椭圆内部:

$$\frac{x_0^2}{9}+\frac{y_0^2}{5}<1 \implies \frac{x_0^2}{9}+\frac{25x_0^2}{81\cdot 5}<1 \implies \frac{x_0^2}{9}+\frac{5x_0^2}{81}<1,$$

$$\frac{9x_0^2+5x_0^2}{81}=\frac{14x_0^2}{81}<1 \implies x_0^2<\frac{81}{14}, \quad |x_0|<\frac{9}{\sqrt{14}}.$$

故中点轨迹为线段 $y=-\frac{5}{9}x$ 落在椭圆内部的部分.

例题 13.115 (双曲线的定斜率弦中点轨迹) 双曲线 $C:\frac{x^2}{4}-\frac{y^2}{3}=1$ 的弦 AB 斜率为 1, 中点为 $M(x_0,y_0)$ $(y_0\neq 0)$. 求 M 满足的关系式, 并说明与椭圆点差的符号差异.

解 由 第13.4节 双曲线弦中点斜率关系 $(a^2=4,b^2=3)$:

$$k=\frac{b^2x_0}{a^2y_0}=\frac{3x_0}{4y_0}.$$

与椭圆 $k=-\frac{b^2x_0}{a^2y_0}$ 相比, 右侧符号相反, 来自双曲线方程中 y^2 项的减号. 令 $k=1$, 得

$$4y_0=3x_0 \implies y_0=\frac{3}{4}x_0.$$

中点须使弦确实存在, 当 M 位于双曲线两支之间的空隙侧时需使 $\Delta>0$. 代入点差相容条件, 粗判

$$\frac{x_0^2}{4}-\frac{y_0^2}{3}=\frac{x_0^2}{4}-\frac{9x_0^2}{48}=\frac{x_0^2}{4}-\frac{3x_0^2}{16}=\frac{x_0^2}{16}.$$

当 $x_0\neq 0$ 时此值 >0, 与椭圆情形要求中点在曲线内部不同; 双曲线中点弦的存在性需回到联立 $\Delta>0$ 检验(细节见 第13.7节).

例题 13.116 (抛物线的定斜率积弦) 抛物线 $C:y^2=4x$ 上两点 A,B 满足 $k_{OA}\cdot k_{OB}=2$. 证明弦 AB 恒过定点, 并求该定点坐标.

解 用旋转设点法中的定斜率积模式,见 第13.5.4小节. 设 $A=(t_1^2,2t_1)$,$B=(t_2^2,2t_2)$ $(t_1t_2\neq 0)$, 则

$$k_{OA}=\frac{2}{t_1}, \qquad k_{OB}=\frac{2}{t_2}.$$

条件 $k_{OA}\cdot k_{OB}=2$ 给出

$$\frac{4}{t_1t_2}=2 \implies t_1t_2=2.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

抛物线 $y^2=4x$ 的两点弦方程为

$$y(t_1+t_2)=2x+2t_1t_2.$$

代入 $t_1t_2=2$, 得

$$y(t_1+t_2)=2x+4.$$

令 $y=0$ 得 $x=-2$, 故弦 AB 恒过定点 $(-2,0)$, 与 t_1+t_2 的取值无关.

注 13.7
对抛物线, 若斜率积 $k_1k_2=c$ 为定值, 则弦 AB 恒过定点; 椭圆与双曲线没有这一结论, 斜率积为定值只约束弦的方向. 点参可写 $A=(x_1,k_1x_1)$ 等, 但展开常繁; 改设线 $y=kx+m$ 联立后用韦达读对称式往往更省. 后者将在 第13.7节 中处理, 习题中也保留一道双曲线斜率积过定点的对照题.

13.6.5 原点与两动点三角形问题

以 O,A,B 为顶点时, 面积

$$S_{\triangle OAB}=\frac{1}{2}|x_1y_2-x_2y_1|,$$

垂直条件

$$OA\perp OB\iff x_1x_2+y_1y_2=0,$$

常与曲线约束联立得定值或最值.

13.6.5.0.1 操作骨架

1. 设 $A(x_1,y_1),B(x_2,y_2)$ 在曲线上, 写出 $F(x_i,y_i)=0$;
2. 垂直则加 $x_1x_2+y_1y_2=0$; 面积则写 $\frac{1}{2}|x_1y_2-x_2y_1|$;
3. 引入方向角或参数方程, 见 第14章, 或设线联立读韦达;
4. 化简得定值 / 最值, 并讨论退化(A,B 共线于过原点的直线时面积为 0).

例题 13.117 (垂直定值) 椭圆 $C:\frac{x^2}{16}+\frac{y^2}{9}=1$ 上两点 A,B 满足 $OA\perp OB$. 证明

$$\frac{1}{|OA|^2}+\frac{1}{|OB|^2}=\frac{1}{16}+\frac{1}{9}=\frac{25}{144}.$$

知乎@余命数

解 设 A 所在射线与 x 轴正半轴夹角为 α（此为方向角，非离心角；系统参数方程见 第14章）. 令 $|OA|=r_1$，则 $A=(r_1\cos\alpha, r_1\sin\alpha)$. 由 $A\in C$：

知乎@余命数

$$\frac{r_1^2\cos^2\alpha}{16}+\frac{r_1^2\sin^2\alpha}{9}=1 \implies \frac{1}{r_1^2}=\frac{\cos^2\alpha}{16}+\frac{\sin^2\alpha}{9}.$$

因 $OA\perp OB$，可取 B 的方向角为 $\alpha+\frac{\pi}{2}$，则

$$\frac{1}{r_2^2}=\frac{\sin^2\alpha}{16}+\frac{\cos^2\alpha}{9}.$$

知乎@余命数

两式相加：

知乎@余命数

$$\frac{1}{r_1^2}+\frac{1}{r_2^2}=\frac{\cos^2\alpha+\sin^2\alpha}{16}+\frac{\sin^2\alpha+\cos^2\alpha}{9}=\frac{1}{16}+\frac{1}{9}=\frac{25}{144}.$$

例题 13.118 (面积最值) 椭圆 $C:\frac{x^2}{9}+\frac{y^2}{4}=1$ 上两点 A,B 满足 $OA\perp OB$. 求 $\triangle OAB$ 面积的最大值.

解 设 A 的方向角为 α、$|OA|=r_1$、$|OB|=r_2$，取

$$A=(r_1\cos\alpha, r_1\sin\alpha),\quad B=(-r_2\sin\alpha, r_2\cos\alpha)$$

（保证 $OA\perp OB$）. 则

知乎@余命数　知乎@余命数　知乎@余命数

$$x_1y_2-x_2y_1=r_1r_2 \implies S=\frac{1}{2}r_1r_2.$$

由 $A,B\in C$：

$$\frac{1}{r_1^2}=\frac{\cos^2\alpha}{9}+\frac{\sin^2\alpha}{4},\qquad \frac{1}{r_2^2}=\frac{\sin^2\alpha}{9}+\frac{\cos^2\alpha}{4}.$$

令 $t=\cos^2\alpha\in[0,1]$，$\sin^2\alpha=1-t$，则

$$\frac{1}{r_1^2r_2^2}=\frac{(9-5t)(4+5t)}{1296},$$

$$S=\frac{18}{\sqrt{g(t)}},$$

$$g(t)=(9-5t)(4+5t)=36+25t-25t^2.$$

$g'(t)=25-50t=0\Rightarrow t=\frac{1}{2}$；$g''=-50<0$，故 g 在 $t=\frac{1}{2}$ 取最大值

知乎@余命数

$$g\left(\tfrac{1}{2}\right)=\frac{169}{4}.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

端点：$g(0)=g(1)=36$. 因 S 与 $\sqrt{g}$ 反变，

知乎@余命数

$$S_{\max}=\frac{18}{\sqrt{36}}=3\quad(t=0\text{ 或}1),\qquad S_{\min}=\frac{18}{13/2}=\frac{36}{13}\quad(t=\tfrac12).$$

故面积最大值为 3，取最大时 A 落在坐标轴方向（$\alpha=0$ 或 $\frac{\pi}{2}$ 的等价配置）.

注 13.8
上例展示了垂直与面积条件并存时的完整点参路径. 若只求最大值，也可用 $S=\frac12|OA|\,|OB|$（因直角）与 $\frac{1}{r_1^2}+\frac{1}{r_2^2}$ 定值配合均值不等式；细节因题而异. 设线法对照见 第13.7节.

知乎@余命数

13.6.6 本节小结

我们回顾五类应用的路径. 对向量条件，先把向量等式写成 X,Y 与 x_0,y_0 的线性关系，反解动点，再用曲线约束消元，必要时由 F 的符号判断内外. 对线段比例，用定比 λ 与点差；已知两端点时可用整式两点式直写弦方程. 对定截距弦，用截距式联立读韦达；中点已知时则用差分求弦长与 $\triangle OAB$ 面积. 对斜率关系，把 k、k_1k_2、k_1+k_2 写成对称式，在椭圆与双曲线上分别消元（点差符号相反）. 对原点三角形，面积 $S=\frac12|x_1y_2-x_2y_1|$，垂直时 $x_1x_2+y_1y_2=0$，可用方向角或参数方程消元.

设点法适合单点或双点坐标条件、定比、中点、垂直与面积行列式. 当条件主要是直线与曲线两交点的对称式（弦长、焦点弦调和平均等），通常改用设线法更为直接. 下一节 第13.7节 从联立运算展开设线法.

知乎@余命数

13.6.7 本节习题

13.27. 椭圆 $\frac{x^2}{25}+\frac{y^2}{16}=1$ 上动点 P，M 为 OP 中点，求 M 的轨迹方程，并说明范围.

13.28. 抛物线 $y^2=4x$ 上 A,B，点 $P(2,0)$ 满足 $\overrightarrow{AP}=\overrightarrow{PB}$，求弦 AB 方程的可能形式.（可用定比 $\lambda=1$ 或中点点差；并尝试两端点整式两点式复核.）

13.29. 椭圆 $\frac{x^2}{9}+\frac{y^2}{4}=1$ 上 A,B 使 $OA\perp OB$，求 $\triangle OAB$ 面积的最大值.

13.30. 判断点 $(2,2)$ 与椭圆 $\frac{x^2}{4}+\frac{y^2}{9}=1$ 的位置关系，并说明依据（F 符号法）.

13.31. 椭圆 $\frac{x^2}{16}+\frac{y^2}{9}=1$，直线 $\frac{x}{8}+\frac{y}{3}=1$ 与椭圆交于 A,B. 求 x_A+x_B，x_Ax_B，并写出 $|AB|$ 的表达式（可用韦达 + 弦长公式，或对照中点差分）.

知乎@余命数

13.32. 抛物线 $y^2=4x$ 上异于原点的两点 A,B 满足 $k_{OA}k_{OB}=\lambda\,(\lambda\neq0)$. 证明弦 AB 过定点，并求该定点.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

13.33.　椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 上 A,B 满足 $OA\perp OB$. 证明 $\frac{1}{|OA|^2}+\frac{1}{|OB|^2}=\frac{1}{a^2}+\frac{1}{b^2}$.

13.34.　已知椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 的弦 AB 中点为 $M(1,1)$. 先用 F 检验 M 是否在内部, 再求 $|AB|$ 与 $S_{\triangle OAB}$.

13.7 联立运算的基础准备

我们在 第13.3节 已经看到联立法的三条基本公式. 韦达定理为 $x_1+x_2=-\dfrac{b}{a}$, $x_1x_2=\dfrac{c}{a}$, 弦长为 $|P_1P_2|=\sqrt{1+k^2}\dfrac{\sqrt{\Delta}}{|a|}$, 焦半径为 $|PF_2|=a-ex_0$. 设点法见 第13.4节 至 第13.6节. 本节在这些结论之上, 整理设线法的联立步骤、判别式化简、弦长与焦半径的选用, 以及面积与对称式的变形.

13.7.1 联立的做法

13.7.1.0.1　为什么联立　解析几何用方程刻画曲线, 点 $P(x_0,y_0)$ 落在曲线上当且仅当它的坐标满足该方程. 把这条对应推广到两条曲线, 交点的含义就清楚了. 交点同时落在两条曲线上, 它的坐标必须同时满足两个方程; 反过来, 同时满足两个方程的 (x,y) 也一定是交点. 因此, 求两条曲线的交点就是求联立方程组的公共解. 以椭圆与直线为例, 我们写出

$$\begin{cases} b^2x^2+a^2y^2=a^2b^2, \\ y=kx+m, \end{cases}$$

方程组的每一组解对应一个交点, 解的组数就是交点的个数 (图 13.34).

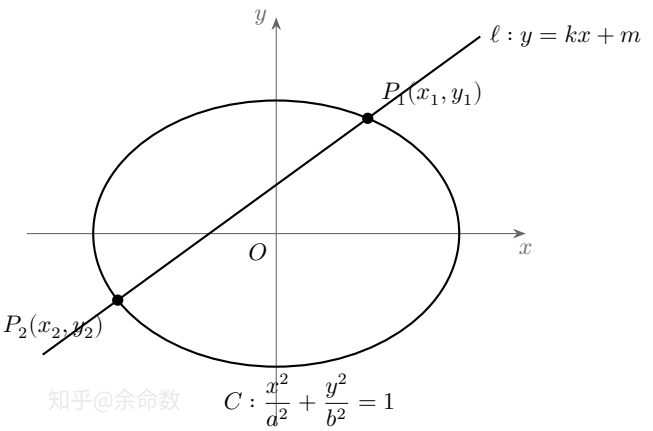


图 13.34　交点 P_1,P_2 的坐标同时满足直线方程与椭圆方程, 是联立方程组的全部解; 解的组数就是交点的个数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

几何直观能判断大概相交, 但交点坐标、相切条件与弦长这类精确答案必须由代数给出. 第一步是消元. 把 $y=kx+m$ 代入椭圆方程, 方程组化为关于 x 的一元二次方程, 交点个数随之变成实根个数. 实根个数由判别式决定. 若 $\Delta>0$, 则有两个交点; 若 $\Delta<0$, 则没有交点. 中间情形 $\Delta=0$ 最为关键. 让直线从相交位置向外平移, 两个交点沿曲线连续靠近, 直到重合为一点, 直线恰好成为切线. 两个根相等这件事, 写在代数上正是 $\Delta=0$. 相切与重根是同一现象的几何与代数两面, 见 图 13.35.

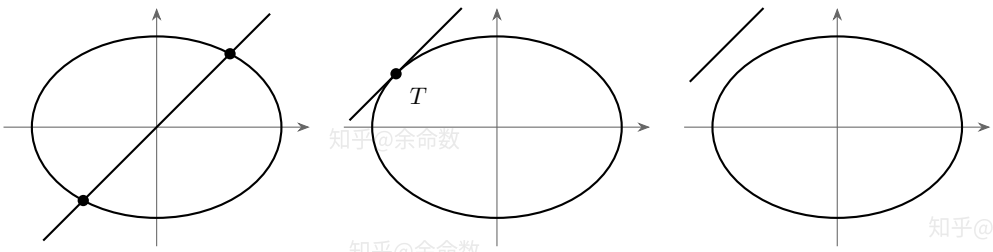


图 13.35 三种位置关系与判别式符号一一对应. 相切是相交的临界情形, 两个交点连续靠近并重合, 代数上对应二重根.

这也解释了为什么我们可以停在系数层面而不解根. 弦长只含 $|x_1-x_2|$, 中点只含 x_1+x_2, 面积只含这两者的组合, 它们都关于两根对称. 由韦达定理, 这些对称式都能用 x_1+x_2 与 x_1x_2 表出, 因此求根可以省去. 整条计算路径至此成形, 见图 13.36. 几何条件经设线联立变成方程组, 消元后变成一元二次方程, 再由系数读出 Δ、S、P, 最后经对称式还原为几何量. 联立是这条路径的起点.

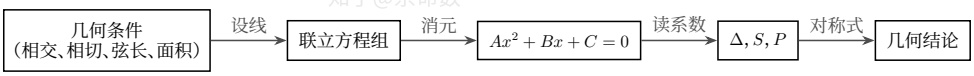


图 13.36 联立的计算路径. 几何条件经设线、消元化为一元二次方程, 由系数读出判别式与韦达两式, 再经对称式还原为几何结论.

我们用设线法把几何条件化为代数二次方程, 再从系数读出位置关系与对称量, 通常不解根. 做法如下. 先依题面设直线 $y=kx+m$ 或 $x=ty+n$ (斜率不存在时优先后者, 见 第13.8节), 再把直线代入曲线方程消去一个变量, 整理成 $Ax^2+Bx+C=0$ 或关于 y 的同形. 整理时我们不再乘额外常数, 以免判别式与韦达分母失真, 见 第13.7.2小节. 随后计算 $\Delta=B^2-4AC$, 判定相交、相切或相离, 并由系数直接写出

$$x_1+x_2=-\frac{B}{A},\qquad x_1x_2=\frac{C}{A}$$

此即 第13.3节 的设而不求. 弦长、中点、面积、距离、斜率等后处理, 一律化为 $S=x_1+x_2$, $P=x_1x_2$ 与已知参数的对称式.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.7.1.0.2　设线　联立的第一步是设直线方程. 标准形有两种, $y=kx+m$ 与 $x=ty+n$, 二者覆盖的直线几乎一样, 运算量却可能相差很大. 我们以过椭圆 $b^2x^2+a^2y^2=a^2b^2$ 右焦点 $F(c,0)$ 的弦为例说明原因. 把 $x=ty+c$ 代入椭圆方程, 得

$$b^2(ty+c)^2+a^2y^2=a^2b^2 \implies (a^2+b^2t^2)y^2+2b^2tc\,y+b^2(c^2-a^2)=0.$$

因为 $c^2-a^2=-b^2$, 常数项直接化简为 $-b^4$, 与参数 t 无关. 改用 $y=k(x-c)$ 代入, 常数项变成 $a^2(k^2c^2-b^2)$, 含两个 k^2 项, 此后的每次通分都要处理它们. 同一条弦, 仅因设线形式不同, 系数就相差一整项. 因此, 直线过 x 轴上的点 $(x_0,0)$ 时, 我们设 $x=ty+x_0$; 过 y 轴上的点 $(0,y_0)$ 或斜率已给定时, 我们设 $y=kx+m$. 第一种写法还把斜率不存在的情形自动收进 $t=0$, 不必另立讨论.

13.7.1.0.3　代入消元　设好直线, 接下来把它代入曲线方程, 消去一个变量. 以椭圆为例, 把 $y=kx+m$ 代入 $b^2x^2+a^2y^2=a^2b^2$, 得到

$$b^2x^2+a^2(kx+m)^2=a^2b^2 \implies b^2x^2+a^2k^2x^2+2a^2km\,x+a^2m^2-a^2b^2=0.$$

这一步唯一的难点在展开 $(kx+m)^2$, 中项 $2kmx$ 最容易在誊写中丢失, 而一旦丢掉, 后面的判别式与韦达全错, 错误还很隐蔽. 稳妥的做法是分两步走, 先写出骨架 $(kx+m)^2=k^2x^2+2kmx+m^2$, 确认三项齐全, 再把 a^2 整体乘进去.

13.7.1.0.4　整理成 $Ax^2+Bx+C=0$　展开之后, 合并同类项并把右端移为零, 得到标准二次形

$$(b^2+a^2k^2)x^2+2a^2km\,x+a^2(m^2-b^2)=0.$$

常数项 $a^2m^2-a^2b^2$ 我们直接因式为 $a^2(m^2-b^2)$, 与 定理 13.23 的形态对齐, 下一步计算 Δ 时 m^2-b^2 会作为一个整块出现. 整理后还要检查二次项系数是否为零. 椭圆情形 $A=b^2+a^2k^2>0$ 恒成立, 检查自动通过; 双曲线情形 $A=b^2-a^2k^2$ 可正可负可为零, 一旦 $A=0$, 方程退化为一次, 对应直线与渐近线平行, 判别式与韦达全都用不上, 见 注 13.9. 整理只允许移项与合并, 不同乘任何常数, 代价在下一段说明.

13.7.1.0.5　计算判别式　方程整理成 $Ax^2+Bx+C=0$ 后, 位置关系由判别式判定. 判别式来自求根公式 $x=\dfrac{-B\pm\sqrt{B^2-4AC}}{2A}$, 根号下的 $\Delta=B^2-4AC$ 决定实根个数, 所以 $\Delta>0$ 时交于两点, $\Delta=0$ 时相切 (重根), $\Delta<0$ 时相离. 题面含参数时, 连展开 Δ 都可以省掉, 定理 13.23 已把椭圆情形整理成因子形式 $\Delta=4a^2b^2(b^2+a^2k^2-m^2)$, 符号直接由括号读出, 这是含参题最省的一步.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.7.1.0.6 读韦达与设而不求 确认交点存在之后, 下一步由韦达定理给出两根的和与积. 设两根为 x_1, x_2, 则

$$Ax^2+Bx+C=A(x-x_1)(x-x_2)=Ax^2-A(x_1+x_2)x+Ax_1x_2,$$

比较两边系数得 $B=-A(x_1+x_2)$, $C=Ax_1x_2$, 于是

$$x_1+x_2=-\frac{B}{A},\qquad x_1x_2=\frac{C}{A}.$$

有了和与积, 根差也随之确定

$$(x_1-x_2)^2=(x_1+x_2)^2-4x_1x_2=\frac{B^2}{A^2}-\frac{4C}{A}=\frac{B^2-4AC}{A^2}=\frac{\Delta}{A^2},$$

因此 $|x_1-x_2|=\dfrac{\sqrt{\Delta}}{|A|}$, 弦长公式 $|P_1P_2|=\sqrt{1+k^2}\,|x_1-x_2|$ 所需的根差由此读出, 全程不解根. 但在使用韦达定理之前, 我们必须确认 $\Delta\geq 0$. 若 $\Delta<0$, 韦达两式在复数范围仍成立, 和与积形式上存在, 但几何上没有交点, 由它们推出的弦长与中点皆无意义. 为自检, 也可取 k 或 m 的特殊值代回, 用真根核对 S, P 的符号与数量级.

整理时为什么不同乘常数, 原因如下. 设把整理好的方程两边同乘 $\lambda\neq 0$, 系数同步放缩为 λA、λB、λC, 此时

$$\Delta'=(\lambda B)^2-4(\lambda A)(\lambda C)=\lambda^2\Delta,\qquad -\frac{\lambda B}{\lambda A}=-\frac{B}{A},\qquad \frac{\lambda C}{\lambda A}=\frac{C}{A}.$$

根与韦达两式不变, 判别式却放缩了 λ^2 倍. 单看方程本身, 放缩无害; 问题出在与现成公式对照的时候.定理 13.23 的 $\Delta=4a^2b^2(b^2+a^2k^2-m^2)$ 与联立面积公式 (13-29) 都按未放缩的 A,B,C 推导, 方程一旦被乘过 λ, 直接代入就差出 λ^2 或 λ 的因子, 检验也要带校正系数. 为免去这类校正, 本节统一采用自然标准化的做法, 即把 $y=kx+m$ 代入 $b^2x^2+a^2y^2=a^2b^2$ 后直接展开为 $(b^2+a^2k^2)x^2+2a^2km\,x+a^2(m^2-b^2)=0$. 有分数系数时只允许去分母一次, 如两边乘 4 化整, 此后不再放缩, 公式与方程一一对应, 无需再校正.

13.7.1.0.7 常见失误清单

- 展开 $(kx+m)^2$ 漏掉中项 $2kmx$, 或移项忘记变号. 展开、乘常数、移项分三步写, 每步只完成一个动作, 可以减少这类错误;
- 双曲线、抛物线未检查 $A\neq 0$ 就计算 Δ, 会漏掉退化一次方程对应的几何情形(渐近线平行, 水平线);
- $\Delta<0$ 时仍代入韦达两式, 和与积形式上算得出, 对应的交点却不存在;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 整理时多乘常数放缩 A,B,C, 再把放缩后的 Δ 代入简化引理或面积公式, 会差出 λ^2 的因子;
- 求面积时先开方计算 $|x_1-x_2|$ 再平方, 会多算一次开方, 直接用 $(x_1-x_2)^2=\dfrac{\Delta}{A^2}$ 可以避免.

13.7.1.0.8　三类曲线的基础联立模板

- **椭圆** $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 上, 代入 $y=kx+m$, 得

$$(b^2+a^2k^2)x^2+2a^2km\,x+a^2(m^2-b^2)=0.$$

知乎@余命数

- **双曲线** $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ 上, 代入 $y=kx+m$, 得

$$(b^2-a^2k^2)x^2-2a^2km\,x-a^2(m^2+b^2)=0$$

(注意二次项系数 $A=b^2-a^2k^2$ 可正可负, 且 $A=0$ 时退化为一次方程, 对应渐近线平行情形).

- **抛物线** $y^2=2px$ 上, 代入 $y=kx+m$, 得

$$k^2x^2+(2km-2p)x+m^2=0$$

(若 $k=0$ 则直线水平, 联立后关于 x 至多一次, 需单独讨论).

例题 13.119 (椭圆基础联立) 椭圆 $C:\dfrac{x^2}{25}+\dfrac{y^2}{16}=1$, 直线 $\ell:y=x+1$. 将直线与椭圆联立, 求 x_1+x_2, x_1x_2 与 Δ, 并判定 ℓ 与 C 的位置关系.

解 取 $k=1,m=1$, 代入 $16x^2+25y^2=400$:

$$16x^2+25(x+1)^2=400\ \Longrightarrow\ 41x^2+50x-375=0.$$

故 $A=41$、$B=50$、$C=-375$. 判别式

$$\Delta=B^2-4AC=50^2-4\cdot41\cdot(-375)=2500+61500=64000>0,$$

故 ℓ 与 C 相交于两点. 由韦达定理 (第13.3节):

$$x_1+x_2=-\frac{50}{41},\qquad x_1x_2=-\frac{375}{41}.$$

不必求根即可得到和与积, 供后续弦长、面积等后处理使用.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.120 (双曲线基础联立) 双曲线 $C:\dfrac{x^2}{9}-\dfrac{y^2}{4}=1$, 直线 $\ell:y=\dfrac{1}{2}x+1$. 写出联立后的一元二次方程, 求 Δ 与韦达两式, 并说明交点个数.

解 此处 $a^2=9$、$b^2=4$、$k=\dfrac{1}{2}$、$m=1$. 代入 $4x^2-9y^2=36$:

$$4x^2-9\Big(\frac{x}{2}+1\Big)^2=36 \implies 4x^2-9\Big(\frac{x^2}{4}+x+1\Big)=36$$
$$\implies 4x^2-\frac{9}{4}x^2-9x-9=36.$$

知乎@余命数

两边乘 4: $16x^2-9x^2-36x-36=144$, 即 $7x^2-36x-180=0$. 故 $A=7$、$B=-36$、$C=-180$. 判别式

知乎@余命数

$$\Delta=(-36)^2-4\cdot 7\cdot(-180)=1296+5040=6336>0,$$

交于两点. 韦达:

$$x_1+x_2=\frac{36}{7},\qquad x_1x_2=-\frac{180}{7}.$$

注意 $A=b^2-a^2k^2=4-9\cdot\dfrac{1}{4}=\dfrac{7}{4}>0$(自然标准化后乘 4 得 $A=7$), 未触及 $A=0$ 的渐近线平行退化.

知乎@余命数

例题 13.121 (抛物线基础联立) 抛物线 $C:y^2=8x$, 即 $2p=8$, $p=4$, 直线 $\ell:y=x+2$. 求交点横坐标的和与积, 并求弦长 $|AB|$.

解 代入 $y=x+2$ 得 $(x+2)^2=8x$, 即 $x^2+4x+4=8x$, 整理 $x^2-4x+4=0$, 即 $(x-2)^2=0$. 故 $\Delta=0$, 直线与抛物线相切于一点 $A(2,4)$(二重根). 此时弦退化为切点, $|AB|=0$.

若改取 $\ell:y=x+1$, 则 $(x+1)^2=8x$、$x^2+2x+1=8x$、$x^2-6x+1=0$.$\Delta=36-4=32>0$,

$$x_1+x_2=6,\qquad x_1x_2=1,\qquad |x_1-x_2|=\sqrt{32}=4\sqrt{2}.$$

知乎@余命数

由弦长公式 (第13.3节, $k=1$):

知乎@余命数

$$|AB|=\sqrt{1+1}\cdot 4\sqrt{2}=4\sqrt{2}\cdot\sqrt{2}=8.$$

抛物线联立展开后常数项常为 m^2, 一次项含 $-2p$, 系数形态与椭圆不同, 但联立与读韦达的步骤相同.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.7.2 判别式的计算与简化

判别式 $\Delta = B^2 - 4AC$ 在联立中反复出现，由符号判定位置关系，由 $\sqrt{\Delta}$ 写弦长，由整式形态讨论定值与存在性 ($\Delta \geq 0$). 直接展开往往冗长，但按固定方式整理后，因子形式通常很清楚. 现给出椭圆与双曲线情形的判别式简化引理，并说明公因子提取与双曲线 $A = 0$ 的退化.

定理 13.23 (椭圆直线判别式简化引理) 设椭圆 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ ($a > b > 0$)，直线 $\ell : y = kx + m$. 联立消 y 得二次方程

$$(b^2 + a^2k^2)x^2 + 2a^2km\,x + a^2(m^2 - b^2) = 0.$$

其判别式简化为

$$\Delta = 4a^2b^2(b^2 + a^2k^2 - m^2). \tag{13-28}$$

特别地，ℓ 与椭圆相切 $\iff \Delta = 0 \iff m^2 = b^2 + a^2k^2$.

证明 联立 $y = kx + m$ 与 $b^2x^2 + a^2y^2 = a^2b^2$：

$$b^2x^2 + a^2(kx + m)^2 - a^2b^2 = 0.$$

展开 $a^2(kx + m)^2 = a^2k^2x^2 + 2a^2km\,x + a^2m^2$，合并得

$$(b^2 + a^2k^2)x^2 + 2a^2km\,x + a^2(m^2 - b^2) = 0.$$

此处 $A = b^2 + a^2k^2$、$B = 2a^2km$、$C = a^2(m^2 - b^2)$. 计算

$$\begin{aligned}\Delta = B^2 - 4AC &= (2a^2km)^2 - 4(b^2 + a^2k^2)\cdot a^2(m^2 - b^2)\\ &= 4a^4k^2m^2 - 4a^2\big(b^2m^2 - b^4 + a^2k^2m^2 - a^2k^2b^2\big)\\ &= 4a^4k^2m^2 - 4a^2b^2m^2 + 4a^2b^4 - 4a^4k^2m^2 + 4a^4k^2b^2\\ &= 4a^2b^2\big(-m^2 + b^2 + a^2k^2\big) = 4a^2b^2\big(b^2 + a^2k^2 - m^2\big).\end{aligned}$$

$m^2 = b^2 + a^2k^2 \iff \Delta = 0$. 二次项系数 $A = b^2 + a^2k^2 > 0$ 恒成立，故相切 $\iff \Delta = 0$. □

注 13.9 (双曲线陷阱)
双曲线 $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$ 的平行结论是将 (13-28) 改为

$$\Delta = 4a^2b^2(m^2 + b^2 - a^2k^2);$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

此时 $A=b^2-a^2k^2$ 可正可负, 且 $A=0$ 时退化一次方程, 对应直线与渐近线平行(斜率 $k=\pm\dfrac{b}{a}$). 故相切 $\iff \Delta=0$ 的结论须附加 $A\neq 0$ 的条件. 凡见 $k=\pm\dfrac{b}{a}$, 先检查 A 是否为零, 再谈 Δ.

知乎@余命数

13.7.2.0.1 公因子提取 展开 $\Delta=B^2-4AC$ 后, 常先提出 $4a^2$ 或 $4b^2$ 等公因子, 再整理括号内的 k,m 多项式. 引理 (13-28) 已把椭圆情形一次完成; 含参题可**跳过展开步骤**直接写因子形式, 然后读符号, 读等式.

例题 13.122 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$, 直线 $\ell:y=x+m$ $(m\in\mathbb{R})$. 用判别式简化引理直接写出 Δ, 并由此分类 ℓ 与 C 的位置关系.

解 此处 $a^2=9$、$b^2=4$、$k=1$. 由 定理 13.23,

$$\Delta=4\cdot 9\cdot 4(4+9-m^2)=144(13-m^2).$$

- 若 $\Delta>0$, 即 $m^2<13$ 亦即 $|m|<\sqrt{13}$, 则 ℓ 与 C 交于两点;
- 若 $\Delta=0$, 即 $|m|=\sqrt{13}$, 则相切;
- 若 $\Delta<0$, 即 $|m|>\sqrt{13}$, 则相离, 无公共点.

若不用引理直接算: 由 $4x^2+9(x+m)^2=36$ 得 $13x^2+18mx+9m^2-36=0$,

$$\Delta=324m^2-4\cdot 13\cdot(9m^2-36)=144(13-m^2).$$

同结果, 但要算三步乘开方; 引理一步直读因子. 对 k 含参的题型, 速差更大.

几何校验如下. 直线 $x-y+m=0$ 与椭圆相切的条件为 $A^2a^2+B^2b^2=C^2$. 此处 $A=1$、$B=-1$、$C=m$, 故 $9+4=m^2\iff m^2=13$, 与引理给出的相切条件 $m^2=b^2+a^2k^2=13$ 完全一致.

13.7.3 弦长公式的灵活选用

第13.3节 的弦长公式有两条等价形式: 对直线 $y=kx+m$ $(k\neq 0)$,

$$|P_1P_2|=\sqrt{1+k^2}\,|x_1-x_2|=\sqrt{1+\frac{1}{k^2}}\,|y_1-y_2|,$$

第二条可写为 $\dfrac{\sqrt{1+k^2}}{|k|}|y_1-y_2|$. 选用哪个形式, 取决于哪种方程的根读得更顺.

选用时可大致依斜率大小: $|k|\le 1$ 时优先消 y 得关于 x 的二次方程(A 中 a^2k^2 不大, 系数较干净); $|k|>1$ 时优先把直线写成 $x=ty+n$ $(t=\dfrac{1}{k})$, 近竖直线时 x 形系数膨胀, y 形更省. 过焦点的弦若目标含 $|PF|$ 或 $|P_1F|+|P_2F|$, 宜先用焦半径

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$|PF_2| = a - ex_0$ (第13.3节) 把距离一阶线性化, 再与韦达配合. 个别题目某一边恰好因式分解时, 仍以具体计算为准.

例题 13.123　椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{9}=1(a^2=16、b^2=9、c=\sqrt{7}$, 右焦点 $F(\sqrt{7},0))$, 过 F 斜率为 4 的直线 $\ell: y=4(x-\sqrt{7})$. 试用两种弦长形式分别计算 $|AB|$, 并比较两条代数路径的难易.

解

把 $y=4(x-\sqrt{7})$ 代入 $9x^2+16y^2=144$:

$$9x^2+16\cdot 16(x-\sqrt{7})^2=144 \implies 265x^2-512\sqrt{7}\,x+1648=0.$$

判别式

$$\Delta_x=(512\sqrt{7})^2-4\cdot 265\cdot 1648=1835008-1746880=88128=5184\cdot 17.$$

故

$$|x_1-x_2|=\frac{\sqrt{\Delta_x}}{265}=\frac{72\sqrt{17}}{265},\qquad |AB|=\sqrt{1+16}\cdot\frac{72\sqrt{17}}{265}=\frac{72\cdot 17}{265}=\frac{1224}{265}.$$

把 ℓ 改写为 $x=\dfrac{1}{4}y+\sqrt{7}$, 代入 $9x^2+16y^2=144$:

$$9\Big(\frac{y}{4}+\sqrt{7}\Big)^2+16y^2=144 \implies \frac{265}{16}y^2+\frac{9\sqrt{7}}{2}y-81=0.$$

自然标准化(乘 16)得 $265y^2+72\sqrt{7}\,y-1296=0$, 判别式

$$\Delta_y=(72\sqrt{7})^2-4\cdot 265\cdot(-1296)=36288+1373760=1410048=16\cdot 88128.$$

注意 $\Delta_y=16\Delta_x$ 来自乘 16 这一放缩. 计算

$$\begin{aligned}|y_1-y_2|&=\frac{\sqrt{\Delta_y}}{265}=\frac{288\sqrt{17}}{265},\\ |AB|&=\sqrt{1+\frac{1}{16}}\cdot\frac{288\sqrt{17}}{265}\\ &=\frac{\sqrt{17}}{4}\cdot\frac{288\sqrt{17}}{265}=\frac{1224}{265}.\end{aligned}$$

两条路径给出同一答案 $\dfrac{1224}{265}$. 这里 $|k|=4$ 偏大, 形式二中间常系数较大; 当 k 近无穷 (近竖直线) 时, 形式二的判别式比形式一小得多, 选形式二才合算. 仅当 $|k|$ 明显偏离 1 时, 改用另一形式才有实际增益.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

过 F 的弦若附加焦半径条件, 如 $|AF|+|BF|$ 为定值, 焦半径公式 $|PF_2|=a-ex_0=4-\dfrac{\sqrt{7}}{4}x_0$(这里 $e=\dfrac{c}{a}=\dfrac{\sqrt{7}}{4}$)会进一步把弦长与中点条件线性化, 详见第13.9节 的韦达后处理专题.

13.7.4 面积的联立表达

三角形面积在联立框架里有三条常用路径, 可视已知条件选用.

13.7.4.0.1 底高式 $S=\dfrac{1}{2}|AB|\cdot d(P,\ell)$, 其中 $|AB|$ 用弦长公式, $d(P,\ell)$ 用点到直线距离. 二者均含因子 $\sqrt{1+k^2}$ 或其倒数, 在乘积中相互抵消, 给出不含根号的联立面积式.

13.7.4.0.2 坐标行列式 已知两端点坐标时, $S=\dfrac{1}{2}|x_1y_2-x_2y_1|$ (第三顶点为原点; 一般第三顶点见 第13.4节 鞋带公式). 联立后若已得 $S=x_1+x_2$, $P=x_1x_2$ 及 $y_i=kx_i+m$, 可化为

$$S_{\triangle OAB}=\frac{1}{2}|m|\,|x_1-x_2|,$$

一般情形则用 y_1y_2, $x_1y_2+x_2y_1$ 的对称展开.

13.7.4.0.3 割补 当三角形与坐标轴或已知矩形/梯形有包含关系时, 用大面积减小面积, 避免直接算斜边. 典型如 $\triangle OAB$ 与坐标轴围成的梯形割补.

13.7.4.0.4 联立面积公式(椭圆, 第三顶点为原点) 椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$, 直线 $\ell: y=kx+m$ 交椭圆于 A,B. 以原点 O 为第三顶点的 $\triangle OAB$ 面积为

$$S_{\triangle OAB}=\frac{ab\,|m|\sqrt{b^2+a^2k^2-m^2}}{b^2+a^2k^2}. \tag{13-29}$$

推导如下(底高式 $+\sqrt{1+k^2}$ 互消).

$$\begin{aligned}S&=\frac{1}{2}|AB|\cdot d(O,\ell)\\&=\frac{1}{2}\cdot\underbrace{\sqrt{1+k^2}\frac{\sqrt{\Delta}}{b^2+a^2k^2}}_{\text{弦长, 见 第13.3节}}\cdot\underbrace{\frac{|m|}{\sqrt{1+k^2}}}_{\text{点线距}}\\&=\frac{|m|\sqrt{\Delta}}{2(b^2+a^2k^2)}.\end{aligned}$$

代入 定理 13.23 的 $\Delta=4a^2b^2(b^2+a^2k^2-m^2)$, 得

$$\sqrt{\Delta}=2ab\sqrt{b^2+a^2k^2-m^2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(取正, 即须 $b^2+a^2k^2-m^2\ge 0$). 整理即得 (13-29).

第三顶点不必是 O, 只需把 $|m|$ 相应换为

$$|kx_0-y_0+m|,$$

仍可同样在弦长乘积里消去 $\sqrt{1+k^2}$.

例题 13.124　椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{4}=1$ ($a^2=16$、$b^2=4$、$c=2\sqrt{3}$, 右焦点 $F(2\sqrt{3},0)$), 过 F 斜率 $k=1$ 的直线 $\ell: y=x-2\sqrt{3}$. 求 $\triangle OAB$ 的面积, 分别用底高式独立计算和联立面积公式 (13-29) 两条路径.

解　此处 $k=1, m=-2\sqrt{3}$, 故 $|m|=2\sqrt{3}$ 且 $b^2+a^2k^2-m^2=4+16-12=8$.

联立得 $(4+16)x^2+2\cdot 16\cdot 1\cdot(-2\sqrt{3})\,x+16(12-4)=0$, 即 $20x^2-64\sqrt{3}\,x+128=0$, 简化为 $5x^2-16\sqrt{3}\,x+32=0$.

$$\Delta_x=(16\sqrt{3})^2-4\cdot 5\cdot 32=768-640=128.$$

故 $|x_1-x_2|=\dfrac{\sqrt{128}}{5}=\dfrac{8\sqrt{2}}{5}$, $|AB|=\sqrt{2}\cdot\dfrac{8\sqrt{2}}{5}=\dfrac{16}{5}$.

点 O 到 $\ell: x-y-2\sqrt{3}=0$ 的距离 $d=\dfrac{|-2\sqrt{3}|}{\sqrt{2}}=\sqrt{6}$. 于是

$$S_{\triangle OAB}=\frac{1}{2}\cdot\frac{16}{5}\cdot\sqrt{6}=\frac{8\sqrt{6}}{5}.$$

代入 $a=4$、$b=2$、$|m|=2\sqrt{3}$、$b^2+a^2k^2-m^2=8$、$b^2+a^2k^2=20$:

$$S_{\triangle OAB}=\frac{4\cdot 2\cdot 2\sqrt{3}\cdot\sqrt{8}}{20}=\frac{16\sqrt{3}\cdot 2\sqrt{2}}{20}=\frac{32\sqrt{6}}{20}=\frac{8\sqrt{6}}{5}.$$

两路径同结果. 区别在路径一需独立算弦长与点线距再相乘 (经历 $\sqrt{1+k^2}$ 与 $\dfrac{1}{\sqrt{1+k^2}}$ 互消的隐性步骤), 路径二把互消预先完成, 常数项直接整体出场.

$a=4, e=\dfrac{c}{a}=\dfrac{\sqrt{3}}{2}$. 对交点 A, B, 由焦半径公式 (第13.3节) $|AF_2|=a-ex_A=4-\dfrac{\sqrt{3}}{2}x_A$, 同样 $|BF_2|=4-\dfrac{\sqrt{3}}{2}x_B$. 焦点为弦内分点, 故

$$|AF|+|BF|=|AB|=\frac{16}{5}\quad\Longleftrightarrow\quad 8-\frac{\sqrt{3}}{2}(x_A+x_B)=\frac{16}{5}.$$

联立得 $x_A+x_B=\dfrac{16\sqrt{3}}{5}$. 由韦达(二次方程 $5x^2-16\sqrt{3}\,x+32=0$) $x_A+x_B=\dfrac{16\sqrt{3}}{5}$, 两式吻合.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

13.7.5 代数变形的基本技巧

联立之后, 许多目标量都是关于两根的对称式. 记 $S=x_1+x_2$, $P=x_1x_2$, 下列几类整式会反复用到.

13.7.5.0.1 常用对称式

- $x_1^2+x_2^2=S^2-2P$;
- $x_1^3+x_2^3=S^3-3PS$;
- $(x_1-x_2)^2=S^2-4P=\dfrac{\Delta}{A^2}$;
- $x_1^2x_2+x_1x_2^2=SP$;
- $\dfrac{1}{x_1}+\dfrac{1}{x_2}=\dfrac{S}{P}\,(P\neq 0)$;
- $(x_1-c)(x_2-c)=P-cS+c^2$.

变形时宜先看清目标式(待证、待求或待比较), 再反向构造 S,P 的线性组合, 避免盲目展开; 多项式先写成 S,P 的整式再代入韦达. 遇到 $x_1^2+x_2^2$ 一类, 常可配上与 m,k 或 c 相关的项使目标齐次, 便于读出定值. 若目标式本身含可分解因子, 如 x_1x_2, x_1+x_2+D, 宜先分解再代入.

知乎@余命数 知乎@余命数

13.7.5.0.2 数值速算提醒

$\Delta=B^2-4AC$ 展开时, 先提公因子 4 或 A 的公共数字; $(x_1-x_2)^2=\dfrac{\Delta}{A^2}$ 可避免先开方再平方的往返. 韦达代入通分时, 分子按 k^2c^2 等同类项归并, 往往一次消掉繁杂的中项.

知乎@余命数

例题 13.125 椭圆 $C:\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1\,(a>b>0,\,c^2=a^2-b^2)$, 右焦点 $F(c,0)$, 直线 $\ell:y=k(x-c)$ 过右焦点, 斜率为 k. 设 ℓ 交 C 于 A,B 两点. 试证

$$(x_A-c)(x_B-c)=-\frac{b^4}{b^2+a^2k^2}.$$

解 ℓ 即 $y=kx-kc$, 对应 $y=kx+m$ 形下 $m=-kc$. 代入椭圆方程按自然标准化整理:

$$b^2x^2+a^2(kx-kc)^2-a^2b^2=0\implies (b^2+a^2k^2)x^2-2a^2ck^2\,x+a^2(k^2c^2-b^2)=0.$$

记 $A_0=b^2+a^2k^2$、$B_0=-2a^2ck^2$、$C_0=a^2(k^2c^2-b^2)$. 由韦达定理 (第13.3节),

$$S=x_A+x_B=\frac{2a^2ck^2}{b^2+a^2k^2},\qquad P=x_Ax_B=\frac{a^2(k^2c^2-b^2)}{b^2+a^2k^2}.$$

知乎@余命数

针对目标式展开

$$(x_A-c)(x_B-c)=P-cS+c^2.$$

代入韦达两式, 通分 $A_0=b^2+a^2k^2$:

$$(x_A-c)(x_B-c)=\frac{a^2(k^2c^2-b^2)-2a^2c^2k^2+c^2(b^2+a^2k^2)}{b^2+a^2k^2}.$$

整理分子(按 $a^2k^2c^2$ 三次同类项归并):

$$\begin{aligned}\text{分子}&=a^2k^2c^2-a^2b^2-2a^2c^2k^2+c^2b^2+c^2a^2k^2\\&=(a^2k^2c^2-2a^2c^2k^2+a^2k^2c^2)+(-a^2b^2+b^2c^2)\\&=0+b^2(c^2-a^2)\\&=b^2\cdot(-b^2)=-b^4,\end{aligned}$$

最后一步用了 $c^2-a^2=(a^2-b^2)-a^2=-b^2$. 故

$$(x_A-c)(x_B-c)=-\frac{b^4}{b^2+a^2k^2}.$$

等式对一切使 $\Delta>0$ 的 k 成立(现 $m=-kc$, $b^2+a^2k^2-m^2=b^2+a^2k^2-k^2c^2=b^2+k^2b^2=b^2(1+k^2)>0$ 恒真).

由于 $b^2+a^2k^2>0$, 右端恒为负. 这正反映 F 严格位于 A,B 之间, 因为 x_A-c 与 x_B-c 异号. 进一步, $|(x_A-c)(x_B-c)|=\dfrac{b^4}{b^2+a^2k^2}$ 仅含 k^2, 当 $k\to\infty$(弦趋于竖直)时渐近 $\dfrac{b^4}{a^2k^2}\to 0$, 与近竖直的过焦弦两端点在 F 上下对称靠拢的图景相一致.

计算时宜先选定目标式 $(x_A-c)(x_B-c)$, 写成 S,P 与常数 c 的组合, 再代入韦达并通分; 分子按 k^2c^2 同类项归并后, 即可显出 $b^2(c^2-a^2)$ 的结构. 先写成整式、再代入, 往往比边代入边展开更易看出 $-b^4$.

13.7.6 弦长与长度和差商积例题

弦长公式是设线法的基本工具, 由此衍生出长度和、差、商、积的各类问题.

例题 13.126 (弦长问题) 已知点 F_1,F_2 分别为椭圆 $\frac{x^2}{2}+y^2=1$ 的左、右焦点, 过 F_2 作倾斜角为 $\frac{\pi}{4}$ 的直线交椭圆于 A,B 两点, 则弦 AB 的长为 __.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

解 答案: $\frac{4\sqrt{2}}{3}$. 圆 $\frac{x^2}{2}+y^2=1$ 的右焦点 $F_2(1,0)$, 因为直线 AB 的倾斜角为 $\frac{\pi}{4}$ 且过点 $F_2(1,0)$, 所以直线 $AB: y=x-1$, 设 $A(x_1,y_1), B(x_2,y_2)$, 联立 $\begin{cases} y=x-1 \\ \frac{x^2}{2}+y^2=1 \end{cases}$, 消去 y 得 $3x^2-4x=0$, 所以 $x_1=0, x_2=\frac{4}{3}$, 所以 $y_1=-1, y_2=\frac{1}{3}$, 所以 $A(0,-1), B\left(\frac{4}{3},\frac{1}{3}\right)$, 所以 $|AB|=\sqrt{\left(\frac{4}{3}-0\right)^2+\left(\frac{1}{3}+1\right)^2}=\frac{4\sqrt{2}}{3}$, 故答案为: $\frac{4\sqrt{2}}{3}$.

例题 13.127 (弦长问题) 已知椭圆 $C: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左、右焦点分别为 F_1, F_2, 离心率为 $\frac{1}{2}$, 椭圆 C 上点 M 满足 $|MF_1|+|MF_2|=4$. (1) 求椭圆 C 的标准方程; (2) 若过坐标原点 $O(0,0)$ 的直线 l 交椭圆 C 于 P, Q 两点, 求线段 PQ 长为 $\sqrt{14}$ 时直线 l 的方程.

知乎@余命数

例题 13.128 (长度和问题) 已知 F 为抛物线 $E: x^2=4y$ 的焦点, 过点 $(0,2)$ 的直线 l 与抛物线 E 交于 A, B 两点, 抛物线 E 在 A, B 两点处的切线交于点 L.

(1) 设 $P(x_0,y_0)$ 是抛物线 E 上一点, 证明: 抛物线 E 在点 P 处的切线方程为 $y=\frac{x_0}{2}x-y_0$, 并利用切线方程求点 L 的纵坐标的值;

(2) 点 C 为抛物线 E 上异于 A, B 的点, 过点 C 作抛物线 E 的切线, 分别与线段 AL, BL 交于 M, N;

(i) 若 $\overrightarrow{LM}=\lambda\overrightarrow{LA}, \overrightarrow{LN}=\mu\overrightarrow{LB}$, 求 $\lambda+\mu$ 的值.

(ii) 证明: $|FA|+|FB|+|FC|>|FL|+|FM|+|FN|$

解 答案: (1) 切线方程 $y=\frac{x_0}{2}x-y_0$, $y_L=-2$; (2)(i) 1; (ii) 见解析. (1) 由题意, 曲线 $y=\frac{x^2}{4}$, 可得 $y'=\frac{x}{2}$, 则 $y'|_{x=x_0}=\frac{x_0}{2}$, 点 $P(x_0,y_0)$ 处的切线方程为 $y=\frac{x_0}{2}(x-x_0)+y_0$, 即 $y=\frac{x_0}{2}x-\frac{x_0^2}{2}+y_0$, 因为 $x_0^2=4y_0$, 代入可得 $y=\frac{x_0}{2}x-y_0$, 设 $A(x_A,y_A), B(x_B,y_B), L(x_L,y_L)$,

例题 13.129 (长度差问题) 已知椭圆 $C: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左、右焦点分别为 F_1, F_2, 且 $|F_1F_2|=2$, 过点 F_2 作两条直线 l_1, l_2, 直线 l_1 与 C 交于 A, B 两点, $\triangle F_1AB$ 的周长为 $4\sqrt{2}$. (1) 求 C 的方程; (2) 若 $\triangle F_1AB$ 的面积为 $\frac{4}{3}$, 求 l_1 的方程; (3) 若 l_2 与 C 交于 M, N 两点, 且 l_1 的斜率是 l_2 的斜率的 2 倍, 求 $|MN|-|AB|$ 的最大值.

知乎@余命数

解 答案: (1) $\frac{x^2}{2}+y^2=1$; (2) $x-y-1=0$ 或 $x+y-1=0$; (3) $\frac{\sqrt{2}}{3}$. (1) 设椭圆的半焦距为 $c(c>0)$, 由题意知 $2c=2$, 所以 $c=1$, $\triangle F_1AB$ 的周长为 $|AF_1|+|AF_2|+|BF_1|+|BF_2|=4a=4\sqrt{2}$,所以 $a=\sqrt{2}$,所以 $b^2=a^2-c^2=1$,故 C 的

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

方程为 $\frac{x^2}{2}+y^2=1$. (2) 易知 l_1 的斜率不为 0,设 $l_1: x=my+1, A(x_1,y_1), B(x_2,y_2)$,联立 $\begin{cases} x=my+1 \\ x^2+2y^2-2=0 \end{cases}$,得 $(m^2+2)y^2+2my-1=0$,所以 $y_1+y_2=-\frac{2m}{m^2+2}, y_1y_2=\frac{-1}{m^2+2}$. 所以 $|y_1-y_2|=\sqrt{(y_1+y_2)^2-4y_1y_2}=\frac{2\sqrt{2}\sqrt{m^2+1}}{m^2+2}$,由 $S_{\triangle F_1AB}=\frac{1}{2}|F_1F_2|\cdot|y_1-y_2|=\frac{2\sqrt{2}\sqrt{m^2+1}}{m^2+2}=\frac{4}{3}$,解得 $m=\pm1$,所以 l_1 的方程为 $x-y-1=0$ 或 $x+y-1=0$. (3) 由 (2) 可知 $|AB|=\sqrt{1+m^2}|y_1-y_2|=\frac{2\sqrt{2}(m^2+1)}{m^2+2}=2\sqrt{2}\left(1-\frac{1}{m^2+2}\right)$, 因为 l_1 的斜率是 l_2 的斜率的 2 倍, 所以 $m\neq0$, 得 $|MN|=2\sqrt{2}\left(1-\frac{1}{4m^2+2}\right)$. 所以 $|MN|-|AB|=2\sqrt{2}\left(\frac{1}{m^2+2}-\frac{1}{4m^2+2}\right)=\frac{3\sqrt{2}m^2}{2m^4+5m^2+2}=\frac{3\sqrt{2}}{2m^2+5+\frac{2}{m^2}}\leq\frac{3\sqrt{2}}{4+5}=\frac{\sqrt{2}}{3}$, 当且仅当 $m=\pm1$ 时, 等号成立, 所以 $|MN|-|AB|$ 的最大值为 $\frac{\sqrt{2}}{3}$.

知乎@余命数

例题 13.130 (长度商问题) 已知抛物线 $C_1: y^2=4x$ 的焦点为椭圆 $C_2: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的右焦点 F, 点 P 为抛物线 C_1 与椭圆 C_2 在第一象限的交点, 且 $|PF|=\frac{5}{3}$. (1) 求椭圆 C_2 的方程; (2) 若直线 l 过点 F, 交抛物线 C_1 于 A,C 两点, 交椭圆 C_2 于 B,D 两点 (A,B,C,D 依次排序), 且 $|AC|-|BD|=\frac{30}{11}$, 求直线 l 的方程.

解 答案: (1) $\frac{x^2}{4}+\frac{y^2}{3}=1$; (2) $\sqrt{2}x\pm y-\sqrt{2}=0$, 即 $y=\pm\sqrt{2}(x-1)$. (1) 由抛物线 $y^2=4x$ 可知: $F(1,0)$, 故由 $|PF|=\frac{5}{3}$ 得: $|PF|=x_P+1=\frac{5}{3}, x_P=\frac{2}{3}$, 故 $y_P^2=\frac{8}{3}$, 则 $P(\frac{2}{3},\frac{2\sqrt{6}}{3})$, 则对于 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 有: $\begin{cases} a^2-b^2=1 \\ \frac{4}{9a^2}+\frac{24}{9b^2}=1 \end{cases}$,解得 $\begin{cases} a^2=4 \\ b^2=3 \end{cases}$,故椭圆方程为: $\frac{x^2}{4}+\frac{y^2}{3}=1$; (2) 过点 F 的直线 l 的斜率不存在时, $l: x=1$、$P(\frac{2}{3},\frac{2\sqrt{6}}{3})$、$1>\frac{2}{3}$, 所以直线 l 在点 P 的右侧, 与两曲线的交点顺序变成 A,B,D,C 的顺序, 不满足题意, 如下图; 所以过点 F 的直线 l 的斜率存在, 故设直线 l 的斜率为 k, 则直线方程为 $y=k(x-1)$, 联立抛物线方程: $\begin{cases} y=k(x-1) \\ y^2=4x \end{cases}$,整理得: $k^2x^2-(2k^2+4)x+k^2=0$, 设 $A(x_1,y_1), C(x_2,y_2)$, 则 $x_1+x_2=2+\frac{4}{k^2}$, 故 $|AC|=x_1+x_2+2=4+\frac{4}{k^2}$, 联立 $\begin{cases} y=k(x-1) \\ \frac{x^2}{4}+\frac{y^2}{3}=1 \end{cases}$,整理得 $(3+4k^2)x^2-8k^2x+4k^2-12=0$, 设 $B(x_3,y_3), D(x_4,y_4)$, 则 $x_3+x_4=\frac{8k^2}{3+4k^2}$, $x_3\cdot x_4=\frac{4k^2-12}{3+4k^2}$, 则 $|BD|=\sqrt{(x_3-x_4)^2+(y_3-y_4)^2}=\sqrt{1+k^2}\cdot\sqrt{(x_3+x_4)^2-4x_3x_4}=\sqrt{1+k^2}\cdot\sqrt{\left(\frac{8k^2}{3+4k^2}\right)^2-\frac{4(4k^2-12)}{3+4k^2}}=\frac{12(1+k^2)}{3+4k^2}$, 又 $|AC|-|BD|=\frac{30}{11}$, 即 $4+\frac{4}{k^2}-\frac{12(1+k^2)}{3+4k^2}=\frac{30}{11}$, 整理得 $38k^4-43k^2-66=0$, 解得 $k^2=2$, 因为 $P(\frac{2}{3},\frac{2\sqrt{6}}{3})$, $F(1,0)$, 而 $1>\frac{2}{3}$, 且 A,B,C,D 依次排序, 所以 $k>0$, 如下图, 故 $k=\sqrt{2}$, 故直线 l 的方程为 $y=\sqrt{2}(x-1)$. 综上, 直线 l 的方程为 $\sqrt{2}x-y-\sqrt{2}=0$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.131 (长度积问题) （2024·高三·北京海淀·开学考试）已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的右顶点为 $A(2,0)$，上顶点为 $B(0,\sqrt{3})$. (1) 求椭圆 C 的方程；(2) 椭圆 C 的左焦点为 F，点 P 为椭圆 C 上不同于顶点的一点，直线 AP,FP 与 y 轴的交点分别为 M,N，若 $|OM|\cdot|ON|=\frac{9}{4}$，求点 P 的横坐标.

解 答案：(1) $\frac{x^2}{4}+\frac{y^2}{3}=1$；(2) 1 或 $-\frac{7}{5}$. (1) 由题设，$\begin{cases}a=2,\\ b=\sqrt{3},\end{cases}$ 所以 C 的方程为 $\frac{x^2}{4}+\frac{y^2}{3}=1$. (2) 法一：设 $P(x_0,y_0)$，所以 $\frac{x_0^2}{4}+\frac{y_0^2}{3}=1\,(-2<x_0<2,x_0\neq 0)$ 所以 $k_{AP}=\frac{y_0}{x_0-2}$，直线 AP 的方程为 $y=\frac{y_0}{x_0-2}(x-2)$. 令 $x=0$，得 $y_M=\frac{-2y_0}{x_0-2}$ 又 $k_{FP}=\frac{y_0}{x_0+1}$，直线 FP 的方程为 $y=\frac{y_0}{x_0+1}(x+1)$ 令 $x=0$，得 $y_N=\frac{y_0}{x_0+1}$，所以 $|OM|\cdot|ON|=\left|\frac{-2y_0}{x_0-2}\right|\left|\frac{y_0}{x_0+1}\right|=\left|\frac{2\cdot3(1-\frac{x_0^2}{4})}{(x_0-2)(x_0+1)}\right|=\frac{9}{4}$，所以 $\left|\frac{3}{2}\frac{(4-x_0^2)}{(x_0-2)(x_0+1)}\right|=\frac{9}{4}$，所以 $\left|\frac{2+x_0}{x_0+1}\right|=\frac{3}{2}$，所以 $x_0=1$ 或 $x_0=-\frac{7}{5}$. 所以点 P 的横坐标为 1 或 $\frac{7}{5}$. 法二：由题意直线 AP 斜率存在，且不为 0. 设直线 AP 的方程为 $y=k(x-2)$. 令 $x=0$，得 $y_M=-2k$ 由 $\begin{cases}y=k(x-2),\\ 3x^2+4y^2=12,\end{cases}$ 得 $(4k^2+3)x^2-16k^2x+16k^2-12=0$. 易得 $\Delta>0$，设 $P(x_0,y_0)\,(-2<x_0<2,x_0\neq 0)$，则 $y_0=k(x_0-2)$，$2x_0=\frac{16k^2-12}{4k^2+3}$，所以 $x_0=\frac{8k^2-6}{4k^2+3}$. 直线 FP 的方程为 $y=\frac{y_0}{x_0+1}(x+1)$.

例题 13.132 (长度的范围与最值问题) 已知双曲线 $C:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的离心率为 2，且经过点 $(2,3)$. (1) 求 C 的方程；(2) 若直线 l 与 C 交于 A，B 两点，且 $\overrightarrow{OA}\cdot\overrightarrow{OB}=0$（点 O 为坐标原点），求 $|AB|$ 的取值范围.

解 答案：(1) $x^2-\frac{y^2}{3}=1$；(2) $[\sqrt{6},+\infty)$. (1) 由题意可得 $\begin{cases}\frac{4}{a^2}-\frac{9}{b^2}=1\\ \frac{a^2+b^2}{a^2}=2\end{cases}$，解得 $a^2=1,b^2=3$，故双曲线方程为 $C:x^2-\frac{y^2}{3}=1$. (2) 当直线 l 斜率不存在时，可设 $A(x_A,y_A),B(x_A,-y_A)$，则 $\overrightarrow{OA}=(x_A,y_A),\overrightarrow{OB}=(x_A,-y_A)$，将其代入双曲线方程 $x_A^2-\frac{y_A^2}{3}=1$，又 $\overrightarrow{OA}\cdot\overrightarrow{OB}=x_A^2-y_A^2=0$，解得 $y_A=\pm\frac{\sqrt{6}}{2}$，此时 $|AB|=2|y_A|=\sqrt{6}$. 当直线 l 斜率存在时，设其方程为 $y=kx+m$，设 $A(x_1,y_1),B(x_2,y_2)$，联立 $\begin{cases}y=kx+m\\ x^2-\frac{y^2}{3}=1\end{cases}\Rightarrow(3-k^2)x^2-2kmx-m^2-3=0,\cdots\cdots$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

例题 13.133 (长度的定值问题) 如图所示, 抛物线 $y^2=2px(p>0)$ 的准线过点 $(-2,3)$, (1) 求抛物线的标准方程; (2) 若角 α 为锐角, 以角 α 为倾斜角的直线经过抛物线的焦点 F, 且与抛物线交于 A、B 两点, 作线段 AB 的垂直平分线 l 交 x 轴于点 P, 证明: $|FP|-|FP|\cos 2\alpha$ 为定值, 并求此定值.

解 (1) 解: 由题意得 $-\frac{p}{2}=-2\therefore p=4\therefore$ 抛物线的方程为 $y^2=8x$ (2) 设 $A\left(x_A,y_A\right),B\left(x_B,y_B\right)$, 直线 AB 的斜率为 $k=\tan\alpha$ 则直线方程为 $y=k\left(x-2\right)$ 将此式代入 $y^2=8x$, 得 $k^2x^2-4\left(k^2+2\right)x+4k^2=0$, 故 $x_A+x_B=\frac{4(k^2+2)}{k^2}$ 设 AB 的中垂线为直线 m, 设直线 m 与 AB 的交点为 $E\left(x_E,y_E\right)$

例题 13.134 (2024·海南模拟预测)已知双曲线 $C:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的实轴长为 $2\sqrt{2}$, 点 $P(2,\sqrt{6})$ 在双曲线 C 上.(1) 求双曲线 C 的标准方程; (2) 过点 P 且斜率为 $2\sqrt{6}$ 的直线与双曲线 C 的另一个交点为 Q, 求 $|PQ|$.

解 答案: (1) $\frac{x^2}{2}-\frac{y^2}{6}=1$; (2) $\frac{20}{7}$. (1) 为双曲线的实轴长为 $2\sqrt{2}$, 所以 $2a=2\sqrt{2}$, 解得: $a=\sqrt{2}$; 又因为点 $P(2,\sqrt{6})$ 在双曲线 C 上, 所以 $\frac{4}{a^2}-\frac{6}{b^2}=1$, 解得: $b=\sqrt{6}$, 所以双曲线的标准方程为: $\frac{x^2}{2}-\frac{y^2}{6}=1$.(2) 设 $P(x_1,y_1),Q(x_2,y_2)$ 由题可得过点 P 且斜率为 $2\sqrt{6}$ 的直线方程为: $y-\sqrt{6}=2\sqrt{6}(x-2)$, 即 $y=2\sqrt{6}x-3\sqrt{6}$, 联立 $\begin{cases}\frac{x^2}{2}-\frac{y^2}{6}=1\\ y=2\sqrt{6}x-3\sqrt{6}\end{cases}$, 消去 y 可得: $7x^2-24x+20=0$, 所以 $x_1+x_2=\frac{24}{7}$, $x_1x_2=\frac{20}{7}$, 所以 $|PQ|=\sqrt{1+k^2}\sqrt{(x_1+x_2)^2-4x_1x_2}=\sqrt{1+24}\sqrt{\left(\frac{24}{7}\right)^2-4\times\frac{20}{7}}=\frac{20}{7}$

例题 13.135 已知抛物线 $E:x^2=2py(p>0)$ 的焦点为 $F(0,2)$.(1) 求 p; (2) 斜率为 2 的直线过点 F, 且与抛物线 E 交于 A,B 两点, 求线段 AB 的长.

解 答案: (1) 4; (2) 40. (1) $\because F(0,2)$ 为抛物线 E 的焦点, $\therefore\frac{p}{2}=2$, 解得: $p=4$.(2) 由 (1) 知: 抛物线 $E:x^2=8y$; $\therefore$ 直线 $AB:y=2x+2$, $\therefore$ 由 $\begin{cases}y=2x+2\\ x^2=8y\end{cases}$ 得: $x^2-16x-16=0$, 设 $A(x_1,y_1),B(x_2,y_2)$, 则 $x_1+x_2=16$、$\therefore y_1+y_2=2x_1+2+2x_2+2=2(x_1+x_2)+4=36$、$\therefore|AB|=y_1+y_2+p=36+4=40$.

例题 13.136 (2024·宁夏银川·银川一中校考模拟) 如图所示, 由半椭圆 $C_1:\frac{x^2}{4}+\frac{y^2}{b^2}=1(y\leqslant 0)$ 和两个半圆 $C_2:(x+1)^2+y^2=1(y\geqslant 0)$、$C_3:(x-1)^2+y^2=1(y\geqslant 0)$ 组成曲线 $C:F(x,y)=0$, 其中点 A_1,A_2 依次为 C_1 的左、右顶点, 点 B 为 C_1 的下顶点, 点 F_1,F_2 依次为 C_1 的左、右焦点. 若点

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

F_1,F_2 分别为曲线 C_2,C_3 的圆心. (1) 求 C_1 的方程;(2) 若过点 F_1,F_2 作两条平行线 l_1,l_2 分别与 C_1,C_2 和 C_1,C_3 交于 M,N 和 P,Q, 求 $|MN|+|PQ|$ 的最小值.

解 答案:(1) $\frac{x^2}{4}+\frac{y^2}{3}=1(y\leqslant 0)$;(2) 5. (1) 由两圆的方程知:圆心分别为 $C_1(-1,0)$、$C_2(1,0)$,即 $F_1(-1,0)$、$F_2(1,0)$.$\therefore b^2+1=4$,解得:$b^2=3$,$\therefore C_1:\frac{x^2}{4}+\frac{y^2}{3}=1(y\leqslant 0)$. (2) 由题意知:$|MN|+|PQ|=|MF_1|+|PF_2|+2$;$\therefore l_1//l_2$,$\therefore$ 由对称性可知:$|MF_1|+|PF_2|$ 为椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 截直线 l_2 的弦长, 设 $l_2:x=my+1$, 其与椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 交于点 (x_1,y_1) 和 (x_2,y_2) 由 $\begin{cases}x=my+1\\ \frac{x^2}{4}+\frac{y^2}{3}=1\end{cases}$ 得: $(3m^2+4)y^2+6my-9=0$, 则 $\Delta=48(3m^2+3)>0$ $\therefore y_1+y_2=\frac{-6m}{3m^2+4}$、$y_1y_2=-\frac{9}{3m^2+4}$、$\therefore |MF_1|+|PF_2|=\sqrt{1+m^2}\cdot\sqrt{(y_1+y_2)^2-4y_1y_2}=\frac{12(m^2+1)}{3m^2+4}=4-\frac{4}{3m^2+4}$,当 $m=0$ 时,$|MF_1|+|PF_2|$ 取得最小值 $4-1=3$,$\therefore |MN|+|PQ|$ 的最小值为 $3+2=5$.

例题 13.137 已知抛物线 $C:y^2=2px$ 经过点 $(2,-2\sqrt{6})$,直线 $l_1:y=kx+m$ $(km\neq 0)$ 与 C 交于 A,B 两点(异于坐标原点 O). (1) 若 $\overrightarrow{OA}\cdot\overrightarrow{OB}=0$,证明:直线 l_1 过定点. (2) 已知 $k=2$,直线 l_2 在直线 l_1 的右侧,$l_1\parallel l_2$,l_1 与 l_2 之间的距离 $d=\sqrt{5}$,l_2 交 C 于 M,N 两点,试问是否存在 m,使得 $|MN|-|AB|=10$? 若存在,求 m 的值;若不存在,说明理由.

解 答案: $m=\frac{31}{24}$. (1) 证明: 将点 $(2,-2\sqrt{6})$ 代入 $y^2=2px$, 得 $24=4p$, 即 $p=6$. 联立 $\begin{cases}y^2=12x,\\ y=kx+m,\end{cases}$ 得 $ky^2-12y+12m=0$, 由 $km\neq 0$, 设 $A(x_1,y_1)$, $B(x_2,y_2)$, 则 $y_1y_2=\frac{12m}{k}$, $x_1x_2=\frac{y_1^2}{12}\cdot\frac{y_2^2}{12}=\frac{(y_1y_2)^2}{144}=\frac{m^2}{k^2}$. 因为 $\overrightarrow{OA}\cdot\overrightarrow{OB}=0$, 所以 $x_1x_2+y_1y_2=\frac{m^2}{k^2}+\frac{12m}{k}=0$ 恒成立,则 $m=-12k$,所以 l_1 的方程为 $y=k(x-12)$, 故直线 l_1 过定点 $(12,0)$. (2) 联立 $\begin{cases}y^2=12x,\\ y=2x+m,\end{cases}$ 得 $4x^2+(4m-12)x+m^2=0$, 则 $\begin{cases}x_1+x_2=-m+3,\\ x_1x_2=\frac{m^2}{4},\end{cases}$ 且 $\Delta=(4m-12)^2-16m^2=48(3-2m)>0$, 即 $m<\frac{3}{2}$, $|AB|=\sqrt{1+2^2}|x_1-x_2|=\sqrt{1+2^2}\sqrt{(x_1+x_2)^2-4x_1x_2}=\sqrt{5}\cdot\sqrt{9-6m}$, 设 $l_2:y=2x+n$, 同理可得 $|MN|=\sqrt{5}\cdot\sqrt{9-6n}$. 因为直线 l_2 在 l_1 的右侧, 所以 $n<m$, 则 $d=\frac{m-n}{\sqrt{5}}=\sqrt{5}$, 即 $n=m-5$. 所以 $|MN|-|AB|=\sqrt{5}\left(\sqrt{9-6(m-5)}-\sqrt{9-6m}\right)=10$, 即 $\sqrt{39-6m}=2\sqrt{5}+\sqrt{9-6m}$, 解得 $m=\frac{31}{24}$,因为 $\frac{31}{24}<\frac{3}{2}$,所以满足条件的 m 存在,$m=\frac{31}{24}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.138　(1) 若 MN 的中点为 H, 直线 OH、MN 的斜率分别为 k_1, k_2, O 为坐标原点, 求 $k_1 \cdot k_2$; (2) 若直线 l_1 与直线 l_2 的交点 T 在直线 $x=\frac{1}{2}$ 上, 且直线 l_1 与直线 l_2 的斜率和为 0, 证明: $\frac{|TP|}{|TM|}=\frac{|TN|}{|TQ|}$.

解 答案:(1) 16. (1) 设双曲线方程为 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$,则 $\begin{cases}\frac{c}{a}=\sqrt{17}\\ \frac{(\sqrt{2})^2}{a^2}-\frac{4^2}{b^2}=1\\ c^2=a^2+b^2\end{cases}$,

解得 $\begin{cases}a=1\\ b=4\end{cases}$,所以 $x^2-\frac{y^2}{16}=1$,设 $M(x_1,y_1)$、$N(x_2,y_2)$、$H(x_0,y_0)$ 因为 M、N 两点都在双曲线 $x^2-\frac{y^2}{16}=1$ 上,所以 $\begin{cases}x_1^2-\frac{y_1^2}{16}=1\\ x_2^2-\frac{y_2^2}{16}=1\end{cases}$,两式作差得 $x_1^2-x_2^2=\frac{y_1^2-y_2^2}{16}$,整理得 $x_0(x_1-x_2)=\frac{y_0(y_1-y_2)}{16}$,则 $k_1\cdot k_2=\frac{y_0(y_1-y_2)}{x_0(x_1-x_2)}=16$; (2) 设 $T(\frac{1}{2},n)$,设直线 MN 的方程为 $y-n=k(x-\frac{1}{2})$、$M(x_1,y_1)$、$N(x_2,y_2)$,联立 $\begin{cases}y-n=k(x-\frac{1}{2})\\ x^2-\frac{y^2}{16}=1\end{cases}$,化简得 $(16-k^2)x^2+(k^2-2kn)x-\frac{1}{4}k^2-n^2+kn-16=0, \Delta=16(4n^2-4kn-3k^2+64)$, 则 $x_1+x_2=-\frac{k^2-2kn}{16-k^2}$, $x_1\cdot x_2=\frac{-\frac{1}{4}k^2-n^2+kn-16}{16-k^2}$, 故 $|TM|=\sqrt{1+k^2}|x_1-\frac{1}{2}|$、$|TN|=\sqrt{1+k^2}|x_2-\frac{1}{2}|$、$|TM|\cdot|TN|=(1+k^2)|x_1-\frac{1}{2}||x_2-\frac{1}{2}|=(1+k^2)|x_1x_2-\frac{1}{2}(x_1+x_2)+\frac{1}{4}|=(1+k^2)|\frac{-\frac{1}{4}k^2-n^2+kn-16}{16-k^2}+\frac{1}{2}\cdot\frac{k^2-2kn}{16-k^2}+\frac{1}{4}|=\frac{(1+k^2)(12+n^2)}{|k^2-16|}$, 由 $k_{PQ}+k_{MN}=0$, 所以 $k_{PQ}=-k$, 从而 $|TP|\cdot|TQ|=\frac{(1+(-k)^2)(12+n^2)}{|(-k)^2-16|}=\frac{(1+k^2)(12+n^2)}{|k^2-16|}$, $\therefore|TM|\cdot|TN|=|TP|\cdot|TQ|$, 即 $\frac{|TP|}{|TM|}=\frac{|TN|}{|TQ|}$.

例题 13.139　(2) 过焦点 F 的直线(斜率存在且不为 0)交抛物线 C 于 A,B 两点, 线段 AB 的中垂线交抛物线的对称轴于点 P, 求 $\frac{|FP|}{|AB|}$

解　答案: $\frac{1}{2}$. (1) 因为抛物线 C 的焦点 F 到准线 l 的距离为 2, 所以 $p=2$, 根据建系方案的不同, 抛物线的标准方程有四种可能, 分别是 $y^2=4x, y^2=-4x, x^2=4y, x^2=-4y$.

(2) 在平面直角坐标系中, 抛物线的位置并不影响 $\frac{|FP|}{|AB|}$ 的取值, 因此不妨取抛物线的方程为 $y^2=4x$, 此时焦点 $F(1,0)$. 根据题意, 直线 AB 的斜率存在且不为 0, 因此设直线 AB 的方程为 $x=my+1$. 与抛物线 $y^2=4x$ 联立, 得关于 y 的一元二次方程 $y^2-4my-4=0$, 则 $\Delta=16m^2+16>0$, 设 $A(x_1,y_1), B(x_2,y_2)$, 则 $y_1+y_2=4m, y_1y_2=-4, |y_1-y_2|=\sqrt{16m^2+16}$.$x_1+x_2=m(y_1+y_2)+2=4m^2+2$, 则 $|AB|=\sqrt{1+m^2}|y_1-y_2|=4(1+m^2)$. 线段 AB 的中点坐标为 $(2m^2+1,2m)$, 中

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

垂线方程为 $y-2m=-m(x-2m^2-1)$, 令 $y=0$, 解得 $x=3+2m^2$, 即中垂线与 x 轴交于 $P(3+2m^2,0)$, 所以 $|FP|=2+2m^2$, 则 $\dfrac{|FP|}{|AB|}=\dfrac{1}{2}$.

例题 13.140 (2024·湖北黄冈·模拟预测) 在生活中, 我们经常看到椭圆, 比如放在太阳底下的篮球, 在地面上的影子就可能是一个椭圆. 已知影子椭圆 $C:\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1(a>b>0)$, C 的上顶点为 A, 两个焦点为 F_1,F_2, 离心率为 $\dfrac{1}{2}$, 过 F_1 且垂直于 AF_2 的直线与 C 交于 D,E 两点, $|DE|=6$, 则 $\dfrac{1}{2DF_1}+\dfrac{1}{AD}$ 的最小值是

解 答案: $\frac{3+2\sqrt{2}}{13}$. $\because$ 椭圆的离心率为 $e=\dfrac{c}{a}=\dfrac{1}{2}$, $\therefore a=2c$, $\therefore b^2=a^2-c^2=3c^2$, $\therefore$ 圆的方程为 $\dfrac{x^2}{4c^2}+\dfrac{y^2}{3c^2}=1$, $3x^2+4y^2=12c^2=0$ 不妨设左焦点为 F_1, 右焦点为 F_2, 如图所示, $\therefore AF_2=a$、$OF_2=c$、$a=2c$, $\therefore\angle AF_2O=\dfrac{\pi}{3}$, $\therefore\triangle AF_1F_2$ 为正三角形, $\therefore$ 过 F_1 且垂直于 AF_2 的直线与 C 交于 D,E 两点, DE 为线段 AF_2 的垂直平分线, $\therefore$ 直线 DE 的斜率为 $\dfrac{\sqrt{3}}{3}$, 斜率倒数为 $\sqrt{3}$, 直线 DE 的方程: $x=\sqrt{3}y-c$, 代入椭圆方程 $3x^2+4y^2-12c^2=0$, 整理得: $13y^2-6\sqrt{3}cy-9c^2=0$, $\Delta=(6\sqrt{3}c)^2+4\times13\times9c^2=6^2\times16\times c^2$,

例题 13.141 (2024·高三·河北·开学考试) 已知椭圆 E 的焦点在 x 轴上, 离心率为 $\frac{\sqrt{5}}{3}$, 对称轴为坐标轴, 且经过点 $\left(2\sqrt{2},\frac{2}{3}\right)$. (1) 求椭圆 E 的方程; (2) 若过 $P(0,1)$ 的直线交椭圆 E 于 C、D 两点, 求 $\left|\frac{CP}{DP}\right|$ 的取值范围.

解 答案: (1) $\frac{x^2}{9}+\frac{y^2}{4}=1$; (2) $[\frac{1}{3},3]$. (1) 依题意, 可设椭圆 E 的方程为 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ $(a>b>0)$. 由 $\frac{c}{a}=\frac{\sqrt{5}}{3}$ 得 $a=\frac{3\sqrt{5}}{5}c$, 又因为 $a^2=b^2+c^2$, 所以 $b=\frac{2\sqrt{5}}{5}c$, 则 $\frac{x^2}{\frac{9}{5}c^2}+\frac{y^2}{\frac{4}{5}c^2}=1$, 因为椭圆经过点 $\left(2\sqrt{2},\frac{2}{3}\right)$, 代入上述方程解得 $c^2=5$, 则 $a^2=9,b^2=4$, 所以椭圆 E 的方程为 $\frac{x^2}{9}+\frac{y^2}{4}=1$.

(2) 由 (1) 可知: $A(0,2),B(0,-2)$, 当斜率不存在时, 若点 C 与 A 重合, D 与 B 重合, 此时 $\left|\frac{CP}{DP}\right|=\left|\frac{AP}{BP}\right|=\frac{1}{3}$. 若点 D 与 A 重合, B 与 C 重合, 则 $\left|\frac{CP}{DP}\right|=\left|\frac{BP}{AP}\right|=3$. 当直线斜率存在时, 设直线 $CD:y=kx+1,C(x_1,y_1),D(x_2,y_2)$, 联立 $\begin{cases}\frac{x^2}{9}+\frac{y^2}{4}=1,\\ y=kx+1,\end{cases}$ 消去 y 可得 $(4+9k^2)x^2+18kx-27=0$, 显然 $\Delta>0$, 则 $x_1+x_2=-\frac{18k}{4+9k^2},x_1x_2=-\frac{27}{4+9k^2}$, 可得 $\frac{(x_1+x_2)^2}{x_1x_2}=-\frac{12k^2}{4+9k^2}=-\frac{4}{3}\left(1-\frac{4}{4+9k^2}\right)$, 整理可得 $\frac{x_1}{x_2}+2+\frac{x_2}{x_1}=-\frac{12k^2}{4+9k^2}=-\frac{4}{3}\left(1-\frac{4}{4+9k^2}\right)$, 因为 $4+9k^2\geqslant4$, 可得 $-\frac{4}{3}\left(1-\frac{4}{4+9k^2}\right)\in\left(-\frac{4}{3},0\right]$, 令 $\frac{x_1}{x_2}=t(t<0)$, 则 $-\frac{4}{3}\leqslant t+\frac{1}{t}+2\leqslant0$, 解得

知乎@余命数

$-3<t<-\frac{1}{3}$, 即 $\frac{x_1}{x_2}\in(-3,-\frac{1}{3})$, 从而 $\left|\frac{x_1}{x_2}\right|\in(\frac{1}{3},3)$. 结合斜率不存在时的端点值, 得 $\left|\frac{CP}{DP}\right|\in[\frac{1}{3},3]$.

例题 13.142　已知抛物线 $C:x^2=2py(p>0)$, F 为 C 的焦点, 过点 F 的直线 l 与 C 交于 H,I 两点, 且在 H,I 两点处的切线交于点 T, 当 l 与 y 轴垂直时, $|HI|=4$. (1) 求 C 的方程; (2) 证明: $|FI|\cdot|FH|=|FT|^2$.

解　答案: (1) $x^2=4y$ (2) 证明见解析. (1) 由题意知 $F\left(0,\frac{p}{2}\right)$, 将 $y=\frac{p}{2}$ 代入 $x^2=2py$, 解得 $x=\pm p$, 所以当 l 与 y 轴垂直时, $|HI|=2p=4$, 所以 $p=2$, 故抛物线 C 的方程为 $x^2=4y$.

(2) 证明: 法一: 根据题意知直线 l 的斜率存在, $F(0,1)$, 设直线 l 的方程为 $y=kx+1$、$H(x_1,y_1)$、$I(x_2,y_2)$, 联立 $\begin{cases}x^2=4y\\y=kx+1\end{cases}$, 得 $x^2-4kx-4=0$, 所以 $\Delta=(-4k)^2+16>0$、$x_1+x_2=4k$、$x_1x_2=-4$. 对 $y=\frac{1}{4}x^2$ 求导, 得 $y'=\frac{1}{2}x$, 所以 $k_{TH}\cdot k_{TI}=\frac{1}{2}x_1\cdot\frac{1}{2}x_2=\frac{1}{4}\times(-4)=-1$, 所以 $TI\perp TH$. 由 $\begin{cases}y-y_1=\frac{1}{2}x_1(x-x_1)\\y-y_2=\frac{1}{2}x_2(x-x_2)\end{cases}$ 得 $\begin{cases}x=2k\\y=-1\end{cases}$, 所以 $T(2k,-1)$. 当 $k=0$ 时, 根据对称性得 $|FI|=|FH|=2$, $|FT|=2$, 所以 $|FI|\cdot|FH|=|FT|^2$; 当 $k\neq0$ 时, $k_{FT}\cdot k_{HI}=-\frac{1}{k}\cdot k=-1$, 所以 $FT\perp HI$, 所以 $\triangle FTI\sim\triangle FHT$, 所以 $\frac{|FT|}{|FH|}=\frac{|FI|}{|FT|}$, 即 $|FI|\cdot|FH|=|FT|^2$. 综上, $|FI|\cdot|FH|=|FT|^2$.

法二: 根据题意知直线 l 的斜率存在, $F(0,1)$, 设直线 l 的方程为 $y=kx+1$、$H(x_1,y_1)$、$I(x_2,y_2)$, 联立 $\begin{cases}x^2=4y\\y=kx+1\end{cases}$, 得 $x^2-4kx-4=0$, 所以 $\Delta=(-4k)^2+16>0$、$x_1+x_2=4k$、$x_1x_2=-4$. 对 $y=\frac{1}{4}x^2$ 求导, 得 $y'=\frac{1}{2}x$, 由 $\begin{cases}y-y_1=\frac{1}{2}x_1(x-x_1)\\y-y_2=\frac{1}{2}x_2(x-x_2)\end{cases}$ 得 $\begin{cases}x=2k\\y=-1\end{cases}$, 所以 $T(2k,-1)$. 因为 $|FH|^2=x_1^2+(y_1-1)^2=4y_1+(y_1-1)^2=(y_1+1)^2$, $|FI|^2=x_2^2+(y_2-1)^2=4y_2+(y_2-1)^2=(y_2+1)^2$, 所以 $|FH|^2\cdot|FI|^2=(y_1+1)^2\cdot(y_2+1)^2=(y_1y_2+y_1+y_2+1)^2=\left(\frac{(x_1x_2)^2}{16}+\frac{x_1^2+x_2^2}{4}+1\right)^2=(4k^2+4)^2$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

例题 13.143 双曲线: $\frac{x^2}{16}-\frac{y^2}{9}=1$ 有一点 P 在双曲线上, 分别过 P 点作渐近线平行线交 x 轴于 A,B, 且 A 在靠近原点的一侧, 过 A 点作 x 轴垂线交以 OB 为直径的圆于点 C, 求 $|OC|$ 的取值范围.

解 答案: 4. 由对称性不妨取点 P 在第一象限, 设 $P(x_0,y_0)$, 则 $\frac{x_0^2}{16}-\frac{y_0^2}{9}=1$, 由双曲线方程可得其渐近线方程为: $y=\pm\frac{3}{4}x$, 则 $PA:y-y_0=\frac{3}{4}(x-x_0)$、$PB:y-y_0=-\frac{3}{4}(x-x_0)$、$\therefore x_A=x_0-\frac{4}{3}y_0$、$x_B=x_0+\frac{4}{3}y_0$、$\because|OC|^2=|OA|\cdot|OB|$、$\therefore|OC|^2=x_A\cdot x_B=x_0^2-\frac{16}{9}y_0^2=16$、$\therefore|OC|=4$.

例题 13.144 已知椭圆 $\omega:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 过点 $A(-2,0)$, 且 $a=2b$. (1) 求椭圆 ω 的方程; (2) 设 O 为原点, 过点 $C(1,0)$ 的直线 l 与椭圆 ω 交于 P, Q 两点, 且直线 l 与 x 轴不重合, 直线 AP, AQ 分别与 y 轴交于 M, N 两点. 求证 $|OM|\cdot|ON|$ 为定值.

解 答案: (1) $\frac{x^2}{4}+y^2=1$; (2) 定值为 $\frac{1}{3}$. (1) 因为椭圆 ω 过点 $A(-2,0)$, 所以 $a=2$. 因为 $a=2b$, 所以 $b=1$. 所以椭圆 ω 的方程为 $\frac{x^2}{4}+y^2=1$. (2) 当直线 l 斜率不存在时, 直线 l 的方程为 $x=1$. 不妨设此时 $P\left(1,\frac{\sqrt{3}}{2}\right)$, $Q\left(1,-\frac{\sqrt{3}}{2}\right)$, 所以直线 AP 的方程为 $y=\frac{\sqrt{3}}{6}(x+2)$, 即 $M\left(0,\frac{\sqrt{3}}{3}\right)$. 直线 AQ 的方程为 $y=-\frac{\sqrt{3}}{6}(x+2)$, 即 $N\left(0,-\frac{\sqrt{3}}{3}\right)$. 所以 $|OM|\cdot|ON|=\frac{1}{3}$. 当直线 l 斜率存在时, 设直线 l 的方程为 $y=k(x-1)$, 由 $\begin{cases}y=k(x-1)\\ \frac{x^2}{4}+y^2=1\end{cases}$, 得 $(4k^2+1)x^2-8k^2x+4k^2-4=0$. 依题意, $\Delta>0$. 设 $P(x_1,y_1)$, $Q(x_2,y_2)$, 则 $x_1+x_2=\frac{8k^2}{4k^2+1}$, $x_1x_2=\frac{4k^2-4}{4k^2+1}$. 又直线 AP 的方程为 $y=\frac{y_1}{x_1+2}(x+2)$, 令 $x=0$, 得点 M 的纵坐标为 $y_M=\frac{2y_1}{x_1+2}$, 即 $M\left(0,\frac{2y_1}{x_1+2}\right)$, 同理 $N\left(0,\frac{2y_2}{x_2+2}\right)$. 所以 $|OM|\cdot|ON|=\left|\frac{4y_1y_2}{(x_1+2)(x_2+2)}\right|=\left|\frac{4k^2(x_1-1)(x_2-1)}{(x_1+2)(x_2+2)}\right|=\left|\frac{4k^2[x_1x_2-(x_1+x_2)+1]}{x_1x_2+2(x_1+x_2)+4}\right|=\left|\frac{4k^2\left(\frac{4k^2-4}{4k^2+1}-\frac{8k^2}{4k^2+1}+1\right)}{\frac{4k^2-4}{4k^2+1}+\frac{16k^2}{4k^2+1}+4}\right|=\left|\frac{4k^2(4k^2-4-8k^2+4k^2+1)}{4k^2-4+16k^2+16k^2+4}\right|=\left|\frac{12k^2}{36k^2}\right|=\frac{1}{3}$. 综上, $|OM|\cdot|ON|$ 为定值, 定值为 $\frac{1}{3}$.

例题 13.145 已知斜率为 2 的直线 l 经过椭圆 $\frac{x^2}{5}+\frac{y^2}{4}=1$ 的右焦点 F_1, 与椭圆相交于 A,B 两点, 求弦 AB 的长.

解 答案: $\frac{5\sqrt{5}}{3}$. 由题意可知: $a=\sqrt{5},b=2,c=\sqrt{a^2-b^2}=1$, 因为直线 l 过椭圆 $\frac{x^2}{5}+\frac{y^2}{4}=1$ 的右焦点 $F_1(1,0)$, 且斜率为 $k=2$, 则直线 l 的方程为 $y=2(x-1)$, 且直线 l 与椭圆必相交, 方法一: 解方程组

知乎@余命数 余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$\begin{cases}\frac{x^2}{5}+\frac{y^2}{4}=1\\ y=2(x-1)\end{cases}$，解得 $\begin{cases}x=0\\ y=-2\end{cases}$ 或 $\begin{cases}x=\frac{5}{3}\\ y=\frac{4}{3}\end{cases}$，不妨令 $A(0,-2)$，$B\left(\frac{5}{3},\frac{4}{3}\right)$，所以 $|AB|=\sqrt{\left(0-\frac{5}{3}\right)^2+\left(-2-\frac{4}{3}\right)^2}=\sqrt{\frac{125}{9}}=\frac{5\sqrt{5}}{3}$；方法二：设 $A(x_1,y_1)$，$B(x_2,y_2)$，联立方程 $\begin{cases}\frac{x^2}{5}+\frac{y^2}{4}=1\\ y=2(x-1)\end{cases}$，消去 y 得 $3x^2-5x=0$，则 $x_1+x_2=\frac{5}{3}$，$x_1x_2=0$. 所以 $|AB|=\sqrt{1+k^2}\sqrt{(x_1+x_2)^2-4x_1x_2}=\sqrt{1+2^2}\sqrt{\left(\frac{5}{3}\right)^2-4\times 0}=\frac{5\sqrt{5}}{3}$；方法三：设 $A(x_1,y_1)$，$B(x_2,y_2)$，联立方程 $\begin{cases}\frac{x^2}{5}+\frac{y^2}{4}=1\\ y=2(x-1)\end{cases}$，消去 x 得 $3y^2+2y-8=0$，则 $y_1+y_2=-\frac{2}{3}$，$y_1y_2=-\frac{8}{3}$.

13.7.7 本节小结

除判别式简化引理外，本节主要把 第13.3节 的韦达与弦长公式组织成可重复的联立步骤，即设线，代入曲线，整理为一元二次，算 Δ，读韦达. 椭圆、双曲线、抛物线共用同一框架，差别只在系数模板.

整理时宜用自然标准化 $b^2x^2+a^2(kx+m)^2=a^2b^2$，避免再乘公分母使 Δ 与韦达失真. 含参判别式可直接写成 $4a^2b^2(b^2+a^2k^2-m^2)$ (定理 13.23)；双曲线还须 $A\neq 0$ (注 13.9). 弦长方面，$|k|\le 1$ 时优先 x 形，$|k|>1$ 时优先 y 形；焦点弦可先用 $|PF_2|=a-ex_0$. 面积可用底高式、坐标行列式或割补，公式 (13-29) 预先消去 $\sqrt{1+k^2}$. 对称式变形常用 S^2-2P、$S^2-4P=\Delta/A^2$、$(x_1-c)(x_2-c)=P-cS+c^2$，并先写成 S,P 的整式再代入.

直线方程取 $y=kx+m$ 还是 $x=ty+n$，往往直接决定运算量. 第13.8节 讨论这两种形式的选用，以及斜率不存在时的写法；韦达之后的代数处理见 第13.9节.

13.7.8 本节习题

13.35. 椭圆 $C:\dfrac{x^2}{25}+\dfrac{y^2}{9}=1$，直线 $\ell:y=x+m$. 试用 定理 13.23 直接写出 Δ，并求使 ℓ 与 C 相切的 m 之值. 再写出相离的 m 范围.

13.36. 双曲线 $C:\dfrac{x^2}{9}-\dfrac{y^2}{16}=1$ ($a^2=9$，$b^2=16$)，直线 $\ell:y=kx+m$ 与 C 交于两点. 用 注 13.9 的 $\Delta=4a^2b^2(m^2+b^2-a^2k^2)$，讨论 $k=\dfrac{b}{a}=\dfrac{4}{3}$ 时 $A=b^2-a^2k^2$ 的取值，并解释此时直线与双曲线的交点个数与渐近线平行性.

13.37. 抛物线 $y^2=4x$，直线 $y=2x-1$. 写出联立后的一元二次方程，求 Δ，韦达两式与弦长 $|AB|$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

13.38. 椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{9}=1$, 过左焦点 $F(-\sqrt{7},0)$, 斜率 k 的弦 $\ell:y=k(x+\sqrt{7})$. 类似 例题 13.125 推证 $(x_A+\sqrt{7})(x_B+\sqrt{7})=-\dfrac{b^4}{b^2+a^2k^2}$, 并解释负号的几何意义, 即焦点内分弦.

13.39. 椭圆 $C:\dfrac{x^2}{25}+\dfrac{y^2}{16}=1$, $\ell:y=x+1$, 此例已在 第13.7.1小节 中计算过韦达两式. 试用联立面积公式 (13-29) 直接求 $\triangle OAB$ 的面积, 并用底高式独立复核.

13.40. 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$, 过右焦点 $F(\sqrt{5},0)$, 斜率为 k 的弦 $\ell:y=k(x-\sqrt{5})$. 用 定理 13.23 直接由公式给出 Δ 的 k 的表达式, 并求使弦长 $|AB|$ 取最大值的 k (提示:弦长 $\sqrt{1+k^2}\dfrac{\sqrt{\Delta}}{4+9k^2}$, 化简为 k 的有理式后再求极值).

13.41. 设二次方程 $Ax^2+Bx+C=0$ 的两根为 x_1,x_2 ($A\neq 0$). 用 $S=-\dfrac{B}{A}$, $P=\dfrac{C}{A}$ 表达 $x_1^2+x_2^2$、$(x_1-x_2)^2$、$x_1^3+x_2^3$, 并验证 $(x_1-x_2)^2=\dfrac{\Delta}{A^2}$.

13.42. 给出一条“面积联立式对第三顶点 $P(x_0,y_0)$ 一般化”的推导: $S_{\triangle PAB}=\dfrac{ab\,|kx_0-y_0+m|\sqrt{b^2+a^2k^2-m^2}}{b^2+a^2k^2}$. 据此公式重新计算 第13.7.4小节 例题中第三顶点改换为左焦点 $F_1(-c,0)$ 时的面积.

13.8 直线方程的形式选择

第13.7节 已介绍设线、联立、读韦达与后处理. 判别式简化引理可直接给出 Δ, 联立面积公式可在乘积中消去 $\sqrt{1+k^2}$. 第一步设线仍有分叉. 同一条非竖直直线既可写为 $y=kx+m$, 也可写为 $x=ty+n$ ($t=1/k$ 为反斜率). 二者只差一个字母, 形式选得不当会显著放大中间系数, 甚至在斜率不存在时直接漏掉一枝.

本节围绕直线形式的选择展开, 先比较斜截式与横截式 (第13.8.1小节), 再处理多动直线时的斜率关系 (第13.8.2小节), 最后讨论以斜截式为主干的含参问题与竖直枝的补验 (第13.8.3小节).

13.8.1 斜截式与横截式的比较

设非竖直直线为 L. 我们把它写成两种等价形式.

- **斜截式** $y=kx+m$, 以 x 为自变量, k 为斜率, m 为 y 截距. 联立椭圆 $b^2x^2+a^2y^2=a^2b^2$ 后二次项为 $b^2+a^2k^2$, 其化简见 定理 13.23; $|k|$ 偏小时系数较小, $|k|$ 偏大时系数膨胀;
- **横截式**, 又称反斜截式, $x=ty+n$, 以 y 为自变量, t 为反斜率 ($t=1/k$), n 为 x 截距. 联立后二次项为 $a^2+b^2t^2$, $|t|$ 偏小时系数较小, 亦即 $|k|$ 偏大时更省.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两式的覆盖范围不同.$y=kx+m$ 覆盖一切非竖直线,但 $k\to\infty$(竖直线)时失效;$x=ty+n$ 覆盖一切非水平线,但 $t\to\infty$ 即水平线时失效. 因此 $|k|<1$ 时宜用 $y=kx+m$,$|k|>1$ 时宜用 $x=ty+n$,$|k|=1$ 时两路径代价相当.

更关键的是斜率不存在的特例. 若题面允许弦取竖直方向,$y=kx+m$ 会漏掉这一枝,须以 $x=x_0$ 独立代入补回,见 第13.8.3小节;而 $x=ty+n$ 在 $t=0$ 时给出 $x=n$,竖直枝一并包含. 对过定点、斜率变化的题型,$x=ty+n$ 通常更稳,代价是横弦 $y=y_0$ 对应 $t\to\infty$ 须另补,但横弦中点多在 x 轴上,较少造成遗漏.

还可结合定点位置. 过定点 P 时,若 $y_P=0$ 则宜选 $x=ty+n$(截距 n 为已知常数),若 $x_P=0$ 则宜选 $y=kx+m$. 下例把同一道过定点弦题用两式各算一遍,对照中间系数,并观察何者可自然容纳竖直弦.

例题 13.146　椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{4}=1$ $(a^2=16,b^2=4)$,动弦 AB 过定点 $P(2,0)$. 分别以 $y=kx+m$ 与 $x=ty+n$ 联立,写出韦达两式与弦中点 M 的坐标,比较两路径的中途系数复杂度,并指出何者自然容纳弦竖直的特例.

知乎@余命

解　因 AB 过 $P(2,0)$,两种设线都据此消去一个参数(P 落在 y 截距或 x 截距之一).

$0=2k+m\implies m=-2k$,即 $y=k(x-2)$. 代入 $4x^2+16y^2=64$,即 $b^2x^2+a^2y^2=a^2b^2$ 形:

$$4x^2+16k^2(x-2)^2=64\implies(4+16k^2)x^2-64k^2\,x+(64k^2-64)=0.$$

知乎@余命数

此处 $A=4+16k^2$、$B=-64k^2$、$C=64k^2-64$. 由 第13.3节 韦达两式:

$$x_A+x_B=\frac{64k^2}{4+16k^2}=\frac{16k^2}{1+4k^2},\qquad x_Ax_B=\frac{64(k^2-1)}{4+16k^2}=\frac{16(k^2-1)}{1+4k^2}.$$

弦中点 M 横坐标 $x_M=\dfrac{x_A+x_B}{2}=\dfrac{8k^2}{1+4k^2}$;因 M 在弦上,$y_M=k(x_M-2)=k\cdot\dfrac{8k^2-2-8k^2}{1+4k^2}=-\dfrac{2k}{1+4k^2}$.

知乎@余命数

$2=t\cdot0+n\implies n=2$,即 $x=ty+2$. 代入 $4x^2+16y^2=64$:

$$4(ty+2)^2+16y^2=64\implies(4t^2+16)y^2+16t\,y+(16-64)=0.$$

即 $(4t^2+16)y^2+16ty-48=0$,自然标准化,除以 4 得 $(t^2+4)y^2+4ty-12=0$. 第13.3节 韦达(y 为根):

$$y_A+y_B=-\frac{4t}{t^2+4},\qquad y_Ay_B=-\frac{12}{t^2+4}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

弦中点 M 纵坐标 $y_M=\dfrac{y_A+y_B}{2}=-\dfrac{2t}{t^2+4}$；因 M 在弦上，$x_M=t\,y_M+2=\dfrac{-2t^2}{t^2+4}+2=\dfrac{8}{t^2+4}$.

$y=kx+m$ 与 $x=ty+n$ 对同一直线满足 $k=1/t$（反斜率关系）. 把 $k=1/t$ 代入路径 A 之 x_M,y_M：

$$x_M=\frac{8/t^2}{1+4/t^2}=\frac{8}{t^2+4},\qquad y_M=-\frac{2/t}{1+4/t^2}=-\frac{2t}{t^2+4},$$

与路径 B 完全吻合. 两路径给出同一答案；代数代价的中途系数复杂度对比如下：

- 路径 A 二次方程 $(1+4k^2)x^2-16k^2\,x+16(k^2-1)=0$. 常数项 $16(k^2-1)$ 含 k^2，三处系数均含 k^2，韦达读后须约去与通分各一次；
- 路径 B 二次方程 $(t^2+4)y^2+4ty-12=0$. 常数项 -12 为静态整数，仅二次项与一次项含参，韦达读后一步即出.

路径 B 中途系数较简，因定点 P 的横坐标 2 直接落在 x 截距 n 上，反斜率 t 才是唯一活跃参数. 定点落在所代入的坐标轴上时，横截式通常更省，与 第13.7.3小节 中按斜率大小选用弦长形式的想法一致.

弦竖直时方程为 $x=2$，路径 A ($y=kx+m$) 要求 $k\to\infty$，系数 $1+4k^2\to\infty$，韦达形式上 $x_A+x_B\to 4$，$x_Ax_B\to 4$，弦长 $|AB|=\sqrt{1+k^2}\,|x_A-x_B|$ 可取极限得 $2\sqrt{3}$，但过程中 $k=\infty$ 形式失效，须独立以 $x=2$ 代入：$4\cdot 4+16y^2=64\implies y^2=3$，弦长 $2\sqrt{3}$. 路径 B 中 $t=0$ 即给出 $x=0\cdot y+2=2$，即 $x=2$ 同一方程，$y_A+y_B=0$，$y_Ay_B=-12/4=-3$，弦长 $|y_A-y_B|=\sqrt{0+4\cdot 3}=2\sqrt{3}$. 路径 B 在 $t=0$ 处自然容纳竖直弦，不另列分类.

13.8.2 多动直线情形下的斜率关系利用

第13.3节 已证明，椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 的弦 AB（斜率 k）与弦中点 M 满足 $k\cdot k_{OM}=-\dfrac{b^2}{a^2}$，其中 O 为椭圆中心. 上式描述单条弦与其弦中点的关系. 题目中的弦往往不止一条. 若同一椭圆内的两条弦被垂直、平行或共点条件串起，我们就可以把上述关系连乘，一次写完整族中点的轨迹.

这条关系中 k 为弦斜率，k_{OM} 为弦中点到中心连线的斜率，二者之积 $-\dfrac{b^2}{a^2}<0$，故 k 与 k_{OM} 反号. 当多条弦由垂直、对称或斜率之和/积为定值等条件相连时，我们就把这种反向关系连乘，以得到中点轨迹之间的结构，不必逐弦联立. 我们先把所有斜率条件写成关于单一参数 k 的约束，再代入中点公式或上述关系一次写完.

我们给出常用的一个网络式.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 13.24 (垂直弦中点的斜率积) 设椭圆 $C:\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1\,(a>b>0)$, l_1,l_2 是过同一椭圆内点 P 的两条弦, 且 $l_1\perp l_2$, 即斜率 $k_1k_2=-1$. 以 M_1,M_2 分别记 l_1,l_2 的中点, 则

$$k_{OM_1}\cdot k_{OM_2}=-\frac{b^4}{a^4}. \tag{13-30}$$

由此 OM_1 与弦 l_1 的方向共轭, OM_2 与弦 l_2 的方向共轭. 两条直径 OM_1、OM_2 相互共轭需要 $k_{OM_1}k_{OM_2}=-\dfrac{b^2}{a^2}$, 由 (13-30) 这要求 $a=b$.

知乎@余命数

证明 由 第13.3节 弦中点斜率关系, 对 l_1,l_2 分别有

$$k_1\cdot k_{OM_1}=-\frac{b^2}{a^2},\qquad k_2\cdot k_{OM_2}=-\frac{b^2}{a^2}.$$

两式相乘, 得

$$(k_1k_2)\cdot(k_{OM_1}k_{OM_2})=\frac{b^4}{a^4}.$$

题设 $l_1\perp l_2\iff k_1k_2=-1$, 这里假设两弦均非竖直与水平, 即斜率均有限非零, 退化情形见本定理后的讨论. 代入上式, 得

$$(-1)\cdot(k_{OM_1}k_{OM_2})=\frac{b^4}{a^4}\implies k_{OM_1}k_{OM_2}=-\frac{b^4}{a^4}.$$

此即 (13-30). 由 $-\dfrac{b^4}{a^4}<0$ 知 OM_1,OM_2 的斜率反号, 二者分别与 l_1,l_2 的方向共轭: 由弦中点斜率关系有 $k_{l_i}k_{OM_i}=-\dfrac{b^2}{a^2}$ $(i=1,2)$. 当 $a=b$ 即椭圆退化为圆时 $-\dfrac{b^4}{a^4}=-1$, 两条半径互相垂直, 与圆的共轭直径互相垂直一致.

讨论退化情形. 若 l_1 竖直, 即 k_1 不存在, 则 $l_2\perp l_1$ 使 l_2 水平 ($k_2=0$). 此时 l_1 为 $x=x_P$, x_P 为 P 之横坐标, 弦 l_1 两端关于 x 轴对称, 中点 $M_1=(x_P,0)$, 故 $k_{OM_1}=0$.l_2 为 $y=y_P$, y_P 为 P 之纵坐标, 两端关于 y 轴对称, 中点 $M_2=(0,y_P)$, 故 k_{OM_2} 不存在 (OM_2 为竖直直径). 形式上 $k_{OM_1}\cdot k_{OM_2}$ 含 $0\cdot\infty$, 解读为 l_1 趋于竖直时 $k_{OM_1}\to0$, 同时 l_2 趋于水平, $k_{OM_2}\to\infty$, 二者乘积的极限仍为 $-b^4/a^4$, 与连续极限意义一致. 此时应退回几何结论, l_1 竖直则中点 M_1 落 x 轴, l_2 水平则中点 M_2 落 y 轴, 不宜强行套用式 (13-30) 的代数形式. □

例题 13.147 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1\,(a^2=9,b^2=4)$, 定点 $P(1,0)$ 在 C 内. 过 P 作两条互相垂直的弦 l_1,l_2 (l_1 斜率 k_1, l_2 斜率 $k_2=-1/k_1$, 假设 k_1 有限非

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

零). 设 M_1, M_2 分别为 l_1, l_2 之中点, 验证 $k_{OM_1}k_{OM_2} = -\dfrac{16}{81}$, 并求 M_1 的轨迹方程.

解

对斜率为 k 的过 $P(1,0)$ 的弦 $y = k(x-1)$. 代入椭圆 $4x^2+9y^2=36$:

$$4x^2 + 9k^2(x-1)^2 = 36 \implies (4+9k^2)x^2 - 18k^2\,x + 9(k^2-4) = 0.$$

此处 $A = 4+9k^2$、$B = -18k^2$、$C = 9(k^2-4)$. 韦达 (第13.3节):

$$x_{A_1} + x_{A_2} = \frac{18k^2}{4+9k^2}, \qquad x_{A_1}x_{A_2} = \frac{9(k^2-4)}{4+9k^2}.$$

中点 M 横坐标 $x_M = \dfrac{x_{A_1}+x_{A_2}}{2} = \dfrac{9k^2}{4+9k^2}$. 由 M 在弦上 $y_M = k(x_M-1) = k\cdot\dfrac{9k^2-4-9k^2}{4+9k^2} = -\dfrac{4k}{4+9k^2}$.

$$k_{OM} = \frac{y_M}{x_M} = \frac{-4k/(4+9k^2)}{9k^2/(4+9k^2)} = -\frac{4}{9k}.$$

套用 定理 13.24, l_1, l_2 垂直 ($k_1k_2=-1$) 时

$$k_{OM_1}k_{OM_2} = -\frac{b^4}{a^4} = -\frac{16}{81}.$$

在本例直接验算: $k_{OM_1} = -\dfrac{4}{9k_1}$, $k_{OM_2} = -\dfrac{4}{9k_2} = -\dfrac{4}{9\cdot(-1/k_1)} = \dfrac{4k_1}{9}$, 故

$$k_{OM_1}k_{OM_2} = -\frac{4}{9k_1}\cdot\frac{4k_1}{9} = -\frac{16}{81}.$$

上述关系可直接读出, 不必分别联立 l_1, l_2.

由 $x_{M_1} = \dfrac{9k_1^2}{4+9k_1^2}$, $y_{M_1} = -\dfrac{4k_1}{4+9k_1^2}$, 把 k_1 视为参数消去. 注意 $\dfrac{y_{M_1}}{x_{M_1}} = -\dfrac{4}{9k_1}$, 即 $k_1 = -\dfrac{4x_{M_1}}{9y_{M_1}}$. 代入 x_{M_1} 表达式, 得

$$x_{M_1} = \frac{9\cdot\dfrac{16x_{M_1}^2}{81y_{M_1}^2}}{4+9\cdot\dfrac{16x_{M_1}^2}{81y_{M_1}^2}} = \frac{\dfrac{16x_{M_1}^2}{9y_{M_1}^2}}{4+\dfrac{16x_{M_1}^2}{9y_{M_1}^2}} = \frac{16x_{M_1}^2}{36y_{M_1}^2+16x_{M_1}^2}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两边乘 $36y_{M_1}^2+16x_{M_1}^2$ 得

$$36x_{M_1}y_{M_1}^2+16x_{M_1}^3=16x_{M_1}^2.$$

若 $x_{M_1}\neq 0$, 两边除 x_{M_1} 得

$$36y_{M_1}^2+16x_{M_1}^2=16x_{M_1},$$

即 $16x_{M_1}^2-16x_{M_1}+36y_{M_1}^2=0$, 配方

$$16\Big(x_{M_1}-\frac{1}{2}\Big)^2+36y_{M_1}^2=4,$$

亦即 $\dfrac{(x_{M_1}-1/2)^2}{1/4}+\dfrac{y_{M_1}^2}{1/9}=1$, 这是以 $(\frac{1}{2},0)$ 为中心, 长半轴 $1/2$, 短半轴 $1/3$ 的椭圆. 因 $P=(1,0)$ 在椭圆 C 之内, 弦须在 l_1 与 C 相交即 $\Delta\geq 0$ 的范围内取. 当 $k_1\in\mathbb{R}\setminus\{0\}$ 取遍时, M_1 走过上述椭圆除去 $x=\frac{1}{2}\pm\frac{1}{2}$ 即 $x=0$ 与 $x=1$ 两端点, 分别对应 $k_1=0$ 与 k_1 不存在的两个退化. 这两点的几何意义, 其中 $k_1=0$ 时 $M_1=O$, 为横弦枝, k_1 不存在时 $M_1=P$, 为竖弦枝, 已在上述 $k_{OM}=-4/(9k)$ 的极限处印证, 不必单独分析(退化情形对照见 定理 13.24 退化情形段).

$P=(1,0)$ 所引垂直弦的中点 M_1 走过一小椭圆, M_2 的轨迹由同一定理刻画, 两轨迹上的对应点满足 $k_{OM_1}k_{OM_2}=-\dfrac{b^4}{a^4}$. 直接代 $k_2=-1/k_1$ 入 x_{M_2},y_{M_2} 公式, 得

$$x_{M_2}=\frac{9/k_1^2}{4+9/k_1^2}=\frac{9}{4k_1^2+9},\qquad y_{M_2}=-\frac{4\cdot(-1/k_1)}{4+9/k_1^2}=\frac{4k_1}{4k_1^2+9}.$$

以 k_2 为参数, 得到形如 $(x_{M_2}=\frac{9k_2^2}{4+9k_2^2},y_{M_2}=-\frac{4k_2}{4+9k_2^2})$ 的参数方程, 即 M_2 走过的是一条同型椭圆.

例题 13.148 (斜率之和定值的统一参数) 椭圆 $C:\dfrac{x^2}{4}+\dfrac{y^2}{1}=1$ ($a^2=4$, $b^2=1$), 过定点 $P(1,0)$ 的两条弦 l_1,l_2 的斜率 k_1,k_2 满足 $k_1+k_2=0$, 即关于 x 轴对称. 用弦中点斜率关系写出 $k_{OM_1}+k_{OM_2}$ 与 $k_{OM_1}k_{OM_2}$ 的关系, 并说明如何把两弦统一到一参 k.

解 由 $k\cdot k_{OM}=-b^2/a^2=-1/4$, 得 $k_{OM}=-1/(4k)$. 对称条件 $k_2=-k_1$ 令 $k_1=k$, $k_2=-k$ ($k\neq 0$), 则

$$k_{OM_1}=-\frac{1}{4k},\qquad k_{OM_2}=-\frac{1}{4(-k)}=\frac{1}{4k},$$

故 $k_{OM_1}+k_{OM_2}=0, k_{OM_1}k_{OM_2}=-1/(16k^2)$. 两弦关于 x 轴对称时, 其中点亦关于 x 轴对称, 这与 $k_{OM_1}+k_{OM_2}=0$ 一致. 我们把它统一到一参, 只对 $l_1:y=k(x-1)$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

联立求 M_1,再由对称直接得 $M_2=(x_{M_1},-y_{M_1})$,不必对 l_2 再联立一次. 当 k 不存在即竖弦 $x=1$ 时,对称条件要求另一弦亦为竖弦,与题面通常要求的两条不同弦冲突,故本条件族默认排除竖直枝. 若题面允许同一竖弦计两次,则中点 $M=(1,0)$ 重合,网络式退化为单点.

13.8.3 斜截式的运用与竖直枝

相对 第13.8.1小节 中的横截式,我们仍以 $y=kx+m$ 为主干,因为它与 定理 13.23 的 $\Delta=4a^2b^2(b^2+a^2k^2-m^2)$ 直接衔接,也与 第13.3节 的弦长 $\sqrt{1+k^2}\,|x_1-x_2|$ 及面积式 (13-29) 同源,后处理不必改形式.

它的覆盖盲点正是斜率不存在的竖直弦. 这也是横截式存在的理由. 若形式上令 $k=\infty$,原公式失效,须另以 $x=x_0$ 代入并独立完成弦长、面积等后处理. 常见失误是只做完有限斜率情形就停笔,把斜率存在当成题面的默认前提;多数题目里竖直枝同样需要检验.

13.8.3.0.1 先一般枝后竖直枝的书写模板

1. **一般枝**. 设 $y=k(x-x_0)$(或 $y=kx+m$ 过定点),联立,读 Δ 与韦达,完成目标量关于 k 或 $u=k^2$ 的表达式与极值分析;
2. **竖直枝**. 独立设 $x=x_0$,代入曲线,求交点弦长 / 面积 / 中点,与一般枝的 $k\to\infty$ 极限对照;
3. **闭环**. 确认竖直枝是否在曲线上取得,是否改写极值结论;不得把 $u\to\infty$ 时仅趋近的情形误读为极值不存在.

下例是一道最短弦题,漏掉竖直枝会得到错误答案,必须独立补验.

例题 13.149 椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{9}=1$ ($a^2=16$, $b^2=9$),过 C 内部定点 $P(2,0)$ 的所有弦中,求弦长 $|AB|$ 的最小值与最大值.

解

弦过 $P(2,0)$,故 $0=2k+m\implies m=-2k$,即弦 $y=k(x-2)$. 代入椭圆 $9x^2+16y^2=144$:

$$9x^2+16k^2(x-2)^2=144\implies(9+16k^2)x^2-64k^2\,x+(64k^2-144)=0.$$

此处 $A=9+16k^2$、$B=-64k^2$、$C=64k^2-144$. 判别式

$$\begin{aligned}\Delta=B^2-4AC&=(-64k^2)^2-4(9+16k^2)(64k^2-144)\\&=4096k^4-4\big(576k^2-1296+1024k^4-2304k^2\big)\\&=4096k^4-4096k^4+6912k^2+5184=6912k^2+5184\end{aligned}$$

$$= 1728(4k^2+3).$$

知乎@余命数

弦长 第13.3节 形式一：

$$|AB| = \sqrt{1+k^2}\cdot\frac{\sqrt{\Delta}}{|A|} = \sqrt{1+k^2}\cdot\frac{\sqrt{1728(4k^2+3)}}{9+16k^2}.$$

写平方(便于稍后求导)：

$$|AB|^2 = (1+k^2)\cdot\frac{1728(4k^2+3)}{(9+16k^2)^2}. \tag{13-31}$$

先看特例.$k=0$ 即弦水平时, 弦就是 x 轴上从 $(-4,0)$ 到 $(4,0)$ 的一段,

知乎@余命数

$$|AB|^2 = 1\cdot\frac{1728\cdot 3}{81} = \frac{5184}{81} = 64, \qquad |AB| = 8,$$

知乎@余命数

与直接令 $y=0$ 代入得 $x^2/16=1 \implies x=\pm 4$, 弦长 8 的结果一致.

令 $u=k^2\geq 0$, $f(u)=(1+u)(3+4u)/(9+16u)^2$, 常数因子 1728 不影响极值, 略去. 对 f 求导, 得

知乎@余命数

$$f'(u) = \frac{(7+8u)(9+16u)^2-(3+7u+4u^2)\cdot 2(9+16u)\cdot 16}{(9+16u)^4}$$
$$= \frac{(7+8u)(9+16u)-32(3+7u+4u^2)}{(9+16u)^3}.$$

知乎@余命数

分子展开：$(7+8u)(9+16u) = 63+112u+72u+128u^2 = 128u^2+184u+63$; $32(3+7u+4u^2) = 96+224u+128u^2$. 二者作差:$128u^2+184u+63-128u^2-224u-96 = -40u-33$. 故

$$f'(u) = \frac{-40u-33}{(9+16u)^3}.$$

对一切 $u\geq 0$, 分子 $-40u-33<0$, 分母 $(9+16u)^3>0$, 故 $f'(u)<0$. 即 $f(u)$ 与相应的 $|AB|^2$ 在 $u\geq 0$ 上严格递减, 因此 $|AB|^2$ 之最大值在 $u=0$ 即 $k=0$ 的水平弦处取得, $|AB|^2_{\max}=1728\cdot 3/81=64$, $|AB|_{\max}=8$; 而 $|AB|^2$ 之最小值在 $u\to+\infty$ 处取得(代数上为极限值, $|AB|^2\to 1728\cdot 4u^2/(16u)^2 = 1728\cdot 4/256 = 27$, $|AB|\to 3\sqrt{3}$).

$u\to+\infty$ 即 k 不存在, 弦趋于竖直 $x=2$. 直接代入 $x=2$ 到椭圆 $9x^2+16y^2=144$：

$$9\cdot 4+16y^2 = 144 \implies 16y^2 = 108 \implies y^2 = \frac{27}{4}, \qquad y = \pm\frac{3\sqrt{3}}{2}.$$

弦两端点为 $(2,\frac{3\sqrt{3}}{2})$ 与 $(2,-\frac{3\sqrt{3}}{2})$, 弦长

知乎@余命数

$$|AB| = 2\cdot\frac{3\sqrt{3}}{2} = 3\sqrt{3}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

与 (13-31) 的极值极限 $u \to \infty$ 完全一致. 故实际最小值被竖直弦 $|AB| = 3\sqrt{3}$ 取得, 有限斜率弦达不到该值.

若只用 $y = kx + m$ 推得 $|AB|^2$ 单调递减, 便会误以为最大值在 $u = 0$ 而最小值不存在(只趋近), 这是典型的漏枝结论. 真实最小值在 $u \to \infty$ 端由竖直枝单独实取, 必须由 $x = 2$ 独立联立验证才能闭环.

例题 13.150 椭圆 $C : \dfrac{x^2}{25} + \dfrac{y^2}{16} = 1$ ($a^2 = 25$, $b^2 = 16$) 与直线 ℓ 交于 A, B, ℓ 过 C 之焦点 $F(3, 0)$. 求 $\triangle OAB$ 的面积, 并涵盖 ℓ 竖直 ($\ell : x = 3$) 的情形.

解

ℓ 过 $(3, 0) \implies m = -3k$, 即 $y = k(x - 3)$. 代入椭圆 $16x^2 + 25y^2 = 400$:

$$16x^2 + 25k^2(x - 3)^2 = 400 \implies (16 + 25k^2)x^2 - 150k^2\,x + (225k^2 - 400) = 0.$$

由 定理 13.23 一步读 $\Delta = 4a^2b^2(b^2 + a^2k^2 - m^2)$, 其中 $m = -3k$, 故 $m^2 = 9k^2$:

$$\Delta = 4 \cdot 25 \cdot 16 \cdot (16 + 25k^2 - 9k^2) = 1600(16 + 16k^2) = 25600(1 + k^2).$$

联立面积公式 (13-29)(顶点为 O):

$$\begin{aligned} S_{\triangle OAB} &= \frac{ab\,|m|\sqrt{b^2 + a^2k^2 - m^2}}{b^2 + a^2k^2} \\ &= \frac{5 \cdot 4 \cdot 3|k| \cdot \sqrt{16(1 + k^2)}}{16 + 25k^2} \\ &= \frac{60|k| \cdot 4\sqrt{1 + k^2}}{16 + 25k^2} \\ &= \frac{240|k|\sqrt{1 + k^2}}{16 + 25k^2}. \end{aligned}$$

A, B 由 $x = 3$ 代入 $16x^2 + 25y^2 = 400$ 给出: $16 \cdot 9 + 25y^2 = 400 \implies 25y^2 = 256 \implies y^2 = \frac{256}{25}$, $y = \pm\frac{16}{5}$. 弦长 $|AB| = 2 \cdot \frac{16}{5} = \frac{32}{5}$; 点 O 到 $\ell : x - 3 = 0$ 之距 $d = 3$. 故

$$S_{\triangle OAB} = \frac{1}{2} \cdot \frac{32}{5} \cdot 3 = \frac{48}{5}.$$

把 $k \to \infty$ 代入非竖直式, 极限为 $\dfrac{240|k|\sqrt{1 + k^2}}{16 + 25k^2} \sim \dfrac{240|k| \cdot |k|}{25k^2} = \dfrac{240}{25} = \dfrac{48}{5}$. 这与竖直情形直接算出的 $\frac{48}{5}$ 一致, 竖直枝是极限之内被代数形式遮蔽的正常值. 我们须独立联立 $x = 3$ 才能确认该值被实际取到. 在本例中, k 任意有限即对应一族过 F 的斜弦, k 不存在对应 $x = 3$ 的竖弦且确实落于 C 上; 在过椭圆内部点的所有弦问题中, 竖直枝是否在 C 内还须再独立检验 ℓ 与 C 是否有交点.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.8.4 本节小结

非竖直直线有两种等价写法 $y=kx+m$ 与 $x=ty+n$ ($t=1/k$). 选用时, 我们先看斜率大小.$|k|<1$ 优先斜截式, $|k|>1$ 优先横截式, 使二次项系数 $b^2+a^2k^2$ 或 $a^2+b^2t^2$ 较小 (第13.8.1小节). 二者盲点互补. 斜截式须另补竖直弦, 横截式须另补水平弦; 过定点且斜率变化时, 定点落在所代入坐标轴上的那一形式通常一并包含竖直或水平枝.

单弦与其弦中点的关系 $k\cdot k_{OM}=-\dfrac{b^2}{a^2}$ (第13.3节) 在多弦问题中可以连乘.定理 13.24 给出, 椭圆内过同点的两垂直弦, 其中点 M_1,M_2 所张直径斜率之积为 $-b^4/a^4$(退化情形须另作几何讨论). 遇到垂直、对称或斜率之和/积为定值等条件, 我们宜先把它们统一到单一参数, 再联立.

以 $y=kx+m$ 为主干配合 定理 13.23 与 式 (13-29), 适合绝大多数有限斜率问题, 但须独立检验竖直枝; 讨论顺序宜先一般后竖直 (第13.8.3小节). 例如最短弦问题 (例题 13.149) 中, 若只看 $u=k^2$ 上 $|AB|^2$ 的单调性, 会误以为最小值不存在, 而竖直枝恰恰取到极小.

联立读出韦达之后, 我们还要处理和与积. 第13.9节 讨论对称式代入、和积关系、点乘双根、整体代换, 以及何时改为直接解点.

知乎@余命数

13.8.5 本节习题

13.43. 椭圆 $C:\dfrac{x^2}{25}+\dfrac{y^2}{9}=1$, 过定点 $P(4,0)$ 的弦 AB. 分别以 $y=kx+m$ 与 $x=ty+n$ 两种形式联立, 求韦达两式与弦中点 M 的轨迹方程; 比较两路径的中途系数复杂度, 并据 P 在椭圆内 ($4^2/25=16/25<1$) 论证 $\Delta\geq 0$ 对一切 k 成立.

13.44. 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$, 定点 $P\left(\frac{3}{5},0\right)$ 在 C 内. 过 P 作两条互相垂直的弦 l_1,l_2, M_1,M_2 分别为其中点. 据 定理 13.24 直接写出 $k_{OM_1}k_{OM_2}$ 之值, 再辅以独立联立验证(不必联立两次, 一次联立得 $k_{OM}=-\frac{4}{9k}$, 由 $k_1k_2=-1$ 即推得结果).

13.45. 椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{9}=1$, 过 C 之右焦点 $F(\sqrt{7},0)$ 的弦 AB. 用 定理 13.23 一步写出 $\Delta(k)$, 并推导 $|AB|^2$ 关于 $u=k^2$ 的表达式; 判定其最大值与最小值分别由何弦取得, 并独立验证竖直弦 $x=\sqrt{7}$ 是否对应 $|AB|$ 的最小或最大值, 即 例题 13.149 的过焦点版本.

13.46. 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$ ($a^2=9$、$b^2=4$、$c=\sqrt{5}$、$e=\sqrt{5}/3$), 过 C 之左焦点 $F_1(-\sqrt{5},0)$ 的弦 AB 设为 $y=k(x+\sqrt{5})$, k 不存在时对应 $x=-\sqrt{5}$. 用 定理 13.23 一步写 $\Delta(k)$, 推 $|AB|^2$ 关于 $u=k^2$ 的表达式. 据 $k=\tan\theta$ 把 $|AB|$ 化为 θ 的有理函数, 验证竖直枝 $\theta=\pi/2$ 即 k 不存在时给出 $|AB|=\dfrac{8}{3}$ (此即椭圆通径

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$2b^2/a$, 一般式证明见第14章 焦弦调和平均定理); 有限枝 $\theta=0$ 即 $k=0$ 的横弦给出 $|AB|=2a=6$, 即长轴长.

13.47. 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$, 动弦 AB 过 $P(1,1)$, P 在 C 内即 $1/9+1/4=13/36<1$. 用 $x=ty+n$ 形式, 由 P 在所设直线上的截距位置得 $1=t\cdot 1+n$, 即 $n=1-t$, 推 $|AB|^2$ 关于 t 的表达式; 验证 $t=0$ 的竖直弦 $x=1$ 与某有限 t 对应的水平弦之弦长. 据此说明横截式如何消去 k 不存在的补枝, 而仍须单独补 $y=y_P=1$ 即 $t=\infty$ 退化枝的可能性.

13.48. 抛物线 $C:y^2=8x$ 开口向右, $2p=8$ 即 $p=4$, 焦点为 $F(2,0)$; 过 F 的弦 AB 斜率为 k, k 不存在时即 $x=2$. 用 $y=k(x-2)$ 联立 $y^2=8x$, 推 $|AB|$ 关于 k 的表达式, 验证其最小值在 k 不存在时取得, 此时为竖直弦 $x=2$, 长 $|AB|=8$, 与 $x=2$ 直接代入 $y^2=16\implies y=\pm 4$ 给出弦长 8 一致. 据此比较抛物线与椭圆的焦弦极值方向, 两者都在过焦且垂直于长轴的弦即通径处取得最短, 抛物线为 $|AB|_{\min}=2p=8$, 椭圆为 $|AB|_{\min}=2b^2/a$, 类比见 例题 13.149; 但抛物线的过焦弦长度无上界, $k\to 0$ 时弦趋近于横轴, 长 $\to\infty$, 而椭圆的过焦弦限于 $2a$(长轴长).

13.9 联立后的代数处理技巧

第13.3节 已给出韦达两式, 第13.7节 介绍设线、联立与读韦达的做法, 第13.8节讨论直线形式的选取. 联立之后, 计算繁简通常取决于对和与积的后续处理. 本节依次讨论下面几类手法. 一是把目标写成对称式直接代入; 二是从几何条件独立得到和积关系; 三是用点乘双根处理非对称式; 四是在点位于曲线上时作整体代换; 五是何时改为直接解出交点. 焦弦的调和平均等极坐标结果见 第14章.

13.9.1 韦达直接代入

目标量若是 x_1+x_2, x_1x_2 的对称多项式, 我们直接把韦达结果代入即可, 不必求根. 做法是先把目标写成 $s=x_1+x_2, p=x_1x_2$ 的整式, 再代入.

这里列出常用公式(第13.7节 已列出, 此处复用).

$$x_1^2+x_2^2=s^2-2p,\qquad (x_1-x_2)^2=s^2-4p=\frac{\Delta}{A^2},\qquad x_1^3+x_2^3=s^3-3sp.$$

若目标含 y_i, 且直线为 $y=kx+m$, 则

$$y_1+y_2=ks+2m,\qquad y_1y_2=k^2p+kms+m^2,$$

从而一切关于 y 的对称式也可转成 s,p 的式子, 仍不解根.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.151　椭圆 $C:\dfrac{x^2}{25}+\dfrac{y^2}{16}=1$, 直线 $\ell:y=x+1$ 与 C 交于 A,B. 求 $x_A^2+x_B^2$ 与 $x_A^3+x_B^3$.

解 由 第13.7节 基础联立例, 得 $41x^2+50x-375=0$, 故

$$s:=x_A+x_B=-\frac{50}{41},\qquad p:=x_Ax_B=-\frac{375}{41}.$$

直接代入, 得

$$x_A^2+x_B^2=s^2-2p=\frac{2500}{1681}+\frac{750}{41}=\frac{2500+750\cdot 41}{1681}=\frac{2500+30750}{1681}=\frac{33250}{1681},$$

$$x_A^3+x_B^3=s^3-3sp=\Big(-\frac{50}{41}\Big)^3-3\Big(-\frac{50}{41}\Big)\Big(-\frac{375}{41}\Big)=-\frac{125000}{68921}-\frac{56250}{1681}.$$

通分 $68921=41^3$:

$$-\frac{125000+56250\cdot 41}{68921}=-\frac{125000+2306250}{68921}=-\frac{2431250}{68921}.$$

全程不解根.

例题 13.152　椭圆 $C:\dfrac{x^2}{16}+\dfrac{y^2}{9}=1$, 直线 $\ell:y=\dfrac{1}{2}x+1$ 与 C 交于 A,B. 求 y_A+y_B 与 y_Ay_B.

解 代入 $9x^2+16y^2=144$, 得

$$\begin{aligned}9x^2+16\Big(\frac{x}{2}+1\Big)^2=144&\\ \Longrightarrow\ 9x^2+16\Big(\frac{x^2}{4}+x+1\Big)=144&\\ \Longrightarrow\ 13x^2+16x-128=0.&\end{aligned}$$

故 $s=-\dfrac{16}{13},p=-\dfrac{128}{13}$. 由 $y=kx+m$ 且 $k=\dfrac{1}{2},m=1$ 可得

$$y_A+y_B=ks+2m=\frac{1}{2}\Big(-\frac{16}{13}\Big)+2=-\frac{8}{13}+2=\frac{18}{13},$$

$$\begin{aligned}y_Ay_B&=k^2p+kms+m^2\\&=\frac{1}{4}\Big(-\frac{128}{13}\Big)+\frac{1}{2}\cdot 1\cdot\Big(-\frac{16}{13}\Big)+1\\&=-\frac{32}{13}-\frac{8}{13}+1\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= -\frac{40}{13} + 1$$
$$= -\frac{27}{13}.$$

也可对 y 消 x 直接读韦达, 结果一致; 当目标本是 y 对称量且 $|k|$ 偏大时, 改用 $x = ty + n$ 更省, 见 第13.8节.

13.9.2 和积关系的独立挖掘

韦达给出的 s, p 来自联立; 题面几何条件往往还能独立给出另一组关于 s, p 的等式. 我们先把条件译成和积约束, 再与韦达联立, 这常比先求 k, m 再回代更短. 我们称这种方法为**和积关系的独立挖掘**.

知乎@余命数 知乎@余命数 知乎@余命数

13.9.2.0.1 标准程序

1. **选定坐标根**. 按设线形式决定以 x 为根还是以 y 为根, 记 $s = x_1 + x_2, p = x_1 x_2$ (或 y 版).
2. **条件 $\to$ 和积**. 把中点, 定比, 斜率, 垂直, 定长等条件写成 s, p 的等式, 暂不引入 k, m.
3. **韦达 $\to$ 参数**. 写出联立后的 $s = -\dfrac{B}{A}, p = \dfrac{C}{A}$ (系数含 k, m).
4. **联立求解**. 用步骤 2 的约束消去未知参数, 或直接求目标对称量.
5. **回写几何量**. 需要时再由 s, p 与直线方程还原斜率, 截距, 弦方程.

两条路径的差别在次序. 先参数后条件, 即设 $y = kx + m$ (或过定点消一参), 展开, 再令中点条件; 先条件后和积, 即由中点坐标直接给出 $s = 2x_M$, 再由弦过中点得斜率关系. 后者通常少一次含参展开.

例题 13.153 椭圆 $C: \dfrac{x^2}{9} + \dfrac{y^2}{5} = 1$ 上弦 AB 的中点为 $M(1, 1)$. 用和积独立挖掘求弦 AB 的斜率 k 与方程, 并与先设线再反解 k 的路径对照.

解

记 $A(x_1, y_1), B(x_2, y_2)$. 中点条件直接给出

$$s_x := x_1 + x_2 = 2, \qquad s_y := y_1 + y_2 = 2.$$

设弦斜率为 k, 则 $y_i - 1 = k(x_i - 1)$, 两端对 $i = 1, 2$ 求和:

$$s_y - 2 = k(s_x - 2) \implies 0 = k \cdot 0,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此式对任意 k 恒成立，故我们改用 第13.3节 的弦中点斜率关系 ($a^2=9, b^2=5$)：

$$k=-\frac{b^2x_M}{a^2y_M}=-\frac{5\cdot 1}{9\cdot 1}=-\frac{5}{9}.$$

弦过 $M(1,1)$：$y-1=-\dfrac{5}{9}(x-1)$，即 $5x+9y-14=0$. 这里由条件到斜率一步完成，未展开联立二次方程.

设 $y=kx+m$ 过 M：$1=k+m\implies m=1-k$. 代入椭圆 $5x^2+9y^2=45$：

$$5x^2+9(kx+1-k)^2=45\implies (5+9k^2)x^2+18k(1-k)x+9(1-k)^2-45=0.$$

中点横坐标 $x_M=\dfrac{s}{2}=-\dfrac{B}{2A}=1$，即 $-B=2A$：

$$\begin{aligned}-18k(1-k)&=2(5+9k^2)\\ \implies -9k+9k^2&=5+9k^2\\ \implies -9k&=5\\ \implies k&=-\frac{5}{9}.\end{aligned}$$

结果相同，但含参展开明显更长. 这里直接调用弦中点斜率关系（它本身是和积约束的几何封装），就是和积关系的捷径.

例题 13.154　椭圆 $C:\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$，斜率 $k=1$ 的弦 AB 满足 $|AB|=\dfrac{4\sqrt{6}}{7}$. 用和积独立挖掘求截距 m 与弦方程，并验证 $s^2-4p=\Delta/A^2$.

解　设 $\ell: y=x+m$，即 $k=1$. 代入 $3x^2+4y^2=12$：

$$3x^2+4(x+m)^2=12\implies 7x^2+8m\,x+(4m^2-12)=0.$$

故 $A=7$、$B=8m$、$C=4m^2-12$，

$$s=-\frac{8m}{7},\qquad p=\frac{4m^2-12}{7}.$$

判别式

$$\begin{aligned}\Delta=B^2-4AC&=(8m)^2-4\cdot 7\cdot(4m^2-12)\\ &=64m^2-112m^2+336\\ &=-48m^2+336\\ &=48(7-m^2).\end{aligned}$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

弦长公式 (第13.3节):

$$|AB| = \sqrt{1+k^2}\,\frac{\sqrt{\Delta}}{|A|} = \sqrt{2}\,\frac{\sqrt{\Delta}}{7}.$$

题设 $|AB| = \dfrac{4\sqrt{6}}{7}$, 故

$$\sqrt{2}\,\frac{\sqrt{\Delta}}{7} = \frac{4\sqrt{6}}{7} \implies \sqrt{\Delta} = 4\sqrt{3} \implies \Delta = 48.$$

知乎@余命数

于是 $48(7-m^2) = 48 \implies m^2 = 6 \implies m = \pm\sqrt{6}$. 弦方程为 $y = x \pm \sqrt{6}$.

取 $m = \sqrt{6}$: $s = -\dfrac{8\sqrt{6}}{7}$, $p = \dfrac{4\cdot 6 - 12}{7} = \dfrac{12}{7}$,

$$s^2 - 4p = \frac{384}{49} - \frac{48}{7} = \frac{384-336}{49} = \frac{48}{49} = \frac{\Delta}{A^2}.$$

知乎@余命数

与 $\Delta = 48$, $A = 7$ 一致. 和积原料 $s^2 - 4p$ 与弦长条件通过 Δ 一一对应.

步骤可归纳为, 先选定 x 为根, 写出含 m 的 s, p 与 $\Delta(m)$; 再由弦长条件得到关于 Δ (因而关于 $s^2 - 4p$) 的等式; 然后解出 m, 回写弦方程. 全程目标是先求出 m. 若先求根再算距离, 代数代价会更大.

例题 13.155 椭圆 $C: \dfrac{x^2}{16} + \dfrac{y^2}{9} = 1$, 过定点 $P(2,0)$ 的弦 AB 满足 $\dfrac{x_A + x_B}{2} = \dfrac{8}{5}$, 即中点横坐标为 $\dfrac{8}{5}$. 用和积挖掘求斜率 k, 并写出弦方程.

解 设 $\ell: y = k(x-2)$ (过 P, $m = -2k$). 代入 $9x^2 + 16y^2 = 144$:

$$9x^2 + 16k^2(x-2)^2 = 144 \implies (9+16k^2)x^2 - 64k^2\,x + (64k^2 - 144) = 0.$$

故

$$s = x_A + x_B = \frac{64k^2}{9+16k^2}.$$

中点横坐标条件 $s/2 = 8/5 \implies s = 16/5$:

知乎@余命数

$$\begin{aligned}
\frac{64k^2}{9+16k^2} &= \frac{16}{5} \\
&\implies 320k^2 = 16(9+16k^2) \\
&\implies 20k^2 = 9 + 16k^2 \\
&\implies 4k^2 = 9 \\
&\implies k = \pm\frac{3}{2}.
\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

弦方程：$y=\frac{3}{2}(x-2)$ 或 $y=-\frac{3}{2}(x-2)$，即 $3x-2y-6=0$ 与 $3x+2y-6=0$.

条件直接约束 s，而韦达把 s 写成 k 的有理式；二者联立一次解出 k. 若改用弦中点斜率关系，中点 $M(\frac{8}{5},y_M)$ 在弦上，于是 $y_M=k(\frac{8}{5}-2)=k(-\frac{2}{5})=-\frac{2k}{5}$. 再由

知乎@余命数

$$k=-\frac{b^2x_M}{a^2y_M}=-\frac{9\cdot(8/5)}{16\cdot(-2k/5)}=\frac{9\cdot 8/5}{32k/5}=\frac{72}{32k}=\frac{9}{4k},$$

知乎@余命数

得 $k^2=\frac{9}{4}$，与上一致. 两种挖掘路径，一是韦达 $s(k)$ 联立给定的 s，二是弦中点斜率关系联立中点在弦上；都是把几何条件先变成和积，或等价的斜率与中点约束，再解参数.

知乎@余命数

13.9.3 点乘双根法

知乎@余命数

对称式可用 s,p 直接代入；非对称式如 $(x_1-a)(x_2-a)$，$\frac{y_1}{x_1-c}+\frac{y_2}{x_2-c}$ 需要把平移后的积一次写出. 点乘双根公式是为此设计.

定理 13.25 (点乘双根公式) 设 $f(x)=Ax^2+Bx+C$ $(A\neq 0)$ 有两根 x_1,x_2. 对任意常数 a，

$$(x_1-a)(x_2-a)=\frac{f(a)}{A}.$$

知乎@余命数

证明 由韦达定理，$x_1+x_2=-\frac{B}{A}$，$x_1x_2=\frac{C}{A}$. 故

$$\begin{aligned}(x_1-a)(x_2-a)&=x_1x_2-a(x_1+x_2)+a^2\\&=\frac{C}{A}+a\frac{B}{A}+a^2\\&=\frac{Aa^2+Ba+C}{A}\\&=\frac{f(a)}{A}.\end{aligned}$$

□

注 13.10
若直线 $\ell:y=kx+m$ 与曲线联立得关于 x 的 $f(x)=0$，且 ℓ 过定点 $A(x_0,y_0)$，则

$$|AP_1|\cdot|AP_2|=(1+k^2)\Big|(x_1-x_0)(x_2-x_0)\Big|=(1+k^2)\Big|\frac{f(x_0)}{A}\Big|.$$

对椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$, 进一步有

$$f(x_0)=a^2b^2\Big(\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}-1\Big)\quad\text{(在自然标准化系数下需按实际 } A,B,C \text{ 核对)},$$

从而

$$|AP_1|\cdot|AP_2|=\frac{a^2b^2(1+k^2)}{b^2+a^2k^2}\Big(\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}-1\Big)$$

当 A 在椭圆外时, 括号为正.$a=b=R$ 时退化为圆幂 $x_0^2+y_0^2-R^2$.

例题 13.156 椭圆 $C:\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$, 定点 $A(3,0)$ 在 C 外. 过 A 斜率为 k 的直线 $\ell:y=k(x-3)$ 与 C 交于 P_1,P_2(设 $\Delta>0$). 求 $|AP_1|\cdot|AP_2|$ 关于 k 的表达式, 并验证 $k=0$ 时结果.

解 $m=-3k$. 联立 $3x^2+4\big(k(x-3)\big)^2=12$:

$$(3+4k^2)x^2-24k^2\,x+(36k^2-12)=0.$$

记 $f(x)=(3+4k^2)x^2-24k^2\,x+(36k^2-12)$, $A=3+4k^2$. 由点乘双根,

$$(x_1-3)(x_2-3)=\frac{f(3)}{A}=\frac{9(3+4k^2)-72k^2+(36k^2-12)}{3+4k^2}=\frac{15}{3+4k^2}.$$

故

$$|AP_1|\cdot|AP_2|=(1+k^2)\cdot\frac{15}{3+4k^2}=\frac{15(1+k^2)}{3+4k^2}.$$

$k=0$ 时 ℓ 为 x 轴, 交点 $(\pm2,0)$, 乘积 $|3-2|\cdot|3+2|=5$, 与 $\dfrac{15}{3}=5$ 一致.

例题 13.157 椭圆 $C:\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$、$F_1(-1,0)$、$F_2(1,0)$. 直线 $\ell:y=k(x+1)$ 过 F_1, 与 C 交于 P_1,P_2 ($k\neq\pm\dfrac{3}{4}$). 求 $k_{F_2P_1}+k_{F_2P_2}$, 即 $\dfrac{y_1}{x_1-1}+\dfrac{y_2}{x_2-1}$.

解 联立 $3x^2+4\big(k(x+1)\big)^2=12$:

$$(3+4k^2)x^2+8k^2\,x+(4k^2-12)=0.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

记 $A=3+4k^2$、$B=8k^2$、$C=4k^2-12$、$f(x)=Ax^2+Bx+C$. 韦达: $s=-B/A$, $p=C/A$. 由点乘双根,

$$(x_1-1)(x_2-1)=\frac{f(1)}{A}=\frac{A+B+C}{A}=\frac{-9+16k^2}{A}.$$

又 $y_i=k(x_i+1)$, 故

$$\frac{y_i}{x_i-1}=k\cdot\frac{x_i+1}{x_i-1}.$$

于是

$$\begin{aligned}\sum\frac{x_i+1}{x_i-1}&=\frac{(x_1+1)(x_2-1)+(x_2+1)(x_1-1)}{(x_1-1)(x_2-1)}\\&=\frac{2p-2}{(x_1-1)(x_2-1)}\\&=\frac{2(p-1)A}{f(1)}.\end{aligned}$$

而 $p-1=\dfrac{C-A}{A}=\dfrac{4k^2-12-(3+4k^2)}{A}=\dfrac{-15}{A}$, $f(1)=-9+16k^2$, 故

$$\sum\frac{x_i+1}{x_i-1}=\frac{2(-15)}{-9+16k^2}=\frac{-30}{16k^2-9}.$$

最终

$$k_{F_2P_1}+k_{F_2P_2}=k\cdot\frac{-30}{16k^2-9}=\frac{-30k}{16k^2-9}.$$

当 $k=1$ 时为 $-\dfrac{30}{7}$. 各 $k_{F_2P_i}$ 本身非对称, 其和却由点乘双根一次写出.

13.9.4 曲线方程的整体代换

点在曲线上时, $b^2x_i^2+a^2y_i^2=a^2b^2$ 可整体替换目标式中的二次组合, 无需先求 s,p 再展开.

例题 13.158　椭圆 $C:\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 上弦 AB 的端点为 $A(x_1,y_1)$, $B(x_2,y_2)$. 证明

$$b^2(x_1^2+x_2^2)+a^2(y_1^2+y_2^2)=2a^2b^2.$$

解　因 $A,B\in C$, 有 $b^2x_1^2+a^2y_1^2=a^2b^2$ 与 $b^2x_2^2+a^2y_2^2=a^2b^2$. 两式相加即得. 这比先求 x_1+x_2, y_1+y_2 再展开更短, 因为曲线方程本身就是整体约束, 无需回到韦达.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 13.159 椭圆 $C:\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$ 外一点 $A(3,0)$，过 A 的动直线与 C 交于 P_1,P_2. 证明

$$|AP_1|\cdot|AP_2|=\frac{15(1+k^2)}{3+4k^2},$$

其中 k 为动直线斜率（设斜率存在）. 并说明 $a=b$ 时退化为圆幂定理.

解 由点乘双根公式与曲线方程整体代换可知，对椭圆外一点 $A(x_0,y_0)$ 与过 A 斜率为 k 的弦，有

$$|AP_1|\cdot|AP_2|=\frac{a^2b^2(1+k^2)}{b^2+a^2k^2}\Big(\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}-1\Big).$$

知乎@余命数

此处 $a^2=4$、$b^2=3$、$x_0=3$、$y_0=0$：$\dfrac{x_0^2}{a^2}-1=\dfrac{9}{4}-1=\dfrac{5}{4}$，故

$$|AP_1|\cdot|AP_2|=\frac{4\cdot3(1+k^2)}{3+4k^2}\cdot\frac{5}{4}=\frac{15(1+k^2)}{3+4k^2}.$$

当 $a=b=R$ 时右端化为 $x_0^2+y_0^2-R^2$，即圆幂，与斜率无关. 这也与 例题 13.156 的直接计算互相印证. 整体代换给出一般公式，点乘双根给出具体验算入口.

知乎@余命数 知乎@余命数 知乎@余命数

13.9.5 设线解点策略

韦达适合处理对称量. 若必须求具体交点坐标，例如代入另一曲线、写两点式、计算非对称函数，我们应直接解二次方程，不必绕道韦达.

知乎@余命数

例题 13.160 (暴力解点) 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$ 与直线 $\ell:x+2y-3=0$ 交于 A,B. 求 A,B 的坐标，并写出弦 AB 的中点.

解 由 ℓ：$x=3-2y$. 代入椭圆：$\dfrac{(3-2y)^2}{9}+\dfrac{y^2}{4}=1$. 乘 36：

$$\begin{aligned}4(3-2y)^2+9y^2&=36\\ \Longrightarrow\ 4(9-12y+4y^2)+9y^2&=36\\ \Longrightarrow\ 36-48y+16y^2+9y^2&=36.\end{aligned}$$

$$25y^2-48y=0\ \Longrightarrow\ y(25y-48)=0.$$

知乎@余命数 知乎@余命数

故 $y=0$ 或 $y=\dfrac{48}{25}$. 对应 $x=3$ 或 $x=3-\dfrac{96}{25}=-\dfrac{21}{25}$. 即 $A=(3,0), B=(-\dfrac{21}{25},\dfrac{48}{25})$ (或对调). 中点

$$M=\Big(\frac{3-\frac{21}{25}}{2},\frac{0+\frac{48}{25}}{2}\Big)=\Big(\frac{54}{50},\frac{24}{25}\Big)=\Big(\frac{27}{25},\frac{24}{25}\Big).$$

题目要求坐标, 解根最直接; 若目标仅是中点, 可读 y 和, 得 $y_A+y_B=\dfrac{48}{25}$, 再由 $x=3-2y$ 得 $x_A+x_B=6-2\cdot\dfrac{48}{25}=\dfrac{54}{25}$, 中点即半和.

例题 13.161 (韦达对照) 仍取椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$ 与直线 $\ell:x+2y-3=0$. 仅求弦 AB 的中点 M, 用韦达直接读出, 并与 例题 13.160 对照.

知乎@余命数

解 由 ℓ:$x=3-2y$. 同上得 $25y^2-48y=0$, 即 $25y^2-48y+0=0$. 韦达(以 y 为根):

$$y_A+y_B=\frac{48}{25},\qquad y_Ay_B=0.$$

中点纵坐标 $y_M=\dfrac{y_A+y_B}{2}=\dfrac{24}{25}$. 又 $x_A+x_B=2\cdot3-2(y_A+y_B)=6-\dfrac{96}{25}=\dfrac{54}{25}$, 故

$$x_M=\frac{x_A+x_B}{2}=\frac{27}{25}.$$

即 $M=(\dfrac{27}{25},\dfrac{24}{25})$, 与暴力解点结果一致, 而且**全程未写出单个交点**. 对照两例可得.

- 目标是坐标 / 两点式 / 非对称函数 $\to$ 暴力解点 (例题 13.160);
- 目标是中点 / 弦长 / 对称幂和 $\to$ 只读韦达 (例题 13.161).

例题 13.162 (必须解点) 椭圆 $C:\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$ 与 $\ell:x+2y-3=0$ 交于 A,B (坐标见 例题 13.160). 求 $\dfrac{y_A}{x_A}+\dfrac{y_B}{x_B}$ 是否可仅用 s,p 完成? 若否, 用已解坐标直接计算.

解 形式上看

$$\frac{y_A}{x_A}+\frac{y_B}{x_B}=\frac{y_Ax_B+y_Bx_A}{x_Ax_B},$$

分子 $y_Ax_B+y_Bx_A$ 一般不是 s_x,s_y,p_x,p_y 的对称式(除非再引入 y_i 与 x_i 的线性关系并整理). 本题有 $x=3-2y$, 故 $y=\dfrac{3-x}{2}$,

$$\frac{y}{x}=\frac{3-x}{2x}=\frac{3}{2x}-\frac{1}{2},$$

$$\sum \frac{y_i}{x_i}=\frac{3}{2}\sum \frac{1}{x_i}-1=\frac{3}{2}\cdot\frac{s}{p}-1.$$

由 例题 13.160：$x_A=3, x_B=-\dfrac{21}{25}$，故

$$s=3-\frac{21}{25}=\frac{54}{25},\qquad p=3\cdot\left(-\frac{21}{25}\right)=-\frac{63}{25},$$

$$\sum \frac{y_i}{x_i}=\frac{3}{2}\cdot\frac{54/25}{-63/25}-1=\frac{3}{2}\cdot\left(-\frac{54}{63}\right)-1=\frac{3}{2}\cdot\left(-\frac{6}{7}\right)-1=-\frac{9}{7}-1=-\frac{16}{7}.$$

直接用坐标：$\dfrac{0}{3}+\dfrac{48/25}{-21/25}=0+\dfrac{48}{-21}=-\dfrac{16}{7}$，一致. 本例说明，经线性关系整理后，某些看似非对称的式子仍可回到 s/p；若整理失败，则暴力解点是正道，不必硬套韦达.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

13.9.6 面积问题例题

圆锥曲线中的三角形面积与四边形面积是设线法代数处理的典型应用，涉及分割法、面积坐标化、共角等角模型、对顶角模型、对角线垂直模型等.

例题 13.163 (三角形的面积问题之分割法) 已知椭圆 C 的对称中心为坐标原点，对称轴为坐标轴，焦点在 y 轴上，离心率 $e=\frac{1}{2}$，且过点 $P(3,2)$，

例题 13.164 (三角形、四边形的面积问题之面积坐标化) 如图，已知双曲线 $C:x^2-\frac{y^2}{3}=1$ 的左右焦点分别为 F_1,F_2，若点 P 为双曲线 C 在第一象限上的一点，且满足 $|PF_1|+|PF_2|=8$，过点 P 分别作双曲线 C 两条渐近线的平行线 PA,PB 与渐近线的交点分别是 A 和 B.

(1) 求四边形 $OAPB$ 的面积；

(2) 若对于更一般的双曲线 $C':\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$，点 P 为双曲线 C' 上任意一点，过点 P' 分别作双曲线 C' 两条渐近线的平行线 $P'A',P'B'$ 与渐近线的交点分别是 A' 和 B'，请问四边形 $OA'P'B'$ 的面积为定值吗？若是定值，求出该定值(用 a,b 表示该定值)；若不是定值，请说明理由.

解 答案：(1) $\frac{\sqrt{3}}{2}$；(2) 是定值，$\frac{1}{2}ab$. (1) 因为双曲线 $C:x^2-\frac{y^2}{3}=1$，由双曲线的定义可得 $|PF_1|-|PF_2|=2$，又因为 $|PF_1|+|PF_2|=8$，$\therefore |PF_1|=5,|PF_2|=3$，因为 $|F_1F_2|=2\sqrt{1+3}=4$，所以，$|PF_1|^2+|F_1F_2|^2=|PF_2|^2$，$\therefore PF_2\perp x$ 轴，$\therefore$ 点 P 的横坐标为 $x_P=2$，所以，$2^2-\frac{y_P^2}{3}=1$，$\therefore y_P>0$，可得 $y_P=3$，即点 $P(2,3)$，过点 P 且与渐近线 $y=-\sqrt{3}x$ 平行的直线的方程为 $y-3=-\sqrt{3}(x-2)$，联立

$\begin{cases}y=\sqrt{3}x\\y-3=-\sqrt{3}(x-2)\end{cases}$，解得 $\begin{cases}x=1+\frac{\sqrt{3}}{2}\\y=\sqrt{3}+\frac{3}{2}\end{cases}$，即点 $B\left(1+\frac{\sqrt{3}}{2},\sqrt{3}+\frac{3}{2}\right)$，直线 OP 的方程为 $3x-2y=0$，点 B 到直线 OP 的距离为 $d=\frac{|3x_B-2y_B|}{\sqrt{13}}=\frac{\sqrt{3}}{2\sqrt{13}}$，且 $|OP|=\sqrt{13}$，因此，四边形 $OAPB$ 的面积为 $S_{\square OAPB}=2S_{\triangle OBP}=|OP|\cdot d=\frac{\sqrt{3}}{2}$；(2) 四边形 $OA'P'B'$ 的面积为定值 $\frac{1}{2}ab$，理由如下：设点 $P'(x_0,y_0)$，双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$ 的渐近线方程为 $y=\pm\frac{b}{a}x$，则直线 $P'B'$ 的方程为 $y-y_0=-\frac{b}{a}(x-x_0)$，联立 $\begin{cases}y-y_0=-\frac{b}{a}(x-x_0)\\y=\frac{b}{a}x\end{cases}$，解得 $\begin{cases}x=\frac{x_0}{2}+\frac{a}{2b}y_0\\y=\frac{y_0}{2}+\frac{b}{2a}x_0\end{cases}$，即点 $B'\left(\frac{x_0}{2}+\frac{a}{2b}y_0,\frac{y_0}{2}+\frac{b}{2a}x_0\right)$，

例题 13.165 (三角形的面积比问题之共角、等角模型)（2024·陕西西安·一模）已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 的短轴长等于焦距，且过点 $(2,1)$ (1) 求椭圆 C 的方程；(2) P 为直线 $y=2\sqrt{3}$ 上一动点，记椭圆 C 的上下顶点为 A,B，直线 PA,PB 分别交椭圆 C 于点 M,N；当 $\triangle PMN$ 与 $\triangle PAB$ 的面积之比为 $\frac{3}{4}$ 时，求直线 MN 的斜率.

知乎@余命数

解 答案：(1) $\frac{x^2}{6}+\frac{y^2}{3}=1$；(2) 0. (1) 由题意可得 $\begin{cases}a^2-b^2=c^2\\2b=2c\\\frac{4}{a^2}+\frac{1}{b^2}=1\end{cases}$，解得 $a=\sqrt{6},b=c=\sqrt{3}$，所以椭圆的方程为 $\frac{x^2}{6}+\frac{y^2}{3}=1$.

(2) 因为 $A(0,\sqrt{3}),B(0,-\sqrt{3})$，设 $P(t,2\sqrt{3}),(t\neq 0),M(x_1,y_1),N(x_2,y_2)$，则直线 PA 的方程为 $y=\frac{\sqrt{3}}{t}x+\sqrt{3}$，联立 $\begin{cases}y=\frac{\sqrt{3}}{t}x+\sqrt{3}\\\frac{x^2}{6}+\frac{y^2}{3}=1\end{cases}$ 消去 y 可得 $(t^2+6)x^2+12tx=0$，$\Delta=(12t)^2>0$，解得 $x_1=\frac{-12t}{t^2+6}$，代入直线方程可得 $y_1=\frac{\sqrt{3}t^2-6\sqrt{3}}{t^2+6}$，故 $M\left(\frac{-12t}{t^2+6},\frac{\sqrt{3}t^2-6\sqrt{3}}{t^2+6}\right)$. 直线 PB 的方程为 $y=\frac{3\sqrt{3}}{t}x-\sqrt{3}$，由 $\begin{cases}y=\frac{3\sqrt{3}}{t}x-\sqrt{3}\\\frac{x^2}{6}+\frac{y^2}{3}=1\end{cases}$ 消去 y 可得 $(t^2+54)x^2-36tx=0,\Delta=(36t)^2>0$，解得 $x_2=\frac{36t}{t^2+54},y_2=\frac{-\sqrt{3}t^2+54\sqrt{3}}{t^2+54}$，故 $N\left(\frac{36t}{t^2+54},\frac{-\sqrt{3}t^2+54\sqrt{3}}{t^2+54}\right)$. 设 $\triangle PMN$ 与 $\triangle PAB$ 的面积分别为 S_1,S_2，则 $\frac{S_1}{S_2}=\frac{\frac{1}{2}|PM|\cdot|PN|\sin\angle MPN}{\frac{1}{2}|PA|\cdot|PB|\sin\angle APB}$. 因为 $\angle MPN=\angle APB$，且 P,A,M 三点共线，P,B,N 三点共线，结合距离公式化简可得 $\frac{S_1}{S_2}=\frac{|PM|\cdot|PN|}{|PA|\cdot|PB|}=\frac{|t-x_1|}{|t|}\cdot\frac{|t-x_2|}{|t|}=\frac{\left(t+\frac{12t}{t^2+6}\right)\left(t-\frac{36t}{t^2+54}\right)}{t^2}=\frac{(t^2+18)^2}{(t^2+6)(t^2+54)}$. 由 $\frac{(t^2+18)^2}{(t^2+6)(t^2+54)}=\frac{3}{4}$ 化简解得 $(t^2-18)^2=0\Rightarrow t^2=18\Rightarrow t=\pm 3\sqrt{2}$. 当 $t=3\sqrt{2}$ 时，$M\left(-\frac{3\sqrt{2}}{2},\frac{\sqrt{3}}{2}\right),N\left(\frac{3\sqrt{2}}{2},\frac{\sqrt{3}}{2}\right)$，$MN$ 的斜率为 $\frac{\frac{\sqrt{3}}{2}-\frac{\sqrt{3}}{2}}{\frac{3\sqrt{2}}{2}+\frac{3\sqrt{2}}{2}}=0$；同理当 $t=-3\sqrt{2}$ 时率也为 0. 故直线 MN 的斜率为 0.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 13.166 (三角形的面积比问题之对顶角模型) (2024·高三·山西吕梁·开学考试) 已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 过点 $A(2,\sqrt{2})$, 且 C 的右焦点为 $F(2,0)$. (1) 求 C 的方程; (2) 设过点 $(4,0)$ 的一条直线与 C 交于 P,Q 两点, 且与线段 AF 交于点 S. (i) 证明: S 到直线 FP 和 FQ 的距离相等; (ii) 若 $\triangle APS$ 的面积等于 $\triangle FQS$ 的面积, 求 Q 的坐标.

解 (1) 根据题意有 $\frac{4}{a^2}+\frac{2}{b^2}=1$, 且由椭圆的几何性质可知 $a^2=b^2+c^2=b^2+4$, 所以 $a^2=8,b^2=4$, 所以 C 的方程为 $\frac{x^2}{8}+\frac{y^2}{4}=1$. (2)(i) 显然 PQ 的斜率存在, 设 PQ 的方程为 $y=k(x-4)$, 代入 C 的方程有: $(2k^2+1)x^2-16k^2x+32k^2-8=0$, 其中 $\Delta>0$. 设 $P(x_1,y_1),Q(x_2,y_2)$, 则 $x_1+x_2=\frac{16k^2}{2k^2+1},x_1x_2=\frac{32k^2-8}{2k^2+1}$,

知乎@余命数

例题 13.167 (四边形的面积问题之对角线垂直模型) (2024·湖北·模拟预测) 已知椭圆 $E:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的上顶点为 B, 右焦点为 F, 点 B,F 都在直线 $3x+\sqrt{3}y-3=0$ 上. (1) 求椭圆 E 的标准方程; (2) 若圆 $x^2+y^2=1$ 的两条相互垂直的切线 l_1,l_2 均不与坐标轴垂直, 且直线 l_1,l_2 分别与 E 相交于点 A,C 和 B,D, 求四边形 $ABCD$ 面积的最小值.

解 (1) 找椭圆 E 的半焦距为 c, 由已知点 B 的坐标为 $(0,b)$, 点 F 的坐标为 $(c,0)$, 因为点 B,F 都在直线 $3x+\sqrt{3}y-3=0$ 上, 所以 $\sqrt{3}b-3=0,3c-3=0$, 又 $a^2=b^2+c^2$, 所以 $a=2,b=\sqrt{3},c=1$, 所以椭圆 E 的方程为: $\frac{x^2}{4}+\frac{y^2}{3}=1$, (2) 由题知 l_1,l_2 的斜率存在且不为 0. 设 $l_1:y=kx+m(k\neq0)$. 因为 l_1 与圆 $x^2+y^2=1$ 相切, 所以 $\frac{|m|}{\sqrt{1+k^2}}=1$, 得 $m^2=1+k^2$. 联立 $l_1:y=kx+m(k\neq0)$ 与 E 的方程 $\frac{x^2}{4}+\frac{y^2}{3}=1$, 可得 $(3+4k^2)x^2+8kmx+4m^2-12=0$, 设 $A(x_1,y_1),C(x_2,y_2)$, $\Delta=64k^2m^2-16(3+4k^2)(m^2-3)=48(3+4k^2-m^2)>0\Leftrightarrow3+4k^2>m^2$, 则 $x_1+x_2=\frac{-8km}{3+4k^2},x_1x_2=\frac{4m^2-12}{3+4k^2}$. 所以 $|AC|=\sqrt{1+k^2}|x_1-x_2|=\sqrt{1+k^2}\sqrt{(x_1+x_2)^2-4x_1x_2}=\frac{4\sqrt{3}\sqrt{(1+k^2)(4k^2+3-m^2)}}{3+4k^2}$, 将 $m^2=1+k^2$ 代入, 可得 $|AC|=\frac{4\sqrt{3}\sqrt{(1+k^2)(3k^2+2)}}{3+4k^2}$. 用 $\frac{1}{k}$ 换 k, 可得 $|BD|=\frac{4\sqrt{3}\sqrt{(1+k^2)(3+2k^2)}}{3k^2+4}$,

例题 13.168 (四边形的面积问题之一解四边形) (2024·辽宁·模拟预测) 给出如下的定义和定理:
定义: 若直线 l 与抛物线 Γ 有且仅有一个公共点 P, 且 l 与 Γ 的对称轴不平行, 则称直线 l 与抛物线 Γ 相切, 公共点 P 称为切点.
定理: 过抛物线 $y^2=2px$ 上一点 (x_0,y_0) 处的切线方程为 $y_0y=px_0+px$.
完成下述问题:

已知抛物线 $\Gamma: y^2=x$, 焦点为 F, 过 Γ 外一点 Q(不在 x 轴上), 作 Γ 的两条切线, 切点分别为 A, B, (A, B 在 x 轴两侧) 直线 QA, QB 分别交 y 轴于 C, D 两点,
(1) 若 $|AF|=\frac{5}{4}$, 求线段 CF 的长度;
(2) 若点 Q 在直线 $x=-1$ 上, 证明直线 AB 过定点, 并求出该定点;
(3) 若点 Q 在曲线 $y^2=\frac{7}{8}(x+1), (-1\leqslant x\leqslant 0)$ 上, 求四边形 $ACDB$ 的面积的范围.

解 答案: $\frac{\sqrt{5}}{4}$. (1) 由题意知 $F(\frac{1}{4},0)$, 直线 QA, QB 的斜率均不为零, 其斜率都存在且异号, 设 $A(x_1,y_1)$, 因为 $|AF|=x_1+\frac{1}{4}=\frac{5}{4}\Rightarrow x_1=1, y^2=x=1\Rightarrow y=\pm 1$, 不妨设 $A(1,1)$, 则 AQ 方程为 $y=\frac{1}{2}(x+1)$, 即 $y=\frac{1}{2}x+\frac{1}{2}$、$C(0,\frac{1}{2})$、$|FC|=\sqrt{(\frac{1}{4})^2+(\frac{1}{2})^2}=\frac{\sqrt{5}}{4}$, 所以线段 CF 的长度为 $\frac{\sqrt{5}}{4}$.

例题 13.169　直线 BA_1, BA_2 的斜率满足 $k_{BA_1}\cdot k_{BA_2}=-\frac{1}{4}$. (1) 求椭圆 C 的方程; (2) 已知椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 上点 (x_0,y_0) 处的切线方程是 $\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1$. 若点 P 为直线 $l: x=\frac{4\sqrt{3}}{3}$ 上的动点, 过点 P 作椭圆 C 的切线 PM, PN, 切点分别为 M, N, 求 $\triangle PMN$ 面积的最小值.

解 答案: $\frac{\sqrt{3}}{6}$. (1) 由椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 上顶点为 $B(0,1)$, 可得 $b=1$, 因为 $k_{BA_1}\cdot k_{BA_2}=-\frac{1}{4}$, $A_1(-a,0), A_2(a,0)$, 所以 $\frac{1}{a}\cdot\frac{1}{-a}=\frac{1}{-a^2}=-\frac{1}{4}$, 所以 $a^2=4$, 所以椭圆 C 的方程为 $\frac{x^2}{4}+y^2=1$. (2) 设 $M(x_1,y_1), N(x_2,y_2)$, $P\left(\frac{4\sqrt{3}}{3},n\right)$, 则椭圆 C 在点 M, N 的切线方程分别为 $\frac{x_1x}{4}+y_1y=1$, $\frac{x_2x}{4}+y_2y=1$, 又 P 在两条切线上, 则 $\frac{\sqrt{3}x_1}{3}+ny_1=1$, $\frac{\sqrt{3}x_2}{3}+ny_2=1$, 则直线 MN 的方程为 $\frac{\sqrt{3}}{3}x+ny=1$, 由 $\begin{cases}\frac{\sqrt{3}}{3}x+ny=1\\ \frac{x^2}{4}+y^2=1\end{cases}$ 整理得 $(3n^2+4)y^2-6ny-1=0$, 则 $y_1+y_2=\frac{6n}{3n^2+4}, y_1y_2=\frac{-1}{3n^2+4}$, 则 $|MN|=\sqrt{1+3n^2}\sqrt{(y_1+y_2)^2-4y_1y_2}=\sqrt{1+3n^2}\sqrt{\left(\frac{6n}{3n^2+4}\right)^2+\frac{4}{3n^2+4}}=\frac{4(3n^2+1)}{3n^2+4}$, 又点 P 到直线 MN 的距离 $d=\frac{\left|\frac{\sqrt{3}}{3}\cdot\frac{4\sqrt{3}}{3}+n^2-1\right|}{\sqrt{1+3n^2}}=\frac{\sqrt{3}}{3}\sqrt{1+3n^2}$, 则 $\triangle PMN$ 的面积为 $S_{\triangle PMN}=\frac{1}{2}|MN|\cdot d=\frac{1}{2}\times\frac{4(3n^2+1)}{3n^2+4}\times\frac{\sqrt{3}}{3}\sqrt{1+3n^2}=\frac{2\sqrt{3}}{3}\cdot\frac{(1+3n^2)\sqrt{3n^2+1}}{3n^2+4}$, 令 $m=\sqrt{1+3n^2}$, $m\geq 1$, 则 $S_{\triangle PMN}=\frac{2\sqrt{3}}{3}\cdot\frac{m^3}{m^2+3}=\frac{2\sqrt{3}}{3}\cdot\frac{1}{\frac{1}{m}+\frac{3}{m^3}}$, $m\geq 1$, 则 $y=\frac{1}{m}+\frac{3}{m^3}$ 在 $[1,+\infty)$ 上单调递减, 则 $S_{\triangle PMN}=\frac{2\sqrt{3}}{3}\cdot\frac{1}{\frac{1}{m}+\frac{3}{m^3}}$ 在 $[1,+\infty)$ 上单调递增, 所以 $S_{\triangle PMN}\geq\frac{2\sqrt{3}}{3}\cdot\frac{1}{1+3}=\frac{\sqrt{3}}{6}$, 当且仅当 $n=0$ 即点 P 坐标为 $P\left(\frac{4\sqrt{3}}{3},0\right)$ 时等号成立, 则 $\triangle PMN$ 面积的最小值为 $\frac{\sqrt{3}}{6}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 13.170 已知双曲线 $\Gamma: x^2-\frac{y^2}{4}=1$ 与直线 $l: y=x+1$ 交于 A、B 两点（A 在 B 左侧），过点 A 的两条关于 l 对称的直线 l_1、l_2 分别交双曲线 Γ 于 C、D 两点，其中 C 在右支，D 在左支. (1) 设直线 l_1 的斜率为 k_1，直线 l_2 的斜率为 k_2，求 $k_1\cdot k_2$ 的值；(2) 若直线 CD 与双曲线 Γ 在点 B 处的切线交于点 P，求 $\triangle ABP$ 的面积.

解 答案：(1) 1；(2) $\frac{32}{15}$. (1) 由题意知直线 l 斜率为 1，$\therefore$ 直线 l 的倾斜角 $\alpha=\frac{\pi}{4}$，设直线 l_1、l_2 的倾斜角分别为 θ_1、θ_2 (θ_1、$\theta_2\in(0,\pi)$)，直线 l_1、l_2 关于直线 l 对称，$\therefore\theta_1+\theta_2=2\alpha=\frac{\pi}{2}$，$\therefore k_1\cdot k_2=\tan\theta_1\cdot\tan\theta_2=\tan\theta_1\cdot\tan(\frac{\pi}{2}-\theta_1)=\frac{\sin\theta_1}{\cos\theta_1}\cdot\frac{\sin(\frac{\pi}{2}-\theta_1)}{\cos(\frac{\pi}{2}-\theta_1)}=1$.

(2) 联立 $\begin{cases}x^2-\frac{y^2}{4}=1\\ y=x+1\end{cases}\Rightarrow A(-1,0),B(\frac{5}{3},\frac{8}{3})$，$\therefore$ 双曲线 Γ 在点 B 处的切线方程为 $\frac{5}{3}x-\frac{2}{3}y=1$. 不妨设直线 CD 为 $m(x+1)+ny=1$、$C(x_1,y_1)$、$D(x_2,y_2)$，联立 $\begin{cases}x^2-\frac{y^2}{4}=1\\ m(x+1)+ny=1\end{cases}$ 得 $\begin{cases}4[(x+1)-1]^2-y^2-4=0\\ m(x+1)+ny=1\end{cases}$，$\Rightarrow 4(x+1)^2-8(x+1)[m(x+1)+ny]-y^2=0$ 整理得 $(\frac{y}{x+1})^2+8n\cdot\frac{y}{x+1}+8m-4=0$，将等式看作关于 $\frac{y}{x+1}$ 的方程：两根之和 $\frac{y_1}{x_1+1}+\frac{y_2}{x_2+1}=-8n$，两根之积 $\frac{y_1}{x_1+1}\cdot\frac{y_2}{x_2+1}=8m-4$，而其中 $k_1\cdot k_2=k_{AC}\cdot k_{AD}=\frac{y_1}{x_1+1}\cdot\frac{y_2}{x_2+1}=8m-4$，由 (1) 得 $k_1\cdot k_2=1$，$\therefore m=\frac{5}{8}$ $\therefore$ 直线 CD 为 $\frac{5}{8}(x+1)+ny=1$，过定点 $(\frac{3}{5},0)$，又 $\because$ 双曲线 Γ 在点 B 处的切线方程为 $\frac{5}{3}x-\frac{2}{3}y=1$，过点 $(\frac{3}{5},0)$，$\therefore P(\frac{3}{5},0)$，$\therefore S_{\triangle ABP}=\frac{1}{2}\cdot|AB|\cdot d_{P-AB}=\frac{1}{2}\cdot\sqrt{2}\cdot\frac{8}{3}\cdot\frac{|\frac{3}{5}-0+1|}{\sqrt{2}}=\frac{32}{15}$.

例题 13.171 已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的短轴长为 2，点 $(1,\frac{\sqrt{2}}{2})$ 在椭圆 C 上. (1) 求椭圆 C 的标准方程；(2) 设点 $T(m,n)$ 在椭圆 C 上（点 T 不在坐标轴上），证明：直线 $\frac{mx}{2}+ny=1$ 与椭圆 C 相切；(3) 设点 P 在直线 $x=-1$ 上，且点 P 在椭圆 C 外，过点 P 作椭圆 C 的两条切线，切点分别为 A,B，O 为坐标原点，若 $\triangle PAB$ 和 $\triangle OAB$ 的面积之和为 1，求直线 AB 的方程.

解 答案：(1) $\frac{x^2}{2}+y^2=1$；(2) 证明见解析；(3) $y=\pm\frac{1}{2}(x+2)$. (1) 由题知，$\begin{cases}\frac{1^2}{a^2}+\frac{(\frac{\sqrt{2}}{2})^2}{b^2}=1\\ 2b=2\end{cases}$，解得 $a=\sqrt{2},b=1$，$\therefore$ 椭圆 C 的标准方程为 $\frac{x^2}{2}+y^2=1$.

(2) 由题意，$T(m,n)$ 在椭圆上，故 $\frac{m^2}{2}+n^2=1$，即 $m^2+2n^2=2$. 联立 $\begin{cases}\frac{x^2}{2}+y^2=1\\ \frac{mx}{2}+ny=1\end{cases}$，消去 y（或 x）并利用 $m^2+2n^2=2$，可得判别式 $\Delta=0$，故直线与椭圆相切.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

(3) 设 $P(-1,t)$, 过 P 的切线为 $\frac{x_0x}{2}+y_0y=1$, 代入 P 得 $-\frac{x_0}{2}+ty_0=1$, 即切点弦 AB 方程为 $-\frac{x}{2}+ty=1$. 设 $AB:x=2ty-2$, 与椭圆联立, 利用面积之和为 1 解得 $t=\pm1$, 故 AB 方程为 $y=\frac{1}{2}(x+2)$ 或 $y=-\frac{1}{2}(x+2)$.

例题 13.172　(1) 求椭圆 C 的标准方程; (2) 若直线 l 与椭圆交于 $A(x_1,y_1)$、$B(x_2,y_2)$ 两点, 且直线 PA、PB 的倾斜角互补, 点 $M(0,8)$, 求三角形 MAB 面积的最大值.

解 答案:(1) $\frac{x^2}{12}+\frac{y^2}{16}=1$;(2) 18. (1) $\because e=\frac{c}{a}=\frac{1}{2}$、$\therefore a=2c$、$b=\sqrt{3}c$. 设椭圆的标准方程为 $\frac{x^2}{3c^2}+\frac{y^2}{4c^2}=1$,即 $4x^2+3y^2=12c^2$,$\therefore$ 过点 $P(3,2)$、$\therefore 36+12=12c^2$、$\therefore c^2=4$.$\therefore$ 椭圆的标准方程为 $\frac{x^2}{12}+\frac{y^2}{16}=1$;(2) 由题意可知直线 l 的斜率存在, 且不过点 $P(3,2)$, 设直线 l 的方程为 $y=kx+m\,(3k+m-2\neq0)$、$A(x_1,y_1)$、$B(x_2,y_2)$,由 $\begin{cases}y=kx+m\\4x^2+3y^2=48\end{cases}$ 消去 y 整理得 $(3k^2+4)x^2+6kmx+3m^2-48=0$、$x_1+x_2=-\frac{6km}{3k^2+4}$、$x_1x_2=\frac{3m^2-48}{3k^2+4}$、$\Delta=48(12k^2+16-m^2)>0$、$\therefore k_{PA}+k_{PB}=\frac{y_1-2}{x_1-3}+\frac{y_2-2}{x_2-3}=0$、$\therefore\frac{kx_1+m-2}{x_1-3}+\frac{kx_2+m-2}{x_2-3}=2k+(3k+m-2)\left(\frac{1}{x_1-3}+\frac{1}{x_2-3}\right)=0$、$\therefore 2k+(3k+m-2)\cdot\frac{x_1+x_2-6}{x_1x_2-3(x_1+x_2)+9}=0$,将 x_1+x_2, x_1x_2 代入整理得 $(k-2)(3k+m-2)=0$,$\therefore k=2$,又因为 $\Delta=48(64-m^2)>0$,解得: $-8<m<8$,三角形 MAB 的面积 $S=\frac{1}{2}|8-m|\cdot|x_1-x_2|=\frac{\sqrt{3}}{8}\cdot\sqrt{(64-m^2)(8-m)^2}$; 令 $f(m)=(64-m^2)(8-m)^2$, $-8<m<8$, 导函数 $f'(m)=-4(8-m)^2(4+m)$, 当 $x\in(-8,-4)$, $f'(m)>0$, 当 $x\in(-4,8)$, $f'(m)<0$, 增区间为 $(-8,-4)$, 减区间为 $(-4,8)$, $\therefore$ 当 $m=-4$ 时, 三角形 MAB 的面积取得最大值, 最大值为 18.

例题 13.173　分别为点 D,E, 直线 AE 与直线 BD 相交于 P 点, (1) 求证: 点 P 在定直线上;(2) 求 $\triangle PAB$ 面积的最大值.

解 答案: (1) 点 P 在直线 $x=3$ 上; (2)$\frac{\sqrt{3}}{4}$. (1) 设椭圆 C 的方程为 $mx^2+ny^2=1(m>0,n>0,m\neq n)$,代入已知点的坐标,得: $\begin{cases}3m+n=1\\4m+\frac{2}{3}n=1\end{cases}$,解得 $\begin{cases}m=\frac{1}{6}\\n=\frac{1}{2}\end{cases}$, 所以椭圆 C 的标准方程为 $\frac{x^2}{6}+\frac{y^2}{2}=1$. (2) 如图. (1) 设直线 l 的方程为 $x=my+2(m\neq0)$, 并记点 $A(x_1,y_1)$、$B(x_2,y_2)$、$P(x_0,y_0)$, 由 $\begin{cases}x=my+2\\\frac{x^2}{6}+\frac{y^2}{2}=1\end{cases}$ 消去 x, 得 $(m^2+3)y^2+4my-2=0$, 易知 $\Delta=16m^2+8(m^2+3)=24(m^2+1)>0$, 则 $y_1+y_2=-\frac{4m}{m^2+3}$, $y_1y_2=-\frac{2}{m^2+3}$. 由条件, $D(x_1,0)$, $E(x_2,0)$, 直线 AE 的方程为 $y=\frac{y_1}{x_1-x_2}(x-x_2)$, 直线 BD 的方程为 $y=\frac{y_2}{x_2-x_1}(x-x_1)$, 联立解得 $x_0=\frac{x_2y_1+x_1y_2}{y_1+y_2}=\frac{my_1y_2+2y_1+my_1y_2+2y_2}{y_1+y_2}=\frac{2my_1y_2}{y_1+y_2}+2=3$, 所以点 P 在定直线

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$x=3$ 上. (2) $S_{\triangle PAB}=\frac{1}{2}|AD|\cdot|x_0-x_2|=\frac{1}{2}|y_1|\cdot|3-x_2|=\frac{1}{2}|y_1|\cdot|1-my_2|=\frac{1}{2}|y_1-my_1y_2|$ 而 $\frac{my_1y_2}{y_1+y_2}=\frac{1}{2}$,所以 $my_1y_2=\frac{1}{2}(y_1+y_2)$,则 $S_{\triangle PAB}=\frac{1}{2}\left|y_1-\frac{y_1+y_2}{2}\right|=\frac{1}{4}|y_1-y_2|=\frac{1}{4}\sqrt{(y_1+y_2)^2-4y_1y_2}=\frac{\sqrt{6(m^2+1)}}{2m^2+6}$. 令 $t=\sqrt{m^2+1}$, 则 $t>1$, 所以 $S_{\triangle PAB}=\frac{\sqrt{6}}{2}\cdot\frac{t}{t^2+2}=\frac{\sqrt{6}}{2}\cdot\frac{1}{t+\frac{2}{t}}\leq\frac{\sqrt{6}}{2}\times\frac{1}{2\sqrt{2}}=\frac{\sqrt{3}}{4}$,当且仅当 $t=\sqrt{2}$ 时,等号成立,所以 $\triangle PAB$ 面积的最大值为 $\frac{\sqrt{3}}{4}$.

例题 13.174 已知抛物线 $y^2=4x$ 与椭圆 $\frac{x^2}{4}+\frac{y^2}{2}=1$,直线 m 与抛物线相切,与椭圆交于 A,B 两点,与 x 轴交于点 D,O 为坐标原点,且 O,A,B 三点不共线,求 $\triangle OAB$ 面积的最大值.

解 答案: $\sqrt{2}$. 当直线 m 斜率不存在时, $m:x=0$ 与抛物线 $y^2=4x$ 相切于 O, 与椭圆交于 A,B 在 y 轴上, 三点共线, 舍去. 设直线 $m:x=ky+b$, 联立抛物线得 $y^2-4ky-4b=0$, 由 $\Delta=16k^2+16b=0$ 得 $b=-k^2<0$. 联立圆得 $(k^2+2)y^2+2kby+b^2-4=0$, 由 $\Delta=4k^2b^2-4(k^2+2)(b^2-4)>0$ 结合 $b=-k^2$ 得 $0<k^2<1+\sqrt{5}$. 设 $A(x_1,y_1),B(x_2,y_2)$, 则 $y_1+y_2=-\frac{2kb}{k^2+2}$, $y_1y_2=\frac{b^2-4}{k^2+2}$. 点 $D(b,0)$,则 $|S_{\triangle AOB}|=\frac{1}{2}|OD||y_1-y_2|=\frac{b\sqrt{4k^2-2b^2+8}}{k^2+2}$. 代入 $b=-k^2$ 得 $S=\frac{k^2\sqrt{-2k^4+4k^2+8}}{k^2+2}$. 令 $t=k^2+2$,则 $2<t<3+\sqrt{5}$, $S=f(t)=\frac{(t-2)\sqrt{-2t^2+12t-8}}{t}$. 求导得 $f'(t)=-\frac{2(t+2)(t-4)(t-1)}{t^2\sqrt{-2t^2+12t-8}}$, 故 $f(t)$ 在 $(2,4)$ 递增, 在 $(4,3+\sqrt{5})$ 递减, $t=4$ 处取最大值. 故 $S_{\max}=\frac{(4-2)\sqrt{-32+48-8}}{4}=\sqrt{2}$.

例题 13.175 (2024·四川达州·二模) 已知抛物线 $\Gamma:y^2=2px(p>0)$, 直线 $l:y=k(x-p)$ 与 Γ 交于 A,B 两点,线段 AB 中点 $M(x_m,y_m)$, $ky_m=2$.
(1) 求抛物线 Γ 的方程;
(2) 直线 l 与 x 轴交于点 C,O 为原点,设 $\triangle BOC,\triangle COM,\triangle MOA$ 的面积分别为 $S_{\triangle BOC},S_{\triangle COM},S_{\triangle MOA}$,若 $S_{\triangle BOC},S_{\triangle COM},S_{\triangle MOA}$ 成等差数列,求 k.

解 答案: $y^2=4x$, $k=\pm\frac{\sqrt{10}}{4}$. 【解法】(1) 设 $A(x_1,y_1),B(x_2,y_2)$,

$$\begin{cases}y^2=2px,\\ y=k(x-p)\end{cases}\Rightarrow ky^2-2py-2p^2k=0,$$

$$\therefore y_1+y_2=\frac{2p}{k},\quad y_m=\frac{p}{k},$$

$$\therefore ky_m=2\Rightarrow p=2,\quad \therefore\Gamma:y^2=4x.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2)

$$\begin{cases} y^2 = 4x \\ y = k(x-2) \end{cases} \Rightarrow ky^2 - 4y - 8k = 0,$$

故 $y_1 + y_2 = \frac{4}{k}, y_1 y_2 = -8$,

$$\therefore S_{\triangle BOC}, S_{\triangle COM}, S_{\triangle MOA} \text{ 成等差数列} \Rightarrow |y_2|, |y_m|, |y_1| - |y_m| \text{ 成等差数列.}$$

$$\therefore 2|y_m| = |y_2| + |y_1| - |y_m| \Rightarrow 3|y_m| = |y_1| + |y_2| \Rightarrow 3|y_m| = |y_1 - y_2|,$$

$$\therefore 9y_m^2 = (y_1 + y_2)^2 - 4y_1 y_2 \Rightarrow 9\left(\frac{2}{k}\right)^2 = \left(\frac{4}{k}\right)^2 + 32,$$

$$\therefore k = \pm\frac{\sqrt{10}}{4}.$$

例题 13.176　求 $\triangle OMN$ 面积 S 的取值范围.

解　答案: $\left(0, \frac{24\sqrt{3}}{49}\right]$. (1) 由题可得点 F 在 x 轴的正半轴, 在直线 $2x+3y-2=0$ 上, 令 $y=0$, 解得 $x=1$, 即焦点 F 坐标为 $(1,0)$, 所以 $\frac{p}{2}=1$, 解得 $p=2$, 所以抛物线的方程为: $y^2=4x$.(2) 设过点 $(3,0)$ 的直线 l 的方程为: $x=my+3$、$P(x_1,y_1)$、$Q(x_2,y_2)$, 联立 $\begin{cases} y^2=4x \\ x=my+3 \end{cases}$ 可得 $y^2-4my-12=0$, $\begin{cases} \Delta = 16m^2+48>0 \\ y_1+y_2=4m \\ y_1y_2=-12 \end{cases}$, 以焦点 F 为圆心, OF 为半径的圆 P 的方程为: $(x-1)^2+y^2=1$. 直线 OP 的方程为: $y=\frac{y_1}{x_1}x=\frac{4}{y_1}x$, 联立 $\begin{cases} y=\frac{4}{y_1}x \\ (x-1)^2+y^2=1 \end{cases}$, 解得 $M\left(\frac{2y_1^2}{y_1^2+16}, \frac{8y_1}{y_1^2+16}\right)$, 同理可得 $N\left(\frac{2y_2^2}{y_2^2+16}, \frac{8y_2}{y_2^2+16}\right)$, 设直线 MN 与 x 轴的交点 $T(n,0)$, 所以 $\overrightarrow{MT} \parallel \overrightarrow{TN}$, 由于 $\overrightarrow{MT}=\left(n-\frac{2y_1^2}{y_1^2+16}, -\frac{8y_1}{y_1^2+16}\right), \overrightarrow{TN}=\left(\frac{2y_2^2}{y_2^2+16}-n, \frac{8y_2}{y_2^2+16}\right)$,所以 $\left(n-\frac{2y_1^2}{y_1^2+16}\right)\left(\frac{8y_2}{y_2^2+16}\right)-\left(\frac{2y_2^2}{y_2^2+16}-n\right)\left(-\frac{8y_1}{y_1^2+16}\right)=0$, 化简可得: $(28n-24)(y_1-y_2)=0$, 由于 $y_1-y_2 eq 0$, 所以 $28n-24=0$, 解得 $n=\frac{6}{7}$, 则直线 MN 恒过点 $T\left(\frac{6}{7},0\right)$, 所以 $S_{\triangle OMN}=\frac{1}{2}\times\frac{6}{7}\times\left|\frac{8y_1}{y_1^2+16}-\frac{8y_2}{y_2^2+16}\right|=\frac{24}{7}\times\left|\frac{y_1y_2^2+16y_1-y_2y_1^2-16y_2}{(y_1^2+16)(y_2^2+16)}\right|$, 将 $\begin{cases} y_1+y_2=4m \\ y_1y_2=-12 \end{cases}$, 代入化简可得: $S_{\triangle OMN}=\frac{24}{7}\times\left|\frac{28(y_1-y_2)}{(y_1^2+16)(y_2^2+16)}\right|=\frac{96\sqrt{16m^2+48}}{784+256m^2}=\frac{6\sqrt{16m^2+48}}{49+16m^2}$, 令 $t=\sqrt{16m^2+48}$, 则 $t=\sqrt{16m^2+48}\geq 4\sqrt{3}$, 因为 $y=t+\frac{1}{t}(t\geq 4\sqrt{3})$ 在 $[4\sqrt{3},+\infty)$ 上单调递增, 所

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

以 $y=t+\frac{1}{t}\geq 4\sqrt{3}+\frac{1}{4\sqrt{3}}=\frac{49\sqrt{3}}{12}$, 则 $0<\frac{1}{y}\leq\frac{12}{49\sqrt{3}}=\frac{4\sqrt{3}}{49}$, 所以 $0<6\times\frac{1}{y}\leq\frac{24\sqrt{3}}{49}$, 即 $0<S_{\triangle OMN}\leq\frac{24\sqrt{3}}{49}$. 故 $\triangle OMN$ 面积 S 的取值范围为 $\left(0,\frac{24\sqrt{3}}{49}\right]$

例题 13.177 设椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左、右顶点分别为 A,B, 离心率为 e, 且 $|AB|=\frac{8\sqrt{3}}{3}e=4$. (1) 求椭圆 C 的标准方程; (2) 设点 P,Q 为椭圆上异于 A,B 的两动点, 记直线 AP 的斜率为 k_1, 直线 QB 的斜率为 k_2, 已知 $k_1=7k_2$. 直线 PQ 与 x 轴相交于点 M, 求 $\triangle APM$ 的面积的最大值.

解 答案: (1) $\frac{x^2}{4}+y^2=1$; (2) $\frac{1}{4}$. (1) 因为 $|AB|=\frac{8\sqrt{3}}{3}e=4$, 所以 $\begin{cases}\frac{8\sqrt{3}}{3}\cdot\frac{c}{a}=4,\\2a=4,\\a^2=b^2+c^2,\end{cases}$ 解得 $\begin{cases}a^2=4,\\b^2=1,\end{cases}$ 所以椭圆 C 的标准方程为 $\frac{x^2}{4}+y^2=1$; (2) 由 $\frac{x^2}{4}+y^2=1$ 可得点 $A(-2,0),B(2,0)$, 设 $P(x_1,y_1),Q(x_2,y_2)$, 直线 $AP:x=t_1y-2$, 直线 $BQ:x=t_2y+2$, 联立 $\begin{cases}x=t_1y-2,\\x^2+4y^2=4,\end{cases}$ 消去 x 得 $(t_1^2+4)y^2-4t_1y=0$, 解得 $y_1=\frac{4t_1}{t_1^2+4}$. 联立 $\begin{cases}x=t_2y+2,\\x^2+4y^2=4,\end{cases}$ 消去 x 得 $(t_2^2+4)y^2+4t_2y=0$, 解得 $y_2=-\frac{4t_2}{t_2^2+4}$. 因为 $k_{AP}=7k_{BQ}$, 且 $k_{AP}=\frac{1}{t_1},k_{BQ}=\frac{1}{t_2}$, 所以 $t_2=7t_1$, 此时 $y_2=-\frac{4t_2}{t_2^2+4}=-\frac{28t_1}{49t_1^2+4}$, 设 $M(x_0,0)$, 由 P,M,Q 三点共线, 所以 $\frac{x_1-x_0}{y_1}=\frac{x_2-x_0}{y_2}$, 则 $x_0=\frac{x_2y_1-x_1y_2}{y_1-y_2}=\frac{(t_2y_2+2)y_1-(t_1y_1-2)y_2}{y_1-y_2}=\frac{(t_2-t_1)y_1y_2+2(y_1+y_2)}{y_1-y_2}=\frac{(7t_1-t_1)\left(-\frac{112t_1^2}{(t_1^2+4)(49t_1^2+4)}\right)+2\left(\frac{4t_1}{t_1^2+4}-\frac{28t_1}{49t_1^2+4}\right)}{\frac{4t_1}{t_1^2+4}-\left(-\frac{28t_1}{49t_1^2+4}\right)}=\frac{-24\times28t_1^3+8t_1(49t_1^2+4)-56t_1(t_1^2+4)}{4t_1(49t_1^2+4)+28t_1(t_1^2+4)}=\frac{-28\times6t_1^2+98t_1^2+8-14t_1^2-56}{49t_1^2+4+7t_1^2+28}=\frac{-84t_1^2-48}{56t_1^2+32}=-\frac{12(7t_1^2+4)}{8(7t_1^2+4)}=-\frac{3}{2}$, 所以 $M\left(-\frac{3}{2},0\right)$. 所以 $S_{\triangle APM}=\frac{1}{2}\times\frac{1}{2}|y_1|=\frac{1}{4}\cdot\frac{4|t_1|}{t_1^2+4}=\frac{1}{|t_1|+\frac{4}{|t_1|}}\leq\frac{1}{4}$, 当且仅当 $|t_1|=\frac{4}{|t_1|}$, 即 $|t_1|=2$ 时等号成立. 所以 $S_{\triangle APM}$ 的最大值为 $\frac{1}{4}$.

例题 13.178 (1) 求 C 的方程;

解 答案: $x^2=4y$. 由抛物线 $C:x^2=2py$ 上一点 $H(-4,y_0)$ 到坐标原点 O 的距离为 $4\sqrt{2}$, 可得 $|OH|=\sqrt{(-4)^2+y_0^2}=4\sqrt{2}$, 解得 $y_0=4,p=2$, 所以抛物线的标准方程为 $x^2=4y$.

例题 13.179 (2) 若 $|PN|=4|PM|$, 求 k 的值;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 答案: $\frac{3\sqrt{2}}{4}$. 由题意, 设直线 l 的方程为 $y=kx+2$, 且 $A(x_1,y_1),B(x_2,y_2)$. 联立方程组 $\begin{cases}y=kx+2\\x^2=4y\end{cases}$ 消去 y 整理得 $x^2-4kx-8=0$, 则 $\Delta=(-4k)^2+4\times8>0$, 所以 $x_1+x_2=4k$, $x_1x_2=-8$, 因为 $MA\perp l, NB\perp l$, 所以 $MA\parallel NB$, 所以 $\triangle APM\sim\triangle BPN$, 又因为 $|PN|=4|PM|$, 所以 $|PB|=4|PA|$, 则 $x_2=-4x_1$, 因为 $k>0$, 所以 $\begin{cases}x_1=-\sqrt{2}\\x_2=4\sqrt{2}\end{cases}$, 则 $k=\frac{x_1+x_2}{4}=\frac{3\sqrt{2}}{4}$.

例题 13.180 (3) 若 $k\in[1,2)$, 记 $\triangle PAM$, $\triangle PBN$ 的面积分别为 S_1, S_2, 求 S_1+S_2 的取值范围.

解 答案: $[32,100)$. 根据抛物线对称性, 不妨令 $y_1<2<y_2$, 由 (2) 中 $A(x_1,y_1),B(x_2,y_2)$, 得直线 NB 的方程为 $y-y_2=-\frac{1}{k}(x-x_2)$, 令 $x=0$, 得 $N\left(0,\frac{x_2}{k}+y_2\right)$, 同理可得 $M\left(0,\frac{x_1}{k}+y_1\right)$, 则 $S_1=\frac{1}{2}\left(2-\frac{x_1}{k}-y_1\right)(-x_1)$, $S_2=\frac{1}{2}\left(\frac{x_2}{k}+y_2-2\right)x_2$, 且 $x_1+x_2=4k$, $x_1x_2=-8$, 故 $S_1+S_2=(x_1+x_2)+\frac{1}{2k}(x_1^2+x_2^2)+\frac{1}{2}(x_1y_1+x_2y_2)=-(x_1+x_2)+\frac{1}{2k}\left[(x_1+x_2)^2-2x_1x_2\right]+\frac{1}{8}(x_1+x_2)\left[(x_1+x_2)^2-3x_1x_2\right]=-4k+\frac{1}{2k}(16k^2+16)+\frac{k}{2}(16k^2+24)=8k^3+16k+\frac{8}{k}$, 令 $f(k)=8k^3+16k+\frac{8}{k}(1\le k\le2)$, 则 $f'(k)=24k^2+16-\frac{8}{k^2}$, 显然 $f'(k)>0$ 在 $[1,2]$ 上恒成立, 所以 $f(k)$ 在 $[1,2]$ 上单调递增, 由 $f(1)=32, f(2)=100$, 可得 S_1+S_2 的取值范围为 $[32,100]$.

例题 13.181 (2024·高三·四川成都·开学考试)已知双曲线 $C:\frac{x^2}{a^2}-y^2=1(a>0)$ 的焦距为 $2\sqrt{5}$ 且左右顶点分别为 A_1,A_2, 过点 $T(4,0)$ 的直线 l 与双曲线 C 的右支交于 M,N 两点. (1) 求双曲线的方程; (2) 记直线 A_2M,A_2N 的斜率分别为 k_1,k_2, 证明 $\frac{k_1}{k_2}$ 是定值; (3) 设 G 为直线 A_1M 和 A_1N 的交点, 记 $\triangle GMN,\triangle GA_1A_2$ 的面积分别为 S_1,S_2, 求 $\frac{S_1}{S_2}$ 的最小值.

解 答案:(1) $\frac{x^2}{4}-y^2=1$;(2) $\frac{k_1}{k_2}=-\frac{1}{3}$;(3) 3. (1) 由双曲线 $C:\frac{x^2}{a^2}-y^2=1(a>0)$ 的焦距为 $2\sqrt{5}$, 得 $a^2+1=(\sqrt{5})^2$, 解得 $a^2=4$, 所以双曲线 C 的方程为 $\frac{x^2}{4}-y^2=1$. (2) 依题意, 设直线 l 的方程为 $x=my+4, M(x_1,y_1),N(x_2,y_2)$, 由 $\begin{cases}x=my+4\\x^2-4y^2=4\end{cases}$ 消去 x 并整理得 $(m^2-4)y^2+8my+12=0$, 由直线 l 与双曲线的右支交于 M,N 两点, 得可得 $\begin{cases}\Delta=64m^2-48(m^2-4)>0\\m^2-4\neq0\\y_1y_2=\frac{12}{m^2-4}<0\end{cases}$, 解得 $-2<m<2$, 则 $y_1+y_2=-\frac{8m}{m^2-4}$,

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

$y_1y_2=\frac{12}{m^2-4}$, 即 $2my_1y_2=-3(y_1+y_2)$, 而 $A_1(-2,0),A_2(2,0)$, 所以 $\frac{k_1}{k_2}=\frac{\frac{y_1}{x_1+2}}{\frac{y_2}{x_2-2}}=\frac{y_1(x_2-2)}{y_2(x_1+2)}=\frac{y_1(my_2+2)}{y_2(my_1+6)}=\frac{my_1y_2+2y_1}{my_1y_2+6y_2}=\frac{-\frac{3}{2}(y_1+y_2)+2y_1}{-\frac{3}{2}(y_1+y_2)+6y_2}=\frac{\frac{1}{2}y_1-\frac{3}{2}y_2}{-\frac{3}{2}y_1+\frac{9}{2}y_2}=-\frac{1}{3}$ 为定值. (3) 由 (2) 知 $k_2=-3k_1$, 直线 $A_1M:y=k_1(x+2)$, 直线 $A_2N:y=-3k_1(x-2)$, 则点 G 的横坐标为 $x_G=1$, 于是 $\frac{S_1}{S_2}=\frac{\frac{1}{2}|GM||GN|\sin\angle MGN}{\frac{1}{2}|GA_1||GA_2|\sin\angle A_1GA_2}=\frac{|GM||GN|}{|GA_1||GA_2|}=\frac{(x_1-1)(x_2-1)}{3}=\frac{(my_1+3)(my_2+3)}{3}=\frac{m^2y_1y_2+3m(y_1+y_2)+9}{3}=\frac{4m^2}{4-m^2}+3\geqslant 3$, 当且仅当 $m=0$ 时取等号,

例题 13.182 (2024·河北·统考模拟预演) 已知抛物线 $C:x^2=2py(p>0)$, 过点 $P(0,2)$ 的直线 l 与 C 交于 A,B 两点, 当直线 l 与 y 轴垂直时, $OA\perp OB$, 其中 O 为坐标原点. (1) 求 C 的准线方程; (2) 若点 A 在第一象限, 直线 l 的倾斜角为锐角, 过点 A 作 C 的切线与 y 轴交于点 T, 连接 TB 交 C 于另一点为 D, 直线 AD 与 y 轴交于点 Q, 求 $\triangle APQ$ 与 $\triangle ADT$ 面积之比的最大值.

解 答案: (1) $y=-\frac{1}{2}$; (2) $3-2\sqrt{2}$. (1) 将 $y=2$ 代入 $x^2=2py(p>0)$, 则 $x=\pm2\sqrt{p}$, 由 $OA\perp OB$, 故 $\triangle AOB$ 为等腰直角三角形, 故 $2\sqrt{p}=2$, 即 $p=1$, 所以 $C:x^2=2y$, 故准线方程为 $y=-\frac{1}{2}$.

(2) 设 $A\left(a,\frac{a^2}{2}\right)$, 直线 $l:y=kx+2(k>0)$, 联立抛物线得 $x^2-2kx-4=0$, 所以 $ax_B=-4$, 则 $x_B=-\frac{4}{a}$, 故 $B\left(-\frac{4}{a},\frac{8}{a^2}\right)$.

由 $y=\frac{x^2}{2}$, 则 $y'=x$, 故 $k_{TA}=a$, 直线 $TA:y-\frac{a^2}{2}=a(x-a)$, 令 $x=0$, 则 $y_T=-\frac{a^2}{2}$, 故 $T\left(0,-\frac{a^2}{2}\right)$.

设直线 $TB:y=k_1x-\frac{a^2}{2}$, 联立物线得 $x^2-2k_1x+a^2=0$, 所以 $\frac{4}{a}x_D=a^2$, 则 $x_D=-\frac{a^3}{4}$, 故 $D\left(-\frac{a^3}{4},\frac{a^6}{32}\right)$.

于是, 直线 $AD:y-\frac{a^2}{2}=\frac{\frac{a^6}{32}-\frac{a^2}{2}}{-\frac{a^3}{4}-a}(x-a)$, 令 $x=0$, 则 $y_Q=\frac{a^4}{8}$, 故 $Q\left(0,\frac{a^4}{8}\right)$.

由直线 l 的倾斜角为锐角, 故 $a>2$, 则 $S_{\triangle APQ}=\frac{1}{2}\left(\frac{a^4}{8}-2\right)a$, $S_{\triangle ADT}=\frac{1}{2}\left(\frac{a^4}{8}+\frac{a^2}{2}\right)\left(a+\frac{a^3}{4}\right)=\frac{a^3}{64}(a^2+4)^2$, 所以 $\frac{S_{\triangle APQ}}{S_{\triangle ADT}}=\frac{4(a^2-4)}{a^2(a^2+4)}$, 令 $t=a^2-4>0$, 则 $a^2=t+4$, 则 $\frac{S_{\triangle APQ}}{S_{\triangle ADT}}=\frac{4t}{(t+4)(t+8)}=\frac{4}{t+\frac{32}{t}+12}\leq\frac{1}{2\sqrt{2}+3}=3-2\sqrt{2}$, 仅当 $t=4\sqrt{2}$, 即 $a^2=4(\sqrt{2}+1)$ 时等号成立, 所以 $\triangle APQ$ 与 $\triangle ADT$ 面积之比的最大值 $3-2\sqrt{2}$.

13.9.7 本节小结

联立读出韦达之后, 常见处理可以概括如下. 目标若是对称式, 我们先把它写成 s,p 的整式, 再代入; y 方向的对称量可经 $y=kx+m$ 转为 x 的和积. 几何条件有时能直接给出关于 s,p 的等式(中点、定长、定比等), 我们宜先写出这些关系, 再与韦达联立. 非对称的 (x_i-a) 型可用点乘双根 $(x_1-a)(x_2-a)=f(a)/A$; 点在曲线上时, 二次组合可直接用曲线方程整体替换. 必须求具体坐标时解方程, 只需对称量时只读韦达.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

直线形式的选取见 第13.8节. 设点与设线的取舍在 第13.16节 讨论; 极坐标与焦弦的更深结构见同卷 第14章.

13.9.8 本节习题

13.49. 椭圆 $\dfrac{x^2}{16}+\dfrac{y^2}{9}=1$, 直线 $y=2x-1$ 与之相交于 A,B. 求 $x_A^2+x_B^2$ 与 y_A+y_B (后者用 $y=2x-1$ 把 y 和转为 x 的韦达).

13.50. 椭圆 $\dfrac{x^2}{9}+\dfrac{y^2}{4}=1$ 外一点 $P(4,0)$, 过 P 斜率为 k 的弦 AB. 用点乘双根写出 $|PA|\cdot|PB|$ 关于 k 的式子, 并验证 $k=0$ 时与直接计算一致.

13.51. 椭圆 $C:\dfrac{x^2}{25}+\dfrac{y^2}{16}=1$, 弦 AB 的中点为 $M(2,2)$. 用和积独立挖掘(弦中点斜率关系)求 k 与弦方程; 再以设 $y=kx+m$ 过 M 后令 $x_M=2$ 的路径对照, 比较代数长度.

13.52. 椭圆 $C:\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$, 斜率 $k=\dfrac{1}{2}$ 的弦满足 $(x_A-x_B)^2=\dfrac{48}{25}$. 用 $s^2-4p=\Delta/A^2$ 独立挖掘截距 m, 并写出弦方程.

13.53. 证明: 对椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 上任意弦 AB, 有 $\dfrac{x_A^2+x_B^2}{a^2}+\dfrac{y_A^2+y_B^2}{b^2}=2$ (曲线方程整体代换).

13.54. 椭圆 $\dfrac{x^2}{25}+\dfrac{y^2}{16}=1$ 与 $x-2y+1=0$ 的交点: (i) 暴力解出坐标; (ii) 仅用韦达求中点; 对照说明何时该解点、何时只读和积.

13.10 一次式与法线式

直线的一般式方程 $Ax+By+C=0$ 的左端即二元一次式 $L(x,y)=Ax+By+C$. 点在直线上当且仅当一次式的值为零; 若点不在直线上, $L(x_0,y_0)$ 的绝对值与点到直线的距离只差一个固定因子, 符号记录点位于直线的哪一侧. 我们以代入一次式为基本动作, 平行与垂直的写法、距离与侧向的判定、角平分线的选取、三线共点的证明, 都可沿同一条主线完成.

13.10.1 过定点的平行线与垂线

回忆 第13.1节. 直线 $Ax+By+C=0$ 以 (A,B) 为法向量. 改变常数项 C 不改变法向量, 只让直线平行移动. 这给出平行线与垂线的快速写法.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 13.26 (过定点的平行线与垂线) 设直线 $l: Ax+By+C=0$,点 $P_0(x_0,y_0)$.

- 过 P_0 且与 l 平行的直线方程为 $A(x-x_0)+B(y-y_0)=0$;
- 过 P_0 且与 l 垂直的直线方程为 $B(x-x_0)-A(y-y_0)=0$.

证明 与 l 平行的直线以 (A,B) 为法向量, 方程形如 $Ax+By+C'=0$. 代入 P_0 得 $C'=-(Ax_0+By_0)$, 整理即得第一式.

与 l 垂直的直线以 l 的法向量 (A,B) 为方向向量,其法向量可取 $(B,-A)$,因为 $(A,B)\cdot(B,-A)=AB-AB=0$. 以 P_0 为定点写出点法式 $B(x-x_0)-A(y-y_0)=0$, 即第二式. □

平行保系数, 垂直换系数并改一处符号. 这一写法省去先求斜率、再用点斜式的中转, 也自动涵盖竖直与水平的情形, 无须讨论斜率是否存在.

知乎@余命数

例题 13.183 求过点 $P(2,-1)$ 且与直线 $l: 3x+4y-5=0$ (i) 平行; (ii) 垂直的直线方程.

解 (i) 平行: $3(x-2)+4(y+1)=0$, 即 $3x+4y-2=0$.

(ii) 垂直: $4(x-2)-3(y+1)=0$, 即 $4x-3y-11=0$.

13.10.2 一次式的符号与点的位置

直线把平面分成两个半平面, 一次式的符号精确记录这件几何事实.

定理 13.27 (一次式的符号判定) 设 $L(x,y)=Ax+By+C$ (A,B 不全为零), l 为直线 $L=0$. 对任意两点 P_1,P_2:

- $L(P_1)$ 与 $L(P_2)$ 同号 $\iff$ P_1,P_2 在 l 的同侧;
- $L(P_1)$ 与 $L(P_2)$ 异号 $\iff$ P_1,P_2 在 l 的异侧.

证明 考察线段 P_1P_2 上的动点 $P(t)=(1-t)P_1+tP_2$, $t\in[0,1]$. 代入一次式, 由 L 的线性性得

$$L(P(t))=(1-t)\,L(P_1)+t\,L(P_2).$$

右端是关于 t 的一次函数, 或常数函数, 端点值分别为 $L(P_1)$ 与 $L(P_2)$.

若 $L(P_1)$ 与 $L(P_2)$ 异号, 由连续性, 存在 $t_0\in(0,1)$ 使 $L(P(t_0))=0$, 即线段 P_1P_2 与 l 有公共点, 两点分居异侧. 若同号, 一次函数在 $[0,1]$ 上单调且端点同号, 故处处同号, 线段与 l 无公共点, 两点位于同侧. □

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

原点落在哪一侧, 由 $L(0,0)=C$ 的符号直接判定. 更一般地, 一次式符号给出点与多边形位置关系的快捷检验. 点 P 位于 $\triangle ABC$ 内部, 当且仅当 P 与 C 在直线 AB 同侧, 与 A 在直线 BC 同侧, 与 B 在直线 CA 同侧, 三个符号条件, 一次算完.

例题 13.184　判断点 $M(1,1)$ 是否在顶点为 $A(0,0)$、$B(4,0)$、$C(0,3)$ 的三角形内部.

解　三边所在直线的方程为 $AB:y=0$; $BC:3x+4y-12=0$; $CA:x=0$.

对 $M(1,1)$ 逐边检验: $y_M=1$ 与 C 的纵坐标 3 同号; $3\cdot 1+4\cdot 1-12=-5$ 与 A 处的值 -12 同号; $x_M=1$ 与 B 的横坐标 4 同号. 三个同侧条件齐备, 故 M 在三角形内部.

知乎@余命数

注 13.11 (与线性规划的联系)
第13.2节 中可行域的每条边界都对应一个一次式. $Ax+By+C\geq 0$ 表示哪个半平面, 是这里的符号判定. 解题时用原点, 或任一便利点试值定向, 这是最快的取域方法.

13.10.3 法线式

点到直线的距离公式 (第13.1节)

$$d=\frac{|Ax_0+By_0+C|}{\sqrt{A^2+B^2}}$$

知乎@余命数

的分母 $\sqrt{A^2+B^2}$ 恰是法向量 (A,B) 的长度. 若一开始就把法向量取成单位向量, 距离公式便不需要分母. 沿这条路走下去, 我们得到直线的另一种标准形式.

设直线 l 不经过原点. 从原点 O 向 l 作垂线, 垂足记为 N, 这条垂线称为 l 的**法线**. 记

$$p=|ON|>0,$$

并记 $\overrightarrow{ON}$ 与 x 轴正向的夹角为 θ $(0\leq\theta<2\pi)$. 于是 $\overrightarrow{n_0}=(\cos\theta,\sin\theta)$ 是 l 的单位法向量, $N=(p\cos\theta,\,p\sin\theta)$. 对 l 上任意一点 $P(x,y)$, 由 $\overrightarrow{NP}\perp\overrightarrow{n_0}$ 得

知乎@余命数

$$(x-p\cos\theta)\cos\theta+(y-p\sin\theta)\sin\theta=0,$$

利用 $\cos^2\theta+\sin^2\theta=1$ 化简, 得

$$x\cos\theta+y\sin\theta-p=0. \tag{13-32}$$

这就是直线的**法线式方程**. 它有两个特征一目了然. 一次项系数组成单位法向量 $(\cos\theta,\sin\theta)$; 常数项 $-p$ 的绝对值是原点到直线的距离.

一般式化为法线式, 只需把法向量单位化, 并调整整体符号使常数项为负.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

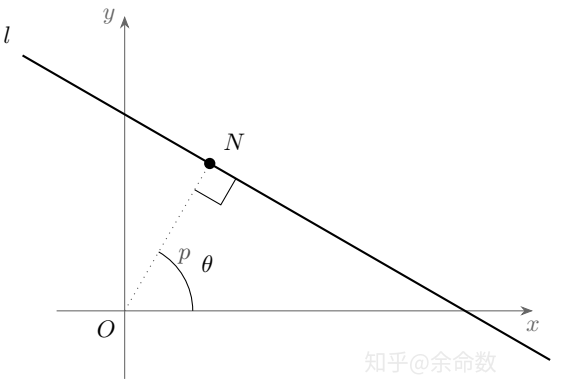


图 13.37 法线式中 p 为原点到直线的距离, θ 为法线的方向角

定理 13.28 (一般式化法线式) 设直线 $Ax+By+C=0$ 中 $C\neq 0$. 以**法线因子**

$$\lambda=-\frac{\operatorname{sgn}(C)}{\sqrt{A^2+B^2}}$$

乘方程两端, 即得法线式. 先除以 $\sqrt{A^2+B^2}$, 再按 C 的相反符号调整整体符号, 使常数项等于 $-p<0$.

证明 除以 $\sqrt{A^2+B^2}$ 后, 一次项系数组成单位向量, 方程已具有 $x\cos\theta+y\sin\theta+$ 常数 $=0$ 的形状. 再乘以 ±1 使常数项为负, 即与 (13-32) 完全一致; 此时常数项 $=-|C|\big/\sqrt{A^2+B^2}$, 故 $p=|C|\big/\sqrt{A^2+B^2}>0$. □

法线式的直接回报是**有向距离**. 把点 $P_0(x_0,y_0)$ 代入法线式左端, 记

$$\delta(P_0)=x_0\cos\theta+y_0\sin\theta-p.$$

$|\delta(P_0)|$ 即 P_0 到 l 的距离 (单位法向量投影的绝对值), 符号指示侧向. 原点处 $\delta(O)=-p<0$, 故 $\delta(P_0)<0$ 时 P_0 与原点同侧, $\delta(P_0)>0$ 时异侧. 一次式的值由此获得完整解释, 大小是距离, 符号是方向.

例题 13.185 把直线 $3x-4y+10=0$ 化为法线式, 求原点到该直线的距离, 并判断点 $Q(1,1)$ 与原点是否在该直线的同侧.

解 $\sqrt{3^2+(-4)^2}=5$, $C=10>0$, 法线因子为 $-\dfrac{1}{5}$. 法线式为

$$-\frac{3}{5}x+\frac{4}{5}y-2=0.$$

读出 $p=2$ (原点到直线的距离), $\cos\theta=-\dfrac{3}{5}$, $\sin\theta=\dfrac{4}{5}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入 $Q(1,1)$, 得 $\delta(Q)=-\dfrac{3}{5}+\dfrac{4}{5}-2=-\dfrac{9}{5}<0$, 故 Q 与原点同侧(且 Q 到直线的距离为 $\dfrac{9}{5}$).

13.10.4 角平分线

两条相交直线 l_1,l_2 的角平分线, 是到两直线距离相等的点的轨迹. 设法线式分别为 $\delta_1(x,y)=0$ 与 $\delta_2(x,y)=0$, 则角平分线的方程为

$$|\delta_1|=|\delta_2| \iff \delta_1=\delta_2 \quad 或 \quad \delta_1=-\delta_2.$$

两条角平分线互相垂直. 问题是哪一条平分哪个角. 采用法线式 (常数项均为负) 时, 规则简明.

定理 13.29 (角平分线的符号规则) 设 l_1,l_2 均不经过原点, $\delta_1=0$, $\delta_2=0$ 为它们的法线式. 则

- $\delta_1=\delta_2$ 是原点所在的那个角(及其对顶角)的平分线;
- $\delta_1=-\delta_2$ 是其余两个角的平分线.

证明 原点处的有向距离 $\delta_1(O)=-p_1<0$, $\delta_2(O)=-p_2<0$. 原点所在的角内, 任一点 P 与 O 在 l_1 同侧且在 l_2 同侧, 由 定理 13.27, $\delta_1(P)<0$ 且 $\delta_2(P)<0$. 两个负数满足 $|\delta_1|=|\delta_2|$, 等价于 $\delta_1=\delta_2$.

在相邻的角内, P 与 O 关于一条直线同侧、关于另一条异侧, $\delta_1(P)$ 与 $\delta_2(P)$ 异号, $|\delta_1|=|\delta_2|$ 等价于 $\delta_1=-\delta_2$. □

例题 13.186 求直线 $3x-4y+1=0$ 与 $5x+12y-2=0$ 的两条角平分线, 并指出哪一条平分原点所在的角.

解 $\sqrt{3^2+(-4)^2}=5$, $\sqrt{5^2+12^2}=13$. 第一条直线常数项为正, 法线因子取 $-\dfrac{1}{5}$; 第二条常数项为负, 法线因子取 $+\dfrac{1}{13}$. 两条法线式分别为

$$\delta_1=\frac{-3x+4y-1}{5}=0, \qquad \delta_2=\frac{5x+12y-2}{13}=0.$$

角平分线由 $\delta_1=\pm\delta_2$ 给出.

(i) $\delta_1=\delta_2$: $13(-3x+4y-1)=5(5x+12y-2)$, 整理得

$$64x+8y-3=0;$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(ii) $\delta_1=-\delta_2$：$13(-3x+4y-1)=-5(5x+12y-2)$，整理得

$$14x-112y+23=0.$$

知乎@余命数

两条平分线的斜率分别为 -8 与 $\dfrac{1}{8}$，乘积为 -1，确实互相垂直. 按符号规则，$64x+8y-3=0$ 平分原点所在的角.

知乎@余命数

若只求两条角平分线而不需区分，可直接写

$$\frac{|A_1x+B_1y+C_1|}{\sqrt{A_1^2+B_1^2}}=\frac{|A_2x+B_2y+C_2|}{\sqrt{A_2^2+B_2^2}},$$

再去绝对值. 需要区分时，法线式的符号规则给出确定的答案. 这一技术在求三角形内心时还会用到 (第13.12节).

知乎@余命数

13.10.5 三线共点的线性组合判据

直线束 (第13.3节) 为三线共点提供了一种代数判据，免去先求交点再验证的迂回.

定理 13.30 (三线共点的线性组合判据) 设直线 $l_1:L_1=0$ 与 $l_2:L_2=0$ 相交. 若存在不全为零的实数 m,n，使

$$L_3=mL_1+nL_2$$

为恒等式，则直线 $l_3:L_3=0$ 经过 l_1 与 l_2 的交点.

证明 设 $P_0=l_1\cap l_2$，则 $L_1(P_0)=L_2(P_0)=0$，从而

$$L_3(P_0)=mL_1(P_0)+nL_2(P_0)=0,$$

即 $P_0\in l_3$. □

这正是曲线系原理在直线情形的复述.$mL_1+nL_2=0$ 是过 $l_1\cap l_2$ 的直线束. 要证三线共点，我们只需把其中一条写成另两条的线性组合，系数由比较 x 项、y 项与常数项解出，全程不解交点. 三角形三线共点问题（中线、高、角平分线）常可沿此路处理，见 第13.12节.

例题 13.187 求证：三直线 $x+y-2=0$、$2x-y+2=0$、$4x+y-2=0$ 共点.

解 设 $4x+y-2=m(x+y-2)+n(2x-y+2)$. 比较系数得方程组

$$m+2n=4,\qquad m-n=1,\qquad -2m+2n=-2.$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

由前两式解得 $m=2, n=1$; 代入第三式, 得 $-2\cdot 2+2\cdot 1=-2$, 成立. 故第三条直线属于前两条所确定的直线束, 三线共点.

知乎@余命数

注 13.12 (行列式观点)
三条直线 $A_ix+B_iy+C_i=0\ (i=1,2,3)$ 两两不平行时共点, 等价于系数行列式

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0.$$

把 $(x,y,1)$ 看作未知量, 这是齐次线性方程组有非零解的条件; 三阶行列式为零即给出共点判据. 齐次线性方程组与行列式的系统讨论见 第22章. 三点共线的行列式判据见 第13.4.5小节, 两者互为对偶.

13.10.6 本节小结

- **平行与垂直**: 平行保系数 $A(x-x_0)+B(y-y_0)=0$; 垂直换系数 $B(x-x_0)-A(y-y_0)=0$.
- **符号判定**: $L(P_1), L(P_2)$ 同号 $\iff$ 同侧, 异号 $\iff$ 异侧; 点在三角形内部 $\iff$ 三个同侧条件.
- **法线式**: $x\cos\theta+y\sin\theta-p=0$, 一次项系数为单位法向量, p 为原点到直线的距离; 化法为除以 $\sqrt{A^2+B^2}$ 再按 C 的相反号调整. 代入点得有向距离, 大小是距离, 符号是侧向.
- **角平分线**: $\delta_1=\pm\delta_2$; 取等号的一条平分原点所在的角.
- **三线共点**: $L_3=mL_1+nL_2$ 的线性组合判据, 不解交点.

13.10.7 本节习题

13.55. 求过点 $(-1,2)$ 且与直线 $2x-3y+4=0$ (i) 平行; (ii) 垂直的直线方程.

13.56. 判断点 $(2,1)$ 与 $(-1,3)$ 在直线 $3x-y+2=0$ 的同侧还是异侧; 两点中哪一个与原点同侧?

13.57. 把直线 $5x+12y-26=0$ 化为法线式, 写出 p 与 $\cos\theta,\sin\theta$, 并求点 $(3,2)$ 到该直线的有向距离.

13.58. 求直线 $x+y-2=0$ 与 $7x-y+4=0$ 的两条角平分线, 并指出平分原点所在角的一条.

13.59. 用线性组合判据证明: 三直线 $2x+y-3=0$、$x-y+3=0$、$x+2y-6=0$ 共点.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

13.60. 求点 $(a,0)$ 到直线 $3x+4y-12=0$ 的有向距离, 并确定 a 的范围, 使该点与原点位于直线同侧.

13.11 直线偶与齐次化

我们来看一次方程与二次方程在几何上的区别. 一次方程表示一条直线, 而能分解为两个一次因式乘积的二次方程则表示两条直线. 特别地, 二元二次**齐次**方程 $Ax^2+Bxy+Cy^2=0$ 总表示过原点的两条直线, 或重合、或退缩为一点, 其斜率关系可以直接从系数读出. 把这一观察与设线法结合, 我们便得到处理过原点弦问题的齐次化技术. 它不求交点坐标, 一步就能读出 OP, OQ 两斜率的和与积.

13.11.1 表示两直线的二次方程

若一个二次方程的左端可以分解为两个一次因式, 则

$$(A_1x+B_1y+C_1)(A_2x+B_2y+C_2)=0$$

的解集是两条直线 $l_1: A_1x+B_1y+C_1=0$ 与 $l_2: A_2x+B_2y+C_2=0$ 的并集. 这样的图形称为**直线偶**, 它是一种退化的二次曲线 (第13.3节).

反过来, 给定一个一般的二元二次方程, 判断它是否表示直线偶并把两条直线找出来, 有一个实用的办法, 按其中一个字母整理成二次三项式. 若判别式恰为另一字母的完全平方, 则求根即得两个一次式.

例题 13.188 判断方程 $x^2-xy-2y^2+x+4y-2=0$ 是否表示两条直线; 若是, 求出这两条直线.

解 先把方程按 x 整理, 得 $x^2+(1-y)x+(-2y^2+4y-2)=0$. 判别式

$$\Delta=(1-y)^2-4(-2y^2+4y-2)=9y^2-18y+9=9(y-1)^2,$$

是完全平方. 求根:

$$x=\frac{(y-1)\pm 3(y-1)}{2} \implies x=2y-2 \text{ 或 } x=-y+1.$$

故原方程即 $(x-2y+2)(x+y-1)=0$, 表示直线 $x-2y+2=0$ 与 $x+y-1=0$.

注 13.13 (一般判据)
$Ax^2+Bxy+Cy^2+Dx+Ey+F=0$ 表示直线偶(允许重合)的充分必要条件, 在线性代数中可用三阶行列式刻画, 见 第22章. 下文处理常见情形, 仍用整理成二次三项式、验证判别式为完全平方的直接分解办法.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.11.2 过原点的两直线

直线偶中结构最整齐的情形，是两条直线都经过原点. 此时方程不含一次项与常数项，成为二元二次**齐次方程**

$$Ax^2 + Bxy + Cy^2 = 0. \tag{13-33}$$

齐次性的含义是，若 (x_0, y_0) 满足方程，则对任意 t，(tx_0, ty_0) 也满足，解集由过原点的直线组成. 设 $C \neq 0$，在 $x \neq 0$ 的点上令 $k = \dfrac{y}{x}$，方程两边除以 x^2 得

$$Ck^2 + Bk + A = 0.$$

两条直线的斜率 k_1, k_2 是这个二次方程的两根，由韦达定理：

$$k_1 + k_2 = -\frac{B}{C}, \qquad k_1 k_2 = \frac{A}{C}. \tag{13-34}$$

判别式 $B^2 - 4AC$ 决定图形. 大于零为两条相交直线，等于零为两条重合直线，小于零则只剩原点一个实点. 由 (13-34) 可读出两条常用推论.

- **垂直条件**. $k_1 k_2 = -1 \iff A + C = 0$；
- **夹角公式**. 设两直线夹角为 φ，则

$$\tan\varphi = \left|\frac{k_1 - k_2}{1 + k_1 k_2}\right| = \frac{\sqrt{(k_1 + k_2)^2 - 4k_1 k_2}}{|1 + k_1 k_2|} = \frac{\sqrt{B^2 - 4AC}}{|A + C|}.$$

例题 13.189　求两直线 $2x^2 - 5xy + 2y^2 = 0$ 的夹角，并判断它们是否垂直.

解　$A = 2、B = -5、C = 2$. $A + C = 4 \neq 0$，故两直线不垂直. $B^2 - 4AC = 25 - 16 = 9$，夹角满足

$$\tan\varphi = \frac{\sqrt{9}}{|2 + 2|} = \frac{3}{4}.$$

（分解验证. $2x^2 - 5xy + 2y^2 = (2x - y)(x - 2y)$，斜率 2 与 $\dfrac{1}{2}$，$\tan\varphi = \left|\frac{2 - \frac{1}{2}}{1 + 2\cdot\frac{1}{2}}\right| = \dfrac{3}{4}$，一致.）

13.11.3 共渐近线的双曲线系

双曲线的渐近线 (第13.3.3小节) 正是直线偶的用武之地. 双曲线

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

的两条渐近线 $y=\pm\dfrac{b}{a}x$, 合成一个齐次方程

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=0.$$

右端的 1 换成 0, 双曲线退化为自己的渐近线偶. 更进一步, 对任意 $\lambda\neq 0$, 方程

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=\lambda \tag{13-35}$$

表示的双曲线都与原双曲线**共渐近线**(渐近方向只由二次项决定); $\lambda>0$ 时焦点在 x 轴一侧, $\lambda<0$ 时在 y 轴一侧, $\lambda=0$ 时退化为渐近线偶. 这是曲线系思想(第13.3节) 的又一成员. 一族曲线共享同一对渐近线, 只由一个参数区分.

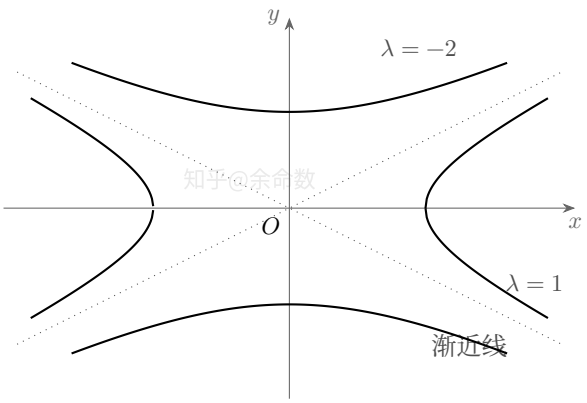


图 13.38 共渐近线的双曲线系 $\dfrac{x^2}{4}-y^2=\lambda$: $\lambda=0$ 退化为渐近线偶

例题 13.190 求与双曲线 $\dfrac{x^2}{4}-y^2=1$ 共渐近线, 且经过点 $(2,\sqrt{3})$ 的双曲线方程.

解 设所求双曲线为 $\dfrac{x^2}{4}-y^2=\lambda$. 代入 $(2,\sqrt{3})$, 得

$$\frac{4}{4}-3=\lambda \implies \lambda=-2.$$

故所求方程为 $\dfrac{x^2}{4}-y^2=-2$, 即 $\dfrac{y^2}{2}-\dfrac{x^2}{8}=1$. 其渐近线 $y=\pm\dfrac{x}{2}$ 与原双曲线相同.

13.11.4 齐次化方法

现在把前面三个结论整合起来. 设直线 l 与二次曲线 Γ 交于 P,Q 两点, 我们关心原点与两交点连线的斜率 k_{OP},k_{OQ} 之间的关系(垂直、和为定值之类). 常规

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

路径是设线联立、韦达、再把斜率关系翻译回 k 与截距 (第13.7节–第13.9节). 齐次化提供一条更短的路径.

13.11.4.0.1　操作程序

1. 把直线 l 的方程写成 $mx+ny=1$ 的形状(这要求 l 不过原点;过原点的情形单独检验, 通常意味着 P 或 Q 与 O 重合而退化);
2. 在曲线 Γ 的方程中, 把常数项与一次项用 $1=mx+ny$ 配齐, 使方程成为关于 x,y 的二次齐次方程;
3. 该齐次方程恰表示 OP,OQ 两直线(它们过原点, 且与 l 的交点即 P,Q), 用 (13-34) 直接读出 $k_{OP}+k_{OQ}$ 与 $k_{OP}\cdot k_{OQ}$.

齐次化后的方程由 Γ 的方程与 l 的方程组合而成, 故 P,Q 的坐标都满足它; 它又是齐二次的, 解集由过原点的直线组成. 这两件事实合在一起, 它表示的正是直线 OP 与 OQ.

例题 13.191 ($OP\perp OQ$ 的弦过定点) 抛物线 $y^2=2px$ $(p>0)$ 上两点 A,B 满足 $OA\perp OB$. 求证:直线 AB 过定点.

解 直线 AB 不过原点(否则 A 或 B 与 O 重合, 与 $OA\perp OB$ 矛盾), 设其方程为 $mx+ny=1$. 用 $1=mx+ny$ 配齐抛物线方程, 得

$$y^2=2px\cdot(mx+ny)\implies 2pm\,x^2+2pn\,xy-y^2=0.$$

这是 OA,OB 两直线的合成方程.$OA\perp OB\iff A+C=0$, 即

$$2pm-1=0\implies m=\frac{1}{2p}.$$

代回直线方程, 得 $\dfrac{x}{2p}+ny=1$, 对任意 n, 它经过点 $(2p,0)$. 故 AB 过定点 $(2p,0)$.

将这里的计算量与设线法路线对比. 那边要联立 $y=kx+b$ 与抛物线、写出两根和积、翻译 $\overrightarrow{OA}\cdot\overrightarrow{OB}=0$、再解出 b 与 k 的关系; 这边一次配齐、一次读出, 条件 $OA\perp OB$ 直接化为 m 的方程. 齐次化把原点与两交点这一结构利用到了极限.

例题 13.192 (垂直弦的弦心距为定值) 直线 l 与椭圆 $\dfrac{x^2}{6}+\dfrac{y^2}{3}=1$ 交于 M,N 两点, 且 $OM\perp ON$. 求证:原点 O 到直线 l 的距离为定值.

解 设 $l:mx+ny=1$(l 不过原点, 理由同上). 配齐椭圆方程, 得

$$\frac{x^2}{6}+\frac{y^2}{3}=(mx+ny)^2\implies\Big(\frac{1}{6}-m^2\Big)x^2-2mn\,xy+\Big(\frac{1}{3}-n^2\Big)y^2=0.$$

知乎@余命数　知乎@余命数　知乎@余命数

它表示 OM, ON 两直线.$OM \perp ON \iff A + C = 0$, 即

$$\frac{1}{6} - m^2 + \frac{1}{3} - n^2 = 0 \implies m^2 + n^2 = \frac{1}{2}.$$

原点到 l 的距离

$$d = \frac{1}{\sqrt{m^2 + n^2}} = \sqrt{2},$$

是与 l 的选取无关的定值.

注 13.14 (何时想到齐次化)
题面信号是原点与两交点同时在场. 条件或目标围绕 OP, OQ 的斜率关系(垂直、和/积为定值、角平分线过 y 轴之类)展开. 若定点不在原点, 可以先平移坐标系把它移到原点, 再齐次化. 若目标是弦长、中点一类与原点无关的量, 齐次化使不上力, 回到设点法或设线法.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

13.11.5 本节小结

- **直线偶**. 二次方程分解为两一次因式 $\iff$ 表示两直线; 判别程序是按一元整理, 判别式须为完全平方.
- **齐次方程**.$Ax^2 + Bxy + Cy^2 = 0$ 表示过原点两直线, $k_1 + k_2 = -B/C$, $k_1 k_2 = A/C$; 垂直 $\iff A + C = 0$, 夹角 $\tan\varphi = \sqrt{B^2 - 4AC}\,/|A + C|$.
- **共渐近线系**.$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = \lambda$, $\lambda = 0$ 即渐近线偶.
- **齐次化**. 直线写成 $mx + ny = 1$, 用 1 配齐曲线方程, 一步读出 OP, OQ 的斜率关系; 条件是直线不过原点.

知乎@余命数 知乎@余命数

13.11.6 本节习题

13.61. 判断 $x^2 - 3xy + 2y^2 = 0$ 表示什么图形; 若为过原点的两直线, 求其夹角.

13.62. 把 $x^2 - xy - 2y^2 + x + 4y - 2 = 0$ 分解为两个一次因式, 并求两直线的交点.

知乎@余命数

13.63. 求与双曲线 $\dfrac{x^2}{9} - \dfrac{y^2}{4} = 1$ 共渐近线且经过点 $(6, 2)$ 的双曲线方程.

13.64. 证明: 方程 $Ax^2 + Bxy + Cy^2 = 0$ 表示的两直线互相垂直, 当且仅当 $A + C = 0$; 并写出它们重合时系数应满足的条件.

知乎@余命数

13.65. 抛物线 $y^2 = 2x$ 上两点 A, B 满足 $OA \perp OB$. 用齐次化证明 AB 过定点, 并写出该定点坐标.

13.66. 直线 $y = -x + b$ 与抛物线 $y^2 = 4x$ 交于 A, B 两点, 且 $OA \perp OB$. 求 b 的值.

13.12 三角形的心

三角形的重心、外心、垂心、内心, 在综合几何中各有定义、各成体系. 坐标法把它们统一到一条主线上. 我们将会看到, 四心的坐标都是顶点坐标的简单代数组合, 四心之间的关系, 例如欧拉线也可以由直接计算证明.

13.12.1 平行四边形的顶点

平行四边形最有用的代数特征, 是对角线互相平分. 用位置向量写, 平行四边形 $ABCD$ 的两条对角线 AC 与 BD 有共同的中点, 因此

$$\frac{A+C}{2}=\frac{B+D}{2}\quad\Longleftrightarrow\quad A+C=B+D.$$

于是已知三个顶点求第四个的问题, 化为一次移项.

例题 13.193 已知 $A(1,2)$、$B(4,1)$、$C(2,5)$. 求使四点成为平行四边形顶点的点 D.

解 题目没有指定顶点的顺序, 我们按 D 与谁相对分三种情形.

- D 与 B 相对, 则 $D=A+C-B=(1+2-4,\,2+5-1)=(-1,6)$;
- D 与 C 相对, 则 $D=A+B-C=(1+4-2,\,2+1-5)=(3,-2)$;
- D 与 A 相对, 则 $D=B+C-A=(4+2-1,\,1+5-2)=(5,4)$.

三个解对应三种不同的平行四边形, 都满足条件.

分类讨论在这里是硬要求. 只写一种情形, 答案就不完整. 同一个中点重合的观点, 在 第11.2节 的定比分点问题中还会反复出现.

13.12.2 重心

三条中线交于一点, 这一点是重心. 用坐标证明, 只需写出一条中线上 $\dfrac{2}{3}$ 处的点, 再看它的对称性.

定理 13.31 (重心坐标公式) 设 $A(x_1,y_1)$、$B(x_2,y_2)$、$C(x_3,y_3)$, 则 $\triangle ABC$ 的重心为

$$G=\Big(\frac{x_1+x_2+x_3}{3},\,\frac{y_1+y_2+y_3}{3}\Big),$$

用位置向量写即 $G=\dfrac{A+B+C}{3}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 边 BC 的中点为 $M=\dfrac{B+C}{2}$. 重心在中线 AM 上, 且 $AG:GM=2:1$, 由定比分点公式 (第11.2节), 得

$$G=\frac{A+2M}{3}=\frac{A+B+C}{3}.$$

所得坐标关于 A,B,C 对称, 故它也在另外两条中线上, 即三条中线共点于 G. □

对称性在这里承担了全部证明工作. 我们只求一条中线上的分点, 得到的表达式关于三顶点对称, 三条中线自然汇于同一点. 重心的进一步性质, 例如奔驰定理中重心对应的面积比 $1:1:1$ 见 第11.2节.

知乎@余命数

13.12.3 外心

三边中垂线交于一点, 即外接圆的圆心, 称为**外心**. 坐标求法有一个简化. 设外心 $O=(x,y)$, 由 $|OA|^2=|OB|^2$, 展开后二次项 x^2+y^2 两边相消, 剩下的是 x,y 的一**次**方程, 这正是 AB 中垂线的方程. 两个这样的方程联立, 即得外心.

例题 13.194 求 $A(0,0)$、$B(4,0)$、$C(1,3)$ 的外接圆圆心与半径.

解 设 $O=(x,y)$. 由 $|OA|^2=|OB|^2$, 得

$$x^2+y^2=(x-4)^2+y^2 \implies 8x=16 \implies x=2.$$

由 $|OA|^2=|OC|^2$, 得

$$x^2+y^2=(x-1)^2+(y-3)^2 \implies 2x+6y=10.$$

代入 $x=2$ 得 $y=1$, 故 $O=(2,1)$, 半径 $R=\sqrt{|OA|^2}=\sqrt{2^2+1^2}=\sqrt{5}$.

方程 $2x+6y=10$ 即 AC 的中垂线 (过中点 $(\frac{1}{2},\frac{3}{2})$ 且法向量为 $\overrightarrow{AC}=(1,3)$), 与几何定义完全吻合.

13.12.4 垂心

三条高交于一点, 称为**垂心**. 求垂心, 写两条高的方程联立即可, 第三条高自动通过交点, 这个事实本身需要证明, 借助外心可直接完成.

定理 13.32 (垂心的向量表示) 设 $\triangle ABC$ 的外心为 O. 以 O 为原点取位置向量, 则垂心

$$H=A+B+C.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 以 O 为原点，有 $|A|=|B|=|C|=R$. 令 $H'=A+B+C$，我们来验证 $AH'\perp BC$.

$$\overrightarrow{AH'}\cdot\overrightarrow{BC}=(H'-A)\cdot(C-B)=(B+C)\cdot(C-B)=|C|^2-|B|^2=0.$$

同理 $BH'\perp CA, CH'\perp AB$. 故 H' 在三条高上，即 $H'=H$. □

数量积 $(B+C)\cdot(C-B)=|C|^2-|B|^2$ 一步归零，外心到三顶点等距的条件用完了全部所需. 这个关系式还给出一条重要推论，见下一小节.

例题 13.195 求 $A(0,0)$、$B(4,0)$、$C(1,3)$ 的垂心.

解 边 BC 的斜率为 $\dfrac{3-0}{1-4}=-1$，故从 A 引出的高斜率为 1，方程为 $y=x$.

边 AC 的斜率为 3，故从 B 引出的高斜率为 $-\dfrac{1}{3}$，方程为 $y=-\dfrac{1}{3}(x-4)$，即 $x+3y=4$.

联立 $y=x$ 与 $x+3y=4$，得 $H=(1,1)$. 我们验证，从 C 引出的高垂直于水平的 AB，方程为 $x=1$，确实过 H.

13.12.5 内心

三条角平分线交于一点，即内切圆的圆心，称为**内心**. 内心的坐标公式以三边长为权.

定理 13.33 (内心坐标公式) 设 $\triangle ABC$ 中 $a=|BC|$、$b=|CA|$、$c=|AB|$，则内心

$$I=\frac{aA+bB+cC}{a+b+c}.$$

证明 由角平分线定理，$\angle A$ 的平分线交 BC 于 D，且 $BD:DC=c:b$，故

$$D=\frac{bB+cC}{b+c}.$$

又 $AI:ID=(b+c):a$（在 $\triangle ABD$ 中对 BI 再用一次角平分线定理），于是

$$I=\frac{aA+(b+c)D}{a+b+c}=\frac{aA+bB+cC}{a+b+c}.$$

表达式对三顶点对称，各顶点乘以对边长，故三条角平分线共点于 I. □

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

权重 a,b,c 恰为对边长度，与奔驰定理中内心对应的面积比 $a:b:c$ 一致(第11.2节). 若只求角平分线的方程而不求交点, 第13.10.4小节 的符号规则可以直接取用.

知乎@余命数

例题 13.196 求 $A(0,0)$、$B(4,0)$、$C(0,3)$ 的内心与内切圆半径.

知乎@余命数

解 三边长 $a=|BC|=5$、$b=|CA|=3$、$c=|AB|=4$, 故

$$I=\frac{5(0,0)+3(4,0)+4(0,3)}{5+3+4}=\left(\frac{12}{12},\frac{12}{12}\right)=(1,1).$$

知乎@余命数

内切圆半径 r 等于 I 到 x 轴(边 AB)的距离, 即 $r=1$. 另作验证, 面积 $\frac{3\cdot 4}{2}=6$, 半周长 6, $r=\frac{6}{6}=1$, 一致.

知乎@余命数

13.12.6 欧拉线

外心、重心、垂心, 三个由不同定义给出的点, 竟永远共线. 这是初等几何中最著名的定理之一, 用坐标法证明几乎只有一行.

定理 13.34 (欧拉线) 设 $\triangle ABC$ 的外心、重心、垂心分别为 O,G,H, 则 O,G,H 共线, 且

$$\overrightarrow{GH}=2\overrightarrow{OG}.$$

这条直线称为三角形的**欧拉线**.

证明 取 O 为原点. 由垂心的向量表示, $H=A+B+C$; 由重心公式, $G=\frac{A+B+C}{3}$. 故

$$H=3G,$$

即 $\overrightarrow{OH}=3\overrightarrow{OG}$, 三点共线, 且 G 在线段 OH 上, $OG:GH=1:2$. □

注 13.15 (九点圆)
欧拉线上还有一位成员. 三边中点、三高垂足、三顶点到垂心连线的中点, 九点共圆, 圆心 N 恰为 OH 的中点, 半径为外接圆半径的一半. 以外心为原点, $N=\frac{H}{2}=\frac{A+B+C}{2}$, 验证各点到 N 的距离都等于 $\frac{R}{2}$, 同样是直接的向量计算.

13.12.7 本节小结

- **平行四边形**.$A+C=B+D$; 第四顶点三种情形, 不可遗漏.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

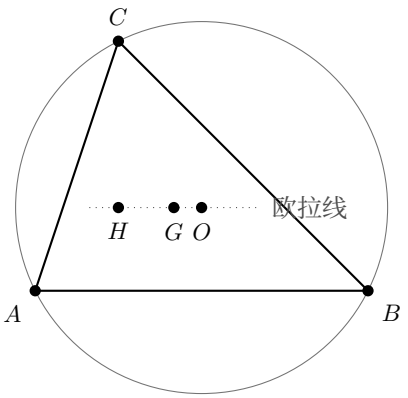


图 13.39　欧拉线: $A(0,0)$、$B(4,0)$、$C(1,3)$ 的外心 $O(2,1)$、重心 $G(\frac{5}{3},1)$、垂心 $H(1,1)$ 共线

- **重心**.$G=\dfrac{A+B+C}{3}$; 对称性保证三线共点.
- **外心**. 距离平方相等 $\to$ 一次方程(中垂线), 两式联立.
- **垂心**. 两条高联立; 以外心为原点时 $H=A+B+C$.
- **内心**.$I=\dfrac{aA+bB+cC}{a+b+c}$, 权重为对边长.
- **欧拉线**.O,G,H 共线, $\overrightarrow{OH}=3\overrightarrow{OG}$; 九点圆圆心为 OH 中点.

13.12.8 本节习题

13.67. 已知三点 $(0,0)$、$(2,1)$、$(1,3)$, 求使四点构成平行四边形的第四顶点(写出全部情形).

13.68. 求 $A(1,1)$、$B(5,1)$、$C(2,4)$ 的外接圆圆心与半径.

13.69. 求上题三角形的垂心, 并验证三条高确实共点.

13.70. 求 $A(0,0)$、$B(3,0)$、$C(0,4)$ 的内心与内切圆半径.

13.71. 设三角形的外心为 $O(2,1)$, 重心为 $G(\frac{5}{3},1)$, 求垂心 H 的坐标.

13.72. 证明: 对任意三角形, $\overrightarrow{OH}=\overrightarrow{OA}+\overrightarrow{OB}+\overrightarrow{OC}$, 其中 O 为外心, H 为垂心; 并说明该式与欧拉线的关系.

13.13 共圆点与共轴圆

圆的方程 $x^2+y^2+Dx+Ey+F=0$ 含三个待定系数, 故三个条件确定一个圆. 但处理圆的问题时, 求出方程往往不是最经济的路径. 四点共圆的判定、两圆的公共弦与根轴、三圆的根心与共轴圆系, 这些问题共同的代数基础只有一条. **两个圆方程相减, 二次项消去, 剩下一个一次式**. 本节的讨论都围绕这个一次式展开.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.13.1 直径式方程与四点共圆

知乎@余命数

以 $A(x_1,y_1)$, $B(x_2,y_2)$ 为直径两端点的圆, 方程可以直接写出.

$$(x-x_1)(x-x_2)+(y-y_1)(y-y_2)=0. \tag{13-36}$$

左端展开即 $\overrightarrow{PA}\cdot\overrightarrow{PB}$ ($P=(x,y)$), 由 Thales 定理 (第13.1节), 点 P 在该圆上当且仅当 $\overrightarrow{PA}\cdot\overrightarrow{PB}=0$. 直径式把直径所对圆周角为直角这一定理翻译成了方程.

一般地, 判定四点共圆, 我们采用标准代数程序. 圆的一般方程关于系数 D,E,F 是**线性**的, 代入三个点的坐标得三元一次方程组, 解出圆方程, 再把第四点代入检验.

例题 13.197 求证: 四点 $A(-1,0)$、$B(1,0)$、$C(0,2)$、$D(0,-\dfrac{1}{2})$ 共圆.

解 设过 A,B,C 的圆为 $x^2+y^2+Dx+Ey+F=0$. 代入三点, 得

$$1-D+F=0,\qquad 1+D+F=0,\qquad 4+2E+F=0.$$

由前两式得 $D=0$, $F=-1$; 代入第三式得 $E=-\dfrac{3}{2}$. 圆方程为

$$x^2+y^2-\frac{3}{2}y-1=0.$$

知乎@余命数

知乎@余命数

代入 $D(0,-\dfrac{1}{2})$, 得 $\dfrac{1}{4}+\dfrac{3}{4}-1=0$, 成立. 故四点共圆.

注 13.16 (几何知识先行)
上题若先观察, AB 与 CD 都过原点, 且 $|OA|\cdot|OB|=1=|OC|\cdot|OD|$ (取有向长度时两侧符号相反), 由相交弦定理的逆定理立得四点共圆, 一个方程都不用解. 代数程序是可靠的底线, 几何观察则经常一步到位, 两者的配合见 第13.15节.

13.13.2 公共弦与根轴

设两圆方程中二次项系数都已化为 1, 得

$$S_1=x^2+y^2+D_1x+E_1y+F_1=0,\qquad S_2=x^2+y^2+D_2x+E_2y+F_2=0.$$

知乎@余命数

两式相减, 得一次式

$$L(x,y)=(D_1-D_2)x+(E_1-E_2)y+(F_1-F_2)=0. \tag{13-37}$$

知乎@余命数

两圆相交时，两个交点的坐标同时满足 $S_1 = 0$ 与 $S_2 = 0$，故满足 $L = 0$. 直线 (13-37) 就是**公共弦所在直线**. 两圆相切时，它是过切点的公切线. 这条直线还有一个名字.

定理 13.35 (根轴) 设两圆不同心(此时 $D_1 - D_2$ 与 $E_1 - E_2$ 不同时为零). 直线 (13-37) 是对两圆**幂相等**的点的轨迹，称为两圆的**根轴**.

证明 点 $P(x_0, y_0)$ 对圆 $S_i = 0$ 的幂，等于把 P 的坐标代入左端所得的值 $S_i(P)$ (第13.1节). 于是

$$S_1(P) = S_2(P) \iff S_1(P) - S_2(P) = 0 \iff L(P) = 0,$$

即 P 在直线 (13-37) 上. □

知乎@余命数

根轴与连心线垂直.(13-37) 的法向量为 $(D_1 - D_2, E_1 - E_2)$，而两圆心为 $(-\frac{D_i}{2}, -\frac{E_i}{2})$，连心方向 $(\frac{D_2 - D_1}{2}, \frac{E_2 - E_1}{2})$ 恰与之平行. 即使两圆不相交，根轴依然存在，幂相等的定义不依赖交点.

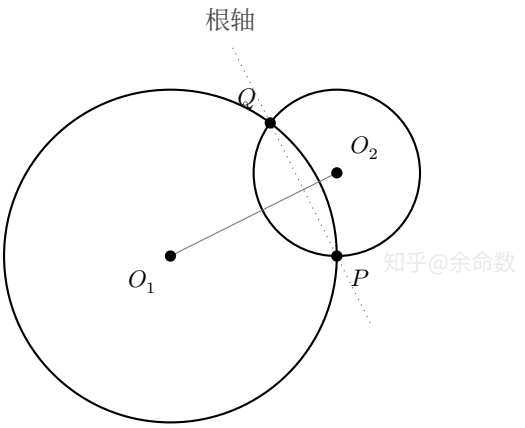


图 13.40　两圆的公共弦 PQ 所在直线即根轴，与连心线 O_1O_2 垂直

例题 13.198　求圆 $x^2 + y^2 = 4$ 与圆 $x^2 + y^2 - 4x - 2y + 4 = 0$ 的公共弦所在直线方程及弦长.

解 两式相减，得 $4x + 2y - 8 = 0$，即

$$2x + y - 4 = 0.$$

知乎@余命数

第一个圆圆心 $O_1(0,0)$, 半径 $R=2$; 圆心到该直线的距离

$$d=\frac{|-4|}{\sqrt{2^2+1^2}}=\frac{4}{\sqrt{5}}<2,$$

故两圆确实相交. 由弦长公式 (第13.1节), 得

$$\text{弦长}=2\sqrt{R^2-d^2}=2\sqrt{4-\frac{16}{5}}=\frac{4\sqrt{5}}{5}.$$

整个过程中没有求出任何一个交点坐标, 这是两式相减的价值.

13.13.3 根心

三个圆两两各有一条根轴, 共三条. 它们的位置关系, 用 第13.10.5小节 的线性组合判据一眼看穿.

定理 13.36 (根心定理) 三个圆(圆心不共线)两两之间的三条根轴交于一点, 称为三圆的**根心**. 根心对三个圆的幂相等.

证明 设三圆为 $S_1=0$、$S_2=0$、$S_3=0$. 三条根轴分别为

$$L_{12}=S_1-S_2=0,\qquad L_{23}=S_2-S_3=0,\qquad L_{13}=S_1-S_3=0.$$

恒等式 $L_{13}=L_{12}+L_{23}$ 表明 L_{13} 是 L_{12} 与 L_{23} 的线性组合, 故 L_{13} 过前两者的交点 P. P 处 $S_1(P)=S_2(P)=S_3(P)$, 即根心对三圆等幂. □

根心定理常用于证明三线共点. 若三条线可看作某三个圆的根轴, 定理直接给出结论.

例题 13.199 求三圆 $x^2+y^2=4$、$x^2+y^2-4x-2y+4=0$、$x^2+y^2+2x-6y+6=0$ 的根心.

解 三条根轴为

$$L_{12}:4x+2y-8=0,\quad L_{23}:-6x+4y-2=0,\quad L_{13}:-2x+6y-10=0.$$

联立 L_{12} 与 L_{23}, 即 $2x+y=4$ 与 $3x-2y=-1$. 由 $y=4-2x$ 代入得 $3x-8+4x=-1$、$x=1$、$y=2$. 代入 L_{13} 验证, 得 $-2+12-10=0$, 成立. 根心为 $(1,2)$.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

13.13.4 共轴圆系

过两圆交点的圆系 $S_1+\lambda S_2=0$, 第13.3节 已给出. 从根轴的观点看, 这个圆系有更清楚的结构. 对任意 λ(除去使二次项消去的 $\lambda=-1$), 圆 $S_1+\lambda S_2=0$ 与 $S_1=0$ 的根轴仍是

$$S_1-\frac{S_1+\lambda S_2}{1+\lambda}=\frac{\lambda}{1+\lambda}(S_1-S_2)=0 \text{ 化简后与 } L=0 \text{ 一致,}$$

即圆系中每个成员都共享同一条根轴. 共享同一根轴的圆的全体, 称为**共轴圆系**. 用根轴 $L=0$ 直接写, 共轴圆系取更灵活的形式

$$S+\lambda L=0,$$

其中 $S=0$ 是系中任一已确定的圆. 系中还有两位退化成员. 根轴 $L=0$ 自身, 即 $\lambda\to\infty$ 的极限, 以及半径为零的**点圆**, 即系中半径最小的位置, 称为极限点. 它们都在其中. 至于同心圆, 圆心相同而半径各异, 可看作根轴退到无穷远处的极端情形.

例题 13.200　求过两圆 $x^2+y^2=4$ 与 $x^2+y^2-4x-2y+4=0$ 的交点, 且经过点 $(0,1)$ 的圆.

解　设所求圆属于共轴圆系 $S_1+\lambda S_2=0$.

$$(x^2+y^2-4)+\lambda(x^2+y^2-4x-2y+4)=0.$$

代入 $(0,1)$, 得 $(1-4)+\lambda(1-2+4)=0$, 即 $-3+3\lambda=0$, $\lambda=1$. 所求圆为

$$2x^2+2y^2-4x-2y=0 \quad \text{即} \quad x^2+y^2-2x-y=0.$$

(验证. 两圆交点满足 $S_1=S_2=0$, 故满足组合方程; $(0,1)$ 代入 $1-1=0$, 成立.)

注 13.17 (正交圆族)
与共轴圆系中每个圆都正交(交点处切线垂直)的圆, 恰好也组成一个共轴圆系, 其根轴是原系的连心线. 一族过两定点, 一族成极限点, 两族互相正交, 这是静电场中电场线与等势线的几何原型.

13.13.5 本节小结

- **直径式**. $(x-x_1)(x-x_2)+(y-y_1)(y-y_2)=0$, 即 $\overrightarrow{PA}\cdot\overrightarrow{PB}=0$.
- **四点共圆**. 三点定圆(关于 D,E,F 线性), 第四点代入; 或先用圆幂观察.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **根轴**.$S_1 - S_2 = 0$, 幂相等点的轨迹, 垂直于连心线; 相交时即公共弦所在直线, 相切时即公切线.
- **根心**. 三圆的三条根轴共点, $L_{13} = L_{12} + L_{23}$.
- **共轴圆系**.$S + \lambda L = 0$, 共享根轴; 退化成员为根轴与点圆.

13.13.6 本节习题

13.73. 写出以 $(1,2),(3,-2)$ 为直径两端点的圆的方程, 并化为标准形式.

13.74. 判断四点 $A(1,1)$、$B(3,1)$、$C(3,3)$、$D(1,3)$ 是否共圆; 若共圆, 求该圆方程.

13.75. 求两圆 $x^2+y^2-2x-4y+1=0$ 与 $x^2+y^2+4x-1=0$ 的根轴, 并验证它与连心线垂直.

13.76. 证明: 两圆 $x^2+y^2+D_1x+E_1y+F_1=0$ 与 $x^2+y^2+D_2x+E_2y+F_2=0$ 的根轴垂直于连心线.

13.77. 验证本节例题中三圆的三条根轴满足 $L_{13}=L_{12}+L_{23}$, 并求根心对三圆的幂.

13.78. 求过两圆 $x^2+y^2=4$ 与 $x^2+y^2-4x-2y+4=0$ 的交点, 且半径最小的圆(提示: 以公共弦为直径).

13.14 轨迹方程

轨迹问题是解析几何的逆问题. 已知动点满足的几何条件, 我们求它所描绘的曲线. 前面各节已零星见过它 (第11.2节 的等和线, 第13.6.1小节 的向量条件求轨迹), 本节把相关方法整理清楚. 求轨迹的基本功是两步. **把几何条件翻译为坐标方程**, 以及**检验所得方程不多不少恰好是轨迹**. 第一步有直接法、定义法、相关点法、交轨法、参数法五种常见路径; 第二步是纯粹性与完备性的讨论, 也最容易被忽略.

13.14.1 直接法

条件本身已经以距离、斜率、数量积等可以直接翻译的形式给出时, 我们设动点 $M(x,y)$, 把条件写成方程, 化简即得. 这是轨迹问题的基准方法.

例题 13.201 动点 M 到两定点 $A(-1,0)$, $B(1,0)$ 的距离的平方和等于 10. 求 M 的轨迹方程, 并说明图形.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 设 $M(x,y)$. 条件即

$$(x+1)^2+y^2+(x-1)^2+y^2=10.$$

展开, 得 $2x^2+2y^2+2=10$, 即

$$x^2+y^2=4.$$

轨迹是以原点为圆心、半径 2 的圆. (一次项 $\pm 2x$ 相互抵消, 这正是对称建系的收益.)

13.14.2 定义法

有时条件翻译到一半, 我们会认出它正是某条圆锥曲线的定义. 到两定点距离之和为常数 ($>|F_1F_2|$) 是椭圆, 距离之差的绝对值为常数 ($<|F_1F_2|$) 是双曲线, 到定点与定直线距离相等是抛物线, 到定点距离为常数是圆. 认出定义, 方程的标准形直接写出, 剩下的只是定系数.

例题 13.202 动圆经过点 $F(2,0)$, 且与 y 轴相切. 求圆心 M 的轨迹方程.

解 设 $M(x,y)$. 动圆过 F, 半径为 $|MF|$; 动圆与 y 轴相切, 半径等于 M 到 y 轴的距离 $|x|$. 故

$$\sqrt{(x-2)^2+y^2}=|x|,$$

两边平方, 得 $(x-2)^2+y^2=x^2$, 化简得

$$y^2=4(x-1).$$

轨迹是抛物线, 顶点 $(1,0)$, 开口向右. 其实题目条件正是抛物线的定义. 焦点 $F(2,0)$, 准线即 y 轴, 定义法认出这一点, 方程法算出的结果与之互相印证.

定义法的要点是**先识别, 后计算**. 条件中出现距离之和、距离之差、等距一类字眼时, 我们先想定义, 再动笔平方.

13.14.3 相关点法

动点 M 由另一主动点 P 带动, 而 P 在已知曲线上运动. 此时把 M 与 P 的坐标关系写出, 反解 P 的坐标用 M 的表示, 代入 P 的曲线方程, 即得 M 的轨迹. 此法也叫**转移法**或**代入法**, 第13.6.1小节 已给出向量形式的操作骨架, 这里看一个对称点式的例子.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 13.203 点 P 在圆 $x^2+y^2=1$ 上运动, $A(2,0)$ 为定点, M 是 A 关于 P 的对称点. 求 M 的轨迹方程.

解 设 $M(x,y)$, $P(x_0,y_0)$.P 是 AM 的中点, 因此

$$x_0=\frac{x+2}{2},\qquad y_0=\frac{y}{2}.$$

由 $x_0^2+y_0^2=1$ 得

$$\left(\frac{x+2}{2}\right)^2+\left(\frac{y}{2}\right)^2=1\implies (x+2)^2+y^2=4.$$

知乎@余命数

轨迹是以 $(-2,0)$ 为圆心、半径 2 的圆 (原圆在以 A 为中心、比例为 2 的位似下的像).

知乎@余命数

13.14.4 交轨法

动点 M 是两条动直线的交点, 两条直线各含同一个参数 t. 把两直线方程联立, 消去 t, 即得 M 的轨迹方程. 消参的手法因题而异. 解出 t 代入, 或两式相乘、相除、平方相加, 把 t 整体消去.

例题 13.204 设 $A(-2,0)$, $B(2,0)$, 动点 P 满足直线 PA 与 PB 的斜率之积为 $-\dfrac{1}{4}$. 求 P 的轨迹方程.

解 设 $P(x,y)$, 两条动直线即 PA 与 PB. 斜率条件给出

$$\frac{y}{x+2}\cdot\frac{y}{x-2}=-\frac{1}{4}\quad(x\neq\pm2).$$

整理, 得 $4y^2=-(x^2-4)$, 即

$$\frac{x^2}{4}+y^2=1\quad(x\neq\pm2).$$

知乎@余命数

轨迹是椭圆 $\dfrac{x^2}{4}+y^2=1$, 但要去掉 $(\pm2,0)$ 两点(该处斜率不存在, 条件无意义).

斜率之积为常数的条件正是椭圆与双曲线的第三定义. 常数为负得椭圆 (去掉两顶点), 常数为正得双曲线. 交轨法把这个几何事实直接算了出来. 当两动直线来自曲线系时, 交轨法与 第13.3节 的曲线系原理相接. 交点满足 $f+\lambda g=0$, 消去 λ 即轨迹.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.14.5 参数法

为动点引进参数(角、斜率、时间), 把 x, y 分别写成参数的函数, 再消去参数. 参数选得贴合运动方式, 消参就只是恒等变形.

例题 13.205　点 P 在圆 $x^2+y^2=4$ 上运动, M 是 P 到 x 轴的垂线段 PQ 的中点. 求 M 的轨迹方程.

知乎@余命数

解 设 $P=(2\cos\theta,\,2\sin\theta)$, 则 $Q=(2\cos\theta,\,0)$, 中点

$$M=(2\cos\theta,\,\sin\theta).$$

记 $M=(x,y)$, 由 $\cos\theta=\dfrac{x}{2}$, $\sin\theta=y$ 及 $\cos^2\theta+\sin^2\theta=1$, 得

$$\frac{x^2}{4}+y^2=1.$$

圆沿竖直方向压缩为一半, 得到椭圆.

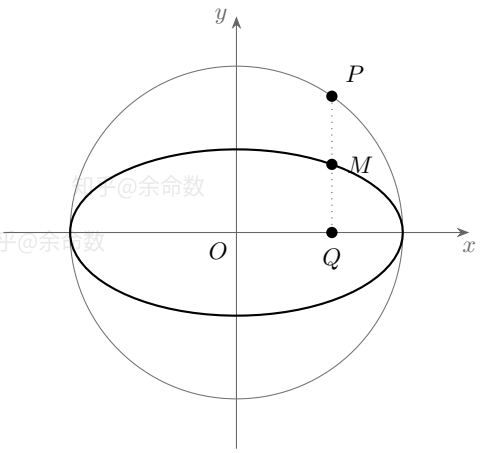


图 13.41　圆 $x^2+y^2=4$ 纵向压缩为椭圆 $\dfrac{x^2}{4}+y^2=1$

13.14.6 纯粹性与完备性

求出方程, 工作只做了一半. 我们还必须回答两个问题.

- **纯粹性**. 检查方程的曲线上每一点是否都满足原条件, 推导中的平方、约分、斜率存在的假设, 都可能引入不满足条件的点, 要挖去;
- **完备性**. 检查满足条件的点是否都在求出的曲线上, 分类讨论中遗漏的情形(斜率不存在、绝对值脱号、参数取不到), 都可能丢失整段轨迹, 要补回.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.206 (斜率假设带来的挖点) 设 $A(-1,0), B(1,0)$, 动点 M 满足 $MA \perp MB$. 求 M 的轨迹.

解 设 $M(x,y)$. 若两斜率都存在, 条件为

$$\frac{y}{x+1}\cdot\frac{y}{x-1}=-1 \implies x^2+y^2=1.$$

知乎@余命数

检验纯粹性. 点 $(\pm1,0)$ 满足方程, 但此时 MA 或 MB 的斜率不存在, M 与 A 或 B 重合, 谈不上垂直, 应挖去. 完备性上, M 与 A, B 重合以外的情形已被斜率写法覆盖(若 M 使 MA 竖直, 则 M 横坐标为 -1, 代入圆方程得 $M=(-1,0)=A$, 已排除).

故轨迹为圆 $x^2+y^2=1$ 去掉 $(\pm1,0)$ 两点.(这正是 Thales 定理的逆命题, 对照第13.1节 与 第13.13.1小节 的直径式. 直接用 $\overrightarrow{MA}\cdot\overrightarrow{MB}=0$ 写, 得到同一方程, 挖点讨论相同.)

例题 13.207 (绝对值脱号带来的补段) 动点 M 到两坐标轴的距离之和等于 2. 求 M 的轨迹.

解 设 $M(x,y)$, 条件为 $|x|+|y|=2$. 分象限脱去绝对值, 得

知乎@余命数

$$\begin{cases} x+y=2, & x\geq 0,\ y\geq 0,\\ -x+y=2, & x<0,\ y\geq 0,\\ -x-y=2, & x<0,\ y<0,\\ x-y=2, & x\geq 0,\ y<0. \end{cases}$$

轨迹是顶点为 $(\pm2,0), (0,\pm2)$ 的正方形边界. 若只写 $x+y=2$, 则丢了三段, 完备性不成立.

方程求出之后, 我们还应对着条件再走一遍, 条件里的每个限制(分母、绝对值、存在性)都应在轨迹上留下相应的刻痕.

13.14.7 轨迹方程例题

知乎@余命数

求轨迹方程的方法: 直接法、定义法、相关点法、交轨法、参数法、点差法, 以及结合向量、复数、立体几何等工具的轨迹问题.

知乎@余命数

例题 13.208 (直接法) 已知 M、N 两点坐标分别 $(-2,0), (2,0)$. 直线 MK、NK 相交于点 K, 且它们的斜率之和是 3, 则点 K 的轨迹方程为(　)

A. $3x^2-2xy-12=0(x\neq\pm2)$　　C. $\frac{x^2}{4}+\frac{y^2}{3}=1(x\neq\pm2)$

B. $3y^2-2yx-12=0(x\neq\pm2)$　　D. $\frac{x^2}{3}-\frac{y^2}{4}=1(x\neq\pm2)$

解 答案: A. 设 $K(x,y)$, 则直线 KM 的斜率为 $k_{KM}=\frac{y}{x+2}$, 直线 KN 的斜率为 $k_{KN}=\frac{y}{x-2}$, 依据题意可知, $k_{KM}+k_{KN}=\frac{y}{x+2}+\frac{y}{x-2}=3$, 化简得: $3x^2-2xy-12=0$, 因为直线 KM、KN 的斜率存在, 所以 $x\neq\pm2$, 所以 $3x^2-2xy-12=0(x\neq\pm2)$, 故选: A.

例题 13.209 (定义法) (2024·浙江温州·一模)动点 $M(x,y)$ 到定点 $F(-4,0)$ 的距离与 M 到定直线 $l:x=-\frac{25}{4}$ 的距离的比等于 $\frac{4}{5}$, 则动点 M 的轨迹方程是(　)

A. $\frac{x^2}{25}+\frac{y^2}{9}=1$　　C. $\frac{y^2}{25}+\frac{x^2}{9}=1$

B. $\frac{x^2}{25}+\frac{y^2}{16}=1$　　D. $\frac{y^2}{25}+\frac{x^2}{16}=1$

解 答案: A. 根据题意可得 $\frac{\sqrt{(x+4)^2+y^2}}{|x+\frac{25}{4}|}=\frac{4}{5}$, 平方化简可得 $9x^2+25y^2=25\times9$, 进而得 $\frac{x^2}{25}+\frac{y^2}{9}=1$, 故选: A.

例题 13.210 (相关点法) 设过点 $P(x,y)$ 的直线分别与 x 轴的正半轴和 y 轴的正半轴交于 A、B 两点, 点 Q 与点 P 关于 y 轴对称, O 为坐标原点. 若 $\overrightarrow{BP}=2\overrightarrow{PA}$, 且 $\overrightarrow{OQ}\cdot\overrightarrow{AB}=1$, 则点 P 的轨迹方程是__.

解 答案: $\frac{3}{2}x^2+3y^2=1\ (x>0,y>0)$. 设点 $P(x,y)$, 则 $Q(-x,y)$, 设 $A(a,0)$, $B(0,b)$, 则 $a>0,b>0$, $\therefore\overrightarrow{BP}=(x,y-b)$, $\overrightarrow{PA}=(a-x,-y)$, $\because\overrightarrow{BP}=2\overrightarrow{PA}$, $\therefore a=\frac{3}{2}x,b=3y$, $\therefore x>0,y>0$, 又 $\because\overrightarrow{AB}=(-a,b)=\left(-\frac{3}{2}x,3y\right)$、$\overrightarrow{OQ}=(-x,y)$、$\overrightarrow{OQ}\cdot\overrightarrow{AB}=1$, $\therefore\left(-\frac{3}{2}x\right)\cdot(-x)+3y\cdot y=1$, 即 $\frac{3}{2}x^2+3y^2=1\ (x>0,y>0)$. 故答案为: $\frac{3}{2}x^2+3y^2=1\ (x>0,y>0)$.

例题 13.211 (交轨法) 已知 A,B 分别为椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 的左、右顶点, 点 M, N 为椭圆上的两个动点, 满足线段 MN 与 x 轴垂直, 则直线 MA 与 NB 交点的轨迹方程为__.

解 答案: $\frac{x^2}{4}-\frac{y^2}{3}=1$. 因为 A, B 分别为椭圆 $\frac{x^2}{4}+\frac{y^2}{3}=1$ 的左、右顶点, 所以 $A(-2,0)$, $B(2,0)$, 设 MA 与 NB 的交点为 P、$P(x,y)$、$M(x_1,y_1)$、$N(x_1,-y_1)$, 由 $k_{PA}=k_{MA},k_{PB}=k_{NB}$, 得 $\frac{y}{x+2}=\frac{y_1}{x_1+2}$, $\frac{y}{x-2}=\frac{-y_1}{x_1-2}$,

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.212 (参数法) 方程 $x^2+y^2-4tx-2ty+3t^2-4=0$, 其中 t 为参数, 所表示的圆的圆心轨迹方程是(结果化为普通方程)

解 答案:$x-2y=0$. 圆 $x^2+y^2-4tx-2ty+3t^2-4=0$ 化为 $(x-2t)^2+(y-t)^2=2t^2+4$, 它表示以 $(2t,t)$ 为圆心, $\sqrt{2t^2+4}$ 为半径的圆, 设圆心坐标为 (x,y), 于是得 $\begin{cases}x=2t\\y=t\end{cases}$ (t 为参数), 消去 t 得:$x-2y=0$, 所以所求圆心轨迹方程是 $x-2y=0$.

例题 13.213 (点差法) 已知椭圆 $\frac{x^2}{4}+\frac{y^2}{9}=1$, 一组平行直线的斜率是 $\frac{3}{2}$, 当它们与椭圆相交时, 这些直线被椭圆截得的线段的中点轨迹方程是__.

解 答案:$y=-\frac{3}{2}x\left(-\sqrt{2}<x<\sqrt{2}\right)$. 设这组平行直线的方程为 $y=\frac{3}{2}x+m$,联立方程组 $\begin{cases}y=\frac{3}{2}x+m\\\frac{x^2}{4}+\frac{y^2}{9}=1\end{cases}$,整理得 $18x^2+12mx+4m^2-36=0$,由 $\Delta=144(18-m^2)>0$ 可得 $-3\sqrt{2}<m<3\sqrt{2}$, 则 $x_1+x_2=-\frac{2}{3}m$, 所以它们与椭圆交点的中点坐标为 $\left(-\frac{1}{3}m,\frac{1}{2}m\right)$, 即这些点均在轨迹 $y=-\frac{3}{2}x\left(-\sqrt{2}<x<\sqrt{2}\right)$ 上, 即直线被椭圆截得的线段的中点轨迹方程是 $y=-\frac{3}{2}x\left(-\sqrt{2}<x<\sqrt{2}\right)$. 故答案为: $y=-\frac{3}{2}x\left(-\sqrt{2}<x<\sqrt{2}\right)$.

例题 13.214 (立体几何与圆锥曲线的轨迹) 已知点 P 是正四面体 $A-BCD$ 内的动点, E 是棱 CD 的中点, 且点 P 到棱 AB 和棱 CD 的距离相等, 则点 P 的轨迹被平面 ABE 所截得的图形为(　)

A. 线段　　C. 双曲线的一部分

B. 椭圆的一部分　　D. 抛物线的一部分

解 答案: D. 在正四面体 $A-BCD$ 中, E 是棱 CD 的中点, 所以 $AE\perp CD$, $BE\perp CD$, 又 $AE\cap BE=E$, $AE,BE\subset$ 平面 ABE, 所以 $CD\perp$ 平面 ABE, 又点 P 的轨迹被平面 ABE 所截, 即点 P 在平面 ABE 内, $\therefore$ 点 P 到棱 CD 的距离为 PE. 在平面 ABE 内过点 P 作 $PF\perp AB$, 则 PF 为点 P 到棱 AB 的距离, 又点 P 到棱 AB 和棱 CD 的距离相等, 即 $PE=PF$, 因此, 在平面 ABE 内, 动点 P 到棱 AB 和到定点 E 的距离相等. 由抛物线的定义得, 动点 P 的轨迹是抛物线.

例题 13.215 (复数与圆锥曲线的轨迹) 已知 i 为虚数单位, 且 $z_0=\frac{1-3i}{1+2i}$, 复数 z 满足 $|z-z_0|=1$, 则复数 z 对应点的轨迹方程为(　)

知乎@余命数

A. $(x-1)^2+(y+1)^2=4$　　C. $(x+1)^2+(y+1)^2=1$

B. $(x-1)^2+(y-1)^2=4$　　D. $(x-1)^2+(y-1)^2=1$

解 答案:C. $z_0=\frac{1-3i}{1+2i}=\frac{(1-3i)(1-2i)}{5}=-1-i$, 表示点 $(-1,-1)$, 故复数 z 的轨迹是以 $(-1,-1)$ 为圆心, 半径为 1 的圆. 故选:C

例题 13.216 (向量与圆锥曲线的轨迹) 已知 O 是平面上一定点, A、B、C 是平面上不共线的三个点, 动点 P 满足 $\vec{OP}=\vec{OA}+\lambda(\vec{AB}+\vec{AC})$, $\lambda\in(0,+\infty)$, 则 P 的轨迹一定通过 $\triangle ABC$ 的(　)

A. 重心　　C. 内心

B. 外心　　D. 垂心

解 答案:A. 由题意 $\vec{AP}=\lambda(\vec{AB}+\vec{AC})$, 当 $\lambda\in(0,+\infty)$ 时, 由于 $\lambda(\vec{AB}+\vec{AC})$ 表示 $\vec{BC}$ 边上的中线所在直线的向量, $\therefore$ 动点 P 的轨迹一定通过 $\triangle ABC$ 的重心, 如图, 故选 A.

例题 13.217 (向量与圆曲线的轨迹) O 是平面内一定点, A,B,C 是平面内不共线三点, 动点 P 满足 $\vec{OP}=\vec{OA}+\lambda(\vec{AB}+\vec{AC})$, $\lambda\in[0,+\infty)$, 则 P 的轨迹一定通过 $\triangle ABC$ 的(　)

A. 外心　　C. 内心

B. 垂心　　D. 重心

解 答案:D. 由题意 $\lambda(\vec{AB}+\vec{AC})=\vec{OP}-\vec{OA}=\vec{AP}$, 而 $\vec{AB}+\vec{AC}$ 所在直线过 BC 中点, 即与 BC 边上的中线重合, 且 $\lambda\in[0,+\infty)$, 所以 P 的轨迹一定通过 $\triangle ABC$ 的重心, 故选:D

例题 13.218 (利用表达式求轨迹方程) (2024·全国模拟预测) 已知椭圆 C : $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ $(a>b>0)$ 的离心率为 $\frac{2}{3}$, 左、右顶点分别为 A_1 和 A_2, M 是椭圆 C 上一点, 且 $\triangle MA_1A_2$ 面积的最大值为 $3\sqrt{5}$. (1) 求椭圆 C 的标准方程; (2) 记 O 为坐标原点, 当点 M 与椭圆 C 的顶点不重合时, 过点 M 分别作直线 OM, MF, 其中直线 MF 不过坐标原点, 且不与坐标轴平行, 直线 OM, MF 与椭圆 C 交于异于点 M 的 E,F 两点, 直线 A_1E 与直线 A_2F 相交于点 D, 直线 OD 与直线 MF 相交于点 N, 求点 N 的轨迹方程.

知乎@余命数

知乎@余命数

解 答案: $\frac{x^2}{9}+\frac{y^2}{5}=1$. (1) 由题可知, $\begin{cases}\frac{c}{a}=\frac{2}{3}\\ \frac{1}{2}\times 2a\times b=3\sqrt{5}\\ a^2=b^2+c^2\end{cases}$, 解得 $\begin{cases}a=3\\ b=\sqrt{5}\\ c=2\end{cases}$, $\therefore$ 椭圆 C 的标准方程为 $\frac{x^2}{9}+\frac{y^2}{5}=1$.

(2) 由 (1) 知 $A_1(-3,0),A_2(3,0)$, 设 $M(x_1,y_1)(x_1,y_1\neq 0),F(x_2,y_2)$, 则 $E(-x_1,-y_1)$, 设直线 MF 的方程为 $x=my+n(m\neq 0,n\neq 0)$, 由 $\begin{cases}x=my+n\\ \frac{x^2}{9}+\frac{y^2}{5}=1\end{cases}$ 消去 x 并整理得 $(5m^2+9)y^2+10mny+5n^2-45=0$、$\therefore\Delta=(-10mn)^2-4\times(5m^2+9)(5n^2-45)=180(5m^2-n^2+9)>0$、$\therefore y_1+y_2=-\frac{10mn}{5m^2+9},y_1y_2=\frac{5n^2-45}{5m^2+9}$, 且 $y_1y_2\neq 0$, $\therefore n\neq\pm 3$, 设点 $D(x_0,y_0)$, 由 D,E,A_1 三点共线得 $k_{DA_1}=k_{EA_1}$, 即 $\frac{y_0}{x_0+3}=\frac{y_1}{x_1-3}$,

知乎@余命数

知乎@余命数

例题 13.219 (四心的轨迹方程) 设点 M、N 分别是不等边 $\triangle ABC$ 的重心与外心, 已知 $A(0,1)$、$B(0,-1)$, 且 $\vec{MN}=\lambda\vec{AB}$, 则动点 C 的轨迹 E__;

解 答案: $\frac{x^2}{3}+y^2=1(xy\neq 0)$. 设点 $C(x,y)$, 则 $\triangle ABC$ 的重心 $M\left(\frac{x}{3},\frac{y}{3}\right)$, $\because\triangle ABC$ 是不等边三角形, $\therefore xy\neq 0$, 再设 $\triangle ABC$ 的外心 $N(n,0)$, $\because$ 已知 $\vec{MN}=\lambda\vec{AB}$, $\therefore MN\parallel AB$, $\therefore n=\frac{x}{3}$, $\therefore$ 点 N 是 $\triangle ABC$ 的外心, $\therefore|NA|=|NC|$, 即 $\sqrt{\left(\frac{x}{3}\right)^2+1}=\sqrt{\left(\frac{x}{3}-x\right)^2+y^2}$, 化简整理得轨迹 E 的方程是 $\frac{x^2}{3}+y^2=1(xy\neq 0)$. $\therefore$ 动点 C 的轨迹 E 是指焦点在轴上的标准位置的一个椭圆(去掉其顶点). 故答案为: $\frac{x^2}{3}+y^2=1(xy\neq 0)$.

知乎@余命数

例题 13.220 在平面直角坐标系中, 若定点 $A(1,2)$ 与动点 $P(x,y)$ 满足向量 $\overrightarrow{OP}$ 在向量 $\overrightarrow{OA}$ 上的投影为 $-\sqrt{5}$, 则点 P 的轨迹方程为( )

A. $x-2y+5=0$　　C. $x+2y+5=0$

B. $x+2y-5=0$　　D. $x-2y-5=0$

解 答案: C. 由已知可得, 向量 $\overrightarrow{OP}$ 在向量 $\overrightarrow{OA}$ 上的投影为 $|\overrightarrow{OP}|\cos\langle\overrightarrow{OA},\overrightarrow{OP}\rangle=|\overrightarrow{OP}|\cdot\frac{\overrightarrow{OA}\cdot\overrightarrow{OP}}{|\overrightarrow{OA}|\cdot|\overrightarrow{OP}|}=\frac{\overrightarrow{OA}\cdot\overrightarrow{OP}}{|\overrightarrow{OA}|}=\frac{x+2y}{\sqrt{5}}=-\sqrt{5}$, 整理可得 $x+2y+5=0$. 因此, 点 P 的轨迹方程为 $x+2y+5=0$. 故选: C.

例题 13.221 已知定点 $P(-4,0)$ 和定圆 $Q:x^2+y^2=8x$, 动圆 M 和圆 Q 外切, 且经过点 P, 求圆心 M 的轨迹方程

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 答案：双曲线 $\frac{x^2}{4}-\frac{y^2}{12}=1$ 的左支. 结合图象可得，$|MQ|-|MP|=4$，可得 $a=2,c=4$，则 $b=2\sqrt{3}$，M 的轨迹为双曲线 $\frac{x^2}{4}-\frac{y^2}{12}=1$ 的左支. 故答案为双曲线 $\frac{x^2}{4}-\frac{y^2}{12}=1$ 的左支.

例题 13.222　已知圆 $C_1:(x+1)^2+y^2=25$，圆 $C_2:(x-1)^2+y^2=1$，动圆 M 与圆 C_2 外切，同时与圆 C_1 内切，则动圆圆心 M 的轨迹方程为(　)

A. $\frac{x^2}{3}+y^2=1$　　C. $\frac{x^2}{9}+y^2=1$

B. $\frac{x^2}{3}+\frac{y^2}{2}=1$　　D. $\frac{x^2}{9}+\frac{y^2}{8}=1$

解 答案：D. 如图，由题意得：$|C_1M|=5-|MQ|$，$|C_2M|=1+|MP|$，其中 $|MQ|=|MP|$，所以 $|C_1M|+|C_2M|=5-|MQ|+1+|MP|=6>2=|C_1C_2|$，由椭圆定义可知，动圆圆心 M 的轨迹为以 C_1,C_2 为焦点的椭圆，设 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$，则 $2a=6,c=1$，解得：$a=3,b^2=a^2-c^2=9-1=8$，故动圆圆心 M 的轨迹方程为 $\frac{x^2}{9}+\frac{y^2}{8}=1$. 故选：D

例题 13.223　已知 $M(-2,0)$，P 是圆 $N:x^2-4x+y^2-32=0$ 上一动点，线段 MP 的垂直平分线交 NP 于点 Q，则动点 Q 的轨迹方程为__.

解 答案：$\frac{x^2}{9}+\frac{y^2}{5}=1$. 由题意，可知圆 N 的标准方程为 $(x-2)^2+y^2=36$，圆心为 $N(2,0)$，半径为 6.∵ 线段 MP 的垂直平分线交 NP 于点 Q，如图，∴$|QP|=|QM|$，∴$|QM|+|QN|=|QP|+|QN|=|PN|=6>|MN|=4$，∴ 点 Q 的轨迹是以 M，N 为焦点的椭圆，∴$a=3$、$c=2$、$b=\sqrt{a^2-c^2}=\sqrt{5}$，∴ 其轨迹方程为 $\frac{x^2}{9}+\frac{y^2}{5}=1$. 故答案为：$\frac{x^2}{9}+\frac{y^2}{5}=1$.

例题 13.224　已知定点 $R(1,0)$，圆 $S:x^2+y^2+2x-15=0$，过 R 点的直线 L_1 交圆于 M，N 两点，过 R 点作直线 $L_2\parallel SN$ 交 SM 于 Q 点，求 Q 点的轨迹方程；

解 答案：$\frac{x^2}{4}+\frac{y^2}{3}=1$. 因为 $S:x^2+y^2+2x-15=0$，即 $(x+1)^2+y^2=16$，所以 $S(-1,0)$，半径为 $r=4$，如图，根据题意可知 $|SM|=|SN|=r=4$，又 $RQ\parallel SN$，所以 $\frac{|RQ|}{|SN|}=\frac{|QM|}{|SM|}$，故 $|RQ|=|QM|$，又 $R(1,0)$，所以 $|QS|+|QR|=|SM|=4>|SR|=2$，故动点 Q 的轨迹是以 S,R 为焦点，长轴长为 4 的椭圆，这里 $a=2,c=1$，故 $a^2=4,b^2=4-1=3$，所以 Q 点的轨迹方程为：$\frac{x^2}{4}+\frac{y^2}{3}=1$.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

例题 13.225 设 O 为坐标原点, $F(2,0)$, 点 A 是直线 $x=-2$ 上一个动点, 连接 AF 并作 AF 的垂直平分线 l, 过点 A 作 y 轴的垂线交 l 于点 P, 则点 P 的轨迹方程为__.

解 答案: $y^2=8x$. 如图, 由垂直平分线的性质可得 $|PA|=|PF|$, 符合抛物线第一定义, 抛物线开口向右, 焦点坐标为 $F(2,0)$, 故 $p=4,2p=8$, 点 P 的轨迹方程为 $y^2=8x$. 故答案为: $y^2=8x$

例题 13.226 已知圆 C 的方程为 $x^2+y^2=16$, 直线 l 为圆 C 的切线, 记 $A(-2,0),B(2,0)$ 两点到直线 l 的距离分别为 d_1,d_2, 动点 P 满足 $|PA|=d_1$, $|PB|=d_2$, 则动点 P 的轨迹方程为( )

A. $x^2+y^2=4$

B. $\frac{x^2}{16}+\frac{y^2}{12}=1$

C. $\frac{x^2}{16}-\frac{y^2}{12}=1$

D. $y^2=4x$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

解 答案: B. 如图, 分别过点 A,O,B 做直线 l 的垂线, 垂足分别为 A_1,O_1,B_1, 则 $AA_1\parallel OO_1\parallel BB_1, d_1=|AA_1|, d_2=|BB_1|$, 切点为 O_1, 因为 $A(-2,0),B(2,0)$, 所以 O 是 AB 的中点, 所以 OO_1 是梯形 ABB_1A_1 的中位线, 所以 $|OO_1|=\frac{|AA_1|+|BB_1|}{2}=\frac{d_1+d_2}{2}$, 又因为圆 C 的方程为 $x^2+y^2=16$, $r=4$, 所以 $|OO_1|=r=4$, 所以 $d_1+d_2=8$, 即 $|PA|+|PB|=8$, 所以动点 P 的轨迹是以 A,B 为焦点, 长轴长为 8 的椭圆, 设椭圆的方程为 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$, 则 $2a=8,c=2$, 所以 $a=4,a^2=16,b^2=a^2-c^2=12$, 所以动点 P 的轨迹方程为 $\frac{x^2}{16}+\frac{y^2}{12}=1$. 故选: B.

例题 13.227 已知点 P 是椭圆 $\frac{x^2}{2}+y^2=1$ 上的动点, $PM\perp x$ 于点 M, 若 $\overrightarrow{PN}=\frac{1}{2}\overrightarrow{NM}$, 则点 N 的轨迹方程为( )

A. $\frac{x^2}{2}+\frac{9y^2}{4}=1$

B. $\frac{x^2}{2}+\frac{y^2}{4}=1$

C. $\frac{x^2}{2}+\frac{4y^2}{9}=1$

D. $\frac{x^2}{2}+\frac{y^2}{9}=1$

解 答案: A. 由于点 P 是椭圆 $\frac{x^2}{2}+y^2=1$ 上的动点, 设 $P(x_0,y_0)$, 则 $\frac{x_0^2}{2}+y_0^2=1$, 又 $PM\perp x$ 于点 M, 则 $M(x_0,0)$; 设 $N(x,y)$, 由 $\overrightarrow{PN}=\frac{1}{2}\overrightarrow{NM}$, 得 $(x-x_0,y-y_0)=\frac{1}{2}(x_0-x,-y)$, 则 $\begin{cases}x_0=x\\ y_0=\frac{3}{2}y\end{cases}$, 代入 $\frac{x_0^2}{2}+y_0^2=1$, 得 $\frac{x^2}{2}+\frac{9y^2}{4}=1$, 即点 N 的轨迹方程为 $\frac{x^2}{2}+\frac{9y^2}{4}=1$, 故选: A

知乎@余命数

13.14.8 本节小结

- **直接法**. 条件直接翻译, 对称建系省力.
- **定义法**. 先识别圆锥曲线定义, 再定系数, 避免盲目平方.
- **相关点法**. 反解主动点坐标, 代入已知曲线.
- **交轨法**. 两动直线联立消参; 与曲线系原理相接.
- **参数法**. 参数贴合运动, 消参用恒等式.
- **纯粹性与完备性**. 挖去增点, 补回漏段; 对条件重走一遍.

13.14.9 本节习题

13.79. 动点 M 到点 $F(3,0)$ 的距离等于它到直线 $x=-3$ 的距离. 求 M 的轨迹方程.

知乎@余命数

13.80. 动点 M 到 $F_1(-3,0)$, $F_2(3,0)$ 的距离之差的绝对值等于 2. 求 M 的轨迹方程.

13.81. 点 P 在抛物线 $y^2=4x$ 上运动, M 是 P 与点 $(-1,0)$ 连线的中点. 求 M 的轨迹方程.

13.82. $A(-1,0)$, $B(1,0)$ 为定点, 动直线 PA 与 PB 的斜率之积为 1. 求 P 的轨迹方程, 并说明需要挖去的点.

13.83. 点 $P=(\cos\theta,\,2\sin\theta)$, 消去参数 θ, 求 P 的轨迹方程.

13.84. 动点 M 到 x 轴的距离与到 y 轴的距离之积等于 1. 求 M 的轨迹, 并说明完备性上需要注意什么.

13.15 解析法证几何题

我们用坐标法既能解计算题, 也能证几何题. 第12.3节 已给出一种保底的建系程序, 把点置于原点、把线贴到轴上, 然后硬算到底. 保底之外还有更精巧的层次. 坐标系可以建得贴合图形的对称性; 题目只涉及平行与分比时, 我们甚至不必坚持两轴垂直; 动笔之前还可以先用综合几何把问题化小.

13.15.1 建系的技艺

我们在 第12.3节 中把建系原则概括为三条, 即原点特殊化、轴线特殊化、利用对称性. 落实到证明题, 还有两件顺手的工具.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **单位化**. 图形只确定到相似时, 把某条关键线段的长直接设为 1(或 2, 看中点是否出现), 消去一个尺度参数;
- **设而不求**. 长度、斜率尽量以平方形式出现, 不急于开方、不急于求出中间点的坐标.

例题 13.228 用坐标法证明: 菱形的对角线互相垂直.

解 菱形首先是平行四边形. 取顶点 A 为原点, AB 所在直线为 x 轴, 设

$$A=(0,0),\quad B=(b,0),\quad D=(d_1,d_2),\quad C=B+D=(b+d_1,\,d_2).$$

菱形的含义是邻边相等, 即 $|AB|^2=|AD|^2$, 也就是

$$b^2=d_1^2+d_2^2.$$

两条对角线的方向向量分别为

$$\overrightarrow{AC}=(b+d_1,\,d_2),\qquad \overrightarrow{BD}=(d_1-b,\,d_2).$$

其数量积

$$\overrightarrow{AC}\cdot\overrightarrow{BD}=(b+d_1)(d_1-b)+d_2^2=d_1^2+d_2^2-b^2=0.$$

故 $AC\perp BD$.

整个证明没有出现一次根号. 等腰条件以平方形式进入, 又以平方形式消耗掉. 若条件与结论都以长度陈述, 则平方往往是更经济的中间语言.

13.15.2 斜坐标

有一类几何命题, 通篇只谈平行、分比、中点、共点、共线, 不涉及长度与角度. 对这类命题, 可以连两轴垂直的要求都放弃.

定义 13.9 (斜坐标系) 在平面上取一点 O 为原点, 取两个不共线的向量 $\overrightarrow{e_1},\overrightarrow{e_2}$ 的方向为坐标轴, 并以它们的长度为各自轴上的单位长. 对任意点 P, 由

$$\overrightarrow{OP}=x\overrightarrow{e_1}+y\overrightarrow{e_2}$$

唯一确定的 (x,y) 称为 P 的**斜坐标**.

斜坐标保留了解析法的大部分代数结构.

- 定比分点公式照旧成立($P=(1-t)P_1+tP_2$, 坐标分别组合), 特别地, 中点坐标是两端坐标之半;

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 直线的方程仍是一次方程 $Ax+By+C=0$;
- 两直线平行 $\iff$ 一次项系数成比例.

失效的只有与度量有关的公式. 距离、垂直、角度都不能按直角坐标的公式计算. 这也划定了斜坐标的适用范围. **仿射性质**(平行、分比、共点、共线)的命题可以放心使用; 涉及度量的命题请回到直角坐标. 第11.2节 的等和线与重心坐标, 第13.3节 的仿射变换, 都与斜坐标同源.

例题 13.229　在 $\triangle ABC$ 中, D 在 BC 上且 $BD:DC=1:2$, E 为 AC 的中点, AD 与 BE 交于 P. 求 $AP:PD$.

解　题目只涉及分比与共点, 我们取斜坐标系. 以 A 为原点, AB, AC 方向为两轴, 则

$$A=(0,0),\quad B=(1,0),\quad C=(0,1).$$

由分点公式, $D=\dfrac{2B+C}{3}=(\dfrac{2}{3},\dfrac{1}{3})$, $E=(0,\dfrac{1}{2})$.

设 P 在 AD 上, 即 $P=tD=(\dfrac{2t}{3},\dfrac{t}{3})$; 又 P 在 BE 上, 即 $P=(1-s)B+sE=(1-s,\dfrac{s}{2})$. 比较坐标, 得

$$\frac{2t}{3}=1-s,\qquad \frac{t}{3}=\frac{s}{2}.$$

由第二式 $s=\dfrac{2t}{3}$, 代入第一式得 $\dfrac{2t}{3}=1-\dfrac{2t}{3}$, 于是 $t=\dfrac{3}{4}$. 故

$$AP:PD=t:(1-t)=3:1.$$

注意坐标 $(1,0),(0,1)$ 中藏着单位化. $|AB|,|AC|$ 各取为所在轴的单位长; 两个单位甚至可以不等长. 正因为题目不问长度, 这种自由是合法的.

13.15.3 几何知识的应用

建系之后立刻硬算, 是解析法的下策. 上策是先用综合几何把问题化简, 再把剩下的部分交给坐标. 下列的对照可以说明差距.

例题 13.230　从圆 $x^2+y^2=r^2$ 外一点 $P(x_0,y_0)$ 引两条切线, 切点为 T_1,T_2. 求直线 T_1T_2(切点弦)的方程.

解　硬算的路径是直接设出切点坐标, 由切点在圆上、半径垂直于切线列两组方程, 再消元. 过程不短.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

先用几何知识. 因为 $\angle OT_1P=\dfrac{\pi}{2}$, 由 Thales 定理的逆 (第13.1节) 可知, T_1 在以 OP 为直径的圆上; 同理 T_2 也在其上. 于是 T_1T_2 是已知圆与以 OP 为直径的圆的**公共弦**. 两圆方程相减即可 (第13.13.2小节).

$$(x^2+y^2-r^2)-(x^2+y^2-x_0x-y_0y)=0,$$

即

$$x_0x+y_0y=r^2.$$

三步以内得到答案, 与 第13.4.6小节 的替换法则一致.

几何知识在这里完成了两件工作. 它把切点转译为某圆上的点 (垂直 $\to$ 直径), 把切点弦转译为公共弦, 即两式相减. 解析法的长处恰恰来自这种几何与代数的配合; 盲目计算等于自断一臂.

13.15.4 牛顿线

完全四边形由四条直线 (每三条不共点) 组成, 它们两两相交得六个交点. 把其中四条边看成四边形的边, 则六个交点中有四边形的四个顶点与两组对边的交点; 四边形的两条对角线, 加上两组对边交点的连线, 共三条线段, 称为完全四边形的三条**对角线**.

定理 13.37 (牛顿线) 完全四边形的三条对角线的中点共线. 这条直线称为**牛顿线**.

证明 结论只涉及中点与共线, 是仿射性质, 我们用斜坐标. 设四边形 $ABCD$ 中, 对边 AB,CD 交于 E, 对边 BC,AD 交于 F. 取 A 为原点, AB,AD 方向为两轴, 则

$$A=(0,0),\quad B=(1,0),\quad D=(0,1),\quad C=(u,v).$$

求两组对边的交点.E 在 AB ($y=0$) 与 CD 上, 算得 $E=(\dfrac{u}{1-v},0)$; F 在 AD ($x=0$) 与 BC 上, 算得 $F=(0,\dfrac{v}{1-u})$.

三条对角线的中点为

$$M_1=\frac{A+C}{2}=\Big(\frac{u}{2},\frac{v}{2}\Big),\quad M_2=\frac{B+D}{2}=\Big(\frac{1}{2},\frac{1}{2}\Big),\quad M_3=\frac{E+F}{2}=\Big(\frac{u}{2(1-v)},\frac{v}{2(1-u)}\Big).$$

计算斜率, 得

$$k_{M_1M_2}=\frac{\frac{1}{2}-\frac{v}{2}}{\frac{1}{2}-\frac{u}{2}}=\frac{1-v}{1-u},$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$k_{M_1M_3}=\frac{\frac{v}{2(1-u)}-\frac{v}{2}}{\frac{u}{2(1-v)}-\frac{u}{2}}=\frac{\frac{uv}{2(1-u)}}{\frac{uv}{2(1-v)}}=\frac{1-v}{1-u}.$$

两斜率相等且有公共点 M_1, 故 M_1,M_2,M_3 共线.(分母为零的情形对应图形退化, 另行检验, 结论不变.) □

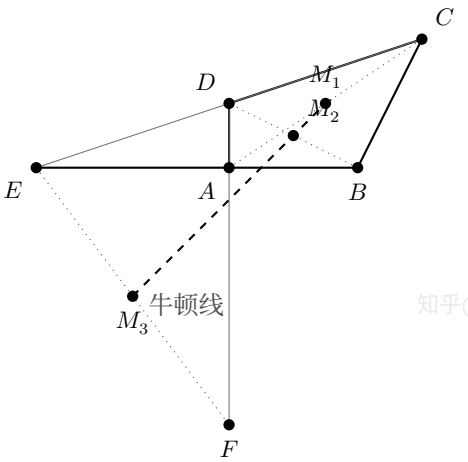


图 13.42　牛顿线:完全四边形三条对角线 AC、BD、EF 的中点 M_1,M_2,M_3 共线

13.15.5 竞赛名题选讲

本节的方法备齐之后, 我们可以看两道名题. 第一道展示曲线系的作用, 第二道展示建系与根轴的配合.

例题 13.231 (蝴蝶定理) 设 M 为圆中弦 PQ 的中点, 过 M 作两弦 AB,CD, 直线 AD,BC 分别交 PQ 于 X,Y. 求证:M 也是 XY 的中点.

解　取 M 为原点, PQ 所在直线为 x 轴, 设 $P=(-p,0)$, $Q=(p,0)$. 将圆的方程 $x^2+y^2+Dx+Ey+F=0$ 代入 P,Q, 得

$$p^2\pm Dp+F=0\implies D=0,\ F=-p^2,$$

故圆为 $x^2+y^2+Ey-p^2=0$.

设 $AB:y=kx$, $CD:y=mx$, 斜率不存在的情形由对称性单独验证. 两直线合成的退化二次曲线

$$(y-kx)(y-mx)=0$$

经过 A,B,C,D 四点. 由二次曲线系原理 (第13.3节), 过这四点的任意二次曲线都属于

$$(y-kx)(y-mx)+\lambda(x^2+y^2+Ey-p^2)=0.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

直线对 AD, BC 合成的退化二次曲线也在其中, 对应某个 λ_0. 它与 x 轴 ($y=0$) 的交点满足

$$kmx^2+\lambda_0(x^2-p^2)=0 \implies (km+\lambda_0)x^2=\lambda_0 p^2,$$

两根互为相反数($km+\lambda_0\neq 0$, 否则直线对过原点, 与构图矛盾). 即 X, Y 关于 M 对称, M 是 XY 的中点.

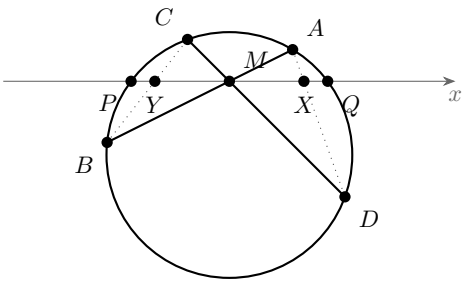


图 13.43 蝴蝶定理: M 是 PQ 的中点, 也是 XY 的中点

证明的要点只有一处. 弦 PQ 与直线对 AD, BC 在同一曲线系中, 它们与 x 轴的交点由同一个只含 x^2 项的方程控制, 故对称性自动成立. 若逐点求坐标, 则这道题会很辛苦.

例题 13.232 (2000 年国际数学奥林匹克第 1 题) 两圆 Γ_1, Γ_2 交于 M, N 两点. 直线 ℓ 分别切 Γ_1, Γ_2 于 A, B, 且 M 比 N 更靠近 ℓ. 过 M 作 ℓ 的平行线, 分别再交 Γ_1, Γ_2 于 C, D. 直线 CA 与 DB 交于 E; 直线 AN 与 CD 交于 P, 直线 BN 与 CD 交于 Q. 求证: $EP=EQ$.

解 MN 是两圆的根轴 (第13.13.2小节). 设它交 ℓ 于 T.T 对两圆的幂相等, 而幂即切线长的平方, 故有

$$TA^2=TB^2 \implies TA=TB,$$

即根轴经过 AB 的中点 T.

取 T 为原点, ℓ 为 x 轴, 并单位化 $TA=TB=1$, 设

$$A=(-1,0), \quad B=(1,0), \quad M=(u,v)\ (v\neq 0).$$

Γ_1 切 x 轴于 A, 圆心为 $(-1,r_1)$; 代入 M 得

$$(u+1)^2+(v-r_1)^2=r_1^2 \implies (u+1)^2+v^2=2vr_1.$$

直线 CD 即 $y=v$.Γ_1 与 $y=v$ 的另一交点 C 满足

$$(x+1)^2=r_1^2-(v-r_1)^2=2vr_1-v^2=(u+1)^2,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $x=u$, 即 M 或 $x=-u-2$, 得 $C=(-u-2,\,v)$. 同理由 Γ_2 得 $D=(2-u,\,v)$.

直线 CA 方向为 $A-C=(u+1,-v)$, 过 $A(-1,0)$, 参数式 $(-1+s(u+1),\,-sv)$; 直线 DB 方向为 $B-D=(u-1,-v)$, 过 $B(1,0)$, 参数式 $(1+t(u-1),\,-tv)$. 交点处纵坐标给出 $s=t$, 横坐标给出

$$-1+s(u+1)=1+s(u-1)\implies 2s=2\implies s=1,$$

代回得 $E=(u,-v)$, 即 E 是 M 关于 x 轴的对称点.

N 在根轴 TM 上, 设 $N=(\lambda u,\lambda v)$. 直线 AN 交 CD $(y=v)$ 于

$$P=\Big(u+\frac{1}{\lambda}-1,\,v\Big),$$

直线 BN 交 CD 于

$$Q=\Big(u+1-\frac{1}{\lambda},\,v\Big).$$

PQ 的中点为

$$\Big(\frac{(u+\frac{1}{\lambda}-1)+(u+1-\frac{1}{\lambda})}{2},\,v\Big)=(u,\,v),$$

恰在过 E 且垂直于 CD 的直线 $x=u$ 上.$CD\parallel\ell$ 水平, 故 E 在 PQ 的垂直平分线上, 于是 $EP=EQ$.

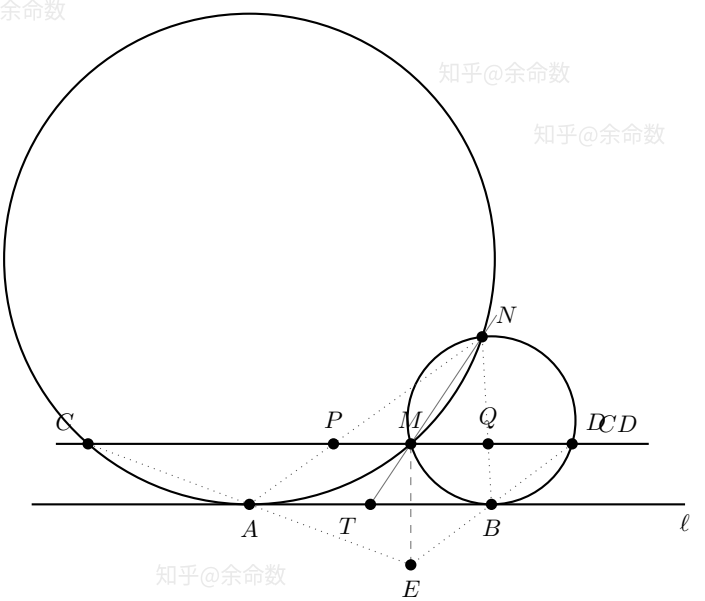


图 13.44　IMO 2000 第 1 题的建系:T 为原点, $TA=TB=1$

回顾全程. 根轴给出 $TA=TB$, 单位化把图形洗净, 四个关键点的坐标依次读出, 结论落在一条竖直线上. 每一步的计算量都不超过一次方程. 这是几何先行、建系随后的标准示范.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

13.15.6 一道题的推广

好定理可以再生长. 以蝴蝶定理为例, 把 M 为 PQ 中点的条件放宽为 M 是 PQ 上任意一点, 看结论如何变化.

我们重走曲线系的计算. 设 $P=(-p,0)$, $Q=(q,0)$ $(p,q>0)$, M 仍为原点. 圆 $x^2+y^2+Dx+Ey+F=0$ 过 P,Q, 得 $D=p-q$, $F=-pq$. 直线 $AB:y=kx$, $CD:y=mx$ 不变. 曲线系与 x 轴的交点满足

$$(km+\lambda)x^2+\lambda Dx+\lambda F=0,$$

设两根为 x_X,x_Y, 即 X,Y 的横坐标, 以 M 为原点看是有向线段. 由韦达定理得

$$\frac{1}{x_X}+\frac{1}{x_Y}=\frac{x_X+x_Y}{x_Xx_Y}=\frac{-\lambda D}{-\lambda F}=\frac{p-q}{pq}=\frac{1}{q}-\frac{1}{p}.$$

写成有向线段的等式, 即

$$\frac{1}{MX}+\frac{1}{MY}=\frac{1}{MQ}-\frac{1}{MP},$$

这就是 **Candy 定理**. 当 M 为 PQ 中点时 MP,MQ 反向等长, 右端为零, 回到蝴蝶定理. 同一个曲线系, 同一个计算, 一般化几乎不增加成本. 对称的结论由此推广为调和的结论.

> **注 13.18 (机器证明)**
> 本节的所有证明遵循同一模式. 我们先建系、设坐标、消元, 再验证多项式恒等式. 吴文俊先生把这一模式推向了机械化. 他将几何命题的条件与结论都代数化为多项式方程组, 再用特征列方法机械地判定前者是否蕴含后者, 于是大量平面几何定理获得了机器证明. 解析几何是这条机械化道路的基础.

13.15.7 本节小结

- **建系**. 对称取轴, 单位化消尺度, 条件与结论以平方形式往来, 不设多余的字母.
- **斜坐标**. 只涉仿射性质(平行、分比、共点、共线)时, 两轴方向与单位长皆可取; 直线仍为一次式, 分点公式照旧, 度量公式不可用.
- **几何先行**. 先用综合几何化简(垂直 $\to$ 直径, 切点弦 $\to$ 公共弦), 再把剩下的交给坐标.
- **牛顿线**. 完全四边形三对角线中点共线, 斜坐标下三次斜率计算.
- **名题**. 蝴蝶定理用曲线系(x 轴截距只含 x^2 项); IMO 2000/1 用根轴 + 单位化 ($TA=TB=1$).
- **推广**. 中点改任意点, 蝴蝶定理生长为 Candy 定理, 同一计算多走半步.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.15.8 本节习题

13.85. 用斜坐标重解：在 $\triangle ABC$ 中，D 在 BC 上且 $BD:DC=2:1$，E 在 CA 上且 $CE:EA=2:1$，AD 与 BE 交于 P. 求 $AP:PD$.

13.86. 用坐标法证明：三角形两边中点的连线平行于第三边，且等于第三边的一半.

13.87. 从点 $P(4,1)$ 向圆 $x^2+y^2=9$ 引两条切线. 求切点弦所在直线的方程与切线长.

13.88. 取 $A(0,0)$、$B(2,0)$、$C(3,2)$、$D(0,1)$，求四边形 $ABCD$ 三条对角线中点的坐标，并验证它们共线（对照本节牛顿线配图）.

13.89. 说明：当 M 为圆心时，蝴蝶定理由图形的对称性立即成立，无须曲线系.

知乎@余命数

13.90. 设 $A(-2,0)$，$B(2,0)$，动点 P 满足直线 PA，PB 的斜率之积为 1. 求 P 的轨迹，并说明这与蝴蝶定理的推广（中点改任意点）的共同思路.

知乎@余命数
知乎@余命数

13.16 解析几何思想与方法论

我们在前十七节中依次讨论了向量与共线，直线与圆，二次曲线，设点法，设线法，以及一次式与法线式、直线偶与齐次化、三角形的心、共圆点与共轴圆、轨迹方程、解析法证几何题等专题. 本节不再引入新公式，转而梳理贯穿始终的几个问题. 坐标语言在说什么，何时设点、何时设线，以及设而不求与整体代换为什么有效. 我们还要给出题型对照与常见陷阱. 更深的射影观点见同卷 第15章（选读），极坐标与焦弦见同卷 第14章.

13.16.1 坐标翻译

解析几何用坐标语言**翻译**几何对象与几何关系.

- 点、直线、曲线 $\mapsto$ 坐标或方程 $F(x,y)=0$；
- 垂直、定比、相切 $\mapsto$ 点乘为零、分点公式、$\Delta=0$.

翻译是否省力，取决于两件事. 一是坐标系选得是否贴合对称性，二是参量选得是否匹配自由度.

知乎@余命数

13.16.1.0.1 **自由度** 一条曲线上的动点有 1 个自由度（曲线方程消掉一个坐标）；平面上自由动点有 2 个自由度. 题中若出现弦的两端，名义上四个坐标，减去两条曲线约束剩两个自由度，再由过定点、定斜率、定中点等条件补足. 设点法与设线

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

法都是在为这 1 至 2 个自由度命名. 前者命名为 (x_0,y_0) 或 $(x_1,y_1),(x_2,y_2)$, 后者命名为 k,m.

13.16.1.0.2 与 第19章 的呼应 先数自由度、再搭建解题框架的思想与数学思想专章一致. 解析几何只是这一原则在坐标语言下的具体化.

13.16.2 设点与设线的决策树

题面信号	优先路径
唯一动点在曲线上; 切线截距, 距离/面积最值	设点法·单动点 (第13.5节)
弦端坐标显式出现; 垂直, 定比, 中点轨迹	设点法·双动点 / 定比点差 (第13.4节–第13.6节)
显含角度, 离心角, 旋转对称	参数方程设点或旋转设点 (第13.5节)
弦长, 焦点弦, 中点横坐标, Δ 与韦达	设线法 · 联立与韦达 (第13.7节–第13.9节)
多条动直线, 斜率关系, 截距定值	设线法·形式选择 (第13.8节)
点对点约束难写成 k 的封闭式, 如 $\overrightarrow{FP}\cdot\overrightarrow{FQ}=0$	先设点展开, 再借韦达回 k(混合)
过原点弦的斜率和/积定值, $OP\perp OQ$ 型条件	齐次化 (第13.11节)
角平分线, 点的侧向与区域, 距离符号讨论	一次式与法线式 (第13.10节)
求轨迹方程; 交点随参数运动	轨迹五法 (第13.14节)
四点共圆, 三圆公共弦, 圆幂定值	共圆点与共轴圆 (第13.13节)
纯几何命题的分比、共点、共线	解析证法与斜坐标 (第13.15节)
调和比, 极点极线, 交比	射影工具 (第15章)
焦半径, 焦弦, 半通径	极坐标 (第14章)

注 13.19 (没有唯一正确答案)
同一题常有两条通路. 判别标准如下.(1) 中间表达式是否对称; (2) 分支(斜率不存在等)是否好写; (3) 计算量是否可控. 大题中途换轨完全正常. 例如先以设线开局, 在需要点积时再设交点坐标.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.16.3 设而不求与整体代换

两条看似相反的原则, 实则同源.

13.16.3.0.1　设而不求　弦的两端 P_1, P_2 的具体坐标往往不必解出. 韦达给出 x_1+x_2, x_1x_2; 点差给出中点与斜率的关系. 目标式若能写成对称多项式, 就应停在和与积的形式.

13.16.3.0.2　整体代换　动点 (x_0, y_0) 在曲线 $F(x_0, y_0)=0$ 上时, 目标式中出现的 x_0^2, y_0^2 组合应优先用 $F=0$ 整块替换, 省去解出 $y_0=\pm\sqrt{\cdots}$ 再代入的弯路. 这是设点法消元的一般做法, 也与设线法中用曲线方程消二次项的手法同源.

13.16.3.0.3　几条原则

能对称则对称, 用韦达或点差; 能整体则整体, 借助曲线约束; 选好直线形式以减少竖直枝讨论; 最后再回代定义域与 Δ.

知乎@余命数

13.16.4 题型速查

题型	推荐骨架
轨迹方程	设点 $\to$ 关系反解 $\to$ 消参; 或设线读中点; 五法见 第13.14节
定值 / 定线	对称式化简; 过定点用 t_1+t_2, t_1t_2
最值	单点参数函数 / 基本不等式 / 设线后对 k 求导
存在性 / 范围	$\Delta\geq 0$ + 参数范围; 必要性探路
弦中点	点差(设点)或韦达(设线)
焦点弦	设线过焦点; 焦半径 $a\pm ex$; 极坐标加深

知乎@余命数

13.16.5 常见陷阱

1. **斜率不存在**. 点斜式 / 斜截式漏掉竖直直线; 用一般式或 $x=ty+n$ 与 $y=kx+m$ 双轨覆盖 (第13.8节).
2. **联立后未检验** Δ. 求出 k 却对应无实交点.
3. **中点在曲线外**. 点差公式在弦不存在时无意义; 先检验中点在曲线内部 (椭圆).
4. **离心角当极角**.$(a\cos\theta, b\sin\theta)$ 中 θ 不是极角 (第14章).
5. **双曲线两支 / 抛物线参数定义域**. 参数覆盖不全导致丢解.
6. **有向与无向**. 定比 λ 外分时符号; 圆幂的有向乘积.
7. **过早求根**. 能用韦达却展开求根公式, 则计算量剧增且易错.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.16.6 本章工具链回顾

向量/共线 $\to$ 直线与圆 $\to$ 二次曲线 $\to$ $\begin{cases}\text{设点法 (sec06〜08)}\\ \text{设线法 (sec09〜11)}\end{cases}$ $\to$ 技巧专题 (sec12〜17) $\to$ 方

再往后, 射影几何把极点极线与交比变成比坐标更基本的不变量; 极坐标把焦点弦写成统一的 $\dfrac{ep}{1-e\cos\theta}$. 坐标法不会过时, 只是会嵌入更大的几何语言.

13.16.7 本节小结

- 解析几何 = 翻译 + 参量选择 + 对称化代数;
- 设点法与设线法是两大范式, 混合使用是常态;
- 设而不求、整体代换、形式选择, 三者决定计算量;
- 决策依据是题面信号与自由度.

13.16.8 本节习题

13.91. 对下列题型各选一种优先方法(设点 / 设线 / 参数 / 混合), 并各写一句理由: (1) 椭圆上一点的切线截距之积的最小值; (2) 过定点的弦长的最小值; (3) 以焦点弦为直径的圆过另一焦点.

13.92. 说明: 在椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 中, 为何中点弦问题用点差往往比联立更短, 而焦点弦长用设线往往更短.

13.93. 抛物线 $y^2=2px$ 上 A,B 满足 $\overrightarrow{OA}\perp\overrightarrow{OB}$. 用本节决策树选择方法, 并证明 AB 过定点(可引用 第13.5.4小节).

13.94. 举一个必须讨论斜率不存在的反例, 若只用 $y=kx+m$ 而不讨论竖直情形, 会漏掉哪类几何结论?

13.95. 将以 PQ 为直径的圆过定点 F 这一条件翻译为坐标形式, 并说明它为何通常导致设点与韦达混合路径.

13.96. 开放题: 从本章任选一道你做过的大题, 画出其自由度计数与参量命名过程(不超过一页).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.17 隐圆

隐圆问题是指题目中没有直接给出圆的方程, 但几何条件背后隐藏着圆的轨迹. 识别隐圆并写出其方程, 是解析几何中一类重要的转化技巧. 隐圆的常见定义有五种, 每种对应一类几何条件.

13.17.1 隐圆的五种定义

- **第一定义**: 到定点的距离等于定长的动点轨迹是圆.
- **第二定义**: 到两定点 A, B 距离的平方和为定值 λ ($\lambda > \frac{1}{2}|AB|^2$) 的动点轨迹是圆(以两定点中点为圆心).
- **第三定义**: 到两定点的夹角为 90° 的动点轨迹是圆(以两定点连线为直径).
- **第四定义**: 对角互补的四边形顶点共圆; 弦垂积为定值亦隐含圆.
- **第五定义**: 到两定点距离之比为定值 ($\neq 1$) 的动点轨迹是圆(阿波罗尼斯圆).

识别隐圆的关键, 是把题目的几何条件与上述五种定义对照, 一旦匹配, 便知动点在某个圆上运动, 后续的范围、最值问题便转化为圆上的优化问题.

例题 13.233 (隐圆的第一定义) 已知 $\vec{a}, \vec{b}$ 是单位向量, $\vec{a}\cdot\vec{b}=0$. 若向量 $\vec{c}$ 满足 $|\vec{c}-\vec{a}+\vec{b}|=1$, 则 $|\vec{c}-\vec{b}|$ 的取值范围是(　)

A. $[\sqrt{2}-1, \sqrt{2}+1]$　　C. $[0, 2]$

B. $[1, \sqrt{2}+1]$　　D. $[\sqrt{5}-1, \sqrt{5}+1]$

解 答案: D.

单位向量 $\vec{a}, \vec{b}$ 满足 $\vec{a}\cdot\vec{b}=0$, 即 $\vec{a}\perp\vec{b}$. 作 $\vec{OA}=\vec{a}, \vec{OB}=\vec{b}$, 以射线 OA, OB 分别作为 x, y 轴非负半轴建立平面直角坐标系, 如图.

则 $\vec{a}=(1,0), \vec{b}=(0,1)$, 设 $\vec{c}=(x,y)$, 则 $\vec{c}-\vec{a}+\vec{b}=(x-1, y+1)$. 由 $|\vec{c}-\vec{a}+\vec{b}|=1$ 得

$$(x-1)^2+(y+1)^2=1,$$

令 $\begin{cases} x=1+\cos\theta \\ y=-1+\sin\theta \end{cases}$ $(0\leqslant\theta<2\pi)$, 即 $\vec{c}=(1+\cos\theta, -1+\sin\theta)$.

则

$$|\vec{c}-\vec{b}|=\sqrt{(1+\cos\theta)^2+(-2+\sin\theta)^2}=\sqrt{6-2(2\sin\theta-\cos\theta)}=\sqrt{6-2\sqrt{5}\sin(\theta-\varphi)},$$

其中锐角 φ 满足 $\begin{cases} \sin\varphi=\frac{1}{\sqrt{5}} \\ \cos\varphi=\frac{2}{\sqrt{5}} \end{cases}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

因此, 当 $\sin(\theta-\varphi)=-1$ 时, $|\vec{c}-\vec{b}|_{\max}=\sqrt{6+2\sqrt{5}}=\sqrt{5}+1$; 当 $\sin(\theta-\varphi)=1$ 时, $|\vec{c}-\vec{b}|_{\min}=\sqrt{6-2\sqrt{5}}=\sqrt{5}-1$.

所以 $|\vec{c}-\vec{b}|$ 的取值范围是 $[\sqrt{5}-1,\sqrt{5}+1]$. 故选: D.

例题 13.234 (隐圆的第一定义) 已知单位向量 $\vec{a}$ 与向量 $\vec{b}=(0,2)$ 垂直, 若向量 $\vec{c}$ 满足 $|\vec{a}+\vec{b}+\vec{c}|=1$, 则 $|\vec{c}|$ 的取值范围为( )

A. $[1,\sqrt{5}-1]$　　C. $[\sqrt{5}-1,\sqrt{5}+1]$

B. $[\frac{\sqrt{3}-1}{2},\frac{\sqrt{3}+1}{2}]$　　D. $[\frac{\sqrt{3}+1}{2},3]$

例题 13.235 (隐圆的第二定义) 在平面直角坐标系 xOy 中, $P(2,2),Q(-4,0)$ 为两个定点, 动点 M 在直线 $x=-1$ 上, 动点 N 满足 $NO^2+NQ^2=16$, 则 $\left|\overrightarrow{PM}+\overrightarrow{PN}\right|$ 的最小值为 ____.

解 答案: 5.

设点 $N(x,y)$, 由 $NO^2+NQ^2=16$ 得

$$x^2+y^2+(x+4)^2+y^2=16,$$

即 $x^2+y^2+4x=0$, 即 $(x+2)^2+y^2=4$, 所以 N 在以 OQ 为直径的圆上. 不妨设 $N(2\cos\theta-2,2\sin\theta)$, $M(-1,m)$, 则

$$\overrightarrow{PM}=(-3,m-2),\quad \overrightarrow{PN}=(2\cos\theta-4,2\sin\theta-2),$$

所以

$$\overrightarrow{PM}+\overrightarrow{PN}=(2\cos\theta-7,2\sin\theta+m-4).$$

则

$$\begin{aligned}|\overrightarrow{PM}+\overrightarrow{PN}|^2&=(2\cos\theta-7)^2+(2\sin\theta+m-4)^2\\&=m^2-8m+69+4[(m-4)\sin\theta-7\cos\theta]\\&=(m-4)^2+53+4\sqrt{(m-4)^2+49}\sin(\theta-\varphi),\end{aligned}$$

其中 φ 为辅助角. 令 $\sqrt{(m-4)^2+49}=t$, $\sin(\theta-\varphi)=a$, 则 $t\geq 7$, $-1\leq a\leq 1$. 所以

$$|\overrightarrow{PM}+\overrightarrow{PN}|^2=t^2+4+4at.$$

令 $f(t)=t^2+4+4at=(t+2a)^2+4-4a^2$、$t\geq 7$、$-1\leq a\leq 1$,

例题 13.236 (圆的第三定义) 已知向量 $\vec{a},\vec{b},\vec{c}$ 满足 $|\vec{a}|=4$、$|\vec{b}|=2\sqrt{2}$、$\vec{a}$ 与 $\vec{b}$ 的夹角为 $\frac{\pi}{4}$, $(\vec{a}-\vec{c})\cdot(\vec{b}-\vec{c})=0$, 则 $|\vec{c}|$ 的最大值为__.

知乎@余命数

解 答案: $\sqrt{10}+\sqrt{2}$.

设 $\vec{OA}=\vec{a},\vec{OB}=\vec{b},\vec{OC}=\vec{c}$, 以 OA 所在的直线为 x 轴, O 为坐标原点建立平面直角坐标系. 因为 $|\vec{a}|=4,|\vec{b}|=2\sqrt{2},\vec{a}$ 与 $\vec{b}$ 的夹角为 $\frac{\pi}{4}$, 所以 $A(4,0),B(2,2)$.

设 $C(x,y)$, 即 $\vec{OA}=\vec{a}=(4,0),\vec{OB}=\vec{b}=(2,2),\vec{OC}=\vec{c}=(x,y)$, 则

$$\vec{a}-\vec{c}=(4-x,-y),\quad \vec{b}-\vec{c}=(2-x,2-y).$$

知乎@余命数

由 $(\vec{a}-\vec{c})\cdot(\vec{b}-\vec{c})=0$ 得

$$x^2-6x+8+y^2-2y=0,$$

知乎@余命数

即

$$(x-3)^2+(y-1)^2=2.$$

圆心坐标为 $D(3,1)$, 半径 $r=\sqrt{2}$.$|\vec{c}|$ 表示点 C 到坐标原点的距离, 即圆上的点到坐标原点的距离.

圆心 $D(3,1)$ 到原点的距离为 $d=\sqrt{3^2+1^2}=\sqrt{10}$, 所以

$$|\vec{c}|_{\max}=d+r=\sqrt{10}+\sqrt{2}.$$

知乎@余命数 知乎@余命数

例题 13.237 (椭圆的第三定义) 已知向量 $\vec{a},\vec{b}$ 为单位向量, 且 $\vec{a}\cdot\vec{b}=0$, 若 $\vec{c}$ 满足 $(\vec{a}-\vec{c})\cdot(\vec{b}-\vec{c})=0$, 则 $|\vec{c}|$ 的最大值是__.

解 答案: $\sqrt{2}$.

向量 $\vec{a},\vec{b}$ 为单位向量, 且 $\vec{a}\cdot\vec{b}=0$, 不妨设 $\vec{a}=(1,0),\vec{b}=(0,1)$. 令 $\vec{c}=(x,y)$, 则

$$\vec{a}-\vec{c}=(1-x,-y),\quad \vec{b}-\vec{c}=(-x,1-y).$$

由 $(\vec{a}-\vec{c})\cdot(\vec{b}-\vec{c})=0$ 得

知乎@余命数

$$-x(1-x)-y(1-y)=0,$$

知乎@余命数

即

$$x^2+y^2-x-y=0,$$

它表示以 $(\frac{1}{2},\frac{1}{2})$ 为圆心, $\frac{\sqrt{2}}{2}$ 为半径的圆.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.238 (隐圆的第四定义) 已知 $\vec{a},\vec{b},\vec{e}$ 是平面向量,$|\vec{e}|=1$. 若非零向量 $\vec{a}$ 与 $\vec{e}$ 的夹角为 $\frac{\pi}{3}$,向量 $\vec{b}$ 满足 $\vec{b}^2-4\vec{b}\cdot\vec{e}+3=0$,则 $|\vec{a}-\vec{b}|$ 的最小值是__.

解 答案: $\sqrt{3}-1$. 设 $\vec{a}=(x,y),\vec{e}=(1,0),\vec{b}=(m,n)$, 则由 $(\vec{a},\vec{e})=\frac{\pi}{3}$ 得 $\vec{a}\cdot\vec{e}=|\vec{a}||\vec{e}|\cos\frac{\pi}{3}, x=\frac{1}{2}\sqrt{x^2+y^2}$, 可得 $y=$

例题 13.239 (椭圆的第五定义) 古希腊著名数学家阿波罗尼奥斯发现, 已知平面内两个定点 A,B 及动点 P, 若 $\left|\frac{PB}{PA}\right|=\lambda(\lambda>0$ 且 $\lambda\neq1)$, 则点 P 的轨迹是圆. 后来人们将这个圆以他的名字命名, 称为阿波罗尼奥斯圆(简称"阿氏圆"). 在平面直角坐标系中,已知 $O(0,0),Q(0,\sqrt{2})$, 直线 $l_1:kx-y+k+3=0$, 直线 $l_2:x+ky+3k+1=0$, 若 P 为 l_1,l_2 的交点, 则 $\frac{3}{2}|PO|+\frac{1}{2}|PQ|$ 的最小值为( )

A. $\frac{\sqrt{66}}{2}$　　C. $9-3\sqrt{2}$

B. $6-3\sqrt{2}$　　D. $\sqrt{66}$

解 答案:A. A

例题 13.240 如果圆 $(x-a)^2+(y-a)^2=8$ 上总存在两个点到原点的距离为 $\sqrt{2}$, 则实数 a 的取值范围是( )

A. $(-3,3)$　　C. $(-3,1)$

B. $(-1,1)$　　D. $(-3,-1)\cup(1,3)$

解 答案:D.

问题可转化为圆 $O:(x-a)^2+(y-a)^2=8$ 和圆 $O_1:x^2+y^2=2$ 相交. 两圆圆心距

$$d=\sqrt{(a-0)^2+(a-0)^2}=\sqrt{2}|a|,$$

由 $R-r<|OO_1|<R+r$ 得 $2\sqrt{2}-\sqrt{2}<\sqrt{2}|a|<2\sqrt{2}+\sqrt{2}$, 解得 $1<|a|<3$, 即 $a\in(-3,-1)\cup(1,3)$.

故选:D.

例题 13.241 设 $m\in R$, 过定点 A 的动直线 $x+2+m(y-7)=0$ 和过定点 B 的动直线 $mx-y-m+3=0$ 交于点 $P(x,y)$, 则 $|PA|+|PB|$ 的取值范围是( )

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

A. $[\sqrt{5},2\sqrt{5}]$　　C. $[2\sqrt{5},4\sqrt{5}]$

B. $[\sqrt{10},4\sqrt{5}]$　　D. $[5,5\sqrt{2}]$

解 答案:D.

由题意可知,动直线 $x+2+m(y-7)=0$ 经过定点 $A(-2,7)$,动直线 $mx-y-m+3=0$ 即 $m(x-1)-y+3=0$,经过定点 $B(1,3)$.

因为 $m\neq 0$ 时,动直线 $x+2+m(y-7)=0$ 和动直线 $mx-y-m+3=0$ 的斜率之积为 -1,始终垂直;$m=0$ 时也垂直,所以两直线始终垂直. 又 P 是两条直线的交点,所以 $PA\perp PB$,从而

$$|PA|^2+|PB|^2=|AB|^2=25.$$

知乎@余命数

设 $\angle ABP=\theta$, 则 $|PA|=5\sin\theta$, $|PB|=5\cos\theta$. 由 $|PA|\geqslant 0$ 且 $|PB|\geqslant 0$, 可得 $\theta\in[0,\frac{\pi}{2}]$.

所以

$$|PA|+|PB|=5(\sin\theta+\cos\theta)=5\sqrt{2}\sin(\theta+\frac{\pi}{4}).$$

因为 $\theta\in[0,\frac{\pi}{2}]$,所以 $\theta+\frac{\pi}{4}\in[\frac{\pi}{4},\frac{3\pi}{4}]$,从而 $\sin(\theta+\frac{\pi}{4})\in[\frac{\sqrt{2}}{2},1]$,因此 $5\sqrt{2}\sin(\theta+\frac{\pi}{4})\in[5,5\sqrt{2}]$.

故选 D.

知乎@余命数 知乎@余命数

例题 13.242 设 $m\in\mathbb{R}$, 过定点 A 的动直线 $x+my=0$ 和过定点 B 的动直线 $mx-y-m+3=0$ 交于点 $P(x,y)$,则 $|PA|^2+|PB|^2$ 的值为(　)

A. 5　　C. $\frac{\sqrt{10}}{2}$

B. 10　　D. $\sqrt{17}$

解 答案:B.

由题意,动直线 $x+my=0$ 经过定点 $(0,0)$,则 $A(0,0)$,动直线 $mx-y-m+3=0$ 变形得 $m(x-1)+(3-y)=0$,则 $B(1,3)$.

由 $\begin{cases}x+my=0\\ mx-y-m+3=0\end{cases}$ 得 $P\left(\frac{m^2-3m}{m^2+1},\frac{3-m}{m^2+1}\right)$,所以

知乎@余命数

$$|PA|^2+|PB|^2=\left(\frac{m^2-3m}{m^2+1}\right)^2+\left(\frac{3-m}{m^2+1}\right)^2+\left(\frac{m^2-3m}{m^2+1}-1\right)^2+\left(\frac{3-m}{m^2+1}-3\right)^2$$

知乎@余命数

$$=\frac{(m^2-3m)^2+(3-m)^2+(3m+1)^2+(3m^2+m)^2}{(m^2+1)^2}$$
$$=\frac{m^4-6m^3+9m^2+9-6m+m^2+9m^2+6m+1+9m^4+6m^3+m^2}{(m^2+1)^2}$$
$$=\frac{10m^4+20m^2+10}{(m^2+1)^2}=10.$$

故选：B.

例题 13.243 （2024·江苏盐城·三模）已知 A,B,C,D 四点共面，$BC=2$，$AB^2+AC^2=20$，且 D 在射线 CA 上满足 $CD=3CA$，则 $|\overrightarrow{BD}|$ 的最大值为__.

解 答案：10.

设 $AC=m$，由题意可得 $DC=3m$，$AB=\sqrt{20-m^2}$，则

$$\cos C=\frac{AC^2+BC^2-AB^2}{2AC\times BC}=\frac{m^2-8}{2m}.$$

$\triangle ABC$ 构成三角形，则 $\begin{cases}m+2>\sqrt{20-m^2}\\|m-2|<\sqrt{20-m^2}\end{cases}$，解得 $2<m<4$.

由余弦定理，

$$BD=\sqrt{BC^2+CD^2-2BC\cdot CD\cdot\cos C}=\sqrt{4+9m^2-2\times2\times3m\times\frac{m^2-8}{2m}}=\sqrt{52+3m^2}.$$

当 $m=4$ 时，$|\overrightarrow{BD}|$ 取得最大值为 10.

例题 13.244 已知圆 $C:(x+1)^2+(y-2)^2=1$，点 $A(-1,0)$，$B(1,0)$，设 P 是圆 C 上的动点，令 $d=|PA|^2+|PB|^2$，则 d 的最小值为__.

解 答案：$14-4\sqrt{5}$.

设 $P(x_0,y_0)$，则

$$|PA|^2=(x_0+1)^2+y_0^2,$$
$$|PB|^2=(x_0-1)^2+y_0^2,$$
$$|PA|^2+|PB|^2=(x_0-1)^2+y_0^2+(x_0+1)^2+y_0^2$$
$$=2x_0^2+2y_0^2+2=2(x_0^2+y_0^2)+2.$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

当 $|OP|$ 取得最小值时，$|PA|^2+|PB|^2$ 取得最小值. 由圆 $C:(x+1)^2+(y-2)^2=1$，则圆心 $C(-1,2)$，半径 $r=1$，易知

$$|OP|_{\min}=|OC|-r=\sqrt{1+4}-1=\sqrt{5}-1,$$

则 $d_{\min}=2(\sqrt{5}-1)^2+2=14-4\sqrt{5}$.

故答案为：$14-4\sqrt{5}$.

例题 13.245 已知圆 $C:(x+1)^2+(y-2)^2=4$，点 $A(-2,0)$，$B(2,0)$，设 P 是圆 C 上的动点，令 $d=|PA|^2+|PB|^2$，则 d 的最小值为__.

解 答案：$26-8\sqrt{5}$.

由 C 知 $C(-1,2)$，$r=2$. 设 $P(x_0,y_0)$、$|PA|=\sqrt{(x_0+2)^2+y_0^2}$、$|PB|=\sqrt{(x_0-2)^2+y_0^2}$，所以

$$d=|PA|^2+|PB|^2=(x_0+2)^2+y_0^2+(x_0-2)^2+y_0^2=2(x_0^2+y_0^2)+8.$$

因为 $|OP|=\sqrt{x_0^2+y_0^2}$，所以当 $|OP|$ 取得最小值时，d 取得最小值. 由 $|OP|$ 的最小值为

$$|OC|-r=\sqrt{(-1)^2+2^2}-2=\sqrt{5}-2,$$

所以 d 的最小值为 $2(\sqrt{5}-2)^2+8=26-8\sqrt{5}$.

例题 13.246 正方形 $ABCD$ 与点 P 在同一平面内，已知该正方形的边长为 1，且 $|PA|^2+|PB|^2=|PC|^2$，则 $|PD|$ 的取值范围为__.

解 答案：$[2-\sqrt{2},2+\sqrt{2}]$.

如图，以 A 为坐标原点，建立平面直角坐标系，则 $A(0,0),B(1,0),C(1,1),D(0,1)$. 设点 $P(x,y)$，则由 $|PA|^2+|PB|^2=|PC|^2$，得

$$x^2+y^2+(x-1)^2+y^2=(x-1)^2+(y-1)^2,$$

整理得

$$x^2+(y+1)^2=2,$$

即点 P 的轨迹是以点 $M(0,-1)$ 为圆心，$\sqrt{2}$ 为半径的圆.

圆心 M 到点 D 的距离为 $|DM|=2$，所以 $|PD|_{\min}=2-\sqrt{2}$，$|PD|_{\max}=2+\sqrt{2}$，所以 $|PD|$ 的取值范围是 $[2-\sqrt{2},2+\sqrt{2}]$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 13.247 已知点 $A(-m,0),B(m,0)$, 若圆 $C:x^2+y^2-6x-8y+24=0$ 上存在点 P, 使得 $PA\perp PB$, 实数 m 的最大值是(　)

A. 4　　C. 6

B. 5　　D. 7

解 答案:C.

圆 $C:x^2+y^2-6x-8y+24=0$ 即为 $(x-3)^2+(y-4)^2=1$, 其圆心为 $(3,4)$, 半径为 1. 设 AB 的中点为 M, 因为点 $A(-m,0),B(m,0)$, 所以 $M(0,0)$, 以 AB 为直径的圆的方程为 $x^2+y^2=m^2$.

$|CM|=\sqrt{3^2+4^2}=5$. 若圆 C 上存在点 P 使得 $PA\perp PB$, 则圆 C 与圆 M 有公共点, 即

$$||m|-1|\leqslant 5\leqslant |m|+1,$$

解得 $4\leqslant |m|\leqslant 6$, 所以实数 m 的最大值是 6.

故选:C.

例题 13.248 已知圆 $C:(x-1)^2+(y+3)^2=10$ 和点 $M(5,t)$, 若圆 C 上存在两点 A,B 使得 $MA\perp MB$, 则实数 t 的取值范围是__.

解 答案:$-5\leqslant t\leqslant -1$.

圆 $C:(x-1)^2+(y+3)^2=10$, 则半径为 $\sqrt{10}$, $C(1,-3)$. 如上图, 对于直线 $x=5$ 上任意一点 $M(5,t)$, 当 AM,BM 均为圆的切线时 $\angle AMB$ 最大. 由题意, $MA\perp MB$ 即 $\angle AMB=90^\circ$ 时, M 为满足条件的临界点, 此时

$$\left|\frac{AC}{|CM|}\right|=\sin\angle AMC\geqslant\frac{\sqrt{2}}{2}.$$

当 M 在临界点之间移动时, 有 $\left|\frac{AC}{|CM|}\right|\geqslant\frac{\sqrt{2}}{2}$, 即

$$\frac{\sqrt{10}}{\sqrt{(5-1)^2+(t+3)^2}}\geqslant\frac{\sqrt{2}}{2},$$

即 $(t+3)^2\leqslant 4$, 解得 $-5\leqslant t\leqslant -1$.

故答案为:$-5\leqslant t\leqslant -1$.

例题 13.249 设向量 $\vec{a},\vec{b},\vec{c}$ 满足 $|\vec{a}|=|\vec{b}|=2$、$\vec{a}\cdot\vec{b}=-2$、$(\vec{a}-\vec{c},\vec{b}-\vec{c})=60^\circ$. 则 $|\vec{c}|$ 的最大值等于(　)

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

A. 4　　C. $\sqrt{2}$

B. 2　　D. 1

解 答案：A.

因为 $|\vec{a}|=|\vec{b}|=2,\vec{a}\cdot\vec{b}=-2$，所以

$$\cos\langle\vec{a},\vec{b}\rangle=\frac{\vec{a}\cdot\vec{b}}{|\vec{a}||\vec{b}|}=-\frac{1}{2},$$

即 $\langle\vec{a},\vec{b}\rangle=120^\circ$.

如图所示，设 $\vec{OA}=\vec{a}$、$\vec{OB}=\vec{b}$、$\vec{OC}=\vec{c}$，则 $\vec{CA}=\vec{a}-\vec{c}$、$\vec{CB}=\vec{b}-\vec{c}$、$\angle AOB=120^\circ$. 所以 $\angle ACB=60^\circ$、$\angle AOB+\angle ACB=180^\circ$、$A,O,B,C$ 四点共圆. 不妨设为圆 M.

因为 $\vec{AB}=\vec{b}-\vec{a}$，所以

$$|\vec{AB}|^2=|\vec{a}|^2-2\vec{a}\cdot\vec{b}+|\vec{b}|^2=12,$$

即 $|\vec{AB}|=2\sqrt{3}$. 由正弦定理可得 $\triangle AOB$ 的外接圆即圆 M 的直径为

$$2R=\frac{|\vec{AB}|}{\sin\angle AOB}=4.$$

所以当 $|\vec{OC}|$ 为圆 M 的直径时，$|\vec{c}|$ 取得最大值 4.

故选：A.

例题 13.250　(2024·天津·一模) 如图，梯形 $ABCD$ 中，$AB\parallel CD$、$AB=2$、$CD=4$、$BC=AD=\sqrt{5}$、E 和 F 分别为 AD 与 BC 的中点，对于常数 λ，在梯形 $ABCD$ 的四条边上恰好有 8 个不同的点 P，使得 $\vec{PE}\cdot\vec{PF}=\lambda$ 成立，则实数 λ 的取值范围是

A. $\left(-\frac{5}{4},-\frac{9}{20}\right)$　　C. $\left(-\frac{1}{4},\frac{11}{4}\right)$

B. $\left(-\frac{5}{4},-\frac{11}{4}\right)$　　D. $\left(-\frac{9}{20},-\frac{1}{4}\right)$

解 答案：D. 以 CD 的中点为坐标原点，CD 所在的直线为 x 轴建立平面直角坐标系，则 $A(-1,2)$、$B(1,2)$、$C(2,0)$、$D(-2,0)$、$E\left(-\frac{3}{2},1\right)$、$F\left(\frac{3}{2},1\right)$.

例题 13.251　(2024·广东广州·一模) 在平面四边形 $ABCD$ 中，连接对角线 BD，已知 $CD=9$、$BD=16$、$\angle BDC=90^\circ$、$\sin A=\frac{4}{5}$，则对角线 AC 的最大值为__.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

解 答案:27.

画出图像如下图所示. 由于 $\sin A=\frac{4}{5}$, $BD=16$ 为定值, 故 A 在以 BD 为弦的圆上运动. 由正弦定理得

$$2R=\frac{16}{\frac{4}{5}}=20,\quad R=10,$$

故圆心的坐标为 $(8,-6)$.AC 的最大值即为 CA' 的值, 也是 $CO+R$ 的值. 由两点间的距离公式有

$$CO+R=\sqrt{8^2+15^2}+10=27.$$

例题 13.252 (2024·江西赣州·模拟预测) 阿波罗尼斯是古希腊著名数学家, 与欧几里得、阿基米德并称为亚历山大时期数学三巨匠, 他对圆锥曲线有深刻而系统的研究, 阿波罗尼斯圆是他的研究成果之一, 指的是已知动点 M 与两定点 A,B 的距离之比为 $\lambda(\lambda>0,\lambda\neq1)$, 那么点 M 的轨迹就是阿波罗尼斯圆, 简称阿氏圆. 已知在平面直角坐标系中, 圆 $O:x^2+y^2=1$ 上点 $A\left(-\frac{1}{2},0\right)$ 和点 $B\left(0,\frac{1}{2}\right)$, M 为圆 O 上的动点, 则 $2|MA|-|MB|$ 的最大值为 (　)

A. $\frac{5}{2}$　　C. $\frac{3}{2}$

B. $\frac{\sqrt{17}}{2}$　　D. $\frac{\sqrt{2}}{2}$

解 答案:B. 设 $M(x,y)$, 令 $2|MA|=|MC|$, 则 $\frac{|MA|}{|MC|}=\frac{1}{2}$, 由题知圆 $x^2+y^2=1$ 是关于点 A,C 的阿波罗尼斯圆, 且 $\lambda=\frac{1}{2}$,

13.17.2 本节小结

- 隐圆的识别:将几何条件与五种定义对照, 找出隐藏的圆心和半径.
- 第一定义看等距, 第二定义看平方和, 第三定义看直角, 第四定义看对角互补, 第五定义看距离比.
- 识别隐圆后, 范围与最值问题转化为圆上的优化问题.

13.17.3 本节习题

13.97. 动点 P 满足 $|PA|^2+|PB|^2=20$, 其中 $A(-1,0)$, $B(1,0)$. 求 P 的轨迹方程, 并指出圆心与半径.

13.98. 动点 P 满足 $\angle APB=90^\circ$, 其中 $A(0,0)$, $B(4,0)$. 求 P 的轨迹方程.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.99. 动点 P 满足 $|PA|=2|PB|$, 其中 $A(0,0)$, $B(3,0)$. 求 P 的轨迹方程(阿波罗尼斯圆).

13.18 直线与圆的综合应用

第13.1节 建立了直线与圆的基本理论, 本节讨论其综合应用. 这类问题的特征是把距离、切线、弦长等基本工具与各种距离定义(切比雪夫距离、曼哈顿距离、闵氏距离)、包络线、阿波罗尼斯圆、反演点等进阶概念结合, 考察在复杂情境下灵活调用基本工具的能力.

知乎@余命数

13.18.1 距离的创新定义

除欧氏距离外, 切比雪夫距离 $d_C(P,Q)=\max(|x_1-x_2|,|y_1-y_2|)$、曼哈顿距离 $d_M(P,Q)=|x_1-x_2|+|y_1-y_2|$、闵氏距离 $d_p(P,Q)=(|x_1-x_2|^p+|y_1-y_2|^p)^{1/p}$ 等在特定问题中会出现. 识别距离定义后, 把圆上的动点代入距离公式求最值.

13.18.2 圆的包络线

一族圆的包络线是与该族每个圆都相切的曲线. 常见的有: 圆心在定直线上、半径随圆心变化的圆族的包络线是两条平行线; 圆心在定圆上、半径等于到定圆的圆心距的圆族的包络线是一个点.

13.18.3 阿波罗尼斯圆与反演

到两定点距离之比为定值 $\lambda\neq 1$ 的动点轨迹是圆, 称为阿波罗尼斯圆. 反演变换把圆映射为圆或直线, 是处理圆系问题的有力工具.

例题 13.253 (距离的创新定义) 数学家华罗庚曾说, “数缺形时少直观, 形少数时难入微”. 事实上, 很多代数问题可以转化为几何问题加以解决, 例如, 与 $\sqrt{(x-a)^2+(y-b)^2}$ 相关的代数问题, 可以转化为点 $A(x,y)$ 与点 $B(a,b)$ 之间距离的几何问题, 结合上述观点, 求求得方程 $\sqrt{x^2+4x+5}+\sqrt{x^2-4x+5}=6$ 的解是(　)

A. $\frac{3\sqrt{62}}{4}$　　C. $\frac{6\sqrt{5}}{5}$

B. $\pm\frac{3\sqrt{62}}{4}$　　D. $\pm\frac{6\sqrt{5}}{5}$

解 答案: D.

因为 $\sqrt{x^2+4x+5}=\sqrt{(x+2)^2+1}=\sqrt{[x-(-2)]^2+(1-0)^2}$, 所以 $\sqrt{x^2+4x+5}$ 可以转化为 $M(x,1)$ 到 $N(-2,0)$ 的距离. 同理, $\sqrt{x^2-4x+5}$ 可以转化为 $M(x,1)$ 到 $P(2,0)$ 的距离.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由 $\sqrt{x^2+4x+5}+\sqrt{x^2-4x+5}=6$, $M(x,1)$ 到两定点 $N(-2,0)$ 和 $P(2,0)$ 的距离之和为 6, 即 $M(x,1)$ 在以 $N(-2,0)$ 和 $P(2,0)$ 为焦点的椭圆上. 设椭圆的标准方程为

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \quad (a>b>0),$$

则 $2a=6$, 即 $a=3$. 又 $a^2-b^2=4$, 所以 $b^2=5$, 椭圆方程为

$$\frac{x^2}{9}+\frac{y^2}{5}=1.$$

知乎@余命数

由 $y=1$, 得 $\frac{x^2}{9}+\frac{1}{5}=1$, 解得 $x=\pm\frac{6\sqrt{5}}{5}$.

故选: D.

例题 13.254 (距离的创新定义) 人脸识别中检测样本之间相似度主要应用距离的测试, 常用测量距离的方式有曼哈顿距离和余弦距离. 若三维空间有两个点 $A(x_1,y_1)$, $B(x_2,y_2)$, 则曼哈顿距离为: $d(A,B)=|x_1-x_2|+|y_1-y_2|$, 余弦相似度为: $\cos(A,B)=\frac{x_1}{\sqrt{x_1^2+y_1^2}}\times\frac{x_2}{\sqrt{x_2^2+y_2^2}}+\frac{y_1}{\sqrt{x_1^2+y_1^2}}\times\frac{y_2}{\sqrt{x_2^2+y_2^2}}$, 余弦距离为 $1-\cos(A,B)$. 若 $A(-1,2)$, $B\left(\frac{3}{5},\frac{4}{5}\right)$, 则 A,B 之间的余距离为( )

例题 13.255 (切比雪夫距离) 在平面直角坐标系中, 定义 $d(A,B)=\max\{|x_1-x_2|,|y_1-y_2|\}$ 为两点 $A(x_1,y_1)$、$B(x_2,y_2)$ 的“切比雪夫距离”, 又设点 P 及 l 上任意一点 Q, 称 $d(P,Q)$ 的最小值为点 P 到直线 l 的“切比雪夫距离”记作 $d(P,l)$; 给出下列四个命题: (1) 对任意三点 A,B,C, 都有 $d(C,A)+d(C,B)\geqslant d(A,B)$; (2) 已知点 $P(3,1)$ 和直线 $l:2x-y-1=0$, 则 $d(P,l)=\frac{4}{3}$; (3) 到原点的切比雪夫距离等于 1 的点的轨迹是正方形; 其中真命题的是( )

A. (1)(2)　　C. (1)(3)

B. (2)(3)　　D. (1)(2)(3)

解 答案: D.

(1) 对任意三点 A,B,C, 若它们共线, 设 $A(x_1,y_1),B(x_2,y_2),C(x_3,y_3)$, 如图, 结合三角形的相似可得 $d(C,A),d(C,B),d(A,B)$ 为 AN,CM,AK, 或 CN,BM,BK, 则 $d(C,A)+d(C,B)\geqslant d(A,B)$; 若 B,C 或 A,C 对调, 可得 $d(C,A)+d(C,B)\geqslant d(A,B)$; 若 A,B,C 不共线, 且三角形 C 为锐角或钝角, 如图, 由矩形 $CMNK$ 或矩形 $BMNK$, $d(C,A)+d(C,B)\geqslant d(A,B)$. 则对任意的三点 A,B,C, 都有 $d(C,A)+d(C,B)\geqslant d(A,B)$, 故 (1) 正确.

(2) 设点 Q 是直线 $y=2x-1$ 上一点, 且 $Q(x,2x-1)$, 可得 $d(P,Q)=\max\{|x-3|,|2-2x|\}$. 由 $|x-3|\geqslant|2-2x|$, 解得 $-1\leqslant x\leqslant\frac{5}{3}$, 即有 $d(P,Q)=|x-3|$, 当

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$x=\frac{5}{3}$ 时，取得最小值 $\frac{4}{3}$；由 $|x-3|\leqslant|2-2x|$，解得 $x>\frac{5}{3}$ 或 $x<-1$，即有 $d(P,Q)=|2x-2|$，$d(P,Q)$ 的范围是 $\left(\frac{4}{3},+\infty\right)$，无最值. 综上可得，$P,Q$ 两点的切比雪夫距离的最小值为 $\frac{4}{3}$，故 (2) 正确.

(3) 由题，到原点 O 的切比雪夫距离为 1 的点 $P(x,y)$ 满足 $d(O,P)=\max\{|x|,|y|\}=1$，即 $|x|\geqslant|y|$ 或 $|x|\leqslant|y|$，显然点 P 的轨迹为正方形，故 (3) 正确.

知乎@余命数

故选：D.

例题 13.256 (曼哈顿距离) 在平面直角坐标系中，定义 $d(A,B)=|x_1-x_2|+|y_1-y_2|$ 为两点 $A(x_1,y_1)$、$B(x_2,y_2)$ 的“曼哈顿距离”，又设点 P 及直线 l 上任意一点 Q，称 $d(P,Q)$ 的最小值为点 P 到直线 l 的“切比雪夫距离”，记作 $d(P,l)$，给出下列三个命题：(1) 对任意三点 A、B、C，都有 $d(C,A)+d(C,B)\geqslant d(A,B)$；

例题 13.257 (曼哈顿距离、折线距离、直角距离问题) （多选题）曼哈顿距离是十九世纪的赫尔曼·闵可夫斯基所创词汇，用以标明两个点在标准坐标系上的绝对轴距总和，其定义如下：在直角坐标平面上任意两点 $A(x_1,y_1),B(x_2,y_2)$ 的曼哈顿距离 $d(A,B)=|x_1-x_2|+|y_1-y_2|$，则下列结论正确的是(　)

A. 若点 $P(2,4),Q(-2,1)$，则 $d(P,Q)=7$

B. 若点 $M(-1,0),N(1,0)$，则在 x 轴上存在点 P，使得 $d(P,M)+d(P,N)=1$

C. 若点 $M(2,1)$，点 P 在直线 $x-2y+6=0$ 上，则 $d(P,M)$ 的最小值是 3

例题 13.258 (闵氏距离问题) （2024·全国·模拟预测）闵氏距离 (Minkowski distance) 是衡量数值点之间距离的一种非常常见的方法，设点 A、B 坐标分别为 (x_1,y_1)、(x_2,y_2)，则闵氏距离 $D_p(A,B)=(|x_1-x_2|^p+|y_1-y_2|^p)^{1/p}(p\in N^*)$. 若点 A、B 分别在 $y=e^x$ 和 $y=x-1$ 的图像上，则 $D_p(A,B)$ 的最小值为(　)

A. $2^{1/p}$　　C. $e^{1/p}$

B. 2^p　　D. e^p

解 答案：A.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

设 $A(x_1,e^{x_1}),B(x_2,x_2-1)$，因为点 A、B 分别在函数 $y=e^x$ 和 $y=x-1$ 的图象上，所以

$$D_p(A,B)=(|x_1-x_2|^p+|e^{x_1}-x_2+1|^p)^{\frac{1}{p}}\geqslant 2^{\frac{1-p}{p}}\left|(x_1-x_2)-(e^{x_1}-x_2+1)\right|=2^{\frac{1-p}{p}}\left|x_1-e^{x_1}-1\right|,$$

当且仅当 $|x_1-x_2|=|e^{x_1}-x_2+1|$ 时等号成立.

设 $g(x)=|x-e^x-1|,h(x)=x-e^x-1$，则 $h'(x)=1-e^x$. 令 $h'(x)\geqslant 0\Rightarrow x\leqslant 0$，$h'(x)<0\Rightarrow x>0$，所以函数 $h(x)$ 在 $(-\infty,0)$ 上单调递增，在 $(0,+\infty)$ 上单调递减，所以 $h(x)_{\max}=h(0)=-2$，即 $h(x)\leqslant -2$，所以 $g(x)=|h(x)|=-h(x)\geqslant 2$，即

$$D_p(A,B)\geqslant 2^{\frac{1}{p}},$$

所以 $D_p(A,B)$ 的最小值为 $2^{\frac{1}{p}}$.

故选：A.

知乎@余命数

例题 13.259 (闵氏距离问题) （2024·高三·安徽阜阳·期末） 闵可夫斯基距离又称为闵氏距离，是两组数据间距离的定义. 设两组数据分别为 $A=(a_1,a_2,\cdots,a_n)$ 和 $B=(b_1,b_2,\cdots,b_n)$，这两组数据间的闵氏距离定义为 $d_{AB}(q)=[\sum_{k=1}^{n}|a_k-b_k|^q]^{\frac{1}{q}}$，其中 q 表示阶数. 现有下列四个命题：(1) 若 $A=(1,2,3,4),B=(0,3,4,5)$，则 $d_{AB}(1)=4$；(2) 若 $A=(a,a+1),B=(b-1,b)$，其中 $a,b\in R$，则 $d_{AB}(1)=d_{AB}(2)$；(3) 若 $A=(a,b),B=(c,d)$，其中 $a,b,c,d\in R$，则 $d_{AB}(1)\geqslant d_{AB}(2)$；(4) 若 $A=(a,a^2),B=(b,b-1)$，其中 $a,b\in R$，则 $d_{AB}(2)$ 的最小值为 $\frac{3\sqrt{2}}{8}$. 其中所有真命题的个数是（　）

例题 13.260 (圆的包络线问题) （多选题）设有一组圆 $C_k:(x-k+1)^2+(y-3k)^2=2k^4(k\in\mathbb{N}^*)$，下列四个命题中真命题的是

A. 存在一条定直线与所有的圆均相切

B. 存在一条定直线与所有的圆均相交

C. 存在一条定直线与所有的圆均不相交

D. 所有的圆均不经过原点

解 答案：BD.

圆心为 $C_k(k-1,3k)$，半径为 $r_k=\sqrt{2}k^2$. $C_1(0,3)$、$r_1=\sqrt{2}$、$C_2(1,6)$、$r_2=4\sqrt{2}$，

$$|C_1C_2|=\sqrt{1^2+3^2}=\sqrt{10}<4\sqrt{2}-\sqrt{2}=3\sqrt{2},$$

圆 C_1 与圆 C_2 内含，因此不可能有直线与这两个圆都相切，从而 A 错误.

圆心在直线 $y=3(x+1)$ 上，此直线与所有圆都相交，B 正确.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

若 $k\to\infty$, 则所有直线都与圆相交, C 错.

将 $(0,0)$ 代入圆方程得 $(k-1)^2+9k^2=2k^4$, 即 $10k^2-2k+1=2k^4$, 左边奇数右边偶数, 无整数解, 故原点不在任一圆上, D 正确.

故选: BD.

知乎@余命数

例题 13.261 (圆的包络线问题) (多选题)设有一组圆 $C_k:(x-1)^2+(y-k)^2=k^4(k\in\mathbb{N}^*)$, 下列四个命题正确的是

A. 存在 k_0 使圆与 x 轴相切

B. 存在一条直线与所有的圆均相交

C. 存在一条直线与所有的圆均不相交

D. 所有的圆均不经过原点

知乎@余命数

解 答案: ABD.

圆心为 $(1,k)$, 半径为 k^2.

选项 A: 当 $k=k^2$ 即 $k=1$ 时, 圆为 $(x-1)^2+(y-1)^2=1$, 与 x 轴相切, 正确.

选项 B: 直线 $x=1$ 过所有圆心, 故与所有圆相交, 正确.

选项 C: 圆 k 与圆 $k+1$ 圆心距 $d=1$, 半径差 $R-r=2k+1>1, k\to\infty$ 时所有直线均相交, 故错误.

选项 D: 将 $(0,0)$ 代入得 $1+k^2=k^4$, 无 $k\in\mathbb{N}^*$ 满足, 故所有圆不过原点, 正确.

故选: ABD.

知乎@余命数

例题 13.262 (阿波罗尼斯圆问题、反演点问题、阿波罗尼斯球问题) 阿波罗尼斯是古希腊著名数学家, 与欧几里得、阿基米德被称为亚历山大时期数学三巨匠, 阿波罗尼斯发现, 平面内到两个定点 A,B 的距离之比为定值 $\lambda(\lambda>0$, 且 $\lambda\neq1)$ 的点的轨迹是圆, 此圆被称为"阿波罗尼斯圆". 在平面直角坐标系 xOy 中, $A(-2,0),B(4,0)$, 点 P 满足 $\frac{|PA|}{|PB|}=\frac{1}{2}$. 设点 P 的轨迹为曲线 C, 则下列说法正确的是( )

A. C 的方程为 $(x+4)^2+y^2=16$

B. 点 A,B 都在曲线 C 内部

C. 当 A,B,P 三点不共线时, 则 $\angle APO=\angle BPO$

D. 若 $D(2,2)$, 则 $|PB|+2|PD|$ 的最小值为 $4\sqrt{5}$

知乎@余命数

解 答案: ACD.

设 $P(x,y)$(P 不与 A,B 重合). 由 $A(-2,0),B(4,0)$, 有 $|PA|=\sqrt{(x+2)^2+y^2}$, $|PB|=\sqrt{(x-4)^2+y^2}$. $\frac{|PA|}{|PB|}=\frac{1}{2}$, 即

$$\frac{\sqrt{(x+2)^2+y^2}}{\sqrt{(x-4)^2+y^2}}=\frac{1}{2},$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

化简得 $(x+4)^2+y^2=16$, 所以点 P 的轨迹曲线 C 是以 $C(-4,0)$ 为圆心, 半径 $r=4$ 的圆, 如图所示.

对于 A 选项, 由曲线 C 的方程为 $(x+4)^2+y^2=16$, 选项 A 正确.

对于 B 选项, 由 $BC=8$, 点 B 在曲线 C 外, 选项 B 错误.

对于 C 选项, 由 $|OA|=2$, $|OB|=4$, 有 $\frac{|OA|}{|OB|}=\frac{1}{2}=\frac{|PA|}{|PB|}$, 则当 A、B、P 三点不共线时, 由三角形内角平分线定理知, PO 是 $\triangle APB$ 内角 $\angle APB$ 的角平分线, 所以 $\angle APO=\angle BPO$, 选项 C 正确.

对于 D 选项, 由 $\frac{|PA|}{|PB|}=\frac{1}{2}$, 得 $|PB|=2|PA|$, 则

$$|PB|+2|PD|=2|PA|+2|PD|=2(|PA|+|PD|)\geq 2|AD|=2\times\sqrt{(-2-2)^2+(0-2)^2}=4\sqrt{5},$$

当且仅当 P 在线段 AD 上时, 等号成立, 则 $|PB|+2|PD|$ 的最小值为 $4\sqrt{5}$, 选项 D 正确.

例题 13.263 (圆中的垂直问题) 已知直线 $l_1: x-3y+1=0$, 直线 l_2 过点 $(1,0)$ 且与直线 l_1 相互垂直, 圆 $C: x^2+y^2-4x-2y-3=0$, 若直线 l_2 与圆 C 交于 M, N 两点, 则 $|MN|=$

解 答案: $\frac{8\sqrt{10}}{5}$.

由直线 $l_1: x-3y+1=0$, 可得斜率 $k_1=\frac{1}{3}$. 因为 $l_1\perp l_2$ 且直线 l_2 过点 $(1,0)$, 所以直线 l_2 的斜率为 $k_2=-3$, l_2 的方程为 $3x+y-3=0$.

又由圆 $C: x^2+y^2-4x-2y-3=0$, 即 $C: (x-2)^2+(y-1)^2=8$, 可得圆 C 的圆心坐标为 $C(2,1)$, 半径为 $r=2\sqrt{2}$.

则圆心 C 到直线 l_2 的距离为

$$d=\frac{|2\times 3+1-3|}{\sqrt{3^2+1^2}}=\frac{4}{\sqrt{10}},$$

所以弦长

$$|MN|=2\sqrt{r^2-d^2}=2\sqrt{(2\sqrt{2})^2-\left(\frac{4}{\sqrt{10}}\right)^2}=\frac{8\sqrt{10}}{5}.$$

故答案为: $\frac{8\sqrt{10}}{5}$.

例题 13.264 (圆中的垂直问题) 已知 AC, BD 为圆 $O: x^2+y^2=9$ 的两条相互垂直的弦, 垂足为 $M(2,2)$, 则 $|AC|\cdot|BD|$ 的最大值为

解 答案: 20.

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设圆心 $O(0,0)$ 到 AC, BD 的距离分别为 m, n. 因为 AC, BD 相互垂直, 所以

$$m^2+n^2=|OM|^2=8.$$

由垂径定理得 $|AC|=2\sqrt{9-m^2}, |BD|=2\sqrt{9-n^2}$, 则

$$|AC|\cdot|BD|=4\sqrt{9-m^2}\cdot\sqrt{9-n^2}=4\sqrt{81-9(m^2+n^2)+m^2n^2}=4\sqrt{9+m^2n^2}.$$

由 $m^2+n^2=8\geq 2mn$, 得 $mn\leq 4$, 当且仅当 $m=n=2$ 时等号成立, 故 $(|AC|\cdot|BD|)_{\max}=20$.

故答案为: 20.

例题 13.265 (圆的存在性问题) 已知圆 $C:(x-1)^2+y^2=80$, 点 P 在直线 $l:y=kx+7(k\in\mathbb{R})$ 上. 若存在过点 P 的直线与圆 C 相交于 A, B 两点, 且 $|AB|=16, \overrightarrow{AP}=\frac{5}{11}\overrightarrow{PB}$, 则 k 的取值范围是__.

解 答案: $(-\infty,-\frac{3}{4}]\cup[\frac{4}{3},+\infty)$.

圆 $C:(x-1)^2+y^2=80$ 圆心 $C(1,0)$, 半径为 $4\sqrt{5}$. 设弦 AB 中点为 Q, 连接 CP, CQ, 由 $|AB|=16, \overrightarrow{AP}=\frac{5}{11}\overrightarrow{PB}$, 可得点 P 在弦 AB 上, 且 $|AP|=5$、$|BP|=11$、$|PQ|=3$. 又圆心 C 到弦 AB 所在直线的距离为

$$|CQ|=\sqrt{80-8^2}=4,$$

则

$$|CP|=\sqrt{|CQ|^2+|PQ|^2}=\sqrt{4^2+3^2}=5.$$

则点 P 在以 C 为圆心半径为 5 的圆上运动, 又点 P 在直线 $l:y=kx+7\,(k\in\mathbb{R})$ 上, 则直线 l 与以 C 为圆心半径为 5 的圆有公共点, 则

$$\frac{|k+7|}{\sqrt{1+k^2}}\leq 5,$$

解之得 $k\geq\frac{4}{3}$ 或 $k\leq-\frac{3}{4}$, 所以 k 的取值范围是 $(-\infty,-\frac{3}{4}]\cup[\frac{4}{3},+\infty)$.

例题 13.266 (圆的存在性问题) 已知圆 $C:(x-1)^2+(y-4)^2=r^2(r>0)$, $A(-3,0), B(-1,0)$, 若 C 上存在点 P, 使得 $\angle APB=90^\circ$, 则 r 的取值范围为__.

解 答案: $[4,6]$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因为点 $A(-3,0),B(-1,0)$，而点 P 满足 $\angle APB=90^\circ$，则点 P 的轨迹是以线段 AB 为直径的圆 M（除点 A，B 外），圆 $M:(x+2)^2+y^2=1\,(y\neq 0)$，半径 $r_1=1$. 又点 P 在圆 $C:(x-1)^2+(y-4)^2=r^2\,(r>0)$ 上，圆 C 的圆心 $C(1,4)$，半径为 r，

$$|CM|=\sqrt{(-2-1)^2+4^2}=5.$$

依题意，圆 M 与圆 C 有公共点，因此

$$|r-r_1|\leq|MC|\leq r+r_1,$$

即 $|r-1|\leq 5\leq r+1$，解得 $4\leq r\leq 6$.

例题 13.267 费马点是指三角形内部到三角形三个顶点距离之和最小的点，当三角形三个内角均小于 120° 时，费马点与三个顶点连线正好三等分费马点所在的周角，即该点所对三角形三边的张角相等，均为 120°. 根据以上性质，已知 $A(-2,0)$、$B(2,0)$、$C(0,4)$、P 为 $\triangle ABC$ 内一点，记 $f(P)=|PA|+|PB|+|PC|$，则 $f(P)$ 的最小值为（　）

A. $2\sqrt{3}$　　C. $4+\sqrt{3}$

B. $4+2\sqrt{3}$　　D. $2+\sqrt{3}$

解 答案：B.

设 $O(0,0)$ 为坐标原点，由 $A(-2,0)$、$B(2,0)$、$C(0,4)$，知 $|AC|=|BC|=2\sqrt{5}$，且 $\triangle ABC$ 为锐角三角形，因此费马点 M 在线段 OC 上. 设 $M(0,h)$，如图，则 $\triangle MAB$ 为顶角是 120° 的等腰三角形，故

$$h=|OB|\tan 30^\circ=\frac{2\sqrt{3}}{3},$$

所以

$$f(P)\geq f(M)=|MA|+|MB|+|MC|=4h+4-h=4+2\sqrt{3},$$

则 $f(P)$ 的最小值为 $4+2\sqrt{3}$.

故选：B.

例题 13.268 以三角形边 BC、CA、AB 为边向形外作正三角形 BCA'、CAB'、ABC'，则 AA'、BB'、CC' 三线共点，该点称为 $\triangle ABC$ 的正等角中心，当 $\triangle ABC$ 的每个内角都小于 120° 时，正等角中心点 P 满足以下性质：(1)$\angle APB=\angle APC=\angle BPC=120^\circ$；(2) 正等角中心是

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

到该三角形三个顶点距离之和最小的点（也即费马点）. 由以上性质得 $\sqrt{x^2+(y-1)^2}+\sqrt{x^2+(y+1)^2}+\sqrt{(x-2)^2+y^2}$ 的最小值为

解 答案:$2+\sqrt{3}$.

根据题意, 在平面直角坐标系中, 令点 $A(0,1)$、$B(0,-1)$、$C(2,0)$, 则 $\sqrt{x^2+(y-1)^2}+\sqrt{x^2+(y+1)^2}+\sqrt{(x-2)^2+y^2}$ 表示坐标系中一点 (x,y) 到点 A、B、C 的距离之和. 因为 $\triangle ABC$ 是等腰三角形, $AC=BC$, 所以 C' 点在 x 轴负半轴上, CC' 与 x 轴重合. 令 $\triangle ABC'$ 的费马点为 $P(a,b)$, 则 P 在 CC' 上, $b=0$. 因为 $\triangle ABC'$ 是锐角三角形, 由性质 (1) 得 $\angle APC=120^\circ$, 所以 $\angle APO=60^\circ$, 所以

$$\frac{1}{a}=\sqrt{3},\quad a=\frac{\sqrt{3}}{3}.$$

$\therefore P\left(\frac{\sqrt{3}}{3},0\right)$ 到 A、B、C 的距离分别为 $PA=PB=\frac{2\sqrt{3}}{3}$, $PC=2-\frac{\sqrt{3}}{3}$, 所以 $\sqrt{x^2+(y-1)^2}+\sqrt{x^2+(y+1)^2}+\sqrt{(x-2)^2+y^2}$ 的最小值即为费马点 P 到点 A、B、C 的距离之和, 则

$$PA+PB+PC=2+\sqrt{3}.$$

故答案为:$2+\sqrt{3}$.

例题 13.269 已知平面上的线段 l 及点 P, 任取 l 上一点 Q, 线段 PQ 长度的最小值称为点 P 到线段 l 的距离, 记作 $d(P,l)$. 请你写出到两条线段 l_1,l_2 距离相等的点的集合 $\Omega=\{P|d(P,l_1)=d(P,l_2)\}$, 其中 $l_1=AB,l_2=CD$; A,B,C,D 是下列两组点中的一组. 对于下列两种情形, 只需选做一种, 满分分别是 (1) 3 分; (2) 5 分.(1) $A(1,3),B(1,0),C(-1,3),D(-1,0)$; (2) $A(1,3),B(1,0),C(-1,3),D(-1,-2)$. 你选择第__ 种情形, 到两条线段 l_1,l_2 距离相等的点的集合 $\Omega=$__

解 答案: (1)y 轴; (2)y 轴非负半轴, 抛物线 $y^2=4x(-2\leq y\leq 0)$, 直线 $y=-x-1(x>1)$.

根据题意从两组点的坐标中选一组, 根据所给的四个点的坐标, 写出两条直线的方程, 从直线方程中看出这两条直线之间的平行关系, 得到要求的结果.

对于 (1), $A(1,3),B(1,0),C(-1,3),D(-1,0)$, 利用两点式写出两条直线的方程 $AB:x=1,CD:x=-1$, 到两条线段 l_1,l_2 距离相等的点的集合 $\Omega=\{P|d(P,l_1)=d(P,l_2)\}$, 根据两条直线的方程可知两条直线之间的关系是平行, $\therefore$ 到两条线段 l_1,l_2 距离相等的点的集合为 $\Omega=\{(x,y)|x=0\}$.

对于 (2), $A(1,3),B(1,0),C(-1,3),D(-1,-2)$: 根据第一组作出的结果, 观察第二组数据的特点, 连接得到线段以后, 可以得到到两条线段距离相等的点是 y

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

轴的非负半轴, 抛物线 $y^2=4x(-2\leq y\leq 0)$, 直线 $y=-x-1(x>1)$, 故满足条件的集合 $\Omega=\{(x,y)|x=0 \text{ 且 } y\geq 0\}\cup\{(x,y)|y^2=4x,0\leq x\leq 1,-2\leq y\leq 0\}\cup\{(x,y)|y=-x-1,x>1\}$.

例题 13.270 (2) 已知点 $P(3,1)$ 和直线 $l:2x-y-1=0$, 则 $d(P,l)=\frac{4}{3}$; (3) 定义 $O(0,0)$, 动点 $P(x,y)$ 满足 $d(P,O)=1$, 则动点 P 的轨迹围成平面图形的面积是 4; 其中真命题的个数(　)

A. 0　　C. 2

B. 1　　D. 3

解 答案: B.

由新定义表示出三点 A,B,C 两两之间的切比雪夫距离, 然后根据绝对值的性质判断 (1), 由新定义计算出 $d(P,l)$, 判断 (2); 根据新定义求出 P 的轨迹方程, 确定其轨迹, 求得轨迹围成的图形面积判断 (3).

(1) 设 $A(x_1,y_1),B(x_2,y_2),C(x_3,y_3)$, 则 $d(A,B)=|x_1-x_2|+|y_1-y_2|$, $d(A,C)+d(B,C)=|x_1-x_3|+|y_1-y_3|+|x_2-x_3|+|y_2-y_3|$, 显然 $|x_1-x_3|+|x_2-x_3|\geq|(x_1-x_3)-(x_2-x_3)|=|x_1-x_2|$, 同理 $|y_1-y_3|+|y_2-y_3|\geq|y_1-y_2|$, $\therefore d(C,A)+d(C,B)\geq d(A,B)$, (1) 正确.

(2) 设 $P(x,y)$ 是直线 l 上任一点, 则 $y=2x-1$,

$$d(P,l)=|x-3|+|y-1|=|x-3|+|2x-2|=\begin{cases}3x-5,x\geq 3\\ x+1,1\leq x<3\\ 5-3x,x<1\end{cases}$$

上是减函数, $\therefore x=1$ 时, $d(P,l)_{\min}=|1-3|+|2-2|=2$, (2) 错.

(3) 由 $d(P,O)=1$ 得 $|x|+|y|=1$, 易知此曲线关于 x 轴, y 轴, 原点都对称, 它是以 $(1,0),(0,1),(-1,0),(0,-1)$ 为顶点的正方形, 其构成图形面积为 $S=\frac{1}{2}\times 2\times 2=2$, (3) 错.

故选: B.

例题 13.271 (2024·上海·二模) 在平面直角坐标系中, 定义 $d(A,B)=\max\{|x_1-x_2|,|y_1-y_2|\}$ 为两点 $A(x_1,y_1)$、$B(x_2,y_2)$ 的“切比雪夫距离”, 又设点 P 及 l 上任意一点 Q, 称 $d(P,Q)$ 的最小值为点 P 到直线 l 的“切比雪夫距离”, 记作 $d(P,l)$, 给出下列三个命题: (1) 对任意三点 A、B、C, 都有 $d(C,A)+d(C,B)\geq d(A,B)$; (2) 已知点 $P(3,1)$ 和直线 $l:2x-y-1=0$, 则 $d(P,l)=\frac{4}{3}$; (3) 定点 $F_1(-c,0)$、$F_2(c,0)$, 动点 $P(x,y)$ 满

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

足 $|d(P,F_1)-d(P,F_2)|=2a(2c>2a>0)$, 则点 P 的轨迹与直线 $y=kx$ 有且仅有 2 个公共点, 其中 k 为常数. 其中真命题的个数是

A. 0　　C. 2

B. 1　　D. 3

解 答案: C. 设 $A(x_A,y_A),B(x_B,y_B),C(x_C,y_C)$, 由题意可得: $d(C,A)+d(C,B)=\max\{|x_A-x_C|,|y_A-y_C|\}+\max\{|x_B-x_C|,|y_B-y_C|\}\geqslant|x_A-x_C|+|x_B-x_C|\geqslant|x_A-x_B|$, 同理可得: $d(C,A)+d(C,B)\geqslant|y_A-y_B|$, 则: $d(C,A)+d(C,B)\geqslant\max\{|x_A-x_B|,|y_A-y_B|\}=d(A,B)$. 命题 (1) 成立; 设点 Q 是直线 $y=2x-1$ 上一点, 且 $Q(x,2x-1)$, 可得 $d(P,Q)=\max\{|x-3|,|2-2x|\}$,

例题 13.272　(2024·高三·上海浦东新·期中) 在平面直角坐标系中, 定义 $d(A,B)=\max\{|x_1-x_2|,|y_1-y_2|\}$ 为两点 $A(x_1,y_1)$、$B(x_2,y_2)$ 的"切比雪夫距离", 又设点 P 及 l 上任意一点 Q, 称 $d(P,Q)$ 的最小值为点 P 到直线 l 的"切比雪夫距离", 记作 $d(P,l)$, 给出四个命题, 正确的是__.(1) 对任意三点 A、B、C, 都有 $d(C,A)+d(C,B)\geqslant d(A,B)$; (2) 到原点的切比雪夫距离等于 1 的点的轨迹是正方形; (3) 已知点 $P(3,1)$ 和直线 $l:2x-y-1=0$, 则 $d(P,l)=\frac{4}{3}$; (4) 定点 $F_1(-c,0)$、$F_2(c,0)$, 动点 $P(x,y)$ 满足 $|d(P,F_1)-d(P,F_2)|=2a.(2c>2a>0)$, 则点 P 的轨迹与直线 $y=k$ 有且仅有 2 个公共点, 其中 k 为常数.

解 答案: (1)(2)(3)(4). (1) 对任意三点 A、B、C, 若它们共线, 设 $A(x_1,y_1)$、$B(x_2,y_2)$、$C(x_3,y_3)$, 如下图, 结合三角形相似可得 $d(C,A)=AN$ 或 CN, $d(C,B)=CM$ 或 BM, $d(A,B)=AK$ 或 BK, 则 $d(C,A)+d(C,B)=d(A,B)$; 若 B、C 或 A、C 对调, 可得 $d(C,A)+d(C,B)>d(A,B)$; 若 A、B、C 不共线, 且 $\triangle ABC$ 中 C 为锐角或钝角, 由矩形 $CMNK$ 或矩形 $BMNK$, ……由 $|x-3|\geqslant|2-2x|$, 解得 $-1\leqslant x\leqslant\frac{5}{3}$, 即有 $d(P,Q)=|x-3|$, 当 $x=\frac{5}{3}$ 时取得最小值 $\frac{4}{3}$; 由 $|x-3|<|2-2x|$, 解得 $x>\frac{5}{3}$ 或 $x<-1$, 即有 $d(P,Q)=|2x-2|$, $d(P,Q)$ 的范围是 $(3,+\infty)\cup(\frac{4}{3},+\infty)=(\frac{4}{3},+\infty)$, 无最小值. 综上可得, P,Q 两点的切比雪夫距离的最小值为 $\frac{4}{3}$. 说法 (2) 正确; 定点 $F_1(-c,0),F_2(c,0)$, 动点 $P(x,y)$ 满足 $|d(P,F_1)-d(P,F_2)|=2a\ (2c>2a>0)$, 则 $|\max\{|x+c|,|y|\}-\max\{|x-c|,|y|\}|=2a$, 显然上述方程所表示的曲线关于原点对称, 故不妨设 $x\geqslant0,y\geqslant0$.(1) 当 $\begin{cases}|x+c|\geqslant y\\ |x+c|-|x-c||=2a\end{cases}$, 得 $\begin{cases}x=a\\ 0\leqslant y\leqslant a-c\end{cases}$;

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(2) 当 $\begin{cases}|x+c|\leqslant y\\|x-c|\leqslant y\end{cases}$，……；(3) 当 $\begin{cases}|x+c|\geqslant y\\|x+c|-y=2a,a<x\\|x-c|<y\end{cases}$，……则点 P 的轨迹是如图所示的以原点为中心的两支折线. 结合图象可知，点 P 的轨迹与直线 $y=k$（k 为常数）有且仅有 2 个公共点，命题 (3) 正确. 综上可得命题 (1)(2)(3) 均正确，真命题的个数是 3. 本题选择 D 选项.

13.18.4 本节小结

- 距离的创新定义：识别距离类型，代入圆上动点求最值.
- 包络线：分析圆族参数变化，找出与所有圆相切的曲线.
- 阿波罗尼斯圆：距离比为定值 $\neq 1$ 时，轨迹是圆.
- 反演：保角保圆，把圆系问题转化为更简单的图形.

知乎@余命数

13.18.5 本节习题

13.100. 动点 $P(x,y)$ 在圆 $x^2+y^2=4$ 上，求切比雪夫距离 $d_C(P,O)$ 的最大值，其中 O 为原点.

13.101. 求圆心在 x 轴上、半径等于圆心到原点距离的圆族的包络线.

13.102. 动点 P 满足 $|PA|=\sqrt{2}|PB|$、$A(0,0)$、$B(4,0)$. 求阿波罗尼斯圆的方程.

13.19 垂直弦问题

垂直弦问题是圆锥曲线中一类特殊的定点与定值问题. 当弦的两端在曲线上运动且满足某种垂直条件（内接直角三角形、两条互相垂直的弦）时，弦所在直线或弦中点所在直线往往过某个定点，或其范围有确定的最值.

13.19.1 内接直角三角形的斜边过定点

椭圆、双曲线、抛物线内接直角三角形的斜边必过某个定点. 设直角顶点在曲线上，两条直角边斜率之积为 -1，联立曲线方程后，斜边方程中的参数被消去，从而得到定点.

13.19.2 两条互相垂直的弦中点所在直线过定点

椭圆、双曲线、抛物线中，过同一个定点 T 的两条互相垂直的弦，其中点所在直线过某个定点. 设两条弦的斜率之积为 -1，分别求出中点坐标，中点连线的方程中参数被消去，得到定点.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

13.19.3 范围与最值

内接直角三角形和两条互相垂直的弦中点都有相应的范围与最值问题, 通常结合弦长公式、韦达定理和不等式求解.

例题 13.273 (椭圆内接直角三角形的斜边必过定点) 已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 的左右焦点分别 F_1,F_2, 若 $__$, 请把以下两个条件中任选一个补充在横线上作答(若都选择, 则按照第一个解答给分) (1) 四点 $F_1(1,2)$、$F_2(-\sqrt{2},0)$、$P_1\left(1,-\frac{\sqrt{2}}{2}\right)$、$P_4\left(-1,\frac{\sqrt{2}}{2}\right)$ 中, 恰有三点在椭圆 C 上. (2) 椭圆 C 经过 $Q(-1,y_0)$, $QF_1\perp x$ 轴, 且 $\sin\angle QF_2F_1=\frac{1}{3}$. (1) 求椭圆 C 的方程;(2) 设点 D 为椭圆 C 的上顶点, 过点 D 作两条互相垂直的直线分别交椭圆于 A、B 两点, 过 D 作直线 AB 的垂线垂足为 M, 判断 y 轴上是否存在定点 N, 使得 $|MN|$ 为定值? 请证明你的结论.

解 【解答】(1) 若选 (1): 因为 $F_1(1,2),F_2(-\sqrt{2},0),P_3\left(1,-\frac{\sqrt{2}}{2}\right),P_4\left(-1,\frac{\sqrt{2}}{2}\right)$ 中有三点在椭圆上, 由于 P_3,P_4 关于原点对称, 所以 P_3,P_4 均在椭圆上,

例题 13.274 (双曲线内接直角三角形的斜边必过定点) 在平面直角坐标系 xOy 中, 动点 P 与定点 $F(2,0)$ 的距离和它到定直线 $l:x=\frac{3}{2}$ 的距离之比是常数 $\frac{2\sqrt{3}}{3}$, 记 P 的轨迹为曲线 E. (1) 求曲线 E 的方程;(2) 设过点 $A(\sqrt{3},0)$ 两条互相垂直的直线分别与曲线 E 交于点 M,N(异于点 A), 求证: 直线 MN 过定点.

解 (1) 设 $P(x,y)$, 因为 P 与定点 $F(2,0)$ 的距离和它到定直线 $l:x=\frac{3}{2}$ 的距离之比是常数 $\frac{2\sqrt{3}}{3}$, 所以 $\frac{\sqrt{(x-2)^2+y^2}}{|x-\frac{3}{2}|}=\frac{2\sqrt{3}}{3}$,

例题 13.275 (抛物线内接直角三角形的斜边必过定点) 已知过抛物线 $E:y^2=2px(p>0)$ 的焦点, 斜率为 $2\sqrt{2}$ 的直线 l 交抛物线于 A, B 两点, 且 $|AB|=9$. (1) 求抛物线 E 的方程;(2) 设过点 $O(0,0)$ 且互相垂直的两条直线与抛物线 E 分别交于点 M,N, 证明: 直线 MN 过定点.

解 答案: (1) $y^2=8x$ (2) 定点 $(8,0)$.

(1) 抛物线 $E:y^2=2px$ 的焦点 $F(\frac{p}{2},0)$, 则直线 AB 的方程为 $y=2\sqrt{2}(x-\frac{p}{2})$. 由

$$\begin{cases}y^2=2px\\ y=2\sqrt{2}(x-\frac{p}{2})\end{cases}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

消去 y 并整理得 $x^2-\frac{5}{4}px+\frac{p^2}{4}=0$，显然 $\Delta>0$. 设 $A(x_1,y_1),B(x_2,y_2)$，则 $x_1+x_2=\frac{5}{4}p$, 因此

$$|AB|=x_1+x_2+p=\frac{9}{4}p=9,$$

解得 $p=4$, 所以抛物线 E 的方程为 $y^2=8x$.

(2) 显然直线 MN 不垂直于 y 轴，设直线 MN 的方程为 $x=my+t$，点 $M(x_3,y_3),N(x_4,y_4)$. 由

$$\begin{cases}x=my+t\\y^2=8x\end{cases}$$

消去 x 得 $y^2-8my-8t=0$, 当 $\Delta'=64m^2+32t>0$ 时, $y_3y_4=-8t$. 由 $OM\perp ON$, 得

$$\overrightarrow{OM}\cdot\overrightarrow{ON}=x_3x_4+y_3y_4=\frac{(y_3y_4)^2}{64}+y_3y_4=\frac{(-8t)^2}{64}-8t=t^2-8t=0,$$

显然 $t\neq0$, 因此 $t=8$, 满足 $\Delta'>0$, 则直线 MN: $x=my+8$, 过定点 $(8,0)$, 所以直线 MN 过定点.

例题 13.276 (椭圆两条互相垂直的弦中点所在直线过定点) 已知椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左右焦点分别为 F_1,F_2, 抛物线 $y^2=4x$ 与椭圆有相同的焦点, 点 P 为抛物线与椭圆在第一象限的交点, 且 $|PF_1|=\frac{7}{3}$. (1) 求椭圆的方程;(2) 过 F 作两条斜率不为 0 且互相垂直的直线分别交椭圆于 A,B 和 C,D, 线段 AB 的中点为 M, 线段 CD 的中点为 N, 证明: 直线 MN 过定点, 并求出该定点的坐标.

解 答案:(1) $\frac{x^2}{4}+\frac{y^2}{3}=1$;(2) 定点坐标为 $\left(\frac{4}{7},0\right)$.

(1) 抛物线焦点坐标为 $(1,0)$, 故 $a^2-b^2=1$. 设 $|PF_2|=t$, 由抛物线定义得点 P 到直线 $x=-1$ 的距离为 t.$\therefore\cos\angle PF_1F_2=\frac{3t}{7}$, 由余弦定理, 得

$$\cos\angle PF_1F_2=\frac{4+\frac{49}{9}-t^2}{2\times2\times\frac{7}{3}}=\frac{3t}{7}.$$

整理，得 $9t^2+36t-85=0$，解得 $t=\frac{5}{3}$ 或 $t=-\frac{17}{3}$（舍去）. 由椭圆定义，得 $|PF_1|+|PF_2|=2a=4$、$\therefore a=2,b=\sqrt{3}$、$\therefore$ 椭圆的方程为 $\frac{x^2}{4}+\frac{y^2}{3}=1$.

(2) 设 $l_{AB}:x=my+1\,(m\neq0)$, 联立

$$\begin{cases}x=my+1\\\frac{x^2}{4}+\frac{y^2}{3}=1\end{cases}\Rightarrow(3m^2+4)y^2+6my-9=0,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即 $y_A+y_B=\frac{-6m}{3m^2+4}$，$\therefore y_M=\frac{y_A+y_B}{2}=\frac{-3m}{3m^2+4}$，代入直线方程得 $x_M=\frac{4}{3m^2+4}$，$\therefore M\left(\frac{4}{3m^2+4},\frac{-3m}{3m^2+4}\right)$. 同理可得 $N\left(\frac{4m^2}{4m^2+3},\frac{3m}{4m^2+3}\right)$、$\therefore k_{MN}=\frac{7m}{4m^2-4}$、$\therefore$

$$l_{MN}:y+\frac{3m}{3m^2+4}=\frac{7m}{4m^2-4}\left(x-\frac{4}{3m^2+4}\right).$$

令 $y=0$, 得

$$x=\frac{4}{3m^2+4}+\frac{12m^2-12}{7(3m^2+4)}=\frac{12m^2+16}{7(3m^2+4)}=\frac{4}{7},$$

所以直线 MN 过定点 $\left(\frac{4}{7},0\right)$.

例题 13.277 (双曲线两条互相垂直的弦中点所在直线过定点) 已知双曲线 $E:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1(a>0,b>0)$ 的一条渐近线方程为 $x-\sqrt{3}y=0$, 焦点到渐近线的距离为 1. (1) 求 E 的方程; (2) 过双曲线 E 的右焦点 F 作互相垂直的两条弦(斜率均存在)AB、CD. 两条弦的中点分别为 P、Q, 那

例题 13.278 (抛物线两条互相垂直的弦中点所在直线过定点) 已知一个边长为 $8\sqrt{3}$ 的等边三角形的一个顶点位于原点, 另外两个顶点在抛物线 C : $x^2=2py(p>0)$ 上. (1) 求抛物线 C 的方程; (2) 过点 $T(0,p)$ 作两条互相垂直的直线 l_1 和 l_2, l_1 交抛物线 C 于 A、B 两点, l_2 交抛物线 C 于 D, E 两点, 若线段 AB 的中点为 M, 线段 DE 的中点为 N, 证明: 直线 MN 过定点.

解 (1) 由对称性可知等边三角形的顶点 $(4\sqrt{3},12)$ 在 $C:x^2=2py(p>0)$ 上, 代入得 $48=24p$, 解得 $p=2$, 所以抛物线方程为 $x^2=4y$.

(2) 由题意知 l_1 和 l_2 斜率均存在, $T(0,2)$, 设直线 l_1 方程为 $y=kx+2$, 则直线 l_2 方程为 $y=-\frac{1}{k}x+2$. 由

$$\begin{cases}x^2=4y\\ y=kx+2\end{cases}$$

联立得 $x^2-4kx-8=0,\Delta=16k^2+32>0$. 设 $A(x_1,y_1),B(x_2,y_2)$, 则 $x_1+x_2=4k$, $y_1+y_2=4k^2+4$, 故 $M(2k,2k^2+2)$, 同理得 $N\left(-\frac{2}{k},\frac{2}{k^2}+2\right)$, 故直线 MN 方程为

$$y-(2k^2+2)=\frac{2k^2-\frac{2}{k^2}}{2k+\frac{2}{k}}(x-2k)=\left(k-\frac{1}{k}\right)(x-2k).$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

例题 13.279 (内接直角三角形范围与最值问题) 设椭圆 $\frac{y^2}{a^2}+\frac{x^2}{b^2}=1(a>b>0)$ 的两焦点为 F_1,F_2, P 为椭圆上任意一点, 点 P 到原点最大距离为 2; 若 $M(3,1)$ 到椭圆右顶点距离为 $\sqrt{5}$.

(1) 求椭圆的方程

(2) 设椭圆的上、下顶点分别为 A、B, 过 A 作两条互相垂直的直线交椭圆于 C、D, 问直线 CD 是否经过定点? 如果是, 请求出定点坐标, 并求出 $\triangle BCD$ 面积的最大值. 如果不是, 请说明理由.

解 答案: $\frac{y^2}{4}+x^2=1$. (1): 点 P 到原点最大距离为 2, 故 $a=2$, $\therefore M(3,1)$ 到椭圆右顶点距离为 $\sqrt{5}$, $\therefore(3-b)^2+1=5$, 解得: $b=1$ 或 5 (舍去 5), $\therefore$ 椭圆的方程为 $\frac{y^2}{4}+x^2=1$.

(2) 设 $CD: y=kx+t$, 联立 $\frac{y^2}{4}+x^2=1$, 得: $(4+k^2)x^2+2ktx+t^2-4=0$, $\therefore x_1+x_2=-\frac{2kt}{4+k^2}, x_1x_2=\frac{t^2-4}{4+k^2}$ $\therefore AC\perp AD, \therefore\overrightarrow{AC}\cdot\overrightarrow{AD}=0$, 即 $(x_1,y_1-2)\cdot(x_2,y_2-2)=x_1x_2+(y_1-2)(y_2-2)=x_1x_2+(kx_1+t-2)(kx_2+t-2)$

例题 13.280 (两条互相垂直的弦中点范围与最值问题) 如图, 抛物线 $\Gamma: y^2=2px(p>0)$, $M(2,1)$ 是抛物线内一点, 过点 M 作两条斜率存在且互相垂直的动直线 l_1,l_2, 设 l_1 与抛物线 Γ 相交于点 A,B, l_2 与抛物线 Γ 相交于点 C,D, 当 M 恰好为线段 AB 的中点时, $|AB|=2\sqrt{6}$. (1) 求抛物线 Γ 的方程; (2) 求 $\overrightarrow{AC}\cdot\overrightarrow{DB}$ 的最小值.

解 (1) 解法一: 设直线 $AB: x-2=m(y-1), A(x_1,y_1), B(x_2,y_2)$; 联立 $\begin{cases}x=my-m+2\\y^2=2px\end{cases}$, 得 $y^2-2pmy+2pm-4p=0$, 所以 $y_1+y_2=2pm, y_1y_2=2pm-4p$. 又因为 $M(2,1)$ 是 AB 的中点, 所以 $\frac{y_1+y_2}{2}=pm=1$, 又 $|AB|=\sqrt{1+m^2}|y_1-y_2|=\sqrt{1+m^2}\sqrt{(y_1+y_2)^2-4y_1y_2}=\sqrt{1+m^2}\sqrt{16p-4}=2\sqrt{6}$.

例题 13.281 过 M 点作两条互相垂直的直线 MA、MB 与椭圆交于 A、B 两点, 证明直线 AB 过定点, 并求 $\triangle DAB$ 面积的最大值.

解 答案: 直线 AB 过定点 $(-\frac{6}{5},0)$, $\triangle DAB$ 面积的最大值为 $\frac{64}{25}$.

(1) 由已知可得: $\frac{c}{a}=\frac{\sqrt{3}}{2}, a^2=1+c^2$, 解得: $a=2, c=\sqrt{3}$, 所以椭圆的方程为 $\frac{x^2}{4}+y^2=1$.

知乎@余命数

(2) 易知点 $M(-2,0)$, 设点 $A(x_1,y_1)$、$B(x_2,y_2)$, 则 $y_1^2=\frac{4-x_1^2}{4}$. 若直线 $AB\perp y$ 轴, 则 $x_2=-x_1, y_2=y_1$, 所以

$$k_{MA}k_{MB}=\frac{y_1}{x_1+2}\cdot\frac{y_1}{-x_1+2}=\frac{y_1^2}{4-x_1^2}=\frac{\frac{4-x_1^2}{4}}{4-x_1^2}=\frac{1}{4},$$

不合乎题意. 设 AB 的直线方程为 $x=ty+m$, 联立

$$\begin{cases}x=ty+m\\x^2+4y^2-4=0\end{cases}$$

整理得 $(t^2+4)y^2+2mty+m^2-4=0$,$\Delta=4m^2t^2-4(t^2+4)(m^2-4)=16(t^2+4-m^2)>0$. 由韦达定理可得

$$y_1+y_2=-\frac{2mt}{t^2+4},\quad y_1y_2=\frac{m^2-4}{t^2+4}.$$

因为 $\angle AMB=\frac{\pi}{2}$, 且 $\overrightarrow{MA}=(x_1+2,y_1)$, $\overrightarrow{MB}=(x_2+2,y_2)$, 所以

$$\overrightarrow{MA}\cdot\overrightarrow{MB}=(x_1+2)(x_2+2)+y_1y_2=(ty_1+m+2)(ty_2+m+2)+y_1y_2=0,$$

即 $(t^2+1)y_1y_2+t(m+2)(y_1+y_2)+(m+2)^2=0$. 代入韦达定理, 整理得

$$5m^2+16m+12=0,$$

解得 $m=-\frac{6}{5}$ 或 $m=-2$(舍去), 所以直线 AB 的方程为 $x=ty-\frac{6}{5}$. 此时

$$y_1+y_2=\frac{12t}{5(t^2+4)},\quad y_1y_2=\frac{-64}{25(t^2+4)},$$

则

$$S_{\triangle DAB}=\frac{1}{2}\cdot\left|2+\frac{6}{5}\right|\cdot|y_1-y_2|=\frac{8}{5}\sqrt{(y_1+y_2)^2-4y_1y_2}=\frac{32}{25}\cdot\frac{\sqrt{25t^2+64}}{t^2+4}.$$

令 $\sqrt{25t^2+64}=u\,(u\geqslant 8)$, 则

$$S_{\triangle DAB}=\frac{32}{25}\cdot\frac{u}{\frac{u^2-64}{25}+4}=\frac{32u}{u^2+36}=\frac{32}{u+\frac{36}{u}}.$$

由对勾函数单调性知, 函数 $y=u+\frac{36}{u}$ 在 $[8,+\infty)$ 上为增函数, 则 $u+\frac{36}{u}\geqslant 8+\frac{36}{8}=\frac{25}{2}$, 所以

$$S_{\triangle DAB}\leqslant\frac{32}{\frac{25}{2}}=\frac{64}{25},$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

当且仅当 $u=8$, 即 $t=0$ 时等号成立, 此时 $S_{\triangle DAB}$ 最大值为 $\frac{64}{25}$.

例题 13.282 已知双曲线 C: $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1\ (a>0,b>0)$ 经过点 $P(2,1)$, 且双曲线 C 的右顶点到一条渐近线的距离为 $\frac{\sqrt{6}}{3}$. (1) 求双曲线 C 的方程; (2) 过点 P 分别作两条互相垂直的直线 PA, PB 与双曲线 C 交于 A, B 两点(A, B 两点均与点 P 不重合), 设直线 AB: $y=kx+m\ (k\neq 0)$, 试求 k 和 m 之间满足的关系式.

解 答案: (1) $\frac{x^2}{2}-y^2=1$; (2) $m=-6k-3$ 或 $m=-2k+1$(舍去).

(1) 已知双曲线 C: $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1\ (a>0,b>0)$ 经过点 $P(2,1)$, 则 $\frac{4}{a^2}-\frac{1}{b^2}=1$. 右顶点为 $(a,0)$, 不妨取渐近线为 $y=-\frac{b}{a}x$, 即 $bx+ay=0$, 则

$$\frac{|ab|}{\sqrt{a^2+b^2}}=\frac{\sqrt{6}}{3},$$

知乎@余命数

从而可解得 $a^2=2, b^2=1$, 所以双曲线 C 的方程为 $\frac{x^2}{2}-y^2=1$.

(2) 设 $A(x_1,y_1), B(x_2,y_2)$, 联立

知乎@余命数

知乎@余命数

$$\begin{cases}\frac{x^2}{2}-y^2=1\\ y=kx+m\end{cases}$$

消 y 得 $(1-2k^2)x^2-4kmx-2m^2-2=0$, 则 $1-2k^2\neq 0$, $\Delta=16k^2m^2-4(1-2k^2)(-2m^2-2)=8(m^2+1-2k^2)>0$, 且

$$x_1+x_2=\frac{4km}{1-2k^2},\quad x_1\cdot x_2=\frac{-2(m^2+1)}{1-2k^2}.$$

因为 $PA\perp PB$, 所以 $\overrightarrow{PA}\cdot\overrightarrow{PB}=(x_1-2)(x_2-2)+(y_1-1)(y_2-1)=0$, 代入化简得

$$(k^2+1)x_1x_2+[k(m-1)-2](x_1+x_2)+(m-1)^2+4=0,$$

整理得 $12k^2+8km+m^2+2m-3=0$, 即 $(6k+m+3)(2k+m-1)=0$, 解得 $m=-6k-3$ 或 $m=-2k+1$. 当 $m=-2k+1$ 时, $MN: y=kx+m=k(x-2)+1$ 过点 $P(2,1)$, 与条件不符, 舍去, 所以 $m=-6k-3$, 即为 k 和 m 之间满足的关系式.

例题 13.283 已知抛物线 $C: y^2=2px(p>0)$, O 是坐标原点, F 是 C 的焦点, M 是 C 上一点, $|FM|=4$, $\angle OFM=120^\circ$. (1) 求抛物线 C 的标准方程; (2) 设点 $Q(x_0,2)$ 在 C 上, 过 Q 作两条互相垂直的直线 QA, QB, 分别交 C 于 A, B 两点(异于 Q 点). 证明: 直线 AB 恒过定点.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 答案：(1) $y^2=4x$ (2) 证明见解析. (1) 由 $|FM|=4,\angle OFM=120°$，可得 $M\left(\frac{p}{2}+2,\pm2\sqrt{3}\right)$，代入 $C:12=2p\left(\frac{p}{2}+2\right)=p^2+4p$. 解得 $p=2$ 或 $p=-6$(舍)，所以抛物线的方程为：$y^2=4x$. (2) 解：由题意可得 $Q(1,2)$，直线 AB 的斜率不为 0，设直线 AB 的方程为 $x=my+n$，设 $A(x_1,y_1),B(x_2,y_2)$，由 $\begin{cases}y^2=4x\\x=my+n\end{cases}$，得 $y^2-4my-4n=0$，从而 $\Delta=16m^2+16n>0$，则 $\begin{cases}y_1+y_2=4m\\y_1y_2=-4n\end{cases}$. 所以 $x_1+x_2=m(y_1+y_2)+2n=4m^2+2n$，$x_1x_2=(my_1+n)(my_2+n)=m^2y_1y_2+mn(y_1+y_2)+n^2=n^2$，$\because\overrightarrow{QA}\perp\overrightarrow{QB}$，$\therefore\overrightarrow{QA}\cdot\overrightarrow{QB}=(x_1-1)(x_2-1)+(y_1-2)(y_2-2)=0$，故 $x_1x_2-(x_1+x_2)+1+y_1y_2-2(y_1+y_2)+4=0$，

例题 13.284　在 (2) 条件下，当 M、N、F 三点可构成三角形时，求 $S_{\triangle MNF}$ 的取值范围.

解 答案：$(0,\frac{36}{49}]$.

由 (2) 知 $M\left(\frac{-8k^2}{3+4k^2},\frac{6k}{3+4k^2}\right)$，$N\left(\frac{-8}{3k^2+4},\frac{-6k}{3k^2+4}\right)$，所以

$$|MF|=\sqrt{\left(\frac{-8k^2}{3+4k^2}+2\right)^2+\left(\frac{6k}{3+4k^2}\right)^2}=\frac{6\sqrt{k^2+1}}{3+4k^2},$$

$$|NF|=\sqrt{\left(\frac{-8}{3k^2+4}+2\right)^2+\left(\frac{-6k}{3k^2+4}\right)^2}=\frac{6\sqrt{k^2+k^4}}{3k^2+4}.$$

由于 $MF\perp NF$，故

$$S_{\triangle MNF}=\frac{1}{2}|MF||NF|=\frac{1}{2}\cdot\frac{6\sqrt{k^2+1}}{3+4k^2}\cdot\frac{6\sqrt{k^2+k^4}}{3k^2+4}=\frac{18(k^2+1)|k|}{(3+4k^2)(3k^2+4)}.$$

由于 $k\neq0$，根据奇偶性不妨只考虑 $k>0$，则 $S_{\triangle MNF}=\frac{18(k^2+1)k}{(3+4k^2)(3k^2+4)}$. 记 $f(k)=\frac{(k^2+1)k}{(3+4k^2)(3k^2+4)}$，$k>0$，则

$$f'(k)=-\frac{(k^2-1)(12k^4+23k^2+12)}{(3+4k^2)^2(3k^2+4)^2}.$$

对于 $12k^4+23k^2+12$ 可知 $\Delta=23^2-4\times144<0$，故 $12k^4+23k^2+12>0$. 当 $0<k<1$ 时 $f'(k)>0$，$f(k)$ 在 $(0,1)$ 单调递增；当 $k>1$ 时 $f'(k)<0$，$f(k)$ 在 $(1,+\infty)$ 单调递减. 故 $f(k)$ 在 $k=1$ 取最大值，$f(1)=\frac{2}{49}$，故此时面积最大值为 $S_{\triangle MNF}=18f(1)=\frac{36}{49}$. 当 $k\to+\infty$ 时，$f(k)\to0$，故 $S_{\triangle MNF}\in(0,\frac{36}{49}]$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 13.285 已知椭圆 $M:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 上任意一点 P 到椭圆 M 两个焦点 F_1,F_2 的距离之和为 4, 且离心率为 $\frac{\sqrt{3}}{2}$. (1) 求椭圆 M 的标准方程; (2) 设 A 为 M 的左顶点, 过 A 点作两条互相垂直的直线 AC,AD 分别与 M 交于 C,D 两点, 证明: 直线 CD 经过定点, 并求这个定点的坐标.

解 答案: (1) $\frac{x^2}{4}+y^2=1$; (2) 定点为 $\left(-\frac{6}{5},0\right)$.

(1) 由椭圆定义知 $|PF_1|+|PF_2|=2a=4$, 解得 $a=2$. 又离心率 $e=\frac{c}{a}=\frac{\sqrt{3}}{2}$、$\therefore c=\sqrt{3}$、$\therefore b^2=a^2-c^2=1$、$\therefore$ 椭圆 M 的标准方程为 $\frac{x^2}{4}+y^2=1$.

(2) 由 (1) 知 $A(-2,0)$. 当直线 CD 斜率存在时, 设 $CD:y=kx+m$、$C(x_1,y_1)$、$D(x_2,y_2)$, 由

$$\begin{cases}y=kx+m\\ \frac{x^2}{4}+y^2=1\end{cases}$$

得 $(1+4k^2)x^2+8kmx+4m^2-4=0$, 则 $\Delta=16(1+4k^2-m^2)>0$, 解得 $m^2<1+4k^2$, $\therefore$

$$x_1+x_2=-\frac{8km}{1+4k^2},\quad x_1x_2=\frac{4m^2-4}{1+4k^2}.$$

$\because k_{AC}\cdot k_{AD}=\frac{y_1}{x_1+2}\cdot\frac{y_2}{x_2+2}=-1$, $\therefore y_1y_2=-x_1x_2-2(x_1+x_2)-4$, 即 $(kx_1+m)(kx_2+m)=k^2x_1x_2+mk(x_1+x_2)+m^2=-x_1x_2-2(x_1+x_2)-4$, $\therefore$

$$(k^2+1)x_1x_2+(mk+2)(x_1+x_2)+m^2+4=0,$$

即 $\frac{(k^2+1)(4m^2-4)}{1+4k^2}-\frac{8km(mk+2)}{1+4k^2}+m^2+4=0$, 整理可得 $5m^2-16km+12k^2=(5m-6k)(m-2k)=0$, $\therefore m=\frac{6}{5}k$ 或 $m=2k$. 当 $m=2k$ 时, 直线 $CD:y=kx+2k=k(x+2)$ 恒过 A 点, 不合题意; 当 $m=\frac{6}{5}k$ 时, 直线 $CD:y=kx+\frac{6}{5}k=k\left(x+\frac{6}{5}\right)$, $\therefore CD$ 恒过定点 $\left(-\frac{6}{5},0\right)$. 当直线 CD 斜率不存在且恒过 $\left(-\frac{6}{5},0\right)$ 时, 即 $CD:x=-\frac{6}{5}$, 由

$$\begin{cases}x=-\frac{6}{5}\\ \frac{x^2}{4}+y^2=1\end{cases}$$

得 $y=\pm\frac{4}{5}$, $\therefore k_{AC}\cdot k_{AD}=\frac{\frac{4}{5}}{-\frac{6}{5}+2}\times\frac{-\frac{4}{5}}{-\frac{6}{5}+2}=-1$, 满足题意. 直线 CD 恒过定点 $\left(-\frac{6}{5},0\right)$.

例题 13.286 已知抛物线 $G:x^2=2py(p>0)$ 焦点为 F, R 为 G 上的动点, $K(1,2)$ 位于 G 的上方区域, 且 $|RK|+|RF|$ 的最小值为 3. (1) 求 G 的方程; (2) 过点 $P(0,2)$ 作两条互相垂直的直线 l_1 和 l_2, l_1 交 G 于 A,B 两点, l_2 交 G

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于 C, D 两点, 且 M, N 分别为线段 AB 和 CD 的中点. 直线 MN 是否恒过一个定点? 若是, 求出该定点坐标; 若不是, 说明理由.

解 答案: (1) $x^2=4y$; (2) 是, 定点为 $(0,4)$.

(1) 抛物线 $G: x^2=2py(p>0)$ 焦点为 $F(0,\frac{p}{2})$, 准线为 $l: y=-\frac{p}{2}$, 设 R 到 l 的距离为 d. 因为 $K(1,2)$ 位于 G 的上方区域, 根据抛物线的定义可知

$$|RK|+|RF|=|RK|+d\geq 2+\frac{p}{2}$$

(当且仅当 $RK\perp l$ 时取等号), 又 $|RK|+|RF|$ 的最小值为 3, 所以 $2+\frac{p}{2}=3$, 解得 $p=2$, 所以抛物线 $G: x^2=4y$.

(2) 依题意直线 l_1 和 l_2 的斜率均存在且不为 0, 设直线 l_1 的方程为 $y=kx+2$, 则直线 l_2 的方程为 $y=-\frac{1}{k}x+2$、$A(x_1,y_1)$、$B(x_2,y_2)$. 联立方程得

$$\begin{cases} y=kx+2 \\ x^2=4y \end{cases}$$

消去 y 并整理得 $x^2-4kx-8=0$, 则 $\Delta=(-4k)^2+32=16k^2+32>0$, 则 $x_1+x_2=4k$, $x_1x_2=-8$, 所以 $y_1+y_2=k(x_1+x_2)+4=4k^2+4$. 因为 M 为 AB 的中点, 所以 $M(2k,2k^2+2)$, 同理 $N(-\frac{2}{k},\frac{2}{k^2}+2)$, 所以直线 MN 的方程为

$$y-(2k^2+2)=\frac{2k^2+2-(\frac{2}{k^2}+2)}{2k+\frac{2}{k}}(x-2k)=(k-\frac{1}{k})(x-2k),$$

整理得 $y=(k-\frac{1}{k})x+4$, 所以直线 MN 恒过点 $(0,4)$.

例题 13.287　过点 $K(0,-1)$ 作抛物线 $G: x^2=2py(p>0)$ 在第一象限部分的切线, 切点为 A, F 为 G 的焦点, O 为坐标原点, $\triangle OAF$ 的面积为 1. (1) 求 G 的方程; (2) 过点 $P(0,2)$ 作两条互相垂直的直线 l_1 和 l_2, l_1 交 G 于 C, D 两点, l_2 交 G 于 P, Q 两点, 且 M, N 分别为线段 CD 和 PQ 的中点. 直线 MN 是否恒过一个定点? 若是, 求出该定点坐标; 若不是, 说明理由.

解 答案: (1) $x^2=4y$; (2) 恒过定点 $(0,4)$.

(1) 由题, $y'=\frac{x}{p}$, 设切点 $A(x_0,y_0)$, 则切线方程为 $y-y_0=\frac{x_0}{p}(x-x_0)$, $x_0^2=2py_0$. $K(0,-1)$ 的坐标代入, 得 $-1-y_0=\frac{x_0}{p}(-x_0)$, 解得 $y_0=1$. 由于 $x_0>0$, 所以 $x_0=\sqrt{2p}$, 由 $\triangle OAF$ 的面积

$$S=\frac{1}{2}\cdot\frac{p}{2}\cdot\sqrt{2p}=1,$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解得 $p=2$, 所以 G 的方程为 $x^2=4y$.

(2) 由题意可知, 直线 l_1 和 l_2 斜率都存在且均不为 0, 设直线 l_1 的方程为 $y=kx+2$, 则直线 l_2 的方程为 $y=-\frac{1}{k}x+2$. 联立方程组

$$\begin{cases} y=kx+2 \\ x^2=4y \end{cases}$$

消去 y 并整理得 $x^2-4kx-8=0$, 则 $\Delta=(-4k)^2+32=16k^2+32>0$. 设 $C(x_1,y_1)$, $D(x_2,y_2)$, 则 $x_1+x_2=4k, x_1\cdot x_2=-8$, 所以 $y_1+y_2=k(x_1+x_2)+4=4k^2+4$. 因为 M 为 CD 中点, 所以 $M(2k,2k^2+2)$, 同理可得 $N\left(-\frac{2}{k},\frac{2}{k^2}+2\right)$, 所以直线 MN 的方程为

$$y-(2k^2+2)=\frac{2k^2+2-\left(\frac{2}{k^2}+2\right)}{2k+\frac{2}{k}}\cdot(x-2k)=\left(k-\frac{1}{k}\right)\cdot(x-2k),$$

整理得 $y=\left(k-\frac{1}{k}\right)x+4$, 所以直线 MN 恒过定点 $(0,4)$.

例题 13.288 72(2024·贵州·模拟预测) 已知抛物线 $C:y^2=2px(p>0)$ 的焦点 F 到双曲线 $\frac{x^2}{3}-y^2=1$ 的渐近线的距离为 $\frac{1}{2}$. (1) 求抛物线 C 的方程; (2) 过点 F 任意作互相垂直的两条直线 l_1,l_2 分别交曲线 C 于点 A,B 和 M,N, 设线段 AB,MN 的中点分别为 P,Q, 求证: 直线 PQ 恒过一个定点.

解 答案: (1) $y^2=4x$; (2) 定点 $E(3,0)$.

(1) 因为抛物线的焦点 F 为 $(\frac{p}{2},0)$, 双曲线的渐近线方程为 $y=\pm\frac{\sqrt{3}}{3}x$, 即 $x\pm\sqrt{3}y=0$, 则

$$\frac{\left|\frac{p}{2}\right|}{\sqrt{1^2+(\pm\sqrt{3})^2}}=\frac{1}{2},$$

解得 $p=2$, 故抛物线的方程为 $y^2=4x$.

(2) 设 A,B 两点坐标分别为 $(x_1,y_1),(x_2,y_2)$, 则点 P 的坐标为 $\left(\frac{x_1+x_2}{2},\frac{y_1+y_2}{2}\right)$. 由题意可设直线 l_1 的方程为 $y=k(x-1)(k\neq 0)$, 由

$$\begin{cases} y^2=4x \\ y=k(x-1) \end{cases}$$

得 $k^2x^2-(2k^2+4)x+k^2=0$, $\Delta=(2k^2+4)^2-4k^4=16k^2+16>0$. 因为直线 l_1 与曲线 C 交于 A,B 两点, 所以 $x_1+x_2=2+\frac{4}{k^2}$, $y_1+y_2=k(x_1+x_2-2)=\frac{4}{k}$, 所以点 P 的坐标为 $(1+\frac{2}{k^2},\frac{2}{k})$. 由题知, 直线 l_2 的斜率为 $-\frac{1}{k}$, 同理可得点 Q 的坐标为 $(1+2k^2,-2k)$. 当 $k\neq\pm1$ 时, 有 $1+\frac{2}{k^2}\neq 1+2k^2$, 此时直线 PQ 的斜率

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$k_{PQ}=\frac{\frac{2}{k}+2k}{1+\frac{2}{k^2}-1-2k^2}=\frac{k}{1-k^2}$, 所以直线 PQ 的方程为 $y+2k=\frac{k}{1-k^2}(x-1-2k^2)$, 整理得 $yk^2+(x-3)k-y=0$, 于是直线 PQ 恒过定点 $E(3,0)$. 当 $k=\pm1$ 时, 直线 PQ 的方程为 $x=3$, 也过定点 $E(3,0)$.

例题 13.289　在平面直角坐标系 xOy 中, 已知椭圆 $\Gamma:\frac{x^2}{3}+\frac{y^2}{2}=1$, 过右焦点 F 作两条互相垂直的弦 AB,CD, 设 AB,CD 中点分别为 M,N.
(1) 写出椭圆右焦点 F 的坐标及该椭圆的离心率; (2) 证明: 直线 MN 必过定点, 并求出此定点坐标; (3) 若弦 AB,CD 的斜率均存在, 求 $\triangle FMN$ 面积的最小值.

解　答案: (1) $F(1,0),e=\frac{\sqrt{3}}{3}$; (2) $\left(\frac{3}{5},0\right)$; (3) 0(下确界, 斜率趋于无穷或零时趋近于 0).
(1) 由椭圆的方程 $\frac{x^2}{3}+\frac{y^2}{2}=1$, 可得 $a^2=3,b^2=2$, 可得 $c^2=a^2-b^2=3-2=1$, 所以 $c=1$, 即右焦点 F 的坐标为 $(1,0)$, 离心率 $e=\frac{c}{a}=\frac{1}{\sqrt{3}}=\frac{\sqrt{3}}{3}$.

(2) 证明: 当直线 AB,CD 的斜率存在且不为 0 时, 设直线 AB 的方程为 $x=my+1$, 设 $A(x_1,y_1),B(x_2,y_2)$ 联立 $\begin{cases}x=my+1\\ \frac{x^2}{3}+\frac{y^2}{2}=1\end{cases}$, 整理可得: $(2m^2+3)y^2+4my-4=0$, 可得 $y_1+y_2=-\frac{4m}{2m^2+3}$, $x_1+x_2=m(y_1+y_2)+2=\frac{6}{2m^2+3}$,

例题 13.290　已知椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 的离心率为 $\frac{\sqrt{3}}{2}$, 左、右顶点分别为 A,B, 上顶点为 D, 坐标原点 O 到直线 AD 的距离为 $\frac{2\sqrt{5}}{5}$. (1) 求椭圆的方程; (2) 过 A 点作两条互相垂直的直线 AP,AQ 与椭圆交于 P,Q 两点, 求 $\triangle BPQ$ 面积的最大值.

解　答案: (1) $\frac{x^2}{4}+y^2=1$. (1) 由已知可得 $\begin{cases}\frac{c}{a}=\frac{\sqrt{3}}{2},\\ ab=\frac{2\sqrt{5}}{5}\cdot\sqrt{a^2+b^2}\end{cases}$ 解得, $a=2,b=1,c=\sqrt{3}$, $a^2=b^2+c^2$. 所以椭圆的方程为 $\frac{x^2}{4}+y^2=1$. (2) 设 PQ 的直线方程为 $x=ty+m$、$P(x_1,y_1)$、$Q(x_2,y_2)$, 联立方程 $\begin{cases}x=ty+m,\\ x^2+4y^2-4=0,\end{cases}$ 整理得 $(t^2+4)y^2+2mty+m^2-4=0$.

例题 13.291　已知抛物线 E 的顶点在原点, 焦点为 $F(2,0)$, 过焦点且斜率为 k 的直线交抛物线于 P,Q 两点, (1) 求抛物线方程; (2) 若 $|FP|=2|FQ|$, 求 k 的值; (3) 过点 $T(t,0)$ 作两条互相垂直的直线分别交抛物线 E 于 A,B,C,D 四点, 且 M,N 分别为线段 AB,CD 的中点, 求 $\triangle TMN$ 的面积最小值.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解 答案: (1) $y^2=8x$; (2) $k=\pm2\sqrt{2}$. (1)·抛物线 E 的顶点在原点, 焦点为 $F(2,0)$, $\therefore$ 抛物线方程为: $y^2=8x$;

(2) 由题意知: $k\neq0$, 可设直线 $PQ: x=\frac{1}{k}y+2$、$P(x_1,y_1)$、$Q(x_2,y_2)$;

$\because|FP|=2|FQ|$, $\therefore y_1=-2y_2$, 即 $\frac{y_1}{y_2}=-2$,

由 $\begin{cases}x=\frac{1}{k}y+2\\ y^2=8x\end{cases}$ 得: $y^2-\frac{8}{k}y-16=0$, $\therefore\begin{cases}y_1+y_2=\frac{8}{k}\\ y_1y_2=-16\end{cases}$

$\therefore\frac{(y_1+y_2)^2}{y_1y_2}=\frac{y_1^2+2y_1y_2+y_2^2}{y_1y_2}=\frac{y_1}{y_2}+\frac{y_2}{y_1}+2=-\frac{4}{k^2}$, 即 $-\frac{4}{k^2}=-2-\frac{1}{2}+2$,

解得: $k^2=8$, $\therefore k=\pm2\sqrt{2}$;

(3) 由题意知: 直线 AB,CD 的斜率均存在,

不妨设 $AB: x=my+t$、$A(x_1,y_1)$、$B(x_2,y_2)$、$C(x_3,y_3)$、$D(x_4,y_4)$;

则 $CD: x=-\frac{1}{m}y+t$;

由 $\begin{cases}x=my+t\\ y^2=8x\end{cases}$ 得: $y^2-8my-8t=0$, 则 $\Delta=64m^2+32t>0$, 即 $2m^2+t>0$;

$\therefore y_1+y_2=8m$、$y_1y_2=-8t$、$\therefore x_1+x_2=m(y_1+y_2)+2t=8m^2+2t$,

例题 13.292 已知椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$, 离心率为 $\frac{\sqrt{3}}{2}$, 点 $\left(\sqrt{3},\frac{1}{2}\right)$ 在圆上. (1) 求 E 的方程; (2) 过点 $K(-1,0)$ 作互相垂直的两条直线 l_1 与 l_2, 设 l_1 交 E 于 A,B 两点, l_2 交 E 于 C,D 两点, AB,CD 的中点分别为 M,N. 探究: $\triangle OMN$ 与 $\triangle KMN$ 的面积之比是否为定值? 若是, 请求出定值; 若不是, 请说明理由.

解 答案: (1) $\frac{x^2}{4}+y^2=1$; (2) 是, 定值为 4. (1) 由题意 $\frac{c}{a}=\frac{\sqrt{3}}{2}$, $\frac{3}{a^2}+\frac{1}{4b^2}=1$, 解得 $a=2,b=1$, 则 E 的方程 $\frac{x^2}{4}+y^2=1$ (2) 设 $l_1:\triangle OMN$ 与 $\triangle KMN$ 面积之比为定值, 定值为 4, 理由如下: 设直线 $AB: x=my-1, A(x_1,y_1), B(x_2,y_2)$, 讨论: (1) 当 $m\neq0$, 且 $m\neq\pm1$ 时, 联立 $\begin{cases}x=my-1\\ \frac{x^2}{4}+y^2=1\end{cases}$, 可得 $(m^2+4)y^2-2my-3=0$, $\Delta=16m^2+48>0$, 则 $y_1+y_2=\frac{2m}{m^2+4}$, 所以 $y_M=\frac{y_1+y_2}{2}=\frac{m}{m^2+4}$, $x_M=my_M-1=m\cdot\frac{m}{m^2+4}-1=\frac{-4}{m^2+4}$ 所以 $M\left(\frac{-4}{m^2+4},\frac{m}{m^2+4}\right)$, 设 $CD: x=\frac{-1}{m}y-1$, 同理可得 $N\left(\frac{-4m^2}{4m^2+1},\frac{-m}{4m^2+1}\right)$ 所以 $k_{MN}=\frac{\frac{-m}{4m^2+1}-\frac{m}{m^2+4}}{\frac{-4m^2}{4m^2+1}-\frac{-4}{m^2+4}}=\frac{5}{4}\frac{m}{m^2-1}$ $(m\neq0$ 且 $m\neq\pm1)$, 所以直线 $MN: y-\frac{m}{m^2+4}=\frac{5}{4}\frac{m}{m^2-1}\left(x+\frac{4}{m^2+4}\right)$, 即 $y=\frac{m}{m^2-1}\left(\frac{5}{4}x+1\right)$, 所以直线 MN 恒过定点 $T\left(-\frac{4}{5},0\right)$; (2) 当 $m=\pm1$ 时, 不妨设直线 $l_1: y=x+1$; $l_2: y=-x-1$, 可发现 $MN\perp x$ 轴, 且 MN 过 $T\left(-\frac{4}{5},0\right)$,

13.19.4 本节小结

- 内接直角三角形: 斜边过定点, 用斜率之积 -1 联立曲线方程消参.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 过同一点的两条互相垂直的弦:中点连线过定点,分别求中点后消参.
- 范围与最值:结合弦长公式与韦达定理,用不等式求极值.

知乎@余命数

13.19.5 本节习题

知乎@余命数

13.103. 设椭圆 $\frac{x^2}{4}+y^2=1$ 内接直角三角形 PAB 的直角顶点 P 在椭圆上,$PA\perp PB$. 证明斜边 AB 过定点,并求该定点.

13.104. 设椭圆 $\frac{x^2}{4}+y^2=1$ 上过点 $T(1,0)$ 的两条弦 AB,CD 互相垂直(斜率之积为 -1). 求 AB 中点与 CD 中点连线所过定点的坐标.

知乎@余命数

13.20 斜率和差商积

知乎@余命数

圆锥曲线中,直线与曲线交于两点后,两点处的切线斜率,或相关斜率的和、差、商、积往往有确定的值或范围. 设直线方程,联立曲线,用韦达定理表示出两个交点的坐标,再把目标斜率表达式代入,消参后得到定值或范围.

13.20.1 斜率和

设直线与圆锥曲线交于 A,B 两点,求曲线在 A,B 处的切线斜率之和. 用韦达定理表示 x_1+x_2,x_1x_2,代入斜率之和的表达式,消参后通常得到定值.

知乎@余命数

13.20.2 斜率差

斜率差的处理方法与斜率和类似,用韦达定理代入后消参.

13.20.3 斜率积

斜率积是圆锥曲线第三定义的延伸:椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上一点 P 与两个顶点 $A_1(-a,0)$、$A_2(a,0)$ 连线的斜率之积为 $-\frac{b^2}{a^2}$. 一般情形下,斜率积为定值或与某个参数相关.

13.20.4 斜率商

斜率商问题把斜率之比作为目标,用韦达定理代入后,消参得到定值或范围.

例题 13.293 (斜率和问题) (2024·山东淄博·二模) 已知椭圆 $\Gamma:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的离心率为 $\frac{\sqrt{3}}{2}$,且四个顶点所围成的菱形的面积为 4. (1) 求椭圆 Γ 的方程;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.294 (斜率差问题) 已知椭圆 $M:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的离心率为 $\frac{1}{2}$, A,B,C 分别为椭圆的左顶点、上顶点和右顶点, F_1 为左焦点, 且 $\triangle ABF_1$ 的面积为 $\frac{\sqrt{3}}{2}$. 若 P 是椭圆 M 上不与顶点重合的动点, 直线 AB 与直线 CP 交于点 Q, 直线 BP 交 x 轴于点 N. (1) 求椭圆 M 的标准方程; (2) 求证: $2k_{QN}-k_{QC}$ 为定值, 并求出此定值, 其中 k_{QN},k_{QC} 分别为直线 QN 和直线 QC 的斜率.

解 答案:(1) $\frac{x^2}{4}+\frac{y^2}{3}=1$;(2) $\frac{\sqrt{3}}{2}$.

(1) 由题意得 $\begin{cases}e=\frac{c}{a}=\frac{1}{2}\\ \frac{1}{2}(a-c)b=\frac{\sqrt{3}}{2}\end{cases}$, 又 $c^2=a^2-b^2$, 解得 $\begin{cases}a=2\\ c=1\\ b=\sqrt{3}\end{cases}$, 所以椭圆 M 的标准方程为 $\frac{x^2}{4}+\frac{y^2}{3}=1$.

(2) 方法一: 直线 $AB:y=\frac{\sqrt{3}}{2}(x+2)$, 设直线 $PC:x=my+2$ ($m\neq 0$ 且 $m\neq\pm\frac{2\sqrt{3}}{3}$, 注:$P$ 不为椭圆顶点). 由

$$\begin{cases}x=my+2\\ y=\frac{\sqrt{3}}{2}(x+2)\end{cases}$$

得 $y_Q=\frac{4\sqrt{3}}{2-\sqrt{3}m}$, 则 $x_Q=my_Q+2=\frac{4+2\sqrt{3}m}{2-\sqrt{3}m}$, 所以 $Q\left(\frac{4+2\sqrt{3}m}{2-\sqrt{3}m},\frac{4\sqrt{3}}{2-\sqrt{3}m}\right)$. 由

$$\begin{cases}x=my+2\\ 3x^2+4y^2-12=0\end{cases}$$

得 $(3m^2+4)y^2+12my=0\Rightarrow y_P=\frac{-12m}{3m^2+4}$, $x_P=my_P+2=\frac{-6m^2+8}{3m^2+4}$, 所以 $P\left(\frac{-6m^2+8}{3m^2+4},\frac{-12m}{3m^2+4}\right)$. 由 B,P,N 三点共线得 $k_{BP}=k_{BN}$, 即 $\frac{y_P-\sqrt{3}}{x_P}=-\frac{\sqrt{3}}{x_N}$, 解得 $x_N=\frac{2(2-\sqrt{3}m)}{\sqrt{3}m+2}$, 故 $k_{QN}=\frac{\sqrt{3}m+2}{4m}$. 于是

$$2k_{QN}-k_{QC}=\frac{\sqrt{3}m+2}{2m}-\frac{1}{m}=\frac{\sqrt{3}}{2}$$

为定值.

例题 13.295 (斜率积问题) (2024·河北保定·三模) 设椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ $(0<b<\sqrt{7})$ 的左、右顶点和椭圆 $F:\frac{x^2}{14}+\frac{y^2}{7}=1$ 的左、右焦点均为 E,F.P 是 C 上异于 E,F 的一个动点. 已知直线 EP 交直线 $l_1:x=\sqrt{7}$ 于点 A, 直线 FP 交直线 $l_2:x=-\sqrt{7}$ 于点 B. 直线 AB 与椭圆 F 交于点 M,N, O 为坐

知乎@余命数
知乎@余命数
知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

标原点. (1) 若 b 为定值, 证明: $\overrightarrow{OA}\cdot\overrightarrow{OB}$ 为定值; (2) 若直线 OM,ON 的斜率之积恒为 $-\frac{1}{2}$, 求 b.

解 (1) 证明: 易知 $E(-\sqrt{7},0),F(\sqrt{7},0)$, 设 $P(m,n)$, 则 $\frac{m^2}{7}+\frac{n^2}{b^2}=1$, 即 $m^2=7-\frac{7n^2}{b^2}$. 直线 $PE:y=\frac{n}{m+\sqrt{7}}(x+\sqrt{7})$, 联立 $l_1:x=\sqrt{7}$, 则 $A\left(\sqrt{7},\frac{2\sqrt{7}n}{m+\sqrt{7}}\right)$, 所以 $\overrightarrow{OA}=\left(\sqrt{7},\frac{2\sqrt{7}n}{m+\sqrt{7}}\right)$. 直线 $PF:y=\frac{n}{m-\sqrt{7}}(x-\sqrt{7}),l_2:x=-\sqrt{7}$, 则 $B\left(-\sqrt{7},-\frac{2\sqrt{7}n}{m-\sqrt{7}}\right)$. 所以 $\overrightarrow{OB}=\left(-\sqrt{7},-\frac{2\sqrt{7}n}{m-\sqrt{7}}\right)$, 所以 $\overrightarrow{OA}\cdot\overrightarrow{OB}=-7-\frac{28n^2}{m^2-7}=-7+4b^2$;

(2) 求直线 $AB:y=kx+t$, 则 $B(-\sqrt{7},t-\sqrt{7}k)$, $A(\sqrt{7},t+\sqrt{7}k)$, 则 $\overrightarrow{OA}\cdot\overrightarrow{OB}=-7+t^2-7k^2=-7+4b^2$, 即 $t^2-7k^2=4b^2$, 令直线 AB 与椭圆 F 联立 $\begin{cases}y=kx+t,\\ \frac{x^2}{14}+\frac{y^2}{7}=1,\end{cases}$ 消去 y, 整理得 $(2k^2+1)x^2+4ktx+2t^2-14=0$, 需满足 $\Delta>0$, 设 $M(x_1,y_1),N(x_2,y_2)$, 则 $x_1+x_2=-\frac{4kt}{2k^2+1},x_1x_2=\frac{2t^2-14}{2k^2+1}$,

知乎@余命数

例题 13.296 (斜率商问题) 已知 $\odot O:x^2+y^2=4$, 圆心 O 是原点, 点 $B(1,0)$, 以线段 AB 为直径的圆内切于 $\odot O$, 动点 A 的轨迹记为曲线 C. (1) 求曲线 C 的方程; (2) 若直线 $l:x=4$, 点 $Q(1,\frac{3}{2})$, 直线 m 过点 B 与曲线 C 交于 M,N 两点, 与直线 l 交于点 K. (1) 若 $|MN|=\frac{10}{3}$, 求直线 m 的斜率; (2) 若记直线 QM,QN,QK 的斜率分别为 k_1,k_2,k_3, 问 $\frac{k_1-k_2}{k_2-k_3}$ 是否为定值? 如果是, 请求出定值; 如果不是, 请说明理由.

解 (1) 设 AB 的中点为 P, 切点为 T, 连接 OP,PT, 取 B 关于 y 轴的对称点 D, 连接 AD, 则 $|AD|=2|OP|$, 故 $|AB|+|AD|=2|OP|+2|PB|=2|OP|+2|PT|=2(|OP|+|PT|)=4>|BD|=2$, 所以点 A 的轨迹是以 B,D 为焦点, 长轴长为 4 的椭圆, 其中 $a=2,c=1$, 则 $b=\sqrt{3}$, 则曲线 C 的方程为 $\frac{x^2}{4}+\frac{y^2}{3}=1$; (2) 设 $M(x_1,y_1),N(x_2,y_2)$, 依题意, 直线 m 的斜率必定存在, 设 $l_1:x=ty+1(t\neq0)$, $\begin{cases}x=ty+1\\ \frac{x^2}{4}+\frac{y^2}{3}=1\end{cases}$, 可得 $(3t^2+4)y^2+6ty-9=0$, $\Delta>0$ 恒成立, 则有 $y_1+y_2=\frac{-6t}{3t^2+4}$, $y_1y_2=\frac{-9}{3t^2+4}$, (1) 若 $|MN|=\frac{10}{3}$, 则有 $|MN|=\sqrt{1+t^2}|y_1-y_2|=\sqrt{1+t^2}\sqrt{(y_1+y_2)^2-4y_1y_2}=\frac{12(t^2+1)}{3t^2+4}=\frac{10}{3}$,

例题 13.297 积为 4. (1) 求椭圆的标准方程; (2) 四边形 $ABCD$ 的顶点在圆上, 且对角线 AC,BD 过原点 O, 设 $A(x_1,y_1),B(x_2,y_2)$, 满足 $x_1x_2=4y_1y_2$. (1) 求证: 直线 AB 和直线 BC 的斜率之和为定值; (2) 求四边形 $ABCD$ 面积的最大值.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

解 答案:(1) $\frac{x^2}{4}+y^2=1$;(2)(1) 0;(2) 4.

(1) 由题意 $e=\frac{c}{a}=\frac{\sqrt{3}}{2}, 2ab=4$, 又 $a^2=b^2+c^2$, 解得 $a^2=4, b^2=1$, 所以椭圆的标准方程为 $\frac{x^2}{4}+y^2=1$.

(2) 如图所示.

(1) 设直线 AB 的方程为 $y=kx+m$, 设 $A(x_1,y_1), B(x_2,y_2)$, 联立

$$\begin{cases} y=kx+m \\ x^2+4y^2=4 \end{cases}$$

得 $(1+4k^2)x^2+8kmx+4(m^2-1)=0$, $\Delta=(8km)^2-4(4k^2+1)\times 4(m^2-1)=16(4k^2-m^2+1)>0$, 则

$$\begin{cases} x_1+x_2=\frac{-8km}{1+4k^2} \\ x_1x_2=\frac{4(m^2-1)}{1+4k^2} \end{cases}$$

知乎@余命数

知乎@余命数

$$y_1y_2=(kx_1+m)(kx_2+m)=k^2x_1x_2+km(x_1+x_2)+m^2=\frac{m^2-4k^2}{1+4k^2}.$$

$\because x_1x_2=4y_1y_2, \therefore \frac{4(m^2-1)}{1+4k^2}=4\times\frac{m^2-4k^2}{1+4k^2}$, 整理得 $4k^2=1, \therefore k=\pm\frac{1}{2}$. 又

知乎@余命数

$$k_{BC}=\frac{y_2+y_1}{x_2+x_1}=k+\frac{2m}{x_2+x_1}=k+\frac{2m}{\frac{-8km}{1+4k^2}}=-\frac{1}{4k},$$

$\therefore k_{AB}+k_{BC}=k-\frac{1}{4k}+\frac{4k^2-1}{4k}=0$, 所以直线 AB 和直线 BC 的斜率之和为定值 0.

(2) 由 (1), 不妨取 $k_{AB}=-\frac{1}{2}$, 则 $x_1+x_2=2m, x_1x_2=2(m^2-1)$. 设原点到直线 AB 的距离为 d, 则

知乎@余命数

$$\begin{aligned} S_{\triangle AOB}&=\frac{1}{2}|AB|\cdot d=\frac{1}{2}\sqrt{1+k^2}|x_2-x_1|\cdot\frac{|m|}{\sqrt{1+k^2}} \\ &=\frac{|m|}{2}\sqrt{(x_1+x_2)^2-4x_1x_2}=\frac{|m|}{2}\sqrt{4m^2-4\cdot 2(m^2-1)}=\sqrt{m^2(2-m^2)}. \end{aligned}$$

又 $\sqrt{m^2(2-m^2)}\le\frac{m^2+2-m^2}{2}=1$, 所以 $S_{\triangle AOB}\le 1$, 当且仅当 $m^2=1$ 时取等号.$S_{四边形ABCD}=4S_{\triangle AOB}\le 4$, 即四边形 $ABCD$ 的面积的最大值为 4.

例题 13.298 (1) 求椭圆 Γ 的方程;

解 答案: $\frac{x^2}{4}+y^2=1$. 由条件 $e=\frac{c}{a}=\frac{\sqrt{3}}{2}$, 四个顶点所围成的菱形面积为 $\frac{1}{2}\cdot 2a\cdot 2b=2ab=4$, 又 $a^2=b^2+c^2$, 解得 $a=2$、$b=1$、$c=\sqrt{3}$, 故椭圆的方程为 $\frac{x^2}{4}+y^2=1$;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.299　(2) 已知直线 $x=1$ 与 x 轴交于点 M, 过点 M 的直线 AB 与 Γ 交于 A,B 两点, 点 P 为直线 $x=1$ 上任意一点, 设直线 AB 与直线 $x=4$ 交于点 N, 记 PA,PB,PN 的斜率分别为 k_1,k_2,k_0, 求证: $k_1+k_2=2k_0$.

解　答案: $k_1+k_2=2k_0$.

设 $A(x_1,y_1),B(x_2,y_2),P(1,t)$, 若直线 AB 与 x 轴不重合时, 设直线 AB 的方程为 $x=my+1$, 点 $N\left(4,\frac{3}{m}\right)$, $k_0=\frac{t-\frac{3}{m}}{-3}$. 代入椭圆方程整理得 $(m^2+4)y^2+2my-3=0$, 显然 $\Delta>0$, 则 $y_1+y_2=-\frac{2m}{m^2+4}$, $y_1y_2=-\frac{3}{m^2+4}$, 所以

$$\begin{aligned}k_1+k_2&=\frac{t-y_1}{1-x_1}+\frac{t-y_2}{1-x_2}=\frac{(t-y_1)(1-x_2)+(t-y_2)(1-x_1)}{(1-x_1)(1-x_2)}\\&=\frac{(t-y_1)(-my_2)+(t-y_2)(-my_1)}{(-my_1)(-my_2)}=\frac{-t(y_1+y_2)+2y_1y_2}{my_1y_2}\\&=\frac{-t\cdot\frac{-2m}{m^2+4}+2\cdot\frac{-3}{m^2+4}}{m\cdot\left(\frac{-3}{m^2+4}\right)}=\frac{2mt-6}{-3m}=2\cdot\frac{t-\frac{3}{m}}{-3}=2k_0.\end{aligned}$$

若直线 AB 与 x 轴重合时, 则 $B(-2,0),A(2,0),N(4,0)$, 此时 $k_1+k_2=\frac{t}{3}+\frac{t}{-1}=-\frac{2}{3}t$, 而 $k_0=-\frac{1}{3}t$, 故 $k_1+k_2=2k_0$.

例题 13.300　(2) 若点 P 为 x 轴正半轴上的任意一点, 过点 P 作直线交抛物线于 A,B 两点, 点 P 关于原点的对称点 M, 连接 MA 交抛物线于点 N, 求证: $k_{AP}+k_{NP}=0$.

解　答案: $k_{AP}+k_{NP}=0$.

(1) 设直线 l 方程为 $y=\frac{4}{3}(x-\frac{p}{2})$, $G(x_1,y_1),H(x_2,y_2)$, 由抛物线的性质可知 $|FG|+|FH|=x_1+x_2+p=\frac{25}{4}$, $\therefore x_1+x_2=\frac{25}{4}-p$. 联立

$$\begin{cases}y=\frac{4}{3}(x-\frac{p}{2})\\y^2=2px\end{cases}$$

消 y 得 $16x^2-34px+4p^2=0$、$\Delta=(34p)^2-64\times4p^2=900p^2>0$、$\therefore x_1+x_2=\frac{34p}{16}=\frac{25}{4}-p$, 解得 $p=2$, $\therefore$ 抛物线 C 的方程为 $y^2=4x$, 焦点 $F(2,0)$.

(2) 设 $P(m,0)$, 则 $M(-m,0)$、$A(x_3,y_3)$、$N(x_4,y_4)$, 直线 AM 的方程为 $x=ty-m$. 联立

$$\begin{cases}x=ty-m\\y^2=2px\end{cases}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

消 x 得 $y^2-2pty+2pm=0$、$\Delta=4p^2t^2-8pm>0$、$\therefore y_3+y_4=2pt, y_3y_4=2pm$. 而

$$k_{AP}+k_{NP}=\frac{y_3}{x_3-m}+\frac{y_4}{x_4-m}=\frac{x_3y_4+x_4y_3-m(y_3+y_4)}{x_3x_4-m(x_3+x_4)+m^2},$$

又知

$$x_3y_4+x_4y_3-m(y_3+y_4)=(ty_3-m)y_4+(ty_4-m)y_3-m(y_3+y_4)=2ty_3y_4-2m(y_3+y_4)=4ptm-4ptm=$$

所以 $k_{AP}+k_{NP}=0$.

例题 13.301 (2024·江西鹰潭·二模)设椭圆 $E:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 经过点 $P(2,-1)$, 且离心率 $e=\frac{\sqrt{2}}{2}$, 直线 $m:x=3$ 垂直 x 轴交 x 轴于 T, 过 T 的直线 l 交椭圆 E 于 $A(x_1,y_1)$, $B(x_2,y_2)$ 两点, 连接 PA,PB,PT.(1) 求椭圆 E 的方程;

知乎@余命数

解 答案: $\frac{x^2}{6}+\frac{y^2}{3}=1$. 由题意知 $\begin{cases}\frac{4}{a^2}+\frac{1}{b^2}=1\\ \frac{c}{a}=\frac{\sqrt{2}}{2}\\ a^2=b^2+c^2\end{cases}$ 解得 $a=\sqrt{6},b=c=\sqrt{3}$ 所以椭圆 E 的方程为 $\frac{x^2}{6}+\frac{y^2}{3}=1$;

例题 13.302 (2) 设直线 PA,PB 的斜率分别为 k_1,k_2.(i) 求 k_1+k_2 的值;

解 答案:2.

易知 $T(3,0),k_{PT}=1,k_1=\frac{y_1+1}{x_1-2},k_2=\frac{y_2+1}{x_2-2}$, 设直线 l 的方程为 $m(x-2)+n(y+1)=1$, 由直线 l 过 $T(3,0)$ 知 $m+n=1$. 联立方程

$$\begin{cases}\frac{x^2}{6}+\frac{y^2}{3}=1\\ m(x-2)+n(y+1)=1\end{cases}$$

得 $(2-4n)(y+1)^2+(4n-4m)(x-2)(y+1)+(1+4m)(x-2)^2=0$, 变形得

$$(2-4n)\left(\frac{y+1}{x-2}\right)^2+(4n-4m)\frac{y+1}{x-2}+1+4m=0,$$

即 $k_1+k_2=\frac{4n-4m}{4n-2}=\frac{4n-4(1-n)}{4n-2}=\frac{8n-4}{4n-2}=2$.

例题 13.303 (2024·重庆渝中·模拟预测)已知椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\ (a>b>0)$ 的离心率为 $\frac{1}{2}$, 点 $A\left(1,\frac{3}{2}\right)$ 在 C 上. (1) 求椭圆 C 的方程; (2) 过点 $T(4,0)$ 的

直线 l 交圆 C 于异于点 A 的 P,Q 两点,过点 P 作 x 轴的垂线与直线 AQ 交于点 M,设直线 AP,AQ 的率分别为 k_1,k_2. 证明:(i) k_1+k_2 为定值;(ii) 直线 AT 过线段 PM 的中点.

解 答案:$\frac{x^2}{4}+\frac{y^2}{3}=1$;$k_1+k_2=-1$;直线 AT 过 PM 中点.

(1) 由题可知 $\begin{cases}\frac{c}{a}=\frac{1}{2}\\ a^2=b^2+c^2\\ \frac{1}{a^2}+\frac{\left(\frac{3}{2}\right)^2}{b^2}=1\end{cases}$,解得 $\begin{cases}a=2\\ b=\sqrt{3}\end{cases}$,所以椭圆 C 的方程为 $\frac{x^2}{4}+\frac{y^2}{3}=1$.

(2)(i) (1) 当直线 l 的斜率为 0 时, 则不妨设 $P(2,0),Q(-2,0)$, 所以 $k_1+k_2=\frac{\frac{3}{2}}{1-2}+\frac{\frac{3}{2}}{1-(-2)}=-\frac{3}{2}+\frac{1}{2}=-1$ 为定值.

(2) 当直线 l 的斜率不为 0 时,设直线 $l:x=my+4\,(m\neq 0)$,$P(x_1,y_1),Q(x_2,y_2)$,联立直线 l 与椭圆 C 的方程

$$\begin{cases}x=my+4\\ \frac{x^2}{4}+\frac{y^2}{3}=1\end{cases}$$

消去 x 整理得 $(3m^2+4)y^2+24my+36=0$, 则 $\Delta>0$、$y_1+y_2=-\frac{24m}{3m^2+4}$、$y_1y_2=\frac{36}{3m^2+4}$,所以 $my_1y_2=-\frac{3}{2}(y_1+y_2)$. 于是

$$k_1+k_2=\frac{y_1-\frac{3}{2}}{x_1-1}+\frac{y_2-\frac{3}{2}}{x_2-1}=\cdots=-1.$$

综上, $k_1+k_2=-1$ 为定值.

(ii) 设线段 PM 的中点为 $N(x_N,y_N)$, $M(x_1,y_M)$, 易得 $\begin{cases}x_N=x_1\\ y_N=\frac{y_1+y_M}{2}\end{cases}$, 可得直线 AQ 的方程为 $y=k_2(x-1)+\frac{3}{2}$,则 $y_M=k_2(x_1-1)+\frac{3}{2}$. 又 $y_1=k_1(x_1-1)+\frac{3}{2}$ 且 $k_1+k_2=-1$, 故

$$y_N=\frac{1}{2}\left[k_1(x_1-1)+\frac{3}{2}+k_2(x_1-1)+\frac{3}{2}\right]=\frac{1}{2}\left[-(x_1-1)+3\right]=-\frac{x_1}{2}+2=-\frac{x_N}{2}+2,$$

即 N 在直线 $AT:y=-\frac{1}{2}x+2$ 上,因此直线 AT 过线段 PM 的中点.

例题 13.304　(2024·四川成都·模拟预测) 已知抛物线 $C:x^2=2py(p>1)$ 的焦点为 F, 过点 $P(1,-1)$ 作抛物线 C 的两条切线, 切点分别为 M,N, $|FM|+|FN|=5$. (1) 求抛物线 C 的方程; (2) 过点 P 作两条倾斜角互补的直线 l_1,l_2, l_1 交抛物线 C 于 A,B 两点, l_2 交抛物线 C 于 C,D 两点, 连接

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

AD, BC, AC, BD, 设 AC, BD 的斜率分别为 k_1, k_2, 求 k_1+k_2 的值; (3) 设 $\angle DBC=\lambda\angle DAC$, 求 λ 的值.

解 答案: (1) $x^2=4y$; (2) 0. (1) 设切点 $M\left(x_1, \frac{x_1^2}{2p}\right), N\left(x_2, \frac{x_2^2}{2p}\right)$, 因为 $x^2=2py$, 所以 $y=\frac{x^2}{2p}$, $y'=\frac{x}{p}$. 以点 M 为切点的切线的斜率为 $\frac{x_1}{p}$, 以点 M 为切点的切线的方程为 $y-\frac{x_1^2}{2p}=\frac{x_1}{p}(x-x_1)$, $\therefore$ 切线过点 $P(1,-1)$, 所以 $-1-\frac{x_1^2}{2p}=\frac{x_1}{p}(1-x_1)$, $\therefore x_1^2-2x_1-2p=0$, 同理 $x_2^2-2x_2-2p=0$, 所以 x_1, x_2 为方程 $x^2-2x-2p=0$ 的两根, $\therefore|FM|+|FN|=\frac{x_1^2}{2p}+\frac{p}{2}+\frac{x_2^2}{2p}+\frac{p}{2}=\frac{(x_1+x_2)^2-2x_1x_2}{2p}+p$、$\therefore p^2-3p+2=0$、$(p-1)(p-2)=0$、$\because p>1$、$\therefore p=2$、$\therefore$ 抛物线方程为 $x^2=4y$. (2) 设 l_1 方程为 $y+1=k(x-1)$,联立 l_1 和抛物线方程,得 $\begin{cases}y=kx-(k+1)\\x^2=4y\end{cases}$、$\therefore x^2-4kx+4(k+1)=0$、$\Delta=(-4k)^2-16(k+1)=16(k^2-k-1)>0$, 解得 $\frac{1-\sqrt{5}}{2}<k<\frac{1+\sqrt{5}}{2}$, 设 $A(x_A, y_A), B(x_B, y_B), C(x_C, y_C), D(x_D, y_D)$, $\therefore x_A+x_B=4k$, 同理 $x_C+x_D=-4k$.$\therefore k_{AC}+k_{BD}=k_1+k_2=\frac{y_C-y_A}{x_C-x_A}+\frac{y_D-y_B}{x_D-x_B}=\frac{x_C^2-x_A^2}{4}\cdot\frac{1}{x_C-x_A}+\frac{x_D^2-x_B^2}{4}\cdot\frac{1}{x_D-x_B}$, $\therefore k_{AC}+k_{BD}=k_1+k_2=\frac{x_C+x_A}{4}+\frac{x_D+x_B}{4}=\frac{(x_A+x_B)+(x_C+x_D)}{4}=0$. (3) $|PA|\cdot|PB|=\sqrt{1+k^2}|x_A-1|\cdot|x_B-1|$,

例题 13.305 在平面直角坐标系 xOy 中, 已知定点 $A(1,0)$, 点 M 在 x 轴上运动, 点 N 在 y 轴上运动, 点 P 为坐标平面内的动点, 且满足 $\overrightarrow{PM}\cdot\overrightarrow{NA}=0$, $\overrightarrow{OM}=2\overrightarrow{ON}+\overrightarrow{PO}$. (1) 求动点 P 的轨迹 C 的方程;(2) 点 Q 为圆 $(x+2)^2+y^2=1$ 上一点, 由 Q 向 C 引切线, 切点分别为 S、T, 记 k_1, k_2 分别为切线 QS, QT 的斜率, 当 Q 运动时, 求 $\left|\frac{1}{k_1}-\frac{1}{k_2}\right|$ 的取值范围.

解 答案:(1) $y^2=4x$;(2) $[2, 2\sqrt{3}]$.

(1) 设 $N(0,b)$、$M(a,0)$、$P(x,y)$. 因为 $\overrightarrow{PM}\cdot\overrightarrow{NA}=0$,所以 $(a-x,-y)\cdot(1,-b)=0$, 即 $a-x+by=0$. 因为 $\overrightarrow{OM}=2\overrightarrow{ON}+\overrightarrow{PO}$, 所以 $(a,0)=(0,2b)+(-x,-y)$, 所以 $x=-a, y=2b$, 所以 $y^2=4x$.

(2) 设 $Q(x_0,y_0)$, $x_0\in[-3,-1]$. 由题意知切线斜率存在, 设为 k, 切线方程为 $y-y_0=k(x-x_0)$. 联立

$$\begin{cases}y-y_0=k(x-x_0)\\y^2=4x\end{cases}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

化简得 $ky^2-4y+4y_0-4kx_0=0,\Delta=16-16k(y_0-kx_0)=0$,即 $x_0k^2-y_0k+1=0$、$\therefore k_1+k_2=\frac{y_0}{x_0}$、$k_1k_2=\frac{1}{x_0}$、$\therefore$

$$\left|\frac{1}{k_1}-\frac{1}{k_2}\right|=\left|\frac{k_2-k_1}{k_1k_2}\right|=\frac{\sqrt{y_0^2-4x_0}}{|x_0|}\cdot|x_0|=\sqrt{y_0^2-4x_0}.$$

将 $y_0^2=1-(x_0+2)^2$ 代入得

$$\left|\frac{1}{k_1}-\frac{1}{k_2}\right|=\sqrt{1-(x_0+2)^2-4x_0}=\sqrt{-x_0^2-8x_0-3},\quad x_0\in[-3,-1].$$

令 $f(x)=-x^2-8x-3$, 对称轴为 $x=-4$, 在 $[-3,-1]$ 上单调递减, $f(-3)=12$、$f(-1)=4$、$\therefore$ 取值范围为 $[2,2\sqrt{3}]$.

例题 13.306　设 M、N 为抛物线 $C:y^2=2px(p>0)$ 上的两点, M 与 N 的中点的纵坐标为 4, 直线 MN 的斜率为 $\frac{1}{2}$. (1) 求抛物线 C 的方程; (2) 已知点 $P(1,2)$, A、B 为抛物线 C(除原点外)上的不同两点, 直线 PA、PB 的斜率分别为 k_1,k_2, 且满足 $\frac{1}{k_1}-\frac{1}{k_2}=2$, 记抛物线 C 在 A、B 处的切线交于点 $S(x_s,y_s)$, 线段 AB 的中点为 $E(x_E,y_E)$, 若 $y_s=\lambda y_E$, 求 λ 的值.

解　答案:(1) $y^2=4x$ (2) $\lambda=1$.

(1) 设 $M(x_1,y_1),N(x_2,y_2)$. 又 M、N 都在抛物线 C 上,即 $y_1^2=2px_1,y_2^2=2px_2$. 由两式相减得 $(y_1+y_2)(y_1-y_2)=2p(x_1-x_2),\because$ 直线 MN 的斜率为 $\frac{1}{2},\therefore\frac{y_1-y_2}{x_1-x_2}=\frac{1}{2}$. 两边同除以 x_1-x_2,且由已知得 $y_1+y_2=8$,所以 $8\cdot\frac{1}{2}=2p$,即 $p=2$,所以抛物线 C 的方程为 $y^2=4x$.

(2) 设 $A(\frac{y_3^2}{4},y_3)$、$B(\frac{y_4^2}{4},y_4)$、$S(x_s,y_s)$. 因为

$$\frac{1}{k_1}-\frac{1}{k_2}=\frac{\frac{y_3^2}{4}-1}{y_3-2}-\frac{\frac{y_4^2}{4}-1}{y_4-2}=\frac{y_3-y_4}{4},$$

所以 $\frac{y_3-y_4}{4}=2$, 所以 $y_3-y_4=8$. 设直线 SA 的斜率为 k, 则直线 $SA:y-y_3=k(x-\frac{y_3^2}{4})$, 由

$$\begin{cases}y-y_3=k(x-\frac{y_3^2}{4})\\ y^2=4x\end{cases}$$

消 x 得 $\frac{ky^2}{4}-y+y_3-\frac{y_3^2}{4}=0$. 由 $\Delta=0$, 得 $(\frac{ky_3}{2}-1)^2=0$, 即 $k=\frac{2}{y_3}$. 所以直线 $SA:y-y_3=\frac{2}{y_3}(x-\frac{y_3^2}{4})$, 同理得直线 $SB:y-y_4=\frac{2}{y_4}(x-\frac{y_4^2}{4})$. 联立以上两个方程解得 $y_s=\frac{y_3+y_4}{2}$, 又 $y_E=\frac{y_3+y_4}{2}$, 所以 $y_s=y_E$, 所以 $\lambda=1$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 13.307 (3) 若过 A_1, A_2 作抛物线 C 的切线 PA_1, PA_2, 交点 P 在直线 $x=-1$ 上, 求 $\triangle A_1OA_2$ 面积的最小值.

解 答案: $\sqrt{3}$. (1) 设 $A_1(x_1,y_1), A_2(x_2,y_2)$, 由题意知, 设直线 $A_1A_2: x=my+t(t>0, m>0)$, 又直线 A_1A_2 与 x 正半轴交于 A_3, 与 y 负半轴交于 A_4, 则 $A_3(t,0), A_4\left(0,-\frac{t}{m}\right)$. 联立 $\begin{cases} y^2=2x \\ x=my+t \end{cases}$, 整理得 $y^2-2my-2t=0$, 所以 $y_1+y_2=2m, y_1y_2=-2t$, 则 $x_1+x_2=my_1+t+my_2+t=m(y_1+y_2)+2t=2m^2+2t$, $x_1x_2=\frac{1}{4}(y_1y_2)^2=t^2$, 若 $\angle A_1OA_2 \geq 90^\circ$, 则 $\overrightarrow{OA_1}\cdot\overrightarrow{OA_2}=x_1x_2+y_1y_2=t^2-2t<0$, 解得 $t\in(0,2)$, 又 $A_3(t,0)$, 则 A_3 横坐标的取值范围为 $(0,2)$;

(2) 因为 G 为重心, 所以 $G\left(\frac{x_1+x_2}{3}, \frac{y_1+y_2}{3}\right)$, 由 (1) 可得 $G\left(\frac{2m^2+2t}{3}, \frac{2m}{3}\right)$, 设 $A_3(x_3,y_3), A_4(x_4,y_4)$, 又直线 A_1A_2 与 x 正半轴交于 A_3, 与 y 负半轴交于 A_4, 则 $y_3=0, x_4=0$, 由直线 $A_1A_2: x=my+t(t>0, m>0)$, 则 $k_1=\frac{1}{m}, A_3(t,0), A_4\left(0,-\frac{t}{m}\right)$, 所以 $k_2=\frac{\frac{2m}{3}-0}{\frac{2m^2+2t}{3}-t}=\frac{2m}{2m^2-t}$, 由 $k_2=2k_1$, 则 $\frac{2m}{2m^2-t}=\frac{2}{m}$, 即 $m^2=t$, 又 $\lambda=\frac{|A_1A_2|}{|A_3A_4|}=\frac{\sqrt{(y_1-y_2)^2-4y_1y_2}}{y_3-y_4}=\frac{\sqrt{(y_1+y_2)^2-4y_1y_2}}{0-(-\frac{t}{m})}=\frac{\sqrt{(2m)^2-4(-2t)}}{0-(-\frac{t}{m})}=\frac{m\sqrt{4m^2+8t}}{t}=\frac{m\sqrt{4m^2+8m^2}}{m^2}=2\sqrt{3}$, 因此 $|A_1A_2|=|A_3A_4|$ 时, λ 为定值.

(3) 设过点 $M(x_0,y_0)(x_0eq0, y_0eq0)$ 的切线方程为 $y-y_0=k(x-x_0)$, 则联立方程 $\begin{cases} y-y_0=k(x-x_0) \\ y^2=2x \end{cases}$, 化简可得 $ky^2-2y+2y_0-2kx_0=0$, 因为直线与抛物线相切, 则 $\Delta=4-4k(2y_0-2kx_0)=0$, 得 $2kx_0k^2-2y_0k+1=0$, 而 $M(x_0,y_0)$ 为抛物线上一点, 则 $y_0^2=2x_0$, 代入可得 $y_0^2k^2-2y_0k+1=0$, 得 $k=\frac{1}{y_0}$, $y-y_0=\frac{1}{y_0}(x-x_0)$, 则 $y\cdot y_0=x-x_0+y_0^2$, 即 $y_0y=x+x_0$, 即过点 M 的切线方程为 $y_0y=x+x_0$. 因此过 $A_1(x_1,y_1)$ 的切线 $PA_1: y_1y=x+x_1$, 过 $A_2(x_2,y_2)$ 的切线 $PA_2: y_2y=x+x_2$, 又切线 PA_1 与切线 PA_2 的交点 P 在直线 $x=-1$ 上, 可设 $P(-1,n)$,

例题 13.308 如图, 已知椭圆 $C: \frac{x^2}{6}+\frac{y^2}{4}=1$ 与顶点 $A(0,-2)$, 经过点 $E(0,1)$ 且斜率存在的直线 l 交椭圆于 Q, N 两点, 点 B 与点 Q 关于坐标原点对称, 连接 AB, AN. 求证: 存在实数 λ 使得 $k_{AN}=\lambda k_{AB}$ 恒成立.

解 答案: $\lambda=3$.

设 $y=kx+1$, 由

$$\begin{cases} y=kx+1, \\ 2x^2+3y^2-12=0, \end{cases}$$

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

可得 $(2+3k^2)x^2+6kx-9=0$. 设 $N(x_1,y_1),Q(x_2,y_2)$, 则

$$\begin{cases}x_1+x_2=-\frac{6k}{2+3k^2},\\ x_1x_2=-\frac{9}{2+3k^2}.\end{cases}$$

知乎@余命数

$x_1y_2+x_2y_1=x_1(kx_2+1)+x_2(kx_1+1)=2kx_1x_2+x_1+x_2$、$2kx_1x_2=3(x_1+x_2)$、$\therefore x_1y_2+x_2y_1=4(x_1+x_2)$. 又 N,E,Q 三点共线, 则 $\frac{y_2-1}{x_2}=\frac{y_1-1}{x_1}$, 即 $x_2y_1-x_1y_2=x_2-x_1$, $\therefore\begin{cases}x_1y_2=\frac{5}{2}x_1+\frac{3}{2}x_2,\\ x_2y_1=\frac{3}{2}x_1+\frac{5}{2}x_2.\end{cases}$ $\therefore A(0,-2),B(-x_2,-y_2)$, 则

$$\lambda=\frac{k_{AN}}{k_{AB}}=\frac{(y_1+2)x_2}{(y_2-2)x_1}=\frac{x_2y_1+2x_2}{x_1y_2-2x_1}=\frac{\frac{3}{2}x_1+\frac{9}{2}x_2}{\frac{1}{2}x_1+\frac{3}{2}x_2}=3,$$

知乎@余命数

$\therefore$ 存在实数 $\lambda=3$, 使得 $k_{AN}=3k_{AB}$ 恒成立.

例题 13.309　(2024·全国·模拟预测) 已知抛物线 $C:y^2=2px(p>0)$ 的焦点 F 是椭圆 $E:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的右焦点, 抛物线 C 与圆 E 在第一象限的公共点 P 的横坐标为 $\frac{2}{3}$, $|PF|=\frac{5}{3}$.
(1) 求抛物线 C 与椭圆 E 的标准方程;
(2) 若 A_1,A_2 分别是椭圆 E 的左、右顶点, M,N 是椭圆 E 上不同于 A_1,A_2 的两点, 直线 A_1M 的斜率是直线 A_2N 的斜率的 3 倍, 证明: 直线 MN 过定点, 并求出定点的坐标.

解　答案: (1) $y^2=4x$, $\frac{x^2}{4}+\frac{y^2}{3}=1$; (2) $(1,0)$. (1) 由抛物线的定义知 $|PF|=\frac{p}{2}+\frac{2}{3}=\frac{5}{3}$, $\therefore p=2$, $\therefore$ 抛物线 C 的标准方程为 $y^2=4x$, $\because$ 抛物线 C 的标准方程为 $y^2=4x$,

例题 13.310　已知椭圆 C: $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的右顶点为 $A(3,0)$, 离心率为 $\frac{\sqrt{5}}{3}$, 过点 $P(3,2)$ 的直线 l 与 C 交于 M,N 两点. (1) 若 C 的上顶点为 B, 直线 BM,BN 的斜率分别为 k_1,k_2, 求 $\frac{1}{k_1}+\frac{1}{k_2}$ 的值; (2) 过点 M 且垂直于 x 轴的直线交直线 AN 于点 Q, 证明: 线段 MQ 的中点在定直线上.

知乎@余命数

解　答案: (1) -3 (2) 证明见解析. (1) 由题意知 $\begin{cases}a=3\\ \frac{c}{a}=\frac{\sqrt{5}}{3}\\ c^2=a^2-b^2\end{cases}$, 解得 $a=3$、$b=2$、$c=\sqrt{5}$, 所以 C 的方程为 $\frac{x^2}{9}+\frac{y^2}{4}=1$, 显然直线 l 的斜率存在, 设直线 l 的方程为: $y=k(x-3)+2$、$M(x_1,y_1)$、$N(x_2,y_2)$, 由 $\begin{cases}\frac{x^2}{9}+\frac{y^2}{4}=1\\ y=k(x-3)+2\end{cases}$, 得 $(4+9k^2)x^2+$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$18k(2-3k)x+81k^2-108k=0$,由方程 $(4+9k^2)x^2+18k(2-3k)x+81k^2-108k=0$ 的判别式 $\Delta>0$,可得 $k>0$,所以 $x_1+x_2=-\frac{18k(2-3k)}{4+9k^2}$,$x_1x_2=\frac{81k^2-108k}{4+9k^2}$,易得 $B(0,2)$,所以 $k_1=\frac{y_1-2}{x_1}=\frac{kx_1-3k}{x_1}$,$k_2=\frac{y_2-2}{x_2}=\frac{kx_2-3k}{x_2}$,所以 $\frac{1}{k_1}+\frac{1}{k_2}=\frac{x_1}{kx_1-3k}+\frac{x_2}{kx_2-3k}=\frac{x_1(x_2-3)+x_2(x_1-3)}{k(x_1-3)(x_2-3)}=\frac{2x_1x_2-3(x_1+x_2)}{k(x_1x_2-3(x_1+x_2)+9)}=\frac{2\cdot\frac{81k^2-108k}{4+9k^2}+\frac{54k(2-3k)}{4+9k^2}}{k\left(\frac{81k^2-108k}{4+9k^2}+\frac{54k(2-3k)}{4+9k^2}+9\right)}=-3$,(2) 证明:设线段 MQ 的中点为 (x_0,y_0),又 $M(x_1,y_1)$,$N(x_2,y_2)$,

例题 13.311 设椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1(a>b>0)$ 的左、右顶点分别为 A,B,离心率为 e,且 $|AB|=\frac{8\sqrt{3}}{3}e=4$. (1) 求椭圆 C 的标准方程;(2) 设点 P,Q 为椭圆上异于 A,B 的两动点,记直线 AP 的斜率为 k_1,直线 QB 的斜率为 k_2,已知 $k_1=7k_2$. 直线 PQ 与 x 轴相交于点 M,求 $\triangle APM$ 的面积的最大值.

解 答案:(1) $\frac{x^2}{4}+y^2=1$;(2) $\frac{1}{4}$.

(1) 因为 $|AB|=\frac{8\sqrt{3}}{3}e=4$,所以

$$\begin{cases}\frac{8\sqrt{3}}{3}\cdot\frac{c}{a}=4,\\ 2a=4,\\ a^2=b^2+c^2,\end{cases}$$

解得 $\begin{cases}a^2=4,\\ b^2=1,\end{cases}$ 所以椭圆 C 的标准方程为 $\frac{x^2}{4}+y^2=1$.

(2) 由 $\frac{x^2}{4}+y^2=1$ 可得点 $A(-2,0),B(2,0)$,设 $P(x_1,y_1),Q(x_2,y_2)$,直线 $AP:x=t_1y-2$,直线 $BQ:x=t_2y+2$. 联立

$$\begin{cases}x=t_1y-2,\\ x^2+4y^2=4,\end{cases}$$

消去 x 得 $(t_1^2+4)y^2-4t_1y=0$,解得 $y_1=\frac{4t_1}{t_1^2+4}$. 同理联立

$$\begin{cases}x=t_2y+2,\\ x^2+4y^2=4,\end{cases}$$

消去 x 得 $(t_2^2+4)y^2+4t_2y=0$,解得 $y_2=-\frac{4t_2}{t_2^2+4}$. 因为 $k_{AP}=7k_{BQ}$,且 $k_{AP}=\frac{1}{t_1}$,$k_{BQ}=\frac{1}{t_2}$,所以 $t_2=7t_1$,此时 $y_2=-\frac{4t_2}{t_2^2+4}=-\frac{28t_1}{49t_1^2+4}$. 设 $M(x_0,0)$,由 P,M,Q 三点共线,所以 $\frac{x_1-x_0}{y_1}=\frac{x_2-x_0}{y_2}$,则

$$x_0=\frac{x_2y_1-x_1y_2}{y_1-y_2}=\frac{(t_2y_2+2)y_1-(t_1y_1-2)y_2}{y_1-y_2}=\frac{(t_2-t_1)y_1y_2+2(y_1+y_2)}{y_1-y_2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
&= \frac{(7t_1 - t_1)\left(-\frac{112t_1^2}{(t_1^2+4)(49t_1^2+4)}\right) + 2\left(\frac{4t_1}{t_1^2+4} - \frac{28t_1}{49t_1^2+4}\right)}{\frac{4t_1}{t_1^2+4} - \left(-\frac{28t_1}{49t_1^2+4}\right)} \\
&= \frac{-24\times 28t_1^3 + 8t_1(49t_1^2+4) - 56t_1(t_1^2+4)}{4t_1(49t_1^2+4) + 28t_1(t_1^2+4)} \\
&= \frac{-84t_1^2 - 48}{56t_1^2 + 32} = -\frac{12(7t_1^2+4)}{8(7t_1^2+4)} = -\frac{3}{2},
\end{aligned}
$$

所以 $M\left(-\frac{3}{2}, 0\right)$. 所以

$$
S_{\triangle APM} = \frac{1}{2}\times\frac{1}{2}|y_1| = \frac{1}{4}\cdot\frac{4|t_1|}{t_1^2+4} = \frac{1}{|t_1| + \frac{4}{|t_1|}} \leq \frac{1}{4},
$$

当且仅当 $|t_1| = \frac{4}{|t_1|}$, 即 $|t_1| = 2$ 时等号成立. 所以 $S_{\triangle APM}$ 的最大值为 $\frac{1}{4}$.

知乎@余命数

例题 13.312　已知椭圆 C: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1(a > b > 0)$ 的离心率为 $\frac{1}{2}$, 点 A, B 在椭圆上运动. 当直线 AB 过椭圆右焦点并垂直于 x 轴时, $\triangle OAB$ 的面积为 $\frac{3}{2}$ (O 为坐标原点). (1) 求椭圆 C 的标准方程; (2) 延长 OA 到 M , 使得 $OM = 3OA$, 且 MB 与椭圆 C 交于点 Q, 若直线 OA, OB 的斜率之积为 $-\frac{3}{4}$, 求 $\frac{BM}{BQ}$ 的值.

知乎@余命数

解　答案: (1) $\frac{x^2}{4} + \frac{y^2}{3} = 1$; (2) 5.

(1) 由题意可得

知乎@余命数

$$
\begin{cases}
a^2 = b^2 + c^2 \\
\frac{1}{2}\times\frac{2b^2}{a}\times c = \frac{3}{2} \\
\frac{c}{a} = \frac{1}{2}
\end{cases}
$$

解得 $a = 2$、$b = \sqrt{3}$、$c = 1$, 所以椭圆的标准方程为 $\frac{x^2}{4} + \frac{y^2}{3} = 1$.

(2) 设点 $A(x_1, y_1)$、$B(x_2, y_2)$、$Q(x_3, y_3)$、$\overrightarrow{BM} = \lambda\overrightarrow{BQ}$, 则 $M(3x_1, 3y_1)$、$(3x_1 - x_2, 3y_1 - y_2) = \lambda(x_3 - x_2, y_3 - y_2)$、$\therefore$

$$
\begin{cases}
x_3 = \frac{3}{\lambda}x_1 + \frac{\lambda-1}{\lambda}x_2, \\
y_3 = \frac{3}{\lambda}y_1 + \frac{\lambda-1}{\lambda}y_2,
\end{cases}
$$

知乎@余命数

$\because k_{OA}\cdot k_{OB} = -\frac{3}{4}$, $\therefore \frac{x_1x_2}{4} + \frac{y_1y_2}{3} = 0$, 且 $\frac{x_1^2}{4} + \frac{y_1^2}{3} = 1$、$\frac{x_2^2}{4} + \frac{y_2^2}{3} = 1$、$\frac{x_3^2}{4} + \frac{y_3^2}{3} = 1$、$\therefore$

$$
\frac{(\frac{3}{\lambda}x_1 + \frac{\lambda-1}{\lambda}x_2)^2}{4} + \frac{(\frac{3}{\lambda}y_1 + \frac{\lambda-1}{\lambda}y_2)^2}{3} = 1,
$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

整理可得

$$\frac{9}{\lambda^2}\left(\frac{x_1^2}{4}+\frac{y_1^2}{3}\right)+\frac{(\lambda-1)^2}{\lambda^2}\left(\frac{x_2^2}{4}+\frac{y_2^2}{3}\right)+\frac{6(\lambda-1)}{\lambda^2}\left(\frac{x_1x_2}{4}+\frac{y_1y_2}{3}\right)=1,$$

$\therefore \frac{9}{\lambda^2}+\frac{(\lambda-1)^2}{\lambda^2}=1$, 即 $\lambda=5$, 故 $\frac{BM}{BQ}=5$.

13.20.5 本节小结

- 斜率和、差、商、积: 设线联立, 韦达定理表示交点坐标, 代入目标斜率表达式消参.
- 斜率积与圆锥曲线第三定义相关, 椭圆上一点与两个顶点连线的斜率积为 $-\frac{b^2}{a^2}$.
- 消参后通常得到定值; 若有参数剩余, 则求范围与最值.

13.20.6 本节习题

13.105. 设直线 $y=kx+1$ 与椭圆 $\frac{x^2}{2}+y^2=1$ 交于 A,B 两点. 求椭圆在 A,B 处切线斜率之和, 用 k 表示.

13.106. 设椭圆 $\frac{x^2}{4}+y^2=1$ 上两点 A,B 满足 $OA\perp OB$(O 为原点). 求 $k_{OA}\cdot k_{OB}$ 的值.

13.21 圆幂与反演

切割线定理与相交弦定理从两个角度处理点与圆的位置关系. 把两者统一起来, 就得到圆幂定理. 以圆幂为工具, 进一步引入反演变换, 即一种把圆内与圆外互相对换的几何变换, 它将直线与圆相互转化, 是纯几何证明的有力工具. 根轴与根心的代数处理见 第13.13节, 本节从几何角度补充.

13.21.1 圆幂定理

定义 13.10 (点对圆的幂) 设 $\odot O$ 的半径为 R, 平面上一点 P 到圆心 O 的距离为 $d=OP$. 称

$$\mathrm{Pow}_{\odot O}(P)=d^2-R^2 \tag{13-38}$$

为点 P 对 $\odot O$ 的**幂**.

圆外的点幂为正, 圆内的点幂为负, 圆上的点幂为零.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

定理 13.38 (圆幂定理) 过点 P 引直线与 $\odot O$ 相交于 A、B 两点(相切时 A 与 B 重合), 则

$$PA \cdot PB = \left|\mathrm{Pow}_{\odot O}(P)\right| \tag{13-39}$$

此乘积与直线方向无关, 仅与点 P 和 $\odot O$ 有关.

证明 设直线 AB 交 $\odot O$ 于 A、B, M 为 AB 中点, 则 $OM \perp AB$. 由 $OA = R$ 及 $OP = d$,

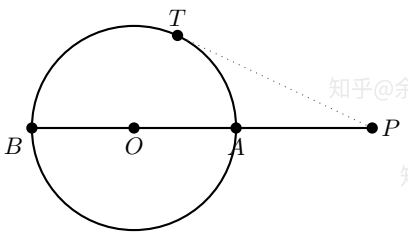
$$MA = \sqrt{R^2 - OM^2}, \quad PM = \sqrt{d^2 - OM^2}$$

当 P 在圆外时, $PA \cdot PB = (PM + MA)(PM - MA) = PM^2 - MA^2 = d^2 - R^2$.
当 P 在圆内时, $PA \cdot PB = (MA - PM)(MA + PM) = MA^2 - PM^2 = R^2 - d^2$.
两种情形合并为 $PA \cdot PB = |d^2 - R^2|$. □

知乎@余命数

推论 13.1 (切割线定理) P 在 $\odot O$ 外, PT 为切线, PAB 为割线, 则 $PT^2 = PA \cdot PB$.

推论 13.2 (相交弦定理) 弦 AB 与 CD 交于圆内点 P, 则 $PA \cdot PB = PC \cdot PD$.



知乎@余命数

图 13.45　圆幂定理, 割线 PAB 与切线 PT 满足 $PA \cdot PB = PT^2$

例题 13.313 (圆幂定理求弦长) P 是圆 $x^2 + y^2 = 25$ 外一点, $P = (7, 1)$. 过 P 引圆的割线, 交圆于 A、B. 求 $PA \cdot PB$.

知乎@余命数

解 圆心 $O = (0, 0)$, 半径 $R = 5$.$OP^2 = 49 + 1 = 50$, $R^2 = 25$. 由圆幂定理 (13-39),

知乎@余命数

$$PA \cdot PB = OP^2 - R^2 = 50 - 25 = 25$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

13.21.2 反演变换

圆幂定理把点与圆的关系浓缩成一个数. 反演变换从这个数出发, 让点 P 与其像 P' 满足 $OP\cdot OP'=R^2$, 即圆幂取定值 R^2. 反演把圆内与圆外对换, 又保持角度, 因而能把交错的圆与直线配置化成简单的.

定义 13.11 (反演变换) 设 $\odot O$ 的半径为 R, P 为平面上异于 O 的点. 在射线 OP 上取点 P' 使 $OP\cdot OP'=R^2$, 则称 P' 为 P 关于 $\odot O$ 的**反演像**, $\odot O$ 称为**反演圆**, O 称为**反演中心**. 映射 $P\mapsto P'$ 称为关于 $\odot O$ 的反演变换.

反演变换把反演圆上的点保持不动, 把圆内点映到圆外, 把圆外点映到圆内.

定理 13.39 (反演的基本性质) 反演变换具有以下性质:

1. **对合性**: 反演的逆变换仍是同一反演, 即 $P''=P$.
2. **保角性**: 反演保持两条曲线在交点处的夹角大小, 但方向反转.
3. **距离公式**: 反演一般不保持距离, 但满足

$$P'Q'=\frac{R^2}{OP\cdot OQ}\cdot PQ \tag{13-40}$$

证明 仅证距离公式. 设 $\angle POQ=\theta$. 由 $OP'=R^2/OP, OQ'=R^2/OQ$,

$$P'Q'^2=OP'^2+OQ'^2-2OP'\cdot OQ'\cos\theta=\frac{R^4}{OP^2}+\frac{R^4}{OQ^2}-\frac{2R^4\cos\theta}{OP\cdot OQ}$$

$$=\frac{R^4}{OP^2\cdot OQ^2}\left(OQ^2+OP^2-2OP\cdot OQ\cos\theta\right)=\frac{R^4}{OP^2\cdot OQ^2}\cdot PQ^2$$

开方即得 (13-40). □

定理 13.40 (反演下的像)

1. 过反演中心的直线, 反演后仍为该直线(点在射线上互换).
2. 不过反演中心的直线, 反演后为过反演中心的圆.
3. 不过反演中心的圆, 反演后为不过反演中心的圆.
4. 过反演中心的圆, 反演后为不过反演中心的直线.
5. 与反演圆正交的圆, 反演后仍为自身.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

证明 仅证第 2 条. 设直线 l 不过 O, O 到 l 的垂足为 H, $OH=h>0$. l 上点 P 的反演像 P' 满足 $OP'=R^2/OP$. 设 $\angle POH=\alpha$, 则 $OP=h/\cos\alpha$, 故 $OP'=R^2\cos\alpha/h$. 这是以 O 为原点、直径 R^2/h(沿 OH 方向)的圆的极坐标方程, 即 P' 在以 OH' 为直径的圆上, 其中 H' 是 H 的反演像, $OH'=R^2/h$. □

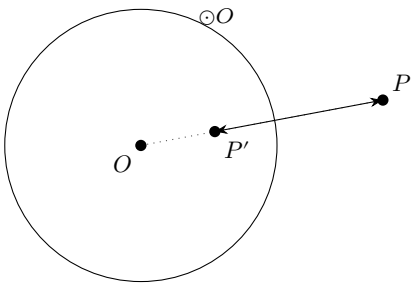


图 13.46　反演变换, $OP\cdot OP'=R^2$

知乎@余命数

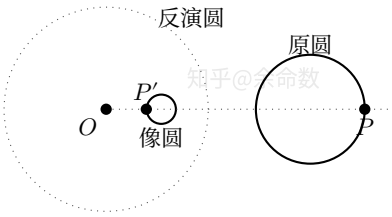


图 13.47　不过反演中心的圆, 反演后仍为圆; 离中心近的像离中心远

例题 13.314 (反演求轨迹) 求直线 $x=2$ 关于单位圆 $x^2+y^2=1$ 的反演像.

解 直线 $x=2$ 不过反演中心 $O=(0,0)$. 直线上点 $P=(2,t)$ 的反演像 $P'=\frac{P}{|P|^2}=\frac{(2,t)}{4+t^2}$. 设 $P'=(x',y')$, 则 $x'=\frac{2}{4+t^2}$, $y'=\frac{t}{4+t^2}$. 消去 t. 由 $t=2y'/x'$, 代入 $x'=\frac{2}{4+4y'^2/x'^2}$ 得 $x'(4+4y'^2/x'^2)=2$, 即 $4x'^2+4y'^2=2x'$, 亦即

$$\left(x'-\frac{1}{4}\right)^2+y'^2=\frac{1}{16}$$

反演像为圆 $\left(x-\frac{1}{4}\right)^2+y^2=\frac{1}{16}$, 圆心 $(1/4,0)$, 半径 $1/4$, 过反演中心.

知乎@余命数

13.21.3 根轴与根心的几何意义

第13.13节 从圆方程相减的代数角度引入了根轴与根心. 这里从圆幂的几何角度重新审视.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 13.41 (根轴的几何刻画) 设两圆 $\odot O_1$ 与 $\odot O_2$ 不同心. 它们的根轴就是对两圆幂相等的点的轨迹, 是一条垂直于连心线 O_1O_2 的直线. 两圆相交时, 根轴是公共弦; 相切时, 根轴是过切点的公切线.

证明 点 P 在根轴上当且仅当 $PO_1^2 - r_1^2 = PO_2^2 - r_2^2$. 以 O_1O_2 为 x 轴建系, $O_1 = (0,0)$、$O_2 = (d,0)$、$P = (x,y)$, 则 $x^2 + y^2 - r_1^2 = (x-d)^2 + y^2 - r_2^2$, 化简得 $x = (d^2 + r_1^2 - r_2^2)/(2d)$, 是垂直于 O_1O_2 的直线. 两圆相交时, 交点对两圆幂均为零, 故交点在根轴上, 根轴即公共弦. □

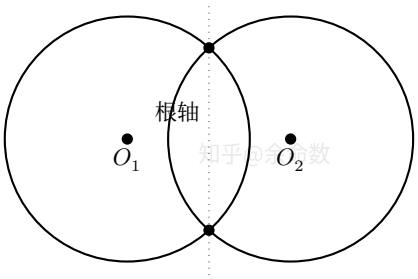


图 13.48 两圆的根轴, 即到两圆幂相等的点的轨迹, 相交时即公共弦

定理 13.42 (根心定理的几何刻画) 三个圆心不共线的圆, 三条根轴交于同一点(圆心共线时三条根轴互相平行). 该点对三圆的幂相等.

证明 设 $\odot O_1$、$\odot O_2$、$\odot O_3$ 的根轴分别为 l_{12}、l_{13}、l_{23}. 设 $P = l_{12} \cap l_{13}$, 则 P 对 $\odot O_1$ 与 $\odot O_2$ 幂相等, 且对 $\odot O_1$ 与 $\odot O_3$ 幂相等, 故对 $\odot O_2$ 与 $\odot O_3$ 幂也相等, 即 P 在 l_{23} 上. □

根心把三个圆的幂统一到一个点, 过根心任作一直线交三圆于 A_i, B_i, 则 $PA_1 \cdot PB_1 = PA_2 \cdot PB_2 = PA_3 \cdot PB_3$. 三圆两两相交时, 三条公共弦交于根心; 三圆共点时, 根心就是那个公共点.

例题 13.315 (用反演化简共圆判定) A、B、C、D 四点共圆, P 为圆外一点, PA、PB、PC、PD 分别交圆于另一点 A'、B'、C'、D'. 证明 $A'B'C'D'$ 四点共圆.

解 以 P 为反演中心, 取反演半径 R 使 $R^2 = PA \cdot PA' = PB \cdot PB' = PC \cdot PC' = PD \cdot PD'$(由圆幂定理, 这些乘积相等). 则 A'、B'、C'、D' 分别是 A、B、C、D 的反演像. 由 定理 13.40 第 3 条, 不过反演中心的圆反演后仍为圆, 故 $A'B'C'D'$ 共圆.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

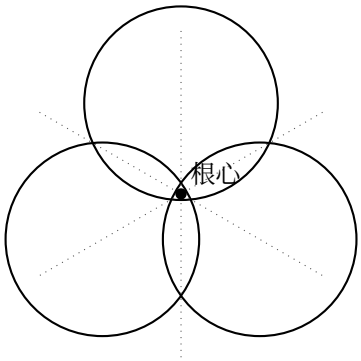


图 13.49　三圆的根轴交于根心, 即到三圆幂相等的点

知乎@余命数

13.22 圆锥曲线的光学性质

第13.3节 证明了椭圆的反射性质, 从一个焦点发出的光经椭圆内壁反射必汇聚于另一焦点. 双曲线与抛物线也有各自的反射性质, 三条定理放在一起, 正好凑成几何光学里最漂亮的一组结论. 接着我们就把另外两条补齐.

13.22.1 双曲线的反射性质与虚焦点

定理 13.43 (双曲线的反射性质) 设 P 是双曲线 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$ 右支上一点, F_1,F_2 是左右焦点, 则 P 处的切线平分 $\angle F_1PF_2$ 的外角. 于是从一个焦点 F_2 射向双曲线镜面的光线, 反射后散开, 但反射线的反向延长线全部穿过另一焦点 F_1, 观察者会觉得光是从 F_1 发出的, F_1 称为**虚焦点**.

知乎@余命数

证明 设 $P(x_0,y_0)$ 在右支上, $c^2=a^2+b^2$, 离心率 $e=c/a$. 由 第13.3节 的焦半径公式, 右支上

$$|PF_1|=ex_0+a,\qquad |PF_2|=ex_0-a.$$

切线 $\dfrac{xx_0}{a^2}-\dfrac{yy_0}{b^2}=1$ 的一个法向量是 $\vec{n}=\left(\dfrac{x_0}{a^2},-\dfrac{y_0}{b^2}\right)$. 用 $\vec{n}$ 分别与两条焦半径向量作数量积. 由曲线方程 $y_0^2/b^2=x_0^2/a^2-1$,

知乎@余命数

$$\vec{n}\cdot\overrightarrow{PF_1}=-\frac{x_0(c+x_0)}{a^2}+\frac{y_0^2}{b^2}=-\frac{cx_0}{a^2}-\frac{x_0^2}{a^2}+\frac{x_0^2}{a^2}-1=-\frac{cx_0+a^2}{a^2},$$

$$\vec{n}\cdot\overrightarrow{PF_2}=\frac{x_0(c-x_0)}{a^2}+\frac{y_0^2}{b^2}=\frac{cx_0}{a^2}-1=\frac{cx_0-a^2}{a^2}.$$

除以各自的焦半径, 并注意 $ex_0+a=\dfrac{cx_0+a^2}{a}$, $ex_0-a=\dfrac{cx_0-a^2}{a}$, 得

$$\vec{n}\cdot\frac{\overrightarrow{PF_1}}{|PF_1|}=-\frac{1}{a},\qquad \vec{n}\cdot\frac{\overrightarrow{PF_2}}{|PF_2|}=\frac{1}{a}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两个单位向量与法向量的夹角余弦大小相等、符号相反, 即切线与两条焦半径成等角, 而且平分的是外角. 按反射定律, 从 F_2 射来的光线在 P 处反射后, 反射线沿 $\overrightarrow{PF_2}$ 的延长方向射出, 其反向延长线经过 F_1. □

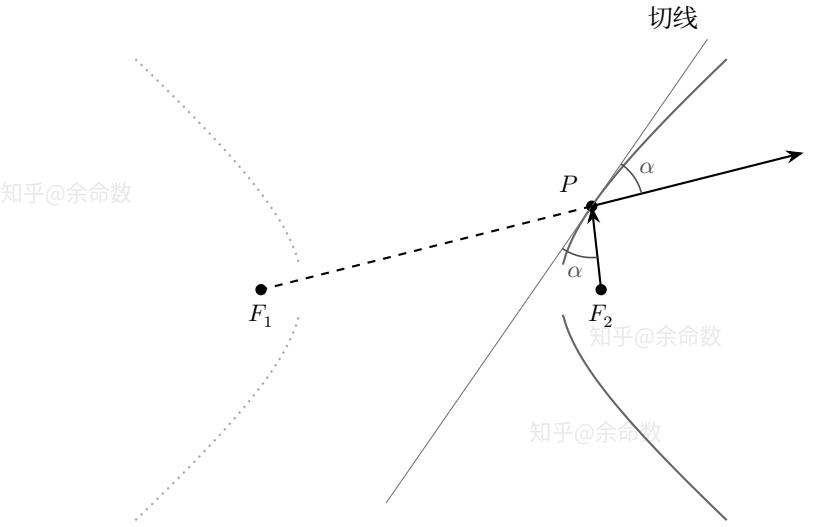


图 13.50 双曲线的反射(P 在 $\frac{x^2}{4}-\frac{y^2}{3}=1$ 右支上): 入射线 F_2P 在 P 处反射, 反射线与切线成等角 α, 虚线为反向延长线, 经过虚焦点 F_1.

13.22.2 抛物线的平行光聚焦性质

抛物线的反射性质有一个出乎意料干净的刻画, 焦点关于切线的对称点可以直接写出来.

定理 13.44 (抛物线的反射性质) 设 $P(x_0,y_0)$ 在抛物线 $y^2=2px$ 上, 焦点 $F\left(\frac{p}{2},0\right)$, 准线 $x=-\frac{p}{2}$, 则 F 关于 P 处切线的对称点恰是 P 在准线上的投影 $A'\left(-\frac{p}{2},\,y_0\right)$. 因此切线平分 PF 与过 P 的轴平行线所成的角, 沿对称轴方向平行的光线经抛物线反射后全部汇聚于焦点.

证明 P 处的切线为 $yy_0=p(x+x_0)$, 写成一般式 $px-y_0y+px_0=0$, 法向量 $\vec{n}=(p,-y_0)$. 设 F 关于切线的对称点为 F', 则

$$\overrightarrow{FF'}=-2\cdot\frac{p\cdot\frac{p}{2}+px_0}{p^2+y_0^2}(p,-y_0).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

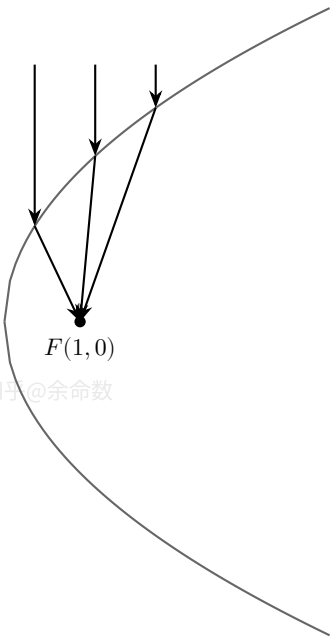
由 $y_0^2=2px_0$ 得分母 $p^2+y_0^2=p(2x_0+p)$, 分子 $p\left(\frac{p}{2}+x_0\right)$, 于是

$$\overrightarrow{FF'}=-\frac{2p(x_0+\frac{p}{2})}{p(2x_0+p)}(p,-y_0)=-(p,-y_0)=(-p,\ y_0),$$

知乎@余命数

$$F'=F+(-p,y_0)=\left(-\frac{p}{2},\ y_0\right),$$

正是 P 在准线上的投影. 由于 $|PF|=|PA'|$(抛物线定义), 又由对称性有 $|PF|=|PF'|$, 切线是等腰三角形 FPF' 顶角平分线所在的直线, 于是切线与 PF、与水平线 PA' 成等角. 沿轴向平行的光线在 P 处反射, 反射角等于入射角, 反射线便沿 PF 射出, 汇聚于 F. □



知乎@余命数
知乎@余命数

图 13.51　抛物线 $y^2=4x$ 的反射: 平行于轴的入射光全部汇聚于焦点, 这是探照灯、太阳灶与抛物面天线的原理.

例题 13.316 (手算一个反射点) 抛物线 $y^2=4x$ 上取 $P(1,2)$, 验证 定理 13.44 的结论.

解 此时 $p=2$, 焦点 $F(1,0)$, 准线 $x=-1$, P 在准线上的投影是 $A'(-1,2)$.P 处切线为 $2y=2(x+1)$, 即 $y=x+1$. 先求 $F(1,0)$ 关于直线 $y=x+1$ 的对称点, 直线写成 $x-y+1=0$, 代入对称公式得

$$F'=(1,0)-2\cdot\frac{1-0+1}{1^2+(-1)^2}(1,-1)=(1,0)-(2,-2)=(-1,2)=A'.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

切线 $y=x+1$ 倾斜角 45°, 恰好平分水平线与竖直线 PF 所成的直角, 从上方平行射入的光在 P 处竖直向下折入焦点.

13.22.3 三条曲线, 一条规律

现在把三条反射性质并排读. 椭圆的切线平分两焦半径夹角的**内**角, 双曲线的切线平分**外**角, 抛物线的切线平分焦半径与轴平行线的夹角. 把抛物线看成另一个焦点退到无穷远的椭圆, 轴平行线正是通向无穷远焦点的焦半径, 三条定理便合成一条, 二次曲线上任意一点处的切线, 平分两条焦半径, 其中一条可以是退化的所成的角. 从这一条规律出发, 椭圆把光从一个焦点搬进另一个焦点, 双曲线把光散射成来自虚焦点的样子, 抛物线则把平行的星光收拢成一点.

这些性质不只是比喻. 探照灯与汽车大灯的反射面是旋转抛物面, 灯丝放在焦点, 出射光便是平行光束, 定理 13.44 反过来用即是; 传说阿基米德在叙拉古围城战中排布铜镜聚阳光焚烧罗马舰队, 用的也是把阳光汇聚到一点的原理. 天文台的射电望远镜与太阳灶把同一原理用于接收, 平行入射的电磁波或阳光被反射面汇聚到焦点处的接收器. 椭圆回声廊里, 一个人站在焦点处低语, 站在另一焦点的人听得清楚, 声波走的正是 第13.3节 的反射路线. 卡塞格伦望远镜更把两面镜子组合起来, 主镜是抛物面, 副镜是双曲面, 让双曲面的一个焦点与抛物面的焦点重合, 光先被抛物面送往这个重合点, 半路被双曲面截住, 散射成来自虚焦点的光, 而 定理 13.43 保证虚焦点正是双曲面的另一个焦点, 光在那里被收集. 一支望远镜, 同时用上了两条反射定理.

这套双曲线定位系统在历史上真实使用过, 台链沿海岸布设, 船只靠测量时间差在双曲线网格上定出自己的位置, 在全球定位卫星出现之前, 它是远洋导航的主力.

例题 13.317 (双曲线导航) 海岸线上有两组电台. 第一组 $A(-30,0)$ 与 $B(30,0)$ 相距 60 km, 第二组 $C(0,30)$ 与 $D(0,-30)$ 相距 60 km. 船上的接收机测得 $|PA|-|PB|=20$ km, $|PC|-|PD|=20$ km(距离差来自信号到达的时间差, 无线电波以光速传播). 求船的位置.

解 到 A,B 距离差为定值 20 的轨迹是双曲线的一支, $2a=20$、$c=30$、$b^2=900-100=800$, 且船离 B 近, 故在右支

$$\frac{x^2}{100}-\frac{y^2}{800}=1\quad(x>0).$$

到 C,D 距离差 20 的轨迹是上下开口的双曲线 $y^2/100-x^2/800=1$ 的一支, 船离 D 近, 故取下半支 $y<0$. 两式联立, 第一式乘 8 得 $8x^2-y^2=800$, 第二式乘 8 得

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$8y^2-x^2=800$, 相减得 $9x^2=9y^2$, 即 $y=\pm x$. 船在下半支与右半支, 取 $y=-x$, 代回 $8x^2-x^2=800$ 得

$$x=\sqrt{\frac{800}{7}}\approx 10.7,\qquad y=-\sqrt{\frac{800}{7}}\approx -10.7.$$

船位于点 $(10.7,-10.7)$ 附近, 与两台核实, $|PA|\approx 42.1$, $|PB|\approx 22.1$, 差恰为 20 km.

这套办法的双曲线定位系统在历史上真实使用过, 台链沿海岸布设, 船只靠测量时间差在双曲线网格上定出自己的位置, 在全球定位卫星出现之前, 它是远洋导航的主力.

练习 13.107　双曲线 $\dfrac{x^2}{9}-\dfrac{y^2}{16}=1$ 上取 $P\left(5,\dfrac{16}{3}\right)$, 求 P 处的切线, 并验证切线与两条焦半径所成的角相等.

练习 13.108　椭圆与双曲线共焦点, 称为共焦曲线族. 证明同一族中的椭圆与双曲线在交点处互相垂直, 即两曲线在交点处的切线互相垂直.

练习 13.109　探照灯的反射面是旋转抛物面, 口径直径 1.2 m, 深度 0.1 m. 灯丝应放在距顶点多远处?

13.23 焦点弦

过焦点的弦叫**焦点弦**. 第13.3节 已经就抛物线算过一条焦点弦的长度, 这里把三类曲线放在一起看, 焦点弦有哪些不依赖具体位置的固有性质. 只消把焦半径写成方向的函数, 一串结论会接连落下.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

13.23.1 焦半径的方向公式

引理 13.3 (焦半径的方向公式) 设过焦点 F 的直线与曲线交出弦 AB, 直线方向与 x 轴(对称轴)夹角为 θ, 记沿该方向到交点的有向距离为 $r(\theta)$, 离心率为 e, 抛物线 $y^2=2px$ 的焦参数为 p, 则

$$r(\theta)=\frac{b^2/a}{1+e\cos\theta}\ (\text{椭圆}),\qquad r(\theta)=\frac{b^2/a}{1-e\cos\theta}\ (\text{双曲线}),$$

$$r(\theta)=\frac{p}{1-\cos\theta}\ (\text{抛物线},\ e=1).$$

证明 以右焦点 $F_2(c,0)$ 与右准线为例. 第13.3节 的统一定义说, 曲线上任一点 P 满足 $|PF_2|=e\cdot d(P,\text{右准线})$.

椭圆的右准线是 $x=a^2/c$. 设从 F_2 出发、方向角为 θ 的射线上的点为 $(c+r\cos\theta,\ r\sin\theta)$, 代入定义,

$$r=e\left(\frac{a^2}{c}-c-r\cos\theta\right)=\frac{eb^2}{c}-er\cos\theta,\qquad r(1+e\cos\theta)=\frac{b^2}{a},$$

末步用了 $eb^2/c=b^2/a$. 抛物线的准线是 $x=-p/2$, 此时 $e=1$, 焦点右侧的准线距是 p, 同样代入得 $r=p+r\cos\theta$, 即 $r(1-\cos\theta)=p$. 双曲线的右准线仍是 $x=a^2/c$, 但它落在顶点左边, 距离要改成 $x-a^2/c$, 代入后符号翻面, 得 $r(1-e\cos\theta)=b^2/a$. 双曲线的公式只在 $1-e\cos\theta>0$ 时给出正的 r, 这正对应直线与右支交出两个点的情形. □

定理 13.45 (焦点弦的调和性质) 过焦点的弦 AB 满足

$$\frac{1}{|AF|}+\frac{1}{|BF|}=\begin{cases}\dfrac{2a}{b^2}, & \text{椭圆与双曲线(双曲线要求两端在同一支)},\\[2mm] \dfrac{2}{p}, & \text{抛物线}.\end{cases}$$

证明 弦的两个端点位于同一条过焦点的直线上, 方向角相差 π, $\cos(\theta+\pi)=-\cos\theta$. 对 引理 13.3 的任何一个公式取倒数相加, 含 $\cos\theta$ 的项成对消去,

$$\frac{1}{r(\theta)}+\frac{1}{r(\theta+\pi)}=\frac{1+e\cos\theta}{b^2/a}+\frac{1-e\cos\theta}{b^2/a}=\frac{2a}{b^2},$$

抛物线情形把 b^2/a 换成 p 即可. □

知乎@余命数

推论 13.3 (通径最短) 椭圆与双曲线中过焦点的弦最短是通径 $2b^2/a$, 抛物线中是通径 $2p$, 且都在弦垂直于轴时取到.

证明 记 $r_1=|AF|, r_2=|BF|$. 由 定理 13.45,

$$|AB|=r_1+r_2=\left(\frac{1}{r_1}+\frac{1}{r_2}\right)r_1r_2=\frac{2a}{b^2}\,r_1r_2\le\frac{2a}{b^2}\left(\frac{r_1+r_2}{2}\right)^2=\frac{a}{2b^2}|AB|^2,$$

中间一步是均值不等式 (第5.3节), 于是 $|AB|\ge 2b^2/a$. 等号当且仅当 $r_1=r_2$, 即焦点是弦的中点, 弦垂直于轴, 恰是通径. 抛物线同理给出 $|AB|\ge 2p$. □

注 13.20
把方向公式两式相加还能一步读出焦点弦长

$$|AB|=r(\theta)+r(\theta+\pi)=\frac{2b^2/a}{1-e^2\cos^2\theta},$$

$\theta=90^\circ$ 时最短, 又见 推论 13.3. 行星沿椭圆轨道运行时, 近日距 $a-c$ 与远日距 $a+c$ 恰好是以半通径 b^2/a 为调和平均的一对数, 这正是 定理 13.45 取 $\theta=0^\circ$ 的情形.

知乎@余命数

例题 13.318 (三类曲线, 同一个角) 抛物线 $y^2=4x$、椭圆 $\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$、双曲线 $\dfrac{x^2}{4}-\dfrac{y^2}{3}=1$ 各有一条过右焦点的弦, 倾角都是 60°. 分别求弦长.

解 抛物线 $p=2$,

$$|AF|=\frac{p}{1-\cos 60^\circ}=\frac{2}{1-1/2}=4,\qquad |BF|=\frac{2}{1+1/2}=\frac{4}{3},\qquad |AB|=\frac{16}{3}.$$

椭圆 $b^2/a=3/2$, $e=1/2$,

$$|AF|=\frac{3/2}{1+1/4}=\frac{6}{5},\qquad |BF|=\frac{3/2}{1-1/4}=2,\qquad |AB|=\frac{16}{5}.$$

双曲线 $e=\sqrt{7}/2$, $\cos 60^\circ=1/2$,

$$|AF|=\frac{3/2}{1-\sqrt{7}/4}=\frac{6}{4-\sqrt{7}}=\frac{2(4+\sqrt{7})}{3},\qquad |BF|=\frac{3/2}{1+\sqrt{7}/4}=\frac{2(4-\sqrt{7})}{3},$$

$$|AB|=\frac{2(4+\sqrt{7})}{3}+\frac{2(4-\sqrt{7})}{3}=\frac{16}{3}.$$

知乎@余命数

三条弦与焦点折出的三对焦半径各自满足调和式, 抛物线与双曲线的弦长同为 $16/3$, 椭圆的短一些, 因为它的 $e<1$ 让分母 $1-e^2\cos^2\theta$ 变大. 核对抛物线一条, $\theta=60^\circ$ 即斜率 $\sqrt{3}$, 第13.3节 的焦点弦长公式给 $4(1+3)/3=16/3$, 一致.

13.23.2 抛物线焦点弦的三件套

抛物线的焦点弦还有一组几何味更浓的性质, 用参数法一次全拿出来. 把 $y^2=2px$ 的点写成 $\left(\dfrac{pt^2}{2},\,pt\right)$, 过参数 t_1,t_2 两点的弦有方程

$$x=\frac{t_1+t_2}{2}\,y-\frac{p\,t_1t_2}{2},$$

把焦点 $\left(\dfrac{p}{2},0\right)$ 代入立得 $t_1t_2=-1$, 这就是焦点弦的参数刻画. 参数 t 处的切线是 $x=ty-\dfrac{pt^2}{2}$, 斜率 $\dfrac{1}{t}$.

定理 13.46 (抛物线焦点弦三性质) 抛物线 $y^2=2px$ 的焦点弦 AB 两端参数为 t_1,t_2, $t_1t_2=-1$. 设 A',B' 是 A,B 在准线上的投影, 则 **(i)** A、B 处的切线交于准线上一点 T, 且两切线互相垂直; **(ii)** $FT\perp AB$; **(iii)** $FA'\perp FB'$.

证明 (i) 联立两条切线 $x=t_1y-\dfrac{pt_1^2}{2}$ 与 $x=t_2y-\dfrac{pt_2^2}{2}$, 两式相减得 $(t_1-t_2)y=\dfrac{p(t_1^2-t_2^2)}{2}$, 故 $y_T=\dfrac{p(t_1+t_2)}{2}$, 代回得 $x_T=\dfrac{pt_1t_2}{2}=-\dfrac{p}{2}$, 交点落在准线 $x=-\dfrac{p}{2}$ 上. 斜率之积 $\dfrac{1}{t_1}\cdot\dfrac{1}{t_2}=\dfrac{1}{t_1t_2}=-1$, 两切线垂直.

(ii) $T\left(-\dfrac{p}{2},\,\dfrac{p(t_1+t_2)}{2}\right)$ 与 $F\left(\dfrac{p}{2},\,0\right)$ 连线的斜率是 $\dfrac{p(t_1+t_2)/2}{-p}=-\dfrac{t_1+t_2}{2}$, 而弦 AB 的斜率由弦方程读出是 $\dfrac{2}{t_1+t_2}$, 乘积为 -1, 垂直.

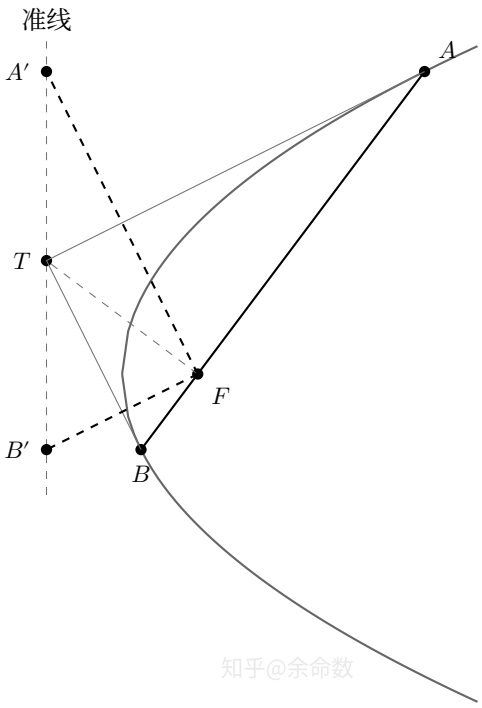
(iii) $A'=\left(-\dfrac{p}{2},\,pt_1\right)$, $B'=\left(-\dfrac{p}{2},\,pt_2\right)$, 于是

$$\overrightarrow{FA'}\cdot\overrightarrow{FB'}=(-p,pt_1)\cdot(-p,pt_2)=p^2(1+t_1t_2)=0.$$

□

反过来, 从准线上任一点 T 向抛物线引两条切线, 设切点为 A,B, 由 (i) 的推导过程倒推, T 的横坐标是 $-p/2$ 迫使两切点参数之积为 -1, 于是 AB 必是焦点弦. 准线、垂直切线对、焦点弦三者一一对应, 定理 13.46 的三幅图景是同一幅.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



知乎@余命数

图 13.52　$y^2=4x$ 的焦点弦 $AB(t_1=2,\ t_2=-\frac{1}{2})$: 切线交点 T 落在准线上且两切线垂直, $FT\perp AB, FA'\perp FB'$.

例题 13.319 (定长焦点弦) 抛物线 $y^2=4x$ 上, 求长度为 5 的焦点弦的斜率, 并问焦点弦长度能否小于 4.

解 设倾角 θ, 由焦点弦长公式 $|AB|=\dfrac{2p}{1-\cos^2\theta}=\dfrac{4}{\sin^2\theta}$. 令 $\dfrac{4}{\sin^2\theta}=5$ 得 $\sin^2\theta=\dfrac{4}{5}$, $\tan^2\theta=\dfrac{\sin^2\theta}{1-\sin^2\theta}=4$, 斜率 $k=\pm2$. 至于下界, $\sin^2\theta\le1$ 给出 $|AB|\ge4$, 等号在 $\theta=90^\circ$ 即通径时取到, 焦点弦长度不会小于 4, 与 推论 13.3 一致.

练习 13.110　椭圆 $\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$ 过右焦点的弦倾角 60°, 用 定理 13.45 验证 $\dfrac{1}{|AF|}+\dfrac{1}{|BF|}=\dfrac{4}{3}$, 并求 $|AB|$.

练习 13.111　双曲线 $\dfrac{x^2}{4}-\dfrac{y^2}{3}=1$ 过右焦点、斜率 1 的弦, 两端是否在同一支上? 求 $\dfrac{1}{|AF|}+\dfrac{1}{|BF|}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

练习 13.112 抛物线 $y^2=4x$ 的焦点弦 AB 满足 $|AF|=4$, 求 $|BF|$ 与 AB 的倾斜角, 并指出 A、B 处切线的交点位置.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 乎@余命数

第 14 章

极坐标与参数方程*

几何主干

「在不确定之中，仍有数的法则可循。」

——雅各布·伯努利



第十四章 极坐标与参数方程 *

自笛卡尔建立解析几何以来, 用代数方程描述几何图形便成为数学中的重要观点之一. 方程 $f(x,y)=0$ 定义了平面上所有满足特定约束条件的点的集合. 这种观点在刻画直线、圆锥曲线等固定的几何形态时十分有效.

然而, 数学的发展, 特别是微积分的诞生, 要求我们既能描述是什么, 也能刻画如何形成. 早在古希腊时期, 数学家们就已经用动态的、运动学的方法构造曲线, 例如阿基米德螺线和尼科米德斯蚌线. 这些曲线的定义本身就蕴含了一种生成过程. 当牛顿与莱布尼茨试图用数学语言描述行星的运行、抛射物的轨迹时, 笛卡尔坐标的内在局限性便显现出来. 它难以表达时间这一关键变量. 一个运动的质点, 其位置坐标 (x,y) 共同受控于一个更根本的独立变量, 即时间 t.

为了捕捉这种动态的生成过程, 一种新的描述方式应运而生, 这就是**参数方程**. 它把点的坐标 x 和 y 分别视为某个第三方参数 t 的函数, 即 $x=f(t), y=g(t)$. 这种方法把曲线看作一个点随参数 t 变化而连续描绘出的轨迹, 恰好对应物理学中的运动学描述, 同时为曲线附加了方向、速度等动态属性, 也简化了复杂曲线的构造与分析.

此时, 数学家们也注意到, 极坐标更适合处理具有旋转对称性或中心相关性的图形. 对于花瓣曲线、心形线以及开普勒定律所描述的行星轨道这类围绕某个中心点展开的几何图形, 使用 (x,y) 坐标描述往往导致方程异常繁琐.

十七世纪的数学家博纳文图拉·卡瓦列里以及后来的雅各布·伯努利等人发现, 用一个点到原点的距离 r 和该点与原点连线相对于基准方向的夹角 θ 来确定点的位置, 更为自然、简洁. 这一思想最终发展成为**极坐标系**.

本章是解析几何的扩展与深化, 系统介绍参数方程与极坐标这两种实用工具. 先引入参数方程, 讨论用它描述曲线生成的方式, 以及与普通方程的互化; 再进入极坐标, 建立极坐标与直角坐标的联系, 研究极坐标下中心对称曲线的表达与分析. 两套工具的引入, 丰富了观察、描述和解决几何问题的视角与手段.

14.0.1 参数方程

14.0.1.1 参数方程的引入

在物理学中, 一个质点在平面上的运动轨迹, 其位置坐标 (x,y) 都是时间 t 的函数.

$$\begin{cases} x=f(t) \\ y=g(t) \end{cases}$$

这里, 时间 t 如同一个独立的控制变量, 随着它的变化, 点 (x,y) 在坐标平面上描绘出一条轨迹. 把这种思想从物理情境中抽象出来, 便得到参数方程的一般概念.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 14.1 (参数方程) 在平面直角坐标系中,如果曲线上任意一点的坐标 x, y 都是某个独立变量 t 的函数.

$$\begin{cases} x = f(t) \\ y = g(t) \end{cases}$$

并且对于 t 在某个允许的取值范围(定义域 I)内的每一个值,由上述方程组所确定的点 (x, y) 都在这条曲线上,那么这个方程组就叫做这条曲线的**参数方程**,联系变量 x, y 的变量 t 叫做**参变数**,简称**参数**.

参数方程把二维平面上点 (x, y) 之间的内在约束关系,拆成两个独立的、关于同一个一维参数 t 的函数关系. 这样可分别研究 x 和 y 随参数变化的规律,从而更精细地理解曲线的动态生成过程.

14.0.1.2 参数方程与普通方程的互化

参数方程与普通方程是描述同一曲线的两种不同形式, 它们之间可以相互转化.

- **由参数方程化为普通方程**. 要点是**消去参数**. 通过代数变形, 从方程组中消去参数 t, 得到一个只含 x, y 的关系式. 消参的方法多种多样, 包括代入消元法、加减消元法、利用三角恒等式等.
- **由普通方程化为参数方程**. 要点是**选取参数**. 选择一个合适的变量作为参数, 将 x, y 表示为该参数的函数. 参数的选取是灵活的,不同的选取方式会得到形式不同的参数方程. 通常我们会选择具有明确几何或物理意义的量作为参数, 如时间、角度、斜率等.

消参时要注意, 消去参数的过程可能扩大或缩小 x, y 的取值范围. 因此, 在得到普通方程后,必须依据参数方程中参数的取值范围, 确定 x, y 的真实取值范围.

例题 14.1 (直线) 将下列参数方程化为普通方程,并说明其轨迹.

$$\begin{cases} x = 2t - 1 \\ y = 3t + 2 \end{cases} \quad (t\text{为参数})$$

知乎@余命数

解 此方程组关于参数 t 线性, 适于用代入消元法消参. 由第一式解出 $t=\frac{x+1}{2}$, 代入第二式 $y=3t+2$, 整理即得普通方程.

知乎@余命数

$$\begin{aligned} y &= 3\cdot\frac{x+1}{2}+2 && \text{（代入消元）}\\ 2y &= 3(x+1)+4=3x+7 && \text{（去分母、合并）}\\ 3x-2y+7&=0 && \text{（移项）} \end{aligned}$$

参数 t 取遍全体实数 $\mathbb{R}$, 故 x,y 也取遍全体实数, 该参数方程表示一条完整的直线 $3x-2y+7=0$.

例题 14.2 (抛物线) 将参数方程

$$\begin{cases} x=2t \\ y=t^2-1 \end{cases} \quad (t\text{为参数})$$

化为普通方程, 并指出其轨迹.

解 仍以代入消元法消参. 由 $x=2t$ 得 $t=\frac{x}{2}$, 代入 $y=t^2-1$ 即得普通方程.

知乎@余命数 知乎@余命数 知乎@余命数

$$\begin{aligned} y &= \left(\frac{x}{2}\right)^2-1 && \text{（代入消元）}\\ &= \frac{1}{4}x^2-1 && \text{（整理）} \end{aligned}$$

参数 t 可取任意实数, 故 $x=2t$ 取遍全体实数; 又由 $t^2\geq 0$ 知 $y=t^2-1\geq -1$. 这恰是抛物线 $y=\frac{1}{4}x^2-1$ 自顶点 $(0,-1)$ 起开口向上的整支, 无截断.

例题 14.3 (圆) 指出由参数方程

$$\begin{cases} x=5\cos\theta \\ y=5\sin\theta \end{cases} \quad (\theta\in[0,2\pi))$$

所表示的曲线.

解 此参数方程含三角函数, 直接代数消元不便, 宜借助恒等式 $\cos^2\theta+\sin^2\theta=1$ 消参. 由参数方程得 $\cos\theta=\frac{x}{5}$, $\sin\theta=\frac{y}{5}$, 代入恒等式整理.

知乎@余命数

$$\begin{aligned} \left(\frac{x}{5}\right)^2+\left(\frac{y}{5}\right)^2&=1 && \text{（代入恒等式）}\\ x^2+y^2&=25 && \text{（去分母）} \end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

θ 在 $[0,2\pi)$ 上取值时, $\cos\theta,\sin\theta$ 各自取遍 $[-1,1]$, 故 x,y 均取遍 $[-5,5]$, 与圆方程的定义域、值域吻合. 该参数方程表示以原点为圆心、半径为 5 的圆; 参数 θ 即从原点指向点 (x,y) 的射线与 x 轴正半轴的夹角, 当 θ 从 0 增至 2π, 点 (x,y) 自 $(5,0)$ 出发沿圆逆时针运行一周.

14.0.1.3 常见曲线的参数方程

任何曲线的普通方程原则上都可以转化为参数方程, 实践中真正被广泛使用的, 是那些代数结构简洁、几何或物理内涵清晰的特定参数化形式. 熟悉这些标准形式是利用参数方程解决问题的基础.

14.0.1.3.1 圆的参数方程 考虑以原点为圆心, 半径为 R 的圆, 其普通方程为 $x^2+y^2=R^2$. 为建立其参数方程, 要找一个能唯一描述圆上点位置的单一变量. 最自然的选择是极角, 即从原点到圆上一点 $P(x,y)$ 的射线 OP 与 x 轴正半轴的夹角, 记为 θ.

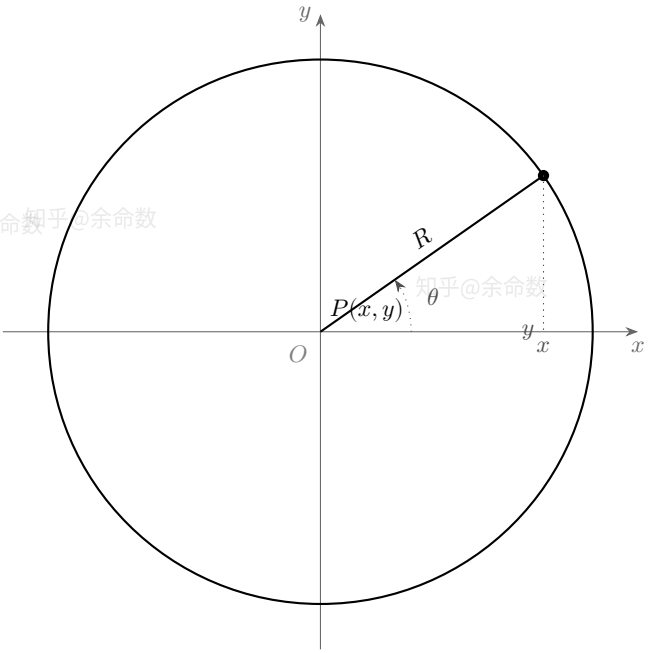


图 14.1 点 P 的坐标由半径 R 和极角 θ 确定.

根据三角函数的定义, 在以 OP 为斜边的直角三角形中, 有.

$$x=R\cos\theta,\quad y=R\sin\theta$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这就构成了圆的参数方程. 随着参数 θ 从 0 变化到 2π, 点 P 沿逆时针方向在圆上运动一周.

$$\begin{cases} x = R\cos\theta \\ y = R\sin\theta \end{cases} \quad (\theta \in [0, 2\pi) \text{ 为参数}) \tag{14-1}$$

若圆心平移至 (x_0, y_0), 其普通方程为 $(x-x_0)^2 + (y-y_0)^2 = R^2$. 其参数方程相应地平移为.

$$\begin{cases} x = x_0 + R\cos\theta \\ y = y_0 + R\sin\theta \end{cases} \quad (\theta \in [0, 2\pi) \text{ 为参数}) \tag{14-2}$$

例题 14.4 (三角形重心的轨迹) 已知定点 $A(2,0)$ 和 $B(4,0)$, 动点 P 在圆 $x^2 + y^2 = 1$ 上运动. 求 $\triangle ABP$ 的重心 G 的轨迹方程.

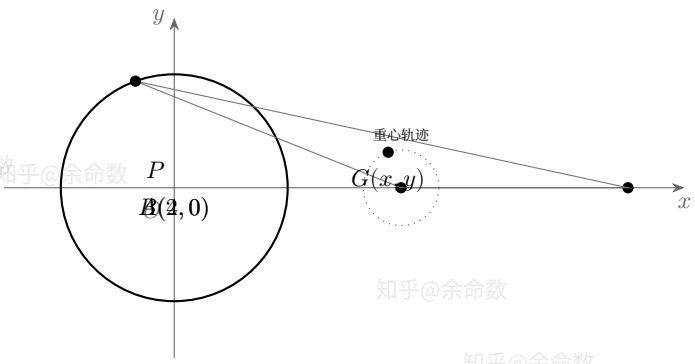


图 14.2 $\triangle ABP$ 的重心 G 的轨迹.

解 解题关键在于建立重心坐标与动点 P 坐标之间的联系, 而动点 P 的位置用参数方程刻画最为方便. 设 $P(\cos\theta, \sin\theta)(\theta \in [0, 2\pi)$ 为参数), 重心 G 的坐标为 (x, y). 由三角形重心坐标公式,

$$\begin{aligned} x &= \frac{x_A + x_B + x_P}{3} = \frac{2 + 4 + \cos\theta}{3} = \frac{6 + \cos\theta}{3}, \\ y &= \frac{y_A + y_B + y_P}{3} = \frac{0 + 0 + \sin\theta}{3} = \frac{\sin\theta}{3}. \end{aligned}$$

此即重心 G 轨迹的参数方程

$$\begin{cases} x = \frac{6+\cos\theta}{3} \\ y = \frac{\sin\theta}{3} \end{cases} \quad (\theta\text{为参数})$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

消去参数 θ 即得普通方程. 由参数方程解出 $\cos\theta = 3x-6,\ \sin\theta = 3y$, 代入三角恒等式 $\cos^2\theta + \sin^2\theta = 1$ 并整理.

$$
\begin{aligned}
(3x-6)^2 + (3y)^2 &= 1 \quad \text{（代入恒等式）} \\
9(x-2)^2 + 9y^2 &= 1 \quad \text{（展开）} \\
(x-2)^2 + y^2 &= \frac{1}{9} \quad \text{（约简）}
\end{aligned}
$$

此方程表示以点 $(2,0)$ 为圆心、半径为 $\frac{1}{3}$ 的圆, 即重心 G 的轨迹.

例题 14.5 (最值问题) 求圆 $x^2+y^2=4$ 上的点到直线 $3x+4y-25=0$ 的距离的最大值与最小值.

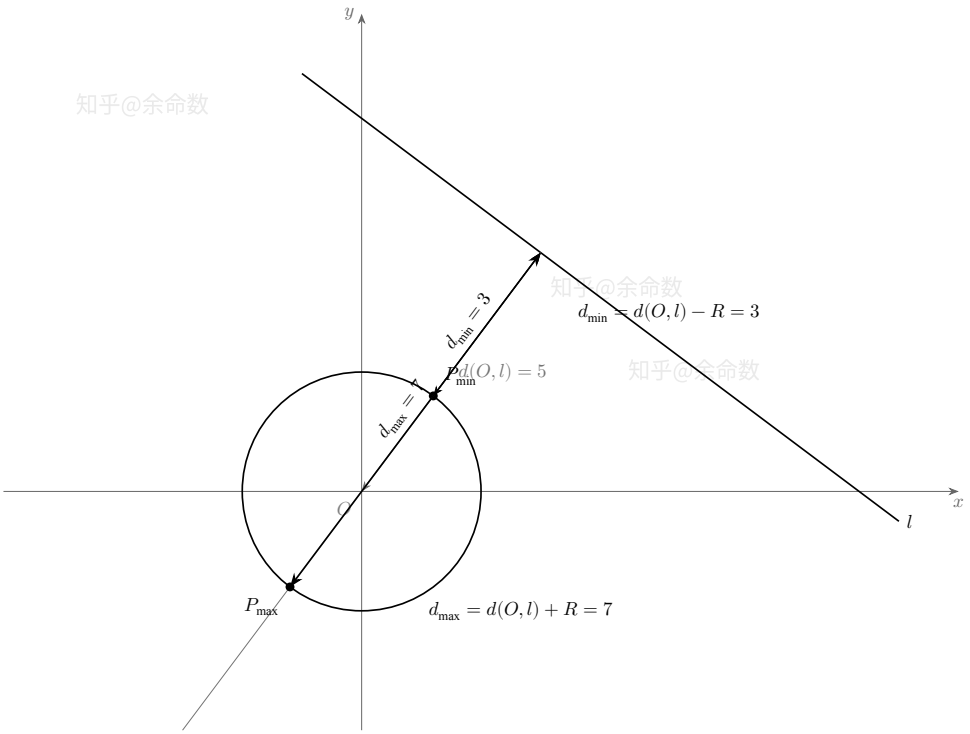


图 14.3 圆上点到直线距离的最值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

解 法一采用几何法. 圆心 $O(0,0)$ 到直线 $l:3x+4y-25=0$ 的距离为 $d=\frac{|3\cdot0+4\cdot0-25|}{\sqrt{3^2+4^2}}=\frac{25}{5}=5$, 而半径 $R=2$. 由圆的几何性质, 圆上点到直线的距离在过圆心且垂直于 l 的直径两端点处取得极值, 故

$$d_{\max}=d+R=5+2=7,$$

$$d_{\min}=d-R=5-2=3.$$

法二采用参数方程法. 此法虽计算上略繁, 但作为普适的代数方法能处理更一般情形. 设圆上点 $P(2\cos\theta,2\sin\theta)$ $(\theta\in[0,2\pi))$, 则 P 到直线的距离为

$$d(\theta)=\frac{|6\cos\theta+8\sin\theta-25|}{5}.$$

对三角部分用辅助角公式. 令 $\cos\varphi=\frac{3}{5}$, $\sin\varphi=\frac{4}{5}$, 则 $6\cos\theta+8\sin\theta=10(\cos\varphi\cos\theta+\sin\varphi\sin\theta)=10\cos(\theta-\varphi)$, 故

$$d(\theta)=\frac{|10\cos(\theta-\varphi)-25|}{5}.$$

知乎@余命数 知乎@余命数 知乎@余命数

由于 $\cos(\theta-\varphi)\in[-1,1]$, 故 $10\cos(\theta-\varphi)-25\in[-35,-15]$ 恒为负, 于是 $|10\cos(\theta-\varphi)-25|=25-10\cos(\theta-\varphi)$. 该式在 $\cos(\theta-\varphi)=-1$ 时取最大值 35, 在 $\cos(\theta-\varphi)=1$ 时取最小值 15, 从而 $d_{\max}=\frac{35}{5}=7$, $d_{\min}=\frac{15}{5}=3$, 与几何法吻合.

例题 14.6 (轨迹问题) 已知圆 $O:x^2+y^2=R^2$. 定点 A,B 在圆 O 上且关于 x 轴对称. 动点 P 是圆 O 上的任意一点. 求 $\triangle ABP$ 的垂心 H 的轨迹.

解 这是求解动点轨迹的经典问题, 纯坐标法需联立两条高线方程, 计算量巨大; 而结合向量法与参数方程则十分简洁. 关键依据是垂心的一个向量性质, 若三角形外心为 O, 则其垂心 H 满足 $\vec{OH}=\vec{OA}+\vec{OB}+\vec{OP}$, 而 P 在圆上宜用参数方程表示.

设 $A(x_0,y_0)$, 由 A,B 关于 x 轴对称知 $B(x_0,-y_0)$; 设垂心 H 的坐标为 (x,y), 即 $\vec{OH}=(x,y)$. 又 $\vec{OA}=(x_0,y_0)$, $\vec{OB}=(x_0,-y_0)$, 而动点 P 在圆上取 $P(R\cos\theta,R\sin\theta)$, 故 $\vec{OP}=(R\cos\theta,R\sin\theta)$. 代入垂心的向量关系式,

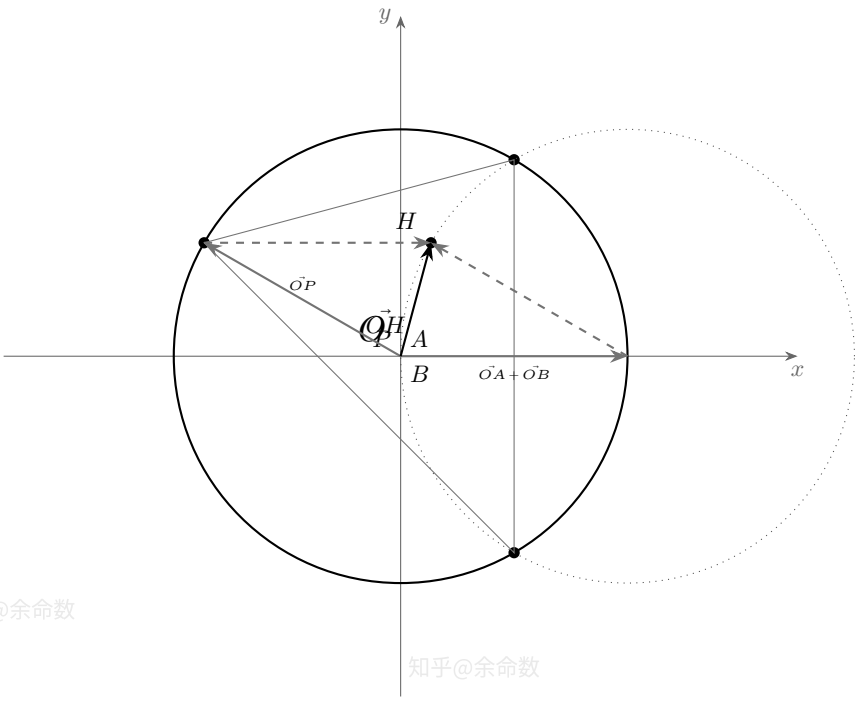
$$\begin{aligned}(x,y)&=(x_0,y_0)+(x_0,-y_0)+(R\cos\theta,R\sin\theta)\\&=(2x_0+R\cos\theta,\ R\sin\theta).\end{aligned}$$

知乎@余命数

此即垂心 H 轨迹的参数方程

$$\begin{cases}x=2x_0+R\cos\theta\\y=R\sin\theta\end{cases}\qquad(\theta\text{为参数})$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



知乎@余命数

图 14.4 垂心 H 的轨迹.

消去参数 θ 得普通方程. 由上式解出 $\cos\theta = \frac{x-2x_0}{R}$, $\sin\theta = \frac{y}{R}$, 代入 $\cos^2\theta + \sin^2\theta = 1$,

$$\left(\frac{x-2x_0}{R}\right)^2 + \left(\frac{y}{R}\right)^2 = 1 \quad \text{（代入恒等式）}$$

$$(x-2x_0)^2 + y^2 = R^2 \quad \text{（去分母）}$$

此方程表示圆心为 $(2x_0, 0)$、半径为 R 的圆，与原圆 O 全等，其圆心恰为向量 $\vec{OA} + \vec{OB}$ 的终点.

14.0.1.3.2 椭圆的参数方程 考虑标准椭圆方程 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. 受圆的参数方程启发, 尝试令

$$\frac{x}{a} = \cos\theta, \quad \frac{y}{b} = \sin\theta$$

这样自然满足 $\cos^2\theta + \sin^2\theta = 1$. 由此得到椭圆的标准参数方程.

$$\begin{cases} x = a\cos\theta \\ y = b\sin\theta \end{cases} \quad (\theta \in [0, 2\pi) \text{ 为参数}) \tag{14-3}$$

这里的参数 θ **不是**点 (x, y) 的极角. 它具有一个特殊的几何意义, 称为**偏心角**. 偏心角的几何构造如下. 作椭圆的辅助圆, 即以长轴为直径的圆 $x^2 + y^2 = a^2$. 对于

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

椭圆上任意一点 $P(x,y)$, 过 P 作 x 轴的垂线, 交辅助圆于点 Q. 连接 OQ. 射线 OQ 与 x 轴正半轴的夹角即为点 P 的偏心角 θ.

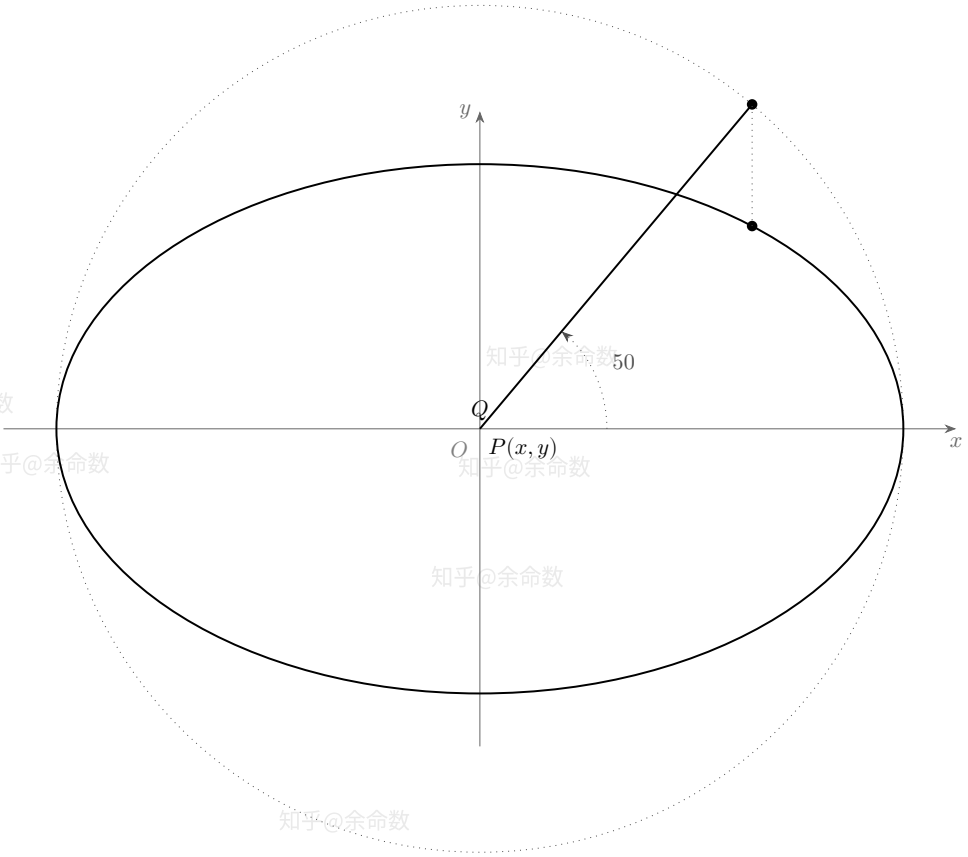


图 14.5 椭圆的偏心角 θ 的几何意义.

例题 14.7 已知椭圆 $\Gamma:\frac{x^2}{4}+\frac{y^2}{3}=1$. 对于 Γ 上的任意两点 P,Q, 我们定义一种运算 $\oplus$ 如下. 过定点 $S(1,\frac{3}{2})$ 作直线 l 平行于直线 PQ. (规定当 $P=Q$ 时, 直线 PQ 为 Γ 在点 P 处的切线). 若 l 与 Γ 有异于 S 的交点 T, 则 $P\oplus Q=T$; 否则 $P\oplus Q=S$. 已知运算 $\oplus$ 满足交换律与结合律. 记 $P_n=\underbrace{P\oplus P\oplus\cdots\oplus P}_{n个}$.

1. 若 $P(2,0),Q(-1,-\frac{3}{2})$, 求 $P\oplus Q$, P_2 以及 P_{2025}.
2. 对于 Γ 上的四点

$$P(2\cos 2\theta,\sqrt{3}\sin 2\theta),\quad Q(2\cos 2\varphi,\sqrt{3}\sin 2\varphi),$$

$$M(2\cos 2\alpha,\sqrt{3}\sin 2\alpha),\quad N(2\cos 2\beta,\sqrt{3}\sin 2\beta),$$

求证: $\overrightarrow{PQ}\parallel\overrightarrow{MN}$ 的充要条件是 $\alpha+\beta=\theta+\varphi+k\pi\ (k\in\mathbb{Z})$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

3. 是否存在异于 S 的点 P, 使得 $P_4 = S$? 若存在, 请求出所有满足条件的点 P 的坐标; 若不存在, 请说明理由.

解 (1) 先验证定点 $S(1, \frac{3}{2})$ 落在椭圆 Γ 上, 由 $\frac{1^2}{4} + \frac{(3/2)^2}{3} = \frac{1}{4} + \frac{9/4}{3} = \frac{1}{4} + \frac{3}{4} = 1$ 即得.

计算 $P \oplus Q$. 点 $P(2,0), Q(-1, -\frac{3}{2})$, 直线 PQ 的斜率 $k_{PQ} = \frac{-\frac{3}{2}-0}{-1-2} = \frac{1}{2}$. 过点 $S(1, \frac{3}{2})$ 且平行于 PQ 的直线 l 的方程为 $y - \frac{3}{2} = \frac{1}{2}(x-1)$, 即 $y = \frac{1}{2}x + 1$. 联立 l 与 Γ, 消去 y 得 $3x^2 + 4(\frac{1}{2}x+1)^2 = 12$, 整理得 $x^2 + x - 2 = 0$, 其解为 $x = 1$(对应点 S) 与 $x = -2$. 当 $x = -2$ 时 $y = \frac{1}{2}(-2) + 1 = 0$, 故异于 S 的交点为 $T(-2, 0)$, 即 $P \oplus Q = (-2, 0)$.

计算 $P_2 = P \oplus P$. 此时直线 PQ 退化为过点 $P(2,0)$ 的切线, 其方程为 $\frac{2x}{4} + \frac{0 \cdot y}{3} = 1$, 即 $x = 2$. 过 S 且平行于该切线的直线 l 为 $x = 1$, 联立 Γ 得 $\frac{1}{4} + \frac{y^2}{3} = 1 \implies y = \pm\frac{3}{2}$, 交点为 $(1, \frac{3}{2})$, 即 S 与 $(1, -\frac{3}{2})$, 故 $P_2 = (1, -\frac{3}{2})$.

为求 P_{2025}, 需探明序列 P_n 的周期性. $P_3 = P_2 \oplus P$, 直线 P_2P 的斜率 $k_{P_2P} = \frac{-3/2-0}{1-2} = \frac{3}{2}$, 过 S 的平行线为 $y = \frac{3}{2}x$, 联立 Γ 解得 $x = 1$ (S 点) 与 $x = -1$, 当 $x = -1$ 时 $y = -3/2$, 故 $P_3 = (-1, -3/2)$.

进而 $P_4 = P_3 \oplus P$, 直线 P_3P 斜率 $k_{P_3P} = \frac{-3/2-0}{-1-2} = \frac{1}{2}$, 过 S 的平行线斜率亦为 $1/2$, 与求 $P \oplus Q$ 时情形相同, 故 $P_4 = (-2, 0)$.

再 $P_5 = P_4 \oplus P$, 直线 P_4P 为 $y = 0$, 过 S 的平行线为 $y = 3/2$, 联立 Γ 解得 $x = 1$(S 点)与 $x = -1$, 故 $P_5 = (-1, 3/2)$.

继而 $P_6 = P_5 \oplus P$, 直线 P_5P 斜率 $k_{P_5P} = \frac{3/2-0}{-1-2} = -\frac{1}{2}$, 过 S 的平行线为 $y - \frac{3}{2} = -\frac{1}{2}(x-1)$, 即 $y = -\frac{1}{2}x + 2$, 联立 Γ 得 $3x^2 + 4(-\frac{1}{2}x+2)^2 = 12 \implies x^2 - 2x + 1 = 0$, 仅有重根 $x = 1$, 即直线 l 与椭圆 Γ 相切于点 S, 依定义无异于 S 的交点, 故 $P_6 = S$.

最后 $P_7 = P_6 \oplus P = S \oplus P$, 此运算涉及直线 SP, 过 S 且平行于 SP 的直线即 SP 本身, 其与 Γ 的交点为 S 与 P, 异于 S 的交点为 P, 故 $S \oplus P = P$.

至此 $P_7 = P$, 序列 P_n 以 6 为周期; 由 $2025 = 6 \times 337 + 3$ 得 $P_{2025} = P_3 = (-1, -\frac{3}{2})$.

(2) 对于椭圆 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, 参数为 u, v 的两点分别为 $(a\cos u, b\sin u)$ 与 $(a\cos v, b\sin v)$. 当两点不重合时, 其连线方向向量为

$$(a(\cos v - \cos u), b(\sin v - \sin u)) = 2\sin\frac{v-u}{2}\left(-a\sin\frac{u+v}{2}, b\cos\frac{u+v}{2}\right).$$

当两点重合时, 约定取切线方向; 而切线方向向量正为 $(-a\sin u, b\cos u)$, 也可视为上式在 $v \to u$ 时的极限. 因而一条弦或切线的方向只由端点参数和 $u + v$ 模 2π 决定.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在本题中, $a=2, b=\sqrt{3}$, 且点的参数为 $2\theta, 2\varphi$ 等形式, 因此 PQ 的方向由 $\theta+\varphi$ 决定, MN 的方向由 $\alpha+\beta$ 决定. 两方向平行的充要条件是这两个方向参数相差 π 的整数倍, 即

$$\alpha+\beta=\theta+\varphi+k\pi \quad (k\in\mathbb{Z})$$

此即**对于该椭圆, 一族平行弦的端点参数之和在模 π 意义下为常数**.

(3)

第 (2) 问的结论把复杂的几何运算 $\oplus$ 归结为参数加法, 暴露其背后的代数结构, 是解决本题的关键. 设点 $P, Q, T=P\oplus Q, S$ 对应的参数分别为 $\theta_P, \theta_Q, \theta_T, \theta_S$. 先确定特殊点 $S(1,3/2)$ 对应的参数 θ_S. 由 $2\cos(2\theta_S)=1 \implies \cos(2\theta_S)=1/2$ 与 $\sqrt{3}\sin(2\theta_S)=3/2 \implies \sin(2\theta_S)=\sqrt{3}/2$, 故 $2\theta_S=\frac{\pi}{3}+2m\pi$, 取主值得 $2\theta_S=\frac{\pi}{3}$, 即 $\theta_S=\frac{\pi}{6}$.

依运算定义, 直线 ST 平行于直线 PQ, 由 (2) 的结论, 端点参数之和模 π 相等:

$$\theta_S+\theta_T\equiv\theta_P+\theta_Q \pmod{\pi}$$

因而运算 $\oplus$ 在参数空间中的映射为

$$\theta_{P\oplus Q}=\theta_P+\theta_Q-\theta_S \pmod{\pi}$$

现求解 $P_4=S$. 设 P 的参数为 θ, 由结合律与上述加法规则,

$$\theta_{P_2}=\theta_{P\oplus P}=\theta+\theta-\theta_S=2\theta-\theta_S \pmod{\pi}$$

$$\theta_{P_3}=\theta_{P_2\oplus P}=(2\theta-\theta_S)+\theta-\theta_S=3\theta-2\theta_S \pmod{\pi}$$

$$\theta_{P_4}=\theta_{P_3\oplus P}=(3\theta-2\theta_S)+\theta-\theta_S=4\theta-3\theta_S \pmod{\pi}$$

$P_4=S$ 的充要条件是其参数 θ_4 与 θ_S 在模 π 意义下相等, 即

$$4\theta-3\theta_S=\theta_S+k\pi$$

$$4\theta=4\theta_S+k\pi \implies \theta=\theta_S+\frac{k\pi}{4}$$

代入 $\theta_S=\pi/6$ 得 $\theta=\frac{\pi}{6}+\frac{k\pi}{4}$.

寻找异于 S 的点, 即 $\theta\neq\theta_S+m\pi$, 这要求 $\frac{k\pi}{4}$ 不能是 π 的整数倍, 亦即 k 不能是 4 的倍数. 只需考虑 $k=1,2,3$ 即得所有不同的解.

当 $k=1$: $\theta=\frac{\pi}{6}+\frac{\pi}{4}=\frac{5\pi}{12} \implies 2\theta=\frac{5\pi}{6}$, P 坐标为 $(2\cos\frac{5\pi}{6}, \sqrt{3}\sin\frac{5\pi}{6})=(-\sqrt{3}, \frac{\sqrt{3}}{2})$.

当 $k=2$: $\theta=\frac{\pi}{6}+\frac{\pi}{2}=\frac{2\pi}{3} \implies 2\theta=\frac{4\pi}{3}$, P 坐标为 $(2\cos\frac{4\pi}{3}, \sqrt{3}\sin\frac{4\pi}{3})=(-1, -\frac{3}{2})$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当 $k=3$: $\theta=\frac{\pi}{6}+\frac{3\pi}{4}=\frac{11\pi}{12} \implies 2\theta=\frac{11\pi}{6}$, P 坐标为 $(2\cos\frac{11\pi}{6},\sqrt{3}\sin\frac{11\pi}{6})=(\sqrt{3},-\frac{\sqrt{3}}{2})$.

存在三个异于 S 的点 P 使得 $P_4=S$, 其坐标分别为

$$(-\sqrt{3},\frac{\sqrt{3}}{2}),\quad (-1,-\frac{3}{2}),\quad (\sqrt{3},-\frac{\sqrt{3}}{2})$$

知乎@余命数

例题 14.8 已知椭圆 $C:\frac{x^2}{4}+\frac{y^2}{9}=1$ 以及定点 $A(-2,0)$. 过点 $D(-2,3)$ 的动直线与椭圆 C 交于 P,Q 两点. 直线 AP,AQ 分别与 y 轴交于 M,N 两点. 证明:线段 MN 的中点为定点, 并求出该定点的坐标.

解 本题涉及动直线与椭圆的交点及由此引发另一族直线的截距, 最终归结为证明一个中点为定点. 传统的联立方程与韦达定理法可行, 但点 $A(-2,0)$ 恰好是椭圆的左顶点, 这提示存在一种更精细、计算更简洁的参数化方法.

设连接顶点 A 与椭圆上任意一点 P 的直线 AP 的斜率 t 为参数. 这种斜率参数法能把椭圆上除 A 之外的所有点与全体实数 t 建立一一对应, 并把点的坐标写成 t 的有理分式, 简化后续的代数运算.

设 t 为直线 AP 的斜率, 则直线 AP 的方程为 $y=t(x+2)$. 代入椭圆 C 的方程 $9x^2+4y^2=36$, 整理得关于 x 的一元二次方程

$$\begin{aligned} 9x^2+4(t(x+2))^2&=36\\ 9x^2+4t^2(x^2+4x+4)&=36\\ (9+4t^2)x^2+16t^2x+(16t^2-36)&=0. \end{aligned}$$

知乎@余命数

此方程的两根为直线 AP 与椭圆 C 交点的横坐标, 其一必为点 A 的横坐标 $x_A=-2$. 设另一交点 P 的横坐标为 x_P. 由韦达定理, 两根之积为 $x_A\cdot x_P=(-2)\cdot x_P=\frac{16t^2-36}{9+4t^2}$, 解得

$$\begin{aligned} x_P&=\frac{18-8t^2}{9+4t^2},\\ y_P&=t(x_P+2)=t\left(\frac{18-8t^2}{9+4t^2}+2\right)=t\left(\frac{18-8t^2+18+8t^2}{9+4t^2}\right)=\frac{36t}{9+4t^2}. \end{aligned}$$

此即椭圆上任意点 P(除 A 外)与参数 t 的对应关系.

知乎@余命数

设直线 DPQ 交椭圆于 P,Q 两点, 直线 AP 的斜率为 t_1, 直线 AQ 的斜率为 t_2, 则 P,Q 的坐标分别由 t_1,t_2 表出. 由点 $D(-2,3),P,Q$ 三点共线, 知斜率 k_{DP} 与 k_{DQ} 相等.

$$\begin{aligned} k_{DP}=\frac{y_P-3}{x_P-(-2)}&=\frac{\dfrac{36t_1}{9+4t_1^2}-3}{\dfrac{18-8t_1^2}{9+4t_1^2}+2}\\ &=\frac{36t_1-3(9+4t_1^2)}{18-8t_1^2+2(9+4t_1^2)}=\frac{36t_1-27-12t_1^2}{36}\\ &=t_1-\frac{1}{3}t_1^2-\frac{3}{4} \end{aligned}$$

因 P,Q 共线, 故 $k_{DP}=k_{DQ}$, 即

$$\begin{aligned} t_1-\frac{1}{3}t_1^2-\frac{3}{4}&=t_2-\frac{1}{3}t_2^2-\frac{3}{4}\\ t_1-t_2=\frac{1}{3}(t_1^2-t_2^2)&=\frac{1}{3}(t_1-t_2)(t_1+t_2). \end{aligned}$$

由 $P\neq Q$ 知 $t_1\neq t_2$, 可约去 t_1-t_2, 得 $1=\frac{1}{3}(t_1+t_2)$, 即 $t_1+t_2=3$. 此简洁关系式是解题的关键.

直线 AP 的方程为 $y=t_1(x+2)$, 与 y 轴的交点 M 为 $(0,2t_1)$; 直线 AQ 的方程为 $y=t_2(x+2)$, 与 y 轴的交点 N 为 $(0,2t_2)$. 故线段 MN 的中点 T 的坐标为

$$T\left(\frac{0+0}{2},\frac{2t_1+2t_2}{2}\right)=(0,t_1+t_2).$$

代入 $t_1+t_2=3$ 得 $T(0,3)$. 因此线段 MN 的中点是与动直线 DPQ 位置无关的定点, 坐标为 $(0,3)$.

例题 14.9 如图 14.6 所示,已知菱形 $ABCD$ 是椭圆 $C:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,(a>b>0)$ 的内接四边形.

1. 求证: $\frac{1}{OA^2}+\frac{1}{OB^2}$ 为定值;
2. 求菱形 $ABCD$ 面积的最大值.

解 (1) 证明 $\frac{1}{OA^2}+\frac{1}{OB^2}$ 为定值.

菱形 $ABCD$ 内接于中心在原点 $O(0,0)$ 的椭圆, 其中心必为椭圆中心, 设菱形中心为 M, 四个顶点可写成 $M\pm u, M\pm v$. 将 $M+u$ 与 $M-u$ 代入椭圆对应的二次型方程并相减, 得 M 与 u 在该二次型意义下正交; 同理 M 与 v 正交. 由 u,v 不共线, 只能有 $M=O$. 故菱形的对角线 AC 和 BD 均过原点 O, 又菱形对角线互相垂直, 因此 OA 与 OB 互相垂直.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

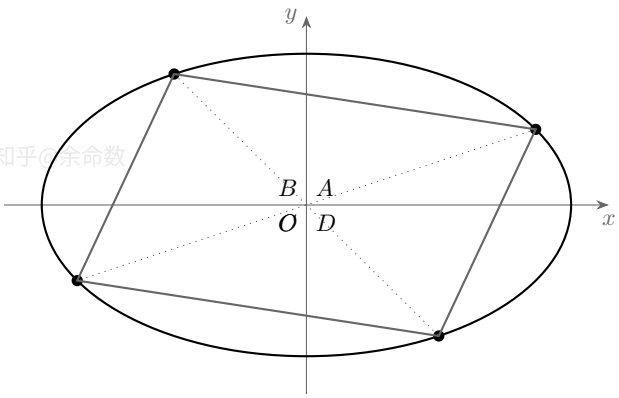


图 14.6 椭圆内接菱形 $ABCD$.

引入极角参数化描述顶点位置. 设 $m = OA$, A 的极角为 α, 则 $A(m\cos\alpha, m\sin\alpha)$; 由 $OA \perp OB$, 设 $n = OB$, 取 B 的极角为 $\alpha + \frac{\pi}{2}$, 则 $B(n\cos(\alpha+\frac{\pi}{2}), n\sin(\alpha+\frac{\pi}{2})) = (-n\sin\alpha, n\cos\alpha)$.

将 A 代入椭圆方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 并整理得 $\frac{1}{m^2}$;同理由 B 得 $\frac{1}{n^2}$.

$$\frac{(m\cos\alpha)^2}{a^2}+\frac{(m\sin\alpha)^2}{b^2}=1 \quad\Longrightarrow\quad \frac{1}{m^2}=\frac{\cos^2\alpha}{a^2}+\frac{\sin^2\alpha}{b^2},$$
$$\frac{(-n\sin\alpha)^2}{a^2}+\frac{(n\cos\alpha)^2}{b^2}=1 \quad\Longrightarrow\quad \frac{1}{n^2}=\frac{\sin^2\alpha}{a^2}+\frac{\cos^2\alpha}{b^2}.$$

两式相加即可定值.

$$\begin{aligned}\frac{1}{OA^2}+\frac{1}{OB^2}=\frac{1}{m^2}+\frac{1}{n^2}&=\left(\frac{\cos^2\alpha}{a^2}+\frac{\sin^2\alpha}{b^2}\right)+\left(\frac{\sin^2\alpha}{a^2}+\frac{\cos^2\alpha}{b^2}\right)\\&=\frac{\cos^2\alpha+\sin^2\alpha}{a^2}+\frac{\sin^2\alpha+\cos^2\alpha}{b^2}=\frac{1}{a^2}+\frac{1}{b^2}\quad(\sin^2\alpha+\cos^2\alpha=1)\end{aligned}$$

a, b 为椭圆半轴长, 系固定值, 故 $\frac{1}{OA^2}+\frac{1}{OB^2}$ 是与参数 α 无关的定值.

(2) 求菱形 $ABCD$ 面积的最大值.

菱形对角线长 $d_1 = AC = 2m, d_2 = BD = 2n$, 面积 $S = \frac{1}{2}d_1d_2 = \frac{1}{2}(2m)(2n) = 2mn$. 求 S 最大值等价于求 mn 最大值, 亦即最小化 $\frac{1}{(mn)^2}$. 将第一问所得 $\frac{1}{m^2}$ 与 $\frac{1}{n^2}$ 相乘并展开,

$$\begin{aligned}\frac{1}{(mn)^2}=\left(\frac{1}{m^2}\right)\left(\frac{1}{n^2}\right)&=\left(\frac{\cos^2\alpha}{a^2}+\frac{\sin^2\alpha}{b^2}\right)\left(\frac{\sin^2\alpha}{a^2}+\frac{\cos^2\alpha}{b^2}\right)\\&=\frac{\cos^2\alpha\sin^2\alpha}{a^4}+\frac{\cos^4\alpha+\sin^4\alpha}{a^2b^2}+\frac{\sin^2\alpha\cos^2\alpha}{b^4}.\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

利用三角恒等式 $\cos^2\alpha\sin^2\alpha=\frac{1}{4}\sin^2(2\alpha)$ 与 $\cos^4\alpha+\sin^4\alpha=1-\frac{1}{2}\sin^2(2\alpha)$ 化简，并重组含 $\sin^2(2\alpha)$ 的项，

$$\begin{aligned}\frac{1}{(mn)^2}&=\frac{1}{4}\sin^2(2\alpha)\left(\frac{1}{a^4}+\frac{1}{b^4}\right)+\frac{1}{a^2b^2}\left(1-\frac{1}{2}\sin^2(2\alpha)\right)\\&=\frac{1}{4}\sin^2(2\alpha)\left(\frac{1}{a^4}+\frac{1}{b^4}-\frac{2}{a^2b^2}\right)+\frac{1}{a^2b^2}\\&=\frac{1}{4}\sin^2(2\alpha)\left(\frac{1}{a^2}-\frac{1}{b^2}\right)^2+\frac{1}{a^2b^2}\quad\text{（配方）}\end{aligned}$$

因 $\sin^2(2\alpha)\in[0,1]$，且由 $a>b>0$ 知 $\frac{1}{a^2}-\frac{1}{b^2}\neq0$，故 $\left(\frac{1}{a^2}-\frac{1}{b^2}\right)^2$ 为正. 要使 $\frac{1}{(mn)^2}$ 最小，需 $\sin^2(2\alpha)=0$，此时 $\frac{1}{(mn)^2}$ 取最小值 $\frac{1}{a^2b^2}$，即 $(mn)^2$ 最大值为 a^2b^2，故 mn 最大值为 $\sqrt{a^2b^2}=ab$.

此情形发生于 $2\alpha=k\pi$ $(k\in\mathbb{Z})$，即 $\alpha=0$ 或 $\alpha=\frac{\pi}{2}$ 时.

- 若 $\alpha=0$，点 A 在 x 轴上，即 $A(m,0)$. 代入椭圆方程得 $\frac{m^2}{a^2}=1\implies m=a$；此时 B 在 y 轴上，即 $B(0,n)$，代入椭圆方程得 $\frac{n^2}{b^2}=1\implies n=b$，故 $mn=ab$.
- 若 $\alpha=\frac{\pi}{2}$，点 A 在 y 轴上，即 $A(0,m)$. 代入椭圆方程得 $\frac{m^2}{b^2}=1\implies m=b$；此时 B 在 x 轴上，即 $B(-n,0)$，代入椭圆方程得 $\frac{n^2}{a^2}=1\implies n=a$，故 $mn=ba$.

此两种情形下菱形对角线与椭圆轴重合，mn 最大值均为 ab. 因此菱形 $ABCD$ 面积最大值为 $S_{\max}=2(mn)_{\max}=2ab$.

传统坐标几何方法直接联立直线与椭圆的方程求解此问题.

解 (1) 证明定值. 设直线 OA 的斜率为 k，当 OA 与 x 轴重合时 $k=0$；当 OA 与 y 轴重合时，可视为 $k\to\infty$ 的极限情况，或单独讨论，则直线 OA 的方程为 $y=kx$. 与椭圆方程 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 联立并消去 y，得

$$\frac{x^2}{a^2}+\frac{(kx)^2}{b^2}=1\implies x^2\left(\frac{1}{a^2}+\frac{k^2}{b^2}\right)=1\implies x^2\left(\frac{b^2+a^2k^2}{a^2b^2}\right)=1.$$

解得点 A 的横坐标平方 $x_A^2=\frac{a^2b^2}{b^2+a^2k^2}$，从而 $y_A^2=k^2x_A^2=\frac{a^2b^2k^2}{b^2+a^2k^2}$，故

$$OA^2=x_A^2+y_A^2=\frac{a^2b^2}{b^2+a^2k^2}+\frac{a^2b^2k^2}{b^2+a^2k^2}=\frac{a^2b^2(1+k^2)}{b^2+a^2k^2}.$$

由 $OB\perp OA$，直线 OB 的斜率为 $-\frac{1}{k}$，将 k 替换为 $-\frac{1}{k}$ 即得

$$OB^2=\frac{a^2b^2(1+(-\frac{1}{k})^2)}{b^2+a^2(-\frac{1}{k})^2}=\frac{a^2b^2(1+\frac{1}{k^2})}{b^2+\frac{a^2}{k^2}}=\frac{a^2b^2\frac{k^2+1}{k^2}}{\frac{b^2k^2+a^2}{k^2}}=\frac{a^2b^2(k^2+1)}{a^2+b^2k^2}.$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

二者倒数之和为

$$\frac{1}{OA^2}+\frac{1}{OB^2}=\frac{b^2+a^2k^2}{a^2b^2(1+k^2)}+\frac{a^2+b^2k^2}{a^2b^2(1+k^2)}$$
$$=\frac{(b^2+a^2)+(a^2+b^2)k^2}{a^2b^2(1+k^2)}=\frac{(a^2+b^2)(1+k^2)}{a^2b^2(1+k^2)}=\frac{a^2+b^2}{a^2b^2}=\frac{1}{a^2}+\frac{1}{b^2}.$$

计算过程虽涉及繁琐代数运算, 最终仍能证明该值为定值.

(2) 求最大面积. 菱形面积 $S=2\cdot OA\cdot OB$,为求其最大值考察 $S^2=4\cdot OA^2\cdot OB^2$, 即

$$S^2=4\cdot\frac{a^2b^2(1+k^2)}{b^2+a^2k^2}\cdot\frac{a^2b^2(1+k^2)}{a^2+b^2k^2}=\frac{4a^4b^4(1+k^2)^2}{(b^2+a^2k^2)(a^2+b^2k^2)}.$$

这是关于斜率 k 的复杂函数. 为求其最值, 令 $u=k^2$ $(u\geq 0)$, 考察函数 $f(u)=\frac{4a^4b^4(1+u)^2}{(b^2+a^2u)(a^2+b^2u)}$ 的单调性, 但求导过程相当繁琐.

更直接的思路是观察分母 $(b^2+a^2u)(a^2+b^2u)=a^2b^2+(a^4+b^4)u+a^2b^2u^2$. 要使 $f(u)$ 最大, 等价于使其分母相对于分子 $(1+u)^2$ 增长最慢. 可验证, 当 $u=0$ 即 $k=0$, 菱形对角线与坐标轴重合, 或 $u\to\infty$ 即 $k\to\infty$ 时函数取最大值. 当 $k=0$ 时 $OA=a$, $OB=b$, 面积 $S=2ab$; 当 $k\to\infty$ 时 $OA=b$, $OB=a$, 面积 $S=2ab$. 因此最大面积为 $2ab$.

对比之下, 参数坐标法把点的位置与角度 α 直接关联, 代数结构更对称、更清晰, 用三角恒等式即可完成简化, 避开了斜率 k 及其趋于无穷的复杂情况.

14.0.1.3.3 双曲线的参数方程 对于标准双曲线方程 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, 需寻找满足 $u^2-v^2=1$ 的函数对. 一种选择是利用三角恒等式 $\sec^2\theta-\tan^2\theta=1$. 令 $\frac{x}{a}=\sec\theta$, $\frac{y}{b}=\tan\theta$, 得到双曲线的三角参数方程.

$$\begin{cases}x=a\sec\theta\\ y=b\tan\theta\end{cases}\quad(\theta\text{ 为参数})\tag{14-4}$$

其中, 当 $\theta\in(-\frac{\pi}{2},\frac{\pi}{2})$ 时, 表示双曲线的右支; 当 $\theta\in(\frac{\pi}{2},\frac{3\pi}{2})$ 时, 表示左支.

另一种方法是使用双曲函数. 定义双曲余弦 $\cosh t=\frac{e^t+e^{-t}}{2}$ 和双曲正弦 $\sinh t=\frac{e^t-e^{-t}}{2}$. 它们满足基本恒等式 $\cosh^2 t-\sinh^2 t=1$. 由此可得双曲线的双曲参数方程.

知乎@余命数

$$\begin{cases}x=a\cosh t\\ y=b\sinh t\end{cases}\quad(t\in\mathbb{R}\text{ 为参数})\tag{14-5}$$

此参数方程仅表示双曲线的右支. 左支的参数方程为 $x=-a\cosh t, y=b\sinh t$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

14.0.1.3.4 抛物线的参数方程 对于标准抛物线方程 $y^2=2px\ (p>0)$, 可以直接选取一个坐标作为参数, 例如令 $y=t$, 则 $x=\frac{t^2}{2p}$. 然而, 为了避免在处理中出现分数, 更常用且形式更整齐的参数化是令 $y=2pt$. 代入原方程 $(2pt)^2=2px \implies 4p^2t^2=2px \implies x=2pt^2$.

$$\begin{cases} x=2pt^2 \\ y=2pt \end{cases} \quad (t\in\mathbb{R}\text{ 为参数}) \tag{14-6}$$

此参数 t 的几何意义是过抛物线上一点 (x,y) 与顶点 $(0,0)$ 的直线 OP 的斜率 $k_{OP}=\frac{y}{x}=\frac{2pt}{2pt^2}=\frac{1}{t}$. 因此参数 t 是过该点与原点连线斜率的倒数.

例题 14.10 在抛物线 $y^2=2px\ (p>0)$ 的对称轴(x 轴)正方向上是否存在一点 K, 使得对于经过 K 点的抛物线的任意一条弦 AB, 总有 $\frac{1}{|KA|^2}+\frac{1}{|KB|^2}$ 为定值? 证明你的结论.

知乎@余命数

解 本题是抛物线焦弦性质的拓展, 判断是否存在一个特殊点使所有过该点的弦都满足一个特定的代数关系. 这类定点与变动弦长度关系的问题, 转换到极坐标系下往往能简化运算.

假设在 x 轴正半轴上存在这样一点 $K(x_0,0)\ (x_0>0)$, 以 K 为极点、x 轴正方向为极轴建立新极坐标系. 设极坐标 (ρ,θ) 与原直角坐标 (x,y) 的转换关系为 $x=x_0+\rho\cos\theta,\ y=\rho\sin\theta$. 代入抛物线方程 $y^2=2px$, 整理得关于极径 ρ 的一元二次方程

$$(\rho\sin\theta)^2=2p(x_0+\rho\cos\theta)$$

$$\rho^2\sin^2\theta-2p\rho\cos\theta-2px_0=0.$$

知乎@余命数

任意一条过点 K 的弦 AB 由固定的极角 θ 及 $\theta+\pi$ 的直线表示, 实际形成两点弦时需 $\sin\theta\neq0$, 其与抛物线的两交点 A,B 到极点 K 的距离 $|KA|,|KB|$ 对应上述二次方程两根的绝对值. 设两根为 ρ_1,ρ_2, 由韦达定理,

$$\rho_1+\rho_2=\frac{2p\cos\theta}{\sin^2\theta},\quad \rho_1\rho_2=-\frac{2px_0}{\sin^2\theta}.$$

考察 $\frac{1}{|KA|^2}+\frac{1}{|KB|^2}=\frac{1}{\rho_1^2}+\frac{1}{\rho_2^2}=\frac{(\rho_1+\rho_2)^2-2\rho_1\rho_2}{(\rho_1\rho_2)^2}$, 代入韦达定理结果,

知乎@余命数

$$\begin{aligned}\frac{1}{|KA|^2}+\frac{1}{|KB|^2}&=\frac{\left(\frac{2p\cos\theta}{\sin^2\theta}\right)^2-2\left(-\frac{2px_0}{\sin^2\theta}\right)}{\left(-\frac{2px_0}{\sin^2\theta}\right)^2}\\&=\frac{\frac{4p^2\cos^2\theta}{\sin^4\theta}+\frac{4px_0}{\sin^2\theta}}{\frac{4p^2x_0^2}{\sin^4\theta}}\end{aligned}$$

知乎@余命数

$$= \frac{4p^2\cos^2\theta + 4px_0\sin^2\theta}{4p^2x_0^2}$$
$$= \frac{p\cos^2\theta + x_0\sin^2\theta}{px_0^2}$$

为使其与变动角度 θ 无关, 将 $\sin^2\theta$ 替换为 $1-\cos^2\theta$ 得 $\frac{(p-x_0)\cos^2\theta+x_0}{px_0^2}$. 此分式不随 θ 变化的唯一可能是分子中 $\cos^2\theta$ 项系数为零, 故 $p-x_0=0$, 即 $x_0=p$. 此时表达式变为

$$\frac{(p-p)\cos^2\theta+p}{p\cdot p^2} = \frac{p}{p^3} = \frac{1}{p^2},$$

确与 θ 无关.

因此在抛物线 $y^2=2px$ 的对称轴正向上存在唯一点 $K(p,0)$, 使对任何过 K 的弦 AB 都有 $\frac{1}{|KA|^2}+\frac{1}{|KB|^2}=\frac{1}{p^2}$ 为定值.

下列方法沿用传统的设直线斜率、联立方程与韦达定理求解. 其计算过程更为繁复, 可作为对比, 以体会不同数学工具的效率差异.

解 假设在 x 轴正半轴上存在满足条件的点 $K(x_0,0)$ $(x_0>0)$. 设过点 K 的任意直线 l 的方程为 $y=m(x-x_0)$. 将 $x=\frac{y}{m}+x_0$ 代入抛物线方程 $y^2=2px$ 消去 x, 整理得关于 y 的一元二次方程

$$y^2 = 2p\left(\frac{y}{m}+x_0\right)$$
$$y^2 - \frac{2p}{m}y - 2px_0 = 0.$$

其两根 y_1, y_2 分别为点 A, B 的纵坐标, 由韦达定理 $y_1+y_2=\frac{2p}{m}$, $y_1y_2=-2px_0$.

表示弦长 $|KA|$ 与 $|KB|$. 对点 $A(x_1,y_1)$, 其到 $K(x_0,0)$ 的距离平方为 $|KA|^2=(x_1-x_0)^2+y_1^2$; 由 A 在直线 l 上, $y_1=m(x_1-x_0)$, 故 $x_1-x_0=\frac{y_1}{m}$, 代入得

$$|KA|^2 = \left(\frac{y_1}{m}\right)^2 + y_1^2 = y_1^2\left(\frac{1}{m^2}+1\right) = y_1^2\frac{1+m^2}{m^2},$$

同理 $|KB|^2 = y_2^2\frac{1+m^2}{m^2}$. 故目标表达式为

$$\frac{1}{|KA|^2}+\frac{1}{|KB|^2} = \frac{m^2}{y_1^2(1+m^2)}+\frac{m^2}{y_2^2(1+m^2)} = \frac{m^2}{1+m^2}\left(\frac{1}{y_1^2}+\frac{1}{y_2^2}\right)$$
$$= \frac{m^2}{1+m^2}\left(\frac{y_1^2+y_2^2}{(y_1y_2)^2}\right).$$

用韦达定理表示 $y_1^2+y_2^2=(y_1+y_2)^2-2y_1y_2=\left(\frac{2p}{m}\right)^2-2(-2px_0)=\frac{4p^2}{m^2}+4px_0$, 代入上式,

$$\frac{1}{|KA|^2}+\frac{1}{|KB|^2} = \frac{m^2}{1+m^2}\cdot\frac{\frac{4p^2}{m^2}+4px_0}{(-2px_0)^2}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

$$
\begin{aligned}
&= \frac{m^2}{1+m^2} \cdot \frac{4p(\frac{p}{m^2}+x_0)}{4p^2x_0^2} \\
&= \frac{m^2}{1+m^2} \cdot \frac{\frac{p+m^2x_0}{m^2}}{px_0^2} \\
&= \frac{p+m^2x_0}{(1+m^2)px_0^2}.
\end{aligned}
$$

要使此式不随斜率 m 变化, 将其视作关于 m^2 的分式函数 $\frac{x_0m^2+p}{(px_0^2)m^2+px_0^2}$. 形如 $\frac{Am+B}{Cm+D}$ 的分式函数若为常数, 须 $\frac{A}{C}=\frac{B}{D}$ $(C,D\neq 0)$. 应用于此分式得 $\frac{x_0}{px_0^2}=\frac{p}{px_0^2}$; 由 $p>0, x_0>0$ 分母不为零, 故 $x_0=p$.

尚需验证直线 AB 垂直于 x 轴(斜率 m 不存在)的特殊情形. 若直线为 $x=x_0$, 交点为 $A(x_0,\sqrt{2px_0})$ 与 $B(x_0,-\sqrt{2px_0})$, 此时 $|KA|^2=2px_0$, $|KB|^2=2px_0$, 故 $\frac{1}{|KA|^2}+\frac{1}{|KB|^2}=\frac{1}{px_0}$. 此值应与斜率存在时的定值相等, 即在 $x_0=p$ 时

$$
\frac{1}{px_0}=\frac{p+m^2x_0}{(1+m^2)px_0^2}, \quad \frac{1}{p^2}=\frac{p+m^2p}{(1+m^2)p^3}=\frac{p(1+m^2)}{p^3(1+m^2)}=\frac{1}{p^2},
$$

等式成立. 因此点 K 坐标为 $(p,0)$ 时结论对所有情形均成立.

14.0.1.3.5 摆线 当一个圆沿一条直线无滑动地滚动时, 摆线指的是圆周上一个定点的轨迹. 建立如下几何模型. 一个半径为 r 的圆在 x 轴上滚动, 其上的定点 P 最初位于原点 O. 取圆滚过的角度 θ 作为参数.

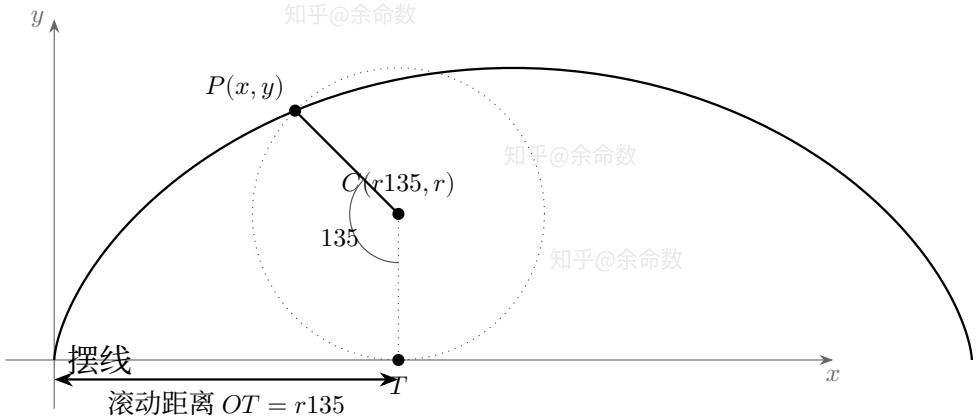


图 14.7 摆线的生成过程.

当圆心从 $(0,r)$ 移动到 $C(x_C,y_C)$ 时, 圆滚过了角度 θ.

- 圆心 C 的坐标. 由于是无滑动滚动, 圆心在 x 轴方向上移动的距离等于圆滚过的弧长, 即 $r\theta$. 圆心高度始终为 r. 所以 $C(r\theta,r)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 点 P 相对于圆心 C 的位置. 设 P 的坐标为 (x, y). 向量 $\vec{CP}$ 的起点为 C, 终点为 P. 相比于竖直向下的位置(对应初始位置), 向量 $\vec{CP}$ 旋转了 θ 角. 因此, 其分量为.

$$x_P - x_C = -r\sin\theta$$

$$y_P - y_C = -r\cos\theta$$

- 点 P 的绝对坐标.

$$x = x_C - r\sin\theta = r\theta - r\sin\theta = r(\theta - \sin\theta)$$

$$y = y_C - r\cos\theta = r - r\cos\theta = r(1 - \cos\theta)$$

最终得到摆线的参数方程.

$$\begin{cases} x = r(\theta - \sin\theta) \\ y = r(1 - \cos\theta) \end{cases} \quad (\theta \text{ 为参数}) \tag{14-7}$$

14.0.2 极坐标系

笛卡尔直角坐标系以水平和竖直的网格定位点, 这种体系在描述平移、矩形等几何结构时十分便利. 但在研究旋转或向心运动生成的轨迹时, 直角坐标的局限性便显现出来. 例如, 描述一张匀速向外扩张同时匀速旋转的蛛网, 或是行星环绕太阳的椭圆轨道, 其直角坐标方程都显得笨拙而缺乏直观.

早在 17 世纪, 艾萨克·牛顿在其《流数法》中就构想了多种坐标系, 其中之一便以角度和长度描述曲线. 此时, 瑞士数学家雅各布·伯努利在研究诸如阿基米德螺线 ($\rho = a\theta$) 和双纽线 ($\rho^2 = a^2\cos(2\theta)$) 等曲线时, 也发现用径向距离和角度表达方程, 形式简洁. 他们意识到, 对于这些以一个中心点为基准的图形, 更自然的提问是它离中心多远、朝哪个方向.

这种源于对运动和自然形态的洞察, 最终演化为一套完整的坐标体系, 即我们现在所熟知的极坐标系. 它的构建基于两个最基本的元素. 一个作为参照的中心点, 和一条作为参照方向的射线.

14.0.2.1 极坐标的定义

- **极点**: 平面内一个固定的点, 记作 O.
- **极轴**: 从极点 O 出发的一条固定的射线, 通常记作 Ox.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 14.2 (极坐标) 平面内任意一点 P 的位置, 可以由一个有序数对 (ρ,θ) 唯一确定, 其中.

- ρ 是从极点 O 到点 P 的距离, 称为点 P 的**极径**. 约定 $\rho \geq 0$.
- θ 是从极轴 Ox 逆时针旋转到射线 OP 所成的角, 称为点 P 的**极角**.

这个数对 (ρ,θ) 就称为点 P 的**极坐标**.

知乎@余命数

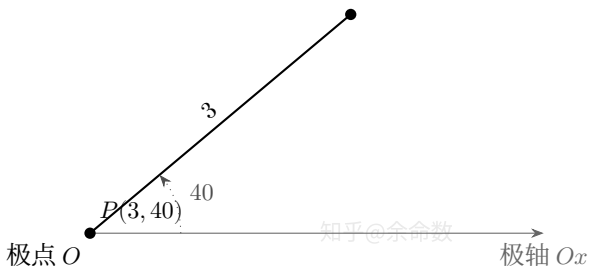


图 14.8 极坐标系的构成.

一个点的极坐标表示不是唯一的. 例如, 点 $(2,\pi/4)$ 的极角若增加 2π 的整数倍, 即 $(2,\pi/4+2k\pi), k\in\mathbb{Z}$, 它们都表示同一个点. 极点 O 的极坐标为 $(0,\theta)$, 其中 θ 可以是任意角. 在没有特殊说明时, 我们通常将极角 θ 的范围限制在 $[0,2\pi)$.

14.0.2.2 极坐标与直角坐标的互化

为了能够在两种坐标系之间灵活切换, 需要建立它们之间的代数联系. 最自然的方式是如下标准对齐.

- 将极坐标系的**极点** O 与直角坐标系的**原点** $(0,0)$ 重合.
- 将极坐标系的**极轴** Ox 与直角坐标系的 x **轴正半轴**重合.

在此设定下, 考察点 P 的两种坐标表示 (ρ,θ) 和 (x,y). 由上图中的直角三角形关系, 可以立即得到转换公式.

知乎@余命数

14.0.2.2.1 由极坐标化为直角坐标

$$x=\rho\cos\theta$$

$$y=\rho\sin\theta$$

14.0.2.2.2 由直角坐标化为极坐标

$$\rho^2=x^2+y^2 \implies \rho=\sqrt{x^2+y^2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

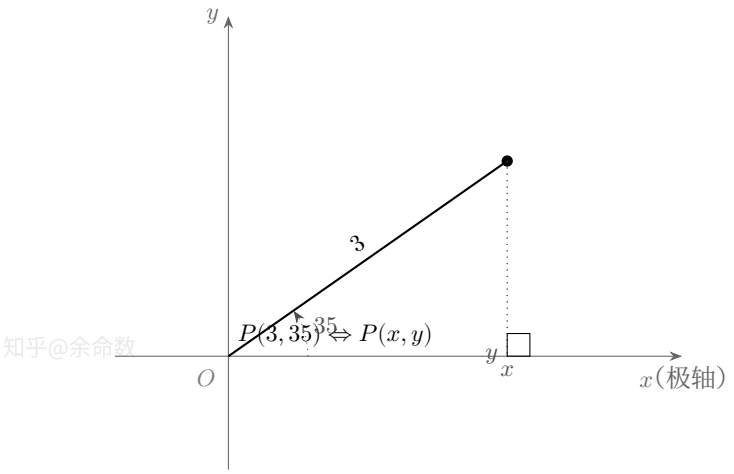


图 14.9 极坐标与直角坐标的几何关系.

$$\tan\theta = \frac{y}{x} \quad (x \neq 0)$$

在计算 θ 时，必须根据点 (x,y) 所在的象限来确定其具体值，因为反正切函数 $\arctan(\frac{y}{x})$ 的值域通常是 $(-\pi/2, \pi/2)$，无法覆盖所有象限.

例题 14.11 将极坐标为 $(4, \frac{2\pi}{3})$ 的点化为直角坐标.

解 已知 $\rho = 4, \theta = \frac{2\pi}{3}$，代入转换公式得

$$x = \rho\cos\theta = 4\cos\frac{2\pi}{3} = 4\cdot\left(-\frac{1}{2}\right) = -2,$$
$$y = \rho\sin\theta = 4\sin\frac{2\pi}{3} = 4\cdot\frac{\sqrt{3}}{2} = 2\sqrt{3}.$$

故该点的直角坐标为 $(-2, 2\sqrt{3})$.

例题 14.12 将直角坐标为 $(-1,-1)$ 的点化为极坐标(要求 $\rho > 0, \theta \in [0, 2\pi)$)
.

解 已知 $x = -1, y = -1$. 先算极径 $\rho = \sqrt{x^2+y^2} = \sqrt{(-1)^2+(-1)^2} = \sqrt{2}$，再定极角 θ. 由

$$\cos\theta = \frac{x}{\rho} = \frac{-1}{\sqrt{2}} = -\frac{\sqrt{2}}{2},$$
$$\sin\theta = \frac{y}{\rho} = \frac{-1}{\sqrt{2}} = -\frac{\sqrt{2}}{2},$$

二者均为负,故 θ 为第三象限角,满足条件者为 $\theta = \frac{5\pi}{4}$. 故该点的极坐标为 $(\sqrt{2}, \frac{5\pi}{4})$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 14.13 将下列方程在直角坐标与极坐标之间进行互化.

1. 直角坐标方程:$x^2+(y-2)^2=4$
2. 极坐标方程:$\rho=\frac{6}{2\cos\theta-3\sin\theta}$

知乎@余命数

解 (1) 将直角坐标方程 $x^2+(y-2)^2=4$ 化为极坐标方程. 展开方程得 $x^2+y^2-4y+4=4$, 即 $x^2+y^2=4y$. 以代换公式 $x^2+y^2=\rho^2$, $y=\rho\sin\theta$ 代入, 得 $\rho^2=4\rho\sin\theta$. 当 $\rho\neq0$ 时两边约去 ρ 得 $\rho=4\sin\theta$;而 $\rho=0$ 即极点这一情形已包含于 $\rho=4\sin\theta$ 中(取 $\theta=0$ 或 $\theta=\pi$ 时 $\rho=0$). 故原方程对应的极坐标方程为 $\rho=4\sin\theta$.

(2) 将极坐标方程 $\rho=\frac{6}{2\cos\theta-3\sin\theta}$ 化为直角坐标方程. 将分母移到左侧得 $\rho(2\cos\theta-3\sin\theta)=6$, 即 $2\rho\cos\theta-3\rho\sin\theta=6$. 以 $x=\rho\cos\theta$, $y=\rho\sin\theta$ 代入得 $2x-3y=6$, 此即所求直角坐标方程, 表示一条直线.

知乎@余命数

14.0.2.3 常见曲线的极坐标方程

将曲线的直角坐标方程通过代换公式转化为极坐标方程, 是研究曲线的一种方法. 更有启发性的是直接在极坐标系下, 依据曲线的几何定义来推导其方程.

14.0.2.3.1 直线的极坐标方程

1. **经过极点的直线**

   若一条直线通过极点 O, 则其上所有点(除极点外)的极角 θ 均等于一个定值 α 或 $\alpha+\pi$. 因此, 其方程极为简洁.

   $$\theta=\alpha\quad(\rho\in\mathbb{R})$$

   其中 α 是直线的倾斜角.

2. **不经过极点的直线**

   对于一条不经过极点的直线 l, 可用其法线段的性质来刻画. 从极点 O 向直线 l 作垂线, 垂足为 N. 设线段 ON 的长度为 p $(p>0)$, 且射线 ON 的极角为 α.

   设 $P(\rho,\theta)$ 是直线 l 上任意一点. 在 $\triangle ONP$ 中, $\angle ONP=\pi/2$、$ON=p$、$OP=\rho$、$\angle NOP=|\theta-\alpha|$. 由三角关系可得

   $$ON=OP\cos(\angle NOP)$$

   即

   $$p=\rho\cos(\theta-\alpha)$$

   这就是不经过极点的直线的一般极坐标方程.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

14.0.2.3.2 圆的极坐标方程

1. **圆心在极点**
   若圆心位于极点 O, 半径为 R, 则圆上任意一点的极径 ρ 恒为 R. 其方程为.

知乎@余命数

$$\rho = R$$

2. **圆经过极点**
   设一个半径为 R 的圆经过极点 O, 其圆心为 $C(R,\alpha)$. 设 $P(\rho,\theta)$ 是圆上异于 O 的任意一点. 在 $\triangle OCP$ 中, $OC=R, CP=R, OP=\rho, \angle COP=|\theta-\alpha|$. 由余弦定理, $CP^2=OC^2+OP^2-2OC\cdot OP\cos(\angle COP)$.

$$R^2=R^2+\rho^2-2R\rho\cos(\theta-\alpha)$$

   化简得 $\rho^2=2R\rho\cos(\theta-\alpha)$. 当 $\rho\neq 0$ 时,

知乎@余命数

$$\rho=2R\cos(\theta-\alpha)$$

   特别地, 若圆心在极轴上, $\alpha=0$, 方程为 $\rho=2R\cos\theta$. 若圆心在与极轴垂直且过极点的射线上, $\alpha=\pi/2$, 方程为 $\rho=2R\sin\theta$.

14.0.2.4 圆锥曲线的统一极坐标表示

极坐标系提供了一个基本而统一的视角来审视所有圆锥曲线. 把极点置于焦点, 这些看似形态各异的曲线便归入同一个简洁方程.

定理 14.1 (圆锥曲线的统一极坐标方程) 以圆锥曲线的一个焦点为极点, 以过该焦点且背离对应准线的方向为极轴. 设焦点到对应准线的距离为 p(焦参数), 离心率为 e. 则该圆锥曲线的极坐标方程为.

$$\rho=\frac{ep}{1-e\cos\theta}$$

在本章约定 $\rho\geq 0$ 的极坐标中, 该式取满足 $1-e\cos\theta>0$ 的方向; 若允许有向极径, 则同一代数式可按有向极径理解.

知乎@余命数

证明 取焦点为极点 O, 极轴正方向背离准线. 在这样的坐标中, 对应准线可写为直线 $x=-p$. 设曲线上一点 P 的极坐标为 (ρ,θ), 则 $OP=\rho$, 且 P 到准线 $x=-p$ 的距离为 $p+\rho\cos\theta$. 根据圆锥曲线的焦点-准线定义,

$$OP=e\cdot d(P,\text{准线}),$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因而

$$\rho = e(p + \rho\cos\theta).$$

移项得到 $\rho(1 - e\cos\theta) = ep$. 因 $e > 0$, $p > 0$, 且本章约定 $\rho \geq 0$, 所以实际点对应的方向满足 $1 - e\cos\theta > 0$. 在这些方向上,

$$\rho = \frac{ep}{1 - e\cos\theta}.$$

当分母趋于 0 时, 对应的是曲线沿该方向趋于无穷远的渐近方向, 与上述方程的极限解释一致. □

此方程是开普勒定律的数学基础, 它精确地描述了天体在引力场中的运行轨道. 以离心率 e 的取值为线索, 分析该方程如何演化为我们所熟知的各种圆锥曲线.

14.0.2.4.1 抛物线 ($e = 1$) 当离心率 $e = 1$ 时, 曲线为抛物线. 代入统一方程, 得其极坐标方程.

$$\rho = \frac{p}{1 - \cos\theta}$$

这个简洁的方程刻画了抛物线的几何特性.

- 当 $\theta \to 0$ 时, $\cos\theta \to 1$, 分母趋于 0, 故 $\rho \to \infty$. 这正说明抛物线开口、无限延伸.
- 当 $\theta = \pi$ 时, $\cos\theta = -1$, 此时 $\rho = \frac{p}{1-(-1)} = \frac{p}{2}$. 这是抛物线的顶点, 是曲线上离焦点(极点)最近的点.

14.0.2.4.2 椭圆 ($0 < e < 1$) 当离心率 $0 < e < 1$ 时, 曲线为椭圆. 其极坐标方程仍为.

$$\rho = \frac{ep}{1 - e\cos\theta}$$

由于 $0 < e < 1$, 无论 θ 取何值, 都有 $|e\cos\theta| < 1$, 因此分母 $1 - e\cos\theta$ 恒为正且永不等于零. 这一代数性质对应其几何特征. 极径 ρ 始终取有限的正值, 对应椭圆作为一条封闭、有界曲线.

- 当 $\theta = 0$ 时, $\cos\theta = 1$, ρ 取最大值, $\rho_{\max} = \frac{ep}{1-e}$. 此为远日点.
- 当 $\theta = \pi$ 时, $\cos\theta = -1$, ρ 取最小值, $\rho_{\min} = \frac{ep}{1+e}$. 此为近日点.

14.0.2.4.3 双曲线 ($e > 1$) 当离心率 $e > 1$ 时, 曲线为双曲线. 其极坐标方程依旧是.

$$\rho = \frac{ep}{1 - e\cos\theta}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

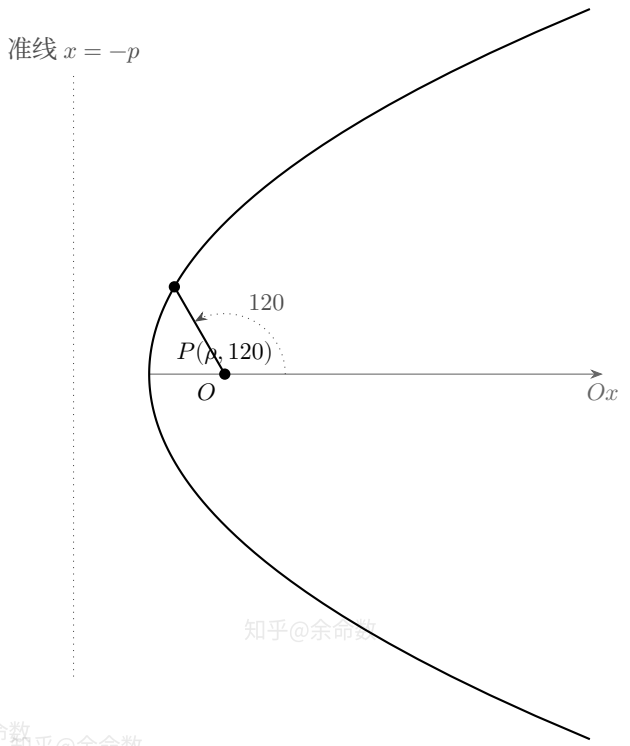


图 14.10　抛物线的极坐标表示 ($e=1$).

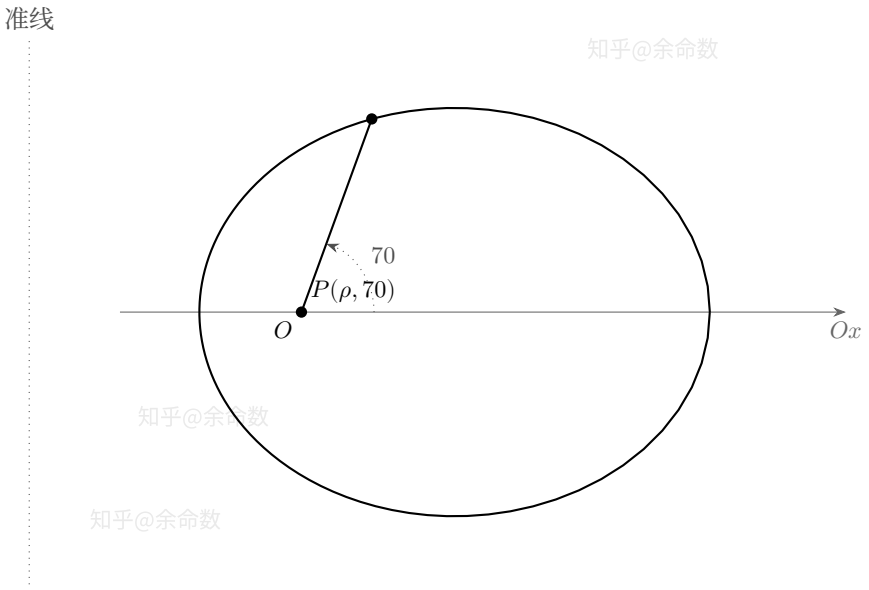


图 14.11　椭圆的极坐标表示 ($0<e<1$).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

此时由于 $e>1$, 存在特定的角 θ_0 使得 $1-e\cos\theta_0=0$, 即 $\cos\theta_0=1/e$. 在这些方向上, $\rho\to\infty$. 这是双曲线渐近线的代数来源, 对应双曲线在无穷远处无限靠近的方向.

- 当 $1-e\cos\theta>0$ 时, $\rho>0$, 描绘出双曲线靠近极点(焦点)的一支.
- 当 $1-e\cos\theta<0$ 时, $\rho<0$. 按照极坐标的约定, 负极径表示在相反方向上取长度为 $|\rho|$ 的点, 恰好描绘出双曲线的另一支.

因此, 这一个看似简单的方程同时包含双曲线的两支, 这是极坐标系的优势所在.

再看极坐标形式的圆锥曲线统一方程如何应用.

14.0.2.4.4 正焦弦长 正焦弦是指过焦点且垂直于主轴（极轴）的弦. 它的长度是圆锥曲线的一个重要内在参数. 设弦的两个端点对应的极角分别为 $\theta=\frac{\pi}{2}$ 和 $\theta=\frac{3\pi}{2}$. 当 $\theta=\frac{\pi}{2}$ 时,

$$\rho_1=\frac{ep}{1-e\cos(\frac{\pi}{2})}=ep$$

知乎@余命数
知乎@余命数

当 $\theta=\frac{3\pi}{2}$ 时,

知乎@余命数
知乎@余命数
知乎@余命数

$$\rho_2=\frac{ep}{1-e\cos(\frac{3\pi}{2})}=ep$$

因此, 正焦弦长为 $\rho_1+\rho_2=2ep$. 这个长度通常记作 $2l$. 量 $l=ep$ 被称为半正焦弦长或焦参数. 于是, 圆锥曲线的统一极坐标方程也可以写成:

$$\rho=\frac{l}{1-e\cos\theta}$$

知乎@余命数

这个形式在天体力学中尤为常用.

知乎@余命数

例题 14.14 (焦弦的调和平均性质) 设一条过圆锥曲线焦点（极点）的直线与曲线交于 P,Q 两点. 证明: 焦半径 $|OP|$ 与 $|OQ|$ 的倒数之和为常数, 即 $\frac{1}{|OP|}+\frac{1}{|OQ|}$ 为定值.

证明 设直线 PQ 与极轴的夹角为 α, 则 P 的极坐标为 (ρ_P,α), Q 的极坐标为 $(\rho_Q,\alpha+\pi)$. 由圆锥曲线的统一极坐标方程 $\rho=\frac{ep}{1-e\cos\theta}$, 两点极径分别为

$$|OP|=\rho_P=\frac{ep}{1-e\cos\alpha},\qquad |OQ|=\rho_Q=\frac{ep}{1-e\cos(\alpha+\pi)}=\frac{ep}{1+e\cos\alpha}.$$

知乎@余命数

两焦半径的倒数之和为

知乎@余命数

$$\frac{1}{|OP|}+\frac{1}{|OQ|}=\frac{1-e\cos\alpha}{ep}+\frac{1+e\cos\alpha}{ep}=\frac{(1-e\cos\alpha)+(1+e\cos\alpha)}{ep}=\frac{2}{ep}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

此结果是与角度 α 无关的常数, 等于半正焦弦长倒数的 2 倍, 亦即半正焦弦长 ep 为所有过焦点弦的两段焦半径的调和平均数. □

例题 14.15 已知行星绕太阳的椭圆轨道方程为 $\rho=\frac{A}{1+\varepsilon\cos\theta}$, 太阳位于焦点. 这里 A 为半正焦弦长, ε 为离心率, 且 θ 是从近日点开始计算的角度. 求此椭圆轨道的半长轴 a 和半短轴 b.

解 题给方程与推导出的标准形式略有差异. 标准形式 $\rho=\frac{ep}{1-e\cos\theta}$ 中 $\theta=0$ 对应远日点, 即离焦点最远的点, 而题中方程 $\rho=\frac{A}{1+\varepsilon\cos\theta}$ 在 $\theta=0$ 时分母最大、ρ 最小, 对应近日点, 这等价于将极轴旋转了 π. 令 $e=\varepsilon,\ ep=A$.

近日点距离, 即 $\theta=0$ 时, 为 $\rho_{\min}=\frac{A}{1+\varepsilon\cos(0)}=\frac{A}{1+\varepsilon}$, 远日点距离, 即 $\theta=\pi$ 时, 为 $\rho_{\max}=\frac{A}{1+\varepsilon\cos(\pi)}=\frac{A}{1-\varepsilon}$. 由椭圆几何性质, 长轴长 $2a$ 为近日点与远日点距离之和, 故

$$2a=\rho_{\min}+\rho_{\max}=\frac{A}{1+\varepsilon}+\frac{A}{1-\varepsilon}=\frac{A(1-\varepsilon)+A(1+\varepsilon)}{(1+\varepsilon)(1-\varepsilon)}=\frac{2A}{1-\varepsilon^2},$$

$$a=\frac{A}{1-\varepsilon^2}.$$

知乎@余命数 知乎@余命数 知乎@余命数

又由 $l=ep=A$ 与 $l=\frac{b^2}{a}$ 知 $A=\frac{b^2}{a}$, 故

$$b^2=aA=\left(\frac{A}{1-\varepsilon^2}\right)\cdot A=\frac{A^2}{1-\varepsilon^2},$$

$$b=\frac{A}{\sqrt{1-\varepsilon^2}}.$$

例题 14.16 求圆锥曲线 $\rho=\frac{ep}{1-e\cos\theta}$ 所有焦弦中点的轨迹方程.

解 设一条焦弦 PQ 的端点为 $P(\rho_1,\alpha)$ 和 $Q(\rho_2,\alpha+\pi)$, 其中 $\rho_1=\frac{ep}{1-e\cos\alpha},\ \rho_2=\frac{ep}{1+e\cos\alpha}$. 设 PQ 的中点为 M, 先求 M 的直角坐标. 点 P 的直角坐标为 $(\rho_1\cos\alpha,\rho_1\sin\alpha)$, 点 Q 的直角坐标为 $(\rho_2\cos(\alpha+\pi),\rho_2\sin(\alpha+\pi))=(-\rho_2\cos\alpha,-\rho_2\sin\alpha)$, 故中点 $M(x_M,y_M)$ 为

$$x_M=\frac{\rho_1\cos\alpha-\rho_2\cos\alpha}{2}=\frac{1}{2}(\rho_1-\rho_2)\cos\alpha,$$

$$y_M=\frac{\rho_1\sin\alpha-\rho_2\sin\alpha}{2}=\frac{1}{2}(\rho_1-\rho_2)\sin\alpha.$$

知乎@余命数 知乎@余命数

计算 $\rho_1-\rho_2$,

$$\rho_1-\rho_2=ep\left(\frac{1}{1-e\cos\alpha}-\frac{1}{1+e\cos\alpha}\right)=ep\left(\frac{2e\cos\alpha}{1-e^2\cos^2\alpha}\right)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{2e^2 p\cos\alpha}{1-e^2\cos^2\alpha}.$$

代入 x_M, y_M 的表达式得

$$x_M = \frac{e^2 p\cos^2\alpha}{1-e^2\cos^2\alpha}, \qquad y_M = \frac{e^2 p\sin\alpha\cos\alpha}{1-e^2\cos^2\alpha}.$$

这是以 α 为参数的参数方程, 为求其直角坐标方程需消去 α. 考察 x_M 与常数 p 的组合

$$x_M + p = \frac{e^2 p\cos^2\alpha}{1-e^2\cos^2\alpha} + p = \frac{e^2 p\cos^2\alpha + p(1-e^2\cos^2\alpha)}{1-e^2\cos^2\alpha} = \frac{p}{1-e^2\cos^2\alpha},$$

利用此关系重新表达 x_M 与 y_M 得

$$x_M = e^2\cos^2\alpha(x_M+p), \qquad y_M = e^2\sin\alpha\cos\alpha(x_M+p).$$

知乎@余命数

由第一式解出 $\cos^2\alpha = \frac{x_M}{e^2(x_M+p)}$;将第二式平方得 $y_M^2 = e^4(1-\cos^2\alpha)\cos^2\alpha(x_M+p)^2$, 代入 $\cos^2\alpha$ 的表达式化简,

知乎@余命数

$$\begin{aligned} y_M^2 &= e^4\left(1-\frac{x_M}{e^2(x_M+p)}\right)\left(\frac{x_M}{e^2(x_M+p)}\right)(x_M+p)^2 \\ &= e^4\left(\frac{e^2(x_M+p)-x_M}{e^2(x_M+p)}\right)\left(\frac{x_M}{e^2(x_M+p)}\right)(x_M+p)^2 \\ &= \left(e^2(x_M+p)-x_M\right)x_M = (e^2x_M + e^2p - x_M)x_M. \end{aligned}$$

知乎@余命数

整理得轨迹的直角坐标方程, 省略下标 M

$$(1-e^2)x^2 - e^2px + y^2 = 0.$$

知乎@余命数

此为新圆锥曲线方程, 配方可定其具体形式. 若 $e \neq 1$,

知乎@余命数

$$(1-e^2)\left(x^2 - \frac{e^2p}{1-e^2}x\right) + y^2 = 0,$$

$$(1-e^2)\left(x - \frac{e^2p}{2(1-e^2)}\right)^2 + y^2 = \frac{(1-e^2)e^4p^2}{4(1-e^2)^2} = \frac{e^4p^2}{4(1-e^2)},$$

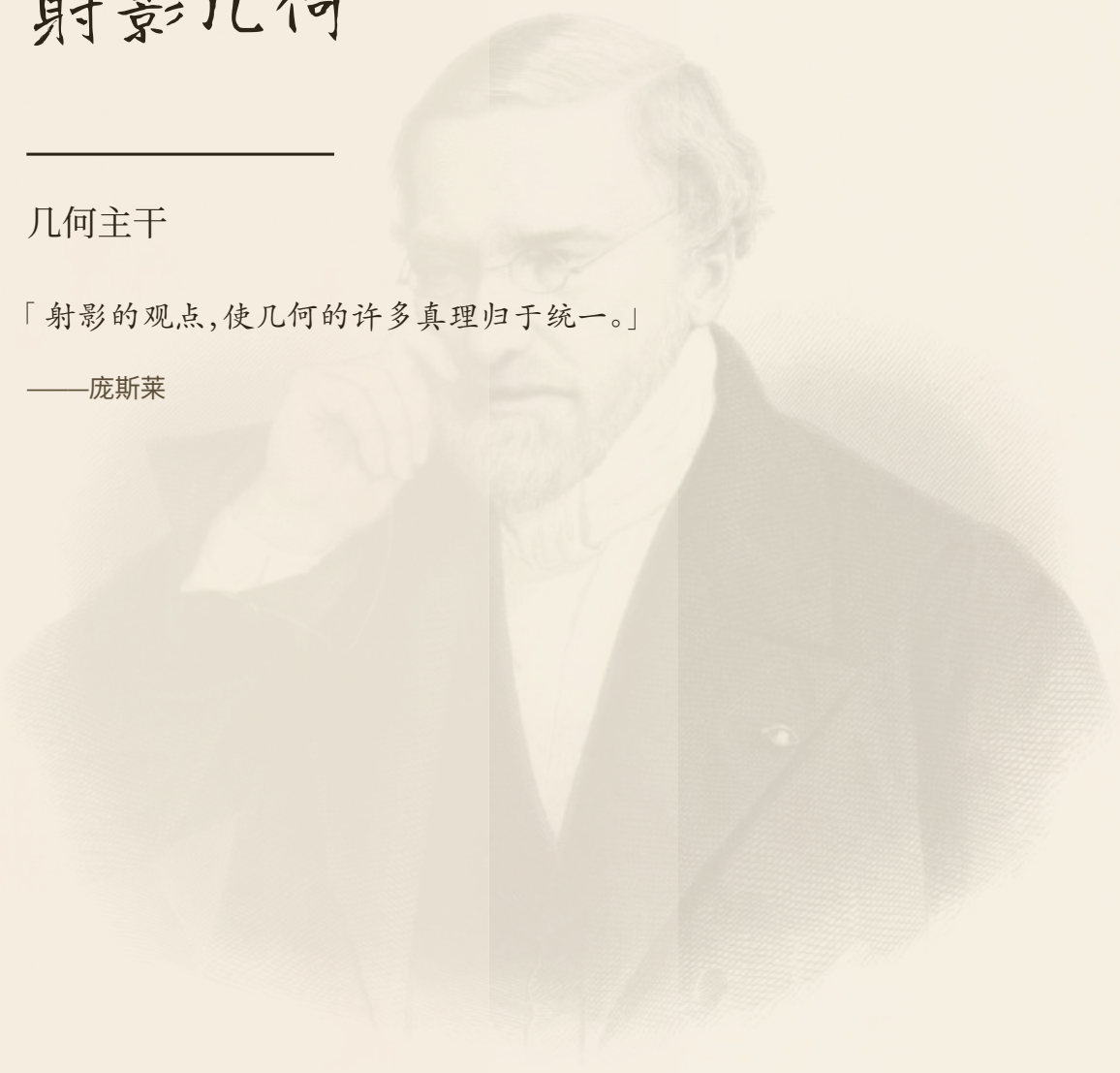
知乎@余命数

此方程表示中心位于 $(\frac{e^2p}{2(1-e^2)}, 0)$ 的圆锥曲线. 若 $e=1$(原曲线为抛物线), 方程退化为 $-px+y^2=0$, 即 $y^2=px$, 表示与原抛物线全等、开口方向相同、顶点在原点, 即原抛物线焦点的抛物线.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第 15 章

射影几何



几何主干

「射影的观点,使几何的许多真理归于统一。」

——庞斯莱

第十五章 射影几何

15.1 源起与观念

欧氏几何以直尺和圆规为基本作图工具. 圆规对应距离的度量, 长度与角度等度量性质构成欧氏几何的理论基础.

若我们舍弃圆规, 只保留能作直线的无刻度直尺, 几何学里就只剩两种基本操作, 一是过两点作一条直线, 二是求两条直线的交点.

第二个操作有一个例外, 那就是当两条直线平行时它们没有交点. 这个例外破坏了对称性. 为消除它, 射影几何的先驱者从艺术中的透视法获得启示. 在透视图中, 平行的铁轨在地平线上交于一个灭点.

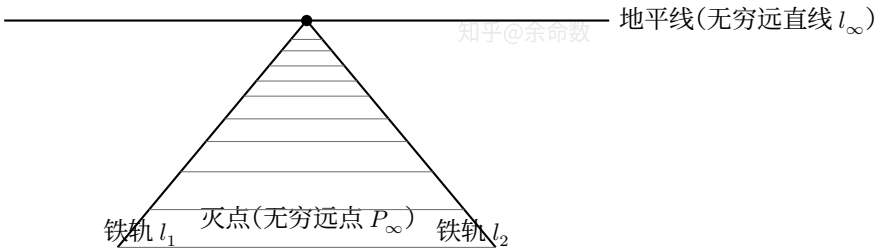


图 15.1 透视法中, 平行的直线在视觉上交于地平线上的灭点.

由此他们提出一个修正, 即假设所有平行线都相交于一个理想的点. 将这一想法系统化, 我们便得到两条规则.

- 对于平面上的每一个方向, 所有与该方向平行的直线(构成一个平行线族)都视为交于同一个**无穷远点**.
- 所有这些新增的无穷远点的集合, 构成一条特殊的直线, 称为**无穷远直线**.

为欧氏平面添加这条无穷远直线, 便得到一个新的几何空间.

定义 15.1 (射影平面) 在欧氏平面的基础上, 添加由所有无穷远点构成的无穷远直线, 所得到的扩充空间称为**射影平面**.

在射影平面中, 关于直线相交的公理变得异常简洁.

- 任意两个不同的点, 唯一确定一条穿过它们的直线.
- 任意两条不同的直线, 唯一确定一个它们的交点.

点与直线地位的这种对称性, 称为**对偶原理**.

射影几何研究在中心投影变换下保持不变的几何性质. 我们可以把中心投影设想为一个点光源 O 将一个平面 π_1 上的图形投射到另一个平面 π_2 上.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

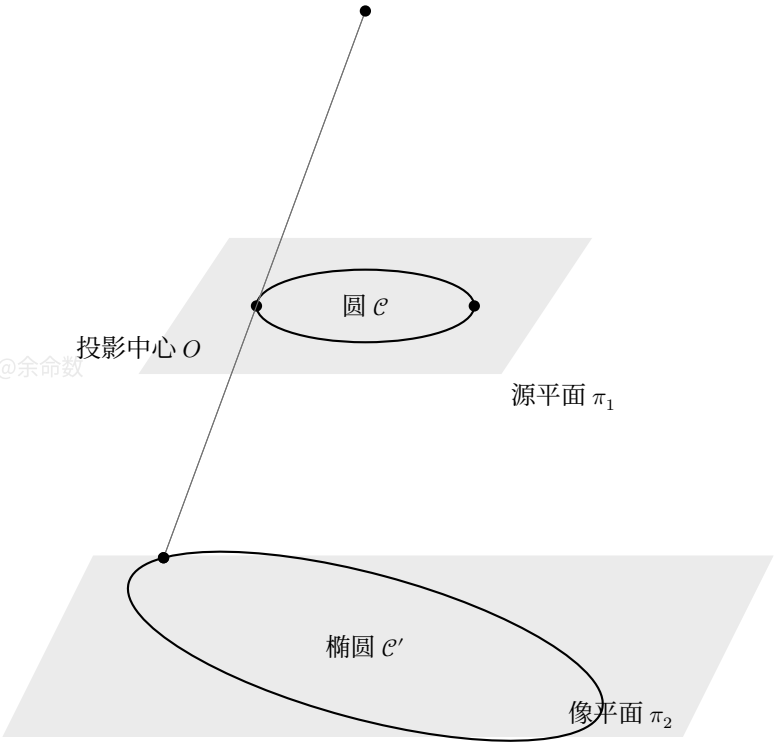


图 15.2　中心投影将一个平面上的圆 $\mathcal{C}$ 映射为另一个平面上的椭圆 $\mathcal{C}'$.

在此变换下, 长度、角度、平行关系通常都会改变. 保持不变的是关联属性, 例如点的共线性与线的共点性. 欧氏几何中性质各异的椭圆、抛物线和双曲线, 在射影几何中被视为同一种对象的不同投影, 即圆锥曲线. 我们可以用它们与无穷远直线的位置关系来区分它们.

- 椭圆与无穷远直线没有实交点.
- 抛物线与无穷远直线相切于一点.
- 双曲线与无穷远直线相交于两点.

为了在此非度量几何中进行定量分析, 我们需要寻找一个在这种投影下保持不变的数值量. 它只依赖于四个共线点的相对位置.

15.2 单比与交比

我们研究在中心投影变换下保持不变的性质. 为了进行定量分析, 我们需要寻找一个在此变换下保持不变的数值量. 长度与角度显然不满足这个要求.

15.2.1 单比

我们先考虑一个更简单的量. 给定共线三点 A, B, C.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 15.2 (单比) 给定共线三点 A, B, C, 点 C 关于基点对 (A, B) 的**单比**定义为有向线段长度之比.

$$r(A, B; C) = \frac{\overrightarrow{AC}}{\overrightarrow{BC}}$$

若在直线上建立坐标系, 设 A, B, C 的坐标为 x_A, x_B, x_C, 则

$$r(A, B; C) = \frac{x_C - x_A}{x_C - x_B}$$

这个值不依赖于坐标系原点或单位长度的选择, 是一个**仿射不变量**. 它在平移、旋转、均匀缩放等仿射变换下保持不变.

然而, 在中心投影下, 单比的值会发生改变. 考虑点列 A, B, C 从投影中心 O 投射到另一条直线上, 得到像点 A', B', C'. 一般而言, $r(A, B; C) \neq r(A', B'; C')$. 投影过程引入了一种非均匀的尺度变化, 破坏了单比的不变性.

15.2.2 交比

单比在投影下会改变, 但两个单比的比值在适当构造下可以抵消这种变化. 这一构造需要四个点.

考虑共线四点 A、B、C、D. 可用基点对 (A, B) 定义两个单比 $r(A, B; C)$ 与 $r(A, B; D)$, 它们的比值即为交比.

定义 15.3 (交比) 给定共线四点 A, B, C, D, 它们的**交比**定义为两个单比的比.

$$(A, B; C, D) = \frac{r(A, B; C)}{r(A, B; D)} = \frac{\overrightarrow{AC}/\overrightarrow{BC}}{\overrightarrow{AD}/\overrightarrow{BD}} = \frac{\overrightarrow{AC} \cdot \overrightarrow{BD}}{\overrightarrow{BC} \cdot \overrightarrow{AD}}$$

在这个定义中, (A, B) 是基准点. 我们分别用 $r(A, B; C)$ 与 $r(A, B; D)$ 度量 C 与 D 相对于这对基准点的位置, 交比就是这两个度量的比值.

中心投影对这两个单比施加相同的尺度扭曲, 取比值后这种扭曲相互抵消, 因此交比是**射影不变量**.

为了严格证明这一点, 我们将交比的概念从点列推广到线束.

定义 15.4 (线束的交比) 以点 O 为中心的四条共点直线 a, b, c, d 构成一个**线束**. 其交比 $O(a, b; c, d)$ 定义为任意一条不经过点 O 的截线 l 与线束交点 A, B, C, D 所构成的点列的交比 $(A, B; C, D)$.

此定义的合理性依赖于交比值不随截线的选择而改变.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定理 15.1 (交比的射影不变性) 对一个由四条互异直线组成的固定共点线束, 任取不经过线束中心且不平行于其中任何一条直线的截线, 所得四个有限截点的交比不依赖于截线的选择.

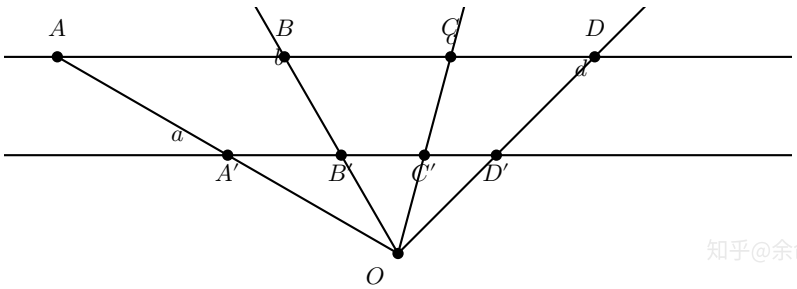


图 15.3　线束 $O(a,b,c,d)$ 被两条不同截线 l 和 l' 所截, 点列交比 $(A,B;C,D)$ 与 $(A',B';C',D')$ 相等.

证明 设线束中心为 O, 四条直线为 a,b,c,d. 截线 l 与之交于 A,B,C,D. 我们用三角形面积来表示有向线段之比.$\triangle OAC$ 和 $\triangle OBC$ 有共同的高, 因此面积比等于底边比.

$$\frac{\overrightarrow{AC}}{\overrightarrow{BC}} = \frac{\text{Area}(\triangle OAC)}{\text{Area}(\triangle OBC)}$$

利用正弦面积公式 $\text{Area}(\triangle OAC) = \frac{1}{2}OA \cdot OC \sin(\angle AOC)$, 可得

$$\frac{\overrightarrow{AC}}{\overrightarrow{BC}} = \frac{\frac{1}{2}OA \cdot OC \sin(\angle AOC)}{\frac{1}{2}OB \cdot OC \sin(\angle BOC)} = \frac{OA \cdot \sin(\angle AOC)}{OB \cdot \sin(\angle BOC)}$$

同理,

$$\frac{\overrightarrow{AD}}{\overrightarrow{BD}} = \frac{OA \cdot \sin(\angle AOD)}{OB \cdot \sin(\angle BOD)}$$

将两式代入交比定义, 便得

$$\begin{aligned}(A,B;C,D) &= \frac{\overrightarrow{AC}}{\overrightarrow{BC}} \div \frac{\overrightarrow{AD}}{\overrightarrow{BD}} \\ &= \left(\frac{OA \cdot \sin(\angle AOC)}{OB \cdot \sin(\angle BOC)}\right) \div \left(\frac{OA \cdot \sin(\angle AOD)}{OB \cdot \sin(\angle BOD)}\right) \\ &= \frac{\sin(\angle AOC) \cdot \sin(\angle BOD)}{\sin(\angle BOC) \cdot \sin(\angle AOD)}\end{aligned}$$

记直线 a,c 间的夹角为 $\angle(a,c)$, 则

$$(A,B;C,D) = \frac{\sin\angle(a,c) \cdot \sin\angle(b,d)}{\sin\angle(b,c) \cdot \sin\angle(a,d)}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由于四条直线互异, 且截线不经过线束中心也不平行于其中任何一条直线, 上述交比中的各分母均非零. 此表达式只与线束的内蕴角度有关, 与截线 l 的位置或截点 A, B, C, D 的具体坐标无关. 这就证明了交比的射影不变性. □

这个不变性是射影几何定量分析的基础. 它也可以揭示不同几何学之间的层级关系. 若点 D 为直线上的无穷远点 P_∞, 则其坐标 $x_D \to \infty$.

$$\lim_{x_D\to\infty} \frac{x_D - x_B}{x_D - x_A} = \lim_{x_D\to\infty} \frac{1 - x_B/x_D}{1 - x_A/x_D} = 1$$

因此, 交比退化为单比.

$$(A, B; C, P_\infty) = \frac{x_C - x_A}{x_C - x_B} = r(A, B; C)$$

故仿射几何的基本不变量单比, 是射影几何基本不变量交比在一个点为无穷远点时的特例. 射影几何是更普遍的框架.

15.2.3 调和共轭

在交比的所有可能取值中, -1 因其对称性与几何意义而最为特殊.

定义 15.5 (调和共轭) 若共线四点 A, B, C, D 的交比 $(A, B; C, D) = -1$, 则称该点列为**调和点列**. 此时, 称点对 (C, D) **调和分割**点对 (A, B), 或称 C 和 D 关于 A 和 B **调和共轭**.

交比为 -1 的线束相应地被称为**调和线束**.

调和共轭的定义式 $(A, B; C, D) = -1$ 等价于

$$\frac{\overrightarrow{AC}}{\overrightarrow{BC}} = -\frac{\overrightarrow{AD}}{\overrightarrow{BD}}$$

这个关系式说明了调和分割的几何内涵.

- 符号的相反意味着, 点 C 和 D 中, 一个必然是线段 AB 的内分点, 另一个是外分点.
- 表达式的绝对值相等, 即 $|\overrightarrow{AC}/\overrightarrow{BC}| = |\overrightarrow{AD}/\overrightarrow{BD}|$, 表明这两个点分割线段 AB 的比的大小是相同的.

这种一内一外、分割比大小相等的关系, 正是调和分割的几何特征.

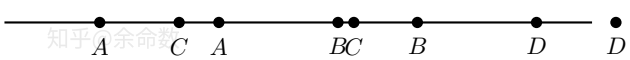


图 15.4 调和点列示意图. 点 C 内分线段 AB, 点 D 外分线段 AB.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

调和共轭的概念将欧氏几何中的一些基本观念统一到射影几何之中. 先考察 C 为线段 AB 中点的特殊情形.

若 C 是 AB 的中点, 则 $\overrightarrow{AC}=-\overrightarrow{BC}$, 因此单比 $r(A,B;C)=\frac{\overrightarrow{AC}}{\overrightarrow{BC}}=-1$. 代入调和条件 $(A,B;C,D)=\frac{r(A,B;C)}{r(A,B;D)}=-1$, 得到

$$\frac{-1}{r(A,B;D)}=-1 \implies r(A,B;D)=1$$

即 $\frac{\overrightarrow{AD}}{\overrightarrow{BD}}=1$, 或 $\overrightarrow{AD}=\overrightarrow{BD}$. 对于两个不同的点 A,B, 在欧氏平面内不存在任何有限远处的点 D 能满足此条件. 唯一的可能是点 D 位于该直线的无穷远处.

由此得一个重要结论, 即线段中点的调和共轭点是该直线上的无穷远点. 反之, 与无穷远点调和共轭的点是线段的中点.

调和点列还满足一个度量关系.

定理 15.2 (调和点列的度量关系) 若有限点列 $(A,B;C,D)$ 为调和点列, 且交比定义中涉及的有向线段均非零 (特别地, $A\neq B$ 且 C,D 均不与 A,B 重合), 则有向线段的倒数满足关系式.

$$\frac{2}{\overrightarrow{AB}}=\frac{1}{\overrightarrow{AC}}+\frac{1}{\overrightarrow{AD}}$$

其中线段的起点均取为 A.

证明 为方便计算, 我们在直线上建立坐标系, 并令点 A 为原点, 即 $x_A=0$. 设 B,C,D 的坐标分别为 b,c,d. 交比的定义式为

$$(A,B;C,D)=\frac{x_C-x_A}{x_C-x_B}\cdot\frac{x_D-x_B}{x_D-x_A}=-1$$

代入坐标, 得到

$$\frac{c-0}{c-b}\cdot\frac{d-b}{d-0}=-1$$

$$c(d-b)=-d(c-b)$$

展开得 $cd-cb=-cd+db$. 移项整理得

$$2cd=cb+db=b(c+d)$$

当 b,c,d 均不为零时, 两边同除以 bcd, 得到

$$\frac{2}{b}=\frac{c+d}{cd}=\frac{1}{d}+\frac{1}{c}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由于原点在 A, 有 $b=\overrightarrow{AB}, c=\overrightarrow{AC}, d=\overrightarrow{AD}$. 代入即得

$$\frac{2}{\overrightarrow{AB}}=\frac{1}{\overrightarrow{AC}}+\frac{1}{\overrightarrow{AD}}$$

□

这个关系表明, $\overrightarrow{AB}$ 的长度是 $\overrightarrow{AC}$ 和 $\overrightarrow{AD}$ 长度的调和平均数.

例题 15.1 在直线上给定三点 A, B, C, 其坐标分别为 $0, 5, 8$. 求点 D 的坐标, 使得 $(A, B; C, D)=-1$.

证明 设点 D 的坐标为 x, 由调和共轭条件 $(A, B; C, D)=-1$ 求其位置. 依交比定义有

$$(A, B; C, D)=\frac{x_C-x_A}{x_C-x_B}\cdot\frac{x_D-x_B}{x_D-x_A}=-1.$$

代入已知坐标 $x_A=0, x_B=5, x_C=8, x_D=x$, 整理得

$$\begin{aligned}\frac{8-0}{8-5}\cdot\frac{x-5}{x-0}&=-1,\\ \frac{8}{3}\cdot\frac{x-5}{x}&=-1,\\ 8(x-5)&=-3x,\\ 8x-40&=-3x,\\ 11x&=40,\\ x&=\frac{40}{11}.\end{aligned}$$

因此点 D 的坐标为 $40/11$.

点 A, B 的坐标为 $0, 5$, 点 C 的坐标为 8, 落在线段 AB 之外为外分点; 所求点 D 的坐标为 $40/11\approx 3.64$, 介于 A, B 之间为内分点. 一内一外的分割方式正符合调和点列的几何特征. □

15.3 二次曲线的极点与极线

调和共轭把直线几何与二次曲线几何联系起来, 并且给出二次曲线的一种内在对称性.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

15.3.1 调和分割与定比点差

先回顾调和点列的定义. 若点列 $(A,B;P,Q)$ 为调和点列, 其交比为 -1.

$$(A,B;P,Q)=\frac{\overrightarrow{AP}}{\overrightarrow{BP}}\div\frac{\overrightarrow{AQ}}{\overrightarrow{BQ}}=-1$$

即

$$\frac{\overrightarrow{AP}}{\overrightarrow{BP}}=-\frac{\overrightarrow{AQ}}{\overrightarrow{BQ}}$$

这个关系式可以用另一种方式来表述. 设存在一个常数 λ, 使得

$$\overrightarrow{AP}=\lambda\overrightarrow{PB}\quad\text{且}\quad\overrightarrow{AQ}=-\lambda\overrightarrow{QB}$$

其中 $\overrightarrow{PB}=-\overrightarrow{BP},\overrightarrow{QB}=-\overrightarrow{BQ}$. 代入后得到

$$\frac{\overrightarrow{AP}}{\overrightarrow{BP}}=-\lambda\quad\text{且}\quad\frac{\overrightarrow{AQ}}{\overrightarrow{BQ}}=\lambda$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

这两个式子合并, 恰好满足交比为 -1 的条件. 因此, 调和分割可用定比点差这一等价形式描述, 即点 P 和 Q 以大小相等、方向相反的比率分割有向线段 AB.

若设 A,B 的坐标分别为 (x_A,y_A) 和 (x_B,y_B), 则点 P,Q 的坐标可以由定比分点公式表示.

$$P=\frac{A+\lambda B}{1+\lambda}=\left(\frac{x_A+\lambda x_B}{1+\lambda},\frac{y_A+\lambda y_B}{1+\lambda}\right)$$

$$Q=\frac{A-\lambda B}{1-\lambda}=\left(\frac{x_A-\lambda x_B}{1-\lambda},\frac{y_A-\lambda y_B}{1-\lambda}\right)$$

这个代数表示给出了二次曲线上点列轨迹的研究工具.

15.3.2 极点与极线的定义与推导

在一条直线上, 调和分割给出四个点之间一种非常刚性的投影不变量, 一旦给定 $(A,B;P,Q)=-1$, 点 Q 便由 A,B,P 唯一决定, 反之亦然. 这是一维情形下点与点之间的精细对应关系.

现在把问题放到平面上的二次曲线 $\mathcal{C}$ 上. 任取一条过定点 P 的割线, 设它与 $\mathcal{C}$ 交于 A,B 两点. 由调和分割的性质, 这条割线上存在唯一的点 Q, 使 $(A,B;P,Q)$ 为调和点列. 让割线绕 P 旋转, 便得到一族点 Q.

不同割线给出不同的 A,B, 因而对应的 Q 也各不相同. 但调和分割是投影不变量, 对图形作任意投影变换, A,B,P,Q 的调和性都保持不变. 因此这些 Q 的集合应具有某种几何结构.

于是提出下列的问题.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

问题. 给定平面上的一个二次曲线 $\mathcal{C}$ 和一个定点 P. 从点 P 引出任意一条割线, 与二次曲线 $\mathcal{C}$ 交于 A, B 两点. 在这条割线上, 存在唯一的点 Q, 使得点列 $(A, B; P, Q)$ 构成调和点列. 当割线绕着点 P 转动时, 点 Q 的轨迹是什么?

所有这样的 Q 都落在同一条直线上, 这条直线只由点 P 与二次曲线 $\mathcal{C}$ 决定, 与割线选择无关. 现引入极点与极线的定义.

定理 15.3 (极点与极线定理) 对于非退化的标准椭圆、双曲线或抛物线 $\mathcal{C}$ 和一点 P,当过 P 的割线与 $\mathcal{C}$ 交于两个不同点 A, B 时,所有满足 $(A, B; P, Q) = -1$ 的有限点 Q 都落在一条固定直线上. 这条直线称为 P 关于 $\mathcal{C}$ 的**极线**, 点 P 称为该极线的**极点**.

证明 以椭圆为例证明. 设椭圆方程为

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

设极点为 $P(x_0, y_0)$, 动点为 $Q(x, y)$. 割线与椭圆的交点记为 $A(x_A, y_A)$ 与 $B(x_B, y_B)$.

根据上一节中调和点列的向量(坐标)表示, 存在常数 $\lambda(\lambda \neq \pm 1)$, 使得 P 和 Q 在割线上的坐标可以表示为

$$\begin{aligned}
x_0 = \frac{x_A + \lambda x_B}{1+\lambda} &\iff x_A + \lambda x_B = x_0(1+\lambda),\\
y_0 = \frac{y_A + \lambda y_B}{1+\lambda} &\iff y_A + \lambda y_B = y_0(1+\lambda),\\
x = \frac{x_A - \lambda x_B}{1-\lambda} &\iff x_A - \lambda x_B = x(1-\lambda),\\
y = \frac{y_A - \lambda y_B}{1-\lambda} &\iff y_A - \lambda y_B = y(1-\lambda).
\end{aligned}$$

因为点 A 和 B 均在椭圆上, 它们满足

$$\frac{x_A^2}{a^2} + \frac{y_A^2}{b^2} = 1, \tag{1}$$

$$\frac{x_B^2}{a^2} + \frac{y_B^2}{b^2} = 1. \tag{2}$$

为刻画 Q 的轨迹, 需消去与具体割线相关的量 $(x_A, y_A, x_B, y_B, \lambda)$, 得到只含 (x, y) 与定点 (x_0, y_0) 的关系式.

用 (1) 减去 λ^2 倍的 (2), 两端相减得

$$\left(\frac{x_A^2}{a^2} + \frac{y_A^2}{b^2}\right) - \lambda^2\left(\frac{x_B^2}{a^2} + \frac{y_B^2}{b^2}\right) = 1 - \lambda^2.$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

即

$$\frac{x_A^2-\lambda^2x_B^2}{a^2}+\frac{y_A^2-\lambda^2y_B^2}{b^2}=1-\lambda^2.$$

再由平方差公式分解, 得

$$\frac{(x_A-\lambda x_B)(x_A+\lambda x_B)}{a^2}+\frac{(y_A-\lambda y_B)(y_A+\lambda y_B)}{b^2}=1-\lambda^2.$$

将前面得到的关系

$$x_A-\lambda x_B=x(1-\lambda),\quad x_A+\lambda x_B=x_0(1+\lambda),$$

以及 y 的对应式代入, 得到

$$\frac{x(1-\lambda)\cdot x_0(1+\lambda)}{a^2}+\frac{y(1-\lambda)\cdot y_0(1+\lambda)}{b^2}=1-\lambda^2.$$

注意 $(1-\lambda)(1+\lambda)=1-\lambda^2$, 于是

$$(1-\lambda^2)\left(\frac{x_0x}{a^2}+\frac{y_0y}{b^2}\right)=1-\lambda^2.$$

由于割线与椭圆有两个不同交点, $\lambda\neq\pm1$, 因此 $1-\lambda^2\neq0$, 可约去, 得到

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1.$$

这是一个关于 (x,y) 的一阶线性方程, 其系数只由椭圆参数 a,b 以及定点 $P(x_0,y_0)$ 决定. 这说明, 随着割线转动, 所有满足调和条件的点 Q 都落在这条固定直线上. 对于双曲线和抛物线, 完全可以作出类似的推导. □

15.3.2.1 消元思路

(1) 与 (2) 给出两个二次约束

$$(1)\quad \frac{x_A^2}{a^2}+\frac{y_A^2}{b^2}=1,\qquad (2)\quad \frac{x_B^2}{a^2}+\frac{y_B^2}{b^2}=1;$$

调和分割则把 $x_A\pm\lambda x_B$ 与 $y_A\pm\lambda y_B$ 表示成 x_0,y_0,x,y 的一次式. 我们的目标是消去 A,B,λ, 得到只含 P,Q 的式子.

为此我们需要把二次式拆成已知一次因子的乘积.$(1)-\lambda^2(2)$ 正好产生

$$x_A^2-\lambda^2x_B^2=(x_A-\lambda x_B)(x_A+\lambda x_B)$$

与 y 的对应式, 使二次信息可以直接用线性关系代入.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其他系数选择都不能做到这一点.$(1)-(2)$ 给出 $x_A^2-x_B^2$, 其一次因子 $x_A\pm x_B$ 与已知的 $x_A\pm\lambda x_B$ 不匹配; $(1)-\lambda(2)$ 给出 $x_A^2-\lambda x_B^2$, 也不能按 $x_A\pm\lambda x_B$ 分解. 只有 λ^2 能满足平方差公式的系数要求.

于是 $(1)-\lambda^2(2)$ 把二次曲线的方程与调和分割的线性关系联系起来, 消去 A,B,λ 后得到

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1,$$

这正是 Q 所满足的线性方程.

定理 15.4 (极线方程) 设点 P 的坐标为 (x_0,y_0).

1. 若二次曲线 $\mathcal{C}$ 为椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$, 则其极线方程为.
$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1$$
2. 若二次曲线 $\mathcal{C}$ 为双曲线 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}=1$, 则其极线方程为.
$$\frac{x_0x}{a^2}-\frac{y_0y}{b^2}=1$$
3. 若二次曲线 $\mathcal{C}$ 为抛物线 $y^2=2px$, 则其极线方程为.
$$y_0y=p(x+x_0)$$

证明 椭圆情形已在极点与极线定理的证明中推出, 结果为

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1.$$

对双曲线 $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, 沿用极点与极线定理证明中的同样记号. 因为交点 A,B 在双曲线上, 有

$$\frac{x_A^2}{a^2}-\frac{y_A^2}{b^2}=1,\qquad \frac{x_B^2}{a^2}-\frac{y_B^2}{b^2}=1.$$

用第一式减去 λ^2 倍第二式, 再代入调和分割给出的 $x_A\pm\lambda x_B$ 与 $y_A\pm\lambda y_B$ 的表达, 得

$$(1-\lambda^2)\left(\frac{x_0x}{a^2}-\frac{y_0y}{b^2}\right)=1-\lambda^2.$$

由于 A,B 为两个不同交点且 Q 为有限点, $\lambda\neq\pm1$, 故可约去 $1-\lambda^2$, 得双曲线极线方程

$$\frac{x_0x}{a^2}-\frac{y_0y}{b^2}=1.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对抛物线 $y^2=2px$, 交点满足 $y_A^2=2px_A$ 与 $y_B^2=2px_B$. 仍用第一式减去 λ^2 倍第二式, 得

$$y_A^2-\lambda^2y_B^2=2p(x_A-\lambda^2x_B).$$

左边分解并代入调和分割关系, 得

$$(y_A-\lambda y_B)(y_A+\lambda y_B)=(1-\lambda^2)y_0y.$$

知乎@余命数

右端的 $x_A-\lambda^2x_B$ 可化为

$$\begin{aligned}x_A-\lambda^2x_B&=\frac{(x_A+\lambda x_B)+(x_A-\lambda x_B)}{2}-\lambda\frac{(x_A+\lambda x_B)-(x_A-\lambda x_B)}{2}\\&=\frac{1-\lambda^2}{2}(x_0+x).\end{aligned}$$

知乎@余命数

因此

$$(1-\lambda^2)y_0y=p(1-\lambda^2)(x+x_0).$$

知乎@余命数

约去 $1-\lambda^2$, 得

$$y_0y=p(x+x_0).$$

□

极线方程也可以直接由二次曲线方程作替换得到.

知乎@余命数

- $x^2\to x_0x$
- $y^2\to y_0y$
- $x\to\frac{1}{2}(x+x_0)$
- $y\to\frac{1}{2}(y+y_0)$
- $xy\to\frac{1}{2}(x_0y+y_0x)$

知乎@余命数

这个过程被称为**配极变换**或**极化**. 它统一了所有二次曲线的极线方程, 并与点 P 位于曲线上时的切线方程在形式上完全一致.

知乎@余命数

例题 15.2 过点 $P(1,1)$ 的直线交椭圆 $\dfrac{x^2}{4}+\dfrac{y^2}{2}=1$ 于 A,B 两点. 在直线 AB 上取异于 P 的点 Q, 使 P,Q 分别为 AB 的内分点与外分点, 且满足

$$\frac{|\overrightarrow{AP}|}{|\overrightarrow{PB}|}=\frac{|\overrightarrow{AQ}|}{|\overrightarrow{QB}|}.$$

求证点 Q 在某条定直线上.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 由 $\dfrac{1^2}{4}+\dfrac{1^2}{2}=\dfrac{3}{4}<1$ 知点 P 在椭圆内部, 故任意过 P 的割线交椭圆于 A,B 时 P 落在线段 AB 上, 而题设中的 Q 为对应的外分点. 题设条件

$$\frac{|\overrightarrow{AP}|}{|\overrightarrow{PB}|}=\frac{|\overrightarrow{AQ}|}{|\overrightarrow{QB}|}$$

即

$$\frac{AP}{PB}=\frac{AQ}{QB},$$

其中 P 取内分比、Q 取外分比且大小相等, 这正说明点列 $(A,B;P,Q)$ 为调和点列, 即

$$(A,B;P,Q)=-1.$$

于是 Q 是 P 在割线 AB 上的调和共轭点.

当 $P(1,1)$ 的割线绕 P 转动时, 点 Q 的轨迹即 P 关于椭圆

$$\frac{x^2}{4}+\frac{y^2}{2}=1$$

的极线, 由极点与极线定理写出其方程即可.

对一般椭圆

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,$$

上一节已经推导出点 $P(x_0,y_0)$ 的极线方程为

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1.$$

本题中 $a^2=4,b^2=2$, 极点为 $P(1,1)$, 代入可得

$$\frac{1\cdot x}{4}+\frac{1\cdot y}{2}=1.$$

两边同乘以 4, 化简为

$$x+2y=4,$$

即

$$x+2y-4=0.$$

这是一条与割线无关的定直线. 因此, 无论过 $P(1,1)$ 的直线如何转动, 只要 Q 满足题中的比例条件, 它都必在这条直线 $x+2y-4=0$ 上.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此, 点 Q 的轨迹是一条定直线, 其方程为

$$x+2y-4=0.$$

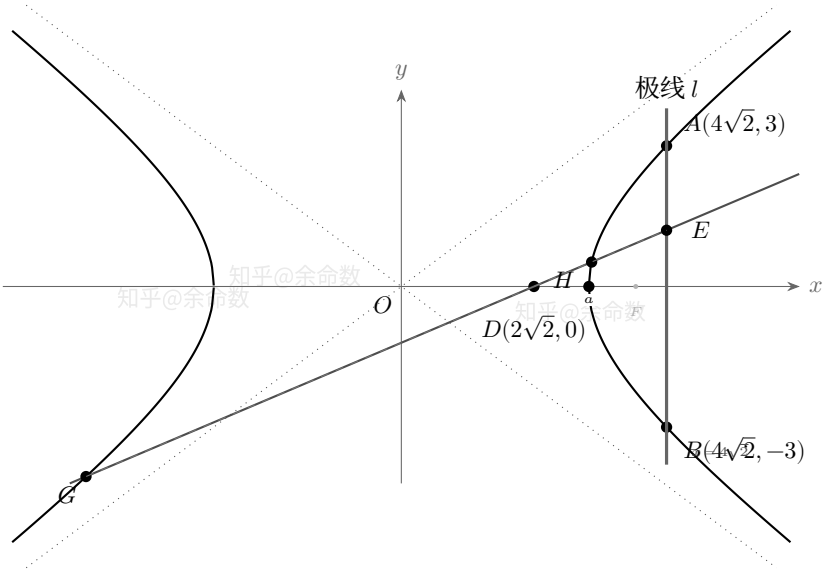
□

知乎@余命数

例题 15.3　已知双曲线 $C:\frac{x^2}{a^2}-\frac{y^2}{b^2}=1\ (a>0,b>0)$ 过点 $A(4\sqrt{2},3)$, 且焦距为 10.

1. 求 C 的方程.
2. 已知点 $B(4\sqrt{2},-3)$、$D(2\sqrt{2},0)$、E 为线段 AB 上一点, 且直线 DE 交 C 于 G,H 两点. 证明 $\dfrac{|GD|}{|GE|}=\dfrac{|HD|}{|HE|}$.

知乎@余命数



知乎@余命数　知乎@余命数　知乎@余命数

证明　(1) 双曲线方程由焦距与点 A 在曲线上两个条件确定. 由焦距 10 知 $2c=10$, 即 $c=5$, 故 $c^2=a^2+b^2=25$. 点 $A(4\sqrt{2},3)$ 在双曲线上, 其坐标满足

$$\frac{(4\sqrt{2})^2}{a^2}-\frac{3^2}{b^2}=1\Longrightarrow\ \frac{32}{a^2}-\frac{9}{b^2}=1$$

由此得到关于 a^2,b^2 的方程组

知乎@余命数

$$\begin{cases}a^2+b^2=25\\32b^2-9a^2=a^2b^2\end{cases}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将 $b^2=25-a^2$ 代入第二式, 整理得

$$32(25-a^2)-9a^2=a^2(25-a^2),$$

$$800-41a^2=25a^2-a^4,$$

$$a^4-66a^2+800=0.$$

这是一个关于 a^2 的二次方程, 分解因式得 $(a^2-16)(a^2-50)=0$. 解得 $a^2=16$ 或 $a^2=50$. 若 $a^2=50$, 则 $b^2=25-50=-25$, 不合题意. 若 $a^2=16$, 则 $b^2=25-16=9$. 因此, 双曲线 C 的方程为 $\dfrac{x^2}{16}-\dfrac{y^2}{9}=1$.

(2) 待证的 $\dfrac{|GD|}{|GE|}=\dfrac{|HD|}{|HE|}$ 等价于 $\dfrac{|GD|}{|HD|}=\dfrac{|GE|}{|HE|}$, 又因 D,E,G,H 共线, 故等价于交比 $(G,H;D,E)=-1$, 即点对 (D,E) 调和分割点对 (G,H).

题中的定点 D、定直线 AB 与过 D 的变动割线 DGH(交 AB 于 E)恰构成点 D 与直线 AB 间的极点极线结构, 故对其加以考察.

我们直接求点 D 关于双曲线 $\dfrac{x^2}{16}-\dfrac{y^2}{9}=1$ 的极线方程. 极点为 $(x_0,y_0)=(2\sqrt{2},0)$, 应用配极变换得

$$\frac{x_0x}{16}-\frac{y_0y}{9}=1$$

代入坐标得

$$\frac{2\sqrt{2}\cdot x}{16}-\frac{0\cdot y}{9}=1$$

$$\frac{\sqrt{2}x}{8}=1\implies x=\frac{8}{\sqrt{2}}=4\sqrt{2}$$

直线 AB 的方程由其上两点确定. 点 A 的坐标为 $(4\sqrt{2},3)$, 点 B 的坐标为 $(4\sqrt{2},-3)$, 故 AB 是垂直于 x 轴的直线, 方程为 $x=4\sqrt{2}$.

点 D 的极线方程与直线 AB 的方程完全相同, 因此直线 AB 就是点 D 的极线.

根据极点与极线的定义, 对任意一条穿过极点 D 的割线, 若其与双曲线交于 G,H 两点, 与极线交于点 E, 则点列 $(G,H;D,E)$ 必构成调和点列, 即 $(G,H;D,E)=-1$.

由交比定义得

$$\frac{\overrightarrow{GD}}{\overrightarrow{HD}}\div\frac{\overrightarrow{GE}}{\overrightarrow{HE}}=-1\implies\frac{\overrightarrow{GD}}{\overrightarrow{HD}}=-\frac{\overrightarrow{GE}}{\overrightarrow{HE}}$$

取两边的绝对值, 得

$$\frac{|\overrightarrow{GD}|}{|\overrightarrow{HD}|}=\frac{|\overrightarrow{GE}|}{|\overrightarrow{HE}|}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

整理后即得所求证的式子.

$$\frac{|GD|}{|GE|}=\frac{|HD|}{|HE|}$$

□

15.4 极点与极线的几何作图

我们考虑如何只用直尺完成极点与极线的相互构造. 这个问题先从调和共轭点的作法入手.

15.4.1 构造调和共轭点

设 A,B,P 为共线三点. 我们作出 P 关于 A,B 的调和共轭点 Q. 作法需要用到**完全四边形**.

一个完全四边形由平面上任意四条直线(无三线共点)构成, 它有六个顶点和三条对角线. 完全四边形的基本性质是, 在任一条对角线上, 这条对角线与另外两条对角线的两个交点, 连同该对角线上的两个原始顶点, 组成调和点列.

15.4.1.0.1 作图步骤

1. 在平面上任取一点 S, 不在直线 ABP 上. 连接 SA,SP,SB.
2. 在直线 SP 上任取一点 R(异于 S,P).
3. 作直线 AR 交 SB 于 D. 作直线 BR 交 SA 于 C.
4. 连接 C,D. 直线 CD 与直线 ABP 的交点即为所求的调和共轭点 Q.

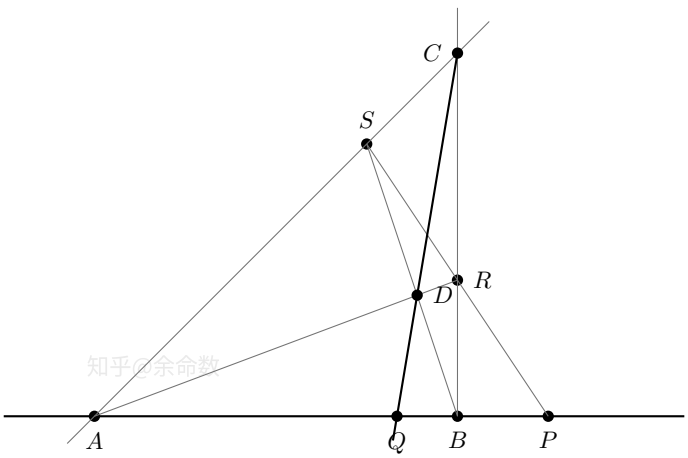


图 15.5　使用完全四边形 $SCDR$(部分顶点为 A,B)构造调和共轭点 Q.

在此构造中, 直线 AB,SP,CD 是完全四边形的三条对角线. 由上述性质, 点列 $(A,B;P,Q)$ 是调和点列.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

15.4.2 由极点作极线

要作出极点 P 的极线, 只需在极线上构造两点.

15.4.2.0.1 作图步骤

1. 从极点 P 引一条割线交二次曲线 $\mathcal{C}$ 于 A_1, B_1.
2. 使用完全四边形法, 构造 P 关于 A_1, B_1 的调和共轭点 Q_1.
3. 从 P 引另一条割线交 $\mathcal{C}$ 于 A_2, B_2.
4. 同样构造调和共轭点 Q_2.
5. 连接 Q_1, Q_2 即得所求的极线 p.

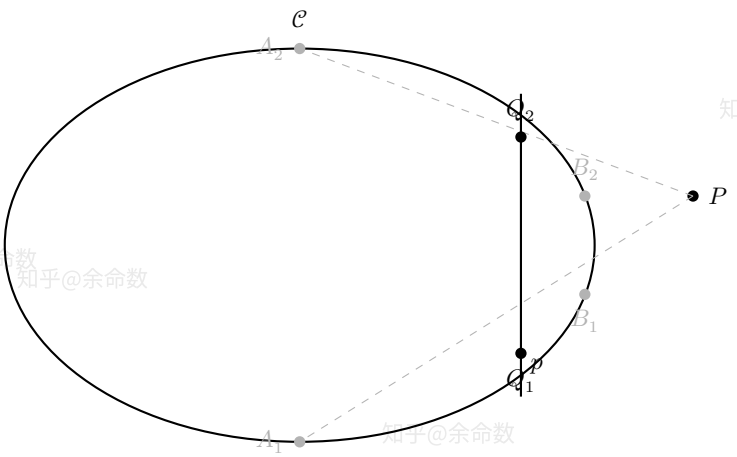


图 15.6 由两个调和共轭点 Q_1, Q_2 确定极线 p.

15.4.3 由极线作极点

下面用拉伊尔定理完成这个作图.

定理 15.5 (拉伊尔定理) 对于本章所讨论的非退化标准二次曲线, 点 P 位于点 Q 关于二次曲线 $\mathcal{C}$ 的极线上, 当且仅当点 Q 位于点 P 关于 $\mathcal{C}$ 的极线上.

证明 我们以标准椭圆为例给出证明. 设 $P(x_1, y_1), Q(x_2, y_2)$. 点 Q 的极线为

$$\frac{x_2 x}{a^2} + \frac{y_2 y}{b^2} = 1.$$

于是 P 在 Q 的极线上, 等价于

$$\frac{x_1 x_2}{a^2} + \frac{y_1 y_2}{b^2} = 1.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

又点 P 的极线为

$$\frac{x_1x}{a^2}+\frac{y_1y}{b^2}=1.$$

将 $Q(x_2,y_2)$ 代入, 得到同一个对称等式

$$\frac{x_1x_2}{a^2}+\frac{y_1y_2}{b^2}=1.$$

故 P 在 Q 的极线上当且仅当 Q 在 P 的极线上. 双曲线与抛物线的极线方程也具有同样的对称双线性形式, 证明完全相同. □

15.4.3.0.1 作图步骤

1. 在给定的极线 p 上任取两点 Q_1,Q_2.
2. 作出点 Q_1 的极线 p_1.
3. 作出点 Q_2 的极线 p_2.
4. 由拉伊尔定理, 所求极点 P 必同时在 p_1 和 p_2 上. 因此, p_1 和 p_2 的交点即为极点 P.

知乎@余命数 知乎@余命数 知乎@余命数

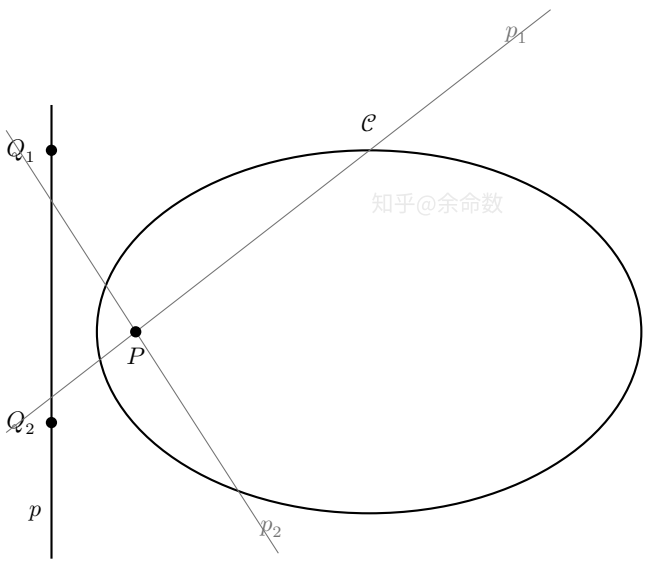


图 15.7 取 p 上两点 Q_1,Q_2, 作其极线 p_1,p_2, 交点即为 p 的极点 P.

知乎@余命数

15.4.4 自极三角形与综合构造

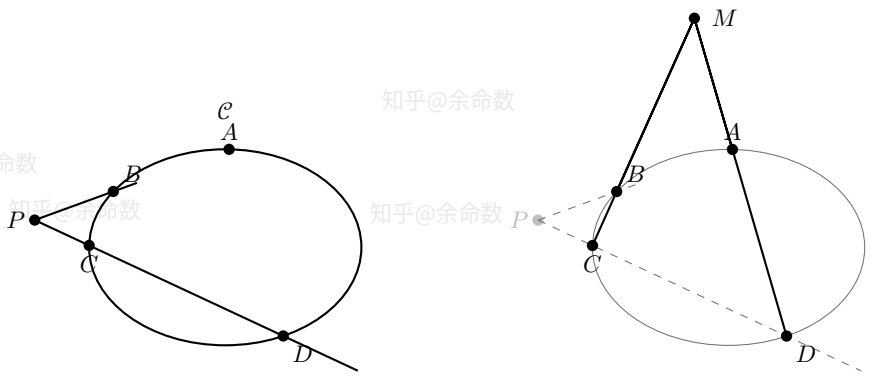
除了用两个独立的调和共轭点确定一条极线之外, 我们还可以直接作出一个自极三角形, 并从中得到极线. 这个构造同时给出极点与极线之间的对偶关系.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设点 Q 在直线 l 上运动. 由拉伊尔定理, Q 的极线 q 恒过直线 l 的极点. 于是, 当 Q 沿 l 移动时, q 绕 l 的极点转动. 利用这一事实, 我们可以只用直尺作出给定一点 P 的极线.

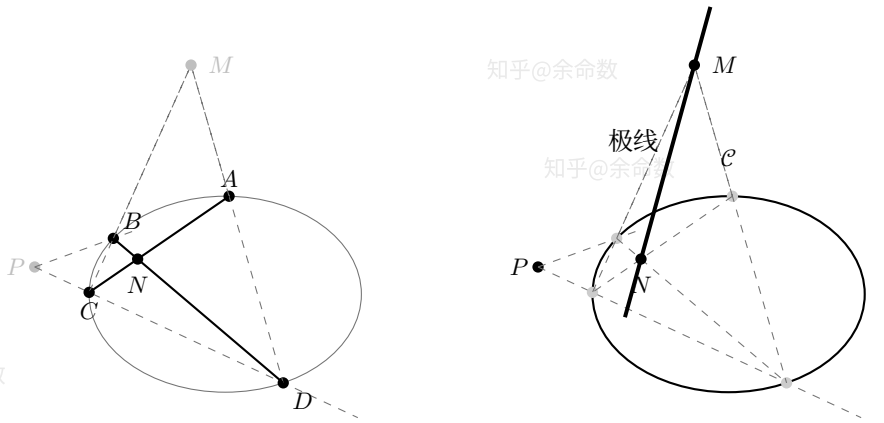
15.4.4.0.1 极线的几何构造 给定二次曲线 $\mathcal{C}$ 和外部一点 P.

1. 从点 P 引两条任意的割线, 分别交二次曲线于点 A, B 和 C, D.
2. 连接对角点, 作直线 AD 和 BC, 设它们的交点为 M.
3. 连接另外一组对角点, 作直线 AC 和 BD, 设它们的交点为 N.
4. 连接 M, N 两点. 直线 MN 即为点 P 关于二次曲线 $\mathcal{C}$ 的极线.



(a) 从点 P 引两条割线交 $\mathcal{C}$ 于 A, B 和 C, D **(b)** 连接对角点 AD 和 BC, 交于点 M

图 15.8 极线构造步骤 1-2



(a) 连接对角点 AC 和 BD, 交于点 N **(b)** 连接 M, N 得到点 P 关于 $\mathcal{C}$ 的极线

图 15.9 极线构造步骤 3-4

这个构造既给出极线 MN, 也给出一个对偶性很强的三角形 $\triangle PMN$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 15.6 (自极三角形) 一个三角形, 如果它的每一个顶点都是其对边的极点, 则称该三角形为关于该二次曲线的**自极三角形**(或自共轭三角形).

在上述构造中, $\triangle PMN$ 就是一个自极三角形. 我们已作出点 P 的极线 MN. 进一步可以证明, 点 M 的极线是 PN, 点 N 的极线是 PM(这里用到交比的射影不变性).

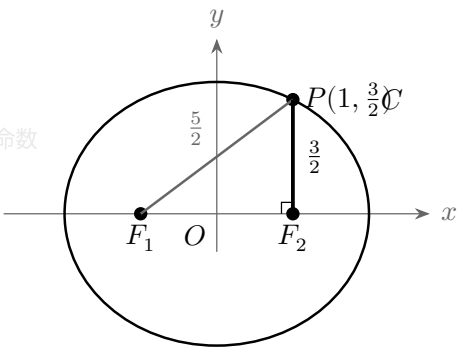
自极三角形把极点与极线的对应关系具体化.

- 取极线 MN 上的任意一点 Q. 由拉伊尔定理, Q 的极线 q 必过 MN 的极点 P. 于是, 点的共线对应着线的共点.
- 再设极线 MN 与二次曲线 $\mathcal{C}$ 交于 E, F. 则 E, F 都在 P 的极线上. 由拉伊尔定理, P 必在 E 的极线和 F 的极线上. 由于 E, F 在 $\mathcal{C}$ 上, 它们在各自处的极线就是切线. 因此 PE 与 PF 是从 P 到 $\mathcal{C}$ 的两条切线, 构成 P 的切点弦.

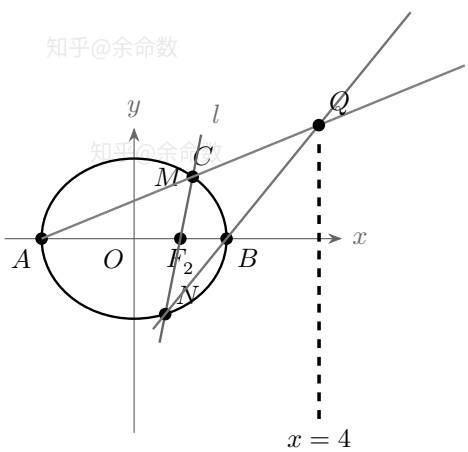
因此, 自极三角形常用于处理与极点、极线、切线相关的几何问题.

例题 15.4 已知椭圆 $C: \frac{x^2}{a^2}+\frac{y^2}{b^2}=1\ (a>b>0)$ 的左右焦点分别为 F_1, F_2, 点 $P(1, \frac{3}{2})$ 在 C 上, 且 $PF_2 \perp F_1F_2$.

1. 求 C 的标准方程.
2. 设 C 的左右顶点分别为 A, B. 直线 l 过右焦点 F_2 且不与坐标轴垂直, l 与 C 交于 M, N 两点. 直线 AM 与直线 BN 相交于点 Q. 证明点 Q 在定直线上.



(a) 问题 (1): 点 P 在椭圆上, $PF_2 \perp F_1F_2$



(b) 问题 (2): AM 与 BN 交于定直线 $x=4$ 上的点 Q

图 15.10　椭圆极点与极线问题示意图

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 (1) 椭圆的焦点在 x 轴上, 坐标为 $F_1(-c,0), F_2(c,0)$. 由 $PF_2 \perp F_1F_2$ 可知点 P 与 F_2 的横坐标相同. 由于 $P(1,\frac{3}{2})$, 可得 $c=1$. 因此 $F_1(-1,0), F_2(1,0)$. 点 P 到左焦点 F_1 的距离为 $|PF_1| = \sqrt{(1-(-1))^2+(\frac{3}{2}-0)^2} = \sqrt{4+\frac{9}{4}} = \frac{5}{2}$. 点 P 到右焦点 F_2 的距离为 $|PF_2| = \frac{3}{2}$. 根据椭圆定义, $2a = |PF_1|+|PF_2| = \frac{5}{2}+\frac{3}{2} = 4$, 所以 $a=2$. 由 $a^2 = b^2+c^2$, 得 $b^2 = a^2-c^2 = 4-1 = 3$. 椭圆 C 的标准方程为 $\dfrac{x^2}{4}+\dfrac{y^2}{3}=1$.

(2) 由题意, 左右顶点分别为 $A(-2,0)$ 与 $B(2,0)$. 设 $M(x_1,y_1)$、$N(x_2,y_2)$、$Q(x_Q,y_Q)$. 直线 l 过右焦点 $F_2(1,0)$, 设其方程为 $y=k(x-1)$. 联立直线与椭圆方程, 得

$$\begin{cases} y=k(x-1) \\ \dfrac{x^2}{4}+\dfrac{y^2}{3}=1 \end{cases} \Longrightarrow (3+4k^2)x^2-8k^2x+4k^2-12=0.$$

x_1, x_2 为此方程的两根, 由韦达定理得

$$x_1+x_2 = \frac{8k^2}{3+4k^2}, \quad x_1x_2 = \frac{4k^2-12}{3+4k^2}$$

由 A, M, Q 三点共线, 斜率相等, 得 $\dfrac{y_Q}{x_Q+2} = \dfrac{y_1}{x_1+2}$ (i). 由 B, N, Q 三点共线, 斜率相等, 得 $\dfrac{y_Q}{x_Q-2} = \dfrac{y_2}{x_2-2}$ (ii).

将 (i) 式与 (ii) 式相除, 得 $\dfrac{x_Q-2}{x_Q+2} = \dfrac{y_1(x_2-2)}{y_2(x_1+2)}$. 两边平方, 得 $\left(\dfrac{x_Q-2}{x_Q+2}\right)^2 = \dfrac{y_1^2(x_2-2)^2}{y_2^2(x_1+2)^2}$.

因为 M, N 在椭圆上, 故 $y_1^2 = 3(1-\frac{x_1^2}{4}) = \frac{3}{4}(4-x_1^2)$, $y_2^2 = \frac{3}{4}(4-x_2^2)$. 代入上式并化简, 得

$$\begin{aligned}\left(\frac{x_Q-2}{x_Q+2}\right)^2 &= \frac{\frac{3}{4}(4-x_1^2)(x_2-2)^2}{\frac{3}{4}(4-x_2^2)(x_1+2)^2} = \frac{(2-x_1)(2+x_1)(x_2-2)^2}{(2-x_2)(2+x_2)(x_1+2)^2} \quad (\text{代入 } y_i^2) \\ &= \frac{(2-x_1)(x_2-2)^2}{(2-x_2)(x_1+2)} = \frac{(2-x_1)(-(2-x_2))^2}{(2-x_2)(x_1+2)} = \frac{(2-x_1)(2-x_2)}{(x_1+2)(x_2+2)} \quad (\text{约去公因子}) \\ &= \frac{4-2(x_1+x_2)+x_1x_2}{4+2(x_1+x_2)+x_1x_2}.\end{aligned}$$

将韦达定理的表达式代入, 得

$$\begin{aligned}\text{分子} &= 4-2\left(\frac{8k^2}{3+4k^2}\right)+\frac{4k^2-12}{3+4k^2} \\ &= \frac{4(3+4k^2)-16k^2+4k^2-12}{3+4k^2} = \frac{12+16k^2-12k^2-12}{3+4k^2} = \frac{4k^2}{3+4k^2}\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\begin{aligned}\text{分母} &= 4+2\left(\frac{8k^2}{3+4k^2}\right)+\frac{4k^2-12}{3+4k^2}\\ &= \frac{4(3+4k^2)+16k^2+4k^2-12}{3+4k^2} = \frac{12+16k^2+20k^2-12}{3+4k^2} = \frac{36k^2}{3+4k^2}\end{aligned}$$

因此,

$$\left(\frac{x_Q-2}{x_Q+2}\right)^2 = \frac{4k^2/(3+4k^2)}{36k^2/(3+4k^2)} = \frac{4}{36} = \frac{1}{9}$$

上式来自平方运算, 还需回到未平方的关系确定符号. 由 $y_i = k(x_i-1)$ 得

$$\frac{x_Q-2}{x_Q+2} = \frac{(x_1-1)(x_2-2)}{(x_2-1)(x_1+2)},$$

再由韦达定理算出

$$3(x_1-1)(x_2-2)-(x_2-1)(x_1+2) = 2x_1x_2-5(x_1+x_2)+8 = 0,$$

故未平方的比值满足

$$\frac{x_Q-2}{x_Q+2} = \frac{1}{3}.$$

由此 $3(x_Q-2) = x_Q+2$, 解得 $x_Q = 4$, 即点 Q 在定直线 $x=4$ 上. □

15.5 笛沙格定理与帕普斯定理

第15.1节 引入了射影平面与中心投影. 在射影平面上, 共线与共点这两类性质经投影后保持不变, 射影几何研究的正是这两类性质. 本节给出射影几何的两条基石, 即笛沙格定理与帕普斯定理, 它们只依赖点线的关联关系, 与距离和角度无关.

15.5.1 笛沙格定理

中心投影把一个图形从一个点投射到另一个平面, 投影前后对应点的连线都经过投影中心. 两个三角形若处于这种投影关系, 即对应顶点连线共点, 它们在射影几何中享有特殊地位. 笛沙格定理刻画这种关系的对偶面貌, 即对应顶点连线共点等价于对应边交点共线.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 15.7 (透视中心与透视轴) 设 $\triangle ABC$ 与 $\triangle A'B'C'$ 是平面上两个三角形. 若三对对应顶点的连线 AA'、BB'、CC' 交于同一点 S, 则称这两个三角形**关于点 S 成透视**, S 称为**透视中心**. 若三对对应边的交点 $P = BC \cap B'C'$、$Q = CA \cap C'A'$、$R = AB \cap A'B'$ 共线, 则称这两个三角形**关于直线 l 成透视**, l 称为**透视轴**.

定理 15.6 (笛沙格定理) 平面上两个三角形关于某点成透视当且仅当它们关于某直线成透视. 即 AA'、BB'、CC' 共点当且仅当 P、Q、R 共线, 其中 $P = BC \cap B'C'$、$Q = CA \cap C'A'$、$R = AB \cap A'B'$.

知乎@余命数

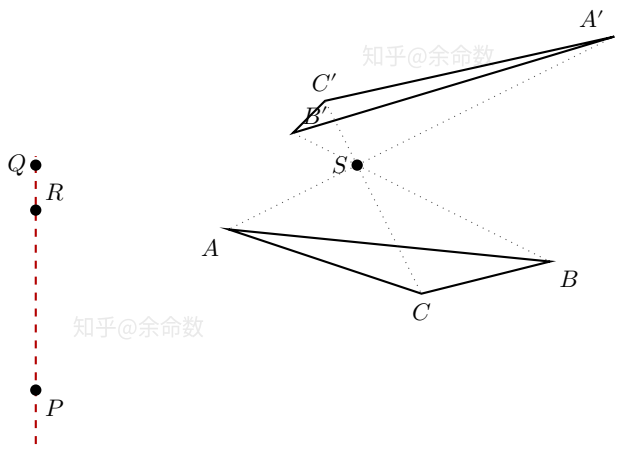


图 15.11 笛沙格定理中的透视中心 S 与透视轴 PQR

证明 仅证充分性(AA'、BB'、CC' 共点 $\Rightarrow$ P、Q、R 共线), 必要性由对偶原理或逆推得到.

设 AA'、BB'、CC' 交于点 S. 在 $\triangle SBC$ 中, B' 在 SB 上, C' 在 SC 上, 且 B'、C'、P 三点共线 ($P = BC \cap B'C'$). 对 $\triangle SBC$ 用截线 $B'C'P$, 由 Menelaus 定理得

$$\frac{SB'}{B'B} \cdot \frac{BP}{PC} \cdot \frac{CC'}{C'S} = -1.$$

同理, 对 $\triangle SCA$ 与截线 $C'A'Q$ 得 Q 满足的 Menelaus 关系, 对 $\triangle SAB$ 与截线 $A'B'R$ 得 R 满足的 Menelaus 关系. 三式相乘, 从 S 出发的三组分比两两相消, 得

$$\frac{BP}{PC} \cdot \frac{CQ}{QA} \cdot \frac{AR}{RB} = -1,$$

由 Menelaus 定理的逆定理, P、Q、R 三点共线. □

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

上述证明在平面内用 Menelaus 定理完成. 笛沙格定理还有一种空间证法, 把两个透视三角形从平面提升到三维, 使对应顶点连线在空间中真正相交于透视中心, 则两个三角形分别落在两个不同的平面上, 而两平面的交线是透视轴, 三对对应边的交点自然落在这条交线上. 笛沙格定理在平面上成立, 根源在于它在三维空间中先成立. 这个空间论证也给出逆定理, 若三对对应边交点共线, 在空间中可构造出使对应顶点连线共点的两个三角形, 再投影回平面即得.

注 15.1
笛沙格定理在射影平面上是自对偶的, 将“点”与“直线”互换, “共点”与“共线”互换, 定理的形式不变. 它是射影几何中少数可以用纯射影方法(不依赖度量)证明的构形定理, 也是建立射影坐标系的逻辑基础.

笛沙格定理处理两个三角形的透视关系, 构形涉及十个点与十条线. 把目光从三角形转向更简单的构形, 即在两条直线上各取三点, 便得到帕普斯定理. 它的假设更少, 结论同样是一条共线关系, 却与笛沙格定理不同, 帕普斯定理无法从笛沙格定理推出, 两条定理在射影几何的公理体系中相互独立.

15.5.2 帕普斯定理

定理 15.7 (帕普斯定理) 设 A、B、C 是直线 l_1 上三点, A'、B'、C' 是直线 l_2 上三点(六点互异, 且 $l_1 \neq l_2$). 记

$$P = BC' \cap B'C, \quad Q = CA' \cap C'A, \quad R = AB' \cap A'B$$

则 P、Q、R 三点共线.

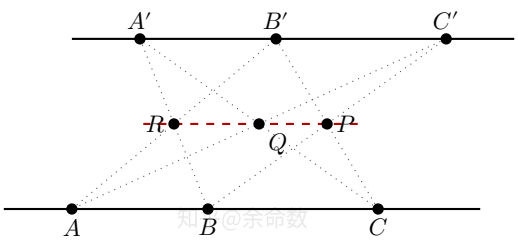


图 15.12　帕普斯定理, 两条直线上各取三点, 交叉连线交点共线

证明 设 l_1 与 l_2 交于点 O(若平行, 取 O 为无穷远点, 以下论证取极限即得). 在 l_1 上取射影坐标使 $O=(0)$、$A=(a)$、$B=(b)$、$C=(c)$; 在 l_2 上取射影坐标使 $O=(0)$、$A'=(a')$、$B'=(b')$、$C'=(c')$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

直线 AB' 的方程由 $A=(a,0)$ 与 $B'=(0,b')$ 给出，其上点可表为 $(ta,(1-t)b')$. 直线 $A'B$ 由 $A'=(0,a')$ 与 $B=(b,0)$ 给出，其上点表为 $(sb,(1-s)a')$. 二者交点 R 满足 $ta=sb$ 且 $(1-t)b'=(1-s)a'$，解得

知乎@余命数

$$R=\left(\frac{ab(a'-b')}{aa'-bb'},\ \frac{a'b'(a-b)}{aa'-bb'}\right)$$

同理

$$P=\left(\frac{bc(b'-c')}{bb'-cc'},\ \frac{b'c'(b-c)}{bb'-cc'}\right),\quad Q=\left(\frac{ca(c'-a')}{cc'-aa'},\ \frac{c'a'(c-a)}{cc'-aa'}\right)$$

验证 P、Q、R 共线等价于行列式

$$\begin{vmatrix} x_P & y_P & 1\\ x_Q & y_Q & 1\\ x_R & y_R & 1 \end{vmatrix}=0$$

知乎@余命数

代入并化简，行列式的值为零，故 P、Q、R 共线. □

注 15.2
帕普斯定理是射影几何中最古老的构形定理，早于笛沙格定理近两千年. 它不涉及任何度量概念（长度、角度、平行），仅依赖点线的关联关系，是射影几何公理体系的出发点之一. 在射影平面上，帕普斯定理等价于坐标乘法的交换律，帕普斯定理成立的射影平面，其坐标环必为交换环.

帕普斯定理中六点分布在两条直线上. 若把两条直线换成一条二次曲线，让六点落在同一条圆锥曲线上，对应的三组对边交点仍然共线，这就是帕斯卡定理. 帕普斯定理是帕斯卡定理在二次曲线退化为两条直线时的特例，也是射影几何从直线构形走向曲线构形的第一步.

例题 15.5 (帕普斯定理的平行情形) 两直线 l_1、l_2 上分别取三个不同点 A、B、C 与 A'、B'、C'. 若 $AB'\parallel A'B$ 且 $AC'\parallel A'C$，证明 $BC'\parallel B'C$.

知乎

知乎@余命数

解 记 $X=AB'\cap A'B$、$Y=AC'\cap A'C$、$Z=BC'\cap B'C$. 由帕普斯定理 定理 15.7，X、Y、Z 三点共线. 前两个条件说明 X 与 Y 都是无穷远点，故这条线就是无穷远直线，Z 也在其上，即 $BC'\parallel B'C$.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

15.6 帕斯卡与布里昂桑

第15.5节 末尾说到, 把帕普斯定理里的两条直线换成一条圆锥曲线, 结论依然成立. 这件事实叫帕斯卡定理, 1640 年, 十六岁的帕斯卡把它写进短文《圆锥曲线论》, 后人把它唤作“神秘六边形”定理. 我们先把圆上的情形证出来, 再借助第15.1节 的中心投影推广到任意圆锥曲线.

15.6.1 圆上的帕斯卡定理

定理 15.8 (帕斯卡定理) 六边形 123456 内接于一条圆锥曲线, 三组对边的交点

$$X = 12 \cap 45, \qquad Y = 23 \cap 56, \qquad Z = 34 \cap 61$$

共线, 这条线称为**帕斯卡线**.

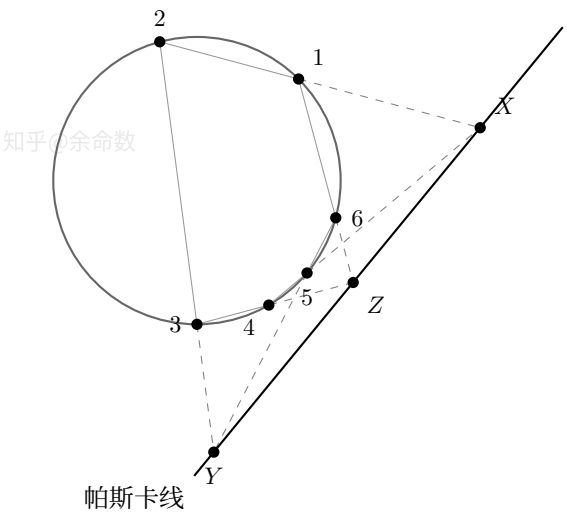


图 15.13 单位圆上的神秘六边形, 对边交点 X、Y、Z 落在同一条帕斯卡线上.

证明 先设曲线是圆, 且三组对边互不平行, 交点都是有限点. 记三条相间的边 $a = 12$、$b = 34$、$c = 56$, 它们两两相交成三角形, 顶点记作

$$U = b \cap c, \qquad V = c \cap a, \qquad W = a \cap b,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

则 $\triangle UVW$ 的三条边依次是 a, 即 VW、b, 即 WU、c, 即 UV, 而六边形的顶点 $1,2$ 在边 a 上, $3,4$ 在边 b 上, $5,6$ 在边 c 上. 剩下的三条边 23、45、61 各截这个三角形一次. 对 $\triangle UVW$ 用截线 23, 它交 VW 于 2、WU 于 3、UV 于 Y, 由 Menelaus 定理

$$\frac{V2}{2W}\cdot\frac{W3}{3U}\cdot\frac{UY}{YV}=-1.$$

同理, 截线 45 与截线 61 给出

$$\frac{W4}{4U}\cdot\frac{U5}{5V}\cdot\frac{VX}{XW}=-1,\qquad \frac{U6}{6V}\cdot\frac{V1}{1W}\cdot\frac{WZ}{ZU}=-1.$$

三式相乘, 并按顶点归并,

$$\frac{V1\cdot V2}{W1\cdot W2}\cdot\frac{W3\cdot W4}{U3\cdot U4}\cdot\frac{U5\cdot U6}{V5\cdot V6}\cdot\frac{UY}{YV}\cdot\frac{VX}{XW}\cdot\frac{WZ}{ZU}=-1.$$

前三个括号用圆幂定理 (定理 13.38) 消去.W 的两条割线分别穿过 $1,2$ 与 $3,4$, 故 $W1\cdot W2=W3\cdot W4$; 同理 $U3\cdot U4=U5\cdot U6$, $V5\cdot V6=V1\cdot V2$. 于是

$$\frac{V1\cdot V2}{W1\cdot W2}\cdot\frac{W3\cdot W4}{U3\cdot U4}\cdot\frac{U5\cdot U6}{V5\cdot V6}=\frac{V1\cdot V2}{W1\cdot W2}\cdot\frac{W1\cdot W2}{U3\cdot U4}\cdot\frac{U3\cdot U4}{V1\cdot V2}=1,$$

剩下

$$\frac{UY}{YV}\cdot\frac{VX}{XW}\cdot\frac{WZ}{ZU}=-1,$$

由 Menelaus 定理的逆定理, 落在 $\triangle UVW$ 三边所在直线上的 Y、X、Z 三点共线. 对边出现平行的情形, 相应交点落到无穷远, 论证沿 第15.5节 处理帕普斯定理的同一取极限思路完成.

曲线是任意圆锥曲线时, 取 第15.1节 的中心投影把它的一个圆原型投成它. 投影保持共线性与共点性, 帕斯卡线原封不动地投成帕斯卡线. □

帕斯卡定理与帕普斯定理的关系也随之清楚了. 一对相交直线是退化的二次曲线, 让椭圆沿长轴方向压扁、最终裂成两条线段, 六个顶点分居两线, 帕斯卡构形恰好化作帕普斯构形. 第15.5节 里那条与笛沙格定理彼此独立的公理, 在更大的图景里是帕斯卡定理的退化特例.

15.6.2 退化, 相邻顶点重合

帕斯卡定理最漂亮的部分藏在退化情形里. 让相邻两个顶点彼此靠近、最后重合, 连接它们的边在极限处变成该点处的**切线**, 于是六边形塌缩成五边形、四边形, 定理仍然有效, 只是对边中出现了切线.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

例题 15.6 (退化帕斯卡的手算验证) 单位圆上取四点 $A(1,0)$、$B\left(\frac{4}{5},\frac{3}{5}\right)$、$C(-1,0)$、$D\left(\frac{3}{5},-\frac{4}{5}\right)$，考虑六边形 A,A,B,C,D,D，两侧的边 AA 与 DD 都换成切线. 验证三组对边交点共线.

解 单位圆在 (x_0,y_0) 处的切线是 $xx_0+yy_0=1$，故 A 处切线为 $x=1$，D 处切线为 $3x-4y=5$. 逐一求交点.CD 过 $C(-1,0)$ 与 $D\left(\frac{3}{5},-\frac{4}{5}\right)$，斜率 $-\frac{1}{2}$，方程 $y=-\frac{x+1}{2}$，与 $x=1$ 交于

$$X=(1,-1).$$

AB 的斜率是 $\frac{3/5-0}{4/5-1}=-3$，方程 $y=-3x+3$，与 $3x-4y=5$ 联立得 $-12x+12=3x-5$，解出 $x=\frac{17}{15}$，

$$Y=\left(\frac{17}{15},-\frac{2}{5}\right).$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

DA 的斜率是 2，方程 $y=2(x-1)$；BC 的斜率是 $\frac{1}{3}$，方程 $y=\frac{x+1}{3}$，联立解出 $x=\frac{7}{5}$，

$$Z=\left(\frac{7}{5},\frac{4}{5}\right).$$

核对共线，直线 XY 的斜率是 $\frac{-\frac{2}{5}+1}{\frac{17}{15}-1}=\frac{3/5}{2/15}=\frac{9}{2}$，直线 XZ 的斜率是 $\frac{\frac{4}{5}+1}{\frac{7}{5}-1}=\frac{9/5}{2/5}=\frac{9}{2}$，两者相同，三点确实共线.

知乎@余命数

四点构形再退化下去，就回到 第15.4节 的自极三角形. 内接四点形的三个对角点构成自极三角形，每个对角点的极线恰是另两个对角点的连线，其根据正是带切线的退化帕斯卡. 那里用调和作图搭起的极线理论，在这里与锥线上的六点构形会师.

15.6.3 逆定理与直尺作图

知乎@余命数

帕斯卡定理可以反过来用，它给出六点落在一条锥线上的一种判别法.

引理 15.1 (五点定锥线) 一般位置(无三点共线)的五个点恰好在一条圆锥曲线上.

证明 二次曲线的方程 $Ax^2+Bxy+Cy^2+Dx+Ey+F=0$ 有六个系数，全体乘同一个非零常数不改变曲线，实际自由度是五. 每让曲线过一个点，系数就多满足

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

一个线性方程. 五个点给出五个齐次线性方程、六个未知数, 必有非零解; 一般位置保证这组解给出的二次曲线不退化成直线对或单点. □

定理 15.9 (帕斯卡逆定理) 六个点 $1,\dots,6$ 中无三点共线, 且 $X = 12\cap 45$、$Y = 23\cap 56$、$Z = 34\cap 61$ 三点共线, 则这六点落在同一条圆锥曲线上.

证明 由 引理 15.1, 过前五点有唯一的锥线 α. 直线 $5Y$ 与 α 除 5 外还交于一点, 记作 6^*. 对 α 上的六边形 $1,2,3,4,5,6^*$ 用帕斯卡定理, 其第一组对边交点仍是 $12\cap 45 = X$; 第二组是 $23\cap 56^*$, 由于 6^* 在直线 $5Y$ 上, 这个交点正是 Y. 于是帕斯卡线是 XY, 第三组对边交点 $34\cap 6^*1$ 也落在直线 XY 上. 另一方面 $Z = 34\cap 61$ 落在直线 34 与直线 XY 上, 一般位置下两线只交于一点, 故 $34\cap 6^*1 = Z$, 即 6^* 在直线 16 上.6 与 6^* 同在直线 $5Y$ 与直线 16 上, 两线不重合, 只能 $6^* = 6$, 六点共锥线. □

例题 15.7 (用直尺作锥线上的第六点) 锥线上给定五个点, 只用直尺作出曲线上更多的点.

解 记给定点为 $1,2,3,4,5$, 它们定出唯一的锥线. 作 $X = 12\cap 45$, 过 X 任画一条直线 l(避开五个给定点), 记 $B = l\cap 23$、$C = l\cap 34$, 再连 $B5$ 与 $C1$, 交点记作 6. 六边形 $1,2,3,4,5,6$ 的三组对边交点分别是 X、B、C, 三点都在 l 上, 由 定理 15.9, 6 在锥线上. 移动直线 l, 每换一条过 X 的直线就得到曲线上一个新点, 全程只用直尺.

15.6.4 布里昂桑, 对偶的定理

把 第15.3节 的极点极线字典逐字翻译帕斯卡构形, 会得到一条新定理. 字典里最要紧的是三句话, 锥线上一点处的极线, 是锥线在该点的切线; 两点连线的极点, 是各自极线的交点; 一组共线点的极线, 全部经过这条线的极点. 顺着字典, 内接六边形翻译成外切六边形, “对边交点共线”翻译成“对角线共点”.

定理 15.10 (布里昂桑定理) 六边形外切于一条圆锥曲线, 三条对角线(相对顶点的连线)共点.

证明 设外切六边形六条边依次切锥线 α 于点 $1,\dots,6$. 内接六边形 123456 的对边交点 $X = 12\cap 45$ 对应过去, X 的极线是连线 (边 12处切线 $\cap$ 边 45处切线), 而这两条切线正是外切六边形的两条对边, 它们的交点是一对相对顶点, 所以 X 的极线就是外切六边形的一条对角线. 同理 Y、Z 的极线是另外两条对角线. 帕斯卡定理说 X、Y、Z 共线, 共线点的极线共点, 三条对角线便交于一点. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

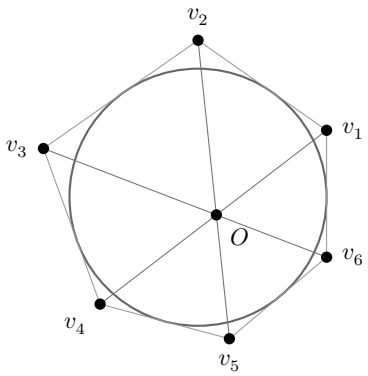


图 15.14　圆的外切六边形, 三条对角线交于同一点 O.

布里昂桑 1806 年发表这条定理时, 用的正是刚刚成形的对偶原理, 它是极点极线理论最早的成功应用. 让外切六边形退化, 把六条边两两重合成三角形的三边, 重合处的顶点塌缩成切点, 就得到一个漂亮的特例. 三角形外切于锥线时, 每个顶点与对边切点的连线三线共点, 对内切圆的情形这正是三角形的 Gergonne 点.

练习 15.1　圆内接六边形 $ABCDEF$ 中, $AB \parallel DE$ 且 $BC \parallel EF$, 证明 $CD \parallel FA$.

练习 15.2　三角形 ABC 外切于锥线 α, 三边上的切点依次为 X、Y、Z, 其中 X 在 BC 上、Y 在 CA 上、Z 在 AB 上. 用退化布里昂桑定理证明 AX、BY、CZ 共点, 再用 Ceva 定理复核一遍.

练习 15.3　锥线上给定四个点 $1, 2, 3, 4$ 与另一点 P, 只用直尺作 P 处的切线.

知乎@余命数

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
@余命数

第四卷 · 分析主干

中学数学之旅

本卷为分析主干：极限与导数，含数列极限、级数与泰勒、积分、微分方程一气呵成，并以数学思想浅谈收束跨章方法. 变化率与累积量在此成对出现.

知乎 @ 余命数 (Midnight Forever)

Project Akari

知乎@余命数
知乎@余命数
知乎@余命数
知乎@
知乎@余命数

本卷导读

本卷线索

第四卷的主线是：**用极限语言刻画变化，用积分语言刻画累积，并整理跨章方法.** 全卷由导数与积分构成分析主干，微分方程与思想方法两章在其上延伸.

知乎@余命数

章节地图

▷ **极限与导数**　从实数完备性与极限定义出发，经连续性进入导数，继而展开中值定理、洛必达法则与泰勒展开，并以方程根、不等式证明与极值点偏移等综合专题收束. 本章是全卷主干，后续各章都建立在它的语言之上.

▷ **简单的积分**　紧随导数展开其逆运算：原函数与不定积分、换元与分部积分，最后由定积分与微积分基本定理把变化与累积两面接通.

▷ **微分方程初步**　以导数为工具处理可分离变量与一阶线性方程等可解模型，并反过来用微分方程求解某些函数方程.

▷ **数学思想浅谈**　总结自由度、解题框架、构造式与满足式、不变量与换元等跨章方法，所用例证大多取自前此各卷的分析与几何训练.

知乎@余命数

向后展望

第五卷把数系扩展到复数，同时讲空间的立体几何，引入矩阵与行列式处理线性结构，并以连分数、三个著名常数与无穷乘积收束数论与常数专题.

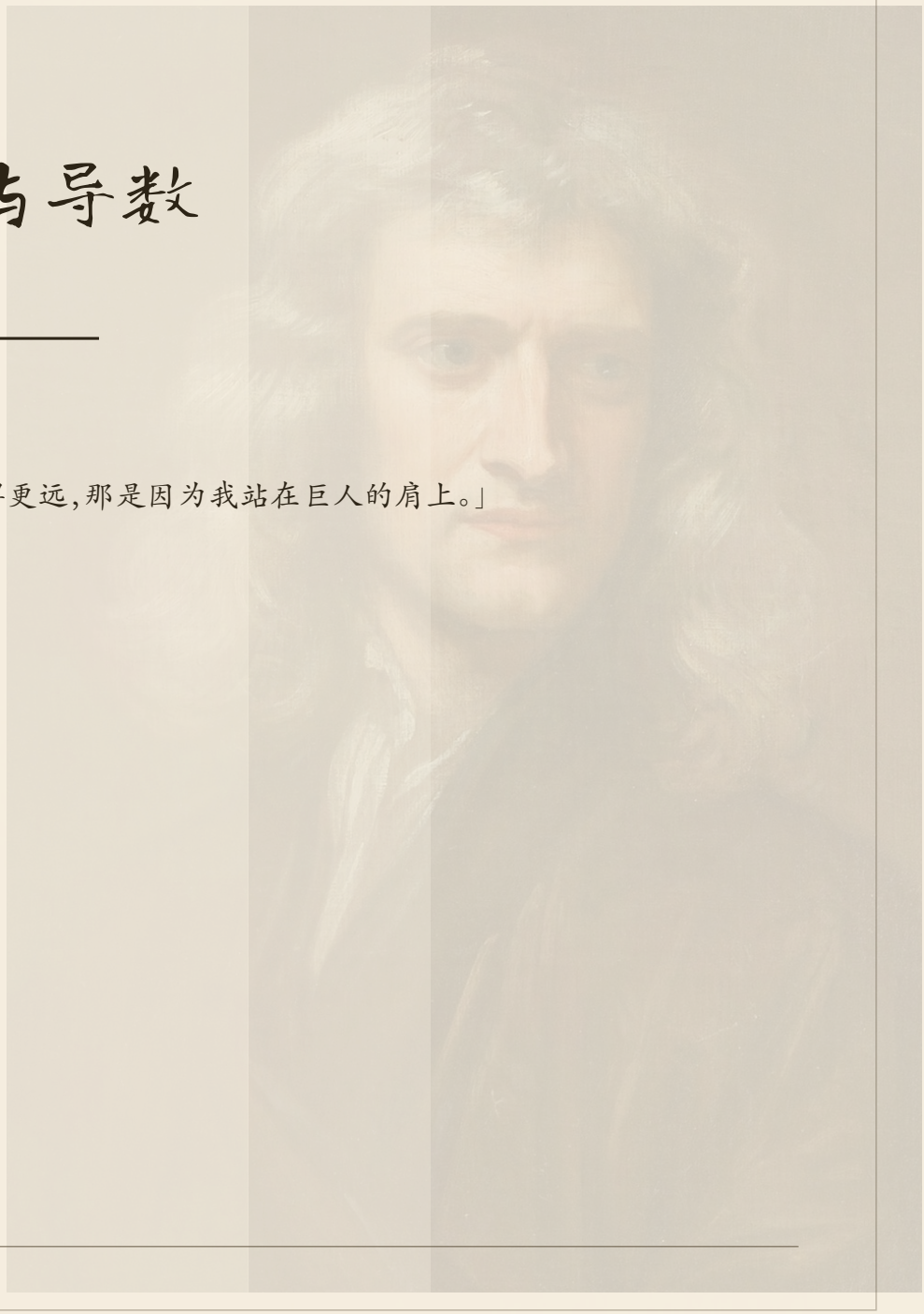
知乎@余命数

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 16 章

极限与导数



分析主干

「如果说我看得更远，那是因为我站在巨人的肩上。」

——牛顿

知乎@余命数

第十六章 极限与导数

知乎@余命数

微积分学处理变化与运动. 我们考虑一个物体沿直线运动, 设其位置由时间 t 的函数 $s = f(t)$ 给出. 为了刻画物体在某一时刻 t_0 的运动状态, 仅知道位置还不够, 我们还需要知道它的瞬时速度. 在时间段 $[t_0, t_0 + \Delta t]$ 内, 物体的平均速度由位移与时间的比值给出.

$$\bar{v} = \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}.$$

在几何上, 若考虑 t-s 平面上的曲线 $s = f(t)$, 该比值对应于连接点 $P(t_0, f(t_0))$ 与点 $Q(t_0 + \Delta t, f(t_0 + \Delta t))$ 的割线斜率.

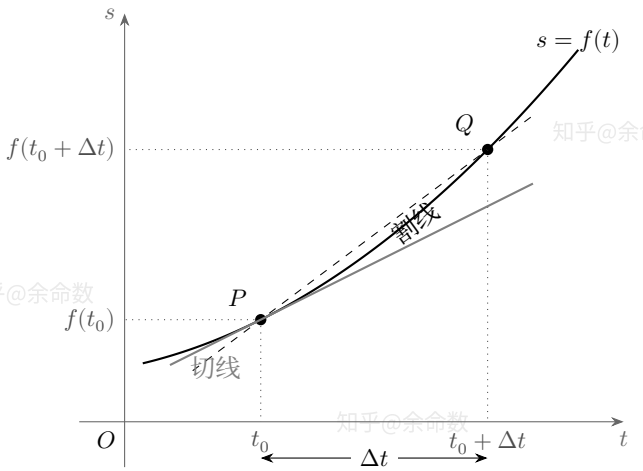


图 16.1 平均速度对应的割线斜率与瞬时速度对应的切线斜率.

物理直觉告诉我们, 当观测的时间间隔 Δt 越来越短时, 平均速度应当逼近某个确定的数值, 即瞬时速度. 几何上, 这对应于点 Q 沿曲线趋近于点 P, 于是割线趋向于曲线在点 P 处的切线. 然而, 若直接令 $\Delta t = 0$, 上述比值变为 $\frac{0}{0}$, 在代数运算中这是未定义的.

为了解决这一矛盾, 我们引入 极限 (limit) 的概念来严格描述无限趋近这一过程. 在定义极限之前, 我们先考察极限运算所依赖的数系. 若数系本身存在空隙, 则极限过程可能指向一个不存在的数, 理论便会失效.

知乎@余命数

16.1 实数系的完备性

知乎@余命数

我们希望在数轴上度量所有连续量, 但有理数系 $\mathbb{Q}$ 做不到这一点. 例如边长为 1 的正方形的对角线长为 $\sqrt{2}$, 它不是有理数. 我们引入完备性公理来弥补这一缺陷, 使实数轴没有空隙, 从而保证极限过程不会指向一个不存在的数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

注 16.1 (记号)

- $\mathbb{N}$: 自然数集 $\{0,1,2,\dots\}$.
- $\mathbb{Z}$: 整数集 $\{\dots,-1,0,1,\dots\}$.
- $\mathbb{Q}$: 有理数集 $\{p/q \mid p\in\mathbb{Z}, q\in\mathbb{Z}\setminus\{0\}\}$.
- $\mathbb{R}$: 实数集.

有理数集 $\mathbb{Q}$ 对加、减、乘、除(除数非零)四则运算封闭, 且其元素在数轴上是稠密的. 但只靠有理数仍不足以度量所有几何量. 我们考虑边长为 1 的正方形, 其对角线长度 l 满足 $l^2=2$.

定理 16.1 ($\sqrt{2}$ 的无理性) 不存在有理数 r, 使得 $r^2=2$.

证明 我们用反证法. 假设存在 $r\in\mathbb{Q}$ 使得 $r^2=2$. 设 $r=p/q$, 其中 $p,q\in\mathbb{Z}^+$ 且 $\gcd(p,q)=1$.

知乎@余命数

由 $(p/q)^2=2$ 得 $p^2=2q^2$. 故 p^2 为偶数, 从而 p 为偶数. 设 $p=2k$. 代入得 $(2k)^2=2q^2 \implies 4k^2=2q^2 \implies q^2=2k^2$. 故 q^2 为偶数, 从而 q 亦为偶数. 故 p,q 至少有公约数 2, 这与 $\gcd(p,q)=1$ 矛盾. □

知乎@余命数

上述定理表明数轴上存在无理数. 若只在 $\mathbb{Q}$ 中讨论, 极限过程可能失效. 例如有理数序列 $1, 1.4, 1.41, \dots$ 无限趋近于 $\sqrt{2}$, 而它的极限不在 $\mathbb{Q}$ 内. 因此我们把数系扩充为实数系 $\mathbb{R}$, 并要求 $\mathbb{R}$ 具有 *完备性*.

为了形式化实数轴没有空隙这一事实, 我们先引入上界与确界.

定义 16.1 (上界) 设 S 是 $\mathbb{R}$ 的非空子集. 若存在 $M\in\mathbb{R}$, 使得对所有 $x\in S$ 都有 $x\le M$, 则称 M 为 S 的一个 *上界* (upper bound).

对于上有界的集合, 上界通常不唯一. 所有上界中最小的那个叫做上确界.

定义 16.2 (上确界) 设 $S\subseteq\mathbb{R}$ 非空且上有界. 若实数 α 满足以下条件, 则称 α 为 S 的 *上确界* (supremum), 记作 $\sup S$:

1. α 是 S 的上界;
2. 任何小于 α 的数都不是 S 的上界, 即 α 是最小的上界.

上确界 α 不一定属于集合 S. 例如 $S=[0,1)$, 则 $\sup S=1\notin S$.

在有理数系 $\mathbb{Q}$ 中, 有界集合的上确界可能不存在. 我们考虑 $A=\{r\in\mathbb{Q}\mid r^2<2\}$.$A$ 在 $\mathbb{Q}$ 中有上界, 例如 2, 但没有最小的有理上界, 因为 $\sqrt{2}\notin\mathbb{Q}$. 完备性公理保证这种情况在 $\mathbb{R}$ 中不会发生.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

公理 16.1 (完备性公理) $\mathbb{R}$ 中任何非空且上有界的子集, 必有上确界.

这一公理在几何上刻画了实数轴的连续性 (continuity).

我们把实数轴看成一条直线. 点集 S 不向右无限延伸时, 它应当有一个最右端点. 完备性公理保证这个端点存在.

所有大于 S 中每个元素的数构成 S 的上界集合.

- 在有理数系 $\mathbb{Q}$ 中, 集合 S 与它的上界集合之间可能存在空隙. 例如 $S=\{x\in\mathbb{Q}\mid x^2<2\}$ 没有最小的有理上界, 因为 $\sqrt{2}\notin\mathbb{Q}$.
- 在实数系 $\mathbb{R}$ 中, 完备性公理填补这些空隙. 每个有上界的集合都有上确界作为精确边界; 任何小于它的数都不是上界, 所以它紧临 S.

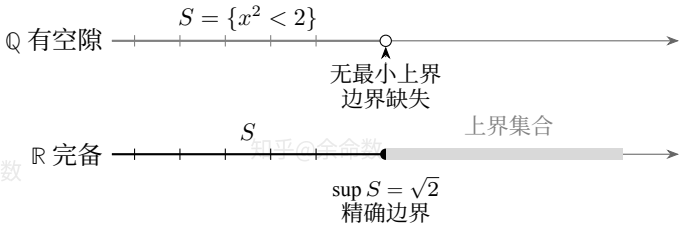


图 16.2 有理数系的空隙与实数系的完备性对比.

16.2 极限的定义 *

知乎@余命数

我们先给出极限的直观定义.

定义 16.3 当变量 x 无限趋近于定值 x_0 且 $x\neq x_0$ 时, 若函数值 $f(x)$ 无限趋近于定值 L, 则称 L 为函数 $f(x)$ 在 x 趋向于 x_0 时的**极限**.

这种描述同时涉及过程和结果两个方面. 记作

$$\lim_{x\to x_0} f(x)=L$$

其中 $\to$ 表示 x 趋近于 x_0.

知乎@余命数

- 符号 lim 是英文 limit 的缩写, 表示这里讨论的是极限.
- 记号 $x\to x_0$ 表示自变量 x 从 x_0 的左右两侧无限靠近, 与 x_0 的距离可以任意小, 但 $x\neq x_0$.
- $f(x)$ 是研究对象, 它的值随 x 变化.
- 等号右侧的 L 表示极限值为确定的常数 L.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个定义关心的是 x 在 x_0 附近的变化过程, 并不关心 $f(x_0)$ 是否存在.

例题 16.1　考察函数 $f(x)=\frac{x^2-1}{x-1}$ 在 $x\to 1$ 时的极限.

解　当 $x=1$ 时, 分母为 0, $f(1)$ 在该点**没有定义**.

极限由 $x=1$ 附近的函数值决定, 与 $f(1)$ 是否存在无关.

当 $x\neq 1$ 时, 函数可化简, 得

$$f(x)=\frac{x^2-1}{x-1}=\frac{(x-1)(x+1)}{x-1}=x+1$$

化简后 $y=f(x)$ 的图像与直线 $y=x+1$ 在 $x\neq 1$ 时相同, 点 $(1,2)$ 处无定义.

我们观察 x 趋近 1 时 $f(x)$ 的变化.

当 x 从 1 的右侧趋近, 例如 $1.1, 1.01, 1.001, \ldots$ 时,

$f(x)=x+1$ 的值依次为$2.1, 2.01, 2.001, \ldots$, 趋近于2.

当 x 从 1 的左侧趋近, 例如 $0.9, 0.99, 0.999, \ldots$ 时,

$f(x)=x+1$ 的值依次为$1.9, 1.99, 1.999, \ldots$, 趋近于2.

虽然 $f(1)$ 不存在, 但 x 从两侧无限趋近 1 时, $f(x)$ 都无限趋近于 2. 因此极限存在且等于 2.

$$\lim_{x\to 1}\frac{x^2-1}{x-1}=2$$

这个例子说明极限值是函数在该点附近的变化趋势.

为了分析 $\lim\limits_{x\to 1}\frac{x^2-1}{x-1}$, 我们分别观察 x 从右侧和左侧趋近 1 时函数值的变化. 两侧的函数值都趋近于 2.

于是我们引入左极限和右极限的定义.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 16.4 (右极限与左极限) • 当 x 从大于 x_0 的一侧无限趋近于 x_0 时,若 $f(x)$ 无限趋近于常数 L,则称 L 为 $f(x)$ 在 x_0 的右极限,记作

$$\lim_{x\to x_0^+} f(x) = L.$$

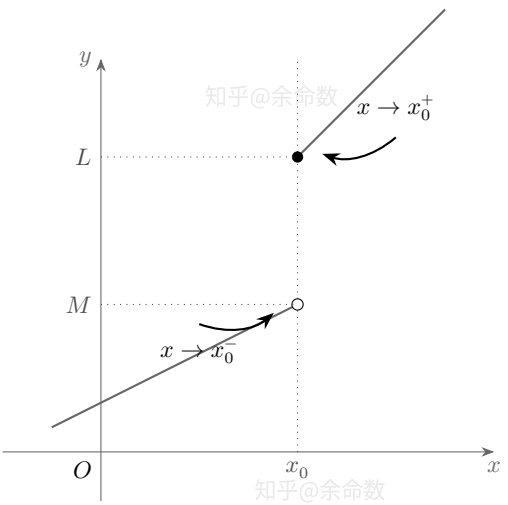
上标 $+$ 表示 x 从右侧趋近 x_0.

• 当 x 从小于 x_0 的一侧无限趋近于 x_0 时,若 $f(x)$ 无限趋近于常数 M,则称 M 为 $f(x)$ 在 x_0 的左极限,记作

$$\lim_{x\to x_0^-} f(x) = M.$$

上标 $-$ 表示 x 从左侧趋近 x_0.

知乎@余命数



函数在某点的极限存在,当且仅当它的左、右极限都存在且相等.

定理 16.2 (极限存在的充要条件) 函数 $f(x)$ 在点 x_0 处的极限存在且等于 L,当且仅当下式成立:

$$\lim_{x\to x_0} f(x) = L \iff \lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x) = L. \tag{16-1}$$

证明 我们证明必要性. 若 $\lim\limits_{x\to x_0} f(x) = L$, 则对任意 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $0 < |x - x_0| < \delta$ 时有 $|f(x) - L| < \varepsilon$. 特别地, 当 $0 < x - x_0 < \delta$ 或 $0 < x_0 - x < \delta$ 时仍满足同一不等式, 因而右极限和左极限都等于 L.

充分性也由定义推出. 若左、右极限都等于 L, 则对任意 $\varepsilon > 0$, 存在 $\delta_1 > 0$, 使得当 $0 < x_0 - x < \delta_1$ 时有 $|f(x) - L| < \varepsilon$; 也存在 $\delta_2 > 0$, 使得当 $0 < x - x_0 < \delta_2$ 时

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

有 $|f(x)-L|<\varepsilon$. 取 $\delta=\min(\delta_1,\delta_2)$. 当 $0<|x-x_0|<\delta$ 时, x 必在 x_0 左侧或右侧之一, 于是由对应的一侧极限估计都可得到 $|f(x)-L|<\varepsilon$. 因此 $\lim\limits_{x\to x_0} f(x)=L$. □

直观地说, 函数在 x_0 处的极限存在, 当且仅当 x 从左右两侧趋近 x_0 时函数值都趋向同一个数.

例题 16.2　考察函数 $f(x)=\dfrac{|x|}{x}$ 在 $x\to 0$ 时的极限.

解　先写成分段形式.

$$f(x)=\frac{|x|}{x}=\begin{cases}1, & x>0\\ -1, & x<0\end{cases}$$

知乎@余命数

再分别计算左、右极限.

右极限(x 从右侧趋近 0, 此时 $x>0$)

$$\lim_{x\to 0^+} f(x)=1$$

知乎@余命数

左极限(x 从左侧趋近 0, 此时 $x<0$)

$$\lim_{x\to 0^-} f(x)=-1$$

知乎@余命数

左、右极限都存在但不相等, 因此

$$\lim_{x\to 0^-} f(x)\neq\lim_{x\to 0^+} f(x)$$

从而

$$\lim_{x\to 0}\frac{|x|}{x}\quad \textbf{不存在}.$$

函数图像在 $x=0$ 两侧分别趋向 1 和 -1, 所以极限不存在.

直观定义中的无限趋近含义模糊, 需要严格的数学表述.

为了把靠近精确化, 我们引入邻域概念.

定义 16.5 (邻域)　以点 x_0 为中心、δ 为半径的开区间称为点 x_0 的**邻域**, 记作 $U(x_0,\delta)$.

$$U(x_0,\delta)=(x_0-\delta,x_0+\delta)=\{x\mid |x-x_0|<\delta\}$$

其中 δ (delta) 是任意给定的正数, 称为邻域的**半径**.

邻域包含中心点, 极限过程需要排除中心点, 由此引入去心邻域.

注 16.2
以下用邻域与 $\varepsilon-\delta$ 语言给出极限的严格表述.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定义 16.6 (去心邻域) 点 x_0 的邻域去掉中心点 x_0 之后, 剩下的部分称为点 x_0 的**去心邻域**或**空心邻域**.

$$U^{\circ}(x_0,\delta)=\{x\mid 0<|x-x_0|<\delta\}$$

去心邻域精确表达 x 趋近于 x_0 但 $x\neq x_0$.

知乎@余命数

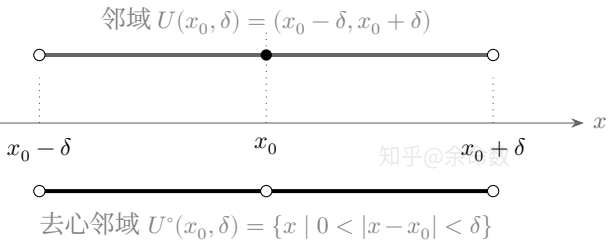


图 16.3 邻域与去心邻域的示意图

借助邻域可将定义改写成更便于使用的形式.

定义 16.7 设函数 $f(x)$ 在点 x_0 的某个**去心邻域**内有定义. 若当 x 在该去心邻域内无限趋近于 x_0 时, $f(x)$ 无限趋近于常数 L, 则称 L 为 $f(x)$ 当 x 趋近于 x_0 时的**极限**, 记作

$$\lim_{x\to x_0} f(x)=L. \tag{16-2}$$

这个定义强调函数只需在 x_0 的去心邻域内有定义, x_0 本身可以没有定义.

邻域仍未消除无限趋近的动态性. 柯西和魏尔斯特拉斯用 $\varepsilon-\delta$ 语言给出了静态定义.

定义 16.8 (函数极限的 $\varepsilon-\delta$ 定义) 设函数 $f(x)$ 在点 x_0 的某个去心邻域内有定义. 若**对于任意给定的正数** ε(无论它有多小), **总存在**一个**正数** δ, 使得当自变量 x 满足不等式 $0<|x-x_0|<\delta$ 时, 对应的函数值 $f(x)$ 都满足不等式 $|f(x)-L|<\varepsilon$, 那么我们就称常数 L 是函数 $f(x)$ 当 $x\to x_0$ 时的极限. 用逻辑符号可表示如下.

$$\lim_{x\to x_0} f(x)=L$$
$$\iff \forall\varepsilon>0,\ \exists\delta>0,\ \text{s.t.}\ \forall x,\ (0<|x-x_0|<\delta \implies |f(x)-L|<\varepsilon).$$

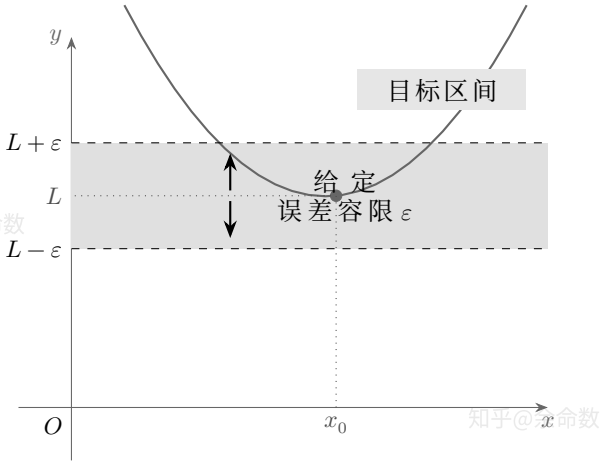
这个定义把动态趋近转化为静态逻辑关系.[1] 下面三幅图把这个定义逐步展开.[2]

[1] 参见《普林斯顿微积分读本》一书附录关于极限定义的说明.

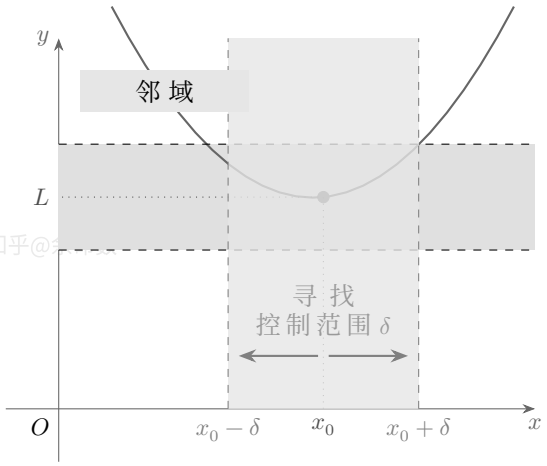
[2] 此图解思路引用自阿德里安·班纳 (Adrian Banner) 所著《普林斯顿微积分读本》一书的附录极限及其证明.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

给定任意正数 ε, 我们就在 y 轴上以 L 为中心确定区间 $(L-\varepsilon, L+\varepsilon)$, 要求 $f(x)$ 的值落在该区间内.



接着寻找正数 δ, 使得去心邻域 $U^\circ(x_0, \delta)$ 内的每个 x 都满足 $|f(x)-L|<\varepsilon$.



最后在去心邻域 $U^\circ(x_0, \delta)$ 内任取 x, 验证 $|f(x)-L|<\varepsilon$. 几何上, 函数图像在垂直带 $x_0-\delta<x<x_0+\delta$ 内的部分落在水平带 $L-\varepsilon<y<L+\varepsilon$ 中.

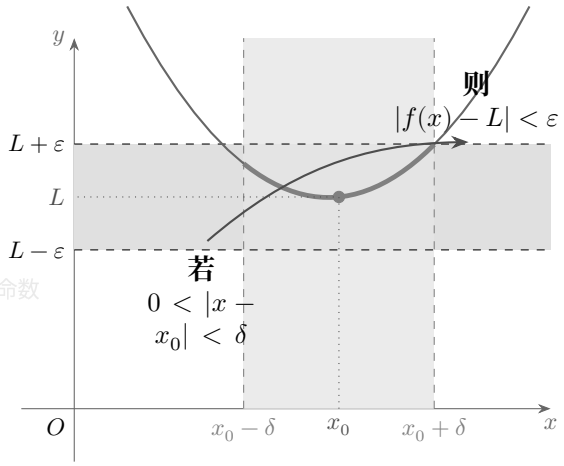
极限存在的含义可以用一个静态条件来描述. 给定任意正数 ε, 我们总能找到正数 δ, 使得只要 $0<|x-x_0|<\delta$, 就有 $|f(x)-L|<\varepsilon$.

依赖关系 $\delta(\varepsilon)$ 是极限定义的灵魂.ε 越小, 所要求的精度越高, 对应的 δ 通常也越小, 但它总存在.

若 $f(x)$ 并不真正趋近于 L, 则当 ε 足够小时, 不存在这样的 δ, 极限便不存在.

例题 16.3　使用 $\varepsilon-\delta$ 定义证明极限 $\lim\limits_{x\to 1} x^2=1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



解 我们用 $\varepsilon-\delta$ 定义直接验证. 对任意 $\varepsilon>0$ 需找到 $\delta>0$, 使当 $0<|x-1|<\delta$ 时恒有 $|x^2-1|<\varepsilon$. 为此, 我们把 $|x^2-1|$ 用 $|x-1|$ 与 $|x+1|$ 表出, 并对因子 $|x+1|$ 给出一个与 x 无关的上界, 从而确定 δ.

将目标式因式分解得 $|x^2-1|=|x-1|\cdot|x+1|$, 其中 $|x-1|$ 可由 δ 直接控制, 而 $|x+1|$ 随 x 变动. 因极限只关心 $x_0=1$ **附近**的行为, 我们预先限定 $\delta\leq 1$ 便可框住 $|x+1|$. 此时 $|x-1|<1$ 蕴含 $-1<x-1<1$, 两边加 2 得 $1<x+1<3$, 故 $|x+1|<3$. 代回变形即有

$$|x^2-1|=|x-1|\cdot|x+1|<\delta\cdot 3,$$

要其小于 ε, 只需 $3\delta\leq\varepsilon$, 即 $\delta\leq\dfrac{\varepsilon}{3}$. 合并两个约束, 取

$$\delta=\min\left\{1,\frac{\varepsilon}{3}\right\},$$

则 $\delta>0$.

对任意给定的 $\varepsilon>0$ 取上述 δ. 当 $0<|x-1|<\delta$ 时, 由 $\delta\leq 1$ 知 $|x-1|<1$, 从而 $1<x+1<3$ 即 $|x+1|<3$; 同时由 $\delta\leq\dfrac{\varepsilon}{3}$ 知 $|x-1|<\dfrac{\varepsilon}{3}$. 两估计结合得

$$|x^2-1|=|x-1|\cdot|x+1|<\frac{\varepsilon}{3}\cdot 3=\varepsilon,$$

即 $|x^2-1|<\varepsilon$ 成立. 由 $\varepsilon-\delta$ 定义, $\lim\limits_{x\to 1}x^2=1$.

16.2.1 本节习题

16.1. 考察函数 $f(x)=x-\lfloor x\rfloor$, 其中 $\lfloor x\rfloor$ 表示不大于 x 的最大整数. 研究函数 $f(x)$ 在 $x=1$ 处的极限是否存在.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.2. 运用极限的 $\varepsilon-\delta$ 定义, 证明 $\lim\limits_{x\to c}(ax+b)=ac+b$, 其中 a,b,c 均为实常数且 $a\neq 0$.

16.3. 运用极限的 $\varepsilon-\delta$ 定义, 证明 $\lim\limits_{x\to c}\sqrt{x}=\sqrt{c}$, 其中 c 为任意正实数.

16.4. 研究函数 $g(x)=\sin\left(\frac{1}{x}\right)$ 在 $x\to 0$ 时的极限.

16.5. 考虑极限 $\lim\limits_{x\to 2}x^3=8$. 对于给定的 $\varepsilon=0.1$, 请求出一个具体的正数 δ, 使得当 $0<|x-2|<\delta$ 时, 恒有 $|x^3-8|<0.1$.

16.6. 证明极限的唯一性. 若 $\lim\limits_{x\to x_0}f(x)=L_1$ 且 $\lim\limits_{x\to x_0}f(x)=L_2$, 则 $L_1=L_2$.

16.7. 运用极限的 $\varepsilon-\delta$ 定义, 证明 $\lim\limits_{x\to 0}\left(x\sin\frac{1}{x}\right)=0$.

16.8. 设函数 $f(x)$ 满足 $\lim\limits_{x\to c}f(x)=L$. 若 $L>0$, 证明存在一个正数 δ, 使得对于所有满足 $0<|x-c|<\delta$ 的 x, 都有 $f(x)>0$.

16.3 数列的极限 *

数列的每一项都对应数轴上的一个点. 当项数无限增大时, 这些点要么无限远离原点, 要么始终落在某个有界范围内. 后者中一部分无限靠近某个固定点, 另一部分在有限范围内持续摆动. 极限理论研究的是无限靠近某个固定点的情形.

芝诺悖论阿喀琉斯追不上乌龟反映了早期人类对无穷过程的困惑. 要精确描述并解决这个悖论, 需要一种能够量化无限逼近的数学工具.

再看一个具体例子. 几何级数 $S_n=\sum_{k=1}^{n}\frac{1}{2^k}=\frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^n}$ 可化为 $S_n=1-\frac{1}{2^n}$. 当 n 增大时, $1/2^n$ 趋近于 0, 整个和 S_n 趋近于 1.

极限理论用严格的逻辑语言把这种直观的无限逼近固定下来. 它连接离散数列与连续函数、有限运算与无限过程, 是微积分和现代数学分析的基础.

16.3.1 逼近的尝试与定义的演进

把无限逼近严格化, 需要剔除一些看似合理却不严谨的说法. 现给出三个逐渐精确的定义尝试.

16.3.1.0.1　越来越近　一种直观想法是, 若数列 $\{a_n\}$ 趋向于 L, 则它到 L 的距离应越来越小. **一个粗糙的定义(版本 1):** 数列 $\{a_n\}$ 的极限是 L, 如果对任意 $n\in\mathbb{N}^*$, 不等式 $|a_{n+1}-L|<|a_n-L|$ 恒成立. 这个定义要求距离严格单调递减. 它适用于 $a_n=1/n$ 这类极限为 0 的简单数列, 但限制太强, 会排除许多应被视为收敛的数列.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

考虑数列 $a_n=\dfrac{(-1)^n}{n}$. 其项为 $-1,\frac{1}{2},-\frac{1}{3},\frac{1}{4},\ldots$, 明显趋近于 0. 距离序列 $|a_n-0|=1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\ldots$ 满足版本 1 的定义. 但再看数列 $b_n=\begin{cases}1/n & \text{若}n\text{为奇数}\\ 1/(2n) & \text{若}n\text{为偶数}\end{cases}$. 其项为 $1,\frac{1}{4},\frac{1}{3},\frac{1}{8},\ldots$, 同样趋近于 0. 然而 $|b_3-0|=\frac{1}{3}>|b_2-0|=\frac{1}{4}$, 违反了版本 1 的要求, 于是版本 1 把这个数列错误地排除在外.

16.3.1.0.2 最终进入任何区间 版本 1 失败的原因在于对如何逼近限制过强. 于是我们放宽条件. 不再要求每一步都更近, 只要求最终进入任意一个包含 L 的区间.

一个粗糙的定义(版本 2): 数列 $\{a_n\}$ 的极限是 L, 如果对任意包含 L 的开区间 (a,b), 数列 $\{a_n\}$ 中只有限项在该区间之外. 这个定义已抓住极限的最终行为特征. 例如对前述反例 b_n, 无论包含 0 的区间多么小, 最终所有 b_n 都会落入其中.

但版本 2 仍不够精确, 也缺乏可操作性. 任意开区间的说法过于模糊, 无法直接代入不等式推导. 同时, 该定义没有突出 L 的中心地位. 区间 $(L-0.1,L+100)$ 与 $(L-0.1,L+0.1)$ 在其中没有区别.

16.3.1.0.3 $\epsilon-N$ 语言 为克服版本 2 的模糊性, 我们引入两个量化工具.

1. 用一个正数 ϵ 表示区间的任意小. 只考虑以 L 为中心的对称区间 $(L-\epsilon,L+\epsilon)$. 任何包含 L 的开区间都包含某个这样的对称区间, 因此进入所有对称区间等价于进入所有普通区间.ϵ 成为衡量误差的标尺.
2. 用一个正整数 N 表示最终. 只有限项在区间外等价于“从第 $N+1$ 项起, 所有项都在区间内”.N 区分初始项与最终规律.

把两者结合. 对任意给定的精度 ϵ, 我们总能找到足够大的项数分界点 N, 使得后续所有项都满足该精度.

16.3.2 极限的严格定义

定义 16.9 (数列极限的 $\epsilon-N$ 定义) 设 $\{a_n\}$ 是一个数列, L 是一个实数. 如果对任意给定的正数 ϵ, 总存在正整数 N, 使得当 $n>N$ 时, 不等式

$$|a_n-L|<\epsilon$$

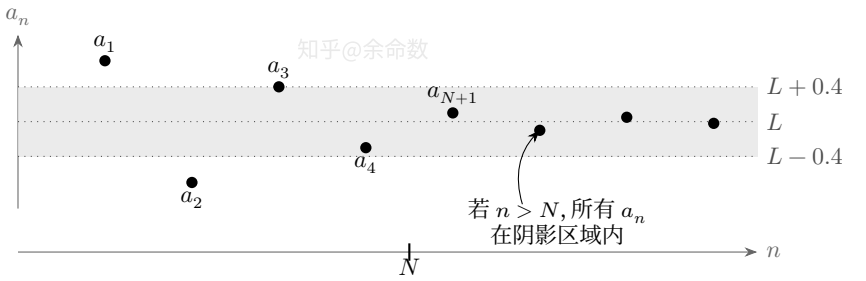
恒成立, 则称数列 $\{a_n\}$ **收敛**于 L, 记作

$$\lim_{n\to\infty}a_n=L \quad \text{或} \quad a_n\to L\ (n\to\infty)$$

如果一个数列不收敛于任何实数, 就称该数列**发散**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个定义完整量化了无限逼近的含义.



例题 16.4　证明: 一个常数数列 $a_n = c$ 收敛于 c.

解　解析: 需要证明 $\lim_{n\to\infty} c = c$. 根据定义, 对任意给定的 $\epsilon > 0$, 需要找到正整数 N, 使得当 $n > N$ 时, $|a_n - c| < \epsilon$ 成立.

对该数列, $|a_n - c| = |c - c| = 0$. 因此不等式变为 $0 < \epsilon$. 由于 ϵ 为正数, $0 < \epsilon$ 恒成立, 与 n 无关.

于是可取任意正整数作为 N, 例如 $N = 1$. 当 $n > 1$ 时, $|a_n - c| = 0 < \epsilon$ 恒成立. 故 $\lim_{n\to\infty} c = c$.

例题 16.5　证明: $\lim_{n\to\infty} \dfrac{1}{n} = 0$.

解　解析: 对任意给定的 $\epsilon > 0$, 需要找到正整数 N, 使得当 $n > N$ 时, $\left|\frac{1}{n} - 0\right| < \epsilon$ 成立.

该不等式化简为 $\dfrac{1}{n} < \epsilon$. 解得 $n > \dfrac{1}{\epsilon}$.

根据实数的阿基米德性质, 对任意正数 $1/\epsilon$, 总存在正整数比它大. 因此可取 $N = \lfloor \frac{1}{\epsilon} \rfloor + 1$, 先对 $1/\epsilon$ 向下取整, 再加 1, 这样 N 必为正整数.

这样, 当 $n > N$ 时, $n \geq N + 1 = \lfloor \frac{1}{\epsilon} \rfloor + 1 > \frac{1}{\epsilon}$. 这就保证了 $|\frac{1}{n} - 0| < \epsilon$ 成立. 证毕.

例题 16.6　证明: 若 $|q| < 1$, 则 $\lim_{n\to\infty} q^n = 0$.

解　解析: 若 $q = 0$, 则数列为 $0, 0, \ldots$, 收敛于 0.

考虑 $0 < |q| < 1$ 的情况. 由于 $1/|q| > 1$, 令 $\dfrac{1}{|q|} = 1 + h$, 其中 $h > 0$. 于是 $|q| = \dfrac{1}{1+h}$.

目标是证明对任意 $\epsilon > 0$, 存在 N 使得当 $n > N$ 时, $|q^n - 0| = |q|^n < \epsilon$. 即 $\left(\dfrac{1}{1+h}\right)^n < \epsilon$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

根据伯努利不等式, $(1+h)^n \geq 1+nh$. 因此,

$$|q|^n = \frac{1}{(1+h)^n} \leq \frac{1}{1+nh}$$

要使 $|q|^n < \epsilon$, 只需保证 $\dfrac{1}{1+nh} < \epsilon$. 解此不等式:

$$1 < \epsilon(1+nh) \iff \frac{1}{\epsilon} < 1+nh \iff nh > \frac{1}{\epsilon} - 1 \iff n > \frac{1}{h}\left(\frac{1}{\epsilon} - 1\right)$$

对任意给定的 $\epsilon > 0$, 总可取正整数 N 大于 $\dfrac{1}{h}\left(\dfrac{1}{\epsilon} - 1\right)$. 当 $n > N$ 时, $|q^n| < \epsilon$ 成立. 证毕.

例题 16.7 证明: 数列 $a_n = (-1)^n$ 是发散的.

解 **解析:** 使用反证法. 假设该数列收敛, 极限为 L.

根据极限定义, 对任意 $\epsilon > 0$, 都存在 N, 使得当 $n > N$ 时, $|a_n - L| < \epsilon$.

知乎@余命数

取 $\epsilon = \dfrac{1}{2}$. 由于数列的项在 1 与 -1 之间跳动, 两者距离为 2, 这个取值足够小.

若极限存在, 则对 $\epsilon = 1/2$, 必存在 N, 使得所有 $n > N$ 的项都满足 $|a_n - L| < 1/2$.

即对某个充分大的偶数 $n_{even} > N$,

$$|a_{n_{even}} - L| = |1 - L| < \frac{1}{2}$$

知乎@余命数

同时, 对某个充分大的奇数 $n_{odd} > N$,

$$|a_{n_{odd}} - L| = |-1 - L| < \frac{1}{2}$$

考察 1 与 -1 之间的距离, 并用三角不等式:

$$\begin{aligned} 2 &= |1 - (-1)| \\ &= |(1 - L) + (L - (-1))| \\ &\leq |1 - L| + |L - (-1)| \\ &= |1 - L| + |-1 - L| \end{aligned}$$

知乎@余命数

知乎@余命数

由上面两个不等式,

$$|1 - L| + |-1 - L| < \frac{1}{2} + \frac{1}{2} = 1$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

结合得到 $2<1$, 矛盾. 这说明最初的假设(数列收敛于某个极限 L)不成立. 因此, 数列 $a_n=(-1)^n$ 发散.

例题 16.8　使用 $\epsilon-N$ 定义证明 $\lim_{n\to\infty}\dfrac{2n-1}{n+1}=2$.

证明　**证明:** 对任意给定的 $\epsilon>0$, 目标是寻找正整数 N, 使得当 $n>N$ 时, 不等式 $\left|\dfrac{2n-1}{n+1}-2\right|<\epsilon$ 恒成立.

先化简不等式左侧:

$$\left|\frac{2n-1}{n+1}-2\right|=\left|\frac{2n-1-2(n+1)}{n+1}\right|=\left|\frac{-3}{n+1}\right|=\frac{3}{n+1}$$

此处利用 n 为正整数, 故 $n+1>0$.

接着确定需要多大的 n 才能保证 $\dfrac{3}{n+1}<\epsilon$. 解此不等式,

$$\frac{3}{n+1}<\epsilon \iff 3<\epsilon(n+1) \iff n+1>\frac{3}{\epsilon} \iff n>\frac{3}{\epsilon}-1$$

因此, 只要 n 大于 $\frac{3}{\epsilon}-1$, 期望的不等式便成立.

于是, 对任意给定的 $\epsilon>0$, 可取正整数 $N=\max\left(1,\lfloor\frac{3}{\epsilon}\rfloor\right)$. 当 $n>N$ 时, 必有 $n>\frac{3}{\epsilon}$, 从而 $n>\frac{3}{\epsilon}-1$ 也成立.

这便完成了证明. 根据定义, $\lim_{n\to\infty}\dfrac{2n-1}{n+1}=2$. □

16.3.3 极限的运算法则

若每次计算极限都直接套用 $\epsilon-N$ 定义, 过程会极其繁琐. 极限运算与代数运算具有良好的相容性, 可以把复杂数列的极限问题分解为若干基本部分. 现逐一证明这些法则.

定理 16.3 (极限的运算法则) 如果 $\lim_{n\to\infty}a_n=A$ 且 $\lim_{n\to\infty}b_n=B$, 那么:

- **和差法则:** $\lim_{n\to\infty}(a_n\pm b_n)=A\pm B$.
- **乘法法则:** $\lim_{n\to\infty}(a_n\cdot b_n)=A\cdot B$.
- **常数倍数法则:** $\lim_{n\to\infty}(c\cdot a_n)=c\cdot A$, 其中 c 为常数.
- **除法法则:** 如果 $B\neq0$, 那么 $\lim_{n\to\infty}\dfrac{a_n}{b_n}=\dfrac{A}{B}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 (和差法则) 仅证明和法则,差法则的证明与之类似. 目标是证明对任意 $\epsilon > 0$, 存在正整数 N, 使得当 $n > N$ 时, $|(a_n + b_n) - (A + B)| < \epsilon$. 从目标表达式出发, 用三角不等式 $|x + y| \leq |x| + |y|$ 放缩:

$$|(a_n + b_n) - (A + B)| = |(a_n - A) + (b_n - B)| \leq |a_n - A| + |b_n - B|$$

要使右侧之和小于 ϵ, 可让两部分都小于 $\epsilon/2$.

因为 $\lim_{n\to\infty} a_n = A$, 所以对正数 $\epsilon/2$, 存在正整数 N_1, 使得当 $n > N_1$ 时,

$$|a_n - A| < \frac{\epsilon}{2}$$

同理, 因为 $\lim_{n\to\infty} b_n = B$, 存在正整数 N_2, 使得当 $n > N_2$ 时,

$$|b_n - B| < \frac{\epsilon}{2}$$

为使两个不等式同时成立, 取 $N = \max(N_1, N_2)$. 当 $n > N$ 时,

$$|(a_n + b_n) - (A + B)| \leq |a_n - A| + |b_n - B| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

□

16.3.3.1 证明乘法法则所需引理

在证明乘法法则之前, 先建立一个引理.

引理 16.1 一个收敛的数列必然是有界的.

证明 设 $\lim_{n\to\infty} a_n = A$. 根据定义, 对 $\epsilon = 1$, 存在正整数 N, 使得当 $n > N$ 时, $|a_n - A| < 1$. 即对 $n > N$, 有 $A - 1 < a_n < A + 1$. 因此从第 $N + 1$ 项起的所有项都落在区间 $(A-1, A+1)$ 内. 令 $M = \max\{|a_1|, |a_2|, \ldots, |a_N|, |A-1|, |A+1|\}$. 则对所有 $n \in \mathbb{N}^*$, 均有 $|a_n| \leq M$. 故 $\{a_n\}$ 有界. □

证明 (乘法法则) 目标是估计 $|a_n b_n - AB|$ 的大小. 技巧是引入中间项 $a_n B$, 再应用三角不等式:

$$\begin{aligned} |a_n b_n - AB| &= |a_n b_n - a_n B + a_n B - AB| \\ &\leq |a_n b_n - a_n B| + |a_n B - AB| \\ &= |a_n(b_n - B)| + |B(a_n - A)| \\ &= |a_n||b_n - B| + |B||a_n - A| \end{aligned}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

由于数列 $\{a_n\}$ 收敛,根据引理,它有界. 即存在常数 $M>0$ 使得对所有 n,$|a_n|\le M$. 因此,

$$|a_nb_n-AB|\le M|b_n-B|+|B||a_n-A|$$

目标是使右端小于给定的 ϵ. 可让两部分都小于 $\epsilon/2$. (为避免分母为零, 在分母中使用 $|B|+1$ 代替 $|B|$,这是常用技巧.)因为 $\lim_{n\to\infty}a_n=A$,所以对正数 $\dfrac{\epsilon}{2(|B|+1)}$,存在 N_1, 当 $n>N_1$ 时,

$$|a_n-A|<\frac{\epsilon}{2(|B|+1)}$$

因为 $\lim_{n\to\infty}b_n=B$, 所以对正数 $\dfrac{\epsilon}{2M}$, 存在 N_2, 当 $n>N_2$ 时,

$$|b_n-B|<\frac{\epsilon}{2M}$$

取 $N=\max(N_1,N_2)$. 当 $n>N$ 时,

$$\begin{aligned}|a_nb_n-AB|&\le M|b_n-B|+|B||a_n-A|\\&<M\cdot\frac{\epsilon}{2M}+|B|\cdot\frac{\epsilon}{2(|B|+1)}\\&=\frac{\epsilon}{2}+\frac{|B|}{|B|+1}\cdot\frac{\epsilon}{2}\\&\le\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon\end{aligned}$$

□

证明 (常数倍数法则) 若 $c=0$, 则 $ca_n=0$ 对所有 n 成立, 结论显然成立. 若 $c\neq0$, 对任意 $\epsilon>0$, 由 $a_n\to A$ 可知存在 N, 当 $n>N$ 时,

$$|a_n-A|<\frac{\epsilon}{|c|}.$$

于是

$$|ca_n-cA|=|c|\,|a_n-A|<\epsilon.$$

因此 $\lim_{n\to\infty}ca_n=cA$. □

证明 (除法法则) 根据乘法法则, 只需证明 $\lim_{n\to\infty}\dfrac{1}{b_n}=\dfrac{1}{B}$. 目标是估计 $\left|\dfrac{1}{b_n}-\dfrac{1}{B}\right|=\left|\dfrac{B-b_n}{b_nB}\right|=\dfrac{|b_n-B|}{|b_n||B|}$. 难点在于分母中的 $|b_n|$ 是变量, 需要为它找到一个正的下界.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因为 $\lim_{n\to\infty} b_n = B$ 且 $B \neq 0$, 所以 $|B| > 0$. 取 $\epsilon_1 = \dfrac{|B|}{2} > 0$. 根据极限定义, 存在正整数 N_1, 使得当 $n > N_1$ 时,

$$|b_n - B| < \frac{|B|}{2}$$

利用反向三角不等式 $|x| - |y| \leq |x - y|$, 得到

$$|B| - |b_n| \leq |B - b_n| < \frac{|B|}{2}$$

移项得 $|b_n| > |B| - \frac{|B|}{2} = \frac{|B|}{2}$. 这就给出了所需下界. 从第 N_1 项起, $|b_n|$ 远离 0.

对 $n > N_1$,

$$\left|\frac{1}{b_n} - \frac{1}{B}\right| = \frac{|b_n - B|}{|b_n||B|} < \frac{|b_n - B|}{(\frac{|B|}{2})|B|} = \frac{2|b_n - B|}{|B|^2}$$

现在, 对任意给定的 $\epsilon > 0$, 希望上式右端小于 ϵ. 即 $|b_n - B| < \dfrac{\epsilon|B|^2}{2}$. 因为 $\lim_{n\to\infty} b_n = B$, 所以对正数 $\dfrac{\epsilon|B|^2}{2}$, 存在正整数 N_2, 使得当 $n > N_2$ 时,

$$|b_n - B| < \frac{\epsilon|B|^2}{2}$$

取 $N = \max(N_1, N_2)$. 当 $n > N$ 时,

$$\left|\frac{1}{b_n} - \frac{1}{B}\right| < \frac{2}{|B|^2} \cdot |b_n - B| < \frac{2}{|B|^2} \cdot \frac{\epsilon|B|^2}{2} = \epsilon$$

这证明了 $\lim_{n\to\infty} \dfrac{1}{b_n} = \dfrac{1}{B}$. 结合乘法法则,

$$\lim_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} \left(a_n \cdot \frac{1}{b_n}\right) = \left(\lim_{n\to\infty} a_n\right) \cdot \left(\lim_{n\to\infty} \frac{1}{b_n}\right) = A \cdot \frac{1}{B} = \frac{A}{B}$$

□

16.3.3.2 幂法则的说明

对任意正整数 k, 可通过对乘法法则进行数学归纳证明 $\lim_{n\to\infty} a_n^k = A^k$. 当 $k = 1$ 时, 结论显然成立. 假设当 $k = m$ 时结论成立, 即 $\lim_{n\to\infty} a_n^m = A^m$. 那么对 $k = m + 1$,

$$\lim_{n\to\infty} a_n^{m+1} = \lim_{n\to\infty} (a_n^m \cdot a_n) = \left(\lim_{n\to\infty} a_n^m\right) \cdot \left(\lim_{n\to\infty} a_n\right) = A^m \cdot A = A^{m+1}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故对所有正整数 k, 该法则成立. 对负整数 k, 可由除法法则导出. 对一般实数指数 k, 严格证明需要借助指数函数与对数函数的连续性, 此处只陈述结论.

例题 16.9 (有理函数极限) 设 $P(n)=a_pn^p+\cdots+a_0$ 和 $Q(n)=b_qn^q+\cdots+b_0$ 是两个关于 n 的多项式, 其中 $a_p\neq0, b_q\neq0$. 分析极限 $\lim_{n\to\infty}\dfrac{P(n)}{Q(n)}$.

解 将分子分母同除以分母的最高次幂 n^q.

知乎@余命数

$$\lim_{n\to\infty}\frac{P(n)}{Q(n)}=\lim_{n\to\infty}\frac{a_pn^{p-q}+a_{p-1}n^{p-q-1}+\cdots+a_0n^{-q}}{b_q+b_{q-1}n^{-1}+\cdots+b_0n^{-q}}$$

分母的极限为 $\lim_{n\to\infty}(b_q+\frac{b_{q-1}}{n}+\ldots)=b_q\neq0$. 因此整个极限的行为由分子的行为决定.

知乎@余命数

分三种情况讨论:

- **当 $p<q$ 时:** $p-q<0$. 分子中 n 的所有幂次均为负. 当 $n\to\infty$ 时, 分子每一项都趋于 0. 故分子的极限为 0. 因此, $\lim_{n\to\infty}\dfrac{P(n)}{Q(n)}=\dfrac{0}{b_q}=0$.
- **当 $p=q$ 时:** $p-q=0$. 分子变为 $a_p+\dfrac{a_{p-1}}{n}+\ldots$. 其极限为 a_p. 因此, $\lim_{n\to\infty}\dfrac{P(n)}{Q(n)}=\dfrac{a_p}{b_p}$.
- **当 $p>q$ 时:** $p-q>0$. 分子的首项 a_pn^{p-q} 趋于 ∞ 或 $-\infty$ (取决于 a_p 的符号). 其余项的增长速度都低于首项. 故整个分子无界. 因此, 数列发散, 极限不存在.

知乎@余命数

由此可知, 两个多项式之比的极限仅取决于它们的最高次幂. 若分母次数更高, 极限为 0; 若次数相同, 极限为最高次项系数之比; 若分子次数更高, 数列发散.

例题 16.10 (根式有理化) 计算极限 $\lim_{n\to\infty}(\sqrt{n^2+an}-n)$, 其中 a 为常数.

解 **解析:** 当 n 足够大时, $n^2+an=n(n+a)>0$, 根式有意义. 这是 $\infty-\infty$ 型的不定式, 不能直接计算. 处理这类根式差的标准做法是乘以共轭表达式, 把减法转化为加法.

知乎@余命数

$$\begin{aligned}\sqrt{n^2+an}-n&=(\sqrt{n^2+an}-n)\cdot\frac{\sqrt{n^2+an}+n}{\sqrt{n^2+an}+n}\\&=\frac{(n^2+an)-n^2}{\sqrt{n^2+an}+n}\\&=\frac{an}{\sqrt{n^2+an}+n}\end{aligned}$$

变换后原不定式化为 $\frac{\infty}{\infty}$ 型, 可用提取主项的方法处理. 分母中 n 的最高次幂等价于 n^1. 将分子分母同除以 n.

$$\frac{an/n}{(\sqrt{n^2+an}+n)/n}=\frac{a}{\frac{\sqrt{n^2+an}}{n}+1}=\frac{a}{\sqrt{\frac{n^2+an}{n^2}}+1}=\frac{a}{\sqrt{1+\frac{a}{n}}+1}$$

知乎@余命数

对变换后的表达式取极限.

知乎@余命数

$$\lim_{n\to\infty}\frac{a}{\sqrt{1+\frac{a}{n}}+1}=\frac{a}{\sqrt{1+0}+1}=\frac{a}{2}$$

结果 $\frac{a}{2}$ 是微积分中常用的经典结论.

例题 16.11 (阶乘与指数的增长率比较) 我们证明, 对任意实数 a, 下列极限成立.

$$\lim_{n\to\infty}\frac{a^n}{n!}=0.$$

解 解析: 这个极限说明阶乘增长的速度远快于指数增长. 若 $a=0$, 结论显然成立. 若 $a\neq 0$, 考虑其绝对值. 令 $x_n=\frac{|a|^n}{n!}$.

考察相邻项的比值:

知乎@余命数

$$\frac{x_{n+1}}{x_n}=\frac{|a|^{n+1}/(n+1)!}{|a|^n/n!}=\frac{|a|}{n+1}$$

当 n 足够大时, 该比值会变得任意小. 取正整数 $N>2|a|$. 那么对所有 $n\geq N$,

$$\frac{|a|}{n+1}<\frac{|a|}{N}<\frac{1}{2}$$

知乎@余命

即从第 N 项起, 数列每一项至多是前一项的一半.

对 $n>N$,

$$x_n=x_N\cdot\frac{|a|}{N+1}\cdot\frac{|a|}{N+2}\cdots\frac{|a|}{n}<x_N\cdot\left(\frac{1}{2}\right)^{n-N}$$

知乎@余命数

令 $C=x_N\cdot(1/2)^{-N}=x_N\cdot 2^N$, 这是与 n 无关的常数. 于是当 $n>N$ 时,

$$0\leq x_n<C\cdot\left(\frac{1}{2}\right)^n$$

由于 $\lim_{n\to\infty}C\cdot\left(\frac{1}{2}\right)^n=0$, 根据夹逼准则(有限项不影响极限), $\lim_{n\to\infty}x_n=0$.

因为 $\lim_{n\to\infty}|\frac{a^n}{n!}|=0$, 原数列的极限也必然为 0.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.3.4 不定型极限的计算策略

直接应用极限运算法则时，常会遇到没有直接定义的表达式，例如 $\frac{\infty}{\infty}, \frac{0}{0}, \infty-\infty, 0\cdot\infty, 1^{\infty}, 0^{0}, \infty^{0}$ 等. 这些形式称为**不定型极限**. 具体结果取决于数列结构，需要进一步变形分析.

16.3.4.1 $\frac{\infty}{\infty}$ 型

知乎@余命数

这种不定型最常见，尤其出现在有理函数或含根式的分式中.

处理办法是**提取主导项**. 在分子和分母中找出 $n\to\infty$ 时增长最快的项（主项），然后将分子分母同除以分母的主项，把原式化为分母极限非零、分子极限易求的形式.

例题 16.12 (有理函数) 计算 $\lim_{n\to\infty}\dfrac{3n^2-2n+5}{2n^2+4n-1}$.

解 解析：分子分母的主项均为 n^2. 同除以 n^2.

$$\lim_{n\to\infty}\frac{3-\frac{2}{n}+\frac{5}{n^2}}{2+\frac{4}{n}-\frac{1}{n^2}}=\frac{3-0+0}{2+0-0}=\frac{3}{2}$$

例题 16.13 (根式分式) 计算 $\lim_{n\to\infty}\dfrac{\sqrt{4n^2+n}+n}{3n-\sqrt{n^2+1}}$.

解 解析：分子中，$\sqrt{4n^2+n}$ 的增长速度与 $\sqrt{4n^2}=2n$ 相当. 因此分子主项等价于 $2n+n=3n$. 分母中，$\sqrt{n^2+1}$ 的增长速度与 $\sqrt{n^2}=n$ 相当. 因此分母主项等价于 $3n-n=2n$. 整个分式的主项等价于 n. 将分子分母同除以 n.

$$\begin{aligned}\lim_{n\to\infty}\frac{\frac{\sqrt{4n^2+n}}{n}+\frac{n}{n}}{\frac{3n}{n}-\frac{\sqrt{n^2+1}}{n}}&=\lim_{n\to\infty}\frac{\sqrt{\frac{4n^2+n}{n^2}}+1}{3-\sqrt{\frac{n^2+1}{n^2}}}\\&=\lim_{n\to\infty}\frac{\sqrt{4+\frac{1}{n}}+1}{3-\sqrt{1+\frac{1}{n^2}}}\\&=\frac{\sqrt{4+0}+1}{3-\sqrt{1+0}}=\frac{2+1}{3-1}=\frac{3}{2}\end{aligned}$$

16.3.4.2 $\infty-\infty$ 型

这种不定型出现在两个都趋于无穷的项相减时，结果取决于两者增长速度的差异.

处理思路是用代数变换把减法不定型转化为除法不定型.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

- **对于分式之差:**进行**通分**, 合并成一个分式, 再化为 $\frac{\infty}{\infty}$ 或 $\frac{0}{0}$ 型.
- **对于根式之差:**用共轭表达式进行分子或分母有理化, 把减法变为加法, 再化成分式.

知乎@余命数

例题 16.14 (根式有理化) 计算 $\lim_{n\to\infty}(\sqrt{n^2+an}-\sqrt{n^2+bn})$.

解 解析: 当 n 足够大时, 两个根式均有意义. 这是 $\infty-\infty$ 型不定式. 乘以其共轭表达式.

$$\begin{aligned}&\lim_{n\to\infty}(\sqrt{n^2+an}-\sqrt{n^2+bn})\cdot\frac{\sqrt{n^2+an}+\sqrt{n^2+bn}}{\sqrt{n^2+an}+\sqrt{n^2+bn}}\\=&\lim_{n\to\infty}\frac{(n^2+an)-(n^2+bn)}{\sqrt{n^2+an}+\sqrt{n^2+bn}}\\=&\lim_{n\to\infty}\frac{(a-b)n}{\sqrt{n^2+an}+\sqrt{n^2+bn}}\end{aligned}$$

知乎@余命数

问题转化为 $\frac{\infty}{\infty}$ 型. 同除以主项 n.

$$=\lim_{n\to\infty}\frac{a-b}{\sqrt{1+\frac{a}{n}}+\sqrt{1+\frac{b}{n}}}=\frac{a-b}{\sqrt{1+0}+\sqrt{1+0}}=\frac{a-b}{2}$$

知乎@余命数

16.3.4.3 1^∞ 型

形如 $\lim_{n\to\infty}(u_n)^{v_n}$, 其中 $\lim u_n=1$ 且 $\lim v_n=\infty$.

处理办法是把表达式恒等变形为重要极限 $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=e$ 的标准形式. 令 $u_n=1+\alpha_n$, 其中 $\lim\alpha_n=0$:

$$\lim_{n\to\infty}(1+\alpha_n)^{v_n}=\lim_{n\to\infty}\left[(1+\alpha_n)^{\frac{1}{\alpha_n}}\right]^{\alpha_n v_n}$$

方括号内的部分趋向于 e, 因此整个极限的结果取决于指数部分 $\lim_{n\to\infty}(\alpha_n v_n)$. 若 $\lim_{n\to\infty}\alpha_n v_n=L$, 则原极限为 e^L.

例题 16.15 计算 $\lim_{n\to\infty}\left(\frac{n+2}{n-1}\right)^{2n+1}$.

知乎@余命数

解 解析: 当 $n\to\infty$ 时, 底数 $\frac{n+2}{n-1}\to1$, 指数 $2n+1\to\infty$. 这是 1^∞ 型. 把底数改写为 $(1+\alpha_n)$ 的形式:

$$\frac{n+2}{n-1}=\frac{(n-1)+3}{n-1}=1+\frac{3}{n-1}$$

其中 $\alpha_n = \dfrac{3}{n-1}$.

原极限变为 $\lim_{n\to\infty}\left(1+\frac{3}{n-1}\right)^{2n+1}$. 进行指数配凑:

$$\lim_{n\to\infty}\left[\left(1+\frac{3}{n-1}\right)^{\frac{n-1}{3}}\right]^{\frac{3}{n-1}\cdot(2n+1)}$$

令 $m=\frac{n-1}{3}$. 当 $n\to\infty$ 时, $m\to\infty$. 方括号内的部分 $\lim_{m\to\infty}\left(1+\frac{1}{m}\right)^m = e$.

只需计算指数部分的极限, 得

$$\lim_{n\to\infty}\frac{3(2n+1)}{n-1} = \lim_{n\to\infty}\frac{6n+3}{n-1} = 6$$

因此, 原极限等于 e^6.

16.3.5 极限判敛准则

在很多情况下, 极限的值无法直接算出, 甚至不确定是否存在. 现两个准则是判断收敛性并求解极限的基本工具.

定理 16.4 (夹逼准则) 设有三个数列 $\{a_n\},\{b_n\},\{c_n\}$. 如果从某一项 $(n>N_0)$ 开始, 恒有

$$a_n \le b_n \le c_n$$

并且 $\{a_n\}$ 与 $\{c_n\}$ 都收敛于同一极限 L, 那么位于它们之间的数列 $\{b_n\}$ 也收敛于 L.

证明 我们要证明,对任意给定的 $\epsilon>0$,存在正整数 N,使得当 $n>N$ 时,$|b_n-L|<\epsilon$ 恒成立. 这个不等式等价于 $L-\epsilon<b_n<L+\epsilon$.

因为 $\lim_{n\to\infty}a_n = L$, 根据极限定义, 对上述 $\epsilon>0$, 存在正整数 N_1, 使得当 $n>N_1$ 时, 恒有

$$|a_n - L| < \epsilon$$

这等价于 $L-\epsilon<a_n<L+\epsilon$. 证明中需要左半部分:

$$L-\epsilon < a_n$$

同理, 因为 $\lim_{n\to\infty}c_n = L$, 存在正整数 N_2, 使得当 $n>N_2$ 时, 恒有

$$|c_n - L| < \epsilon$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

这等价于 $L-\epsilon<c_n<L+\epsilon$. 证明中需要右半部分:

$$c_n<L+\epsilon$$

根据题设, 还存在正整数 N_0, 使得当 $n>N_0$ 时, $a_n\le b_n\le c_n$ 成立.

为使所有不等式同时成立, 取 $N=\max(N_0,N_1,N_2)$.

当 $n>N$ 时,

$$L-\epsilon<a_n\quad(\text{因为}n>N\ge N_1)$$

$$a_n\le b_n\quad(\text{因为}n>N\ge N_0)$$

$$b_n\le c_n\quad(\text{因为}n>N\ge N_0)$$

$$c_n<L+\epsilon\quad(\text{因为}n>N\ge N_2)$$

串联得到

$$L-\epsilon<a_n\le b_n\le c_n<L+\epsilon$$

因此

$$L-\epsilon<b_n<L+\epsilon$$

即 $|b_n-L|<\epsilon$.

证明完成. 根据极限定义, $\lim_{n\to\infty}b_n=L$. □

夹逼准则的含义是, 若一个数列被两个收敛于同一极限的数列从上下两侧夹住, 则它也必须收敛于该极限. 它的作用在于用两个易处理的数列控制一个复杂的数列, 从而确定后者的极限.

例题 16.16 求极限 $\lim_{n\to\infty}\sum_{k=1}^{n}\dfrac{1}{\sqrt{n^2+k}}$.

解 解析: 该和式的项数随 n 变化, 不能直接用运算法则. 这是应用夹逼准则的典型情形. 令 $S_n=\dfrac{1}{\sqrt{n^2+1}}+\dfrac{1}{\sqrt{n^2+2}}+\cdots+\dfrac{1}{\sqrt{n^2+n}}$. 通过对和式中的每一项进行统一放缩, 构造 S_n 的上下界.

在和式各项中, 分母最小的是 $\sqrt{n^2+1}$, 故 $\dfrac{1}{\sqrt{n^2+1}}$ 是最大项. 分母最大的是 $\sqrt{n^2+n}$, 故 $\dfrac{1}{\sqrt{n^2+n}}$ 是最小项.

对上界, 把每一项都放大为最大项:

$$S_n=\sum_{k=1}^{n}\frac{1}{\sqrt{n^2+k}}\le\sum_{k=1}^{n}\frac{1}{\sqrt{n^2+1}}=n\cdot\frac{1}{\sqrt{n^2+1}}=\frac{n}{\sqrt{n^2+1}}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

对下界, 把每一项都缩小为最小项:

$$S_n = \sum_{k=1}^{n} \frac{1}{\sqrt{n^2+k}} \geq \sum_{k=1}^{n} \frac{1}{\sqrt{n^2+n}} = n \cdot \frac{1}{\sqrt{n^2+n}} = \frac{n}{\sqrt{n^2+n}}$$

于是得到

$$\frac{n}{\sqrt{n^2+n}} \leq S_n \leq \frac{n}{\sqrt{n^2+1}}$$

接着计算两边数列的极限.

$$\lim_{n\to\infty} \frac{n}{\sqrt{n^2+n}} = \lim_{n\to\infty} \frac{n}{n\sqrt{1+1/n}} = \lim_{n\to\infty} \frac{1}{\sqrt{1+1/n}} = 1$$

$$\lim_{n\to\infty} \frac{n}{\sqrt{n^2+1}} = \lim_{n\to\infty} \frac{n}{n\sqrt{1+1/n^2}} = \lim_{n\to\infty} \frac{1}{\sqrt{1+1/n^2}} = 1$$

由于上下界的极限均为 1, 根据夹逼准则, 原极限值为 1.

知乎@余命数

定理 16.5 (单调收敛定理) 一个单调且有界的数列必然收敛.

- 非减且有上界的数列收敛.
- 非增且有下界的数列收敛.

知乎@余命数
@余命数

证明 这个定理体现了实数系的完备性, 也说明极限与序结构之间的紧密联系. 先证第一种情况, 第二种情况的证明与之对偶.

知乎@余命数

设 $\{a_n\}$ 是非减且有上界的数列. 非减意味着对所有 $n \in \mathbb{N}^*$, 有 $a_{n+1} \geq a_n$. 有上界意味着存在实数 M, 使得对所有 $n \in \mathbb{N}^*$, 有 $a_n \leq M$.

考虑由所有项构成的集合 $S = \{a_1, a_2, a_3, \dots\}$. 集合 S 非空, 且有上界 M.

根据实数系的**完备性公理**(或称最小上界公理), 任何非空且有上界的实数集都有最小上界, 即**上确界**. 令 $L = \sup S$. 现证 $\lim_{n\to\infty} a_n = L$.

知乎@余命数

要证明 $\lim_{n\to\infty} a_n = L$, 需依据 $\epsilon - N$ 定义. 对任意给定的 $\epsilon > 0$, 需要找到正整数 N, 使得当 $n > N$ 时, $|a_n - L| < \epsilon$ 成立.

由上确界定义, L 是集合 S 的上界, 因此对所有 n, 都有 $a_n \leq L$. 由于 $\epsilon > 0$, $L < L + \epsilon$, 故 $a_n < L + \epsilon$ 自动成立.

知乎@余命数

再证明 $a_n > L - \epsilon$. 由于 $\epsilon > 0$, 有 $L - \epsilon < L$. 根据上确界的最小性, 任何严格小于 L 的数都不是集合 S 的上界. 因此 $L - \epsilon$ 不是 S 的上界, 集合 S 中必存在某个元素大于 $L - \epsilon$. 即存在正整数 N, 使得 $a_N > L - \epsilon$.

知乎@余命数

由于 $\{a_n\}$ 非减, 对任何 $n > N$, 都有 $a_n \geq a_N$.

于是对任何 $n > N$:

$$L - \epsilon < a_N \leq a_n$$

结合 $a_n \le L < L+\epsilon$, 得到对所有 $n > N$:

$$L - \epsilon < a_n < L + \epsilon$$

这等价于 $|a_n - L| < \epsilon$.

因此, 对任意 $\epsilon > 0$, 都能找到对应的 N. 根据极限定义, 数列 $\{a_n\}$ 收敛于 L.

对非增且有下界的数列, 证明完全类似, 只需考虑集合的**下确界**, 并证明数列收敛于下确界. 证毕. □

单调收敛定理是**实数系 $\mathbb{R}$ 的标志性特征**, 其证明完全依赖实数系的完备性公理. 若把范围限制在有理数系 $\mathbb{Q}$ 中, 该定理不再成立.

考虑有理数列 $\{a_n\}$:

$$a_1 = 1, a_2 = 1.4, a_3 = 1.41, a_4 = 1.414, ...$$

这是 $\sqrt{2}$ 的十进制逐次逼近. 该数列满足:

- **单调递增**: $a_{n+1} > a_n$.
- **有上界**: 例如所有项都小于 2.

在有理数系中它满足定理的全部条件, 却不收敛, 因为它的极限 $\sqrt{2}$ 不属于 $\mathbb{Q}$. 有理数轴上存在这样的空洞, 单调有界数列可能通向一个不存在的点.

这个反例说明实数系的完备性填补了有理数轴上的这些空洞. 正因如此, 单调有界数列的极限总能在实数系中找到, 这个极限就是它的上确界或下确界.

例题 16.17 已知数列 $\{a_n\}$ 满足 $a_1 = \sqrt{2}, a_{n+1} = \sqrt{2 + a_n}$. 证明该数列收敛, 并求其极限.

解 解析: 该数列由递推关系定义, 通项不易求出. 现用单调收敛定理分析其有界性与单调性.

先用数学归纳法证明数列有上界 2. 当 $n = 1$ 时, $a_1 = \sqrt{2} < 2$, 结论成立. 假设当 $n = k\,(k \ge 1)$ 时, $a_k < 2$ 成立. 则对 $n = k+1$, $a_{k+1} = \sqrt{2 + a_k} < \sqrt{2+2} = 2$. 故对所有正整数 n, $a_n < 2$ 恒成立. 又显然 $a_n > 0$, 所以数列 $\{a_n\}$ 有界.

再判断单调性. 考察相邻两项的平方之差:

$$a_{n+1}^2 - a_n^2 = (2 + a_n) - a_n^2 = -(a_n^2 - a_n - 2) = -(a_n - 2)(a_n + 1)$$

由于 $0 < a_n < 2$, 可知 $a_n - 2 < 0$ 且 $a_n + 1 > 0$. 因此 $a_{n+1}^2 - a_n^2 > 0$, 即 $a_{n+1}^2 > a_n^2$. 因为数列各项均为正, 所以 $a_{n+1} > a_n$. 故数列 $\{a_n\}$ 严格递增.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此 $\{a_n\}$ 单调递增且有上界. 根据单调收敛定理, 该数列收敛. 设 $\lim_{n\to\infty} a_n = L$. 对递推关系式 $a_{n+1} = \sqrt{2+a_n}$ 两边同时取极限. 由于 $\lim_{n\to\infty} a_{n+1} = L$, 并且平方根函数连续,

$$L = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \sqrt{2+a_n} = \sqrt{2+\lim_{n\to\infty} a_n} = \sqrt{2+L}$$

得到关于 L 的方程 $L = \sqrt{2+L}$. 两边平方得 $L^2 = 2+L$, 即 $L^2 - L - 2 = 0$, 解得 $L = 2$ 或 $L = -1$. 由于数列 $\{a_n\}$ 的所有项均为正数, 其极限必为非负数, 故舍去 $L = -1$.

于是该数列收敛, 且极限为 2.

16.3.6 两个重要的极限

以下两个极限在更复杂的极限计算中常被作为已知结论引用.

定理 16.6

$$\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e$$

其中 $e \approx 2.71828\ldots$ 是自然对数的底数. 这个极限是指数函数 e^x 的一个基本定义. 在已经建立指数函数与对数函数连续性的框架下, 其更一般的形式为

$$\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = e^x \quad (x \in \mathbb{R}).$$

证明 先说明 $\left(1+\frac{1}{n}\right)^n$ 的极限确实存在. 记

$$x_n = \left(1+\frac{1}{n}\right)^n, \qquad y_n = \left(1+\frac{1}{n}\right)^{n+1}.$$

由二项式展开可得

$$x_n = \sum_{k=0}^{n} \frac{1}{k!} \prod_{j=0}^{k-1} \left(1-\frac{j}{n}\right),$$

其中 $k=0$ 时约定空积为 1. 对固定的 k, 乘积中的每个因子随 n 增大而不减, 且 x_{n+1} 比 x_n 多出一个非负项, 因而 $\{x_n\}$ 单调递增. 又 $x_n < y_n$, 而 $\{y_n\}$ 单调递减; 后一结论可由

$$\frac{y_{n+1}}{y_n} = \left(1-\frac{1}{(n+1)^2}\right)^{n+1} \left(1+\frac{1}{n+1}\right) < 1$$

得到, 其中最后一步使用伯努利不等式 $(1-t)^{-(n+1)} > 1+(n+1)t$ $(0<t<1)$. 因此 $\{x_n\}$ 单调递增且以 $y_1 = 4$ 为上界, 必收敛. 其极限定义为自然常数 e.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

对一般实数 x, 当 n 足够大时 $1+x/n>0$. 在采用连续的指数、对数函数后,

$$\ln\left(1+\frac{x}{n}\right)^n = n\ln\left(1+\frac{x}{n}\right)\to x,$$

这里使用基本极限 $\lim_{u\to 0}\dfrac{\ln(1+u)}{u}=1$. 再由指数函数连续性得

$$\left(1+\frac{x}{n}\right)^n\to e^x.$$

□

定理 16.7

$$\lim_{n\to\infty}\sqrt[n]{n}=1$$

证明 **证明:** 令 $x_n=\sqrt[n]{n}-1$. 显然当 $n>1$ 时, $\sqrt[n]{n}>1$, 故 $x_n>0$. 有 $\sqrt[n]{n}=1+x_n$, 两边取 n 次幂, 得到 $n=(1+x_n)^n$.

根据二项式定理, 对 $n\geq 2$:

$$n=(1+x_n)^n=\binom{n}{0}+\binom{n}{1}x_n+\binom{n}{2}x_n^2+\cdots+\binom{n}{n}x_n^n$$

由于所有项均为正, 可只取其中一项放缩:

$$n>\binom{n}{2}x_n^2=\frac{n(n-1)}{2}x_n^2$$

对 $n\geq 2, 1>\frac{n-1}{2}x_n^2$, 整理得

$$x_n^2<\frac{2}{n-1}$$

因此得到 x_n 的界:

$$0<x_n<\sqrt{\frac{2}{n-1}}$$

当 $n\to\infty$ 时, $\lim_{n\to\infty}\sqrt{\frac{2}{n-1}}=0$.

根据夹逼准则,$\lim_{n\to\infty}x_n=0$. 因此,$\lim_{n\to\infty}\sqrt[n]{n}=\lim_{n\to\infty}(1+x_n)=1+0=1$.

□

这个极限刻画了不同增长类型在开 n 次方后的行为.

考虑两个函数: 常数函数 $f(n)=c$ 和线性函数 $g(n)=n$. 线性函数的增长速度显然远快于常数函数. 对两者同时施加 $\sqrt[n]{\cdot}$ 运算:

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 对常数 $c>0$, $\lim_{n\to\infty}\sqrt[n]{c}=1$. 因为 $\sqrt[n]{c}=c^{1/n}$, 当 $n\to\infty$ 时 $1/n\to 0$, 故 $c^{1/n}\to c^0=1$.
- 对线性函数 n, $\lim_{n\to\infty}\sqrt[n]{n}=1$.

尽管 n 的增长速度远超任何常数 c, 但在开 n 次方后, 它们的极限行为完全相同.

这说明开 n 次方根运算的削弱作用极强. 它能将任何多项式速率增长的序列, 如 n, n^2, n^k 的极限都压回到 1. 与指数函数 a^n $(a>1)$ 的爆炸式增长相比, 任何多项式 $P(n)$ 的增长都可忽略; 考察 $\sqrt[n]{P(n)}$ 时, 其极限也总是 1.

这个极限是级数收敛性根值判别法的基础, 用来衡量一个数列增长速率是否快于几何级数.

知乎@余命数

16.4 级数

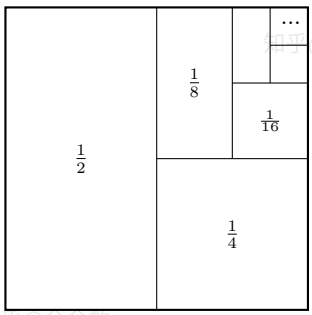


图 16.4　无穷级数 $1/2+1/4+1/8+\cdots=1$ 的几何直观.

阿喀琉斯与乌龟的赛跑

芝诺曾提出一个著名悖论. 阿喀琉斯永远追不上前面爬行的乌龟, 理由是每次阿喀琉斯到达乌龟所在位置时, 乌龟又向前移动了一小段; 待他跑完这段新距离, 乌龟又移动了一点. 这个过程可以无限细分, 阿喀琉斯似乎永远在完成超越.

这个悖论提出的问题是, 包含无穷多个步骤的过程能否在有限时间内完成, 换成更具体的说法, 把 1 米长的绳子依次切成 1/2 米、1/4 米、1/8 米……得到的这一串无穷片段能否重新拼成原来的 1 米.

直觉上, 无穷多个正数相加似乎结果应为无穷大. 但经验(阿喀琉斯确实能追上乌龟, 绳子也确实长 1 米)表明, 在特定条件下, 无穷多个数之和可以是一个有限值.

这种无穷项之和的数学结构称为**级数**. 它是理解连续、逼近与函数的重要工具.

16.4.1 定义与收敛性

无穷次加法无法直接执行. 为了给无穷和赋予数学意义, 需要借助部分和与极限.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

定义 16.10 (级数) 给定一个数列 $\{a_n\}$. 表达式

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \dots$$

称为**无穷级数**, 简称**级数**. 其中 a_n 称为级数的**通项**.
为计算其值, 构造新的数列 $\{S_n\}$, 称为级数的**部分和数列**, 其中

$$S_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n$$

如果部分和数列 $\{S_n\}$ 收敛于实数 S, 即 $\lim_{n\to\infty} S_n = S$, 则称该级数**收敛**, 并称 S 为该级数的**和**. 如果 $\{S_n\}$ 发散, 则称该级数**发散**.

级数的和就是部分和数列的极限. 若前 n 项和随 n 增大而趋于某个固定数值, 这个数值就被当作无穷多项相加的结果.

知乎@余命数

16.4.2 级数收敛的必要条件

直观上, 要把无穷多个数加成一个有限值, 构成级数的数列 $\{a_n\}$ 必须越来越小, 最终趋于零.

上一节已经证明, 若两个数列都有极限, 则它们和与差的极限等于极限的和与差. 据此即证项趋零性.

知乎@余命数

定理 16.8 (项趋零性) 如果级数 $\sum_{n=1}^{\infty} a_n$ 收敛, 那么其通项 a_n 随 n 趋于无穷时的极限必为零, 即

$$\lim_{n\to\infty} a_n = 0$$

知乎@余命

证明 设级数收敛于 S, 即 $\lim_{n\to\infty} S_n = S$. 通项与部分和的关系为 $a_n = S_n - S_{n-1}$ (当 $n \geq 2$). 又因为 $S_n \to S$, 去掉或平移有限项不改变极限, 所以 $S_{n-1} \to S$. 对等式两边取极限, 得

知乎@余命数

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1} = S - S = 0$$

证毕. □

这只是一个**必要条件**. 即使 $\lim_{n\to\infty} a_n = 0$, 级数也未必收敛. 这一点在后续关于调和级数的讨论中将得到证实. 但如果 $\lim_{n\to\infty} a_n \neq 0$, 则可以断定级数发散.

例题 16.18　判断级数 $\sum_{n=1}^{\infty} \frac{n}{n+1}$ 的敛散性.

解　考察通项的极限.

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n}{n+1} = 1$$

由于极限不为 0, 根据必要条件, 该级数发散.

16.4.3　重要的级数模型

16.4.3.1　几何级数(等比级数)

几何级数是最基本且应用最广泛的级数. 考虑等比数列 $\{aq^{n-1}\}$ 构成的级数.

$$\sum_{n=1}^{\infty} aq^{n-1} = a + aq + aq^2 + \ldots \quad (a \neq 0)$$

其部分和 S_n 由等比数列求和公式给出.

$$S_n = \begin{cases} \frac{a(1-q^n)}{1-q} & q \neq 1 \\ na & q = 1 \end{cases}$$

分析 S_n 的极限行为.

- 当 $|q| < 1$ 时, $\lim_{n\to\infty} q^n = 0$, 故 $\lim_{n\to\infty} S_n = \frac{a}{1-q}$. 级数收敛.
- 当 $|q| \geq 1$ 时, 通项 aq^{n-1} 不趋于 0, 故级数发散; 对应的部分和可能无界, 也可能振荡而不收敛.

综合两种情形, 几何级数 $\sum_{n=1}^{\infty} aq^{n-1}$ 收敛当且仅当 $|q| < 1$, 其和为 $\frac{a}{1-q}$.

例题 16.19　计算循环小数 $0.999\ldots$ 的值, 并判断它与 1 的关系.

解　首先明确 $0.999\ldots$ 的含义. 它表示一个极限过程. 依次取

$$s_1 = 0.9, \quad s_2 = 0.99, \quad s_3 = 0.999, \quad \ldots$$

一般地, 第 n 个近似写作

$$s_n = 0.\underbrace{99\ldots9}_{n\text{个}9}.$$

把数列 $\{s_n\}$ 的极限(若存在)记作 $0.999\ldots$:

$$0.999\ldots := \lim_{n\to\infty} s_n.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

接下来把 s_n 写成有限级数. 每个 s_n 可表示为

$$s_n = \frac{9}{10} + \frac{9}{10^2} + \cdots + \frac{9}{10^n} = \sum_{k=1}^{n} \frac{9}{10^k}.$$

这是等比级数的前 n 项和, 首项为 $a = \dfrac{9}{10}$, 公比为 $q = \dfrac{1}{10}$. 由等比数列求和公式

$$\sum_{k=1}^{n} aq^{k-1} = a\,\frac{1-q^n}{1-q}$$

可得

$$s_n = \frac{9}{10} \cdot \frac{1-(\frac{1}{10})^n}{1-\frac{1}{10}} = \frac{9}{10} \cdot \frac{1-10^{-n}}{\frac{9}{10}} = 1 - 10^{-n}.$$

于是

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} (1-10^{-n}) = 1 - \lim_{n\to\infty} 10^{-n} = 1 - 0 = 1.$$

根据 $0.999\ldots$ 的定义, 有

$$0.999\cdots = \lim_{n\to\infty} s_n = 1.$$

从差的角度再看一次. 由 $s_n = 1 - 10^{-n}$ 可得

$$1 - s_n = 10^{-n}.$$

对两边取极限, 得

$$1 - 0.999\cdots = 1 - \lim_{n\to\infty} s_n = \lim_{n\to\infty} (1 - s_n) = \lim_{n\to\infty} 10^{-n} = 0.$$

两数之差为 0, 说明 $0.999\ldots$ 与 1 表示同一个实数, 它们之间没有间隙.

上面的计算还给出了另一个常用写法的依据, 得

$$0.00999\cdots = \frac{9}{10^3} + \frac{9}{10^4} + \cdots = \sum_{k=3}^{\infty} \frac{9}{10^k} = \frac{9/10^3}{1-\frac{1}{10}} = \frac{1}{10^2} = 0.01.$$

因此

$$0.5 = 0.4999\ldots, \quad 0.25 = 0.24999\ldots, \quad \text{等等.}$$

任何有限小数的十进制展开都有一种等价表示. 把最后一位减一, 再接无穷多个 9; $0.999\cdots = 1$ 是其中最简单的一例.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

16.4.3.2 调和级数

考察通项为 $a_n = \frac{1}{n}$ 的级数.

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

虽然 $\lim_{n\to\infty} \frac{1}{n} = 0$, 满足收敛的必要条件, 但该级数**发散**, 正印证了通项趋于零并非充分条件.

证明 (Oresme 的证明) 将级数的项分组放缩.

$$\begin{aligned} S &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots \\ &> 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots \\ &= 1 + \frac{1}{2} + \frac{2}{4} + \frac{4}{8} + \dots \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots \end{aligned}$$

通过把 2^k 项的分母放大为 2^{k+1} 的倒数, 构造出了无穷多个 $1/2$. 这个和趋向于无穷大. 因此, 调和级数发散. □

16.4.3.3 裂项级数

若级数的通项 a_n 可写成 $a_n = b_n - b_{n+1}$, 则部分和 S_n 中会出现中间项抵消. 这种级数称为裂项级数.

$$S_n = (b_1 - b_2) + (b_2 - b_3) + \dots + (b_n - b_{n+1}) = b_1 - b_{n+1}$$

此时级数的敛散性完全取决于数列 $\{b_n\}$ 的极限.

$$\sum_{n=1}^{\infty} (b_n - b_{n+1}) = b_1 - \lim_{n\to\infty} b_{n+1}$$

例题 16.20 计算级数 $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ 的和.

解 利用部分分式分解.

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$$

部分和为.

$$S_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \dots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

取极限.

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(1 - \frac{1}{n+1}\right) = 1$$

16.4.4 绝对收敛与条件收敛

项的符号发生变化的级数需要更细致的分析. 考虑交错调和级数.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

调和级数发散, 但由于正负项相互抵消, 这个级数收敛(收敛于 $\ln 2$).

定义 16.11 对级数 $\sum a_n$:

- 若各项绝对值构成的级数 $\sum |a_n|$ 收敛, 则称原级数**绝对收敛**.
- 若原级数 $\sum a_n$ 收敛, 但 $\sum |a_n|$ 发散, 则称原级数**条件收敛**.

绝对收敛的级数具有很好的性质, 例如可以任意交换求和顺序而不改变和的值. 条件收敛级数则不然. 黎曼重排定理指出, 对条件收敛级数, 可以通过重排项的顺序使其收敛于任意实数, 甚至发散. 这说明处理无穷级数时, 有限运算的直觉, 如交换律需要经过严格检验.

16.4.5 典型级数与函数展开

建立级数收敛概念后, 需要几个标准参照系. 判断陌生级数的敛散性, 常用方法是把它与已知典型级数比较. 同时, 级数也可以表示函数, 从而把研究对象从数字之和提升到函数.

16.4.5.1 几何级数

几何级数(等比级数)是计算中最友好的级数, 也是衡量其他级数收敛速度的参照.

考虑级数 $\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \dots$. 前文已知, 当且仅当 $|x| < 1$ 时, 该级数收敛于 $\frac{1}{1-x}$.

这个求和公式说明了一个基本代数事实. 函数 $f(x) = \frac{1}{1-x}$ 可以展开为一个无穷多项式.

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \quad (|x| < 1) \tag{16-3}$$

这个等式是现代分析中幂级数理论的起点. 在收敛区间内, 复杂的函数可以被简单的幂函数之和替代.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

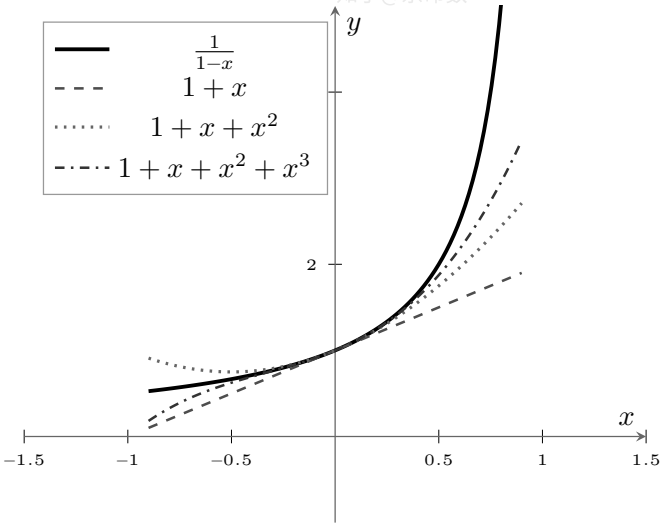


图 16.5　函数 $y=\frac{1}{1-x}$ 与其部分和多项式的逼近情况. 注意在 $|x|<1$ 区间内的拟合效果.

16.4.5.2 调和级数的推广

调和级数 $\sum \frac{1}{n}$ 的发散性划定了敛散性的边界. 若通项趋于 0 的速度比 $1/n$ 更慢, 级数必发散; 若快得多, 级数可能收敛. 为量化这种快慢, 引入 p-级数.

定义 16.12 (p-级数) 形如以下的级数称为 **p-级数**.

$$\sum_{n=1}^{\infty}\frac{1}{n^p}=1+\frac{1}{2^p}+\frac{1}{3^p}+\dots$$

其中 p 为实常数.

其敛散性完全由指数 p 决定.

- 当 $p\le 1$ 时, 级数**发散**.($p=1$ 即为调和级数)
- 当 $p>1$ 时, 级数**收敛**.

例如, $\sum \frac{1}{n^2}$ 收敛(欧拉证明了其和为 $\pi^2/6$), 而 $\sum \frac{1}{\sqrt{n}}$ ($p=1/2$) 发散.p-级数常被用作比较判别法中的参照物.

16.4.5.3 交错级数

数学中常出现正负项交替出现的级数, 称为**交错级数**. 其一般形式为.

$$\sum_{n=1}^{\infty}(-1)^{n-1}a_n=a_1-a_2+a_3-a_4+\dots\quad(a_n>0)$$

对这类级数, 莱布尼茨给出了一个简洁实用的判别法.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 16.9 (莱布尼茨判别法) 如果交错级数 $\sum_{n=1}^{\infty}(-1)^{n-1}a_n$ 中 $a_n>0$, 且满足以下两个条件.

1. **单调递减**, 即 $a_{n+1}\le a_n$ 对所有 n 成立;
2. **趋于零**, 即 $\lim_{n\to\infty}a_n=0$.

那么该级数**收敛**.

证明 记部分和

$$S_n=\sum_{k=1}^{n}(-1)^{k-1}a_k.$$

将奇数项、偶数项的部分和分开.

$$S_{2n}=a_1-a_2+\cdots+a_{2n-1}-a_{2n},\qquad S_{2n+1}=a_1-a_2+\cdots-a_{2n}+a_{2n+1}.$$

偶数部分和单调递增.

$$\begin{aligned}S_{2n+2}-S_{2n}&=(a_1-\cdots-a_{2n}+a_{2n+1}-a_{2n+2})\\&\quad-(a_1-\cdots-a_{2n})\\&=a_{2n+1}-a_{2n+2}\ge 0,\end{aligned}$$

因为 $\{a_n\}$ 单调递减, 所以 $a_{2n+1}\ge a_{2n+2}$. 因此 $\{S_{2n}\}$ 单调递增.

奇数部分和单调递减.

$$S_{2n+3}-S_{2n+1}=(S_{2n+1}-a_{2n+2}+a_{2n+3})-S_{2n+1}=-a_{2n+2}+a_{2n+3}\le 0,$$

同样由于 $a_{2n+2}\ge a_{2n+3}$, 故 $\{S_{2n+1}\}$ 单调递减.

奇偶部分和互相夹逼. 一方面,

$$S_{2n+1}=S_{2n}+a_{2n+1}>S_{2n},$$

所以对所有 n, 都有 $S_{2n}\le S_{2n+1}$, 即

$$S_2\le S_4\le\cdots\le S_{2n}\le S_{2n+1}\le\cdots\le S_3\le S_1.$$

于是 $\{S_{2n}\}$ 有上界 S_1, $\{S_{2n+1}\}$ 有下界 S_2. 结合第 1、2 点, 得到.

$$\{S_{2n}\}\text{ 单调递增有上界 }\Rightarrow\ \{S_{2n}\}\text{ 收敛},$$

$$\{S_{2n+1}\}\text{ 单调递减有下界 }\Rightarrow\ \{S_{2n+1}\}\text{ 收敛}.$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

设

$$\lim_{n\to\infty} S_{2n} = L_1, \qquad \lim_{n\to\infty} S_{2n+1} = L_2.$$

利用 $a_n \to 0$ 说明两极限相同：注意到

$$S_{2n+1} - S_{2n} = a_{2n+1} \xrightarrow[n\to\infty]{} 0.$$

两边取极限可得

$$L_2 - L_1 = \lim_{n\to\infty}(S_{2n+1} - S_{2n}) = \lim_{n\to\infty} a_{2n+1} = 0,$$

所以 $L_1 = L_2 =: L$.

这说明奇数、偶数部分和收敛到同一极限 L，于是整体部分和 S_n 也收敛到 L，即原交错级数收敛.

直观理解. 每加一项就向左右交替摆动，由于加上的正负项绝对值越来越小，部分和在一段越来越短的区间内震荡；区间长度约为 a_n，而 $a_n \to 0$，所以震荡最终集中到同一个点，这个点就是级数的和. □

例题 16.21　判断级数 $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ 的敛散性.

解　这是交错调和级数. 令 $a_n = \frac{1}{n}$. 1. 显然 $a_{n+1} = \frac{1}{n+1} < \frac{1}{n} = a_n$，满足单调递减. 2. $\lim_{n\to\infty}\frac{1}{n} = 0$，满足极限为零. 根据莱布尼茨判别法，该级数收敛.

例题 16.22　判断级数 $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{\sqrt{n}}$ 的敛散性，并说明是绝对收敛还是条件收敛.

解　令 $a_n = \dfrac{1}{\sqrt{n}}$.

1. 单调性. 对 $n \geq 1$，

$$a_{n+1} = \frac{1}{\sqrt{n+1}} < \frac{1}{\sqrt{n}} = a_n,$$

所以 $\{a_n\}$ 单调递减；

2. 极限. $\displaystyle\lim_{n\to\infty}\frac{1}{\sqrt{n}} = 0$.

满足莱布尼茨判别法的两个条件，因此

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{\sqrt{n}}$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

收敛.

再看绝对收敛性.

$$\sum_{n=1}^{\infty}\left|(-1)^{n-1}\frac{1}{\sqrt{n}}\right|=\sum_{n=1}^{\infty}\frac{1}{\sqrt{n}}$$

是 p 级数 ($p=\frac{1}{2}\leq 1$), 发散.

因此原级数**条件收敛**, 绝对值级数发散.

例题 16.23 判断级数 $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\dfrac{1}{n(n+1)}$ 的敛散性, 并说明是绝对收敛还是条件收敛.

知乎@余命数

解 令 $a_n=\dfrac{1}{n(n+1)}$.

1. 单调性. 对 $n\geq 1$,

$$a_{n+1}=\frac{1}{(n+1)(n+2)}.$$

知乎@余命数

比较 a_{n+1} 与 a_n:

$$\begin{aligned}a_{n+1}<a_n &\iff \frac{1}{(n+1)(n+2)}<\frac{1}{n(n+1)}\\ &\iff n(n+1)<(n+1)(n+2),\end{aligned}$$

知乎@余命数

而 $n(n+1)<(n+1)(n+2)$ 显然成立, 所以 $a_{n+1}<a_n$, 序列单调递减;

2. 极限. $\lim\limits_{n\to\infty}\dfrac{1}{n(n+1)}=0$.

故由莱布尼茨判别法,

知乎@余命数

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n(n+1)}$$

知乎@余命数

收敛.

再看绝对收敛性.

$$\sum_{n=1}^{\infty}\left|(-1)^{n-1}\frac{1}{n(n+1)}\right|=\sum_{n=1}^{\infty}\frac{1}{n(n+1)}.$$

注意

$$\frac{1}{n(n+1)}=\frac{1}{n}-\frac{1}{n+1},$$

因此

$$\sum_{n=1}^{N}\frac{1}{n(n+1)}=\sum_{n=1}^{N}\left(\frac{1}{n}-\frac{1}{n+1}\right)=1-\frac{1}{N+1}\xrightarrow{N\to\infty}1.$$

所以 $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n(n+1)}$ 收敛, 从而原级数**绝对收敛**.

例题 16.24　判断级数 $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\sin\frac{1}{n}$ 的敛散性, 并说明是绝对收敛还是条件收敛.

解　令 $a_n=\sin\frac{1}{n}$.

1. 单调性. 当 n 增大时, $\dfrac{1}{n}$ 单调递减; 而在区间 $(0,1]$ 上, $\sin x$ 单调递增 (因为 $\cos x>0$), 因此复合函数

$$a_n=\sin\left(\frac{1}{n}\right)$$

随 n 增大而单调递减, 即 $a_{n+1}\le a_n$;

2. 极限. $\displaystyle\lim_{n\to\infty}\sin\frac{1}{n}=\sin 0=0$.

满足莱布尼茨判别法条件, 因此

$$\sum_{n=1}^{\infty}(-1)^{n-1}\sin\frac{1}{n}$$

收敛.

考察绝对收敛性.

$$\sum_{n=1}^{\infty}\left|\sin\frac{1}{n}\right|.$$

对大 n, $\dfrac{1}{n}$ 很小, 利用极限

$$\lim_{x\to0}\frac{\sin x}{x}=1\Rightarrow\lim_{n\to\infty}\frac{\sin\frac{1}{n}}{\frac{1}{n}}=1.$$

于是由极限比较判别法,

$$\sum_{n=1}^{\infty}\left|\sin\frac{1}{n}\right|\quad\text{与}\quad\sum_{n=1}^{\infty}\frac{1}{n}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

有相同的敛散性, 而调和级数 $\sum \frac{1}{n}$ 发散, 因此

$$\sum_{n=1}^{\infty} \left|\sin \frac{1}{n}\right|$$

发散.

故原级数**条件收敛**, 绝对值级数发散.

16.4.5.4 作为函数的级数

把几何级数 $\sum x^n$ 的思想推广. 若级数中的常数项替换为变量 x 的函数, 级数本身就定义了一个函数. 最重要的一类是**幂级数**.

定义 16.13 (幂级数) 形如

$$\sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + \dots$$

的级数称为关于 x 的幂级数, 其中 c_n 为常数系数.

幂级数可看作次数无限的多项式. 对一个幂级数, 存在非负实数 R, 称为**收敛半径**, 使得

- 当 $|x| < R$ 时, 级数绝对收敛;
- 当 $|x| > R$ 时, 级数发散.

许多常见初等函数都可以表示为幂级数, 称为函数的**泰勒展开**. 这种表示把复杂的函数运算, 如求导、积分转化为简单的多项式系数运算.

例题 16.25 指数函数 e^x 的幂级数展开如下.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

该级数对所有实数 x 收敛, 即 $R = +\infty$.

例题 16.26 正弦函数 $\sin x$ 的幂级数展开如下.

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

这也是一个交错级数, 且对所有实数 x 收敛.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

借助这些展开式, 计算机可用有限的加减乘除高精度计算 e^2 或 $\sin(1)$ 等数值. 这是级数理论在现代计算中的主要用途之一.

16.5 切线问题

有了极限的初步概念, 我们就可以严格定义曲线在某一点的切线, 这是最初激发微积分的几何问题.

16.5.1 古典定义

古希腊几何学家对圆、椭圆等二次曲线的切线已有基本认识. 他们所依赖的要点是, 若一条直线与曲线在某点附近**仅有一个交点**, 则此直线为该曲线在该点的切线. 这种基于交点个数的判别法对凸性良好的简单曲线直观有效.

然而, 当我们把定义用于更复杂的函数曲线时, 它暴露出根本缺陷. 我们考察函数 $y = x^3$ 在原点 $(0,0)$ 处的切线. 从直观上看, 水平的 x 轴是最佳候选. 但用古典定义检验它, 会立即陷入矛盾.

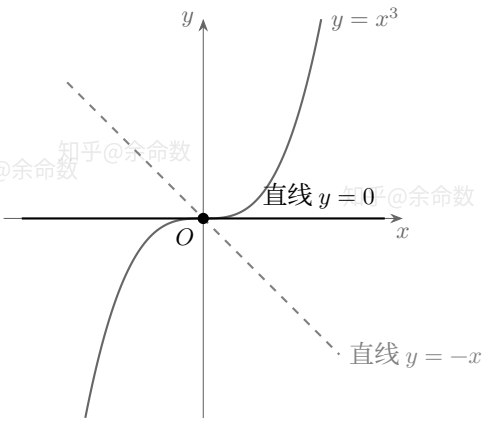


图 16.6　曲线 $y = x^3$ 在原点的切线问题

直线 $y = 0$ 即 x 轴, 与曲线 $y = x^3$ 在原点接触. 但当 $x > 0$ 时 $x^3 > 0$, 曲线在直线上方; 当 $x < 0$ 时 $x^3 < 0$, 曲线在直线下方. 于是该直线在接触点处**穿过**了曲线, 这与切线应支撑于曲线一侧的直观感受矛盾.

更严重的是唯一性问题. 我们考察任意一条斜率为负的直线 $y = mx$ $(m < 0)$. 方程 $x^3 = mx$ 仅有唯一实数解 $x = 0$. 若依据唯一交点准则, 则任何一条过原点的斜率为负的直线都是切线, 这显然是荒谬的.

这些矛盾说明, 依赖曲线**全局**交点个数的静态定义不够用. 我们需要一种能精确捕捉曲线在某点附近**局部趋势**的工具, 而极限理论正好提供这一工具.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.5.2 割线的极限观点

我们要求曲线在点 $P(x_0, f(x_0))$ 处的切线. 确定一条直线需要一个点与一个斜率. 点 P 已知, 剩下的问题是求斜率.

斜率定义 $\frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$ 需要两个不同点, 仅凭点 P 无法直接算出任何斜率. 因此我们在曲线点 P 的邻近另取一个不同点 $Q(x, f(x))$. 连接 P 与 Q 的直线称为曲线的一条**割线**. 其斜率 k_{PQ} 为

$$k_{PQ} = \frac{f(x) - f(x_0)}{x - x_0}$$

此值量化了函数在 $[x_0, x]$ 这段区间内的**平均变化率**.

这条割线是对切线的一个近似. 点 Q 越接近点 P, 割线 PQ 对点 P 处切线的近似就越精确. 我们用**极限** 严格表述这一逼近过程. 让点 Q 沿曲线向点 P 移动, 即横坐标 x 无限趋近于 x_0.

从几何上看, 割线 PQ 在此过程中不断旋转. 当 Q 无限接近 P 时, 割线稳定于一个确定的极限位置. 这个极限位置的直线就是曲线在点 P 处的**切线**.

几何上的极限位置对应着代数上割线斜率的极限值. 因此, 点 P 处切线的斜率就是割线斜率 k_{PQ} 当 $x \to x_0$ 时的极限.

16.5.3 切线斜率的严格定义

我们将上述直观思想写成数学语言. 曲线在点 P 处的切线斜率, 就是当 $x \to x_0$ 时割线斜率 k_{PQ} 的极限.

定义 16.14 (切线斜率) 曲线 $y = f(x)$ 在点 $P(x_0, f(x_0))$ 处的**切线斜率** $k_{\tan}$ 定义为

$$k_{\tan} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{16-4}$$

若令 $x = x_0 + \Delta x$, 则 $x \to x_0$ 等价于增量 $\Delta x \to 0$. 于是定义式可改写为

$$k_{\tan} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}. \tag{16-5}$$

若此极限存在, 则称曲线在该点处的切线存在.

基于极限的定义解决了古典定义的困境, 并给 $\frac{0}{0}$ 型不定式赋予了几何意义. 它考察的是当 Δx 无限趋近于 0 时整个比值的极限.

回到 $f(x) = x^3$ 在原点的例子. 连接原点 $(0, 0)$ 与曲线上另一点 (x, x^3) 的割线斜率为

$$k_{\text{割线}} = \frac{x^3 - 0}{x - 0} = x^2 \quad (x \neq 0)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

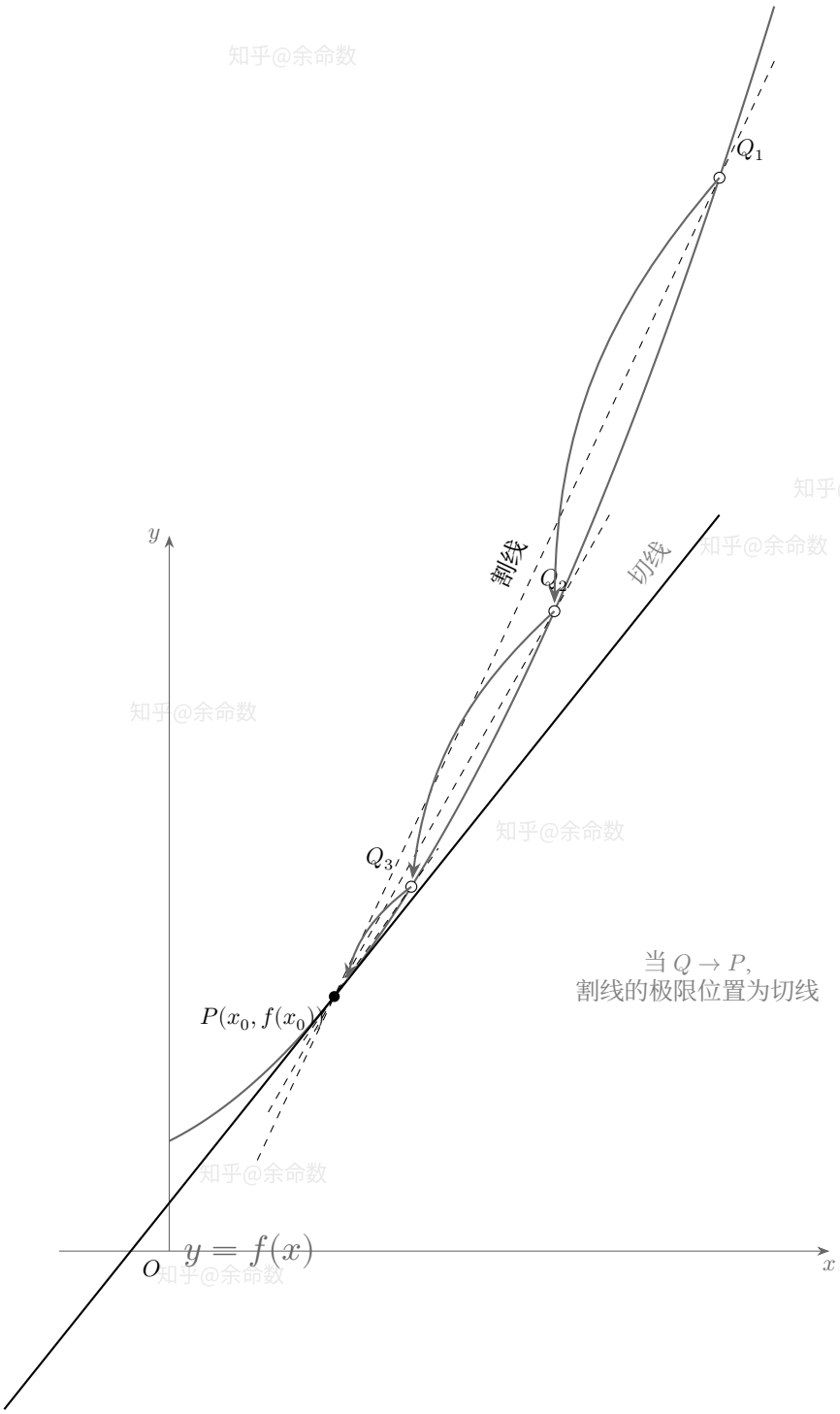


图 16.7　当点 Q 沿曲线趋近于点 P 时, 割线 PQ 的极限位置即为在点 P 处的切线.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们考察当 $x \to 0$ 时此斜率的极限

$$k_{\tan} = \lim_{x \to 0} x^2 = 0$$

这个极限唯一且确定. 无论动点从左侧还是右侧趋近原点, 割线斜率都趋向于 0. 因此原点处切线的斜率为 0, 对应切线方程为 $y - 0 = 0(x - 0)$, 即 $y = 0$. 定义由此解决了之前的矛盾.

极限提供了一个局部放大的视角. 在无限小的尺度下, 任何光滑曲线都近似于一条直线, 这条直线就是它的切线. 切线是曲线在该点附近的**最佳线性近似**.

例题 16.27 求曲线 $f(x) = x^2$ 在点 $P(1, 1)$ 处的切线斜率, 并写出切线方程.

解 所求切线斜率, 是连接定点 $P(1, 1)$ 与其邻近动点 $Q(1 + \Delta x, (1 + \Delta x)^2)$ 的割线斜率在 $\Delta x \to 0$ 时的极限.

割线 PQ 的斜率为 $k_{PQ} = \dfrac{(1 + \Delta x)^2 - 1}{\Delta x}$. 若直接令 $\Delta x = 0$ 将得到 $\frac{0}{0}$ 的无意义形式, 因此我们在取极限前先作代数化简. 将分子展开并在 $\Delta x \neq 0$ 时提取公因式约分, 得

$$k_{PQ} = \frac{1 + 2\Delta x + (\Delta x)^2 - 1}{\Delta x} = \frac{2\Delta x + (\Delta x)^2}{\Delta x} = \frac{\Delta x(2 + \Delta x)}{\Delta x} = 2 + \Delta x \quad (\Delta x \neq 0).$$

这表明只要 Q 点不与 P 点重合, 割线斜率就等于 $2 + \Delta x$. 令 $\Delta x \to 0$, 得切线斜率

$$k_{\tan} = \lim_{\Delta x \to 0} (2 + \Delta x) = 2,$$

即曲线 $y = x^2$ 在点 $(1, 1)$ 处的切线斜率为 2.

由点斜式方程, 所求切线方程为 $y - 1 = 2(x - 1)$, 即 $y = 2x - 1$.

切线问题说明, $\lim\limits_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$ 这一形式的极限既解决了几何问题, 又描述瞬时变化. 我们给它取名为**导数**.

16.5.4 本节习题

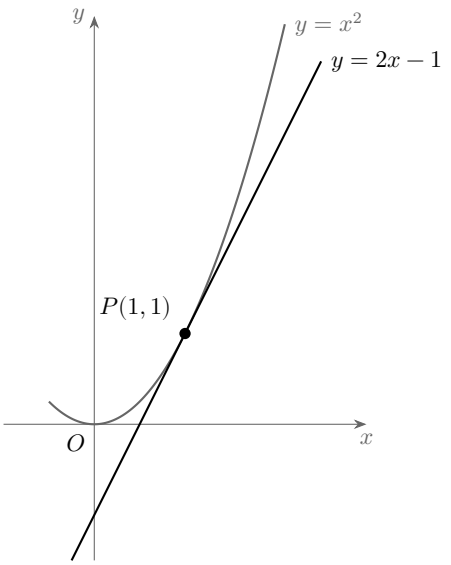
16.9. 考察函数 $g(x) = \sqrt{x}$. 运用极限的定义, 证明其在任意点 $P(x_0, \sqrt{x_0})$ $(x_0 > 0)$ 处的切线斜率为 $\dfrac{1}{2\sqrt{x_0}}$.

16.10. 求所有通过点 $A(1, -3)$ 且与抛物线 $y = x^2$ 相切的直线方程.

16.11. 考虑函数 $h(x) = |x^2 - 4|$. 判断该函数在点 $x = 2$ 处是否存在切线. 若存在, 求出其方程; 若不存在, 请阐述其几何与代数原因.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数



16.12. 确定参数 a 的值, 使得抛物线 $y=ax^2$ 与对数曲线 $y=\ln x$ 恰好相切.

16.13. 确定常数 b 的值, 使得直线 $y=x+b$ 与对数曲线 $y=\ln x$ 相切.

16.14. 求抛物线 $y=x^2-1$ 上的点, 使得该点处的法线通过坐标原点 $(0,0)$.

16.15. 证明对于任意实数 x, 不等式 $e^x \geq x+1$ 恒成立.

16.16. 考虑双曲线 $y=1/x$ 在第一象限的分支. 证明由该曲线上任意一点的切线与两坐标轴所围成的三角形的面积是一个与切点位置无关的常数.

知乎@余命数

16.5.5 高考真题选讲

以下高考真题围绕切线的构造与判别展开, 涵盖“过点”与“在点”的区分、公切线求解以及切线倾斜角的取值范围.

例题 16.28 (2004 年重庆卷文科数学第 15 题) 已知曲线 $y=\dfrac{1}{3}x^3+\dfrac{4}{3}$, 求过点 $P(2,4)$ 的切线方程.

解 点 $P(2,4)$ 在曲线上, 但“过点”的切线未必以 P 为切点. 设切点为 $\left(x_0, \frac{1}{3}x_0^3+\frac{4}{3}\right)$, 则切线斜率为 $k=y'(x_0)=x_0^2$, 切线方程为

$$y-\left(\frac{1}{3}x_0^3+\frac{4}{3}\right)=x_0^2(x-x_0)$$

将 $P(2,4)$ 代入, 得

$$4-\frac{1}{3}x_0^3-\frac{4}{3}=x_0^2(2-x_0)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

整理得 $x_0^3-3x_0^2+4=0$, 因式分解为 $(x_0-2)^2(x_0+1)=0$, 解得 $x_0=2$ 或 $x_0=-1$.

当 $x_0=2$ 时, 斜率 $k=4$, 切线方程为 $y-4=4(x-2)$, 即 $4x-y-4=0$. 当 $x_0=-1$ 时, 斜率 $k=1$, 切点为 $(-1,1)$, 切线方程为 $y-1=x+1$, 即 $y=x+2$.

故过点 $P(2,4)$ 的切线有两条: $4x-y-4=0$ 与 $y=x+2$.

例题 16.29 (2016 年全国 □ 卷理科数学第 16 题) 若直线 $y=kx+b$ 是曲线 $y=\ln x+2$ 的切线, 也是曲线 $y=\ln(x+1)$ 的切线, 求 b 的值.

解 设直线与两曲线的切点分别为 $(x_1,\ln x_1+2)$ 和 $(x_2,\ln(x_2+1))$. 由导数的几何意义, 切线斜率同时满足

$$k=\frac{1}{x_1}=\frac{1}{x_2+1}$$

故 $x_1=x_2+1$. 切点在直线 $y=kx+b$ 上, 给出

$$kx_1+b=\ln x_1+2$$

$$kx_2+b=\ln(x_2+1)$$

两式相减, 并利用 $x_1=x_2+1$ 与 $\ln x_1=\ln(x_2+1)$, 得

$$k(x_1-x_2)=2$$

即 $k=2$. 代回得 $x_1=\frac{1}{2}$, $x_2=-\frac{1}{2}$.

由 $kx_1+b=\ln x_1+2$ 得 $b=\ln\frac{1}{2}+2-2\cdot\frac{1}{2}=1-\ln 2$.

例题 16.30 (2021 年全国乙卷文科数学第 21 题) 已知函数 $f(x)=x^3-x^2+ax+1$.

1. 讨论 $f(x)$ 的单调性;
2. 求曲线 $y=f(x)$ 过坐标原点的切线与曲线 $y=f(x)$ 的公共点的坐标.

解 **(1) 单调性分析** $f'(x)=3x^2-2x+a$, 判别式 $\Delta=4-12a$.

当 $\Delta\le 0$ 即 $a\ge\frac{1}{3}$ 时, $f'(x)\ge 0$ 恒成立, $f(x)$ 在 $\mathbb{R}$ 上单调递增.

当 $\Delta>0$ 即 $a<\frac{1}{3}$ 时, 令 $f'(x)=0$ 解得 $x_1=\frac{1-\sqrt{1-3a}}{3}$, $x_2=\frac{1+\sqrt{1-3a}}{3}$. $f(x)$ 在 $(-\infty,x_1)$ 和 $(x_2,+\infty)$ 上单调递增, 在 (x_1,x_2) 上单调递减.

(2) 过原点的切线设切点为 $(x_0,f(x_0))$, 切线斜率 $f'(x_0)=3x_0^2-2x_0+a$. 切线方程为

$$y-f(x_0)=f'(x_0)(x-x_0)$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

切线过原点, 代入 $(0,0)$ 得 $f(x_0)=x_0f'(x_0)$, 即

$$x_0^3-x_0^2+ax_0+1=x_0(3x_0^2-2x_0+a)$$

整理得 $2x_0^3-x_0^2-1=0$, 因式分解为 $(x_0-1)(2x_0^2+x_0+1)=0$. 二次因子判别式 $1-8=-7<0$, 故唯一实根 $x_0=1$.

切线斜率 $f'(1)=a+1$, 切线方程为 $y=(a+1)x$.

求切线与曲线的交点, 令 $f(x)=(a+1)x$, 即 $x^3-x^2-x+1=0$, 因式分解为 $(x-1)^2(x+1)=0$, 解得 $x=1$ 或 $x=-1$.

公共点坐标为 $(1,a+1)$ 与 $(-1,-a-1)$.

例题 16.31 (2010 年辽宁卷理科数学第 10 题) 已知点 P 在曲线 $y=\dfrac{4}{e^x+1}$ 上, α 为曲线在点 P 处的切线的倾斜角, 则 α 的取值范围是

A. $\left[0,\dfrac{\pi}{4}\right)$

B. $\left[\dfrac{\pi}{4},\dfrac{\pi}{2}\right)$

C. $\left(\dfrac{\pi}{2},\dfrac{3\pi}{4}\right]$

D. $\left[\dfrac{3\pi}{4},\pi\right)$

解 求导得

$$y'=-\frac{4e^x}{(e^x+1)^2}=-\frac{4}{e^x+e^{-x}+2}$$

由基本不等式 $e^x+e^{-x}\geq 2$, 故分母 $e^x+e^{-x}+2\geq 4$, 从而 $y'\in[-1,0)$.

切线倾斜角 $\alpha\in[0,\pi)$, 且 $\tan\alpha=y'\in[-1,0)$, 因此 $\alpha\in\left[\frac{3\pi}{4},\pi\right)$.

故选 D.

16.6 函数的连续性

我们前面研究的是, 当 x 趋近于 x_0 时, $f(x)$ 的极限. 我们刻意把极限值 $\lim\limits_{x\to x_0}f(x)$ 与函数在该点的取值 $f(x_0)$ 当作两个独立的对象. 它们可能相等, 也可能不相等, 甚至 $f(x_0)$ 可能没有定义.

然而, 许多常见函数满足另一个关系. 例如, 对 $f(x)=x^2$, 我们有 $\lim\limits_{x\to 2}x^2=4$, 而 $f(2)=2^2=4$. 这说明极限值恰好等于函数值. 从几何上看, 函数图像在该点连续不断, 没有孔洞、断裂或跳跃.

连续不断是一个几何直观. 下面我们用极限给出它的严格定义.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.6.1 连续性的定义

定义 16.15 (函数在一点的连续性) 设函数 $y=f(x)$ 在点 x_0 及其某个邻域内有定义. 如果 $\lim\limits_{x\to x_0} f(x)$ 存在, 并且等于 $f(x_0)$, 那么我们就称 $f(x)$ 在点 x_0 处**连续**.

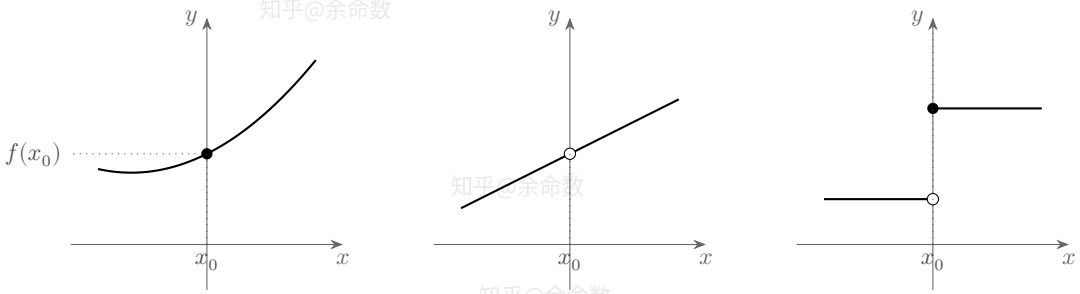
这一定义要求三个条件同时满足.

1. 函数在点 x_0 处有定义, 即 $f(x_0)$ 存在.
2. 函数在点 x_0 处的极限存在, 即 $\lim\limits_{x\to x_0} f(x)$ 存在.
3. 极限值等于函数值, 即 $\lim\limits_{x\to x_0} f(x)=f(x_0)$.

如果上述三个条件中任何一个不满足, 我们就称函数 $f(x)$ 在点 x_0 处**不连续**或**间断**.

注 16.3 (连续的关键)
若函数在某点连续, 则极限过程的终点恰好是函数在该点的取值. 这件事可写成

$$\lim_{x\to x_0} f(x)=f(x_0).$$



(a) 在 x_0 处连续　　(b) 可去间断点　　(c) 跳跃间断点

图 16.8 连续与间断的几何直观

图 (a) 展示连续函数的情形. 当 x 从两侧逼近 x_0 时, 函数值 $f(x)$ 逼近的高度正好等于 $f(x_0)$. 图 (b) 展示一种间断. 虽然 $\lim\limits_{x\to x_0} f(x)$ 存在, 但 $f(x_0)$ 在该点无定义, 或取了别的值, 于是形成一个孔洞. 这种间断称为**可去间断点**. 图 (c) 展示另一种间断. 左极限与右极限虽然都存在, 但二者不相等, 于是函数图像在此处发生跳跃. 这种间断称为**跳跃间断点**.

16.6.1.0.1 区间上的连续性 我们把点的连续性推广到区间上.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 若 $f(x)$ 在开区间 (a,b) 内的**每一点**都连续, 则称 $f(x)$ 在**开区间** (a,b) **上连续**.
- 若 $f(x)$ 在开区间 (a,b) 上连续, 并且在左端点 a 处**右连续** (即 $\lim\limits_{x\to a^+} f(x) = f(a)$), 在右端点 b 处**左连续**, 即 $\lim\limits_{x\to b^-} f(x) = f(b)$, 则称 $f(x)$ 在**闭区间** $[a,b]$ **上连续**.

常见连续函数

所有**基本初等函数**在其**定义域**内的每一个区间上都连续. 它们包括

- 幂函数 $y = x^a$
- 指数函数 $y = a^x$
- 对数函数 $y = \log_a x$
- 三角函数 $y = \sin x, y = \cos x, y = \tan x, \ldots$
- 常数函数 $y = c$

此外, 由这些基本初等函数经过有限次四则运算和复合运算所构成的函数, 在其定义域内也连续.

我们下一节将直接利用这个结论. 若函数 $f(x)$ 在点 x_0 连续, 那么根据定义, 求 $\lim\limits_{x\to x_0} f(x)$ 就等价于计算函数值 $f(x_0)$.

16.7 极限的运算

知乎@余命数

来龙去脉

我们已对极限有一个直观认识. 极限描述的是函数值在一点附近的归宿. 当遇到由多个简单函数通过加、减、乘、除组合而成的复杂函数时, 就需要一套求极限的运算法则. 以 $\lim\limits_{x\to 1}(x^2+3x-1)$ 为例. 一个自然的想法是, 当 x 趋近于 1 时, x^2 趋近于 $1^2=1$, $3x$ 趋近于 $3\cdot 1=3$, 于是整个表达式的极限就是这些部分极限的简单加减, 即 $1+3-1=3$.

这个猜想成立. 极限的运算性质允许我们把一个复杂函数的极限问题**分解**为若干个简单函数极限的四则运算问题, 从而简化求解.

16.7.1 极限的四则运算法则

引理 16.2 (三角不等式) 对任意实数 a 和 b, 以下不等式恒成立.

$$|a+b| \le |a| + |b| \tag{16-6}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 设 x 为任意实数, 则 $-|x| \le x \le |x|$. 令 $x = a$ 和 $x = b$, 分别得到

$$-|a| \le a \le |a|,$$

$$-|b| \le b \le |b|.$$

将这两式相加, 得到 $-(|a| + |b|) \le a + b \le |a| + |b|$. 不等式 $-k \le y \le k$ 与 $|y| \le k$ 等价, 因此 $|a + b| \le |a| + |b|$. □

定理 16.10 (极限的运算法则) 设 $\lim\limits_{x \to x_0} f(x) = L$ 且 $\lim\limits_{x \to x_0} g(x) = M$. 那么

1. **和差法则**.

$$\lim_{x \to x_0} [f(x) \pm g(x)] = \lim_{x \to x_0} f(x) \pm \lim_{x \to x_0} g(x) = L \pm M$$

(和的极限等于极限的和)

2. **乘积法则**.

$$\lim_{x \to x_0} [f(x) \cdot g(x)] = \left[\lim_{x \to x_0} f(x)\right] \cdot \left[\lim_{x \to x_0} g(x)\right] = L \cdot M$$

(积的极限等于极限的积)

3. **常数倍法则**(乘积法则的特例).

$$\lim_{x \to x_0} [c \cdot f(x)] = c \cdot \lim_{x \to x_0} f(x) = c \cdot L \quad (c \text{ 为常数})$$

4. **商法则**. 如果 $M \ne 0$, 那么

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to x_0} f(x)}{\lim\limits_{x \to x_0} g(x)} = \frac{L}{M}$$

(商的极限等于极限的商, 前提是分母极限不为零)

证明 由题设, $\lim\limits_{x \to x_0} f(x) = L$ 且 $\lim\limits_{x \to x_0} g(x) = M$. 根据极限的 $\varepsilon - \delta$ 定义,

- 对任意 $\varepsilon_1 > 0$, 存在 $\delta_1 > 0$, 使得当 $0 < |x - x_0| < \delta_1$ 时, 有 $|f(x) - L| < \varepsilon_1$;
- 对任意 $\varepsilon_2 > 0$, 存在 $\delta_2 > 0$, 使得当 $0 < |x - x_0| < \delta_2$ 时, 有 $|g(x) - M| < \varepsilon_2$.

下面我们用这些条件证明四条法则.

先证明和法则, 差法则只需把 g 换成 $-g$ 即可. 现证, 对任意给定的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $0 < |x - x_0| < \delta$ 时, 有 $|[f(x) + g(x)] - (L + M)| < \varepsilon$.

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

考虑目标表达式, 根据引理 16.2,

$$|[f(x)+g(x)]-(L+M)| = |(f(x)-L)+(g(x)-M)| \le |f(x)-L|+|g(x)-M|$$

为了使上式最终小于 ε, 我们让右侧的每一项都小于 $\varepsilon/2$.

根据极限的定义,

- 取 $\varepsilon_1 = \varepsilon/2$, 则存在 $\delta_1 > 0$, 使得当 $0 < |x-x_0| < \delta_1$ 时, 有 $|f(x)-L| < \varepsilon/2$;
- 取 $\varepsilon_2 = \varepsilon/2$, 则存在 $\delta_2 > 0$, 使得当 $0 < |x-x_0| < \delta_2$ 时, 有 $|g(x)-M| < \varepsilon/2$.

取 $\delta = \min(\delta_1, \delta_2)$. 当 $0 < |x-x_0| < \delta$ 时, 上述两个不等式同时成立, 因此

$$|[f(x)+g(x)]-(L+M)| \le |f(x)-L|+|g(x)-M| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon$$

这就证明了和法则.

接着证明乘积法则. 由于 $\lim\limits_{x\to x_0} g(x) = M$, 取 $\varepsilon_2 = 1$ 可知, 存在 $\delta_0 > 0$, 使得当 $0 < |x-x_0| < \delta_0$ 时有 $|g(x)-M| < 1$, 从而 $|g(x)| \le |M|+1$. 对任意给定的 $\varepsilon > 0$, 再取 $\delta_1 > 0$ 使 $|f(x)-L| < \frac{\varepsilon}{2(|M|+1)}$, 取 $\delta_2 > 0$ 使 $|g(x)-M| < \frac{\varepsilon}{2(|L|+1)}$. 当 $0 < |x-x_0| < \delta = \min(\delta_0, \delta_1, \delta_2)$ 时,

$$\begin{aligned} |f(x)g(x)-LM| &= |(f(x)-L)g(x)+L(g(x)-M)| \\ &\le |f(x)-L|\,|g(x)|+|L|\,|g(x)-M| \\ &< \frac{\varepsilon}{2(|M|+1)}(|M|+1)+\frac{|L|\varepsilon}{2(|L|+1)} < \varepsilon. \end{aligned}$$

因此乘积法则成立. 常数倍法则令 $g(x) \equiv c$ 即为乘积法则的特例.

最后证明商法则. 设 $M \neq 0$. 因 $g(x) \to M$, 存在 $\delta_0 > 0$, 使得当 $0 < |x-x_0| < \delta_0$ 时 $|g(x)-M| < \frac{|M|}{2}$, 从而 $|g(x)| > \frac{|M|}{2}$, 故分母不为零. 由乘积法则可知 $g(x)\cdot\frac{1}{g(x)} = 1$ 的极限若存在应满足 $M \cdot \lim\frac{1}{g(x)} = 1$; 直接验证如下.

$$\left|\frac{1}{g(x)}-\frac{1}{M}\right| = \frac{|g(x)-M|}{|M|\,|g(x)|} \le \frac{2|g(x)-M|}{|M|^2}.$$

因此 $\lim\limits_{x\to x_0}\frac{1}{g(x)} = \frac{1}{M}$. 再由乘积法则,

$$\lim_{x\to x_0}\frac{f(x)}{g(x)} = \lim_{x\to x_0} f(x)\cdot\lim_{x\to x_0}\frac{1}{g(x)} = \frac{L}{M}.$$

□

这些法则的要点是, **极限符号可以穿透四则运算**. 设每个部分的极限都存在, 并且除法中分母的极限不为零, 那么我们就可以先对每个部分求极限, 再进行运算.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 16.32 试利用极限的运算法则，计算 $\lim\limits_{x\to 2}(x^2+3x)$.

解 我们把这个极限一步步分解到最基本的极限形式.

$$\begin{aligned}\lim_{x\to 2}(x^2+3x) &= \lim_{x\to 2}(x^2)+\lim_{x\to 2}(3x) && \text{(和法则)}\\ &= \lim_{x\to 2}(x\cdot x)+3\cdot\lim_{x\to 2}(x) && \text{(积法则与常数倍法则)}\\ &= \left(\lim_{x\to 2}x\right)\cdot\left(\lim_{x\to 2}x\right)+3\cdot\left(\lim_{x\to 2}x\right) && \text{(再次应用积法则)}\end{aligned}$$

对于 $\lim\limits_{x\to 2}x$，其值为 2. 将此结果代入，得到

$$\begin{aligned}\text{原式} &= (2)\cdot(2)+3\cdot(2)\\ &= 4+6\\ &= 10\end{aligned}$$

注 16.4 (注记) 这个例题说明，对任意多项式函数 $P(x)$，极限 $\lim\limits_{x\to x_0}P(x)$ 的值都等于 $P(x_0)$. 这就是直接代入法的理论基础.

对于多项式函数、有理分式函数（分母不为零时）、三角函数、指数函数和对数函数，它们在其定义域内都连续. 若函数 $f(x)$ 在点 x_0 连续，那么 $\lim\limits_{x\to x_0}f(x)=f(x_0)$.

求极限的黄金法则

对任意初等函数 f，在其**定义域内**任取一点 x_0，求 $\lim\limits_{x\to x_0}f(x)$ 时，第一步总是**把 $x=x_0$ 直接代入函数表达式**.

- 如果得到一个确定的数值，那么这个数值就是所求的极限.
- 如果代入后出现**分母为零**，那么直接代入法失效，需要改用其他方法.

例题 16.33 求下列极限. (1) $\lim\limits_{x\to 2}(x^3-5x+1)$ (2) $\lim\limits_{x\to\frac{\pi}{4}}\dfrac{\sin x+\cos x}{x}$

解 (1) 该函数是多项式函数，在 $x=2$ 处连续. 直接代入，得到

$$\lim_{x\to 2}(x^3-5x+1)=2^3-5(2)+1=8-10+1=-1$$

(2) 该函数在 $x=\frac{\pi}{4}$ 处有定义且连续（分母不为零），故可直接代入，得到

$$\lim_{x\to\frac{\pi}{4}}\frac{\sin x+\cos x}{x}=\frac{\sin\frac{\pi}{4}+\cos\frac{\pi}{4}}{\frac{\pi}{4}}=\frac{\frac{\sqrt{2}}{2}+\frac{\sqrt{2}}{2}}{\frac{\pi}{4}}=\frac{\sqrt{2}}{\frac{\pi}{4}}=\frac{4\sqrt{2}}{\pi}$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

16.7.2 本节习题

16.17. 计算极限 $\lim\limits_{x\to 2}\dfrac{x^3-x^2-x-2}{x^2-4}$.

16.18. 计算极限 $\lim\limits_{x\to 1}\dfrac{\sqrt{x+3}-2}{\sqrt[3]{x}-1}$.

16.19. 计算极限 $\lim\limits_{x\to+\infty}(\sqrt{x^2+ax}-\sqrt{x^2+bx})$, 其中 a,b 为常数.

16.20. 计算极限 $\lim\limits_{x\to-\infty}\dfrac{\sqrt{x^2+1}}{x}$.

16.21. 已知 $\lim\limits_{x\to 2}\dfrac{x^2+ax+b}{x-2}=5$, 求常数 a,b 的值.

16.22. 考虑函数

$$f(x)=\begin{cases}\dfrac{\sin(\pi x)}{x-1}, & x\neq 1\\ k, & x=1\end{cases}$$

若函数 $f(x)$ 在点 $x=1$ 处连续, 求常数 k 的值.

16.23. 计算极限 $\lim\limits_{x\to 1}\dfrac{|x-1|}{x^2-1}$.

16.24. 确定常数 a,b 的值, 使得 $\lim\limits_{x\to+\infty}\left(\dfrac{x^2+1}{x+1}-ax-b\right)=0$.

16.25. 计算极限 $\lim\limits_{x\to 0}\dfrac{(1+x)^a-1}{x}$, 其中 a 为任意实数.

16.26. 计算极限 $\lim\limits_{x\to 0}\dfrac{\tan x-\sin x}{x^3}$.

16.27. 求函数 $f(x)=\sqrt{x^2+3x}$ 的所有渐近线.

16.28. 设函数 $f(x)$ 满足 $\lim\limits_{x\to 0}\dfrac{f(x)}{x}=L$, 其中 L 为非零常数. 求极限 $\lim\limits_{x\to 0}\dfrac{f(x^2)}{\sin^2 x}$.

16.29. 设函数 $f(x)$ 定义为 $f(x)=\lim\limits_{n\to\infty}\dfrac{x^{2n-1}+ax}{x^{2n}+1}$, 其中 n 为正整数. 试确定常数 a 的值, 使得 $f(x)$ 是一个连续函数.

16.30. 证明方程 $x^3-3x+1=0$ 在区间 $(0,1)$ 内至少存在一个实数根.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.8 无穷远处的极限与无穷小量

我们之前讨论的极限多是当 x 趋近于某个**有限值** x_0 时的情形. 但很多应用关心自变量 x 变得**无限大**或无限小时函数的长期变化趋势.

我们引入 $x \to +\infty$ 时的极限来描述这种最终趋势. 它用于研究函数的**渐近线**、比较函数的**增长速度**.

16.8.1 无穷远处的极限

定义 16.16 ($x \to +\infty$ 时的极限) 如果当 x 沿正方向无限增大时, 函数值 $f(x)$ 无限趋近于一个确定的常数 L, 那么我们就称 L 是函数 $f(x)$ 当 x 趋于正无穷时的极限, 记作

$$\lim_{x\to+\infty} f(x) = L \tag{16-7}$$

几何上, 这表示函数 $y = f(x)$ 的图像有一条**水平渐近线** $y = L$.

无穷远极限的一个基本结论是, 幂函数的倒数在无穷远处趋于零.

定理 16.11 (最重要的无穷极限) 对于任意正数 $n > 0$ 和任意实常数 c, 当 x 取正实数并趋于正无穷时, 有

$$\lim_{x\to+\infty} \frac{c}{x^n} = 0 \tag{16-8}$$

证明 按定义, 需证明对任意 $\varepsilon > 0$ 存在 $M > 0$, 使当 $x > M$ 时恒有 $\left|\dfrac{c}{x^n} - 0\right| < \varepsilon$. 现分 $c = 0$ 与 $c \neq 0$ 两种情形.

若 $c = 0$, 则对任意 $x \neq 0$, $\dfrac{c}{x^n} = 0$, 不等式化为 $0 < \varepsilon$, 由 $\varepsilon > 0$ 恒成立, 任取 $M > 0$, 例如 $M = 1$ 即可.

若 $c \neq 0$, 由目标不等式 $\left|\dfrac{c}{x^n}\right| < \varepsilon$ 出发定出 M. 因 $x \to +\infty$ 时只关心 $x > 0$, 此时 $x^n > 0$ 故 $|x^n| = x^n$, 不等式化为 $\dfrac{|c|}{x^n} < \varepsilon$. 由 $\varepsilon > 0$ 且 $x^n > 0$ 交叉相乘(不改变不等号方向)得 $|c| < \varepsilon\, x^n$, 即 $x^n > \dfrac{|c|}{\varepsilon}$; 因 $n > 0$, 两边取 n 次方根得 $x > \left(\dfrac{|c|}{\varepsilon}\right)^{1/n}$. 故取 $M = \left(\dfrac{|c|}{\varepsilon}\right)^{1/n}$.

验证此 M: 当 $x > M$ 时, $x > \left(\dfrac{|c|}{\varepsilon}\right)^{1/n}$, 两边取 n 次方(因 $n > 0$ 且两边为正, 合法)得 $x^n > \dfrac{|c|}{\varepsilon}$, 由 x^n, ε 均为正整理得 $\varepsilon > \dfrac{|c|}{x^n} = \left|\dfrac{c}{x^n}\right|$, 即 $\left|\dfrac{c}{x^n} - 0\right| < \varepsilon$.

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

综合两种情形, 对任意 $\varepsilon>0$ 都存在相应 $M>0$ 使当 $x>M$ 时 $|\frac{c}{x^n}-0|<\varepsilon$ 成立. 由定义, 定理成立. 证毕. □

这条公式是求有理函数在无穷远处极限的基本工具.

16.8.2 高次项、低次项与抓高次原则

我们处理多项式或有理分式在无穷远处的极限时, 一个有用的观察是, **在无穷的尺度下, 只有增长最快的那一项才重要**.

定义 16.17 (高次项与低次项) 在一个多项式中, 我们称次数最高的项为**最高次项**或**高阶项**, 其余的项为**低阶项**.
例如, 在多项式 $P(x)=3x^4-100x^2+5x-2024$ 中.

- $3x^4$ 是**最高次项**.
- $-100x^2, 5x, -2024$ 都是**低次项**.

当 x 变得非常大时, 例如 $x=10^9$, $3x^4$ 的值将远远超过所有低次项之和, 低次项的影响变得可以忽略不计. 这就是所谓的抓高次原则.

阅读提示

当 $x\to+\infty$ 时, 一个多项式或有理分式的极限, 完全由其分子和分母的**最高次项**所决定.

$$\lim_{x\to+\infty}\frac{a_nx^n+\cdots+a_0}{b_mx^m+\cdots+b_0}=\lim_{x\to+\infty}\frac{a_nx^n}{b_mx^m} \tag{16-9}$$

例题 16.34 求极限 $\lim\limits_{x\to+\infty}\dfrac{2x^3-4x^2+5}{7x^3+3x-1}$.

解 法一直接将分子、分母同除以最高次幂 x^3.

$$\lim_{x\to+\infty}\frac{2x^3-4x^2+5}{7x^3+3x-1}=\lim_{x\to+\infty}\frac{2-\frac{4}{x}+\frac{5}{x^3}}{7+\frac{3}{x^2}-\frac{1}{x^3}}=\frac{2}{7}.$$

法二使用抓高次原则.$x\to+\infty$ 时极限由分子、分母的最高次项决定,

$$\lim_{x\to+\infty}\frac{2x^3-4x^2+5}{7x^3+3x-1}=\lim_{x\to+\infty}\frac{2x^3}{7x^3}=\frac{2}{7}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

注 16.5 (有理函数无穷极限的快速判断)
对于有理函数 $f(x)=\dfrac{a_nx^n+\cdots}{b_mx^m+\cdots}$, 我们按以下方式判断它在 $x\to+\infty$ 时的极限.

- 若 $n>m$(分子次数 > 分母次数), 则函数趋于 ∞ 或 $-\infty$, 在有限实数意义下极限**不存在**.
- 若 $n<m$(分子次数 < 分母次数), 则极限为 **0**.
- 若 $n=m$(分子次数 = 分母次数), 则极限为最高次项系数之比 $\dfrac{a_n}{b_m}$.

16.8.3 高阶无穷大与低阶无穷大

16.8.3.1 抓高次原则的升维

我们在前面遇到的抓高次原则, 实质上是在比较不同次幂的幂函数 x^n 在 $x\to+\infty$ 时的增长速度. 可见 x^3 比 x^2 增长得快, x^2 又比 x 增长得快.

但如果问题变成比较 e^x 和 x^{10000} 哪个增长更快, 或者比较 $\ln x$ 和 $\sqrt[100]{x}$, 仅仅比较次数就不够用了. 我们需要一个更普适的工具来描述和比较**任意函数**在趋于无穷时的增长速率, 这就是阶的概念.

定义 16.18 (无穷大的阶) 假设 $\lim\limits_{x\to+\infty}f(x)=\infty$ 且 $\lim\limits_{x\to+\infty}g(x)=\infty$. 考察它们的比值在 $x\to+\infty$ 时的极限以比较增长速度.

- 如果 $\lim\limits_{x\to+\infty}\dfrac{f(x)}{g(x)}=\infty$, 我们称 $f(x)$ 是比 $g(x)$ **更高阶的无穷大**. 这说明 $f(x)$ 增长得比 $g(x)$ **快得多**.
- 如果 $\lim\limits_{x\to+\infty}\dfrac{f(x)}{g(x)}=0$, 我们称 $f(x)$ 是比 $g(x)$ **更低阶的无穷大**. 这说明 $f(x)$ 增长得比 $g(x)$ **慢得多**.
- 如果 $\lim\limits_{x\to+\infty}\dfrac{f(x)}{g(x)}=c$($c$ 为非零常数), 我们称 $f(x)$ 与 $g(x)$ 是**同阶无穷大**. 这说明它们的增长速度在同一量级. 例如 $2x^2$ 和 $10x^2$ 就是同阶无穷大.

注 16.6 (常见函数的增长速度阶梯)
当 $x\to+\infty$ 时, 我们熟知的几类函数, 其增长速度存在清晰的层级.
对数函数 $\ll$ 幂函数 $\ll$ 指数函数 $\ll$ 更高阶的函数
更具体地说, 对于任意的 $a>1, n>0$, 下面两个极限极为重要.

$$\lim_{x\to+\infty}\frac{\log_a x}{x^n}=0 \qquad \text{(幂函数增速 远快于对数函数)}$$

$$\lim_{x\to+\infty}\frac{x^n}{a^x}=0 \qquad \text{(指数函数增速 远快于幂函数)}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

一个更形象的排序是(从慢到快).

$$\cdots \ll \ln x \ll \sqrt{x} \ll x \ll x^2 \ll \cdots \ll (1.01)^x \ll e^x \ll 3^x \ll \cdots \ll x^x \ll \ldots$$

抓高次原则的终极形态

当 $x \to +\infty$ 时, 求一个由不同阶无穷大**相加减**构成的分式极限时, 我们只需要保留分子和分母中**最高阶的无穷大**, 其余低阶项均可忽略.

例题 16.35 求极限 $\lim\limits_{x\to+\infty} \dfrac{e^x - 100x^{10} + \ln x}{2e^x + 500x^2}$.

解 按抓高次原则, 关键是定出分子与分母中各自最高阶的无穷大. 依增长速度阶梯, 在分子 $e^x - 100x^{10} + \ln x$ 中指数函数 e^x 是最高阶无穷大, x^{10} 与 $\ln x$ 是其低阶无穷大; 分母 $2e^x + 500x^2$ 中同样以 e^x 为最高阶, x^2 为低阶. 保留最高阶项, 极限即 $\dfrac{e^x}{2e^x} = \dfrac{1}{2}$.

为使论证严格, 我们仿照有理函数的做法, 将分子、分母同除以最高阶的无穷大 e^x.

$$\lim_{x\to+\infty} \frac{e^x - 100x^{10} + \ln x}{2e^x + 500x^2} = \lim_{x\to+\infty} \frac{1 - 100\frac{x^{10}}{e^x} + \frac{\ln x}{e^x}}{2 + 500\frac{x^2}{e^x}}.$$

由增长速度阶梯, 当 $x \to +\infty$ 时 $\frac{x^{10}}{e^x} \to 0$、$\frac{\ln x}{e^x} \to 0$、$\frac{x^2}{e^x} \to 0$, 故

$$\text{原式} = \frac{1 - 100(0) + 0}{2 + 500(0)} = \frac{1}{2}.$$

16.8.4 常见函数的极限图像分析

我们通过观察函数图像, 总结几类基本初等函数在无穷远处或定义域边界处的极限行为.

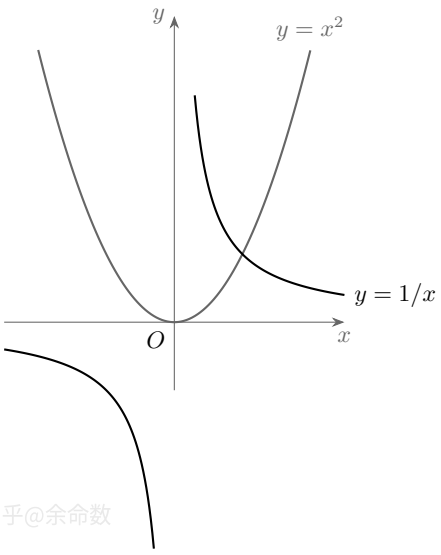
16.8.4.0.1 幂函数 $y = x^n$

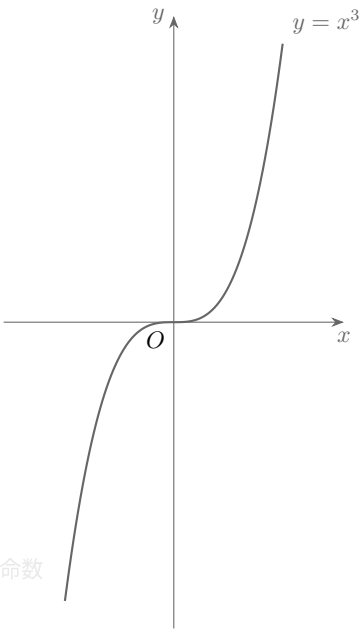
- 当 n 为**正整数**时, 图像向两侧无限延伸.

$$\lim_{x\to+\infty} x^n = +\infty$$

$$\lim_{x\to-\infty} x^n = \begin{cases} +\infty & (n \text{ 为偶数}) \\ -\infty & (n \text{ 为奇数}) \end{cases}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



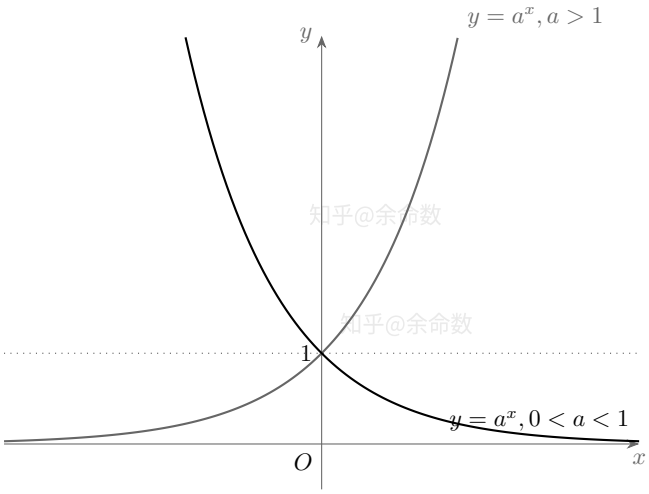


知乎@余命数
知乎@余命数
知乎@余命数

- 当 n 为**负整数**时, 如 $y = x^{-1} = \frac{1}{x}$, 图像以 x 轴为水平渐近线.

$$\lim_{x\to\pm\infty} \frac{1}{x^k} = 0 \quad (k \text{ 为正整数})$$

知乎@余命数



知乎@余命数
知乎@余命数
知乎@余命数

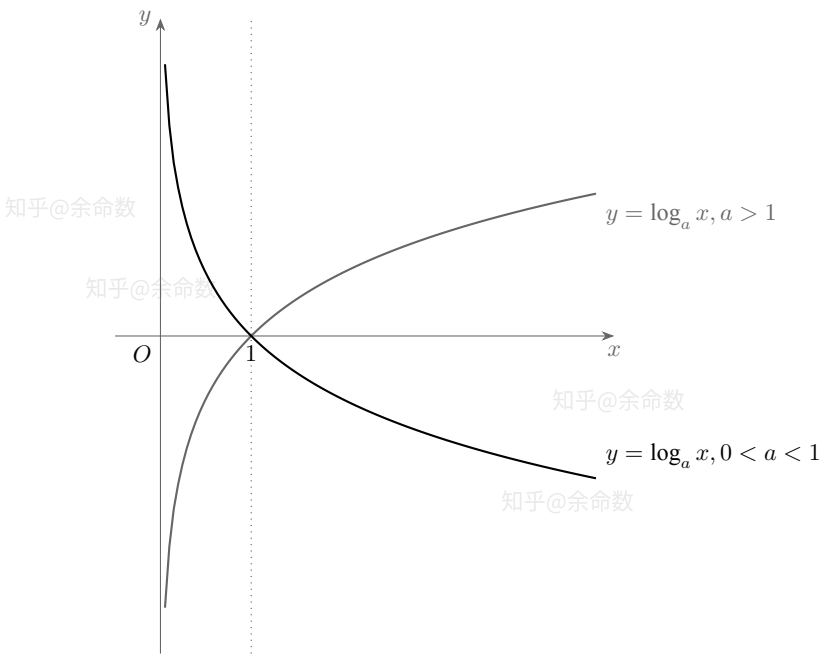
16.8.4.0.2 指数函数 $y = a^x$

- 当底数 $a > 1$ 时, 函数单调递增, 向右侧无限增长, 向左侧无限趋近于 0.

$$\lim_{x\to+\infty} a^x = +\infty, \quad \lim_{x\to-\infty} a^x = 0$$

- 当底数 $0<a<1$ 时, 函数单调递减, 向左侧无限增长, 向右侧无限趋近于 0.

$$\lim_{x\to+\infty} a^x=0,\quad \lim_{x\to-\infty} a^x=+\infty$$



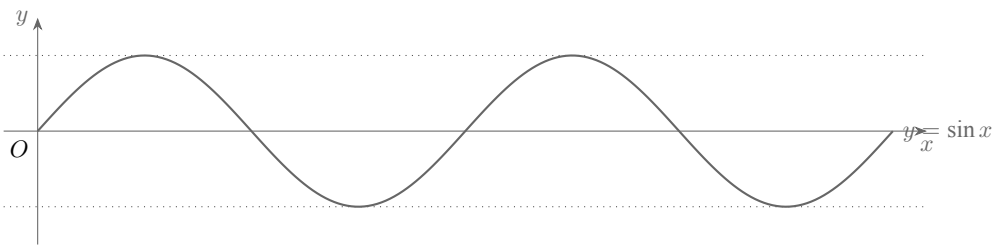
16.8.4.0.3　对数函数 $y=\log_a x$

- 当底数 $a>1$ 时, 函数单调递增. 图像向右无限缓慢上升, 以 y 轴为垂直渐近线向下无限延伸.

$$\lim_{x\to+\infty} \log_a x=+\infty,\quad \lim_{x\to0^+} \log_a x=-\infty$$

- 当底数 $0<a<1$ 时, 函数单调递减. 图像向右无限缓慢下降, 以 y 轴为垂直渐近线向上无限延伸.

$$\lim_{x\to+\infty} \log_a x=-\infty,\quad \lim_{x\to0^+} \log_a x=+\infty$$



知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.8.4.0.4 **三角函数** $y=\sin x, y=\cos x$ 从图像可见, 当 x 无限增大时, $\sin x$ 和 $\cos x$ 的值在 -1 和 1 之间永无休止地**振荡**, 它们并不趋近于任何一个**确定的**常数.

$$\lim_{x\to+\infty}\sin x \text{ 不存在}, \quad \lim_{x\to+\infty}\cos x \text{ 不存在}$$

16.8.5 本节习题

16.31. 分析函数 $f(x)=\dfrac{3x+5}{x^2+x+1}$ 在 $x\to\pm\infty$ 处的渐近行为, 并绘制其函数图像的示意图.

16.32. 分析函数 $f(x)=\dfrac{2x^2-3x}{x^2-1}$ 的渐近线(包含水平与垂直渐近线), 并绘制其函数图像的示意图.

16.33. 分析函数 $f(x)=\sqrt{x^2+6x}-x$ 在 $x\to+\infty$ 时的极限, 阐述其几何意义并绘制示意图.

16.34. 分析函数 $f(x)=\sqrt{x^2+6x}+x$ 在 $x\to-\infty$ 时的极限, 阐述其几何意义并绘制示意图.

16.35. 分析函数 $f(x)=\dfrac{x^3}{(x-1)^2}$ 的渐近线, 并绘制其函数图像的示意图.

16.36. 分析函数 $f(x)=\dfrac{x^2+\cos x}{2x^2-\sin x}$ 的水平渐近线, 并绘制其示意图.

16.37. 分析函数 $f(x)=\dfrac{e^x-e^{-x}}{e^x+e^{-x}}$ 在 $x\to+\infty$ 与 $x\to-\infty$ 处的极限, 绘制示意图以展示其两条不同的水平渐近线.

16.38. 分析函数 $f(x)=\dfrac{\ln(x^2+1)}{x}$ 的水平渐近线, 并绘制示意图.

16.39. 分析函数 $f(x)=x(e^{1/x}-1)$ 的水平渐近线, 并绘制其示意图.

16.40. 分析函数 $f(x)=\left(1+\dfrac{1}{x}\right)^{2x}$ 的水平渐近线, 并绘制其示意图.

16.41. 分析函数 $f(x)=x+\arctan x$ 的渐近行为. 阐述该函数为何有两条不同的斜渐近线并求出其方程, 随后绘制示意图.

16.42. 求函数 $f(x)=\sqrt{4x^2+x}$ 的所有斜渐近线, 并绘制其示意图.

16.43. 确定常数 a,b 的值, 使得 $\lim\limits_{x\to+\infty}\left(\sqrt{4x^2+bx}-(ax-1)\right)=3$, 并解释 a,b 取值背后的数学原理.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.44.　考虑函数 $f(x)=\lim\limits_{n\to\infty}\dfrac{x^{2n+1}+2x}{x^{2n}+1}$, 其中 n 为正整数. 请分段写出 $f(x)$ 的解析式, 分析其间断点, 并绘制其函数图像的示意图.

16.9 导数的概念

在极限的帮助下, 我们终于可以触及一个在科学与哲学上困扰了人类近两千年的古老难题, 即如何精确地描述一个事物在**某一瞬间**的变化速度.

古希腊的芝诺悖论就说明了这一困境. 一支飞行的箭, 在任何一个确定的瞬间, 它都占据着一个确定的位置, 似乎是静止的. 那么, 它的运动究竟体现在哪里, 这个问题实质上是混淆了瞬间和时间段的区别. 在有限的时间段内, 可计算平均速度, 但在一个没有长度的瞬间, 速度的概念似乎失效了.

直到 17 世纪, 科学的发展迫切需要解决这一难题. 天文学家开普勒需要研究行星运动的瞬时速度, 物理学家伽利略需要描述落体在每一时刻的速度. 他们都遇到了同样的问题, 却只能用近似和模糊的语言来描述. 这一挑战最终由牛顿和莱布尼茨以相似的思路攻克.

16.9.1 牛顿与莱布尼茨的逼近思想

牛顿从物理学的视角出发, 关心物体的**瞬时速度**; 莱布尼茨则从几何学的视角入手, 醉心于曲线的**切线问题**. 尽管切入点不同, 他们都意识到面对的是同一个矛盾, 只能计算两点之间的量(平均变化率), 却想知道一点处的量, 即瞬时变化率. 直接计算会导致 $\frac{0}{0}$ 的困境.

他们的解决方法是, **既然无法直接到达那一点, 就用无限逼近来代替.**

我们跟随莱布尼茨的几何思路来重温这一过程.

16.9.1.0.1　从割线斜率出发　要研究曲线 $y=f(x)$ 在点 $P(x_0,f(x_0))$ 的倾斜程度, 无法直接下手, 可在曲线上另取一点 $Q(x_0+\Delta x,f(x_0+\Delta x))$, 然后计算连接 P,Q 两点的**割线**的斜率. 这是一个简单的初中数学问题.

$$k_{\text{割线}}=\frac{\text{纵坐标之差}}{\text{横坐标之差}}=\frac{\Delta y}{\Delta x}=\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$$

这个值, 代表了函数在 x_0 到 $x_0+\Delta x$ 这段区间内的**平均变化率**.

16.9.1.0.2　动态逼近的过程　接着, 让点 Q 沿着曲线向点 P 移动. 这个过程的实质, 就是让横坐标的增量 Δx 越来越小, 趋向于 0.

随着 $\Delta x\to 0$, 点 Q 无限逼近点 P, 连接它们的割线也在不断旋转. 其最终归宿, 就是那条在点 P 处与曲线相切的**切线**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

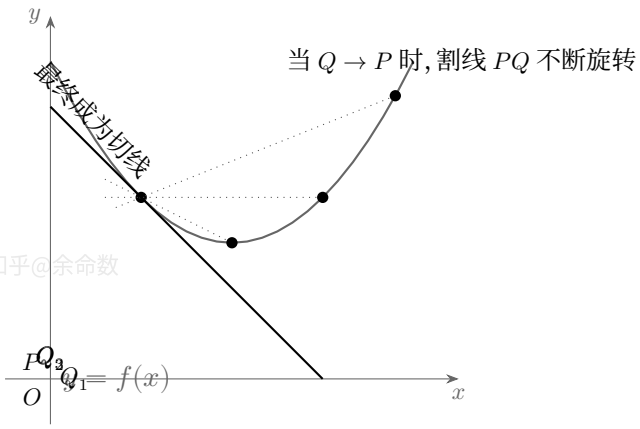


图 16.9 割线旋转逼近切线的过程

16.9.1.0.3 用极限锁定最终结果 割线的斜率 $k_{\text{割线}} = \frac{\Delta y}{\Delta x}$ 在这个过程中不断变化. 它所无限逼近的那个**极限值**就是切线的斜率. 这一逼近思想绕开了 Δx 直接等于 0 的障碍, 不问当 $\Delta x = 0$ 时是什么, 只问当 $\Delta x \to 0$ 时比值的**极限**是什么.

$$k_{\text{切线}} = \lim_{\Delta x \to 0} k_{\text{割线}} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

至此, 几何上的切线斜率问题和物理上的瞬时速度问题, 都通过这个统一的、基于极限的思想得到了解决.

16.9.2 导数的诞生

这个通过平均变化率求极限得到的瞬时变化率是一个基本而实用的概念, 数学家们给它取名为**导数**.

定义 16.19 (函数在一点的导数) 函数 $y = f(x)$ 在点 x_0 处的**导数** $f'(x_0)$, 定义为函数在该点的平均变化率当自变量增量趋于零时的极限, 即

$$f'(x_0) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \tag{16-10}$$

如果这个极限存在, 我们就说函数 $f(x)$ 在点 x_0 处是**可导的**.

注 16.7 (导数的三种意义)

- **几何意义**. $f'(x_0)$ 就是曲线 $y = f(x)$ 在点 $(x_0, f(x_0))$ 处**切线的斜率**.
- **物理意义**. 如果 $s(t)$ 是位移关于时间的函数, 那么 $s'(t_0)$ 就是物体在时刻 t_0 的**瞬时速度**.
- **一般意义**. 导数 $f'(x_0)$ 衡量了函数 $f(x)$ 在点 x_0 处变化的**快慢程度**, 是函数的**瞬时变化率**.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

16.9.3 用定义求导

我们把求瞬时变化率这个模糊的想法精确地固化为导数的极限定义.

$$f'(x)=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

这就是导数的定义. 接下来运用它, 实际计算一个函数的导数.

下面沿着早期微积分先驱的思路, 不借助现成的求导公式, 仅凭极限的定义计算一些简单函数的导数. 这一过程会使导数的瞬时变化率含义更加具体.

注 16.8 (用定义求导)
根据导数的定义, 求解函数 $f(x)$ 的导数 $f'(x)$ 的过程分为三步.

1. **求增量**. 计算当自变量从 x 变为 $x+\Delta x$ 时, 函数值的增量 $\Delta y=f(x+\Delta x)-f(x)$.
2. **作商**. 计算平均变化率 $\dfrac{\Delta y}{\Delta x}$. 这一步要进行代数化简, 尤其是通过因式分解、有理化等技巧, **约分掉分母中的** Δx, 从而消除 $\Delta x\to 0$ 时 $\frac{0}{0}$ 的不定形式.
3. **取极限**. 计算 $\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}$, 所得的结果就是导函数 $f'(x)$.

例题 16.36　求函数 $f(x)=c$(c 为常数)的导数.

解 这是一个最简单的例子. 它代表了一条水平直线, 直觉上认为它的变化率(斜率)处处为 0.

求增量

$$\Delta y=f(x+\Delta x)-f(x)=c-c=0$$

作商当 $\Delta x\neq 0$ 时,

$$\frac{\Delta y}{\Delta x}=\frac{0}{\Delta x}=0$$

取极限

$$f'(x)=\lim_{\Delta x\to 0}0=0$$

于是 $(c)'=0$, 印证了直觉, 常数函数没有变化, 其瞬时变化率恒为零.

例题 16.37　求函数 $f(x)=x^2$ 的导数.

解 这是第一个非线性函数, 它的倾斜程度是随 x 变化的.

$$\begin{aligned}\text{求增量}\Delta y&=f(x+\Delta x)-f(x)\\&=(x+\Delta x)^2-x^2\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

$$= (x^2 + 2x\Delta x + (\Delta x)^2) - x^2$$

$$= 2x\Delta x + (\Delta x)^2$$

知乎@余命数

作商
$$\frac{\Delta y}{\Delta x} = \frac{2x\Delta x + (\Delta x)^2}{\Delta x}$$
$$= \frac{\Delta x(2x + \Delta x)}{\Delta x} \quad \text{(关键一步:提取公因式并约分)}$$
$$= 2x + \Delta x$$

这里的差商只在 $\Delta x \neq 0$ 时讨论, 因而上述约分是合法的.

取极限
$$f'(x) = \lim_{\Delta x \to 0} (2x + \Delta x)$$
$$= 2x$$

知乎@余命数

于是 $(x^2)' = 2x$. 这个基本结果表明, 抛物线 $y = x^2$ 在任意一点 x 处的切线斜率恰好是该点横坐标的两倍. 例如在 $x = 1$ 处斜率为 2, 在 $x = -3$ 处斜率为 -6. 导数本身是一个新的函数.

例题 16.38 求函数 $f(x) = \dfrac{1}{x}$ 的导数.

解 这个例子需要运用分式通分的技巧. 以下计算在 $x \neq 0$ 处进行, 且取 Δx 足够小时有 $x + \Delta x \neq 0$.

知乎@余命数

$$\Delta y = \frac{1}{x + \Delta x} - \frac{1}{x}$$
$$= \frac{x - (x + \Delta x)}{x(x + \Delta x)} \quad \text{(通分)}$$
$$= \frac{-\Delta x}{x(x + \Delta x)}$$

知乎@余命数

$$\frac{\Delta y}{\Delta x} = \frac{-\Delta x}{x(x + \Delta x)} \cdot \frac{1}{\Delta x}$$
$$= \frac{-1}{x(x + \Delta x)}$$

知乎@余命数

知乎@余命数

$$f'(x) = \lim_{\Delta x \to 0} \frac{-1}{x(x + \Delta x)}$$
$$= \frac{-1}{x(x + 0)} = -\frac{1}{x^2}$$

知乎@余命数

于是当 $x \neq 0$ 时, $\left(\frac{1}{x}\right)' = -\frac{1}{x^2}$.

以上例子表明, 导数的定义本身蕴含着实用的计算能力. 然而这一过程虽然逻辑清晰, 对于稍复杂一点的函数(比如 $f(x) = x^5$ 或 $f(x) = \sin x$), 代数变形将变得异常繁琐.

这正是数学家们曾经面临的困境. 为摆脱每次都从极限定义出发的低效做法, 他们基于定义推导出了一系列普适的、可以直接使用的**求导法则与公式**.

16.9.4 基本求导公式的诞生

16.9.4.1 来龙去脉

在上一节中, 我们用导数定义直接计算了导数. 这个过程虽然严谨, 但对每一个新函数都必须重复一遍复杂的极限运算, 无疑是低效的. 正如我们不会每次计算乘法都去重复地做加法一样, 数学家们也渴望拥有更直接、更高效的求导工具.

这一节的任务是进行最后一次艰苦的劳动, 利用导数的极限定义, 系统地推导出高中阶段遇到的所有基本初等函数的导数. 一旦这些**基本求导公式**建立起来, 套用它们就能替代每次重复的极限计算, 把繁琐的运算压缩成可以直接使用的结论.

知乎@余命数

在处理三角函数、指数函数与对数函数的求导之前, 要先解决两个以 $\frac{0}{0}$ 或 1^∞ 形式出现的特定极限, 直接使用极限的四则运算法则对它们无能为力. 这两个极限就是

$$\lim_{x\to 0} \frac{\sin x}{x} \quad 和 \quad \lim_{x\to 0} (1+x)^{\frac{1}{x}}$$

它们沟通了代数与超越函数世界. 为了攻克它们, 必须先掌握一个直观且实用的极限法则, 即三明治定理.

16.9.4.2 三明治定理 *(选读内容)

引理 16.3　如果在点 x_0 的某个去心邻域内, 有三个函数 $g(x), f(x), h(x)$ 始终满足

$$g(x) \leq f(x) \leq h(x)$$

并且, $g(x)$ 与 $h(x)$ 的极限存在且相等

$$\lim_{x\to x_0} g(x) = \lim_{x\to x_0} h(x) = L$$

那么, 被夹在中间的 $f(x)$ 的极限也必然存在, 且等于 L:

$$\lim_{x\to x_0} f(x) = L$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

证明 现证 $\lim\limits_{x\to x_0} f(x)=L$. 根据极限的 $\varepsilon-\delta$ 定义, 需证明

> 对于任意给定的 $\varepsilon>0$, 存在实数 $\delta>0$, 使得当 $0<|x-x_0|<\delta$ 时, 不等式 $|f(x)-L|<\varepsilon$ 恒成立.

由已知条件, $\lim\limits_{x\to x_0} g(x)=L$, 即对上述给定的 $\varepsilon>0$, 存在 $\delta_1>0$, 使当 $0<|x-x_0|<\delta_1$ 时 $|g(x)-L|<\varepsilon$, 即 $L-\varepsilon<g(x)<L+\varepsilon$. 同理, $\lim\limits_{x\to x_0} h(x)=L$, 即对同一 $\varepsilon>0$, 存在 $\delta_2>0$, 使当 $0<|x-x_0|<\delta_2$ 时 $|h(x)-L|<\varepsilon$, 即 $L-\varepsilon<h(x)<L+\varepsilon$. 又已知 $g(x)\le f(x)\le h(x)$ 在 x_0 的某去心邻域内成立, 设其半径为 $\delta_3>0$.

接着, 让这三个条件同时生效, 需选取使 x 同时位于三个邻域之内的范围. 取

$$\delta=\min(\delta_1,\delta_2,\delta_3),$$

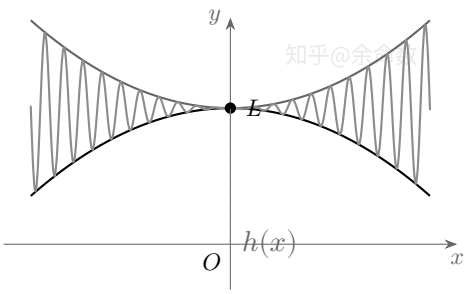
由 $\delta_1,\delta_2,\delta_3$ 均为正知 $\delta>0$.

考察任一满足 $0<|x-x_0|<\delta$ 的 x, 上述三个条件同时成立, 串联起来, 由条件 1 有 $L-\varepsilon<g(x)$, 由条件 3 有 $g(x)\le f(x)$, 故 $L-\varepsilon<g(x)\le f(x)$; 又由条件 3 有 $f(x)\le h(x)$, 由条件 2 有 $h(x)<L+\varepsilon$, 故 $f(x)\le h(x)<L+\varepsilon$. 合并左右两半得不等式链 $L-\varepsilon<f(x)<L+\varepsilon$, 即 $|f(x)-L|<\varepsilon$.

至此, 已证明, 对于任意给定的 $\varepsilon>0$, 已找到 $\delta>0$, 即 $\min(\delta_1,\delta_2,\delta_3)$, 使得当 $0<|x-x_0|<\delta$ 时, 有 $|f(x)-L|<\varepsilon$.

这完全符合极限的 $\varepsilon-\delta$ 定义.

证毕. □

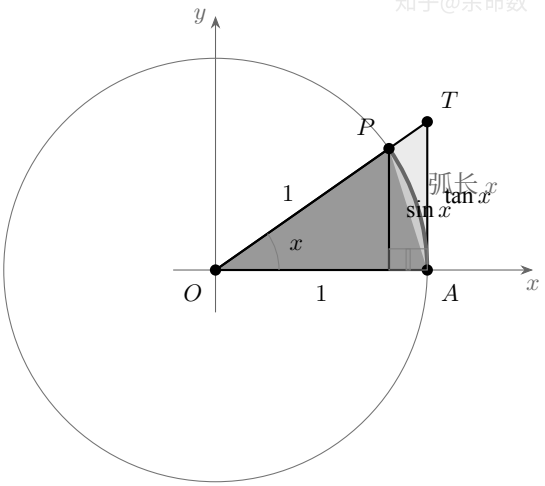


这个定理的直观意义是, 如果一个函数被另外两个在某点收敛于同一值的函数夹住, 那么它只能收敛到同一个值.

例题 16.39 证明 $\lim\limits_{x\to 0}\frac{\sin x}{x}=1$.

解 这个极限的证明是三明治定理最经典的几何应用. 考虑在单位圆中, 一个很小的正角 x(弧度制).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



比较图中三角形、扇形与大三角形三块区域的面积, 得

$$\text{面积}(\triangle OAP) < \text{面积}(\text{扇形}OAP) < \text{面积}(\triangle OAT)$$

分别计算这三个面积(单位圆中半径 $r=1$), 得

- 面积 $(\triangle OAP) = \frac{1}{2}|OA| \cdot |PC| = \frac{1}{2} \cdot 1 \cdot \sin x = \frac{1}{2}\sin x$
- 面积 $(\text{扇形}OAP) = \frac{1}{2}r^2 x = \frac{1}{2} \cdot 1^2 \cdot x = \frac{1}{2}x$
- 面积 $(\triangle OAT) = \frac{1}{2}|OA| \cdot |AT| = \frac{1}{2} \cdot 1 \cdot \tan x = \frac{1}{2}\tan x$

代入面积不等式, 并同乘以 2, 得

$$\sin x < x < \tan x$$

因为考虑的是一个很小的正角 $x \to 0^+$, 所以 $\sin x > 0$. 将不等式同除以 $\sin x$:

$$1 < \frac{x}{\sin x} < \frac{\tan x}{\sin x} = \frac{1}{\cos x}$$

取倒数, 不等号反向, 得

$$\cos x < \frac{\sin x}{x} < 1$$

至此夹逼的结构已经形成, 考察两侧函数的极限, 得

$$\lim_{x \to 0^+} \cos x = \cos 0 = 1$$

$$\lim_{x \to 0^+} 1 = 1$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据三明治定理, 被夹在中间的函数极限也必须是 1, 即

$$\lim_{x\to 0^+}\frac{\sin x}{x}=1$$

对于 $x\to 0^-$ 的情况, 由于 $\frac{\sin(-x)}{-x}=\frac{-\sin x}{-x}=\frac{\sin x}{x}$ 是一个偶函数, 其左极限等于右极限. 因此, 证得第一个重要极限.

$$\lim_{x\to 0}\frac{\sin x}{x}=1 \tag{16-11}$$

第二个重要极限的推导, 更像是一个发现与定义的过程. 它源于对复利问题的思考.

假设银行年利率为 100

- 如果每年计息一次, 一年后本利和为 $1\times(1+1)=2$ 元.
- 如果每半年计息一次, 一年后本利和为 $1\times(1+\frac{1}{2})^2=2.25$ 元.
- 如果每月计息一次, 一年后本利和为 $1\times(1+\frac{1}{12})^{12}\approx 2.613$ 元.
- 如果每天计息一次, 一年后本利和为 $1\times(1+\frac{1}{365})^{365}\approx 2.715$ 元.

可见, 随着计息周期越来越短, 即计息次数 n 越来越大, 本利和 $(1+\frac{1}{n})^n$ 似乎在逼近一个确定的数值. 数学家们证明了这个极限确实存在, 并把这个极限值定义为一个基本的无理数, 即**自然常数** e.

定义 16.20 (自然常数 e) 自然常数 e 定义为当 n 趋于无穷时, 表达式 $(1+\frac{1}{n})^n$ 的极限值.

$$e=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\approx 2.71828\ldots \tag{16-12}$$

严格证明还需说明当连续变量趋于无穷时 $(1+\frac{1}{t})^t$ 也趋于同一个数 e. 在这个结论基础上, 令 $x=\frac{1}{t}$, 则当 $t\to\infty$ 时 $x\to 0^+$, 再结合左侧极限可得到第二个重要极限.

$$\lim_{x\to 0}(1+x)^{\frac{1}{x}}=e \tag{16-13}$$

这两个重要极限, 是后续求导推导的必备工具.

16.9.4.3 幂函数 $f(x)=x^n$ (n 为正整数)

这是最重要、最基础的求导公式. 推导它的要点是二项式定理 $(a+b)^n=C_n^0a^n+C_n^1a^{n-1}b+C_n^2a^{n-2}b^2+\cdots+C_n^nb^n$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
(x^n)' &= \lim_{\Delta x\to 0}\frac{(x+\Delta x)^n - x^n}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{(C_n^0x^n + C_n^1x^{n-1}\Delta x + C_n^2x^{n-2}(\Delta x)^2 + \dots) - x^n}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{nx^{n-1}\Delta x + \frac{n(n-1)}{2}x^{n-2}(\Delta x)^2 + \dots}{\Delta x} \quad (\text{因为} C_n^0 = 1, C_n^1 = n)\\
&= \lim_{\Delta x\to 0}\left(nx^{n-1} + \frac{n(n-1)}{2}x^{n-2}\Delta x + \dots\right)
\end{aligned}
$$

当 $\Delta x\to 0$ 时, 除了第一项 nx^{n-1} 外, 所有其余项都含有 Δx 因子, 因此它们的极限都为 0.

知乎@余命数

$$(x^n)' = nx^{n-1}$$

上面的二项式推导证明了正整数指数情形. 更一般地, 在相应可导定义域内, 幂函数公式对实数指数也成立, 例如我们之前求过的 $(x^2)' = 2x^{2-1} = 2x$ 和 $(\frac{1}{x})' = (x^{-1})' = -1\cdot x^{-2} = -\frac{1}{x^2}$ 都符合该规律.

知乎@余命数

16.9.4.4 三角函数 $f(x) = \sin x$

推导三角函数的导数, 需要用到一个重要的极限 $\lim\limits_{u\to 0}\dfrac{\sin u}{u} = 1$. 以及和差化积公式 $\sin A - \sin B = 2\cos\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)$.

知乎@余命数

$$
\begin{aligned}
(\sin x)' &= \lim_{\Delta x\to 0}\frac{\sin(x+\Delta x) - \sin x}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{2\cos\left(\frac{x+\Delta x+x}{2}\right)\sin\left(\frac{x+\Delta x-x}{2}\right)}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{2\cos\left(x+\frac{\Delta x}{2}\right)\sin\left(\frac{\Delta x}{2}\right)}{\Delta x}\\
&= \lim_{\Delta x\to 0}\cos\left(x+\frac{\Delta x}{2}\right)\cdot\frac{\sin\left(\frac{\Delta x}{2}\right)}{\frac{\Delta x}{2}} \quad (\text{凑重要极限的形式})
\end{aligned}
$$

知乎@余命数

由于 $\lim\limits_{\Delta x\to 0}\cos(x+\frac{\Delta x}{2}) = \cos x$ 且 $\lim\limits_{\Delta x\to 0}\frac{\sin(\frac{\Delta x}{2})}{\frac{\Delta x}{2}} = 1$, 于是得

$$(\sin x)' = \cos x$$

同理可证(作为练习).

$$(\cos x)' = -\sin x$$

16.9.4.5 指数函数 $f(x)=a^x$

推导指数函数的导数, 需要用到另一个重要的极限, 它与自然对数的底 e 有关, $\lim\limits_{h\to 0}\dfrac{a^h-1}{h}=\ln a$.

$$\begin{aligned}(a^x)'&=\lim_{\Delta x\to 0}\frac{a^{x+\Delta x}-a^x}{\Delta x}\\&=\lim_{\Delta x\to 0}\frac{a^x(a^{\Delta x}-1)}{\Delta x}\\&=a^x\cdot\lim_{\Delta x\to 0}\frac{a^{\Delta x}-1}{\Delta x}\end{aligned}$$

根据重要极限的定义, 得

$$(a^x)'=a^x\ln a$$

当底数是特殊的 e 时, 由于 $\ln e=1$, 得最简洁的求导公式

$$(e^x)'=e^x$$

这个函数 e^x 的特殊之处在于, 它在任意一点的变化率恰好等于它在该点本身的函数值.

16.9.4.6 对数函数 $f(x)=\log_a x$

这是推导中技巧性最强的一个, 它需要用到对数运算法则以及 e 的极限定义 $\lim\limits_{u\to 0}(1+u)^{\frac{1}{u}}=e$.

$$\begin{aligned}(\log_a x)'&=\lim_{\Delta x\to 0}\frac{\log_a(x+\Delta x)-\log_a x}{\Delta x}\\&=\lim_{\Delta x\to 0}\frac{1}{\Delta x}\log_a\left(\frac{x+\Delta x}{x}\right)=\lim_{\Delta x\to 0}\frac{1}{\Delta x}\log_a\left(1+\frac{\Delta x}{x}\right)\\&=\lim_{\Delta x\to 0}\log_a\left(\left(1+\frac{\Delta x}{x}\right)^{\frac{1}{\Delta x}}\right)\quad(\text{利用}c\log A=\log A^c)\end{aligned}$$

为了凑出 e 的定义, 我们进行代换, 令 $t=\frac{\Delta x}{x}$, 则 $\Delta x=xt$, 当 $\Delta x\to 0$ 时 $t\to 0$.

$$\begin{aligned}(\log_a x)'&=\lim_{t\to 0}\log_a\left((1+t)^{\frac{1}{xt}}\right)=\lim_{t\to 0}\log_a\left(\left((1+t)^{\frac{1}{t}}\right)^{\frac{1}{x}}\right)\\&=\log_a\left(\lim_{t\to 0}\left((1+t)^{\frac{1}{t}}\right)^{\frac{1}{x}}\right)\quad(\text{对数函数连续, 极限可以进到内部})\\&=\log_a\left(e^{\frac{1}{x}}\right)=\frac{1}{x}\log_a e\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

利用对数的换底公式 $\log_a e = \frac{\ln e}{\ln a} = \frac{1}{\ln a}$，得最终结果

$$(\log_a x)' = \frac{1}{x \ln a}$$

同样，当底数是 e 时，由于 $\ln a = \ln e = 1$，得一个非常简洁的形式

$$(\ln x)' = \frac{1}{x}$$

经过这一系列推导，即得基本求导公式，汇总如下表.

阅读提示

函数	导数
$f(x) = c$（常数）	$f'(x) = 0$
$f(x) = x^n$（$n \in \mathbb{R}$，在相应定义域内）	$f'(x) = nx^{n-1}$
$f(x) = \sin x$	$f'(x) = \cos x$
$f(x) = \cos x$	$f'(x) = -\sin x$
$f(x) = a^x$	$f'(x) = a^x \ln a$
$f(x) = e^x$	$f'(x) = e^x$
$f(x) = \log_a x$	$f'(x) = \frac{1}{x \ln a}$
$f(x) = \ln x$	$f'(x) = \frac{1}{x}$

至此已有一张基本求导公式表，它在求导中的作用正如算术中的乘法口诀表. 接下来需要学习如何运用这些基本公式，去处理由它们通过加、减、乘、除等运算组合而成的更复杂的函数.

告别了每次都要从极限定义出发的低效做法，我们进入高效的公式化阶段. 首先学习求导运算中最基础的两个法则.

阅读提示

设函数 $f(x), g(x)$ 均可导，c 为常数，则：

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

1. **常数倍法则**

$$[c \cdot f(x)]' = c \cdot f'(x)$$

求导时, 常数系数可以原封不动地提到导数符号外面.

2. **加减法则**

$$[f(x) \pm g(x)]' = f'(x) \pm g'(x)$$

和的导数等于导数的和, 差的导数等于导数的差. 即可以将复杂的函数拆开, 逐项求导, 最后再合并结果.

证明 (常数倍法则的证明) 令 $h(x) = c \cdot f(x)$, 并利用导数的定义来求 $h'(x)$.

根据导数的定义, 得

$$\begin{aligned}
[c \cdot f(x)]' &= \lim_{\Delta x \to 0} \frac{h(x+\Delta x) - h(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{c \cdot f(x+\Delta x) - c \cdot f(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{c \cdot [f(x+\Delta x) - f(x)]}{\Delta x} \quad (\text{在分子中提取公因式} c) \\
&= c \cdot \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \quad (\text{利用极限的常数倍运算法则}) \\
&= c \cdot f'(x) \quad (\text{根据} f'(x) \text{ 的定义})
\end{aligned}$$

这就证明了常数系数在求导时可以被直接提到求导符号的外面. □

证明 (加减法则的证明 (以加法为例)) 令 $h(x) = f(x) + g(x)$, 并利用导数的定义来求 $h'(x)$.

根据导数的定义, 得

$$\begin{aligned}
[f(x) + g(x)]' &= \lim_{\Delta x \to 0} \frac{h(x+\Delta x) - h(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{[f(x+\Delta x) + g(x+\Delta x)] - [f(x) + g(x)]}{\Delta x}
\end{aligned}$$

重新组合分子中的项, 得

$$= \lim_{\Delta x \to 0} \frac{[f(x+\Delta x) - f(x)] + [g(x+\Delta x) - g(x)]}{\Delta x}$$

将分数拆分成两部分, 并利用极限的加法运算法则, 得

$$= \lim_{\Delta x \to 0} \left[\frac{f(x+\Delta x) - f(x)}{\Delta x} + \frac{g(x+\Delta x) - g(x)}{\Delta x} \right]$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} + \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x}$$

最后, 根据 $f'(x)$ 和 $g'(x)$ 的定义, 得

$$= f'(x) + g'(x)$$

和的导数等于导数的和得证. 减法法则的证明过程完全类似, 只需将上述过程中的加号替换为减号即可. □

结合基本求导公式和这两个运算法则, 已能够解决一大类常见函数的求导问题了.

例题 16.40　求多项式函数 $f(x) = 2x^3 - 5x^2 + 7x - 1$ 的导数.

解

$$\begin{aligned}
f'(x) &= (2x^3 - 5x^2 + 7x - 1)' \\
&= (2x^3)' - (5x^2)' + (7x)' - (1)' && \text{(利用加减法则, 逐项求导)} \\
&= 2(x^3)' - 5(x^2)' + 7(x^1)' - (1)' && \text{(提出常数系数)} \\
&= 2(3x^2) - 5(2x) + 7(1) - 0 && \text{(套用幂函数和常数求导公式)} \\
&= 6x^2 - 10x + 7 && \text{(最后化简整理)}
\end{aligned}$$

这是求导运算中最基本的应用, 一个复杂的多项式求导被分解为一系列简单的幂函数求导的组合.

例题 16.41　求函数 $y = 3\sqrt{x} - \dfrac{2}{x^2}$ 在定义域 $x > 0$ 上的导数.

解 要点是将所有项都改写为幂函数 x^n 的形式. 以下在函数定义域 $x > 0$ 内求导.

$$y = 3x^{1/2} - 2x^{-2}$$

接着, 可像处理多项式一样来求导, 得

$$\begin{aligned}
y' &= (3x^{1/2} - 2x^{-2})' \\
&= (3x^{1/2})' - (2x^{-2})' \\
&= 3(x^{1/2})' - 2(x^{-2})' \\
&= 3\left(\frac{1}{2}x^{\frac{1}{2}-1}\right) - 2\left(-2x^{-2-1}\right)
\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{3}{2}x^{-1/2} - (-4)x^{-3}$$
$$= \frac{3}{2\sqrt{x}} + \frac{4}{x^3}$$

(应用幂函数公式时, 指数为小数和负数要格外小心; 最后将结果写回根式与分式形式更清晰.)

例题 16.42 求函数 $f(x) = 4\cos x + e^x - \ln 2$ 的导数.

解 这个例子混合了三角函数、指数函数和一个常数项($\ln 2$ 是一个常数).

$$\begin{aligned} f'(x) &= (4\cos x + e^x - \ln 2)' \\ &= (4\cos x)' + (e^x)' - (\ln 2)' \\ &= 4(\cos x)' + (e^x)' - (\ln 2)' \\ &= 4(-\sin x) + e^x - 0 \quad \text{(套用三角函数、指数函数和常数求导公式)} \\ &= -4\sin x + e^x \end{aligned}$$

可见, 掌握了基本求导公式和加减、数乘运算法则后, 已经能够处理任何由基本函数加减组合而成的函数.

但现实世界中的函数关系远比这复杂, 函数还会相乘、相除, 比如 $f(x) = x^2\sin x$, 或者相互复合, 如 $y = \sin(2x+1)$.

对于这些情况, 简单的加减法则已无能为力, 还需要**乘法法则、除法法则和链式法则**.

16.9.5 乘、除、复合求导法则

16.9.5.1 三大运算法则

阅读提示

设函数 $u(x)$ 和 $v(x)$ 在点 x 处均可导.

1. **乘法法则**

$$[u(x)v(x)]' = u'(x)v(x) + u(x)v'(x)$$

用语言说, 两函数乘积的导数, 等于第一个函数的导数乘以第二个函数, 加上第一个函数乘以第二个函数的导数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

2. **除法法则**

$$\left[\frac{u(x)}{v(x)}\right]' = \frac{u'(x)v(x) - u(x)v'(x)}{[v(x)]^2} \quad (v(x) \neq 0)$$

用语言说, 两函数商的导数, 等于分子的导数乘以分母, 减去分子乘以分母的导数, 所得的差再除以分母的平方.

3. **链式法则**设函数 $y = f(u)$ 的自变量 u 本身是另一个关于 x 的函数 $u = g(x)$, 则复合函数 $y = f(g(x))$ 对 x 的导数为

$$[f(g(x))]' = f'(g(x)) \cdot g'(x)$$

用语言说, 复合函数的导数, 等于外层函数对内层函数求导, 乘以内层函数对自变量求导.

16.9.5.1.1　链式法则的莱布尼茨表示　莱布尼茨符号把链式法则书写为:若记 $y = f(u), u = g(x)$, 则链式法则写作

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

此形式在结构上酷似分数的约分, 非常直观. 它提醒我们, 导数的传递过程是逐层相乘.

这些法则均可由导数的极限定义严格推导.

知乎@余命数

16.9.5.2　乘法法则的证明

证明的要点在于构造项, 即添一项, 减一项的技巧.

$$\begin{aligned}
[u(x)v(x)]' &= \lim_{\Delta x \to 0} \frac{u(x+\Delta x)v(x+\Delta x) - u(x)v(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{1}{\Delta x}\Big(u(x+\Delta x)v(x+\Delta x) - u(x)v(x+\Delta x) \\
&\qquad + u(x)v(x+\Delta x) - u(x)v(x)\Big) \\
&= \lim_{\Delta x \to 0}\left(\frac{u(x+\Delta x) - u(x)}{\Delta x}\, v(x+\Delta x)\right. \\
&\qquad \left. + u(x)\,\frac{v(x+\Delta x) - v(x)}{\Delta x}\right) \\
&= \left(\lim_{\Delta x \to 0} \frac{u(x+\Delta x) - u(x)}{\Delta x}\right) \cdot \left(\lim_{\Delta x \to 0} v(x+\Delta x)\right)
\end{aligned}$$

知乎@余命数　知乎@余命数　知乎@余命数

$$+u(x)\cdot\left(\lim_{\Delta x\to 0}\frac{v(x+\Delta x)-v(x)}{\Delta x}\right)$$

由于函数 $v(x)$ 可导则必然连续,所以 $\lim\limits_{\Delta x\to 0}v(x+\Delta x)=v(x)$. 其余两项分别是 $u'(x)$ 和 $v'(x)$ 的定义. 因此,

$$[u(x)v(x)]'=u'(x)v(x)+u(x)v'(x)$$

证毕.

16.9.5.3 除法法则的证明

证明过程与乘法法则类似, 首先对分数作差进行通分, 再应用构造项的技巧.

$$\begin{aligned}\left[\frac{u(x)}{v(x)}\right]'&=\lim_{\Delta x\to 0}\frac{1}{\Delta x}\left(\frac{u(x+\Delta x)}{v(x+\Delta x)}-\frac{u(x)}{v(x)}\right)\\&=\lim_{\Delta x\to 0}\frac{u(x+\Delta x)v(x)-u(x)v(x+\Delta x)}{\Delta x\cdot v(x)v(x+\Delta x)}\\&=\lim_{\Delta x\to 0}\frac{1}{\Delta x\cdot v(x)v(x+\Delta x)}\Big(u(x+\Delta x)v(x)-u(x)v(x)\\&\qquad+u(x)v(x)-u(x)v(x+\Delta x)\Big)\\&=\lim_{\Delta x\to 0}\frac{1}{v(x)v(x+\Delta x)}\left(\frac{u(x+\Delta x)-u(x)}{\Delta x}v(x)\right.\\&\qquad\left.-u(x)\frac{v(x+\Delta x)-v(x)}{\Delta x}\right)\\&=\frac{1}{[v(x)]^2}[u'(x)v(x)-u(x)v'(x)]\end{aligned}$$

证毕.

16.9.5.4 链式法则的证明

链式法则的严格证明涉及对内层函数增量是否为零的讨论,过程较为精细,证明从略, 以下关注其应用.

16.9.6 三大法则的应用与实战

16.9.6.1 识别最外层结构

掌握了这三个基本法则后,由基本初等函数构成的任意复杂函数都能求导,关键在于**准确识别函数最外层的代数结构**.

最外层是两个函数**相乘**的, 应用乘法法则; **相除**的, 应用除法法则; 一个函数**复合**另一个函数的, 应用链式法则. 施加法则时, 如果遇到某个组成部分需要求导,

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再对该部分重复这个识别过程，直至所有求导问题都化归为基本求导公式. 这个过程可以看作是一个递归分析的过程.

例题 16.43 (乘法法则) 求函数 $f(x)=x^2\sin x$ 的导数.

解 结构分析. 函数 $f(x)$ 是 $u(x)=x^2$ 与 $v(x)=\sin x$ 的**乘积**. 因此，首要应用乘法法则.

根据乘法法则 $[uv]'=u'v+uv'$：

$$\begin{aligned} f'(x) &= (x^2)'\cdot\sin x + x^2\cdot(\sin x)' \\ &= (2x)\cdot\sin x + x^2\cdot(\cos x) \\ &= 2x\sin x + x^2\cos x \end{aligned}$$

例题 16.44 (除法法则) 求函数 $y=\dfrac{e^x}{x^2+1}$ 的导数.

解 结构分析. 函数 y 是 $u(x)=e^x$ 与 $v(x)=x^2+1$ 的**商**. 应用除法法则.

根据除法法则 $\left[\dfrac{u}{v}\right]'=\dfrac{u'v-uv'}{v^2}$：

$$\begin{aligned} y' &= \frac{(e^x)'(x^2+1)-e^x(x^2+1)'}{(x^2+1)^2} \\ &= \frac{e^x(x^2+1)-e^x(2x)}{(x^2+1)^2} \\ &= \frac{e^x(x^2-2x+1)}{(x^2+1)^2} \\ &= \frac{e^x(x-1)^2}{(x^2+1)^2} \end{aligned}$$

在分子上提取公因式 e^x 使表达式更为清晰，这是一个良好的代数习惯.

例题 16.45 (链式法则) 求函数 $y=\cos(x^3-4x)$ 的导数.

解 结构分析. 函数 y 的结构是 $\cos(\cdot)$ 函数**复合**一个多项式 x^3-4x. 这是一个典型的复合函数，必须使用链式法则，从外层向内层逐层处理.

- 外层函数为 $y=\cos(u)$
- 内层函数为 $u=x^3-4x$
- 外导，即 y 对 u 求导，$\dfrac{dy}{du}=(\cos u)'=-\sin u$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

- 内导, 即 u 对 x 求导, $\dfrac{du}{dx} = (x^3 - 4x)' = 3x^2 - 4$

  根据链式法则 $\dfrac{dy}{dx} = \dfrac{dy}{du} \cdot \dfrac{du}{dx}$:

$$\begin{aligned}\frac{dy}{dx} &= (-\sin u) \cdot (3x^2 - 4) \\ &= -\sin(x^3 - 4x) \cdot (3x^2 - 4) \\ &= -(3x^2 - 4)\sin(x^3 - 4x)\end{aligned}$$

(最后一步必须将中间变量 u 换回关于 x 的表达式.)

例题 16.46 (多层复合) 求函数 $f(x) = \ln(\sin(e^{x^2}))$ 的导数.

解 这是一个典型的多层复合函数, 结构层次清晰, 需要从最外层开始, 逐层向内求导, 并将每一层的导数相乘. 以下计算均在函数有定义的点进行.

层次分解我们把这个函数分解为四层.

- 最外层为 $y = \ln(u)$
- 第二层为 $u = \sin(v)$
- 第三层为 $v = e^w$
- 最内层为 $w = x^2$

逐层求导并应用链式法则根据链式法则 $[f(g(h(k(x))))]' = f' \cdot g' \cdot h' \cdot k'$, 有

$$\begin{aligned}f'(x) &= \frac{1}{\sin(e^{x^2})} \cdot [\sin(e^{x^2})]' \\ &= \frac{1}{\sin(e^{x^2})} \cdot \cos(e^{x^2}) \cdot [e^{x^2}]' \quad \text{(对第二层求导)} \\ &= \frac{\cos(e^{x^2})}{\sin(e^{x^2})} \cdot e^{x^2} \cdot [x^2]' \quad \text{(对第三层求导)} \\ &= \cot(e^{x^2}) \cdot e^{x^2} \cdot (2x) \quad \text{(对最内层求导)}\end{aligned}$$

整理, 得最终结果

$$f'(x) = 2xe^{x^2}\cot(e^{x^2})$$

例题 16.47 求函数 $y = x^3\sqrt{1 - x^2}$ 的导数.

解 函数在 $-1 < x < 1$ 内可导. 其**最外层结构**是两个部分 $u(x) = x^3$ 和 $v(x) = \sqrt{1 - x^2}$ 的**乘积**. 因此, 须首先应用乘法法则. 计算 $v'(x)$ 时, 需要应用链式法则.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

根据乘法法则 $y'=u'v+uv'$, 先写出求导框架

$$y'=(x^3)'\sqrt{1-x^2}+x^3(\sqrt{1-x^2})'$$

第一部分导数很简单.

$$(x^3)'=3x^2$$

第二部分 $(\sqrt{1-x^2})'$ 需要使用链式法则. 将根式写为幂函数形式 $(1-x^2)^{\frac{1}{2}}$. 外层函数是 $w^{\frac{1}{2}}$, 内层函数是 $w=1-x^2$.

$$(\sqrt{1-x^2})'=\frac{1}{2}(1-x^2)^{\frac{1}{2}-1}\cdot(1-x^2)'=\frac{1}{2}(1-x^2)^{-\frac{1}{2}}\cdot(-2x)=\frac{-x}{\sqrt{1-x^2}}$$

将各部分代入乘法法则的框架中, 得

$$\begin{aligned}y'&=3x^2\sqrt{1-x^2}+x^3\left(\frac{-x}{\sqrt{1-x^2}}\right)\\&=3x^2\sqrt{1-x^2}-\frac{x^4}{\sqrt{1-x^2}}\end{aligned}$$

知乎@余命数

为了得到更简洁的形式, 进行通分, 得

$$\begin{aligned}y'&=\frac{3x^2(1-x^2)-x^4}{\sqrt{1-x^2}}\\&=\frac{3x^2-3x^4-x^4}{\sqrt{1-x^2}}\\&=\frac{3x^2-4x^4}{\sqrt{1-x^2}}=\frac{x^2(3-4x^2)}{\sqrt{1-x^2}}\end{aligned}$$

知乎@余命数

例题 16.48　求函数 $f(x)=\frac{x^2e^{2x}}{\ln(x^2+1)}$ 的导数.

解　此函数在 $x\neq 0$ 时有定义, 最外层结构是**商**, 因此主框架是除法法则. 然而, 分子 $u(x)=x^2e^{2x}$ 本身是一个**乘积**, 求其导数需要乘法法则. 分子中的 e^{2x} 和分母 $v(x)=\ln(x^2+1)$ 求导则需要**链式法则**.

知乎@余命数

令 $u(x)=x^2e^{2x}$, $v(x)=\ln(x^2+1)$. 先求出 $u'(x)$ 和 $v'(x)$.

求 $u'(x)$(应用乘法法则):

$$\begin{aligned}u'(x)&=(x^2)'e^{2x}+x^2(e^{2x})'\\&=2x\cdot e^{2x}+x^2\cdot(e^{2x}\cdot 2)\quad(\text{对}e^{2x}\text{ 应用链式法则})\\&=2xe^{2x}(1+x)\end{aligned}$$

求 $v'(x)$(应用链式法则):

$$v'(x)=\frac{1}{x^2+1}\cdot(x^2+1)'=\frac{2x}{x^2+1}$$

接着,将 u,v,u',v' 代入除法法则 $\left[\frac{u}{v}\right]'=\frac{u'v-uv'}{v^2}$:

$$\begin{aligned}f'(x)&=\frac{[2xe^{2x}(1+x)]\cdot\ln(x^2+1)-(x^2e^{2x})\cdot\left(\frac{2x}{x^2+1}\right)}{[\ln(x^2+1)]^2}\\&=\frac{2xe^{2x}\left((1+x)\ln(x^2+1)-\frac{x^2}{x^2+1}\right)}{[\ln(x^2+1)]^2}\end{aligned}$$

知乎@余命数

例题 16.49 求函数 $g(x)=\left(\ln(x\sin^2 x)\right)^3$ 的导数.

解 复合结构由外到内依次为 u^3、$\ln v$、$v=x\sin^2 x$、以及 w^2 $(w=\sin x)$. 在 $x\sin^2 x>0$ 的定义域内,

$$g'(x)=3\big(\ln(x\sin^2 x)\big)^2\cdot\frac{1}{x\sin^2 x}\cdot(x\sin^2 x)'.$$

由乘法法则与链式法则,

$$(x\sin^2 x)'=\sin^2 x+x\cdot 2\sin x\cdot\cos x=\sin^2 x+x\sin(2x).$$

知乎@余命数

代回得

$$g'(x)=\frac{3(\sin^2 x+x\sin(2x))}{x\sin^2 x}\big(\ln(x\sin^2 x)\big)^2.$$

知乎@余命数 知乎@

例题 16.50 求函数 $y=(x^2+1)^{\ln x}$ 的导数.

解 函数形式为 $f(x)^{g(x)}$,底数和指数均为变量,必须使用对数求导法.

知乎@余命数

$$\ln y=\ln((x^2+1)^{\ln x})=(\ln x)\cdot\ln(x^2+1)$$

知乎@余命数

左边应用链式法则,右边应用乘法法则.

$$(\ln y)'=[(\ln x)\cdot\ln(x^2+1)]'$$

$$\frac{1}{y}\cdot\frac{dy}{dx}=(\ln x)'\cdot\ln(x^2+1)+(\ln x)\cdot(\ln(x^2+1))'$$

分别计算右侧的导数,得

$$(\ln x)'=\frac{1}{x}$$

$$(\ln(x^2+1))' = \frac{1}{x^2+1}\cdot(x^2+1)' = \frac{2x}{x^2+1}$$

代回方程, 得

$$\frac{1}{y}\frac{dy}{dx} = \frac{1}{x}\ln(x^2+1) + (\ln x)\frac{2x}{x^2+1}$$

将 y 乘到右边, 并代入其原始表达式, 得

$$\frac{dy}{dx} = y\left(\frac{\ln(x^2+1)}{x} + \frac{2x\ln x}{x^2+1}\right)$$

知乎@余命数

$$\frac{dy}{dx} = (x^2+1)^{\ln x}\left(\frac{\ln(x^2+1)}{x} + \frac{2x\ln x}{x^2+1}\right)$$

例题 16.51　求函数 $f(x) = \ln\left(\frac{\sqrt{x^2+1}-x}{\sqrt{x^2+1}+x}\right)$ 的导数.

知乎@余命数

解　直接对复合式求导会陷入冗长的商法则与链式法则, 宜先化简对数内部. 分子有理化, 得

$$\frac{\sqrt{x^2+1}-x}{\sqrt{x^2+1}+x} = \frac{(\sqrt{x^2+1}-x)^2}{(x^2+1)-x^2} = (\sqrt{x^2+1}-x)^2.$$

故 $f(x) = \ln((\sqrt{x^2+1}-x)^2) = 2\ln(\sqrt{x^2+1}-x)$. 求导, 得

$$\begin{aligned} f'(x) &= 2\cdot\frac{1}{\sqrt{x^2+1}-x}\cdot\left(\frac{x}{\sqrt{x^2+1}}-1\right) \\ &= \frac{2}{\sqrt{x^2+1}-x}\cdot\frac{x-\sqrt{x^2+1}}{\sqrt{x^2+1}} = -\frac{2}{\sqrt{x^2+1}}. \end{aligned}$$

知乎@余

若不化简而直接求导, 令 $u = \frac{\sqrt{x^2+1}-x}{\sqrt{x^2+1}+x}$, 则

$$f'(x) = \frac{1}{u}u' = \frac{\sqrt{x^2+1}+x}{\sqrt{x^2+1}-x}\cdot\left(\frac{\sqrt{x^2+1}-x}{\sqrt{x^2+1}+x}\right)'.$$

记 $P = \sqrt{x^2+1}-x, Q = \sqrt{x^2+1}+x$, 则

$$P' = \frac{x}{\sqrt{x^2+1}} - 1, \qquad Q' = \frac{x}{\sqrt{x^2+1}} + 1.$$

由商法则 $\left(\frac{P}{Q}\right)' = \frac{P'Q-PQ'}{Q^2}$:

$$\left(\frac{\sqrt{x^2+1}-x}{\sqrt{x^2+1}+x}\right)'$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$
\begin{aligned}
&= \frac{\left(\frac{x}{\sqrt{x^2+1}}-1\right)(\sqrt{x^2+1}+x)-(\sqrt{x^2+1}-x)\left(\frac{x}{\sqrt{x^2+1}}+1\right)}{(\sqrt{x^2+1}+x)^2} \\
&= \frac{\frac{x(\sqrt{x^2+1}+x)}{\sqrt{x^2+1}}-(\sqrt{x^2+1}+x)-\frac{x(\sqrt{x^2+1}-x)}{\sqrt{x^2+1}}-(\sqrt{x^2+1}-x)}{(\sqrt{x^2+1}+x)^2} \\
&= \frac{\frac{x\sqrt{x^2+1}+x^2-x\sqrt{x^2+1}+x^2}{\sqrt{x^2+1}}-(\sqrt{x^2+1}+x+\sqrt{x^2+1}-x)}{(\sqrt{x^2+1}+x)^2} \\
&= \frac{\frac{2x^2}{\sqrt{x^2+1}}-2\sqrt{x^2+1}}{(\sqrt{x^2+1}+x)^2} \\
&= \frac{\frac{2x^2-2(x^2+1)}{\sqrt{x^2+1}}}{(\sqrt{x^2+1}+x)^2} \\
&= \frac{\frac{2x^2-2x^2-2}{\sqrt{x^2+1}}}{(\sqrt{x^2+1}+x)^2} \\
&= \frac{-2}{\sqrt{x^2+1}(\sqrt{x^2+1}+x)^2}
\end{aligned}
$$

将这个复杂的结果代回 $f'(x)$ 的表达式, 得

$$
\begin{aligned}
f'(x) &= \frac{\sqrt{x^2+1}+x}{\sqrt{x^2+1}-x}\cdot\frac{-2}{\sqrt{x^2+1}(\sqrt{x^2+1}+x)^2} \\
&= \frac{-2}{(\sqrt{x^2+1}-x)\sqrt{x^2+1}(\sqrt{x^2+1}+x)} \\
&= \frac{-2}{\sqrt{x^2+1}[(\sqrt{x^2+1}-x)(\sqrt{x^2+1}+x)]} \\
&= \frac{-2}{\sqrt{x^2+1}[(x^2+1)-x^2]} \\
&= \frac{-2}{\sqrt{x^2+1}[1]} \\
&= -\frac{2}{\sqrt{x^2+1}}
\end{aligned}
$$

虽然最终也能化简到 $-\frac{2}{\sqrt{x^2+1}}$, 但过程的繁琐程度和出错风险远高于法一.

可见, 面对含嵌套分式或根式的结构时, 先化简再求导通常更稳妥. 直接套用商法则与链式法则虽然可行, 但计算量大、出错风险高.

例题 16.52 (综合应用) 求函数 $y = e^{-x}\ln(x^2)$ 的导数.

解 结构分析. 此函数在 $x \neq 0$ 时有定义, 最外层的结构是 e^{-x} **乘以** $\ln(x^2)$, 因此首要步骤是应用**乘法法则**. 在计算乘法法则的各个部分时, 可见 $(e^{-x})'$ 和 $(\ln(x^2))'$ 需要用到**链式法则**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

应用乘法法则, 先写出框架

$$y' = (e^{-x})' \ln(x^2) + e^{-x}(\ln(x^2))'$$

接着, 分别计算两个需要链式法则求导的部分, 得

- 求 $(e^{-x})'$: 外层为 e^u, 内层为 $u = -x$. 导数为 $e^u \cdot u' = e^{-x} \cdot (-1) = -e^{-x}$.
- 求 $(\ln(x^2))'$: 外层为 $\ln u$, 内层为 $u = x^2$. 导数为 $\dfrac{1}{u} \cdot u' = \dfrac{1}{x^2} \cdot 2x = \dfrac{2}{x}$.

最后, 将计算结果代回乘法法则的框架中, 得

$$\begin{aligned} y' &= (-e^{-x}) \cdot \ln(x^2) + e^{-x} \cdot \left(\frac{2}{x}\right) \\ &= e^{-x}\left(\frac{2}{x} - \ln(x^2)\right) \end{aligned}$$

注 16.9 (常见错误与关键技巧)

- **混淆运算法则**
  一个根本性错误是将乘法或除法法则误用为线性法则, 例如认为 $[u(x)v(x)]' = u'(x)v'(x)$. 必须明确, 导数运算对于乘法和除法不具有这种简单的分配性质.
- **链式法则遗漏内导**
  这是应用链式法则时最频发的错误, 即**只对外层函数求导, 而忘记乘以内层函数的导数**.
  例如, 将 $(\sin(2x))'$ 错误地计算为 $\cos(2x)$.
  正确过程 $(\sin(2x))' = \cos(2x) \cdot (2x)' = 2\cos(2x)$. 内层的导数是链条中必要的一环.
- **复合层次识别不清**
  对于 $y = \sin^3 x$, 其真实结构是 $y = (\sin x)^3$. 最外层是**幂函数** u^3, 内层是三角函数 $u = \sin x$.
  求导 $y' = 3(\sin x)^{3-1} \cdot (\sin x)' = 3\sin^2 x \cos x$.
  而对于 $y = \sin(x^3)$, 最外层是**正弦函数** $\sin u$, 内层是幂函数 $u = x^3$.
  求导 $y' = \cos(x^3) \cdot (x^3)' = \cos(x^3) \cdot 3x^2 = 3x^2\cos(x^3)$.
- **先化简, 再求导**
  在启动复杂的求导法则之前, 应先审视函数表达式是否可以化简. 例如, 对函数 $y = \ln(e^{2x})$ 求导, 若直接使用链式法则, 则 $y' = \frac{1}{e^{2x}} \cdot (e^{2x})' = \frac{1}{e^{2x}} \cdot 2e^{2x} = 2$. 若先化简 $y = \ln(e^{2x}) = 2x$, 则 $y' = 2$, 过程更为简捷.

16.9.7 求导技巧选讲 *

至此, 已构建一套实用的、普适的求导法则体系. 对于任何能明确写成 $y = f(x)$ 形式的函数, 似乎都能通过乘法、除法与链式法则解决. 然而数学的世界充满了更多奇特的结构:

- 当一个函数的结构是幂的幂或者一长串的乘积时, 比如 $y = \dfrac{\sqrt{x+1}(x-2)^3}{(x+5)^4}$, 直接使用乘除法和链式法则会非常繁琐.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

- 当函数关系隐藏在一个方程中, 比如圆的方程 $x^2+y^2=25$, 就需要另想办法求出曲线上任意一点的切线斜率 $\frac{dy}{dx}$.

16.9.7.1 取对数求导法

在处理导数问题时, 若函数表达式呈现变量的变量次幂或多项因子的乘除开方结构, 常规的求导法则往往导致运算冗杂. 取对数求导法利用对数运算将乘除转化为加减、将幂转化为乘积的代数性质, 从而简化求导结构.

注 16.10 (取对数求导法)
设 $y=f(x)$ 在某个区间上满足 $f(x)\neq 0$(幂指结构 $u(x)^{v(x)}$ 还要求 $u(x)>0$). 若 f 的表达式由多项因子的积、商或幂指结构构成, 可按如下路径求导:

1. 考虑等式两边绝对值的自然对数 $\ln|y|=\ln|f(x)|$;
2. 利用对数性质展开右端, 将其化为各项的代数和;
3. 视 y 为 x 的函数, 两边对 x 求导, 左端应用链式法则得到 $\frac{y'}{y}$;
4. 解出 y', 即 $y'=y\cdot[\ln|f(x)|]'$.

例题 16.53 求幂指函数 $y=x^{\sin x}$ $(x>0)$ 的导数.

解 观察函数结构, 底数 x 与指数 $\sin x$ 均随变量变化, 既非单纯的幂函数也非指数函数, 无法直接套用基本公式. 考虑引入对数以分离底数与指数.

在 $x>0$ 的假设下, 两边取自然对数, 得

$$\ln y=\ln(x^{\sin x})=\sin x\cdot\ln x$$

此时, 复杂的幂指关系转化为三角函数与对数函数的乘积.

方程两边对 x 求导. 左边应用链式法则, 右边应用乘法法则, 得

$$\frac{1}{y}\cdot y'=(\sin x)'\ln x+\sin x(\ln x)'$$

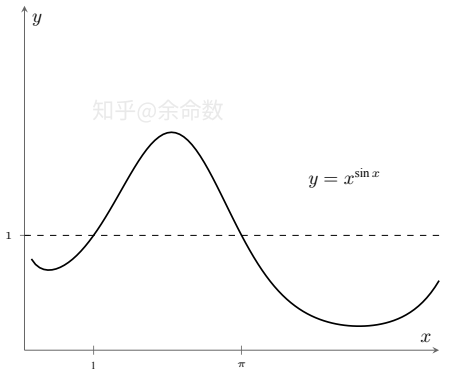
即

$$\frac{y'}{y}=\cos x\ln x+\frac{\sin x}{x}$$

将 $y=x^{\sin x}$ 代回整理, 得最终结果

$$y'=x^{\sin x}\left(\cos x\ln x+\frac{\sin x}{x}\right)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



知乎@余命数

例题 16.54　求函数 $y=\dfrac{\sqrt{x+1}(x-2)^3}{(x+5)^4}$ 的导数.

解　此函数由三个代数式的幂通过乘除构成. 若直接使用商的求导法则, 分子部分需嵌套乘法法则, 且涉及根式求导, 计算路径冗长. 利用对数性质将运算线性化更加自然.

先在 $x>-1$ 且 $x\neq 2$ 的区域考察绝对值的对数, 利用 $\ln(abc)=\ln a+\ln b+\ln c$ 及 $\ln(a^n)=n\ln a$ 展开, 得

$$\begin{aligned}\ln|y| &= \ln\left|\frac{(x+1)^{1/2}(x-2)^3}{(x+5)^4}\right| \\ &= \frac{1}{2}\ln|x+1|+3\ln|x-2|-4\ln|x+5|\end{aligned}$$

两边同时对 x 求导. 注意 $\frac{d}{dx}(\ln|u|)=\frac{u'}{u}$, 运算被简化为简单的分式求和, 得

$$\frac{y'}{y}=\frac{1}{2(x+1)}+\frac{3}{x-2}-\frac{4}{x+5}$$

最后解出 y', 即用原函数表达式乘以对数导数项, 得

$$y'=\frac{\sqrt{x+1}(x-2)^3}{(x+5)^4}\left[\frac{1}{2(x+1)}+\frac{3}{x-2}-\frac{4}{x+5}\right]$$

该式适用于 $x>-1$ 且 $x\neq 2$. 当 $x=2$ 时, 原函数可写为 $y=A(x)(x-2)^3$, 其中 $A(x)=\dfrac{\sqrt{x+1}}{(x+5)^4}$ 在 $x=2$ 附近可导, 故 $y'(2)=0$.

通过这种变换, 原本需要多重嵌套法则的运算被分解为独立的对数求导, 结构清晰且易于检验.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.9.7.2 隐函数求导法

当 x 和 y 的关系由一个方程 $F(x,y)=0$ 确定, 而没有, 或很难写成 $y=f(x)$ 的形式时, 我们就称这个关系定义了一个**隐函数**. 求这种函数的导数, 需要直接在方程上操作.

注 16.11 (隐函数求导法要点)
将方程中的 y 看作是 x 的一个未知函数 $y(x)$. 然后, 将方程两边**同时对 x 求导**. 在这个过程中, 凡是遇到与 y 相关的项, 都要应用**链式法则**. 最后, 从得到的方程中解出 y'.

例题 16.55 求由圆的方程 $x^2+y^2=25$ 所确定的函数在任意点 (x,y) 的导数 $\dfrac{dy}{dx}$.

解 这是典型的隐函数. 将方程两边同时对 x 求导.

$$\frac{d}{dx}(x^2+y^2)=\frac{d}{dx}(25)$$

利用加法法则展开, 得

$$\frac{d}{dx}(x^2)+\frac{d}{dx}(y^2)=0$$

- 对 x^2 求导, 即 $2x$.
- 对 y^2 求导则不同. 因为 y 是 x 的函数, 所以 y^2 是一个复合函数, 外层是平方运算, 内层是 $y(x)$. 根据链式法则, 得

$$\frac{d}{dx}(y^2)=2y\cdot\frac{dy}{dx}=2y\cdot y'$$

将求导结果代回方程, 得

$$2x+2y\cdot y'=0$$

接着, 把它当作一个关于 y' 的普通代数方程来解, 得

$$2y\cdot y'=-2x$$

$$y'=-\frac{2x}{2y}=-\frac{x}{y}\quad(y\neq 0)$$

这个结果符合几何直观, 它表明, 圆上满足 $y\neq 0$ 的任意一点 (x,y) 的切线斜率就是 $-\frac{x}{y}$. 例如在点 $(3,4)$, 切线斜率为 $-\frac{3}{4}$. 当 $y=0$ 时, 对应点为 $(\pm 5,0)$, 切线为竖直直线, 此时不能写成有限的 $\dfrac{dy}{dx}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

16.9.8 单调性

导数 $f'(x)$ 精确地刻画了函数在一点的瞬时变化率, 其几何意义为该点切线的斜率, 这是一个局部性质. 本节利用函数在每一点的瞬时变化信息, 推断其在整个区间上的增减趋势, 即单调性.

导数的正负符号直观上预示了函数的走向, 正的导数意味着上升趋势, 负的导数意味着下降趋势. 然而这还不够严谨, 需要为这一直观建立严格的理论基础, 并形成一套分析方法.

16.9.8.1 从理论基石到判定法则

此探索始于对函数图像中转折点的精确描述, 这些点是函数单调性可能发生改变的地方, 我们称之为极值点.

定义 16.21 (函数的极值) 设函数 $y=f(x)$ 在点 x_0 的某个邻域内有定义.

- 如果对于该邻域内**任意**异于 x_0 的点 x, 都恒有 $f(x)<f(x_0)$, 那么我们称 $f(x_0)$ 是函数 $f(x)$ 的一个**极大值**, 称 x_0 为**极大值点**.
- 如果对于该邻域内**任意**异于 x_0 的点 x, 都恒有 $f(x)>f(x_0)$, 那么我们称 $f(x_0)$ 是函数 $f(x)$ 的一个**极小值**, 称 x_0 为**极小值点**.

极大值与极小值统称为**极值**, 极大值点与极小值点统称为**极值点**.

@余命数

极值是一个局部概念, 它只关心一个点在其邻域中的地位, 而最值是全局概念. 此定义告诉了我们什么是极值, 却未指明如何寻找它们. 直接用定义去比较无穷多个点的大小是不现实的. 直观提供了线索, 在任何光滑曲线的峰顶或谷底, 切线必然是水平的. 这一观察被精确地表述为费马引理.

引理 16.4 (费马引理) 如果函数 $f(x)$ 在点 x_0 处满足以下两个条件.

1. $f(x)$ 在点 x_0 处取得极值;
2. $f(x)$ 在点 x_0 处可导;

那么, 函数在该点的导数必为零, 即 $f'(x_0)=0$.

证明 以 $f(x_0)$ 为极大值为例. 根据极大值的定义, 存在一个以 x_0 为中心的邻域, 对于该邻域内任意异于 x_0 的点 x, 都有 $f(x)<f(x_0)$, 因而对充分小且非零的 Δx, 有 $f(x_0+\Delta x)-f(x_0)\leq 0$.

考察导数 $f'(x_0)$ 的定义式. 当自变量从右侧逼近 ($\Delta x\to 0^+$) 时, 分母 $\Delta x>0$, 分子 $f(x_0+\Delta x)-f(x_0)\leq 0$. 故比值非正, 其极限, 即右导数满足 $f'_+(x_0)\leq 0$. 当自

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

变量从左侧逼近 ($\Delta x \to 0^-$) 时, 分母 $\Delta x < 0$, 分子非正. 故比值非负, 其极限, 即左导数满足 $f'_-(x_0) \geq 0$.

由于函数在 x_0 处可导, 其左、右导数必须存在且相等, 即 $f'(x_0) = f'_+(x_0) = f'_-(x_0)$. 联立 $f'(x_0) \leq 0$ 与 $f'(x_0) \geq 0$, 唯一可能的情况便是 $f'(x_0) = 0$. □

注 16.12 (必要而不充分)
费马引理仅为必要条件.
它的真正含义是, **可导函数的极值点必定是导数为零的点.** 于是寻找极值点时, 只需在所有导数为零的点, 也称**驻点**或**稳定点**以及导数不存在的点(**奇点**)中筛选.
然而反之不成立, 即 $f'(x_0) = 0$ **并不能保证** x_0 **就是极值点.** 经典反例是 $f(x) = x^3$ 在 $x = 0$ 处, 其导数为零但函数持续递增, 该点是一个**拐点**.

费马引理联系了单个点的极值性质与导数性质. 要从点的性质推广到区间的性质, 需要中值定理. 先从其特殊形式, 即罗尔定理, 开始.

引理 16.5 (罗尔定理) 设 $a < b$. 如果函数 $f(x)$ 满足下列条件.(1) 在闭区间 $[a, b]$ 上连续; (2) 在开区间 (a, b) 内可导; (3) 在端点处的函数值相等, 即 $f(a) = f(b)$. 那么, 在 (a, b) 内至少存在一点 ξ, 使得 $f'(\xi) = 0$.

证明 根据闭区间上连续函数的最值定理, $f(x)$ 在 $[a, b]$ 上必取得最大值 M 和最小值 m. 若 $M = m$, 则 $f(x)$ 为常数函数, 其导数在 (a, b) 内处处为零, 结论显然成立. 若 $M > m$, 由于 $f(a) = f(b)$, 则 M 与 m 中至少有一个必定在区间的**内部**某点 $\xi \in (a, b)$ 处取得. 即点 ξ 是一个极值点, 且函数在该点可导. 根据费马引理, 可导函数在极值点处的导数必为零, 故 $f'(\xi) = 0$. □

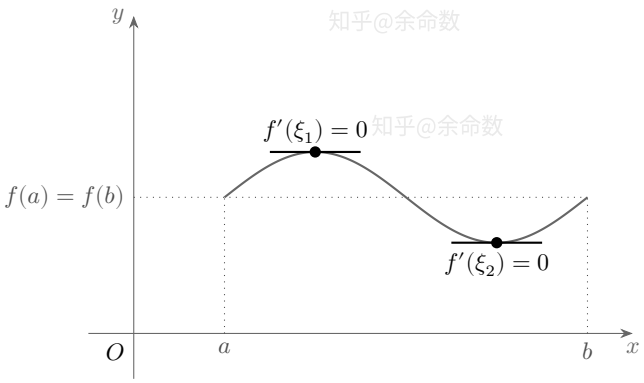


图 16.10 罗尔定理的几何意义是, 若曲线两端点等高, 则其间必有水平切线.

罗尔定理要求 $f(a) = f(b)$, 条件较为苛刻. 拉格朗日将其推广到一般情形, 思想是通过构造辅助函数, 将任意函数旋转至满足罗尔定理条件的状态.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 16.12 (拉格朗日中值定理) 设 $a<b$. 如果函数 $f(x)$ 满足下列条件.(1) 在 $[a,b]$ 上连续;(2) 在 (a,b) 内可导. 那么, 在 (a,b) 内至少存在一点 ξ, 使得

$$f'(\xi)=\frac{f(b)-f(a)}{b-a} \tag{16-14}$$

证明 证明的思路是构造一个辅助函数, 使其恰好满足罗尔定理的条件. 连接端点 $(a,f(a))$ 和 $(b,f(b))$ 的割线方程为 $y=f(a)+\frac{f(b)-f(a)}{b-a}(x-a)$. 构造辅助函数 $\varphi(x)=f(x)-\left[f(a)+\frac{f(b)-f(a)}{b-a}(x-a)\right]$. 此函数 $\varphi(x)$ 代表了曲线 $y=f(x)$ 与其端点割线之间的竖直距离. 验证 $\varphi(a)=0$ 且 $\varphi(b)=0$, 故 $\varphi(x)$ 满足罗尔定理的全部条件. 因此, 存在一点 $\xi\in(a,b)$ 使得 $\varphi'(\xi)=0$. 对 $\varphi(x)$ 求导得 $\varphi'(x)=f'(x)-\frac{f(b)-f(a)}{b-a}$. 令 $x=\xi$ 并代入 $\varphi'(\xi)=0$, 移项后即得定理结论. □

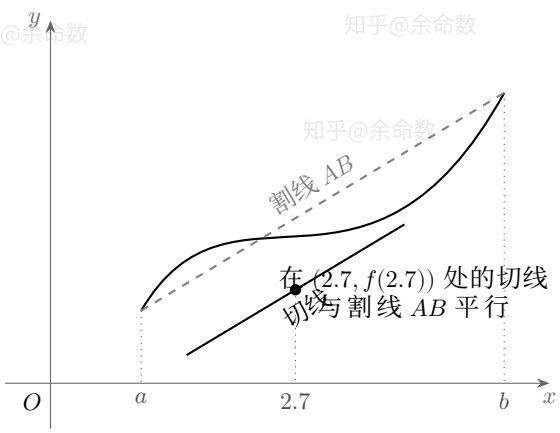


图 16.11　拉格朗日中值定理的几何意义

拉格朗日中值定理把瞬时变化率 $f'(\xi)$ 与区间平均变化率 $\frac{f(b)-f(a)}{b-a}$ 联系起来. 至此所有理论工具已准备就绪, 可以严谨地证明单调性判定法则了.

定理 16.13 (函数单调性的判定法) 设 I 是一个区间, 函数 $y=f(x)$ 在 I 上连续, 在 I 的内部可导.

- 如果在 I 的内部恒有 $f'(x)>0$, 那么函数 $y=f(x)$ 在 I 上是**单调递增**的.
- 如果在 I 的内部恒有 $f'(x)<0$, 那么函数 $y=f(x)$ 在 I 上是**单调递减**的.
- 如果在 I 的内部恒有 $f'(x)=0$, 那么函数 $y=f(x)$ 在 I 上是一个**常数函数**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 设 x_1, x_2 是区间 I 内的任意两点, 且 $x_1 < x_2$. 由于 I 是区间, 闭区间 $[x_1, x_2]$ 包含于 I; 由题设知 f 在 $[x_1, x_2]$ 上连续, 在 (x_1, x_2) 内可导. 在闭区间 $[x_1, x_2]$ 上对函数 $f(x)$ 应用拉格朗日中值定理, 可知在开区间 (x_1, x_2) 内必定存在一点 ξ, 使得

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

变形, 得核心关系式

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$$

因为假设 $x_1 < x_2$, 所以因子 $(x_2 - x_1)$ 恒为正数. 因此, 函数值的差 $f(x_2) - f(x_1)$ 的符号, 完全由导数值 $f'(\xi)$ 的符号决定.

- **若在 I 内** $f'(x) > 0$: 由于 $\xi \in (x_1, x_2) \subset I$, 所以 $f'(\xi) > 0$. 此时 $f(x_2) - f(x_1) =$ (正数) $\times$ (正数) > 0, 即 $f(x_2) > f(x_1)$. 根据定义, 函数单调递增.
- **若在 I 内** $f'(x) < 0$: 同理, 此时 $f'(\xi) < 0$. 于是 $f(x_2) - f(x_1) =$ (负数) $\times$ (正数) < 0, 即 $f(x_2) < f(x_1)$. 根据定义, 函数单调递减.
- **若在 I 内** $f'(x) = 0$: 同理, 此时 $f'(\xi) = 0$. 于是 $f(x_2) - f(x_1) = 0$, 即 $f(x_2) = f(x_1)$. 由于 x_1, x_2 是任意两点, 故函数为常数函数.

至此, 由导数符号判定单调性的法则得证. □

16.9.8.2 单调性分析的标准流程与应用

根据上述定理, 我们把求解函数单调区间的流程整理为四步.

1. **确定定义域**. 这是最基础也是最容易被忽略的一步. 后续所有讨论都必须在定义域内进行.
2. **求出导数**. 计算出导函数 $f'(x)$.
3. **求解驻点与奇点**. 求解方程 $f'(x) = 0$ 的根(驻点), 以及使 $f'(x)$ 不存在的奇点. 这些点是划分单调区间的**分界点**.
4. **列表分析**. 用表格清晰地展示出定义域被分界点划分成的各个子区间, 以及每个子区间上 $f'(x)$ 的符号和 $f(x)$ 的单调性.

例题 16.56 求函数 $f(x) = 2x^3 - 9x^2 + 12x - 3$ 的单调区间.

解 **定义域** 该函数为多项式函数, 定义域为 $(-\infty, +\infty)$.

求导数

$$f'(x) = 6x^2 - 18x + 12 = 6(x^2 - 3x + 2) = 6(x - 1)(x - 2)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

求分界点, 令 $f'(x)=0$, 解得 $x=1$ 或 $x=2$. 这两个驻点将定义域 $(-\infty,+\infty)$ 划分为三个区间.

列表分析

区间	$(-\infty,1)$	1	$(1,2)$	2	$(2,+\infty)$
$f'(x)$ 的符号	$+$	0	$-$	0	$+$
$f(x)$ 的单调性	↗(增)	极大值	↘(减)	极小值	↗(增)

因此, 函数 $f(x)$ 的单调递增区间为 $(-\infty,1)$ 和 $(2,+\infty)$, 单调递减区间为 $(1,2)$.

例题 16.57　求函数 $f(x)=\dfrac{\ln x}{x}$ 的单调区间.

解 **定义域**函数 $f(x)$ 的定义域为 $(0,+\infty)$.

求导数利用除法法则, 得

$$f'(x)=\frac{(\ln x)'\cdot x-\ln x\cdot(x)'}{x^2}=\frac{\frac{1}{x}\cdot x-\ln x\cdot 1}{x^2}=\frac{1-\ln x}{x^2}$$

求分界点, 由于 $x\in(0,+\infty)$, 分母 x^2 恒为正, 所以 $f'(x)$ 的符号完全由分子 $1-\ln x$ 决定. 令 $f'(x)=0$, 得 $1-\ln x=0$, 解得 $x=e$.

列表分析

区间	$(0,e)$	e	$(e,+\infty)$
$f'(x)$ 的符号	$+$	0	$-$
$f(x)$ 的单调性	↗(增)	极大值	↘(减)

因此, 函数 $f(x)$ 的单调递增区间为 $(0,e)$, 单调递减区间为 $(e,+\infty)$.

例题 16.58　已知函数 $f(x)=x\ln x$.

1. 讨论函数 $g(x)=f(x)-x+1$ 的单调性, 并证明对任意 $x>0$, 恒有 $x\ln x\geq x-1$.
2. 证明存在唯一的 $x_0\in(1,e)$, 使得曲线 $y=x\ln x$ 在点 x_0 处的切线, 与连接点 $(1,0)$ 和 (e,e) 的直线平行.

解　第 (1) 问把不等式 $x\ln x\geq x-1$ 转化为 $g(x)=x\ln x-x+1\geq 0$ 的非负性, 通过求导定其最小值即可; 第 (2) 问要求存在唯一的 $x_0\in(1,e)$ 使曲线在 x_0 处的切线与连接 $(1,0)$、(e,e) 的直线平行, 等价于方程 $f'(x_0)=\dfrac{f(e)-f(1)}{e-1}$ 在 $(1,e)$ 内有

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

唯一解, 这正是拉格朗日中值定理保证存在、由导函数单调性保证唯一性的典型情形.

(1) 函数 $g(x)=x\ln x-x+1$ 的定义域为 $(0,+\infty)$. 对其求导得 $g'(x)=(\ln x+1)-1=\ln x$. 令 $g'(x)=0$ 解得 $x=1$. 当 $x\in(0,1)$ 时 $g'(x)=\ln x<0$, $g(x)$ 单调递减; 当 $x\in(1,+\infty)$ 时 $g'(x)=\ln x>0$, $g(x)$ 单调递增. 故 $g(x)$ 在 $x=1$ 处取得极小值, 亦为全局最小值, $g(1)=1\ln 1-1+1=0$. 由最小值为 0, 对任意 $x>0$ 有 $g(x)\geq 0$, 即 $x\ln x-x+1\geq 0$, 移项得 $x\ln x\geq x-1$.

(2) 如上分析, 本问等价于证明方程 $\ln x+1=\frac{e}{e-1}$ 在 $(1,e)$ 内有唯一解. 构造辅助函数 $H(x)=\ln x+1-\frac{e}{e-1}$, 证明其在 $(1,e)$ 内有唯一零点.

先证**存在性**. $H(x)$ 在闭区间 $[1,e]$ 上连续, 其端点值为 $H(1)=\ln 1+1-\frac{e}{e-1}=-\frac{1}{e-1}<0$, $H(e)=\ln e+1-\frac{e}{e-1}=\frac{e-2}{e-1}>0$. 因 $H(x)$ 在 $[1,e]$ 两端异号, 由零点存在性定理, 在 $(1,e)$ 内至少存在一点 x_0 使 $H(x_0)=0$.

再证**唯一性**. 对 $H(x)$ 求导得 $H'(x)=\frac{1}{x}$, 对任意 $x\in(1,e)$ 有 $H'(x)>0$, 故 $H(x)$ 在 $(1,e)$ 上严格单调递增, 其图像与 x 轴至多有一个交点. 结合已证的存在性, $H(x)$ 在 $(1,e)$ 内有且仅有一个零点. 证毕.

16.10 中值定理的应用与推广

前一节建立了费马引理、罗尔定理及拉格朗日中值定理. 这些定理构成微分学的主要理论框架, 并把导数符号与函数单调性联系起来. 本节从定性判断转向定量分析. 我们利用中值定理建立函数增量与导数取值范围的联系, 证明不等式; 估计线性近似的误差; 并把拉格朗日中值定理推广为处理两函数比值的柯西中值定理.

16.10.1 利用中值定理证明不等式

我们先写出拉格朗日中值定理的等式形式.

$$f(b)-f(a)=f'(\xi)(b-a),\quad \xi\in(a,b)$$

该等式给出函数在区间两端的增量 $\Delta y=f(b)-f(a)$ 与导数在中间点的值 $f'(\xi)$、自变量增量 $\Delta x=b-a$ 之间的精确关系.

ξ 的具体位置未知, 但由 $\xi\in(a,b)$ 可估计 $f'(\xi)$ 的取值范围, 从而得到函数增量的上界或下界. 这样便由局部变化率推出了整体改变量.

16.10.1.1 方法与实例

证明 $g(x)>h(x)$ 这类不等式, 可转化为证明 $g(x)-h(x)>0$. 把 x 看作区间的一个端点, 构造函数并在相应区间上应用中值定理.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 16.59　证明当 $x>0$ 时, 不等式 $e^x>1+x$ 成立.

证明　不等式可变形为 $e^x-1>x$. 考虑函数 $f(t)=e^t$. 该函数在闭区间 $[0,x]$ 上连续, 在 $(0,x)$ 内可导. 根据拉格朗日中值定理, 存在 $\xi\in(0,x)$, 使得

$$f(x)-f(0)=f'(\xi)(x-0)$$

知乎@余命数

代入函数具体的表达式, 得

$$e^x-1=e^{\xi}\cdot x$$

由于 $\xi>0$, 利用指数函数的单调性可知 $e^{\xi}>e^0=1$. 又因为 $x>0$, 所以 $e^{\xi}\cdot x>1\cdot x=x$. 即 $e^x-1>x$, 亦即 $e^x>1+x$. □

对于包含对数或分式的不等式, 直接比较往往较难, 而中值定理可将其化为倒数或线性函数的比较.

例题 16.60　证明对于任意 $a>b>0$, 有

$$\frac{a-b}{a}<\ln\frac{a}{b}<\frac{a-b}{b}$$

知乎@余命数

证明　首先观察中间项 $\ln\frac{a}{b}=\ln a-\ln b$, 这提示我们利用对数函数的增量公式. 构造函数 $f(x)=\ln x$. 函数 $f(x)$ 在区间 $[b,a]$ 上满足拉格朗日中值定理的条件. 于是存在 $\xi\in(b,a)$, 使得

$$\ln a-\ln b=f'(\xi)(a-b)=\frac{1}{\xi}(a-b)$$

即

$$\ln\frac{a}{b}=\frac{a-b}{\xi}$$

知乎@余命数　知乎@余命数　知乎@余命数

由于 $b<\xi<a$, 且函数 $y=\frac{1}{x}$ 在 $(0,+\infty)$ 上单调递减, 故有

$$\frac{1}{a}<\frac{1}{\xi}<\frac{1}{b}$$

知乎@余命数

因为 $a>b$, 所以 $a-b>0$. 在上述不等式各项同时乘以 $(a-b)$, 得

$$\frac{a-b}{a}<\frac{a-b}{\xi}<\frac{a-b}{b}$$

代入 $\ln\frac{a}{b}=\frac{a-b}{\xi}$, 即证得结论. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.10.2 中值定理与近似计算的误差分析

工程与科学计算中常用切线近似曲线, 我们称这种近似为线性近似. 设函数 $f(x)$ 在点 x_0 处可导, 则 x 在 x_0 附近的**线性近似**(或一阶泰勒近似)定义为

$$L(x) = f(x_0) + f'(x_0)(x - x_0)$$

知乎@余命数

几何上, 这是用切线 $y = L(x)$ 代替曲线 $y = f(x)$. 我们把近似带来的**误差**记为

$$R(x) = f(x) - L(x) = f(x) - f(x_0) - f'(x_0)(x - x_0)$$

当 $x \to x_0$ 时, 误差 $R(x)$ 是比 $(x - x_0)$ 高阶的无穷小. 应用中值定理可进一步给出 $R(x)$ 的具体表达式, 从而量化近似的精确度.

定理 16.14 (线性近似的误差估计) 设函数 $f(x)$ 在包含 x_0 的区间上具有二阶导数. 则对于区间内任意一点 x, 当 $x \neq x_0$ 时, 存在介于 x_0 与 x 之间的 ξ, 使得

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(\xi)(x - x_0)^2 \tag{16-15}$$

当 $x = x_0$ 时, 上式两边均等于 $f(x_0)$, 结论直接成立. 即误差项为 $R(x) = \frac{1}{2}f''(\xi)(x - x_0)^2$.

证明 若 $x = x_0$, 公式直接成立. 以下设 $x \neq x_0$. 我们引入参数 K, 使得等式在 x 处成立.

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + K(x - x_0)^2$$

令 $\varphi(t)$ 为函数值与近似值在点 t 处的差.

$$\varphi(t) = f(t) - [f(x_0) + f'(x_0)(t - x_0) + K(t - x_0)^2]$$

直接有 $\varphi(x_0) = 0$. 由 K 的定义可知 $\varphi(x) = 0$. 根据罗尔定理, 在 x_0 与 x 之间存在一点 η, 使得 $\varphi'(\eta) = 0$. 对 $\varphi(t)$ 求导得 $\varphi'(t) = f'(t) - [f'(x_0) + 2K(t - x_0)]$, 注意 $\varphi'(x_0) = f'(x_0) - f'(x_0) = 0$. 现有 $\varphi'(\eta) = 0$ 和 $\varphi'(x_0) = 0$. 再次对 $\varphi'(t)$ 在区间 $[x_0, \eta]$ (或 $[\eta, x_0]$) 上应用罗尔定理, 可知存在 ξ 介于 x_0 与 η 之间, 也就是同时介于 x_0 与 x 之间, 使得 $\varphi''(\xi) = 0$. 计算 φ 的二阶导数 $\varphi''(t) = f''(t) - 2K$, 代入 $\varphi''(\xi) = 0$ 得 $f''(\xi) - 2K = 0$, 即 $K = \frac{1}{2}f''(\xi)$. 将 K 代回原式, 即得证. □

知乎@余命数

该结论表明, 线性近似的误差大小取决于函数二阶导数 $f''(x)$ 的界. 二阶导数越大, 曲线弯曲越甚, 切线近似的误差也越大.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.10.3 柯西中值定理

拉格朗日中值定理建立了单个函数增量与导数的关系. 考察两个函数 $f(x)$ 和 $g(x)$ 的增量比值, 就得到柯西中值定理. 它是拉格朗日中值定理的推广, 也是推导洛必达法则的工具.

定理 16.15 (柯西中值定理) 设 $a<b$. 如果函数 $f(x)$ 和 $g(x)$ 满足下列条件.

1. 在闭区间 $[a,b]$ 上连续;
2. 在开区间 (a,b) 内可导;
3. 对任意 $x\in(a,b)$, $g'(x)\neq 0$.

那么, 在 (a,b) 内至少存在一点 ξ, 使得

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)} \tag{16-16}$$

知乎@余命数

注 16.13
我们对此定理作三点说明. 1. 若取 $g(x)=x$, 则 $g'(x)=1$, 定理退化为拉格朗日中值定理. 2. 条件 $g'(x)\neq 0$ 保证了分母 $g(b)-g(a)\neq 0$(否则由罗尔定理知存在一点导数为零). 3. 该式不能简单地通过对 f 和 g 分别应用拉格朗日定理得到, 因为那样会得到两个不同的中间点 ξ_1 和 ξ_2, 分别属于 f 和 g, 因此不能直接把两个等式相除.

证明 我们仍用罗尔定理证明, 关键是构造端点相等的辅助函数. 将结论变形为 $[f(b)-f(a)]g'(\xi)-[g(b)-g(a)]f'(\xi)=0$. 构造辅助函数 $\Phi(x)$, 得

$$\Phi(x)=[f(b)-f(a)]g(x)-[g(b)-g(a)]f(x)$$

其端点值为

$$\Phi(a)=f(b)g(a)-f(a)g(a)-g(b)f(a)+g(a)f(a)=f(b)g(a)-g(b)f(a)$$

$$\Phi(b)=f(b)g(b)-f(a)g(b)-g(b)f(b)+g(a)f(b)=f(b)g(a)-g(b)f(a)$$

可见 $\Phi(a)=\Phi(b)$. 由 f,g 的性质知 $\Phi(x)$ 在 $[a,b]$ 上连续, 在 (a,b) 内可导. 根据罗尔定理, 存在 $\xi\in(a,b)$ 使得 $\Phi'(\xi)=0$. 求导得 $\Phi'(\xi)=[f(b)-f(a)]g'(\xi)-[g(b)-g(a)]f'(\xi)=0$. 由条件 $g'(x)\neq 0$ 与罗尔定理可知 $g(b)-g(a)\neq 0$; 同时 $g'(\xi)\neq 0$. 因而可以相除. 移项并整理, 即得

$$\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(\xi)}{g'(\xi)}$$

□

知乎@余命数

知乎@余命数

16.10.3.1 几何意义

柯西中值定理有几何解释. 我们用参数方程表示这条曲线.

$$\begin{cases} x = g(t) \\ y = f(t) \end{cases}, \quad t \in [a, b]$$

此时, 比值 $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{\Delta y}{\Delta x}$ 代表连接曲线起点 $A(g(a), f(a))$ 和终点 $B(g(b), f(b))$ 的**弦的斜率**. 而比值 $\frac{f'(\xi)}{g'(\xi)} = \frac{dy/dt}{dx/dt}\Big|_{t=\xi} = \frac{dy}{dx}\Big|_{t=\xi}$ 代表曲线在参数 $t = \xi$ 对应点处的**切线斜率**.

定理表明, 对一条光滑的参数曲线, 在起点和终点之间, 必有一点, 该点处的切线与连接两端点的弦平行.

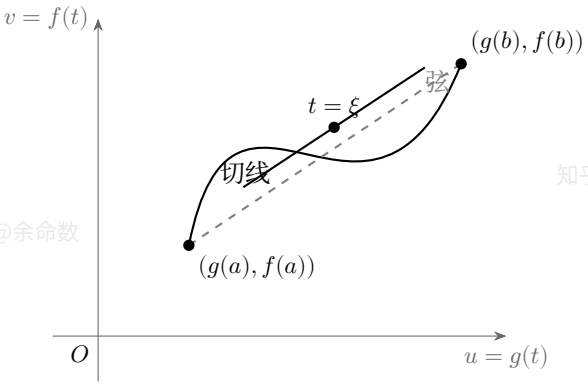


图 16.12 柯西中值定理的几何意义是, 参数曲线上的切线平行于弦.

当 $f(a) = g(a) = 0$ 时, $\frac{f(x)}{g(x)} = \frac{f(x)-f(a)}{g(x)-g(a)} = \frac{f'(\xi)}{g'(\xi)}$. 令 $x \to a$, 则 $\xi \to a$, 于是函数之比便过渡到导数之比. 这正是洛必达法则的雏形.

16.11 导数的综合应用

我们已能用导数判断函数的单调性与极值. 本节把它们用来处理三类问题,即方程根的分布、函数零点个数、以及含参不等式的恒成立. 这三类问题的共同之处是都需要把代数命题翻译成函数图像的几何特征, 再用导数进行计算.

讨论零点就是研究函数图像与 x 轴的交点. 一条曲线怎样穿越或接触 x 轴,取决于它的起伏与位置. 我们用导数确定起伏, 再用端点值与极值确定位置.

16.11.1 函数零点问题

函数 f 的零点就是方程 $f(x) = 0$ 的实根. 如果解析式里没有参数, 我们只研究一条固定的曲线; 如果含有参数, 曲线会随参数移动, 零点个数也会改变. 本节分别处理这两种情形.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

16.11.1.1 基本判定准则

零点的存在性由介值定理保证. 如果连续函数 $f(x)$ 在区间 $[a,b]$ 端点取值异号, 即 $f(a)f(b)<0$, 那么 (a,b) 内至少存在一个零点. 导数进一步约束了零点的个数.

定理 16.16 (单调性与零点) 若 $f(x)$ 在区间 I 内严格单调, 则 $f(x)$ 在 I 内至多有一个零点.

证明 我们用反证法. 假设 f 在 I 内存在两个不同零点 x_1,x_2, 不妨设 $x_1<x_2$. 若 f 严格递增, 则 $f(x_1)<f(x_2)$; 若 f 严格递减, 则 $f(x_1)>f(x_2)$. 但 x_1,x_2 都是零点, 所以 $f(x_1)=f(x_2)=0$, 这与严格单调性矛盾. 因此零点至多一个. □

对于非单调函数, 我们先把定义域分成若干单调区间, 再分别考察每个区间端点处的函数值与极值. 极值的符号决定曲线在该区间是否穿过 x 轴.

16.11.1.2 定系数函数的零点

解析式确定后, 零点个数由极值符号与区间端点函数值共同决定. 我们按求导、找极值、看边界的顺序操作.

例题 16.61 确定函数 $f(x)=x^3-3x+1$ 的零点个数.

解 函数的定义域为 $\mathbb{R}$. 求导得 $f'(x)=3x^2-3=3(x-1)(x+1)$. 令 $f'(x)=0$, 解得驻点 $x_1=-1,x_2=1$.

导数符号如下.

- 当 $x\in(-\infty,-1)$ 时, $f'(x)>0$, f 单调递增;
- 当 $x\in(-1,1)$ 时, $f'(x)<0$, f 单调递减;
- 当 $x\in(1,+\infty)$ 时, $f'(x)>0$, f 单调递增.

极大值为 $f(-1)=-1+3+1=3>0$. 极小值为 $f(1)=1-3+1=-1<0$.

再考察边界趋势. $\lim_{x\to-\infty}f(x)=-\infty$, $\lim_{x\to+\infty}f(x)=+\infty$.

由 $\lim\limits_{x\to-\infty}f(x)<0$ 且 $f(-1)>0$, 结合该区间上的单调性, 在 $(-\infty,-1)$ 存在唯一零点; 由 $f(-1)>0>f(1)$, 在 $(-1,1)$ 存在唯一零点; 由 $f(1)<0$ 且 $\lim\limits_{x\to+\infty}f(x)>0$, 结合该区间上的单调性, 在 $(1,+\infty)$ 存在唯一零点. 故该函数共有 3 个零点.

例题 16.62 (2024· 重庆南开中学)设函数 $f(x)=\sin x-x\cos x$.

1. 当 $x\in[-\pi,\pi]$ 时, 求 $f(x)$ 的最值;

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

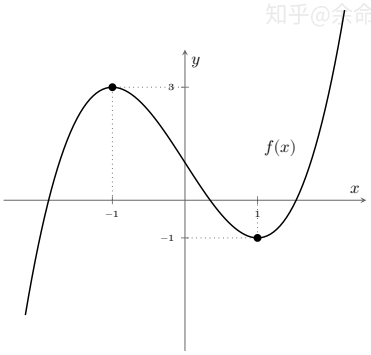


图 16.13 利用极大值与极小值的符号判定三次函数的零点个数

2. 讨论函数 $g(x)=\ln(x+1)-f'(x)$ 在区间 $(-1,2\pi)$ 内的零点个数.

解 **(1) 单调性分析**对 $f(x)$ 求导, 得

$$f'(x)=\cos x-(\cos x-x\sin x)=x\sin x.$$

当 $x\in[-\pi,\pi]$ 时,

- 若 $x\in[-\pi,0)$, 则 $x<0,\sin x<0$, 于是 $f'(x)>0$;
- 若 $x\in(0,\pi]$, 则 $x>0,\sin x>0$, 于是 $f'(x)>0$.

故 $f(x)$ 在 $[-\pi,\pi]$ 上单调递增. 最小值为 $f(-\pi)=\sin(-\pi)-(-\pi)\cos(-\pi)=-\pi$. 最大值为 $f(\pi)=\sin\pi-\pi\cos\pi=\pi$.

(2) 分段考察零点函数解析式为 $g(x)=\ln(x+1)-x\sin x$, 定义域为 $(-1,2\pi)$. 由于含三角函数, 我们按它的符号周期性分区间讨论.

区间 $(-1,0)$. 当 $x\in(-1,0)$ 时, $0<x+1<1$, 故 $\ln(x+1)<0$. 同时 $x<0$, $\sin x<0$, 于是 $x\sin x>0$. 因此 $g(x)=$ 负 $-$ 正 <0, 该区间无零点.

点 $x=0$. $g(0)=\ln 1-0=0$, 故 $x=0$ 是一个零点.

区间 $(\pi,2\pi)$. 当 $x\in(\pi,2\pi)$ 时, $\sin x<0$, 且 $x>0$, 故 $-x\sin x>0$. 又 $\ln(x+1)>\ln(\pi+1)>0$. 两项均为正, 故 $g(x)>0$, 该区间无零点.

区间 $(0,\frac{\pi}{2})$. 将方程 $g(x)=0$ 改写为

$$\frac{\ln(1+x)}{x}=\sin x\quad(x>0).$$

令 $H(x)=\frac{\ln(1+x)}{x}-\sin x$. 则

$$H'(x)=\frac{\frac{x}{1+x}-\ln(1+x)}{x^2}-\cos x<0,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其中 $\ln(1+x) > \frac{x}{1+x}$，且在 $(0, \frac{\pi}{2})$ 上 $\cos x > 0$. 因此 H 在该区间严格递减. 又 $\lim\limits_{x\to 0^+} H(x) = 1 > 0$, 且

$$H\left(\frac{\pi}{2}\right) = \frac{\ln(1+\frac{\pi}{2})}{\frac{\pi}{2}} - 1 < 0,$$

故 $(0, \frac{\pi}{2})$ 内恰有一个零点.

区间 $(\frac{\pi}{2}, \pi)$. 先看端点符号.

$$g\left(\frac{\pi}{2}\right) = \ln\left(1+\frac{\pi}{2}\right) - \frac{\pi}{2} < 0, \quad g(\pi) = \ln(1+\pi) > 0.$$

因此至少存在一个零点. 为证明唯一性, 考察

$$g'(x) = \frac{1}{x+1} - \sin x - x\cos x, \quad g''(x) = -\frac{1}{(x+1)^2} - 2\cos x + x\sin x.$$

当 $x \in (\frac{\pi}{2}, \frac{2\pi}{3}]$ 时, $x\sin x > 1 > \frac{1}{(x+1)^2}$, 且 $-2\cos x \geq 0$; 当 $x \in [\frac{2\pi}{3}, \pi)$ 时, $-2\cos x \geq 1 > \frac{1}{(x+1)^2}$, 且 $x\sin x > 0$. 故 $g''(x) > 0$ 在 $(\frac{\pi}{2}, \pi)$ 上恒成立. 于是 $g'(x)$ 严格递增. 又

$$g'\left(\frac{\pi}{2}\right) = \frac{1}{1+\frac{\pi}{2}} - 1 < 0, \quad g'(\pi) = \frac{1}{1+\pi} + \pi > 0,$$

故 g' 在该区间仅有一个零点, 从而 g 先减后增. 结合 $g(\frac{\pi}{2}) < 0$ 与 $g(\pi) > 0$, 可知 $(\frac{\pi}{2}, \pi)$ 内恰有一个零点.

综合结论 $(0, \frac{\pi}{2})$ 与 $(\frac{\pi}{2}, \pi)$ 内各有一个零点, $x = 0$ 也是零点; 而 $(-1, 0)$ 与 $(\pi, 2\pi)$ 内无零点.

$g(x)$ 在 $(-1, 2\pi)$ 上共有 3 个零点(分别为 $0, x_1, x_2$).

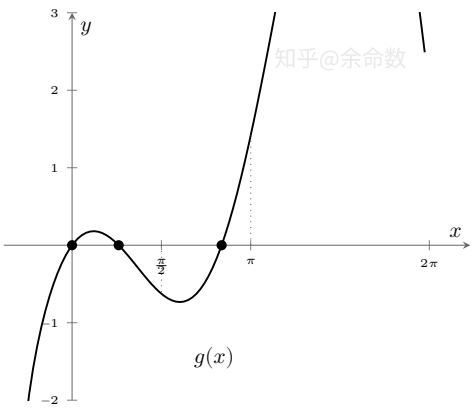


图 16.14　超越函数 $g(x)$ 的零点分布示意图

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.11.1.3 含参方程的零点

当方程含有参数时, 零点个数随参数变化而改变. 此时需对参数进行分类讨论. 常用方法有两种.

1. **直接分析法**. 研究含参函数 $f(x,a)$ 的极值点位置及极值大小与参数的关系.
2. **参变分离法**. 将方程变形为 $g(x)=a$, 转化为固定曲线 $y=g(x)$ 与水平直线 $y=a$ 的交点问题.

下例采用直接分析法演示.

例题 16.63 讨论关于 x 的方程 $x-a\ln x=0\ (a>0)$ 的实根个数.

解 **构造与求导**令 $f(x)=x-a\ln x$, 定义域为 $(0,+\infty)$. 求导得 $f'(x)=1-\frac{a}{x}=\frac{x-a}{x}$.

单调性与极值由于 $x>0$, $a>0$, 导数符号仅取决于 $x-a$. 当 $x\in(0,a)$ 时, $f'(x)<0$, f 递减; 当 $x\in(a,+\infty)$ 时, $f'(x)>0$, f 递增. 故 $f(x)$ 在 $x=a$ 处取得极小值, 亦为最小值

$$f(a)=a-a\ln a=a(1-\ln a).$$

边界趋势

$$\lim_{x\to0^+}(x-a\ln x)=+\infty,\quad \lim_{x\to+\infty}x(1-a\frac{\ln x}{x})=+\infty.$$

函数图像呈 U 型, 两端趋于正无穷.

分类讨论零点个数取决于最小值 $f(a)$ 与 0 的关系.

1. 若 $f(a)>0$, 即 $a(1-\ln a)>0$, 则 $0<a<e$. 图像整体位于 x 轴上方, **无实根**.
2. 若 $f(a)=0$, 即 $a=e$, 图像与 x 轴相切于顶点, **有唯一实根**.
3. 若 $f(a)<0$, 即 $a>e$, 极小值小于 0 且两端趋于 $+\infty$, 分别在 $(0,a)$ 与 $(a,+\infty)$ 各有**一个实根**, 共两个.

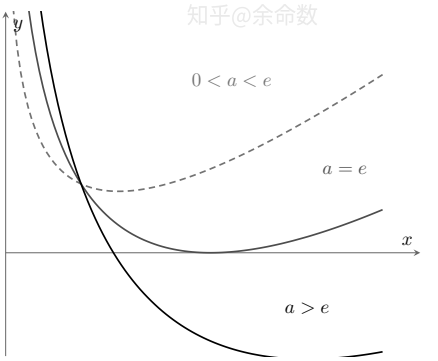


图 16.15 参数 a 改变了 $f(x)$ 极小值的位置与大小, 从而改变零点个数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 16.64　(2024· 成都七中三诊)已知函数 $f(x)=e^x-ax\sin x-x-1$, 其中 $x\in(0,\pi)$.

1. 若 $a=\frac{1}{2}$, 证明 $f(x)>0$;
2. 若函数 $f(x)$ 在 $(0,\pi)$ 内有唯一零点, 求实数 a 的取值范围.

解　**(1) 二阶导数定性**当 $a=\frac{1}{2}$ 时, $f(x)=e^x-\frac{1}{2}x\sin x-x-1$. 求导得

$$f'(x)=e^x-\frac{1}{2}(\sin x+x\cos x)-1.$$

由于直接判断符号困难, 我们考察 $f'(x)$ 的导数, 即原函数二阶导.

$$f''(x)=e^x-\frac{1}{2}(2\cos x-x\sin x)=e^x-\cos x+\frac{1}{2}x\sin x.$$

当 $x\in(0,\pi)$ 时, $e^x>1\geqslant\cos x$, 且 $x\sin x>0$, 故 $f''(x)>0$. 故 $f'(x)$ 在 $(0,\pi)$ 上单调递增.

又 $f'(0)=e^0-0-1=0$, 故当 $x>0$ 时, $f'(x)>f'(0)=0$. 进而 $f(x)$ 在 $(0,\pi)$ 上单调递增. 又 $f(0)=1-0-0-1=0$, 故对于 $x\in(0,\pi)$, 恒有 $f(x)>f(0)=0$.

(2) 分类与比较由题意, 我们讨论方程 $f(x)=0$ 在 $(0,\pi)$ 上的根.

情形一 ($a\leqslant\frac{1}{2}$). 利用第 (1) 问结论进行缩放.

$$f(x)=\left(e^x-\frac{1}{2}x\sin x-x-1\right)+\left(\frac{1}{2}-a\right)x\sin x.$$

由 (1) 知第一部分大于 0; 由 $a\leqslant\frac{1}{2}$ 及 $x\in(0,\pi)$ 知 $(\frac{1}{2}-a)x\sin x\geqslant 0$. 故 $f(x)>0$, 函数无零点.

情形二 ($a>\frac{1}{2}$). 由于 $x\sin x>0$ 在 $(0,\pi)$ 内恒成立, 方程 $f(x)=0$ 等价于

$$a=\Phi(x):=\frac{e^x-x-1}{x\sin x}.$$

现证 $\Phi(x)$ 在 $(0,\pi)$ 上严格递增, 且值域为 $(\frac{1}{2},+\infty)$.

将其写成

$$\Phi(x)=\frac{\frac{e^x-x-1}{x^2}}{\frac{\sin x}{x}}.$$

记 $A(x)=\frac{e^x-x-1}{x^2}$. 则

$$A'(x)=\frac{(x-2)e^x+x+2}{x^3}.$$

令 $q(x)=(x-2)e^x+x+2$, 则 $q(0)=0$、$q'(x)=(x-1)e^x+1$、$q''(x)=xe^x>0$. 因而 $q'(x)>0$, $q(x)>0$ 在 $x>0$ 上成立, 所以 $A(x)$ 严格递增.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

又 $B(x)=\frac{\sin x}{x}$ 在 $(0,\pi)$ 上为正且严格递减, 因为

$$B'(x)=\frac{x\cos x-\sin x}{x^2}<0,$$

其中 $\sin x-x\cos x$ 的导数为 $x\sin x>0$, 且在 $x=0$ 处极限为 0. 故 $1/B(x)$ 严格递增. 于是正函数 $A(x)$ 与 $1/B(x)$ 的乘积 $\Phi(x)$ 严格递增.

再由

$$\lim_{x\to 0^+}\Phi(x)=\frac{1}{2},\qquad \lim_{x\to\pi^-}\Phi(x)=+\infty,$$

可知当且仅当 $a>\frac{1}{2}$ 时, 方程 $a=\Phi(x)$ 在 $(0,\pi)$ 内有唯一解, 即 $f(x)$ 在 $(0,\pi)$ 内有唯一零点.

故满足题意的 a 的取值范围是 $\left(\frac{1}{2},+\infty\right)$.

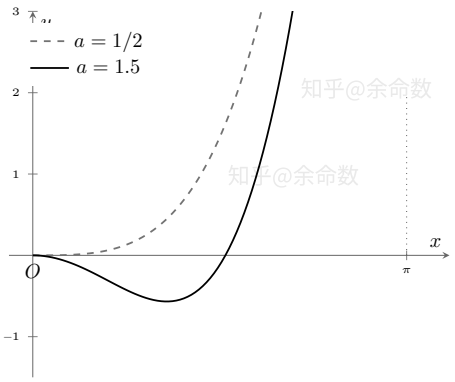


图 16.16 参数 a 改变了函数在 0 附近的凹凸性与单调性

例题 16.65 (2022· 全国乙卷)已知函数 $f(x)=\ln(1+x)+axe^{-x}$.

1. 当 $a=1$ 时, 求曲线 $y=f(x)$ 在点 $(0,f(0))$ 处的切线方程;
2. 若 $f(x)$ 在区间 $(-1,0)$ 和 $(0,+\infty)$ 各恰有一个零点, 求 a 的取值范围.

解 **(1) 切线求解**当 $a=1$ 时, $f(x)=\ln(1+x)+xe^{-x}$. $f(0)=\ln 1+0=0$. 切点为 $(0,0)$. 求导得

$$f'(x)=\frac{1}{1+x}+e^{-x}-xe^{-x}=\frac{1}{1+x}+e^{-x}(1-x).$$

故切线斜率 $k=f'(0)=1+1(1)=2$. 切线方程为 $y-0=2(x-0)$, 即 $y=2x$.

(2) 转化与分类函数定义域为 $(-1,+\infty)$. 求导并通分, 得

$$f'(x)=\frac{1}{1+x}+ae^{-x}(1-x)=\frac{e^x+a(1-x^2)}{e^x(1+x)}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设辅助函数 $h(x)=e^x+a(1-x^2)$. 当 $x>-1$ 时, 分母恒正, 故 $f'(x)$ 与 $h(x)$ 同号.

情形一 ($a\geq 0$). 当 $x\in(-1,0)$ 时, $e^x>0$ 且 $1-x^2>0$. 若 $a\geq 0$, 则 $h(x)>0$ 恒成立. 此时 $f(x)$ 在 $(-1,0)$ 单调递增. 又 $f(0)=0$, 故当 $x\in(-1,0)$ 时, $f(x)<f(0)=0$, 不存在零点. 不合题意.

情形二 ($-1\leq a<0$). 考察 $h(x)$ 的导数, 得 $h'(x)=e^x-2ax$. 当 $x\in(0,+\infty)$ 时, $e^x>1$, 且 $-2ax\geq 0$ (因为 a 为负、x 为正), 故 $h'(x)>0$. $h(x)$ 单调递增, 且 $h(0)=1+a\geq 0$. 故在 $(0,+\infty)$ 上 $h(x)>0$, 从而 $f'(x)>0$. $f(x)$ 单调递增, 且 $f(0)=0$, 故 $f(x)>0$, 无零点. 不合题意.

情形三 ($a<-1$). 此时 $h(0)=1+a<0$.

1. 区间 $(-1,0)$ 的分析. $x\to -1^+$ 时, $h(x)\to e^{-1}>0$. 结合 $h(0)<0$, 存在 $x_1\in(-1,0)$ 使得 $h(x_1)=0$. 为说明该零点唯一, 我们注意 $h''(x)=e^x-2a>0$, 故 $h'(x)$ 严格递增; 又 $h'(-1^+)=e^{-1}+2a<0$, $h'(0)=1>0$, 所以 h 先减后增. 由于 $h(0)<0$, 增段上的函数值仍小于 0, 因而 h 在 $(-1,0)$ 内只与 x 轴相交一次. 当 $x\in(-1,x_1)$ 时 $h(x)>0$, f 递增; 当 $x\in(x_1,0)$ 时 $h(x)<0$, f 递减. $f(x_1)$ 为极大值. 我们注意 $\lim_{x\to -1^+}f(x)=-\infty$, 且 $f(0)=0$. 函数图像先由 $-\infty$ 升至 $f(x_1)$, 再降至 0. 于是必然有 $f(x_1)>0$, 且在 $(-1,x_1)$ 内存在唯一零点.

2. 区间 $(0,+\infty)$ 的分析. $h'(x)=e^x-2ax$. 因 $a<-1$, 故 $-2a>2$, 所以 $h'(x)>0$, $h(x)$ 单调递增. $h(0)<0$, 且 $x\to+\infty$ 时 $h(x)\to+\infty$. 故存在唯一 $x_2\in(0,+\infty)$ 使得 $h(x_2)=0$. 当 $x\in(0,x_2)$ 时 $h(x)<0$, f 递减; 当 $x\in(x_2,+\infty)$ 时 $h(x)>0$, f 递增. $f(x_2)$ 为极小值. 由于 $f(0)=0$, 函数从 0 开始下降至 $f(x_2)$, 再上升至 $+\infty$. 显然 $f(x_2)<0$, 故在 $(x_2,+\infty)$ 内存在唯一零点.

当 $a<-1$ 时, 满足题设要求.

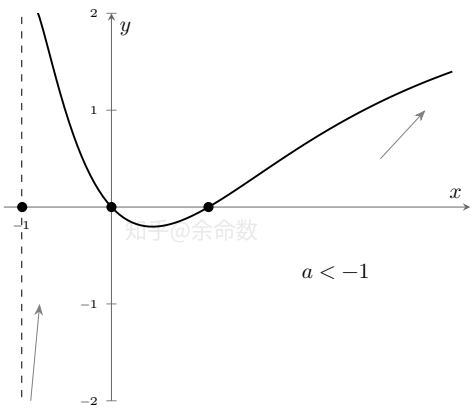


图 16.17　$a<-1$ 时函数 $f(x)$ 呈现出的双谷形态与零点分布

16.11.2 不等式恒成立问题

证明 “对于区间 I 内的任意 x, 不等式 $f(x)\geq g(x)$ 恒成立”, 即全称命题 $\forall x\in I, f(x)\geq g(x)$. 此类问题通常用最值法或参数分离法处理.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.11.2.1 最值法

最值法将不等式变形为 $h(x)=f(x)-g(x)\geq 0$. 要保证处处非负, 只需保证最低点非负, 即转化为求 $h(x)$ 在 I 上的最小值并验证 $h(x)_{\min}\geq 0$.

例题 16.66 求实数 a 的取值范围, 使得 $e^x\geq ax+1$ 对所有 $x\in\mathbb{R}$ 恒成立.

证明 构造函数 $h(x)=e^x-ax-1$. 问题转化为求 a 使得 $h(x)_{\min}\geq 0$. 求导得 $h'(x)=e^x-a$.

情形 1 设 $a\leq 0$. 此时 $h'(x)=e^x-a>0$ 恒成立, 函数 $h(x)$ 在 $\mathbb{R}$ 上单调递增. 当 $a=0$ 时, $x\to-\infty$ 有 $h(x)\to-1<0$; 当 $a<0$ 时, $x\to-\infty$ 有 $h(x)=e^x-ax-1\to-\infty$. 这说明必然存在 x 使得不等式不成立. 故 $a\leq 0$ 不合题意.

情形 2 设 $a>0$. 令 $h'(x)=0$, 解得唯一的驻点 $x_0=\ln a$.

- 当 $x<\ln a$ 时, $h'(x)<0$, h 单调递减;
- 当 $x>\ln a$ 时, $h'(x)>0$, h 单调递增.

故 $h(x)$ 在 $x=\ln a$ 处取得全局最小值, 即

$$h(\ln a)=e^{\ln a}-a\ln a-1=a-a\ln a-1$$

要使恒成立, 需 $h(\ln a)\geq 0$, 即 $a(1-\ln a)-1\geq 0$. 观察辅助函数 $\varphi(a)=a-a\ln a-1$. 易知 $\varphi(1)=0$. 求导 $\varphi'(a)=-\ln a$. 当 $a=1$ 时, $\varphi(a)$ 取得最大值 0. 因此, 只有当 $a=1$ 时, 最小值 $h(\ln a)=0\geq 0$ 成立; 若 $a\neq 1$, 则最小值必然小于 0.

故符合条件的 a 只有 1. (几何上, 这对应于 $y=x+1$ 是曲线 $y=e^x$ 在 $(0,1)$ 处的切线, 切线始终在凸曲线下方.) □

16.11.2.2 参数分离法

参数分离法适用于能将参数 a 与变量 x 完全分离的情形, 例如变形为 $a\leq k(x)$. 此时 a 始终小于等于 $k(x)$, 等价于 a 不超过 $k(x)$ 的下确界, 或最小值, 即 $a\leq k(x)_{\min}$. 此法避免了对含参函数求导时复杂的分类讨论.

例题 16.67 若对任意 $x>1$, 不等式 $(x-1)\ln x-a(x-1)^2\geq 0$ 恒成立, 求 a 的取值范围.

解 由于 $x>1$, 故 $(x-1)^2>0$. 不等式可变形为分离参数的形式

$$a\leq\frac{(x-1)\ln x}{(x-1)^2}=\frac{\ln x}{x-1}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $k(x)=\frac{\ln x}{x-1}$ $(x>1)$. 问题转化为求 $a\le \inf_{x>1}k(x)$. 对 $k(x)$ 求导, 得

$$k'(x)=\frac{\frac{1}{x}(x-1)-\ln x\cdot 1}{(x-1)^2}=\frac{x-1-x\ln x}{x(x-1)^2}$$

令分子为 $g(x)=x-1-x\ln x$. 则 $g'(x)=1-(1+\ln x)=-\ln x$. 当 $x>1$ 时, $g'(x)<0$, 故 $g(x)$ 单调递减. 又 $g(1)=0$, 所以当 $x>1$ 时, $g(x)<0$. 进而可知 $k'(x)<0$, 函数 $k(x)$ 在 $(1,+\infty)$ 上单调递减.

再考察函数的值域边界. 利用基本极限可得

$$\lim_{x\to 1^+}\frac{\ln x}{x-1}=1,\quad \lim_{x\to+\infty}\frac{\ln x}{x-1}=0$$

函数从 1 单调递减趋向于 0. 因此, 要使 $a\le k(x)$ 对一切 $x>1$ 成立, 必须有 $a\le 0$. (若取 a 为正数, 当 x 足够大时 $k(x)$ 会趋近于 0 从而小于 a, 导致不等式失效.)

例题 16.68 (2018· 浙江卷)已知函数 $f(x)=\sqrt{x}-\ln x$.

1. 若 $f(x)$ 在 $x=x_1,x_2(x_1\ne x_2)$ 处的导数相等, 证明 $f(x_1)+f(x_2)>8-8\ln 2$;
2. 若 $a\le 3-4\ln 2$, 证明对于任意 $k>0$, 直线 $y=kx+a$ 与曲线 $y=f(x)$ 有唯一公共点.

解 (1) 换元与构造

求导得 $f'(x)=\frac{1}{2\sqrt{x}}-\frac{1}{x}=\frac{\sqrt{x}-2}{2x}$. 令 $t=\sqrt{x}>0$, 则 $f'(x)$ 对应 $h(t)=\frac{t-2}{2t^2}=\frac{1}{2t}-\frac{1}{t^2}$. 由 $f'(x_1)=f'(x_2)$ 且 $x_1\ne x_2$, 可知 t_1,t_2 是方程 $h(t)=C$ 的两个不同根.

$$\frac{1}{2t_1}-\frac{1}{t_1^2}=\frac{1}{2t_2}-\frac{1}{t_2^2}\implies \frac{1}{2}\left(\frac{1}{t_1}-\frac{1}{t_2}\right)=\frac{1}{t_1^2}-\frac{1}{t_2^2}.$$

约去 $\frac{1}{t_1}-\frac{1}{t_2}$, 得 $\frac{1}{2}=\frac{1}{t_1}+\frac{1}{t_2}$. 由此可知 $\frac{t_1+t_2}{t_1t_2}=\frac{1}{2}\Rightarrow t_1t_2=2(t_1+t_2)$. 又 $(t_1+t_2)^2>4t_1t_2=8(t_1+t_2)\implies t_1+t_2>8$. 由此推得 $t_1t_2=2(t_1+t_2)>16$.

目标式 $S=f(x_1)+f(x_2)=(t_1-2\ln t_1)+(t_2-2\ln t_2)=(t_1+t_2)-2\ln(t_1t_2)$. 代入关系式得 $S=\frac{1}{2}t_1t_2-2\ln(t_1t_2)$. 令 $u=t_1t_2>16$, 构造 $\phi(u)=\frac{u}{2}-2\ln u$. $\phi'(u)=\frac{1}{2}-\frac{2}{u}=\frac{u-4}{2u}$. 当 $u>16$ 时, $\phi'(u)>0$, ϕ 单调递增. 故 $S>\phi(16)=8-2\ln 16=8-8\ln 2$. 证毕.

(2) 转化零点问题直线与曲线的公共点等价于方程 $g(x)=f(x)-kx-a=0$ 的根.

$$g(x)=\sqrt{x}-\ln x-kx-a,\quad x\in(0,+\infty).$$

边界分析表明, $\lim_{x\to0^+}g(x)=+\infty$(主导项 $-\ln x$), 而 $\lim_{x\to+\infty}g(x)=-\infty$, 主导项为 $-kx$. 由介值定理, 至少存在一个零点. 要证唯一性, 需分析函数起伏.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

求导得 $g'(x)=\frac{1}{2\sqrt{x}}-\frac{1}{x}-k$. 令 $t=\sqrt{x}$, 导数符号由 $q(t)=-kt^2+\frac{1}{2}t-1$ 决定 ($g'(x)=\frac{q(t)}{t^2}$). 方程 $k=\frac{1}{2t}-\frac{1}{t^2}$ 的右边设为 $\varphi(t)$. $\varphi'(t)=\frac{4-t}{2t^3}$.$\varphi(t)$ 在 $(0,4)$ 递增, 在 $(4,+\infty)$ 递减. 最大值为 $\varphi(4)=\frac{1}{16}$.

情形一 ($k\geq\frac{1}{16}$). 此时 $g'(x)\leq 0$ 恒成立(仅在 $k=\frac{1}{16}, x=16$ 处为 0).$g(x)$ 单调递减, 必有唯一零点.

情形二 ($0<k<\frac{1}{16}$). 方程 $g'(x)=0$ 有两个根, 对应 $t_1\in(2,4)$ 和 $t_2\in(4,+\infty)$. $g(x)$ 在 $(0,x_1)$ 递减, 在 (x_1,x_2) 递增, 在 $(x_2,+\infty)$ 递减. 图像走势为从 $+\infty$ 下降到极小值, 再上升, 最后再下降到 $-\infty$. 要保证只有唯一零点, 必须确保极小值 $g(x_1)>0$(从而图像在 x_1 处不穿过 x 轴).

计算极小值 $g(x_1)$, 其中 $k=\varphi(t_1)=\frac{1}{2t_1}-\frac{1}{t_1^2}$:

$$g(x_1)=t_1-2\ln t_1-\left(\frac{1}{2t_1}-\frac{1}{t_1^2}\right)t_1^2-a=\frac{t_1}{2}-2\ln t_1+1-a.$$

令 $m(t)=\frac{t}{2}-2\ln t+1$. 需证明对于 $t_1\in(0,4)$, 恒有 $m(t_1)>a$. $m'(t)=\frac{1}{2}-\frac{2}{t}=\frac{t-4}{2t}$. 在 $(0,4)$ 上 $m'(t)<0$, m 单调递减. 故 $m(t)>m(4)=2-2\ln 4+1=3-4\ln 2$. 已知 $a\leq 3-4\ln 2$, 故 $g(x_1)=m(t_1)-a>m(4)-a\geq 0$.

因为极小值 $g(x_1)>0$, 函数在 $(0,x_2]$ 区间内恒正, 仅在 $(x_2,+\infty)$ 单调下降穿过 x 轴一次. 故公共点唯一.

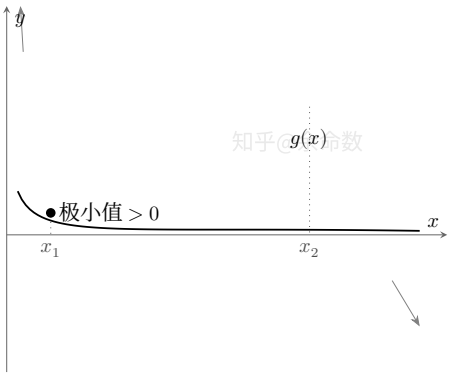


图 16.18 当 $0<k<1/16$ 时, 利用极小值大于 0 保证零点唯一

16.11.3 零点与含参问题的常用策略

处理零点问题时, 我们综合考察函数的形态与位置. 导数确定形态(单调性、凹凸性), 边界值与极值确定位置.

当 $f'(x)=0$ 为超越方程难以直接求解时, 不可强行计算, 而应分析导数表达式的结构.

1. **提取公因式**. 若 $f'(x)=A(x)\cdot h(x)$ 且 $A(x)$ 恒正, 则 $f(x)$ 的单调性完全由辅助函数 $h(x)$ 决定.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

2. **高阶求导**. 若 $f'(x)$ 的符号不明, 可继续对 $f'(x)$ 求导. 二阶导数 $f''(x)$ 的符号能确定 $f'(x)$ 的单调性, 进而确定 $f'(x)$ 的零点, 即原函数的极值点.
3. **隐零点代换**. 若 $f'(x_0)=0$ 的解 x_0 无法用初等函数显式表达, 保留 x_0 的形式, 利用 $f'(x_0)=0$ 导出的代数关系, 如 $e^{x_0}=\frac{a}{x_0}$ 对极值 $f(x_0)$ 进行降次或化简.

而在处理含参范围问题时, 直接讨论往往分支繁杂. 有效的策略是先寻找**必要条件**.

- **端点与特殊点限制**. 若要求 $f(x)$ 有零点, 函数在区间端点或特殊点, 如 $0,1,e$ 的值往往受到限制.
- **趋势分析**. 考察 $x\to\infty$ 或定义域边界时的极限. 若两端趋向于 $+\infty$, 则函数必有下界, 零点个数取决于最小值的符号.

确定参数的必要范围后, 再在缩小的范围内证明充分性.

此外, 当精确计算困难时, 可利用不等式对函数进行夹逼或放缩.

定理 16.17 (零点存在性的比较判别) 设 I 是一个区间, $f(x)$ 在 I 上连续.

- 若存在定义在 I 上的函数 $g(x)$ 使得在 I 上恒有 $f(x)\geq g(x)$, 且 $g(x)$ 的最小值为正, 则 $f(x)$ 无零点.
- 若存在点 $x_1,x_2\in I$ 使得 $f(x_1)>0$ 且 $f(x_2)<0$, 则 $f(x)$ 必有零点. 在寻找 x_1,x_2 时, 常用切线放缩, 如 $\ln x\leq x-1$ 或泰勒展开简化函数表达式.

证明 第一条中, 设 m 为 g 在 I 上的正最小值. 对任意 $x\in I$, 有 $f(x)\geq g(x)\geq m>0$, 因而 $f(x)$ 不可能等于 0.

第二条中, 不妨设 $x_1<x_2$. 因 I 是区间, $[x_1,x_2]\subset I$. 函数 f 在 $[x_1,x_2]$ 上连续, 且 $f(x_1)>0, f(x_2)<0$. 由介值定理, 存在 $\xi\in(x_1,x_2)$ 使得 $f(\xi)=0$. □

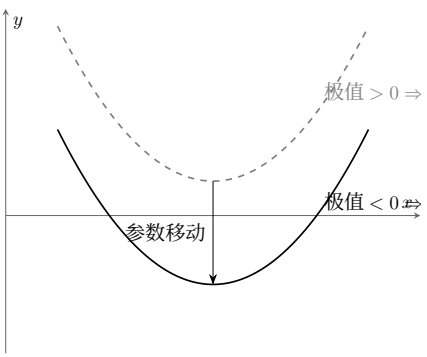


图 16.19　参数变化导致函数图像竖向平移, 极值点相对于 x 轴的位置决定零点个数

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

16.11.4 不等式中的存在性问题

与恒成立对应的是存在性问题, 即特称命题 $\exists x_0 \in I, f(x_0) \geq g(x_0)$.
我们有如下对应关系.

- 恒成立 $f(x) \geq 0$ 等价于 $[f(x)]_{\min} \geq 0$(最坏的情况都要达标);
- 存在性 $f(x) \geq 0$ 等价于 $[f(x)]_{\max} \geq 0$(只要最好的情况达标即可).

例题 16.69 已知函数 $f(x) = x^3 - 3x$. 问是否存在区间 $[0,2]$ 内的 x, 使得 $f(x) > a$? 求 a 的范围.

解 问题等价于 $a < [f(x)]_{\max}$,其中 $x \in [0,2]$. 求导得 $f'(x) = 3x^2-3 = 3(x-1)(x+1)$. 在 $[0,2]$ 上, 驻点为 $x = 1$.

- $x \in (0,1)$、$f'(x) < 0$、f 递减;
- $x \in (1,2)$、$f'(x) > 0$、f 递增.

知乎@余命数

极小值 $f(1) = 1 - 3 = -2$. 端点值 $f(0) = 0$, $f(2) = 8 - 6 = 2$. 比较可知, 最大值 $M = \max\{0, -2, 2\} = 2$. 因此, 只要 $a < 2$, 就存在这样的 x.

知乎@余命数

16.11.5 隐零点与双变量问题初步

在处理高阶导数或复杂函数时, 我们有时会遇到导函数的零点 x_0 无法通过方程 $f'(x) = 0$ 显式解出的情况, 如方程 $e^x = 2x + 3$. 此时需采用设而不求的策略.

知乎@余命数

16.11.5.1 隐零点代换(设而不求)

若 x_0 是方程 $f'(x) = 0$ 的根, 则有关系式 $f'(x_0) = 0$. 在后续求最值或证明不等式时, 将目标式中的参数, 或超越项用含有 x_0 的代数式替换, 从而把双变量问题(x 与参数)转化为关于 x_0 的单变量函数问题.

例题 16.70 已知函数 $f(x) = e^x - ax$ $(a > e)$ 有两个零点 x_1, x_2 $(x_1 < x_2)$. 证明 $x_1 + x_2 > 2$.

知乎@余命数

证明 由 $f(x) = 0$ 可知 $e^{x_1} = ax_1$ 且 $e^{x_2} = ax_2$. 即 x_1, x_2 是方程 $\frac{e^x}{x} = a$ 的两个根.

令 $g(t) = \frac{e^t}{t}$. 求导得 $g'(t) = \frac{e^t(t-1)}{t^2}$. $t = 1$ 是 $g(t)$ 的极小值点. 由于 $a > e = g(1)$, 直线 $y = a$ 与曲线 $y = g(t)$ 确有两个交点, 分布在 $t = 1$ 两侧, 即 $0 < x_1 < 1 < x_2$.

我们用对称点法. 要证 $x_1 + x_2 > 2$,即证 $x_2 > 2 - x_1$. 由于 $x_1 < 1$,故 $2 - x_1 > 1$. x_2 和 $2 - x_1$ 均落在区间 $(1, +\infty)$ 内. 在该区间上, 函数 $g(t)$ 单调递增. 因此, 证明 $x_2 > 2 - x_1$ 等价于证明 $g(x_2) > g(2 - x_1)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

已知 $g(x_2)=a=g(x_1)$, 故只需证明 $g(x_1)>g(2-x_1)$. 代入表达式, 即证

$$\frac{e^{x_1}}{x_1}>\frac{e^{2-x_1}}{2-x_1}\iff\frac{2-x_1}{x_1}>e^{2-2x_1}$$

令 $t=x_1\in(0,1)$. 不等式两边取对数, 变形为 $\ln(2-t)-\ln t>2-2t$. 设 $H(t)=\ln(2-t)-\ln t-2+2t$. 求导得

$$H'(t)=\frac{-1}{2-t}-\frac{1}{t}+2=\frac{-2+2t(2-t)}{t(2-t)}=\frac{-2(t-1)^2}{t(2-t)}$$

对于 $t\in(0,1)$, 恒有 $H'(t)<0$. 故 $H(t)$ 在 $(0,1)$ 上单调递减.

$$H(t)>H(1)=\ln 1-\ln 1-2+2=0$$

证毕. □

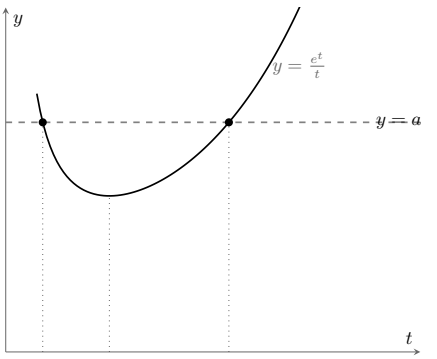


图 16.20　利用函数的单调性与对称性处理隐零点问题

16.11.6 高考真题选讲

以下四题选自历年高考真题, 覆盖零点分析、不等式恒成立与分类讨论的常见考法.

例题 16.71 (2017 年全国 □ 卷理科数学第 21 题) 已知函数 $f(x)=ax^2-ax-x\ln x$, 且 $f(x)\geq 0$.

1. 求 a;
2. 证明: $f(x)$ 存在唯一的极大值点 x_0, 且 $e^{-2}<f(x_0)<2^{-2}$.

解 (1) 求 a

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

$f(x)=x(ax-a-\ln x)\,(x>0)$, $f(x)\geq 0$ 等价于 $h(x)=ax-a-\ln x\geq 0$.$h'(x)=a-\dfrac{1}{x}$. 当 $a\leq 0$ 时 h 在 $(0,+\infty)$ 上单调递减, $h(x)<h(1)=0$ 对 $x>1$ 成立, 矛盾, 故 $a>0$.h 在 $x=\dfrac{1}{a}$ 处取最小值. 由 $h(1)=0$ 知 $\dfrac{1}{a}=1$, 故 $a=1$.

(2) 证明 $e^{-2}<f(x_0)<2^{-2}$

$f(x)=x^2-x-x\ln x$, $f'(x)=2x-2-\ln x$. 令 $t(x)=2x-2-\ln x$, $t'(x)=2-\dfrac{1}{x}$.t 在 $(0,\frac{1}{2})$ 上递减, 在 $(\frac{1}{2},+\infty)$ 上递增, $t(\frac{1}{2})=\ln 2-1<0$, $t(\frac{1}{e^2})=\frac{2}{e^2}>0$, 故 t 在 $(0,\frac{1}{2})$ 上有唯一零点 x_0, 在 $(\frac{1}{2},+\infty)$ 上有唯一零点 x_2.$f'(x)>0$ 在 $(0,x_0)$ 上成立, $f'(x)<0$ 在 (x_0,x_2) 上成立, 故 x_0 是 $f(x)$ 的唯一极大值点.

由 $2x_0-2-\ln x_0=0$ 得 $\ln x_0=2x_0-2$, 故

$$f(x_0)=x_0^2-x_0-x_0(2x_0-2)=x_0-x_0^2$$

由 $x_0<\dfrac{1}{2}$ 得 $f(x_0)<\dfrac{1}{4}=2^{-2}$. 由 $f'(\frac{1}{e})<0$ 知 $x_0<\frac{1}{e}$, 故 $f(x_0)>f(\frac{1}{e})=\frac{1}{e^2}=e^{-2}$.

例题 16.72 (2019 年全国 □ 卷理科数学第 20 题) 已知函数 $f(x)=\sin x-\ln(1+x)$, $f'(x)$ 为 $f(x)$ 的导数. 证明:

1. $f'(x)$ 在区间 $(-1,\frac{\pi}{2})$ 存在唯一极大值点;
2. $f(x)$ 有且仅有 2 个零点.

解 (1)$f'(x)$ 的极大值点

$f'(x)=\cos x-\dfrac{1}{1+x}$, $f''(x)=-\sin x+\dfrac{1}{(1+x)^2}$. 令 $g(x)=f''(x)$, 则 $g'(x)=-\cos x-\dfrac{2}{(1+x)^3}<0$ 在 $(-1,\frac{\pi}{2})$ 上恒成立, 故 $f''(x)$ 在 $(-1,\frac{\pi}{2})$ 上单调递减. 由 $f''(0)=1>0$ 与 $f''(\frac{\pi}{2})<0$ 及零点存在定理, $f''(x)$ 在 $(-1,\frac{\pi}{2})$ 上存在唯一零点 x_0. 因此 $f'(x)$ 在 $(-1,x_0)$ 上递增, 在 $(x_0,\frac{\pi}{2})$ 上递减, x_0 为唯一极大值点.

(2)$f(x)$ 的零点个数

在 $(-1,0)$ 上 $f'(x)$ 递增, $f'(x)<f'(0)=0$, f 递减, $f(x)>f(0)=0$.

在 $(0,x_0)$ 上 $f'(x)>f'(0)=0$, f 递增. 在 $(x_0,\frac{\pi}{2})$ 上 f' 递减, $f'(x_0)>0$, $f'(\frac{\pi}{2})<0$, 故 f' 在 $(x_0,\frac{\pi}{2})$ 上有唯一零点 x_1.f 在 (x_0,x_1) 上递增, 在 $(x_1,\frac{\pi}{2})$ 上递减.$f(\frac{\pi}{2})=1-\ln(1+\frac{\pi}{2})>0$.

在 $(\frac{\pi}{2},\pi)$ 上 $f'(x)=\cos x-\frac{1}{1+x}<0$, f 递减.$f(\pi)=-\ln(1+\pi)<0$, 故 f 在 $(\frac{\pi}{2},\pi)$ 上有唯一零点.

当 $x\geq\pi$ 时 $\sin x\leq 1<\ln(1+x)$, $f(x)<0$, 无零点.

$f(x)$ 的零点为 $x=0$ 及 $(\frac{\pi}{2},\pi)$ 中的唯一零点, 共 2 个.

知乎@余命数 知乎@余命数 知乎@余命数

例题 16.73 (2019 年全国 □ 卷理科数学第 20 题) 已知函数 $f(x)=2x^3-ax^2+b$.

1. 讨论 $f(x)$ 的单调性;
2. 是否存在 a,b, 使得 $f(x)$ 在区间 $[0,1]$ 的最小值为 -1 且最大值为 1? 若存在, 求出 a,b 的所有值; 若不存在, 说明理由.

解 (1) 单调性

$f'(x)=6x^2-2ax=6x\left(x-\dfrac{a}{3}\right)$.

$a=0$: $f'(x)\geq 0$, f 在 $\mathbb{R}$ 上单调递增.

$a>0$: f 在 $(-\infty,0)$ 和 $(\frac{a}{3},+\infty)$ 上单调递增, 在 $(0,\frac{a}{3})$ 上单调递减.

$a<0$: f 在 $(-\infty,\frac{a}{3})$ 和 $(0,+\infty)$ 上单调递增, 在 $(\frac{a}{3},0)$ 上单调递减.

(2) 存在性讨论

$a\leq 0$: f 在 $[0,1]$ 上单调递增.$f(0)=b=-1$, $f(1)=2-a+b=1$, 解得 $a=0,b=-1$.

$a\geq 3$ ($\frac{a}{3}\geq 1$): f 在 $[0,1]$ 上单调递减.$f(0)=b=1$, $f(1)=2-a+b=-1$, 解得 $a=4,b=1$.

$0<a<3$ ($0<\frac{a}{3}<1$): f 在 $[0,\frac{a}{3}]$ 上递减, 在 $[\frac{a}{3},1]$ 上递增. 最小值 $f(\frac{a}{3})=-\frac{a^3}{27}+b=-1$. 最大值为 b 或 $2-a+b$. 若 $b=1$ 则 $a=3\sqrt[3]{2}>3$, 矛盾. 若 $2-a+b=1$ 则 $a=0$ 或 $\pm 3\sqrt{3}$, 均与 $0<a<3$ 矛盾.

存在两组解: $(a,b)=(0,-1)$ 或 $(4,1)$.

例题 16.74 (2017 年全国 □ 卷理科数学第 21 题) 已知函数 $f(x)=ae^{2x}+(a-2)e^x-x$.

1. 讨论 $f(x)$ 的单调性;
2. 若 $f(x)$ 有两个零点, 求 a 的取值范围.

解 (1) 单调性

$f'(x)=2ae^{2x}+(a-2)e^x-1$.

$a=0$: $f'(x)=-2e^x-1<0$, f 在 $\mathbb{R}$ 上单调递减.

$a>0$: $f'(x)=2a\left(e^x+\frac{1}{2}\right)\left(e^x-\frac{1}{a}\right)$, 零点为 $x=\ln\frac{1}{a}$. f 在 $(-\infty,\ln\frac{1}{a})$ 上单调递减, 在 $(\ln\frac{1}{a},+\infty)$ 上单调递增.

$a<0$: $f'(x)<0$ 恒成立, f 在 $\mathbb{R}$ 上单调递减.

(2) 求 a 的取值范围

$a\leq 0$: f 单调递减, 至多一个零点.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$a>0$：f 在 $x=\ln\frac{1}{a}$ 处取最小值. 当 $x\to\pm\infty$ 时 $f(x)\to+\infty$，故有两个零点当且仅当 $f\left(\ln\frac{1}{a}\right)<0$.

$$f\left(\ln\frac{1}{a}\right)=\frac{1}{a}+1-\frac{2}{a}-\ln\frac{1}{a}=1-\frac{1}{a}-\ln\frac{1}{a}$$

令 $t=\frac{1}{a}$，需 $1-t-\ln t<0$，即 $\ln t+t-1>0$. 令 $g(t)=\ln t+t-1$、$g'(t)=\frac{1}{t}+1>0$、$g(1)=0$，故 $g(t)>0$ 当且仅当 $t>1$，即 $0<a<1$.

所求 a 的取值范围是 $(0,1)$.

16.12 高阶导数与泰勒展开初步

前几章主要用一阶导数 $f'(x)$ 研究函数的单调性与极值. 一阶导数刻画瞬时变化率，几何上对应切线斜率. 本章引入高阶导数，特别是二阶导数 $f''(x)$，用它分析函数图像的弯曲程度（凹凸性）.

泰勒公式是用多项式逼近复杂函数的经典工具. 它把初等数学里的一次函数、二次函数推广到任意次数，也为不等式证明提供了一种通用方法.

16.12.1 高阶导数与函数的凹凸性

16.12.1.1 高阶导数的定义

设函数 $f(x)$ 的导函数 $f'(x)$ 仍然可导. 我们称 $f'(x)$ 的导数为 $f(x)$ 的**二阶导数**，记作 $f''(x)$ 或 $\frac{d^2y}{dx^2}$. 类似地，若二阶导数仍可导，则其导数称为 $f(x)$ 的**三阶导数**，记作 $f'''(x)$ 或 $f^{(3)}(x)$.

若用 $s(t)$ 表示物体运动的位移，则 $s'(t)$ 是瞬时速度，$s''(t)$ 是瞬时加速度.

从几何上看，$f''(x)$ 描述了切线斜率 $f'(x)$ 变化的快慢.

- 若 $f''(x)>0$，说明斜率 $f'(x)$ 单调递增，曲线向上弯曲.
- 若 $f''(x)<0$，说明斜率 $f'(x)$ 单调递减，曲线向下弯曲.

16.12.1.2 函数的凹凸性

二阶导数直接决定函数图像的弯曲方向，我们把它称为凹凸性.

定义 16.22 (凹凸性) 设函数 $f(x)$ 在区间 I 上连续.

- **下凸 (Convex)**. 若图像上任意两点连成的弦始终位于这段曲线的**上方**，则称 f 在 I 上是**下凸**的，图像形如 $\cup$.
- **上凸 (Concave)**. 若图像上任意两点连成的弦始终位于这段曲线的**下方**，则称 f 在 I 上是**上凸**的，图像形如 $\cap$.

注 16.14
国内外教材对“凹”与“凸”的中文命名常有差异. 为避免歧义，本节主要关注图像的几何形状，即 $\cup$ 型与 $\cap$ 型.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 16.18 (凹凸性判定法则) 设 $f(x)$ 在区间 I 上具有二阶导数.

- 若在 I 内恒有 $f''(x)>0$, 则 $f(x)$ 的图像是**下凸**的 ($\cup$).
- 若在 I 内恒有 $f''(x)<0$, 则 $f(x)$ 的图像是**上凸**的 ($\cap$).

证明 任取 $x_1,x_2\in I$ 且 $x_1<x_2$, 再取 $x\in(x_1,x_2)$. 若 $f''(x)>0$ 在 I 内恒成立, 则由函数单调性的判定法, $f'(x)$ 在 I 上严格递增. 分别在 $[x_1,x]$ 与 $[x,x_2]$ 上应用拉格朗日中值定理, 可得

$$\frac{f(x)-f(x_1)}{x-x_1}=f'(\eta_1),\qquad \frac{f(x_2)-f(x)}{x_2-x}=f'(\eta_2),$$

其中 $\eta_1\in(x_1,x),\eta_2\in(x,x_2)$. 因 $\eta_1<\eta_2$, 故 $f'(\eta_1)<f'(\eta_2)$, 从而

$$\frac{f(x)-f(x_1)}{x-x_1}<\frac{f(x_2)-f(x)}{x_2-x}.$$

整理得

$$f(x)<\frac{x_2-x}{x_2-x_1}f(x_1)+\frac{x-x_1}{x_2-x_1}f(x_2),$$

即曲线在任意弦的下方, 图像为下凸. 若 $f''(x)<0$, 则 f' 严格递减, 上述不等式方向全部反向, 曲线在任意弦的上方, 图像为上凸. □

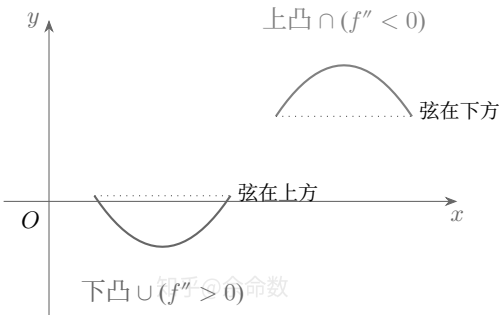


图 16.21　函数凹凸性与弦的位置关系

16.12.1.3 拐点

定义 16.23 (拐点) 连续曲线 $y=f(x)$ 上凹凸性发生改变的点, 称为**拐点**.

若 $f''(x)$ 存在, 则 $f''(x_0)=0$ 是点 $(x_0,f(x_0))$ 为拐点的**必要条件**. 判定拐点时, 我们检查 $f''(x)$ 在 x_0 两侧是否**变号**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.12.1.4 极值的第二充分条件

利用二阶导数, 我们可以在不列出单调性表格的情况下判断一个驻点是极大值点还是极小值点.

定理 16.19 (极值的第二充分条件) 设函数 f 在 x_0 的某邻域内一阶可导, $f'(x_0)=0$, 且 $f''(x_0)$ 存在.

- 若 $f''(x_0)>0$, 则 $f(x_0)$ 是**极小值**(图像 $\cup$ 且切线水平).
- 若 $f''(x_0)<0$, 则 $f(x_0)$ 是**极大值**(图像 $\cap$ 且切线水平).
- 若 $f''(x_0)=0$, 该判定法失效, 需回到单调性分析.

证明 我们用二阶导数的定义. 若 $f''(x_0)>0$, 则

$$\lim_{h\to 0}\frac{f'(x_0+h)-f'(x_0)}{h}=f''(x_0)>0.$$

因此存在 $\delta>0$, 使得当 $0<|h|<\delta$ 时,

$$\frac{f'(x_0+h)-f'(x_0)}{h}>0.$$

又 $f'(x_0)=0$, 所以当 $0<h<\delta$ 时 $f'(x_0+h)>0$, 当 $-\delta<h<0$ 时 $f'(x_0+h)<0$. 于是 f 在 x_0 左侧单调递减, 在 x_0 右侧单调递增, 故 $f(x_0)$ 为极小值.

若 $f''(x_0)<0$, 同理可得 $f'(x_0+h)>0$ 对 $h<0$ 成立, $f'(x_0+h)<0$ 对 $h>0$ 成立, 因而 $f(x_0)$ 为极大值. 当 $f''(x_0)=0$ 时, 上述符号推断失效; 例如 x^4 在 0 处有极小值, 而 x^3 在 0 处无极值. 因此需改用单调性分析. □

16.12.2 泰勒展开

一次函数与二次函数形式简单、易于计算. 对于 e^x、$\sin x$、$\ln x$ 等超越函数, 基本想法是用多项式去逼近, 让多项式在某一点处的函数值以及各阶导数值都与原函数相等. 次数越高, 近似程度越好.

16.12.2.1 逼近的逻辑

设我们要用一个多项式在 $x=0$ 附近逼近函数 $f(x)$.

- **零阶近似**. $P_0(x)=f(0)$, 只保证高度相同.
- **一阶近似**(切线). $P_1(x)=f(0)+f'(0)x$, 保证高度与斜率相同.
- **二阶近似**(抛物线). $P_2(x)=f(0)+f'(0)x+\frac{f''(0)}{2}x^2$, 保证高度、斜率、弯曲程度相同.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

沿着这个逻辑，若要求多项式 $P_n(x)$ 在 $x=0$ 处的 0 到 n 阶导数都与 $f(x)$ 相同，就得到泰勒公式.

16.12.2.2 带有佩亚诺余项的麦克劳林公式

设函数 $f(x)$ 在 $x=0$ 处具有 n 阶导数. 那么

$$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+o(x^n) \tag{16-17}$$

其中 $o(x^n)$ 称为**佩亚诺余项**，表示当 $x\to 0$ 时，误差是比 x^n 更高阶的无穷小.

16.12.2.3 常用函数的展开式

以下几个展开式在 $x\to 0$ 时成立.

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+o(x^n) \tag{16-18}$$

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots+(-1)^{m-1}\frac{x^{2m-1}}{(2m-1)!}+o(x^{2m-1}) \tag{16-19}$$

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots+(-1)^{m}\frac{x^{2m}}{(2m)!}+o(x^{2m}) \tag{16-20}$$

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots+(-1)^{n-1}\frac{x^n}{n}+o(x^n) \tag{16-21}$$

16.12.3 泰勒近似在不等式证明中的应用

泰勒公式把超越函数写成多项式，我们还可以控制余项的符号，从而得到不等式.

16.12.3.1 切线放缩的再审视

我们已在 定理 5.14 中证明过 $e^x\geq 1+x$. 这其实是 e^x 的一阶泰勒展开.

- 因为 $(e^x)''=e^x>0$，函数下凸，图像位于切线 $y=1+x$ 上方，故 $e^x\geq 1+x$.
- 推广到一般情形，若 $f''(x)>0$，则 $f(x)\geq f(x_0)+f'(x_0)(x-x_0)$.

16.12.3.2 高阶展开的应用实例

用更高阶的展开，可以得到更精确的不等式.

例题 16.75　证明当 $x>0$ 时，$\sin x>x-\frac{x^3}{6}$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

证明 我们比较 $\sin x$ 与它的三阶泰勒多项式. 构造函数 $f(x)=\sin x-\left(x-\frac{x^3}{6}\right)$. 要证当 $x>0$ 时 $f(x)>0$. 对 $f(x)$ 逐次求导, 得

知乎@余命数

$$f'(x)=\cos x-1+\frac{x^2}{2}$$

$$f''(x)=-\sin x+x$$

$$f'''(x)=-\cos x+1$$

先分析 $f'''(x)$. 对于任意 $x\in\mathbb{R}$, 都有 $\cos x\leq 1$, 故 $f'''(x)=1-\cos x\geq 0$. 所以 $f''(x)$ 在 $x\geq 0$ 上单调递增. 又 $f''(0)=-\sin 0+0=0$, 因此当 $x>0$ 时, $f''(x)>0$.

同理, 由 $f''(x)>0$ 可知 $f'(x)$ 单调递增. 又 $f'(0)=\cos 0-1+0=0$, 所以当 $x>0$ 时, $f'(x)>0$.

再由 $f'(x)>0$ 可知 $f(x)$ 单调递增. 又 $f(0)=\sin 0-0=0$, 所以当 $x>0$ 时, $f(x)>0$. 即 $\sin x>x-\frac{x^3}{6}$ 成立. □

注 16.15
这个例子说明, 反复求导可以把关于 $\sin x$ 的问题化为常数与 0 的比较, 再反向推回. 这正是泰勒公式处理不等式时的常用路径.

16.13 泰勒公式与泰勒级数

知乎@余命数

多项式仅涉及有限次加法与乘法运算, 易于计算、求导与积分. 对于复杂的非多项式函数, 如 $e^x,\sin x,\ln x$, 可以找一个多项式在局部替代它.

我们的思路是, 若多项式 $P(x)$ 在某点 x_0 处与原函数 $f(x)$ 数值相等, 且一阶导数、二阶导数……直到 n 阶导数都相等, 则 $P(x)$ 在 x_0 附近与 $f(x)$ 高度一致.

知乎@余命数

16.13.1 多项式逼近的构造

考虑在点 $x_0=0$ 附近逼近函数 $f(x)$, 假设 $f(x)$ 在该点具有任意阶导数. 构造 n 次多项式

$$P_n(x)=c_0+c_1x+c_2x^2+\cdots+c_nx^n$$

使得 $P_n(x)$ 与 $f(x)$ 在 $x=0$ 处的 0 到 n 阶导数相等.

知乎@余命数

$$P_n^{(k)}(0)=f^{(k)}(0),\quad k=0,1,\ldots,n$$

对 $P_n(x)$ 逐次求导并令 $x=0$, 得

知乎@余命数

$$P_n(0)=c_0$$

$$P_n'(0)=c_1$$

知乎@余命数

$$P_n''(0)=2\cdot 1\cdot c_2$$

$$P_n'''(0)=3\cdot 2\cdot 1\cdot c_3$$

$$\vdots$$

$$P_n^{(k)}(0) = k! \cdot c_k$$

令 $P_n^{(k)}(0) = f^{(k)}(0)$, 解得系数 c_k:

$$c_k = \frac{f^{(k)}(0)}{k!}$$

由此得到逼近多项式的显式表达. 把它推广到任意展开点 x_0, 即得到泰勒多项式.

定义 16.24 (泰勒多项式) 设函数 $f(x)$ 在点 x_0 处具有 n 阶导数. 称多项式

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

为 $f(x)$ 在 x_0 处的 n **次泰勒多项式**. 当 $x_0 = 0$ 时, 称为 **麦克劳林多项式**.

16.13.2 泰勒公式与余项

多项式 $T_n(x)$ 与 $f(x)$ 之间的差为误差项(余项)$R_n(x)$:

$$f(x) = T_n(x) + R_n(x)$$

该等式构成**泰勒公式**. 余项的大小决定逼近的精度.

定理 16.20 (泰勒中值定理) 如果函数 $f(x)$ 在包含 x_0 与 x 的闭区间上具有连续的 n 阶导数, 并在其内部具有 $n+1$ 阶导数, 则存在介于 x 与 x_0 之间的某个 ξ, 使得

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

其中最后一项称为**拉格朗日余项**.

证明 当 $x = x_0$ 时结论显然成立. 以下设 $x \neq x_0$, 并不妨先考虑 $x > x_0$; 若 $x < x_0$, 只需把区间端点对调.

记

$$P_n(t) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(t - x_0)^k, \qquad R = f(x) - P_n(x).$$

构造辅助函数

$$\Phi(t) = f(t) - P_n(t) - R\left(\frac{t - x_0}{x - x_0}\right)^{n+1}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由构造可知 $\Phi(x_0)=\Phi(x)=0$. 进一步地, 由泰勒多项式的系数选取可知

$$\Phi(x_0)=\Phi'(x_0)=\cdots=\Phi^{(n)}(x_0)=0.$$

对 Φ 在区间 $[x_0,x]$ 上反复应用罗尔定理. 第一次得到 Φ' 在 (x_0,x) 内有零点; 结合 $\Phi'(x_0)=0$ 再应用一次, 得到 Φ'' 在 (x_0,x) 内有零点; 如此重复 $n+1$ 次, 可得存在 $\xi\in(x_0,x)$ 使得

$$\Phi^{(n+1)}(\xi)=0.$$

而

$$\Phi^{(n+1)}(t)=f^{(n+1)}(t)-R\frac{(n+1)!}{(x-x_0)^{n+1}}.$$

代入 $t=\xi$ 得

$$R=\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

将 $R=f(x)-P_n(x)$ 代回, 即得所需公式. □

该定理是微分中值定理的推广 ($n=0$ 时即为拉格朗日中值定理). 余项表明, 若 $f^{(n+1)}(x)$ 有界, 且 $(x-x_0)$ 较小或 n 较大, 误差会迅速衰减.

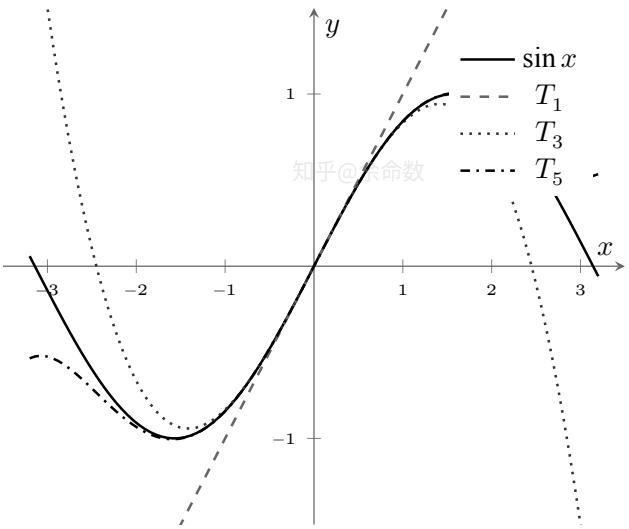


图 16.22 $\sin x$ 及其麦克劳林多项式 T_1,T_3,T_5 的图像. 随阶数 n 增加, 多项式与原函数的重合范围逐渐扩大.

16.13.3 泰勒级数

当 n 趋于无穷大时, 如果余项 $R_n(x)$ 的极限为 0, 则泰勒公式转化为一个无穷级数, 此时函数被级数**精确表示**.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 16.25 (泰勒级数) 设 $f(x)$ 在 x_0 处具有任意阶导数. 级数

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

称为 $f(x)$ 在 x_0 处的**泰勒级数**.

泰勒级数的收敛域可能仅限于 x_0 附近的一个区间(收敛区间). 在该区间外, 级数发散, 或虽收敛但不等于原函数.

16.13.4 常见函数的麦克劳林展开

以下展开式是分析学中的常用结论, 展开点均为 $x_0=0$.

16.13.4.1 指数函数

设 $f(x)=e^x$. 由 $f^{(n)}(x)=e^x$, 故 $f^{(n)}(0)=1$.

$$e^x=1+x+\frac{x^2}{2!}+\frac{x^3}{3!}+\cdots=\sum_{n=0}^{\infty}\frac{x^n}{n!}\quad(x\in\mathbb{R})$$

此级数对所有实数收敛, 且收敛速度快.

知乎@余命数

16.13.4.2 正弦与余弦函数

对于 $f(x)=\sin x$, 各阶导数循环为 $\cos x$、$-\sin x$、$-\cos x$、$\sin x$, ……; 在 $x=0$ 处依次取值 0、1、0、-1. 偶数项系数为 0, 奇数项符号交替.

$$\begin{aligned}\sin x&=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots\\&=\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!}\quad(x\in\mathbb{R}).\end{aligned}$$

同理可得余弦函数的展开(仅含偶次项).

$$\begin{aligned}\cos x&=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots\\&=\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n)!}\quad(x\in\mathbb{R}).\end{aligned}$$

这也解释了 $\sin x$ 为何是奇函数而 $\cos x$ 为何是偶函数, 前者仅含奇次幂, 后者仅含偶次幂.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.13.4.3 二项式级数

考虑 $f(x)=(1+x)^{\alpha}$, 其中 α 为任意实数. 其系数满足 $c_n=\frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}$, 记为广义组合数 $\binom{\alpha}{n}$.

$$(1+x)^{\alpha}=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots=\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n \quad (|x|<1)$$

当 α 为正整数时, 级数截断为有限项, 即牛顿二项式定理; 当 α 为负数或分数时, 得到无穷级数.

例题 16.76 利用泰勒展开计算 $\sqrt{1+x}$ 的近似多项式(保留至二次项).

解 取 $\alpha=1/2$.

- $n=0$: $c_0=1$.
- $n=1$: $c_1=\alpha=1/2$.
- $n=2$: $c_2=\frac{(1/2)(1/2-1)}{2!}=\frac{-1/4}{2}=-1/8$.

因此

$$\sqrt{1+x}\approx 1+\frac{1}{2}x-\frac{1}{8}x^2$$

当 x 较小时, 该多项式是平方根函数的一个良好近似.

16.13.5 欧拉公式的导出

观察 $e^x,\sin x,\cos x$ 的级数形式. 若把虚数单位 $i(i^2=-1)$ 代入 e^x 的级数,得到.

$$\begin{aligned}e^{ix}&=1+ix+\frac{(ix)^2}{2!}+\frac{(ix)^3}{3!}+\frac{(ix)^4}{4!}+\ldots\\&=1+ix-\frac{x^2}{2!}-i\frac{x^3}{3!}+\frac{x^4}{4!}+\ldots\\&=\left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-\ldots\right)+i\left(x-\frac{x^3}{3!}+\ldots\right)\\&=\cos x+i\sin x\end{aligned}$$

这就是**欧拉公式**, $e^{ix}=\cos x+i\sin x$. 它通过泰勒级数把指数函数与三角函数统一在复数域中.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.14 洛必达法则与极限计算的系统策略

我们在 第16.3节 中已学过极限的多种计算方法. 但面对 $\frac{0}{0}$、$\frac{\infty}{\infty}$ 这类不定式极限, 除少数特殊形式外, 我们还没有普适的处理工具.定理 16.15 给出的柯西中值定理是处理这类问题的理论基础.

本节从该定理导出**洛必达法则** (L'Hôpital's Rule), 用它把不定式极限的计算化为求导问题, 并给出按复杂度分层的极限求解策略.

16.14.1 洛必达法则

我们先从该法则的几何与代数背景出发, 说明导数的比值为什么能取代原函数的比值来求极限.

设 $x \to a$ 时 $f(x)$ 与 $g(x)$ 同时趋向 0(或 ∞). 此时比值的极限取决于两者趋向目标的相对快慢, 而度量变化快慢的工具是**导数**.

当 $f(a) = g(a) = 0$ 时, 直接比较两个函数值, 即计算 $0/0$ 没有意义, 但我们可以比较两者趋近 0 的速度. 若 $f(x)$ 趋向 0 的速度是 $g(x)$ 的两倍, 则在 a 附近 $f(x)$ 的值大约也是 $g(x)$ 的两倍, 极限为 2. 而 $f(x)$ 在 $x = a$ 处的变化速度就是 $f'(a)$. 于是我们提出猜想.

$$\lim_{x \to a} \frac{f(x)}{g(x)} \approx \frac{f'(a)}{g'(a)}$$

我们用导数的定义对这一猜想作初步验证. 设 $f(a) = g(a) = 0$, f, g 在 $x = a$ 处可导, 且 $g'(a) \neq 0$. 由导数的定义, 当 x 充分接近 a 时, 有近似关系 $f(x) \approx f'(a)(x - a)$ 与 $g(x) \approx g'(a)(x - a)$. 两式相除, 公因子 $(x - a)$ 被消去, 得

$$\frac{f(x)}{g(x)} \approx \frac{f'(a)(x - a)}{g'(a)(x - a)} = \frac{f'(a)}{g'(a)}$$

于是洛必达法则的直观图像是, 用函数在 a 点附近的**切线(线性近似)**替代原函数来计算.

然而在实际问题中, $f'(a)$ 或 $g'(a)$ 可能不存在, 比值本身也可能仍是不定式. 这时我们需要更坚实的理论基础, 即**柯西中值定理**. 根据柯西中值定理, 对区间 $[a, x]$, 必存在一点 ξ 介于 a 与 x 之间, 使得

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi)}{g'(\xi)}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

利用条件 $f(a)=g(a)=0$, 等式左边即简化为 $\frac{f(x)}{g(x)}$. 现在对等式两边取 $x\to a$ 的极限. 关键在于变量的传递.ξ 被 a 与 x 夹在中间 $(a<\xi<x)$, 当 x 逼近 a 时, ξ 也随之逼近 a. 因此 $\lim_{x\to a}$ 对 ξ 的作用等价于 $\lim_{\xi\to a}$. 于是得

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\lim_{\xi\to a}\frac{f'(\xi)}{g'(\xi)}$$

于是原函数的极限问题转化为导函数的极限问题, 其依据是柯西中值定理建立的函数增量比与导数比值之间的恒等关系.

我们把洛必达法则严格叙述如下. 在特定条件下, 两个函数之比的极限等于它们导数之比的极限, 我们借助导数(切线斜率)的信息推断函数值的趋近行为.

定理 16.21 (洛必达法则 (有限导数比版本)) 设函数 $f(x)$ 和 $g(x)$ 在有限点 a 的某去心邻域内可导, 且在该邻域内 $g(x)\neq 0, g'(x)\neq 0$. 若满足以下两类条件之一:

1. $\frac{0}{0}$ **型**: $\lim_{x\to a}f(x)=0$ 且 $\lim_{x\to a}g(x)=0$;
2. $\frac{\infty}{\infty}$ **型**: $\lim_{x\to a}f(x)=\infty$ 且 $\lim_{x\to a}g(x)=\infty$;

并且极限 $\lim_{x\to a}\frac{f'(x)}{g'(x)}=A$ 存在且为有限实数, 那么

$$\lim_{x\to a}\frac{f(x)}{g(x)}=A=\lim_{x\to a}\frac{f'(x)}{g'(x)} \tag{16-22}$$

对单侧极限同样成立; $x\to\infty$ 的情形可通过变量代换 $t=1/x$ 化为有限点的单侧极限.

证明 我们用柯西中值定理来证明. 只证右极限 $x\to a^+$ 的情形, 左极限的证明完全类似, 双侧极限由左右极限一致即得.

对于 $\frac{0}{0}$ 型, 补充定义 $f(a)=g(a)=0$. 当 $x>a$ 且充分靠近 a 时, f,g 在 $[a,x]$ 上连续、在 (a,x) 内可导. 由柯西中值定理, 存在 $\xi\in(a,x)$, 使得

$$\frac{f(x)}{g(x)}=\frac{f(x)-f(a)}{g(x)-g(a)}=\frac{f'(\xi)}{g'(\xi)}.$$

当 $x\to a^+$ 时有 $\xi\to a^+$, 因而上式右端趋于 A.

对于 $\frac{\infty}{\infty}$ 型, 任取 $\varepsilon>0$. 由 $\frac{f'}{g'}\to A$, 存在 $\delta>0$, 使得当 $a<u<a+\delta$ 时,

$$\left|\frac{f'(u)}{g'(u)}-A\right|<\varepsilon.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

固定一点 $c \in (a, a+\delta)$. 当 $x \in (a, c)$ 时，在 $[x, c]$ 上应用柯西中值定理，存在 $\xi \in (x, c) \subset (a, a+\delta)$, 使得

$$\frac{f(x)-f(c)}{g(x)-g(c)} = \frac{f'(\xi)}{g'(\xi)},$$

因而该增量比与 A 的距离小于 ε. 又

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(c)}{g(x)-g(c)} \cdot \frac{g(x)-g(c)}{g(x)} + \frac{f(c)}{g(x)}.$$

当 $x \to a^+$ 时, 由 $g(x) \to \infty$ 得 $\frac{g(x)-g(c)}{g(x)} \to 1$ 且 $\frac{f(c)}{g(x)} \to 0$. 因此

$$\begin{aligned}\left|\frac{f(x)}{g(x)} - A\right| \leq \left|\frac{f(x)-f(c)}{g(x)-g(c)} - A\right|\left|\frac{g(x)-g(c)}{g(x)}\right| \\ + |A|\left|\frac{g(x)-g(c)}{g(x)} - 1\right| + \left|\frac{f(c)}{g(x)}\right|.\end{aligned}$$

令 $x \to a^+$ 后右端上极限不超过 ε. 由于 ε 任意, 得 $\frac{f(x)}{g(x)} \to A$.

若讨论 $x \to \infty$, 令 $t = 1/x$, 则 $x \to \infty$ 化为 $t \to 0^+$, 且

$$\frac{\mathrm{d}}{\mathrm{d}t} f(1/t) \Big/ \frac{\mathrm{d}}{\mathrm{d}t} g(1/t) = \frac{f'(1/t)}{g'(1/t)}.$$

因此无穷远处的有限导数比情形也归结为上述结论. □

16.14.1.1 基本应用实例

例题 16.77 计算 $\lim_{x\to 0} \frac{\sin x}{x}$.

解 这是 $\frac{0}{0}$ 型不定式, 我们用它演示洛必达法则的操作. 要注意这里有循环论证之嫌, 三角函数的求导公式本身已依赖这个重要极限.

$$\lim_{x\to 0} \frac{\sin x}{x} \overset{L'H}{=} \lim_{x\to 0} \frac{(\sin x)'}{(x)'} = \lim_{x\to 0} \frac{\cos x}{1} = 1$$

例题 16.78 计算 $\lim_{x\to+\infty} \frac{\ln x}{x}$.

解 这是 $\frac{\infty}{\infty}$ 型不定式.

$$\lim_{x\to+\infty} \frac{\ln x}{x} \overset{L'H}{=} \lim_{x\to+\infty} \frac{1/x}{1} = 0$$

这表明, 对数函数的增长速度远慢于幂函数.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 16.79 计算 $\lim_{x\to+\infty}\frac{x^n}{e^x}$（$n$ 为正整数）.

解 这是 $\frac{\infty}{\infty}$ 型. 我们连续使用 n 次洛必达法则, 直至分子变为常数.

$$\lim_{x\to+\infty}\frac{x^n}{e^x}\overset{L'H}{=}\lim_{x\to+\infty}\frac{nx^{n-1}}{e^x}=\cdots=\lim_{x\to+\infty}\frac{n!}{e^x}=0$$

这表明, 指数函数的增长速度远快于任何多项式函数.

16.14.2 法则的滥用与陷阱

洛必达法则的适用范围有限, 盲目使用会导致错误或死循环.

16.14.2.1 非不定式误用

法则只适用于 $\frac{0}{0}$ 或 $\frac{\infty}{\infty}$ 型. 若把它用于非不定式, 就会给出错误结果.

- **错误示例**: 计算 $\lim_{x\to0}\frac{x+1}{x+2}$.
- **误解**: $\lim\frac{(x+1)'}{(x+2)'}=\lim\frac{1}{1}=1$.
- **正解**: 直接代入得 $\frac{0+1}{0+2}=\frac{1}{2}$.

16.14.2.2 导数极限不存在

若 $\lim\frac{f'(x)}{g'(x)}$ 不存在（振荡）, 原极限仍可能存在. 此时洛必达法则失效, 我们改用夹逼定理或定义法.

例题 16.80 计算 $\lim_{x\to\infty}\frac{x+\sin x}{x}$.

解 原极限存在且等于 1, 由 $|\sin x|\le1$ 与夹逼定理即得. 但若改用洛必达法则, 则得到

$$\lim_{x\to\infty}\frac{1+\cos x}{1}$$

由于 $\cos x$ 在 ∞ 处振荡, 该极限不存在. 此时法则失效.

16.14.2.3 循环与繁琐

对某些函数, 如 $\sqrt{x^2+1}$, 求导后形式可能更复杂, 甚至循环出现, 算不出结果. 而对因式分解就能解决的问题, 如 $\frac{x^2-1}{x-1}$, 用洛必达法则反而笨拙.

16.14.3 其他类型的不定式

对于 $0\cdot\infty$、$\infty-\infty$、1^∞、0^0、∞^0 等非标准型不定式, 必须先通过代数变形将其转化为 $\frac{0}{0}$ 或 $\frac{\infty}{\infty}$ 型.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.14.3.1 $0\cdot\infty$ 型

此型的处理办法是将乘积 $f\cdot g$ 变形为商 $\frac{f}{1/g}$ 或 $\frac{g}{1/f}$. 通常选择求导较简单的函数放在分子, 较复杂的倒数放在分母.

例题 16.81 计算 $\lim_{x\to 0^+} x\ln x$.

解 这是 $0\cdot(-\infty)$ 型. 将 x 下放:

$$\lim_{x\to 0^+}\frac{\ln x}{1/x}\quad(\text{型为}\frac{\infty}{\infty})$$

$$\overset{L'H}{=}\lim_{x\to 0^+}\frac{1/x}{-1/x^2}=\lim_{x\to 0^+}(-x)=0$$

16.14.3.2 $\infty-\infty$ 型

此型的处理办法是通分、提取公因式或有理化, 合并为一个分式.

知乎@余命数

例题 16.82 计算 $\lim_{x\to 0}(\frac{1}{\sin x}-\frac{1}{x})$.

解

知乎@余命数

$$\lim_{x\to 0}\frac{x-\sin x}{x\sin x}\quad(\text{型为}\frac{0}{0})$$

连续使用洛必达法则可得

$$\lim_{x\to 0}\frac{x-\sin x}{x\sin x}=\lim_{x\to 0}\frac{1-\cos x}{\sin x+x\cos x}=\lim_{x\to 0}\frac{\sin x}{2\cos x-x\sin x}=0.$$

16.14.3.3 幂指函数型 ($1^\infty, 0^0, \infty^0$)

此型的处理办法是**取对数**, 利用恒等式 $y=e^{\ln y}$. 设 $y=f(x)^{g(x)}$, 则 $\ln y=g(x)\ln f(x)$, 转化为 $0\cdot\infty$ 型.

例题 16.83 计算 $\lim_{x\to\infty}(1+\frac{1}{x})^x$.

知乎@余命数

解 令 $y=(1+\frac{1}{x})^x$, 取对数:

$$\ln y=x\ln(1+\frac{1}{x})=\frac{\ln(1+1/x)}{1/x}$$

令 $t=1/x$, 当 $x\to\infty$ 时 $t\to 0$:

$$\lim_{t\to 0}\frac{\ln(1+t)}{t}\overset{L'H}{=}\lim_{t\to 0}\frac{1/(1+t)}{1}=1$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $\lim \ln y = 1$, 原极限 $\lim y = e^1 = e$.

16.14.4 极限求解的策略总览

不同类型的极限问题适用不同的解法. 按方法的复杂度与适用场景, 常用的求解策略可分为三个层次.

知乎@余命数

16.14.4.1 方法分层

1. 整体观察层
   - **直接代入**: 先检查是否为不定式. 若分母不为零, 连续性保证极限值等于函数值.
   - **初等变形**: 包括因式分解、分子/分母有理化、提取最高次幂等, 适用于有理分式或简单根式.
   - **重要极限**: 利用标准结果, 如 $\lim_{x\to 0} \frac{\sin x}{x} = 1$ 或 $\lim_{x\to\infty}(1+\frac{1}{x})^x = e$.
2. 精度替换层
   - **等价无穷小替换**: $x \to 0$ 时利用 $\sin x \sim x, \ln(1+x) \sim x$ 等替换. 计算简便, 但通常仅适用于**乘除因子**, 加减法中使用需谨慎, 以防精度丢失.
   - **夹逼定理**: 处理含振荡项, 如 $\sin \frac{1}{x}$ 或无法直接计算的极限, 通过放缩确立上下界.
3. 系统求解层
   - **洛必达法则**: 适用于满足条件的 $\frac{0}{0}$ 或 $\frac{\infty}{\infty}$ 型不定式. 求导后形式简化时最为有效.
   - **泰勒展开**: 适用于 $x \to 0$ 且涉及加减消项的复杂组合. 相比等价无穷小, 能精确控制误差阶数.

16.14.4.2 实例对比

例题 16.84 计算 $\lim_{x\to 0} \frac{x-\sin x}{x^3}$.

解 法一采用洛必达法则. 该极限为 $\frac{0}{0}$ 型, 满足法则条件.

$$\lim_{x\to 0} \frac{x-\sin x}{x^3} = \lim_{x\to 0} \frac{1-\cos x}{3x^2} = \lim_{x\to 0} \frac{\sin x}{6x} = \lim_{x\to 0} \frac{\cos x}{6} = \frac{1}{6}$$

知乎@余命数

此过程需要连续使用三次洛必达法则.

法二采用泰勒展开. 利用 $\sin x = x - \frac{x^3}{6} + o(x^3)$ 代入分子:

$$\lim_{x\to 0} \frac{x-(x-\frac{x^3}{6}+o(x^3))}{x^3} = \lim_{x\to 0} \frac{\frac{x^3}{6}+o(x^3)}{x^3} = \frac{1}{6}$$

通过保留三阶项, 直接观察到分子的阶数特征.

16.14.4.3 方法的选择建议

知乎@余命数

- 对于**乘除形式**的因子, 优先考虑**等价无穷小替换**, 以简化表达式.
- 对于**加减形式**的不定式, 若求导后各项变得简单, 可使用**洛必达法则**.
- 对于**高阶无穷小相抵消**, 如 $x-\sin x$ 或涉及复合函数的复杂极限, **泰勒展开**通常是最稳健和系统的方法.

16.15 必要性探路与端点分析法 *

在运用导数处理不等式恒成立问题时, 例如求参数范围使得 $f(x)\geq k$ 在区间 I 上恒成立, 经典做法是求出 $f(x)$ 在区间 I 上的全局最小值 $f_{\min}$, 并要求 $f_{\min}\geq k$. 这种做法属于**充分性**构造, 直接验证最不利情形下的不等式成立.

许多问题中函数的导函数繁琐, 驻点不易求解, 或最值点落在端点. 此时全局最小值的求取变得困难或不可行, 需要换一种角度. **必要性分析**是这样的途径.

必要性分析反向提问, 若不等式恒成立, 它在哪些特殊点必须满足何种条件. 这种从结论反推初始条件的思路, 关注函数在边界点处的初值与变化趋势. 当临界情形发生在区间端点时, 该思路尤为有效, 称为**端点分析法**.

16.15.1 端点分析法则

考察如下模型. 已知函数 $g(x)$ 在区间 $[m,+\infty)$ 上满足 $g(m)=0$, 要求证或求参使得 $g(x)\geq 0$ 在此区间上恒成立.

此时函数图像从点 $(m,0)$ 出发. 要使后续函数值不小于零, 函数在出发点不能有向下的趋势, 即初始变化率必须为非负. 进一步地, 若函数的二阶导数恒正, 即变化率本身单调递增, 则这个初始趋势就足以决定函数的全程行为.

定理 16.22 (端点分析法则) 设函数 $g(x)$ 在区间 $[m,+\infty)$ 上连续, 在 $(m,+\infty)$ 上可导, 右导数 $g'_+(m)$ 存在, 且 $g'(x)$ 在 $(m,+\infty)$ 上单调递增并满足 $g'(x)\geq g'_+(m)$. 若 $g(m)=0$, 则不等式 $g(x)\geq 0$ 在区间 $[m,+\infty)$ 上恒成立的充分必要条件是 $g'_+(m)\geq 0$.

证明 导函数 $g'(x)$ 在 $(m,+\infty)$ 上单调递增且不低于端点右导数的条件, 可以通过 $g''(x)\geq 0$ 在 $(m,+\infty)$ 上恒成立并结合 $g'_+(m)$ 存在来保证.

充分性证明（证明 $g'_+(m)\geq 0 \implies g(x)\geq 0$）. 已知 $g'(x)\geq g'_+(m)$, 且 $g'_+(m)\geq 0$. 对于任意 $x>m$, 可得 $g'(x)\geq 0$. 即函数 $g(x)$ 在整个区间 $[m,+\infty)$ 上单调不减. 因此, 对于任意 $x\in[m,+\infty)$, 恒有 $g(x)\geq g(m)$. 结合已知条件 $g(m)=0$, 得 $g(x)\geq 0$ 在 $[m,+\infty)$ 上恒成立.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

必要性证明(证明 $g(x)\geq 0 \implies g'_+(m)\geq 0$). 回归右导数的定义. 函数 $g(x)$ 在端点 m 处的右导数定义为

$$g'_+(m)=\lim_{x\to m^+}\frac{g(x)-g(m)}{x-m}$$

根据题设, 对于所有 $x>m$, 恒有 $g(x)\geq 0$ 且 $g(m)=0$. 因此差商的分子 $g(x)-g(m)=g(x)\geq 0$, 分母 $x-m>0$. 故整个差商表达式 $\frac{g(x)-g(m)}{x-m}\geq 0$. 若该差商的极限为负数 $-\alpha$ ($\alpha>0$), 则按极限定义, 当 x 充分靠近 m 且 $x>m$ 时差商应小于 $-\alpha/2$, 这与差商恒非负矛盾. 因此 $g'_+(m)=\lim\limits_{x\to m^+}\frac{g(x)-g(m)}{x-m}\geq 0$.

充分性与必要性均得以证明. □

注 16.16 (物理学释义)
用物理学语言说, 一个从原点出发的质点, 若其初始速度非负, 且其加速度始终为正, 即速度单调递增, 则该质点将永远不会运动到坐标轴的负半轴. 此定理把涉及整个区间的恒成立问题, 转化为只需检验端点处导数值的问题, 简化了分析.

16.15.2 应用与思辨

例题 16.85 (2024 年全国新高考 I 卷) 已知函数 $f(x)=\ln\frac{x}{2-x}+ax+b(x-1)^3$.

1. 若 $b=0$, $f'(x)\geq 0$, 求 a 的取值范围.
2. 证明 $f(x)$ 的图像是中心对称图形. 若不等式 $f(x)>-2$ 成立的充分必要条件是 $1<x<2$, 求 b 的取值范围.

解 函数的定义域由 $\frac{x}{2-x}>0$ 确定, 解得 $x\in(0,2)$.

当 $b=0$ 时, $f(x)=\ln x-\ln(2-x)+ax$. 其导函数为 $f'(x)=\frac{1}{x}+\frac{1}{2-x}+a$. 要使 $f'(x)\geq 0$ 在 $(0,2)$ 上恒成立, 等价于 $a\geq -(\frac{1}{x}+\frac{1}{2-x})$ 恒成立. 令 $h(x)=-(\frac{1}{x}+\frac{1}{2-x})$. 需求 a 大于等于 $h(x)$ 的最大值. 对 $h(x)$ 求导得 $h'(x)=\frac{1}{x^2}-\frac{1}{(2-x)^2}$. 令 $h'(x)=0$, 得 $x=1$. 易知 $h(x)$ 在 $x=1$ 处取得最大值, $h_{\max}=h(1)=-2$. 故 a 的取值范围是 $[-2,+\infty)$.

为证明 $f(x)$ 的图像是中心对称图形, 考察 $f(x)+f(2-x)$.

$$\begin{aligned}f(x)+f(2-x)&=\left(\ln\frac{x}{2-x}+ax+b(x-1)^3\right)\\&\quad+\left(\ln\frac{2-x}{x}+a(2-x)+b(1-x)^3\right)\\&=0+(ax+2a-ax)+(b(x-1)^3-b(x-1)^3)\\&=2a\end{aligned}$$

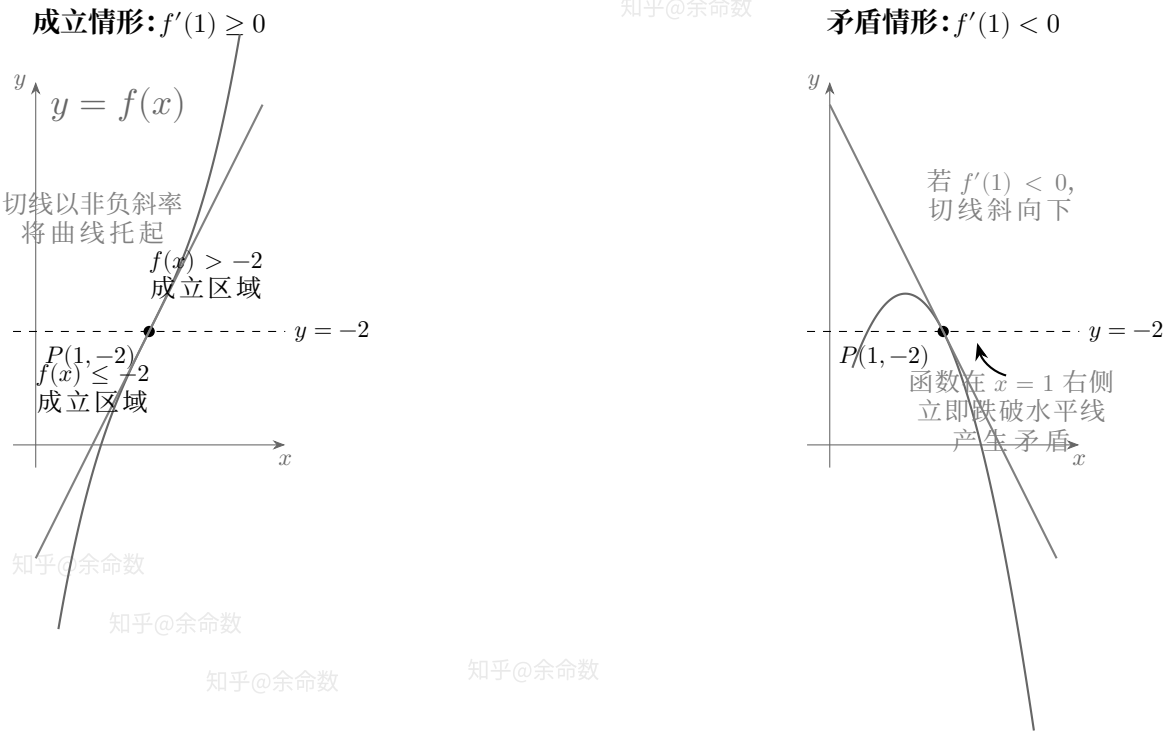


图 16.23　在临界点 $P(1,-2)$, 函数的初始趋势 $f'(1)$ 决定了其能否满足题设条件

因此, $f(x)$ 的图像关于点 $(1,a)$ 中心对称.

接着分析不等式条件. 题设“$f(x)>-2$ 成立的充要条件是 $x\in(1,2)$”蕴含了三个信息. (1) 对所有 $x\in(1,2)$, $f(x)>-2$ 成立. (2) 对所有 $x\in(0,1]$, $f(x)\leq -2$ 成立. (3) 在临界点 $x=1$ 处, 由连续性可知必有 $f(1)=-2$.

由 $f(1)=-2$ 可确定参数 a 的值. $f(1)=\ln\frac{1}{1}+a(1)+b(1-1)^3=a$. 故 $a=-2$.

此时条件明确. 当 $a=-2$ 时, 函数 $f(x)$ 在 $x=1$ 处函数值为 -2, 在 $(1,2)$ 上函数值大于 -2, 在 $(0,1]$ 上函数值不大于 -2. 故函数 $f(x)$ 必须在其定义域 $(0,2)$ 内是严格单调递增的.

因此, 问题转化为求解参数 b 的范围, 使得当 $a=-2$ 时, 函数 $f(x)$ 在 $(0,2)$ 上严格单调递增. 这等价于其导函数 $f'(x)\geq 0$ 在 $(0,2)$ 上恒成立, 且等号仅在孤立点处取得.

$f'(x)=\frac{1}{x}+\frac{1}{2-x}+a+3b(x-1)^2$. 代入 $a=-2$ 得, $f'(x)=\left(\frac{1}{x}+\frac{1}{2-x}-2\right)+3b(x-1)^2\geq 0$.

化简括号内的表达式. $\frac{1}{x}+\frac{1}{2-x}-2=\frac{2-x+x-2x(2-x)}{x(2-x)}=\frac{2(1-2x+x^2)}{x(2-x)}=\frac{2(x-1)^2}{x(2-x)}$.

于是, 不等式变为 $\frac{2(x-1)^2}{x(2-x)}+3b(x-1)^2\geq 0$. 当 $x\in(0,2)$ 且 $x\neq 1$ 时, 可约去正因子 $(x-1)^2$, 得到 $\frac{2}{x(2-x)}+3b\geq 0\implies 3b\geq -\frac{2}{x(2-x)}$.

此不等式需对所有 $x\in(0,1)\cup(1,2)$ 恒成立. 因此, $3b$ 必须大于或等于函数 $k(x)=-\frac{2}{x(2-x)}$ 在该定义域上的最大值(上确界).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $p(x)=x(2-x)$, 其在 $x=1$ 处取得最大值 $p(1)=1$. 故 $p(x)$ 在 $(0,2)$ 内的值域为 $(0,1]$. 因此, 函数 $k(x)=-2/p(x)$ 的值域为 $(-\infty,-2]$. 其上确界为 -2.

故须有 $3b\geq -2$, 即 $b\geq -\frac{2}{3}$.

最后检验, 当 $b\geq -2/3$ 时, $f'(x)=(x-1)^2\left(\frac{2}{x(2-x)}+3b\right)$. 由于 $\frac{2}{x(2-x)}\geq 2$, 故括号内表达式 $\geq 2+3(-\frac{2}{3})=0$. 等号仅在 $x=1$ 且 $b=-2/3$ 时成立. 因此 $f'(x)\geq 0$ 恒成立, 且仅在 $x=1$ 取零, 满足严格单调递增的条件.

b 的取值范围是 $[-\frac{2}{3},+\infty)$.

例题 16.86 (2023 年全国高考甲卷) 已知函数 $f(x)=ax-\dfrac{\sin x}{\cos^3 x}$, $x\in(0,\frac{\pi}{2})$.

1. 若 $a=8$, 讨论 $f(x)$ 的单调性;
2. 若 $f(x)<\sin 2x$ 恒成立, 求 a 的取值范围.

解 函数 $f(x)$ 的定义域为 $(0,\frac{\pi}{2})$. 注意到 $\frac{\sin x}{\cos^3 x}=\tan x\sec^2 x$, 其导数为 $\left(\frac{\sin x}{\cos^3 x}\right)'=\frac{1+2\sin^2 x}{\cos^4 x}$, 故

$$f'(x)=a-\frac{1+2\sin^2 x}{\cos^4 x}.$$

当 $a=8$ 时, $f'(x)=8-\dfrac{1+2\sin^2 x}{\cos^4 x}$. 令 $g(x)=\dfrac{1+2\sin^2 x}{\cos^4 x}$, 并作代换 $t=\sin x\in(0,1)$, 得 $h(t)=\dfrac{1+2t^2}{(1-t^2)^2}$. 求导得 $h'(t)=\frac{4t(2+t^2)}{(1-t^2)^3}>0$ 于 $(0,1)$, 故 h 严格递增, 从而 g 在 $(0,\frac{\pi}{2})$ 上严格递增.

令 $f'(x)=0$, 即 $g(x)=8$, 得

$$\frac{1+2\sin^2 x}{(1-\sin^2 x)^2}=8\implies 8\sin^4 x-18\sin^2 x+7=0.$$

解得 $\sin^2 x=\frac{1}{2}$(另一根 $\sin^2 x=7/4$ 舍去), 在定义域内唯一解为 $x=\frac{\pi}{4}$. 由 g 单调递增可知, 当 $x\in(0,\frac{\pi}{4})$ 时 $g(x)<8$, $f'(x)>0$; 当 $x\in(\frac{\pi}{4},\frac{\pi}{2})$ 时 $g(x)>8$, $f'(x)<0$. 故 f 在 $(0,\frac{\pi}{4})$ 上递增, 在 $(\frac{\pi}{4},\frac{\pi}{2})$ 上递减.

不等式 $f(x)<\sin 2x$ 恒成立, 即 $ax-\frac{\sin x}{\cos^3 x}<2\sin x\cos x$ 在 $(0,\frac{\pi}{2})$ 上恒成立. 构造 $G(x)=ax-\frac{\sin x}{\cos^3 x}-2\sin x\cos x$, 问题化为 $G(x)<0$ 恒成立. $\lim_{x\to 0^+}G(x)=0$; 虽 0 不在定义域内, 可将 G 连续延拓为 $G(0)=0$. 要求右侧函数值恒为负, 这给出右导数极限非正的必要条件.

$$G'(x)=a-\frac{1+2\sin^2 x}{\cos^4 x}-2\cos(2x),\qquad \lim_{x\to 0^+}G'(x)=a-3.$$

因而必须 $a-3\leq 0$, 即 $a\leq 3$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再证 $a \le 3$ 充分. 只需验证 $a = 3$ 时 $G(x) < 0$ 恒成立(若 $a < 3$ 则 G 更小). 当 $a = 3$ 时, $G'(x) = 3 - \frac{1+2\sin^2 x}{\cos^4 x} - 2\cos(2x)$. 令 $t = \cos^2 x \in (0,1)$, 则

$$k(t) = 5 - \frac{3-2t}{t^2} - 4t = 5 - \frac{3}{t^2} + \frac{2}{t} - 4t.$$

$k'(t) = \frac{6}{t^3} - \frac{2}{t^2} - 4 = \frac{p(t)}{t^3}$, 其中 $p(t) = -4t^3 - 2t + 6$. 因 $p'(t) = -12t^2 - 2 < 0$ 且 $p(1) = 0$, 故对 $t \in (0,1)$ 有 $p(t) > 0$, 从而 $k'(t) > 0$, k 严格递增. 于是 $k(t) < \lim_{t\to 1^-} k(t) = 0$, 即 $G'(x) = k(\cos^2 x) < 0$. 故 G 在 $(0, \frac{\pi}{2})$ 上严格递减; 结合 $\lim_{x\to 0^+} G(x) = 0$, 得 $G(x) < 0$.

因此 a 的取值范围是 $(-\infty, 3]$.

16.15.3 极限的局部保号性 *

时常需要由函数在某点的极限性质, 推断其在极限点附近的性态. 一个基本事实是, 若函数无限趋近于一个非零常数, 则在充分靠近极限点的邻域内, 函数值的符号与极限值的符号一致. 这一结论把极限与函数的局部行为联系起来, 现给出严格证明.

想象一个变量 $f(x)$ 的值正在无限地逼近一个正数 $L = 5$. 即 $f(x)$ 与 5 之间的距离 $|f(x)-5|$ 可以变得任意小. 如果我们要求这个距离小于 1, 即 $|f(x)-5| < 1$, 那么 $f(x)$ 的值就必须被限制在区间 $(4,6)$ 之内. 这个区间内的所有数显然都是正数.

该原理表明, 只要函数收敛于一个非零极限, 总能取一个足够小的误差范围, 将函数值限制在一个不包含零的区间内. 将此直观形式化即得极限的局部保号性定理.

定理 16.23 (极限的局部保号性) 设函数 $f(x)$ 在 $x \to x_0$ 时的极限存在且等于 L.

- 若 $L > 0$, 则必然存在点 x_0 的一个去心邻域 $U^\circ(x_0, \delta)$, 使得对于任意属于此邻域的 x, 恒有 $f(x) > 0$.
- 若 $L < 0$, 则必然存在点 x_0 的一个去心邻域 $U^\circ(x_0, \delta)$, 使得对于任意属于此邻域的 x, 恒有 $f(x) < 0$.

证明 以 $L > 0$ 的情形为例证明, $L < 0$ 的情况完全类似.

目标是证明存在一个正数 δ, 使得当 $0 < |x - x_0| < \delta$ 时, 恒有 $f(x) > 0$.

证明的要点在于恰当地选取 $\varepsilon - \delta$ 定义中的 ε. 取一个足够小的 ε, 使由不等式 $|f(x) - L| < \varepsilon$ 确定的区间 $(L-\varepsilon, L+\varepsilon)$ 完全位于 x 轴的上方, 即 $L - \varepsilon > 0$.

一个简洁有效的选择是令 $\varepsilon = \dfrac{L}{2}$. 由于 $L > 0$, ε 是一个合法的正数.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

根据极限的 $\varepsilon-\delta$ 定义, 对于选定的 $\varepsilon=L/2$, **必然存在**一个正数 δ, 使得当自变量 x 满足 $0<|x-x_0|<\delta$ 时, 不等式

$$|f(x)-L|<\frac{L}{2}$$

恒成立.

将此绝对值不等式展开并在三边同加 L、化简, 得

$$\begin{aligned}-\frac{L}{2}<f(x)-L<\frac{L}{2} &\iff L-\frac{L}{2}<f(x)<L+\frac{L}{2} \quad (\text{两边加 } L)\\ &\iff \frac{L}{2}<f(x)<\frac{3L}{2}.\end{aligned}$$

知乎@余命数

由其左半部分及 $L>0$ 知 $\frac{L}{2}>0$, 故 $f(x)>\frac{L}{2}>0$, 直接证明了 $f(x)>0$.

若 $L<0$, 取 $\varepsilon=-\frac{L}{2}>0$. 同理可得在某个去心邻域内

$$\frac{3L}{2}<f(x)<\frac{L}{2}.$$

知乎@余命数

由于 $\frac{L}{2}<0$, 故 $f(x)<0$.

至此, 两个符号情形均已证明. 证毕. □

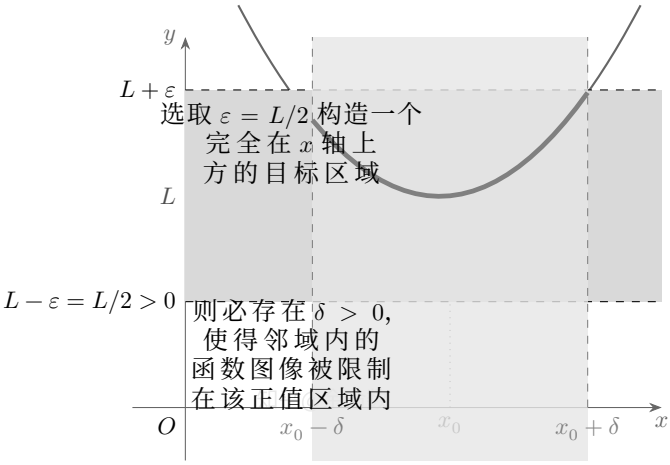


图 16.24 极限局部保号性的几何诠释

极限的局部保号性定理有一个重要推论, 处理的是极限与不等号的关系.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

推论 16.1 (极限的保不等式性) 设函数 $f(x)$ 在 $x \to x_0$ 时的极限存在. 如果在点 x_0 的某个去心邻域内, 恒有 $f(x) \geq 0$, 那么

$$\lim_{x \to x_0} f(x) \geq 0$$

同理, 若恒有 $f(x) \leq 0$, 那么 $\lim\limits_{x \to x_0} f(x) \leq 0$.

证明 采用反证法. 假设 $\lim\limits_{x \to x_0} f(x) = L < 0$. 根据刚刚证明的局部保号性定理, 若极限 L 为负, 则必然存在 x_0 的一个去心邻域, 使得对于此邻域内的所有 x, 都有 $f(x) < 0$.

将这个邻域与题设中使 $f(x) \geq 0$ 成立的去心邻域取交, 仍可得到一个非空的去心邻域. 在该交邻域内同一个 x 应同时满足 $f(x) < 0$ 与 $f(x) \geq 0$, 矛盾. 因此初始假设"$L < 0$"必定是错误的. 故必有 $L \geq 0$. 对 $f(x) \leq 0$ 的情形, 对 $-f(x)$ 应用刚证结论即可. □

注 16.17 (严格不等号的退化)
一个必须注意的细节是, 即使在邻域内恒有严格不等式 $f(x) > 0$, 其极限也只能保证是**非负的**, 即 $\lim\limits_{x \to x_0} f(x) \geq 0$, 而不能保证是严格大于零的.
例如, 考虑函数 $f(x) = x^2$ 在 $x \to 0$ 时的极限. 对于任意去心邻域 $U^\circ(0, \delta)$, 恒有 $f(x) = x^2 > 0$. 然而其极限为 $\lim\limits_{x \to 0} x^2 = 0$.

16.15.3.1 应用与思辨

极限的保不等式性虽然简单, 在许多证明中却是基石. 前文端点分析法则的论证, 也可用此推论重新表述其必要性部分.

在证明端点分析法则的必要性时, 曾用到下述事实, 若 $g(x) \geq 0$ 在 $[m, +\infty)$ 上恒成立且 $g(m) = 0$, 则右导数 $g'_+(m) \geq 0$. 证明过程如下.

$$g'_+(m) = \lim_{x \to m^+} \frac{g(x) - g(m)}{x - m}$$

我们论证了对于所有 $x > m$, 差商 $\frac{g(x)-g(m)}{x-m}$ 是一个非负的量. 接着, 由极限的保不等式性可得, 一个非负函数在 $x \to m^+$ 时的极限, 也必然是一个非负数.

$$\lim_{x \to m^+} \underbrace{\frac{g(x) - g(m)}{x - m}}_{\geq 0} \geq 0$$

这就是 $g'_+(m) \geq 0$ 的严格数学依据.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.15.4 方法的局限性与全局分析的回归

已认识到, 端点分析法的有效性严格依赖于函数一阶导数的单调性. 当此前提条件不被满足时, 该方法将由充分必要条件退化为仅具有必要性, 从而可能导出错误的结论.

端点分析法失效的根本原因在于函数的一阶导数 $g'(x)$ 在所考察的区间上不具备单调性. 考虑在区间 $[m,+\infty)$ 上, 函数 $g(x)$ 满足 $g(m)=0$ 的情形. 即使在端点处满足必要条件 $g'_+(m)\geq 0$, 若 $g'(x)$ 在区间 $(m,+\infty)$ 内出现下降, 则可能存在以下过程:

- $g'(x)$ 的值在初始为正或零之后, 随 x 的增大而减小.
- 在区间的某个内部点, $g'(x)$ 的值可能由正转负, 即原函数 $g(x)$ 的单调性由增转减, 形成一个局部极大值.
- 若此后 $g'(x)$ 持续为负, 函数 $g(x)$ 将进入单调递减阶段. 如果此单调递减的趋势足够明显, 函数 $g(x)$ 的值完全可能下降至小于其初始值 $g(m)=0$.

最终, 函数可能在区间的内部某点 x_0 取得一个极小值 $g(x_0)<0$. 这个极小值在端点分析中是不可见的, 我们称之为**隐蔽极小值**.

定理 16.24 (端点分析法适用性判别法则) 设函数 $g(x)$ 在区间 $[m,+\infty)$ 上连续, 在 $(m,+\infty)$ 上二阶可导, 右导数 $g'_+(m)$ 存在, 且满足 $g(m)=0$.

1. 若 $g''(x)$ 在区间 $(m,+\infty)$ 内恒有 $g''(x)\geq 0$, 则 $g'(x)$ 在该区间上单调递增. 此时, 不等式 $g(x)\geq 0$ 在 $[m,+\infty)$ 上恒成立的充分必要条件是 $g'_+(m)\geq 0$.
2. 若不能保证 $g'(x)$ 在该区间上单调递增, 则端点分析法一般失效, 条件 $g'_+(m)\geq 0$ 仍为不等式 $g(x)\geq 0$ 恒成立的**必要条件**, 但不再能直接作为充分条件. 问题的解决必须转为对 $g(x)$ 在区间 $[m,+\infty)$ 上的全局最小值分析.

证明 **证明第 1 条(适用情形)**: 只证明 $g''(x)\geq 0$ 的情况. **充分性**: 假设 $g'_+(m)\geq 0$. 由于 $g''(x)\geq 0$, 故 $g'(x)$ 在 $(m,+\infty)$ 上单调递增. 对任意 $x>m$, 取 $h\in(0,x-m)$. 由拉格朗日中值定理, 存在 $\xi_h\in(m,m+h)$ 使得

$$\frac{g(m+h)-g(m)}{h}=g'(\xi_h)\leq g'(x).$$

令 $h\to 0^+$, 得 $g'_+(m)\leq g'(x)$. 因而 $g'(x)\geq 0$ 对所有 $x>m$ 成立, g 在 $[m,+\infty)$ 上单调不减. 所以对于任意 $x\geq m$, 都有 $g(x)\geq g(m)=0$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

必要性：假设 $g(x)\geq 0$ 在 $[m,+\infty)$ 上恒成立. 对所有 $x>m$，有

$$\frac{g(x)-g(m)}{x-m}\geq 0.$$

令 $x\to m^+$，由右导数存在可知 $g'_+(m)\geq 0$.

证明第 2 条（失效情形）：必要性不依赖于 g' 的单调性，上段差商论证已经说明只要 $g(x)\geq 0$ 恒成立，就必有 $g'_+(m)\geq 0$. 但充分性会失效. 例如在 $[0,+\infty)$ 上取 $g(x)=x-x^2$. 则 $g(0)=0$，$g'_+(0)=1>0$，但 $g(2)=-2<0$. 因此，在缺少 g' 单调递增保障时，端点导数非负不能保证全区间非负. □

例题 16.87　已知函数 $f(x)=e^x+ax^2-x$. 当 $x\geq 0$ 时，$f(x)\geq \frac{1}{2}x^3+1$，求 a 的取值范围.

解　原不等式等价于 $e^x+ax^2-x-\frac{1}{2}x^3-1\geq 0$ 在区间 $[0,+\infty)$ 上恒成立. 构造辅助函数 $g(x)=e^x+ax^2-x-\frac{1}{2}x^3-1$. 问题转化为求解参数 a 的范围，使得 $g(x)\geq 0$ 对所有 $x\in[0,+\infty)$ 成立.

计算函数在端点 $x=0$ 的值，$g(0)=e^0-1=0$. 此点对任意 a 均满足不等式，因此需要分析 $x>0$ 的情形.

当 $x>0$ 时，原不等式可等价变形为

$$a\geq A(x):=\frac{1+x+\frac{1}{2}x^3-e^x}{x^2}.$$

因此，所求条件就是 a 不小于 $A(x)$ 在 $(0,+\infty)$ 上的最大值.

对 $A(x)$ 求导，得

$$A'(x)=\frac{N(x)}{x^3},\qquad N(x)=(2-x)e^x+\frac{1}{2}x^3-x-2.$$

由于 $x^3>0$，$A'(x)$ 的符号由 $N(x)$ 决定. 继续求导得

$$N'(x)=(1-x)e^x+\frac{3}{2}x^2-1,\qquad N''(x)=x(3-e^x).$$

因此 $N''(x)>0$ 在 $(0,\ln 3)$ 上成立，$N''(x)<0$ 在 $(\ln 3,+\infty)$ 上成立，所以 $N'(x)$ 先增后减. 又 $N'(0)=0$、$N'(1)=\frac{1}{2}>0$、$N'(2)=5-e^2<0$，故 $N'(x)$ 在 $(\ln 3,2)$ 内有唯一零点. 于是 $N(x)$ 先增后减. 结合 $N(0)=0$ 与 $N(2)=0$，可得 $N(x)>0$ 在 $(0,2)$ 上成立，$N(x)<0$ 在 $(2,+\infty)$ 上成立.

所以 $A(x)$ 在 $(0,2)$ 上递增，在 $(2,+\infty)$ 上递减，最大值为

$$A(2)=\frac{1+2+\frac{1}{2}\cdot 8-e^2}{4}=\frac{7-e^2}{4}.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

因此原不等式恒成立的充分必要条件为

$$a \geq \frac{7-e^2}{4}.$$

所求 a 的取值范围是 $[\frac{7-e^2}{4}, +\infty)$.

16.15.5 何时有效, 何时失效

已看到, 端点分析法的有效性依赖于一阶导数的单调性. 应用之前检验适用性, 是判断结论正确的前提.

回顾端点分析法所处理的模型, 即证明 $g(x) \geq 0$ 在 $[m, +\infty)$ 上恒成立, 其中已知 $g(m) = 0$. 其要点在于, 若函数在出发点 $(m, 0)$ 的初始速度 $g'(m)$ 非负, 且其加速度始终使该速度不为负, 则函数值永远不会跌破零.

这里的加速度由二阶导数 $g''(x)$ 刻画, 它决定了一阶导数 $g'(x)$ 的单调性.

注 16.18 (端点分析法适用性判别三步法)
对于端点值为零的不等式恒成立问题, 判断端点分析法是否适用的可操作步骤如下.

1. **构造函数与求导.** 构造满足 $g(m) = 0$ 的辅助函数 $g(x)$, 并计算其一阶导数 $g'(x)$ 和二阶导数 $g''(x)$.
2. **检验二阶导数符号.** 分析 $g''(x)$ 在区间 $(m, +\infty)$ 上的符号.
   - **若 $g''(x)$ 恒为非负.** 故 $g'(x)$ 在该区间上单调递增. 此时, **端点分析法完全适用**. 不等式 $g(x) \geq 0$ 恒成立的**充分必要条件**就是 $g'(m) \geq 0$.
   - **若 $g''(x)$ 不能保证 $g'(x)$ 单调递增.** 例如 $g''(x)$ 符号不定时, **端点分析法失效**. 条件 $g'(m) \geq 0$ 仅仅是一个**必要条件**, 不再是充分条件.
3. **确定做法.** 若适用, 则检验 $g'(m)$ 的符号直接得出结论; 若失效, 则必须放弃端点分析, 回归到求解 $g'(x)$ 的零点, 分析 $g(x)$ 的单调性以寻找其在 $[m, +\infty)$ 上的**全局最小值**.

用一个具体例子说明. 当 $g'(x)$ 不单调时, 端点分析法会给出错误结论.

例题 16.88 已知函数 $f(x) = \sin x - x + \frac{1}{6}x^3$. 若不等式 $f(x) \geq ax^4$ 在区间 $[0, +\infty)$ 上恒成立, 求实数 a 的取值范围.

解 原不等式等价于 $\sin x - x + \frac{1}{6}x^3 - ax^4 \geq 0$ 在 $[0, +\infty)$ 上恒成立. 构造辅助函数 $g(x) = \sin x - x + \frac{1}{6}x^3 - ax^4$, 则 $g(0) = 0$, 这是端点值为零的恒成立问题.

为使 g 从 $g(0) = 0$ 出发后始终非负, 可先在端点处逐阶考察导数, 以获取参数的必要条件.

$$g'(x) = \cos x - 1 + \frac{1}{2}x^2 - 4ax^3, \qquad g'(0) = 0,$$

$$g''(x) = -\sin x + x - 12ax^2, \qquad g''(0) = 0,$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$g'''(x)=-\cos x+1-24ax, \qquad g'''(0)=0,$$
$$g''''(x)=\sin x-24a, \qquad g''''(0)=-24a.$$

g 在 $x=0$ 附近的局部性态由第一个非零高阶导数决定;要使 $x=0$ 成为局部极小值点,需 $g''''(0)\geq 0$,即 $a\leq 0$.

上述结论仅来自端点的局部信息. 当 $a>0$ 时 $g''(x)$ 在 $(0,+\infty)$ 上符号不定,端点分析法所依赖的导函数单调性并不成立,因而 $a\leq 0$ 只是必要条件,是否充分尚须全局论证. 下按 a 的符号分两种情形.

若 $a\leq 0$,则对一切 $x\geq 0$ 有 $-ax^4\geq 0$,从而

$$g(x)=\left(\sin x-x+\frac{1}{6}x^3\right)-ax^4\geq \sin x-x+\frac{1}{6}x^3.$$

记 $\phi(x)=\sin x-x+\frac{1}{6}x^3$. 逐阶求导.

$$\phi(0)=0,$$
$$\phi'(x)=\cos x-1+\frac{1}{2}x^2, \qquad \phi'(0)=0,$$
$$\phi''(x)=-\sin x+x, \qquad \phi''(0)=0,$$
$$\phi'''(x)=1-\cos x.$$

对 $x\geq 0$ 有 $\cos x\leq 1$,故 $\phi'''(x)\geq 0$,从而 ϕ'' 在 $[0,+\infty)$ 上单调递增,$\phi''(x)\geq \phi''(0)=0$;同理 $\phi'(x)\geq 0$,$\phi(x)\geq 0$. 于是 $a\leq 0$ 时 $g(x)\geq \phi(x)\geq 0$,不等式恒成立.

若 $a>0$,需说明必有某处 g 取负值. 将 g 写成

$$g(x)=x^4\left(\frac{\sin x}{x^4}-\frac{1}{x^3}+\frac{1}{6x}-a\right).$$

由 $-1\leq \sin x\leq 1$ 与夹逼定理,$\lim_{x\to+\infty}\frac{\sin x}{x^4}=0$;又 $\lim_{x\to+\infty}\frac{1}{x^3}=\lim_{x\to+\infty}\frac{1}{6x}=0$,故括号内表达式的极限为 $-a<0$. 由极限的保号性,存在 $M>0$,使得当 $x>M$ 时括号内为负;此时 $x^4>0$,因而 $g(x)<0$. 故 $a>0$ 时不等式不恒成立.

综合两种情形,所求 a 的取值范围是 $(-\infty,0]$.

例题 16.89　求实数 a 的取值范围,使得对于任意 $x\geq 0$,恒有不等式 $x^3-3x^2+ax\geq 0$ 成立.

解　先构造辅助函数并作初步分析. 构造辅助函数 $g(x)=x^3-3x^2+ax$. 考察函数在端点 $x=0$ 的值.

$$g(0)=0^3-3(0)^2+a(0)=0$$

这是一个端点值为零的恒成立问题.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再检验方法的适用性. 计算一阶与二阶导数, 得

$$g'(x) = 3x^2 - 6x + a$$

$$g''(x) = 6x - 6 = 6(x-1)$$

二阶导数 $g''(x)$ 在区间 $(0,+\infty)$ 上的符号是不定的. 当 $x \in (0,1)$ 时, $g''(x) < 0$, 导函数 $g'(x)$ 单调递减. 当 $x \in (1,+\infty)$ 时, $g''(x) > 0$, 导函数 $g'(x)$ 单调递增.

由于 $g'(x)$ 在考察的区间上不具备单调性, 故**端点分析法在此问题中失效**.

最后剖析失效原因并回到正确方法. 从"$g(x) \geq 0$ 恒成立"只能推出必要条件 $g'(0) \geq 0$.

$$g'(0) = 3(0)^2 - 6(0) + a = a \geq 0$$

若误将此必要条件当作充分条件, 会得到错误答案 $a \geq 0$.

验证此答案不成立. 取 $a = 2$ (满足 $a \geq 0$), 则 $g(x) = x^3 - 3x^2 + 2x = x(x-1)(x-2)$. 当 $x \in (1,2)$ 时, $g(x) < 0$, 与题设 $g(x) \geq 0$ 恒成立相矛盾.

这个矛盾的根源在于, 初始速度 $g'(0) = a \geq 0$ 虽然保证函数在 $x = 0$ 附近是增加的, 但由于 $g'(x)$ 在 $(0,1)$ 上递减, 速度 $g'(x)$ 持续减小, 甚至可能在某点变为负值, 从而使函数 $g(x)$ 转为递减. 后续的递减可能使函数值跌至 x 轴之下, 形成一个端点处不可见的隐蔽极小值.

因此, 须放弃端点分析, 回归到对函数全局最小值的分析. 原不等式即 $x(x^2 - 3x + a) \geq 0$. 由于 $x \geq 0$, 问题等价于 $h(x) = x^2 - 3x + a \geq 0$ 在 $[0,+\infty)$ 上恒成立.

函数 $h(x)$ 是一个开口向上的二次函数, 其对称轴为 $x = \frac{3}{2}$. 由于对称轴位于区间 $[0,+\infty)$ 内部, 所以函数 $h(x)$ 在此区间上的最小值在顶点处取得.

$$h_{\min} = h\left(\frac{3}{2}\right) = \left(\frac{3}{2}\right)^2 - 3\left(\frac{3}{2}\right) + a = \frac{9}{4} - \frac{9}{2} + a = a - \frac{9}{4}$$

为使 $h(x) \geq 0$ 恒成立, 须有其最小值非负,

$$a - \frac{9}{4} \geq 0 \implies a \geq \frac{9}{4}$$

所求实数 a 的取值范围是 $[\frac{9}{4},+\infty)$.

注 16.19 (结论与警示)
此例说明机械套用端点分析法的危险性. 在处理端点值为零的恒成立问题时, 分析二阶导数 $g''(x)$ 的符号是判断方法是否适用的判据. 若 $g''(x)$ 符号恒定, 则端点分析可通往正确答案; 若 $g''(x)$ 符号不定, 则须警惕隐蔽极小值的存在, 并回归到全局最值分析.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数



16.15.6 本节习题

知乎@余命数

16.45. 求实数 a 的取值范围, 使得对于任意 $x \geq 0$, 恒有不等式 $e^x \geq ax+1$ 成立.

16.46. 求实数 a 的取值范围, 使得对于任意 $x \geq 1$, 恒有不等式 $a(x-1) \geq \ln x$ 成立.

16.47. 求实数 a 的取值范围, 使得对于任意 $x \in [0, \frac{\pi}{2}]$, 恒有不等式 $ax \geq \sin x$ 成立.

16.48. 已知函数 $f(x) = e^x - \frac{1}{2}x^2 - ax - 1$. 若 $f(x) \geq 0$ 对任意 $x \geq 0$ 恒成立, 求实数 a 的取值范围.

16.49. 求实数 a 的取值范围, 使得对于任意 $x \geq 0$, 恒有不等式 $x^3 - 3x^2 + ax \geq 0$ 成立. 请首先检验端点分析法在此问题中能用与否, 然后再求解.

知乎@余命数

16.50. 已知函数 $f(x) = e^x + ax^2 - x$. 若当 $x \geq 0$ 时, 不等式 $f(x) \geq \frac{1}{2}x^3 + 1$ 恒成立, 求实数 a 的取值范围. 请首先检验端点分析法在此问题中能用与否, 然后再求解.

16.15.7 高考真题选讲

知乎@余命数 知乎@余命数

以下三题选自历年高考真题, 涉及端点效应与切线放缩两类方法.

例题 16.90 (2010 年海南宁夏卷理科数学第 21 题) 设 $f(x) = e^x - 1 - x - ax^2$.

1. 当 $a = 0$ 时, 求 $f(x)$ 的单调区间;
2. 若当 $x \geq 0$ 时 $f(x) \geq 0$, 求 a 的取值范围.

解 (1)$a = 0$ 时的单调区间

知乎@余命数

$f(x) = e^x - 1 - x$, $f'(x) = e^x - 1$. 当 $x < 0$ 时 $f'(x) < 0$, 当 $x > 0$ 时 $f'(x) > 0$, 故 $f(x)$ 在 $(-\infty, 0)$ 上单调递减, 在 $(0, +\infty)$ 上单调递增.

知乎@余命数

(2) 求 a 的取值范围

$f(0) = 0$、$f'(x) = e^x - 1 - 2ax$、$f'(0) = 0$. 由 $e^x \geq 1 + x$ 得

$$f'(x) \geq (1+x) - 1 - 2ax = (1-2a)x$$

当 $a \leq \dfrac{1}{2}$ 时, $(1-2a)x \geq 0$ 对 $x \geq 0$ 成立, 故 $f'(x) \geq 0$ 在 $[0, +\infty)$ 上成立, $f(x) \geq f(0) = 0$.

当 $a > \dfrac{1}{2}$ 时, $f''(0) = 1 - 2a < 0$, $f'(x)$ 在 $x = 0$ 附近递减. 由 $f'(0) = 0$ 知存在 $\delta > 0$ 使 $f'(x) < 0$ 在 $(0, \delta)$ 上成立, 从而 $f(x) < f(0) = 0$, 矛盾.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

所求 a 的取值范围是 $\left(-\infty,\dfrac{1}{2}\right]$.

知乎@余命数

例题 16.91 (2018 年全国 □ 卷文科数学第 21 题) 已知函数 $f(x)=ae^x-\ln x-1$.

1. 设 $x=2$ 是 $f(x)$ 的极值点, 求 a 并求 $f(x)$ 的单调区间;
2. 证明: 当 $a\geq\dfrac{1}{e}$ 时, $f(x)\geq 0$.

解 (1) 求 a 与单调区间

$f'(x)=ae^x-\dfrac{1}{x}$. 由 $f'(2)=0$ 得 $a=\dfrac{1}{2e^2}$, 此时 $f'(x)=\dfrac{e^{x-2}}{2}-\dfrac{1}{x}$. 当 $0<x<2$ 时 $f'(x)<0$, 当 $x>2$ 时 $f'(x)>0$, 故 $f(x)$ 在 $(0,2)$ 上单调递减, 在 $(2,+\infty)$ 上单调递增.

知乎@余命数

(2) 证明 $f(x)\geq 0$

当 $a\geq\dfrac{1}{e}$ 时, $ae^x\geq e^{x-1}$, 故

$$f(x)\geq e^{x-1}-\ln x-1$$

知乎@余命数

由 $e^{x-1}\geq x$ 与 $\ln x\leq x-1$ 得

$$e^{x-1}-\ln x-1\geq x-(x-1)-1=0$$

知乎@余命数　知乎@余命数

故 $f(x)\geq 0$.

例题 16.92 (2020 年山东卷数学第 21 题) 已知函数 $f(x)=ae^{x-1}-\ln x+\ln a$.

1. 当 $a=e$ 时, 求曲线 $y=f(x)$ 在点 $(1,f(1))$ 处的切线与两坐标轴围成的三角形的面积;
2. 若 $f(x)\geq 1$, 求 a 的取值范围.

知乎@余命数

解 (1) 切线与面积

$a=e$ 时 $f(x)=e^x-\ln x+1$、$f(1)=e+1$、$f'(x)=e^x-\dfrac{1}{x}$、$f'(1)=e-1$. 切线方程为

$$y-(e+1)=(e-1)(x-1)$$

令 $x=0$ 得 $y=2$, 令 $y=0$ 得 $x=-\dfrac{2}{e-1}$, 三角形面积为 $\dfrac{2}{e-1}$.

(2) 求 a 的取值范围

必要性: $f(1)=a+\ln a\geq 1$. 令 $g(a)=a+\ln a-1$、$g'(a)=1+\dfrac{1}{a}>0$、$g(1)=0$, 故 $a\geq 1$.

充分性：当 $a \geq 1$ 时，$ae^{x-1} \geq e^{x-1}$ 且 $\ln a \geq 0$，故

$$f(x) \geq e^{x-1} - \ln x$$

由 $e^{x-1} \geq x$ 与 $\ln x \leq x-1$ 得

$$e^{x-1} - \ln x \geq x - (x-1) = 1$$

故 $f(x) \geq 1$.

所求 a 的取值范围是 $[1, +\infty)$.

16.16 极值点偏移问题

处理双零点或双极值点问题时，常遇到一类偏离对称性的难题.

例如，已知函数 $f(x)$ 有两个零点 x_1, x_2，题目要求证明 $x_1 + x_2 > 2x_0$，其中 x_0 为极值点. 若函数关于 $x = x_0$ 对称，如二次函数，有 $x_1 + x_2 = 2x_0$. 多数超越函数不具备这种对称性，中点 $\frac{x_1+x_2}{2}$ 因此偏离极值点 x_0. 这一现象称为极值点偏移.

现给出三种方法，从通用的构造法到变量代换.

16.16.1 问题的引入与识别

16.16.1.1 典型情景

极值点偏移问题通常以以下两种形式出现.

1. **零点偏移**. 函数 $f(x)$ 有两个零点 x_1, x_2，证明 $x_1 + x_2 > 2x_0$ 或 $x_1 x_2 > x_0^2$.
2. **极值偏移**. 函数 $f(x)$ 有两个极值点 x_1, x_2，即导函数的零点，证明 $f(x_1) + f(x_2) > C$.

16.16.1.2 特征与典例

此类问题源于函数的伪对称性，极值点两侧的图像相似，增长或下降的速度却并不一致. 常见的模型函数如 xe^x、$x\ln x$、$\frac{e^x}{x}$ 等，其导数包含指数或对数项，由此产生非对称性.

例题 16.93 已知函数 $f(x) = xe^x$. 若 $f(x_1) = f(x_2)$ 且 $x_1 \neq x_2$，求证 $x_1 + x_2 < -2$.

解 求导得 $f'(x) = (x+1)e^x$. 函数在 $x = -1$ 处取得极小值. 因为 $f(x)$ 在 $(-\infty, -1)$ 上单调递减，在 $(-1, +\infty)$ 上单调递增，且 $f(-1) = -\frac{1}{e}$，$\lim\limits_{x\to-\infty} f(x) = 0^-$，所以若 $f(x_1) = f(x_2)$ 且 $x_1 \neq x_2$，不妨设 $x_1 < x_2$，则必有 $x_1 < -1 < x_2 < 0$. 若 $f(x)$ 的图像关于直线 $x = -1$ 对称，则 $x_1 + x_2 = 2(-1) = -2$. 结论 < -2 暗示两根的中点偏

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

向极值点的左侧. 在 $x=-1$ 左侧 $(x\to-\infty)$ 函数趋于 0, 较为平缓; 右侧则增长极快. 为达到相同高度, 左侧的点 x_1 须比右侧的点 x_2 走得更远.

函数本身不对称时, 可人为构造一个对称结构作比较. 设 $x_1<x_0<x_2$, 其中 x_0 是极值点, 比较 x_2 与 x_1 关于 x_0 的对称点 $2x_0-x_1$ 的大小, 并构造辅助函数 $F(x)=f(x)-f(2x_0-x)$, 借助单调性推导 $f(x_1)$ 与 $f(2x_0-x_1)$ 的关系.

以典例 $f(x)=xe^x$ 为例, 证明 $x_1+x_2<-2$. 极值点 $x_0=-1$.

1. **设而不求**. 由上面的单调性分析,不妨设 $x_1<-1<x_2<0$. 要证 $x_1+x_2<-2$, 即证 $x_1<-2-x_2$. 由于 $f(x)$ 在 $(-\infty,-1)$ 上单调递减, 上述不等式等价于 $f(x_1)>f(-2-x_2)$. 又因为 $f(x_1)=f(x_2)$, 问题转化为证明, 对 $x_2\in(-1,0)$, 有 $f(x_2)>f(-2-x_2)$.
2. **构造辅助函数**. 令 $g(x)=f(x)-f(-2-x)$, 其中 $x>-1$. 目标是证明 $g(x)>0$.
3. **分析单调性**. 求导, 得

$$g'(x)=f'(x)-f'(-2-x)\cdot(-1)=f'(x)+f'(-2-x)$$

代入 $f'(x)=(x+1)e^x$, 得

$$\begin{aligned}g'(x)&=(x+1)e^x+(-2-x+1)e^{-2-x}\\&=(x+1)e^x-(x+1)e^{-2-x}\\&=(x+1)(e^x-e^{-2-x})\end{aligned}$$

因为 $x>-1$, 所以 $x+1>0$. 同时 $x>-2-x\iff 2x>-2\iff x>-1$, 由 e^x 单调递增可知 $e^x>e^{-2-x}$. 因此对 $x>-1$, 恒有 $g'(x)>0$.

4. **完成论证**. 函数 $g(x)$ 在 $(-1,+\infty)$ 上单调递增.

$$g(x)>g(-1)=f(-1)-f(-2-(-1))=f(-1)-f(-1)=0$$

即 $f(x)>f(-2-x)$. 令 $x=x_2$, 则 $f(x_2)>f(-2-x_2)$. 代回 $f(x_1)=f(x_2)$, 得 $f(x_1)>f(-2-x_2)$. 因为 x_1 和 $-2-x_2$ 都在单调递减区间 $(-\infty,-1)$ 内, 所以 $x_1<-2-x_2$, 即 $x_1+x_2<-2$. 证毕.

16.16.2 对数均值不等式与变换主元法

对于指数或对数型函数, 常可通过代数变形, 把双变量问题转化为单变量不等式.

16.16.2.1 变换主元的思路

利用 $f(x_1)=f(x_2)$, 把 x_1,x_2 用一个新变量 t(通常设 $t=x_1/x_2$ 或 $t=x_1-x_2$)表示, 从而将待证不等式转化为关于 t 的函数不等式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.16.2.2 操作步骤

仍以 $f(x)=xe^x$ 为例.

1. **代数变形**. 由 $x_1e^{x_1}=x_2e^{x_2}$, 两边取对数, 注意 x_1,x_2 均为负数, 取对数需取绝对值, 或移项变形. 变形为 $\frac{x_1}{x_2}=e^{x_2-x_1}$. 或更直接地: $x_1+\ln(-x_1)=x_2+\ln(-x_2)$(设 $x_1,x_2<0$). 整理得 $x_2-x_1=\ln(-x_1)-\ln(-x_2)=\ln(\frac{x_1}{x_2})$.
2. **变换主元**. 令 $t=\frac{x_1}{x_2}$. 因为 $x_1<-1<x_2<0$, 所以 $t>1$. 将 $x_2-x_1=\ln t$ 与 $\frac{x_1}{x_2}=t$ 联立求解.

$$x_2=\frac{\ln t}{1-t},\quad x_1=\frac{t\ln t}{1-t}$$

3. **构造新函数**, 得要证 $x_1+x_2<-2$, 即证

$$\frac{(t+1)\ln t}{1-t}<-2.$$

由于 $1-t<0$, 两边同乘 $1-t$ 后不等号反向, 得到

$$(t+1)\ln t>2(t-1).$$

设 $h(t)=(t+1)\ln t-2(t-1)$ $(t>1)$. 则

$$h'(t)=\ln t+\frac{1}{t}-1,\quad h''(t)=\frac{t-1}{t^2}>0.$$

又 $h'(1)=0$, 故 $h'(t)>0$ $(t>1)$, 从而 $h(t)>h(1)=0$. 因此 $(t+1)\ln t>2(t-1)$, 即 $x_1+x_2<-2$.

16.16.3 泰勒展开与高阶导数分析法

极值点偏移源于函数在极值点两侧弯曲程度不同, 而弯曲程度由高阶导数控制. 以下从这一角度给出一种分析方法.

16.16.3.1 泰勒展开的思路

把函数 $f(x)$ 在极值点 x_0 处展开. 设 $x_1=x_0-\delta_1, x_2=x_0+\delta_2$ $(\delta_1,\delta_2>0)$. 由泰勒公式.

$$f(x)\approx f(x_0)+\frac{f''(x_0)}{2}(x-x_0)^2+\frac{f'''(x_0)}{6}(x-x_0)^3$$

代入 x_1,x_2 并利用 $f(x_1)=f(x_2)$:

$$\frac{f''(x_0)}{2}\delta_1^2-\frac{f'''(x_0)}{6}\delta_1^3\approx\frac{f''(x_0)}{2}\delta_2^2+\frac{f'''(x_0)}{6}\delta_2^3$$

这暗示 δ_1 与 δ_2 的大小关系取决于三阶导数 $f'''(x_0)$ 的符号.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

16.16.3.2 判定准则

- 若 $f'''(x_0)=0$, 函数在极值点附近局部对称性强, 偏移极小.
- 若 $f'''(x_0)>0$, 通常意味着极值点右侧更陡峭, 为达到相同高度, 右侧所需的距离 δ_2 更小, 即 $\delta_2<\delta_1$. 此时 $x_1+x_2=2x_0-\delta_1+\delta_2<2x_0$.
- 若 $f'''(x_0)<0$, 同理可得 $x_1+x_2>2x_0$.

注 16.20
此方法不能直接作为解答步骤, 但能迅速指明不等式的方向, 可用来检验猜想并指导构造辅助函数. 例如在典例中, $f'''(-1)$ 易算得为正, 故可判断 $x_1+x_2<-2$.

16.16.4 方法选择

- **构造对称函数法**. 逻辑严密, 适用范围广, 是解答题的标准写法.
- **变换主元法**. 技巧性强, 对 $x, e^x, \ln x$ 组合的问题往往有效, 可作为备选.
- **泰勒与高阶导数分析**. 用于预判结论方向与洞察偏移成因.

16.16.5 高考真题选讲

例题 16.94 (2016 年全国 □ 卷理科数学第 21 题) 已知函数 $f(x)=(x-2)e^x+a(x-1)^2$ 有两个零点.

1. 求 a 的取值范围;
2. 设 x_1,x_2 是 $f(x)$ 的两个零点, 证明 $x_1+x_2<2$.

解 (1) 求 a 的取值范围

$x=1$ 不是 $f(x)$ 的零点. 当 $x\neq 1$ 时, $f(x)=0$ 等价于

$$-a=\frac{(x-2)e^x}{(x-1)^2}$$

设 $g(x)=\dfrac{(x-2)e^x}{(x-1)^2}$, 求导得

$$g'(x)=\frac{e^x(x^2-4x+5)}{(x-1)^3}$$

其中 $x^2-4x+5=(x-2)^2+1>0$, 故 $g(x)$ 在 $(-\infty,1)$ 上单调递减, 在 $(1,+\infty)$ 上单调递增.

$g(x)$ 在 $(-\infty,1)$ 上的值域为 $(-\infty,0)$, 在 $(1,+\infty)$ 上的值域为 $(-\infty,+\infty)$. 方程 $g(x)=-a$ 有两个实根当且仅当 $-a<0$, 即 $a>0$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

(2) 证明 $x_1+x_2<2$

由 (1) 不妨设 $x_1<1<x_2$, 需证 $x_2<2-x_1$. 因 $g(x)$ 在 $(1,+\infty)$ 上单调递增, 只需证 $g(x_2)<g(2-x_1)$, 即 $g(x_1)<g(2-x_1)$.

构造 $h(x)=(x-1)^2(g(x)-g(2-x))$, 即 $h(x)=e^x(x-2)+e^{2-x}x$, 其中 $x<1$. 求导得

$$h'(x)=(x-1)\left(e^x-e^{2-x}\right)$$

当 $x<1$ 时, $x<2-x$, 故 $e^x<e^{2-x}$, 从而 $h'(x)>0$, $h(x)$ 在 $(-\infty,1)$ 上单调递增. 因此 $h(x)<h(1)=0$, 即 $g(x)<g(2-x)$ 对一切 $x<1$ 成立.

取 $x=x_1$ 即得 $g(x_1)<g(2-x_1)$, 故 $x_1+x_2<2$.

16.17 牛顿迭代法

求方程 $f(x)=0$ 的根贯穿代数与数值分析. 五次及以上的多项式方程不存在一般的求根公式, 超越方程如 $e^x+x=0$ 更无法用初等运算求解. 牛顿迭代法借助导数与切线, 把求根问题转化为一个可以反复执行的迭代公式, 在根附近以极快的速度收敛.

第7.11节 已经从不动点的角度提及了牛顿法, 并指出巴比伦算法是其特例. 本节从切线逼近的几何直观出发, 推导迭代公式, 分析收敛速度, 并讨论其失效情形.

16.17.1 从切线到迭代

16.17.1.1 迭代公式的推导

设 $f(x)$ 在根 r 附近可导, 且 $f'(r)\neq 0$. 我们已知一个近似值 x_n, 希望找到一个更好的近似值 x_{n+1}.

在 第16.5节 中, 曲线 $y=f(x)$ 在点 $(x_n,f(x_n))$ 处的切线斜率为 $f'(x_n)$, 切线方程为

$$y-f(x_n)=f'(x_n)(x-x_n) \tag{16-23}$$

切线在点 $(x_n,f(x_n))$ 附近是曲线的最佳线性近似. 当 x_n 接近根 r 时, 切线与 x 轴的交点应当比 x_n 更接近 r.

令 (16-23) 中 $y=0$, 解出切线与 x 轴的交点, 作为下一次的近似值 x_{n+1}:

$$-f(x_n)=f'(x_n)(x_{n+1}-x_n)$$

整理得牛顿迭代公式.

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)} \tag{16-24}$$

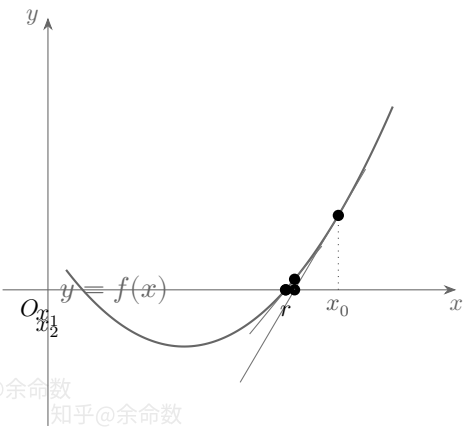
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

注 16.21 (迭代函数的不动点)
牛顿迭代 (16-24) 是一个形如 $x_{n+1}=g(x_n)$ 的迭代, 其中迭代函数为 $g(x)=x-\frac{f(x)}{f'(x)}$. 若 r 是 $f(x)=0$ 的根且 $f'(r)\neq 0$, 则 $g(r)=r-\frac{0}{f'(r)}=r$, 故根 r 是 g 的不动点. 这与 第7.11节 中不动点迭代的框架一致.

知乎@余命数

16.17.1.2 几何直观

牛顿法的每一步, 都是用切线的根代替曲线的根. 当初始值足够接近真实根, 且函数在该处性质良好时, 切线的根会迅速逼近曲线的根.



知乎@余命数　知乎@余命数　知乎@余命数

图 16.25　牛顿迭代法的几何过程, 在 x_0 处作切线, 其与 x 轴的交点 x_1 更接近根 r; 重复此过程, x_2 进一步逼近 r.

知乎@余命数

16.17.2 计算平方根的实例

16.17.2.1 迭代公式的导出

求 $\sqrt{2}$ 等价于解方程 $f(x)=x^2-2=0$. 由 $f'(x)=2x$, 代入 (16-24) 得

$$x_{n+1}=x_n-\frac{x_n^2-2}{2x_n}=\frac{1}{2}\left(x_n+\frac{2}{x_n}\right) \tag{16-25}$$

知乎@余命数

这是 第7.11节 中提到的巴比伦算法. 第16.1节 已证明 $\sqrt{2}$ 是无理数, 无法写成有限小数或分数, 而迭代 (16-25) 给出了一条逼近它的具体路径.

知乎@余命数

16.17.2.2 迭代数据

取初始值 $x_0=1$, 逐次代入 (16-25).

知乎@余命数

步数 n	近似值 x_n	误差 $\lvert x_n-\sqrt{2}\rvert$	正确位数
0	1	4.14×10^{-1}	0
1	$\dfrac{3}{2}=1.5$	8.58×10^{-2}	1
2	$\dfrac{17}{12}\approx1.41667$	2.45×10^{-3}	2
3	$\dfrac{577}{408}\approx1.414216$	2.12×10^{-6}	5
4	$\dfrac{665857}{470832}\approx1.414213562$	1.59×10^{-12}	11

$\sqrt{2}\approx1.41421356237$. 仅迭代四步，误差已小于 10^{-11}. 正确位数从 0 增长到 $1,2,5,11$, 每步大约翻倍. 这种误差按平方缩小的收敛速度称为**二阶收敛**.

16.17.3 收敛速度

16.17.3.1 误差递推的推导

设 r 是 $f(x)=0$ 的单根, 即 $f(r)=0$ 且 $f'(r)\neq0$. 记误差 $e_n=x_n-r$. 将 $f(x_n)$ 在 r 处作泰勒展开, 见 第16.12节:

$$f(x_n)=f(r)+f'(r)\,e_n+\frac{f''(r)}{2}\,e_n^2+O(e_n^3) \tag{16-26}$$

因 $f(r)=0$, 上式化为

$$f(x_n)=f'(r)\,e_n+\frac{f''(r)}{2}\,e_n^2+O(e_n^3)$$

对 $f'(x_n)$ 同样在 r 处展开:

$$f'(x_n)=f'(r)+f''(r)\,e_n+O(e_n^2)$$

代入牛顿迭代公式 (16-24), $x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}$, 即 $e_{n+1}=e_n-\frac{f(x_n)}{f'(x_n)}$:

$$\begin{aligned} e_{n+1}&=e_n-\frac{f'(r)\,e_n+\frac{f''(r)}{2}\,e_n^2+O(e_n^3)}{f'(r)+f''(r)\,e_n+O(e_n^2)}\\ &=e_n-\frac{e_n\left(f'(r)+\frac{f''(r)}{2}\,e_n+O(e_n^2)\right)}{f'(r)\left(1+\frac{f''(r)}{f'(r)}\,e_n+O(e_n^2)\right)} \end{aligned}$$

利用 $\frac{1}{1+u}=1-u+O(u^2)$, 分母的倒数展开为

$$\frac{1}{f'(r)}\left(1-\frac{f''(r)}{f'(r)}\,e_n+O(e_n^2)\right)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代回并整理, e_n 的一阶项相互抵消:

$$\begin{aligned} e_{n+1} &= e_n - e_n\left(1+\frac{f''(r)}{2f'(r)}\,e_n\right)\left(1-\frac{f''(r)}{f'(r)}\,e_n+O(e_n^2)\right) \\ &= e_n - e_n\left(1-\frac{f''(r)}{2f'(r)}\,e_n+O(e_n^2)\right) \\ &= \frac{f''(r)}{2f'(r)}\,e_n^2+O(e_n^3) \end{aligned}$$

当 e_n 足够小, 高阶项 $O(e_n^3)$ 可忽略, 得到误差递推的主项.

定理 16.25 (牛顿法的二阶收敛) 设 r 是 $f(x)=0$ 的单根 ($f'(r)\neq 0$), $f''(r)$ 存在. 当初始值 x_0 足够接近 r 时, 牛顿迭代 (16-24) 的误差满足

$$e_{n+1} \approx \frac{f''(r)}{2f'(r)}\,e_n^2 \tag{16-27}$$

即误差按平方缩小, 称为**二阶收敛** (quadratic convergence).

(16-27) 中的常数 $C=\frac{f''(r)}{2f'(r)}$ 由函数在根处的曲率与斜率决定. 若 C 不太大, 每步迭代将误差从 e_n 压缩到 $C\,e_n^2$. 当 $e_n<1$ 时, e_n^2 远小于 e_n, 误差迅速趋于零.

16.17.3.2 正确位数翻倍

二阶收敛的一个直接推论, 每步迭代后, 近似值的有效位数大约翻倍. 若 x_n 有 d 位正确数字 ($e_n\sim 10^{-d}$), 则 $e_{n+1}\sim C\cdot 10^{-2d}$, 即 x_{n+1} 约有 $2d$ 位正确数字, 当 C 接近 1 时.

计算 $\sqrt{2}$ 的数据印证了这一点, 正确位数从 $0\to 1\to 2\to 5\to 11$, 基本呈翻倍增长.

16.17.3.3 与不动点迭代的对比

第7.11节 讨论了一般的不动点迭代 $x_{n+1}=g(x_n)$. 当 g 在不动点 r 处可导且 $g'(r)\neq 0$ 时, 误差满足

$$e_{n+1}\approx g'(r)\,e_n$$

误差按固定比例线性缩小, 称为**一阶收敛** (linear convergence). 每步将误差乘以常数 $|g'(r)|<1$, 正确位数按固定增量增长.

牛顿法的迭代函数 $g(x)=x-\frac{f(x)}{f'(x)}$ 在单根 r 处满足 $g'(r)=0$ (因 $g'(x)=1-\frac{f'(x)^2-f(x)f''(x)}{f'(x)^2}$, 代入 $f(r)=0$ 得 $g'(r)=1-1=0$). 由于 $g'(r)=0$, 误差递推中一阶项消失, 收敛速度从线性提升到二阶.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

	一阶收敛(不动点迭代)	二阶收敛(牛顿法)
误差递推	$e_{n+1} \approx g'(r)\, e_n$	$e_{n+1} \approx C\, e_n^2$
正确位数增长	每步固定增量	每步约翻倍
收敛条件	$\lvert g'(r)\rvert < 1$	$f'(r) \neq 0$, x_0 足够接近 r

16.17.4 失效情形

牛顿法并非万能. 以下三种情形会导致迭代失败或收敛变慢.

16.17.4.1 切线水平导致的迭代中断

当迭代到某点 x_n 恰好使 $f'(x_n) = 0$ 时, 切线水平, 与 x 轴平行, 没有交点. 公式 (16-24) 中分母为零, 迭代无法进行.

例如, 对 $f(x) = x^3 - x$ 用牛顿法求根 $r = 0$. 若初始值取 $x_0 = 0$, 则 $f'(0) = 3(0)^2 - 1 = -1 \neq 0$, 迭代正常. 但若取 $x_0 = 1/\sqrt{3}$, 则 $f'(x_0) = 3 \cdot \frac{1}{3} - 1 = 0$, 切线水平, 迭代中断.

16.17.4.2 初始值远离根时不收敛或收敛到错误的根

牛顿法只保证在根附近收敛. 当初始值 x_0 离真实根太远时, 切线可能将下一步引向远离根的方向, 甚至发散.

此外, 当 $f(x) = 0$ 有多个根时, 牛顿法收敛到哪一个根取决于初始值的位置. 从 $x_0 = 10$ 出发求 $f(x) = x^3 - 2x - 5 = 0$ 的根, 迭代会收敛到某个根, 但未必是所求的那个.

16.17.4.3 重根处的收敛退化为线性

当 r 是 $f(x) = 0$ 的 m 重根 ($m \geq 2$), 即 $f(r) = f'(r) = \cdots = f^{(m-1)}(r) = 0$ 但 $f^{(m)}(r) \neq 0$ 时, 牛顿法的二阶收敛条件 $f'(r) \neq 0$ 不再满足.

以 $f(x) = x^2$ 为例, $r = 0$ 是二重根, $f'(0) = 0$. 牛顿迭代公式为

$$x_{n+1} = x_n - \frac{x_n^2}{2x_n} = \frac{x_n}{2}$$

误差 $e_{n+1} = \frac{e_n}{2}$, 退化为线性收敛, 每步误差减半. 达到 10^{-12} 精度需要约 40 步, 远多于单根情形的 4 步.

对重根情形, 可采用修正牛顿法 $x_{n+1} = x_n - m\,\frac{f(x_n)}{f'(x_n)}$, 其中 m 为根的重数, 恢复二阶收敛. 但这要求事先知道根的重数, 实际应用中往往难以满足.

注 16.22 (实践中的策略)

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

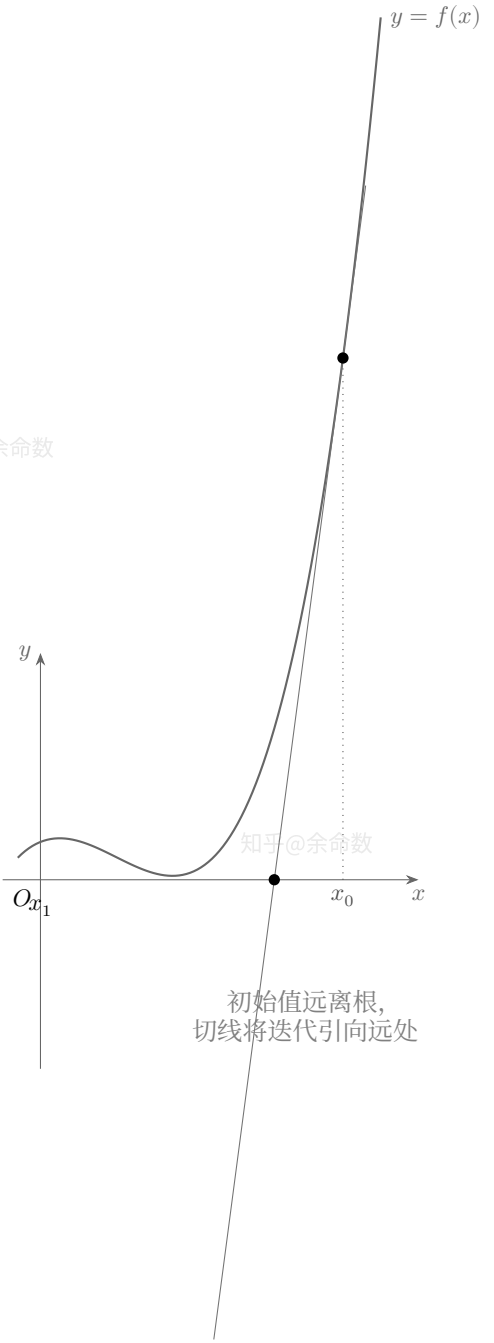


图 16.26　初始值远离根时, 切线的交点可能比 x_0 更远离根, 迭代发散.

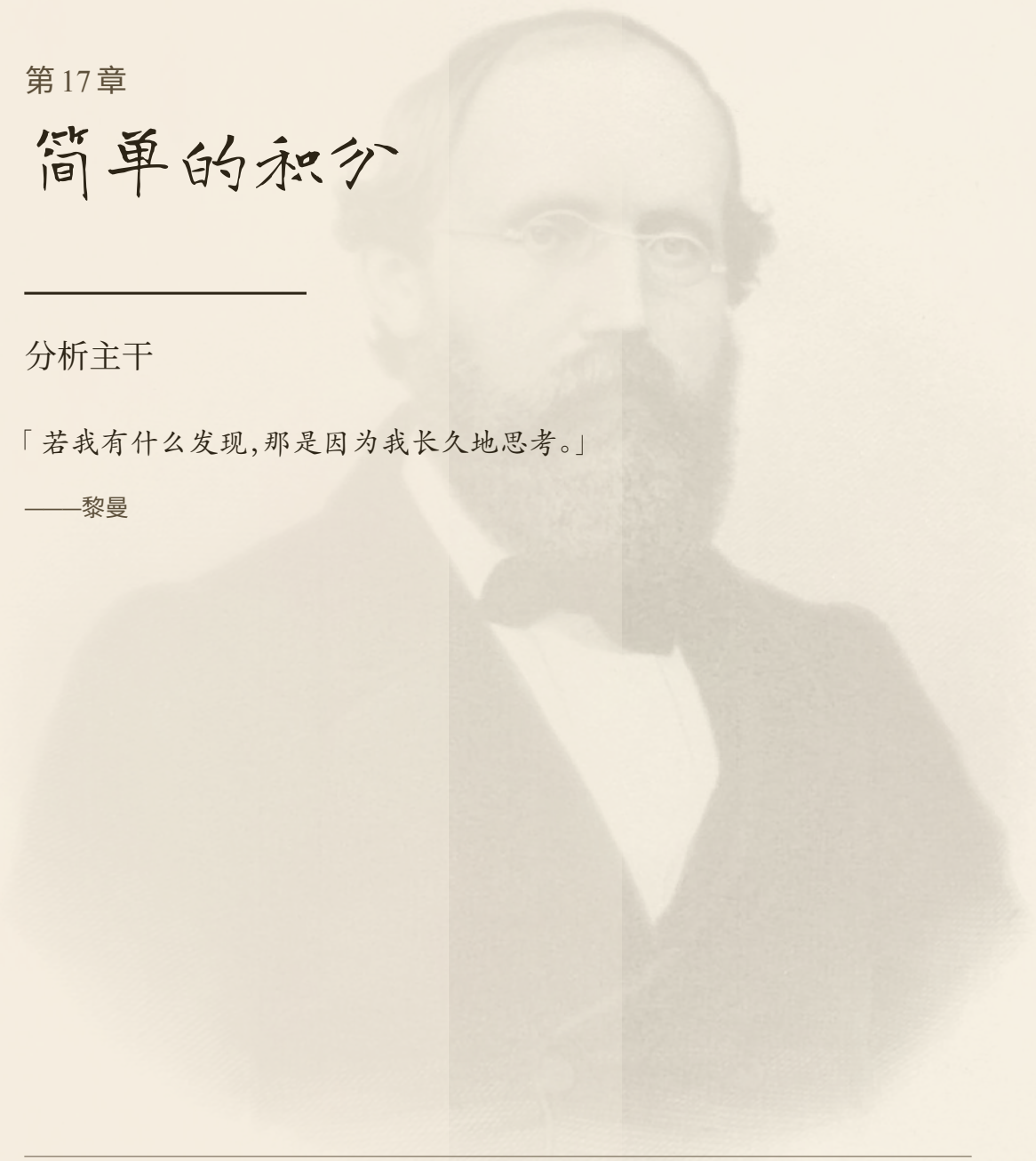
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

牛顿法要求计算 $f'(x)$，当导数难以解析求出时，可用差商 $\frac{f(x_n)-f(x_{n-1})}{x_n-x_{n-1}}$ 代替，得到**割线法**. 割线法的收敛阶约为 1.618（黄金比例），介于一阶与二阶之间. 当导数计算成本高时，割线法是牛顿法的常用替代.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 17 章

简单的积分



分析主干

「若我有什么发现,那是因为我长久地思考。」

——黎曼

第十七章 简单的积分

在微分学里我们做过这样的事, 已知一辆车的位移 $s(t)$, 对它求导, 就得到瞬时速度 $v(t)=s'(t)$, 从位置在哪儿算出跑得多快.

现在把这个问题掉过头来, 只看一张速度记录 $v(t)$, 要反推出位移 $s(t)$.

试一个最简单的例子. 一辆车以 $v(t)=2t$ 的速度匀加速行驶, 凭对求导公式的熟悉, 立刻能猜到位移是 $s(t)=t^2$, 因为 $(t^2)'=2t$. 同理 t^2+1, t^2-3 也都说得通, 给 t^2 加上任意常数, 求导后常数消去, 得到的还是 $2t$.

但如果速度是 $v(t)=\sin t$ 或者 $v(t)=e^{-t^2}$ 这种钟形速度曲线, 情形就不同了. 前者还好猜, $-\cos t$ 就是一个答案; 后者则不然, 我们手中所有熟悉的函数都不足以描述它的位移.

这种从变化率反推总量的运算, 是求导的逆运算, 叫做**积分**. 与求导几乎总能套公式不同, 积分更像一门手艺, 它有一本基本的辞典, 有若干分类技巧, 也有相当一批做不出来的情形. 后一点会反复提醒我们, 积分远比微分来得丰富.

17.1 原函数与不定积分

17.1.1 原函数

从章首的例子继续. 已知速度 $v(t)=2t$, 想反推位移 $s(t)$, 就是找一个函数, 它的导数恰好是 $2t$. 我们立刻想到 t^2, 因为 $(t^2)'=2t$. 但也不止 t^2 一个: t^2+1、t^2-3、t^2+100 都对, 给 t^2 加上任意常数后再求导, 常数都会消掉, 结果仍然是 $2t$. 既然找导数等于已知函数的函数这件事如此自然, 我们给它一个名字.

定义 17.1 (原函数) 如果在区间 I 上, 可导函数 $F(x)$ 的导函数为 $f(x)$, 即对于任意 $x\in I$, 都有

$$F'(x)=f(x)$$

那么, 我们称函数 $F(x)$ 为 $f(x)$ 在区间 I 上的一个**原函数**.

举几个简单例子. 我们熟悉 $(-\cos x)'=\sin x$, 所以 $-\cos x$ 是 $\sin x$ 的一个原函数. 同理 e^x 是 e^x 自己的原函数, 因为指数函数求导不变; $\frac{x^{n+1}}{n+1}$ 是 x^n 的原函数 ($n\neq -1$), 因为幂函数求导一次指数降一次, 系数 $n+1$ 恰好被消去. 这些都是从求导公式倒着读得到的.

回看章首那个速度例子, $v(t)=2t$ 的原函数不止一个, t^2、t^2+1、t^2-3、... 都是. 背后是一条普遍规律, 一个函数的原函数若存在, 就一定有无穷多个, 而且彼

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此之间只差一个常数. 所有原函数是否都具有这种形式, 下面这条定理给出肯定的回答.

定理 17.1 (原函数的结构) 若函数 $F(x)$ 是 $f(x)$ 在一个**区间** I 上的一个原函数, 则 $f(x)$ 在该区间上的任意其他原函数 $G(x)$ 必定具有形式 $G(x) = F(x) + C$, 其中 C 为某个常数.

证明 设 $F(x)$ 和 $G(x)$ 是 $f(x)$ 在区间 I 上的任意两个原函数, 由定义有 $F'(x) = f(x)$ 与 $G'(x) = f(x)$. 考虑辅助函数 $\Phi(x) = G(x) - F(x)$, 对其求导得

$$\Phi'(x) = G'(x) - F'(x) = f(x) - f(x) = 0, \quad x \in I.$$

在导数章节中已证(作为拉格朗日中值定理的推论), 在连通区间上导数恒为零的函数必为常数, 故存在常数 C 使对所有 $x \in I$ 有 $\Phi(x) = C$, 即 $G(x) - F(x) = C$, 亦即 $G(x) = F(x) + C$. □

定理强调区间二字, 因为它依赖的事实是导数恒为零的函数必为常数, 而后者只在连成一片的定义域上才成立. 若定义域断成不相交的几段, 那每一段上原函数的差都可以是不同的常数, 拼起来就不再是单常数差异了. 例如 $f(x) = \frac{1}{x^2}$, 它的定义域 $(-\infty, 0) \cup (0, \infty)$ 就是这样断开的两段. 一个原函数是 $F(x) = -\frac{1}{x}$. 那么 $G(x) = \begin{cases} -\frac{1}{x} + 2 & x > 0 \\ -\frac{1}{x} - 5 & x < 0 \end{cases}$ 也是 $f(x)$ 的原函数, 但两段上的常数差异不同, $G - F$ 已经不是一个统一的常数 C 了.

17.1.2 不定积分

既然一个函数的原函数构成一整族 $F(x) + C$, 我们就用一个专门的符号把它们打包写在一起.

定义 17.2 (不定积分) 函数 $f(x)$ 在区间 I 上的全体原函数所构成的集合, 称为 $f(x)$ 在 I 上的**不定积分**, 记作

$$\int f(x)\,dx$$

如果 $F(x)$ 是 $f(x)$ 的一个原函数, 那么我们写出

$$\int f(x)\,dx = F(x) + C \tag{17-1}$$

这里的 C 被称为**积分常数**, 它代表了函数族中所有可能的常数项.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

积分与微分是一对互逆运算, 但互逆的方式在两个方向上不对称.

$$\frac{d}{dx}\left(\int f(x)\,dx\right)=\frac{d}{dx}(F(x)+C)=f(x)$$

先积分再微分, 常数 C 在求导时被消掉, 完全还原为 $f(x)$;

$$\int\left(\frac{d}{dx}F(x)\right)\,dx=\int F'(x)\,dx=F(x)+C$$

但反过来, 先微分再积分, 由于求导那一头本来就会丢掉常数信息, 末端就多出一个待定的 C. 这种一头干净、一头多出常数的不对称, 正是微分与积分作为逆运算的具体表现.

几何上, 不定积分 $y=F(x)+C$ 既是一族曲线, 又能这样看, 在平面上每一点 (x_0,y_0) 画一个斜率恰好为 $f(x_0)$ 的小箭头, 全部箭头组成一片**方向场**, 又称斜率场. 而 $y=F(x)+C$ 这一族积分曲线, 正是顺着这片方向场走出来的曲线, 曲线在每一点的切线方向, 都与该点箭头的方向吻合. 图 17.1 中 $y'=2x$ 的方向场与积分曲线族 $y=x^2+C$ 显示了这一点, 把其中任意一条曲线沿 y 方向平移, 都还是这个方向场的解.

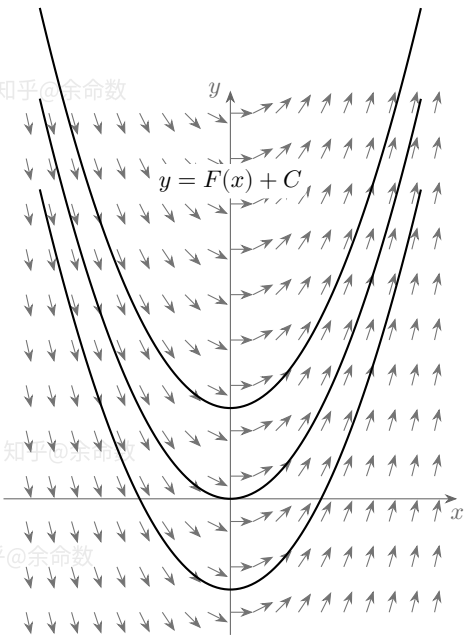


图 17.1 方向场 $y'=2x$ 与积分曲线族 $y=x^2+C$

17.1.3 达布定理与原函数不存在判据

已定义原函数的概念, 但尚未讨论其存在的条件, 何种函数必定拥有原函数、何种函数绝无原函数, 由导函数的一条基本性质给出判据.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

达布定理揭示了导函数的一个基本性质, 即便导函数不连续, 它也必须具备介值性质. 这条性质提供了一个原函数不存在的判据.

定理 17.2 (达布定理 (导数介值定理)) 若函数 $F(x)$ 在闭区间 $[a,b]$ 上可导, 则其导函数 $F'(x)$ 在 $[a,b]$ 上具有介值性质. 对于任意 $y_1, y_2 \in [a,b]$ 且 $y_1 < y_2$, 以及任意介于 $F'(y_1)$ 与 $F'(y_2)$ 之间的实数值 k, 必存在 $\xi \in (y_1, y_2)$, 使得

$$F'(\xi) = k.$$

证明 不妨设 $F'(y_1) < k < F'(y_2)$. 引入辅助函数 $G(x) = F(x) - kx$. 则 G 在 $[a,b]$ 上可导, 且

$$G'(y_1) = F'(y_1) - k < 0, \quad G'(y_2) = F'(y_2) - k > 0.$$

由导数符号的局部行为, 存在 y_1 右侧附近的点 x_1 使 $G(x_1) < G(y_1)$, 存在 y_2 左侧附近的点 x_2 使 $G(x_2) < G(y_2)$. 因此 G 在 $[a,b]$ 上的最小值不在端点 y_1 或 y_2 处取得, 而必落在内部某点 $\xi \in (y_1, y_2)$ 处. 由费马引理, 在该内部极值点处 $G'(\xi) = 0$, 即 $F'(\xi) = k$. □

注 17.1
达布定理的等价表述是. 导函数的像集 $F'([a,b])$ 必为一个区间; 因此导函数即便不连续, 也不能跳跃取值, 必须填满其值之间的所有空隙. 这一性质比连续性弱, 却比任意取值更强.

推论 17.1 (含第一类间断点的函数无原函数) 若函数 $f(x)$ 在区间 I 上存在第一类(可去或跳跃)间断点 c, 则 $f(x)$ 在 I 上**不存在原函数**, 即不存在任何在 I 上处处可导的 F 使得 $F'(x) = f(x)$ 在 I 上处处成立.

证明 反设存在这样的 F 满足 $F'(x) = f(x)$ 在 I 上处处成立. 由达布定理, F' 在 I 的任何子闭区间上必具有介值性质. 但 f 在 c 处有跳跃或可去间断点, 意味着 f 在 c 左右两侧的极限值, 或极限与函数值之间存在间隙, 无法填满该处的所有介值, 与 F' 必具介值性相矛盾, 故假设不成立. □

此推论给出了何种函数无原函数的负向判据. 配合 第17.1.4小节 的正向存在定理, 我们便对原函数何时存在有了较为完整的认识. 连续性保证原函数存在, 含第一类间断点则排除原函数的存在. 需补充说明, 含第二类间断点的函数可能存在原函数, 例如 $f(x) = 2x\sin(1/x) - \cos(1/x)(x \neq 0)$, $f(0) = 0$, 在 $x = 0$ 处有第二类间断, 却仍存在原函数 $F(x) = x^2\sin(1/x)(F(0) = 0)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

17.1.4 原函数的存在性

已定义原函数, 并获得了达布定理给出的非存在判据. 存在性的**充分条件**仍待确立. 并非每个函数都拥有原函数, 但微积分的一个基石性结论(微积分基本定理的第一部分)保证了以下事实.

定理 17.3 (原函数存在定理) 若函数 $f(x)$ 在区间 I 上**连续**, 则 $f(x)$ 在该区间上必存在原函数.

证明 取定区间 I 中一点 a. 对任意 $x \in I$, 定义

$$F(x) = \int_a^x f(t)\,dt,$$

其中当 $x < a$ 时按约定 $\int_a^x = -\int_x^a$. 由于 f 在区间 I 上连续, 微积分基本定理第一部分保证由上述变上限定积分定义的函数 F 在 I 上可导, 且

$$F'(x) = f(x).$$

因而 F 就是 f 在 I 上的一个原函数. 这说明连续性足以保证原函数存在. □

这个定理将函数的分析性质(连续性)与积分运算的可能性直接联系起来, 确保对一切遇到的连续函数谈论其不定积分都有意义.

17.2 不定积分

上一节写下了不定积分的记号 $\int f(x)\,dx = F(x) + C$, 也就是 $f(x)$ 的全体原函数. 有了记号, 还需要弄清它的来历, 并建立一份由求导公式反推而来的基本积分表.

17.2.1 记号的来历与基本积分表

莱布尼茨引入的积分符号 $\int f(x)\,dx$ 具有内在含义. 积分号 $\int$ 是字母 S 的拉长, 代表求和 (Summa), 这预示了不定积分与定积分(黎曼和的极限)之间的内在联系. 被积表达式 $f(x)dx$ 则暗示变量的微分 dx 在运算中的重要地位, 这一地位在换元积分法中将充分显现.

正如求导对加法、数乘可以分配, 即 $(kF)' = kF'$, $(F+G)' = F' + G'$, 积分也享有这两条性质, 称为**线性性**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 17.4 (不定积分的线性性质) 若函数 $f(x)$ 和 $g(x)$ 在同一区间上的原函数都存在, $F'(x)=f(x)$、$G'(x)=g(x)$、k 为任意常数, 则

1. **齐次性**: $\int kf(x)\,dx = kF(x)+C$
2. **可加性**: $\int [f(x)+g(x)]\,dx = F(x)+G(x)+C$

等价地, 在不定积分记号中, $\int kf(x)\,dx = k\int f(x)\,dx$ 与 $\int (f+g)\,dx = \int f\,dx + \int g\,dx$ 均按积分常数可重新合并的意义理解.

证明 证明依赖求导的线性性 $(kF)'=kF'$ 与 $(F+G)'=F'+G'$. 对 $kF(x)+C$ 求导得 $\frac{d}{dx}(kF(x)+C)=kF'(x)=kf(x)$, 故 $kF(x)+C$ 是 $kf(x)$ 的原函数; 由原函数结构定理, $kf(x)$ 的任意两个原函数只差常数, 此即全体原函数($k=0$ 时给出 $\int 0\,dx=C$), 齐次性得证. 同理, 对 $F(x)+G(x)+C$ 求导得 $\frac{d}{dx}(F(x)+G(x)+C)=F'(x)+G'(x)=f(x)+g(x)$, 故 $F(x)+G(x)+C$ 是 $f(x)+g(x)$ 的原函数, 再由原函数结构定理给出全体原函数, 可加性得证. □

知乎@余命数

这个定理表明, 在把相差常数的原函数视为同一类的意义下, 不定积分具有**线性结构**. 这允许将复杂函数的积分问题分解为若干个简单函数积分的线性组合, 是进行积分计算的根本法则.

17.2.2 基本积分表

既然积分是微分的逆运算, 那么已掌握的每一个求导公式, 反向审视, 都对应着一条积分公式. 这些基本公式是求解更复杂积分问题的基础.

17.2.2.1 幂函数与对数函数

幂函数求导法则 $(x^\alpha)'=\alpha x^{\alpha-1}$ 是我们最熟悉的法则之一. 为将其逆转以求 $\int x^n dx$, 需寻找一个函数, 其导数恰为 x^n. 注意到 $\left(x^{n+1}\right)'=(n+1)x^n$. 为抵消系数 $n+1$, 我们构造 $\frac{x^{n+1}}{n+1}$, 其导数为 $\left(\frac{x^{n+1}}{n+1}\right)'=x^n$. 此推导要求分母 $n+1\neq 0$, 即 $n\neq -1$.

知乎@余命数

$$\int x^n\,dx = \frac{x^{n+1}}{n+1}+C \quad (n\in\mathbb{R}, n\neq -1)$$

幂函数积分公式中 $n=-1$ 的情况构成了一个独特的缺口.$\int x^{-1}dx=\int \frac{1}{x}dx$ 的答案由一个全新的超越函数, 即自然对数函数, 来填补. 我们已知当 $x>0$ 时, $(\ln x)'=\frac{1}{x}$. 当 $x<0$ 时, $-x>0$, 函数 $\ln(-x)$ 有定义. 根据链式法则, $(\ln(-x))'=$

知乎@余命数

$\frac{1}{-x}\cdot(-1)=\frac{1}{x}$. 故, 无论 x 在哪个半轴, $\frac{1}{x}$ 的原函数都与对数相关. 为统一表达, 我们引入绝对值.

$$\int \frac{1}{x}\,dx = \ln|x| + C$$

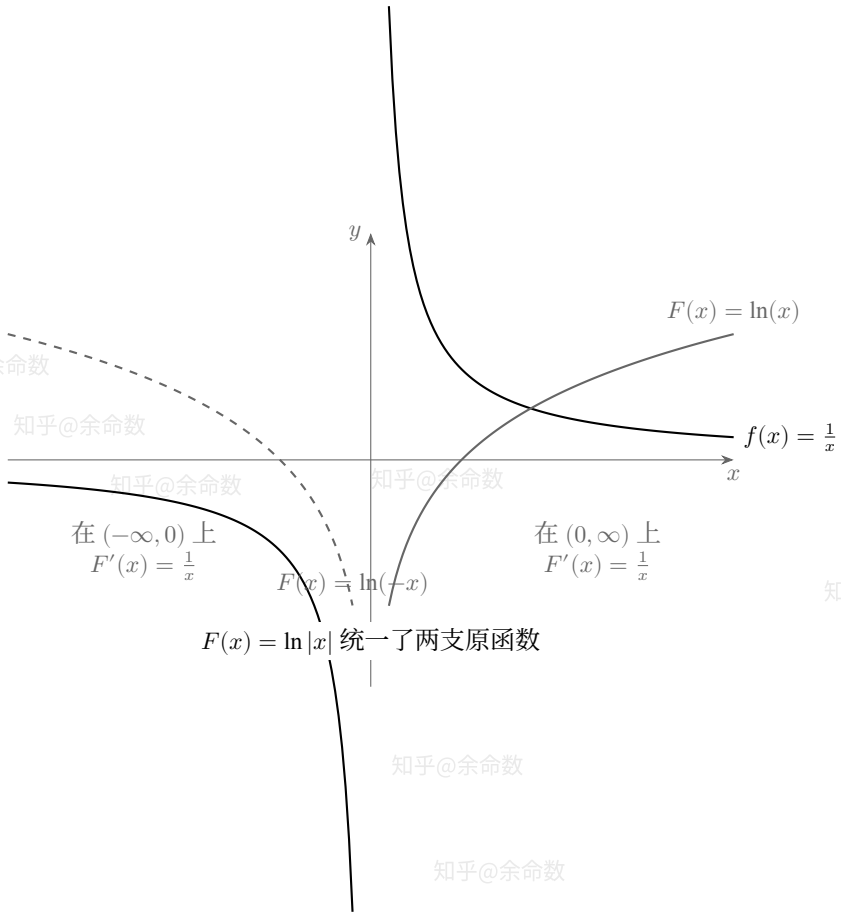


图 17.2 函数 $\ln|x|$ 作为 $\frac{1}{x}$ 在其整个定义域上的原函数

17.2.2.2 指数与三角函数

指数函数 e^x 在微积分中因其导数的不变性而地位特殊, 这一性质原样延续到积分中. 对于一般底数的指数函数 a^x, 求导时产生的 $\ln a$ 因子, 在积分时需要被抵消. 三角函数的积分公式同样是其导数关系的直接逆转, 需特别注意符号的变化.

$$\int e^x\,dx = e^x + C$$

$$\int a^x\,dx = \frac{a^x}{\ln a} + C \quad (a>0, a\neq 1)$$

$$\int \cos x\,dx = \sin x + C$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\int \sin x\,dx = -\cos x + C$$

$$\int \sec^2 x\,dx = \tan x + C$$

$$\int \csc^2 x\,dx = -\cot x + C$$

知乎@余命数

注 17.2 (基本不定积分表)

1. $\int 0\,dx = C$
2. $\int x^n\,dx = \dfrac{x^{n+1}}{n+1} + C \quad (n \neq -1)$
3. $\int \dfrac{1}{x}\,dx = \ln|x| + C$
4. $\int e^x\,dx = e^x + C$
5. $\int a^x\,dx = \dfrac{a^x}{\ln a} + C$
6. $\int \cos x\,dx = \sin x + C$
7. $\int \sin x\,dx = -\cos x + C$
8. $\int \sec^2 x\,dx = \tan x + C$
9. $\int \csc^2 x\,dx = -\cot x + C$
10. $\int \tan x\,dx = -\ln|\cos x| + C = \ln|\sec x| + C$
11. $\int \cot x\,dx = \ln|\sin x| + C$
12. $\int \sec x\tan x\,dx = \sec x + C$
13. $\int \csc x\cot x\,dx = -\csc x + C$
14. $\int \dfrac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + C$
15. $\int \dfrac{1}{1+x^2}\,dx = \arctan x + C$
16. $\int \sinh x\,dx = \cosh x + C$
17. $\int \cosh x\,dx = \sinh x + C$

知乎@余命数　知乎@余命数

上述四组新公式分别由以下求导关系逆转得到: $(\tan x)' = \sec^2 x$ 与 $(\cot x)' = -\csc^2 x$ 配套出现; $(\sec x)' = \sec x\tan x$ 与 $(\csc x)' = -\csc x\cot x$ 配套出现; $(\ln|\cos x|)' = -\tan x$ 给出 $\int \tan x$, $(\ln|\sin x|)' = \cot x$ 给出 $\int \cot x$; 最后两行反三角公式由 $(\arcsin x)' = 1/\sqrt{1-x^2}$ 与 $(\arctan x)' = 1/(1+x^2)$ 直接逆转. 双曲公式由 $(\sinh x)' = \cosh x$ 与 $(\cosh x)' = \sinh x$ 逆转.

知乎@余命数

17.2.3 积分技巧

知乎@余命数　知乎@余命数　余命数

在实际问题中, 很少有被积函数能直接与基本积分表中的某一项完全对应. 求解积分的真正挑战, 在于通过代数或三角恒等变换, 将被积函数**化归**为基本形式的线性组合.

例题 17.1　求不定积分 $\int (4x^3 - \dfrac{3}{\sqrt{x}} + 2e^x)\,dx$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 将根式与负指数统一改写为幂函数, 被积函数即 $4x^3-3x^{-1/2}+2e^x$, 于是可由线性性质逐项套用基本积分表.

$$\begin{aligned}\int\left(4x^3-3x^{-1/2}+2e^x\right)dx &= 4\int x^3\,dx-3\int x^{-1/2}\,dx+2\int e^x\,dx\\ &= 4\cdot\frac{x^4}{4}-3\cdot\frac{x^{1/2}}{1/2}+2e^x+C \quad \text{（幂函数与指数函数积分公式）}\\ &= x^4-6\sqrt{x}+2e^x+C\end{aligned}$$

各项原函数中的常数已合并入总的积分常数 C.

例题 17.2 求不定积分 $\displaystyle\int\frac{(x-1)^2}{x\sqrt{x}}\,dx$.

解 分母 $x\sqrt{x}=x^{3/2}$, 将分子展开后按幂函数整理, 被积函数即化为可逐项积分的形式.

$$\begin{aligned}\frac{(x-1)^2}{x\sqrt{x}} &= \frac{x^2-2x+1}{x^{3/2}}\\ &= \frac{x^2}{x^{3/2}}-\frac{2x}{x^{3/2}}+\frac{1}{x^{3/2}}\\ &= x^{1/2}-2x^{-1/2}+x^{-3/2}\end{aligned}$$

展开后各项均为幂函数, 逐项积分得

$$\begin{aligned}\int(x^{1/2}-2x^{-1/2}+x^{-3/2})\,dx &= \frac{x^{3/2}}{3/2}-2\frac{x^{1/2}}{1/2}+\frac{x^{-1/2}}{-1/2}+C\\ &= \frac{2}{3}x^{3/2}-4x^{1/2}-2x^{-1/2}+C\\ &= \frac{2}{3}x\sqrt{x}-4\sqrt{x}-\frac{2}{\sqrt{x}}+C\end{aligned}$$

例题 17.3 求不定积分 $\displaystyle\int\tan^2 x\,dx$.

解 $\tan^2 x$ 不在基本积分表中, 但勾股恒等式 $1+\tan^2 x=\sec^2 x$ 将其与表中已有的 $\sec^2 x$ 联系起来, 即 $\tan^2 x=\sec^2 x-1$. 代入后两项均可直接积分,

$$\begin{aligned}\int\tan^2 x\,dx &= \int(\sec^2 x-1)\,dx\\ &= \int\sec^2 x\,dx-\int 1\,dx\\ &= \tan x-x+C \quad \left(\int\sec^2 x=\tan x \text{ 与 } \int 1=x\right)\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

17.2.3.1 直接积分法

所谓**直接积分法**, 是指仅依靠代数或三角恒等变形, 配合线性性质与基本积分表, 直接求出积分的方法. 这是最基础的积分技巧, 但在被积函数形式复杂时, 难点在于看穿被积函数的代数结构.

例题 17.4 (多项式除法) 求不定积分 $\displaystyle\int \frac{x^4}{x^2+1}\,dx$.

解 被积函数分子次数高于分母, 属假分式. 多项式除法给出 $x^4=(x^2+1)(x^2-1)+1$, 故

$$\frac{x^4}{x^2+1}=x^2-1+\frac{1}{x^2+1},$$

三项分别对应基本积分公式. 由线性性质,

$$\begin{aligned}\int \frac{x^4}{x^2+1}\,dx&=\int x^2\,dx-\int 1\,dx+\int \frac{1}{x^2+1}\,dx\\&=\frac{x^3}{3}-x+\arctan x+C\quad\left(\int \frac{dx}{x^2+1}=\arctan x\right).\end{aligned}$$

当被积函数是分子次数不低于分母次数的有理函数时, 先做多项式除法, 再分别处理多项式部分与余下的真分式.

例题 17.5 (分子加项减项) 求不定积分 $\displaystyle\int \frac{x}{x+1}\,dx$.

解 被积函数 $\frac{x}{x+1}$ 形式简单, 但分子比分母少一项, 而少的恰是分母本身, 故可用**加项减项**把分子改写为与分母相关的形式.

$$\frac{x}{x+1}=\frac{(x+1)-1}{x+1}=1-\frac{1}{x+1}.$$

原积分由此拆为两项可直接积分的形式,

$$\begin{aligned}\int \frac{x}{x+1}\,dx&=\int 1\,dx-\int \frac{1}{x+1}\,dx\\&=x-\ln|x+1|+C.\end{aligned}$$

将分子表达为分母的导数倍再加减适当项, 是处理许多有理积分的开端.

例题 17.6 (和差化积) 求不定积分 $\displaystyle\int \sin 3x\cos 2x\,dx$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 被积函数是两个不同频率三角函数的乘积, 不直接匹配基本积分表. 由**积化和差**公式 $\sin A\cos B=\frac{1}{2}[\sin(A+B)+\sin(A-B)]$ 得 $\sin 3x\cos 2x=\frac{1}{2}[\sin 5x+\sin x]$, 各项均为基本形式, 故

$$\begin{aligned}\int \sin 3x\cos 2x\,dx &= \frac{1}{2}\int \sin 5x\,dx+\frac{1}{2}\int \sin x\,dx\\ &= \frac{1}{2}\left(-\frac{\cos 5x}{5}\right)+\frac{1}{2}(-\cos x)+C \quad \left(\int \sin kx\,dx=-\frac{\cos kx}{k}\right)\\ &= -\frac{\cos 5x}{10}-\frac{\cos x}{2}+C.\end{aligned}$$

知乎@余命数

17.2.4 初等函数积分的局限性

已在 第17.1节 中引入了达布定理及其推论, 给出了何种函数无原函数的负向判据. 然而, 存在性与能否用所熟悉的函数形式表达是两个截然不同的问题, 一个原函数可能存在, 却不属于任何**初等函数**族.

所熟悉的函数, 如多项式、有理函数、指数、对数、三角函数及其反函数, 以及由它们经过有限次四则运算与复合得到的函数, 统称为**初等函数**. 然而, 一个初等函数的原函数未必也是初等函数, 大量看似简单的初等函数, 其原函数无法用初等函数表示. 这反映了积分运算与微分运算的一个根本性不对称, 即**初等函数的导数必为初等函数, 反之不然**.

一个简短的历史注记. 法国数学家刘维尔 (Liouville, 1835) 证明了一个一般性结论, 若一个初等函数 f 的原函数能用初等函数写出, 则该原函数必然具有*有理函数加几个对数项的常数倍*这一特定形式 (用符号写为 $F=R+\sum_i c_i\ln R_i$, 其中 R, R_i 都是初等构件的有理函数) . 这一定理的证明属于微分代数 (differential algebra) 范畴, 此处只介绍结论. 反过来, 只要能确认 f 的原函数不属于上述有理加对数的形式, 即可断言它不是初等函数.

知乎@余命数

刘维尔定理最常用的推论是, 若 f 的原函数不能用上述有理加对数形式表达, 则 $\int f\,dx$ 不是初等函数. 这里举几个著名的原函数不是初等函数的例子, 它们的原函数分别定义了重要的特殊函数:

知乎@余命数

- $\int e^{-x^2}\,dx$: 高斯积分的被积函数, 其原函数与误差函数 $\operatorname{erf}(x)=\frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$ 相关, 在概率论与统计学中频繁出现.
- $\int \frac{\sin x}{x}\,dx$: 正弦积分 $\operatorname{Si}(x)=\int_0^x \frac{\sin t}{t}\,dt$, 在信号处理与傅里叶分析中频繁出现.
- $\int \frac{1}{\ln x}\,dx$: 对数积分 $\operatorname{Li}(x)=\int_0^x \frac{dt}{\ln t}$ (约定绕过奇点), 在素数分布研究中与素数定理紧密相关.
- $\int \sqrt{1-k^2\sin^2 x}\,dx$: 第二类完全/不完全椭圆积分 $\mathrm{E}(k)$, 在最速降线、单摆周期等问题中扮演关键角色.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

基本积分公式加上换元、分部等技巧所能解出的初等积分，仅仅是全部初等函数中的沧海一粟. 这一事实说明数值积分方法与特殊函数理论同样属于积分学的组成部分.

17.3 换元积分法

我们现有的工具，即基本积分公式和线性运算法则，使我们能够对基本函数的和与差进行积分. 然而这套工具在面对复合函数时则显得无能为力. 例如，我们能够求解 $\int \cos x\,dx$，但面对 $\int \cos(2x)\,dx$ 却束手无策；可求解 $\int \sqrt{x}\,dx$，但对于 $\int 2x\sqrt{x^2+1}\,dx$ 却感到棘手.

这些更复杂的被积函数源于链式法则，往往具有一个共同的结构特征，它们看起来像是一个复合函数经过链式法则求导后的产物. 这提示我们，处理复合函数的积分，关键在微分学的链式法则之中.

再看链式法则.

$$\frac{d}{dx}[F(g(x))] = F'(g(x)) \cdot g'(x)$$

如果对这个等式的两边同时进行不定积分，根据微分与积分的互逆关系，我们会得到.

$$\int F'(g(x)) \cdot g'(x)\,dx = F(g(x)) + C$$

这个恒等式是换元技巧的依据. 它说明了一个基本的模式，如果一个被积函数可以被识别为一个外层函数的导数 F' 作用于一个内层函数 $g(x)$，再乘以这个内层函数自身的导数 $g'(x)$ 的形式，那么它的积分结果就是那个外层函数的原函数 F 作用于内层函数 $g(x)$.

为了系统地利用这一模式，我们引入一种形式上的变量替换，即令 $u = g(x)$. 这种变量替换的技巧，就是所谓的**换元积分法**，它在积分运算中的地位，完全等同于链式法则在微分运算中的地位. 它分为两种方向，第一类换元法通过 $u = g(x)$ 将复杂结构收缩为简单形式；第二类换元法通过 $x = \phi(t)$ 将变量展开为一个函数，以期简化被积函数的内在结构.

17.3.1 第一类换元法（凑微分法）

第一类换元法的策略是由内向外识别结构. 需在被积函数中辨认出 $f(g(x))g'(x)$ 的形式，然后通过替换 $u = g(x)$ 将其转化为更易处理的 $\int f(u)du$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定理 17.5 (第一类换元积分法则) 设函数 $f(u)$ 在区间 J 上具有原函数 $F(u)$, 即 $F'(u)=f(u)$. 如果 $u=g(x)$ 是区间 I 上的可导函数, 且 $g(I)\subset J$, 那么有

$$\int f(g(x))g'(x)\,dx=\int f(u)\,du=F(u)+C=F(g(x))+C \tag{17-2}$$

知乎@余命数

证明 依不定积分的定义, 只需验证 $F(g(x))+C$ 是被积函数 $f(g(x))g'(x)$ 的原函数, 即证

$$\frac{d}{dx}[F(g(x))+C]=f(g(x))g'(x).$$

$F(g(x))$ 由外层 $F(u)$ 与内层 $u=g(x)$ 复合而成, 由链式法则

$$\begin{aligned}\frac{d}{dx}F(g(x))&=\frac{dF}{du}\cdot\frac{du}{dx}\\&=F'(u)\cdot g'(x)=f(u)\cdot g'(x)\quad\left(\text{由 }F'(u)=f(u)\text{ 与 }\frac{du}{dx}=g'(x)\right).\end{aligned}$$

回代 $u=g(x)$, 上式即 $f(g(x))g'(x)$; 常数 C 求导为零, 故

$$\frac{d}{dx}[F(g(x))+C]=f(g(x))g'(x),$$

知乎@余命数

这表明 $F(g(x))$ 是 $f(g(x))g'(x)$ 的一个原函数. 再由假设 $\int f(u)\,du=F(u)+C$ 代入 $u=g(x)$, 定理得证. □

在实际操作中, 将莱布尼茨的微分符号 $du=g'(x)dx$ 作为一个形式上的整体来进行替换. 这种将微分 dx 视为可分离、可组合的代数实体的做法, 在严格的分析基础上需要更深的理论支撑(微分形式), 但在计算层面是一种极为有效且直观的启发式工具.

注 17.3 (第一类换元法战略思想)

1. **识别结构**. 审视被积函数, 尝试寻找一个复合函数的内层 $u=g(x)$, 其导数 $g'(x)$(或其常数倍)也在被积表达式中作为因子出现.
2. **换元**.
   - 声明替换 $u=g(x)$, 并计算其微分 $du=g'(x)dx$.
   - 将被积表达式中所有与 x 相关的部分完全替换为关于 u 和 du 的表达式. 积分 $\int f(g(x))g'(x)\,dx$ 此时应转化为更简洁的 $\int f(u)\,du$. 这个过程也常被称为凑微分.
3. **求解新积分**. 对关于 u 的新积分求解. 这一步通常可以直接套用基本积分公式.
4. **回代**. 将积分结果中的变量 u 替换回原来的表达式 $g(x)$, 确保最终答案是关于 x 的函数.

知乎@余命数

例题 17.7　求不定积分 $\int 2x\sqrt{x^2+1}\,dx$.

解　被积函数含有复合结构 $\sqrt{x^2+1}$, 取内层 $u=x^2+1$, 其微分 $du=2x\,dx$ 恰好是根号外剩余的因子, 故为一可凑微分的换元结构. 令 $u=x^2+1$, 则 $du=2x\,dx$, 原积分化为关于 u 的幂函数积分,

$$\int \sqrt{x^2+1}\,(2x\,dx)=\int u^{1/2}\,du=\frac{2}{3}u^{3/2}+C.$$

回代 $u=x^2+1$, 得

$$\frac{2}{3}(x^2+1)^{3/2}+C.$$

例题 17.8　求不定积分 $\int \tan x\,dx$.

解　$\tan x$ 不在基本积分公式表中, 但改写为分式 $\tan x=\frac{\sin x}{\cos x}$ 后, 其分母 $\cos x$ 可取作内层 $u=\cos x$, 微分 $du=(\cos x)'\,dx=-\sin x\,dx$, 与分子 $\sin x\,dx$ 仅差常数因子 -1. 凑微分得 $\sin x\,dx=-du$, 于是

$$\int \tan x\,dx=\int \frac{\sin x\,dx}{\cos x}=\int \frac{-du}{u}=-\int \frac{du}{u}=-\ln|u|+C.$$

回代 $u=\cos x$,得 $-\ln|\cos x|+C$;由对数性质亦可写成 $\ln|\cos x|^{-1}+C=\ln|\sec x|+C$.

17.3.2 第二类换元法

第一类换元法的策略是令 $u=g(x)$, 将被积函数向标准形式收缩. 而第二类换元法则反其道而行之, 它通过令 $x=\phi(t)$, 将被积函数进行扩张和变形. 这种方法的目标是, 选择一个巧妙的函数 $\phi(t)$, 使得原先棘手的被积函数(尤其是含有根式的), 在新的变量 t 下, 能够利用代数或三角恒等式得到根本性的简化, 从而达到有理化或消除复杂结构的目的.

这种方法最典型的应用场景是处理含有二次根式的积分, 特别是形如 $\sqrt{a^2-x^2}$, $\sqrt{a^2+x^2}$ 和 $\sqrt{x^2-a^2}$ 的被积函数. 这些根式结构与勾股定理以及三角函数的基本恒等式在结构上一一对应.

- $\sqrt{a^2-x^2}\longleftrightarrow$ 直角边 $\sqrt{c^2-b^2}\longleftrightarrow \cos t=\sqrt{1-\sin^2 t}$
- $\sqrt{a^2+x^2}\longleftrightarrow$ 斜边 $\sqrt{b^2+c^2}\longleftrightarrow \sec t=\sqrt{1+\tan^2 t}$
- $\sqrt{x^2-a^2}\longleftrightarrow$ 直角边 $\sqrt{c^2-b^2}\longleftrightarrow \tan t=\sqrt{\sec^2 t-1}$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

利用这些联想, 引入三角函数充当新变量即可彻底消除根号, 把复杂的无理函数积分转化为相对简单的有理三角函数积分. 这就是三**角换元法**的做法.

在进入操作细节之前, 我们先把第二类换元法的理论依据陈述清楚, 它的证明正是链式法则的一次直接应用.

定理 17.6 (第二类换元积分法则) 设 $f(x)$ 在区间 I 上有定义, $x=\psi(t)$ 在区间 J 上单调可导, $\psi'(t)\neq 0$ 在 J 上恒成立, 且 $\psi(J)\subset I$. 若 $f(\psi(t))\psi'(t)$ 在 J 上有原函数 $\Phi(t)$, 则 $f(x)$ 在 $\psi(J)$ 上有不定积分, 且

$$\int f(x)\,dx=\Phi(\psi^{-1}(x))+C.$$

证明 由 ψ 在 J 上单调可导且 $\psi'(t)\neq 0$, 反函数 $t=\psi^{-1}(x)$ 在 $\psi(J)$ 上可导, 且

$$\frac{d}{dx}\psi^{-1}(x)=\left.\frac{1}{\psi'(t)}\right|_{t=\psi^{-1}(x)}.$$

知乎@余命数

考虑复合函数 $F(x)=\Phi(\psi^{-1}(x))$, 由链式法则对其关于 x 求导,

$$F'(x)=\Phi'(t)\cdot\frac{d}{dx}\psi^{-1}(x)=\Phi'(\psi^{-1}(x))\cdot\frac{1}{\psi'(\psi^{-1}(x))}.$$

将假设 $\Phi'(t)=f(\psi(t))\psi'(t)$ 代入, 即 $\Phi'(\psi^{-1}(x))=f(x)\,\psi'(\psi^{-1}(x))$, $\psi'(\psi^{-1}(x))$ 与其倒数相乘抵消, 得

$$F'(x)=f(x).$$

因此 $F(x)=\Phi(\psi^{-1}(x))$ 是 $f(x)$ 在 $\psi(J)$ 上的一个原函数. 由不定积分定义,

$$\int f(x)\,dx=\Phi(\psi^{-1}(x))+C.$$

□

知乎@余命数

17.3.2.1 三角换元的操作策略

根据被积函数中根式的不同形式, 我们采用三种标准化的替换策略.

注 17.4 (三角换元法则)

1. **若含 $\sqrt{a^2-x^2}$ ($a>0$):** 令 $x=a\sin t$, 其中 $t\in[-\frac{\pi}{2},\frac{\pi}{2}]$. 此时 $dx=a\cos t\,dt$. 根式化为 $\sqrt{a^2-a^2\sin^2 t}=a\sqrt{\cos^2 t}=a|\cos t|$. 由于 t 的取值范围保证了 $\cos t\geq 0$, 故 $\sqrt{a^2-x^2}=a\cos t$.

余命数

知乎@余命数

2. **若含 $\sqrt{a^2+x^2}$ $(a>0)$：**令 $x=a\tan t$，其中 $t\in(-\frac{\pi}{2},\frac{\pi}{2})$. 此时 $dx=a\sec^2 t\,dt$. 根式化为 $\sqrt{a^2+a^2\tan^2 t}=a\sqrt{\sec^2 t}=a|\sec t|$. 由于 t 的取值范围保证了 $\sec t>0$，故 $\sqrt{a^2+x^2}=a\sec t$.
3. **若含 $\sqrt{x^2-a^2}$ $(a>0)$：**令 $x=a\sec t$，其中 $t\in[0,\frac{\pi}{2})\cup(\frac{\pi}{2},\pi]$. 此时 $dx=a\sec t\tan t\,dt$. 根式化为 $\sqrt{a^2\sec^2 t-a^2}=a\sqrt{\tan^2 t}=a|\tan t|$. 在所选的 t 定义域内，通过选取合适的象限可保证 $\tan t$ 的符号，从而去掉绝对值.

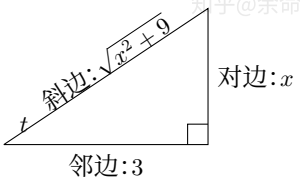
三条法则的要点相同，选取适当的变量代换，利用三角恒等式 $1-\sin^2 t=\cos^2 t$，$1+\tan^2 t=\sec^2 t$ 或 $\sec^2 t-1=\tan^2 t$ 完成开方运算，使被积函数有理化.

例题 17.9　求不定积分 $\displaystyle\int\frac{1}{(x^2+9)^{3/2}}\,dx$.

解　被积函数含有 $(x^2+9)^{3/2}=(\sqrt{x^2+9})^3$，取 $a=3$ 即呈 $\sqrt{x^2+a^2}$ 的形式，适用第二类三角换元. 令 $x=3\tan t$，$t\in(-\frac{\pi}{2},\frac{\pi}{2})$，则 $dx=3\sec^2 t\,dt$，$\sqrt{x^2+9}=\sqrt{9\tan^2 t+9}=3\sec t$，故 $(x^2+9)^{3/2}=27\sec^3 t$. 代入原积分，

$$\begin{aligned}\int\frac{1}{27\sec^3 t}\cdot(3\sec^2 t\,dt)&=\frac{3}{27}\int\frac{\sec^2 t}{\sec^3 t}\,dt\\&=\frac{1}{9}\int\frac{1}{\sec t}\,dt=\frac{1}{9}\int\cos t\,dt=\frac{1}{9}\sin t+C.\end{aligned}$$

最后需将结果从 t 域转换回 x 域. 由换元关系 $x=3\tan t$，即 $\tan t=\frac{x}{3}$，构造辅助直角三角形几何化此关系. 由勾股定理，斜边长为 $\sqrt{x^2+3^2}=\sqrt{x^2+9}$，由此读出



$\sin t=\frac{x}{\sqrt{x^2+9}}$. 代回积分结果得

$$\frac{1}{9}\frac{x}{\sqrt{x^2+9}}+C.$$

17.3.2.2 根式整体代换与降幂倒代换

三角换元是处理一类二次根式的常见代换手段，适用于 $\sqrt{a^2\pm x^2}$ 与 $\sqrt{x^2-a^2}$ 等根式形式. 但当根式结构与 x 的幂次交错出现，例如根号外含 x 的高次因子时，直接令 $x=1/t$ 作**降幂倒代换**，往往能把根号下 x 的高次项一次性化为低次. 此技巧常与根式**整体代换**配合，目标同为消去根号.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

例题 17.10 求不定积分 $\displaystyle\int \frac{dx}{x\sqrt{a^2-x^2}}(a>0, 0<x<a, \text{故 } x>0)$.

解 令 $x=1/t$(降幂倒代换), 则 $dx=-dt/t^2$. 由 $x>0$ 知 $t>0$, 故

$$\sqrt{a^2-x^2}=\sqrt{a^2-\frac{1}{t^2}}=\frac{\sqrt{a^2t^2-1}}{t}.$$

被积表达式变为

$$\frac{1}{(1/t)\cdot\sqrt{a^2t^2-1}/t}\cdot\left(-\frac{dt}{t^2}\right)=-\frac{dt}{\sqrt{a^2t^2-1}}.$$

令 $u=at$, 则 $du=a\,dt$,

$$-\int\frac{dt}{\sqrt{a^2t^2-1}}=-\frac{1}{a}\int\frac{du}{\sqrt{u^2-1}}=-\frac{1}{a}\operatorname{arcosh}|u|+C,$$

其中用到反双曲余弦公式 $\displaystyle\int\frac{du}{\sqrt{u^2-1}}=\operatorname{arcosh}|u|+C=\ln\left|u+\sqrt{u^2-1}\right|+C$, 该公式在 $|u|>1$ 时成立. 由 $0<x<a$ 得 $u=a/x>1$, 定义域自动满足. 回代 $u=at=a/x$,

$$\int\frac{dx}{x\sqrt{a^2-x^2}}=-\frac{1}{a}\operatorname{arcosh}\left(\frac{a}{x}\right)+C=-\frac{1}{a}\ln\left(\frac{a+\sqrt{a^2-x^2}}{x}\right)+C.$$

17.3.2.3 双曲代换

双曲函数由恒等式 $\cosh^2 t-\sinh^2 t=1$ 紧密联结, 这一等式与三角恒等式 $\cos^2 t+\sin^2 t=1$ 在结构上对偶. 差一个符号正是它能消去 $\sqrt{x^2\pm a^2}$ 类根号的关键: 处理 $\sqrt{a^2+x^2}$ 时令 $x=a\sinh t$, 此时 $dx=a\cosh t\,dt$, $\sqrt{a^2+x^2}=a\cosh t$; 处理 $\sqrt{x^2-a^2}$ 时令 $x=a\cosh t\,(t>0)$, 此时 $dx=a\sinh t\,dt$, $\sqrt{x^2-a^2}=a\sinh t$. 两种代换都能将无理根式直接化为双曲函数, 之后的积分只需结合 $\cosh^2 t=(\cosh 2t+1)/2$ 等恒等式即可完成.

例题 17.11 求不定积分 $\displaystyle\int\sqrt{x^2+a^2}\,dx\,(a>0)$.

解 令 $x=a\sinh t$, 则 $dx=a\cosh t\,dt$, $\sqrt{x^2+a^2}=\sqrt{a^2\sinh^2 t+a^2}=a\cosh t$. 原积分化为

$$\int(a\cosh t)(a\cosh t)\,dt=a^2\int\cosh^2 t\,dt.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

利用恒等式 $\cosh^2 t=\dfrac{\cosh 2t+1}{2}$，

$$a^2\int \cosh^2 t\,dt=a^2\int \frac{\cosh 2t+1}{2}\,dt=a^2\left(\frac{t}{2}+\frac{\sinh 2t}{4}\right)+C.$$

将 $\sinh 2t=2\sinh t\cosh t$ 代入，并回代 $t=\operatorname{arsinh}(x/a)$、$\sinh t=x/a$、$\cosh t=\sqrt{x^2+a^2}/a$，

$$a^2\left(\frac{t}{2}+\frac{\sinh t\,\cosh t}{2}\right)+C=\frac{a^2}{2}\operatorname{arsinh}\left(\frac{x}{a}\right)+\frac{x}{2}\sqrt{x^2+a^2}+C.$$

再以 $\operatorname{arsinh}(x/a)=\ln\left(x/a+\sqrt{x^2/a^2+1}\right)=\ln\left(x+\sqrt{x^2+a^2}\right)-\ln a$ 整理，其中 $-\dfrac{a^2}{2}\ln a$ 并入常数 C，

$$\int\sqrt{x^2+a^2}\,dx=\frac{x}{2}\sqrt{x^2+a^2}+\frac{a^2}{2}\ln\left(x+\sqrt{x^2+a^2}\right)+C.$$

17.3.2.4 万能代换与欧拉代换

17.3.2.5 万能代换（Weierstrass 代换）

对于仅由 $\sin x,\cos x$ 经四则运算复合而成的**有理三角积分** $\int R(\sin x,\cos x)\,dx$，万能代换 $t=\tan(x/2)$ 将其统一化为有理代数积分. 关键三角恒等式为

$$\sin x=\frac{2t}{1+t^2},\quad \cos x=\frac{1-t^2}{1+t^2},\quad dx=\frac{2\,dt}{1+t^2}.$$

这一代换的几何实质是将单位圆上的点 $(\cos x,\sin x)$ 以斜率参数 t 重新参数化，据此任何 $\int R(\sin x,\cos x)\,dx$ 都可转化为关于 t 的有理函数积分 $\int \widetilde{R}(t)\,dt$.

例题 17.12　求不定积分 $\displaystyle\int\frac{dx}{2+\sin x}$.

解　令 $t=\tan(x/2)$，则 $\sin x=\dfrac{2t}{1+t^2}$，$dx=\dfrac{2\,dt}{1+t^2}$. 代入原积分，

$$\int\frac{1}{2+\dfrac{2t}{1+t^2}}\cdot\frac{2\,dt}{1+t^2}=\int\frac{2\,dt}{2(1+t^2)+2t}=\int\frac{dt}{t^2+t+1}.$$

将分母完成平方，$t^2+t+1=\left(t+\dfrac{1}{2}\right)^2+\dfrac{3}{4}$，故

$$\int\frac{dt}{\left(t+\frac{1}{2}\right)^2+\frac{3}{4}}=\frac{2}{\sqrt{3}}\arctan\left(\frac{2\left(t+\frac{1}{2}\right)}{\sqrt{3}}\right)+C=\frac{2}{\sqrt{3}}\arctan\left(\frac{2t+1}{\sqrt{3}}\right)+C.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

回代 $t=\tan(x/2)$,

$$\int \frac{dx}{2+\sin x}=\frac{2}{\sqrt{3}}\arctan\left(\frac{2\tan(x/2)+1}{\sqrt{3}}\right)+C.$$

知乎@余命数

17.3.2.6 欧拉代换

对于含 $\sqrt{ax^2+bx+c}$ 的积分, 欧拉 (Euler) 给出三种代换以彻底消去根号, 将其化为有理函数积分. 根号下二次式的判别式情况不同, 选用不同的代换:

- 第一种 ($a>0$): 令 $\sqrt{ax^2+bx+c}=t\pm\sqrt{a}\,x$. 两边平方后 x^2 项系数正负相消, x 可被 t 的有理式表示. 适用于 ax^2+bx+c 无实根的情形.
- 第二种 ($c>0$): 令 $\sqrt{ax^2+bx+c}=xt\pm\sqrt{c}$. 两边平方后常数项正负相消, x 可被 t 的有理式表示. 适用于 $c>0$(常与无实根情形配合).
- 第三种 (ax^2+bx+c 有实根 α,β, 即 $ax^2+bx+c=a(x-\alpha)(x-\beta)$): 令 $\sqrt{ax^2+bx+c}=t(x-\alpha)$. 两边平方后 $(x-\alpha)$ 因子直接约去, 化 x 为 t 的有理式. 适用于判别式 $b^2-4ac>0$ 时必用.

例题 17.13 求不定积分 $\displaystyle\int\frac{dx}{x\sqrt{x^2+x+1}}$ (设 $x>0$, 此时 $\sqrt{x^2+x+1}$ 有意义).

解 此处 $c=1>0$, 适用第二种欧拉代换. 令 $\sqrt{x^2+x+1}=xt+1$(取正号, 配合 $x>0$). 平方两边,

$$x^2+x+1=x^2t^2+2xt+1\Longrightarrow x^2(1-t^2)+x(1-2t)=0.$$

由 $x>0$ 约去 x, 得 $x(1-t^2)=2t-1$, 即

$$x=\frac{2t-1}{1-t^2}.$$

对 x 关于 t 求导,

$$dx=\frac{d}{dt}\left(\frac{2t-1}{1-t^2}\right)dt=\frac{2(t^2-t+1)}{(1-t^2)^2}\,dt.$$

同时由 $\sqrt{x^2+x+1}=xt+1$ 得

$$\sqrt{x^2+x+1}=\frac{(2t-1)t}{1-t^2}+1=\frac{t^2-t+1}{1-t^2},$$

因此

$$x\sqrt{x^2+x+1}=\frac{(2t-1)(t^2-t+1)}{(1-t^2)^2}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

将被积表达式在 t 变量下整理,

$$\frac{dx}{x\sqrt{x^2+x+1}}=\frac{2(t^2-t+1)\,dt/(1-t^2)^2}{(2t-1)(t^2-t+1)/(1-t^2)^2}=\frac{2\,dt}{2t-1}.$$

故

$$\int\frac{dx}{x\sqrt{x^2+x+1}}=\int\frac{2\,dt}{2t-1}=\ln|2t-1|+C.$$

由 $\sqrt{x^2+x+1}=xt+1$ 解出 $t=\dfrac{\sqrt{x^2+x+1}-1}{x}$, 回代,

$$\int\frac{dx}{x\sqrt{x^2+x+1}}=\ln\left|\frac{2\sqrt{x^2+x+1}-x-2}{x}\right|+C.$$

17.4 分部积分法

已掌握了不定积分的线性性质, 它允许我们处理函数和与差的积分. 然而, 在面对两个函数乘积的积分时, 线性法则便无能为力. 微分学中的乘法法则 $(uv)'=u'v+uv'$ 在积分学中存在一个对应的结构, 它是一种基本的转化策略, 这就是**分部积分法**.

其思想的根源,正在于对乘法法则的逆向应用. 我们从两个可导函数 $u=u(x)$ 与 $v=v(x)$ 的乘积的导数法则出发.

$$\frac{d}{dx}[u(x)v(x)]=u'(x)v(x)+u(x)v'(x)$$

这是一个关于函数及其导数的恒等式. 根据微积分基本定理的推论, 等式两边函数的原函数族必然只相差一个常数. 于是对该恒等式两边同时取不定积分.

$$\int\frac{d}{dx}[u(x)v(x)]\,dx=\int u'(x)v(x)\,dx+\int u(x)v'(x)\,dx$$

等式左边, 积分与微分互为逆运算, 因此结果就是原函数本身.

$$u(x)v(x)=\int u'(x)v(x)\,dx+\int u(x)v'(x)\,dx$$

通过简单的移项, 便能用一个积分来表达另一个积分, 说明一种转化关系.

$$\int u(x)v'(x)\,dx=u(x)v(x)-\int v(x)u'(x)\,dx$$

为计算上的便利与形式上的统一, 我们采用莱布尼茨的微分记号, 令 $du=u'(x)dx$ 以及 $dv=v'(x)dx$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 17.7 (分部积分法) 若函数 $u(x)$ 与 $v(x)$ 均可导，且不定积分 $\int v(x)u'(x)\,dx$ 存在，则 $\int u(x)v'(x)\,dx$ 也存在，并且

$$\int u\,dv = uv - \int v\,du \tag{17-3}$$

证明 由乘积求导法则，

$$(uv)' = u'v + uv'.$$

移项得 $uv' = (uv)' - u'v$. 对两边取不定积分，并把积分常数并入最终的常数项，得

$$\int u(x)v'(x)\,dx = uv - \int u'(x)v(x)\,dx.$$

用微分记号 $du = u'(x)dx$, $dv = v'(x)dx$ 书写，即

$$\int u\,dv = uv - \int v\,du.$$

知乎@余命数

□

分部积分法是一种**积分的变换**. 它将计算 $\int u\,dv$ 的问题，转化为计算 $\int v\,du$ 的问题. 这种变换的价值在于，它把作用在 v 上的微分算子转移到了 u 上. 如果这个转移过程能使新的被积函数 $v\,du$ 比原来的 $u\,dv$ 更易于积分，那么策略就成功了.

注 17.5 (分部积分法的取舍原则)
将被积表达式 $f(x)dx$ 分解为 u 和 dv 两部分时，需要进行权衡. 一个成功的分解遵循以下原则.

1. **简化原则**. 所选取的 u 在微分后得到的 du 应当是形式上更简单的函数. 多项式函数是这一原则的典型受益者，其次数每经过一次微分便降低一阶.
2. **可积原则**. 所选取的 dv 必须是能够被直接积分以求得 v 的. 若 v 本身无法求出，则该方法无法继续.

例题 17.14 求不定积分 $\int xe^x\,dx$.

解 被积函数是多项式 x 与指数函数 e^x 的乘积. 为利用分部积分法简化积分，取多项式 $u = x$(经微分降幂)，$dv = e^x\,dx$，则 $du = dx$ 且 $v = \int e^x\,dx = e^x$. 由分部积分公式 $\int u\,dv = uv - \int v\,du$，

$$\int xe^x\,dx = xe^x - \int e^x\,dx$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

$$= xe^x - e^x + C.$$

原积分由此化为可直接求解的基本积分.

例题 17.15　求不定积分 $\int \ln x\,dx$.

解　被积函数虽只有一个因子，但可视为 $\ln x \cdot 1$ 的乘积，从而套用分部积分. 取 $u = \ln x$, 其导数 $\frac{1}{x}$ 为代数函数, 形式上更简单; 取 $dv = 1 \cdot dx$, 则 $du = \frac{1}{x}\,dx$, $v = x$. 由分部积分公式，

$$\begin{aligned}\int \ln x\,dx &= (\ln x)\cdot x - \int x \cdot \frac{1}{x}\,dx \\ &= x\ln x - \int 1\,dx \\ &= x\ln x - x + C.\end{aligned}$$

反向取 $u = 1, dv = \ln x\,dx$ 将使 $v = \int \ln x\,dx$ 退回原问题, 故此处须令对数部分作 u.

例题 17.16　求不定积分 $\int e^x \sin x\,dx$.

解　指数函数与三角函数的导数和积分都保持其函数类型不变，一次分部积分不会直接简化被积函数，但连续两次应用会产生一个含原积分本身的代数方程，由此可解出 I. 记 $I = \int e^x \sin x\,dx$. 取 $u = e^x, dv = \sin x\,dx$, 则 $du = e^x\,dx, v = -\cos x$, 第一次分部积分给出

$$\begin{aligned}I &= e^x(-\cos x) - \int (-\cos x)e^x\,dx \\ &= -e^x\cos x + \int e^x \cos x\,dx.\end{aligned}$$

对新出现的 $\int e^x \cos x\,dx$ 再次分部，保持选择一致（仍取 e^x 作 u），即 $u = e^x$、$dv = \cos x\,dx$、$du = e^x\,dx$、$v = \sin x$,

$$\begin{aligned}\int e^x \cos x\,dx &= e^x \sin x - \int (\sin x)e^x\,dx \\ &= e^x \sin x - I.\end{aligned}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

把后一式代回前一式，得关于 I 的方程 $I = -e^x\cos x + (e^x\sin x - I)$，即 $2I = e^x(\sin x - \cos x)$. 解出并补上积分常数，

知乎@余命数

$$I = \frac{e^x}{2}(\sin x - \cos x) + C.$$

注 17.6 (反对幂三指)
分部积分中选 u 时，通行的优先次序是反对幂三指，即按以下顺序优先取 u：

- **反**：反三角函数，如 $\arcsin x$, $\arctan x$.
- **对**：对数函数，如 $\ln x$, $\log_a x$.
- **幂**：幂函数与代数函数，含多项式，如 x^n, $\sqrt{x}$.
- **三**：三角函数，如 $\sin x$, $\cos x$.
- **指**：指数函数，如 e^x, a^x.

这样排序的理由是，排在前面的函数微分后结构会简化（反三角微分变为代数型、对数变有理式、幂函数降次），排在后面的函数微分或积分后类型不变，故前者取 u、后者取 dv.

17.4.1 广义分部积分法与循环技巧

知乎@余命数

当被积函数 $u(x)v(x)$ 中的 u 在反复微分后降阶迅速，而 v 在反复积分后保持函数类不变，如 e^x, $\sin x$, $\cos x$ 时，一次分部积分不足以解决问题，但反复分部会得到一长串的加减相间项. 此时有一种系统化的列表法，称为 **DI 表**（微分-积分表），可以条理清楚地组织这种反复操作.

注 17.7 (DI 表方法)
将 u 与 dv 写在表头两列，微分列 D 由 u 开始向下逐次求导，积分列 I 由 dv 开始向下逐次积分，然后沿斜对角配对相乘并交替符号：

微分列 (D)	u	$-u'$	u''	$-u'''$
积分列 (I)	$v_1 = \int dv$	$v_2 = \int v_1$	$v_3 = \int v_2$	$v_4 = \int v_3$

沿斜对角线将 D 行某一项与下一行 I 列对应项相乘，符号取 D 行的符号：

$$\int u\,dv = uv_1 - u'v_2 + u''v_3 - u'''v_4 + \dots$$

知乎@余命数

当 D 列出现 0 时，即 $u^{(n)} = 0$，例如 u 为多项式，求和自然终止. 这是多项式-指数/三角类积分的高效工具.

知乎@余命数

定理 17.8 (广义分部积分公式) 设 u 在区间 I 上 n 阶可导, $v_1, v_2, \dots, v_n$ 满足 $v_{k+1}' = v_k$, 即可逐次求其原函数. 记 $u^{(0)} = u$. 则

$$\int u\,dv = \sum_{k=0}^{n-1} (-1)^k\, u^{(k)} v_{k+1} + (-1)^n \int u^{(n)} v_n\,dt.$$

证明 对 n 归纳. 当 $n = 1$ 时即普通分部积分公式

$$\int u\,dv = uv_1 - \int u' v_1\,dt.$$

设 $n = m$ 时结论已成立. 当 $n = m+1$ 时, 对余项 $(-1)^m \int u^{(m)} v_m\,dt$ 再次应用普通分部积分公式(取 $u^{(m)}$ 作新 u, $v_m\,dt$ 作新 dv). 余项化为

$$(-1)^m\big[u^{(m)} v_{m+1} - \int u^{(m+1)} v_{m+1}\,dt\big] = (-1)^m u^{(m)} v_{m+1} + (-1)^{m+1} \int u^{(m+1)} v_{m+1}\,dt.$$

合入主和即得 $n = m+1$ 公式. □

例题 17.17 求不定积分 $\displaystyle\int x^3 e^x\,dx$.

解 这是一个典型适用 DI 表的问题. 取 $u = x^3$, 经 4 次微分归零; 取 $dv = e^x\,dx$, 经反复积分仍为 e^x. 列表:

D	x^3	$3x^2$	$6x$	6	0
signs	$+$	$-$	$+$	$-$	$+$
I	e^x	e^x	e^x	e^x	e^x

按斜对角线读出(符号取 D 列的符号, 乘下一行 I 列同位项):

$$\int x^3 e^x\,dx = x^3 e^x - 3x^2 e^x + 6x\,e^x - 6e^x + C = (x^3 - 3x^2 + 6x - 6)\,e^x + C.$$

验证: 对右侧求导得 $(x^3 - 3x^2 + 6x - 6)' e^x + (x^3 - 3x^2 + 6x - 6)e^x = (3x^2 - 6x + 6 + x^3 - 3x^2 + 6x - 6)e^x = x^3 e^x$, 与被积函数一致.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

17.4.2 积分递推公式

许多含参量 n 的积分 I_n 可以通过分部积分建立与低阶 I_{n-2}(或 I_{n-1})之间的关系, 最终问题降阶为某起点 I_0, I_1 的直接计算. 这就是所谓**积分递推公式**. 现给出四类经典递推.

定理 17.9 ($\sin^n x$ 与 $\cos^n x$ 的降阶递推) 设 $n \geq 2$. 记 $I_n = \int \sin^n x\,dx$ (或 $I_n = \int \cos^n x\,dx$), 则

$$I_n = -\frac{1}{n}\sin^{n-1} x\cos x + \frac{n-1}{n} I_{n-2} \quad (\sin \text{ 版}),$$

$$I_n = \frac{1}{n}\cos^{n-1} x\sin x + \frac{n-1}{n} I_{n-2} \quad (\cos \text{ 版}).$$

证明 仅证 sin 版. 拆 $\sin^n x = \sin^{n-1} x \cdot \sin x$, 取 $u = \sin^{n-1} x$, $dv = \sin x\,dx$, 则 $du = (n-1)\sin^{n-2} x\cos x\,dx$, $v = -\cos x$. 由分部积分,

$$I_n = -\sin^{n-1} x\cos x + (n-1)\int \sin^{n-2} x\cos^2 x\,dx.$$

代入 $\cos^2 x = 1 - \sin^2 x$,

$$I_n = -\sin^{n-1} x\cos x + (n-1)\int \sin^{n-2} x\,dx - (n-1)\int \sin^n x\,dx = -\sin^{n-1} x\cos x + (n-1)I_{n-2} - (n-1)I_n.$$

移项得 $nI_n = -\sin^{n-1} x\cos x + (n-1)I_{n-2}$, 整理即得 sin 版递推.cos 版推导对偶 (取 $u = \cos^{n-1} x$, $dv = \cos x\,dx$). □

注 17.8 (技巧点)
把 $\sin^n x$ 拆为 $\sin^{n-1} x \cdot \sin x$, 选 $u = \sin^{n-1} x$, 微分降一次, $dv = \sin x\,dx$, 积分变 $-\cos x$. 则 $du = (n-1)\sin^{n-2} x\cos x\,dx$, $v = -\cos x$. 由 $\int u\,dv = uv - \int v\,du$, 得 $I_n = -\sin^{n-1} x\cos x + (n-1)\int \sin^{n-2} x\cos^2 x\,dx$. 把 $\cos^2 x = 1 - \sin^2 x$ 拆回, 得 $nI_n = -\sin^{n-1} x\cos x + (n-1)I_{n-2}$, 整理即得递推公式.cos 版的推导对偶.

例题 17.18 求不定积分 $\int \sin^4 x\,dx$.

解 使用 sin 版递推 ($n = 4$, $I_0 = \int 1\,dx = x + C$):

$$I_4 = -\frac{\sin^3 x\cos x}{4} + \frac{3}{4} I_2.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对 I_2 再用一次 $(n=2)$:

$$I_2=-\frac{\sin x\cos x}{2}+\frac{1}{2}I_0=-\frac{\sin 2x}{4}+\frac{x}{2}.$$

整理

$$I_4=-\frac{\sin^3 x\cos x}{4}+\frac{3}{4}\left(-\frac{\sin 2x}{4}+\frac{x}{2}\right)=-\frac{\sin^3 x\cos x}{4}-\frac{3\sin 2x}{16}+\frac{3x}{8}+C.$$

作为对照, 直接用半角公式 $\sin^2 x=(1-\cos 2x)/2$ 展开: $\sin^4 x=(1-2\cos 2x+\cos^2 2x)/4=\frac{1}{4}-\frac{\cos 2x}{2}+\frac{1+\cos 4x}{8}$, 积分得 $\frac{3x}{8}-\frac{\sin 2x}{4}+\frac{\sin 4x}{32}+C$. 二者经恒等化简一致.

定理 17.10 ($\sec^n x$ 的递推) 设 $n\geq 3$. 记 $J_n=\int \sec^n x\,dx$, 则

$$J_n=\frac{\sec^{n-2}x\,\tan x}{n-1}+\frac{n-2}{n-1}J_{n-2}.$$

证明 拆 $\sec^n x=\sec^{n-2}x\cdot\sec^2 x$, 取 $u=\sec^{n-2}x$, $dv=\sec^2 x\,dx$, 则 $du=(n-2)\sec^{n-2}x\tan x\,dx$, $v=\tan x$. 由分部积分,

$$J_n=\sec^{n-2}x\tan x-(n-2)\int \sec^{n-2}x\tan^2 x\,dx.$$

代入 $\tan^2 x=\sec^2 x-1$,

$$J_n=\sec^{n-2}x\tan x-(n-2)\int \sec^n x\,dx+(n-2)\int \sec^{n-2}x\,dx=\sec^{n-2}x\tan x-(n-2)J_n+(n-2)J_{n-2}.$$

移项得 $(n-1)J_n=\sec^{n-2}x\tan x+(n-2)J_{n-2}$, 整理即得递推公式. □

注 17.9 (技巧点)
拆 $\sec^n x=\sec^{n-2}x\cdot\sec^2 x$. 取 $u=\sec^{n-2}x$, 微分得 $(n-2)\sec^{n-2}x\tan x$, $dv=\sec^2 x\,dx$, 积分得 $\tan x$. 用 $\sec^2 x=1+\tan^2 x$ 拆回, 得到同时含 J_n 与 J_{n-2} 的方程, 可按相似方式解出 J_n. 起点 $J_1=\int\sec x\,dx=\ln|\sec x+\tan x|+C$, $J_2=\tan x+C$.

定理 17.11 ($\int\frac{dx}{(x^2+a^2)^n}$ 的递推) 设 $a>0, n\geq 2$. 记 $K_n=\int\frac{dx}{(x^2+a^2)^n}$, 则

$$K_n=\frac{1}{2a^2(n-1)}\left(\frac{x}{(x^2+a^2)^{n-1}}+(2n-3)\,K_{n-1}\right).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 对恒等式 $\dfrac{d}{dx}\left[\dfrac{x}{(x^2+a^2)^{n-1}}\right]=\dfrac{1}{(x^2+a^2)^{n-1}}-\dfrac{2(n-1)x^2}{(x^2+a^2)^n}$ 两端积分, 得

$$\frac{x}{(x^2+a^2)^{n-1}}=K_{n-1}-2(n-1)\int\frac{x^2}{(x^2+a^2)^n}\,dx.$$

拆 $x^2=(x^2+a^2)-a^2$, 则 $\displaystyle\int\frac{x^2}{(x^2+a^2)^n}\,dx=\int\frac{dx}{(x^2+a^2)^{n-1}}-a^2\int\frac{dx}{(x^2+a^2)^n}=K_{n-1}-a^2K_n$. 代回,

$$\frac{x}{(x^2+a^2)^{n-1}}=K_{n-1}-2(n-1)(K_{n-1}-a^2K_n)=(3-2n)K_{n-1}+2(n-1)a^2K_n.$$

解出 K_n,

$$K_n=\frac{1}{2a^2(n-1)}\left(\frac{x}{(x^2+a^2)^{n-1}}+(2n-3)K_{n-1}\right).$$

□

注 17.10 (技巧点)
起点 $K_1=\frac{1}{a}\arctan\frac{x}{a}+C$. 推导思路是对 $\int\frac{dx}{(x^2+a^2)^n}$ 用分部 $\int\frac{1}{(x^2+a^2)^{n-1}}\cdot\frac{1}{x^2+a^2}\,dx$, 取 $u=1/(x^2+a^2)^{n-1}$, $dv=dx/(x^2+a^2)$. 利用 $\frac{d}{dx}\left(\frac{x}{(x^2+a^2)^{n-1}}\right)$ 的代数关系整理后即可解出 K_n 与 K_{n-1} 的线性关系. 也可借助三角代换 $x=a\tan\theta$ 化简后再分部.

例题 17.19 求不定积分 $K_2=\displaystyle\int\frac{dx}{(x^2+1)^2}$.

解 直接令 $x=\tan\theta,\theta\in(-\pi/2,\pi/2)$, 则 $dx=\sec^2\theta\,d\theta$、$x^2+1=\sec^2\theta$、$(x^2+1)^2=\sec^4\theta$. 故

$$K_2=\int\frac{\sec^2\theta\,d\theta}{\sec^4\theta}=\int\cos^2\theta\,d\theta=\int\frac{1+\cos2\theta}{2}\,d\theta=\frac{\theta}{2}+\frac{\sin2\theta}{4}+C.$$

回代: $\theta=\arctan x$, $\sin2\theta=2\sin\theta\cos\theta=\dfrac{2x}{\sqrt{1+x^2}}\cdot\dfrac{1}{\sqrt{1+x^2}}=\dfrac{2x}{1+x^2}$. 故

$$K_2=\frac{1}{2}\arctan x+\frac{x}{2(x^2+1)}+C.$$

作为验证, 用递推公式重算:

$$\begin{aligned}K_2&=\frac{1}{2\cdot1\cdot1}\left(\frac{x}{x^2+1}+1\cdot K_1\right)\\&=\frac{1}{2}\left(\frac{x}{x^2+1}+\arctan x\right)+C.\end{aligned}$$

知乎@余命数

与上述结果完全一致.

17.4.3 分部积分与分部求和的对偶

分部积分法的思想并不局限于连续的积分领域, 它在离散的求和世界中也有对应, 即**分部求和法**, 或称**阿贝尔求和公式**. 从一个更高的视角看, 两种方法都是把一个算子 (微分或差分) 从一个因子转移到另一个因子上, 只是一个用微分, 一个用差分.

先引入离散微积分中的**前向差分算子** Δ. 对于一个数列 $\{a_k\}$,其定义为 $\Delta a_k = a_{k+1} - a_k$. 这个算子是微分算子 $\frac{d}{dx}$ 的离散模拟. 其逆运算是求和算子 $\sum$, 这构成了**离散微积分基本定理**.

$$\sum_{k=m}^{n-1} \Delta a_k = \sum_{k=m}^{n-1} (a_{k+1} - a_k) = a_n - a_m$$

这与 $\int_a^b F'(x)dx = F(b) - F(a)$ 的结构完全一致.

接着, 我们推导离散世界中的乘法法则. 考虑两个数列乘积 $a_k B_k$ 的差分.

$$\begin{aligned}
\Delta(a_k B_k) &= a_{k+1}B_{k+1} - a_k B_k \\
&= a_{k+1}B_{k+1} - a_{k+1}B_k + a_{k+1}B_k - a_k B_k \\
&= a_{k+1}(B_{k+1} - B_k) + (a_{k+1} - a_k)B_k \\
&= a_{k+1}(\Delta B_k) + (\Delta a_k)B_k
\end{aligned}$$

移项整理, 得 $\sum$ 算子下分部法则的雏形.

$$a_{k+1}(\Delta B_k) = \Delta(a_k B_k) - (\Delta a_k)B_k$$

对此恒等式两边从 $k = 1$ 到 $n - 1$ 求和.

$$\sum_{k=1}^{n-1} a_{k+1}(\Delta B_k) = \sum_{k=1}^{n-1} \Delta(a_k B_k) - \sum_{k=1}^{n-1} (\Delta a_k)B_k$$

右边的第一项是裂项和, 等于 $a_n B_n - a_1 B_1$. 经过指标变换并定义 $b_k = \Delta B_{k-1}$, 即得其最常用的形式.

定理 17.12 (分部求和法) 给定两个数列 $\{a_k\}$ 和 $\{b_k\}$. 令 $B_n = \sum_{k=1}^{n} b_k$ 为 $\{b_k\}$ 的部分和, 且约定 $B_0 = 0$. 则对于任意正整数 $n \geq 1$,

$$\sum_{k=1}^{n} a_k b_k = a_n B_n - \sum_{k=1}^{n-1} B_k (a_{k+1} - a_k) \tag{17-4}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 由 $B_k - B_{k-1} = b_k$ 可得

$$\sum_{k=1}^{n} a_k b_k = \sum_{k=1}^{n} a_k (B_k - B_{k-1}) = \sum_{k=1}^{n} a_k B_k - \sum_{k=1}^{n} a_k B_{k-1}.$$

第二个和式中 $B_0 = 0$, 故

$$\sum_{k=1}^{n} a_k B_{k-1} = \sum_{k=2}^{n} a_k B_{k-1} = \sum_{j=1}^{n-1} a_{j+1} B_j.$$

因而

$$\begin{aligned}\sum_{k=1}^{n} a_k b_k &= \left(\sum_{k=1}^{n-1} a_k B_k + a_n B_n\right) - \sum_{k=1}^{n-1} a_{k+1} B_k \\ &= a_n B_n - \sum_{k=1}^{n-1} B_k (a_{k+1} - a_k).\end{aligned}$$

这是所需公式. □

17.4.3.1 连续与离散的对偶性

阿贝尔求和公式与分部积分公式之间存在一一对应关系, 揭示了连续数学与离散数学之间基本的结构同构性.

分部积分(连续)		分部求和(离散)
$\int u\,dv$	$\longleftrightarrow$	$\sum a_k b_k$
u	$\longleftrightarrow$	a_k
$dv = v'(x)dx$	$\longleftrightarrow$	b_k
$v = \int dv$(积分)	$\longleftrightarrow$	$B_k = \sum b_i$(求和)
$du = u'(x)dx$(微分)	$\longleftrightarrow$	$\Delta a_k = a_{k+1} - a_k$(差分)
$uv\Big\vert_a^b$(边界项)	$\longleftrightarrow$	$a_n B_n$(边界项)
$\int v\,du$(转化后的积分)	$\longleftrightarrow$	$\sum B_k \Delta a_k$(转化后的求和)

例题 17.20 利用分部求和法, 计算等差比数列之和 $S_n = \sum_{k=1}^{n} kq^k \quad (q \neq 1)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解　与分部积分中对多项式求导以降幂的做法对偶，此处取等差数列作差分. 令 $a_k=k, b_k=q^k$, 则 $\Delta a_k=(k+1)-k=1$, 而 b_k 的部分和为几何级数

$$B_k=\sum_{i=1}^{k} q^i=\frac{q(1-q^k)}{1-q}.$$

由阿贝尔求和公式 $\sum_{k=1}^{n} a_k b_k=a_n B_n-\sum_{k=1}^{n-1} B_k(a_{k+1}-a_k)$, 得

$$\begin{aligned} S_n &= n\cdot B_n-\sum_{k=1}^{n-1} B_k\cdot 1 \\ &= n\frac{q(1-q^n)}{1-q}-\sum_{k=1}^{n-1}\frac{q(1-q^k)}{1-q}. \end{aligned}$$

将常数因子 $\frac{q}{1-q}$ 提出, 剩余和式为

$$\begin{aligned} \sum_{k=1}^{n-1}(1-q^k) &= \sum_{k=1}^{n-1} 1-\sum_{k=1}^{n-1} q^k \\ &= (n-1)-\frac{q(1-q^{n-1})}{1-q}. \end{aligned}$$

代回原式整理,

$$\begin{aligned} S_n &= n\frac{q(1-q^n)}{1-q}-\frac{q}{1-q}\left[(n-1)-\frac{q(1-q^{n-1})}{1-q}\right] \\ &= \frac{nq-nq^{n+1}}{1-q}-\frac{q(n-1)}{1-q}+\frac{q^2(1-q^{n-1})}{(1-q)^2} \\ &= \frac{nq-nq^{n+1}-nq+q}{1-q}+\frac{q^2-q^{n+1}}{(1-q)^2} \\ &= \frac{q-nq^{n+1}}{1-q}+\frac{q^2-q^{n+1}}{(1-q)^2} \\ &= \frac{(q-nq^{n+1})(1-q)+q^2-q^{n+1}}{(1-q)^2} \\ &= \frac{q-q^2-nq^{n+1}+nq^{n+2}+q^2-q^{n+1}}{(1-q)^2} \\ &= \frac{q-(n+1)q^{n+1}+nq^{n+2}}{(1-q)^2}. \end{aligned}$$

此结果与错位相减法所得一致, 而分部求和公式给出了一个更高观点的推导.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

17.5 定积分与微积分基本定理

前面的讨论把不定积分定义为微分的逆运算. 它提供一个函数族, 回答了何种函数的导数是 $f(x)$ 这一纯粹的分析问题. 然而, 积分学的历史根源和其在物理、几何等领域的广泛应用, 都指向一个更具体的问题, 即**积累问题**.

无论是计算曲线下的面积、变速直线运动在某段时间内的位移, 还是变力在某个过程中所作的功, 它们都是对一个连续变化的量在指定区间上进行无限求和, 以得到总积累量. 这一思想, 经过黎曼的严格化, 形成了定积分的现代理论.

17.5.1 定积分

设想我们要计算一个连续函数 $y=f(x)$ 在区间 $[a,b]$ 上与 x 轴所围成的曲边梯形的面积. 阿基米德的穷竭法启发我们, 可以用一系列规则图形, 如矩形的面积来逼近这个不规则图形的面积.

1. **分割**. 将区间 $[a,b]$ 任意地分割成 n 个小区间. 这个分割记为 $\mathcal{P}=\{x_0,x_1,\dots,x_n\}$, 其中 $a=x_0<x_1<\cdots<x_n=b$. 第 i 个小区间的长度为 $\Delta x_i=x_i-x_{i-1}$.
2. **近似**. 在每个小区间 $[x_{i-1},x_i]$ 内, 任意选取一个**样本点** ξ_i. 接着, 用一个高为 $f(\xi_i)$、宽为 Δx_i 的矩形面积来近似该小区间上曲边梯形的面积.
3. **求和**. 将所有这些小矩形的面积加总, 得到总面积的一个近似值. 这个和式被称为**黎曼和**.

$$S_{\mathcal{P}}=\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

这个黎曼和的值, 依赖于我们如何分割区间 $\mathcal{P}$ 以及如何在每个小区间内选取样本点 ξ_i.

直观上, 当我们让分割无限地细密, 所有小区间都变得无限窄时, 黎曼和就应该逼近一个确定的值, 即曲边梯形的精确面积. 用分割的**模** $\|\mathcal{P}\|=\max_i\{\Delta x_i\}$ 来度量分割的精细程度.

定义 17.3 (定积分) 若当分割的模 $\|\mathcal{P}\|\to 0$ 时, 黎曼和 $S_{\mathcal{P}}$ 的极限存在, 且此极限值与样本点 ξ_i 的选取方式无关, 则称函数 $f(x)$ 在区间 $[a,b]$ 上是**可积的**. 将这个唯一的极限值定义为 $f(x)$ 从 a 到 b 的**定积分**, 记作

$$\int_a^b f(x)\,dx \tag{17-5}$$

这里的积分号 $\int$ 是 Summa(求和)首字母 S 的拉长形式, 它与求和号 $\sum$ 在实质上遥相呼应.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

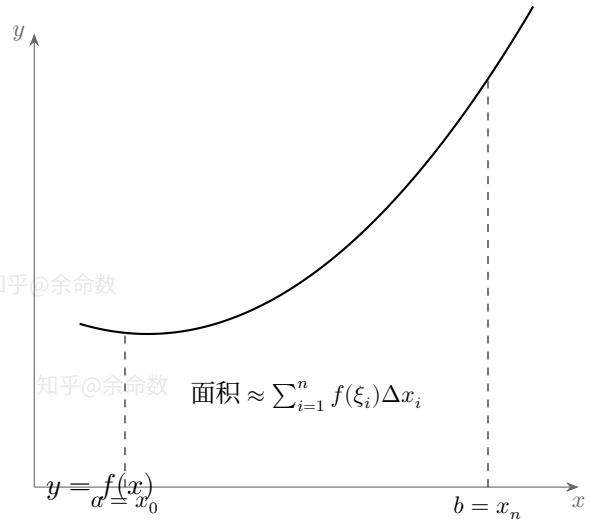


图 17.3　对非均匀分割的黎曼和

什么函数可积, 是可积性理论首先要回答的问题. 连续性足以保证可积性, 这是微积分学的一个基石性结论.

定理 17.13 (连续函数的可积性) 设 $a < b$. 若函数 $f(x)$ 在闭区间 $[a, b]$ 上连续, 则 $f(x)$ 在该区间上必定可积.

证明 回到可积性的严格判据: $f(x)$ 在 $[a, b]$ 上可积当且仅当对任意 $\epsilon > 0$, 存在分割 $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ 使上和与下和之差

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i < \epsilon,$$

其中 $M_i = \sup_{[x_{i-1}, x_i]} f$, $m_i = \inf_{[x_{i-1}, x_i]} f$. 故须用 f 的连续性控制 $M_i - m_i$.

闭区间上的连续函数必一致连续 (海涅-康托尔定理, 实数完备性的推论). 取 $\epsilon' = \frac{\epsilon}{b-a}$, 则存在 $\delta > 0$, 使对 $[a, b]$ 中任意 x', x'', 只要 $|x' - x''| < \delta$ 就有 $|f(x') - f(x'')| < \epsilon'$, 且 δ 只依赖 ϵ' 而与点的位置无关.

取正整数 n 使 $\frac{b-a}{n} < \delta$, 对 $[a, b]$ 作 n 等分, 每个小区间宽度 $\Delta x_i = \frac{b-a}{n} < \delta$. 由极值定理, f 在 $[x_{i-1}, x_i]$ 上取到最大值 $M_i = f(x_i'')$ 与最小值 $m_i = f(x_i')$, 而 $|x_i'' - x_i'| \le \Delta x_i < \delta$, 故由一致连续性

$$M_i - m_i = |f(x_i'') - f(x_i')| < \epsilon' = \frac{\epsilon}{b-a}, \quad i = 1, 2, \dots, n.$$

于是

$$U(f, \mathcal{P}) - L(f, \mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$< \frac{\epsilon}{b-a}\sum_{i=1}^{n}\Delta x_i = \frac{\epsilon}{b-a}(b-a) = \epsilon.$$

对任意 $\epsilon > 0$ 均可造出这样的分割, 由黎曼可积的充要条件, f 在 $[a,b]$ 上可积. □

17.5.2 可积的必要条件与充分条件

至此已看到连续函数是可积的, 但这只是一个**充分条件**. 什么样的函数*必定*可积, 什么样的函数*必定不可积*, 下列条件在判别一个陌生的被积函数时极为有用.

定理 17.14 (可积的必要条件) 若函数 f 在闭区间 $[a,b]$ 上(Riemann 意义下)可积, 则 f 在 $[a,b]$ 上必有界.

证明 反证. 设 f 在 $[a,b]$ 上无界. 对任意给定的实数 $M > 0$ 与任意分割 $\mathcal{P}$, 至少存在某个小区间 $[x_{j-1}, x_j]$ 使 f 在其上无界. 在该区间内选取样本点 ξ_j 使 $|f(\xi_j)|$ 充分大(具体地, $|f(\xi_j)| > \left(M + \sum_{i\neq j}|f(\xi_i)|\,\Delta x_i\right)/\Delta x_j$); 其它小区间任意取样本点. 此时

$$\left|\sum_{i=1}^{n} f(\xi_i)\,\Delta x_i\right| \geq |f(\xi_j)|\Delta x_j - \sum_{i\neq j}|f(\xi_i)|\Delta x_i > M.$$

因 M 任意大, 黎曼和 $S_{\mathcal{P}}$ 不可能收敛到任何确定的极限, 与 f 可积相矛盾. 故 f 必有界. □

注 17.11
此必要条件反过来并不成立. 狄利克雷函数 $D(x)$ 在 $[0,1]$ 上有界 ($|D| \leq 1$), 但在前文已证其黎曼不可积. 故有界严格弱于可积.

定理 17.15 (有限间断点函数的可积性) 若 f 在 $[a,b]$ 上有界, 且只有**有限个**第一类间断点(跳跃或可去), 则 f 在 $[a,b]$ 上可积.

证明 还将在 第17.5.3小节 中给出此结论的严格论述. 关键在于有限个坏点可用总长度任意小的区间隔离, 在此区间外 f 连续从而可积, 而在隔离区间内 f 有界故其上下和之差的贡献可任意小. 因此整体 $U - L$ 可任意小, f 黎曼可积. □

定理 17.16 (单调函数的可积性) 若 f 在 $[a,b]$ 上单调, 则 f 在 $[a,b]$ 上可积.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 不妨设 f 单调递增. 对任意分割 $\mathcal{P}$, 在每个小区间 $[x_{i-1}, x_i]$ 上由单调性 $M_i - m_i \le f(x_i) - f(x_{i-1})$, 故

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \le \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\Delta x_i \le (f(b) - f(a))\,\|\mathcal{P}\|.$$

若 f 为常数则平凡. 否则 $f(b) - f(a) > 0$, 对任给 $\varepsilon > 0$, 取 $\delta = \varepsilon/(f(b) - f(a))$, 则 $\|\mathcal{P}\| < \delta$ 时 $U(f,\mathcal{P}) - L(f,\mathcal{P}) < \varepsilon$, 故 f 可积. 单调递减情形对偶. □

注 17.12 (技巧点)
单调函数不一定连续, 甚至可以有可数无穷多个间断点, 例如阶梯型函数 $f(x) = \lfloor x \rfloor$ 在 $[0, n]$ 上有 n 个间断点. 但单调性便于控制振荡, 对任意分割 $\mathcal{P}$, 在每个小区间 $[x_{i-1}, x_i]$ 上由单调性必有 $M_i = \max f$ 与 $m_i = \min f$ 取在端点附近, 故 $M_i - m_i \le |f(x_i) - f(x_{i-1})|$. 于是

$$U(f,\mathcal{P}) - L(f,\mathcal{P}) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i \le \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \cdot \Delta x_i \le |f(b) - f(a)| \cdot \|\mathcal{P}\|.$$

只要 $\|\mathcal{P}\| \to 0$, 右端 $\to 0$, 即 $U - L$ 可任意小, 故 f 可积. 完整的 ε-δ 推导即取 $\delta = \varepsilon/|f(b) - f(a)|$ 当 f 非常数, 即得 $U - L < \varepsilon$.

17.5.3 当函数不连续时

连续性是可积的一个**充分条件**, 同时存在不连续的可积函数. 当一个函数不再连续时, 其可积性取决于不连续点的数量与分布.

17.5.3.1 有限个间断点

若函数 $f(x)$ 在 $[a, b]$ 上有界, 且只有有限个间断点, 例如跳跃间断点, 则函数**仍然是可积的**. 其直观思想是, 这些有限个坏点的影响, 可以通过将其隔离在一些总长度任意小的区间内而被控制住. 例如, 对于函数 $f(x) = \begin{cases} 1, & x \ge 0 \\ -1, & x < 0 \end{cases}$ 在 $[-1, 1]$ 上的积分. 它仅在 $x = 0$ 处有一个间断点. 可将积分拆分为 $\int_{-1}^{0} (-1)dx + \int_{0}^{1} (1)dx = -1 + 1 = 0$. 严格的证明是, 可用一个极窄的区间 $(-\delta, \delta)$ 将 $x = 0$ 包围起来. 在这个区间内, 函数的振荡 $M - m = 1 - (-1) = 2$ 是固定的, 但这个区间对 $U - L$ 的贡献是 $2 \cdot (2\delta)$, 可以随 $\delta \to 0$ 而任意小. 在区间 $[-1, -\delta]$ 和 $[\delta, 1]$ 上, 函数是连续的, 因此可积, 其 $U - L$ 之差也可以做得任意小. 两部分合起来, 总的 $U - L$ 仍然可以小于任意给定的 ϵ.

17.5.3.2 无穷个间断点

当间断点的数量达到无穷时, 情况变得极为微妙. 此时, 函数可能可积, 也可能不可积.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

17.5.3.3 狄利克雷函数(不可积)

考虑定义在 $[0,1]$ 上的狄利克雷函数:

知乎@余命数

$$D(x)=\begin{cases}1, & x\in\mathbb{Q}\\ 0, & x\notin\mathbb{Q}\end{cases}$$

知乎@余命数

这个函数在每一点都不连续. 对于 $[0,1]$ 上的任何一个分割 $\mathcal{P}$, 在任意一个小区间 $[x_{i-1},x_i]$ 内, 由于有理数和无理数的稠密性, 我们总能找到有理数和无理数. 因此, 每个小区间上的上确界 M_i 永远是 1, 而下确界 m_i 永远是 0. 于是, 对于**任何**分割 $\mathcal{P}$:

$$U(D,\mathcal{P})=\sum_{i=1}^{n}1\cdot\Delta x_i=1-0=1$$

$$L(D,\mathcal{P})=\sum_{i=1}^{n}0\cdot\Delta x_i=0$$

上和与下和之差恒为 1, 永远无法小于任意正数 ϵ. 因此, 狄利克雷函数是黎曼不可积的.

知乎@余命数

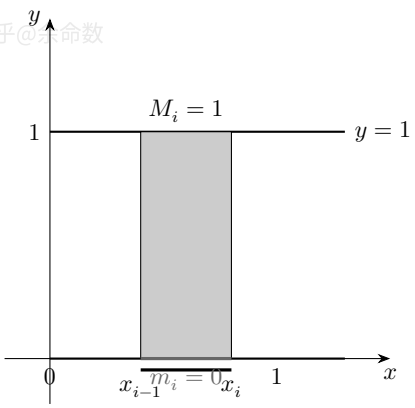


图 17.4 狄利克雷函数在任意子区间上的上确界为 1、下确界为 0, 故不可积

知乎@余命数

17.5.3.4 可积的托玛函数

然而, 有些含无穷间断点的函数仍可积. 考虑托玛函数 $t(x)$:

知乎@余命数

知乎@余命数

$$t(x)=\begin{cases}1/q, & x=p/q\in\mathbb{Q}\text{(最简分数)}\\ 0, & x\notin\mathbb{Q}\\ 1, & x=0\end{cases}$$

这个函数在所有有理点上间断, 在所有无理点上连续. 尽管其间断点是无穷且稠密的, 但它在 $[0,1]$ 上却是黎曼可积的, 其积分值为 0. 其深层原因是, 对于任何 $\epsilon>0$, 只有限个有理点的函数值 $t(x)=1/q$ 大于 ϵ(因为这要求分母 $q<1/\epsilon$). 可将这有限个主要间断点用总长度任意小的区间隔离起来, 而在剩下的区间上, 函数的振荡 $M-m$ 小于 ϵ. 从而, 总的 $U-L$ 之差可以被控制.

知乎@余命数

17.5.4 微积分基本定理

至此, 我们面对着两个在起源、定义和形式上都截然不同的积分概念.

- **不定积分** 是一个纯粹的分析概念, 它作为求导的逆运算, 给出一个函数族 $F(x)+C$.
- **定积分**是一个源于几何与物理的概念, 它作为黎曼和的极限, 给出一个数值.

数学史上最伟大的发现之一, 是牛顿和莱布尼茨发现了这两个概念之间存在着内在的联系. 这种联系就是**微积分基本定理**. 它包含两个部分, 把微分与积分联系起来.

17.5.4.1 第一基本定理

知乎@余命数

为了探寻二者的联系, 我们构造一个辅助函数, 即**变上限积分函数**. 对于一个在 $[a,b]$ 上连续的函数 $f(t)$, 可定义一个新函数 $A(x)$:

$$A(x)=\int_a^x f(t)\,dt \quad (x\in[a,b])$$

这个函数 $A(x)$ 的几何意义是 $f(t)$ 的图像从 a 到 x 所累积的有向面积. 微积分第一基本定理精确地给出了这个累积函数 $A(x)$ 的瞬时变化率.

定理 17.17 (微积分第一基本定理) 若 $f(t)$ 在区间 $[a,b]$ 上连续, 则由 $A(x)=\int_a^x f(t)\,dt$ 定义的函数 $A(x)$ 在 (a,b) 上可导, 并且

$$A'(x)=\frac{d}{dx}\left(\int_a^x f(t)\,dt\right)=f(x) \tag{17-6}$$

证明 考察 $A(x)$ 的导数定义式.

$$A'(x)=\lim_{h\to0}\frac{A(x+h)-A(x)}{h}$$

知乎@余命数

根据 $A(x)$ 的定义和定积分的区间可加性,

$$A(x+h)-A(x)=\int_a^{x+h}f(t)\,dt-\int_a^x f(t)\,dt=\int_x^{x+h}f(t)\,dt$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个积分值表示了在极窄区间 $[x, x+h]$ 上的面积. 现给出严格估计. 当 $h \neq 0$ 且

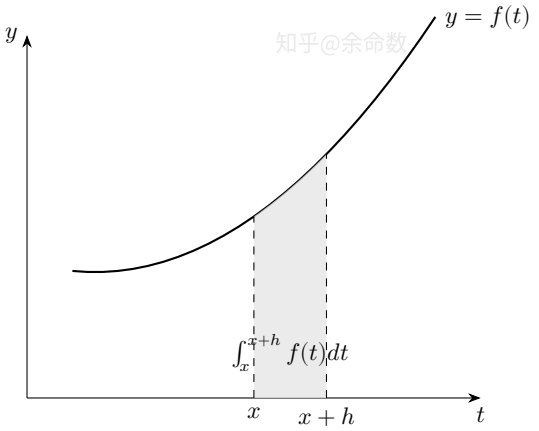


图 17.5 第一基本定理的几何直观

$x+h \in [a,b]$ 时,

$$\begin{aligned}\frac{A(x+h)-A(x)}{h}-f(x) &= \frac{1}{h}\int_x^{x+h} f(t)\,dt - f(x)\\ &= \frac{1}{h}\int_x^{x+h} (f(t)-f(x))\,dt.\end{aligned}$$

由 f 在点 x 处连续,对任意 $\varepsilon>0$,存在 $\delta_1>0$,使得当 $|t-x|<\delta_1$ 时,$|f(t)-f(x)|<\varepsilon$. 又因 $x \in (a,b)$, 可取

$$\delta = \min\{\delta_1,\, x-a,\, b-x\} > 0.$$

当 $0<|h|<\delta$ 时, $x+h \in [a,b]$, 且积分区间 $[x, x+h]$ (或按有向积分理解的 $[x+h, x]$)内所有 t 都满足该不等式, 因而

$$\left|\frac{A(x+h)-A(x)}{h}-f(x)\right| \le \frac{1}{|h|}\int_{\min(x,x+h)}^{\max(x,x+h)} |f(t)-f(x)|\,dt < \varepsilon.$$

故极限存在且等于 $f(x)$, 即 $A'(x)=f(x)$. □

第一基本定理说明, 对连续函数 f, 变上限积分 $A(x)=\int_a^x f(t)\,dt$ 给出它的一个原函数. 这证明了连续函数必有原函数, 也表明微分与积分分别作为变化率与总积累的互逆关系.

17.5.4.2 第二基本定理

第一基本定理是一个存在性与结构性的定理, 而第二基本定理则是一个实用的计算工具. 它是我们通常所说的微积分基本定理.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 17.18 (微积分第二基本定理 (牛顿-莱布尼茨公式)) 如果函数 $f(x)$ 在区间 $[a,b]$ 上连续, 且 $F(x)$ 是 $f(x)$ 在该区间上的**任意一个**原函数, 即 $F'(x)=f(x)$, 那么

$$\int_a^b f(x)\,dx = F(b)-F(a) \tag{17-7}$$

证明 由第一基本定理, $A(x)=\int_a^x f(t)\,dt$ 是 $f(x)$ 的一个原函数. 设 $F(x)$ 是 $f(x)$ 的任一原函数, 由原函数结构定理, 二者只差一个常数 C, 即 $A(x)=F(x)+C$. 令 $x=a$ 确定 C, 因 $A(a)=\int_a^a f(t)\,dt=0$, 故 $0=F(a)+C$, 即 $C=-F(a)$, 从而 $\int_a^x f(t)\,dt=F(x)-F(a)$. 令 $x=b$ 即得 $\int_a^b f(x)\,dx=F(b)-F(a)$. □

为方便书写, 我们通常引入记号 $F(x)\big|_a^b$ 来表示 $F(b)-F(a)$. 于是, 公式可以写为 $\int_a^b f(x)\,dx=F(x)\big|_a^b$.

这个定理的伟大之处在于, 它将一个原则上需要计算无穷和的极限问题 (求定积分), 转化为了一个简单的代数问题.

1. 找到被积函数 $f(x)$ 的一个原函数 $F(x)$(通过不定积分).
2. 计算原函数在积分区间两个端点处的值, 并求其差.

这是数学思想的一次巨大飞跃, 它使得曾经只有像阿基米德那样的天才才能解决的面积问题, 变成了任何掌握了基本积分公式的人都能完成的常规计算.

17.5.4.3 变限积分的复合求导法则

第一基本定理对变上限 $\int_a^x f(t)\,dt$ 处处可导给出了 $A'(x)=f(x)$. 在应用中,积分上下限常以更一般的函数给出, 例如 $\int_0^{x^2} f(t)\,dt$ 或 $\int_{\sin x}^{e^x} f(t)\,dt$. 下列推论 (复合变限求导)是第一基本定理与链式法则的直接组合.

推论 17.2 (复合变限求导) 设 f 在 $[a,b]$ 上连续, α,β 在某区间 I 上可导且 $\alpha(I)\subset[a,b]$, $\beta(I)\subset[a,b]$. 定义

$$\Phi(x)=\int_{\alpha(x)}^{\beta(x)} f(t)\,dt.$$

则 Φ 在 I 上可导, 且

$$\Phi'(x)=f(\beta(x))\,\beta'(x)-f(\alpha(x))\,\alpha'(x).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 令 $F(u)=\int_a^u f(t)\,dt$，由第一基本定理知 F 在 $[a,b]$ 上可导且 $F'(u)=f(u)$. 注意到 $\Phi(x)=F(\beta(x))-F(\alpha(x))$，由链式法则分别求两项：

$$\frac{d}{dx}F(\beta(x))=F'(\beta(x))\beta'(x)=f(\beta(x))\beta'(x),$$

$$\frac{d}{dx}F(\alpha(x))=F'(\alpha(x))\alpha'(x)=f(\alpha(x))\alpha'(x).$$

两式相减即得证. □

例题 17.21 求导数 $\dfrac{d}{dx}\displaystyle\int_0^{x^2}\sin(t^2)\,dt$.

解 在此取常数下限 $\alpha(x)=0$（此时 $\alpha'(x)=0$），上限 $\beta(x)=x^2$. 被积函数 $f(t)=\sin(t^2)$. 由复合变限求导法则

$$\frac{d}{dx}\int_0^{x^2}\sin(t^2)\,dt=\sin\big((x^2)^2\big)\cdot(x^2)'-\sin(0^2)\cdot 0=\sin(x^4)\cdot 2x=2x\sin(x^4).$$

例题 17.22 计算定积分 $\displaystyle\int_0^{\pi}\sin x\,dx$.

解 被积函数 $f(x)=\sin x$ 的一个原函数为 $F(x)=-\cos x$（$F'(x)=\sin x$）. 由微积分基本定理，

$$\int_0^{\pi}\sin x\,dx=(-\cos x)\Big|_0^{\pi}=(-\cos\pi)-(-\cos 0)=1+1=2.$$

正弦曲线在 $[0,\pi]$ 上的一个拱形所围区域的精确面积恰为 2.

例题 17.23 计算定积分 $\displaystyle\int_1^{e}\frac{\ln x}{x}\,dx$.

解 被积函数 $f(x)=\frac{\ln x}{x}$ 的不定积分宜用换元法求出. 令 $u=\ln x$，则 $du=\frac{1}{x}dx$；定积分换元须同步变换上下限，$x=1$ 时 $u=0$，$x=e$ 时 $u=1$. 于是

$$\int_1^{e}\frac{\ln x}{x}\,dx=\int_0^{1}u\,du=\frac{u^2}{2}\Big|_0^{1}=\frac{1}{2}.$$

直接变换积分限免去了求出原函数后回代 x 的步骤，是处理定积分换元的标准做法.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

17.5.5 定积分的极限应用

黎曼和的极限定义了定积分，也给出了将求和极限化为定积分的实用方法. 当某数列极限形如 n 等分区间上的函数求和形式 $\sum f(\frac{i}{n})\,\frac{1}{n}$（或更一般地 $\sum f(a+i\Delta x)\Delta x, \Delta x=(b-a)/n$），我们便可反向利用定积分定义，把这个极限识别为某个连续函数在某区间上的定积分，再用牛顿-莱布尼茨公式计算.

定理 17.19 (由黎曼和还原定积分) 设 f 在 $[a,b]$ 上连续，$\Delta x=(b-a)/n$. 则

$$\lim_{n\to\infty}\sum_{i=1}^{n} f(a+i\,\Delta x)\,\Delta x=\int_a^b f(x)\,dx.$$

证明 对每个正整数 n，取 n 等分分割 $\mathcal{P}_n=\{a,a+\Delta x,a+2\Delta x,\dots,b\}$，其模 $\|\mathcal{P}_n\|=\Delta x=(b-a)/n$，当 $n\to\infty$ 时它趋于 0. 在每个小区间 $[a+(i-1)\Delta x,a+i\Delta x]$ 上选取样本点为右端点 $\xi_i=a+i\Delta x$. 此分割下相应的黎曼和为

$$S_n=\sum_{i=1}^{n} f(\xi_i)\Delta x=\sum_{i=1}^{n} f(a+i\Delta x)\Delta x.$$

由于 f 在 $[a,b]$ 上连续，故可积（前文已证），可积函数的黎曼和对任意细分分割与任意样本点选取都收敛到同一极限 $\int_a^b f(x)\,dx$. 当 $\|\mathcal{P}_n\|\to 0$ 时即得 $S_n\to\int_a^b f(x)\,dx$. □

例题 17.24 求极限 $L=\lim_{n\to\infty}\dfrac{1}{n}\sum_{i=1}^{n}\sqrt{1+\dfrac{i}{n}}$.

解 取 $\Delta x=1/n$、$a=0$、$b=1$，被积函数 $f(x)=\sqrt{1+x}$，则样本点 $\xi_i=a+i\Delta x=i/n$，黎曼和

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{1+\frac{i}{n}}=\sum_{i=1}^{n} f(a+i\Delta x)\,\Delta x.$$

由定理，该极限即 $\int_0^1\sqrt{1+x}\,dx$. 直接积分得

$$\int_0^1\sqrt{1+x}\,dx=\frac{2}{3}(1+x)^{3/2}\Big|_0^1=\frac{2}{3}(2^{3/2}-1)=\frac{2}{3}(2\sqrt{2}-1)=\frac{4\sqrt{2}-2}{3},$$

故 $L=\frac{2}{3}(2\sqrt{2}-1)\approx 1.219$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

17.6 定积分的性质

微积分基本定理是计算定积分的实用工具, 而定积分本身来自黎曼和的极限. 因此, 它继承了求和运算 ($\sum$) 的结构, 并表现出与积分区间、被积函数大小相关的几何与分析性质. 掌握这些性质, 既能简化计算, 也能看清定积分作为算子的行为.

17.6.1 代数性质

定积分算子 $I(f)=\int_a^b f(x)\,dx$ 作用于一个函数空间之上. 在这个空间中, 最基本的运算是函数的加法与数乘. 定积分算子与这些运算兼容, 展现出**线性性质**.

知乎@余命数

定理 17.20 (定积分的线性性) 设 $a<b$. 若函数 $f(x)$ 与 $g(x)$ 均在 $[a,b]$ 上可积, k 为任意常数, 则 $kf(x)$ 与 $f(x)+g(x)$ 也在 $[a,b]$ 上可积, 且

1. **齐次性**: $\displaystyle\int_a^b kf(x)\,dx=k\int_a^b f(x)\,dx$
2. **可加性**: $\displaystyle\int_a^b [f(x)+g(x)]\,dx=\int_a^b f(x)\,dx+\int_a^b g(x)\,dx$

证明 先说明可积性在这些运算下保持. 若 f 可积, 则对任意分割 $\mathcal{P}$,

$$U(kf,\mathcal{P})-L(kf,\mathcal{P})=|k|(U(f,\mathcal{P})-L(f,\mathcal{P})),$$

因而 kf 可积. 若 f,g 可积, 对任意 $\varepsilon>0$ 可分别取分割使两者的上下和差小于 $\varepsilon/2$; 取这两个分割的公共加细后, 利用

$$U(f+g,\mathcal{P})-L(f+g,\mathcal{P})\le (U(f,\mathcal{P})-L(f,\mathcal{P}))+(U(g,\mathcal{P})-L(g,\mathcal{P}))$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

可得 $f+g$ 可积.

再由求和算子 $\sum$ 的线性性给出积分等式. 对 $kf(x)$ 的黎曼和有 $\sum_{i=1}^n kf(\xi_i)\Delta x_i=k\sum_{i=1}^n f(\xi_i)\Delta x_i$, 取极限 $\|\mathcal{P}\|\to 0$ 即得齐次性. 对 $f(x)+g(x)$ 的黎曼和有 $\sum_{i=1}^n[f(\xi_i)+g(\xi_i)]\Delta x_i=\sum_{i=1}^n f(\xi_i)\Delta x_i+\sum_{i=1}^n g(\xi_i)\Delta x_i$, 取极限即得可加性. □

从一个更抽象的视角看, 此定理表明, 在给定区间 $[a,b]$ 上的可积函数构成一个**线性空间**(或称向量空间), 而定积分算子是从这个函数空间到实数域 $\mathbb{R}$ 的一个**线性映射**, 也称线性泛函.

17.6.2 积分区间的性质

这些性质阐明了定积分的值如何依赖于其积分域 $[a,b]$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

1. **积分域退化**：若积分的上下限重合，则积分区间 $[a,a]$ 的长度为零.

$$\int_a^a f(x)\,dx = 0$$

这既可以从黎曼和的角度理解（所有 $\Delta x_i = 0$），也可以由牛顿-莱布尼茨公式直接得出 $F(a) - F(a) = 0$.

2. **积分域反向**：交换定积分的上下限，积分值反号.

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$$

此性质最初作为约定引入，它使得牛顿-莱布尼茨公式 $F(a) - F(b) = -(F(b) - F(a))$ 在 $a > b$ 时依然保持形式上的统一.

3. **积分域的可加性**：积分的区间可以被拆分或合并.

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

该性质对于 a, b, c 的任意排列顺序均成立，只要 $f(x)$ 在包含这三点的最大区间上可积. 其几何意义极为直观. 从 a 到 b 的总（有向）面积，等于从 a 到 c 的面积与从 c 到 b 的面积之和.

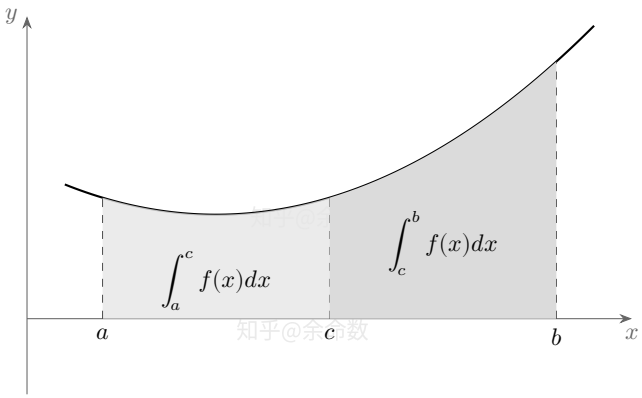


图 17.6　定积分的区间可加性

17.6.3 序性质与估值

这类性质将函数的大小关系（序关系）与积分值的大小关系联系起来，是进行积分估算的理论基础.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

4. **保序性**：若在区间 $[a,b]$ 上恒有 $f(x)\geq g(x)$，那么

$$\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx$$

特别地，若 $f(x)\geq 0$，则 $\int_a^b f(x)\,dx\geq 0$. 从黎曼和看，由于每个 $f(\xi_i)\geq g(\xi_i)$ 且 $\Delta x_i>0$，故和式的大小关系得以保持，取极限后亦然. 几何上，这表示更高的函数曲线所围成的（有向）面积也更大. 从算子的角度看，定积分是一个**保序算子**.

5. **估值定理**：设函数 $f(x)$ 在 $[a,b]$ 上的最大值为 M，最小值为 m. 那么

$$m(b-a)\leq \int_a^b f(x)\,dx \leq M(b-a)$$

证明 因为在 $[a,b]$ 上恒有 $m\leq f(x)\leq M$. 根据保序性，对这个不等式链进行积分.

$$\int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx$$

由于 $\int_a^b k\,dx=k(b-a)$，定理得证. □

几何上，这表示曲线下的真实面积一定被两个矩形的面积所夹住. 一个是以最小值为高的内接矩形，另一个是以最大值为高的外切矩形.

6. **积分的绝对值不等式**：

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx \quad (a<b)$$

证明 已知 $-|f(x)|\leq f(x)\leq |f(x)|$. 根据保序性，对这个不等式积分，得到

$$\int_a^b -|f(x)|\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b |f(x)|\,dx$$

即

$$-\int_a^b |f(x)|\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b |f(x)|\,dx$$

这是绝对值不等式 $|A|\leq B \iff -B\leq A\leq B$ 的形式. □

17.6.4 积分中值定理

估值定理给出了积分值的一个界. 在这个界之间，存在一个值能以平均的方式代表整个积分.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

定义 17.4 (函数的平均值) 函数 $f(x)$ 在区间 $[a,b]$ 上的**平均值**定义为

$$\bar{f}=\frac{1}{b-a}\int_a^b f(x)\,dx$$

这个定义的合理性在于,它推广了有限个数算术平均值的概念. 黎曼和 $\sum f(\xi_i)\Delta x_i$ 是对函数值的加权求和, 除以总长度 $b-a=\sum\Delta x_i$ 后再取极限, 就得到了连续情形下的平均值.

那么, 对于一个连续函数, 它是否一定能在某一点取到它的平均值, 积分中值定理给出了肯定的回答.

定理 17.21 (积分中值定理) 设 $a<b$. 若函数 $f(x)$ 在闭区间 $[a,b]$ 上连续, 则在 $[a,b]$ 内至少存在一点 ξ, 使得

$$f(\xi)=\frac{1}{b-a}\int_a^b f(x)\,dx \tag{17-8}$$

或者写成等价形式 $\int_a^b f(x)\,dx=f(\xi)(b-a)$.

证明 设 m 和 M 是 $f(x)$ 在 $[a,b]$ 上的最小值与最大值. 根据估值定理,

$$m(b-a)\le\int_a^b f(x)\,dx\le M(b-a)$$

两边同除以 $(b-a)$, 得

$$m\le\frac{1}{b-a}\int_a^b f(x)\,dx\le M$$

故函数的平均值 $\bar{f}$ 介于函数的最小值与最大值之间. 由于 $f(x)$ 在闭区间上连续, 根据**介值定理**, 对于任何一个介于 m 与 M 之间的数值, 函数 $f(x)$ 必定能在区间内的某一点取到该值. 因此, 在 $[a,b]$ 内至少存在一点 ξ, 使得 $f(\xi)=\bar{f}$. □

此定理的几何意义极为基本. 对于任意连续曲线下的面积, 总能找到一个高度为 $f(\xi)$ 的矩形, 其宽度为区间长度 $(b-a)$, 使得这个矩形的面积与曲线下的面积相等.

17.6.5 推广的积分第一中值定理

积分第一中值定理中, 被积函数是单个函数 $f(x)$. 在应用中常常遇到两个函数的乘积 $f(x)g(x)$ 的积分. 当其中一个因子 g 不变号时, 中值定理有相应的推广形式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

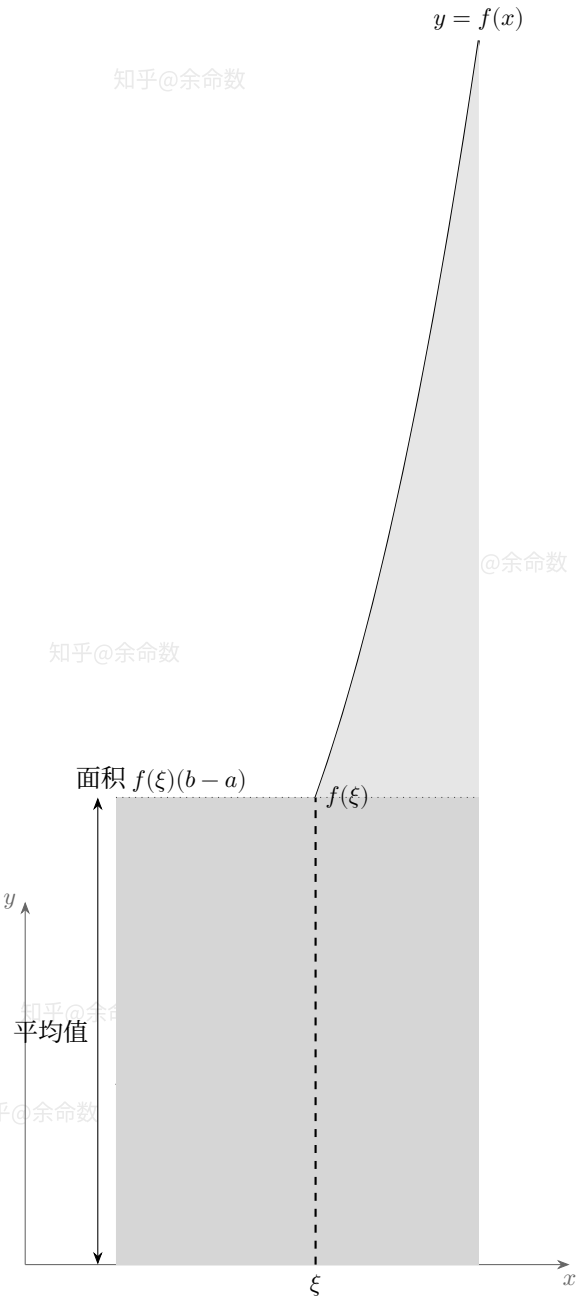


图 17.7 积分中值定理的几何诠释

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 17.22 (推广的积分第一中值定理) 设 f,g 均在 $[a,b]$ 上连续, 且 g 在 $[a,b]$ 上不变号, 即恒有 $g(x)\geq 0$ 或恒有 $g(x)\leq 0$. 则存在 $\xi\in[a,b]$, 使得

$$\int_a^b f(x)g(x)\,dx = f(\xi)\int_a^b g(x)\,dx.$$

证明 不妨设 $g(x)\geq 0$ 在 $[a,b]$ 上恒成立 (否则考虑 $-g$). 设 $m=\min_{[a,b]} f$, $M=\max_{[a,b]} f$, 由连续性极值定理知 m,M 存在. 由 $m\leq f(x)\leq M$ 与 $g(x)\geq 0$ 得

$$m\,g(x)\leq f(x)g(x)\leq M\,g(x).$$

保序性已在 第17.6.3小节 中证明过, 由它积分得

$$m\int_a^b g\leq \int_a^b fg\leq M\int_a^b g.$$

若 $\int_a^b g=0$, 则不等式两边均为 0, $\int_a^b fg=0$, 任意 ξ 都成立, 定理平凡. 否则

$$m\leq \frac{\int_a^b fg}{\int_a^b g}\leq M.$$

知乎@余命数　　知乎@余命数

由 f 在 $[a,b]$ 上连续, 介值定理保证存在 $\xi\in[a,b]$ 使 $f(\xi)=\frac{\int_a^b fg}{\int_a^b g}$, 即 $\int_a^b fg = f(\xi)\int_a^b g$. □

注 17.13
当 $g\equiv 1$ 时, 即退化为前文的(普通)积分中值定理 $\int_a^b f=f(\xi)(b-a)$. 推广形式允许处理带权函数 g 的平均值问题.

例题 17.25 估计 $\int_0^1 x\,e^{-x}\,dx$ 的范围.

解 在 $[0,1]$ 上 $e^{-x}>0$ 不变号, 视作权函数 $g(x)=e^{-x}$, $f(x)=x$. 由推广中值定理存在 $\xi\in[0,1]$, 使

$$\int_0^1 x\,e^{-x}\,dx=\xi\int_0^1 e^{-x}\,dx=\xi\cdot(1-e^{-1}).$$

知乎@余命数

因 $0 \le \xi \le 1$, 故

$$0 \le \int_0^1 x\,e^{-x}\,dx \le 1 - \frac{1}{e} \approx 0.6321.$$

(精确值用分部积分易得 $\int_0^1 xe^{-x}\,dx = 1 - 2/e \approx 0.2642$, 确实在估计区间内.)

17.6.6 对称区间与周期函数的积分性质

对于具有对称性或周期性的积分, 利用函数本身的结构可以大幅简化计算. 下列两条定理是利用结构简化积分的范例.

定理 17.23 (奇偶函数在对称区间上的积分) 设 f 在 $[-a, a]$ 上可积. 若 f 为奇函数 ($f(-x) = -f(x)$), 则 $\displaystyle\int_{-a}^{a} f(x)\,dx = 0$. 若 f 为偶函数 ($f(-x) = f(x)$), 则 $\displaystyle\int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx$.

证明 拆分 $\int_{-a}^{a} f = \int_{-a}^{0} f + \int_0^a f$. 在 $\int_{-a}^{0} f(x)\,dx$ 中作换元 $u = -x$, $du = -dx$, 当 $x = -a$ 时 $u = a$, $x = 0$ 时 $u = 0$, 故

$$\int_{-a}^{0} f(x)\,dx = -\int_a^0 f(-u)\,du = \int_0^a f(-u)\,du.$$

奇函数情形 $f(-u) = -f(u)$, 故 $\int_{-a}^{0} f = -\int_0^a f$, 总积分 $= 0$. 偶函数情形 $f(-u) = f(u)$, 故 $\int_{-a}^{0} f = \int_0^a f$, 总积分 $= 2\int_0^a f$. □

定理 17.24 (周期函数在一个周期上的积分) 设 f 以 $T > 0$ 为周期, 且在任何有限区间上可积. 则对任意 $a \in \mathbb{R}$,

$$\int_a^{a+T} f(x)\,dx = \int_0^T f(x)\,dx.$$

证明 假定 $0 \le a \le T$ (其余情况由区间可加性及周期性化归于此). 由区间可加性

$$\int_a^{a+T} f = \int_a^T f + \int_T^{a+T} f.$$

知乎@余命数

在第二项中作换元 $u=x-T$, $du=dx$, 当 $x=T$ 时 $u=0$, $x=a+T$ 时 $u=a$. 由周期性 $f(u+T)=f(u)$, 即

$$\int_T^{a+T} f(x)\,dx=\int_0^a f(u+T)\,du=\int_0^a f(u)\,du.$$

合并: $\int_a^{a+T} f=\int_a^T f+\int_0^a f=\int_0^T f$. □

例题 17.26 求 $\displaystyle\int_0^{2\pi}\sin^2 x\,dx$.

解 利用 $\sin^2 x=\frac{1-\cos 2x}{2}$. 因 $\cos 2x$ 的周期为 π, 故 $\sin^2 x$ 的最小周期也为 π. 由周期性定理(取 $a=0, T=\pi$, 注意 $2\pi=2T$),

$$\int_0^{2\pi}\sin^2 x\,dx=2\int_0^{\pi}\sin^2 x\,dx.$$

在 $[0,\pi]$ 上对称地, $\sin^2 x$ 关于中点 $\pi/2$ 对称, 故 $\int_0^{\pi}\sin^2 x\,dx=2\int_0^{\pi/2}\sin^2 x\,dx$. 应用将要给出的华里士公式 ($n=2$), 得 $\int_0^{\pi/2}\sin^2 x\,dx=\frac{1!!}{2!!}\cdot\frac{\pi}{2}=\frac{\pi}{4}$, 故

$$\int_0^{2\pi}\sin^2 x\,dx=2\cdot 2\cdot\frac{\pi}{4}=\pi.$$

也可直接积分验证, $\int_0^{2\pi}\frac{1-\cos 2x}{2}\,dx=\left[\frac{x}{2}-\frac{\sin 2x}{4}\right]_0^{2\pi}=\frac{2\pi}{2}-\frac{0-0}{4}=\pi$.

17.6.7 区间反射性质与特殊变换

对积分区间作线性代换 $u=a+b-x$, 可以把积分反射回原区间, 但被积函数 f 换成 $f(a+b-x)$. 这一定积分本身不变的性质(King 性质)是处理对称积分的有效工具.

定理 17.25 (King 性质 (区间反射)) 对 $[a,b]$ 上连续函数 f,

$$\int_a^b f(x)\,dx=\int_a^b f(a+b-x)\,dx.$$

证明 令 $u=a+b-x$, $du=-dx$. 当 $x=a$ 时 $u=b$, $x=b$ 时 $u=a$, 故

$$\int_a^b f(x)\,dx=-\int_b^a f(a+b-u)\,du=\int_a^b f(a+b-u)\,du.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

把积分变量名 u 改回 x 即得证. □

知乎@余命数

推论 17.3 (半周期消去) 若 f 在 $[0,\pi]$ 上连续, 则

$$\int_0^{\pi} x\, f(\sin x)\,dx = \frac{\pi}{2}\int_0^{\pi} f(\sin x)\,dx.$$

证明 记 $I=\int_0^{\pi} x\, f(\sin x)\,dx$, 应用 King 性质(取 $a=0,b=\pi$, 故 $a+b=\pi$). 作代换 $u=\pi-x$, 则 $\sin(\pi-u)=\sin u$, 故 $f(\sin(\pi-u))=f(\sin u)$.

$$I=\int_0^{\pi}(\pi-u)\,f(\sin u)\,du=\pi\int_0^{\pi} f(\sin u)\,du-\int_0^{\pi} u\,f(\sin u)\,du.$$

最后一项即 $-I$. 故 $I=\pi\int_0^{\pi} f(\sin u)\,du - I$, 即 $2I=\pi\int_0^{\pi} f(\sin u)\,du$, 解出

$$I=\frac{\pi}{2}\int_0^{\pi} f(\sin x)\,dx.$$

□

知乎@余命数 知乎@余命数

例题 17.27 求 $\displaystyle\int_0^{\pi} x\sin^3 x\,dx$.

命数 知乎@余命数

解 取 $f(t)=t^3$, 则 $f(\sin x)=\sin^3 x$. 由半周期消去推论得

$$\int_0^{\pi} x\sin^3 x\,dx=\frac{\pi}{2}\int_0^{\pi}\sin^3 x\,dx.$$

知乎@余命数

计算 $\int_0^{\pi}\sin^3 x\,dx=\int_0^{\pi}\sin x(1-\cos^2 x)\,dx$. 令 $u=\cos x$, $du=-\sin x\,dx$, 当 $x=0$ 时 $u=1$, $x=\pi$ 时 $u=-1$.

知乎@余命数

$$\int_0^{\pi}\sin^3 x\,dx=\int_1^{-1}(1-u^2)(-du)=\int_{-1}^{1}(1-u^2)\,du=2-\frac{2}{3}=\frac{4}{3}.$$

知乎@余命数

故 $\int_0^{\pi} x\sin^3 x\,dx=\frac{\pi}{2}\cdot\frac{4}{3}=\frac{2\pi}{3}$.

17.6.8 华里士公式 (Wallis' Formula)

已在 第17.4.2小节 的 $\sin^n x$ 递推公式中,掌握了把 $I_n=\int\sin^n x\,dx$ 降阶到 I_{n-2} 的方法. 当积分区间限定在 $[0,\pi/2]$ 时, 把同一递推重复应用, 最终归约到 I_0 或 I_1, 即可得到一个统一封闭的公式, 这就是华里士公式.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 17.26 (华里士公式) 对正整数 n,

$$\int_0^{\pi/2}\sin^n x\,dx=\int_0^{\pi/2}\cos^n x\,dx=\begin{cases}\dfrac{(n-1)!!}{n!!}\cdot\dfrac{\pi}{2}, & n\text{ 为偶数},\\[2mm] \dfrac{(n-1)!!}{n!!}, & n\text{ 为奇数}.\end{cases}$$

其中 $n!!=n(n-2)(n-4)\cdots$ 为双阶乘, 约定 $(-1)!!=0!!=1$.

证明 记 $I_n=\displaystyle\int_0^{\pi/2}\sin^n x\,dx$. 由 第17.4.2小节 的 sin 版递推应用于 $[0,\pi/2]$,

$$I_n=-\frac{\sin^{n-1}x\cos x}{n}\bigg|_0^{\pi/2}+\frac{n-1}{n}I_{n-2}=\frac{n-1}{n}I_{n-2},$$

知乎@余命数

端点项因 $\cos(\pi/2)=0$ 与 $\sin(0)=0$ 而消失. 起点 $I_0=\pi/2$, $I_1=1$. $n=2m$ 为偶数时, 重复递推至 I_0:

$$I_{2m}=\frac{2m-1}{2m}\cdot\frac{2m-3}{2m-2}\cdots\frac{1}{2}\cdot I_0=\frac{(2m-1)!!}{(2m)!!}\cdot\frac{\pi}{2}.$$

知乎@余命数

$n=2m+1$ 为奇数时, 重复递推至 I_1:

$$I_{2m+1}=\frac{2m}{2m+1}\cdot\frac{2m-2}{2m-1}\cdots\frac{2}{3}\cdot I_1=\frac{(2m)!!}{(2m+1)!!}.$$

知乎@余命数　知乎@余命数

$\cos^n$ 版由换元 $x\to\pi/2-x$ 化为 $\sin^n$ 版, 结果一致. □

例题 17.28 求 $\displaystyle\int_0^{\pi/2}\sin^6 x\,dx$.

知乎@余命数　知乎@余命数

解 $n=6$ 为偶数, 由公式得 $I_6=\frac{5!!}{6!!}\cdot\frac{\pi}{2}$. 计算双阶乘: $5!!=5\cdot3\cdot1=15$, $6!!=6\cdot4\cdot2=48$. 故

$$I_6=\frac{15}{48}\cdot\frac{\pi}{2}=\frac{5\pi}{32}.$$

作为验证, $\int_0^{\pi/2}\sin^7 x\,dx$($n=7$ 奇)应为 $\frac{6!!}{7!!}=\frac{48}{105}=\frac{16}{35}$, 对偶可用同公式.

17.6.9 经典积分不等式

最后用两条著名不等式结束本章.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 17.27 (柯西-施瓦茨积分不等式) 设 f,g 在 $[a,b]$ 上可积, 则

$$\left(\int_a^b f(x)g(x)\,dx\right)^2 \le \int_a^b f(x)^2\,dx \cdot \int_a^b g(x)^2\,dx.$$

等号成立当且仅当 f 与 g 线性相关, 即 $f=\lambda g$ 或 $g=\lambda f$. 在 g 不恒为零的连续情形下, 即存在常数 c 使 $f(x)=cg(x)$ 在 $[a,b]$ 上恒成立.

证明 对任意 $\lambda\in\mathbb{R}$, $\int_a^b (f(x)+\lambda g(x))^2 dx \ge 0$. 展开得关于 λ 的二次式

$$\left(\int_a^b g^2\right)\lambda^2 + 2\left(\int_a^b fg\right)\lambda + \int_a^b f^2 \ge 0.$$

该二次式非负, 故判别式 $\Delta = 4\left(\int_a^b fg\right)^2 - 4\left(\int_a^b f^2\right)\left(\int_a^b g^2\right) \le 0$, 即 $\left(\int_a^b fg\right)^2 \le \left(\int_a^b f^2\right)\left(\int_a^b g^2\right)$. 等号当 $\Delta=0$ 时成立, 即存在 λ_0 使 $\int_a^b (f+\lambda_0 g)^2 = 0$, 亦即 $f+\lambda_0 g = 0$ 几乎处处, 故 $f=cg$. □

例题 17.29 估计 $\left(\int_0^1 x\sqrt{1-x}\,dx\right)^2$ 的上界.

解 取 $f(x)=x$、$g(x)=\sqrt{1-x}$、$[a,b]=[0,1]$. 由柯西-施瓦茨不等式

$$\left(\int_0^1 x\sqrt{1-x}\,dx\right)^2 \le \int_0^1 x^2\,dx \cdot \int_0^1 (1-x)\,dx = \frac{1}{3}\cdot\frac{1}{2} = \frac{1}{6}.$$

故 $\left|\int_0^1 x\sqrt{1-x}\,dx\right| \le \frac{1}{\sqrt{6}} \approx 0.408$. (精确值可用换元 $x=1-t^2$ 或直接由 $\int_0^1 x\sqrt{1-x}\,dx = \int_0^1 (1-u)\sqrt{u}\,du = \int_0^1 (u^{1/2}-u^{3/2})\,du = \frac{2}{3}-\frac{2}{5} = \frac{4}{15} \approx 0.267$ 求得.)

定理 17.28 (切比雪夫积分不等式) 设 f,g 都在 $[a,b]$ 上可积且**同单调**, 即同时单调递增或同时单调递减. 记 $\bar f = \frac{1}{b-a}\int_a^b f$, $\bar g = \frac{1}{b-a}\int_a^b g$. 则

$$\frac{1}{b-a}\int_a^b f(x)g(x)\,dx \ge \bar f\cdot\bar g.$$

若 f,g **反单调**(一个单增一个单减), 则不等式反向.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 考虑二重积分 $\displaystyle\iint_{[a,b]\times[a,b]}(f(x)-f(y))(g(x)-g(y))\,dx\,dy$. 当 f,g 同单调时，对任意 $x,y\in[a,b]$，$(f(x)-f(y))(g(x)-g(y))\geq 0$，故二重积分非负. 展开被积函数，得

$$\begin{aligned}&\iint (f(x)g(x)-f(x)g(y)-f(y)g(x)+f(y)g(y))\,dx\,dy\\ =\;& 2(b-a)\int_a^b fg\,dx-2\left(\int_a^b f\,dx\right)\left(\int_a^b g\,dx\right)\geq 0.\end{aligned}$$

两端除以 $2(b-a)^2$ 得 $\dfrac{1}{b-a}\displaystyle\int_a^b fg\,dx\geq \bar f\cdot\bar g$. 反单调时 $(f(x)-f(y))(g(x)-g(y))\leq 0$，不等式反向. □

例题 17.30 验证 $\displaystyle\int_0^1 x\,e^x\,dx\geq\left(\int_0^1 x\,dx\right)\left(\int_0^1 e^x\,dx\right)$.

解 在 $[0,1]$ 上，$f(x)=x$ 与 $g(x)=e^x$ 都单调递增（后者因 e^x 严格单增），故由切比雪夫不等式

$$\int_0^1 xe^x\,dx\geq\left(\int_0^1 x\,dx\right)\left(\int_0^1 e^x\,dx\right)=\frac{1}{2}(e-1)\approx 0.8591.$$

精确值用分部积分计算，$\int_0^1 xe^x\,dx=\left[(x-1)e^x\right]_0^1=0-(-1)=1\geq 0.859$. 不等式成立.

17.7 定积分的几何与物理应用

定积分源于黎曼和的极限，它对一个连续变化的量在区间上做无限求和，得到总积累量. 前几节建立了积分的计算工具，包括换元法、分部法、微积分基本定理. 本节把这些工具应用于几何与物理中的经典积累问题：面积、体积、弧长、做功与平均值.

每一个应用都遵循同一个模式，把待求的总量分割成若干微元，每个微元用已知的公式算出近似值，再将所有微元求和并取极限，得到一个定积分. 这个方法称为**微元法**（或元素法）.

17.7.1 平面图形的面积

第17.5节 已经用黎曼和定义了曲边梯形的面积，连续函数 $f(x)\geq 0$ 在 $[a,b]$ 上与 x 轴围成区域的面积为 $\int_a^b f(x)\,dx$. 当 $f(x)$ 在某些区间取负值时，有向面积用积分的符号自然处理；而几何面积则需取绝对值.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

更一般的情形是两条曲线 $y=f(x)$ 与 $y=g(x)$ 之间的区域. 设 f,g 在 $[a,b]$ 上连续, 且在 $[a,b]$ 上 $f(x)\geq g(x)$, 则两曲线之间的区域面积为

$$A=\int_a^b [f(x)-g(x)]\,dx. \tag{17-9}$$

当 f 与 g 在 $[a,b]$ 上有交叉时, 需要找到所有交点, 在每个子区间上用上方的曲线减下方的曲线, 分段积分. 等价地, 统一写为

知乎@余命数

$$A=\int_a^b |f(x)-g(x)|\,dx. \tag{17-10}$$

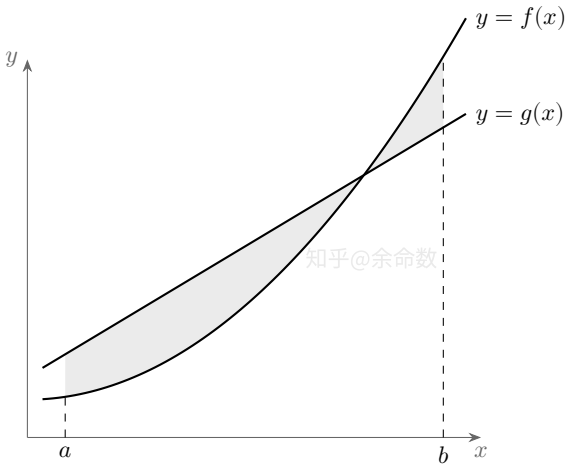


图 17.8 两曲线之间的区域面积

若曲线以参数方程 $\begin{cases} x=\varphi(t) \\ y=\psi(t) \end{cases}$ $(t\in[t_1,t_2])$ 给出, 且 φ 单调可导, 则可作换元 $x=\varphi(t)$, 面积公式化为

$$A=\int_{t_1}^{t_2} \psi(t)\,\varphi'(t)\,dt. \tag{17-11}$$

此即参数方程下的有向面积公式(详见 第14章 对参数方程的讨论).

例题 17.31 求抛物线 $y=x^2$ 与直线 $y=x$ 所围成区域的面积.

解 先求交点. 由 $x^2=x$ 得 $x=0$ 或 $x=1$. 在 $[0,1]$ 上 $x\geq x^2$, 故上方曲线为 $f(x)=x$, 下方曲线为 $g(x)=x^2$. 由 (17-9),

$$A=\int_0^1 (x-x^2)\,dx=\left[\frac{x^2}{2}-\frac{x^3}{3}\right]_0^1=\frac{1}{2}-\frac{1}{3}=\frac{1}{6}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 17.32　求椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\ (a,b>0)$ 的面积.

解　用参数方程 $x=a\cos t,\ y=b\sin t\ (t\in[0,2\pi])$. 由对称性，面积等于第一象限部分的四倍. 在第一象限 t 从 0 到 $\pi/2$，x 从 a 递减到 0. 由 (17-11)，注意 $\varphi'(t)=(a\cos t)'=-a\sin t$,

$$A=4\left|\int_0^{\pi/2} b\sin t\cdot(-a\sin t)\,dt\right|=4ab\int_0^{\pi/2}\sin^2 t\,dt.$$

知乎@余命数

由华里士公式 (第17.6节), $\int_0^{\pi/2}\sin^2 t\,dt=\frac{1}{2}\cdot\frac{\pi}{2}=\frac{\pi}{4}$. 故 $A=4ab\cdot\frac{\pi}{4}=\pi ab$.

17.7.2 极坐标下面积

在极坐标下, 曲线由 $r=r(\theta)$ 给出（详见 第14章）. 考虑由极坐标曲线 $r(\theta)$ 与两条射线 $\theta=\alpha,\theta=\beta\ (\alpha<\beta)$ 所围成的扇形区域.

将角度区间 $[\alpha,\beta]$ 分割成 n 个小扇形,第 i 个小扇形的圆心角为 $\Delta\theta_i$. 当 $\Delta\theta_i$ 很小时, 该小扇形近似为一个半径 $r(\theta_i)$、圆心角 $\Delta\theta_i$ 的圆扇形, 其面积为 $\frac{1}{2}r(\theta_i)^2\Delta\theta_i$. 对所有小扇形求和并取极限, 得到极坐标下的面积公式

$$A=\frac{1}{2}\int_\alpha^\beta r(\theta)^2\,d\theta. \tag{17-12}$$

若区域由内曲线 $r_1(\theta)$ 与外曲线 $r_2(\theta)$ $(r_2(\theta)\ge r_1(\theta)\ge 0)$ 及射线 $\theta=\alpha,\theta=\beta$ 围成, 则面积为两个扇形面积之差

$$A=\frac{1}{2}\int_\alpha^\beta\left[r_2(\theta)^2-r_1(\theta)^2\right]d\theta. \tag{17-13}$$

例题 17.33　求心形线 $r=a(1+\cos\theta)\ (a>0)$ 所围成区域的面积.

解　心形线关于极轴对称, 故面积等于上半平面部分 $(\theta\in[0,\pi])$ 的两倍. 由 (17-12),

知乎@余命数

$$A=2\cdot\frac{1}{2}\int_0^\pi a^2(1+\cos\theta)^2\,d\theta=a^2\int_0^\pi(1+2\cos\theta+\cos^2\theta)\,d\theta.$$

其中 $\int_0^\pi\cos^2\theta\,d\theta=\frac{\pi}{2}$（由半角公式或对称性可得）, $\int_0^\pi\cos\theta\,d\theta=[\sin\theta]_0^\pi=0$. 故

$$A=a^2(\pi+0+\tfrac{\pi}{2})=\tfrac{3}{2}\pi a^2.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

17.7.3 旋转体的体积

设连续函数 $f(x) \geq 0$ 在 $[a,b]$ 上定义, 曲线 $y = f(x)$ 与直线 $x = a$, $x = b$ 及 x 轴所围成的曲边梯形绕 x 轴旋转一周, 得到一个旋转体. 下面用微元法推导其体积.

在点 x 处取厚度为 Δx 的薄片, 当 Δx 很小时, 该薄片近似为一个半径 $f(x)$、高 Δx 的圆柱, 其体积为 $\pi f(x)^2 \Delta x$. 对所有薄片求和并取极限, 得到旋转体体积公式

$$V = \pi \int_a^b f(x)^2 \, dx. \tag{17-14}$$

若旋转体的截面为圆环, 例如两曲线之间的区域绕轴旋转, 则微元为圆环面积 $\pi[R(x)^2 - r(x)^2]$ 乘以厚度, 其中 $R(x)$ 为外半径、$r(x)$ 为内半径. 这给出圆环法, 也称垫圈法的体积公式

$$V = \pi \int_a^b \left[R(x)^2 - r(x)^2\right] dx. \tag{17-15}$$

当旋转轴为 y 轴时, 若曲线以 $y = f(x)$ 给出且 f 单调, 可作换元反解出 $x = g(y)$, 再对 y 积分; 也可直接对 x 积分, 此时微元为半径 x、高 $f(x)$ 的圆柱壳(柱壳法), 其体积微元为 $2\pi x f(x)\, dx$, 于是

$$V = 2\pi \int_a^b x f(x) \, dx. \tag{17-16}$$

三种方法的选择取决于被积函数的形式与计算便利, 圆盘法与圆环法沿旋转轴方向切片, 柱壳法则沿垂直于旋转轴的方向切片.

例题 17.34 求由曲线 $y = \sqrt{x}$ $(0 \leq x \leq 4)$ 与 x 轴所围区域绕 x 轴旋转所得旋转体的体积.

解 由 (17-14),

$$V = \pi \int_0^4 (\sqrt{x})^2 \, dx = \pi \int_0^4 x \, dx = \pi \left[\frac{x^2}{2}\right]_0^4 = \pi \cdot 8 = 8\pi.$$

17.7.4 平行截面面积为已知的立体的体积

旋转体是截面为圆盘的特殊情形. 更一般地, 设一个立体夹在垂直于 x 轴的两平面 $x = a$ 与 $x = b$ 之间, 且过点 x 处垂直于 x 轴的截面面积为已知的连续函

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

数 $A(x)$. 用微元法, 在点 x 处取厚度 Δx 的薄片, 其体积近似为 $A(x)\Delta x$, 求和取极限得

$$V=\int_a^b A(x)\,dx. \tag{17-17}$$

公式 (17-14) 是上式的特例, 当截面为半径 $f(x)$ 的圆盘时, $A(x)=\pi f(x)^2$.

例题 17.35 设一立体的底面是 xy 平面上由直线 $x=0$、$x=4$、$y=0$、$y=2$ 围成的矩形, 垂直于 x 轴的截面均为正方形(边长等于底面上该处的宽度). 求该立体的体积.

解 在点 x 处, 底面上的截面宽度恒为 2(沿 y 方向从 0 到 2), 故正方形截面的边长为 2, 面积 $A(x)=4$. 由 (17-17),

$$V=\int_0^4 4\,dx=16.$$

17.7.5 平面曲线的弧长

设曲线 $y=f(x)$ 在 $[a,b]$ 上有连续导数, 即 $f\in C^1[a,b]$. 将区间 $[a,b]$ 分成 n 段, 对应曲线上的分点连成折线, 弧长定义为折线总长当分割无限加密时的极限.

每小段折线的长度为 $\sqrt{\Delta x_i^2+\Delta y_i^2}$. 由拉格朗日中值定理, $\Delta y_i=f'(\xi_i)\Delta x_i$, 故折线总长为 $\sum\sqrt{1+f'(\xi_i)^2}\,\Delta x_i$, 取极限得直角坐标下的弧长公式

$$L=\int_a^b \sqrt{1+f'(x)^2}\,dx. \tag{17-18}$$

若曲线以参数方程 $\begin{cases}x=\varphi(t)\\ y=\psi(t)\end{cases}$ $(t\in[\alpha,\beta])$ 给出, 且 φ,ψ 在 $[\alpha,\beta]$ 上有连续导数, 则弧长微元为 $\sqrt{\varphi'(t)^2+\psi'(t)^2}\,dt$, 于是

$$L=\int_\alpha^\beta \sqrt{\varphi'(t)^2+\psi'(t)^2}\,dt. \tag{17-19}$$

若曲线以极坐标 $r=r(\theta)$ $(\theta\in[\alpha,\beta])$ 给出且 r 可导, 利用极坐标与直角坐标的关系 $x=r(\theta)\cos\theta$, $y=r(\theta)\sin\theta$, 代入参数式并化简得极坐标下的弧长公式

$$L=\int_\alpha^\beta \sqrt{r(\theta)^2+r'(\theta)^2}\,d\theta. \tag{17-20}$$

三个公式的弧长微元相同, 均为 $\sqrt{\dot{x}^2+\dot{y}^2}$ 乘以参数微分, 只是在不同坐标系下表达形式不同.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 17.36 求抛物线 $y=\frac{x^2}{2}$ 在 $0\le x\le 1$ 上的一段弧的长度.

解 由 $y'=x$, 代入 (17-18),

$$L=\int_0^1 \sqrt{1+x^2}\,dx.$$

用换元法令 $x=\tan t$, 则 $\sqrt{1+x^2}=\sec t$, $dx=\sec^2 t\,dt$. 当 $x=0$ 时 $t=0$, 当 $x=1$ 时 $t=\pi/4$. 故

$$L=\int_0^{\pi/4}\sec^3 t\,dt.$$

由分部积分, 或查积分表, $\int_0^{\pi/4}\sec^3 t\,dt=\frac{1}{2}\left[\sec t\tan t+\ln|\sec t+\tan t|\right]_0^{\pi/4}$. 代入端值: $\sec\frac{\pi}{4}=\sqrt{2}$, $\tan\frac{\pi}{4}=1$, 故

$$L=\tfrac{1}{2}\left[\sqrt{2}+\ln(1+\sqrt{2})-(0+0)\right]=\tfrac{\sqrt{2}+\ln(1+\sqrt{2})}{2}.$$

17.7.6 变力做功

物理学中, 恒力 F 沿力的方向位移 s 所做的功为 $W=Fs$. 当力的大小随位置变化时, 需要用积分处理.

设物体沿 x 轴从 a 运动到 b, 所受的力为位置的连续函数 $F(x)$ (方向沿 x 轴). 将区间 $[a,b]$ 分割, 在每一小段 $[x_i,x_{i+1}]$ 上力近似为常数 $F(x_i)$, 该段做功近似为 $F(x_i)\Delta x_i$. 求和取极限得变力做功公式

$$W=\int_a^b F(x)\,dx. \tag{17-21}$$

抽取液体做功是另一类典型应用. 设一垂直放置的容器内盛有密度 ρ 的液体, 液面距容器底部高为某定值, 要将液体全部抽出容器顶部 (高出容器底部 H), 需计算克服重力所做的功. 在高度 y 处取厚度 Δy 的薄层, 其体积为截面积 $A(y)\Delta y$, 质量为 $\rho A(y)\Delta y$, 该层液体被提升的距离为 $H-y$, 故做功微元为 $\rho g A(y)(H-y)\Delta y$, 总功为

$$W=\rho g\int_{y_1}^{y_2} A(y)(H-y)\,dy, \tag{17-22}$$

其中 $[y_1,y_2]$ 为液体占据的高度区间.

例题 17.37 一弹簧的自然长度为 $1\,\mathrm{m}$, 用力拉长时弹力满足胡克定律 $F=kx$, 其中 $k=20\,\mathrm{N/m}$, x 为伸长量. 求将弹簧从自然长度拉长 $0.5\,\mathrm{m}$ 所做的功.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 以伸长量 x 为积分变量, 范围 $[0, 0.5]$. 由 (17-21),

$$W = \int_0^{0.5} 20x\,dx = 10\big[x^2\big]_0^{0.5} = 10 \times 0.25 = 2.5\,\mathrm{J}.$$

17.7.7 函数的平均值

n 个离散数 $y_1, \dots, y_n$ 的算术平均值为 $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. 连续函数 $f(x)$ 在区间 $[a, b]$ 上有无穷多个取值, 其平均值用积分的极限来定义.

将 $[a, b]$ 等分为 n 段, 每段长 $\Delta x = \frac{b-a}{n}$, 取分点 $x_i = a + i\Delta x$, 则离散平均值为 $\frac{1}{n}\sum_{i=1}^{n} f(x_i)$. 注意到 $\frac{1}{n} = \frac{\Delta x}{b-a}$, 令 $n \to \infty$ 取极限得连续函数的平均值

$$\bar{f} = \frac{1}{b-a}\int_a^b f(x)\,dx. \tag{17-23}$$

该公式有明确的几何意义, 以 $[a, b]$ 为底、$\bar{f}$ 为高的矩形面积等于曲线 $y = f(x)$ 下方的曲边梯形面积.

关于平均值的一个重要结论是积分中值定理, 若 f 在 $[a, b]$ 上连续, 则存在 $\xi \in [a, b]$ 使得

$$f(\xi) = \frac{1}{b-a}\int_a^b f(x)\,dx. \tag{17-24}$$

即连续函数在某区间上必能取到其平均值. 该定理的证明依赖闭区间上连续函数的介值定理.

例题 17.38 求正弦交流电 $i(t) = I_m \sin \omega t$ 在半个周期 $[0, \pi/\omega]$ 上的电流平均值与有效值.

解 由 (17-23), 平均值为

$$\bar{i} = \frac{\omega}{\pi}\int_0^{\pi/\omega} I_m \sin(\omega t)\,dt = \frac{\omega I_m}{\pi}\left[-\frac{\cos(\omega t)}{\omega}\right]_0^{\pi/\omega} = \frac{\omega I_m}{\pi}\cdot\frac{2}{\omega} = \frac{2I_m}{\pi}.$$

有效值(方均根值)定义为 $I_{\mathrm{rms}} = \sqrt{\overline{i^2}}$, 即

$$I_{\mathrm{rms}} = \sqrt{\frac{\omega}{\pi}\int_0^{\pi/\omega} I_m^2 \sin^2(\omega t)\,dt}.$$

由半角公式 $\sin^2(\omega t) = \frac{1-\cos(2\omega t)}{2}$,

$$\int_0^{\pi/\omega} \sin^2(\omega t)\,dt = \frac{1}{2}\int_0^{\pi/\omega}\big[1 - \cos(2\omega t)\big]\,dt = \frac{1}{2}\left[t - \frac{\sin(2\omega t)}{2\omega}\right]_0^{\pi/\omega} = \frac{\pi}{2\omega}.$$

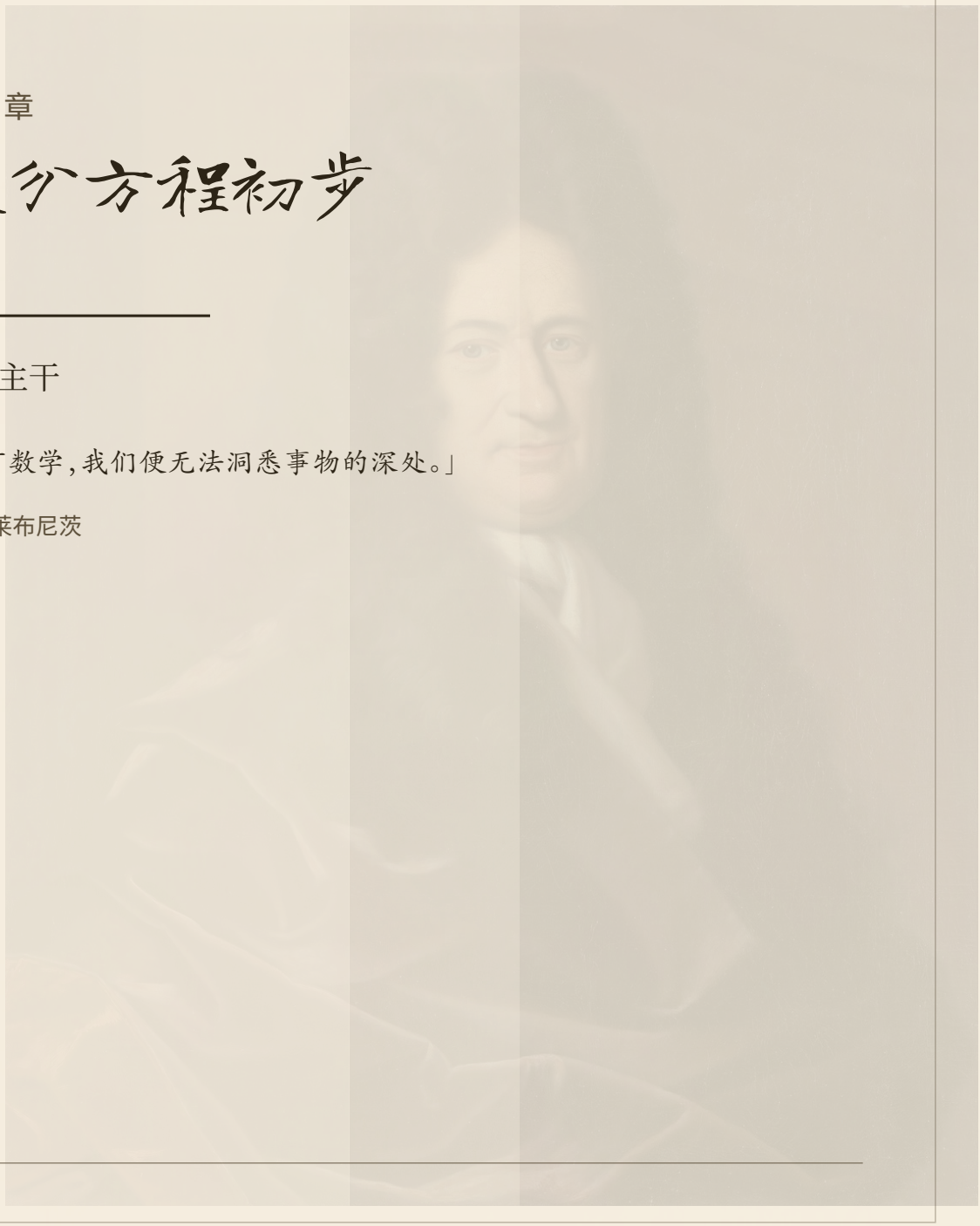
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $I_{\text{rms}}=I_m\sqrt{\frac{\omega}{\pi}\cdot\frac{\pi}{2\omega}}=\frac{I_m}{\sqrt{2}}$. 这正是正弦交流电有效值等于峰值除以 $\sqrt{2}$ 的由来.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 18 章

微分方程初步



分析主干

「没有数学，我们便无法洞悉事物的深处。」

——莱布尼茨

第十八章 微分方程初步

知乎@余命数

18.1 微分方程、阶与解

我们研究两类对象, 不变的量和变化的量. 代数方程(如 $x^2-1=0$)的解是一些孤立的数, 它们描述静止的状态. 但自然科学与工程技术中的许多问题关心的是变化过程. 要刻画这些过程, 我们需要把未知对象从**数**转向**函数**, 即用一个函数描述整个过程的演变.

微分学告诉我们, 一个函数在区间上的整体变化由它在每一点附近的变化率累积而成. 系统在某一瞬间的变化率就是它的导数. 在实验里, 局部的变化法则往往比全局规律更容易测量. 例如, 物体的冷却速率在每一时刻都与它和环境的温差成正比. 这个观察把温度函数 $T(t)$ 与它的导数 $T'(t)$ 直接联系起来.

我们把未知函数与其导数之间的关系写成等式, 就得到**微分方程**. 它要求未知函数在定义域中的每一点都满足自身与变化率之间的特定关系.

方向场与积分曲线的几何图像

一阶微分方程最常见的形式是

$$\frac{\mathrm{d}y}{\mathrm{d}x}=f(x,y).$$

左端是未知曲线 $y=y(x)$ 在横坐标 x 处的切线斜率, 右端只依赖于点的位置 (x,y). 于是方程给出明确的指令. 平面上每一点都规定了曲线在该点允许的切线方向.

我们在平面上取任意一点 $P=(x_0,y_0)$. 若 f 在该点有定义, 则任何经过 P 的解曲线在 P 处的切线斜率都必须等于 $f(x_0,y_0)$. 解曲线在 P 处只能沿这个斜率穿过, 不能取别的倾角.

我们把同一要求加到平面上许多点. 在每一格点画一小段短线, 方向由 $f(x,y)$ 决定. 这些短线铺满平面, 构成**方向场**, 也称**斜率场**. 方向场在每一点都标出该朝哪个方向前进, 就像一张只标风向的地图.

在这张地图上手描一条曲线, 就是从某一位置出发, 始终顺着每一点的短线方向前进, 使曲线在每一点的切线都与该点的短线方向重合. 这样画出的曲线称为方程的**积分曲线**. 我们也可以把方向场想象成河面的流速方向, 把积分曲线想象成小船的航迹; 航迹在每一点都顺着水流, 才能处处与场相切.

图 18.1 以方程 $\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{1}{2}(x-y)$ 为例. 图中每个小箭头的倾角由右端 $f(x,y)=\frac{1}{2}(x-y)$ 决定. 在区域 $x>y$ 内斜率为正, 曲线向右上方走; 在区域 $x<y$ 内斜率为负, 曲线向右下方走; 靠近直线 $y=x$ 时斜率接近 0, 场几乎水平. 图中的几条连续曲线都是积分曲线, 它们与箭头处处相切, 不会横穿场的方向.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

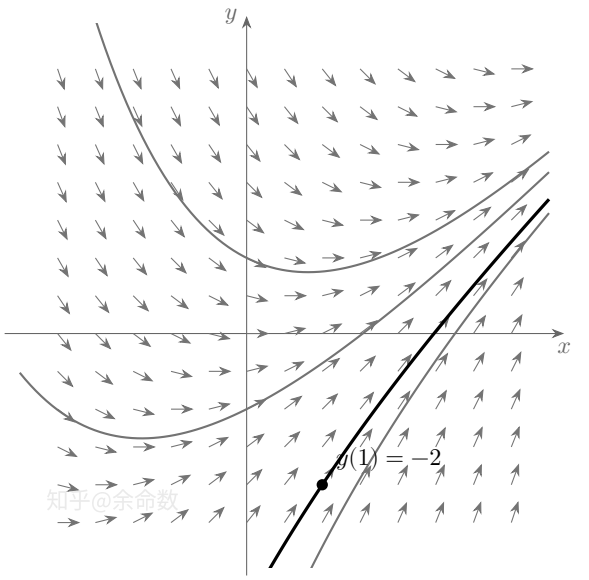


图 18.1　微分方程 $\dfrac{\mathrm{d}y}{\mathrm{d}x} = \dfrac{1}{2}(x-y)$ 的方向场与积分曲线. 灰色短箭头在每一点给出斜率 $f(x,y)=\frac{1}{2}(x-y)$, 构成方向场. 灰色与黑色曲线是积分曲线, 在每一点都与场方向相切. 加粗的黑色曲线经过点 $(1,-2)$, 由初始条件 $y(1)=-2$ 唯一确定.

从不同起点出发, 可以画出无穷多条积分曲线. 它们共同构成方程的**通解**, 是一族填满方向场的曲线. 若额外指定曲线必须经过某一点 (x_0,y_0), 即给定**初始条件** $y(x_0)=y_0$, 则通常只有一条积分曲线满足要求, 称为**特解**. 图中加粗的曲线经过 $(1,-2)$, 就是由 $y(1)=-2$ 挑出的那一条.

几何图像与后文的代数写法互相印证. 一阶方程的通解通常含有一个任意常数, 这个常数对应在方向场中选取哪一条积分曲线. 一个初值问题由微分方程和曲线过定点这两个条件组成, 在方向场里通常只有一条解. 场在每一点只给出一个方向, 从定点沿场走下去, 航迹便被唯一确定. 严格的存在唯一性定理需要用分析语言证明, 方向场给出同一事实的直观版本.

我们用正式语言把上面的对象整理成定义.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 18.1 (微分方程及其相关概念)

- **微分方程**. 建立自变量 x、未知函数 $y=y(x)$ 以及其直至 n 阶的导数 $y', y'', \dots, y^{(n)}$ 之间关系的方程, 一般形式为 $F(x, y, y', y'', \dots, y^{(n)})=0$.
- **阶**. 方程中出现的导数的最高阶数 n.
- **解**. 若将函数 $y=\phi(x)$ 及其各阶导数代入微分方程后得到关于 x 的恒等式, 则称 $y=\phi(x)$ 是该方程的一个解.
- **通解**. 含有 n 个相互独立任意常数的 n 阶方程的解, 代表满足方程约束的整个函数族.
- **初始条件**. 为确定通解中任意常数而给定的附加条件,通常形如 $y(x_0)=y_0$、$y'(x_0)=y_1$、$\dots$、$y^{(n-1)}(x_0)=y_{n-1}$.
- **特解**. 在初始条件约束下, 由通解确定全部任意常数后得到的解.

18.2 可分离变量的微分方程

我们从结构最简单的一类一阶微分方程开始研究. 设一阶微分方程 $\frac{dy}{dx}=f(x,y)$ 的右端可以分解为一个只与 x 有关的函数 $g(x)$ 和一个只与 y 有关的函数 $h(y)$ 的乘积, 即

$$\frac{dy}{dx}=g(x)h(y)$$

我们就称它为**可分离变量的微分方程**. 这种方程的特征是, 自变量和未知函数对变化率的影响可以分开.

求解这类方程的方法来自积分学. 设 $h(y)\neq 0$, 我们先对原方程作代数变形, 把含 y 的项放到等式左端, 把含 x 的项放到右端.

$$\frac{1}{h(y)}\frac{dy}{dx}=g(x)$$

这个等式断言, 对于解函数 $y=y(x)$, 复合函数 $\frac{1}{h(y(x))}\frac{dy}{dx}$ 与函数 $g(x)$ 恒等. 两个相等的函数关于 x 的不定积分也相等(至多差一个常数). 于是我们对等式两端关于 x 积分.

$$\int \frac{1}{h(y(x))}\frac{dy}{dx}\,dx=\int g(x)\,dx$$

观察左端的积分. 令 $u=y(x)$, 则 $du=\frac{dy}{dx}dx$. 根据换元法, 左端的积分可以化为对 y 的积分. 于是

$$\int \frac{1}{h(y)}\,dy=\int g(x)\,dx$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这一步也说明, 把 $\frac{dy}{dx}$ 看作分数并交叉相乘, 得到 $\frac{dy}{h(y)} = g(x)\,dx$. 这个看似不严格的符号操作是正确的. 求出上述两个积分后, 我们就得到一个联系 x 与 y 的代数方程, 它通常以隐函数形式给出原微分方程的通解.

例题 18.1　求解微分方程 $\frac{dy}{dx} = \frac{x^2}{y}$.

解　方程的右端是 $g(x) = x^2$ 与 $h(y) = \frac{1}{y}$ 的乘积, 因此它是可分离变量方程. 以下在 $y \neq 0$ 的区间上求解. 两端乘以 y 后对 x 积分, 左端用换元 $u = y(x)$ 化为对 y 的积分. 依次演算得

$$\begin{aligned}
\int y\,\frac{dy}{dx}\,dx &= \int x^2\,dx \quad \text{（两端乘 } y \text{ 后对 } x \text{ 积分）}\\
\int y\,dy &= \int x^2\,dx \quad \text{（左端令 } u = y(x) \text{ 换元）}\\
\tfrac{1}{2}y^2 &= \tfrac{1}{3}x^3 + C_1 \quad \text{（求不定积分, } C_1 \text{ 为任意常数）}\\
3y^2 &= 2x^3 + C \quad \text{（两端乘 6, 令 } C = 6C_1\text{）.}
\end{aligned}$$

$3y^2 - 2x^3 = C$ 就是方程在不穿过 $y = 0$ 的区间上的通解, 不同的常数 C 对应平面上一族积分曲线. 当积分曲线碰到 $y = 0$ 时, 原方程无定义, 因此解只能取在避开这些点的局部区间上.

该方程的通解为 $3y^2 - 2x^3 = C$.

定理 18.1 (指数增长与衰减模型)　描述一个量的变化率正比于其自身大小的微分方程

$$\frac{dy}{dx} = ky \quad (k \neq 0 \text{ 为常数})$$

其通解为

$$y = Ce^{kx}$$

其中 C 为任意常数.

证明　方程 $\frac{dy}{dx} = ky$ 描述种群增长、放射性衰变等动态过程. 为了不讨论 $y = 0$ 处除以 y 的问题, 我们把它改写成乘积导数的形式. 设 y 是任意可导解, 则

$$\frac{d}{dx}\left(e^{-kx}y(x)\right) = e^{-kx}y'(x) - ke^{-kx}y(x) = e^{-kx}\left(y'(x) - ky(x)\right) = 0.$$

因此 $e^{-kx}y(x)$ 在所讨论的区间上为常数, 记为 C. 于是

$$y(x) = Ce^{kx}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

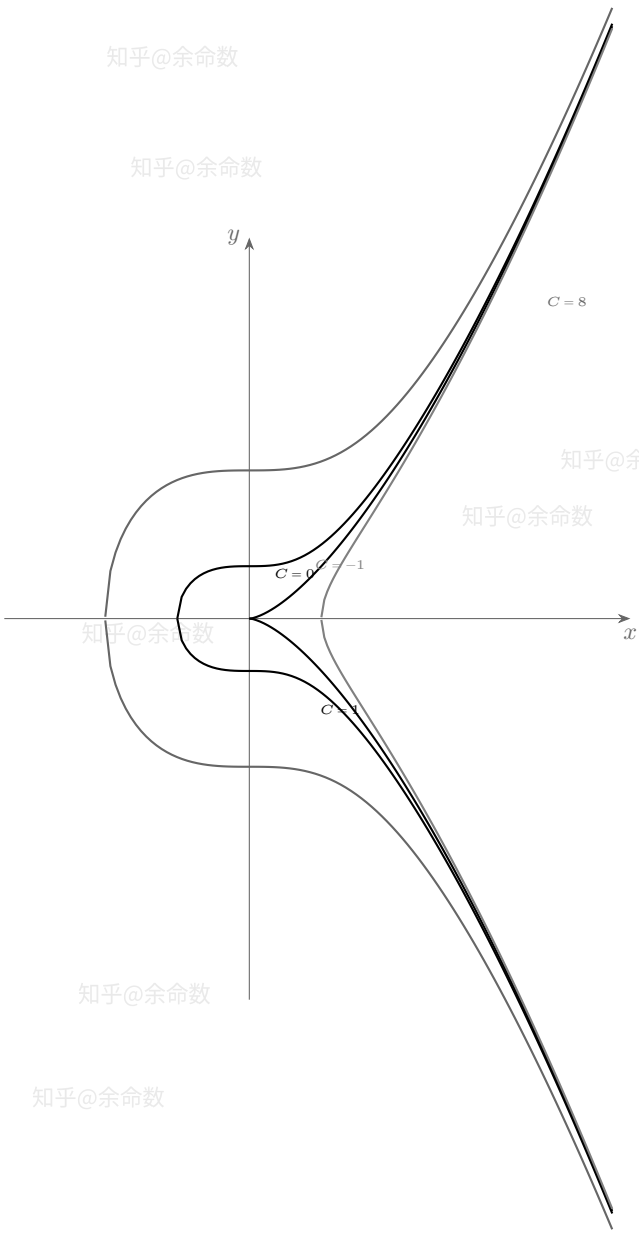


图 18.2 微分方程 $\frac{dy}{dx} = \frac{x^2}{y}$ 的积分曲线族 $3y^2 - 2x^3 = C$.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

反过来, 对任意常数 C, 函数 $y = Ce^{kx}$ 满足

$$y' = kCe^{kx} = ky,$$

因而都是方程的解. 当 $C = 0$ 时即包含平凡解 $y \equiv 0$.

方程 $\frac{dy}{dx} = ky$ 的通解为 $y = Ce^{kx}$, 其中 C 是任意实数. □

18.3 一阶线性微分方程

另一类重要的一阶微分方程是线性方程. 下面给出它的标准形式.

定义 18.2 (一阶线性微分方程) 形如

$$\frac{dy}{dx} + P(x)y = Q(x)$$

的方程, 被称为**一阶线性微分方程**.

我们引入一个辅助函数**积分因子**来求解这类方程. 具体做法是, 在方程两边同乘某个函数 $\mu(x)$, 使新方程的左边正好是一个乘积的导数.

考虑标准形式 $y' + P(x)y = Q(x)$. 两边同乘 $\mu(x)$, 得到 $\mu(x)y' + \mu(x)P(x)y = \mu(x)Q(x)$. 我们希望左端能够写成 $(\mu(x)y)'$. 由乘法求导法则, $(\mu(x)y)' = \mu'(x)y + \mu(x)y'$. 比较两式, 只需 $\mu'(x) = \mu(x)P(x)$ 即可.

方程 $\mu'(x) = \mu(x)P(x)$ 正是上一节讨论的指数增长模型, 它的解是 $\mu(x) = e^{\int P(x)\,dx}$. 这个 $\mu(x)$ 就是我们要找的积分因子.

例题 18.2 求解微分方程 $y' - 2xy = e^{x^2}$.

解 方程已经具有标准形 $y' + P(x)y = Q(x)$, 其中 $P(x) = -2x$, $Q(x) = e^{x^2}$. 按积分因子法, 取 $\mu(x) = e^{\int P(x)\,dx} = e^{-x^2}$, 使左端凑成 $(\mu y)'$. 两端乘 $\mu(x)$ 后依次演算.

$$\begin{aligned}
e^{-x^2}y' - 2xe^{-x^2}y &= 1 && \text{(两端乘 } e^{-x^2}\text{, 右端 } e^{-x^2}e^{x^2} = 1\text{)}\\
(e^{-x^2}y)' &= 1 && \text{(左端恰为乘积导数)}\\
e^{-x^2}y &= x + C && \text{(两端对 } x \text{ 积分)}\\
y &= (x + C)e^{x^2} && \text{(两端除以 } e^{-x^2}\text{)}.
\end{aligned}$$

方程的通解为 $y = (x + C)e^{x^2}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

阅读提示

可分离变量方程与一阶线性方程的求解方法, 是用微分方程处理函数方程的工具. 尤其是方程 $y' = ky$ 与它的解 $y = Ce^{kx}$, 在 第18.4节 中会反复出现.

知乎@余命数

18.4 应用微分方程求解函数方程

18.4.1 从代数恒等式到分析微分方程

在前面的章节里, 我们已经学过赋值法、模型法等处理抽象函数方程的代数工具. 这些工具对特定结构的方程很有效. 但遇到形式更新或更复杂的方程时, 纯代数变换可能不够.

函数方程是在定义域上处处成立的**恒等式**. 微积分的一个基本事实是, 若两个函数在某区间上恒等且都可导, 那么它们在该区间上的导函数也恒等.

据此我们可以对函数方程这个恒等式两边求导, 得到关于函数导数的信息. 这个过程常常把一个含多个变量、关系复杂的函数方程, 化成一个只含单个变量、结构清晰的**微分方程**. 一旦转化成功, 我们就可以用微积分的整套工具求出原函数的精确表达式. 这是处理某些函数方程的有效路径.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

18.4.2 偏导数

在应用这个策略之前, 我们先需要一个预备工具. 函数方程通常含有多个变量, 例如 $f(x+y)$. 我们熟悉的导数定义针对单变量函数, 对一个多变量表达式求导则需要**偏导数**.

它的想法是直观而直接的. 要对其中一个变量, 例如 x 求导, 就把其余变量, 例如 y 暂时**看作常数**, 然后按单变量求导法则运算.

定义 18.3 (偏导数的操作性定义) 对于一个含有变量 x 和 y 的表达式 $F(x,y)$, 其对 x 的偏导数(记作 $\frac{\partial}{\partial x}F(x,y)$)的计算方式是把 y 看作一个常数, 然后对 x 进行常规的求导.

在函数方程的应用里, 最常见的操作是对复合函数求偏导. 这需要链式法则. 我们举例计算 $f(x+y)$ 对 x 的偏导数. 设 $u = x+y$. 由链式法则,

$$\frac{\partial}{\partial x} f(x+y) = \frac{df}{du} \cdot \frac{\partial u}{\partial x} = f'(u) \cdot \frac{\partial}{\partial x}(x+y)$$

对 x 求偏导时, y 是常数, 因此 $\frac{\partial}{\partial x}(x+y) = 1+0 = 1$. 于是

知乎@余命数

$$\frac{\partial}{\partial x} f(x+y) = f'(x+y)$$

同理, 对于乘法形式 $f(xy)$, 对 x 求偏导时, y 是常数, 则

$$\frac{\partial}{\partial x}f(xy)=f'(xy)\cdot\frac{\partial}{\partial x}(xy)=f'(xy)\cdot y$$

18.4.3 求解的四步流程

利用偏导数这个工具, 我们可以把函数方程问题转成微分方程问题.

例题 18.3 (再探柯西指数方程) 若可导函数 $f(x)$ 满足 $f(x+y)=f(x)f(y)$, 且 $f(x)$ 不恒为零, 求 $f(x)$.

解 方程 $f(x+y)=f(x)f(y)$ 关于 x,y 恒等. 固定 y, 把 $f(y)$ 看作常数, 两端对 x 求偏导. 由链式法则得到

$$\begin{aligned}\frac{\partial}{\partial x}f(x+y)&=\frac{\partial}{\partial x}[f(x)f(y)]\\ f'(x+y)&=f'(x)f(y)\quad(f(y)\text{ 视为常数}).\end{aligned}$$

上式仍含两个变量. 令 $x=0$ 消去 x, 得到 $f'(y)=f'(0)f(y)$. 记 $k=f'(0)$, 把变量改记为 x, 就得到一阶方程 $f'(x)=kf(x)$. 这正是指数增长模型, 它的通解是 $f(x)=Ce^{kx}$.

为确定 C, 我们把通解代回原方程检验.

$$\begin{aligned}f(x+y)&=Ce^{k(x+y)}=Ce^{kx}e^{ky},\\ f(x)f(y)&=(Ce^{kx})(Ce^{ky})=C^2e^{kx}e^{ky}.\end{aligned}$$

两端相等当且仅当 $C=C^2$, 即 $C=0$ 或 $C=1$. 若 $C=0$, 则 $f(x)\equiv0$, 与题设中 f 不恒为零矛盾, 因此 $C=1$. 令 $a=e^k$ (a 由 $f'(0)$ 决定), 我们便得到非零可导解.

因此, 唯一的非零可导解是 $f(x)=e^{kx}=(e^k)^x=a^x$.

例题 18.4 设可导函数 $f(x)$ 对任意 $x,y>0$ 满足 $f(xy)=f(x)+f(y)$, 求 $f(x)$.

解 方程 $f(xy)=f(x)+f(y)$ 对所有 $x,y>0$ 成立. 固定 y 并把它看作常数, 两端对 x 求偏导. 由链式法则得

$$\begin{aligned}\frac{\partial}{\partial x}f(xy)&=\frac{\partial}{\partial x}[f(x)+f(y)]\\ yf'(xy)&=f'(x)\quad(y\text{ 视为常数}, f(y)\text{ 对 }x\text{ 导数为 }0).\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为了消去 xy 中的复合变量, 令 $x=1$, 得到 $yf'(y)=f'(1)$. 记 $k=f'(1)$, 把变量改记为 x, 就有 $xf'(x)=k$, 也就是 $f'(x)=\frac{k}{x}$. 两端积分(定义域 $x>0$, 不用绝对值), 得到通解 $f(x)=k\ln x+C$.

把通解代回原方程检验.

$$f(xy)=k\ln(xy)+C=k\ln x+k\ln y+C,$$

$$f(x)+f(y)=(k\ln x+C)+(k\ln y+C)=k\ln x+k\ln y+2C.$$

两端相等要求 $C=2C$, 即 $C=0$. 若 $k\neq 0$, 令 $a=e^{1/k}$, 则 $k\ln x=\frac{\ln x}{\ln a}=\log_a x$; 若 $k=0$, 则解退化为零函数.

故该方程的可导解为 $f(x)=k\ln x$; 其中 $k\neq 0$ 时也可写作 $f(x)=\log_a x$, $k=0$ 时为 $f(x)\equiv 0$.

例题 18.5 (达朗贝尔方程) 若二阶可导函数 $f(x)$ 满足 $f(x+y)+f(x-y)=2f(x)f(y)$, 且 f 不恒为 0. 求 $f(x)$.

解 我们对方程 $f(x+y)+f(x-y)=2f(x)f(y)$ 两边求偏导并分离变量来求解. 先令 $x=y=0$, 得到 $2f(0)=2f(0)^2$, 于是 $f(0)\in\{0,1\}$. 若 $f(0)=0$, 再令 $y=0$, 得 $2f(x)=2f(x)f(0)=0$, 即 $f(x)\equiv 0$. 因此非平凡解必有 $f(0)=1$.

对原方程分别关于 x、y 求偏导, 得

$$f'(x+y)+f'(x-y)=2f'(x)f(y),\quad (\text{对 } x \text{ 求偏导})$$

$$f'(x+y)-f'(x-y)=2f(x)f'(y).\quad (\text{对 } y \text{ 求偏导})$$

对前一式再关于 x 求偏导、后一式再关于 y 求偏导, 两式左端同为 $f''(x+y)+f''(x-y)$, 故右端相等, 得 $2f''(x)f(y)=2f(x)f''(y)$. 令 $y=0$(非平凡情形 $f(0)=1$)并记 $k=f''(0)$, 即得二阶常系数齐次方程 $f''(x)=kf(x)$.

按 k 的符号分为三种情形.

- $k=0$. 方程 $f''(x)=0$ 的通解为 $f(x)=Ax+B$, 由 $f(0)=1$ 得 $B=1$. 在原方程中令 $x=0$ 得 $f(y)+f(-y)=2f(0)f(y)=2f(y)$, 即 f 为偶函数, 故 $A=0$, 解为 $f(x)=1$.
- $k>0$. 令 $k=\omega^2$ ($\omega>0$), 特征方程 $r^2-\omega^2=0$ 给出 $r=\pm\omega$, 通解 $f(x)=C_1e^{\omega x}+C_2e^{-\omega x}$. 偶性要求 $f'(0)=0$, 联立

$$f(0)=C_1+C_2=1,$$

$$f'(0)=\omega(C_1-C_2)=0,$$

解得 $C_1=C_2=\frac{1}{2}$, 故 $f(x)=\frac{e^{\omega x}+e^{-\omega x}}{2}=\cosh(\omega x)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- $k<0$. 令 $k=-\omega^2$ ($\omega>0$), 特征方程 $r^2+\omega^2=0$ 给出 $r=\pm i\omega$, 通解 $f(x)=C_1\cos(\omega x)+C_2\sin(\omega x)$. 同样由 $f(0)=1$ 与 $f'(0)=0$ 得

$$f(0)=C_1=1,$$
$$f'(0)=\omega C_2=0,$$

即 $C_1=1, C_2=0$, 故 $f(x)=\cos(\omega x)$.

该方程的所有非平凡二阶可导解为 $f(x)=1$, $f(x)=\cos(\omega x)$, 或 $f(x)=\cosh(\omega x)$.

例题 18.6 (正切加法模型) 若可导函数 $f(x)$ 在包含 0 的区间内满足 $f(x+y)=\frac{f(x)+f(y)}{1-f(x)f(y)}$, 其中 $x, y, x+y$ 均在该区间内, 且分母不为零, 求 $f(x)$.

解 方程的结构与正切和角公式一致. 先确定 $f(0)$. 令 $x=y=0$, 得 $f(0)=\frac{2f(0)}{1-f(0)^2}$, 即 $f(0)(1+f(0)^2)=0$. 在实数域内唯一解为 $f(0)=0$.

把方程两边对 x 求偏导 ($f(y)$ 视为常数). 左端是 $f'(x+y)$; 右端按商的求导法则展开分子并化简.

$$\frac{\partial}{\partial x}\left(\frac{f(x)+f(y)}{1-f(x)f(y)}\right)=\frac{f'(x)(1-f(x)f(y))-(f(x)+f(y))(-f'(x)f(y))}{(1-f(x)f(y))^2}$$
$$=\frac{f'(x)(1+f(y)^2)}{(1-f(x)f(y))^2}\quad\text{(分子合并同类项)}.$$

故 $f'(x+y)=\dfrac{f'(x)(1+f(y)^2)}{(1-f(x)f(y))^2}$. 由原方程关于 x, y 对称, 对 y 求偏导得对称式 $f'(x+y)=\dfrac{f'(y)(1+f(x)^2)}{(1-f(x)f(y))^2}$. 比较两式, 得 $f'(x)(1+f(y)^2)=f'(y)(1+f(x)^2)$, 分离变量即

$$\frac{f'(x)}{1+f(x)^2}=\frac{f'(y)}{1+f(y)^2}.$$

左端仅依赖 x, 右端仅依赖 y, 故两端同为常数 ω, 得可分离变量方程 $\dfrac{df}{1+f^2}=\omega\,dx$. 两端积分得

$$\arctan(f(x))=\omega x+C\quad\left(\int\frac{df}{1+f^2}=\arctan f\right),$$

即 $f(x)=\tan(\omega x+C)$. 代入 $f(0)=0$ (取 arctan 主值) 得 $C=0$; 若用多值反正切, 整数倍 π 由正切周期性吸收, 表达式不变.

故该方程在不跨越正切极点的区间内的可导解为 $f(x)=\tan(\omega x)$. 若要求在整个 $\mathbb{R}$ 上处处有定义并满足原式, 则只能取 $\omega=0$, 即 $f(x)\equiv 0$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

18.5 增长与衰减模型

第18.2节 证明了方程 $\frac{dy}{dx} = ky$ 的通解是 $y = Ce^{kx}$. 这个简单的方程出现在大量互不相干的领域里:银行存款、细菌繁殖、放射性衰变、物体冷却、溶液混合. 它们的共同点是, 变化率与现存量成正比. 接着我们就把这些模型一个一个建立起来.

18.5.1 连续复利

设本金为 P, 年利率为 $100k\%$, 如 $k = 0.06$ 对应 6%, 一年计息 n 次, 每次利率 $\frac{k}{n}$. 由 第18.1节 的复利知识, t 年后的本利和为

$$A = P\left(1 + \frac{k}{n}\right)^{nt}. \tag{18-1}$$

计息周期越短, 结算越频繁. 让 $n \to \infty$, 利用 $\left(1 + \frac{k}{n}\right)^{n/k} \to e$(第24.1节 的极限), 得到

$$A = Pe^{kt}. \tag{18-2}$$

这就是**连续复利**公式. 对 式 (18-2) 求导, $\frac{dA}{dt} = kA$, 本利和恰好满足增长率正比于现存量的微分方程.k 的含义也随之清晰, 它是单位时间内的相对增长率.

例题 18.7 本金 10000 元, 年利率 6%, 存期一年. 比较一年计息一次与连续复利的本利和.

解 一年计息一次, 本利和为 $10000 \times 1.06 = 10600$ 元. 连续复利为

$$10000\, e^{0.06} \approx 10000 \times 1.061837 = 10618.37 \text{ 元}.$$

连续复利多出约 18 元. 计息频率的收益有上限, 这个上限正是 e 设定的.

一年计息一次得 10600 元, 连续复利得约 10618.37 元.

18.5.2 人口增长与倍增期

设 $t = 0$ 时刻细菌数为 x_0, 营养与环境不受限制, 任一时刻的繁殖率与当时的细菌总数成正比. 记 t 时刻总数为 $x(t)$, 则

$$\frac{dx}{dt} = kx, \qquad x(0) = x_0,$$

解出

$$x(t) = x_0 e^{kt}. \tag{18-3}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

数量翻倍所需的时间称为**倍增期**. 在 式 (18-3) 中令 $x(T)=2x_0$, 得 $e^{kT}=2$, 故

$$T=\frac{\ln 2}{k}. \tag{18-4}$$

倍增期只依赖增长率 k, 与初值无关.

例题 18.8　某地区人口以每年 2% 的速率净增长. 按指数增长模型, 人口翻一番需要多少年?

解　此时 $k=0.02=\frac{1}{50}$, 由 式 (18-4),

$$T=50\ln 2\approx 50\times 0.6931=34.7 \text{ 年}.$$

约 34.7 年.

倍增期公式还顺手解释了一条粗估法则, 年增长率百分之 r 的量, 倍增大约要 $\dfrac{70}{r}$ 年. 拿 $r=2$ 验证, $\dfrac{70}{2}=35$, 与 $50\ln 2$ 相差无几, 原因是 $\ln 2\approx 0.693\approx 0.7$.

18.5.3 放射性衰变与半衰期

放射性元素的原子核随机衰变. 单个原子核何时衰变无法预测, 但大量原子核合起来服从一条统计规律, 单位时间内发生衰变的原子核数与现存原子核数成正比. 设 t 时刻现存质量为 $x(t)$, 则

$$-\frac{dx}{dt}=kx\quad(k>0),\qquad \text{即}\qquad x(t)=x_0e^{-kt}.$$

负号表示质量在减少. 现存量减半所需的时间称为**半衰期** $T_{1/2}$. 令 $x(T_{1/2})=\frac{x_0}{2}$, 得

$$k\,T_{1/2}=\ln 2. \tag{18-5}$$

半衰期与倍增期是同一个公式的两面, 前者刻画衰减到一半, 后者刻画增长到两倍.

例题 18.9 (碳十四测年)　碳十四 (^{14}C) 的半衰期约为 5730 年. 活体生物与环境交换碳, 体内 ^{14}C 的比例与环境一致; 死亡后交换停止, 体内的 ^{14}C 按半衰期衰减. 某古墓木制品的 ^{14}C 放射性只有同类活体材料的 25%, 求它的年代.

解　放射性强度正比于 ^{14}C 的现存量. 设制作年代为 t 年前, 则

$$e^{-kt}=0.25=\frac{1}{4}.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

由 式 (18-5), $k=\frac{\ln 2}{5730}$, 故

$$t=\frac{\ln 4}{k}=\frac{2\ln 2}{\ln 2}\times 5730=2\times 5730=11460 \text{ 年}.$$

该木制品距今约 11460 年.

注 18.1
碳十四测年由化学家 Willard Libby 在 1940 年代末建立, 他因此获得 1960 年诺贝尔化学奖. 半衰期 5730 年恰好落在几千年的尺度上, 适合考古测年; 更久远的年代则改用铀铅法(半衰期约 44.7 亿年)等慢时钟.

18.5.4 Newton 冷却定律

热学实验给出另一条正比规律, 物体温度向环境温度收敛的速率与二者温差成正比. 设物体温度为 $T(t)$, 环境温度为常数 T_s, 则

$$\frac{dT}{dt}=-k(T-T_s),\qquad k>0. \tag{18-6}$$

温差 $T-T_s$ 大时散热快, 温差小则散热慢, 收敛的极限是环境温度. 令 $\theta(t)=T(t)-T_s$, 方程化为 第18.2节 的衰减模型 $\theta'=-k\theta$, 解为

$$T(t)=T_s+(T(0)-T_s)e^{-kt}.$$

例题 18.10 一杯 100 °C 的热水放在 20 °C 的房间里, 30 分钟后降至 50 °C. 再过 30 分钟, 水温是多少?

解 记温差 $\theta(t)=T(t)-20$, 初值 $\theta(0)=80$. 由衰减解 $\theta(t)=80e^{-kt}$. $t=30$ 时 $\theta=30$, 即 $80e^{-30k}=30$, $e^{-30k}=\frac{3}{8}$. 于是

$$T(60)=20+80\,e^{-60k}=20+80\,(e^{-30k})^2=20+80\times\frac{9}{64}=31.25\,^\circ\text{C}.$$

再过 30 分钟水温为 31.25 °C.

18.5.5 混合问题

一池液体, 外部以固定流量注入含盐浓度为常数 c_{in} 的盐水, 池内搅拌均匀后以相同流量流出, 池内水量保持不变. 设池内盐量为 $x(t)$, 水量为 V, 流量为 r, 则

$$\frac{dx}{dt}=\underbrace{r\,c_{\text{in}}}_{\text{注入速率}}-\underbrace{r\cdot\frac{x}{V}}_{\text{流出速率}}.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这是一个一阶线性方程 (第18.3节), 也可分离变量求解.

例题 18.11　池内有 100 L 清水.$t=0$ 起注入浓度为 2 g/L 的盐水, 流量为 3 L/min, 同时以相同流量流出搅拌均匀的混合液. 问池内盐量何时达到 100 g? 长时间后盐量趋于多少?

知乎@余命数

解 注入速率为 $3\times 2=6$ g/min, 流出速率为 $3\cdot\frac{x}{100}$ g/min, 故

$$\frac{dx}{dt}=6-\frac{3x}{100}=\frac{3}{100}(200-x),\qquad x(0)=0.$$

分离变量:

$$\frac{dx}{200-x}=\frac{3}{100}\,dt\implies -\ln(200-x)=\frac{3}{100}t+C.$$

代入 $x(0)=0$ 得 $C=-\ln 200$, 于是

$$x(t)=200\left(1-e^{-3t/100}\right).$$

知乎@余命数

令 $x=100$ 得 $e^{-3t/100}=\frac{1}{2}$, $t=\frac{100\ln 2}{3}\approx 23.1$ 分钟.$t\to\infty$ 时 $x\to 200$ g, 与直觉一致, 盐量最终应等于注入浓度乘以池内容积, 即 2 g/L × 100 L.

约 23.1 分钟后盐量达 100 g, 长期盐量趋于 200 g.

练习 18.1　某放射性同位素经 8 天后剩余 30%. 求它的半衰期.

练习 18.2　一杯饮料在冰箱中从 25 °C 降到 10 °C 用了 40 分钟, 冰箱内温度恒为 4 °C. 它降到 6 °C 还需多久?

知乎@余命数
知乎@余命数
知乎@余命数

18.6 落体与极限速度

18.6.1 自由落体

取竖直向下为正方向, y 为下落距离. 无空气阻力时, 落体以恒定加速度 $g\approx 9.8$ m/s^2 下落, 运动满足二阶方程

$$\frac{d^2y}{dt^2}=g.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

令 $v=\dfrac{dy}{dt}$，方程变为 $\dfrac{dv}{dt}=g$，积分得

$$v=gt+v_0,\qquad y=\frac{1}{2}gt^2+v_0t+y_0,$$

其中 v_0,y_0 为初始速度与初始位置. 从静止开始下落时 $v=gt,y=\frac{1}{2}gt^2$，消去 t 得

$$v=\sqrt{2gy}. \tag{18-7}$$

式 (18-7) 也可由能量守恒导出，动能的增加等于重力势能的减少，$\frac{1}{2}mv^2=mgy$. 两条路径都归结为同一个关系 $v^2=2gy$.

18.6.2 空气阻力与极限速度

真实的下落伴随空气阻力. 低速情形下，阻力的大小近似与速度成正比、方向与运动相反. 设阻力为 $-cv(c>0$，由物体形状与介质决定)，速度方程为

$$\frac{dv}{dt}=g-cv. \tag{18-8}$$

这是 第18.5节 混合问题中出现过的方程，常数项 g 持续加速，比例项 $-cv$ 持续减速. 分离变量求解：

$$\frac{dv}{g-cv}=dt\implies -\frac{1}{c}\ln(g-cv)=t+C_1\implies g-cv=C_2e^{-ct}.$$

代入初值 $v(0)=0$ 得 $C_2=g$，故

$$v(t)=\frac{g}{c}\left(1-e^{-ct}\right). \tag{18-9}$$

当 $t\to\infty$，$v\to\dfrac{g}{c}$. 速度不会无限增大，而是单调逼近上界 $\dfrac{g}{c}$，称为**极限速度**. 物理图像很直白. 速度小时阻力也小，净加速接近 g；速度增大后阻力跟上来，把净加速一点点吃掉，等到 $cv=g$ 时加速归零，速度就稳定在 $\dfrac{g}{c}$.

例题 18.12 跳伞员张开降落伞后，每千克质量受到的空气阻力约为 $2v$(单位 N，速度 v 以 m/s 计)，即 $c=2\ \mathrm{s}^{-1}$. 求张伞后的极限速度，以及达到极限速度 95% 所需的时间.

解 极限速度为

$$\frac{g}{c}=\frac{9.8}{2}=4.9\ \mathrm{m/s}.$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

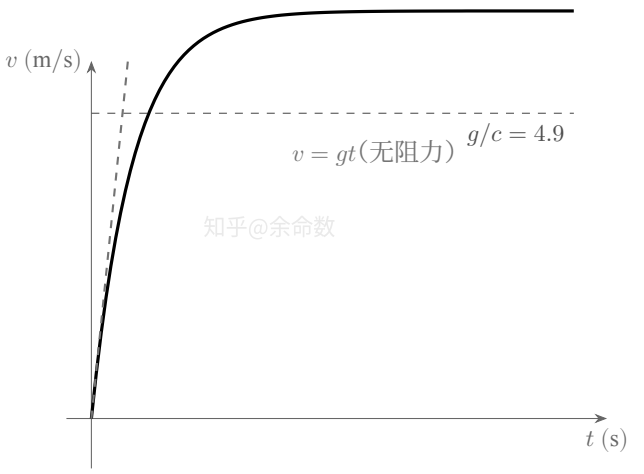


图 18.3　受阻落体的速度曲线 $v(t)=\frac{g}{c}(1-e^{-ct})$(实线, $c=2$), 单调逼近极限速度 g/c; 虚线为无阻力时的 $v=gt$, 短时间内二者几乎重合.

张伞后 $v(t)=4.9\,(1-e^{-2t})$. 达到 95% 即 $e^{-2t}=0.05$, 故

$$t=\frac{\ln 20}{2}\approx 1.5\ \text{s}.$$

两秒不到, 跳伞员便以近似恒定的速度下降.4.9 m/s 相当于从约 1.2 米高处跳下的落地速度, 这正是降落伞救命的原因.

极限速度 4.9 m/s, 约 1.5 秒后达到其 95%.

> **注 18.2**
> 阻力与速度成正比只在低速下成立. 速度较高时, 阻力更接近与 v^2 成正比, 方程 $\frac{dv}{dt}=g-cv^2$ 依然可分离变量, 解为 $v=\sqrt{\frac{g}{c}}\tanh\left(\sqrt{gc}\,t\right)$, 极限速度为 $\sqrt{g/c}$. 处理手法与 第18.2节 一致, 积分时用到 $\int\frac{dv}{1-v^2}=\operatorname{artanh} v$.

18.6.3　单摆的小角度摆动 *(选读)

一根长为 a 的轻杆, 下端固定质量为 m 的小球, 构成单摆. 设摆角为 $\theta(t)$, 重力沿圆弧切线方向的分量为 $-mg\sin\theta$, 由牛顿第二定律

$$a\,\frac{d^2\theta}{dt^2}=-g\sin\theta. \tag{18-10}$$

这是非线性方程. 当摆角很小时, 第6.10节 的倍角关系给出 $\sin\theta\approx\theta$(弧度制), 方程近似为

$$\frac{d^2\theta}{dt^2}=-\frac{g}{a}\,\theta.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第18.4节 求解 d'Alembert 方程时，$k<0$ 的情形正是 $f''(x)=-\omega^2 f(x)$，通解为 $C_1\cos\omega t+C_2\sin\omega t$. 取 $\omega=\sqrt{g/a}$, 设初始摆角为 α、静止释放, 得

$$\theta(t)=\alpha\cos\left(\sqrt{\frac{g}{a}}\,t\right).$$

知乎@余命数

摆角按余弦规律振荡, 周期

$$T=2\pi\sqrt{\frac{a}{g}}, \tag{18-11}$$

在小角度近似下与振幅 α 无关. 这个等时性是摆钟的理论基础, 伽利略据传由教堂吊灯的摆动发现了它, 惠更斯则据此造出了第一座摆钟. 大角度下 式 (18-10) 的周期依赖振幅, 摆幅越大周期越长, 摆钟工程师称之为**圆误差**; 其精确周期由一个不能表示为初等函数的积分(椭圆积分)给出.

> **练习 18.3** 要把单摆的周期从 2 秒(秒摆)调整为 3 秒, 摆长应变为原来的多少倍?

知乎@余命数

18.7 悬链线与追逐曲线 *(选读)

18.7.1 降阶法

形如 $y''=f(x,y')$ 的二阶方程不含 y 本身. 令 $p=y'$, 则 $y''=\dfrac{dp}{dx}$, 方程降为关于 p 的一阶方程

$$\frac{dp}{dx}=f(x,p).$$

解出 $p(x)$ 后再积分一次便得 y. 前面两节的落体问题(令 $v=y'$)用的正是这一手法, 现在拿它来处理两个经典的几何问题.

知乎@余命数

18.7.2 悬链线

一条密度均匀的软链悬挂于两点, 静止时下垂成曲线. 伽利略猜测它是抛物线, 莱布尼茨、惠更斯与约翰·伯努利在 1690 年前后各自证明它另有其形. 这条曲线称为**悬链线**(catenary, 源自拉丁语 catena, 链).

知乎@余命数

取 y 轴过链的最低点. 考察最低点与任意点 (x,y) 之间的一段链, 把它想成被三个力拽住而保持平衡的一小段, 一端是最低点处的水平张力 T_0, 另一端是 (x,y) 处的张力 T (链柔软, 张力只能沿切线方向), 中间还有这段链自身的重力. 设线

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

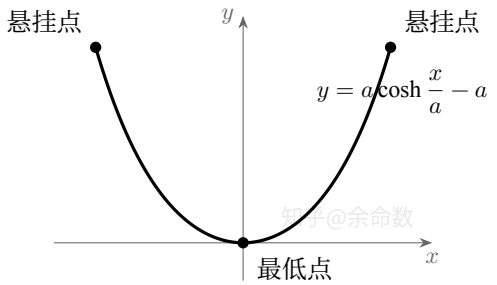


图 18.4　悬链线 $y = a\cosh(x/a) - a$: 坐标原点取在最低点, 纵轴过最低点竖直向上.

密度为常数 w_0, 这段弧长为 s, 重力便是 $w_0 s$. 水平与竖直方向的合力都为零, 写出来是

$$T\cos\theta = T_0, \qquad T\sin\theta = w_0 s,$$

其中 θ 为切线倾角, $\tan\theta = y'$. 两式相除, 再对 x 求导消去 s, 其中用到 $\frac{ds}{dx} = \sqrt{1+(y')^2}$, 见 第17.7节:

$$T_0 y' = w_0 s \implies T_0 y'' = w_0 \frac{ds}{dx} = w_0\sqrt{1+(y')^2}.$$

记 $a = \dfrac{T_0}{w_0}$, 其中 a 为张力与线密度之比, 量纲为长度, 得微分方程

$$y'' = \frac{1}{a}\sqrt{1+(y')^2}. \tag{18-12}$$

令 $p = y'$, 降阶后分离变量:

$$\frac{dp}{\sqrt{1+p^2}} = \frac{1}{a}\,dx \implies \ln\left(p + \sqrt{1+p^2}\right) = \frac{x}{a} + C.$$

最低点处切线水平, $p(0) = 0$, 故 $C = 0$. 现在从 $p + \sqrt{1+p^2} = e^{x/a}$ 反解 p. 由恒等式 $(p+\sqrt{1+p^2})(\sqrt{1+p^2}-p) = 1$ 得 $\sqrt{1+p^2} - p = e^{-x/a}$, 两式相减除以 2,

$$p = \frac{1}{2}\left(e^{x/a} - e^{-x/a}\right) = \sinh\frac{x}{a},$$

其中 $\sinh u = \frac{1}{2}(e^u - e^{-u})$ 是 第18.4节 中与 cosh 一同出现的双曲正弦. 再积分一次, 取最低点为原点:

$$y = a\cosh\frac{x}{a} - a. \tag{18-13}$$

悬链线由双曲余弦给出, 与抛物线只在小范围内近似重合, 靠近最低点处 $\cosh u \approx 1 + \frac{u^2}{2}$, 悬链线近似于抛物线 $y = \frac{x^2}{2a}$; 离最低点越远, 二者差距越大. 第18.4节 求解 d'Alembert 方程时出现的 cosh, 在这里找到了几何原型.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

18.7.3 追逐曲线

兔子从原点出发沿 y 轴以速度 a 逃跑, 狗从 $(c,0)$ 出发以速度 b 追击, 方向始终指向兔子. 狗的轨迹称为**追逐曲线**.

设时刻 t 兔在 $R=(0,at)$, 狗在 $D=(x,y)$. 追击方向恒指向兔子, 即 DR 是轨迹的切线:

$$\frac{dy}{dx}=\frac{y-at}{x}\implies xy'-y=-at. \tag{18-14}$$

对 x 求导消去 t:

$$xy''=-a\frac{dt}{dx}.$$

狗以速度 b 沿轨迹奔跑, $\frac{ds}{dt}=b$; x 随时间减少, $\frac{ds}{dx}=-\sqrt{1+(y')^2}$, 故

$$\frac{dt}{dx}=\frac{dt}{ds}\cdot\frac{ds}{dx}=-\frac{1}{b}\sqrt{1+(y')^2},\qquad xy''=\frac{a}{b}\sqrt{1+(y')^2}.$$

记 $k=\dfrac{a}{b}$, 再用降阶法. 令 $p=y'$:

$$\frac{dp}{\sqrt{1+p^2}}=k\,\frac{dx}{x}\implies \ln\left(p+\sqrt{1+p^2}\right)=k\ln x+C.$$

初值 $p(c)=0$ 给出 $C=-k\ln c$, 于是

$$p+\sqrt{1+p^2}=\left(\frac{x}{c}\right)^k\implies p=\frac{1}{2}\left[\left(\frac{x}{c}\right)^k-\left(\frac{c}{x}\right)^k\right].$$

$$\frac{dy}{dx}=\frac{1}{2}\left[\left(\frac{x}{c}\right)^k-\left(\frac{c}{x}\right)^k\right],\qquad k=\frac{a}{b}. \tag{18-15}$$

例题 18.13 (狗能否追上兔子) 按式 (18-15) 判断: $b>a(k<1)$ 与 $b\le a(k\ge 1)$ 两种情形各有什么结局?

解 沿曲线积分式 (18-15) 得狗在横坐标 x 处的高度

$$y(x)=\int_c^x p\,du=\frac{1}{2}\int_x^c\left[\left(\frac{c}{u}\right)^k-\left(\frac{u}{c}\right)^k\right]du.$$

右端在 $x\to 0^+$ 处的行为由 $\int_x^c (c/u)^k\,du$ 决定. $k<1$ 时积分收敛, 曲线在有限高度到达 y 轴, 狗在有限时间内追上兔子. $k\ge 1$ 时积分发散, 狗的高度随 $x\to 0$ 无限增长, 轨迹贴着 y 轴向上延伸, 追击不会结束.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等速情形 $k=1$ 的结局可以精确算出. 此时 $p=\frac{1}{2}\left(\frac{x}{c}-\frac{c}{x}\right)$, 积分得

$$y=\frac{x^2}{4c}-\frac{c}{2}\ln\frac{x}{c}.$$

由 式 (18-14), 兔子此刻的高度为 $t=y-xy'$(取 $a=b=1$), 于是二者的纵向距离

$$t-y=-xp=\frac{c}{2}-\frac{x^2}{2c}\longrightarrow\frac{c}{2}\qquad(x\to0^+).$$

狗与兔同速时, 间距恰好稳定在初始横向距离的一半, 越追越近, 近到 $c/2$ 为止.

$b>a$ 时狗在有限时间内追上; 等速时间距趋于 $\dfrac{c}{2}$; $b<a$ 时追击不结束.

注 18.3

同一条技巧链(受力分析或几何约束给出方程, 降阶化为可分离变量, 积分出曲线)还处理过另一个古典问题. 沿 y 轴拖动长度为 a 的细绳, 绳端物体在平面上留下的轨迹称为**曳物线** (tractrix), 方程为 $y=a\ln\dfrac{a+\sqrt{a^2-x^2}}{x}-\sqrt{a^2-x^2}$. 曳物线绕 y 轴旋转得到的曲面具有常数负曲率, 是罗巴切夫斯基非欧几何的一个经典模型, 称为伪球面.

18.8 章末习题

习题

18.1. (分离变量) 解方程 $\dfrac{dy}{dx}=\dfrac{y^2}{x^2}\cdot\dfrac{y+1}{x+1}$ 的通解. 提示: 两端同乘 $\dfrac{dx}{y^2}$ 后, 左端可写成 $\dfrac{dy}{(y+1)y^2}$ 之类的分式, 部分分式分解后再积分; 也可先取倒数 $\dfrac{dx}{dy}$.

18.2. (积分因子) 解初值问题 $y'+y\tan x=\sec x,\ y(0)=0$. 提示: 积分因子取 $\mu(x)=e^{\int\tan x\,dx}=\dfrac{1}{\cos x}$, 两端乘开后左端恰为 $\left(\dfrac{y}{\cos x}\right)'$.

18.3. (复利) 年利率 r 按半年复利与按连续复利计算, 一年后本利和之比为 $\left(1+\frac{r}{2}\right)^2:e^r$. 当 $r=0.12$ 时, 连续复利比半年复利多百分之几? 提示: 计算 $e^{0.12}$ 与 1.06^2, 差额再除以半年复利的结果.

18.4. (倍增期) 某培养基中的细菌每小时增长 18%. 按指数模型, 数量增至十倍需要多少小时? 提示: 解 $e^{0.18t}=10$, 即 $t=\dfrac{\ln 10}{0.18}$.

18.5. (半衰期) 钚-239 的半衰期约为 24100 年. 一枚样品中钚-239 已衰变掉 90%, 它经过了多少年? 提示: 剩余 10%, 即 $e^{-kt}=0.1$, 结合 $kT_{1/2}=\ln 2$.

18.6. (冷却) 傍晚 6 时测得路面温度为 32 °C, 气温恒为 22 °C. 已知路面温度按 Newton 冷却定律下降, 且每小时温差缩减为原来的 80%. 午夜时路面温度是多少? 提示: 温差 $\theta(t)=\theta_0\cdot0.8^t$, 六小时后温差乘以 0.8^6.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

18.7. (混合)池内有 200 L 浓度 1 g/L 的盐水. 以 4 L/min 注入清水, 同时以 4 L/min 流出混合液. 盐量何时降到 50 g? 提示: $\dfrac{dx}{dt}=-\dfrac{4x}{200}$, 盐量按 $e^{-t/50}$ 衰减, 初值 $x_0=200$.

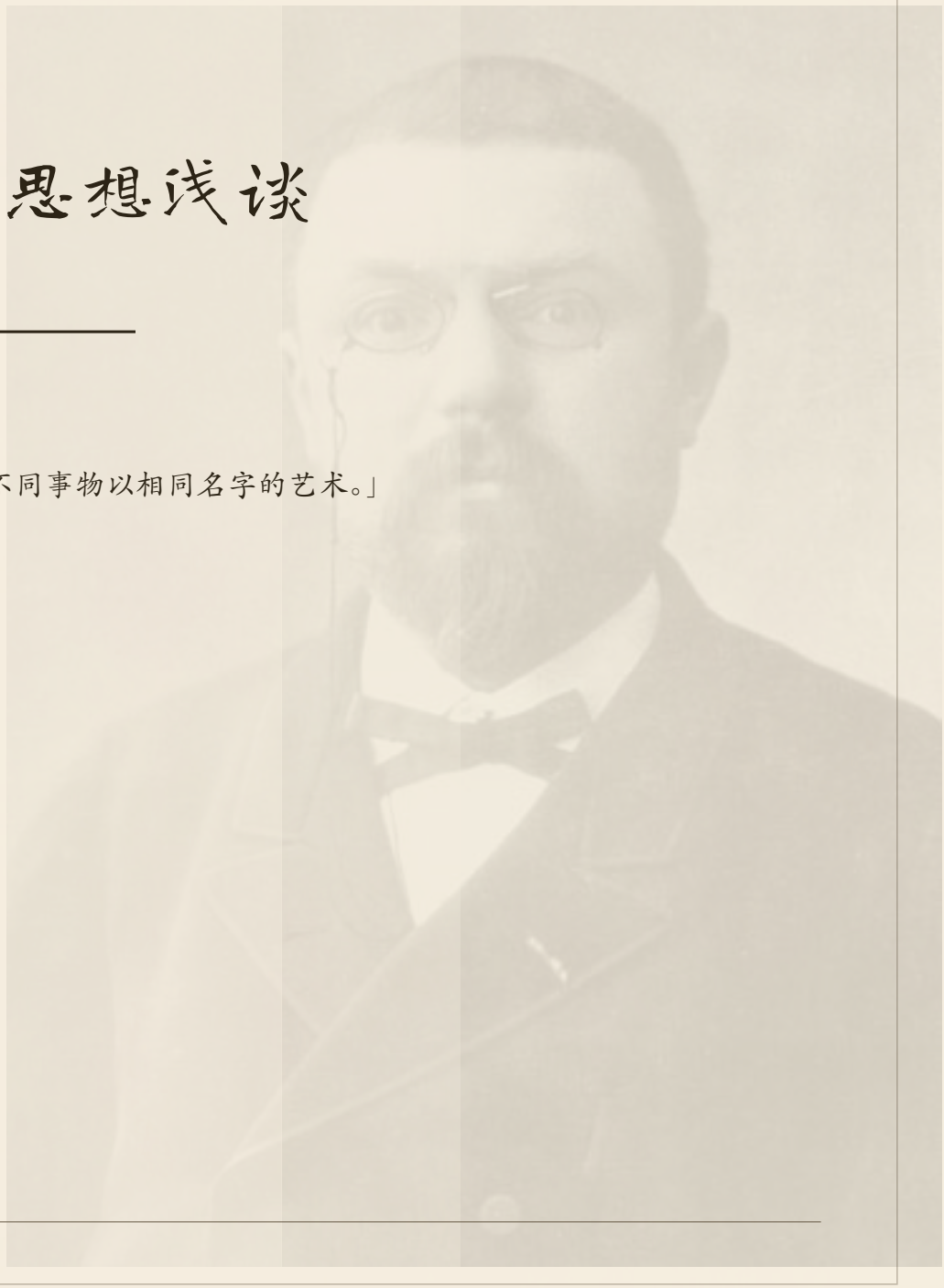
18.8. (极限速度)质量为 m 的雨滴下落,阻力为 $-cv$. 写出速度满足的方程并求极限速度;若雨滴初始静止, 速度达到极限速度一半所需的时间是多少? 提示: $v=\dfrac{mg}{c}(1-e^{-ct/m})$, 令 $v=\dfrac{1}{2}\dfrac{mg}{c}$ 解出 $t=\dfrac{m\ln 2}{c}$.

18.9. (悬链线, 选做) 悬链线 $y=a\cosh\dfrac{x}{a}$ 在最低点附近与抛物线 $y=\dfrac{x^2}{2a}$ 近似. 利用 e^u 在 $u=0$ 处的二阶泰勒展开说明理由, 并估计 $x=a$ 处两曲线的相对误差. 提示: $\cosh u=1+\dfrac{u^2}{2}+\dfrac{u^4}{24}+\cdots$; $x=a$ 处悬链线高度为 $a(\cosh 1-1)$, 抛物线高度为 $\dfrac{a}{2}$.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

第 19 章

数学思想浅谈



分析主干

「数学是给不同事物以相同名字的艺术。」

——庞加莱

第十九章 数学思想浅谈

19.1 数学对象与问题三要素

解题之前先审题. 审题要做的第一件事, 是把一道题拆成三个部分, 即**研究对象、已知条件、待求目标**.

- **研究对象**是题目所谈论的那个对象. 它可以是一个数、一个式子、一个点、一条曲线、一个函数, 也可以是一组对象构成的结构(一个数列、一个三角形、一个圆系).
- **已知条件**是对研究对象的约束. 题面明确写出的等式、不等式、几何关系属于条件; 定义域、整数性、连续性、隐含的取值范围同样是条件, 只是写得隐蔽.
- **待求目标**规定了解题的方向. 同一组对象和条件, 对应的目标可以是求值、求范围、证明一个断言、判断存在性、求轨迹, 问法不同, 路径随之改变.

例题 19.1 (同一个条件, 三种问法) 设正数 a, b 满足 $a+b=1$. 考虑三个问题. (1) 求 ab 的最大值; (2) 证明 $a^2+b^2 \geq \dfrac{1}{2}$; (3) 是否存在 a, b 使 $a^2+b^2=1$?

解 三个问题共享同一组研究对象(正数 a, b)和同一个已知条件 ($a+b=1$), 目标却各异.

(1) 由均值不等式 $ab \leq \left(\dfrac{a+b}{2}\right)^2 = \dfrac{1}{4}$, 等号在 $a=b=\dfrac{1}{2}$ 取到, 故 ab 的最大值为 $\dfrac{1}{4}$.

(2) $a^2+b^2=(a+b)^2-2ab=1-2ab$. 由 (1) 知 $ab \leq \dfrac{1}{4}$, 故 $a^2+b^2 \geq 1-2\cdot\dfrac{1}{4}=\dfrac{1}{2}$.

(3) $a^2+b^2=1-2ab$. 因为 a, b 为正数, $ab>0$, 故 $a^2+b^2<1$ 严格成立, 等号取不到. 因此不存在这样的 a, b.

研究对象与条件完全相同, 仅因目标从求最值换成证明不等式再换成判断存在性, 三段论证的依据与落点各不相同.

19.1.1 显条件与隐条件

题面写出的等式与不等式是**显条件**. 还有一类**隐条件**不写在题面上, 却同样约束着解的形态.

- 定义域与值域. $y=\log_a x$ 隐含 $a>0$、$a\neq 1$、$x>0$;
- 整数性. n 为正整数把连续变量截断为离散点, 自由度骤降;
- 几何存在性. 三角形隐含两边之和大于第三边; 圆的半径非负;

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 连续性与可导性. 题目说“函数 f 在 $\mathbb{R}$ 上可导”, 这就附带 f 连续这一隐藏条件.

隐条件常常是解题的关键, 也是丢解与增根的根源. 例如求过圆外一点引圆的切线, 若用斜率 k 参数化直线, 斜率不存在的竖直直线落在参数化范围之外, 这是一条必须单独检验的隐条件, 见 例题 19.17. 把隐条件写成显条件, 是审题的实质工作.

19.1.2 对象个体与对象族

知乎@余命数

研究对象有时是一个**个体**(“已知圆 $C:x^2+y^2=4$”), 有时是一个**族**, 比如过点 P 的直线或椭圆上的动点 Q. 区分二者, 决定了该不该设参数、该不该求范围.

一族对象由若干自由参数刻画, 参数取遍允许值就生成整族. 面对一族, 题目通常问的是这族的共同性质, 如求轨迹、求范围、求最值, 或问其中某个特殊成员(求满足额外条件的那个). 面对个体, 问的则是它的某个具体属性, 例如求值. 误把个体当族, 会无端引入参数徒增计算; 误把族当个体, 会漏掉一族解或答非所问.

19.1.3 确定的程度决定问题的形态

知乎@余命数

条件约束对象, 约束的强弱决定对象被锁定的程度, 锁定的程度又决定问题的形态.

- 对象被**完全确定**时, 问题表现为求值, 答案是一个或有限多个具体的数值;
- 对象只被**部分确定**时, 它仍保留变动的余地, 问题表现为求范围、求最值或求轨迹, 答案是一个区间或一条曲线.

下列一对只差一个条件的对照题, 直观展示这种差别.

例题 19.2 (三边给定的三角形) 已知 $\triangle ABC$ 的三边长分别为 $5,6,7$, 求它的面积.

知乎@余命数

解 三边长是三个独立的实数, 它们把三角形完全锁定(边边边全等). 剩下的只是代入公式. 半周长 $s=\dfrac{5+6+7}{2}=9$, 由海伦公式

$$S=\sqrt{s(s-5)(s-6)(s-7)}=\sqrt{9\times 4\times 3\times 2}=6\sqrt{6}.$$

对象被完全确定, 答案是一个确定的数值.

例题 19.3 (两边给定、夹角自由的三角形) 已知 $\triangle ABC$ 中 $AB=6$, $AC=5$, 而夹角 $\angle A$ 未给定, 求它的面积的取值范围.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

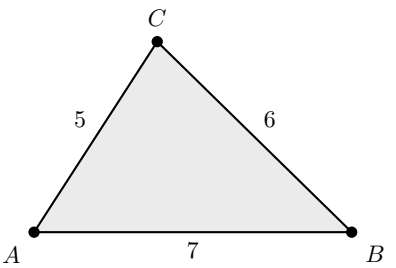
解 与上题相比, 条件从三个减为两个, 三角形未被锁定, 它仍可绕顶点 A 自由开合. 面积公式恰好把这种不确定性显示出来.

$$S = \frac{1}{2} \cdot AB \cdot AC \cdot \sin A = 15 \sin A.$$

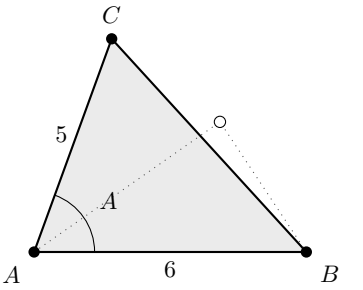
由 $0 < A < \pi$ 知 $0 < \sin A \le 1$, 故 $0 < S \le 15$, 当 $A = \dfrac{\pi}{2}$ 时取到最大值 15.

对象保留了一份变动的余地, 问题便从求值变成求范围与最值.

知乎@余命数



(a) 三边给定, 三角形被完全锁定.



(b) 仅两边给定, 三角形仍可绕 A 开合(虚线为另一可能位置).

图 19.1 确定的程度决定问题的形态

知乎@余命数

例题 19.2 与 例题 19.3 的差别只有一条边, 条件少了, 对象反而多了变动的余地. 这份余地的多少可以被精确计量, 计量它的工具是**自由度**, 我们在 第19.2节 中引入它.

19.2 自由度与变量、约束、独立参数

一个数学对象可以由若干个独立的实数来刻画. 完全确定这个对象所需的最少实数个数, 就是它的**自由度**. 自由度衡量一个数学对象内在的信息量. 几种熟悉的几何图形足以直观展示这一概念.

除了直观的几何图形, 更抽象的数学对象也同样具有自由度.

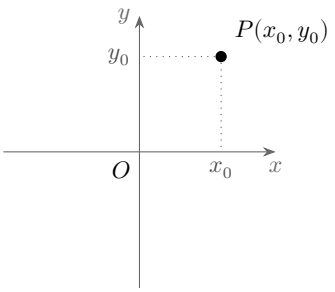
知乎@余命数

知乎@余命数

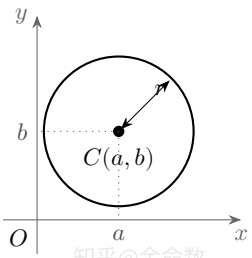
- **一个实数**是最基本的数学对象, 拥有 **1** 个自由度. 它的值可以是具体的数值, 如 $-\frac{7}{3}, e^3, 1+\ln 5$; 也可以是由题目参数决定的值, 如 $\sin a, a(a+1)$; 还可以是某个方程的隐式解. 例如, 锐角 θ 满足 $\cos\theta = \frac{1}{3}$, 虽然写不出 θ 的精确值, 它已被唯一确定, 其所有三角函数值也随之确定, 可视为已确定的对象.
- **平面上的直线**可由斜率 k 和截距 b 两个实数决定 $(y = kx + b)$, 拥有 **2** 个自由度. 垂直于 x 轴的直线是特殊情况, 不改变一般情形的自由度个数.
- **二次函数**由其三个系数 a, b, c 决定 $(y = ax^2 + bx + c)$, 拥有 **3** 个自由度.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

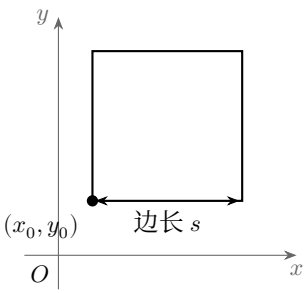
知乎@余命数



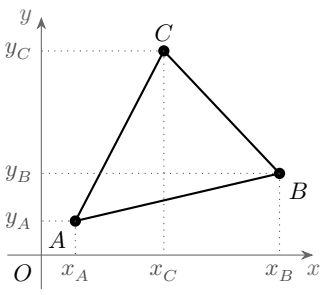
(a) 平面上的点, 由 (x_0, y_0) 确定, **2** 个自由度.



(b) 平面上的圆, 由 (a, b, r) 确定, **3** 个自由度.



(c) 轴平行正方形, 由 (x_0, y_0, s) 确定, **3** 个自由度.



(d) 任意三角形, 由 $(x_A, y_A, x_B, y_B, x_C, y_C)$ 确定, **6** 个自由度.

图 19.2　常见几何对象的自由度

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

19.2.1 变量、约束与独立参数

把对象写成一组参数, 参数的个数就是**变量数**; 每一条约束(一个等式或等价的不等式)消去一个自由度. 于是

$$\text{自由度} = \text{变量数} - \text{独立约束数}.$$

这里的**独立**二字不可省去, 只有彼此独立的约束才各消去一个自由度, 重复或可由其他约束推出的约束不提供新信息.

我们常说两点确定一条直线. 一个点有 2 个自由度, 两个点合起来有 $2+2=4$ 个自由度, 它们确定的直线却只有 2 个自由度, 差额出在参数的过剩上.

用两个点 (x_1, y_1) 和 (x_2, y_2) 描述一条直线时, 这四个参数提供的信息是**过剩**的. 它们既固定了直线的**位置**, 用去 2 个自由度, 也把这两点在直线上的**具体位置**固定下来(每个点沿直线需要 1 个自由度, 共 2 个自由度). 因此直线本身的自由度是 $4-2=2$. 多余的参数描述了直线之外的信息, 自由度的总数并不改变.

给对象设参数时, 应优先选择**独立**的参数, 让参数个数恰等于自由度. 过剩的参数会引入冗余方程, 徒增消元的负担; 欠少的参数又无法覆盖全部对象, 造成漏解.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

19.2.2 例题

例题 19.4 (相切的圆) 求所有满足以下条件的圆的方程. 它与 y 轴相切, 经过点 $P(1,2)$, 且圆心在直线 $y=x$ 上.

解 分析对象是一个平面上的圆.

该对象的自由度为 3, 可由其圆心坐标 (a,b) 和半径 r 完全确定, 标准方程为 $(x-a)^2+(y-b)^2=r^2$. 目标是用题目的三个条件确定这三个参数.

条件一为圆与 y 轴相切, 即圆心到 y 轴(直线 $x=0$)的距离等于半径, $|a|=r$. 由于 $r>0$, 得 $r^2=a^2$.

条件二为圆经过点 $P(1,2)$, $(1-a)^2+(2-b)^2=r^2$.

条件三为圆心 (a,b) 在直线 $y=x$ 上, $a=b$.

三个参数 (a,b,r) 对应三个方程.

$$\begin{cases} r^2=a^2 \\ (1-a)^2+(2-b)^2=r^2 \\ a=b \end{cases}$$

知乎@余命数 知乎@余命数

将 $b=a$ 和 $r^2=a^2$ 代入第二个方程，得

$$(1-a)^2+(2-a)^2=a^2$$

$$(1-2a+a^2)+(4-4a+a^2)=a^2$$

$$a^2-6a+5=0$$

$$(a-1)(a-5)=0$$

解得 $a=1$ 或 $a=5$.

当 $a=1$ 时，$b=1, r=1$，圆的方程为 $(x-1)^2+(y-1)^2=1$.

当 $a=5$ 时，$b=5, r=5$，圆的方程为 $(x-5)^2+(y-5)^2=25$.

因此存在两个满足条件的圆. 此例说明，当独立条件的个数等于对象的自由度时，通常得到离散的解集，可能恰有一解、有限多解，或无解. 条件个数与自由度的其他大小关系在 第19.3节 中讨论.

例题 19.5 (三次函数) 已知三次函数 $f(x)$ 满足 $f(0)=1$、$f(1)=0$、$f'(1)=0$，且 $f(2)=9$. 求 $f(x)$ 的表达式.

解 分析对象是一个三次函数.

其一般形式为 $f(x)=ax^3+bx^2+cx+d$ $(a\neq 0)$，由四个系数 (a,b,c,d) 完全确定，自由度为 4. 题目恰好给出四个独立的条件.

由条件 $f(1)=0$ 与 $f'(1)=0$ 可知 $f(x)$ 在 $x=1$ 处有一个重根，即 $f(x)$ 含有因子 $(x-1)^2$.

据此把三次多项式按 $(x-1)$ 的幂展开，即在 $x=1$ 处的泰勒形式，见 第16.13节，并入 第16章，让重根条件自动转化为系数的取值.

$$f(x)=k(x-1)^3+m(x-1)^2+n(x-1)+p$$

$f(1)=0 \implies p=0$. $f'(x)=3k(x-1)^2+2m(x-1)+n$，故 $f'(1)=0 \implies n=0$.

此时函数形式简化为 $f(x)=k(x-1)^3+m(x-1)^2$，只剩两个待定系数 (k,m)，自由度降为 2. 利用剩下两个条件.

$f(0)=1 \implies k(-1)^3+m(-1)^2=1 \implies -k+m=1$.

$f(2)=9 \implies k(1)^3+m(1)^2=9 \implies k+m=9$.

解关于 (k,m) 的线性方程组

$$\begin{cases} -k+m=1 \\ k+m=9 \end{cases}$$

得 $m=5, k=4$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

因此 $f(x)=4(x-1)^3+5(x-1)^2=4x^3-7x^2+2x+1$.

例题 19.6 (确定圆锥曲线) 平面直角坐标系中的一条圆锥曲线由一般二次方程 $Ax^2+Bxy+Cy^2+Dx+Ey+F=0$ 定义. 确定一条圆锥曲线需要几个自由度? 至少需要几个(处于一般位置的)点才能唯一确定它?

解 分析对象是一条圆锥曲线.

该对象的代数表示有 A,B,C,D,E,F 六个系数. 将整个方程乘以一个非零常数 λ, 即

$$(\lambda A)x^2+(\lambda B)xy+(\lambda C)y^2+(\lambda D)x+(\lambda E)y+(\lambda F)=0$$

所代表的曲线与原方程完全相同. 因此六个系数不同时为零, 且只在整体非零倍数的意义下唯一.

为消除这种尺度上的冗余, 选定其中一个非零系数归一化为 1. 例如 $F\neq 0$ 时, 两边同除以 F, 方程写为

$$A'x^2+B'xy+C'y^2+D'x+E'y+1=0$$

无论选取哪个非零系数作归一化, 都是从六个系数中扣除一个整体比例自由度. 因此一条圆锥曲线的自由度为 $6-1=5$.

要确定这五个参数, 需要五个独立的线性方程. 一个点 (x_i,y_i) 位于曲线上, 给出一个关于五个未知参数的线性方程.

$$A'x_i^2+B'x_iy_i+C'y_i^2+D'x_i+E'y_i=-1$$

因此, 唯一确定一条圆锥曲线至少需要五个点.

一般位置这个限制是必要的, 它排除退化情形. 例如五个点共线时, 以这条直线为因子的退化二次曲线有无穷多条经过它们, 无法唯一确定.

例题 19.7 (点的轨迹) 设圆 C 的方程为 $x^2+y^2=R^2$.A,B 是圆 C 上两个固定的不同点. 点 P 在圆 C 上运动(P 不与 A,B 重合). 求 $\triangle PAB$ 的垂心 H 的轨迹.

解 分析对象是 $\triangle PAB$ 的垂心 H.

先分析自由度. 圆 C 和点 A,B 固定, 没有自由度. 动点 P 在圆周上运动, 其位置可用一个参数, 如极角 θ 描述, 故 P 有 1 个自由度.

由 A,B,P 三点构成的 $\triangle PAB$ 也因此拥有 1 个自由度. 所求的垂心 H 完全由这三个顶点决定, 其位置也由这 1 个自由度控制. 一个自由度通常对应一条曲线, 因此预期 H 的轨迹是一条曲线.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为确定这条曲线, 采用向量法. 设圆心为坐标原点 $O(0,0)$, 令 $\vec{a},\vec{b},\vec{p}$ 分别为点 A,B,P 的位置向量.

对于以原点 O 为外心的三角形 PAB, 其垂心 H 的位置向量 $\vec{h}$ 满足

$$\vec{h}=\vec{a}+\vec{b}+\vec{p}$$

本题中 A,B 固定, $\vec{a}+\vec{b}$ 是固定向量, 记为 $\vec{s}=\vec{a}+\vec{b}$. 于是

$$\vec{h}=\vec{s}+\vec{p}$$

这个关系式清晰说明了 H 的轨迹. 点 P 在以 O 为圆心、R 为半径的圆 C 上运动, 其位置向量 $\vec{p}$ 的终点轨迹是圆 C. 向量 $\vec{h}$ 是固定向量 $\vec{s}$ 与动向量 $\vec{p}$ 的和, 表示 H 的轨迹是将 P 的轨迹(圆 C)按向量 $\vec{s}$ 平移所得的图形.

因此垂心 H 的轨迹位于一个半径同样为 R 的圆上, 圆心为 S, 位置向量为 $\vec{s}=\vec{a}+\vec{b}$. 严格地说, 由于题设要求 P 不与 A,B 重合, 轨迹应去掉与 $P=A$ 和 $P=B$ 对应的两个点;若讨论轨迹的闭包, 则是整个圆.

19.3 自由度如何决定解的个数

第19.2节 给出了自由度的计算方式, 即变量数减去独立约束数. 接着我们把它当作动笔之前的判断工具使用, 看自由度的数值对应着怎样的解的形态.

19.3.1 圆过点的序列

平面内的几何对象, 其自由度有直观的几何意义. 以圆为例层层加约束, 能看清自由度如何决定解的个数.

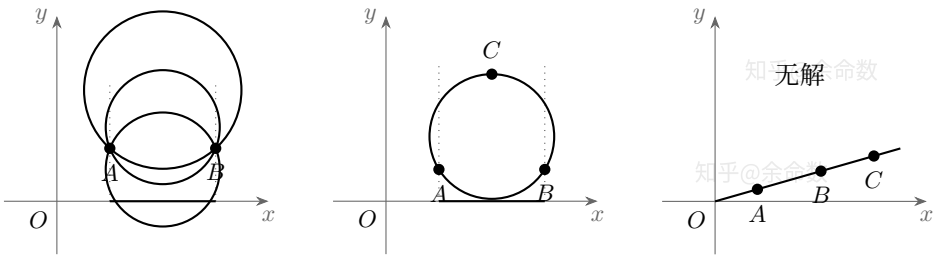
- **平面内一点**有 2 个自由度, 对应横、纵坐标两个独立实数.
- **一个圆**有 3 个自由度, 对应圆心两个坐标和半径.
- **过两个定点的圆**剩下 1 个自由度, 圆心必须落在两点的垂直平分线上, 沿这条直线滑动即得一族圆(半径随之改变).
- **过三个定点的圆**一般唯一确定, 自由度降为 0; 若三点共线, 则不存在这样的圆, 约束相容性被破坏.

自由度从正降到零, 解从一族收缩为有限个; 约束再加一重, 若新约束与已有约束相容, 不过是冗余确认, 若不相容, 则解集体消失.

19.3.2 三种形态

设待定对象的总自由度为 n, 独立的约束(满足式条件)个数为 m. 比较 m 与 n, 可在动笔之前预判解的形态.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



(a) 过两点的圆, 圆心在垂直平分线上滑动, 一族解.

(b) 过三点的圆, 唯一确定(圆心同时落在两条垂直平分线上).

(c) 三点共线,不存在过它们的圆, 约束不相容.

图 19.3 圆过点的序列, 自由度从 1 到 0 到不相容

- $m < n$ 即**欠定**, 对象无法被完全确定, 解是含 $n - m$ 个参数的一**族**对象. 此时问题的合理形态是求轨迹、求范围或求最值, 第19.1节 中那个夹角自由的三角形就是如此, $n - m = 1$ 的余地留给 $\sin A$ 去变化. 误以为能解出唯一答案, 就是误判了问题的形态.
- $m = n$ 即**恰定**, 通常得到**离散解集**, 恰有一解、有限多解, 或因条件隐性相斥而无解. 第19.2节 中相切圆有两解、三次函数有唯一解, 两者都属于这种形态.
- $m > n$ 即**超定**, 条件可能互相矛盾, 此时无解, 例如不存在过三个共线点的圆; 若条件彼此相容, 多余的等价条件不构成新信息, 可用于**检验**或提供计算的捷径.

19.3.3 自由度分析在解题中的作用

把三种形态翻译成解题动作.

- **自由度为零**($m = n$ 恰定). 原则上所有量都可以求出. 此时直接设参数、列方程、解出离散值, 最后回代检验相容性与定义域.
- **自由度为正**($m < n$ 欠定). 对象成族, 适合用参数化或函数观点处理. 题目问的是整族的共同性质, 即求范围、求最值、求轨迹, 目标是把族参数与所求量联系起来, 刻画整族的共同形态.
- **自由度为负**($m > n$ 超定). 约束过强, 首要任务是判断条件之间是否相容、是否真有解. 贸然求解可能得到形式解却无实际对象对应; 确认相容后, 多余的约束可作为检验或简化计算的工具.

19.3.4 三个常见陷阱

自由度分析简明, 却有几处容易失足.

约束未必独立. 计入 m 的只能是**独立**约束. 不独立的约束会造成看似足够、实则欠定的陷阱. 第19.2节 中确定圆锥曲线需要五个点, 前提是五点处于一般位置, 若有四点共线, 第五个条件就不再独立, 曲线也就无法唯一确定. 判断独立性没有机械的规则, 一个有效的自检是看这个条件是否真的缩小了解集.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

隐条件会改变自由度. 一道题表面上给出 m 个方程, 隐条件却可能暗中增减自由度. 例如 x 为正整数把连续的自由度压缩到离散点;“a,b,c 为三角形三边”附带三角不等式, 限制了参数的可行域. 自由度的计算必须把隐条件一并计入, 否则预判的形态会与实际不符.

整数解问题尤需谨慎. 整数约束使自由度概念从连续的实数降到离散的整数计数. 方程 $x+y=10$ 在实数域有 1 个自由度(一族解), 在正整数域却只有 9 个解, 自由度的连续直觉不再直接适用. 处理整数解问题时, 自由度分析只能作粗略定位, 精确的解的个数需要数论工具, 见 第9章 的计数与整除.

这样, 自由度分析在动笔之前就能给出方向的预判, 欠定的问题不必空求唯一解, 超定的问题先验相容性再动手.

19.4 解题框架的四种基本路径

第19.3节 用自由度预判了解的形态. 形态既定, 接下来是动手推进. 推进之前还有一段调查工作, 先把题目变成可以动手的材料;调查完成之后, 再把条件分类, 并在四条基本路径中选择与组合.

19.4.1 动手调查

自由度分析回答的是问题有没有确定的答案, 它不说明第一步该写什么. 从读懂题意到选定路径之间隔着一段摸索, 这段摸索可以拆成几项具体动作, 分别是动手列举、缩小规模、幻想答案, 以及在举例时守住猜想的边界. 它们的共同作用是先制造出可以操作的材料, 把抽象的题面变成具体的例子, 解题思路常常在这些例子中间浮现出来.

19.4.1.0.1　动手列举　题面条件少而怪的时候, 最有效的起步是把所有可能的情形列出来. 数据一旦摆到桌面上, 隐藏的结构就有了着落.

> **例题 19.8 (普查员的难题)** 人口普查员询问一位母亲三个女儿的年龄. 母亲给出两条信息, 女儿们的年龄都是正整数, 乘积等于 36. 普查员说这不够. 母亲又说, 她们年龄之和恰好是自家的门牌号. 普查员看了一眼门牌号, 仍然说不够. 母亲想了想, 补充一句, 大女儿养了一条狗. 普查员马上得出了三个年龄. 求三个女儿的年龄, 并说明门牌号是多少.

解 乘积等于 36 的正整数三元组一共八组, 把它们连同各自的年龄之和一并列出.

年龄	$(1,1,36)$	$(1,2,18)$	$(1,3,12)$	$(1,4,9)$	$(1,6,6)$	$(2,2,9)$	$(2,3,6)$	$(3,3,4)$
之和	38	21	16	14	13	13	11	10

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

普查员知道了门牌号, 却依然无法确定年龄, 说明他所看到的那一个和数对应不止一组解. 八组数据当中只有和数 13 出现了两次, 对应 $(1,6,6)$ 与 $(2,2,9)$, 门牌号因此是 13. 第三条信息说的是“大女儿”养狗, 可见年龄最大的人只有一个, $(1,6,6)$ 里有两个孩子同为六岁, 不符合这一条. 于是三个年龄依次是 $2,2,9$.

整道题只能靠列举推进. 乘积条件展开成八行数据之后, “信息仍然不够”这句话立刻暴露出重复的和. 列举把语言描述的线索转成可以比对的数据, 这一步没有别的办法替代.

19.4.1.0.2 缩小规模 一般情形无从下手时, 先缩小问题的规模. 把数字改小, 把条件删减, 让结构在简单版本里先出现, 再带着这份认识回到原题.

例题 19.9 (连续正整数之和) 哪些正整数可以写成至少两个连续正整数之和? 例如 $6=1+2+3, 36=11+12+13$.

解 一般情形无从下手, 先缩小规模, 从 1 到 16 逐个检验, 把能写出的数写出一种分解.

- 能写出的数有 $3=1+2$、$5=2+3$、$6=1+2+3$、$7=3+4$、$9=2+3+4$、$10=1+2+3+4$、$11=5+6$、$12=3+4+5$、$13=6+7$、$14=2+3+4+5$、$15=7+8$.
- 写不出的数恰好是 $1,2,4,8,16$, 即 2 的幂.

分界线落在 2 的幂上, 于是猜想, 能写成至少两个连续正整数之和的正整数, 恰好是全体不是 2 的幂的正整数.

猜想分两半, 先证 2 的幂写不出. 设从 a 起共写了 k 项, 其中 $k\geq 2$ 且 $a\geq 1$, 由等差数列求和公式 (第7.4节)

$$2n=k\,(2a+k-1).$$

右边两个因子 k 与 $2a+k-1$ 恰好一奇一偶. 若 n 是 2 的幂, 则 $2n$ 也是 2 的幂, 而 2 的幂的因子只能是 2 的幂, 一奇一偶便迫使那个奇因子等于 1. 若 $k=1$, 与 $k\geq 2$ 矛盾; 若 $2a+k-1=1$, 与 $a\geq 1$ 且 $k\geq 2$ 矛盾. 故 2 的幂写不出, 这一半成立.

再证非 2 的幂都写得出. 设 n 不是 2 的幂, 把 $2n$ 分解成一个大于 1 的奇数与一个 2 的幂的乘积, 设其中较小的因子为 u, 较大的因子为 v. 取项数 $k=u$, 并令 $2a+k-1=v$, 解得

$$a=\frac{v-u+1}{2}.$$

两个因子一奇一偶, 故 $v-u+1$ 是偶数, a 是整数; $v>u$ 说明 $v-u+1\geq 2$, 即 $a\geq 1$; 而 u 是 2 的幂或者奇因子, 都满足 $k\geq 2$ 且各项为正. 这组 a 与 k 给出了合法的分解, 后一半也成立.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

两半合起来,能写成至少两个连续正整数之和的正整数,恰好是全体不是 2 的幂的正整数. 以 $n=36$ 验算,$2n=72=8\times 9$,取 $u=8, v=9$,得 $k=8$ 与 $a=1$,于是 $36=1+2+\cdots+8$.

19.4.1.0.3　幻想答案　还有一种起步方式看起来不合规矩,却经常奏效. 假装题目已经解出,凭直觉先写下一个答案,再回头逐条检验它. 检验中出现的差距就是题目缺少的部分,把差距写成方程,方程组随即出现.

例题 19.10 (培训住宿) 公司送 216 名员工外出参加培训,需要在酒店住一晚. 酒店提供双人房与四人房两种房型,每间双人房 100 元,每间四人房 150 元,住宿预算共 9600 元. 两种房各订多少间,才能恰好住满全部员工并把预算用完?

解　先假装答案是双人房 30 间,四人房 39 间. 这两个数字是随手写的,不必正确,要紧的是它们带着单位,可以拿去检验.

第一条检验是人数,$30\times 2+39\times 4=216$,恰好住满,这条通过. 第二条检验是预算,$30\times 100+39\times 150=8850$,比 9600 少 750,这条不通过.

两条检验正好就是题目的两条约束. 设双人房 x 间,四人房 y 间,把检验写成方程

$$\begin{cases} 2x+4y=216 \\ 100x+150y=9600 \end{cases} \Longrightarrow \begin{cases} x+2y=108 \\ 2x+3y=192 \end{cases} \Longrightarrow \quad y=24,\ x=60.$$

验算人数 $60\times 2+24\times 4=216$,预算 $60\times 100+24\times 150=9600$,两条同时满足. 答案是双人房 60 间,四人房 24 间.

写下一个具体答案,等于提前给自由度命名 (第19.2节). 未知数 x 与 y 在代入检验的那一刻就位,两条检验对应两条约束,方程组的形态先于列方程一步出现. 应用题正面列式感到困难时,从答案这一端倒着走,常常比从条件那端硬凑更快.

19.4.1.0.4　例证与猜想　列举与实验会给出猜想,而猜想有时并不成立.

例题 19.11 (两个失效的猜想) (1) 多项式 $P(x)=x^2+x+41$ 在 $x=0,1,2,\ldots,39$ 处的取值都是素数,能否据此断言它对一切正整数都取素数?
(2) 圆周上取 n 个点,两两连线,设任意三条弦不共点.$n=1,2,3,4,5$ 时圆被分成的区域数依次为 $1,2,4,8,16$,能否据此断言区域数总等于 2^{n-1}?

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 (1) 不能. 前四十个值确实都是素数, 但 $P(40)=1600+40+41=41^2$ 是合数. 这个多项式在一段不短的范围里始终保持素数, 原因在初等层面看不出来, 恰好说明有限次检验的说服力十分有限.

(2) 不能.$n=6$ 时区域数是 31, 与 $2^5=32$ 相差一个. 精确的计数也不难做出. 把弦一条一条添进圆里, 每一条新弦被它与此前各弦的交点分成若干段, 每一段都把所在的一个区域一分为二,所以新弦增加的区域数等于它上面的交点数加一. 全部累加, 区域总数等于 1 加上弦的条数 $\binom{n}{2}$, 再加上相交的弦对数. 每一对相交的弦由圆上四个点确定, 反过来每四个点恰好给出唯一一对相交的弦, 相交弦对数为 $\binom{n}{4}$. 于是区域数为

$$\binom{n}{2}+\binom{n}{4}+1,$$

代入 $n=5$ 得 16, 代入 $n=6$ 得 31, 数列在第六项处离开了 2 的幂.

例证无论积累多少, 覆盖的都只是有限个情形. 一个全称命题的反例只需要一个, 而检验每做一次, 只排除有限多个整数, 这就是列举与实验无法代替证明的原因. 第7.1节 的数学归纳法把无穷多个情形压缩成两个有限的步骤, 正因如此它值得单独讨论.

调查得到的材料, 接下来交给四条基本路径使用.

19.4.2 构造式与满足式两类条件

定义 19.1 题目中的条件可分为两类.

- **构造式条件**是直接表明如何构建出数学对象, 或其一部分的条件. 例如“点 O 的坐标为 $(0,0)$”、“抛物线方程为 $y^2=4x$”、“点 A 是 C_1 和 C_2 的交点”.
- **满足式条件**是对象必须恰好满足的约束, 通常表现为方程或等式, 是需利用和求解的目标. 例如 $a_5+a_6=16$、直线与圆相切.

构造式条件回答对象从哪里来, 满足式条件回答对象还要满足什么. 一道题的全部条件几乎都可以归入这两类; 剩下要做的, 是让它们各司其职. 这一分类是后续 第19.5节 与 第19.7节 两节的方法论起点, 前者是纯构造式, 后者含满足式.

19.4.3 四条基本路径

19.4.3.0.1 正向推进 从已知条件出发, 逐步推出能推出的结论, 看能否抵达目标. 适用于条件较多、结论相对简单的情形. 风险在于可能推出大量中间结果, 却不知哪一条通往目标, 陷入漫无目的的推导.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 19.12 (正向推进的缩影) 已知 a,b,c 为正数且 $abc=1$, 证明 $\dfrac{1}{a}+\dfrac{1}{b}+\dfrac{1}{c}\geq 3$.

解 从条件 $abc=1$ 出发, $\dfrac{1}{a}+\dfrac{1}{b}+\dfrac{1}{c}=bc+ca+ab$. 由三元均值不等式 $ab+bc+ca\geq 3\sqrt[3]{a^2b^2c^2}=3$, 得证. 条件单一而结论是熟悉的均值形式, 顺势正向推进即可.

19.4.3.0.2　反向分析　从结论出发, 思考"要结论成立, 只需什么先成立", 一步步倒退到已知条件. 适用于结论复杂、条件较少的情形. 风险在于反向步骤未必可逆, 由结论推出的必要条件未必充分, 需在写正式证明时正向验证每一步的等价性.

例题 19.13 (反向分析的缩影) 设 n 为任意正整数, 证明 $\sqrt{n+1}-\sqrt{n}<\dfrac{1}{2\sqrt{n}}$.

解 要证 $\sqrt{n+1}-\sqrt{n}<\dfrac{1}{2\sqrt{n}}$, 只需证其有理化形式 $\dfrac{1}{\sqrt{n+1}+\sqrt{n}}<\dfrac{1}{2\sqrt{n}}$, 只需 $\sqrt{n+1}+\sqrt{n}>2\sqrt{n}$, 只需 $\sqrt{n+1}>\sqrt{n}$, 而 $n+1>n$ 且算术平方根保持不等号方向, 此式成立. 反向找到的链条每一步都可逆, 正向重写即得证明.

19.4.3.0.3　倒数第二步　反向分析常以倒数第二步的形式出现. 遇到链条很长、起点又推不出来的情形, 不去回溯整条链条, 只问最后一步需要什么前提, 把这个前提当作独立的命题先解决掉, 再回头解决原来的问题.

例题 19.14 (分数不可约) 证明对任意正整数 n, 分数 $\dfrac{21n+4}{14n+3}$ 都不能约分.

解 要证不能约分, 只需证明分子与分母的最大公因数是 1. 要证这个最大公因数是 1, 只需找到分子与分母的一个整数组合等于 1. 组合并不难找

$$3(14n+3)-2(21n+4)=42n+9-42n-8=1,$$

于是分子与分母的最大公因数整除 1, 只能等于 1. 若把最后一步写成一般形式, 它是 **??** 中裴蜀定理的充分方向.

证明只有三步, 每一步都在回答上一步提出的前提, 原命题始终没有从正面碰过. 倒数第二步把目标换成一个更小也更具体的问题, 是反向分析中最常用的一种压缩方式.

19.4.3.0.4　双向夹击　同时从条件和结论出发, 在中间寻找相遇点. 条件端推出一批性质, 结论端倒退一批前提, 两边靠近直到衔接. 几何证明中尤为常见, 从已知关系推一端, 从待证关系推另一端, 会在某个中间等式或比例上汇合.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 19.15 (双向夹击的缩影) 在 $\triangle ABC$ 中, D 为 BC 中点, $AB = AC$. 用双向夹击说明 $AD \perp BC$.

解 从条件端看, 等腰且 D 为底中点, 可推 $\triangle ABD$ 与 $\triangle ACD$ 三边对应相等($AB = AC$、$BD = CD$、AD 公共), 故 $\angle ADB = \angle ADC$. 从结论端看, 要证 $AD \perp BC$, 即 $\angle ADB = \angle ADC = 90^\circ$, 即此二角相等且互补. 两端在"$\angle ADB = \angle ADC$"相遇, 补上平角 180° 即得垂直.

19.4.3.0.5 模式识别 识别题目背后的深层结构, 如对称多项式、递推关系、几何变换、典型模型, 把陌生题归入熟悉模式. 这是最省力的路径, 也是最容易出错的路径. 模式识别依赖经验的厚度, 一旦僵化套用, 形似而质异的题目便被误塞进同一模板.

例题 19.16 (模式识别的缩影) 设 x, y, z 满足 $x + y + z = 0$, $x^2 + y^2 + z^2 = 6$, 求 $x^4 + y^4 + z^4$.

解 识别出对称多项式的模式, 令 $s = x + y + z = 0$、$p = xy + yz + zx$、$q = xyz$. 由 $x^2 + y^2 + z^2 = s^2 - 2p = 6$ 及 $s = 0$ 得 $p = -3$. 又

$$x^4 + y^4 + z^4 = (x^2 + y^2 + z^2)^2 - 2(x^2y^2 + y^2z^2 + z^2x^2),$$

其中 $x^2y^2 + y^2z^2 + z^2x^2 = p^2 - 2sq = p^2 = 9$. 故 $x^4 + y^4 + z^4 = 36 - 18 = 18$. 识别对称结构后, 所有运算沿对称多项式的标准通道展开.

19.4.4 四条路径的交织

四条路径可以交织使用, 多数题目都需要如此. 先模式识别定位结构, 再正向推进展开条件, 遇阻时反向分析寻找突破口, 过程中双向夹击收拢两端. 综合例题最能体现这种交织.

例题 19.17 (过定点引圆的切线) 求过点 $P(3, 1)$ 且与圆 $x^2 + y^2 = 4$ 相切的直线的方程.

解 先让框架空转一遍, 看清形态.

清点对象, 已知对象有圆 $x^2 + y^2 = 4$(由方程直接给出)与点 $P(3, 1)$, 待定对象是一条直线. 标注自由度, 平面上的直线有 2 个自由度, $n = 2$. 分类条件, "直线过点 P"与"直线与圆相切"都没有直接给出直线, 都是满足式条件, $m = 2$. 预判形态, $m = n = 2$, 恰定, 预期得到有限条切线; 由 $3^2 + 1^2 = 10 > 4$ 知 P 在圆外, 应当恰有两条.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这一步是**反向分析**与**模式识别**的交汇, 识别出“切线 + 外点”的标准模式, 预判出离散解集的个数, 避免动手时心里没底.

正向求解时, 先为直线的自由度命名. 设斜率为 k, 直线为 $y-1=k(x-3)$, 即 $kx-y+1-3k=0$.“过 P”已用在参数形式里, 剩下“相切”翻译为圆心到直线的距离等于半径.

$$\frac{|1-3k|}{\sqrt{k^2+1}}=2 \implies (1-3k)^2=4(k^2+1) \implies 5k^2-6k-3=0$$

解得 $k=\dfrac{3\pm 2\sqrt{6}}{5}$. 对应的两条切线为

$$y-1=\frac{3+2\sqrt{6}}{5}(x-3) \quad 与 \quad y-1=\frac{3-2\sqrt{6}}{5}(x-3).$$

最后回代检验. 用斜率 k 为直线命名时, 默认了斜率存在, 竖直直线 $x=3$ 被排除在参数化之外, 它也可能是被漏掉的解. 点 $P(3,1)$ 确在 $x=3$ 上, 但圆心到 $x=3$ 的距离为 $3>2$, 不相切, 本次没有遗漏. 这一步绝不能省. 若把题中的点换成 $(2,1)$, 竖直直线 $x=2$ 恰是一条切线, 漏参就会丢掉一个解.

19.4.5 为自由度命名

例题 19.17 中还有一处自由度参数的选择. 我们选用斜率 k(配合定点), 也可以改用**切点** (x_0,y_0) 命名, 此时切线方程为 $xx_0+yy_0=4$, 条件过 P 给出 $3x_0+y_0=4$, 与 $x_0^2+y_0^2=4$ 联立, 同样是 $m=n=2$ 的封闭方程组. 为自由度选择哪个名字, 是计算量的分水岭, 第19.7节 将专门展开.

19.4.6 框架早已在手中

这套框架只是把前几卷已有经验**统一命名**.

- 解析几何中的**设点法与设线法**, 是为自由度命名的两种范式（第13.4节 至 第13.7节, 决策见 第13.16节）;
- 恒成立问题中的**必要性探路与端点分析**, 是欠定情形下先用特例压缩候选集、再证充分性的策略 (第16.15节);
- 数列中“首项 + 递推关系确定一个数列”, 是用构造式条件锁定一个无穷对象 (第7.9节);
- 函数方程的**赋值法**, 是主动代入构造式取值、迫使抽象对象显形 (第4.1节).

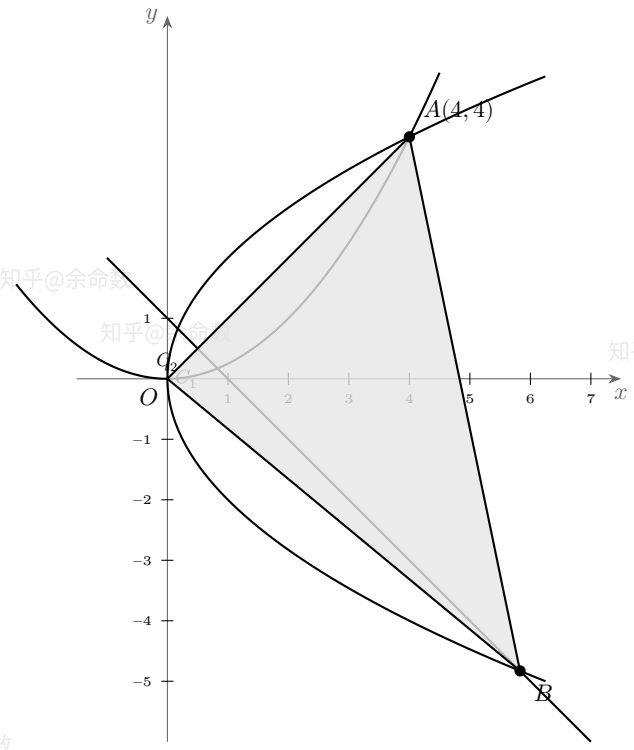
面对新题时不妨先让框架空转一遍, 清点对象、标注自由度、给条件分类、选好路径, 再调用各章的具体技术去执行. 解题框架根据题目结构动态调整, 没有固定流程. 条件多则正向, 结论繁则反向, 几何题常双向夹击, 结构熟则模式识别.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

19.5 纯构造式问题的解法

当所有条件都是构造式时, 解题路径非常直接, 沿着构造链条一步步确定出所有对象即可.

例题 19.18 (抛物线交点与面积) 已知原点 $O(0,0)$, 抛物线 $C_1: y^2=4x$ 和抛物线 $C_2: x^2=4y$ 在第一象限的交点为 A, C_1 与直线 $l: x+y-1=0$ 在第四象限的交点为 B, 求 $\triangle OAB$ 的面积.



解 本题是典型的纯构造式问题. 已知条件直接给出几何对象 O, C_1, C_2, l, 而点 A 作为 C_1 与 C_2 的交点、点 B 作为 C_1 与 l 的交点随之被构造出来, 于是三个确定的点 O, A, B 便构造出 $\triangle OAB$ 及其面积. 解题沿构造链分三段推进, 先定 A 的坐标, 再定 B 的坐标, 最后用坐标求面积.

联立 C_1 与 C_2 的方程定出点 A.

$$\begin{cases} y^2=4x \\ x^2=4y \end{cases} \Longrightarrow x^4=(4y)^2=16y^2=16(4x)=64x$$

即 $x(x^3-64)=0$, 解得 $x=0$ 或 $x=4$. 因为点 A 在第一象限, 取 $x=4$ 代入得 $y=4$, 故 $A(4,4)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

联立 C_1 与 l 的方程定出点 B.

$$\begin{cases} y^2=4x \\ x+y-1=0 \end{cases} \implies y^2=4(1-y) \implies y^2+4y-4=0$$

解得 $y=\frac{-4\pm\sqrt{16-4(-4)}}{2}=-2\pm2\sqrt{2}$. 因为点 B 在第四象限, $y<0$, 取 $y=-2-2\sqrt{2}$, 此时 $x=1-y=3+2\sqrt{2}$, 故 $B(3+2\sqrt{2},-2-2\sqrt{2})$.

将两点的坐标代入坐标面积公式 $S=\frac{1}{2}|x_Ay_B-x_By_A|$ 求出三角形面积.

$$\begin{aligned} S_{\triangle OAB} &= \frac{1}{2}\left|4(-2-2\sqrt{2})-(3+2\sqrt{2})\cdot 4\right| \\ &= \frac{1}{2}\left|-8-8\sqrt{2}-12-8\sqrt{2}\right| \\ &= \frac{1}{2}\left|-20-16\sqrt{2}\right| \\ &= 10+8\sqrt{2} \end{aligned}$$

沿构造链推进时, 还有一处自由度常被忽略, 即**构造的次序**. 目标对象未必需要先显式定出全部中间对象; 有时可以越过它们, 直接由已知对象构造出目标.

例题 19.19 (两圆的公共弦所在直线) 设圆 $C_1:x^2+y^2=4$ 与圆 $C_2:(x-4)^2+y^2=9$ 交于 A,B 两点, 求直线 AB 的方程.

解 沿构造链走到底, 可直接求出两交点再连线. 联立两圆方程, 消去 x^2+y^2 得 $x=\frac{11}{8}$, 代回得 $y=\pm\frac{3\sqrt{15}}{8}$, 即

$$A\left(\tfrac{11}{8},\tfrac{3\sqrt{15}}{8}\right),\quad B\left(\tfrac{11}{8},-\tfrac{3\sqrt{15}}{8}\right).$$

两点横坐标相同, 直线 AB 为 $x=\frac{11}{8}$. 这条路径万无一失, 但要穿越根式.

越过交点, 可由两圆方程直接构造出公共弦. 交点 A,B 的坐标同时满足两个圆的方程, 因此也满足两方程之**差**.

$$[(x-4)^2+y^2-9]-[x^2+y^2-4]=0 \implies -8x+11=0.$$

差是一个一次方程, 表示一条直线; A,B 两点都在其上, 而两点确定一条直线 (第19.2节), 故直线 AB 即 $8x-11=0$. 整个过程没有求出任何一个交点坐标.

次序二的实质, 是识别出目标对象即公共弦所在直线, 可以由两圆方程这个已知对象**直接组合而成**, 于是构造链被短路了. 这种确定对象而不求出对象的姿态, 正是解析几何中**设而不求**思想的萌芽 (第13.16节 有系统讨论); 两方程相减所得的这条直线也被称为两圆的**根轴**, 在圆幂问题中反复出现.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

纯构造式问题有两条经验, 一条是沿构造链步步确定对象, 另一条是推进之前先问一句, 目标能否由已有对象直接构造, 哪些中间对象其实不必显式求出.

19.6 直接构造困难时的反证法

知乎@余命数

第19.5节 的构造法直接造出对象, 证明它存在. 当正面构造受阻, 题目要求证明*不存在*某对象、某对象*唯一*、某类对象*有无穷多个*时, 反证法常常更省力. 它把结论的否定作为额外条件引入, 反而打开了新的推理空间.

19.6.1 反证法的逻辑结构

要证明 P 推出 Q, 反证法改为证明"P 与非 Q 同时成立会推出矛盾". 既然 P 已知为真, 非 Q 又引出矛盾, 非 Q 必假, Q 必真. 形式上多假设了一个非 Q, 实质上是给推理增添了可用的结构. 证明不存在时, 可以假设它存在并加以使用; 证明唯一时, 可以假设有两个并加以比较.

反证法给出的是**判定**, 它说明对象不存在、唯一或有无穷多, 却不直接造出对象. 需要具体对象时, 往往仍要另寻构造途径.

19.6.2 按适用场景分类

19.6.2.0.1 证明不存在某对象 假设该对象存在, 导出矛盾. 这是最自然的反证场景, 因为存在一旦被假设, 就提供了可操作的实体.

例题 19.20 ($\sqrt{2}$ 的无理性) 证明 $\sqrt{2}$ 不是有理数.

解 假设 $\sqrt{2}=\dfrac{p}{q}$, 其中 p,q 为互素的正整数(既约分数). 则 $p^2=2q^2$, 故 p^2 为偶数, p 为偶数. 设 $p=2k$, 代入得 $4k^2=2q^2$, 即 $q^2=2k^2$, 故 q 也为偶数.p,q 同为偶数与它们互素矛盾. 故 $\sqrt{2}$ 不是有理数.

19.6.2.0.2 证明唯一性 假设存在两个不同的对象都满足条件, 推出它们其实相等.

例题 19.21 (方程解的唯一性) 证明方程 $x^3+x+1=0$ 在 $\mathbb{R}$ 上恰有一个实根.

解 $f(x)=x^3+x+1$ 连续, $f(0)=1>0$, $f(-1)=-1<0$, 由介值定理至少有一实根. 假设有两个不同实根 $x_1\neq x_2$. 由罗尔定理, $f'(x)=3x^2+1$ 在 (x_1,x_2) 内有零点. 但 $3x^2+1\geq 1>0$ 恒成立, f' 无零点, 矛盾. 故实根唯一.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

19.6.2.0.3　证明某类对象有无穷多个　假设只有有限个, 据此构造出一个不在其中的新对象, 矛盾.

例题 19.22 (素数无穷多) 证明素数有无穷多个.

解 假设素数只有限个 $p_1, p_2, \dots, p_n$. 令 $N = p_1 p_2 \cdots p_n + 1$. 对每个 p_i, N 除以 p_i 余 1, 故 $p_i \nmid N$. 但 $N > 1$ 必有素因子, 该素因子不在 $p_1, \dots, p_n$ 之中, 与素数已全部列出矛盾. 故素数无穷多.

19.6.2.0.4　证明至少型命题　形如至少有一个对象满足某性质的命题, 假设全部对象都不满足该性质, 导出与已知条件的矛盾.

例题 19.23 (抽屉原理的至少型) 把 5 个点放入一个边长为 2 的正方形, 证明至少有两个点的距离不超过 $\sqrt{2}$.

解 将正方形等分为 4 个边长 1 的小正方形. 假设任意两点距离都大于 $\sqrt{2}$, 则每个小正方形内至多一个点(同一小正方形内两点距离 $\le \sqrt{2}$). 4 个小正方形至多容纳 4 个点, 却要放入 5 个, 矛盾. 故至少有一对点距离不超过 $\sqrt{2}$.

19.6.3 反证法的局限

反证法并非万能. 它擅长给出存在性、唯一性、无穷性的判定, 却有两处局限.

其一, 否定结论后若无可用结构, 反证便无从展开. 假设对象不存在时手里没有可操作的对象, 假设对象不唯一时若两对象的差异无法比较, 推理都会卡住. 此时正面构造或直接论证可能更顺手.

其二, 反证法即便成功, 也只回答有没有、是否唯一、是否无穷, 不回答是什么. 证明了解的存在之后, 要具体求出或构造该对象, 仍需借助构造法或满足式策略(第19.5节、第19.7节). 把反证法当作存在性判定的工具, 把构造法当作造出对象的工具, 两者配合使用.

19.7 含满足式问题的通用策略

题目里出现满足式条件时, 我们**用自由度参数表达出满足式条件, 然后解方程**.

例题 19.24 (椭圆上的角与斜率) 设椭圆 $C: \frac{x^2}{4} + \frac{y^2}{3} = 1$ 的左、右顶点分别为 $A(-2,0)$, $B(2,0)$. P 为 C 上一点且在 x 轴上方. 求 $\tan\angle PAB + 2\tan\angle PBA$ 的最小值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

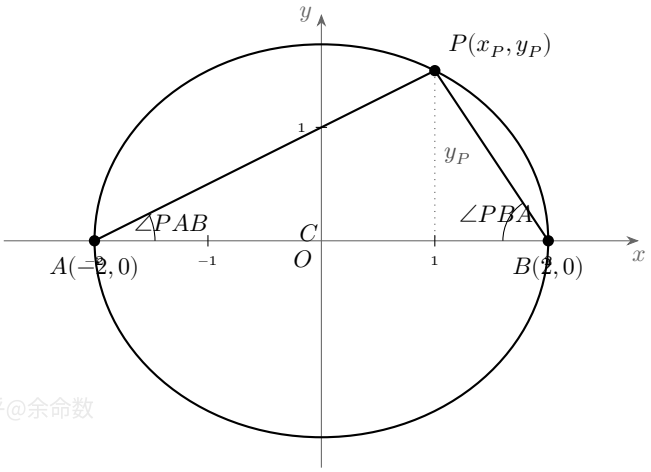


图 19.4 椭圆上点 P 与两顶点连线的斜率及夹角.

面对此题, 第一反应通常是为动点 P 设定变量. 无论设坐标 $P(x,y)$ 还是用参数方程 $P(2\cos\theta, \sqrt{3}\sin\theta)$, 都会将问题导向含根式或复杂三角函数的表达式, 后续求最值异常繁琐. 这提示存在一个更优的视角来描述点 P 的运动.

椭圆的第三定义指出, 对椭圆 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ 上任意一点 P(非左右顶点), 其与左右顶点 A, B 连线的斜率之积为定值.

$$k_{AP}\cdot k_{BP}=-\frac{b^2}{a^2}$$

本题中 $a^2=4, b^2=3$, 故 $k_{AP}\cdot k_{BP}=-\frac{3}{4}$.

接着把原问题中的几何语言翻译为斜率语言.$\angle PAB$ 是直线 AP 的倾斜角, 故 $\tan\angle PAB = k_{AP}$. 而 $\angle PBA$ 是直线 BP 与 x 轴**负向**的夹角. 设 BP 的倾斜角为 α_{BP}, 则 $\angle PBA=\pi-\alpha_{BP}$, 故

$$\tan\angle PBA=\tan(\pi-\alpha_{BP})=-\tan\alpha_{BP}=-k_{BP}$$

于是所求表达式化为

$$S=k_{AP}-2k_{BP}$$

点 P 在 x 轴上方, $k_{AP}>0$ 且 $k_{BP}<0$. 令 $k_A=k_{AP}>0$ 及 $k'_B=-k_{BP}>0$, 则 $S=k_A+2k'_B$, 约束变为

$$k_A\cdot(-k'_B)=-\frac{3}{4}\quad\Longrightarrow\quad k_A\cdot k'_B=\frac{3}{4}$$

应用基本不等式

$$S=k_A+2k'_B\geq 2\sqrt{k_A\cdot(2k'_B)}=2\sqrt{2(k_Ak'_B)}=2\sqrt{2\cdot\frac{3}{4}}=\sqrt{6}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等号当且仅当 $k_A = 2k'_B$ 时成立, 联立 $k_A k'_B = 3/4$ 可解出 k_A, k'_B, 表明最值可以取到. 所求最小值为 $\sqrt{6}$.

二级结论实质上是基本的变量代换. 它把“点 $P(x,y)$ 在椭圆上”这一关于**坐标变量** (x,y) 的约束, **翻译**成“连接 P 与顶点的两条直线斜率满足某关系”这一关于**斜率变量** (k_{AP}, k_{BP}) 的约束. 完成这一翻译, 原理便随之清晰.

给定两个斜率 $k_A = k_{AP}$ 和 $k_B = k_{BP}$, 写出两直线方程并求其交点 $P(x_P, y_P)$.

$$\begin{cases} y = k_A(x+2) \\ y = k_B(x-2) \end{cases}$$

联立解得

$$x_P = \frac{2(k_A + k_B)}{k_B - k_A}, \quad y_P = \frac{4k_A k_B}{k_B - k_A}$$

将 P 的坐标代入椭圆方程 $\frac{x^2}{4} + \frac{y^2}{3} = 1$, 即强加点 P 必须在椭圆上的约束.

$$\frac{1}{4}\left(\frac{2(k_A + k_B)}{k_B - k_A}\right)^2 + \frac{1}{3}\left(\frac{4k_A k_B}{k_B - k_A}\right)^2 = 1$$

两边同乘 $3(k_B - k_A)^2$ 并化简, 最终剩下

$$12k_A k_B + 16k_A^2 k_B^2 = 0 \quad \Longrightarrow \quad 4k_A k_B(3 + 4k_A k_B) = 0$$

由于 P 不在 x 轴上, $y_P \neq 0$, 所以 k_A, k_B 均不为零. 因此必有

$$3 + 4k_A k_B = 0 \quad \Longrightarrow \quad k_A k_B = -\frac{3}{4}$$

这一推导说明结论的由来. 在复杂的 (x,y) 坐标系下求解绕远, 切换到约束条件极其简洁的 (k_A, k_B) 斜率坐标系里更直接.

整个过程沿同一条策略展开. 先识别出用斜率 (k_{AP}, k_{BP}) 描述问题可能更简单; 再把 P 在椭圆上这一条件翻译为斜率坐标系下的简洁关系 $k_{AP} \cdot k_{BP} = -3/4$; 最后将目标函数也用斜率表达, 并在新变量体系下结合几何约束 (P 在 x 轴上方) 求出最值. 此题的要领是**借助变量代换选择有利的表达, 使约束尽可能简单**.

再看一个代数形式的例子, 策略仍然是同一条.

例题 19.25 (和积换元求范围) 设实数 x, y 满足 $x^2 + xy + y^2 = 3$, 求 $x^2 + y^2$ 的取值范围.

解 先沿解题流水线过一遍. 对象是实数对 (x,y), 共 2 个自由度; 唯一的条件 $x^2 + xy + y^2 = 3$ 是满足式, 化为 1 个方程, $m = 1 < n = 2$, 欠定. 于是 (x,y) 在一条曲线上变动, 求目标式的范围正是与欠定形态相配的问题 (第19.3.2小节).

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

直接用 (x,y) 处理并不顺手, 约束与目标都是二次式, 消元会引入根式. 约束与目标关于 x,y 都是对称的, 改用和与积命名的新变量.

$$s=x+y,\qquad p=xy.$$

由 $x^2+xy+y^2=s^2-p$, 约束化为 $s^2-p=3$, 即 $p=s^2-3$.

目标随之改写为 $x^2+y^2=s^2-2p=s^2-2(s^2-3)=6-s^2$.

还需刻画像集, 否则翻译不完备. 实数对 (x,y) 经和积换元后只落在 (s,p) 平面的一个子集里, x,y 是方程 $t^2-st+p=0$ 的两个实根, 当且仅当判别式非负.

$$s^2-4p\geq 0 \implies s^2-4(s^2-3)\geq 0 \implies s^2\leq 4.$$

这一步绝不能省. 若漏掉它, s^2 将无上界, 便会错误地断言 $6-s^2$ 没有最小值.

在新变量体系下, 由 $0\leq s^2\leq 4$ 得

$$2\leq 6-s^2\leq 6.$$

端点均可取到. $s=\pm 2$ 时 $x=y=\pm 1$; $s=0$ 时 $x=-y=\pm\sqrt{3}$, 均满足原约束.

因此 x^2+y^2 的取值范围是 $[2,6]$.

椭圆题把点在曲线上翻译成斜率之间的关系, 本题把同一个坐标约束翻译成和积之间的关系, **处理满足式条件的通用手法是把它翻译成更好用的变量语言**, 在新语言里目标式也随之变得简单. 此外, (x,y) 与 (y,x) 被映到同一对 (s,p), 而目标 x^2+y^2 对此并不敏感, 这个细节正是 第19.11节 要剖析的多对一映射与不变量的雏形.

19.8 不变量、单调量、极端原理、抽屉原理与对称

第19.6节 的反证法处理存在性判定. 另有一大类问题涉及**过程**与**操作**, 棋盘上移动棋子、数列按规则递推、算法逐步执行. 这类问题的复杂度随步数膨胀, 控制它的手段有四种: 在变化中寻找**不变**或**单调**的量, 在有限对象中取**极端**元素, 用**抽屉原理**从计数中读出存在性. 这四件工具之外再补一件适用范围更广的, 即在结构里寻找**对称**.

19.8.1 用不变量排除不可达状态

不变量是操作过程中始终保持不变的量. 若目标状态与初始状态在某不变量上取值不同, 则目标状态不可达, 这是不变量最常见的用法, 常配反证法使用.

例题 19.26 (残缺棋盘的覆盖) 8×8 棋盘去掉左上角与右下角两格, 能否用 31 块 1×2 多米诺骨牌(每块恰好覆盖两格)铺满?

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

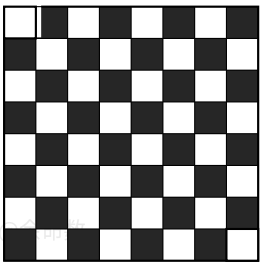


图 19.5　去掉两个同色对角格的棋盘, 黑白格数不等, 无法用多米诺铺满.

解　将棋盘黑白相间染色 (图 19.5). 完整棋盘有 32 黑 32 白, 而左上角与右下角是同色格, 去掉后剩下 30 格一色、32 格另一色. 每块 1×2 骨牌必覆盖一黑一白, 故任何铺法覆盖的黑格数等于白格数.$30\neq 32$, 铺满不可能.

这里的不变量是“已覆盖区域中黑格数与白格数之差”, 始终为 0; 目标区域该差为 ± 2, 不在同一不变量类, 故不可达.

例题 19.27 (函数对称性与不变量)　已知函数 $f(x)$ 对一切 x 满足 $f(x)+f(2-x)=4$. 证明 f 的图像关于点 $(1,2)$ 中心对称.

解　操作 $x\mapsto 2-x$ 是关于 $x=1$ 的反射. 在此操作下, 量 $f(x)+f(2-x)$ 始终等于 4, 这就是不变量.

令 $x=1$ 得 $2f(1)=4$, 故 $f(1)=2$. 对任意 x, 点 $(x,f(x))$ 在图像上, 其反射点 $(2-x,f(2-x))=(2-x,4-f(x))$ 也在图像上, 二者中点为 $(1,2)$. 因此图像关于 $(1,2)$ 中心对称.

第16.15节 中 2023 年全国甲卷题是用 $f(x)+f(2-x)=2a$ 这一不变量确定参数 a.

不变量也可用于判定两个状态是否属于同一操作可达的等价类. 若某量在一切允许操作下不变, 取值不同的状态互不可达. 奇偶性是最常见的不变量, 模某数的余数是它的推广.

19.8.1.0.1　用染色呈现不变量　例题 19.26 的解法把棋盘黑白相间染色, 让黑白格数相等这个不变量变得一目了然. 染色法是不变量原理的几何变体, 给棋盘格子、多边形顶点、格点这类离散对象按特定规律涂色或标数, 把抽象的不变量转化为颜色计数. 选取染色方案的关键, 是让每一步操作恰好覆盖各色一个, 于是颜色之差成为不变量; 目标状态的颜色之差一旦与之不同, 便不可达.

19.8.1.0.2　几何中的不变量　不变量不限于离散的染色. 给定一个圆与一个定点 P, 过 P 任作一条直线与圆交于 X 与 Y, 乘积 $PX\cdot PY$ 不随直线的选取改变, 这就是 第13.21节 的圆幂定理. 直线在动, 交点随之改变, 乘积却恒定. 相交弦定理与割

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

线定理都可以看作这一条不变量的具体形态, 区别只在 P 落在圆内还是圆外. 解题时先认出哪个量在变化中保持不变, 比逐条记忆定理更省力.

19.8.1.0.3 变形问题中的不变量 有些过程问题的状态空间很大, 逐态计算无从下手, 把注意力放到某个守恒量上, 问题立刻变短.

知乎@余命数

例题 19.28 (兑在一起的咖啡与牛奶) 瓶 A 装有一升牛奶, 瓶 B 装有一升咖啡. 从 B 中取一勺咖啡倒入 A, 搅匀后再从 A 中取同样容量的一勺混合液倒回 B. 此时 A 里的咖啡与 B 里的牛奶, 哪个更多?

解 逐次计算浓度需要跟踪两次混合, 式子很快变长. 改用守恒的眼光, 两个瓶子的容积自始至终都是一升.A 中的咖啡只能来自 B 少掉的那一部分, 而 B 少掉的咖啡由等量的牛奶补上, 因此 A 中的咖啡恰好等于 B 中的牛奶.

两者一样多, 这个结论与勺子的大小无关, 与有没有搅匀也无关. 只要两次取走的体积相同, 交换就是等量的.

19.8.1.0.4 只取两个值的不变量 有的过程里, 不变量只有奇偶两种取值. 这类二值信息可以当作暗号使用, 让一个人把看到的情形传给其他人.

知乎@余命数

例题 19.29 (黑白帽与奇偶传信) 64 人排成一列, 每人头上戴一顶黑帽或白帽, 帽色由旁人任意安排. 每人只能看见自己前面那些人的帽子. 游戏从队尾开始, 依次询问每人自己帽子的颜色, 被问者只能说“黑”或“白”两个字, 所有人都能听见这些回答. 什么方案能保证存活的人最多, 最多能保证多少人?

解 队列里的人按视野排成一条链, 队尾的人能看见前面 63 顶帽子, 队首的人什么都看不见. 若各人只顾自己猜, 后面的人得不到任何有用的信息, 只能碰运气.

关键是把队尾那一次发言变成一句公用的情报. 约定“黑”表示他所看到的黑帽个数为奇数, “白”表示黑帽个数为偶数. 队尾的人按看到的黑帽数奇偶发言, 他本人只有一半的可能恰好说中, 而在他前面的每个人都因此有了可以使用的信息.

第二个人看得见后面 62 顶帽子, 把自己看到的三者与队尾报出的两者相比. 两人所数黑帽只在第二个人的头上相差一顶, 两份奇偶相同就说明他是白帽, 不同就说明他是黑帽.

此后每个人照同样的办法办理. 第 i 个人已经听到前面 $i-2$ 个正确的回答, 又看得见后面所有人的帽子, 只有自己的帽色未知. 他把已知的黑帽个数从队尾报出的总数奇偶中减掉, 剩下的奇偶就决定自己的帽色.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

最多能保证 63 人存活. 队尾那个人的回答被用作情报, 他无法保证自己也说对, 其余 63 人全部获救. 整个方案只用了一个取值为奇偶两态的不变量, 却把每个人的可见信息接成了一条链.

19.8.2 用单调量证明过程终止

知乎@余命数

单调量是操作过程中严格单调变化且有界的量. 若某量每步严格递减且下有界,如非负整数,则过程必在有限步终止. 这是证明算法终止的标准手段. 单调量的形态不限于逐次减一的计数, 两个都在变的量相除, 比值常常就是那个单调的量.

例题 19.30 (欧几里得算法的终止性) 说明欧几里得算法(辗转相除法)对任意正整数输入必在有限步终止.

解 算法从正整数对 (a,b) 出发, 反复用 $(a,b)\to(b,a \bmod b)$, 直到余数为 0. 考察第二个分量 b 的序列. 每一步中新的第二分量 $a \bmod b$ 满足 $0 \le a \bmod b < b$. 只要余数非零, 第二分量就是非负整数且严格递减. 非负整数列不能无限严格递减, 故有限步内余数必为 0, 算法终止.

这里的单调量是当前数对中较小者的序列, 严格递减的非负整数保证了终止. 同样的论证适用于任何状态空间有限、每步状态严格按某序下降的过程.

另一类用法出现在计数过程里. 两个量各自都在变化, 看不出规律, 把二者相除, 比值却每一步减少一个定值.

例题 19.31 (病毒与细菌) 某容器在 $t=0$ 时放入 2009 个细菌与 1 个病毒. 每分钟每个病毒先杀死 1 个细菌, 随后所有存活的细菌与病毒各自分裂为 2 个. 细菌会不会灭绝? 若会, 在什么时候?

解 设某一时刻细菌数为 b, 病毒数为 v. 一分钟后每个病毒先杀死一个细菌, 再连同细菌一起分裂加倍, 于是

知乎@余命数

$$b' = 2(b-v), \qquad v' = 2v.$$

单独观察 b 与 v, 两者都在变化, 规律并不明显. 把二者相除, 规律立刻出现

$$\frac{b'}{v'} = \frac{2b-2v}{2v} = \frac{b}{v} - 1.$$

知乎@余命数

比值 $\dfrac{b}{v}$ 每分钟恰好减少 1, 是一个单调递减的量. 初始比值为 $\dfrac{2009}{1} = 2009$, 每过一分钟减 1, 经过 2009 分钟减到 0, 细菌全部消失.

知乎@余命数

整个过程不需要解递推式, 只要盯住这一个比值. 题面通常不直接给出单调量, 它需要从几个已知的量组合出来.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 19.32 (迭代数列的收敛) 已知 $a_1=2$, $a_{n+1}=\sqrt{a_n+2}$. 证明 $\{a_n\}$ 收敛并求其极限.

解 先证 $a_n\le 2$ 对一切 n 成立:$a_1=2$;若 $a_n\le 2$, 则 $a_{n+1}=\sqrt{a_n+2}\le\sqrt{4}=2$.

再证 $a_{n+1}\ge a_n$:$a_{n+1}^2-a_n^2=(a_n+2)-a_n^2=-(a_n-2)(a_n+1)\ge 0$(因 $a_n\le 2$).

$\{a_n\}$ 单调递增且上有界,故收敛. 设极限为 L,由 $L=\sqrt{L+2}$ 得 $L^2-L-2=0$, 正根 $L=2$.

这里的单调量是 a_n 本身, 严格递增且有上界, 保证了数列逐步逼近极限值.

例题 19.33 (比利时华夫饼) 6×8 格的华夫饼, 两人轮流沿格线切开, 每一刀只沿一条完整的格线把其中一块一分为二, 轮到切不动的人输. 先手有没有必胜策略?

解 局面分析看起来要枚举大量切法, 实际上根本不存在策略问题. 每切一刀, 饼的块数增加 1. 切到最后, 每一小格都彼此分开, 共有 $6\times 8=48$ 块, 所以无论两人怎么切, 总刀数都等于 47.

胜负只取决于最后一刀归谁.47 是奇数, 先手走的是第 $1,3,\dots,47$ 刀, 最后一刀由先手切出, 先手必胜. 单调量不必交代每一步的细节, 它只说明总步数固定, 胜负随之确定.

单调量证明过程必终止, 并能给出步数上界. 若量从 N 起每步至少减 1 且不降至负, 则至多 N 步终止.

19.8.3 用极端原理取极值元素

面对存在某对象满足某性质的命题, 取极端元素, 即最大、最小、最近、最远的那个, 让极端性自动带来所需性质. 极端原理把存在性证明转化为取极值 + 用极值性质导出结论.

例题 19.34 (整数的素因子分解) 证明每个大于 1 的整数都能写成素数之积(素数本身视为一积).

解 假设存在大于 1 的整数不能写成素数之积, 由极端原理, 自然数集的非空子集必有最小元, 取其中最小者 n.n 不是素数(否则它是自己的一积), 故 $n=ab$ 且 $1<a,b<n$. 由 n 的最小性, a 与 b 都能写成素数之积, 二者之积即 n 的素数之积, 与 n 的取法矛盾. 故不存在这样的 n.

知乎@余命数

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

最小反例是极端原理的典型形态. 取最小的坏对象, 用更小的对象已被处理来迫使坏对象其实不坏. 它与反证法天然配合, 假设坏对象存在, 取最小坏对象, 导出矛盾.

极端原理不止于数论. 在几何中, 取距离最短的(点, 直线)对, 同样能用极端性导出结论.

例题 19.35 (西尔维斯特问题) 平面上有限个点不全共线. 证明存在一条直线恰好经过其中两个点.

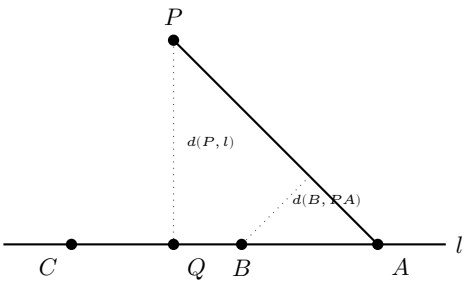


图 19.6 极端原理取距离最短的(点, 直线)对 (P,l), 导出 l 只含两个点.

解 全部共线的情形已排除, 故存在点 P 不在经过至少两个给定点的直线 l 上. 在所有这样的(点, 直线)对中, 取 $d(P,l)$ 最小的一对. 断言 l 恰好经过两个给定点.

设 l 上至少三个点, Q 为 P 到 l 的垂足.l 上的点分布在 Q 的两侧, 含 Q 本身, 总能取出同侧两点 A,B 使 B 比 A 离 Q 更近. 考虑直线 PA, P,A 都是给定点, 故 PA 经过至少两个点.B 不在 PA 上. 由 $\triangle PAB$ 面积的两种算法,

$$d(B,PA)=\frac{|AB|}{|PA|}\,d(P,l).$$

因 $|AB|<|PA|$ (直角 $\triangle PQA$ 中 $|PA|^2=|PQ|^2+|QA|^2>|QA|^2\geq|AB|^2$), 故 $d(B,PA)<d(P,l)$, 与 (P,l) 的最小性矛盾. 故 l 上恰好两个点.

极端原理有时以比较相邻对象的形式出现. 下面这道题里, 每个人朝离自己最近的人射击, 最近这一条是可以反复使用的极端性.

例题 19.36 (水枪) 平坦的场地上站着 n 个人 $(n>1)$, 每人到其他各人的距离互不相同. 一声令下, 每人朝离自己最近的人射击. 证明当 n 为奇数时至少有一个人不会中枪, 并说明 n 为偶数时结论不成立.

解 把每个人画成一个点, 朝谁射击就从谁那里引一条有向边出去, 得到一个每点出度都是 1 的有向图. 某人被射中一次或多次, 相当于相应点的入度不小于 1.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

先看 n 为偶数时的反例. 把人两两配成 $n/2$ 对, 同一对内两人相距 1 米, 不同对的人相距 100 米. 每个配对里, 两人的最近之人都是对方, 于是两人互射, 所有人都中枪.

再证奇数情形. 假设所有人都被射中, 则每个点的入度不小于 1. 全部点的出度之和等于 n, 入度之和也等于 n, 于是每个点的入度恰好是 1. 这时每个点恰有一条入边与一条出边, 图被分成若干个有向环.

环的长度不能超过 2. 若有一个长度为 $k \geq 3$ 的环 $v_1 \to v_2 \to \cdots \to v_k \to v_1$, 记 e_i 为 v_i 到 v_{i+1} 的距离(下标按环循环).v_i 射击的对象 v_{i+1} 是离它最近的人, 而 v_{i-1} 是另一个人, 故

$$e_i < e_{i-1}.$$

把这条不等式沿环走一圈, 得到

$$e_1 < e_k < e_{k-1} < \cdots < e_2 < e_1,$$

首尾是同一个数, 矛盾, 故不存在长度不小于 3 的环.

于是每个环都是长度为 2 的互射对. 点被两两配对, 总数 n 必为偶数, 与 n 为奇数矛盾. 因此奇数情形下至少有一人不中枪.

例题 19.37 (必要性探路与极端原理) 已知 $f(x) = e^x - ax$ 在 $[0, +\infty)$ 上单调递增, 求 a 的取值范围.

解 $f'(x) = e^x - a \geq 0$ 需在 $[0, +\infty)$ 上恒成立.e^x 在 $[0, +\infty)$ 上递增, 故 $f'(x)$ 的最小值在左端点 $x = 0$ 处取得: $f'(0) = 1 - a$.

取极端元素, 即 $f'(x)$ 的最小值, 得到必要条件 $1 - a \geq 0$, 即 $a \leq 1$. 当 $a \leq 1$ 时, $f'(x) = e^x - a \geq e^0 - a = 1 - a \geq 0$, 条件也充分.

第16.15节 的“必要性探路”是极端原理在高考导数题中的运用, 取边界点或极值点这一极端元素, 从中读出参数的必要条件.

19.8.4 用抽屉原理证明存在性

抽屉原理说, 把 $n + 1$ 个物体放入 n 个抽屉, 至少有一个抽屉里有两个或更多物体. 这条计数事实是组合数学中最强的存在性证明工具, 它从数量直接读出必然有重合, 无需构造具体的重合对象.

例题 19.38 (五整数取三, 其和为三的倍数) 任意给定 5 个整数, 证明必能从中选出 3 个使其和为 3 的倍数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 每个整数按模 3 的余数 $0,1,2$ 归入三类. 若某类至少 3 个, 取同类的 3 个, 其和 $\equiv 3r \equiv 0 \pmod 3$. 若每类至多 2 个, 则 $5 = 2+2+1$, 三类各非空, 从每类各取一个, 和 $\equiv 0+1+2 = 3 \equiv 0 \pmod 3$. 两种情形都成立.

知乎@余命数

例题 19.39 (数列周期性与抽屉原理) 已知数列 $\{a_n\}$ 满足 $a_1 = 1$, $a_{n+1} = 2a_n + 1 \pmod 7$. 证明 $\{a_n\}$ 是周期数列并求其周期.

解 a_n 只取 $\{0,1,\ldots,6\}$ 中的值, 共 7 个可能. 前 8 项中必有两项相等: $a_i = a_j$ $(i < j \le 8)$. 递推关系确定, 从 a_i 起的序列与从 a_j 起的序列完全相同, 故周期为 $j-i$.

直接计算: $a_1 = 1, a_2 = 3, a_3 = 0, a_4 = 1 = a_1$, 周期为 3.

抽屉原理保证了周期性的存在, 7 个抽屉装 8 个数, 必有重合, 而递推的确定性把重合转化为周期. 第7.14节 对此有一般性论述.

例题 19.40 (六人相识或互不相识) 任取 6 个人, 证明其中总能找到 3 个人两两相识, 或者找到 3 个人两两不识.

解 把 6 个人画成 6 个点, 两人相识就把相应两点连成红边, 不识就连成蓝边, 得到一个完全图 K_6 的红蓝二染色. 命题等价于, 任意红蓝二染色之下都存在一个单色三角形.

任取一个顶点 P, 从它引出 5 条边, 只有红蓝两种颜色, 由抽屉原理, 把 5 条边放进两个抽屉, 至少有一个抽屉装着 3 条或更多. 设 P 到 A,B,C 的三条边都是红色. 若 A,B,C 之间存在一条红边, 它与 P 构成红色三角形; 若 A,B,C 之间没有红边, 则三条边全为蓝色, A,B,C 自身构成蓝色三角形. 两种情形都出现单色三角形.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

6 是最小的数. 把 K_5 的边染成红蓝两色, 让红边构成一个五边形, 蓝边构成剩下的五条对角线, 逐一检查全部十个三角形, 没有一个是单色的. 记 $R(3,3) = 6$. 这类只保证模式存在而不管具体位置的结论属于拉姆齐理论, 它的入口就是抽屉原理.

构造抽屉的关键, 是把“存在性”翻译成“某类必有两个以上”. 模某数的余数是最常用的抽屉; 第7.14节 是用模 m 的 m^k 个状态作抽屉, 证明了整数递推数列的最终周期性. 第19.6节 中“5 点放入边长 2 的正方形”一题, 把正方形等分为 4 个小正方形作抽屉, 也是同一手法.

19.8.5 用对称导出结论

对称是比前面几件工具更普遍的观念. 一个对象在某个变换下保持不变, 就称它关于这个变换对称. 图形绕中心旋转半圈仍与原图重合, 多项式对调两个字母形式不变, 都是对称的具体形态. 对称有两种用法. 题面里已经含有对称时, 把

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

它直接读出来就能省掉大半计算；题面里看不出对称时，可以主动改造条件，造出一份对称来.

例题 19.41 (倒数方程) 解方程 $x^4-4x^3+5x^2-4x+1=0$.

解 方程左端的系数正着读与倒着读相同，这类方程称为倒数方程. 直接求根要处理四次式，先把这份对称用出来. 注意 $x=0$ 不是方程的根，两边同除以 x^2，再按 x 的次数从高到低整理，得

$$x^2-4x+5-\frac{4}{x}+\frac{1}{x^2}=0.$$

x^2 与 $\dfrac{1}{x^2}$ 成对出现，x 与 $\dfrac{1}{x}$ 也成对出现，引入新变量 $y=x+\dfrac{1}{x}$，此时

$$x^2+\frac{1}{x^2}=\left(x+\frac{1}{x}\right)^2-2=y^2-2.$$

原方程化为 $y^2-2-4y+5=0$，即 $y^2-4y+3=0$，解得 $y=1$ 或 $y=3$.

换元把四次方程降成了二次，但 y 还不是答案. 对实数 x，由 $x+\dfrac{1}{x}=y$ 得 $x^2-yx+1=0$，判别式 y^2-4 必须非负，所以 $y=1$ 不合，舍去. 对 $y=3$，解 $x^2-3x+1=0$ 得

$$x=\frac{3\pm\sqrt{5}}{2}.$$

原方程共有四个根，实根恰好是这两个，它们互为倒数，另外两个是复根.

除以 x^2 与换元 $y=x+\dfrac{1}{x}$ 都是围绕 x 与 $\dfrac{1}{x}$ 的对换造出来的对称. 换元之后还要用判别式核对新变量的允许范围，这一步不能省，漏掉它就会把 $y=1$ 也算进答案.

例题 19.42 (四只虫子) 正方形的四个顶点各有一只虫子，A 追 B，B 追 C，C 追 D，D 追 A，每只都以 1 米每秒的速率朝自己目标当前的位置前进. 求四只虫子相遇所需的时间.

解 逐只跟踪轨迹要把位置写成时间的函数，方程很快变得复杂. 换一个着眼点，只看相邻两只虫子的距离.

四只虫子的初始位置与运动规则在绕中心旋转 90° 的操作下完全不变，只是四只虫子彼此换了名字. 这份整体对称保证四条轨迹在任一时刻仍是一个正方形，各人的速度方向也与相邻两人的连线垂直. 于是 A 朝 B 的全速都用在缩短 AB 上，B 的运动方向与 AB 垂直，不改变 AB 的长度.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

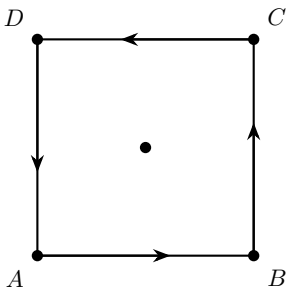


图 19.7 四只虫子沿各自的目标方向前进, 四条轨迹保持四重旋转对称, 正方形在收缩的同时整体转过相同的角度.

知乎@余命数

AB 以速率 1 米每秒缩短, 初始长度就是正方形的边长 a 米, 相遇需要 a 秒. 若正方形边长为 1 米, 答案是 1 秒. 虫子们的坐标都在变化, 正方形的形状作为整体始终不变, 这份整体对称把逐点追踪的问题约化成一条直线上的追及.

相遇点在正方形中心. 相遇之前, 四只虫子的位置始终构成一个绕中心旋转并不断收缩的正方形, 中心是唯一在整场运动中不动的点.

例题 19.43 (储物柜) 1000 个储物柜排成一排, 编号 1 到 1000, 起初全部关着. 第一个人把编号为 2 的倍数的柜子全部打开, 第二个人改变编号为 3 的倍数的柜子的开合状态, 第三个人改变编号为 4 的倍数的柜子, 依次进行到没有柜子可改变为止. 最后哪些柜子是关着的?

解 直接跟踪一千个柜子的状态没有意义, 先把问题换成一句数论的话. 编号为 n 的柜子, 每有一个不小于 2 的因数 d, 它就在第 $d-1$ 个人经过时改变一次状态. 于是柜子 n 改变状态的次数等于 n 的正因数个数减 1. 与 第19.4节 的倒数第二步一样, 原问题被换成了一个更小也更清楚的问题.

只要比较因数个数的奇偶.n 的正因数可以按 d 与 $\dfrac{n}{d}$ 两两配对, 当 $d \neq \dfrac{n}{d}$ 时这一对是两个不同的数, 只有 $d = \dfrac{n}{d}$ 时它配不上对. 后者要求 n 是完全平方数. 所以 n 是完全平方数时因数个数为奇数, 改变状态的次数为偶数, 柜子保持关着; n 不是完全平方数时因数个数为偶数, 改变状态的次数为奇数, 柜子最终开着.

关着的柜子恰好是编号为完全平方数的那些, 即 $1^2, 2^2, \ldots, 31^2$ 这 31 个($31^2 = 961$, 而 $32^2 = 1024 > 1000$), 其余 969 个柜子都开着.

找对称也有次序可循. 代数式先看系数或形式的排布是否对称, 几何图形先看有没有在运动过程中保持不变的形状或变换, 计数对象先看能不能把自己的结构两两配对. 都不成立时, 再动手改造条件, 把题目改成对称的形状, 倒数方程里的除法与换元就是这样的改造.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

19.8.6 归纳法与良序原理

第7.1节 的数学归纳法证明命题 $P(n)$ 对所有正整数 n 成立, 先验证 $P(1)$, 再证 $P(k)\Rightarrow P(k+1)$. 极端原理中的最小反例, 是同一条逻辑的另一面.

良序原理断言, 自然数的任一非空子集有最小元. 用它重读 例题 19.34. 假设存在不能分解为素数之积的整数, 取其中最小者 n, 良序原理保证了 n 的存在.n 不是素数, 故 $n=ab$ 且 $1<a,b<n$; 由 n 的最小性, a,b 都能分解, 合起来给出 n 的分解, 矛盾.

归纳法与最小反例是同一原理的两种表述. 归纳法说“没有反例”(从 1 起逐步排除); 最小反例说“若有反例, 取最小的, 导出更小的反例, 矛盾”. 两者等价, 都建立在良序原理之上. 面对涉及自然数的命题, 归纳法给出正面推进的路径, 最小反例给出反面排除的路径, 择优使用.

例题 19.44 (数列通项的猜想与归纳证明) 已知 $a_1=1$, $a_{n+1}=\dfrac{2a_n}{a_n+2}$. 猜想 a_n 的通项公式并用数学归纳法证明.

解 计算前几项: $a_1=1, a_2=\frac{2}{3}, a_3=\frac{1}{2}, a_4=\frac{2}{5}$. 观察规律: $a_n=\dfrac{2}{n+1}$.

基础步: $n=1$ 时, $\dfrac{2}{1+1}=1=a_1$, 成立.

归纳步: 设 $a_k=\dfrac{2}{k+1}$, 则

$$a_{k+1}=\frac{2a_k}{a_k+2}=\frac{2\cdot\frac{2}{k+1}}{\frac{2}{k+1}+2}=\frac{\frac{4}{k+1}}{\frac{2+2k+2}{k+1}}=\frac{4}{2k+4}=\frac{2}{k+2}.$$

故 $a_{k+1}=\frac{2}{(k+1)+1}$, 归纳完成.

从前几项猜想规律, 再用归纳法严格证明, 是高考数列题的标准路径. 归纳法把有限的观察推广为对所有自然数成立的结论.

19.8.7 这些原理如何组合

不变量、 单调量、 极端原理、 抽屉原理与对称常与 第19.6节 的反证法和第19.5节 的构造法组合使用, 形成完整的论证回路.

- **极端原理取极值元素**. 在有限或良序集合中取最大、最小、最近的元素, 为后续论证提供具有极端性质的对象, 如最小反例、最短点-线对.
- **不变量或单调量导出矛盾**. 用极值元素配合不变量或单调量, 推出与已知条件或极端性本身的矛盾, 从而否定假设.
- **抽屉原理读出存在性**. 从计数中直接得到 “某类必有两个以上”, 无需构造具体对象.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **对称化简计算**. 用保持不变的形状、变换或配对, 把待处理的对象减少到最小的一组.
- **构造法补充存在性**. 反证与极端只给出判定, 需要具体对象时再回到构造.

以 例题 19.34 为例. 极端原理取出最小反例 n, 反证法假设其存在, 组合后由 n 的分解把矛盾推回假设本身, 三者在同一道题里各司其职. 解操作类与过程类问题时, 依次检查四件事: 什么量不变, 什么量单调, 能否取极端, 能否构造抽屉; 再看题目里有没有现成的对称. 这五件事问过之后, 往往比直接枚举所有状态高效得多.

知乎@余命数

19.9 谈谈二级结论

解题时常会遇到二级结论. 看清题目结构, 才能用好它们. 有些结论功能明确, 比如用顶点坐标的行列式求三角形面积, 它是坐标确定面积这一构造过程的快速算法, 可自然整合进解题流程, 作为计算捷径.

另一些结论的作用则显得较为晦涩. 例如, 等比数列的前 n 项和 S_n 满足“$S_m, S_{2m}-S_m, S_{3m}-S_{2m}, \ldots$ 成等比数列”. 不理解它的来源, 生搬硬套反而扰乱对问题结构的整体把握.

知乎@余命数

许多高阶的二级结论, 都是某种**视角的转换**或**变量的代换**. 它们将原问题中对某个数学对象 A 的约束条件, 翻译成了对另一个相关对象 B 的、形式上更简洁的约束条件. 洞察这一翻译过程, 使用这些结论简化计算时才不至于失去对题目结构的把握.

例题 19.45 (等比数列的打包重组) 设正项等比数列 $\{a_n\}$ 的前 n 项和为 S_n, 已知 $a_5+a_6=16$, 且 $S_6=21$, 求 S_2.

此题若用基本量法 (设 a_1, q) 求解, 会面临求解高次方程的困境. 题目的条件 a_5+a_6 和 S_6 提示把数列的项**两两分组**. 这个洞察是解题的要点.

我们构造一个新数列来简化问题.

知乎@余命数

知乎@余命数

19.9.0.0.1　定义新变量, 翻译已知条件　将原数列 $\{a_n\}$ 的项两两打包, 定义新数列 $\{B_n\}$.

$$B_k = a_{2k-1}+a_{2k}$$

接着把原问题的所有信息用新数列的语言翻译.

- **已知条件一**.$a_5+a_6=16 \xrightarrow{\text{翻译}} B_3=16$.
- **已知条件二**.$S_6=(a_1+a_2)+(a_3+a_4)+(a_5+a_6)=21 \xrightarrow{\text{翻译}} B_1+B_2+B_3=21$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

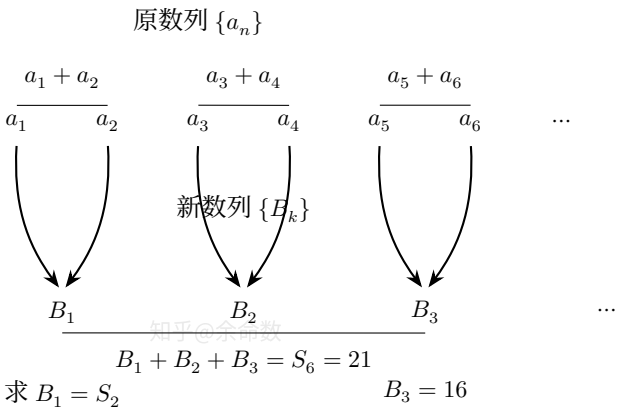


图 19.8 通过打包重组将数列 $\{a_n\}$ 映射为新数列 $\{B_k\}$

- **求解目标**.$S_2=a_1+a_2 \xrightarrow{\text{翻译}}$ 求B_1.

经翻译,原问题转化为一个关于新数列 $\{B_n\}$ 的问题. 已知数列 $\{B_n\}$ 满足 $B_3=16$ 且 $B_1+B_2+B_3=21$, 求 B_1. 仅凭这些条件尚不足, 还需翻译最重要的一个性质.

19.9.0.0.2 翻译关键性质, 说明新序列的规律 还未使用的关键条件是“$\{a_n\}$ 是等比数列”. 设 $\{a_n\}$ 的首项为 a_1, 公比为 q. 由于是正项数列, $a_1>0, q>0$. 考察 $\{B_n\}$ 的通项.

$$\begin{aligned} B_k &= a_{2k-1}+a_{2k}=a_1q^{2k-2}+a_1q^{2k-1} \\ &= a_1q^{2k-2}(1+q)=[a_1(1+q)]\cdot(q^2)^{k-1} \end{aligned}$$

这表明**新数列** $\{B_n\}$ **自身也是一个等比数列**,首项为 $B_1=a_1(1+q)$,公比为 $Q=q^2$. 由于 a_1, q 均为正, $\{B_n\}$ 也是正项数列. 打包重组的过程完整继承了等比的结构.

19.9.0.0.3 在新框架内解决问题 问题已经完全清晰. 已知一个正项等比数列 $\{B_n\}$, 其中 $B_3=16$ 且 $B_1+B_2+B_3=21$, 求 B_1. 由等比中项 $B_2^2=B_1B_3$ 及正项知 $B_2=4\sqrt{B_1}$, 代入和式

$$B_1+4\sqrt{B_1}+16=21$$

整理得关于 $\sqrt{B_1}$ 的一元二次方程

$$(\sqrt{B_1}-1)(\sqrt{B_1}+5)=0$$

$\sqrt{B_1}$ 必为正, 故 $\sqrt{B_1}+5\neq 0$、$\sqrt{B_1}=1$、$B_1=1$. 所求 $S_2=B_1=1$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

二级结论“$S_m, S_{2m}-S_m, S_{3m}-S_{2m}, \dots$ 成等比数列”就是刚才打包重组的直接结果. 完成这一推导, 问题得以解决, 也理清了**构造新序列以简化问题内在结构**的思路.

二级结论是高效的翻译工具, 能把问题从一个变量体系转换到另一个更利于求解的体系. 遇到一个二级结论时, 我们追问三件事.

知乎@余命数

- **它在转换什么?** 从坐标到斜率, 还是从单项到数组?
- **它在翻译什么条件?** 点在曲线上, 还是数列成等比?
- **这个翻译是双向的吗?** 这决定了代换的完备性.

19.10 二级结论的适用边界

第19.9节 把二级结论还原为视角转换与变量代换. 真正需要掌握的, 是**使用时机与适用边界**. 二级结论是对常见模式的凝练, 价值在于减少重复推导; 边界一旦模糊, 凝练就成了陷阱.

19.10.1 使用前的三项检验

知乎@余命数

套用一个二级结论之前, 须依次完成三项检验.

19.10.1.0.1　条件是否完全满足　每个二级结论都附带成立条件, 包括正负号、定义域、取等条件、几何位置. 结论的形式简洁, 条件却容易在题面里隐去. 条件缺一项, 结论就未必成立. 使用前先把条件逐条对照题目, 确认每一条都落到实处.

19.10.1.0.2　是否存在更直接的解法　二级结论是捷径, 捷径只有在原路更长时才有价值. 若直接展开、直接求导、直接代入比调用结论还省事, 套用结论反而引入额外负担. 用之前估一下直接解法的代价, 再决定是否走捷径.

知乎@余命数

19.10.1.0.3　结论能否现场重现　若一条结论能用一两行从基本定理重现, 其含义与条件也随之清晰; 若其推导本身需大段步骤, 则使用时更需留意其成立条件.

19.10.2 滥用的三种形态

例题 19.46 (对角线相等的误用)　平行四边形有结论“对角线平方和等于四边平方和”.

$$|\vec{a}+\vec{b}|^2+|\vec{a}-\vec{b}|^2=2(|\vec{a}|^2+|\vec{b}|^2).$$

有人据此在一般平行四边形中直接写对角线相等, 套用了矩形才有的性质.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 对角线平方和的恒等式对一切平行四边形成立, 但对角线相等 $|\vec{a}+\vec{b}|=|\vec{a}-\vec{b}|$ 等价于 $\vec{a}\cdot\vec{b}=0$, 即邻边垂直, 只对矩形成立. 把恒等式记成了相等关系, 等于偷偷加进了垂直这一条件. 在非矩形的菱形上使用, 结论失真. 几何二级结论的适用位置(矩形、菱形、一般平行四边形)必须与题目对象对齐.

例题 19.47 (均值不等式取等条件失效) 求 $f(x)=x+\dfrac{1}{x}(x\in(0,1))$ 的最小值. 有人用均值不等式 $x+\dfrac{1}{x}\geq 2$, 得最小值 2.

解 均值不等式 $x+\dfrac{1}{x}\geq 2$ 的取等条件是 $x=\dfrac{1}{x}$, 即 $x=1$. 但 $x\in(0,1)$, $x=1$ 不在定义域内, 等号取不到. 函数在 $(0,1)$ 上单调递减, 最小值在 $x\to 1^-$ 时趋近 2 但达不到. 结论的形式正确, 定义域把取等点排除在外, 机械套用就给出错误的最小值. 定义域、整数性、开闭区间, 都是结论条件的一部分.

例题 19.48 (简单题套重型结论) 估计 $\ln(1.2)$ 的范围. 有人调用对数均值不等式

$$\frac{x-y}{\ln x-\ln y}\leq\frac{x+y}{2}$$

代入 $x=1.2, y=1$ 求界.

解 对数均值不等式给出的是双向夹逼, 代入后还要解对 $\ln(1.2)$ 的不等式, 步骤不轻. 而题目只需粗估, 直接用 $\ln(1+x)\leq x(x>-1)$ 得 $\ln(1.2)\leq 0.2$, 一行即得所需上界. 重型结论在不需精细估计时反成负担. 精度需求与结论复杂度匹配, 才是用对二级结论.

19.10.3 理解推导, 把握场景

二级结论的可靠使用依赖两点. 其一, 清楚结论如何由基本定理推得, 既知其含义也知其条件; 其二, 清楚结论适用于哪类场景, 见到相应结构即可调用. 只记形式而不知来源, 一旦条件偏差便无从察觉; 只记公式而不知何处可用, 等于背了一串没有触发条件的符号.

用典型题时, 可把“这题把 X 翻译成了 Y”的表述连同其条件与方向一并记下, 这样的结论自然携带推导与适用场景.

19.11 换元法为何行之有效 *(选读)

无论把点的坐标转换为连线斜率, 还是把数列的单项组合为新序列, 用到的都是同一种**变量代换**, 我们称之为**换元法**. 换元法可靠而实用, 背后支撑它的数学

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

原理是**映射**. 理解了映射, 就明白为何可以在新变量的体系中解决问题, 并得到原问题的正确答案.

每一次成功的换元都建立了一个从**原变量空间**到**新变量空间**的**映射关系**.

19.11.0.0.1　换元法的基本流程　换元法把一个不易处理的问题通过映射 f 投射到结构更清晰、更易于分析的新空间. 这个过程分解为以下步骤.

1. **建立映射**. 定义新变量, 明确其与旧变量的函数关系, 即 $y = f(x)$, 由此建立从原变量空间到新变量空间的映射.
2. **转化约束**. 将原问题中对旧变量 x 的约束条件全部翻译成对新变量 y 的约束条件, 精确描绘原问题解集在映射下的**像集**. 之前推导二级结论或新变量关系的努力, 正是在完成这一步.
3. **转化目标**. 将原问题的求解目标也用新变量 y 来表示.
4. **在新空间求解**. 在新变量的约束条件下, 完成对新目标的求解.

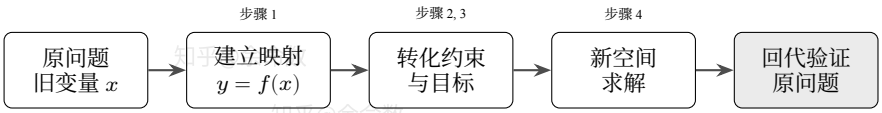


图 19.9　换元法的四步流程: 原问题经映射投到新空间, 在新空间求解后回代. 能否回代, 取决于映射的性质.

在新空间中得到的结果能否作为原问题的答案, 取决于映射 f 的性质, 以及所求解目标与这个映射的关系. 步骤 2 中的**像集**刻画尤为关键, 新变量的取值范围是旧变量经映射 f 后真正能到达的值的全体, 需要精确刻画. 像集刻画不完整, 等于在新空间里求解了一个比原问题更大的问题, 得到的解可能对原问题无效.

19.11.1　保证换元有效的两种典型情形

19.11.1.0.1　一一映射　当建立的映射 f 在我们关心的定义域和值域上是**一一对应**的, 即原变量空间中的每一个解都与新变量空间中的一个解恰好配对, 不多不少.

在这种情况下, 原问题和新问题在结构上**完全等价**. 它们就像同一本书的中文版和英文版, 语言(变量)不同, 讲述的故事在数学结构上完全一致. 在新变量空间中找到的任何解, 如最大值、最小值、特定值, 都唯一地对应着原变量空间中的一个解.

19.11.1.0.1.1　回顾椭圆问题　在 例题 19.24 中, 我们关注上半椭圆上(除顶点外)的点 $P(x_P, y_P)$. 对于每一个这样的点, 都唯一对应一对斜率 (k_{AP}, k_{BP}), 满足 $k_{AP} > 0, k_{BP} < 0$ 且 $k_{AP}k_{BP} = -3/4$. 反之, 每一对满足这些条件的斜率也唯一确定一个上半椭圆上的点 P.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

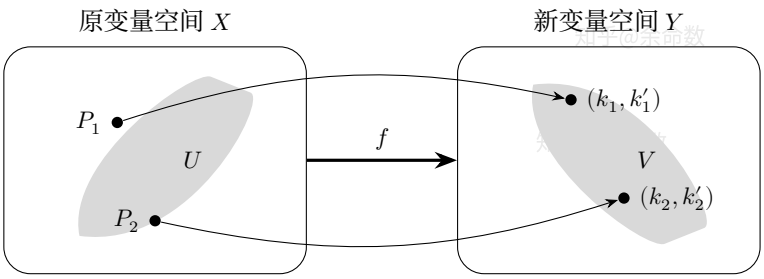


图 19.10 一一映射：原解集 $U \subseteq X$ 中的每个元素，如点 P 与像集 $V \subseteq Y$ 中的每个元素，如斜率对 (k, k') 一一对应. 问题在结构上完全等价.

$$P(x_P, y_P) \quad \overset{\text{一一对应}}{\longleftrightarrow} \quad (k_{AP}, k_{BP})$$

因此，从点到斜率的映射是**一一映射**. 求解关于 k_{AP}, k_{BP} 的表达式 $k_{AP} - 2k_{BP}$ 的最小值，就完全等价于求解原问题. 在新空间中走到终点，就意味着在原空间中也到达了终点.

19.11.1.0.2 多对一映射与不变量 在更多情况下，我们建立的映射是**多对**一的，即原变量空间中多个不同的对象被映射到新变量空间的同一个对象上. 此时，换元法是否依然有效，取决于一个关键条件. **所求解的目标必须是该映射下的一个不变量**.

不变量是指，所有被映射到同一个像的那些原像，在所关心的属性上完全一致. 所求的量其值只依赖于新变量的像，而与它究竟来自哪个原像无关.

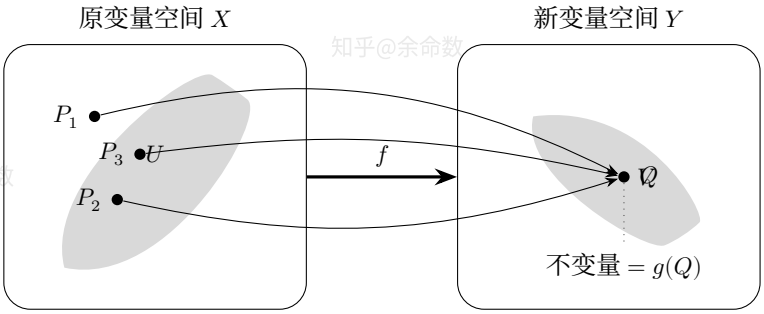


图 19.11 多对一映射：P_1, P_2, P_3 三个原像经 f 映到同一个像 Q. 若所求目标在 P_1, P_2, P_3 上取值相同，即不变量，则只需在 Q 处求值.

图 19.11 与 图 19.10 的对比，概括了换元法的关键判断. 一一映射下，所有问题都能回代；多对一映射下，只有不变量能回代. 动手换元之前先问清两件事：建立的映射是几对几，所求的量是否在合并的那些原像上取值一致.

19.11.1.0.2.1 回顾数列问题 将数列 $\{a_n\}$ 映射到新数列 $\{B_k\}$，其中 $B_k = a_{2k-1} + a_{2k}$. 这个映射是多对一的，不同的首项与公比组合 (a_1, q) 可能产生相同的新序列 $\{B_n\}$.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

$$\{a_n\}^{(1)},\{a_n\}^{(2)},\ldots \xrightarrow{\text{多对一映射}} \{B_n\}$$

然而,要求解的目标是 S_2. 换元时,我们已将 S_2 与新变量直接挂钩.$S_2=a_1+a_2=B_1$. 无论原空间中有多少个不同的数列 $\{a_n\}$ 能映射到同一个 $\{B_k\}$,它们的前两项和 S_2 必定都等于这个像的第一项 B_1.S_2 的值是这个映射的**不变量**.

因此可以在新空间中放心求解 B_1. 一旦求出 $B_1=1$,就知道所有满足题设条件的原数列 $\{a_n\}$,其前两项和 S_2 必然等于 1.

19.11.1.0.2.2　倒数换元

例题 19.49 ($x+\frac{1}{x}$ 型求值) 已知 $x+\dfrac{1}{x}=3$,求 $x^3+\dfrac{1}{x^3}$ 的值.

解 映射 $x\mapsto t=x+\dfrac{1}{x}$ 是典型的二对一映射.x 与 $\dfrac{1}{x}$ 被映到同一个 t(它们是方程 $u^2-tu+1=0$ 的两个根). 由 $t=3$ 反解 x,会得到 $x^2-3x+1=0$ 的两个互为倒数的根,原变量并未被唯一确定.

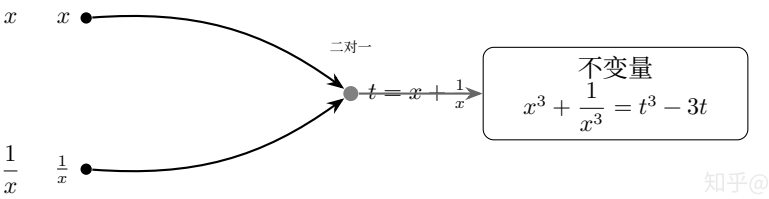


图 19.12　x 与 $1/x$ 映到同一个 t,所求 x^3+1/x^3 在两者上取值相同,是该映射的不变量.

但所求目标在原像对换 $x\leftrightarrow\dfrac{1}{x}$ 下纹丝不动,且可以只用 t 表达.

$$x^3+\frac{1}{x^3}=\left(x+\frac{1}{x}\right)^3-3\left(x+\frac{1}{x}\right)=t^3-3t.$$

代入 $t=3$,得 $27-9=18$.

目标是该映射下的不变量,故换元有效,甚至不必知道 x 究竟是哪一个根.

此例还能把不变量的边界照得更清楚. 若把目标换成 $x-\dfrac{1}{x}$,它在 $x\leftrightarrow\dfrac{1}{x}$ 下变号,不是不变量;此时仅凭 $t=3$ 无法确定其值(两个原像分别给出 $\pm\sqrt{5}$). **换元是否有效,答案随目标而变**,这正是动手之前就要问清的问题.

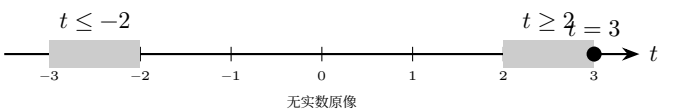


图 19.13　$t=x+1/x$ 的像集为 $|t|\geq 2$(阴影区域).$t=3$ 在像集内,换元有效;若题设给 $t=1$,落在空白区域,方程 $x^2-x+1=0$ 无实根,换元失效.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

像集的刻画在此例中同样是要务. 实数 x 经 $t=x+\dfrac{1}{x}$ 所得的像集是 $|t|\geq 2$(图 19.13);若题设误给 $t=1$, 方程 $x^2-x+1=0$ 无实根, 一切关于实数 x 的推论都将落空. 这与 例题 19.25 中的完备性检查是同一条要求, 都要求新变量的取值范围恰好是原变量经映射能到达的全体.

换元法借助一次精心设计的映射, 把问题投到更易处理的空间, 图 19.9 画出了其中的四步.

换元之前, 逐条检验下面三件事.

1. **映射是否确定**. 新变量 $y=f(x)$ 与旧变量 x 之间的函数关系是否明确, 能否双向翻译约束与目标.
2. **像集是否完备**. 新变量的取值范围是否恰好是旧变量经映射后能到达的值的全体, 不多不少. 像集过宽, 会在新空间中引入原问题没有的解; 像集过窄, 会丢失原问题的解. 图 19.13 展示了 $t=x+1/x$ 的像集 $|t|\geq 2$, 题设值落在像集外则换元失效.
3. **目标是否可回代**. 一一映射下一切目标可回代; 多对一映射下, 只有映射的不变量可回代.例题 19.49 中 x^3+1/x^3 是不变量可回代, $x-1/x$ 不是不变量不可回代, 同一映射下目标不同, 结论就不同.

三条标准对应换元法的三个环节: 建立映射、转化约束、回代验证. 任何一环出问题, 换元就会给出错误答案. 养成换元前逐条检查的习惯, 比事后纠错省力得多.

19.12 换元法的典型模式

第19.11节 从映射机理上说明了换元为何有效. 落到操作层面, 换元有六种典型模式, 每种针对一类结构. 识别出题目属于哪种模式, 换元就有了切入点.

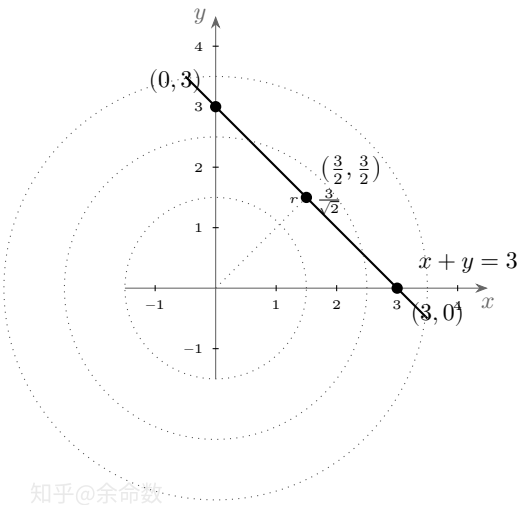
19.12.1 对称式换元

当条件与目标关于若干变量对称时, 改用和与积, 或其他对称的初等对称多项式作为新变量, 往往能把多变量问题降为单变量.

例题 19.50 (对称式换元) 已知实数 x,y 满足 $x+y=3$, 求 x^2+y^2 的最小值.

解 令 $s=x+y=3$, $p=xy$. 由 $x^2+y^2=s^2-2p=9-2p$, 最小化 x^2+y^2 即最大化 p. 实数 x,y 存在要求判别式 $s^2-4p\geq 0$, 即 $p\leq \dfrac{s^2}{4}=\dfrac{9}{4}$, 等号在 $x=y=\dfrac{3}{2}$ 取到. 故 $x^2+y^2\geq 9-2\cdot\dfrac{9}{4}=\dfrac{9}{2}$. 对称结构在 (s,p) 下变成单变量 p 的线性函数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



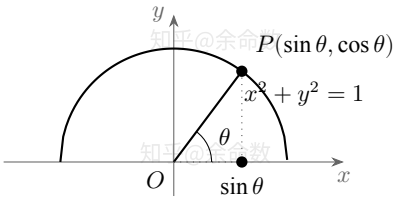
知乎@余命数

图 19.14　对称式换元, 直线 $x+y=3$ 上离原点最近的点是 $(\frac{3}{2},\frac{3}{2})$, 此处 x^2+y^2 取最小值 $\frac{9}{2}$. 同心圆表示 $x^2+y^2=r^2$ 的等值线, 与直线相切时 r 最小.

19.12.2 三角换元

当约束含 $\sqrt{a^2-x^2}$、$\sqrt{x^2+a^2}$ 或 $x^2+y^2=r^2$ 等结构时, 用三角恒等式 $\sin^2+\cos^2=1$ 或 $1+\tan^2=\sec^2$ 消去根号.

例题 19.51 (三角换元) 求 $f(x)=x\sqrt{1-x^2}(|x|\le 1)$ 的最大值.



令 $x=\sin\theta$, 则 $\sqrt{1-x^2}=\cos\theta$

图 19.15　三角换元, 令 $x=\sin\theta$ 把 $|x|\le 1$ 上的根式 $\sqrt{1-x^2}$ 化为 $\cos\theta$. 点 P 在单位圆上半部分, 横坐标即 x, 纵坐标即 $\sqrt{1-x^2}$.

解　令 $x=\sin\theta$ ($\theta\in[-\frac{\pi}{2},\frac{\pi}{2}]$ 保证可逆), 则 $\sqrt{1-x^2}=\cos\theta$, $f=\sin\theta\cos\theta=\dfrac{1}{2}\sin 2\theta\le\dfrac{1}{2}$, 等号在 $\theta=\dfrac{\pi}{4}$, 即 $x=\dfrac{\sqrt{2}}{2}$ 取到. 三角换元把根式化为整式, 最值一眼可读.

19.12.3 整体换元

当某个复合式在题中重复出现, 把它整体设为新变量, 可降低嵌套层级.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 19.52 (整体换元) 解方程 $(x^2-3x)^2-4(x^2-3x)-12=0$.

解 令 $t=x^2-3x$, 原方程化为 $t^2-4t-12=0$, 即 $(t-6)(t+2)=0$, 故 $t=6$ 或 $t=-2$. 回代.$x^2-3x=6 \implies x=\dfrac{3\pm\sqrt{33}}{2}$; $x^2-3x=-2 \implies x=1$ 或 $x=2$. 四个解均经检验成立. 重复出现的 x^2-3x 被整体替换后, 四次方程降为二次.

知乎@余命数

19.12.4 同构换元

当不等式或方程两端呈同构形式 $F(a)$ 与 $F(b)$ 时, 构造统一函数 F, 把比较两端归结为研究 F 的单调性.

例题 19.53 (同构换元) 设 $a>b>e$, 比较 a^b 与 b^a 的大小.

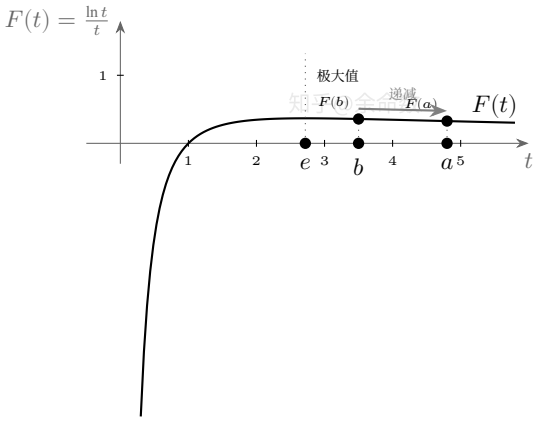


图 19.16 同构换元, 构造 $F(t)=\frac{\ln t}{t}$, 在 $t>e$ 处取极大值后严格递减.$a>b>e$ 时 $F(a)<F(b)$, 即 $\frac{\ln a}{a}<\frac{\ln b}{b}$, 故 $a^b<b^a$.

知乎@余命数

解 两边取对数, 比较 $b\ln a$ 与 $a\ln b$, 即比较 $\dfrac{\ln a}{a}$ 与 $\dfrac{\ln b}{b}$. 构造同构函数 $F(t)=\dfrac{\ln t}{t}$, 则 $F'(t)=\dfrac{1-\ln t}{t^2}$. 当 $t>e$ 时 $\ln t>1$、$F'(t)<0$、F 严格递减. 由 $a>b>e$ 得 $F(a)<F(b)$, 即 $\dfrac{\ln a}{a}<\dfrac{\ln b}{b}$, 亦即 $b\ln a<a\ln b$, 故 $a^b<b^a$. 两端的同构形式被统一函数 F 一并收纳, 比较化为单调性的直接应用.

19.12.5 参数化换元

当几何约束把对象限制在曲线上时, 用曲线的参数方程显式地把对象化为单参数, 把隐式约束化为参数的自然取值范围.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 19.54 (参数化换元) 椭圆 $\dfrac{x^2}{4}+y^2=1$ 上的点表为 $(2\cos t,\sin t)$, 求 $x+y$ 的取值范围.

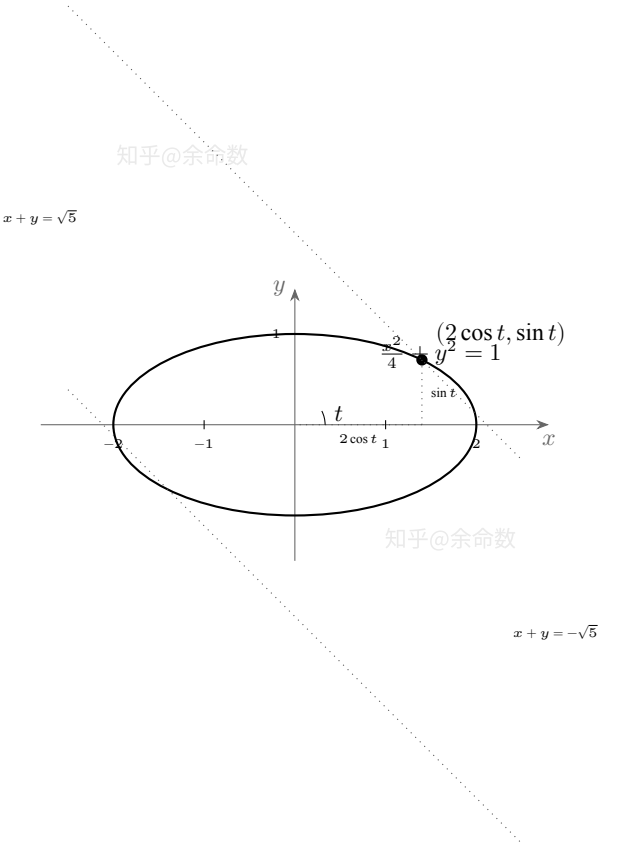


图 19.17　参数化换元, 椭圆上的点用参数 t 表示为 $(2\cos t,\sin t)$. 直线族 $x+y=c$ 与椭圆相切时取到极值 $c=\pm\sqrt{5}$, 对应 $x+y=\sqrt{5}\sin(t+\varphi)$ 的振幅.

解　$x+y=2\cos t+\sin t=\sqrt{5}\sin(t+\varphi)$, 其中 $\tan\varphi=2$. 由正弦的取值范围, $x+y\in[-\sqrt{5},\sqrt{5}]$, 两端都能取到. 参数化把点在椭圆上这一隐式约束变成单参数 t 的自由取值, 目标式也随之化为标准振幅型.

19.12.6 参考系换元

当题目里有两个以上的运动对象, 速度互相叠加时, 换一个参考系常常能让其中一个对象静止下来, 相对运动随之变成简单的追及或相遇.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 19.55 (漂流中的帽子) 一人沿河游泳, 逆流游了 30 分钟时发现帽子不见了, 帽子在他落水时就已掉入水中, 一直随水漂流. 他立即掉头顺流去追, 恰好在下游 1 千米处追上帽子. 设他在静水中的游速保持不变, 求水流速度.

解 以河岸为参考系, 人逆流时的速度是 $v-u$, 顺流时是 $v+u$, 帽子以 u 漂流. 三个速度纠缠在一起, 列方程解方程能把答案算出来, 计算量却不小.

改以河水为参考系, 帽子在这个参考系里始终静止不动. 人离开帽子的速率是静水游速 v, 追回来的速率也是 v, 来回路程相同, 所以去程与回程耗时相同. 去程用了 30 分钟, 回程也是 30 分钟, 从落水到追上总共 1 小时.

这一个小时里帽子相对河岸移动了 1 千米, 水流速度因此是 1 千米每小时. 答案与水速、游速的具体数值都无关, 只由漂流距离和总时间决定.

换参考系与 例题 19.42 中只盯住相邻两只虫子的距离, 做法是同一件事, 都是把不容易处理的绝对运动换成容易处理的相对运动.

19.12.7 换元后的三项检查

无论哪种模式, 换元落定后都要做三项检查.

新变量的范围. 新变量能取哪些值, 取决于原变量与映射的像集. 第19.11节 反复强调的完备性检查, 在这里就是确定新变量的定义域. 漏掉范围, 最值会错取在不可达的端点上, 见 例题 19.47.

可逆性. 若问题要求最终回到原变量, 如解方程要回代求原解, 换元映射必须在所关心范围内可逆; 多对一映射下, 一个新变量的值可能对应多个原变量的值, 回代时须全部找回, 不可只取其一.例题 19.49 中 x 与 $1/x$ 同映到一个 t, 若题目要的是 x 本身, 仅靠 t 就不够.

是否真正简化. 换元的目的是简化. 若新变量下的约束与目标反而更繁(根号换成三角却出现高次、整体换元后方程次数反升), 说明这次换元并不合适, 应换一种参数化或干脆回到原变量. 三角换元尤其要警惕. 消去根号的代价是引入三角恒等式, 只有当目标式能随之化为简洁振幅型或齐次型时才划算.

19.13 本章小结与习题

19.13.1 本章脉络

本章始终在讲同一件事, 即用条件去确定数学对象.

第19.1节 把任何题目拆成研究对象、已知条件、待求目标三要素, 区分显条件与隐条件、对象个体与对象族. 确定的程度决定问题的形态, 完全确定则求值, 部分确定则求范围、最值或轨迹.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第19.2节 用自由度计量这种确定. 完全确定一个对象所需的最少独立实数个数, 等于变量数减独立约束数. 设参数就是给自由度命名; 过剩的参数不带来过剩的信息. 第19.3节 把自由度提升为解题判断工具, 自由度为零对应离散解, 为正对应一族解, 为负提示约束过强需先验相容.

第19.4节 先给出动手调查的四项动作, 即列举、缩小规模、幻想答案、例证与猜想, 再把条件分成构造式与满足式, 给出正向推进、反向分析(含倒数第二步)、双向夹击、模式识别四条基本路径, 并组装成可动态调整的解题框架. 第19.5节 处理纯构造式问题, 沿构造链步步确定对象; 目标若能由已有对象直接组合, 便可越过中间对象, 这是设而不求的萌芽. 第19.6节 处理正面构造受阻时的反证法, 把结论的否定作为额外条件引入, 导出矛盾以判定存在性、唯一性或无穷性. 第19.7节 处理含满足式的问题, 把约束翻译成更好用的变量语言, 再在新语言中表达目标、求解; 翻译必须完备, 像集要刻画清楚.

第19.8节 用不变量排除不可达状态, 染色法是其视觉化; 用单调量证明过程终止; 用极端原理取极值元素, 西尔维斯特问题即其一例; 用抽屉原理从计数读出存在性; 用对称化简计算; 并以良序原理把归纳法与最小反例统一起来, 五者常与反证法、构造法配合.

第19.9节 把二级结论还原为视角转换与变量代换, 看清它在转换什么、翻译什么条件、是否双向, 才能驾驭它. 第19.10节 把二级结论从记忆推进到使用, 使用前检验条件是否齐备、有无更直接解法、能否当场推导; 记推导与场景胜过记结论形式.

第19.11节 从映射的角度说明换元为何有效, 一一映射时新旧问题等价, 多对一映射时目标须为不变量. 第19.12节 给出对称式、三角、整体、同构、参数化、参考系六种换元模式, 以及换元后核范围、查可逆、验简化的三项检查.

本章的工具都在回答两个层次的问题. 一个层次是结论为什么成立, 它解释结构的来源, 比如倒数方程同除以 x^2 之后四次式必然降为二次, 再如 例题 19.40 中抽屉原理为什么恰好给出同色三角形. 另一个层次是步骤怎么走通, 它保证书写与计算的正确, 比如倒数方程解出 y 之后还要回到 x, 六人问题染色之后还要把三条边中是否有红边的两种情形都讨论到.

19.13.2 与前文的联系

本章的概念在全书各处都有对应的技术化身, 回读时不妨对照.

映射的语言见 第2.2节, 它为 第19.11节 提供了底座. 为自由度命名的具体技术, 是 第13.4节 至 第13.7节 的设点法与设线法, 决策原则见 第13.16节. 欠定情形下的策略, 可回看 第16.15节 的必要性探路与端点分析. 用构造式条件锁定对象的典型例子, 包括 第7.9节 中“首项 + 递推确定数列”, 以及 第4.1节 中函数方程的赋值法. 第19.2节 的三次函数例所采用的展开形式, 其思想来自 第16.13节 (现并入第16章) 的泰勒公式. 极端原理与最小反例在 第9章 的整除与计数讨论中多次出

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

现. 第19.8节 的良序原理把最小反例与 第7.1节 的数学归纳法统一为同一原理的两种表述;抽屉原理在 第7.14节 证明了整数递推数列的最终周期性.

知乎@余命数

习题

知乎@余命数

19.1. (三要素辨析)对 例题 19.1 的三个问法, 分别写出研究对象、已知条件、待求目标, 并说明三者在哪一步开始分叉.
提示:对象与条件相同, 分叉发生在目标决定的方向上.

19.2. (自由度计数)确定下列对象各需几个自由度, 并说明理由. (1) 定圆的内接三角形, 三个顶点都在该圆上; (2) 椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 的一条切线; (3) 一个等差数列; (4) 抛物线 $y=ax^2+bx+c$.
提示:曲线上的一个动点只有 1 个自由度.

19.3. (条件分类与预判)对问题"已知等差数列 $\{a_n\}$ 满足 $a_1=2$, 且 a_3,a_5,a_8 成等比数列, 求其通项公式", 完成解题流水线的前三步:清点对象、标注自由度、给条件分类, 并预判解的形态.
提示:等差数列有 2 个自由度.

知乎@余命数 知乎@余命数

19.4. (三种形态)只考虑三角形的**形状**, 不计位置与朝向, 此时三角形有 3 个自由度, 判断下列各组条件确定三角形的形态(欠定 / 恰定 / 超定或不相容), 并说明理由. (1) 两角及一条边; (2) 满足三角不等式的三边; (3) 三个角; (4) 两边及其中一边的对角.
提示:(3) 确定的是相似一族; (4) 是恰定, 但离散解集可能有两解、一解或无解.

19.5. (框架实操)求过点 $P(2,3)$ 且与圆 $x^2+y^2=4$ 相切的直线方程.
提示:用斜率参数化后, 务必检验被参数化排除在外的竖直直线; 本题恰有一解藏在其中.

19.6. (构造次序)求圆 $x^2+y^2=9$ 与圆 $(x-5)^2+y^2=16$ 的公共弦所在直线的方程.
提示:不必先求交点;两方程之差即所求.

知乎@余命数

19.7. (反证法)用反证法证明 $\sqrt{3}$ 不是有理数.
提示:仿照 例题 19.20 中 $\sqrt{2}$ 的证明, 注意 $p^2=3q^2$ 推出 p 被 3 整除.

知乎@余命数

19.8. (反证法与单调量)证明方程 $x^5+x^3+x=0$ 在 $\mathbb{R}$ 上只有 $x=0$ 一个解.
提示:假设有两个不同实根, 用罗尔定理与 $f'(x)=5x^4+3x^2+1>0$ 导出矛盾.

19.9. (满足式与完备性)设实数 x,y 满足 $x^2-xy+y^2=1$, 求 x^2+y^2 的最大值与最小值.
提示:令 $s=x+y, p=xy$ 翻译约束与目标, 并用判别式刻画像集.

19.10. (不变量)8×8 棋盘去掉左上角与右下角两格, 两格同色, 能否用 31 块 1×2 多米诺骨牌铺满? 说明理由.
提示:黑白相间染色下, 骨牌必覆盖一黑一白;比较剩余的黑白格数.

知乎@余命数

19.11. (极端原理)证明每个大于 1 的整数要么是素数, 要么有不超过 $\sqrt{n}$ 的素因子.
提示:设 n 为合数, $n=ab, 1<a\le b$, 则 $a\le\sqrt{n}$; a 的素因子即所求.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

19.12. (极端原理·西尔维斯特问题) 平面上 5 个点不全共线, 用 例题 19.35 的方法证明存在一条直线恰好经过其中两个点.

提示: 取不在此直线上的一点与此直线的距离为最小的(点, 直线)对, 反证此直线上只有两个点.

19.13. (抽屉原理) 任取 6 个整数, 证明必有两个之差为 5 的倍数.

提示: 按模 5 的余数分 5 个抽屉, 6 个数必有两个同余.

19.14. (抽屉原理) 在边长为 1 的正三角形内任取 5 个点, 证明至少有两点距离不超过 $\frac{1}{2}$.

提示: 连接三边中点把正三角形分为 4 个边长 $\frac{1}{2}$ 的小正三角形, 每个内两点距离 $\leq \frac{1}{2}$.

19.15. (归纳法与最小反例) 用最小反例法证明: 对任意正整数 n, 3^n-1 能被 2 整除. 再与第7.1节 的数学归纳法证法对照, 体会两种表述的等价性.

提示: 假设存在不满足的正整数, 取最小者 n; $n\neq 1(3^1-1=2)$, 故 $n\geq 2$, $3^n-1=3\cdot 3^{n-1}-1=3(3^{n-1}-1)+2$.

19.16. (二级结论的翻译) 设等比数列 $\{a_n\}$ 的前 n 项和为 S_n, 已知 $S_3=2$, $S_6=6$, 求 S_9.

提示: 把每三项打包为一组, 新数列仍成等比.

19.17. (适用边界) 判断下列使用是否成立, 并指出错误所在. 求 $f(x)=x+\dfrac{4}{x}$ 在 $x\in(0,1]$ 上的最小值, 用均值不等式得 $f(x)\geq 4$, 故最小值为 4.

提示: 检查取等条件 $x=\dfrac{4}{x}$ 即 $x=2$ 是否在定义域内.

19.18. (映射观点) 已知 $x+\dfrac{1}{x}=-3$, 求 $x^5+\dfrac{1}{x^5}$ 的值.

提示: 记 $a_n=x^n+x^{-n}$, 则 $a_{n+1}=\left(x+\frac{1}{x}\right)a_n-a_{n-1}$; 目标是 $x\leftrightarrow\frac{1}{x}$ 下的不变量.

19.19. (不变量的边界) 考虑换元 $(s,p)=(x+y,\,xy)$. 判断下列目标能否仅由 s,p 确定, 并说明理由. (1) x^2+y^2; (2) x^3y+xy^3; (3) $x-y$.

提示: 检验目标在对换 $x\leftrightarrow y$ 下是否不变, 以及能否用 s,p 表出.

19.20. (换元模式识别) 指出下列各题应优先采用哪种换元模式, 并写出换元. (1) 求 $y=x+\sqrt{1-x^2}$ 的最大值; (2) 解 $x^4-5x^2+4=0$; (3) 已知 $x+y=4$, 求 xy 的最大值; (4) 比较 e^{π} 与 π^e 的大小.

提示: 依次对应三角换元、整体换元、对称式换元、同构换元.

19.21. (开放回看) 从 第13章 或 第16章 中任选一道你完整做过的大题, 画出它的解题流水线, 写清对象清单、自由度标注、条件分类, 以及 m 与 n 的比较. 无标准答案, 完成即正确.

19.22. (动手列举·变体) 三个正整数的乘积是 72, 它们之和是某人已知的门牌号. 该人知道门牌号后仍不能确定这三个数, 又被告知最大的数只有一个, 随即确定. 求这三个数与门牌号.

提示: 列出乘积为 72 的全部三元组, 重复出现的和数对应两组, 其中一组有两个数并列最大.

知乎@余命数

19.23. (例证与猜想)验证 $P(n)=n^2+n+41$ 在 $n=0,1,\ldots,39$ 时都取素数, 并说明 $n=40$ 时 $P(n)$ 是合数.

提示: $P(40)=40^2+40+41=41\times 41$.

19.24. (连续正整数之和)证明 1024 不能写成两个以上连续正整数之和, 并求 1 到 100 中不能这样写出的正整数.

提示: 由 例题 19.9 的结论, 不能写出的恰是 2 的幂.

19.25. (幻想答案)某书店出售单价 3 元与 5 元的两种练习册, 某人用 50 元恰好买了 12 本. 先用幻想答案法写一个答案再列方程检验, 求两种各买了多少本.

提示: 设 3 元本为 x 本, 则 $3x+5(12-x)=50$.

19.26. (不变量)一只杯子盛满水, 另一只盛满酒. 从水杯取一勺水倒入酒杯, 搅匀后再从酒杯取同容量的一勺混合液倒回水杯, 重复若干次. 证明水杯中的酒与酒杯中的水始终等量.

提示: 两次转移的容量相同, 交换就是等量的; 也可以用水的总量守恒来说明.

19.27. (奇偶传信·变体)帽子有红、黄、蓝三色, 每人只能说一个 0 到 2 的整数, 其余规则与例题 19.29 相同. 证明除队尾外所有人都能保证说对.

提示: 把三种颜色依次编号 $0,1,2$, 队尾报出他看到的所有帽子编号之和模 3 的余数.

19.28. (对称·倒数方程)求方程 $x^4-x^3+x^2-x+1=0$ 的实根个数.

提示: 两边除以 x^2, 令 $y=x+\dfrac{1}{x}$, 方程化为 $y^2-y-1=0$; 两个根都满足 $|y|<2$.

19.29. (参考系换元)自动扶梯以恒定速率上行. 某人以每秒 1 级的速度向上走, 到顶时数得 20 级; 以每秒 2 级的速度向上走, 数得 32 级. 求扶梯静止时可见的级数.

提示: 设扶梯每秒上行 s 级, 行走时间为 $\dfrac{N}{v+s}$, 数得的级数是 $v\cdot\dfrac{N}{v+s}$; 两式消去 N 得 $s=3$, 再回代求 N.

19.30. (单调量·必胜策略)把 例题 19.33 的华夫饼换成 $m\times n$ 格, 说明胜负与两人的切法无关, 只由 $mn-1$ 的奇偶决定.

提示: 每刀使块数增加 1, 终局的块数恒为 mn.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

第五卷 · 扩展与线性结构

中学数学之旅



本卷扩展对象与结构：复数、立体几何与矩阵行列式.

在几何与分析底座上打开数系、空间与线性变换的大门.

知乎 @ 余命数 (Midnight Forever)

Project Akari

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

本卷导读

本卷线索

知乎@余命数

第五卷在几何与分析的底座之上扩展研究对象，数系向复数延伸，几何从平面进入空间，矩阵与行列式则把线性结构纳入统一的处理框架. 末三章转入数论与常数，连分数给出最佳有理逼近的判据，e、π、γ 三个常数借级数与积分登场，无穷乘积把乘法版的收敛理论接到 ζ 函数与素数上.

章节地图

知乎@余命数

- ▷ **复数**　从代数形式与四则运算讲起，继而在复平面上给出几何意义，三角形式把复数乘法与旋转、角的加法统一起来，最后抵达代数基本定理.
- ▷ **立体几何**　以空间中直线与平面的位置关系为主干，从公理出发建立平行与垂直的判定和性质，再引入空间向量与坐标法，处理距离、角度与体积等度量问题.
- ▷ **矩阵与行列式**　沿运算、行列式、线性变换、逆矩阵与克拉默法则、秩与消元法逐节推进，最终以二次型统一圆锥曲线的分类.
- ▷ **连分数**　从辗转相除法的商序列出发，建立有限与无限连分数的表示与收敛理论，给出最佳有理逼近的判据，并以黄金比说明收敛最慢的极端情形.
- ▷ **常数 e、π 与 γ**　用级数与积分工具处理三个著名常数：e 的无理性、Niven 对 π 无理性的初等证明，以及欧拉常数 γ 的基本性质与仍未解决的问题，末节回到巴塞尔问题.
- ▷ **无穷乘积**　把级数收敛理论搬到无穷乘积，依次讨论 Wallis 乘积、Vieta 公式与欧拉乘积，让 ζ 函数与素数的分布进入视野.

知乎@余命数

知乎@余命数

向后展望

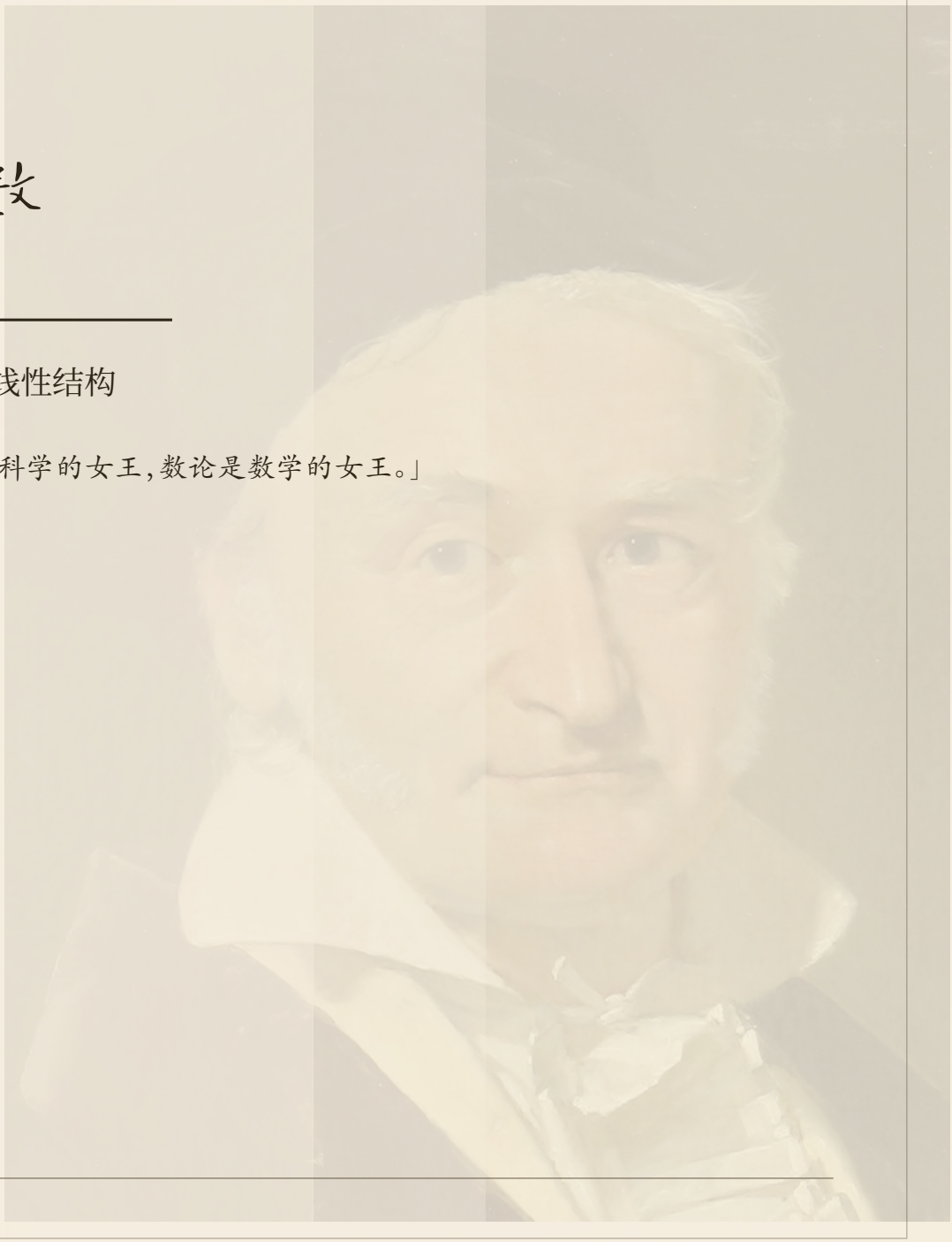
至此，函数、几何、分析与线性结构这些中学数学的主要对象连成一张可导航的地图，大学课程将在这些门口继续深入.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 20 章

复数



扩展与线性结构

「数学是科学的女王，数论是数学的女王。」

——高斯

第二十章 复数

数系的每一次扩充都由同一个问题驱动. 某种运算在原有体系内无解, 我们就把体系扩大, 使它有解. 从自然数 $\mathbb{N}$ 到整数 $\mathbb{Z}$, 是为了让减法封闭; 从整数 $\mathbb{Z}$ 到有理数 $\mathbb{Q}$, 是为了让除法封闭; 从有理数 $\mathbb{Q}$ 到实数 $\mathbb{R}$, 是为了让开方与极限封闭.

然而, 即便是实数集 $\mathbb{R}$ 中, 一个最基础的代数方程仍然说明了它的局限性.

$$x^2 = -1$$

在实数范围内, 任何数的平方都非负, 因此该方程无解. 这解释了求根公式在判别式 $\Delta < 0$ 时失效的原因, 也说明实数系在代数完备性上存在根本缺陷. 为了弥合这一缺陷, 数学家们必须探寻对数这一概念的再次延拓.

暂时搁置纯粹的代数求解, 从几何视角看数的运算. 在实数轴上, 乘以 -1 对应着一个明确的几何变换, 即以原点为中心旋转 180°. 例如, $2 \times (-1) = -2$, 点 $(2, 0)$ 旋转 180° 到达点 $(-2, 0)$. 一个根本问题由此产生, 是否存在一种运算, 当它连续作用两次后, 其效果恰好等同于“旋转 180°”, 又能否为“旋转 90°”这一几何变换, 寻找到一个对应的代数实体.

这个几何问题提供一条构造性的路径. 若要将代表 1 的点进行 90° 的旋转, 其像必然会脱离一维的实数轴, 而进入一个二维的平面. 我们不妨将这个“逆时针旋转 90°”的算子记为 i. 那么, 从 1 出发, 施以 i 的运算就意味着将点 $(1, 0)$ 旋转至新位置. 再施以一次 i 的运算, 即连续旋转两次 90°, 就等同于旋转 180°, 最终到达 -1 的位置.

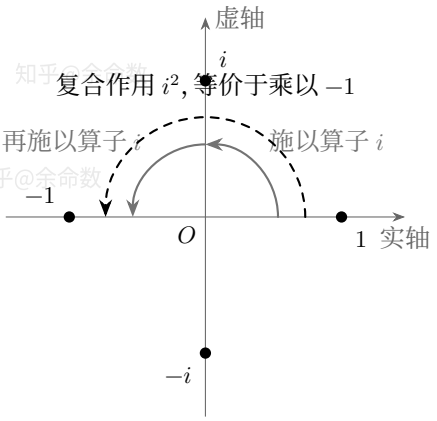


图 20.1 从代数到几何的统一: $i^2 = -1$

这一过程说明了代数结构与几何变换之间内在的对应关系. 我们为“逆时针旋转 90°”这一几何变换所寻找的代数算子 i, 其二次幂 (i^2) 的代数效果, 必然等同

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于乘以 -1. 这便导出了新数系的核心定义, 它同时满足了代数与几何的双重逻辑要求.

$$i^2 = -1$$

因此, 复数的概念是代数需求与几何直观内在统一的必然结果. 它将对数的认知从一维的数轴拓展至二维的平面.

知乎@余命数

20.1 复数的概念与几何意义

20.1.1 复数概念

定义 20.1 (复数) 我们定义形如 $z = a + bi$ 的数称为**复数**, 其中 a, b 均为实数 ($\mathbb{R}$), i 称为**虚数单位**, 它满足 $i^2 = -1$.

- a 称为复数 z 的**实部**, 记作 $\mathrm{Re}(z) = a$.
- b 称为复数 z 的**虚部**, 记作 $\mathrm{Im}(z) = b$.

所有复数构成的集合称为**复数集**, 记作 $\mathbb{C}$.

知乎@余命数

实部与虚部也可记作 $\mathrm{Re}(z) = a$, $\mathrm{Im}(z) = b$.

性质 (复数的分类) 设复数 $z = a + bi$, 得

- 当 $b = 0$ 时, $z = a$ 是一个**实数**.
- 当 $b \neq 0$ 时, z 称为**虚数**.
- 当 $a = 0$ 且 $b \neq 0$ 时, $z = bi$ 称为**纯虚数**.

两个复数相等, 当且仅当它们的实部和虚部分别相等.

$$a + bi = c + di \iff a = c \text{ 且 } b = d$$

知乎@余命数

证明 前三条分类直接来自虚部 b 是否为 0 以及实部 a 是否为 0 的定义. 对相等判定, 若 $a = c$ 且 $b = d$, 则显然 $a + bi = c + di$. 反过来, 若 $a + bi = c + di$, 则 $(a - c) + (b - d)i = 0 = 0 + 0i$. 由复数代数形式的唯一性, 得 $a - c = 0$ 且 $b - d = 0$, 即 $a = c$ 且 $b = d$. □

知乎@余命数

20.1.2 复数的几何表示

定义 20.2 (复平面) 建立了直角坐标系来表示复数的平面, 称为**复平面**或**阿尔冈图**.

- x 轴称为**实轴**.
- y 轴称为**虚轴**.

复数 $z=a+bi$ 与复平面上的点 $Z(a,b)$ 以及从原点指向点 Z 的向量 $\overrightarrow{OZ}$ 是一一对应的. 这种对应关系是数形结合思想在复数中的基石.

复平面给出复数的几何位置, 模与共轭则刻画其长度与对称关系.

定义 20.3 (模与共轭) 设复数 $z=a+bi$.

- **模**: 复数 z 对应的向量 $\overrightarrow{OZ}$ 的长度, 记作 $|z|$.

$$|z|=\sqrt{a^2+b^2}$$

- **共轭复数**: 实部相等, 虚部互为相反数的两个复数. z 的共轭复数记作 $\bar{z}$.

$$\bar{z}=a-bi$$

在复平面上, 共轭复数对应的点关于实轴对称.

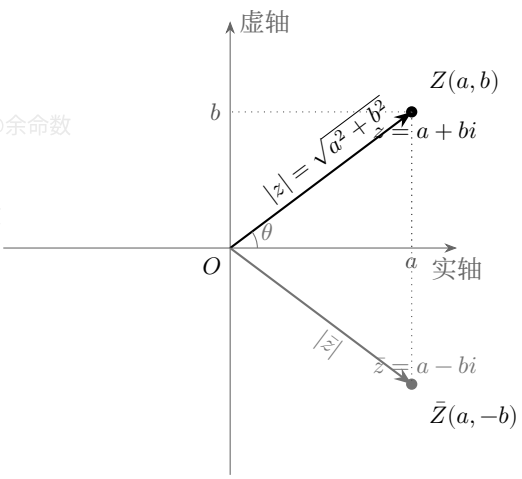


图 20.2 复数、模与共轭的几何意义

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

性质 (模与共轭的性质) 设 $z, z_1, z_2 \in \mathbb{C}$, 得

1. $z \cdot \bar{z} = |z|^2$. 此式连接代数运算与几何模长.
2. $|z_1 z_2| = |z_1||z_2|$, $\quad |\frac{z_1}{z_2}| = \frac{|z_1|}{|z_2|}$ $(z_2 \neq 0)$.
3. $\overline{z_1 \pm z_2} = \bar{z_1} \pm \bar{z_2}$. 共轭运算对加减法可分配.
4. $\overline{z_1 z_2} = \bar{z_1}\bar{z_2}$, $\quad \overline{(\frac{z_1}{z_2})} = \frac{\bar{z_1}}{\bar{z_2}}$ $(z_2 \neq 0)$. 共轭运算对乘除法亦可分配.

证明

1. 设 $z = a + bi$, 其中 $a, b \in \mathbb{R}$, 则 $\bar{z} = a - bi$. 展开得

$$\begin{aligned} z \cdot \bar{z} &= (a+bi)(a-bi) \\ &= a^2 - (bi)^2 && \text{(平方差公式)} \\ &= a^2 - b^2 i^2 \\ &= a^2 - b^2(-1) && (i^2 = -1) \\ &= a^2 + b^2 \end{aligned}$$

知乎@余命数

又 $|z| = \sqrt{a^2+b^2}$, 故 $|z|^2 = \left(\sqrt{a^2+b^2}\right)^2 = a^2+b^2$, 与上式比较即得 $z \cdot \bar{z} = |z|^2$.

2. 设 $z_1 = a + bi$, $z_2 = c + di$, 则 $z_1 z_2 = (ac - bd) + (ad + bc)i$, 进而

$$\begin{aligned} |z_1 z_2|^2 &= (ac-bd)^2 + (ad+bc)^2 \\ &= a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + 2abcd + b^2c^2 \\ &= (a^2+b^2)(c^2+d^2) = |z_1|^2|z_2|^2. \end{aligned}$$

复数的模为非负实数, $|z_1 z_2|$ 与 $|z_1||z_2|$ 均非负, 两边取非负平方根即得 $|z_1 z_2| = |z_1||z_2|$.

当 $z_2 \neq 0$ 时, 令 $w = \frac{z_1}{z_2}$, 则 $z_1 = wz_2$, 由已证乘法性质得 $|z_1| = |w||z_2|$. 又 $z_2 \neq 0$ 蕴含 $|z_2| > 0$, 因此 $\left|\frac{z_1}{z_2}\right| = |w| = \frac{|z_1|}{|z_2|}$.

3. 设 $z_1 = a + bi, z_2 = c + di$. 以加法为例, 比较等式两边

$$\overline{z_1 + z_2} = \overline{(a+c) + (b+d)i} = (a+c) - (b+d)i,$$

$$\bar{z_1} + \bar{z_2} = (a - bi) + (c - di) = (a+c) - (b+d)i,$$

知乎@余命数

两边相同, 故 $\overline{z_1 + z_2} = \bar{z_1} + \bar{z_2}$. 减法同理.

4. 设 $z_1 = a + bi, z_2 = c + di$. 比较乘法等式两边

知乎@余命数

$$\overline{z_1 z_2} = \overline{(ac-bd) + (ad+bc)i} = (ac-bd) - (ad+bc)i,$$

$$\bar{z_1}\cdot\bar{z_2}=(a-bi)(c-di)=ac-adi-bci+bdi^2=(ac-bd)-(ad+bc)i,$$

两边相同, 故 $\overline{z_1z_2}=\bar{z_1}\bar{z_2}$.

除法可由乘法结论转化. 令 $w=\frac{z_1}{z_2}$, 则 $z_1=w\cdot z_2$, 两边取共轭得 $\bar{z_1}=\overline{w\cdot z_2}$. 由已证的乘法性质, $\bar{z_1}=\bar{w}\cdot\bar{z_2}.z_2\neq 0$ 蕴含 $\bar{z_2}\neq 0$, 解得 $\bar{w}=\frac{\bar{z_1}}{\bar{z_2}}$. 将 $w=\frac{z_1}{z_2}$ 代回, 即 $\overline{\left(\frac{z_1}{z_2}\right)}=\frac{\bar{z_1}}{\bar{z_2}}$.

知乎@余命数

□

例题 20.1 已知复数 z,w 均不为 0, 则下列结论正确的是(　)

A. $z^2=|z|^2$

B. $\frac{z}{\bar{z}}=\frac{z^2}{|z|^2}$

C. $\overline{z-w}=\bar{z}-\bar{w}$

D. $\left|\frac{z}{w}\right|=\frac{|z|}{|w|}$

解 四个选项均围绕恒等式 $|z|^2=z\bar{z}$ 展开, 逐一判定.

选项 A 错误. 取 $z=i$, 左边 $z^2=i^2=-1$, 右边 $|z|^2=|i|^2=1$, 两边不等.

选项 B 正确. 由 $|z|^2=z\bar{z}$,

$$\frac{z}{\bar{z}}=\frac{z}{\bar{z}}\cdot\frac{z}{z}=\frac{z\cdot z}{\bar{z}\cdot z}=\frac{z^2}{z\bar{z}}=\frac{z^2}{|z|^2}.$$

选项 C 正确. 设 $z=a+bi$, $w=c+di$, 比较两边

$$\overline{z-w}=\overline{(a-c)+(b-d)i}=(a-c)-(b-d)i,$$

$$\bar{z}-\bar{w}=(a-bi)-(c-di)=(a-c)-(b-d)i,$$

两边相等.

选项 D 正确. 由 $|u|^2=u\bar{u}$ 与 $\overline{(u/v)}=\bar{u}/\bar{v}$,

$$\left|\frac{z}{w}\right|^2=\left(\frac{z}{w}\right)\cdot\overline{\left(\frac{z}{w}\right)}=\frac{z}{w}\cdot\frac{\bar{z}}{\bar{w}}=\frac{z\bar{z}}{w\bar{w}}=\frac{|z|^2}{|w|^2}.$$

模为非负数, 两边开方得 $\left|\frac{z}{w}\right|=\frac{|z|}{|w|}$.

故正确的选项为 B、C、D.

注 20.1 (经验总结)

四个选项围绕恒等式 $|z|^2=z\bar{z}$ 展开, 它连接复数 z、共轭 $\bar{z}$ 与模 $|z|$.

- B 通过构造 $z\bar{z}$ 得到分母 $|z|^2$; D 对模取平方后归约为同一恒等式.

- $|z|^2 = z\bar{z}$ 是处理复数恒等变换的常用入口.

20.2 复数的运算

知乎@余命数

20.2.1 代数运算

设 $z_1 = a + bi, z_2 = c + di$.

- **加法/减法**:

$$z_1 \pm z_2 = (a \pm c) + (b \pm d)i$$

几何意义:遵循向量加/减法的平行四边形/三角形法则.

- **乘法**:

$$z_1 \cdot z_2 = (a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

知乎@余命数

(像多项式一样展开, 然后用 $i^2 = -1$ 化简)

- **除法**:

$$\frac{z_1}{z_2} = \frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{ac + bd}{c^2 + d^2} + \frac{bc - ad}{c^2 + d^2}i$$

分子分母同乘以分母的共轭复数, 实现分母实数化.

知乎@余命数

例题 20.2 (复数四则运算) 计算 $z = \frac{(1-2i)^2}{2+i} + i^{2023}$.

解 按先乘方、后除法、再加法的顺序计算, 并利用 i 的周期性.

知乎@余命数

先化简分子 $(1 - 2i)^2$, 由平方展开得

$$(1 - 2i)^2 = 1^2 - 2(1)(2i) + (2i)^2 = 1 - 4i + 4i^2 = 1 - 4i - 4 = -3 - 4i$$

将其代入除法, 分子分母同乘分母的共轭 $2 - i$ 实现分母实数化:

$$\begin{aligned}
\frac{-3 - 4i}{2 + i} &= \frac{(-3 - 4i)(2 - i)}{(2 + i)(2 - i)} \\
&= \frac{-6 + 3i - 8i + 4i^2}{4 + 1} \qquad (i^2 = -1) \\
&= \frac{-6 - 5i - 4}{5} \\
&= \frac{-10 - 5i}{5} = -2 - i
\end{aligned}$$

知乎@余命数

再处理 i 的高次幂.i 的幂周期为 4,

$$i^1=i,\quad i^2=-1,\quad i^3=-i,\quad i^4=1$$

由 $2023=4\times505+3$ 得 $i^{2023}=i^3=-i$.

两部分相加得 $z=(-2-i)+(-i)=-2-2i$.

20.3 二次方程的复数解

复数系的代数框架已经建立,现在可以求解最初激发这一理论发展的问题,即实数域 $\mathbb{R}$ 内无解的二次方程.

20.3.0.1 负实数的平方根

实数系中, 形如 $x^2=-p\ (p>0)$ 的方程无解, 原因在于没有实数的平方为负数. 虚数单位 i 由 $i^2=-1$ 定义, 正为此提供了代数基础.

据此, 对任意正实数 p, $-p$ 的平方根可定义:

$$\sqrt{-p}=\sqrt{p\cdot(-1)}=\sqrt{p}\cdot\sqrt{-1}$$

将 i 视作 $\sqrt{-1}$ 的代数记号, 则

$$\sqrt{-p}=\sqrt{p}\,i$$

例如:

$$\sqrt{-4}=\sqrt{4}\,i=2i$$

$$\sqrt{-16}=\sqrt{16}\,i=4i$$

$$\sqrt{-3}=\sqrt{3}\,i$$

有了这一定义, 求根公式 $x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$ 可推广至判别式 $\Delta<0$ 的情形.

20.3.0.2 求根公式的应用

例题 20.3 解方程 $x^2+1=0$.

解 方程 $x^2+1=0$ 即 $x^2+0x+1=0$. 其判别式 $\Delta=b^2-4ac=0^2-4(1)(1)=-4$, 故 $\sqrt{\Delta}=\sqrt{-4}=2i$. 代入求根公式得

$$x=\frac{-b\pm\sqrt{\Delta}}{2a}=\frac{-0\pm2i}{2(1)}=\pm i$$

该方程在复数集 $\mathbb{C}$ 内的解为 $x_1=i$ 和 $x_2=-i$, 是一对共轭的纯虚数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

代入验证：当 $x=i$ 时，$x^2+1=i^2+1=-1+1=0$；当 $x=-i$ 时，$x^2+1=(-i)^2+1=i^2+1=-1+1=0$. 两根均满足原方程.

例题 20.4 解方程 $x^2-2x+5=0$.

解 方程 $x^2-2x+5=0$ 的判别式 $\Delta=(-2)^2-4(1)(5)=4-20=-16$，故 $\sqrt{\Delta}=\sqrt{-16}=4i$. 代入求根公式得

$$x=\frac{-(-2)\pm 4i}{2(1)}=\frac{2\pm 4i}{2}=1\pm 2i$$

方程的两个解为 $x_1=1+2i$ 和 $x_2=1-2i$，是一对共轭复数.

20.3.0.3 一般性结论

上述例子蕴含一个普遍规律，归结为下述定理，将一元二次方程的求根理论在复数集 $\mathbb{C}$ 内补全.

定理 20.1 (实系数二次方程的共轭虚根) 对于任意实系数一元二次方程 $ax^2+bx+c=0$ $(a\neq 0)$，当其判别式 $\Delta=b^2-4ac<0$ 时，方程在复数集 $\mathbb{C}$ 内有两个共轭的虚数根.

$$x_{1,2}=\frac{-b\pm\sqrt{|\Delta|}\,i}{2a} \tag{20-1}$$

证明 由配方法，

$$ax^2+bx+c=0 \iff (2ax+b)^2=b^2-4ac=\Delta.$$

当 $\Delta<0$ 时，记 $|\Delta|=-\Delta>0$. 因为 $i^2=-1$，所以

$$\left(\sqrt{|\Delta|}\,i\right)^2=-|\Delta|=\Delta.$$

因而 $(2ax+b)^2=\Delta$ 的两个解由

$$2ax+b=\pm\sqrt{|\Delta|}\,i$$

给出，即

$$x=\frac{-b\pm\sqrt{|\Delta|}\,i}{2a}.$$

由于 a,b,c 为实数且 $\sqrt{|\Delta|}\neq 0$，这两个根虚部非零且互为共轭. □

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

此定理表明, 复数域中任何实系数一元二次方程皆有解. 引入复数后, 此类方程在复数域内总能求解.

20.4 复数运算的几何意义 *

复数的代数形式 $z=a+bi$ 处理加减法时, 几何意义与向量的平行四边形法则完全对应. 乘除法则不然, 代数形式 $(ac-bd)+(ad+bc)i$ 难以显露其几何意义. 为探明乘除法在复平面上对应的几何变换, 我们引入复数的三角形式.

定义 20.4 (三角形式与辐角) 任何一个非零复数 $z=a+bi$ 都可以表示为.

$$z=r(\cos\theta+i\sin\theta)$$

其中 $r=|z|=\sqrt{a^2+b^2}$ 是复数的**模**, θ 是以实轴正半轴为始边, 向量 $\overrightarrow{OZ}$ 为终边的角, 称为**辐角**, 记作 $\operatorname{Arg}(z)$.
在所有可能的辐角中, 我们通常选取满足 $0\leq\theta<2\pi$(或 $-\pi<\theta\leq\pi$)的唯一值, 称之为**辐角主值**, 记作 $\arg(z)$.

三角形式把复数分解为模与辐角, 乘除法在此形式下有简洁的几何法则.

定理 20.2 (乘除法的三角表示) 设 $z_1=r_1(\cos\theta_1+i\sin\theta_1)$, $z_2=r_2(\cos\theta_2+i\sin\theta_2)$ 为非零复数.

$$\begin{aligned} z_1z_2&=r_1r_2[\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2)]\\ \frac{z_1}{z_2}&=\frac{r_1}{r_2}[\cos(\theta_1-\theta_2)+i\sin(\theta_1-\theta_2)] \end{aligned}$$

该定理给出了复数乘除法的几何含义, 两条法则:

- **复数相乘:模相乘, 辐角相加.**
- **复数相除:模相除, 辐角相减.**

从几何变换看, 用复数 z_2 乘 z_1, 相当于把代表 z_1 的向量 $\overrightarrow{OZ_1}$ 的长度变为原来的 r_2 倍, 再逆时针旋转 θ_2 角.

证明 以乘法为例, 除法类似. 将 $z_1=r_1(\cos\theta_1+i\sin\theta_1)$ 与 $z_2=r_2(\cos\theta_2+i\sin\theta_2)$ 直接相乘.

$$\begin{aligned} z_1z_2&=[r_1(\cos\theta_1+i\sin\theta_1)]\cdot[r_2(\cos\theta_2+i\sin\theta_2)]\\ &=r_1r_2(\cos\theta_1+i\sin\theta_1)(\cos\theta_2+i\sin\theta_2) \end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

展开并利用 $i^2=-1$:

$$=r_1r_2[(\cos\theta_1\cos\theta_2-\sin\theta_1\sin\theta_2)+i(\sin\theta_1\cos\theta_2+\cos\theta_1\sin\theta_2)]$$

括号内的实部与虚部恰是和角公式:

$$=r_1r_2[\cos(\theta_1+\theta_2)+i\sin(\theta_1+\theta_2)]$$

乘法法则得证.

对于除法法则, 需有 $z_2\neq 0$, 即 $r_2>0$. 先注意到

$$\frac{1}{\cos\theta_2+i\sin\theta_2}=\cos\theta_2-i\sin\theta_2=\cos(-\theta_2)+i\sin(-\theta_2),$$

因为两者相乘为 $\cos^2\theta_2+\sin^2\theta_2=1$. 因此

$$\begin{aligned}\frac{z_1}{z_2}&=\frac{r_1}{r_2}(\cos\theta_1+i\sin\theta_1)(\cos(-\theta_2)+i\sin(-\theta_2))\\&=\frac{r_1}{r_2}[\cos(\theta_1-\theta_2)+i\sin(\theta_1-\theta_2)],\end{aligned}$$

其中最后一步使用已证的乘法法则. 除法法则得证. □

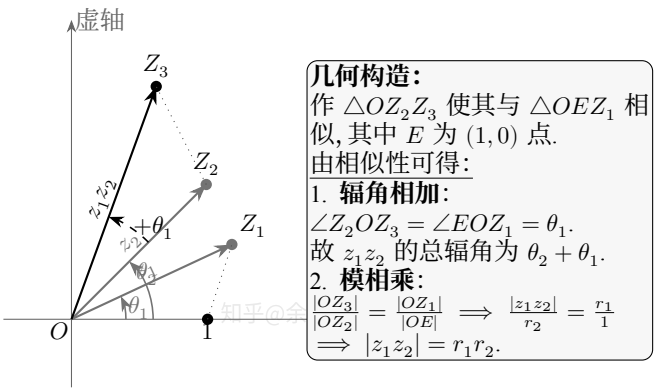


图 20.3　复数乘法的几何意义

例题 20.5　已知平面直角坐标系中的点 A 的坐标为 (x,y). 将点 A 绕原点 O 逆时针旋转 90° 得到点 B, 将点 A 绕原点 O 顺时针旋转 90° 得到点 C, 请求出点 B 和点 C 的坐标.

解　平面上的点用复数表示, 绕原点的旋转对应于乘以一个模为 1 的复数.

点 $A(x,y)$ 对应复数 $z_A=x+yi$.

将一个复数乘以 $\cos\theta+i\sin\theta$, 即把对应向量绕原点逆时针旋转 θ 角.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

点 B 由 A 绕原点逆时针旋转 90° 得到, 对应算子

$$\cos(90^\circ)+i\sin(90^\circ)=0+1\cdot i=i$$

故

$$\begin{aligned} z_B &= z_A\cdot i \\ &= (x+yi)i \\ &= xi+yi^2=-y+xi \end{aligned}$$

点 B 的坐标为 $(-y,x)$.

点 C 由 A 绕原点顺时针旋转 90°, 即逆时针旋转 -90°, 对应算子

$$\cos(-90^\circ)+i\sin(-90^\circ)=0-1\cdot i=-i$$

故

$$\begin{aligned} z_C &= z_A\cdot(-i) \\ &= (x+yi)(-i) \\ &= -xi-yi^2=y-xi \end{aligned}$$

点 C 的坐标为 $(y,-x)$.

点 B 的坐标为 $(-y,x)$, 点 C 的坐标为 $(y,-x)$.

本例展示了复数的一项应用, **将几何变换代数化**. 由此直接得到平面点绕原点旋转的坐标公式.

- 点 (x,y) 绕原点逆时针旋转 90° 后的坐标为 $(-y,x)$.
- 点 (x,y) 绕原点顺时针旋转 90° 后的坐标为 $(y,-x)$.

其代数根源在于复数乘法的几何意义. 复数把旋转、伸缩等几何变换归结为乘法运算.

复数乘法法则的一个直接推论是关于乘方运算的棣莫弗公式.

定理 20.3 (棣莫弗公式) 对于任意整数 n 和任意正实数 r,

$$[r(\cos\theta+i\sin\theta)]^n=r^n[\cos(n\theta)+i\sin(n\theta)]$$

证明 以 n 为正整数为例, 用数学归纳法证明, 实为多次应用模相乘, 辐角相加法则.

当 $n=1$ 时, 左边 $=r(\cos\theta+i\sin\theta)$, 右边 $=r^1(\cos(1\theta)+i\sin(1\theta))$, 等式成立.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设 $n=k$ ($k\in\mathbb{N}^*$) 时命题成立, 即 $[r(\cos\theta+i\sin\theta)]^k=r^k[\cos(k\theta)+i\sin(k\theta)]$. 往证 $n=k+1$ 时亦成立:

$$\begin{aligned}[r(\cos\theta+i\sin\theta)]^{k+1}&=[r(\cos\theta+i\sin\theta)]^k\cdot[r(\cos\theta+i\sin\theta)]\\&\overset{\text{归纳假设}}{=}(r^k[\cos(k\theta)+i\sin(k\theta)])\cdot(r[\cos\theta+i\sin\theta])\\&\overset{\text{乘法法则}}{=}(r^k\cdot r)[\cos(k\theta+\theta)+i\sin(k\theta+\theta)]\\&=r^{k+1}[\cos((k+1)\theta)+i\sin((k+1)\theta)]\end{aligned}$$

故 $n=k+1$ 时命题也成立.

由数学归纳法, 公式对所有正整数 n 成立. 当 $n=0$ 时两边均为 1; 当 $n<0$ 时, 令 $m=-n>0$. 由除法三角表示知

$$z^{-1}=r^{-1}(\cos(-\theta)+i\sin(-\theta)).$$

于是

$$z^n=(z^{-1})^m=r^{-m}[\cos(-m\theta)+i\sin(-m\theta)]=r^n[\cos(n\theta)+i\sin(n\theta)].$$

故公式对所有整数 n 成立. □

例题 20.6 (棣莫弗公式的应用) 计算 $(1+i)^{16}$.

解 直接代数展开繁琐, 改用棣莫弗公式, 先把底数化为三角形式.

设 $z=1+i$. 其模 $|z|=\sqrt{1^2+1^2}=\sqrt{2}$; 又 z 对应点 $(1,1)$, 位于第一象限, 由 $\cos\theta=\frac{a}{r}=\frac{1}{\sqrt{2}}$ 及 $\sin\theta=\frac{b}{r}=\frac{1}{\sqrt{2}}$ 得辐角主值 $\theta=\frac{\pi}{4}$. 故 $1+i=\sqrt{2}(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4})$.

应用棣莫弗公式并化简得

$$\begin{aligned}(1+i)^{16}&=\left[\sqrt{2}\left(\cos\frac{\pi}{4}+i\sin\frac{\pi}{4}\right)\right]^{16}\\&=(\sqrt{2})^{16}\left[\cos\left(16\cdot\frac{\pi}{4}\right)+i\sin\left(16\cdot\frac{\pi}{4}\right)\right]\\&=(2^{1/2})^{16}[\cos(4\pi)+i\sin(4\pi)]\\&=2^8[1+i\cdot0]\\&=256\end{aligned}$$

即 $(1+i)^{16}=256$.

例题 20.7 利用复数方法, 将 $\cos(3\theta)$ 和 $\sin(3\theta)$ 表示为 $\cos\theta$ 和 $\sin\theta$ 的多项式.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解 对复数 $(\cos\theta + i\sin\theta)^3$ 做两次计算, 一次用棣莫弗公式, 一次用二项式定理展开. 比较两次结果的实部与虚部, 即得所需三角恒等式.

据棣莫弗公式:

$$(\cos\theta + i\sin\theta)^3 = \cos(3\theta) + i\sin(3\theta) \tag{20-2}$$

知乎@余命数

另一方面, 用二项式定理展开 $(\cos\theta + i\sin\theta)^3$:

$$\begin{aligned}(\cos\theta + i\sin\theta)^3 &= \binom{3}{0}\cos^3\theta + \binom{3}{1}\cos^2\theta(i\sin\theta) \\ &\quad + \binom{3}{2}\cos\theta(i\sin\theta)^2 + \binom{3}{3}(i\sin\theta)^3 \\ &= \cos^3\theta + 3i\cos^2\theta\sin\theta + 3\cos\theta(i^2\sin^2\theta) + i^3\sin^3\theta \\ &= \cos^3\theta + 3i\cos^2\theta\sin\theta - 3\cos\theta\sin^2\theta - i\sin^3\theta\end{aligned}$$

(此处利用 $i^2 = -1$ 与 $i^3 = -i$.)

整理其实部与虚部得

$$(\cos\theta + i\sin\theta)^3 = (\cos^3\theta - 3\cos\theta\sin^2\theta) + i(3\cos^2\theta\sin\theta - \sin^3\theta) \tag{20-3}$$

知乎@余命数

比较 (20-2) 与 (20-3) 的实部与虚部, 即得

$$\cos(3\theta) = \cos^3\theta - 3\cos\theta\sin^2\theta$$

$$\sin(3\theta) = 3\cos^2\theta\sin\theta - \sin^3\theta$$

为得到只含 $\cos\theta$ 的 $\cos(3\theta)$ 表达式, 代入 $\sin^2\theta = 1 - \cos^2\theta$:

$$\begin{aligned}\cos(3\theta) &= \cos^3\theta - 3\cos\theta(1 - \cos^2\theta) \\ &= 4\cos^3\theta - 3\cos\theta\end{aligned}$$

知乎@余命数

同理, 为得到只含 $\sin\theta$ 的 $\sin(3\theta)$ 表达式, 代入 $\cos^2\theta = 1 - \sin^2\theta$:

$$\begin{aligned}\sin(3\theta) &= 3(1 - \sin^2\theta)\sin\theta - \sin^3\theta \\ &= 3\sin\theta - 4\sin^3\theta\end{aligned}$$

于是所求三角恒等式为

知乎@余命数

$$\cos(3\theta) = 4\cos^3\theta - 3\cos\theta$$

$$\sin(3\theta) = 3\sin\theta - 4\sin^3\theta$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

棣莫弗公式处理复数高次幂与开方, 把多项式形式的**代数乘方运算**归约为**模的实数乘方**与**辐角的线性数乘**.

20.4.1 历史纵横

棣莫弗公式 $[r(\cos\theta+i\sin\theta)]^n=r^n[\cos(n\theta)+i\sin(n\theta)]$ 看似只是计算技巧, 其背后另有深意.

公式以法国数学家亚伯拉罕·棣莫弗命名. 他长期客居英国, 与牛顿、哈雷交往密切, 在概率论和解析几何领域贡献卓著. 约 1707 年他最早注意到复数与三角函数之间的这种关系, 并就正整数次幂作了阐述. 其后 **莱昂哈德·欧拉**在 1748 年提出**欧拉公式**

$$e^{i\theta}=\cos\theta+i\sin\theta \tag{20-4}$$

后, 棣莫弗公式便有了解释, 它实为 $(a^x)^y=a^{xy}$ 的自然推论:

$$[r(\cos\theta+i\sin\theta)]^n=(re^{i\theta})^n=r^n(e^{i\theta})^n=r^ne^{i(n\theta)}=r^n[\cos(n\theta)+i\sin(n\theta)]$$

棣莫弗发现了现象, 欧拉给出了根本解释.

棣莫弗公式的主要用途如下. 最直接的应用, 是把涉及二项式展开的**乘方运算**转化为**辐角乘法**. 除求高次幂外, 它还是**复数开方**, 例如 $z^n=c$ 的理论基础.

棣莫弗公式可推导任意倍角的三角函数公式. 例如推导 $\cos(2\theta)$ 与 $\sin(2\theta)$, 只需展开 $(\cos\theta+i\sin\theta)^2$:

$$(\cos\theta+i\sin\theta)^2=\cos(2\theta)+i\sin(2\theta)$$

另一方面, 按二项式展开:

$$\begin{aligned}(\cos\theta+i\sin\theta)^2&=\cos^2\theta+2i\sin\theta\cos\theta+(i\sin\theta)^2\\&=(\cos^2\theta-\sin^2\theta)+i(2\sin\theta\cos\theta)\end{aligned}$$

比较两式的实部和虚部:

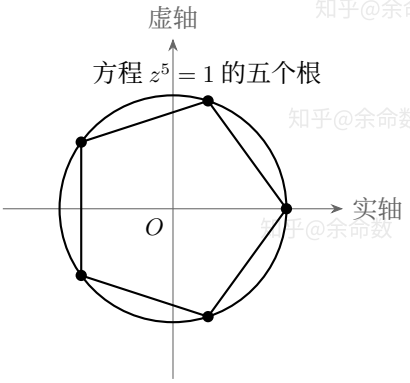
$$\cos(2\theta)=\cos^2\theta-\sin^2\theta\quad\text{且}\quad\sin(2\theta)=2\sin\theta\cos\theta$$

这给出了不依赖几何法、纯代数推导三角恒等式的途径.

模相乘, 辐角相加这一法则精确刻画了几何上的**旋转与伸缩**. 求解方程 $z^n=1$ (单位根)时, 棣莫弗公式表明其 n 个解均匀分布在复平面的单位圆上, 构成**正 n 边形**的顶点.

1. 已知复数 $z=3-4i$, 求它的共轭复数 $\bar{z}$ 和模 $|z|$. 并在复平面内指出 z 和 $\bar{z}$ 对应的点的位置.
2. 计算.(1) $(2-i)(1+3i)-4i$　(2) $\frac{1+i}{1-i}+i^5$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵



3. 设复数 $z=(m^2-1)+(m-1)i$, 其中 $m\in\mathbb{R}$. 分别求当 m 取何值时. (1) z 是实数; (2) z 是虚数; (3) z 是纯虚数; (4) z 对应的点在第四象限.
4. 若复数 z 满足 $(1+i)z=2-4i$, 求复数 z.

20.5 代数基本定理 *

已看到, 复数域 $\mathbb{C}$ 中所有实系数一元二次方程都有解. 对更高次的方程, 例如 $z^3-2z+5=0$ 或 $z^5+iz^2-(3+i)=0$, 复数系已足以应对.

无需为此再发明新的数.

定理 20.4 (代数基本定理) 任何一个次数 $n\geq 1$ 的复系数多项式方程

$$P(z)=a_nz^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0=0$$

(其中系数 $a_k\in\mathbb{C}$ 且 $a_n\neq 0$)在复数集 $\mathbb{C}$ 中至少有一个根.

由代数基本定理, 经因式分解得到一个更强的推论:

推论 20.1 任何一个 n 次复系数多项式在复数集 $\mathbb{C}$ 中恰好有 n 个根(计入重数). 即, 任何复系数多项式都可以被完全分解为 n 个一次因式的乘积.

$$P(z)=a_n(z-z_1)(z-z_2)\cdots(z-z_n)$$

其中 $z_1,z_2,\cdots,z_n$ 是 $P(z)=0$ 的 n 个复数根.

注 20.2 (证明说明) 代数基本定理的严格证明需要复分析、拓扑或高等代数工具, 例如刘维尔定理或绕数思想. 本书把它作为引用定理使用.

证明 对次数 n 作归纳. 当 $n=1$ 时, 一次多项式 a_1z+a_0 的唯一根为 $z=-a_0/a_1$, 结论成立.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

假设次数小于 n 的复系数多项式均可完全分解. 对 n 次多项式 $P(z)$, 由代数基本定理, 存在 $z_1 \in \mathbb{C}$ 使 $P(z_1)=0$. 由因式定理, $z-z_1$ 整除 $P(z)$, 因而存在一个 $n-1$ 次复系数多项式 $Q(z)$, 使得

$$P(z)=(z-z_1)Q(z).$$

由归纳假设, $Q(z)$ 可分解为 $a_n(z-z_2)\cdots(z-z_n)$. 于是

$$P(z)=a_n(z-z_1)(z-z_2)\cdots(z-z_n).$$

因式分解中的每个一次因式给出一个根, 重复因式按出现次数计入重数, 总数正好为 n. □

回顾数系扩张的历程, 每一步都是为了让某种方程有解.

- 为了解 $x+5=2$, 将自然数 $\mathbb{N}$ 扩展到整数 $\mathbb{Z}$.
- 为了解 $3x=2$, 将整数 $\mathbb{Z}$ 扩展到有理数 $\mathbb{Q}$.
- 为了解 $x^2=2$, 将有理数 $\mathbb{Q}$ 扩展到实数 $\mathbb{R}$.
- 为了解 $x^2=-1$, 将实数 $\mathbb{R}$ 扩展到复数 $\mathbb{C}$.

代数基本定理表明, 为求解**多项式方程**而扩张数系的旅程, 到复数 $\mathbb{C}$ 即告终结. 不必再为任何多项式方程发明新的数. 因此复数集 $\mathbb{C}$ 是**代数封闭**的.

代数基本定理关于**实系数**多项式的一个推论如下.

性质 (实系数多项式的虚根成对定理) 若实系数多项式 $P(z)$ 有一个虚数根 $z_0=a+bi\ (b\neq 0)$, 则其共轭复数 $\bar{z_0}=a-bi$ 也必是它的一个根.

证明 设实系数多项式 $P(z)=a_nz^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0$, 所有系数 a_k 为实数. 已知 z_0 是 $P(z)$ 的根, 即 $P(z_0)=a_nz_0^n+a_{n-1}z_0^{n-1}+\cdots+a_1z_0+a_0=0$. 对该等式两边取共轭, 逐项应用共轭对加减、乘法可分配的性质 ($\overline{u\pm v}=\bar{u}\pm\bar{v}$, $\overline{uv}=\bar{u}\,\bar{v}$, 故 $\overline{a_kz_0^k}=\overline{a_k}(\bar{z}_0)^k$), 得

$$\begin{aligned}
0=\overline{P(z_0)}&=\overline{a_nz_0^n}+\overline{a_{n-1}z_0^{n-1}}+\cdots+\overline{a_1z_0}+\overline{a_0} && \text{(共轭对加法可分配)}\\
&=\overline{a_n}(\bar{z}_0)^n+\overline{a_{n-1}}(\bar{z}_0)^{n-1}+\cdots+\overline{a_1}\,\bar{z}_0+\overline{a_0} && \text{(共轭对乘法可分配)}\\
&=a_n(\bar{z}_0)^n+a_{n-1}(\bar{z}_0)^{n-1}+\cdots+a_1\bar{z}_0+a_0 && (a_k\in\mathbb{R},\overline{a_k}=a_k).
\end{aligned}$$

末行恰是把 $\bar{z}_0$ 代入 $P(z)$ 所得, 即 $P(\bar{z}_0)=0$, 故 $\bar{z}_0$ 也是 $P(z)$ 的根. □

虚根成对定理解释了实系数一元二次方程判别式为负时,两虚根恒为共轭,例如 $1\pm 2i$, 此即实系数多项式根的共轭对称性.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

20.6 曲线方程的复数表示 *

复数模的几何意义提供了一种简洁描述平面上几何轨迹的方式. 其依据是:

$$\text{复数}z_1\text{与}z_2\text{所对应点之间的距离} = |z_1 - z_2|$$

据此, 许多基于距离定义的几何图形可直接翻译为复数方程. 设变量 z 对应轨迹上任意一点 P, 而 z_0, z_1, z_2 等对应平面上的定点.

- **圆**:
  几何定义:平面上到定点 $C(z_0)$ 的距离等于定长 r 的所有点 $P(z)$ 的集合.
  这一关系可以直接表述为 $PC = r$, 即.

$$|z - z_0| = r \quad (r > 0)$$

  其中, z_0 是圆心对应的复数, r 是圆的半径.

- **线段的中垂线**:
  几何定义:平面上到两个定点 $A(z_1)$ 和 $B(z_2)$ 距离相等的所有点 $P(z)$ 的集合.
  该定义等价于 $PA = PB$, 即.

$$|z - z_1| = |z - z_2|$$

- **椭圆**:
  几何定义:平面上到两个定点(焦点)$F_1(z_1)$ 和 $F_2(z_2)$ 的距离之和为常数 $2a$ 的所有点 $P(z)$ 的集合, 其中该常数大于两焦点间的距离.
  此定义可直接翻译为 $PF_1 + PF_2 = 2a$, 即.

$$|z - z_1| + |z - z_2| = 2a \quad (2a > |z_1 - z_2|)$$

- **双曲线**:
  几何定义:平面上到两个定点(焦点)$F_1(z_1)$ 和 $F_2(z_2)$ 的距离之差的绝对值为常数 $2a$ 的所有点 $P(z)$ 的集合, 其中该常数小于两焦点间的距离.
  此定义可翻译为 $|PF_1 - PF_2| = 2a$, 即.

$$\big||z - z_1| - |z - z_2|\big| = 2a \quad (0 < 2a < |z_1 - z_2|)$$

  若去掉外层的绝对值, 方程 $|z - z_1| - |z - z_2| = 2a$ 仅表示双曲线靠近 z_2 的一支.

这种表示在处理涉及这些曲线的几何问题时, 把几何关系转化为代数运算.

20.6.0.1 有趣的应用

看一些有趣的应用.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 20.8　已知复数 z 满足方程 $|z-(4+3i)|=1$. 求 $|z|$ 的最大值与最小值.

解　问题归于复数模的几何意义: 方程 $|z-z_0|=r$ 描述以 z_0 为圆心、半径为 r 的圆, 而 $|z|$ 表示点 z 到原点的距离, 故需在已知圆上求到原点距离最大与最小的点. 这里方程 $|z-(4+3i)|=1$ 表示复平面上以 $C(4,3)$(对应 $z_0=4+3i$)为圆心、半径 $r=1$ 的圆.

连接原点 O 与圆心 C, 该连线与圆交于两点, 据几何直观距原点最近和最远的点必在此连线上, 即到原点的最小距离为 $|OC|-r$, 最大距离为 $|OC|+r$. 又圆心到原点的距离

$$|z_0|=|4+3i|=\sqrt{4^2+3^2}=\sqrt{16+9}=\sqrt{25}=5.$$

为严格化此估计, 对复数模用反向三角不等式 $\big||z_1|-|z_2|\big|\le|z_1-z_2|$, 取 $z_2=z_0$ 得 $\big||z|-|z_0|\big|\le|z-z_0|$. 代入 $|z-z_0|=1$ 与 $|z_0|=5$, 得

$$\begin{aligned}\big||z|-5\big|&\le 1,\\ -1\le|z|-5&\le 1,\\ 4\le|z|&\le 6.\end{aligned}$$

两个端点恰在连线 OC 与圆的交点处取得, 故 $|z|$ 的最小值为 4, 最大值为 6.

例题 20.9　在复平面内, 复数 z 满足方程 $|z-(1+2i)|=|z-(3-i)|$. (1) 描述该方程所表示的几何图形. (2) 求该图形的直角坐标方程.

解　把复数模的几何意义直接翻译: $|z-z_1|=|z-z_2|$ 对应到两定点距离相等的点的集合, 即线段的中垂线.

(1) 令 $z_1=1+2i$, $z_2=3-i$, 在复平面上对应点 $A(1,2)$ 与 $B(3,-1)$. 方程 $|z-z_1|=|z-z_2|$ 表示点 $P(z)$ 到 A 的距离等于到 B 的距离, 轨迹是线段 AB 的中垂线.

(2) 设 $z=x+yi, x,y\in\mathbb{R}$, 代入原方程得 $|(x-1)+(y-2)i|=|(x-3)+(y+1)i|$. 由复数模的定义两边平方并展开化简

$$\begin{aligned}(x-1)^2+(y-2)^2&=(x-3)^2+(y+1)^2,\\ x^2-2x+1+y^2-4y+4&=x^2-6x+9+y^2+2y+1,\\ -2x-4y+5&=-6x+2y+10,\\ 4x-6y-5&=0,\end{aligned}$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

即线段 AB 的中垂线直角坐标方程为 $4x-6y-5=0$.

例题 20.10 (阿波罗尼斯圆) 求复数 z 满足方程 $|z-1|=2|z+1|$ 时, 其在复平面上形成的轨迹.

解 形如 $|z-z_1|=k|z-z_2|\,(k>0,k\neq 1)$ 的方程定义阿波罗尼斯圆, 即到两定点距离之比为常数的点的轨迹. 通过代数方法化为直角坐标方程, 圆心与半径便一目了然.

设 $z=x+yi$, 代入得 $|(x-1)+yi|=2|(x+1)+yi|$. 两边平方去根号后展开整理

$$\begin{aligned}(x-1)^2+y^2&=4[(x+1)^2+y^2],\\ x^2-2x+1+y^2&=4(x^2+2x+1+y^2),\\ x^2-2x+1+y^2&=4x^2+8x+4+4y^2,\\ 3x^2+10x+3y^2+3&=0.\end{aligned}$$

两边除以 3 并配方, 化为圆的标准方程

$$\begin{aligned}x^2+\frac{10}{3}x+y^2+1&=0,\\ \left(x^2+\frac{10}{3}x+\left(\frac{5}{3}\right)^2\right)+y^2&=-1+\left(\frac{5}{3}\right)^2,\\ \left(x+\frac{5}{3}\right)^2+y^2&=-1+\frac{25}{9}=\frac{16}{9}=\left(\frac{4}{3}\right)^2.\end{aligned}$$

轨迹是以 $(-\frac{5}{3},0)$ 为圆心、半径 $\frac{4}{3}$ 的圆.

例题 20.11 已知复数 z 满足 $|z-3|+|z+3|=10$. 求 $|z|$ 的最大值和最小值.

解 先识别复数方程的几何图形. 方程 $|z-z_1|+|z-z_2|=2a$ 是椭圆的标准复数形式, 问题化为求椭圆上点到原点距离的最值.

方程 $|z-3|+|z+3|=10$ 表示点 z 到定点 $F_1(3,0)$ 与 $F_2(-3,0)$ 的距离之和为 10, 两定点分别对应复数 3 与 -3.

这是以 F_1,F_2 为焦点的椭圆. 由定义焦距 $2c=|3-(-3)|=6$, 故 $c=3$; 长轴长 $2a=10$, 故 $a=5$; 短半轴 $b=\sqrt{a^2-c^2}=\sqrt{5^2-3^2}=\sqrt{16}=4$.

椭圆中心在原点, 焦点在实轴上, 直角坐标方程为 $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, 即 $\frac{x^2}{25}+\frac{y^2}{16}=1$.

求 $|z|$ 的最值, 即椭圆上点到原点 $(0,0)$ 距离的最值. 因椭圆中心即原点, 直接考察长半轴与短半轴, 即点 z 位于长轴顶点 $(\pm5,0)$ 时距原点最远, $|z|=5$; 位于短轴顶点 $(0,\pm4)$ 时距原点最近, $|z|=4$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $|z|$ 的最大值为 5, 最小值为 4.

例题 20.12　在复平面上, 点 z 满足方程 $z\bar{z}+i(z-\bar{z})-2=0$. 求点 z 的轨迹.

解 复数方程同时含 z 与 $\bar{z}$ 时, 用 $z=x+yi$, $\bar{z}=x-yi$ 代换, 转化为关于 x,y 的直角坐标方程.

由 $z\bar{z}=|z|^2=x^2+y^2$ 及 $z-\bar{z}=(x+yi)-(x-yi)=2yi$, 代入原方程 $z\bar{z}+i(z-\bar{z})-2=0$ 得

$$(x^2+y^2)+i(2yi)-2=0,$$

$$x^2+y^2+2yi^2-2=0,$$

$$x^2+y^2-2y-2=0.$$

对 y 配方化为圆的标准方程

$$x^2+(y^2-2y)=2,$$

$$x^2+(y^2-2y+1)=2+1,$$

$$x^2+(y-1)^2=3.$$

点 z 的轨迹是以 $(0,1)$ 为圆心、半径 $\sqrt{3}$ 的圆.

20.6.0.2 复数和平面几何

复数方法可为一些平面几何定理提供简洁的代数证明.

例题 20.13　设复平面上的三点 A,B,C 对应的复数分别为 a,b,c, 且 A,B,C 构成非退化三角形. 证明 $\triangle ABC$ 为正三角形的充要条件是

$$a^2+b^2+c^2=ab+bc+ca.$$

证明 把正三角形的几何性质, 即一个顶点可由另一顶点绕第三个顶点旋转 $\pm 60^\circ$ 得到, 转为代数运算. 旋转 60° 对应乘以复数 $e^{i\pi/3}$.

$\triangle ABC$ 为正三角形, 等价于其内角均为 60°, 即向量 $\overrightarrow{AC}$ 可由向量 $\overrightarrow{AB}$ 绕点 A 旋转 $\pm 60^\circ$ 得到.

$\overrightarrow{AB}$ 对应复数 $b-a$, $\overrightarrow{AC}$ 对应复数 $c-a$. 旋转 $\pm 60^\circ$ 的算子为 $\cos(\pm 60^\circ)+i\sin(\pm 60^\circ)=\frac{1\pm i\sqrt{3}}{2}$.

故 $\triangle ABC$ 为正三角形的充要条件是

$$c-a=(b-a)\left(\frac{1+i\sqrt{3}}{2}\right)\quad 或\quad c-a=(b-a)\left(\frac{1-i\sqrt{3}}{2}\right),$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即 $\frac{c-a}{b-a} \in \left\{\frac{1+i\sqrt{3}}{2}, \frac{1-i\sqrt{3}}{2}\right\}$, 二者恰为二次方程 $x^2-x+1=0$ 的两根(两根之和为 1, 两根之积为 $\frac{1-i^2(3)}{4}=1$). 于是该几何条件等价于

$$\left(\frac{c-a}{b-a}\right)^2-\left(\frac{c-a}{b-a}\right)+1=0.$$

因 A,B,C 构成非退化三角形, $b\neq a$, 两边同乘 $(b-a)^2$ 去分母得 $(c-a)^2-(c-a)(b-a)+(b-a)^2=0$. 将其展开并合并同类项

$$(c^2-2ac+a^2)-(bc-ac-ab+a^2)+(b^2-2ab+a^2)=0,$$

$$c^2-2ac+a^2-bc+ac+ab-a^2+b^2-2ab+a^2=0,$$

$$a^2+b^2+c^2-ab-bc-ca=0,$$

$$a^2+b^2+c^2=ab+bc+ca.$$

非退化前提下上述每一步可逆, 故此代数条件是 $\triangle ABC$ 为正三角形的充要条件. □

例题 20.14 (拿破仑定理) 在任意 $\triangle ABC$ 的三边外侧分别作等边三角形 $\triangle BCD$, $\triangle CAE$, $\triangle ABF$. 设这三个等边三角形的中心分别为 P,Q,R. 证明 $\triangle PQR$ 是等边三角形.

证明 将正三角形的旋转关系转为复数乘法, 以三次单位根 $\omega=e^{i2\pi/3}$ 简化表示, 证明 P,Q,R 满足等边三角形的代数充要条件.

设点 A,B,C 对应复数 a,b,c. 设 $A\to B\to C$ 为顺时针方向; 若为逆时针, 下文只需将 ω 与 ω^2 对换, 证明相同. 设点 P,Q,R 对应复数 p,q,r. 令 $\omega=e^{i2\pi/3}=\cos(120^\circ)+i\sin(120^\circ)=-\frac{1}{2}+i\frac{\sqrt{3}}{2}$, 满足 $\omega^3=1$ 与 $1+\omega+\omega^2=0$.

考虑边 BC 外侧的等边 $\triangle BCD$, 顶点顺序 $B\to C\to D$ 逆时针. 向量 $\overrightarrow{BD}$ 由 $\overrightarrow{BC}$ 逆时针旋转 60° 得到. 旋转 60° 的算子 $e^{i\pi/3}=\cos(60^\circ)+i\sin(60^\circ)=-\omega^2$. 点 D 对应复数 d 满足:

$$d-b=(c-b)(-\omega^2)\implies d=b-\omega^2(c-b)$$

中心 P 是 $\triangle BCD$ 的重心, 故

$$\begin{aligned}p&=\frac{1}{3}(b+c+d)\\&=\frac{1}{3}(b+c+b-\omega^2(c-b))\\&=\frac{1}{3}(2b+c-\omega^2c+\omega^2b)\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \frac{1}{3}(b(2+\omega^2)+c(1-\omega^2))$$

由 $1+\omega+\omega^2=0, 2+\omega^2=1+(1+\omega^2)=1-\omega$, 故

$$p = \frac{1}{3}(b(1-\omega)+c(1-\omega^2))$$

同理, 对 CA 边外侧的 $\triangle CAE$, 顶点 $C\to A\to E$ 逆时针, 中心 Q:

$$q = \frac{1}{3}(c(1-\omega)+a(1-\omega^2))$$

对 AB 边外侧的 $\triangle ABF$, 顶点 $A\to B\to F$ 逆时针, 中心 R:

$$r = \frac{1}{3}(a(1-\omega)+b(1-\omega^2))$$

由复数 p,q,r 表示的三角形为正三角形, 充要条件是 $p^2+q^2+r^2=pq+qr+rp$, 等价地 $p+\omega q+\omega^2 r=0$ ($P\to Q\to R$ 逆时针) 或 $p+\omega^2 q+\omega r=0$. 检验第二个条件.

$$\begin{aligned}
&3(p+\omega^2 q+\omega r)\\
&=[b(1-\omega)+c(1-\omega^2)]+\omega^2[c(1-\omega)+a(1-\omega^2)]\\
&\quad+\omega[a(1-\omega)+b(1-\omega^2)]\\
&=b(1-\omega)+c(1-\omega^2)+c(\omega^2-\omega^3)\\
&\quad+a(\omega^2-\omega^4)+a(\omega-\omega^2)+b(\omega-\omega^3)
\end{aligned}$$

由 $\omega^3=1$ 与 $\omega^4=\omega$:

$$\begin{aligned}
&=b(1-\omega)+c(1-\omega^2)+c(\omega^2-1)+a(\omega^2-\omega)+a(\omega-\omega^2)+b(\omega-1)\\
&=a(\omega^2-\omega+\omega-\omega^2)+b(1-\omega+\omega-1)+c(1-\omega^2+\omega^2-1)\\
&=a(0)+b(0)+c(0)\\
&=0
\end{aligned}$$

由 $p+\omega^2 q+\omega r=0$, 三点 P,Q,R 构成正三角形. □

拿破仑定理的复数证明展示了几何问题如何代数化, 它把涉及多重几何构造的命题归约为一次代数验算.

整个构型译为复数语言后, 点 A,B,C 由复数 a,b,c 标识, 向量 $\overrightarrow{AB}$ 由复数之差 $b-a$ 表示, 而等边三角形的构造即一次旋转. 在逆时针方向的 $\triangle ABC$ 中, $\overrightarrow{AC}$ 即 $\overrightarrow{AB}$ 绕 A 逆时针旋转 60° 的结果, 对应于乘以 $e^{i\pi/3}$. 直接用 $e^{i\pi/3}$ 亦可, 但引入三

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

次单位根 $\omega = e^{i2\pi/3}$ 更为简洁, 其满足 $\omega^3 = 1$ 与 $1+\omega+\omega^2 = 0$, 处理正三角形对称性时不断消项; 旋转 60° 的算子写作 $-\omega^2$, 化简时反复用 $1+\omega+\omega^2 = 0$ 消去项, 避开三角函数的纠缠.

化简结果恒为零, 且与 a, b, c 的选取无关, 故无论 $\triangle ABC$ 形状如何, 其外侧等边三角形的中心必构成新等边三角形. 一个几何定理由此化为关于任意复数的代数恒等式, 显出几何图形背后的代数结构.

20.7 本章小结与思想方法

注 20.3 (本章知识网络)

- **概念与代数形式**
  - **定义**: $z = a + bi$ $(a, b \in \mathbb{R})$, 虚数单位 i 满足 $i^2 = -1$.
  - **分类**: 实数 $(b = 0)$, 虚数 $(b \neq 0)$, 纯虚数 $(a = 0, b \neq 0)$.
  - **相等**: $a + bi = c + di \iff a = c, b = d$.
  - **共轭**: $\bar{z} = a - bi$.
  - **模**: $|z| = \sqrt{a^2 + b^2}$. 重要性质: $z\bar{z} = |z|^2$.
- **复数运算**
  - **加/减法**: $(a + bi) \pm (c + di) = (a \pm c) + (b \pm d)i$.(几何上遵循向量法则)
  - **乘法**: $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$.
  - **除法**: $\frac{a+bi}{c+di} = \frac{(a+bi)(c-di)}{c^2+d^2}$.(要点是分母实数化)
  - **i 的周期性**: $i^{4k+1} = i, i^{4k+2} = -1, i^{4k+3} = -i, i^{4k} = 1$.
- **几何意义与三角形式**
  - **复平面**: 实轴与虚轴, 复数 $z \leftrightarrow$ 点 $Z(a, b) \leftrightarrow$ 向量 $\overrightarrow{OZ}$.
  - **三角形式**: $z = r(\cos\theta + i\sin\theta)$、$r = |z|$、$\theta = \arg(z)$.
  - **乘除运算**: $z_1 z_2 = r_1 r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)]$.(模相乘, 辐角相加)
  - **棣莫弗公式**: $[r(\cos\theta + i\sin\theta)]^n = r^n[\cos(n\theta) + i\sin(n\theta)]$.
  - **几何方程**: 圆 $|z - z_0| = r$, 直线 $|z - z_1| = |z - z_2|$.

注 20.4 (本章思想方法)

- **数形结合思想**: 贯穿本章. 复数把代数与几何联系起来.
  - **表示的结合**: 复数 $\leftrightarrow$ 点 $\leftrightarrow$ 向量, 模 $|z| \leftrightarrow$ 点到原点距离, 辐角 $\arg(z) \leftrightarrow$ 向量与实轴正向夹角.
  - **运算的结合**: 复数加减 $\leftrightarrow$ 向量平行四边形/三角形法则; 复数乘除 $\leftrightarrow$ 旋转与伸缩.
  - **问题的结合**: 许多代数问题, 如求模的最值, 可化为几何问题, 如求点到点或线的距离最值; 反之亦然, 如求轨迹方程.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

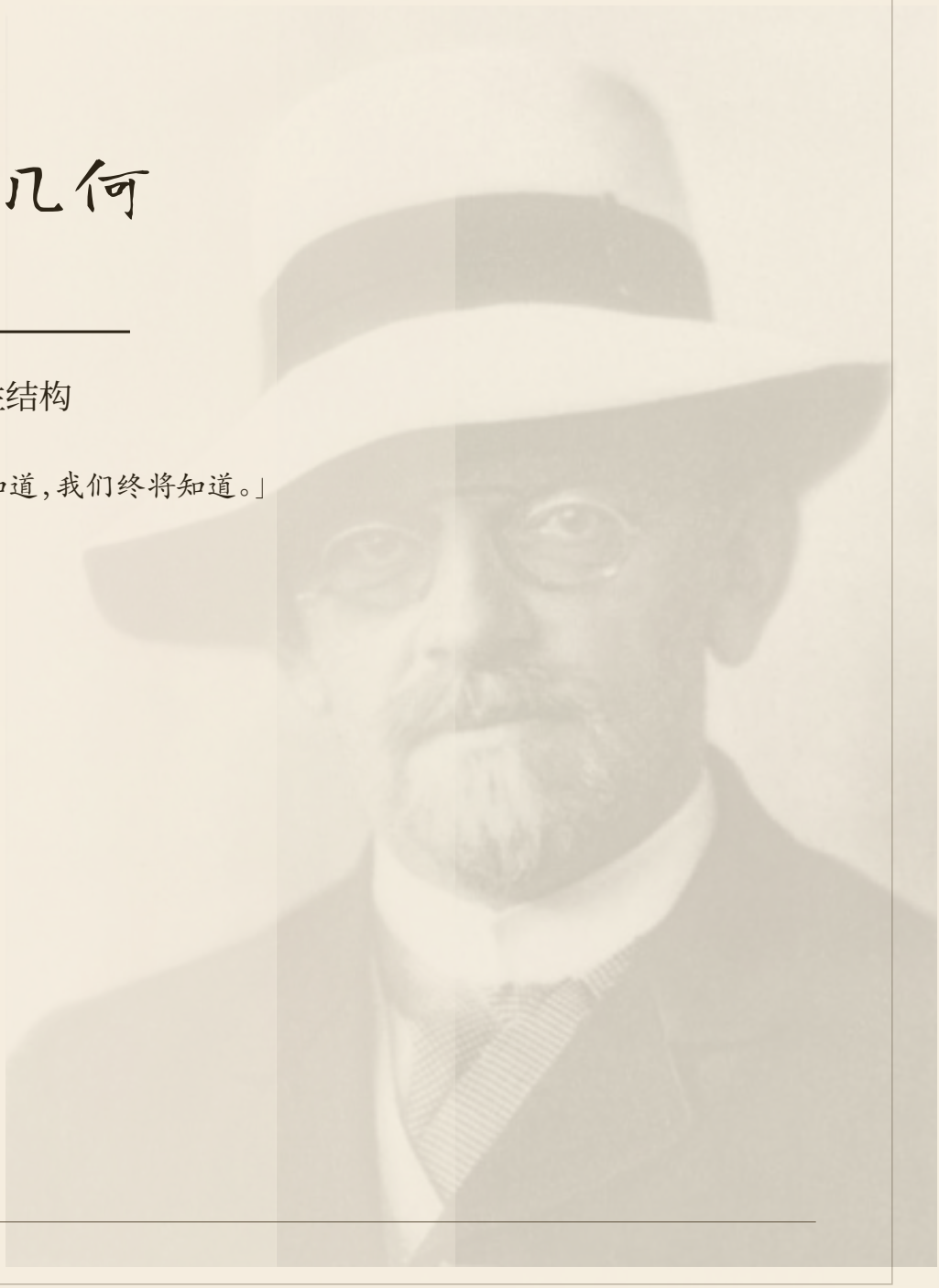
- **分类讨论思想**：用于含参复数问题. 按题设条件对参数讨论，确定复数是实数、虚数还是纯虚数，或确定其在复平面上的位置. 讨论须不重不漏.
- **转化与化归思想**：解决复杂问题的入口.
  - **形式转化**：代数形式与三角形式，或指数形式相互转化. 高次幂运算优先化为三角形式；加减或求轨迹时代数形式往往更方便.
  - **问题转化**：复数除法通过乘以共轭化为乘法；复数高次幂通过棣莫弗公式化为实数乘法.
  - **领域转化**：三角函数恒等式的代数证明化为复数乘方问题；代数方程求根化为几何图形顶点问题.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第21章

立体几何



扩展与线性结构

「我们必须知道，我们终将知道。」

——希尔伯特

第二十一章 立体几何

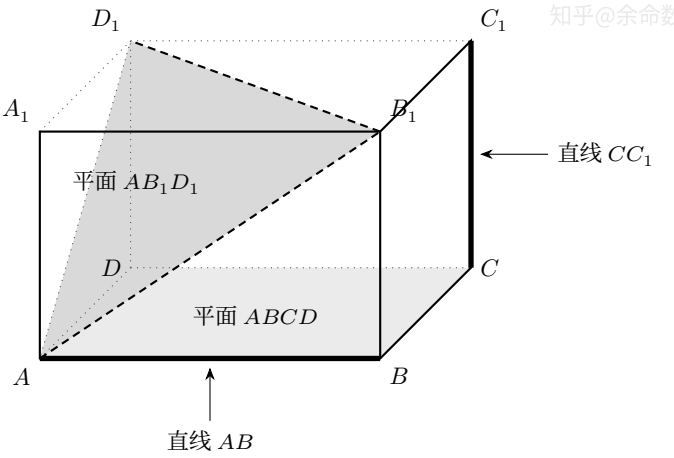


图 21.1 长方体中的直线与平面示意图

平面几何用公理体系描述二维空间. 在那里, 两条直线的位置关系只有相交与平行两种. 我们感知的物理世界是三维的. 要把几何从平面推广到空间, 我们必须引入新的基本元素, 并考察由此产生的更复杂的位置关系.

观察 图 21.1 中的平行六面体. 直线 AB 与直线 CC_1 不共面, 因此既不平行也不相交, 我们把这样的两条直线称为**异面直线**. 异面直线如何严格定义并研究, 两个平面在空间中相对倾斜的程度即二**面角**又如何刻画, 都是有赖于新体系才能说清的问题.

回答这些问题, 单靠直观不够. 我们扩展平面几何的公理基础, 建立一套适用于三维空间的演绎体系, 用它研究点、直线、平面三种基本元素之间位置关系的判定、性质与度量. 本章前半部分建立公理并给出位置关系的证明策略, 后半部分用空间向量与坐标法处理度量与计算.

21.1 基本公理与空间元素

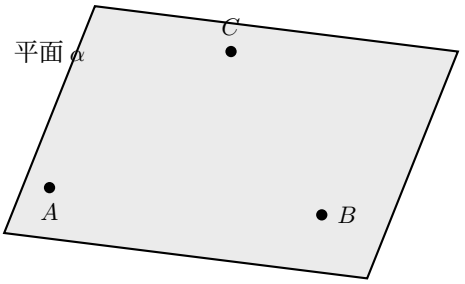
平面几何从点与直线这两个不加定义的基本元素出发. 立体几何引入第三个基本元素, 即**平面**, 我们用 α, β, γ 等希腊字母表示它. 这三种元素之间的基本关系由下面几条公理规定, 它们是一切逻辑推演的起点.

公理 21.1 不共线的三点确定唯一一个平面.

在空间中确定一个平面的位置, 依据的就是这条公理. 稳固的三脚架是它在物理世界中的对应物. 这条命题无法证明, 因为平面是不加定义的基本元素, 要证

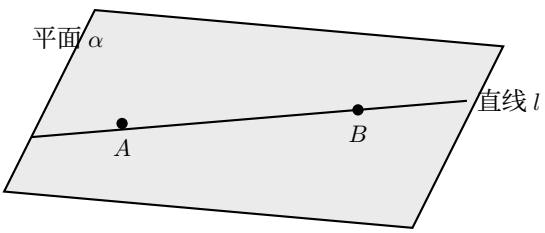
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

明它就得先有关于平面的更基本性质可用, 用公理证明公理是循环论证. 它之所以被接受, 在于与我们对平面的直接经验一致. 两点确定一条直线, 但过两点的平面有无穷多个(绕连线旋转即得); 一点更不够; 四点又可能不共面. 恰好三个不共线的点, 才把平面的位置钉死.



公理 21.2　若一条直线上有两点在一个平面内, 则该直线在这个平面内.

这条公理给出直线与平面之间的从属关系. 它保证了平面的平整性, 即平面上任意两点间的连线不会逃出该平面. 平面之所以为平面, 就在于它处处平直, 没有起伏. 若直线上已有两点落在某平面内, 而直线却有一部分翘出平面之外, 这个平面在该处就出现了弯折, 不再是平面. 要证明这条公理, 同样需要更基本的平面性质, 而这样的性质并不存在, 平整性本身就是我们对平面的原始认识.



公理 21.3　若两个不同平面有一个公共点, 则它们有且只有一条通过该点的公共直线.

这条公理排除了两个平面仅交于一点的可能, 并规定了两个不同平面相交的形态. 平面向各方向无限延伸, 两个这样的无限平面一旦在某点相遇, 必然沿着某个方向彼此贯穿, 贯穿的轨迹是一条直线. 若只交于一点, 则在这一点之外两平面立刻分离, 这要求平面在公共点处有某种边界或折痕, 与平面的无限延伸性相矛盾. 这条公理的依据是前两条公理所确立的无限延伸性与平直性, 三条公理彼此呼应, 共同锁定平面这一对象的全部基本特征.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

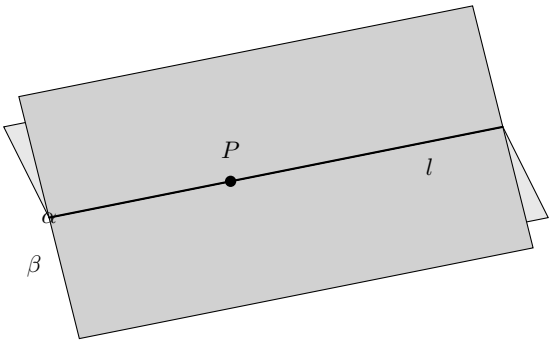


图 21.2 两平面相交公理的立体示意图

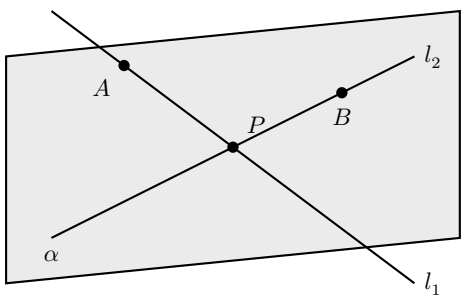
21.1.1 公理的推论

我们从这三条公理出发, 推导关于空间元素位置关系的基本定理.

定理 21.1 (相交直线确定平面) 两条相交直线确定唯一一个平面.

证明 设两直线 l_1, l_2 交于点 P. 在直线 l_1 上取异于 P 的一点 A, 在直线 l_2 上取异于 P 的一点 B. 由于 l_1 与 l_2 是不同的相交直线, 点 A, B, P 必然不共线. 据公理 1, 不共线的三点 A, B, P 确定唯一一个平面, 记为 α.

点 A 与点 P 均在平面 α 内, 且二者同在直线 l_1 上, 据公理 2 直线 l_1 落在平面 α 内. 同理, 点 B 与点 P 均在平面 α 内, 且二者同在直线 l_2 上, 据公理 2 直线 l_2 落在平面 α 内. 故存在一个包含 l_1, l_2 的平面 α. 任一包含 l_1, l_2 的平面都必须包含点 A, B, P, 而这三点仅能确定一个平面, 因而该平面唯一. □



定理 21.2 (平行直线确定平面) 两条平行直线确定唯一一个平面.

证明 设两直线 l_1, l_2 相互平行. 据平行直线的定义, 二者共面, 因而至少存在一个平面同时包含 l_1, l_2.

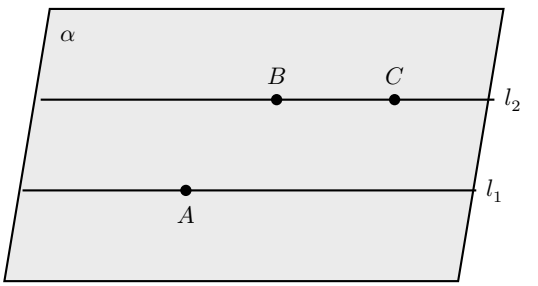
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

为证唯一, 在直线 l_1 上取一点 A, 在直线 l_2 上取两点 B, C. 由 $l_1 \parallel l_2$ 知两直线无公共点, 故 $A \notin l_2$, 从而 A, B, C 不共线. 据公理 1, 不共线三点 A, B, C 确定唯一平面 α.

点 B, C 在 α 内, 由公理 2 得 $l_2 \subset \alpha$. 又平行直线 l_1, l_2 共面, 且过直线 l_2 与直线外一点 A 的平面唯一, 故包含 l_2 与 A 的共面平面只能是 α, 因而 $l_1 \subset \alpha$.

任一同时包含 l_1, l_2 的平面都含 A, B, C 三点, 故只能是 α. 两条平行直线确定唯一一个平面. □

知乎@余命数



21.1.2 空间中直线与平面的位置关系

有了这些公理, 我们就可以对空间中基本元素之间可能的位置关系进行分类.

三维空间比平面多一个维度, 位置关系也更复杂. 我们依次考察直线与直线、直线与平面、平面与平面这三组基本关系.

21.1.2.1 直线与直线的位置关系

在平面几何中, 两条不同直线的位置关系只有两种, 即有公共点时相交于一点, 无公共点时相互平行. 这种划分隐含了一个前提, 即两条直线位于同一个平面上. 进入三维空间后, 这一前提不再必然成立, 于是出现了新的可能.

知乎@余命数

因此, 我们对空间中两条直线 l_1, l_2 分类时, 第一个判据是**共面性**.

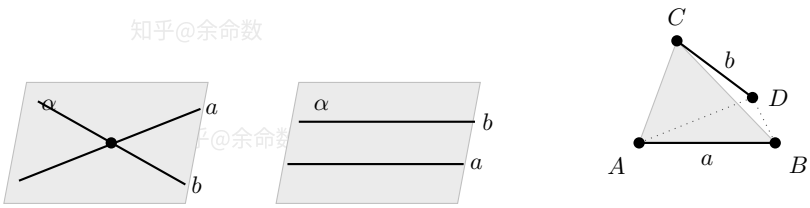
定义 21.1 (空间中两直线的位置关系) 设 l_1, l_2 为空间中两条不同的直线. 我们按它们是否共面, 把位置关系分成两层来定义.

1. **共面直线**: 若存在一个平面 α 同时包含 l_1 和 l_2. 在此前提下, 其位置关系与平面几何中的情况完全一致.
   - 若 l_1, l_2 有且仅有一个公共点, 则称它们**相交**.
   - 若 l_1, l_2 没有公共点, 则称它们**平行**.
2. **异面直线**: 若不存在任何一个平面能同时包含 l_1 和 l_2, 则称这两条直线**异面**.

这样分类之后, 判断两条直线的关系时, 我们先问它们是否共面.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 若共面, 则问题降维为熟知的二维平面问题, 后续只需考察公共点的个数.
- 若不共面, 则无需再考察公共点, 这两条直线的位置关系便唯一确定为异面. 异面直线既不平行, 因为平行线必须共面, 也不相交, 否则两相交直线必确定唯一平面, 与不共面矛盾.



相交　　平行　　异面

图 21.3　空间中两直线的三种位置关系

在具体问题中判定两条直线 l_1, l_2 的位置关系, 可以按下面的顺序进行. 先判断它们是否共面. 我们用几何构造来判断. 在 l_1 上取一点 A, 在 l_2 上取一点 B. 若 $l_1 \parallel l_2$ 或 l_1, l_2 相交, 则它们共面. 若不属于这两种情况, 则在 l_1 上另取一点 C. 如果 A, B, C 以及 l_2 均在同一个平面内, 则 l_1, l_2 共面.

若已确定 l_1, l_2 共面, 则问题退化为平面几何问题, 只需判断它们是否有公共点. 若有, 则相交; 若无, 则平行.

若通过任何构造都无法将 l_1, l_2 置于同一平面内, 例如, 在四面体中, 底面的边 AB 与其上方的棱 CD, 则根据定义, 它们是异面直线.

于是, 相交、平行、异面给出了空间中任意两条不同直线位置关系的完备且互斥的分类.

21.1.2.2 直线与平面的位置关系

设 l 是直线, α 是平面. 我们考察它们公共点的个数. 这个数只能是 0, 1 或无穷, 它给出了对位置关系分类的自然依据.

定义 21.2 (直线与平面的位置关系) 设 l 为直线, α 为平面.

- 若 l 与 α 有且仅有一个公共点, 则称直线与平面**相交**.
- 若 l 与 α 没有公共点, 则称直线与平面**平行**, 记作 $l \parallel \alpha$.
- 若 l 与 α 有无穷多个公共点, 则称**直线在平面内**, 记作 $l \subset \alpha$.

公共点的个数不能是 2 或更大的有限数. 根据公理 2, 若一条直线上有两点在平面内, 则该直线必在平面内. 于是公共点一旦有两个, 就必有无穷多个.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

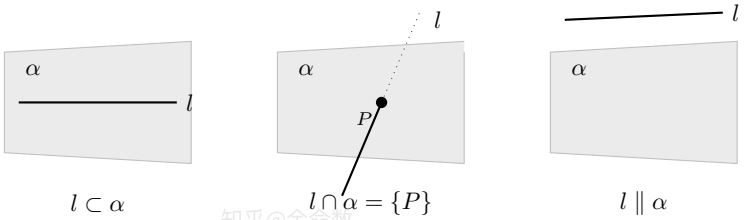


图 21.4　直线与平面的三种位置关系

21.1.2.3 平面与平面的位置关系

最后, 我们考察两个平面 α 和 β 在空间中的位置关系. 这种关系由它们的公共点集的形态决定.

定义 21.3 (平面与平面的位置关系) 设 α, β 为两个不同的平面.

- 若 α 与 β 没有公共点, 则称这两个平面**平行**, 记作 $\alpha \parallel \beta$.
- 若 α 与 β 有公共点, 则称这两个平面**相交**.

公理 3 直接规定了相交平面的形态. 两个不同平面若有公共点, 则它们必交于一条通过该点的直线. 于是两平面不可能仅交于一点或一条线段. 交集若非空, 必为一条无限延伸的直线. 至此, 我们定义完了空间基本元素之间的位置关系.



图 21.5　平面与平面的两种位置关系

21.1.3 本节习题

21.1.　判断以下论断是否为真, 并基于本节公理与推论阐述理由. 空间中任意四个点必然共面.

21.2.　证明一条直线与不经过该直线的一个点, 确定唯一一个平面.

21.3.　设四点 A, B, C, D 不共面. 证明连接 A, B 的直线与连接 C, D 的直线是异面直线.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

21.4. 在空间四边形 $ABCD$, 即四顶点不共面的四边形中, 点 M,N,P,Q 分别是棱 AB,BC,CD,DA 上的点. 若直线 MN 与直线 PQ 相交于点 K. 证明: A,C,K 三点共线.

21.5. 考察空间中三条不同的直线 l_1,l_2,l_3. 若它们两两相交, 证明这三条直线或者共点, 或者共面.

知乎@余命数

21.6. 考察空间中四条不同的直线, 其中任意三条都不共面. 若这四条直线中的每一条都与其他三条相交, 证明这四条直线必共点.

21.7. 设三个互不重合的平面 α,β,γ 两两相交, 且其三条交线互不平行. 证明这三条交线必共点.

21.8. (空间中的迪沙格定理)设 O,A,B,C,A',B',C' 为空间中七个不同的点, 且 A,B,C 不共线, A',B',C' 亦不共线. 若三条直线 AA',BB',CC' 交于一点 O. 设平面 ABC 与平面 $A'B'C'$ 相交于直线 l. 证明:两组对应边的交点,即 $P=AB\cap A'B'$、$Q=BC\cap B'C'$、$R=CA\cap C'A'$ 三点均在直线 l 上, 因而共线.

21.2 证明线线平行的基本策略

知乎@余命数

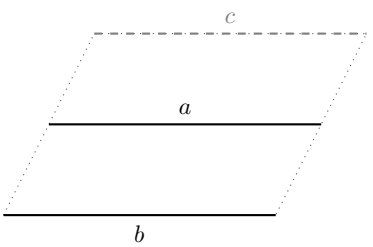
给定空间中两条直线 a,b, 证明它们相互平行(记作 $a\parallel b$)的基本思路有两条. 一条是把空间问题化为平面问题, 另一条是借助线面平行、面面平行这类更高层级的平行关系来推导. 由此得到三种常用策略.

21.2.1 利用平行公理的传递性

此法的理论基础是平行公理的传递性, 即平行于同一直线的两条直线互相平行.

这是最直接的方法. 设 c 是一条中间直线. 如果我们分别证明了 $a\parallel c$ 与 $b\parallel c$, 那么由传递性即得 $a\parallel b$. 这条中间直线 c 通常取某个辅助平面图形, 如三角形、平行四边形的边或对角线.

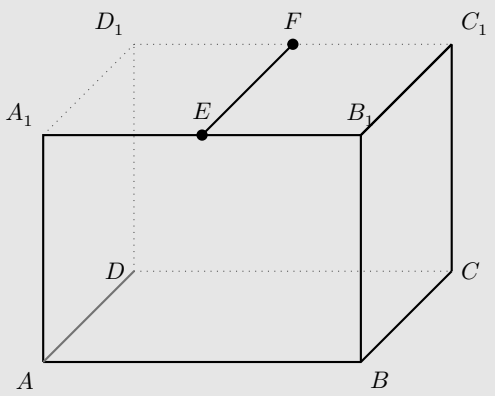
知乎@余命数



在梯形 $ABCD$ 中, 中位线 EF 平行于底边 DC 和 AB.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 21.1　在平行六面体 $ABCD-A_1B_1C_1D_1$ 中，设 E,F 分别是棱 A_1B_1 和 C_1D_1 的中点. 求证：$EF \parallel AD$.



证明　我们用传递性. 只要找到一条与 EF 和 AD 都平行的中间直线即可.

平行六面体的顶面 $A_1B_1C_1D_1$ 是平行四边形，故

$$A_1B_1 \parallel D_1C_1 \quad 且 \quad A_1B_1 = D_1C_1.$$

知乎@余命数

点 E,F 分别是 A_1B_1 与 C_1D_1 的中点，故

$$A_1E = \tfrac{1}{2}A_1B_1, \quad D_1F = \tfrac{1}{2}D_1C_1,$$

从而 $A_1E \parallel D_1F$ 且 $A_1E = D_1F$，四边形 A_1EFD_1 是平行四边形，对边 $EF \parallel A_1D_1$.

侧面 ADD_1A_1 同为平行四边形，故 $AD \parallel A_1D_1$. 以直线 A_1D_1 为中转，据平行公理的传递性，

$$EF \parallel AD.$$

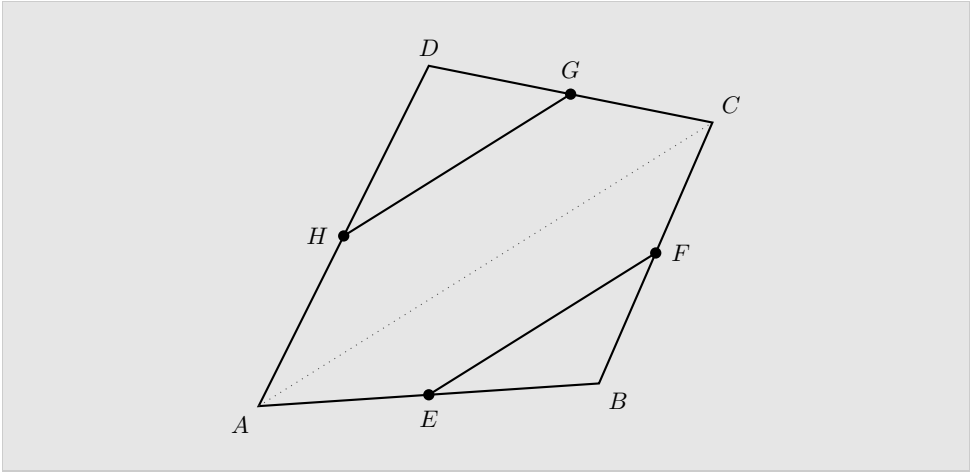
知乎@余命数

□

例题 21.2　在空间四边形 $ABCD$ 中，点 E,F,G,H 分别是边 AB,BC,CD,DA 的中点. 求证：$EF \parallel HG$.

知乎@余命数

知乎@余命数



证明 四点 A,B,C,D 不共面. 连接对角线 AC, 就把空间问题放进平面几何的框架内.

在 $\triangle ABC$ 中, E,F 分别为 AB,BC 的中点, 据三角形中位线定理,

$$EF \parallel AC.$$

在 $\triangle ADC$ 中, H,G 分别为 DA,CD 的中点, 同理

$$HG \parallel AC.$$

AC 作为 EF 与 HG 的中介. 由 $EF \parallel AC$ 与 $HG \parallel AC$, 据平行公理的传递性,

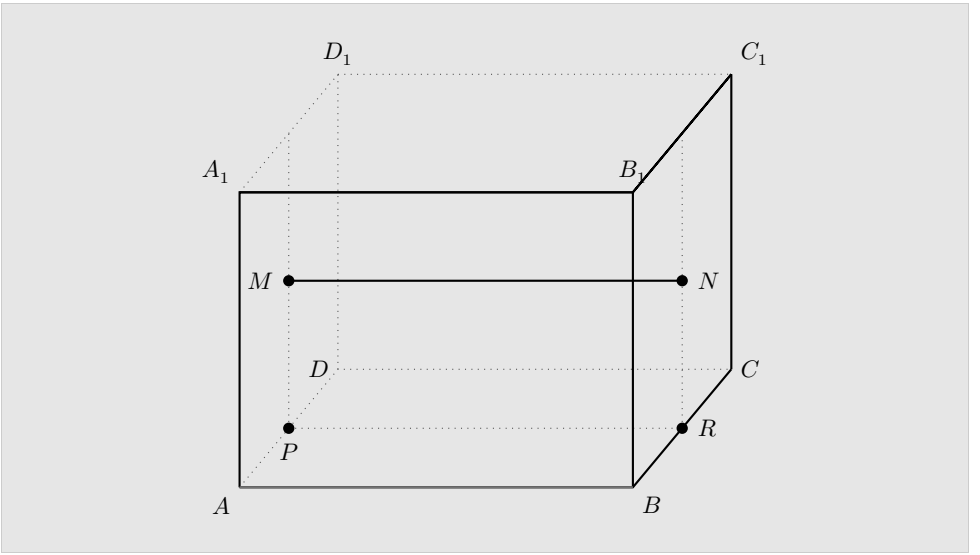
$$EF \parallel HG.$$

由此并得四边形 $EFGH$ 是平行四边形. □

知乎@余命数

例题 21.3 在平行六面体 $ABCD-A_1B_1C_1D_1$ 中, 设 M,N 分别是侧面 ADD_1A_1 与 BCC_1B_1 的中心. 求证: $MN \parallel AB$.

知乎@余命数
知乎@余命数



证明 平行四边形的中心是两条对角线的交点, 也是对角线的中点. 于是 M 是对角线 AD_1 与 A_1D 的中点, N 是对角线 BC_1 与 B_1C 的中点. 面中心的问题就化为边中点的问题.

在平行六面体中, $AB \parallel D_1C_1$ 且 $AB = D_1C_1$, 故 ABC_1D_1 是平行四边形. 其中 M, N 分别是边 AD_1, BC_1 的中点. 连接 AC_1, 设 O 为其中点. 在 $\triangle ABC_1$ 中, O, N 分别为 AC_1, BC_1 的中点, 故

$$ON \parallel AB.$$

在 $\triangle AD_1C_1$ 中, O, M 分别为 AC_1, AD_1 的中点, 故

$$OM \parallel D_1C_1.$$

由 $D_1C_1 \parallel AB$ 与平行公理的传递性, 得 $OM \parallel AB$.

$ON \parallel AB$ 与 $OM \parallel AB$ 同时成立, 而 OM, ON 皆过点 O. 据过直线外一点有且只有一条直线与已知直线平行的公理, M, O, N 共线, 故 $MN \parallel AB$. □

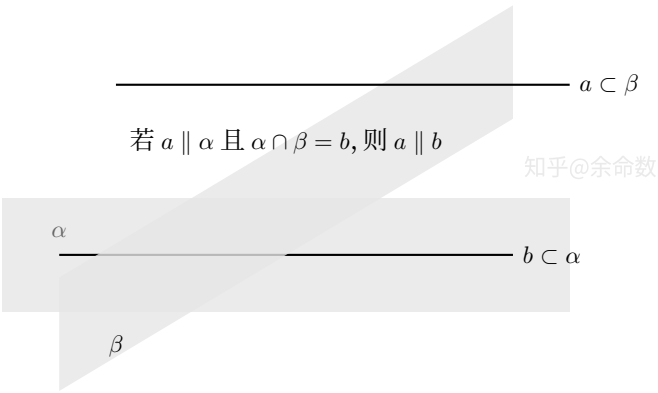
21.2.2 利用线面平行的性质定理

此法的理论基础是线面平行的性质定理. 设直线 a 与平面 α 平行, β 是过 a 的任意一个平面, 且 β 与 α 相交于直线 b, 那么 $a \parallel b$.

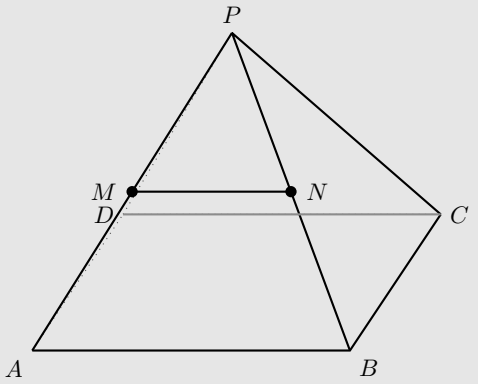
我们用降维的办法. 要证 $a \parallel b$, 先构造一个包含 b 的平面 α, 证明 $a \parallel \alpha$. 再指出 a 与 b 同在另一个平面 β 内, 且 β 与 α 的交线恰好是 b. 根据线面平行的性质定理, 即得 $a \parallel b$.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



例题 21.4 设四棱锥 $P-ABCD$ 的底面是平行四边形, 点 M, N 分别是棱 PA, PB 的中点. 求证: MN 平行于平面 PCD 与平面 $ABCD$ 的交线.



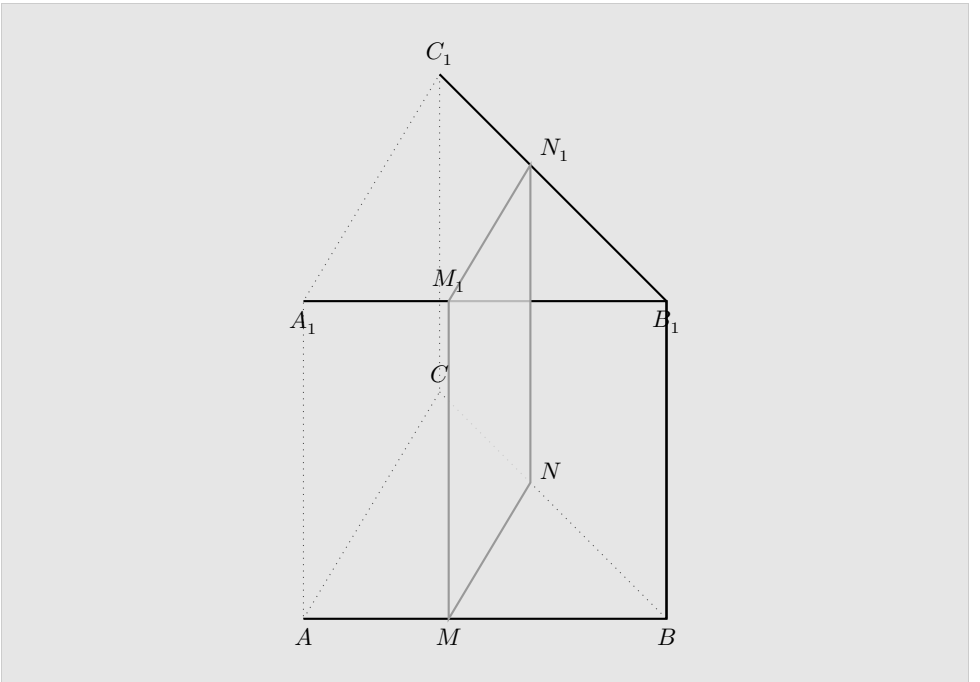
证明 题中已给直线 MN, 另一条直线是平面 PCD 与平面 $ABCD$ 的交线. 这两个平面都过棱 CD, 故交线就是 CD, 问题归为证 $MN \parallel CD$.

在 $\triangle PAB$ 中, M, N 分别为 PA, PB 的中点, 据三角形中位线定理 $MN \parallel AB$. 又底面 $ABCD$ 是平行四边形, 故 $AB \parallel CD$. 据平行公理的传递性, 由 $MN \parallel AB$ 与 $AB \parallel CD$ 推得 $MN \parallel CD$.

而 CD 正是平面 PCD 与平面 $ABCD$ 的交线, 故 MN 平行于二者的交线. □

例题 21.5 在三棱柱 $ABC-A_1B_1C_1$ 中, 一个平面 β 与棱 AC 平行, 且与棱 AB, BC, A_1B_1, B_1C_1 分别交于点 M, N, M_1, N_1. 求证: $MN \parallel M_1N_1$.

知乎@余命数



证明 我们用传递性来证 $MN \parallel M_1N_1$. 题设的关键条件是线面平行关系 平面$\beta \parallel AC$, 即 $AC \parallel$ 平面β.

直线 AC 含于底面 ABC, 可以应用线面平行的性质定理. 已知 AC 与平面 β (即平面 MNN_1M_1)平行, 平面 ABC 过 AC, 它与 β 的交线是 MN. 根据定理,

知乎@余命数

$$MN \parallel AC.$$

在三棱柱中, 侧面 ACC_1A_1 是平行四边形, 故 $AC \parallel A_1C_1$.

知乎@余命数

若 A_1C_1 落在平面 β 内, 则平面 $A_1B_1C_1$ 与 β 的交线为 A_1C_1, 即 $M_1N_1 = A_1C_1$, 由 $MN \parallel AC \parallel A_1C_1$ 已得结论. 否则, 由 $AC \parallel$ 平面β 与 $AC \parallel A_1C_1$ 推出 $A_1C_1 \parallel$ 平面β.

对顶面重复同一论证.A_1C_1 与平面 β 平行, 平面 $A_1B_1C_1$ 过 A_1C_1, 它与 β 的交线是 M_1N_1. 根据同一性质定理,

$$M_1N_1 \parallel A_1C_1.$$

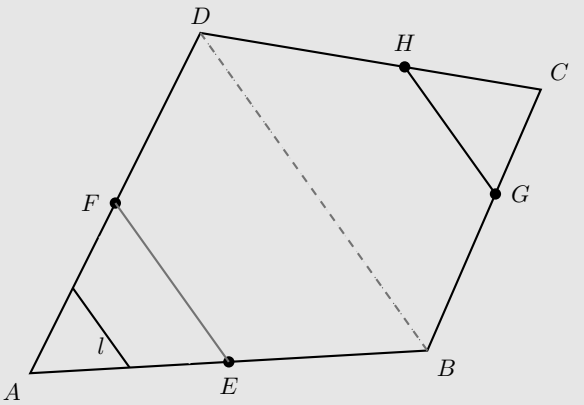
现有 $MN \parallel AC$, $M_1N_1 \parallel A_1C_1$, 且 $AC \parallel A_1C_1$. 据平行公理的传递性,

$$MN \parallel M_1N_1.$$

□

知乎@余命数

例题 21.6 在一个空间四边形 $ABCD$ 中, E, F 分别是边 AB, AD 的中点. 一个平面 α 经过 GH, 其中 G, H 分别是边 BC, CD 上的点, 且平面 α 平行于直线 EF. 若平面 α 与平面 ABD 的交线为 l, 求证: $l \parallel BD$.



证明 我们用线面平行的性质定理来证 $l \parallel BD$. 题设给出平面 $\alpha \parallel$ 直线 EF, 即 $EF \parallel \alpha$; 又 $l =$ 平面$\alpha \cap$ 平面ABD, 定理的条件正好齐备.

直线 EF 与平面 α 平行, 且含于平面 ABD 之中. 根据定理, 过 EF 的平面 ABD 与 α 的交线 l 必与 EF 平行, 即

$$l \parallel EF.$$

问题由 $l \parallel BD$ 归为 $EF \parallel BD$.

后者是平面几何问题. 在 $\triangle ABD$ 中, E, F 分别为 AB, AD 的中点, 据三角形中位线定理,

$$EF \parallel BD.$$

结合 $l \parallel EF$ 与 $EF \parallel BD$, 据平行公理的传递性,

$$l \parallel BD.$$

□

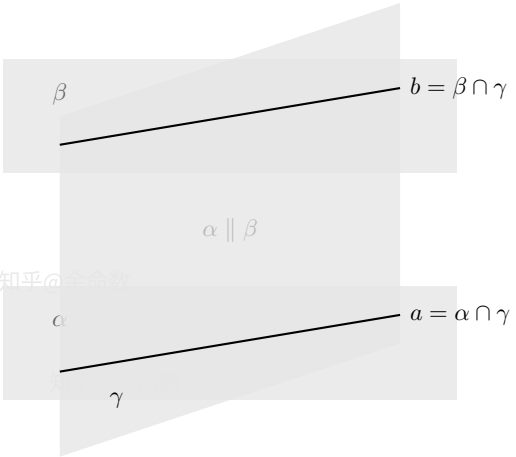
21.2.3 利用面面平行的性质定理

此法的理论基础是面面平行的性质定理, 即若两个平行平面同时与第三个平面相交, 那么它们的交线平行.

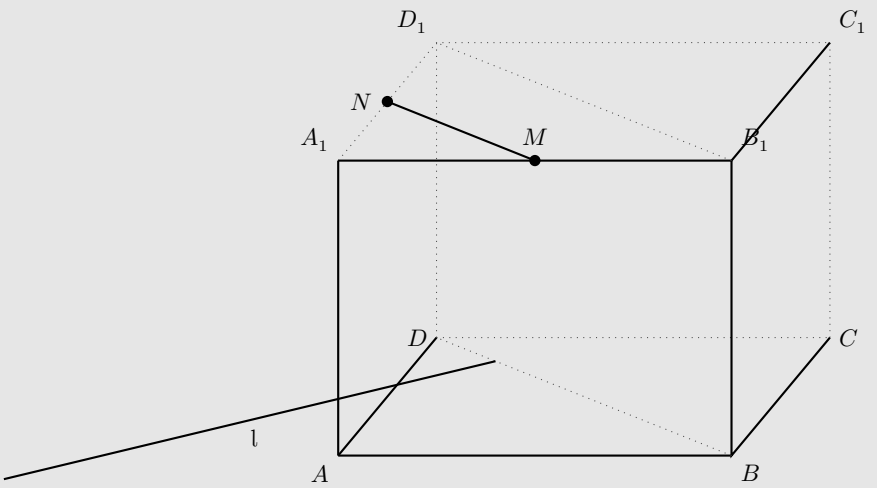
要证 $a \parallel b$, 我们构造两个相互平行的平面 α 与 β, 使 a 在 α 内, b 在 β 内. 再找第三个平面 γ, 它与 α 和 β 都相交, 且交线恰好分别是 a 与 b, 即 $\alpha \cap \gamma = a$ 且 $\beta \cap \gamma = b$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此时根据面面平行的性质定理, $a \parallel b$. 作截面时也常用这个定理. 几何体被一族平行平面所截, 截面轮廓线中那些平行的边就来自这里.



例题 21.7　在正方体 $ABCD-A_1B_1C_1D_1$ 中, 设 M, N 分别是棱 A_1B_1, A_1D_1 的中点. 一个过点 A, M, N 的平面 γ 与底面 $ABCD$ 相交于直线 l. 求证: $l \parallel BD$.



证明　我们用面面平行的性质定理来证 $l \parallel BD$. 顶面 $A_1B_1C_1D_1$ 与底面 $ABCD$ 平行. 记顶面为 β, 底面为 α, 则

$$\alpha \parallel \beta.$$

平面 γ(即平面 AMN)与这两个平面相交. 题设 γ 与底面 α 的交线为 l, 即

$$l = \gamma \cap \alpha.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

点 M,N 分别在棱 A_1B_1,A_1D_1 上,故同在平面 β 内. 它们又同在平面 γ 内,所以直线 MN 就是 γ 与 β 的交线,即

$$MN=\gamma\cap\beta.$$

两个平行平面 α,β 被第三平面 γ 所截,交线分别为 l 与 MN,据面面平行性质定理,

$$l\parallel MN.$$

问题归为证 $MN\parallel BD$.

在顶面 $\triangle A_1B_1D_1$ 中,M,N 分别为 A_1B_1,A_1D_1 的中点,据三角形中位线定理,

$$MN\parallel B_1D_1.$$

正方体中侧棱平行且相等,四边形 BDD_1B_1 是矩形,故

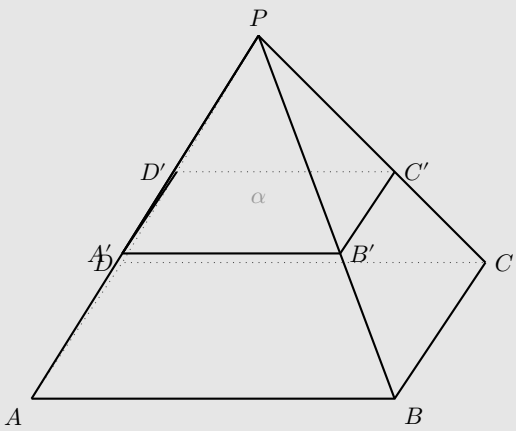
$$BD\parallel B_1D_1.$$

合 $l\parallel MN,MN\parallel B_1D_1,B_1D_1\parallel BD$,据传递性,

$$l\parallel BD.$$

□

例题 21.8 设有一棱锥 $P-ABCD$, 其底面 $ABCD$ 为平行四边形. 一个平行于底面的平面 α 与棱 PA,PB,PC,PD 分别相交于点 A',B',C',D'. 求证:四边形 $A'B'C'D'$ 是一个平行四边形.



证明 即证四边形 $A'B'C'D'$ 的两组对边分别平行, 即证 $A'B'\parallel D'C'$ 且 $A'D'\parallel B'C'$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

题设平面 $A'B'C'D'$(即平面 α)与底面 $ABCD$ 平行, 即

$$\text{平面}A'B'C'D' \parallel \text{平面}ABCD.$$

侧面 PAB 同时与这两个平面相交, 与 $A'B'C'D'$ 交于 $A'B'$, 与 $ABCD$ 交于 AB. 根据面面平行的性质定理,

$$A'B' \parallel AB.$$

侧面 PCD 同样与这两个平面相交, 与 $A'B'C'D'$ 交于 $D'C'$, 与 $ABCD$ 交于 DC. 根据同一性质定理,

$$D'C' \parallel DC.$$

底面 $ABCD$ 是平行四边形, 故 $AB \parallel DC$. 合 $A'B' \parallel AB$, $AB \parallel DC$, $DC \parallel D'C'$, 据传递性,

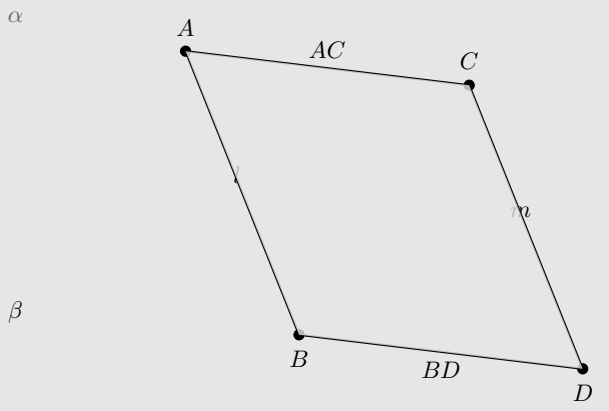
$$A'B' \parallel D'C'.$$

同理, 侧面 PAD 给出 $A'D' \parallel AD$, 侧面 PBC 给出 $B'C' \parallel BC$, 底面给出 $AD \parallel BC$. 据传递性,

$$A'D' \parallel B'C'.$$

四边形 $A'B'C'D'$ 两组对边分别平行, 故为平行四边形. □

例题 21.9 设 α 与 β 为两个相互平行的平面. 一条直线 l 与它们都相交, 交点分别为 $A \in \alpha$ 和 $B \in \beta$. 另有一点 C 在平面 α 内, 且 $C \neq A$. 过点 C 作直线 m, 使得 $m \parallel l$, 且 m 与平面 β 交于点 D. 求证: $AC \parallel BD$.



证明 已知 $l \parallel m$. 由 $C \neq A$ 且 l 与平面 α 仅交于点 A, 可知 m 与 l 不重合. 两条平行直线共面, 设含 l 与 m 的平面为 γ.

点 A, B 在直线 l 上, 点 C, D 在直线 m 上, 故四点皆在平面 γ 内.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

已知 $\alpha \parallel \beta$. 平面 γ 与平面 α 相交, 其交线是含全部公共点的直线; 点 A, C 既在 α 内又在 γ 内, 故为公共点, 从而

$$AC = \gamma \cap \alpha.$$

同理, 点 B, D 既在 β 内又在 γ 内, 故

$$BD = \gamma \cap \beta.$$

两平行平面 α, β 被第三平面 γ 所截, 交线分别为 AC 与 BD. 据面面平行性质定理,

$$AC \parallel BD.$$

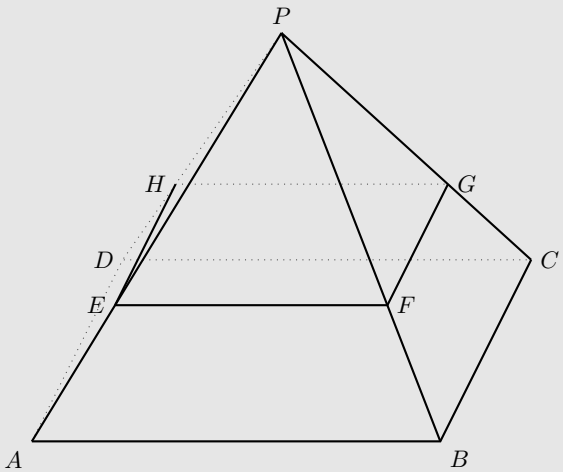
由此并得四边形 $ACDB$ 是平行四边形. □

21.2.4 计算问题

与平行有关的计算问题, 大多归结为求线段长度. 我们的做法是先用上面的平行判定定理与性质定理, 从线面平行或面面平行推出所需的线线平行.

有了线线平行, 就能构造一个包含目标线段的平面图形, 如三角形或梯形, 把空间中的长度计算降为平面几何问题, 再用相似、全等、三角函数或勾股定理求解.

例题 21.10 在一个底面为平行四边形的四棱锥 $P-ABCD$ 中, 一个平行于底面的平面 α 与棱 PA, PB, PC, PD 分别交于点 E, F, G, H. 若 $PE : EA = 2 : 1$, 且底面边长 $AB = 9$, $BC = 6$, 求截面四边形 $EFGH$ 的周长.



知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 即求 $EF+FG+GH+HE$. 为此, 我们要建立截面边长与底面边长的定量关系. 题设截面 $EFGH$ 与底面 $ABCD$ 平行, 即

$$\text{平面}EFGH \parallel \text{平面}ABCD,$$

这正好可以应用面面平行的性质定理.

侧面 PAB 与二平行平面相交, 与 $EFGH$ 交于 EF, 与 $ABCD$ 交于 AB, 据定理 $EF \parallel AB$. 在 $\triangle PAB$ 中, 由 $EF \parallel AB$ 得 $\triangle PEF \sim \triangle PAB$, 故

$$\frac{EF}{AB}=\frac{PE}{PA}.$$

题设 $PE:EA=2:1$, 即 $PA=PE+EA$ 合 3 份而 PE 占 2 份, 故 $\frac{PE}{PA}=\frac{2}{3}$. 结合 $AB=9$,

$$EF=\frac{2}{3}AB=\frac{2}{3}\times 9=6.$$

侧面 PBC 同样截二平行平面, 交线为 FG 与 BC, 故 $FG \parallel BC$ 且 $\triangle PFG \sim \triangle PBC$. 平行平面族截过同一点的直线族, 所得线段成比例. 在 $\triangle PAB$ 中, 由 $EF \parallel AB$ 已得 $\frac{PF}{PB}=\frac{PE}{PA}=\frac{2}{3}$, 故

$$FG=\frac{2}{3}BC=\frac{2}{3}\times 6=4.$$

底面 $ABCD$ 是平行四边形, 截面 $EFGH$ 亦为平行四边形, 故 $GH=EF=6$, $HE=FG=4$.

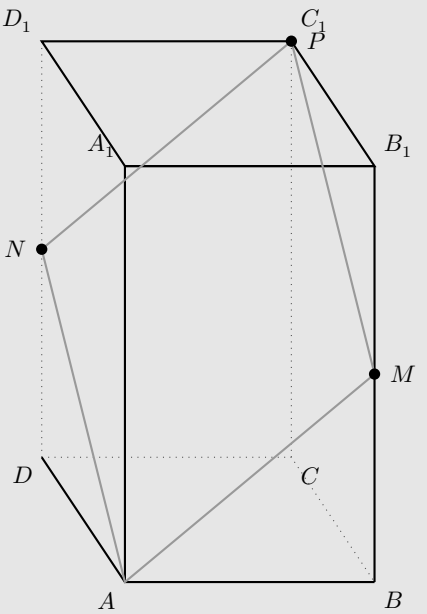
各边相加, 截面四边形 $EFGH$ 的周长为

$$C_{EFGH}=EF+FG+GH+HE=6+4+6+4=20.$$

□

例题 21.11 在长方体 $ABCD-A_1B_1C_1D_1$ 中,已知 $AB=3, AD=4, AA_1=5$. 点 M,N 分别是棱 BB_1, DD_1 的中点. 一个过点 A,M,N 的平面 α 与棱 CC_1 交于点 P. 求线段 MP 的长度.

知乎@余命数



证明 即求 MP 的长度, 我们把它放进一个可计算的平面图形中. 长方体的侧面 ADD_1A_1 与侧面 BCC_1B_1 平行, 即

$$\text{平面}ADD_1A_1 \parallel \text{平面}BCC_1B_1.$$

平面 α(即平面 AMN)与这两个平行侧面相交, 与 ADD_1A_1 交于 AN, 与 BCC_1B_1 交于 MP. 由平行关系可知交点实为 $P = C_1$, 即棱 CC_1 的端点. 根据面面平行的性质定理,

$$AN \parallel MP.$$

同理, 前后两侧面 ABB_1A_1 与 DCC_1D_1 平行, 平面 α 与之分别交于 AM 与 NP, 故

$$AM \parallel NP.$$

四边形 $AMNP$ 两组对边分别平行, 是平行四边形, 对边等长, 即

$$MP = AN.$$

空间长度问题转化为平面 ADD_1A_1 内的计算问题.

在矩形 ADD_1A_1 中求对角线 AN. 长方体的侧棱 DD_1 垂直于底面, 故 $DD_1 \perp AD$, $\triangle ADN$ 是以 $\angle D$ 为直角的直角三角形. 题设 $AD = 4$; 点 N 是棱 DD_1 的中点, 且 $DD_1 = AA_1 = 5$, 故 $DN = \frac{1}{2}DD_1 = \frac{5}{2}$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在 Rt△ADN 中应用勾股定理,

$$AN^2 = AD^2 + DN^2 = 4^2 + \left(\frac{5}{2}\right)^2 \quad \text{（勾股定理）}$$
$$= 16 + \frac{25}{4} = \frac{64+25}{4} = \frac{89}{4},$$

故

知乎@余命数

$$AN = \sqrt{\frac{89}{4}} = \frac{\sqrt{89}}{2}.$$

由 $MP = AN$ 得

$$MP = \frac{\sqrt{89}}{2}.$$

□

21.2.5 高考真题选讲

以下高考真题考查线线平行的判定与线面平行的证明.

例题 21.12 (2021 年浙江卷第 6 题) 已知正方体 $ABCD$-$A_1B_1C_1D_1$、M、N 分别是 A_1D, D_1B 的中点, 则

A. 直线 A_1D 与直线 D_1B 垂直, 直线 $MN \parallel$ 平面 $ABCD$

B. 直线 A_1D 与直线 D_1B 平行, 直线 $MN \perp$ 平面 BDD_1B_1

C. 直线 A_1D 与直线 D_1B 相交, 直线 $MN \parallel$ 平面 $ABCD$

D. 直线 A_1D 与直线 D_1B 异面, 直线 $MN \perp$ 平面 BDD_1B_1

解 答案: A. 连接 AD_1.$A_1D \perp AD_1$ 且 $A_1D \perp AB$, 故 $A_1D \perp$ 平面 ABD_1, 从而 $A_1D \perp D_1B$.

MN 是 $\triangle D_1AB$ 的中位线, $MN \parallel AB$. 由 $AB \subset$ 平面 $ABCD$, $MN \not\subset$ 平面 $ABCD$, 得 $MN \parallel$ 平面 $ABCD$.

$A_1 \notin$ 平面 BDD_1B_1 而 $D_1B \subset$ 平面 BDD_1B_1, 且 A_1D 与平面 BDD_1B_1 的交点 D 不在 D_1B 上, 故 A_1D 与 D_1B 异面, 排除 B、C.$MN \parallel AB$, AB 不垂直于平面 BDD_1B_1, 故 MN 也不垂直于该平面, 排除 D.

例题 21.13 (2020 年北京卷第 16 题) 在正方体 $ABCD$-$A_1B_1C_1D_1$ 中, E 为 BB_1 的中点.

1. 求证: $BC_1 \parallel$ 平面 AD_1E;
2. 求直线 AA_1 与平面 AD_1E 所成角的正弦值.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 (1) 由正方体的性质, $AB \parallel C_1D_1$ 且 $AB = C_1D_1$, 故四边形 ABC_1D_1 是平行四边形, $BC_1 \parallel AD_1$. 又 $BC_1 \not\subset$ 平面 AD_1E, $AD_1 \subset$ 平面 AD_1E, 依线面平行判定定理, $BC_1 \parallel$ 平面 AD_1E.

(2) 以 A 为原点, AB、AD、AA_1 分别为 x、y、z 轴建立空间直角坐标系. 设正方体棱长为 a, 则 $A(0,0,0)$、$A_1(0,0,a)$、$D_1(0,a,a)$、$E(a,0,\frac{a}{2})$.

$\overrightarrow{AA_1} = (0,0,a)$、$\overrightarrow{AD_1} = (0,a,a)$、$\overrightarrow{AE} = (a,0,\frac{a}{2})$.

设平面 AD_1E 的法向量 $\vec{n} = (x,y,z)$, 由 $\vec{n} \cdot \overrightarrow{AD_1} = 0$ 及 $\vec{n} \cdot \overrightarrow{AE} = 0$ 得

$$\begin{cases} y + z = 0, \\ x + \frac{1}{2}z = 0. \end{cases}$$

令 $z = 2$, 则 $\vec{n} = (-1,-2,2)$.

设直线 AA_1 与平面 AD_1E 所成角为 θ, 则

$$\sin\theta = \frac{|\vec{n} \cdot \overrightarrow{AA_1}|}{|\vec{n}|\,|\overrightarrow{AA_1}|} = \frac{2a}{3a} = \frac{2}{3}.$$

21.3 空间中的平行关系

平面几何中我们已经熟悉线线平行. 把它推广到空间, 要回答的是一条直线何时与一个平面平行, 以及两个平面何时相互平行. 我们先给出严格定义, 再为每种平行各配一个判定定理与一个性质定理.

21.3.1 直线与平面平行

直线与平面的位置关系有三种, 我们先研究其中的平行. 一条直线与一个平面平行, 直观上看, 这条直线无论怎样延伸都不会与该平面相遇. 用集合的语言说, 直线与平面作为点集, 其交集为空.

定义 21.4 (直线与平面平行) 若直线 l 与平面 α 的交集为空集, 即 $l \cap \alpha = \emptyset$, 则称直线 l 与平面 α **平行**, 记作 $l \parallel \alpha$.

直接用定义判定平行是困难的, 它要求我们验证直线与平面这两个无穷点集没有任何公共元素. 下面的判定定理把这件事简化, 只需在平面内找到一条与已知直线平行的直线.

定理 21.3 (线面平行判定定理) 若一个平面之外的一条直线与此平面内的某一条直线平行, 则该直线与此平面平行.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 设直线 $l \not\subset \alpha$, 且平面 α 内存在直线 m 使 $l \parallel m$, 即证 $l \cap \alpha = \emptyset$.

用反证法. 设 l 与 α 不平行, 则有公共点 P:

$$P \in l \quad \text{且} \quad P \in \alpha.$$

据平行直线的定义, l 与 m 共面, 设此唯一平面为 β. 由 $P \in l \subset \beta$ 知 $P \in \beta$; 又 $m \subset \alpha$ 且 $m \subset \beta$, 即 m 是 α 与 β 的一条公共直线.

点 P 同时属于 α, β, 故落在二者的交线上. 两平面若相交则交于唯一一条直线(此处 $l \not\subset \alpha$, 故 α, β 不重合), 而 m 已是其一公共直线, 故交线即 m, 从而 $P \in m$.

于是点 P 同时在 l 与 m 上, 与 $l \parallel m$(两平行线无公共点)矛盾. 故 l 与 α 无公共点, 即 $l \parallel \alpha$. □

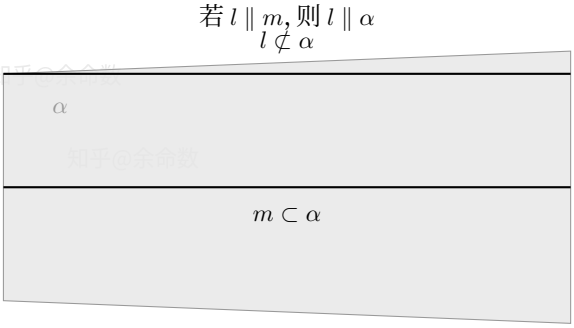
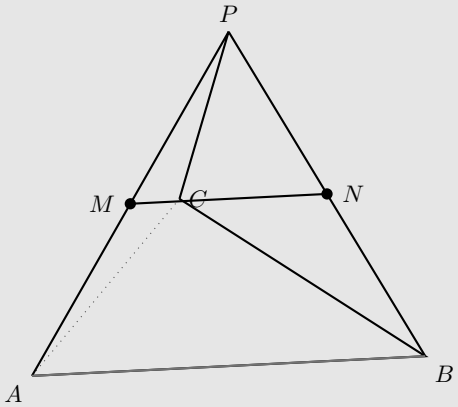


图 21.6　线面平行判定定理的几何直观

用这个定理证明线面平行, 归结为在平面内寻找一条与已知直线平行的直线.

例题 21.14　在四面体 $P-ABC$ 中, 设点 M, N 分别是棱 PA, PB 的中点. 求证: $MN \parallel$ 平面 ABC.



证明 即证直线 MN 与平面 ABC 平行. 据线面平行判定定理, 只需在平面 ABC 内找到一条与 MN 平行的直线.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

在 $\triangle PAB$ 中, M, N 分别为 PA, PB 的中点, 据三角形中位线定理, MN 平行于第三边 AB:

$$MN \parallel AB.$$

直线 AB 含于平面 ABC,而 MN 上的点 M, N 在棱 PA, PB 上,只要四点 P, A, B, C 不共面, MN 即不在平面 ABC 内.

平面 ABC 外的直线 MN 与该平面内直线 AB 平行, 判定定理的全部条件已满足, 故

$$MN \parallel \text{平面} ABC.$$

□

线面平行一经确认, 就蕴含进一步的几何事实. 下面的性质定理给出这种平行结构被其他平面截取时的规律.

定理 21.4 (线面平行性质定理) 若一条直线与一个平面平行, 则经过这条直线且与此平面相交的任意一个平面, 其交线与该直线平行.

证明 设直线 $l \parallel \alpha$. 又设平面 β 含直线 l ($l \subset \beta$) 且与 α 相交, 交线 m 即 $\alpha \cap \beta = m$, 即证 $l \parallel m$.

由交线定义, $m \subset \alpha$ 且 $m \subset \beta$; 又 $l \subset \beta$, 故 l, m 同在平面 β 内, 是共面直线. 欲证二者平行, 只需证其无公共点.

用反证法. 设 l 与 m 有公共点 Q. 由 $Q \in m \subset \alpha$ 得 $Q \in \alpha$; 此时点 Q 同时在直线 l 与平面 α 上, 即为二者的公共点, 与前提 $l \parallel \alpha$(无公共点)矛盾. 故 l 与 m 无公共点; 二者共面且无公共点, 故 $l \parallel m$. □

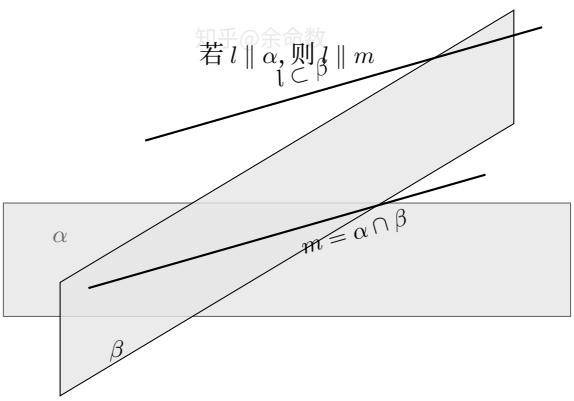
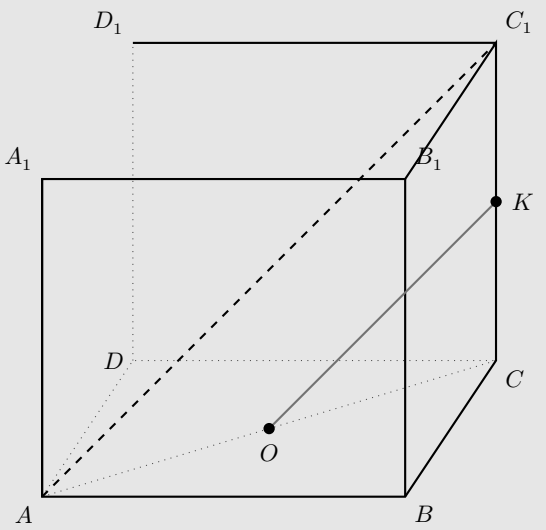


图 21.7 线面平行性质定理的几何直观

这个性质定理把空间中的平行关系转移到截面上, 化为线线平行的问题.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 21.15　在正方体 $ABCD-A_1B_1C_1D_1$ 中, 平面 α 经过对角线 BD 且平行于对角线 AC_1. 平面 α 与棱 CC_1 交于点 K. 求证 $CK=\frac{1}{2}CC_1$.



知乎@余命数

证明　我们要确定 K 在棱 CC_1 上的位置,依据是题给的线面平行条件. 已知 $AC_1\parallel\alpha$, 其中 α 即平面 BDK. 取含对角线 AC_1 与 AC 的对角面 ACC_1A_1 作过 AC_1 的平面.

设底面对角线 AC 与 BD 交于 O.O 在 AC 上, 故 $O\in$ 平面ACC_1A_1;O 在 BD 上, 故 $O\in$ 平面$BDK(\alpha)$.K 在 CC_1 上, 故 $K\in$ 平面ACC_1A_1;K 又在 α 上.O,K 同为 ACC_1A_1 与 α 的公共点, 连接二者即得交线 OK.

应用线面平行性质定理, 直线 $AC_1\parallel\alpha$, 平面 ACC_1A_1 经过 AC_1, 与 α 的交线为 OK, 据定理

$$OK\parallel AC_1.$$

空间问题遂归为平面 ACC_1A_1 内的平面几何问题. 在 $\triangle ACC_1$ 中, O 是对角线 AC 的中点; 由 $OK\parallel AC_1$, 据平行线分线段成比例定理的推论, 即三角形中位线的逆定理, OK 是 $\triangle ACC_1$ 的中位线.

故 K 是边 CC_1 的中点, 即 $CK=\frac{1}{2}CC_1$. □

21.3.2　平面与平面平行

平面内两条直线若不相交, 则必然平行. 把平面看作二维子空间并推广到三维, 两个平面的位置关系也只有两种, 相交或平行. 两个平面相互平行, 它们在空间中无限延伸却永不相交.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定义 21.5 (平面与平面平行) 若两个平面 α 与 β 的交集为空集, 即 $\alpha \cap \beta = \emptyset$, 则称这两个平面**相互平行**, 记作 $\alpha \parallel \beta$.

与线面平行一样, 直接用定义判定两个无限延展的平面没有公共点是困难的. 我们需要一个从有限构造出发的判定定理.

定理 21.5 (面面平行判定定理) 若一个平面内有两条相交直线分别平行于另一个平面, 则此二平面相互平行.

证明 设平面 α 内有两条相交于点 P 的直线 l_1, l_2, 且 $l_1 \parallel \beta, l_2 \parallel \beta$, 即证 $\alpha \parallel \beta$.

用反证法. 设 α 与 β 不平行, 则二者相交于一条直线 m.

对 l_1: 由 $l_1 \subset \alpha$ 且 $l_1 \parallel \beta$, 据线面平行性质定理, 任何含 l_1 的平面若与 β 相交, 交线必与 l_1 平行; α 即这样一个平面, 故交线 $m \parallel l_1$.

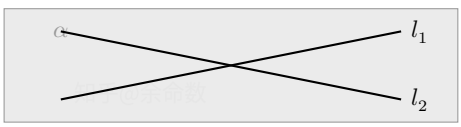
$$m \parallel l_1$$

对 l_2 同理: 由 $l_2 \subset \alpha$ 且 $l_2 \parallel \beta$, 据同一性质定理, $m \parallel l_2$.

$$m \parallel l_2$$

于是 $l_1 \parallel m$ 且 $l_2 \parallel m$, 即平面 α 内过同一点 P 的两条不同直线 l_1, l_2 都平行于同一直线 m, 违背平面几何的平行公理.

故假设"α 与 β 相交"不成立, 二平面相互平行. □



若 $l_1 \parallel \beta$ 且 $l_2 \parallel \beta$, 则 $\alpha \parallel \beta$



图 21.8 面面平行判定定理的几何模型

面面平行确立之后, 任何截面都在这两个平面上留下平行的交线. 这就是下面的性质定理.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 21.6 (面面平行性质定理) 若两个相互平行的平面被第三个平面所截，则它们的交线相互平行.

证明 设 $\alpha \parallel \beta$, 第三平面 γ 与 α, β 分别交于直线 l_1, l_2, 即 $l_1 = \alpha \cap \gamma, l_2 = \beta \cap \gamma$, 即证 $l_1 \parallel l_2$.

由交线定义, l_1, l_2 均在平面 γ 内, 是共面直线. 再证其无公共点, 用反证法, 设 l_1, l_2 有公共点 P. 由 $P \in l_1 \subset \alpha$ 得 $P \in \alpha$; 由 $P \in l_2 \subset \beta$ 得 $P \in \beta$. 即点 P 是 α, β 的公共点, 与前提 $\alpha \parallel \beta$(无公共点)矛盾.

故 l_1, l_2 无公共点. 二者共面且无公共点, 故相互平行. □

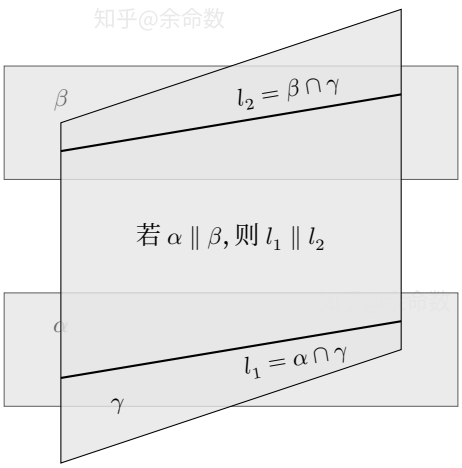


图 21.9　面面平行性质定理的几何模型

线线平行有传递性, 面面平行也有.

定理 21.7 (平行平面的传递性) 平行于同一个平面的两个不同平面相互平行.

证明 设平面 $\alpha \parallel \gamma$ 且平面 $\beta \parallel \gamma$, 且 $\alpha \neq \beta$, 即证 $\alpha \cap \beta = \emptyset$.

我们用坐标法证明. 取平面 γ 内两条相交直线的方向作为前两个坐标方向, 并取垂直于 γ 的方向为第三坐标方向. 在这样的空间直角坐标系中, 平面 γ 可写为 $z = 0$. 由于 $\alpha \parallel \gamma$, 平面 α 上所有点到 γ 的有向距离相同, 故它可写为 $z = c_1$; 同理 β 可写为 $z = c_2$. 若 $c_1 = c_2$, 则两平面重合, 与 $\alpha \neq \beta$ 矛盾. 因此 $c_1 \neq c_2$. 方程 $z = c_1$ 与 $z = c_2$ 无公共解, 所以 $\alpha \cap \beta = \emptyset$, 即 $\alpha \parallel \beta$. □

21.3.3 高考真题选讲

以下高考真题考查平面平行的充要条件及线面平行的条件关系.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 21.16 (2019 年全国 □ 卷第 7 题) 设 α, β 为两个平面, 则 $\alpha \parallel \beta$ 的充要条件是

A. α 内有无数条直线与 β 平行　　C. α, β 平行于同一条直线

B. α 内有两条相交直线与 β 平行　　D. α, β 垂直于同一平面

解 答案: B. 一个平面内有无数条直线与另一平面平行时, 两平面可能相交 (交线以外的直线都平行于另一平面), A 不充分. 一个平面内有两条相交直线与另一平面平行时, 由面面平行判定定理两平面平行, B 充分且必要. 两平面平行于同一条直线时仍可能相交, C 不充分. 两平面垂直于同一平面时也可能相交, D 不充分.

例题 21.17 (2018 年浙江卷第 6 题) 已知平面 α, 直线 m, n 满足 $m \not\subset \alpha, n \subset \alpha$, 则"$m \parallel n$"是"$m \parallel \alpha$"的

A. 充分不必要条件　　C. 充分必要条件

B. 必要不充分条件　　D. 既不充分也不必要条件

解 答案: A. 由 $m \not\subset \alpha$, $n \subset \alpha$, 若 $m \parallel n$, 依线面平行判定定理得 $m \parallel \alpha$, 充分性成立. 反之, $m \parallel \alpha$ 时 m 平行于 α 内无数条直线, 但不一定平行于特定的 n, 必要性不成立.

21.4 空间中的垂直关系

上一节研究了平行, 本节研究垂直. 我们仍按定义、判定定理、性质定理的顺序, 依次讨论直线与平面垂直、平面与平面垂直, 中间插入一个关于斜线与射影的定理.

21.4.1 直线与平面垂直

定义 21.6 若一条直线的方向与一个平面内任意直线的方向都垂直, 则称这条直线与这个平面互相**垂直**.

这个定义要求检验平面内的每一条直线, 直接用它判定并不方便. 下面的判定定理把检验对象缩减为两条相交直线.

定理 21.8 (线面垂直判定定理) 若一条直线与一个平面内的两条相交直线都垂直, 则该直线与此平面垂直.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 设平面 α 内两条相交直线 m_1, m_2 的方向向量分别为 $\mathbf{u}_1, \mathbf{u}_2$, 直线 l 的方向向量为 $\mathbf{v}$. 因 m_1, m_2 相交, 它们不平行, 故 $\mathbf{u}_1, \mathbf{u}_2$ 张成平面 α 内的全部方向. 由题设 $l \perp m_1$ 且 $l \perp m_2$, 得

$$\mathbf{v} \cdot \mathbf{u}_1 = 0, \qquad \mathbf{v} \cdot \mathbf{u}_2 = 0.$$

对平面 α 内任意一条直线 m, 其方向向量可写成 $\mathbf{u} = a\mathbf{u}_1 + b\mathbf{u}_2$. 于是

$$\mathbf{v} \cdot \mathbf{u} = a(\mathbf{v} \cdot \mathbf{u}_1) + b(\mathbf{v} \cdot \mathbf{u}_2) = 0.$$

知乎@余命数

因此 l 与平面 α 内任意直线的方向都垂直, 按定义 $l \perp \alpha$. □

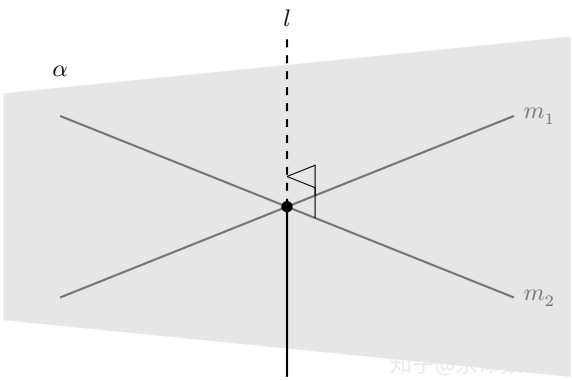


图 21.10　线面垂直判定定理

21.4.2 三垂线定理

设 PO 是平面的一条斜线, $P'O$ 是它在平面内的射影. 平面内一条直线若与射影垂直, 按三垂线定理, 它与斜线本身也垂直.

> **定理 21.9 (三垂线定理)** 在平面内的一条直线, 若它与此平面的一条斜线的射影垂直, 则它也与这条斜线垂直.

证明 设 l 是平面 α 内的一条直线, PO 是平面 α 的一条斜线, $P'O$ 是其在 α 内的射影, 且 $l \perp P'O$, 即证 $l \perp PO$.

知乎@余命数

由射影定义, $PP' \perp \alpha$. 取 $\mathbf{n} = \overrightarrow{P'P}$, $\mathbf{q} = \overrightarrow{P'O}$, 取 $\mathbf{u}$ 为直线 l 的方向向量. 由 $l \subset \alpha$ 与 $PP' \perp \alpha$, 得 $\mathbf{u} \cdot \mathbf{n} = 0$; 又由题设 $l \perp P'O$, 得 $\mathbf{u} \cdot \mathbf{q} = 0$. 斜线 PO 的方向向量

$$\overrightarrow{PO} = \overrightarrow{P'O} - \overrightarrow{P'P} = \mathbf{q} - \mathbf{n}.$$

故

$$\mathbf{u} \cdot \overrightarrow{PO} = \mathbf{u} \cdot \mathbf{q} - \mathbf{u} \cdot \mathbf{n} = 0,$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即 $l \perp PO$. □

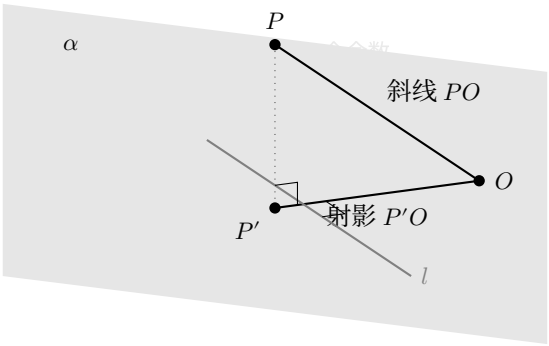


图 21.11 三垂线定理

21.4.3 平面与平面垂直

定义 21.7 若两个相交平面所成的二面角是直二面角, 则称这两个平面互相**垂直**.

这个定义用二面角表述, 直接检验二面角是否为直角并不方便. 下面给出一个更直接的判定方法.

定理 21.10 (面面垂直判定定理) 若一个平面经过另一个平面的一条垂线, 则这两个平面互相垂直.

证明 设平面 β 经过平面 α 的一条垂线 l, 且 $\alpha \cap \beta = m$, 即证 $\alpha \perp \beta$.

设 $l \cap \alpha = O$, 则 O 必在交线 m 上. 在平面 α 内作直线 $n \perp m$ 于点 O. 由 $l \perp \alpha$ 及线面垂直的性质, l 与 α 内任意直线都垂直, 故 $l \perp m$ 且 $l \perp n$.

角 $\angle(l, n)$ 是两平面所成二面角的平面角. 由 $l \perp n$, 该平面角为 $90°$, 据定义两平面互相垂直. □

定理 21.11 (面面垂直性质定理) 若两个平面互相垂直, 则在一个平面内垂直于它们交线的直线垂直于另一个平面.

证明 设 $\alpha \perp \beta$ 且 $\alpha \cap \beta = m$, 直线 $l \subset \alpha$ 且 $l \perp m$ 于点 O, 即证 $l \perp \beta$.

在平面 β 内作直线 $n \perp m$ 于点 O. 由 $\alpha \perp \beta, l, n$ 构成的二面角的平面角为 $90°$, 即 $l \perp n$(面面垂直定义). 又题设 $l \perp m.m, n$ 是平面 β 内两条相交直线, l 同时与二者垂直, 据线面垂直判定定理, $l \perp \beta$. □

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

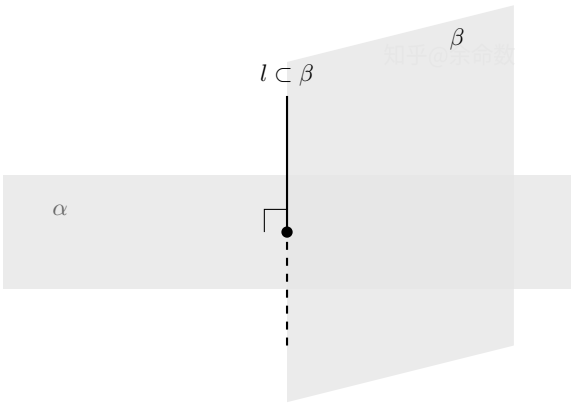


图 21.12　面面垂直判定定理

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

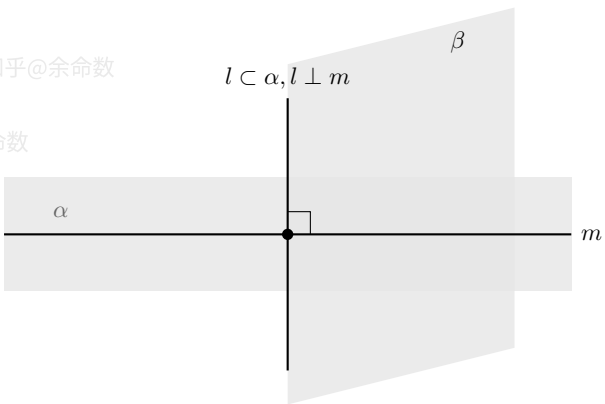


图 21.13　面面垂直性质定理

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

21.4.4 高考真题选讲

以下高考真题考查线面垂直与面面垂直的判定及性质.

例题 21.18 (2017 年全国 Ⅱ 卷文科数学第 10 题) 在正方体 $ABCD$-$A_1B_1C_1D_1$ 中, E 为棱 CD 的中点, 则

A. $A_1E \perp DC_1$　　C. $A_1E \perp BC_1$

B. $A_1E \perp BD$　　D. $A_1E \perp AC$

解 答案: C. 连接 B_1C. 正方体中 $BC_1 \perp B_1C$ 且 $A_1B_1 \perp B_1C$, 故 $BC_1 \perp$ 平面 A_1B_1CE.$A_1E \subset$ 平面 A_1B_1CE, 故 $A_1E \perp BC_1$.

例题 21.19 (2021 年全国 Ⅱ 卷第 20 题) 在三棱锥 A-BCD 中, 平面 $ABD \perp$ 平面 BCD、$AB = AD$、O 为 BD 的中点.

1. 求证: $OA \perp CD$;
2. 若 $\triangle OCD$ 是边长为 1 的等边三角形, 点 E 在棱 AD 上, $DE = 2EA$, 且二面角 E-BC-D 的大小为 45°, 求三棱锥 A-BCD 的体积.

解 (1) 由 $AB = AD$ 及 O 为 BD 中点, 得 $AO \perp BD$. 平面 $ABD \perp$ 平面 BCD, 交线为 BD, $AO \subset$ 平面 ABD, 故 $AO \perp$ 平面 BCD.$CD \subset$ 平面 BCD, 故 $AO \perp CD$.

(2) 以 O 为原点, OD 为 y 轴, OA 为 z 轴, 过 O 平行于 $\triangle OCD$ 中 OD 边上的高为 x 轴建系. 设 $OA = t$, 则 $B(0,-1,0)$、$C(\frac{\sqrt{3}}{2},\frac{1}{2},0)$、$D(0,1,0)$、$A(0,0,t)$、$E(0,\frac{1}{3},\frac{2t}{3})$.

平面 BCD 的法向量为 $\overrightarrow{OA} = (0,0,t)$. 设平面 BCE 的法向量 $\vec{n} = (x,y,z)$, 由 $\vec{n} \cdot \overrightarrow{BC} = 0$ 及 $\vec{n} \cdot \overrightarrow{BE} = 0$ 得

$$\begin{cases} \frac{\sqrt{3}}{2}x + \frac{3}{2}y = 0, \\ \frac{4}{3}y + \frac{2t}{3}z = 0. \end{cases}$$

令 $x = \sqrt{3}$, 则 $y = -1$、$z = \frac{2}{t}$、$\vec{n} = (\sqrt{3},-1,\frac{2}{t})$.

二面角 E-BC-D 为 45°, 故

$$\frac{|\vec{n} \cdot \overrightarrow{OA}|}{|\vec{n}|\,|\overrightarrow{OA}|} = \frac{2}{t\sqrt{4+4/t^2}} = \frac{1}{\sqrt{t^2+1}} = \frac{\sqrt{2}}{2},$$

解得 $t = 1$, 即 $OA = 1$.

$S_{\triangle BCD} = \frac{1}{2} \times BD \times h_C = \frac{1}{2} \times 2 \times \frac{\sqrt{3}}{2} = \frac{\sqrt{3}}{2}$, 故

$$V_{A\text{-}BCD} = \frac{1}{3} \cdot \frac{\sqrt{3}}{2} \cdot 1 = \frac{\sqrt{3}}{6}.$$

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 21.20 (2018 年全国 □ 卷理科数学第 18 题) 四边形 $ABCD$ 为正方形, E, F 分别为 AD, BC 的中点, 以 DF 为折痕把 $\triangle DFC$ 折起, 使点 C 到达点 P 的位置, 且 $PF \perp BF$.

1. 求证: 平面 $PEF \perp$ 平面 $ABFD$;
2. 求 DP 与平面 $ABFD$ 所成角的正弦值.

解 (1) 在正方形中 $EF \parallel AB$, 故 $EF \perp BC$, $EF \perp BF$. 又 $PF \perp BF$, $EF \cap PF = F$, 故 $BF \perp$ 平面 PEF. $BF \subset$ 平面 $ABFD$, 故平面 $PEF \perp$ 平面 $ABFD$.

(2) 设正方形边长为 $2a$. 折叠保持 $\triangle DFP \cong \triangle DFC$, 故 $PD = CD = 2a$、$PF = CF = a$、$DE = a$. 由 $BF \perp$ 平面 PEF 及 $DE \parallel BF$, 得 $DE \perp$ 平面 PEF, 故 $DE \perp EP$. 在 $\triangle PDE$ 中, $PE = \sqrt{PD^2 - DE^2} = \sqrt{3}\,a$.

在平面 PEF 中过 P 作 $PH \perp EF$ 于 H. 由平面 $PEF \perp$ 平面 $ABFD$, 交线为 EF, 得 $PH \perp$ 平面 $ABFD$, $\angle PDH$ 即为所求角.

$PF \perp DE (DE \parallel BF, PF \perp BF)$ 且 $PF \perp PD (\angle FPD = \angle FCD = 90^\circ)$, 故 $PF \perp$ 平面 PDE, $V_{F\text{-}PDE} = \frac{1}{3} PF \cdot S_{\triangle PDE} = \frac{1}{3} \cdot a \cdot \frac{\sqrt{3}}{2} a^2 = \frac{\sqrt{3}}{6} a^3$.

又 $S_{\triangle DEF} = \frac{1}{2} \cdot DE \cdot EF = \frac{1}{2} \cdot a \cdot 2a = a^2$, 由 $V_{P\text{-}DEF} = \frac{1}{3} PH \cdot S_{\triangle DEF}$ 得 $PH = \frac{\sqrt{3}}{2} a$.

$$\sin \angle PDH = \frac{PH}{PD} = \frac{\frac{\sqrt{3}}{2} a}{2a} = \frac{\sqrt{3}}{4}.$$

知乎@余命数

21.5 空间向量与坐标法

前几节用公理与判定定理处理位置关系, 还有一类问题要求夹角、距离与轨迹, 纯综合的路往往很绕. **空间向量与坐标法**把这类问题化为坐标计算. 平面向量已在 第11章 中建立, 本节把它推广到三维, 并固定一套建立坐标系的习惯.

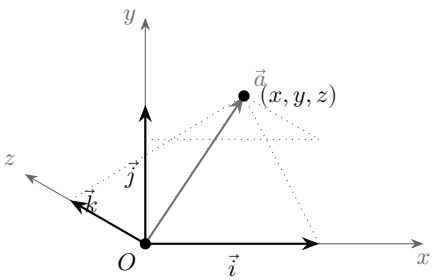


图 21.14　空间标准正交基与向量 $\vec{a} = x\vec{i} + y\vec{j} + z\vec{k}$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

21.5.1 从平面到空间

定义 21.8 (空间向量) 空间中的向量仍是有大小、有方向的量. 取定点 O 为原点后, 任一向量 $\vec{a}$ 可唯一写成

$$\vec{a} = x\,\vec{i} + y\,\vec{j} + z\,\vec{k},$$

其中 $\vec{i},\vec{j},\vec{k}$ 为两两垂直的单位向量 (空间标准正交基), 有序三元组 (x,y,z) 称为 $\vec{a}$ 的**坐标**.

加法与数乘仍按分量进行.

$$(x_1,y_1,z_1)+(x_2,y_2,z_2)=(x_1+x_2,y_1+y_2,z_1+z_2),\quad \lambda(x,y,z)=(\lambda x,\lambda y,\lambda z).$$

模长

$$|\vec{a}| = \sqrt{x^2+y^2+z^2}.$$

定义 21.9 (数量积) 设 $\vec{a}=(x_1,y_1,z_1),\vec{b}=(x_2,y_2,z_2)$, 定义

$$\vec{a}\cdot\vec{b} = x_1x_2+y_1y_2+z_1z_2 = |\vec{a}||\vec{b}|\cos\theta,$$

其中 θ 为两向量的夹角. 因而

$$\vec{a}\perp\vec{b} \iff \vec{a}\cdot\vec{b}=0,\qquad \cos\theta=\frac{\vec{a}\cdot\vec{b}}{|\vec{a}||\vec{b}|}\quad(|\vec{a}||\vec{b}|\neq 0).$$

注 21.1 (与平面向量的衔接)
平面向量的平行四边形法则、定比分点、共线判定在空间中同样成立. 共面则多一个条件. $\vec{a},\vec{b},\vec{c}$ 共面当且仅当存在不全为零的实数 λ,μ,ν 使 $\lambda\vec{a}+\mu\vec{b}+\nu\vec{c}=\vec{0}$(三维中亦可用混合积为零刻画, 见下一节).

21.5.2 直线与平面的方向数据

- 与直线平行的任一非零向量称为直线的**方向向量**.
- 与平面垂直的任一非零向量称为平面的**法向量**. 若平面内两不共线向量为 $\vec{u},\vec{v}$, 则可选

$$\vec{n}=\vec{u}\times\vec{v}$$

(叉积/外积; 坐标公式见下一节体积讨论, 或直接用垂直于 $\vec{u},\vec{v}$ 的向量求解线性方程).

知乎@余命数

于是位置关系可翻译为向量条件.

线线平行	方向向量平行(对应分量成比例)
线线垂直	方向向量数量积为零
线面平行	方向向量 $\cdot$ 法向量 $=0$, 且直线不在平面内
线面垂直	方向向量 $\parallel$ 法向量
面面平行	法向量平行
面面垂直	法向量垂直(数量积为零)

这些向量条件与前面公理化的判定定理**结论一致**; 向量语言把位置关系的证明变成计算.

知乎@余命数

21.5.3 空间直角坐标系与建系

定义 21.10 (空间直角坐标系) 取两两垂直的坐标轴 Ox, Oy, Oz, 单位长度相同, 构成**空间直角坐标系** O-xyz. 点 P 的坐标 (x,y,z) 满足 $\overrightarrow{OP}=(x,y,z)$.

建系的常见选择. 原则是让尽可能多的棱落在坐标轴或坐标平面上.

1. 正方体或长方体, 取一个顶点为原点, 三条棱为轴.
2. 正四面体或正三棱锥, 把底面放在 xOy 平面, 对称轴对齐 z 轴.
3. 有互相垂直的棱时, 把这组棱取为坐标轴方向.

知乎@余命数　知乎@余命数　知乎@余命数

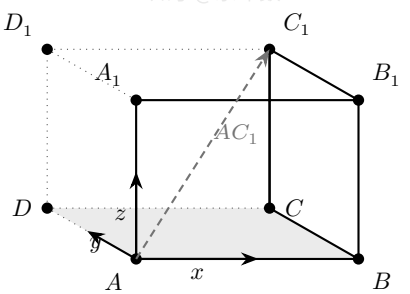


图 21.15　长方体建系: 以 A 为原点, 三棱为轴; 虚线为体对角线 AC_1

例题 21.21 (长方体中的线面角) 长方体 $ABCD$-$A_1B_1C_1D_1$ 中 $AB=a$、$AD=b$、$AA_1=c$. 求直线 AC_1 与平面 $ABCD$ 所成角的正弦.

解 建系. 令 A 为原点, AB, AD, AA_1 分别为 x,y,z 轴正方向. 则

$$A=(0,0,0),\quad C=(a,b,0),\quad C_1=(a,b,c).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

平面 $ABCD$ 的法向量可取 $\vec{n}=(0,0,1)$, 直线 AC_1 的方向向量 $\vec{v}=\overrightarrow{AC_1}=(a,b,c)$. 线面角 θ 满足

$$\sin\theta=\left|\frac{\vec{v}\cdot\vec{n}}{|\vec{v}||\vec{n}|}\right|=\frac{|c|}{\sqrt{a^2+b^2+c^2}}.$$

例题 21.22 (线线夹角的向量计算) 正方体 $ABCD$-$A_1B_1C_1D_1$ 棱长为 1. 判断 A_1C 与 BD_1 是否垂直, 并求这两条异面直线所成角的余弦.

知乎@余命数

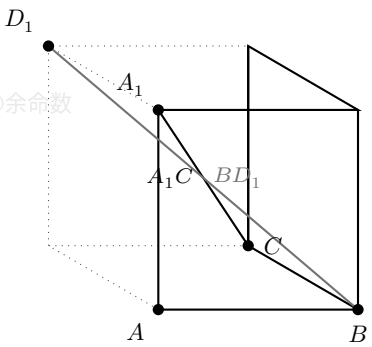


图 21.16 正方体中异面直线 A_1C 与 BD_1

解 建系同前, A 为原点, AB, AD, AA_1 为轴. 则

$$A_1=(0,0,1),\ C=(1,1,0),\ B=(1,0,0),\ D_1=(0,1,1).$$

方向向量

$$\overrightarrow{A_1C}=(1,1,-1),\qquad \overrightarrow{BD_1}=(-1,1,1).$$

数量积

$$\overrightarrow{A_1C}\cdot\overrightarrow{BD_1}=-1+1-1=-1\neq 0,$$

故**不垂直**. 夹角 θ 满足

$$\cos\theta=\frac{-1}{\sqrt{3}\cdot\sqrt{3}}=-\frac{1}{3},$$

异面直线所成角取锐角(与方向向量夹角或其补角中的锐角相同), 其余弦为 $\frac{1}{3}$.

例题 21.23 (法向量求二面角) 在棱长为 2 的正方体中, 求平面 A_1BC_1 与底面 $ABCD$ 所成二面角(锐角)的余弦.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 建系：$A=(0,0,0)$、$B=(2,0,0)$、$D=(0,2,0)$、$A_1=(0,0,2)$、$C_1=(2,2,2)$. 底面法向量 $\vec{n}_1=(0,0,1)$. 平面 A_1BC_1 内取

$$\overrightarrow{A_1B}=(2,0,-2),\qquad \overrightarrow{A_1C_1}=(2,2,0).$$

法向量 $\vec{n}_2=\overrightarrow{A_1B}\times\overrightarrow{A_1C_1}$：

$$\vec{n}_2=\big(0\cdot 0-(-2)\cdot 2,\ (-2)\cdot 2-2\cdot 0,\ 2\cdot 2-2\cdot 0\big)=(4,-4,4),$$

可取 $\vec{n}_2=(1,-1,1)$. 则

$$\cos\langle\vec{n}_1,\vec{n}_2\rangle=\frac{1}{\sqrt{3}}.$$

知乎@余命数

两平面法向量夹角等于二面角或其补角. 设 $\alpha=\langle\vec{n}_1,\vec{n}_2\rangle$，则 $\cos\alpha=\dfrac{1}{\sqrt{3}}$. 锐二面角 φ 取 α 与 $\pi-\alpha$ 中的锐角，故

$$\cos\varphi=\frac{1}{\sqrt{3}}.$$

注 21.2 (综合法与坐标法)
综合法直接调用判定定理与性质定理；坐标法把同一问题化为代数计算. 上面的第二、三例显示，数量积可以快速否定看似垂直的猜测，法向量把二面角化成向量的夹角.

21.5.4 高考真题选讲

以下高考真题考查空间直角坐标系的建立及向量法求线面角与二面角.

例题 21.24 (2021 年全国 □ 卷第 19 题) 在四棱锥 Q-$ABCD$ 中，底面 $ABCD$ 是正方形，$AD=2$、$QD=QA=\sqrt{5}$、$QC=3$.

1. 求证：平面 $QAD\perp$ 平面 $ABCD$；
2. 求二面角 B-QD-A 的平面角的余弦值.

解 (1) 取 AD 中点 O，连接 QO，CO. 由 $QA=QD$ 得 $QO\perp AD$，$QO=\sqrt{5-1}=2$. 在正方形中 $CO=\sqrt{1+4}=\sqrt{5}$.$QC^2=9=4+5=QO^2+OC^2$，故 $QO\perp OC$.$OC\cap AD=O$，故 $QO\perp$ 平面 $ABCD$.$QO\subset$ 平面 QAD，故平面 $QAD\perp$ 平面 $ABCD$.

(2) 以 O 为原点，过 O 平行于 CD 的方向为 x 轴，OD 为 y 轴，OQ 为 z 轴建系. 则 $D(0,1,0)$、$Q(0,0,2)$、$B(2,-1,0)$、$\overrightarrow{BQ}=(-2,1,2)$、$\overrightarrow{BD}=(-2,2,0)$.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

设平面 QBD 的法向量 $\vec{n}=(x,y,z)$，由 $\vec{n}\cdot\overrightarrow{BQ}=0$ 及 $\vec{n}\cdot\overrightarrow{BD}=0$ 得

$$\begin{cases}-2x+y+2z=0,\\-2x+2y=0.\end{cases}$$

取 $x=1$，则 $\vec{n}=(1,1,\frac{1}{2})$.

平面 QAD 的法向量 $\vec{m}=(1,0,0)$，故

$$\cos\langle\vec{m},\vec{n}\rangle=\frac{1}{1\times\frac{3}{2}}=\frac{2}{3}.$$

二面角 B-QD-A 为锐角，余弦值为 $\frac{2}{3}$.

例题 21.25 (2021 年天津卷第 17 题) 在棱长为 2 的正方体 $ABCD$-$A_1B_1C_1D_1$ 中，E 为棱 BC 的中点，F 为棱 CD 的中点.

1. 求证：$D_1F\parallel$ 平面 A_1EC_1；
2. 求直线 AC_1 与平面 A_1EC_1 所成角的正弦值；
3. 求二面角 A-A_1C_1-E 的正弦值.

解 以 A 为原点，AB、AD、AA_1 分别为 x、y、z 轴建系. 则 $A(0,0,0)$、$A_1(0,0,2)$、$C_1(2,2,2)$、$D_1(0,2,2)$、$E(2,1,0)$、$F(1,2,0)$.

(1) $\overrightarrow{D_1F}=(1,0,-2)$、$\overrightarrow{A_1C_1}=(2,2,0)$、$\overrightarrow{A_1E}=(2,1,-2)$. 设平面 A_1EC_1 的法向量 $\vec{m}=(x,y,z)$，由 $\vec{m}\cdot\overrightarrow{A_1C_1}=0$ 及 $\vec{m}\cdot\overrightarrow{A_1E}=0$ 得

$$\begin{cases}2x+2y=0,\\2x+y-2z=0.\end{cases}$$

令 $x=2$，则 $\vec{m}=(2,-2,1)$.$\overrightarrow{D_1F}\cdot\vec{m}=2-2=0$，且 $D_1F\not\subset$ 平面 A_1EC_1，故 $D_1F\parallel$ 平面 A_1EC_1.

(2) $\overrightarrow{AC_1}=(2,2,2)$、$|\vec{m}|=3$、$|\overrightarrow{AC_1}|=2\sqrt{3}$、$\vec{m}\cdot\overrightarrow{AC_1}=4-4+2=2$.

$$\sin\theta=\frac{|\vec{m}\cdot\overrightarrow{AC_1}|}{|\vec{m}|\,|\overrightarrow{AC_1}|}=\frac{2}{6\sqrt{3}}=\frac{\sqrt{3}}{9}.$$

(3) 平面 AA_1C_1 的法向量 $\overrightarrow{DB}=(2,-2,0)$.

$$\cos\langle\overrightarrow{DB},\vec{m}\rangle=\frac{4+4+0}{2\sqrt{2}\times3}=\frac{2\sqrt{2}}{3},$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$\sin\varphi=\sqrt{1-\frac{8}{9}}=\frac{1}{3}.$$

例题 21.26 (2021 年全国甲卷理科数学第 19 题) 已知直三棱柱 $ABC\text{-}A_1B_1C_1$ 中, 侧面 AA_1B_1B 为正方形, $AB=BC=2$、E、F 分别为 AC 和 CC_1 的中点, D 为棱 A_1B_1 上的点, $BF\perp A_1B_1$.

1. 求证: $BF\perp DE$;
2. 当 B_1D 为何值时, 平面 BB_1C_1C 与平面 DFE 所成二面角的正弦值最小?

解 (1) 连接 AF.F 为 CC_1 中点, $CF=1$, $BF=\sqrt{5}$. 由 $BF\perp A_1B_1$ 及 $AB\parallel A_1B_1$, 得 $BF\perp AB$, 故 $AF=\sqrt{4+5}=3$, $AC=\sqrt{9-1}=2\sqrt{2}$.$AC^2=AB^2+BC^2$, 故 $BA\perp BC$.

以 B 为原点, BA、BC、BB_1 分别为 x、y、z 轴建系. 则 $A(2,0,0)$、$C(0,2,0)$、$E(1,1,0)$、$F(0,2,1)$. 设 $B_1D=m$, 则 $D(m,0,2)$.

$\overrightarrow{BF}=(0,2,1)$、$\overrightarrow{DE}=(1-m,1,-2)$、$\overrightarrow{BF}\cdot\overrightarrow{DE}=0+2-2=0$, 故 $BF\perp DE$.

(2) 平面 BB_1C_1C 的法向量 $\vec{m}=(1,0,0)$.$\overrightarrow{DE}=(1-m,1,-2)$, $\overrightarrow{EF}=(-1,1,1)$. 设平面 DEF 的法向量 $\vec{n}=(x,y,z)$, 由

$$\begin{cases}(1-m)x+y-2z=0,\\ -x+y+z=0\end{cases}$$

令 $x=3$, 则 $y=m+1$、$z=2-m$、$\vec{n}=(3,m+1,2-m)$.

$$\cos\langle\vec{m},\vec{n}\rangle=\frac{3}{\sqrt{9+(m+1)^2+(2-m)^2}}=\frac{3}{\sqrt{2m^2-2m+14}}=\frac{3}{\sqrt{2(m-\frac{1}{2})^2+\frac{27}{2}}}.$$

当 $m=\frac{1}{2}$ 时分母最小, 余弦最大, 正弦最小. 故 $B_1D=\frac{1}{2}$ 时二面角的正弦值最小.

21.6 空间中的度量

前面几节讨论平行与垂直, 现在讨论的是多远、多斜、多大. 这类问题称为**度量**. 我们用 第21.5节 的向量与坐标语言, 给出距离、夹角与体积的公式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

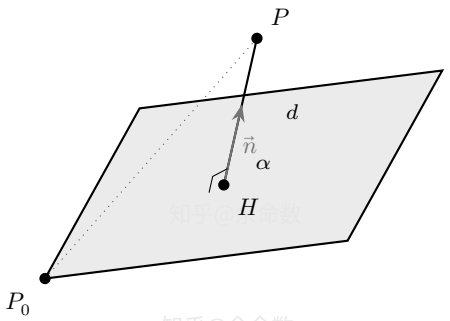


图 21.17 点 P 到平面 α 的距离：垂线 $PH\perp\alpha$，方向沿法向量 $\vec{n}$

21.6.1 距离

定理 21.12 (点到平面的距离) 平面 α 过点 P_0，法向量为 $\vec{n}$；点 P 到 α 的距离为

$$d=\frac{|\overrightarrow{P_0P}\cdot\vec{n}|}{|\vec{n}|}.$$

若平面方程为 $Ax+By+Cz+D=0$，点 $P(x_0,y_0,z_0)$，则

$$d=\frac{|Ax_0+By_0+Cz_0+D|}{\sqrt{A^2+B^2+C^2}}.$$

证明 设垂足为 H，则 $\overrightarrow{HP}$ 与 $\vec{n}$ 共线，且 $\overrightarrow{P_0P}=\overrightarrow{P_0H}+\overrightarrow{HP}$. 两边点乘 $\vec{n}$，由 $\overrightarrow{P_0H}\subset\alpha$ 而 $\vec{n}\perp\alpha$ 知 $\overrightarrow{P_0H}\cdot\vec{n}=0$，故 $\overrightarrow{P_0P}\cdot\vec{n}=\overrightarrow{HP}\cdot\vec{n}$，因而 $|\overrightarrow{HP}|=\dfrac{|\overrightarrow{P_0P}\cdot\vec{n}|}{|\vec{n}|}$. 坐标形式由 $\vec{n}=(A,B,C)$ 与点法式方程直接得到. □

定理 21.13 (异面直线的距离 (方向向量已知)) 设直线 l_1 过 A、方向 $\vec{d}_1$，直线 l_2 过 B、方向 $\vec{d}_2$，且 $\vec{d}_1\nparallel\vec{d}_2$. 则异面距离

$$d=\frac{|(\overrightarrow{AB},\vec{d}_1,\vec{d}_2)|}{|\vec{d}_1\times\vec{d}_2|},$$

其中分子为混合积的绝对值 $|\overrightarrow{AB}\cdot(\vec{d}_1\times\vec{d}_2)|$.

证明 l_1 与 l_2 的公垂线方向垂直于 $\vec{d}_1$ 与 $\vec{d}_2$，故平行于 $\vec{d}_1\times\vec{d}_2$. 异面距离等于 $\overrightarrow{AB}$ 在公垂线方向上的投影长度. 取单位公垂向量 $\vec{n}=\vec{d}_1\times\vec{d}_2/|\vec{d}_1\times\vec{d}_2|$，则

$$d=|\overrightarrow{AB}\cdot\vec{n}|=\frac{|\overrightarrow{AB}\cdot(\vec{d}_1\times\vec{d}_2)|}{|\vec{d}_1\times\vec{d}_2|}=\frac{|(\overrightarrow{AB},\vec{d}_1,\vec{d}_2)|}{|\vec{d}_1\times\vec{d}_2|}.$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

□

21.6.2 夹角

- **线线角**取为方向向量的夹角(锐角或题设规定的角).
- **线面角** θ 由方向 $\vec{v}$ 与法向量 $\vec{n}$ 给出, $\sin\theta=\dfrac{|\vec{v}\cdot\vec{n}|}{|\vec{v}||\vec{n}|}(\theta\in[0,\frac{\pi}{2}])$.
- **二面角**由两平面法向量 $\vec{n}_1,\vec{n}_2$ 的夹角给出, 它等于二面角或其补角; 需结合图形判断取锐角还是钝角, 或由平面角定义直接取平面角.

注 21.3
若两直线平行, 则距离退化为点到直线的距离. 取 l_2 上一点 B, $d=\dfrac{|\overrightarrow{AB}\times\vec{d}|}{|\vec{d}|}$.

21.6.3 体积与叉积、混合积

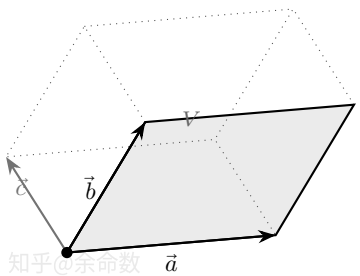


图 21.18　以 $\vec{a},\vec{b},\vec{c}$ 为棱的平行六面体; 体积为混合积绝对值

定义 21.11 (向量积 (叉积)) 在右手系中, $\vec{a}\times\vec{b}$ 垂直于 $\vec{a},\vec{b}$, 模长 $|\vec{a}||\vec{b}|\sin\theta$ 等于以二者为邻边的平行四边形面积, 方向由右手定则确定. 坐标公式为

$$\vec{a}\times\vec{b}=\begin{vmatrix}\vec{i}&\vec{j}&\vec{k}\\a_1&a_2&a_3\\b_1&b_2&b_3\end{vmatrix}=(a_2b_3-a_3b_2,\ a_3b_1-a_1b_3,\ a_1b_2-a_2b_1).$$

向量积给出平行四边形面积, 混合积把它推广到平行六面体体积.

定理 21.14 (平行六面体体积) 以 $\vec{a},\vec{b},\vec{c}$ 为棱的平行六面体体积

$$V=|(\vec{a},\vec{b},\vec{c})|=|\vec{a}\cdot(\vec{b}\times\vec{c})|=\left|\det\begin{pmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{pmatrix}\right|.$$

三向量共面 $\Longleftrightarrow$ 混合积为零.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 以 $\vec{b},\vec{c}$ 为邻边的底面平行四边形面积为 $|\vec{b}\times\vec{c}|$. 高为 $\vec{a}$ 在法向 $\vec{b}\times\vec{c}/|\vec{b}\times\vec{c}|$ 上投影的绝对值, 即 $|\vec{a}\cdot(\vec{b}\times\vec{c})|/|\vec{b}\times\vec{c}|$. 故

$$V=|\vec{b}\times\vec{c}|\cdot\frac{|\vec{a}\cdot(\vec{b}\times\vec{c})|}{|\vec{b}\times\vec{c}|}=|\vec{a}\cdot(\vec{b}\times\vec{c})|=|(\vec{a},\vec{b},\vec{c})|.$$

行列式形式由混合积的坐标公式直接得到. 三向量共面时高为零, 故混合积为零; 反之混合积为零则高为零, 三向量共面. □

将平行六面体体积公式应用于四面体, 得到三棱锥体积公式.

推论 21.1 (三棱锥体积) 以 O 为顶点、A,B,C 为底面顶点的四面体体积

$$V=\frac{1}{6}|(\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC})|.$$

证明 以 $\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC}$ 为棱的平行六面体可分割为六个等体积的三棱锥, 四面体 $OABC$ 是其中之一. 由平行六面体体积公式,

$$V=\frac{1}{6}|(\overrightarrow{OA},\overrightarrow{OB},\overrightarrow{OC})|.$$

□

例题 21.27 正方体棱长为 1, 求以一顶点为公共点、沿三条两两异面的面对角线方向所张成的平行六面体体积.

解 建系后三棱向量可取 $\vec{a}=(1,1,0)$、$\vec{b}=(1,0,1)$、$\vec{c}=(0,1,1)$. 则

$$V=\left|\det\begin{pmatrix}1&1&0\\1&0&1\\0&1&1\end{pmatrix}\right|=|1\cdot(0-1)-1\cdot(1-0)+0|=2.$$

例题 21.28 (点到平面距离) 正方体 $ABCD\text{-}A_1B_1C_1D_1$ 棱长为 2. 求点 A_1 到平面 AB_1D_1 的距离.

解 建系. 取 $A=(0,0,0)$、$B=(2,0,0)$、$D=(0,2,0)$、$A_1=(0,0,2)$、$B_1=(2,0,2)$、$D_1=(0,2,2)$. 平面 AB_1D_1 内

$$\overrightarrow{AB_1}=(2,0,2),\qquad \overrightarrow{AD_1}=(0,2,2).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

法向量

$$\vec{n}=\overrightarrow{AB_1}\times\overrightarrow{AD_1}=(0\cdot 2-2\cdot 2,\ 2\cdot 0-2\cdot 2,\ 2\cdot 2-0\cdot 0)=(-4,-4,4),$$

取 $\vec{n}=(1,1,-1)$. 平面过 A, 故点法式 $x+y-z=0$. 点 $A_1(0,0,2)$ 到平面距离

$$d=\frac{|0+0-2|}{\sqrt{1+1+1}}=\frac{2}{\sqrt{3}}=\frac{2\sqrt{3}}{3}.$$

例题 21.29 (球的截面) 球心为 $O(0,0,0)$、半径 $R=3$ 的球面, 被平面 $z=1$ 所截. 求截面圆的圆心、半径, 以及截面圆的面积.

解 球心到平面 $z=1$ 的距离 $d=1<3$, 故相交成圆. 圆心为球心在平面上的射影 $O'=(0,0,1)$, 半径

$$r=\sqrt{R^2-d^2}=\sqrt{9-1}=2\sqrt{2}.$$

截面圆面积 $\pi r^2=8\pi$.

球的定义、截面性质、切接问题与球面几何见 第21.8节.

21.6.4 高考真题选讲

以下高考真题考查空间几何体的体积、距离与夹角计算.

例题 21.30 (2020 年全国 □ 卷理科数学第 3 题) 埃及胡夫金字塔的形状可视为一个正四棱锥. 以该四棱锥的高为边长的正方形面积等于该四棱锥一个侧面三角形的面积, 则其侧面三角形底边上的高与底面正方形的边长的比值为

A. $\frac{\sqrt{5}-1}{4}$　　C. $\frac{\sqrt{5}+1}{4}$

B. $\frac{\sqrt{5}-1}{2}$　　D. $\frac{\sqrt{5}+1}{2}$

解 答案: C. 设正四棱锥的高为 h, 底面边长为 a, 侧面三角形底边上的高为 h'. 依题意 $h^2=\frac{1}{2}ah'$. 由勾股定理 $h^2=h'^2-(\frac{a}{2})^2$. 联立得 $h'^2-(\frac{a}{2})^2=\frac{1}{2}ah'$, 令 $t=\frac{h'}{a}$, 则 $4t^2-2t-1=0, t=\frac{1+\sqrt{5}}{4}$(负值舍去).

例题 21.31 (2019 年天津卷理科数学第 11 题) 已知四棱锥的底面是边长为 $\sqrt{2}$ 的正方形, 侧棱长均为 $\sqrt{5}$. 若圆柱的一个底面的圆周经过四棱锥四条侧棱的中点, 另一个底面的圆心为四棱锥底面的中心, 则该圆柱的体积为 ____.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

解 答案:$\frac{\pi}{4}$. 底面对角线长为 2,半对角线为 1. 由侧棱长 $\sqrt{5}$ 得棱锥高 $h=\sqrt{5-1}=2$. 侧棱中点位于半高处, 故圆柱高为 1. 侧棱中点构成正方形, 其对角线为底面对角线的一半, 即 1, 故圆柱底面半径为 $\frac{1}{2}$.$V=\pi(\frac{1}{2})^2\cdot 1=\frac{\pi}{4}$.

例题 21.32 (2021 年全国甲卷理科数学第 11 题) 已知 A、B、C 是半径为 1 的球 O 的球面上的三个点, 且 $AC\perp BC$, $AC=BC=1$, 则三棱锥 O-ABC 的体积为

A. $\frac{\sqrt{2}}{12}$　　C. $\frac{\sqrt{2}}{4}$

B. $\frac{\sqrt{3}}{12}$　　D. $\frac{\sqrt{3}}{4}$

解 答案: A. $\triangle ABC$ 为等腰直角三角形, $AB=\sqrt{2}$, 其外接圆半径 $r=\frac{\sqrt{2}}{2}$. 球心到平面 ABC 的距离 $d=\sqrt{1-r^2}=\frac{\sqrt{2}}{2}$.$S_{\triangle ABC}=\frac{1}{2}$, $V=\frac{1}{3}\cdot\frac{1}{2}\cdot\frac{\sqrt{2}}{2}=\frac{\sqrt{2}}{12}$.

例题 21.33 (2018 年港澳台华侨联考卷第 11 题) 若四面体棱长都相等, 则相邻两侧面所成的二面角的余弦值为

A. $\frac{1}{4}$　　C. $\frac{1}{2}$

B. $\frac{1}{3}$　　D. $\frac{2}{3}$

解 答案: B. 设正四面体棱长为 2, 取 CD 中点 E, 连接 AE, BE.$AE=BE=\sqrt{3}$、$AE\perp CD$、$BE\perp CD$、$\angle AEB$ 为二面角的平面角. 在 $\triangle ABE$ 中,

知乎@余命数

$$\cos\angle AEB=\frac{AE^2+BE^2-AB^2}{2\cdot AE\cdot BE}=\frac{3+3-4}{2\cdot 3}=\frac{1}{3}.$$

知乎@余命数
知乎@余命数

例题 21.34 (2005 年湖南卷文科数学第 4 题) 正方体 $ABCD$-$A_1B_1C_1D_1$ 的棱长为 1, O 是底面 $A_1B_1C_1D_1$ 的中心, 则 O 到平面 ABC_1D_1 的距离为

A. $\frac{1}{2}$　　C. $\frac{\sqrt{2}}{2}$

B. $\frac{\sqrt{2}}{4}$　　D. $\frac{\sqrt{3}}{2}$

解 答案: B. 以 A 为原点, AB、AD、AA_1 分别为 x、y、z 轴建系.$O(\frac{1}{2},\frac{1}{2},1)$. 平面 ABC_1D_1 含 $A(0,0,0)$、$B(1,0,0)$、$D_1(0,1,1)$, 法向量 $\vec{n}=\overrightarrow{AB}\times\overrightarrow{AD_1}=(0,-1,1)$, 方程为 $z-y=0$. 点 O 到该平面的距离

$$d=\frac{|1-\frac{1}{2}|}{\sqrt{1+1}}=\frac{\sqrt{2}}{4}.$$

知乎@余命数

例题 21.35 (2011 年全国大纲卷理科数学第 6 题) 已知直二面角 α-l-β，点 $A\in\alpha$、$AC\perp l$、C 为垂足，$B\in\beta$、$BD\perp l$、D 为垂足，若 $AB=2$，$AC=BD=1$，则 D 到平面 ABC 的距离等于

A. $\frac{\sqrt{2}}{3}$　　C. $\frac{\sqrt{6}}{3}$

B. $\frac{\sqrt{3}}{3}$　　D. 1

解 答案：C. 在直二面角中 $AC\perp l$，故 $AC\perp$ 平面 β，$AC\perp CD$ 且 $AC\perp BD$. 以 C 为原点，l 为 x 轴，AC 为 y 轴，BD 为 z 轴建系. 设 $CD=d$，则 $A(0,1,0)$、$D(d,0,0)$、$B(d,0,1)$.$AB=\sqrt{d^2+2}=2$，故 $d=\sqrt{2}$.

$V_{D\text{-}ABC}=V_{B\text{-}ACD}=\frac{1}{3}\cdot\frac{1}{2}\cdot AC\cdot CD\cdot BD=\frac{\sqrt{2}}{6}$.

$\overrightarrow{AB}=(\sqrt{2},-1,1)$、$\overrightarrow{AC}=(0,-1,0)$、$\overrightarrow{AB}\times\overrightarrow{AC}=(1,0,-\sqrt{2})$、$S_{\triangle ABC}=\frac{1}{2}\sqrt{3}$.

$$h=\frac{3V}{S}=\frac{\sqrt{2}/2}{\sqrt{3}/2}=\frac{\sqrt{6}}{3}.$$

21.7 三面角与四面体

第21.4节 引入了二面角，第21.6节 给出空间度量的基本公式. 本节研究从一点引出三条射线所构成的图形，即三面角，建立面角与二面角之间的余弦关系与正弦关系，并应用于球面距离的计算.

21.7.1 三面角的定义

定义 21.12 (三面角) 从空间一点 O 引出三条不共面的射线 OA、OB、OC，这三条射线连同它们之间的平面区域构成的图形称为**三面角**，记作 O-ABC. 点 O 称为三面角的**顶点**，三条射线称为**棱**. 两棱之间的夹角 $\angle BOC$、$\angle COA$、$\angle AOB$ 称为**面角**，分别记为 α、β、γ. 两面之间的二面角称为**二面角**，棱 OA 处的二面角记为 A，棱 OB 处的记为 B，棱 OC 处的记为 C.

三面角的三个面角满足类似于三角形三边的大小关系，任意两个面角之和大于第三个面角，任意两个面角之差小于第三个面角.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

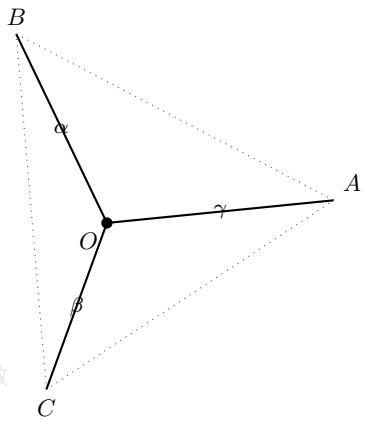


图 21.19 三面角 O-ABC, 三条棱 OA、OB、OC 与三个面角 α、β、γ

21.7.2 三面角的余弦定理

定理 21.15 (三面角的余弦定理) 设三面角的三个面角为 α、β、γ, 与面角 γ 相对的二面角为 θ, 即棱 OC 处的二面角 C, 则

$$\cos\gamma = \cos\alpha\cos\beta + \sin\alpha\sin\beta\cos\theta \tag{21-1}$$

或等价地

$$\cos\theta = \frac{\cos\gamma - \cos\alpha\cos\beta}{\sin\alpha\sin\beta} \tag{21-2}$$

证明 先设面角 α、β 均为锐角. 在棱 OC 上取异于 O 的点 H, 过 H 分别在面 OAC 与面 OBC 内作棱 OC 的垂线, 分别交射线 OA 于 A、交射线 OB 于 B. 面角非锐角时, 交点落在射线的反向延长线上, 以有向长度代入同一计算即可. 记 $\angle AOH = \beta$, 即面角 $\angle COA$, $\angle BOH = \alpha$, 即面角 $\angle BOC$, 二面角 $\angle AHB = \theta$, 面角 $\angle AOB = \gamma$.

设 $OH = 1$, 则 $AO = 1/\cos\beta$、$BO = 1/\cos\alpha$、$AH = \tan\beta$、$BH = \tan\alpha$. 在 $\triangle AHB$ 中由余弦定理:

$$AB^2 = AH^2 + BH^2 - 2\cdot AH\cdot BH\cdot\cos\theta = \tan^2\beta + \tan^2\alpha - 2\tan\alpha\tan\beta\cos\theta$$

在 $\triangle AOB$ 中由余弦定理:

$$\cos\gamma = \frac{AO^2 + BO^2 - AB^2}{2\cdot AO\cdot BO}$$

代入 AO、BO、AB^2 并化简:

$$\cos\gamma = \frac{\sec^2\beta + \sec^2\alpha - \tan^2\beta - \tan^2\alpha + 2\tan\alpha\tan\beta\cos\theta}{2\sec\alpha\sec\beta}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

利用 $\sec^2 x-\tan^2 x=1$, 分子化为 $2+2\tan\alpha\tan\beta\cos\theta$, 故

$$\cos\gamma=\frac{2+2\tan\alpha\tan\beta\cos\theta}{2\sec\alpha\sec\beta}=\cos\alpha\cos\beta+\sin\alpha\sin\beta\cos\theta$$

□

证明 (向量证法) 取 OA、OB、OC 方向上的单位向量 $\mathbf{a}$、$\mathbf{b}$、$\mathbf{c}$, 则 $\mathbf{a}\cdot\mathbf{b}=\cos\gamma$、$\mathbf{b}\cdot\mathbf{c}=\cos\alpha$、$\mathbf{c}\cdot\mathbf{a}=\cos\beta$. 由二重外积公式 $\mathbf{c}\times(\mathbf{c}\times\mathbf{a})=(\mathbf{c}\cdot\mathbf{a})\mathbf{c}-(\mathbf{c}\cdot\mathbf{c})\mathbf{a}=\cos\beta\,\mathbf{c}-\mathbf{a}$, 故

$$(\mathbf{b}\times\mathbf{c})\cdot(\mathbf{c}\times\mathbf{a})=\mathbf{b}\cdot[\mathbf{c}\times(\mathbf{c}\times\mathbf{a})]=\mathbf{b}\cdot(\cos\beta\,\mathbf{c}-\mathbf{a})=\cos\alpha\cos\beta-\cos\gamma$$

另一方面,$|\mathbf{b}\times\mathbf{c}|=\sin\alpha$,$|\mathbf{c}\times\mathbf{a}|=\sin\beta$,设两者夹角为 θ,即二面角,则 $(\mathbf{b}\times\mathbf{c})\cdot(\mathbf{c}\times\mathbf{a})=\sin\alpha\sin\beta\cos\theta$. 比较两式即得 (21-1). □

余弦定理把面角与二面角联系起来, 已知三个面角可求二面角, 已知两个面角与夹二面角可求第三面角. 向量证法用二重外积公式绕开了几何构造, 直接从向量代数得到结果, 体现了外积在立体几何中的计算效率.

定理 21.16 (二面角为直角的三面角) 设三面角 O-ABC 中与面角 $\gamma=\angle AOB$ 相对的二面角(棱 OC 处)为直角, 则面角满足

$$\cos\gamma=\cos\alpha\cos\beta \tag{21-3}$$

若三个二面角均为直角, 则三棱两两垂直, 三个面角同为直角, 上式成为恒等式 $0=0$.

证明 在 (21-1) 中取 $\theta=\pi/2$, $\cos\theta=0$ 即得. □

例题 21.36 (面角的取值范围) 三面角的三个面角 α、β、γ 满足

$$\left|\frac{\cos\gamma-\cos\alpha\cos\beta}{\sin\alpha\sin\beta}\right|\le 1$$

解 由 (21-2),$|\cos\theta|\le 1$ 给出此不等式. 从面角的和差关系看,$|\alpha-\beta|\le\gamma\le\min\{\alpha+\beta,\,2\pi-\alpha-\beta\}$,故 $\cos(\alpha-\beta)\ge\cos\gamma\ge\cos(\alpha+\beta)$,即 $|\cos\gamma-\cos\alpha\cos\beta|\le\sin\alpha\sin\beta$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

21.7.3 三面角的正弦定理

定理 21.17 (三面角的正弦定理) 设三面角的三个面角为 α、β、γ, 三个二面角为 A、B、C, 则

$$\frac{\sin\alpha}{\sin A}=\frac{\sin\beta}{\sin B}=\frac{\sin\gamma}{\sin C}=\frac{\sin\alpha\sin\beta\sin\gamma}{V} \tag{21-4}$$

其中

$$V=\sqrt{1-\cos^2\alpha-\cos^2\beta-\cos^2\gamma+2\cos\alpha\cos\beta\cos\gamma}$$

证明 由 (21-2), $\sin^2\theta=1-\frac{(\cos\gamma-\cos\alpha\cos\beta)^2}{\sin^2\alpha\sin^2\beta}$. 展开分子:

$$\begin{aligned}&\sin^2\alpha\sin^2\beta-(\cos\gamma-\cos\alpha\cos\beta)^2\\&=(1-\cos^2\alpha)(1-\cos^2\beta)-\cos^2\gamma+2\cos\alpha\cos\beta\cos\gamma-\cos^2\alpha\cos^2\beta\\&=1-\cos^2\alpha-\cos^2\beta-\cos^2\gamma+2\cos\alpha\cos\beta\cos\gamma=V^2\end{aligned}$$

故 $\sin^2\theta/\sin^2\gamma=V^2/(\sin^2\alpha\sin^2\beta\sin^2\gamma)$. 此式关于 α、β、γ 轮换对称, 故 $\sin A/\sin\alpha=\sin B/\sin\beta=\sin C/\sin\gamma=V/(\sin\alpha\sin\beta\sin\gamma)$, 取倒数即得 (21-4). □

21.7.4 球面距离的计算

球面上两点的最短路径是过两点的大圆弧, 其弧长由两点对球心的张角决定. 把球心 O、两地面点 A、B 与北极 N 视为三面角 O-ABN, 面角 $\angle AOB$ 就是所求的张角, 三面角余弦定理直接给出答案.

例题 21.37 (经纬度求球面距离) 设地球半径为 R, 球面上两点 A、B 的经纬度分别为 (u_1,v_1)、(u_2,v_2)(u 为经度, v 为纬度). 求 A、B 间的球面距离.

解 取地心 O 与北极 N. 射线 OA、OB、ON 构成三面角, 面角 $\angle AOB$ 即 A、B 对地心的张角, 面角 $\angle AON=\pi/2-v_1$, $\angle BON=\pi/2-v_2$ (纬度角的余角), 二面角 A-ON-B 即经度差 u_1-u_2. 由三面角余弦定理 (21-1):

$$\cos\angle AOB=\sin v_1\sin v_2+\cos v_1\cos v_2\cos(u_1-u_2)$$

球面距离为

$$d_{AB}=R\arccos(\sin v_1\sin v_2+\cos v_1\cos v_2\cos(u_1-u_2))$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

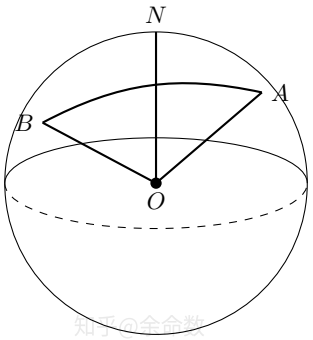


图 21.20　球面距离, 地心 O、两地点 A、B 与北极 N 构成三面角

例题 21.38 (面角与二面角的关系) 三面角 O-ABC 中与面角 γ 相对的二面角为直角, 记 $\angle BOC=\alpha$、$\angle COA=\beta$、$\angle AOB=\gamma$. 已知 $\alpha=\beta=\pi/3$, 求 γ.

解 由 (21-3), $\cos\gamma=\cos\alpha\cos\beta=\cos(\pi/3)\cos(\pi/3)=\frac{1}{4}$, 故 $\gamma=\arccos(1/4)$.

21.8 球

第21.6节 给出了球的定义与三条常用事实. 本节把球作为独立对象, 展开它的截面、切接、局部构造与球面几何.

21.8.1 截面与球面距离

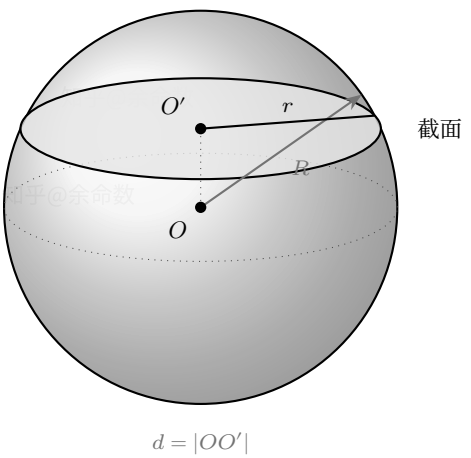


图 21.21　平面截球, 圆心为球心在平面上的射影 O', 半径 $r=\sqrt{R^2-d^2}$

平面截球所得的截面恒为圆. 截面经过球心时, 半径等于球半径 R, 称为**大圆**; 不经过球心时, 设球心到截面的距离为 d, 截面半径

$$r=\sqrt{R^2-d^2}\qquad(0\le d\le R).\tag{21-5}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

地球上的赤道与所有经线圈是大圆, 纬线圈(除赤道外)是小圆.

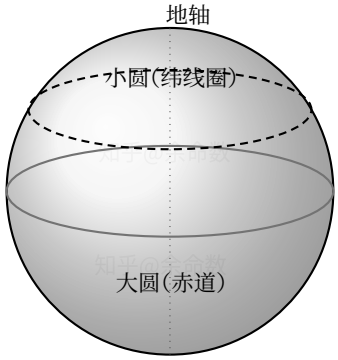


图 21.22 大圆与小圆, 赤道是过球心的大圆, 纬线圈是不过球心的小圆

定义 21.13 (球面距离) 球面上两点 A, B 之间的**球面距离**, 定义为经过这两点的大圆上不超过半圆的弧长; A、B 为对径点时过这两点的大圆不唯一, 但每段半圆弧长都是 πR, 球面距离仍为 πR.

设球半径为 R, 大圆上 A, B 所张圆心角为 θ(弧度, $0 \le \theta \le \pi$), 则球面距离

$$\ell = R\,\theta. \tag{21-6}$$

两点间的空间直线距离为弦长 $|AB| = 2R\sin\frac{\theta}{2}$, 它穿过球体内部, 与球面距离是两个不同的量.

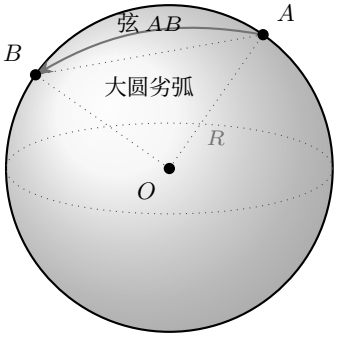


图 21.23 球面距离(大圆劣弧)与弦长的区别, 弦长即空间直线距离

例题 21.39 (球面距离) 设地球为半径 R 的球, 北纬 45° 圈上有两点 A, B, 经度差为 90°. 求 A, B 的球面距离.

解 北纬 45° 圈是小圆, 半径 $r = R\cos 45^\circ = \frac{\sqrt{2}}{2}R$. 在纬线圈上, 经度差 90° 对应圆心角 90°, 故纬线圈上的弧长为 $\frac{\sqrt{2}}{2}R \cdot \frac{\pi}{2} = \frac{\sqrt{2}\pi R}{4}$. 但这不是球面距离.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

求 A, B 所张的球心角 θ. 设球心 O, 地轴为 z 轴.A, B 在北纬 45°, 即 $z = \frac{\sqrt{2}}{2}R$, 投影到赤道平面的半径为 $\frac{\sqrt{2}}{2}R$. 取 A 在经度 0°, B 在经度 90°:

$$\overrightarrow{OA} = \Big(\frac{\sqrt{2}}{2}R,\, 0,\, \frac{\sqrt{2}}{2}R\Big), \quad \overrightarrow{OB} = \Big(0,\, \frac{\sqrt{2}}{2}R,\, \frac{\sqrt{2}}{2}R\Big).$$

$$\cos\theta = \frac{\overrightarrow{OA}\cdot\overrightarrow{OB}}{R^2} = \frac{0+0+\frac{1}{2}R^2}{R^2} = \frac{1}{2}, \quad \theta = \frac{\pi}{3}.$$

球面距离 $\ell = R\cdot\frac{\pi}{3} = \frac{\pi R}{3}$.

知乎@余命数

21.8.2 外接球与内切球

球的表面若经过多面体的所有顶点, 这个球就称为多面体的**外接球**. 球心到各顶点距离相等, 通常由各面外心所作垂线的交点定位.

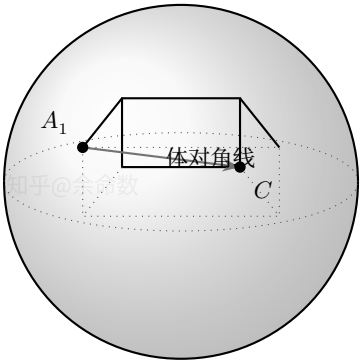


图 21.24　长方体的外接球, 体对角线是它的直径

定理 21.18 (长方体的外接球) 长方体的长、宽、高为 a, b, c, 其体对角线是外接球的直径, 故外接球半径

$$R = \frac{\sqrt{a^2+b^2+c^2}}{2}. \tag{21-7}$$

证明 以长方体的中心(体对角线交点)为原点, 三条棱所在直线为坐标轴建立空间直角坐标系. 此时八个顶点的坐标为 $(\pm\frac{a}{2}, \pm\frac{b}{2}, \pm\frac{c}{2})$ 的所有组合.

任意一个顶点到原点 O 的距离平方为

$$\Big(\frac{a}{2}\Big)^2 + \Big(\frac{b}{2}\Big)^2 + \Big(\frac{c}{2}\Big)^2 = \frac{a^2+b^2+c^2}{4},$$

与正负号选取无关, 故八个顶点到 O 的距离都等于 $\frac{1}{2}\sqrt{a^2+b^2+c^2}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于是 O 到所有顶点等距, 以 O 为球心、$\frac{1}{2}\sqrt{a^2+b^2+c^2}$ 为半径的球面经过全部顶点, 即长方体的外接球. 体对角线两端点的距离为 $\sqrt{a^2+b^2+c^2}$, 恰为该球直径, 故体对角线是外接球的直径. □

例题 21.40 (正四面体的外接球) 棱长为 a 的正四面体, 求其外接球半径.

解 正四面体 $ABCD$ 的外接球球心到四个顶点等距. 设球心为 O, 半径 R. 由对称性, O 到每个面的距离相同.

正四面体的高 $h=\sqrt{a^2-\left(\frac{\sqrt{3}}{3}a\right)^2}=\frac{\sqrt{6}}{3}a$. 球心在高的延长线上, 到底面三角形外心(重心)的距离记为 d. 由 $R^2=d^2+\left(\frac{\sqrt{3}}{3}a\right)^2$ 与 $R+d=h=\frac{\sqrt{6}}{3}a$, 解得

$$R=\frac{\sqrt{6}}{4}a.$$

球面若与多面体的每个面都相切, 这个球就称为多面体的**内切球**. 设多面体体积为 V, 表面积为 S, 以球心为顶点、各面为底面把多面体分割成若干棱锥, 每个棱锥的高都是内切球半径 r, 故

$$V=\frac{1}{3}rS \quad\Longrightarrow\quad r=\frac{3V}{S}. \tag{21-8}$$

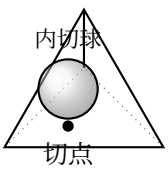


图 21.25 正四面体的内切球, 球心到每个面的距离都等于内切球半径

例题 21.41 (正四面体的内切球) 棱长为 a 的正四面体, 求其内切球半径.

解 正四面体体积 $V=\frac{\sqrt{2}}{12}a^3$, 表面积 $S=4\cdot\frac{\sqrt{3}}{4}a^2=\sqrt{3}\,a^2$. 由 (21-8),

$$r=\frac{3V}{S}=\frac{3\cdot\frac{\sqrt{2}}{12}a^3}{\sqrt{3}\,a^2}=\frac{\sqrt{6}}{12}a.$$

与外接球半径 $R=\frac{\sqrt{6}}{4}a$ 相比, $r:R=1:3$.

注 21.4 (外接球与内切球半径之比) 正四面体的外接球半径与内切球半径之比恒为 $3:1$. 这一比值由对称性决定, 与棱长无关.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

21.8.3 两球的位置关系

球心为 (x_0,y_0,z_0)、半径 R 的球面方程为

$$(x-x_0)^2+(y-y_0)^2+(z-z_0)^2=R^2. \tag{21-9}$$

设两球球心距为 d, 半径 R_1,R_2(不妨设 $R_1\geq R_2$), 位置关系由 d 与 R_1,R_2 的比较决定:

关系	几何条件	交集
相离	$d>R_1+R_2$	空
外切	$d=R_1+R_2$	一个点
相交	$\lvert R_1-R_2\rvert<d<R_1+R_2$	一个圆
内切	$d=R_1-R_2$	一个点
内含	$d<R_1-R_2$	空

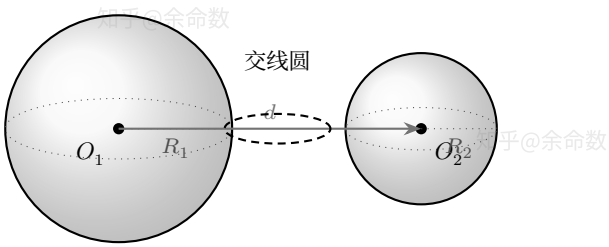


图 21.26　两球相交, 交线是垂直于球心连线的空间圆

两球相交时, 交线是一个空间中的圆. 该圆所在平面垂直于两球心连线, 圆心在两球心连线上.

例题 21.42 (两球相交的截面圆) 球 O_1 球心 $(0,0,0)$、半径 $R_1=3$, 球 O_2 球心 $(4,0,0)$、半径 $R_2=2$. 求两球相交圆的圆心与半径.

解 球心距 $d=4$, 满足 $|3-2|<4<3+2$, 故相交. 交线圆在平面 $x=x_0$ 上, 圆心 $C=(x_0,0,0)$ 在 O_1O_2 连线上. 由

$$x_0^2+r^2=R_1^2=9,\qquad (x_0-4)^2+r^2=R_2^2=4,$$

两式相减得 $8x_0-16=5$, 故 $x_0=\frac{21}{8}$. 代回得

$$r^2=9-\left(\frac{21}{8}\right)^2=\frac{576-441}{64}=\frac{135}{64},\quad r=\frac{3\sqrt{15}}{8}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

21.8.4 球冠与球缺

用平面截去球的一部分，截下的曲面与几何体各有其公式. 这一节给出球冠与球缺的完整定义、公式与直观推导.

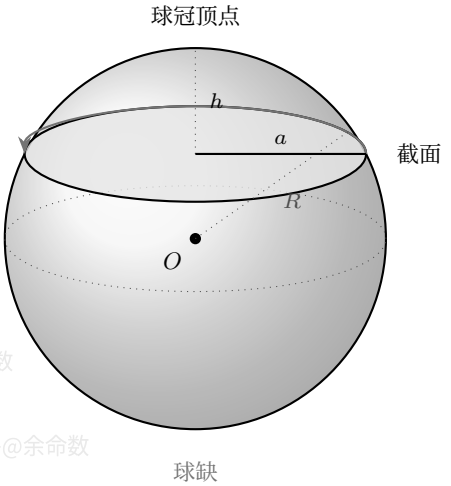


图 21.27 球冠与球缺，截面半径 a，高度 h，球半径 R

定义 21.14 (球冠) 球被平面所截，截面与球面之间的**曲面部分**称为**球冠**. 截面圆的圆心到球冠最高点的距离称为球冠的**高**，记为 h.

定义 21.15 (球缺) 球被平面所截，截面与球冠曲面共同围成的**几何体**称为**球缺**. 球缺的高与对应球冠的高相同，仍记为 h.

球冠是曲面（二维），球缺是几何体（三维）. 球缺 = 球冠曲面 + 截面圆盘 + 内部体积.

注 21.5 (截面半径与高度、球半径的关系)
设球半径为 R，球冠高为 h，截面圆半径为 a，球心到截面的有向距离 $d = R - h$，当 $h > R$ 时截面越过球心，d 取负值. 由勾股定理

$$a^2 = R^2 - d^2 = R^2 - (R-h)^2 = 2Rh - h^2 = h(2R-h). \tag{21-10}$$

这一关系在球冠面积与球缺体积的推导中反复使用.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

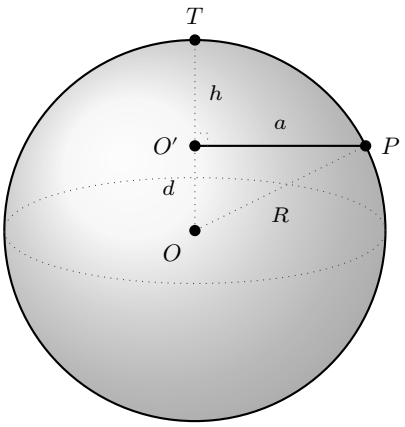


图 21.28　截面半径公式 $a^2 = h(2R - h)$ 的几何关系

定理 21.19 (球冠的面积) 半径为 R 的球, 高为 h 的球冠面积为

$$S_{\text{冠}} = 2\pi R\, h. \tag{21-11}$$

证明 (直观推导) 把球冠沿高度方向切成 n 个窄带. 在高度 z 处 $(R - h \le z \le R)$, 窄带的半径为 $a(z) = \sqrt{R^2 - z^2}$, 周长 $2\pi a(z)$. 窄带的斜高(沿球面方向的高度)近似为 Δs, 面积近似为 $2\pi a(z)\,\Delta s$.

当 $n \to \infty$ 时, 窄带面积之和趋于球冠面积. 由勾股定理, 窄带斜高 Δs 与竖直高度 Δz 的关系为 $\Delta s = \frac{R}{a(z)}\Delta z$ (球面在高度 z 处的倾斜程度). 代入得窄带面积 $2\pi a(z) \cdot \frac{R}{a(z)}\Delta z = 2\pi R\,\Delta z$. 求和:

$$S_{\text{冠}} \approx \sum 2\pi R\,\Delta z = 2\pi R \sum \Delta z = 2\pi R \cdot h.$$

取极限即得 $S_{\text{冠}} = 2\pi R\, h.$ □

注 21.6 (球冠面积公式的特点)
球冠面积只与球半径 R 和高度 h 有关, 与截面位置无关. 当 $h = R$ 时(半球), $S = 2\pi R^2$, 恰为球面积 $4\pi R^2$ 的一半. 当 $h = 2R$ 时, 即整个球面, $S = 4\pi R^2$, 与球面积公式一致.

定理 21.20 (球缺的体积) 半径为 R 的球, 高为 h 的球缺体积为

$$V_{\text{缺}} = \frac{\pi h^2}{3}(3R - h). \tag{21-12}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

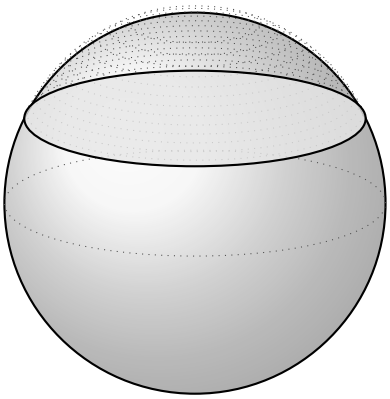


图 21.29 球冠面积的直观推导, 沿高度方向切成窄带, 求和取极限

证明 (直观推导) 把球缺沿高度方向切成 n 个薄圆片. 在高度 z 处 ($R-h \le z \le R$), 薄片的半径为 $a(z)=\sqrt{R^2-z^2}$, 面积为 $\pi a(z)^2=\pi(R^2-z^2)$, 厚度为 Δz, 体积近似为 $\pi(R^2-z^2)\,\Delta z$.

求和取极限, 球缺体积等于这些薄片体积之和:

$$V_{\text{缺}}=\lim_{n\to\infty}\sum_{i=1}^{n}\pi(R^2-z_i^2)\,\Delta z.$$

这个求和的精确计算需要定积分工具. 将变量替换 $t=z-(R-h)$, 使积分下限变为 0, 上限变为 h. 在新坐标下, 薄片半径为 $\sqrt{R^2-(R-h+t)^2}=\sqrt{2Rh-h^2-2(R-h)t-t^2}$, 化简后仍较复杂.

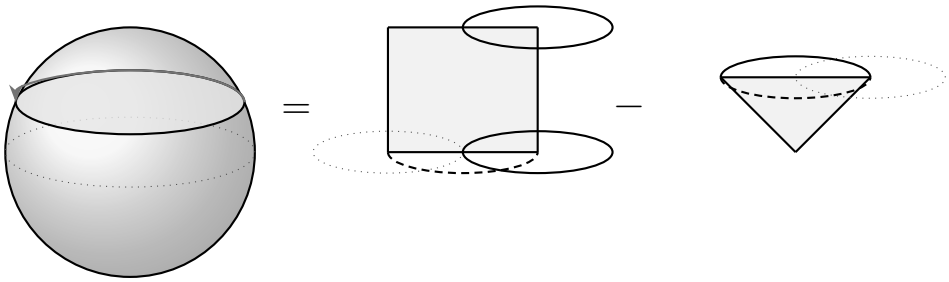
直接验证结果, 即把 $V=\frac{\pi h^2}{3}(3R-h)$ 展开,

$$V=\pi Rh^2-\frac{\pi h^3}{3}.$$

第一项是底半径为 h、高为 R 的圆柱体积, 第二项是底半径为 h、高为 h 的圆锥体积. 球缺体积 = 圆柱 − 圆锥, 这一关系在 图 21.30 中直观展示. □

注 21.7 (球缺体积的严格证明)
上述推导中, 薄圆片体积之和的精确计算需要定积分 $\int_{R-h}^{R}\pi(R^2-z^2)\,dz$. 定积分是 第17章 的主题, 此处不给出严格推导. 有了定积分工具后, 这个积分只需三行即可完成, 结果与 (21-12) 一致.

注 21.8 (球缺公式的极限验证)



球缺　　圆柱 $\pi h^2 R$　　圆锥 $\frac{\pi h^2 h}{3}$

图 21.30　球缺 = 圆柱 − 圆锥: 体积关系 $V_{缺} = \pi h^2 R - \frac{\pi h^3}{3}$ 的直观展示

当 $h = 2R$ 时, 球缺就是整个球, 体积

$$V = \frac{\pi(2R)^2}{3}(3R - 2R) = \frac{4\pi R^3}{3},$$

与球体积公式一致. 当 $h = R$ 时(半球),

$$V = \frac{\pi R^2}{3}(3R - R) = \frac{2\pi R^3}{3},$$

恰为球体积的一半.

例题 21.43 (球缺体积计算) 半径 $R = 5$ 的球, 被距球心 $d = 3$ 的平面所截, 求较小部分(球缺)的体积.

解 球心到截面的距离 $d = 3$, 故球缺的高 $h = R - d = 5 - 3 = 2$. 截面半径 $a = \sqrt{R^2 - d^2} = \sqrt{25 - 9} = 4$.

由 (21-12),

$$V = \frac{\pi h^2}{3}(3R - h) = \frac{\pi \cdot 4}{3}(15 - 2) = \frac{52\pi}{3}.$$

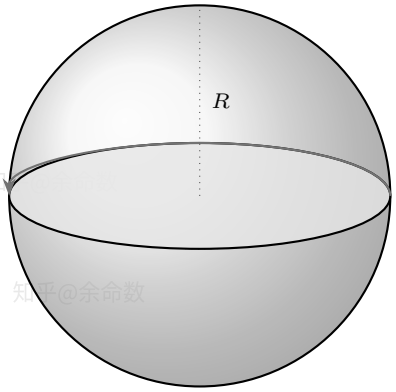
例题 21.44 (球冠面积计算) 半径 $R = 1$ 的球, 球冠的高 $h = \frac{1}{2}$, 求球冠面积与它占球面积的比例.

解 由 (21-11),

$$S = 2\pi R\, h = 2\pi \cdot 1 \cdot \frac{1}{2} = \pi.$$

球面积 $S_{球} = 4\pi R^2 = 4\pi$, 故球冠占球面积的比例为 $\frac{\pi}{4\pi} = \frac{1}{4}$.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵



半球 $h = R$

图 21.31 半球是球冠与球缺在 $h = R$ 时的特例

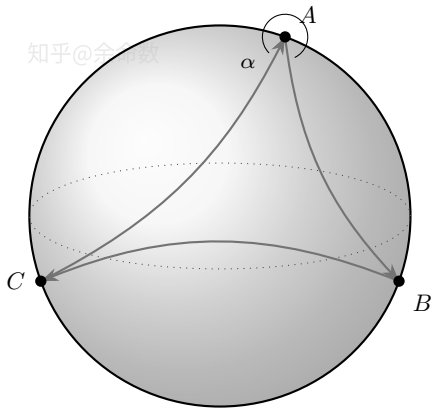


图 21.32 球面三角形, 三边为大圆弧, 内角和大于 180°

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

21.8.5 球面几何初步

在欧几里得平面上, 三角形内角和恒为 180°. 球面上的几何遵循不同的规则.

定义 21.16 (球面三角形) 球面上由三段大圆弧围成的区域称为**球面三角形**. 三段大圆弧称为它的边, 两边的夹角称为球面三角形的角.

球面三角形的角定义为两段大圆弧在顶点处的切线夹角. 由于球面弯曲, 球面三角形的内角和严格大于 180° 且小于 540°:

$$180^\circ < \alpha + \beta + \gamma < 540^\circ. \tag{21-13}$$

例题 21.45 (球面三角形的内角和) 在单位球面上, 取三段大圆弧各为 90°. 从北极沿经线到赤道的弧记为 AB, 沿赤道走 90° 的弧记为 BC, 再沿经线回到北极的弧记为 CA. 这构成一个球面三角形, 三个角都是 90°, 内角和 270°.

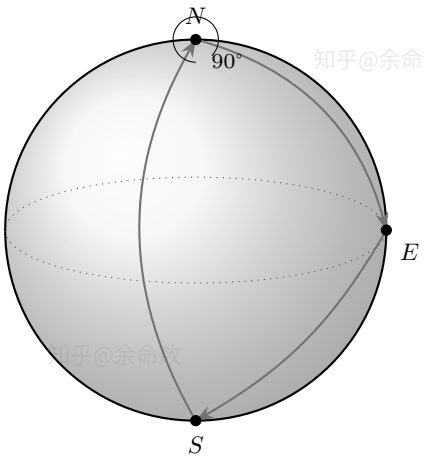


图 21.33　三直角球面三角形, 三个角都是 90°, 内角和 270°

注 21.9 (球面几何的应用)
球面几何是导航学、天文学以及广义相对论中弯曲时空的几何基础. 球面三角形的内角和超过 180° 的部分, 称为**球面过剩**, 与三角形面积成正比. 这一关系在 第17章 的积分应用中还会再次出现.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

知乎@余命数

知乎@余命数

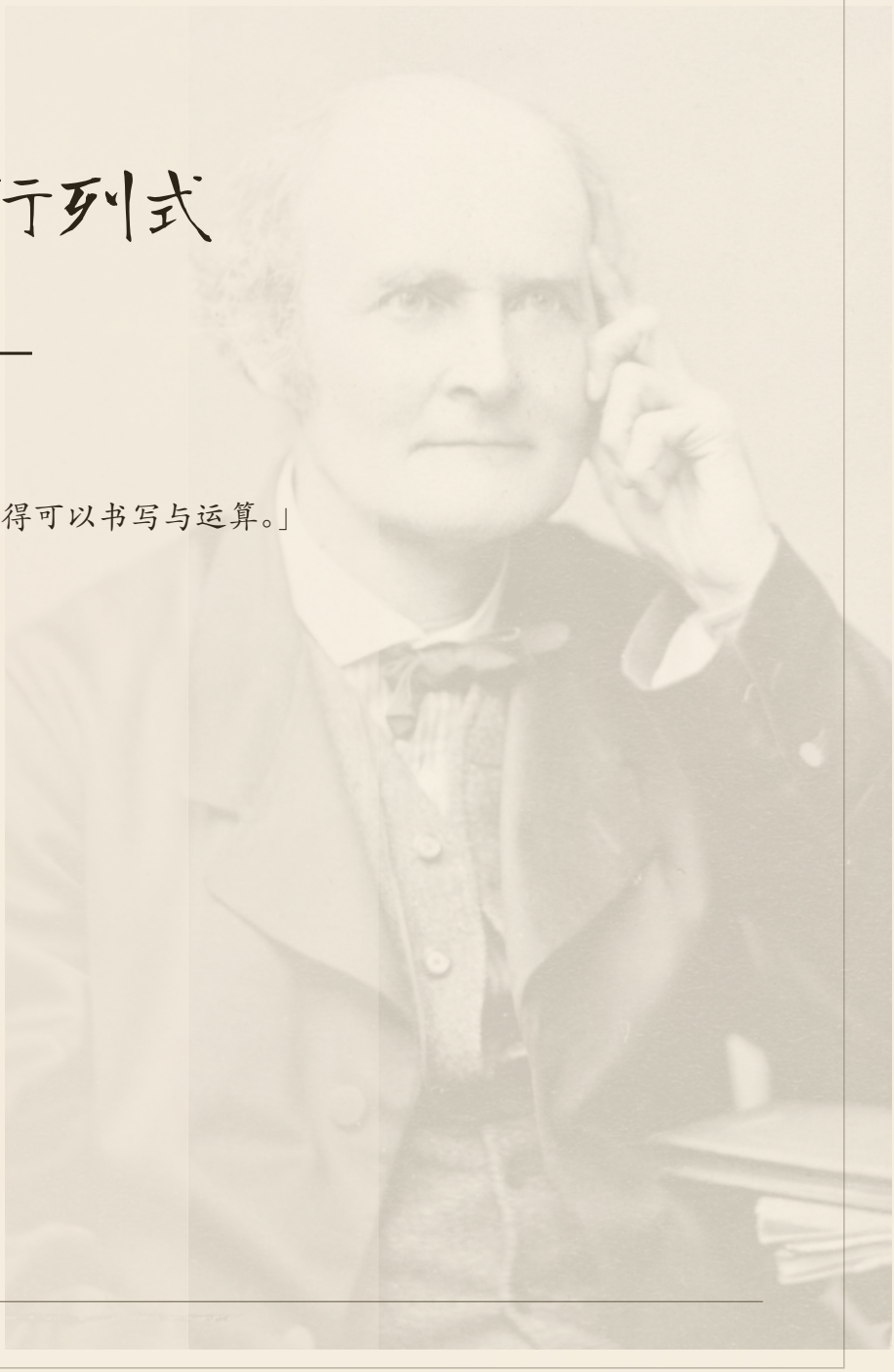
知乎@余命数

知乎@余命数

知乎@余命数

第 22 章

矩阵与行列式



扩展与线性结构

「矩阵使线性关系变得可以书写与运算。」

——凯莱

第二十二章 矩阵与行列式

知乎@余命数

在数学的众多领域中, 我们常常需要处理由数字构成的矩形阵列. 为了理解这种结构的起源与作用, 先回到一个熟悉的问题. 求解线性方程组.

考虑一个三元一次方程组.

$$2x+3y-z=5$$

$$4x-y+2z=1$$

$$x+5y+3z=10$$

当我们运用加减消元法求解时, 操作实质上是对方程的系数和常数项组合与变换. 变量 x,y,z 在整个过程中仅仅是占位符, 它们的位置是固定的. 真正承载系统全部信息的是那些数字.

$$\left(\begin{array}{ccc|c} 2 & 3 & -1 & 5 \\ 4 & -1 & 2 & 1 \\ 1 & 5 & 3 & 10 \end{array}\right)$$

这种洞察提示, 可以把这些数字信息从方程的形式中剥离出来, 作为一个独立的数学对象研究. 这个被剥离出来的、由数字构成的矩形阵列, 就是下面要研究的**矩阵**. 它将线性方程组的求解、几何空间的变换以及数据的统计分析等诸多问题, 统一在了一个共同的代数框架之下.

22.0.1 约束与自由度

在系统研究矩阵之前, 有必要从一个更基本的层面来理解线性方程组. 在没有任何限制的情况下, 三个变量 x,y,z 可以在三维空间中自由取值. 可独立地为 x,y,z 中的任意一个或多个变量赋值, 而不会影响其他变量的可能性. 我们称这种情况拥有三个**自由度**.

然而, 当我们引入第一个方程 $2x+3y-z=5$ 时, 情况发生了根本性的变化. 这个方程对变量 x,y,z 施加了一个**约束**. 它们不再能够独立地自由取值. 例如, 一旦我们确定了 x 和 y 的值, z 的值就必须是 $z=2x+3y-5$, 它被唯一地确定了. 此时, 我们只能自由地选择两个变量的值, 系统的自由度从三个降至两个.

从几何上看, 三个自由度对应着整个三维空间. 而一个形如 $ax+by+cz=d$ 的线性方程, 在三维空间中定义的是一个平面. 这个平面是一个二维的几何对象. 因此, 施加一个约束, 等价于将解的活动范围从三维空间限制到了一个二维平面上, 自由度减一.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当我们引入第二个方程 $4x-y+2z=1$ 时，我们施加了第二个约束. 现在，解 (x,y,z) 必须同时位于两个平面上. 两个不平行的平面相交，其交集是一条直线. 直线是一个一维的几何对象. 即此时只剩下一个自由度. 只要我们确定了直线上一点的某个坐标(例如 x)，其余的坐标 y,z 就会随之确定.

知乎@余命数

最后，引入第三个方程 $x+5y+3z=10$，这是第三个约束. 这个解现在必须同时位于三个平面上. 在通常情况下，一条直线与一个不平行于它的平面相交，其交集是一个唯一的点. 点是零维的几何对象. 此时，我们没有任何自由度了，变量 x,y,z 的值被这三个约束完全锁定，这就是方程组有**唯一解**的情形.

这个过程可以写成一个关系式.

$$\text{解的自由度} = \text{变量的个数} - \text{有效约束的个数}$$

然而，有些约束会提供冗余信息. 如果第三个方程是前两个方程的线性组合，例如，第一个方程乘以 3，那么它并没有提供任何新的信息，也就不是一个有效的独立约束. 在几何上，即第三个平面恰好也通过前两个平面的交线，因此交集仍然是一条直线，解有无穷多个. 如果第三个方程与前两个方程矛盾，例如 $2x+3y-z=100$，那么三个平面将没有公共交点，方程组无解.

因此，求解线性方程组的首要任务，转化为一个更为透彻的问题.

$$\textit{判断一个方程组中的 } n \textit{ 个约束是否相互独立.}$$

要系统地解决这个问题，需一种能够描述并量化这些系数之间关系的工具. 这正是矩阵和行列式即将派上用场的地方.

注 22.1

本章顺序. 向量 → 矩阵定义与运算 → 二/三阶行列式 → 逆与克拉默 → 消元求秩与解方程组；线性变换与齐次坐标、二次曲线的矩阵分类见后文，并与 第13章 对照. 第11章 中的平面向量与本节有序数组语言对照阅读.

知乎@余命数

22.1 向量

已在 第11章 与 第21.5节 中把向量当作几何对象使用过. 本节换一个视角，把向量写成有序数组，作为矩阵的行/列与线性方程组的未知量载体.

知乎@余命数

知乎@余命数

22.1.1 定义与表示

定义 22.1 (向量) 一个**向量**是一个有序的数字列表. 它可以表示为一个**行向量**或一个**列向量**.

- **列向量**: $\mathbf{v}=\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$
- **行向量**: $\mathbf{v}=\begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$

其中 $v_1, v_2, \ldots, v_n$ 称为向量的分量. 向量的维数等于其分量的数量.

知乎@余命数

22.1.2 几何解释与基本运算

22.1.2.0.1 几何解释 在坐标系中, 一个向量理解为一个从原点出发指向某一点的有向线段, 其分量对应着点的坐标.

- 二维向量 (x, y) 对应平面上的一个点.
- 三维向量 (x, y, z) 对应三维空间中的一个点.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

向量的**大小**或**模长**通常用双竖线表示, 例如 $\|\mathbf{v}\|$. 在二维和三维空间中, 它可以用勾股定理计算, 得

$$\left\| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\| = \sqrt{v_1^2 + v_2^2}, \quad \left\| \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

22.1.2.0.2 向量加法 向量加法的几何意义是将两个向量首尾相接, 其和向量是从第一个向量起点到第二个向量终点的向量, 即三角形法则; 或将两个向量平移到同一起点, 其和向量是以这两个向量为邻边构成的平行四边形的对角线, 即平行四边形法则. 在代数上, 向量加法是对应分量相加. 设 $\mathbf{u}=\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \mathbf{v}=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, 则

$$\mathbf{u}+\mathbf{v}=\begin{pmatrix} u_1+v_1 \\ u_2+v_2 \end{pmatrix}.$$

知乎@余命数

22.1.2.0.3　标量乘法　向量的标量乘法是改变向量的长度和/或方向. 标量 k 乘以向量 $\mathbf{v}$, 就是将 $\mathbf{v}$ 的每个分量都乘以 k. 设 $\mathbf{v}=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, k 为标量, 则 $k\mathbf{v}=\begin{pmatrix} kv_1 \\ kv_2 \end{pmatrix}$. 如果 $k>0$, 向量方向不变, 长度变为原来的 k 倍; 如果 $k<0$, 向量方向反向, 长度变为原来的 $|k|$ 倍.

22.1.2.0.4　数量积(点积)　对同维列向量 $\mathbf{u}, \mathbf{v}$, 定义

$$\mathbf{u}\cdot\mathbf{v}=u_1v_1+\cdots+u_nv_n.$$

二维、三维时它与夹角公式 $\mathbf{u}\cdot\mathbf{v}=|\mathbf{u}||\mathbf{v}|\cos\theta$ 一致, 并给出正交判定 $\mathbf{u}\cdot\mathbf{v}=0$. 更高维则把正交直接定义为点积为零.

立体几何中的空间向量侧重位置关系与度量 (第21章); 本章把同一对象用于矩阵乘法与行列式. 二者对象相同, 处理角度不同.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

22.2 矩阵的定义与运算

22.2.1 定义与记法

定义 22.2 (矩阵) 一个 $m \times n$ 的**矩阵**是一个由 $m \cdot n$ 个数排成的长方形数表, 它有 m 条**行**与 n 条**列**, 行自上而下, 列自左向右. 本书中, 矩阵里的数一般取实数, 必要时会说明为复数或其他数.
我们通常用大写字母表示矩阵, 例如 A. 为了精确地指明矩阵中的每一个数, 我们给这些数编上行、列两个下标. 记

$$a_{ij} \text{ 为矩阵} A \text{ 中第} i \text{ 行、第} j \text{ 列的元素} \quad (1 \le i \le m,\ 1 \le j \le n).$$

于是一个 $m \times n$ 的矩阵 A 可以写成

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

为了强调由哪些元素组成和大小是多少, 我们也常将其简记为

$$A = (a_{ij})_{m\times n} \quad \text{或} \quad A = (a_{ij})_{1\le i\le m,\ 1\le j\le n}.$$

在矩阵的大小 $m \times n$ 已经由上下文明确时, 也常简写为 $A = (a_{ij})$.

- 2×3 **矩阵**.

$$A = \begin{pmatrix} 1 & -2 & 0 \\ 4 & 5 & 9 \end{pmatrix}.$$

它有 2 行、3 列, 所以是一个 2×3 矩阵. 按上面的记号, 矩阵元素 a_{21} 等于 4, 它位于第 2 行第 1 列; a_{13} 等于 0, 它位于第 1 行第 3 列; 等等.

- 3×3 **的单位矩阵**.

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个矩阵的行数与列数都为 3, 它是一个 3×3 矩阵. 这类行数和列数相同的矩阵统称为**方阵**, 在后续的行列式、逆矩阵和特征值等内容中会反复出现.

- 3×1 **的列向量**.

$$C=\begin{pmatrix}7\\-5\\1\end{pmatrix}.$$

只有一列的矩阵被称为**列向量**, 可以看作是竖着写的一组数, 常用来表示空间中一点的坐标或一个方向.

- 1×4 **的行向量**.

$$D=\begin{pmatrix}8 & 0 & -2 & 6\end{pmatrix}.$$

只有一行的矩阵被称为**行向量**, 可以看作是横着写的一组数.

定义 22.3 (方阵) 当一个矩阵的行数与列数相等时, 即 $m=n$, 我们称之为一个 n **阶方阵**, 也常称为 n 阶矩阵.

方阵在线性代数中处处出现. 行列式、逆矩阵、特征值和特征向量, 都是首先在方阵上定义的.

22.2.1.1 矩阵作为平面或空间的变换

上面的定义把矩阵看作排成表格的数. 从几何的角度看, 更重要的是, 矩阵可以用来描述平面或空间中的线性几何变换.

例如, 在平面上, 每个点都可以用一个列向量来表示.

$$(x,y)\in\mathbb{R}^2 \quad\longleftrightarrow\quad \begin{pmatrix}x\\y\end{pmatrix}.$$

给定一个 2×2 矩阵

$$A=\begin{pmatrix}a & b\\c & d\end{pmatrix},$$

可让它作用在这个向量上.

$$A\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}ax+by\\cx+dy\end{pmatrix}.$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

几何上, 这表示平面上点 (x,y) 被矩阵 A 送到新点

$$(x',y')=(ax+by,\ cx+dy).$$

于是, 矩阵 A 描述了一次把整个平面统一移动的过程, 例如整体拉伸、压缩、剪切或者绕原点旋转等.

例题 22.1 (拉伸与压缩) 考虑

$$S=\begin{pmatrix}2 & 0\\ 0 & \frac{1}{2}\end{pmatrix}.$$

则

$$S\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}2x\\ \frac{1}{2}y\end{pmatrix}.$$

几何上, 这个矩阵把所有点的 x 坐标放大 2 倍, 把 y 坐标缩小为原来的一半. 整个平面在水平方向被拉伸, 在竖直方向被压缩.

例题 22.2 (绕原点旋转 90°) 考虑

$$R=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}.$$

对任意点 (x,y),

$$R\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}-y\\ x\end{pmatrix}.$$

在平面中, 把点 (x,y) 绕原点逆时针旋转 90° 后的坐标正好是 $(-y,x)$. 因此, 矩阵 R 准确地表示了绕原点逆时针旋转 90° 这一几何变换.

以后, 当我们写 $A\mathbf{x}$(矩阵 A 乘以列向量 $\mathbf{x}$)时, 可以同时借助两个图像来理解.

- 代数上. 按行列相乘的规则得到一个新的向量;
- 几何上.$\mathbf{x}$ 对应的点在平面, 或空间中被某个几何变换送到了新位置.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

22.2.2 矩阵的运算

有了记号 $A=(a_{ij})$ 之后, 可非常简洁地描述矩阵之间的运算. 从几何的角度看, 这些运算也会对应到把变换叠加、把变换整体缩放等操作(后面在矩阵乘法中会看到).

22.2.2.1 矩阵加法与标量乘法

矩阵的加法与数乘运算都是逐元素进行的. 设 $A=(a_{ij})$ 和 $B=(b_{ij})$ 是两个**大小相同**的 $m\times n$ 矩阵, k 是一个常数, 也称为**标量**. 定义

$$A+B:=(a_{ij}+b_{ij})_{m\times n},\qquad kA:=(ka_{ij})_{m\times n}.$$

知乎@余命数

即:

- **矩阵加法**. 两个矩阵相加, 得到一个同大小的新矩阵, 新矩阵的每一个元素, 都是原矩阵对应位置元素之和;
- **标量乘法**. 一个标量乘以一个矩阵, 得到一个同大小的新矩阵, 新矩阵的每一个元素, 都是原矩阵对应位置元素与该标量的乘积.

矩阵的减法 $A-B$ 定义为加法与标量乘法的组合.

$$A-B:=A+(-1)B.$$

例题 22.3 设 $A=\begin{pmatrix}1&-2\\3&0\end{pmatrix}, B=\begin{pmatrix}4&5\\-1&6\end{pmatrix}$. 计算 $A+B$.

知乎@余命数

解 矩阵加法按对应位置元素求和, 故

$$\begin{aligned}A+B&=\begin{pmatrix}1&-2\\3&0\end{pmatrix}+\begin{pmatrix}4&5\\-1&6\end{pmatrix}\\&=\begin{pmatrix}1+4&-2+5\\3+(-1)&0+6\end{pmatrix}\\&=\begin{pmatrix}5&3\\2&6\end{pmatrix}.\end{aligned}$$

知乎@余命数

例题 22.4 设矩阵 $C=\begin{pmatrix}2 & 0 & -1\\ 4 & 5 & 1\end{pmatrix}$. 计算 $3C$.

解 标量乘法将标量 3 分配到矩阵的每个元素, 故

$$\begin{aligned}3C &= 3\begin{pmatrix}2 & 0 & -1\\ 4 & 5 & 1\end{pmatrix}\\ &= \begin{pmatrix}3\cdot 2 & 3\cdot 0 & 3\cdot(-1)\\ 3\cdot 4 & 3\cdot 5 & 3\cdot 1\end{pmatrix}\\ &= \begin{pmatrix}6 & 0 & -3\\ 12 & 15 & 3\end{pmatrix}.\end{aligned}$$

例题 22.5 设 $X=\begin{pmatrix}8 & 1\\ 0 & 4\end{pmatrix}, Y=\begin{pmatrix}-2 & 3\\ 5 & 1\end{pmatrix}$. 计算 $2X-3Y$.

解 $2X-3Y$ 是标量乘法与矩阵减法的组合. 由标量乘法逐元素进行,

$$2X=2\begin{pmatrix}8 & 1\\ 0 & 4\end{pmatrix}=\begin{pmatrix}16 & 2\\ 0 & 8\end{pmatrix},\qquad 3Y=3\begin{pmatrix}-2 & 3\\ 5 & 1\end{pmatrix}=\begin{pmatrix}-6 & 9\\ 15 & 3\end{pmatrix}.$$

再将两者按对应位置相减即得

$$\begin{aligned}2X-3Y &= \begin{pmatrix}16 & 2\\ 0 & 8\end{pmatrix}-\begin{pmatrix}-6 & 9\\ 15 & 3\end{pmatrix}\\ &= \begin{pmatrix}16-(-6) & 2-9\\ 0-15 & 8-3\end{pmatrix}\\ &= \begin{pmatrix}22 & -7\\ -15 & 5\end{pmatrix}.\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

22.2.3 矩阵的乘法

矩阵的加法与标量乘法是逐元素定义的, 这非常直观; 但矩阵*乘法*采用线性变换复合对应的规则. 这一定义源自线性变换的复合与几何变形的叠加.

假设有两组变量, 其中第一组变量 $\{y_1, y_2\}$ 可以由第二组变量 $\{x_1, x_2, x_3\}$ 的线性组合来表示.

$$\begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + a_{13}x_3, \\ y_2 &= a_{21}x_1 + a_{22}x_2 + a_{23}x_3. \end{aligned}$$

线性关系可以写成矩阵形式

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \text{记作} y = Ax.$$

A 把 3 维空间 $\mathbb{R}^3$ 中的向量 x 映射到 2 维空间 $\mathbb{R}^2$ 中的向量 y.

再假设有第三组变量 $\{z_1, z_2\}$ 由 $\{y_1, y_2\}$ 线性表示.

$$\begin{aligned} z_1 &= b_{11}y_1 + b_{12}y_2, \\ z_2 &= b_{21}y_1 + b_{22}y_2, \end{aligned}$$

同样写成矩阵形式

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \text{记作} z = By.$$

现在我们要找一个*直接*的线性关系, 把 $\{z_1, z_2\}$ 写成 $\{x_1, x_2, x_3\}$ 的线性组合. 为此, 将 $y = Ax$ 代入到 $z = By$ 中推导.

$$\begin{aligned} z_1 &= b_{11}(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) + b_{12}(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) \\ &= (b_{11}a_{11} + b_{12}a_{21})x_1 + (b_{11}a_{12} + b_{12}a_{22})x_2 + (b_{11}a_{13} + b_{12}a_{23})x_3, \\ z_2 &= b_{21}(a_{11}x_1 + a_{12}x_2 + a_{13}x_3) + b_{22}(a_{21}x_1 + a_{22}x_2 + a_{23}x_3) \\ &= (b_{21}a_{11} + b_{22}a_{21})x_1 + (b_{21}a_{12} + b_{22}a_{22})x_2 + (b_{21}a_{13} + b_{22}a_{23})x_3. \end{aligned}$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

于是 z 与 x 的直接关系也可以写成矩阵形式 $z = Cx$, 其中

$$C = \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} & b_{11}a_{13} + b_{12}a_{23} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} & b_{21}a_{13} + b_{22}a_{23} \end{pmatrix}.$$

注意到 $z = By = B(Ax) = (BA)x = Cx$, 因此我们自然地把 C 记作 BA, 称为 B 与 A 的**乘积**.(这里之所以记为 BA, 原因是从变量关系看, 是先对 x 施加 A, 再施加 B; 在矩阵算符的表达中, 作用顺序总是从右往左的.)

仔细观察 C 的元素, 例如

$$c_{11} = b_{11}a_{11} + b_{12}a_{21} = \begin{pmatrix} b_{11} & b_{12} \end{pmatrix} \cdot \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix},$$

它是 B 的第一行与 A 的第一列的内积. 一般地, c_{ij} 是 B 的第 i 行与 A 的第 j 列的内积. 这就是行乘列求和计算规则的由来.

定义 22.4 (矩阵乘法) 设 A 为 $m \times n$ 矩阵, B 为 $p \times m$ 矩阵. 若 B 的列数 (m) 等于 A 的行数 (m), 则乘积 $C = BA$ 有定义, 并且是一个 $p \times n$ 矩阵. 它的第 i 行第 j 列元素定义为

$$c_{ij} = (B \text{ 的第 } i \text{ 行}) \cdot (A \text{ 的第 } j \text{ 列}) = \sum_{k=1}^{m} b_{ik}a_{kj}.$$

上述推导给出了代数上的严谨定义. 不过单看代数, 这条规则仍显生硬; 转到**几何**视角, 把它读作线性变换的复合, 动机就清楚了.

在 $\mathbb{R}^2$ 中, 标准基向量为

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

任何向量 x 都可以写成 $x = x_1e_1 + x_2e_2$. 当取一个 2×2 矩阵 A 作用在 x 上时, 利用线性性, 有

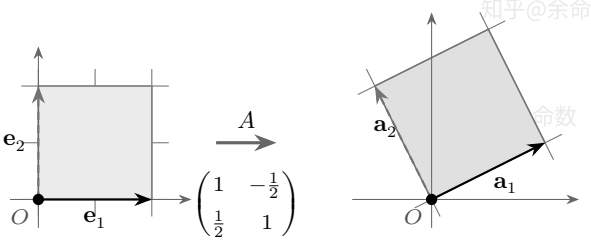
$$Ax = A(x_1e_1 + x_2e_2) = x_1(Ae_1) + x_2(Ae_2).$$

直接计算 A 对基向量的作用, 得

$$Ae_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \qquad Ae_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

说明了一个基本的几何事实. **矩阵 A 的第 j 列, 是基向量 e_j 变换后的像**. 即, 矩阵是对空间怎么被拉伸、旋转、剪切的说明书, 列向量 a_1, a_2 描述了变换后新坐标轴的位置.



(a) 变换前　　(b) 变换后

图 22.1　矩阵的几何意义: A 的第 j 列 $\mathbf{a}_j = A\mathbf{e}_j$ 是标准基向量 $\mathbf{e}_j$ 的像, 单位正方形变换为平行四边形.

既然矩阵代表了线性变换, 两个矩阵相乘 BA 就对应两次变换的复合.

设 $A = (a_1 \dots a_n)$, 将 A 看作第一个变换, 它把标准基 e_j 变成了 a_j. 随后, B 作为第二个变换, 继续作用在这些新轴上. 于是, 复合变换 BA 对 e_j 的作用结果就是 $B(Ae_j) = Ba_j$. 即, BA 的第 j 列, 等于 B 作用在 A 的第 j 列上:

$$BA = \begin{pmatrix} | & & | \\ Ba_1 & \cdots & Ba_n \\ | & & | \end{pmatrix}.$$

从几何上看, 矩阵乘法实质上就是**连续做两次几何变换**. 先按照 A 描述的方式拉伸或旋转, 再按照 B 描述的方式继续变形.

除了上述的列视角, 还可以从**行视角**来审视乘法. 把 B 的第 i 行 r_i 看作一个线性测量函数, 或共变向量, 那么元素 $c_{ij} = r_i \cdot a_j$ 就代表先用 A 把基向量 e_j 送到空间中的某个位置 a_j, 再用 B 的第 i 个测量尺 r_i 去测量这个位置的投影或数值. 列视角读作基变换, 行视角读作对像的线性测量, 两种视角落在同一运算上.

这种变换复合的观点, 在二维几何操作中体现得尤为清楚. 例如, 逆时针旋转 90° 对应的矩阵是 $R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 而放大 2 倍对应的矩阵是 $S = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$. 如果我们想先放大, 再旋转, 其整体效果就由乘积 $M = RS$ 一次性描述. 这种思想也是计算机图形学与神经网络的基础, 无论是屏幕上的像素坐标变换, 还是深度学习中特征向量在层与层之间的流动, 都是一连串矩阵乘法的叠加.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

然而, 空间的变换顺序往往是不能随意的. 考虑剪切变换.

$$S_x = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} (\text{水平剪切}), \quad S_y = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} (\text{竖直剪切}).$$

知乎@余命数

若是先水平剪切再竖直剪切 (S_yS_x), 网格会变成一种形态; 若是反过来 (S_xS_y), 网格则会变成完全不同的另一种形态. 计算结果也证实了这一点, 得

$$S_yS_x = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = S_xS_y.$$

这便解释了矩阵乘法**不满足交换律** ($AB \neq BA$) 的几何根源. 在多维空间中, 变换的次序决定了最终的姿态.

既然矩阵乘法对应着线性变换的复合,这自然就对参与运算的矩阵的**形状**,即维度提出了严格的匹配要求. 我们不能随意地将两个矩阵相乘, 除非它们的维度相容.

从几何上讲, 如果矩阵 A 代表从 n 维空间到 m 维空间的变换, 记作 $A : \mathbb{R}^n \to \mathbb{R}^m$, 而矩阵 B 代表从 q 维空间到 p 维空间的变换, 记作 $B : \mathbb{R}^q \to \mathbb{R}^p$, 那么要进行复合运算 $B(Ax)$, 必须要求 A 的输出能够毫无阻碍地成为 B 的输入.A 的输出空间维度 m 必须严格等于 B 的输入空间维度 q.

这就引出了矩阵乘法**何时可行**的代数铁律.

只有当左边矩阵的**列数**等于右边矩阵的**行数**时, 乘法才有定义.

可用一个简单的符号图示来表示这个规则. 设 A 是 $m \times n$ 矩阵, B 是 $n \times p$ 矩阵, 计算乘积 AB, 得

$$\underbrace{(m \times \mathbf{n})}_{\text{左矩阵}A} \quad \times \quad \underbrace{(\mathbf{n} \times p)}_{\text{右矩阵}B} \quad = \quad \underbrace{(m \times p)}_{\text{结果矩阵}C}$$

注意中间相邻的两个数字, 即 **n** 与 **n**, 它们必须相等, 并在运算过程中消去, 最终结果矩阵的形状由两端剩下的数字 m 和 p 决定.

不能相乘的情形. 一旦中间的这两个维度不匹配, 乘法就是未定义的. 原因可以从几何与代数两方面看.

- **几何上:** 维度错位.A 输出的向量维数与 B 接受的输入维数不同, 复合便无从谈起.
- **代数上:** 内积无法计算. 矩阵乘法的要点是左行乘右列. 如果左矩阵的一行有 3 个元素, 即左矩阵有 3 列, 而右矩阵的一列只有 2 个元素, 即右矩阵有 2 行, 当我们试图将它们对应相乘时, 会发现元素个数不一致, 无法完成点积运算.

知乎@余命数

用一个具体的例子来辨析这一点. 设

$$A=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\ (2\times 3),\qquad B=\begin{pmatrix}1&0\\0&1\end{pmatrix}\ (2\times 2).$$

1. **计算 BA：** B 是 $2\times\mathbf{2}$, A 是 $\mathbf{2}\times 3$. 中间维度均为 2, 匹配成功. 结果将是一个 2×3 的矩阵.
2. **计算 AB：** A 是 $2\times\mathbf{3}$, B 是 $\mathbf{2}\times 2$. 中间维度 $3\neq 2$, 匹配失败. 因此, AB 是不存在的. 我们无法用 A 的一行 $(1,2,3)$ 去和 B 的一列 $(1,0)^T$ 做内积, 因为前者长度为 3, 后者长度为 2.

这个例子再次印证了矩阵乘法的非交换性. 很多时候, BA 有意义, 而 AB 甚至连定义都没有.

理解了背后的几何意义, 我们再回过头来看具体的计算过程. 实际上, 每一个元素的计算 $c_{ij}=\sum b_{ik}a_{kj}$ 都可以通过下图这种直观的方式进行索引.

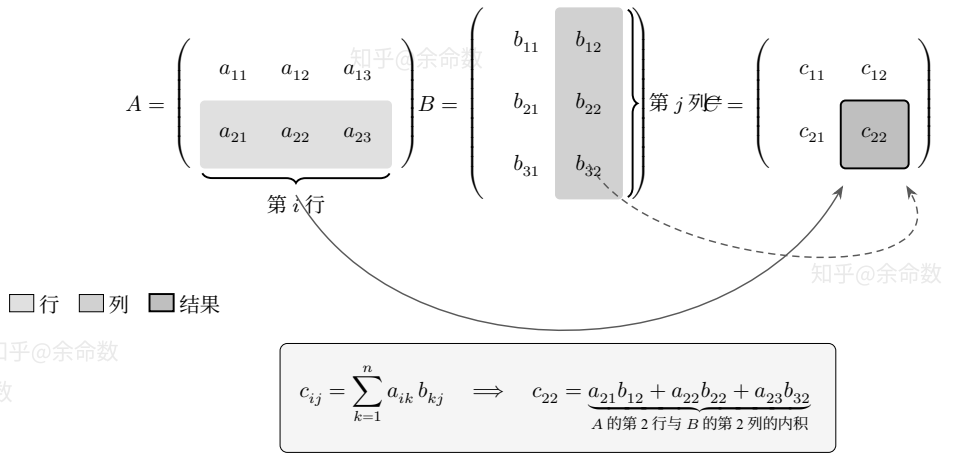


图 22.2　矩阵乘法的计算规则中, C 的第 (i,j) 元素等于 A 的第 i 行与 B 的第 j 列的内积.

再以一个具体的数值计算演示这一过程. 设

$$A=\begin{pmatrix}1&0&2\\3&1&0\end{pmatrix},\quad B=\begin{pmatrix}4&1\\0&5\\2&3\end{pmatrix}.$$

知乎@余命数
知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由于 A 是 2×3, B 是 3×2, 乘积 AB 将是一个 2×2 矩阵. 按行乘列求和, 第一行第一列元素为 $1\cdot 4+0\cdot 0+2\cdot 2=8$, 第二行第一列元素为 $3\cdot 4+1\cdot 0+0\cdot 2=12$, 余下各元同理计算, 最终得到

$$AB=\begin{pmatrix}8 & 7\\ 12 & 8\end{pmatrix}.$$

知乎@余命数

最后,这种看似复杂的乘法规则是 19 世纪数学家们在解决同一根本问题时的发现. 早在 19 世纪初, 高斯在处理线性方程组的变量代换时, 就已经在手稿中使用了这种行乘列的组合方式, 但他当时仅将其视为一种计算技巧.

真正赋予其独立生命的是英国数学家亚瑟·凯莱 (Arthur Cayley). 凯莱在研究不变量理论时, 为了描述坐标变换的复合, 即我们前文提到的将 $x\to x'\to x''$ 合并为一步, 发现必须定义这样一种乘法规则.1858 年, 凯莱引入矩阵这一概念, 将几何变换的复合与代数上的矩阵乘法对应起来. 自此, 矩阵超越了方程组系数的缩写, 成为描述线性结构与空间变换的代数工具.

知乎@余命数

22.2.4 矩阵的基本性质

与我们熟悉的实数运算相似, 矩阵的运算也遵循一套明确的代数法则. 掌握这些性质, 是进行更复杂的矩阵计算与推导的基础.

22.2.4.1 矩阵加法与标量乘法的性质

矩阵的加法与标量乘法性质与向量的性质非常相似, 它们共同构成了一个线性空间. 设 A,B,C 为同阶矩阵, k,l 为标量.

知乎@余命数

定义 22.5 (零矩阵) 所有元素均为 0 的矩阵称为**零矩阵**, 记作 O.

22.2.4.2 矩阵乘法的性质

矩阵乘法的性质更为特殊, 特别是它不满足交换律. 设 A,B,C 为阶数适合于下列运算的矩阵, k 为标量.

知乎@余命数

定义 22.6 (单位矩阵) 一个 n 阶方阵, 如果其主对角线上的元素全为 1, 其余元素全为 0, 则称其为 n **阶单位矩阵**, 记作 I 或 I_n.

$$I_2=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad I_3=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

知乎@余命数

即便 AB 和 BA 都有定义且阶数相同, 也绝不能默认为 $AB=BA$.

22.2.4.3 矩阵的转置

转置将矩阵的行与列进行互换.

定义 22.7 (转置矩阵) 将一个 $m\times n$ 矩阵 $A=(a_{ij})$ 的所有元素沿主对角线进行翻转, 得到的 $n\times m$ 矩阵称为 A 的**转置矩阵**, 记作 A^T. 其元素满足 $(A^T)_{ij}=a_{ji}$.

$$\text{例如, } A=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix},\quad \text{则}\quad A^T=\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}$$

知乎@余命数

知乎@余命数

例题 22.6 验证性质 $(AB)^T=B^TA^T$. 设 $A=\begin{pmatrix}1&2\\0&3\end{pmatrix}, B=\begin{pmatrix}4&1\\2&5\end{pmatrix}$.

命数

解 性质 $(AB)^T=B^TA^T$ 的验证归结为分别计算等式两侧的转置乘积并比较. 由矩阵乘法按行乘列求和,

$$AB=\begin{pmatrix}1&2\\0&3\end{pmatrix}\begin{pmatrix}4&1\\2&5\end{pmatrix}=\begin{pmatrix}1\cdot4+2\cdot2&1\cdot1+2\cdot5\\0\cdot4+3\cdot2&0\cdot1+3\cdot5\end{pmatrix}=\begin{pmatrix}8&11\\6&15\end{pmatrix},$$

沿主对角线翻转即得左端 $(AB)^T=\begin{pmatrix}8&6\\11&15\end{pmatrix}$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

转置将行列互换, 故 $A^T=\begin{pmatrix}1&0\\2&3\end{pmatrix}$, $B^T=\begin{pmatrix}4&2\\1&5\end{pmatrix}$, 右端按 B^T 的行乘 A^T 的列展开为

$$B^TA^T=\begin{pmatrix}4&2\\1&5\end{pmatrix}\begin{pmatrix}1&0\\2&3\end{pmatrix}=\begin{pmatrix}4\cdot1+2\cdot2&4\cdot0+2\cdot3\\1\cdot1+5\cdot2&1\cdot0+5\cdot3\end{pmatrix}=\begin{pmatrix}8&6\\11&15\end{pmatrix}.$$

两侧结果相同, 性质得到验证.

方阵配备的行列式既是可逆性的判据, 也是面积与体积的代数编码, 见第22.3节.

22.3 行列式

上一节把矩阵建成可运算的对象. 本节给每个方阵配一个标量, 即行列式, 用来判断可逆性, 并读出面积/体积的符号信息.

引入

对于每一个**方阵**, 我们都可以计算一个与之对应的唯一的标量, 这个标量被称为该方阵的**行列式**. 行列式是一个基本的概念, 它浓缩了方阵的相关信息, 例如与该方阵对应的线性方程组是否有唯一解, 以及它所代表的几何变换是放大、缩小还是反转了空间.

先从最简单的二阶方阵出发来定义行列式.

定义 22.8 (二阶行列式) 对于一个二阶方阵

$$A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

我们定义, 从左上角到右下角的元素序列 (a,d) 称为**主对角线**, 从右上角到左下角的元素序列 (b,c) 称为**副对角线**.

该方阵的**行列式**的值, 定义为主对角线元素之积减去副对角线元素之积, 记作 $\det(A)$ 或 $|A|$.

$$\det(A)=\begin{vmatrix}a&b\\c&d\end{vmatrix}=ad-bc \tag{22-1}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这个定义来自二元一次方程组的通解公式. 考虑方程组

$$a_1x+b_1y=c_1$$

$$a_2x+b_2y=c_2$$

其解为 $x=\frac{c_1b_2-c_2b_1}{a_1b_2-a_2b_1}, y=\frac{a_1c_2-a_2c_1}{a_1b_2-a_2b_1}$. 分母 $a_1b_2-a_2b_1$ 是系数矩阵 $A=\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ 的行列式.

知乎@余命数

定义 22.9 (三阶行列式) 对于一个三阶方阵

$$A=\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

其**行列式**的值定义为

$$\det(A)=\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}=a_1b_2c_3+b_1c_2a_3+c_1a_2b_3-a_3b_2c_1-b_3c_2a_1-c_3a_2b_1 \tag{22-2}$$

对于三阶行列式的计算, 一种直观的记忆法则是**对角线法则** (Sarrus' Rule). 将矩阵的前两列复制到右侧, 然后将三条主对角线方向的元素之积相加, 再减去三条副对角线方向的元素之积.

知乎@余命数

$$\begin{array}{ccc|cc} a_1 & b_1 & c_1 & a_1 & b_1 \\ a_2 & b_2 & c_2 & a_2 & b_2 \\ a_3 & b_3 & c_3 & a_3 & b_3 \end{array}$$

$$\det(A)=(a_1b_2c_3+b_1c_2a_3+c_1a_2b_3)-(a_3b_2c_1+b_3c_2a_1+c_3a_2b_1)$$

对角线法则仅适用于三阶行列式, 不能推广至更高阶的行列式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 22.7 计算矩阵 $A=\begin{pmatrix}1 & 2 & 3\\ 0 & 4 & 5\\ -1 & 0 & 6\end{pmatrix}$ 的行列式.

解 由对角线法则, 三条主对角线方向的元素之积相加再减去三条副对角线方向的元素之积,

$$\begin{aligned}\det(A) &= (1\cdot 4\cdot 6+2\cdot 5\cdot(-1)+3\cdot 0\cdot 0)-((-1)\cdot 4\cdot 3+0\cdot 5\cdot 1+6\cdot 0\cdot 2)\\ &= (24-10+0)-(-12+0+0)\\ &= 14-(-12)\\ &= 26.\end{aligned}$$

对更高阶的行列式, 通用的计算方法是**代数余子式展开**. 任意 n 阶行列式的值, 都等于其任意一行, 或一列的各元素与其对应的代数余子式乘积之和. 对于三阶行列式, 沿第一行展开的形式为.

$$\begin{vmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{vmatrix}=a_1\begin{vmatrix}b_2 & c_2\\ b_3 & c_3\end{vmatrix}-b_1\begin{vmatrix}a_2 & c_2\\ a_3 & c_3\end{vmatrix}+c_1\begin{vmatrix}a_2 & b_2\\ a_3 & b_3\end{vmatrix}$$

展开项的符号遵循 $(-1)^{i+j}$ 的棋盘格规则, 其中 i 和 j 分别是元素的行号和列号. 第一行的符号即为 $+,-,+$.

22.3.1 行列式的几何应用

行列式最耐人寻味之处在于它的几何含义.

22.3.1.1 三角形面积

先阐述公式的几何原理, 再通过代数推导给出严格证明.

22.3.1.1.0.1 几何原理 公式的几何要点是把三角形的面积与两条边向量所张成的平行四边形面积联系起来. 考虑由三个顶点 $A(x_1,y_1)$、$B(x_2,y_2)$、$C(x_3,y_3)$ 构成的三角形. 可从顶点 A 出发, 构造两条边向量, 得

$$\vec{AB}=(x_2-x_1,y_2-y_1)$$

$$\vec{AC}=(x_3-x_1,y_3-y_1)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

三角形 ABC 的面积, 恰好是由向量 $\vec{AB}$ 和 $\vec{AC}$ 作为邻边所构成的平行四边形面积的一半.

在二维坐标系中, 由两个向量 $\vec{u}=(u_x,u_y)$ 和 $\vec{v}=(v_x,v_y)$ 所张成的平行四边形的**有向面积**, 恰好由二阶行列式给出.

$$\text{平行四边形的面积}=\begin{vmatrix}u_x & u_y\\ v_x & v_y\end{vmatrix}=u_xv_y-u_yv_x$$

这个值的正负号取决于从 $\vec{u}$ 旋转到 $\vec{v}$ 的方向(逆时针为正, 顺时针为负). 因此, 三角形 ABC 的面积可以表示为.

$$S=\frac{1}{2}\left\|\begin{vmatrix}x_2-x_1 & y_2-y_1\\ x_3-x_1 & y_3-y_1\end{vmatrix}\right\| \tag{22-3}$$

我们的任务就是证明, 那个三阶行列式的值实质上等于这个二阶行列式.

我们从目标三阶行列式出发, 利用行列式的一条基本性质化简, 即把某一行, 或列的倍数加到另一行, 或列, 行列式的值不变.

令目标行列式为

$$D=\begin{vmatrix}x_1 & y_1 & 1\\ x_2 & y_2 & 1\\ x_3 & y_3 & 1\end{vmatrix}.$$

为在行列式中构造出向量 $\vec{AB}$ 和 $\vec{AC}$ 的分量, 把第 1 行乘以 -1 分别加到第 2 行和第 3 行 ($R_2\to R_2-R_1, R_3\to R_3-R_1$), 得

$$D=\begin{vmatrix}x_1 & y_1 & 1\\ x_2-x_1 & y_2-y_1 & 1-1\\ x_3-x_1 & y_3-y_1 & 1-1\end{vmatrix}=\begin{vmatrix}x_1 & y_1 & 1\\ x_2-x_1 & y_2-y_1 & 0\\ x_3-x_1 & y_3-y_1 & 0\end{vmatrix}.$$

第三列已含两个零元素, 沿此列作代数余子式展开便十分简洁:

$$\begin{aligned}D&=1\cdot\begin{vmatrix}x_2-x_1 & y_2-y_1\\ x_3-x_1 & y_3-y_1\end{vmatrix}-0+0\\ &=\begin{vmatrix}x_2-x_1 & y_2-y_1\\ x_3-x_1 & y_3-y_1\end{vmatrix}.\end{aligned}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

这表明三阶行列式 D 的值与基于向量法得到的二阶行列式完全相等, 代入方程 (22-3) 即完成公式证明.

$$S = \frac{1}{2}|D| = \frac{1}{2}\left\|\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}\right\| \tag{22-4}$$

行列式本身计算的是一个**有向面积**. 其正负取决于顶点 $A \to B \to C$ 的排列顺序 (逆时针为正, 顺时针为负). 因此, 取其绝对值来获得通常意义下的面积. 将此行列式沿第三列展开, 得

$$\begin{aligned} \det &= 1\begin{vmatrix} x_2 & y_2 \\ x_3 & y_3 \end{vmatrix} - 1\begin{vmatrix} x_1 & y_1 \\ x_3 & y_3 \end{vmatrix} + 1\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \\ &= (x_2y_3 - x_3y_2) - (x_1y_3 - x_3y_1) + (x_1y_2 - x_2y_1) \\ &= (x_1y_2 + x_2y_3 + x_3y_1) - (x_2y_1 + x_3y_2 + x_1y_3) \end{aligned}$$

这个表达式是著名的**鞋带公式**. 它是该三阶行列式经过代数展开后的自然结果.

22.3.1.2 三点共线与直线方程

面积公式立即给出两个推论. 其一, 三点 $A(x_1, y_1)$、$B(x_2, y_2)$、$C(x_3, y_3)$ 共线的充要条件是它们所构成的三角形面积为零.

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0$$

其二, 给定两点 $A(x_1, y_1)$ 和 $B(x_2, y_2)$, 平面上任意一点 $P(x, y)$ 在直线 AB 上的充要条件是 P, A, B 三点共线. 由此, 经过 A, B 两点的直线方程可以写成.

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

例题 22.8 求经过点 $A(1, 2)$ 和 $B(3, -4)$ 的直线方程.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 由行列式形式的直线方程,

$$\begin{vmatrix} x & y & 1 \\ 1 & 2 & 1 \\ 3 & -4 & 1 \end{vmatrix} = 0$$

知乎@余命数

沿第一行展开此行列式得

$$x\begin{vmatrix} 2 & 1 \\ -4 & 1 \end{vmatrix} - y\begin{vmatrix} 1 & 1 \\ 3 & 1 \end{vmatrix} + 1\begin{vmatrix} 1 & 2 \\ 3 & -4 \end{vmatrix} = 0,$$

即

$$x(2-(-4)) - y(1-3) + 1(-4-6) = 0,$$

化简得 $6x + 2y - 10 = 0$, 亦即直线方程 $3x + y - 5 = 0$.

22.3.1.3 向量外积与法向量

在三维空间中, 行列式是定义向量外积的标准工具. 对于两个向量 $\vec{a} = (a_1, a_2, a_3)$ 和 $\vec{b} = (b_1, b_2, b_3)$, 它们的外积 $\vec{a} \times \vec{b}$ 是一个同时垂直于 $\vec{a}$ 和 $\vec{b}$ 的新向量, 其定义为一个形式上的行列式.

$$\vec{a} \times \vec{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \quad (22\text{-}5)$$

其中 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 是沿 x, y, z 轴的单位向量. 这个工具在解析几何中用于直接计算平面法向量. 一个由三点 A, B, C 确定的平面, 其法向量 $\vec{n}$ 必定垂直于平面内的任意向量, 例如 $\vec{AB}$ 和 $\vec{AC}$. 因此, 计算这两个向量的外积即得法向量.

$$\vec{n} = \vec{AB} \times \vec{AC}$$

例题 22.9 求经过三点 $A(1,0,0)$、$B(0,1,0)$、$C(0,0,1)$ 的平面的一个法向量.

解 在平面内取两个向量, 由顶点坐标相减得

$$\vec{AB} = B - A = (0-1, 1-0, 0-0) = (-1, 1, 0), \qquad \vec{AC} = C - A = (0-1, 0-0, 1-0) = (-1, 0, 1).$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

平面的法向量 $\vec{n}$ 即这两个向量的外积. 按形式行列式展开,

$$\begin{aligned}\vec{n} = \vec{AB} \times \vec{AC} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{vmatrix} \\ &= \mathbf{i}\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - \mathbf{j}\begin{vmatrix} -1 & 0 \\ -1 & 1 \end{vmatrix} + \mathbf{k}\begin{vmatrix} -1 & 1 \\ -1 & 0 \end{vmatrix} \\ &= \mathbf{i}(1-0) - \mathbf{j}(-1-0) + \mathbf{k}(0-(-1)) \\ &= 1\mathbf{i} + 1\mathbf{j} + 1\mathbf{k}.\end{aligned}$$

因此该平面的一个法向量是 $\vec{n} = (1, 1, 1)$.

22.3.1.4 平行六面体的体积

行列式的几何意义可以自然地推广到三维体积. 出发点是一个基本的几何事实. 任何柱体的体积都等于其底面积乘以高. 在这里, 将使用向量代数的工具来精确地表达底面积和高.

一个平行六面体的体积 V 可以通过其底面平行四边形的面积 $S_{\text{底}}$ 与垂直于该底面的高 h 的乘积来计算.

$$V = S_{\text{底}} \cdot h$$

我们不妨选择由向量 $\vec{b}$ 和 $\vec{c}$ 所张成的平行四边形作为底面. 根据向量外积的几何意义, 这个底面平行四边形的面积等于外积向量 $\vec{b} \times \vec{c}$ 的模长.

$$S_{\text{底}} = \|\vec{b} \times \vec{c}\|$$

根据外积的定义, 向量 $\vec{n} = \vec{b} \times \vec{c}$ 的方向是同时垂直于 $\vec{b}$ 和 $\vec{c}$ 的, 即垂直于底面. 平行六面体的高 h, 就是第三个向量 $\vec{a}$ 在底面法向量 $\vec{n}$ 方向上的投影的长度. 一个向量 $\vec{a}$ 在另一个向量 $\vec{n}$ 上的标量投影 (用 proj 表示) 由点积给出.$\text{proj}_{\vec{n}}\vec{a} = \frac{\vec{a}\cdot\vec{n}}{\|\vec{n}\|}$. 因此, 高 h 是这个投影值的绝对值, 因为高度必须为非负.

$$h = \left|\frac{\vec{a} \cdot \vec{n}}{\|\vec{n}\|}\right| = \frac{|\vec{a} \cdot (\vec{b} \times \vec{c})|}{\|\vec{b} \times \vec{c}\|}$$

接着, 将底面积和高的表达式代入体积公式.

$$V = S_{\text{底}} \cdot h = \|\vec{b} \times \vec{c}\| \cdot \frac{|\vec{a} \cdot (\vec{b} \times \vec{c})|}{\|\vec{b} \times \vec{c}\|}$$

知乎@余命数

分子与分母中的 $\|\vec{b}\times\vec{c}\|$ 项相互抵消, 得一个极为简洁的表达式.

$$V=|\vec{a}\cdot(\vec{b}\times\vec{c})|$$

这个表达式 $\vec{a}\cdot(\vec{b}\times\vec{c})$ 被称为三个向量的**混合积**. 它在几何上代表了由这三个向量所张成的平行六面体的**有向体积**.

最后一步是证明混合积的代数计算结果与三阶行列式的值完全相同. 设 $\vec{a}=(a_1,a_2,a_3)$、$\vec{b}=(b_1,b_2,b_3)$、$\vec{c}=(c_1,c_2,c_3)$. 由外积的形式行列式定义,

$$\vec{b}\times\vec{c}=(b_2c_3-b_3c_2)\mathbf{i}+(b_3c_1-b_1c_3)\mathbf{j}+(b_1c_2-b_2c_1)\mathbf{k}.$$

将 $\vec{a}$ 与此结果作点积, 按分量相乘求和得

$$\begin{aligned}\vec{a}\cdot(\vec{b}\times\vec{c})&=a_1(b_2c_3-b_3c_2)+a_2(b_3c_1-b_1c_3)+a_3(b_1c_2-b_2c_1)\\&=a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1.\end{aligned}$$

另一方面, 将以这三个向量为行构成的三阶行列式沿第一行展开 (按代数余子式展开定理),

$$\begin{aligned}\begin{vmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{vmatrix}&=a_1\begin{vmatrix}b_2&b_3\\c_2&c_3\end{vmatrix}-a_2\begin{vmatrix}b_1&b_3\\c_1&c_3\end{vmatrix}+a_3\begin{vmatrix}b_1&b_2\\c_1&c_2\end{vmatrix}\\&=a_1(b_2c_3-b_3c_2)-a_2(b_1c_3-b_3c_1)+a_3(b_1c_2-b_2c_1)\\&=a_1b_2c_3-a_1b_3c_2-a_2b_1c_3+a_2b_3c_1+a_3b_1c_2-a_3b_2c_1.\end{aligned}$$

将两个展开式各项重排后完全相同, 这就证明了混合积与行列式在数值上等价. 因此, 平行六面体的体积可以表示为.

$$V=\left\|\begin{vmatrix}a_1&a_2&a_3\\b_1&b_2&b_3\\c_1&c_2&c_3\end{vmatrix}\right\| \tag{22-6}$$

由此可知, 三个向量共面的充要条件是它们无法张成一个具有体积的实体, 即它们所张成的平行六面体体积为零. 这等价于它们的混合积为零, 即上述行列式的值为零.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

例题 22.10 (平面截距式方程) 设平面 π 与 x,y,z 坐标轴分别交于三点

$$A(a,0,0),\quad B(0,b,0),\quad C(0,0,c)$$

其中 a,b,c 均为非零实数. 求该平面的方程.

解 平面方程可由共面条件确定. 设 $P(x,y,z)$ 是该平面上的任意一点, 则向量 $\vec{AP}$、$\vec{AB}$、$\vec{AC}$ 必共面; 共面的充要条件是它们张成的平行六面体体积为零, 即三者构成的行列式为零. 由顶点坐标相减,

$$\vec{AP}=(x-a,y-0,z-0)=(x-a,y,z),\quad \vec{AB}=(0-a,b-0,0-0)=(-a,b,0),\quad \vec{AC}=(0-a,0-0,c-0)$$

令其构成的行列式等于零, 得

$$\begin{vmatrix} x-a & y & z \\ -a & b & 0 \\ -a & 0 & c \end{vmatrix}=0.$$

沿第一行展开此行列式,

$$(x-a)\begin{vmatrix} b & 0 \\ 0 & c \end{vmatrix}-y\begin{vmatrix} -a & 0 \\ -a & c \end{vmatrix}+z\begin{vmatrix} -a & b \\ -a & 0 \end{vmatrix}=0,$$

即

$$(x-a)(bc)-y(-ac)+z(0-(-ab))=0,$$

整理得

$$bcx-abc+acy+abz=0.$$

将常数项 $-abc$ 移到等式右边,

$$bcx+acy+abz=abc.$$

由于 a,b,c 均非零, 以 abc 同除等式两边, 便得到平面方程的著名形式, 即**截距式方程**:

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

22.3.2 行列式与多项式因式分解

行列式本身是元素的多线性组合, 因此一旦元素中含有未知量, 行列式自然就是关于这些未知量的一个多项式. 这一点使得行列式成为因式分解 (尤其是高次、多元、具有对称性或轮换对称性的多项式)的一个有效工具.

总体思路可以概括为两种互补的方向.

- *从行列式出发做因式分解*. 把行列式看成多项式, 利用代入使行列式为零来找因式;
- *从多项式出发构造行列式*. 先把一个给定多项式巧妙地改写成一个行列式的值, 再利用行列式的初等变换和公因式提取来实现因式分解.

这两种方法的理论基础都来自下列两条简单却非常有力的原理.

22.3.2.1 根与因式的关系(因式定理)

对于一个关于变量 x 的一元多项式 $P(x)$, 因式定理指出.

$$\text{如果} P(a)=0, \text{那么} (x-a) \text{是} P(x) \text{的一个因式}.$$

多元情形下, 若把某个变量看作未知, 其余变量看作参数, 这一原理同样成立. 例如, 设 $P(a,b,c)$ 是关于 a 的多项式(b,c 只是系数所在的参数). 若

$$P(b,b,c)=0,$$

则在多项式环 $\mathbb{R}[a,b,c]$ 中, $(a-b)$ 就是 $P(a,b,c)$ 的一个因式. 把 a 换成 b 会使多项式变成 0, 即 $(a-b)$ 整除这个多项式.

22.3.2.2 行列式的零值条件

若行列式的任意两行或两列成比例(特别是完全相同), 则该行列式的值为零.

这源于行列式对每一行, 或列的线性性和交错性. 当两行 (列) 相同或成比例时, 行列式中相应项彼此抵消, 结果必为 0.

把这两个原理结合, 就得到利用行列式寻找多项式因式的基本逻辑链条.

1. 把一个含有变量的行列式 $D(x,y,z,\dots)$ 视作关于这些变量的多项式;
2. 通过对变量赋特定值, 如令 $x=y$ 或令某变量为 0,1 等, 观察此时行列式的两行或两列是否会变得相同或成比例;
3. 一旦在这种代入下出现两行, 或两列相同, 或成比例, 就知道此时行列式的值必为 0;

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

4. 因此, 按因式定理, 对应的表达式, 如 $x-y$、$a+b+c$ 等就是该行列式所代表多项式的一个因式.

对于范德蒙行列式、轮换对称行列式等高度对称的结构, 这一逻辑链条往往能系统地逼出所有本应出现的线性因式, 如 $(x-y)$、$(a-b)$、$(a+b+c)$ 等, 再配合次数比较和系数比较, 就可以完成因式分解.

22.3.2.3 典型用途概览

行列式方法在因式分解中的常见用途包括.

- 处理范德蒙类行列式, 得到简洁的积形式, 如 $\prod_{i<j}(x_i-x_j)$;
- 对轮换(循环)对称多项式/行列式自动检出因式 $(a+b+c)$、$(a-b)(b-c)(c-a)$ 等;
- 在题目给出的是普通多项式时, 逆向构造行列式, 通过初等变换和公因式提取完成分解;
- 证明某些多项式恒等式, 例如 $a^3+b^3+c^3-3abc=(a+b+c)(a^2+b^2+c^2-ab-bc-ca)$ 时, 行列式方法往往给出结构清晰的证明.

分别从两种方向整理具体操作步骤.

22.3.2.4 一、从行列式出发做因式分解

设 $D(a,b,c,\dots)$ 是由若干多项式元素构成的行列式.

1. **寻找线性因式(通过造零)**
仔细观察行列式的结构, 尝试通过代入

$$a=b,\quad b=c,\quad a=-b,\quad a+b+c=0,\quad x=0,1,\dots$$

等特殊关系, 使得两行或两列变得相同/成比例, 从而保证 $D=0$. 每得到一次这样的必为 0 的代入关系, 就对应找到了一个因式.
例如.

   - 若令 $a=b$ 时 $D=0$, 则 $(a-b)$ 是 D 的一个因式;
   - 对于轮换对称行列式, 如果 $(a-b)$ 是因式, 那么通常 $(b-c)$、$(c-a)$ 也是因式;
   - 若总能在 $a+b+c=0$ 时得到 $D=0$, 往往 $(a+b+c)$ 就是一个因式.

2. **比较次数**
把 D 展开成多项式后, 可以讨论它的总次数(多元情形可以看作齐次多项式的次数, 或专门看作关于某一指定变量的次数).

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对大多数有规律的行列式, 只需观察每行中次数最高的项, 从每行各取一次就能得到最高次项. 例如, 对

$$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix}$$

每一行的最高次数分别为 $0,1,2$, 故整个行列式是一个三次齐次多项式.
同时, 计算前一步已找到的所有因式之积的总次数. 若两者次数相同, 则行列式与这些因式之积至多只差一个常数因子.

3. **确定常数因子**
   若已知

$$D = k \cdot (\text{因子}_1) \cdot (\text{因子}_2) \cdots,$$

且左右两边次数相同, 则只需要比较某个具体单项式的系数即可确定 k.
常用做法是选取行列式展开中较容易定位的某一单项式, 例如主对角线上元素乘积产生的那一项, 在因式乘积一侧也找到同一单项式, 比较系数即可. 若 $k=1$, 则说明因式分解已经完备.

22.3.2.5 二、从多项式出发构造行列式

当题目一开始给出的是普通多项式 $P(x)$ 或 $P(a,b,c,\dots)$ 时, 也可以*主动*把它写成行列式的形式, 然后用上面的行列式技巧来说明因式结构.

1. **构造行列式**
   这是最具创造性的一步. 目标是把 P 通过分组、配凑等方式写成

$$P = A(x)D(x) - B(x)C(x)$$

这样的 $AD-BC$ 结构, 从而可以写成一个二阶行列式.

$$P(x) = \begin{vmatrix} A(x) & B(x) \\ C(x) & D(x) \end{vmatrix}.$$

有时也可以构造成更高阶行列式, 但能够用 2×2 解决时往往最简洁.

2. **初等变换制造公因式**
   利用行列式的初等变换性质.

   - 某一行(列)乘以一个数加到另一行(列)上, 行列式的值不变;
   - 调换两行(列)只会改变符号;
   - 某一行(列)乘以常数 k, 行列式整体乘以 k.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

通过这些变换，力图使某一行或某一列的每一项都含有同一个因式，例如 $(x+3)$、$(x-1)$、$(a+b+c)$ 等.

3. **提取公因式**
一旦某行或某列出现了公因式，就可根据行列式的多线性性质把它提出行列式之外.

$$\begin{vmatrix} f(x)\,g(x) & f(x)\,h(x) \\ \cdots & \cdots \end{vmatrix} = f(x) \begin{vmatrix} g(x) & h(x) \\ \cdots & \cdots \end{vmatrix}.$$

4. **化简并重复**
对提取了公因式后剩下的、更简单的行列式，继续做初等变换与公因式提取，直至可以直接计算其值. 最终就得到

$$P = (\text{若干因式}) \times (\text{一个较简单的多项式}),$$

这样 P 的因式结构就显现出来了.

下列例题展示这两种思路在具体问题中的运用，既有标准范德蒙行列式及其变体，也有轮换对称行列式以及经巧妙代数构造得到的 2×2 行列式.

例题 22.11 因式分解多项式 $P(x)=x^3+4x^2+x-6$.

解 把 $P(x)$ 写成一个 2×2 行列式. 对多项式分组，

$$P(x)=(x^3+4x^2)+(x-6)=x^2(x+4)-(-1)(x-6),$$

于是

$$P(x)=\begin{vmatrix} x^2 & x-6 \\ -1 & x+4 \end{vmatrix},$$

因为

$$\begin{vmatrix} x^2 & x-6 \\ -1 & x+4 \end{vmatrix} = x^2(x+4)-(x-6)(-1)=x^3+4x^2+x-6.$$

对行列式做初等变换以制造公因式. 做列变换 $C_2\to C_2+C_1$ 后，

$$P(x)=\begin{vmatrix} x^2 & x^2+x-6 \\ -1 & x+3 \end{vmatrix},$$

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

其中二次三项式 $x^2+x-6=(x+3)(x-2)$, 故

$$P(x)=\begin{vmatrix} x^2 & (x+3)(x-2) \\ -1 & x+3 \end{vmatrix}.$$

此时第二列有公因式 $(x+3)$, 按行列式列的多线性提出,

$$P(x)=(x+3)\begin{vmatrix} x^2 & x-2 \\ -1 & 1 \end{vmatrix}.$$

再做行变换 $R_1 \to R_1+R_2$ 进一步化简,

$$P(x)=(x+3)\begin{vmatrix} x^2-1 & x-1 \\ -1 & 1 \end{vmatrix}.$$

由于 $x^2-1=(x-1)(x+1)$, 第一行有公因式 $(x-1)$, 提出,

$$P(x)=(x+3)(x-1)\begin{vmatrix} x+1 & 1 \\ -1 & 1 \end{vmatrix}.$$

计算剩下的 2×2 行列式,

$$\begin{vmatrix} x+1 & 1 \\ -1 & 1 \end{vmatrix}=(x+1)\cdot 1-1\cdot(-1)=x+2,$$

因此

$$P(x)=(x-1)(x+2)(x+3).$$

本例属于从多项式出发构造行列式的思路. 先凑出 $AD-BC$, 再用行列式变换与公因式提取完成因式分解.

例题 22.12 (范德蒙行列式) 因式分解三阶行列式

$$D=\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix}.$$

解 把 D 视为关于 x,y,z 的一个多项式.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

先通过造零寻找因式. 若令 $x=y$, 则第一列与第二列完全相同, 得

$$\begin{vmatrix} 1 & 1 & 1 \\ y & y & z \\ y^2 & y^2 & z^2 \end{vmatrix} = 0,$$

即 $x=y$ 时 $D=0$, 由因式定理知 $(x-y)$ 是 D 的一个因式. 同理, 由于 x,y,z 的地位完全对称, 令 $y=z$ 得 $(y-z)$ 是一个因式, 令 $z=x$ 得 $(z-x)$ 是一个因式. 于是 D 至少含有因式

$$(x-y)(y-z)(z-x).$$

再比较次数. 第 k 行的元素是 $x^{k-1}, y^{k-1}, z^{k-1}$, 其次数分别为 $0,1,2$, 故整个行列式是关于 x,y,z 的三次齐次多项式; 而因式乘积 $(x-y)(y-z)(z-x)$ 的总次数也是 3. 因此存在常数 k 使得

$$D = k(x-y)(y-z)(z-x).$$

最后比较系数确定 k. 在左边的展开式中, yz^2 项只可能来自主对角线 $1\cdot y\cdot z^2$, 故其系数为 $+1$. 右边展开

$$(x-y)(y-z)(z-x) = -xz^2 - y^2z + yz^2 - x^2y + x^2z + xy^2,$$

其中 yz^2 的系数也是 $+1$, 故 $1 = k\cdot 1$, 即 $k=1$.

故

$$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix} = (x-y)(y-z)(z-x).$$

本例是从行列式出发的标准范例. 用行列式的零值条件找线性因式, 用次数与系数比较确定常数因子.

例题 22.13 因式分解行列式

$$D = \begin{vmatrix} x & y & z \\ x^2 & y^2 & z^2 \\ x^3 & y^3 & z^3 \end{vmatrix}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 每一列都含公因式：第 1 列有 x，第 2 列有 y，第 3 列有 z. 由行列式列的多线性将这三个因式提出，

$$D=xyz\begin{vmatrix}1 & 1 & 1\\ x & y & z\\ x^2 & y^2 & z^2\end{vmatrix}.$$

括号内是三阶范德蒙行列式，其因式分解刚刚已求得为 $(x-y)(y-z)(z-x)$，故

$$D=xyz(x-y)(y-z)(z-x).$$

本例先利用公因式把行列式简化，再套用已知模型.

例题 22.14　因式分解轮换对称行列式

$$D=\begin{vmatrix}a & b & c\\ b & c & a\\ c & a & b\end{vmatrix}.$$

解 该行列式具有明显的轮换对称性，把 $a\to b$, $b\to c$, $c\to a$ 作循环置换，行列式的形式保持不变. 这种对称性提示因式中很可能包含对称的表达式，例如 $(a+b+c)$.

为造出 $(a+b+c)$，将第 2 列与第 3 列同时加到第 1 列上 $(C_1\to C_1+C_2+C_3)$，得

$$D=\begin{vmatrix}a+b+c & b & c\\ b+c+a & c & a\\ c+a+b & a & b\end{vmatrix}.$$

此时第一列每个元素都是 $(a+b+c)$，按行列式列的多线性将其提出，

$$D=(a+b+c)\begin{vmatrix}1 & b & c\\ 1 & c & a\\ 1 & a & b\end{vmatrix}.$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

再做行变换 $R_2 \to R_2 - R_1,\ R_3 \to R_3 - R_1$ 制造零元以降阶,

$$D = (a+b+c)\begin{vmatrix} 1 & b & c \\ 0 & c-b & a-c \\ 0 & a-b & b-c \end{vmatrix},$$

沿第一列展开便化为一个二阶行列式,

$$D = (a+b+c)\cdot 1 \cdot \begin{vmatrix} c-b & a-c \\ a-b & b-c \end{vmatrix}.$$

计算此 2×2 行列式并整理,

$$\begin{aligned} \begin{vmatrix} c-b & a-c \\ a-b & b-c \end{vmatrix} &= (c-b)(b-c) - (a-c)(a-b) \\ &= -(b-c)^2 - (a^2 - ab - ac + bc) \\ &= -(a^2+b^2+c^2 - ab - bc - ca), \end{aligned}$$

于是

$$D = -(a+b+c)(a^2+b^2+c^2-ab-bc-ca).$$

另一方面, 利用恒等式

$$a^3+b^3+c^3-3abc = (a+b+c)(a^2+b^2+c^2-ab-bc-ca),$$

可知

$$D = -(a^3+b^3+c^3-3abc) = 3abc - a^3 - b^3 - c^3.$$

本例的要领是用恰当的列(行)变换主动造出 $(a+b+c)$ 这样的对称因式, 再经降阶与整理, 轮换对称行列式的因式结构便清晰可见.

行列式给出体积缩放因子. 第22.4节 把矩阵当作几何变换的算子. 旋转、反射、切变, 以及齐次坐标下的平移.

22.4 线性变换与几何应用

在 第22.2节 至 第22.3节 中, 矩阵作为数表与行列式的载体出现. 本节转向几何, 矩阵同时是空间变换的算子, 也用来记录坐标.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

旋转、镜像、切变、缩放等几何变换可由矩阵统一描述; 行列式度量面积的变化与翻转; 矩阵语言可直接判断一条二次曲线是圆、椭圆、双曲线还是抛物线; 中学解析几何里的配方与旋转坐标系, 在高观点下是线性代换(矩阵变换).

22.4.0.1 线性变换的矩阵表示

当我们说一个映射 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是**线性变换**时, 意味着它保持了向量加法和标量乘法的结构. 原点保持不动, 直线变换后仍为直线, 或退化为点, 且平行的刻度保持均匀.

任何平面上的线性变换都可以由一个二阶方阵 A 唯一确定. 如前所述, 寻找这个矩阵的要点在于追踪**基向量**的去向. 设标准基向量

$$e_1=\begin{pmatrix}1\\0\end{pmatrix},\qquad e_2=\begin{pmatrix}0\\1\end{pmatrix},$$

若它们在变换 T 下的像分别为

$$T(e_1)=v_1,\qquad T(e_2)=v_2,$$

则该变换对应的矩阵为.

$$A=\begin{pmatrix}v_1 & v_2\end{pmatrix},\qquad T(x)=Ax.$$

即. **知道基向量往哪儿去, 就等于知道了整个线性变换**.

22.4.0.2 典型变换

在解析几何中, 几种最常用的线性变换可以统一写成矩阵.

- **各向同性缩放**(等比放缩).

$$S_\lambda=\begin{pmatrix}\lambda & 0\\0 & \lambda\end{pmatrix},\quad (x,y)^T\mapsto(\lambda x,\lambda y)^T.$$

- **各向异性缩放**(在不同方向上拉伸不同倍数).

$$D_{a,b}=\begin{pmatrix}a & 0\\0 & b\end{pmatrix},\quad (x,y)^T\mapsto(ax,by)^T.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- **关于坐标轴的镜像**.

$$x\text{-轴对称：}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\qquad y\text{-轴对称：}\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}.$$

知乎@余命数

- **关于原点的旋转**(逆时针角 θ).

知乎@余命数

$$R_\theta=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

- **切变** (Shear).

$$\text{水平切变：}\ S_k=\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix},\qquad (x,y)^T\mapsto(x+ky,y)^T.$$

知乎@余命数

下列例子展示矩阵与几何的对应.

知乎@余命数

例题 22.15 (镜像变换) 求平面上关于直线 $y=x$ 对称的变换矩阵.

解 追踪基向量的像即可确定变换矩阵.x 轴上的单位向量 $e_1=(1,0)^T$ 关于直线 $y=x$ 对称后落在 y 轴上, 变为 $v_1=(0,1)^T$; y 轴上的单位向量 $e_2=(0,1)^T$ 对称后落在 x 轴上, 变为 $v_2=(1,0)^T$. 以这两个像作为列向量, 变换矩阵为

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

$$P=\begin{pmatrix}v_1 & v_2\end{pmatrix}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}.$$

验证. 对任意点 $(a,a)^T$,

$$P\begin{pmatrix}a\\ a\end{pmatrix}=\begin{pmatrix}a\\ a\end{pmatrix},$$

保持不动; 点 $(1,0)^T$ 变成 $(0,1)^T$, 与几何直观一致.

例题 22.16 (旋转变换) 求绕原点逆时针旋转角 θ 的线性变换矩阵.

知乎@余命数

解　同样考察基向量的像. 点 $(1,0)$ 绕原点逆时针旋转 θ 后落在单位圆上的点 $(\cos\theta,\sin\theta)$, 故

$$T(e_1)=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}.$$

知乎@余命数

点 $(0,1)$ 绕原点逆时针旋转 θ 后, 其极角从 $\frac{\pi}{2}$ 变为 $\frac{\pi}{2}+\theta$, 坐标为

知乎@余命数

$$(\cos(\tfrac{\pi}{2}+\theta),\ \sin(\tfrac{\pi}{2}+\theta))=(-\sin\theta,\ \cos\theta),$$

故

知乎@余命数

$$T(e_2)=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}.$$

以这两个像作为列向量, 旋转矩阵为

$$R_\theta=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

这个矩阵在高中解析几何中非常重要, 它描述坐标系的旋转, 并暗含了三角恒等式, 例如

$$R_\alpha R_\beta=R_{\alpha+\beta}$$

知乎@余命数

从矩阵乘法写开, 就得到熟悉的

知乎@余命数

知乎@余命数

$$\cos(\alpha+\beta)=\cos\alpha\cos\beta-\sin\alpha\sin\beta,\quad \sin(\alpha+\beta)=\sin\alpha\cos\beta+\cos\alpha\sin\beta.$$

例题 22.17 (切变与面积不变性)　考察水平切变矩阵 $S_k=\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}$ 对单位正方形的作用, 验证它保持面积不变.

解　单位正方形的顶点为 $(0,0),(1,0),(1,1),(0,1)$. 变换后四点分别为

$$(0,0),\ (1,0),\ (1+k,1),\ (k,1),$$

它们构成一个平行四边形, 其两条邻边向量为

$$\vec{u}=(1,0)^T,\quad \vec{v}=(k,1)^T.$$

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

平行四边形面积等于

$$|\det(\vec{u},\vec{v})| = \begin{Vmatrix} 1 & k \\ 0 & 1 \end{Vmatrix} = 1,$$

与原单位正方形面积相同. 这与 $|\det(S_k)| = 1$ 完全一致, 见下一小节.

22.4.0.3 行列式的几何意义

在 第22.3节 中, 二阶行列式给出两个向量围成的平行四边形的有向面积.

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc.$$

如果将这两列看作两个向量 $\vec{u} = (a,b)^T$, $\vec{v} = (c,d)^T$, 则

$$|\det(\vec{u},\vec{v})|$$

就是它们围成平行四边形的面积.

更一般地, 对于线性变换 $y = Ax$, 如果输入空间中的一个图形具有面积 S, 那么变换后的图形面积 S' 与原面积之间存在确定的倍率关系.

定理 22.1 (面积变换定理) 设 A 为二阶方阵, Ω 是平面上的一个可测图形, 其面积为 S_Ω. 经矩阵 A 变换后, 图形 $\Omega' = \{Ax \mid x \in \Omega\}$ 的面积为.

$$S_{\Omega'} = |\det(A)| \cdot S_\Omega.$$

证明 先看单位正方形. 设 A 的两列分别为

$$\mathbf{u} = (a,c)^T, \qquad \mathbf{v} = (b,d)^T.$$

单位正方形在变换 $x \mapsto Ax$ 下变为由 $\mathbf{u},\mathbf{v}$ 张成的平行四边形, 其面积为

$$|\det(\mathbf{u},\mathbf{v})| = |ad - bc| = |\det(A)|.$$

任意平行四边形可由单位正方形再经一个线性变换得到, 因而同样按 $|\det(A)|$ 的比例改变面积. 将平行四边形沿对角线分成两个三角形, 可知三角形面积也按同一比例改变; 再把多边形剖分为有限个三角形, 结论对任意多边形成立.

对一般可测图形, 采用面积定义中的多边形逼近. 可用内接多边形和外接多边形把它的面积夹住, 且两者面积差任意小. 若 $\det(A) = 0$, 则 A 的像至多是一条

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

直线, 面积为 0, 公式成立. 若 $\det(A)\neq 0$, 线性变换保持包含关系; 对夹逼的多边形使用上一步的结论, 再令多边形逼近误差趋于 0, 即得

$$S_{\Omega'}=|\det(A)|S_{\Omega}.$$

□

因此, 行列式绝对值的几何含义是. **它是线性变换改变面积的比例系数.**

- $|\det(A)|>1$: 变换产生放大效果.
- $0<|\det(A)|<1$: 变换产生缩小效果.
- $\det(A)=0$: 图形被压扁, 面积变为 0(退化为直线或点).
- $\det(A)<0$: 变换包含了一个镜像翻转操作, 改变了图形的手性(定向).

例题 22.18 (三角形面积的行列式公式) 已知三角形三个顶点

$$P_1(x_1,y_1),\quad P_2(x_2,y_2),\quad P_3(x_3,y_3)$$

证明其面积为

$$S=\frac{1}{2}\left|\begin{vmatrix}x_2-x_1 & x_3-x_1\\ y_2-y_1 & y_3-y_1\end{vmatrix}\right|.$$

解 以 P_1 为原点, 向量

$$\vec{u}=P_2-P_1=(x_2-x_1,\ y_2-y_1)^T,\quad \vec{v}=P_3-P_1=(x_3-x_1,\ y_3-y_1)^T$$

构成三角形的两条邻边. 它们围成的平行四边形面积为

$$|\det(\vec{u},\vec{v})|=\left|\begin{vmatrix}x_2-x_1 & x_3-x_1\\ y_2-y_1 & y_3-y_1\end{vmatrix}\right|.$$

三角形面积是该平行四边形的一半, 故结论成立.

例题 22.19 (椭圆面积公式) 利用单位圆的变换推导长半轴为 a、短半轴为 b 的椭圆面积.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 单位圆 $x^2+y^2\le 1$ 的面积已知为 π, 标准椭圆方程为 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}\le 1$. 构造对角矩阵

$$A=\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

将单位圆拉伸为该椭圆, 对单位圆上的任意向量 $\mathbf{u}=(u_1,u_2)^T$, 变换后的向量

$$\mathbf{x}=A\mathbf{u}=(au_1,bu_2)^T$$

满足

$$\left(\frac{x}{a}\right)^2+\left(\frac{y}{b}\right)^2=u_1^2+u_2^2\le 1,$$

故该变换确实将单位圆双射到了椭圆区域.

变换矩阵的行列式 $\det(A)=ab$, 由面积变换定理,

$$S_{\text{椭圆}}=|\det(A)|\cdot S_{\text{单位圆}}=ab\cdot\pi=\pi ab.$$

这个推导说明了一种思路. **通过线性变换把复杂图形拉回到熟悉的标准图形, 再用行列式把面积缩放回来**. 这是许多高等积分技巧的雏形.

22.4.0.4 齐次坐标与平移

线性变换有一个固有的局限. 它必须固定原点 ($A\mathbf{0}=\mathbf{0}$). 即单纯的 2×2 矩阵乘法无法表示**平移** (Translation).

为了统一处理平移与旋转, 解析几何中引入了**齐次坐标**.

定义 22.10 (齐次坐标) 平面点 (x,y) 的齐次坐标表示为三维列向量

$$\tilde{x}=\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

通过增加一个维度, 可用 3×3 的矩阵来描述二维平面上的**仿射变换** (Affine Transformation). 例如.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

- 平移向量 (h,k):

$$T_{h,k}=\begin{pmatrix}1&0&h\\0&1&k\\0&0&1\end{pmatrix},\qquad T_{h,k}\begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}x+h\\y+k\\1\end{pmatrix}.$$

- 先旋转再平移.

$$\tilde{A}=\begin{pmatrix}\cos\theta&-\sin\theta&h\\\sin\theta&\cos\theta&k\\0&0&1\end{pmatrix},$$

知乎@余命数

则

$$\tilde{A}\begin{pmatrix}x\\y\\1\end{pmatrix}=\begin{pmatrix}x\cos\theta-y\sin\theta+h\\x\sin\theta+y\cos\theta+k\\1\end{pmatrix}.$$

这样, 所有旋转 + 缩放 + 切变 + 平移的组合都可以表示为一个 3×3 矩阵的一次相乘. 计算机图形学中, 三维空间里的物体变换同理使用 4×4 齐次坐标矩阵.

例题 22.20 (组合变换一次完成) 已知点 $P(1,0)$, 先绕原点逆时针旋转 $\frac{\pi}{2}$, 再向右平移 2 个单位, 求变换后的坐标, 并给出相应的齐次矩阵.

解 绕原点旋转 $\frac{\pi}{2}$ 的矩阵为

$$R_{\frac{\pi}{2}}=\begin{pmatrix}0&-1\\1&0\end{pmatrix}.$$

向右平移 2 的矩阵为

$$T_{2,0}=\begin{pmatrix}1&0&2\\0&1&0\\0&0&1\end{pmatrix}.$$

知乎@余命数

对应的齐次形式(先旋转后平移)是

$$\tilde{A} = \begin{pmatrix} 0 & -1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

代入 P 的齐次坐标 $(1,0,1)^T$:

$$\tilde{A}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix},$$

因此变换后点为 $(2,1)$.

22.4.0.5 二次曲线的矩阵分类 *

中学解析几何中的圆锥曲线, 一般从如下二元二次方程出发.

$$Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F = 0,$$

其中 A, B, C 不全为零. 交叉项 xy 的存在使得曲线发生旋转, 难以直接辨识其类型(椭圆、双曲线或抛物线).

矩阵提供一种更高视角. 先**只看最高次项**, 忽略线性项和常数项, 关注在方程的二次部分, 即二次型.

22.4.0.6 二次型与系数矩阵

方程的二次部分可以写成矩阵形式.

$$q(x,y) = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} A & B \\ B & C \end{pmatrix}}_{M} \begin{pmatrix} x \\ y \end{pmatrix},$$

其中 M 是一个实对称矩阵, 称为**系数矩阵**, q 称为**二次型**.

几何上, M 决定了曲线的开闭形态(椭圆型 / 双曲线型 / 抛物线型). 线性项 $Dx + Ey$ 与常数项 F 会影响曲线的位置、大小以及退化或是否有实点, 但不改变由 $\delta = AC - B^2$ 判定的三大类型.

知乎@余命数

22.4.0.7 线性代换下的不变量

设 $x = Ay$, 其中 A 是任意可逆的 2×2 矩阵, 这相当于做了一个线性坐标变换(旋转、缩放、切变等的组合). 则

$$q(x) = x^T Mx = y^T (A^T MA)y.$$

知乎@余命数

这说明.

- 在新坐标中, 二次项的系数矩阵变成 $M' = A^T MA$;
- $\det(M') = \det(A^T)\det(M)\det(A) = \det(A)^2\det(M)$, 因而**符号**不变.

所以, $\det(M)$ 的符号是一个几何不变量. 不管你怎么选择坐标系, 即怎么做线性代换, 它都不会变号.

22.4.0.8 椭圆/双曲线/抛物线判别

记

$$M = \begin{pmatrix} A & B \\ B & C \end{pmatrix}, \qquad \delta = \det(M) = AC - B^2.$$

知乎@余命数

定理 22.2 (二次曲线分类准则) 对于二次曲线

$$Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F = 0,$$

其中 A, B, C 不全为 0. 令 $\delta = AC - B^2$.

1. 若 $\delta > 0$: 曲线为**椭圆型**(包括圆、点或虚椭圆);
2. 若 $\delta < 0$: 曲线为**双曲线型**(包括相交双直线);
3. 若 $\delta = 0$: 曲线为**抛物线型**(包括平行双直线、重合直线或虚退化情形).

证明 二次项的系数矩阵

$$M = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

是实对称矩阵. 二阶实对称矩阵可以经正交变换(旋转或镜像)对角化, 这里引用线性代数的一个基本事实, 即存在正交矩阵 P, 使得

知乎@余命数

$$P^T MP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

知乎@余命数

作正交换元 $(x,y)^T = P(u,v)^T$ 后, 二次部分化为

$$\lambda_1 u^2 + \lambda_2 v^2,$$

且 $\delta = \det(M) = \lambda_1\lambda_2$.

若 $\delta > 0$, 两个特征值同号且均非零. 将方程必要时整体乘以 -1, 可把二次部分化为正定型. 对 u, v 分别配方, 方程成为

$$\alpha(u-u_0)^2 + \beta(v-v_0)^2 = K \qquad (\alpha,\beta>0),$$

因而按 $K>0, K=0, K<0$ 分别为椭圆、点或虚椭圆.

若 $\delta < 0$, 两个特征值异号. 配方后得到

$$\alpha(u-u_0)^2 - \beta(v-v_0)^2 = K \qquad (\alpha,\beta>0),$$

其中 $K \neq 0$ 时为双曲线, $K=0$ 时退化为两条相交直线, 故为双曲线型.

若 $\delta = 0$, 且 A, B, C 不全为零, 则恰有一个特征值为 0. 旋转后方程形如

$$\lambda u^2 + pu + qv + r = 0 \qquad (\lambda \neq 0).$$

当 $q \neq 0$ 时可解出 v 为 u 的二次函数, 是抛物线; 当 $q=0$ 时方程只含 u, 按一元二次方程判别式退化为两条平行直线、重合直线或无实点的虚退化情形. 因而属于抛物线型. □

这是高中常见判别式 $B^2 - AC$ 的矩阵版. 因为

$$\delta = AC - B^2 = -(B^2 - AC),$$

于是

$$\delta > 0 \Longleftrightarrow B^2 - AC < 0, \quad \delta < 0 \Longleftrightarrow B^2 - AC > 0, \quad \delta = 0 \Longleftrightarrow B^2 - AC = 0,$$

与熟悉的判别法则($B^2 - AC < 0$ 为椭圆, $=0$ 为抛物线, >0 为双曲线)完全一致, 只是写法不同而已.

$\det(M)$ 的符号已经足以判断曲线类型. 如果用一点特征值(此处仅引用事实), 实对称矩阵 M 可以正交对角化.

$$P^T M P = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

知乎@余命数 知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

其中 P 是一个旋转或旋转与镜像复合的正交矩阵, λ_1, λ_2 是实特征值. 换变量

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \begin{pmatrix} u \\ v \end{pmatrix},$$

知乎@余命数

二次项就变成

$$q(x, y) = \lambda_1 u^2 + \lambda_2 v^2.$$

知乎@余命数

此时

$$\delta = \det(M) = \lambda_1 \lambda_2.$$

于是.

- $\delta > 0 \Rightarrow \lambda_1, \lambda_2$ 同号.u^2、v^2 同号, 形如 $\alpha u^2 + \beta v^2 + \cdots = 0$, 为椭圆型;
- $\delta < 0 \Rightarrow \lambda_1, \lambda_2$ 异号. 形如 $u^2 - v^2 + \cdots = 0$, 为双曲线型;
- $\delta = 0 \Rightarrow$ 有一个特征值为 0. 形如 $u^2 + \cdots = 0$ 或 $v^2 + \cdots = 0$, 缺少一个二次方向, 得到抛物线型.

知乎@余命数

这就是定理 22.2 背后的高观点.

例题 22.21 (用行列式快速判别类型) 判断曲线

$$x^2 + 4xy + y^2 = 1$$

的类型.

解 对应的系数矩阵为

$$M = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

其中交叉项 $2B = 4$, 故 $B = 2$. 于是

$$\delta = \det(M) = 1 \cdot 1 - 2 \cdot 2 = -3 < 0,$$

该曲线为**双曲线型**.

如要验证, 可配方.

$$x^2 + 4xy + y^2 = (x + 2y)^2 - 3y^2 = 1,$$

知乎@余命数

令 $u = x + 2y,\ v = \sqrt{3}\, y$, 方程变为 $u^2 - v^2 = 1$, 是标准双曲线.

知乎@余命数

例题 22.22 (椭圆型消去 xy 项) 判断并化简曲线

$$3x^2 - 4xy + 3y^2 = 12.$$

解 先判别类型. 二次项矩阵为

$$M = \begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix}, \qquad \delta = 3 \cdot 3 - (-2)^2 = 9 - 4 = 5 > 0,$$

故为**椭圆型**.

接下来用旋转坐标系消去 xy 项, 在线性代数语言下即找一个正交矩阵 R 使 $M' = R^T M R$ 为对角矩阵. 一个自然的选择是 45° 旋转.

$$R = R_{\frac{\pi}{4}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

按矩阵乘法展开得

$$M' = R^T M R = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}.$$

作变量代换

$$\begin{pmatrix} x \\ y \end{pmatrix} = R \begin{pmatrix} u \\ v \end{pmatrix}, \quad \text{即} \quad \begin{cases} x = \dfrac{u - v}{\sqrt{2}}, \\ y = \dfrac{u + v}{\sqrt{2}}, \end{cases}$$

代入原方程, 二次部分变为二次型

$$\begin{pmatrix} u & v \end{pmatrix} M' \begin{pmatrix} u \\ v \end{pmatrix} = u^2 + 5v^2,$$

故方程化为

$$u^2 + 5v^2 = 12.$$

这是一个以原点为中心、长短轴分别沿 u、v 轴的椭圆.

$$\frac{u^2}{12} + \frac{v^2}{12/5} = 1.$$

与传统先求旋转角 θ 再代入的做法相同, 只是用矩阵 R 更为简洁统一.

知乎@余命数 知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

例题 22.23 (抛物线型 $\delta=0$) 判断并化简曲线

$$x^2+2xy+y^2+2x+2y=0.$$

解 二次项矩阵为

$$M=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix},\qquad \delta=1\cdot 1-1\cdot 1=0,$$

故为**抛物线型**(或其退化形式).

注意到 $x^2+2xy+y^2=(x+y)^2$, 方程可化为

$$(x+y)^2+2(x+y)=0.$$

令 $u=x+y$, 则

$$u^2+2u=0 \iff u(u+2)=0,$$

即原方程是两条直线

$$x+y=0,\quad x+y=-2$$

的并集, 是抛物线型的一个退化情形(平行双直线).

这里以 $x+y$ 为新变量 u, 实质上是做了一次线性代换

$$\begin{pmatrix}u\\ v\end{pmatrix}=P\begin{pmatrix}x\\ y\end{pmatrix}$$

的第一行, 其中 P 的第一行是 $(1,1)$.

22.4.0.9 从配方法到矩阵语言

中学里处理圆锥曲线, 常用两类方法.

1. **配方**消去一次项, 找到中心或顶点, 如

$$x^2+4x+y^2-2y+1=0 \ \Rightarrow\ (x+2)^2+(y-1)^2=4.$$

2. **旋转坐标系**消去 xy 项, 用公式 $\tan 2\theta=\dfrac{2B}{A-C}$.

在矩阵视角下.

- 配方对应**平移**. 对方程

$$x^TMx+b^Tx+F=0,$$

知乎@余命数

令 $x=z+x_0$, 选取 x_0 使得线性项消失, 即 $2Mx_0+b=0$, 就完成了把中心移到原点的操作;

- 旋转对应**正交变换**. 选正交矩阵 P, 令 $x=Py$, 则二次项矩阵变为 P^TMP, 可选 P 使其对角化, 从而消去 xy 项.

二者组合(先平移再旋转)就把任意圆锥曲线化成最熟悉的标准方程. 整个过程只是在对向量做线性代换, 这就是圆锥曲线的线性代数高观点.

例题 22.24 (用矩阵思路整理一条椭圆) 设曲线

$$4x^2+y^2-4x-2y-2=0.$$

用矩阵加线性代换的语言, 把它化成标准椭圆方程, 并求椭圆中心与半轴长.

@余命数

解 二次项矩阵为

知乎@余命数

$$M=\begin{pmatrix}4&0\\0&1\end{pmatrix},\quad \delta=4\cdot1-0=4>0,$$

没有 xy 项且为椭圆型, 故无需旋转, 只需平移(配方)即可.

原方程即

知乎@余命数

$$\begin{pmatrix}x&y\end{pmatrix}\begin{pmatrix}4&0\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}+\begin{pmatrix}-4&-2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}-2=0,$$

知乎@余命数

亦即 $4x^2+y^2-4x-2y-2=0$. 对 x、y 各自配方,

知乎@余命数

$$4x^2-4x=4\Big(x^2-x\Big)=4\Big[(x-\tfrac12)^2-\tfrac14\Big],\qquad y^2-2y=(y-1)^2-1.$$

代回原式并整理常数项,

$$\begin{aligned}4\Big[(x-\tfrac12)^2-\tfrac14\Big]+(y-1)^2-1-2&=0,\\ 4(x-\tfrac12)^2+(y-1)^2-1-1-2&=0\\ \iff 4(x-\tfrac12)^2+(y-1)^2&=4.\end{aligned}$$

两边同除以 4,

$$\frac{(x-\frac12)^2}{1}+\frac{(y-1)^2}{4}=1.$$

这就是以 $C(\frac12,1)$ 为中心、横向半轴长 1、纵向半轴长 2 的椭圆(对应横轴长 2、纵轴长 4).

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在线性代数语言里, 平移

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}$$

就是把新坐标 (u, v) 的原点移到椭圆中心, 然后方程在 (u, v) 坐标下变为

$$4u^2 + v^2 = 4.$$

第22.5节 用逆矩阵与克拉默法则判定变换何时可逆,并给出 $Ax = b$ 在 $\det A \neq 0$ 时的显式解.

22.5 逆矩阵与克拉默法则

知乎@余命数

22.5.1 逆矩阵

上一节把矩阵写成线性变换. 自然要追问, 矩阵 A 对空间做了某种变换后, 能否由另一个矩阵 B 把空间还原回初始状态.

这就引出了**逆矩阵**的概念. 在实数算术中, 非零数 a 存在倒数 a^{-1}, 使得 $a \cdot a^{-1} = 1$. 对矩阵寻求类似的结构, 但必须留意零元素(零矩阵或零行列式)带来的奇异性.

知乎@余命数

22.5.2 变换的可逆性

知乎@余命数

考虑线性方程组 $Ax = b$. 这里的 A 是一个算子, 将输入向量 x 映射为输出向量 b. 如果我们要从 b 反推 x, 实质上是在寻找一个逆算子 A^{-1}, 使得 $x = A^{-1}b$.

若 A 代表逆时针旋转 90°, 那么 A^{-1} 显然应当是顺时针旋转 90°. 若 A 代表放大 2 倍, A^{-1} 则是缩小 2 倍. 然而, 如果 A 是一个投影变换, 例如将三维空间压扁到一个平面上, 或者将所有点都映射到原点(零矩阵), 这种变换显然丢失了信息, 无法复原.

定义 22.11 (逆矩阵) 对于 n 阶方阵 A, 如果存在 n 阶方阵 B, 使得

$$AB = BA = I_n$$

其中 I_n 是单位矩阵, 则称 A 是**可逆的** (Invertible) 或**非奇异的** (Non-singular). 矩阵 B 称为 A 的**逆矩阵**, 记作 A^{-1}.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

注 22.2
定义中要求 $AB=BA$ 看似多余, 因为对于有限维方阵, 仅需 $AB=I$ 即可推出 $BA=I$. 但这一性质并不平凡, 它依赖于线性空间的维数理论. 在几何上, 即左逆(能够撤销变换的算子)必定也是右逆.

22.5.3 伴随矩阵与逆的构造

具体计算一个矩阵的逆, 需要建立矩阵元素与其行列式之间的联系. 回忆代数余子式的展开定理.

$$\sum_{k=1}^{n} a_{ik}A_{ik} = \det(A)$$

其中 A_{ik} 是元素 a_{ik} 对应的代数余子式. 把某一行的元素与另一行的代数余子式相乘, 得到下面的和式. 考虑

$$\sum_{k=1}^{n} a_{ik}A_{jk} \quad (i \neq j)$$

这等价于计算一个新矩阵的行列式, 该矩阵的第 j 行被替换成了第 i 行的复制品. 根据行列式性质, 含有两行相同的行列式值为零. 总结上述两条规律, 引入 Kronecker 记号 δ_{ij}:

$$\sum_{k=1}^{n} a_{ik}A_{jk} = \delta_{ij}\det(A)$$

这提示我们, 如果构造一个由代数余子式组成的矩阵, 并将其转置, 使得行元素与行余子式在乘法中能够对应相碰, 便能构造出单位矩阵的倍数.

定义 22.12 (伴随矩阵) 设 $A=(a_{ij})$ 为 n 阶方阵, 行列式 $|A|$ 的各个元素的代数余子式 A_{ij} 所构成的矩阵的**转置矩阵**.

$$\operatorname{adj}(A) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}^{T} = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

称为矩阵 A 的**伴随矩阵**.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

利用上述的正交性关系, 可直接验证.

$$A\cdot\operatorname{adj}(A)=\begin{pmatrix}\det(A) & 0 & \cdots & 0\\ 0 & \det(A) & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \det(A)\end{pmatrix}=\det(A)I$$

由此导出了逆矩阵存在的充要条件及显式公式.

定理 22.3 (逆矩阵公式) 方阵 A 可逆的充要条件是 $\det(A)\neq 0$. 当此条件满足时,

$$A^{-1}=\frac{1}{\det(A)}\operatorname{adj}(A) \tag{22-7}$$

证明 先证充分性. 由代数余子式展开的正交关系可得

$$A\operatorname{adj}(A)=\det(A)I,\qquad \operatorname{adj}(A)A=\det(A)I.$$

若 $\det(A)\neq 0$, 两边同除以 $\det(A)$, 得

$$A\left(\frac{1}{\det(A)}\operatorname{adj}(A)\right)=I,\qquad \left(\frac{1}{\det(A)}\operatorname{adj}(A)\right)A=I.$$

因而 A 可逆, 且逆矩阵正是公式所给矩阵.

再证必要性. 若 A 可逆, 存在 A^{-1} 使 $AA^{-1}=I$. 取行列式并用行列式乘法公式,

$$\det(A)\det(A^{-1})=\det(I)=1,$$

故 $\det(A)\neq 0$. □

这个定理连接了代数与几何.

- **几何上**.$\det(A)=0$ 意味着变换将空间压扁了, 例如体积变为 0. 被压扁的维度丢失了信息, 无法通过逆变换复原.
- **代数上**.$\det(A)$ 出现在分母位置. 除数不能为零.

例题 22.25 求二阶矩阵 $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ 的逆矩阵.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 行列式 $\det(A) = ad - bc$. 各元素的代数余子式为 $A_{11} = d, A_{12} = -c, A_{21} = -b, A_{22} = a$, 按代数余子式的转置排列得到伴随矩阵

$$\operatorname{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

知乎@余命数

因此, 若 $ad - bc \neq 0$,

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

知乎@余命数

二阶矩阵求逆即主对角元素对调、副对角元素变号, 再除以行列式.

乎@余命数

22.5.4 克拉默法则

对于 n 个未知数 n 个方程的线性方程组 $Ax = b$, 如果系数矩阵 A 可逆, 可直接写出解的代数形式 $x = A^{-1}b$. 将逆矩阵的显式公式代入, 可以得到一个仅仅依赖于行列式的解法(克拉默法则).

22.5.4.1 法则陈述

知乎@余命数
知乎@余命数

定理 22.4 (克拉默法则) 设线性方程组 $Ax = b$ 中 A 为 n 阶方阵、b 为 n 维列向量, 且系数行列式 $D = \det(A) \neq 0$. 则方程组有唯一解, 且第 j 个分量 x_j 为.

$$x_j = \frac{D_j}{D} = \frac{\det(A_j)}{\det(A)} \tag{22-8}$$

其中 D_j 是将矩阵 A 的第 j 列替换为常数项向量 b 后得到的矩阵的行列式.

证明 因 $D = \det(A) \neq 0$, 由逆矩阵公式知 A 可逆, 故方程组有唯一解.

设 A 的列向量为 $\alpha_1, \alpha_2, \ldots, \alpha_n$. 方程 $Ax = b$ 等价于

$$b = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n.$$

知乎@余命数

将 A 的第 j 列换成 b, 并利用行列式对每一列的线性性,

知乎@余命数

$$\begin{aligned} D_j &= \det(\alpha_1, \ldots, \alpha_{j-1}, b, \alpha_{j+1}, \ldots, \alpha_n) \\ &= \sum_{i=1}^{n} x_i \det(\alpha_1, \ldots, \alpha_{j-1}, \alpha_i, \alpha_{j+1}, \ldots, \alpha_n). \end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当 $i \neq j$ 时, 上式中的行列式有两列相同, 故为 0; 只有 $i = j$ 的项保留, 得

$$D_j = x_j \det(\alpha_1, \dots, \alpha_n) = x_j D.$$

由于 $D \neq 0$, 故 $x_j = D_j / D$. □

知乎@余命数

22.5.4.2 几何证明与直观

代数推导(利用伴随矩阵展开)是克拉默法则的标准证法, 但它掩盖了几何意义. 这里给出一个基于体积变换的几何证明.

考察方程 $Ax = b$. 设 A 的列向量为 $\alpha_1, \alpha_2, \dots, \alpha_n$. 方程实际上是说向量 b 是列向量的线性组合.

$$x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = b$$

我们以二维情况为例. 系数行列式 $D = \det(\alpha_1, \alpha_2)$ 代表由 α_1, α_2 构成的平行四边形的有向面积. 考虑 D_1, 即把第一列 α_1 换成 b 所得的行列式

$$D_1 = \det(b, \alpha_2),$$

将 $b = x_1\alpha_1 + x_2\alpha_2$ 代入并利用行列式对每一列的线性性,

$$\begin{aligned} D_1 &= \det(x_1\alpha_1 + x_2\alpha_2,\ \alpha_2) \\ &= x_1 \det(\alpha_1, \alpha_2) + x_2 \det(\alpha_2, \alpha_2). \end{aligned}$$

含有两列相同的行列式 $\det(\alpha_2, \alpha_2) = 0$(由行列式的交错性), 故仅剩第一项

$$D_1 = x_1 \det(\alpha_1, \alpha_2) = x_1 D,$$

由此立即得到 $x_1 = \frac{D_1}{D}$.

这个几何推导可以推广到 n 维.D_j 实际上是将 n 维平行六面体在第 j 个方向上的边长拉伸了 x_j 倍, 同时引入了其他方向的剪切 (Shear). 由于行列式对剪切变换是不变的, 即 $\det(\alpha_i, \alpha_i) = 0$, 唯一改变体积的因素就是第 j 个方向的伸缩系数 x_j.

以上推导说明, **分量 x_j 是两个体积的比率**.

$\det A = 0$ 时不可逆仍有层次之分,空间可能被压成平面、直线或一点. 第22.6节用**秩**精确描述维度塌缩, 并用高斯消元计算它.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

22.6 秩与高斯消元

22.6.1 秩与线性相关性

已在 第22.3节 中了解过. 行列式是否为零区分了方阵的可逆与奇异. 当 $\det(A) \neq 0$ 时矩阵满秩可逆; 当 $\det(A) = 0$ 时矩阵奇异, 但奇异二字内部仍有不同的维数塌缩. 然而, 奇异二字掩盖了太多的细节. 一个 3×3 的矩阵如果是奇异的, 它可能把三维空间压缩成一个二维平面, 也可能压成一条一维直线, 还可能塌缩成一个零维的点. 仅凭行列式为零这一信息, 我们无法区分这些情形. 为了精确描述矩阵所携带信息的**维度**, 需引入**秩** (Rank) 的概念.

22.6.2 线性相关性

矩阵的列向量实质上是一组生成元, 它们通过线性组合生成了像空间. 如果某些列向量可以由其他列向量线性表示, 说明它们提供了冗余的信息.

定义 22.13 (线性相关与线性无关) 设 $v_1, v_2, \dots, v_k$ 为一组向量. 如果存在一组**不全为零**的标量 $c_1, c_2, \dots, c_k$, 使得

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = \mathbf{0}$$

则称这组向量是**线性相关**的. 反之, 如果上述等式仅在 $c_1 = c_2 = \cdots = c_k = 0$ 时成立, 则称这组向量是**线性无关**的.

注 22.3
线性相关的几何意义是扁平化. 两个线性相关的向量共线; 三个线性相关的向量共面. 从数据的角度看, 线性相关意味着冗余. 例如, 在统计表中, 如果总分列等于语文列加数学列, 那么总分这一列就是线性相关的, 它没有提供任何新的独立维度.

22.6.3 秩的定义与几何意义

对于一个 $m \times n$ 的矩阵 A, 可将其看作由 n 个列向量组成, 也可以看作由 m 个行向量组成.

定义 22.14 (秩) 矩阵 A 的**列秩**定义为其线性无关的列向量的极大数量. 矩阵 A 的**行秩**定义为其线性无关的行向量的极大数量.

这里有线性代数中的一个重要定理, 即**秩定理**. 尽管行向量属于 $\mathbb{R}^n$, 列向量属于 $\mathbb{R}^m$, 它们处于完全不同的空间, 但它们的有效维度却是严格相等的.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 22.5 (秩定理) 对于任意矩阵 A, 其行秩等于列秩. 将其统称为矩阵 A 的**秩**, 记作 $\mathrm{rank}(A)$ 或 $r(A)$.

证明 设 A 是 $m \times n$ 矩阵, 行秩为 r. 取 r 个线性无关的行向量

$$R_{p_1}, R_{p_2}, \dots, R_{p_r}$$

作为行空间的一组基. 对任意第 i 行 R_i, 存在常数 $c_{i1}, \dots, c_{ir}$, 使得

$$R_i = \sum_{s=1}^{r} c_{is} R_{p_s}.$$

取第 j 列并比较第 i 个分量, 得

$$a_{ij} = \sum_{s=1}^{r} c_{is} a_{p_s j}.$$

因而第 j 列向量可写为

$$C_j = \sum_{s=1}^{r} a_{p_s j} \, (c_{1s}, c_{2s}, \dots, c_{ms})^T.$$

这说明所有列向量都落在至多 r 个向量张成的空间中, 所以列秩不超过行秩.

对转置矩阵 A^T 重复同样论证, 等价地交换行与列, 可得行秩不超过列秩. 两个不等式合并, 行秩等于列秩. □

秩最直观的解释是**图像的维度**. 若 A 是一个 3×3 矩阵, 它代表一个线性变换 $T : \mathbb{R}^3 \to \mathbb{R}^3$.

- $\mathrm{rank}(A) = 3$: 图像充满了整个三维空间. 变换是可逆的, 体积变换率 (行列式) 非零.
- $\mathrm{rank}(A) = 2$: 整个三维空间被压缩到了一个二维平面上.
- $\mathrm{rank}(A) = 1$: 整个三维空间被压缩到了一条直线上.
- $\mathrm{rank}(A) = 0$: 整个空间被压缩到了原点 (仅当 A 为零矩阵).

秩与行列式的关系可以概括为. 对于 n 阶方阵 A,$\mathrm{rank}(A) = n \iff \det(A) \neq 0$. 更一般地, 矩阵的秩等于其**最高阶非零子式**, 即非零子行列式的阶数.

22.6.4 初等变换与高斯消元

定义了秩之后, 我们面临的实际问题是如何计算一个 10×10 矩阵的秩. 直接寻找最大线性无关组或计算所有子行列式显然是不切实际的. 需一种手段, 在简

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

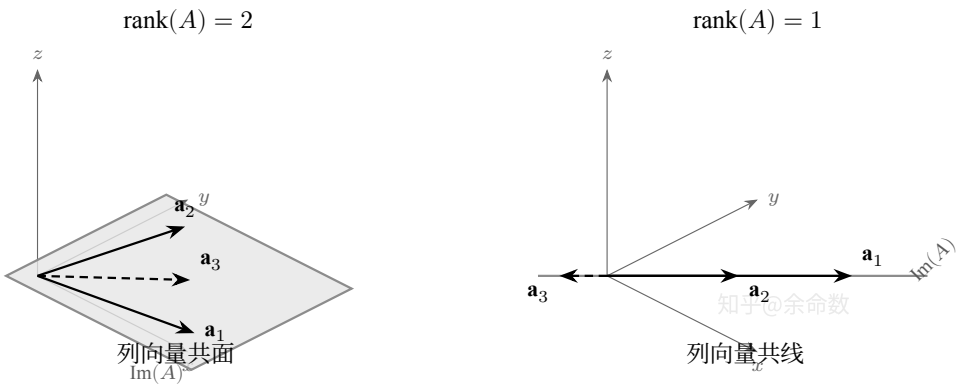


图 22.3 矩阵秩的几何意义：设 $A=[\mathbf{a}_1 \mid \mathbf{a}_2 \mid \mathbf{a}_3]$，则 $\operatorname{rank}(A)=\dim(\operatorname{span}\{\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3\})$. 虚线向量可由其他向量线性表出.

化矩阵形式的同时，严格保持矩阵的秩（以及对应方程组的解集）不变. 这就是**初等变换**.

22.6.5 三种基本操作

考虑求解方程组时的操作. 交换两个方程的位置、将一个方程乘以非零常数、将一个方程的倍数加到另一个方程上. 这些操作不会改变方程组的解. 对应到矩阵上，就是**初等行变换**.

定义 22.15 (初等行变换) 矩阵 A 的三种初等行变换为.

1. **交换**：交换第 i 行和第 j 行的位置. 记作 $R_i \leftrightarrow R_j$.
2. **倍乘**：将第 i 行乘以非零常数 k. 记作 $R_i \leftarrow kR_i$.
3. **倍加**：将第 j 行的 k 倍加到第 i 行上. 记作 $R_i \leftarrow R_i + kR_j$.

同理可定义初等列变换.

三种初等变换都是可逆操作，可逆性保证了秩不变.

命题 22.1 (秩的不变性) 初等变换不改变矩阵的秩. 若矩阵 A 经过初等变换得到 B，则称 A 与 B **等价**，记作 $A \sim B$.

证明 先看初等行变换. 交换两行只改变行向量的排列；某一行乘以非零常数不改变该行所张成的一维方向；把一行替换为它与另一行倍数之和时，新行属于原行向量组的张成空间. 反过来，原来的该行也可由新行减去另一行的相同倍数得到，因而原行空间与新行空间完全相同. 所以初等行变换保持行秩不变.

知乎@余命数

初等列变换同理保持列秩不变. 由秩定理, 矩阵的行秩与列秩相等, 因此无论做初等行变换还是初等列变换, 矩阵的秩都不变. □

这一命题之所以成立, 是因为初等变换实质上只是对原来的行向量组或列向量组做了一个可逆的线性组合. 相应张成空间的维度自然不会变.

22.6.6 矩阵乘法视角的初等变换

知乎@余命数

初等变换不只是操作规则, 它们本身就是矩阵. **对矩阵 A 做初等行变换, 等价于在 A 的左侧乘以一个特定的初等矩阵**.

定义 22.16 (初等矩阵) 对单位矩阵 I 施行一次初等变换所得到的矩阵, 称为**初等矩阵**.

- **交换矩阵** E_{ij}: 由 I 交换第 i, j 行得到.

$$I=\begin{pmatrix}1&0\\0&1\end{pmatrix}\xrightarrow{R_1\leftrightarrow R_2}E_{12}=\begin{pmatrix}0&1\\1&0\end{pmatrix}$$

- **倍乘矩阵** $E_i(k)$: 由 I 的第 i 行乘 k 得到.

$$I=\begin{pmatrix}1&0\\0&1\end{pmatrix}\xrightarrow{R_2\times 3}E_2(3)=\begin{pmatrix}1&0\\0&3\end{pmatrix}$$

知乎@余命数　知乎@余命数

- **倍加矩阵** $E_{ij}(k)$: 由 I 的第 j 行的 k 倍加到第 i 行得到.

$$I=\begin{pmatrix}1&0\\0&1\end{pmatrix}\xrightarrow{R_2+2R_1}E_{21}(2)=\begin{pmatrix}1&0\\2&1\end{pmatrix}$$

验证一下倍加操作. 计算 $E_{21}(2)\cdot A$, 得

$$\begin{pmatrix}1&0\\2&1\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}1\cdot a+0\cdot c&1\cdot b+0\cdot d\\2\cdot a+1\cdot c&2\cdot b+1\cdot d\end{pmatrix}=\begin{pmatrix}a&b\\c+2a&d+2b\end{pmatrix}$$

结果恰好是将 A 的第 1 行的 2 倍加到了第 2 行.

注 22.4
这是一个普遍规律.

知乎@余命数

- **行变换** $\Longleftrightarrow$ **左乘**初等矩阵.
- **列变换** $\Longleftrightarrow$ **右乘**初等矩阵.

这个性质将算法操作(消元)与代数结构(矩阵乘法)统一了起来.

22.6.7 阶梯形矩阵与秩的计算

通过反复使用初等行变换, 可将任何矩阵化简为一种标准形态, 从而一眼看出其秩的大小. 这种形态称为**行阶梯形** (Row Echelon Form, REF).

定义 22.17 (行阶梯形) 一个矩阵处于行阶梯形, 如果.

1. 所有的零行(元素全为 0 的行)都位于矩阵的底部.
2. 每一非零行的第一个非零元素, 称为**主元**, Pivot 位于上一行主元的右侧.

例如.

$$\begin{pmatrix} \mathbf{2} & 3 & 1 & 5 \\ 0 & \mathbf{1} & 4 & -2 \\ 0 & 0 & 0 & \mathbf{3} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

是一个行阶梯形矩阵. 其中加粗的元素 $\mathbf{2}, \mathbf{1}, \mathbf{3}$ 为主元.

定理 22.6 (行阶梯形矩阵的秩) 行阶梯形矩阵的秩, 等于其**非零行的行数**(也就是主元的个数).

证明 设行阶梯形矩阵有 r 个非零行, 它们的主元列号依次为

$$p_1 < p_2 < \cdots < p_r.$$

只需证明这 r 个非零行线性无关. 假设它们有线性组合为零, 得

$$c_1 R_1 + c_2 R_2 + \cdots + c_r R_r = 0.$$

若系数不全为零, 取最小的指标 s 使 $c_s \neq 0$. 观察第 p_s 列. 第 s 行在此列为非零主元;所有更靠下的非零行主元都在 p_s 右侧, 因而在第 p_s 列为 0;而更靠上的行对应系数已为 0. 所以该列的线性组合值为 c_s 乘以一个非零主元, 不可能等于 0, 矛盾.

因此非零行线性无关, 行秩为 r. 由秩定理, 矩阵秩也等于 r. □

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 @余命数

在上例中, 非零行有 3 行, 因此该矩阵的秩为 3.

例题 22.26 (求秩与化简) 求矩阵 $A=\begin{pmatrix}1&2&1\\2&4&0\\3&6&1\end{pmatrix}$ 的秩.

解 利用初等行变换消去主元下方的元素以化矩阵为阶梯形. 先以第 1 行为主元消去其下方元素, 做行变换 $R_2\leftarrow R_2-2R_1$, $R_3\leftarrow R_3-3R_1$:

$$\begin{pmatrix}1&2&1\\2&4&0\\3&6&1\end{pmatrix}\sim\begin{pmatrix}1&2&1\\0&0&-2\\0&0&-2\end{pmatrix}.$$

再以第 2 行为主元消去其下方元素, 做行变换 $R_3\leftarrow R_3-R_2$:

$$\sim\begin{pmatrix}1&2&1\\0&0&-2\\0&0&0\end{pmatrix}.$$

此时矩阵已达到阶梯形, 非零行有 2 行(第一行和第二行), 因此 $\operatorname{rank}(A)=2$.

这也意味着, 原矩阵的第 3 行可以由前两行线性表示(实际上 $R_3=R_1+R_2$), 它没有提供新的信息.

22.6.8 解线性方程组的理论结构

利用秩的概念, 我们可以完整地回答关于线性方程组 $Ax=b$ 的解的存在性与结构.

考虑方程组 $Ax=b$, 其中 A 为 $m\times n$ 矩阵. 构造**增广矩阵** (Augmented Matrix) $\bar{A}=(A|b)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

定理 22.7 (线性方程组解的判定定理) 1. **无解** $\iff r(A)<r(\bar{A})$.
(意味着向量 b 不在列向量张成的空间内, 引入 b 增加了维度.)

2. **有唯一解** $\iff r(A)=r(\bar{A})=n$, 其中 n 为未知数个数.
(意味着列向量线性无关, 且 b 在张成空间内, 坐标唯一.)

3. **有无穷多解** $\iff r(A)=r(\bar{A})<n$.
(意味着解存在, 但列向量之间存在冗余, 导致坐标表示不唯一. 自由度为 $n-r(A)$.)

证明 对增广矩阵 $\bar{A}=(A\mid b)$ 进行初等行变换, 化为行阶梯形矩阵. 初等行变换不改变方程组的解集, 也不改变 A 与 $\bar{A}$ 的秩.

若 $r(A)<r(\bar{A})$, 则在阶梯形增广矩阵中必出现一行. 系数部分全为 0, 但最后一列非零. 它对应矛盾方程

$$0=c\qquad (c\neq 0),$$

因而原方程组无解. 反之, 若出现这样的矛盾行, 最后一列增加了一个主元, 必有 $r(A)<r(\bar{A})$.

若 $r(A)=r(\bar{A})=r$, 则没有矛盾行, 方程组有解. 在阶梯形中, r 个主元对应 r 个基本未知量, 剩余 $n-r$ 个未知量为自由未知量. 当 $r=n$ 时没有自由未知量, 解唯一; 当 $r<n$ 时至少有一个自由未知量, 在实数域上它可取无穷多个值, 故有无穷多解. □

这个定理是线性代数的基础. 它表明. 求解方程组就是比较系数矩阵携带的信息量与包含常数项后所需的总信息量. 当秩小于未知数个数时, 多出的维度转化为了解空间的**自由度**, 解集构成了一个仿射子空间.

注 22.5
尽管克拉默法则给出了解的解析结构, 但在数值计算中, 对于高阶方程组, 例如 $n>3$, 计算 $n+1$ 个行列式的计算量远大于高斯消元法. 因此, 它主要用于理论推导及低阶矩阵的手算.

矩阵运算、行列式、线性变换、逆矩阵与克拉默法则, 以及秩与高斯消元构成了本章的内容. 几何向量语言见 第11章 与 第21章.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第 23 章

连分数

扩展与线性结构

「数学是打开科学大门的钥匙。」

——佚名

第二十三章 连分数

圆周率 π 的近似分数 $\dfrac{22}{7}$ 只用两个小整数, 误差却只有约 0.0013. 这样准的分数来自一个古老的算法.

辗转相除法求最大公因数时, 每一步都产生一个商; 把这些商按顺序排好, 就得到一个层层嵌套的分数, 即**连分数**. 它给每个实数配上一串最佳近似分数, $\dfrac{22}{7}$ 与 $\dfrac{355}{113}$ 之所以出色, 正因为它们是 π 的连分数渐近分数.

连分数把有理数与无理数放在同一框架下, 有理数对应有限连分数, 无理数对应无限连分数. 其中系数全部为 1 的一个就是黄金比, 它收敛最慢, 也最难逼近.

23.1 辗转相除法与连分数

?? 用辗转相除法求最大公因数, 即对 $a > b > 0$, 反复做带余除法 $a = bq + r$, 把 $\gcd(a, b)$ 化为 $\gcd(b, r)$, 直到余数为零. 每一步的商 q 构成一个序列, 这个序列正是连分数的全部信息.

23.1.1 从辗转相除到嵌套分数

以 $\dfrac{62}{23}$ 为例. 辗转相除给出

$$
\begin{aligned}
62 &= 2 \times 23 + 16,\\
23 &= 1 \times 16 + 7,\\
16 &= 2 \times 7 + 2,\\
7 &= 3 \times 2 + 1,\\
2 &= 2 \times 1.
\end{aligned}
$$

商依次为 $2, 1, 2, 3, 2$. 把每一步的余数回代, 就能把 $\dfrac{62}{23}$ 写成层层嵌套的形式. 由第一行

$$
\frac{62}{23} = 2 + \frac{16}{23}.
$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

把 $\dfrac{16}{23}$ 翻转:由第二行 $23=1\times16+7$ 得 $\dfrac{23}{16}=1+\dfrac{7}{16}$,故

$$\frac{16}{23}=\frac{1}{1+\dfrac{7}{16}},\qquad \frac{62}{23}=2+\frac{1}{1+\dfrac{7}{16}}.$$

继续由第三行 $16=2\times7+2$ 得 $\dfrac{16}{7}=2+\dfrac{2}{7}$,代入

$$\frac{62}{23}=2+\cfrac{1}{1+\cfrac{1}{2+\dfrac{2}{7}}}.$$

由第四行 $7=3\times2+1$ 得 $\dfrac{7}{2}=3+\dfrac{1}{2}$,代入

$$\frac{62}{23}=2+\cfrac{1}{1+\cfrac{1}{2+\cfrac{1}{3+\dfrac{1}{2}}}}.$$

最后一步 $2=2\times1$ 恰好除尽,嵌套到此终止. 这就是 $\dfrac{62}{23}$ 的**连分数**展开.

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

23.1.2 连分数的定义

定义 23.1 (简单连分数) 设 a_0 为整数, $a_1, a_2, \dots, a_n$ 为正整数. 称

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}} \tag{23-1}$$

为**简单连分数**, 简称**连分数**, 记作

$$[a_0;\, a_1,\, a_2,\, \dots,\, a_n].$$

a_0 称为**整数部分**, $a_1, \dots, a_n$ 称为**部分商**. 当 n 有限时称为**有限连分数**; 当序列无限延伸时称为**无限连分数**, 记 $[a_0;\, a_1,\, a_2,\, \dots]$.

"简单"二字指所有分子都是 1 且部分商为正整数. 若允许分子取其他值或部分商取实数, 则得到**广义连分数**, 本书只讨论简单连分数.

按此记号, $\dfrac{62}{23} = [2;\, 1,\, 2,\, 3,\, 2]$.

知乎@余命数 知乎@余命数

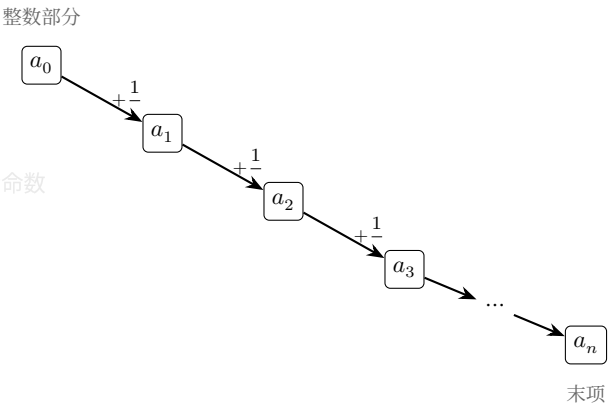


图 23.1 连分数 $[a_0;\, a_1, a_2, \dots, a_n]$ 的嵌套结构: 每一层都是整数加一的倒数.

例题 23.1 展开 $\dfrac{43}{30}$ 为连分数.

解 辗转相除:

$$43 = 1 \times 30 + 13, \quad 30 = 2 \times 13 + 4, \quad 13 = 3 \times 4 + 1, \quad 4 = 4 \times 1.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

商为 $1,2,3,4$, 故 $\dfrac{43}{30}=[1;\,2,\,3,\,4]$.

注 23.1
辗转相除法的每一步都把大数除以小数化为小数除以余数, 连分数的每一层都把整数加分数化为整数加一的倒数. 二者步骤一一对应, 第 k 个商 q_k 就是连分数的第 k 个部分商 a_k.

知乎@余命数

知乎@余命数

23.2 有限连分数与有理数

23.2.1 有理数与有限连分数的对应

第23.1节 的例子暗示了一个一般事实, 即每个有理数都能展开成有限连分数. 反过来, 任何有限连分数显然都等于一个有理数.

定理 23.1 (有理数的连分数表示) 每个有理数都有有限简单连分数表示. 约定末项 $a_n>1(n\geq 1)$ 时, 该表示唯一.

证明 存在性. 设 $x=p/q$、$q>0$、$\gcd(p,q)=1$. 对 p,q 做辗转相除法

$$p=a_0q+r_1,\quad q=a_1r_1+r_2,\quad r_1=a_2r_2+r_3,\quad \dots,\quad r_{n-2}=a_{n-1}r_{n-1}+r_n,$$

知乎@余命数

其中 $r_n=\gcd(p,q)=1$, $r_{n+1}=0$. 逐层回代即得 $x=[a_0;\,a_1,\,\dots,\,a_n]$.

唯一性. 末项为 1 时, $[a_0;\,\dots,\,a_{n-1},\,1]=[a_0;\,\dots,\,a_{n-1}+1]$, 故同一有理数有两种写法. 约定 $a_n>1$ 后, 表示唯一. □

知乎@余命数

例题 23.2 展开 $\dfrac{355}{113}$ 为连分数.

解 辗转相除:

$$355=3\times 113+16,$$

$$113=7\times 16+1,$$

$$16=16\times 1.$$

商为 $3,7,16$, 故

$$\frac{355}{113}=[3;\,7,\,16].$$

知乎@余命数

注 23.2
$\frac{355}{113}$ 是中国古代数学家祖冲之给出的圆周率**密率**, 误差小于 10^{-6}. 它的连分数 $[3;\,7,\,16]$ 只有三项, 却达到了惊人的精度, 原因见 第23.4节.

23.2.2 渐近分数

截取连分数的前若干项, 得到的分数称为**渐近分数**.

定义 23.2 (渐近分数) 设 $\alpha=[a_0;\,a_1,\,a_2,\,\ldots]$(有限或无限). 第 k 个**渐近分数**定义为

$$\frac{p_k}{q_k}=[a_0;\,a_1,\,\ldots,\,a_k]=a_0+\cfrac{1}{a_1+\cfrac{1}{\ddots+\cfrac{1}{a_k}}}. \tag{23-2}$$

渐近分数可用递推公式高效计算, 无需逐层展开嵌套.

性质 (渐近分数递推) 设 $p_{-2}=0,\,p_{-1}=1,\,q_{-2}=1,\,q_{-1}=0$. 对 $k\ge 0$,

$$p_k=a_kp_{k-1}+p_{k-2},\qquad q_k=a_kq_{k-1}+q_{k-2}. \tag{23-3}$$

证明 对 k 归纳.$k=0$: $p_0=a_0\cdot 1+0=a_0$, $q_0=a_0\cdot 0+1=1$, 即 $p_0/q_0=a_0$, 成立.$k=1$: $p_1=a_1a_0+1$, $q_1=a_1$, 即 $p_1/q_1=a_0+1/a_1$, 成立.

设 $k\le m$ 成立.$[a_0;\,\ldots,\,a_m,\,a_{m+1}]$ 等于把 $[a_0;\,\ldots,\,a_{m-1}]$ 中的 a_m 替换为 a_m+1/a_{m+1}. 由归纳假设, 分子分母都是 a_m 的线性函数, 替换后

$$p_{m+1}=\Big(a_m+\frac{1}{a_{m+1}}\Big)p_{m-1}+p_{m-2}=\frac{(a_ma_{m+1}+1)p_{m-1}+a_{m+1}p_{m-2}}{a_{m+1}}.$$

分子应等于 $a_{m+1}p_m+p_{m-1}$, 验证: $a_{m+1}(a_mp_{m-1}+p_{m-2})+p_{m-1}=(a_ma_{m+1}+1)p_{m-1}+a_{m+1}p_{m-2}$, 吻合. 分母同理. □

例题 23.3 计算 $\frac{62}{23}=[2;\,1,\,2,\,3,\,2]$ 的全部渐近分数.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 按式 (23-3) 递推, 列表如下.

k	a_k	p_k	q_k	p_k/q_k
-2		0	1	
-1		1	0	
0	2	2	1	2
1	1	3	1	3
2	2	8	3	$8/3$
3	3	27	10	$27/10$
4	2	62	23	$62/23$

渐近分数依次为 $2,\ 3,\ \dfrac{8}{3},\ \dfrac{27}{10},\ \dfrac{62}{23}$, 末项即原数.

23.2.3 渐近分数的交替逼近

定理 23.2 (交替逼近) 设 $\alpha=[a_0;\,a_1,\,\ldots]$ 的渐近分数为 p_k/q_k. 则

1. 渐近分数从两侧交替逼近 α: 偶数阶 $p_0/q_0,\,p_2/q_2,\,p_4/q_4,\,\ldots$ 单调递增, 奇数阶 $p_1/q_1,\,p_3/q_3,\,\ldots$ 单调递减, 且
$$\frac{p_0}{q_0}<\frac{p_2}{q_2}<\frac{p_4}{q_4}<\cdots<\alpha<\cdots<\frac{p_5}{q_5}<\frac{p_3}{q_3}<\frac{p_1}{q_1}.$$
2. 每个渐近分数比前一个更接近 α:
$$\left|\alpha-\frac{p_k}{q_k}\right|<\left|\alpha-\frac{p_{k-1}}{q_{k-1}}\right|.$$

相邻渐近分数之差有一个简洁公式, 它是上述定理的基石.

引理 23.1 $p_kq_{k-1}-p_{k-1}q_k=(-1)^{k-1}$, 故 $\dfrac{p_k}{q_k}-\dfrac{p_{k-1}}{q_{k-1}}=\dfrac{(-1)^{k-1}}{q_kq_{k-1}}$.

证明 对 k 归纳.$k=0$: $p_0q_{-1}-p_{-1}q_0=a_0\cdot0-1\cdot1=-1=(-1)^{-1}$, 成立. 设 k 成立, 由式 (23-3)

$$p_{k+1}q_k-p_kq_{k+1}=(a_{k+1}p_k+p_{k-1})q_k-p_k(a_{k+1}q_k+q_{k-1})=p_{k-1}q_k-p_kq_{k-1}=-(-1)^{k-1}=(-1)^k.$$

□

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

证明 (证明 定理 23.2) 由 引理 23.1， 相邻渐近分数之差 $p_k/q_k - p_{k-1}/q_{k-1} = \dfrac{(-1)^{k-1}}{q_k q_{k-1}}$, 符号交替, 绝对值随 q_k 递增而递减.k 为奇时差为正 ($p_k/q_k > p_{k-1}/q_{k-1}$), k 为偶时差为负 ($p_k/q_k < p_{k-1}/q_{k-1}$), 跳跃幅度越来越小, 故偶数阶递增、奇数阶递减, 二者从两侧夹逼 α, 误差单调递减. □

引理 23.1 还给出一个有用的推论：$\gcd(p_k, q_k) = 1$, 即渐近分数总是既约分数.

知乎@余命数

23.3 无限连分数与无理数

有理数的连分数是有限的. 对无理数, 辗转相除法永不终止 (余数永不为零), 连分数无限延伸.

知乎@余命数

定理 23.3 (无理数的连分数表示) 每个无理数有唯一的无限简单连分数表示; 反之, 每个无限简单连分数收敛于一个无理数.

证明思路与有理数情形平行, 对无理数 α, 取 $a_0 = \lfloor \alpha \rfloor$, 令 $\alpha_1 = 1/(\alpha - a_0) > 1$, 再取 $a_1 = \lfloor \alpha_1 \rfloor$, 如此反复. 因 α 无理, 每一步的 α_k 都无理, 过程永不终止, 得到无限序列 $[a_0;\, a_1,\, a_2,\, \ldots]$. 唯一性由每步取整的唯一性保证.

23.3.1 $\sqrt{2}$ 的连分数

知乎@余命数

定理 23.4 $\sqrt{2} = [1;\, \overline{2}]$, 即

$$\sqrt{2} = 1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \ddots}}}.$$

证明 设 $x = \sqrt{2} - 1 > 0$, 则 $\sqrt{2} = 1 + x$. 由 $(\sqrt{2}-1)(\sqrt{2}+1) = 1$ 得

$$x = \sqrt{2} - 1 = \frac{1}{\sqrt{2}+1} = \frac{1}{2 + (\sqrt{2}-1)} = \frac{1}{2+x}.$$

故 x 是 $2 + x$ 的倒数, 反复代入:

$$x = \frac{1}{2+x} = \frac{1}{2 + \dfrac{1}{2+x}} = \frac{1}{2 + \dfrac{1}{2 + \dfrac{1}{2+x}}} = \cdots$$

知乎@余命数

每展开一次, x 的位置就被 $\dfrac{1}{2+x}$ 替换, 产生一个新的系数 2. 系数为 2 的来历是 $\dfrac{1}{x}=\sqrt{2}+1$, 其整数部分 $\lfloor\sqrt{2}+1\rfloor=2$. 由 $\sqrt{2}=1+x$ 展开即得 $\sqrt{2}=[1;\,2,\,2,\,2,\,\ldots]$.

换用 $\alpha=\sqrt{2}$ 本身. $a_0=\lfloor\sqrt{2}\rfloor=1$, $\alpha-a_0=\sqrt{2}-1$, 取倒数 $\dfrac{1}{\sqrt{2}-1}=\sqrt{2}+1$, 故 $a_1=\lfloor\sqrt{2}+1\rfloor=2$, 余下 $\sqrt{2}+1-2=\sqrt{2}-1$, 取倒数又回到 $\sqrt{2}+1$. 过程循环, $a_1=a_2=\cdots=2$. □

注 23.3
$\sqrt{2}$ 的无理性已在 第16.1节 中用反证法证明. 连分数给出了另一个视角: $\sqrt{2}$ 的连分数 $[1;\,\overline{2}]$ 是无限的, 故 $\sqrt{2}$ 无理. 有限连分数必为有理数, 无限连分数必为无理数, 二者泾渭分明.

23.3.2 $\sqrt{3}$ 的连分数

例题 23.4 求 $\sqrt{3}$ 的连分数.

解 $a_0=\lfloor\sqrt{3}\rfloor=1$, $\sqrt{3}-1$ 取倒数得 $\dfrac{1}{\sqrt{3}-1}=\dfrac{\sqrt{3}+1}{2}\approx 1.366$, 故 $a_1=1$. 余下 $\dfrac{\sqrt{3}+1}{2}-1=\dfrac{\sqrt{3}-1}{2}$, 取倒数得 $\dfrac{2}{\sqrt{3}-1}=\sqrt{3}+1\approx 2.732$, 故 $a_2=2$. 余下 $\sqrt{3}+1-2=\sqrt{3}-1$, 回到起点. 故

$$\sqrt{3}=[1;\,\overline{1,\,2}].$$

23.3.3 周期连分数与二次无理数

$\sqrt{2}$ 和 $\sqrt{3}$ 的连分数都是**周期**的, 从某一项起, 部分商序列不断重复. 上划线标记一个循环节, 如 $[1;\,\overline{2}]$ 表示 $[1;\,2,\,2,\,2,\,\ldots]$.

定义 23.3 (周期连分数) 若存在 $m\ge 0$ 与 $T\ge 1$ 使 $a_{k+T}=a_k$ 对所有 $k>m$ 成立, 则称 $[a_0;\,a_1,\,\ldots]$ 为**周期连分数**, 记作 $[a_0;\,\ldots,\,a_m,\,\overline{a_{m+1},\,\ldots,\,a_{m+T}}]$. 从 a_0 起就周期的称为**纯周期连分数**.

定理 23.5 (拉格朗日定理) 实数 α 的连分数是周期的, 当且仅当 α 是二次无理数, 即 α 是整系数二次方程 $ax^2+bx+c=0(a\neq 0)$ 的无理根.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

注 23.4
拉格朗日定理把连分数的周期性与二次方程的无理根这两个看似无关的性质等同起来：$\sqrt{2}, \sqrt{3}, \sqrt{5}, \ldots$ 都是二次无理数，连分数都周期；π 和 e 不是任何整系数多项式的根（超越数），连分数不周期.

例题 23.5 几个常见二次无理数的连分数：

$$\sqrt{2} = [1; \overline{2}], \quad \sqrt{3} = [1; \overline{1, 2}], \quad \sqrt{5} = [2; \overline{4}], \quad \sqrt{7} = [2; \overline{1, 1, 1, 4}].$$

练习 23.1 用本节的方法求 $\sqrt{6}$ 的连分数，并验证它是周期的.

知乎@余命数

23.4 最佳有理逼近

23.4.1 渐近分数的最优性

第23.2节 的渐近分数不仅是越来越好的近似，而且在分母不超过某值的所有分数中是最优的.

定理 23.6 (最佳逼近) 设 p_k/q_k 是 α 的第 k 个渐近分数. 若 $0 < q' \le q_k$ 且 $p'/q' \ne p_k/q_k$，则

$$\left|\alpha - \frac{p_k}{q_k}\right| < \left|\alpha - \frac{p'}{q'}\right|.$$

即 p_k/q_k 是分母不超过 q_k 的分数中，最接近 α 的那一个.

知乎@余命数

证明 (证明大意) 关键不等式：对任意 $k \ge 0$，

$$\frac{1}{q_k(q_{k+1} + q_k)} < \left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k q_{k+1}}. \tag{23-4}$$

左端说明误差不会太小，右端给出上界. 设 p'/q' 满足 $0 < q' \le q_k$ 且 $|p'/q' - p_k/q_k| < |\alpha - p_k/q_k|$. 由 引理 23.1，$|p_k/q_k - p_{k-1}/q_{k-1}| = 1/(q_k q_{k-1})$. 可以证明 p'/q' 必落在 p_k/q_k 与 p_{k-1}/q_{k-1} 之间，但该区间内分母不超过 q_k 的既约分数只有 p_k/q_k 本身，矛盾. □

注 23.5
定理 23.6 的精确含义是，渐近分数即**最佳有理逼近**. 给定分母上界 q_k，找不到比 p_k/q_k 更接近 α 的分数.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

23.4.2 圆周率的渐近分数

例题 23.6　圆周率 π 的连分数前几项为

$$\pi = [3;\, 7,\, 15,\, 1,\, 292,\, \ldots].$$

计算前四个渐近分数.

解　按式 (23-3) 递推:

k	a_k	p_k	q_k	p_k/q_k
0	3	3	1	3
1	7	22	7	$22/7$
2	15	333	106	$333/106$
3	1	355	113	$355/113$

故 π 的前四个渐近分数为

$$3,\quad \frac{22}{7},\quad \frac{333}{106},\quad \frac{355}{113}.$$

表 23.1　π 的渐近分数及其误差.

渐近分数	小数值	误差	有效位数
3	3.0	0.14	0
$22/7$	$3.1428\ldots$	1.3×10^{-3}	2
$333/106$	$3.14150\ldots$	8.3×10^{-5}	4
$355/113$	$3.1415929\ldots$	2.7×10^{-7}	6

表 23.1 回答了章首的问题.$\dfrac{22}{7}$ 之所以准, 因为它是 π 的第一个非平凡渐近分数;$\dfrac{355}{113}$ 之所以更准,因为它是第三个. 二者都是**最佳逼近**, 在分母不超过 7(113) 的所有分数中, 没有比 $\dfrac{22}{7}(\dfrac{355}{113})$ 更接近 π 的.

注 23.6
π 的第四个部分商 $a_4 = 292$ 异常大. 由式 (23-4), $q_{k+1} = a_{k+1}q_k + q_{k-1}$, 大的 a_{k+1} 使 q_{k+1} 暴增, 误差上界 $1/(q_kq_{k+1})$ 骤降. 这解释了 $\dfrac{355}{113}$ 为何精度惊人, 它后面紧跟一个 292, 把下一个渐近分数的分母推到 33102, 误差骤降两个数量级.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

23.4.3 与牛顿迭代法的对照

第7.11节 介绍的牛顿迭代法用切线逼近曲线的根, 每步迭代使有效位数大致翻倍, 收敛极快. 连分数渐近分数则用整数之比逼近实数, 每步增加一个部分商, 分母按递推增长.

两种方法服务于不同目的. 牛顿迭代给出**实数**近似值, 即小数, 精度随迭代指数增长, 但结果不是精确的分数. 连分数渐近分数给出**有理数**近似值, 即分数, 精度由部分商决定, 结果是简洁的 p/q. 当需要用最简单的分数表示某个实数时, 连分数无可替代; 当需要高精度小数时, 牛顿迭代更高效.

例题 23.7 用连分数解释: 为何 $\dfrac{355}{113}$ 比 $\dfrac{22}{7}$ 好得多, 却只多用了两个整数?

解 $\dfrac{22}{7}=[3;\,7]$ 是 π 的第一个渐近分数, $\dfrac{355}{113}=[3;\,7,\,15,\,1]$ 是第三个. 从 $[3;\,7]$ 到 $[3;\,7,\,15,\,1]$, 分母从 7 增长到 113, 误差从 10^{-3} 降到 10^{-7}. 增长来自部分商 15 和 1: $q_2=15\times 7+1=106$, $q_3=1\times 106+7=113$. 仅两个额外整数, 分母扩大 16 倍, 误差缩小近 5000 倍.

知乎@余命数

23.5 黄金比

23.5.1 黄金比的连分数

第7.7节 在求斐波那契数列的通项时, 遇到了特征方程 $x^2-x-1=0$, 其正根

$$\varphi=\frac{1+\sqrt{5}}{2}\approx 1.618$$

称为**黄金比**. 这个数在连分数理论中地位特殊, 它的连分数是所有无限连分数中最简单的.

定理 23.7 $\varphi=[1;\,\overline{1}]$, 即

$$\varphi=1+\cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{1+\ddots}}}.$$

证明 由 $\varphi^2=\varphi+1$(特征方程), 两边除以 φ 得

$$\varphi=1+\frac{1}{\varphi}.$$

右端的 φ 再次用 $1+1/\varphi$ 替换：

$$\varphi = 1 + \cfrac{1}{1+\cfrac{1}{\varphi}} = 1 + \cfrac{1}{1+\cfrac{1}{1+\cfrac{1}{\varphi}}} = \cdots$$

每替换一次，产生一个新的系数 1，过程永不终止. 故 $\varphi = [1;\,1,\,1,\,1,\,\ldots] = [1;\,\overline{1}]$. □

知乎@余命数

23.5.2 斐波那契比与渐近分数

φ 的渐近分数有一个优美的性质，它们恰是斐波那契数列的相邻项之比.

定理 23.8 设 F_n 为斐波那契数列 ($F_0 = 0,\, F_1 = 1,\, F_n = F_{n-1} + F_{n-2}$). 则 $\varphi = [1;\,\overline{1}]$ 的第 k 个渐近分数为

$$\frac{p_k}{q_k} = \frac{F_{k+2}}{F_{k+1}} \quad (k \geq 0).$$

证明 φ 的所有部分商 $a_k = 1$. 由 式 (23-3), $p_k = p_{k-1} + p_{k-2},\, q_k = q_{k-1} + q_{k-2}$, 二者都满足斐波那契递推. 初值 $p_0 = 1 = F_2,\, q_0 = 1 = F_1;\, p_1 = 2 = F_3,\, q_1 = 1 = F_2$. 由归纳法 $p_k = F_{k+2},\, q_k = F_{k+1}$. □

例题 23.8 列出 φ 的前几个渐近分数.

解

$$\frac{p_0}{q_0} = \frac{1}{1},\quad \frac{p_1}{q_1} = \frac{2}{1},\quad \frac{p_2}{q_2} = \frac{3}{2},\quad \frac{p_3}{q_3} = \frac{5}{3},\quad \frac{p_4}{q_4} = \frac{8}{5},\quad \frac{p_5}{q_5} = \frac{13}{8},\quad \ldots$$

分子分母都是斐波那契数：$\frac{F_2}{F_1},\, \frac{F_3}{F_2},\, \frac{F_4}{F_3},\, \frac{F_5}{F_4},\, \frac{F_6}{F_5},\, \frac{F_7}{F_6},\, \ldots$

第7.7节 用比内公式证明了 $\lim_{n\to\infty} F_{n+1}/F_n = \varphi$. 连分数给出了同一结论的另一条路径：$\varphi$ 的渐近分数恰为 F_{n+1}/F_n，而渐近分数必收敛于连分数本身，故 $F_{n+1}/F_n \to \varphi$.

知乎@余命数

23.5.3 最无理的数

定理 23.9 (最慢收敛) 在所有无理数中，φ 的连分数渐近分数收敛最慢.

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

收敛速度由 式 (23-4) 控制, 误差 $|\alpha - p_k/q_k| < \dfrac{1}{q_k q_{k+1}}$, 而 $q_{k+1} = a_{k+1}q_k + q_{k-1}$. 部分商 a_{k+1} 越大, q_{k+1} 越大, 误差越小, 收敛越快.

φ 的所有部分商都是 1, 这是正整数中最小的可能值. 这使 $q_{k+1} = q_k + q_{k-1}$ 增长最慢 (恰为斐波那契增长), 误差上界最大. 任何其他无理数至少有一个部分商大于 1, 在该处收敛加速.

知乎@余命数

注 23.7

φ 被称为"最无理的数", 正因为它的连分数最难逼近. 用分数逼近 φ 的效率最低, 要达到 n 位精度, 所需分母比逼近其他任何无理数都大. 这一性质使 φ 在自然界中频繁出现, 例如向日葵种子的螺旋排列、松果鳞片的布局都遵循黄金角 ($2\pi/\varphi^2$ 弧度), 正因为黄金比最难有理逼近, 螺旋之间不会周期性重合, 排列最均匀.

23.5.4 Conway 的士兵问题 *(选读)

黄金比与斐波那契数常在意想不到的地方出现. 下面这个问题由 Conway 提出, 解决它要用一列以 $\dfrac{1}{\varphi}$ 为公比的几何级数.

例题 23.9 (士兵问题) 在坐标系里, 满足 $y \le 0$ 的每个整点上都放一枚棋子. 允许的操作是水平或竖直的跳跃, 一枚棋子越过紧邻的另一枚棋子, 落到它后面紧接着的空格上, 被越过的棋子随即拿掉. 只做有限多次操作, 能否让一枚棋子到达点 $C(0,5)$?

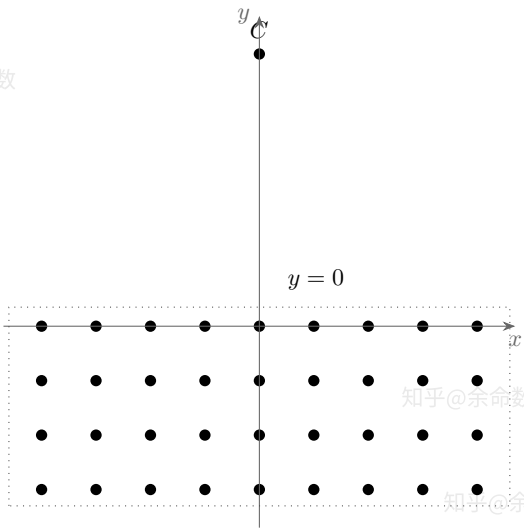


图 23.2 初始局面为下半平面的整点满格, 目标是让棋子到达 $C(0,5)$. 每枚棋子的权是 z^d, 其中 $z = 1/\varphi$, d 为它到 C 的曼哈顿距离.

解 把棋子换成一份可以求和的账. 记

$$z=\frac{\sqrt{5}-1}{2}=\frac{1}{\varphi},$$

它满足

$$z^2+z=1.$$

对平面上任一整点 P, 记 $d(P)=|x_P|+|y_P-5|$, 即 P 到 C 的曼哈顿距离, 并给 P 处的棋子记上权 $z^{d(P)}$. 某个局面下所有棋子权的和称为这个局面的 **Conway 和**.

先算初始局面的和.$y=0$ 这一行到 C 的距离依次为 $|x|+5$, 该行贡献

$$z^5+2\left(z^6+z^7+\cdots\right)=z^5+\frac{2z^6}{1-z}=z^5+2z^4,$$

其中用到了 第16.4节 的无穷几何级数求和公式. 再由 $z^2+z=1$ 化简,

$$z^5+2z^4=z^3\left(z^2+z\right)+z^4=z^3+z^4=z^2\left(z+z^2\right)=z^2.$$

同理, $y=-k$ 这一行贡献 z^{k+2}. 全盘的初始和为

$$z^2+z^3+z^4+\cdots=\frac{z^2}{1-z}=\frac{z^2}{z^2}=1,$$

最后一步又用了 $1-z=z^2$.

再看每次跳跃如何改变这份和. 设被越过的棋子到 C 的距离为 d.

- 朝着 C 跳时, 跳跃者从距离 $d+1$ 移到 $d-1$, 变化量为

$$z^{d-1}-z^d-z^{d+1}=z^{d-1}\left(1-z-z^2\right)=0;$$

- 背着 C 跳时, 变化量为

$$z^{d+2}-z^d-z^{d+1}=z^d\left(z^2-1-z\right)=-2z^{d+1}<0;$$

- 跳跃者到 C 的距离不变时, 它的权不变, 被越过的棋子连同它的权一并消失, 变化量为负.

三种情形合起来说明, 每一次跳跃都不会让 Conway 和变大, 它自初始值 1 起单调不增.

假设有一枚棋子到达了 C. 该点到 C 的距离为 0, 权为 $z^0=1$, 这一个 1 已经用尽全部可用的份额. 又因为只做了有限多次操作, 盘上仍有无限多枚棋子, 它们的权都是正数, 于是局面的 Conway 和严格大于 1, 与它始终不超过 1 矛盾.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

故有限次跳跃无法把棋子送到 $C(0,5)$, 同样的论证对任何 $C(0,n)$ 都成立. 这份账本以黄金比的倒数为公比, 它满足 $z^2+z=1$, 朝着目标的三项 z^{d-1}, z^d, z^{d+1} 才会恰好抵消.

练习 23.2 验证 φ 的前五个渐近分数满足 定理 23.2 的交替逼近, 即偶数阶递增, 奇数阶递减, 二者从两侧夹逼 φ.

知乎@余命数

23.6 逼近的阶

23.6.1 误差该与什么比较

比较两个近似分数的好坏, 只看误差本身有失公允, $\dfrac{355}{113}$ 的误差是 $\dfrac{3}{1}$ 的几十万分之一, 但它的分母也大了一百多倍. 公允的尺度是**误差乘以分母的平方**

$$q^2\left|\alpha-\frac{p}{q}\right|.$$

知乎@余命数

分母每扩大一格, 误差按比例缩减一格, 才算真正的改善. 式 (23-4) 的右端给出, 渐近分数恒满足

$$\left|\alpha-\frac{p_k}{q_k}\right|<\frac{1}{q_k^2},\qquad \text{即}\qquad q_k^2\left|\alpha-\frac{p_k}{q_k}\right|<1.$$

知乎@余命数 知乎@余命数

这个界还能不能收紧, 收紧的下限在哪里, 接着我们就来回答这两个问题, 最后用渐近分数设计历法里的闰年规则.

知乎@余命数

23.6.2 相邻渐近分数, 二者必有其一

定理 23.10 (两取一定理) 设 α 无理, p_k/q_k 与 p_{k-1}/q_{k-1} 是相邻两个渐近分数. 则至少有一个满足

$$\left|\alpha-\frac{p}{q}\right|<\frac{1}{2q^2}.$$

证明 α 夹在相邻渐近分数之间 (定理 23.2), 故两个误差之和恰等于两渐近分数之差:

$$\left|\alpha-\frac{p_k}{q_k}\right|+\left|\alpha-\frac{p_{k-1}}{q_{k-1}}\right|=\left|\frac{p_k}{q_k}-\frac{p_{k-1}}{q_{k-1}}\right|=\frac{1}{q_kq_{k-1}}.$$

知乎@余命数 知乎@余命数

由 $2ab \le a^2+b^2$（第5.3节 的均值不等式）且 $q_k \ne q_{k-1}$，

$$\frac{1}{q_k q_{k-1}} < \frac{1}{2q_k^2} + \frac{1}{2q_{k-1}^2}.$$

两个误差之和小于两个阈值之和，二者之中至少有一个小于各自的阈值. □

反过来，误差足够小的分数必然是渐近分数.

定理 23.11 (Legendre 判据) 设既约分数 $\frac{a}{b}$ 满足 $\left|\alpha - \frac{a}{b}\right| < \frac{1}{2b^2}$. 则 $\frac{a}{b}$ 是 α 的某个渐近分数.

证明 只需证 $\frac{a}{b}$ 是最佳逼近，再由 定理 23.6 的结构即得结论. 设既约分数 $\frac{c}{d} \ne \frac{a}{b}$ 比 $\frac{a}{b}$ 更接近 α 且 $0 < d \le b$. 一方面，两个不同分数之差的绝对值至少为 $\frac{1}{bd}$（分子 $|ad-bc|$ 是非零整数）；另一方面，

$$\left|\frac{c}{d} - \frac{a}{b}\right| \le \left|\frac{c}{d} - \alpha\right| + \left|\alpha - \frac{a}{b}\right| < \left|\alpha - \frac{a}{b}\right| + \left|\alpha - \frac{a}{b}\right| < \frac{1}{b^2}.$$

两式联立得 $\frac{1}{bd} < \frac{1}{b^2}$，即 $d > b$，与 $d \le b$ 矛盾. 故分母不超过 b 的分数中 $\frac{a}{b}$ 最接近 α，它是渐近分数. □

23.6.3 Hurwitz 定理与 $\sqrt{5}$ 的临界性

定理 23.12 (Hurwitz, 1891) 对任意无理数 α，相邻三个渐近分数 p_k/q_k、p_{k-1}/q_{k-1}、p_{k-2}/q_{k-2} 中至少有一个满足

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{\sqrt{5}\,q^2}.$$

特别地，每个无理数都有无穷多个分数 $\frac{p}{q}$ 满足此式.

证明的骨架与 定理 23.10 同类，把三个误差通过完全商联系起来，若三者都不达标，可推出某个部分商 $a_k < 1$，与其为正整数矛盾. 界中的常数 $\sqrt{5}$ 来自何处，下面的取等例子给出答案.

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

定理 23.13 (黄金比取等) 设 $\varphi = \dfrac{1+\sqrt{5}}{2} = [1;\,\overline{1}]$, 其渐近分数为 $\dfrac{F_{k+1}}{F_k}$. 则

$$\lim_{k\to\infty} F_k^2 \left|\varphi - \frac{F_{k+1}}{F_k}\right| = \frac{1}{\sqrt{5}}.$$

证明 φ 的每个完全商都是 φ 自身, 把 φ 按 $[1;\,1,\,1,\,\ldots]$ 展开到任意层, 剩下的尾段仍是同一个无限连分数. 于是

$$\varphi = \frac{p_k\varphi + p_{k-1}}{q_k\varphi + q_{k-1}}, \qquad \left|\varphi - \frac{p_k}{q_k}\right| = \frac{1}{q_k(q_k\varphi + q_{k-1})} = \frac{1}{q_k^2\left(\varphi + \dfrac{q_{k-1}}{q_k}\right)}.$$

由 定理 23.8, $q_k = F_k$; 由 第23.5节, $\dfrac{q_{k-1}}{q_k} = \dfrac{F_{k-1}}{F_k} \to \dfrac{1}{\varphi}$. 因此

$$F_k^2\left|\varphi - \frac{F_{k+1}}{F_k}\right| \longrightarrow \frac{1}{\varphi + \dfrac{1}{\varphi}} = \frac{1}{\sqrt{5}},$$

最后一步用了 $\varphi + \dfrac{1}{\varphi} = \dfrac{1+\sqrt{5}}{2} + \dfrac{\sqrt{5}-1}{2} = \sqrt{5}$. □

定理 23.13 给出 Hurwitz 定理中常数 $\dfrac{1}{\sqrt{5}}$ 的临界性, 即把右端收紧为 $\dfrac{c}{q^2}(c < \dfrac{1}{\sqrt{5}})$, 则当 k 充分大后 $\varphi + \dfrac{q_{k-1}}{q_k} < \dfrac{1}{c}$, 渐近分数的误差全部越界; 再由 定理 23.11 ($c < \dfrac{1}{2}$ 时满足不等式的分数必为渐近分数), 其他分数也帮不上忙. 于是对黄金比而言, 更强的界只剩有限多组解. 黄金比正是逼近意义下最难有理逼近的数, 第23.5节 的说法在这里成为定理.

23.6.4 闰年规则的连分数解释

回归年长约 365.2421897 天, 小数部分 0.2421897 决定闰年的频率. 它的连分数展开为

$$0.2421897 = [0;\,4,\,7,\,1,\,3,\,27,\,\ldots],$$

渐近分数依次为

$$\frac{1}{4},\quad \frac{7}{29},\quad \frac{8}{33},\quad \frac{31}{128},\quad \cdots$$

每个渐近分数给出一条闰年方案, 误差如下.

知乎@余命数

表 23.2　闰年方案与误差.

方案	闰年频率	每年误差(天)	折合
儒略历:每 4 年 1 闰	$1/4 = 0.25$	$+0.0078103$	约 11 分钟
格里高利历:每 400 年 97 闰	$97/400 = 0.2425$	$+0.0003103$	约 27 秒
渐近分数 $31/128$:每 128 年 31 闰	$31/128 = 0.2421875$	-0.0000022	约 0.2 秒

格里高利历的 $\dfrac{97}{400}$ 位于渐近分数 $\dfrac{8}{33}$ 与 $\dfrac{31}{128}$ 之间, 误差比二者都大. 连分数直接给出的 $\dfrac{31}{128}$ 分母更小、误差缩小一百多倍, 每 128 年置 31 个闰年, 一万年才偏离约半天. 历法不必苛求到此精度, 但若问"分母不超过四百的分数里谁最接近真实的闰年频率", 答案由 定理 23.6 唯一给出, 就是渐近分数 $\dfrac{31}{128}$.

23.6.5 代数数逼近的下界与刘维尔数 *(选读)

Hurwitz 定理说, 任何无理数都能被分数以 $\dfrac{1}{\sqrt{5}\,q^2}$ 的精度逼近. 对代数数, 逼近精度**远好于** $\dfrac{1}{q^2}$ 是不可能的, 例如达到 $\dfrac{1}{q^3}$ 或 $\dfrac{1}{q^{10}}$; Liouville 在 1844 年证明了这一点.

定义 23.4 (代数数与超越数) 若实数 α 是某个整系数多项式方程

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0 \quad (a_n \neq 0)$$

的根, 则称 α 为**代数数**; 满足的多项式次数的最小值称为 α 的**次数**. 非代数数的实数称为**超越数**.

有理数 $\dfrac{a}{b}$ 是一次代数数(方程 $bx - a = 0$ 的根); $\sqrt{2}$ 是二次代数数; 定理 23.5 里出现过的二次无理数都是二次代数数.

定理 23.14 (Liouville 逼近定理) 设 α 是 n 次 ($n \geq 2$) 代数无理数. 则存在常数 $C > 0$, 使得对一切既约分数 $\dfrac{p}{q}$,

$$\left|\alpha - \frac{p}{q}\right| > \frac{C}{q^n}.$$

知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数　知乎@余命数

证明 取整系数多项式 $f(x)$, 使 α 是 f 的根, 且 α 不是任何更低次整系数多项式的根. 先说明 α 是 f 的单根. 若 $f(x)$ 含因子 $(x-\alpha)^2$, 求导后 $f'(x)$ 仍含因子 $x-\alpha$, α 便是次数更低的多项式 $f'(x)$ 的根, 与取法矛盾. 于是可写

$$f(x)=(x-\alpha)h(x),$$

其中 h 是 $n-1$ 次多项式且 $h(\alpha)=f'(\alpha)\neq 0$. 由连续性, 存在 $\delta>0$ 与 $M>0$, 使 $|h(x)|\leq M$ 在 $|\alpha-x|\leq\delta$ 内成立.

对任意的既约分数 p/q 分两种情形. 若 $|\alpha-p/q|>\delta$, 结论显然, 此时取 $C<\delta$ 即可. 若 $|\alpha-p/q|\leq\delta$, 由分解式

$$\left|\alpha-\frac{p}{q}\right|=\frac{|f(p/q)|}{|h(p/q)|}\geq\frac{1}{M}\left|f\left(\frac{p}{q}\right)\right|.$$

另一方面, $q^n f(p/q)=a_n p^n+a_{n-1}p^{n-1}q+\cdots+a_1 pq^{n-1}+a_0 q^n$ 是整数, 且不为零 (否则 p/q 是 f 的有理根, 与 α 无理且次数最低矛盾), 故 $|f(p/q)|\geq\dfrac{1}{q^n}$. 取 $C=\min\{\delta,\dfrac{1}{M}\}$ 即得

$$\left|\alpha-\frac{p}{q}\right|>\frac{C}{q^n}.$$

□

二次无理数因此不可能有优于 $\dfrac{C}{q^2}$ 的无穷逼近序列, Hurwitz 的 $\dfrac{1}{\sqrt{5}\,q^2}$ 恰好贴着这条下界运行. 更高次的代数数下界更松, 但依然存在. 沿此思路, 一个能被**任意好**地逼近的数必然是超越数. 这样的数现在就能造出来.

定理 23.15 (刘维尔数) 级数

$$L=\sum_{k=1}^{\infty}\frac{1}{10^{k!}}=0.110001000000000000000001\ldots$$

是超越数.

证明 (证明梗概) 取前 n 项的和 $\dfrac{p_n}{q_n}(q_n=10^{n!})$, 余项

$$0<L-\frac{p_n}{q_n}<2\cdot\frac{1}{10^{(n+1)!}}.$$

对任意固定的正整数 N, 当 $n>N$ 时 $(n+1)!-N\cdot n!=n!(n+1-N)\geq n!$, 故

$$q_n^N\left(L-\frac{p_n}{q_n}\right)<2\cdot 10^{N\cdot n!-(n+1)!}\leq\frac{2}{10^{n!}}\to 0.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

即对任何 N 与任何 $C>0$, 总有分数使 $\left|L-\dfrac{p}{q}\right|<\dfrac{C}{q^N}$. 若 L 是 N 次代数数, 与 定理 23.14 矛盾; 故 L 是超越数. □

L 的小数展开在 $k!$ 位上放 1、其余位放 0, 间隙增长极快, 每一步截断都给出精度惊人的近似分数, 快到任何代数数都消受不起. 这是人类构造出的第一批超越数. 自然出现的常数随后跟上, Hermite 在 1873 年证明 e 是超越数, Lindemann 在 1882 年证明 π 是超越数, 后者一并宣判了化圆为方问题的死刑.

✎ **练习 23.3**　利用 定理 23.11 证明: 若 $\left|\alpha-\dfrac{a}{b}\right|<\dfrac{1}{2b^2}$ 对无穷多组 (a,b) 成立, 则 α 是无理数.

知乎@余命数

知乎@余命数
知乎@余命数
@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

第 24 章

常数 e、π 与 γ

扩展与线性结构

「数学是打开科学大门的钥匙。」

——佚名

第二十四章 常数 e、π 与 γ

数学里最重要的三个常数, 恰好来自三个不同的源头.

$e \approx 2.71828\ldots$ 是自然对数的底. 它的出场方式很朴素: 把年利率 100% 的利息按一年 n 次复利计算, 每次利率 $1/n$, 年末倍数是 $(1+1/n)^n$; 让 n 趋于无穷, 倍数趋于 e. 放射性衰变、种群增长、连续贴现, 凡是瞬时连续变化的模型, 底数都落在 e 上.

$\pi \approx 3.14159\ldots$ 来自几何, 即圆周与直径之比. 但它渗透到分析与数论的深处: 一个纯几何常数, 却出现在概率、级数与复变函数的公式里. 本章给出 Niven 的初等证明, 确认 π 是无理数.

$\gamma \approx 0.5772\ldots$ 是欧拉常数, 刻画调和级数 $H_n = 1 + 1/2 + \cdots + 1/n$ 与 $\ln n$ 之间的差距. 调和级数发散, $\ln n$ 也发散, 二者之差却收敛到一个有限值. 这个值是否为无理数, 至今没有答案.

三个常数的共同点是: 它们都由极限或级数定义, 性质的研究离不开 第7章 的级数工具与 第17章 的积分工具. 本章用这些工具给出 e 与 π 的无理性证明, 建立 γ 的基本性质, 并以巴塞尔问题 $\sum 1/n^2 = \pi^2/6$ 收尾, 这个结果把级数、函数与常数编织在一起.

24.1 自然常数 e

24.1.1 e 的定义

第16.4节 已经给出指数函数的幂级数展开 $e^x = \sum_{n=0}^{\infty} x^n/n!$, 并指出 $e \approx 2.71828\ldots$. 接着我们就从极限定义出发, 把这个常数正式建立起来.

定义 24.1 (自然常数 e)

$$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

这个极限的存在性将在 第24.1.3小节 用二项式定理直接证明, 无需预先接受. 先考察它的直观来源.

复利模型. 设年利率为 100%. 若一年计息一次, 年末倍数为 $1 + 1 = 2$. 若半年计息一次, 每次利率 $1/2$, 年末倍数为 $(1 + 1/2)^2 = 2.25$. 若一年计息 n 次, 每次利率 $1/n$, 年末倍数为

$$\left(1 + \frac{1}{n}\right)^n.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

计息越频繁, 年末倍数越大, 但增长有顶, 它趋于 e. 把离散的计息间隔压缩到连续, 就得到自然常数.

n	$(1+1/n)^n$	误差
1	2.00000	0.71828
10	2.59374	0.12454
100	2.70481	0.01347
1000	2.71692	0.00136
$\to\infty$	$\to e\approx 2.71828$	$\to 0$

表 24.1　复利倍数 $(1+1/n)^n$ 随 n 增大趋于 e.

24.1.2 复利史话

利息问题是最古老的数学应用之一. 现藏于卢浮宫的一块巴比伦泥板 (约公元前 1700 年)问过一个干脆的问题, 年利率 20%、按年复利, 本金何时翻倍?

例题 24.1 (巴比伦的翻倍问题) 按年复利, 年利率 20%, 本金翻倍需要多少年? 巴比伦人用线性插值得到 3.7870 年, 与真值差多少?

解 设需要 x 年, 则 $1.2^x=2$, 取对数 (第3.9节) 得

$$x=\frac{\ln 2}{\ln 1.2}=\frac{0.6931}{0.1823}\approx 3.8018.$$

巴比伦人没有对数, 他们注意到 $1.2^3=1.728$, $1.2^4=2.0736$, 真值夹在 3 与 4 之间, 便按比例内插:

$$x\approx 3+\frac{2-1.728}{2.0736-1.728}=3+\frac{0.272}{0.3456}\approx 3.7870,$$

与真值 3.8018 只差 0.015 年, 约五天半. 泥板上以六十进制记为 $3;47,13,20$, 即 $3+\dfrac{47}{60}+\dfrac{13}{60^2}+\dfrac{20}{60^3}$.

真值约 3.8018 年, 线性插值 3.7870 年, 误差约 0.4%.

复利的“复”字在于利息也生利息, 计息越频繁, 同样的年利率滚出越多的钱, 但这条增长有顶. 以本金 100 元、年利率 5% 为例, 各档计息周期一年后的本利和如下.

从一年一次到一日一次, 收益只多出一毛三分. 把计息间隔压缩到零, 收益逼近一个确定的数.1683 年, 雅各布·伯努利在研究复利问题时证明了这一点.$(1+1/n)^n$ 随 n 增大而增大, 却始终被 3 压住. 增长与压制的平衡点就是 e.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

计息周期	一年末本利和(元)
一年一次	105.00
半年一次	105.06
一季一次	105.09
一月一次	105.12
一周一次	105.12
一日一次	105.13
连续复利	$100\,e^{0.05}=105.13$

表 24.2 同样 5% 的年利率, 计息越频繁收益越高, 但增量迅速枯竭, 上限是 $100e^{0.05}$ 元.

24.1.3 极限的存在性与级数表示

伯努利的论断可以补全成一条定理, 证明只需要二项式定理 (第9.3节) 和单调有界原理 (第16.3节). 记 $s_n=\left(1+\dfrac{1}{n}\right)^n$, 按二项式定理展开, 第 k 项可以整理成

$$s_n=\sum_{k=0}^{n}\frac{1}{k!}\prod_{j=0}^{k-1}\left(1-\frac{j}{n}\right), \tag{24-1}$$

其中 $k=0$ 与 $k=1$ 两项都是 1, 从 $k=2$ 起每个因子 $1-j/n$ 都落在 0 与 1 之间.

先看单调性. 固定 k, 当 n 变大时每个因子 $1-j/n$ 都不减, 第 k 项便随 n 增大而增大, 项数还从 $n+1$ 个变成 $n+2$ 个, 多出来的又是一个正项. 逐项比较 s_{n+1} 与 s_n 的展开式, 得 $s_{n+1}>s_n$, 数列严格递增.

再看上界. 每个因子不超过 1, 故第 k 项不超过 $\dfrac{1}{k!}$; 而 $k\ge 1$ 时 $k!=1\cdot 2\cdots k\ge 2^{k-1}$, 于是

$$s_n\le\sum_{k=0}^{n}\frac{1}{k!}\le 1+\sum_{k=1}^{n}\frac{1}{2^{k-1}}<3.$$

递增而有上界的数列必有极限, e 的存在性就此落定, 顺带也得到了伯努利的判断, e 夹在 2 与 3 之间.

这番展开还白送一件礼物. 在式 (24-1) 中固定 $m\le n$ 只看前 m 项, 令 $n\to\infty$, 每个因子 $1-j/n\to 1$, 第 k 项趋于 $\dfrac{1}{k!}$, 于是 $\lim s_n\ge\sum_{k=0}^{m}\dfrac{1}{k!}$; 再令 $m\to\infty$, 得 $\lim s_n\ge\sum_{k=0}^{\infty}\dfrac{1}{k!}$(右端的和存在, 因为 $1/k!\le 1/2^{k-1}$, 与几何级数同敛散). 反向的不等式刚刚已经算过, 就是 $s_n\le\sum_{k=0}^{n}1/k!\le\sum_{k=0}^{\infty}1/k!$. 两头夹住, 极限只能是级数的和.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

定理 24.1 (e 的级数表示)

$$e=\sum_{n=0}^{\infty}\frac{1}{n!}=1+1+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\cdots$$

这个等式是 e^x 的泰勒展开 (第16.4节) 在 $x=1$ 的特例. 阶乘 $n!$ 增长极快, 通项 $1/n!$ 迅速趋于零, 级数收敛速度远超极限定义.

知乎@余命数

n	$1/n!$	部分和 S_n
0	1.000000	1.000000
1	1.000000	2.000000
2	0.500000	2.500000
3	0.166667	2.666667
4	0.041667	2.708333
5	0.008333	2.716667
6	0.001389	2.718056
7	0.000198	2.718254
8	0.000025	2.718279

表 24.3　e 的级数部分和, 仅取前 9 项即精确到五位小数.

24.1.4 e 的无理性

知乎@余命数
知乎@余命数

级数表示计算方便, 同时直接给出 e 是无理数的证明.

知乎@余命数

定理 24.2 (e 是无理数)　e 是无理数.

证明　反设 $e=p/q$, 其中 p,q 为正整数. 取整数 $n\geq q$. 由 定理 24.1,

$$e=\sum_{k=0}^{n}\frac{1}{k!}+\sum_{k=n+1}^{\infty}\frac{1}{k!}.$$

两端乘以 $n!$:

$$n!\cdot e=\underbrace{\sum_{k=0}^{n}\frac{n!}{k!}}_{\text{整数}}+\underbrace{\sum_{k=n+1}^{\infty}\frac{n!}{k!}}_{\text{余项}}.$$

第一部分是整数, 因为当 $k\leq n$ 时 $n!/k!$ 是整数.

知乎@余命数
知乎@余命数

估计余项. 对 $k \geq n+1$,

$$\frac{n!}{k!} = \frac{1}{(n+1)(n+2)\cdots k} \leq \frac{1}{(n+1)^{k-n}}.$$

因此

$$0 < \sum_{k=n+1}^{\infty} \frac{n!}{k!} \leq \sum_{k=n+1}^{\infty} \frac{1}{(n+1)^{k-n}} = \sum_{j=1}^{\infty} \frac{1}{(n+1)^j} = \frac{1}{n}.$$

知乎@余命数

最后一个等号用了几何级数求和 $\sum_{j=1}^{\infty} r^j = r/(1-r)$, 此处 $r = 1/(n+1)$. 当 $n \geq 1$ 时 $1/n \leq 1$, 故余项严格介于 0 与 1 之间.

另一方面, 若 $e = p/q$ 且 $n \geq q$, 则 $n! \cdot e = n! \cdot p/q$ 是整数(因 $q \mid n!$). 于是

$$\underbrace{n! \cdot e}_{\text{整数}} = \underbrace{\sum_{k=0}^{n} \frac{n!}{k!}}_{\text{整数}} + \underbrace{\sum_{k=n+1}^{\infty} \frac{n!}{k!}}_{\in(0,1)}.$$

整数等于整数加上一个严格介于 0 与 1 之间的数, 矛盾. 故 e 是无理数. □

证明的机制是, 级数 $\sum 1/k!$ 的前 n 项乘以 $n!$ 后是整数, 而余项乘以 $n!$ 后被几何级数控制在 $(0,1)$ 之内. 有理数假设 $e = p/q$ 迫使 $n! \cdot e$ 为整数, 与余项的严格小性冲突. 这种整数等于整数加严格小数的矛盾, 是无理性证明的经典模式.

注 24.1 (e 的超越性)
e 的无理性之上, 还有一个更强的结论. 第23.6节 的 Liouville 定理说, 代数数(整系数多项式的根)不可能被有理数逼近得太狠, 而 Hermite 在 1873 年证明 e 不属于代数数, 即 e 是超越数. 同一年代前后, Lindemann 用类似的方法证明 π 也是超越数, 一并宣判了化圆为方问题的死刑.

注 24.2 (e 的连分数)
Euler 在 1737 年算出

$$e = [2;\, 1,\, 2,\, 1,\, 1,\, 4,\, 1,\, 1,\, 6,\, 1,\, 1,\, 8,\, \ldots],$$

除去开头的 2, 部分商里每隔两个 1 站着一个递增的偶数 $2, 4, 6, 8, \ldots$, 规律之整齐, 在无理常数中独一份. 相应的渐近分数 (第23.1节) 依次是 $\frac{8}{3}, \frac{11}{4}, \frac{19}{7}, \frac{87}{32}, \frac{193}{71}, \frac{1264}{465}, \ldots$, 其中 $\frac{193}{71} \approx 2.71831$ 已准到小数点后四位.

练习 24.1 证明 $\left(1 - \frac{1}{n}\right)^n \to \frac{1}{e}$. 借此说明 第10.2节 错排问题的极限 $\frac{D_n}{n!} \to \frac{1}{e}$, 信封全装错的概率恰好是 e 的倒数.

知乎@余命数

24.2 π 的无理性

24.2.1 Niven 的证明

e 的无理性证明依赖级数，π 的无理性证明则依赖积分.1947 年，Ivan Niven 给出了一个只用多项式、积分与 sin 函数的初等证明，全部工具都在本书范围内（第17章 的分部积分，第6章 的三角函数）.

定理 24.3 (π 是无理数) π 是无理数.

知乎@余命数

证明 反设 $\pi = a/b$，其中 a, b 为正整数. 对任意正整数 n，定义多项式

$$f_n(x) = \frac{x^n(a-bx)^n}{n!}.$$

知乎@余命数

考虑积分

$$I_n = \int_0^\pi f_n(x)\sin x\,dx.$$

先看 I_n 的符号. 在区间 $(0,\pi)$ 上，$x>0$，$\sin x>0$. 又因 $\pi = a/b$，故 $a-bx = b(\pi-x)>0$ 当 $0<x<\pi$. 因此 $f_n(x)>0$ 在 $(0,\pi)$ 上成立，从而 $I_n>0$.

再看 I_n 的整数性. 对 $f_n(x)\sin x$ 反复使用分部积分 (第17.4节). 注意 $f_n(x)$ 是 $x^n(a-bx)^n/n!$，展开后是关于 x 的 $2n$ 次多项式，系数形如 $c_k x^k/n!$，其中 c_k 是整数.

知乎@余命数

分部积分公式 $\int u\,dv = uv - \int v\,du$ 中，取 $u = f_n(x)$，$dv = \sin x\,dx$，则 $v = -\cos x$：

$$I_n = [-f_n(x)\cos x]_0^\pi + \int_0^\pi f_n'(x)\cos x\,dx.$$

知乎@余命数

边界项：$f_n(0) = 0$（因 x^n 因子），$f_n(\pi) = \pi^n(a-b\pi)^n/n! = \pi^n\cdot 0^n/n! = 0$，因 $a-b\pi = 0$. 故边界项为零，得

$$I_n = \int_0^\pi f_n'(x)\cos x\,dx.$$

再分部，取 $u = f_n'(x)$、$dv = \cos x\,dx$、$v = \sin x$：

$$I_n = [f_n'(x)\sin x]_0^\pi - \int_0^\pi f_n''(x)\sin x\,dx.$$

边界项：$\sin 0 = \sin\pi = 0$，故边界项为零，得

$$I_n = -\int_0^\pi f_n''(x)\sin x\,dx.$$

知乎@余命数

每两次分部积分, sin 与 cos 互换, f_n 的导数升高两阶, 符号交替变化. 重复 n 轮共 $2n$ 次分部积分后, f_n 的 $2n$ 阶导数是常数, 因 f_n 是 $2n$ 次多项式, 更高阶导数为零. 最终

$$I_n = \pm \int_0^{\pi} f_n^{(2n)}(x) \cdot (\sin \text{ 或} \cos)\, dx + \text{边界项}.$$

每一步的边界项都形如 $f_n^{(k)}(0) \cdot (\sin \text{ 或} \cos)(0)$ 或 $f_n^{(k)}(\pi) \cdot (\sin \text{ 或} \cos)(\pi)$. 由于 $\sin 0 = \sin \pi = 0$、$\cos 0 = \pm 1$、$\cos \pi = \pm 1$, 边界项只涉及 $f_n^{(k)}(0)$ 与 $f_n^{(k)}(\pi)$.

现在计算这些导数值.$f_n(x) = x^n(a-bx)^n/n!$. 对 $k < n$, $f_n^{(k)}(0) = 0$(因 x^n 因子使 0 处低于 n 阶的导数为零). 对 $n \le k \le 2n$, $f_n^{(k)}(0)$ 是 $x^n(a-bx)^n$ 在 $x=0$ 处的 k 阶导数除以 $n!$, 展开后每一项系数都是整数, 除以 $n!$ 后仍为整数, 因 $k \ge n$ 时组合数 $\binom{k}{n}$ 等因子吸收了 $n!$. 同理, $f_n^{(k)}(\pi) = f_n^{(k)}(a/b)$, 代入 $a - bx = b(\pi - x)$, 由 $\pi = a/b$ 的结构, 这些值也是整数.

因此所有边界项都是整数, I_n 是整数. 结合 $I_n > 0$, I_n 是**正整数**.

最后估计 I_n 的大小, 导出矛盾. 在 $[0,\pi]$ 上, $x \le \pi$ 且 $a - bx = b(\pi - x) \le b\pi = a$, 故

$$0 \le f_n(x) = \frac{x^n(a-bx)^n}{n!} \le \frac{(\pi \cdot a)^n}{n!} = \frac{(a\pi)^n}{n!}.$$

又 $\sin x \le 1$, 故

$$0 < I_n \le \pi \cdot \frac{(a\pi)^n}{n!}.$$

阶乘 $n!$ 的增长速度远超指数 $(a\pi)^n$(比值判别: $(a\pi)^{n+1}/(n+1)! \div (a\pi)^n/n! = a\pi/(n+1) \to 0$), 故 $(a\pi)^n/n! \to 0$, 从而 $I_n \to 0$.

但 I_n 是正整数, 正整数序列不可能趋于零. 矛盾. 故 π 是无理数. □

注 24.3 (证明的结构)

Niven 证明的精妙之处在于构造了一个恰到好处的多项式 $f_n(x) = x^n(a-bx)^n/n!$:

- x^n 与 $(a-bx)^n$ 保证 f_n 在 0 与 π 处有足够高的零点, 使分部积分的边界项可控;
- $1/n!$ 保证各阶导数在端点处取整数值;
- $2n$ 次多项式的结构使反复分部积分最终终止.

这条矛盾与 第24.1节 中 e 的无理性证明同属一个模式, 两者都造出一个既必须是整数、又被限制在 $(0,1)$ 内的量.

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

24.3 欧拉常数 γ

24.3.1 γ 的定义

第16.4节 证明了调和级数发散:

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \to \infty.$$

对数函数 $\ln n$ 也随 $n \to \infty$ 发散. 二者的发散速度接近, 差距 $H_n - \ln n$ 却趋于一个有限值.

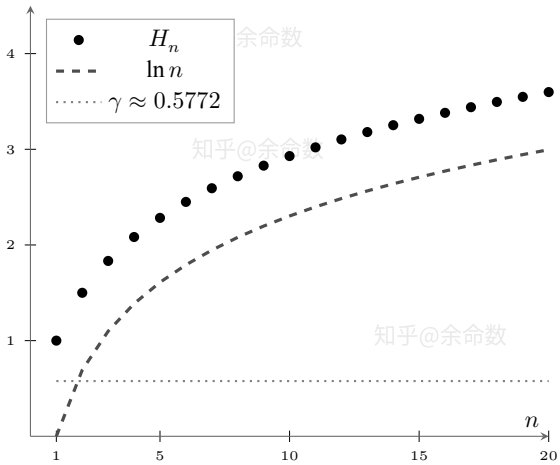


图 24.1　调和级数部分和 H_n 与 $\ln n$ 的差趋于欧拉常数 γ.

定理 24.4 (欧拉常数 γ 的存在性) 极限 $\lim_{n\to\infty}(H_n - \ln n)$ 存在且有限.

证明 记 $a_n = H_n - \ln n$. 用积分估计 $1/k$ 与 $\ln(k+1) - \ln k$ 的差. 对 $k \geq 1$,

$$\ln(k+1) - \ln k = \int_k^{k+1} \frac{1}{x}\, dx.$$

在 $[k, k+1]$ 上, $1/x$ 单调递减, 故 $1/(k+1) \leq 1/x \leq 1/k$, 从而

$$\frac{1}{k+1} \leq \ln(k+1) - \ln k \leq \frac{1}{k}.$$

因此

$$0 \leq \frac{1}{k} - \big(\ln(k+1) - \ln k\big) \leq \frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

对 k 从 1 到 $n-1$ 求和：

$$0 \le H_{n-1} - \ln n = \sum_{k=1}^{n-1} \left[\frac{1}{k} - (\ln(k+1) - \ln k) \right] \le \sum_{k=1}^{n-1} \frac{1}{k(k+1)}.$$

右端是裂项级数，$\sum_{k=1}^{n-1} 1/[k(k+1)] = 1 - 1/n < 1$. 故 $H_{n-1} - \ln n$ 有界. 又每一项 $1/k - (\ln(k+1) - \ln k) \ge 0$，故 $H_{n-1} - \ln n$ 单调递增有上界，极限存在.

最后，$a_n = H_n - \ln n = (H_{n-1} - \ln n) + 1/n$，而 $1/n \to 0$，故 a_n 与 $H_{n-1} - \ln n$ 同极限. □

知乎@余命数

定义 24.2 (欧拉常数)

$$\gamma := \lim_{n\to\infty} (H_n - \ln n) \approx 0.5772\ldots$$

γ 的存在性证明只用到积分比较与单调有界原理，两者都在本书范围内.$1/k$ 是 $[k, k+1]$ 上 $1/x$ 的上矩形，$\ln(k+1) - \ln k$ 是精确积分，二者之差正且可被 $1/[k(k+1)]$ 控制，裂项求和后有界.

余命数

24.3.2 γ 的性质

知乎@余命数

定理 24.5 (γ 的积分表示)

$$\gamma = \int_0^1 \left(\frac{1}{t} - \frac{1}{-\ln(1-t)} \right) dt$$

这个积分表示的推导需要变量代换与极限交换，属于 第17章 积分技巧的延伸，此处作为结论引用. 它把 γ 从离散的 $H_n - \ln n$ 转化为一个连续的积分，在渐近分析中频繁出现.

定理 24.6 (γ 与 ζ 函数的关系)

$$\gamma = \sum_{n=2}^{\infty} \frac{(-1)^n \zeta(n)}{n}$$

其中 $\zeta(s) = \sum_{k=1}^{\infty} 1/k^s$ 是黎曼 ζ 函数.

ζ 函数在 $s > 1$ 时收敛，$\zeta(2) = \pi^2/6$ 是 第24.4节 巴塞尔问题的结论.γ 与 ζ 的这一联系，把调和级数、对数与 π 串联在同一条链上，为 第25章 的无穷乘积与 ζ 函数的系统讨论埋下伏笔.

知乎关注@余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

注 24.4 (γ 的无理性)
e 与 π 的无理性都已证明 (第24.1节, 第24.2节), γ 是否为无理数, 至今没有答案. 数学家已经证明, 若 γ 是有理数 p/q, 则 q 必须超过 10^{242080}, 一个天文数字. 这强烈暗示 γ 是无理数, 但严格证明尚未找到.γ 是数学中最基本的常数之一, 其无理性却仍是开放问题.

24.4 巴塞尔问题

24.4.1 问题与历史

第16.4节 指出 p-级数 $\sum 1/n^p$ 在 $p>1$ 时收敛, 并提到 $p=2$ 时和为 $\pi^2/6$. 这个结论的历史可以追溯到 1644 年 Pietro Mengoli 提出的问题.

问题 24.1 巴塞尔问题 求无穷级数的精确值:

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=1+\frac{1}{4}+\frac{1}{9}+\frac{1}{16}+\cdots=\ ?$$

这个级数收敛, 因为通项 $1/n^2$ 比 $1/n$ 衰减得快得多, 精确值却迟迟无人算出.1735 年, 28 岁的 Euler 宣布了答案.

$$\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}.$$

一个纯几何常数 π 出现在一个纯级数的和里, 这个结果震惊了当时的数学界. 它把离散的求和 $\sum 1/n^2$、连续的函数 $\sin x$ 与几何常数 π 绑在了一起.

24.4.2 Euler 的推导思路

Euler 的原始推导基于一个大胆的假设, 把多项式的根与因式分解关系推广到无穷级数, 即“无穷多项式”. 严格证明需要复分析与无穷乘积理论, 属于 第25章 的范畴. 这里给出 Euler 的推导, 展示其惊人的洞察力.

先写出 $\sin x/x$ 的零点. 函数 $\sin x$ 的零点为 $x=0,\pm\pi,\pm2\pi,\pm3\pi,\ldots$. 因此 $\sin x/x$ 的零点（去掉 $x=0$ 这个可去奇点）为 $\pm\pi,\pm2\pi,\pm3\pi,\ldots$, 即 $x=\pm n\pi$, $n=1,2,3,\ldots$.

再看因式分解的类比. 一个 d 次多项式 $P(x)$ 若有根 $r_1,r_2,\ldots,r_d$ 且 $P(0)=1$, 则可分解为

$$P(x)=\prod_{k=1}^{d}\left(1-\frac{x}{r_k}\right).$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

Euler 假设这个关系对“无穷多项式”$\sin x/x$ 同样成立, 把 $\pm n\pi$ 配对后得到

$$\frac{\sin x}{x} = \prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2\pi^2}\right) = \left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{4\pi^2}\right)\left(1-\frac{x^2}{9\pi^2}\right)\cdots \tag{24-2}$$

这就是 $\sin x$ 的无穷乘积展开. 它的严格证明由 Weierstrass 在近一个世纪后完成.

最后比较系数. 另一方面, $\sin x/x$ 的泰勒展开 (第16.4节) 为

$$\frac{\sin x}{x} = 1-\frac{x^2}{3!}+\frac{x^4}{5!}-\frac{x^6}{7!}+\cdots = 1-\frac{x^2}{6}+\frac{x^4}{120}-\cdots \tag{24-3}$$

把 式 (24-2) 展开, 收集 x^2 项. 每个因子贡献 $-x^2/(n^2\pi^2)$, 交叉项是 x^4 及更高次, 故 x^2 系数为

$$-\sum_{n=1}^{\infty}\frac{1}{n^2\pi^2} = -\frac{1}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

与 式 (24-3) 中 x^2 系数 $-1/6$ 比较:

$$-\frac{1}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2} = -\frac{1}{6}.$$

解出

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}. \tag{24-4}$$

注 24.5 (Euler 推导的地位)
Euler 的推导在 1735 年并不严格, 无穷乘积 式 (24-2) 的合法性需要复分析工具才能证明, 把多项式因式分解推广到无穷这一步在当时没有依据. 但这个推导揭示了 $\sin x$、级数 $\sum 1/n^2$ 与 π 之间的深层联系, 直接催生了无穷乘积理论与复分析的发展. 第25章 将系统讨论无穷乘积.
比较系数的方法还可以继续, 比较 x^4 系数得 $\sum 1/n^4 = \pi^4/90$, 比较 x^6 系数得 $\sum 1/n^6 = \pi^6/945$. 一般地, Euler 得到了偶数 ζ 值的公式

$$\zeta(2k) = \sum_{n=1}^{\infty}\frac{1}{n^{2k}} = (-1)^{k+1}\frac{B_{2k}(2\pi)^{2k}}{2\cdot(2k)!},$$

其中 B_{2k} 是伯努利数. 奇数 ζ 值 $\zeta(3),\zeta(5),\ldots$ 是否为无理数, 至今大部分仍是未解之谜, 目前 $\zeta(3)$ 的无理性已由 Apéry 于 1978 年证明, $\zeta(5)$ 及更高奇数 ζ 值的无理性仍属开放问题.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

第 25 章

无穷乘积

扩展与线性结构

「数学是打开科学大门的钥匙。」

——佚名

第二十五章 无穷乘积

级数研究无穷次加法, 无穷乘积研究无穷次乘法. 二者通过取对数联系起来, 对数把乘积变成加法, 级数的收敛理论就能搬过来用.

第16.4节 已经建立了级数的收敛框架, 本章把它推广到乘积. 三个经典乘积将依次登场: Wallis 乘积把 π 写成无穷多个有理分数相乘, Vieta 公式用嵌套根号给出 π 的乘积表示, 欧拉乘积则把素数分布与 ζ 函数编织进同一个等式.

知乎@余命数

25.1 无穷乘积的收敛性

25.1.1 部分积与收敛

定义 25.1 (无穷乘积) 给定数列 $\{a_n\}$, 称表达式

$$\prod_{n=1}^{\infty} a_n = a_1 \cdot a_2 \cdot a_3 \cdots$$

为**无穷乘积**. 构造**部分积数列** $\{P_N\}$:

$$P_N = \prod_{n=1}^{N} a_n = a_1 a_2 \cdots a_N.$$

若 $\lim_{N\to\infty} P_N$ 存在且等于非零实数 P, 则称该无穷乘积**收敛**于 P; 若极限不存在或等于零, 则称其**发散**.

收敛值要求非零, 这与级数不同. 若 $P_N \to 0$, 乘积提供的信息太少, 任何一项趋于零都会让整个乘积趋于零, 这种情形归入发散.

25.1.2 收敛的必要条件是通项趋于 1

级数收敛要求通项趋于 0; 乘积收敛要求通项趋于 1.

定理 25.1 (通项趋于 1) 若 $\prod_{n=1}^{\infty} a_n$ 收敛, 则 $\lim_{n\to\infty} a_n = 1$.

证明 设部分积 $P_N \to P \neq 0$. 当 $n \geq 2$ 时, $a_n = P_n / P_{n-1}$. 由极限的除法法则,

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{P_n}{P_{n-1}} = \frac{P}{P} = 1.$$

□

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

通项趋于 1 只是必要条件, 并不充分. 判断充分条件需要借助对数.

25.1.3 取对数把乘积化为级数

取对数把乘法变成加法. 当 $a_n>0$ 时,

$$\ln\left(\prod_{n=1}^{N} a_n\right)=\sum_{n=1}^{N}\ln a_n.$$

知乎@余命数

部分积 P_N 的对数恰是对数项的部分和. 由此得到核心判据.

定理 25.2 (对数判据) 设 $a_n>0$ 对所有 n 成立. 无穷乘积 $\prod_{n=1}^{\infty}a_n$ 收敛, 当且仅当级数 $\sum_{n=1}^{\infty}\ln a_n$ 收敛. 此时

$$\prod_{n=1}^{\infty} a_n=\exp\left(\sum_{n=1}^{\infty}\ln a_n\right).$$

证明 记 $S_N=\sum_{n=1}^{N}\ln a_n$, 则 $P_N=e^{S_N}$. 指数函数连续且严格单调, 故 P_N 收敛于非零极限 P 当且仅当 S_N 收敛于 $\ln P$. □

正项级数的全部判别法, 包括比较判别法、比值判别法、积分判别法, 现在都能用于判断无穷乘积的收敛性.

知乎@余命数

25.1.4 两个基本例子

知乎@余命数

例题 25.1 判断 $\prod_{n=1}^{\infty}\left(1+\dfrac{1}{n}\right)$ 的敛散性.

解 对数级数为

$$\sum_{n=1}^{\infty}\ln\left(1+\frac{1}{n}\right)=\sum_{n=1}^{\infty}\ln\frac{n+1}{n}.$$

部分和裂项:

$$\sum_{n=1}^{N}\ln\frac{n+1}{n}=\ln\frac{2}{1}+\ln\frac{3}{2}+\cdots+\ln\frac{N+1}{N}=\ln(N+1).$$

当 $N\to\infty$ 时 $\ln(N+1)\to\infty$, 对数级数发散, 故乘积发散.

例题 25.2 计算 $\prod_{n=2}^{\infty}\left(1-\dfrac{1}{n^2}\right)$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

解 对每个因子做因式分解:

$$1-\frac{1}{n^2}=\frac{n^2-1}{n^2}=\frac{(n-1)(n+1)}{n\cdot n}=\frac{n-1}{n}\cdot\frac{n+1}{n}.$$

部分积为

$$P_N=\prod_{n=2}^{N}\frac{n-1}{n}\cdot\prod_{n=2}^{N}\frac{n+1}{n}.$$

第一个乘积裂项:

$$\prod_{n=2}^{N}\frac{n-1}{n}=\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{3}{4}\cdots\frac{N-1}{N}=\frac{1}{N}.$$

第二个乘积裂项:

$$\prod_{n=2}^{N}\frac{n+1}{n}=\frac{3}{2}\cdot\frac{4}{3}\cdot\frac{5}{4}\cdots\frac{N+1}{N}=\frac{N+1}{2}.$$

两式相乘:

$$P_N=\frac{1}{N}\cdot\frac{N+1}{2}=\frac{N+1}{2N}.$$

取极限:

$$\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right)=\lim_{N\to\infty}\frac{N+1}{2N}=\frac{1}{2}.$$

例题 25.1 中通项 $1+1/n\to1$ 满足必要条件, 乘积仍发散; 例题 25.2 中通项 $1-1/n^2\to1$ 同样满足必要条件, 乘积却收敛于 $1/2$. 区别在于 $\ln(1+1/n)\sim1/n$, 对应发散的调和级数, 而 $\ln(1-1/n^2)\sim-1/n^2$, 对应 $p=2$ 的收敛 p-级数. 对数判据把乘积的敛散性归结为级数的敛散性, 第16.4节 的全部工具在此生效.

25.2 Wallis 乘积

25.2.1 从积分递推到乘积

第17.4.2小节 建立了 $\sin^n x$ 的积分递推, 第17.6节 据此得到华里士公式. 现在把同一递推推向极限, 得到 π 的第一个无穷乘积表示.

记

$$I_n=\int_0^{\pi/2}\sin^n x\,dx.$$

知乎@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由分部积分递推 (第17.4.2小节), 当 $n \geq 2$ 时

$$I_n = \frac{n-1}{n} I_{n-2}. \tag{25-1}$$

起点为 $I_0 = \pi/2$, $I_1 = 1$. 反复应用递推:

$n = 2m$ 为偶数时, 递推至 I_0:

$$I_{2m} = \frac{2m-1}{2m} \cdot \frac{2m-3}{2m-2} \cdots \frac{1}{2} \cdot I_0 = \frac{(2m-1)!!}{(2m)!!} \cdot \frac{\pi}{2}. \tag{25-2}$$

$n = 2m+1$ 为奇数时, 递推至 I_1:

$$I_{2m+1} = \frac{2m}{2m+1} \cdot \frac{2m-2}{2m-1} \cdots \frac{2}{3} \cdot I_1 = \frac{(2m)!!}{(2m+1)!!}. \tag{25-3}$$

25.2.2 夹逼与极限

在 $[0, \pi/2]$ 上 $0 \leq \sin x \leq 1$, 故 $\sin^{2m+1} x \leq \sin^{2m} x \leq \sin^{2m-1} x$, 积分保号:

$$I_{2m+1} \leq I_{2m} \leq I_{2m-1}. \tag{25-4}$$

将 (25-2) 与 (25-3) 代入 (25-4):

$$\frac{(2m)!!}{(2m+1)!!} \leq \frac{(2m-1)!!}{(2m)!!} \cdot \frac{\pi}{2} \leq \frac{(2m-2)!!}{(2m-1)!!}.$$

两端同除以 $\dfrac{(2m-1)!!}{(2m)!!}$:

$$\frac{1}{2m+1}\left[\frac{(2m)!!}{(2m-1)!!}\right]^2 \leq \frac{\pi}{2} \leq \frac{1}{2m}\left[\frac{(2m)!!}{(2m-1)!!}\right]^2. \tag{25-5}$$

左右两端的比值:

$$\frac{\dfrac{1}{2m}\left[\dfrac{(2m)!!}{(2m-1)!!}\right]^2}{\dfrac{1}{2m+1}\left[\dfrac{(2m)!!}{(2m-1)!!}\right]^2} = \frac{2m+1}{2m} \to 1 \quad (m \to \infty).$$

两端夹逼同一值, 故

$$\frac{\pi}{2} = \lim_{m\to\infty} \frac{1}{2m+1}\left[\frac{(2m)!!}{(2m-1)!!}\right]^2. \tag{25-6}$$

知乎@余命数 知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

25.2.3 Wallis 乘积公式

把 (25-6) 中的双阶乘展开.$(2m)!! = 2\cdot 4\cdot 6\cdots 2m$, $(2m-1)!! = 1\cdot 3\cdot 5\cdots(2m-1)$, 故

$$\left[\frac{(2m)!!}{(2m-1)!!}\right]^2 = \frac{2^2\cdot 4^2\cdots(2m)^2}{1^2\cdot 3^2\cdots(2m-1)^2} = \prod_{n=1}^{m}\frac{(2n)^2}{(2n-1)(2n+1)}\cdot(2m+1).$$

代入 (25-6) 后 $2m+1$ 约去, 得到:

定理 25.3 (Wallis 乘积)

$$\frac{\pi}{2} = \prod_{n=1}^{\infty}\frac{(2n)(2n)}{(2n-1)(2n+1)} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots \tag{25-7}$$

每个因子 $\dfrac{(2n)^2}{(2n-1)(2n+1)} = \dfrac{4n^2}{4n^2-1} = 1+\dfrac{1}{4n^2-1} > 1$, 通项趋于 1, 满足第25.1节 的必要条件. 对数 $\ln\left(1+\dfrac{1}{4n^2-1}\right)\sim\dfrac{1}{4n^2}$, 对应级数收敛, 乘积确实收敛.

Wallis 乘积是 π 的第一个无穷乘积表示(1655 年). 它把超越数 π 表为无穷多个有理数的乘积, 每个因子都简单到只需整数四则运算. 下一节将看到, Vieta 公式给出了更早的、用嵌套根号表示 π 的乘积.

25.3 Vieta 公式

25.3.1 倍角公式的逆向使用

半角公式 $\sin x = 2\sin\dfrac{x}{2}\cos\dfrac{x}{2}$ 把一个正弦拆成两个. 对 $\sin\dfrac{x}{2}$ 再拆, 对 $\sin\dfrac{x}{4}$ 再拆, 如此反复, $\sin x$ 就展开为一系列余弦的乘积.

取 $x=\pi/2$, 此时 $\sin(\pi/2)=1$. 反复套用半角公式:

$$\begin{aligned}
1 = \sin\frac{\pi}{2} &= 2\sin\frac{\pi}{4}\cos\frac{\pi}{4}\\
&= 2\left(2\sin\frac{\pi}{8}\cos\frac{\pi}{8}\right)\cos\frac{\pi}{4} = 2^2\sin\frac{\pi}{8}\cos\frac{\pi}{4}\cos\frac{\pi}{8}\\
&= 2^3\sin\frac{\pi}{16}\cos\frac{\pi}{4}\cos\frac{\pi}{8}\cos\frac{\pi}{16}\\
&\ \ \vdots\\
&= 2^n\sin\frac{\pi}{2^{n+1}}\prod_{k=1}^{n}\cos\frac{\pi}{2^{k+1}}.
\end{aligned}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

移项得

$$\frac{1}{2^n \sin(\pi/2^{n+1})} = \prod_{k=1}^{n} \cos\frac{\pi}{2^{k+1}}. \tag{25-8}$$

25.3.2 取极限

当 $n \to \infty$ 时, $\pi/2^{n+1} \to 0$. 由 $\lim_{\theta\to 0} \dfrac{\sin\theta}{\theta} = 1$,

$$2^n \sin\frac{\pi}{2^{n+1}} = \frac{\pi}{2} \cdot \frac{\sin(\pi/2^{n+1})}{\pi/2^{n+1}} \to \frac{\pi}{2}.$$

故 (25-8) 左端趋于 $\dfrac{2}{\pi}$, 右端趋于无穷乘积 $\prod_{k=1}^{\infty} \cos\dfrac{\pi}{2^{k+1}}$.

定理 25.4 (Vieta 公式)

$$\frac{2}{\pi} = \prod_{n=1}^{\infty} \cos\frac{\pi}{2^{n+1}} = \cos\frac{\pi}{4} \cdot \cos\frac{\pi}{8} \cdot \cos\frac{\pi}{16} \cdots \tag{25-9}$$

25.3.3 嵌套根号形式

每个余弦因子可用半角公式 $\cos\dfrac{\theta}{2} = \sqrt{\dfrac{1+\cos\theta}{2}}$ 化为根号嵌套. 由 $\cos\dfrac{\pi}{4} = \dfrac{\sqrt{2}}{2}$ 出发:

$$\begin{aligned}
\cos\frac{\pi}{4} &= \frac{\sqrt{2}}{2},\\
\cos\frac{\pi}{8} &= \sqrt{\frac{1+\cos(\pi/4)}{2}} = \frac{\sqrt{2+\sqrt{2}}}{2},\\
\cos\frac{\pi}{16} &= \sqrt{\frac{1+\cos(\pi/8)}{2}} = \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2},\\
&\vdots
\end{aligned}$$

第 n 个因子的分子是 2 的 n 重嵌套平方根. 代入 (25-9):

$$\frac{2}{\pi} = \frac{\sqrt{2}}{2} \cdot \frac{\sqrt{2+\sqrt{2}}}{2} \cdot \frac{\sqrt{2+\sqrt{2+\sqrt{2}}}}{2} \cdots \tag{25-10}$$

Vieta 公式 (1593 年) 是数学史上第一个无穷乘积, 也是 π 的第一个分析表达式. 在它之前, π 只能通过几何作图或多边形逼近来理解; Vieta 公式把 π 写成无穷多个可用尺规作图构造的根号的乘积, 首次用代数运算逼近 π.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

25.3.4 与连分数的对照

第23章 用连分数给无理数配上了嵌套的*加法*结构, 黄金比 $\varphi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cdots}}$, 分子恒为 1, 分母层层嵌套.Vieta 公式给 π 配上了嵌套的*乘法*结构, 每个因子是余弦, 余弦又化为根号, 根号层层嵌套. 连分数的嵌套靠除法传递, Vieta 乘积的嵌套靠开方传递, 二者从不同方向把无理数拆解为可计算的无限过程.

25.4 欧拉乘积

知乎@余命数

25.4.1 黎曼 ζ 函数

第16.4节 引入了 p-级数 $\sum \dfrac{1}{n^p}$, 当 $p > 1$ 时收敛. 把 p 换成连续变量 s, 就得到黎曼 ζ 函数.

余命数 知乎@余命数

定义 25.2 (黎曼 ζ 函数) 对实数 $s > 1$, 定义

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \cdots \tag{25-11}$$

$\zeta(2) = \pi^2/6$ 是巴塞尔问题的解, 这一结果将在后续章节给出.$\zeta(s)$ 在 $s > 1$ 时是 s 的连续函数, 随 s 增大而递减, $\zeta(s) \to 1$ $(s \to \infty)$; 当 $s \to 1^+$ 时 $\zeta(s) \to +\infty$(调和级数发散).

知乎@余命数

25.4.2 欧拉乘积公式

定理 25.5 (欧拉乘积) 对 $s > 1$,

$$\zeta(s) = \prod_{p \text{ 素数}} \frac{1}{1 - p^{-s}} = \prod_{p \text{ 素数}} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \cdots\right) \tag{25-12}$$

其中乘积遍取所有素数 $p = 2, 3, 5, 7, 11, \ldots$.

证明 从 $\zeta(s)$ 出发, 逐步筛去每个素数的倍数.

以 $p = 2$ 为例.$\zeta(s) = 1 + \dfrac{1}{2^s} + \dfrac{1}{3^s} + \dfrac{1}{4^s} + \dfrac{1}{5^s} + \cdots$. 两端乘 $\dfrac{1}{2^s}$:

$$\frac{1}{2^s}\zeta(s) = \frac{1}{2^s} + \frac{1}{4^s} + \frac{1}{6^s} + \frac{1}{8^s} + \cdots$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

相减：

$$\left(1-\frac{1}{2^s}\right)\zeta(s)=1+\frac{1}{3^s}+\frac{1}{5^s}+\frac{1}{7^s}+\frac{1}{9^s}+\cdots \tag{25-13}$$

右端不再含 2 的倍数对应的项.

对 (25-13) 重复同样的操作，两端乘 $\frac{1}{3^s}$：

$$\frac{1}{3^s}\left(1-\frac{1}{2^s}\right)\zeta(s)=\frac{1}{3^s}+\frac{1}{9^s}+\frac{1}{15^s}+\cdots$$

相减：

$$\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\zeta(s)=1+\frac{1}{5^s}+\frac{1}{7^s}+\frac{1}{11^s}+\cdots \tag{25-14}$$

右端不再含 2 或 3 的倍数对应的项.

对每个素数 p 重复这一做法，用 $\left(1-\frac{1}{p^s}\right)$ 乘当前表达式，筛去剩余项中 p 的倍数. 这正是埃拉托色尼筛法，依次用 $2,3,5,7,11,\ldots$ 筛去各自的倍数.

对前 k 个素数 $p_1,p_2,\ldots,p_k$ 执行筛法后：

$$\prod_{i=1}^{k}\left(1-\frac{1}{p_i^s}\right)\zeta(s)=\sum_{\substack{n\geq 1\\ p_1,\ldots,p_k\nmid n}}\frac{1}{n^s}.$$

右端求和遍取不被 $p_1,\ldots,p_k$ 整除的正整数. 当 $k\to\infty$ 时，由 **??** 的算术基本定理，每个大于 1 的整数 n 都有唯一的素因子分解，故 n 必被某个素数整除. 剩余求和中唯一不被任何素数整除的数只有 $n=1$，右端趋于 1：

$$\prod_{p\text{ 素数}}\left(1-\frac{1}{p^s}\right)\zeta(s)=1.$$

移项即得 (25-12). □

25.4.3 几何级数展开

(25-12) 的第二个等号来自几何级数. 对每个素数 p，

$$\frac{1}{1-p^{-s}}=1+p^{-s}+p^{-2s}+p^{-3s}+\cdots=\sum_{k=0}^{\infty}\frac{1}{p^{ks}}.$$

把所有素数对应的几何级数相乘，展开后每一项形如

$$\frac{1}{p_1^{k_1 s}\,p_2^{k_2 s}\cdots p_r^{k_r s}}=\frac{1}{(p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r})^s}=\frac{1}{n^s},$$

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

其中 $n = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$. 由算术基本定理, 每个正整数 n 对应唯一的素数幂分解, 故展开式恰好给出 $\sum \dfrac{1}{n^s} = \zeta(s)$, 每项出现且仅出现一次.

25.4.4 素数与分析的桥梁

欧拉乘积把两个看似无关的对象联系在一起. 左端 $\zeta(s)$ 是连续变量 s 的函数, 由全体正整数构成, 属于分析学; 右端是遍取所有素数的无穷乘积, 素数是离散的、由整除性定义的对象, 属于数论. 一个等式把连续与离散、分析与数论焊接在一起.

?? 定义了素数并给出算术基本定理, 即每个大于 1 的正整数唯一分解为素数之积. 欧拉乘积正是这一唯一性在分析层面的投影, $\zeta(s)$ 对全体正整数求和, 乘积形式对全体素数求积, 二者相等, 靠的就是每个整数的素因子分解唯一.

欧拉乘积还给出了素数无穷的另一个证明: 若素数只有有限个, 乘积 $\prod_p \dfrac{1}{1-p^{-s}}$ 是有限乘积, 对所有 $s > 1$ 取有限值; 但当 $s \to 1^+$ 时 $\zeta(s) \to +\infty$, 有限乘积无法发散到无穷, 矛盾. 它与 **??** 中欧几里得的证明结论相同.

黎曼把 $\zeta(s)$ 的定义域从实数 $s > 1$ 延拓到整个复平面, 发现 $\zeta(s)$ 的非平凡零点全部落在复平面上 $\mathrm{Re}(s) = 1/2$ 这条直线上, 这就是黎曼猜想, 至今未被证明. 零点的分布决定了素数计数的精确渐近行为, 欧拉乘积是这条道路的起点.

知乎@余命数

习题解答

第 1 章预备知识

1.1. **解**

(a) $P(x,y)=x^2+xy+y^2$: 交换 x,y 不变, 为对称式; 在两元轮换下不变, 为轮换式; 各项为二次, 为齐次式.

(b) $P(a,b,c)=a^2b+b^2c+c^2a$: 在 $a\to b\to c\to a$ 轮换下不变, 为轮换式; 交换 a,b 时变式, 非对称式; 各项为三次, 为齐次式.

(c) $P(x,y,z)=x^3+y^3+z^2$: 交换 x,z (或作轮换) 时式子改变, 非对称、非轮换; 各项次数不全相等, 非齐次式.

(d) $P(a,b,c)=(a-b)(b-c)(c-a)$: 在轮换 $a\to b\to c\to a$ 下因子重排而积不变, 为轮换式; 交换 a,b 后整体变号, 非对称式; 展开后各项为三次, 为齐次式.

1.2. **解** 在轮换 $a\to b\to c\to a$ 下, 因式 $(a-b)$ 依次对应 $(b-c)$ 与 $(c-a)$. 由 P 的轮换性, 这两个因式亦为 P 的因式.

1.3. **解** 记原式为 $P(x,y,z)$. 各项为三次项, P 为三次齐次式; 在轮换 $x\to y\to z\to x$ 下不变, 为轮换式.

令 $x=y$, 得

$$P(y,y,z)=y(y^2-z^2)+y(z^2-y^2)+0=0,$$

故 $(x-y)$ 为因式. 由轮换性, $(y-z)$ 与 $(z-x)$ 亦为因式. 于是存在常数 k 使

$$P(x,y,z)=k(x-y)(y-z)(z-x).$$

取 $(x,y,z)=(1,0,-1)$, 有

$$P(1,0,-1)=1(0-1)+0(1-1)+(-1)(1-0)=-2,$$

而

$$(1-0)(0-(-1))(-1-1)=-2,$$

从而 $k=1$. 因而

$$x(y^2-z^2)+y(z^2-x^2)+z(x^2-y^2)=(x-y)(y-z)(z-x).$$

1.4. **解** 以 $y\neq 0$ 为例. 由齐次性, 有

$$F(x,y)=F(ty,y)=y^dF(t,1),$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于是 $F(x,y)=0$ 与 $F(t,1)=0$ 等价 ($y\neq 0$). 原方程的非零解对应于一元方程 $F(t,1)=0$ 的根 $t=\dfrac{x}{y}$, 再由 $x:y=t:1$ 还原.

1.5. 解

$$(x+2y)^2-(x-y)^2=[(x+2y)-(x-y)][(x+2y)+(x-y)]=3y(2x+y).$$

1.6. 解

$$8a^3+27b^3=(2a)^3+(3b)^3=(2a+3b)(4a^2-6ab+9b^2).$$

1.7. 解

$$\left(x+\frac{1}{x}\right)^2=x^2+\frac{1}{x^2}+2=16,$$

故 $x^2+\dfrac{1}{x^2}=14$. 再用

$$x^3+\frac{1}{x^3}=\left(x+\frac{1}{x}\right)^3-3\cdot x\cdot\frac{1}{x}\left(x+\frac{1}{x}\right)=64-12=52.$$

1.8. 解

$$x^4-16=(x^2-4)(x^2+4)=(x-2)(x+2)(x^2+4),$$

$$x^4+4x^2+16=(x^2+2x+4)(x^2-2x+4).$$

1.9. 解

$$2x^2-5x-3=(2x+1)(x-3),$$

$$6x^2+5xy-6y^2=(3x-2y)(2x+3y).$$

1.10. 解 对 $P(x)=2x^3-x^2-8x+4$, 有 $a_3=2, a_0=4$. 根据有理根定理, 可能的有理根为

$$\pm 1,\pm 2,\pm 4,\pm\frac{1}{2}.$$

代入 $x=1$ 得 $P(1)=2-1-8+4=-3\neq 0$, 再代入 $x=2$, 得

$$P(2)=16-4-16+4=0,$$

故 $(x-2)$ 为因式. 综合除法(系数 $2,-1,-8,4$ 与根 2)得商

$$2x^2+3x-2.$$

继续分解

$$2x^2+3x-2=(2x-1)(x+2),$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于是

$$P(x)=(x-2)(2x-1)(x+2).$$

1.11. **解** 先提取二次项系数 2.

$$y=2(x^2-4x)+5$$

再对括号内的部分配方. 一次项系数 -4 的一半是 -2, 其平方为 4.

$$\begin{aligned}y&=2(x^2-4x+4-4)+5\\&=2[(x-2)^2-4]+5\\&=2(x-2)^2-8+5\\&=2(x-2)^2-3\end{aligned}$$

因为 $(x-2)^2\geq 0$, 所以 $y\geq -3$. 当 $x=2$ 时, 函数取得最小值 -3.

1.12. **解** 原方程可变形为 $(2^x)^2-3\cdot 2^x+2=0$. 令 $t=2^x$. 因为指数函数 $y=2^x$ 的值域为 $(0,+\infty)$, 所以新变量 t 必须满足 $t>0$. 原方程化为关于 t 的一元二次方程 $t^2-3t+2=0$. 解得 $(t-1)(t-2)=0$, 即 $t=1$ 或 $t=2$. 这两个 t 值均满足 $t>0$ 的范围限制. 当 $t=1$ 时, $2^x=1\implies x=0$. 当 $t=2$ 时, $2^x=2\implies x=1$. 所以原方程的解为 $x=0$ 或 $x=1$.

1.13. **解** 利用裂项公式 $\dfrac{1}{k(k+1)}=\dfrac{1}{k}-\dfrac{1}{k+1}$, 原式化为

$$\begin{aligned}S&=\sum_{k=1}^{n}\left(\frac{1}{k}-\frac{1}{k+1}\right)\\&=\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\cdots+\left(\frac{1}{n}-\frac{1}{n+1}\right)\\&=1-\frac{1}{n+1}=\frac{n}{n+1}\end{aligned}$$

1.14. **解** 分子次数为 2, 分母次数为 1, 先做多项式除法.

$$x^2+3x+5=(x+1)(x+2)+3$$

故

$$\frac{x^2+3x+5}{x+1}=x+2+\frac{3}{x+1}$$

1.15. **解** 直接求 a,b 再平方求和较繁琐. 注意到

$$a^2+b^2=(a+b)^2-2ab=3^2-2\times 2=5$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第 8 章统计

8.1. **解** $\hat{p}=432/900=0.48$, $\widehat{\mathrm{SE}}=\sqrt{0.48\times0.52/900}\approx0.0167$, 置信区间 $48\%\pm3.3\%$, 即 $[44.7\%,51.3\%]$.

8.2. **解** 标准误与 $\sqrt{n}$ 成反比, 减半需 n 扩大四倍, 即 3600 人. 一般地按最保守的 $p=0.5$ 估计, ±1.5 个百分点为两倍标准误, $2\sqrt{0.25/n}=0.015$ 给出 $n\approx4400$.

8.3. **解** 50 万份回复的随机误差确实极小, 但问卷的样本是自愿点击进入的活跃用户, 不回复者与回复者对收费调整的态度通常恰恰不同, 抽样偏差可能远大于随机误差.《文学文摘》的教训正是样本巨大而偏差更大, 结论是否可信取决于样本如何产生, 与样本量的大小关系不大.

8.4. **解** $\bar{x}=\dfrac{21}{6}=3.5$, $\bar{y}=\dfrac{19}{6}$, 于是

$$d_x=\left(-\tfrac{5}{2},-\tfrac{3}{2},-\tfrac{1}{2},\tfrac{1}{2},\tfrac{3}{2},\tfrac{5}{2}\right),\qquad d_y=\tfrac{1}{6}\left(-7,-13,5,-1,11,5\right).$$

逐项计算得 $\sum d_xd_y=\dfrac{63}{6}=10.5$、$\sum d_x^2=\dfrac{35}{2}=17.5$、$\sum d_y^2=\dfrac{65}{6}$. 代入得

$$r=\frac{10.5}{\sqrt{17.5\times\frac{65}{6}}}\approx0.76,\qquad b=\frac{10.5}{17.5}=\frac{3}{5},\qquad a=\frac{19}{6}-\frac{3}{5}\times\frac{7}{2}=\frac{16}{15}.$$

回归直线为 $\hat{y}=\dfrac{3}{5}x+\dfrac{16}{15}$, 预测 $\hat{y}(8)=\dfrac{88}{15}\approx5.87$.

8.5. **解** $\bar{x}=3$、$d_x=(-2,-1,0,1,2)$、$d_y=(5,0,-3,0,5)$, 于是 $\sum d_xd_y=-10+10=0$, $r=0$. 但每个 x 唯一决定 $y=(x-3)^2$, 关系是完全确定的非线性关系. 相关为零只排除直线关系, 不排除其他形态的依赖.

8.6. **解** 其一, 重视教育的家庭既多买书又多陪伴, 家庭条件与教育态度是混杂变量, 藏书量只是它们的外显标志. 其二, 观察性数据里的斜率只描述人群关联, 因果结论需要随机化对照实验, 而随机指定家庭买书并不现实. 相关为预测提供依据, 不为干预背书.

第 12 章解三角形

12.1. **解** 中线 CD 把 $\triangle ABC$ 分成 $\triangle ADC$ 与 $\triangle BDC$. 它们共享边 CD, 且 $\angle ADC$ 与 $\angle BDC$ 互补. 我们对这两个三角形分别用余弦定理, 消去未知角, 建立关于边长的方程.

对于第 (1) 问.

我们利用 $\cos\angle ADC=-\cos\angle BDC$ 对两个小三角形分别使用余弦定理, 消去未知角.

设 $\angle ADC=\alpha$, 则 $\angle BDC=\pi-\alpha$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

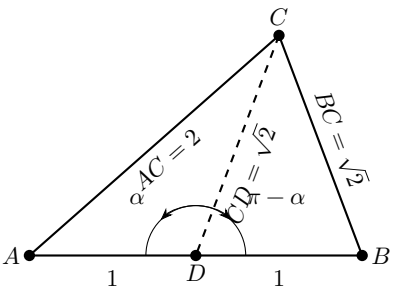


图 25.1 第一问

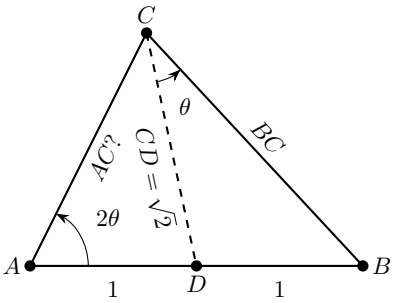


图 25.2 第二问

在 $\triangle ADC$ 中, D 是 AB 中点, 故 $AD=1$. 应用余弦定理

$$AC^2 = AD^2 + CD^2 - 2(AD)(CD)\cos\alpha$$

$$AC^2 = 1^2 + (\sqrt{2})^2 - 2(1)(\sqrt{2})\cos\alpha = 3 - 2\sqrt{2}\cos\alpha$$

在 $\triangle BDC$ 中, $BD=1$, $BC=\sqrt{2}$. 应用余弦定理

$$BC^2 = BD^2 + CD^2 - 2(BD)(CD)\cos(\pi-\alpha)$$

$$(\sqrt{2})^2 = 1^2 + (\sqrt{2})^2 - 2(1)(\sqrt{2})(-\cos\alpha)$$

$$2 = 1 + 2 + 2\sqrt{2}\cos\alpha$$

从 $\triangle BDC$ 的方程中, 可解出 $\cos\alpha$:

$$2 = 3 + 2\sqrt{2}\cos\alpha \implies 2\sqrt{2}\cos\alpha = -1 \implies \cos\alpha = -\frac{1}{2\sqrt{2}} = -\frac{\sqrt{2}}{4}$$

将 $\cos\alpha$ 的值代回 $\triangle ADC$ 的方程中, 求解 AC:

$$AC^2 = 3 - 2\sqrt{2}\left(-\frac{\sqrt{2}}{4}\right) = 3 + \frac{4}{4} = 4$$

所以, $AC=2$.

对于第 (2) 问.

本问的条件跨越两个三角形. 除余弦定理外, 我们用正弦定理建立角与边的关系, 得到求解 AC 与 BC 所需的两个方程.

先用余弦定理建立边之间的关系.

这一步与第 (1) 问相同, 得到关于边长的普适方程. 在 $\triangle ADC$ 中,

$$\cos\angle ADC = \frac{AD^2 + CD^2 - AC^2}{2(AD)(CD)} = \frac{1 + 2 - AC^2}{2\sqrt{2}} = \frac{3 - AC^2}{2\sqrt{2}}.$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

在 $\triangle BDC$ 中,

$$\cos\angle BDC=\frac{BD^2+CD^2-BC^2}{2(BD)(CD)}=\frac{1+2-BC^2}{2\sqrt{2}}=\frac{3-BC^2}{2\sqrt{2}}.$$

因为 $\cos\angle ADC=-\cos\angle BDC$, 所以:

$$\frac{3-AC^2}{2\sqrt{2}}=-\frac{3-BC^2}{2\sqrt{2}}$$

整理得 $3-AC^2=-3+BC^2$, 即第一个方程:

$$AC^2+BC^2=6\quad\cdots\cdots(1)$$

再建立角与边的关系.

设 $\angle BCD=\theta$, 则由题意 $\angle A=2\theta$. 需在不同的三角形中应用正弦定理, 借公共角或等角传递比例关系.

在 $\triangle BCD$ 中, 对 $\angle B$ 和 $\angle BCD$ 应用正弦定理:

$$\frac{CD}{\sin B}=\frac{BD}{\sin\theta}\implies\frac{\sqrt{2}}{\sin B}=\frac{1}{\sin\theta}\implies\sin B=\sqrt{2}\sin\theta$$

在 $\triangle ABC$ 中, 对 $\angle A$ 和 $\angle B$ 应用正弦定理:

$$\frac{BC}{\sin A}=\frac{AC}{\sin B}\implies BC\sin B=AC\sin A$$

将已得的关系 $\sin B=\sqrt{2}\sin\theta$ 和 $A=2\theta$ 代入上式, 得

$$BC(\sqrt{2}\sin\theta)=AC\sin(2\theta)$$

利用二倍角公式 $\sin(2\theta)=2\sin\theta\cos\theta$:

$$\sqrt{2}BC\sin\theta=2AC\sin\theta\cos\theta$$

由于 θ 是三角形内角的一部分, $\sin\theta\neq0$, 可约去 $\sqrt{2}\sin\theta$, 得到第二个方程:

$$BC=\sqrt{2}AC\cos\theta\quad\cdots\cdots(2)$$

联立求解, 消去未知角 θ.

此时方程组仍含三个未知数 AC, BC, θ. 我们消去 θ. 回到 $\triangle BCD$, 对角 θ 应用余弦定理, 用边长表示 $\cos\theta$:

$$\cos\theta=\frac{BC^2+CD^2-BD^2}{2(BC)(CD)}=\frac{BC^2+2-1}{2(BC)\sqrt{2}}=\frac{BC^2+1}{2\sqrt{2}BC}$$

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

将这个 $\cos\theta$ 的表达式代入方程 (2) 中:

$$BC = \sqrt{2}AC\left(\frac{BC^2+1}{2\sqrt{2}BC}\right)$$

化简得

$$2BC^2 = AC(BC^2+1) \quad \cdots\cdots(3)$$

解这个代数方程组.

接着已有两个只含 AC 和 BC 的方程,联立 (1) 和 (3) 求解. 设 $x = AC, y = BC$. 方程组为

$$\begin{cases} x^2+y^2=6 \\ 2y^2 = x(y^2+1) \end{cases}$$

由第二式得 $x = \frac{2y^2}{y^2+1}$. 代入第一式:

$$\left(\frac{2y^2}{y^2+1}\right)^2 + y^2 = 6$$

$$\frac{4y^4}{(y^2+1)^2} + y^2 = 6$$

令 $u = y^2$. 则 $\frac{4u^2}{(u+1)^2} + u = 6$.

$$4u^2 + u(u+1)^2 = 6(u+1)^2$$

$$4u^2 + u(u^2+2u+1) = 6(u^2+2u+1)$$

$$4u^2 + u^3 + 2u^2 + u = 6u^2 + 12u + 6$$

$$u^3 - 11u - 6 = 0$$

通过试根法,可见 $u=-3$ 是一个根,故 $(u+3)$ 是一个因式. $(u+3)(u^2-3u-2)=0$. 由于 $u=y^2$ 必须为正,$u=-3$ 舍去. 解 $u^2-3u-2=0$ 得 $u = \frac{3\pm\sqrt{9-4(-2)}}{2} = \frac{3\pm\sqrt{17}}{2}$. 由于 $u>0$,故 $u = y^2 = BC^2 = \frac{3+\sqrt{17}}{2}$.

最后,代入方程 (1) 求 AC^2:

$$AC^2 = 6 - BC^2 = 6 - \frac{3+\sqrt{17}}{2} = \frac{12-3-\sqrt{17}}{2} = \frac{9-\sqrt{17}}{2}$$

所以

$$AC = \sqrt{\frac{9-\sqrt{17}}{2}} = \frac{\sqrt{18-2\sqrt{17}}}{2} = \frac{\sqrt{(\sqrt{17}-1)^2}}{2} = \frac{\sqrt{17}-1}{2}$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

(1) $AC=2$.(2) $AC=\frac{\sqrt{17}-1}{2}$.

12.2. **解** 条件 $\angle CAD=\angle B$ 跨越两个三角形, 提示我们寻找相似三角形. 一旦发现相似, 边长关系便可直接由比例得到.

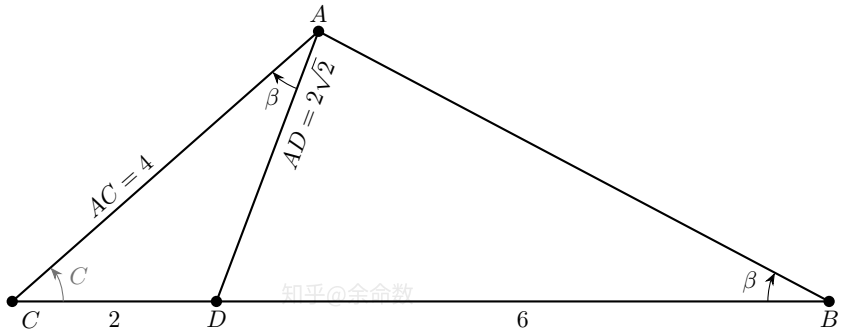


图 25.3 根据第 (2) 问数据按比例精确绘制的图示

知乎@余命数

对于第 (1) 问.

考察 $\triangle ADC$ 和 $\triangle BAC$.

- $\angle CAD=\angle B$(已知条件)
- $\angle ACD=\angle BCA$(公共角)

根据 AA 相似判定定理, $\triangle BAC\sim\triangle ADC$.

由相似三角形的性质, 对应边成比例.

$$\frac{BC}{AC}=\frac{AC}{DC}=\frac{BA}{AD}$$

取前两项, 交叉相乘即得

$$AC^2=BC\cdot CD$$

对于第 (2) 问.

求 $\triangle ABD$ 的面积, 可用 $S_{\triangle ABD}=S_{\triangle ABC}-S_{\triangle ADC}$ 或 $\frac{S_{\triangle ABD}}{S_{\triangle ADC}}=\frac{BD}{CD}$. 两种方法都需要先求 $\triangle ADC$ 的元素与 BD.

首先, 在 $\triangle ADC$ 中, 已知两边 AD,CD 和一角 $\cos C$. 可用余弦定理求出边 AC.

$$AD^2=AC^2+CD^2-2(AC)(CD)\cos C$$

代入已知值 $AD=2\sqrt{2},CD=2,\cos C=\frac{3}{4}$:

$$(2\sqrt{2})^2=AC^2+2^2-2(AC)(2)\left(\frac{3}{4}\right)$$

$$8=AC^2+4-3AC$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

整理得一个关于 AC 的一元二次方程:

$$AC^2 - 3AC - 4 = 0$$

因式分解得 $(AC-4)(AC+1)=0$. 由于 AC 为边长, 必为正数, 故 $AC=4$.

接下来, 利用第 (1) 问的结论求出 BC:

$$BC = \frac{AC^2}{CD} = \frac{4^2}{2} = 8$$

从而可以求出 BD 的长度:

$$BD = BC - CD = 8 - 2 = 6$$

接着用面积比来求解. 首先计算 $\triangle ADC$ 的面积. 已知 $\cos C = \frac{3}{4}$, 且 $C \in (0,\pi)$, 故 $\sin C = \sqrt{1-\cos^2 C} = \sqrt{1-\frac{9}{16}} = \frac{\sqrt{7}}{4}$.

$$S_{\triangle ADC} = \frac{1}{2} AC \cdot CD \sin C = \frac{1}{2}(4)(2)\left(\frac{\sqrt{7}}{4}\right) = \sqrt{7}$$

由于 $\triangle ABD$ 和 $\triangle ADC$ 共用从顶点 A 出发的高线, 它们的面积之比等于底边之比:

$$\frac{S_{\triangle ABD}}{S_{\triangle ADC}} = \frac{BD}{CD} = \frac{6}{2} = 3$$

所以

$$S_{\triangle ABD} = 3 \cdot S_{\triangle ADC} = 3\sqrt{7}$$

(1) 证明见上.(2) $\triangle ABD$ 的面积为 $3\sqrt{7}$.

第 13 章解析几何

13.1. **解** $a=3, b=4, c=5$. 切线 $\frac{5x}{9} - \frac{y}{3} = 1$, 化为 $5x-3y-9=0$, 法向量 $\vec{n}=(5,-3)$. 两焦半径 $|PF_1| = \frac{25}{3}+3 = \frac{34}{3}$, $|PF_2| = \frac{25}{3}-3 = \frac{16}{3}$. 算数量积, $\vec{n}\cdot\overrightarrow{PF_1} = 5\cdot(-10) - 3\cdot\left(-\frac{16}{3}\right) = -34$, 除以 $|PF_1| = \frac{34}{3}$ 得 -3; $\vec{n}\cdot\overrightarrow{PF_2} = 5\cdot 0 - 3\cdot\left(-\frac{16}{3}\right) = 16$, 除以 $|PF_2| = \frac{16}{3}$ 得 3. 两个夹角余弦大小相等, 切线与两焦半径成等角, 且平分的是外角.

13.2. **解** 设它们交于 P, 共同焦点为 F_1, F_2. 椭圆在 P 处的切线是 $\angle F_1PF_2$ 的内角平分线, 双曲线在 P 处的切线是外角平分线(第13.3节 与 定理 13.43), 而同一顶

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

角的内、外角平分线互相垂直, 故两条切线垂直. 设两条焦半径的方向角为 0° 与 θ, 内角平分线方向 $\theta/2$, 外角平分线方向 $\theta/2+90^\circ$, 两者总是差 90°.

13.3. **解** 过轴的截面是抛物线, 顶点为原点、轴向为 x 轴时, 口径边缘的点为 $(0.1, 0.6)$, 代入 $y^2=2px$ 得 $0.36=0.2p$, $p=1.8$. 灯丝放在焦点, 距顶点 $\dfrac{p}{2}=0.9$ m.

13.4. **解** 例题已算出 $|AF|=\dfrac{6}{5}$, $|BF|=2$, 于是 $\dfrac{5}{6}+\dfrac{1}{2}=\dfrac{4}{3}=\dfrac{2a}{b^2}$, 与定理一致, $|AB|=\dfrac{16}{5}$. 也可以不走例题, 直接对 $r(\theta)=\dfrac{3/2}{1+\frac{1}{2}\cos\theta}$ 取 $\theta=60^\circ$ 与 240° 相加.

13.5. **解** 渐近线斜率 $\pm\dfrac{\sqrt{3}}{2}$, 弦的斜率 1 比渐近线更陡, 与右支交出两个点, 两端在同一支上, 调和式适用, 直接得

$$\frac{1}{|AF|}+\frac{1}{|BF|}=\frac{2a}{b^2}=\frac{4}{3}.$$

若对同一支的判断不放心, 可解直线 $y=x-\sqrt{7}$ 与双曲线的联立, 两个交点的横坐标均为正, 确在右支上.

13.6. **解** 由 $\dfrac{1}{|AF|}+\dfrac{1}{|BF|}=\dfrac{2}{p}=1$ 得 $\dfrac{1}{|BF|}=\dfrac{3}{4}$, $|BF|=\dfrac{4}{3}$. 设 A 在倾斜角 θ 方向, $|AF|=\dfrac{p}{1-\cos\theta}=4$ 给 $\cos\theta=\dfrac{1}{2}$, $\theta=60^\circ$. 此时 A 的参数 $t_1=\sqrt{3}$(由 $\cos\theta=\dfrac{t^2-1}{t^2+1}$), $t_2=-\dfrac{1}{\sqrt{3}}$, 切线交点由 定理 13.46 落在准线上, 坐标 $T\left(-1,\ \dfrac{p(t_1+t_2)}{2}\right)=\left(-1,\ \dfrac{2}{\sqrt{3}}\right)$.

第 15 章射影几何

15.1. **解** 记 $X=AB\cap DE$、$Y=BC\cap EF$、$Z=CD\cap FA$. 前两对平行使 X、Y 都是无穷远点, 帕斯卡线 (定理 15.8) 只能是无穷远直线, 于是 Z 也在无穷远直线上, 即 $CD\parallel FA$. 这与 第15.5节 末尾帕普斯定理的平行情形如出一辙.

15.2. **解** 把六边形的边依次取成 AB, AB, BC, BC, CA, CA, 相邻两边重合处, 顶点塌缩成切点, 六个顶点依次为 Z, B, X, C, Y, A. 布里昂桑定理给出三条对角线 ZC、BY、XA 共点, 即 AX、BY、CZ 共点.Ceva 复核时注意从每个顶点引出的两条切线等长, $AZ=AY$、$BZ=BX$、$CX=CY$, 于是

$$\frac{BX}{XC}\cdot\frac{CY}{YA}\cdot\frac{AZ}{ZB}=\frac{BZ}{CY}\cdot\frac{CY}{AZ}\cdot\frac{AZ}{BZ}=1,$$

乘积为一, 三线共点.

15.3. **解** 把 P 视作六边形相邻两顶点的重合, 切线是退化的那条边. 取六边形 $P, P, 1, 2, 3, 4$, 对边交点依次为 $X=$ 切线 $_P\cap 23$、$Y=P1\cap 34$、$Z=12\cap 4P$, 三者

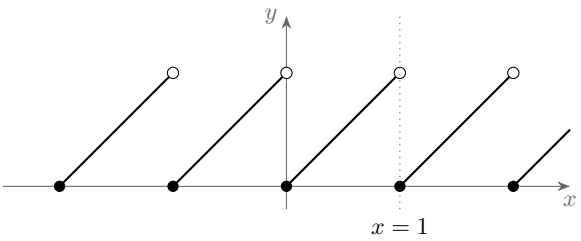
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

共线. 于是先作 $Y = P1 \cap 34$ 与 $Z = 12 \cap 4P$, 连 YZ 交直线 23 于 X, 连 PX 即为切线, 依据是退化帕斯卡.

第 16 章极限与导数

16.1. **解** 先题

函数 $f(x) = x - \lfloor x \rfloor$ 表示 x 的小数部分. 为了判断在 $x = 1$ 处的极限是否存在, 我们分别考察其左极限与右极限.



当 x 从右侧趋近于 1 时, 记作 $x \to 1^+$. 此时 x 的取值略大于 1, 例如 $1.1, 1.01, \ldots$. 对于这些 x 值, $\lfloor x \rfloor = 1$. 因此右极限为

$$\lim_{x \to 1^+} (x - \lfloor x \rfloor) = \lim_{x \to 1^+} (x - 1) = 1 - 1 = 0$$

当 x 从左侧趋近于 1 时, 记作 $x \to 1^-$. 此时 x 的取值略小于 1, 例如 $0.9, 0.99, \ldots$. 对于这些 x 值, $\lfloor x \rfloor = 0$. 因此左极限为

$$\lim_{x \to 1^-} (x - \lfloor x \rfloor) = \lim_{x \to 1^-} (x - 0) = 1 - 0 = 1$$

由于左极限与右极限存在但不相等 ($0 \neq 1$), 根据极限存在的充要条件, 函数 $f(x)$ 在 $x = 1$ 处的极限不存在.

16.2. **解** 再题

按 $\varepsilon - \delta$ 定义, 我们需对任意 $\varepsilon > 0$ 找到 $\delta > 0$, 使当 $0 < |x - c| < \delta$ 时恒有 $|(ax + b) - (ac + b)| < \varepsilon$. 将目标式化简, 可知误差仅由 $|x - c|$ 决定.

$$|(ax + b) - (ac + b)| = |a(x - c)| = |a| \cdot |x - c|.$$

因 $a \neq 0$, $|a| > 0$, 两边除以 $|a|$ 即得 $|x - c| < \dfrac{\varepsilon}{|a|}$, 这给出了 δ 的选取.

对任意 $\varepsilon > 0$, 取 $\delta = \dfrac{\varepsilon}{|a|}$. 由 $\varepsilon > 0$ 且 $|a| > 0$ 知 $\delta > 0$. 当 $0 < |x - c| < \delta$ 时, $|x - c| < \dfrac{\varepsilon}{|a|}$, 两边乘 $|a|$ 得

$$|a| \cdot |x - c| < \varepsilon, \quad \text{即} \quad |(ax + b) - (ac + b)| < \varepsilon,$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

此即所需不等式. 由定义, $\lim\limits_{x\to c}(ax+b)=ac+b$.

16.3. **解** 接着题

按 $\varepsilon-\delta$ 定义, 我们需对任意 $\varepsilon>0$ 找到 $\delta>0$, 使当 $0<|x-c|<\delta$ 时恒有 $|\sqrt{x}-\sqrt{c}|<\varepsilon$. 对目标式有理化, 把它与可控的 $|x-c|$ 相联系.

$$|\sqrt{x}-\sqrt{c}|=\left|\frac{x-c}{\sqrt{x}+\sqrt{c}}\right|=\frac{|x-c|}{\sqrt{x}+\sqrt{c}}.$$

分母中的 $\sqrt{x}$ 非常数, 但因 x 趋近于 c, 我们先限制 x 的范围以获取分母下界. 预先要求 $\delta\leqslant c/2$. 若 $|x-c|<c/2$, 则 $x>c/2$, 从而 $\sqrt{x}>\sqrt{c/2}$, 故 $\sqrt{x}+\sqrt{c}>\sqrt{c}$ (取此宽松下界以简化估计). 于是

$$|\sqrt{x}-\sqrt{c}|=\frac{|x-c|}{\sqrt{x}+\sqrt{c}}<\frac{|x-c|}{\sqrt{c}},$$

要其小于 ε, 只需 $|x-c|\leqslant\varepsilon\sqrt{c}$. 合并两个约束, 取 $\delta=\min\{c/2,\,\varepsilon\sqrt{c}\}$.

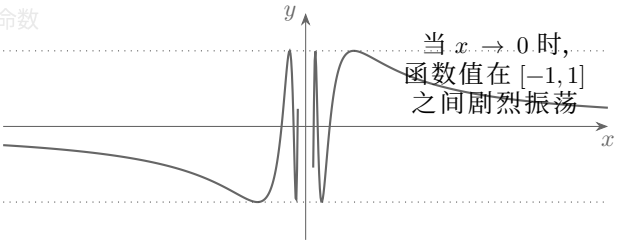
对任意 $\varepsilon>0$, 取上述 δ. 由 $c>0,\varepsilon>0$ 知 $\delta>0$. 当 $0<|x-c|<\delta$ 时, 同时有 $|x-c|<c/2$ 与 $|x-c|<\varepsilon\sqrt{c}$. 由前者得 $x>c/2$, 故 $\sqrt{x}+\sqrt{c}>\sqrt{c}$, 从而

$$|\sqrt{x}-\sqrt{c}|=\frac{|x-c|}{\sqrt{x}+\sqrt{c}}<\frac{|x-c|}{\sqrt{c}}<\frac{\varepsilon\sqrt{c}}{\sqrt{c}}=\varepsilon.$$

由定义, $\lim\limits_{x\to c}\sqrt{x}=\sqrt{c}$.

16.4. **解** 最后题

函数 $g(x)=\sin(1/x)$ 在 $x\to0$ 时的极限不存在. 我们构造两个都趋近于 0 的序列 $\{x_n\}$ 和 $\{y_n\}$, 使对应的函数值序列 $\{g(x_n)\}$ 和 $\{g(y_n)\}$ 收敛到不同的极限.



先构造函数值始终为 1 的序列. 令 $\frac{1}{x_n}=2n\pi+\frac{\pi}{2}$, 其中 n 为正整数. 则 $x_n=\dfrac{1}{2n\pi+\pi/2}$. 由 $x_n=\dfrac{1}{2n\pi+\pi/2}$ 知, 当 $n\to\infty$ 时 $x_n\to0$. 对应的函数值为

$$g(x_n)=\sin(2n\pi+\pi/2)=\sin(\pi/2)=1$$

因此 $\lim\limits_{n\to\infty}g(x_n)=1$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

再构造函数值始终为 0 的序列. 令 $\frac{1}{y_n} = n\pi$, 其中 n 为正整数. 则 $y_n = \dfrac{1}{n\pi}$. 由 $y_n = \dfrac{1}{n\pi}$ 知, 当 $n \to \infty$ 时 $y_n \to 0$. 对应的函数值为

$$g(y_n) = \sin(n\pi) = 0$$

因此 $\lim\limits_{n\to\infty} g(y_n) = 0$.

我们找到了两条趋近于 0 的路径, 对应的函数值分别趋于 1 和 0. 因此 $g(x)$ 在 $x \to 0$ 时没有确定的极限. 故 $\lim\limits_{x\to 0} \sin\left(\frac{1}{x}\right)$ 不存在.

16.5. **解** 第五题

我们需找到一个 $\delta > 0$, 使得当 $0 < |x-2| < \delta$ 时, 恒有 $|x^3 - 8| < 0.1$.

先分析不等式 $|x^3 - 8| < 0.1$.

$$|x^3 - 8| = |(x-2)(x^2+2x+4)| = |x-2| \cdot |x^2+2x+4|$$

我们希望这个乘积小于 0.1.

能直接控制的项是 $|x-2|$. 麻烦在于 $|x^2+2x+4|$ 随 x 变化. 但 x 在 2 附近, 我们预先限制 δ 的范围即可框住 $|x^2+2x+4|$ 的大小.

为限定范围, 先取 $\delta \le 1$. 若 $|x-2| < 1$, 则 $1 < x < 3$. 在此范围内, 我们估计 $|x^2+2x+4|$ 的上界. 由于 $x^2, 2x, 4$ 在此区间均为正的增函数, 故其和在 $x = 3$ 处取最大值.

$$|x^2+2x+4| < 3^2 + 2(3) + 4 = 9 + 6 + 4 = 19$$

结合该上界, 不等式变为

$$|x^3 - 8| = |x-2| \cdot |x^2+2x+4| < \delta \cdot 19$$

为了满足 $|x^3 - 8| < 0.1$, 只需要 $19\delta \le 0.1$. 这要求 $\delta \le \dfrac{0.1}{19}$.

于是 δ 需同时满足 $\delta \le 1$ 和 $\delta \le \frac{0.1}{19}$. 为同时满足, 取二者中较小的一个即可. 因为 $\frac{0.1}{19} < 1$, 故可取 $\delta = \frac{0.1}{19}$.

例如, 取 $\delta = \frac{0.1}{19} \approx 0.00526$. 这是一个满足条件的具体正数.

16.6. **解** 第六题

我们采用反证法. 假设极限存在但不唯一, 即 $\lim\limits_{x\to x_0} f(x) = L_1$ 且 $\lim\limits_{x\to x_0} f(x) = L_2$, 其中 $L_1 \ne L_2$.

由 $L_1 \ne L_2$ 知, $|L_1 - L_2|$ 是一个正数. 选取 $\varepsilon = \dfrac{|L_1 - L_2|}{2}$. 由 $|L_1 - L_2| > 0$ 知 $\varepsilon > 0$.

由极限的定义可知, 因为 $\lim\limits_{x\to x_0} f(x) = L_1$, 所以对上述 ε, 存在 $\delta_1 > 0$, 使得当 $0 < |x - x_0| < \delta_1$ 时, 有

$$|f(x) - L_1| < \varepsilon$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

又因为 $\lim\limits_{x\to x_0} f(x)=L_2$,所以对同一个 ε,也存在 $\delta_2>0$,使得当 $0<|x-x_0|<\delta_2$ 时,有

$$|f(x)-L_2|<\varepsilon$$

取 $\delta=\min\{\delta_1,\delta_2\}$. 这是一个正数. 当 x 满足 $0<|x-x_0|<\delta$ 时,上述两个关于 $f(x)$ 的不等式将同时成立.

对于这样的 x,我们考察 $|L_1-L_2|$,并拆项.

$$\begin{aligned}|L_1-L_2| &= |(L_1-f(x))+(f(x)-L_2)|\\ &\le |L_1-f(x)|+|f(x)-L_2| \quad (\text{三角不等式})\\ &= |f(x)-L_1|+|f(x)-L_2|\\ &< \varepsilon+\varepsilon=2\varepsilon\end{aligned}$$

将 $\varepsilon=\dfrac{|L_1-L_2|}{2}$ 代入上式,得到

$$|L_1-L_2|<2\cdot\frac{|L_1-L_2|}{2}=|L_1-L_2|$$

我们得到一个严格的不等式 $|L_1-L_2|<|L_1-L_2|$. 这个矛盾源于最初假设 $L_1\neq L_2$. 因此该假设必定是错误的.

故必有 $L_1=L_2$. 极限的唯一性得证.

16.7. 解 第七题

按 $\varepsilon-\delta$ 定义,我们需对任意 $\varepsilon>0$ 找到 $\delta>0$,使当 $0<|x|<\delta$ 时恒有 $\left|x\sin\frac{1}{x}\right|<\varepsilon$. 我们对振荡因子 $\sin\frac{1}{x}$ 用一致界 $|\sin u|\le 1$ 放缩,把估计归结为 $|x|$.

$$\left|x\sin\frac{1}{x}\right|=|x|\cdot\left|\sin\frac{1}{x}\right|\le|x|\cdot 1=|x|.$$

故只要保证 $|x|<\varepsilon$ 即可,这给出 δ 的直接选取.

对任意 $\varepsilon>0$,取 $\delta=\varepsilon$,由 $\varepsilon>0$ 知 $\delta>0$. 当 $0<|x|<\delta$ 时,

$$\left|x\sin\frac{1}{x}\right|=|x|\left|\sin\frac{1}{x}\right|\le|x|<\varepsilon.$$

由定义,$\lim\limits_{x\to0}\left(x\sin\frac{1}{x}\right)=0$.

16.8. 解 第八题

如果 $f(x)$ 的极限为正数,那么它在 c 的某个去心邻域内恒为正.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

已知 $\lim\limits_{x\to c} f(x) = L$ 且 $L > 0$. 由 $\varepsilon - \delta$ 定义, 对任意 $\varepsilon > 0$ 都存在 $\delta > 0$, 使当 $0 < |x - c| < \delta$ 时恒有 $|f(x) - L| < \varepsilon$. 取 $\varepsilon = \dfrac{L}{2} > 0$, 则相应存在 $\delta > 0$, 使当 $0 < |x - c| < \delta$ 时

$$|f(x) - L| < \frac{L}{2}.$$

将该绝对值不等式展开并在两边同加 L、化简, 得

$$\begin{aligned} -\frac{L}{2} < f(x) - L < \frac{L}{2} &\iff L - \frac{L}{2} < f(x) < L + \frac{L}{2} \quad \text{(两边加 } L\text{)} \\ &\iff \frac{L}{2} < f(x) < \frac{3L}{2}. \end{aligned}$$

由其左半部分得 $f(x) > \dfrac{L}{2}$. 因 $L > 0$, $\dfrac{L}{2}$ 亦为正数, 故对所有满足 $0 < |x - c| < \delta$ 的 x 都有 $f(x) > 0$.

因此存在正数 δ, 使在 c 的去心邻域 $U^\circ(c, \delta)$ 内 $f(x)$ 恒为正. 证毕.

16.9. **解** 先题

设 x_0 为定义域内任意正实数. 需计算曲线在点 $P(x_0, \sqrt{x_0})$ 处的切线斜率.

根据切线斜率的定义, 我们考察连接点 P 与邻近动点 $Q(x_0 + \Delta x, \sqrt{x_0 + \Delta x})$ 的割线斜率, 并令 $\Delta x \to 0$.

$$k_{PQ} = \frac{\sqrt{x_0 + \Delta x} - \sqrt{x_0}}{\Delta x}$$

此为 $\frac{0}{0}$ 型不定式. 为消去不定性, 我们对分子乘以共轭表达式

$$\begin{aligned} k_{PQ} &= \frac{\sqrt{x_0 + \Delta x} - \sqrt{x_0}}{\Delta x} \cdot \frac{\sqrt{x_0 + \Delta x} + \sqrt{x_0}}{\sqrt{x_0 + \Delta x} + \sqrt{x_0}} \\ &= \frac{(x_0 + \Delta x) - x_0}{\Delta x(\sqrt{x_0 + \Delta x} + \sqrt{x_0})} \\ &= \frac{\Delta x}{\Delta x(\sqrt{x_0 + \Delta x} + \sqrt{x_0})} \end{aligned}$$

当 $\Delta x \neq 0$ 时约去公因子 Δx, 得 $k_{PQ} = \dfrac{1}{\sqrt{x_0 + \Delta x} + \sqrt{x_0}}$. 令 $\Delta x \to 0$, 由 $\sqrt{x}$ 在 x_0 处的连续性(根函数在正数处连续), 得

$$k_{\tan} = \lim_{\Delta x \to 0} \frac{1}{\sqrt{x_0 + \Delta x} + \sqrt{x_0}} = \frac{1}{\sqrt{x_0 + 0} + \sqrt{x_0}} = \frac{1}{2\sqrt{x_0}}.$$

此结论对于任意 $x_0 > 0$ 均成立. 证毕.

16.10. **解** 再题

所求切线的切点位于曲线外, 因此我们先设切点为曲线上的点 $P(x_0, x_0^2)$.

知乎@余命数

曲线 $y=x^2$ 在点 P 处的切线斜率为 $2x_0$. 于是过点 P 的切线方程为

$$y-x_0^2=2x_0(x-x_0)$$

整理得

$$y=2x_0x-x_0^2$$

此切线过点 $A(1,-3)$, 于是点 A 的坐标满足该方程

$$-3=2x_0(1)-x_0^2$$

知乎@余命数

得到关于 x_0 的一个二次方程

$$x_0^2-2x_0-3=0$$

因式分解得 $(x_0-3)(x_0+1)=0$, 解得 $x_0=3$ 或 $x_0=-1$.

即存在两条满足条件的切线.

当 $x_0=3$ 时, 切点为 $(3,9)$, 切线斜率为 $2(3)=6$. 切线方程为

$$y-9=6(x-3)\quad 即\quad y=6x-9$$

当 $x_0=-1$ 时, 切点为 $(-1,1)$, 切线斜率为 $2(-1)=-2$. 切线方程为

$$y-1=-2(x-(-1))\quad 即\quad y=-2x-1$$

故所求的切线方程为 $y=6x-9$ 与 $y=-2x-1$.

16.11. **解** 接着题

函数 $h(x)=|x^2-4|$ 可写为分段函数

$$h(x)=\begin{cases}x^2-4, & x\leq -2\text{ 或}x\geq 2\\ -(x^2-4), & -2<x<2\end{cases}$$

切线存在的充分必要条件是定义切线斜率的极限存在. 这要求割线斜率的左、右极限均存在且相等. 我们需在点 $x=2$ 处分别考察其左、右两侧割线斜率的极限, 其中分子里的 $h(2)=|2^2-4|=0$.

从右侧考察, 当 $x\to 2^+$ 时, $x>2$, 此时 $h(x)=x^2-4$.

$$\lim_{x\to 2^+}\frac{h(x)-h(2)}{x-2}=\lim_{x\to 2^+}\frac{(x^2-4)-0}{x-2}=\lim_{x\to 2^+}\frac{(x-2)(x+2)}{x-2}$$

$$=\lim_{x\to 2^+}(x+2)=4$$

知乎@余命数

从左侧考察, 当 $x \to 2^-$ 时, $x < 2$, 此时 $h(x) = -(x^2-4)$.

$$\lim_{x\to 2^-} \frac{h(x)-h(2)}{x-2} = \lim_{x\to 2^-} \frac{-(x^2-4)-0}{x-2} = \lim_{x\to 2^-} \frac{-(x-2)(x+2)}{x-2}$$
$$= \lim_{x\to 2^-} -(x+2) = -4$$

由于左、右极限不相等 ($4 \neq -4$), 因此函数在点 $x = 2$ 处割线斜率的极限 $\lim\limits_{x\to 2} \frac{h(x)-h(2)}{x-2}$ 不存在.

所以函数 $h(x) = |x^2-4|$ 在点 $x = 2$ 处不存在切线. 其代数原因为割线斜率的左、右极限不一致; 几何原因为该点是一个尖点, 曲线在此处方向发生剧烈转折, 无法用唯一的直线进行线性近似.

16.12. **解** 最后题

两条曲线在某点相切, 要求它们在该点有共同的函数值和切线斜率. 设切点为 $P(x_0, y_0), x_0 > 0$.

函数值相等给出 $ax_0^2 = \ln x_0 \quad (1)$. 切线斜率亦须相等. 对于 $y = ax^2$, 在 x_0 处的斜率为

$$k_1 = \lim_{\Delta x\to 0} \frac{a(x_0+\Delta x)^2 - ax_0^2}{\Delta x} = \lim_{\Delta x\to 0} \frac{a(2x_0\Delta x + (\Delta x)^2)}{\Delta x}$$
$$= \lim_{\Delta x\to 0} a(2x_0+\Delta x) = 2ax_0$$

对于 $y = \ln x$, 在 x_0 处的斜率为(此过程依赖重要极限 $\lim\limits_{t\to 0}(1+t)^{1/t} = e$)

$$k_2 = \lim_{\Delta x\to 0} \frac{\ln(x_0+\Delta x) - \ln x_0}{\Delta x} = \lim_{\Delta x\to 0} \frac{1}{\Delta x} \ln\left(\frac{x_0+\Delta x}{x_0}\right)$$
$$= \lim_{\Delta x\to 0} \ln\left(\left(1+\frac{\Delta x}{x_0}\right)^{\frac{1}{\Delta x}}\right) = \ln\left[\lim_{\Delta x\to 0}\left(\left(1+\frac{\Delta x}{x_0}\right)^{\frac{x_0}{\Delta x}}\right)^{\frac{1}{x_0}}\right]$$
$$= \ln(e^{1/x_0}) = \frac{1}{x_0}$$

令斜率相等, 则 $2ax_0 = \frac{1}{x_0} \quad (2)$.

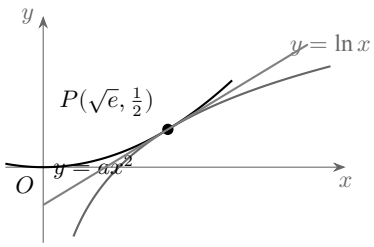
从 (2) 解得 $a = \frac{1}{2x_0^2}$. 代入 (1) 得

$$\left(\frac{1}{2x_0^2}\right) x_0^2 = \ln x_0 \implies \frac{1}{2} = \ln x_0$$

解得切点横坐标 $x_0 = e^{1/2} = \sqrt{e}$. 将 x_0 代回 a 的表达式: $a = \frac{1}{2(\sqrt{e})^2} = \frac{1}{2e}$. 故所求参数 $a = \frac{1}{2e}$.

16.13. **解** 第五题

知乎@余命数



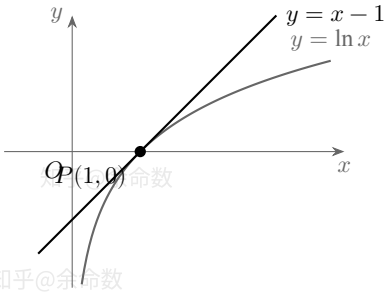
若直线 $y=x+b$ 与曲线 $y=\ln x$ 相切于点 $P(x_0,y_0)$, 则在切点处二者斜率相等. 直线 $y=x+b$ 的斜率恒为 1. 曲线 $y=\ln x$ 在点 x_0 处的斜率已在上一题中通过极限定义求得, 为 $1/x_0$. 令两斜率相等, 得

$$1=\frac{1}{x_0} \implies x_0=1$$

由此确定切点的横坐标为 $x_0=1$. 在该点两曲线的函数值相等, 因此

$$y_0=1+b \quad \text{且} \quad y_0=\ln(1)=0$$

联立可得 $1+b=0$, 解得 $b=-1$. 故当 $b=-1$ 时, 直线 $y=x-1$ 与曲线 $y=\ln x$ 在点 $(1,0)$ 处相切.



16.14. **解** 第六题

法线是与切线在切点处相互垂直的直线. 设切点为 $P(x_0,x_0^2-1)$. 我们先计算抛物线 $y=x^2-1$ 在点 P 处的切线斜率

$$k_{\tan}=\lim_{\Delta x\to 0}\frac{((x_0+\Delta x)^2-1)-(x_0^2-1)}{\Delta x}=\lim_{\Delta x\to 0}\frac{2x_0\Delta x+(\Delta x)^2}{\Delta x}=2x_0$$

当 $x_0\neq 0$ 时, 法线的斜率 $k_{\text{norm}}=-\dfrac{1}{k_{\tan}}=-\dfrac{1}{2x_0}$. 过点 P 的法线方程为

$$y-(x_0^2-1)=-\frac{1}{2x_0}(x-x_0)$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

此法线通过原点 $(0,0)$, 于是其坐标满足

$$0-(x_0^2-1)=-\frac{1}{2x_0}(0-x_0) \implies -(x_0^2-1)=\frac{1}{2}$$

解得 $x_0^2=1/2$, 即 $x_0=\pm\frac{\sqrt{2}}{2}$. 对应的点为 $(\frac{\sqrt{2}}{2},-\frac{1}{2})$ 和 $(-\frac{\sqrt{2}}{2},-\frac{1}{2})$.

当 $x_0=0$ 时, 切线斜率为 0(水平线), 法线为竖直直线 $x=0$. 此法线通过原点. 此时切点为 $(0,-1)$.

故存在三个满足条件的点: $(0,-1)$, $(\frac{\sqrt{2}}{2},-\frac{1}{2})$ 和 $(-\frac{\sqrt{2}}{2},-\frac{1}{2})$.

16.15. 解 第七题

我们从切线的角度证明这个不等式. 先验证直线 $y=x+1$ 是曲线 $y=e^x$ 的切线. 在 $x=0$ 处, $e^0=1$, 直线 $y=0+1=1$, 故点 $(0,1)$ 是公共点. 曲线 $y=e^x$ 在 $x=0$ 的切线斜率(依赖重要极限 $\lim\limits_{h\to0}(e^h-1)/h=1$)为

$$k=\lim_{\Delta x\to0}\frac{e^{0+\Delta x}-e^0}{\Delta x}=\lim_{\Delta x\to0}\frac{e^{\Delta x}-1}{\Delta x}=1$$

直线 $y=x+1$ 的斜率也是 1. 故 $y=x+1$ 是 $y=e^x$ 在点 $(0,1)$ 的切线.

为了证明不等式, 我们构造辅助函数 $g(x)=e^x-(x+1)$. 原不等式等价于证明 $g(x)\geq0$ 恒成立. 我们考察函数 $g(x)$ 在任意点 x 处的斜率

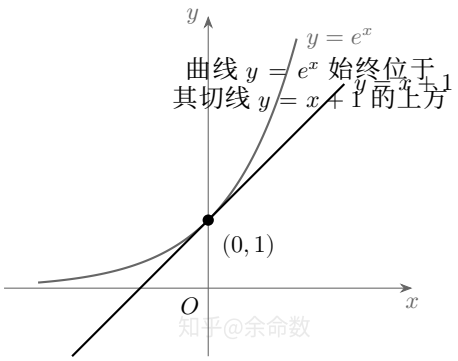
$$\begin{aligned}
k_g(x)&=\lim_{\Delta x\to0}\frac{g(x+\Delta x)-g(x)}{\Delta x}\\
&=\lim_{\Delta x\to0}\frac{(e^{x+\Delta x}-(x+\Delta x+1))-(e^x-(x+1))}{\Delta x}\\
&=\lim_{\Delta x\to0}\frac{e^x(e^{\Delta x}-1)-\Delta x}{\Delta x}\\
&=e^x\left(\lim_{\Delta x\to0}\frac{e^{\Delta x}-1}{\Delta x}\right)-\lim_{\Delta x\to0}\frac{\Delta x}{\Delta x}\\
&=e^x\cdot1-1=e^x-1
\end{aligned}$$

令此斜率为 0, 即 $e^x-1=0$, 解得 $x=0$. 当 $x<0$ 时, $e^x<1$, 故斜率 $k_g(x)<0$, 函数 $g(x)$ 在此区间减小. 当 $x>0$ 时, $e^x>1$, 故斜率 $k_g(x)>0$, 函数 $g(x)$ 在此区间增加.

因此, 函数 $g(x)$ 在 $x=0$ 处取得唯一的全局最小值. 该最小值为 $g(0)=e^0-(0+1)=1-1=0$. 函数的最小值是 0, 于是对任意实数 x, 都有 $g(x)\geq0$, 即 $e^x\geq x+1$. 证毕.

16.16. 解 第八题

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



设 $P(x_0, 1/x_0)$ 为双曲线上 $x > 0$ 分支的任意一点. 我们先计算曲线 $y = 1/x$ 在点 P 处的切线斜率, 得

$$k = \lim_{\Delta x \to 0} \frac{\frac{1}{x_0+\Delta x} - \frac{1}{x_0}}{\Delta x} = \lim_{\Delta x \to 0} \frac{x_0 - (x_0 + \Delta x)}{x_0(x_0 + \Delta x)\Delta x}$$
$$= \lim_{\Delta x \to 0} \frac{-\Delta x}{x_0(x_0 + \Delta x)\Delta x} = -\frac{1}{x_0^2}$$

过点 P 的切线方程为

$$y - \frac{1}{x_0} = -\frac{1}{x_0^2}(x - x_0)$$

令 $x = 0$, 得 y 轴截距 $y = \frac{2}{x_0}$. 截点为 $A(0, 2/x_0)$. 令 $y = 0$, 得 x 轴截距 $x = 2x_0$. 截点为 $B(2x_0, 0)$.

所求三角形是由原点 $O(0, 0)$、点 A 和点 B 构成的直角三角形. 其面积为

$$\text{Area} = \frac{1}{2} \cdot |OA| \cdot |OB| = \frac{1}{2} \cdot \left|\frac{2}{x_0}\right| \cdot |2x_0|$$

由于 $x_0 > 0$(第一象限),

$$\text{Area} = \frac{1}{2} \cdot \frac{2}{x_0} \cdot 2x_0 = 2$$

此面积值是一个常数 2, 与切点 x_0 的选取无关. 证毕.

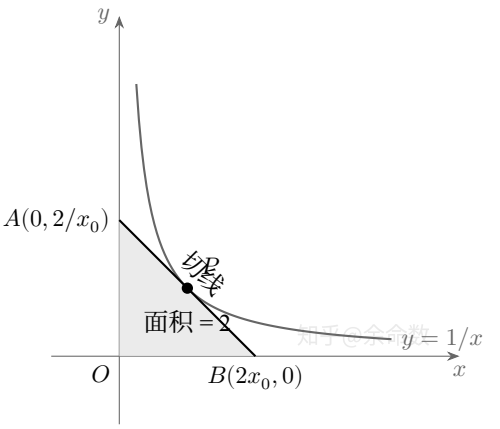
16.17. **解** 第 1 题

当 $x \to 2$ 时, 分母 $x^2 - 4 \to 0$, 分子 $x^3 - x^2 - x - 2 \to 2^3 - 2^2 - 2 - 2 = 0$.

这是一个 $\frac{0}{0}$ 型不定式极限. 因此分子与分母都含有因子 $(x - 2)$, 需用因式分解消去这个不定形式.

分母的因式分解是直接的.$x^2 - 4 = (x - 2)(x + 2)$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



对于分子, 由于 $x=2$ 是其根, 由多项式除法或综合除法得 $x^3-x^2-x-2=(x-2)(x^2+x+1)$. 于是

$$\lim_{x\to 2}\frac{(x-2)(x^2+x+1)}{(x-2)(x+2)}.$$

当 $x\to 2$ 时 $x\neq 2$, 故可约去公因子 $(x-2)$, 得到 $\lim\limits_{x\to 2}\dfrac{x^2+x+1}{x+2}$. 此时分母的极限不再为零, 直接代入 $x=2$ 得 $\dfrac{2^2+2+1}{2+2}=\dfrac{7}{4}$.

16.18. **解** 第 2 题

当 $x\to 1$ 时, 这是一个 $\dfrac{0}{0}$ 型不定式. 分子与分母都含有无理式, 需分别进行有理化.

对于分子的平方根, 我们乘以它的共轭表达式 $\sqrt{x+3}+2$.

对于分母的立方根, 我们利用立方差公式 $a^3-b^3=(a-b)(a^2+ab+b^2)$, 乘以 $(\sqrt[3]{x})^2+\sqrt[3]{x}+1=\sqrt[3]{x^2}+\sqrt[3]{x}+1$.

$$\begin{aligned}\lim_{x\to 1}\frac{\sqrt{x+3}-2}{\sqrt[3]{x}-1}&=\lim_{x\to 1}\left[\frac{\sqrt{x+3}-2}{\sqrt[3]{x}-1}\cdot\frac{\sqrt{x+3}+2}{\sqrt{x+3}+2}\cdot\frac{\sqrt[3]{x^2}+\sqrt[3]{x}+1}{\sqrt[3]{x^2}+\sqrt[3]{x}+1}\right]\\&=\lim_{x\to 1}\left[\frac{(x+3)-4}{(\sqrt[3]{x})^3-1^3}\cdot\frac{\sqrt[3]{x^2}+\sqrt[3]{x}+1}{\sqrt{x+3}+2}\right]\\&=\lim_{x\to 1}\left[\frac{x-1}{x-1}\cdot\frac{\sqrt[3]{x^2}+\sqrt[3]{x}+1}{\sqrt{x+3}+2}\right]\end{aligned}$$

约去因子 $(x-1)$ 后, 直接代入 $x=1$ 求值, 得到

$$\frac{\sqrt[3]{1^2}+\sqrt[3]{1}+1}{\sqrt{1+3}+2}=\frac{1+1+1}{2+2}=\frac{3}{4}$$

16.19. **解** 第 3 题

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

当 $x \to +\infty$ 时, 这是一个 $\infty - \infty$ 型不定式. 我们用分子有理化来转化它.

$$\begin{aligned}
&\lim_{x\to+\infty}(\sqrt{x^2+ax}-\sqrt{x^2+bx}) \\
&\quad=\lim_{x\to+\infty}\frac{(\sqrt{x^2+ax}-\sqrt{x^2+bx})(\sqrt{x^2+ax}+\sqrt{x^2+bx})}{\sqrt{x^2+ax}+\sqrt{x^2+bx}} \\
&\quad=\lim_{x\to+\infty}\frac{(x^2+ax)-(x^2+bx)}{\sqrt{x^2+ax}+\sqrt{x^2+bx}} \\
&\quad=\lim_{x\to+\infty}\frac{(a-b)x}{\sqrt{x^2(1+a/x)}+\sqrt{x^2(1+b/x)}}
\end{aligned}$$

由于 $x \to +\infty$, x 为正, 故 $\sqrt{x^2} = |x| = x$.

$$\begin{aligned}
&=\lim_{x\to+\infty}\frac{(a-b)x}{x\sqrt{1+a/x}+x\sqrt{1+b/x}} \\
&=\lim_{x\to+\infty}\frac{a-b}{\sqrt{1+a/x}+\sqrt{1+b/x}}
\end{aligned}$$

当 $x \to +\infty$ 时, $a/x \to 0$ 且 $b/x \to 0$.

$$\frac{a-b}{\sqrt{1+0}+\sqrt{1+0}}=\frac{a-b}{2}$$

16.20. **解** 第 4 题

这里 $x \to -\infty$, 因此 x 是绝对值很大的负数.

当 $x < 0$ 时, 根据绝对值的定义, $\sqrt{x^2} = |x| = -x$.

我们将分子与分母同时除以 x.

$$\lim_{x\to-\infty}\frac{\sqrt{x^2+1}}{x}=\lim_{x\to-\infty}\frac{\sqrt{x^2(1+1/x^2)}}{x}=\lim_{x\to-\infty}\frac{|x|\sqrt{1+1/x^2}}{x}$$

由于 $x < 0$, $|x| = -x$.

$$\lim_{x\to-\infty}\frac{-x\sqrt{1+1/x^2}}{x}=\lim_{x\to-\infty}-\sqrt{1+\frac{1}{x^2}}$$

当 $x \to -\infty$ 时, $1/x^2 \to 0$.

$$-\sqrt{1+0}=-1$$

16.21. **解** 第 5 题

当 $x \to 2$ 时, 分母 $x - 2 \to 0$.

由于极限值为有限数 5, 分子在 $x \to 2$ 时也必须趋于 0, 否则整个分式的极限将是无穷大.

于是第一个条件是 $\lim\limits_{x\to2}(x^2+ax+b)=0$.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

由于多项式函数是连续的, 直接代入可得

$$2^2 + a(2) + b = 0 \implies 4 + 2a + b = 0 \quad (1)$$

由这个条件解出 $b = -2a - 4$. 将它代回原极限表达式.

$$\begin{aligned}\lim_{x\to 2} \frac{x^2 + ax - 2a - 4}{x - 2} &= \lim_{x\to 2} \frac{(x^2 - 4) + a(x - 2)}{x - 2} \\ &= \lim_{x\to 2} \frac{(x - 2)(x + 2) + a(x - 2)}{x - 2} \\ &= \lim_{x\to 2} (x + 2 + a)\end{aligned}$$

这个极限的值为 $2 + 2 + a = 4 + a$. 根据题设, 它等于 5.

$$4 + a = 5 \implies a = 1$$

将 $a = 1$ 代入 (1) 式, 得到 $4 + 2(1) + b = 0 \implies b = -6$.

因此所求常数为 $a = 1, b = -6$.

16.22. **解** 第 6 题

函数 $f(x)$ 在点 $x = 1$ 处连续的定义是 $\lim\limits_{x\to 1} f(x) = f(1)$.

根据函数的定义, $f(1) = k$. 我们先计算 $\lim\limits_{x\to 1} f(x)$.

$$\lim_{x\to 1} f(x) = \lim_{x\to 1} \frac{\sin(\pi x)}{x - 1}$$

这是一个 $\dfrac{0}{0}$ 型不定式. 为了利用重要极限 $\lim\limits_{u\to 0} \frac{\sin u}{u} = 1$, 我们进行变量代换.

令 $u = x - 1$, 则 $x = u + 1$. 当 $x \to 1$ 时, $u \to 0$.

$$\begin{aligned}\lim_{u\to 0} \frac{\sin(\pi(u + 1))}{u} &= \lim_{u\to 0} \frac{\sin(\pi u + \pi)}{u} \\ &= \lim_{u\to 0} \frac{-\sin(\pi u)}{u}\end{aligned}$$

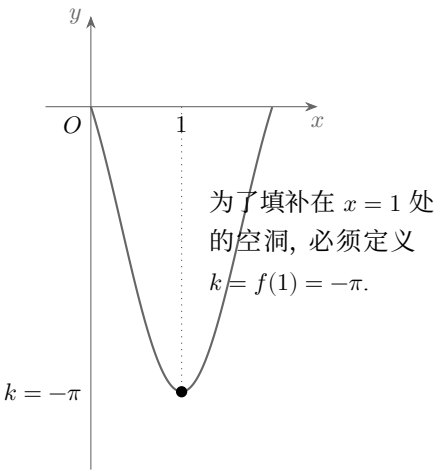
为了凑出重要极限的形式, 我们对分母进行调整.

$$\lim_{u\to 0} \left(-\pi \cdot \frac{\sin(\pi u)}{\pi u} \right) = -\pi \cdot \lim_{u\to 0} \frac{\sin(\pi u)}{\pi u} = -\pi \cdot 1 = -\pi$$

因此 $\lim\limits_{x\to 1} f(x) = -\pi$.

要使函数连续, 必须有 $k = -\pi$.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数



16.23. **解** 第 7 题

由于绝对值函数的存在, 我们分别考察 x 从 1 的左侧和右侧趋近时的极限.

当 $x\to 1^+$ 时, $x>1$, 故 $x-1>0$, 此时 $|x-1|=x-1$.

$$\lim_{x\to 1^+}\frac{|x-1|}{x^2-1}=\lim_{x\to 1^+}\frac{x-1}{(x-1)(x+1)}=\lim_{x\to 1^+}\frac{1}{x+1}=\frac{1}{2}$$

当 $x\to 1^-$ 时, $x<1$, 故 $x-1<0$, 此时 $|x-1|=-(x-1)$.

$$\lim_{x\to 1^-}\frac{|x-1|}{x^2-1}=\lim_{x\to 1^-}\frac{-(x-1)}{(x-1)(x+1)}=\lim_{x\to 1^-}\frac{-1}{x+1}=-\frac{1}{2}$$

由于左极限与右极限不相等,

$$\lim_{x\to 1^-}\frac{|x-1|}{x^2-1}\neq\lim_{x\to 1^+}\frac{|x-1|}{x^2-1}$$

根据极限存在的充要条件, 该极限不存在.

16.24. **解** 第 8 题

我们先对括号内的表达式通分, 把它化为单一的有理分式.

$$\begin{aligned}\frac{x^2+1}{x+1}-ax-b&=\frac{x^2+1-(ax+b)(x+1)}{x+1}\\&=\frac{x^2+1-(ax^2+ax+bx+b)}{x+1}\\&=\frac{(1-a)x^2-(a+b)x+(1-b)}{x+1}\end{aligned}$$

我们要求这个分式在 $x\to+\infty$ 时的极限为 0.

这是有理分式在无穷远处的极限问题, 分母的最高次幂为 1.

要使极限为 0, 分子的次数必须严格小于分母的次数.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

因此分子中 x^2 的系数必须为零.

$$1-a=0 \implies a=1$$

当 $a=1$ 时, 表达式变为 $\dfrac{-(1+b)x+(1-b)}{x+1}$.

此时分子与分母的最高次幂都是 1. 根据有理函数在无穷远处的极限法则, 其极限等于最高次项系数之比.

$$\lim_{x\to+\infty}\frac{-(1+b)x+(1-b)}{x+1}=\frac{-(1+b)}{1}=-1-b$$

知乎@余命数

根据题设, 此极限为 0.

$$-1-b=0 \implies b=-1$$

因此所求常数为 $a=1, b=-1$.

16.25. **解** 第 9 题

我们求这个极限时用到第二个重要极限 $\lim\limits_{u\to0}(1+u)^{1/u}=e$ 及其推论.

利用对数与指数的互逆关系, 我们把表达式重写为

知乎@余命数

$$\lim_{x\to0}\frac{(1+x)^a-1}{x}=\lim_{x\to0}\frac{e^{a\ln(1+x)}-1}{x}$$

这是一个 $\dfrac{0}{0}$ 型不定式.

通过如下恒等变形, 我们凑出已知的极限形式, 得到

知乎@余命数

知乎@余命数

$$\begin{aligned}\lim_{x\to0}\frac{e^{a\ln(1+x)}-1}{x}&=\lim_{x\to0}\left(\frac{e^{a\ln(1+x)}-1}{a\ln(1+x)}\cdot\frac{a\ln(1+x)}{x}\right)\\&=\left(\lim_{x\to0}\frac{e^{a\ln(1+x)}-1}{a\ln(1+x)}\right)\cdot\left(\lim_{x\to0}a\frac{\ln(1+x)}{x}\right)\end{aligned}$$

对于第一个极限, 令 $u=a\ln(1+x)$. 当 $x\to0$ 时, $u\to0$, 因此该极限为 $\lim\limits_{u\to0}\frac{e^u-1}{u}=1$.

对于第二个极限,

$$\begin{aligned}\lim_{x\to0}a\frac{\ln(1+x)}{x}&=a\lim_{x\to0}\ln(1+x)^{1/x}=a\ln\left(\lim_{x\to0}(1+x)^{1/x}\right)\\&=a\ln(e)=a\cdot1=a\end{aligned}$$

因此, 原极限的值为 $1\cdot a=a$.

16.26. **解** 第 10 题

这是一个 $\dfrac{0}{0}$ 型不定式, 涉及三角函数. 我们利用重要极限 $\lim\limits_{x\to0}\frac{\sin x}{x}=1$.

知乎@余命数

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

先对分子进行三角恒等变形.

$$\tan x-\sin x=\frac{\sin x}{\cos x}-\sin x=\sin x\left(\frac{1}{\cos x}-1\right)=\sin x\frac{1-\cos x}{\cos x}$$

利用恒等式 $1-\cos x=2\sin^2\frac{x}{2}$.

$$\tan x-\sin x=\sin x\frac{2\sin^2(x/2)}{\cos x}$$

代回原极限表达式, 得到

$$\begin{aligned}\lim_{x\to0}\frac{\sin x\cdot 2\sin^2(x/2)}{x^3\cos x}&=\lim_{x\to0}\left(\frac{\sin x}{x}\cdot\frac{2}{\cos x}\cdot\frac{\sin^2(x/2)}{x^2}\right)\\&=\left(\lim_{x\to0}\frac{\sin x}{x}\right)\cdot\left(\lim_{x\to0}\frac{2}{\cos x}\right)\cdot\left(\lim_{x\to0}\frac{\sin^2(x/2)}{x^2}\right)\end{aligned}$$

第一个极限为 1, 第二个极限为 $\frac{2}{1}=2$.

对于第三个极限, 我们凑形式如下.

$$\lim_{x\to0}\frac{\sin^2(x/2)}{x^2}=\lim_{x\to0}\frac{\sin^2(x/2)}{4(x/2)^2}=\frac{1}{4}\lim_{x\to0}\left(\frac{\sin(x/2)}{x/2}\right)^2=\frac{1}{4}\cdot1^2=\frac{1}{4}$$

将各部分结果相乘, 得到

$$\text{原极限}=1\cdot2\cdot\frac{1}{4}=\frac{1}{2}$$

16.27. **解** 第 11 题

函数 $f(x)=\sqrt{x^2+3x}$ 的定义域由 $x^2+3x\geq0$ 给出,即 $x\in(-\infty,-3]\cup[0,+\infty)$.

垂直渐近线. 函数在其定义域内连续, 因此不存在垂直渐近线.

水平或斜渐近线. 我们分别考察 $x\to+\infty$ 和 $x\to-\infty$ 两种情况.

当 $x\to+\infty$ 时. 我们先计算斜渐近线 $y=mx+b$ 的斜率 m, 得到

$$\begin{aligned}m&=\lim_{x\to+\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}\frac{\sqrt{x^2+3x}}{x}\\&=\lim_{x\to+\infty}\frac{\sqrt{x^2(1+3/x)}}{x}=\lim_{x\to+\infty}\frac{x\sqrt{1+3/x}}{x}=1\end{aligned}$$

接着计算截距 b, 得到

$$\begin{aligned}b&=\lim_{x\to+\infty}(f(x)-mx)=\lim_{x\to+\infty}(\sqrt{x^2+3x}-x)\\&=\lim_{x\to+\infty}\frac{(\sqrt{x^2+3x}-x)(\sqrt{x^2+3x}+x)}{\sqrt{x^2+3x}+x}\\&=\lim_{x\to+\infty}\frac{x^2+3x-x^2}{\sqrt{x^2+3x}+x}\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

$$= \lim_{x\to+\infty} \frac{3x}{x\sqrt{1+3/x}+x} = \lim_{x\to+\infty} \frac{3}{\sqrt{1+3/x}+1} = \frac{3}{2}$$

因此, 当 $x \to +\infty$ 时, 函数有一条斜渐近线 $y = x + \frac{3}{2}$.

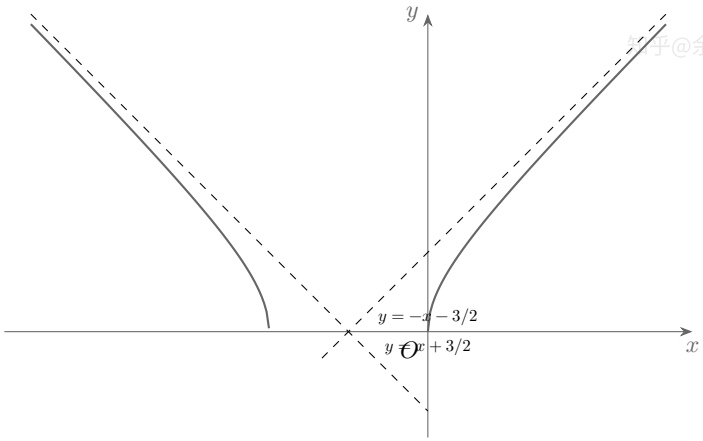
当 $x \to -\infty$ 时.

$$m = \lim_{x\to-\infty} \frac{f(x)}{x} = \lim_{x\to-\infty} \frac{\sqrt{x^2+3x}}{x}$$
$$= \lim_{x\to-\infty} \frac{|x|\sqrt{1+3/x}}{x} = \lim_{x\to-\infty} \frac{-x\sqrt{1+3/x}}{x} = -1$$

$$b = \lim_{x\to-\infty} (f(x) - mx) = \lim_{x\to-\infty} (\sqrt{x^2+3x} - (-1)x)$$
$$= \lim_{x\to-\infty} (\sqrt{x^2+3x} + x)$$
$$= \lim_{x\to-\infty} \frac{(\sqrt{x^2+3x}+x)(\sqrt{x^2+3x}-x)}{\sqrt{x^2+3x}-x}$$
$$= \lim_{x\to-\infty} \frac{x^2+3x-x^2}{\sqrt{x^2(1+3/x)}-x}$$
$$= \lim_{x\to-\infty} \frac{3x}{|x|\sqrt{1+3/x}-x} = \lim_{x\to-\infty} \frac{3x}{-x\sqrt{1+3/x}-x}$$
$$= \lim_{x\to-\infty} \frac{3}{-\sqrt{1+3/x}-1} = \frac{3}{-1-1} = -\frac{3}{2}$$

因此, 当 $x \to -\infty$ 时, 函数有一条斜渐近线 $y = -x - \frac{3}{2}$.

故函数有两条渐近线.$y = x + \frac{3}{2}$ 和 $y = -x - \frac{3}{2}$.



16.28. **解** 第 12 题

我们计算 $\lim\limits_{x\to0} \dfrac{f(x^2)}{\sin^2 x}$.

这是一个 $\dfrac{0}{0}$ 型不定式. 我们通过代数变形, 凑出已知的极限形式 $\lim\limits_{u\to0} \frac{f(u)}{u} = L$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

我们把表达式改写为

$$\frac{f(x^2)}{\sin^2 x}=\frac{f(x^2)}{x^2}\cdot\frac{x^2}{\sin^2 x}=\frac{f(x^2)}{x^2}\cdot\left(\frac{x}{\sin x}\right)^2$$

根据极限的乘法法则,我们可以分别计算每一部分的极限.

对于第一部分 $\lim\limits_{x\to0}\dfrac{f(x^2)}{x^2}$,我们作变量代换. 令 $u=x^2$. 当 $x\to0$ 时,有 $u\to0$.

$$\lim_{x\to0}\frac{f(x^2)}{x^2}=\lim_{u\to0}\frac{f(u)}{u}=L$$

对于第二部分 $\lim\limits_{x\to0}\left(\dfrac{x}{\sin x}\right)^2$,我们利用重要极限.

$$\lim_{x\to0}\left(\frac{x}{\sin x}\right)^2=\left(\lim_{x\to0}\frac{x}{\sin x}\right)^2=\left(\frac{1}{\lim\limits_{x\to0}\frac{\sin x}{x}}\right)^2=\left(\frac{1}{1}\right)^2=1$$

将两部分极限值相乘,得到最终结果.

$$\lim_{x\to0}\frac{f(x^2)}{\sin^2 x}=L\cdot1=L$$

16.29. **解** 第 13 题

函数 $f(x)$ 用极限定义,其具体表达式依赖于 x 的取值. 我们分情况讨论 $|x|$ 与 1 的关系.

当 $|x|<1$ 时. 当 $n\to\infty$ 时,$x^{2n}\to0$ 且 $x^{2n-1}\to0$.

$$f(x)=\lim_{n\to\infty}\frac{x^{2n-1}+ax}{x^{2n}+1}=\frac{0+ax}{0+1}=ax$$

当 $|x|>1$ 时.$x^{2n}\to\infty$. 我们将分子与分母同除以最高次项 x^{2n}.

$$f(x)=\lim_{n\to\infty}\frac{\frac{x^{2n-1}}{x^{2n}}+\frac{ax}{x^{2n}}}{\frac{x^{2n}}{x^{2n}}+\frac{1}{x^{2n}}}=\lim_{n\to\infty}\frac{\frac{1}{x}+\frac{ax}{x^{2n}}}{1+\frac{1}{x^{2n}}}=\frac{\frac{1}{x}+0}{1+0}=\frac{1}{x}$$

当 $x=1$ 时,

$$f(1)=\lim_{n\to\infty}\frac{1^{2n-1}+a(1)}{1^{2n}+1}=\frac{1+a}{1+1}=\frac{1+a}{2}$$

当 $x=-1$ 时,

$$f(-1)=\lim_{n\to\infty}\frac{(-1)^{2n-1}+a(-1)}{(-1)^{2n}+1}=\frac{-1-a}{1+1}=\frac{-1-a}{2}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

于是得到 $f(x)$ 的分段表达式.

$$f(x)=\begin{cases}1/x, & x<-1\\ (-1-a)/2, & x=-1\\ ax, & -1<x<1\\ (1+a)/2, & x=1\\ 1/x, & x>1\end{cases}$$

要使 $f(x)$ 为连续函数, $f(x)$ 必须在分界点 $x=1$ 和 $x=-1$ 处连续.

在 $x=1$ 处连续, 要求左极限、右极限与函数值相等.

$$\lim_{x\to 1^-} f(x)=\lim_{x\to 1^-}(ax)=a$$

$$\lim_{x\to 1^+} f(x)=\lim_{x\to 1^+}(1/x)=1$$

$$f(1)=(1+a)/2$$

因此必须有 $a=1=(1+a)/2$. 解得 $a=1$.

在 $x=-1$ 处连续, 要求左极限、右极限与函数值相等.

$$\lim_{x\to -1^-} f(x)=\lim_{x\to -1^-}(1/x)=-1$$

$$\lim_{x\to -1^+} f(x)=\lim_{x\to -1^+}(ax)=-a$$

$$f(-1)=(-1-a)/2$$

因此必须有 $-1=-a=(-1-a)/2$. 解得 $a=1$.

两个条件给出相同的 a 值. 因此当 $a=1$ 时, $f(x)$ 是连续函数.

16.30. **解** 第 14 题

我们用零点存在性定理证明.

设 $f(x)=x^3-3x+1$. 这是多项式函数, 在实数域上连续, 特别地在闭区间 $[0,1]$ 上连续. 其端点值为

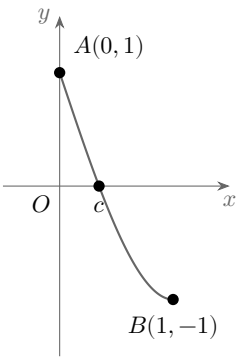
$$f(0)=0^3-3(0)+1=1,\qquad f(1)=1^3-3(1)+1=-1,$$

二者异号($f(0)>0$ 且 $f(1)<0$).

根据零点存在性定理, 若一个函数在闭区间上连续, 且在端点处函数值异号, 则其图像必在该开区间内至少穿越 x 轴一次. 因此在开区间 $(0,1)$ 内至少存在一个实数 c, 使得 $f(c)=0$.

于是方程 $x^3-3x+1=0$ 在区间 $(0,1)$ 内至少存在一个实数根.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

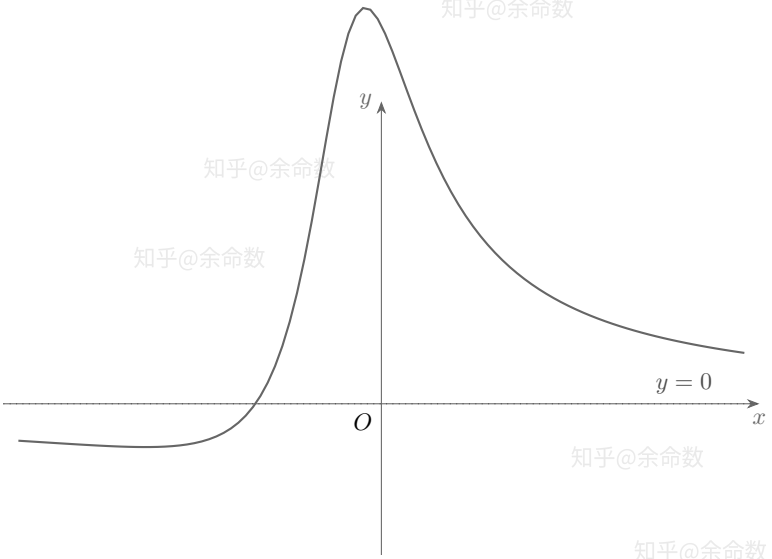


16.31. **解** 第 1 题

函数 $f(x)=\dfrac{3x+5}{x^2+x+1}$ 的定义域为 $\mathbb{R}$,因为分母 $x^2+x+1=(x+1/2)^2+3/4>0$ 恒成立. 考察 $x\to+\infty$ 时的极限. 分子的最高次幂为 1, 分母的最高次幂为 2. 由于分子次数严格小于分母次数, 故极限为 0.

$$\lim_{x\to+\infty}\frac{3x+5}{x^2+x+1}=\lim_{x\to+\infty}\frac{3/x+5/x^2}{1+1/x+1/x^2}=\frac{0+0}{1+0+0}=0$$

即直线 $y=0$, 即 x 轴是函数图像的水平渐近线.



16.32. **解** 第 2 题

函数 $f(x)=\dfrac{2x^2-3x}{x^2-1}$ 的定义域为 $x\neq\pm1$.

水平渐近线: 考察 $x\to+\infty$. 分子与分母的最高次幂均为 2. 极限值为最高次项系数之比.

$$\lim_{x\to+\infty}\frac{2x^2-3x}{x^2-1}=\lim_{x\to+\infty}\frac{2-3/x}{1-1/x^2}=2$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $y=2$ 为水平渐近线.

垂直渐近线: 考察分母为零的点 $x=\pm1$.

$$\lim_{x\to1^+}\frac{x(2x-3)}{(x-1)(x+1)}=\frac{1(-1)}{0^+\cdot2}=-\infty$$

$$\lim_{x\to1^-}\frac{x(2x-3)}{(x-1)(x+1)}=\frac{1(-1)}{0^-\cdot2}=+\infty$$

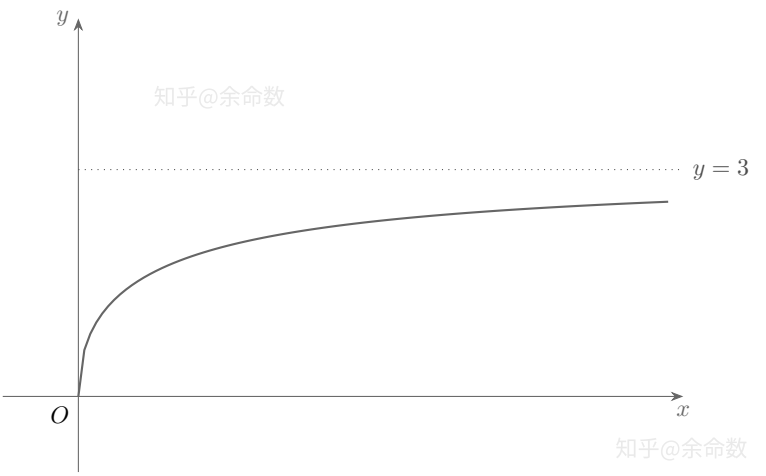
故 $x=1$ 为垂直渐近线. 同理可分析 $x=-1$ 亦为垂直渐近线.

16.33. **解** 第 3 题

当 $x\to+\infty$ 时, 这是一个 $\infty-\infty$ 型的不定式. 采用分子有理化的方法.

$$\begin{aligned}\lim_{x\to+\infty}(\sqrt{x^2+6x}-x)&=\lim_{x\to+\infty}\frac{(\sqrt{x^2+6x}-x)(\sqrt{x^2+6x}+x)}{\sqrt{x^2+6x}+x}\\&=\lim_{x\to+\infty}\frac{(x^2+6x)-x^2}{\sqrt{x^2(1+6/x)}+x}\\&=\lim_{x\to+\infty}\frac{6x}{x\sqrt{1+6/x}+x}=\lim_{x\to+\infty}\frac{6}{\sqrt{1+6/x}+1}\\&=\frac{6}{\sqrt{1+0}+1}=\frac{6}{2}=3\end{aligned}$$

此极限的几何意义是, 当 x 趋于正无穷时, 函数 $f(x)$ 的图像无限趋近于直线 $y=3$. 故 $y=3$ 是函数在正无穷方向的水平渐近线.



16.34. **解** 第 4 题

当 $x\to-\infty$ 时, 这依然是 $\infty-\infty$ 型的不定式. 令 $t=-x$, 则当 $x\to-\infty$ 时, $t\to+\infty$.

$$\lim_{x\to-\infty}(\sqrt{x^2+6x}+x)=\lim_{t\to+\infty}(\sqrt{(-t)^2+6(-t)}+(-t))$$

知乎关注 @ 余命数谢谢喵, 有帮助的话去爱发电支持一下谢谢喵

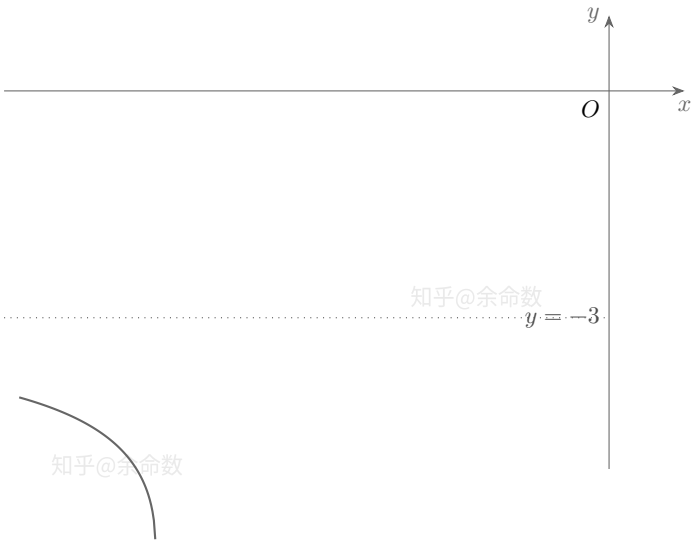
$$= \lim_{t\to+\infty}(\sqrt{t^2-6t}-t)$$

接着对此 t 的极限表达式进行有理化.

$$\begin{aligned}\lim_{t\to+\infty}\frac{(t^2-6t)-t^2}{\sqrt{t^2-6t}+t} &= \lim_{t\to+\infty}\frac{-6t}{\sqrt{t^2(1-6/t)}+t}\\ &= \lim_{t\to+\infty}\frac{-6t}{t\sqrt{1-6/t}+t} = \lim_{t\to+\infty}\frac{-6}{\sqrt{1-6/t}+1}\\ &= \frac{-6}{\sqrt{1-0}+1} = \frac{-6}{2} = -3\end{aligned}$$

@余命数 知乎@余命数

此极限的几何意义是，当 x 趋于负无穷时，函数 $f(x)$ 的图像无限趋近于直线 $y=-3$. 故 $y=-3$ 是函数在负无穷方向的水平渐近线.



知乎@余命数

16.35. **解** 第 5 题

函数 $f(x)=\dfrac{x^3}{(x-1)^2}$ 定义域为 $x\neq 1$. **垂直渐近线**: $\lim\limits_{x\to1}\dfrac{x^3}{(x-1)^2}=\dfrac{1}{0^+}=+\infty$. 故 $x=1$ 是垂直渐近线.

斜渐近线：分子次数比分母次数高一，存在斜渐近线 $y=mx+b$. $m=\lim\limits_{x\to+\infty}\dfrac{f(x)}{x}=\lim\limits_{x\to+\infty}\dfrac{x^2}{(x-1)^2}=1$.

$$\begin{aligned}b &= \lim_{x\to+\infty}(f(x)-mx) = \lim_{x\to+\infty}\left(\frac{x^3}{(x-1)^2}-x\right)\\ &= \lim_{x\to+\infty}\frac{x^3-x(x-1)^2}{(x-1)^2} = \lim_{x\to+\infty}\frac{x^3-x(x^2-2x+1)}{(x-1)^2}\\ &= \lim_{x\to+\infty}\frac{2x^2-x}{x^2-2x+1} = 2\end{aligned}$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故斜渐近线为 $y=x+2$.

16.36. **解** 第 6 题

函数 $f(x)=\dfrac{x^2+\cos x}{2x^2-\sin x}$ 的分子分母中均含有振荡项. 然而, 当 $x\to+\infty$ 时, x^2 是无穷大, 而 $\cos x$ 与 $\sin x$ 均是有界函数. 将分子分母同除以最高次项 x^2.

$$\lim_{x\to+\infty}\frac{x^2+\cos x}{2x^2-\sin x}=\lim_{x\to+\infty}\frac{1+\frac{\cos x}{x^2}}{2-\frac{\sin x}{x^2}}$$

根据夹逼定理, 因为 $-1\leq\cos x\leq 1$, 有 $-\frac{1}{x^2}\leq\frac{\cos x}{x^2}\leq\frac{1}{x^2}$. 当 $x\to+\infty$ 时, $\lim\limits_{x\to+\infty}(\pm\frac{1}{x^2})=0$, 故 $\lim\limits_{x\to+\infty}\frac{\cos x}{x^2}=0$. 同理 $\lim\limits_{x\to+\infty}\frac{\sin x}{x^2}=0$. 因此, 原极限为.

$$\frac{1+0}{2-0}=\frac{1}{2}$$

函数有水平渐近线 $y=\frac{1}{2}$.

16.37. **解** 第 7 题

当 $x\to+\infty$ 时: $e^{-x}\to 0$. 将分子分母同除以 e^x.

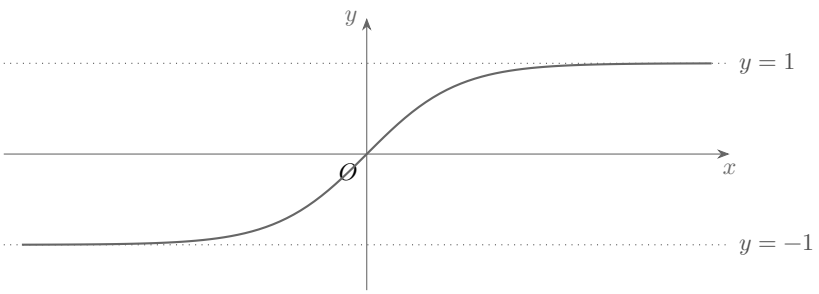
$$\lim_{x\to+\infty}\frac{e^x-e^{-x}}{e^x+e^{-x}}=\lim_{x\to+\infty}\frac{1-e^{-2x}}{1+e^{-2x}}=\frac{1-0}{1+0}=1$$

故 $y=1$ 为正无穷方向的水平渐近线.

当 $x\to-\infty$ 时: $e^{x}\to 0$. 将分子分母同除以 e^{-x}.

$$\lim_{x\to-\infty}\frac{e^x-e^{-x}}{e^x+e^{-x}}=\lim_{x\to-\infty}\frac{e^{2x}-1}{e^{2x}+1}=\frac{0-1}{0+1}=-1$$

故 $y=-1$ 为负无穷方向的水平渐近线.



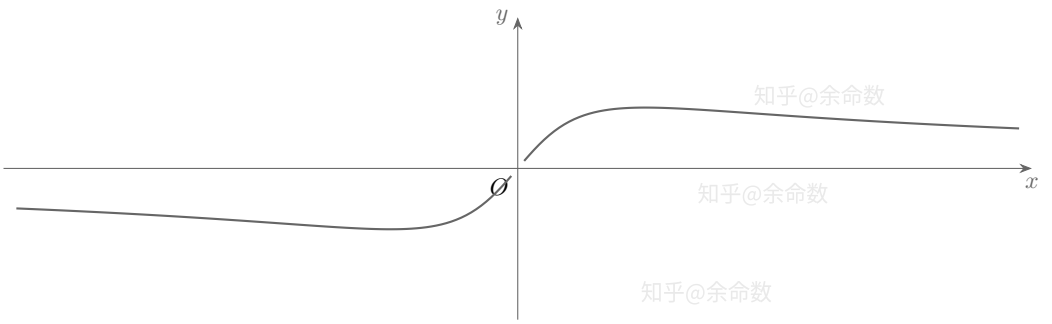
16.38. **解** 第 8 题

函数 $f(x)=\dfrac{\ln(x^2+1)}{x}$ 定义域为 $x\neq 0$. 考察 $x\to+\infty$ 时的极限. 这是一个 $\dfrac{\infty}{\infty}$ 型不定式. 根据函数增长阶梯, 幂函数 x 的增长速度快于对数函数 $\ln(x^2+1)$. 因此, 极限值为 0.

$$\lim_{x\to+\infty}\frac{\ln(x^2+1)}{x}=0$$

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

故 $y=0$ 是函数的水平渐近线.



16.39. **解** 第 9 题

考察 $x\to+\infty$ 时的极限, 以确定水平渐近线.

$$\lim_{x\to+\infty} x(e^{1/x}-1)$$

令 $t=1/x$, 当 $x\to+\infty$ 时, $t\to 0$.

$$\lim_{t\to 0}\frac{1}{t}(e^t-1)=\lim_{t\to 0}\frac{e^t-1}{t}$$

这是一个重要的极限, 其值为 1. 故函数存在水平渐近线 $y=1$.

16.40. **解** 第 10 题

此极限是重要极限 $\lim\limits_{u\to\infty}(1+\frac{1}{u})^u=e$ 的变形. 可将其重写为.

$$\begin{aligned}\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^{2x}&=\lim_{x\to+\infty}\left[\left(1+\frac{1}{x}\right)^{x}\right]^2\\&=\left[\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^{x}\right]^2\\&=e^2\end{aligned}$$

故函数存在水平渐近线 $y=e^2\approx 7.389$.

16.41. **解** 第 11 题

函数 $f(x)=x+\arctan x$ 的渐近行为在正负无穷方向不同.

当 $x\to+\infty$ 时: $m=\lim\limits_{x\to+\infty}\frac{x+\arctan x}{x}=\lim\limits_{x\to+\infty}(1+\frac{\arctan x}{x})=1+0=1$. $b=\lim\limits_{x\to+\infty}(f(x)-mx)=\lim\limits_{x\to+\infty}(x+\arctan x-x)=\lim\limits_{x\to+\infty}\arctan x=\frac{\pi}{2}$. 故有斜渐近线 $y=x+\frac{\pi}{2}$.

当 $x\to-\infty$ 时: $m=\lim\limits_{x\to-\infty}\frac{x+\arctan x}{x}=1$. $b=\lim\limits_{x\to-\infty}(x+\arctan x-x)=\lim\limits_{x\to-\infty}\arctan x=-\frac{\pi}{2}$. 故有斜渐近线 $y=x-\frac{\pi}{2}$.

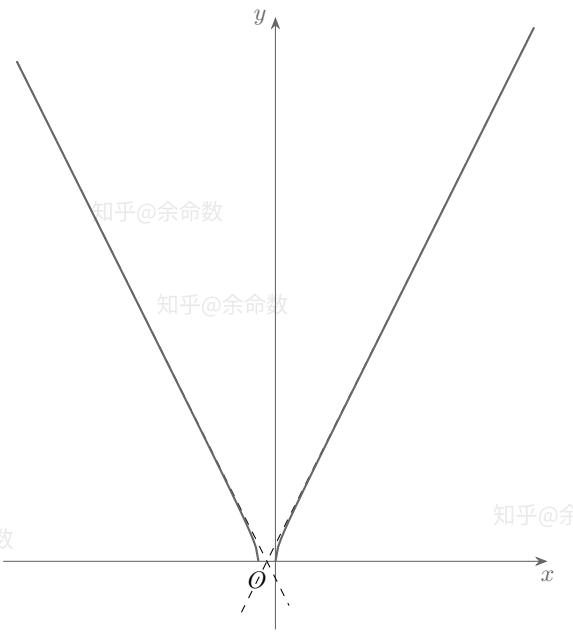
16.42. **解** 第 12 题

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

函数 $f(x)=\sqrt{4x^2+x}$ 定义域为 $x(4x+1)\geq 0$, 即 $(-\infty,-1/4]\cup[0,\infty)$.

当 $x\to+\infty$ 时: $m=\lim\limits_{x\to+\infty}\frac{\sqrt{4x^2+x}}{x}=\lim\limits_{x\to+\infty}\frac{x\sqrt{4+1/x}}{x}=\sqrt{4}=2$. $b=\lim\limits_{x\to+\infty}(\sqrt{4x^2+x}-2x)=\lim\limits_{x\to+\infty}\frac{4x^2+x-4x^2}{\sqrt{4x^2+x}+2x}=\lim\limits_{x\to+\infty}\frac{x}{x(\sqrt{4+1/x}+2)}=\frac{1}{4}$. 故斜渐近线为 $y=2x+\frac{1}{4}$.

当 $x\to-\infty$ 时: $m=\lim\limits_{x\to-\infty}\frac{\sqrt{4x^2+x}}{x}=\lim\limits_{x\to-\infty}\frac{|x|\sqrt{4+1/x}}{x}=-2$. $b=\lim\limits_{x\to-\infty}(\sqrt{4x^2+x}+2x)$. 令 $t=-x, t\to+\infty$. $b=\lim\limits_{t\to+\infty}(\sqrt{4t^2-t}-2t)=\lim\limits_{t\to+\infty}\frac{4t^2-t-4t^2}{\sqrt{4t^2-t}+2t}=\lim\limits_{t\to+\infty}\frac{-t}{t(\sqrt{4-1/t}+2)}=-\frac{1}{4}$. 故斜渐近线为 $y=-2x-\frac{1}{4}$.



16.43. **解** 第 13 题

分析极限 $\lim\limits_{x\to+\infty}\left(\sqrt{4x^2+bx}-ax+1\right)=3$. 这是一个 $\infty-\infty$ 型不定式.

首先, 为使极限为一个有限数, 两部分的无穷大必须是同阶的, 且能够相互抵消. $\sqrt{4x^2+bx}\sim\sqrt{4x^2}=2x$(当 $x\to+\infty$). $ax-1\sim ax$. 为使它们能够抵消, 必须有 $ax=2x$, 故 $a=2$.

当 $a=2$ 时, 极限变为 $\lim\limits_{x\to+\infty}(\sqrt{4x^2+bx}-(2x-1))=3$. 对此表达式进行有理化.

$$\lim_{x\to+\infty}\frac{(4x^2+bx)-(2x-1)^2}{\sqrt{4x^2+bx}+(2x-1)}=\lim_{x\to+\infty}\frac{4x^2+bx-(4x^2-4x+1)}{\sqrt{4x^2+bx}+2x-1}$$

$$=\lim_{x\to+\infty}\frac{(b+4)x-1}{\sqrt{4x^2+bx}+2x-1}$$

知乎@余命数

这是一个有理分式形式,分子分母的最高次幂均为 1. 其极限为最高次项系数之比.

$$\frac{b+4}{\sqrt{4}+2}=\frac{b+4}{4}$$

根据题设, 此极限为 3.

$$\frac{b+4}{4}=3 \implies b+4=12 \implies b=8$$

故所求常数为 $a=2, b=8$.

16.44. **解** 第 14 题

函数表达式依赖于 x 的取值, 需分情况讨论 $|x|$ 与 1 的关系.

当 $|x|<1$ **时**: $\lim\limits_{n\to\infty} x^{2n}=0$ 且 $\lim\limits_{n\to\infty} x^{2n+1}=0$.

$$f(x)=\frac{0+2x}{0+1}=2x$$

当 $|x|>1$ **时**: $\lim\limits_{n\to\infty} x^{2n}=\infty$. 将分子分母同除以 x^{2n}.

$$f(x)=\lim_{n\to\infty}\frac{x+2x/x^{2n}}{1+1/x^{2n}}=\frac{x+0}{1+0}=x$$

当 $x=1$ **时**:

$$f(1)=\lim_{n\to\infty}\frac{1^{2n+1}+2(1)}{1^{2n}+1}=\frac{1+2}{1+1}=\frac{3}{2}$$

当 $x=-1$ **时**:

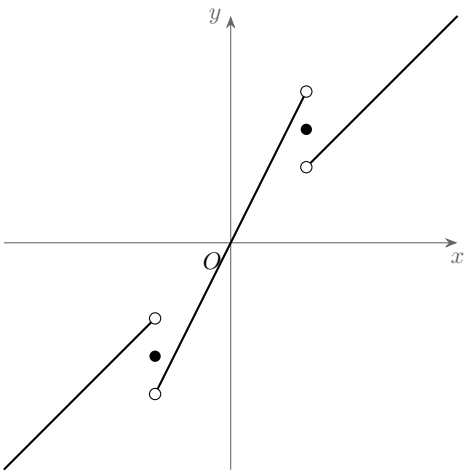
$$f(-1)=\lim_{n\to\infty}\frac{(-1)^{2n+1}+2(-1)}{(-1)^{2n}+1}=\frac{-1-2}{1+1}=-\frac{3}{2}$$

故 $f(x)$ 的分段表达式为.

$$f(x)=\begin{cases}x, & x<-1\\ -3/2, & x=-1\\ 2x, & -1<x<1\\ 3/2, & x=1\\ x, & x>1\end{cases}$$

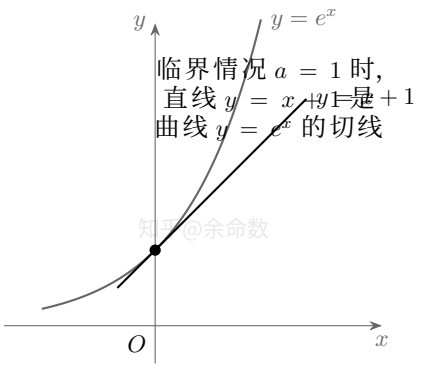
间断点为 $x=1$ 和 $x=-1$.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



16.45. **解** 先题

原不等式等价于 $e^x-ax-1\geq 0$ 在 $[0,+\infty)$ 上恒成立. 构造 $g(x)=e^x-ax-1$, 则 $g(0)=0$, 属端点值为零的恒成立问题. $g'(x)=e^x-a$, $g''(x)=e^x>0$ 于 $(0,+\infty)$, 故 g' 在 $[0,+\infty)$ 上严格递增, 端点分析法适用. 于是 $g\geq 0$ 恒成立当且仅当 $g'(0)\geq 0$, 即 $1-a\geq 0$, $a\leq 1$. 所求范围为 $(-\infty,1]$.



16.46. **解** 再题

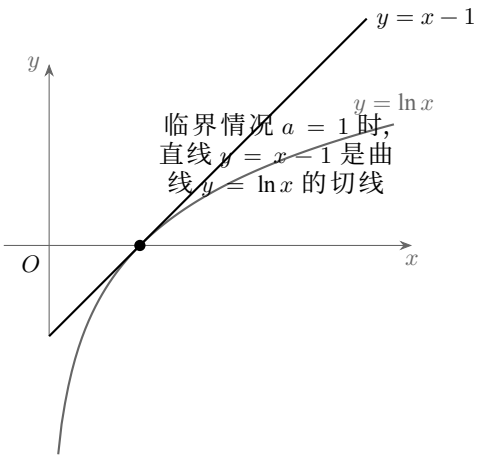
原不等式等价于 $a(x-1)-\ln x\geq 0$ 在 $[1,+\infty)$ 上恒成立. 构造 $g(x)=a(x-1)-\ln x$, 则 $g(1)=0$. $g'(x)=a-\frac{1}{x}$, $g''(x)=\frac{1}{x^2}>0$ 于 $(1,+\infty)$, 故 g' 严格递增, 端点分析法适用. $g\geq 0$ 恒成立当且仅当 $g'(1)=a-1\geq 0$, 即 $a\geq 1$. 所求范围为 $[1,+\infty)$.

16.47. **解** 接着题

原不等式等价于 $ax-\sin x\geq 0$ 在 $[0,\frac{\pi}{2}]$ 上恒成立. 构造 $g(x)=ax-\sin x$, 则 $g(0)=0$. $g'(x)=a-\cos x$, $g''(x)=\sin x>0$ 于 $(0,\frac{\pi}{2})$, 故 g' 严格递增, 端点分析法适用. $g\geq 0$ 恒成立当且仅当 $g'(0)=a-1\geq 0$, 即 $a\geq 1$. 参数分离亦可验证: $x>0$ 时 $a\geq\frac{\sin x}{x}$, 而 $\frac{\sin x}{x}$ 在 $(0,\frac{\pi}{2}]$ 上递减, 上确界为 1, 故同样得 $a\geq 1$. 所求范围为 $[1,+\infty)$.

16.48. **解** 最后题

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



不等式 $f(x) = e^x - \frac{1}{2}x^2 - ax - 1 \geq 0$ 在 $[0, +\infty)$ 上恒成立. 取 $g = f$, 则 $g(0) = 0$、$g'(x) = e^x - x - a$、$g''(x) = e^x - 1 > 0$ 于 $(0, +\infty)$, 故 g' 严格递增, 端点分析法适用. $g \geq 0$ 恒成立当且仅当 $g'(0) = 1 - a \geq 0$, 即 $a \leq 1$. 所求范围为 $(-\infty, 1]$.

16.49. **解** 第五题

原不等式 $x^3 - 3x^2 + ax \geq 0$ 在 $[0, +\infty)$ 上恒成立. 构造 $g(x) = x^3 - 3x^2 + ax$, 则 $g(0) = 0$. $g'(x) = 3x^2 - 6x + a$, $g''(x) = 6(x - 1)$, 在 $(0, 1)$ 上 $g'' < 0$, 在 $(1, +\infty)$ 上 $g'' > 0$, 故 g' 不单调, **端点分析法失效**. $g'(0) = a \geq 0$ 仅为必要条件. 例如 $a = 2$ 时 $g(x) = x(x - 1)(x - 2)$, 在 $(1, 2)$ 上取负, 说明条件不充分.

因 $x \geq 0$, 原不等式等价于 $h(x) = x^2 - 3x + a \geq 0$ 在 $[0, +\infty)$ 上恒成立. h 开口向上, 对称轴 $x = \frac{3}{2}$ 落在区间内, 最小值在顶点处取得,

$$h\Big(\frac{3}{2}\Big) = a - \frac{9}{4}.$$

需 $a - \frac{9}{4} \geq 0$, 即 $a \geq \frac{9}{4}$.

所求实数 a 的取值范围是 $[\frac{9}{4}, +\infty)$.

16.50. **解** 第六题

原不等式等价于 $e^x + ax^2 - x - \frac{1}{2}x^3 - 1 \geq 0$ 在 $[0, +\infty)$ 上恒成立. 构造 $g(x) = e^x + ax^2 - x - \frac{1}{2}x^3 - 1$, 则 $g(0) = 0$, $g'(x) = e^x + 2ax - 1 - \frac{3}{2}x^2$, 且 $g'(0) = 0$ 对一切 a 成立, 必要条件不给出关于 a 的约束. 当 $x > 0$ 时, 不等式等价于

$$a \geq A(x) := \frac{1 + x + \frac{1}{2}x^3 - e^x}{x^2},$$

故 a 须不小于 A 在 $(0, +\infty)$ 上的最大值.

对 $A(x)$ 求导, 得

$$A'(x) = \frac{N(x)}{x^3}, \qquad N(x) = (2 - x)e^x + \frac{1}{2}x^3 - x - 2.$$

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵



知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数

继续求导得

$$N'(x)=(1-x)e^x+\frac{3}{2}x^2-1,\qquad N''(x)=x(3-e^x).$$

因此 $N''(x)>0$ 在 $(0,\ln 3)$ 上成立, $N''(x)<0$ 在 $(\ln 3,+\infty)$ 上成立, 所以 $N'(x)$ 先增后减. 又 $N'(0)=0$、$N'(1)=\frac{1}{2}>0$、$N'(2)=5-e^2<0$, 故 $N'(x)$ 在 $(\ln 3,2)$ 内有唯一零点. 于是 $N(x)$ 先增后减. 结合 $N(0)=0$ 与 $N(2)=0$, 可得 $N(x)>0$ 在 $(0,2)$ 上成立, $N(x)<0$ 在 $(2,+\infty)$ 上成立.

所以 $A(x)$ 在 $(0,2)$ 上递增, 在 $(2,+\infty)$ 上递减, 最大值为

$$A(2)=\frac{7-e^2}{4}.$$

综合充分性与必要性, 所求 a 的取值范围是 $[\frac{7-e^2}{4},+\infty)$.

@余命数 知乎@余命数

第 18 章微分方程初步

18.1. **解** 设 $x(t)=x_0e^{-kt}$.$x(8)=0.3x_0$ 给出 $8k=\ln\frac{10}{3}$. 半衰期 $T_{1/2}=\frac{\ln 2}{k}=\frac{8\ln 2}{\ln(10/3)}\approx\frac{8\times0.6931}{1.2040}\approx 4.6$ 天.

18.2. **解** 温差 $\theta=T-4$ 按指数衰减.$\theta(0)=21$, $\theta(40)=6$, 故 $e^{-40k}=\frac{6}{21}=\frac{2}{7}$. 降到 6 °C 即 $\theta=2$, 需 $e^{-kt}=\frac{2}{21}$, 于是 $t=40\,\frac{\ln(21/2)}{\ln(7/2)}\approx 40\times\frac{2.3499}{1.2528}\approx 75.0$ 分钟, 还需约 35 分钟.

18.3. **解** 由式 (18-11), $T\propto\sqrt{a}$, 故 $\frac{a_2}{a_1}=\left(\frac{T_2}{T_1}\right)^2=\left(\frac{3}{2}\right)^2=2.25$. 摆长应变为原来的 2.25 倍.

知乎@余命数

第 21 章立体几何

21.1. **解** 先题

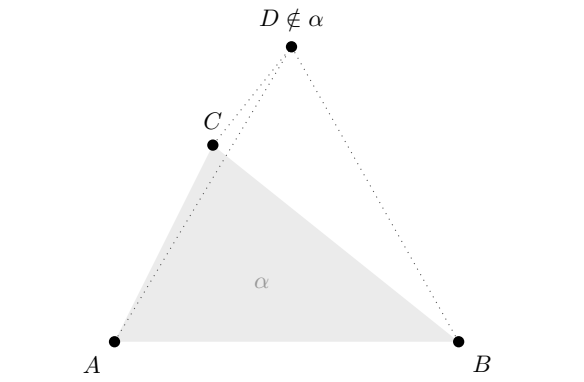
该论断为假. 取不共线的三个点 A,B,C, 据公理 1 其确定唯一一个平面, 记为 α. 在平面 α 之外任取一点 D(这样的点必然存在, 否则空间将退化为平面).

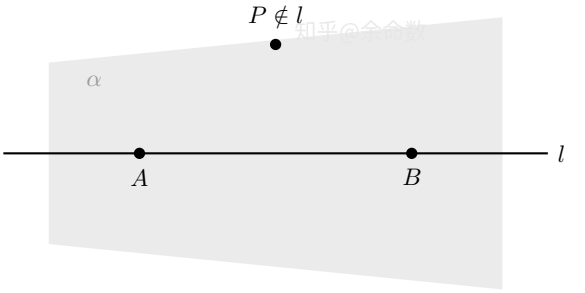
假设四点 A,B,C,D 共面于某个平面 β. 由 A,B,C 均在 β 内且不共线, 据公理 1 的唯一性, 平面 β 必须就是平面 α, 从而点 D 落在平面 α 内. 这与 D 取自 α 之外相矛盾, 故假设不成立. 空间中可以取到四个不共面的点, 它们构成的几何体即四面体.

21.2. **解** 再题

设已知直线为 l, 点为 P, 且 $P\notin l$. 在直线 l 上任取两个不同的点 A,B. 由 P 不在直线 l 上, 三点 P,A,B 不共线.

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵





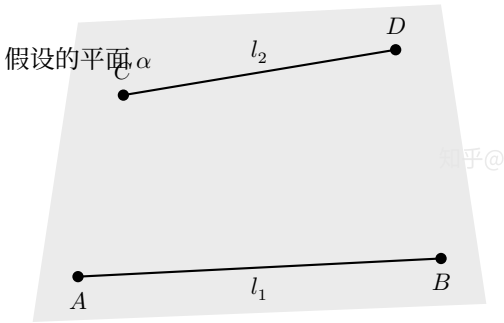
据公理 1, 不共线的三点 P, A, B 确定唯一一个平面 α. 点 P 为确定 α 的三点之一, 故 $P \in \alpha$; 点 A, B 同在直线 l 上, 又是确定 α 的点, 故 $A \in \alpha$ 且 $B \in \alpha$. 据公理 2, 直线 l 落在平面 α 内. 存在性得证.

任一包含直线 l 与点 P 的平面都必含 A, B, P 三点, 而这三个不共线的点仅能确定一个平面, 故满足条件的平面唯一.

21.3. **解** 接着题

设直线 AB 为 l_1, 直线 CD 为 l_2, 即证不存在任何平面能同时包含 l_1 与 l_2.

用反证法. 假设存在一个平面 α 同时包含 l_1 与 l_2.



由 $l_1 \subset \alpha$ 与点的从属关系, 直线上的所有点都在平面内, 故 $A \in \alpha$ 且 $B \in \alpha$; 同理由 $l_2 \subset \alpha$ 得 $C \in \alpha$ 且 $D \in \alpha$. 于是 A, B, C, D 四点皆在 α 内, 即四点共面.

此结论与题设四点 A, B, C, D 不共面相矛盾, 故假设不成立. 据异面直线的定义, 直线 AB 与 CD 是异面直线.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

21.4. 解 最后题

点 K 既在直线 MN 上, 又在直线 PQ 上. 考察由不共面的四点 A, B, C, D 所确定的两平面 ABC 与 ADC. 据公理 3, 二者必相交于一条直线; 点 A 与点 C 是其公共点, 故交线即直线 AC.

点 M 在棱 AB 上, 点 N 在棱 BC 上, 由 $AB, BC \subset$ 平面ABC 知 $M, N \in$ 平面ABC. 据公理 2, 直线 MN 落在平面 ABC 内, 由 $K \in MN$ 得 $K \in$ 平面ABC.

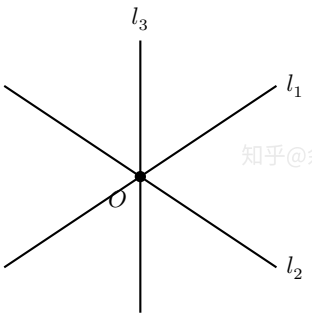
同理, 点 P 在棱 CD 上, 点 Q 在棱 DA 上, 由 $CD, DA \subset$ 平面ADC 知 $P, Q \in$ 平面ADC. 据公理 2, 直线 PQ 落在平面 ADC 内, 由 $K \in PQ$ 得 $K \in$ 平面ADC.

点 K 同属两平面, 即为它们的公共点, 故必落在交线上. 而交线为 AC, 故 K 在直线 AC 上, A, C, K 三点共线.

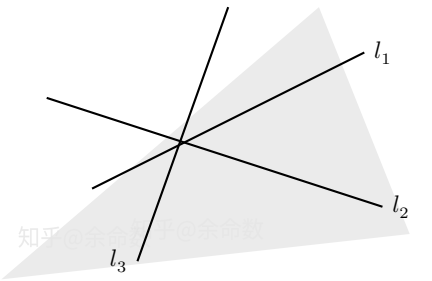
21.5. 解 第五题

设三条直线为 l_1, l_2, l_3, 两两相交, 设 $l_1 \cap l_2 = P_3$、$l_2 \cap l_3 = P_1$、$l_3 \cap l_1 = P_2$.

若 $P_1 = P_2 = P_3$, 设此公共点为 O, 则 l_1, l_2, l_3 皆过点 O, 据定义三直线共点.



共点情形



共面情形

否则三个交点不完全重合, 不失一般性设 $P_1 \neq P_2$. l_1, l_3 是两条相交直线(交于 P_2), 据本节推论确定唯一平面 α, 故 $l_1 \subset \alpha$ 且 $l_3 \subset \alpha$.

点 P_1 是 l_2 与 l_3 的交点, 由 $l_3 \subset \alpha$ 知 $P_1 \in \alpha$; 点 P_3 是 l_2 与 l_1 的交点, 由 $l_1 \subset \alpha$ 知 $P_3 \in \alpha$. 若 $P_1 = P_3$, 则 l_1, l_3 都与 l_2 交于同一点, 化归前一情形; 故此时 P_1, P_3 是 l_2 上两个不同的点. 直线 l_2 上有两点 P_1, P_3 在 α 内, 据公理 2, $l_2 \subset \alpha$.

l_1, l_2, l_3 三条直线皆位于平面 α 内, 据定义三直线共面. 故三条两两相交的直线, 或共点, 或共面.

21.6. 解 第六题

设这四条不同的直线为 l_1, l_2, l_3, l_4. 取其中三条, 例如 l_1, l_2, l_3, 它们两两相交, 据上一题结论, 三者或共点, 或共面. 题设规定任意三条都不共面, 故 l_1, l_2, l_3 只能共点, 设公共点为 O.

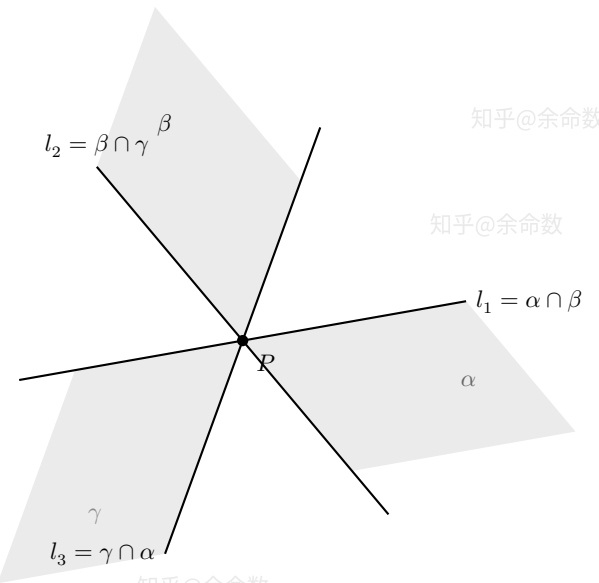
对另取的三条 l_1, l_2, l_4, 它们同样两两相交且不共面, 故亦共点, 公共点即 l_1 与 l_2 的交点 O. 因而 l_4 也经过 O.

l_1, l_2, l_3, l_4 四条直线皆过点 O, 即四直线共点.

21.7. 解 第七题

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

设 $\alpha\cap\beta=l_1$、$\beta\cap\gamma=l_2$、$\gamma\cap\alpha=l_3$. 由 $l_1=\alpha\cap\beta$ 知 $l_1\subset\alpha$, 由 $l_3=\gamma\cap\alpha$ 知 $l_3\subset\alpha$, 故 l_1,l_3 同在平面 α 内, 是共面直线. 题设三交线互不平行, 故 l_1 与 l_3 相交, 设交点为 P.



$P\in l_1$, 而 l_1 是平面 α,β 的交线, 故 $P\in\alpha$ 且 $P\in\beta$; 又 $P\in l_3$, 而 l_3 是平面 γ,α 的交线, 故 $P\in\gamma$ 且 $P\in\alpha$. 合之, 点 P 同时属于平面 β 与平面 γ.

据公理 3, 两个有公共点的平面必交于一条通过该点的直线. 点 P 是平面 β 与 γ 的公共点, 故落在二者的交线即 l_2 上, 即 $P\in l_2$.

l_1 与 l_3 的交点 P 既在 l_2 上, 三条交线 l_1,l_2,l_3 便交于同一点 P, 即三交线共点.

21.8. 解 第八题

设平面 ABC 为 π_1, 平面 $A'B'C'$ 为 π_2, 交线为 l. 即证 P,Q,R 均在 l 上.

对点 $P=AB\cap A'B'$: P 在直线 AB 上, 而 $AB\subset$ 平面ABC, 即 π_1, 故 $P\in\pi_1$; P 又在 $A'B'$ 上, 而 $A'B'\subset$ 平面$A'B'C'$, 即 π_2, 故 $P\in\pi_2$.P 同时属于 π_1,π_2, 故位于其交线 l 上.

至于 P 之存在, 它等价于直线 AB 与 $A'B'$ 相交, 即二者共面且不平行. 考察由点 O,A,B 确定的平面 π_{OAB}: A,A' 都在直线 OA 上, B,B' 都在直线 OB 上, 故 A,B,A',B' 四点共在此平面内, AB 与 $A'B'$ 共面; 只要二者不平行, P 即存在; 题设并不保证这一点(取 $A'=2A$、$B'=2B$、$C'=3C$ 即得 $AB\parallel A'B'$), 以下讨论以 AB 与 $A'B'$ 不平行即 P 为有穷远点为前提.

对点 $Q=BC\cap B'C'$: 由 $Q\in BC\subset\pi_1$ 与 $Q\in B'C'\subset\pi_2$, 得 Q 在交线 l 上; 其存在性由直线 BC 与 $B'C'$ 在平面 OBC 内共面且不平行所保证.

对点 $R=CA\cap C'A'$: 由 $R\in CA\subset\pi_1$ 与 $R\in C'A'\subset\pi_2$, 得 R 在交线 l 上; 其存在性由直线 CA 与 $C'A'$ 在平面 OCA 内共面且不平行所保证.

P,Q,R 三点同在直线 l, 即 π_1 与 π_2 的交线上, 故共线.

知乎关注 @ 余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

第 23 章连分数

23.1. **解** $a_0 = \lfloor\sqrt{6}\rfloor = 2.\sqrt{6}-2$ 取倒数: $\dfrac{1}{\sqrt{6}-2} = \dfrac{\sqrt{6}+2}{2} \approx 2.225$, $a_1 = 2$. 余 $\dfrac{\sqrt{6}+2}{2} - 2 = \dfrac{\sqrt{6}-2}{2}$, 取倒数 $\dfrac{2}{\sqrt{6}-2} = \sqrt{6}+2 \approx 4.449$, $a_2 = 4$. 余 $\sqrt{6}+2-4 = \sqrt{6}-2$, 取倒数又得 $\dfrac{\sqrt{6}+2}{2}$, 回到 a_1 的起点. 故 $\sqrt{6} = [2;\overline{2,4}]$, 周期为 2.

23.2. **解** 前五个渐近分数为 $\dfrac{1}{1}$、$\dfrac{2}{1}$、$\dfrac{3}{2}$、$\dfrac{5}{3}$、$\dfrac{8}{5}$, 数值约 1.000、2.000、1.500、1.667、1.600.$\varphi \approx 1.618$. 偶数阶 $(k = 0, 2, 4)$: $1.000 < 1.500 < 1.600$, 递增, 均小于 φ; 奇数阶 $(k = 1, 3)$: $2.000 > 1.667$, 递减, 均大于 φ. 二者从两侧夹逼 φ, 符合 定理 23.2.

23.3. **解** 反设 $\alpha = \dfrac{m}{n}$ 为有理数. 由 定理 23.11, 每个满足该不等式的 $\dfrac{a}{b}$ 都是 α 的渐近分数; 而有理数的渐近分数只有有限个(定理 23.1, 其连分数有限), 矛盾.

知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数 知乎@余命数

术语索引

本索引按**篇章模块**归类：先见「基础篇 / 进阶篇 / 拓展篇·章名」，其下为该章术语与关键词；页码指向正文首次或主要出现处。

基础篇·预备知识

- 中国剩余定理 (Chinese remainder theorem), 58
- 勒让德定理 (Legendre's formula), 56
- 取整函数 (Floor function), 56
- 基本对称式 (Elementary symmetric polynomial), 8
- 对换 (Transposition), 7
- 对称式 (Symmetric polynomial), 7
- 对称群 (Symmetric group), 7
- 欧拉函数 (Euler's totient function), 57
- 欧拉定理 (Euler's theorem), 57
- 算术基本定理 (Fundamental theorem of arithmetic), 55
- 素数 (Prime), 55
- 置换 (Permutation), 6
- 群 (Group), 7
- 费马小定理 (Fermat's little theorem), 58
- 轮换式 (Cyclic polynomial), 8
- 高斯函数 (Gauss function), 56

基础篇·集合与映射

- 值域 (Range), 78
- 单射 (Injection), 79
- 双射 (Bijection), 79
- 定义域 (Domain), 78
- 映射 (Mapping), 78
- 满射 (Surjection), 79

基础篇·函数

- Bézout 等式 (Bézout's identity), 214
- Cardano 公式 (Cardano's formula), 221
- Ferrari 降次法 (Ferrari's method), 224
- Lagrange 插值 (Lagrange interpolation), 218
- Vieta 替换 (Vieta's substitution), 222
- 三次方程三角解法 (trigonometric solution of cubic), 223
- 不可约多项式 (irreducible polynomial), 215
- 值域 (Range), 90
- 凹凸性 (Convexity), 161
- 函数 (Function), 89, 180
- 切比雪夫多项式 (Chebyshev polynomial), 219
- 切比雪夫多项式的根 (roots of Chebyshev polynomial), 220
- 初等对称多项式 (elementary symmetric polynomial), 217
- 单射 (Injection), 128
- 单调性 (Monotonicity), 154, 161, 163
- 反函数 (Inverse function), 127
- 因式分解唯一定理 (unique factorization), 215
- 复合函数 (Composite function), 131, 132

知乎关注 @ 余命数谢谢喵，有帮助的话去爱发电支持一下谢谢喵

多项式 (polynomial), 213
奇偶性 (Parity), 161
定义域 (Domain), 89, 161
对数函数 (Logarithmic function), 163, 209
对称多项式基本定理 (fundamental theorem of symmetric polynomials), 217
带余除法 (polynomial division), 214
幂函数 (Power function), 160
指数函数 (Exponential function), 152, 209
数域 (number field), 213
映射 (Mapping), 88
最大公因式 (greatest common divisor), 214
牛顿恒等式 (Newton's identities), 218
范数 (Norm), 199
连续 (Continuity), 182
重因式 (multiple factor), 216
重因式判别 (multiple factor criterion), 216
基础篇·函数方程
Babbage 方程 (Babbage's equation), 244
Jensen 方程 (Jensen's equation), 240
Pexider 方程 (Pexider's equation), 242
对合 (Involution), 243
对数函数 (Logarithmic function), 230
幂函数 (Power function), 230
指数函数 (Exponential function), 230
迭代 (Iteration), 243
迭代根 (Iterative root), 244
基础篇·不等式
Schur 不等式 (Schur's inequality), 325
切比雪夫不等式 (Chebyshev's inequality), 323
嵌入不等式 (Embedding inequality), 322
幂平均不等式 (Power mean inequality), 322
柯西不等式, 254
赫尔德不等式 (Hölder's inequality), 320
闵可夫斯基不等式 (Minkowski inequality), 321
基础篇·三角函数
值域 (Range), 341, 345, 349, 354, 405, 437
函数 (Function), 339, 356
单调性 (Monotonicity), 337, 345, 351, 355, 437
周期性 (Periodicity), 337, 339, 345, 351, 355
奇偶性 (Parity), 337, 345, 351, 355, 437
定义域 (Domain), 341, 345, 349, 354, 437
魏尔斯特拉斯 (Weierstrass), 388
魏尔斯特拉斯代换 (Weierstrass substitution), 388
齐次式 (Homogeneous polynomial), 358, 388
基础篇·数列
单调性 (Monotonicity), 549
子列 (Subsequence), 560
数学归纳法 (Mathematical induction), 444, 463

知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

等差数列 (Arithmetic progression), 461, 478
等比数列 (Geometric progression), 444, 465
基础篇·概率
值域 (Range), 659
函数 (Function), 659
定义域 (Domain), 659
几何篇·平面向量
二重外积公式 (vector triple product), 766
共面判定 (coplanarity), 768
向量外积 (cross product), 764
外积运算律, 765
拉格朗日恒等式 (Lagrange's identity), 766
混合积 (scalar triple product), 767
混合积几何意义, 767
混合积轮换对称性, 768
范数 (Norm), 711
雅可比恒等式 (Jacobi identity), 767
几何篇·解三角形
九点圆 (Nine-point circle), 850
余弦定理 (Law of cosines), 773, 777, 797
塞瓦定理 (Ceva's theorem), 844
张角定理, 845
托勒密不等式 (Ptolemy's inequality), 853
托勒密定理 (Ptolemy's theorem), 852
斯特瓦尔特定理 (Stewart's theorem), 847
旁心 (Excenter), 849
梅涅劳斯定理 (Menelaus' theorem), 843
欧拉公式 (Euler's formula), 851
正弦定理 (Law of sines), 773, 777, 797
西姆松定理 (Simson's theorem), 854
西姆松线 (Simson line), 855
费马点 (Fermat point), 849
几何篇·解析几何
共轴圆 (Coaxial circles), 1077
内心 (Incenter), 1071
切割线定理 (Secant-tangent theorem), 1149
反演变换 (Inversion), 1150
圆幂 (Power of a point), 1148
圆幂定理 (Power of a point theorem), 1149
垂心 (Orthocenter), 1070
外心 (Circumcenter), 1070
斜坐标 (Oblique coordinates), 1090
旋转设点 (Rotational point-setting), 989
有向距离 (Signed distance), 1060
极点极线, 987
根心 (Radical center), 1076
根轴 (Radical axis), 1075
欧拉线 (Euler line), 1072
法线式 (Normal form), 1059
满足式条件, 983
牛顿线 (Newton line), 1092
直线偶 (Line pair), 1064
相交弦定理 (Intersecting chords theorem), 1149
相关点代入, 988
蝴蝶定理 (Butterfly theorem), 1093
设点法 (Point-setting method), 858
设线法 (Line-setting method), 858
几何篇·射影几何

知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

布里昂桑定理 (Brianchon's theorem), 1222
帕斯卡定理 (Pascal's theorem), 1219
帕普斯定理 (Pappus' theorem), 1217
笛沙格定理 (Desargues' theorem), 1216
透视中心 (Perspective center), 1216
透视轴 (Perspective axis), 1216
分析篇·极限与导数
函数 (Function), 1299
定义域 (Domain), 1277, 1318
对数函数 (Logarithmic function), 1280
导数 (Derivative), 1272, 1290, 1349
幂函数 (Power function), 1311
指数函数 (Exponential function), 1280
极限 (Limit), 1232, 1239, 1270, 1290
欧拉公式 (Euler formula), 1348
泰勒公式 (Taylor formula), 1345
洛必达法则 (L'Hôpital's rule), 1349
连续 (Continuity), 1276
分析篇·积分
不定积分 (Indefinite integral), 1383, 1417
定积分 (Definite integral), 1412
连续 (Continuity), 1386
分析篇·微分方程
函数 (Function), 1442
微分方程 (Differential equation), 1442, 1448
方向场 (Direction field), 1442
积分曲线 (Integral curve), 1442
分析篇·数学思想
映射 (Mapping), 1501
拓展篇·复数
复数 (Complex number), 1519
欧拉公式 (Euler formula), 1531
拓展篇·立体几何
三面角 (Trihedral angle), 1587
三面角余弦定理 (Trihedral cosine theorem), 1588
三面角正弦定理 (Trihedral sine theorem), 1590
拓展篇·矩阵与行列式
矩阵 (Matrix), 1604, 1608
行列式 (Determinant), 1620

知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎@余命数
知乎关注@余命数谢谢喵,有帮助的话去爱发电支持一下谢谢喵

PROJECT AKARI

AKARI

内容简介

《中学数学之旅》把初等数学当作一个完整的体系，从初中运算的起点一路讲到大学数学的门口. 全书五卷按数学自身的生长顺序编排，从代数语言与函数出发，经周期、离散与随机，走过几何主干，拧成极限、导数与积分的分析主干，最后扩到复数、空间与线性结构. 每条结论尽量给出推导；高考与竞赛的主干题材全部覆盖，取舍只看一条，它在数学结构里是否承重.

无论你是准备升学的学生、带学生的老师，还是想把数学重新学一遍的人，这本书都用得上.

本系列分卷

第 1 卷：预备知识、集合与映射、函数、函数方程、不等式
第 2 卷：三角函数、数列、统计、计数原理、概率
第 3 卷：平面向量、解三角形、解析几何、极坐标与参数方程、射影几何
第 4 卷：极限与导数、简单的积分、微分方程初步、数学思想浅谈
第 5 卷：复数、立体几何、矩阵与行列式、连分数、常数 e、π 与 γ、无穷乘积

公开版 • 10 月 7 日

Project Akari