

微积分基础教程

(Calculus Fundamentals)

您的姓名

2025 年 12 月 26 日

版权信息

本书仅提供电子版，固定版式。

版权所有 © 2025 by 您的姓名。保留一切权利。

前言

自牛顿与莱布尼茨独立创设微积分(1660–1680年代)以来,这门“变化的数学”已走过三个半世纪。三百年间,微积分从“计算变化率”演化为“描述一切可微结构”的元语言:数学不再只是工具,它已经从一门关于量的科学,演变为一门关于模式、结构和形式推理的宏大体系。

现代微积分以其严谨性而闻名,但其基石——极限的 $\epsilon - \delta$ 定义和由此构建的标准分析体系,正在向哲学化甚至宗教化方向发展:非直观性、逻辑严重形式化,过度僵化、挑战常识,和现实严重脱节正在演变成数学空想主义。

有鉴于此,格洛克以极度自下而上的方式重新思考整个微积分体系,成功建立可视化的格洛克代数空间,可视化范围从 $(-\infty^\infty, \infty^\infty)$ 每一个数都可见,并将 ∞, ϵ 定义为未解析数,分别表示无穷大和无穷小,突然之间,一切都变得如此简单,并由此带来以下颠覆性的数学结果:

极限被重新定义,简单而直观,极限的计算只是简单的数学计算,洛必达法则不再必要,神秘性已全部消失。

开闭区间被重新描述,无穷小的数值化定义使开区间基本消失,在极限状态下都是闭区间。

连续性、间断点等相关理论已消失,基本只保留描述性概念和常识。

数列与极限的收敛相关内容是为了传统趋近极限的表述,而无穷大和无穷小不再是一个趋近的概念,这部分内容已经没有意义。因此只保留数列的通项公式并作为统一的函数对待。

导数不再是一个莱布尼兹定义的神秘符号 $\frac{d}{dx} F(x)$,现在只是一个重新定义的点 x 位置的点微分的商,即 $dF(x)/dx$,和小学的除法运算没有什么不同,更强调计算结果是瞬时变化率或几何图形上的斜率。



图 1: 另一张图片, 位于右侧



图 2: 示例插图: 函数的几何图形展示

积分被重新定义, 删除积分定理等内容, 重新定义积分为微分的逆运算。具体变化为点积分简称为积分取代不定积分, 保留定积分并重新命名为区间积分对应区间微分。

自然常数 e 被重新定义为 $e = (1 + \epsilon)^\infty$, 清晰而直观, 并拓展为自然常数序列 $e_n = (1 + n\epsilon)^\infty$, 使其不再具有神秘性。

定义和定理中的函数只进行最小化覆盖, 即只包含一个定义域区间, 消除传统定义和定理的繁杂限制条件, 简单而清晰。

然而, 更严肃而重要的思考是, 我们为什么要接受数学基础/数理逻辑/集合论/数学分析这些所谓的数学家小圈子理论, 无论对于普通人还是科研人员既看不懂也没有任何实用价值。回归常识和简单, 我们真的需要数学家小圈子给我们解释为什么 $1 + 1 = 2$ 吗?

考德 · 格洛克

2025 年 11 月

重要提示

请确保您已安装 tcolorbox 宏包。

风险提示

不要在数学环境中使用
换行，请改用 align 或 split 环境。

目录

第一章 格洛克代数空间	1
1.1 初步理解无穷大和无穷小	1
1.2 格洛克代数空间	2
1.3 无穷大和无穷小的定义	4
第二章 函数和极限	9
2.1 理解函数	9
2.2 参数方程	12
2.3 函数的极限	14
2.4 自然常数和自然对数	17
2.5 泰勒展开在极限计算中的应用	20
第三章 点微分和导数	23
3.1 极限变化量与点微分	23
3.2 导数和导数公式	26
3.3 导数的运算法则	31
3.4 参数方程求导	37
3.5 高阶导数	39
3.6 泰勒级数和麦克劳林级数	39
3.7 欧拉公式	46
3.8 微分公式和运算法则	49
第四章 导数和微分的应用	53
4.1 极限计算	53
4.2 函数单调性	54
4.3 函数的极值	56
4.4 函数的渐近线	59
4.5 牛顿法解方程	61

第五章 不定积分和积分方法	65
5.1 不定积分的基本概念	65
5.2 不定积分的主要性质	67
5.3 基本积分公式	68
5.4 积分技术	70
第六章 常微分方程	81
6.1 一阶常微分方程	81
6.2 二阶线性常微分方程	88
第七章 定积分	95
7.1 区间微分和区间积分	95
7.2 定积分的主要性质	98
第八章 定积分的应用	105
8.1 曲线的弧长	105
8.2 曲线之间的面积	109
8.3 极坐标下的面积和弧长	113
8.4 旋转体的体积	117
第九章 积分和积分方法	121
9.1 导数的定义与几何意义	121
第十章 多变量函数	123
10.1 从平面到空间：定义域与可视化	123
10.2 偏导数：固定维度的变化	126
10.3 全微分与切平面	126
10.4 链式法则：变量之间的传递链	127
第十一章 空间曲线	129
11.1 向量微分和曲线弧长	129
11.2 切向量 \mathbf{T} 、法向量 \mathbf{N} 和副法向量 \mathbf{B}	129
11.3 机器学习中的常用函数及其导函数	133
11.4 通量和散度	140
11.5 环量和旋度	141
11.6 格林定理	142

第十二章 向量微分	145
12.1 微分向量分析	145

第一章 格洛克代数空间

本章首先从常识入手理解无穷大和无穷小，并得出基本的运算法则。在此基础上实现格洛克代数空间。最后给出无穷大和无穷小的正式定义和运算（规则）列表。

1.1 初步理解无穷大和无穷小

自微积分诞生以来，如何阐释、理解无穷大和无穷小以及由此引出的函数的极限一直是问题的核心，所以让我们以此为切入点，开始问题的探索。

格洛克：无穷大和无穷小的直观理解

无穷大和无穷小是具有特殊作用的两个数，分别用 ∞, ϵ 表示，对于 ∞ ，所有实数都是它的无穷小；对于 ϵ ，所有实数都是它的无穷大。

由此产生如下运算法则：对于任意实数 C

$$\infty + C = \infty$$

$$\epsilon + C = C$$

这是基于 ∞, ϵ 本身的直观语义得出的。更进一步，基于数的性质，我们得到：

$$C_1\infty + C_2 = C_1 \left(\infty + \frac{C_2}{C_1} \right) = C_1\infty$$

$$C_1\epsilon + C_2 = C_1 \left(\epsilon + \frac{C_2}{C_1} \right) = C_2$$

$$C_1\infty^2 + C_2\infty = \infty(C_1\infty + C_2) = C_1\infty^2$$

$$C_1\epsilon^2 + C_2\epsilon = \epsilon(C_1\epsilon + C_2) = C_2\epsilon$$

这和 200 多年来的传统解释完全不同，甚至相反。传统解释中并不认为无穷大和无穷小是数，因此不能直接进行数学运算，强调它是一个无限趋近的“动态”过程，因此理解起来要困难复杂得多。

格洛克的直观解释如果可以成立，会使微积分学变得简单而清晰，这就需要建立一个直观而简单的可视化代数模型来阐释其合理性。

1.2 格洛克代数空间

格洛克代数环轴

和传统的实数数轴不同，格洛克数轴是一个圆环，圆环长度固定为 ∞ ，并将其分割为相等的 ∞ 段，每段长度为 1，将数轴逆时针依次标注 $0, 1, 2, \dots$ ，顺时针标注为 $-1, -2, \dots$ ，这样我们就得到了一个覆盖 $(-\infty, \infty)$ 数字空间的环形数轴。

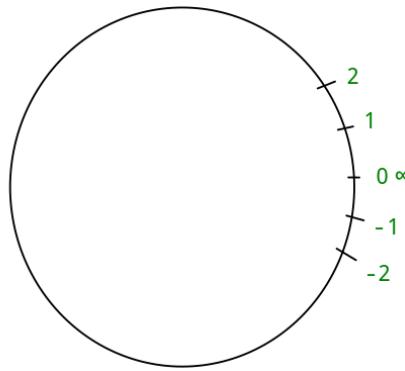


图 1.1: 格洛克代数环轴

下图是格洛克环轴展开后的样子，在数轴的右端标注的是单位刻度的长度量级 $\infty^0 = 1$ ，因此称为格洛克 0 阶数轴。

$$0 \boxed{1 \ 2 \ 3} \dots \boxed{-2 \ -1} \ \infty^0$$

图 1.2: 格洛克 0 阶数轴

需要注意的是， $\infty, -\infty$ 和原点 0 重合，这意味着数值大小达到 ∞ 时需要进位，以表示 ∞ 量级或更大的数。基于同样的原理，我们还需要量级更小的单位，即 ϵ ，用来进位到格洛克 0 阶数轴。

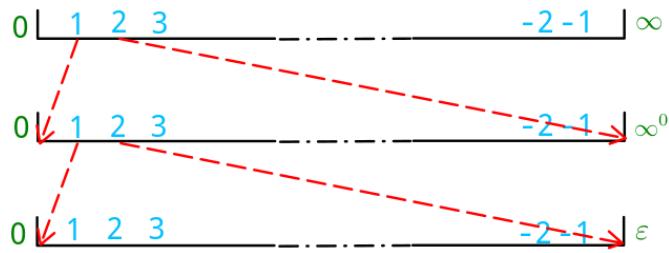


图 1.3: 宏空间和微空间

∞ 以及更大的量级称为格洛克宏空间，对应的， ϵ 以及更小的量级称为格洛克微空间，图 1.3 表示了格洛克宏空间 1 轴，格洛克 0 轴和格洛克微空间 1 轴。

宏空间 1 轴覆盖区间 $[0, \infty^2)$, 单位刻度大小为 1∞ , 将 1 单位刻度放大 ∞ 倍, 得到格洛克 0 轴。

格洛克 0 轴覆盖区间 $[0, \infty)$, 单位刻度大小为 1, 将 1 单位刻度放大 ∞ 倍, 得到微空间 1 轴。

微空间 1 轴覆盖区间 $[0, 1)$, 单位刻度大小为 $\frac{1}{\infty}$, 即 1ϵ 。

将宏空间和微空间以同样的方式分别向上和向下延展, 形成完整的格洛克代数空间。

宏空间和微空间

这是一个相对的概念, 如 ∞^{n+1} 是 ∞^n 的宏空间, ∞^n 是 ∞^{n+1} 的微空间; ϵ^n 是 ϵ^{n+1} 的宏空间, ϵ^{n+1} 是 ϵ^n 的微空间。

宏空间是微空间的无穷大, 微空间是宏空间的无穷小。

格洛克代数空间

完整格洛克代数空间如图 1.4 所示。

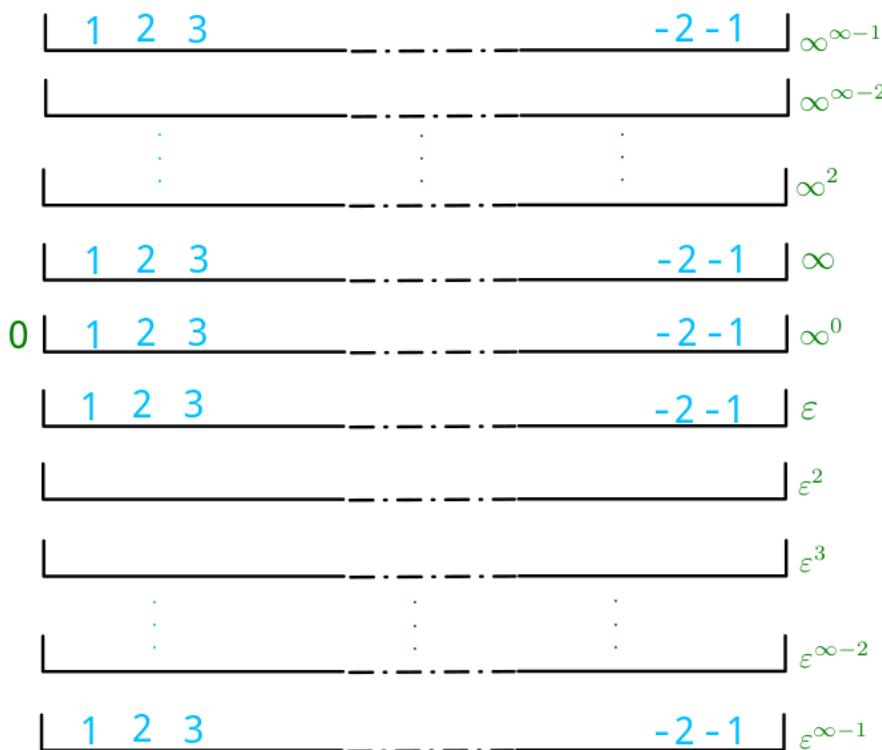


图 1.4: 格洛克代数空间

由于指数进位回 $0, \infty^\infty$ 和 ϵ^∞ 回归原点 0, 图中未予显示。

格洛克代数空间最大覆盖范围 $[0, \infty^\infty)$, 最小单位刻度 $1\epsilon^{\infty-1}$, 可以满足任何数值的可视化标注。

需要注意的是，由于所有可描述的实数都是 ∞ 的无穷小，正数 C 只能标注在一轴的最左侧的相对无穷小区间；同样的，负数 $-C$ 只能标注在最右侧的相对无穷小区间。

例 1.2.1. 在格洛克代数空间上标注：

$$(1) a = 0.5\infty + 1, \quad (2) b = 2\infty - 2, \quad (3) c = a + b.$$

解：

$$\begin{aligned} c &= a + b \\ &= (0.5\infty + 1) + (2\infty - 2) \\ &= (0.5\infty + 2\infty) + (1 - 2) \\ &= 2.5\infty - 1 \end{aligned}$$

a, b, c 如图 1.5 所示。

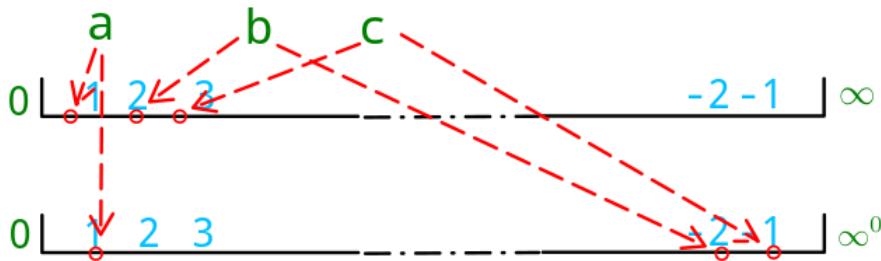


图 1.5: 数值标注

1.3 无穷大和无穷小的定义

经过前面两节关于无穷大和无穷小的讨论和格洛克代数空间的形成，下面给出无穷大和无穷小的正式定义。

定义 1.1. 无穷大和无穷小

设 ∞, ϵ 是未解析数 (*unsolved number*)， ∞ 是正整数，表示无穷大， ϵ 表示无穷小。

性质列表

- (1) ∞, ϵ 互为倒数，即 $\infty \cdot \epsilon = 1$ 。
- (2) 对于任意实数 C ，整数 n ， $\infty^{n+1} + C\infty^n = \infty^{n+1}$ ， $\epsilon^n + C\epsilon^{n+1} = \epsilon^n$ 。
- (3) ϵ 是最小的正数， $-\epsilon$ 是最大的负数， $0^+ = \epsilon$ ， $0^- = -\epsilon$ 。

- (4) 对于任意实数 C , $C^+ = C + \epsilon$, $C^- = C - \epsilon$ 。
- (5) 等效重定义。如果关于 ∞ 的表达式 $g(\infty)$ 的值仍是无穷大, 那么 ∞ 可以重定义为 $g(\infty)$, 记作: $\infty \leftarrow g(\infty)$, $\epsilon \leftarrow \frac{1}{g(\infty)}$ 。
- (6) 对于 $n > 1$ 有, $\log_n \infty < \sqrt[n]{\infty} < \infty < \infty^n < n^\infty < \infty!$ 。

性质 (1) 对无穷大和无穷小的关系进行了标准化, 使得数学属性更加清晰, 高阶无穷大和无穷小的比较和转换变成了数学计算, 而不是逻辑分析。

对于性质 (2), 当 $n = 0$ 时有

$$\infty + C = \infty \quad (1.1)$$

$$1 + C\epsilon = 1 \quad (1.2)$$

将 (1.1) 两边同时减去 ∞ , 将 (1.2) 两边同时减去 1, 得到

$$C = 0 \quad (1.3)$$

$$C\epsilon = 0 \quad (1.4)$$

(1.3) 说明在格洛克宏空间 1 轴 (1 阶无穷大) 的视角来看, 无论 C 多么大, 都是 ∞ 的无穷小, 格洛克宏空间 1 轴上的值都为 0; (1.4) 说明了在格洛克 0 轴 (实数轴) 的视角来看, 无论 n 多么大, $n\epsilon$ 都是无穷小, 在格洛克 0 轴上的值都为 0。这也是后续章节进行极限计算时的正确结果。

性质 (2) 本质上是将本章第一节关于无穷大和无穷小的运算规则进行了合并。

性质 (3) 可以合并到性质 (4), 分开是为了突出 0 的特殊性。 C^+ 表示实数 C 右侧最靠近它的数, C^- 表示 C 左侧最靠近它的数。这可以用格洛克微空间展开的术语进行理解。

格洛克微空间展开

对于展开点 C , 取单位长度区间 $[C, C + 1]$ 并放大 ∞ 倍数, 当我们将 C 点与微空间 1 轴的 0 对齐, 将会“看到”最靠近 C 右侧的坐标刻度是 1ϵ 。类似地, 取单位长度区间 $[C - 1, C]$ 并放大 ∞ 倍数, 将 $C - 1$ 点与微空间 1 轴的 0 对齐, 将会“看到”最靠近 C 左侧的坐标刻度是 -1ϵ 。如图 1.6 所示。

事实上, 我们也可以指定 $C^+ = C + n\epsilon$, $C^- = C - n\epsilon$, 结果也是正确的。从格洛克 0 轴 (或实数轴) 的角度来看, 无论我们指定靠近 C 的距离是多么小, 无论是百万分之一还是千万分之一, 都是 $n\epsilon$ 的无穷大。

关于性质 (5), 传统的无穷大和无穷小并没有明确的等阶划分, 事实上覆盖了整个格洛克代数空间。通过等效重定义, 我们可以得到一个明确的计算结果。如

$$f(\infty) = \log_n \infty$$

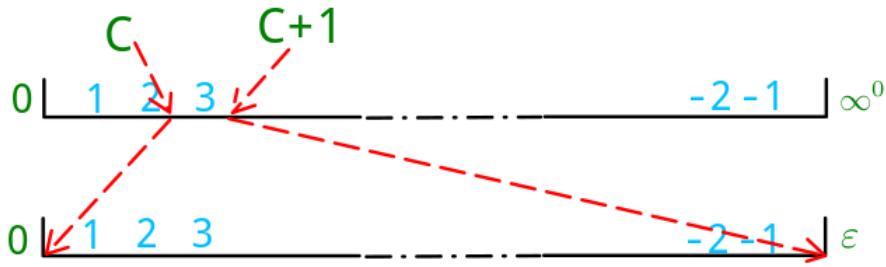


图 1.6: 格洛克微空间展开

$$\begin{aligned}
 &= \log_n n^\infty \quad (\infty \leftarrow n^\infty) \\
 &= \infty
 \end{aligned}$$

观察对数函数的图形，它是一个增长异常缓慢的函数，因此 $\log_n \infty$ 是一个非常低阶的无穷大，通过等效重定义，我们得到了一个明确的无穷大结果，本质上我们改变了无穷大在宏空间的位置，但没有改变传统无穷大的结果，这在极限计算时很有用。

性质（6）明确了不等式的各项位于不同阶的格洛克宏空间，不等式的左侧都是右侧的无穷小，体现了如下运算规则：

$$\begin{aligned}
 &\log_n \infty + \sqrt[n]{\infty} + \infty + \infty^n + n^\infty + \infty! \\
 &= \sqrt[n]{\infty} + \infty + \infty^n + n^\infty + \infty! \\
 &= \infty + \infty^n + n^\infty + \infty! \\
 &= \infty^n + n^\infty + \infty! \\
 &= n^\infty + \infty! \\
 &= \infty!
 \end{aligned}$$

虽然不等式的每一项都是无穷大，但它们的增长速度比值都达到了无穷大的量级，严格的证明需要后续知识的支撑，这里只进行说明性的证明：

增长顺序链：对数函数 \ll 幂函数 \ll 指数函数 \ll 阶乘函数

对数函数增长得极其缓慢，它是“加法”的逆运算（将乘法转化为加法），而幂函数仍涉及乘法。通过将无穷大“放大”，可以看出两者的差距巨大：

$$\begin{aligned}
 \frac{\sqrt[n]{\infty}}{\log_n \infty} &= \frac{\sqrt[n]{\infty^n}}{\log_n \infty^n} \quad (\infty \leftarrow \infty^n) \\
 &= \frac{\infty}{n \log_n \infty} \quad (\infty \leftarrow n^\infty) \\
 &= \frac{1}{n} \frac{n^\infty}{\infty}
 \end{aligned}$$

在幂函数家族内部，指数越大，增长越快，即 $\sqrt[n]{x} \ll x \ll x^n$

幂函数 x^n 只进行有限次的乘法 (n 次)，而指数函数 b^x 涉及无限次的乘法 (x 次 b 相乘)。

阶乘函数 $x!$ 的乘数是不断增大的 ($1, 2, 3, \dots, x$)，而指数函数 n^x 的乘数是固定的 n ，即

$$\frac{n^\infty}{\infty!} = \underbrace{\left(\frac{n}{1} \cdot \frac{n}{2} \cdot \dots \cdot \frac{n}{n} \right)}_{C_n \text{ (常数)}} \cdot \underbrace{\left(\frac{n}{n+1} \cdot \frac{n}{n+2} \cdot \dots \cdot \frac{n}{\infty} \right)}_{P_x \text{ (收敛部分)}}$$

可以通过函数的图形陡峭程度直观理解这种增长速度差异。

例 1.3.1. 数列用符号 $\{f(n)\}$ 表示，其中 $f(n)$ 是通项公式。如果 $n \rightarrow \infty$, $f(n) = C$ 那么数列是收敛的；如果 $n \rightarrow \infty$, $f(n) = \infty$ 那么数列是发散的；否则数列既不收敛也不发散。

判断下列数列的收敛性：

$$(1) \left\{ (-1)^n \frac{1}{n} \right\} \quad (2) \left\{ (-1)^n \frac{n+1}{n} \right\} \quad (3) \left\{ n - \frac{1}{n} \right\}$$

解：

(1)

$$\begin{aligned} f(n) &= (-1)^n \frac{1}{n} \\ f(\infty) &= (-1)^\infty \frac{1}{\infty} \\ &= (-1)^\infty \epsilon \end{aligned}$$

$f(\infty) = \epsilon = 0$, 如果 ∞ 为偶数； $f(\infty) = -\epsilon = 0$, 如果 ∞ 为奇数。

综合奇偶两种情况， $f(\infty) = 0$ ，数列收敛。

(2)

$$\begin{aligned} f(n) &= (-1)^n \frac{n+1}{n} \\ f(\infty) &= (-1)^\infty \frac{\infty+1}{\infty} \\ &= (-1)^\infty \frac{\infty}{\infty} \\ &= (-1)^\infty \end{aligned}$$

数列在 1 和 -1 之间来回振荡，既不收敛也不发散。

(3)

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

$$\begin{aligned}
 f(\infty) &= \infty - \frac{1}{\infty} \\
 &= \infty - \epsilon \\
 &= \infty
 \end{aligned}$$

数列发散。

通过上面的例子可以看出，关于无穷大和无穷小的计算完全符合数学运算法则，比传统的基于数学分析的形式化方法简单而清晰。

例 1.3.2.

$$\begin{aligned}
 f(\infty) &= \frac{\sqrt{\infty + 3}}{\infty} \\
 &= \frac{\sqrt{(\infty^2 - 3) + 3}}{\infty^2 - 3} \quad (\infty \leftarrow \infty^2 - 3) \\
 &= \frac{\infty}{\infty^2} \\
 &= \frac{1}{\infty} \\
 &= \epsilon \\
 &= 0
 \end{aligned}$$

通过等效重定义，我们精确计算出了表达式的结果，无需进行繁琐的形式化数学分析。

格洛克观点

格洛克代数空间的建立可以使我们抛弃传统的基于语言的数学逻辑，转而专注于简单精确的数学计算。后续章节我们将会看到微积分学是多么的简单，大量晦涩的定义、定理将被抛弃，一切变得如此不可思议。

第二章 函数和极限

函数是数学中一个核心概念，它描述了变量之间的一种对应关系。简单来说，函数将一个输入值（自变量）映射到一个输出值（因变量）。函数的概念起源于 17 世纪的莱布尼茨和牛顿，用于描述物理现象，如运动和变化。

传统的函数表示是一种泛化的表示，试图包容一切，解决一切问题，自变量和函数值（因变量）之间的关系抽象成规则，试图解决一切假想中的问题；因此在描述定义、定理时需要晦涩的前提限制条件，并形成抽象的形式化定义。本文或格洛克描述的函数只有一种，自变量和函数值之间的关系由自变量表达式确定，保持具体而简单。

极限是微积分的基础概念，它描述了函数在某个点附近的行为，即使函数在该点未定义或不连续。极限的概念由柯西和魏尔斯特拉斯在 19 世纪正式化，用于处理无穷小和无穷大。

基于格洛克代数空间，定义域的区间和函数的连续性进行了重新表述，极限概念被重新定义，一切变得简单而直观。

2.1 理解函数

函数是数学中描述对应关系的一种基本且非常重要的概念。通常写作 $y = f(x)$ 。

函数的关键特点

唯一性对应 (Uniqueness)：对于定义域内的每一个输入值 x ，函数都只能对应一个唯一的输出值 $f(x)$ 。

定义域 (Domain)：所有允许的输入值 x 的集合。

例如，在函数 $f(x) = \frac{1}{x}$ 中， x 不能为 0，所以定义域是除 0 以外的所有实数。

值域 (Range)：值域是实际输出值的集合。

函数的输入输出关系通过自变量表达式来描述，通常有几何图形来对应表示。例如函数 $f(x) = x^2 + 1$ ，在几何图形上表现为抛物线。

定义域区间

函数的定义域通常是数轴上的一个或多个连续的数值范围，因此我们常用区间符号来表示它。

圆括号 (a, b) 表示不包含端点（开区间）。表示 $a < x < b$ ，以及区间边界是 ∞ 或 $-\infty$ 的情况。

方括号 $[a, b]$ 表示包含端点（闭区间）。表示 $a \leq x \leq b$ 。

开区间转换为闭区间

依据格洛克无穷大和无穷小定义的性质 (4)，有

$$(a, b) \Rightarrow [a + \epsilon, b - \epsilon] \quad (a, \infty) \Rightarrow [a + \epsilon, \infty)$$

∞ 既可以理解为开区间也可以理解为闭区间，根据具体的问题进行不同的理解。

分段函数

分段函数将定义域分成若干互不相交的子区间，并在每个子区间上定义一个子函数。

分段函数常用大括号表示，例如：

$$f(x) = |x| = \begin{cases} -x, & \text{if } x < 0 \\ x, & \text{if } x \geq 0 \end{cases}$$

$$f(x) = \frac{1}{x} = \begin{cases} \frac{1}{x}, & \text{if } x < 0 \\ \frac{1}{x}, & \text{if } x > 0 \end{cases}$$

$$f(x) = \begin{cases} x^2 + 1, & \text{if } x < 1 \\ 3x - 1, & \text{if } x \geq 1 \end{cases}$$

还有一些常见的特殊分段函数：

阶跃函数 (Step Function)：分段函数由常数函数组成。例如，单位阶跃函数 (Unit Step Function)。

分段线性函数 (Piecewise Linear Function)：分段函数由线段（线性函数）组成。通过以上讨论，我们现在可以明确：**函数的概念和性质总是作用于单一定义域区间。**

复合函数

复合函数指的是将两个或多个函数组合起来，形成一个新的函数。具体来说，如果有函数 (f) 和 (g) ，则它们的复合函数记作 $f \circ g$ ，定义为 $(f \circ g)(x) = f(g(x))$ ，其中，先计算 $g(x)$ ，再计算 f 。

复合函数 $f \circ g$ 的定义域同时受到 g 和 f 定义域的双重约束。

$f \circ g$ 和 $g \circ f$ 通常并不相同，也就是说复合函数的顺序很重要。

例 2.1.1. 假设 $f(x) = x^2$, $g(x) = x + 1$ 。

$$(f \circ g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2 = x^2 + 2x + 1$$

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

反函数

反函数就是“逆转”原函数的操作，将输出变回输入。函数 $f(x)$ 的反函数通常表示为 $f^{-1}(x)$ 。

反函数存在条件

- (1) 一对一性：每个 x 对应唯一的 y ，并且每个 y 对应唯一的 x 。
- (2) 覆盖性：函数的值域覆盖整个定义域。

反函数的性质

- (1) $f(f^{-1}(x)) = f^{-1}(f(x)) = x$ 。
- (2) 原函数和反函数的图形关于直线 $y = x$ 对称。

例 2.1.2. 假设 $f(x) = 2x + 1$, $g(x) = \frac{1}{x}$, 求反函数并验证恒等性质。

解：设 $y = f(x) = 2x + 1$, 那么 $x = \frac{y-1}{2}$, 交换变量 x, y , 得到 $y = \frac{x-1}{2}$, 反函数

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

恒等式

$$f(f^{-1}(x)) = f\left(\frac{x-1}{2}\right) = 2 \cdot \frac{x-1}{2} + 1 = x$$

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

所以 $f(f^{-1}(x)) = f^{-1}(f(x)) = x$ 。

设 $y = g(x) = \frac{1}{x}$, 那么 $x = \frac{1}{y}$, 交换变量 x, y , 得到 $y = \frac{1}{x}$, 反函数

$$f^{-1}(x) = \frac{1}{x}$$

恒等式

$$f(f^{-1}(x)) = f\left(\frac{1}{x}\right) = \frac{1}{1/x} = x$$

$$f^{-1}(f(x)) = f^{-1}\left(\frac{1}{x}\right) = \frac{1}{1/x} = x$$

所以 $f(f^{-1}(x)) = f^{-1}(f(x)) = x$ 。

2.2 参数方程

参数方程是一个非常重要的数学概念，尤其在描述运动轨迹和复杂曲线时非常方便。

函数关系是参数方程的一种特殊情况，而参数方程提供了描述更广泛几何图形和运动的更灵活方式。

在平面直角坐标系中，我们通常用一个含有 x 和 y 的方程（普通方程）来表示一条曲线。

普通方程（或笛卡尔方程）： $y = f(x)$ 或 $F(x, y) = 0$ 。

参数方程（Parametric Equation）则是引入一个参数（通常用 t 或 θ 表示），将曲线上点的坐标 (x, y) 分别表示成这个参数的函数：

$$\begin{cases} x = f(t) \\ y = g(t) \end{cases} \quad (t \text{ 为参数})$$

当参数 t 在一个指定的范围内变化时，对应的 (x, y) 坐标点就描绘出一条曲线。在物理学中，参数 t 常常代表时间。

使用参数方程有几个显著的优点：

描述运动轨迹更自然：如果 t 代表时间，那么 $x = f(t)$ 和 $y = g(t)$ 就直接描述了一个物体在 t 时刻的横坐标和纵坐标，非常适合描述质点运动。

表示复杂曲线更简单：有些曲线的普通方程非常复杂，甚至无法写出，但它们的参数方程却相对简单。例如摆线（Cycloid）。

表示非函数关系：普通方程 $y = f(x)$ 只能表示一个 x 对应一个 y 的函数关系。但对于圆这样的曲线（一个 x 对应两个 y ），用参数方程表示更为简洁。

常见曲线的参数方程

普通方程与参数方程的互化

(1) 参数方程化为普通方程：核心思想是消去参数 t 。

例 2.2.1. 将 $\begin{cases} x = 2t + 1 \\ y = 4t^2 \end{cases}$ 化为普通方程。

从第一个方程解出 t ： $t = \frac{x-1}{2}$

代入第二个方程： $y = 4 \left(\frac{x-1}{2} \right)^2$

化简： $y = 4 \cdot \frac{(x-1)^2}{4} = (x-1)^2$

结果：普通方程为 $y = (x-1)^2$ （开口向上的抛物线）。

曲线类型	普通方程	参数方程	参数的范围
直线	$y - y_0 = m(x - x_0)$	$\begin{cases} x = x_0 + at \\ y = y_0 + bt \end{cases}$	$t \in (-\infty, \infty)$
圆	$(x - h)^2 + (y - k)^2 = r^2$	$\begin{cases} x = h + r \cos \theta \\ y = k + r \sin \theta \end{cases}$	$\theta \in [0, 2\pi)$
椭圆	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	$\begin{cases} x = a \cos \theta \\ y = b \sin \theta \end{cases}$	$\theta \in [0, 2\pi)$
抛物线	$y^2 = 2px$	$\begin{cases} x = \frac{p}{2}t^2 \\ y = pt \end{cases}$	$t \in (-\infty, \infty)$

例 2.2.2. 将 $\begin{cases} x = 3 \cos \theta \\ y = 3 \sin \theta \end{cases}$ 化为普通方程。

利用三角恒等式: $\cos^2 \theta + \sin^2 \theta = 1$

解出 $\cos \theta$ 和 $\sin \theta$: $\cos \theta = x/3$, $\sin \theta = y/3$

代入恒等式: $(x/3)^2 + (y/3)^2 = 1$

化简: $x^2 + y^2 = 9$

结果: 普通方程为 $x^2 + y^2 = 9$ (圆心在原点, 半径为 3 的圆)。

(2) 普通方程化为参数方程: 核心思想是选择一个合适的参数 t 。

选择参数没有固定方法, 但通常选择 x 或 y 或它们的组合, 或选择一个有几何意义的角。

例 2.2.3. 将直线 $y = 3x - 5$ 化为参数方程。

设 $x = t$:

代入原方程: $y = 3t - 5$

结果: $\begin{cases} x = t \\ y = 3t - 5 \end{cases}$

例 2.2.4. 将椭圆 $\frac{x^2}{16} + \frac{y^2}{9} = 1$ 化为参数方程。

改写方程: $(\frac{x}{4})^2 + (\frac{y}{3})^2 = 1$

利用三角恒等式, 设 $\frac{x}{4} = \cos \theta$ 和 $\frac{y}{3} = \sin \theta$:

结果: $\begin{cases} x = 4 \cos \theta \\ y = 3 \sin \theta \end{cases}$ (θ 为参数)

参数方程是用一个独立的变量 t 来间接描述 x 和 y 的关系。

2.3 函数的极限

函数极限是微积分中的一个基本概念，它描述了一个函数在自变量（输入值）接近某一给定值时，它的函数值（输出值）所表现出的趋势。

函数极限的直观理解

- (1) 函数 $f(x)$ 在 x 趋近于 a 时的极限是 L ，记作 $\lim_{x \rightarrow a} f(x) = L$ 。
- (2) 当 x 无限地接近 a (但不等于 a) 时，函数值 $f(x)$ 会无限地接近一个确定的值 L 。
- (3) $f(x)$ 在 $x = a$ 处是否有定义，以及 $f(a)$ 的值是多少，并不影响 $\lim_{x \rightarrow a} f(x)$ 的结果。极限关注的是趋近过程，而非终点的值。

从上面直觉出发，逐渐演变完善了极限的 $\epsilon - \delta$ 定义。 $\epsilon - \delta$ 定义为极限提供了逻辑严密性，但代价是牺牲了直觉性、增加了学习难度和操作复杂性。

格洛克极限定义

格洛克采用逆向思维，并以可视化的格洛克代数空间和无穷大无穷小的定义作为逻辑支撑，给出极限的定义。

定义 2.1 (函数的极限). 对于函数 $f(x)$, $x \rightarrow a \iff x = a \pm \epsilon$, 函数 $f(x)$ 在 x 趋近于 a 时的极限

- (1) $\lim_{x \rightarrow a} f(x) = f(a \pm \epsilon) = f(a)$, 如果 a 位于定义域区间内。
- (2) $\lim_{x \rightarrow a^+} f(x) = f(a + \epsilon)$, 如果 a 位于定义域区间左边界。
- (3) $\lim_{x \rightarrow a^-} f(x) = f(a - \epsilon)$, 如果 a 位于定义域区间右边界。

对于 (2) 和 (3)，如果 a 是闭区间边界，极限值应用 $f(a)$ 仍然成立。

格洛克将传统极限的动态趋近描述转换成了静态描述，因为在格洛克微空间“看到了”最接近 a 的极限位置。由此不再需要晦涩难解的逻辑推理描述，简单直接地进行极限计算。

例 2.3.1. 计算极限

$$(1) \lim_{x \rightarrow 1} (2x + 1) \quad (2) \lim_{x \rightarrow \infty} \frac{3x^3 - x + 5}{2x^3 + 4x^2 - 1} \quad (3) \lim_{x \rightarrow 2} \frac{x^2 - 4}{x - 2} \quad (4) \lim_{x \rightarrow 0} \frac{\sqrt{1+x} - 1}{x}$$

解：

$$(1) \lim_{x \rightarrow 1} (2x + 1) = 2 \cdot (1) + 1 = 3$$

$$(2) \lim_{x \rightarrow \infty} \frac{3x^3 - x + 5}{2x^3 + 4x^2 - 1} = \frac{3\infty^3 - \infty + 5}{2\infty^3 + 4\infty^2 - 1} = \frac{3\infty^3}{2\infty^3} = \frac{3}{2}$$

$$(3) \lim_{x \rightarrow 2} \frac{x^2 - 4}{x - 2} = \frac{(2 + \epsilon)^2 - 4}{(2 + \epsilon) - 2} = \frac{4\epsilon + \epsilon^2}{\epsilon} = 4 + \epsilon = 4$$

$$\begin{aligned}
 (4) \lim_{x \rightarrow 0} \frac{\sqrt{1+x} - 1}{x} &= \frac{\sqrt{1+\epsilon} - 1}{\epsilon} \\
 &= \frac{(\sqrt{1+\epsilon} - 1)(\sqrt{1+\epsilon} + 1)}{\epsilon(\sqrt{1+\epsilon} + 1)} \\
 &= \frac{1}{\sqrt{1+\epsilon} + 1} = \frac{1}{2}
 \end{aligned}$$

上例 (3) 和 (4) 略去了计算左极限, 因为左右极限相等。

函数的连续性

(1) 在定义域区间内函数总是连续的, 在闭区间左边界右连续, 在闭区间右边界左连续。

(2) 如果分段函数的分段点有定义, 分段子函数分别为 f_1, f_2 , 并且有 $f_1(a-\epsilon) = f_2(a+\epsilon) = f(a)$, 那么分段函数在分段点处连续。

函数的连续性是显而易见的, 换句话说, 函数 $f(x)$ 的值随着自变量 x 的连续变化而连续变化。函数图形可视化的证明了函数的连续性。事实上, 证明函数的连续性很容易进入逻辑上的循环证明。

讨论函数的连续性只有在函数的两个相邻区间的分段点处才有意义, 格洛克采用分段函数的方式描述了这种情况, 分段子函数 f_1, f_2 可以相同, 也可以不同。具有苛刻定义的非常规函数因失去一般性不在本节进行讨论。

例 2.3.2. 计算极限

$$(1) \lim_{x \rightarrow 0} \frac{1}{x} \quad (2) \lim_{x \rightarrow 0} |x|$$

解:

$$(1) \lim_{x \rightarrow 0^-} \frac{1}{x} = \frac{1}{-\epsilon} = -\infty \quad \lim_{x \rightarrow 0^+} \frac{1}{x} = \frac{1}{\epsilon} = \infty$$

左右极限不相等, 极限不存在。

$$(2) \lim_{x \rightarrow 0^-} |x| = |- \epsilon| = \epsilon = 0 \quad \lim_{x \rightarrow 0^+} |x| = |\epsilon| = 0$$

分段点 0 左右极限相等, 极限存在, 极限值为 0。

分段点 0 有定义且函数值为 0, 所以函数在分段点 0 处连续。

例 2.3.3. 计算极限

$$\lim_{x \rightarrow 0} \frac{\sin x}{x}$$

解: 这是一个著名的极限问题, 传统上使用夹逼定理进行求解。

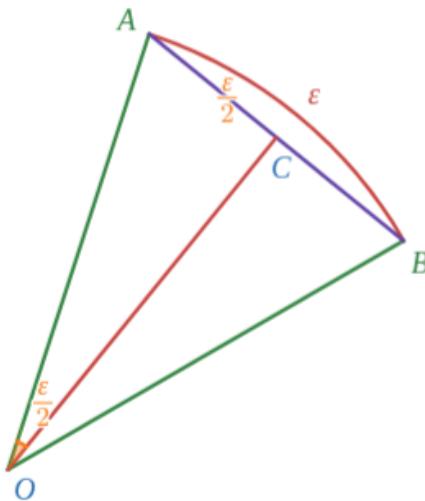


图 2.1: 单位圆上的弧

当光滑曲线上的两点 A, B 越来越接近时, 曲线弧 AB 与割线 AB 逐渐重合, 换句话说, 只要 A, B 两点的距离足够小, $\widehat{AB} = |AB|$, 如图 2.1 所示。

我们取弧长 $\widehat{AB} = \epsilon$, 得到如下关系:

$$|AB| = \epsilon, \quad \angle AOB = \epsilon, \quad |AC| = \angle AOC = \frac{\epsilon}{2}$$

根据三角公式

$$\begin{aligned} \sin \frac{\epsilon}{2} &= \frac{AC}{OA} = AC = \frac{\epsilon}{2} \\ \cos \frac{\epsilon}{2} &= \frac{OC}{OA} = \sqrt{1 - \left(\frac{\epsilon}{2}\right)^2} \end{aligned}$$

现在我们可以计算极限

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\sin x}{x} &= \frac{\sin \epsilon}{\epsilon} = \frac{2 \sin \frac{\epsilon}{2} \cos \frac{\epsilon}{2}}{\epsilon} \\ &= \frac{2 \cdot \frac{\epsilon}{2} \cdot \sqrt{1 - \left(\frac{\epsilon}{2}\right)^2}}{\epsilon} \\ &= \sqrt{1 - \left(\frac{\epsilon}{2}\right)^2} \\ &= 1 \end{aligned}$$

例 2.3.4. 计算极限

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x}$$

解: 利用半角公式有 $\cos x = 1 - 2 \sin^2 \left(\frac{x}{2}\right)$, 所以

$$\lim_{x \rightarrow 0} \frac{1 - \cos x}{x} = \frac{1 - \cos \epsilon}{\epsilon}$$

$$\begin{aligned}
 &= \frac{1 - 1 + 2 \sin^2\left(\frac{\epsilon}{2}\right)}{\epsilon} \\
 &= \frac{2 \cdot \left(\frac{\epsilon}{2}\right)^2}{\epsilon} \\
 &= \frac{1}{2}\epsilon = 0
 \end{aligned}$$

极限的 $\epsilon - \delta$ 定义

设函数 $f(x)$ 在点 x_0 的某个去心邻域内有定义。如果存在一个实数 L , 使得:

对于任意给定的正数 ϵ (无论它多么小), 总存在一个正数 δ (它通常依赖于 ϵ), 使得当 x 满足 $0 < |x - x_0| < \delta$ 时, 都有 $|f(x) - L| < \epsilon$ 成立。

则称 L 是函数 $f(x)$ 当 x 趋近于 x_0 时的极限, 记作:

$$\lim_{x \rightarrow x_0} f(x) = L$$

2.4 自然常数和自然对数

自然常数 e 是数学中一个重要的无理数, 常作为自然对数的底数出现。它大约等于 2.71828, 并在许多自然现象、增长模型 (如复利) 和极限计算中扮演关键角色。

e 最早由瑞士数学家雅各布 · 伯努利 (Jacob Bernoulli) 在研究复利问题时发现, 后来由莱昂哈德 · 欧拉 (Leonhard Euler) 在 18 世纪正式引入符号”e” (源自”exponential”, 指数的)。

e 可以定义为以下极限:

$$e = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = \left(1 + \frac{1}{\infty}\right)^\infty = (1 + \epsilon)^\infty$$

从格洛克代数空间的角度看, 自然常数是一个未解析的无穷大无穷小计算表达式。使用无穷大等效重定义 $\infty \leftarrow f(\infty)$, 格洛克给出自然常数的定义。

定义 2.2. 自然常数

设 $f(\infty)$ 是关于 ∞ 的表达式, 即 $f(\infty)$ 位于格洛克宏空间; $f(\epsilon)$ 是关于 ϵ 的表达式, 即 $f(\epsilon)$ 位于格洛克微空间。自然常数 e 定义为

$$e = \left(1 + \frac{1}{f(\infty)}\right)^{f(\infty)} \tag{2.1}$$

$$e = (1 + f(\epsilon))^{1/f(\epsilon)} \tag{2.2}$$

需要明确的是，即使 $f(\infty), f(\epsilon)$ 是负值时，自然常数的定义仍然有效。

$$\begin{aligned} \left(1 - \frac{1}{\infty}\right)^{-\infty} &= \left(\frac{\infty - 1}{\infty}\right)^{-\infty} = \left(\frac{\infty}{\infty - 1}\right)^\infty \\ &= \left(1 + \frac{1}{\infty - 1}\right)^{\infty-1} \cdot \left(1 + \frac{1}{\infty - 1}\right) \\ &= e \cdot (1 + \epsilon) \\ &= e \end{aligned}$$

自然对数，记作 $\ln(x)$ ，是以 e 为底的对数函数，它是 $\log_e(x)$ 的简写。

定理 2.1. 自然恒等式

$$(1 + a \cdot \epsilon)^b = e^{ab \cdot \epsilon} = 1 + ab \cdot \epsilon \quad (2.3)$$

$$a^\epsilon = e^{\ln a \cdot \epsilon} = 1 + \ln a \cdot \epsilon \quad (2.4)$$

$$\ln(1 + a \cdot \epsilon) = a \cdot \epsilon \quad (2.5)$$

证明

$$\begin{aligned} (1 + a \cdot \epsilon)^b &= (1 + a \cdot \epsilon)^{\frac{1}{a\epsilon} \cdot a\epsilon \cdot b} \\ &= e^{ab \cdot \epsilon} \\ &= (1 + ab \cdot \epsilon)^{\frac{1}{ab\epsilon} \cdot (ab \cdot \epsilon)} \\ &= 1 + ab \cdot \epsilon \end{aligned}$$

$$\begin{aligned} a^\epsilon &= e^{\ln a^\epsilon} = e^{\epsilon \cdot \ln a} \\ &= (1 + \epsilon \cdot \ln a)^{\frac{1}{\epsilon \cdot \ln a} \cdot (\epsilon \cdot \ln a)} \\ &= 1 + \ln a \cdot \epsilon \end{aligned}$$

$$\begin{aligned} \ln(1 + a \cdot \epsilon) &= \ln e^{a \cdot \epsilon} \\ &= a \cdot \epsilon \cdot \ln e \\ &= a \cdot \epsilon \end{aligned}$$

例 2.4.1. 计算极限

$$\begin{aligned} \lim_{x \rightarrow \infty} \left(1 + \frac{2}{x}\right)^x &= \left(1 + \frac{2}{\infty}\right)^\infty \\ &= \left(1 + \frac{2}{\infty}\right)^{\frac{\infty}{2} \cdot 2} \end{aligned}$$

$$= e^2$$

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{\ln(1+3x)}{x} &= \frac{\ln(1+3\epsilon)}{\epsilon} \\ &= \frac{\ln e^{3\epsilon}}{\epsilon} = \frac{3\epsilon}{\epsilon} \\ &= 3\end{aligned}$$

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{e^x - 1}{x} &= \frac{e^\epsilon - 1}{\epsilon} \\ &= [(1+\epsilon)^{\infty \cdot \epsilon} - 1] \cdot \frac{1}{\epsilon} \\ &= \frac{(1+\epsilon) - 1}{\epsilon} = \frac{\epsilon}{\epsilon} \\ &= 1\end{aligned}$$

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{a^x - 1}{x} &= \frac{a^\epsilon - 1}{\epsilon} \\ &= \frac{(1 + \ln a \cdot \epsilon) - 1}{\epsilon} \\ &= \frac{\ln a \cdot \epsilon}{\epsilon} \\ &= \ln a\end{aligned}$$

例 2.4.2. 计算极限 $\lim_{x \rightarrow 1} \frac{\ln x}{1-x}$, $\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt[n]{x}}$

$$\begin{aligned}\lim_{x \rightarrow 1} \frac{\ln x}{1-x} &= \frac{\ln(1+\epsilon)}{1-(1+\epsilon)} \\ &= \frac{\ln e^\epsilon}{-\epsilon} \\ &= \frac{\epsilon}{-\epsilon} = -1\end{aligned}$$

$$\begin{aligned}\lim_{x \rightarrow \infty} \frac{\ln x}{\sqrt[n]{x}} &= \frac{\ln \infty}{\sqrt[n]{\infty}} \\ &= \frac{\ln \infty^n}{\sqrt[n]{\infty^n}} \quad (\infty \leftarrow \infty^n) \\ &= \frac{n \ln \infty}{\infty} = 0\end{aligned}$$

自然恒等式是极限运算

显然, 定理 2.1 的恒等结果是在舍弃更高阶无穷小后的极限运算结果, 例如极限计算

$$\frac{e^{\epsilon^2} - 1 - \epsilon^2}{\epsilon^4}$$

我们不能简单的应用定理 2.1, 那样分子将为 0, 此时正确的方法是使用“平滑”的泰勒展开来得到更高阶无穷小, 从而计算出正确的结果。

定理 2.1 的意义在于简化指数函数导数公式的推导, 免去复杂的逻辑推理。

2.5 泰勒展开在极限计算中的应用

泰勒展开 (Taylor Series Expansion) 是计算涉及 $e, \ln(x), \sin(x), \cos(x)$ 等基本函数的极限时, 一个非常强大且高效的工具。

它主要利用函数在某一点 (通常是 $x = 0$, 即麦克劳林展开) 附近的多项式近似来简化分子和分母。

常用函数的麦克劳林展开式 ($x \rightarrow 0$)

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (2.6)$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \quad (2.7)$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \dots \quad (2.8)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad (2.9)$$

$$\ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad (2.10)$$

$$(1 + x)^a = 1 + ax + \frac{a(a-1)}{2!}x^2 + \frac{a(a-1)(a-2)}{3!}x^3 + \dots \quad (2.11)$$

例 2.5.1. 计算极限

$$\lim_{x \rightarrow 0} \frac{\ln(1 + x)}{x} = \frac{\ln(1 + \epsilon)}{\epsilon} = \frac{\epsilon}{\epsilon} = 1$$

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\ln(\cos x)}{x^2} &= \frac{\ln(\cos \epsilon)}{\epsilon^2} = \frac{\ln(1 - \frac{1}{2}\epsilon^2)}{\epsilon^2} \\ &= \frac{-\frac{1}{2}\epsilon^2}{\epsilon^2} = -\frac{1}{2} \end{aligned}$$

$$\lim_{x \rightarrow 0} \frac{e^{x^2} - 1 - x^2}{x^4} = \frac{e^{\epsilon^2} - 1 - \epsilon^2}{\epsilon^4}$$

$$\begin{aligned}
 &= \frac{(1 + \epsilon^2 + \frac{1}{2}\epsilon^4) - 1 - \epsilon^2}{\epsilon^4} \\
 &= \frac{\frac{1}{2}\epsilon^4}{\epsilon^4} = \frac{1}{2}
 \end{aligned}$$

例 2.5.2. 计算极限

$$\begin{aligned}
 \lim_{x \rightarrow 0} \frac{\sin x}{x} &= \frac{\sin \epsilon}{\epsilon} = \frac{\epsilon}{\epsilon} = 1 \\
 \lim_{x \rightarrow 0} \frac{1 - \cos x}{x} &= \frac{1 - \cos \epsilon}{\epsilon} = \frac{1 - (1 - \frac{1}{2}\epsilon^2)}{\epsilon} = \frac{1}{2}\epsilon = 0
 \end{aligned}$$

使用泰勒展开计算极限的要点：

- (1) 确定阶数：根据分母的无穷小阶数来确定分子和分母需要展开到的阶数。目标是让分子中最低次项（非零项）的阶数与分母的阶数相等。
- (2) 抵消低阶项：展开后，分子中的低阶项（如常数项、一次项等）通常会相互抵消，只剩下与分母同阶或更高阶的无穷小项。
- (3) 提取主部：保留分子分母中最低次项的系数，即可求出极限。

我们在学习泰勒级数之前，提前利用泰勒级数的展开式来简化复杂极限的计算，避免进行不必要的复杂逻辑分析。

循环论证

复杂多项式的泰勒级数展开在计算极限时很强大和直观，但是注意不能用它来证明极限和导数的推导，因为泰勒级数是在极限和导数的理论框架下推导出来的。否则可能会出现循环论证的逻辑错误。

第三章 点微分和导数

传统上，微分是描述函数在某一点附近变化的线性近似。这也是微积分学混乱而难学的根源之一。

格洛克利用全新定义的无穷小概念，对微分进行了重新定义：微分是函数在某一点的极限变化量。并和导数进行了统一，形成微分表达式。

导数是描述函数在某一点变化率的精确数值。传统的极限理解方式和计算方法让初学者对导数的学习曲线非常陡峭，莱布尼兹的导数定义符号 $\frac{d}{dx}F(x)$ 大大增加了导数的神秘性。

格洛克弃用莱布尼兹导数表示符号，回归常识，重新描述导数为微分表达式的变换，即导数是函数微分和自变量微分的商。与传统导数概念描述不同的是，导数本质上是一个二级定义，它不是原子定义，而是建立在微分定义之上。

3.1 极限变化量与点微分

通过前面二章的学习，我们具备了解决瞬时速度和切线问题的理论框架，以此为切入点，引出点微分的概念。

瞬时速度问题

我们首先考察平均速度的问题，如果位移函数 $S = F(x)$ ，其中自变量 x 表示时间，那么在 t_1, t_2 时间间隔内的平均速度表示为

$$\bar{v} = \frac{F(t_2) - F(t_1)}{t_2 - t_1}$$

如果我们用 Δt 表示时间间隔，即 $\Delta t = t_2 - t_1$ ，那么平均速度又可以表示为

$$\bar{v} = \frac{F(t + \Delta t) - F(t)}{\Delta t}$$

可以看出，计算平均速度的公式很简单，计算也很容易。

我们现在的问题是如何计算 $x = t$ 时刻的瞬时速度？问题的难点在于速度的计算要求时间间隔必须存在，唯一的办法是不断缩小时间间隔 Δt ，以此来获取一个近似的 t 时刻的瞬时速度。 Δt 越小，得到的平均速度就越接近瞬时速度。

基于格洛克代数空间，我们在数值 t 处进行格洛克微空间展开，获取最小的时间间隔 ϵ ，得到 t 时刻的瞬时速度

$$v = \frac{F(t + \epsilon) - F(t)}{\epsilon}$$

对于任意时刻 x ，我们得到数学描述的速度函数

$$v(x) = \frac{F(x + \epsilon) - F(x)}{\epsilon} \quad (3.1)$$

这是一个极限运算，可以用我们在第二章的极限计算方法计算出 $v(x)$ 函数。

例 3.1.1. 假设一个物体沿直线运动，其位移 S （单位：米，m）随时间 t （单位：秒，s）变化的函数关系为：

$$S(t) = 3t^2 + 5t - 2$$

计算该物体在 $t = 2$ 秒时的瞬时速度。

解：

首先求解瞬时速度函数 $v(t)$ ：

$$\begin{aligned} v(t) &= \frac{S(t + \epsilon) - S(t)}{(t + \epsilon) - t} \\ &= \frac{[3(t + \epsilon)^2 + 5(t + \epsilon) - 2] - (3t^2 + 5t - 2)}{\epsilon} \\ &= \frac{(3t^2 + 6t \cdot \epsilon + 3\epsilon^2 + 5t + 5\epsilon - 2) - (3t^2 + 5t - 2)}{\epsilon} \\ &= \frac{6t \cdot \epsilon + 5\epsilon + 3\epsilon^2}{\epsilon} \\ &= 6t + 5 + 3\epsilon \\ &= 6t + 5 \end{aligned}$$

有了速度函数，现在我们可以计算物体在 $t = 2$ 秒时的瞬时速度

$$v(2) = 6 \cdot (2) + 5 = 17$$

切线斜率问题

几何上，直线的斜率可通过已知的两点坐标计算

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

切线是一条“刚好触碰”曲线上某一点的直线。对于曲线 $y = F(x)$ 上的某一点 P ，如果我们要计算经过 P 点切线的斜率，会遇到和上面计算瞬时速度相同的

问题，我们还需要曲线上额外的一点 Q ，即计算割线 PQ 的斜率来近似 P 点切线的斜率。 Q 越接近 P ，计算的割线斜率就越接近 P 点切线的斜率。

$$m_{PQ} = \frac{y_Q - y_P}{x_Q - x_P}$$

令 $\Delta x = x_Q - x_P$ ，并使 Δx 足够小，曲线 $y = F(x)$ 上经过 P 点切线的斜率可近似表示为

$$m_P \approx \frac{F(x_P + \Delta x) - F(x_P)}{\Delta x}$$

我们在数值 x_P 处进行格洛克微空间展开，获取最小的 PQ 间隔 ϵ ，得到 P 点切线的斜率

$$m_P = \frac{F(x_P + \epsilon) - F(x_P)}{\epsilon}$$

对于曲线上的任意点 $(x, F(x))$ ，切线的斜率函数表示为

$$m(x) = \frac{F(x + \epsilon) - F(x)}{\epsilon} \quad (3.2)$$

例 3.1.2. 计算抛物线 $y = x^2$ 在点 $(2, 4)$ 处的切线斜率，并求出通过该点的切线方程。

解：

首先求解抛物线的切线斜率函数 $m(x)$:

$$\begin{aligned} m(x) &= \frac{F(x + \epsilon) - F(x)}{\epsilon} \\ &= \frac{(x + \epsilon)^2 - x^2}{\epsilon} \\ &= \frac{(x^2 + 2x \cdot \epsilon + \epsilon^2) - x^2}{\epsilon} \\ &= \frac{2x \cdot \epsilon + \epsilon^2}{\epsilon} \\ &= 2x + \epsilon \\ &= 2x \end{aligned}$$

有了斜率函数，现在我们可以计算抛物线 $y = x^2$ 在点 $(2, 4)$ 处的切线斜率

$$m(2) = 2 \cdot (2) = 4$$

切线方程 (点斜式):

$$y - y_0 = m(x - x_0)$$

$$y - 4 = 4(x - 2)$$

$$y = 4x - 8 + 4$$

$$y = 4x - 4$$

从上面两个例子可以看出，速度问题和切线斜率问题虽然完全不同，但在解决问题的数学方法上却完全相同：我们需要找到自变量的极限变化量和对应函数的极限变化量，通过除法得到要解决问题的极限变化率。由此引出和传统定义完全不同的微分概念和定义：

定义 3.1. 微分

对于函数 $F(x)$ ，自变量 x 通过格洛克微空间展开的方式获得 $(x, F(x))$ 极限变化量的过程称为微分。如果 $\delta = f(\epsilon)$ 是无穷小的表达式，微分用符号 d 表示，那么

- (1) $dx = (x + \delta) - x$ ，称为微分自变量。
- (2) $dF(x) = F(x + \delta) - F(x)$ ，称为微分表达式。
- (3) 基于最简计算原则，通常取 $\delta = \epsilon$ 。

对于 (1) 包含两层含义，首先定义了微分点 x ，然后是自变量极限变化量的大小 $\delta = \epsilon$ ，这是最简单的表示，也可以是无穷小的表达式。

对于 (2) 首先是一个微分表达式而不是一个直接计算结果的值，因为进行极限计算

$$F(x + \epsilon) - F(x) = F(x) - F(x) = 0$$

其次 $dF(x)$ 是与 x 相对应的极限变化量，也就是随着 x 的变化， $dF(x)$ 也随之变化。最后，定义中给出的是 x 右侧的极限变化量，取左侧或同时取两侧也是正确的，定义中未明确是为了保持简单。因此 $dF(x)$ 可以表示为

$$dF(x) = F(x + \epsilon) - F(x) = F(x) - F(x - \epsilon) = F(x + \epsilon/2) - F(x - \epsilon/2)$$

对于定义域的闭区间边界只能取右侧或左侧极限变化量。

因此，(3.1) 和 (3.2) 可以统一表示为

$$f(x) = \frac{dF(x)}{dx}$$

微分表达式与微分自变量的比值称为微分变化率，极限变化率或瞬时变化率，传统上统一定义为导函数，简称导数。

3.2 导数和导数公式

有了微分的定义，我们在此基础上定义导函数。导函数简称导数。

定义 3.2. 导函数

微分表达式与微分自变量的比值函数称为导函数。如果原函数用 $f(x)$ 表示，那么导函数用 $f'(x)$ 表示。

$$f'(x) = \frac{df(x)}{dx} = \frac{d}{dx} f(x)$$

需要注意的是，对于 $\frac{d}{dx} f(x)$ ，我们不使用 $\frac{d}{dx}$ 作为导数的表示符号，而是作为除法表达式。

导数用来描述函数相对其自变量的微分变化率，极限变化率或瞬时变化率，或者几何上的曲线切线的斜率。

如果 $F(x)$ 表示原函数，那么通常用 $f(x)$ 表示导函数，有时也用 $F'(x)$ 表示导函数。

例 3.2.1. 已知函数 $f(x) = \frac{1}{x}$ ，求函数在点 $(1, f(1))$ 处的切线斜率，并求出通过该点的切线方程。

解：

首先求导函数 $f'(x)$:

$$\begin{aligned} f'(x) &= \frac{df(x)}{dx} \\ &= \frac{f(x + \epsilon) - f(x)}{(x + \epsilon) - x} \\ &= \frac{1/(x + \epsilon) - 1/x}{\epsilon} \\ &= \frac{x - (x + \epsilon)}{x(x + \epsilon)} \cdot \frac{1}{\epsilon} \\ &= \frac{-\epsilon}{x(x + \epsilon)} \cdot \frac{1}{\epsilon} \\ &= -\frac{1}{x^2 + x \cdot \epsilon} \\ &= -\frac{1}{x^2} \end{aligned}$$

有了导函数，现在我们可以计算函数在点 $(1, f(1)) = (1, 1)$ 处的切线斜率

$$f'(1) = -\frac{1}{(1)^2} = -1$$

切线方程 (点斜式):

$$\begin{aligned} y - y_0 &= m(x - x_0) \\ y - 1 &= (-1) \cdot (x - 1) \\ y &= -x + 1 + 1 \\ y &= -x + 2 \end{aligned}$$

常用导函数公式

求解导函数，需要同时进行代数运算和极限运算，从示例可以看出，即使是简单的原函数，求解导函数也相当繁琐。因此有必要推导一些常用函数的导函数形成导函数公式，避免重复进行极限的繁琐计算。

(1) 常数函数 $f(x) = C$

$$\frac{dC}{dx} = \frac{f(x + \epsilon) - f(x)}{\epsilon} = \frac{C - C}{\epsilon} = 0$$

所以，常数函数的导函数为 0，即

$$f'(x) = \frac{dC}{dx} = 0 \quad (3.3)$$

(2) 幂函数 $f(x) = x^n$

当 n 为整数时，考虑二项式定理

$$\begin{aligned} (x + \epsilon)^n &= \sum_{k=0}^n \binom{n}{k} x^{n-k} \epsilon^k \\ &= x^n + \binom{n}{1} x^{n-1} \epsilon + \binom{n}{2} x^{n-2} \epsilon^2 + \cdots + \epsilon^n \end{aligned}$$

$$\begin{aligned} \frac{dx^n}{dx} &= \frac{(x + \epsilon)^n - x^n}{\epsilon} \\ &= \frac{[x^n + nx^{n-1}\epsilon + \binom{n}{2}x^{n-2}\epsilon^2 + \cdots + \epsilon^n] - x^n}{\epsilon} \\ &= nx^{n-1} + \binom{n}{2}x^{n-2}\epsilon + \cdots + \epsilon^{n-1} \\ &= nx^{n-1} \end{aligned}$$

当 n 为一般实数时上面的公式也成立，通常需要依赖于指数和对数函数的导数公式，本节稍后作为例题进行推导。

所以，幂函数的导数公式

$$f'(x) = \frac{dx^n}{dx} = nx^{n-1} \quad (3.4)$$

(3) 指数函数 $f(x) = e^x$, $f(x) = a^x$

使用上一章的自然恒等式 $e^\epsilon = 1 + \epsilon$, $a^\epsilon = 1 + \ln a \cdot \epsilon$

$$\begin{aligned} \frac{de^x}{dx} &= \frac{e^{x+\epsilon} - e^x}{\epsilon} \\ &= \frac{e^x \cdot (e^\epsilon - 1)}{\epsilon} \end{aligned}$$

$$\begin{aligned}
&= e^x \cdot \frac{e^\epsilon - 1}{\epsilon} \\
&= e^x \cdot \frac{(1 + \epsilon) - 1}{\epsilon} \\
&= e^x \cdot 1 = e^x
\end{aligned}$$

$$\begin{aligned}
\frac{da^x}{dx} &= \frac{a^{x+\epsilon} - a^x}{\epsilon} \\
&= \frac{a^x \cdot (a^\epsilon - 1)}{\epsilon} \\
&= a^x \cdot \frac{a^\epsilon - 1}{\epsilon} \\
&= a^x \cdot \frac{(1 + \ln a \cdot \epsilon) - 1}{\epsilon} \\
&= a^x \ln a
\end{aligned}$$

所以，指数函数的导数公式

$$f'(x) = \frac{e^x}{dx} = e^x \quad (3.5)$$

$$f'(x) = \frac{a^x}{dx} = a^x \ln a \quad (3.6)$$

(4) 对数函数 $f(x) = \ln x$, $f(x) = \log_a x$

使用上一章的自然恒等式 $\ln(1 + a\epsilon) = a\epsilon$

$$\begin{aligned}
\frac{d \ln x}{dx} &= \frac{\ln(x + \epsilon) - \ln x}{\epsilon} \\
&= \ln \frac{x + \epsilon}{x} \cdot \frac{1}{\epsilon} \\
&= \ln \left(1 + \frac{1}{x} \cdot \epsilon\right) \cdot \frac{1}{\epsilon} \\
&= \frac{1}{x} \cdot \epsilon \cdot \frac{1}{\epsilon} \\
&= \frac{1}{x}
\end{aligned}$$

$$\begin{aligned}
\frac{d \log_a x}{dx} &= \frac{\log_a(x + \epsilon) - \log_a x}{\epsilon} \\
&= \log_a \frac{x + \epsilon}{x} \cdot \frac{1}{\epsilon} \\
&= \log_a \left(1 + \frac{\epsilon}{x}\right) \cdot \frac{1}{\epsilon} \\
&= \frac{\ln(1 + \frac{\epsilon}{x})}{\ln a} \cdot \frac{1}{\epsilon}
\end{aligned}$$

$$\begin{aligned} &= \frac{1}{\ln a} \cdot \frac{\epsilon}{x} \cdot \frac{1}{\epsilon} \\ &= \frac{1}{x \ln a} \end{aligned}$$

所以，对数函数的导数公式

$$f'(x) = \frac{d \ln x}{dx} = \frac{1}{x} \quad (3.7)$$

$$f'(x) = \frac{d \log_a x}{dx} = \frac{1}{x \ln a} \quad (3.8)$$

例 3.2.2. 推导幂函数 $f(x) = x^n$ 的导数公式。

$$\begin{aligned} \frac{dx^n}{dx} &= \frac{(x + \epsilon)^n - x^n}{\epsilon} \\ &= (e^{n \ln(x + \epsilon)} - e^{n \ln x}) \cdot \frac{1}{\epsilon} \\ &= (e^{n \ln x + n \ln(1 + \frac{\epsilon}{x})} - e^{n \ln x}) \cdot \frac{1}{\epsilon} \\ &= (e^{n \ln x + \frac{n\epsilon}{x}} - e^{n \ln x}) \cdot \frac{1}{\epsilon} \\ &= e^{n \ln x} \cdot (e^{\frac{n\epsilon}{x}} - 1) \cdot \frac{1}{\epsilon} \\ &= x^n \cdot \frac{(1 + \frac{n\epsilon}{x}) - 1}{\epsilon} \\ &= x^n \cdot \frac{n}{x} \\ &= nx^{n-1} \end{aligned}$$

(5) 三角函数 $f(x) = \sin x, f(x) = \cos x$

对于 $f(x) = \sin x$, 考虑使用和角公式 $\sin(x + \epsilon) = \sin x \cdot \cos \epsilon + \cos x \cdot \sin \epsilon$ 。

$$\begin{aligned} \frac{d \sin x}{dx} &= \frac{\sin(x + \epsilon) - \sin x}{\epsilon} \\ &= \frac{(\sin x \cdot \cos \epsilon + \cos x \cdot \sin \epsilon) - \sin x}{\epsilon} \\ &= \frac{\sin x(\cos \epsilon - 1) + \cos x \cdot \sin \epsilon}{\epsilon} \\ &= \sin x \cdot \frac{\cos \epsilon - 1}{\epsilon} + \cos x \cdot \frac{\sin \epsilon}{\epsilon} \\ &= \sin x \cdot 0 + \cos x \cdot 1 \\ &= \cos x \end{aligned}$$

对于 $f(x) = \cos x$, 考虑使用和角公式 $\cos(x + \epsilon) = \cos x \cos \epsilon - \sin x \sin \epsilon$ 。

$$\frac{d \cos x}{dx} = \frac{\cos(x + \epsilon) - \cos x}{\epsilon}$$

$$\begin{aligned}
&= \frac{\cos x \cos \epsilon - \sin x \sin \epsilon - \cos x}{\epsilon} \\
&= \frac{\cos x(\cos \epsilon - 1) - \sin x \sin \epsilon}{\epsilon} \\
&= \cos x \cdot \frac{\cos \epsilon - 1}{\epsilon} - \sin x \cdot \frac{\sin \epsilon}{\epsilon} \\
&= \cos x \cdot 0 - \sin x \cdot 1 \\
&= -\sin x
\end{aligned}$$

所以，三角函数的导数公式

$$f'(x) = \frac{d \sin x}{dx} = \cos x \quad (3.9)$$

$$f'(x) = \frac{d \cos x}{dx} = -\sin x \quad (3.10)$$

这些是常用的基本函数导数公式，其它直接推导相对复杂的函数，我们在后续章节中利用导数的运算法则进行推导。

基本函数导数公式

函数 $f(x)$	导数 $f'(x)$
C (常数)	0
x^n	nx^{n-1}
e^x	e^x
a^x	$a^x \ln a$
$\ln x$	$\frac{1}{x}$
$\log_a x$	$\frac{1}{x \ln a}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\sec^2 x$
$\cot x$	$-\csc^2 x$
$\sec x$	$\sec x \tan x$
$\csc x$	$-\csc x \cot x$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
$\arctan x$	$\frac{1}{1+x^2}$
$\text{arccot } x$	$-\frac{1}{1+x^2}$

3.3 导数的运算法则

假设 $u = u(x)$ 和 $v = v(x)$ 都是可导函数， C 是常数。

(1) 常数与函数积的导数 $(Cu)' = Cu'$

$$\frac{dCu(x)}{dx} = \frac{Cu(x + \epsilon) - Cu(x)}{\epsilon} = C \cdot \frac{u(x + \epsilon) - u(x)}{\epsilon} = C \cdot \frac{du(x)}{dx}$$

(2) 和或差的导数 $(u \pm v)' = u' \pm v'$

$$\begin{aligned}\frac{d[u(x) + v(x)]}{dx} &= \frac{[u(x + \epsilon) + v(x + \epsilon)] - [u(x) + v(x)]}{\epsilon} \\ &= \frac{u(x + \epsilon) - u(x)}{\epsilon} + \frac{v(x + \epsilon) - v(x)}{\epsilon} \\ &= \frac{du(x)}{dx} + \frac{dv(x)}{dx} = u' + v'\end{aligned}$$

例 3.3.1. 求 $f(x) = x^5 + \ln x$ 的导数。

$$f'(x) = (x^5)' + (\ln x)' = 5x^{5-1} + \frac{1}{x} = 5x^4 + \frac{1}{x}$$

(3) 积的导数 (乘法法则) $(uv)' = u'v + uv'$

积的导数可参考图 3.1 进行计算。

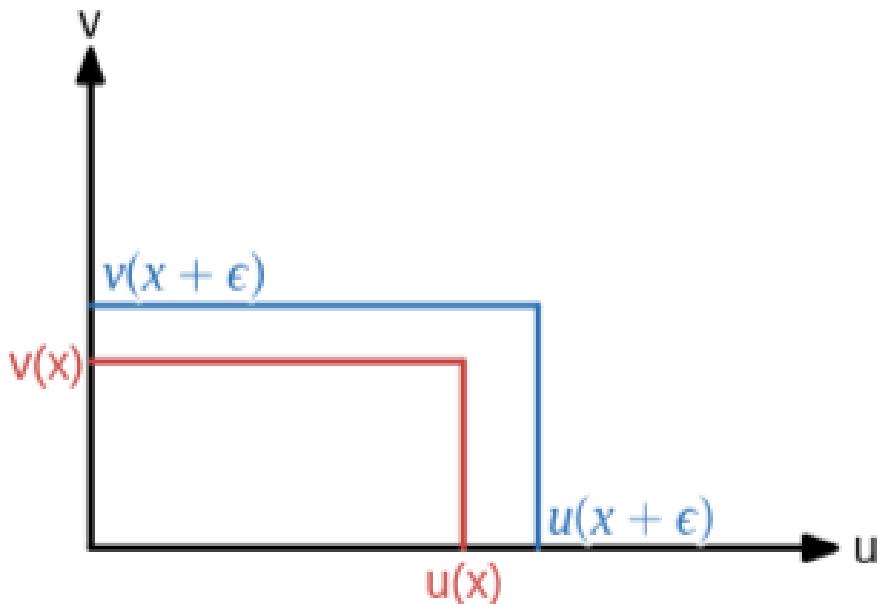


图 3.1: 积的导数

$$\begin{aligned}\frac{d[u(x)v(x)]}{dx} &= \frac{u(x + \epsilon)v(x + \epsilon) - u(x)v(x)}{\epsilon} \\ &= \frac{[u(x + \epsilon) - u(x)]v(x) + u(x)[v(x + \epsilon) - v(x)] + [u(x + \epsilon) - u(x)] \cdot [v(x + \epsilon) - v(x)]}{\epsilon} \\ &= \frac{du(x)}{dx} \cdot v(x) + u(x) \cdot \frac{dv(x)}{dx} + \frac{[u(x + \epsilon) - u(x)] \cdot [v(x + \epsilon) - v(x)]}{\epsilon^2} \cdot \epsilon\end{aligned}$$

$$\begin{aligned} &= u'(x)v(x) + u(x)v'(x) + u'(x)v'(x) \cdot \epsilon \\ &= u'v + uv' \end{aligned}$$

例 3.3.2. 求 $g(x) = x \cos x$ 的导数。

$$g'(x) = (x)' \cos x + x(\cos x)' = 1 \cdot \cos x + x(-\sin x) = \cos x - x \sin x$$

(4) 商的导数 (除法法则)

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \quad (v \neq 0)$$

仍然参考图 3.1 进行面积变换。

$$\begin{aligned} \left(\frac{u}{v}\right)' &= \left[\frac{u(x+\epsilon)}{v(x+\epsilon)} - \frac{u(x)}{v(x)} \right] \cdot \frac{1}{\epsilon} \\ &= \frac{u(x+\epsilon)v(x) - u(x)v(x+\epsilon)}{v(x+\epsilon)v(x) \cdot \epsilon} \\ &= \frac{[u(x+\epsilon) - u(x)] \cdot v(x) - u(x) \cdot [v(x+\epsilon) - v(x)]}{v(x+\epsilon)v(x) \cdot \epsilon} \\ &= \frac{u'(x) \cdot v(x) - u(x) \cdot v'(x)}{v(x+\epsilon)v(x)} \\ &= \frac{u'v - uv'}{v^2} \end{aligned}$$

例 3.3.3. 求 $f(x) = \tan x$ 的导数。

$$\begin{aligned} (\tan x)' &= \left(\frac{\sin x}{\cos x} \right)' = \frac{(\sin x)' \cdot \cos x - \sin x \cdot (\cos x)'}{\cos^2 x} \\ &= \frac{\cos^2 x - \sin x \cdot (-\sin x)}{\cos^2 x} \\ &= \frac{1}{\cos^2 x} = \sec^2 x \end{aligned}$$

所以

$$(\tan x)' = \sec^2 x \quad (3.11)$$

例 3.3.4. 求 $f(x) = \sec x$ 的导数。

$$\begin{aligned} (\sec x)' &= \left(\frac{1}{\cos x} \right)' = \frac{(1)' \cdot \cos x - (1) \cdot (\cos x)'}{\cos^2 x} \\ &= \frac{0 \cdot \cos x - (-\sin x)}{\cos^2 x} \\ &= \frac{1}{\cos x} \cdot \frac{\sin x}{\cos x} = \sec x \tan x \end{aligned}$$

所以

$$(\sec x)' = \sec x \tan x \quad (3.12)$$

(5) 链式法则 (复合函数求导)

如果 $y = f(u)$ 且 $u = g(x)$, 那么 y 对 x 的导数:

$$\frac{dy}{dx} = \frac{df(u)}{du} \cdot \frac{du}{dx} \quad \text{或} \quad [f(g(x))]' = f'(u) \cdot g'(x)$$

链式法则是除法运算的简单算术变换, 可以用相对速度的概念进行理解。

例 3.3.5. 求 $h(x) = e^{2x}$ 的导数。

令 $u = 2x$, 则 $h(x) = e^u$ 。

$$\frac{dh}{dx} = \frac{d(e^u)}{du} \cdot \frac{d(2x)}{dx} = e^u \cdot 2 = 2e^{2x}$$

(6) 反函数的导数

假设 f 是一个可导的单调函数, 并且 $u(x) = f(x)$ 有反函数 $v(x) = f^{-1}(x)$, 那么 $u(x)$ 和 $v(x)$ 互为反函数。我们想要推导出 $v'(x)$ 用 $u'(x)$ 表示的公式。

首先, 由反函数的定义, 知道

$$u(v(x)) = x$$

等式两边关于 x 求导 (使用链式法则):

$$[u(v(x))]' = (x)'$$

左边应用链式法则, 右边直接求导得到 1。所以有:

$$u'(v) \cdot v'(x) = 1$$

将 $v'(x)$ 单独提出来:

$$v'(x) = \frac{1}{u'(v)}$$

所以如果 $u(x), v(x)$ 互为反函数, 则用反函数表示的导数公式为:

$$v'(x) = \frac{1}{u'(v)} \tag{3.13}$$

例 3.3.6. 假设 $f(x) = x^3$, 那么反函数为 $f(x) = x^{1/3}$ 。求反函数的导数。

首先, 表示为标准公式形式, $u(x) = x^3, v(x) = x^{1/3}$

那么,

$$u'(x) = 3x^2, u'(v) = 3[x^{1/3}]^2 = 3x^{2/3}$$

反函数的导数

$$v'(x) = \frac{1}{u'(v)} = \frac{1}{3x^{2/3}}$$

例 3.3.7. 求 $\arctan x$ 的导数。

首先, 表示为标准公式形式, $v(x) = \arctan x$, $u(x) = \tan x$

那么, $u(v) = \tan(\arctan x) = x$,

$$\begin{aligned} u'(x) &= \sec^2 x = 1 + \tan^2 x = 1 + u^2(x) \\ u'(v) &= 1 + u^2(v) = 1 + x^2 \\ v'(x) &= \frac{1}{u'(v)} = \frac{1}{1 + x^2} \end{aligned}$$

函数 $\arctan x$ 的导数

$$(\arctan x)' = \frac{1}{1 + x^2} \quad (3.14)$$

(7) 隐函数求导

如果一个函数关系不能写成 $y = f(x)$ (或 $z = f(x, y)$ 等) 的显式形式, 而是以一个包含自变量和因变量的方程 $F(x, y) = 0$ (或 $F(x, y, z) = 0$ 等) 给出, 那么 y (或 z) 就是由该方程所确定的 x (或 x, y) 的隐函数。

例如: $x^2 + y^2 = 1$ $e^y + xy = 5$

对于由方程 $F(x, y) = 0$ 确定的隐函数 $y = y(x)$, 求导的基本思路是利用链式法则, 将 y 视为 x 的函数, 对原方程两边同时关于 x 求导。

隐函数求导步骤

(1) 视为复合函数: 将方程 $F(x, y) = 0$ 中的因变量 y 看作是自变量 x 的函数, 即 $y = y(x)$ 。

(2) 两边求导: 对原方程 $F(x, y) = 0$ 的等号两边同时关于 x 求导。

在求导过程中, 所有包含 y 的项 (如 $y^2, xy, \sin(y)$ 等) 都必须使用链式法则。例如,

$$\frac{d(y^2)}{dx} = 2y \cdot \frac{dy}{dx};$$

$\frac{d(xy)}{dx}$ 需使用乘积法则和链式法则:

$$\frac{d(x)}{dx} \cdot y + x \cdot \frac{d(y)}{dx} = 1 \cdot y + x \cdot \frac{dy}{dx}.$$

(3) 分离导数: 整理求导后的方程, 将含有 $\frac{dy}{dx}$ (或 y') 的项移到方程的一边, 不含 $\frac{dy}{dx}$ 的项移到另一边。

(4) 解出导数: 解出 $\frac{dy}{dx}$ 的表达式。

例 3.3.8. 求由方程 $x^2 + y^2 = 25$ 确定的隐函数 y 的导数 $\frac{dy}{dx}$ 。

两边求导: 对 $x^2 + y^2 = 25$ 两边关于 x 求导:

$$\frac{d(x^2)}{dx} + \frac{d(y^2)}{dx} = \frac{d(25)}{dx}$$

$\frac{d(x^2)}{dx} = 2x$, $\frac{d(y^2)}{dx} = 2y \cdot \frac{dy}{dx}$ (链式法则), $\frac{d(25)}{dx} = 0$, 得到方程:

$$2x + 2y \cdot \frac{dy}{dx} = 0$$

分离导数: 将含 $\frac{dy}{dx}$ 的项保留, 其余项移项:

$$2y \cdot \frac{dy}{dx} = -2x$$

解出导数:

$$\frac{dy}{dx} = -\frac{2x}{2y} = -\frac{x}{y}$$

例 3.3.9. 求 $\arcsin x$ 的导数。

设 $y = \arcsin x$, 对等式两边取 \sin :

$$\sin y = x$$

对等式两边关于 x 求导:

$$\begin{aligned}\frac{d(\sin y)}{dx} &= \frac{d(x)}{dx} \\ \cos y \cdot \frac{dy}{dx} &= 1\end{aligned}$$

解出 $\frac{dy}{dx}$:

$$\frac{dy}{dx} = \frac{1}{\cos y}$$

现在我们需要用 x 来表示 $\cos y$ 。我们使用三角恒等式 $\sin^2 y + \cos^2 y = 1$:

$$\cos^2 y = 1 - \sin^2 y$$

$$\cos y = \pm \sqrt{1 - \sin^2 y}$$

因为我们最初定义 $\sin y = x$, 所以代入:

$$\cos y = \pm \sqrt{1 - x^2}$$

$\arcsin x$ 的值域是 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 。在这个区间内, $\cos y$ 总是非负数 ($\cos y \geq 0$)。因此, 我们只取正根:

$$\cos y = \sqrt{1 - x^2}$$

将此结果代回:

$$\frac{dy}{dx} = \frac{1}{\sqrt{1 - x^2}}$$

函数 $\arcsin x$ 的导数

$$(\arcsin x)' = \frac{1}{\sqrt{1 - x^2}} \quad (3.15)$$

3.4 参数方程求导

当一个函数 y 不直接表示为 x 的函数 $y = f(x)$, 而是通过第三个变量 (参数 t) 来表示 x 和 y 时, 即:

$$\begin{cases} x = f(t) \\ y = g(t) \end{cases}$$

其中 t 是参数。我们要求的是 $\frac{dy}{dx}$, 即 y 对 x 的导数。

根据链式法则 (Chain Rule), 我们可以将 $\frac{dy}{dx}$ 表示为关于参数 t 的导数的比值:

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} \quad (3.16)$$

其中:

$\frac{dy}{dt}$ 是 y 对参数 t 的导数, 即 $g'(t)$ 。

$\frac{dx}{dt}$ 是 x 对参数 t 的导数, 即 $f'(t)$ 。

注意: 进行求导的前提是 $\frac{dx}{dt} \neq 0$ 。

例 3.4.1. 求参数方程 $\begin{cases} x = 2t + 1 \\ y = t^2 \end{cases}$ 的导数 $\frac{dy}{dx}$ 。

步骤 1: 分别求 $\frac{dx}{dt}$ 和 $\frac{dy}{dt}$

$$\begin{aligned} \frac{dx}{dt} &= \frac{d}{dt}(2t + 1) = 2 \\ \frac{dy}{dt} &= \frac{d}{dt}(t^2) = 2t \end{aligned}$$

步骤 2: 利用公式求 $\frac{dy}{dx}$

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{2t}{2} = t$$

例 3.4.2. 由参数方程定义的圆:

$$\begin{cases} x = 3 \cos t \\ y = 3 \sin t \end{cases}$$

求出当参数 $t = \frac{\pi}{4}$ 时, 该曲线 (圆) 的切线斜率。

确定切点坐标:

$$x_0 = 3 \cos\left(\frac{\pi}{4}\right) = 3 \cdot \frac{\sqrt{2}}{2} = \frac{3\sqrt{2}}{2}$$

$$y_0 = 3 \sin\left(\frac{\pi}{4}\right) = 3 \cdot \frac{\sqrt{2}}{2} = \frac{3\sqrt{2}}{2}$$

切点是 $\left(\frac{3\sqrt{2}}{2}, \frac{3\sqrt{2}}{2}\right)$ 。

分别对 x 和 y 关于参数 t 求导并计算导数 $\frac{dy}{dx}$:

$$\begin{aligned}\frac{dx}{dt} &= \frac{d}{dt}(3 \cos t) = -3 \sin t \\ \frac{dy}{dt} &= \frac{d}{dt}(3 \sin t) = 3 \cos t \\ \frac{dy}{dx} &= \frac{3 \cos t}{-3 \sin t} = -\frac{\cos t}{\sin t} = -\cot t\end{aligned}$$

结果 $\frac{dy}{dx} = -\cot t$ 是曲线的切线斜率关于参数 t 的表达式。

将 $t = \frac{\pi}{4}$ 代入 $\frac{dy}{dx}$ 的表达式:

$$m = \left. \frac{dy}{dx} \right|_{t=\frac{\pi}{4}} = -\cot\left(\frac{\pi}{4}\right) = -1$$

当 $t = \frac{\pi}{4}$ 时, 该参数曲线 (圆) 在点 $\left(\frac{3\sqrt{2}}{2}, \frac{3\sqrt{2}}{2}\right)$ 处的切线斜率为 -1 。

摆线是圆在平面上滚动时, 圆周上一点的轨迹, 它在微积分中非常经典。

摆线的参数方程通常定义为 (假设圆的半径为 R):

$$\begin{cases} x = R(\theta - \sin \theta) \\ y = R(1 - \cos \theta) \end{cases}$$

其中 θ 是参数, 代表圆滚动的角度。

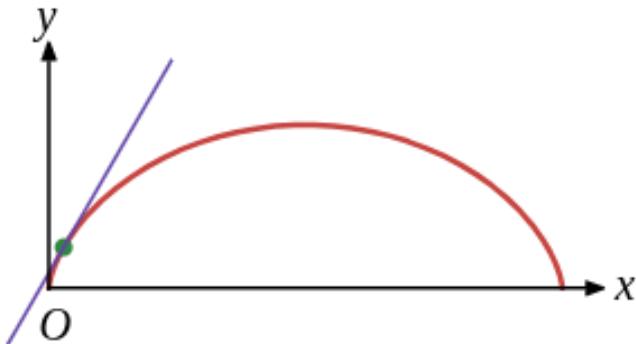


图 3.2: 摆线在对应点的切线斜率

我们来求当 $\theta = \frac{\pi}{3}$ 时, 摆线在对应点的切线斜率。

$$\begin{aligned}\frac{dx}{d\theta} &= \frac{dR(\theta - \sin \theta)}{d\theta} = R(1 - \cos \theta) \\ \frac{dy}{d\theta} &= \frac{dR(1 - \cos \theta)}{d\theta} = R(0 - (-\sin \theta)) = R \sin \theta \\ \frac{dy}{dx} &= \frac{R \sin \theta}{R(1 - \cos \theta)} = \frac{\sin \theta}{1 - \cos \theta}\end{aligned}$$

将 $\theta = \frac{\pi}{3}$ 代入 $\frac{dy}{dx}$ 的表达式：

$$\sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}, \cos\left(\frac{\pi}{3}\right) = \frac{1}{2}$$

$$m = \left. \frac{dy}{dx} \right|_{\theta=\frac{\pi}{3}} = \frac{\sin \frac{\pi}{3}}{1 - \cos \frac{\pi}{3}} = \frac{\frac{\sqrt{3}}{2}}{1 - \frac{1}{2}} = \frac{\frac{\sqrt{3}}{2}}{\frac{1}{2}} = \sqrt{3}$$

当 $\theta = \frac{\pi}{3}$ 时，摆线在对应点的切线斜率为 $\sqrt{3}$ 。

3.5 高阶导数

高阶导数 (Higher-order derivative) 是指一个函数的导数的导数。

一阶导数：函数 $f(x)$ 的导数 $f'(x)$ 。它表示函数 $f(x)$ 随变量 x 的变化率。

二阶导数：如果一阶导数 $f'(x)$ 仍然可导，那么 $f'(x)$ 的导数就是函数 $f(x)$ 的二阶导数，记作 $f''(x)$ 或 $f^{(2)}(x)$ 。它表示函数变化率的变化率（即曲率或在物理学中的加速度）。

n 阶导数：一般地，如果一个函数的 $(n-1)$ 阶导数是可导的，那么其导数就是该函数的 n 阶导数，记作 $f^{(n)}(x)$ 。

二阶及二阶以上的导数统称为高阶导数。

如果 $s(t)$ 是物体在时刻 t 的位移，则：

一阶导数 $s'(t)$ 是速度 $v(t)$ 。

二阶导数 $s''(t)$ 是加速度 $a(t)$ 。

三阶导数 $s'''(t)$ 是加加速度 (Jerk)。

二阶导数可以用来判断函数图形的凹凸性：

若 $f''(x) > 0$ ，则函数图形在该区间上是凹的（向上弯曲）。

若 $f''(x) < 0$ ，则函数图形在该区间上是凸的（向下弯曲）

高阶导数是构造函数的泰勒级数 (Taylor series) 的关键。

高阶导数保证了多项式在局部与原函数在每一个细节（斜率、弯曲度、更复杂的弯曲变化）上都吻合，实现了在局部区间内对原函数的最优多项式逼近。

3.6 泰勒级数和麦克劳林级数

我们在上一章的末尾展示了泰勒级数在计算极限时的强大应用。

对于一个在某点 a 处无限可导的函数 $f(x)$ ，我们希望用一个多项式来近似表示它，并且要求这个多项式在 $x = a$ 及其附近与原函数 $f(x)$ 有相同的值和相同的

各阶导数值。这个多项式就是泰勒多项式 (Taylor Polynomial)，当其项数趋于无穷时，就变成了泰勒级数。

泰勒级数的推导

我们假设要找一个 n 次多项式 $P_n(x)$ 来近似 $f(x)$ ，以 $x = a$ 为中心点：

$$P_n(x) = c_0 + c_1(x - a) + c_2(x - a)^2 + c_3(x - a)^3 + \cdots + c_n(x - a)^n$$

目标是确定系数 $c_0, c_1, c_2, \dots, c_n$ 。

(1) 确定 c_0

我们要求多项式在 $x = a$ 处的值与函数值相等：

$$P_n(a) = f(a)$$

代入 $P_n(x)$ 的表达式：

$$\begin{aligned} P_n(a) &= c_0 + c_1(a - a) + c_2(a - a)^2 + \cdots \\ P_n(a) &= c_0 \end{aligned}$$

因此，第一个系数 c_0 为：

$$c_0 = f(a)$$

(2) 确定 c_1

我们要求多项式在 $x = a$ 处的一阶导数值与函数的一阶导数值相等：

$$P'_n(a) = f'(a)$$

先求 $P_n(x)$ 的一阶导数：

$$P'_n(x) = c_1 + 2c_2(x - a) + 3c_3(x - a)^2 + 4c_4(x - a)^3 + \cdots$$

代入 $x = a$ ：

$$\begin{aligned} P'_n(a) &= c_1 + 2c_2(0) + 3c_3(0) + \cdots \\ P'_n(a) &= c_1 \end{aligned}$$

因此，第二个系数 c_1 为：

$$c_1 = f'(a) = \frac{f'(a)}{1!}$$

(3) 确定 c_2

我们要求多项式在 $x = a$ 处的二阶导数值与函数的二阶导数值相等：

$$P''_n(a) = f''(a)$$

先求 $P'_n(x)$ 的导数，即 $P_n(x)$ 的二阶导数：

$$P''_n(x) = 2c_2 + 2 \cdot 3c_3(x - a) + 3 \cdot 4c_4(x - a)^2 + \dots$$

代入 $x = a$ ：

$$\begin{aligned} P''_n(a) &= 2c_2 + 0 + 0 + \dots \\ P''_n(a) &= 2c_2 \end{aligned}$$

因此，第三个系数 c_2 为：

$$c_2 = \frac{f''(a)}{2} = \frac{f''(a)}{2!}$$

(4) 确定 c_k (一般项)

通过归纳法，我们可以看到一般的规律。对于第 k 阶导数 $P_n^{(k)}(x)$ ，当 $x = a$ 时，所有包含 $(x - a)$ 因子的项都为零，只剩下常数项：

$$P_n^{(k)}(x) = k!c_k + \text{包含}(x - a)\text{的项}$$

因此，我们要求：

$$\begin{aligned} P_n^{(k)}(a) &= f^{(k)}(a) \\ k!c_k &= f^{(k)}(a) \end{aligned}$$

解出系数 c_k ：

$$c_k = \frac{f^{(k)}(a)}{k!}$$

将所有系数 c_k 代回原来的多项式表达式，并将项数 n 延伸至无穷大，就得到了泰勒级数的表达式。

对于一个在点 a 处无限可导的函数 $f(x)$ ，其泰勒级数为：

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n \quad (3.17)$$

$$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \frac{f'''(a)}{3!}(x - a)^3 + \dots \quad (3.18)$$

a : 级数的中心点或展开点。级数在离 a 越近的地方，近似效果越好。

$f^{(n)}(a)$: 函数 $f(x)$ 的 n 阶导数在点 $x = a$ 处的值。

$n!$: n 的阶乘。

麦克劳林级数

麦克劳林级数是泰勒级数的一个特殊情况，即展开点 $a = 0$ 时的泰勒级数。

麦克劳林级数是函数 $f(x)$ 在 $x = 0$ 处的泰勒级数：

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} (x)^n \quad (3.19)$$

$$f(x) = f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \frac{f'''(0)}{3!} x^3 + \dots \quad (3.20)$$

泰勒级数本质上是利用函数在一点 (a) 处的值和各阶导数值来构造一个无限阶多项式，使该多项式能够完全匹配原函数在该点处的局部行为。

例 3.6.1. 函数 $f(x) = e^x$ 的麦克劳林级数

(1) 计算函数及其各阶导数

我们需要计算函数 $f(x) = e^x$ 在 $x = 0$ 处的各阶导数值 $f^{(k)}(0)$ 。

k	$f^{(k)}(x)$ (第 k 阶导数)	$f^{(k)}(0)$ (在 $x = 0$ 处的值)
0	$f(x) = e^x$	$f(0) = e^0 = 1$
1	$f'(x) = e^x$	$f'(0) = e^0 = 1$
2	$f''(x) = e^x$	$f''(0) = e^0 = 1$
3	$f'''(x) = e^x$	$f'''(0) = e^0 = 1$
k	$f^{(k)}(x) = e^x$	$f^{(k)}(0) = e^0 = 1$

可以看出，对于函数 $f(x) = e^x$ ，在 $x = 0$ 处的任意阶导数值都等于 1，即 $f^{(k)}(0) = 1$ 。

(2) 代入麦克劳林级数公式

将计算得到的导数值 $f^{(k)}(0) = 1$ 代入麦克劳林级数公式：

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$$

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

(3) 展开级数

将求和符号展开，列出级数的前几项：

$$e^x = \frac{f^{(0)}(0)}{0!} x^0 + \frac{f^{(1)}(0)}{1!} x^1 + \frac{f^{(2)}(0)}{2!} x^2 + \frac{f^{(3)}(0)}{3!} x^3 + \dots$$

$$e^x = \frac{1}{0!} x^0 + \frac{1}{1!} x^1 + \frac{1}{2!} x^2 + \frac{1}{3!} x^3 + \dots$$

由于 $0! = 1$ 且 $x^0 = 1$ (对于 $x \neq 0$)，我们得到最终的麦克劳林级数 (即 e^x 的泰勒级数)：

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots + \frac{x^n}{n!} + \dots \quad (3.21)$$

如果想以任意点 a 为中心展开 $f(x) = e^x$, 则各阶导数值: $f^{(k)}(a) = e^a$ 。泰勒级数:

$$\begin{aligned} e^x &= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k \\ e^x &= \sum_{k=0}^{\infty} \frac{e^a}{k!} (x-a)^k \\ e^x &= e^a + \frac{e^a}{1!}(x-a) + \frac{e^a}{2!}(x-a)^2 + \frac{e^a}{3!}(x-a)^3 + \dots \end{aligned}$$

可以看出, 这个形式是 (e^a) 乘以 $\left(1 + (x-a) + \frac{(x-a)^2}{2!} + \dots\right)$, 这与 e^{x-a} 的麦克劳林级数相乘, 符合指数函数的性质 $e^x = e^a \cdot e^{x-a}$ 。

例 3.6.2. 函数 $f(x) = \sin x$ 的麦克劳林级数

(1) 计算函数及其各阶导数在 $x = 0$ 处的值

我们需要计算 $f(x) = \sin x$ 在 $x = 0$ 处的各阶导数的值 $f^{(n)}(0)$:

第 0 阶:	$f(x) = \sin x$	$f(0) = \sin(0) = 0$
第 1 阶:	$f'(x) = \cos x$	$f'(0) = \cos(0) = 1$
第 2 阶:	$f''(x) = -\sin x$	$f''(0) = -\sin(0) = 0$
第 3 阶:	$f'''(x) = -\cos x$	$f'''(0) = -\cos(0) = -1$
第 4 阶:	$f^{(4)}(x) = \sin x$	$f^{(4)}(0) = \sin(0) = 0$
第 5 阶:	$f^{(5)}(x) = \cos x$	$f^{(5)}(0) = \cos(0) = 1$

(2) 观察周期规律

我们可以看到 $f^{(n)}(0)$ 的值以 $0, 1, 0, -1$ 为周期重复出现:

n	$f^{(n)}(x)$	$f^{(n)}(0)$
0	$\sin x$	0
1	$\cos x$	1
2	$-\sin x$	0
3	$-\cos x$	-1
4	$\sin x$	0
5	$\cos x$	1
\dots	\dots	\dots

这意味着只有当 n 是奇数 ($n = 2k+1$) 时, 系数才非零。

当 $n = 1, 5, 9, \dots$ 时, $f^{(n)}(0) = 1$

当 $n = 3, 7, 11, \dots$ 时, $f^{(n)}(0) = -1$

(3) 将计算出的 $f^{(n)}(0)$ 值代入麦克劳林级数公式:

$$\begin{aligned} f(x) &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \frac{f^{(5)}(0)}{5!}x^5 + \dots \\ \sin x &= 0 + 1 \cdot x + \frac{0}{2!}x^2 + \frac{-1}{3!}x^3 + \frac{0}{4!}x^4 + \frac{1}{5!}x^5 + \dots \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \dots \end{aligned}$$

(4) 写出级数的求和形式

该级数只包含奇数次幂项, 且符号正负交替, 我们可以用一个求和符号来表示它:

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

这就是函数 $f(x) = \sin x$ 的麦克劳林级数。

常用的麦克劳林级数

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \quad (3.22)$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (3.23)$$

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \quad (3.24)$$

$$\tan x = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \frac{62x^9}{2835} + \dots \quad (3.25)$$

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots \quad (3.26)$$

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad (3.27)$$

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} x^{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots \quad (3.28)$$

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!} x^2 + \dots \quad (3.29)$$

例 3.6.3. 计算极限

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\cos(x) - 1 + \frac{1}{2}x^2}{x^4} &= \frac{\cos(\epsilon) - 1 + \frac{1}{2}\epsilon^2}{\epsilon^4} \\ &= \frac{\left(1 - \frac{\epsilon^2}{2!} + \frac{\epsilon^4}{4!}\right) - 1 + \frac{1}{2}\epsilon^2}{\epsilon^4} \end{aligned}$$

$$= \frac{\frac{\epsilon^4}{4!}}{\epsilon^4} = \frac{1}{4!}$$

例 3.6.4. 使用泰勒展开式来近似计算 $\sqrt{2}$ 。

我们希望近似计算 $f(x) = \sqrt{x}$ 在 $x = 2$ 处的值。

为了使计算简化且收敛更快，我们应该选择一个离 2 近且容易计算其函数值和导数值的点 a 。

最佳选择是 $a = 1$ 或 $a = 4$ ，因为 $\sqrt{1} = 1$ 和 $\sqrt{4} = 2$ 都是整数。我们选择 $a = 1$ 。

所以，我们对 $f(x) = \sqrt{x}$ 在 $a = 1$ 处进行泰勒展开，并令 $x = 2$ 进行计算。

我们需要计算函数 $f(x)$ 及其各阶导数在 $a = 1$ 处的值：

阶数 n	导数 $f^{(n)}(x)$	$f^{(n)}(1)$
0	$f(x) = x^{1/2}$	$f(1) = 1$
1	$f'(x) = \frac{1}{2}x^{-1/2}$	$f'(1) = \frac{1}{2}$
2	$f''(x) = -\frac{1}{4}x^{-3/2}$	$f''(1) = -\frac{1}{4}$
3	$f'''(x) = \frac{3}{8}x^{-5/2}$	$f'''(1) = \frac{3}{8}$
4	$f^{(4)}(x) = -\frac{15}{16}x^{-7/2}$	$f^{(4)}(1) = -\frac{15}{16}$

将 $f(x) = \sqrt{x}$ 和 $a = 1$ 代入公式：

$$\begin{aligned}\sqrt{x} &= f(1) + \frac{f'(1)}{1!}(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3 + \dots \\ \sqrt{x} &= 1 + \frac{1/2}{1!}(x-1) + \frac{-1/4}{2!}(x-1)^2 + \frac{3/8}{3!}(x-1)^3 + \dots \\ \sqrt{x} &= 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2 + \frac{3/8}{6}(x-1)^3 - \dots \\ \sqrt{x} &= 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2 + \frac{1}{16}(x-1)^3 - \dots\end{aligned}$$

将 $x = 2$ 代入展开式，则 $x - 1 = 1$:

$$\begin{aligned}\sqrt{2} &\approx 1 + \frac{1}{2}(1) - \frac{1}{8}(1)^2 + \frac{1}{16}(1)^3 - \dots \\ \sqrt{2} &\approx 1 + \frac{1}{2} - \frac{1}{8} + \frac{1}{16} - \dots\end{aligned}$$

我们使用三阶泰勒多项式 $P_3(2)$ 来近似：

$$\begin{aligned}P_3(2) &= 1 + \frac{1}{2} - \frac{1}{8} + \frac{1}{16} \\ &= \frac{16}{16} + \frac{8}{16} - \frac{2}{16} + \frac{1}{16} \\ &= \frac{16 + 8 - 2 + 1}{16}\end{aligned}$$

$$= \frac{23}{16} = 1.4375$$

泰勒近似值 $P_3(2)$: 1.4375

$\sqrt{2}$ 的真实值 (精确到五位小数): 1.41421

可以看到, 三阶泰勒展开式已经给出了一个相对接近的近似值。如果我们计算更高阶的项, 近似值会更精确。例如, 使用四阶近似 $P_4(2)$ (加上 $-\frac{5}{128}$) 得到:

$$\frac{23}{16} - \frac{5}{128} = \frac{184}{128} - \frac{5}{128} = \frac{179}{128} \approx 1.3984375$$

虽然四阶的近似值反而比三阶的更差, 这是因为泰勒展开的误差项并非单调的, 并且 $x = 2$ 离展开中心 $a = 1$ 较远, 导致展开式收敛速度较慢。

更好的方法是使用二项式级数:

这是一个二项式级数 (Binomial Series) 的展开, 它就是 $\sqrt{1+u}$ 在 $u = 0$ 处的麦克劳林展开 ($a = 1$ 处的泰勒展开):

$$\sqrt{x} = \sqrt{1 + (x - 1)}$$

令 $u = x - 1$, 则

$$\sqrt{1+u} = 1 + \frac{1}{2}u - \frac{1}{8}u^2 + \frac{1}{16}u^3 - \frac{5}{128}u^4 + \dots$$

代入 $x = 2$, 得 $u = 1$, 所以:

$$\sqrt{2} = 1 + \frac{1}{2} - \frac{1}{8} + \frac{1}{16} - \frac{5}{128} + \frac{7}{256} - \frac{21}{1024} + \dots$$

这个级数收敛于 $\sqrt{2}$, 但由于 $u = 1$ 是级数的收敛半径 $|u| < 1$ 的边界, 所以收敛速度较慢。

3.7 欧拉公式

如果 θ 是实数, 即某个角度的弧度度量, 则复指数函数 $e^{i\theta}$ 定义为

$$e^{i\theta} = \cos \theta + i \sin \theta \tag{3.30}$$

这是最著名的欧拉公式, 它建立了自然对数的底数 e 、虚数单位 i 和三角函数之间的深刻联系。

在复平面上, 欧拉公式 $e^{i\theta}$ 表示一个模长为 1 的复数, 其位置在单位圆上, 并且与正实轴的夹角为 θ (逆时针为正)。

欧拉公式可以通过将麦克劳林级数 e^x 中的 x 正式替换为 $i\theta$ 并推导得出

$$\begin{aligned} e^{i\theta} &= 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \frac{(i\theta)^6}{6!} + \dots \\ &= 1 + i\theta - \frac{\theta^2}{2!} - i\frac{\theta^3}{3!} + \frac{\theta^4}{4!} + i\frac{\theta^5}{5!} - \frac{\theta^6}{6!} - \dots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots\right) + i\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \dots\right) \\ &= \cos\theta + i\sin\theta \end{aligned}$$

其中最后一步来自 $\cos\theta$ 和 $\sin\theta$ 的麦克劳林级数。

欧拉公式是复数从直角坐标形式 $x + iy$ 转换到极坐标形式 $re^{i\theta}$ 的关键：

$$z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}$$

欧拉恒等式 (Euler's Identity)

当我们在欧拉公式中取 $\theta = \pi$ (即 180°) 时, 得到一个被誉为“数学中最美的公式”的恒等式:

$$e^{i\pi} = \cos\pi + i\sin\pi$$

因为 $\cos\pi = -1$ 且 $\sin\pi = 0$, 所以:

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

这个恒等式以简洁的形式将数学中五个最重要的常数 $e, i, \pi, 1, 0$ 以及三种基本运算 (加法、乘法、指数) 完美地结合在一起。

例 3.7.1. 证明和角公式

$$\cos(a + b) = \cos a \cos b - \sin a \sin b \quad \sin(a + b) = \cos a \sin b + \sin a \cos b$$

要证明和角公式, 我们可以利用欧拉公式 $e^{i\theta} = \cos\theta + i\sin\theta$

$$e^{i(a+b)} = \cos(a + b) + i\sin(a + b)$$

应用欧拉公式于 a 和 b

$$e^{ia} = \cos a + i\sin a \quad e^{ib} = \cos b + i\sin b$$

将这两个表达式相乘:

$$e^{i(a+b)} = e^{ia} \cdot e^{ib} = (\cos a + i\sin a)(\cos b + i\sin b)$$

$$\begin{aligned} e^{i(a+b)} &= \cos a \cos b + i \cos a \sin b + i \sin a \cos b + i^2 \sin a \sin b \\ e^{i(a+b)} &= (\cos a \cos b - \sin a \sin b) + i(\cos a \sin b + \sin a \cos b) \end{aligned}$$

与欧拉公式的形式对比: $e^{i(a+b)} = \cos(a+b) + i \sin(a+b)$

实部:

$$\cos(a+b) = \cos a \cos b - \sin a \sin b \quad (3.31)$$

虚部:

$$\sin(a+b) = \cos a \sin b + \sin a \cos b \quad (3.32)$$

由此, 我们证明了余弦(实部)和正弦(虚部)和角公式。

这个证明利用了复数的性质和欧拉公式, 将三角函数的和角问题简化到了复数乘法的框架内。

例 3.7.2. $a \cos x + b \sin x$ 转换为复指数形式。

欧拉公式定义了复指数 $e^{i\theta}$ 和三角函数之间的关系:

$$\begin{aligned} e^{i\theta} &= \cos \theta + i \sin \theta \\ e^{-i\theta} &= \cos(-\theta) + i \sin(-\theta) = \cos \theta - i \sin \theta \end{aligned}$$

将上面两个公式左右同时相加和相减, 分别得到 $\cos \theta$ 和 $\sin \theta$ 复指数表示:

$$\begin{aligned} e^{i\theta} + e^{-i\theta} &= (\cos \theta + i \sin \theta) + (\cos \theta - i \sin \theta) \\ e^{i\theta} + e^{-i\theta} &= 2 \cos \theta \end{aligned}$$

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}$$

$$\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

$a \cos x + b \sin x$ 的复指数表示:

$$\begin{aligned} a \cos x + b \sin x &= a \left(\frac{e^{ix} + e^{-ix}}{2} \right) + b \left(\frac{e^{ix} - e^{-ix}}{2i} \right) \\ &= \frac{a}{2}(e^{ix} + e^{-ix}) + \frac{b}{2i}(e^{ix} - e^{-ix}) \\ &= \frac{a}{2}e^{ix} + \frac{a}{2}e^{-ix} - \frac{bi}{2}e^{ix} + \frac{bi}{2}e^{-ix} \end{aligned}$$

$$\begin{aligned}
 &= \left(\frac{a}{2} - \frac{bi}{2} \right) e^{ix} + \left(\frac{a}{2} + \frac{bi}{2} \right) e^{-ix} \\
 &= \left(\frac{a - bi}{2} \right) e^{ix} + \left(\frac{a + bi}{2} \right) e^{-ix}
 \end{aligned}$$

如果复数：

$$c = \frac{a - bi}{2}$$

那么其共轭复数：

$$\bar{c} = \frac{a + bi}{2}$$

事实上，上例完成了傅立叶三角级数到复指数级数的变换。

3.8 微分公式和运算法则

如果导数已知，那么通过变换可以得到原函数的微分公式，即

$$f'(x) = \frac{df(x)}{dx} \iff df(x) = f'(x) dx \quad (3.33)$$

因为 $dx = \epsilon \neq 0$ ，所以式 3.21 的变换恒成立。

函数在单一定义域区间内总是可微（可导）的，在闭区间的左边界向右可微（可导），在闭区间的右边界向左可微（可导）。

根据前面的基本函数导数公式，很容易得到微分公式。

基本函数微分公式

$dF(x)$	$f'(x) dx$
dC (常数)	0
dx^n	$nx^{n-1} dx$
de^x	$e^x dx$
da^x	$a^x \ln a dx$
$d \ln x$	$\frac{1}{x} dx$
$d \log_a x$	$\frac{1}{x \ln a} dx$
$d \sin x$	$\cos x dx$
$d \cos x$	$-\sin x dx$
$d \tan x$	$\sec^2 x dx$
$d \cot x$	$-\csc^2 x dx$
$x \rightarrow 0 \sin x dx$	$\sec x \tan x dx$
$d \csc x$	$-\csc x \cot x dx$
$d \arcsin x$	$\frac{1}{\sqrt{1-x^2}} dx$
$d \arccos x$	$-\frac{1}{\sqrt{1-x^2}} dx$
$d \arctan x$	$\frac{1}{1+x^2} dx$
$d \operatorname{arccot} x$	$-\frac{1}{1+x^2} dx$

类似地，我们可以通过导数运算法则得到微分运算法则：

令 $u = u(x)$, $v = v(x)$, 那么

$$dCu = Cdu = Cu' dx \quad (3.34)$$

$$d(u \pm v) = du \pm dv = u' dx \pm v' dx \quad (3.35)$$

$$d(uv) = vdu + udv + du \cdot dv = vu' dx + uv' dx \quad (3.36)$$

$$d\frac{u}{v} = \frac{vdu - udv + du \cdot dv}{v^2} = \frac{vdu - udv}{v^2} \quad (3.37)$$

$$du(v(x)) = u'(v) \cdot dv = u'(v) \cdot v' dx \quad (3.38)$$

$d(uv)$ 和 $d\frac{u}{v}$ 的中间项中，我们保留了 $du \cdot dv = u' dx \cdot v' dx = u' v' d^2x$ ，这是一个二阶无穷小，正常情况下极限计算后为 0，如公式右侧结果所示。考虑到微分减法，如 $d(uv) - vdu - udv = du \cdot dv$ ，在与二阶无穷小比值的情况下，这会得到正确结果。

例 3.8.1. 复合函数 $\sin^3 x$ 微分展开

$$u(x) = x^3, \quad v(x) = \sin x$$

$$\begin{aligned} u'(x) &= 3x^2 \rightarrow u'(v) = 3 \sin^2 x \\ v'(x) &= \cos x \\ d\sin^3 x &= u'(v) \cdot v' dx = 3 \sin^2 x \cos x dx \end{aligned}$$

例 3.8.2. 复合函数 $e^{\cos(2x)}$ 微分展开

$$\begin{aligned} de^{\cos(2x)} &= e^{\cos(2x)} \cdot d\cos(2x) \\ &= e^{\cos(2x)} \cdot (-\sin(2x)) \cdot d2x \\ &= -2 \sin(2x) e^{\cos(2x)} dx \end{aligned}$$

因为常数的导数为 0，所以我们有

$$d[f(x) + C] = df(x) + dC = df(x) = f'(x) dx$$

对于本节的微分公式，如果从公式的右侧向左侧计算

$$f'(x) dx = d[f(x) + C] \quad (3.39)$$

第四章 导数和微分的应用

本章我们利用导数和微分来解决一些数学上的问题，如函数的单调性、渐近线和极值等。

4.1 极限计算

1. $\frac{0}{0}$ 型不定式的极限

由微分的基本定义 $df(x) = f(x + \epsilon) - f(x)$ 得到

$$f(x + \epsilon) = df(x) + f(x)$$

如果 $f(x)$ 在 $x = c$ 处为 0，即 $f(c) = 0$ ，则

$$f(c + \epsilon) = df(c) + f(c) = df(c) \quad \text{if } f(c) = 0$$

同时，基于 $df(x) = f'(x) dx \iff df(c) = f'(c) \cdot \epsilon$ 。

结合以上两点，我们考虑当 $f(c) = g(c) = 0$, $x \rightarrow c$ 时函数 $\frac{f(x)}{g(x)}$ 的极限：

$$\begin{aligned} \lim_{x \rightarrow c} \frac{f(x)}{g(x)} &= \frac{f(c + \epsilon)}{g(c + \epsilon)} \\ &= \frac{df(c) + f(c)}{dg(c) + g(c)} \\ &= \frac{df(c)}{dg(c)} \quad (\text{if } f(c) = g(c) = 0) \\ &= \frac{f'(c) \cdot \epsilon}{g'(c) \cdot \epsilon} \\ &= \frac{f'(c)}{g'(c)} \end{aligned}$$

将上面的内容进行总结，我们得到：

对于比值函数 $\frac{f(x)}{g(x)}$ ，如果 $f(c) = g(c) = 0$ ，那么 $x \rightarrow c$ 时函数的极限

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \frac{f(c + \epsilon)}{g(c + \epsilon)} = \frac{f'(c)}{g'(c)} \quad (\text{if } f(c) = g(c) = 0) \quad (4.1)$$

例 4.1.1. 求 $\lim_{x \rightarrow 0} \frac{\sin x}{x}$ 。

当 $x \rightarrow 0$ 时, 分子 $f(0) = \sin 0 = 0$, 分母 $g(0) = (0) = 0$ 。

$$f'(x) = (\sin x)' = \cos x \quad g'(x) = (x)' = 1$$

$$f'(0) = \cos 0 = 1 \quad g'(0) = 1$$

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = \frac{f'(0)}{g'(0)} = \frac{1}{1} = 1$$

例 4.1.2. 求 $\lim_{x \rightarrow 0} \frac{e^x - 1 - x}{x^2}$ 。

当 $x \rightarrow 0$ 时, 分子 $f(0) = e^0 - 1 - (0) = 0$, 分母 $g(0) = (0)^2 = 0$ 。

$$\lim_{x \rightarrow 0} \frac{e^x - 1 - x}{x^2} = \lim_{x \rightarrow 0} \frac{(e^x - 1 - x)'}{(x^2)'} = \lim_{x \rightarrow 0} \frac{e^x - 1}{2x}$$

当 $x \rightarrow 0$ 时, 分子 $f(0) = e^0 - 1 = 0$, 分母 $g(0) = 2(0) = 0$ 。

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{e^x - 1}{2x} &= \lim_{x \rightarrow 0} \frac{(e^x - 1)'}{(2x)'} \\ &= \lim_{x \rightarrow 0} \frac{e^x}{2} = \frac{e^0}{2} = \frac{1}{2} \end{aligned}$$

2. $\frac{\infty}{\infty}$ 型不定式的极限

当求取极限的表达式难解时, 很多时候进行格洛克无穷大等效重定义能够简化极限问题, 方便进一步的极限计算。

例 4.1.3. 求 $\lim_{x \rightarrow 0^+} x \ln x$ 。

$$\begin{aligned} \lim_{x \rightarrow 0^+} x \ln x &= \epsilon \cdot \ln \epsilon \\ &= e^{-\infty} \cdot \ln e^{-\infty} \quad (\epsilon \leftarrow e^{-\infty}) \\ &= e^{-\infty} \cdot (-\infty) = -\frac{\infty}{e^\infty} = 0 \end{aligned}$$

例 4.1.4. 求 $\lim_{x \rightarrow \infty} \frac{\ln x}{x}$ 。

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{\ln x}{x} &= \frac{\ln \infty}{\infty} \\ &= \frac{\ln e^\infty}{e^\infty} \quad (\infty \leftarrow e^\infty) \\ &= \frac{\infty}{e^\infty} = 0 \end{aligned}$$

4.2 函数单调性

单调函数是指在给定区间上是增函数或减函数的函数。具有单调性的区间 I 称为单调区间。

定义 4.1. 单调函数

设函数 $y = f(x)$ 的单一区间定义域为 D , 区间 $I \subseteq D$ 。对于区间 I 上自变量的任意值 x , 我们有以下分类:

(1) 单调递增, 如果微分 $df(x) > 0$ 。

(2) 单调递减, 如果微分 $df(x) < 0$ 。

$df(x) = f(x + \epsilon) - f(x) = f(x) - f(x - \epsilon)$, 在区间 I 的左边界右微分, 在区间 I 的右边界左微分。

单调函数的主要特点:

(1) 最值: 单调函数在闭区间上的最大值和最小值一定在区间的端点处取到。

(2) 零点: 如果单调函数在其区间端点处函数值的符号相反, 则该函数在区间内有且仅有一个零点。

(3) 可逆性: 严格单调函数存在反函数。

例 4.2.1. 函数 $f(x) = x^3$ 在区间 $I = (-\infty, +\infty)$ 上的单调性。

$$\begin{aligned} df(x) &= (x + \epsilon)^3 - x^3 = 3x^2\epsilon + 3x\epsilon^2 + \epsilon^3 \\ &= 3x^2\epsilon + \epsilon^3 > 0 \end{aligned}$$

因为 $x^2 \geq 0$, 可以看出, 对于任意 $x \in I$, $df(x) > 0$ 恒成立, 所以函数 $f(x) = x^3$ 单调递增。

在一般情况下, 我们可以根据单调性的定义直接进行计算来判断函数的单调性。因为 $df(x) = f'(x) dx$, $dx = \epsilon > 0$ 不会改变正负符号, 因此当函数相对复杂时, 可以考虑利用导函数 $f'(x)$ 的符号来判断:

(1) 若在区间 I 上 $f'(x) > 0$, 则函数 $f(x)$ 在 I 上单调递增。

(2) 若在区间 I 上 $f'(x) < 0$, 则函数 $f(x)$ 在 I 上单调递减。

(3) 若在区间 I 上 $f'(x) \geq 0$, 且仅在有限个点上 $f'(x) = 0$, 则函数 $f(x)$ 在 I 上单调递增。

(4) 若在区间 I 上 $f'(x) \leq 0$, 且仅在有限个点上 $f'(x) = 0$, 则函数 $f(x)$ 在 I 上单调递减。

对于 (3) 和 (4) 的理解要点:

$f'(x) \geq 0$ (非负): 这意味着函数 $f(x)$ 在区间 I 上是单调不减的 (即增函数)。函数值 $f(x)$ 永远不会随着 x 的增大而减小。

$f'(x) = 0$ (有限点): 这意味着函数曲线只有在孤立的几个点上有水平切线 (即瞬时增长率为零)。

例 4.2.2. 判定函数 $f(x) = xe^{-x}$ 的单调性。

(3) 求导数

首先, 使用乘积法则 $(uv)' = u'v + uv'$ 对函数 $f(x)$ 求导:

设 $u = x$ 和 $v = e^{-x}$ 。则 $u' = 1$ 和 $v' = -e^{-x}$ 。

$$\begin{aligned} f'(x) &= (1) \cdot e^{-x} + x \cdot (-e^{-x}) \\ &= e^{-x} - xe^{-x} = e^{-x}(1 - x) \end{aligned}$$

(2) 分析导数的符号

由于 e^{-x} 对于所有的实数 x 恒为正 ($e^{-x} > 0$), 因此导数 $f'(x)$ 的符号完全由因子 $(1 - x)$ 决定。

令 $f'(x) = 0$ 找到临界点:

$$e^{-x}(1 - x) = 0$$

因为 $e^{-x} \neq 0$, 所以我们解 $1 - x = 0$, 得到:

$$x = 1$$

(3) 确定单调区间

当 $x < 1$ 时单调递增: 取 $x = 0$ 检验。 $1 - x = 1 - 0 = 1 > 0$ 。因此, $f'(x) = e^{-x}(1 - x) > 0$ 。函数 $f(x)$ 在区间 $(-\infty, 1)$ 上单调递增。

当 $x > 1$ 时单调递减: 取 $x = 2$ 检验。 $1 - x = 1 - 2 = -1 < 0$ 。因此, $f'(x) = e^{-x}(1 - x) < 0$ 。函数 $f(x)$ 在区间 $(1, +\infty)$ 上单调递减。

函数在 $x = 1$ 处达到局部最大值, $f(1) = 1 \cdot e^{-1} = \frac{1}{e}$ 。

4.3 函数的极值

函数极值 (Extremum) 是局部极大值 (Local Maximum) 和局部极小值 (Local Minimum) 的统称, 反映了函数在其定义域的局部范围内的最大或最小性。

定义 4.2. 函数极值

对于函数 $f(x)$ 区间内某一点 $x = c$:

(1) $f(c)$ 是局部极大值, 如果: $f(c) > f(c - \epsilon)$ 并且 $f(c) > f(c + \epsilon)$

(2) $f(c)$ 是局部极小值, 如果: $f(c) < f(c - \epsilon)$ 并且 $f(c) < f(c + \epsilon)$

我们在 $x = c$ 的最小微区间 $[c - \epsilon, c + \epsilon]$ 进行极值定义和判断, 彻底消除了传统极值定义的模糊化表述。

从图像上看, 局部极大值点就像是函数图形上的山顶, 局部极小值点就像是函数图形上的山谷底部。

例 4.3.1. 判断 $x = 0$ 是否是函数的极值点, $f(x) = x^2$, $g(x) = x^3$ 。

对于 $f(x) = x^2$:

$$f(0) - f(0 - \epsilon) = (0)^2 - (-\epsilon)^2 = -\epsilon^2 < 0$$

$$f(0) - f(0 + \epsilon) = (0)^2 - (\epsilon)^2 = -\epsilon^2 < 0$$

所以 $f(0) < f(0 - \epsilon)$ 并且 $f(0) < f(0 + \epsilon)$, $f(0)$ 是函数 $f(x) = x^2$ 的局部极小值。

对于 $g(x) = x^3$:

$$g(0) - g(0 - \epsilon) = (0)^3 - (-\epsilon)^3 = \epsilon^3 > 0$$

$$g(0) - g(0 + \epsilon) = (0)^3 - (\epsilon)^3 = -\epsilon^3 < 0$$

所以 $x = 0$ 不是函数 $g(x) = x^3$ 的极值点。

定理 4.1. 极值判定

对于函数 $f(x)$ 区间内某一点 $x = c$, 如果 $f'(c) = 0$ 并且 $f'(c - \epsilon) \cdot f'(c + \epsilon) < 0$, 那么 c 是函数 $f(x)$ 的极值点。并且

(1) $f(c)$ 是局部极大值, 如果: $f'(c - \epsilon) > 0$ 或者 $f'(c + \epsilon) < 0$

(2) $f(c)$ 是局部极小值, 如果: $f'(c - \epsilon) < 0$ 或者 $f'(c + \epsilon) > 0$

如果: $f'(c - \epsilon) > 0$, 说明 c 点左侧单调递增, 如果 $f'(c + \epsilon) < 0$, 说明 c 点右侧单调递减, 因此 $f(c)$ 是极大值。极小值的情况正好相反。

例 4.3.2. 求函数 $f(x) = x^3 - 3x + 2$ 的所有极值点及极值。

求导数: $f'(x) = (x^3 - 3x + 2)' = 3x^2 - 3$

求驻点 (令 $f'(x) = 0$):

$$3x^2 - 3 = 0$$

$$3(x^2 - 1) = 0$$

$$x^2 = 1$$

解得驻点: $x_1 = -1$, $x_2 = 1$ 。

判断极值:

$$f'(-1 - \epsilon) = 3(-1 - \epsilon)^2 - 3 = 6\epsilon + 3\epsilon^2 > 0$$

$$f'(-1 + \epsilon) = 3(-1 + \epsilon)^2 - 3 = -6\epsilon + 3\epsilon^2 < 0$$

$$f'(1 - \epsilon) = 3(1 - \epsilon)^2 - 3 = -6\epsilon + 3\epsilon^2 < 0$$

$$f'(1 + \epsilon) = 3(1 + \epsilon)^2 - 3 = 6\epsilon + 3\epsilon^2 > 0$$

$x_1 = -1$ 是极大值点, $f(-1) = (-1)^3 - 3(-1) + 2 = 4$

$x_2 = 1$ 是极小值点, $f(1) = (1)^3 - 3(1) + 2 = 0$

例 4.3.3. 求函数 $f(x) = x \ln x$ 的极值。

由于函数定义域要求 $x > 0$, 因此我们只在开区间 $(0, \infty)$ 上求解。

求导数: 使用乘积法则 $(uv)' = u'v + uv'$, 其中 $u = x$, $v = \ln x$.

$$\begin{aligned} f'(x) &= (x)' \cdot \ln x + x \cdot (\ln x)' \\ &= 1 \cdot \ln x + x \cdot \frac{1}{x} = \ln x + 1 \end{aligned}$$

求驻点 (令 $f'(x) = 0$): $\ln x + 1 = 0 \Rightarrow \ln x = -1$

解得驻点: $x = e^{-1} = \frac{1}{e}$ 。

判断极值:

$$\begin{aligned} f'\left(\frac{1}{e} - \epsilon\right) &= \ln\left(\frac{1}{e} - \epsilon\right) + 1 = \ln\left(\frac{1 - e\epsilon}{e}\right) + 1 \\ &= \ln(1 - e\epsilon) - \ln e + 1 \\ &= -e\epsilon < 0 \end{aligned}$$

$$\begin{aligned} f'\left(\frac{1}{e} + \epsilon\right) &= \ln\left(\frac{1}{e} + \epsilon\right) + 1 = \ln\left(\frac{1 + e\epsilon}{e}\right) + 1 \\ &= \ln(1 + e\epsilon) - \ln e + 1 \\ &= e\epsilon > 0 \end{aligned}$$

$x = e^{-1} = \frac{1}{e}$ 是极小值点, $f\left(\frac{1}{e}\right) = \frac{1}{e} \cdot \ln\left(\frac{1}{e}\right) = -\frac{1}{e}$

极值与最值（全局极值）的区别

极值是一个局部性的概念, 而最值是全局性的概念。

可以有多个局部极值, 最多只有一个最大值和一个最小值。

局部极大值不一定大于局部极小值, 最大值必然是所有函数值的最大。

最值需要评估所有极值和区间边界的值, 并选取最大值或最小值。

求解闭区间最值需要比较函数在所有驻点和闭区间端点处的函数值。

例 4.3.4. 求函数 $f(x) = x - 2 \sin x$ 在闭区间 $[0, \frac{\pi}{2}]$ 上的最大值和最小值。

求一阶导数: $f'(x) = (x - 2 \sin x)' = 1 - 2 \cos x$

求驻点 (令 $f'(x) = 0$): $1 - 2 \cos x = 0 \Rightarrow \cos x = \frac{1}{2}$

在区间 $[0, \frac{\pi}{2}]$ 内, 满足 $\cos x = 1/2$ 的点为:

$$x = \frac{\pi}{3}$$

计算函数值: 比较区间端点和驻点处的函数值。

$$\begin{aligned} f(0) &= 0 - 2 \sin(0) = 0 - 0 = 0 \\ f\left(\frac{\pi}{2}\right) &= \frac{\pi}{2} - 2 \sin\left(\frac{\pi}{2}\right) = \frac{\pi}{2} - 2(1) \\ &\approx 1.57 - 2 = -0.43 \\ f\left(\frac{\pi}{3}\right) &= \frac{\pi}{3} - 2 \sin\left(\frac{\pi}{3}\right) = \frac{\pi}{3} - 2\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{3} - \sqrt{3} \\ &\approx 1.047 - 1.732 = -0.685 \end{aligned}$$

比较结果: 最大值: $f(0) = 0$, 最小值: $f\left(\frac{\pi}{3}\right) \approx -0.685$

4.4 函数的渐近线

渐近线 (Asymptote) 是一条直线, 当函数曲线上的点无限远离原点 (即 x 或 y 坐标趋于无穷大) 时, 该曲线与这条直线之间的距离趋近于零。

简单来说, 渐近线就是函数图形在无穷远处无限靠近却永远不会相交的直线。

渐近线主要分为三类: 垂直渐近线, 水平渐近线和斜渐近线。

垂直渐近线 (Vertical Asymptote)

垂直渐近线是形如 $x = a$ 的竖直线。

定义 4.3. 垂直渐近线

如果当 x 从左侧或右侧趋近于某个有限值 a 时, 函数 $f(x)$ 的值趋向于正无穷 (∞) 或负无穷 ($-\infty$), 即:

$$\lim_{x \rightarrow a^-} f(x) = f(a - \epsilon) = \pm\infty \quad (4.2)$$

$$\lim_{x \rightarrow a^+} f(x) = f(a + \epsilon) = \pm\infty \quad (4.3)$$

那么直线 $x = a$ 就是函数 $f(x)$ 的垂直渐近线。

对于有理函数（即两个多项式的比 $f(x) = \frac{P(x)}{Q(x)}$ ），垂直渐近线通常出现在使分母 $Q(x)$ 为零，但分子 $P(x)$ 不为零的点 $x = a$ 处。

例 4.4.1. 求函数 $f(x) = \frac{1}{x-2}$ 的垂直渐近线。

分母 $x - 2 = 0$ 时， $x = 2$ ，此时分子 $1 \neq 0$ 。

检查极限：

$$\begin{aligned}\lim_{x \rightarrow 2^-} f(x) &= f(2 - \epsilon) = \frac{1}{(2 - \epsilon) - 2} \\ &= \frac{1}{-\epsilon} = -\infty\end{aligned}$$

$$\begin{aligned}\lim_{x \rightarrow 2^+} f(x) &= f(2 + \epsilon) = \frac{1}{(2 + \epsilon) - 2} \\ &= \frac{1}{\epsilon} = \infty\end{aligned}$$

所以，直线 $x = 2$ 是垂直渐近线。

水平渐近线 (Horizontal Asymptote)

水平渐近线是形如 $y = b$ 的水平线。

定义 4.4. 水平渐近线

如果当 x 趋向于正无穷 (∞) 或负无穷 ($-\infty$) 时，函数 $f(x)$ 的值趋近于某个有限值 b ，即：

$$\lim_{x \rightarrow \infty} f(x) = f(\infty) = b \quad (4.4)$$

$$\lim_{x \rightarrow -\infty} f(x) = f(-\infty) = b \quad (4.5)$$

那么直线 $y = b$ 就是函数 $f(x)$ 的水平渐近线。

例 4.4.2. 求函数 $f(x) = \frac{3x^2 - 5}{2x^2 + x}$ 的水平渐近线。

检查极限：

$$\begin{aligned}\lim_{x \rightarrow \infty} f(x) &= f(\infty) = \frac{3\infty^2 - 5}{2\infty^2 + x} \\ &= \frac{3\infty^2}{2\infty^2} = \frac{3}{2}\end{aligned}$$

所以水平渐近线为： $y = \frac{3}{2}$ 。

斜渐近线 (Slant/Oblique Asymptote))

斜渐近线是形如 $y = kx + b$ 的斜直线 (其中 $k \neq 0$)。

定义 4.5. 斜渐近线

如果函数 $f(x)$ 满足：

$$k = \lim_{x \rightarrow \pm\infty} \frac{f(x)}{x} = \frac{f(\pm\infty)}{\pm\infty} \quad (4.6)$$

$$b = \lim_{x \rightarrow \pm\infty} [f(x) - kx] = f(\pm\infty) - k(\pm\infty) \quad (4.7)$$

且 k 和 b 都是有限的实数，那么直线 $y = kx + b$ 就是函数 $f(x)$ 的斜渐近线。

对于有理函数 $f(x) = \frac{P(x)}{Q(x)}$ ，只有当分子的次数恰好比分母的次数高 1 ($n = m+1$) 时，函数才可能存在斜渐近线。

例 4.4.3. 求函数 $f(x) = \frac{x^2+1}{x-1}$ 的斜渐近线。

分子次数 $n = 2$ ，分母次数 $m = 1$ ， $n = m + 1$ ，存在斜渐近线。

$$\begin{aligned} k &= \frac{f(\infty)}{\infty} = \frac{\infty^2 + 1}{(\infty - 1) \cdot \infty} \\ &= \frac{\infty^2 + 1}{\infty^2 - \infty} = \frac{\infty^2}{\infty^2} \\ &= 1 \\ b &= f(\infty) - k(\infty) = \frac{\infty^2 + 1}{\infty - 1} - \infty \\ &= \frac{\infty^2 + 1}{\infty - 1} - \frac{\infty^2 - \infty}{\infty - 1} \\ &= \frac{\infty + 1}{\infty - 1} \\ &= 1 \end{aligned}$$

因此，直线 $y = x + 1$ 是斜渐近线。

4.5 牛顿法解方程

牛顿法是一种高效的数值计算方法，广泛应用于各种需要求解非线性方程的工程和科学计算领域。

牛顿法的核心思想是用切线来逐步逼近函数 $f(x)$ 的根 (即函数图形与 x 轴的交点)：

(1) 初始点：首先，选择一个接近真实解的初始猜测值 x_0 。

- (2) 切线逼近：在函数曲线 $y = f(x)$ 上，找到点 $(x_n, f(x_n))$ 。
 - (3) 局部线性化：用该点的切线来代替曲线在 x_n 附近的走势（这是对函数进行局部线性化的操作）。
 - (4) 求新近似解：找到这条切线与 x 轴的交点 x_{n+1} 。这个 x_{n+1} 通常会比 x_n 更接近方程的真实解。
 - (5) 迭代：重复以上步骤，用 x_{n+1} 作为新的猜测值，不断迭代，直到相邻两次近似值的差小于预设的精度要求，或者函数值 $|f(x_{n+1})|$ 足够接近零。
- 牛顿法背后的几何结构如图所示。

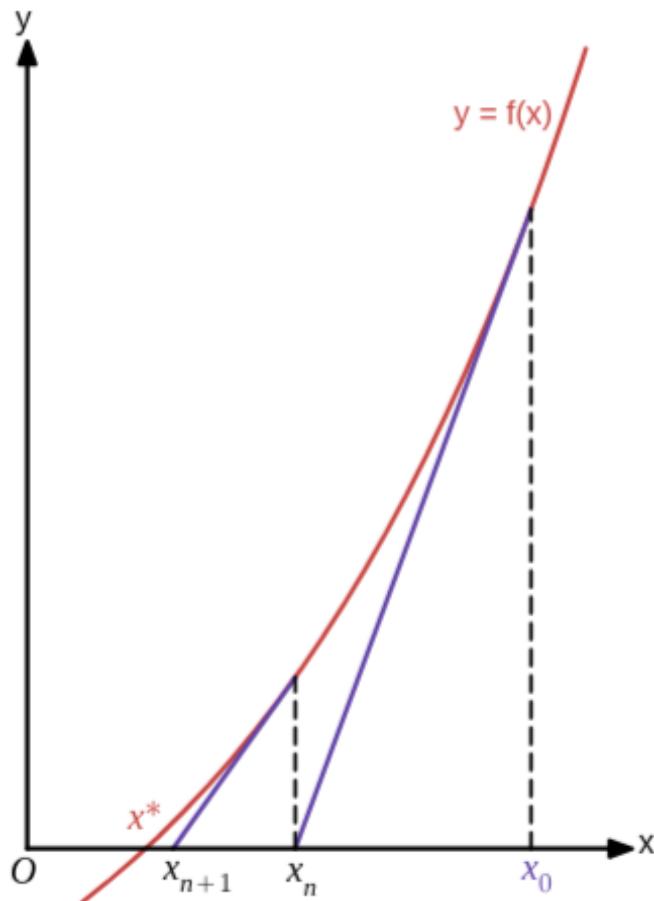


图 4.1: 牛顿法解方程

迭代公式代数推导

从代数的角度，牛顿法是基于函数的泰勒级数展开。

我们假设 x^* 是方程 $f(x) = 0$ 的精确解，即 $f(x^*) = 0$ ，而 x_n 是当前的一个近似解。令 $\Delta x = x^* - x_n$ ，则 $x^* = x_n + \Delta x$ 。

将 $f(x)$ 在 x_n 处进行泰勒级数展开：

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(x_n)}{2!}(x - x_n)^2 + \dots$$

由于我们寻找的是根 x^* , 所以 $f(x^*) = 0$ 。代入 $x = x^*$:

$$0 = f(x_n) + f'(x_n)(x^* - x_n) + \frac{f''(x_n)}{2!}(x^* - x_n)^2 + \dots$$

牛顿法的关键在于截断: 我们只保留泰勒级数的前两项 (即用切线进行线性近似), 忽略高阶项 (二阶及更高阶的项):

$$0 \approx f(x_n) + f'(x_n)(x^* - x_n)$$

将 x^* 用下一个近似值 x_{n+1} 代替, 并解出 x_{n+1} :

$$\begin{aligned} 0 &= f(x_n) + f'(x_n)(x_{n+1} - x_n) \\ f'(x_n)(x_{n+1} - x_n) &= -f(x_n) \\ x_{n+1} - x_n &= -\frac{f(x_n)}{f'(x_n)} \\ x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \end{aligned}$$

最终得到牛顿法的迭代公式:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

其中, $f'(x_n)$ 是函数 $f(x)$ 在 x_n 处的导数 (切线的斜率)。

牛顿法优点与限制

收敛速度快: 在解的附近, 通常具有二阶收敛的特性, 这意味着每次迭代正确的小数位数大致会翻倍, 收敛速度非常快。

需要导数: 要求函数 $f(x)$ 可导, 并且每次迭代都需要计算导数值 $f'(x_n)$ 。

对初值敏感: 如果初始值 x_0 离真实根太远, 或者在迭代过程中 $f'(x_n)$ 接近于零 (即切线接近水平), 方法可能不收敛或收敛到错误的根。

例 4.5.1. 求解 $e^{-x} = x$ 的根

这等价于求解函数 $f(x) = e^{-x} - x = 0$ 的根。

原函数: $f(x) = e^{-x} - x$

导数: $f'(x) = (e^{-x} - x)' = -e^{-x} - 1$

确定迭代公式

将 $f(x)$ 和 $f'(x)$ 代入牛顿法的迭代公式 $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$:

$$x_{n+1} = x_n - \frac{e^{-x_n} - x_n}{-e^{-x_n} - 1} = x_n + \frac{e^{-x_n} - x_n}{e^{-x_n} + 1}$$

开始迭代

首先，通过观察或作图，我们可以找到一个合理的初始猜测。

当 $x = 0$ 时， $f(0) = e^0 - 0 = 1$ ；当 $x = 1$ 时， $f(1) = e^{-1} - 1 \approx 0.368 - 1 = -0.632$ 。由于函数值在 0 和 1 之间变号，所以根在 $[0, 1]$ 区间内。我们取初始值 $x_0 = 0.5$ ：

$$\begin{aligned}x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} = 0.5 - \frac{f(0.5)}{f'(0.5)} \approx 0.566311003197 \\x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} \approx 0.567143165035 \\x_3 &= x_2 - \frac{f(x_2)}{f'(x_2)} \approx 0.56714329041\end{aligned}$$

经过 3 次迭代，我们得到的近似解 $x_3 \approx 0.56714329041$ 已经达到了极高的精度。

这个例子强调了牛顿法的优势：对于具有良好导数性质的函数，它能以二次收敛速度迅速找到方程的根。

第五章 不定积分和积分方法

不定积分的概念起源于 17 世纪，由艾萨克·牛顿和戈特弗里德·莱布尼兹独立发展，作为微积分学的基础。

不定积分也称为反导函数或原函数，它本质上是求导数的逆运算，用于找到一个函数的“原始形式”。

5.1 不定积分的基本概念

考慮微分表达式 $dF(x) = F(x + \epsilon) - F(x)$ ，我们希望用数学的方法由微分 $dF(x)$ 返回原函数 $F(x)$ ，这时我们可以定义一个专用符号如 \int 来表示这种运算操作，即

$$\int dF(x) = F(x)$$

由此我们可以进行一些基本运算：

$$\begin{aligned}\int dx &= x & \int d(\ln x + \sin x) &= \ln x + \sin x \\ F(x) = x^2 + 2x + 3 &\Rightarrow \int dF(x) = x^2 + 2x + 3\end{aligned}$$

接下来考慮微分的导数展开形式：

$$dF(x) = f(x) dx \tag{5.1}$$

如果等式 (5.1) 从右向左计算：

$$f(x) dx = d(F(x) + C) = [F(x + \epsilon) + C] - [F(x) + C] = F(x + \epsilon) - F(x) \tag{5.2}$$

其中 C 为任意常数。也就是说导函数 $f(x)$ 对应的不是唯一原函数 $F(x)$ ，而是一个原函数集合 $F(x) + C$ 。所以我们有

$$\int f(x) dx = \int d(F(x) + C) = F(x) + C \tag{5.3}$$

如果我们将等式 (5.2) 看作众所周知的常识，为了避免计算中间过程的繁琐性，作为一种约定，我们省略微分表达式中的常数项 C ，而只是在最终结果中列出。这样等式 (5.3) 就变成：

$$\int f(x) dx = \int dF(x) = F(x) + C \quad (5.4)$$

这样，我们就完成了由导函数推导计算原函数的过程，并可据此总结出不定积分的定义。

定义 5.1. 不定积分

不定积分是微分的逆运算，作用于微分表达式并解析出原函数。用符号 \int 表示不定积分，我们有

$$\int dF(x) = F(x) \quad (5.5)$$

对于以导函数表示的微分展开形式，不定积分的运算结果是原函数（相对于导函数）的集合，即

$$\int f(x) dx = \int dF(x) = F(x) + C \quad (5.6)$$

在不定积分表达式中微分自变量 dx 也称为积分自变量。

不定积分与微分的互逆关系

微分：给定一个函数 $F(x)$ ，求其导数 $\frac{dF(x)}{dx} = f(x)$ 。

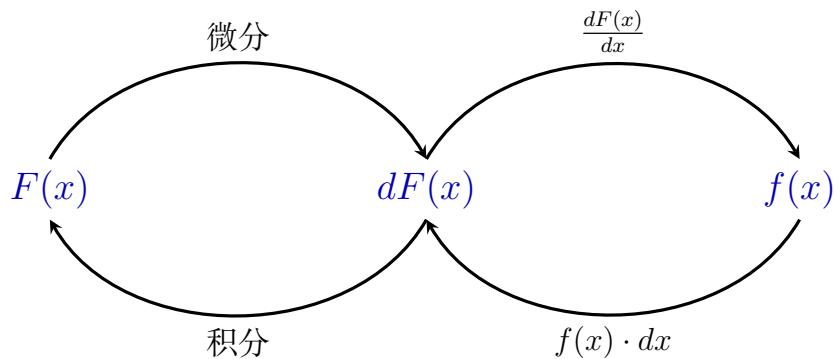
不定积分：给定一个函数 $f(x)$ ，求其原函数 $F(x)$ ，即 $\int f(x) dx = F(x) + C$ 。

逆运算体现：不定积分正是微分（求导数）的逆过程。如果对一个函数先求导，再求不定积分，会回到原来的函数（相差一个常数 C ）。

$$\int \frac{dF(x)}{dx} dx = \int f(x) dx = F(x) + C$$

如果对一个函数先求不定积分，再求导，就会回到原来的函数。

$$\frac{d}{dx} \left(\int f(x) dx \right) = \frac{d}{dx} (F(x) + C) = f(x)$$



5.2 不定积分的主要性质

1. 导数与不定积分的互逆关系

不定积分与微分（求导数）是互逆运算。

不定积分的导数等于被积函数：

$$\left(\int f(x) dx \right)' = f(x)$$

函数的微分（导数）的不定积分等于该函数与任意常数之和：

$$\int F'(x) dx = F(x) + C$$

其中， C 是积分常数。

2. 积分的基本运算法则

(1) 常数倍的积分

常数因子可以提到积分号的外面（其中 k 是不为零的常数）：

$$\int k f(x) dx = k \int f(x) dx \quad (5.7)$$

(2) 和（差）的积分

有限个函数的和（差）的积分等于各个函数的积分的和（差）：

$$\int [f(x) \pm g(x)] dx = \int f(x) dx \pm \int g(x) dx \quad (5.8)$$

(3) 乘积的链式法则

基于微分的链式法则：

$$\int u'(v) \cdot v' dx = \int u'(v) dv = \int du(v) = u(v(x)) + C \quad (5.9)$$

例 5.2.1. 计算积分 $\int (2 \cos x + \sin x) dx$ 。

$$\begin{aligned} \int (2 \cos x + \sin x) dx &= \int 2 \cos x dx + \int \sin x dx \\ &= 2 \cdot \int d \sin x + \int d(-\cos x) \\ &= (2 \sin x + C_1) - (\cos x + C_2) \\ &= 2 \sin x - \cos x + C \end{aligned}$$

在上面的计算中 C_1 和 C_2 被合并成了一个常数项 C ，在实际的计算中， C_1 和 C_2 可以省略。

例 5.2.2. 计算积分 $\int \tan x dx$ 。

$$\begin{aligned}\int \tan x dx &= \int \frac{\sin x}{\cos x} dx \\ &= \int \frac{1}{\cos x} d(-\cos x) = - \int \frac{1}{\cos x} d\cos x\end{aligned}$$

对照链式法则公式：

$$\begin{aligned}u' &= \frac{1}{x} & u &= \ln x & v &= \cos x \\ \Rightarrow u'(v) &= \frac{1}{\cos x} & u(v) &= \ln(\cos x)\end{aligned}$$

所以

$$\int \tan x dx = - \int \frac{1}{\cos x} d\cos x = - \int d\ln |\cos x| = -\ln |\cos x| + C$$

3. 不定积分的几何意义

不定积分 $\int f(x)dx = F(x) + C$ 表示由函数 $f(x)$ 的所有原函数组成的曲线族。

对于曲线族中的任意一条曲线 $y = F(x) + C_1$ 和 $y = F(x) + C_2$, 在横坐标 x_0 处的切线斜率都等于 $f(x_0)$, 因此它们是互相平行的。

这个曲线族中的任意一条曲线, 都可以通过将其中任一条曲线沿 y 轴方向作平移得到。

5.3 基本积分公式

将第三章的导数公式或微分公式进行适当变换, 即可得到常用积分公式。

基本积分公式

- (1) $\int k du = ku + C$ (其中 k 是常数)
- (2) $\int u^n du = \frac{1}{n+1} u^{n+1} + C$ ($n \neq -1$)
- (3) $\int \frac{1}{u} du = \ln |u| + C$
- (4) $\int e^u du = e^u + C$
- (5) $\int a^u du = \frac{a^u}{\ln a} + C$ ($a > 0, a \neq 1$)
- (6) $\int \sin u du = -\cos u + C$

- $$(7) \int \cos u \, du = \sin u + C$$
- $$(8) \int \sec^2 u \, du = \tan u + C$$
- $$(9) \int \csc^2 u \, du = -\cot u + C$$
- $$(10) \int \sec u \tan u \, du = \sec u + C$$
- $$(11) \int \csc u \cot u \, du = -\csc u + C$$
- $$(12) \int \tan u \, du = -\ln |\cos u| + C = \ln |\sec u| + C$$
- $$(13) \int \cot u \, du = \ln |\sin u| + C = -\ln |\csc u| + C$$
- $$(14) \int \sec u \, du = \ln |\sec u + \tan u| + C$$
- $$(15) \int \csc u \, du = -\ln |\csc u + \cot u| + C = \ln |\csc u - \cot u| + C$$
- $$(16) \int \frac{1}{\sqrt{a^2 - u^2}} \, du = \arcsin \left(\frac{u}{a} \right) + C \quad (a > 0)$$
- $$(17) \int \frac{1}{a^2 + u^2} \, du = \frac{1}{a} \arctan \left(\frac{u}{a} \right) + C \quad (a \neq 0)$$
- $$(18) \int \frac{1}{u\sqrt{u^2 - a^2}} \, du = \frac{1}{a} \operatorname{arcsec} \left(\frac{|u|}{a} \right) + C \quad (a > 0)$$

对于公式 (10):

$$\begin{aligned} \int \sec u \tan u \, du &= \int \frac{1}{\cos u} \cdot \frac{\sin u}{\cos u} \, dx \\ &= \int \frac{d(-\cos u)}{(\cos u)^2} = -\frac{1}{-2+1} (\cos u)^{-2+1} \\ &= \frac{1}{\cos u} = \sec u \end{aligned}$$

对于公式 (14):

$$\begin{aligned} \int \sec u \, du &= \int \sec u \cdot \frac{\sec u + \tan u}{\sec u + \tan u} \, du \\ &= \int \frac{\sec u(\sec u + \tan u)}{\sec u + \tan u} \, du \\ &= \int \frac{\sec^2 u + \sec u \tan u}{\sec u + \tan u} \, du \end{aligned}$$

令 $w = \sec u + \tan u$, 回忆导数公式:

$$\frac{d \sec u}{du} = \sec u \tan u \quad \frac{d \tan u}{du} = \sec^2 u$$

$$\begin{aligned}\frac{dw}{du} &= \frac{d}{du}(\sec u + \tan u) = \sec u \tan u + \sec^2 u \\ dw &= (\sec^2 u + \sec u \tan u) du\end{aligned}$$

将换元结果 w 和 dw 代入原积分表达式：

$$\begin{aligned}\int \frac{\sec^2 u + \sec u \tan u}{\sec u + \tan u} du &= \int \frac{1}{w} dw \\ &= \ln |w| + C \\ &= \ln |\sec u + \tan u| + C\end{aligned}$$

因此，我们推导出：

$$\int \sec u du = \ln |\sec u + \tan u| + C$$

对于公式 (16), (17), (18)，我们可以利用微分公式从右侧向左侧推导出结果。积分公式的直接推导我们在稍后的积分方法中给出。

5.4 积分技术

本节主要聚焦于不定积分的求解方法。

1. 换元法

换元法用于简化积分变量，通过引入新变量 $u = g(x)$ ，使积分形式更简单。步骤：

- (1) 选择 $u = g(x)$ ，计算 $du = g'(x) dx$ 。
- (2) 用 u 和 du 替换原积分。
- (3) 求新积分，然后反代回 x 。
- (4) 添加常数 C 。

适用场景：复合函数，如 $\int f(g(x))g'(x) dx$ 。

注意：替换后积分应完全用 u 表示。

例 5.4.1. 计算不定积分。

$$(1) \int x\sqrt{x^2 + 1} dx$$

令 $u = x^2 + 1$ ，则 $du = d(x^2 + 1) = 2x dx$ ，所以 $x dx = \frac{1}{2} du$

$$\begin{aligned}\int x\sqrt{x^2 + 1} dx &= \int \sqrt{x^2 + 1} \cdot x dx = \frac{1}{2} \int \sqrt{u} du \\ &= \frac{1}{2} \cdot \frac{1}{\frac{1}{2} + 1} u^{\frac{1}{2} + 1} + C\end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \cdot \frac{2}{3} u^{\frac{3}{2}} + C \\
 &= \frac{1}{3} (x^2 + 1)^{\frac{3}{2}} + C
 \end{aligned}$$

$$(2) \int e^{3x-5} dx$$

令 $u = 3x - 5$, 则 $du = d(3x - 5) = 3dx$, 所以 $dx = \frac{1}{3}du$

$$\begin{aligned}
 \int e^{3x-5} dx &= \int e^u \cdot \frac{1}{3} du = \frac{1}{3} \int e^u du \\
 &= \frac{1}{3} e^u + C \\
 &= \frac{1}{3} e^{3x-5} + C
 \end{aligned}$$

$$(3) \int \frac{x}{\sqrt{x^2+4}} dx$$

令 $u = x^2 + 4$, 则 $du = d(x^2 + 4) = 2x dx$, 所以 $x dx = \frac{1}{2} du$

$$\begin{aligned}
 \int \frac{x}{\sqrt{x^2+4}} dx &= \int \frac{1}{\sqrt{u}} \left(\frac{1}{2} du \right) \\
 &= \frac{1}{2} \int u^{-1/2} du = \frac{1}{2} \left(\frac{u^{1/2}}{1/2} \right) + C \\
 &= u^{1/2} + C = \sqrt{x^2+4} + C
 \end{aligned}$$

替换法的关键在于选择正确的 u , 使 du 能够简化或消除原积分式中剩余的 x 变量。

2. 分部积分法

分部积分法源于乘积求导法则:

$$(uv)' = u'v + uv'$$

对等式两边求积分:

$$\int (uv)' dx = \int u'v dx + \int uv' dx$$

等式的左边 $\int (uv)' dx = \int d(uv) = uv$, 因此:

$$uv = \int v du + \int u dv$$

稍作移项, 就得到了分部积分公式:

$$\int u dv = uv - \int v du \tag{5.10}$$

分部积分法的核心在于将一个难以求解的积分 $\int u dv$ 转化为一个相对更容易求解的积分 $\int v du$ 。

例 5.4.2. 求 $\int xe^x dx$ 。

首先将积分转化为 $\int u dv$ 的形式: $\int xe^x dx = \int x de^x$ 。

对照公式, $u = x$, $v = e^x$, 所以

$$\begin{aligned}\int xe^x dx &= \int x de^x \\ &= x \cdot e^x - \int e^x dx \\ &= xe^x - e^x + C\end{aligned}$$

例 5.4.3. 求 $\int \ln x dx$ 。

对照公式, $u = \ln x$, $v = x$, 所以

$$\begin{aligned}\int \ln x dx &= \ln x \cdot x - \int x d \ln x \\ &= x \ln x - \int x \cdot \left(\frac{1}{x}\right) dx \\ &= x \ln x - x + C\end{aligned}$$

例 5.4.4. 求解 $\int x \cos x dx$ 。

首先将积分转化为 $\int u dv$ 的形式: $\int x \cos x dx = \int x d \sin x$ 。

对照公式, $u = x$, $v = \sin x$, 所以

$$\begin{aligned}\int x \cos x dx &= \int x d \sin x \\ &= x \cdot \sin x - \int \sin x dx \\ &= x \sin x - (-\cos x) \\ &= x \sin x + \cos x + C\end{aligned}$$

例 5.4.5. 求解 $\int x^2 e^x dx$ 。

$$\begin{aligned}\int x^2 e^x dx &= \int x^2 de^x = x^2 \cdot e^x - \int e^x dx^2 \\ &= x^2 e^x - \int e^x \cdot 2x dx = x^2 e^x - 2 \int x de^x \\ &= x^2 e^x - 2 \left[x \cdot e^x - \int e^x dx \right] \\ &= x^2 e^x - 2xe^x + 2e^x + C \\ &= (x^2 - 2x + 2)e^x + C\end{aligned}$$

分部积分法重要的不是用 u 和 v 进行变量替换, 而是形成 $\int u dv$ 的形式, 然后写出紧凑并且逻辑连贯的解题过程。

3. 部分分式法

部分分式法是一种代数技巧，用于将有理函数（即两个多项式的比值）分解为几个更简单的分式之和。这种分解是积分有理函数时非常重要的一步，因为它能将复杂的积分问题转化为更容易解决的基本积分形式。

(1) 适用条件与预处理

部分分式法仅适用于真有理函数的积分，即分子多项式的次数低于分母多项式的次数。

如果是非真有理函数（分子次数 \geq 分母次数），必须先使用多项式长除法将其分解为一个多项式（可直接积分）和一个真有理函数，然后对剩下的真有理函数部分应用部分分式法。例如

$$\int \frac{x^3 + x}{x^2 - 4} dx$$

需要先进行长除法： $\frac{x^3+x}{x^2-4} = x + \frac{5x}{x^2-4}$ 。接下来只需要对 $\frac{5x}{x^2-4}$ 应用部分分式法。

(2) 部分分式分解的步骤

部分分式法的核心在于分解分母。一旦分母被分解成因子，就可以根据这些因子的类型来构建部分分式。

步骤一：分解分母

将分母多项式分解为实系数的一次因子（形如 $ax + b$ ）和不可再分解的二次因子（形如 $ax^2 + bx + c$ ，其中 $b^2 - 4ac < 0$ ）的乘积。

步骤二：构建部分分式根据分母分解的因子类型，构造待定的部分分式形式。主要有以下四种情况：

分母因子类型	对应的部分分式形式
I. 不同的线性因子 $(ax + b)$	$\frac{A}{ax+b}$
II. 重复的线性因子 $(ax + b)^n$	$\frac{A_1}{ax+b} + \frac{A_2}{(ax+b)^2} + \dots$
III. 不同的不可分解二次因子 $(ax^2 + bx + c)$	$\frac{Ax+B}{ax^2+bx+c}$
IV. 重复的不可分解二次因子 $(ax^2 + bx + c)^n$	$\frac{A_1x+B_1}{ax^2+bx+c} + \frac{A_2x+B_2}{(ax^2+bx+c)^2} + \dots$

步骤三：确定待定系数

将构建的部分分式相加，然后与原始的分子多项式设为相等。有以下两种常用方法确定待定系数 (A, B, C, \dots) ：

特值法对于线性因子，选择能使分母中某个因子为零的 x 值代入方程，从而快速求解一些系数。

系数比较法：将方程两边多项式按 x 的幂次展开并合并同类项，然后比较等式两边 x^n 项的系数，得到一个线性方程组，求解该方程组。

步骤四：进行积分

将确定系数后的部分分式代回积分式中，然后分别对每一项进行积分。

线性因子项 $\int \frac{A}{ax+b} dx$ 通常使用对数 $\frac{A}{a} \ln |ax+b| + C$ 。

不可分解二次因子项 $\int \frac{Ax+B}{ax^2+bx+c} dx$ 通常需要配方，拆分成 \ln 形式和 \arctan 形式的组合。

例 5.4.6. 求积分 $\int \frac{1}{x^2-5x+6} dx$ 。

步骤一：分解分母

分母 $x^2 - 5x + 6$ 可分解为 $(x-2)(x-3)$ 。

步骤二：构建部分分式（类型 I）

$$\frac{1}{x^2-5x+6} = \frac{1}{(x-2)(x-3)} = \frac{A}{x-2} + \frac{B}{x-3}$$

步骤三：确定待定系数

通分并比较分子：

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

所以，

$$\frac{1}{x^2-5x+6} = \frac{-1}{x-2} + \frac{1}{x-3}$$

步骤四：进行积分

$$\begin{aligned} \int \frac{1}{x^2-5x+6} dx &= \int \left(\frac{1}{x-3} - \frac{1}{x-2} \right) dx \\ &= \int \frac{1}{x-3} dx - \int \frac{1}{x-2} dx \\ &= \ln|x-3| - \ln|x-2| + C \\ &= \ln \left| \frac{x-3}{x-2} \right| + C \end{aligned}$$

例 5.4.7. 求积分：

$$\int \frac{x^3+4x^2-x+1}{x(x^2+1)^2} dx$$

步骤一：分解分母与构建部分分式

分母已经分解： $x(x^2+1)^2$ 。

x : 不同的线性因子。

$(x^2+1)^2$: 重复的不可分解二次因子。根据分解规则，构造部分分式：

$$\frac{x^3+4x^2-x+1}{x(x^2+1)^2} = \frac{A}{x} + \frac{Bx+C}{x^2+1} + \frac{Dx+E}{(x^2+1)^2}$$

步骤二：确定待定系数

通分并比较分子：

$$x^3 + 4x^2 - x + 1 = A(x^2 + 1)^2 + (Bx + C)x(x^2 + 1) + (Dx + E)x$$

展开并分组：

$$\begin{aligned} & x^3 + 4x^2 - x + 1 \\ &= A(x^4 + 2x^2 + 1) + (Bx^2 + Cx)(x^2 + 1) + Dx^2 + Ex \\ &= Ax^4 + 2Ax^2 + A + Bx^4 + Bx^2 + Cx^3 + Cx + Dx^2 + Ex \\ &= (A + B)x^4 + (C)x^3 + (2A + B + D)x^2 + (C + E)x + (A) \end{aligned}$$

比较等式两边 x^n 项的系数：

x 的幂次	左边系数	右边系数	得到的方程
x^4	0	$A + B$	(1) $A + B = 0$
x^3	1	C	(2) $C = 1$
x^2	4	$2A + B + D$	(3) $2A + B + D = 4$
x^1	-1	$C + E$	(4) $C + E = -1$
x^0 (常数项)	1	A	(5) $A = 1$

求解系数：

由 (5) 得 $A = 1$ 。

由 (2) 得 $C = 1$ 。

将 $A = 1$ 代入 (1) 得 $1 + B = 0 \Rightarrow B = -1$ 。

将 $C = 1$ 代入 (4) 得 $1 + E = -1 \Rightarrow E = -2$ 。

将 $A = 1$ 和 $B = -1$ 代入 (3) 得 $2(1) + (-1) + D = 4 \Rightarrow 1 + D = 4 \Rightarrow D = 3$ 。

分解结果：

$$\frac{x^3 + 4x^2 - x + 1}{x(x^2 + 1)^2} = \frac{1}{x} + \frac{-x + 1}{x^2 + 1} + \frac{3x - 2}{(x^2 + 1)^2}$$

步骤三：进行积分

$$\int \frac{1}{x} dx + \int \frac{1-x}{x^2+1} dx + \int \frac{3x-2}{(x^2+1)^2} dx$$

第一项 (线性因子)

$$\int \frac{1}{x} dx = \ln|x|$$

第二项 (不同的二次因子) 将分子拆开：

$$\int \frac{1-x}{x^2+1} dx = \int \frac{1}{x^2+1} dx - \int \frac{x}{x^2+1} dx$$

$$\begin{aligned}
&= \arctan x - \frac{1}{2} \int \frac{1}{x^2 + 1} d(x^2 + 1) \\
&= \arctan x - \frac{1}{2} \ln(x^2 + 1)
\end{aligned}$$

第三项 (重复的二次因子) 将分子拆开:

$$\int \frac{3x - 2}{(x^2 + 1)^2} dx = \int \frac{3x}{(x^2 + 1)^2} dx - \int \frac{2}{(x^2 + 1)^2} dx$$

$\int \frac{3x}{(x^2 + 1)^2} dx$: 使用代换 $u = x^2 + 1$, $du = 2x dx$

$$\frac{3}{2} \int \frac{1}{u^2} du = \frac{3}{2} \left(-\frac{1}{u} \right) = \frac{3}{2(x^2 + 1)}$$

$\int \frac{2}{(x^2 + 1)^2} dx$: 需要使用三角代换 $x = \tan \theta$, $dx = \sec^2 \theta d\theta$

$$\begin{aligned}
\int \frac{2}{(\tan^2 \theta + 1)^2} \sec^2 \theta d\theta &= \int \frac{2}{(\sec^2 \theta)^2} \sec^2 \theta d\theta \\
&= 2 \int \frac{1}{\sec^2 \theta} d\theta = 2 \int \cos^2 \theta d\theta \\
&= 2 \int \frac{1 + \cos(2\theta)}{2} d\theta = \int (1 + \cos(2\theta)) d\theta \\
&= \theta + \frac{1}{2} \sin(2\theta) = \theta + \sin \theta \cos \theta
\end{aligned}$$

因为 $x = \tan \theta$, 所以 $\theta = \arctan x$ 。

由三角函数关系可知 $\sin \theta = \frac{x}{\sqrt{x^2 + 1}}$ 且 $\cos \theta = \frac{1}{\sqrt{x^2 + 1}}$ 。所以

$$\int \frac{2}{(x^2 + 1)^2} dx = \arctan x + \frac{x}{x^2 + 1}$$

最终结果

将所有部分相加:

$$\begin{aligned}
&\int \frac{x^3 + 4x^2 - x + 1}{x(x^2 + 1)^2} dx \\
&= \ln|x| + \arctan x - \frac{1}{2} \ln(x^2 + 1) - \frac{3}{2(x^2 + 1)} - \left(\arctan x + \frac{x}{x^2 + 1} \right) \\
&= \ln|x| - \frac{1}{2} \ln(x^2 + 1) - \frac{3}{2(x^2 + 1)} - \frac{x}{x^2 + 1} + C \\
&= \ln \left| \frac{x}{\sqrt{x^2 + 1}} \right| - \frac{2x + 3}{2(x^2 + 1)} + C
\end{aligned}$$

可以看到, 包含重复因子和二次因子的分解步骤与基础情况相同, 但确定系数和积分求解过程会复杂很多。

部分分式法是微积分中处理有理函数积分的基础工具，关键在于正确分解分母并准确构造部分分式。

4. 三角函数积分法

三角函数的积分通常需要结合三角恒等式和换元法来求解。

(1) 降幂与倍角恒等式积分（针对偶次幂）

当积分中包含 $\sin x$ 或 $\cos x$ 的偶次幂时，核心策略是使用降幂公式将其转化为 $\cos(2nx)$ 的线性组合。

核心公式：

$$\sin^2 x = \frac{1 - \cos 2x}{2} \quad \cos^2 x = \frac{1 + \cos 2x}{2}$$

例 5.4.8. 求积分 $\int \cos^2 x dx$ 。

$$\begin{aligned} \int \cos^2 x dx &= \int \frac{1 + \cos 2x}{2} dx && \text{(使用降幂公式)} \\ &= \frac{1}{2} \int (1 + \cos 2x) dx \\ &= \frac{1}{2} \left(\int 1 dx + \int \cos 2x dx \right) \\ &= \frac{1}{2} \left(x + \frac{1}{2} \sin 2x \right) + C && \text{(使用 } u = 2x \text{ 换元)} \\ &= \frac{1}{2}x + \frac{1}{4} \sin 2x + C \end{aligned}$$

例 5.4.9. 求积分 $\int \sin^4 x dx$ 。

第一次降幂：将 $\sin^4 x$ 视为 $(\sin^2 x)^2$ ，使用降幂公式 $\sin^2 x = \frac{1-\cos 2x}{2}$ 。

$$\begin{aligned} \int \sin^4 x dx &= \int (\sin^2 x)^2 dx \\ &= \int \left(\frac{1 - \cos 2x}{2} \right)^2 dx \\ &= \frac{1}{4} \int (1 - 2\cos 2x + \cos^2 2x) dx \end{aligned}$$

第二次降幂：对 $\cos^2 2x$ 使用降幂公式 $\cos^2 \theta = \frac{1+\cos 2\theta}{2}$ ，其中 $\theta = 2x$ 。

$$\cos^2 2x = \frac{1 + \cos(2 \cdot 2x)}{2} = \frac{1 + \cos 4x}{2}$$

代入并积分：

$$\begin{aligned} \int \sin^4 x dx &= \frac{1}{4} \int \left(1 - 2\cos 2x + \frac{1 + \cos 4x}{2} \right) dx \\ &= \frac{1}{4} \int \left(\frac{3}{2} - 2\cos 2x + \frac{1}{2} \cos 4x \right) dx \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{4} \left(\frac{3}{2}x - 2 \cdot \frac{1}{2} \sin 2x + \frac{1}{2} \cdot \frac{1}{4} \sin 4x \right) + C \\
 &= \frac{3}{8}x - \frac{1}{4} \sin 2x + \frac{1}{32} \sin 4x + C
 \end{aligned}$$

(2) 奇次幂的积分 (u 换元法)

当积分形如 $\int \sin^m x \cos^n x dx$, 且 m 或 n 至少有一个是正奇数时, 我们使用 u 换元法。

留下一个奇数次幂的因子 (如 $\sin x$ 或 $\cos x$) 作为 du 的一部分。

将剩余的偶次幂因子利用 $\sin^2 x + \cos^2 x = 1$ 的关系, 转化成另一个三角函数。

例 5.4.10. 求积分 $\int \sin^3 x \cos^2 x dx$ 。

$m = 3$ 是奇数, 我们留下一个 $\sin x$:

变形: $\sin^3 x \cos^2 x = \sin^2 x \cos^2 x \cdot \sin x = (1 - \cos^2 x) \cos^2 x \cdot \sin x$

换元: 令 $u = \cos x$, 则 $du = -\sin x dx$, 即 $\sin x dx = -du$ 。

代入:

$$\begin{aligned}
 \int \sin^3 x \cos^2 x dx &= \int (1 - \cos^2 x) \cos^2 x \cdot \sin x dx \\
 &= \int (1 - u^2) u^2 (-du) \\
 &= \int (u^4 - u^2) du \\
 &= \frac{u^5}{5} - \frac{u^3}{3} + C \\
 &= \frac{\cos^5 x}{5} - \frac{\cos^3 x}{3} + C
 \end{aligned}$$

例 5.4.11. 求积分 $\int \sin^2 x \cos^5 x dx$ 。

$n = 5$ 是奇数, 我们留下一个 $\cos x$ 作为 du 的部分。

变形: 提取一个 $\cos x$, 将剩余的 $\cos^4 x$ 转化为 $\sin x$ 的形式。

$$\begin{aligned}
 \sin^2 x \cos^5 x &= \sin^2 x \cos^4 x \cdot \cos x \\
 &= \sin^2 x (\cos^2 x)^2 \cdot \cos x \\
 &= \sin^2 x (1 - \sin^2 x)^2 \cdot \cos x
 \end{aligned}$$

换元: 令 $u = \sin x$, 则 $du = \cos x dx$ 。

代入并积分:

$$\int \sin^2 x \cos^5 x dx = \int u^2 (1 - u^2)^2 du$$

$$\begin{aligned}
&= \int u^2(1 - 2u^2 + u^4) du \\
&= \int (u^2 - 2u^4 + u^6) du \\
&= \frac{1}{3}u^3 - \frac{2}{5}u^5 + \frac{1}{7}u^7 + C
\end{aligned}$$

替换回 x :

$$\int \sin^2 x \cos^5 x dx = \frac{1}{3} \sin^3 x - \frac{2}{5} \sin^5 x + \frac{1}{7} \sin^7 x + C$$

(3) 万能换元法 (针对三角有理函数)

对于任何形如 $\int R(\sin x, \cos x) dx$ 的有理函数积分 (R 是关于 $\sin x$ 和 $\cos x$ 的多项式的比)，万能换元法保证能将其转化为关于新变量 t 的有理函数积分，理论上总能求解。

核心换元：令 $t = \tan \frac{x}{2}$ 。

核心替换关系：

$$\sin x = \frac{2t}{1+t^2} \quad \cos x = \frac{1-t^2}{1+t^2} \quad dx = \frac{2}{1+t^2} dt$$

例 5.4.12. 求积分 $\int \frac{1}{3+5 \cos x} dx$ 。

代入万能换元公式：

$$\begin{aligned}
\int \frac{1}{3+5 \cos x} dx &= \int \frac{1}{3+5 \left(\frac{1-t^2}{1+t^2} \right)} \cdot \frac{2}{1+t^2} dt \\
&= \int \frac{1}{\frac{3(1+t^2)+5(1-t^2)}{1+t^2}} \cdot \frac{2}{1+t^2} dt \\
&= \int \frac{1+t^2}{(3+3t^2)+(5-5t^2)} \cdot \frac{2}{1+t^2} dt \\
&= \int \frac{2}{8-2t^2} dt \\
&= \int \frac{1}{4-t^2} dt
\end{aligned}$$

部分分式分解：

$$\frac{1}{4-t^2} = \frac{1}{(2-t)(2+t)} = \frac{1}{4} \left(\frac{1}{2+t} + \frac{1}{2-t} \right)$$

积分：

$$\int \frac{1}{4-t^2} dt = \frac{1}{4} \int \left(\frac{1}{2+t} + \frac{1}{2-t} \right) dt$$

$$\begin{aligned}
 &= \frac{1}{4} (\ln |2+t| - \ln |2-t|) + C \\
 &= \frac{1}{4} \ln \left| \frac{2+t}{2-t} \right| + C
 \end{aligned}$$

替换回 x :

$$\int \frac{1}{3+5\cos x} dx = \frac{1}{4} \ln \left| \frac{2+\tan \frac{x}{2}}{2-\tan \frac{x}{2}} \right| + C$$

(4) 其它 u 换元应用

对于涉及 $\sec x, \tan x, \csc x, \cot x$ 的积分，常常通过恒等式将其转化为 $\sin x$ 和 $\cos x$ 的形式，或直接使用特定的 u 换元。

例 5.4.13. 求积分 $\int \tan^3 x dx$ 。

变形: $\tan^3 x = \tan x \cdot \tan^2 x = \tan x (\sec^2 x - 1)$

$$\int \tan^3 x dx = \int \tan x \sec^2 x dx - \int \tan x dx$$

第一个积分 (u 换元): $\int \tan x \sec^2 x dx$

令 $u = \tan x$, 则 $du = \sec^2 x dx$

$$\int u du = \frac{u^2}{2} = \frac{1}{2} \tan^2 x$$

第二个积分 (基础公式): $\int \tan x dx = -\ln |\cos x|$

合并:

$$\int \tan^3 x dx = \frac{1}{2} \tan^2 x + \ln |\cos x| + C$$

相对来说微分很简单，但积分却不是。在求函数的导数时，通常很清楚必须应用哪个公式，但应该使用哪种方法来积分给定函数可能并不明显。由于本节的积分问题都集中在对应的方法上，因此通常很清楚对给定的积分时使用什么方法。

无论如何，熟练掌握基础公式的推导过程和基本应用是提高积分技术的基础。

第六章 常微分方程

常微分方程 (Ordinary Differential Equations, ODE) 是数学中描述变化率的重要工具，广泛应用于物理、工程、生物等领域。

常微分方程涉及未知函数及其导数的方程，其中未知函数只有一个自变量（通常是时间或空间变量）。例如， $\frac{dy}{dx} = f(x, y)$ 就是一个一阶常微分方程。

阶数：方程中最高导数的阶数。一阶涉及一阶导数，二阶涉及二阶导数，以此类推。

线性 vs. 非线性：线性 ODE 的未知函数及其导数以线性形式出现（无乘积或幂次），如 $y'' + p(x)y' + q(x)y = r(x)$ 。

齐次 vs. 非齐次：齐次方程的右边为 0，非齐次有非零项。

通解 vs. 特解：通解包含任意常数，特解通过初始条件确定。

6.1 一阶常微分方程

一阶 ODE 的形式为 $\frac{dy}{dx} = f(x, y)$ 。常见类型包括可分离、线性、齐次等。

可分离方程

如果方程可写成 $\frac{dy}{dx} = g(x)h(y)$ ，则可分离。

求解步骤：

(1) 分离变量

假设 $h(y) \neq 0$ ，将 $h(y)$ 除到等号左边，并将 dx 乘到等号右边，得到微分形式：

$$\frac{1}{h(y)}dy = g(x)dx$$

为了简化表示，我们通常用 $p(y) = \frac{1}{h(y)}$ 来表示左边只含 y 的部分。方程变为：

$$p(y)dy = g(x)dx$$

(2) 积分

对等式的两边分别进行不定积分：

$$\int p(y)dy = \int g(x)dx + C$$

其中 C 是任意常数。

(3) 得到通解

执行积分后，通常可以得到一个包含 x, y 和常数 C 的隐式解，即方程的通解。如果可能，也可以进一步解出 y 关于 x 的显式解。

例 6.1.1. 求解 $\frac{dy}{dx} = \frac{y}{x}$ (假设 $x \neq 0, y \neq 0$)。

分离变量得 $\frac{dy}{y} = \frac{dx}{x}$ 。

两边积分：

$$\begin{aligned} \int \frac{dy}{y} &= \int \frac{dx}{x} \\ \ln |y| &= \ln |x| + C \\ e^{\ln |y|} &= e^{\ln |x|+C} \\ |y| &= e^C \cdot |x| \\ y &= Ax \quad (A = \pm e^C, \text{ 其中 } A \text{ 为任意非零常数}) \end{aligned}$$

方程通解： $y(x) = Ax$ 。

这是一个比例增长模型，如人口增长率与人口成正比。

例 6.1.2. 求解微分方程： $\frac{dy}{dx} = x^2y$ 。

分离变量：假设 $y \neq 0$ ，将 y 除到左边， dx 乘到右边：

$$\frac{1}{y}dy = x^2dx$$

积分：对两边分别积分：

$$\begin{aligned} \int \frac{1}{y}dy &= \int x^2dx + C \\ \ln |y| &= \frac{1}{3}x^3 + C \end{aligned}$$

得到通解（显式解）：将通解转化为显式形式（解出 y ）：

$$|y| = e^{\frac{1}{3}x^3+C} = e^C e^{\frac{1}{3}x^3}$$

令 $A = \pm e^C$ (A 为任意非零常数)，则：

$$y = Ae^{\frac{1}{3}x^3}$$

如果考虑到 $y = 0$ 也是原方程的一个解（即 $y \equiv 0$ ），并且 $y \equiv 0$ 可以包含在上述通解中（令 $A = 0$ ），所以最终的通解为：

$$y = Ae^{\frac{1}{3}x^3} \quad (\text{其中 } A \text{ 是任意常数})$$

线性方程

一阶线性微分方程的标准形式可以表示为：

$$\frac{dy}{dx} + P(x)y = Q(x)$$

其中：

y 是未知函数（因变量），它是 x 的函数。

$\frac{dy}{dx}$ 是 y 对 x 的一阶导数。

$P(x)$ 和 $Q(x)$ 是只与自变量 x 有关的已知函数（或常数）。

求解步骤：

求解一阶线性微分方程的标准方法是利用积分因子法。

(1) 确定积分因子 $I(x)$

积分因子 $I(x)$ 定义为：

$$I(x) = e^{\int P(x) dx}$$

$$I'(x) = e^{\int P(x) dx} \cdot \left(\int P(x) dx \right)' = I(x)P(x)$$

注意：在计算 $\int P(x) dx$ 时，可以不加任意常数 C ，因为 C 最终会被消去。

(2) 求解过程

将标准形式的微分方程 $\frac{dy}{dx} + P(x)y = Q(x)$ 两边同乘以积分因子 $I(x)$ ：

$$I(x)\frac{dy}{dx} + I(x)P(x)y = I(x)Q(x)$$

关键在于，等式的左边正好是函数乘积 $I(x)y$ 的导数：

$$\frac{d}{dx}[I(x)y] = I(x)\frac{dy}{dx} + I(x)P(x)y$$

所以原方程变为：

$$\frac{d}{dx}[I(x)y] = I(x)Q(x)$$

(3) 积分并得到通解

对上式两边关于 x 进行积分：

$$I(x)y = \int I(x)Q(x) dx + C$$

其中 C 是积分常数。

最后, 解出 y , 得到方程的通解:

$$y = \frac{1}{I(x)} \left[\int I(x)Q(x) dx + C \right]$$

例 6.1.3. 求解微分方程:

$$\frac{dy}{dx} + \frac{2}{x}y = x$$

确定 $P(x)$ 和 $Q(x)$: $P(x) = \frac{2}{x}$, $Q(x) = x$

计算积分因子 $I(x)$:

$$\begin{aligned} \int P(x) dx &= \int \frac{2}{x} dx = 2 \ln|x| = \ln(x^2) \\ I(x) &= e^{\int P(x) dx} = e^{\ln(x^2)} = x^2 \end{aligned}$$

应用公式或直接求解: 将原方程乘以 $I(x) = x^2$:

$$x^2 \frac{dy}{dx} + 2xy = x^3$$

左边可以写成 $\frac{d}{dx}(x^2y)$:

$$\frac{d}{dx}(x^2y) = x^3$$

积分:

$$\begin{aligned} \int \frac{d}{dx}(x^2y) dx &= \int x^3 dx \\ x^2y &= \frac{1}{4}x^4 + C \end{aligned}$$

得到通解:

$$\begin{aligned} y &= \frac{1}{x^2} \left(\frac{1}{4}x^4 + C \right) \\ y &= \frac{1}{4}x^2 + \frac{C}{x^2} \end{aligned}$$

一阶线性方程的求解公式是:

$$y = e^{-\int P(x) dx} \left[\int Q(x)e^{\int P(x) dx} dx + C \right]$$

这个公式是通用且可靠的。

齐次方程

一个一阶常微分方程可以写成：

$$\frac{dy}{dx} = f(x, y)$$

如果函数 $f(x, y)$ 是一个零次齐次函数，它总是可以写成只依赖于 $\frac{y}{x}$ 的某个函数 $g\left(\frac{y}{x}\right)$ 。

因此，一阶齐次微分方程的标准形式可以写成：

$$\frac{dy}{dx} = g\left(\frac{y}{x}\right)$$

一阶齐次微分方程通过代换 $u = \frac{y}{x}$ 转化为可分离变量方程来求解。

求解步骤：

齐次微分方程的求解方法是利用一个巧妙的变量代换，将其转化为可分离变量方程来求解。

(1) 变量代换

令一个新的变量 u 为 $\frac{y}{x}$ ：

$$u = \frac{y}{x} \implies y = ux$$

(2) 转换微分

对 $y = ux$ 两边关于 x 求导（使用乘积求导法则）：

$$\frac{dy}{dx} = \frac{d}{dx}(ux) = \frac{du}{dx}x + u\frac{dx}{dx} = x\frac{du}{dx} + u$$

(3) 代入原方程

将 $\frac{dy}{dx}$ 和 $\frac{y}{x}$ 代入原方程 $\frac{dy}{dx} = g\left(\frac{y}{x}\right)$ ：

$$x\frac{du}{dx} + u = g(u)$$

(4) 转化为可分离变量方程

将 u 移到等式右边：

$$x\frac{du}{dx} = g(u) - u$$

假设 $g(u) - u \neq 0$ ，现在方程变成了关于 x 和 u 的可分离变量方程：

$$\frac{1}{g(u) - u} du = \frac{1}{x} dx$$

(5) 积分求解

对两边积分：

$$\int \frac{1}{g(u) - u} du = \int \frac{1}{x} dx + C$$

解出关于 u 和 x 的关系。

(6) 还原变量

用 $u = \frac{y}{x}$ 替换回 u , 得到原方程关于 y 和 x 的通解。

例 6.1.4. 求解微分方程:

$$\frac{dy}{dx} = \frac{y^2 + x^2}{xy}$$

判断齐次性:

$$\frac{dy}{dx} = \frac{y^2}{xy} + \frac{x^2}{xy} = \frac{y}{x} + \frac{x}{y}$$

由于右边可以表示成 $\frac{y}{x}$ 的函数

$$g\left(\frac{y}{x}\right) = \frac{y}{x} + \frac{1}{\frac{y}{x}}$$

所以是齐次方程。

变量代换: 令 $u = \frac{y}{x}$, 则 $\frac{dy}{dx} = x\frac{du}{dx} + u$ 。

代入方程:

$$x\frac{du}{dx} + u = u + \frac{1}{u}$$

分离变量: u 被抵消:

$$\begin{aligned} x\frac{du}{dx} &= \frac{1}{u} \\ udu &= \frac{1}{x}dx \end{aligned}$$

积分求解:

$$\begin{aligned} \int u du &= \int \frac{1}{x} dx + C \\ \frac{1}{2}u^2 &= \ln|x| + C \end{aligned}$$

还原变量: 用 $u = \frac{y}{x}$ 替换:

$$\begin{aligned} \frac{1}{2}\left(\frac{y}{x}\right)^2 &= \ln|x| + C \\ \frac{y^2}{2x^2} &= \ln|x| + C \\ y^2 &= 2x^2(\ln|x| + C) \end{aligned}$$

例 6.1.5. 求解微分方程: $\frac{dy}{dx} = \frac{1}{x+y}$

我们可以使用变量替换法来求解。

变量替换

设 $u = x + y$ 。对 $u = x + y$ 关于 x 求导：

$$\frac{du}{dx} = \frac{d}{dx}(x + y) = \frac{dx}{dx} + \frac{dy}{dx}$$

即

$$\frac{du}{dx} = 1 + \frac{dy}{dx}$$

因此，我们可以将 $\frac{dy}{dx}$ 表示为：

$$\frac{dy}{dx} = \frac{du}{dx} - 1$$

代入原方程

将 $u = x + y$ 和 $\frac{dy}{dx} = \frac{du}{dx} - 1$ 代入原方程：

$$\frac{du}{dx} - 1 = \frac{1}{u}$$

分离变量

将方程整理为可分离变量的形式：

$$\frac{du}{dx} = 1 + \frac{1}{u} = \frac{u+1}{u}$$

分离变量，得到：

$$\frac{u}{u+1} du = dx$$

积分

对等式两边进行积分：

$$\int \frac{u}{u+1} du = \int dx$$

对于左边的积分，可以将 $\frac{u}{u+1}$ 改写为 $1 - \frac{1}{u+1}$ ：

$$\int \left(1 - \frac{1}{u+1}\right) du = \int dx$$

进行积分：

$$\int 1 du - \int \frac{1}{u+1} du = \int dx$$

$$u - \ln|u+1| = x + C$$

其中 C 是任意常数。

还原变量

最后，用 $x + y$ 替换 u ：

$$(x + y) - \ln|x + y + 1| = x + C$$

两边消去 x :

$$y - \ln|x + y + 1| = C$$

原微分方程 $\frac{dy}{dx} = \frac{1}{x+y}$ 的通解为:

$$y - \ln|x + y + 1| = C$$

这是一个隐式解。

6.2 二阶线性常微分方程

一个二阶线性常微分方程的一般形式可以表示为:

$$y'' + P(x)y' + Q(x)y = R(x)$$

其中:

x 是自变量 (通常是时间 t 或空间坐标)。

y 是因变量, 它是 x 的函数, 即 $y = y(x)$ 。

y' 和 y'' 分别是 y 对 x 的一阶和二阶导数。

“二阶” 指的是方程中出现的最高阶导数是二阶导数 y'' 。

齐次 (Homogeneous): 如果 $R(x) = 0$, 方程为齐次方程。

非齐次 (Non-homogeneous): 如果 $R(x) \neq 0$, 方程为非齐次方程。

齐次线性常系数方程

方程的一般形式为:

$$y'' + Py' + Qy = 0$$

其中 P 和 Q 都是常数。

求解这类方程的关键步骤是引入一个特征方程。我们假设方程有一个形式为 $y = e^{rx}$ 的解 (其中 r 是待定常数), 代入原方程:

$$y' = re^{rx} \quad y'' = r^2e^{rx}$$

代入原方程:

$$r^2e^{rx} + P(re^{rx}) + Q(e^{rx}) = 0$$

因为 $e^{rx} \neq 0$, 两边除以 e^{rx} , 得到特征方程:

$$r^2 + Pr + Q = 0$$

解出这个一元二次代数方程的两个根 r_1 和 r_2 , 它们被称为特征根。根据 $\Delta = P^2 - 4Q$ 的符号, 特征根有三种情况, 从而决定了微分方程的通解形式。

三种通解情况

微分方程的通解 $y(x)$ 是由两个线性无关的基本解 $y_1(x)$ 和 $y_2(x)$ 的线性组合构成: $y(x) = C_1y_1(x) + C_2y_2(x)$ (其中 C_1, C_2 为任意常数)。

(1) 两不相等实根 ($\Delta > 0$)

特征根: r_1 和 r_2 是两个不相等的实数。

基本解: $y_1(x) = e^{r_1 x}$ 和 $y_2(x) = e^{r_2 x}$

通解:

$$y(x) = C_1e^{r_1 x} + C_2e^{r_2 x}$$

物理意义 (例如阻尼振动): 通常对应于过阻尼情况, 系统衰减到平衡位置, 不发生振荡。

(2) 两相等实根 (重根) ($\Delta = 0$)

特征根: $r_1 = r_2 = r$ (二重实根)。

基本解: $y_1(x) = e^{rx}$ 和 $y_2(x) = xe^{rx}$

通解:

$$y(x) = (C_1 + C_2x)e^{rx}$$

物理意义 (例如阻尼振动): 对应于临界阻尼情况, 系统以最快速度衰减到平衡位置, 不发生振荡。

(3) 共轭复根 ($\Delta < 0$)

特征根: $r_1, r_2 = \alpha \pm i\beta$ (其中 $\alpha = -P/2, \beta = \sqrt{4Q - P^2}/2, i^2 = -1$)。

基本解: $y_1(x) = e^{\alpha x} \cos(\beta x)$ 和 $y_2(x) = e^{\alpha x} \sin(\beta x)$

通解:

$$y(x) = e^{\alpha x}(C_1 \cos(\beta x) + C_2 \sin(\beta x))$$

物理意义 (例如阻尼振动): 对应于欠阻尼情况, 系统在衰减过程中会发生振荡。

例 6.2.1. 求解微分方程: $y'' + 4y' - 12y = 0$ 。

首先, 根据微分方程写出特征方程 $r^2 + Pr + Q = 0$ 。

这里 $P = 4$, $Q = -12$, 所以特征方程是:

$$r^2 + 4r - 12 = 0$$

使用因式分解或求根公式来求解 r :

$$(r + 6)(r - 2) = 0$$

得到两个特征根:

$$r_1 = -6 \quad \text{和} \quad r_2 = 2$$

由于我们得到了两不相等实根(即情况 1), 通解的形式为 $y(x) = C_1 e^{r_1 x} + C_2 e^{r_2 x}$ 。最终通解为:

$$y(x) = C_1 e^{-6x} + C_2 e^{2x}$$

例 6.2.2. 求解微分方程: $y'' - 6y' + 9y = 0$ 。

特征方程: $r^2 - 6r + 9 = 0$

特征根: $(r - 3)^2 = 0 \implies r_1 = r_2 = 3$ (重根)

通解 (情况 2):

$$y(x) = (C_1 + C_2 x)e^{3x}$$

例 6.2.3. 求解微分方程: $y'' + 2y' + 5y = 0$ 。

特征方程: $r^2 + 2r + 5 = 0$

特征根 (使用求根公式 $r = \frac{-P \pm \sqrt{P^2 - 4Q}}{2}$):

$$\begin{aligned} r &= \frac{-2 \pm \sqrt{2^2 - 4(1)(5)}}{2} = \frac{-2 \pm \sqrt{4 - 20}}{2} \\ &= \frac{-2 \pm \sqrt{-16}}{2} = \frac{-2 \pm 4i}{2} \\ &= -1 \pm 2i \end{aligned}$$

即 $\alpha = -1$ 和 $\beta = 2$ 。

通解 (情况 3):

$$y(x) = e^{\alpha x}(C_1 \cos(\beta x) + C_2 \sin(\beta x))$$

$$y(x) = e^{-x}(C_1 \cos(2x) + C_2 \sin(2x))$$

如果需要确定通解中的常数 C_1 和 C_2 , 这通常需要初始条件或边界条件才能确定。

非齐次线性常系数方程

非齐次二阶常微分方程的求解是建立在齐次方程求解基础之上的, 因为它引入了外部驱动项 (非齐次项)。

非齐次线性常微分方程的一般形式为:

$$y'' + p(x)y' + q(x)y = r(x)$$

其中 $r(x)$ 是非齐次项 (或驱动项), 且 $r(x) \neq 0$ 。

求解结构: 通解的叠加原理

非齐次方程的通解 $y(x)$ 总是由两部分组成, 即叠加原理:

$$y(x) = y_h(x) + y_p(x)$$

(1) 齐次通解 (y_h): 对应于 $r(x) = 0$ 时 (即齐次方程) 的通解。

$$y'' + p(x)y' + q(x)y = 0$$

这一部分包含了两个任意常数 C_1 和 C_2 , 我们已在之前的讨论中解决了常系数齐次方程的求解。

(2) 特解 (y_p): 满足原非齐次方程的任意一个特定解。

$$y_p'' + p(x)y_p' + q(x)y_p = r(x)$$

这一部分不包含任意常数。

求解非齐次方程的关键和挑战在于找到这个特解 $y_p(x)$ 。

待定系数法求解特解 y_p

这种方法适用于非齐次项 $r(x)$ 具有特定形式 (多项式、指数函数、正弦/余弦函数或它们的乘积) 的情况。

核心思想: 猜测 $y_p(x)$ 的形式应与 $r(x)$ 的形式相似。

$r(x)$ 的形式	$y_p(x)$ 的初始猜测形式
Ae^{ax}	Ke^{ax}
Ax^n	$K_n x^n + K_{n-1} x^{n-1} + \cdots + K_0$
$A \cos(\omega x)$ 或 $B \sin(\omega x)$	$K_1 \cos(\omega x) + K_2 \sin(\omega x)$
乘积: 例如 xe^{ax}	乘积: 例如 $(K_1 x + K_0)e^{ax}$

修正规则 (当猜测形式与 y_h 重复时):

如果 y_p 的猜测形式的任何一项与 y_h 中的某一项相同 (即是齐次方程的解), 则必须将猜测形式乘以 x^s , 其中 s 是使新的 y_p 猜测形式中所有项都不是齐次方程解的最小非负整数 (通常 $s = 1$ 或 $s = 2$)。

例 6.2.4. 求解微分方程: $y'' - 3y' - 4y = 2 \sin x$ 。

第 1 步: 求解齐次通解 (y_h)

齐次方程: $y'' - 3y' - 4y = 0$

特征方程: $r^2 - 3r - 4 = 0$

特征根: $(r - 4)(r + 1) = 0 \implies r_1 = 4, r_2 = -1$

齐次通解:

$$y_h(x) = C_1 e^{4x} + C_2 e^{-x}$$

第 2 步: 求解特解 (y_p)

非齐次项 $r(x) = 2 \sin(x)$ 。根据表格, 猜测特解形式为:

$$y_p = A \cos(x) + B \sin(x)$$

(注意：由于 y_h 中不包含 $\cos(x)$ 或 $\sin(x)$ 项，不需要进行修正。)

求导：

$$y'_p = -A \sin(x) + B \cos(x)$$

$$y''_p = -A \cos(x) - B \sin(x)$$

代入原方程 $y'' - 3y' - 4y = 2 \sin(x)$ ：

$$(-A \cos x - B \sin x) - 3(-A \sin x + B \cos x) - 4(A \cos x + B \sin x) = 2 \sin x$$

$$(-A - 3B - 4A) \cos x + (-B + 3A - 4B) \sin x = 2 \sin x$$

$$(-5A - 3B) \cos x + (3A - 5B) \sin x = 2 \sin x$$

系数对比：

$$\cos(x) \text{ 系数: } -5A - 3B = 0$$

$$\sin(x) \text{ 系数: } 3A - 5B = 2$$

$$\text{解线性方程组: } A = \frac{3}{17}, B = -\frac{5}{17}$$

特解：

$$y_p(x) = \frac{3}{17} \cos(x) - \frac{5}{17} \sin(x)$$

将 y_h 和 y_p 相加写出通解：

$$y(x) = C_1 e^{4x} + C_2 e^{-x} + \frac{3}{17} \cos(x) - \frac{5}{17} \sin(x)$$

当非齐次项 $r(x)$ 的形式与齐次通解 y_h 中的某一项线性相关时，直接代入 $r(x)$ 的形式作为特解猜测会导致矛盾（即代入后左侧为零），因此必须应用修正规则。

例 6.2.5. 求解微分方程： $y'' - 2y' - 3y = 5e^{3x}$ 。

第 1 步：求解齐次通解 (y_h)

齐次方程： $y'' - 2y' - 3y = 0$

特征方程： $r^2 - 2r - 3 = 0$

特征根： $(r - 3)(r + 1) = 0 \implies r_1 = 3, r_2 = -1$

齐次通解：

$$y_h(x) = C_1 e^{3x} + C_2 e^{-x}$$

第 2 步：求解特解 (y_p)

非齐次项 $r(x) = 5e^{3x}$ 。

初始猜测：如果没有重合，我们会猜测 $y_p = Ae^{3x}$ 。

检查重合：但是，我们的 y_h 中包含 $C_1 e^{3x}$ 这一项。这意味着 e^{3x} 是齐次方程的解，代入原方程左侧将得到 0，无法等于 $5e^{3x}$ 。

应用修正规则：由于 e^{3x} 是齐次解，我们需要将初始猜测乘以 x ，即：

$$y_p = Axe^{3x}$$

(因为 xe^{3x} 不是齐次方程的解。)

求导并代入：现在我们必须计算 y'_p 和 y''_p ：

$$\begin{aligned} y'_p &= A(1 \cdot e^{3x} + x \cdot 3e^{3x}) = Ae^{3x}(1 + 3x) \\ y''_p &= A[3e^{3x}(1 + 3x) + e^{3x}(3)] = Ae^{3x}(3 + 9x + 3) \\ &= Ae^{3x}(6 + 9x) \end{aligned}$$

将 y_p, y'_p, y''_p 代入原方程 $y'' - 2y' - 3y = 5e^{3x}$ ：

$$\begin{aligned} Ae^{3x}(6 + 9x) - 2 \cdot Ae^{3x}(1 + 3x) - 3 \cdot Axe^{3x} &= 5e^{3x} \\ Ae^{3x}[(6 + 9x) - 2(1 + 3x) - 3x] &= 5e^{3x} \\ Ae^{3x}[6 + 9x - 2 - 6x - 3x] &= 5e^{3x} \\ 4Ae^{3x} &= 5e^{3x} \\ A &= \frac{5}{4} \end{aligned}$$

特解：

$$y_p(x) = \frac{5}{4}xe^{3x}$$

将 y_h 和 y_p 相加写出通解：

$$y(x) = y_h(x) + y_p(x) = C_1e^{3x} + C_2e^{-x} + \frac{5}{4}xe^{3x}$$

单重根重合： $r(x) = 5e^{3x}$ 与 y_h 中的 C_1e^{3x} 重合，则 $s = 1$ ，特解形式为 Axe^{3x} 。

双重根(重根)重合：如果齐次方程的特征根是重根 $r = 3$ ，即 $y_h = C_1e^{3x} + C_2xe^{3x}$ ，而 $r(x) = 5e^{3x}$ 。此时 e^{3x} 和 xe^{3x} 都是齐次解，则需乘以 x^2 ，特解形式为 Ax^2e^{3x} 。

求解注意事项

(1) 检验方程类型和适用方法

在开始求解前，必须准确判断方程的类型（一阶/高阶、线性/非线性、齐次/非齐次）。

只有在确定方程属于某种经典类型时，才尝试使用相应的解析求解方法。

(2) 初始条件和边界条件的完整性

n 阶常微分方程的通解中包含 n 个待定常数。

求解初值问题或边值问题时，必须利用给定的初始值或边界条件来确定这些常数，从而得到特解。

(3) 解析解的隐式性与定义域

有些微分方程只能得到隐式解，即因变量和自变量的关系 $F(x, y) = C$ ，而不是显式的 $y = f(x)$ 。需要注意在什么区域内 y 可以被视为 x 的函数。

检查解（特别是特解）在整个定义域内是否有效，是否存在奇点或解的突然中断。

第七章 定积分

定积分的起源可以追溯到古希腊的“穷竭法”，它主要用来解决求不规则图形面积的问题。

现代定积分的系统理论是由牛顿 (Isaac Newton) 和莱布尼茨 (Gottfried Wilhelm Leibniz) 在 17 世纪独立发展的微积分子学奠定的基础。19 世纪，黎曼 (Bernhard Riemann) 提出了黎曼积分的定义，使其更严谨。

格洛克提出区间微分的概念，并据此进行区间积分，计算在格洛克微空间进行，极限计算得出微积分的基本定理，在理解上更加简单直接。

7.1 区间微分和区间积分

考虑函数 $F(x)$ 的点微分： $dF(x) = f(x) dx$ ，在几何上，左侧是线性的极限变化量，右侧是一个高度为 $f(x)$ ，宽度为无穷小的微矩形的面积，如图 7.1 所示。

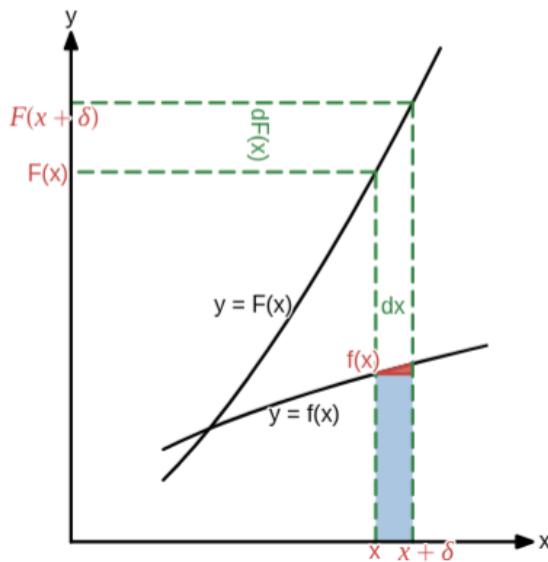


图 7.1：微分的几何表示

需要注意的是，微矩形的面积和对应点的 $F(x)$ 线性极限变化量精确相等。

接下来我们考虑图中函数 $f(x)$ 在区间 $[x, x + \delta]$ 投影到 x 轴（曲边梯形）的面积 S :

$$\begin{aligned} S &= f(x) dx + \frac{1}{2}[f(x + \delta) - f(x)] \cdot dx \\ &= f(x) dx + \frac{1}{2} df(x) \cdot dx \\ &= f(x) dx + \frac{1}{2} f'(x) d^2x \end{aligned}$$

很明显，在格洛克微空间，曲边和红色三角形斜边重合，如图 7.1 所示。 $d^2x = \delta^2$ ，微三角形的面积是一个二阶无穷小，因此可以忽略，所以

$$S = f(x) dx = dF(x) = F(x + \delta) - F(x)$$

沿着这个思路，我们着手解决求不规则图形面积的问题。

考虑一个函数 $f(x)$ ，我们想要求曲线 $y = f(x)$ 、直线 $x = a$ 、 $x = b$ 以及 x 轴所围成的曲边梯形的面积 A ，如下图。

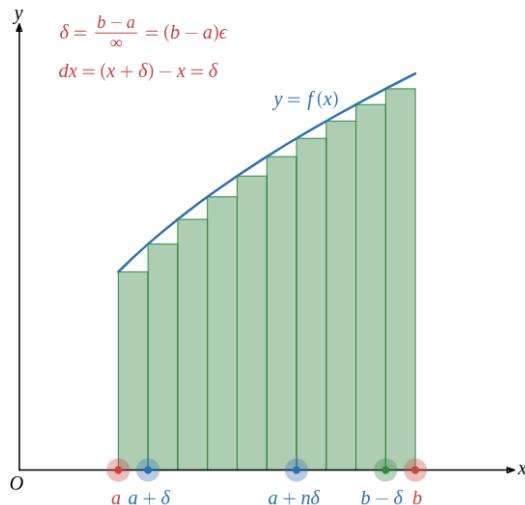


图 7.2: 区间微分

计算这个面积的基本思路是：

(1) 区间微分

分割：将区间 $[a, b]$ 分割成 ∞ 个小区间，区间宽度 $\delta = \frac{b-a}{\infty} = (b-a)\epsilon$ 。

右微分：对于区间 $[a, b - \delta]$ 内任意分割点 x ，微分形式： $dx = (x + \delta) - x = \delta$ 。

左微分：对于区间 $[a + \delta, b]$ 内任意分割点 x ，微分形式： $dx = x - (x - \delta) = \delta$ 。

微分面积： $dA = f(x) dx$

(2) 面积计算

右微分 x 取值:

$$x = a, a + \delta, a + 2\delta, \dots, a + (\infty - 2)\delta, a + (\infty - 1)\delta$$

注意: $a + (\infty - 1)\delta = b - \delta, a + (\infty - 2)\delta = b - 2\delta, \dots$

左微分 x 取值:

$$x = a + \delta, a + 2\delta, \dots, a + (\infty - 2)\delta, a + (\infty - 1)\delta, b$$

微分面积 dA 和对应点的 $F(x)$ 线性极限变化量精确相等:

$$dA = f(x) dx = dF(x) \quad (7.1)$$

对应关系及面积计算如下表所示:

n 值	x 值	dA	$dF(x)$
0	$a + 0\delta$	$f(a + 0\delta) dx$	$F(a + 1\delta) - F(a + 0\delta)$
1	$a + 1\delta$	$f(a + 1\delta) dx$	$F(a + 2\delta) - F(a + 1\delta)$
2	$a + 2\delta$	$f(a + 2\delta) dx$	$F(a + 3\delta) - F(a + 2\delta)$
\vdots	\vdots	\vdots	\vdots
$\infty - 2$	$a + (\infty - 2)\delta$	$f(a + (\infty - 2)\delta) dx$	$F(b - \delta) - F(b - 2\delta)$
$\infty - 1$	$a + (\infty - 1)\delta$	$f(a + (\infty - 1)\delta) dx$	$F(b) - F(b - \delta)$

表中面积列 dA 的每行和对应的 $F(x)$ 线性变化量 $dF(x)$ 列的每行都精确相等, 两列分别相加, 有

$$\begin{aligned} & f(a) dx + f(a + \delta) dx + \cdots + f(a + (\infty - 2)\delta) dx + f(a + (\infty - 1)\delta) dx \\ &= F(a + \delta) - F(a) + F(a + 2\delta) - F(a + \delta) + \cdots + F(b) - F(b - \delta) \\ &= F(b) - F(a) \end{aligned}$$

$$\sum_{n=0}^{\infty-1} f(a + n\delta) dx = F(b) - F(a)$$

曲边梯形的面积 A 等于面积列的总和, 所以

$$\delta = \frac{b - a}{\infty} = (b - a)\epsilon \quad (7.2)$$

$$dx = (x + \delta) - x = x - (x - \delta) = \delta \quad (7.3)$$

$$A = \sum_{n=0}^{\infty-1} f(a + n\delta) dx = F(b) - F(a) \quad (\text{右微分}) \quad (7.4)$$

$$A = \sum_{n=1}^{\infty} f(a + n\delta) dx = F(b) - F(a) \quad (\text{左微分}) \quad (7.5)$$

(3) 积分表示

应用积分符号，并限定积分区间，我们得到定积分计算曲边梯形面积的表示形式：

$$\int_a^b f(x) dx = \sum_{n=1}^{\infty} f(a + n\delta) dx = F(b) - F(a) = F(x) \Big|_{x=a}^{x=b} \quad (7.6)$$

a 和 b 分别是积分下限和积分上限，定义了积分的区间。

定积分 $\int_a^b f(x) dx$ 在几何上表示函数 $y = f(x)$ 的图像、直线 $x = a$ 、 $x = b$ 和 x 轴所围成的有向面积。

定积分的结果是一个确定的数值。

在 x 轴上方的面积取正值。

在 x 轴下方的面积取负值。

定积分的无限和形式（及变种）主要用于近似计算，数学精确计算采用不定积分计算原函数再应用积分区间，如公式右侧的形式。

传统上定义的定积分也可称作区间积分。

7.2 定积分的主要性质

1. 微积分基本定理

$$\int_a^b f(x) dx = F(b) - F(a)$$

其中 $F(x)$ 是 $f(x)$ 的任意一个原函数（即 $F'(x) = f(x)$ ）。

这是连接不定积分（原函数）与定积分的最重要定理。

定积分与积分变量无关：定积分的结果是一个数值，与积分变量的符号无关。

$$\int_a^b f(x) dx = \int_a^b f(t) dt = \int_a^b f(u) du$$

2. 线性性质

齐次性（常数乘法）：被积函数乘以一个常数 k 的定积分，等于这个常数乘以原函数的定积分。

$$\int_a^b k \cdot f(x) dx = k \int_a^b f(x) dx \quad (k \text{ 是常数})$$

加减性：两个函数的和（或差）的定积分，等于它们各自定积分的和（或差）。

$$\int_a^b [f(x) \pm g(x)] dx = \int_a^b f(x) dx \pm \int_a^b g(x) dx$$

3. 区间性质

定积分上下限的对调：对调定积分的上下限，积分值会改变符号。

$$\int_a^b f(x) dx = - \int_b^a f(x) dx$$

零长度区间：如果积分区间的上下限相同，则定积分的值为零。

$$\int_a^a f(x) dx = 0$$

区间可加性（对积分区间的拆分与合并）：如果 c 是 $[a, b]$ 区间上的任意一点（不一定非要在 a 和 b 之间），定积分可以在该点处进行拆分或合并。

$$\int_a^c f(x) dx + \int_c^b f(x) dx = \int_a^b f(x) dx$$

对称区间：

偶函数在对称区间上的定积分等于它在 $[0, a]$ 区间上定积分的两倍。

$$\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx \quad (\text{当 } f(x) \text{ 为偶函数时})$$

奇函数在对称区间上的定积分恒等于零。

$$\int_{-a}^a f(x) dx = 0 \quad (\text{当 } f(x) \text{ 为奇函数时})$$

例 7.2.1. 计算定积分。

$$\begin{aligned} \int_{-1}^1 (x^2 + 1) dx &= 2 \int_0^1 (x^2 + 1) dx \\ &= 2 \left[\frac{x^3}{3} + x \right]_0^1 = 2 \left[\left(\frac{1^3}{3} + 1 \right) - 0 \right] \\ &= 2 \left(\frac{1}{3} + 1 \right) = 2 \cdot \frac{4}{3} = \frac{8}{3} \end{aligned}$$

$$\begin{aligned} \int_{-2}^2 x^3 dx &= \left[\frac{x^4}{4} \right]_{-2}^2 \\ &= \frac{2^4}{4} - \frac{(-2)^4}{4} = \frac{16}{4} - \frac{16}{4} = 0 \end{aligned}$$

在遇到复杂的定积分时，如果积分区间是对称的，可以将函数拆分为奇函数部分和偶函数部分，然后分别应用上述性质进行简化：

$$\begin{aligned}\int_{-a}^a [f_{\text{even}}(x) + f_{\text{odd}}(x)] dx &= \int_{-a}^a f_{\text{even}}(x) dx + \int_{-a}^a f_{\text{odd}}(x) dx \\ &= 2 \int_0^a f_{\text{even}}(x) dx + 0\end{aligned}$$

例 7.2.2. 计算定积分 $\int_{-\pi}^{\pi} (x^5 + \cos x) dx$ 。

x^5 是奇函数，其积分在 $[-\pi, \pi]$ 上为 0。

$\cos x$ 是偶函数。因此

$$\begin{aligned}\int_{-\pi}^{\pi} (x^5 + \cos x) dx &= 0 + \int_{-\pi}^{\pi} \cos x dx \\ &= 2 \int_0^{\pi} \cos x dx = 2[\sin x]_0^{\pi} \\ &= 2(\sin \pi - \sin 0) = 2(0 - 0) = 0\end{aligned}$$

4. 比较性质

这些性质允许我们比较不同函数的定积分大小，通常假设 $a \leq b$ 。

非负性：如果在积分区间 $[a, b]$ 上，被积函数 $f(x)$ 恒大于等于零 ($f(x) \geq 0$)，那么它的定积分也大于等于零。

$$\text{若 } f(x) \geq 0 \text{ 且 } a \leq b, \text{ 则 } \int_a^b f(x) dx \geq 0$$

保序性：如果在积分区间 $[a, b]$ 上，一个函数 $f(x)$ 始终小于等于另一个函数 $g(x)$ ($f(x) \leq g(x)$)，那么 $f(x)$ 的定积分也小于等于 $g(x)$ 的定积分。

$$\text{若 } f(x) \leq g(x) \text{ 且 } a \leq b, \text{ 则 } \int_a^b f(x) dx \leq \int_a^b g(x) dx$$

有界性：如果函数 $f(x)$ 在区间 $[a, b]$ 上的最大值为 M ，最小值为 m ，那么定积分的值介于 $m(b-a)$ 和 $M(b-a)$ 之间。

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a)$$

5. 变限积分函数

如果定积分的上限（或下限）是变量，这个定积分就变成了一个以该变量为自变量的函数。我们称这种函数为变限积分函数。

$$F(x) = \int_a^x f(t) dt$$

其中 a 是常数下限, x 是变量上限, t 是积分变量 (哑变量)。

变限积分函数 $F(x)$ 实际上是 $f(x)$ 的一个原函数。

变限积分函数最重要的性质是它的导数:

$$F'(x) = \frac{d}{dx} \int_a^x f(t) dt = f(x) \quad (7.7)$$

链式法则应用: 如果上限不是 x , 而是关于 x 的函数 $u(x)$

$$\frac{d}{dx} \int_a^{u(x)} f(t) dt = f(u(x)) \cdot u'(x) \quad (7.8)$$

根据上面的求导法则, $F'(x) = f(x)$, 这证明了任何连续函数都存在原函数。

变限积分函数是许多求解微分方程的方法的基础。

例 7.2.3. 给定变限积分函数 $F(x)$:

$$F(x) = \int_1^x (t^3 + 2t) dt$$

求 $F(x)$ 的解析表达式和导函数 $F'(x)$ 。

$$\begin{aligned} \int_1^x (t^3 + 2t) dt &= \left[\frac{t^4}{4} + t^2 + C \right]_1^x \\ &= \left(\frac{x^4}{4} + x^2 \right) - \left(\frac{1^4}{4} + 1^2 \right) \\ &= \frac{x^4}{4} + x^2 - \frac{5}{4} \end{aligned}$$

$F(x)$ 的解析表达式:

$$F(x) = \frac{x^4}{4} + x^2 - \frac{5}{4}$$

导函数 $F'(x)$:

$$\begin{aligned} F'(x) &= \frac{d}{dx} \left(\frac{1}{4}x^4 + x^2 - \frac{5}{4} \right) \\ &= \frac{1}{4} \cdot 4x^3 + 2x - 0 \\ &= x^3 + 2x \end{aligned}$$

如果例子中上限是 $u(x) = \sin x$, 则

$$H(x) = \int_1^{\sin x} (t^3 + 2t) dt = \frac{1}{4} \sin^4 x + \sin^2 x - \frac{5}{4}$$

$$\begin{aligned}
 H'(x) &= \frac{d}{dx} \int_1^{\sin x} (t^3 + 2t) dt \\
 &= (\sin^3 x + 2 \sin x) \cdot \frac{d \sin x}{dx} \\
 &= (\sin^3 x + 2 \sin x) \cos x
 \end{aligned}$$

6. 广义积分

定积分的上限（或下限）是开区间或无穷大时，这被称为广义积分。

广义积分（Improper Integral），也常被称为瑕积分或反常积分。

在格洛克代数空间，开区间可以转换为闭区间，因此广义积分和普通积分没有什么不同，除了最后需要进行极限计算。

$$\begin{aligned}
 \int_a^\infty f(x) dx &= F(\infty) - F(a) \\
 \int_{-\infty}^b f(x) dx &= F(b) - F(-\infty) \\
 \int_{-\infty}^\infty f(x) dx &= \int_{-\infty}^C f(x) dx + \int_C^\infty f(x) dx = F(\infty) - F(-\infty)
 \end{aligned}$$

$$\begin{aligned}
 \int_a^{b-\epsilon} f(x) dx &= F(b-\epsilon) - F(a) \\
 \int_{a+\epsilon}^b f(x) dx &= F(b) - F(a+\epsilon)
 \end{aligned}$$

例 7.2.4. 计算广义积分 $\int_1^\infty \frac{1}{x^2} dx$ 。

$$\begin{aligned}
 \int_1^\infty \frac{1}{x^2} dx &= \int_1^\infty x^{-2} dx \\
 &= \left[\frac{x^{-1}}{-1} \right]_1^\infty = \left[-\frac{1}{x} \right]_1^\infty \\
 &= \left(-\frac{1}{\infty} - \left(-\frac{1}{1} \right) \right) \\
 &= -\epsilon + 1 = 1
 \end{aligned}$$

极限存在且为有限值 1，因此该广义积分收敛，其值为 1。

例 7.2.5. 计算广义积分 $\int_0^1 \frac{1}{\sqrt{x}} dx$ 。

被积函数 $f(x) = \frac{1}{\sqrt{x}}$ 在积分区间的下限 $x = 0$ 处没有定义，且 $\lim_{x \rightarrow 0^+} \frac{1}{\sqrt{x}} = +\infty$ 。因此 $x = 0$ 是一个瑕点。

$$\begin{aligned}\int_0^1 \frac{1}{\sqrt{x}} dx &= \int_{\epsilon}^1 x^{-1/2} dx \\&= \left[\frac{x^{1/2}}{1/2} \right]_{\epsilon}^1 = 2\sqrt{x} \Big|_{\epsilon}^1 \\&= 2\sqrt{1} - 2\sqrt{\epsilon} = 2\end{aligned}$$

极限存在且为有限值 2，因此该广义积分收敛，其值为 2。

第八章 定积分的应用

定积分作为微积分的核心工具，在数学、物理、工程、经济和生物等多个领域有广泛应用。它通过累积微小变化来计算总量，体现了极限思想的实际价值。

在几何方面，定积分常用于计算不规则形状的几何量，如弧长、面积、体积和曲面面积。

在物理中，定积分用于计算连续变化的量，如功、力矩、质心和流体静力。

在工程学、经济学、生物学与医学等方面都有广泛的应用。

8.1 曲线的弧长

曲线的弧长，简单来说，就是曲线上两点之间那段路径的精确长度。

要计算平面曲线 $y = f(x)$ 从 $x = a$ 到 $x = b$ 之间的弧长 s ，我们需要使用定积分。

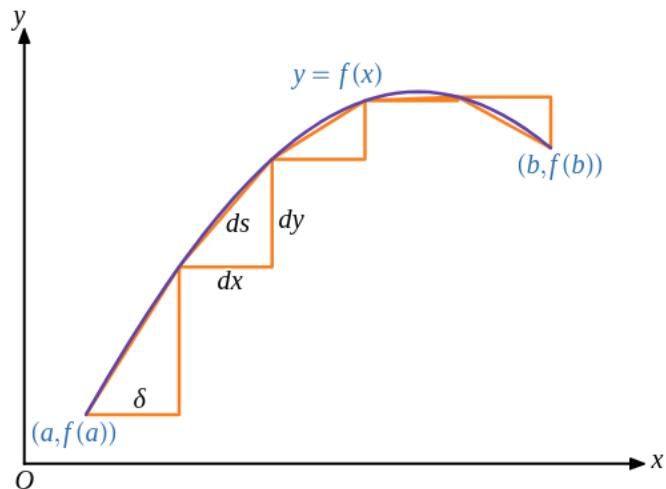


图 8.1: 弧长微分

弧长公式的推导

(1) 区间微分:

$$\delta = \frac{b-a}{\infty} = (b-a)\epsilon$$

$$dx = (x + \delta) - x = \delta$$

$$dy = df(x) = f'(x) dx$$

(2) 弧长微分 ds : 根据勾股定理

$$d^2s = d^2x + d^2y = d^2x + [f'(x)]^2 \cdot d^2x$$

$$ds = \sqrt{d^2x + [f'(x)]^2 \cdot d^2x} = \sqrt{1 + [f'(x)]^2} dx$$

(3) 定积分求和: 将这些微小的弧长 ds 从 $x = a$ 累加到 $x = b$, 得到总弧长 s

$$s = \int_a^b ds = \int_a^b \sqrt{1 + [f'(x)]^2} dx \quad (8.1)$$

$$= \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \quad (8.2)$$

曲线 $x = g(y)$ 的弧长

类似地, 如果曲线由函数 $x = g(y)$ 定义, 且 $g'(y)$ 在闭区间 $[c, d]$ 上连续, 则从 $y = c$ 到 $y = d$ 的弧长 s 为:

$$s = \int_c^d \sqrt{1 + [g'(y)]^2} dy \quad (8.3)$$

$$s = \int_c^d \sqrt{1 + \left(\frac{dx}{dy}\right)^2} dy \quad (8.4)$$

弧长计算步骤

确定曲线的函数形式 ($y = f(x)$ 或 $x = g(y)$), 并求出它的导数 $\frac{dy}{dx}$ 或 $\frac{dx}{dy}$ 。

将导数代入相应的弧长积分公式。

确定积分的上下限 (x 或 y 的范围)。

计算定积分以得出最终的弧长值。

例 8.1.1. 计算曲线 $y = \frac{2}{3}x^{3/2}$ 从 $x = 0$ 到 $x = 3$ 的弧长。

求导:

$$\frac{dy}{dx} = \frac{2}{3} \cdot \frac{3}{2} x^{3/2-1} = x^{1/2} = \sqrt{x}$$

代入公式:

$$\left(\frac{dy}{dx}\right)^2 = (\sqrt{x})^2 = x$$

$$\sqrt{1 + \left(\frac{dy}{dx}\right)^2} = \sqrt{1+x}$$

设置限值并构建积分：

$$s = \int_0^3 \sqrt{1+x} dx$$

计算积分：

$$\begin{aligned} \int_0^3 \sqrt{1+x} dx &= \int_0^3 (1+x)^{\frac{1}{2}} d(1+x) \\ &= \left[\frac{1}{1+1/2} (1+x)^{1+1/2} \right]_0^3 \\ &= \left[\frac{2}{3} (1+x)^{\frac{3}{2}} \right]_0^3 \\ &= \frac{2}{3} (1+3)^{\frac{3}{2}} - \frac{2}{3} (1+0)^{\frac{3}{2}} \\ &= \frac{2}{3} \cdot 8 - \frac{2}{3} = \frac{14}{3} \end{aligned}$$

因此，该曲线在 $x = 0$ 到 $x = 3$ 之间的弧长是 $\frac{14}{3}$ 。

弧长计算的难点通常在于积分的计算，因为 $\sqrt{1+[f'(x)]^2}$ 形式的被积函数往往会产生难以求出反导数的表达式。

弧长函数

弧长函数定义为从曲线上的固定起点 $A(a, f(a))$ 到曲线上任意一点 $P(x, f(x))$ 的弧长。

它是一个以 x 为变量的变限积分函数：

$$s(x) = \int_a^x \sqrt{1+[f'(t)]^2} dt \quad (8.5)$$

其中：

a 是起点的横坐标。

x 是终点的横坐标（也是函数的变量）。

t 是积分中的哑变量（dummy variable）。

弧长函数最主要的用途之一是作为曲线的自然参数。当曲线的参数是其弧长 s 时，曲线的参数化被称为弧长参数化。在这种参数化下，曲线的切向量的模（长度）恒为 1，这极大地简化了微分几何中的许多公式和计算，例如曲率和挠率的计算。

在物理学中，弧长函数可以用来确定一个物体沿着给定路径（曲线）移动的距离。

参数方程表示的曲线弧长

假设一条曲线由参数方程定义：

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad (\alpha \leq t \leq \beta)$$

那么

$$\begin{aligned} d^2s &= d^2x + d^2y \\ d^2s &= d^2x(t) + d^2y(t) \\ &= [x'(t)]^2 \cdot d^2t + [y'(t)]^2 \cdot d^2t \end{aligned}$$

$$ds = \sqrt{[x'(t)]^2 + [y'(t)]^2} dt$$

那么该曲线从 $t = \alpha$ 到 $t = \beta$ 的弧长 s 可以用以下定积分表示：

$$s = \int_{\alpha}^{\beta} \sqrt{[x'(t)]^2 + [y'(t)]^2} dt \quad (8.6)$$

或者写成更常见的形式：

$$s = \int_{\alpha}^{\beta} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} dt \quad (8.7)$$

例 8.1.2. 求半径为 R 的圆的周长：圆的参数方程为：

$$\begin{cases} x = R \cos t \\ y = R \sin t \end{cases} \quad (0 \leq t \leq 2\pi)$$

求导： $\frac{dx}{dt} = -R \sin t$, $\frac{dy}{dt} = R \cos t$

圆的周长：

$$\sqrt{(-R \sin t)^2 + (R \cos t)^2} = \sqrt{R^2(\sin^2 t + \cos^2 t)} = R$$

$$s = \int_0^{2\pi} R dt = Rt \Big|_0^{2\pi} = 2\pi R$$

在实际计算中，被积函数 $\sqrt{[x'(t)]^2 + [y'(t)]^2}$ 往往带有根号，计算可能会比较复杂，有时需要用到三角代换或特殊的积分技巧。

8.2 曲线之间的面积

如果想计算两条曲线 $y = f(x)$ 和 $y = g(x)$ 在区间 $x = a$ 到 $x = b$ 之间所围成的面积 A , 且在这个区间内 $f(x) \geq g(x)$ (即 $f(x)$ 的曲线始终在 $g(x)$ 的曲线上方或与它相交), 如图所示

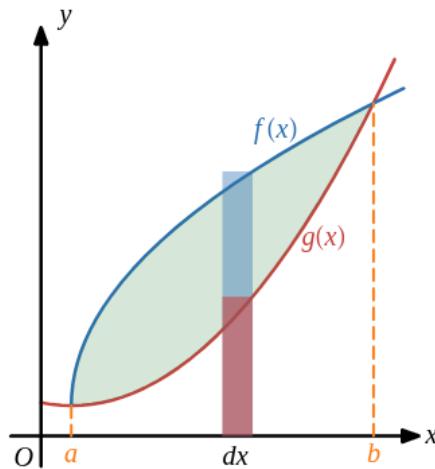


图 8.2: 曲线之间的面积

那么这个面积 A 可以通过以下定积分来计算:

$$\begin{aligned} dA_1 &= f(x) dx \\ dA_2 &= g(x) dx \\ dA &= dA_1 - dA_2 = f(x) dx - g(x) dx \\ dA &= [f(x) - g(x)] dx \end{aligned}$$

等式两边进行区间积分:

$$A = \int_a^b [f(x) - g(x)] dx$$

面积计算步骤

(1) 找到交点 (确定上下限 a 和 b):

将两个函数 $f(x)$ 和 $g(x)$ 设置相等, $f(x) = g(x)$, 解出 x 的值。这些 x 值就是曲线的交点, 它们通常作为积分的上下限 a 和 b 。如果问题已经给出了区间, 则使用给定的区间端点。

(2) 确定哪个函数在上方:

在积分区间 (a, b) 内, 选择一个测试点 c , 计算 $f(c)$ 和 $g(c)$ 的值。值较大的函数就是上方的函数 $f_{top}(x)$, 值较小的函数是下方的函数 $f_{bottom}(x)$ 。

如果在整个区间内，两条曲线有交叉（即哪个函数在上方发生了变化），您需要将积分分成多个部分，并在每个部分重新确定上下函数。面积 A 总是 $\int |f(x) - g(x)| dx$ 。

(3) 设置并计算定积分：

使用正确的上方函数 $f_{top}(x)$ 和下方函数 $f_{bottom}(x)$ 设置积分

$$A = \int_a^b [f_{top}(x) - f_{bottom}(x)] dx \quad (8.8)$$

然后计算这个定积分，得到面积。

例 8.2.1. 计算曲线 $y = x^2$ 和 $y = x$ 之间的面积。

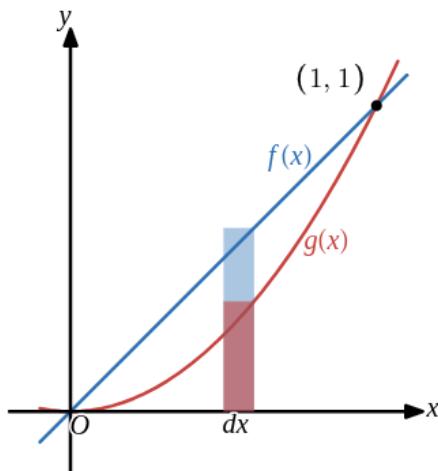


图 8.3: 面积计算

找交点： $x^2 = x \iff x = 0, x = 1$ 。因此， $a = 0, b = 1$ 。

确定上下函数：

$(0.5)^2 < (0.5)$ ，所以 $y = x$ 是上方的函数 ($f_{top}(x) = x$)，而 $y = x^2$ 是下方的函数 ($f_{bottom}(x) = x^2$)。

计算面积：

$$\begin{aligned} A &= \int_0^1 (x - x^2) dx \\ &= \left[\frac{1}{2}x^2 - \frac{1}{3}x^3 \right]_0^1 = \left(\frac{1}{2} - \frac{1}{3} \right) \\ &= \frac{1}{6} \end{aligned}$$

曲线之间的面积是 $1/6$ 平方单位。

在某些情况下，如果曲线方程是 x 作为 y 的函数（即 $x = f(y)$ 和 $x = g(y)$ ），或者如果用垂直矩形（对 x 积分）计算面积会要求将积分分解成多段，那么使用水平矩形（对 y 积分）会更简单。

$$A = \int_c^d [f(y) - g(y)] dy$$

$f(y)$: 右侧的函数（曲线）。

$g(y)$: 左侧的函数（曲线）。

d 和 c : 积分的上下限， y 的值。

例 8.2.2. 计算由抛物线 $x = 4 - y^2$ 和直线 $x = y$ 所围成的面积。

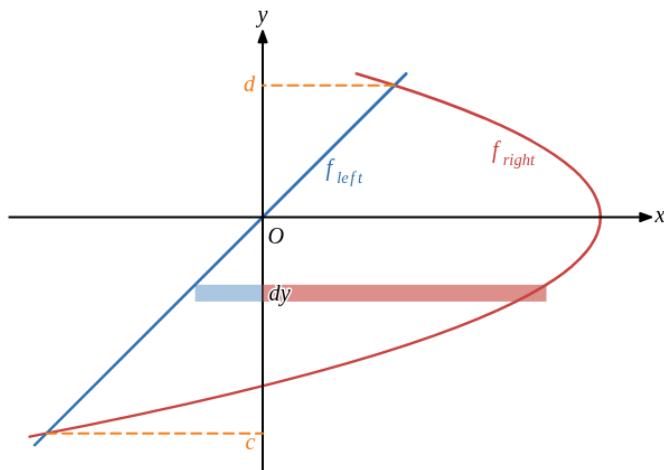


图 8.4: 水平矩形

找到交点（确定上下限 c 和 d ）：

将两个函数的 x 设置相等，以找到它们相交的 y 值：

$$4 - y^2 = y$$

$$y^2 + y - 4 = 0$$

使用求根公式 $y = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ ：

$$\begin{aligned} y &= \frac{-1 \pm \sqrt{1^2 - 4(1)(-4)}}{2(1)} \\ &= \frac{-1 \pm \sqrt{1 + 16}}{2} \\ &= \frac{-1 \pm \sqrt{17}}{2} \end{aligned}$$

所以，积分的下限 $c = \frac{-1-\sqrt{17}}{2}$ (约 -2.56)，上限 $d = \frac{-1+\sqrt{17}}{2}$ (约 1.56)。

确定哪个函数在右侧 ($f_{right}(y)$)：

在 y 的区间 (c, d) 内，选择一个测试点，例如 $y = 0$ 。

对于 $x = 4 - y^2$ (抛物线): $x(0) = 4 - 0^2 = 4$

对于 $x = y$ (直线): $x(0) = 0$

设置并计算定积分：

$$\begin{aligned} A &= \int_c^d [f_{right}(y) - f_{left}(y)] dy \\ &= \int_c^d [(4 - y^2) - y] dy \\ &= \int_c^d (4 - y - y^2) dy \\ &= \left[4y - \frac{y^2}{2} - \frac{y^3}{3} \right]_c^d \end{aligned}$$

这个具体的代入计算会非常复杂，但我们可以利用公式来简化。对于二次方程 $ay^2 + by + c = 0$ 所围成的面积，如果 y_1 和 y_2 是它的两个根，则面积 A 为：

$$A = \frac{|\Delta|^{3/2}}{6a^2}$$

在这个例子中，被积函数是 $-(y^2 + y - 4)$ 。我们的原始二次方程是 $y^2 + y - 4 = 0$ 。

这里 $a = 1, b = 1, c = -4$ 。判别式 $\Delta = b^2 - 4ac = 1^2 - 4(1)(-4) = 17$ 。

因此，面积 A 为：

$$A = \frac{(17)^{3/2}}{6(1)^2} = \frac{17\sqrt{17}}{6}$$

椭圆的面积

考虑中心在原点的标准椭圆方程：

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

其中 a 是长半轴， b 是短半轴。

椭圆在四个象限内是完全对称的。因此，我们只需要计算第一象限（即 x 从 0 到 a ）的面积，然后乘以 4 即可。

在第二章我们学习到椭圆的参数方程

$$\begin{cases} x = a \cos \theta \\ y = b \sin \theta \end{cases} \quad (\theta \in [0, 2\pi])$$

在第一象限，注意到 x 从 0 到 a 时， θ 由 $\frac{\pi}{2}$ 变化到 0。

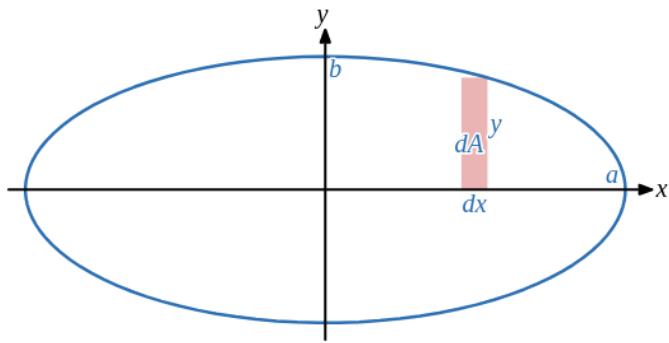


图 8.5: 椭圆的面积

椭圆的面积:

$$\begin{aligned}
 dA &= y dx = b \sin \theta da \cos \theta \\
 &= b \sin \theta \cdot (-a \sin \theta d\theta) = -ab \sin^2 \theta d\theta \\
 A &= 4 \int_{\frac{\pi}{2}}^0 (-ab \sin^2 \theta) d\theta = 4ab \int_0^{\frac{\pi}{2}} \sin^2 \theta d\theta \\
 &= 4ab \int_0^{\frac{\pi}{2}} \frac{1 - \cos 2\theta}{2} d\theta \\
 &= 2ab \int_0^{\frac{\pi}{2}} d\theta - ab \int_0^{\frac{\pi}{2}} \cos 2\theta d2\theta \\
 &= \pi ab
 \end{aligned}$$

椭圆面积公式: $S = \pi ab$

如果 $a = b = r$ (圆的情况), 公式变为 $S = \pi r^2$, 这验证了公式的准确性。

直观理解: 椭圆可以看作是将半径为 a 的圆在 y 轴方向上压缩 (或拉伸) 了 b/a 倍。

8.3 极坐标下的面积和弧长

极坐标系 (Polar Coordinate System) 是一种不同于我们常用的直角坐标系 (笛卡尔坐标系) 的定位方式。它不使用 “左右、上下” 来定位, 而是通过 “距离和角度” 来确定点的位置。

这种坐标系在处理圆形、螺旋线或者与中心点相关的运动 (如雷达、卫星轨道) 时非常高效。

1. 基本概念

在极坐标系中, 平面上的任何一点 P 都可以用一对坐标 (r, θ) 来表示:

极点 (Pole): 相当于直角坐标系的原点 $(0, 0)$ 。

极轴 (Polar Axis): 从极点引出的一条水平射线，通常指向右侧（相当于 x 轴的正方向）。

极径 (r): 点 P 到极点的距离。

极角 (θ): 极轴按逆时针方向旋转到极径 OP 所成的角度。

2. 坐标表示法

通常极角 θ 使用弧度 (Radians) 表示，但在初学者教程中也常用角度 ($^{\circ}$)。

逆时针旋转: 角度为正。

顺时针旋转: 角度为负。

示例：

$(3, 45^{\circ})$ 表示距离中心 3 个单位，角度为 45 度。

$(2, \pi)$ 表示距离中心 2 个单位，角度为 180 度。

3. 极坐标与直角坐标的转换

如果需要在两种系统之间切换，可以使用以下三角函数公式：

从极坐标 (r, θ) 转为直角坐标 (x, y) :

$$x = r \cos(\theta) \quad (8.9)$$

$$y = r \sin(\theta) \quad (8.10)$$

从直角坐标 (x, y) 转为极坐标 (r, θ) :

$$r = \sqrt{x^2 + y^2} \quad (8.11)$$

$$\theta = \arctan\left(\frac{y}{x}\right) \quad (8.12)$$

计算 θ 时需根据 (x, y) 所在的象限调整角度。

4. 常见的极坐标方程图形

极坐标能用非常简单的公式画出极其复杂的曲线：

方程类型	公式示例	图形特征
圆	$r = a$	以原点为圆心，半径为 a 的圆
玫瑰线	$r = a \sin(n\theta)$	像花瓣一样的图形， n 决定花瓣数量
阿基米德螺旋线	$r = a\theta$	随着角度增大，距离匀速增加的螺旋
心形线	$r = a(1 - \sin \theta)$	形状像一颗心

在计算某些圆形的面积或重积分时，使用极坐标可以极大减少计算的复杂度。

5. 面积公式

在极坐标系中推导面积公式，其核心思想与直角坐标系下的定积分一致：将不规则图形切割成无数个微小的“基本单元”，求和后再取极限。

在直角坐标系中，我们使用的是“小矩形”；而在极坐标系中，最基本的单元是“小扇形”。

(1) 基本单元：扇形的面积

首先，我们需要回顾几何学中扇形的面积公式。对于一个半径为 r ，圆心角为 θ （弧度制）的扇形，其面积 A 为：

$$A = \frac{1}{2}r^2\theta$$

假设我们要计算由曲线 $r = f(\theta)$ 以及射线 $\theta = \alpha$ 和 $\theta = \beta$ 所围成的图形面积。

(2) 区间微分

将角度区间 $[\alpha, \beta]$ 等分成 ∞ 个小区间，微分小扇形的面积 dA ：

$$\begin{aligned}\delta &= \frac{\beta - \alpha}{\infty} = (\beta - \alpha)\epsilon \\ d\theta &= (\theta + \delta) - \theta = \theta - (\theta - \delta) = \delta \\ dA &= \frac{1}{2}f^2(\theta) d\theta\end{aligned}$$

(3) 区间积分

将所有小扇形的面积相加，得到总面积 A ：

$$A = \sum_{n=1}^{\infty} \frac{1}{2}f^2(\alpha + n\delta) d\theta = \int_{\alpha}^{\beta} \frac{1}{2}f^2(\theta) d\theta$$

(3) 极坐标下面积公式

极坐标下曲线 $r = f(\theta)$ 围成的面积公式为：

$$A = \frac{1}{2} \int_{\alpha}^{\beta} f^2(\theta) d\theta = \frac{1}{2} \int_{\alpha}^{\beta} r^2 d\theta \quad (8.13)$$

例 8.3.1. 计算心形线 $r = a(1 + \cos \theta)$ 的面积

由于心形线在 θ 从 0 变到 2π 时刚好绕原点一周，形成一个封闭图形，因此积分限为：

下限 $\alpha = 0$

上限 $\beta = 2\pi$

观察心形线的图形可以发现它关于极轴（ x 轴）对称。因此，我们可以先计算上半部分（0 到 π ），再将结果乘以 2。

根据公式 $A = \frac{1}{2} \int_{\alpha}^{\beta} r^2 d\theta$, 代入 $r = a(1 + \cos \theta)$:

$$\begin{aligned} A &= \frac{1}{2} \int_0^{2\pi} [a(1 + \cos \theta)]^2 d\theta \\ &= \frac{a^2}{2} \int_0^{2\pi} (1 + 2\cos \theta + \cos^2 \theta) d\theta \end{aligned}$$

用三角恒等式降次: $\cos^2 \theta = \frac{1+\cos(2\theta)}{2}$:

$$\begin{aligned} A &= \frac{a^2}{2} \int_0^{2\pi} \left(1 + 2\cos \theta + \frac{1 + \cos(2\theta)}{2} \right) d\theta \\ &= \frac{a^2}{2} \int_0^{2\pi} \left(\frac{3}{2} + 2\cos \theta + \frac{1}{2} \cos(2\theta) \right) d\theta \end{aligned}$$

现在逐项积分:

$$\begin{aligned} \int_0^{2\pi} \frac{3}{2} d\theta &= \frac{3}{2} \theta \Big|_0^{2\pi} = 3\pi \\ \int_0^{2\pi} 2\cos \theta d\theta &= 2 \sin \theta \Big|_0^{2\pi} = 0 \\ \int_0^{2\pi} \frac{1}{2} \cos(2\theta) d\theta &= \frac{1}{4} \int_0^{2\pi} \cos(2\theta) d2\theta = \frac{1}{4} \sin(2\theta) \Big|_0^{2\pi} = 0 \end{aligned}$$

将上述结果加总并乘以系数:

$$A = \frac{a^2}{2} \cdot (3\pi + 0 + 0) = \frac{3}{2}\pi a^2$$

心形线 $r = a(1 + \cos \theta)$ 所围成的面积为 $\frac{3}{2}\pi a^2$ 。

6. 弧长公式

在直角坐标系中, 微元长度 $ds = \sqrt{dx^2 + dy^2}$ 。在极坐标系中, 我们有转换关系:

$$x = r \cos \theta, \quad y = r \sin \theta$$

其中 r 是 θ 的函数 $r = f(\theta)$ 。对 x, y 关于 θ 求导:

$$\begin{aligned} x &= r \cos \theta, \quad y = r \sin \theta \\ dx &= dr \cos \theta = \cos \theta dr + r d \cos \theta \\ &= \cos \theta dr - r \sin \theta d\theta \\ dy &= dr \sin \theta = \sin \theta dr + r d \sin \theta \\ &= \sin \theta dr + r \cos \theta d\theta \end{aligned}$$

$$\begin{aligned}
 d^2x + d^2y &= (\cos \theta dr - r \sin \theta d\theta)^2 + (\sin \theta dr + r \cos \theta d\theta)^2 \\
 &= d^2r + r^2 d^2\theta \\
 &= \left[r^2 + \left(\frac{dr}{d\theta} \right)^2 \right] \cdot d^2\theta \\
 ds &= \sqrt{dx^2 + dy^2} = \sqrt{r^2 + \left(\frac{dr}{d\theta} \right)^2} d\theta
 \end{aligned}$$

如果曲线 $r = f(\theta)$ 在区间 $[\alpha, \beta]$ 上连续可导，则其弧长 s 为：

$$s = \int_{\alpha}^{\beta} \sqrt{r^2 + \left(\frac{dr}{d\theta} \right)^2} d\theta = \int_{\alpha}^{\beta} \sqrt{r^2 + (r')^2} d\theta \quad (8.14)$$

例 8.3.2. 计算心形线 $r = a(1 + \cos \theta)$ 的周长。

$$\begin{aligned}
 \frac{dr}{d\theta} &= -a \sin \theta \\
 r^2 + \left(\frac{dr}{d\theta} \right)^2 &= [a(1 + \cos \theta)]^2 + (-a \sin \theta)^2 \\
 &= a^2(1 + 2 \cos \theta + \cos^2 \theta + \sin^2 \theta) \\
 &= a^2(2 + 2 \cos \theta) = 2a^2(1 + \cos \theta)
 \end{aligned}$$

利用半角公式 $1 + \cos \theta = 2 \cos^2 \frac{\theta}{2}$ ：

$$\sqrt{r^2 + \left(\frac{dr}{d\theta} \right)^2} = \sqrt{4a^2 \cos^2 \frac{\theta}{2}} = 2a \left| \cos \frac{\theta}{2} \right|$$

考虑到对称性，我们计算上半部分（0 到 π ）再乘以 2。在 $0 \leq \theta \leq \pi$ 时， $\cos \frac{\theta}{2} \geq 0$ ，所以：

$$\begin{aligned}
 s &= 2 \int_0^{\pi} 2a \cos \frac{\theta}{2} d\theta \\
 &= 4a \cdot \left[2 \sin \frac{\theta}{2} \right]_0^{\pi} \\
 &= 4a \cdot (2 - 0) = 8a
 \end{aligned}$$

心形线 $r = a(1 + \cos \theta)$ 的总弧长（周长）为 $8a$ 。

8.4 旋转体的体积

旋转体的体积，简单来说，就是将一个平面图形绕着某条轴旋转 360° ，求所形成的几何体的体积。

最常用的方法有两种：磁盘法（Disk Method）和外壳法（Shell Method）。

1. 磁盘法

这种方法适用于切片垂直于旋转轴的情况。可以把旋转体想象成由无数个极薄的圆柱体（圆盘）堆叠而成。

如果曲线 $y = f(x)$ 绕 x 轴旋转，从 $x = a$ 到 $x = b$ ：

$$V = \int_a^b \pi f^2(x) dx \quad (8.15)$$

每个圆盘的半径是 $R = f(x)$ ，圆盘面积是 $A = \pi R^2$ ，厚度是 dx 。

如果是绕 y 轴旋转：公式变为

$$V = \int_c^d \pi g^2(x) dy \quad (8.16)$$

当平面图形是由两条曲线围成，且旋转轴不在图形边缘时，中间会产生空心。这就像是在大圆盘中间挖掉了一个小圆盘。

$$V = \int_a^b \pi [f^2(x) - g^2(x)] dx \quad (8.17)$$

R : 外半径（距离旋转轴较远的函数）。

r : 内半径（距离旋转轴较近的函数）。

例 8.4.1. 计算由曲线 $y = \sin(x)$ 在区间 $[0, \pi/2]$ 上，以及 $x = 0$ (y 轴) 和 $y = 1$ 围成的图形，绕 x 轴旋转一周所得旋转体的体积。

因为绕 x 轴旋转，切片垂直于 x 轴，变量为 x 。

这个图形在旋转时，上方由 $y = 1$ 限制，下方由 $y = \sin(x)$ 限制。

外半径 $R(x)$: 1

内半径 $r(x)$: $\sin(x)$

区间: $x \in [0, \pi/2]$

代入公式计算体积:

$$\begin{aligned} V &= \int_0^{\pi/2} \pi (1^2 - \sin^2 x) dx = \pi \int_0^{\pi/2} \cos^2 x dx \\ &= \frac{\pi}{2} \int_0^{\pi/2} (1 + \cos(2x)) dx \\ &= \frac{\pi}{2} \left[x + \frac{1}{2} \sin(2x) \right]_0^{\pi/2} \\ &= \frac{\pi}{2} \cdot \frac{\pi}{2} = \frac{\pi^2}{4} \end{aligned}$$

2. 圆柱外壳法

这种方法适用于切片平行于旋转轴的情况。可以把几何体想象成由一层层“洋葱皮”（圆柱壳）嵌套而成。

如果曲线 $y = f(x)$ 绕 y 轴旋转：

$$V = \int_a^b 2\pi x f(x) dx \quad (8.18)$$

每个外壳的半径是 x , 高度是 $f(x)$, 展开后的面积是 $2\pi \cdot \text{radius} \cdot \text{height}$ 。

当函数 $y = f(x)$ 很难反解成 $x = g(y)$ 时, 绕 y 轴旋转用外壳法通常更容易。

例 8.4.2. 求曲线 $y = 3x - x^2$ 与 x 轴围成的区域, 绕 y 轴旋转一周所得的体积。

分析图形

边界：这是一个开口向下的抛物线，交 x 轴于 $(0, 0)$ 和 $(3, 0)$ 。

旋转轴： y 轴。

在外壳法中, 想象我们在图形内部取一条垂直于 x 轴的“细条”进行旋转。

半径 (Radius): 细条到旋转轴 (y 轴) 的距离, 即 $r = x$ 。

高度 (Height): 细条的高度, 即曲线的纵坐标, $h = y = 3x - x^2$ 。

积分区间：从图形的左端点 $x = 0$ 到右端点 $x = 3$ 。

代入公式计算体积：

$$\begin{aligned} V &= \int_a^b 2\pi x f(x) dx \\ &= \int_0^3 2\pi \cdot x \cdot (3x - x^2) dx = 2\pi \int_0^3 (3x^2 - x^3) dx \\ &= 2\pi \left[x^3 - \frac{1}{4}x^4 \right]_0^3 \\ &= 2\pi \left((3^3 - \frac{1}{4} \cdot 3^4) - 0 \right) = 2\pi \left(27 - \frac{81}{4} \right) \\ &= 2\pi \left(\frac{108 - 81}{4} \right) = 2\pi \cdot \frac{27}{4} \\ &= \frac{27\pi}{2} \end{aligned}$$

旋转轴是 y 轴, 但函数是 $y = f(x)$: 这种“交叉配合”通常意味着外壳法更简单。

3. 磁盘法和外壳法的比较

在数学本质上, 它们是同一种思想的不同表现形式; 但在计算实践中, 它们是“切蛋糕”的两种不同刀法。

磁盘法：

切法：垂直于旋转轴切。

微元形状：像薄圆片（或带孔的垫圈）。

外壳法：

切法：平行于旋转轴切。

微元形状：像薄圆柱壳（类似洋葱圈或俄罗斯套娃）。

例 8.4.3. 计算由曲线 $y = x^2$ 、 $x = 1$ 和 $y = 0$ （即 x 轴）围成的图形，绕 y 轴旋转一周所得旋转体的体积。

方法 1：磁盘法

在磁盘法中，切片是垂直于旋转轴（ y 轴）的，所以切片是水平的，变量必须为 y 。

因为图形绕 y 轴旋转，且左侧有空隙，这会形成一个“带孔的垫圈”：

外半径 R ：固定的，最右边是 $x = 1$ 。

内半径 r ：由曲线决定。既然 $y = x^2$ ，那么 $x = \sqrt{y}$ 。所以内半径 $r = \sqrt{y}$ 。

积分区间：从 $y = 0$ 到 $y = 1$ （当 $x = 1$ 时， $y = 1^2 = 1$ ）。

代入公式计算体积：

$$\begin{aligned} V &= \int_0^1 \pi(R^2 - r^2) dy \\ &= \int_0^1 \pi(1^2 - (\sqrt{y})^2) dy = \pi \int_0^1 (1 - y) dy \\ &= \pi \left[y - \frac{1}{2}y^2 \right]_0^1 = \pi(1 - \frac{1}{2}) = \frac{\pi}{2} \end{aligned}$$

方法 2：圆柱外壳法

在外壳法中，切片是平行于旋转轴（ y 轴）的，所以切片是垂直的，变量为 x 。

半径 r ：离旋转轴的距离，即 x 。

高度 h ：函数的高度，即 $f(x) = x^2$ 。

积分区间：从 $x = 0$ 到 $x = 1$ 。

代入公式计算体积：

$$\begin{aligned} V &= \int_a^b 2\pi x f(x) dx \\ &= \int_0^1 2\pi \cdot x \cdot x^2 dx = 2\pi \int_0^1 x^3 dx \\ &= 2\pi \left[\frac{1}{4}x^4 \right]_0^1 = 2\pi \cdot \frac{1}{4} = \frac{\pi}{2} \end{aligned}$$

第九章 积分和积分方法

9.1 导数的定义与几何意义

函数 $f(x)$ 在点 x_0 处的导数定义为：

$$f'(x_0) = \lim_{\Delta x \rightarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

导数表示函数在某点变化率的精确度量。

第十章 多变量函数

二元函数的图像（曲面）：建立 3D 空间感。

偏导数：仅作为“固定其他方向、只看一个方向”的变化率。

全微分（线性近似）：强调 $df \approx f_x dx + f_y dy$ 。这是积分应用中 dA 、 ds 元素分解的理论根基。

针对“积分应用”的极简前置清单

A. 投影与区域描述 (Domain Description)

重积分最难的不是积分本身，而是确定积分限。

裁剪重点：教会读者如何把空间中的曲面投影到 xy 平面上，并用不等式描述这个区域。

前置意义：这是二重积分、三重积分的直接基础。

B. 局部线性化 (Local Linearity)

裁剪重点：不要讲二阶导数判别法，重点讲“切平面”的近似。

前置意义：在计算曲面积分时，我们需要把弯曲的表面切割成无数个微小的“小平面 (dS)”。如果读者理解了全微分，理解 $dS = \sqrt{1 + f_x^2 + f_y^2} dA$ 就会容易得多。

10.1 从平面到空间：定义域与可视化

核心目标：建立 3D 空间感，学会“描述区域”。

关键内容：二元函数的直观理解： $z = f(x, y)$ 是空间中的一张曲面。

等高线 (Level Curves)：重点讲解，因为它是后续理解“梯度垂直于等值线”的直观来源。

区域描述（核心必备）：学习用不等式描述 xy 平面上的区域（如圆、矩形、三角形），这是重积分定限的直接前置。

在单变量微积分中，我们研究的是 $y = f(x)$ ，它在平面上是一条线。当我们引入第二个变量 y ，研究 $z = f(x, y)$ 时，数学世界从“线”跃迁到了“面”。

1. 二元函数：空间中的“地形图”

想象你正站在一座山上，你所处的海拔高度 z 取决于你所在的经度 x 和纬度 y 。

单变量函数：像是一条过山车的轨道。

二元函数：则是整座山的表面。

每一个坐标点 (x, y) 都对应唯一的高度 z 。这种直观的物理对应，是我们理解后续所有偏导数和积分应用的基础。

当我们谈论 $z = f(x, y)$ 时，你可以把它想象成在 xy 平面（地面）上每一个点 (x, y) 处，都垂直竖起一根高度为 z 的柱子。所有这些柱子的顶端连在一起，就形成了一张曲面。

例子 A：平面 (The Plane) —— 坡度恒定的斜坡

方程： $z = ax + by + c$

视觉特征：就像一张斜铺在空中的无限大的硬纸板。

教学意义：这是最简单的地形。无论你在哪一点，坡度都是恒定的。它是以后理解切平面（全微分）的基石。

积分关联：如果我们要对这个平面进行积分，其实就是在求一个棱柱体的体积。

例子 B：抛物面 (The Paraboloid) —— 碗状山谷或山峰

方程： $z = x^2 + y^2$

视觉特征：一个完美的圆底大碗。

特征分析：在原点 $(0, 0)$ 处，它是平坦的（导数为 0）。

离原点越远，高度增加得越快（坡度越来越陡）。

坐标衔接：引导读者观察这个形状的轮廓：如果从上方俯视，它的等高线是同心圆。

例子 C：马鞍面 (The Saddle Surface) —— 复杂的交汇点

方程： $z = x^2 - y^2$

视觉特征：就像一个马鞍，或者山口 (Pass)。

特征分析：沿着 x 轴看，它是向上弯曲的谷底；沿着 y 轴看，它是向下弯曲的山脊。

教学意义：这告诉读者，多元函数的变化极其复杂，仅仅看一个方向（偏导数）是不够的，同一个点在不同方向可能有完全相反的变化趋势。

如何从方程“读出”地形？

固定 y ：想象用一把大刀沿着平行于 x 轴的方向切下去。剩下的曲线就是一个关于 x 的单变量函数。这正是偏导数正在做的事情。

固定 z ：想象用水平面去拦截这个地形。截出来的轮廓线就是等高线。这正是重积分限时寻找边界的方法。

2. 定义域：积分的“地基”

在单变量微积分里，定义域通常只是 x 轴上的一段区间 $[a, b]$ 。但在多变量微积分里，定义域变成了 xy 平面上的一块区域 (Region)。

用数学不等式准确地描述这块区域：

矩形区域： $a \leq x \leq b, c \leq y \leq d$ 。

圆盘区域： $x^2 + y^2 \leq R^2$ 。

一般区域：例如由两条曲线 $y = g(x)$ 和 $y = h(x)$ 围成的部分。

在处理二元函数 $z = f(x, y)$ 时，我们不仅关心高度，更关心高度下面的那块地基。在微积分里，这块地基被称为积分区域 D 。

复杂的区域归纳为两类：

A. X-型区域（垂直切割）

如果一个区域的左、右边界是两条直线，而上、下边界是两条曲线，我们称之为 X-型区域。

特征： x 的范围是常数，而 y 的范围取决于 x 。

数学描述：

$$a \leq x \leq b$$

$$g_1(x) \leq y \leq g_2(x)$$

B. Y-型区域（水平切割）

如果区域的上、下边界是直线，而左、右边界是曲线。

特征： y 的范围是常数，而 x 的范围取决于 y 。

数学描述：

$$c \leq y \leq d$$

$$h_1(y) \leq x \leq h_2(y)$$

经典案例：圆盘区域 (The Disk)

方程： $x^2 + y^2 \leq R^2$

直角坐标描述（麻烦但必须懂）：

$$-R \leq x \leq R$$

$$-\sqrt{R^2 - x^2} \leq y \leq \sqrt{R^2 - x^2}$$

3. 等高线 (Level Curves)

理解 3D 曲面最有效的手段是“等高线”——就像地理地图上标注海拔高度的线一样。

令 $f(x, y) = C$ （常数），在 xy 平面上画出的曲线就是等高线。

密集的等高线意味着山势陡峭（变化率大，即梯度大）。

稀疏的等高线意味着地势平缓。

10.2 偏导数：固定维度的变化

核心目标：降维理解变化率。

关键内容：

偏导的计算：把不相关的变量看作常数（算法层）。

几何意义：切线的斜率。只需强调：偏导数只关心沿坐标轴方向的变化。

裁剪掉：高阶导数、克莱罗定理（除非涉及极值判定）。

当你站在山坡上，想要描述地势的倾斜程度时，问题来了：你可以向四面八方走，每一个方向的坡度可能都不一样。

为了简化问题，微积分采取了“分而治之”的策略：先只看**正东（ x 轴）方向的变化，再只看正北（ y 轴）**方向的变化。

符号表示：我们用 $\frac{\partial f}{\partial x}$ 或 f_x 来表示对 x 的偏导数。

例子：设海拔高度方程为 $z = x^2 + 3xy + y^2$

求 f_x ：把 y 当成常数。 x^2 变成 $2x$, $3xy$ 变成 $3y$ （因为 x 被导掉了）， y^2 变成 0。

结果： $f_x = 2x + 3y$

求 f_y ：把 x 当成常数。 x^2 变成 0, $3xy$ 变成 $3x$, y^2 变成 $2y$ 。结果： $f_y = 3x + 2y$

几何意义：切线的斜率

这是最关键的可视化步骤。

f_x 的直观：想象用一把平行于 x 轴的刀，垂直切开山体。切口是一条曲线，这条曲线在某一点的切线斜率就是 f_x 。

f_y 的直观：同理，用平行于 y 轴的刀切开，得到的切线斜率就是 f_y 。

偏导数告诉我们，如果 x 挪动了一点点 (dx)，高度 z 会变化多少。

变化量 $\approx f_x \cdot dx$

偏导数本质上是“假装其他变量不存在”的单变量微积分。

哪些不讲？

不要讲：复杂的偏导数定义（极限形式）。只需让读者会算、懂几何意义即可。

不要讲：高阶偏导数的混合求导顺序证明。只需口头告知“通常情况下 $f_{xy} = f_{yx}$ ”。

不要讲：隐函数求导的复杂公式。在应用层，直接两边求导更符合直觉。

10.3 全微分与切平面

核心目标：为积分中的“微元法”和向量分析中的“切平面”打底。

关键内容：

全微分公式： $dz = f_x dx + f_y dy = \partial f(x) + \partial f(y)$ 。

线性近似的直观：“以直代曲”。解释为什么在极小尺度下，曲面可以看作平面。

切平面方程：为后续向量分析中寻找“法向量”埋下伏笔。

在单变量微积分中，我们学会了用切线来近似一条曲线。到了多变量世界，我们要用一个切平面来近似一块曲面。

从偏导数到全微分

偏导数 f_x 和 f_y 分别告诉了我们沿两个轴方向的“坡度”。如果我们把这两个变化结合起来，当 x 变动了 dx ，同时 y 也变动了 dy 时，函数高度 z 的总变化量是多少？

全微分公式：

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

直观理解：

总高度的变化，等于“ x 方向贡献的变化”加上“ y 方向贡献的变化”。

这里的 dz 是切平面上的高度变化，而 Δz 是真实曲面上的高度变化。当 dx, dy 极小时，这两者几乎相等。这就是线性近似。

切平面

想象你把一个足球无限放大，当你站在足球表面的一点时，你会觉得地面是平的。这个“平地”就是切平面。

切平面方程：利用全微分公式，如果我们已知点 (x_0, y_0, z_0) ，切平面的方程就是：

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

观察切平面方程，如果我们把它重新排列成：

$$f_x(x - x_0) + f_y(y - y_0) - (z - z_0) = 0$$

这完全符合向量代数中平面的点法式方程。这意味着，向量 $\mathbf{n} = (f_x, f_y, -1)$ 恰好是垂直于该平面的法向量。

现在，我们不需要通过复杂的几何观察去寻找法向量了。只要你会求偏导数，你就拥有了描述空间中任何曲面方向的能力。这种能力，我们称之为梯度。

10.4 链式法则：变量之间的传递链

在实际应用中，变量往往是层层嵌套的。例如：

空间中某点的温度 T 取决于坐标 (x, y, z) 。

而由于你正在运动，你的坐标 (x, y, z) 又取决于时间 t 。那么，随着时间的推移，你感受到的温度变化率 $\frac{dT}{dt}$ 是多少？

1. 核心工具：路径图法 (Tree Diagram)

不要试图背诵公式。处理链式法则最直观的方法是画变量关系图。

场景：设 $z = f(x, y)$, 而 $x = x(t), y = y(t)$ 。

第一层：画出 z 。

第二层：从 z 连出两条线，指向 x 和 y (代表偏导数 $\frac{\partial z}{\partial x}$ 和 $\frac{\partial z}{\partial y}$)。

第三层：从 x 连出一条线指向 t (代表 $\frac{dx}{dt}$), 从 y 连出一条线指向 t (代表 $\frac{dy}{dt}$)。

求导法则：走到底：沿着每一条路径走到终点 t , 并把路径上的导数相乘。

加总：把所有到达 t 的路径结果相加。

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}$$

坐标变换在二重积分中，我们经常需要从直角坐标 (x, y) 切换到极坐标 (r, θ) 。这时， $x = r \cos \theta, y = r \sin \theta$ 。如果我们想知道函数 f 随着半径 r 如何变化：

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r}$$

裁剪说明：

裁剪掉：隐函数求导公式 (如 $dy/dx = -F_x/F_y$)。直接告诉读者用全微分或链式法则对等式两边同时求导即可，没必要让读者多背一个容易记混正负号的公式。

裁剪掉：复杂的二阶偏导链式法则。除非读者要研究高深的物理波动方程，否则一阶链式法则已经涵盖了 90

第十一章 空间曲线

11.1 向量微分和曲线弧长

11.2 切向量 \mathbf{T} 、法向量 \mathbf{N} 和副法向量 \mathbf{B}

在微分几何中，Frenet-Serret 标架 $(\mathbf{T}, \mathbf{N}, \mathbf{B})$ 是描述曲线局部特征的核心。当曲线按弧长参数化时，切向量的模长恒为 1，这简化了所有的导数关系。

1. 切向量 \mathbf{T} 的推导

定义曲线为 $\mathbf{r}(s)$ 。切向量定义为位置对弧长的变化率：

$$\mathbf{T} = \frac{d\mathbf{r}}{ds}$$

由于 s 是弧长，根据定义 $\|\mathbf{T}\| = 1$ 。

2. 主法向量 \mathbf{N} 与曲率 κ

由于 \mathbf{T} 是单位向量，其模长平方为常数： $\mathbf{T} \cdot \mathbf{T} = 1$ 。两边对 s 求导：

$$\frac{d\mathbf{T}}{ds} \cdot \mathbf{T} + \mathbf{T} \cdot \frac{d\mathbf{T}}{ds} = 0 \implies 2\mathbf{T} \cdot \frac{d\mathbf{T}}{ds} = 0$$

这证明了导向量 $d\mathbf{T}/ds$ 始终垂直于 \mathbf{T} 。

曲率 κ ：定义为切线方向随弧长变化的速率，即 $\kappa = \| \frac{d\mathbf{T}}{ds} \|$ 。

主法向量 \mathbf{N} ：定义为 $d\mathbf{T}/ds$ 方向上的单位向量：

$$\frac{d\mathbf{T}}{ds} = \kappa \mathbf{N}$$

3. 副法向量 \mathbf{B} 与挠率 τ

为了构成右手正交标架，我们定义副法向量为：

$$\mathbf{B} = \mathbf{T} \times \mathbf{N}$$

挠率 τ 的得出：

我们要考察 \mathbf{B} 随弧长的变化。对 \mathbf{B} 求导：

$$\frac{d\mathbf{B}}{ds} = \frac{d\mathbf{T}}{ds} \times \mathbf{N} + \mathbf{T} \times \frac{d\mathbf{N}}{ds}$$

由于 $d\mathbf{T}/ds = \kappa\mathbf{N}$, 而 $\mathbf{N} \times \mathbf{N} = 0$, 第一项消失:

$$\frac{d\mathbf{B}}{ds} = \mathbf{T} \times \frac{d\mathbf{N}}{ds}$$

这意味着 $d\mathbf{B}/ds$ 垂直于 \mathbf{T} 。同时, 由于 \mathbf{B} 是单位向量, $d\mathbf{B}/ds$ 也垂直于 \mathbf{B} 。既然它同时垂直于 \mathbf{T} 和 \mathbf{B} , 它必须在 \mathbf{N} 的方向上。

因此, 我们定义:

$$\frac{d\mathbf{B}}{ds} = -\tau\mathbf{N}$$

这里的 τ 称为挠率 (Torsion), 负号是几何学上的约定, 表示当 $\tau > 0$ 时, 曲线随 s 增加向副法向量定义的右手螺旋方向扭转。

总结

\mathbf{T} 描述前进方向。

\mathbf{N} 描述向哪弯曲, 曲率 κ 是弯曲程度 (偏离直线的程度)。

\mathbf{B} 描述运动平面的法线, 挠率 τ 是扭曲程度 (偏离平面的程度)。对于平面曲线 (如 $y = x^2$), \mathbf{r}''' 依然在 xy 平面, 外积后与 z 轴垂直, 因此 τ 恒等于 0。

曲率 κ 的求解公式

曲率 κ 的基本定义是切向量对弧长的变化率的模:

$$\kappa = \left\| \frac{d\mathbf{T}}{ds} \right\|$$

由于我们通常使用的是参数 t (如时间), 根据链式法则:

$$\frac{d\mathbf{T}}{dt} = \frac{d\mathbf{T}}{ds} \cdot \frac{ds}{dt} = \frac{d\mathbf{T}}{ds} \cdot \|\mathbf{r}'(t)\|$$

因此有:

$$\kappa = \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|}$$

我们知道切向量 $\mathbf{T} = \frac{\mathbf{r}'}{\|\mathbf{r}'\|}$, 所以速度向量可以写为:

$$\mathbf{r}' = \|\mathbf{r}'\|\mathbf{T}$$

对上式两边关于 t 求导 (使用乘法法则):

$$\mathbf{r}'' = (\|\mathbf{r}'\|)' \mathbf{T} + \|\mathbf{r}'\| \mathbf{T}'$$

利用外积 (叉乘) 消项

现在我们将 \mathbf{r}' 和 \mathbf{r}'' 做外积:

$$\mathbf{r}' \times \mathbf{r}'' = (\|\mathbf{r}'\|\mathbf{T}) \times ((\|\mathbf{r}'\|)' \mathbf{T} + \|\mathbf{r}'\| \mathbf{T}')$$

利用外积的分配律：

$$\mathbf{r}' \times \mathbf{r}'' = \|\mathbf{r}'\|(\|\mathbf{r}'\|)'(\mathbf{T} \times \mathbf{T}) + \|\mathbf{r}'\|^2(\mathbf{T} \times \mathbf{T}')$$

因为任何向量与自身的外积为零 ($\mathbf{T} \times \mathbf{T} = 0$)，所以：

$$\mathbf{r}' \times \mathbf{r}'' = \|\mathbf{r}'\|^2(\mathbf{T} \times \mathbf{T}')$$

取两边的模长：

$$\|\mathbf{r}' \times \mathbf{r}''\| = \|\mathbf{r}'\|^2 \cdot \|\mathbf{T} \times \mathbf{T}'\|$$

因为 \mathbf{T} 是单位向量，且我们已知 \mathbf{T}' 垂直于 \mathbf{T} (见前文推导)，所以 $\|\mathbf{T} \times \mathbf{T}'\| = \|\mathbf{T}\| \|\mathbf{T}'\| \sin(90^\circ) = \|\mathbf{T}'\|$ 。代入上式：

$$\|\mathbf{r}' \times \mathbf{r}''\| = \|\mathbf{r}'\|^2 \|\mathbf{T}'\|$$

解出 $\|\mathbf{T}'\|$ ：

$$\|\mathbf{T}'\| = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^2}$$

最后，将这个结果代回最初的曲率定义：

$$\kappa = \frac{\|\mathbf{T}'\|}{\|\mathbf{r}'\|} = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^2 \cdot \|\mathbf{r}'\|} = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^3}$$

这个推导巧妙地利用了 \mathbf{r}'' 在 \mathbf{T} 方向 (切向加速度) 和 \mathbf{N} 方向 (法向加速度) 的分解。外积操作自动过滤掉了不改变方向的切向部分，只留下了反映“弯曲”的法向部分，从而直接提取出了曲率。

T, N, B 求解步骤：

第一步：切向量 \mathbf{T} (Tangent)

对曲线 $\mathbf{r}(t)$ 求导并单位化：

$$\mathbf{T} = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$$

第二步：副法向量 \mathbf{B} (Binormal)

在 3D 中，通常先求 \mathbf{B} 反而更容易。利用加速度 (二阶导) 和速度的外积：

$$\mathbf{B} = \frac{\mathbf{r}'(t) \times \mathbf{r}''(t)}{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|}$$

原理：速度和加速度构成的平面被称为密切平面， \mathbf{B} 就是该平面的法向量。

第三步：主法向量 \mathbf{N} (Normal)

有了 \mathbf{T} 和 \mathbf{B} ， \mathbf{N} 可以通过叉乘直接得出，确保右手系：

$$\mathbf{N} = \mathbf{B} \times \mathbf{T}$$

例 11.2.1. 抛物线 $y = x^2$ 的切向量 \mathbf{T} 、法向量 \mathbf{N} 和副法向量 \mathbf{B} 。

参数化方程为 $\mathbf{r}(x) = (x, x^2)$ 。

为了方便计算空间中的 B 向量，我们可以将其视为在 $z = 0$ 平面上的三维曲线，即 $\mathbf{r}(x) = (x, x^2, 0)$ 。

首先计算位置向量的一阶和二阶导数：

$$\mathbf{r}'(x) = (1, 2x, 0)$$

$$\mathbf{r}''(x) = (0, 2, 0)$$

一阶导的模长： $\|\mathbf{r}'(x)\| = \sqrt{1 + (2x)^2} = \sqrt{1 + 4x^2}$

切向量 \mathbf{T} (*Tangent*)

直接对一阶导进行单位化：

$$\mathbf{T} = \frac{\mathbf{r}'(x)}{\|\mathbf{r}'(x)\|} = \left(\frac{1}{\sqrt{1+4x^2}}, \frac{2x}{\sqrt{1+4x^2}}, 0 \right)$$

副法向量 \mathbf{B} (*Binormal*)

对于平面曲线，副法向量总是垂直于平面。我们通过叉乘计算：

$$\mathbf{r}' \times \mathbf{r}'' = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2x & 0 \\ 0 & 2 & 0 \end{vmatrix} = (0, 0, 2)$$

单位化后得到：

$$\mathbf{B} = (0, 0, 1)$$

(这说明抛物线始终在 xy 平面内弯曲)

主法向量 \mathbf{N} (*Normal*)

利用 $\mathbf{N} = \mathbf{B} \times \mathbf{T}$ (满足右手系且指向凹侧)：

$$\mathbf{N} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{1+4x^2}} & \frac{2x}{\sqrt{1+4x^2}} & 0 \end{vmatrix} = \left(-\frac{2x}{\sqrt{1+4x^2}}, \frac{1}{\sqrt{1+4x^2}}, 0 \right)$$

计算曲率 κ (*Curvature*)

曲率公式为：

$$\kappa = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^3}$$

代入已知量：

$$\|\mathbf{r}' \times \mathbf{r}''\| = 2$$

$$\|\mathbf{r}'\| = (1 + 4x^2)^{1/2}$$

得到：

$$\kappa(x) = \frac{2}{(1+4x^2)^{3/2}}$$

在顶点 $(0, 0)$ 处： $\kappa = 2$ 。此时曲率最大，密切圆半径 $R = 1/\kappa = 0.5$ 。

当 $x \rightarrow \infty$ 时： $\kappa \rightarrow 0$ 。这符合直觉，因为抛物线远端越来越趋于直线。

11.3 机器学习中的常用函数及其导函数

在机器学习中，激活函数（Activation Functions）和损失函数（Loss Functions）是模型学习的核心。了解它们的导函数（梯度）至关重要，因为反向传播算法正是利用导数来更新权重。

1. 激活函数 (Activation Functions)

激活函数为神经网络引入非线性，使其能够拟合复杂的函数。

Sigmoid 函数

Sigmoid 常用于二分类任务的输出层，将输入映射到 $(0, 1)$ 之间。

函数表达式： $f(x) = \frac{1}{1+e^{-x}}$

导函数： $f'(x) = f(x)(1-f(x))$

我们将使用导函数的商法则 (Quotient Rule) 或链式法则 (Chain Rule) 来完成推导

基础导数公式：

$$\frac{d}{dx}(e^u) = e^u \cdot \frac{du}{dx}$$

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

Sigmoid 函数定义：

$$f(x) = \frac{1}{1+e^{-x}} = (1+e^{-x})^{-1}$$

我们可以将 $f(x)$ 看作一个复合函数，设 $u = 1+e^{-x}$ ，则 $f(x) = u^{-1}$ 。

对外部函数求导：

$$\frac{df}{du} = -1 \cdot u^{-2} = -\frac{1}{(1+e^{-x})^2}$$

对内部函数 u 求导：

$$\frac{du}{dx} = \frac{d}{dx}(1+e^{-x}) = 0 + e^{-x} \cdot (-1) = -e^{-x}$$

应用链式法则

将上述两部分相乘：

$$f'(x) = \frac{df}{du} \cdot \frac{du}{dx} = \left(-\frac{1}{(1+e^{-x})^2}\right) \cdot (-e^{-x})$$

$$f'(x) = \frac{e^{-x}}{(1 + e^{-x})^2}$$

转化为自身表示：

$$\begin{aligned} f'(x) &= \frac{1}{1 + e^{-x}} \cdot \frac{e^{-x}}{1 + e^{-x}} \\ &= \frac{1}{1 + e^{-x}} \cdot \frac{(1 + e^{-x}) - 1}{1 + e^{-x}} \\ &= \frac{1}{1 + e^{-x}} \cdot \left(\frac{1 + e^{-x}}{1 + e^{-x}} - \frac{1}{1 + e^{-x}} \right) \\ &= \frac{1}{1 + e^{-x}} \cdot \left(1 - \frac{1}{1 + e^{-x}} \right) \end{aligned}$$

代回原函数：因为 $f(x) = \frac{1}{1+e^{-x}}$

$$f'(x) = f(x)(1 - f(x))$$

Sigmoid 函数的优缺点：

计算优势：在神经网络反向传播时，我们已经计算过了前向传播的激活值 $f(x)$ 。有了这个公式，求导就不再需要重新计算复杂的指数幂，只需一次减法和一次乘法，极大地提升了效率。

梯度消失：观察公式 $f(x)(1 - f(x))$ 。当 x 很大时， $f(x) \approx 1$ ，导数趋于 $1(1 - 1) = 0$ ；当 x 很小时， $f(x) \approx 0$ ，导数趋于 $0(1 - 0) = 0$ 。这就是为什么神经元在这些区域会“饱和”，导致权重无法更新。

ReLU 函数

ReLU 函数 (Rectified Linear Unit, 修正线性单元) 是目前深度学习中最受欢迎、应用最广泛的激活函数。它的结构虽然极其简单，但却解决了深度神经网络训练中的许多难题。

ReLU 函数本质上是一个分段线性函数，它会将所有负值归零，而保留所有正值。数学表达式：

$$f(x) = \max(0, x)$$

分段形式：

$$f(x) = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}$$

由于 ReLU 在 $x > 0$ 时是线性函数，在 $x < 0$ 时是常数函数，其导数非常容易计算。

导数表达式：

$$f'(x) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x < 0 \end{cases}$$

关于 $x = 0$ 点：在数学上，ReLU 在 0 点是不可导的（左导数为 0，右导数为 1）。但在机器学习的实际实现中，通常人为定义 $f'(0) = 0$ 或 1。

在 ReLU 出现之前，Sigmoid 是主流，但 ReLU 凭借以下优势改变了局面：

缓解梯度消失：在 $x > 0$ 的区域，ReLU 的导数恒为 1。这意味着在反向传播时，梯度可以无损地传递，不会像 Sigmoid 那样因为多次相乘而迅速趋近于 0。这使得训练深层神经网络变得更加容易。

计算速度极快：Sigmoid 涉及指数运算 (e^{-x})，计算量大；而 ReLU 只需要一个阈值判断 (if $x > 0$)，在大规模模型中能显著节省计算资源。

稀疏激活：在同一个批次的数据中，通常会有大量的神经元输出为 0。这种“稀疏性”使得模型更具代表性，且能起到一定的正则化作用，防止过拟合。

2. 均方误差 (MSE)

均方误差 (Mean Squared Error, MSE) 是回归任务中最常用的损失函数。它衡量的是模型预测值与真实值之间差异的平方的平均值。

假设我们有 n 个样本，对于每个样本 i ，真实值为 y_i ，模型预测值为 \hat{y}_i 。

单个样本的误差： $e_i = (y_i - \hat{y}_i)^2$

全样本的 MSE 表达式：

$$L = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

注意：在机器学习推导（如梯度下降）中，为了求导后能抵消掉平方项产生的系数 2，通常会在公式前加上 $\frac{1}{2}$ ，写成：

$$L = \frac{1}{2n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

导函数推导

我们通常需要对预测值 \hat{y}_i 求导，以确定如何调整模型参数来减小损失。

步骤 1：应用链式法则设 $u_i = y_i - \hat{y}_i$ ，则 $L = \frac{1}{2n} \sum u_i^2$ 。根据链式法则：

$$\frac{\partial L}{\partial \hat{y}_i} = \frac{\partial L}{\partial u_i} \cdot \frac{\partial u_i}{\partial \hat{y}_i}$$

步骤 2：分别求导

对外部平方项求导： $\frac{\partial L}{\partial u_i} = \frac{1}{2n} \cdot 2u_i = \frac{1}{n}u_i$

对内部差值项求导： $\frac{\partial u_i}{\partial \hat{y}_i} = \frac{\partial(y_i - \hat{y}_i)}{\partial \hat{y}_i} = -1$

步骤 3：组合结果

$$\frac{\partial L}{\partial \hat{y}_i} = \frac{1}{n}(y_i - \hat{y}_i) \cdot (-1) = \frac{1}{n}(\hat{y}_i - y_i)$$

对于第 i 个样本，MSE 对预测值的梯度就是预测值与真实值之差（缩放 $1/n$ 倍）。

MSE 的特性

惩罚离群点：由于存在平方项，当预测值与真实值差距较大时，误差会呈二次方增长。这意味着 MSE 对异常值（Outliers）非常敏感。

凸性 (Convexity)：MSE 是一个凸函数。在线性回归中，这意味着它只有一个全局最小值，使用梯度下降法很容易找到最优解。

物理意义：MSE 对应于高斯分布下的最大似然估计 (MLE)。

3. Softmax 函数

Softmax 函数是机器学习中处理多分类问题的核心函数。它通常位于神经网络的最后一层，负责将原始的得分 (Logits) 转换为概率分布。

函数定义

假设我们有一个向量 \mathbf{z} ，包含 K 个类别的原始得分 z_1, z_2, \dots, z_K 。Softmax 函数将其中第 i 个分量映射为一个 $(0, 1)$ 之间的概率值 a_i 。

数学表达式：

$$a_i = \text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

核心特性：

归一化：所有输出概率之和恒等于 1，即 $\sum_{i=1}^K a_i = 1$ 。

放大差异：通过指数函数 e^z ，Softmax 会“拉开”得分之间的差距，使得得分最高的类别在概率分布中占据主导地位。

导函数推导

Softmax 的求导比 Sigmoid 复杂一些，因为它是一个“多对多”的函数——输出向量中的每一个元素 a_i 都依赖于输入向量中的每一个元素 z_j 。

我们需要计算偏导数 $\frac{\partial a_i}{\partial z_j}$ 。根据 i 是否等于 j ，分为两种情况：

准备工作：商法则

回顾导数商法则： $\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$ 。在 Softmax 中：

$$u = e^{z_i}$$

$$v = \sum_{k=1}^K e^{z_k}$$

情况 1： $i = j$ (对自身得分求导)

$$\text{此时 } u' = \frac{\partial e^{z_i}}{\partial z_i} = e^{z_i}, \quad v' = \frac{\partial (\sum e^{z_k})}{\partial z_i} = e^{z_i}.$$

$$\frac{\partial a_i}{\partial z_i} = \frac{e^{z_i} \cdot \sum -e^{z_i} \cdot e^{z_i}}{(\sum)^2}$$

$$\frac{\partial a_i}{\partial z_i} = \frac{e^{z_i}}{\sum} \cdot \frac{\sum -e^{z_i}}{\sum} = a_i(1 - a_i)$$

情况 2: $i \neq j$ (对其他得分求导)

此时 $u = e^{z_i}$ 对 z_j 求导为 0 (因为 $i \neq j$), $v' = \frac{\partial(\sum e^{z_k})}{\partial z_j} = e^{z_j}$ 。

$$\frac{\partial a_i}{\partial z_j} = \frac{0 \cdot \sum -e^{z_i} \cdot e^{z_j}}{(\sum)^2}$$

$$\frac{\partial a_i}{\partial z_j} = -\frac{e^{z_i}}{\sum} \cdot \frac{e^{z_j}}{\sum} = -a_i a_j$$

综合结论:

$$\frac{\partial a_i}{\partial z_j} = \begin{cases} a_i(1 - a_j), & \text{if } i = j \\ -a_i a_j, & \text{if } i \neq j \end{cases}$$

与交叉熵结合

在深度学习框架中, Softmax 通常与交叉熵损失 (Cross-Entropy Loss) 配合使用。

有趣的是, 当它们结合在一起时, 最终的梯度形式会变得异常简洁。

设 y_i 为真实标签 (One-hot 编码), a_i 为 Softmax 输出, 总损失为 $L = -\sum y_i \ln a_i$ 。经过链式法则推导, 损失 L 对输入 z_j 的梯度简单到令人惊讶:

$$\frac{\partial L}{\partial z_j} = a_j - y_j$$

直观理解: 梯度就是“预测概率与真实标签的差值”。如果预测概率比真实值大, 梯度为正, 减小该权重; 反之亦然。这种完美的抵消特性正是 Softmax + 交叉熵成为分类任务标配的原因。

在计算 Softmax 时, 如果 z_i 非常大, e^{z_i} 会导致数值溢出。在计算前减去向量中的最大值, 即 $e^{z_i - \max(z)}$ 。这不会改变输出结果, 但能保证计算安全。

4. 交叉熵损失 (Cross-Entropy Loss)

交叉熵损失是分类问题中最核心的损失函数。它源于信息论, 用于衡量两个概率分布之间的“距离”。

在机器学习中, 我们希望模型输出的概率分布 P 尽可能接近数据的真实概率分布 Q 。

函数定义

二分类 (Binary Cross-Entropy, BCE)

当我们只有两个类别 (如: 猫或狗) 时, 模型输出一个值 $\hat{y} \in (0, 1)$ 。

数学表达式:

$$L = -[y \ln(\hat{y}) + (1 - y) \ln(1 - \hat{y})]$$

其中 y 为真实标签 (0 或 1), \hat{y} 为模型预测概率。

多分类 (Categorical Cross-Entropy)

当有 K 个类别时, 使用 Softmax 输出概率分布 $\hat{\mathbf{y}}$ 。

数学表达式:

$$L = - \sum_{i=1}^K y_i \ln(\hat{y}_i)$$

其中 y_i 是真实标签的 One-hot 编码 (正确类为 1, 其余为 0)。

二分类导函数推导

在神经网络中, \hat{y} 通常由 Sigmoid 函数产生, 即 $\hat{y} = \sigma(z) = \frac{1}{1+e^{-z}}$ 。我们要求对 z 的梯度。

步骤 1: 利用链式法则

$$\frac{\partial L}{\partial z} = \frac{\partial L}{\partial \hat{y}} \cdot \frac{\partial \hat{y}}{\partial z}$$

步骤 2: 对损失函数求导 $\frac{\partial L}{\partial \hat{y}}$

对 $L = -y \ln \hat{y} - (1-y) \ln(1-\hat{y})$ 求偏导:

$$\frac{\partial L}{\partial \hat{y}} = -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}} = \frac{\hat{y}-y}{\hat{y}(1-\hat{y})}$$

步骤 3: 结合 Sigmoid 的导数

前面推导过 $\frac{\partial \hat{y}}{\partial z} = \hat{y}(1-\hat{y})$ 。将两者相乘:

$$\begin{aligned}\frac{\partial L}{\partial z} &= \frac{\hat{y}-y}{\hat{y}(1-\hat{y})} \cdot \hat{y}(1-\hat{y}) \\ \frac{\partial L}{\partial z} &= \hat{y} - y\end{aligned}$$

梯度同样非常简洁, 即预测概率与真实标签的差值。

多分类导函数推导

在多分类任务中, Softmax 激活函数与 ** 交叉熵损失 (Cross-Entropy Loss) ** 的结合是标准配置。它们的组合推导过程展示了数学上的巧妙抵消。

Softmax 输出: $a_i = \frac{e^{z_i}}{\sum_{k=1}^K e^{z_k}}$

多分类交叉熵损失: $L = - \sum_{i=1}^K y_i \ln a_i$

其中 y_i 是标签的 One-hot 编码。对于一个样本, 只有一个 $y_{true} = 1$, 其余均为 0。

Softmax 自身的偏导 (前文已证):

$$\frac{\partial a_i}{\partial z_j} = \begin{cases} a_i(1-a_i) & i = j \\ -a_i a_j & i \neq j \end{cases}$$

根据链式法则，由于 L 是关于所有 a_i 的函数，而每个 a_i 又是关于 z_j 的函数：

$$\frac{\partial L}{\partial z_j} = \sum_{i=1}^K \frac{\partial L}{\partial a_i} \cdot \frac{\partial a_i}{\partial z_j}$$

对 $L = -\sum y_i \ln a_i$ 求导：

$$\frac{\partial L}{\partial a_i} = -\frac{y_i}{a_i}$$

由于 $\frac{\partial a_i}{\partial z_j}$ 在 $i = j$ 和 $i \neq j$ 时表现不同，我们将求和式拆分为两部分：

$$\frac{\partial L}{\partial z_j} = \left(\frac{\partial L}{\partial a_j} \cdot \frac{\partial a_j}{\partial z_j} \right) + \sum_{i \neq j}^K \left(\frac{\partial L}{\partial a_i} \cdot \frac{\partial a_i}{\partial z_j} \right)$$

代入具体导数

$$\frac{\partial L}{\partial z_j} = \left(-\frac{y_j}{a_j} \cdot a_j(1 - a_j) \right) + \sum_{i \neq j}^K \left(-\frac{y_i}{a_i} \cdot (-a_i a_j) \right)$$

左边项简化： $-y_j(1 - a_j) = -y_j + y_j a_j$

右边项简化： $\sum_{i \neq j}^K (y_i a_j) = a_j \sum_{i \neq j}^K y_i$

合并与最终简化

$$\frac{\partial L}{\partial z_j} = -y_j + y_j a_j + a_j \sum_{i \neq j}^K y_i$$

$$\frac{\partial L}{\partial z_j} = -y_j + a_j \left(y_j + \sum_{i \neq j}^K y_i \right)$$

因为 y 是 One-hot 编码，所有概率之和 $\sum_{i=1}^K y_i = 1$ ，所以括号内的部分等于 1：

$$\frac{\partial L}{\partial z_j} = a_j - y_j$$

这个结果非常优雅：梯度 = 预测值 - 真实值。

如果预测很准：比如类别 j 是正确的 ($y_j = 1$)，模型预测 $a_j = 0.99$ ，那么梯度 $0.99 - 1 = -0.01$ ，权重只会有极小的修正。

如果预测很差：比如类别 j 是正确的 ($y_j = 1$)，但模型预测 $a_j = 0.1$ ，那么梯度 $0.1 - 1 = -0.9$ ，巨大的梯度会强制模型大幅度调整权重。

在编写深度学习框架（如 PyTorch 的 CrossEntropyLoss）时，通常会将 Softmax 和 CrossEntropy 合并计算。这不仅是为了求导公式的简洁，更是为了数值稳定性。直接计算 $a_j - y_j$ 避免了中间步骤中可能出现的 $\log(0)$ 或极大指数幂溢出的问题。

11.4 通量和散度

通量的物理直觉是：衡量矢量场 \mathbf{F} 在单位时间内穿过某个微小表面的“净流量”。在直角坐标系下，我们推导矢量场 $\mathbf{F} = (F_1, F_2, F_3)$ 流出一个微元体（体积 $dV = dx dy dz$ ）的净通量。

1. 建立微元模型

想象一个中心位于 (x, y, z) 的微小长方体，六个面分别垂直于坐标轴。我们以穿过垂直于 x 轴的两个表面的通量为例进行推导。

右侧面（位于 $x + \frac{dx}{2}$ ）：法向量指向 $+x$ 方向。该面中心处的场分量 F_1 约为：

$$F_1(x + \frac{dx}{2}, y, z) \approx F_1 + \frac{\partial F_1}{\partial x} \frac{dx}{2}$$

穿出通量： $\Phi_{right} = (F_1 + \frac{\partial F_1}{\partial x} \frac{dx}{2}) \cdot dy dz$

左侧面（位于 $x - \frac{dx}{2}$ ）：法向量指向 $-x$ 方向（流出体积的方向）。该面中心处的场分量 F_1 约为：

$$F_1(x - \frac{dx}{2}, y, z) \approx F_1 - \frac{\partial F_1}{\partial x} \frac{dx}{2}$$

左侧面（位于 $x - \frac{dx}{2}$ ）：法向量指向 $-x$ 方向（流出体积的方向）。该面中心处的场分量 F_1 约为：

$$F_1(x - \frac{dx}{2}, y, z) \approx F_1 - \frac{\partial F_1}{\partial x} \frac{dx}{2}$$

穿出通量（由于法向量向左，需乘面积向量 $-dy dz \mathbf{i}$ ）： $\Phi_{left} = -(F_1 - \frac{\partial F_1}{\partial x} \frac{dx}{2}) \cdot dy dz$
 x 方向的净通量（两面之和）：

$$d\Phi_x = \Phi_{right} + \Phi_{left} = \left(\frac{\partial F_1}{\partial x} dx \right) dy dz = \frac{\partial F_1}{\partial x} dV$$

3. 汇总三个方向

同理，我们可以得到穿过另外两对面的净通量：

y 方向净通量： $d\Phi_y = \frac{\partial F_2}{\partial y} dy (dx dz) = \frac{\partial F_2}{\partial y} dV$

z 方向净通量： $d\Phi_z = \frac{\partial F_3}{\partial z} dz (dx dy) = \frac{\partial F_3}{\partial z} dV$

总净通量（流出该微元体的总和）：

$$d\Phi = d\Phi_x + d\Phi_y + d\Phi_z = \left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \right) dx dy dz$$

4. 结论：从通量到散度

通过上面的推导，我们发现了一个极其重要的物理量：

特征量：括号内的 $\left(\frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} \right)$ 描述了该点处通量的源性。

散度的诞生：如果我们定义单位体积的净通量为该点的“散度”，那么：

$$\operatorname{div} \mathbf{F} = \frac{d\Phi}{dV} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

算子表示：这恰好对应 ∇ 算子与 \mathbf{F} 的点乘： $\nabla \cdot \mathbf{F}$ 。

11.5 环量和旋度

1. 物理直觉：环量密度

如果一个场在某处有旋转趋势，那么你沿着一个极小的闭合回路 L 走一圈，矢量场沿路径的累积效应（环量） $\oint_L \mathbf{F} \cdot d\mathbf{l}$ 就不应该为零。

我们要寻找一个量，用来描述 \mathbf{F} 在某点附近的旋转强度。这个量不应该依赖于我们取的回路大小，因此我们要计算的是单位面积的环量极限。

1. 考察 xy 平面内的旋转（绕 z 轴）

我们在 xy 平面上取一个微小矩形，中心为 (x, y, z) ，边长为 dx, dy 。我们计算沿边界逆时针走一圈的环量 $\Gamma_z = \oint \mathbf{F} \cdot d\mathbf{l}$ 。

底边 (1): 路径 dx ，位于 $y - \frac{dy}{2}$ 。 F_1 的贡献：

$$\left(F_1(x, y - \frac{dy}{2}, z) \right) dx \approx (F_1 - \frac{\partial F_1}{\partial y} \frac{dy}{2}) dx$$

顶边 (2): 路径 $-dx$ （向左），位于 $y + \frac{dy}{2}$ 。 F_1 的贡献：

$$-\left(F_1(x, y + \frac{dy}{2}, z) \right) dx \approx -(F_1 + \frac{\partial F_1}{\partial y} \frac{dy}{2}) dx$$

右边 (3): 路径 dy （向上），位于 $x + \frac{dx}{2}$ 。 F_2 的贡献：

$$\left(F_2(x + \frac{dx}{2}, y, z) \right) dy \approx (F_2 + \frac{\partial F_2}{\partial x} \frac{dx}{2}) dy$$

左边 (4): 路径 $-dy$ （向下），位于 $x - \frac{dx}{2}$ 。 F_2 的贡献：

$$-\left(F_2(x - \frac{dx}{2}, y, z) \right) dy \approx -(F_2 - \frac{\partial F_2}{\partial x} \frac{dx}{2}) dy$$

累加总环量 Γ_z ：

$$\Gamma_z = \underbrace{\left[(F_2 + \frac{\partial F_2}{\partial x} \frac{dx}{2}) - (F_2 - \frac{\partial F_2}{\partial x} \frac{dx}{2}) \right]}_{\text{垂直边贡献}} dy + \underbrace{\left[(F_1 - \frac{\partial F_1}{\partial y} \frac{dy}{2}) - (F_1 + \frac{\partial F_1}{\partial y} \frac{dy}{2}) \right]}_{\text{水平边贡献}} dx$$

简化后得到：

$$\Gamma_z = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy$$

由此定义绕 z 轴的旋转强度（单位面积环量）：

$$\omega_z = \frac{\Gamma_z}{dxdy} = \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}$$

2. 推广到其他两个平面

利用同样的逻辑（轮换对称性），我们可以求出在其他两个正交平面上的旋转强度：

在 yz 平面内（绕 x 轴）：考察 F_2 和 F_3 随 y 和 z 的变化：

$$\omega_x = \frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}$$

在 zx 平面内（绕 y 轴）：考察 F_3 和 F_1 随 z 和 x 的变化：

$$\omega_y = \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}$$

3. 合成旋度矢量

至此，我们发现这三个量 $(\omega_x, \omega_y, \omega_z)$ 完整地描述了矢量场在这一点三个维度的旋转特性。我们将这三个分量组合成一个新的矢量，并命名为 $\text{curl } \mathbf{F}$ ：

$$\text{curl } \mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) \mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) \mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) \mathbf{k}$$

4. 符号化的最终形式

为了方便记忆这种复杂的偏微分组合，我们引入算子 $2\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$ 。观察 3 发现，上述结果恰好等于 ∇ 与 \mathbf{F} 的叉乘结果：

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix}$$

总结比较

概念微元基础物理意义最终形成的算子

环量微元

闭合线积分 $\oint \mathbf{F} \cdot d\mathbf{l}$ 描述局部的旋转强度 $\nabla \times \mathbf{F}$

通量微元

闭合面积分 $\oint \mathbf{F} \cdot d\mathbf{S}$ 描述局部的发散/汇聚散度 $\nabla \cdot \mathbf{F}$

11.6 格林定理

格林定理（Green's Theorem）是向量分析中一个非常优美的结论，它将闭合曲线上的线积分与该曲线所围区域上的面积分联系了起来。

实际上，格林定理可以看作是旋度微元原理在二维平面上的宏观累加。

1. 物理直觉：小漩涡合成大漩涡

想象一个平面区域 D ，被一条闭合曲线 C 包围。我们将区域 D 划分为无数个极其微小的矩形元（即我们之前推导过的“环量微元”）。

内部抵消：对于两个相邻的微元，它们相交的边会被经过两次，但方向相反。因此，沿这些内部公共边的线积分会全部相互抵消。

边界残留：唯一没有被抵消掉的，只有最外层紧贴曲线 C 的那些微元边。

结论：所有微元环量的总和，等于绕最外圈大边界的环量。

2. 数学推导过程

设平面矢量场为 $\mathbf{F} = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ 。我们要证明：

$$\oint_C (Pdx + Qdy) = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA$$

第一步：处理 Q 分量的部分

我们先考察面积分中的 $\iint_D \frac{\partial Q}{\partial x} dA$ 。假设区域 D 是一个简单的 x -型区域（左边界为 $x = g_1(y)$ ，右边界为 $x = g_2(y)$ ）：

$$\iint_D \frac{\partial Q}{\partial x} dxdy = \int_c^d \left[\int_{g_1(y)}^{g_2(y)} \frac{\partial Q}{\partial x} dx \right] dy$$

据微积分基本定理，内部积分结果为：

$$\int_{g_1(y)}^{g_2(y)} \frac{\partial Q}{\partial x} dx = Q(g_2(y), y) - Q(g_1(y), y)$$

代回原式：

$$\iint_D \frac{\partial Q}{\partial x} dA = \int_c^d Q(g_2(y), y) dy - \int_c^d Q(g_1(y), y) dy$$

这两个积分正好对应曲线 C 的右半部分和左半部分对 dy 的积分。合并后即为：

$$\iint_D \frac{\partial Q}{\partial x} dA = \oint_C Qdy$$

第二步：处理 P 分量的部分

同理，假设区域 D 是一个 y -型区域（下边界 $y = f_1(x)$ ，上边界 $y = f_2(x)$ ），计算 $\iint_D \frac{\partial P}{\partial y} dA$ ：

$$\iint_D \frac{\partial P}{\partial y} dydx = \int_a^b [P(x, f_2(x)) - P(x, f_1(x))] dx$$

注意，当我们沿着曲线 C 逆时针走时，上边界是从右向左走的（ dx 为负），下边界是从左向右走的（ dx 为正）。因此：

$$\oint_C P dx = \int_a^b P(x, f_1(x)) dx + \int_b^a P(x, f_2(x)) dx = - \iint_D \frac{\partial P}{\partial y} dA$$

第三步：组合结果

将上述两部分相加：

$$\oint_C P dx + \oint_C Q dy = \iint_D \frac{\partial Q}{\partial x} dA - \iint_D \frac{\partial P}{\partial y} dA$$

合并后得到格林定理的标准形式：

$$\oint_C (P dx + Q dy) = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$$

右边的被积函数： $\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)$ 。这正是我们之前推导的旋度在 z 方向的分量（平面旋转强度）吗。

所以格林定理的本质就是：“区域内所有微小旋转的总和 = 边界上的总环流”。

第十二章 向量微分

12.1 微分向量分析

16. 向量的微分。让我们考虑如下向量场：

$$\mathbf{u} = \alpha(x, y, z, t)\mathbf{i} + \beta(x, y, z, t)\mathbf{j} + \gamma(x, y, z, t)\mathbf{k} \quad (42)$$

在任意点 $P(x, y, z)$ 及任意时间 t , 式 (42) 定义了一个向量。如果我们固定点 P , 由于分量 α, β, γ 具有时间相关性, 向量 \mathbf{u} 仍会发生变化。如果我们固定时间, 我们会注意到在点 $P(x, y, z)$ 处的向量通常与在以下点处的向量不同:

$$Q(x + dx, y + dy, z + dz)$$

在微积分中, 学生已经学过如何计算 x, y, z, t 的单函数变化。那么在向量的情况下, 我们会遇到什么困难吗? 实际上完全没有, 因为我们很容易注意到, 当且仅当向量的分量发生变化时, \mathbf{u} 才会发生变化。因此, $\alpha(x, y, z, t)$ 的变化会产生 \mathbf{u} 在 x 方向上的变化, 同理, β 和 γ 的变化分别产生 \mathbf{u} 在 y 和 z 方向上的变化。由此, 我们得出以下定义:

$$d\mathbf{u} = d\alpha \mathbf{i} + d\beta \mathbf{j} + d\gamma \mathbf{k} \quad (43)$$

$$\begin{aligned} d\mathbf{u} &= \left(\frac{\partial \alpha}{\partial x} dx + \frac{\partial \alpha}{\partial y} dy + \frac{\partial \alpha}{\partial z} dz + \frac{\partial \alpha}{\partial t} dt \right) \mathbf{i} \\ &\quad + \left(\frac{\partial \beta}{\partial x} dx + \frac{\partial \beta}{\partial y} dy + \frac{\partial \beta}{\partial z} dz + \frac{\partial \beta}{\partial t} dt \right) \mathbf{j} \\ &\quad + \left(\frac{\partial \gamma}{\partial x} dx + \frac{\partial \gamma}{\partial y} dy + \frac{\partial \gamma}{\partial z} dz + \frac{\partial \gamma}{\partial t} dt \right) \mathbf{k} \end{aligned}$$

例如, 设 $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ 为三维空间中运动质点 $P(x, y, z)$ 的位置向量。则:

$$d\mathbf{r} = dx \mathbf{i} + dy \mathbf{j} + dz \mathbf{k}$$

以及

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k} \quad (44)$$

$$\mathbf{a} = \frac{d^2\mathbf{r}}{dt^2} = \frac{d^2x}{dt^2}\mathbf{i} + \frac{d^2y}{dt^2}\mathbf{j} + \frac{d^2z}{dt^2}\mathbf{k} \quad (45)$$

根据定义，方程 (44) 和 (45) 分别是质点的速度和加速度。我们假设向量 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 在空间中保持固定。

如果向量 \mathbf{u} 仅取决于单个变量 t ，我们可以定义：

$$\frac{d\mathbf{u}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{u}(t + \Delta t) - \mathbf{u}(t)}{\Delta t} \quad (46)$$

很容易验证 (46) 与 (43) 是等价的。

例 14。考虑一个质点 P 在半径为 r 的圆上以恒定角速度 $\omega = \frac{d\theta}{dt}$ 运动（见图 28）。我们注意到：

$$\mathbf{r} = r \cos \theta \mathbf{i} + r \sin \theta \mathbf{j}$$

因此：

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = (-r \sin \theta \mathbf{i} + r \cos \theta \mathbf{j}) \frac{d\theta}{dt}$$

以及：

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = (-r \cos \theta \mathbf{i} - r \sin \theta \mathbf{j}) \left(\frac{d\theta}{dt} \right)^2$$

所以加速度为：

$$\mathbf{a} = -\omega^2 \mathbf{r} \quad (47)$$

点 P 具有指向原点的加速度，其大小为恒定的 $\omega^2 r$ 。这个加速度是由于速度向量以恒定速率改变方向而产生的；它被称为向心加速度。

例 15。设 P 为空间曲线上的任意一点（见图 29）：

$$x = x(s)$$

$$y = y(s)$$

$$z = z(s)$$

其中 s 是从某个固定点 Q 开始测量的弧长。那么：

$$\mathbf{r} = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k} \quad (48)$$

因此：

$$\frac{d\mathbf{r}}{ds} = \frac{dx}{ds}\mathbf{i} + \frac{dy}{ds}\mathbf{j} + \frac{dz}{ds}\mathbf{k} \quad (49)$$

并且由微积分可知：

$$\begin{aligned}\frac{d\mathbf{r}}{ds} \cdot \frac{d\mathbf{r}}{ds} &= \left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 \\ &= \frac{dx^2 + dy^2 + dz^2}{ds^2} \equiv 1\end{aligned}$$

因此 $\frac{d\mathbf{r}}{ds}$ 是一个单位向量。当 $\Delta s \rightarrow 0$ 时， $\frac{\Delta\mathbf{r}}{\Delta s}$ 的位置趋近于点 P 处的切线。因此，(49) 代表了空间曲线 (48) 的单位切向量。

17. 微分规则。考虑：

$$\begin{aligned}\varphi(t) &= \mathbf{u}(t) \cdot \mathbf{v}(t) \\ \varphi(t + \Delta t) - \varphi(t) &= \mathbf{u}(t + \Delta t) \cdot \mathbf{v}(t + \Delta t) - \mathbf{u}(t) \cdot \mathbf{v}(t)\end{aligned}$$

现在：

$$\begin{aligned}\mathbf{u}(t + \Delta t) &= \mathbf{u}(t) + \Delta\mathbf{u} \\ \mathbf{v}(t + \Delta t) &= \mathbf{v}(t) + \Delta\mathbf{v}\end{aligned}$$

(见图 27)，所以：

$$\frac{\varphi(t + \Delta t) - \varphi(t)}{\Delta t} = \mathbf{u} \cdot \frac{\Delta\mathbf{v}}{\Delta t} + \frac{\Delta\mathbf{u}}{\Delta t} \cdot \mathbf{v} + \frac{\Delta\mathbf{u}}{\Delta t} \cdot \Delta\mathbf{v}$$

取极限，我们得到：

$$\frac{d(\mathbf{u} \cdot \mathbf{v})}{dt} = \mathbf{u} \cdot \frac{d\mathbf{v}}{dt} + \frac{d\mathbf{u}}{dt} \cdot \mathbf{v} \quad (50)$$

类似地：

$$\frac{d(\mathbf{u} \times \mathbf{v})}{dt} = \mathbf{u} \times \frac{d\mathbf{v}}{dt} + \frac{d\mathbf{u}}{dt} \times \mathbf{v} \quad (51)$$

$$\frac{d(f\mathbf{u})}{dt} = f \frac{d\mathbf{u}}{dt} + \frac{df}{dt} \mathbf{u} \quad (52)$$

请注意这些公式与微积分规则的一致性。

例 16。设 $\mathbf{u}(t)$ 为一个模（大小）恒定的向量。因此：

$$\mathbf{u} \cdot \mathbf{u} = u^2 = \text{常数}$$

通过微分，我们得到：

$$\begin{aligned}\mathbf{u} \cdot \frac{d\mathbf{u}}{dt} + \frac{d\mathbf{u}}{dt} \cdot \mathbf{u} &= 0 \\ \mathbf{u} \cdot \frac{d\mathbf{u}}{dt} &= 0\end{aligned}$$

因此，要么 $\frac{d\mathbf{u}}{dt} = 0$ ，要么 $\frac{d\mathbf{u}}{dt}$ 与 \mathbf{u} 垂直。这是一个重要的结果，学生应当充分理解。读者应给出该定理的几何证明。

例 17。在所有情况下， $\mathbf{u} \cdot \mathbf{u} = u^2$ ，其中 u 是 \mathbf{u} 的长度。微分得：

$$2\mathbf{u} \cdot \frac{d\mathbf{u}}{dt} = 2u \frac{du}{dt}$$

以及：

$$\mathbf{u} \cdot \frac{d\mathbf{u}}{dt} = u \frac{du}{dt} \quad (53)$$

这一结果并非显而易见，因为 $|d\mathbf{u}| \neq du$ 。

例 18：平面运动。现在 $\mathbf{r} = r\mathbf{R}$ ，其中 \mathbf{R} 是一个单位向量（见图 30）。因此：

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{dr}{dt}\mathbf{R} + r \frac{d\mathbf{R}}{dt}$$

由于 \mathbf{R} 是单位向量，根据例 16， $\frac{d\mathbf{R}}{dt}$ 与 \mathbf{R} 垂直。同时，通过对 $\mathbf{R} = \cos \theta \mathbf{i} + \sin \theta \mathbf{j}$ 进行微分，我们可以很容易验证 $|\frac{d\mathbf{R}}{dt}| = \frac{d\theta}{dt}$ 。因此 $\mathbf{v} = \frac{dr}{dt}\mathbf{R} + r \frac{d\theta}{dt} \mathbf{P}$ ，其中 \mathbf{P} 是垂直于 \mathbf{R} 的单位向量。再次微分得：

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2r}{dt^2}\mathbf{R} + \frac{dr}{dt} \frac{d\mathbf{R}}{dt} + \frac{dr}{dt} \frac{d\theta}{dt} \mathbf{P} + r \frac{d^2\theta}{dt^2} \mathbf{P} + r \frac{d\theta}{dt} \frac{d\mathbf{P}}{dt}$$

或：

$$\mathbf{a} = \left[\frac{d^2r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right] \mathbf{R} + \left[2 \frac{dr}{dt} \frac{d\theta}{dt} + r \frac{d^2\theta}{dt^2} \right] \mathbf{P}$$

因为：

$$\frac{d\mathbf{P}}{dt} = - \frac{d\theta}{dt} \mathbf{R} \quad (54)$$

因此：

$$\mathbf{a} = \left[\frac{d^2r}{dt^2} - r \left(\frac{d\theta}{dt} \right)^2 \right] \mathbf{R} + \frac{1}{r} \frac{d}{dt} \left(r^2 \frac{d\theta}{dt} \right) \mathbf{P} \quad (55)$$

18. 梯度

设 $\varphi(x, y, z)$ 为任意连续可微的空间函数。根据微积分：

$$d\varphi = \frac{\partial \varphi}{\partial x} dx + \frac{\partial \varphi}{\partial y} dy + \frac{\partial \varphi}{\partial z} dz \quad (56)$$

现在设 \mathbf{r} 为指向点 $P(x, y, z)$ 的位置向量：

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

如果我们移动到点 $Q(x+dx, y+dy, z+dz)$ （见图 32），则：

$$d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$$

现在注意到式 (56) 包含了项 dx, dy, dz 以及项 $\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z}$ 。我们定义一个由 φ 的三个偏导数组成的新向量。令 $\operatorname{del} \varphi \equiv \nabla \varphi \equiv \text{gradient } \varphi$ (φ 的梯度) 定义为：

$$\nabla\varphi = \frac{\partial\varphi}{\partial x}\mathbf{i} + \frac{\partial\varphi}{\partial y}\mathbf{j} + \frac{\partial\varphi}{\partial z}\mathbf{k} \quad (57)$$

我们立即可以看到：

$$d\varphi = d\mathbf{r} \cdot \nabla\varphi \quad (58)$$

梯度的几何解释

我们现在给出 $\nabla\varphi$ 的几何解释。在点 $P(x_0, y_0, z_0)$ 处， φ 的值为 $\varphi(x_0, y_0, z_0)$ ，因此：

$$\varphi(x, y, z) = \varphi(x_0, y_0, z_0)$$

表示一个显然包含点 $P(x_0, y_0, z_0)$ 的曲面（等值面）。

只要我们沿着这个曲面移动， φ 就具有恒定值 $\varphi(x_0, y_0, z_0)$ ，且 $d\varphi = 0$ 。因此，根据式 (58)：

$$d\mathbf{r} \cdot \nabla\varphi = 0 \quad (59)$$

由于 $\nabla\varphi$ 是一个在 φ 被求导后就完全确定的向量，而式 (59) 表明，只要 $d\mathbf{r}$ 表示从 P 到 Q 的位移且 Q 仍在 $\varphi = \text{常数}$ 的曲面上， $\nabla\varphi$ 就与 $d\mathbf{r}$ 垂直。

因此， $\nabla\varphi$ 垂直于曲面在 P 点处所有可能的切线，所以 $\nabla\varphi$ 必然正交（垂直）于曲面 $\varphi(x, y, z) = \text{常数}$ 。

(接上页图 33)。现在让我们回到公式 $d\varphi = d\mathbf{r} \cdot \nabla\varphi$ 。向量 $\nabla\varphi$ 在任何给定点 $P(x, y, z)$ 都是固定的，因此 $d\varphi$ (φ 的变化量) 在很大程度上取决于 $d\mathbf{r}$ 。

显然，当 $d\mathbf{r}$ 与 $\nabla\varphi$ 平行时， $d\varphi$ 取得最大值，因为 $d\mathbf{r} \cdot \nabla\varphi = |d\mathbf{r}| |\nabla\varphi| \cos\theta$ ，而 $\cos\theta$ 在 $\theta = 0^\circ$ 时达到最大值。因此， $\nabla\varphi$ 的方向是函数 $\varphi(x, y, z)$ 增加最快的方向。

设 $|d\mathbf{r}| = ds$ ，则有：

$$\frac{d\varphi}{ds} = \mathbf{u} \cdot \nabla\varphi \quad (60)$$

其中 \mathbf{u} 是沿 $d\mathbf{r}$ 方向的单位向量。因此， φ 在任何方向上的变化率，就是 $\nabla\varphi$ 在该方向单位向量上的投影。

示例解析例

19：求曲面 $x^2 + y^2 - z = 1$ 在点 $P(1, 1, 1)$ 处的单位法向量。

这里， $\varphi(x, y, z) = x^2 + y^2 - z$ 。

$$\nabla\varphi = 2x\mathbf{i} + 2y\mathbf{j} - \mathbf{k}.$$

在点 $P(1, 1, 1)$ 处， $\nabla\varphi = 2\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ 。

因此，单位法向量为：

$$\mathbf{N} = \frac{2\mathbf{i} + 2\mathbf{j} - \mathbf{k}}{3}$$

(注: 分母 3 是向量长度 $\sqrt{2^2 + 2^2 + (-1)^2}$)

例 20: 若 $r = (x^2 + y^2 + z^2)^{\frac{1}{2}}$, 求 ∇r 。

曲面 $r = \text{常数}$ 是一个球面。

因此 ∇r 与球面正交, 也就是说它与位置向量 \mathbf{r} 平行。

故 $\nabla r = k\mathbf{r}$ 。根据公式 (53) 有 $dr = d\mathbf{r} \cdot \nabla r = k d\mathbf{r} \cdot \mathbf{r} = kr dr$ 。

因此 $k = \frac{1}{r}$, 且:

$$\nabla r = \frac{\mathbf{r}}{r} = \mathbf{R}$$

(此处 \mathbf{R} 表示径向单位向量)

例 21 (链式法则证明):

$$\begin{aligned}\nabla f(u) &= \frac{\partial f}{\partial x} \mathbf{i} + \frac{\partial f}{\partial y} \mathbf{j} + \frac{\partial f}{\partial z} \mathbf{k} \\ &= f'(u) \frac{\partial u}{\partial x} \mathbf{i} + f'(u) \frac{\partial u}{\partial y} \mathbf{j} + f'(u) \frac{\partial u}{\partial z} \mathbf{k} = f'(u) \nabla u\end{aligned}$$

22 (多元复合函数):

$$\nabla f(u_1, u_2, \dots, u_n) = \sum_{\alpha=1}^n \frac{\partial f}{\partial u_\alpha} \nabla u_\alpha \quad (62)$$

梯度的几何应用:

椭圆例 23: 考虑由 $r_1 + r_2 = \text{常数}$ 定义的椭圆 (见图 34)。此时 $\nabla(r_1 + r_2)$ 与椭圆正交。设 \mathbf{T} 为椭圆的单位切向量, 则有:

$$\nabla(r_1 + r_2) \cdot \mathbf{T} = 0, \quad \text{即} \quad \nabla r_1 \cdot \mathbf{T} = -\nabla r_2 \cdot \mathbf{T} \quad (63)$$

根据例 20, ∇r_1 是平行于向量 \vec{AP} 的单位向量, 而 ∇r_2 是平行于向量 \vec{BP} 的单位向量。这表明向量 \vec{AP} 和 \vec{BP} 与椭圆切线的夹角相等。

19. 向量算子 ∇ 。我们定义:

$$\nabla \equiv \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \quad (64)$$

请注意, ∇ 是一个算子, 正如 $\frac{d}{dx}$ 是微分学中的一个算子一样。因此:

$$\begin{aligned}\nabla \varphi &= \left(\mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \right) \varphi \\ &= \mathbf{i} \frac{\partial \varphi}{\partial x} + \mathbf{j} \frac{\partial \varphi}{\partial y} + \mathbf{k} \frac{\partial \varphi}{\partial z}\end{aligned}$$

我们称 ∇ (读作 del) 为向量算子, 因为它的分量是 $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ 。在未来, 记住 ∇ 既表现为微分算子, 又表现为向量, 将对我们有所帮助。

例 24

$$\begin{aligned}\nabla(uv) &= \mathbf{i} \frac{\partial(uv)}{\partial x} + \mathbf{j} \frac{\partial(uv)}{\partial y} + \mathbf{k} \frac{\partial(uv)}{\partial z} \\ &= \left(\mathbf{i} \frac{\partial v}{\partial x} + \mathbf{j} \frac{\partial v}{\partial y} + \mathbf{k} \frac{\partial v}{\partial z} \right) u + \left(\mathbf{i} \frac{\partial u}{\partial x} + \mathbf{j} \frac{\partial u}{\partial y} + \mathbf{k} \frac{\partial u}{\partial z} \right) v\end{aligned}$$

$$\nabla(uv) = u\nabla v + v\nabla u \quad (65)$$

如果我们记住 ∇ 是一个微分算子，从而可以应用普通的微积分法则，那么这个结果就很容易记住了。

20. 向量的散度。

让我们考虑密度为 $\rho(x, y, z)$ 的流体的运动。我们假设其速度场由 $\mathbf{f} = u(x, y, z)\mathbf{i} + v(x, y, z)\mathbf{j} + w(x, y, z)\mathbf{k}$ 给出。由于 ρ 和 \mathbf{f} 显式地独立于时间（不随时间变化），这种类型的运动被称为定常运动（steady motion）。我们现在集中研究流经一个尺寸为 dx, dy, dz 的微小平行六面体 $ABCDEFGH$ （图 35）的流量。

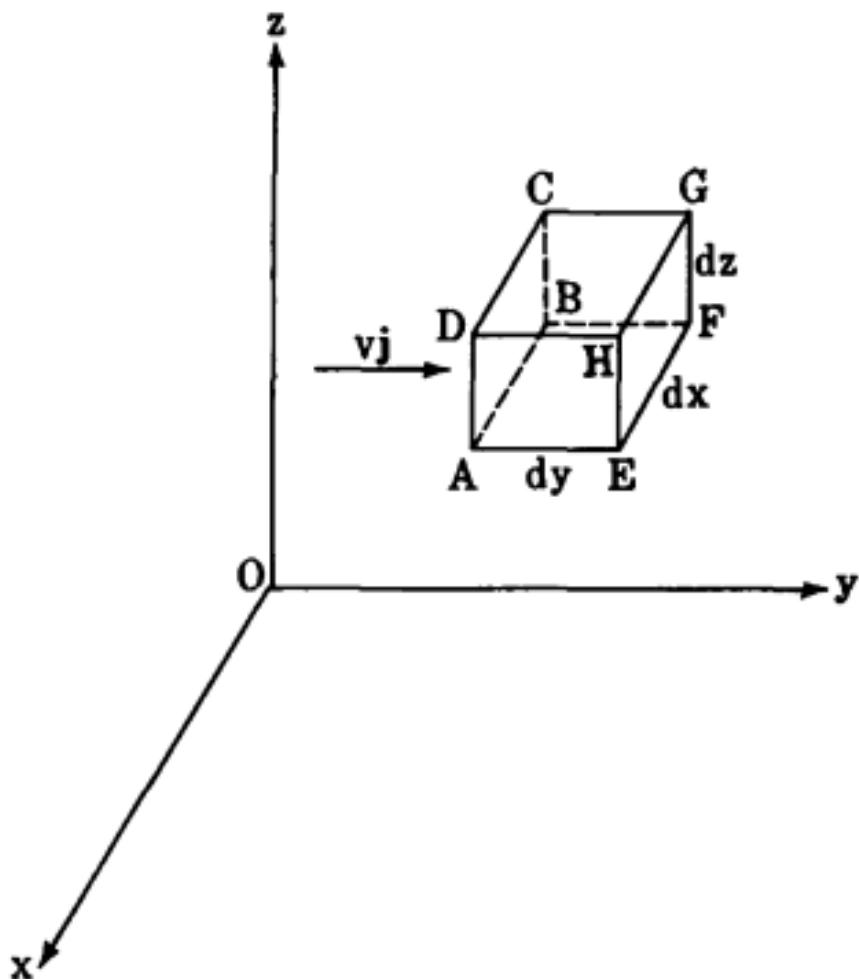


图 12.1: 向量的散度

1. 核心逻辑：空间的变化率

假设流体沿着 y 轴方向流动。在进入面 $ABCD$ 时，单位时间内流过的质量流量（密度 \times 速度 \times 截面积）可以表示为一个关于坐标的函数 $F(y)$ ，其中 $F = \rho v$ 。当流体经过一段极小的距离 dy 到达出口面 $EFGH$ 时，由于流场是不均匀的，出口处的流量 $F(y + dy)$ 通常不等于入口处的流量 $F(y)$ 。

2. 数学上的推导根据导数的定义，函数在 $y + dy$ 处的值可以通过其在 y 处的值加上变化量来表示：

$$F(y + dy) \approx F(y) + \frac{dF}{dy} \cdot dy$$

$$F(y + dy) = F(y) + dF(y) = \frac{dF}{dy} \cdot dy$$

将我们的物理量代入：

入口流量 (y 处): ρv (这里省略了截面积 $dx dz$)。

变化率: $\frac{\partial(\rho v)}{\partial y}$ ，这表示单位长度上质量流量的变化量。

总变化量: 变化率 \times 距离 dy ，即 $\frac{\partial(\rho v)}{\partial y} dy$ 。

所以，出口处的流量就等于：

$$\text{入口量} + \text{这一段路程中产生的增量} = \rho v + \frac{\partial(\rho v)}{\partial y} dy$$

3. 为什么这个“增量”导致了“损失”？

理解这个问题的关键在于净流量的概念：

净流出量 = 流出 - 流入。

代入公式: $\left[\rho v + \frac{\partial(\rho v)}{\partial y} dy \right] - \rho v = \frac{\partial(\rho v)}{\partial y} dy$ 。

如果这个结果是正数，说明“流出的比流入的多”。因为质量守恒，多出来的流体只能来自于这个小方块内部存储的流体，所以对这个方块本身来说，它损失了质量。

4. 举个直观的例子

想象一条高速公路：

入口 (y): 每分钟进来 100 辆车。

路段 (dy): 在这段路内，车速变快了或者车流变稀疏了。

出口 ($y+dy$): 每分钟出去了 110 辆车。

结果：出口比入口多了 10 辆。这多出来的 10 辆车（增量）必然导致了这段公路上的车辆总数在减少。这 10 辆车就是该路段的“车辆损失”。

首先，让我们计算单位时间内通过面 $ABCD$ 的流体量。速度 \mathbf{f} 的 x 和 z 分量对流经 $ABCD$ 的流量没有贡献。单位时间内进入面 $ABCD$ 的流体质量由 $\rho v dx dz$ 给出。单位时间内离开面 $EFGH$ 的流体质量为：

$$\left[\rho v + \frac{\partial(\rho v)}{\partial y} dy \right] dx dz$$

因此，单位时间内的质量损失等于：

$$\frac{\partial(\rho v)}{\partial y} dx dy dz$$

如果我们也考虑其他两个面，我们会发现单位时间内的总质量损失为：

$$\left[\frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} \right] dx dy dz$$

因此

$$\frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} \quad (66)$$

表示单位时间、单位体积内的质量损失。这个量被称为向量 $\rho \mathbf{f}$ 的散度。我们立刻可以看到：

$$\nabla \cdot (\rho \mathbf{f}) = \operatorname{div} (\rho \mathbf{f}) = \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = \frac{1}{V} \frac{dM}{dt} \quad (67)$$

由于 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 是常向量。 M 和 V 分别代表流体的质量和体积。

公式最后提到的 $\frac{1}{V} \frac{dM}{dt}$ 实际上表达了连续性方程：

$$\nabla \cdot (\rho \mathbf{f}) = -\frac{\partial \rho}{\partial t}$$

这里的散度代表了单位体积内质量流出的速率。

任意向量 \mathbf{f} 的散度定义为 $\nabla \cdot \mathbf{f}$ 。我们现在计算 $\varphi(x, y, z) \mathbf{f}$ 的散度：

$$\begin{aligned} \nabla \cdot (\varphi \mathbf{f}) &= \frac{\partial(\varphi u)}{\partial x} + \frac{\partial(\varphi v)}{\partial y} + \frac{\partial(\varphi w)}{\partial z} \\ &= \varphi \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) + \left(u \frac{\partial \varphi}{\partial x} + v \frac{\partial \varphi}{\partial y} + w \frac{\partial \varphi}{\partial z} \right) \end{aligned}$$

由此得到：

$$\nabla \cdot (\varphi \mathbf{f}) = \varphi \nabla \cdot \mathbf{f} + \mathbf{f} \cdot \nabla \varphi \quad (68)$$

如果我们把 ∇ 看作一个向量微分算子，就能很容易地记住这个结果。因此，当对 $\varphi \mathbf{f}$ 进行运算时，我们首先保持 φ 不变让 ∇ 作用于 \mathbf{f} ，然后保持 \mathbf{f} 不变让 ∇ 作用于 φ （注意： $\nabla \cdot \varphi$ 是没有意义的），由于 \mathbf{f} 和 $\nabla \varphi$ 都是向量，我们通过取它们的点积（内积）来完成乘法。

例 25 计算 $\nabla \cdot \mathbf{f}$ ，其中 $\mathbf{f} = \mathbf{r}/r^3$ （平方反比力）。

$$\begin{aligned}
\nabla \cdot (r^{-3} \mathbf{r}) &= r^{-3} \nabla \cdot \mathbf{r} + \mathbf{r} \cdot \nabla r^{-3} \\
&= 3r^{-3} + \mathbf{r} \cdot (-3r^{-4} \nabla r) \\
&= 3r^{-3} - 3r^{-5} \mathbf{r} \cdot \mathbf{r} = 3r^{-3} - 3r^{-3} = 0
\end{aligned}$$

$\nabla \cdot (r^{-3} \mathbf{r}) = 0 \quad (69)$

这是一个重要的结果：平方反比力的散度为零。我们注意到：

$$\nabla \cdot \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$$

公式 (69) 说明对于像引力或静电力这样的平方反比力场，在除源点 ($r = 0$) 之外的任何地方，场线的“流出”与“流入”是平衡的。

第一步：套用乘法法则（公式 68）

根据公式 (68): $\nabla \cdot (\varphi \mathbf{f}) = \varphi \nabla \cdot \mathbf{f} + \mathbf{f} \cdot \nabla \varphi$ 。在这里，我们令标量 $\varphi = r^{-3}$, 向量 $\mathbf{f} = \mathbf{r}$ 。带入后得到：

$$\nabla \cdot (r^{-3} \mathbf{r}) = r^{-3} (\nabla \cdot \mathbf{r}) + \mathbf{r} \cdot (\nabla r^{-3})$$

第二步：计算两个关键项

我们需要分别算出括号里的两个部分：

计算 $\nabla \cdot \mathbf{r}$ ：由于 $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, 根据散度定义：

$$\nabla \cdot \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 1 + 1 + 1 = 3$$

所以第一项变成了 $3r^{-3}$ 。

计算 ∇r^{-3} (标量的梯度)：这用到复合函数求导（链式法则）。由于 $r = (x^2 + y^2 + z^2)^{1/2}$:

$$\nabla r^n = nr^{n-1} \nabla r$$

当 $n = -3$ 时：

$$\nabla r^{-3} = -3r^{-4} \nabla r$$

而 ∇r (距离函数的梯度) 等于单位向量 $\frac{\mathbf{r}}{r}$ 。所以： $\nabla r^{-3} = -3r^{-4}(\frac{\mathbf{r}}{r}) = -3r^{-5}\mathbf{r}$ 。

在向量运算中，一个向量与其自身的点积等于其模长的平方，即 $\mathbf{r} \cdot \mathbf{r} = r^2$ 。

例 26 梯度的散度是什么？

$$\begin{aligned}
\nabla \cdot (\nabla \varphi) &= \nabla \cdot \left(\frac{\partial \varphi}{\partial x} \mathbf{i} + \frac{\partial \varphi}{\partial y} \mathbf{j} + \frac{\partial \varphi}{\partial z} \mathbf{k} \right) \\
&= \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2}
\end{aligned}$$

这个重要的量（梯度的散度）被称为 φ 的拉普拉斯算子（Laplacian）：

$$\text{Lap } \varphi = \nabla \cdot (\nabla \varphi) = \nabla^2 \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} \quad (70)$$

21. 向量的旋度。

我们暂时搁置旋度的物理含义，直接给出定义：

$$\text{curl } \mathbf{f} = \nabla \times \mathbf{f} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ u & v & w \end{vmatrix}$$

$$\nabla \times \mathbf{f} = \mathbf{i} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) + \mathbf{j} \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right) + \mathbf{k} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \quad (71)$$

例 27 径向量

$$\nabla \times \mathbf{r} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x & y & z \end{vmatrix} = 0$$

(注：这说明向径向量场是无旋的)

例 28

$$\begin{aligned} \nabla \times (\varphi \mathbf{f}) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \varphi u & \varphi v & \varphi w \end{vmatrix} \\ &= \mathbf{i} \left[\frac{\partial(\varphi w)}{\partial y} - \frac{\partial(\varphi v)}{\partial z} \right] + \mathbf{j} \left[\frac{\partial(\varphi u)}{\partial z} - \frac{\partial(\varphi w)}{\partial x} \right] + \mathbf{k} \left[\frac{\partial(\varphi v)}{\partial x} - \frac{\partial(\varphi u)}{\partial y} \right] \\ &= \varphi \left[\mathbf{i} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) + \dots \right] + \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} & \frac{\partial \varphi}{\partial z} \\ u & v & w \end{vmatrix} \end{aligned}$$

最终得到公式：

$$\nabla \times (\varphi \mathbf{f}) = \varphi \nabla \times \mathbf{f} + \nabla \varphi \times \mathbf{f} \quad (72)$$

通过将 ∇ 视为一个向量微分算子，可以很容易地得到这个结果。

例 29. 证明梯度的旋度为零。

$$\nabla \times (\nabla \varphi) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial \varphi}{\partial x} & \frac{\partial \varphi}{\partial y} & \frac{\partial \varphi}{\partial z} \end{vmatrix} = \mathbf{i} \left(\frac{\partial^2 \varphi}{\partial y \partial z} - \frac{\partial^2 \varphi}{\partial z \partial y} \right) + \mathbf{j} \left(\frac{\partial^2 \varphi}{\partial z \partial x} - \frac{\partial^2 \varphi}{\partial x \partial z} \right) + \mathbf{k} \left(\frac{\partial^2 \varphi}{\partial x \partial y} - \frac{\partial^2 \varphi}{\partial y \partial x} \right)$$

因此：

$$\nabla \times \nabla \varphi = 0 \quad (73)$$

前提是 φ 具有连续的二阶导数。

(注：根据偏导数无关次序的原则，交叉偏导数相等，故结果为零)

例 30. 证明旋度的散度为零。

$$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{f}) &= \frac{\partial}{\partial x} \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) + \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right) + \frac{\partial}{\partial z} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) \\ &= \frac{\partial^2 w}{\partial y \partial x} - \frac{\partial^2 v}{\partial z \partial x} + \frac{\partial^2 u}{\partial z \partial y} - \frac{\partial^2 w}{\partial x \partial y} + \frac{\partial^2 v}{\partial x \partial z} - \frac{\partial^2 u}{\partial y \partial z} \end{aligned}$$

由此得到：

$$\nabla \cdot (\nabla \times \mathbf{f}) = 0 \quad (74)$$

例 31. $(\mathbf{u} \cdot \nabla) \mathbf{v}$ 是什么意思？我们首先计算 \mathbf{u} 与 ∇ 的点积。这产生了一个标量微分算子：

$$u_x \frac{\partial}{\partial x} + u_y \frac{\partial}{\partial y} + u_z \frac{\partial}{\partial z}$$

然后我们用它作用于向量 \mathbf{v} ，得到：

$$(\mathbf{u} \cdot \nabla) \mathbf{v} = u_x \frac{\partial \mathbf{v}}{\partial x} + u_y \frac{\partial \mathbf{v}}{\partial y} + u_z \frac{\partial \mathbf{v}}{\partial z}$$

因此：

$$\begin{aligned} d\mathbf{f} &= \frac{\partial \mathbf{f}}{\partial x} dx + \frac{\partial \mathbf{f}}{\partial y} dy + \frac{\partial \mathbf{f}}{\partial z} dz \\ &= dx \frac{\partial \mathbf{f}}{\partial x} + dy \frac{\partial \mathbf{f}}{\partial y} + dz \frac{\partial \mathbf{f}}{\partial z} \end{aligned}$$

(注：此处 $d\mathbf{f}$ 描述了向量场 \mathbf{f} 随位移 $d\mathbf{r}$ 的全微分变化量)

重点概念解析

两个为零的恒等式：

$\nabla \times \nabla \varphi = 0$: 梯度的场一定是“无旋”的。在物理上，这意味着保守场（如重力场）没有涡流。

$\nabla \cdot (\nabla \times \mathbf{f}) = 0$: 旋度的场一定是“无源”的。这意味着涡旋线既没有起点也没有终点，总是闭合的。

方向导数算子 $(\mathbf{u} \cdot \nabla)$:

这个算子在流体力学中极其重要，它描述了物理量随流体运动而产生的变化（即对流项）。例如在纳维-斯托克斯方程中，速度场的自对流项就是用这种形式表示的。

23. 曲线坐标系 (Curvilinear Coordinates)

数学家、物理学家或工程师经常发现，使用除了熟悉的直角笛卡尔坐标系以外的其他坐标系会更加方便。如果他正在处理球体问题，他可能会发现用球面坐标 r, θ, φ 来描述空间中点的位置更为得当（见图 31）。

让我们注意以下定义：

$$\text{球面: } x^2 + y^2 + z^2 = r^2$$

$$\text{圆锥面: } z/(x^2 + y^2 + z^2)^{1/2} = \cos \theta$$

$$\text{平面: } y/x = \tan \varphi$$

这些面都通过点 $P(r, \theta, \varphi)$ 。我们可以考虑如下从 $x - y - z$ 坐标系到 $r - \theta - \varphi$ 坐标系的变换：

$$r = (x^2 + y^2 + z^2)^{1/2}$$

$$\theta = \cos^{-1} \frac{z}{(x^2 + y^2 + z^2)^{1/2}}$$

$$\varphi = \tan^{-1} \frac{y}{x}$$

表面 $r = c_1, \theta = c_2, \varphi = c_3$ 分别代表球面、圆锥面和平面。空间中除了原点外的任何点 P ，都恰好会有每种类型的一个面通过它，点 P 的坐标由常数 c_1, c_2, c_3 决定。

球面与圆锥面的交线是一个圆，即纬线圈，在该圆上点 P 具有单位切向量 \mathbf{e}_φ 。这个圆被称为 φ -曲线，因为在该曲线上 r 和 θ 保持不变，只有坐标 φ 随着我们沿曲线移动而改变。

球面与平面的交线产生 θ -曲线（经线圈）；而圆锥面与平面的交线产生从原点通过 P 点的直线，即 r -曲线。

在 P 点的三个单位向量 $\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{e}_\varphi$ 是互相垂直的，可以被视为在 P 点邻域内形成了一个坐标系的基底。与 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 不同的是，它们不是固定不变的，因为当我们从一点移动到另一点时，它们的方向会发生改变。

因此，当处理球面坐标时，我们可以预见到梯度 (gradient)、散度 (divergence)、旋度 (curl) 和拉普拉斯算子 (Laplacian) 的公式会变得更加复杂。

24. Frenet-Serret 公式

欧几里得空间中的三维曲线可以用位置向量终点的轨迹表示，如下所示：

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} \quad (93)$$

其中 t 是一个在值集 $t_0 \leq t \leq t_1$ 范围内变化的参数。我们假设 $x(t), y(t), z(t)$ 具有各阶连续导数，并且可以在曲线任何点的邻域内展开为泰勒级数。

我们在第 2 章第 16 节中已经看到， $\frac{dt}{ds}$ 是曲线的单位切向量。令 $\mathbf{t} = \frac{d\mathbf{r}}{ds}$ 。现在 \mathbf{t} 是一个单位向量，因此它的导数垂直于 \mathbf{t} 。此外，该导数 $\frac{d\mathbf{t}}{ds}$ 告诉我们当我们沿曲线移动时，单位切向量改变方向的速度。曲线的主法线相应地由下式定义：

$$\frac{dt}{ds} = \kappa n \quad (94)$$

其中 κ 是 $\frac{dt}{ds}$ 的模，被称为曲率 (curvature)。曲率的倒数 $\rho = 1/\kappa$ 被称为曲率半径 (radius of curvature)。重要的是要注意，方程 (94) 同时定义了 κ 和 n ， κ 是 $\frac{dt}{ds}$ 的长度，而 n 是平行于 $\frac{dt}{ds}$ 的单位向量。在曲线的任意点 P ，我们现在拥有两个彼此成直角的向量 t, n (见图 37)。

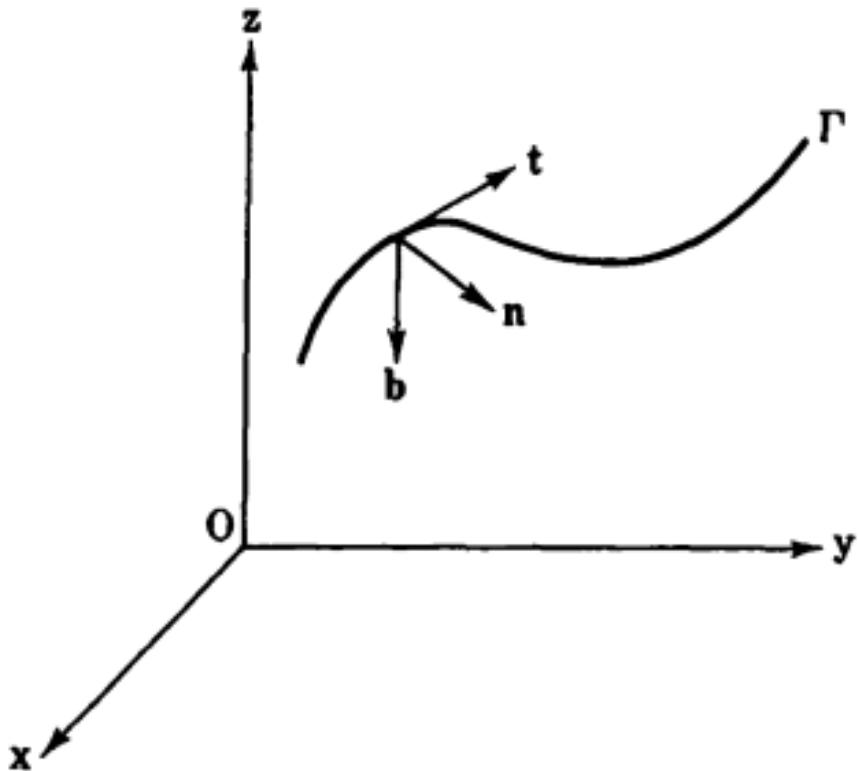


FIG. 37.

图 12.2: 曲线的主法线

在曲线的任意点 P ，我们现在拥有两个彼此垂直的向量 t, n (见图 37)。通过定义第三个与 t 和 n 均垂直的向量，我们可以在 P 点建立一个局部坐标系。我们将该向量定义为副法线 (binormal) 向量：

$$b = t \times n$$

所有与曲线在 P 点相关的向量都可以写成这三个基本向量 t, n, b 的线性组合，它们在 P 点构成了一个三面体 (triangular)。

现在让我们计算 $\frac{db}{ds}$ 和 $\frac{dn}{ds}$ 。由于 b 是单位向量，其导数垂直于 b ，因此位于 t 和 n 构成的平面内。此外，由于 $b \cdot t = 0$ ，通过求导我们得到：

$$\frac{d\mathbf{b}}{ds} \cdot \mathbf{t} + \kappa \mathbf{b} \cdot \mathbf{n} = 0$$

即 $\frac{d\mathbf{b}}{ds} \cdot \mathbf{t} = 0$ 。因此 $\frac{d\mathbf{b}}{ds}$ 也垂直于 \mathbf{t} , 这意味着 $\frac{d\mathbf{b}}{ds}$ 必须平行于 \mathbf{n} 。由此可知:

$$\frac{d\mathbf{b}}{ds} = \tau \mathbf{n}$$

其中 τ 被定义为 $\frac{d\mathbf{b}}{ds}$ 的模 (带有方向符号), 被称为曲线的挠率 (torsion)。

主法线导数 $\frac{d\mathbf{n}}{ds}$ 的推导

利用关系式 $\mathbf{n} = \mathbf{b} \times \mathbf{t}$, 对其求导:

$$\frac{d\mathbf{n}}{ds} = \mathbf{b} \times \frac{dt}{ds} + \frac{d\mathbf{b}}{ds} \times \mathbf{t}$$

代入已知的 $\frac{dt}{ds} = \kappa \mathbf{n}$ 和 $\frac{d\mathbf{b}}{ds} = \tau \mathbf{n}$:

$$\frac{d\mathbf{n}}{ds} = \mathbf{b} \times (\kappa \mathbf{n}) + (\tau \mathbf{n}) \times \mathbf{t}$$

根据叉乘性质 $\mathbf{b} \times \mathbf{n} = -\mathbf{t}$ 以及 $\mathbf{n} \times \mathbf{t} = -\mathbf{b}$:

$$\frac{d\mathbf{n}}{ds} = -\kappa \mathbf{t} - \tau \mathbf{b}$$

总结: 著名的 Frenet-Serret 公式组

$$\begin{cases} \frac{dt}{ds} = \kappa \mathbf{n} \\ \frac{d\mathbf{n}}{ds} = -\kappa \mathbf{t} - \tau \mathbf{b} \\ \frac{d\mathbf{b}}{ds} = \tau \mathbf{n} \end{cases} \quad (95)$$

这些公式表明, 曲线在空间中的形态完全由曲率 κ (反映弯曲程度) 和挠率 τ (反映扭曲程度) 决定。

例 42: 圆柱螺旋线 (Circular Helix) 的代数推导

给定圆柱螺旋线的参数方程为:

$$\mathbf{r}(t) = a \cos t \mathbf{i} + a \sin t \mathbf{j} + bt \mathbf{k}$$

1. 计算弧长变化率 $\frac{ds}{dt}$

首先, 我们需要求出速度向量 (对参数 t 求导):

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = -a \sin t \mathbf{i} + a \cos t \mathbf{j} + b \mathbf{k}$$

弧长 s 对时间 t 的变化率为该向量的模:

$$\frac{ds}{dt} = \left| \frac{d\mathbf{r}}{dt} \right| = \sqrt{(-a \sin t)^2 + (a \cos t)^2 + b^2} = \sqrt{a^2(\sin^2 t + \cos^2 t) + b^2} = \sqrt{a^2 + b^2}$$

由此得: $\frac{dt}{ds} = (a^2 + b^2)^{-1/2}$ 。

2. 求单位切向量 \mathbf{t}

根据定义 $\mathbf{t} = \frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt} \cdot \frac{dt}{ds}$:

$$\mathbf{t} = (-a \sin t \mathbf{i} + a \cos t \mathbf{j} + b \mathbf{k})(a^2 + b^2)^{-1/2}$$

验证其大小: $\mathbf{t} \cdot \mathbf{t} = (a^2 \sin^2 t + a^2 \cos^2 t + b^2)(a^2 + b^2)^{-1} = 1$ 。

3. 求曲率 κ 和主法向量 \mathbf{n}

利用 Frenet 公式 $\frac{d\mathbf{t}}{ds} = \kappa \mathbf{n}$ 。我们先对 t 求导再转换:

$$\frac{d\mathbf{t}}{ds} = \frac{d\mathbf{t}}{dt} \cdot \frac{dt}{ds} = [(-a \cos t \mathbf{i} - a \sin t \mathbf{j})(a^2 + b^2)^{-1/2}] \cdot (a^2 + b^2)^{-1/2}$$

$$\frac{d\mathbf{t}}{ds} = (-a \cos t \mathbf{i} - a \sin t \mathbf{j})(a^2 + b^2)^{-1}$$

由于 κ 是该向量的模, 且 \mathbf{n} 是单位向量:

$$\text{曲率: } \kappa = \left| \frac{d\mathbf{t}}{ds} \right| = \sqrt{a^2 \cos^2 t + a^2 \sin^2 t}(a^2 + b^2)^{-1} = a(a^2 + b^2)^{-1}$$

$$\text{主法向量: } \mathbf{n} = \frac{1}{\kappa} \frac{d\mathbf{t}}{ds} = -\cos t \mathbf{i} - \sin t \mathbf{j}$$

4. 求副法向量 \mathbf{b}

通过叉乘计算 $\mathbf{b} = \mathbf{t} \times \mathbf{n}$:

$$\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -a \sin t & a \cos t & b \\ -\cos t & -\sin t & 0 \end{vmatrix} (a^2 + b^2)^{-1/2}$$

展开行列式:

$$\mathbf{b} = [(0 - (-b \sin t))\mathbf{i} - (0 - (-b \cos t))\mathbf{j} + (a \sin^2 t - (-a \cos^2 t))\mathbf{k}](a^2 + b^2)^{-1/2}$$

$$\mathbf{b} = (b \sin t \mathbf{i} - b \cos t \mathbf{j} + a \mathbf{k})(a^2 + b^2)^{-1/2}$$

5. 求挠率 τ

利用公式 $\frac{d\mathbf{b}}{ds} = \tau \mathbf{n}$ 。对 \mathbf{b} 求导:

$$\frac{d\mathbf{b}}{ds} = \frac{d\mathbf{b}}{dt} \cdot \frac{dt}{ds} = [(b \cos t \mathbf{i} + b \sin t \mathbf{j})(a^2 + b^2)^{-1/2}] \cdot (a^2 + b^2)^{-1/2}$$

$$\frac{d\mathbf{b}}{ds} = (b \cos t \mathbf{i} + b \sin t \mathbf{j})(a^2 + b^2)^{-1}$$

观察 $\mathbf{n} = -\cos t \mathbf{i} - \sin t \mathbf{j}$, 可以写成:

$$\frac{d\mathbf{b}}{ds} = -b(a^2 + b^2)^{-1}\mathbf{n}$$

因此，挠率为：

$$\tau = -b(a^2 + b^2)^{-1}$$

25. 基本平面 (Fundamental Planes)

在曲线上的 P 点，由 $\mathbf{t}, \mathbf{n}, \mathbf{b}$ 向量两两确定的三个平面非常重要：

密切平面 (Osculating Plane)：

包含切向量 \mathbf{t} 和主法线 \mathbf{n} 的平面。

因为它垂直于副法线 \mathbf{b} ，若 \mathbf{s} 为平面内任意一点的变化向量，则方程为：

$$(\mathbf{s} - \mathbf{r}) \cdot \mathbf{b} = 0 \quad (96)$$

法平面 (Normal Plane)：

通过 P 点并垂直于切向量 \mathbf{t} 的平面。

方程为：

$$(\mathbf{s} - \mathbf{r}) \cdot \mathbf{t} = 0 \quad (97)$$

从切平面/展直平面 (Rectifying Plane)：

通过 P 点并垂直于主法线 \mathbf{n} 的平面。

方程为：

$$(\mathbf{s} - \mathbf{r}) \cdot \mathbf{n} = 0 \quad (98)$$