

计算与应用数学

zdd

2025 年 11 月 17 日



目录

第一部分 基本知识	1
1 ODE	1
1.1 ODE	1
1.2 初值问题与边值问题	2
1.3 齐次线性方程	3
1.4 齐次线性方程解法	6
1.5 非齐次线性函数	8
1.6 一阶非线性微分方程	12
1.7 高阶非线性微分方程	14
1.8 特征值问题	15
1.9 复平面中的微分方程	18
2 差分方程	19
2.1 差分	19
2.2 初等差分方程	20
2.3 齐次线性差分方程	22
2.4 非齐次线性差分方程	27
2.5 非线性差分方程	29
第二部分 局部分析	30
3 线性微分方程的近似解法	30
3.1 齐次线性方程的奇点分类	30
3.2 齐次线性方程常点附近的局部级数展开	33
3.3 齐次线性方程正则奇点附近的局部级数展开	34
3.4 齐次线性方程非正则奇点处的局部级数展开	40
3.5 无穷处的非正则奇点	46
3.6 非齐次线性方程的局部分析	50
3.7 渐近关系	51
3.8 渐进级数	56
4 非线性微分方程的近似解法	65
5 差分方程的近似解法	66
6 积分的渐近展开	67
第三部分 扰动法	68
7 扰动级数	68

8 求和级数	69
第四部分 全局分析	70
9 边界层理论	70
10 WKB 理论	71
11 多尺度分析	72
第五部分 凸优化	73
12 基础定义	73
12.1 一些符号约定	73
12.2 凸集与凸函数	73
12.3 可微凸函数	76
12.4 凸包、仿射包和相对内点	79
12.5 凸函数的性质	83
12.6 局部与全局最小值	87
12.7 渐进序列	89
12.8 超平面	91
12.9 最小公共点和最大相交点问题	95
12.10 零和游戏	98
第六部分 PDE 数值解	99
13 基础定义	99
13.1 二阶 PDE 的分类	99
13.2 Fourier 分析	101
14 椭圆 PDE 的有限差分方法	102

第一部分 基本知识

1 ODE

1.1 ODE

定义 1.1

一个 n 阶微分方程具有如下形式

$$y^{(n)}(x) = F[x, y(x), y'(x), \dots, y^{(n-1)}(x)], \quad (1.1.1)$$

其中 $y^{(k)} = d^k y / dx^k$ 。

如果 F 是 y 及其导数的线性函数，则称方程 (1.1.1) 是线性微分方程。

如果 (1.1.1) 是线性的，则其通解 $y(x)$ 依赖于 n 个独立的参数（称为积分常数）且其所有解都可以通过适当选择这些常数得到。如果 (1.1.1) 是非线性微分方程，则它也有一个包含 n 个积分常数的通解。然而，非线性微分方程有时存在特殊的额外解，这些解无法通过选择积分常数从通解中得到。我们省略对这些微分方程基本性质的严格讨论，在接下来的三个例子中说明它们。

例 1.1 (可分方程)

可分方程 (Separable equations) 是最容易求解的微分方程。如果一个方程是一阶的，且方程 (1.1.1) 中的 F 关于 x 和 y 的依赖可以分离，则称它是可分的。

最一般的可分方程是

$$y'(x) = a(x)b(y) \quad (1.1.2)$$

直接积分给出通解

$$\int^y \frac{dt}{b(t)} = \int^x a(s)ds + c_1 \quad (1.1.3)$$

其中 c_1 是积分常数。

形式上，线性微分方程可写为

$$Ly(x) = f(x) \quad (1.1.4)$$

其中 L 是线性微分算子：

$$L = p_0(x) + p_1(x)\frac{d}{dx} + \dots + p_{n-1}(x)\frac{d^{n-1}}{dx^{n-1}} + \frac{d^n}{dx^n} \quad (1.1.5)$$

定义 1.2

若 $f(x) \equiv 0$ ，则称微分方程 (1.1.4) 是齐次 (homogeneous) 的，否则是非齐次 (inhomogeneous) 的。

例 1.2 (线性方程的解)

齐次线性方程

$$y'' - \frac{1+x}{x}y' + \frac{1}{x}y = 0 \quad (1.1.6)$$

的通解为 $y(x) = c_1 e^x + c_2(1+x)$, 这显式地依赖于两个积分常数 c_1 和 c_2 , 这也给出了方程 (1.1.6) 的所有解。

例 1.3 (非线性方程的解)

两个常规可解的非线性微分方程是 Riccati 方程

$$y' = \frac{A^2}{x^4} - y^2, \quad A \text{ 是常数} \quad (1.1.7)$$

其通解为

$$y(x) = \frac{1}{x} + \frac{A}{x^2} \frac{c_1 - e^{2A/x}}{c_1 + e^{2A/x}} \quad (1.1.8)$$

和等维方程

$$y'' = yy'/x \quad (1.1.9)$$

其通解为

$$y(x) = 2c_1 \tan(c_1 \ln x + c_2) - 1 \quad (1.1.10)$$

方程 (1.1.9) 存在一个特解 $y = c_3$, 该解无法通过选择 c_1 和 c_2 从通解 (1.1.10) 中得到。

1.2 初值问题与边值问题

微分方程的解 $y(x)$ 不能仅由方程本身唯一确定, 必须指定 n 个独立积分常数的值。这些积分常数可通过多种不同方式指定。

在**初值问题 (initial-value problem)**中, 在单点 $x = x_0$ 处指定 y 及其前 $n-1$ 阶导数 $y', \dots, y^{(n-1)}$ 的值:

$$\begin{aligned} y(x_0) &= a_0, \\ y'(x_0) &= a_1, \dots, \\ y^{(n-1)}(x_0) &= a_{n-1}. \end{aligned} \quad (1.2.1)$$

其中 $a_0, \dots, a_{n-1}$ 是初值问题的 n 个积分常数。在**边值问题 (boundary value problem)**中, 在两个或多个点处总共指定 n 个量。例如, 对于五阶微分方程, 可施加条件: $y(x_1) = a_1, y'(x_2) = a_2, y''(x_3) = a_3, y'''(x_4) = a_4, y'''(x_1) + [y(x_2)]^2 = a_5$, 其中一些点 x_i 可能位于求解区域 $y(x)$ 的内部而非边界上。但无论如何, a_i 仍是该问题的积分常数。

关于初值和边值问题解的存在唯一性定理为应用分析的两个层次提供了基础。局部分析研究解在某点邻域的性质, 当邻域足够小时结果有效。全局分析则试图在有限区域内确定解的性质, 这一般更为困难。例如, 渐近级数仅局部有效, 而 WKB 近似是全局的。

初值问题本质上比边值问题简单得多。可通过局部分析判断在 x_0 的足够小邻域内是否存在唯一解 $y(x)$ 。例如, 可重复微分方程 (1.1.1) 以计算 $y(x)$ 在 x_0 处 Taylor 展开的系数。一般而言, 若 $F, \partial F/\partial y, \partial F/\partial y', \dots, \partial F/\partial y^{(n-1)}$ 在 $x_0, a_0, a_1, \dots, a_{n-1}$ 的某邻域内均为 $x, y, y', \dots, y^{(n-1)}$ 的连续函数, 则在包含 x_0 的某区间上存在满足 (1.1.1) 和 (1.2.1) 的唯一解 $y(x)$ 。

例 1.4 (存在唯一解的初值问题)

由于满足存在唯一性定理的条件, 以下五个初值问题均有唯一解 (当 x 充分接近 0 时存在):

- (a) $y' = \sin(xy), \quad y(0) = 1;$
 (b) $y' = (x+y)x^2y^2, \quad y(0) = 1;$
 (c) $y' = e^x + x/y, \quad y(0) = 1;$
 (d) $y'' = y^2 + e^x, \quad y(0) = y'(0) = 0;$
 (e) $y''' = e^{xy'}, \quad y(0) = y'(0) = y''(0) = 0.$

遗憾的是, 这些解的存在性仅是理论上的, 因为我们无法用已知函数表示其中任何一个。

若 F 不是其自变量的光滑函数, 则存在唯一性定理可能不成立。

例 1.5 (具有非唯一解的微分方程)

初值问题 $y' = y^{1/3}$ [$y(0) = 0$] 中, F 在 $x = 0, y = 0$ 附近连续但 $\partial F / \partial y$ 不连续。因此该问题存在两个解: $y = 0$ 和 $y = (2x/3)^{3/2}$ 。注意到一阶方程 $y' = y^{1/3}$ 的通解为 $y(x) = \eta(2x/3 + c_1)^{3/2}$, 其中 $\eta = \pm 1, 0$, c_1 任意。此通解不仅依赖于积分常数 c_1 , 还依赖于离散参数 η 。一般地, 若 (1.1.1) 中的 F 不光滑, 则解 $y(x)$ 可能依赖于超过 n 个积分常数。

例 1.6 (存在唯一解的微分方程)

初值问题 $y' = x^{-1/2}$ [$y(0) = 1$] 确有唯一解 $y = 2x^{1/2} + 1$, 尽管 F 在 $x = 0$ 不连续。但这并不违背唯一性定理, 因该定理给出的是存在唯一性的充分非必要条件。

若初值问题在 x_0 附近存在解, 下一步是确定该解的全局性质, 特别是确定局部解存在的区间范围。这对线性方程较容易, 可以证明: 与初始条件无关, 该区间至少与不包含 $p_0(x), p_1(x), \dots, p_{n-1}(x)$ 间断点的最大区间一样大, 作进一步分析时可能发现区间更大。

例 1.7 (线性方程解的存在区间)

习题 1.4 的结果保证初值问题 $y' = (\tan x)y + 1$ [$y(0) = 1$] 的解在 $\tan x$ 连续的区间内存在, 即 $(-\pi/2, \pi/2)$ 。然而, 其解 $y(x) = (1 + \sin x)/\cos x$ 实际存在于 $-3\pi/2 < x < \pi/2$ 区间。

非线性微分方程初值问题的全局分析要困难得多。

边值问题本质上是全局的。解的存在唯一性定理必须在包含所有点 $x_1, x_2, \dots$ 的足够大区间上证明。线性微分方程的全局性质相对容易得到, 但对非线性微分方程边值问题解的存在唯一性建立严格准则非常困难。

1.3 齐次线性方程

本节回顾齐次线性微分方程初值和边值问题的理论。

解的线性无关性

命题 1.1

n 阶齐次线性方程

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \cdots + p_0(x)y = 0 \quad (1.3.1)$$

的通解具有形式:

$$y(x) = \sum_{j=1}^n c_j y_j(x) \quad (1.3.2)$$

其中 c_j 是任意积分常数, $\{y_j(x)\}$ 是线性无关函数集, 且每个函数均满足 (1.3.1)。在系数 $p_0(x), p_1(x), \cdots, p_{n-1}(x)$ 连续的任意区域内, 方程 (1.3.1) 总存在恰好 n 个线性无关解。

若存在不全为零的数集 $\{c_j\}$ 满足 $\sum_{j=1}^n c_j y_j(x) = 0$, 则称函数集 $\{y_j(x)\}$ 线性相关 (linearly dependent)。若不存在这样的集合, 则 $\{y_j(x)\}$ 线性无关 (linearly independent)。

注: 由于线性无关是全局概念, 故在单点讨论线性无关无意义。

对于可微函数集我们有简单的线性相关性检验方法。

朗斯基行列式

定义 1.3 (朗斯基行列式)

朗斯基行列式 (Wronskian) $W(x)$ 定义为行列式:

$$W(x) = W[y_1(x), y_2(x), \cdots, y_n(x)] \equiv \det \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \vdots & \vdots & \ddots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix} \quad (1.3.3)$$

命题 1.2

$W(x)$ 在区间上恒为零当且仅当 $\{y_j(x)\}$ 是线性相关函数集。若 $\{y_j(x)\}$ 在某个区间线性无关, 则 $W(x)$ 不恒为零 (可能在孤立点为零)。

例 1.8 (线性相关)

对所有 x 有 $W[e^x, e^{-x}, \cosh x] = 0$, 故 $[e^x, e^{-x}, \cosh x]$ 是线性相关函数集。

例 1.9 (线性无关)

为验证方程 $y'' - y'(1+x)/x + y/x = 0$ 的解 $y(x) = c_1 e^x + c_2(1+x)$ 确为通解, 计算朗斯基行列式: $W[1+x, e^x] = xe^x$ 。而 xe^x 仅在 $x=0$ 处为零, 因此对所有 x , $\{1+x, e^x\}$ 是线性无关集。

齐次线性方程具有如下性质:

命题 1.3

方程 (1.3.1) 任意 n 个解的朗斯基行列式 $W(x)$ 满足简单一阶方程:

$$W'(x) = -p_{n-1}(x)W(x) \quad (1.3.4)$$

方程 (1.3.4) 的解称为阿贝尔公式:

$$W(x) = \exp \left[- \int^x p_{n-1}(t) dt \right]. \quad (1.3.5)$$

因此, 我们可以在得到微分方程的任何解之前计算 $W(x)$ 。

现在使用这些理论结果讨论初值和边值问题的适定性。此处定义**适定**问题为存在唯一解的问题。

初值问题

求解初值问题需选择 (1.3.2) 中的 c_j 以满足 (1.2.1) 的初始条件。 c_j 由 n 个联立代数方程组 $\sum_{j=1}^n c_j y_j^{(i)}(x_0) = a_i$ ($i = 0, 1, \dots, n-1$) 确定。根据 Cramer 法则, 这些方程有唯一解仅当:

$$\det [y_j^{(i)}(x_0)] = W(x_0) \neq 0. \quad (1.3.6)$$

因此, 朗斯基行列式自然出现在初值问题研究中。它实际有两种应用。首先, 可通过检验 $\{y_j\}$ 是否线性无关集, **全局**判断 (1.3.2) 形式的解是否为 (1.3.1) 的通解。

其次, 朗斯基行列式可**局部**用于发现不适定初值问题, 无需实际求解微分方程, 仅需计算 (1.3.5) 并参考 (1.3.6)。若齐次初值问题的初始条件给定在朗斯基行列式为零的点 x_0 (由阿贝尔公式计算), 则该问题不适定。

例 1.10 (不适定问题——朗斯基行列式为零)

方程 (1.1.6) 的朗斯基行列式仅在 $x = 0$ 处为零。因此初值问题 $y'' - y'(1+x)/x + y/x = 0$ [$y(1) = 1, y'(1) = 2$] 适定并不意外。当初始条件替换为 $y(0) = 1, y'(0) = 2$ 时, 无解。当替换为 $y(0) = 1, y'(0) = 1$ 时, 解 $y(x) = e^x + c(e^x - 1 - x)$ 不唯一。

遗憾的是, 朗斯基行列式非零不意味着问题适定。

例 1.11 (不适定问题——朗斯基行列式非零)

初值问题 $y'' - 6y/x^2 = 0$ [$y(0) = 6, y'(0) = 6$] 不适定, 因其通解 $y(x) = c_1 x^3 + c_2 x^{-2}$ 在 $x = 0$ 处或发散或为零。阿贝尔公式 (1.3.5) 表明朗斯基行列式处处为常数 (含 $x = 0$), 且该常数非零 (因 $y(x)$ 是通解)。因此, (1.2.1) 中 $W(y)$ 在 $x = x_0$ 非零是初值问题适定的必要非充分条件。当然, 此问题不适定是显然的, 初始条件给在 $p_0(x)$ 的间断点。

因此, 尽管朗斯基行列式在理论上是有趣的对象, 但作为初值问题局部分析工具有时不实用。

边值问题

边值问题的解由两个或多个不同点处的条件确定。与初值问题不同，系数函数的连续性不足以保证边值问题有解：

例 1.12 (无解的边值问题)

考虑边值问题 $y'' + y = 0$ [$y'(0) = 0, y(\pi/2) = 1$]。微分方程通解为 $y(x) = c_1 \sin x + c_2 \cos x$ 。朗斯基行列式 $W(\sin x, \cos x) = -1$ 非零且微分方程系数函数连续。然而，边值问题无解：条件 $y'(0) = 0$ 蕴含 $c_1 = 0$ ，剩余解 $y(x) = c_2 \cos x$ 无法满足 $y(\pi/2) = 1$ 。

上例可能暗示边值问题很少有解。但事实正相反：多数边值问题存在唯一解。考虑一般二阶线性微分方程 $y'' + p_1(x)y' + p_0(x)y = 0$ 及边界条件 $y'(x_0) = a, y(x_1) = b$ 。若 $y_1(x)$ 和 $y_2(x)$ 是微分方程的两个线性无关解，则通解可写为 $y(x) = Ay_1(x) + By_2(x)$ ，边界条件要求： $Ay_1'(x_0) + By_2'(x_0) = a, Ay_1(x_1) + By_2(x_1) = b$ 。当行列式非零时，可由 a, b 唯一确定 A, B ：

$$\det \begin{vmatrix} y_1'(x_0) & y_2'(x_0) \\ y_1(x_1) & y_2(x_1) \end{vmatrix} \neq 0$$

此条件是 (1.3.6) 中朗斯基条件的推广，通常成立（保证边值问题有唯一解）。但若选择点 x_0 和 x_1 使得 $y_1'(x_0)y_2(x_1) - y_2'(x_0)y_1(x_1) = 0$ ，则边值问题不适定：或无解，或有无穷多解。若 $by_1'(x_0) = ay_1(x_1)$ 且 $y_1'(x_0) \neq 0, y_1(x_1) \neq 0$ 时，后者可能出现。此时边值问题有无穷多解，它们相差函数 $z(x) = y_1(x)y_2'(x_0) - y_2(x)y_1'(x_0)$ 的任意常数倍。

1.4 齐次线性方程解法

本节讨论如何解方程 $Ly = 0$ 。一阶方程易解（可分）。但当阶数 ≥ 2 时，显式解仅在少数情况下存在。我们简要分类并讨论部分可解方程。

常系数方程

常系数方程的特征是 $p_0, p_1, \dots, p_{n-1}$ 与 x 无关。寻求形如 $y(x) = e^{rx}$ 的解，代入微分方程得 $L[e^{rx}] = P(r)e^{rx}$ ，其中

$$P(r) = r^n + \sum_{j=0}^{n-1} p_j r^j$$

是 n 次多项式。对应 $P(r) = 0$ 的不同根 $r_1, r_2, \dots$ 的解为：

$$y = e^{r_1 x}, e^{r_2 x}, \dots \quad (1.4.1)$$

但若有重根，(1.4.1) 不是完全解集。设 r_1 是 m 重根，则

$$L[e^{r_1 x}] = e^{r_1 x} (r - r_1)^m Q(r) \quad (1.4.2)$$

其中 Q 是 $n - m$ 次多项式。容易验证函数 $x^k e^{r_1 x}$ 是解，这给出 m 个解：

$$y = e^{r_1 x}, x e^{r_1 x}, x^2 e^{r_1 x}, \dots, x^{m-1} e^{r_1 x}. \quad (1.4.3)$$

(1.4.1) 和 (1.4.3) 中所有解的线性组合构成微分方程的通解。

等维 (Equidimensional) 方程

等维 (或欧拉) 方程在变换 $x \rightarrow ax$ 下不变, 其系数形式为 $p_j(x) = q_j/x^{n-j}$, 其中 q_j 与 x 无关。可通过变量替换

$$x = e^t, \quad x \frac{d}{dx} = \frac{d}{dt} \quad (1.4.4)$$

将等维方程转化为关于 t 的常系数方程。或者, 直接代试探函数 $y = x^r$ 得 $L[x^r] = P(r)x^{r-n}$, 这里 $P(r)$ 是 n 次多项式。因此当 $r_1, r_2, \dots$ 是 $P(r)$ 的不同根时, 解为:

$$y = x^{r_1}, x^{r_2}, \dots \quad (1.4.5)$$

当 $P(r)$ 有重根 r_1 时, 完全解集形式为:

$$y = x^{r_1}, x^{r_1} \ln x, x^{r_1} (\ln x)^2, \dots \quad (1.4.6)$$

$P(r)$ 的根称为指标指数 (indicial exponents)。

恰当 (Exact) 方程

恰当方程是低阶方程的导数: $Ly = \frac{d}{dx}(My) = 0$ 。此方程的解为 $My = c_1$ (所得方程不再齐次)。
积分因子 (Integrating factor) 是使微分方程乘以该函数后变为恰当方程的函数。

降阶法 (Reduction of Order)

是通过找到的解来简化线性微分方程以降为低阶的技术。设 $y_1(x)$ 是 $Ly = 0$ 的解, 寻求如下形式的其他线性无关解:

$$y(x) = u(x)y_1(x) \quad (1.4.7)$$

将 $y(x)$ 代入 $Ly = 0$ 得新方程 $Mu = 0$ 。此替换的精妙在于 Mu 不含 $p_0(x)u$ 项, 故 $Mu = 0$ 是关于 $v(x) = u'(x)$ 的 $n-1$ 阶线性齐次方程。

变换为已知方程

若其他方法均不适用, 可将微分方程变换为其他经典方程。

定义 1.4 (Airy 方程)

$$y'' = xy, \quad (1.4.8)$$

定义 1.5 (抛物柱面方程 (Weber-Hermite 方程))

$$y'' + \left(\nu + \frac{1}{2} - \frac{1}{4}x^2 \right) y = 0, \quad (1.4.9)$$

定义 1.6 (Bessel 方程)

$$y'' + \frac{1}{x}y' + \left(1 - \frac{\nu^2}{x^2}\right)y = 0, \quad (1.4.10)$$

1.5 非齐次线性函数

非齐次线性微分方程仅比齐次方程略复杂。这是因为 $Ly = f(x)$ 的任意两解之差满足 $Ly = 0$ 。因此, $Ly = f(x)$ 的通解是其任一特解与 $Ly = 0$ 的通解之和。

所有一阶线性非齐次方程均可解, 因为总存在仅依赖于 x 的积分因子。方程

$$y'(x) + p_0(x)y(x) = f(x) \quad (1.5.1)$$

的积分因子为 $I(x) = \exp \left[\int p_0(t) dt \right]$ 。乘以 $I(x)$ 得 $\frac{d}{dx} [I(x)y(x)] = f(x)I(x)$, 故 (1.5.1) 的解为:

$$y(x) = \frac{c_1}{I(x)} + \frac{1}{I(x)} \int f(t)I(t) dt \quad (1.5.2)$$

求解高阶非齐次线性方程有以下标准方法。

参数变异法 (Variation of Parameters)

当齐次方程可解时, 求解非齐次方程的难点在于寻找特解。参数变异法是确定特解的通用可靠方法, 可视为高阶降阶法。

以二阶方程为例。设 $y_1(x)$ 和 $y_2(x)$ 是齐次方程 $Ly = 0$ 的两个线性无关解, 其中 $L = d^2/dx^2 + p_1(x)d/dx + p_0(x)$ 。我们来求 $Ly = f(x)$ 的对称形式特解:

$$y(x) = u_1(x)y_1(x) + u_2(x)y_2(x) \quad (1.5.3)$$

其中 u_1, u_2 待定, 可施加约束简化方程:

$$u_1'(x)y_1(x) + u_2'(x)y_2(x) = 0 \quad (1.5.4)$$

对 (1.5.3) 求导两次并代入 $Ly = f(x)$, 结合 $Ly_1 = Ly_2 = 0$ 和 (1.5.4) 得:

$$u_1'(x)y_1'(x) + u_2'(x)y_2'(x) = f(x) \quad (1.5.5)$$

联立 (1.5.4) 和 (1.5.5) 解出:

$$\begin{aligned} u_1'(x) &= -\frac{f(x)y_2(x)}{W(x)}, \\ u_2'(x) &= \frac{f(x)y_1(x)}{W(x)} \end{aligned} \quad (1.5.6)$$

其中 $W(x) = W[y_1(x), y_2(x)]$ 为朗斯基行列式, 因 y_1, y_2 线性无关作为分母 W 非零。积分 (1.5.6) 即得特解:

$$y(x) = -y_1(x) \int \frac{f(t)y_2(t)}{W(t)} dt + y_2(x) \int \frac{f(t)y_1(t)}{W(t)} dt + c_1y_1(x) + c_2y_2(x) \quad (1.5.7)$$

格林函数 (Green's Function)

构造非齐次线性微分方程解有另一种通用方法，等价于参数变异法。该方法将解表示为格林函数的积分。定义格林函数需引入 Dirac 函数 $\delta(x-a)$ ：

定义 1.7

定义 Dirac 函数为满足如下两个性质的函数：

$$\delta(x-a) = 0, \quad x \neq a \quad (1.5.8a)$$

$$\int_{-\infty}^{\infty} \delta(x-a) dx = 1 \quad (1.5.8b)$$

此函数可视为单位脉冲：在 $x=a$ 处无限大的尖峰，且面积为 1。

命题 1.4

若 $f(x)$ 在 a 连续，则

$$\int_{-\infty}^{\infty} \delta(x-a) f(x) dx = f(a) \quad (1.5.9)$$

命题 1.5

δ 函数有如下多种表示：

$$\delta(x-a) = \lim_{\varepsilon \rightarrow 0} F_{\varepsilon}(x) \quad (1.5.10a)$$

$$\text{其中 } F_{\varepsilon}(x) = \begin{cases} 0, & x < a - \frac{1}{2}\varepsilon \\ 1/\varepsilon, & a - \frac{1}{2}\varepsilon \leq x \leq a + \frac{1}{2}\varepsilon \\ 0, & a + \frac{1}{2}\varepsilon < x \end{cases}$$

$$\text{或 } \delta(x-a) = \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{\pi[(x-a)^2 + \varepsilon^2]} \quad (1.5.10b)$$

$$\text{或 } \delta(x-a) = \lim_{\varepsilon \rightarrow 0^+} (\pi\varepsilon)^{-1/2} e^{-(x-a)^2/\varepsilon} \quad (1.5.10c)$$

$$\text{或 } \delta(x-a) = \lim_{L \rightarrow +\infty} \frac{1}{2\pi} \int_{-L}^L e^{i(x-a)t} dt \quad (1.5.10d)$$

δ 函数也可视为不连续函数的导数。设 $h(x-a)$ 为 Heaviside 阶跃函数：

$$h(x-a) = \begin{cases} 0, & x < a \\ \frac{1}{2}, & x = a \\ 1, & x > a \end{cases}$$

则

$$\delta(x-a) \equiv \frac{d}{dx} h(x-a) \quad (1.5.11)$$

定义 1.8

非齐次方程 $Ly = f(x)$ 的格林函数 $G(x, a)$ 满足微分方程

$$LG(x, a) = \delta(x - a) \quad (1.5.12)$$

已知 $G(x, a)$ 后, $Ly = f(x)$ 的解可表示为积分:

$$y(x) = \int_{-\infty}^{\infty} da f(a)G(x, a) \quad (1.5.13)$$

剩余问题是从 (1.5.12) 中解出 $G(x, a)$ 。当齐次方程 $Ly = 0$ 的解已知时, 这是简单的。以二阶方程为例:

$$LG(x, a) = \left[\frac{d^2}{dx^2} + p_1(x) \frac{d}{dx} + p_0(x) \right] G(x, a) = \delta(x - a) \quad (1.5.14)$$

设 $y_1(x), y_2(x)$ 为 $Ly = 0$ 的线性无关解。则当 $x \neq a$ 时, (1.9.1) 右边为零:

$$G(x, a) = A_1 y_1(x) + A_2 y_2(x), \quad x < a$$

$$G(x, a) = B_1 y_1(x) + B_2 y_2(x), \quad x > a$$

为关联 $x < a$ 和 $x > a$ 的格林函数解, 我们来论证 $G(x, a)$ 在 $x = a$ 连续且 $\partial G / \partial x$ 在 $x = a$ 有幅度为 1 的有限跳跃间断。格林函数方程 (1.9.1) 左侧最奇异项必为 $\partial^2 G / \partial x^2$, 进一步有 $\partial^2 G / \partial x^2 - \delta(x - a)$ 在 $x = a$ 处奇异度低于 δ 函数。积分 $\partial^2 G / \partial x^2 - \delta(x - a)$ 得 $\partial G / \partial x - h(x - a)$ 处处连续。因此 $\partial G / \partial x$ 在 $x = a$ 的间断性与函数 $h(x - a)$ 相同:

$$\lim_{\varepsilon \rightarrow 0^+} \left[\left. \frac{\partial G}{\partial x} \right|_{x=a+\varepsilon} - \left. \frac{\partial G}{\partial x} \right|_{x=a-\varepsilon} \right] = 1 \quad (1.5.15)$$

且因 $\partial G / \partial x$ 仅有有限跳跃间断, 其原函数 $G(x, a)$ 在 $x = a$ 连续, 这给出条件:

$$A_1 y_1(a) + A_2 y_2(a) = B_1 y_1(a) + B_2 y_2(a) \quad (1.5.16)$$

(1.5.15) 要求:

$$B_1 y_1'(a) + B_2 y_2'(a) - A_1 y_1'(a) - A_2 y_2'(a) = 1 \quad (1.5.17)$$

解出 $B_1 - A_1$ 和 $B_2 - A_2$:

$$B_1 - A_1 = -\frac{y_2(a)}{W[y_1(a), y_2(a)]} \quad (1.5.18)$$

$$B_2 - A_2 = \frac{y_1(a)}{W[y_1(a), y_2(a)]} \quad (1.5.19)$$

此式与 (1.5.6) 高度相似。令 $A_1 = A_2 = 0$ 并用 (1.5.18)、(1.5.19) 确定 B_1, B_2 得:

$$G(x, a) = \begin{cases} -\frac{y_2(a)y_1(x)}{W[y_1(a), y_2(a)]}, & x < a \\ \frac{y_1(a)y_2(x)}{W[y_1(a), y_2(a)]}, & x \geq a \end{cases} \quad (1.5.20)$$

代入 (1.5.13) 即得与参数变异法 (1.5.7) 相同结果。

当需对固定算子 L 和边界条件, 对不同 f 求解 $Ly = f$ 时, 格林函数法优于参数变异法, 这是因为格林函数可反复利用。齐次边界条件下分析尤简。

例 1.13 (格林函数解边值问题)

边值问题 $y'' = f(x)$ [$y(0) = 0, y(1) = 0$] 的格林函数满足:

$$\frac{\partial^2 G}{\partial x^2}(x, a) = \delta(x - a), \quad G(0, a) = 0, \quad \frac{\partial G}{\partial x}(1, a) = 0$$

解为:

$$G(x, a) = \begin{cases} -x, & x < a \\ -a, & x \geq a \end{cases} \quad (0 < a < 1)$$

故对任意 $f(x)$, 解为 $y(x) = \int_0^1 G(x, a)f(a) da$ ($0 \leq x \leq 1$)。积分限为 $[0, 1]$ 而非 $\mathbb{R}$ 是因为边界条件限制在有限区间。

例 1.14 (格林函数解边值问题)

边值问题 $y'' - y = f(x)$ [$y(\pm\infty) = 0$] 的格林函数满足:

$$\frac{\partial^2 G}{\partial x^2} - G(x, a) = \delta(x - a), \quad G(0, a) = 0, \quad G(\pm\infty, a) = 0$$

解为:

$$G(x, a) = -\frac{1}{2}e^{-|x-a|}$$

故对任意 $f(x)$, 解为 $y(x) = -\frac{1}{2} \int_{-\infty}^{\infty} e^{-|x-a|} f(a) da$ 。

降阶法

例 1.15 (非齐次方程的降阶法)

齐次方程 $a(x)y'' + xy' - y = 0$ 的一个解为 $y_1(x) = x$ 。为用降阶法解非齐次方程 $a(x)y'' + xy' - y = f(x)$, 设解形如 $y(x) = xu(x)$ 。代入得关于 $u'(x)$ 的一阶方程:

$$xa(x)u'' + (2a(x) + x^2)u' = f(x)$$

待定系数法 (Method of Undetermined Coefficients)

此法本质上是系统猜测, 通常适用于常系数方程且 $f(x)$ 为 e^x 、 $\sin x$ 、 $\cos x$ 与多项式之组合, 或等维方程且 $f(x)$ 为多项式。

例 1.16

(a) 解 $y''' + y = e^x \sin x$, 猜测 $y = ae^x \sin x + be^x \cos x$, 代入确定系数 $a = -\frac{1}{5}, b = -\frac{2}{5}$ 。

(b) 解 $y'' - y = e^x$, 因 e^x 已满足齐次方程, 猜特解 $y = axe^x$, 得 $a = \frac{1}{2}$ 。

(c) 解 $y'' - \frac{y}{x^2} = x^4 + x^3$, 猜多项式特解 $y = ax^6 + bx^5$, 得 $a = \frac{1}{25}, b = \frac{1}{15}$ 。

(d) 解 $y'' + xy' + 2y = 1$, 猜特解 $y = a$, 得 $y = \frac{1}{2}$ 。

(e) 解 $y'' + xy' + 2y = x^4$, 猜特解 $y = ax^4 + bx^2 + c$, 得 $y = \frac{x^4}{6} - \frac{x^2}{2} + \frac{1}{4}$ 。

1.6 一阶非线性微分方程

尽管多数非线性微分方程难以显式求解, 但识别可解方程类型并掌握相应解法至关重要。一阶方程的常规策略是通过代换转化为线性或恰当方程。

伯努利方程

标准形式为:

$$y' = a(x)y + b(x)y^P \quad (1.6.1)$$

其中 $a(x), b(x)$ 为任意函数, P 为任意常数。当 $P = 0$ 时为线性方程, $P = 1$ 时为可分方程。解法为利用代换

$$y(x) = [u(x)]^{1-P} \quad (1.6.2)$$

得线性方程:

$$u'(x) = (1-P)a(x)u(x) + (1-P)b(x) \quad (1.6.3)$$

Riccati 方程

标准形式为:

$$y'(x) = a(x)y^2(x) + b(x)y(x) + c(x) \quad (1.6.4)$$

当 $a(x) = 0$ 时为线性方程, $c(x) = 0$ 时为伯努利方程。解法为利用代换

$$y(x) = -\frac{w'(x)}{a(x)w(x)} \quad (1.6.5)$$

转化为二阶线性方程:

$$w''(x) - \left[\frac{a'(x)}{a(x)} + b(x) \right] w'(x) + a(x)c(x)w(x) = 0 \quad (1.6.6)$$

此变换可逆。这表明方程与二阶齐次线性方程等价, 但我们目前并未得到二阶线性微分方程的通解。回到开头, 若已知特解 $y_1(x)$, 令

$$y(x) = y_1(x) + u(x) \quad (1.6.7)$$

则转化为伯努利方程:

$$u'(x) = [b(x) + 2a(x)y_1(x)]u(x) + a(x)u^2(x)$$

例 1.17 (因式分解法)

Riccati 方程源于二阶线性方程的因式分解尝试, 希望将 n 阶线性算子分解为:

$$Ly = \left[\frac{d}{dx} + a_1(x) \right] \left[\frac{d}{dx} + a_2(x) \right] \cdots \left[\frac{d}{dx} + a_n(x) \right] y(x) \quad (1.6.8)$$

要注意这些因子既不是可交换的:

$$\left(\frac{d}{dx} + \frac{1}{x} \right) \left(\frac{d}{dx} + x \right) \neq \left(\frac{d}{dx} + x \right) \left(\frac{d}{dx} + \frac{1}{x} \right)$$

也不是唯一的:

$$y'' - y = \left(\frac{d}{dx} + 1 \right) \left(\frac{d}{dx} - 1 \right) y = \left(\frac{d}{dx} + \tanh x \right) \left(\frac{d}{dx} - \tanh x \right) y$$

若分解成功, 求解 $Ly = f(x)$ 转化为一系列一阶线性方程:

$$\begin{aligned} w'_1 + a_1 w_1 &= f(x) \\ w'_2 + a_2 w_2 &= w_1(x) \\ &\vdots \\ y' + a_n y &= w_{n-1}(x) \end{aligned} \quad (1.6.9)$$

最终解含 n 个积分常数。

当然这里的难点在于分解算子。以二阶算子 $L = \frac{d^2}{dx^2} + p_1(x) \frac{d}{dx} + p_0(x)$ 为例:

$$Ly = \left[\frac{d}{dx} + a_1(x) \right] \left[\frac{d}{dx} + a_2(x) \right] y$$

比较系数消去 a_1 得 a_2 的 Riccati 方程:

$$a'_2 = a_2^2 - p_1 a_2 + p_0$$

其难度与原二阶方程相当。代换 $a_2(x) = -w'(x)/w(x)$ 后恰好还原为原方程 $Lw = 0$ 。

恰当方程

一阶恰当方程可表示为:

$$M[x, y(x)] + N[x, y(x)]y'(x) = \frac{d}{dx} f[x, y(x)] = 0$$

其解为 $f[x, y(x)] = c_1$ 。恰当性的充要条件为:

$$\frac{\partial M}{\partial y}(x, y) = \frac{\partial N}{\partial x}(x, y) \quad (1.6.10)$$

解存在性定理表明每个非线性方程均可通过积分因子化为恰当方程, 但寻找积分因子可能极困难。设方程 $a(x, y) + b(x, y)y'(x) = 0$, 乘以待定积分因子 $I(x, y)$ 后要求满足 (1.6.10), 得 I 的线性

偏微分方程:

$$a(x, y) \frac{\partial I}{\partial y} - b(x, y) \frac{\partial I}{\partial x} = \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right) I$$

此方程的标准解法(特征线法)的特征曲线方程为 $dy/dx = -a/b$, 恰为原微分方程!

代换法

有时可通过代换将非线性方程转化为可直接求解的形式。线性代换因易识别而应优先考虑。

例 1.18

- (a) 代换 $u = x + y$ 使 $y' = \cos(x + y)$ 可分: $u' = 1 + \cos u$ 。
- (b) 线性变换 $x = av + bw + c, y = dv + ew + f$ 在适当选择参数后可将 $y' = (Ax + By + C)/(Dx + Ey + F)$ 转化为关于 $w(v)$ 的可分方程。
- (c) 方程 $y' = (y - x)/(y + x)$ 的解为对数螺线。使用极坐标代换:

$$x = r \cos \theta, \quad dx = \cos \theta dr - r \sin \theta d\theta,$$

$$y = r \sin \theta, \quad dy = \sin \theta dr + r \cos \theta d\theta,$$

得可分方程 $dr/d\theta = -r$, 解为 $r = c_1 e^{-\theta}$ 。

- (d) 代换 $u = x + y$ 将 $y' = y/x + 1/(y + x)$ 转化为伯努利方程: $u' = u/x + 1/u$ 。
- (e) 对形如 $y' = F(y/x)$ 的方程, 代换 $u = y/x$ 得可分方程: $u' = [F(u) - u]/x$ 。
- (f) 代换 $u = xy$ 将 $y' = \cos(xy)$ 转化为 $u' = x \cos u + u/x$, 其求解难度未降低。乘法型代换 ($u = xy, u = x^2 y$ 等) 通常对 $y' = F(xy)$ 类方程无效。

1.7 高阶非线性微分方程

本节简要总结求解 n 阶非线性微分方程的方法, 大多数方法试图降低微分方程的阶数。

自治方程 (Autonomous Equations)

定义 1.9

自治方程是指自变量不显式出现的方程, 如 $y'' + y' + y = 0, y''' + y'' = 0, yy' = y''y''$ 。自治方程在平移变换 $x \rightarrow x + a$ 下保持不变。

一个 n 阶自治方程总可以用一个 $(n - 1)$ 阶非自治方程替代。标准技巧是将 $u \equiv y'(x)$ 表示为 y 的函数, 并找出 u 及其导数满足的方程。因此, 我们令

$$y'(x) = u(y), \tag{1.7.1a}$$

$$y''(x) = \frac{du}{dx} = \frac{du}{dy} \frac{dy}{dx} = u'(y)u(y), \tag{1.7.1b}$$

$$y'''(x) = \frac{d}{dx}[u'(y)u(y)] = u(y)[u'(y)]^2 + [u(y)]^2 u''(y) \tag{1.7.1c}$$

以此类推得到关于自变量 y 的新微分方程，且其中 u 关于 y 的最高阶导数总是比原方程中 y 关于 x 的最高阶导数低一阶。

关于 x 的齐次方程 (Equidimensional-in- x Equations)

定义 1.10

如果一个方程在尺度变换 $x \rightarrow ax$ 下保持不变，则称该方程是关于 x 齐次的。

所有关于 x 齐次的方程都可以转化为同阶的自治方程。所需的变量变换已在 (1.4.4) 中给出：

$$x = e^t, \quad (1.7.2a)$$

$$\frac{d}{dx} = e^{-t} \frac{d}{dt}, \quad (1.7.2b)$$

$$x^2 \frac{d^2}{dx^2} = \frac{d^2}{dt^2} - \frac{d}{dt}. \quad (1.7.2c)$$

尺度不变方程 (Scale-Invariant Equations)

定义 1.11

如果存在 p 使得尺度变换

$$x \rightarrow ax, \quad y \rightarrow a^p y \quad (1.7.3)$$

保持原微分方程不变，则该微分方程是尺度不变的。

识别微分方程为尺度不变是求解的重要进展，因为所有尺度不变方程可通过代换

$$y(x) = x^p v(x) \quad (1.7.4)$$

转化为关于 x 的齐次方程。

关于 y 的齐次方程 (Equidimensional-in- y Equations)

定义 1.12

若方程在尺度变换 $y \rightarrow ay$ 下不变，则称为关于 y 齐次。

此类方程可通过变换

$$y(x) = e^{u(x)} \quad (1.7.5)$$

将方程阶数降低一阶。

1.8 特征值问题

特征值问题是一类边界值问题，仅当问题中的参数 E 取特定值（称为特征值）时才存在非平凡解。

例 1.19 (简单特征值问题)

考虑边值问题

$$y'' + Ey = 0, \quad y(0) = 0, \quad y(1) = 0. \quad (1.8.1)$$

对任意参数 E , 存在平凡解 $y(x) = 0$ ($0 \leq x \leq 1$)。但当 E 取特定值时, (1.8.1) 存在额外的非零解:

$$E = (n\pi)^2, \quad y(x) = A_n \sin(n\pi x), \quad n = 1, 2, 3, \dots \quad (1.8.2)$$

其中 A_n 为任意常数, $E = \pi^2, 4\pi^2, 9\pi^2, \dots$ 称为特征值, 对应的非平凡解 $\sin \pi x, \sin 2\pi x, \sin 3\pi x, \dots$ 称为特征函数。且易证 (1.8.1) 无其他特征值。

对齐次线性边值问题, 当 E 为特征值时, 解不唯一: 除 $y(x) = 0$ 外, 存在特征函数的任意常数倍解。反之, 当 E 非特征值时, 平凡解唯一。

例 1.20 (另一简单特征值问题)

考虑特征值问题

$$y'' + Ey = 0, \quad y(0) = 0, \quad y'(1) = 0. \quad (1.8.3)$$

对任意参数 E , 存在平凡解 $y(x) = 0$ 。满足 $y(0) = 0$ 的通解为 $y(x) = A \sin(\sqrt{E}x)$, 由 $y'(1) = 0$ 得 $A\sqrt{E} \cos \sqrt{E} = 0$, 故特征值为 $E = \left(\frac{\pi}{2}\right)^2, \left(\frac{3\pi}{2}\right)^2, \left(\frac{5\pi}{2}\right)^2, \dots$

例 1.21 (无限域上的特征值问题)

特征值问题

$$y'' + (E - x^2)y = 0, \quad -\infty < x < \infty, \quad (1.8.4a)$$

$$y(x) \rightarrow 0 \quad \text{当} \quad |x| \rightarrow \infty \quad (1.8.4b)$$

称为量子谐振子 (**quantum harmonic oscillator**)。对任意 E 存在平凡解 $y(x) = 0$, 但对特定 E 存在非平凡解。(1.8.4) 的特征值为 $E = n + \frac{1}{2}$, 对应的特征函数为

$$y_n(x) = A \cdot \text{He}_n(x) e^{-x^2/4}, \quad (1.8.5)$$

其中 A 为任意常数, $\text{He}_n(x)$ 为 n 阶 Hermite 多项式, $\text{He}_0(x) = 1, \text{He}_1(x) = x, \text{He}_2(x) = x^2 - 1, \dots$ 。

例 1.22 (含超越特征值的问题)

考虑特征值问题

$$y'' + (E - x)y = 0 \quad (0 < x < \infty), \quad y(0) = 0, \quad y(\infty) = 0. \quad (1.8.6)$$

后文中将证明: 满足 $x \rightarrow \infty$ 时 $y \rightarrow 0$ 的通解为

$$y(x) = A \cdot \text{Ai}(x - E), \quad (1.8.7)$$

其中 $\text{Ai}(z)$ 为 Airy 函数 (见 (1.4.8))。边界条件 $y(0) = 0$ 给出特征值条件 $\text{Ai}(-E) = 0$, 这是超越函数, 其零点可数值计算, 得到特征值序列:

$$E_0 \approx 2.338, \quad E_1 \approx 4.088, \quad E_2 \approx 5.521, \quad E_3 \approx 6.787, \quad E_4 \approx 7.944, \quad E_5 \approx 9.023, \dots$$

例 1.23

考虑特征值问题

$$y'' + (E + v \sec^2 x)y = 0, \quad y \rightarrow 0 \quad \text{当} \quad |x| \rightarrow \infty. \quad (1.8.8)$$

仅存在有限个离散特征值: $E = -\frac{1}{4}(2n+1-\sqrt{1+4v})^2$ ($0 \leq n \leq V$), 其中 $V = (\sqrt{1+4v}-1)/2$ 且 n 为整数。

例 1.24 (Sturm-Liouville 特征值问题)

特征值问题

$$a(x)y''(x) + b(x)y'(x) + c(x)y(x) + d(x)Ey(x) = 0, \quad y(\alpha) = y(\beta) = 0 \quad (1.8.9)$$

总可化为 Sturm-Liouville 标准形式:

$$\frac{d}{dx} \left[p(x) \frac{dy}{dx} \right] + q(x)y + Er(x)y = 0, \quad y(\alpha) = y(\beta) = 0 \quad (1.8.10)$$

方程 (1.8.10) 称为 Sturm-Liouville 问题。其性质为: 当

$$p(x) > 0, \quad q(x) \leq 0, \quad r(x) > 0, \quad \alpha \leq x \leq \beta$$

时, 存在无限多个实数正特征值 $E = E_0, E_1, E_2, \dots$ 。此外, 对应特征值 E_n 的特征函数 $y_n(x)$ ($n = 0, 1, 2, \dots$) 可关于权函数 $r(x)$ 正交归一化:

$$\int_{\alpha}^{\beta} r(x)y_n(x)y_m(x)dx = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases}$$

例 1.25 (薛定谔特征值问题)

若选取满足 $|x| \rightarrow \infty$ 时 $V(x) \rightarrow +\infty$ 的正函数 $V(x)$, 则薛定谔特征值问题

$$-y''(x) + V(x)y(x) = Ey(x), \quad y(\pm\infty) = 0 \quad (1.8.11)$$

有无限多个实数正特征值 $E_0, E_1, E_2, \dots$ 。物理上, 这些特征值表示势场 $V(x)$ 中粒子的允许能级。(1.8.4) 的问题是特例。仅对特殊 $V(x)$ 存在 (1.8.11) 的显式解。

1.9 复平面中的微分方程

此前我们使用字母 x 表示微分方程的自变量, 并假设 x 为实数。但在后续分析中, 我们将关注复平面中的微分方程性质, 此时自变量记为 z 。当微分方程从实轴 x 推广到复平面 z 时, 所有导数必须按复导数解释。这导出一个重要结论 (实变函数不适用): 在复导数 $y'(z)$ 存在的区域, 函数 $y(z)$ 是解析的, 该结论限制了可在复平面中构造的微分方程类型。

例 1.26 (一阶微分方程)

微分方程 $dy/dx = |x|$ 在实轴上有意义, 其解为 $y(x) = \frac{1}{2}|x|^2 + c$ 。但方程 $dy/dz = |z|$ 在复平面中无意义, 因为 $y'(z)$ 要求 $y(z)$ 解析, 而 $|z|$ 处处不解析。

一般地, 为使复微分方程 $dy/dz = g(z)$ 有意义, 需存在某个复区域使 $g(z)$ 解析。

例 1.27 (二阶微分方程)

若微分方程 $y'' + p_1(z)y' + p_0(z)y = 0$ 在某个区域存在两个线性无关解, 则 $p_0(z)$ 和 $p_1(z)$ 除可能孤立点外必须是解析函数。设 y_1 和 y_2 为两个线性无关解, 因其复导数存在故解析, 则

$$p_0(z) = \frac{y_1''y_2' - y_2''y_1'}{y_1'y_2 - y_1y_2'}, \quad p_1(z) = \frac{y_2''y_1 - y_1''y_2}{y_1'y_2 - y_1y_2'}, \quad (1.9.1)$$

其中分母仅在孤立点为零。

具有非解析系数的二阶复微分方程可能有解, 但解可能不含两个独立任意参数。例如方程 $y'' - |z|y' + (|z| - 1)y = 0$ 有解 $y = ce^z$, 但无其他解。

除上述限制外, 实微分方程与复微分方程无本质区别。本章介绍的求解方法同样适用于复微分方程。

2 差分方程

差分方程的研究基于离散微积分概念，旨在总结其求解的基本方法。其常用于计算递归定义的量，例如：Taylor 级数或 Fourier 展开的第 n 项系数，通过余子式展开的 $n \times n$ 矩阵的行列式。在数值分析中，差分方程频繁出现于用离散系统逼近连续系统的场景。

2.1 差分

定义 2.1

定义在整数集上的函数将每个整数 n 映射到数 a_n ，函数 a_n 的离散导数定义为：

$$Da_n \equiv a_{n+1} - a_n.$$

二阶导数 D^2a_n 是导数的导数：

$$D^2a_n = (a_{n+2} - a_{n+1}) - (a_{n+1} - a_n) = a_{n+2} - 2a_{n+1} + a_n.$$

k 阶导数通过连续取 k 次导数定义。

定义 2.2

函数 a_n 的离散原函数（或离散积分） b_n 定义为和式：

$$b_n = \sum_{j=n_0}^n a_j.$$

例 2.1

连续函数 $f(x) = x^k$ 对应的离散函数为：

$$f_n = n(n+1) \cdots (n+k-1)$$

- 连续导数： $f'(x) = kx^{k-1}$
- 离散导数：

$$f_{n+1} - f_n = k \cdot (n+1)(n+2) \cdots (n+k-1)$$

含 $k-1$ 个因子。

2.2 初等差分方程

定义 2.3

同 1.1 节 N 阶微分方程的定义, N 阶差分方程的自然定义为:

$$a_n^{(N)} = F[n, a_n, a_n^{(1)}, \dots, a_n^{(N-1)}]$$

其中 $a_n^{(j)}$ 是 a_n 的 j 阶离散导数。为避免 $a_n, a_{n+1}, \dots$ 的重复, 采用更简洁的常规定义:

$$a_{n+N} = G[n, a_n, a_{n+1}, \dots, a_{n+N-1}] \quad (2.2.1)$$

此形式可显式转换为离散导数形式。

N 阶差分方程 (2.2.1) 的通解含 N 个独立参数 $c_1, c_2, \dots, c_N$ 。

例 2.2 (阶乘函数)

一阶线性齐次差分方程 $a_{n+1} = na_n$ 的解为:

$$a_n = c_1(n-1)!$$

其中常数 c_1 为求和常数。

例 2.3 (伽马函数 $\Gamma(z)$)

Gamma 函数是最广泛使用的高等超越函数。当 $z = n$ 为正整数时, $\Gamma(n) = (n-1)!$ 。对 $\operatorname{Re} z > 0$, 其积分定义为:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt \quad (2.2.2)$$

其满足差分方程:

$$\Gamma(z+1) = z\Gamma(z) \quad (2.2.3)$$

此差分方程可将 $\Gamma(z)$ 解析延拓至 $\operatorname{Re} z \leq 0$ 。 $\Gamma(z)$ 是复平面上的单值解析函数, 在 $z = 0, -1, -2, \dots$ 处有单极点。

例 2.4 (一阶线性齐次方程)

方程 $a_{n+1} = p(n)a_n$ 的解为乘积形式:

$$a_n = a_1 \prod_{j=1}^{n-1} p(j) \quad (2.2.4)$$

其中 a_1 为任意常数。类比微分方程 $y' = p(x)y$ 的解为:

$$y(x) = y(0) \exp\left(\int_0^x p(t) dt\right)$$

例 2.5 (一阶线性非齐次方程)

解方程 $a_{n+1} = p(n)a_n + q(n)$ 需用求和因子:

$$\left[\prod_{j=1}^n p(j) \right]^{-1}$$

乘两边得:

$$\frac{a_{n+1}}{\prod_{j=1}^n p(j)} - \frac{a_n}{\prod_{j=1}^{n-1} p(j)} = \frac{q(n)}{\prod_{j=1}^n p(j)}$$

左式为离散导数。从 1 到 $n-1$ 求和得解:

$$a_n = \prod_{j=1}^{n-1} p(j) \left[\sum_{k=1}^{n-1} \frac{q(k)}{\prod_{j=1}^k p(j)} + a_1 \right], \quad n \geq 2$$

例 2.6 (带边界条件的一阶方程)

要求方程 $a_{n+1} = \frac{n}{n+1}a_n + \frac{n-1}{(n+1)^2(n+2)(n+3)}$ 在 $n=0$ 处有界。乘求和因子 $n+1$ 后求和得通解:

$$a_n = \frac{a_1 + \frac{1}{6}}{n} - \frac{1}{(n+1)(n+2)} \quad (n > 0)$$

由 a_0 有界得 $a_1 = -\frac{1}{6}$, 故唯一解:

$$a_n = -\frac{1}{(n+1)(n+2)} \quad (n \geq 0)$$

类比微分方程 $y' + (\cot x)y = -1$ 在 $x=0$ 处有界的解为:

$$y(x) = \frac{\cos x - 1}{\sin x}$$

非线性差分方程通常难解, 但以下特例可求显式解。

例 2.7 ($a_{n+1} = a_n^2$)

取对数得线性方程:

$$\ln a_{n+1} = 2 \ln a_n \implies \ln a_n = 2^{n-1} \ln a_1 \implies a_n = a_1^{2^{n-1}}$$

例 2.8 ($a_{n+2} = a_{n+1}^2/a_n$)

取对数得:

$$\ln a_{n+2} - 2 \ln a_{n+1} + \ln a_n = 0$$

二阶离散导数为零。两次求和得:

$$\ln a_n = c_1 n + c_2 \implies a_n = e^{c_1 n + c_2}$$

例 2.9 ($a_{n+1} = 2a_n^2 - 1$)

$$a_n = \begin{cases} \cos(2^{n-1} \cos^{-1} a_1), & |a_1| \leq 1 \\ \cosh(2^{n-1} \cosh^{-1} a_1), & |a_1| \geq 1 \end{cases}$$

2.3 齐次线性差分方程

N 阶齐次线性差分方程的一般形式为:

$$a_{n+N} + p_{N-1}(n)a_{n+N-1} + \cdots + p_0(n)a_n = 0 \quad (2.3.1)$$

其中 $p_0, p_1, \cdots, p_{N-1}$ 为 n 的任意函数。与微分方程类似, (2.3.1) 的通解是其 N 个线性无关解的任意线性组合。

常数系数方程

当系数 p_j 与 n 无关时, 称为常数系数方程。借鉴微分方程解法, 指数形式解 $a_n = r^n$, 代入 (2.3.1) 得特征方程:

$$r^N + p_{N-1}r^{N-1} + \cdots + p_0 = 0 \quad (2.3.2)$$

若特征根 $r_1, r_2, \cdots, r_N$ 互异, 则通解为 $a_n = \sum_{k=1}^N c_k r_k^n$ 。若 r_1 为 m 重根, 则对应解为 $r_1^n, nr_1^n, \cdots, n^{m-1}r_1^n$ 。

线性无关性与 Wronskian

函数集 $\{a_n, b_n, \cdots\}$ 在区间 $n_1 \leq n \leq n_2$ 线性无关当且仅当方程 $k_1 a_n + k_2 b_n + \cdots = 0$ 仅有零解。

定义 2.4

离散 Wronskian (Casoratian) 定义为:

$$W_n = \det \begin{pmatrix} a_n & b_n & \cdots \\ a_{n+1} & b_{n+1} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad (2.3.3)$$

命题 2.1

函数集线性相关当且仅当对所有 n 有 $W_n \equiv 0$ 。

Wronskian 的递推关系

对 N 阶齐次线性差分方程, 解的 Wronskian 满足一阶方程 (类比 Abel 公式):

例 2.10 (二阶方程情形)

设 a_n, b_n 为方程

$$a_{n+2} + p_1(n)a_{n+1} + p_0(n)a_n = 0 \quad (2.3.4)$$

的解, 则:

$$W_{n+1} = W_n p_0(n) \implies W_n = W_{n_0} \prod_{j=n_0}^{n-1} p_0(j) \quad (2.3.5)$$

当 $p_0(n) \neq 0$ 时, 非零 Wronskian 保持非零。

初值问题与边值问题

N 阶齐次线性差分方程的通解含 N 个参数, 可通过两种方式确定:

1. **初值问题:** 在定义域 $n_1 \leq n \leq n_2$ 内指定单点 n_0 的 N 个离散导数值 (或 N 个连续函数值)。解存在唯一的充要条件:

- n_0 处 Wronskian 非零
- $p_0(n)$ 在 $n_1 \leq n \leq n_0$ 非零

2. **边值问题:** 在定义域内任意点指定 N 个函数值或导数值。解的存在唯一性无简单判别准则。

降阶法

若已知方程的一个特解 A_n , 可通过代换

$$a_n = A_n x_n \quad (2.3.6)$$

将方程降阶。

欧拉方程

回顾 1.4 节, 除常数系数微分方程外, 还有一类可求解的方程称为欧拉方程。欧拉微分方程的形式为:

$$x^N \frac{d^N}{dx^N} y + p_{N-1} x^{N-1} \frac{d^{N-1}}{dx^{N-1}} y + \cdots + p_1 x \frac{dy}{dx} + p_0 y = 0 \quad (2.3.7)$$

其中 $p_0, p_1, \dots, p_{N-1}$ 是常数。为构造欧拉方程的差分方程类比, 回顾 2.1 节例 2.1 中 x 幂的离散类比: $x \leftrightarrow n$, $x^2 \leftrightarrow n(n+1)$, $x^3 \leftrightarrow n(n+1)(n+2)$, $\dots$, $x^N \leftrightarrow n(n+1)(n+2) \cdots (n+N-1)$ 。我们利用 Gamma 函数重写:

$$\frac{\Gamma(n+N)}{\Gamma(n)} = n(n+1)(n+2) \cdots (n+N-1) \quad (2.3.8)$$

定义 2.5 (离散欧拉方程)

定义离散欧拉方程为:

$$\frac{\Gamma(n+N)}{\Gamma(n)} a_n^{(N)} + p_{N-1} \frac{\Gamma(n+N-1)}{\Gamma(n)} a_n^{(N-1)} + \cdots + p_1 n a_n^{(1)} + p_0 a_n = 0 \quad (2.3.9)$$

现回顾欧拉微分方程 (2.3.7) 的解法, 取代换

$$y(x) = x^r \quad (2.3.10)$$

将方程 (2.3.7) 化为关于 r 的方程:

$$r(r-1)(r-2)\cdots(r-N+1) + p_{N-1}r(r-1)(r-2)\cdots(r-N+2) + \cdots + p_1r + p_0 = 0. \quad (2.3.11)$$

为求解离散欧拉方程 (2.3.9), 需对实数 r 代换 x^r 的离散类比。定义 x^r 的离散类比为 $\Gamma(n+r)/\Gamma(n)$ 。其一阶离散导数为:

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

此即 rx^r/x 的离散类比。二阶导数为: $r(r-1)\frac{\Gamma(n+r)}{\Gamma(n)n(n+1)}$, 对应 $r(r-1)x^r/x^2$ 。 N 阶导数为:

$$\left[\frac{r(r-1)\cdots(r-N+1)}{\Gamma(n+N)/\Gamma(n)} \right] \left[\frac{\Gamma(n+r)}{\Gamma(n)} \right],$$

对应 $r(r-1)\cdots(r-N+1)x^r/x^N$ 。将 $a_n = \Gamma(n+r)/\Gamma(n)$ 代入 (2.3.9) 得关于 r 的多项式方程:

$$r(r-1)\cdots(r-N+1) + p_{N-1}r(r-1)\cdots(r-N+2) + \cdots + p_1r + p_0 = 0. \quad (2.3.12)$$

此方程的根决定 (2.3.9) 的解。

例 2.11 (离散欧拉方程求解)

(a) 求解欧拉方程:

$$n(n+1)a_n^{(2)} + na_n^{(1)} - \frac{1}{4}a_n = 0 \quad (2.3.13)$$

代换 $a_n = \Gamma(n+r)/\Gamma(n)$ 得特征方程: $r^2 - \frac{1}{4} = 0$ 根为 $r_1 = \frac{1}{2}$, $r_2 = -\frac{1}{2}$ 。通解为:

$$a_n = c_1 \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)} + c_2 \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n)}$$

(b) 求解欧拉方程:

$$n(n+1)a_n^{(2)} + \frac{1}{4}a_n = 0 \quad (2.3.14)$$

同上代换得特征方程: $r^2 - r + \frac{1}{4} = 0$ 有二重根 $r = \frac{1}{2}$, 仅得一个特解:

$$a_n = \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}$$

设 $a_n = \varphi(n) \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)}$, 通过降阶法得: $\varphi(n) = \sum_{k=1}^n \frac{1}{k-\frac{1}{2}}$, 通解为:

$$a_n = c_1 \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)} + c_2 \varphi(n) \frac{\Gamma(n+\frac{1}{2})}{\Gamma(n)} \quad (2.3.15)$$

生成函数法

若差分方程的解 a_n 是 $F(x)$ Taylor 展开的第 n 项系数, 则 $F(x)$ 可能满足易解的微分方程。

例 2.12 (生成函数法求解差分方程)

解差分方程:

$$(n+1)(n+2)a_{n+2} - 2(n+1)a_{n+1} - 3a_n = 0, \quad n \geq 0 \quad (2.3.16)$$

初值条件: $a_0 = 2, a_1 = 2$ 。定义生成函数: $F(x) = \sum_{n=0}^{\infty} a_n x^n$, 逐项求导得:

$$F'(x) = \sum_{n=0}^{\infty} (n+1)a_{n+1}x^n, \quad F''(x) = \sum_{n=0}^{\infty} (n+2)(n+1)a_{n+2}x^n$$

代入 (2.3.16) 得微分方程:

$$F''(x) - 2F'(x) - 3F(x) = 0$$

初值条件: $F(0) = 2, F'(0) = 2$ 。解得:

$$F(x) = e^{3x} + e^{-x}$$

由 Taylor 公式得到

$$a_n = \frac{1}{n!} \frac{d^n}{dx^n} F(x) \Big|_{x=0} = \frac{3^n + (-1)^n}{n!}$$

例 2.13 (降阶的生成函数法)

二阶差分方程:

$$a_{n+2} = a_{n+1} + (n+1)a_n, \quad a_0 = a_1 = 1 \quad (2.3.17)$$

定义生成函数:

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n x^n}{n!}$$

将 (2.3.17) 乘以 $\frac{x^{n+1}}{(n+1)!}$ 后求和得微分方程:

$$F'(x) - 1 = (F(x) - 1) + xF(x) \implies F'(x) = (x+1)F(x)$$

初值条件 $F(0) = 1$, 解得:

$$F(x) = e^{x+\frac{x^2}{2}}$$

最终解为:

$$a_n = \frac{d^n}{dx^n} e^{x+\frac{x^2}{2}} \Big|_{x=0}$$

特征值问题

与微分方程类似, 差分方程也存在边值问题以确定特征值和特征函数。

例 2.14

考虑差分方程:

$$(n+8)(n+2)a_{n+2} - (n+1)(2n+11)a_{n+1} + (n+1)(n+2)a_n = Ea_{n+1} \quad (2.3.18)$$

满足边界条件:

$$a_{-1} < \infty \quad (2.3.19)$$

$$S = \sum_{n=0}^{\infty} a_n < \infty \quad (2.3.20)$$

仅当参数 E 取特定值 (特征值) 时, 方程组 (2.3.18)-(2.3.20) 有解, 下面我们来逐步分析。

代入 $n = -1$ 得: $7a_1 = Ea_0 \implies a_1 = \frac{E}{7}a_0$, 这表明以 a_0, a_1 为初值, 序列由递推唯一确定。

不妨 $a_0 = 1$, 则序列完全由 E 决定。

在 $n \rightarrow \infty$ 时有渐近形式:

$$a_n = c_1(E)n^{-3+\sqrt{2+E}} \text{ 或 } c_2(E)n^{-3-\sqrt{2+E}} \quad (2.3.21)$$

当 $E \geq 2$ 时, 仅当 $c_1(E) = 0$ 时级数收敛; 当 $E < 2$ 时, 级数可能收敛。

对 (2.3.18) 从 $n = -1$ 到 ∞ 求和得

$$2 \sum_{n=0}^{\infty} a_n = 2S \quad (2.3.22)$$

故有 $E = 2$ 或 $S = 0$ 。

• $E = 2$ 的特征函数:

$$a_n = \frac{720}{(n+2)(n+3)(n+4)(n+5)(n+6)} \quad (2.3.23)$$

• $E = -1$ 的特征函数:

$$a_n = \frac{(1 - 3n/2) \cdot 720}{(n+2)(n+3)(n+4)(n+5)(n+6)} \quad (2.3.24)$$

• $E = -2$ 的特征函数:

$$a_n = \frac{(n^2 - 13n + 6) \cdot 120}{(n+2)(n+3)(n+4)(n+5)(n+6)} \quad (2.3.25)$$

不同特征值 $E_a \neq E_b$ 对应的特征函数 a_n, b_n 满足正交关系:

$$\sum_{n=0}^{\infty} (n+2)(n+3)(n+4)(n+5)(n+6)a_nb_n = 0 \quad (2.3.26)$$

可通过直接计算验证 (2.3.23)-(2.3.25) 满足此条件。

2.4 非齐次线性差分方程

非齐次线性差分方程的通解由特解与对应齐次方程通解的线性组合构成。求解方法与微分方程类似。

参数变异法

参数变异法用于构造 N 阶非齐次差分方程

$$a_{n+N} + p_{N-1}(n)a_{n+N-1} + \cdots + p_1(n)a_{n+1} + p_0(n)a_n = q(n) \quad (2.4.1)$$

的解，其解的基为相关联的齐次方程

$$a_{n+N} + p_{N-1}(n)a_{n+N-1} + \cdots + p_1(n)a_{n+1} + p_0(n)a_n = 0 \quad (2.4.2)$$

的 N 个线性无关解。此处展示二阶方程的求解过程。设 A_n 和 B_n 是方程

$$a_{n+2} + p_1(n)a_{n+1} + p_0(n)a_n = 0 \quad (2.4.3)$$

的两个已知线性无关解。为求解

$$a_{n+2} + p_1(n)a_{n+1} + p_0(n)a_n = q(n), \quad (2.4.4)$$

我们寻求形如

$$a_n = A_n x_n + B_n y_n \quad (2.4.5)$$

的解 a_n 。将 (2.4.5) 代入 (2.4.4) 得

$$A_{n+2}x_{n+2} + B_{n+2}y_{n+2} + p_1(n)A_{n+1}x_{n+1} + p_1(n)B_{n+1}y_{n+1} + p_0(n)A_n x_n + p_0(n)B_n y_n = q(n).$$

利用 A_n 和 B_n 满足 (2.4.3) 的性质，可从此方程中消去 A_{n+2} 和 B_{n+2} ，得到：

$$\begin{aligned} & -x_{n+2}p_1(n)A_{n+1} - x_{n+2}p_0(n)A_n - y_{n+2}p_1(n)B_{n+1} - y_{n+2}p_0(n)B_n \\ & + p_1(n)A_{n+1}x_{n+1} + p_1(n)B_{n+1}y_{n+1} + p_0(n)A_n x_n + p_0(n)B_n y_n = q(n). \end{aligned} \quad (2.4.6)$$

至此，我们视 x_n 和 y_n 为 n 的两个独立函数。保持 x_n 任意性，施加约束

$$A_{n+1}(x_{n+2} - x_{n+1}) + B_{n+1}(y_{n+2} - y_{n+1}) = 0 \quad (2.4.7)$$

(2.4.6) 化简为：

$$-A_n p_0(n)(x_{n+2} - x_{n+1}) - p_0(n)B_n(y_{n+2} - y_{n+1}) = q(n). \quad (2.4.8)$$

联立 (2.4.7) 和 (2.4.8) 解得 $x_{n+1} - x_n$ 和 $y_{n+1} - y_n$ ：

$$\begin{aligned} x_{n+2} - x_{n+1} &= -\frac{q(n)B_{n+1}}{p_0(n)W_n}, \\ y_{n+2} - y_{n+1} &= \frac{q(n)A_{n+1}}{p_0(n)W_n}, \end{aligned}$$

其中 $W_n = A_n B_{n+1} - B_n A_{n+1}$ 为齐次方程两解的 Wronskian 行列式。由此得到通解 a_n :

$$a_n = c_1 A_n + c_2 B_n - A_n \sum_{j=0}^{n-2} \frac{q(j) B_{j+1}}{W_{j+1}} + B_n \sum_{j=0}^{n-2} \frac{q(j) A_{j+1}}{W_{j+1}}, \quad (2.4.9)$$

此处求和下限保留任意性, c_1 和 c_2 为任意常数, 并利用了关系式 $W_{j+1} = p_0(j)W_j$ 。

例 2.15 (二阶方程的参数变易法)

考虑方程

$$(n+4)a_{n+2} + a_{n+1} - (n+1)a_n = 1. \quad (2.4.10)$$

的特解。系数函数 $p_1(n) = 1/(n+4)$, $p_0(n) = -(n+1)/(n+4)$, $q(n) = 1/(n+4)$ 。两个线性无关解已知:

$$A_n = \frac{1}{(n+1)(n+2)},$$

$$B_n = \frac{(-1)^{n+1}(2n+3)}{4(n+1)(n+2)}.$$

利用 A_n 和 B_n 的表达式计算 Wronskian 行列式 W_n :

$$W_n = A_n B_{n+1} - B_n A_{n+1} = \frac{(-1)^n}{(n+1)(n+2)(n+3)}.$$

由此可得

$$\frac{B_{n+1}}{W_{n+1}} = -\frac{(n+4)(2n+5)}{4}, \quad \frac{A_{n+1}}{W_{n+1}} = (-1)^{n+1}(n+4).$$

将上述结果代入 (2.4.9) 得特解公式:

$$\begin{aligned} a_n &= -A_n \sum_{j=0}^{n-2} \frac{q(j) B_{j+1}}{W_{j+1}} + B_n \sum_{j=0}^{n-2} \frac{q(j) A_{j+1}}{W_{j+1}} \\ &= \frac{1}{4(n+1)(n+2)} \left[(n+1)^2 + (2n+3) \frac{1+(-1)^n}{2} \right]. \end{aligned}$$

可舍弃括号中的 $(-1)^n(2n+3)$ 项, 因其与齐次方程的解 B_n 成比例; 亦可舍弃括号中的常数项, 因其与 A_n 成比例。最终得特解 $a_n = n(n+3)/[4(n+1)(n+2)]$ 。易验证此解满足 (2.4.10)。
通解:

$$a_n = c_1 \frac{1}{(n+1)(n+2)} + c_2 \frac{(-1)^{n+1}(2n+3)}{4(n+1)(n+2)} + \frac{n(n+3)}{4(n+1)(n+2)}.$$

降阶法

若已知 N 阶齐次线性差分方程的一个解, 可用降阶法分离该解并将方程降至 $N-1$ 阶。降阶法亦适用于非齐次方程。对于二阶方程

$$a_{n+2} + p_1(n)a_{n+1} + p_0(n)a_n = q(n), \quad (2.4.11)$$

设 A_n 是相关联齐次方程的已知解。仿照 (2.3.6), 寻求形如

$$a_n = x_n A_n \quad (2.4.12)$$

的非齐次方程通解。将 (2.4.12) 代入 (2.4.11) 得

$$A_{n+2}(x_{n+2} - x_{n+1}) - p_0(n)A_n(x_{n+1} - x_n) = q(n). \quad (2.4.13)$$

此式为关于差分 $x_n^{(1)} = x_{n+1} - x_n$ 的一阶非齐次方程。

待定系数法

待定系数法并非一种通用方法，该方法极少适用于更复杂类型的方程，故此处不赘述。

2.5 非线性差分方程

除已提及的方法外，非线性差分方程求显式解的方法极少，下例简要概述几种应用较广的方法。

例 2.16 (代换法)

方程 $a_{n+1} = 2a_n(1 - a_n)$ 可通过代换 $a_n = (1 - b_n)/2$ 求解。新方程 $b_{n+1} = b_n^2$ 的解为 $b_n = b_1^{(2^{n-1})}$ 。因此，

$$a_n = \frac{1 - (1 - 2a_1)^{2^{n-1}}}{2}.$$

例 2.17 (利用非线性函数关系求解方程)

方程 $a_{n+1} = 4a_n^3 - 3a_n$ ($|a_0| \leq 1$) 可通过代换 $a_n = \cos \theta_n$ 简化为 $\cos \theta_{n+1} = \cos 3\theta_n$ 。故解为 $a_n = \cos(3^n \cos^{-1} a_0)$ 。若 $|a_0| > 1$ ，如何求解此差分方程？

例 2.18 (生成函数法)

卷积方程

$$a_{n+1} = K \sum_{j=0}^n a_j a_{n-j}, \quad a_0 = 1, \quad (2.5.1)$$

最适合通过生成函数 $f(x) = \sum_{n=0}^{\infty} a_n x^n$ 求解。将 (2.5.1) 乘以 x^n 并对 $n = 0$ 至 ∞ 求和，得 $f(x)$ 的方程：

$$\frac{f(x) - 1}{x} = K[f(x)]^2.$$

解为 $f(x) = (1 \pm \sqrt{1 - 4xK})/2xK$ ，为确保 $f(0) = a_0 = 1$ 需取负号。Taylor 展开得到：

$$f(x) = \sum_{n=0}^{\infty} \frac{(4xK)^n \Gamma(n + \frac{1}{2})}{(n+1)! \Gamma(\frac{1}{2})},$$

故 x^n 的系数 a_n 为

$$a_n = \frac{(4K)^n \Gamma(n + \frac{1}{2})}{(n+1)! \Gamma(\frac{1}{2})}.$$

第二部分 局部分析

当精确解无法获得或过于复杂时, 需采用近似方法。局部分析是基础步骤, 旨在用初等函数在单点邻域内近似表示解。这些局部结果可通过全局分析拼接, 形成解在整个区间上的均匀近似, 为困难方程提供了实用且高效的初步处理工具。

3 线性微分方程的近似解法

3.1 齐次线性方程的奇点分类

本章沿用第一章的齐次线性微分方程形式:

$$y^{(n)}(x) + p_{n-1}(x)y^{(n-1)}(x) + \cdots + p_1(x)y'(x) + p_0(x)y(x) = 0, \quad (3.1.1)$$

其中 $y^{(k)}(x) = d^k y/dx^k$ 。下述分类方案要求系数函数 $p_0, \dots, p_{n-1}$ 在复平面上有定义。

定义 3.1

若系数函数 $p_0(x), \dots, p_{n-1}(x)$ 均在复平面点 $x_0 \neq \infty$ 的邻域解析, 则称 x_0 为 (3.1.1) 的常点 (ordinary point)。

Fuchs (1866) 证明了在常点邻域内 (3.1.1) 的所有 n 个线性无关解均解析。进一步, 若在常点 x_0 将解展为 Taylor 级数 $y(x) = \sum_{i=0}^{\infty} a_i(x-x_0)^i$, 则收敛半径至少等于复平面上距最近系数函数奇点的距离。解的奇点位置必与某系数函数奇点重合。

例 3.1

方程 $(x^2+1)y' + 2xy = 0$ 以 0 为常点。解 $y = (x^2+1)^{-1}$ 的 Taylor 级数收敛半径为 1, 恰为 (3.1.1) 形式下系数奇点 $x = \pm i$ 到常点的最小距离。

定义 3.2

若系数 $p_0(x), \dots, p_{n-1}(x)$ 不全解析, 但函数 $(x-x_0)^n p_0(x), (x-x_0)^{n-1} p_1(x), \dots, (x-x_0) p_{n-1}(x)$ 均在 $x_0 \neq \infty$ 邻域解析, 则称 x_0 为 (3.1.1) 的正则奇点 (regular singular point)。

(3.1.1) 的解在正则奇点可能解析; 若非解析, 其奇性必为极点、代数或对数支点。Fuchs 证明: 至少存在一个形式如下的解:

$$y = (x-x_0)^\alpha A(x), \quad (3.1.2)$$

其中 α 为指标指数 (indicial exponent), $A(x)$ 在 x_0 解析, 且其 Taylor 级数收敛半径至少等于复平面上距最近系数奇点的距离。

例 3.2

方程 $y' = y/\sinh x$ 在 0 处有正则奇点。解 $y(x) = c \tanh(x/2)$ 在 $x=0$ 解析, 但在 $x = \pm i\pi$ 有极点。故 $y(x)$ Taylor 展开的收敛半径为 π , 即系数函数 $1/\sinh x$ 奇点到常点的最小距离。

若 (3.1.1) 阶数 $n \geq 2$, 则存在第二个线性无关解, 形式为以下两种之一:

$$y = (x - x_0)^\beta B(x) \quad (3.1.3)$$

或

$$y = (x - x_0)^\alpha A(x) \ln(x - x_0) + C(x)(x - x_0)^\gamma. \quad (3.1.4)$$

(3.1.3) 和 (3.1.4) 是 Euler 方程解在指标指数无重根和有重根情形的推广。 $A(x)$ 与 (3.1.2) 中函数相同, $B(x)$ 和 $C(x)$ 为新解析函数, 收敛半径同样受限于最近系数奇点距离。

一般地, 每个新线性无关解对应一个新解析函数, 并伴有一个新指标指数或额外的 $\ln(x - x_0)$ 幂次。因此, 第 n 个解的最复杂形式为:

$$y(x) = (x - x_0)^\delta \prod_{i=0}^{n-1} [\ln(x - x_0)]^i A_i(x), \quad (3.1.5)$$

所有函数 $A_i(x)$ 均在 x_0 解析。反之, Fuchs 证明, 若所有 n 个解均具有 (3.1.2) 至 (3.1.5) 的形式, 则 x_0 至多为方程的正则奇点。

定义 3.3

若点 $x_0 \neq \infty$ 既非常点也非正则奇点, 则称为 (3.1.1) 的非正则奇点 (irregular singular point)。

虽无完备理论, 但可断言在非正则奇点处, 至少存在一个解不满足 (3.1.2) 至 (3.1.5) 的形式。至此有限复平面内的点分类已完成, 对于无穷远点 $x_0 = \infty$, 通过反演变换

$$x = \frac{1}{t}, \quad \frac{d}{dx} = -t^2 \frac{d}{dt}, \quad \frac{d^2}{dx^2} = t^4 \frac{d^2}{dt^2} + 2t^3 \frac{d}{dt}, \quad (3.1.6)$$

将无穷远点解析映射至原点, 再对 $t = 0$ 分类。

定义 3.4

若 $t = 0$ 分别为常点、正则奇点或非正则奇点, 则对应称 $x_0 = \infty$ 为常点、正则奇点或非正则奇点。

以下示例展示解在常点、正则奇点及非正则奇点附近的行为。

例 3.3 (常点、正则奇点与非正则奇点对比)

考虑方程:

$$\frac{dy}{dx} - \frac{1}{2}y = 0, \quad (3.1.7)$$

$$\frac{dy}{dx} - \frac{1}{2x}y = 0, \quad (3.1.8)$$

$$\frac{dy}{dx} - \frac{1}{2x^2}y = 0. \quad (3.1.9)$$

变换 $x = 1/t$ 得新方程: $dy/dt + y/2t^2 = 0$, $dy/dt + y/2t = 0$, $dy/dt + y/2 = 0$ 。

(3.1.7) 除 ∞ (非正则奇点) 外均为常点。解 $y(x) = e^{-x/2}$ 解析, 但在 $x = \infty$ 有本性奇点。

(3.1.8) 除 0 和 ∞ (正则奇点) 外均为常点。解 $y(x) = x^{-1/2}$ 解析, 但在 $x = 0$ 和 $x = \infty$ 有支点。

(3.1.9) 除 0 (非正则奇点) 外均为常点。解 $y(x) = e^{1/(2x)}$ 在扩充复平面解析, 但在 $x = 0$ 有本性奇点。

例 3.4 (常点处的 Taylor 级数)

方程 $y' + y/(x-1) = 0$ 在 1 和 ∞ 有正则奇点。解 $y(x) = (x-1)^{-1}$ 在 $x = 1$ 有极点, 在 ∞ 解析。解在 $x = 0$ 的 Taylor 级数 $y(x) = \sum_{k=0}^{\infty} x^k$ 收敛半径为 1 (距正则奇点 $x = 1$ 的距离)。

例 3.5 (收敛范围超越最近奇点的 Taylor 解)

方程 $(x-1)(2x-1)y'' + 2xy' - 2y = 0$ 在 $1/2, 1, \infty$ 有正则奇点。一解 $y(x) = 1/(x-1)$ 在常点 $x = 0$ 的 Taylor 展开可超越最近奇点 $x = 1/2$, 但在 $|x| = 1$ 处发散。线性无关解 $y(x) = x$ 在 $x = 0$ 的 Taylor 级数对所有 x 收敛。

例 3.6 (非正则奇点附近的本性奇性)

方程 $y' + 3y/2x + y/4x^3 = 0$ 在 0 有非正则奇点, 在 ∞ 有正则奇点。两线性无关解为 $y(x) = \sin x^{-1/2}$ 和 $y(x) = \cos x^{-1/2}$ 。前者在原点有本性奇点 (含 $x = 0$ 和 $x = \infty$ 的支点), 后者无支割线且在 $x = \infty$ 解析。

例 3.7 (奇点附近的解析解)

在正则或非正则奇点处, 一个甚至多个线性无关解可能解析。

方程 $y'' + (1-x)y'/x - y/x^2 = 0$ 在 0 有正则奇点, 在 ∞ 有非正则奇点。解 $y(x) = (e^x - 1 - x)/x$ 在 $x = 0$ 解析, 但在 ∞ 有本性奇点; 线性无关解 $y(x) = (1+x)/x$ 在 $x = 0$ 有极点, 在 ∞ 解析。1.3 节例 1.10 的方程

$$y'' - \frac{1+x}{x}y' + \frac{1}{x}y = 0, \quad (3.1.10)$$

在 0 有正则奇点, 在 ∞ 有非正则奇点。但两线性无关解

$$y(x) = e^x, \quad y(x) = 1 + x \quad (3.1.11)$$

均在 $x = 0$ 解析 (前者在 ∞ 有本性奇点, 后者在 ∞ 有极点)。一般地, 所有线性无关解在正则奇点可能解析, 但在非正则奇点至少有一解奇异。

例 3.8 (自变量变换消除奇点)

方程 $y' - 2x^{-1/2}y = 0$ 在 0 的非正则奇点可通过新自变量 $t = x^{1/2}$ 消除。新方程 $dy/dt - y = 0$ 在 $t = 0$ 为常点。

例 3.9 (因变量变换消除奇点)

方程

$$y'' + \frac{2}{x}y' - y = 0 \quad (3.1.12)$$

在 0 有正则奇点, 在 ∞ 有非正则奇点。变换 $y(x) = w(x)/x$ 可消除 $x = 0$ 奇点, 但新方程 $w'' - w = 0$ 仍在 ∞ 有非正则奇点。

例 3.10 (转换为线性系统消除奇点)

(3.1.12) 的奇点可消除, 因两解在 $x = 0$ 的奇性阶数相同: $y \sim c_1/x$, $y \sim c_2e^x/x$ 。但对 (3.1.10) (解 $y = e^x$ 和 $y = 1 + x$ 均在 $x = 0$ 解析), 难以直接消除其正则奇点。唯一方法是转换为线性系统:

$$y'(x) = a(x)y(x) + b(x)z(x), \quad z'(x) = c(x)y(x) + d(x)z(x). \quad (3.1.13)$$

若二阶微分方程通解在 $x = 0$ 解析, 则存在形如 (3.1.13) 的等价齐次线性系统, 其系数在 $x = 0$ 解析。(3.1.10) 的适当系统为 $y' = y + z$, $z' = 0$ 。能否找到另一系数解析的等价系统?

3.2 齐次线性方程常点附近的局部级数展开

当精确显式解未知时, 为获得常点处的级数解, 我们将 Taylor 级数 $y(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ 代入微分方程, 并通过递推关系确定系数 a_n 。

例 3.11 (一阶微分方程的 Taylor 级数解)

考虑初值问题 $y' = 2xy$ [$y(0) = 1$]。因 $x = 0$ 是常点, 设解 $y(x) = \sum_{n=0}^{\infty} a_n x^n$, 则:

$$\sum_{n=0}^{\infty} n a_n x^{n-1} = 2x \sum_{n=0}^{\infty} a_n x^n = 2 \sum_{n=0}^{\infty} a_{n-2} x^{n-1}.$$

由 Taylor 系数唯一性, 得:

$$n a_n = 0, \quad n = 0, 1;$$

$$n a_n = 2 a_{n-2}, \quad n = 2, 3, \dots$$

由 $n = 0$ 知 a_0 为任意常数; 由 $n = 1$ 知 $a_1 = 0$; 解得 $a_{2n} = a_0/n!$ ($n \geq 1$), $a_{2n+1} = 0$ ($n \geq 1$)。故 $y' = 2xy$ 的通解为:

$$y(x) = a_0 \sum_{n=0}^{\infty} \frac{x^{2n}}{n!} = a_0 e^{x^2}.$$

可代回方程验证。

例 3.12 (Airy 方程在 $x = 0$ 的局部分析)

为求 Airy 方程 (1.4.8) $y'' = xy$ 在 $x = 0$ 附近的局部行为, 代入级数 $y(x) = \sum_{n=0}^{\infty} a_n x^n$ 得:

$$\sum_{n=0}^{\infty} a_n n(n-1)x^{n-2} = \sum_{n=0}^{\infty} a_n x^{n+1}.$$

联立系数:

$$a_n n(n-1) = 0, \quad n = 0, 1, 2;$$

$$a_n n(n-1) = a_{n-3}, \quad n = 3, 4, \dots$$

$n = 0, 1$ 时得 a_0, a_1 任意, $a_2 = 0$, 解得:

$$\begin{aligned} a_{3n} &= \frac{a_0}{3^n n! 3^n (n - \frac{1}{3})(n - 1 - \frac{1}{3}) \cdots (\frac{2}{3})} = \frac{a_0 \Gamma(\frac{2}{3})}{9^n n! \Gamma(n + \frac{2}{3})}, \\ a_{3n+1} &= \frac{a_1}{3^n n! 3^n (n - \frac{1}{3})(n - 1 + \frac{1}{3}) \cdots (\frac{4}{3})} = \frac{a_1 \Gamma(\frac{4}{3})}{9^n n! \Gamma(n + \frac{4}{3})}, \\ a_{3n+2} &= 0 \end{aligned}$$

定义 $c_1 = a_0 \Gamma(\frac{2}{3})$, $c_2 = a_1 \Gamma(\frac{4}{3})$, 则 Airy 方程通解为:

$$y(x) = c_1 \sum_{n=0}^{\infty} \frac{x^{3n}}{9^n n! \Gamma(n + \frac{2}{3})} + c_2 \sum_{n=0}^{\infty} \frac{x^{3n+1}}{9^n n! \Gamma(n + \frac{4}{3})}.$$

此通解可通过快速收敛的 Taylor 展开近似表示于所有有限 x 。定义两个特解 (Airy 函数):

$$\text{Ai}(x) \equiv 3^{-2/3} \sum_{n=0}^{\infty} \frac{x^{3n}}{9^n n! \Gamma(n + \frac{2}{3})} - 3^{-4/3} \sum_{n=0}^{\infty} \frac{x^{3n+1}}{9^n n! \Gamma(n + \frac{4}{3})}, \quad (3.2.1)$$

$$\text{Bi}(x) \equiv 3^{-1/6} \sum_{n=0}^{\infty} \frac{x^{3n}}{9^n n! \Gamma(n + \frac{2}{3})} + 3^{-5/6} \sum_{n=0}^{\infty} \frac{x^{3n+1}}{9^n n! \Gamma(n + \frac{4}{3})}. \quad (3.2.2)$$

当 $|x|$ 不大时, (3.2.1) 和 (3.2.2) 的 Taylor 级数仅需几项即可高精度逼近 Airy 函数 (误差估计见问题 3.7)。

3.3 齐次线性方程正则奇点附近的局部级数展开

若尝试用 Taylor 级数表示解在正则奇点附近的局部行为, 会发生什么?

例 3.13 (正则奇点处 Taylor 级数表示的失效)

对方程 $y'' + y/4x^2 = 0$ 在 $x = 0$ 处寻求 Taylor 级数解 $y = \sum_{n=0}^{\infty} a_n x^n$ 。代入方程并逐项微分后, 得到关于 a_n 的方程组: $(n - \frac{1}{2})^2 a_n = 0$ ($n = 0, 1, 2, \dots$)。其解为 $a_n = 0$, 仅得到平凡解 $y(x) = 0$ 。此方法无效。

上例中 Taylor 级数失效的原因是: Taylor 级数不足以描述解在正则奇点附近的局部行为。但是 3.1 节 Fuchs 定理告诉我们解具有比 Taylor 级数更一般的结构: 若 x_0 是齐次线性微分方程的正则奇点, 则至少存在一个形如 (3.1.2) 的解: $y(x) = (x - x_0)^\alpha A(x)$, 其中 $A(x)$ 在 x_0 解析。若 α 非整数,

解在 x_0 有代数支点；若 α 为负整数，则有极点；若 $\alpha = 0, 1, 2, \dots$ ，则解在 x_0 解析。 n 阶方程的其余 $n-1$ 个线性无关解可能具有 (3.1.3)、(3.1.4) 或 (3.1.5) 的形式，故可能含对数支点。

因 $A(x)$ 解析，故可展为 Taylor 级数：

$$y(x) = (x - x_0)^\alpha A(x) = (x - x_0)^\alpha \sum_{n=0}^{\infty} a_n (x - x_0)^n. \quad (3.3.1)$$

称 (3.3.1) 右侧为 **Frobenius 级数**， α 为 **指标指数**。通常假设 Frobenius 级数中 $a_0 \neq 0$ （通过适当选择 α 实现）。Taylor 级数是 Frobenius 级数的特例。

例 3.14 (正则奇点处的局部分析)

对上例的方程，在 $x_0 = 0$ 处尝试 (3.3.1) 的 Frobenius 级数，得：

$$\left[(n + \alpha)(n + \alpha - 1) + \frac{1}{4} \right] a_n = 0 \quad (n = 0, 1, 2, \dots).$$

由 $a_0 \neq 0$ 得：

$$\alpha(\alpha - 1) + \frac{1}{4} = 0 \implies \alpha = \frac{1}{2}.$$

代入 $\alpha = \frac{1}{2}$ 后，其余方程给出 $a_1 = a_2 = \dots = 0$ 。故 Frobenius 级数 $y(x) = a_0 \sqrt{x}$ (a_0 任意)，此为精确解。

由 1.4 节 Euler 方程的研究，知该方程的第二个解为 $y(x) = b\sqrt{x} \ln x$ （不具有 (3.3.1) 形式）。一般理论仅保证存在一个 Frobenius 级数形式的解，其余解可能更复杂。

下面我们将详述 Fuchs 和 Frobenius 提出的系统方法，用于计算指标指数 α 和展开系数 a_n 。此方法亦可求解形如 (3.1.4) 和 (3.1.5) 的含对数解，方法的核心是判断其余解应取 (3.1.2)-(3.1.5) 中何种形式的准则。

二阶方程的 Frobenius 方法

若微分方程

$$y'' + \frac{p(x)}{x - x_0} y' + \frac{q(x)}{(x - x_0)^2} y = 0 \quad (3.3.2)$$

在 x_0 有正则奇点，则 $p(x)$ 和 $q(x)$ 在 x_0 解析。故可将 $p(x)$ 和 $q(x)$ 展为 x_0 处的 Taylor 级数：

$$p(x) = \sum_{n=0}^{\infty} p_n (x - x_0)^n, \quad q(x) = \sum_{n=0}^{\infty} q_n (x - x_0)^n.$$

代入 (3.3.2) 并将 Frobenius 级数解 (3.3.1) 代入，对比系数得：

$$(x - x_0)^{\alpha-2} : [\alpha^2 + (p_0 - 1)\alpha + q_0] a_0 = 0, \quad (3.3.3a)$$

$$\begin{aligned} (x - x_0)^{\alpha+n-2} : & [(\alpha + n)^2 + (p_0 - 1)(\alpha + n) + q_0] a_n \\ & = - \sum_{k=0}^{n-1} [(\alpha + k)p_{n-k} + q_{n-k}] a_k, \quad n = 1, 2, \dots \end{aligned} \quad (3.3.3b)$$

因 $a_0 \neq 0$, (3.3.3a) 要求 α 为指标多项式 $P(\alpha)$ 的根:

$$P(\alpha) = \alpha^2 + (p_0 - 1)\alpha + q_0. \quad (3.3.4)$$

对每个 α 值, 需由递推关系 (3.3.3b) 解出 a_n (以 a_0 表示)。仅当 $P(\alpha + n) \neq 0$ 时, (3.3.3b) 可递推求解 a_n 。若此条件对所有正整数 n 成立, 则级数 (3.3.1) 的收敛半径至少等于 $p(x)$ 或 $q(x)$ 的最小复奇点距离。设 α_1 和 α_2 为指标多项式 $P(\alpha)$ 的两根满足 $\operatorname{Re} \alpha_1 \geq \operatorname{Re} \alpha_2$ 。取 $\alpha = \alpha_1$ 时, 因 α_2 是 $P(\alpha)$ 的唯一另一根, 故 $P(\alpha_1 + n) \neq 0$ ($n = 1, 2, \dots$), 此即 Frobenius 级数解恒存在的原因。

例 3.15 (ν 阶修正 Bessel 方程)

修正 Bessel 方程

$$y'' + \frac{1}{x}y' - \left(1 + \frac{\nu^2}{x^2}\right)y = 0 \quad (3.3.5)$$

在 $x = 0$ 有正则奇点。代入 Frobenius 级数 $y = \sum_{n=0}^{\infty} a_n x^{\alpha+n}$ 得:

$$\sum_{n=0}^{\infty} (\alpha + n)(\alpha + n - 1)a_n x^{\alpha+n-2} + \sum_{n=0}^{\infty} (\alpha + n)a_n x^{\alpha+n-2} - \sum_{n=0}^{\infty} a_n x^{\alpha+n} - \nu^2 \sum_{n=0}^{\infty} a_n x^{\alpha+n-2} = 0.$$

将 $\sum_{n=0}^{\infty} a_n x^{\alpha+n}$ 改写为 $\sum_{n=1}^{\infty} a_{n-1} x^{\alpha+n-2}$, 联立 x 的幂次系数:

$$\begin{aligned} x^{\alpha-2}: & (\alpha^2 - \nu^2)a_0 = 0, \\ x^{\alpha-1}: & [(\alpha + 1)^2 - \nu^2]a_1 = 0, \\ x^{\alpha+n-2}: & [(\alpha + n)^2 - \nu^2]a_n = a_{n-1}, \quad n = 2, 3, \dots \end{aligned} \quad (3.3.6)$$

由 $a_0 \neq 0$, α 须满足指标方程 $P(\alpha) = \alpha^2 - \nu^2 = 0$, 故 $\alpha = \pm\nu$ 。不妨设 $\operatorname{Re} \nu \geq 0$, 则 $\alpha_1 = \nu$, $\alpha_2 = -\nu$ 。因 $P(\alpha_1 + n) \neq 0$ ($n = 1, 2, \dots$), 可由 (3.3.6) 解得 $a_1 = a_3 = \dots = 0$,

$$a_{2n} = \frac{a_{2n-2}}{2^2 n(\nu + n)} = \frac{a_{2n-4}}{2^4 n(n-1)(\nu + n)(\nu + n - 1)} = \dots = \frac{a_0 \Gamma(\nu + 1)}{2^{2n} n! \Gamma(\nu + n + 1)}. \quad (3.3.7)$$

故

$$y(x) = a_0 \Gamma(\nu + 1) x^\nu \sum_{n=0}^{\infty} \frac{(x/2)^{2n}}{n! \Gamma(\nu + n + 1)}.$$

取 $a_0 = 2^{-\nu} / \Gamma(\nu + 1)$, 即得修正 Bessel 函数 $I_\nu(x)$ 的 Frobenius 级数:

$$I_\nu(x) = \sum_{n=0}^{\infty} \frac{(x/2)^{2n+\nu}}{n! \Gamma(\nu + n + 1)}. \quad (3.3.8)$$

比值判别法表明级数收敛半径无穷, 对应 (3.3.5) 在有限复平面仅 $x = 0$ 一个奇点。

当 ν 非整数时, 取 $\alpha = -\nu$ 可得第二解。若 2ν 亦非整数, 则 $P(-\nu + n) \neq 0$ (对所有正整数 n), 故存在第二个 Frobenius 级数。只要 2ν 非偶数, 由 (3.3.7) 可确定 a_n , 且 $I_{-\nu}(x)$ 的级数 (3.3.8) 收敛。当 ν 非整数时, 因 $I_\nu(x)$ 与 $I_{-\nu}(x)$ 的首项幂次不同, 二者线性无关。

故当 $\nu \neq 0, 1, 2, 3, \dots$ 时, Frobenius 方法给出完备的线性无关解集。当 ν 为整数时, 求第二解较复杂, 本节后续将阐明相应方法。

总的来说, 对二阶方程 (3.3.2), 若指标指数差 $\alpha_1 - \alpha_2$ 非整数, 则取 $\alpha = \alpha_1$ 和 $\alpha = \alpha_2$ 可构造

两个线性无关的 Frobenius 级数解。若 $\alpha_1 = \alpha_2$, 则取 $\alpha = \alpha_2$ 仅得与第一解成比例的解, 需进一步分析求第二解。若 $\alpha_1 - \alpha_2 = N$ 为正整数, 则当 $n = N$ 时递推关系 (3.3.3b) 化为:

$$0 \cdot a_N = - \sum_{k=0}^{N-1} [(\alpha + k)p_{N-k} + q_{N-k}]a_k. \quad (3.3.9)$$

此时有两种可能:

1. 若 (3.3.9) 右端非零, 则 a_N 不存在, 需额外方法求第二线性无关解;
2. 若 (3.3.9) 右端恰为零, 则方程退化为恒等式 $0 = 0$, a_N 为任意常数。递推关系可越过 $n = N$ 继续求解 $a_{N+1}, a_{N+2}, \dots$, 得到含两个任意常数 a_0 和 a_N 的第二线性无关解。

综上我们有分类:

I $\alpha_1 \neq \alpha_2$ 且 $\alpha_1 - \alpha_2$ 非整数: 存在两个 Frobenius 级数解。

II $\alpha_1 - \alpha_2 = N = 0, 1, 2, \dots$, 分两子类:

(a) $\alpha_1 = \alpha_2$: 仅一个 Frobenius 级数解。

(b) $\alpha_1 - \alpha_2 = 1, 2, 3, \dots$, 再分两情形:

(i) (3.3.9) 右端非零: 仅一个 Frobenius 级数解。

(ii) (3.3.9) 右端为零: 存在两个 Frobenius 级数解。

情形 II(b)(ii)

此情形由 (3.3.9) 右端为零刻画, 确保存在指标指数 α_2 的 Frobenius 级数解。

例 3.16 (常点处的 Frobenius 级数)

在微分方程的常点 x_0 应用 Frobenius 理论。此时 $p_0 = q_0 = 0$, 指标方程为 $P(\alpha) = \alpha(\alpha-1) = 0$, 指标指数 $\alpha_1 = 1, \alpha_2 = 0$ 。对应解:

$\alpha_1 = 1$: 解为 Taylor 级数 $y(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^{n+1}$, 满足初值 $y(x_0) = 0, y'(x_0) = a_0 \neq 0$ 。

$\alpha_2 = 0$: 存在形式 (3.3.1) 的第二解, 亦为 Taylor 级数。

例 3.17 (半整数阶修正 Bessel 方程)

修正 Bessel 方程 (3.3.5) 中取 $\nu = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$ (半整数), 则指标指数差 $\alpha_1 - \alpha_2 = 2\nu = 1, 3, 5, \dots$ 。此处 $p_0 = 1, q_0 = -\nu^2$, 其余 $p_n = q_n = 0$ ($n \geq 1$)。可证 (3.3.9) 右端为零, 故两解均有 Frobenius 级数展开, 即 (3.3.8) 给出的 $I_\nu(x)$ 和 $I_{-\nu}(x)$ 。

情形 II(a)

将 a_n 视为 a_0 和 α 的函数, 记所得 Frobenius 级数为:

$$y(x, \alpha) = (x - x_0)^\alpha \sum_{n=0}^{\infty} a_n(\alpha)(x - x_0)^n. \quad (3.3.10)$$

注意: 仅当 $\alpha = \alpha_1$ 时 $y(x, \alpha)$ 才是 (3.3.2) 的解。

定义微分算子:

$$L \equiv \frac{d^2}{dx^2} + \frac{p(x)}{x - x_0} \frac{d}{dx} + \frac{q(x)}{(x - x_0)^2}.$$

(3.3.2) 的解满足 $Ly(x) = 0$ 。由构造, $a_n(\alpha)$ 满足递推关系 (3.3.3b), 故 $y(x, \alpha)$ 在 $\alpha \neq \alpha_1$ 时满足:

$$Ly(x, \alpha) = a_0(x - x_0)^{\alpha-2}P(\alpha). \quad (3.3.11)$$

取 $\alpha = \alpha_1$ 时, (3.3.11) 右端为零, $y(x, \alpha_1)$ 即为 Frobenius 级数解。但若对 (3.3.11) 关于 α 求导后取 $\alpha = \alpha_1$, 因 $\alpha_1 = \alpha_2$ 时 $P(\alpha) = (\alpha - \alpha_1)^2$, 右端仍为零。因此, $\frac{\partial}{\partial \alpha} y(x, \alpha)|_{\alpha=\alpha_1}$ 是 (3.3.2) 的新解:

$$L \left[\frac{\partial}{\partial \alpha} y(x, \alpha) \Big|_{\alpha=\alpha_1} \right] = 0.$$

对 (3.3.3b) 关于 α 求导, 得新解形式:

$$\frac{\partial}{\partial \alpha} y(x, \alpha) \Big|_{\alpha=\alpha_1} = y(x, \alpha_1) \ln(x - x_0) + \sum_{n=0}^{\infty} b_n (x - x_0)^{\alpha_1+n}, \quad (3.3.12)$$

其中

$$b_n = \frac{\partial}{\partial \alpha} a_n(\alpha) \Big|_{\alpha=\alpha_1}. \quad (3.3.13)$$

由此得到正则奇点附近两个线性无关解的级数展开, (3.3.12) 的收敛半径至少等于方程 (3.3.2) 的最近奇点距离。

例 3.18 (零阶修正 Bessel 方程)

(3.3.5) 中取 $\nu = 0$, 指标多项式 $P(\alpha) = \alpha^2$ 在 $\alpha = 0$ 有二重根。第一解 (3.3.8) 为:

$$I_0(x) = \sum_{n=0}^{\infty} \frac{(x/2)^{2n}}{(n!)^2}. \quad (3.3.14)$$

线性无关解形如 (3.3.12)。由 (3.3.6) 得: $a_0(\alpha) = a_0, a_{2k+1}(\alpha) = 0 (k = 0, 1, 2, \dots)$, 且

$$a_{2n}(\alpha) = \frac{a_0}{2^{2n}(\alpha + 2n)^2(\alpha + 2n - 2)^2 \cdots (\alpha + 2)^2}, \quad n \geq 1.$$

故 $b_0 = \frac{\partial}{\partial \alpha} a_0(\alpha) \Big|_{\alpha=0} = 0, b_{2k+1} = 0$, 且

$$b_{2n} = \frac{\partial}{\partial \alpha} a_{2n}(\alpha) \Big|_{\alpha=0} = -\frac{a_0}{2^{2n}(n!)^2} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} \right), \quad n \geq 1.$$

代入 (3.3.12) 得 $I_0(x)$ 的线性无关解:

$$\frac{\partial}{\partial \alpha} y(x, \alpha) \Big|_{\alpha=0} = a_0 \ln x \sum_{n=0}^{\infty} \frac{(x/2)^{2n}}{(n!)^2} - a_0 \sum_{n=1}^{\infty} \frac{(x/2)^{2n}}{(n!)^2} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} \right).$$

取 $a_0 = -1$ 并加上 $\ln 2 - \gamma$ 倍的 $I_0(x)$ 得特解 $K_0(x)$:

$$K_0(x) = -\left(\ln \frac{x}{2} + \gamma\right) I_0(x) + \sum_{n=1}^{\infty} \frac{(x/2)^{2n}}{(n!)^2} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} \right). \quad (3.3.15)$$

如此定义 $K_0(x)$ 的原因是其在 $x \rightarrow \infty$ 有优美的渐近行为, 后文我们将提到。

情形 II(b)(i)

当 $\alpha_1 - \alpha_2 = N$ 且 (3.3.9) 右端非零时, 尝试类似情形 II(a) 的方法。对 (3.3.11) 关于 α 求导后取 $\alpha = \alpha_1$, 得 $\frac{\partial}{\partial \alpha} y(x, \alpha) \Big|_{\alpha=\alpha_1}$ 满足非齐次方程:

$$L \left[\frac{\partial}{\partial \alpha} y(x, \alpha) \Big|_{\alpha=\alpha_1} \right] = a_0 P'(\alpha_1) (x - x_0)^{\alpha_2 + N - 2}. \quad (3.3.16)$$

为构造齐次方程 (3.3.2) 的解, 需先求 (3.3.16) 的第二个特解。可证存在标准 Frobenius 展开的特解:

$$\sum_{n=0}^{\infty} c_n (x - x_0)^{\alpha_2 + n}. \quad (3.3.17)$$

代入 (3.3.16) 并联立 $(x - x_0)^{\alpha_2 + n - 2}$ 系数:

$$(x - x_0)^{\alpha_2 - 2}: \quad P(\alpha_2) c_0 = 0, \quad (3.3.18a)$$

$$(x - x_0)^{\alpha_2 + n - 2}: \quad P(\alpha_2 + n) c_n + \sum_{k=0}^{n-1} [(\alpha_2 + k) p_{n-k} + q_{n-k}] c_k = 0, \quad n \neq 0, N, \quad (3.3.18b)$$

$$(x - x_0)^{\alpha_2 + N - 2}: \quad P(\alpha_2 + N) c_N + \sum_{k=0}^{N-1} [(\alpha_2 + k) p_{N-k} + q_{N-k}] c_k = a_0 P'(\alpha_1). \quad (3.3.18c)$$

因 $P(\alpha_2 + N) = 0$, (3.3.18c) 不决定 c_N , 而给出约束:

$$a_0 = \frac{1}{P'(\alpha_1)} \sum_{k=0}^{N-1} [(\alpha_2 + k) p_{N-k} + q_{N-k}] c_k. \quad (3.3.19)$$

构造 (3.3.2) 的第二解 $y(x)$ 为两非齐次解之差:

$$y(x) = \sum_{n=0}^{\infty} c_n (x - x_0)^{\alpha_2 + n} - \frac{\partial}{\partial \alpha} y(x, \alpha) \Big|_{\alpha=\alpha_1}.$$

其中系数 c_n 满足 (3.3.18b); c_0 和 c_N 任意, (3.3.19) 决定 a_0 。展开式 (3.3) 的收敛半径至少为 $p(x)$ 和 $q(x)$ 的最近奇点距离。

例 3.19 (整数阶修正 Bessel 方程 ($\nu = 1$))

当 $\nu = 1$ 时, (3.3.5) 的一解为 Frobenius 级数 $I_1(x)$ 。取 $c_0 = 1, c_2 = \frac{1}{2} - \frac{1}{2} \ln 2 - \frac{\gamma}{4}$, 定义修正 Bessel 函数 $K_1(x)$:

$$K_1(x) = \left(\gamma + \ln \frac{x}{2} \right) I_1(x) + \frac{1}{x} - \frac{x}{4} - \sum_{n=1}^{\infty} \frac{(x/2)^{2n}}{n!(n+1)!} \left(\sum_{k=1}^n \frac{1}{k} + \frac{1}{2n+2} \right). \quad (3.3.20)$$

至此, 我们完整叙述了二阶方程的 Frobenius 方法。

高阶方程的 Frobenius 级数

Frobenius 方法可推广至 n 阶齐次线性微分方程在正则奇点 x_0 的情形:

$$\frac{d^n y}{dx^n} + \frac{q_{n-1}(x)}{x-x_0} \frac{d^{n-1} y}{dx^{n-1}} + \cdots + \frac{q_0(x)}{(x-x_0)^n} y = 0,$$

其中 $q_0(x), \dots, q_{n-1}(x)$ 在 x_0 解析。指标方程为:

$$\alpha(\alpha-1)\cdots(\alpha-n+1) + q_{n-1}(x_0)\alpha(\alpha-1)\cdots(\alpha-n+2) + \cdots + q_0(x_0) = 0.$$

若 n 个根 α 的差非整数, 则存在 n 个形如 (3.3.1) 的线性无关解, 否则需推广至 (3.1.4) 或 (3.1.5) 的形式。

3.4 齐次线性方程非正则奇点处的局部级数展开

前两节方法的固化风格反映了理论的完备性, 几乎没有想象性数学的空间。但从此处开始, 风格转变。分析非正则奇点附近解的局部行为需发展全新数学工具, 整体构成近似演算体系。

本节将看到, 描述非正则奇点附近解局部行为的公式是 Frobenius 级数的推广, 如同 Frobenius 级数是 Taylor 级数的推广。

首先我们来看 Frobenius 级数为何不足以描述非正则奇点附近行为。

例 3.20

微分方程 $y' = x^{1/2}y$ 在 $x = 0$ 有非正则奇点。解 $y(x)$ 可表为 $x = 0$ 处的收敛级数:

$$y(x) = a_0 e^{2x^{3/2}/3} = a_0 \sum_{n=0}^{\infty} \frac{(2x^{3/2}/3)^n}{n!},$$

此级数并非 Frobenius 级数。

例 3.21

微分方程

$$x^3 y'' = y \tag{3.4.1}$$

若存在形式解 $y = \sum_{n=0}^{\infty} a_n x^{\alpha+n}$ ($a_0 \neq 0$), 则

$$\sum_{n=0}^{\infty} (\alpha+n)(\alpha+n-1) a_n x^{\alpha+n+1} = \sum_{n=0}^{\infty} a_n x^{\alpha+n}.$$

联立 x^α 项系数得 $a_0 = 0$, 矛盾!

例 3.22

方程

$$x^2 y'' + (1 + 3x)y' + y = 0 \quad (3.4.2)$$

在 0 有非正则奇点。代入 $y = \sum_{n=0}^{\infty} a_n x^{\alpha+n}$:

$$x^2 \sum_{n=0}^{\infty} (\alpha+n)(\alpha+n-1)a_n x^{\alpha+n-2} + (1+3x) \sum_{n=0}^{\infty} (\alpha+n)a_n x^{\alpha+n-1} + \sum_{n=0}^{\infty} a_n x^{\alpha+n} = 0.$$

联立 x 幂次系数:

$$x^{\alpha-1}: \quad \alpha a_0 = 0,$$

$$x^{\alpha+n}: \quad (\alpha+n+1)a_{n+1} + (\alpha+n+1)^2 a_n = 0, \quad n = 0, 1, \dots$$

因 Frobenius 级数要求 $a_0 \neq 0$, 指标方程为 $\alpha = 0$. 解得 $a_{n+1} = -(n+1)a_n$, 故 $a_n = (-1)^n n! a_0$, 解为:

$$y(x) = a_0 \sum_{n=0}^{\infty} (-1)^n n! x^n. \quad (3.4.3)$$

该级数收敛半径为零, 对所有 $x \neq 0$ 发散, 矛盾!

发散级数 (3.4.3) 是应用数学中常见的级数展开——**渐近级数**。其虽发散, 但可得到微分方程解的精确近似。

渐近分析简介

引入两个新符号以描述函数相对行为:

定义 3.5记号 $f(x) \ll g(x), x \rightarrow x_0$ 定义为:

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0.$$

记号 $f(x) \sim g(x), x \rightarrow x_0$ 定义为:

$$f(x) - g(x) \ll g(x), \quad x \rightarrow x_0,$$

或等价于

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 1.$$

注意到若 $f(x) \sim g(x) (x \rightarrow x_0)$, 则 $g(x) \sim f(x) (x \rightarrow x_0)$ 。

非正则奇点附近的行为

当前首要任务是求解本节开头两个无法用 Frobenius 级数求解的方程。先观察已知结果: (3.4.1) 的两解:

$$y(x) \sim c_1 x^{3/4} e^{2x^{-1/2}}, \quad x \rightarrow 0^+; \quad (3.4.4a)$$

$$y(x) \sim c_2 x^{3/4} e^{-2x^{-1/2}}, \quad x \rightarrow 0^+. \quad (3.4.4b)$$

(3.4.2) 的缺失解:

$$y(x) \sim c_3 x^{-1} e^{x^{-1}}, \quad x \rightarrow 0^+. \quad (3.4.5)$$

这些渐进行为均在 $x = 0$ 有本性奇点。

定义 3.6

上述渐近行为实为解局部行为无穷级数表示的首项, 称为**主导行为 (leading behavior)**。主导行为在 $x \rightarrow x_0$ 时变化最快的部分称为**控制因子 (controlling factor)**。例如上三例的控制因子分别为 $e^{2x^{-1/2}}, e^{-2x^{-1/2}}, e^{x^{-1}}$ 。级数 (3.4.3) 的主导行为为 a_0 , Frobenius 级数 $\sum a_n x^{\alpha+n}$ 的主导行为为 $a_0 x^\alpha$ 。

指数代换法

要得知非正则奇点附近解的主导行为, 首先要识别控制因子。因控制因子常呈指数形式, 故采用代换:

$$y(x) = e^{S(x)}. \quad (3.4.6)$$

此技巧将 n 阶线性微分方程化为 $S(x)$ 的近似一阶方程, 以二阶方程为例:

$$y'' + p(x)y' + q(x)y = 0 \quad (3.4.7)$$

有非正则奇点 x_0 。代入 $y = e^S$ 得:

$$S'' + (S')^2 + p(x)S' + q(x) = 0. \quad (3.4.8)$$

若 x_0 为非正则奇点, 通常有:

$$S'' \ll (S')^2, \quad x \rightarrow x_0. \quad (3.4.9)$$

略去 S'' 得渐近微分方程:

$$(S')^2 \sim -p(x)S' - q(x), \quad x \rightarrow x_0. \quad (3.4.10)$$

此方程易解, 且可验证近似 (3.4.9) 的有效性。

例 3.23 (非正则奇点 0 附近的控制因子)

回顾微分方程 (3.4.1) $x^3 y'' = y$ 在非正则奇点 0 附近。渐近微分方程 (3.4.10) 给出: $(S')^2 \sim x^{-3}$, $x \rightarrow 0^+$. 解得 $S'(x) \sim \pm x^{-3/2}$ ($x \rightarrow 0^+$), 积分得:

$$S(x) \sim \pm 2x^{-1/2}, \quad x \rightarrow 0^+. \quad (3.4.11)$$

此即 (3.4.4a), (3.4.4b) 中的控制因子。

例 3.24 (非正则奇点 0 附近的主导行为)

继续考虑解 $S \sim 2x^{-1/2}$ ($x \rightarrow 0^+$)。设:

$$S(x) = 2x^{-1/2} + C(x), \quad C(x) \ll 2x^{-1/2}, \quad x \rightarrow 0^+. \quad (3.4.12)$$

代入 (3.4.8) 得: $\frac{3}{2}x^{-5/2} + C'' - 2x^{-3/2}C' + (C')^2 = 0$. 由 $C' \ll x^{-3/2}$ 得渐近方程:

$$\frac{3}{2}x^{-5/2} \sim 2x^{-3/2}C', \quad x \rightarrow 0^+, \quad (3.4.13)$$

解得 $C \sim \frac{3}{4} \ln x$, 且满足 $C'' \sim -\frac{3}{4}x^{-2} \ll x^{-5/2}$ 成立。设:

$$S(x) = 2x^{-1/2} + \frac{3}{4} \ln x + D(x) \quad (3.4.14)$$

其中

$$D(x) \ll \ln x, \quad x \rightarrow 0^+. \quad (3.4.15)$$

代入 (3.4.8) 得: $-\frac{3}{16}x^{-2} + D'' + (D')^2 - 2x^{-3/2}D' + \frac{3}{2}x^{-1}D' = 0$. 由

$$\frac{3}{2}x^{-1}D' \ll 2x^{-3/2}D', (D')^2 \ll x^{-1}D', D'' \ll x^{-2}, \quad x \rightarrow 0^+.$$

得渐近方程:

$$-2x^{-3/2}D' \sim \frac{3}{16}x^{-2}, \quad x \rightarrow 0^+.$$

解得:

$$D' \sim -\frac{3}{32}x^{-1/2} \implies D(x) - d \sim \frac{3}{16}x^{1/2}$$

其中 d 为常数, 换句话说存在 $\delta(x) \sim -3x^{1/2}/16, x \rightarrow 0^+$:

$$D(x) = d + \delta(x). \quad (3.4.16)$$

例 3.25 (数值验证)

我们先检验已有结果。设 $e^{\delta(x)} = 1 + \varepsilon(x)$, 其中 $\varepsilon(x) \rightarrow 0$ 当 $x \rightarrow 0^+$ 。为验证 $\varepsilon(x)$ 在 $x \rightarrow 0^+$ 时趋近于零, 取初值 $y(1) = 1, y'(1) = 0$ 数值求解 (3.4.1)。通过拟合数值解可检验主导行为预测:

$$y(x) = c_1 x^{3/4} e^{2x^{-1/2}} [1 + \varepsilon(x)]. \quad (3.4.17)$$

$y(x)/(x^{3/4}e^{2x^{-1/2}})$ 在 $x \rightarrow 0^+$ 时趋近极限 $c_1 \approx 0.1432 \neq 0$ 。

主导平衡法

主导平衡法 (Method of dominant balance) 用于识别在渐近极限下可能被忽略的方程中的那些项, 这包含三个步骤:

1. 丢弃所有小项, 并用渐近关系取代精确方程。

2. 渐近关系替换为一个方程, 把 $\sim$ 号换成等号, 并精确求解所得方程。

3. 检查所得到的解是否与步骤 (1) 中所做的近似一致, 并证明: 通过从精确解本身中分解出主导平衡解而得到的关于剩余函数的方程, 其解的变动比主导平衡解更缓慢。

例 3.26 (修正原点附近的主导行为)

我们对 (3.4.1) 式解的行为的分析仍不完整。我们的策略是计算 (3.4.17) 式中 $\varepsilon(x)$ 的逐次近似, 首先找到 $1 + \varepsilon(x)$ 满足的微分方程, 然后将以形式幂级数求解该方程,

$$1 + \varepsilon(x) = 1 + a_1 x^\alpha + a_2 (x^\alpha)^2 + a_3 (x^\alpha)^3 + \cdots, \quad (3.4.18)$$

其中 α 是一个正数。我们将 (3.4.17) 代入 (3.4.1) 来计算系数 a_n 与 α , 得到 $w(x) = 1 + \varepsilon(x)$ 满足的方程:

$$w'' + \left(\frac{3}{2x} - \frac{2}{x^{3/2}} \right) w' - \frac{3}{16x^2} w = 0. \quad (3.4.19)$$

我们将 $w(x) = 1 + \varepsilon(x)$ 代入 (3.4.19), 得到

$$\varepsilon'' + (3/2x - 2/x^{3/2})\varepsilon' - 3\varepsilon/16x^2 = 0.$$

注意到当 $x \rightarrow 0+$ 时 $\varepsilon \ll 1$, 从而

$$3\varepsilon/16x^2 \ll 3/16x^2, \quad x \rightarrow 0+,$$

并且

$$3x^{-1}/2 \ll 2x^{-3/2}, \quad x \rightarrow 0+,$$

这给出了渐近微分方程

$$\varepsilon'' - \frac{2}{x^{3/2}}\varepsilon' \sim \frac{3}{16x^2}, \quad x \rightarrow 0+, \quad (3.4.20)$$

(a) 当 $x \rightarrow 0+$ 时 $\varepsilon'' \sim 3x^{-2}/16$ 且 $2x^{-3/2}\varepsilon' \ll 3x^{-2}/16$ 。对前一个渐近关系积分两次得 $\varepsilon(x) \sim -3 \ln x / 16 (x \rightarrow 0+)$, 这与 $\varepsilon \rightarrow 0, x \rightarrow 0+$ 矛盾。

(b) 当 $x \rightarrow 0+$ 时 $\varepsilon''/\varepsilon' \sim 2x^{-3/2}$ 且 $3x^{-2}/16 \ll 2x^{-3/2}\varepsilon'$ 。对前一个渐近关系积分得 $\ln \varepsilon' \sim -4x^{-1/2} (x \rightarrow 0+)$, 这与后者矛盾。

(c) 当 $x \rightarrow 0+$ 时, (3.4.20) 中没有一项相对于其他两项是可忽略的, 留给读者证明矛盾。

(d) 当 $x \rightarrow 0+$ 时 $-2x^{-3/2}\varepsilon' \sim 3x^{-2}/16$ 且 $\varepsilon'' \ll 2x^{-3/2}\varepsilon'$ 。对前一个渐近关系积分得 $\varepsilon(x) \sim -3x^{1/2}/16 (x \rightarrow 0+)$, 满足后者条件。

因此存在 $\varepsilon_1 \ll 3x^{1/2}/16 (x \rightarrow 0+)$ 满足:

$$w(x) = 1 - 3x^{1/2}/16 + \varepsilon_1(x),$$

例 3.27 (续)

我们现在可以通过类似刚才的论证来确定 $\varepsilon_1(x)$ 的主导行为:

$$\varepsilon_1(x) = -15x/512 + \varepsilon_2(x), \quad \varepsilon_2 \ll 15x/512, \quad x \rightarrow 0+.$$

显然, 我们可以永远继续这个过程, 并生成 $w(x)$ 的一个级数表示。索性将完整的形式级数表示 (3.4.18) (其中 $\alpha = \frac{1}{2}$, $w(x) = \sum_{n=0}^{\infty} a_n x^{n/2}$ ($a_0 = 1$)) 直接代入 (3.4.19)。我们得到

$$\sum_{n=1}^{\infty} \frac{n}{2} \left(\frac{n}{2} - 1 \right) a_n x^{n/2-2} + \frac{3}{2} \sum_{n=1}^{\infty} \frac{n}{2} a_n x^{n/2-2} - 2 \sum_{n=1}^{\infty} \frac{n}{2} a_n x^{n/2-5/2} - \frac{3}{16} \sum_{n=0}^{\infty} a_n x^{n/2-2} = 0.$$

由此得递推关系:

$$a_{n+1} = \frac{(2n-1)(2n+3)}{16(n+1)} a_n, \quad n = 0, 1, 2, \dots \quad (3.4.21)$$

故有通项公式为:

$$a_n = -[\Gamma(n - \frac{1}{2})\Gamma(n + \frac{3}{2})]/\pi 4^n n!, \quad n = 0, 1, 2, \dots$$

由此完成 (3.4.4a) 主导行为的解之局部分析:

$$y(x) \sim -c_1 x^{3/4} e^{2x^{-1/2}} \sum_{n=0}^{\infty} \frac{\Gamma(n - \frac{1}{2})\Gamma(n + \frac{3}{2})}{\pi 4^n n!} x^{n/2}, \quad x \rightarrow 0+. \quad (3.4.22a)$$

级数对所有 x 发散, 但作为渐近级数在 $x \rightarrow 0+$ 给出精确近似。类似分析 (3.4.4b) 得:

$$y(x) \sim -c_2 x^{3/4} e^{-2x^{-1/2}} \sum_{n=0}^{\infty} (-1)^n \frac{\Gamma(n - \frac{1}{2})\Gamma(n + \frac{3}{2})}{\pi 4^n n!} x^{n/2}, \quad x \rightarrow 0+. \quad (3.4.22b)$$

例 3.28 (不规则奇点附近的局部行为)

在 (3.4.2) 作指数替换 $y = e^S$ 得:

$$x^2 S'' + x^2 (S')^2 + (1 + 3x)S' + 1 = 0. \quad (3.4.23)$$

忽略 $x^2 S''$ 和 $3xS'$, 解 $S' + 1 \sim -x^2 (S')^2$ 得两根:

$$S' \sim -1 \quad (x \rightarrow 0+), \quad (3.4.24a)$$

$$S' \sim -x^{-2} \quad (x \rightarrow 0+). \quad (3.4.24b)$$

积分 (3.4.24a) 得 $y(x) \sim c_1$, 积分 (3.4.24b) 得 $S \sim x^{-1}$, 代入 $S(x) = x^{-1} + C(x)$ ($C \ll x^{-1}$) 得:

$$x^2 C'' + x^2 (C')^2 - (1 - 3x)C' - x^{-1} + 1 = 0. \quad (3.4.25)$$

渐近近似给出

$$C'(x) \sim -x^{-1} \quad (3.4.26)$$

积分得 $C(x) \sim -\ln x$ 。

例 3.29 (n 阶薛定谔方程在奇点附近的局部行为)

对方程

$$\frac{d^n y}{dx^n} = Q(x)y \quad (3.4.27)$$

在奇点 x_0 附近, 指数替换 $y = e^S$ 及渐近近似 $d^k S/dx^k \ll (S')^k$ ($k \geq 2$) 得:

$$(S')^n \sim Q(x) \quad (x \rightarrow x_0).$$

故 $S(x) \sim \omega \int^x [Q(t)]^{1/n} dt$, 其中 $\omega^n = 1$ 。主导行为为:

$$y(x) \sim c[Q(x)]^{(1-n)/2n} \exp\left(\omega \int^x [Q(t)]^{1/n} dt\right), \quad x \rightarrow x_0. \quad (3.4.28)$$

当 $x_0 \neq \infty$ 时要求 $|(x-x_0)^n Q(x)| \rightarrow \infty$; 当 $x_0 = \infty$ 时要求 $|x^n Q(x)| \rightarrow \infty$ 。

此公式是 WKB 理论基石, 示例:

$$(a) \ y'' = y/x^5: \ y(x) \sim cx^{5/4}e^{\pm 2x^{-3/2}/3} \quad (x \rightarrow 0+)$$

$$(b) \ y''' = xy: \ y(x) \sim cx^{-1/3}e^{3\omega x^{4/3}/4} \quad (x \rightarrow +\infty) \quad (\omega^3 = 1)$$

$$(c) \ d^4 y/dy^4 = (x^4 + \sin x)y: \ y(x) \sim cx^{-3/2}e^{\omega x^2/2} \quad (x \rightarrow +\infty) \quad (\omega = \pm 1, \pm i)$$

3.5 无穷处的非正则奇点

许多著名微分方程在 ∞ 处存在不规则奇点。应用中, 掌握这些方程解在大 $|x|$ 时的局部行为至关重要。前几节方法建议通过反演变换 $x = 1/t$ 将奇点映射至原点, 再用 3.4 节方法分析新方程在原点处的行为。但更高效的方式是直接推导原方程在 ∞ 处有效的渐近级数表示。

渐近幂级数**定义 3.7**

若对任意 N 满足 $y(x) - \sum_{n=0}^N a_n(x-x_0)^n \ll (x-x_0)^N$ ($x \rightarrow x_0$), 则称幂级数 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ 在 $x \rightarrow x_0$ 时渐近于函数 $y(x)$, 记为 $y(x) \sim \sum_{n=0}^{\infty} a_n(x-x_0)^n$ ($x \rightarrow x_0$)。

我们有等价定义, 若 a_M 是 a_N 后首个非零系数, 则 $y(x) - \sum_{n=0}^N a_n(x-x_0)^n \sim a_M(x-x_0)^M$ ($x \rightarrow x_0$)。当非零系数无限时, 渐近幂级数等价于无限渐近关系序列。

将其推广到分数幂: 级数 $\sum_{n=0}^{\infty} a_n(x-x_0)^{m_n}$ ($m_n > 0$) 渐近于 $y(x)$ 的条件是: 对任意 N 有 $y(x) - \sum_{n=0}^N a_n(x-x_0)^{m_n} \ll (x-x_0)^{M_N}$ ($x \rightarrow x_0$)。

当 $x_0 = \infty$ 时: $y(x) \sim \sum_{n=0}^{\infty} a_n x^{-m_n}$ ($x \rightarrow \infty$) 需满足 $y(x) - \sum_{n=0}^N a_n x^{-m_n} \ll x^{-M_N}$ ($x \rightarrow \infty$)。

注意非所有函数可展为渐近幂级数, 且若存在则唯一。所有极限存在是函数具有渐近幂级数展开的条件。

例 3.30 (泰勒级数在大 x 时的行为)

考虑 Taylor 定义的函数:

$$y(x) = \sum_{n=0}^{\infty} \frac{x^n}{(n!)^2}. \quad (3.5.1)$$

该级数对所有 x 收敛, 但小 x 时收敛快, 大 x 时收敛慢. 为求大 x 行为, 先构造 $y(x)$ 满足的二阶微分方程. 由 $y' = \sum_{n=1}^{\infty} x^{n-1}/[n!(n-1)!]$ 和 $(xy')' = \sum_{n=1}^{\infty} x^{n-1}/[(n-1)!]^2 = y$ 得:

$$xy'' + y' = y. \quad (3.5.2)$$

此方程在 ∞ 处有不规则奇点. 代入 $y = e^S$ 求主导行为:

$$xS'' + x(S')^2 + S' = 1. \quad (3.5.3)$$

假设 $xS'' \ll x(S')^2$ ($x \rightarrow +\infty$), 得渐近方程 $x(S')^2 \sim 1 - S'$. 解出:

$$S'(x) \sim [-1 \pm (1 + 4x)^{1/2}]/2x, \quad x \rightarrow +\infty.$$

大 x 时近似为:

$$S'(x) \sim \pm x^{-1/2}, \quad x \rightarrow +\infty.$$

积分得:

$$S(x) = \pm 2x^{1/2} + C(x), \quad (3.5.4)$$

其中 $C(x) \ll 2x^{1/2}$ ($x \rightarrow +\infty$). 因 (3.5.1) 系数全正, $y(x)$ 单调增, 故取正号: $S(x) = 2x^{1/2} + C(x)$. 代入 (3.5.3) 得 $C(x)$ 方程:

$$xC'' + x(C')^2 + (2x^{1/2} + 1)C' + \frac{1}{2}x^{-1/2} = 0.$$

由大 x 条件 $1 \ll 2x^{1/2}$, $C' \ll x^{-1/2}$, $C'' \ll x^{-3/2}$ 得渐近方程:

$$C' \sim -\frac{1}{4}x^{-1}, \quad x \rightarrow +\infty,$$

解为:

$$C(x) \sim -\frac{1}{4} \ln x, \quad x \rightarrow +\infty.$$

故 $y(x)$ 主导行为:

$$y(x) \sim cx^{-1/4}e^{2x^{1/2}}, \quad x \rightarrow +\infty. \quad (3.5.5)$$

可计算出 $c = \frac{1}{2}\pi^{-1/2}$, 比值:

$$\frac{1}{2}\pi^{-1/2}x^{-1/4}e^{2x^{1/2}}/y(x) \quad (3.5.6a)$$

在 $x \rightarrow +\infty$ 时趋近 1, 且在小 x 时也接近 1, 表明主导行为在大区间有效. 部分和比值:

$$\left[\sum_{n=0}^{10} \frac{x^n}{(n!)^2} \right] / y(x). \quad (3.5.6b)$$

是良好多项式近似.

例 3.31 (修正 Bessel 函数在大 x 时的行为)

修正 Bessel 方程

$$x^2 y'' + xy' - (x^2 + v^2)y = 0$$

在 ∞ 处有不规则奇点, 解的主导行为:

$$y(x) \sim c_1 x^{-1/2} e^x, \quad x \rightarrow +\infty, \quad (3.5.7a)$$

$$y(x) \sim c_2 x^{-1/2} e^{-x}, \quad x \rightarrow +\infty. \quad (3.5.7b)$$

上主导行为虽与 v 无关, 但完整渐近展开的系数含 v , 代入剥离形式:

$$y(x) = c_1 x^{-1/2} e^x w(x) \quad (3.5.8a)$$

得 $w(x)$ 的方程:

$$x^2 w'' + 2x^2 w' + \left(\frac{1}{4} - v^2\right)w = 0.$$

设 $w(x) = 1 + \varepsilon(x)$, 由方程:

$$x^2 \varepsilon'' + 2x^2 \varepsilon' + \left(\frac{1}{4} - v^2\right)\varepsilon + \frac{1}{4} - v^2 = 0,$$

利用近似:

$$\left(\frac{1}{4} - v^2\right)\varepsilon \ll \frac{1}{4} - v^2, \quad x^2 \varepsilon'' \ll x^2 \varepsilon', \quad x \rightarrow +\infty$$

得渐近方程:

$$2x^2 \varepsilon'(x) \sim v^2 - \frac{1}{4}, \quad x \rightarrow +\infty.$$

因 $\varepsilon(x) \ll 1$, 积分常数 $c = 0$, 解得:

$$\varepsilon(x) \sim \left(\frac{1}{8} - \frac{1}{2}v^2\right)x^{-1}, \quad x \rightarrow +\infty.$$

此分析可迭代得 $w(x)$ 的完整渐近展开. 代入 $w(x) \sim \sum_{n=0}^{\infty} a_n x^{-n}$ ($a_0 = 1$) 方程得:

$$0 \sim \sum_{n=0}^{\infty} n(n+1)a_n x^{-n} - 2 \sum_{n=0}^{\infty} n a_n x^{1-n} + \left(\frac{1}{4} - v^2\right) \sum_{n=0}^{\infty} a_n x^{-n}, \quad x \rightarrow +\infty.$$

于是有:

$$\left[\left(n + \frac{1}{2}\right)^2 - v^2\right] a_n - 2(n+1)a_{n+1} = 0, \quad n = 0, 1, 2, \dots$$

由 $a_0 = 1$ 解递推得:

$$w(x) \sim 1 - \frac{(4v^2 - 1^2)}{1! 8x} + \frac{(4v^2 - 1^2)(4v^2 - 3^2)}{2! (8x)^2} - \dots, \quad x \rightarrow +\infty. \quad (3.5.9a)$$

收敛半径为 0. 同理得 (3.5.7b) 的解为

$$y(x) = c_2 x^{-1/2} e^{-x} w(x) \quad (3.5.8b)$$

$$w(x) \sim 1 + \frac{(4v^2 - 1^2)}{1! 8x} + \frac{(4v^2 - 1^2)(4v^2 - 3^2)}{2! (8x)^2} + \dots, \quad x \rightarrow +\infty. \quad (3.5.9b)$$

例 3.32 (渐近级数数值计算 $I_v(x)$)

上例得到 $I_v(x)$ 在 $x \rightarrow +\infty$ 时的渐近展开。取 $c_1 = (2\pi)^{-1/2}$ 和 $v = 5$ 代入 (3.5.8a) 和 (3.5.9a), 得 $I_s(x)$ 展开式:

$$I_s(x) \sim (2\pi)^{-1/2} e^x x^{-1/2} \left[1 - \frac{(100-1)}{1!8x} + \frac{(100-1)(100-9)}{2!(8x)^2} - \dots \right], \quad x \rightarrow +\infty. \quad (3.5.10)$$

此级数为渐近级数 (未证明但成立)。由本节定义, 若在 x^{-N} 项截断, 余项 (函数值与部分和之差) 渐近于级数第 $N+2$ 项。当 x 固定时, 首项被略项仅为误差估计。

数值计算规则: 观察级数项——通常项先递减, 但因级数发散最终递增。对每个 x 定位最小项, 将其前所有项相加 (不包括最小项)。此有限和给出最优估计, 因下一项 (误差估计) 是级数最小项, 称为最优渐近近似。

与收敛幂级数 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ 对比: 收敛级数可通过增加项提高精度; 发散渐近级数对固定 x 存在精度上限, 偏离最优项数会降低精度。若不满足精度, 则需增大 x 。

例 3.33 (抛物柱函数在大 x 时的行为)

考察抛物柱方程

$$y'' + \left(v + \frac{1}{2} - \frac{1}{4}x^2\right)y = 0 \quad (3.5.11)$$

的解 $y(x)$ 在 $x \rightarrow +\infty$ 时的行为, 其中 v 为参数。因方程在 ∞ 处有不规则奇点, 作指数替换 $y = e^S$ 得:

$$S'' + (S')^2 + v + \frac{1}{2} - \frac{1}{4}x^2 = 0.$$

由近似 $S'' \ll (S')^2, v + \frac{1}{2} \ll \frac{1}{4}x^2$ ($x \rightarrow +\infty$) 得渐近方程 $(S')^2 \sim \frac{1}{4}x^2$, 解为:

$$S(x) \sim \pm \frac{1}{4}x^2, \quad x \rightarrow +\infty.$$

故 $y(x)$ 的控制因子可能为 $e^{\pm x^2/4}$, 可计算出可能的主导行为为:

$$y(x) \sim c_1 x^{-v-1} e^{x^2/4}, \quad x \rightarrow +\infty, \quad (3.5.12a)$$

$$y(x) \sim c_2 x^v e^{-x^2/4}, \quad x \rightarrow +\infty. \quad (3.5.12b)$$

抛物柱函数 (parabolic cylinder) $D_v(x)$ 定义为满足 (3.5.11) 且渐近行为为 (3.5.12b) 取 $c_2 = 1$ 的解。该条件唯一确定 $D_v(x)$ (因 c_1 必为 0)。对其进一步应用需知 $D_v(x)$ 在 $x \rightarrow +\infty$ 的完整渐近展开。代入:

$$D_v(x) = x^v e^{-x^2/4} w(x) \quad (3.5.13)$$

可解得:

$$w(x) \sim 1 - \frac{v(v-1)}{2^1 \cdot 1! \cdot x^2} + \frac{v(v-1)(v-2)(v-3)}{2^2 \cdot 2! \cdot x^4} - \dots, \quad x \rightarrow +\infty. \quad (3.5.14)$$

若 $v \notin \{0, 1, 2, 3, \dots\}$, 由比值审敛法得级数收敛半径 $R = 0$ 。但当 $v = N$ 时级数有限且等于 Hermite 多项式乘衰减指数:

$$D_N(x) = e^{-x^2/4} H_N(x).$$

例 3.34 (Airy 函数在大 x 时的行为)

Airy 方程

$$y'' = xy \quad (3.5.15)$$

在 ∞ 处有不规则奇点。由 (3.4.28) 式取 $n = 2$, $Q(x) = x$ 得大 x 时解的主导行为:

$$y(x) \sim c_1 x^{-1/4} e^{-2x^{3/2}/3}, \quad x \rightarrow +\infty, \quad (3.5.16a)$$

$$y(x) \sim c_2 x^{-1/4} e^{2x^{3/2}/3}, \quad x \rightarrow +\infty. \quad (3.5.16b)$$

定义 Airy 函数 $\text{Ai}(x)$ 与 $\text{Bi}(x)$ 分别是唯一满足上两式的解 (取 $c_1 = \frac{1}{2}\pi^{-1/2}$, $c_2 = \pi^{-1/2}$):

$$\text{Ai}(x) \sim \frac{1}{2}\pi^{-1/2} x^{-1/4} e^{-2x^{3/2}/3}, \quad x \rightarrow +\infty. \quad (3.5.17a)$$

$$\text{Bi}(x) \sim \pi^{-1/2} x^{-1/4} e^{2x^{3/2}/3}, \quad x \rightarrow +\infty. \quad (3.5.17b)$$

前者条件唯一, 后者条件不唯一。完整渐近展开需剥离主导行为。对 $\text{Ai}(x)$ 代入:

$$y(x) = x^{-1/4} e^{-2x^{3/2}/3} w(x). \quad (3.5.18)$$

则 $w(x)$ 满足方程:

$$x^2 w'' - (2x^{5/2} + \frac{1}{2}x)w' + \frac{5}{16}w = 0 \quad (3.5.19)$$

且 $w(x) \rightarrow 1$ ($x \rightarrow +\infty$)。我们这里不采用常规方法, 而直接假设渐近展开

$$w(x) \sim \sum_{n=0}^{\infty} a_n x^{\alpha n} \quad (\alpha < 0, a_0 = 1).$$

代入 (3.5.19) 得:

$$\frac{5}{16}a_0 - 2\alpha a_1 x^{\alpha+3/2} + (\alpha^2 - \frac{3}{2}\alpha + \frac{5}{16})a_1 x^{\alpha} - \dots \sim 0.$$

由 $a_0 = 1$ 和 $\alpha < 0$ 得 $\alpha = -\frac{3}{2}$, $a_1 = -\frac{5}{48}$ 。一般地, 令 $(x^{-3/2})^n$ 系数为零得递推关系:

$$a_{n+1} = -\frac{3}{4} \frac{(n + \frac{5}{6})(n + \frac{1}{6})}{n+1} a_n, \quad n \geq 0.$$

解为:

$$a_n = \frac{1}{2\pi} \left(-\frac{3}{4}\right)^n \frac{\Gamma(n + \frac{5}{6})\Gamma(n + \frac{1}{6})}{n!}.$$

3.6 非齐次线性方程的局部分析

对于非齐次微分方程解的局部行为的方法, 我们只需找到一个特解的局部行为即可, 因为通解的局部行为可通过叠加关联齐次方程解的行为得到。我们倾向于直接对微分方程应用主导平衡法。需要考虑几种情形。

若 x_0 是关联齐次微分方程的常点, 且是非齐次项的解析点。在这种情况下, 通解具有泰勒级数表示。

例 3.35

考虑在 $x = 0$ 附近的微分方程

$$y' + xy = x^3$$

若 x_0 是微分方程的常点，但不是非齐次项的解析点，则尽管齐次方程的所有解均可展开为泰勒级数，非齐次方程的特解却不具有泰勒级数展开。在这种情况下，我们使用主导平衡法确定特解的行为。

例 3.36

考虑在 $x = 0$ 附近的微分方程

$$y' + xy = 1/x^4$$

主导平衡法同样适用于关联齐次方程的奇点处。

例 3.37

考虑在 $x = \infty$ 附近的微分方程

$$y'' = xy - 1$$

使用主导平衡法时可能会遇到一个略显有趣的细微之处：所有两项平衡都可能是矛盾的。当这种情况发生时，我们必须考虑三项或更多项之间的可能主项平衡。

例 3.38

考虑在 $x = 0$ 附近的微分方程

$$y' - y/x = (\cos x)/x^2$$

3.7 渐近关系

本节详细研究渐近关系。

振荡函数的渐近关系

为说明渐近关系的一个微妙之处，我们考察 Airy 方程 $y'' = xy$ 在大 $|x|$ 时解的主导行为。我们可以计算得到 Airy 函数 $\text{Ai}(x)$ 和 $\text{Bi}(x)$ 的主导行为为

$$\text{Ai}(x) \sim \frac{1}{2} \pi^{-1/2} x^{-1/4} e^{-2x^{3/2}/3}, \quad x \rightarrow +\infty \quad (3.7.1)$$

$$\text{Bi}(x) \sim \pi^{-1/2} x^{-1/4} e^{2x^{3/2}/3}, \quad x \rightarrow +\infty \quad (3.7.2)$$

使用代换 $y = e^S$ ，得到 $S'' + (S')^2 = x$ 。忽略 S'' ，得到渐近微分方程 $(S')^2 \sim x$ ，由此推出 $S \sim \pm \frac{2}{3} i (-x)^{3/2}$ 。这给出了 y 在 $x \rightarrow -\infty$ 时主导行为的控制因子 $\exp[\pm \frac{2}{3} i (-x)^{3/2}]$ ，进而我们可以得到主导行为

$$y \sim c(-x)^{-1/4} \exp\left[\pm \frac{2}{3} i (-x)^{3/2}\right], \quad x \rightarrow -\infty. \quad (3.7.3)$$

由于 Airy 函数在 x 为实数时是实值的，我们必须将 (3.7.3) 中两个主导行为拆分成实线性组合：

$$y(x) \sim c_1(-x)^{-1/4} \sin \left[\frac{2}{3}(-x)^{3/2} \right] + c_2(-x)^{-1/4} \cos \left[\frac{2}{3}(-x)^{3/2} \right], \quad x \rightarrow -\infty.$$

事实上， $\text{Ai}(x)$ 在大负 x 时的图像非常接近

$$\pi^{-1/2}(-x)^{-1/4} \sin \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right]$$

的图像，而 $\text{Bi}(x)$ 的图像非常接近

$$\pi^{-1/2}(-x)^{-1/4} \cos \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right]$$

的图像。然而写成

$$\text{Ai}(x) \sim \pi^{-1/2}(-x)^{-1/4} \sin \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right], \quad x \rightarrow -\infty \quad (3.7.4)$$

和

$$\text{Bi}(x) \sim \pi^{-1/2}(-x)^{-1/4} \cos \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right], \quad x \rightarrow -\infty \quad (3.7.5)$$

是错误的，因为对于大的负 x ，左右两边的零点并不完全重合。接下来的两个例子将说明如何解决零点不重合的问题。

例 3.39

考虑函数 $\sin x$ 和 $\sin(x + x^{-1})$ ，我们倾向于说这两个函数当 $x \rightarrow \infty$ 时是渐近的。然而，根据渐近关系的定义，这些函数并非渐近的。

问题在于如何描述 $\sin x$ 和 $\sin(x + x^{-1})$ 行为的高度相似性，同时避免除以零。一种方法是指出，当 $x \rightarrow \infty$ 时，正弦函数的自变量是渐近的： $x + x^{-1} \sim x$ ，这避开了零点不重合的问题。

例 3.40

考虑函数 $\sin x$ 和 $(1 + x^{-1})\sin(x + x^{-1})$ ，将其写成

$$(1 + x^{-1})\sin(x + x^{-1}) = w_1(x)\sin x + w_2(x)\cos x,$$

其中 $w_1(x) = (1 + x^{-1})\cos x^{-1}$ ， $w_2(x) = (1 + x^{-1})\sin x^{-1}$ 。当 $x \rightarrow \infty$ 时， $w_1(x) \sim 1$ ， $w_2(x) \ll 0$ 。

总的来说，两个不含相同零点的函数并非渐近的。上例提示我们将 Airy 函数写成

$$y(x) = w_1(x)(-x)^{-1/4} \sin \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right] + w_2(x)(-x)^{-1/4} \cos \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right] \quad (3.7.6)$$

然后去考虑 $w_1(x)$ 与 $w_2(x)$ 的行为，我们希望 Airy 函数 $\text{Ai}(x)$ 可以由 $x \rightarrow -\infty$ 时， $w_1(x) \sim \pi^{-1/2}$ ， $w_2(x) \ll w_1(x)$ 来表示， $\text{Bi}(x)$ 可以由 $x \rightarrow -\infty$ 时， $w_2(x) \sim \pi^{-1/2}$ ， $w_1(x) \ll w_2(x)$ 来表示。当然，这些行为中的常数 $\pi^{-1/2}$ 无法仅通过 Airy 方程的局部分析来确定。

将 (3.7.6) 代入 Airy 方程 $y'' = xy$, 经化简后得到:

$$\begin{aligned} & \left[w_1'' - \frac{1}{2}x^{-1}w_1' + 2(-x)^{1/2}w_2' + \frac{5}{16}x^{-2}w_1 \right] (-x)^{-1/4} \sin \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right] \\ & + \left[w_2'' - \frac{1}{2}x^{-1}w_2' - 2(-x)^{1/2}w_1' + \frac{5}{16}x^{-2}w_2 \right] (-x)^{-1/4} \cos \left[\frac{2}{3}(-x)^{3/2} + \frac{1}{4}\pi \right] = 0. \end{aligned} \quad (3.7.7)$$

(3.7.6) 中 $w_1(x)$ 和 $w_2(x)$ 的定义仅给出了一个约束条件, 我们选择第二个约束条件为

$$w_1'' - \frac{1}{2}x^{-1}w_1' + 2(-x)^{1/2}w_2' + \frac{5}{16}x^{-2}w_1 = 0. \quad (3.7.8)$$

则式 (3.7.7) 意味着

$$w_2'' - \frac{1}{2}x^{-1}w_2' - 2(-x)^{1/2}w_1' + \frac{5}{16}x^{-2}w_2 = 0. \quad (3.7.9)$$

由此, 我们得到了不含 (3.7.6) 中的振荡项的关于 w_1 和 w_2 的一对微分方程, 分析可知, 存在当 $x \rightarrow -\infty$ 时趋近于常数的解, 而 $w_1(x)$ 和 $w_2(x)$ 的完整渐近展开式如下:

$$w_1(x) \sim \sum_{n=0}^{\infty} a_n (-x)^{-3n/2}, \quad x \rightarrow -\infty, \quad (3.7.10)$$

$$w_2(x) \sim \sum_{n=0}^{\infty} b_n (-x)^{-3n/2}, \quad x \rightarrow -\infty. \quad (3.7.11)$$

将这些级数代入 (3.7.8) 和 (3.7.9), 解得

$$a_{2n} = a_0 (-1)^n c_{2n}, \quad n = 0, 1, \dots,$$

$$a_{2n+1} = b_0 (-1)^n c_{2n+1}, \quad n = 0, 1, \dots,$$

$$b_{2n} = b_0 (-1)^n c_{2n}, \quad n = 0, 1, \dots,$$

$$b_{2n+1} = a_0 (-1)^{n+1} c_{2n+1}, \quad n = 0, 1, \dots,$$

其中

$$c_n = \frac{(2n+1)(2n+3)\cdots(6n-1)}{144^n n!} = \frac{1}{2\pi} \left(\frac{3}{4} \right)^n \frac{\Gamma(n + \frac{5}{6}) \Gamma(n + \frac{1}{6})}{n!}.$$

Airy 函数 $\text{Ai}(x)$ 当 $x \rightarrow -\infty$ 时的渐近行为可由 $a_0 = \pi^{-1/2}$, $b_0 = 0$ 表示, 其中

$$w_1(x) \sim \pi^{-1/2} \sum_{n=0}^{\infty} c_{2n} x^{-3n}, \quad x \rightarrow -\infty, \quad (3.7.12a)$$

$$w_2(x) \sim -\pi^{-1/2} (-x)^{-3/2} \sum_{n=0}^{\infty} c_{2n+1} x^{-3n}, \quad x \rightarrow -\infty. \quad (3.7.12b)$$

另一方面, $\text{Bi}(x)$ 当 $x \rightarrow -\infty$ 时的渐近行为可由 $a_0 = 0$, $b_0 = \pi^{-1/2}$ 表示, 其中

$$w_1(x) \sim -\pi^{-1/2} (-x)^{-3/2} \sum_{n=0}^{\infty} c_{2n+1} x^{-3n}, \quad x \rightarrow -\infty, \quad (3.7.13a)$$

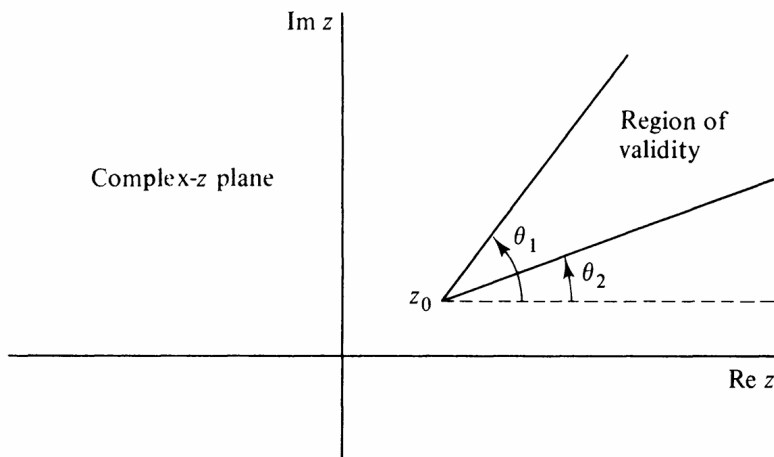
$$w_2(x) \sim \pi^{-1/2} \sum_{n=0}^{\infty} c_{2n} x^{-3n}, \quad x \rightarrow -\infty. \quad (3.7.13b)$$

在许多其他情况下, 必须使用此类分析来寻找渐近表示。

复平面上的渐近关系

到目前为止, 我们仅在实轴上讨论了渐近关系, 但要充分理解渐近关系的结构, 必须在复平面中对其进行分析。将渐近关系的定义推广到复函数并非易事, 若简单地定义 $f(z) \sim g(z) (z \rightarrow z_0)$ 为 $\lim_{z \rightarrow z_0} f(z)/g(z) = 1$, 其中极限 $z \rightarrow z_0$ 沿复平面中趋近于 z_0 的任意路径取值, 则会遇到严重困难。一个令人满意的复渐近关系定义必须是路径无关且唯一的, 这需要小心排除那些在趋近 z_0 过程中绕 z_0 旋转的复路径。此类路径可能破坏渐近关系的单侧性, 并导致非唯一极限。

消除这类路径的最简单方法是要求路径完全位于复平面的一个扇形区域内。若扇形的边界在 z_0 邻域内为光滑曲线, 则在局部可由它们在 z_0 处的切线近似。这些切线之间的夹角 $\theta_2 - \theta_1$ 称为扇形的张角。当 $z \rightarrow z_0$ 时两个函数在复平面内渐近的表述, 必须满足 z 沿位于扇形区域内的路径趋近于 z_0 , 其中张角的大小取决于渐近的函数。



以下是复渐近关系的一些例子。

例 3.41

$$\sinh(z^{-1}) \sim \frac{1}{2}e^{z^{-1}}, \quad z \rightarrow 0; \quad |\arg z| < \frac{1}{2}\pi, \quad (3.7.14)$$

$$\sinh(z^{-1}) \sim -\frac{1}{2}e^{-z^{-1}}, \quad z \rightarrow 0; \quad \frac{1}{2}\pi < \arg z < \frac{3}{2}\pi. \quad (3.7.15)$$

例 3.42

$\text{Bi}(z)$ 是 Airy 方程 $y''(z) = zy(z)$ 的一个解, 当 $z \rightarrow +\infty$ ($\text{Im } z = 0$) 时指数增长, 其在复平面中的主导行为为

$$\text{Bi}(z) \sim \pi^{-1/2} z^{-1/4} e^{2z^{3/2}/3}, \quad z \rightarrow \infty; \quad |\arg z| < \frac{\pi}{3}. \quad (3.7.16)$$

例 3.43

渐近关系 $\sin z \sim z$ ($z \rightarrow 0$) 在复平面中沿任意路径 $z \rightarrow 0$ 时均有效。

例 3.44

微分方程 $x^3 y'' = y$ 的解在 $x \rightarrow 0^+$ 时的主导行为 $y(z) \sim c_1 z^{1/4} e^{2z^{-1/2}}$ (3.4.4a) 在 $|\arg z| < \pi$ 内有效。当 $|\arg z| = \pi$ 时, 该渐近关系失效。

斯托克斯现象与次支配性

如果两个函数 $f(z)$ 和 $g(z)$ 在 $z \rightarrow z_0$ 时是渐近的, 那么差值 $f(z) - g(z)$ 与 $g(z)$ 相比是很小的, 即 $\lim_{z \rightarrow z_0} [f(z) - g(z)]/g(z) = 0$ 。这个差值 $f(z) - g(z)$ 正是我们在构建一个渐近关系时所忽略的; 如果我们考虑等式: $f(z) = g(z) + [f(z) - g(z)]$, 当 z 位于扇形区域内部时, $f(z) - g(z)$ 被称为次主导的或隐性的, 而 $f(z)$ 或 $g(z)$ 被称为主导的。

当 z 接近渐近线时, 次主导项 $f - g$ 与主导项 g 相比会增长。扇形区域的边缘称为**斯托克斯线**, 在其上不再能有效地区分主导项和次主导项, 因为 f 和 $f - g$ 具有相同的数量级。当 z 穿过斯托克斯线时, 主导项和次主导项交换了它们的身份, 这称为**斯托克斯现象**。

例 3.45

函数 $\sinh(z^{-1})$ 及其渐近行为 $\frac{1}{2}e^{z^{-1}}$ 在右半平面 $\operatorname{Re} z > 0$ 中呈指数级增长, 两者差值是 $-\frac{1}{2}e^{-z^{-1}}$, 它在右半平面 $\operatorname{Re} z > 0$ 中是指数级小的。如果我们将 z 旋转到左半平面, $\sinh(z^{-1}) - \frac{1}{2}e^{z^{-1}} (\rightarrow 0)$ 的渐近关系就不再有效, 因为次主导的误差项 $-\frac{1}{2}e^{-z^{-1}}$ 不再小。通常, 任何渐近关系有效的扇形的张角是左右两侧相对误差最小的最大角度。

注意, 当 $z \rightarrow 0$ 时, $\sinh(z^{-1})$ 有两个可能的主导渐近行为, 它们在斯托克斯线上时的量级上是最不相等的, 而在纯实数时的量级上是最相等的。主导行为最不相等的线被称为**反斯托克斯线**。

斯托克斯现象是指数函数的一个性质。如果次主导项比主导项小一个幂次 $z - z_0$, 那么当 $z \rightarrow z_0$ 在复平面的所有方向上, 就不会有斯托克斯现象。

如果一个二阶微分方程的解的主导行为是 $e^{S_1(z)}$ 和 $e^{S_2(z)}$ 当 $z \rightarrow z_0$ 时, 那么斯托克斯线是曲线的渐近线

$$\operatorname{Re}[S_1(z) - S_2(z)] = 0$$

反斯托克斯线是曲线的渐近线

$$\operatorname{Im}[S_1(z) - S_2(z)] = 0$$

注意, 我们将斯托克斯线和反斯托克斯线定义为渐近线, 因为这些渐近线决定了扇形区域的张角。斯托克斯和反斯托克斯线是函数行为的局部属性, 它们仅在 z_0 的邻域中有意义。

例 3.46 (Airy 方程的斯托克斯线)

Airy 方程 $y'' - zy = 0$ 的解在大 z 时的控制因子由 $e^{\pm 2z^{3/2}/3}$ 给出。因此, 斯托克斯线出现在 $\operatorname{Re} z^{3/2} = 0$, 即 $\arg z = \pm \frac{\pi}{3}, \pi (|z| \rightarrow \infty)$; 反斯托克斯线出现在 $\operatorname{Im} z^{3/2} = 0$, 即 $\arg z = 0, \pm \frac{2\pi}{3} (|z| \rightarrow \infty)$ 。

$\operatorname{Bi}(z)$ 的主导渐近行为在斯托克斯线 $\arg z = \pm \frac{\pi}{3}$ 之间 (3.7.16) 是有效的。另一方面, $\operatorname{Ai}(z)$ 的主导渐近行为在一个更大的楔形区域内有效。

$$\operatorname{Ai}(z) \sim \frac{1}{2} \pi^{-1/2} z^{-1/4} e^{-\frac{2}{3} z^{3/2}}, \quad z \rightarrow \infty; \quad |\arg z| < \pi. \quad (3.7.17)$$

例 3.47 (抛物柱函数方程的斯托克斯线)

抛物柱函数方程 $y'' + (v + \frac{1}{2} - \frac{1}{4}z^2)y = 0$ 解的行为的控制因子是 $e^{\pm z^2/4}$ ($z \rightarrow \infty$)。因此, 斯托克斯线出现在 $\arg z = \pm \frac{\pi}{4}, \pm \frac{3\pi}{4}$ ($|z| \rightarrow \infty$), 反斯托克斯线出现在 $\arg z = 0, \pm \frac{\pi}{2}, \pi$ ($|z| \rightarrow \infty$)。抛物柱函数 $D_v(z)$ 的主导渐近行为是

$$D_v(z) \sim z^v e^{-z^2/4}, \quad z \rightarrow \infty; \quad |\arg z| < \frac{3\pi}{4}. \quad (3.7.18)$$

上式沿正实轴指数衰减, 在斯托克斯线 $\arg z = \pm \frac{3\pi}{4}$ 处失效。

一般来说, 当 z 沿着最靠近正实轴的斯托克斯线旋转时, 衰减最快的解 (例如 $\text{Ai}(z)$ 或 $D_v(z)$) 会增长, 对于这类解, 将其主导渐近行为解析延拓到最靠近正实轴的斯托克斯线之外是不正确的; 继续解析延拓那些在正实轴附近增长的解 (例如 $\text{Bi}(z)$) 的主导渐近行为, 越过最靠近实轴的斯托克斯线也是不正确的。下一节中, 我们会展示如何将一个完整的渐近级数延拓越过斯托克斯线, 解释了为什么一些渐近关系在斯托克斯线之外仍然有效, 并讨论了在 $x \rightarrow x_0+$ 和 $x \rightarrow x_0-$ 时具有不同行为的解之间的联系。

正如我们反复说过的, 使用渐近近似的原因是用更简单的表达式来取代复杂的超越函数, 为这种便利付出的代价之一是必须处理这些超越函数的复杂性。斯托克斯现象不是 $\text{Ai}(z)$ 的内在属性, 而是用于在渐近近似中逼近 $\text{Ai}(z)$ 的函数的属性, 其反映了指数函数在渐近近似中存在。关系 $\text{Ai}(z) \sim \text{Ai}(z)$ 在 $z \rightarrow \infty$ 时在所有方向上都有效, 它不表现出斯托克斯现象, 但这个关系不包含任何有用的信息。相比之下, 关系

$$\text{Ai}(z) \sim \frac{1}{2\sqrt{\pi}} z^{-1/4} e^{-2z^{3/2}/3}, \quad z \rightarrow \infty; \quad |\arg z| < \pi, \quad (3.7.19)$$

取得了进展, 但同时它在 $|\arg z| = \pi$ 时表现出斯托克斯现象。

3.8 渐进级数

在本节中, 我们用严谨的数学分析语言叙述前文所用到的渐进方法。

收敛与发散幂级数

在 3.5 节中, 我们将 $f(x) \sim \sum_{n=0}^{\infty} a_n(x-x_0)^n$ ($x \rightarrow x_0$) 定义为: 对于每一个 N , 当 $x \rightarrow x_0$ 时, 余项 $\varepsilon_N(x) = f(x) - \sum_{n=0}^N a_n(x-x_0)^n \ll (x-x_0)^N$ ($x \rightarrow x_0$)。

容易看出渐近级数是 Taylor 级数的推广。一个级数不必是收敛的才能是渐近的。收敛是一个绝对的概念, 它代表的是系数 a_n 的性质, 而渐近性是展开系数和级数所属函数 $f(x)$ 的一个相对属性。事实上, 每个幂级数对于某个连续函数 $f(x)$ 都是渐近的。

命题 3.1

给定一个形式幂级数 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$, 我们定义一个连续函数 $\phi(x; \alpha)$ 如下:

$$\phi(x; \alpha) = \begin{cases} 1, & |x| \leq \frac{1}{2}\alpha, \\ 2\left(1 - \frac{|x|}{\alpha}\right), & \frac{1}{2}\alpha < |x| < \alpha, \\ 0, & \alpha \leq |x|. \end{cases}$$

再定义 $\alpha_n = \min(1/|a_n|, 2^{-n})$, 其中 a_n 是任意级数 $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ 的系数. 定义:

$$f(x) = \sum_{n=0}^{\infty} a_n \phi(x-x_0; \alpha_n) (x-x_0)^n$$

则 f 是有限、连续的, 且满足

$$f(x) \sim \sum_{n=0}^{\infty} a_n (x-x_0)^n, \quad x \rightarrow x_0.$$

证明. 对于任何 $x \neq x_0$, 因为该级数在最多 N 项后截断, 其中 N 是满足 $2^{-N} \leq |x-x_0|$ 的最小整数, 故它是有限且连续的. 若 $|x-x_0| \leq \min(2^{-n}, 1/|a_n|, \dots, 1/|a_N|) = R_N$, 则 $\phi(x-x_0; \alpha_n) = 1, n=0, \dots, N$. 因此, 若 $|x-x_0| \leq R_N$,

$$f(x) = a_0 + a_1(x-x_0) + \dots + a_N(x-x_0)^N + \sum_{n=N+1}^{\infty} a_n \phi(x-x_0; \alpha_n) (x-x_0)^n.$$

由定义任意 x 与 n 有 $|a_n \phi(x-x_0; \alpha_n)| \leq 1, |\phi(x-x_0; \alpha_n)| \leq 1$, 因此若 $|x-x_0| \leq R_N$, 则

$$\begin{aligned} |\varepsilon_N(x)| &= \left| f(x) - \sum_{n=0}^N a_n (x-x_0)^n \right| \\ &= \left| a_{N+1} \phi(x-x_0; \alpha_{N+1}) (x-x_0)^{N+1} + \sum_{n=N+2}^{\infty} a_n \phi(x-x_0; \alpha_n) (x-x_0)^n \right| \\ &\leq |a_{N+1}| |x-x_0|^{N+1} + \sum_{n=N+2}^{\infty} |x-x_0|^{n-1} \\ &= |x-x_0|^{N+1} (|a_{N+1}| + (1-|x-x_0|)^{-1}) \ll (x-x_0)^N, \quad x \rightarrow x_0. \end{aligned}$$

因此, $\sum_{n=0}^{\infty} a_n (x-x_0)^n$ 对于 $x \rightarrow x_0$ 是 $f(x)$ 的渐近级数. □

渐近级数的例子**定义 3.8 (Stieltjes 级数)**

Stieltjes 级数 定义为

$$\sum_{n=0}^{\infty} (-1)^n n! x^n. \quad (3.8.1)$$

上文提到这是微分方程 (3.4.2) 的形式幂级数解。我们现在将利用它来渐近该方程的解。

为了对级数进行“求和”，我们利用积分恒等式 $n! = \int_0^\infty e^{-t} t^n dt$ ：

$$\sum_{n=0}^{\infty} (-x)^n n! \rightarrow \sum_{n=0}^{\infty} (-x)^n \int_0^\infty e^{-t} t^n dt.$$

无视任何收敛定理所要求的条件，交换求和与积分的顺序：

$$\sum_{n=0}^{\infty} (-x)^n \int_0^\infty e^{-t} t^n dt \rightarrow \int_0^\infty dt e^{-t} \sum_{n=0}^{\infty} (-xt)^n,$$

然后对 $\sum_{n=0}^{\infty} (-xt)^n \rightarrow 1/(1+xt)$ 求和，即使对于 $|xt| \geq 1$ 的 t 是发散的，至此我们得到：

$$y(x) = \int_0^\infty \frac{e^{-t}}{1+xt} dt, \quad (3.8.2)$$

这被称为Stieltjes 积分，存在且在 $x > 0$ 上解析，并且 $y(x)$ 满足微分方程 (3.4.2)。

命题 3.2

Stieltjes 级数对于 Stieltjes 积分是渐近的。

证明. 反复利用分部积分我们得到：

$$y(x) = 1 - x + 2!x^2 - 3!x^3 + \cdots + (-1)^N N!x^N + \varepsilon_N(x),$$

其中

$$\varepsilon_N(x) = (-1)^{N+1} (N+1)! x^{N+1} \int_0^\infty (1+xt)^{-N-2} e^{-t} dt.$$

利用不等式

$$\int_0^\infty (1+xt)^{-N-2} e^{-t} dt \leq \int_0^\infty e^{-t} dt = 1, \quad x \rightarrow 0+,$$

得到

$$|\varepsilon_N(x)| \leq (N+1)! x^{N+1} \ll x^N, \quad x \rightarrow 0+.$$

□

定义 3.9 (广义 Stieltjes 级数和积分)

广义 Stieltjes 积分定义为

$$f(x) = \int_0^\infty \frac{\rho(t)}{1+xt} dt, \quad (3.8.3)$$

其中权重函数 $\rho(t)$ 是非负的, 并且当 $t \rightarrow \infty$ 时迅速趋于零, 使得积分

$$a_n = \int_0^\infty t^n \rho(t) dt \quad (3.8.4)$$

对于所有正整数 n 都存在。对应的广义 Stieltjes 级数定义为

$$f(x) \sim \sum_{n=0}^{\infty} (-1)^n a_n x^n, \quad x \rightarrow 0+. \quad (3.8.5)$$

命题 3.3

广义 Stieltjes 级数对于广义 Stieltjes 积分是渐近的。

证明. 注意到对任意 N , 有

$$\begin{aligned} \varepsilon_N(x) &= f(x) - \sum_{n=0}^N (-1)^n a_n x^n \\ &= \int_0^\infty \rho(t) \left[\frac{1}{1+xt} - \sum_{n=0}^N (-xt)^n \right] dt \\ &= \int_0^\infty \frac{\rho(t)}{1+xt} (-xt)^{N+1} dt \end{aligned} \quad (3.8.6)$$

因此,

$$|\varepsilon_N(x)| \leq x^{N+1} \int_0^\infty \rho(t) t^{N+1} dt = a_{N+1} x^{N+1} \ll x^N, \quad x \rightarrow 0+, \quad (3.8.7)$$

□

命题 3.4

如果 $\rho(t) = K_0(t)$, 即零阶修正 Bessel 函数 (3.3.15), 那么 (3.8.5) 变为

$$\int_0^\infty \frac{K_0(t)}{1+xt} dt \sim \frac{1}{2} \sum_{n=0}^{\infty} (-2x)^n \left[\Gamma\left(\frac{1}{2}n + \frac{1}{2}\right) \right]^2 \quad (3.8.8)$$

使用渐进级数的数值近似: 最优截断法则

对于 Stieltjes 级数, 我们可以提供一个精确的渐近估计, 用于描述当 $x \rightarrow 0+$ 时 Stieltjes 积分的精确值与 Stieltjes 级数的最优截断之间的差异。

例 3.48

根据 (3.8.6), 对于权重函数 $\rho(t) = e^{-t}$ 的 Stieltjes 级数 (3.8.1), 其 N 项后的误差为

$$(-x)^N \int_0^\infty \frac{t^N}{1+xt} e^{-t} dt.$$

因为第 $n+1$ 项与第 n 项的比值是 $-nx$, 故 Stieltjes 级数 (3.8.1) 的最优截断是选择 N 为 $1/x$ 的向下取整。如果用这个积分表示来近似最优误差 $\varepsilon_{\text{optimal}}(x)$, 我们有

$$|\varepsilon_{\text{optimal}}(x)| \sim \left(\frac{\pi}{2x}\right)^{1/2} e^{-1/x}, \quad x \rightarrow 0+.$$
 (3.8.9)

当 $x \rightarrow 0+$ 时, 两者比值接近 1。

渐进级数的性质**(a) 非唯一性**

一个发散幂级数的“和”并不是唯一确定的。例如, 如果 $f(x) \sim \sum_{n=0}^\infty a_n(x-x_0)^n$ ($x \rightarrow x_0$), 那么 $f(x) + e^{-(x-x_0)^{-2}} \sim \sum_{n=0}^\infty a_n(x-x_0)^n$ ($x \rightarrow x_0$) 也同样成立。事实上, 级数 $\sum_{n=0}^\infty a_n(x-x_0)^n$ 对于任何与 $f(x)$ 相差一个函数 $g(x)$ 的函数都是渐近的, 只要当 $x \rightarrow x_0$ 时 $g(x)$ 比 $f(x)$ 的所有幂次 $x-x_0$ 都更快地趋于 0。这样的函数 $g(x)$ 被称为相对于渐近幂级数是次主导的, 其渐近展开是

$$g(x) \sim \sum_{n=0}^\infty 0 \cdot (x-x_0)^n, \quad x \rightarrow x_0.$$

(b) 唯一性

虽然有许多不同的函数可以渐近于一个给定的幂级数, 但对于每个函数只有一个渐近幂级数。这已在 3.5 节中给出。

(c) 渐进级数中的系数等同

$\sum_{n=0}^\infty a_n(x-x_0)^n \sim \sum_{n=0}^\infty b_n(x-x_0)^n$ ($x \rightarrow x_0$) 并不严格, 因为幂级数只能渐近于函数。这个记号的意思是 $\sum a_n(x-x_0)^n$ 和 $\sum b_n(x-x_0)^n$ 为渐近的函数类别是相同的。容易看出两个幂级数是渐近的, 当且仅当对于所有 n , $a_n = b_n$ 。因此, 我们可以将幂级数中的系数等同起来。

(d) 渐进级数的算术运算

算术运算可以逐项对渐近级数进行。具体来说, 假设

$$f(x) \sim \sum_{n=0}^\infty a_n(x-x_0)^n, \quad x \rightarrow x_0,$$

$$g(x) \sim \sum_{n=0}^\infty b_n(x-x_0)^n, \quad x \rightarrow x_0.$$

那么

$$\alpha f(x) + \beta g(x) \sim \sum_{n=0}^{\infty} (\alpha a_n + \beta b_n)(x - x_0)^n, \quad x \rightarrow x_0,$$

$$f(x)g(x) \sim \sum_{n=0}^{\infty} c_n(x - x_0)^n, \quad x \rightarrow x_0,$$

$$\frac{f(x)}{g(x)} \sim \sum_{n=0}^{\infty} d_n(x - x_0)^n, \quad x \rightarrow x_0,$$

其中

$$c_n = \sum_{m=0}^n a_m b_{n-m}, d_n = \frac{a_n - \sum_{m=0}^{n-1} d_m b_{n-m}}{b_0},$$

后者当然要求 $b_0 \neq 0$ 。证明留给读者。

(e) 渐近级数的积分

命题 3.5

任何渐近级数 $f(x) \sim \sum_{n=0}^{\infty} a_n(x - x_0)^n$ ($x \rightarrow x_0$)，如果 $f(x)$ 在 x_0 附近可积，则可以逐项积分：

$$\int_{x_0}^x f(t) dt \sim \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x - x_0)^{n+1}, \quad x \rightarrow x_0.$$

证明. 由 $|f(x) - \sum_{n=0}^N a_n(x - x_0)^n| \ll (x - x_0)^N$ ($x \rightarrow x_0$) 可知，对于任何 $\varepsilon > 0$ ，存在 $\delta, R > 0$ 使得任意 $|x - x_0| \leq R$ ，有

$$\left| f(x) - \sum_{n=0}^N a_n(x - x_0)^n \right| \leq \varepsilon |x - x_0|^N,$$

因此，

$$\begin{aligned} \left| \int_{x_0}^x \left(f(t) - \sum_{n=0}^N a_n(t - x_0)^n \right) dt \right| &\leq \left| \int_{x_0}^x \left| f(t) - \sum_{n=0}^N a_n(t - x_0)^n \right| dt \right| \\ &\leq \varepsilon \left| \int_{x_0}^x |t - x_0|^N dt \right| \\ &= \frac{\varepsilon}{N+1} |x - x_0|^{N+1} \end{aligned}$$

于是有，

$$\left| \frac{\int_{x_0}^x f(t) dt - \sum_{n=0}^N \frac{a_n}{n+1} (x - x_0)^{n+1}}{(x - x_0)^{N+1}} \right| \leq \frac{\varepsilon}{N+1}$$

由于 $\varepsilon > 0$ 是任意的，故

$$\int_{x_0}^x f(t) dt - \sum_{n=0}^N \frac{a_n}{n+1} (x - x_0)^{n+1} \ll (x - x_0)^{N+1}, \quad x \rightarrow x_0,$$

这表明渐近级数可以逐项积分。 □

如果我们对一个在无穷远处的渐近级数进行积分, 则上述论证可以推广到, 如果 $f(x) \sim \sum_{n=0}^{\infty} a_n x^{-n}$ ($x \rightarrow \infty$), 那么

$$\int_x^{\infty} [f(t) - a_0 - a_1 t^{-1}] dt \sim \sum_{n=2}^{\infty} \frac{a_n}{n-1} x^{1-n}, \quad x \rightarrow \infty.$$

(f) 渐近级数的微分

渐近级数通常不能逐项微分, 这与次主导性有关。考虑

$$g(x) = f(x) + e^{-1/(x-x_0)^2} \sin(e^{1/(x-x_0)^2})$$

与 $f(x)$ 相差一个次主导函数, 因此当 $x \rightarrow x_0$ 时具有相同的渐近级数展开。然而, $f'(x)$ 和

$$g'(x) = f'(x) - 2(x-x_0)^{-3} \cos(e^{1/(x-x_0)^2}) + 2(x-x_0)^{-3} e^{-1/(x-x_0)^2} \sin(e^{1/(x-x_0)^2})$$

不具有相同的渐近幂级数展开。我们有如下几种定理来处理渐近级数的微分。

定理 3.1

设 $f'(x)$ 存在且可积, 且 $f(x) \sim \sum_{n=0}^{\infty} a_n (x-x_0)^n$ ($x \rightarrow x_0$), 那么

$$f'(x) \sim \sum_{n=1}^{\infty} n a_n (x-x_0)^{n-1}, \quad x \rightarrow x_0.$$

证明. 设 $f'(x) \sim \sum_{n=0}^{\infty} b_n (x-x_0)^n$ ($x \rightarrow x_0$)。逐项积分给出

$$f(x) = f(x_0) + \int_{x_0}^x f'(t) dt \sim f(x_0) + \sum_{n=0}^{\infty} \frac{b_n}{n+1} (x-x_0)^{n+1}, \quad x \rightarrow x_0.$$

由渐近级数的唯一性即得。 □

定理 3.2

设 $f(x) \sim x^p$ ($x \rightarrow +\infty$), 其中 $p \geq 1$ 且 $f''(x)$ 是正的, 那么 $f'(x) \sim p x^{p-1}$ ($x \rightarrow +\infty$)。

这个结果只涉及 $f(x)$ 的主导行为, 并不能证明渐近级数的逐项微分。渐近级数的逐项微分在复数域中要清晰得多。

定理 3.3

设 $f(z)$ 在扇区 $\theta_1 \leq \arg(z-z_0) \leq \theta_2$, $0 < |z-z_0| < R$ 内是解析的, 并且 $f(z) \sim \sum_{n=0}^{\infty} a_n (z-z_0)^n$ [$z \rightarrow z_0; \theta_1 \leq \arg(z-z_0) \leq \theta_2$], 那么

$$f'(z) \sim \sum_{n=1}^{\infty} n a_n (z-z_0)^{n-1}, \quad z \rightarrow z_0; \theta_1 < \arg(z-z_0) < \theta_2.$$

即使渐近级数的微分定理比较复杂, 但我们指出, 如果已知一个函数满足一个线性微分方程, 其系数可以展开为渐近级数, 那么整个情况就会大大简化。

微分方程解的渐近展开

上文中给出的形式化方法需要有其合理性的证明, 这有如下几点困难要处理。首先, 我们一直在假设主导行为被分解出来后, 解的剩余部分存在一个渐近级数展开。然而, 容易看到并非所有函数都能被展开为渐近级数。其次, 渐近级数通常不能逐项微分。我们考虑 $y(x)$ 是 $y'' + py' + qy = 0$ 的解, 其中 $p(x), q(x)$ 当 $x \rightarrow x_0$ 时可展开为渐近幂级数, 则若 $y(x)$ 当 $x \rightarrow x_0$ 时可展开为渐近幂级数, 就有 $y'(x)$ 的渐近幂级数是对 $y(x)$ 的渐近幂级数逐项微分得到的。这个结论对于一般的 n 阶微分方程也是成立的。

剩下的问题是要证明解具有渐近级数表示, 我们通过一个例子来说明。

例 3.49 ($K_\nu(x)$ 渐近展开的存在性)

回忆, $K_\nu(x)$ 被定义为修正 Bessel 方程

$$x^2 y'' + xy' - (x^2 + \nu^2)y = 0 \quad (3.8.10)$$

的解, 其主导行为是

$$K_\nu(x) \sim \left(\frac{\pi}{2x}\right)^{1/2} e^{-x}, \quad x \rightarrow +\infty. \quad (3.8.11)$$

将 $w(x) = (\pi/2x)^{-1/2} e^x w(x)$ 代入 (3.8.10), 得到 $w(x)$ 的微分方程:

$$w'' - 2w' + \frac{\lambda}{x^2} w = 0, \quad (3.8.12)$$

其中 $\lambda = \frac{1}{4} - \nu^2$ 。我们想证明 $w(x)$ 具有 $w(x) \sim 1 + a_1/x + a_2/x^2 + \dots$ 形式的渐近展开, 容易验证它必须满足积分方程:

$$w(x) = 1 + \frac{\lambda}{2} \int_x^\infty (e^{2(x-t)} - 1) \frac{w(t)}{t^2} dt. \quad (3.8.13)$$

可以证明积分方程 (3.8.13) 有一个对 $x > a$ 有界 B_0 的解: $|w(x)| \leq B_0$ ($x > a$), 进而有:

$$\frac{1}{2} \left| \int_x^\infty (e^{2(x-t)} - 1) \frac{w(t)}{t^2} dt \right| \leq \frac{1}{2} \int_x^\infty \frac{|w(t)|}{t^2} dt \leq \frac{B_0}{2x}, \quad x > a.$$

上式与 (3.8.13) 表明 $w(x) \equiv 1 + w_1(x)$ 满足 $|w_1(x)| \leq B_1/x$ ($x > a$), 其中 $B_1 = \frac{1}{2}\lambda B_0$, 代入 (3.8.13) 给出

$$w_1(x) = \frac{\lambda}{2} \int_x^\infty (e^{2(x-t)} - 1) \frac{1}{t^2} dt + \frac{\lambda}{2} \int_x^\infty (e^{2(x-t)} - 1) \frac{w_1(t)}{t^2} dt.$$

对第一个积分分部积分得到

$$\frac{\lambda}{2} \int_x^\infty (e^{2(x-t)} - 1) \frac{1}{t^2} dt = -\frac{\lambda}{2x} + \frac{\lambda^2}{2} \int_x^\infty \frac{e^{2x-2t}}{t^2} dt.$$

因此, 存在一个正常数 B_2 使得 $w(x) = 1 - \lambda/2x + w_2(x)$, 其中 $|w_2(x)| < B_2/x^2$ ($x > a$), 这等价于 $w(x) \sim 1 - \lambda/2x$ ($x \rightarrow \infty$)。反复进行分部积分, 我们可以证明 $w(x)$ 具有渐近展开。

复平面中的渐近幂级数：不同有效扇形区域之间的联系

在级数对一个函数的有效扇形区域之外，通常还有其他级数对该函数渐近，且它们有自己的有效扇形区域。例如，我们在 3.5 节和 3.7 节中看到，函数 $\text{Ai}(x)$ 和 $\text{Bi}(x)$ 对大 $|x|$ 有不同的渐近展开，这取决于 x 的符号。因此，对于 Airy 方程 $y'' = zy$ ，一般的渐近表示为

$$y(z) \sim a_1 z^{-1/4} e^{-\frac{2}{3}z^{3/2}} \sum_{n=0}^{\infty} (-1)^n c_n z^{-3n/2} + a_2 z^{-1/4} e^{\frac{2}{3}z^{3/2}} \sum_{n=0}^{\infty} c_n z^{-3n/2}, \quad z \rightarrow \infty, \quad (3.8.14)$$

其中 a_1 和 a_2 是常数。特殊解 $\text{Ai}(z)$ 和 $\text{Bi}(z)$ 的渐近展开可以通过指定 a_1 和 a_2 的值来确定。斯托克斯现象仅表明在复平面的不同扇形区域中， $\text{Ai}(z)$ 或 $\text{Bi}(z)$ 的 a_1 和 a_2 的值可能不同。对于 $\text{Ai}(z)$:

$$a_1 = \frac{1}{2\sqrt{\pi}}, \quad a_2 = 0, \quad |\arg z| < \pi; \quad (3.8.15)$$

$$a_1 = \frac{1}{2\sqrt{\pi}}, \quad a_2 = \frac{i}{2\sqrt{\pi}}, \quad \frac{\pi}{3} < \arg z < \frac{5\pi}{3}; \quad (3.8.16)$$

$$a_1 = \frac{1}{2\sqrt{\pi}}, \quad a_2 = \frac{-i}{2\sqrt{\pi}}, \quad -\frac{5\pi}{3} < \arg z < -\frac{\pi}{3}. \quad (3.8.17)$$

对于 $\text{Bi}(z)$:

$$a_1 \text{ 任意}, \quad a_2 = \frac{1}{\sqrt{\pi}}, \quad |\arg z| < \frac{\pi}{3}; \quad (3.8.18)$$

$$a_1 = \frac{i}{2\sqrt{\pi}}, \quad a_2 = \frac{1}{\sqrt{\pi}}, \quad \frac{\pi}{3} < \arg z < \frac{5\pi}{3}; \quad (3.8.19)$$

$$a_1 = \frac{-i}{2\sqrt{\pi}}, \quad a_2 = \frac{1}{\sqrt{\pi}}, \quad -\frac{5\pi}{3} < \arg z < -\frac{\pi}{3}. \quad (3.8.20)$$

在 (3.8.18) 中， a_1 是任意的，这是因为其为在 $|\arg z| < \pi/3$ 时为次主导的函数。尽管 Airy 函数的渐近展开是唯一的，但人们不能解析地将其渐近展开从一个有效扇形区域延拓到另一个。那么，这些对大 $|z|$ 的不同渐近展开之间有什么联系呢？要回答这个问题，考虑：

$$\text{Ai}(z) = -\omega \text{Ai}(\omega z) - \omega^2 \text{Ai}(\omega^2 z), \quad (3.8.21a)$$

$$\text{Bi}(z) = i\omega \text{Ai}(\omega z) - i\omega^2 \text{Ai}(\omega^2 z), \quad (3.8.21b)$$

其中 $\omega = e^{-2\pi i/3}$ 。我们使用该函数关系来推导 (3.8.15) 和 (3.8.16) 与 (3.8.19) 之间的联系。若 $\frac{1}{3}\pi < \arg z < \frac{5}{3}\pi$ ，那么 $-\pi/3 < \arg(\omega z) < \pi$ 并且 $-\pi < \arg(\omega^2 z) < \pi/3$ 。因此，(3.8.14) 中具有 a_1 和 a_2 的 (3.8.15) 给出了 $\text{Ai}(\omega z)$ 和 $\text{Ai}(\omega^2 z)$ 的有效渐近展开。使用 (3.8.14)，我们得到 (3.8.16)。方程 (3.8.21a,b) 允许我们将渐近展开旋转 120 度。 $\text{Ai}(z)$ 的渐近展开由 (3.8.14) 和 (3.8.16) 给出。然而， $\text{Bi}(z)$ 的渐近展开 (3.8.18) 与 (3.8.20) 给出的区域没有重叠。为了验证 $\text{Bi}(z)$ 在 $-\pi < \arg z < -\pi/3$ 和 $\pi/3 < \arg z < \pi$ 时的有效性，可能需要获得额外的渐近展开。通过推导这些不同的渐近展开，我们解释了 $\text{Ai}(x)$ 和 $\text{Bi}(x)$ 对大正 x 的指数行为与对大负 x 的振荡行为之间的联系。

4 非线性微分方程的近似解法

5 差分方程的近似解法

6 积分的渐近展开

第三部分 扰动法

7 扰动级数

8 求和级数

第四部分 全局分析

9 边界层理论

10 WKB 理论

11 多尺度分析

第五部分 凸优化

12 基础定义

12.1 一些符号约定

- 我们只考虑 $\mathbb{R}^n$ 空间上的优化问题。
- 所有的向量约定为列向量。
- $'$ 表示转置。
- $x'y$ 为 $\sum_{i=1}^n$ 表示 x 和 y 的内积。
- $\|x\|_2 = (x'x)^{1/2}$ 为 x 的欧几里得范数。

12.2 凸集与凸函数

定义 12.1 (凸集)

设 $C \subseteq \mathbb{R}^n$ ，如果对于任意 $x, y \in C$ 和 $\alpha \in [0, 1]$ ，都有

$$\alpha x + (1 - \alpha)y \in C,$$

则称 C 是凸集 (convex set)。

例 12.1

- **Polyhedron sets:** 非空的满足

$$C = \{x \mid a_j'x \leq b_j, j = 1, \dots, r\}$$

- **Cones:** 满足对任意的 $x \in C, \lambda > 0$ 均有 $\lambda x \in C$ 的集合。

定义 12.2 (凸函数)

设 $C \subseteq \mathbb{R}^n$ 是凸集， $f: C \rightarrow \mathbb{R}$ ，如果对于任意 $x, y \in C$ 和 $\alpha \in [0, 1]$ ，都有

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y),$$

则称 f 是定义在 C 上的凸函数 (convex function)。若等号当且仅当 $x = y$ 时成立，则称 f 是严格凸函数 (strictly convex function)。

命题 12.1

若 f 是凸函数，则 $\{x \in C \mid f(x) \leq \gamma\}$ 与 $\{x \in C \mid f(x) < \gamma\}$ 均为凸集。

定义 12.3

$f: X \rightarrow [-\infty, +\infty]$ 的 **epigraph** 定义为

$$\text{epi}(f) = \left\{ (x, w) \mid x \in X, w \in \mathbb{R}, w \geq f(x) \right\}.$$

定义 12.4

f 的 **effective domain** 定义为

$$\text{dom}(f) = \left\{ x \in X \mid f(x) < +\infty \right\}.$$

定义 12.5

我们称 f 是 **凸的**, 若 $\text{epi}(f)$ 是凸集。

注意到, 若 $f(x) \in \mathbb{R}$, 则该定义与之前的定义等价。

定义 12.6

我们称 f 是 **闭的**, 若 $\text{epi}(f)$ 是闭集。

定义 12.7

我们称 f 在 $x \in X$ 处是 **下半连续的 (lower semicontinuous)**, 若对于任意 $\{x_k\} \subseteq X$ 满足 $x_k \rightarrow x$ 时, 有

$$\liminf_{k \rightarrow \infty} f(x_k) \geq f(x).$$

命题 12.2

对函数 $f: \mathbb{R}^n \rightarrow [-\infty, +\infty]$, 下列命题等价:

- 对任意 $\gamma \in \mathbb{R}$, 集合 $V_\gamma = \left\{ x \in \mathbb{R}^n \mid f(x) \leq \gamma \right\}$ 是闭集;
- f 在任意 $x \in \mathbb{R}^n$ 处是下半连续的;
- f 是闭的。

证明. (1) $\implies$ (2): 设 $\{x_k\} \subseteq \mathbb{R}^n$ 且 $x_k \rightarrow x$, 则对任意 $\varepsilon > 0$, 存在 $N > 0$ 使得当 $k > N$ 时, 有 $f(x_k) > f(x) - \varepsilon$ 。因此,

$$\liminf_{k \rightarrow \infty} f(x_k) \geq f(x) - \varepsilon.$$

由于 $\varepsilon > 0$ 是任意的, 故 $\liminf_{k \rightarrow \infty} f(x_k) \geq f(x)$ 。

(2) $\implies$ (3): 设 $\{(x_k, w_k)\} \subseteq \text{epi}(f)$ 且 $(x_k, w_k) \rightarrow (x, w)$, 则 $w_k \geq f(x_k)$ 。由下半连续性可知

$$w = \lim_{k \rightarrow \infty} w_k \geq \liminf_{k \rightarrow \infty} f(x_k) \geq f(x),$$

因此, $(x, w) \in \text{epi}(f)$ 。

(3) $\implies$ (1): 设 $\{x_k\} \subseteq V_\gamma$ 且 $x_k \rightarrow x$, 则 $(x_k, \gamma) \in \text{epi}(f)$ 。由闭性可知 $(x, \gamma) \in \text{epi}(f)$,

即 $f(x) \leq \gamma$, 所以 $x \in V_\gamma$ 。

□

命题 12.3

设 $f: X \rightarrow [-\infty, +\infty]$, 则若 $\text{dom}(f)$ 是闭集且 f 在 $\text{dom}(f)$ 上下半连续, 则 f 是闭的。

证明. 设 $\{(x_k, w_k)\} \subseteq \text{epi}(f)$ 且 $(x_k, w_k) \rightarrow (x, w)$, 则 $w_k \geq f(x_k)$ 。若 $x \in \text{dom}(f)$, 则由下半连续性可知

$$w = \lim_{k \rightarrow \infty} w_k \geq \liminf_{k \rightarrow \infty} f(x_k) \geq f(x),$$

因此, $(x, w) \in \text{epi}(f)$ 。若 $x \notin \text{dom}(f)$, 则 $f(x) = +\infty$, 显然 $(x, w) \in \text{epi}(f)$ 。

□

定义 12.8

我们称 f 是 **proper** 若 $\text{dom}(f) \neq \emptyset$ 且 $f(x) > -\infty$ 对任意 $x \in X$ 成立。我们称 f 是 **improper** 若 f 不是 proper。

命题 12.4

- $g: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 由下式给出

$$g(x) = \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x),$$

其中 $\lambda_i > 0$, 若 $f_i: \mathbb{R}^n \rightarrow [-\infty, +\infty]$ 是凸 (闭) 函数, 则 g 是凸 (闭) 函数。

- $g: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 由下式给出

$$g(x) = f(Ax)$$

其中 $A \in \mathbb{R}^{m \times n}$, 若 $f: \mathbb{R}^m \rightarrow [-\infty, +\infty]$ 是凸 (闭) 函数, 则 g 是凸 (闭) 函数。

- 考虑 $f_i: \mathbb{R}^n \rightarrow (-\infty, \infty]$, $i \in I$, $g: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 由下式给出

$$g(x) = \max_{i \in I} f_i(x)$$

若 f_i 是凸 (闭) 函数, 则 g 是凸 (闭) 函数。

证明. 证明留给读者。

□

12.3 可微凸函数

命题 12.5

设 $C \subset \mathbb{R}^n$ 是凸集, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是可微函数, 则 f 是 C 上的凸函数当且仅当

$$f(z) \geq f(x) + (z - x)' \nabla f(x), \forall x, z \in C$$

若等号当且仅当 $x = z$ 时成立, 则 f 是严格凸函数。

证明. “ $\Leftarrow$ ”: 设 $x, y \in C$, $\alpha \in [0, 1]$, $z = \alpha x + (1 - \alpha)y$, 则

$$\begin{aligned} f(z) &\geq f(x) + (z - x)' \nabla f(x) \\ &= f(x) + (1 - \alpha)(y - x)' \nabla f(x) \\ &\geq f(x) + (1 - \alpha)(f(y) - f(x)) \\ &= \alpha f(x) + (1 - \alpha)f(y) \end{aligned}$$

“ $\Rightarrow$ ”: 设 $x, z \in C$, $\alpha \in (0, 1)$, $y = \frac{z - \alpha x}{1 - \alpha}$, 则

$$\begin{aligned} f(z) &= f(\alpha x + (1 - \alpha)y) \\ &\leq \alpha f(x) + (1 - \alpha)f(y) \end{aligned}$$

即

$$f(y) \geq \frac{f(z) - \alpha f(x)}{1 - \alpha}.$$

由 $y = \frac{z - \alpha x}{1 - \alpha}$ 可知

$$z = \alpha x + (1 - \alpha)y.$$

对上式两边关于 α 求导, 得

$$0 = x - y + (1 - \alpha) \frac{\partial y}{\partial \alpha},$$

即

$$\frac{\partial y}{\partial \alpha} = \frac{y - x}{1 - \alpha}.$$

因此,

$$\frac{\partial}{\partial \alpha} f(y) = \nabla f(y)' \frac{\partial y}{\partial \alpha} = \nabla f(y)' \frac{y - x}{1 - \alpha}.$$

对不等式两边关于 α 求导, 得

$$0 \leq -\frac{f(z) - f(x)}{(1 - \alpha)^2} + \frac{1}{1 - \alpha} \nabla f(y)' (y - x).$$

令 $\alpha \rightarrow 0$, 则 $y \rightarrow z$, 从而

$$f(z) \geq f(x) + (z - x)' \nabla f(z).$$

□

定理 12.1 (optimality condition)

设 C 是 $\mathbb{R}^n$ 中的非空凸集, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是凸函数且在 C 上可微, 则 $x^* \in C$ 是 $\min_{x \in C} f(x)$ 的解当且仅当

$$(x - x^*)' \nabla f(x^*) \geq 0, \forall x \in C.$$

证明. “ $\Rightarrow$ ”: 设 $x \in C$, $\alpha \in (0, 1)$, 则 $\alpha x + (1 - \alpha)x^* \in C$, 因此

$$f(\alpha x + (1 - \alpha)x^*) \geq f(x^*).$$

由凸性可知

$$f(\alpha x + (1 - \alpha)x^*) \leq \alpha f(x) + (1 - \alpha)f(x^*),$$

故

$$\alpha f(x) + (1 - \alpha)f(x^*) \geq f(x^*).$$

即

$$f(x) - f(x^*) \geq 0.$$

对上式两边关于 α 求导, 得

$$0 \leq f(x) - f(x^*) = (x - x^*)' \nabla f(x^*).$$

“ $\Leftarrow$ ”: 设 $x \in C$, $\alpha \in [0, 1]$, 则 $\alpha x + (1 - \alpha)x^* \in C$, 由凸性可知

$$f(\alpha x + (1 - \alpha)x^*) \leq \alpha f(x) + (1 - \alpha)f(x^*).$$

由假设可知

$$f(x) - f(x^*) \geq 0.$$

因此,

$$f(\alpha x + (1 - \alpha)x^*) \leq \alpha f(x) + (1 - \alpha)f(x^*) \leq \alpha f(x^*) + (1 - \alpha)f(x^*) = f(x^*).$$

□

定理 12.2 (projection theorem)

设 C 是 $\mathbb{R}^n$ 中的非空闭凸集。

- 对于任意 $z \in \mathbb{R}^n$, 存在唯一的 $x^* \in C$ 使得 $f(x) = \|x - z\|^2$ 在 C 上取得最小值, x^* 称为 z 在 C 上的**投影 (projection)**。
- x^* 是 z 在 C 上的投影当且仅当

$$(x - x^*)'(z - x^*) \leq 0, \forall x \in C.$$

证明. (a): 设 $\{x_k\} \subseteq C$ 且 $f(x_k) \rightarrow \inf_{x \in C} f(x)$, 则 $\{x_k\}$ 有界, 因此存在 $\{x_{k_i}\} \subseteq \{x_k\}$ 和 $x^* \in \mathbb{R}^n$ 使得 $x_{k_i} \rightarrow x^*$. 由于 C 是闭集, 故 $x^* \in C$. 由 f 的连续性可知

$$f(x^*) = \lim_{i \rightarrow \infty} f(x_{k_i}) = \inf_{x \in C} f(x).$$

若 $y^* \in C$ 且 $y^* \neq x^*$, 则

$$\|y^* - z\|^2 = \|y^* - x^* + x^* - z\|^2 = \|y^* - x^*\|^2 + \|x^* - z\|^2 + 2(y^* - x^*)'(x^* - z) > \|x^* - z\|^2,$$

因此, x^* 是唯一的。

(b): 由上一定理可知, x^* 是 z 在 C 上的投影当且仅当

$$(x - x^*)' \nabla f(x^*) \geq 0, \forall x \in C.$$

由于 $\nabla f(x) = 2(x - z)$, 因此结论成立。 $\square$

命题 12.6

设 C 是 $\mathbb{R}^n$ 中的非空闭凸集, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是 $\mathbb{R}^n$ 上的二阶连续可微函数。

- 若 $\nabla^2 f(x)$ 在 C 上半正定, 则 f 是 C 上的凸函数。
- 若 $\nabla^2 f(x)$ 在 C 上正定, 则 f 是 C 上的严格凸函数。
- 若 C 是开的且 f 是 C 上的凸函数, 则 $\nabla^2 f(x)$ 在 C 上半正定。

证明. (a): 由中值定理可知对 $x, y \in C$, 存在 $\alpha \in (0, 1)$ 使得

$$f(y) = f(x) + (y - x)' \nabla f(x) + \frac{1}{2} (y - x)' \nabla^2 f(x + \alpha(y - x)) (y - x).$$

由于 $\nabla^2 f(x + \alpha(y - x))$ 半正定, 故

$$f(y) \geq f(x) + (y - x)' \nabla f(x).$$

由前命题可知, f 是 C 上的凸函数。

(b): 同 (a), 由于 $\nabla^2 f(x + \alpha(y - x))$ 正定, 故

$$f(y) > f(x) + (y - x)' \nabla f(x).$$

由前命题可知, f 是 C 上的严格凸函数。

(c): 由反证法, 设存在 $x \in C$ 和 $d \in \mathbb{R}^n$ 使得 $d' \nabla^2 f(x) d < 0$, 则存在 $\alpha > 0$ 使得

$$f(x + \alpha d) < f(x) + \alpha d' \nabla f(x).$$

由于 C 是开的, 故 $x + \alpha d \in C$. 由前命题可知, f 不是 C 上的凸函数, 这与假设矛盾。 $\square$

12.4 凸包、仿射包和相对内点

给定集合 $X \subset \mathbb{R}^n$ 。

定义 12.9

X 中元素的一个凸组合 (convex combination) 是形如

$$\sum_{i=1}^k \alpha_i x_i,$$

其中 $x_i \in X$, $\alpha_i \geq 0$ 且 $\sum_{i=1}^k \alpha_i = 1$ 的向量。

定义 12.10

X 的凸包 (convex hull) 是包含 X 的所有凸集的交集, 记为 $\text{conv}(X)$ 。

定义 12.11

一个仿射集 (affine set) 形如 $\bar{x} + S$, 其中 $\bar{x} \in \mathbb{R}^n$, S 是 $\mathbb{R}^n$ 的子空间。

定义 12.12

X 的仿射包 (affine hull) 是包含 X 的所有仿射集的交集, 记为 $\text{aff}(X)$ 。

定义 12.13

X 中元素的一个非负线性组合 (nonnegative linear combination) 是形如 $\sum_{i=1}^k \alpha_i x_i$, 其中 $x_i \in X$, $\alpha_i \geq 0$ 的向量。

定义 12.14

X 生成的锥 (cone) 是所有 X 中元素的非负线性组合的集合, 记为 $\text{cone}(X)$ 。

定理 12.3 (Caratheodory's theorem)

设 X 是 $\mathbb{R}^n$ 的子集。

- 任意非零元 $x \in \text{cone}(X)$ 可表示为 X 中至多 n 个元素的非负线性组合。
- 任意非零元 $x \in \text{conv}(X)$ 可表示为 X 中至多 $n + 1$ 个元素的凸组合。

证明. (a): 设 $x \in \text{cone}(X)$, 则存在 $x_i \in X$ 和 $\alpha_i \geq 0$ 使得

$$x = \sum_{i=1}^m \alpha_i x_i.$$

若 $m \leq n$, 则结论成立。若 $m > n$, 则 $\{x_1, \dots, x_m\}$ 线性相关, 故存在不全为零的 β_i 使得

$$\sum_{i=1}^m \beta_i x_i = 0.$$

设 $I = \{i | \beta_i > 0\}$, $J = \{i | \beta_i < 0\}$, 则 $I \cup J = \{1, \dots, m\}$ 。由于 $x \neq 0$, 故 $I \neq \emptyset$ 。令

$$\theta = \min_{i \in I} \frac{\alpha_i}{\beta_i} > 0,$$

则对任意 $i \in I$, 都有 $\alpha_i - \theta\beta_i \geq 0$, 且存在 $i_0 \in I$ 使得 $\alpha_{i_0} - \theta\beta_{i_0} = 0$ 。因此,

$$x = \sum_{i=1}^m \alpha_i x_i = \sum_{i=1}^m (\alpha_i - \theta\beta_i) x_i,$$

其中 $\alpha_i - \theta\beta_i \geq 0$ 且 $\alpha_{i_0} - \theta\beta_{i_0} = 0$ 。由此可去掉 x_{i_0} , 得到 x 的一个新的表示。反复进行该过程, 直到表示中不超过 n 个元素为止。

(b): 将 (a) 应用于 $X' = \{(x, 1) | x \in X\} \subseteq \mathbb{R}^{n+1}$ 即可。 $\square$

命题 12.7

紧集的凸包是紧集。

证明. 设 X 是 $\mathbb{R}^n$ 中的紧集, $\{y_k\} \subseteq \text{conv}(X)$ 且 $y_k \rightarrow y$, 则存在 $x_{k,i} \in X$ 和 $\alpha_{k,i} \geq 0$ 使得

$$y_k = \sum_{i=1}^{m_k} \alpha_{k,i} x_{k,i}, \quad \sum_{i=1}^{m_k} \alpha_{k,i} = 1.$$

由 Caratheodory 定理可知, $m_k \leq n+1$ 。因此, 存在 $m \leq n+1$ 和 $\{y_{k_j}\} \subseteq \{y_k\}$ 使得

$$y_{k_j} = \sum_{i=1}^m \alpha_{k_j,i} x_{k_j,i}, \quad \sum_{i=1}^m \alpha_{k_j,i} = 1.$$

由于 X 是紧集, 故存在 $\{x_i\} \subseteq X$ 和 $\{\alpha_i\}$ 使得 $x_{k_j,i} \rightarrow x_i$ 和 $\alpha_{k_j,i} \rightarrow \alpha_i$ 。由 $\sum_{i=1}^m \alpha_{k_j,i} = 1$ 可知 $\sum_{i=1}^m \alpha_i = 1$ 。因此,

$$y = \lim_{j \rightarrow \infty} y_{k_j} = \sum_{i=1}^m \alpha_i x_i,$$

即 $y \in \text{conv}(X)$ 。 $\square$

定义 12.15

称 x 是 C 的相对内点 (relative interior), 记为 $\text{ri}(C)$, 若 x 是一个 C 在 $\text{aff}(C)$ 中的内点。

定义 12.16

记 $\text{ri}(C)$ 为 C 所有相对内点的集合。

定理 12.4 (Line Segment Principle)

设 C 是 $\mathbb{R}^n$ 中的非空凸集, 则对任意 $x \in \text{ri}(C)$ 和 $\bar{x} \in \text{cl}(C)$, 都有

$$\alpha x + (1 - \alpha)\bar{x} \in \text{ri}(C), \forall \alpha \in [0, 1).$$

证明. 固定 $\alpha \in (0, 1]$. 考虑 $x_\alpha = \alpha x + (1 - \alpha)\bar{x}$.

令 L 为平行于 $\text{aff}(C)$ 的子空间. 定义 $B_L(0, \epsilon) := \{z \in L \mid \|z\| < \epsilon\}$.

由于 $\bar{x} \in \text{cl}(C)$, 对于所有 $\epsilon > 0$, 有 $\bar{x} \in C + B_L(0, \epsilon)$. 那么

$$\begin{aligned} B(x_\alpha; \epsilon) \cap \text{aff}(C) &= \{\alpha x + (1 - \alpha)\bar{x}\} + B_L(0; \epsilon) \\ &\subset \{\alpha x\} + (1 - \alpha)C + (2 - \alpha)B_L(0; \epsilon) \\ &= (1 - \alpha)C + \alpha \left[x + B_L\left(0; \frac{2 - \alpha}{\alpha} \epsilon\right) \right] \end{aligned}$$

由于 $x \in \text{ri}(C)$, 对于足够小的 ϵ , 有 $x + B_L\left(0; \frac{2 - \alpha}{\alpha} \epsilon\right) \subset C$.

因此 $B(x_\alpha; \epsilon) \cap \text{aff}(C) \subset \alpha C + (1 - \alpha)C = C$ (因为 C 是凸集). 所以, $x_\alpha \in \text{ri}(C)$. $\square$

命题 12.8

设 C 是非空凸集.

- $\text{ri}(C)$ 是非空凸集, 且其仿射包与 C 相同.
- **Prolongation Lemma:** $x \in \text{ri}(C)$ 当且仅当对任意以 x 为端点的所有线段可以延伸超过 x 的另一端点而仍在 C 内.

证明. (a): 设 $0 \in C$. 取 m 个线性独立向量 $z_1, \dots, z_m \in C$, 其中 m 是 $\text{aff}(C)$ 的维数, 设

$$X = \left\{ \sum_{i=1}^m \alpha_i z_i \mid \sum_{i=1}^m \alpha_i < 1, \alpha_i > 0, i = 1, \dots, m \right\}$$

则 $X \subseteq \text{ri}(C)$ 且 $\text{aff}(X) = \text{aff}(C)$. 因此, $\text{ri}(C)$ 是非空的. 由凸性可知, $\text{ri}(C)$ 是凸集.

(b): 必要性显然. 充分性: 任取 $\bar{x} \in C$, 由 Line Segment Principle 可知

$$\alpha x + (1 - \alpha)\bar{x} \in \text{ri}(C), \forall \alpha \in [0, 1).$$

因此, $x \in \text{ri}(C)$. $\square$

命题 12.9

一个凹函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在 X 上的最小值在 $x^* \in \text{ri}(X)$ 处取得, 则 f 在 X 上为常值.

证明. 反证法, 设存在 $x \in X$ 使得 $f(x) > f(x^*)$. 由 Line Segment Principle 可知

$$\alpha x + (1 - \alpha)x^* \in \text{ri}(X), \forall \alpha \in [0, 1).$$

由凸性可知

$$f(\alpha x + (1 - \alpha)x^*) \leq \alpha f(x) + (1 - \alpha)f(x^*) < f(x^*),$$

矛盾. $\square$

推论 12.1

一个非常值线性函数在其定义域的相对内点上没有极值。

命题 12.10

- $\text{cl}(C) = \text{cl}(\text{ri}(C))$ 且 $\text{ri}(\text{cl}(C)) = \text{ri}(C)$ 。
- 设 $\bar{C}$ 是 $\mathbb{R}^n$ 中的非空闭凸集, 则
 - C 和 $\bar{C}$ 有相同的相对内点;
 - C 和 $\bar{C}$ 有相同的闭包;
 - $\text{ri}(C) \subset \bar{C} \subset \text{cl}(C)$ 。

证明. (a): 由定义可知 $\text{cl}(\text{ri}(C)) \subseteq \text{cl}(C)$ 。反过来, 设 $\bar{x} \in \text{cl}(C)$, $x \in \text{ri}(C)$, 则由 Line Segment Principle 可知

$$\alpha x + (1 - \alpha)\bar{x} \in \text{ri}(C), \forall \alpha \in (0, 1].$$

因此, $\bar{x} \in \text{cl}(\text{ri}(C))$, 从而 $\text{cl}(C) \subseteq \text{cl}(\text{ri}(C))$ 。

其余同理。

□

命题 12.11

设 C 是 $\mathbb{R}^n$ 中的非空凸集且 A 是 $m \times n$ 矩阵, 则

- $A \cdot (\text{ri}(C)) = \text{ri}(A \cdot C)$;
- $A \cdot (\text{cl}(C)) \subset \text{cl}(A \cdot C)$ 。若 C 是有界的, 则等号成立。

证明. (a): 由 C 中 (关于仿射包) 的邻域映射为 $A \cdot C$ 中的邻域即得。

(b): 我们有 $A \cdot \text{cl}(C) \subseteq \text{cl}(A \cdot C)$, 这是因为对序列 $\{x_k\} \subset C$, 若 $x_k \rightarrow x \in \text{cl}(C)$, 则 $Ax_k \rightarrow Ax \in \text{cl}(A \cdot C)$ 。

反过来, 设 C 是有界的, 取任意 $z \in \text{cl}(A \cdot C)$, 则存在 $\{x_k\} \subseteq A \cdot C$ 使得 $x_k \rightarrow z$ 。由 C 有界可知存在 $\{x_{k_j}\} \subseteq \{x_k\}$ 和 $x \in \mathbb{R}^n$ 使得 $x_{k_j} \rightarrow x$ 。由于 C 是闭集, 故 $x \in C$, 于是有 $z = Ax$ 。因此, $z \in A \cdot \text{cl}(C)$ 。

□

命题 12.12

设 C_1 和 C_2 是 $\mathbb{R}^n$ 中的非空凸集, 则

•

$$\text{ri}(C_1) + \text{ri}(C_2) = \text{ri}(C_1 + C_2);$$

$$\text{cl}(C_1) + \text{cl}(C_2) \subset \text{cl}(C_1 + C_2).$$

若 C_1 和 C_2 都是有界的, 则等号成立。

•

$$\text{ri}(C_1) \cap \text{ri}(C_2) \subset \text{ri}(C_1 \cap C_2) \quad \text{cl}(C_1 \cap C_2) \subset \text{cl}(C_1) \cap \text{cl}(C_2).$$

若 $\text{ri}(C_1) \cap \text{ri}(C_2) \neq \emptyset$, 则等号成立。

证明. 证明留给读者。 □

命题 12.13 (Cartesian product)

设 C 是 $\mathbb{R}^{n+m}$ 中的非空凸集, 对 $x \in \mathbb{R}^n$, 定义

$$C_x = \{y \in \mathbb{R}^m \mid (x, y) \in C\}.$$

与

$$D = \{x \mid C_x \neq \emptyset\}$$

则

$$\text{ri}(C) = \{(x, y) \mid x \in \text{ri}(D), y \in \text{ri}(C_x)\}$$

证明. 由于 D 是 C 的投影, 故

$$\text{ri}(D) = \{x \mid \exists y \in \mathbb{R}^m, (x, y) \in \text{ri}(C)\}$$

于是

$$\text{ri}(C) = \bigcup_{x \in \text{ri}(D)} (M_x \cap \text{ri}(C))$$

其中 $M_x = \{(x, y) \mid y \in \mathbb{R}^m\}$ 。对任意 $x \in \text{ri}(D)$, $M_x \cap \text{ri}(C) = \text{ri}(M_x \cap C) = \{(x, y) \mid y \in \text{ri}(C_x)\}$ □

12.5 凸函数的性质**命题 12.14**

若 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是凸的, 则 f 连续。

证明. 我们来证明 f 在 0 处连续。由凸性, f 在单位正方体上有界且最大值在顶点处取得。

考虑 $x_k \rightarrow 0$, 令 $y_k = x_k/\|x_k\|_\infty$, $z_k = -x_k/\|x_k\|_\infty$, 则

$$f(x_k) \leq (1 - \|x_k\|_\infty)f(0) + \|x_k\|_\infty f(y_k)$$

$$f(0) \leq \frac{\|x_k\|_\infty}{\|x_k\|_\infty + 1} f(z_k) + \frac{1}{\|x_k\|_\infty + 1} f(x_k)$$

令 $k \rightarrow \infty$ 即得。 □

不仅如此, 我们还可以将连续性延拓到 $\text{ri}(\text{dom}(f))$ 上。

定义 12.17

函数 $f: X \rightarrow [-\infty, +\infty]$ 的闭包为函数 $\text{cl}f: \mathbb{R}^n \rightarrow [-\infty, +\infty]$ 定义为

$$\text{epi}(\text{cl}f) = \text{cl}(\text{epi}(f)).$$

定义 12.18

f 的凸包为函数 $\check{\text{cl}}f$ 定义为

$$\text{epi}(\check{\text{cl}}f) = \text{cl}(\text{conv}(\text{epi}(f))).$$

命题 12.15

对任意 $f: X \rightarrow [-\infty, +\infty]$,

$$\inf_{x \in X} f(x) = \inf_{x \in \mathbb{R}^n} (\text{cl}f)(x) = \inf_{x \in \mathbb{R}^n} (\check{\text{cl}}f)(x).$$

命题 12.16

对任意 $f: X \rightarrow [-\infty, +\infty]$, 则

- $\text{cl}f$ (或 $\check{\text{cl}}f$) 是最大的闭 (或闭凸) 函数满足被 f 优化。
- 若 f 是凸的, 则 $\text{cl}f$ 是凸的。其 proper 当且仅当 f proper。且

$$(\text{cl}f)(x) = f(x), \forall x \in \text{ri}(\text{dom}(f)).$$

且若 $x \in \text{ri}(\text{dom}(f))$, $y \in \text{dom}(\text{cl}f)$, 则

$$(\text{cl}f)(y) = \lim_{\alpha \rightarrow 0^+} f(y + \alpha(x - y)).$$

定义 12.19

给定非空凸集 C , 向量 d 称为回收方向 (a direction of recession), 若从任意点 $x \in C$ 沿 d 方向出发的射线都包含在 C 中, 即

$$x + \alpha d \in C, \forall x \in C, \forall \alpha \geq 0.$$

定义 12.20

C 的回收锥 (recession cone) 定义为所有回收方向构成的集合, 记为 R_C 。

容易看到 R_C 是包含原点的锥。

命题 12.17

设 C 是非空闭凸集。

- R_C 是闭凸锥。
- 向量 $d \in R_C$ 当且仅当存在向量 $x \in C$ 使得 $x + \alpha d \in C, \forall \alpha \geq 0$ 。
- R_C 含非零向量当且仅当 C 是无界的。
- C 和 $\text{ri}(C)$ 有相同的回收锥。
- 若 D 为另一个闭凸集满足 $C \cap D \neq \emptyset$, 则

$$R_{C \cap D} = R_C \cap R_D$$

更一般的对一族闭凸集 $\{C_i\}_{i \in I}$, 若 $\cap_{i \in I} C_i \neq \emptyset$, 则

$$R_{\cap_{i \in I} C_i} = \cap_{i \in I} R_{C_i}.$$

(b) 的证明. 设 $d \neq 0$, 存在 $x \in C$ 使得 $x + \alpha d \in C, \forall \alpha \geq 0$ 。固定 $\bar{x} \in C, \alpha > 0$, 我们来证明 $\bar{x} + \alpha d \in C$ 。通过单位化 d 只需证明 $\bar{x} + d \in C$ 。

对 $k = 1, 2, \dots$, 令

$$z_k = x_k + kd, d_k = \frac{(z_k - \bar{x})}{\|z_k - \bar{x}\|} \|d\|$$

我们有

$$\frac{d_k}{\|d_k\|} = \frac{z_k - \bar{x}}{\|z_k - \bar{x}\|} \frac{d}{\|d\|} + \frac{x - \bar{x}}{\|z_k - \bar{x}\|}, \frac{z_k - \bar{x}}{\|z_k - \bar{x}\|} \rightarrow 1, \frac{x - \bar{x}}{\|z_k - \bar{x}\|} \rightarrow 0$$

故 $d_k \rightarrow d$ 且 $\bar{x} + d_k \rightarrow \bar{x} + d$ 。由于 C 是闭凸集, 故 $\bar{x} + d \in C$ 。

□

定义 12.21

凸集 C 的线性空间 (lineality space) 定义为 C 的回收锥与其负回收锥的交集, 记为

$$L_C = R_C \cap (-R_C).$$

若 $d \in L_C$, 这表明 C 包含经过任意点 $x \in C$ 的方向为 d 的直线。由此我们可以将凸集 C 分解为

$$C = L_C + (C \cap L_C^\perp).$$

这使得我们可以将证明关于 C 的结论归结为关于 $C \cap L_C^\perp$ 的结论。该分解将 L_C 换成其子空间 S 同样成立。

接下来我们的目标是刻画凸函数的单调递减方向。

命题 12.18

设 $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 是闭且 proper 的凸函数, 考虑水平集 $V_\gamma = \{x \mid f(x) \leq \gamma\}$, 其中 γ 是标量, 则:

- 所有非空水平集 V_γ 有相同的回收锥:

$$R_{V_\gamma} = \left\{ d \mid (d, 0) \in R_{\text{epi}(f)} \right\}.$$

- 若某个非空水平集 V_γ 是紧的, 则所有水平集都是紧的。

证明. (a): 显然。

(b): 由 (a) 即得。 □

定义 12.22

对闭且 proper 的凸函数 $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$, V_γ 的回收锥称为称为 f 的回收锥 (**recession cone**), 记为 R_f . 同上定义 $d \in R_f$ 为 f 的回收方向, $L_f = R_f \cap (-R_f)$ 为 f 的线性空间, $d \in L_f$ 称为 f 的不变空间。

例 12.2

对半正定矩阵 Q , 函数 $f(x) = x'Qx + a'x + b$, 我们有

$$R_f = \left\{ d \mid Qd = 0, a'd \leq 0 \right\}, L_f = \left\{ d \mid Qd = 0, a'd = 0 \right\}.$$

定义 12.23

定义 f 的回收函数 (**recession function**) 为 $r_f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$, 其满足

$$\text{epi}(r_f) = R_{\text{epi}(f)}.$$

命题 12.19

$$R_f = \left\{ d \mid r_f(d) \leq 0 \right\}, L_f = \left\{ d \mid r_f(d) = r_f(-d) = 0 \right\}.$$

命题 12.20

$$r_f(d) = \sup_{\alpha} \frac{f(x + \alpha d) - f(x)}{\alpha} = \lim_{\alpha \rightarrow +\infty} \frac{f(x + \alpha d) - f(x)}{\alpha}, \forall x \in \text{dom}(f).$$

因此 $r_f(d)$ 描述了 f 的渐进线。事实上若 f 可微,

$$r_f(d) = \lim_{\alpha \rightarrow +\infty} \nabla f(x + \alpha d)' d, \forall x, d \in \mathbb{R}^n.$$

12.6 局部与全局最小值

考虑在集合 $X \subset \mathbb{R}^n$ 上极小化 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 的问题。

定义 12.24

- 称 x 是可行的 (**feasible**) 若 $x \in X \cap \text{dom}(f)$ 。
- 称 x^* 是 f 在 X 上的 (全局) 最小值点 (**minimum**), 若 x^* 可行且 $f(x^*) = \inf_{x \in X} f(x)$ 。
- 称 x^* 是 f 在 X 上的局部最小值点 (**local minimum**), 若存在某个 $\epsilon > 0$ 使得 x^* 是 $X \cap B(x^*; \epsilon)$ 上的最小值点。

命题 12.21

若 X 是凸集且 f 是凸的, 则

- 任意局部最小值点都是全局最小值点;
- 若 f 是严格凸的, 则最小值点唯一。

下面我们考虑优化问题的解的存在性问题。

注意到 f 最小值点的集合是 f 的水平集的交集。若某个水平集是紧的, 则所有水平集都是紧的, 且最小值点存在。

定理 12.5 ((An Extension of the) Weierstrass's theorem)

f 最小值点的集合是非空且紧的, 若 X 是闭的, f 在 X 上下半连续, 且下述条件有一者成立:

- X 有界;
- 某个水平集 $V_\gamma = \{x \in X | f(x) \leq \gamma\}$ 是非空的且有界。
- 对任意子序列 $\{x_k\} \subset X$, 若 $\|x_k\| \rightarrow +\infty$, 则 $f(x_k) \rightarrow +\infty$ 。

证明. 在每种情形下, f 的水平集 $\cap X$ 均为紧集。 □

命题 12.22 (Weierstrass' Theorem specialized to convex function)

设 X 是 $\mathbb{R}^n$ 中的非空闭凸集, $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 是闭且凸的函数满足 $X \cap \text{dom}(f) \neq \emptyset$, 则 f 在 X 上的最小值点存在当且仅当 $R_X \cap R_f = \{0\}$ 。

证明. 设 $f^* = \inf_{x \in X} f(x)$ 。注意到由于 $X \cap \text{dom}(f) \neq \emptyset$, 故 $f^* < \infty$ 。设 $\{\gamma_k\}$ 是一个递减到 f^* 的数列, 考虑

$$V_k = \{x \in X | f(x) \leq \gamma_k\}$$

则 f 在 X 上的最小值点为

$$X^* = \bigcap_{k=1}^{\infty} (X \cap V_k).$$

其中 $X \cap V_k$ 是非空的且 $R_f \cap R_X$ 是其回收锥, 当 $X \neq \emptyset$ 时这也是 X^* 的回收锥。这表明 X^*

是非空且紧的当且仅当 $R_f \cap R_X = \{0\}$ 。 $\square$

命题 12.23

设 $f_i : \mathbb{R}^n \rightarrow (-\infty, +\infty]$, $i = 1, 2$ 是闭且 proper 的凸函数, 且满足

$$f = f_1 + \cdots + f_m$$

是 proper 的。若存在一个 f_i 满足 $r_{f_i}(d) = \infty$ 对任意非零 $d \in R_f$, 则 f 在 $\mathbb{R}^n$ 上的最小值点集非空且紧。

证明. 由假设可知 $r_f(d) = \infty$ 对任意非零 $d \in R_f$ 。因此, $R_f = \{0\}$, 由 Weierstrass 定理即得。 $\square$

将 f 改为 $\max\{f_1, \cdots, f_m\}$ 结论同样成立。

考虑一个基础的问题: 给定 $\mathbb{R}^n$ 上的非空闭集序列 $\{C_k\}$ 满足 $C_{k+1} \subseteq C_k$, 则 $\bigcap_{k=1}^{\infty} C_k$ 是否非空? 由此可以借由 $X \cap V_k$ 的交集非空判断 f 在 X 上的最小值点集是否非空。

还有一个相关的问题: 若 C 是闭的, A 是矩阵, 则 AC 是否是闭的? 更特别的若 C_1, C_2 是闭的, $C_1 + C_2$ 是否是闭的? 注意到 $C_1 + C_2 = A(C_1 \times C_2)$, 其中 $A(x_1, x_2) = x_1 + x_2$ 。

命题 12.24

设 C 是非空闭凸集, A 是矩阵, $N(A)$ 是其零空间。则 AC 是闭的当且仅当 $N(A) \cap R_C = \{0\}$ 。

证明. 设 $\{y_k\} \subset AC$ 满足 $y_k \rightarrow \bar{y}$ 。定义 $C_k = C \cap N_k$, 其中

$$N_k = \left\{x \mid Ax \in W_k\right\}, \quad W_k = \left\{z \mid \|z - \bar{y}\| \leq \|y_k - \bar{y}\|\right\}$$

我们有 $R_{N_k} = N(A)$, 故 C_k 是紧的, 且 C_k 的交不空。取 $x \in \bigcap_{k=1}^{\infty} C_k$, 则 $Ax = \bar{y}$, 从而 $\bar{y} \in AC$ 。 $\square$

于是我们有: $C_1 + C_2$ 是闭的当且仅当 C_1, C_2 是闭的且有一者是紧的。

定理 12.6

若 X 是多面体, 则 AX 是闭的。

设 $F : \mathbb{R}^{n+m} \rightarrow (-\infty, +\infty]$ 是闭且 proper 的凸函数, 定义 $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 为

$$f(x) = \inf_{z \in \mathbb{R}^m} F(x, z).$$

命题 12.25

若 F 是凸的, 则 f 是凸的。

若 $F(x, z)$ 是闭的, $f(x)$ 不一定是闭的。考虑闭的凸函数

$$F(x, z) = \begin{cases} e^{-\sqrt{xz}} & \text{if } x \geq 0, z \geq 0 \\ \infty & \text{otherwise} \end{cases}$$

但是

$$f(x) = \inf_{z \in \mathbb{R}^m} F(x, z) = \begin{cases} 0 & x > 0 \\ 1 & x = 0 \\ \infty & x < 0 \end{cases}$$

不是闭的

命题 12.26

$$P(\text{epi}(F)) \subset \text{epi}(f) \subset \text{cl}(P(\text{epi}(F))),$$

其中 $P(x, z, w) = (x, w)$ 。

定理 12.7 (Preservation of Closedness Under Compactness)

设 $F : \mathbb{R}^{n+m} \rightarrow (-\infty, +\infty]$ 是闭且 proper 的凸函数, 定义 $f : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ 为 $f(x) = \inf_{z \in \mathbb{R}^m} F(x, z)$. 若存在某个 $\bar{x} \in \mathbb{R}^n, \bar{\gamma} \in \mathbb{R}$ 满足

$$\left\{ z \mid F(\bar{x}, z) \leq \bar{\gamma} \right\}$$

是非空且紧的, 则 f 是闭且 proper 的凸函数。同样的, 对任意 $x \in \text{dom}(f)$, 都有 $F(x, \cdot)$ 的极小值点是非空且紧的。

12.7 渐进序列

定义 12.25 (asymptotic sequence)

给定闭凸集序列 $\{C_k\}$, $\{x_k\}$ 称为 **渐进序列 (asymptotic sequence)**, 若

$$0 \neq x_k \in C_k, \|x_k\| \rightarrow +\infty, \lim_{k \rightarrow \infty} \frac{x_k}{\|x_k\|} = \frac{d}{\|d\|}$$

其中 d 是 C_k 共有的非零回收方向。

命题 12.27

任意无界序列 $\{x_k\}$ 中存在一个渐进子序列。

定义 12.26 (retractive)

$\{x_k\}$ 称为 **可收缩的 (retractive)**, 若存在某个非零回收方向 d , 存在某个 $\bar{k} > 0$ 使得

$$x_k - d \in C_k, \forall k \geq \bar{k}.$$

定义 12.27

一个闭凸集序列 $\{C_k\}$ 称为 **可收缩的 (retractive)**, 若其任意渐近序列都是可收缩的。

命题 12.28

设 $\{C_k\}$ 是闭凸集序列, X 是可收缩集, 即 $\bar{C}_k = X \cap C_k$ 是非空集。设

$$R_X \cap R \subset L$$

, 其中

$$R = \bigcap_{k=0}^{\infty} R_{C_k}, \quad L = \bigcap_{k=0}^{\infty} L_{C_k}$$

则 $\{\bar{C}_k\}$ 是可收缩的且 $\bigcap_{k=0}^{\infty} \bar{C}_k$ 非空。

证明. $\{\bar{C}_k\}$ 共有的回收方向为 $R_X \cap R$ 。对任意 $\{x_k\}$, 其渐近方向 $d \in R_X \cap R$, 由 $d \in L$ 有 $x_k - d \in C_k$, 由 X 可收缩有 $x_k - d \in X$, 从而 $x_k - d \in \bar{C}_k$ 。因此, $\{\bar{C}_k\}$ 是可收缩的。由 Weierstrass 定理即得。□

定理 12.8

设

$$f(x) = x'Qx + c'x, X = \left\{x \mid a'_j x + b_j \leq 0, j = 1, \dots, r\right\}$$

其中 Q 是半正定对称矩阵。若 f 在 X 上的最小值有限, 则 f 在 X 上存在最小值点。

证明. 我们有

$$\text{最小值点集} = \bigcap_{k=0}^{\infty} \left(X \cap \left\{x \mid x'Qx + c'x \leq \gamma_k\right\} \right)$$

其中 $\gamma_k \downarrow f^* = \inf_{x \in X} f(x)$ 。设 R_X 和 R_f 分别为 X 和 f 的回收锥, 则共有的回收锥为 $R_X \cap R_f$ 。注意到

$$R_f = \left\{d \mid Qd = 0, c'd \leq 0\right\}, L_f = \left\{d \mid Qd = 0, c'd = 0\right\}.$$

由 $f^* > -\infty$ 可知 $R_X \cap R_f \subseteq L_f$, 由此可知该闭凸集序列是可收缩的。由此由前一命题即得。□

命题 12.29

设 C 是非空闭凸集, 设 A 是矩阵其零空间为 $N(A)$ 。

- 若 $R_C \cap N(A) \subset L_C$, 则 AC 是闭的。
- 若 X 是可收缩的且 $R_X \cap R_C \cap N(A) \subset L_C$, 则 $A(X \cap C)$ 是可收缩的。

证明. 设 $\{y_k\} \subset AC$ 满足 $y_k \rightarrow \bar{y}$ 。我们来证明 $\bigcap_{k=0}^{\infty} C_k \neq \emptyset$, 其中

$$C_k = C \cap N_k, \quad N_k = \left\{x \mid Ax \in W_k\right\}, \quad W_k = \left\{z \mid \|z - \bar{y}\| \leq \|y_k - \bar{y}\|\right\}.$$

由于 $R_{N_k} = N(A)$, 故共有的回收锥为 $R_C \cap N(A)$ 。由假设可知该闭凸集序列是可收缩的。由 Weierstrass 定理即得。

□

命题 12.30

设 $C_1, \dots, C_m$ 是 $\mathbb{R}^n$ 的非空闭凸子集满足若 $d_1 + \dots + d_m = 0$ 且 $d_i \in R_{C_i}$, 则有 $d_i = 0$ 。则 $C_1 + \dots + C_m$ 是闭的。

证明. $C = C_1 \times \dots \times C_m$ 是 $\mathbb{R}^{nm}$ 的非空闭凸子集, 且其回收锥 $R_C = R_{C_1} \times \dots \times R_{C_m}$ 。定义 A 满足:

$$A(x_1, \dots, x_m) = x_1 + \dots + x_m$$

则 $N(A) = \{(x_1, \dots, x_m) \mid x_1 + \dots + x_m = 0\}$ 。由假设可知 $R_C \cap N(A) = \{0\}$, 由前一命题即得。

□

12.8 超平面

定义 12.28

一个超平面 (hyperplane) 是形如

$$H = \{x \mid a'x = b\}$$

的集合, 其中 $a \in \mathbb{R}^n, b \in \mathbb{R}$ 且 $a \neq 0$ 。

定义 12.29

我们称两个集合 C_1 和 C_2 被超平面 H 分离 (separated), 若

$$a'x \leq b, \forall x \in C_1, \quad a'x \geq b, \forall x \in C_2.$$

或者

$$a'x \geq b, \forall x \in C_1, \quad a'x \leq b, \forall x \in C_2.$$

若上述不等式均严格成立, 则称 C_1 和 C_2 被 H 严格分离 (strictly separated)。

定义 12.30

若 $\bar{x}$ 属于 C 的闭包和一个分离 C 超平面 H , 则称单点集 $\{\bar{x}\}$ 是 C 在 $\bar{x}$ 处的支撑 (supporting)。

定理 12.9 (supporting hyperplane theorem)

设 C 是凸集, $\bar{x}$ 不是 C 的内点, 则存在一个超平面经过 $\bar{x}$ 且支撑 C 。

证明. 取 $\{x_k\}$ 不属于 $\text{cl}(C)$ 且 $x_k \rightarrow \bar{x}$ 。设 $\hat{x}_k$ 为 x_k 在 $\text{cl}(C)$ 上的投影, 我们有对任意 $x \in \text{cl}(C)$,

$$a'_k x \geq a'_k x_k, \forall k = 0, 1, \dots$$

其中 $a_k = (x_k - \hat{x}_k) / \|x_k - \hat{x}_k\|$ 。令 $k \rightarrow \infty$ 即得。 $\square$

定理 12.10 (separating hyperplane theorem)

设 C_1 和 C_2 是 $\mathbb{R}^n$ 中的非空凸集且 $\text{ri}(C_1) \cap \text{ri}(C_2) = \emptyset$ ，则存在一个超平面分离 C_1 和 C_2 。

证明. 考虑凸集

$$C_1 - C_2 = \{x_1 - x_2 | x_1 \in C_1, x_2 \in C_2\}.$$

由于 C_1, C_2 是不交的，故 $0 \notin \text{ri}(C_1 - C_2)$ 。由支持超平面定理可知，存在 $0 \neq a \in \mathbb{R}^n$ 使得

$$a'x \geq 0, \forall x \in C_1 - C_2.$$

$\square$

推论 12.2 (strict separation theorem)

设 C_1 和 C_2 是 $\mathbb{R}^n$ 中的非空闭凸集且 $C_1 \cap C_2 = \emptyset$ 。若 C_1 是闭的， C_2 是紧的，则存在一个超平面严格分离 C_1 和 C_2 。

证明. 考虑 $C_1 - C_2$ ，由于 C_1 闭 C_2 紧， $C_1 - C_2$ 是闭的。设 $\bar{x}_1 - \bar{x}_2$ 是 0 在 $C_1 - C_2$ 上的投影，则 $\|\bar{x}_1 - \bar{x}_2\| > 0$ 。令 $a = \bar{x}_1 - \bar{x}_2$ ，则

$$a'x \geq \|a\|^2, \forall x \in C_1 - C_2.$$

$\square$

推论 12.3 (fundamental characterization)

集合 C 的凸包的闭包为所有包含 C 的闭半平面的交。

定义 12.31

我们称超平面 properly 分离两个集合 C_1 和 C_2 ，若其分离 C_1 和 C_2 且不完全包含 C_1 或 C_2 。

定理 12.11 (proper separation theorem)

设 C_1 和 C_2 是 $\mathbb{R}^n$ 中的非空闭凸集且 $C_1 \cap C_2 = \emptyset$ 。则存在一个超平面 properly 分离 C_1 和 C_2 当且仅当 $\text{ri}(C_1) \cap \text{ri}(C_2) = \emptyset$ 。

定义 12.32 (nonvertical)

一个 $\mathbb{R}^{n+1}$ 中的超平面伴有法向量 (μ, β) 称为非垂直的 (nonvertical)，若 $\beta \neq 0$ 。

其与第 $n+1$ 维轴交于 $\xi = (\mu/\beta)' \bar{u} + \bar{w}$ ，其中 $(\bar{u}, \bar{w})$ 是超平面上的任意点。

定理 12.12 (novertical hyperplane theorem)

设 C 是 $\mathbb{R}^{n+1}$ 中的非空闭凸集且 C 不包含某个垂直方向的直线，则：

- C 包含于某个非垂直超平面的一侧，即存在 $\mu \in \mathbb{R}^n, \beta \in \mathbb{R} \setminus \{0\}$ 使得

$$\mu' u + \beta w \geq \gamma, \forall (u, w) \in C.$$

- 若 $(\bar{u}, \bar{w}) \notin \text{cl}(C)$ ，则存在一个非垂直超平面严格分离 C 和 $(\bar{u}, \bar{w})$ 。

证明. 由于 C 不包含任意方向的垂线，故 $\text{ri}(C)$ 不包含任意方向的垂线，而 $\text{ri}(C)$ 和 $\text{cl}(C)$ 有相同的回收锥，故 $\text{cl}(C)$ 也不包含任意方向的垂线。故我们只考虑 C 是闭的。

(a): C 为所有包含 C 的闭半平面的交，若这些半平面均垂直，则 C 包含某个垂直方向的直线，矛盾。

(b): 由支持超平面定理可知，存在一个超平面分离 C 和 $(\bar{u}, \bar{w})$ 。若该超平面垂直，则 C 包含某个垂直方向的直线，矛盾。 $\square$

定义 12.33 (conjugate convex function)

对任意 $f: \mathbb{R}^n \rightarrow [-\infty, +\infty]$ ，其共轭凸函数 (conjugate convex function) 定义为

$$f^*(y) = \sup_{x \in \mathbb{R}^n} (x'y - f(x)), \forall y \in \mathbb{R}^n.$$

命题 12.31

f^* 是闭且凸的函数。

定理 12.13 (conjugacy theorem)

若 f 是闭且 proper 的凸函数，则 $f^{**} = f$ 。

定理 12.14 (conjugacy theorem)

设 $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$ ，设 $\check{\text{cl}}f$ 为其凸闭包，设 f^* 是其共轭凸函数，考虑 f^* 的共轭

$$f^{**}(x) = \sup_{y \in \mathbb{R}^n} (x'y - f^*(y)), \forall x \in \mathbb{R}^n.$$

- $f(x) \geq f^{**}(x), \forall x \in \mathbb{R}^n$.
- 若 f 是凸的，则 f, f^*, f^{**} 中任意一者是 proper 给出另两者是 proper。
- 若 f 是闭且 proper 的凸函数，则 $f = f^{**}$ 。
- 若 $\check{\text{cl}}f(x) > -\infty$ ，则 $\check{\text{cl}}f = f^{**}$ 。

证明. (a): 显然。

(b): 由定义即得。

(c): 设 $g = f^{**}$, 则 g 是闭且 proper 的凸函数且 $f(x) \geq g(x), \forall x \in \mathbb{R}^n$. 由支持超平面定理可知, f 的任意支撑超平面也是 g 的支撑超平面, 故 $f = g$.

(d): 由 (a) 和 (c) 即得。 $\square$

例 12.3

- 对 $\frac{1}{p} + \frac{1}{q} = 1, p > 1$ 的 $f(x) = \frac{1}{p}\|x\|^p$, 其共轭函数为 $f^*(y) = \frac{1}{q}\|y\|^q$.
- $f(x) = \frac{1}{2}x'Qx + a'x + b$, Q 为半正定矩阵, 则 $f^*(y) = \frac{1}{2}(y-a)'Q^{-1}(y-a) - b$.
- $f(x) = p(A(x-c)) + a'x + b$, 则 $f^*(y) = q((A')^{-1}(y-a)) + c'y + d$, 其中 q 是 p 的共轭函数, $d = -(c'a + b)$.

定义 12.34 (support function)

X 的示性函数 δ_X

$$\sigma_X(y) = \sup_{x \in X} y'x$$

称为 X 的支撑函数 (support function)。

命题 12.32

$\text{epi}(\sigma_X)$ 是闭凸锥。

命题 12.33

$X, \text{cl}(X), \text{conv}(X), \text{cl}(\text{conv}(X))$ 有相同的支撑函数。

命题 12.34

若 C 是锥,

$$\sigma_C(y) = \begin{cases} 0 & \text{if } y'x \leq 0, \forall x \in C \\ +\infty & \text{otherwise} \end{cases}$$

即 σ_C 是 C^* 的示性函数, 其中

$$C^* = \{y \mid y'x \leq 0, \forall x \in C\}$$

称为 C 的极锥 (polar cone)。

由共轭定理可知 C^* 的极锥为 $\text{cl}(\text{conv}(C))$, 这称为极锥定理 (polar cone theorem)。

例 12.4

若 $C = \text{cone}(\{a_1, \dots, a_r\})$, 则

$$C^* = \{x \mid a'_j x \leq 0, j = 1, \dots, r\}.$$

引理 12.1 (Farkas' lemma)

$$(C^*)^* = C$$

我们已经将对偶的概念从集合的对偶延伸至函数的对偶，现在我们考虑问题的对偶。

12.9 最小公共点和最大相交点问题

我们介绍一对基本问题。

定义 12.35

设 M 是 $\mathbb{R}^{n+1}$ 的非空子集。

- **最小公共点问题 (minimum common point problem):** 寻找点 $(x, w) \in M$ 使得 w 最小化。
- **最大相交点问题 (maximum intersecting point problem):** 考虑所有包含了 M 的非垂直超平面，寻找其中与第 $n+1$ 维轴交点最大的超平面。

对最小公共问题 $w^* = \inf\{w | (x, w) \in M\}$ ，对最大相交点问题，我们考虑法向量为 $(\mu, 1)$ 的超平面其相交点 ξ 满足

$$\xi \leq w + \mu'u, \forall (u, w) \in M$$

从而最大相交点问题可表述为最大化 $\xi \leq \inf_{(u, w) \in M} (w + \mu'u), \mu \in \mathbb{R}^n$ ，即

$$\text{maximize } q(\mu) = \inf_{(u, w) \in M} (w + \mu'u), \quad \text{subject } \mu \in \mathbb{R}^n.$$

容易得到如下的弱对偶 (weak duality) 结果：

$$q(\mu) = \inf_{(u, w) \in M} (w + \mu'u) \leq \inf_{(0, w) \in M} w = w^*, \forall \mu \in \mathbb{R}^n.$$

我们称强对偶 (strong duality) 成立，若 $q^* = w^*$ 。

例 12.5

$$M = \text{epi}(p)$$

其中 $p: \mathbb{R}^n \rightarrow [-\infty, \infty]$ ，则 $w^* = p(0)$ ，且

$$q(\mu) = \inf_{(u, w) \in \text{epi}(p)} (w + \mu'u) = \inf_{u \in \mathbb{R}^n} (p(u) + \mu'u) = -p^*(-\mu).$$

进而有

$$q^* = \sup_{\mu \in \mathbb{R}^n} q(\mu) = \sup_{\mu \in \mathbb{R}^n} \{0 \cdot (-\mu) - p^*(-\mu)\} = p^{**}(0)$$

例 12.6

考虑更一般的最小化 $f: \mathbb{R}^n \rightarrow [-\infty, \infty]$ 的问题。设 $F: \mathbb{R}^{n+r} \rightarrow [-\infty, \infty]$ 为

$$f(x) = F(x, 0), \forall x \in \mathbb{R}^n.$$

考虑扰动函数 (perturbation function)

$$p(u) = \inf_{x \in \mathbb{R}^n} F(x, u), \forall u \in \mathbb{R}^r.$$

且 MC/MC 框架为 $M = \text{epi}(p)$ 。

其最小公共值

$$w^* = p(0) = \inf_{x \in \mathbb{R}^n} F(x, 0) = \inf_{x \in \mathbb{R}^n} f(x).$$

其对偶函数为

$$q(\mu) = \inf_{u \in \mathbb{R}^r} (p(u) + \mu'u) = \inf_{(x, u) \in \mathbb{R}^{n+r}} (F(x, u) + \mu'u).$$

故 $q(\mu) = -F^*(0, -\mu)$, 其中 F^* 是 F 的共轭函数, 视作 (x, u) 的函数。

我们有

$$q^* = \sup_{\mu \in \mathbb{R}^r} q(\mu) = - \inf_{\mu \in \mathbb{R}^r} F^*(0, -\mu) = - \inf_{\mu \in \mathbb{R}^r} F^*(0, \mu)$$

且弱对偶性形如:

$$w^* = \inf_{x \in \mathbb{R}^n} F(x, 0) \geq - \inf_{\mu \in \mathbb{R}^r} F^*(0, \mu) = q^*.$$

例 12.7

考虑约束优化: 最小化 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在

$$C = \{x \in X \mid g(x) \leq 0\}$$

其中 $X \subset \mathbb{R}^n$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^r$ 。

考虑扰动约束集 (perturbed constraint set)

$$C_u = \{x \in X \mid g(x) \leq u\}, \forall u \in \mathbb{R}^r.$$

与函数

$$F(x, u) = \begin{cases} f(x) & \text{if } x \in C_u \\ +\infty & \text{otherwise} \end{cases}$$

这满足 $F(x, 0) = f(x), \forall x \in C$ 。

例 12.8

考虑扰动函数

$$p(u) = \inf_{x \in \mathbb{R}^n} F(x, u) = \inf_{x \in C_u} f(x), \forall u \in \mathbb{R}^r.$$

且 MC/MC 框架为 $M = \text{epi}(p)$ 。

设 $L(x, \mu) = f(x) + \mu'g(x)$ 为 Lagrange 函数, 则

$$q(\mu) = \inf_{u \in \mathbb{R}^r} \{p(u) + \mu'u\} = \inf_{u \in \mathbb{R}^r, x \in X, g(x) \leq u} \{f(x) + \mu'u\} = \begin{cases} \inf_{x \in X} L(x, \mu) & \text{if } \mu \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

例 12.9

考虑线性规划问题: 最小化 $c'x$ 在

$$C = \left\{x \in \mathbb{R}^n \mid a'_j x \geq b_j, j = 1, \dots, r\right\}$$

其中 $c \in \mathbb{R}^n$, $a_j \in \mathbb{R}^n$, $b_j \in \mathbb{R}$ 。

对 $\mu \geq 0$, 对偶函数

$$q(\mu) = \inf_{x \in \mathbb{R}^n} \left\{ c'x + \sum_{j=1}^r \mu_j (b_j - a'_j x) \right\} = \begin{cases} b'\mu & \text{if } c = \sum_{j=1}^r \mu_j a_j \\ -\infty & \text{otherwise} \end{cases}$$

因此其对偶问题为

$$\text{maximize } b'\mu, \quad \text{subject to } A\mu = c, \mu \geq 0.$$

定义 12.36 (minimax problem)

给定函数 $\phi: X \times Z \rightarrow \mathbb{R}$, 其中 $X \subset \mathbb{R}^n$, $Z \subset \mathbb{R}^m$, 考虑

$$\text{minimize } \sup_{z \in Z} \phi(x, z), \quad \text{subject to } x \in X.$$

或

$$\text{maximize } \inf_{x \in X} \phi(x, z), \quad \text{subject to } z \in Z.$$

我们总是有

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) \leq \inf_{x \in X} \sup_{z \in Z} \phi(x, z).$$

一个关键问题是: 何时等号成立?

例 12.10

对于约束问题

$$\text{minimize } f(x), \quad \text{subject to } g(x) \leq 0, x \in X$$

其 Lagrange 函数为 $L(x, \mu) = f(x) + \mu'g(x)$, 其原始问题为

$$\min_{x \in X} \sup_{\mu \geq 0} L(x, \mu)$$

其对偶问题为

$$\max_{\mu \geq 0} \inf_{x \in X} L(x, \mu).$$

其等号成立需考虑何时

$$\sup_{\mu \geq 0} \inf_{x \in X} L(x, \mu) = w^* \stackrel{?}{=} q^* = \inf_{x \in X} \sup_{\mu \geq 0} L(x, \mu).$$

12.10 零和游戏**定义 12.37 (zero sum games)**

两个玩家：第一个选择 $i \in \{1, \dots, n\}$, 第二个选择 $j \in \{1, \dots, m\}$ 。支付矩阵为 $A \in \mathbb{R}^{n \times m}$, 则第一个玩家从第二个玩家处获得 a_{ij} 。两个玩家选择

$$x = (x_1, \dots, x_n) \quad z = (z_1, \dots, z_m)$$

其中 x_i 和 z_j 分别为两个玩家选择 i 和 j 的概率。则第一个玩家的期望收益为

$$\phi(x, z) = x'Az = \sum_{i=1}^n \sum_{j=1}^m a_{ij}x_iz_j.$$

第一个玩家的目标是最大化其最小收益: $\text{minimize } \max_z \phi(x, z)$

第二个玩家的目标是最小化其最大损失: $\text{maximize } \min_x \phi(x, z)$

第六部分 PDE 数值解

13 基础定义

13.1 二阶 PDE 的分类

我们考虑一般形式的二阶 PDE:

$$\sum_{i,j=1}^n a_{ij}(x)u_{x_i x_j} + \sum_{i=1}^n b_i(x)u_{x_i} + c(x)u = f(x)$$

其中 $u = u(x_1, \dots, x_n)$ 是未知解, a_{ij}, b_i, c, f 为 $x = (x_1, \dots, x_n)$ 的已知函数, 且 $a_{ij} = a_{ji}$ 。设 $A = [a_{ij}]_{i,j=1}^n$, 则 PDE 通过 A 分为如下几类:

定义 13.1

- 若 A 的特征值符号为 $\{+, +, \dots, +\}$ 或 $\{-, -, \dots, -\}$, 则称为 **椭圆型 (elliptic)** PDE。
- 若 A 的特征值符号为 $\{+, +, \dots, +, 0\}$ 或 $\{-, -, \dots, -, 0\}$, 则称为 **抛物型 (parabolic)** PDE。
- 若 A 的特征值符号为 $\{+, +, \dots, +, -\}$ 或 $\{-, -, \dots, -, +\}$, 则称为 **双曲型 (hyperbolic)** PDE。

考虑 $n = 2$ 与 $u = u(x, y)$, 则 PDE 可写成

$$au_{xx} + 2bu_{xy} + cu_{yy} + du_x + eu_y + fu = g$$

我们有

$$\|A\| = ac - b^2 = \begin{cases} > 0 & \text{elliptic} \\ = 0 & \text{parabolic} \\ < 0 & \text{hyperbolic} \end{cases}$$

为了简化书写, 我们定义如下记号:

定义 13.2

- 对 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 定义 $\nabla f = \text{grad} f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)^T$
- 对 $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, 定义 $\nabla \cdot F = \text{div} F = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}$
- 对 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 定义 $\Delta f = \nabla \cdot \nabla f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}$

例 13.1 (Elliptic)

考虑带边界条件的一般形式

$$\begin{cases} \mathcal{L}u = f, & x \in \Omega \\ u = g, & x \in \partial\Omega \end{cases}$$

其中 $\mathcal{L}$ 是椭圆型算子。进一步的例子有 $\Delta u = 0, x \in \Omega$ (Laplace 方程), $\Delta u = f, x \in \Omega$ (Poisson 方程)。

例 13.2 (Parabolic)

考虑带边界条件的一般形式

$$\begin{cases} u_t = \mathcal{L}u + f, \\ u|_{t=0} = u_0, \\ u|_{x \in \partial\Omega} = g \end{cases}$$

其中 $\mathcal{L}$ 是椭圆型算子。进一步的例子有 $u_t = \beta \Delta u + f$ (热方程)。

例 13.3 (Hyperbolic)

考虑带边界条件的一般形式

$$\begin{cases} u_{tt} = \mathcal{L}u + f, \\ u|_{t=0} = u_0, u_t|_{t=0} = \hat{u}_0, \\ u|_{x \in \partial\Omega} = g \end{cases}$$

其中 $\mathcal{L}$ 是椭圆型算子。进一步的例子有 $u_{tt} = c^2 \Delta u + f$ (波动方程)。

下面我们介绍另一种分类方法。

设 $\vec{u}(t, x) : \mathbb{R} \times \mathbb{R}^m$ 且 $\vec{u}_t + A\vec{u}_x = \vec{F}, A \in \mathbb{R}^{m \times m}$ 。

定义 13.3

- 若 A 无实特征值, 则称为 **椭圆型 (elliptic)** PDE。
- 若 A 有 m 个互不相同的实特征向量和 m 个实特征值, 即可对角化, 则称为 **双曲型 (hyperbolic)** PDE。
- 若 A 有 m 个互不相同的实特征值, 则称为 **严格双曲型 (strictly hyperbolic)** PDE。

事实上我们可以从物理角度考虑, 视 $u(x, t)$ 为密度函数, 由如下方式改变 u :

$$\begin{cases} \text{flux: } F, \\ \text{source: } G, \end{cases}$$

则 $u_t + F_x = G$ 。于是有如下三种有物理意义的例子:

例 13.4

- Advection: 若 $F = au$, $G = 0$, 则 $u_t + au_x = 0$ 。
- Diffusion: 若 $F = -\beta u_x$, $G = 0$, 则 $u_t = \beta u_{xx}$ 。
- Steady: 若 $F = -\beta u_x$, $G \neq 0$, $u_t \rightarrow 0 (t \rightarrow \infty)$, 则 $-\beta u_{xx} = G$ 。

13.2 Fourier 分析

回忆, 我们有

命题 13.1

- 对 $e^{i\xi x}$, 有 $\frac{d}{dx}e^{i\xi x} = i\xi e^{i\xi x}$ 。
- 定义 $D_h V_j := \frac{1}{2h}(V_{j+1} - V_{j-1})$ 是定义在 $\{V_j\}$ 上的算子。对 $\omega_j = e^{ijh\xi}$, $D_h \omega_j = \frac{1}{2h}(\omega_{j+1} - \omega_{j-1}) = \frac{1}{h} \sin(h\xi) \omega_j$ 。

于是我们可以将 $e^{i\xi x}$ 视作 $\frac{d}{dx}$ 与 D_h 的特征函数。

定义 13.4

设 $V(x) \in L^2(\mathbb{R})$, 即 $\|V\|_{L^2}^2 = \int_{\mathbb{R}} \|V(x)\|^2 dx < \infty$, 则 **Fourier 变换 (Fourier Transformation)** 定义为

$$\mathcal{F}(V(x)) = \hat{V}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} V(x) e^{-i\xi x} dx$$

反过来, $\hat{V}(\xi)$ 有逆 Fourier 变换:

$$\mathcal{F}^{-1}(\hat{V}(\xi)) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{V}(\xi) e^{i\xi x} d\xi = V(x)$$

定理 13.1 (Parseval equation)

$$\|V(x)\|_{L^2} = \|\hat{V}(\xi)\|_{L^2}$$

我们可以将其运用到 PDE 中。

对 $u(x, t) \in L^2(\mathbb{R})$, 则 $u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{u}(\xi, t) e^{i\xi x} d\xi$, 则

$$\begin{cases} u_t(t, x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{u}(\xi, t) e^{i\xi x} d\xi \\ u_x(t, x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} i\xi \hat{u}(\xi, t) e^{i\xi x} d\xi \end{cases}$$

例 13.5

对于 Advection 方程

$$\begin{cases} u_t + au_x = 0 \\ u(x, 0) = \eta(x) \end{cases}$$

由 FT 得 $\hat{u}_t + a \cdot i\xi \hat{u} = 0$ 。解 ODE 得 $\hat{u}(\xi, t) = e^{-i\xi x a t} \hat{\eta}(\xi)$ 再用 IFT 得

$$\begin{aligned} u(x, t) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{u}(\xi, t) e^{i\xi x} d\xi \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi(x-at)} \hat{\eta}(\xi) d\xi \\ &= \eta(x - at) \end{aligned}$$

例 13.6

对于热方程

$$\begin{cases} u_t = \beta u_{xx} \\ u(x, 0) = \eta(x) \end{cases}$$

由 FT 得 $\hat{u}_t = -\beta \xi^2 \hat{u}$ 。解 ODE 得 $\hat{u}(\xi, t) = e^{-\beta \xi^2 t} \hat{\eta}(\xi)$

例 13.7

对于 Dispersive 方程

$$\begin{cases} u_t = cu_{xxx} \\ u(x, 0) = \eta(x) \end{cases}$$

由 FT 得 $\hat{u}_t = -i\xi^3 c \hat{u}$ 。解 ODE 得 $\hat{u}(\xi, t) = e^{-i\xi^3 ct} \hat{\eta}(\xi)$ ，进一步 IFT 得：

$$u(x, t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi(x-c\xi^2 t)} \hat{\eta}(\xi) d\xi$$

14 椭圆 PDE 的有限差分方法

首先我们给出推广的定义。考虑高维 PDE

$$-\mathcal{L}u = f, \quad \mathcal{L} = \sum_{k=1}^{2m} \sum_{i_1, \dots, i_k=1}^n a_{i_1 \dots i_k}(x) \frac{\partial^k}{\partial x_{i_1} \dots \partial x_{i_k}} + \dots + a_0$$

定义 14.1

我们称其为椭圆型的，若

$$\forall x \in \Omega, \forall \xi \in \mathbb{R}^n \setminus \{0\}, \quad \sum_{i_1 \dots i_{2m}=1}^n a_{i_1 \dots i_{2m}}(x) \xi_{i_1} \dots \xi_{i_{2m}} \geq \alpha(x) \sum_{i=1}^n \xi_i^{2m}$$

其中 $\alpha(x) > 0$ 。若 $\inf_{x \in \Omega} \alpha(x) > 0$ ，则称其为一致严格椭圆型。对 $n = 2$ 的情形，该定义等价于 $A \succ 0$ 。

定义 14.2

对 $p \geq 1$, 向量 $a \in \mathbb{R}^n$ 的 p -范数定义为 $\|a\|_p = (\sum_{i=1}^n |a_i|^p)^{\frac{1}{p}}$ 。对矩阵 $A \in \mathbb{R}^{n \times m}$, 其 p -范数定义为

$$\|A\|_p = \max \left\{ \|Ax\|_p \mid x \in \mathbb{R}^m, \|x\|_p = 1 \right\}$$

A 的 Schatten 范数定义为:

$$\|A\|_{S_p} = \left[\text{Tr} \left(\left(\sqrt{A^T A} \right)^p \right) \right]^{\frac{1}{p}} = \left[\sum_i s_i^p \right]^{\frac{1}{p}}$$

其中 s_i 为 A 的奇异值。