

数值分析期末精要

杨锷 2313878

2026 年 1 月 24 日

没啥好说的，只能说计算量非常大，其他的倒也没什么，考试大家都做不完哈哈。。。

1 引论

1.1 进制

定义 1.1 (进制转换).

$$\begin{aligned}x &= \pm(a_{n-1}\beta^{n-1} + a_{n-2}\beta^{n-2} + \cdots + a_0\beta^0 + a_{-1}\beta^{-1} + a_{-2}\beta^{-2} + \cdots + a_{-m}\beta^{-m}) \\ &= \pm(a_{n-1}a_{n-2}\cdots a_0 \cdot a_{-1}a_{-2}\cdots a_{-m})_\beta\end{aligned}$$

 每年都考.

具体来说就是 n 进制转十进制按照整数部分和小数部分分部分按此方展开即可:

$$(10101)_2 = 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 = (21)_{10},$$

$$-(10.101)_2 = -(1 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3}) = -(2.625)_{10}.$$

十进制转 n 进制可以用短除法或者挨个挨个减来试.

特殊的: 2 进制和 8 进制之间的转换可以 3 位 3 位来看.

1.2 误差

定义 1.2 (舍入和截断). 舍入就是四舍五入, 截断就是之间全部去掉.

 我们很容易可以想到舍入的精度最多不超过精度的一半, 但是截断只能做到不超过精度区间.

定义 1.3 (误差理论). 设数 x 是某个量的精确值, 数 x^* 是该量的已知近似值, 记

$$E(x) = x - x^*, \quad (1)$$

称 $E(x)$ 为近似数 x^* 的绝对误差, 简称误差.

我们可以估计出绝对误差的上限, 即可求出一个正数 η , 使得

$$|x - x^*| \leq \eta. \quad (2)$$

η 称为近似数 x^* 的绝对误差限. 有了绝对误差限就可得到 x (精确值) 的范围.

近似数 x^* 的绝对误差和精确值的比, 即

$$E_r(x) = \frac{x - x^*}{x}, \quad (3)$$

称为 x^* 的**相对误差**。

我们往往可以估计出相对误差的上限, 即可以求出一个正数 δ , 使得

$$|E_r(x)| \leq \delta. \quad (4)$$

δ 称为 x^* 的**相对误差限**。



定义 1.4 (有效数字). 设精确值 x 的近似值为


$$x^* = \pm 10^m \times 0.a_1a_2 \cdots a_k \cdots a_n, \quad (5)$$

其中, $a_i (i = 1, 2, \dots, n)$ 是 0 到 9 中的某一整数, 并且 $a_1 \neq 0$ 。若 x^* 的绝对误差限满足

$$|x - x^*| \leq \frac{1}{2} \times 10^{m-k}, \quad (6)$$

则称 a_k 为**有效数字**。显然, 若 a_k 为有效数字, 那么 a_1, \dots, a_{k-1} 都是有效数字。如果 x^* 的每一位都是有效数字, 那么 x^* 称为具有 n 位有效数字的**有效数**。



 **注.** 从定义可知, 若 x^* 是 x 经“四舍五入”得到的, 那么它肯定是有有效数。但 x 的近似数 x^* 的每一位数字不一定是有效数字。

例如,

$$e = 2.7182818284 \cdots$$

对于 e , 如果直接截取它的前 2~5 位, 四个数是

$$2.7, \quad 2.71, \quad 2.718, \quad 2.7182,$$

其中 2.7 与 2.71 均有二位有效数字, 2.718 与 2.7182 均有四位有效数字; 相对的, 2.72 是三位有效数字, 2.7183 是五位有效数字。

 我们通过一个简单的例子来回忆一下误差传播:

例 1.1. 计算积分

$$I_n = \int_0^1 x^n e^{x-1} dx, \quad n = 1, 2, \dots$$

利用分部积分得

$$\begin{aligned} \int_0^1 x^n e^{x-1} dx &= x^n e^{x-1} \Big|_0^1 - n \int_0^1 x^{n-1} e^{x-1} dx \\ &= 1 - nI_{n-1}, \end{aligned}$$

即得递推公式

$$I_n = 1 - nI_{n-1}, \quad n = 2, 3, \dots \quad (7)$$

考虑误差传播情况。由式 (7) 可得

$$I_2 = 1 - 2I_1.$$



设近似值为 $\tilde{I}_2 = 1 - 2\tilde{I}_1$, 则

$$I_2 - \tilde{I}_2 = -2(I_1 - \tilde{I}_1).$$

记 $I_1 - \tilde{I}_1 = \varepsilon$ 为 I_1 值的误差。式 (1.31) 表明, \tilde{I}_1 的误差传播到 \tilde{I}_2 时, 误差已增加了 2 倍。一般地, 有

$$I_n - \tilde{I}_n = (-1)^{n-1} n! (I_1 - \tilde{I}_1) = (-1)^{n-1} n! \varepsilon.$$

这表明若 \tilde{I}_1 有 ε 的误差, 那么 \tilde{I}_n 就有 $(-1)^{n-1} n!$ 倍 ε 的误差, 由此可见当 n 充分大时计算严重不可靠。

另一方面, 若使用逆递推形式 (由 (7) 改写为)

$$I_{n-1} = \frac{1 - I_n}{n}, \quad n = N, N-1, \dots, 2$$

则误差传播情况完全不同。此时有

$$I_1 - \tilde{I}_1 = -\frac{1}{2}(I_2 - \tilde{I}_2) = \frac{1}{3!}(I_3 - \tilde{I}_3) = \dots = (-1)^{n-1} \frac{1}{n!}(I_n - \tilde{I}_n).$$

这说明误差在逆推过程中传播时不仅没有增大, 反而被因子 $\frac{1}{n!}$ 缩小, 故 I_n 的逆推过程是稳定的, 计算的结果可靠。

 注. 顺递推过程误差被放大, 但是逆递推过程误差反而减小.

定义 1.5 (误差传播因子). 设数值计算过程中某一步的输出 y 可表示为输入 x_1, x_2, \dots, x_n 的函数:

$$y = f(x_1, x_2, \dots, x_n),$$

且 x_i 具有误差 $\varepsilon_i = \tilde{x}_i - x_i$ (\tilde{x}_i 为近似值)。若误差 ε_i 足够小, 则 y 的误差 $\varepsilon_y = \tilde{y} - y$ 可近似由一阶泰勒展开表示为:

$$\varepsilon_y \approx \sum_{i=1}^n \frac{\partial f}{\partial x_i} \varepsilon_i.$$

称系数 $\frac{\partial f}{\partial x_i}$ 为误差 ε_i 传播到 y 的**误差传播因子** (或**误差放大因子**)。

特别地, 若 $y = g(x)$ 为单变量函数, 则误差传播因子为 $g'(x)$, 且

$$\varepsilon_y \approx g'(x) \varepsilon_x.$$

当 $|g'(x)| > 1$ 时, 误差被放大; 当 $|g'(x)| < 1$ 时, 误差被抑制; 当 $|g'(x)| = 1$ 时, 误差等幅传播。

 这个书上没有, 但不知道为什么 24-25 年考了.

 注. 实际上就是求一阶导的得到的系数, 例如求 $x^{1/3}$ 的误差传播因子:

现考虑函数

$$y = f(x) = x^{1/3}.$$

其导数为

$$f'(x) = \frac{d}{dx} (x^{1/3}) = \frac{1}{3} x^{-2/3}.$$

因此, 误差传播因子为

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

进而, x 的误差 ε_x 传播到 y 的近似误差为:

$$\varepsilon_y \approx \frac{1}{3}x^{-2/3} \cdot \varepsilon_x.$$

若用相对误差表示, 记 x 的相对误差为 $\delta_x = \frac{\varepsilon_x}{x}$, y 的相对误差为 $\delta_y = \frac{\varepsilon_y}{y}$. 由 $y = x^{1/3}$ 得

$$\delta_y \approx \frac{1}{3}\delta_x.$$

即相对误差传播因子为 $\frac{1}{3}$.

2 插值逼近

这一章设计的主要插值问题: Lagrange 型多项式插值、Hermite 型多项式插值、样条插值、最佳一致、最佳平方逼近、最小二乘逼近.

定理 2.1. 设 n 为一非负整数, $f(x)$ 在点 x_0 的某一邻域 $I_\delta = (x_0 - \delta, x_0 + \delta)$ 有 $n+1$ 阶连续导数, 则对 $\forall x \in I_\delta$, 有

$$f(x) = p_n(x) + R_n(x), \quad (8)$$

其中, n 次 Taylor 逼近多项式 $p_n(x)$ 和误差余项 $R_n(x)$ 分别为

$$p_n(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n, \quad (9)$$

$$R_n(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_0)^{n+1}, \quad (10)$$

其中, ξ 为 x_0 与 x 之间的某个点。



2.1 Lagrange 型多项式插值

设函数 $y = f(x)$ 在区间 $[a, b]$ 上有定义, 且已知它在 $n+1$ 个互异点 $a \leq x_0 < x_1 < \cdots < x_n \leq b$ 上的函数值 y_0, y_1, \dots, y_n , 即

$$y_i = f(x_i), \quad i = 0, 1, \dots, n.$$

寻求一个次数不超过 n 的多项式 $P_n(x)$, 使得

$$P_n(x_i) = y_i, \quad i = 0, 1, \dots, n, \quad (11)$$

则称 $P_n(x)$ 为 $f(x)$ 的 n 次插值多项式, 点 x_0, x_1, \dots, x_n 称为插值节点. 条件 (11) 称为插值条件.

定义 2.1 (Lagrange 基函数). 定义 $n+1$ 个 n 次多项式

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x-x_j}{x_i-x_j}, \quad i = 0, 1, \dots, n,$$



称 $l_i(x)$ 为关于节点 x_0, x_1, \dots, x_n 的 **Lagrange 插值基函数**。它们满足性质

$$l_i(x_j) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

定理 2.2 (Lagrange 插值多项式). 满足插值条件 (11) 的 n 次插值多项式存在且唯一, 并可表示为

$$L_n(x) = \sum_{i=0}^n y_i l_i(x) = \sum_{i=0}^n y_i \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}. \quad (12)$$

$L_n(x)$ 称为 **Lagrange 插值多项式**。

定义 2.2 (插值余项). 若 $f(x)$ 在 $[a, b]$ 上 $n+1$ 阶连续可导, 则对任意 $x \in [a, b]$, Lagrange 插值余项为

$$R_n(x) = f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i), \quad (13)$$

其中 $\xi \in (a, b)$ 且与 x 有关。

2.1.1 差商

定义 2.3 (差商). 设函数 $f(x)$ 在互异点 x_0, x_1, \dots, x_n 上有定义, 则

1. 零阶差商定义为函数值本身:

$$f[x_i] = f(x_i), \quad i = 0, 1, \dots, n.$$

2. 一阶差商定义为:

$$f[x_i, x_{i+1}] = \frac{f[x_{i+1}] - f[x_i]}{x_{i+1} - x_i}.$$

3. 一般地, k 阶差商可由 $k-1$ 阶差商递归定义为:

$$f[x_i, x_{i+1}, \dots, x_{i+k}] = \frac{f[x_{i+1}, \dots, x_{i+k}] - f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}.$$

性质 2.1 (差商的性质).

1. **对称性**: 差商的值与节点的排列顺序无关, 即

$$f[x_{i_0}, x_{i_1}, \dots, x_{i_k}] = f[x_0, x_1, \dots, x_k],$$

其中 (i_0, i_1, \dots, i_k) 是 $(0, 1, \dots, k)$ 的任一排列。

2. **差商与导数的关系**: 若 $f(x)$ 在区间 $[a, b]$ 上 k 阶连续可导, 且 $x_0, x_1, \dots, x_k \in [a, b]$ 互异, 则存在 $\xi \in (\min x_i, \max x_i)$, 使得

$$f[x_0, x_1, \dots, x_k] = \frac{f^{(k)}(\xi)}{k!}.$$

3. **线性性**: 对于任意常数 α, β ,

$$(\alpha f + \beta g)[x_0, x_1, \dots, x_k] = \alpha f[x_0, x_1, \dots, x_k] + \beta g[x_0, x_1, \dots, x_k].$$

2.1.2 Neville 插值公式

定义 2.4 (Neville 插值公式). 设 $P_{i,i+1,\dots,i+k}(x)$ 表示以 $x_i, x_{i+1}, \dots, x_{i+k}$ 为插值节点的 k 次插值多项式, 则 Neville 公式给出了这些插值多项式之间的递推关系:

$$P_i(x) = f(x_i),$$
$$P_{i,i+1,\dots,i+k}(x) = \frac{(x-x_i)P_{i+1,\dots,i+k}(x) - (x-x_{i+k})P_{i,\dots,i+k-1}(x)}{x_{i+k} - x_i}, \quad k \geq 1.$$

特别地, 当 $k = n$ 时, $P_{0,1,\dots,n}(x)$ 即为以所有 $n+1$ 个节点 x_0, x_1, \dots, x_n 为插值节点的 n 次插值多项式。

注. 即两个低次的插值多项式可以通过“线性插值”得到高一次的插值多项式。

定理 2.3 (Neville 公式的等价形式). Neville 公式也可表示为如下显式形式:

$$P_{i,\dots,i+k}(x) = \frac{(x-x_i)f[x_{i+1}, \dots, x_{i+k}] - (x-x_{i+k})f[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i},$$

并且最终插值多项式可写为

$$P_{0,\dots,n}(x) = \sum_{i=0}^n f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x-x_j),$$

这正是 Newton 插值多项式的形式。

2.1.3 Newton 插值

定义 2.5 (Newton 插值多项式). 设 $f(x)$ 在 $n+1$ 个互异节点 x_0, x_1, \dots, x_n 上的函数值已知。形如

$$N_n(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \dots + a_n(x-x_0)(x-x_1)\dots(x-x_{n-1})$$

的多项式称为 **Newton 插值多项式**, 其中 a_0, a_1, \dots, a_n 为待定系数, 由插值条件

$$N_n(x_i) = f(x_i), \quad i = 0, 1, \dots, n$$

唯一确定。

定理 2.4 (Newton 插值公式). Newton 插值多项式可显式表示为

$$N_n(x) = f[x_0] + f[x_0, x_1](x-x_0) + f[x_0, x_1, x_2](x-x_0)(x-x_1) + \dots + f[x_0, x_1, \dots, x_n] \prod_{j=0}^{n-1} (x-x_j), \quad (14)$$

其中 $f[x_0, x_1, \dots, x_k]$ 表示 $f(x)$ 关于节点 x_0, x_1, \dots, x_k 的 k 阶差商。



定义 2.6 (Newton 插值基函数). 称

$$\begin{aligned}\pi_0(x) &= 1, \\ \pi_k(x) &= \prod_{j=0}^{k-1} (x - x_j), \quad k = 1, 2, \dots, n\end{aligned}$$

为 *Newton 插值基函数*。Newton 插值多项式可写为

$$N_n(x) = \sum_{k=0}^n a_k \pi_k(x),$$

其中 $a_k = f[x_0, x_1, \dots, x_k]$ 。



2.1.4 差分

定义 2.7 (向前差分). 一阶向前差分定义为

$$\Delta f_k = f_{k+1} - f_k,$$

其中 $f_k = f(x_k)$ 。 m 阶向前差分递归定义为

$$\Delta^m f_k = \Delta^{m-1} f_{k+1} - \Delta^{m-1} f_k.$$



定义 2.8 (向后差分). 一阶向后差分定义为

$$\nabla f_k = f_k - f_{k-1}.$$

m 阶向后差分递归定义为

$$\nabla^m f_k = \nabla^{m-1} f_k - \nabla^{m-1} f_{k-1}.$$



定义 2.9 (中心差分). 一阶中心差分定义为

$$\delta f_{k+\frac{1}{2}} = f_{k+1} - f_k \quad \text{或} \quad \delta f_k = f_{k+\frac{1}{2}} - f_{k-\frac{1}{2}}.$$

m 阶中心差分递归定义为

$$\delta^m f_k = \delta^{m-1} f_{k+\frac{1}{2}} - \delta^{m-1} f_{k-\frac{1}{2}}.$$



性质 2.2 (差分与差商的关系). 对于等距节点 $x_k = x_0 + kh$, 差商与差分有如下关系:

$$\begin{aligned} f[x_k] &= f_k, \\ f[x_k, x_{k+1}] &= \frac{\Delta f_k}{h}, \\ f[x_k, x_{k+1}, x_{k+2}] &= \frac{\Delta^2 f_k}{2!h^2}, \\ &\vdots \\ f[x_k, x_{k+1}, \dots, x_{k+m}] &= \frac{\Delta^m f_k}{m!h^m}. \end{aligned}$$

类似地, 向后差商与向后差分的关系为

$$f[x_k, x_{k-1}, \dots, x_{k-m}] = \frac{\nabla^m f_k}{m!h^m}.$$



定理 2.5 (Newton 向前插值公式). 令 $x = x_0 + th$, $t \in \mathbb{R}$. 则 Newton 向前插值公式为

$$N_n(x) = f_0 + \frac{t}{1!}\Delta f_0 + \frac{t(t-1)}{2!}\Delta^2 f_0 + \dots + \frac{t(t-1)\cdots(t-n+1)}{n!}\Delta^n f_0. \quad (15)$$

该公式适用于在 x_0 附近插值, 且 t 较小的情况。



定理 2.6 (Newton 向后插值公式). 令 $x = x_n + th$, 则 Newton 向后插值公式为

$$N_n(x) = f_n + \frac{t}{1!}\nabla f_n + \frac{t(t+1)}{2!}\nabla^2 f_n + \dots + \frac{t(t+1)\cdots(t+n-1)}{n!}\nabla^n f_n. \quad (16)$$

该公式适用于在 x_n 附近插值, 且 t 为负的小量。



最后, 我们有下面两个性质:

定理 2.7 (差分性质). 设 $p(x)$ 是一个 n 次多项式 ($n \geq 0$), 节点为等距节点 $x_k = x_0 + kh$ ($h \neq 0$). 则:

1. $p(x)$ 的 n 阶向前 (或向后) 差分 $\Delta^n p_k$ 为常数;
2. $p(x)$ 的 $n+1$ 阶及更高阶差分恒为零, 即

$$\Delta^{n+1} p_k = 0, \quad \nabla^{n+1} p_k = 0, \quad \delta^{n+1} p_k = 0$$

对所有 k 成立。



定理 2.8 (差商性质). 设 $p(x)$ 是一个 n 次多项式, x_0, x_1, \dots, x_m 为任意互异节点. 则:

1. $p(x)$ 的 n 阶差商 $p[x_0, x_1, \dots, x_n]$ 为常数 (等于 n 次项系数 a_n 乘以 $n!$ 的适当组合);
2. $p(x)$ 的 $n+1$ 阶及更高阶差商恒为零, 即

$$p[x_0, x_1, \dots, x_{n+1}] = 0, \quad p[x_0, x_1, \dots, x_{n+2}] = 0, \quad \dots$$



2.2 Hermite 插值

定义 2.10 (Hermite 插值问题). 已知函数 $f(x)$ 在 $n+1$ 个互异节点 x_0, x_1, \dots, x_n 上的函数值 $f(x_i)$ 及导数值 $f'(x_i)$ ($i = 0, 1, \dots, n$), 寻求一个次数不超过 $2n+1$ 的多项式 $H_{2n+1}(x)$, 使得

$$H_{2n+1}(x_i) = f(x_i), \quad (17)$$

$$H'_{2n+1}(x_i) = f'(x_i), \quad i = 0, 1, \dots, n. \quad (18)$$

称 $H_{2n+1}(x)$ 为 $f(x)$ 的 *Hermite 插值多项式*, 条件 (17) 与 (18) 称为 *Hermite 插值条件*. 

2.3 Hermite 插值公式的构造

设 $l_i(x)$ 为关于节点 x_0, x_1, \dots, x_n 的 Lagrange 基函数, 即

$$l_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j}, \quad i = 0, 1, \dots, n.$$

定义辅助函数

$$\alpha_i(x) = [1 - 2l'_i(x_i)(x - x_i)]l_i^2(x), \quad \beta_i(x) = (x - x_i)l_i^2(x).$$

定理 2.9 (Hermite 插值公式). 满足 *Hermite 插值条件* 的 $2n+1$ 次多项式可唯一表示为

$$H_{2n+1}(x) = \sum_{i=0}^n f(x_i)\alpha_i(x) + \sum_{i=0}^n f'(x_i)\beta_i(x). \quad (19)$$

例 2.1 (两点三次 Hermite 插值). 设节点为 x_0 和 x_1 , 已知 $f(x_0), f(x_1), f'(x_0), f'(x_1)$. 相应的 Lagrange 基函数为

$$l_0(x) = \frac{x - x_1}{x_0 - x_1}, \quad l_1(x) = \frac{x - x_0}{x_1 - x_0}.$$

计算导数:

$$l'_0(x) = \frac{1}{x_0 - x_1}, \quad l'_1(x) = \frac{1}{x_1 - x_0}.$$

于是

$$\alpha_0(x) = \left[1 - \frac{2}{x_0 - x_1}(x - x_0)\right] \left(\frac{x - x_1}{x_0 - x_1}\right)^2, \quad \beta_0(x) = (x - x_0) \left(\frac{x - x_1}{x_0 - x_1}\right)^2,$$

$$\alpha_1(x) = \left[1 - \frac{2}{x_1 - x_0}(x - x_1)\right] \left(\frac{x - x_0}{x_1 - x_0}\right)^2, \quad \beta_1(x) = (x - x_1) \left(\frac{x - x_0}{x_1 - x_0}\right)^2.$$

因此, 三次 Hermite 插值多项式为

$$H_3(x) = f(x_0)\alpha_0(x) + f(x_1)\alpha_1(x) + f'(x_0)\beta_0(x) + f'(x_1)\beta_1(x). \quad (20)$$

2.4 误差分析

定理 2.10 (Lagrange 插值余项). 若 $f \in C^n[a, b]$, $f^{(n+1)}(x)$ 在 (a, b) 上存在, 则对任意 $x \in [a, b]$, 关于节点 x_0, x_1, \dots, x_n 的 Lagrange 插值多项式 $L_n(x)$ 的误差为

$$R_n(x) = f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i), \quad (21)$$

其中 $\xi \in (a, b)$, 且依赖于 x .

证明. 当 $x = x_i$ ($i = 0, 1, \dots, n$) 时, (21) 两边均为零, 结论成立. 现设 $x \neq x_i$ 对所有 i 成立. 对固定的 $x \in [a, b]$, 定义辅助函数

$$F(t) = f(t) - L_n(t) - K \prod_{i=0}^n (t - x_i),$$

其中 K 为待定常数. 选择 K 使得 $F(x) = 0$, 即

$$K = \frac{f(x) - L_n(x)}{\prod_{i=0}^n (x - x_i)}.$$

于是 $F(t)$ 在点 x, x_0, x_1, \dots, x_n 这 $n+2$ 个点上取零值. 由 Rolle 定理, $F'(t)$ 在 (a, b) 内至少存在 $n+1$ 个零点. 反复应用 Rolle 定理, 最终存在 $\xi \in (a, b)$ 使得 $F^{(n+1)}(\xi) = 0$.

计算 $F^{(n+1)}(t)$:

$$F^{(n+1)}(t) = f^{(n+1)}(t) - L_n^{(n+1)}(t) - K \cdot (n+1)!$$

由于 $L_n(t)$ 是 n 次多项式, $L_n^{(n+1)}(t) \equiv 0$. 代入 $t = \xi$ 得

$$0 = f^{(n+1)}(\xi) - K \cdot (n+1)!$$

解出 K 并代回 K 的表达式, 即得

$$f(x) - L_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \prod_{i=0}^n (x - x_i).$$

□

定理 2.11 (两点三次 Hermite 插值余项). 若 $f \in C^3[a, b]$, $f^{(4)}(x)$ 在 (a, b) 上存在, 则对任意 $x \in [a, b]$, 两点 Hermite 插值多项式 $H_3(x)$ 的误差为

$$R_3(x) = f(x) - H_3(x) = \frac{f^{(4)}(\xi)}{4!} (x - x_0)^2 (x - x_1)^2, \quad (22)$$

其中 $\xi \in (a, b)$, 且依赖于 x .

证明. 当 $x = x_0$ 或 $x = x_1$ 时, (22) 两边均为零, 结论成立. 现设 $x \neq x_0, x_1$. 对固定的 $x \in [a, b]$, 定义辅助函数

$$F(t) = f(t) - H_3(t) - K(t - x_0)^2 (t - x_1)^2,$$

其中 K 为待定常数. 选择 K 使得 $F(x) = 0$, 即

$$K = \frac{f(x) - H_3(x)}{(x - x_0)^2 (x - x_1)^2}.$$

由于 $H_3(t)$ 满足 Hermite 插值条件

$$H_3(x_i) = f(x_i), \quad H_3'(x_i) = f'(x_i), \quad i = 0, 1,$$

易知 $F(t)$ 在 x_0 和 x_1 处均有二重零点 (因 H_3 与 f 的函数值、一阶导数值均相同, 且 $(t - x_i)^2$ 也给出二重零点)。此外 $F(x) = 0$ 。故 $F(t)$ 在 $[a, b]$ 上至少有 4 个零点 (计重数)。

反复应用 Rolle 定理, 可知存在 $\xi \in (a, b)$ 使得 $F^{(4)}(\xi) = 0$ 。

计算 $F^{(4)}(t)$:

$$F^{(4)}(t) = f^{(4)}(t) - H_3^{(4)}(t) - K \cdot 4!.$$

由于 $H_3(t)$ 是三次多项式, $H_3^{(4)}(t) \equiv 0$ 。代入 $t = \xi$ 得

$$0 = f^{(4)}(\xi) - K \cdot 4!.$$

解出 K 并代回 K 的表达式, 即得

$$f(x) - H_3(x) = \frac{f^{(4)}(\xi)}{4!} (x - x_0)^2 (x - x_1)^2.$$

□

2.5 分段低次多项式插值

设区间 $[a, b]$ 有剖分 $\Delta: a = x_0 < x_1 < \cdots < x_n = b$, 记单元 $e_i = [x_{i-1}, x_i]$ 。定义关于剖分 Δ 的分段多项式插值函数空间为

$$Sp(k; l; \Delta) = \{s \in C^l[a, b] : s|_{e_i} \in P_k, i = 1, 2, \dots, n\},$$

它表示在每个剖分单元上为 k 次多项式, 且在整個区间 $[a, b]$ 上 l 阶连续可微的函数全体。称 $Sp(k; l; \Delta)$ 为关于剖分 Δ 、重数为 $k - l$ 的 k 次多项式样条空间。

更一般地, 引入重指标向量 $M = (m_1, \dots, m_{n-1})$, 定义多项式样条空间

$$Sp(k; M; \Delta) = \{s : s|_{e_i} \in P_k, s \in C^{k-m_i}(x_i), i = 1, \dots, n-1\}.$$

该空间中的函数在内部分节点 x_i 处要求满足 $k - m_i$ 阶连续可微。特别地, 当 $m_1 = \cdots = m_{n-1} = k - l$ 时, $Sp(k; M; \Delta) = Sp(k; l; \Delta)$ 。当 $l = k - 1$ 时, 简记 $Sp(k; l; \Delta)$ 为 $Sp(k; \Delta)$, 称为 (完全) 样条函数空间。

定理 2.12 (样条空间维数)。样条空间 $Sp(k; M; \Delta)$ 的维数可由以下分析得到:

1. 若不考虑段间连接条件, 分段 k 次多项式函数的维数为 $(k + 1)n$ 。
2. 在每个内节点 x_i 处要求 $s \in C^{k-m_i}$, 这给出 $k - m_i + 1$ 个约束条件, 总约束数为 $\sum_{i=1}^{n-1} (k - m_i + 1)$ 。

因此维数公式为

$$\dim Sp(k; M; \Delta) = (k + 1)n - \sum_{i=1}^{n-1} (k - m_i + 1). \quad (23)$$



2.5.1 分段一次多项式插值

取插值函数空间为 $Sp(1; \Delta) := Sp(1; 0; \Delta)$, 即分段线性连续函数空间。由 (23) 知该空间维数为 $2n - (n - 1) = n + 1$ 。

分段一次插值问题: 求 $s \in Sp(1; \Delta)$, 满足

$$s(x_i) = f_i, \quad i = 0, 1, \dots, n.$$

解存在唯一, 且可局部表示为

$$s(x) = \frac{x - x_i}{x_{i-1} - x_i} f_{i-1} + \frac{x - x_{i-1}}{x_i - x_{i-1}} f_i, \quad x \in e_i, \quad i = 1, \dots, n.$$

误差估计: 若 $f \in C^2[a, b]$, 则当 $x \in e_i$ 时,

$$f(x) - s(x) = \frac{f''(\xi_i)}{2!} (x - x_{i-1})(x - x_i), \quad \xi_i \in (x_{i-1}, x_i).$$

于是

$$|f(x) - s(x)| \leq \frac{1}{8} h^2 \max_{a \leq x \leq b} |f''(x)|, \quad h = \max_i h_i$$

2.5.2 分段三次 Hermite 多项式插值

取插值函数空间为 $Sp(3; 1; \Delta)$, 由 (23) 知其维数为 $4n - 2(n - 1) = 2n + 2$ 。

分段三次 Hermite 插值问题: 求 $H \in Sp(3; 1; \Delta)$, 满足

$$H(x_i) = f_i, \quad H'(x_i) = f'_i, \quad i = 0, 1, \dots, n.$$

解存在唯一, 且可局部确定。在每个单元 e_i 上, 利用两点 Hermite 插值公式:

$$H(x) = f_{i-1} h_{i-1,0}(x) + f_i h_{i,0}(x) + f'_{i-1} h_{i-1,1}(x) + f'_i h_{i,1}(x), \quad x \in e_i,$$

其中

$$\begin{aligned} h_{i-1,0}(x) &= \phi_0\left(\frac{x - x_{i-1}}{h_i}\right), & h_{i,0}(x) &= \phi_0\left(\frac{x_i - x}{h_i}\right), \\ h_{i-1,1}(x) &= h_i \phi_1\left(\frac{x - x_{i-1}}{h_i}\right), & h_{i,1}(x) &= -h_i \phi_1\left(\frac{x_i - x}{h_i}\right), \end{aligned}$$

而 $\phi_0(t) = (1 + 2t)(1 - t)^2$, $\phi_1(t) = t(1 - t)^2$, $t \in [0, 1]$ 。

误差估计: 若 $f \in C^4[a, b]$, 则当 $x \in e_i$ 时,

$$f(x) - H(x) = \frac{f^{(4)}(\xi_i)}{4!} (x - x_{i-1})^2 (x - x_i)^2, \quad \xi_i \in (x_{i-1}, x_i).$$

因此

$$|f(x) - H(x)| \leq \frac{h^4}{384} \max_{a \leq x \leq b} |f^{(4)}(x)|, \quad h = \max_i h_i$$

2.5.3 三次样条插值

取插值函数空间为 $Sp(3; \Delta)$, 即分段三次多项式且在整个区间 $[a, b]$ 上二阶连续可微的函数全体:

$$Sp(3; \Delta) = \{s \in C^2[a, b] : s|_{e_i} \in P_3, \quad i = 1, 2, \dots, n\}.$$

由样条空间维数公式可得

$$\dim Sp(3; \Delta) = 4n - 3(n - 1) = n + 3.$$

定义三次插值样条函数 $s \in Sp(3; \Delta)$ ，首先要求在内节点满足插值条件：

$$s(x_i) = f_i, \quad i = 1, 2, \dots, n - 1. \quad (24)$$

这些条件共有 $n - 1$ 个。为唯一确定 $s(x)$ ，还需补充 4 个条件，通常给出以下三种边界条件提法：

1. **I 型边界条件**（给定端点函数值与一阶导数值）：

$$s(x_0) = f_0, \quad s(x_n) = f_n; \quad s'(x_0) = f'_0, \quad s'(x_n) = f'_n.$$

2. **II 型边界条件**（给定端点函数值与二阶导数值）：

$$s(x_0) = f_0, \quad s(x_n) = f_n; \quad s''(x_0) = f''_0, \quad s''(x_n) = f''_n.$$

3. **III 型边界条件**（周期边界条件）：

$$s(x_0) = s(x_n) = f_0; \quad s'(x_0) = s'(x_n); \quad s''(x_0) = s''(x_n).$$

在内节点插值条件 (24) 与任一上述边界条件共同构成 $n + 3$ 个条件，从而可唯一确定三次样条插值函数 $s(x)$ 。

2.6 数值积分、数值微分

定义 2.11 (代数精度). 设数值积分公式

$$\int_a^b f(x) dx \approx \sum_{k=0}^n A_k f(x_k)$$

如果对所有次数 $\leq m$ 的多项式精确成立，但对 $m + 1$ 次多项式不精确，则称该公式具有 m 次代数精度。

例如，梯形公式

$$\int_a^b f(x) dx \approx \frac{b-a}{2} [f(a) + f(b)]$$

具有 1 次代数精度；Simpson 公式

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]$$

具有 3 次代数精度。

为使公式(2.11)具有 m 次代数精度，只需令它对 $f(x) = 1, x, \dots, x^m$ 均精确成立，即要求系数 A_k 与节点 x_k 满足：

$$\begin{cases} \sum_{k=0}^n A_k = b - a, \\ \sum_{k=0}^n A_k x_k = \frac{1}{2}(b^2 - a^2), \\ \vdots \\ \sum_{k=0}^n A_k x_k^m = \frac{1}{m+1}(b^{m+1} - a^{m+1}). \end{cases} \quad (25)$$

2.6.1 复化求积公式

定理 2.13 (复化梯形公式). 将区间 $[a, b]$ n 等分, 步长 $h = \frac{b-a}{n}$, 节点 $x_k = a + kh$. 在每个子区间上使用梯形公式, 则

$$\int_a^b f(x) dx = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(x) dx \approx \sum_{k=0}^{n-1} \frac{h}{2} [f(x_k) + f(x_{k+1})],$$

即得

$$\int_a^b f(x) dx \approx \frac{h}{2} \left[f(a) + 2 \sum_{k=1}^{n-1} f(x_k) + f(b) \right] := T_n. \quad (26)$$

定理 2.14 (复化 Simpson 公式). 在每个子区间 $[x_k, x_{k+1}]$ 上使用 *Simpson* 公式, 记中点 $x_{k+1/2} = x_k + h/2$, 则

$$\int_a^b f(x) dx \approx \frac{h}{6} \left[f(a) + 4 \sum_{k=0}^{n-1} f(x_{k+1/2}) + 2 \sum_{k=1}^{n-1} f(x_k) + f(b) \right] := S_n. \quad (27)$$

定理 2.15 (复化梯形公式误差). 若 $f \in C^2[a, b]$, 则复化梯形公式(26)的误差满足

$$\left| \int_a^b f(x) dx - T_n \right| \leq \frac{b-a}{12} \cdot h^2 \max_{a \leq x \leq b} |f''(x)|. \quad (28)$$

定理 2.16 (复化 Simpson 公式误差). 若 $f \in C^4[a, b]$, 则复化 *Simpson* 公式(27)的误差满足

$$\left| \int_a^b f(x) dx - S_n \right| \leq \frac{b-a}{2880} \cdot h^4 \max_{a \leq x \leq b} |f^{(4)}(x)|. \quad (29)$$

复化 Simpson 公式常写为等价形式:

$$\int_a^b f(x) dx \approx \frac{h}{6} \left\{ f(a) - f(b) + \sum_{k=1}^n [4f(x_{k-1/2}) + 2f(x_k)] \right\}.$$

3 最佳逼近

3.1 最佳逼近

定理 3.1 (Weierstrass 定理). 设函数 $f(x)$ 在区间 $[a, b]$ 上连续, 则对任意 $\varepsilon > 0$, 总存在代数多项式 $p(x)$, 使得

$$\max_{a \leq x \leq b} |f(x) - p(x)| < \varepsilon. \quad (3.1)$$

该定理保证了连续函数可以用多项式任意精度一致逼近。由此启发我们寻求在给定次数 n 下, 使逼近误差最小的多项式, 即最佳一致逼近多项式。

定义 3.1 (最佳一致逼近). 设 $f(x)$ 是 $[a, b]$ 上的连续函数, 记 P_n 为次数不超过 n 的多项式全体. 定义

$$E_n(f) = \inf_{p_n \in P_n} \max_{a \leq x \leq b} |f(x) - p_n(x)| \quad (3.2)$$


为 $f(x)$ 在 P_n 中的**最佳一致逼近** (或称 *Chebyshev* 逼近). 若存在 $p_n^* \in P_n$ 使得

$$E_n(f) = \max_{a \leq x \leq b} |f(x) - p_n^*(x)|,$$

则称 p_n^* 为 $f(x)$ 的 n 次**最佳一致逼近多项式**. 


定义 3.2 (交错点组). 设 $f \in C[a, b]$, 点集 $\{x_i\}_{i=0}^k$ 满足 $a \leq x_0 < x_1 < \dots < x_k \leq b$. 若存在 $\sigma = 1$ 或 -1 , 使得

$$f(x_i) = (-1)^i \sigma \|f\|, \quad i = 0, 1, \dots, k,$$

其中 $\|f\| = \max_{a \leq x \leq b} |f(x)|$, 则称 $\{x_i\}_{i=0}^k$ 为 $f(x)$ 在 $[a, b]$ 上的**交错点组**, x_i 称为**交错点** (简称 (e) 点). 

定理 3.2 (Vallee-Poussin 定理). 设 $f \in C[a, b]$, 若存在多项式 $p \in P_n$, 使得 $f(x) - p(x)$ 在 $[a, b]$ 上至少有 $n+2$ 个点 x_0, x_1, \dots, x_{n+1} 处取值正负相间, 则

$$E_n(f) \geq \lambda = \min_{0 \leq i \leq n+1} |f(x_i) - p(x_i)|. \quad \text{img alt="robot icon" data-bbox="854 462 874 477"/>$$

定理 3.3 (Chebyshev 定理). 设 $f \in C[a, b]$ 且 $f \notin P_n$, 则 $p \in P_n$ 是 f 的 n 次最佳一致逼近多项式的充要条件是: $f - p$ 在 $[a, b]$ 上存在一个至少包含 $n+2$ 个点的交错点组. 

3.2 Chebyshev 多项式与最小零偏差问题

考虑在区间 $[-1, 1]$ 上, 子空间 P_{n-1} 对函数 x^n 的最佳一致逼近问题: 求 $p_{n-1}^* \in P_{n-1}$, 使得

$$E_{n-1}(x^n) = \min_{p_{n-1} \in P_{n-1}} \|x^n - p_{n-1}\| = \|x^n - p_{n-1}^*\|,$$

其中范数 $\|f\| = \max_{-1 \leq x \leq 1} |f(x)|$.

记 P_n^1 为首项系数为 1 的 n 次代数多项式全体. 上述问题等价于: 求 $p_n^* \in P_n^1$, 使得

$$\|p_n^* - 0\| = \min_{p_n \in P_n^1} \|p_n - 0\|.$$

此即与**零偏差最小问题**.

定义 3.3 (Chebyshev 多项式). 作变换 $x = \cos \theta$, 它为 $[0, \pi] \rightarrow [-1, 1]$ 的一一映射. 定义 n 次 *Chebyshev* 多项式为

$$T_n(x) = \cos n\theta = \cos(n \arccos x), \quad n = 0, 1, \dots \quad (3.9) \quad \text{img alt="robot icon" data-bbox="854 821 874 836"/>$$

性质 3.1 (Chebyshev 多项式的递推关系). 利用三角恒等式

$$\cos n\theta + \cos(n-2)\theta = 2 \cos \theta \cos(n-1)\theta, \quad \text{img alt="robot icon" data-bbox="854 896 874 911"/>$$

可得递推公式:

$$\begin{cases} T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), & n = 2, 3, \dots, \\ T_0(x) = 1, \quad T_1(x) = x. \end{cases}$$

由此可知, $T_n(x)$ 是首项系数为 2^{n-1} 的 n 次代数多项式, 且 $T_{2k}(x)$ 只含 x 的偶次幂, $T_{2k-1}(x)$ 只含 x 的奇次幂。

定理 3.4. 在首项系数为 1 的所有 n 次多项式中, 多项式

$$p_n^*(x) = 2^{1-n}T_n(x)$$

对零的偏差最小, 即

$$\min_{p_n \in P_n^1} \|p_n\| = \|p_n^*\| = 2^{1-n}. \quad (30)$$

推论 3.5. 设 $p_n(x)$ 是首项系数为 1 的 n 次多项式, 则

$$\|p_n\| = \max_{|x| \leq 1} |p_n(x)| \geq 2^{1-n}.$$

性质 3.2.

1. **正交性:** 序列 $\{T_k(x)\}_{k=0}^n$ 在 $[-1, 1]$ 上关于权函数 $\frac{1}{\sqrt{1-x^2}}$ 正交, 即

$$\int_{-1}^1 \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}} dx = \begin{cases} 0, & m \neq n, \\ \pi/2, & m = n \neq 0, \\ \pi, & m = n = 0. \end{cases}$$

2. **零点:** $T_n(x)$ 在 $[-1, 1]$ 上恰有 n 个不同的实根

$$x_k = \cos\left(\frac{(2k-1)\pi}{2n}\right), \quad k = 1, 2, \dots, n.$$

3. **极值点:** $T_n(x)$ 在 $[-1, 1]$ 上取得 $n+1$ 个极值点 $x_j = \cos\left(\frac{j\pi}{n}\right)$, 且 $|T_n(x_j)| = 1$ 。

例 3.1. 由第 2 章可知, 函数 $f(x)$ 在区间 $[-1, 1]$ 上关于互异节点 x_0, x_1, \dots, x_n 的 n 次插值多项式 $p_n(x)$ 有余项估计

$$|f(x) - p_n(x)| \leq \frac{\|f^{(n+1)}\|}{(n+1)!} \max_{|x| \leq 1} |(x-x_0)(x-x_1)\cdots(x-x_n)|,$$

其中 $\|f^{(n+1)}\| = \max_{|x| \leq 1} |f^{(n+1)}(x)|$ 。

所谓**代数插值多项式余项的极小化问题**是指: 如何选取插值节点 x_0, x_1, \dots, x_n , 使得

$$\max_{|x| \leq 1} |(x-x_0)(x-x_1)\cdots(x-x_n)|$$

达到最小。换句话说, 是在首项系数为 1 的 $n+1$ 次多项式中, 寻找与零偏差最小的多项式。

由 Chebyshev 最佳逼近定理(30)可知, 当插值节点取为 $n+1$ 次 Chebyshev 多项式 $T_{n+1}(x)$ 的零点时, 上述最大值达到最小。 $T_{n+1}(x)$ 的零点为

$$x_k = \cos \frac{(2k-1)\pi}{2(n+1)}, \quad k = 1, 2, \dots, n+1. \quad (3.10)$$

此时有

$$\max_{|x| \leq 1} \left| \prod_{k=0}^n (x - x_k) \right| = \min_{q_{n+1} \in P_{n+1}^1} \max_{|x| \leq 1} |q_{n+1}(x)| = 2^{-n}.$$

3.3 最佳平方逼近与正交多项式

3.3.1 连续情形的最佳平方逼近

考虑带权内积空间 $L_\rho^2[a, b]$, 内积与范数定义为

$$(f, g) = \int_a^b \rho(x) f(x) g(x) dx, \quad \|f\| = \sqrt{(f, f)}, \quad \forall f, g \in L_\rho^2[a, b].$$

当 $\rho(x) \equiv 1$ 时简记为 $L^2[a, b]$ 。

对任意 $f \in L_\rho^2[a, b]$, 在 $n+1$ 维多项式子空间 P_n 中存在唯一的最佳平方逼近多项式

$$p_n^*(x) = \sum_{j=0}^n a_j^* x^j,$$

其系数 $\{a_i^*\}_{i=0}^n$ 满足法方程组

$$\sum_{j=0}^n (x^i, x^j) a_j^* = (f, x^i), \quad i = 0, 1, \dots, n, \quad (31)$$

其中

$$(x^i, x^j) = \int_a^b \rho(x) x^{i+j} dx, \quad (f, x^i) = \int_a^b \rho(x) f(x) x^i dx.$$

例 3.2 (一次最佳平方逼近). 设 $f(x) = \sqrt{x} \in L^2[1/4, 1]$, 求其一次最佳平方逼近多项式 $p_1^*(x) = a_0^* + a_1^* x$ 。

计算内积:

$$\begin{aligned} (1, 1) &= \frac{3}{4}, & (1, x) &= (x, 1) = \frac{15}{32}, \\ (x, x) &= \frac{21}{64}, & (f, 1) &= \frac{7}{12}, & (f, x) &= \frac{31}{80}. \end{aligned}$$

代入法方程组(31)得

$$\begin{cases} \frac{3}{4} a_0^* + \frac{15}{32} a_1^* = \frac{7}{12}, \\ \frac{15}{32} a_0^* + \frac{21}{64} a_1^* = \frac{31}{80}. \end{cases}$$

解得 $a_0^* = \frac{10}{27}$, $a_1^* = \frac{88}{135}$, 故

$$p_1^*(x) = \frac{10}{27} + \frac{88}{135} x.$$

3.3.2 常用带权正交多项式

1. Legendre 多项式

取 $\rho(x) \equiv 1$, 区间 $[-1, 1]$ 上的正规正交多项式称为 Legendre 多项式, 表达式为

$$l_k(x) = \frac{1}{2^k k!} \sqrt{\frac{2k+1}{2}} \cdot \frac{d^k}{dx^k} [(x^2-1)^k], \quad k = 0, 1, \dots, n.$$

递推公式:

$$l_{k+1}(x) = \frac{\sqrt{(2k+1)(2k+3)}}{k+1} x l_k(x) - \frac{k}{k+1} \sqrt{\frac{2k+3}{2k-1}} l_{k-1}(x), \quad k \geq 1.$$

$l_k(x)$ 的奇偶性: k 偶为偶函数, k 奇为奇函数。

2. Chebyshev 多项式 (第一类)

由 $T_j(x) = \cos(j \arccos x)$ 定义, 在 $[-1, 1]$ 上带权 $\rho(x) = (1-x^2)^{-1/2}$ 正交。

3. 第二类 Chebyshev 多项式

$$U_n(x) = \frac{\sin((1+n) \arccos x)}{\sqrt{1-x^2}}, \quad n = 0, 1, \dots$$

在 $[-1, 1]$ 上带权 $(1-x^2)^{1/2}$ 正交。

4. Laguerre 多项式

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x}), \quad n = 0, 1, \dots$$

在 $[0, +\infty)$ 上带权 e^{-x} 正交。

5. Hermite 多项式

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} (e^{-x^2}), \quad n = 0, 1, \dots$$

在 $(-\infty, +\infty)$ 上带权 e^{-x^2} 正交。

3.3.3 Chebyshev 最佳平方逼近

利用带权 $\rho(x) = (1-x^2)^{-1/2}$ 的 Chebyshev 正交多项式, $f \in L^2_\rho[-1, 1]$ 的 n 次最佳平方逼近为

$$S_n(x) = a_0 \frac{T_0(x)}{2} + \sum_{j=1}^n a_j T_j(x), \quad (32)$$

其中

$$a_j = \frac{(f, T_j)}{(T_j, T_j)} = \frac{2}{\pi} \int_{-1}^1 \frac{f(x) T_j(x)}{\sqrt{1-x^2}} dx, \quad j = 0, 1, \dots, n.$$

$S_n(x)$ 称为 $f(x)$ 的 Chebyshev 级数部分和。若 $f \in C^1[-1, 1]$, 则 $S_n(x)$ 一致收敛于 $f(x)$, 且系数 a_k 随 k 增大迅速衰减。

例 3.3 (e^x 的三次 Chebyshev 最佳平方逼近). 求 $S_3(x)$ 。由 (32),

$$S_3(x) = \frac{a_0}{2} + a_1T_1(x) + a_2T_2(x) + a_3T_3(x),$$

其中

$$a_j = \frac{2}{\pi} \int_0^\pi e^{\cos \theta} \cos(j\theta) d\theta.$$

数值计算得

$$a_0 \approx 2.5321318, a_1 \approx 1.1303182, a_2 \approx 0.2714953, a_3 \approx 0.0443369.$$

将 $T_j(x)$ 用幂函数表示后整理得

$$S_3(x) \approx 0.994571 + 0.997308x + 0.542991x^2 + 0.177347x^3.$$



3.4 Gauss 型数值积分

考虑带权积分

$$I = \int_a^b \rho(x)f(x) dx,$$

其中权函数 $\rho(x) \geq 0$ 。特别地, $\rho(x) \equiv 1$ 时为普通积分。讨论如下形式的数值积分公式

$$\int_a^b \rho(x)f(x) dx \approx \sum_{k=1}^n A_k f(x_k), \quad (33)$$

其中 $x_i \neq x_j$ ($i \neq j$)。可以证明, 公式 (33) 至少具有 $n-1$ 次代数精度的充要条件是它为插值型, 即

$$A_k = \int_a^b \rho(x)l_k(x) dx, \quad l_k(x) = \prod_{\substack{j=1 \\ j \neq k}}^n \frac{x - x_j}{x_k - x_j}. \quad (34)$$

3.5 Gauss 型求积公式的定义

定义 3.4 (Gauss 型求积公式). 若形如 (33) 的求积公式具有 $2n-1$ 次代数精度, 则称它为 n 点 **Gauss 型求积公式**, 节点 $\{x_k\}$ 称为 **Gauss 点**。



为求 $\{x_k\}$ 与 $\{A_k\}$, 令 (33) 对 $f(x) = x^i$ ($i = 0, 1, \dots, 2n-1$) 精确成立, 得方程组

$$\sum_{k=1}^n A_k x_k^i = \int_a^b \rho(x)x^i dx, \quad i = 0, 1, \dots, 2n-1. \quad (35)$$

例 3.4 (一点 Gauss 公式). 当 $n=1$ 时, (35) 化为

$$A_1 = \int_a^b \rho(x) dx, \quad A_1 x_1 = \int_a^b x \rho(x) dx.$$

解得

$$A_1 = \int_a^b \rho(x) dx, \quad x_1 = \frac{\int_a^b x \rho(x) dx}{\int_a^b \rho(x) dx}.$$



特别地, 当 $\rho(x) \equiv 1$ 时, 有

$$\int_a^b f(x) dx \approx (b-a)f\left(\frac{a+b}{2}\right),$$

即中矩形公式, 它具有 1 次代数精度。



定理 3.6 (Gauss 点的充要条件). 互异节点 $x_k (k = 1, \dots, n)$ 是 n 点 Gauss 求积公式的 Gauss 点的充要条件是多项式

$$\omega_n(x) = \prod_{k=1}^n (x - x_k)$$

与任何次数 $\leq n-1$ 的多项式关于权 $\rho(x)$ 正交, 即

$$\int_a^b \rho(x)\omega_n(x)x^j dx = 0, \quad j = 0, 1, \dots, n-1. \quad (36)$$



证明. **必要性:** 因为 Gauss 公式具有 $2n-1$ 次代数精度, 故对 $f(x) = \omega_n(x)x^j (j = 0, \dots, n-1)$ 精确成立:

$$\int_a^b \rho(x)\omega_n(x)x^j dx = \sum_{k=1}^n A_k \omega_n(x_k)x_k^j.$$

由 $\omega_n(x_k) = 0$ 即得 (36)。

充分性: 以 x_k 为节点构造插值型公式 (33)。对任意次数 $\leq 2n-1$ 的多项式 $f(x)$, 用 $\omega_n(x)$ 除 $f(x)$ 得

$$f(x) = \omega_n(x)p(x) + q(x), \quad \deg p \leq n-1, \quad \deg q \leq n-1.$$

于是 $f(x_k) = q(x_k)$ 。计算积分:

$$\int_a^b \rho(x)f(x) dx = \int_a^b \rho(x)\omega_n(x)p(x) dx + \int_a^b \rho(x)q(x) dx.$$

由正交性 (36), 第一项为零; 第二项因插值型公式至少具有 $n-1$ 次精度, 故

$$\int_a^b \rho(x)q(x) dx = \sum_{k=1}^n A_k q(x_k) = \sum_{k=1}^n A_k f(x_k).$$

因此公式对 $f(x)$ 精确成立, 即具有 $2n-1$ 次代数精度, 从而 x_k 为 Gauss 点。□

3.5.1 正交多项式与 Gauss 点的关系

关于权函数 $\rho(x)$ 的正交多项式系 $\{\varphi_m(x)\}$ 满足:

1. $\deg \varphi_m = m$, 且在 (a, b) 内有 m 个互异零点;
2. $\int_a^b \rho(x)\varphi_m(x)q(x) dx = 0$ 对任意 $\deg q \leq m-1$ 成立;
3. 若 $\{g_m(x)\}$ 也是正交多项式系, 则 $g_m(x) = c_m \varphi_m(x)$ (c_m 为常数)。

设 x_1, \dots, x_n 是 $\varphi_n(x)$ 的零点, $\omega_n(x) = \prod_{k=1}^n (x - x_k)$, 则 $\varphi_n(x) = a_n \omega_n(x)$ (a_n 为首项系数)。于是对任意 $\deg q \leq n-1$,

$$\int_a^b \rho(x)\omega_n(x)q(x) dx = a_n^{-1} \int_a^b \rho(x)\varphi_n(x)q(x) dx = 0.$$

由定理 (36) 知 x_1, \dots, x_n 为 n 点 Gauss 公式的 Gauss 点。

反之, 若 x_1, \dots, x_n 是 Gauss 点, 则 (36) 成立, 从而对正交基 $\{\varphi_k\}_{k=0}^n$ 有

$$(\omega_n, \varphi_k) = \int_a^b \rho(x) \omega_n(x) \varphi_k(x) dx = 0, \quad 0 \leq k \leq n-1.$$

将 $\omega_n(x)$ 按正交基展开:

$$\omega_n(x) = \sum_{j=0}^n \alpha_j \varphi_j(x),$$

则

$$(\omega_n, \varphi_k) = \sum_{j=0}^n \alpha_j (\varphi_j, \varphi_k) = \alpha_k (\varphi_k, \varphi_k).$$

于是 $\alpha_k = 0$ ($0 \leq k \leq n-1$), 故 $\omega_n(x) = \alpha_n \varphi_n(x)$, 即 x_1, \dots, x_n 是 $\varphi_n(x)$ 的零点。

由此得到:

定理 3.7 (Gauss 公式的存在性与构造). 对任意正整数 n , n 点 Gauss 求积公式存在, 其 Gauss 点就是区间 $[a, b]$ 上关于权函数 $\rho(x)$ 的正交多项式 $\varphi_n(x)$ 的零点。



 **注.** 关于 Gauss 型求积公式的余项和稳定性, 这个我感觉不会考, 就没写进来了。

3.5.2 特殊 Gauss 型求积公式

对一般区间 $[a, b]$ 的积分, 可通过线性变换 $x = \frac{a+b}{2} + \frac{b-a}{2}t$ 化为 $[-1, 1]$ 上的积分:

$$\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{a+b}{2} + \frac{b-a}{2}t\right) dt.$$

因此下面只讨论 $[-1, 1]$ 上的 Gauss 求积公式。

Gauss-Legendre 求积公式

Legendre 多项式

$$L_n(x) = \frac{1}{2^n \cdot n!} \frac{d^n}{dx^n} [(x^2 - 1)^n], \quad n \geq 1, \quad L_0(x) = 1$$

是 $[-1, 1]$ 上关于权函数 $\rho(x) \equiv 1$ 的正交多项式。 n 点 Gauss-Legendre 求积公式为

$$\int_{-1}^1 f(x) dx \approx \sum_{k=1}^n A_k f(x_k), \quad (37)$$

其中 x_1, \dots, x_n 是 $L_n(x)$ 的 n 个互异零点,

$$A_k = \int_{-1}^1 l_k(x) dx, \quad l_k(x) = \prod_{\substack{i=1 \\ i \neq k}}^n \frac{x - x_i}{x_k - x_i}, \quad k = 1, \dots, n.$$

误差公式为

$$E_n[f] = \frac{2^{2n+1}(n!)^4}{(2n+1)[(2n)!]^3} f^{(2n)}(\xi), \quad -1 < \xi < 1.$$

 下面我们来具体算一下具体的值:

例 3.5 ($n = 3$ 的 Gauss-Legendre 公式计算). 构造三点 Gauss-Legendre 求积公式:

$$\int_{-1}^1 f(x) dx \approx A_1 f(x_1) + A_2 f(x_2) + A_3 f(x_3),$$

其中节点 x_k 为三次 Legendre 多项式 $L_3(x)$ 的零点, 权重 A_k 由相应公式给出。



1. 计算节点 x_k 。三次 Legendre 多项式为

$$L_3(x) = \frac{1}{2}(5x^3 - 3x).$$

令 $L_3(x) = 0$, 得

$$\frac{1}{2}(5x^3 - 3x) = 0 \Rightarrow x(5x^2 - 3) = 0.$$

解得三个实根:

$$x_1 = -\sqrt{\frac{3}{5}}, \quad x_2 = 0, \quad x_3 = \sqrt{\frac{3}{5}}.$$

数值为

$$x_1 \approx -0.7745966692, \quad x_2 = 0, \quad x_3 \approx 0.7745966692.$$

2. 计算权重 A_k 。利用公式

$$A_k = \frac{2}{(1 - x_k^2)[L_3'(x_k)]^2}.$$

先计算导数:

$$L_3'(x) = \frac{1}{2}(15x^2 - 3) = \frac{3}{2}(5x^2 - 1).$$

• 对于 $x_1 = -\sqrt{3/5}$:

$$L_3'(x_1) = \frac{3}{2} \left(5 \cdot \frac{3}{5} - 1 \right) = \frac{3}{2}(3 - 1) = 3.$$

$$A_1 = \frac{2}{\left[1 - \left(-\sqrt{3/5} \right)^2 \right] \cdot 3^2} = \frac{2}{\left(1 - \frac{3}{5} \right) \cdot 9} = \frac{2}{\frac{2}{5} \cdot 9} = \frac{2}{18/5} = \frac{5}{9} \approx 0.5555555556.$$

• 对于 $x_2 = 0$:

$$L_3'(x_2) = \frac{3}{2}(0 - 1) = -\frac{3}{2}.$$

$$A_2 = \frac{2}{(1 - 0) \cdot \left(-\frac{3}{2} \right)^2} = \frac{2}{1 \cdot \frac{9}{4}} = \frac{8}{9} \approx 0.8888888889.$$

• 对于 $x_3 = \sqrt{3/5}$: 由对称性 $A_3 = A_1 = 5/9 \approx 0.5555555556$ 。

验证: 通过积分 Lagrange 基函数 $l_2(x)$ 计算 A_2 :

$$l_2(x) = \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} = \frac{(x + \sqrt{3/5})(x - \sqrt{3/5})}{(0 + \sqrt{3/5})(0 - \sqrt{3/5})} = \frac{x^2 - 3/5}{-3/5} = -\frac{5}{3} \left(x^2 - \frac{3}{5} \right).$$

$$A_2 = \int_{-1}^1 l_2(x) dx = -\frac{5}{3} \int_{-1}^1 \left(x^2 - \frac{3}{5} \right) dx = -\frac{5}{3} \left[\frac{x^3}{3} - \frac{3}{5}x \right]_{-1}^1 = \frac{8}{9},$$

与前述结果一致。

3. 最终公式。三点 Gauss-Legendre 求积公式为

$$\int_{-1}^1 f(x) dx \approx \frac{5}{9}f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9}f(0) + \frac{5}{9}f\left(\sqrt{\frac{3}{5}}\right),$$

或写为数值形式:

$$\int_{-1}^1 f(x) dx \approx 0.5555555556 f(-0.7745966692) + 0.8888888889 f(0) + 0.5555555556 f(0.7745966692).$$

4. 验证代数精度。取 $f(x) = x^5$ (5 次多项式, $2n - 1 = 5$): 精确积分:

$$\int_{-1}^1 x^5 dx = 0.$$

近似积分:

$$\frac{5}{9} \left(-\sqrt{\frac{3}{5}}\right)^5 + \frac{8}{9} \cdot 0^5 + \frac{5}{9} \left(\sqrt{\frac{3}{5}}\right)^5 = \frac{5}{9} \left[-\left(\frac{3}{5}\right)^{5/2} + \left(\frac{3}{5}\right)^{5/2}\right] = 0.$$

因此公式对 x^5 精确成立, 验证了其具有 $2n - 1 = 5$ 次代数精度。

 这个在 24-25 考过判断, 具体是让你判断一个公式是不是 Gauss 积分公式, 主要那个公式是在 $[-1, 1]$ 上且权函数为 $\rho(x) \equiv 1$, 刚好就是 Gauss-Legendre 积分公式, 此时我们只要验证一下系数是相同即可。

Gauss-Chebyshev 求积公式

Chebyshev 多项式定义为

$$T_n(x) = \cos(n \arccos x), \quad -1 \leq x \leq 1,$$

它是 $[-1, 1]$ 上关于权函数 $\rho(x) = (1 - x^2)^{-1/2}$ 的正交多项式。其 n 个互异零点为

$$x_k = \cos\left(\frac{2k-1}{2n}\pi\right), \quad k = 1, 2, \dots, n.$$

n 点 Gauss-Chebyshev 求积公式为

$$\int_{-1}^1 \frac{1}{\sqrt{1-x^2}} f(x) dx \approx \frac{\pi}{n} \sum_{k=1}^n f(x_k), \quad (38)$$

其中 x_k 同上。

误差公式为

$$E_n[f] = \frac{\pi}{2^{2n-1}(2n)!} f^{(2n)}(\xi), \quad -1 < \xi < 1.$$

 下面我们来看一个例子:

例 3.6 (带权 \sqrt{x} 的 Gauss 公式). 构造形如

$$\int_0^1 \sqrt{x} f(x) dx \approx A_0 f(x_0) + A_1 f(x_1) \quad (39)$$

的两点 Gauss 公式。令 (39) 对 $f(x) = 1, x, x^2, x^3$ 精确成立, 得方程组

$$\begin{cases} A_0 + A_1 = \int_0^1 \sqrt{x} dx = \frac{2}{3}, \\ A_0 x_0 + A_1 x_1 = \int_0^1 x^{3/2} dx = \frac{2}{5}, \\ A_0 x_0^2 + A_1 x_1^2 = \int_0^1 x^{5/2} dx = \frac{2}{7}, \\ A_0 x_0^3 + A_1 x_1^3 = \int_0^1 x^{7/2} dx = \frac{2}{9}. \end{cases} \quad (40)$$

利用 (40) 中前两式可得

$$\frac{2}{3}x_0 + (x_1 - x_0)A_1 = \frac{2}{5}.$$

类似地, 分别用第一式与第三式、第一式与第四式组合得

$$\frac{2}{5}x_0 + (x_1 - x_0)x_1 A_1 = \frac{2}{7}, \quad \frac{2}{7}x_0 + (x_1 - x_0)x_1^2 A_1 = \frac{2}{9}.$$

消去 $(x_1 - x_0)A_1$ 得

$$\frac{2}{5}(x_0 + x_1) - \frac{2}{3}x_0 x_1 = \frac{2}{7}, \quad \frac{2}{7}(x_0 + x_1) - \frac{2}{5}x_0 x_1 = \frac{2}{9}.$$

解得

$$x_0 x_1 = \frac{5}{21}, \quad x_0 + x_1 = \frac{10}{9}.$$

进而解出

$$x_0 \approx 0.821162, \quad x_1 \approx 0.289949, \quad A_0 \approx 0.389111, \quad A_1 \approx 0.277556.$$

于是所求 Gauss 公式为

$$\int_0^1 \sqrt{x} f(x) dx \approx 0.389111 f(0.821162) + 0.277556 f(0.289949).$$



3.6 离散情形的最佳平方逼近与最小二乘法

3.6.1 离散最佳平方逼近问题

给定向量 $X = (x_1, x_2, \dots, x_n)^T$, $Y = (y_1, y_2, \dots, y_n)^T \in \mathbb{R}^n$ 及权系数 $w_i > 0, i = 1, \dots, n$, 定义内积与范数:

$$(X, Y) = \sum_{i=1}^n w_i x_i y_i, \quad \|X\| = \sqrt{(X, X)}. \quad (41)$$

设 $X_i = (x_{1i}, x_{2i}, \dots, x_{ni})^T, i = 1, \dots, l (l \leq n)$ 为 \mathbb{R}^n 中一组线性无关向量, 记子空间

$$H_l = \text{span}\{X_1, X_2, \dots, X_l\}.$$

对任意 $Y \in \mathbb{R}^n$, 考虑最佳逼近问题: 求 $X^* = \sum_{k=1}^l c_k^* X_k \in H_l$, 使得

$$\|Y - X^*\| = \inf_{X \in H_l} \|Y - X\|. \quad (42)$$

称 X^* 为 Y 在离散意义下的最佳平方逼近向量。

由正交投影原理，系数 $\{c_i^*\}_{i=1}^l$ 由法方程组唯一确定：

$$\sum_{j=1}^l (X_i, X_j) c_j^* = (Y, X_i), \quad i = 1, \dots, l. \quad (43)$$

记 $d_i = (Y, X_i)$ ，则 (43) 可写为矩阵形式

$$Ac^* = d, \quad (44)$$

其中

$$c^* = (c_1^*, \dots, c_l^*)^T, \quad d = (d_1, \dots, d_l)^T, \quad A = (a_{ij})_{l \times l}, \quad a_{ij} = (X_i, X_j).$$

由 X_1, \dots, X_l 线性无关可知 A 对称正定，从而 (44) 存在唯一解。

3.6.2 多变量最小二乘拟合算法步骤

1. 建立数学模型。确定拟合函数空间

$$\Phi_l = \text{span}\{\phi_1, \phi_2, \dots, \phi_l\}, \quad l \leq n,$$

其中 $\phi_i(x_1, x_2, \dots, x_m)$ ($i = 1, \dots, l$) 为线性无关多元函数组。

2. 定义优化问题。给定测量数据 $(x_1^{(i)}, x_2^{(i)}, \dots, x_m^{(i)}, y_i)$, $i = 1, \dots, n$ ，求系数 $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_l^*)^T$ 使得

$$\sum_{i=1}^n w_i [y_i - \phi^*(x_1^{(i)}, \dots, x_m^{(i)})]^2 = \min_{\alpha \in \mathbb{R}^l} \sum_{i=1}^n w_i [y_i - \phi(x_1^{(i)}, \dots, x_m^{(i)})]^2, \quad (45)$$

其中

$$\phi^*(x_1, \dots, x_m) = \sum_{k=1}^l \alpha_k^* \phi_k(x_1, \dots, x_m), \quad \phi(x_1, \dots, x_m) = \sum_{k=1}^l \alpha_k \phi_k(x_1, \dots, x_m).$$

3. 构造法方程组。定义离散内积：

$$(\phi_j, \phi_k) = \sum_{i=1}^n w_i \phi_j(x_1^{(i)}, \dots, x_m^{(i)}) \phi_k(x_1^{(i)}, \dots, x_m^{(i)}),$$

$$(y, \phi_j) = \sum_{i=1}^n w_i y_i \phi_j(x_1^{(i)}, \dots, x_m^{(i)}).$$

则法方程组为

$$\sum_{k=1}^l (\phi_j, \phi_k) \alpha_k^* = (y, \phi_j), \quad j = 1, \dots, l.$$

4. 解法方程组。求解上述线性方程组得系数 α^* ，即得最小二乘拟合函数 ϕ^* 。

该算法将多变量最小二乘问题转化为离散最佳平方逼近问题，其中内积由 (41) 定义在数据点上的加权和实现。

4 线性代数方程组求解

4.1 预备知识

4.1.1 矩阵范数

定义 4.1 (从属范数). 设 $\|\cdot\|$ 是 \mathbb{R}^n 上的向量范数, 对任一 $A \in \mathbb{R}^{n \times n}$, 定义

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1} \|Ax\|.$$

称 $\|A\|$ 为由向量范数 $\|\cdot\|$ 导出的**矩阵从属范数**。



定理 4.1 (从属范数的性质). 矩阵从属范数满足:

1. $\|A\| \geq 0$, 且 $\|A\| = 0$ 当且仅当 $A = 0$;
2. $\|\alpha A\| = |\alpha| \cdot \|A\|$, $\forall \alpha \in \mathbb{R}$;
3. $\|A + B\| \leq \|A\| + \|B\|$, $\forall A, B \in \mathbb{R}^{n \times n}$;
4. $\|AB\| \leq \|A\| \cdot \|B\|$ (次可乘性)。



1. **列和范数 (1-范数):**

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|.$$

2. **行和范数 (∞ -范数):**

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

3. **谱范数 (2-范数):**

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

其中 $\lambda_{\max}(A^T A)$ 表示 $A^T A$ 的最大特征值。

例 4.1 (矩阵范数计算). 设 $A = \begin{bmatrix} 1 & -2 \\ -3 & 4 \end{bmatrix}$, 则

$$\|A\|_1 = \max\{1+3, 2+4\} = 6, \quad \|A\|_\infty = \max\{1+2, 3+4\} = 7.$$

计算 $A^T A = \begin{bmatrix} 10 & -14 \\ -14 & 20 \end{bmatrix}$, 其特征方程为

$$\lambda^2 - 30\lambda + 4 = 0, \quad \lambda = 15 \pm \sqrt{221}.$$

于是

$$\|A\|_2 = \sqrt{15 + \sqrt{221}} \approx 5.46.$$



4.1.2 谱半径

定义 4.2 (谱半径). 设 $A \in \mathbb{R}^{n \times n}$, 称

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$$

为 A 的谱半径, 其中 $\sigma(A)$ 表示 A 的特征值全体。



显然有 $\|A\|_2 = \sqrt{\rho(A^T A)}$ 。若 A 对称, 则 $\|A\|_2 = \rho(A)$ 。

定理 4.2 (谱半径与范数的关系). 设 $A \in \mathbb{R}^{n \times n}$, 则

1. 对任意从属范数 $\|\cdot\|$, 有 $\rho(A) \leq \|A\|$;
2. 对任意 $\varepsilon > 0$, 存在从属范数 $\|\cdot\|_\varepsilon$, 使得

$$\|A\|_\varepsilon \leq \rho(A) + \varepsilon.$$

4.1.3 矩阵收敛

定义 4.3 (矩阵序列收敛). 矩阵序列 $A^{(k)} = (a_{ij}^{(k)}) \in \mathbb{R}^{n \times n}$ 收敛于 $A = (a_{ij})$, 若

$$\lim_{k \rightarrow \infty} a_{ij}^{(k)} = a_{ij}, \quad i, j = 1, \dots, n.$$

矩阵序列 $A^{(k)} \rightarrow A$ 的充要条件是 $\|A^{(k)} - A\| \rightarrow 0$ (对任意从属范数)。

定理 4.3 (矩阵幂收敛的谱条件). $A^k \rightarrow 0$ (当 $k \rightarrow \infty$) 的充要条件是 $\rho(A) < 1$ 。

定理 4.4 (Neumann 引理). 矩阵幂级数 $\sum_{k=0}^{\infty} A^k$ 收敛的充要条件是 $\rho(A) < 1$, 且此时

$$\sum_{k=0}^{\infty} A^k = (I - A)^{-1}.$$

定理 4.5 (Banach 引理). 若 $A \in \mathbb{R}^{n \times n}$ 非奇异, $E \in \mathbb{R}^{n \times n}$ 满足 $\|A^{-1}\| \cdot \|E\| < 1$, 则 $A + E$ 非奇异, 且

$$\|(A + E)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|E\|}.$$

4.1.4 条件数

定义 4.4 (条件数). 对非奇异矩阵 $A \in \mathbb{R}^{n \times n}$, 定义其条件数为

$$\text{cond}(A) = \|A^{-1}\| \cdot \|A\|.$$

当采用 p -范数时, 记为 $\text{cond}_p(A)$ 。

显然 $\text{cond}(A) \geq \|A^{-1}A\| = \|I\| = 1$ 。若 A 对称正定, 其特征值为 $0 < \lambda_{\min} \leq \lambda_{\max}$, 则

$$\|A\|_2 = \lambda_{\max}, \quad \|A^{-1}\|_2 = \frac{1}{\lambda_{\min}},$$

于是

$$\text{cond}_2(A) = \frac{\lambda_{\max}}{\lambda_{\min}}.$$

对一般非奇异矩阵, 常定义其谱条件数为

$$\text{cond}_{sp}(A) = \frac{\max\{|\lambda_i(A)| : i = 1, \dots, n\}}{\min\{|\lambda_i(A)| : i = 1, \dots, n\}}. \quad (46)$$

4.1.5 对角占优矩阵

定义 4.5 (对角占优). 矩阵 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ 称为**按行对角占优**, 若

$$\sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \leq |a_{ii}|, \quad i = 1, \dots, n,$$

且至少对一个 i 严格成立。若对所有 i 严格成立, 则称 A 按行严格对角占优。类似可定义按列对角占优。

定理 4.6 (严格对角占优矩阵非奇异). 若 A 严格对角占优, 则 A 非奇异。

证明. 仅证按行严格对角占优情形。假设 $Ax = 0$ 有非零解 $x = (x_1, \dots, x_n)^T$, 设 $|x_i| = \max_j |x_j| > 0$ 。由第 i 行方程

$$\sum_{j=1}^n a_{ij}x_j = 0,$$

得

$$|a_{ii}||x_i| \leq \sum_{j \neq i} |a_{ij}||x_j| \leq |x_i| \sum_{j \neq i} |a_{ij}|,$$

从而 $|a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$, 与严格对角占优矛盾。 \square

4.1.6 可约与不可约矩阵

定义 4.6 (可约矩阵). 设 A 为 n 阶方阵 ($n \geq 2$), 若存在置换矩阵 P 使得

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ O & A_{22} \end{bmatrix},$$

其中 A_{11} 为 r 阶方阵 ($1 \leq r < n$), 则称 A 为**可约矩阵**; 否则称为不可约矩阵。

定理 4.7 (对角占优且不可约则非奇异). 若 A 为对角占优且不可约, 则 A 非奇异。

证明思路. 设 $A\alpha = 0$, 分情况讨论 $|\alpha_i|$ 的相对大小, 利用对角占优和不可约条件导出矛盾。 \square

4.1.7 其他重要矩阵类

定义 4.7 (性质 A). 若存在置换矩阵 P 使得

$$PAP^T = \begin{bmatrix} D_1 & H \\ R & D_2 \end{bmatrix},$$

其中 D_1, D_2 为对角矩阵, 则称 A 具有**性质 A**。

定义 4.8 (稀疏矩阵). 若矩阵 $A \in \mathbb{R}^{n \times n}$ 的非零元个数为 $O(n)$, 则称 A 为**稀疏矩阵**。

定义 4.9 (非负矩阵). 若 A 的所有元素非负, 则称 A 为**非负矩阵**, 记作 $A \geq 0$ 。

定义 4.10 (M 矩阵). 矩阵 $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ 称为 M 矩阵, 如果满足:

1. $a_{ii} > 0, i = 1, \dots, n$;
2. $a_{ij} \leq 0, i \neq j$;
3. A 非奇异;
4. $A^{-1} \geq 0$.



4.2 Gauss 消去与矩阵分解

定义 4.11 (列主元). 在 Gauss 消去法的第 k 步 ($1 \leq k \leq n-1$), 从第 k 列的第 k 行至第 n 行中选取绝对值最大的元素作为主元, 并通过行交换将其换到 (k, k) 位置。



定义 4.12 (全主元). 在第 k 步, 从右下 $(n-k+1) \times (n-k+1)$ 子矩阵中选取绝对值最大的元素作为主元, 并通过行交换和列交换将其换到 (k, k) 位置。



4.2.1 LU 分解

Gauss 消去法等价于将系数矩阵 A 分解为一个单位下三角矩阵 L 与一个上三角矩阵 U 的乘积:

$$A = LU = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_{21} & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{bmatrix} \cdot \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & u_{nn} \end{bmatrix}.$$

称此分解为矩阵 A 的 **LU 分解**。

利用 LU 分解求解 $Ax = b$ 的步骤为:

1. 分解 $A = LU$;
2. 解下三角方程组 $Lz = b$ (前代);
3. 解上三角方程组 $Ux = z$ (回代)。

定理 4.8 (LU 分解的充要条件). 当且仅当 A 的所有顺序主子阵 A_k ($k = 1, \dots, n$) 均非奇异时, A 存在唯一的 LU 分解 (L 为单位下三角阵, U 为上三角阵)。



定理 4.9 (列主元下的 PLU 分解). 设 A 非奇异, 则存在置换矩阵 P 、单位下三角阵 L (元素满足 $|l_{ij}| \leq 1$) 和上三角阵 U , 使得

$$PA = LU.$$

该分解可由列主元 Gauss 消去法得到。



Doolittle 分解与 Crout 分解

 这个重点记！年年考！

定义 4.13 (Doolittle 分解). 若 A 的 LU 分解中 L 为单位下三角阵, U 为一般上三角阵, 即

$$A = \begin{bmatrix} 1 & & & 0 \\ l_{21} & 1 & & \\ \vdots & \ddots & \ddots & \\ l_{n1} & \cdots & l_{n,n-1} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix},$$

则称该分解为 *Doolittle 分解*。

定义 4.14 (Crout 分解). 若分解中 L 为一般下三角阵, U 为单位上三角阵, 即

$$A = \begin{bmatrix} l_{11} & & & 0 \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & \cdots & u_{1n} \\ & 1 & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix},$$

则称该分解为 *Crout 分解*。

两种分解均可通过比较等式两边对应元素, 用待定系数法递推计算。

LDU 分解

更精细地, 可将 A 分解为单位下三角阵 L 、对角阵 D 和单位上三角阵 U 的乘积:

$$A = LDU, \quad D = \text{diag}(d_1, d_2, \dots, d_n). \quad (47)$$

这是 LU 分解的进一步变形, 当 A 对称时可简化为 $A = LDL^T$ 。

例 4.2 (LU 分解与 LDU 分解). 设

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \\ -1 & -3 & 0 \end{bmatrix}.$$

其 LU 分解为

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

进一步可得 LDU 分解:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & \frac{1}{2} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

对于 Doolittle 分解, 元素可按如下递推计算:

$$\begin{aligned} u_{1j} &= a_{1j}, \quad j = 1, \dots, n, \\ l_{i1} &= \frac{a_{i1}}{u_{11}}, \quad i = 2, \dots, n, \end{aligned}$$

对 $k = 2, \dots, n$:

$$\begin{aligned} u_{kj} &= a_{kj} - \sum_{r=1}^{k-1} l_{kr} u_{rj}, \quad j = k, \dots, n, \\ l_{ik} &= \frac{1}{u_{kk}} \left(a_{ik} - \sum_{r=1}^{k-1} l_{ir} u_{rk} \right), \quad i = k+1, \dots, n. \end{aligned}$$

 下面针对两种特殊情形, 做进一步的讨论。

定理 4.10 (对称正定矩阵的 Cholesky 分解). 当 A 为对称正定矩阵时, A 存在三角分解

$$A = L \cdot L^T, \quad (48)$$

其中 L 为下三角矩阵。



证明. 由于 A 对称正定, 其顺序主子式均大于零, 因此存在唯一的 LDU 分解 $A = LDU$, 其中 L 是单位下三角矩阵, D 是对角矩阵, U 是单位上三角矩阵。由 $A^T = A$ 可得

$$U^T D L^T = L D U.$$

由三角分解的唯一性, 必有 $U = L^T$, 于是

$$A = L D L^T. \quad (49)$$

设 $D = \text{diag}(d_1, d_2, \dots, d_n)$ 。由于 A 正定, 对任意非零向量 x 有 $x^T A x > 0$ 。取 $x = L^{-1} e_i$ (其中 e_i 是第 i 个单位向量), 由式(49)得 $d_i > 0, i = 1, 2, \dots, n$ 。

定义 $D^{1/2} = \text{diag}(\sqrt{d_1}, \sqrt{d_2}, \dots, \sqrt{d_n})$, 并令 $\tilde{L} = L D^{1/2}$, 则 \tilde{L} 仍为下三角矩阵, 且

$$A = L D^{1/2} D^{1/2} L^T = (L D^{1/2})(L D^{1/2})^T = \tilde{L} \tilde{L}^T.$$

式(48)称为 Cholesky 分解。 □

下面利用待定系数法给出 $A = L L^T$ 分解的具体算法。设

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \cdot \begin{bmatrix} l_{11} & l_{21} & \cdots & l_{n1} \\ 0 & l_{22} & \cdots & l_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & l_{nn} \end{bmatrix}.$$

比较等式两边对应元素, 依次可得:

$$\begin{aligned} l_{11}^2 &= a_{11}, \\ l_{21} l_{11} &= a_{21}, \quad l_{21}^2 + l_{22}^2 = a_{22}, \\ &\vdots \end{aligned}$$

一般地, 对 $i = 1, 2, \dots, n$, 有

$$\begin{aligned} l_{ij} &= \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk} \right) / l_{jj}, \quad j = 1, 2, \dots, i-1, \\ l_{ii} &= \left(a_{ii} - \sum_{k=1}^{i-1} l_{ik}^2 \right)^{1/2}. \end{aligned} \quad (50)$$

由 A 的正定性可以证明式(50)中平方根内的表达式为正, 从而 $l_{ii} > 0$ 。

对称正定矩阵 A 的 Cholesky 分解在计算 l_{ii} 时需要开平方运算。若采用 $A = LDL^T$ 分解 (其中 L 为单位下三角矩阵, D 为对角矩阵), 则可避免开方。下面举例说明。

例 4.3. 考虑矩阵

$$A = \begin{bmatrix} 3 & 3 & 5 \\ 3 & 5 & 9 \\ 5 & 9 & 17 \end{bmatrix}.$$

将其分解为 $A = LDL^T$, 其中

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}, \quad D = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix}.$$

通过矩阵乘法比较元素可得:

$$\begin{aligned} d_1 &= 3, \\ l_{21}d_1 &= 3 \Rightarrow l_{21} = 1, \\ l_{31}d_1 &= 5 \Rightarrow l_{31} = \frac{5}{3}, \\ d_2 + l_{21}^2d_1 &= 5 \Rightarrow d_2 = 2, \\ l_{32}d_2 + l_{31}l_{21}d_1 &= 9 \Rightarrow l_{32} = 2, \\ d_3 + l_{31}^2d_1 + l_{32}^2d_2 &= 17 \Rightarrow d_3 = \frac{2}{3}. \end{aligned}$$



定理 4.11 (解三对角方程组的追赶法). 设线性方程组 $Ax = b$ 的系数矩阵 A 为三对角矩阵

$$A = \begin{bmatrix} e_1 & f_1 & 0 & \cdots & 0 \\ d_2 & e_2 & f_2 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & d_{n-1} & e_{n-1} & f_{n-1} \\ 0 & \cdots & 0 & d_n & e_n \end{bmatrix}, \quad (51)$$

且满足 $|e_1| > |f_1| > 0$, $|e_n| > |d_n| > 0$, 以及 $|e_i| \geq |d_i| + |f_i|$ ($i = 2, \dots, n-1$) 的强对角占



优条件, 则 A 可作如下 LU 分解:

$$A = LU = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ l_2 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & l_n & 1 \end{bmatrix} \cdot \begin{bmatrix} r_1 & f_1 & 0 & \cdots & 0 \\ 0 & r_2 & f_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & r_n & \end{bmatrix}. \quad (52)$$

证明. 将 A 与 LU 的乘积对应元素进行比较, 可得递推关系:

$$\begin{aligned} r_1 &= e_1, \\ l_i r_{i-1} = d_i &\Rightarrow l_i = \frac{d_i}{r_{i-1}}, \quad i = 2, 3, \dots, n, \\ r_i + l_i f_{i-1} = e_i &\Rightarrow r_i = e_i - l_i f_{i-1}, \quad i = 2, 3, \dots, n. \end{aligned}$$

在所述对角占优条件下, 可证所有 $r_i \neq 0$, 分解可行且数值稳定。□

基于分解(52), 求解 $Ax = b$ 的追赶法步骤如下:

步骤 1 (分解, 即“追”过程):

$$\begin{aligned} r_1 &= e_1, \\ \text{对 } i = 2, 3, \dots, n \text{ 计算 } l_i &= \frac{d_i}{r_{i-1}}, \quad r_i = e_i - l_i f_{i-1}. \end{aligned}$$

步骤 2 (解 $Ly = b$, 即“赶”过程):

$$\begin{aligned} y_1 &= b_1, \\ \text{对 } i = 2, 3, \dots, n \text{ 计算 } y_i &= b_i - l_i y_{i-1}. \end{aligned}$$

步骤 3 (解 $Ux = y$, 即“回代”过程):

$$\begin{aligned} x_n &= \frac{y_n}{r_n}, \\ \text{对 } i = n-1, n-2, \dots, 1 \text{ 计算 } x_i &= \frac{y_i - f_i x_{i+1}}{r_i}. \end{aligned}$$

4.3 扰动分析、Gauss 消去法的舍入误差

 注. 感觉没法考啊。。。考这个的话真的老冯飞升了吧。。。

4.4 迭代方法

考虑线性方程组

$$Ax = b, \quad (53)$$

其中 $A \in \mathbb{R}^{n \times n}$ 非奇异, $b \in \mathbb{R}^n$. 将 A 分解为

$$A = D - L - U, \quad (54)$$

其中 $D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$ 为 A 的对角部分, $-L$ 和 $-U$ 分别为 A 的严格下三角部分和严格上三角部分 (即 L 和 U 的主对角元素均为 0)。

4.4.1 Jacobi 迭代法

定义 4.15 (Jacobi 迭代法). 设 A 的对角元素 $a_{ii} \neq 0$ ($i = 1, 2, \dots, n$), 由式(54), 将方程组 $Ax = b$ 改写为

$$Dx = (L + U)x + b.$$

由此构造迭代格式

$$x^{(k+1)} = D^{-1}(L + U)x^{(k)} + D^{-1}b, \quad k = 0, 1, 2, \dots \quad (55)$$

式(55)称为 **Jacobi 迭代法** (又称简单迭代法)。记 $B_J = D^{-1}(L + U)$ 为 Jacobi 迭代矩阵, $f_J = D^{-1}b$, 则迭代格式可简写为

$$x^{(k+1)} = B_J x^{(k)} + f_J.$$



分量形式推导. 将方程组 $Ax = b$ 的第 i 个方程

$$\sum_{j=1}^n a_{ij}x_j = b_i$$

分离出 x_i 项, 得到

$$a_{ii}x_i = b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j.$$

由于 $a_{ii} \neq 0$, 解出 x_i 并建立迭代格式:

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j^{(k)} \right), \quad i = 1, 2, \dots, n. \quad (56)$$

式(56)即 Jacobi 迭代法的分量计算公式。可见 Jacobi 迭代法每次更新 $x_i^{(k+1)}$ 时, 全部使用上一轮迭代值 $x^{(k)}$ 的分量, 因此更新可以同时进行 (并行计算)。□

4.4.2 Gauss-Seidel 迭代法

定义 4.16 (Gauss-Seidel 迭代法). 在 Jacobi 迭代法分量形式(56)中, 当计算 $x_i^{(k+1)}$ 时, 实际上 $x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_{i-1}^{(k+1)}$ 已经得到。Gauss-Seidel (G-S) 迭代法利用这些最新值代替旧值, 得到迭代格式

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right), \quad i = 1, 2, \dots, n. \quad (57)$$




矩阵形式推导. 由 $A = D - L - U$, 将 $Ax = b$ 写为

$$(D - L)x = Ux + b.$$

由于 $D - L$ 为下三角矩阵且对角元 $a_{ii} \neq 0$, 故可逆。由此构造迭代格式

$$x^{(k+1)} = (D - L)^{-1}Ux^{(k)} + (D - L)^{-1}b. \quad (58)$$

式(58)即 Gauss-Seidel 迭代法的矩阵形式。记 $B_{GS} = (D - L)^{-1}U$ 为 G-S 迭代矩阵, $f_{GS} = (D - L)^{-1}b$ 。□

 注. 与 *Jacobi* 迭代法相比, *G-S* 迭代法在计算 $x_i^{(k+1)}$ 时立即使用已更新的分量, 通常收敛更快, 且节省存储空间 (只需存储一个向量 x). 但更新顺序依赖, 不利于并行计算.

4.4.3 逐次超松弛迭代法 (SOR)

定义 4.17 (逐次超松弛迭代法). 将 *G-S* 迭代格式(57)视为中间结果:

$$\tilde{x}_i^{(k+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right).$$

引入松弛因子 $\omega > 0$, 将迭代值定义为当前值与中间值的加权平均:

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \omega\tilde{x}_i^{(k+1)}, \quad i = 1, 2, \dots, n. \quad (59)$$

代入 $\tilde{x}_i^{(k+1)}$ 得

$$x_i^{(k+1)} = (1 - \omega)x_i^{(k)} + \frac{\omega}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right). \quad (60)$$

式(59)称为 **逐次超松弛迭代法** (*Successive Over-Relaxation Method, SOR*). 当 $\omega = 1$ 时即为 *G-S* 迭代法. 

矩阵形式推导. 将 SOR 迭代的分量形式改写为

$$a_{ii}x_i^{(k+1)} + \omega \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} = (1 - \omega)a_{ii}x_i^{(k)} - \omega \sum_{j=i+1}^n a_{ij}x_j^{(k)} + \omega b_i.$$


写成矩阵形式为

$$(D - \omega L)x^{(k+1)} = [(1 - \omega)D + \omega U]x^{(k)} + \omega b.$$

由于 $D - \omega L$ 可逆 (因其为对角元非零的下三角矩阵), 得到 SOR 迭代的矩阵形式:

$$x^{(k+1)} = (D - \omega L)^{-1}[(1 - \omega)D + \omega U]x^{(k)} + \omega(D - \omega L)^{-1}b. \quad (61)$$

记 $B_{SOR}(\omega) = (D - \omega L)^{-1}[(1 - \omega)D + \omega U]$ 为 SOR 迭代矩阵. □

定理 4.12 (SOR 方法收敛的必要条件). *SOR* 方法收敛的必要条件是松弛因子满足 $0 < \omega < 2$. 

证明概要. 设 $B_{SOR}(\omega)$ 的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$. 由矩阵行列式关系可得

$$\det(B_{SOR}(\omega)) = \prod_{i=1}^n \lambda_i = (1 - \omega)^n.$$

若 SOR 方法收敛, 则谱半径 $\rho(B_{SOR}(\omega)) < 1$, 从而 $|\lambda_i| < 1$. 于是

$$|\det(B_{SOR}(\omega))| = \prod_{i=1}^n |\lambda_i| < 1.$$

结合 $\det(B_{SOR}(\omega)) = (1 - \omega)^n$, 得 $|1 - \omega| < 1$, 即 $0 < \omega < 2$. □


 注. • 当 $0 < \omega < 1$ 时称为 **欠松弛** (*Under-Relaxation*), 可用于改善某些发散迭代的收敛性.

- 当 $\omega = 1$ 时即为 G - S 迭代。
- 当 $1 < \omega < 2$ 时称为 **超松弛 (Over-Relaxation)**，通常可加速 G - S 迭代的收敛速度。
- 最优松弛因子 ω_{opt} 的选取依赖于系数矩阵 A 的性质。对于一类重要的矩阵（如对称正定矩阵或具有性质 A 的矩阵），存在理论公式可估算 ω_{opt} 。

为了便于理论分析，下面给出 Jacobi 迭代法、Gauss-Seidel (G-S) 迭代法和逐次超松弛 (SOR) 迭代法的统一矩阵形式。一般的线性迭代法可以通过对原系数矩阵 A 进行矩阵分裂来构造。

定义 4.18 (矩阵分裂与迭代格式). 设 $A \in \mathbb{R}^{n \times n}$ 非奇异。将 A 分解为

$$A = M - N, \quad (62)$$

其中 M 为 n 阶非奇异矩阵，通常要求以 M 为系数矩阵的线性方程组易于求解（例如 M 为对角矩阵或三角矩阵）。 

利用分裂式(62)，方程组 $Ax = b$ 可等价地写为

$$Mx = Nx + b \quad \Rightarrow \quad x = M^{-1}Nx + M^{-1}b.$$

由此建立迭代格式

$$x^{(k+1)} = M^{-1}Nx^{(k)} + M^{-1}b = Gx^{(k)} + M^{-1}b, \quad k = 0, 1, 2, \dots, \quad (63)$$

其中矩阵 $G = M^{-1}N$ 称为迭代法(63)的**迭代矩阵**。

将矩阵 A 按其对角部分、严格下三角部分和严格上三角部分分解为

$$A = D - L - U, \quad (64)$$

其中 $D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn})$ ， $-L$ 和 $-U$ 分别为 A 的严格下三角部分和严格上三角部分（即 L 和 U 的对角元素均为 0）。下面给出三种经典迭代法对应的矩阵分裂与迭代矩阵。

1. **Jacobi 迭代法**：

$$M = D, \quad N = L + U, \quad G_J = D^{-1}(L + U) = I - D^{-1}A.$$

2. **Gauss-Seidel (G-S) 迭代法**：

$$M = D - L, \quad N = U, \quad G_{GS} = (D - L)^{-1}U.$$

3. **逐次超松弛 (SOR) 迭代法**：考虑等价方程组 $\omega Ax = \omega b$ ($\omega \neq 0$)，对系数矩阵 ωA 作如下分裂：

$$M = D - \omega L, \quad N = (1 - \omega)D + \omega U,$$

其迭代矩阵为

$$G_{SOR}(\omega) = (D - \omega L)^{-1}[(1 - \omega)D + \omega U]. \quad (65)$$

当 $\omega = 1$ 时， $G_{SOR}(1) = (D - L)^{-1}U = G_{GS}$ ，即退化为 G-S 迭代法。

迭代格式(63)还可写成一种实用的**校正型**形式:

$$x^{(k+1)} = x^{(k)} + B^{-1}(b - Ax^{(k)}), \quad (66)$$

其中 B 为 n 阶非奇异矩阵, 同样要求方程组 $By = r$ 易于求解。比较式(63)与式(66), 可知 B 与迭代矩阵 G 的关系为

$$G = I - B^{-1}A, \quad \text{或等价地} \quad B = M. \quad (67)$$

在校正式(66)中, $r^{(k)} = b - Ax^{(k)}$ 为第 k 步的残差向量, $B^{-1}r^{(k)}$ 为校正量。

定理 4.13 (迭代法收敛的充要条件). 对于任意初始向量 $x^{(0)}$, 迭代格式(63)收敛的充分必要条件是迭代矩阵 G 的谱半径满足

$$\rho(G) < 1. \quad (68)$$

定理 4.14 (对角占优矩阵的收敛性). 若 A 为严格对角占优矩阵 (即 $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$ 对所有 i 成立) 或不可约对角占优矩阵, 则求解 $Ax = b$ 的 *Jacobi* 迭代法和 *SOR* 迭代法 (其中 $0 < \omega \leq 1$) 均收敛。

定理 4.15 (对称正定矩阵的收敛性). 设 A 为实对称正定矩阵, 则

1. 若 $2D - A$ 正定, 则 *Jacobi* 迭代法收敛;
2. 当 $0 < \omega < 2$ 时, *SOR* 迭代法收敛。

例 4.4. 考虑线性方程组 $Ax = b$ 的系数矩阵

$$A = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 1 & -1 \\ -2 & -2 & 1 \end{bmatrix}.$$

分析 *Jacobi* 迭代法和 *G-S* 迭代法的收敛性。

Jacobi 迭代矩阵为

$$G_J = D^{-1}(L + U) = \begin{bmatrix} 0 & 2 & -2 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \end{bmatrix}.$$

其特征多项式为 $\det(\lambda I - G_J) = \lambda^3$, 故特征值全为 0, 谱半径 $\rho(G_J) = 0 < 1$, 因此 *Jacobi* 迭代法收敛。

G-S 迭代矩阵为

$$G_{GS} = (D - L)^{-1}U = \begin{bmatrix} 0 & 2 & -2 \\ 0 & 2 & -1 \\ 0 & 8 & -6 \end{bmatrix}.$$

计算其特征多项式:

$$\begin{aligned} \det(\lambda I - G_{GS}) &= \lambda \begin{vmatrix} \lambda - 2 & 1 \\ -8 & \lambda + 6 \end{vmatrix} \\ &= \lambda[(\lambda - 2)(\lambda + 6) + 8] = \lambda(\lambda^2 + 4\lambda - 4). \end{aligned}$$

解得特征值 $\lambda_1 = 0$, $\lambda_{2,3} = -2 \pm 2\sqrt{2}$ 。谱半径 $\rho(G_{GS}) = |-2 + 2\sqrt{2}| = 2\sqrt{2} - 2 \approx 0.828$?
 注意 $-2 - 2\sqrt{2} \approx -4.828$, 其绝对值约为 4.828, 实际上 $\rho(G_{GS}) = 2 + 2\sqrt{2} \approx 4.828 > 1$ 。因此 G-S 迭代法发散。



注. 本例表明, 即使对同一个方程组, 不同迭代法的收敛性可能截然不同。*Jacobi* 迭代收敛并不意味着 *G-S* 迭代一定收敛, 反之亦然。在实际计算中, 应结合矩阵特性选择合适的迭代法, 并利用谱半径或矩阵范数进行收敛性判断。

4.5 共轭梯度法

注. 这个考了我吃。

5 非线性方程的数值解法

注. 关于重根和二分法部分不做过多赘述。。。

5.1 不动点迭代法

求非线性方程

$$f(x) = 0 \tag{69}$$

的根, 常将其转化为等价形式

$$x = \varphi(x), \tag{70}$$

其中 $\varphi(x)$ 称为**迭代函数**。

定义 5.1 (不动点). 若 x^* 满足 $\varphi(x^*) = x^*$, 则称 x^* 为迭代函数 φ 的一个**不动点**。此时显然有 $f(x^*) = 0$, 反之亦然。因此求 $f(x) = 0$ 的根等价于求 φ 的不动点。



给定一个初始近似值 x_0 , 构造迭代序列

$$x_{k+1} = \varphi(x_k), \quad k = 0, 1, 2, \dots \tag{71}$$

式(71)称为**不动点迭代法** (亦称简单迭代法或逐次逼近法)。若序列 $\{x_k\}$ 收敛到 x^* , 则 x^* 即为 φ 的不动点, 从而也是方程 $f(x) = 0$ 的根。

注. 不动点迭代法的关键在于选取合适的迭代函数 $\varphi(x)$ 。同一个方程 $f(x) = 0$ 可以转化为多种等价形式 $x = \varphi(x)$, 不同的 φ 可能导致迭代序列收敛性、收敛速度的显著差异。例如, 对 $f(x) = x^2 - 2 = 0$, 可以构造 $\varphi_1(x) = x - (x^2 - 2)$, 或 $\varphi_2(x) = \frac{1}{2}(x + \frac{2}{x})$ (即 *Newton* 法对应的迭代函数)。前者可能发散, 后者则平方收敛。

定理 5.1 (不动点迭代法的收敛性). 设迭代函数 $\varphi(x)$ 满足:

1. $\varphi(x) \in C[a, b] \cap C^1(a, b)$;
2. $\varphi(x)$ 将区间 $[a, b]$ 映射到自身, 即 $a \leq \varphi(x) \leq b$ 对所有 $x \in [a, b]$ 成立;
3. 存在常数 $0 < L < 1$, 使得 $|\varphi'(x)| \leq L$ 对所有 $x \in (a, b)$ 成立。



则有如下结论:

1. 对任意初始值 $x_0 \in [a, b]$, 由不动点迭代

$$x_{k+1} = \varphi(x_k), \quad k = 0, 1, 2, \dots$$

产生的序列 $\{x_k\}$ 必收敛到 φ 在 $[a, b]$ 上的唯一不动点 x^* (即 $x^* = \varphi(x^*)$)。

2. 序列 $\{x_k\}$ 有如下误差估计:

$$|x_k - x^*| \leq \frac{1}{1-L} |x_{k+1} - x_k|, \quad (72)$$

$$|x_k - x^*| \leq \frac{L^k}{1-L} |x_1 - x_0|. \quad (73)$$



5.1.1 收敛速度与收敛阶

为定量描述迭代序列 $\{x_k\}$ 收敛于极限 x^* 的快慢, 引入收敛阶的概念。

定义 5.2 (收敛阶). 设序列 $\{x_k\}$ 收敛到 x^* , 记误差 $e_k = x_k - x^*$. 若存在实数 $p \geq 1$ 及非零常数 $c > 0$, 使得


$$\lim_{k \rightarrow \infty} \frac{|e_{k+1}|}{|e_k|^p} = c, \quad (74)$$

则称序列 $\{x_k\}$ 是 p 阶收敛的, c 称为渐近误差常数。特别地,

- 当 $p = 1$ 且 $0 < c < 1$ 时, 称为线性收敛;
- 当 $p > 1$ 时, 称为超线性收敛;
- 当 $p = 2$ 时, 称为平方收敛或二次收敛。

若由迭代格式(71)产生的序列 $\{x_k\}$ 是 p 阶收敛的, 则称该迭代法是 p 阶收敛的。



 **注.** p 的大小直观反映了收敛速度: p 越大, 收敛越快。线性收敛意味着误差按比例 c 逐次衰减, 而超线性收敛 (尤其是平方收敛) 的误差衰减速度更快, 是高效迭代法追求的目标。

 下面的定理给出了不动点迭代法收敛阶与迭代函数 φ 在不动点处导数值的直接关系。

定理 5.2 (收敛阶的充要条件). 设迭代函数 $\varphi(x)$ 在不动点 x^* 的某邻域内具有直到 p 阶的连续导数 ($p \geq 1$)。则不动点迭代法(71)是 p 阶收敛的充分必要条件是

$$\varphi(x^*) = x^*, \quad \varphi^{(l)}(x^*) = 0, \quad l = 1, 2, \dots, p-1, \quad \varphi^{(p)}(x^*) \neq 0. \quad (75)$$

此时, 渐近误差常数为

$$\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k^p} = \frac{1}{p!} \varphi^{(p)}(x^*) \neq 0. \quad (76)$$



证明. 由 Taylor 展开, 将 $\varphi(x_k)$ 在 x^* 处展开至 p 阶:

$$\begin{aligned}x_{k+1} &= \varphi(x_k) \\&= \varphi(x^*) + \varphi'(x^*)(x_k - x^*) + \frac{\varphi''(x^*)}{2!}(x_k - x^*)^2 + \cdots \\&\quad + \frac{\varphi^{(p-1)}(x^*)}{(p-1)!}(x_k - x^*)^{p-1} + \frac{\varphi^{(p)}(\xi_k)}{p!}(x_k - x^*)^p,\end{aligned}$$

其中 ξ_k 介于 x_k 与 x^* 之间. 利用 $\varphi(x^*) = x^*$ 及条件(75), 得

$$e_{k+1} = x_{k+1} - x^* = \frac{\varphi^{(p)}(\xi_k)}{p!} e_k^p.$$

由于 $\varphi^{(p)}$ 连续且 $\varphi^{(p)}(x^*) \neq 0$, 当 $k \rightarrow \infty$ 时 $\xi_k \rightarrow x^*$, 故

$$\lim_{k \rightarrow \infty} \frac{e_{k+1}}{e_k^p} = \frac{1}{p!} \varphi^{(p)}(x^*),$$

满足 p 阶收敛定义. 反之, 若迭代法 p 阶收敛, 则必可推出条件(75)成立 (否则展开式中会出现更低阶的非零项, 导致收敛阶低于 p). \square

5.2 Newton 类迭代法

用迭代法求解方程 $f(x) = 0$, 关键在于选取合适的迭代函数 $\varphi(x)$, 使得迭代过程 $x_{n+1} = \varphi(x_n)$ 收敛到方程的根 α . Newton 法是一种经典而高效的构造方法.

设 $f(x)$ 在根 α 附近二次连续可微, 取初始近似值 x_0 位于 α 的邻域内. 将 $f(x)$ 在 x_0 处作 Taylor 展开:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi_0)}{2!}(x - x_0)^2, \quad (77)$$

其中 ξ_0 介于 x 与 x_0 之间. 令 $x = \alpha$, 则

$$0 = f(\alpha) = f(x_0) + f'(x_0)(\alpha - x_0) + \frac{f''(\xi_0)}{2!}(\alpha - x_0)^2.$$


忽略二阶小量 $(\alpha - x_0)^2$ 项, 并设 $f'(x_0) \neq 0$, 得到 α 的一个新的近似值:

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$


重复这一过程, 得到一般的迭代格式:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots \quad (78)$$

式(78)称为 **Newton 迭代法** (又称 Newton-Raphson 方法).

 **注.** Newton 法具有清晰的几何解释: 在曲线 $y = f(x)$ 上点 $(x_n, f(x_n))$ 处作切线, 该切线与 x 轴的交点的横坐标即为 x_{n+1} . 因此 Newton 法是用切线的零点来逼近曲线的零点, 属于切线法.

 Newton 法在适当条件下具有高阶收敛性.

定理 5.3 (Newton 法的局部收敛性). 设 $f(x)$ 在根 α 的某邻域内具有 m ($m \geq 2$) 阶连续导数, 且 α 是 $f(x) = 0$ 的单根 (即 $f'(\alpha) \neq 0$). 则当初始值 x_0 充分接近 α 时, Newton 迭代法(78)收敛, 且至少是二阶收敛的. 

证明概要. 将 Newton 法的迭代函数记为

$$\varphi(x) = x - \frac{f(x)}{f'(x)}.$$

计算其导数:

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


由于 α 是单根, $f(\alpha) = 0$ 且 $f'(\alpha) \neq 0$, 故 $\varphi'(\alpha) = 0$. 由前节定理 (收敛阶的充要条件) 知, 迭代法至少是二阶收敛的. 进一步, 利用 Taylor 展开可证当 x_0 充分接近 α 时, 迭代序列收敛到 α . \square

定理 5.4 (Newton 法的全局收敛性). 设函数 $f(x)$ 在有限区间 $[a, b]$ 上二阶可导, 且满足:

1. $f(a)f(b) < 0$ (即 $f(x)$ 在区间两端异号);
2. $f'(x) \neq 0$ 对任意 $x \in [a, b]$ 成立, 且 $f''(x)$ 在 $[a, b]$ 上不变号;
3. 区间端点满足

$$\left| \frac{f(a)}{f'(a)} \right| < b - a, \quad \left| \frac{f(b)}{f'(b)} \right| < b - a.$$

则对任意初始值 $x_0 \in [a, b]$, Newton 迭代法(78)均收敛于方程 $f(x) = 0$ 在 $[a, b]$ 内的唯一解 p , 且收敛阶为 2.

 **注.** 定理 5.4 的条件保证了函数 $f(x)$ 在 $[a, b]$ 上单调且凸性不变, 从而 Newton 迭代产生的序列不会跳出区间, 且每一步迭代都更接近真根. 条件 3 确保了初始切线交点仍落在区间内.

 下面我们来看一个利用牛顿法开方的例子:

对于给定正数 C , 考虑方程 $f(x) = x^2 - C = 0$, 其正根即为 \sqrt{C} . 应用 Newton 迭代公式(78), $f'(x) = 2x$, 得

$$x_{k+1} = x_k - \frac{x_k^2 - C}{2x_k} = \frac{1}{2} \left(x_k + \frac{C}{x_k} \right), \quad k = 0, 1, 2, \dots \quad (79)$$

式(79)是计算平方根的经典算法 (亦称巴比伦算法).

定理 5.5 (开方 Newton 法的全局收敛性). 对任意初始值 $x_0 > 0$, 迭代格式(79)产生的序列 $\{x_k\}$ 收敛到 \sqrt{C} , 且为平方收敛.

证明. 直接计算相邻迭代值与真值的关系:

$$\begin{aligned} x_{k+1} - \sqrt{C} &= \frac{1}{2} \left(x_k + \frac{C}{x_k} \right) - \sqrt{C} = \frac{1}{2x_k} (x_k^2 - 2\sqrt{C}x_k + C) = \frac{1}{2x_k} (x_k - \sqrt{C})^2, \\ x_{k+1} + \sqrt{C} &= \frac{1}{2} \left(x_k + \frac{C}{x_k} \right) + \sqrt{C} = \frac{1}{2x_k} (x_k^2 + 2\sqrt{C}x_k + C) = \frac{1}{2x_k} (x_k + \sqrt{C})^2. \end{aligned}$$

两式相除得

$$\frac{x_{k+1} - \sqrt{C}}{x_{k+1} + \sqrt{C}} = \left(\frac{x_k - \sqrt{C}}{x_k + \sqrt{C}} \right)^2. \quad (80)$$

记 $q_k = \frac{x_k - \sqrt{C}}{x_k + \sqrt{C}}$, 则式(80)给出 $q_{k+1} = q_k^2$. 递推得

$$q_k = q_0^{2^k}, \quad \text{其中 } q_0 = \frac{x_0 - \sqrt{C}}{x_0 + \sqrt{C}}.$$

由于 $x_0 > 0$, 易证 $|q_0| < 1$. 因此 $|q_k| = |q_0|^{2^k} \rightarrow 0$ 当 $k \rightarrow \infty$, 从而 $x_k \rightarrow \sqrt{C}$. 由 $x_{k+1} - \sqrt{C} = \frac{1}{2x_k}(x_k - \sqrt{C})^2$ 可知收敛阶为 2. \square


5.2.1 修正的 Newton 迭代法


应用 Newton 法每一步需计算 $f'(x_n)$, 若导数计算困难或代价较高, 可将 $f'(x_n)$ 固定为初始点的导数值 $f'(x_0)$, 得到修正的 Newton 迭代法:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_0)}, \quad n = 0, 1, 2, \dots \quad (81)$$

其迭代函数为

$$\varphi(x) = x - \frac{f(x)}{f'(x_0)}.$$

 **注.** 修正的 Newton 法在几何上是用过点 $(x_n, f(x_n))$ 且斜率为 $f'(x_0)$ 的直线 (即平行于初始切线的直线) 来逼近曲线, 取该直线与 x 轴的交点作为新的近似值. 该方法每步只需计算函数值 $f(x_n)$, 而导数 $f'(x_0)$ 只需计算一次, 节省了计算量, 但收敛速度通常降为线性收敛 (若 $f'(x_0) \neq 0$ 且 $f'(x_0) \approx f'(\alpha)$).

定理 5.6 (修正 Newton 法的局部收敛性). 设 $f(x)$ 在根 α 的邻域内二阶连续可导, $f'(\alpha) \neq 0$, 且初始值 x_0 充分接近 α . 则修正的 Newton 迭代法(81)局部收敛, 且为线性收敛. 

5.2.2 多重根的处理


设 α 为 $f(x) = 0$ 的 m 重根 ($m \geq 2$), 即

$$f(\alpha) = f'(\alpha) = \dots = f^{(m-1)}(\alpha) = 0, \quad f^{(m)}(\alpha) \neq 0.$$

此时若直接应用 Newton 迭代法

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

其收敛速度将减慢.

定理 5.7 (Newton 法在重根处的收敛速度). 设 α 是 $f(x) = 0$ 的 m 重根 ($m \geq 2$), 且 f 在 α 的邻域内充分光滑. 则 Newton 迭代法在 α 附近局部收敛, 但仅为线性收敛, 且收敛因子为 $1 - \frac{1}{m}$. 

证明. 记迭代函数 $\varphi(x) = x - \frac{f(x)}{f'(x)}$, 显然 $\varphi(\alpha) = \alpha$. 为计算 $\varphi'(\alpha)$, 令 $x = \alpha + h$ (h 充分小). 由 Taylor 展开:

$$\begin{aligned} f(\alpha + h) &= \frac{f^{(m)}(\alpha)}{m!} h^m + O(h^{m+1}), \\ f'(\alpha + h) &= \frac{f^{(m)}(\alpha)}{(m-1)!} h^{m-1} + O(h^m). \end{aligned}$$

于是

$$\begin{aligned}\varphi(\alpha+h) &= \alpha+h-\frac{f(\alpha+h)}{f'(\alpha+h)} \\ &= \alpha+h-\frac{\frac{f^{(m)}(\alpha)}{m!}h^m+O(h^{m+1})}{\frac{f^{(m)}(\alpha)}{(m-1)!}h^{m-1}+O(h^m)} \\ &= \alpha+h-\frac{h}{m}\cdot\frac{1+O(h)}{1+O(h)} \\ &= \alpha+\left(1-\frac{1}{m}\right)h+O(h^2).\end{aligned}$$

因此

$$\varphi'(\alpha)=\lim_{h\rightarrow 0}\frac{\varphi(\alpha+h)-\varphi(\alpha)}{h}=1-\frac{1}{m}.$$

当 $m \geq 2$ 时, $0 < 1 - \frac{1}{m} < 1$, 故迭代法线性收敛, 收敛因子为 $1 - \frac{1}{m}$. □

为了恢复平方收敛性, 可对 Newton 法进行修改。

定理 5.8 (已知重数的修正 Newton 法). 若已知根 α 的重数 m , 则迭代格式

$$x_{n+1}=x_n-m\cdot\frac{f(x_n)}{f'(x_n)} \quad (82)$$

在 α 附近局部平方收敛。

证明. 此时迭代函数为 $\psi(x)=x-m\frac{f(x)}{f'(x)}$. 类似上述推导可得

$$\psi(\alpha+h)=\alpha+h-m\cdot\frac{h}{m}+O(h^2)=\alpha+O(h^2),$$

从而 $\psi'(\alpha)=0$, 由收敛阶定理知迭代法至少二阶收敛。 □

然而实际中根的重数 m 往往未知. 另一种不依赖于 m 的修正方法如下。

定理 5.9 (未知重数的修正 Newton 法). 设 α 是 $f(x)=0$ 的 m 重根, 令

$$u(x)=\frac{f(x)}{f'(x)}.$$

则 α 是 $u(x)=0$ 的单根. 应用 Newton 法于 $u(x)$, 得迭代格式

$$x_{n+1}=x_n-\frac{u(x_n)}{u'(x_n)}=x_n-\frac{f(x_n)f'(x_n)}{[f'(x_n)]^2-f(x_n)f''(x_n)}. \quad (83)$$

此迭代法在 α 附近局部平方收敛。

证明. 由于 $f(x)=(x-\alpha)^mg(x)$, 其中 $g(\alpha)\neq 0$, 计算得

$$\begin{aligned}f'(x) &= m(x-\alpha)^{m-1}g(x)+(x-\alpha)^mg'(x), \\ u(x) &= \frac{f(x)}{f'(x)} = \frac{(x-\alpha)g(x)}{mg(x)+(x-\alpha)g'(x)}.\end{aligned}$$

可见 $u(\alpha)=0$, 但 $u'(\alpha)=\frac{g(\alpha)}{mg(\alpha)}=\frac{1}{m}\neq 0$, 故 α 是 $u(x)=0$ 的单根. 对 $u(x)$ 应用 Newton 法即得式(83), 且因其为单根情形, 故平方收敛。 □

5.3 Newton 下山法

在实际计算中, Newton 法的收敛性依赖于初始值 x_0 的选取。为了扩大收敛区域, 可引入松弛因子, 得到 Newton 下山法。

定义 5.3 (Newton 下山法). 在 Newton 迭代格式中引入下山因子 $\lambda > 0$:

$$x_{n+1} = x_n - \lambda \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, \dots \quad (84)$$

选取 λ 的原则是保证函数值绝对值单调下降:

$$|f(x_{n+1})| < |f(x_n)|. \quad (85)$$

式(84)称为 *Newton 下山法*。

 **注** (下山因子的选取策略). 通常采用以下自适应策略调整 λ :

1. 从 $\lambda = 1$ 开始, 用式(84)计算 x_{n+1} 。
2. 若 $|f(x_{n+1})| < |f(x_n)|$, 则接受该步, 并尝试将 λ 加倍 (如取 $\lambda := 2\lambda$) 进行下一步迭代, 以加速收敛。
3. 若 $|f(x_{n+1})| \geq |f(x_n)|$, 则逐步将 λ 减半 (如取 $\lambda := \lambda/2$), 重新计算 x_{n+1} , 直至下降条件(85)满足为止。

该策略确保每次迭代都使残差 $|f(x_n)|$ 严格减小, 同时尽可能采用较大的 λ 以保持较快的收敛速度。

定理 5.10 (Newton 下山法的性质). 设 $f(x)$ 在根 α 的某邻域内二阶连续可导, $f'(\alpha) \neq 0$ 。则存在 $\delta > 0$, 使得对任意 x_0 满足 $|x_0 - \alpha| < \delta$, 总可选取合适的下山因子 λ , 使 Newton 下山法产生的序列 $\{x_n\}$ 收敛到 α 。

证明思路. 当 x_n 接近 α 时, 有近似关系

$$f(x_{n+1}) \approx f(x_n) - \lambda f'(x_n)(x_n - \alpha) + O((x_n - \alpha)^2) = (1 - \lambda)f(x_n) + O((x_n - \alpha)^2).$$

因此只要选取 λ 使 $|1 - \lambda| < 1$ (即 $0 < \lambda < 2$), 并忽略高阶项, 便有 $|f(x_{n+1})| < |f(x_n)|$ 。通过适当调整 λ , 可保证下降条件成立, 从而迭代序列单调趋向于根。 \square

例 5.1. 用 Newton 下山法求 $f(x) = x^3 - x - 1 = 0$ 在 $x = 1.5$ 附近的根。取 $x_0 = 1.5$, $f'(x) = 3x^2 - 1$ 。

- 第一步: $\lambda = 1$, $x_1 = 1.5 - \frac{1-1.25}{5.75} \approx 1.3043$, $|f(x_1)| \approx 0.3704 < |f(x_0)| = 0.875$, 接受并尝试增大 λ 。
- 第二步: 取 $\lambda = 2$, $x_2 = 1.3043 - 2 \cdot \frac{0.3704}{4.104} \approx 1.1249$, $|f(x_2)| \approx 0.2493 < 0.3704$, 继续。
- 后续步骤调整 λ 使 $|f(x_n)|$ 持续下降, 直至满足精度要求。

5.4 非线性方程组

 注. 这里不能考吧。。。