

科学计算笔记

任云玮

目录

1	绪论	3
1.1	计算机数值计算基本原理	3
1.1.1	实数的存贮方法	3
1.1.2	实数的基本运算原理	4
1.2	误差的来源与估计	5
1.2.1	误差的来源	5
1.2.2	误差与有效数字	5
1.2.3	数值运算的误差估计	6
1.2.4	数字求和的舍入误差分析	7
1.3	避免算法失效的基本原则	8
2	函数的多项式插值	10
2.1	问题的提出	10
2.2	Lagrange 插值法	11
2.3	Runge 现象	12
2.4	Newton 插值法	12
2.5	Hermite 插值	14
2.6	分段低次多项式插值	15
2.7	三次样条插值	16
3	函数的多项式逼近	17
3.1	绪论	17
3.2	最佳平方逼近	18
3.3	正交多项式 · 绪论	20
3.4	Legendre 多项式	22
3.5	Chebyshev 多项式	23

4	变分方法与数据拟合	25
4.1	绪论	25
4.2	变分方法	26
4.3	曲线拟合的正则化方法	29
5	数值积分与数值微分	32
5.1	绪论	32
5.2	Newton-Cotes 公式	36
5.3	复合求积公式	36
5.4	Gauss 求积公式	37
5.5	Romberg 求积公式	38
5.6	自适应求积公式	39
5.7	数值微分	40
6	非线性方程求根	42
6.1	二分法	42
6.2	不动点法	42
6.3	Newton 法	45
7	解线性方程组的直接/间接方法	48
7.1	绪论	48
8	附录	50
8.1	不等式	50
8.2	积分相关公式	51
8.3	特殊函数	52
8.4	Euler-Maclaurin 公式	53
8.4.1	广义 Euler 常数	53
8.4.2	Euler 求和公式	54
8.4.3	余项分析	55
8.4.4	Bernoulli 数和 Euler 求和公式的一般形式	56

1 绪论

1.1 计算机数值计算基本原理

1.1.1 实数的存贮方法

1 定义 (二进制浮点数系)¹ 实数在计算机内部为近似存贮, 采用二进制浮点数系

$$F(2, n, L, U) = \{\pm 0.a_1a_2 \dots a_n \times 10^m\} \cup \{0\}$$

其中 $a_1 = 1$, $a_i \in \{0, 1\}$. 指数 m 满足 $L \leq m \leq U$. 称 n 为其字长, 2 表示采用二进制.

2 标准 (IEEE)

1. 单精度: $t = 24, L = -126, U = 127$

2. 双精度: $t = 53, L = -1022, U = 1023$

3. Underflow Limit: $UFL = 0.1 \times 2^L$. 若 $0 < x < UFL$, 则 $fl(x) = 0$.

4. Overflow Limit: $OFL = 0.11 \dots 1 * 2^U$. 若 $x > OFL$, 则 $fl(x) = \infty$.

5. 舍入: 若 $UFL \leq x \leq OFL$, 则 $fl(x)$ 为舍入所得浮点数. 舍入规则如下: 设 $x = 0.a_1a_2 \dots a_n \dots \times 2^m$. 若 $a_{n+1} = 1$, 则 $d_t + 1$ 并舍弃其后项; 否则直接舍弃其后项.

3 定义 (机器精度) 下仅考虑舍去的情况.

$$\begin{aligned} x - fl(x) &= 2^m \times 0.0 \dots 0a_{n+2} \dots \\ &= 2^m \times [2^{-(t+2)} + 2^{-(t+3)} + \dots] \\ &= 2^m \times 2^{-(t+1)} \end{aligned}$$

其相对误差满足

$$\frac{x - fl(x)}{x} < \frac{x - fl(x)}{0.5 \times 2^m} = 2^{-t}$$

记为 ε , 称之为机器精度.

4 命题

$$fl(x) = x(1 + \delta), \text{ 其中 } |\delta| \leq \varepsilon$$

¹floating Number System

1.1.2 实数的基本运算原理

加法 + 硬件实现 \Rightarrow 四则运算.

5 实现 $(x + y)$ 设 x, y 为浮点数, 则 $x + y$ 的实现方式如下:

1. 对阶: 将指数 m 化为两者中较大者;
2. 尾数相加;
3. 舍入;
4. 溢出分析等……
5. 结果输出.

评注 由 $fl(x) + fl(y) = x(1 + \delta_x) + y(1 + \delta_y)$ 可知, 当一个大数与一个小数相加时, 小数有可能被忽略, 所以应当避免大数小数间的相加.

1.2 误差的来源与估计

1.2.1 误差的来源

1. 模型问题. 例: 近似地球为球体来计算.
2. 测量误差. 例: 测量地球半径时的误差.
3. 方法误差 (截断误差). 例: 对于 $y = f(x)$, 求 $f(x^*)$ 时使用 Taylor 展开.
4. 舍入误差 (rounding-off). 例: 计算机计算时的误差.

1.2.2 误差与有效数字

6 定义 (绝对误差) 设 x 为给定实数, x^* 为其近似值. 定义绝对误差为

$$e(x^*) = x^* - x.$$

称 ε^* 为其误差上界, 若

$$|e(x^*)| \leq \varepsilon^*$$

7 定义 (相对误差) 对于同上的 x 和 x^* , 定义其相对误差

$$e_r(x^*) = \frac{x^* - x}{x}$$

称 ε_r^* 为其相对误差界, 若

$$|e_r(x^*)| \leq \varepsilon_r^*$$

评注 在实际应用中, x 通常是未知的, 所以会采用

$$\bar{e}_r(x^*) = \frac{x^* - x}{x^*}$$

来代替相对误差. 对于分子, 使用绝对误差界来替代, 有如下不等式

$$|\bar{e}_r(x^*)| \leq \frac{\varepsilon^*}{|x^*|}.$$

这两种相对误差界间的差别, 当 $\varepsilon^* \ll 1$ 时, 满足

$$|e_r - \bar{e}_r| = O((\varepsilon_r^*)^2)$$

8 定义 (有效数字) 设 $x \in R$, $x^* = 0.a_1a_2 \cdots a_k \times 10^m$ 为其近似值. 称 x^* 相对于 x 有 n ($n \leq k$) 位有效数字, 若 n 是满足下式的 n 的最大值.

$$|x^* - x| \leq \frac{1}{2} \times 10^{m-n}$$

评注 在实践中, 一般可以采用更加简便的方法, 对于归一化以后的 x^* , 在尾数部分有 n 位, 则称其有 n 位有效数字. 注意, 此方法对于错误的舍入结果是不适用的, 对于错误的情况, 需要再减去一位有效数字.

9 定理 (误差与有效数字) 若 $x = 0.a_1a_2 \dots a_n \times 10^m$ 有 n 位有效数字, 则

$$\varepsilon_r^* \leq \frac{1}{2a_1} \times 10^{1-n}.$$

反之, 若

$$\varepsilon_r^* \leq \frac{1}{2(1+a_1)} \times 10^{1-n},$$

则 x^* 至少有 n 位有效数字.

证明 对于前者, 只需利用有效数字的定义, 以及利用 $x \geq 0.a_1$ (仅考虑 $a_1 \neq 0$ 的情况). 对于后者, 证明是类似的.

1.2.3 数值运算的误差估计

以下内容都假设运算无误差.

10 定理 (四则运算误差估计)

1. 加/减法: $\varepsilon(x^* \pm y^*) \leq \varepsilon_x^* + \varepsilon_y^*$
2. 乘法: $\varepsilon(x^* y^*) \leq |x^*| \varepsilon_y^* + |y^*| \varepsilon_x^*$
3. 除法: $\varepsilon\left(\frac{x^*}{y^*}\right) \leq \frac{|x^*| \varepsilon_y^* + |y^*| \varepsilon_x^*}{|y^*|^2}$

证明 考虑加法的误差估计. 对于 x, y 及其近似值 x^*, y^* , 计算 $x^* \pm y^*$ 和 $x \pm y$ 间的误差.

$$\begin{aligned} |x^* \pm y^* - (x \pm y)| &\leq |x^* - x| + |y^* - y| \leq \varepsilon_x^* + \varepsilon_y^* \\ \Rightarrow \varepsilon(x^* \pm y^*) &\leq \varepsilon_x^* + \varepsilon_y^* \end{aligned}$$

对于其他的运算, 证明是类似的. (证明中可用 $+1-1$ 技巧)

11 定理 (运算的误差估计) 设 $A = f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$, \mathbf{x}^* 是 \mathbf{x} 的估计值. 利用带 Peano 余项的 Taylor 展开, 可知 A 的绝对误差满足

$$\begin{aligned} e(A^*) &= f(\mathbf{x}^*) - f(\mathbf{x}) \\ &= \sum_{p=1}^q d^p f(\mathbf{x}^*) + o(\|\mathbf{x}^* - \mathbf{x}\|^q) \\ \text{取 } q=1, \text{ 则} \\ &= \sum_{k=1}^n \partial_k f(\mathbf{x}^*)(x_k^* - x_k) + o(\|\mathbf{x}^* - \mathbf{x}\|) \end{aligned}$$

利用上式, 可知

$$\begin{aligned}\varepsilon(A^*) &\approx \sum_{k=1}^n |\partial_k f(\mathbf{x}^*)| \varepsilon(x^*) \\ \varepsilon_r(A^*) &= \frac{\varepsilon(A^*)}{|A^*|}\end{aligned}$$

评注 对于定义在 \mathbf{R} 上的函数, 即为

$$\varepsilon(f(x^*)) \approx |f'(x^*)| \varepsilon(x^*)$$

1.2.4 数字求和的舍入误差分析

12 命题 n 个浮点数相加, 若将它们从小到大排列后相加, 则可以减小舍入误差.

证明 考虑浮点数的求和 $S_n = \sum_i^n a_i$, 在计算机中的过程表现为

$$\begin{aligned}S_2^* &= fl(a_1 + a_2) = (a_1 + a_2)(1 + \varepsilon_2), \quad |\varepsilon_2| \leq \varepsilon = 2^{-t} \\ &\dots\dots\dots \\ S_n^* &= fl(S_{n-1}^* + a_n)(1 + \varepsilon_n), \quad |\varepsilon_n| \leq \varepsilon\end{aligned}$$

对于 S_n^* 的误差, 若定义 $\varepsilon_1 = 0$, 则

$$S_n^* = \sum_{k=1}^n a_k \prod_{p=k}^n (1 + \varepsilon_p)$$

对误差进行估计, 舍去高阶无穷小, 有

$$\prod_{i=k}^n (1 + \varepsilon_k) \approx 1 + \sum_{i=k}^n \varepsilon_k$$

综合上两式, 有

$$\begin{aligned}S_n^* &\approx \sum_{k=1}^n a_k (1 + \sum_{p=k}^n \varepsilon_p) \\ &= S_n + \sum_{k=1}^n a_k \sum_{p=k}^n \varepsilon_p\end{aligned}$$

进行移项, 并取绝对值, 再利用三角不等式, 以及 $|\varepsilon_i| \leq \varepsilon$, 得

$$|S_n^* - S_n| \leq \sum_{k=1}^n |a_k| \sum_{p=k}^n |\varepsilon_p| \leq \varepsilon \sum_{k=1}^n |a_k| (n - k + 1)$$

其中 $n - k + 1$ 关于 k 单调减少, 所以根据排序不等式 [引理 135], 即可知命题成立. ■

1.3 避免算法失效的基本原则

13 定理 (原则)

1. 避免两数相除/相减, 否则会严重损失有效数字.
2. 避免两相近数相减.
3. 避免绝对值很小的数做除数.
4. 避免大数与小数相加;
5. 简化计算步骤.

14 算法 (高效计算 e^A) 高效计算 e^A , 其中 $A \in \mathbf{R}^{n \times n}$. 首先有

$$e^A = e^{(A/2^n)2^n} = B^{2^n}$$

只需要得到 B , 即可以利用倍乘的方法快速得到 B^{2^n} . 下对于 B 进行估计. 当 $x \rightarrow 0$ 时, e^x 有 Taylor 展开

$$e^x = 1 + x + \cdots + \frac{x^n}{n!} + \cdots$$

而取足够大的 n , 即可以使得 $A/2^n \approx 0$, 则可以对它展开得

$$B \approx I + C + \frac{1}{2}C^2, \text{ 其中 } C = A/2^n$$

而对于倍乘, 考虑 B^2 , 展开平方得

$$B^2 \approx I + 2(C + \frac{1}{2}C^2) + (C + \frac{1}{2}C^2)^2$$

从右至左相加即可.

15 算法 (秦九韶, 多项式估值) 设有多项式 (1), 计算 $p(z), z \in \mathbf{R}$ 的值.

$$p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \quad (1)$$

定义 b_n 满足

$$b_0 = a_0, \quad b_k = a_k + b_{k-1}z$$

则 b_n 即为所要求的值. 并且成立

$$p'(z) = \sum_{k=0}^{n-1} b_k z^{n-1-k}$$

证明 用 $x - z$ 去除 $p(x)$, 记所得余数为 $b_n(x)$, 即

$$p(x) = (x - z)q(x) + b_n(x),$$

代入 $x = z$, 则左侧第一项为 0, 可知 $p(z) = b_n(z)$. 将两边的式子展开, 利用对应系数相等, 即可得算法中 b_n 的递推式.

16 定理 (外推法) 设 x_0, x_1 是 x 的两个估计值, 且 x_1 相较于 x_0 更接近 x , 则可以通过恰当的权值 ω , 使得它们的加权平均

$$\bar{x} = x_1 + \omega(x_1 - x_0)$$

更加接近精确值 x .

17 算法 (π 的估计) 考虑单位圆, 其面积为 π , 设 π_n 为单位圆的内接正 $2n$ 边形的面积, 以及

$$\tilde{\pi}_n = \frac{1}{3}(4\pi_{2n} - \pi_n)$$

则 π_n 与 $\tilde{\pi}_n$ 与 π 的误差满足

$$|\pi_n - \pi| = O\left(\frac{1}{n^2}\right), \quad |\tilde{\pi}_n - \pi| = O\left(\frac{1}{n^4}\right)$$

证明 对于 π_n .

$$\pi_n = n \sin \frac{\pi}{n} = \pi - \frac{\pi^3}{3!} \frac{1}{n^2} + \frac{\pi^5}{5!} \frac{1}{n^4} - \cdots \Rightarrow |\pi_n - \pi| = O\left(\frac{1}{n^2}\right)$$

对于 $\tilde{\pi}_n$.

$$\begin{aligned} \tilde{\pi}_n &= \pi_{2n} + k(\pi_{2n} - \pi_n) = (1+k)\pi_{2n} - k\pi_n \\ &= (1+k)\left(\pi - \frac{\pi^3}{3!} \frac{1}{4n^2} + \cdots\right) - k\left(\pi - \frac{\pi^3}{3!} \frac{1}{n^2} + \cdots\right) \\ &= \pi - \left(\frac{k+1}{4} - k\right) \frac{\pi^3}{3!} \frac{1}{n^2} + O\left(\frac{1}{n^4}\right) \end{aligned}$$

为使式子的第二项为零, 取 $k = \frac{1}{3}$, 则成立

$$|\tilde{\pi}_n - \pi| = O\left(\frac{1}{n^4}\right) \quad \blacksquare$$

评注 在实际中, π_n 也是没有办法直接计算而得的, 但是对于 $n = 3$, 即 6 边形的情况, 可以知道 $\pi_3 = 3\sqrt{3}/2$. 同时有递推公式

$$\pi_{2n} = \sqrt{2n(n - \sqrt{n^2 - \pi_n^2})},$$

而开平方可以通过迭代的方式实现, 从而即计算得到足够精确的 π_{2n} 和 π_n .

2 函数的多项式插值

2.1 问题的提出

18 定义 (插值) 设函数 $y = f(x)$ 在 $[a, b]$ 上有定义, 且已知在点 $a \leq x_0 < x_1 < \cdots < x_n \leq b$ 处的函数值 $y_i = f(x_i)$, 若存在一简单函数 $P(x)$, 成立

$$P(x_i) = y_i,$$

则称 $P(x)$ 为 $f(x)$ 的插值函数, 点 x_1, x_2, \dots, x_n 称为插值节点, $[a, b]$ 称为插值区间, 求 $P(x)$ 的方法被称为插值法.

若 $P(x) \in P_n$ 为次数不超过 n 的多项式, 则称为多项式插值.

19 定理 (唯一性) 给定满足定义 18 的 $n + 1$ 个点上的函数值, 则次数不超过 n 的插值多项式 $P_n(x)$ 存在且唯一.

证明 利用待定系数法, 设多项式的系数为 a_0, \dots, a_n , 则有线性方程组

$$\begin{cases} a_0 + a_1x_0 + \cdots + a_nx_0^n = y_0, \\ a_0 + a_1x_1 + \cdots + a_nx_1^n = y_1, \\ \quad \quad \quad \dots\dots\dots \\ a_0 + a_1x_n + \cdots + a_nx_n^n = y_n, \end{cases}$$

其系数矩阵为 Vandermonde 矩阵

$$A = \begin{pmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & & \vdots \\ 1 & x_n & \cdots & x_n^n \end{pmatrix}$$

根据定义 18 中对于 x_i 的要求, 矩阵行列式成立

$$\det A = \prod_{i,j=0, i>j}^n (x_i - x_j) \neq 0.$$

所以该方程组有唯一解.

评注 虽然插值多项式是唯一的, 但是根据基函数的选取的不同, 系数是不相同的, 所以才需要不同的插值方法.

2.2 Lagrange 插值法

20 定理 (Lagrange 插值法) 定义

$$l_i(x) = \frac{\prod_{j \neq i}(x - x_j)}{\prod_{j \neq i}(x_i - x_j)}$$
$$(L_n f)(x) = \sum_{i=0}^n y_i l_i(x),$$

则 $L_n f$ 即为 f 的插值多项式.

证明 考虑构造 $l_i \in P_n$, 满足条件 $l_i(x_j) = \delta_{ij}$, 这样 $L_n f = \sum y_i l_i$ 满足要求. 改写条件为 (以 l_0 为例)

$$l_0(x) = \alpha(x - x_1) \cdots (x - x_n)$$
$$l_0(x_0) = 1$$

解得

$$\alpha = \frac{1}{(x_0 - x_1) \cdots (x_0 - x_n)} \quad \blacksquare$$

评注 这样构造插值多项式的动机在于在取定插值节点后, 插值实际上相当于构造一个从 $\mathbf{y} = (y_0, \dots, y_n) \in \mathbf{R}^{n+1}$ 到 $y^*(x) \in P_n$ 的一个映射 \mathcal{F} , 并且可以证明, \mathcal{F} 是线性的. 因此成立

$$\mathcal{F}(\mathbf{y}) = \mathcal{F}\left(\sum_{i=0}^n y_i \mathbf{e}_i\right) = \sum_{i=0}^n y_i \mathcal{F}(\mathbf{e}_i) = \sum_{i=0}^n y_i l_i(x).$$

21 定理 (Lagrange 余项公式) 设符号含义同定理 20 且 f 充分光滑, 则对于每一个固定的 x 成立

$$f(x) - (L_n f)(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x)$$

其中 $\xi \in (a, b)$ 且

$$\omega_{n+1}(x) = (x - x_0) \cdots (x - x_n).$$

证明 固定 $x \neq x_i$, 定义 $R(x)$ 满足

$$f(x) - (L_n f)(x) = R(x) \omega_{n+1}(x).$$

构造辅助函数 $g(t)$

$$g(t) = f(t) - (L_n f)(t) - R(x) \omega_{n+1}(t).$$

根据插值法与 $R(x)$ 的定义, 成立

$$g(x_i) = 0, \quad g(x) = 0,$$

即函数 $g(t)$ 有 $n+2$ 个零点. 反复应用 Rolle 定理, 可知存在 $\xi \in (a, b)$, 成立 $g^{(n+1)} = 0$, 即

$$\begin{aligned} g^{(n+1)}(\xi) &= f^{(n+1)}(\xi) - R(x)(n+1)! = 0 \\ \Rightarrow R(x) &= \frac{f^{(n+1)}(\xi)}{(n+1)!} \end{aligned}$$

结合 $R(x)$ 的定义式即可知命题成立. ■

评注 当已知 $f^{(n+1)}$ 有界时, 可以使用此公式进行估计.

2.3 Runge 现象

22 定理 对于复函数 $f(z)$, 如果存在 $r_0 > \frac{3}{2}(b-a)$, 使得 $f(z)$ 在 $B_{r_0}(\frac{a+b}{2})$ 内解析, 则 $P_n(x) = L_n(x)$ 在 $[a, b]$ 上一致收敛与 $f(z)$. 这里 $B_{r_0}(\frac{a+b}{2})$ 为以 $\frac{a+b}{2}$ 为圆心, r_0 为半径的圆.

2.4 Newton 插值法

23 例 $n=2$ 时问题的求解.

设 $y^*(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1)$, 根据 $y^*(x_0) = y_0$, $y^*(x_1) = y_0$, 得 $a_0 = f(x_0)$, $a_0 + a_1(x_1-x_0) = f(x_1)$, 即

$$a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

可以发现 a_1 为割线的斜率. 同理可知

$$a_2 = \frac{f(x_2) - f(x_0) - a_1(x_2 - x_0)}{(x_2 - x_0)(x_2 - x_1)} = \frac{\frac{f(x_2) - f(x_0)}{x_2 - x_0} - \frac{f(x_1) - f(x_0)}{x_1 - x_0}}{x_2 - x_1}$$

为斜率的斜率.

24 定义 (差商) 递归 $f(x)$ 在 x_i, \dots, x_{i+n} 的各阶差商为: $f[x_i] = f(x_i)$, 第 k 阶差商为

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

25 定理 (n 次 Newton 插值法) x_0, \dots, x_n 为互异插值点, 则函数 $f(x)$ 满足

$$\begin{aligned} f(x) &= f[x_0] + f[x_0, x_1](x-x_0) + f[x_0, x_1, x_2](x-x_0)(x-x_1) + \dots \\ &\quad + f[x_0, \dots, x_n](x-x_0)\cdots(x-x_{n-1}) + R_n(x). \end{aligned}$$

其中 $R_n(x)$ 为其 Newton 插值多项式的余项, 为

$$R_n(x) = f[x_0, \dots, x_n, x](x-x_0)\cdots(x-x_n).$$

证明 根据差商的定义, 有

$$\begin{aligned} f(x) &= f[x_0] + f[x_0, x](x - x_0) \\ f[x_0, x] &= f[x_0, x_1] + f[x_0, x_1, x](x - x_1) \\ &\dots\dots\dots \end{aligned}$$

$$f[x_0, \dots, x_{n-1}, x] = f[x_0, \dots, x_n] + f[x_0, \dots, x_n, x](x - x_n)$$

将上述式子反复代入它上面的式子, 即得 Newton 插值公式.

评注 Newton 插值法的优点在于, 当插值点的个数增加时, 无需重新计算原有的系数, 即 Newton 插值多项式是可以递归计算的.

26 定理 根据 Newton 插值公式, 可以得到如下差商的性质.

1.

$$f[x_0, \dots, x_m] = \sum_{i=0}^m \frac{f(x_i)}{\prod_{j \neq i} (x_i - x_j)}$$

2. 设 i_0, \dots, i_m 为 $0, \dots, m$ 的任意一个排列, 则

$$f[x_0, \dots, x_m] = f[x_{i_0}, \dots, x_{i_m}].$$

3. 广义 Lagrange 中值定理

$$f[x_0, \dots, x_m] = \frac{f^{(m)}(\xi)}{m!}, \quad \xi \in (\min\{x_i\}, \max\{x_i\})$$

证明 交换插值节点的顺序后, n 次 Newton 插值多项式的 n 次项系数不变, 所以 [2.] 成立. 同 $m-1$ 次的 Lagrange 插值多项式比较第 $m-1$ 次项系数及其余项, 即可得到 [1.] 和 [3.]

评注 根据 [3.] 可知, 对于一个 n 次多项式, n 阶差商即为其 n 次项系数, $k(k > n)$ 阶差商为零.

27 定义 给定序列 $\{f_k\}$, f_k 表示 f 在 $x = x_k$ 处的值, 定义

1. 前向差分算符 $\Delta f_k = f_{k+1} - f_k$,

2. 移位算符 $E f_k = f_{k+1}$,

3. 恒等算符 $I f_k = f_k$.

28 定理 (算符二项式定理) 对于算符 A, B , 若它们可交换, 则成立二项式定理

$$(A + B)^m = \sum_{k=0}^m \binom{m}{k} A^k B^{m-k}$$

29 命题 根据定义 27 和定理 28 可知

$$\begin{aligned}\Delta &= E - I \\ \Delta^n &= \sum_{k=0}^n \binom{n}{k} (-I)^{n-k} E^k \\ E^n &= \sum_{k=0}^n \binom{n}{k} \Delta^k\end{aligned}$$

30 定理 (均匀插值) 设 x_0, x_1, \dots, x_n 满足 $x_k = x_0 + kh$, 则有

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

31 定理 (Newton 前插公式) 设记号同定理 30, 另 $x = x_0 + th$, $t \in \mathbf{R}$, 代入定理 25, 则成立

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

其余项为

$$R_n(x) = \frac{t(t-1)\dots(t-n)}{(n+1)!}h^{n+1}f^{(n+1)}(\xi), \quad \xi \in (x_0, x_n)$$

评注 利用广义二项式定理, 有

$$\begin{aligned}f(x) &= f(x_0 + th) = E^t f(x_0) = (I + \Delta)^t f(x_0) \\ &= \sum_{k=0}^{\infty} \binom{t}{k} \Delta^k f_0\end{aligned}$$

2.5 Hermite 插值

32 定理 设 $f \in C^n[a, b]$, x_0, x_1, \dots, x_n 为 $[a, b]$ 上的互异节点, 则 $f[x_0, \dots, x_n, x]$ 在 $[a, b]$ 上连续.

33 定理 (Hermite 插值) 若给出 $m+1$ 个插值条件 (含函数值和导数值) 可构造出次数不超过 m 次的多项式.

评注 可以利用待定系数法或者基函数法求的 Hermite 插值多项式.

2.6 分段低次多项式插值

思路 局部入手，整体分析². 化整为零，以直代曲³.

34 引理 (ω_n 的估计) 任给节点 $x_0 < x_1 < \cdots < x_n$, 记 $h = \max x_{i+1} - x_i : i = 0, 1, \dots, n$, 则对于任意 $x \in [x_0, x_n]$, 成立

$$|(x - x_0) \cdots (x - x_n)| \leq \frac{n!h^{n+1}}{4}.$$

35 定理 (分段线性插值) 记 $a = x_0 < \cdots < x_n = b$, $e_k = (x_k, x_{k+1})$, $h_k = x_{k+1} - x_k$, $h = \max h_k$. 找函数 $y = f_h(x)$, 逼近原有函数, 使得

1. 满足插值条件, $f_h(x_k) = f(x_k)$,
2. $f_h(x)$ 连续,
3. $f_h(x) \in P_1$, $x \in e_k$.

则 f_h 的结果为

$$f_h(x) = f_h(x_k) + f_h[x_k, x_{k+1}](x - x_k), \quad (x \in e_k).$$

设 M_2 表示二阶导数的上界, 则误差 $R(x)$ 满足,

$$R(x) = \left| \frac{f^{(2)}(\xi)}{2} (x - x_k)(x - x_{k+1}) \right| \leq \frac{1}{8} M_2 h^2.$$

评注 若作整体的 Lagrange 插值, 则余项 $R_L(x)$ 满足

$$\left| \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1} \right| \leq \frac{M_{n+1} h^{n+1}}{n+1}.$$

而求高阶导数后容易出现 Runge 现象.

36 定理 (分段三次 Hermite 插值) 节点同上, 构造 $f_h(x)$ 使得

1. $f_h(x_k) = f(x_k)$, $f'_h(x_k) = f'(x_k)$
2. $f \in \mathcal{C}^1[a, b]$
3. $f_h(x) \in P_3$, $x \in e_k$

对于两插值点的情况, f 的结果为

$$f_h(x) = f(x_k)\alpha_k + f(x_{k+1})\alpha_{k+1} + f'(x_k)\beta_k + f'(x_{k+1})\beta_{k+1}.$$

其余项 $R(x)$ 满足

$$R(x) = \left| \frac{f^{(4)}(\xi)}{4} (x - x_k)^2 (x - x_{k+1})^2 \right| \leq \frac{M_4}{4!} \times \left(\frac{h}{4} \right)^4 = \frac{1}{384} M_4 h^4.$$

²例: 微分流形

³例: Riemann 积分

2.7 三次样条插值

37 定义 给定控制点 $a = x_0 < \cdots < x_n = b$, 设函数 $y = y^*(x)$ 满足

1. $y^*(x) = y_k$,
2. $(y^*)^{(4)} = 0, x \in e_k \Leftrightarrow y^* \in P_3, x \in e_k$,
3. $y^* \in \mathcal{C}^2[a, b]$.

称满足后两个条件的函数为**三次样条函数**, 称满足上述三个条件的函数为**三次样条插值函数**.

38 定义 (边界条件)

1. 转角条件: $S'(x_0) = f'_0, S'(x_n) = f'_n$,
2. 弯矩条件: $S''(x_0) = f''_0, S''(x_n) = f''_n$, 称 $S''(x_0) = S''(x_n) = 0$ 的特例为自然边界条件,
3. 周期条件: $S(x_0+0) = S(x_n-0), S'(x_0+0) = S'(x_n-0), S''(x_0+0) = S''(x_n-0)$,
4. 非纽结条件: $S'''(x)$ 在 $x = x_1$ 和 $x = x_{n-1}$ 处连续.⁴

评注 根据定义 37, 在每个小区间上有 4 个待定系数, 所以总共有 $4n$ 个待定系数. 而所给条件仅有 $n+1$ 个插值条件, 以及在中间 $n-1$ 个插值节点处二阶导数连续 (从而原函数与一阶导数也连续), 有 $3n-3$ 个光滑性条件, 共 $4n-2$ 个条件, 因此需要额外的边界条件来确定剩余两个系数.

39 定理 (样条插值的求解) 设 $S''(x_i) = M_i$, 通过求解 M_i 来确定插值多项式. 由于 $S(x) \in \mathcal{C}^2$ 且在每一段上 $S(x) \in P_3$, 所以 $S''(x)$ 是分段的线性函数. 设在每一段上,

$$S''(x) = M_j \frac{x_{j+1} - x}{x_{j+1} - x_j} + M_{j+1} \frac{x - x_j}{x_{j+1} - x_j}.$$

对 S'' 积分一次、二次, 并分别利用光滑性条件、插值条件, 以及所给定的边界条件, 即可求得 M_i 的值.

40 定理 设 $f(x) \in \mathcal{C}^4[a, b]$, 则三次样条插值函数 $S_3(x)$,

$$\max_{a \leq x \leq b} |f^{(m)}(x) - S^{(m)}(x)| \leq C_m \max_{a \leq x \leq b} |f^{(4)}(x)| h^{4-m}, \quad m = 0, 1, 2,$$

其中 $C_0 = 5/384, C_1 = 1/24, C_2 = 3/8$.

⁴ (Not-a-knot end Condition) 这是 Matlab 中 spline 在 X 和 Y 长度相同时所应用的边界条件.

3 函数的多项式逼近

3.1 绪论

41 **定义 (逼近)** 对函数 f 逼近, 即找一简单函数 g , 使得在某种度量的意义下, 它们之间的误差最小或足够小.

42 **定理 (Weierstrass)** 对于定义在 $[a, b]$ 上的连续复函数, 存在一系列复多项式 $\{P_n\}$, 成立

$$\lim_{n \rightarrow \infty} P_n = f,$$

且是一致的. 若 f 是实函数, 则 P_n 的系数也为实数.

评注 Stone-Weierstrass 定理⁵ 保证了至少在最大模的意义下, 用多项式来逼近函数是可能的.

43 **定义 (常用范数)** 对于 \mathbf{R}^n , 常用的范数有

1. $\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|,$

2. $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|,$

3. $\|\mathbf{x}\|_2 = (\sum_{i=1}^n x_i^2)^{1/2}.$

对于 $\mathcal{C}[a, b]$, 常用的范数有

1. $\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|,$

2. $\|f\|_1 = \int_a^b |f(x)| dx,$

3. $\|f\|_2 = (\int_a^b f^2(x) dx)^{1/2}.$

评注 通常对于内积空间 X , 可以定义范数为

$$\|\mathbf{x}\| = \sqrt{(x, x)}.$$

44 **定义 (权函数)** 设 $[a, b]$ 为有限或无限区间⁶, 非负函数 $\rho(x)$ 称为 $[a, b]$ 上的权函数, 若满足

1. $\int_a^b \rho(x) x^k dx < \infty, k = 0, 1, 2, \dots,$

2. 对任意非负 $g \in \mathcal{C}[a, b]$, 若 $\int_a^b \rho(x) g(x) dx = 0$, 则 $g = 0$.

评注 利用权函数, 可以定义带权内积和范数.

⁵ **Theorem(Stone)** Suppose \mathcal{A} is a self-adjoint algebra of complex continuous functions on a compact set K , \mathcal{A} separates points on K , and \mathcal{A} vanishes at no point of K . Then the uniform closure \mathcal{B} of \mathcal{A} consists of all complex continuous functions on K . In other words, \mathcal{A} is dense in $\mathcal{C}(K)$.

⁶例子中说 $\rho = 1$ 是一个常用的权函数, 但我没有明白, 在无限区间的时候 [1.] 是如何成立的.

3.2 最佳平方逼近

45 定义 (最佳平方逼近) 给定 $f \in \mathcal{C}[a, b]$ 和线性无关的函数列 $\varphi_0, \dots, \varphi_n \in \mathcal{C}[a, b]$, 定义 $S_n = \text{span}\{\varphi_0, \dots, \varphi_n\}$, 称 $f^* \in S_n$ 为最佳平方逼近函数, 若

$$\|f^* - f\| = \min_{g \in S_n} \|f - g\|_2.$$

即 f^* 是在 2-范数的含义下, S_n 中与 f 最接近的函数.

评注 对于离散的情况⁷, 可以描述为: 给定 x_0, \dots, x_n 处的函数值 $f(x_k)$, 求 f^* , 成立

$$\sum_{i=0}^n \rho(x_i) |f(x_j) - f^*(x_j)|^2 = \min_{g \in S_n} \sum_{i=0}^n \rho(x_i) |f(x_j) - g(x_j)|^2$$

46 定理 (最佳平方逼近的求解) 设记号同定义 45, 设

$$g(x) = \sum_{i=0}^n a_i \varphi_i(x),$$

则可以定义关于 $\mathbf{a} = (a_0, \dots, a_n)^\top$ 的函数

$$I(\mathbf{a}) = \|f - g\|_2^2 = \left\| f - \sum_{i=0}^n a_i \varphi_i \right\|_2^2 = \int_a^b \rho \left(f - \sum_{i=0}^n a_i \varphi_i \right)^2 dx.$$

根据定义, $I(\mathbf{a})$ 在 f^* 处取极值, 根据 Fermat 定理, 在该点各偏导数为零, 通常假设 f 的条件足够好, 极限和积分可以换序, 即有

$$\begin{aligned} \frac{\partial I}{\partial a_j} &= \int_a^b \frac{\partial}{\partial a_i} \rho \left(f - \sum_{i=0}^n a_i \varphi_i \right)^2 dx \\ &= -2 \int_a^b \rho \left(f - \sum_{i=0}^n a_i \varphi_i \right) \varphi_j dx = 0. \end{aligned}$$

即有线性方程组,

$$\begin{cases} (\varphi_0, \varphi_0)a_0 + (\varphi_0, \varphi_1)a_1 + \dots + (\varphi_0, \varphi_n)a_n = (f, \varphi_0) \\ (\varphi_1, \varphi_0)a_0 + (\varphi_1, \varphi_1)a_1 + \dots + (\varphi_1, \varphi_n)a_n = (f, \varphi_1) \\ \dots\dots\dots \\ (\varphi_n, \varphi_0)a_0 + (\varphi_n, \varphi_1)a_1 + \dots + (\varphi_n, \varphi_n)a_n = (f, \varphi_n) \end{cases} \quad (2)$$

⁷ 实际上我们可以利用 Riemann-Stieltjes 积分定义内积,

$$\begin{aligned} (f, g) &= \int_a^b f dG, \\ G(x) &= \begin{cases} \int_a^b g dx, & g \text{ 为函数,} \\ \sum_{i=0}^n f(x_i) I(x - x_i), & g \text{ 为离散点} \end{cases} \end{aligned}$$

其中 $I(x)$ 为单位阶跃函数. 可以发现, 这两种描述的方式是等价的. 在这样的描述下, 对于离散点的 G 实际上是阶梯函数.

由于 $\{\varphi_k\}$ 线性无关, 所以方程组 (2) 有唯一解. 设其解为 \mathbf{a}^* , 则最佳平方逼近函数即为

$$f^* = \sum_{i=0}^n a_i^* \varphi_i(x),$$

评注 实际上在计算的时候一般采用 Legendre 多项式来计算, 而非解法方程. 见定理 55.

几何描述 可以从几何的角度来理解最佳平方逼近. S_n 是 $\{\varphi_k\}$ 张成的空间, 而 f 是 S_n 内或 S_n 外的一个向量, 最佳平方逼近即找 S_n 中找 f^* , 使得 $\|f - f^*\|$ 最小. 根据几何上的直观, $f - f^*$ 应该和 S_n “垂直”, 即与张成 S_n 的向量组中的向量分别垂直. 而垂直可以被描述为内积为零. 从而就得到了式 (2). (见图 1)

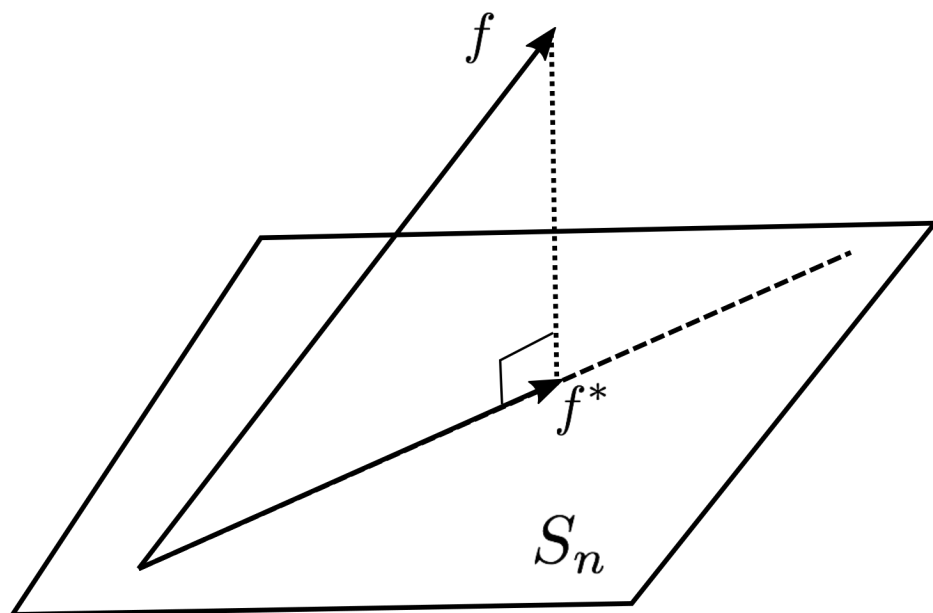


图 1: 最佳平方逼近几何含义

3.3 正交多项式 · 绪论

47 定义 (正交) 设函数 $f, g \in \mathcal{C}[a, b]$, ρ 为 $[a, b]$ 上的权函数且满足

$$(f, g) = \int_a^b \rho f g dx = 0,$$

则称 f 和 g 在 $[a, b]$ 上带权 ρ 正交. 若函数组 $\{\varphi_k\}_{k=0}^\infty$ 满足

$$(\varphi_i, \varphi_j) = \begin{cases} 0, & i \neq j \\ A_k > 0, & i = j \end{cases}$$

则称 $\{\varphi_k\}$ 为 $[a, b]$ 上的带权 ρ 的正交函数组. 若 $A_k = 1$, 则称为标准正交函数组.

48 定义 (正交多项式) 设 $\{\varphi_k\}_{k=0}^\infty$ 是首项系数 $a_n \neq 0$ 的 n 次多项式序列. 若它们正交, 则称它们为正交多项式序列.

49 算法 (Gram-Schmidt 正交化) 设 $\{\varphi_k\}$ 是内积空间 V 的一组基, 定义

$$\begin{aligned} \psi_0 &= \varphi_0, \\ \psi_n &= \varphi_n - \sum_{i=0}^{n-1} (\varphi_n, \psi_i) \eta_i \end{aligned}$$

其中 $\eta_i = \psi_i / \|\psi_i\|^2$. 则 $\{\psi_k\}$ 为 V 的一组正交基.

评注 要求 n 次正交多项式组, 只需另 $\varphi_k = x^k$, 再进行 Gram-Schmidt 正交化即可.

50 定理 设 $\{\varphi_n\}_{n=0}^\infty$ 是一列正交多项式, 根据正交性 (从而线性无关) 可以得到正交多项式的如下性质,

1. $P_n \subset \text{span}\{\varphi_0, \dots, \varphi_n\}$,
2. 设 $P \in P_{n-1}$, 则 φ_n 与 P 正交.

51 定理 设 $\{\varphi_n\}_{n=0}^\infty$ 是 $[a, b]$ 上带权 ρ 的正交多项式, 则成立

$$\varphi_{n+1} = (x - \alpha_n)\varphi_n - \beta_n\varphi_{n-1}, \quad n = 0, 1, \dots,$$

其中

$$\begin{aligned} \varphi_0 &= 1, \quad \varphi_{-1} = 0, \\ \alpha_n &= (x\varphi_n, \varphi_n) / (\varphi_n, \varphi_n), \\ \beta_n &= (\varphi_n, \varphi_n) / (\varphi_{n-1}, \varphi_{n-1}). \end{aligned}$$

证明 由于齐次性, 不妨设 φ_n 首项系数为 1. 所以成立

$$\varphi_{n+1} - x\varphi_n = \sum_{k=0}^n \gamma_k \varphi_k,$$

对于系数 γ_k , 成立⁸

$$\gamma_k = \frac{(\varphi_{n+1} - x\varphi_n, \varphi_k)}{(\varphi_k, \varphi_k)} = \frac{(\varphi_{n+1}, \varphi_k) - (\varphi_n, x\varphi_k)}{(\varphi_k, \varphi_k)}.$$

由于 $\{\varphi_n\}$ 正交, 所以当 $k < n-1$ 时, 成立 $\gamma_k = 0$. 所以有

$$\varphi_{n+1} - x\varphi_n = \gamma_n \varphi_n + \gamma_{n-1} \varphi_{n-1}.$$

再进行一些代换, 即可以得到原递推式. ■

52 定理 设 $\{\varphi_n\}_{n=0}^{\infty}$ 为 $[a, b]$ 上带权 ρ 的正交多项式, 则 φ_n 在区间 (a, b) 上有 n 个不同的零点.

证明 首先利用权函数的定义, 证明零点不可能都是偶数重的. 再假设 x_1, \dots, x_l 是 φ_n 的奇数重零点, 则

$$(\varphi_n, (x - x_1) \cdots (x - x_l)) \neq 0,$$

再利用正交性可得 $l = n$. ■

⁸ 设 B 是内积空间 V 的一组正交基, 则对于任意 $x \in V$, 成立

$$x = \sum_{\beta \in B} \frac{(x, \beta)}{(\beta, \beta)} \beta$$

另外, 根据这里内积的定义, 成立 $(\varphi_i, \varphi_j) = (\varphi_i/x, x\varphi_j)$.

3.4 Legendre 多项式

53 定义 (Legendre 多项式) 取区间 $[-1, 1]$, $\rho(x) \equiv 1$, 称由 $\{1, x, \dots, x^n, \dots\}$ 正交化而得的多项式为 Legendre 多项式. 其表达式为

$$P_0(x) = 1, \quad P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n, \quad n = 1, 2, \dots$$

首项系数为 1 的 Legendre 多项式为

$$\tilde{P}_n(x) = \frac{n!}{(2n)!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

54 定理 (Legendre 多项式的性质) Legendre 多项式有如下性质,

1. 正交性:

$$\int_{-1}^1 P_n(x) P_m(x) dx = \begin{cases} 0, & m \neq n, \\ \frac{2}{2n+1}, & m = n. \end{cases}$$

2. 奇偶性:

$$P_n(x) = (-1)^n P_n(-x)$$

3. 递推关系:

$$(n+1)P_{n+1} = (2n+1)xP_n - nP_{n-1}, \quad n = 1, 2, \dots$$

证明 todo

55 定理 (Legendre 多项式的逼近性质) 在区间 $[-1, 1]$ 上, 设 \tilde{L}_n 是首项系数为 1 的 Legendre 多项式, 则

$$\|\tilde{L}_n\|_2 = \min_{P \in P_n} \|P(x)\|_2.$$

评注 应用方法和说明可以参考 Chebyshev 多项式的逼近性质. (定理 59)

56 引理 (前 4 项 Legendre 多项式)

$$\begin{aligned} P_0 &= 1, \\ P_1 &= x, \\ P_2 &= \frac{3}{2}x^2 - \frac{1}{2}, \\ P_3 &= \frac{5}{2}x^3 - \frac{3}{2}x \end{aligned}$$

3.5 Chebyshev 多项式

57 定义 (Chebyshev 多项式) 取区间 $[-1, 1]$, $\rho(x) = (1-x^2)^{-1/2}$, 称由 $\{1, x, \dots, x^n, \dots\}$ 正交化而得的多项式为 Chebyshev 多项式. 其表达式为

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

58 定理 (Chebyshev 多项式的性质) Chebyshev 多项式有如下性质,

1. 递推关系:

$$\begin{aligned} T_0(x) &= 1, \quad T_1(x) = x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, \dots \end{aligned}$$

2. 正交性:

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

3. $T_{2k}(x)$ 只含 x 的偶次幂, $T_{2k+1}(x)$ 只含 x 的奇次幂.

4. T_n 在区间 $[-1, 1]$ 上的 n 个零点为

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

5. T_n 的首项系数为 2^{n-1} .

59 定理 (Chebyshev 多项式的逼近性质) 在区间 $[-1, 1]$ 上, 设 \tilde{T}_n 是首项系数为 1 的 Chebyshev 多项式, 则

$$\|\tilde{T}_n\|_\infty = \min_{P \in \tilde{P}_n} \|P(x)\|_\infty = \frac{1}{2^{n-1}}.$$

评注 这一定理意味着, 取区间 $[-1, 1]$, n 次 Chebyshev 多项式是所有次数小于等于 n 的首项为 1 的多项式中, 绝对值的最大值最小的一个. 从而, 若想用 P_{n-1} 中的多项式来逼近 n 次多项式 f , 只需找 $f^* \in P_{n-1}$, 使得

$$f - f^* = a_n \tilde{T}_n.$$

其中 a_n 为 f 的 n 次项系数. 对于一般的在区间 $[a, b]$ 上的情况, 只需利用平移和伸缩映射到 $[-1, 1]$ 上即可.

60 定理 (Chebyshev 零点插值) 设插值节点 x_0, \dots, x_n 为 Chebyshev 多项式 T_{n+1} 的零点, 被插值函数 $f \in \mathcal{C}^{n+1}[-1, 1]$, 则多项式插值的余项 R_n 满足

$$|R_n| \leq \frac{1}{2^n(n+1)!} \|f^{(n+1)}(x)\|_\infty.$$

证明 由于插值点是 Chebyshev 多项式的零点, 所以 $\omega_{n+1} = \tilde{T}_n$, 所以根据定理 59, 成立

$$\omega_{n+1} \leq \frac{1}{2^n}. \quad \blacksquare$$

评注 这一定理保证了使用 Chebyshev 多项式的零点插值, 至少可以使得误差的最大值最小.

4 变分方法与数据拟合

4.1 绪论

若要求函数 $f(\mathbf{x})$, $x \in \mathbf{R}^n$ 的最值点, 根据 Fermat 引理, 只需要求出所有成立 $\nabla f = 0$ 的点再逐一验证即可. 变分是这一思想的推广, 它所处理的是过程的优化. 下给出一个优化过程的例子, 完整的解答见之后的章节. (todo: ref)

61 例 (最速降线) 给定空中的一点 $A = (0, 0)$, 地上一点 $B = (x_1, y_1)$, 求一条连接 A 和 B 的轨迹, 使得假设在无阻力情况下, 有小球沿轨道从 A 到 B 所需要的时间最短.

假设轨道曲线充分光滑, 则问题可以转换为, 设滑行轨道 $y \in \mathcal{C}^2$ 的方程为

$$y = y(x), \quad y(0) = 0, \quad y(x_1) = y_1.$$

试确定曲线 y 使得时间 $T(y)$ 最小. 根据机械能守恒, 小球在 $(*, -y)$ 处 (见图 2) 的速度, 为

$$v = \sqrt{2gy},$$

同时由速度的定义知

$$v = \frac{ds}{dt} = \sqrt{1 + y'^2(x)} \frac{dx}{dt},$$

根据上两式, 即有

$$\sqrt{2gy} dt = \sqrt{1 + y'^2(x)} dx \Rightarrow dt = \left(\frac{1 + y'^2}{2gy} \right)^{1/2} dx$$

设过程为 $A \rightarrow B$, $0 \rightarrow t_1$, $0 \rightarrow x_1$, 则

$$T(y) = t_1 = \int_0^{x_1} \left(\frac{1 + y'^2}{2gy} \right)^{1/2} dx,$$

即为所要最小化的 $T(y)$.

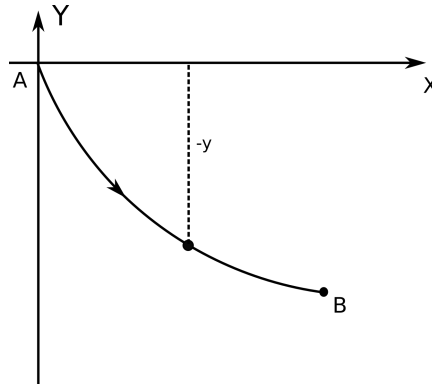


图 2: 最速降线问题

4.2 变分方法

62 定义 (过程优化) 过程的优化即求解

$$y^* = \arg \min_{y \in K} J(y)$$

的过程, 其中函数集合

$$K = \{y \in \mathcal{C}^2[x_0, x_1] : y(x_0) = y_0, y(x_1) = y_1\}.$$

63 定义 定义函数集合

$$K = \{f \in \mathcal{C}^2[x_0, x_1] : f(x_0) = y_0, f(x_1) = y_1\},$$

$$K_0 = \{f \in \mathcal{C}^2[x_0, x_1] : f(x_0) = f(x_1) = 0\}.$$

对于任意 $f_0 \in K$, $\eta \in K_0$ 定义集合

$$K(f_0, \eta) = \{f_0 + \varepsilon\eta : \varepsilon \in \mathbf{R}\}.$$

64 定义 (泛函的方向导数) 记号同定义 63. 定义泛函

$$J(f) = \int_a^b L(x, f(x), f'(x))dx, \quad f \in K.$$

如果 f^* 是函数 $J(f)$ 在集合 K 中的最小值点, 即

$$f^* = \arg \min_{f \in K} J(f),$$

则对于任意 $\eta \in K_0$, f^* 也是 $J(f)$ 在集合 $K(f^*, \eta)$ 中的最小值点. 所以成立

$$f^* = \arg \min_{\varepsilon \in \mathbf{R}} J(f^* + \varepsilon\eta).$$

由于函数 $J(f^* + \varepsilon\eta)$ 是关于实数 ε 的一元函数, 所以在它最小值点, 即 $\varepsilon = 0$ 处成立

$$\left. \frac{d}{d\varepsilon} J(f^* + \varepsilon\eta) \right|_{\varepsilon=0} = 0. \quad (3)$$

称左侧为泛函 J 在 f^* 处沿 η 方向的**变分**或**方向导数**. 即为 $\delta J(f^*, \eta)$ ⁹.

评注 这里采用的是分析学中一个常见思想, 将一个在高维空间中的问题转化为一个低维空间中的问题. 一个更加简单的例子是, 证明若 k 维欧式空间中的函数 f 在凸集 K 中的各偏导数恒为零, 则它在 K 中为常量. 一个证法是取定 K 中任意一点 (向量) \mathbf{x}_0 , 则对于任意 $\mathbf{x} \in K$, 则对于任意直线段 xx_0 , 有方程

$$y = \mathbf{x}_0 t + \mathbf{x}(1 - t).$$

上式是一个一元实函数, 对它求导并利用一元函数微分学中的知识, 可以知道在这条直线段上 f 的函数值不变. 由于 \mathbf{x}_0 和 \mathbf{x} 的选取是任意的, 所以在 K 上 f 的值不变.

在这里采用的是同样的思想, 只是把欧式空间换成了一个函数集合而已.

⁹一般来讲, f^* 可以是集合中的任意一点.

65 引理 (变分引理 I) 设 $f \in \mathcal{C}[a, b]$, 且对任意满足 $g(a) = g(b) = 0$ 的 $g \in \mathcal{C}^2[a, b]$, 有

$$\int_a^b f g dx = 0,$$

则在 $[a, b]$ 上成立 $f \equiv 0$.

66 引理 (变分引理 II) 设 $f \in \mathcal{C}[a, b]$, 且对于任意满足 $\eta(a) = \eta(b) = 0$ 的 $\eta \in \mathcal{C}^1(a, b)$ 都成立

$$\int_a^b f \eta' dx = 0,$$

则在 $[a, b]$ 上成立 $f \equiv \text{Const.}$

67 定理 (Euler-Lagrange 方程) 记号同定义 64. 泛函 $J(f)$ 在极值点满足 Euler-Lagrange 方程

$$\frac{\partial L}{\partial f} - \frac{d}{dx} \frac{\partial L}{\partial f'} = 0. \quad (4)$$

证明 假设满足求导积分换序的条件, 则泛函 $J(f)$ 在 $f_* + \varepsilon \eta$ 处沿 η 方向的方向导数为¹⁰

$$\begin{aligned} \frac{d}{d\varepsilon} J(f_* + \varepsilon \eta) &= \frac{d}{d\varepsilon} \int_{x_0}^{x_1} L(x, f_* + \varepsilon \eta, f'_* + \varepsilon \eta') dx \\ &= \int_{x_0}^{x_1} \frac{\partial L}{\partial (f_* + \varepsilon \eta)} \eta + \frac{\partial L}{\partial (f'_* + \varepsilon \eta')} \eta' dx. \end{aligned}$$

代入 $\varepsilon = 0$, 即得 J 在最小值点 f_* 处的方向导数, 为

$$\begin{aligned} \frac{d}{d\varepsilon} J(f_*) &= \int_{x_0}^{x_1} \frac{\partial L}{\partial f} \eta + \frac{\partial L}{\partial f'} \eta' dx \\ &= \int_{x_0}^{x_1} \frac{\partial L}{\partial f} \eta dx + \left. \frac{\partial L}{\partial f'} \eta \right|_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta \frac{d}{dx} \frac{\partial L}{\partial f'} dx \\ &= \int_{x_0}^{x_1} \eta \left(\frac{\partial L}{\partial f} - \frac{d}{dx} \frac{\partial L}{\partial f'} \right) dx = 0. \end{aligned}$$

注意由于 $\eta \in K_0$, 即有 $\eta(x_0) = \eta(x_1) = 0$. 由于 η 的选取是任意的, 所以根据引理 65, 式子 (4) 成立. ■

评注 对于 $L = L(f, f')$, 即 L 不显含 x 的情况, (4) 是可以精确求解的.

68 定理 (守恒律定理) 设 $L = L(f, f')$, 则沿着 (62) 的解曲线 $y^* = f^*(x)$, 成立

$$H = f' \frac{\partial L}{\partial f'} - L = \text{Const.}$$

¹⁰为了记号的清晰, 这里用 f_* 替代之前的 f^* . 并且这里关于偏导数的记号, 应理解为关于分母所表示的那一分量的偏导数.

证明

$$\begin{aligned}\frac{dH}{dx} &= \frac{d}{dx} \left(f' \frac{\partial L}{\partial f'} - L(f, f') \right) \\ &= f'' \frac{\partial L}{\partial f'} + f' \frac{d}{dx} \frac{\partial L}{\partial f'} - \frac{\partial L}{\partial f} f' - \frac{\partial L}{\partial f'} f''.\end{aligned}$$

根据定理 67, $dH/dx = 0$, 所以命题成立. ■

4.3 曲线拟合的正则化方法

69 定义 (Tikhonov 正则化) 对于给定的数据 Y , 定义数据拟合项为 $J_1(f)$, 用于表示拟合结果相较于原数据的接近程度, 同时要求拟合的结果尽可能满足对于结果的要求, 用 $J_2(f)$ 来描述 f 满足要求的程度, 则求解拟合结果的过程即为求解

$$f_* = \arg \min_{f \in K} (J_1(f) + \alpha J_2(f)).$$

其中 α 为**正则化参数**, 用于表示拟合的过程中, 应更接近原数据或是更满足拟合要求. 若取 $\alpha = 0$, 即为插值.

70 问题 给定函数 y 在样本点 $0 = x_0 < x_1 < \cdots < x_n = 1$ 处的近似值 \tilde{y}_i , 误差满足

$$|\tilde{y}_i - y(x_i)| \leq \delta,$$

试重构 y 的近似函数 f_* .

按照定义 69 的思想, 定义

$$J_1(f) = \sum_{i=1}^{n-1} \frac{h_i + h_{i+1}}{2} (\tilde{y}_i - f(x_i))^2$$
$$J_2(f) = \int_0^1 (f'')^2 dx$$

其中

$$h_i = x_i - x_{i-1}, \quad i = 1, 2, \dots, n$$
$$h = \max_{1 \leq i \leq n} h_i$$

则问题转换为求解

$$f_* = \arg \min_{f \in K} (J_1(f) + \alpha J_2(f)). \quad (5)$$

评注 不失一般性的, 可以设 $\tilde{y}_0 = f(x_0)$ 且 $\tilde{y}_n = f(x_n)$. 否则只需要用

$$Y(x) = y(x) + \tilde{y}_0 - y(0) + (\tilde{y}_n - y(1) + y(0) - \tilde{y}_n)x$$

来替代 y 即可. 可以证明

1. $Y(0) = \tilde{y}_0$ 且 $Y(1) = \tilde{y}_n$,
2. $|\tilde{y}_i - Y(x_i)| \leq 4\delta$.

71 定理 对于任意 $\alpha > 0$, (5) 的解为三次样条函数.

证明 设 f_* 为式 (5) 的解.

对于任意 $\eta \in K_0$, $\varepsilon \in R$,

$$J(f_* + \varepsilon\eta) = \sum_{i=1}^{n-1} \frac{h_i + h_{i+1}}{2} [\tilde{y}_i - f_*(x_i) - \varepsilon\eta(x_i)]^2 + \alpha \int_0^1 (f_*'' + \varepsilon\eta'')^2 dx.$$

求它关于 ε 的导数, 成立

$$\frac{d}{d\varepsilon} J(f_* + \varepsilon\eta) = -2 \sum_{i=1}^{n-1} \frac{h_i + h_{i+1}}{2} [\tilde{y}_i - f_*(x_i) - \varepsilon\eta(x_i)] \eta(x_i) + 2\alpha \int_0^1 (f_*'' + \varepsilon\eta'') \eta'' dx.$$

根据定义 64 中的 (3), 上式在 $\varepsilon = 0$ 时值为零, 即

$$\left. \frac{d}{d\varepsilon} J(f_* + \varepsilon\eta) \right|_{\varepsilon=0} = -2 \sum_{i=1}^{n-1} \frac{h_i + h_{i+1}}{2} [\tilde{y}_i - f_*(x_i)] \eta(x_i) + 2\alpha \int_0^1 f_*'' \eta'' dx = 0. \quad (6)$$

接下来分两步构造出 f_* .

Step 1. 由于 η 的选取是任意的, 所以我们选择恰当的 $\eta \in \mathcal{C}^\infty(x_i, x_{i+1})$, 使得成立

$$\eta(x_i) = 0, \quad i = 1, 2, \dots, n-1$$

所以根据 (6), 成立

$$\int_{x_i}^{x_{i+1}} f_*'' \eta'' = 0.$$

对于上式进行两次分部积分, 得到

$$0 = f_*'' \eta' \Big|_{x_i}^{x_{i+1}} - f_*^{(3)} \eta \Big|_{x_i}^{x_{i+1}} + \int_{x_i}^{x_{i+1}} f_*^{(4)} \eta dx.$$

同样因为 η 的选取是任意的, 所以可以在原来的基础上, 选取 η 使得成立

$$\eta(x_i) = \eta'(x_i) = 0, \quad i = 1, 2, \dots, n-1$$

所以有

$$\int_{x_i}^{x_{i+1}} f_*^{(4)} \eta dx = 0.$$

根据变分引理, 即成立

$$f_*^{(4)}(x) = 0 \quad \Rightarrow \quad f \in P_3, \quad x \in (x_i, x_{i+1}). \quad (7)$$

Step 2. 取满足 $\eta(0) = \eta(1) = 0$ 的 $\eta \in \mathcal{C}^\infty[0, 1]$, 分部积分得 (注意 $f_*^{(4)} = 0$)

$$\begin{aligned} \int_0^1 f_*'' \eta'' dx &= \sum_{i=1}^{n-1} \left(-f_*^{(3)} \eta \Big|_{x_{i-1}}^{x_i} + f_*'' \eta' \Big|_{x_{i-1}}^{x_i} \right) \\ &= \sum_{i=1}^{n-1} (f_*^{(3)}(x_i+) - f_*^{(3)}(x_i-)) \eta(x_i) + \sum_{i=1}^{n-1} (f_*''(x_i+) - f_*''(x_i-)) \eta'(x_i) \\ &\quad + f_*''(1) \eta'(1) - f_*''(0) \eta'(0) \end{aligned}$$

将结果代入 (6), 即

$$\begin{aligned}
0 = & f''_*(1)\eta'(1) - f''_*(0)\eta'(0) \\
& + \sum_{i=1}^{n-1} \left(-\frac{h_i + h_{i+1}}{2} [\tilde{y}_i - f_*(x_i)] + \alpha(f_*^{(3)}(x_{i+}) - f_*^{(3)}(x_{i-})) \right) \eta(x_i) \\
& + \alpha \sum_{i=1}^{n-1} (f''_*(x_{i+}) - f''_*(x_{i-})) \eta'(x_i)
\end{aligned} \tag{8}$$

由于 η 的选取是任意的, 即意味着 $\eta(x_i)$, $(i = 1, 2, \dots, n-1)$ 和 $\eta'(x_i)$, $(i = 0, 1, \dots, n)$ 的选取是任意的, 所以需要选取适当的 f 才可以使得 (8) 中的每一项都恒为零. 为使第 1、3 项为零, 需要满足

$$\begin{aligned}
f''_*(1) &= f''_*(0) = 0, \\
f''_*(x_{i+}) &= f''_*(x_{i-}).
\end{aligned}$$

同样的, 根据不同的 α , 选取恰当的 $f(x_i)$ 或 $f^{(3)}(x_i)$ 使得第二项为零. 综合上式以及 (7), 可知 f_* 为三次样条函数, 且当 $\alpha = 0$ 时为样条插值函数. ■

5 数值积分与数值微分

5.1 绪论

72 定理 (N-L 公式) 设 f 和 F 定义在 $[a, b]$ 上, $F \in \mathcal{C}[a, b]$ 且在 (a, b) 上成立 $F' = f$, 则

$$\int_a^b f(x)dx = F(b) - F(a).$$

评注 N-L 公式是求解积分的基本方法. 但是 $F(x)$ 通常是难以求解的, 所以需要数值方法.

73 定理 (积分第一中值定理) 设 $f \in \mathcal{C}[a, b]$, $g \in \mathcal{R}[a, b]$ 且不变号, 则存在 $\xi \in [a, b]$, 成立

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx.$$

证明 不妨设 $g(x) \geq 0$. 若 $g(x) \equiv 0$, 则结论是显然的. 考虑 g 不恒为零的情况. 由于 $f \in \mathcal{C}$, 所以 f 在 $[a, b]$ 上可以取到最小值 m 和最大值 M , 则成立

$$\begin{aligned} mg(x) &\leq f(x)g(x) \leq Mg(x) \\ \Rightarrow m \int_a^b g(x)dx &\leq \int_a^b f(x)g(x)dx \leq M \int_a^b g(x)dx \\ \Rightarrow m &\leq \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \leq M \end{aligned}$$

由于连续函数有介质性, 所以存在 $\xi \in [a, b]$ 成立

$$f(\xi) = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx}.$$

即成立

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx. \quad \blacksquare$$

评注 注意, 定理要求 $f \in \mathcal{C}[a, b]$, 这在很多情况下是难以满足的. 但是有些时候可以利用证明中的思路, 在 f 不连续的情况下得出同样的结论.

74 命题 设 $f \in \mathcal{C}[a, b]$, $g \in \mathcal{R}[a, b]$ 且不变号, $\psi: [a, b] \rightarrow [a, b]$, 则存在 $\xi \in [a, b]$, 使得

$$\int_a^b f(\psi(x))g(x)dx = f(\xi) \int_a^b g(x)dx.$$

评注 证明同定理 73 是几乎一样的。

75 命题 (中矩形公式) 设 f 足够光滑, 则

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

且存在 $\xi \in [a, b]$, 其误差 $R(f) = \text{L.H.S} - \text{R.H.S}$ 满足

$$R(f) = \frac{1}{24}(b-a)^3 f''(\xi).$$

证明

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

在 $x = (a+b)/2$ 处带 Lagrange 余项 Taylor 展开, 有

$$\begin{aligned} R(f) &= \int_a^b \left[f'\left(\frac{a+b}{2}\right) \left(x - \frac{a+b}{2}\right) + \frac{1}{2} f''(\xi(x)) \left(x - \frac{a+b}{2}\right)^2 \right] dx \\ &= \frac{1}{2} \int_a^b f''(\xi(x)) \left(x - \frac{a+b}{2}\right)^2 dx \end{aligned}$$

由于 $\xi(x)$ 的连续性是无法保证的, 所以无法直接使用积分中值定理. 但是可以根据命题 74 得, 存在 $\zeta \in [a, b]$, 成立

$$R(f) = \frac{1}{2} f''(\zeta) \int_a^b \left(x - \frac{a+b}{2}\right)^2 dx = \frac{1}{24} (b-a)^3 f''(\zeta). \quad \blacksquare$$

76 命题 (梯形公式) 设 f 足够光滑, 则

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

且存在 $\xi \in [a, b]$, 其误差 $R(f) = \text{L.H.S} - \text{R.H.S}$ 满足

$$R(f) = -\frac{1}{12} (b-a)^3 f''(\xi).$$

评注 证明同命题 75 的证明是相似的. 实际上梯形公式可以理解为对原有的函数进行线性插值, 并用插值结果的积分来近似原函数的积分. 按照这一思路推广, 即发现可以用插值多项式的积分来近似原来函数的积分.

77 命题 (Simpson 公式) 设 f 足够光滑, 则

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

存在 $\xi \in [a, b]$, 其误差 $R(f) = \text{L.H.S} - \text{R.H.S}$ 满足

$$R(f) = -\frac{(b-a)^5}{2880} f^{(4)}(\xi).$$

评注 实际上这是以 $a, b, (a+b)/2$ 作插值节点作二次多项式插值后的函数积分后的结果. 但是若按照之前的思路证明, 会发现 $g(x) = (x-a)(x-(a+b)/2)(x-b)$ 在 $[a, b]$ 上是不保号的, 所以不能直接沿用之前的做法. 为了让它保号, 我们可以采用 Hermite 插值, 具体证法见下. 另外, 有人或许会尝试把它拆分到两个区间上, 让 g 分别保号, 但这样的做法一般来说是错误的, 问题出现在最后一步合并两个区间结果的时候, 权重有可能会出现负值.

证明 设 Hermite 插值多项式 $H \in P_3$ 满足 $H(a) = f(a), H(b) = f(b), H((a+b)/2) = f((a+b)/2), H'((a+b)/2) = f'((a+b)/2)$. 则存在 $\xi \in [a, b]$, 使其误差满足

$$f(x) - H(x) = \frac{f^{(4)}(\xi)}{4!}(x-a)\left(\frac{a+b}{2}\right)^2(x-b).$$

它在 $[a, b]$ 上是保号的. 对 $H(x)$ 积分即可得 R.H.S. 剩下的内容和之前的证明是相同的. ■

评注 根据余项公式可以发现, Simpson 公式对于 $f \in P_3$ 都是精确成立的, 基于此思想, 定义代数精度.

78 定义 (代数精度) 记¹¹

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

若数值积分公式

$$I(f) \approx Q(f)$$

在 $f \in P_n$ 时精确成立, 在 $f \in P_{n+1}$ 时不精确成立, 则称该求积公式有 n 阶代数精度.

评注 代数精度是求积公式精确程度的一个度量, 通常希望求积公式有更高的代数精度.

79 定义 (机械求积公式) 称求积公式为机械求积公式, 若它满足

$$I(f) = \sum_{k=0}^n A_k f(x_k).$$

称 $x_k \in [a, b]$ 为求积节点, 称 A_k 为求积系数, 它的选取与被积函数无关, 至于求积节点有关.

¹¹在以后若不特殊说明, 此记号都表示在 $[a, b]$ 上的积分.

评注 可以发现这一节中出现的求积公式都是机械求积公式，但它们的误差各不相同. 我们希望通过恰当地选取求积节点 x_k 和求积系数 A_k ，使得求积公式的代数精度尽可能高.

若我们希望机械求积公式有 m 次代数精度，即意味着对于 $1, x, \dots, x^m$ 的积分都是精确成立的，即成立非线性方程组

$$\begin{aligned}\sum A_k &= b - a, \\ \sum A_k x_k &= \frac{1}{2}(b^2 - a^2), \\ &\dots\dots\dots \\ \sum A_k x_k^m &= \frac{1}{m+1}(b^{m+1} - a^{m+1}).\end{aligned}$$

这一方程组求解通常是困难的.

80 定义 (插值求积公式) 给定节点 $a \leq x_0 < \dots < x_n \leq b$ 处的函数值 $f(x_i)$ ，构造 n 次 Lagrange 多项式，则有

$$I(f) \approx \sum_{k=0}^n \left(\int_a^b l_k dx \right) f(x_k).$$

l_k 的定义见定理 20. 其误差满足

$$R(f) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega_{n+1}(x)$$

81 定义 (收敛) 对于求积公式

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

记 $h = \max\{x_i - x_{i-1}\}$ ，若当 $h \rightarrow 0$ 时，L.H.S \rightarrow R.H.S，则称求积公式是收敛的.

82 定义 (稳定) 设 \tilde{f}_k 是 $f(x_k)$ 的近似值. 对于任意的 $\varepsilon > 0$ ，如果存在 $\delta > 0$ ，只需 $|\tilde{f}_k - f(x_k)| < \delta$ ，就成立

$$\left| Q(f) - Q(\tilde{f}) \right| = \left| \sum_{k=0}^n A_k (f(x_k) - \tilde{f}_k) \right| < \varepsilon,$$

则称求积公式是稳定的.

评注 一个求积公式稳定，意味着测量误差并不会随着计算而扩大.

83 定理 (稳定性条件) 如果机械求积公式的系数 $A_k > 0$ ，则该求积公式是稳定的.

5.2 Newton-Cotes 公式

84 定理 (Newton-Cotes 公式) 将积分区间 $[a, b]$ n 等分, 步长 $h = \frac{b-a}{n}$, 构造插值型求积公式, 即

$$Q(f) = (b-a) \sum_{k=0}^n C_k^{(n)} f(x_k)$$

称为 Newton-Cotes 公式, 其中 $C_k^{(n)}$ 称为 Cotes 系数, 为

$$C_k^{(n)} = \frac{h}{b-a} \int_0^n \prod_{j=0, j \neq k}^n \frac{t-j}{k-j} dt = \frac{(-1)^{n-k}}{nk!(n-k)!} \int_0^n \prod_{j=0, j \neq k}^n (t-j) dt$$

评注 当 $n \geq 8$ 的时候, Cotes 系数出现负值, 意味着测量误差会随着计算而增大, 所以 $n \geq 8$ 的 Newton-Cotes 公式是不用的.

85 定理 (插值型求积公式的代数精度) 形如

$$Q(f) = \sum_{k=0}^n A_k f(x_k)$$

的求积公式有 n 次或以上代数精度的充要条件为, 它是插值型的.

证明 充分性是显然的. 对于必要性, 用 $Q(f)$ 计算 l_k 来得到 A_k , 即

$$\int_a^b l_k(x) dx = \sum_{j=0}^n A_j l_k(x_j) \Rightarrow A_k = \int_a^b l_k(x) dx. \quad \blacksquare$$

86 定理 (偶阶求积公式的代数精度) 当 n 为偶数时, n 阶插值型求积公式至少有 $n+1$ 次代数精度.

5.3 复合求积公式

即对原积分区间 n 等分, 使得每一段长度 $h < 1$, 之后在每一段上分别求积.

87 定理 (复合梯形公式余项)

$$T_n(f) = \frac{h}{2} [f(a) + 2 \sum_{k=1}^{n-1} f(x_k) + f(b)]$$

其余项为

$$R_n(f) = -\frac{b-a}{12} h^2 f''(\eta), \quad \eta \in [a, b]$$

88 定理 (复合 Simpson 公式)

$$S_n = \frac{h}{6} [f(a) + 4 \sum_{k=0}^{n-1} f(x_{k+1/2}) + 2 \sum_{k=1}^{n-1} f(x_k) + f(b)]$$

其余项为

$$R_n(f) = -\frac{b-a}{180} \left(\frac{h}{2}\right)^4 f^{(4)}(\eta), \quad \eta \in [a, b]$$

5.4 Gauss 求积公式

89 命题 (求积公式代数精度上限) 形如下式的求积公式的代数精度至多为 $2n + 1$.

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

评注 根据本节之后的构造, 可以发现这个上限是取得到的.

90 定义 (Gauss 求积公式) 设

$$I(f) = \int_a^b f(x)\rho(x)dx \approx \sum_{k=0}^n A_k f(x_k) = Q(f). \quad (9)$$

若 (9) 有 $2n + 1$ 次代数精度, 则称其节点 x_k 为 Gauss 点, 称该公式为 Gauss 型求积公式.

评注 若要解权函数对应的 Gauss 求积公式, 则只需要取 $f(x) = x^k, k = 0, 1, \dots, 2n+1$, 解对应的方程组即可. 但由于它是关于 x_k 和 A_k 非线性的方程组, 所以需要先确定 x_k , 得到关于 A_k 的线性方程组.

91 定理 节点为 $a \leq x_0 < \dots < x_n = b$ 的带权机械求积公式 $Q(f)$ 有 $n+k$ ($0 \leq k \leq n+1$) 次代数精度的充要条件为:

1. $Q(f)$ 为插值型求积公式,
2. 对任意 $p \in P_{k-1}$, 成立

$$\int_a^b \rho(x)p(x)\omega_{n+1}(x)dx = 0$$

证明 todo

92 推论 (9) 有 $2n + 1$ 次代数精度, 当且仅当节点 $\{x_i\}$ 是 $n + 1$ 次正交多项式的根.

评注 若积分区间为 $[a, b]$, 一般会先变换到 $[-1, 1]$ 上. 之后的讨论中的区间都选取 $[-1, 1]$.

93 命题 Gauss 求积公式的系数全是正的, 从而它是稳定的.

94 定理 Gauss 求积公式是收敛的, 即

$$\lim_{n \rightarrow \infty} \sum_{k=0}^n A_k f(x_k) = \int_a^b \rho f dx.$$

5.5 Romberg 求积公式

动机 在利用复合求积公式的时候, 如果需要增加精度, 则需要再二分一遍区间. 即增加了节点 $x_{k+\frac{1}{2}} = \frac{1}{2}(x_k + x_{k+1})$. 如果每次增加节点后都需要按照求积公式重新计算, 则显然计算量过大并且浪费了先前计算好的结果, 所以希望可以充分地利用先前的计算结果来简化计算.

95 算法 (Richardson 外推方法) 设 Q 为需要计算的值的精确结果, $Q_1(h)$ 是通过某一算法计算得的 Q 的近似, 若误差满足

$$Q - Q_1(h) = c_1 h^{p_1} + c_2 h^{p_2} + \cdots = o(h^{p_1-1}) \quad (10)$$

其中 $0 < p_1 < p_2 < \cdots$ 与 h 无关. 则令

$$Q_{k+1}(h) = \frac{Q_k(h/2) - 2^{-p_k} Q_k(h)}{1 - 2^{-p_k}}.$$

其误差满足

$$Q - Q_{k+1}(h) = c_2^* h^{p_2} + c_3^* h^{p_3} + \cdots = o(h^{p_2-1})$$

评注 Richardson 外推方法意味着如果有二分前的结果 $Q(h)$ 和二分后的结果 $Q(h/2)$, 那么可以通过这两个结果做一次外推从而得到更高的精度. 并且二分了多少次, 就可以外推多少次.

证明 根据 (10), 成立

$$\begin{aligned} Q - Q(h/2) &= c_1 \left(\frac{h}{2}\right)^{p_1} + c_2 \left(\frac{h}{2}\right)^{p_2} + \cdots \\ 2^{-p_1}(Q - Q(h)) &= c_1 \left(\frac{h}{2}\right)^{p_1} + c_2 2^{-p_1} h^{p_2} + \cdots \end{aligned}$$

将上述两式相减, 即成立

$$\begin{aligned} Q - Q(h/2) - 2^{-p_1}(Q - Q(h)) &= c_2^* h^{p_2} + c_3^* h^{p_3} + \cdots \\ \Rightarrow Q_2(h) &= \frac{Q(h/2) - 2^{-p_1} Q(h)}{1 - 2^{-p_1}} \quad \blacksquare \end{aligned}$$

96 命题 复合梯形公式外推一次即得 Simpson 公式,

97 命题 (复合梯形公式余项) 对于 $[a, b]$ 进行 n 等分, 令 $h = (b - a)/n$, 考虑复合梯形求积公式

$$T(h) = h \left[\frac{f(x_0)}{2} + f(x_1) + \cdots + f(x_{n-1}) + \frac{f(x_n)}{2} \right].$$

存在与 h 无关的 a_i , 成立

$$T(h) = \int_a^b f dx + \sum_{k=1}^{\infty} a_{2k} h^{2k}.$$

98 算法 (Romberg 算法) 记 n 次二等分第 m 次外推后的结果是 $R(n, m)$, 记 $h_n = \frac{1}{2^n}(b-a)$, 则

$$\begin{aligned} R(0, 0) &= h_1(f(a) + f(b)) \\ R(n, 0) &= \frac{1}{2}R(n-1, 0) + h_n \sum_{k=1}^{2^{n-1}} f(a + (2k-1)h_n) \\ R(n, m) &= \frac{1}{4^m - 1}(4^m R(n, m-1) - R(n-1, m-1)) \end{aligned}$$

其余项是 $O(h_n^{2(m-1)})$ 的.

评注 $R(n, 0)$ 是二等分 n 次后的复合梯形公式的结果, 它可以通过重复利用 $R(n-1, 0)$ 中的结果来较快得到. 由于 $R(n, m-1)$ 和 $R(n-1, m-1)$ 的余项满足算法 95 中的要求, 从而可以进行 $n-1$ 次外推.

5.6 自适应求积公式

99 定义 (后验估计) 如果误差的上界是可计算量, 则称为**后验估计**, 否则则称为**先验估计**.¹²

100 例 (刘徽割圆术) 设圆的面积为 S , 它是一个不可计算量. 设圆的内接正 n 边形面积为 S_n , 它是一个可计算量. 则成立刘徽不等式

$$S_{2n} < S < S_{2n} + (S_{2n} - S_n) \Rightarrow 0 < S - S_{2n} < S_{2n} - S_n.$$

它是对于割圆术误差的一个后验估计.

101 定理 (Simpson 公式的后验估计) 记 $S_n(a, b)$ 为对 $[a, b]$ 区间进行 $n-1$ 次二等分, 在每一区间上利用 Simpson 公式数值积分的结果, 记 $I(a, b)$ 为在 $[a, b]$ 上积分 f 的结果, 则成立

$$|I(a, b) - S_2(a, b)| \leq \frac{1}{15} |S_2(a, b) - S_1(a, b)|.$$

证明 Simpson 公式的误差满足

$$I(a, b) - S_1(a, b) = -\frac{b-a}{180} \left(\frac{h}{2}\right)^2 f^{(4)}(\xi) \quad (11)$$

其中 $h = b - a$, $\xi \in [a, b]$. 显然它是一个先验估计. 将 $[a, b]$ 二等分, 在每段上利用 Simpson 公式计算, 则有

$$\begin{aligned} I\left(a, \frac{a+b}{2}\right) - S_1\left(a, \frac{a+b}{2}\right) &= -\frac{b-a}{360} \left(\frac{h}{4}\right)^2 f^{(4)}(\eta_1) \\ I\left(\frac{a+b}{2}, b\right) - S_1\left(\frac{a+b}{2}, b\right) &= -\frac{b-a}{360} \left(\frac{h}{4}\right)^2 f^{(4)}(\eta_2) \end{aligned}$$

¹² 见例 100.

可以发现两个等式的右侧系数是同号的，所以将它们相加，可得

$$I(a, b) - S_2(a, b) = -\frac{b-a}{180} \left(\frac{h}{4}\right)^2 f^{(4)}(\eta) \quad (12)$$

我们可以假设 h 足够的小，从而可以有 $\xi \approx \eta$ ，从而根据 (11) 和 (12)，有

$$\begin{aligned} 16(I(a, b) - S_2(a, b)) &\approx I(a, b) - S_1(a, b) \\ \Rightarrow |I(a, b) - S_2(a, b)| &\leq \frac{1}{15} |S_2(a, b) - S_1(a, b)|. \quad \blacksquare \end{aligned}$$

102 定理 (停机准则) 设所要满足的精度为 ε ，即与精确值间误差小于等于 ε ，则可设置停机准则为

$$\begin{aligned} |I(a, b) - S_2(a, b)| &\leq \varepsilon \\ |I(c, d) - S_2(c, d)| &\leq \frac{d-c}{b-a} \varepsilon. \end{aligned}$$

103 算法 (自适应方法) 首先利用 Simpson 公式的后验估计和停机准则，选出划分 $a = x_0 < \cdots < x_n = b$ 。再用 Romberg 计算各个子区间上的满足一定精度要求的积分值，最后相加。

5.7 数值微分

104 定义 (插值型数值微分) 给定函数 $f(x)$ 在 $a = x_0 < \cdots < x_n = b$ 处的函数值 $f(x_k)$ ，设 $P_n(x)$ 为该插值节点的 n 次插值多项式，用插值多项式在 x_k 处的微分来近似原函数在该点的微分，即

$$P'_n(x_k) \approx f'(x_k).$$

注意，我们仅仅考虑插值节点处的导数。

105 定义 (含重节点的差商) 定义

$$f[x_0, \dots, x_n, x_n] = \lim_{\Delta x \rightarrow 0} f[x_0, \dots, x_n, x_n + \Delta x].$$

评注 根据定义，可以用差商来表示差商的导数，即

$$\frac{d}{dx} f[x_0, \dots, x_n, x] = f[x_0, \dots, x_n, x, x].$$

106 定理 (插值型数值微分的误差) 设符号同定义 104，则在节点 x_k 处的误差满足

$$f'(x_k) - P'_n(x_k) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \omega'_{n+1}(x_k), \quad \xi \in [a, b].$$

证明

$$\begin{aligned} f(x) &= P_n(x) + f[x_0, \dots, x_n, x]\omega_{n+1}(x) \\ \Rightarrow f'(x) &= P'_n(x) + \{f[x_0, \dots, x_n, x, x]\omega_{n+1}(x) + f[x_0, \dots, x_n, x]\omega'_{n+1}(x)\}. \end{aligned}$$

令 $x = x_k$, 有 $\omega_{n+1}(x_k) = 0$, 从而成立

$$f'(x_k) - P'_n(x_k) = f[x_0, \dots, x_n, x]\omega'_{n+1}(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}\omega'_{n+1}(x_k). \quad \blacksquare$$

107 推论 (中点方法) 设有 $x_{k-1} = x_k - h$, x_k , $x_{k+1} = x_k + h$, 则

$$f'(x_k) \approx \frac{f(x_{k+1}) - f(x_{k-1}))}{2h}.$$

其余项满足

$$R(f, x_k) = \frac{f^{(3)}(\xi)}{6}h^2.$$

108 命题 (待定系数法) 考虑数值微分公式

$$f'(x_k) = \alpha_i \sum_{i=0}^n f(x_i).$$

利用待定系数法确定 $\{\alpha_i\}$, 使得 R.H.S - L.H.S 的余项阶数尽可能高.

109 命题 (外推) 中点方法的余项满足 Ricahrdson 外推的条件, 所以可以有

$$\begin{aligned} G_0(h) &= \frac{f(x+h) - f(x-h)}{2h}, \\ G_m(h) &= \frac{4^m G_{m-1}(h/2) - G_{m-1}(h)}{4^m - 1}. \end{aligned}$$

110 例 (测量误差的影响) 对于函数 $y = f(x)$, 设测量得的函数值为 $y_n = f(x) + \frac{1}{n^2} \sin(n^4 x)$, 则测量误差满足 $\|y - y_n\|_\infty = \frac{1}{n^2}$, 但是对于误差, 成立

$$\|y' - y'_n\|_\infty = \left\| \frac{1}{n^2} n^4 \cos(n^4 x) \right\|_\infty = n^2.$$

可以发现节点越多, 误差可能反而越大, 这说明了求导运算是不稳定的.

todo 测量误差与误差分析.

6 非线性方程求根

6.1 二分法

111 定理 (闭区间套定理) 设 $\{[a_n, b_n]\}$ 满足

1. $[a_n, b_n] \subset [a_{n-1}, b_{n-1}]$,
2. 当 $n \rightarrow \infty$ 时, $|b_n - a_n| \rightarrow 0$.

则存在 ξ 成立

$$\bigcap_{k=0}^{\infty} [a_n, b_n] = \{\xi\}.$$

112 定理 (连续函数零点定理) 设 $f \in \mathcal{C}[a, b]$ 且 $f(a)f(b) < 0$, 则存在 $c \in [a, b]$, 成立 $f(c) = 0$.

评注 这一定理是多种方程求根方法的基础, 下给出构造性的证明, 这一证明本身实际上描述了二分法求根的过程.

证明 构造如下闭区间套, 令 $[a_0, b_0] = [a, b]$, $c_n = (a_n + b_n)/2$, 如果 $f(c_n) = 0$, 则结论成立, 否则 c_n 至少与 a_n 和 b_n 中的一个异号, 取该半个区间为 $[a_{n+1}, b_{n+1}]$. 由上述构造可知

$$[a_{n+1}, b_{n+1}] \subset [a_n, b_n], \quad |b_n - a_n| = \frac{b-a}{2^n} \rightarrow 0.$$

从而存在 $\xi \in \bigcap_{k=0}^{\infty} [a_n, b_n]$. 若 $f(\xi) \neq 0$, 不妨设 $f(\xi) = r > 0$, 则存在 $\delta > 0$, 在 $[\xi - \delta, \xi + \delta]$ 上 $f(x) > 0$ 恒成立. 取足够大的 n , 即可使 $[a_n, b_n] \subset [\xi - \delta, \xi + \delta]$, 其中 $f(a_n)f(b_n) < 0$, 与 f 在 $[\xi - \delta, \xi + \delta]$ 保号矛盾, 从而 $f(\xi) = 0$. ■

113 算法 (二分法求解零点) 对区间 n 等分, 对每个满足端点函数值异号的区间 $[a_k, a_{k+1}]$, 按照定理 112 证明中的方法二分, 直到满足 $b_n - a_n < \varepsilon$, 即与零点误差小于 ε 为止.

评注 二分法实现简单, 但是在高维情况下, 由于不再有“区间端点”的概念, 所以难以推广.

6.2 不动点法

114 定义 (压缩映射) 设 f 将 $E \subset \mathbf{R}$ 映射到 E 上. 称 f 为压缩映射, 若存在常数 $l < 1$, 对任意 $x, y \in E$ 成立

$$|f(x) - f(y)| \leq l|x - y|.$$

115 定理 (压缩映射定理) 设 φ 是 $[a, b]$ 上的连续压缩映射, 则存在唯一的 $x \in [a, b]$, 成立 $\varphi(x) = x$.

证明 对于唯一性. 设成立 $x_1 = \varphi(x_1)$, $x_2 = \varphi(x_2)$, 则

$$|x_1 - x_2| = |\varphi(x_1) - \varphi(x_2)| \leq l|x_1 - x_2| \Rightarrow x_1 = x_2.$$

对于存在性. 令 $F(x) = x - \varphi(x) \in \mathcal{C}[a, b]$. 若 $a = \varphi(a)$ 或 $\varphi(b)$, 则得证. 否则由于 f 映射到 $[a, b]$ 自身, 所以成立 $a < \varphi(a), \varphi(b) < b$. 则成立 $F(a)F(b) < 0$, 由定理 112 可知, 存在 $\xi \in [a, b]$, 成立 $F(\xi) = 0$, 即 $\xi = \varphi(\xi)$. ■

评注 此定理对于任意的完备度量空间都是成立的, 且条件中的连续性要求可以略去, 证明方法是构造迭代数列 $x_{n+1} = \varphi(x_n)$.

116 定理 (余项估计) 设 φ 是 $[a, b]$ 上压缩常数为 l 的连续压缩映射. 则数列

$$x_0 = a, \quad x_{n+1} = \varphi(x_n)$$

收敛至 x_* . 且有估计式

$$|x_n - x_*| \leq \frac{l^n}{1-l} |x_1 - x_0|.$$

证明 下证明 $\{x_n\}$ 为 Cauchy 序列.

$$|x_{k+1} - x_k| = |\varphi(x_k) - \varphi(x_{k-1})| \leq l|x_k - x_{k-1}| \leq \cdots \leq l^k |x_1 - x_0|.$$

所以对于任意的 n 和 $p > 0$, 成立

$$|x_{n+p} - x_n| \leq \sum_{k=0}^{p-1} |x_{n+k+1} - x_{n+k}| \leq |x_1 - x_0| \sum_{k=0}^{p-1} l^{n+k}.$$

由于 $l > 0$, 根据几何级数的性质, 成立

$$|x_{n+p} - x_n| \leq \frac{l^n}{1-l} |x_1 - x_0|.$$

当 $n \rightarrow \infty$ 时, R.H.S $\rightarrow 0$, 从而 $\{x_n\}$ 是 Cauchy 序列, 所以收敛. 令 $p \rightarrow \infty$, 即得估计式

$$|x_n - x_*| \leq \frac{l^n}{1-l} |x_1 - x_0|. \quad \blacksquare$$

117 定义 (局部收敛) 设 $\varphi(x)$ 有不动点 x_* , 则对于迭代法

$$x_{n+1} = \varphi(x_n). \quad (13)$$

如果存在 x_* 的某个领域 $O_\delta(x)$, 使得任意 $x_0 \in O_\delta(x)$, (13) 收敛至 x_* , 则称 (13) 局部收敛.

118 定理 (局部收敛的条件) 设 x_* 是 $\varphi \in \mathcal{C}^1$ 的不动点, 且 $|\varphi'(x_*)| < 1$, 则 (13) 在 x_* 处局部收敛.

证明 由于 φ' 是连续的, 所以存在 $O_\delta(x_*)$, 对任意的 $x \in O_\delta(x_*)$, 成立

$$|\varphi'(x)| < l < 1.$$

从而对于任意的 $x, y \in O_\delta(x_*)$, 根据中值定理, 成立

$$|\varphi(x) - \varphi(y)| = |\varphi'(\xi)||x - y| < l|x - y|.$$

即 $\varphi(x)$ 在 $O_\delta(x_*)$ 上是压缩映射, 从而对任意 $x_0 \in O_\delta(x_*)$, (13) 收敛至 x_* . ■

119 命题 (不局部收敛的条件) 设 x_* 是 $\varphi(x) \in \mathcal{C}^1[a, b]$ 的不动点, 若 $\varphi(x)$ 在 $[a, b]$ 上单调且 $|\varphi'(x_*)| \geq h > 1$, 则 $\varphi(x)$ 不局部收敛.

证明 由于 $\varphi'(x_*) > 1$ 且连续, 所以存在 $\delta > 0$, 使得在 $O_\delta(x_*)$ 中成立

$$|\varphi(x) - x_*| = |\varphi'(\xi)(x - x_*)| \geq h|x - x_*|.$$

所以对任意的 $0 < \varepsilon < \delta$, 对任意的 $x_0 \in O_\delta(x_*) \setminus \{x_*\}$, 迭代足够多次后成立 $|x_n - x_*| > \varepsilon$. 由于 $\varphi(x)$ 单调, 所以仅有 $x = x_*$ 满足 $\varphi(x) = x_*$. 从而若存在 $x_0 \in [a, b]$, 使迭代数列收敛至 x_* , 则对任意 $\varepsilon > 0$, 存在 n , 使得 $x_n \in O_\varepsilon(x_*) \setminus \{x_*\}$. 对于这样的 n , 继续迭代足够多次后, 成立 $|x_m - x_*| > \varepsilon$, 与收敛矛盾, 从而不存在 $x_0 \in [a, b]$, 使得迭代数列收敛. ■

120 定义 (p 阶收敛) 设 x_* 是 $\varphi(x)$ 的不动点, 记迭代误差 $e_k = x_k - x_*$, 若当 $k \rightarrow \infty$ 时, 成立

$$\frac{e_{k+1}}{e_k^p} \rightarrow C \neq 0.$$

则称 (13) p 阶收敛.

评注 此定义描述了迭代式收敛的速度.

121 定理 (p 阶收敛条件) 设 x_* 是迭代过程 $x_{n+1} = \varphi(x_n)$ 的不动点, 若对正整数 p , $\varphi^{(p)}$ 在 x_* 附加连续, 且成立

$$\varphi'(x_*) = \cdots = \varphi^{(p-1)}(x_*) = 0, \quad \varphi^{(p)}(x_*) \neq 0.$$

则 (13) 在 x_* 附近 p 阶收敛.

证明 在 x_* 处将 φ Taylor 展开即可. 存在 ξ 在 x_k 和 x_* 之间, 成立

$$\frac{x_{k+1} - x_*}{(x_k - x_*)^p} = \frac{\varphi^{(p)}(\xi)}{p!}. \quad \blacksquare \quad (14)$$

评注 经实践证明, 由于当 $k \rightarrow \infty$ 时, $\xi \rightarrow x_*$, 所以对于前 $p-1$ 阶导数为零的迭代式, 使用 (14) 来计算定义 120 中的常数是十分方便的.

122 算法 (不动点法) 对于方程 $f(x) = 0$, 将其变形为等价的 $x = \varphi(x)$ 的形式, 且 φ 满足在零点处局部收敛, 则可以利用迭代数列 $x_{n+1} = \varphi(x_n)$ 来求解方程的根.

123 算法 (Aitken Δ^2 加速法) 设 $\{x_n\}$ 为一收敛的迭代数列, 令

$$\bar{x}_{k+1} = x_k - \frac{(x_{k+1} - x_k)^2}{x_k - 2x_k x_{k+1} + x_{k+2}} = x_k - \frac{(\Delta x_k)^2}{\Delta^2 x_k}.$$

$\{\bar{x}_k\}$ 收敛得比 $\{x_k\}$ 更快, 即满足

$$\lim_{k \rightarrow \infty} \frac{\bar{x}_{k+1} - x_*}{x_k - x_*} = 0.$$

评注 这一加速法的思路在于

$$x_{n+1} - x_* = \varphi(x_n) - \varphi(x_*) = \varphi'(\xi)(x_n - x_*).$$

当区间足够小以至 φ' 变化不大时, 可以近似地将 $\varphi'(\xi)$ 看作常量. 另外, 这一做法实际上只是数据的后处理, 实际上如果我需要的是 \bar{x}_{n+1} , 那无需计算 $\bar{x}_1, \dots, \bar{x}_n$, 而只需利用 x_n, x_{n+1}, x_{n+2} 来计算 \bar{x}_{n+1} 即可.

124 算法 (Steffensen 迭代法)

$$\begin{cases} y_k = \varphi(x_k), & z_k = \varphi(y_k), \\ x_{k+1} = x_k - \frac{(y_k - x_k)^2}{z_k - 2y_k + x_k}. \end{cases}$$

评注 这一算法实际上只是将 Aitken 加速法中的 \bar{x}_n 也加入了迭代序列中.

125 定理 (Steffensen 迭代法) 若 x_* 是迭代函数

$$\psi = x - \frac{(\varphi - x)^2}{\varphi \circ \varphi - 2\varphi + x}$$

的不动点, 则 x_* 也是 φ 的不动点. 反之, 若 x_* 是 φ 的不动点, φ'' 存在, 且 $\varphi'(x_*) \neq 1$, 则 x_* 也是 ψ 的不动点, 且 Steffensen 迭代法二阶收敛.

评注 这一定理表明, 就算 φ 所对应的迭代序列不收敛, 但是对应的 Steffensen 迭代法的迭代序列还是有可能收敛的.

6.3 Newton 法

126 定理 (Newton 法) 设 x_* 是 $f(x)$ 的零点且 $f'(x_*) \neq 0$, 则

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

当 $f'(x_*) \neq 0$ 时, 在 x_* 处局部平方收敛, 否则线性收敛.

评注 Newton 法的动机在于在零点附近, 用一个线性函数来近似原函数, 即

$$f(x) \approx f(x_k) + f'(x_k)(x - x_k) \Rightarrow x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

则该线性方程的解是原方程的一个近似解. 证明只需利用 定理 121 即可.

127 算法 (Newton 法) Newton 法求根的步骤如下:

1. 选定初值 x_0 , 计算 $f(x_0)$ 和 $f'(x_0)$;
2. 按照公式 $x_{n+1} = x_n - f_n/f'_n$, 并计算 $f(x_{n+1})$ 和 $f'(x_{n+1})$;
3. 如果 x_n 满足 $|\delta| < \varepsilon_1$ 或 $|f| < \varepsilon_2$, 则终止迭代, 以 x_n 作为所求根, 否则进入步骤 4. 此处 $\varepsilon_1, \varepsilon_2$ 是允差, 而

$$\delta = \begin{cases} |x_n - x_{n-1}|, & |x_1| < C \\ \frac{|x_n - x_{n-1}|}{x_n}, & |x_1| \geq C \end{cases}$$

其中 C 是取绝对误差或相对误差的控制常数, 一般取 $C = 1$.

4. 如果迭代次数达到预先规定的 N , 或 $f'_n = 0$, 则方法失败, 否则继续迭代.

128 定理 (重根情况的 Newton 法) 设 x_* 是 $f(x)$ 的 m 重根, 则根据

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

所得的迭代序列, 在 x_* 附近平方收敛.

129 定理 (重根情况的 Newton 法) 设 x_* 是 $f(x)$ 的根, 则根据

$$\varphi(x) = x - \frac{ff'}{(f')^2 - ff''}$$

所得的迭代序列在 x_* 附近平方收敛.

评注 这一函数实际上是令

$$\mu(x) = \frac{(x - x_*)g(x)}{mg(x) + (x - x_*)g'(x)}.$$

在考虑 $\varphi(x) = x - \mu/\mu'$ 所得到的. 其缺点在于涉及了 f 的二阶导数.

130 定理 (Newton 法的收敛区间) 设 $f \in \mathcal{C}^2[a, b]$, 且满足 $f(a)f(b) < 0$, 在 $[a, b]$ 上 $f'(x) \neq 0$, 且 f'' 不变号, 以及初值 x_0 使得 $f(x_0)f''(x_0) > 0$ 成立. 则 Newton 法所得序列单调收敛于唯一根 x_* .

131 命题 (方程组情况的推广) 对于 $\mathbf{F}: \mathbf{R}^n \rightarrow \mathbf{R}^n$, 设 \mathbf{x}_0 是它的近似根, 则可以构造方程组的 Newton 方法

$$\mathbf{x}_{k+1} = \mathbf{x}_k - [\mathbf{F}'(\mathbf{x}_k)]^{-1}\mathbf{F}(\mathbf{x}_k).$$

其中 \mathbf{F}' 为 \mathbf{F} 的 Jacobi 矩阵.

评注 这一算法的最大工作量在于求解 Jacobi 矩阵的逆，即求解线性方程组.

7 解线性方程组的直接/间接方法

7.1 绪论

本章讨论的问题是线性方程组

$$A\mathbf{x} = B, \quad A \in \mathbf{R}^{m \times n}, \mathbf{x} \in \mathbf{R}^n, B \in \mathbf{R}^m \quad (15)$$

的求解. 讨论问题在于

1. (15) 是否可解.
2. 若可解, 则如何求解.
3. 若无解, 则如何求近似.

其中不可解的情况对应 B 为带有误差的测量数据, 求解 X 使得 $A\mathbf{x}$ 在特定含义下距离 B 尽可能接近. 例如考虑最小二乘法, 即求解

$$\mathbf{x}^* = \arg \min \frac{1}{2} \|A\mathbf{x} - B\|_2^2 \quad (16)$$

这一问题实际上就是求解极值.

132 定理 (可解性条件) 设 $A = (\alpha_1 \dots \alpha_n)$, 其中 α_i 为列向量, 则以下命题等价

1. $A\mathbf{x} = B$ 有解.
2. $r(A) = r(\bar{A})$.
3. $B \in \text{span}\{\alpha_1, \dots, \alpha_n\}$.

其中 $r(A)$ 为 A 的秩, \bar{A} 为 A 的增广矩阵.

133 引理 (梯度的定价定义) f 的梯度 ∇f 为对任意方向向量 v 都满足如下条件的函数

$$\langle \nabla f, v \rangle = \mathbf{D}_v f = \left. \frac{d}{d\varepsilon} f(x + \varepsilon v) \right|_{\varepsilon=0}.$$

134 定理 $f(\mathbf{x}) = \frac{1}{2} \|A\mathbf{x} - B\|_2^2$ 的梯度可以被表示为

$$\nabla f(\mathbf{x}) = A^T(A\mathbf{x} - B).$$

证明 下利用引理 133 证明. 已知

$$f(\mathbf{x} + \varepsilon v) = \frac{1}{2} \langle A\mathbf{x} - B + \varepsilon v, A\mathbf{x} - B + \varepsilon v \rangle = \frac{1}{2} \|A\mathbf{x} - B\|_2^2 + \varepsilon \langle A\mathbf{x} - B, Av \rangle + \frac{1}{2} \varepsilon^2 \|Av\|_2^2.$$

所以, 在该处对任意方向向量 v 成立

$$\langle \nabla f(\mathbf{x}), v \rangle = \left. \frac{d}{d\varepsilon} f(\mathbf{x} + \varepsilon v) \right|_{\varepsilon=0} = \langle A\mathbf{x} - B, Av \rangle = \langle A^T(A\mathbf{x} - B), v \rangle.$$

比较上式的 L.H.S 和 R.H.S, 可以得

$$\nabla f = A^T(A\mathbf{x} - B). \quad \blacksquare$$

todo $m \ll n$ 情况

8 附录

8.1 不等式

135 引理 (排序不等式) 对于满足下述条件的 $\{a_n\}, \{b_n\}$,

$$0 \leq a_1 \leq a_2 \leq \cdots \leq a_n$$

$$0 \leq b_1 \leq b_2 \leq \cdots \leq b_n$$

则同序相乘求和值最大, 逆序最小, 即

$$\sum_{i=1}^n a_i b_i \geq \sum_{i=1}^n a_i b_{k_i} \geq \sum_{i=1}^n a_i b_{n-i+1}$$

136 引理 (算数-几何均值不等式)

$$(a_1 a_2 \cdots a_n)^{1/n} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}$$

当且仅当 $a_1 = a_2 = \cdots = a_n$ 时等号成立.

证明 因为有齐次性, 所以不妨设 $\prod a_i = 1$, 并令

$$a_1 = \frac{\alpha_1}{\alpha_2}, \quad \dots, \quad a_{n-1} = \frac{\alpha_{n-1}}{\alpha_n}, \quad a_n = \frac{\alpha_n}{\alpha_1}$$

则只需证明下式即可.

$$\frac{\alpha_1}{\alpha_2} + \cdots + \frac{\alpha_n}{\alpha_1} \geq n$$

不妨设 $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$, 则根据排序不等式

$$\text{L.H.S} \geq \alpha_1 \frac{1}{\alpha_1} + \cdots + \alpha_n \frac{1}{\alpha_n} = n \quad \blacksquare$$

8.2 积分相关公式

137 引理 (分部积分) 设 $u, v \in \mathcal{C}^{n+1}[a, b]$, 则成立

$$\int_a^b uv^{(n+1)}dx = [uv^{(n)} - u'v^{(n-1)} + \cdots + (-1)^n u^{(n)}v] \Big|_a^b + (-1)^{n+1} \int_a^b u^{(n+1)}v dx.$$

8.3 特殊函数

138 定义 (处处连续且不可导) 称

$$f(x) = \sum_{n=0}^{\infty} a^n \cos(b^n \pi x)$$

为 Weierstrass 函数, 其中 $0 < a < 1$, b 为正奇数, 且满足

$$ab > 1 + \frac{3}{2}\pi.$$

它处处连续且处处不可导.

8.4 Euler-Maclaurin 公式

绪论 此节内容的主要是对 Apostol, T. M. (1 May 1999). "An Elementary View of Euler's Summation Formula" 的翻译和整理.

8.4.1 广义 Euler 常数

绪论 本章节仅考虑在 $[1, \infty)$ 上满足 $f > 0$, 且严格单调递减的函数.

139 定义 (d_n) 定义 d_n 为用积分近似离散求和的误差, 即

$$d_n = \sum_{k=1}^{n-1} f(k) - \int_1^n f(x) dx \quad (17)$$

评注 注意, d_n 的定义中, k 仅遍历 $[1, n-1]$, 而不包含 n . 图 3 为 d_n 的几何解释, 其

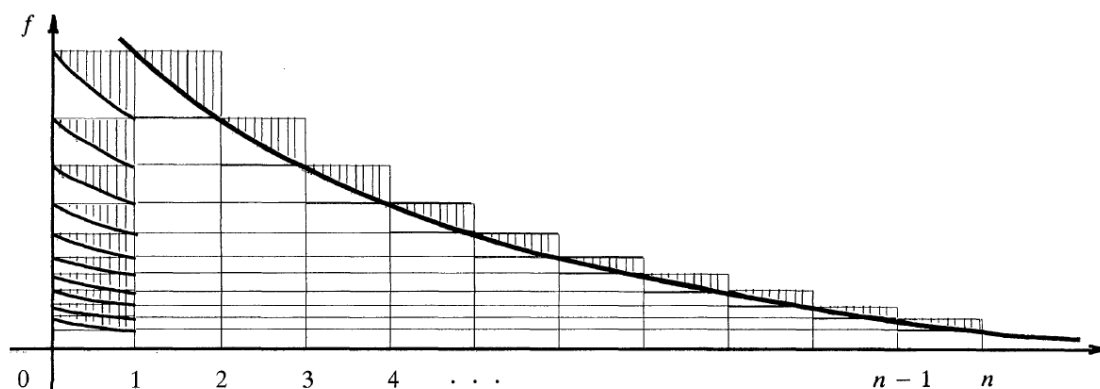


图 3: d_n 的几何解释

中在曲线上方的阴影部分即为 d_n , 可以将它们统一移动到图像的最左侧.

140 定义 (广义 Euler 常数) 已知 $0 < d_n < d_{n+1} < f(1)$, 所以可以定义关于 f 的广义 Euler 常数 $C(f)$ 为

$$C(f) = \lim_{n \rightarrow \infty} d(n)$$

141 命题 (余项估计) $0 < C(f) - d_n < f(n)$.

142 定理 设 f 在 $[1, \infty)$ 上为正且严格单调递减, 则存在一系列 $\{E_f(n)\}$, 满足 $0 < E_f(n) < f(n)$, 成立

$$\sum_{k=1}^n f(k) = \int_1^n f dx + C(f) + E_f(n), \quad n = 2, 3, \dots \quad (18)$$

评注 只需令 $L.H.S = \sum_{k=1}^{n-1} f(k) + f(n)$, 在根据 d_n 的定义即可证明. 这一定理意味着求和与积分之间的误差仅仅为一个与 f 有关的常数以及一个比 $f(n)$ 还要小的正数. 从而如果 $f(n) \rightarrow 0$, 则 $E_f(n) \rightarrow 0$, 此时成立

$$C(f) = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n f(k) - \int_1^n f(x) dx \right). \quad (19)$$

143 定义 (经典 Euler 常数) 取 $f(x) = 1/x$, 它满足 $\lim_{x \rightarrow \infty} f(x) = 0$, 则有

$$\gamma = C\left(\frac{1}{x}\right) = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \ln n \right).$$

8.4.2 Euler 求和公式

绪论 在此章节不再限制 f 需要为正且单调减, 在目前仅要求 $f \in \mathcal{R}[1, n]$ (在之后会添加诸如连续可导等条件). 在此基础上重新定义 d_n (与上一节保持一致) 并推导出类似于上一节中的结论.

144 定义 (d_n) 定义 d_n 为

$$d_n = \sum_{k=1}^{n-1} I(k) = \sum_{k=1}^{n-1} \int_k^{k+1} \{f(k) - f(x)\} dx. \quad (20)$$

评注 可以发现这一定义和上一节中的定义, 在 $f > 0$ 且单调减的情况下是一致的. 实际上这一定义只是将每一点处的差值积分起来而已.

145 命题 对于任意常数 c , 由于成立 $dx = d(x+c)$, 所以可以取 $c = -(k+1)$, 并分部积分, 则有

$$I(k) = \int_k^{k+1} (x - [x] - 1) f'(x) dx.$$

从而

$$d_n = \int_1^n (x - [x]) f'(x) dx + f(1) - f(n).$$

146 定理 (一次导数形式的 Euler 求和公式) 对于任意 $f \in \mathcal{C}^1[1, n]$, 成立

$$\sum_{k=1}^n f(k) = \int_1^n f(x) dx + \int_1^n (x - [x]) f'(x) dx + f(1). \quad (21)$$

评注 对于误差项中的 f' , 如果它不变号, 即 f 单调的情况下, 可以考虑让它乘一个变号的因数, 从而使得误差更小. 例如, 可以用 $x - [x] - \frac{1}{2}$ 来代替 $x - [x]$. 则有了如下推论.

147 推论 对于任意 $f \in \mathcal{C}^1[1, n]$, 定义一次 Bernoulli 函数为

$$P_1(x) = \begin{cases} x - [x] - \frac{1}{2}, & x \notin \mathbf{N} \\ 0, & x \in \mathbf{N}. \end{cases}$$

则可以重写 (21) 为

$$\sum_{k=1}^n f(k) = \int_1^n f(x)dx + \int_1^n P_1(x)f'(x)dx + \frac{1}{2}(f(n) + f(1)). \quad (22)$$

若假设 $\int_1^\infty P_1 f dx$ 存在, 那么上式可以继续写为

$$\begin{aligned} \sum_{k=1}^n f(k) &= \int_1^n f(x)dx + C(f) + E_f(n) \\ C(f) &= \frac{1}{2}f(1) + \int_1^\infty P_1 f' dx \\ E_f(n) &= \frac{1}{2}f(n) - \int_n^\infty P_1 f' dx \end{aligned}$$

评注 注意, 在此处 $C(f)$ 和 $E_f(n)$ 的定义和之前依然是一致的. 并且只需要 $f(x)$ 在 $[1, \infty)$ 上反常绝对可积, 就可以保证 $\int_1^\infty P_1 f dx$ 的存在性.¹³

148 推论 (Euler 常数) 将 $f(x) = 1/x$, 代入推论 147, 得经典 Euler 常数 γ 为

$$\gamma = \frac{1}{2} = \int_1^\infty \frac{P_1(x)}{x} dx.$$

8.4.3 余项分析

绪论 这一章节通过对

$$\int_n^{n+1} P_1(x)f(x)dx \quad (23)$$

分部积分来进一步分析余项, 并为此引入了更高次的 Bernoulli 函数.

149 定义 (二次 Bernoulli 函数) 定义二次 Bernoulli 函数 $P_2(x)$ 为

$$P_2(x) = 2 \int_0^x P_1(t)dt + \frac{1}{6}.$$

评注 如此定义二次 Bernoulli 函数的动机来自于对 (23) 的进一步分部积分. 这样就需要有

$$P_2(x) = 2 \int_0^x P_1(t)dt + c$$

¹³我暂时并没有明白, 是什么条件保证了 $f(n) \rightarrow 0$, 从而 $C(f) = \lim_{n \rightarrow \infty} d_n$ 存在.

之所以要有系数 2 是为了之后公式的简洁. 由于 $P_1(x)$ 有周期性, 从而

$$P_2(x+1) - P_2(x) = 2 \int_x^{x+1} P_1(t) dt = 0,$$

即 $P_2(x)$ 也有周期性. 为了让按照同样方式定义的 P_3 也具有周期性, 所以希望有

$$\int_0^1 P_2(x) dx = 0 \Rightarrow c = \frac{1}{6}. \quad \blacksquare$$

150 定理 (二次导数形式的 Euler 求和公式) 设 $f \in \mathcal{C}^2[1, n]$, 成立

$$\sum_{k=1}^n f(k) = \int_1^n f dx - \frac{1}{2} \int_1^n P_2 f'' dx + \frac{1}{2} P_2(0)(f'(n) - f'(1)) + \frac{1}{2}(f(n) + f(1)).$$

并且, 如果 f'' 在 $[1, \infty)$ 上反常绝对可积, 则有

$$\begin{aligned} \sum_{k=1}^n f(k) &= \int_1^n f dx + C(f) + E_f(n) \\ C(f) &= \frac{1}{2} f(1) - \frac{1}{2} P_2(0) f'(1) - \frac{1}{2} \int_1^\infty P_2 f'' dx \\ E_f(n) &= \frac{1}{2} f(n) + \frac{1}{2} P_2(0) f'(n) + \frac{1}{2} \int_n^\infty P_2 f'' dx. \end{aligned}$$

8.4.4 Bernoulli 数和 Euler 求和公式的一般形式

151 定义 定义 Bernoulli 周期函数为

$$P_k(x) = k \int_0^x P_{k-1}(t) dt + B_k, \quad k \geq 2 \quad (24)$$

其中 B_k 称为 **Bernoulli 数**, 它使得下式成立

$$\int_0^1 P_k(x) dx = 0.$$

显然 $P_k(x)$ 在 $k \geq 2$ 的时候在 $[0, 1]$ 上是 k 次多项式, 称它为 **Bernoulli 多项式**.

152 定理 (等价定义) Bernoulli 数和 Bernoulli 多项式的另一种常见定义为,

$$\begin{aligned} B_0 &= 1, \quad \sum_{k=0}^{m+1} \binom{m+1}{k} B_k = 0 \\ B_n(x) &= \sum_{k=0}^n \binom{n}{k} B_{n-k} x^k. \end{aligned}$$

评注 对于大于 1 的奇数 k , $B_k = 0$, 且有

$$|B_{2k}(x)| \leq |B_{2k}|, \quad |B_{2k+1}(x)| \leq (2k+1)|B_{2k}|.$$

153 定理 (Euler 求和公式的一般形式) 设 $f \in \mathcal{C}^{2m+1}[1, n]$, 成立

$$\begin{aligned} \sum_{k=1}^n f(k) &= \int_1^n f dx + \frac{1}{(2m+1)!} \int_1^n P_{2m+1}(x) f^{(2m+1)}(x) dx \\ &\quad + \sum_{r=1}^m \frac{B_{2r}}{(2r)!} (f^{(2r-1)}(n) - f^{(2r-1)}(1)) + \frac{1}{2} (f(1) + f(n)). \end{aligned}$$

且若 $f^{(2m+1)}$ 在 $[1, \infty)$ 上反常绝对可积, 则成立

$$\begin{aligned} \sum_{k=1}^n f(k) &= \int_1^n f dx + C(f) + E_f(n) \\ C(f) &= \frac{1}{2} f(1) - \sum_{r=1}^m \frac{B_{2r}}{(2r)!} f^{(2r-1)}(1) + \frac{1}{(2m+1)!} \int_1^\infty P_{2m+1} f^{(2m+1)} dx \\ E_f(n) &= \frac{1}{2} f(n) + \sum_{r=1}^m \frac{B_{2r}}{(2r)!} f^{(2r-1)}(n) - \frac{1}{(2m+1)!} \int_n^\infty P_{2m+1} f^{(2m+1)} dx \end{aligned}$$