# 基础知识

# 1.1 二进制十进制转换

例如  $(53.7)_{10} = (110101.1\overline{0110})$ , 小数点左边倒序, 小数点右边正序。

$53 \div 2 = 26 \cdots 1$	$6 \div 2 = 3 \cdots 0$	$0.7 \times 2 = 0.4 + 1$	$0.6 \times 2 = 0.2 + 1$
$26 \div 2 = 13 \dots 0$	3 ÷ 2 = 1 1	$0.4 \times 2 = 0.8 \pm 0$	$0.2 \times 2 = 0.4 + 0$

$$13 \div 2 = 6 \cdots 1$$
  $1 \div 2 = 0 \cdots 1$   $0.8 \times 2 = 0.6 + 1$   $0.4 \times 2 = 0.8 + 0$ 

# 1.2 浮点数、误差、加减法

标准化 IEEE 浮点数  $+1.bbb \cdots b \times 2^p$ , IEEE 最近舍入法则: 对 53 位, 如果为 0,则舍去,如果为1,则进位。舍入误差的计算

精度	符号	指数	尾数	$fl(8.3) = 8.3 + [1 - (0.\overline{1001})_2] \times 2^{-52} \times 2^3$
单精度	1	8	23	$= 8.3 + 2^{-49} - (0.0\overline{1001})_2 \times 2 \times 2^{-49}$
双精度	1	11	52	$= 8.3 + 2^{-49} - 0.3 \times 2 \times 2^{-49}$
长双精度	1	15	64	- 0.3 + 2 - 0.3 \ 2 \ \ 2

加减法时需要根据大的数对齐,再用0补齐到小的数对应的最后一位。

$$(8.3)_{10} = (1000.0\overline{1001})_2 = 1.0000\overline{1001} \times 2^3 = 1. \dots 10011010 \times 2^3$$
  
 $(7.3)_{10} = (111.0\overline{1001})_2 = 1.110\overline{1001} \times 2^2 = 1. \dots 00110011 \times 2^2$ 

(向 52 位之后的位置补齐)  $(8.3)_{10}$  -  $(7.3)_{10}$  =  $1.00001001 \cdots 100110100 \times 2^3$ 

(对齐,保留 52 位之后的数字) - 0.11101001 ··· 100110011 × 23

(52 位之后的数字也要进行计算) = 0.00100000 ··· 000000001 × 23

(移动到 0 位为 1, 舍入到 52 位) = 1.00000000 ··· 00000100 × 20

# 解方程

# 2.1 二分法、误差

初始区间 [a,b] 使 f(a)f(b) < 0,取中点 c = (a+b)/2 若 f(a)f(c) < 0,下一迭 代区间为 [a,c],若 f(c)f(b) < 0,下一迭代区间为 [c,b],输出近似解为  $x_a = (a+b)/2$ (线性收敛)。近似解  $x_a$  的前向误差为  $|r-x_a|$ ,后向误差为  $|f(x_a)|$ 。二分法满足

$$|r-x_a| < rac{b-a}{2^{n+1}}$$
 误差放大因子  $\kappa = rac{ ext{ ext{ ext{ ext{H}}}} ext{ ext{ ext{H}}} ext{ ext{ in}} ext{ in} ext{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{\text{H}}}}}} ext{\text{\text{\text{\text{\text{\text{\text{\text{\text{H}}}}} ext{\text{2}}}\text{\texi}\text{\text{$ 

## 2.2 不动点迭代

不动点迭代将方程 f(x) = 0 变换为 g(x) = x, 满足条件可以达到线性收敛。

$$x_{i+1} = g(x_i)$$
  $\lim_{i \to \infty} \frac{e_{i+1}}{e_i} = S = |g'(r)| < 1$   $x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)}$ 

牛顿方法,可以达到二次收敛,出现m>2重根时变为线性收敛。

$$\lim_{i \to \infty} \frac{e_{i+1}}{e_i^2} = \left| \frac{f''(r)}{2f'(r)} \right| \qquad \lim_{i \to \infty} \frac{e_{i+1}}{e_i} = \frac{m-1}{m} \qquad x_{i+1} = x_i - \frac{mf(x_i)}{f'(x_i)}$$

如果  $0 = f(r) = f'(r) = \cdots = f^{(m-1)}(r)$ ,但是  $f^{(m)}(r) \neq 0$ ,则 r 为 m 重根,可 用改进的牛顿方法,恢复二次收敛速度。

#### 解方程组 3

### 3.1 高斯消元法、LU 分解

高斯消元的操作次数  $\frac{2}{3}n^3$ , LU 分解的操作次数  $\frac{2}{3}n^3$ , 每次回代的操作次数:  $n^2$ 。

### 3.2 PA=LU 分解

部分主元:每一行根据第一列元素从大到小排序,避免湮灭问题。

$$\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \qquad P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 \\ \frac{2}{3} & 1 \end{bmatrix}$$

$$P = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

$$L = \begin{bmatrix} 1 & 0 \\ \frac{2}{3} & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \rightarrow \textcircled{1} \leftrightarrow \textcircled{2} \rightarrow \begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \rightarrow \textcircled{2} - \textcircled{1} \times \frac{2}{3} \rightarrow \begin{bmatrix} 3 & 2 \\ 0 & \frac{5}{3} \end{bmatrix} = U$$

$$Lc = Pb \Rightarrow \begin{bmatrix} 1 & 0 \\ \frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} \qquad Ux = c \Rightarrow \begin{bmatrix} 3 & 2 \\ 0 & \frac{5}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

# 3.3 迭代方法、误差

要求矩阵 A 为严格对角占优矩阵, $|a_{ii}|>\sum_{i\neq j}|a_{ij}|$ ,D 表示 A 的主对角线矩阵, L 表示 A 的下三角矩阵,U 表示 A 的上三角矩阵。

雅可比方法, 高斯-赛德尔方法, 连续过松弛  $\omega > 1$ ,  $\omega = 1$  时, 退化

$$\begin{split} x_{k+1} &= D^{-1}(b - (L+U)x_k), \qquad x_{k+1} &= D^{-1}(b - Ux_k - Lx_{k+1}) \\ x_{k+1} &= (\omega L + D)^{-1}[(1-\omega)Dx_k - \omega Ux_k] + \omega (D + \omega L)^{-1}b \end{split}$$

近似解 
$$x_a$$
 的余项为  $r=b-Ax_a$ ,前向误差为  $\|r\|_{\infty}$ ,后向误差为  $\|b-Ax_a\|_{\infty}$ 。

误差放大因子可能具有条件数的数量级,||A||∞ 为每行元素绝对值之和的最大值。

误差放大因子 
$$\kappa = \frac{\|x - x_a\|_{\infty}/\|x\|_{\infty}}{\|r\|_{\infty}/\|b\|_{\infty}}$$
 条件数  $cond = \|A\|_{\infty} \cdot \|A^{-1}\|_{\infty}$ 

#### 插值 4

# 4.1 拉格朗日插值

对 
$$n$$
 个数据点  $(x_1,y_1),\cdots,(x_n,y_n)$ ,有  $n-1$  阶多项式

$$L_k(x) = \frac{(x-x_1)\cdots(x-x_{k-1})(x-x_{k+1})\cdots(x-x_n)}{(x_k-x_1)\cdots(x_k-x_{k-1})(x_k-x_{k+1})\cdots(x_k-x_n)}$$

$$P_{n-1}(x) = y_1 L_1(x) + \dots + y_n L_n(x)$$

$$P_2(x) = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_3)(x - x_1)}{(x_2 - x_3)(x_2 - x_1)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}$$

### 4.2 牛顿差商

$$\begin{aligned} x_1 & & f[x_k] = f(x_k) \\ & & f[x_1] & & f[x_k x_{k+1}] = \frac{f[x_{k+1}] - f[x_k]}{x_{k+1} - x_k} \\ x_2 & & f[x_2] & & f[x_1 x_2 x_3] \\ & & & f[x_2 x_3] & & f[x_k x_{k+1} x_{k+2}] = \frac{f[x_{k+1} x_{k+2}] - f[x_k x_{k+1}]}{x_{k+2} - x_k} \\ x_3 & & f[x_3] & & & P_{n-1}(x) = \sum_{i=1}^n f[x_1 \cdots x_i](x - x_1) \cdots (x - x_{i-1}) \end{aligned}$$

# 4.3 切比雪夫插值

在区间 [a,b] 内,切比雪夫插值节点,使以不等式在 [a,b] 内成立

$$x_i = \frac{a+b}{2} + \frac{b-a}{2} \cos \frac{(2i-1)\pi}{2n} \quad i = 1, 2, \dots, n \quad \left| (x-x_1) \cdots (x-x_n) \right| \leq \frac{\left(\frac{b-a}{2}\right)^n}{2^{n-1}}$$

### 4.4 插值误差

插值误差公式,以及误差上界,其中 $x_{min} < c < x_{max}$ 

$$|f(x) - P(x)| = \frac{(x - x_1) \cdots (x - x_n)}{n!} f^{(n)}(c) \le \frac{|(x - x_1) \cdots (x - x_n)|}{n!} |f^{(n)}(x_i)|_{max}$$

# 最小二乘法

### 5.1 不一致方程组、误差

最小二乘解 $\bar{x}$ 的余项为 $r = b - A\bar{x}$ ,后向误差(二范数误差,欧式长度)为

对于不一致系统  $Ax = b \Rightarrow A^T A \bar{x} = A^T b$  $||r||_2 = \sqrt{r_1^2 + \dots + r_m^2}$ 平方误差,均方根误差,以及三者直接的关系分别为

$$SE = r_1^2 + \dots + r_m^2$$
  $RMSE = \sqrt{\frac{r_1^2 + \dots + r_m^2}{m}}$   $RMSE = \frac{\sqrt{SE}}{\sqrt{m}} = \frac{\|r\|_2}{\sqrt{m}}$ 

### 5.2 数据拟合

多项式(线性)模型,周期型数据,指数模型的线性化。

$$y = c_1 + c_2 t + c_3 t^2 + \cdots$$

$$y = c_1 + c_2 \cos 2\pi t + c_3 \sin 2\pi t$$

$$y = c_1 e^{c_2 t} \Rightarrow \ln y = \ln c_1 + c_2 t$$

$$y = c_1 t e^{c_2 t} \Rightarrow \ln y - \ln t = \ln c_1 + c_2 t$$

### 5.3 QR 分解

使用格拉姆-斯密特正交找出  $A=(A_1|\cdots|A_n)$  的消减 QR 分解

$$y_j = A_j - q_1(q_1^T A_j) - q_2(q_2^T A_j) - \dots - q_{j-1}(q_{j-1}^T A_j) \quad q_j = \frac{y_j}{\|y_j\|_2}$$

$$\begin{cases} r_{jj} = \left\| y_j \right\|_2 \\ r_{ij} = q_i^T A_j \end{cases} \qquad A = (q_1 | \cdots | q_n) \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \end{bmatrix} = QR$$

对  $A = (A_1 | \cdots | A_n)$  补齐 (m-n) 个线性无关向量, 计算  $q_{n+1}, \cdots, q_{n+m}$ , 得到完 全 QR 分解,并通过回代  $R\bar{x}=Q^Tb$ ,求解最小二乘问题,虚线以下部分直接扔掉,可 以直接得到误差  $||e||_2 = \sqrt{d_{n+1} + \cdots d_{n+m}}$ 

$$A = (q_1|\cdots|q_m) \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} = QR \quad \begin{bmatrix} r_{11} & \cdots & r_{1n} \\ & \ddots & \vdots \\ & & r_{nn} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = Q^T \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \\ \vdots \\ d_m \end{bmatrix}$$

#### 数值微分积分 6

### 6.1 数值微分

二点前向差分公式(一阶导一阶公式),三点中心差分公式(一阶导二阶公式)  $f'(x) = \frac{f(x+h) - f(x)}{h} - \frac{h}{2}f''(c) \quad f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{h^2}{6}f'''(c)$ 

$$f''(x) = \frac{f(x-h) - 2f(x) + f(x+h)}{h^2} - \frac{h^2}{12}f^{(4)}(c)$$

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{3!}f'''(x) + \dots + \frac{h^k}{k!}f^{(k)}(x) + \dots$$

# 6.2 理查德外推

对 n 阶近似公式  $F_n(x)$  外推, 至少可以得到 (n+1) 阶近似 Q 的公式 对三点中心差分公式的外推结果为 4 阶

# 6.3 数值积分、求积法则、精度

梯形法则 (精度 1),中点法则 (精度 1),辛普森法则 (精度 3),及其复合法则

梯形法则(精度 1),甲点法则(精度 1),辛普森法则(精度 3),及具复合法 
$$\int_{x_0}^{x_1} f(x) \, \mathrm{d}x = \frac{h}{2} (y_0 + y_1) - \frac{h^3}{12} f''(c) \qquad \qquad h = x_1 - x_0$$
 
$$\int_a^b f(x) \, \mathrm{d}x = \frac{h}{2} (y_0 + y_m + 2 \sum_{i=1}^{m-1} y_i) - \frac{(b-a)h^2}{12} f''(c) \qquad h = \frac{b-a}{m}$$
 
$$\int_{x_0}^{x_1} f(x) \, \mathrm{d}x = h f(x_0 + \frac{h}{2}) + \frac{h^3}{24} f''(c) \qquad h = x_1 - x_0$$
 
$$\int_a^b f(x) \, \mathrm{d}x = h \sum_{i=0}^{m-1} f(x_i + \frac{h}{2}) + \frac{(b-a)h^2}{24} f''(c) \qquad h = \frac{b-a}{m}$$

$$\int_{x_0}^{x_2} f(x) dx = \frac{h}{3} (y_0 + 4y_1 + y_2) - \frac{h^5}{90} f^{(4)}(c) \qquad h = x_1 - x_0 = x_2 - x_1$$

$$\int_{a}^{b} f(x) dx = \frac{h}{3} (y_0 + y_{m2} + 4 \sum_{i=1}^{m} y_{2i-1} + 2 \sum_{i=1}^{m-1} y_{2i}) - \frac{(b-a)h^4}{180} f^{(4)}(c) \quad h = 0$$

对形如下的积分公式代入  $f(x) = 1, x, x^2, \cdots x^n$ ,来确定  $w_0, w_1, \cdots, w_n$ ,使得其有 最高精度 n,成为 n阶积分法则。对积分法则 I,有积分误差公式。

$$\int_{a}^{b} f(x) dx = w_{0} f(x_{0}) + w_{1} f(x_{1}) + \dots + w_{n} f(x_{n})$$

$$\int_{a}^{b} f(x) dx - I = \left| \frac{f^{n+1}(c)}{(n+1)!} \int_{a}^{b} (x - x_{0})(x - x_{1}) \cdots (x - x_{n}) dx \right|$$

### 6.4 龙贝格积分

添加数据进行扩展,直到达到指定的精度,龙贝格积分是对复合梯形法则的外推。

$$h_j = \frac{b-a}{2^{j-1}} \qquad R_{11} = \frac{h_1}{2} (f(a) + f(b)) \qquad R_{j1} = \frac{1}{2} R_{j-1,1} + h_j \sum_{i=1}^{2^{j-1}} f(a + (2i-1)h_j)$$

 $R_{j+1,1}$  是对  $R_{j1}$  步长减半的结果(向下), $R_{j,k+1}$  是对  $R_{jk}$  理查德外推结果(向右)。

### 6.5 高斯积分

对于积分区间不在 [-1,1] 上的积分,利用换元归一化到 [-1,1]。

$$\int_{-1}^1 f(x) \, \mathrm{d}x \approx \sum_{i=1}^n c_i f(x_i) \qquad \int_a^b f(x) \, \mathrm{d}x = \int_{-1}^1 f\left(\frac{(b-a)t+b+a}{2}\right) \frac{b-a}{2} \, \mathrm{d}t$$

# 常微分方程

## 7.1 初值问题,方程组,高阶

欧拉方法,中点方法,显示梯形方法,k阶泰勒方法,偏导展开  $w_{i+1} = w_i + hf(t_i, w_i) + O(h), \quad w_{i+1} = w_i + hf\left(t_i + \frac{h}{2}, w_i + \frac{h}{2}f(t_i, w_i)\right) + O(h^2)$  $W_{i+1} = W_i + \frac{h}{2}f(t_i, W_i) + \frac{h}{2}f(t_i + h, W_i + hf(t_i, W_i)) + O(h^2)$  $w_{i+1} = w_i + h f(t_i, w_i) + \frac{h^2}{2} f'(t_i, w_i) + \dots + \frac{h^k}{k!} f^{(k-1)}(t_i, w_i) + O(h^{+1})$ 

$$w_{i+1} = w_i + hf(t_i, w_i) + \frac{h^2}{2} \left[ \frac{\partial f}{\partial t}(t_i, w_i) + \frac{\partial f}{\partial w}(t_i, w_i) f(t_i, w_i) \right] + \frac{h^3}{6} y'''(c)$$

二阶龙格库塔,第一种解为显示梯形方法,第二种解为中点方法

$$w_{i+1} = w_i + af(t_i, w_i) + bf\left(t_i + \alpha h, w_i + \beta f(t_i, w_i)\right)$$

$$w_{i+1} = w_i + (a+b)f(t_i, w_i) + b \left[ \alpha h \frac{\partial f}{\partial t}(t_i, w_i) + \beta \frac{\partial f}{\partial w}(t_i, w_i)f(t_i, w_i) + O(h^2) \right]$$

$$\begin{cases} a+b=h \\ b\alpha h = b\beta = \frac{h^2}{2} \end{cases} \Rightarrow \begin{cases} a=b=\frac{h}{2} \\ \alpha = 1, \beta = h \end{cases} \text{ or } \begin{cases} a=0, b=h \\ \alpha = \frac{1}{2}, \beta = \frac{h}{2} \end{cases}$$

四阶龙格库塔 
$$w_{i+1} = w_i + \frac{h}{6}(s_1 + 2s_2 + 2s_3 + s_4)$$
  $s_1 = f(t_i, w_i)$ 

$$s_2 = f\left(t_i + \frac{h}{2}, w_i + \frac{h}{2}s_1\right) \quad s_3 = f\left(t_i + \frac{h}{2}, w_i + \frac{h}{2}s_2\right) \quad s_4 = f\left(t_i + h, w_i + hs_3\right)$$

对于方程组  $y_1'=f_1(t,y_1,y_2,\cdots),y_2'=f_2(t,y_1,y_2,\cdots),\cdots$  直接套用以上方法。 对于高阶 y'' = f(t, y, y') 变换为  $y_1 = y, y_2 = y$  则有  $y_1' = y_2, y_2' = f(t, y_1, y_2)$ 。

# 利普希茨连续、唯一性

 $\frac{b-a}{2m}$  当存在利普希茨常熟 L 对  $S=[a,b]\times[\alpha,\beta]$  中的每一对  $(t,y_1),(t,y_2)$  都满足

$$\frac{2m}{|f(t,y_1) - f(t,y_2)|} \le L|y_1 - y_2| \qquad L = \left|\frac{\partial f}{\partial y}(t,c)\right|_{max} \qquad \begin{cases} y' = f(t,y) \\ y(a) = y_0 \end{cases} \quad t \in [a,c]$$

则f(t, y) 相对于 y 在 S 上利普希茨连续,也可以使用求导的方法得到利普希茨常数。 满足以上条件,且  $\alpha < y_0 < \beta$ ,则存在  $c \in (a,b)$  使初值问题有唯一解。如果 f(t,y) 相对于 y 在  $S = [a,b] \times (-\infty,\infty)$  利普希茨连续,则在 [a,b] 上存在唯一解。 满足以上条件,如果 Y(t), Z(t) 是 y' = f(t,y) 在 S 上的解,分别有初始值

$$|Y(a), Z(a)| \leq e^{L(t-a)} |Y(a) - Z(a)|$$

### 7.3 边值问题,有限差分方法

利用差分公式替换微分方程中的导数,组装矩阵,估计初值,解方程组。

$$\begin{cases} y'' = 4y + t^2 \\ y(a) = y_a \\ y(b) = y_b \end{cases} \Rightarrow \begin{cases} y_a + (-4h^2 - 2)w_1 + w_2 + h^2t_1^2 = 0 \\ w_{i-1} + (-4h^2 - 2)w_i + w_{i+1} + h^2t_i^2 = 0 \\ w_{n-1} + (-4h^2 - 2)w_n + y_b + h^2t_n^2 = 0 \end{cases} \quad t_1 = a + h$$

$$\begin{bmatrix} -4h^2 - 2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -4h^2 - 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & \ddots & 0 & 0 & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \ddots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -4h^2 - 2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -4h^2 - 2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \vdots \\ w_{n-2} \\ w_{n-1} \\ w_n \end{bmatrix} = \begin{bmatrix} -h^2t_1^2 - y_a \\ -h^2t_2^2 \\ -h^2t_2^2 \\ -h^2t_{n-2}^2 \\ -h^2t_{n-1}^2 \\ -h^2t_{n-1}^2 - y_b \end{bmatrix}$$

$$y'' = 4y + t^2 \Rightarrow \frac{w_{i-1} - 2w_i + w_{i+1}}{h^2} = 4w_i + t_i^2 \qquad w_0 = y_a \qquad w_{n+1} = y_i$$