习题课

1　迭代法求线性方程组

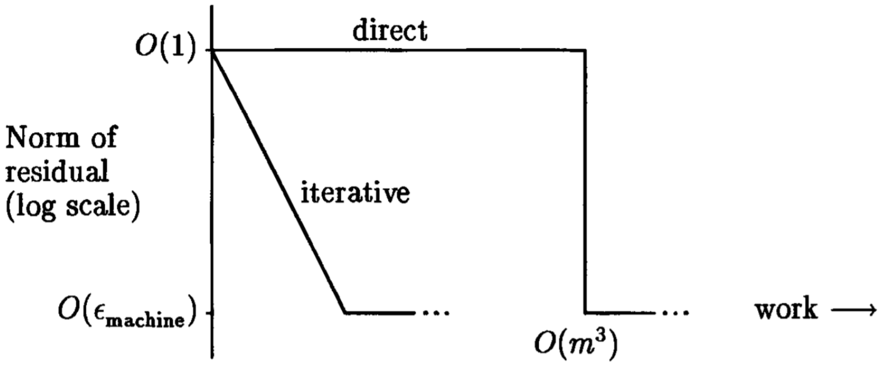

图 1: 求解线性方程组的直接法与迭代法的工作量

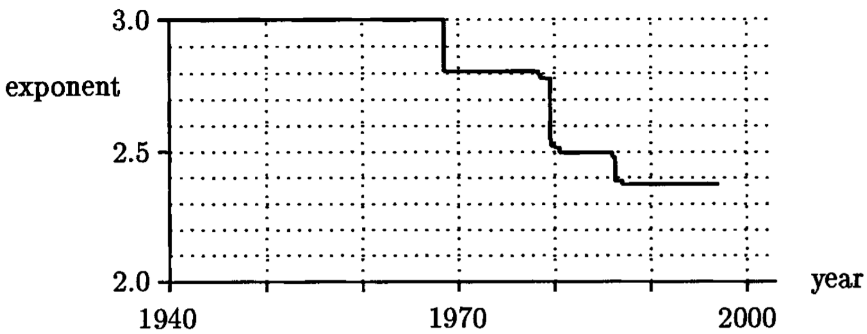

图 2: 直接法的复杂度中的指数

1.1　收敛的充要条件

基本思想：将方程 $Ax = b$ 重写为

$$Mx = Nx + b, \tag{1}$$

其中 $A = M - N$。迭代格式为

$$Mx^{k+1} = Nx^k + b, \tag{2}$$

或

$$x^{k+1} = M^{-1}Nx^k + M^{-1}b. \tag{3}$$

若方程组有解，设解为 x^*，则有

$$x^* = M^{-1}Nx^* + M^{-1}b. \tag{4}$$

将式(3)与(4)相减并记 $e^k = x^k - x^*$，我们有

$$e^{k+1} = Ge^k, \tag{5}$$

其中

$$G = M^{-1}N = I - M^{-1}A. \tag{6}$$

一般地，我们有

$$e^k = G^k e^0. \tag{7}$$

假定 G 可对角化，故有 n 个线性无关的特征向量 $v_1, v_2, \ldots, v_n$，并记 $\lambda_i, i = 1, 2, \ldots$ 为其相对应的特征值。我们有

$$e^0 = c_1 v_1 + c_2 v_2 + \cdots c_n v_n, \tag{8}$$

从而有

$$e^k = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots c_n \lambda_n^k v_n. \tag{9}$$

因此

$$\|e^k\| \leq |c_1||\lambda_1|^k \|v_1\| + |c_2||\lambda_2|^k \|v_2\| + \cdots |c_n||\lambda_n|^k \|v_n\|. \tag{10}$$

因为 $|\lambda_i|^k \to 0$ 当且仅当 $|\lambda_i| < 1$，所以

$$\|e^k\| \to 0 \quad \Leftrightarrow \quad \max_i |\lambda_i| < 1. \tag{11}$$

矩阵 G 的谱半径为

$$\rho(G) = \max_{\lambda \in \sigma(G)} |\lambda|. \tag{12}$$

因为我们有

$$\|G\|_p > \rho(G), \tag{13}$$

对任意矩阵范数 $\|\cdot\|_p$ 成立，故若 $\|G\|_p < 1$ 对某个范数成立，则矩阵谱半径小于 1，迭代格式收敛。注意这个条件只是充分条件，而非必要条件。学有余力：严格对角占优以及正定的情况（非充要条件，只是充分条件）。

1.2 常用矩阵范数

- $\|\cdot\|_1$, 竖着算
- $\|\cdot\|_2$, SVD
- $\|\cdot\|_\infty$. 横着算

$\|\cdot\|_2$ 的计算一般用矩阵 A^TA 的最大特征值的平方根得到。

1.3 迭代格式

- Jacobi: $M = D$, $G = I - D^{-1}A$,
- G-S: $M = D + L$, $G = I - (D + L)^{-1}A$,
- SOR: $M = \frac{1}{\omega}D + L$, $G = I - (\frac{1}{\omega}D + L)^{-1}A$.

要求会写分量形式！需要注意 SOR 迭代格式中取加权平均的步骤的位置。

for $i = 1, \ldots, n$ **do**

$$\begin{cases} \widetilde{x} \leftarrow \dfrac{1}{a_{ii}}\left(b_i - \displaystyle\sum_{j\neq i} a_{ij}x_j\right), \\ \delta \leftarrow \widetilde{x} - x_i, \\ x_i \leftarrow x_i + \omega\delta. \end{cases}$$

end for

以下为错误示范

for $i = 1, \ldots, n$ **do**

$$\begin{cases} \widetilde{x}_i \leftarrow \dfrac{1}{a_{ii}}\left(b_i - \displaystyle\sum_{j\neq i} a_{ij}x_j\right), \\ \delta_i \leftarrow \widetilde{x}_i - x_i. \end{cases}$$

end for

$x \leftarrow x + \omega\delta.$

1.4 牛顿法

牛顿迭代法求解非线性方程：

- 收敛阶数及其证明，
- 重根的情况：一阶方法。如何修正得到二阶方法？

- 收敛速度高于二阶；收敛速度为三阶的条件：$f(\alpha)=0, f'(\alpha)\neq 0, f''(\alpha)=0,$
- 求解非线性方程组的牛顿法：Jacobian 的构造。

1.5 典型例题

例 1. 讨论用 *Jacobi* 迭代和 *Gauss-Seidel* 迭代解 $Ax=b$ 时的收敛性，如果收敛，比较那种方法收敛较快，其中

$$A=\begin{pmatrix} 3 & 0 & 2 \\ 0 & 2 & 1 \\ -2 & 1 & 2 \end{pmatrix}. \tag{14}$$

解答：若用 Jacobi 迭代格式，迭代矩阵 $G_J=I-D^{-1}A$，谱半径为 $\rho(G_J)\approx 0.6455$。若用 G-S 迭代格式，迭代矩阵 $G_G=I-(D+L)^{-1}A$，谱半径为 $\rho(G_G)\approx 0.4167$。

例 2. 给定线性方程组 $Ax=b$，其中

$$A=\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}, \quad b=\begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

若使用迭代公式

$$x^{k+1}=x^k+\alpha(b-Ax^k), \alpha\in\mathbb{R},$$

求解方程，则迭代矩阵为？α 取何值时收敛？何值收敛最快？

解答：迭代矩阵为

$$G=I-\alpha A=\begin{pmatrix} 1-3\alpha & -2\alpha \\ -\alpha & 1-2\alpha \end{pmatrix},$$

计算其特征值得到 $\lambda_1=1-4\alpha, \lambda_2=1-\alpha$。若迭代格式收敛，则有

$$|1-4\alpha|<1, |1-\alpha|<1,$$

故有

$$0\leq\alpha\leq\frac{1}{2}.$$

迭代矩阵谱半径为 $\rho(G)=\max(|\lambda_1|,|\lambda_2|)$，若迭代收敛速度最快，则有

$$1-4\alpha=-(1-\alpha),$$

即 $\alpha=\frac{2}{5}$ 时收敛速度最快。

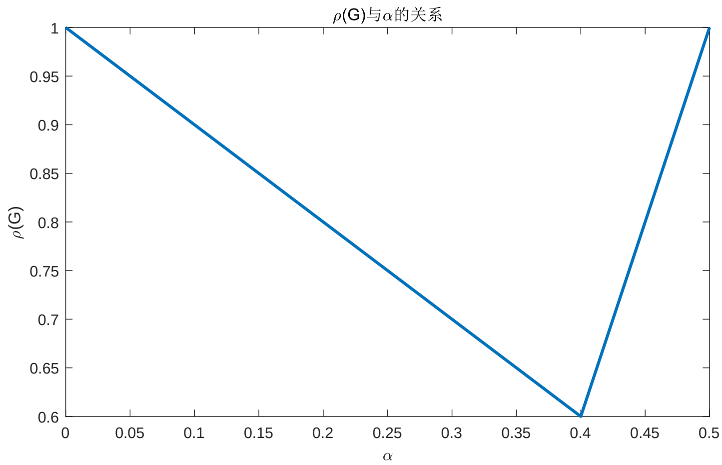

图 3: $\rho(G)$ 与 α 的关系

例 3. 判断迭代序列：$x_{k+1} = 1 + \frac{1}{x_k^2}$ 对初值 $x_0 = 1.5$ 的收敛性，并简述理由。

解答：我们由 $x_k \geq 1.3$ 可以得到 $x_{k+1} = 1 + \frac{1}{x_k^2} \leq 1 + \frac{1}{1.3^2} \approx 1.59 < 1.6$。有了上界估计，我们可以得到下界估计：由 $x_k \leq 1.6$ 得到 $x_{k+1} = 1 + \frac{1}{x_k^2} \geq 1 + \frac{1}{1.6^2} \approx 1.39 > 1.3$。从而迭代产生的序列所在区间为 $[1.3, 1.6]$。而又有

$$\left|\left(1 + \frac{1}{x^2}\right)'\right| = \left|\frac{-2}{x^3}\right| < 1, \quad \forall x \in [1.3, 1.6].$$

再用泰勒展开（或书上定理）。

例 4. 已知参数 $C > 0$，设有线性代数方程组 $Ax = b$，其中

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & C \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

1 分别给出 *Jacobi* 迭代和 *Gauss-Seidel* 迭代的分量形式；

2 求 *Jacobi* 迭代的迭代矩阵；

3 试确定参数 $C(C > 0)$ 的取值范围，使 *Jacobi* 迭代收敛。

解答：

1 注意 b 要变换为 $M^{-1}b$。

2 迭代矩阵为 $I - D^{-1}A$，计算得到

$$\begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{C} & 0 \end{pmatrix} \tag{15}$$

3 求解(15)的特征值得到 $\lambda^2 = \frac{1+C\pm\sqrt{1+C^2}}{4C}$。故我们要求

$$\frac{1+C+\sqrt{1+C^2}}{4C} < 1 \text{ 且} \frac{1+C-\sqrt{1+C^2}}{4C} < 1.$$

注意 $C > 0$，解得 $C > \frac{3}{4}$。

2 迭代法求解矩阵的特征值

为什么用迭代法？Galois 理论：一般高于四次的代数方程不能用根式求解。

2.1 幂法

Algorithm 1 幂法求矩阵的模意义下最大的特征值

输入: [矩阵 A，迭代次数 m，误差限 ε]

输出: [A 的模最大的特征值]

1 初始化。置 $a = b = (1, 1, \dots, 1)^T, \bar{a} = a/\|a\|_\infty, \bar{b} = b/\|b\|_\infty, c = Aa, \bar{c} = c/\|c\|_\infty$。

2 迭代计算 $x^{k+1} = Ax^k$。

for $k = 1, 2, \dots, m$ **do**

3 将 A 作用到向量上。$d = A\bar{c}, \bar{d} = d/\|d\|_\infty$

4 判断模最大特征值是否为单特征值。

if $\|\bar{d} - \bar{c}\| < \varepsilon$ 或 $\|\bar{d} + \bar{c}\| < \varepsilon$ **then**

5 模最大特征值为单特征值，计算 $\lambda = \frac{d_j}{c_j}$ 对某个 $j \in \{1, 2, \dots, n\}$。

return 特征值 λ，特征向量 $\bar{d}$。

end if

6 判断模最大特征值是否为互为反号的一对特征值。

if $\|\bar{d} - \bar{b}\| < \varepsilon$ 且 $\|\bar{c} - \bar{a}\| < \varepsilon$ **then**

7 模最大特征值为互为反号的一对特征值，计算 $\lambda_1 = \sqrt{\|d\|_\infty \|c\|_\infty}$ 对某个 $j \in \{1, 2, \dots, n\}$（或 $\lambda_1 = \sqrt{\frac{d_j}{b_j}}$），计算特征向量 $v_1 = \lambda_1 \bar{c} + d, v_1 = \lambda_1 \bar{c} - d$，归一化特征向量 $v_1 = v_1/\|v_1\|_\infty, v_2 = v_2/\|v_2\|_\infty$。

return 特征值 $\lambda_1, -\lambda_1$，特征向量 v_1, v_2，其中 v_1 对应于正特征值 λ_1，v_2 对应于负特征值 λ_2。

end if

8 更新存储变量的值。$\bar{a} = \bar{b}, \bar{b} = \bar{c}, c = d, \bar{c} = \bar{d}$（若用 $\lambda_1 = \sqrt{\frac{d_j}{b_j}}$ 计算，需更新 $b = c$）。

end for

9 **return** 幂法不收敛。

2.2 反幂法

设矩阵 A 可逆，我们求解 A^{-1} 的模最大特征值来得到 A 的模最小特征值，或求解 $(A - \rho I)^{-1}$ 的最大特征值，来求解 A 最靠近 ρ 的特征值，$\rho \in \mathbb{C}$（$\rho = 0$ 即为求解 A 模最小的特征值）。

Algorithm 2 反幂法求解矩阵特征值

输入: [矩阵 A，迭代次数 m，误差限 ε，$\rho \in \mathbb{C}$]

输出: [A 的最接近 ρ 的特征值]

1 初始化。置 $a = (1, 1, \ldots, 1)^T$。求出矩阵 $A - \rho I$ 的 LU 分解（为了方便求解 $(A - \rho I)x = b$，即 $x = (A - \rho I)^{-1}b$）。

2 迭代计算 $(A - \rho I)x^{k+1} = x^k$。

for $k = 1, 2, \ldots, m$ **do**

3 将 $(A - \rho I)^{-1}$ 作用到向量上。$q = L\backslash a, q = U\backslash q$。

4 求出 q 按分量最大的元素序号并归一化。$j = \arg\max_{i=1,2,\ldots,n} |q_i|$ 存储 $s = q_j$，计算归一化向量 $q = q/q_j$。

if $\|q - a\| < \varepsilon$ **then**

return 特征值 $\rho + \frac{1}{s}$，特征向量 q。

end if

5 更新存储变量的值。$a = q$。

end for

9

return 反幂法不收敛。

2.3 实对称矩阵的 Jacobi 方法

基本思想：消除非对角线元素。基本方法：Givens 旋转变换。

$$Q(p, q, \theta) = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & \cos\theta & & & \sin\theta & & \\ & & & 1 & & & & \\ & & & & \ddots & & & \\ & & & & & 1 & & \\ & & -\sin\theta & & & \cos\theta & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \tag{16}$$

记 $B=Q^TAQ$，则有下列公式

$$\begin{cases} b_{ip}=b_{pi}=a_{pi}\cos\theta-a_{pi}\sin\theta, & i\neq p,q, \\ b_{iq}=b_{qi}=a_{pi}\sin\theta+a_{qi}\cos\theta, & i\neq p,q, \\ b_{pp}=a_{pp}\cos^2\theta+a_{qq}\sin^2\theta-a_{pq}\sin 2\theta, \\ b_{qq}=a_{pp}\sin^2\theta+a_{qq}\cos^2\theta+a_{pq}\sin 2\theta, \\ b_{pq}=b_{qp}=a_{pq}\cos 2\theta+\dfrac{a_{pp}-a_{qq}}{2}\sin 2\theta. \end{cases} \tag{17}$$

目标是使得 $b_{pq}=b_{qp}=0$，即求解 $a_{pq}\cos 2\theta+\frac{a_{pp}-a_{qq}}{2}\sin 2\theta=0$。记 $s=-\frac{a_{pp}-a_{qq}}{2a_{pq}},t=\tan\theta$，则

$$t=\begin{cases} t^2+2st-1=0\text{的绝对值较小根，} & s\neq 0, \\ 1, & s=0. \end{cases} \tag{18}$$

例 5. 当使用 *Givens* 旋转变换将 $A=\begin{pmatrix} 4 & -2 & 4 \\ -2 & 10 & -2 \\ 4 & -2 & 8 \end{pmatrix}$ 的非对角线按模最大元素消为零时，所使用的的 *Givens* 旋转矩阵是？变换后得到的矩阵是？

解答：模最大的非对角线元素为 $A(1,3)=A(3,1)=4,p=1,q=3$，故 $s=-\frac{a_{pp}-a_{qq}}{2a_{pq}}=-\frac{4-8}{2\times 4}=0.5$。求解方程

$$t^2+2st-1=0$$

得到

$$t=\frac{-1\pm\sqrt{5}}{2}.$$

故我们取 $t=\frac{-1+\sqrt{5}}{2}$。从而

$$\cos\theta=\frac{1}{\sqrt{1+t^2}}=\sqrt{\frac{5+\sqrt{5}}{10}},\quad \sin\theta=\frac{t}{\sqrt{1+t^2}}=\frac{-1+\sqrt{5}}{\sqrt{10-2\sqrt{5}}}.$$

所需的 Givens 旋转矩阵是

$$\begin{pmatrix} \sqrt{\frac{5+\sqrt{5}}{10}} & 0 & \frac{-1+\sqrt{5}}{\sqrt{10-2\sqrt{5}}} \\ 0 & 1 & 0 \\ -\frac{-1+\sqrt{5}}{\sqrt{10-2\sqrt{5}}} & 0 & \sqrt{\frac{5+\sqrt{5}}{10}} \end{pmatrix}$$

变换后得到的矩阵是

$$\begin{pmatrix} 6-2\sqrt{5} & (-3+\sqrt{5})\sqrt{\frac{5+\sqrt{5}}{10}} & 0 \\ (-3+\sqrt{5})\sqrt{\frac{5+\sqrt{5}}{10}} & 10 & -(1+\sqrt{5})\sqrt{\frac{5+\sqrt{5}}{10}} \\ 0 & -(1+\sqrt{5})\sqrt{\frac{5+\sqrt{5}}{10}} & 6+2\sqrt{5} \end{pmatrix}$$

3 差商

- 重节点差商。课本 36 页例 10.
- 差商的含义：在点 $\{x_1, x_2, \ldots, x_{n+1}\}$ 处插值于函数 f 的 n 次多项式的最高次项的系数（归纳法证明）。若 $f = 2x^7 + x^3 + 1$，则 $f[2^0, 2^1, \ldots, 2^7], f[2^0, 2^1, \ldots, 2^8]$ 为？