第 3 章 线性方程组的数值解法

本章主要内容：

(1) Gauss 消去法, Gauss 列主元消去法

(2) 方程组的性态和误差估计

(3) 迭代法 (Jacobi、Gauss-Seidel、SOR 迭代)

(4) 幂法和反幂法

1 Gauss 消去法

思想：利用线性代数中的初等行变换将方程组化为等价的三角形方程组.

1.1 三角形方程组的回代法

考虑三角形方程组

$$
\begin{array}{rrcrrl}
u_{11}x_1+ & u_{12}x_2+ & \cdots + & u_{1,n-1}x_{n-1}+ & u_{1n}x_n & = y_1 \\
 & u_{22}x_2+ & \cdots + & u_{2,n-1}x_{n-1}+ & u_{2n}x_n & = y_2 \\
 & & \ddots & \cdots & \cdots\cdots & \\
 & & & u_{n-1,n-1}x_{n-1}+ & u_{n-1,n}x_n & = y_{n-1} \\
 & & & & u_{nn}x_n & = y_n
\end{array}
$$

其中 $u_{ii} \neq 0(i = 1, 2, \cdots, n)$. 用下面的回代法求解:

$$
\begin{aligned}
x_n &= y_n/u_{nn}, \\
x_i &= \left(y_i - \sum_{j=i+1}^{n} u_{ij}x_j \right) /u_{ii}, \quad i = n-1, n-2, \cdots, 1.
\end{aligned}
$$

1.2 Gauss 消去法

考虑一般的线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{1.1}$$

其中

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

将方程 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 用增广矩阵表示, 记

$$\bar{\boldsymbol{A}}^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & a_{2,n+1}^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & a_{n,n+1}^{(1)} \end{bmatrix},$$

其中

$$a_{ij}^{(1)} = a_{ij}, \quad i, j = 1, 2 \cdots n,$$

$$a_{i,n+1}^{(1)} = b_i, \quad i = 1, 2 \cdots, n.$$

下面用 $n-1$ 步消元（初等行变换）将矩阵 $\bar{\boldsymbol{A}}^{(1)}$ 化为上三角矩阵.

(1) 第 1 步消元： 假设 $a_{11}^{(1)} \neq 0$ (否则交换两行的位置), 记 $l_{i1} = \dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}}$,

第 1 行乘 $-l_{i1}$ 加到第 i $(2 \le i \le n)$ 行得

$$\bar{\boldsymbol{A}}^{(1)} \longrightarrow \bar{\boldsymbol{A}}^{(2)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{bmatrix}$$

其中

$$a_{ij}^{(2)} = a_{ij}^{(1)} - l_{i1} a_{1j}^{(1)}, \quad i = 2, 3, \cdots, n, \quad j = 2, 3, \cdots, n+1.$$

(2) 第 2 步消元： 假设 $a_{22}^{(2)} \neq 0$ (否则交换两行的位置), 记 $l_{i2} = \dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}}$,

第 2 行乘 $-l_{i2}$ 加到第 i $(3 \le i \le n)$ 行得

$$\bar{\boldsymbol{A}}^{(2)} \longrightarrow \bar{\boldsymbol{A}}^{(3)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & a_{3n}^{(3)} & \cdots & a_{nn}^{(3)} & a_{n,n+1}^{(3)} \end{bmatrix}$$

其中

$$a_{ij}^{(3)} = a_{ij}^{(2)} - l_{i2}a_{2j}^{(2)}, \quad i = 3, 4, \cdots, n, \quad j = 3, 4, \cdots, n+1.$$

(3) 假设按上面进行了 $k-1$ 步消元, 即有

$$\bar{\boldsymbol{A}}^{(1)} \longrightarrow \bar{\boldsymbol{A}}^{(2)} \longrightarrow \bar{\boldsymbol{A}}^{(3)} \longrightarrow \cdots \longrightarrow \bar{\boldsymbol{A}}^{(k)},$$

其中

$$\bar{\boldsymbol{A}}^{(k)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1,k-1}^{(1)} & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2,k-1}^{(2)} & a_{2,k}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3,k-1}^{(3)} & a_{3k}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & \cdots & a_{k-1,n}^{(k-1)} & a_{k-1,n+1}^{(k-1)} \\ 0 & 0 & 0 & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & a_{k,n+1}^{(k)} \\ 0 & 0 & 0 & \cdots & 0 & a_{k+1,k}^{(k)} & \cdots & a_{k+1,n}^{(k)} & a_{k+1,n+1}^{(k)} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & a_{n,n+1}^{(k)} \end{bmatrix}$$

(4) 第 k 步消元： 假设 $a_{kk}^{(k)} \neq 0$ (否则交换两行的位置), 记 $l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$, 第 k 行乘 $-l_{ik}$ 加到第 i $(k+1 \leq i \leq n)$ 行得

$$
\begin{aligned}
&\bar{\boldsymbol{A}}^{(k)} \longrightarrow \bar{\boldsymbol{A}}^{(k+1)} \\
&= \begin{bmatrix}
a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1,k}^{(1)} & a_{1,k+1}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\
0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2,k}^{(2)} & a_{2,k+1}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\
0 & 0 & a_{33}^{(3)} & \cdots & a_{3,k}^{(3)} & a_{3,k+1}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \cdots & a_{k,n}^{(k)} & a_{k,n+1}^{(k)} \\
0 & 0 & 0 & \cdots & 0 & a_{k+1,k+1}^{(k+1)} & \cdots & a_{k+1,n}^{(k+1)} & a_{k+1,n+1}^{(k+1)} \\
0 & 0 & 0 & \cdots & 0 & a_{k+2,k+1}^{(k+1)} & \cdots & a_{k+2,n}^{(k+1)} & a_{k+2,n+1}^{(k+1)} \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & a_{n,k+1}^{(k+1)} & \cdots & a_{nn}^{(k+1)} & a_{n,n+1}^{(k+1)}
\end{bmatrix}
\end{aligned}
$$

其中

$$\begin{aligned}&a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik}a_{kj}^{(k)},\\ &i = k+1, k+2, \cdots, n, \quad j = k+1, k+2, \cdots, n+1.\end{aligned} \tag{1.2}$$

(5) 总共进行 $n-1$ 步 $(k = 1, 2, \cdots, n-1)$ 消元后,

$$\bar{\boldsymbol{A}}^{(1)} \longrightarrow \bar{\boldsymbol{A}}^{(2)} \longrightarrow \bar{\boldsymbol{A}}^{(3)} \longrightarrow \cdots \longrightarrow \bar{\boldsymbol{A}}^{(k)} \longrightarrow \cdots \longrightarrow \bar{\boldsymbol{A}}^{(n)},$$

其中

$$\bar{\boldsymbol{A}}^{(n)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & & \ddots & \vdots & \vdots & \vdots \\ & & & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} & a_{n-1,n+1}^{(n-1)} \\ & & & & a_{nn}^{(n)} & a_{n,n+1}^{(n)} \end{bmatrix}. \tag{1.3}$$

若记

$$
\boldsymbol{U}=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} \\ & & \ddots & \vdots & \vdots \\ & & & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} \\ & & & & a_{nn}^{(n)} \end{bmatrix}, \quad \boldsymbol{y}=\begin{bmatrix} a_{1,n+1}^{(1)} \\ a_{2,n+1}^{(2)} \\ \vdots \\ a_{n-1,n+1}^{(n-1)} \\ a_{n,n+1}^{(n)} \end{bmatrix},
$$

则 (1.3) 对应的线性方程组为 $\boldsymbol{Ux}=\boldsymbol{y}$, 用回代可求得 $\boldsymbol{x}$.

可以估计, Gauss 消元法解线性方程组 (1.1) 需要加减法和乘除法次数均为 $O(n^3)$.

定理 1.1 给定线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$. 如果 $\boldsymbol{A}$ 的各阶顺序主子式非零, 则 Gauss 消去法中的各主元 $a_{kk}^{(k)}(k=1,2,\cdots,n)$ 均非零.

1.3　三对角方程组的追赶法

考虑三对角方程组

$$\begin{bmatrix} b_1 & c_1 & & & & \\ a_2 & b_2 & c_2 & & & \\ & a_3 & b_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix} \tag{1.4}$$

其系数矩阵元素满足

(1) $|b_1| > |c_1| > 0$;

(2) $|b_i| \geq |a_i| + |c_i|, \quad a_i c_i \neq 0 (i = 2, 3, \cdots, n-1)$;

(3) $|b_n| > |a_n| > 0$.

易证 (1.4) 的系数矩阵非奇异. 利用 Gauss 消去法解方程组 (1.4), 每步消元只要消一个元素.

消元过程算法如下:

$$\begin{aligned}
&\beta_1 = b_1, \quad y_1 = d_1, \\
&\text{对} i = 2, 3, \cdots, n \text{ 计算} \\
&l_i = \frac{a_i}{\beta_{i-1}}, \quad \beta_i = b_i - l_i c_{i-1}, \quad y_i = d_i - l_i y_{i-1}.
\end{aligned}$$

回代算法:

$$\begin{aligned}
&x_n = y_n/\beta_n, \\
&\text{对} i = n-1, n-2, \cdots, 1 \text{ 计算} \\
&x_i = (y_i - c_i x_{i+1})/\beta_i.
\end{aligned}$$

1.4 列主元 Gauss 消去法

由 (1.2) 看出, 在 Gauss 消去法中, 当 $|l_{ik}|$ 很大 (如 $|a_{kk}^{(k)}|$ 很小) 时, 元素 $a_{kj}^{(k)}$ 很小的误差, 将导致元素 $a_{ij}^{(k+1)}$ 较大的误差. 所以希望 $|l_{ik}| \le 1$.

设进行了 $k-1$ 步消元.

$$
\bar{\boldsymbol{A}}^{(k)} = \begin{bmatrix}
a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1,k-1}^{(1)} & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\
0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2,k-1}^{(2)} & a_{2,k}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\
0 & 0 & a_{33}^{(3)} & \cdots & a_{3,k-1}^{(3)} & a_{3k}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & a_{k-1,k-1}^{(k-1)} & a_{k-1,k}^{(k-1)} & \cdots & a_{k-1,n}^{(k-1)} & a_{k-1,n+1}^{(k-1)} \\
0 & 0 & 0 & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} & a_{k,n+1}^{(k)} \\
0 & 0 & 0 & \cdots & 0 & a_{k+1,k}^{(k)} & \cdots & a_{k+1,n}^{(k)} & a_{k+1,n+1}^{(k)} \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} & a_{n,n+1}^{(k)}
\end{bmatrix}
$$

在第 k 步消元之前, 从第 k 列位于对角线以下的元素中选绝对值最大

者作为主元. 如果

$$|a_{sk}^{(k)}| = \max_{k\le i\le n} |a_{ik}^{(k)}|,$$

即 (s,k) 元素绝对值最大, 则交换第 s 行和第 k 行对应元素, 然后进行消元. 显然此时有

$$|l_{ik}| = \left|\frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}\right| \le 1.$$

例 1.1 用列主元 Gauss 消去法求下列方程组的解.

$$\begin{bmatrix} 3 & 1 & -1 \\ 4 & 0 & 4 \\ 12 & -3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \\ 9 \end{bmatrix}.$$

2 方程组的性态与误差分析

2.1 向量范数

设 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T \in \mathbf{R}^n$.

定义 2.1 设 $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ 是 $\mathbf{R}^n$ 上的函数, 如果满足:

(1) (非负性) $\forall \boldsymbol{x} \in \mathbf{R}^n$, 有 $\|\boldsymbol{x}\| \geq 0$, 且 $\|\boldsymbol{x}\| = 0 \Leftrightarrow \boldsymbol{x} = \mathbf{0}$;

(2) (齐次性) $\forall \boldsymbol{x} \in \mathbf{R}^n$, $\lambda \in \mathbf{R}$, 有 $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$;

(3) (三角不等式) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$, 有 $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.

则称 $\|\cdot\|$ 为 $\mathbf{R}^n$ 上的**向量范数**.

注 2.1 用 $\mathbf{C}^n$ 表示所有复的 n 维向量组成的复线性空间, 类似可以定义对应的向量范数.

常用的三个向量范数:

(1) $1-$ 范数: $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{n} |x_i|$;

(2) $\infty-$ 范数: $\|\boldsymbol{x}\|_\infty = \max_{1 \leq i \leq n} \{|x_i|\}$;

(3) 2− 范数: $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\boldsymbol{x}^T\boldsymbol{x}} = \sqrt{(\boldsymbol{x},\boldsymbol{x})}$.

由范数定义的三角不等式易得

$$\big|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|\big| \le \|\boldsymbol{x} - \boldsymbol{y}\|.$$

定理 2.1 (向量范数的连续性) 设 $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ 为 $\mathbf{R}^n$ 上的任意一个向量范数, 则 $f(\boldsymbol{x})$ 为 $\boldsymbol{x}$ 分量的连续函数. 即

$$\lim_{\boldsymbol{x}\to\boldsymbol{y}} f(\boldsymbol{x}) = f(\boldsymbol{y}).$$

定义 2.2 设 $\|\cdot\|_p$ 和 $\|\cdot\|_q$ 是 $\mathbf{R}^n$ 上两个范数. 如果存在正常数 c_1, c_2 使得对 $\forall \boldsymbol{x} \in \mathbf{R}^n$, 有

$$c_1\|\boldsymbol{x}\|_p \le \|\boldsymbol{x}\|_q \le c_2\|\boldsymbol{x}\|_p,$$

则称范数 $\|\cdot\|_p$ 和 $\|\cdot\|_q$ 等价.

定理 2.2 (向量范数的等价性) $\mathbf{R}^n$ 上任意两个范数都等价.

定义 2.3 设 $\|\cdot\|$ 是 $\mathbf{R}^n$ 上的范数, $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n$, 称 $\|\boldsymbol{x} - \boldsymbol{y}\|$ 为 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 之间的距离.

利用距离可以研究 2 个向量之间的绝对误差和相对误差. 设 $\boldsymbol{x}^*, \tilde{\boldsymbol{x}}$ 是方程组 $A\boldsymbol{x}=b$ 的精确解和近似解. 则可用 $\|\boldsymbol{x}^*-\tilde{\boldsymbol{x}}\|$ 和 $\dfrac{\|\boldsymbol{x}^*-\tilde{\boldsymbol{x}}\|}{\|\boldsymbol{x}^*\|}$ 或 $\dfrac{\|\boldsymbol{x}^*-\tilde{\boldsymbol{x}}\|}{\|\tilde{\boldsymbol{x}}\|}$ 表示近似解 $\tilde{\boldsymbol{x}}$ 的**绝对误差**和**相对误差**.

定义 2.4 设 $\|\cdot\|$ 是 $\mathbf{R}^n$ 中的一个范数, $\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots$ 是 $\mathbf{R}^n$ 中的一个向量序列, $\boldsymbol{c}\in\mathbf{R}^n$ 为常向量. 如果

$$\lim_{k\to\infty}\|\boldsymbol{x}^{(k)}-\boldsymbol{c}\|=0,$$

则称向量序列 $\{\boldsymbol{x}^{(k)}\}_{k=0}^{\infty}$ 收敛于 $\boldsymbol{c}$, 记为

$$\lim_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{c}.$$

2.2 矩阵范数

设

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

$\|\cdot\|$ 为 $\mathbf{R}^n$ 上的一个范数.

定义 2.5 称

$$\max_{\substack{x\in\mathbf{R}^n \\ x\neq\mathbf{0}}} \frac{\|Ax\|}{\|x\|}$$

为矩阵 A 的范数, 记为 $\|A\|$. 即

$$\|A\| = \max_{\substack{x\in\mathbf{R}^n \\ x\neq\mathbf{0}}} \frac{\|Ax\|}{\|x\|}.$$

问题: 最大值能否取到?

矩阵范数的性质: 设 $\|\cdot\|$ 是一个矩阵范数, 则有

(1) $\forall A \in \mathbf{R}^{n\times n}$, $\|A\| \geq 0$, 且 $\|A\| = 0 \Leftrightarrow A = \mathbf{0}$.

(2) $\forall \boldsymbol{A} \in \mathbf{R}^{n\times n}, \lambda \in \mathbf{R}$, 有 $\|\lambda \boldsymbol{A}\| = |\lambda|\|\boldsymbol{A}\|$.

(3) $\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbf{R}^{n\times n}$, 有 $\|\boldsymbol{A}+\boldsymbol{B}\| \le \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$.

(4) $\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbf{R}^{n\times n}$, 有 $\|\boldsymbol{A}\boldsymbol{B}\| \le \|\boldsymbol{A}\|\|\boldsymbol{B}\|$.

(5) $\forall \boldsymbol{A} \in \mathbf{R}^{n\times n}, \boldsymbol{x} \in \mathbf{R}^n$, 有 $\|\boldsymbol{A}\boldsymbol{x}\| \le \|\boldsymbol{A}\|\|\boldsymbol{x}\|$.

定义 2.6 设矩阵 $\boldsymbol{B} \in \mathbf{R}^{n\times n}$, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 $\boldsymbol{B}$ 的 n 个特征值. 称

$$\rho(\boldsymbol{B}) = \max_{1\le i\le n}\{|\lambda_i|\}$$

为矩阵 $\boldsymbol{B}$ 的谱半径.

定理 2.3 设 $\boldsymbol{A} \in \mathbf{R}^{n\times n}$, 则

(1)

$$\|\boldsymbol{A}\|_1 = \max_{\substack{\boldsymbol{x}\in\mathbf{R}^n\\ \boldsymbol{x}\ne\mathbf{0}}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_1} = \max_{1\le j\le n}\left\{\sum_{i=1}^n |a_{ij}|\right\}.$$

(2)

$$\|\boldsymbol{A}\|_\infty = \max_{\substack{\boldsymbol{x}\in\mathbf{R}^n\\ \boldsymbol{x}\ne\mathbf{0}}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} = \max_{1\le i\le n}\left\{\sum_{j=1}^n |a_{ij}|\right\}.$$

(3)

$$\|\boldsymbol{A}\|_2=\max_{\substack{\boldsymbol{x}\in\mathbf{R}^n\\ \boldsymbol{x}\neq\mathbf{0}}}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}=\sqrt{\rho(\boldsymbol{A}^T\boldsymbol{A})}.$$

定理 2.4 设 $\|\cdot\|$ 是 $\mathbf{R}^{n\times n}$ 上的任意一个矩阵范数, $\boldsymbol{A}\in\mathbf{R}^{n\times n}$, 则有

$$\rho(\boldsymbol{A})\leq\|\boldsymbol{A}\|.$$

定理 2.5 如果 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 为对称矩阵, 则 $\rho(\boldsymbol{A})=\|\boldsymbol{A}\|_2$.

定义 2.7 设 $\|\cdot\|_p$ 和 $\|\cdot\|_q$ 为 $\mathbf{R}^{n\times n}$ 上的两个范数, 如果存在正常数 c_1,c_2, 使得对 $\forall\boldsymbol{A}\in\mathbf{R}^{n\times n}$ 有

$$c_1\|\boldsymbol{A}\|_p\leq\|\boldsymbol{A}\|_q\leq c_2\|\boldsymbol{A}\|_p.$$

则称 $\|\cdot\|_p$ 和 $\|\cdot\|_q$ 等价.

定理 2.6 $\mathbf{R}^{n\times n}$ 上任意两个矩阵范数都等价.

定义 2.8 设 $\|\cdot\|$ 是 $\mathbf{R}^{n\times n}$ 上的矩阵范数, $\boldsymbol{A},\boldsymbol{B}\in\mathbf{R}^{n\times n}$, 称 $\|\boldsymbol{A}-\boldsymbol{B}\|$ 为 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之间的距离.

定义 2.9 设 $\|\cdot\|$ 是 $\mathbf{R}^{n\times n}$ 上的矩阵范数, $\boldsymbol{A}^{(0)},\boldsymbol{A}^{(1)},\cdots,\boldsymbol{A}^{(k)},\cdots$ 为 $\mathbf{R}^{n\times n}$ 上的一个矩阵序列, $\boldsymbol{A}\in\mathbf{R}^{n\times n}$. 如果

$$\lim_{k\to\infty}\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\|=0,$$

则称矩阵序列 $\{\boldsymbol{A}^{(k)}\}_{k=0}^{\infty}$ 收敛于矩阵 $\boldsymbol{A}$, 记为

$$\lim_{k\to\infty} \boldsymbol{A}^{(k)} = \boldsymbol{A}.$$

定理 2.7 设 $\boldsymbol{B} \in \mathrm{R}^{n\times n}$, 则矩阵序列 $\boldsymbol{I}, \boldsymbol{B}, \boldsymbol{B}^2, \cdots, \boldsymbol{B}^k, \cdots$ 收敛于零矩阵 (即 $\lim\limits_{k\to\infty} \boldsymbol{B}^k = \boldsymbol{O}$) 的充要条件是 $\rho(\boldsymbol{B}) < 1$.

2.3 方程组的性态和条件数

例 2.1 考虑方程组

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

其解为 $x_1 = x_2 = 1$.

设方程组系数有小扰动, 方程组成为

$$\begin{bmatrix} 1 & -1 \\ 1 & 1.0005 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

其解为 $x_1 = x_2 = 0.99975006$.

该例子说明矩阵误差对解的影响不大.

例 2.2 考虑方程组

$$\begin{bmatrix} 10 & -10 \\ -1 & 1.001 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.001 \end{bmatrix},$$

其解为 $x_1 = x_2 = 1$.

同样, 若系数有误差, 方程组变为

$$\begin{bmatrix} 10 & -10 \\ -1 & 1.0015 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.001 \end{bmatrix},$$

其解为 $x_1 = x_2 = \dfrac{2}{3}$.

这个例子说明方程组的系数矩阵的误差对解的影响很大. 问题: 怎么判别?

设 $\boldsymbol{A}$ 非奇异, $\boldsymbol{b} \neq \mathbf{0}$, 方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 的精确解为 $\boldsymbol{x}^*$.

(1) 设 $\boldsymbol{b}$ 有很小的扰动 $\delta\boldsymbol{b}$, 此时解 $\boldsymbol{x}^*$ 变为 $\boldsymbol{x}^* + \delta\boldsymbol{x}^*$, 即有

$$\boldsymbol{A}(\boldsymbol{x}^* + \delta\boldsymbol{x}^*) = \boldsymbol{b} + \delta\boldsymbol{b}.$$

注意到 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$, 则得

$$\boldsymbol{A}\delta\boldsymbol{x}^* = \delta\boldsymbol{b},$$

即

$$\delta\boldsymbol{x}^* = \boldsymbol{A}^{-1}\delta\boldsymbol{b}.$$

两边取范数并由 矩阵范数性质 得

$$\|\delta\boldsymbol{x}^*\| \leq \|\boldsymbol{A}^{-1}\|\|\delta\boldsymbol{b}\|.$$

另一方面, 由方程 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$ 可得

$$\|\boldsymbol{b}\| \le \|\boldsymbol{A}\|\|\boldsymbol{x}^*\|.$$

从而有

$$\frac{\|\delta\boldsymbol{x}^*\|}{\|\boldsymbol{x}^*\|} \le \|\boldsymbol{A}^{-1}\|\|\boldsymbol{A}\|\frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}. \tag{2.1}$$

(2) 设 $\boldsymbol{A}$ 有微小变化 $\delta\boldsymbol{A}$, 解变为 $\boldsymbol{x}^* + \delta\boldsymbol{x}^*$. 即

$$(\boldsymbol{A} + \delta\boldsymbol{A})(\boldsymbol{x}^* + \delta\boldsymbol{x}^*) = \boldsymbol{b}.$$

利用方程 $\boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{b}$ 得

$$\delta\boldsymbol{A}(\boldsymbol{x}^* + \delta\boldsymbol{x}^*) + \boldsymbol{A}\delta\boldsymbol{x}^* = 0,$$

即

$$\delta\boldsymbol{x}^* = -\boldsymbol{A}^{-1}\delta\boldsymbol{A}(\boldsymbol{x}^* + \delta\boldsymbol{x}^*).$$

两边取范数得

$$\frac{\|\delta\boldsymbol{x}^*\|}{\|\boldsymbol{x}^* + \delta\boldsymbol{x}^*\|} \le \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|\frac{\|\delta\boldsymbol{A}\|}{\|\boldsymbol{A}\|}. \tag{2.2}$$

定义 2.10 设 $\boldsymbol{A}$ 为非奇异矩阵, 称 $\|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|$ 为矩阵 $\boldsymbol{A}$ 的条件数, 记为

$$\mathrm{cond}(\boldsymbol{A}) = \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|.$$

常用的条件数

(1) $\mathrm{cond}(\boldsymbol{A})_\infty = \|\boldsymbol{A}\|_\infty \|\boldsymbol{A}^{-1}\|_\infty$.

(2) $\mathrm{cond}(\boldsymbol{A})_2 = \|\boldsymbol{A}\|_2 \|\boldsymbol{A}^{-1}\|_2 = \sqrt{\dfrac{\lambda_{\max}(\boldsymbol{A}^T\boldsymbol{A})}{\lambda_{\min}(\boldsymbol{A}^T\boldsymbol{A})}}$.

当 $\boldsymbol{A}$ 对称正定时, $\mathrm{cond}(\boldsymbol{A})_2 = \dfrac{\lambda_1}{\lambda_n}$

其中,$\lambda_{\max}(\boldsymbol{A}^T\boldsymbol{A})$ 和 $\lambda_{\min}(\boldsymbol{A}^T\boldsymbol{A})$ 分别为 $\boldsymbol{A}^T\boldsymbol{A}$ 的最大特征值和最小特征值, λ_1 和 λ_n 分别为 $\boldsymbol{A}$ 的最大特征值和最小特征值.

定义 2.11 对于方程组 $\boldsymbol{Ax} = \boldsymbol{b}$, 其中 $\boldsymbol{A}$ 非奇异. 如果 $\boldsymbol{A}$ 的条件数很大, 则称该方程组为病态（坏条件）方程组; 如果 $\boldsymbol{A}$ 的条件数比较小, 则称该方程组为良态（好条件）方程组.

可以计算例 (2.1) 中矩阵的条件数 $\mathrm{cond}(\boldsymbol{A})_\infty = 2$, 故例 (2.1) 中的方程组是良态的. 而例 (2.2) 中矩阵的条件数 $\mathrm{cond}(\boldsymbol{A})_\infty = 22002$, 故例例 (2.1) 中的方程组是病态的.

在实际中, 由于计算 $\boldsymbol{A}^{-1}$ 较困难, 所以 $\mathrm{cond}(\boldsymbol{A})$ 不容易计算. 但可以通过下面情形来判断病态矩阵:

(1) 用列主元 Gauss 消去法时出现绝对值很小的主元.

(2) 系数矩阵某些行 (列) 近似线性相关.

(3) 系数矩阵元素间数量级相差很大, 且没有一定规律.

对于病态方程组, 一般可以采用

(1) 双精度计算, 减少舍入误差;

(2) 对方程组进行预处理. 即选择非奇异矩阵 $\boldsymbol{D}, \boldsymbol{C}$, 将方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 化为等价的方程组 $\boldsymbol{DAC}[\boldsymbol{C}^{-1}\boldsymbol{x}] = \boldsymbol{Db}$, 使 $\boldsymbol{DAC}$ 的条件数比较小.

(3) 正则化方法.

2.4 方程组近似解可靠性的判别

设 $\tilde{\boldsymbol{x}}$ 为方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 的近似解, 记 $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\tilde{\boldsymbol{x}}$, 若 $\boldsymbol{r} = \boldsymbol{0}$, 则 $\tilde{\boldsymbol{x}}$ 为精确解. 一般 $\boldsymbol{r} \neq \boldsymbol{0}$.

问题 能否根据 $\|\boldsymbol{r}\|$ 的大小来判断近似解 $\tilde{\boldsymbol{x}}$ 的精确程度?

定理 2.8 设 $\tilde{\boldsymbol{x}}$ 是方程组 $\boldsymbol{Ax} = \boldsymbol{b}$ 的的一个近似解, $\boldsymbol{x}^*$ 为精确解, $\boldsymbol{r} = \boldsymbol{b} - \boldsymbol{A}\tilde{\boldsymbol{x}}$. 则

$$\frac{\|\boldsymbol{x}^* - \tilde{\boldsymbol{x}}\|}{\|\boldsymbol{x}^*\|} \leq \operatorname{cond}(\boldsymbol{A})\frac{\|\boldsymbol{r}\|}{\|\boldsymbol{b}\|}.$$

3 线性方程组的迭代法

3.1 迭代格式的构造

设 $\boldsymbol{A}$ 非奇异, $\boldsymbol{b} \neq \mathbf{0}$. 考虑方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, 设其解为 $\boldsymbol{x}^*$. 将方程组改写为等价的方程

$$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{x} + \boldsymbol{f}$$

这里 $\boldsymbol{B}$ 为 n 阶方阵, $\boldsymbol{f} \in \mathbf{R}^n$. 任取一个向量 $\boldsymbol{x}^{(0)} \in \mathbf{R}^n$, 产生迭代

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{B}\boldsymbol{x}^{(k)} + \boldsymbol{f}, \quad k = 0, 1, 2 \cdots . \tag{3.1}$$

上式产生一个向量序列 $\{\boldsymbol{x}^{(k)}\}_{k=0}^{\infty}$. 若它收敛于 $\bar{\boldsymbol{x}}$, 即

$$\lim_{k \to \infty} \boldsymbol{x}^{(k)} = \bar{\boldsymbol{x}},$$

(3.1) 两边取极限得 $\bar{\boldsymbol{x}} = \boldsymbol{B}\bar{\boldsymbol{x}} + \boldsymbol{f}$, 它等价于 $\boldsymbol{A}\bar{\boldsymbol{x}} = \boldsymbol{b}$. 故 $\bar{\boldsymbol{x}}$ 是原方程组的解, 即 $\boldsymbol{x}^* = \bar{\boldsymbol{x}}$.

(3.1) 称为**迭代格式**, $\boldsymbol{B}$ 为**迭代矩阵**, 称 $\boldsymbol{x}^{(k)}$ 为**第 $\boldsymbol{k}$ 次迭代近似解**, $\boldsymbol{e}^{(k)} = \boldsymbol{x}^* - \boldsymbol{x}^{(k)}$ 为第 k 次**迭代误差**. 如果迭代格式 (3.1) 对任意初始向量 $\boldsymbol{x}^{(0)}$ 产生的迭代序列 $\{\boldsymbol{x}^{(k)}\}_{k=0}^{\infty}$ 均收敛, 称该迭代收敛.

问题

(1) 怎样构造迭代格式? (2) 迭代何时收敛?

3.2 三个常用的迭代格式

将线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 写为分量形式

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n = b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \cdots + a_{3n}x_n = b_3, \\ \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 \cdots + a_{nn}x_n = b_n. \end{cases}$$

假设 $a_{ii} \neq 0 \, (i = 1, 2 \cdots, n)$.

(1) Jacobi 迭代格式

在上述方程组中第 i 个方程求出 x_i 得

$$\begin{cases} x_1 = (b_1 - a_{12}x_2 - a_{13}x_3 - \cdots - a_{1n}x_n)/a_{11}, \\ x_2 = (b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n)/a_{22}, \\ x_3 = (b_3 - a_{31}x_1 - a_{32}x_2 - \cdots - a_{3n}x_n)/a_{33}, \\ \qquad \vdots \\ x_n = (b_n - a_{n1}x_1 - a_{n2}x_2 - \cdots - a_{n,n-1}x_{n-1})/a_{nn}. \end{cases}$$

从而产生下面的 Jacobi 迭代格式

$$\begin{cases} x_1^{(k+1)} = (b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)})/a_{11}, \\ x_2^{(k+1)} = (b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)})/a_{22}, \\ x_3^{(k+1)} = (b_3 - a_{31}x_1^{(k)} - a_{32}x_2^{(k)} - \cdots - a_{3n}x_n^{(k)})/a_{33}, \\ \qquad\qquad \vdots \\ x_n^{(k+1)} = (b_n - a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)})/a_{nn}, \end{cases}$$

$$k = 0, 1, \cdots .$$

下面将 Jacobi 迭代用矩阵和向量表示. 记

$$\boldsymbol{L} = \begin{bmatrix} 0 & & & & & \\ a_{21} & 0 & & & & \\ a_{31} & a_{32} & 0 & & & \\ \vdots & \vdots & \vdots & \ddots & & \\ a_{n-1,1} & a_{n-1,2} & a_{n-1,3} & \cdots & 0 & \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{n,n-1} & 0 \end{bmatrix},$$

$$
\boldsymbol{D} = \begin{bmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & a_{33} & & \\ & & & \ddots & \\ & & & & a_{nn} \end{bmatrix}, \quad \boldsymbol{U} = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & a_{n,n-1} \\ & & & & 0 \end{bmatrix},
$$

则

$$
\begin{aligned}
&\boldsymbol{A} = \boldsymbol{L} + \boldsymbol{D} + \boldsymbol{U} \quad \Longrightarrow \\
&\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Longleftrightarrow (\boldsymbol{L} + \boldsymbol{D} + \boldsymbol{U})\boldsymbol{x} = \boldsymbol{b} \\
&\Longrightarrow \boldsymbol{D}\boldsymbol{x} = -(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} + \boldsymbol{b} \\
&\Longrightarrow \quad \boldsymbol{x} = -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})\boldsymbol{x} + \boldsymbol{D}^{-1}\boldsymbol{b}.
\end{aligned}
$$

Jacobi 迭代格式:

$$
\boldsymbol{x}^{(k+1)} = \boldsymbol{J}\boldsymbol{x}^{(k)} + \boldsymbol{f}_J, \quad k = 0, 1, \cdots,
$$

其中

$$
\boldsymbol{J} = -\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U}), \quad \boldsymbol{f}_J = \boldsymbol{D}^{-1}\boldsymbol{b}.
$$

(2) Gauss-Seidel 迭代格式

在 Jacobi 迭代中将已经求出的分量直接参加下一个分量的计算, 得到下面的 Gauss-Seidel 迭代格式:

$$\begin{cases} x_1^{(k+1)} = (b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)})/a_{11}, \\ x_2^{(k+1)} = (b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)})/a_{22}, \\ x_3^{(k+1)} = (b_3 - a_{31}x_1^{(k+1)} - a_{32}x_2^{(k+1)} - a_{34}x_4^{(k)} - \cdots - a_{3n}x_n^{(k)})/a_{33}, \\ \qquad \vdots \\ x_n^{(k+1)} = (b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)})/a_{nn}, \end{cases}$$

$$k = 0, 1 \cdots .$$

用矩阵表示为:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{D}^{-1}(\boldsymbol{b} - \boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{U}\boldsymbol{x}^{(k)}),$$
$$\Longrightarrow (\boldsymbol{D} + \boldsymbol{L})\boldsymbol{x}^{(k+1)} = \boldsymbol{b} - \boldsymbol{U}\boldsymbol{x}^{(k)}.$$

Gauss-Seidel 迭代格式:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{G}\boldsymbol{x}^{(k)} + \boldsymbol{f}_G, \quad k = 0, 1, \cdots ,$$

其中

$$\boldsymbol{G} = -(\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{U}, \quad \boldsymbol{f}_G = (\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{b}.$$

(3) SOR(逐次超松弛) 迭代格式

已知 $\boldsymbol{x}^{(k)} = (\boldsymbol{x}_1^{(k)}, \boldsymbol{x}_2^{(k)}, \cdots, \boldsymbol{x}_n^{(k)})^T$, 将 Gauss-Seidel 迭代中的 $\boldsymbol{x}^{(k+1)}$ 与 $\boldsymbol{x}^{(k)}$ 加权平均, 得到 SOR 迭代:

$$\begin{cases} x_1^{(k+1)} = (1-\omega)x_1^{(k)} + \omega(b_1 - a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)})/a_{11}, \\ x_2^{(k+1)} = (1-\omega)x_2^{(k)} + \omega(b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)})/a_{22}, \\ \vdots \\ x_n^{(k+1)} = (1-\omega)x_n^{(k)} + \omega(b_n - a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)})/a_{nn}, \end{cases}$$

$$k = 0, 1, \cdots,$$

其中 ω 称为松弛因子. 当 $\omega = 1$, 上式成为 Gauss-Seidel 迭代.

将 SOR 迭代用矩阵表示:

$$\begin{aligned} x^{(k+1)} &= (1-\omega)x^{(k)} + \omega \boldsymbol{D}^{-1}(\boldsymbol{b} - \boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{U}\boldsymbol{x}^{(k)}), \Longrightarrow \\ \boldsymbol{D}\boldsymbol{x}^{(k)} &= (1-\omega)\boldsymbol{D}\boldsymbol{x}^{(k)} + \omega(\boldsymbol{b} - \boldsymbol{L}\boldsymbol{x}^{(k+1)} - \boldsymbol{U}\boldsymbol{x}^{(k)}), \Longrightarrow \\ (\boldsymbol{D} + \omega\boldsymbol{L})\boldsymbol{x}^{(k+1)} &= [(1-\omega)\boldsymbol{D} - \omega\boldsymbol{U}]\boldsymbol{x}^{(k)} + \omega\boldsymbol{b}, \Longrightarrow \\ \boldsymbol{x}^{(k+1)} &= \boldsymbol{S}_\omega \boldsymbol{x}^{(k)} + \boldsymbol{f}_\omega, \end{aligned}$$

其中

$$\boldsymbol{S}_\omega = (\boldsymbol{D} + \omega\boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D} - \omega\boldsymbol{U}], \quad \boldsymbol{f}_\omega = \omega(\boldsymbol{D} + \omega\boldsymbol{L})^{-1}\boldsymbol{b}.$$

3.3 迭代格式的收敛性

(1) 迭代法基本定理

定理 3.1 迭代格式 $\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{f} \quad (k=0,1,2\cdots)$ 收敛 $\Longleftrightarrow \rho(\boldsymbol{B})<1.$

定理 3.2 若迭代矩阵 $\|\boldsymbol{B}\|<1$, 则迭代 $\boldsymbol{x}^{(k+1)}=\boldsymbol{B}\boldsymbol{x}^{(k)}+\boldsymbol{f}$ 收敛.

我们还可以证明：若 $\|\boldsymbol{B}\|<1$, 则对 $k=1,2,3,\cdots$ 有

$$\left\|\boldsymbol{x}^*-\boldsymbol{x}^{(k)}\right\| \leq \frac{\|\boldsymbol{B}\|}{1-\|\boldsymbol{B}\|}\left\|\boldsymbol{x}^{(k)}-\boldsymbol{x}^{(k-1)}\right\|,$$

$$\left\|\boldsymbol{x}^*-\boldsymbol{x}^{(k)}\right\| \leq \frac{\|\boldsymbol{B}\|^k}{1-\|\boldsymbol{B}\|}\left\|\boldsymbol{x}^{(1)}-\boldsymbol{x}^{(0)}\right\|.$$

(2) Jacobi 迭代和 Gauss-Seidel 迭代的收敛性

由定理 3.1, Jacobi 迭代收敛 $\Longleftrightarrow \rho(\boldsymbol{J})<1$. $\boldsymbol{J}$ 的特征方程为

$$|\lambda \boldsymbol{I}-\boldsymbol{J}|=|\lambda \boldsymbol{I}+\boldsymbol{D}^{-1}(\tilde{\boldsymbol{L}}+\boldsymbol{U})|=|\boldsymbol{D}^{-1}(\lambda \boldsymbol{D}+\boldsymbol{L}+\boldsymbol{U})|=0.$$

$$\Longleftrightarrow |\lambda \boldsymbol{D}+\boldsymbol{L}+\boldsymbol{U}|=0.$$

例 3.1 给定线性方程组

$$\begin{bmatrix} 8 & -1 & 1 \\ 2 & 10 & -1 \\ 1 & 1 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}.$$

(1) 写出 Jocobi 迭代格式.

(2) 分析 Jocobi 迭代格式的收敛性.

解 (1) Jacobi迭代格式为

$$\begin{cases} x_1^{(k+1)} = (1 + x_2^{(k)} - x_3^{(k)})/8, \\ x_2^{(k+1)} = (4 - 2x_1^{(k)} + x_3^{(k)})/10, \\ x_3^{(k+1)} = (3 - x_1^{(k)} - x_2^{(k)})/(-5). \end{cases}$$

(2) Jacobi 迭代矩阵 $\boldsymbol{J}$ 的特征方程为

$$\begin{vmatrix} 8\lambda & -1 & 1 \\ 2 & 10\lambda & -1 \\ 1 & 1 & -5\lambda \end{vmatrix} = 0,$$

展开得

$$400\lambda^3 + 12\lambda - 3 = 0,$$

用 Newton 迭代法求得一个实根为 $\lambda_1 = 0.146084$. 记另两根为 λ_2, λ_3, 通过待定系数可得

$$(\lambda - \lambda_1)(400\lambda^3 + b\lambda + c) = 0,$$

其中

$$b = \frac{1}{400\lambda_1}, \quad c = \frac{3}{\lambda_1}.$$

易知 λ_1, λ_2 为共轭复根, 且有

$$|\lambda_2\lambda_3| = \frac{3}{400\lambda_1},$$

$$|\lambda_1| = |\lambda_2| = \sqrt{\frac{3}{400\lambda_1}} = 0.226584,$$

所以

$$\rho(\boldsymbol{J}) = \max_{i=1,2,3}\{|\lambda_i|\} = 0.226584 < 1,$$

Jacobi 迭代收敛.

同样由定理 3.1, Gauss-Seidel 迭代收敛 $\Longleftrightarrow \rho(\boldsymbol{G}) < 1$. $\boldsymbol{G}$ 的特征方程为

$$
\begin{aligned}
&|\lambda \boldsymbol{I} - \boldsymbol{G}| = |\lambda \boldsymbol{I} + (\boldsymbol{D} + \boldsymbol{L})^{-1}\boldsymbol{U}| \\
&= |(\boldsymbol{D} + \boldsymbol{L})^{-1}(\lambda(\boldsymbol{D} + \boldsymbol{L}) + \boldsymbol{U}| = 0. \\
&\Longleftrightarrow |(\lambda(\boldsymbol{D} + \boldsymbol{L}) + \boldsymbol{U}| = 0.
\end{aligned}
$$

例 3.2 给定线性方程组

$$\begin{bmatrix} 8 & -1 & 1 \\ 2 & 10 & -1 \\ 1 & 1 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}.$$

(1) 写出 Gauss-Seidel 迭代格式.

(2) 分析 Gauss-Seidel 迭代格式的收敛性.

解 (1) Gauss-Seidel 迭代格式为

$$\begin{cases} x_1^{(k+1)} = (1 + x_2^{(k)} - x_3^{(k)})/8, \\ x_2^{(k+1)} = (4 - 2x_1^{(k+1)} + x_3^{(k)})/10, \\ x_3^{(k+1)} = (3 - x_1^{(k+1)} - x_2^{(k+1)})/(-5). \end{cases}$$

(2) Gauss-Seidel 迭代矩阵 $\boldsymbol{G}$ 的特征方程为

$$\begin{vmatrix} 8\lambda & -1 & 1 \\ 2\lambda & 10\lambda & -1 \\ \lambda & \lambda & -5\lambda \end{vmatrix} = 0,$$

展开得

$$\lambda(400\lambda^2 + 10\lambda - 1),$$

求得

$$\lambda_1 = 0, \quad \lambda_2 = \frac{-1+\sqrt{17}}{80}, \quad \lambda_2 = \frac{-1-\sqrt{17}}{80},$$
$$\rho(\boldsymbol{G}) = \max_{i=1,2,3}\{|\lambda_i|\} = 0.0640388 < 1,$$

Gauss-Seidel 迭代收敛.

定义 3.1 设

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

如果

$$|a_{ii}| > \sum_{\substack{j=1\\ i\neq j}}^{n} |a_{ij}|, \quad i=1,2,\cdots,n$$

则称 $\boldsymbol{A}$ 按行严格对角占优；如果

$$|a_{jj}| > \sum_{\substack{i=1\\ j\neq i}}^{n} |a_{ij}|, \quad j=1,2,\cdots,n$$

则称 $\boldsymbol{A}$ 按列严格对角占优。按行严格对角占优或按列严格对角占优统称为严格对角占优。

引理 3.1 设 $\boldsymbol{A}$ 是严格对角占优矩阵, 则 $|\boldsymbol{A}| \neq 0$.

证 这里仅证 A 是按行严格对角占优的情况. 用反证法.

设 $|\boldsymbol{A}| \neq 0$, 则齐次方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 有非零解 $\boldsymbol{x}^* = (x_1^*, x_2^*, \cdots, x_n^*)^T$.

设 $\|\boldsymbol{x}^*\|_\infty = |x_k^*| \neq 0$. 由第 k 个方程

$$a_{kk}x_k^* + \sum_{j=1,j\neq k}^{n} a_{kj}x_j^* = 0$$

可得

$$\begin{aligned} |a_{kk}| \cdot |x_k^*| &= |\sum_{j=1,j\neq k}^{n} a_{kj}x_j^*| \leq \sum_{j=1,j\neq k}^{n} |a_{kj}| \cdot |x_j^*| \\ &\leq \sum_{j=1,j\neq k}^{n} |a_{kj}| \cdot |x_k^*| \end{aligned}$$

两边约去 $|x_k^*|$, 得

$$|a_{kk}| \leq \sum_{j=1,j\neq k}^{n} |a_{kj}|$$

与按行严格对角占优矛盾. 因而 $|\boldsymbol{A}| \neq 0$.

定理 3.3 给定线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, 如果 $\boldsymbol{A}$ 是严格对角占优矩阵, 则 Jacobi 迭代格式收敛.

证明 记

$$\boldsymbol{B}(\lambda)=\begin{pmatrix}\lambda a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & \lambda a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & \lambda a_{nn}\end{pmatrix}$$

则 Jacobi 迭代矩阵 J 的特征方程为 $|\boldsymbol{B}(\lambda)|=0$. 设 $\boldsymbol{A}$ 是按行严格对角占优的, 则当 $|\lambda|\geq 1$ 时, 有

$$|\lambda a_{ii}|\geq|a_{ii}|>\sum_{j=1,j\neq k}^{n}|a_{ij}|\ \ (i=1,2,\cdots,n)$$

即 $\boldsymbol{B}(\lambda)$ 是严格按行对角占优的。由上面的一个引理可知, 当 $|\lambda|\geq 1$ 时, $|\boldsymbol{B}(\lambda)|\neq 0$. 换句话说, 方程 $|\boldsymbol{B}(\lambda)|=0$ 的 n 个根 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都应满足 $|\lambda_i|<1$, $i=1,2,\cdots,n$. 于是 $\rho(\boldsymbol{J})<1$. 同样可证当 $\boldsymbol{A}$ 按列严格对角占优的情况。因而 Jacobi 迭代格式收敛。

定理 3.4 给定线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, 如果 $\boldsymbol{A}$ 是严格对角占优矩阵, 则 Gauss-Seidel 迭代格式收敛.

证 记

$$C(\lambda)=\begin{pmatrix}\lambda a_{11} & a_{12} & \cdots & a_{1n}\\ \lambda a_{21} & \lambda a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nn}\end{pmatrix}$$

则 Gauss-Seidel 迭代矩阵 G 的特征方程为 $|C(\lambda)|=0$. 设 A 是按行严格对角占优的, 则当 $|\lambda|\geq 1$ 时, 有

$$\begin{aligned}|\lambda a_{ii}| &= |\lambda||a_{ii}|>|\lambda|\sum_{j=1,j\neq k}^{n}|a_{ij}|\\ &\geq |\lambda|\sum_{j=1}^{i-1}|a_{ij}|+\sum_{j=i+1}^{n}|a_{ij}| \quad i=1,2,\cdots,n\end{aligned}$$

即 $C(\lambda)$ 是严格按行对角占优的。由上面的一个引理可知, 当 $|\lambda|\geq 1$ 时, $|C(\lambda)|\neq 0$. 换句话说, 方程 $|C(\lambda)|=0$ 的 n 个根 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都应满足 $|\lambda_i|<1, i=1,2,\cdots,n$. 于是 $\rho(G)<1$. 同样可证当 A 按列严格对角占优的情况。因而 Gauss-Seidel 迭代格式收敛。

注 3.1 由定理3.3 和定理3.4 知, 例3.1 和例3.2 中的系数矩阵严格对角占优, 故 Jacobi 迭代和 Gauss-Seidel 迭代均收敛.

(3) SOR 迭代的收敛性

SOR 迭代的迭代矩阵为

$$S_\omega = (\boldsymbol{D} + \omega \boldsymbol{L})^{-1}[(1-\omega)\boldsymbol{D} - \omega \boldsymbol{U}],$$

由定理 3.1, SOR 迭代收敛 $\Longleftrightarrow \rho(S_\omega) < 1.$

定理 3.5 SOR 迭代收敛的必要条件是 $0 < \omega < 2.$

证 设 S_ω 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则由线性代数知

$$|\det(S_\omega)| = |\lambda_1\lambda_2\cdots\lambda_n| \le \rho(S_\omega)^n.$$

另一方面, 由定理 3.1, 若 SOR 收敛, 则 $\rho(S_\omega) < 1$, 从而有 $|\det(S_\omega)| < 1$. 而行列式

$$\begin{aligned}\det(S_\omega) &= \det[(\boldsymbol{D}+\omega\boldsymbol{L})^{-1}]\det[(1-\omega)\boldsymbol{D}-\omega\boldsymbol{U}] \\ &= \left(\prod_{i=1}^n a_{ii}\right)^{-1}\prod_{i=1}^n[(1-\omega)a_{ii}] \\ &= (1-\omega)^n.\end{aligned}$$

从而有

$$|(1-\omega)^n| < 1, \quad \Longrightarrow 0 < \omega < 2.$$

定理 3.6 给定线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$. 如果 $\boldsymbol{A}$ 对称正定, 且 $0<\omega<2$, 则 SOR 迭代收敛.

注 3.2 如果 $\boldsymbol{A}$ 对称正定, 则 Gauss-Seidel 迭代收敛.

例 3.3 给定线性方程组

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

(a) 分别写出解该方程组的 Jacobi 迭代和 Gauss-Seidel 迭代格式.

(b) 分析这两种迭代格式的收敛性.

解

(a) Jacobi 迭代格式为

$$\begin{cases} x_1^{(k+1)} = (-x_2^{(k)} - x_3^{(k)} + 1)/2, \\ x_2^{(k+1)} = (-x_1^{(k)} - x_3^{(k)} + 2)/2, \qquad k = 0, 1, \cdots . \\ x_3^{(k+1)} = (-x_1^{(k)} - x_2^{(k)} + 3)/2. \end{cases}$$

Gauss-Seidel 迭代格式为

$$\begin{cases} x_1^{(k+1)} = (-x_2^{(k)} - x_3^{(k)} + 1)/2, \\ x_2^{(k+1)} = (-x_1^{(k+1)} - x_3^{(k)} + 2)/2, \quad k = 0, 1, \cdots . \\ x_3^{(k+1)} = (-x_1^{(k+1)} - x_2^{(k+1)} + 3)/2. \end{cases}$$

(b) Jacobi 迭代矩阵 $\boldsymbol{J}$ 的特征方程为

$$\begin{vmatrix} 2\lambda & 1 & 1 \\ 1 & 2\lambda & 1 \\ 1 & 1 & 2\lambda \end{vmatrix} = 0,$$

展开得

$$4\lambda^3 - 3\lambda + 1 = 0,$$

即有

$$(2\lambda - 1)^2(\lambda + 1) = 0,$$

求得

$$\lambda_{1,2} = \frac{1}{2}, \quad \lambda_3 = -1,$$

所以 $\rho(\boldsymbol{J}) = 1$, 故 Jacobi 迭代发散.

Gauss-Seidel 迭代矩阵 $\boldsymbol{G}$ 的特征方程是

$$\begin{vmatrix} 2\lambda & 1 & 1 \\ \lambda & 2\lambda & 1 \\ \lambda & \lambda & 2\lambda \end{vmatrix} = 0,$$

展开得

$$\lambda(8\lambda^2 - 2\lambda + 1) = 0,$$

求得

$$\lambda_1 = 0, \quad \lambda_{2,3} = \frac{1 \pm \sqrt{7}\mathbf{i}}{8},$$

所以 $\rho(\boldsymbol{G}) = \dfrac{\sqrt{2}}{4} < 1$, 故 Gauss-Seidel 迭代收敛.

例 3.4 给定线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, $\boldsymbol{A}$ 为 n 阶非奇异矩阵. 构造迭代

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\omega(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(k)}),\quad (k=0,1,2,\cdots.)$$

其中 $\omega\neq 0$ 为常数.

(a) 证明: 如果迭代收敛, 则迭代序列 $\{\boldsymbol{x}^{(k)}\}_{k=0}^{\infty}$ 收敛于方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解.

(b) 设 $\boldsymbol{A}=\begin{bmatrix}2&1&1\\1&2&1\\1&1&2\end{bmatrix}$, 则 ω 取何值时迭代收敛?

解

(a) 证　设 $\lim\limits_{k\to\infty}\boldsymbol{x}^{(k)}=\boldsymbol{x}^*$. 在迭代格式两边取极限

$$\lim_{k\to\infty}\boldsymbol{x}^{(k+1)}=\lim_{k\to\infty}[\boldsymbol{x}^{(k)}+\omega(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(k)})].$$

得 $\boldsymbol{x}^*=\boldsymbol{x}^*+\omega(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^*)$. 因为 $\omega\neq 0$, 所以 $\boldsymbol{A}\boldsymbol{x}^*=\boldsymbol{b}$, 即 $\boldsymbol{x}^*$ 是方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解.

(b) 迭代格式的迭代矩阵为 $\boldsymbol{I}-\omega\boldsymbol{A}$, 迭代收敛 $\Longleftrightarrow$ $\rho(\boldsymbol{I}-$

$\omega \boldsymbol{A}) < 1$. 矩阵 $\boldsymbol{I} - \omega \boldsymbol{A}$ 的特征方程为

$$
\begin{aligned}
& |\lambda \boldsymbol{I} - \boldsymbol{I} - \omega \boldsymbol{A}| \\
= & \begin{vmatrix} (\lambda - 1 + 2\omega) & \omega & \omega \\ \omega & (\lambda - 1 + 2\omega) & \omega \\ \omega & \omega & (\lambda - 1 + 2\omega) \end{vmatrix} = 0, \\
& \text{记}\mu = (\lambda - 1 + 2\omega), \Longrightarrow \mu^3 - 3\omega^2\mu + 2\omega^3 = 0, \\
& \Longrightarrow (\mu - \omega)^2(\mu + 2\omega) = 0. \\
& \Longrightarrow \mu = \omega \text{或} \mu = -2\omega, \Longrightarrow \lambda = 1 - \omega \text{或} \lambda = 1 - 4\omega. \\
& \begin{cases} |1 - \omega| < 1 \\ |1 - 4\omega| < 1 \end{cases} \Longrightarrow 0 < \omega < \frac{1}{2}.
\end{aligned}
$$

4 幂法与反幂法

幂法和反幂法是迭代法. 幂法用于求矩阵按模最大的特征值和对应的特征向量. 当特征值非零时, 反幂法用于求按模最小的特征值和对应的特征向量.

4.1 求主特征值的幂法

设 $\boldsymbol{A} = (a_{ij})_{n\times n}$ 是 n 阶方阵, 它有 n 个线性无关的特征向量 $\boldsymbol{x}_1$, $\boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$, 对应的特征值为 $\lambda_j(j = 1, 2 \cdots, n)$, 按模的大小排列

$$|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|,$$

其中 λ_1 是主特征值. 给定初值非零向量 $\boldsymbol{v}_0$, 构造迭代

$$\boldsymbol{v}_k = \boldsymbol{A}\boldsymbol{v}_{k-1}, \quad k = 1, 2, \cdots .$$

$\boldsymbol{v}_0$ 可由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ 线性表示, 设表示为

$$\boldsymbol{v}_0 = \sum_{i=1}^{n} a_i \boldsymbol{x}_i,$$

且 $a_1 \neq 0$. 因此有

$$\boldsymbol{v}_k = \boldsymbol{A}\boldsymbol{v}_{k-1} = \boldsymbol{A}^k\boldsymbol{v}_0 = \boldsymbol{A}^k\sum_{i=1}^{n} a_i\boldsymbol{x}_i = \sum_{i=1}^{n} a_i\lambda_i^k\boldsymbol{x}_i, \quad k = 1, 2, \cdots.$$

(1) 特征值满足 $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. 则

$$\boldsymbol{v}_k = \lambda_1^k\left[a_1\boldsymbol{x}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\boldsymbol{x}_2 + \cdots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\boldsymbol{x}_n\right]$$

$$\boldsymbol{v}_{k+1} = \lambda_1^{k+1}\left[a_1\boldsymbol{x}_1 + a_2\left(\frac{\lambda_2}{\lambda_1}\right)^{k+1}\boldsymbol{x}_2 + \cdots + a_n\left(\frac{\lambda_n}{\lambda_1}\right)^{k+1}\boldsymbol{x}_n\right]$$

因为

$$|\lambda_1| > |\lambda_i| \quad i = 2, 3, \cdots, n,$$

故

$$\lim_{k\to\infty}\left(\frac{\lambda_i}{\lambda_1}\right)^k = 0, \quad i = 2, 3, \cdots, n.$$

从而当 k 充分大, 有

$$\boldsymbol{v}_k \approx a_1\lambda_1^k\boldsymbol{x}_1, \quad \boldsymbol{v}_{k+1} \approx a_1\lambda_1^{k+1}\boldsymbol{x}_1 \approx \lambda_1\boldsymbol{v}_k. \tag{4.1}$$

由上式知 $\boldsymbol{v}_{k+1}=\boldsymbol{A}\boldsymbol{v}_k\approx\lambda_1\boldsymbol{v}_k$. 该式说明 $\boldsymbol{v}_k$ 是 λ_1 对应的近似特征向量, $\boldsymbol{v}_k$ 和 $\boldsymbol{v}_{k+1}$ 近似线性相关. 所以

$$\lambda_1=\lim_{k\to\infty}\frac{(\boldsymbol{v}_{k+1})_i}{(\boldsymbol{v}_k)_i},$$

其中 $(\boldsymbol{v}_k)_i$ 表示 $\boldsymbol{v}_k$ 的第 i 个分量. 实际计算时, 为了避免 $\boldsymbol{v}_k$ 的分量产生溢出, 可以采用“归一化”, 具体将算法改为:

$$\begin{cases}\boldsymbol{u}_0=\boldsymbol{v}_0,\\ \boldsymbol{v}_k=\boldsymbol{A}\boldsymbol{u}_{k-1},\\ m_k=\max(\boldsymbol{v}_k),\\ \boldsymbol{u}_k=\boldsymbol{v}_k/m_k,\end{cases}\qquad k=1,2,\cdots, \tag{4.2}$$

其中 $m_k=\max(\boldsymbol{v}_k)$ 表示 $\boldsymbol{v}_k$ 中 (首次出现的) 绝对值最大的分量.

定理 4.1 设 $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$, 则由算法 (4.2) 产生的序列 $\{\boldsymbol{u}_k\}$ 和 $\{m_k\}$ 均收敛, 且

$$\lim_{k\to\infty} \boldsymbol{u}_k = \frac{\boldsymbol{x}_1}{\max(\boldsymbol{x}_1)}, \qquad \lim_{k\to\infty} m_k = \lambda_1.$$

证

$$\boldsymbol{u}_k = \frac{1}{m_k}\boldsymbol{v}_k = \frac{1}{m_k}\boldsymbol{A}\boldsymbol{u}_{k-1}, \quad k = 1, 2, \cdots ,$$

利用上式递推得

$$\begin{aligned}
\boldsymbol{u}_k = & \frac{1}{m_k}\boldsymbol{A}\left(\frac{1}{m_{k-1}}\boldsymbol{A}\boldsymbol{u}_{k-2}\right) \\
= & \frac{1}{m_k m_{k-1}}\boldsymbol{A}^2\boldsymbol{u}_{k-2} \\
= & \cdots \\
= & \frac{1}{m_k m_{k-1}\cdots m_1}\boldsymbol{A}^k\boldsymbol{u}_0.
\end{aligned}$$

由 $\boldsymbol{u}_k$ 的归一化, $m_k m_{k-1}\cdots m_1 = \dfrac{\max(\boldsymbol{A}^k\boldsymbol{u}_0)}{\max(\boldsymbol{u}_k)} = \max(\boldsymbol{A}^k\boldsymbol{u}_0)$,

$$\boldsymbol{u}_k = \frac{\boldsymbol{A}^k\boldsymbol{u}_0}{\max(\boldsymbol{A}^k\boldsymbol{u}_0)}$$

$$= \frac{\lambda_1^k \left(a_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \boldsymbol{x}_i \right)}{\max \left(\lambda_1^k \left(a_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \boldsymbol{x}_i \right) \right)}$$

$$= \frac{a_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \boldsymbol{x}_i}{\max \left(a_1 \boldsymbol{x}_1 + \sum\limits_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \boldsymbol{x}_i \right)}.$$

从而

$$\lim_{k \to \infty} \boldsymbol{u}_k = \frac{\boldsymbol{x}_1}{\max(\boldsymbol{x}_1)}.$$

同样

$$\boldsymbol{v}_k = \quad \boldsymbol{A} \boldsymbol{u}_{k-1} = \frac{\boldsymbol{A}^k \boldsymbol{u}_0}{\max(\boldsymbol{A}^{k-1} \boldsymbol{u}_0)}$$

$$= \frac{\lambda_1 \max\left(a_1\boldsymbol{x}_1 + \sum_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{x}_i\right)}{\max\left(a_1\boldsymbol{x}_1 + \sum_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k-1} \boldsymbol{x}_i\right)}.$$

因此

$$\lim_{k\to\infty} \max(\boldsymbol{v}_k) = \lambda_1.$$

从以上分析知幂法的收敛速度与 $\left|\frac{\lambda_2}{\lambda_1}\right|$ 有关, 称比值 $\left|\frac{\lambda_2}{\lambda_1}\right|$ 为幂法的收敛速率.

例 4.1 用幂法计算矩阵

$$\boldsymbol{A}=\begin{bmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{bmatrix}$$

的主特征值及对应的特征向量.

解 取 $\boldsymbol{u}_0=(0,0,1)^T$, 由 (4.2) 构造迭代

$$\begin{aligned} \boldsymbol{u}_0 &= \boldsymbol{v}_0, \\ \boldsymbol{v}_k &= \boldsymbol{A}\boldsymbol{u}_{k-1}, \\ m_k &= \max(\boldsymbol{v}_k), \\ \boldsymbol{u}_k &= \boldsymbol{v}_k/m_k, \quad k=1,2,\cdots. \end{aligned}$$

计算结果见表6.1.

表 6.1　幂法求主特征值

k	$\boldsymbol{u}_k^T$	m_k
0	$(0, 0, 1)$	1
1	$(0.5, 1.0, 0.25)$	4
2	$(0.5, 1.0, 0.8611)$	9
3	$(0.5, 1.0, 0.7306)$	11.4400
4	$(0.5, 1.0, 0.7535)$	10.9224
5	$(0.5, 1.0, 0.7493)$	11.0140
6	$(0.5, 1.0, 0.7501)$	10.9927
7	$(0.5, 1.0, 0.7500)$	11.0004
8	$(0.5, 1.0, 0.7500)$	11.0000

(2) 当 $|\lambda_1| = |\lambda_2|$, 且 $|\lambda_2| > |\lambda_3|$ 时.

(a) $\lambda_1 = \lambda_2$.

$$\boldsymbol{u}_k = \frac{\boldsymbol{A}^k \boldsymbol{u}_0}{\max(\boldsymbol{A}^k \boldsymbol{u}_0)} = \frac{a_1 \boldsymbol{x}_1 + a_2 \boldsymbol{x}_2 + \sum_{i=3}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{x}_i}{\max\left(a_1 \boldsymbol{x}_1 + a_2 \boldsymbol{x}_2 + \sum_{i=3}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{x}_i\right)}$$

因此

$$\lim_{k\to\infty} \boldsymbol{u}_k = \frac{a_1 \boldsymbol{x}_1 + a_2 \boldsymbol{x}_2}{\max(a_1 \boldsymbol{x}_1 + a_2 \boldsymbol{x}_2)}.$$

$$\lim_{k\to\infty} m_k = \lim_{k\to\infty} \max(\boldsymbol{v}_k) = \lambda_1.$$

(b) $\lambda_1 = -\lambda_2$

(c) $\lambda_1 = \bar{\lambda}_2$

4.2 反幂法

当 $\boldsymbol{A}$ 非奇异时, $\boldsymbol{A}$ 的特征值非零. $\boldsymbol{A}^{-1}$ 的按模最大特征值就是 $\boldsymbol{A}$ 的按模最小的特征值. 因此用幂法可以求 $\boldsymbol{A}^{-1}$ 的按模最大的特征值即 $\boldsymbol{A}$ 的按模最小的特征值, 这就是反幂法. 算法如下:

$$\begin{cases} \boldsymbol{u}_0 = \boldsymbol{v}_0 \\ \boldsymbol{A}\boldsymbol{v}_k = \boldsymbol{u}_{k-1} \\ m_k = \max(\boldsymbol{v}_k) \\ \boldsymbol{u}_k = \boldsymbol{v}_k / m_k \end{cases} \qquad k = 1, 2, \cdots .$$

易知

$$\lim_{k\to\infty} \boldsymbol{u}_k = \frac{\boldsymbol{x}_n}{\max(\boldsymbol{x}_n)}, \qquad \lim_{k\to\infty} m_k = \frac{1}{\lambda_n}.$$

解

$$
\begin{bmatrix} 3 & 1 & -1 & 4 \\ 4 & 0 & 4 & 8 \\ 12 & -3 & 3 & 9 \end{bmatrix}
\xrightarrow{r_1 \leftrightarrow r_3}
\begin{bmatrix} 12 & -3 & 3 & 9 \\ 4 & 0 & 4 & 8 \\ 3 & 1 & -1 & 4 \end{bmatrix}
\xrightarrow[-\frac{1}{4}r_1+r_3]{-\frac{1}{3}r_1+r_2}
$$

$$
\begin{bmatrix} 12 & -3 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & \dfrac{7}{4} & -\dfrac{7}{4} & \dfrac{7}{4} \end{bmatrix}
\xrightarrow{r_2 \leftrightarrow r_3}
\begin{bmatrix} 12 & -3 & 3 & 9 \\ 0 & \dfrac{7}{4} & -\dfrac{7}{4} & \dfrac{7}{4} \\ 0 & 1 & 3 & 5 \end{bmatrix}
\xrightarrow{-\frac{4}{7}r_2+r_3}
$$

$$
\begin{bmatrix} 12 & -3 & 3 & 9 \\ 0 & \dfrac{7}{4} & -\dfrac{7}{4} & \dfrac{7}{4} \\ 0 & 0 & 4 & 4 \end{bmatrix}.
$$

回代得 $x_1 = 1, \quad x_2 = 2, \quad x_3 = 1.$

最大值$\max\limits_{\substack{\boldsymbol{x}\in\mathbf{R}^n\\ \boldsymbol{x}\neq 0}}\dfrac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$能否取到? 由定理 2.1 知, 对 $\forall\boldsymbol{x}\in\mathbf{R}^n$, $\|\boldsymbol{A}\boldsymbol{x}\|$ 在 $\mathbf{R}^n$ 上连续. 记

$$\mathbf{S}=\{\boldsymbol{y}|\boldsymbol{y}\in\mathbf{R}^n,\|\boldsymbol{y}\|=1\},$$

$\mathbf{S}$ 是有界闭集. $\|\boldsymbol{A}\boldsymbol{x}\|$ 在 $\mathbf{S}$ 上能取到最大值, 设最大值为 M. 即存在 $\boldsymbol{y}_0\in\mathbf{S}$ 使得

$$\max_{\boldsymbol{y}\in\mathbf{S}}\|\boldsymbol{A}\boldsymbol{y}\|=\|\boldsymbol{A}\boldsymbol{y}_0\|=M.$$

故对 $\forall\boldsymbol{x}\in\mathbf{R}^n,\boldsymbol{x}\neq 0$, 有 $\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\in\mathbf{S}$, 令 $\boldsymbol{y}=\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$, 有

$$\max_{\substack{\boldsymbol{x}\in\mathbf{R}^n\\ \boldsymbol{x}\neq 0}}\frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}=\max_{\boldsymbol{y}\in S}\|\boldsymbol{A}\boldsymbol{y}\|=M.$$

5　习题

习题 3 p.120

1, 5(2), 9, 17, 19, 20, 23, 25, 27, 28, 30, 32(1)(2), 34, 39(上机题)