计算方法

北京师范大学数学科学学院
主讲人：纪光华
Email：ghji @bnu.edu.cn

第三章、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. 向量和矩阵的范数，方程组的性态和误差估计
3. Jacobi迭代、Gauss-Seidel迭代格式，收敛性的判别
4. 幂法和反幂法

3.1 Cramer法则

考虑一般的线性方程组

$$\mathbf{Ax} = \mathbf{b}, \tag{1}$$

其中

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

如果$det\mathbf{A} \neq 0$则方程组(1)有唯一解$x_i = \dfrac{D_i}{D}(i = 1, \cdots, n)$.

- 计算量：$(n+1) * n!(n-1) = (n-1)(n+1)!$次乘法，当$n = 100$时，$100! \times 99 \approx 9.24 \times 10^{159}$；
- 神威·太湖之光（4）：93015 TFlop/s，3.15×10^{135}年；
- Fugaku（富岳）：442010 TFlop/s;
- Summit: 148600 TFlop/s;
- https://www.chinastor.com/hpc-top500/

3.2 Gauss消去法

思想：利用线性代数中学习过的初等行变换将方程组化为等价的三角形方程组.

★Gauss消去法

将方程$\mathbf{Ax}=\mathbf{b}$用增广矩阵表示，记

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

其中

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

下面用$n-1$步消元（初等行变换）将矩阵$\bar{\mathbf{A}}^{(1)}$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$-l_{i1}$加到第行$i$ $(2 \leq i \leq n)$行得

$$\bar{\mathbf{A}}^{(1)} \longrightarrow \bar{\mathbf{A}}^{(2)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{i2}^{(2)} & \cdots & a_{in}^{(2)} & a_{i,n+1}^{(2)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{bmatrix}$$

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) 假设按上面进行了$k-1$步消元, 即有

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

其中$\bar{\mathbf{A}}^{(k)}$

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

3) 第k步消元：假设$a_{kk}^{(k)} \neq 0$ (否则交换两行的位置)，

记$l_{ik} = \dfrac{a_{ik}^{(k)}}{a_{kk}^{(k)}}$，第$k$行乘$-l_{ik}$加到第$i$ $(k+1 \leq i \leq n)$行得

$$\bar{\mathbf{A}}^{(k)} \longrightarrow \bar{\mathbf{A}}^{(k+1)}$$

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

其中

$$a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik}a_{kj}^{(k)},$$
$$i = k+1, k+2, \cdots, n, \quad j = k+1, k+2, \cdots, n+1. \qquad (2)$$

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

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

其中

$$\bar{\mathbf{A}}^{(n)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,n-1}^{(1)} & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2,n-1}^{(2)} & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & & \ddots & \vdots & \vdots & \vdots \\ & & & a_{n-1,n-1}^{(n-1)} & a_{n-1,n}^{(n-1)} & a_{n-1,n+1}^{(n-1)} \\ & & & & a_{nn}^{(n)} & a_{n,n+1}^{(n)} \end{bmatrix}. \qquad (3)$$

若记

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

则(3)对应的线性方程组为$\mathbf{U}\mathbf{x}=\mathbf{y}$, 用回代可求得$\mathbf{x}$.

消元过程乘除法次数$\dfrac{n^3}{3}+\dfrac{n^2}{2}-\dfrac{5n}{6}$, 加减法次数$\dfrac{n^3}{3}-\dfrac{n}{3}$.

★ 三角形方程组的回代法

考虑三角形方程组

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

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

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

计算量：乘除法：$\dfrac{n(n+1)}{2}$、加减法：$\dfrac{n(n-1)}{2}$.

可以估计，Gauss消元法解线性方程组(1)需要加减法和乘除法次数均为$O(n^3)$.

定理3.1

给定线性方程组$\mathbf{Ax}=\mathbf{b}$. 如果$\mathbf{A}$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.

★三对角方程组的追赶法 考虑三对角方程组

$$\begin{bmatrix} b_1 & c_1 & & & & \\ a_2 & b_2 & c_2 & & & \\ & a_3 & b_3 & c_3 & & \\ & & \ddots & \ddots & \ddots & \\ & & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_{n-1} \\ d_n \end{bmatrix} \tag{4}$$

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. $|b_n| > |a_n| > 0.$

方程组(4)的系数矩阵非奇异. 利用Gauss消去法解方程组(4), 每步消元只要消一个元素.

三消元过程算法如下:

$$\beta_1 = b_1, \quad y_1 = d_1,$$
$$l_i = \frac{a_i}{\beta_{i-1}}, \quad \beta_i = b_i - l_i c_{i-1}, \quad y_i = d_i - l_i y_{i-1} \quad (i = 2, 3, \cdots, n).$$

$$\begin{bmatrix} \beta_1 & c_1 & & & & \\ & \beta_2 & c_2 & & & \\ & & \beta_3 & c_3 & & \\ & & & \ddots & \ddots & \\ & & & & \beta_{n-1} & c_{n-1} \\ & & & & & \beta_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{n-1} \\ y_n \end{bmatrix}$$

回代算法:

$$x_n = y_n / \beta_n,$$
$$x_i = (y_i - c_i x_{i+1}) / \beta_i \quad (n = n-1, n-2, \cdots, 1).$$

追赶法计算次数，乘除法：5n-4次，加减法：3n-3次；

★列主元Gauss消去法 **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

$$\bar{\mathbf{A}}^{(k)} \longrightarrow \bar{\mathbf{A}}^{(k+1)}$$

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

其中

$$
\begin{aligned}
&a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)}, \\
&\quad i = k+1, k+2, \cdots, n, \quad j = k+1, k+2, \cdots, n+1. \text{(5)}
\end{aligned}
$$

由(5)看出，在Gauss消去法中，当$|l_{ik}|$很大(如$|a_{kk}^{(k)}|$很小)时，元素a_{kj}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.

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

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3.1

用列主元Gauss消去法求下列方程组的解.

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

解:

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

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

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

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

3.2 矩阵的LU分解

Gauss消去法第一步消元相当于$\mathbf{L}_1\bar{\mathbf{A}}^{(1)} = \bar{\mathbf{A}}^{(2)}$,其中矩阵

$$\mathbf{L}_1 = \begin{bmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & 0 & 1 & & \\ \vdots & \vdots & & \ddots & \\ -l_{n1} & 0 & & \cdots & 1 \end{bmatrix}$$

以此类推得$\mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_1\bar{\mathbf{A}}^{(1)} = \bar{\mathbf{A}}^{(n)}$,即$\mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_1(\mathbf{A}|b) = (\mathbf{U}|y)$.记$\mathbf{L} = \mathbf{L}_1^{-1}\cdots\mathbf{L}_{n-1}^{-1}$可得$\mathbf{A} = \mathbf{LU}$

定理3.2

如果$\mathbf{A}$的各阶顺序主子式非零,则对$\mathbf{A}$可作唯一的LU分解。

★LU分解

$$
\begin{aligned}
u_{1j} &= a_{1j}, (j = 1, 2, \cdots, n+1) \\
l_{i1} &= \frac{a_{i1}}{a_{11}}, (i = 2, 3, \cdots, n), \\
u_{kj} &= a_{kj} - \sum_{q=1}^{k-1} l_{kq} u_{qj}, (j = k, k+1, \cdots, n+1), \\
l_{ik} &= \left(a_{ik} - \sum_{q=1}^{k-1} l_{iq} u_{qk} \right) / u_{kk}, (i = k+1, k+2, \cdots, n)
\end{aligned}
$$

Remark:

1. 对称矩阵的LU分解–计算量减少一半.
2. 列选主元LU分解.

3.3 方程组的性态与误差分析

例3.2

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3.3

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$x_1 = x_2 = \frac{2}{3}$.

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?

★向量范数

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

定义3.1

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

1. (非负性) $\forall\mathbf{x}\in\mathbf{R}^n$, 有$\|\mathbf{x}\|\geq 0$, 且$\|\mathbf{x}\|=0\Leftrightarrow\mathbf{x}=\mathbf{0}$;
2. (齐次性) $\forall\mathbf{x}\in\mathbf{R}^n$, $\lambda\in\mathbf{R}$, 有$\|\lambda\mathbf{x}\|=|\lambda|\|\mathbf{x}\|$;
3. (三角不等式) $\forall\mathbf{x},\mathbf{y}\in\mathbf{R}^n$, 有$\|\mathbf{x}+\mathbf{y}\|\leq\|\mathbf{x}\|+\|\mathbf{y}\|$.

则称$\|\cdot\|$为$\mathbf{R}^n$上的范数.

常用的三个向量范数:

1. $1-$范数: $\|\mathbf{x}\|_1=\sum\limits_{i=1}^{n}|x_i|$;
2. $\infty-$范数: $\|\mathbf{x}\|_\infty=\max\limits_{1\leq i\leq n}|x_i|$;
3. $2-$范数: $\|\mathbf{x}\|_2=\sqrt{\sum\limits_{i=1}^{n}x_i^2}=\sqrt{\mathbf{x}^T\mathbf{x}}=\sqrt{(\mathbf{x},\mathbf{x})}$.

由范数定义的三角不等式易得

$$\big|\|\mathbf{x}\| - \|\mathbf{y}\|\big| \leq \|\mathbf{x} - \mathbf{y}\|.$$

定理3.3

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

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

定义3.2

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$\forall \mathbf{x} \in \mathbf{R}^n$，有

$$c_1\|\mathbf{x}\|_p \leq \|\mathbf{x}\|_q \leq c_2\|\mathbf{x}\|_p,$$

则称范数$\|\cdot\|_p$和$\|\cdot\|_q$等价．

定理3.4

$\mathbf{R}^n$上任意两个范数都等价.

定义3.3

设$\|\cdot\|$是$\mathbf{R}^n$上的范数，$\mathbf{x},\mathbf{y}\in\mathbf{R}^n$，称$\|\mathbf{x}-\mathbf{y}\|$为$\mathbf{x}$和$\mathbf{y}$之间的距离. 利用距离可以研究向量的绝对误差和相对误差. 设$\mathbf{x}^*$,是精确值(向量)，$\mathbf{x}$是$\mathbf{x}^*$的近似值(向量) 则称$\|\mathbf{x}^*-\mathbf{x}\|$和$\dfrac{\|\mathbf{x}^*-\mathbf{x}\|}{\|\mathbf{x}^*\|}$或$\dfrac{\|\mathbf{x}^*-\mathbf{x}\|}{\|\mathbf{x}\|}$ 表示近似解$\mathbf{x}$ 的绝对误差和相对误差.

定义3.4

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

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

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

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

★矩阵范数 设

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

$\|\cdot\|$为$\mathbf{R}^n$上的一个范数.

定义3.5

称

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

为矩阵$\mathbf{A}$的范数，记为$\|\mathbf{A}\|$. 即

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

最大值$\max\limits_{\substack{\mathbf{x}\in\mathbf{R}^n\\ \mathbf{x}\neq 0}}\dfrac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}$能否取到？由定理3.3知，对$\forall\mathbf{x}\in\mathbf{R}^n$，$\|\mathbf{A}\mathbf{x}\|$在$\mathbf{R}^n$上连续．记

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

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

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

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

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

矩阵范数的性质： 设$\|\cdot\|$是一个矩阵范数，则有

1. $\forall \mathbf{A} \in \mathbf{R}^{n\times n}$, $\|\mathbf{A}\| \geq 0$, 且$\|\mathbf{A}\| = 0 \Leftrightarrow \mathbf{A} = \mathbf{0}$.
2. $\forall \mathbf{A} \in \mathbf{R}^{n\times n}, \lambda \in \mathbf{R}$, 有$\|\lambda\mathbf{A}\| = |\lambda|\|\mathbf{A}\|$.
3. $\forall \mathbf{A}, \mathbf{B} \in \mathbf{R}^{n\times n}$, 有$\|\mathbf{A}+\mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$.
4. $\forall \mathbf{A}, \mathbf{B} \in \mathbf{R}^{n\times n}$, 有$\|\mathbf{AB}\| \leq \|\mathbf{A}\|\|\mathbf{B}\|$.
5. $\forall \mathbf{A} \in \mathbf{R}^{n\times n}, \mathbf{x} \in \mathbf{R}^{n}$, 有$\|\mathbf{Ax}\| \leq \|\mathbf{A}\|\|\mathbf{x}\|$.

定义3.6

设$\mathbf{B} \in \mathbf{R}^{n\times n}$, $\lambda_1, \lambda_2, \cdots, \lambda_n$为$\mathbf{B}$的$n$个特征值. 称

$$\rho(\mathbf{B}) = \max_{1\leq i\leq n}\{|\lambda_i|\}$$

为矩阵$\mathbf{B}$的谱半径.

定理3.5

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

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

2.
$$\|\mathbf{A}\|_\infty = \max_{\substack{\mathbf{x}\in\mathbf{R}^n\\ \mathbf{x}\neq\mathbf{0}}} \frac{\|\mathbf{A}\mathbf{x}\|_\infty}{\|\mathbf{x}\|_\infty} = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|.$$

3.
$$\|\mathbf{A}\|_2 = \max_{\substack{\mathbf{x}\in\mathbf{R}^n\\ \mathbf{x}\neq\mathbf{0}}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sqrt{\rho(\mathbf{A}^T\mathbf{A})}.$$

定理3.6

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

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

定理3.7

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

定义3.7

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$\forall\mathbf{A}\in\mathbf{R}^{n\times n}$有

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

则称$\|\cdot\|_p$和$\|\cdot\|_q$等价.

定理3.8

$\mathbf{R}^{n\times n}$上任意两个矩阵范数都等价.

定义3.8
设$\|\cdot\|$是$\mathbf{R}^{n\times n}$上的矩阵范数, $\mathbf{A},\mathbf{B}\in\mathbf{R}^{n\times n}$, 称$\|\mathbf{A}-\mathbf{B}\|$为$\mathbf{A}$与$\mathbf{B}$之间的距离.

定义3.9
设$\|\cdot\|$是$\mathbf{R}^{n\times n}$上的矩阵范数, $\mathbf{A}^{(0)},\mathbf{A}^{(1)},\cdots,\mathbf{A}^{(k)},\cdots$ 为 $\mathbf{R}^{n\times n}$ 上的一个矩阵序列, $\mathbf{A}\in\mathbf{R}^{n\times n}$. 如果

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

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

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

定理3.9

设$\mathbf{B} \in \mathbf{R}^{n\times n}$, 则矩阵序列$\mathbf{I}, \mathbf{B}, \mathbf{B}^2, \cdots, \mathbf{B}^k, \cdots$收敛于零矩阵(即$\lim\limits_{k\to\infty} \mathbf{B}^k = \mathbf{0}$)的充要条件是$\rho(\mathbf{B}) < 1$.

条件数

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

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

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

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

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

即

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

两边取范数并由矩阵范数性质得

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

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

$$\|\mathbf{b}\|\leq\|\mathbf{A}\|\|\mathbf{x}^*\|.$$

从而有

$$\frac{\|\delta\mathbf{x}^*\|}{\|\mathbf{x}^*\|}\leq\|\mathbf{A}^{-1}\|\|\mathbf{A}\|\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}\|}. \tag{6}$$

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

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

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

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

即

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

两边取范数得

$$\frac{\|\delta\mathbf{x}^*\|}{\|\mathbf{x}^*+\delta\mathbf{x}^*\|}\leq\|\mathbf{A}\|\|\mathbf{A}^{-1}\|\frac{\|\delta\mathbf{A}\|}{\|\mathbf{A}\|}. \tag{7}$$

定义3.10

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

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

常用的条件数

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

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

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

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

定义3.11

对于方程组$\mathbf{Ax} = \mathbf{b}$, 其中$\mathbf{A}$非奇异. 如果$\mathbf{A}$的条件数很大, 则称该方程组为病态（坏条件）方程组; 如果$\mathbf{A}$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3.2中矩阵的条件数$\mathrm{cond}(\mathbf{A})_\infty = 2$, 故例3.2中的方程组是良态的. 而例3.3中矩阵的条件数$\mathrm{cond}(\mathbf{A})_\infty = 22002$, 故例3.3中的方程组是病态的.

在实际中, 由于计算$\mathbf{A}^{-1}$很麻烦, 所以$\mathrm{cond}(\mathbf{A})$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$\mathbf{D}, \mathbf{C}$, 将方程组$\mathbf{Ax} = \mathbf{b}$ 化为等价的方程组$\mathbf{DAC}[\mathbf{C}^{-1}\mathbf{x}] = \mathbf{Db}$, 使$\mathbf{DAC}$ 的条件数比较小.

★<u>方程组近似解可靠性的判别</u> 设$\tilde{\mathbf{x}}$为方程组$\mathbf{A}\mathbf{x}=\mathbf{b}$的近似解，记$\mathbf{r}=\mathbf{b}-\mathbf{A}\tilde{\mathbf{x}}$, 若$\mathbf{r}=\mathbf{0}$，则$\tilde{\mathbf{x}}$为精确解. 一般$\mathbf{r}\neq\mathbf{0}$.

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

定理3.10

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

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

3.3 线性方程组的迭代法

★<u>迭代格式的构造</u>

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

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

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

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}, \quad k = 0, 1, 2 \cdots. \tag{8}$$

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

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

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

(8)称为迭代格式, $\mathbf{B}$为迭代矩阵. 若迭代格式(8)对任意初始向量$\mathbf{x}^{(0)}$产生的迭代序列$\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$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?

★三个常用的迭代格式 将线性方程组$\mathbf{Ax}=\mathbf{b}$写为分量形式

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
$$

假设$a_{ii} \neq 0$ $(i = 1, 2 \cdots, n)$.
在上述方程组中第i个方程求出x_i得

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

(1) Jacobi迭代格式

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

例子 用Jacobi迭代格式解方程组

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

精确至3位有效数字.

解：Jacobi迭代格式为

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

取初始迭代向量$x^{(0)} = (0, 0, 0)^T$,各次迭代结果如下：

k	0	1	2	3	4	5	6
$x_1^{(k)}$	0.0000	0.1250	0.2500	0.2263	0.2235	0.2251	0.225
$x_2^{(k)}$	0.0000	0.4000	0.3150	0.3005	0.3060	0.3058	0.305
$x_3^{(k)}$	0.0000	-0.6000	-0.4950	-0.4870	-0.4946	-0.4941	-0.493

所以满足精度要求的近似解为$x^* = (0.225, 0.306, -0.494)^T$.

下面将Jacobi迭代用矩阵和向量表示. 记

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

$$\mathbf{D} = \begin{bmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & a_{33} & & \\ & & & \ddots & \\ & & & & a_{nn} \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ & 0 & a_{23} & \cdots & a_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & a_{n,n-1} \\ & & & & 0 \end{bmatrix}.$$

则

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

Jacobi迭代：

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

其中

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

(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.

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

例子 用Gauss-Seidel迭代格式解方程组

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

精确至3位有效数字.

解

Gauss-seidel迭代格式为

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

取初始迭代向量$x^{(0)} = (0, 0, 0)^T$,各次迭代结果如下:

k	0	1	2	3	4
$x_1^{(k)}$	0.0000	0.1250	0.2344	0.2245	0.2250
$x_2^{(k)}$	0.0000	0.3750	0.3031	0.3059	0.3056
$x_3^{(k)}$	0.0000	-0.4925	-0.4925	-0.4939	-0.4936

所以满足精度要求的近似解为$x^* = (0.225, 0.306, -0.494)^T$.

用矩阵表示为:

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

Gauss-Seidel迭代:

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

其中

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

★SOR（超松弛）迭代格式 已知$\mathbf{x}^{(k)} = (\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \cdots, \mathbf{x}_n^{(k)})^T$，将Gauss-Seidel迭代得到的$\mathbf{x}^{(k+1)}$ 与$\mathbf{x}^{(k)}$ 加权平均，得到SOR迭代：

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

其中ω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:

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

其中

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

★迭代格式的收敛性

迭代法基本定理

定理3.11

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

证明:

例3.4

给定线性方程组

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

(a) 分别写出求解该方程组的*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:

$$\begin{aligned} x_1^{(k+1)} &= -\frac{1}{2}x_2^{(k)} - \frac{1}{2}x_3^{(k)}, \\ x_2^{(k+1)} &= -\frac{1}{2}x_1^{(k)} - \frac{1}{2}x_3^{(k)} + 1, \quad k = 0, 1, \cdots . \\ x_3^{(k+1)} &= -\frac{1}{2}x_1^{(k)} - \frac{1}{2}x_2^{(k)} + 1. \end{aligned}$$

Gauss-Seidel迭代格式为:

$$\begin{aligned} x_1^{(k+1)} &= -\frac{1}{2}x_2^{(k)} - \frac{1}{2}x_3^{(k)}, \\ x_2^{(k+1)} &= -\frac{1}{2}x_1^{(k+1)} - \frac{1}{2}x_3^{(k)} + 1, \quad k = 0, 1, \cdots . \\ x_3^{(k+1)} &= -\frac{1}{2}x_1^{(k+1)} - \frac{1}{2}x_2^{(k+1)} + 1. \end{aligned}$$

(b) Jacobi迭代矩阵**J**的特征方程是

$$\begin{vmatrix} \lambda & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \lambda \end{vmatrix} = 0. \Longrightarrow 4\lambda^3 - 3\lambda + 1 = 0.$$

$$\Longrightarrow (2\lambda - 1)(2\lambda^2 + \lambda - 1) = 0, \Longrightarrow \rho(\mathbf{J}) = 1.$$

故Jacobi迭代发散.

Gauss-Seidel迭代矩阵**G**的特征方程是

$$\begin{vmatrix} \lambda & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}\lambda & \lambda & \frac{1}{2} \\ \frac{1}{2}\lambda & \frac{1}{2}\lambda & \lambda \end{vmatrix} = 0. \Longrightarrow \lambda(8\lambda^2 - 2\lambda + 1) = 0.$$

$$\Longrightarrow \lambda_1 = 0, \lambda_{2,3} = \frac{1 \pm \sqrt{7}i}{8}, \quad \rho(\mathbf{G}) = \frac{1}{\sqrt{8}} < 1.$$

所以Gauss-Seidel迭代收敛.

例3.5

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

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

其中$\omega\neq 0$为常数.

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

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

定义3.12

设

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

如果

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

则称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 设A是严格对角占优的，则$|A| \neq 0$.

证明: 这里仅证A是按行严格对角占优的情况.用反证法.

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

设$\|x^*\|_\infty = |x_k^*| \neq 0$. 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$|A| \neq 0$.

定理3.12

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

证明: 记

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

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

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

即$B(\lambda)$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，$|B(\lambda)|\neq 0$. 换句话说，方程$|B(\lambda)|=0$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$\rho(J)<1$. 同样可证当A按列严格对角占优的情况。因而Jacobi迭代格式收敛。

定理3.13

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

证明: 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。

3) SOR迭代的收敛性

SOR迭代的迭代矩阵为

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

由定理3.11, SOR迭代收敛$\Longleftrightarrow \rho(S_\omega) < 1$.

定理3.14

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3.11, 若SOR收敛, 则$\rho(S_\omega) < 1$, 从而有$|\det(S_\omega)| < 1$. 而行列式

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

从而有

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

定理3.15

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

注1 如果$\mathbf{A}$对称正定, 则Gauss-Seidel迭代收敛.

注2 SSOR:

$$\begin{aligned}
\mathbf{x}^{(k+1/2)} &= (1-w)\mathbf{x}^{(k)}+\omega\mathbf{D}^{-1}(\mathbf{b}-\mathbf{L}\mathbf{x}^{(k+1/2)}-\mathbf{L}^T\mathbf{x}^{(k)}),\\
\mathbf{x}^{(k+1)} &= (1-w)\mathbf{x}^{(k+1/2)}+\omega\mathbf{D}^{-1}(\mathbf{b}-\mathbf{L}\mathbf{x}^{(k+1/2)}-\mathbf{L}^T\mathbf{x}^{(k+1)})
\end{aligned}$$

3.4 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.

★<u>求主特征值的幂法</u> 设$\mathbf{A} = (a_{ij})_{n\times n}$ 是n 阶方阵, 它有n 个线性无关的特征向量$\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$, 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$\mathbf{v}_0$, 构造迭代

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

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

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

且$a_1 \neq 0$. 因此有

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

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

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

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

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

$$\mathbf{v}_k \approx a_1 \lambda_1^k \mathbf{x}_1, \quad \mathbf{v}_{k+1} \approx a_1 \lambda_1^{k+1} \mathbf{x}_1 \approx \lambda_1 \mathbf{v}_k. \tag{9}$$

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

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

实际计算时，为了避免$\mathbf{v}_k$的分量产生溢出，可以采用“归一化”，具体将算法改为：

$$\left\{\begin{aligned} \mathbf{u}_0 &= \mathbf{v}_0 \\ \mathbf{v}_k &= \mathbf{A}\mathbf{u}_{k-1} \\ m_k &= \max(\mathbf{v}_k) \\ \mathbf{u}_k &= \mathbf{v}_k / m_k \end{aligned}\right. \qquad k = 1, 2, \cdots. \tag{10}$$

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

定理3.16

设$|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$，则由算法(10)产生的序列$\{\mathbf{u}_k\}$和$\{m_k\}$均收敛，且

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

证

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

利用上式递推得

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

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

$$
\begin{aligned}
\mathbf{u}_k = & \quad \frac{\mathbf{A}^k\mathbf{u}_0}{\max(\mathbf{A}^k\mathbf{u}_0)} \\
= & \quad \frac{\lambda_1^k\left(a_1\mathbf{x}_1 + \sum_{i=2}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k\mathbf{x}_i\right)}{\max\left(\lambda_1^k\left(a_1\mathbf{x}_1 + \sum_{i=2}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k\mathbf{x}_i\right)\right)} \\
= & \quad \frac{a_1\mathbf{x}_1 + \sum_{i=2}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k\mathbf{x}_i}{\max\left(a_1\mathbf{x}_1 + \sum_{i=2}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k\mathbf{x}_i\right)}.
\end{aligned}
$$

从而

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

同样

$$
\begin{aligned}
\mathbf{v}_k = \quad & \mathbf{A}\mathbf{u}_{k-1} = \frac{\mathbf{A}^k \mathbf{u}_0}{\max(\mathbf{A}^{k-1}\mathbf{u}_0)} \\
= \quad & \frac{\lambda_1 \left(a_1 \mathbf{x}_1 + \sum_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i \right)}{\max \left(a_1 \mathbf{x}_1 + \sum_{i=2}^{n} a_i \left(\frac{\lambda_i}{\lambda_1}\right)^{k-1} \mathbf{x}_i \right)}.
\end{aligned}
$$

因此

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

收敛速率：$\left|\dfrac{\lambda_2}{\lambda_1}\right|$

例3.6

用幂法计算矩阵

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

的主特征值及对应的特征向量。

解： 取$\mathbf{u}_0 = (0, 0, 1)^T$，由幂法公式$v_1 = Au_0 = (2, 4, 1)^T$，可知

$$m_1 = 4, u_1 = (0.5, 1.0, 0.25)^T.$$

依次迭代,得如下结果

k	$\mathbf{u}_k^T$(归一化向量)	$m_k = \max(\mathbf{v}_k)$
0	(0.0000, 0.0000, 1.0000)	1.0000
1	(0.5000, 1.0000, 0.2500)	4.0000
2	(0.5000, 1.0000, 0.8611)	9.0000
3	(0.5000, 1.0000, 0.7306)	11.4400
4	(0.5000, 1.0000, 0.7535)	10.9224
5	(0.5000, 1.0000, 0.7493)	11.0140
6	(0.5000, 1.0000, 0.7501)	10.9927
7	(0.5000, 1.0000, 0.7500)	11.0004
8	(0.5000, 1.0000, 0.7500)	11.0000

Table: 幂法求矩阵A的主特征值

矩阵A的三个特征值为$11, -3, -2$.

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

(a) $\lambda_1 = \lambda_2$.

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

因此

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

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

收敛率: $\left|\frac{\lambda_3}{\lambda_1}\right|$.

(b) $\lambda_1 = -\lambda_2$

$$\mathbf{u}_k = \frac{\mathbf{A}^k\mathbf{u}_0}{\max(\mathbf{A}^k\mathbf{u}_0)} = \frac{a_1\mathbf{x}_1 + (-1)^k a_2\mathbf{x}_2 + \sum_{i=3}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i}{\max\left(a_1\mathbf{x}_1 + (-1)^k a_2\mathbf{x}_2 + \sum_{i=3}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i\right)}$$

其中$|a_1| + |a_2| \neq 0$，当$a_2 \neq 0$时$\mathbf{u}_k$不收敛。

$$\mathbf{A}^2\mathbf{u}_k = \lambda_1^2 \frac{a_1\mathbf{x}_1 + (-1)^k a_2\mathbf{x}_2 + \sum_{i=3}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k+2} \mathbf{x}_i}{\max\left(a_1\mathbf{x}_1 + (-1)^k a_2\mathbf{x}_2 + \sum_{i=3}^{n} a_i\left(\frac{\lambda_i}{\lambda_1}\right)^k \mathbf{x}_i\right)}$$

$$\lim_{k\to\infty}\max(\mathbf{A}^2\mathbf{u}_k)=\lim_{k\to\infty}\lambda_1^2\frac{\max\left(a_1\mathbf{x}_1+(-1)^k a_2\mathbf{x}_2+\sum\limits_{i=3}^n a_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k+2}\mathbf{x}_i\right)}{\max\left(a_1\mathbf{x}_1+(-1)^k a_2\mathbf{x}_2+\sum\limits_{i=3}^n a_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k}\mathbf{x}_i\right)}$$

收敛率：$\left|\dfrac{\lambda_3}{\lambda_1}\right|$.

$$\begin{aligned}
\mathbf{A}\mathbf{u}_k+\lambda_1\mathbf{u}_k &= \frac{2\lambda_1 a_1\mathbf{x}_1+\sum\limits_{i=3}^n a_i(\lambda_i+\lambda_1)\left(\frac{\lambda_i}{\lambda_1}\right)^k\mathbf{x}_i}{\max\left(a_1\mathbf{x}_1+(-1)^k a_2\mathbf{x}_2+\sum\limits_{i=3}^n a_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k}\mathbf{x}_i\right)}, \\
\mathbf{A}\mathbf{u}_k-\lambda_1\mathbf{u}_k &= \frac{(-1)^{k+1}2\lambda_1 a_2\mathbf{x}_2+\sum\limits_{i=3}^n a_i(\lambda_i-\lambda_1)\left(\frac{\lambda_i}{\lambda_1}\right)^k\mathbf{x}_i}{\max\left(a_1\mathbf{x}_1+(-1)^k a_2\mathbf{x}_2+\sum\limits_{i=3}^n a_i\left(\frac{\lambda_i}{\lambda_1}\right)^{k}\mathbf{x}_i\right)}
\end{aligned}$$

可以分别作为$\lambda_1,-\lambda_1$所对应的特征向量。

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

$$\mathbf{v}_k = \lambda_1^k \left[a_1 \mathbf{x}_1 + a_2 \left(\frac{\bar{\lambda}_1}{\lambda_1} \right)^k \bar{\mathbf{x}}_1 + \sum_{i=3}^{n} a_i \left(\frac{\lambda_i}{\lambda_1} \right)^k \mathbf{x}_i \right],$$

因为$|\frac{\lambda_i}{\lambda_1}| < 1$，当$k \to \infty$有

$$\begin{aligned} \mathbf{v}_k &\approx a_1 \lambda_1^k \mathbf{x}_1 + a_2 \bar{\lambda}_1^k \bar{\mathbf{x}}_1, \\ \mathbf{v}_{k+1} &\approx a_1 \lambda_1^{k+1} \mathbf{x}_1 + a_2 \bar{\lambda}_1^{k+1} \bar{\mathbf{x}}_1, \\ \mathbf{v}_{k+2} &\approx a_1 \lambda_1^{k+2} \mathbf{x}_1 + a_2 \bar{\lambda}_1^{k+2} \bar{\mathbf{x}}_1, \end{aligned}$$

分别乘以$q = \lambda_1 \bar{\lambda}_1, p = -(\lambda_1 + \bar{\lambda}_1), 1,$相加

$$\mathbf{v}_{k+2} + p\mathbf{v}_{k+1} + q\mathbf{v}_k \approx 0$$

特征值，特征向量

$$\lambda_1 = -\frac{p}{2} + \sqrt{\frac{p^2}{4} - q},$$
$$\bar{\lambda}_1 = -\frac{p}{2} - \sqrt{\frac{p^2}{4} - q}.$$

对应的特征向量

$$\begin{aligned} \mathbf{v}_{k+1} - \bar{\lambda}_1 \mathbf{v}_k &= \lambda_1^k (\lambda_1 - \bar{\lambda}_1) a_1 \mathbf{x}_1, \\ \mathbf{v}_{k+1} - \lambda_1 \mathbf{v}_k &= \bar{\lambda}_1^k (\bar{\lambda}_1 - \lambda_1) a_1 \bar{\mathbf{x}}_1 \end{aligned}$$

幂法优缺点

1. 方法简单，对大型稀疏矩阵比较合适；
2. 收敛情况复杂，需预先进行分析；
3. 初值($a_1 \neq 0, |a_1| + |a_2| \neq 0$)选取不好，会影响收敛速度。

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

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

易知

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