第三章 线性代数方程组的解法

§1 引言

§2 Gauss消去法

§3 矩阵三角分解法

§4 解线性方程组的迭代法

§1 引言

设线性方程组为
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

或写为矩阵形式 $\mathbf{Ax} = \mathbf{b}$，其中$\mathbf{A}$为非奇异矩阵，且

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

注：$\Delta_k = \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}$ 称为$\mathbf{A}$的k阶顺序主子式。

一 研究数值解法的必要性

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$ 的解 $x_k = \dfrac{D_k}{D}, k = 1,2,\ldots,n$

计算一个n阶行列式需要做$(n-1)n!$个乘法，求解上述方程共做$N=(n+1)(n-1)n!+n$次乘除法。

$n=20$，$N \approx 9.7 \times 10^{20}$，工作24小时的计算机上：

每秒上亿次（10^8）浮点乘除法运算的计算机大致需要30万年。

每秒千亿次（10^{11}）浮点乘除法运算的计算机大致需要300年。

求解线性方程组的Gram法则理论上非常完美，但其计算工作量大的惊人，失去实用价值。

二 直接法与迭代法

1 直接法： 只包含有限次四则运算，若在计算过程中不发生舍入误差的假定下，计算结果就是原方程组的精确解。

2 迭代法： 将方程组的解看作是某极限过程的极限值，且计算这一极限值的每一步是利用前一步所得结果施行相同的演算步骤而进行。

§2 Gauss消去法

一 Gauss顺序消去法

将方程组的增广矩阵记为

$$
(\mathbf{A}|\mathbf{b})=\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & a_{1,n+1} \\ a_{21} & a_{22} & \cdots & a_{2n} & a_{2,n+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & a_{n,n+1} \end{array}\right)=\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & a_{2,n+1}^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & a_{n,n+1}^{(1)} \end{array}\right)
$$

基本思想：对增广矩阵进行一系列初等行变换，以使**A**对角线以下的元素化为零，从而使原方程组等价转化为容易求解的上三角方程组。

具体方法

第一步： 设$a_{11}^{(1)} \neq 0$， 令 $l_{i1} = a_{i1}^{(1)} \big/ a_{11}^{(1)} \quad (i = 2,3,\ldots,n)$ 。

在增广矩阵中，用 $-l_{i1}$ 乘以第一行再加到第i行 (i=2,3,…,n)，则第i行的第j个元素化为

$$a_{ij}^{(1)} - \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \cdot a_{1j}^{(1)} \ \underline{\underline{\text{记为}}}\ a_{ij}^{(2)} \quad j = 2,3,\cdots,n+1$$

增广矩阵相应地化为

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & \cdots & \cdots & \cdots & \cdots \\ & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{array}\right) \qquad \left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & a_{2,n+1}^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & a_{n,n+1}^{(1)} \end{array}\right)$$

第二步: 设 $a_{22}^{(2)} \neq 0$， 令 $l_{i2} = a_{i2}^{(2)} \big/ a_{22}^{(2)}$ $(i=3,4,\ldots,n)$。

在增广矩阵中，用 $-l_{i2}$ 乘以第二行再加到第i行 (i=3,4,…,n)，则第i行的第j个元素化为

$$a_{ij}^{(2)} - \frac{a_{i2}^{(2)}}{a_{22}^{(1)}} \cdot a_{2j}^{(2)} \underline{\underline{\text{记为}}}\ a_{ij}^{(3)} \qquad j=3,4,\cdots,n+1$$

增广矩阵相应地化为

$$\left(\begin{array}{ccccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\
 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\
 & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\
 & & \vdots & \cdots & \cdots & \cdots \\
 & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & a_{n,n+1}^{(3)}
\end{array}\right)$$

$$\left(\begin{array}{cccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\
 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\
 & \cdots & \cdots & \cdots & \cdots \\
 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)}
\end{array}\right)$$

完成第3 ～ n–1步后，增广矩阵化为

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & & \ddots & \cdots & \cdots \\ & & & a_{nn}^{(n)} & a_{n,n+1}^{(n)} \end{array}\right)$$

即原方程组被等价转化成为上三角方程组

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots\cdots\cdots\cdots + a_{1n}^{(1)}x_n = a_{1,n+1}^{(1)} \\ \qquad a_{22}^{(2)}x_2 + \cdots\cdots\cdots\cdots + a_{2n}^{(2)}x_n = a_{2,n+1}^{(2)} \\ \qquad\qquad a_{33}^{(3)}x_3 + \cdots + a_{3n}^{(3)}x_n = a_{3,n+1}^{(3)} \\ \qquad\qquad \cdots\cdots\cdots \\ \qquad\qquad\qquad a_{nn}^{(n)}x_n = a_{n,n+1}^{(n)} \end{cases}$$

设$a_{kk}^{(k)} \neq 0, k = n, n-1, \cdots, 1$，逐步回代得原方程组的解

$$\begin{cases} x_n = \dfrac{a_{n,n+1}^{(n)}}{a_{nn}^{(n)}} \\ x_k = \dfrac{1}{a_{kk}^{(k)}}\left[a_{k,n+1}^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)} x_j\right], \quad k = n-1, n-2, \cdots, 1 \end{cases}$$

Remarks

① 如果**A**为n (n较大) 阶矩阵，Gauss顺序消去法的乘除法与加减法次数分别为

$$N_{MD} = n^3/3 + n^2 - n/3 \approx n^3/3 \quad N_{AS} = n(n-1)(2n+5)/6 \approx n^3/3$$

而Gram法则需要$(n+1)(n-1)n!+n$次乘除法。

当n=20，用每秒完成千亿次浮点乘除法运算的计算机求解，耗时不足1秒。

② **定理**：Gauss顺序消去法能够求解**Ax=b**的解之充要条件为**A**的各阶顺序主子式均不为零。

例（第一章）：在四位计算机上求解

$$\begin{cases} 0.00001x_1 + 2x_2 = 1 \\ 2x_1 + 3x_2 = 2 \end{cases}$$

精确解 $x_1 = 0.250001875\cdots \quad x_2 = 0.499998749\cdots$

直接消元

$$\begin{cases} 10^{-4} \times 0.1000x_1 + 10^1 \times 0.2000x_2 = 10^1 \times 0.1000 \\ \qquad\qquad -10^6 \times 0.4000x_2 = -10^6 \times 0.2000 \end{cases}$$

近似解为 $x_2^* = 0.5;\ \ x_1^* = 0$

③ 当 $\left|a_{kk}^{(k)}\right| \approx 0$或$\left|a_{kk}^{(k)}\right|$相对于$\left|a_{ik}^{(k)}\right|$ $(i=k+1,k+2,\ldots,n)$比较小时，计算产生的舍入误差可能导致计算结果误差较大。

二 Gauss主元消去法

如果计算中舍入误差增大迅速，造成计算解与真解相差甚远，则相应使用的计算方法不稳定法。反之，计算中的舍入误差增大能得到控制，该方法就是稳定的。

$a_{kk}^{(k)}\ (a_{kk}^{(k)} \neq 0), k = 1,2,\cdots,n$ 称为第k步的**主元（素）**。

$\left|a_{kk}^{(k)}\right|$很小是Gauss顺序消去法不稳定的根源，因此需要采用“**选主元素**”技术，即**选取绝对值最大的元素作为主元素**。

1. Gauss列主元消去法

第一步　在A的第一列中选取绝对值最大的元素作为主元素，如：

$$\left|a_{i_1 1}^{(1)}\right| = \max_{1\le i\le n}\left|a_{i1}^{(1)}\right| \quad (a_{i_1,1}^{(1)}\text{称之为第1列的主元素})$$

如果$i_1 \neq 1$，则调换第一行与第i_1行，这时的$a_{11}^{(1)}$就是原来的$a_{i_1,1}^{(1)}$。

用新的$a_{11}^{(1)}$将第一列后$n-1$个元素$a_{i1}^{(1)}\ (i=2,3,\cdots,n)$按Gauss顺序消元过程化零，得增广矩阵

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & \cdots & \cdots & \cdots & \cdots \\ & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{array}\right)$$

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & a_{2,n+1}^{(1)} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & a_{n,n+1}^{(1)} \end{array}\right)$$

第二步 在上述增广矩阵第二列后$n-1$个元素中选主元$a_{i_2,2}^{(2)}$，即

$$\left|a_{i_2,2}^{(2)}\right| = \max_{2\le i\le n}\left|a_{i2}^{(2)}\right|$$

如果$i_2 \neq 2$，则调换第二行与第i_2行，这时的$a_{22}^{(2)}$就是原来的$a_{i_2,2}^{(2)}$。

用新的$a_{22}^{(2)}$将第二列后$n-2$个元素$a_{i2}^{(2)}\ (i=3,4,\cdots,n)$按Gauss顺序消元过程化零，得增广矩阵

$$\left(\begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & a_{3,n+1}^{(3)} \\ & & \vdots & \cdots & \cdots & \cdots \\ & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & a_{n,n+1}^{(3)} \end{array}\right)$$

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & \cdots & \cdots & \cdots & \cdots \\ & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{array}\right)$$

以后每一步都类似地在右下角方阵中的第一列中选主元。再经行的对调将主元换到右下角方阵中左上角的位置，并按Gauss顺序消元过程计算$a_{ij}^{(k)}$（$i,j=k,\ldots,n$），……。直至将方程组化成上三角方程组，再进行回代就可得解。

列主元消去法除了每步需要按列选主元并可能作行交换外，其消去过程与Gauss顺序消去法的消去过程完全一样。

2. Gauss全主元消去法

第一步　在A的所有元素中选取绝对值最大的元素作为主元素，如：

$$\left|a_{i_1,j_1}^{(1)}\right| = \max_{\substack{1\le i\le n\\1\le j\le n}}\left|a_{ij}^{(1)}\right| \quad (a_{i_1,j_1}^{(1)}\text{称之为第一步的主元素})$$

如果 $i_1 \ne 1$ 且 $j_1 \ne 1$，则调换第一行与第 i_1 行，第一列与第 j_1 列。这时的 $a_{11}^{(1)}$ 就是原来的 $a_{i_1,j_1}^{(1)}$ 。

用新的 $a_{11}^{(1)}$ 将第一列后 $n-1$ 个元素 $a_{i1}^{(1)}\ (i=2,3,\cdots,n)$ 按Gauss顺序消元过程化零，得增广矩阵

$$\left(\begin{array}{cccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\
 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\
 & \cdots & \cdots & \cdots & \cdots \\
 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)}
\end{array}\right)
\qquad
\left(\begin{array}{cccc|c}
a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\
a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & a_{2,n+1}^{(1)} \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & a_{n,n+1}^{(1)}
\end{array}\right)$$

第二步 在上述增广矩阵中右下角的$n-1$阶矩阵中选取绝对值最大的元素作为主元素，再经行的对换和列的对换把主元素移到$a_{22}^{(2)}$的位置。

用新的$a_{22}^{(2)}$将第二列后$n-2$个元素$a_{i2}^{(2)}\ (i=3,4,\cdots,n)$按Gauss顺序消元过程化零，得增广矩阵

$$\left(\begin{array}{cccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} & & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} & & a_{2,n+1}^{(2)} \\ & & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} & & a_{3,n+1}^{(3)} \\ & & \vdots & \cdots & \cdots & & \cdots \\ & & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} & & a_{n,n+1}^{(3)} \end{array}\right)$$

$$\left(\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} & a_{1,n+1}^{(1)} \\ & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} & a_{2,n+1}^{(2)} \\ & \cdots & \cdots & \cdots & \cdots \\ & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} & a_{n,n+1}^{(2)} \end{array}\right)$$

以后每一步都类似地在右下角方阵所有元素中选取主元。再经行、列对调将主元换到右下角方阵中左上角的位置，并按Gauss顺序消元过程计算 $a_{ij}^{(k)}$（$i,j=k,\ldots,n$），……。直至将方程组化成上三角方程组，再进行回代就可得解。

Remarks

① 完全主元素消去法在选主元素时要花费较多的计算时间，且未知数交换的顺序需要存储与还原。

② 实用中常用列主元消去法。

例(第一章)：在四位计算机上用列主元素消去法求解

$$\begin{cases} 0.00001x_1 + 2x_2 = 1 \\ 2x_1 + 3x_2 = 2 \end{cases}$$

精确解 $x_1 = 0.250001875 \quad x_2 = 0.499998749$

解：按列选主元并消元，有

$$B = \begin{bmatrix} 10^{-4}\times 0.1000 & 10^{1}\times 0.2000 & 10^{1}\times 0.1000 \\ 10^{1}\times 0.2000 & 10^{1}\times 0.3000 & 10^{1}\times 0.2000 \end{bmatrix} \to \begin{bmatrix} 10^{1}\times 0.2000 & 10^{1}\times 0.3000 & 10^{1}\times 0.2000 \\ 10^{-4}\times 0.1000 & 10^{1}\times 0.2000 & 10^{1}\times 0.1000 \end{bmatrix}$$

$$\to \begin{bmatrix} 10^{1}\times 0.2000 & 10^{1}\times 0.3000 & 10^{1}\times 0.2000 \\ & 10^{1}\times 0.2000 - 10^{1}\times 0.3000 \times \dfrac{10^{-4}\times 0.1000}{10^{1}\times 0.2000} & 10^{1}\times 0.1000 - 10^{1}\times 0.2000 \times \dfrac{10^{-4}\times 0.1000}{10^{1}\times 0.2000} \end{bmatrix}$$

$$= \begin{bmatrix} 10^{1}\times 0.2000 & 10^{1}\times 0.3000 & 10^{1}\times 0.2000 \\ & 10^{1}\times 0.2000 - 10^{-4}\times 0.15 & 10^{1}\times 0.1000 - 10^{-4}\times 0.1 \end{bmatrix}$$

$$\xrightarrow{\text{对阶}} \begin{bmatrix} 10^{1}\times 0.2000 & 10^{1}\times 0.3000 & 10^{1}\times 0.2000 \\ & 10^{1}\times 0.2000 & 10^{1}\times 0.1000 \end{bmatrix}$$

近似解 $x_2^* = 0.5; \quad x_1^* = 0.25$

§3 矩阵三角分解法

一 矩阵三角分解的基本概念

1. Gauss顺序消去法的矩阵形式

设 $\mathbf{Ax}=\mathbf{b}$ 中$\mathbf{A}$的各阶顺序主子式均不为零。

Gauss顺序消去法

Step 1 设$a_{11}^{(1)}\neq 0$，令$l_{i1}=a_{i1}^{(1)}/a_{11}^{(1)}\ (i=2,3,\ldots,n)$。对增广矩阵，

用$-l_{i1}$乘以第一行再加到第i行 $(i=2,3,\ldots,n)$。

Step 2 设$a_{22}^{(2)}\neq 0$，令$l_{i2}=a_{i2}^{(2)}/a_{22}^{(2)}\ (i=3,4,\ldots,n)$。对增广矩阵，

用$-l_{i2}$乘以第二行再加到第i行 $(i=3,4,\ldots,n)$。

……

Gauss顺序消去法的矩阵形式　　令

$$
\mathbf{L}_1=\begin{bmatrix} 1 & & & & \\ -l_{21} & 1 & & & \\ -l_{31} & 0 & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ -l_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}
\mathbf{L}_2=\begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & -l_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & -l_{n2} & \cdots & \cdots & 1 \end{bmatrix}
\cdots \mathbf{L}_{n-1}=\begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & -l_{n,n-1} & 1 \end{bmatrix}
$$

则$n-1$步消元过程为　　　得

$$\mathbf{L}_1[\mathbf{A}^{(1)}\vdots\mathbf{b}^{(1)}]=[\mathbf{A}^{(2)}\vdots\mathbf{b}^{(2)}]$$

$$\mathbf{L}_2[\mathbf{A}^{(2)}\vdots\mathbf{b}^{(2)}]=[\mathbf{A}^{(3)}\vdots\mathbf{b}^{(3)}]$$

$$\vdots$$

$$\mathbf{L}_{n-1}[\mathbf{A}^{(n-1)}\vdots\mathbf{b}^{(n-1)}]=[\mathbf{A}^{(n)}\vdots\mathbf{b}^{(n)}]$$

$$\mathbf{L}_{n-1}\mathbf{L}_{n-2}[\mathbf{A}^{(n-2)}\vdots\mathbf{b}^{(n-2)}]=[\mathbf{A}^{(n)}\vdots\mathbf{b}^{(n)}]$$

$$\mathbf{L}_{n-1}\mathbf{L}_{n-2}\mathbf{L}_{n-3}[\mathbf{A}^{(n-3)}\vdots\mathbf{b}^{(n-3)}]=[\mathbf{A}^{(n)}\vdots\mathbf{b}^{(n)}]$$

$$\vdots$$

$$\mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_2\mathbf{L}_1[\mathbf{A}\vdots\mathbf{b}]=[\mathbf{A}^{(n)}\vdots\mathbf{b}^{(n)}]$$

其中 $\mathbf{L}_i\ (i=1,2,\cdots,k)$ 均为单位下三角阵; $\mathbf{A}^{(n)}$ 为上三角阵。

$$l_{i1}=a_{i1}^{(1)}\big/a_{11}^{(1)}\ (i=2,3,\ldots,n)$$

$$l_{i2}=a_{i2}^{(2)}\big/a_{22}^{(2)}\ (i=3,4,\ldots,n)$$

$$l_{n,n-1}=a_{n,n-1}^{(n-1)}\big/a_{n-1,n-1}^{(n-1)}$$

记 $[\mathbf{A}^{(n)} \vdots \mathbf{b}^{(n)}] = [\mathbf{U} \vdots \mathbf{g}]$ 得 $\mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_2\mathbf{L}_1[\mathbf{A} \vdots \mathbf{b}] = [\mathbf{U} \vdots \mathbf{g}]$

其中 $\mathbf{L}_i \ (i = 1,2,\cdots,k)$ 均为单位下三角阵；**U**为上三角阵。

即 $$\begin{cases} \mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_2\mathbf{L}_1\mathbf{A} = \mathbf{U} \\ \mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_2\mathbf{L}_1\mathbf{b} = \mathbf{g} \end{cases}$$

Gauss顺序消去法：用一连串初等变换（单位下三角矩阵左乘A）将系数方阵A化成上三角方阵U，同时把右端向量b化为向量g。

Gauss顺序消去法能够实施时，有

$$\boxed{\mathbf{A}} = (\mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_2\mathbf{L}_1)^{-1}\mathbf{U} = \boxed{\mathbf{LU}}$$

其中 $\mathbf{L} = (\mathbf{L}_{n-1}\mathbf{L}_{\mathbf{k-1}}\cdots\mathbf{L}_{\mathbf{2}}\mathbf{L}_{\mathbf{1}})^{-1}$，**U**为上三角阵。

可以证明：L为单位下三角矩阵。

$$\mathbf{L}_{n-1}\mathbf{L}_{n-2}\cdots\mathbf{L}_2\mathbf{L}_1[\mathbf{A} \vdots \mathbf{b}] = [\mathbf{A}^{(n)} \vdots \mathbf{b}^{(n)}]$$

2. 矩阵的三角分解及其条件

定义1 设**A**为n阶矩阵（$n\geq 2$）， **L**是下三角矩阵，**U**是上三角矩阵，称**A=LU**为**矩阵A的三角分解**。

A=LU的分解**形式并不惟一。**

定义2 如果**L是单位下三角矩阵**，**U**是上三角矩阵，则称三角分解**A=LU**为**Doolittle分解**；如果**L**是下三角矩阵，**U是单位上三角矩阵**，则称三角分解**A=LU**为**Crout分解**。

Gauss顺序消去法对应了**A=LU**的Doolittle分解。

定理 设**A**为n阶可逆矩阵，则**A**有唯一Doolittle (或Crout) 分解的充要条件为：**A**的前$n-1$阶顺序主子式不为零。（证明略）

3. 矩阵三角分解的目的

若A=LU, 则 $\mathbf{LUx}=\mathbf{b} \Leftrightarrow \begin{cases}\mathbf{Ux}=\mathbf{y}\\ \mathbf{Ly}=\mathbf{b}\end{cases}$

① 求解 $\mathbf{Ly}=\mathbf{b}$

$$\begin{pmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \cdots & \cdots & \cdots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \\ \cdots \\ b_n \end{pmatrix} \Rightarrow y_1, y_2, \cdots, y_n$$

$$\sum_{m=1}^{0} = 0$$

$$y_k = b_k - \sum_{m=1}^{k-1} l_{km} y_m, k = 1,2,\cdots,n$$

② 求解 $\mathbf{Ux}=\mathbf{y}$

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \cdots & \cdots \\ & & & u_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}=\begin{pmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{pmatrix} \Rightarrow x_n, x_{n-1}, \cdots, x_1$$

$$\sum_{k=n+1}^{n} = 0$$

$$x_k = (y_k - \sum_{j=k+1}^{n} u_{kj} x_j) \Big/ u_{kk}, i = n, n-1, \cdots, 1$$

二 Doolittle分解(直接三角分解法)

对给定矩阵进行三角分解时，通常由**A=LU**按矩阵乘法直接进行。下面以**Doolittle分解**为例。

设**A**的各阶顺序主子式不为零，则有**A=LU**，即

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & l_{n3} & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix}$$

第一步

比较第一行元素： $a_{1j} = u_{1j} \Rightarrow u_{1j} \ (j = 1,2,\cdots n)$

比较第一列元素： $a_{i1} = l_{i1}u_{11} \Rightarrow l_{i1} = a_{i1} / u_{11} \ (i = 2,3,\cdots n)$

第k步　设**L**的前k−1列，**U**的前k−1行已求出，则

比较第k行元素：

$$a_{kj}=\sum_{m=1}^{n}l_{km}u_{mj}=\sum_{m=1}^{k-1}l_{km}u_{mj}+l_{kk}u_{kj}+\sum_{m=k+1}^{n}\underline{l_{km}}u_{mj}=\sum_{m=1}^{k-1}l_{km}u_{mj}+u_{kj}$$

$$\Rightarrow\quad u_{kj}=a_{kj}-\sum_{m=1}^{k-1}l_{km}u_{mj}\qquad (j=k,k+1\cdots,n)$$

比较第k列元素：

$$a_{ik}=\sum_{m=1}^{n}l_{im}u_{mk}=\sum_{m=1}^{k-1}l_{im}u_{mk}+l_{ik}u_{kk}+\sum_{m=k+1}^{n}l_{im}\underline{u_{mk}}=\sum_{m=1}^{k-1}l_{im}u_{mk}+l_{ik}u_{kk}$$

$$\Rightarrow\quad l_{ik}=\left(a_{ik}-\sum_{m=1}^{k-1}l_{im}u_{mk}\right)\Big/u_{kk}\qquad (i=k+1,\cdots,n)$$

$$\mathbf{L}=\begin{bmatrix}1&&&\\ l_{21}&1&&\\ \vdots&\vdots&\ddots&\\ l_{n1}&l_{n2}&\cdots&1\end{bmatrix}\qquad \mathbf{U}=\begin{bmatrix}u_{11}&u_{12}&\cdots&u_{1n}\\ &u_{22}&\cdots&u_{2n}\\ &&\ddots&\vdots\\ &&&u_{nn}\end{bmatrix}$$

已知**A=LU**，则**Ax=b** 的求解过程为: $\begin{cases}\mathbf{Ly}=\mathbf{b}\\ \mathbf{Ux}=\mathbf{y}\end{cases}$

求解 $\mathbf{Ly}=\mathbf{b}$ 的递归公式

$$y_k = b_k - \sum_{m=1}^{k-1} l_{km} y_m \quad k=1,2,\cdots,n \quad \sum_{m=1}^{0} = 0$$

求解 $\mathbf{Ux}=\mathbf{y}$ 的递归公式

$$x_k = \left(y_k - \sum_{j=k+1}^{n} u_{kj} x_j\right) \Big/ u_{kk} \quad i=n,n-1,\cdots,1 \quad \sum_{k=n+1}^{n} = 0$$

对比计算 u_{kj} 和 y_k 公式

$$u_{kj} = a_{kj} - \sum_{m=1}^{k-1} l_{km} u_{mj} \quad (j=k,k+1\cdots,n)$$

发现二者计算中的规律类似。

$\mathbf{LUx}=\mathbf{b}$

上述公式计算的数据可用**紧凑格式**存储：

$$
\begin{array}{cccccc}
u_{11} & u_{12} & u_{13} & \cdots & u_{1n} & y_1 \\
l_{21} & u_{22} & u_{23} & \cdots & u_{2n} & y_2 \\
l_{31} & l_{32} & u_{33} & \cdots & u_{3n} & y_3 \\
l_{41} & l_{42} & l_{43} & \ddots & & \\
\vdots & \vdots & \vdots & & \ddots & \\
l_{n1} & l_{n2} & l_{n3} & & & \ddots
\end{array}
$$

解 $\mathbf{Ux}=\mathbf{y}$ 即得方程组的解

$$u_{kj}=a_{kj}-\sum_{m=1}^{k-1} l_{km}u_{mj} \quad (j=k,k+1\cdots,n) \qquad y_k=b_k-\sum_{m=1}^{k-1} l_{km}y_m$$

$$k=1,2,\cdots,n$$

$$l_{ik}=(a_{ik}-\sum_{m=1}^{k-1} l_{im}u_{mk})\Big/u_{kk} \quad (i=k+1,\cdots,n)$$

三 平方根法（Cholesky分解）

定义 设**A**为n阶$(n\geq 2)$对称正定矩阵，**L**是非奇异下三角矩阵，称 $\mathbf{A}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$为**A**的**Cholesky分解**。

$$\mathbf{L}=\begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix}$$

定理 n阶$(n\geq 2)$对称正定矩阵**A**定有如下分解式

$$\mathbf{A}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$$

当限定**L**的对角元为正时，$\mathbf{A}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$（**Cholesky分解**）的分解式唯一。

称**A=LU**为**矩阵A的三角分解**。

平方根法的计算公式

设 $\mathbf{A}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$，其中$\mathbf{L}$为下三角矩阵。

令

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

按列比较两端矩阵中主对角线元素与下三角部分元素

第一步

$$a_{11}=l_{11}^2 \quad \Rightarrow \quad l_{11}=\sqrt{a_{11}}$$

$$a_{i1}=l_{i1}l_{11} \quad \Rightarrow \quad l_{i1}={a_{i1}}/{l_{11}} \qquad (i \geq 2,3,\cdots,n)$$

第k步

设**L**的前k-1列元素已求出，比较第k列元素，有

$$a_{ik}=\sum_{m=1}^{n}l_{im}l_{mk}^{T}=\sum_{m=1}^{n}l_{im}l_{km}=\sum_{m=1}^{k-1}l_{im}l_{km}+l_{ik}l_{kk}+\underline{\sum_{m=k+1}^{n}l_{im}l_{km}}$$

当$i=k$，有 $a_{kk}=\sum_{m=1}^{k-1}l_{km}^{2}+l_{kk}^{2}\Rightarrow l_{kk}=\sqrt{a_{kk}-\sum_{m=1}^{k-1}l_{km}^{2}}$

当$i>k$，有 $a_{ik}=\sum_{m=1}^{k-1}l_{im}l_{km}+l_{ik}l_{kk}\Rightarrow l_{ik}=(a_{ik}-\sum_{m=1}^{k-1}l_{im}l_{km})\Big/l_{kk}$

$$i=k+1,k+2,\cdots,n$$

上述公式是对称正定矩阵**A**的**Choledsky分解法**（**平方根法**）的计算公式。

$$\mathbf{A}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$$

求解： 若$\mathbf{A}=\mathbf{L}\mathbf{L}^{\mathrm{T}}$，则 $\mathbf{L}\mathbf{L}^{\mathrm{T}}\mathbf{x}=\mathbf{b} \Leftrightarrow \begin{cases} \mathbf{L}^{\mathrm{T}}\mathbf{x}=\mathbf{y} \\ \mathbf{L}\mathbf{y}=\mathbf{b} \end{cases}$

当n比较大时，平方根法求解方程组的计算量约为

$$\frac{1}{6}\left(n^3+9n^2+cn\right)\approx\frac{1}{6}n^3 \quad \text{乘除法}$$

但直接三角分解法求解方程组的计算量约为

$$\frac{1}{3}n^3+n^2-\frac{1}{3}n\approx\frac{1}{3}n^3 \quad \text{乘除法}$$

用平方根法求解对称正定方程组不需选主元。

例 用平方根法求解如下方程组 $\begin{pmatrix} 3 & 3 & 5 \\ 3 & 5 & 9 \\ 5 & 9 & 17 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 16 \\ 30 \end{pmatrix}$

解 可验证系数矩阵**A**对称正定（各阶顺序主子式为正）。

设 $\begin{pmatrix} 3 & 3 & 5 \\ 3 & 5 & 9 \\ 5 & 9 & 17 \end{pmatrix} = \begin{pmatrix} l_{11} & & \\ l_{21} & l_{22} & \\ l_{31} & l_{32} & l_{33} \end{pmatrix}\begin{pmatrix} l_{11} & l_{21} & l_{31} \\ & l_{22} & l_{32} \\ & & l_{33} \end{pmatrix}$ 按列比较矩阵对应元素，得

$l_{11}^2 = 3 \Rightarrow l_{11} = \sqrt{3}$ $\quad l_{21}l_{11} = 3 \Rightarrow l_{21} = \sqrt{3}$ $\quad l_{31}l_{11} = 5 \Rightarrow l_{31} = \dfrac{5}{\sqrt{3}}$

$l_{21}^2 + l_{22}^2 = 5 \Rightarrow l_{22} = \sqrt{2}$ $\quad l_{31}l_{21} + l_{32}l_{22} = 9 \Rightarrow l_{32} = 2\sqrt{2}$ $\quad l_{31}^2 + l_{32}^2 + l_{33}^2 = 17 \Rightarrow l_{33} = \sqrt{\dfrac{2}{3}}$

解下三角方程组 $\begin{pmatrix} \sqrt{3} & & \\ \sqrt{3} & \sqrt{2} & \\ \dfrac{5}{\sqrt{3}} & 2\sqrt{2} & \sqrt{\dfrac{2}{3}} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 10 \\ 16 \\ 30 \end{pmatrix}$ 得 $y_1 = \dfrac{10}{\sqrt{3}}, y_2 = \dfrac{6}{\sqrt{2}}, y_3 = \dfrac{2\sqrt{2}}{\sqrt{3}}$

解上三角方程组 $\begin{pmatrix} \sqrt{3} & \sqrt{3} & \dfrac{5}{\sqrt{3}} \\ & \sqrt{2} & 2\sqrt{2} \\ & & \sqrt{\dfrac{2}{3}} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \dfrac{10}{\sqrt{3}} \\ \dfrac{6}{\sqrt{2}} \\ 2\sqrt{\dfrac{2}{3}} \end{pmatrix}$ 得 $x_3 = 2, x_2 = -1, x_1 = 1$

四 解三对角方程组的追赶法

1. 矩阵严格对角占优的概念

设 $\mathbf{A}=(a_{ij})_{n\times n}$

若对于i=1,2,…,n，均有 $|a_{ii}|>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}|$，则称矩阵**A 按行严格对角占优**。

按列严格对角占优定义类似。

例 $\mathbf{A}=\begin{pmatrix}4 & -3 & 0\\ -2 & 10 & -8\\ 1 & -6 & 9\end{pmatrix}$ **非按行严格对角占优**

按列严格对角占优。

当**A严格对角占优**，可证A的各阶顺序主子式（含行列式）非零。故**A有惟一的Doolittle（Crout）分解。**

2. 追赶法

设 $\mathbf{Ax}=\mathbf{d}$

A 三对角矩阵

即

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

且 $a_i \neq 0, i=2,\cdots,n,\ \ c_i \neq 0, i=1,\cdots,n-1$

$$|b_1|>|c_1|>0 \quad |b_i|>|a_i|+|c_i| \ \ i=2,\cdots,n-1 \quad |b_n|>|a_n|>0$$

（满足**严格对角占优**条件）

当**A严格对角占优**，可证**A**的各阶顺序主子式非零。故**A有惟一的Doolittle（Crout）分解。**

当**A**为上述三对角阵时，**A**有特殊的三角分解形式。

下面以Doolittle分解为例

$$\begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ l_2 & 1 & & & \\ & l_3 & 1 & & \\ & & \ddots & \ddots & \\ & & & l_n & 1 \end{bmatrix} \begin{bmatrix} u_1 & c_1 & & & \\ & u_2 & c_2 & & \\ & & u_3 & c_3 & \\ & & & \ddots & \ddots \\ & & & & u_n \end{bmatrix}$$

比较两端第i行的第$i+1$列，立刻可得**U**的次上三角线元素为$c_i\ (i=1,2,\cdots,n-1)$。计算次序$u_1, l_2, u_2, l_3, \cdots, l_n, u_n$

比较第1行第1列，得 $b_1 = u_1$

比较第2行第1列，得 $a_2 = l_2 u_1$

按行比较第i行i−1、i列

比较第2行第2列，得 $b_2 = l_2 c_1 + u_2$

比较第i行第i−1列，得 $a_i = l_i u_{i-1}$

比较第i行第i列，得 $b_i = l_i c_{i-1} + u_i$

解Ax=d 等价于 **LUx=d** 令 **y=Ux** 即 $\begin{cases}\mathbf{Ly}=\mathbf{d}\\ \mathbf{Ux}=\mathbf{y}\end{cases}$

由
$$\begin{bmatrix}1 & & & \\ l_2 & 1 & & \\ & \ddots & \ddots & \\ & & l_n & 1\end{bmatrix}\begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix}=\begin{bmatrix}d_1\\ d_2\\ \vdots\\ d_n\end{bmatrix}$$

得
$$\begin{cases}y_1=d_1\\ y_i=d_i-l_iy_{i-1}\ \ i=2,3,\cdots,n\end{cases}$$
追的过程

由
$$\begin{bmatrix}u_1 & c_1 & & \\ & u_2 & \ddots & \\ & & \ddots & c_{n-1}\\ & & & u_n\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}=\begin{bmatrix}y_1\\ x_2\\ \vdots\\ y_n\end{bmatrix}$$

得
$$\begin{cases}x_n=y_n/u_n\\ x_i=(y_i-c_ix_{i+1})/a_i\ \ i=n-1,\cdots,2,1\end{cases}$$
赶的过程

上述计算过程称为解三对角方程组的**追赶法**或**Thomas方法**。

可以证明：只要三对角矩阵按行严格对角占优，则追赶法定能进行下去，且计算过程是稳定的（不必选主元素）。

追赶法的乘除法运算次数为O(n)。

例 用追赶法求解如下方程组
$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 3 & -7 & 3 \\ 0 & 0 & 2 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -10 \\ 2 \end{pmatrix}$$

解 $\mathbf{A}$严格对角占优，设
$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 3 & -7 & 3 \\ 0 & 0 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ l_2 & 1 & & \\ & l_3 & 1 & \\ & & l_4 & 1 \end{pmatrix}\begin{pmatrix} u_1 & 1 & & \\ & u_2 & 0 & \\ & & u_3 & 3 \\ & & & u_4 \end{pmatrix}$$

$\boxed{\text{按行比较}}$矩阵对应元素，得

$$u_1 = 2$$

$$l_2 u_1 = 1 \Rightarrow l_2 = \frac{1}{2} \qquad l_2 \times 1 + u_2 = 2 \Rightarrow u_2 = \frac{3}{2}$$

$$l_3 u_2 = 3 \Rightarrow l_3 = 2 \qquad l_3 \times 0 + u_3 = -7 \Rightarrow u_3 = -7$$

$$l_4 u_3 = 2 \Rightarrow l_4 = -\frac{2}{7} \qquad l_4 \times 3 + u_4 = 5 \Rightarrow u_4 = \frac{41}{7}$$

解下三角方程组
$$\begin{pmatrix} 1 & & & \\ 1/2 & 1 & & \\ & 2 & 1 & \\ & & -\frac{2}{7} & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ -10 \\ 2 \end{pmatrix}$$

得 $\mathbf{y} = (3, -\frac{3}{2}, -7, 0)^T$

解上三角方程组
$$\begin{pmatrix} 2 & 1 & & \\ & \frac{3}{2} & 0 & \\ & & -7 & 3 \\ & & & \frac{41}{7} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 \\ -\frac{3}{2} \\ -7 \\ 0 \end{pmatrix}$$

得 $\mathbf{x} = (2, -1, 1, 0)^T$

§4 解线性方程组的迭代法

一 向量范数、矩阵范数与谱半径

范数：度量向量与矩阵误差“大小”的工具。

1 向量的范数

(1) 定义

若 $\mathbf{x}\in\mathbf{R}^n$ 的某个实值函数 $\|\mathbf{x}\|$ 满足

(a) 正定性 $\|\mathbf{x}\|\ge 0$ 当且仅当 $\mathbf{x}=\mathbf{0}$ 时，有 $\|\mathbf{x}\|=0$

(b) 齐次性 $\|\alpha\mathbf{x}\|=|\alpha|\|\mathbf{x}\|$，$\alpha$ 为任意实数

(c) 三角不等式 $\|\mathbf{x}+\mathbf{y}\|\le\|\mathbf{x}\|+\|\mathbf{y}\|\quad \mathbf{x},\mathbf{y}\in\mathbf{R}^n$

则称 $\|\mathbf{x}\|$ 是 $\mathbf{R}^n$ 上的一个向量范数。

(2) 常用的向量范数

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \cdots + |x_n|$$

$$\|\mathbf{x}\|_2 = (\mathbf{x}, \mathbf{x})^{1/2} = \{\sum_{i=1}^{n} x_i^2\}^{1/2} = \sqrt{x_1^2 + \cdots + x_n^2}$$

$$\|\mathbf{x}\|_\infty = \max_{1 \le i \le n} |x_i|$$

可验证:

它们均满足向量范数的三个条件。

(3) 向量序列收敛

设 $\{\mathbf{x}^{(k)}\}_{k=0}^{+\infty}$ 是 $\mathbf{R}^n$ 中的向量序列，若有向量 $\mathbf{x}^* \in \mathbf{R}^n$，

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

则称 $\{\mathbf{x}^{(k)}\}_{k=0}^{+\infty}$ 收敛于 $\mathbf{x}^*$。记为 $\lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^*$

可以证明向量范数具有等价性，故 $\{\mathbf{x}^{(k)}\}_{k=0}^{+\infty}$ 的收敛性和范数的选择无关。

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}^* \Leftrightarrow \lim_{k\to\infty} \|\mathbf{x}^{(k)} - \mathbf{x}^*\| = \lim_{k\to\infty} \max_{1\le i\le n} |x_i^{(k)} - x_i^*| = 0$$

2 矩阵的范数

(1) 定义

若矩阵 $\mathbf{A}\in\mathbf{R}^{n\times n}$ 的某个非负的实值函数 $\|\mathbf{A}\|$，满足

(a) 正定性 $\|\mathbf{A}\|\geq 0$ 当且仅当A=O时，有 $\|\mathbf{A}\|=0$

(b) 齐次性 $\|\alpha\mathbf{A}\|=|\alpha|\|\mathbf{A}\|$， α为任意实数

(c) 三角不等式 $\|\mathbf{A}+\mathbf{B}\|\leq\|\mathbf{A}\|+\|\mathbf{B}\|\quad \mathbf{A},\mathbf{B}\in\mathbf{R}^{n\times n}$

(d) 自相容性 $\|\mathbf{AB}\|\leq\|\mathbf{A}\|\|\mathbf{B}\|\quad \mathbf{A},\mathbf{B}\in\mathbf{R}^{n\times n}$

则称 $\|\mathbf{A}\|$ 是A的一个矩阵范数。

(2) 常用的矩阵范数

$\|\mathbf{A}\|_\infty=\max\limits_{1\leq i\leq n}\sum\limits_{j=1}^{n}|a_{ij}|=\max\limits_{1\leq i\leq n}(|a_{i1}|+|a_{i2}|+\cdots+|a_{in}|)$ 称为A的行范数

$\|\mathbf{A}\|_1=\max\limits_{1\leq j\leq n}\sum\limits_{i=1}^{n}|a_{ij}|=\max\limits_{1\leq j\leq n}(|a_{1j}|+|a_{2j}|+\cdots+|a_{nj}|)$ 称为A的列范数

$$\|\mathbf{A}\|_2=\sqrt{\lambda_{\max}(\mathbf{A}^T\mathbf{A})}$$ 称为A的2-范数（谱范数）

其中 $\lambda_{\max}(\mathbf{A}^T\mathbf{A})$表示 $\mathbf{A}^T\mathbf{A}$的最大特征值。

$$\|\mathbf{A}\|_F=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}^2}$$ 称为A的Frobenius范数

可验证:它们均满足矩阵范数的四个条件。

例 $\mathbf{A}=\begin{pmatrix}4 & -3 & 0\\ -2 & 10 & -8\\ 1 & -6 & 9\end{pmatrix}$

$$\|\mathbf{A}\|_\infty=\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|=\max\{7,20,16\}=20$$

$$\|\mathbf{A}\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|=\max\{7,19,17\}=19$$

3 矩阵的谱半径

设$\mathbf{B}\in\mathbf{R}^{n\times n}$, $\rho(\mathbf{B})=\max\limits_{1\le i\le n}|\lambda_i(\mathbf{B})|$ 叫做矩阵B的谱半径。

若数λ满足 $\mathbf{Bx}=\lambda\mathbf{x},\mathbf{x}\neq\mathbf{0}$, 则$\lambda$ 称为B的一个特征值。求B的特征值，可通过求det(λI –B)=0 （或det(B –λI)=0 ）的n个根得到。

4 矩阵范数与谱半径的关系

定理 对任意矩阵 $\mathbf{B}\in\mathbf{R}^{n\times n}$，有 $\rho(\mathbf{B})\le\|\mathbf{B}\|$

其中$\|\cdot\|$是 $\mathbf{B}\in\mathbf{R}^{n\times n}$上的任何一种矩阵范数。

证明 设λ为$\mathbf{B}$的任一特征值，$\mathbf{x}\neq 0$为相应于λ的特征向量，则有 $\mathbf{Bx}=\lambda\mathbf{x}$ 。

注： $|\lambda|\|\mathbf{x}\|=\|\lambda\mathbf{x}\|=\|\mathbf{Bx}\|\le\|\mathbf{B}\|\cdot\|\mathbf{x}\| \Rightarrow |\lambda|\le\|\mathbf{B}\|$ 错误

显然存在$\mathbf{y}\in\mathbf{R}^{n}$，使$\mathbf{xy}^T$为非零矩阵，故有

$$\mathbf{Bxy}^T=\lambda\mathbf{xy}^T$$

取矩阵范数 $|\lambda|\left\|\mathbf{xy}^T\right\|=\left\|\lambda\mathbf{xy}^T\right\|=\left\|\mathbf{Bxy}^T\right\|\le\|\mathbf{B}\|\cdot\left\|\mathbf{xy}^T\right\|$

得 $|\lambda|\le\|\mathbf{B}\|$ 因λ为$\mathbf{B}$的任一特征值，故 $\rho(\mathbf{B})\le\|\mathbf{B}\|$

5 $\lim\limits_{k\to\infty}\mathbf{B}^k=\mathbf{O}$ 与谱半径的关系

定理 设$\mathbf{B}\in\mathbf{R}^{n\times n}$，则$\lim\limits_{k\to\infty}\mathbf{B}^k=\mathbf{O}$ 的充要条件是$\rho(\mathbf{B})<1$ 。

证明略

二 简单迭代法

1 简单迭代格式的构造

给定 $\mathbf{Ax}=\mathbf{b}$，$\mathbf{A}$为n阶非奇异矩阵， $\mathbf{b}$为n维向量。

将该方程组等价变形为 $\mathbf{x}=\mathbf{Bx}+\mathbf{g}$

构造简单迭代格式 $\mathbf{x}^{(k+1)}=\mathbf{Bx}^{(k)}+\mathbf{g}, k=0,1,\cdots$

取定初始向量$\mathbf{x}^{(0)}$，则产生向量序列$\{\mathbf{x}^{(k)}\}$ 。

若$\{\mathbf{x}^{(k)}\}$收敛于确定的向量 $\mathbf{x}^*$，则有 $\mathbf{x}^*=\mathbf{Bx}^*+\mathbf{g}$

即 $\mathbf{Ax}^*=\mathbf{b}$ 所以$\mathbf{x}^*$就是方程组的解。

此时称简单迭代 $\mathbf{x}^{(k+1)}=\mathbf{Bx}^{(k)}+\mathbf{g}, k=0,1,\cdots$ 关于初始向量$\mathbf{x}^{(0)}$收敛，其中$\mathbf{B}$称为简单迭代格式的迭代矩阵。

简单迭代法求$f(x)=0$的根：由等价方程$x=\varphi(x)$构造 $x_{k+1}=\varphi(x_k), k=0,1,\cdots$

Remarks

① 一个简单迭代格式可以关于某个 $\mathbf{x}^{(0)}$收敛，而关于另一个$\mathbf{x}^{(0)}$不收敛。

> 谈到格式收敛，是指对任意 $\mathbf{x}^{(0)}$，均有$\lim\limits_{k\to\infty}\mathbf{x}^{(k)}=\mathbf{x}^*$

② **Ax=b** 变形为 $\mathbf{x}=\mathbf{Bx}+\mathbf{g}$的方式不唯一。因此，对同一个方程组可以构造不同的简单迭代格式。其中可能有的收敛，而有的不收敛（发散）。

③ 当收敛时，只要k充分大，则可用 $\mathbf{x}^{(k+1)}$作为$\mathbf{x}^*$的近似值。

问题:

① **如何建立（收敛的）简单迭代格式**。

② **简单迭代格式收敛的条件**。

2 简单迭代收敛的充要条件

定理 简单迭代 $\mathbf{x}^{(k+1)}=\mathbf{B}\mathbf{x}^{(k)}+\mathbf{g},\ k=0,1,\cdots$

对任意初始向量 $\mathbf{x}^{(0)}$ 都收敛的充要条件是：$\lim\limits_{k\to\infty}\mathbf{B}^k=\mathbf{O}$

证明 设 $\mathbf{x}^*$ 是 $\mathbf{x}=\mathbf{B}\mathbf{x}+\mathbf{g}$ 的唯一解，即 $\boxed{\mathbf{x}^*=\mathbf{B}\mathbf{x}^*+\mathbf{g}}$

与相应的简单迭代 $\boxed{\mathbf{x}^{(k+1)}=\mathbf{B}\mathbf{x}^{(k)}+\mathbf{g},\ k=0,1,\cdots}$ 相减

$$\mathbf{x}^{(k+1)}-\mathbf{x}^*=\mathbf{B}(\mathbf{x}^{(k)}-\mathbf{x}^*)=\mathbf{B}^2(\mathbf{x}^{(k-1)}-\mathbf{x}^*)=\cdots=\mathbf{B}^{k+1}(\mathbf{x}^{(0)}-\mathbf{x}^*)$$

由 $\mathbf{x}^{(0)}$ 的任意性，上式右端趋于零向量的充要条件是

$$\lim_{k\to\infty}\mathbf{B}^{k+1}=\mathbf{O}$$

定理 简单迭代 $\mathbf{x}^{(k+1)}=\mathbf{B}\mathbf{x}^{(k)}+\mathbf{g},\ k=0,1,\cdots$

对任意初始向量 $\mathbf{x}^{(0)}$ 都收敛的充要条件是：$\rho(\mathbf{B})<1$

证明 根据 $\lim\limits_{k\to\infty}\mathbf{B}^k=\mathbf{O} \quad \Leftrightarrow \quad \rho(\mathbf{B})<1$ 即得

Remarks $\rho(\mathbf{B})\geq 1$ 时，有可能存在某个初始向量 $\mathbf{x}^{(0)}$，使简单迭代格式收敛。

3 简单迭代收敛的充分条件

定理 若 $\|\mathbf{B}\|<1$，则简单迭代对任意 $\mathbf{x}^{(0)}$ 收敛。其中 $\|\cdot\|$ 是 $\mathbf{B}\in\mathbf{R}^{n\times n}$ 上的任何一种矩阵范数。

证明 对任意矩阵 $\mathbf{B}\in\mathbf{R}^{n\times n}$，有 $\rho(\mathbf{B})\le\|\mathbf{B}\|$

当 $\|\mathbf{B}\|<1$，则 $\rho(\mathbf{B})<1$，故迭代收敛。

常用的矩阵范数（谱范数计算不便）

$$\|\mathbf{B}\|_\infty=\max_{1\le i\le n}\sum_{j=1}^{n}|b_{ij}|=\max_{1\le i\le n}(|b_{i1}|+|b_{i2}|+\cdots+|b_{in}|)\quad \mathbf{B}\text{的行范数}$$

$$\|\mathbf{B}\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}|b_{ij}|=\max_{1\le j\le n}(|b_{1j}|+|b_{2j}|+\cdots+|b_{nj}|)\quad \mathbf{B}\text{的列范数}$$

Remarks

① 如果 $\|\mathbf{B}\|<1$，则用 $\|\mathbf{B}\|$ 判断迭代的收敛性比用 $\rho(\mathbf{B})<1$ 更为方便，但此结论仅为充分条件。

如果 $\|\mathbf{B}\|\ge1$，则判断迭代的收敛性需考察 $\rho(\mathbf{B})<1$ 是否成立。

② 当迭代格式收敛时，常用

$$\left\|\mathbf{x}^{(k)}-\mathbf{x}^{(k-1)}\right\|_\infty=\max_{1\le i\le n}\left|x_i^{(k)}-x_i^{(k-1)}\right|<\varepsilon$$

作为迭代终止的控制条件,并取$\mathbf{x}^{(k)}$作为$\mathbf{x}^*$的近似值。

③ 收敛的迭代格式其最终计算结果与初始向量的选取无关，但初始向量的选取对迭代工作量有影响。

若对初始向量的选取无任何工程背景，则对

$$\begin{cases} x_1^{(k+1)}=b_{11}x_1^{(k)}+b_{12}x_2^{(k)}+b_{13}x_3^{(k)}+\cdots+b_{1n}x_n^{(k)}+g_1 \\ x_2^{(k+1)}=b_{21}x_1^{(k)}+b_{22}x_2^{(k)}+b_{23}x_3^{(k)}+\cdots+b_{2n}x_n^{(k)}+g_2 \\ \vdots \\ x_n^{(k+1)}=b_{n1}x_1^{(k)}+b_{n2}x_2^{(k)}+b_{n3}x_3^{(k)}+\cdots+b_{n,n}x_n^{(k)}+g_n \end{cases} \quad k=0,1,2,\cdots$$

可取$\mathbf{x}^{(0)}=\mathbf{0}$。

④ 如果需证明迭代发散，则必需证明$\rho(\mathbf{B})\ge 1$。

4 Jacobi迭代（特殊的简单迭代法）

① 迭代格式

设线性方程组为
$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{cases}$$

其系数矩阵非奇异，且 $a_{ii}\neq 0, i=1,2,\cdots,n$ 。

将方程组变形为
$$\begin{cases} x_1=\dfrac{1}{a_{11}}(-a_{12}x_2-a_{13}x_3-\cdots-a_{1n}x_n+b_1) \\ x_2=\dfrac{1}{a_{22}}(-a_{21}x_1-a_{23}x_3-\cdots-a_{2n}x_n+b_2) \\ \cdots \\ x_n=\dfrac{1}{a_{nn}}(-a_{n1}x_1-a_{n2}x_2-\cdots-a_{n,n-1}x_{n-1}+b_n) \end{cases}$$

Jacobi迭代格式的分量形式

$$
\begin{cases}
x_1^{(k+1)} = \dfrac{1}{a_{11}}(-a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)} + b_1) \\
x_2^{(k+1)} = \dfrac{1}{a_{22}}(-a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)} + b_2) \\
\cdots \\
x_n^{(k+1)} = \dfrac{1}{a_{nn}}(-a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)} + b_n)
\end{cases}
\qquad k = 0, 1, 2, \cdots
$$

该迭代格式称为**Jacobi迭代格式**。

Jacobi迭代格式的分量形式也可写为

$$
x_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij}x_j^{(k)}) \quad i = 1, 2, \cdots, n \quad k = 0, 1, 2, \cdots
$$

Jacobi迭代格式的矩阵形式

设**A**=**D**+**L**+**U**

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & \ddots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}
$$

$$
+ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ a_{21} & 0 & \ddots & \ddots & 0 \\ a_{31} & a_{32} & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ 0 & \ddots & 0 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

则 $\mathbf{Ax}=\mathbf{b} \quad \Leftrightarrow \quad (\mathbf{L}+\mathbf{D}+\mathbf{U})\mathbf{x}=\mathbf{b}$

$$\Rightarrow \quad \mathbf{Dx}=\mathbf{b}-(\mathbf{L}+\mathbf{U})\mathbf{x}$$

因 $a_{ii}\neq 0$，故**D**可逆，于是

$$\Rightarrow \quad \mathbf{x}=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})\mathbf{x}+\mathbf{D}^{-1}\mathbf{b}$$

Jacobi迭代格式的矩阵形式为

$$\boxed{\mathbf{x}^{(k+1)}=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})\mathbf{x}^{(k)}+\mathbf{D}^{-1}\mathbf{b}} \qquad k=0,1,2,\cdots$$

令 $\boxed{\mathbf{B}_J=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})} \qquad \mathbf{g}_J=\mathbf{D}^{-1}\mathbf{b}$

则得 $\mathbf{x}^{(k+1)}=\mathbf{B}_J\mathbf{x}^{(k)}+\mathbf{g}_J \quad k=0,1,2,\cdots$

$\mathbf{B}_J$ 称为Jacobi迭代格式的迭代矩阵。

$$x_i^{(k+1)}=\frac{1}{a_{ii}}(b_i-\sum_{\substack{j=1\\ j\neq i}}^{n}a_{ij}x_j^{(k)}) \quad i=1,2,\cdots,n \quad k=0,1,2,\cdots$$

② 收敛条件

充要条件

Jacobi迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 收敛的充要条件是 $\rho(\mathbf{B}_J)<1$

充分条件

（I）若 $\|\mathbf{B}_J\|<1$，则Jacobi迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 都收敛。

（II）设系数矩阵 $\mathbf{A}=(a_{ij})_{n\times n}$ 严格对角占优，

即 $$|a_{ii}|>\sum_{\substack{j=1\\j\neq i}}^{n}|a_{ij}| \quad i=1,2,\cdots,n \qquad \text{（按行）}$$

或 $$|a_{jj}|>\sum_{\substack{i=1\\i\neq j}}^{n}|a_{ij}| \quad j=1,2,\cdots,n \qquad \text{(按列)}$$

则Jacobi迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 都收敛。

$\mathbf{B}_J=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})$

证明 思路为证明 $\rho(\mathbf{B}_J)<1$，其中 $\mathbf{B}_J=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})$。

设λ为$\mathbf{B}_J$的任意特征值，则有

$$\det(\lambda\mathbf{I}-\mathbf{B}_J)=\det\left(\lambda\mathbf{I}+\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})\right)=0$$

即 $$\det\left(\lambda\mathbf{I}+\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})\right)=\det\left(\mathbf{D}^{-1}(\lambda\mathbf{D}+\mathbf{L}+\mathbf{U})\right)$$

$$=\det\left(\mathbf{D}^{-1}\right)\det(\lambda\mathbf{D}+\mathbf{L}+\mathbf{U})=0$$

于是 $\det(\lambda\mathbf{D}+\mathbf{L}+\mathbf{U})=0$

若$|\lambda|\geq 1$，因系数矩阵 $\mathbf{A}=\mathbf{L}+\mathbf{D}+\mathbf{U}$严格对角占优，

故$\lambda\mathbf{D}+(\mathbf{L}+\mathbf{U})$也严格对角占优。

于是，其行列式不为零。

这与 $\det(\lambda\mathbf{D}+\mathbf{L}+\mathbf{U})=0$矛盾。所以$|\lambda|<1$，即 $\rho(\mathbf{B}_J)<1$。

故Jacobi迭代对任意 $\mathbf{x}^{(0)}$收敛。

当A严格对角占优，A的各阶顺序主子式（含行列式）非零。

三 基于简单迭代的Gauss—Seidel迭代法

1 Gauss—Seidel 格式的构造

设有简单迭代格式 $\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{g} \quad k = 0,1,\cdots$

简单迭代格式的分量形式为

$$\begin{cases} x_1^{(k+1)} = b_{11}x_1^{(k)} + b_{12}x_2^{(k)} + b_{13}x_3^{(k)} + \cdots + b_{1n}x_n^{(k)} + g_1 \\ x_2^{(k+1)} = b_{21}x_1^{(k)} + b_{22}x_2^{(k)} + b_{23}x_3^{(k)} + \cdots + b_{2n}x_n^{(k)} + g_2 \\ x_3^{(k+1)} = b_{31}x_1^{(k)} + b_{32}x_2^{(k)} + b_{33}x_3^{(k)} + \cdots + b_{3n}x_n^{(k)} + g_3 \qquad k = 0,1,2,\cdots \\ \cdots \\ x_n^{(k+1)} = b_{n1}x_1^{(k)} + b_{n2}x_2^{(k)} + b_{n3}x_3^{(k)} + \cdots + b_{n,n}x_n^{(k)} + g_n \end{cases}$$

将得到的分量用最新计算的数值取代，得

$$\begin{cases} x_1^{(k+1)} = b_{11}x_1^{(k)} + b_{12}x_2^{(k)} + b_{13}x_3^{(k)} + \cdots + b_{1n}x_n^{(k)} + g_1 \\ x_2^{(k+1)} = b_{21}x_1^{(k+1)} + b_{22}x_2^{(k)} + b_{23}x_3^{(k)} + \cdots + b_{2n}x_n^{(k)} + g_2 \\ x_3^{(k+1)} = b_{31}x_1^{(k+1)} + b_{32}x_2^{(k+1)} + b_{33}x_3^{(k)} + \cdots + b_{3n}x_n^{(k)} + g_3 \quad k = 0,1,2,\cdots \\ \cdots \\ x_n^{(k+1)} = b_{n1}x_1^{(k+1)} + b_{n2}x_2^{(k+1)} + b_{n3}x_3^{(k+1)} + \cdots + b_{n,n}x_n^{(k)} + g_n \end{cases}$$

称为**由简单迭代格式导出的Gauss-Seidel迭代格式**。

它的分量形式也可写为

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} b_{ij}x_j^{(k+1)} + \sum_{j=i}^{n} b_{ij}x_j^{(k)} + g_i \quad i = 1,2,\cdots,n$$

对简单迭代格式 $\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{g} \quad k = 0,1,\cdots$

将**B**分解为 $\mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2$

即
$$\mathbf{B} = \begin{bmatrix} 0 & & & \\ b_{21} & 0 & & \\ \vdots & & \ddots & \\ b_{n1} & b_{n2} & \cdots & 0 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{bmatrix}$$

则
$$\mathbf{x}^{(k+1)} = \mathbf{B}_1\mathbf{x}^{(k)} + \mathbf{B}_2\mathbf{x}^{(k)} + \mathbf{g} \quad k = 0,1,\cdots$$

由简单迭代导出的Gauss-Seidel迭代格式的矩阵形式

$$\boxed{\mathbf{x}^{(k+1)} = \mathbf{B}_1\mathbf{x}^{(k+1)} + \mathbf{B}_2\mathbf{x}^{(k)} + \mathbf{g} \quad k = 0,1,\cdots}$$

若 $\mathbf{I} - \mathbf{B}_1$ 可逆，则与简单迭代格式相应的Gauss-Seidel迭代格式可化为

$$\mathbf{x}^{(k+1)} = (\mathbf{I} - \mathbf{B}_1)^{-1}\mathbf{B}_2\mathbf{x}^{(k)} + (\mathbf{I} - \mathbf{B}_1)^{-1}\mathbf{g} \quad k = 0,1,\cdots$$

$$\boxed{x_i^{(k+1)} = \sum_{j=1}^{i-1} b_{ij}x_j^{(k+1)} + \sum_{j=i}^{n} b_{ij}x_j^{(k)} + g_i \quad i = 1,2,\cdots,n}$$

2 Gauss—Seidel迭代收敛的充要条件

若 $\mathbf{I}-\mathbf{B}_1$ 可逆，则与简单迭代格式相应的Gauss-Seidel迭代收敛的充要条件为 $\rho((\mathbf{I}-\mathbf{B}_1)^{-1}\mathbf{B}_2)<1$

3 Gauss—Seidel迭代收敛的充分条件

（I） 若 $\left\|(\mathbf{I}-\mathbf{B}_1)^{-1}\mathbf{B}_2\right\|<1$，则对于任意初始向量 $\mathbf{x}^{(0)}$，与简单迭代相应的Gauss-Seidel迭代收敛。

（II） 设 $\mathbf{B}=\mathbf{B}_1+\mathbf{B}_2$，若 $\|\mathbf{B}\|_\infty<1$ 或 $\|\mathbf{B}\|_1<1$，则对于任意的初始向量 $\mathbf{x}^{(0)}$，与简单迭代相应的Gauss-Seidel迭代收敛。

对于由简单迭代导出的Gauss-Seidel迭代：

$$\mathbf{x}^{(k+1)}=\mathbf{B}_1\mathbf{x}^{(k+1)}+\mathbf{B}_2\mathbf{x}^{(k)}+\mathbf{g}\quad k=0,1,\cdots$$

将其改写为简单迭代：

$$\mathbf{x}^{(k+1)}=\left(\mathbf{I}-\mathbf{B}_1\right)^{-1}\mathbf{B}_2\mathbf{x}^{(k)}+\left(\mathbf{I}-\mathbf{B}_1\right)^{-1}\mathbf{g}\quad k=0,1,\cdots$$

证明 仅以$\|\mathbf{B}\|_\infty = \max\limits_{1\le i\le n}\sum\limits_{j=1}^{n}|b_{ij}| < 1$为例。

思路： 证明$\lim\limits_{k\to\infty} x_i^{(k+1)} = x_i^* \Rightarrow \lim\limits_{k\to\infty}\mathbf{x}^{(k+1)} = \mathbf{x}^*$

由
$$\begin{cases}\mathbf{x}^{(k+1)} = \mathbf{B}_1\mathbf{x}^{(k+1)} + \mathbf{B}_2\mathbf{x}^{(k)} + \mathbf{g} \quad k = 0,1,\cdots \\ \mathbf{x}^* = \mathbf{B}\mathbf{x}^* + \mathbf{g}\end{cases}$$

的分量形式

$$x_i^{(k+1)} = \sum_{j=1}^{i-1} b_{ij}x_j^{(k+1)} + \sum_{j=i}^{n} b_{ij}x_j^{(k)} + g_i \qquad x_i^* = \sum_{j=1}^{n} b_{ij}x_j^* + g_i$$

相减有 $x_i^{(k+1)} - x_i^* = \sum\limits_{j=1}^{i-1} b_{ij}(x_j^{(k+1)} - x_j^*) + \sum\limits_{j=i}^{n} b_{ij}(x_j^{(k)} - x_j^*)$

得 $\left|x_i^{(k+1)} - x_i^*\right| \le \sum\limits_{j=1}^{i-1}|b_{ij}|\left|x_j^{(k+1)} - x_j^*\right| + \sum\limits_{j=i}^{n}|b_{ij}|\left|x_j^{(k)} - x_j^*\right|$

记 $\delta_k = \max\limits_{1\le i\le n}\left|x_i^{(k)} - x_i^*\right| \quad \alpha_i = \sum\limits_{j=1}^{i-1}\left|b_{ij}\right| \quad \beta_i = \sum\limits_{j=i}^{n}\left|b_{ij}\right|$

则 $\left|x_i^{(k+1)} - x_i^*\right| \le \delta_{k+1}\alpha_i + \delta_k\beta_i \quad i = 1,2,\cdots,n$

上式对i=1,2,…,n都成立，故有 $i = i_0$，使

$$\left|x_{i_0}^{(k+1)} - x_{i_0}^*\right| = \delta_{k+1}$$

于是，取 $i = i_0$，有 $\delta_{k+1} \le \delta_{k+1}\alpha_{i_0} + \delta_k\beta_{i_0}$

即 $$\delta_{k+1} \le \left(\frac{\beta_{i_0}}{1-\alpha_{i_0}}\right)\delta_k \le \left(\frac{\beta_{i_0}}{1-\alpha_{i_0}}\right)\left(\frac{\beta_{i_1}}{1-\alpha_{i_1}}\right)\delta_{k-1}$$

$$\le \cdots \le \left(\max_{i_k}\frac{\beta_{i_k}}{1-\alpha_{i_k}}\right)^{k+1}\delta_0$$

其中 $\delta_0 = \max\limits_{1\le i\le n}\left|x_i^{(0)} - x_i^*\right| \ne 0$

$$\left|x_i^{(k+1)} - x_i^*\right| \le \sum_{j=1}^{i-1}\left|b_{ij}\right|\left|x_j^{(k+1)} - x_j^*\right| + \sum_{j=i}^{n}\left|b_{ij}\right|\left|x_j^{(k)} - x_j^*\right|$$

注意： $\alpha_i + \beta_i = \sum_{j=1}^{n} |b_{i,j}| \le \max_{1\le i\le n} \sum_{j=1}^{n} |b_{i,j}| = \|\mathbf{B}\|_\infty < 1$

即 $1-\alpha_i > \beta_i > 0 \quad i=1,2,\cdots,n$ 所以 $0 \le \max_{i_k} \frac{\beta_{i_k}}{1-\alpha_{i_k}} < 1$

故当$k\to\infty$时，则 $(\max_{i_k} \frac{\beta_{i_k}}{1-\alpha_{i_k}})^{k+1} \to 0$

即当$k\to\infty$时，有 $\delta_{k+1} = \max_{1\le i\le n} |x_i^{(k+1)} - x_i^*| \to 0$

即 $\lim_{k\to\infty} x_i^{(k+1)} = x_i^*$

由序列收敛的等价性定理，知 $\lim_{k\to\infty} \mathbf{x}^{(k+1)} = \mathbf{x}^*$ 对任意的 $\mathbf{x}^{(0)}$ 都成立。所以，与简单迭代相应的Gauss-Seidel迭代收敛。

$\alpha_i = \sum_{j=1}^{i-1} |b_{ij}|$ $\beta_i = \sum_{j=i}^{n} |b_{ij}|$ $\delta_{k+1} \le (\max_{i_k} \frac{\beta_{i_k}}{1-\alpha_{i_k}})^{k+1} \delta_0$

4 与Jacobi迭代相应的Gauss-Seidel（JGS）迭代

① 迭代格式

与Jacobi迭代相应的Gauss-Seidel迭代的分量形式

由Jacobi迭代格式：

$$\begin{cases} x_1^{(k+1)} = \dfrac{1}{a_{11}}(-a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)} + b_1) \\ x_2^{(k+1)} = \dfrac{1}{a_{22}}(-a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)} + b_2) \\ \cdots \\ x_n^{(k+1)} = \dfrac{1}{a_{nn}}(-a_{n1}x_1^{(k)} - a_{n2}x_2^{(k)} - \cdots - a_{n,n-1}x_{n-1}^{(k)} + b_n) \end{cases} \qquad k = 0,1,2,\cdots$$

将Jacobi迭代格式中得到的分量用最新计算的数值取代，即得**与Jacobi迭代相应的Gauss-Seidel迭代格式**。

$$\begin{cases} x_1^{(k+1)} = \dfrac{1}{a_{11}}(-a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)} + b_1) \\ x_2^{(k+1)} = \dfrac{1}{a_{22}}(-a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)} + b_2) \\ \cdots \\ x_n^{(k+1)} = \dfrac{1}{a_{nn}}(-a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)} + b_n) \end{cases} \quad k = 0,1,2,\cdots$$

该格式简称为**JGS迭代格式**。它的分量形式也可写为

$$x_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}) \quad i = 1,2,\cdots,n$$

与Jacobi迭代相应的Gauss-Seidel迭代的矩阵形式

设**A**=**D**+**L**+**U**

$$
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & \ddots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & 0 & a_{nn} \end{bmatrix}
$$

$$
+\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ a_{21} & 0 & \ddots & \ddots & 0 \\ a_{31} & a_{32} & 0 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ 0 & \ddots & 0 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}
$$

则 $\mathbf{Ax}=\mathbf{b} \Leftrightarrow (\mathbf{L}+\mathbf{D}+\mathbf{U})\mathbf{x}=\mathbf{b} \Leftrightarrow \mathbf{Dx}=-\mathbf{Lx}-\mathbf{Ux}+\mathbf{b}$

JGS迭代格式的矩阵形式为

$$\boxed{\mathbf{x}^{(k+1)}=-\mathbf{D}^{-1}\mathbf{L}\mathbf{x}^{(k+1)}-\mathbf{D}^{-1}\mathbf{U}\mathbf{x}^{(k)}+\mathbf{D}^{-1}\mathbf{b}} \quad k=0,1,2,\cdots$$

变形 $\mathbf{Dx}^{(k+1)}=-\mathbf{Lx}^{(k+1)}-\mathbf{Ux}^{(k)}+\mathbf{b} \qquad k=0,1,2,\cdots$

即 $(\mathbf{D}+\mathbf{L})\mathbf{x}^{(k+1)}=-\mathbf{Ux}^{(k)}+\mathbf{b} \qquad k=0,1,2,\cdots$

因 $a_{ii}\neq 0$，故$\mathbf{D}+\mathbf{L}$可逆，于是

JGS迭代格式的矩阵形式为

$$\mathbf{x}^{(k+1)}=-(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U}\mathbf{x}^{(k)}+(\mathbf{D}+\mathbf{L})^{-1}\mathbf{b} \qquad k=0,1,2,\cdots$$

令 $\boxed{\mathbf{B}_{JGS}=-(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U}} \qquad \mathbf{g}_{JGS}=(\mathbf{D}+\mathbf{L})^{-1}\mathbf{b}$

则得 $\mathbf{x}^{(k+1)}=\mathbf{B}_{JGS}\mathbf{x}^{(k)}+\mathbf{g}_{JGS} \qquad k=0,1,2,\cdots$

$\mathbf{B}_{JGS}$ 称为JGS迭代格式的迭代矩阵。

$$\boxed{x_i^{(k+1)}=\frac{1}{a_{ii}}\left(b_i-\sum_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum_{j=i+1}^{n}a_{ij}x_j^{(k)}\right) \quad i=1,2,\cdots,n}$$

② 收敛条件

充要条件

JGS迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 收敛的充要条件是 $\rho(\mathbf{B}_{JGS})<1$

充分条件

（I）若 $\|\mathbf{B}_{JGS}\|<1$ ，则JGS迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 都收敛。

（II）设系数矩阵 $\mathbf{A}=(a_{ij})_{n\times n}$ 严格对角占优，则JGS迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 都收敛。

（证明与Jacobi迭代类似）

（III）设A为对称正定矩阵，则JGS迭代关于任意初始向量 $\mathbf{x}^{(0)}$ 都收敛。（证明略）

$$\mathbf{B}_{JGS}=-(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U}$$

简单迭代 → **与简单迭代格式相应的Gauss -Seidel迭代**

$$\mathbf{x}^{(k+1)}=\mathbf{B}\mathbf{x}^{(k)}+\mathbf{g}$$

充要 $\rho(\mathbf{B})<1$

充分 $\|\mathbf{B}\|<1$

$$\mathbf{x}^{(k+1)}=\mathbf{B}_1\mathbf{x}^{(k+1)}+\mathbf{B}_2\mathbf{x}^{(k)}+\mathbf{g}$$

充要 $\rho((\mathbf{I}-\mathbf{B}_1)^{-1}\mathbf{B}_2)<1$

充分 $\left\|(\mathbf{I}-\mathbf{B}_1)^{-1}\mathbf{B}_2\right\|<1$

充分 $\|\mathbf{B}_1+\mathbf{B}_2\|_1<1$ 或 $\|\mathbf{B}_1+\mathbf{B}_2\|_\infty<1$

$\|\cdot\|$ 是任何一种矩阵范数。

Jacobi迭代 → **与Jacobi迭代相应的Gauss-Seidel迭代**

$$\boxed{\mathbf{B}_J=-\mathbf{D}^{-1}(\mathbf{L}+\mathbf{U})}$$

充要 $\rho(\mathbf{B}_J)<1$

充分 $\|\mathbf{B}_J\|<1$

充分 **A**严格对角占优

$$\boxed{\mathbf{B}_{JGS}=-(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U}}$$

充要 $\rho(\mathbf{B}_{JGS})<1$

充分 $\|\mathbf{B}_{JGS}\|<1$

充分 **A**严格对角占优

充分 **A**对称正定

例 判断对错：对简单迭代 $\mathbf{x}^{(k+1)}=\mathbf{B}\mathbf{x}^{(k)}+\mathbf{g}$ 与相应的G-S迭代 $\mathbf{x}^{(k+1)}=\mathbf{B}_1\mathbf{x}^{(k+1)}+\mathbf{B}_2\mathbf{x}^{(k)}+\mathbf{g}$, $\mathbf{B}=\mathbf{B}_1+\mathbf{B}_2$,

(a) 当 $\|\mathbf{B}\|<1$，两者关于任意 $\mathbf{x}^{(0)}$ 收敛。$\|\mathbf{B}\|$ 为任何一种矩阵范数。 错

(b) 当 $\|\mathbf{B}\|_\infty<1$ 或 $\|\mathbf{B}\|_1<1$，两者关于任意 $\mathbf{x}^{(0)}$ 收敛。 对

Remarks

① 书面作业：对Jacobi迭代与JGS迭代收敛性首先考察系数矩阵A是否严格对角占优。对于JGS迭代其次考察系数矩阵A是否对称正定.然后判断 $\|\mathbf{B}_J\|<1$、$\|\mathbf{B}_{JGS}\|<1$ 是否成立。最后考察 $\rho(\mathbf{B}_J)<1$、$\rho(\mathbf{B}_{JGS})<1$ 是否成立。

② Jacobi迭代与Gauss-Seidel迭代的收敛性没有确定关系。两者可以同时收敛、同时发散或其中某个收敛而另一个发散。若Jacobi迭代与Gauss-Seidel迭代同时收敛，一般Gauss-Seidel迭代收敛速度较快。

例题：给定如下线性方程组

$$\begin{cases} 2x + y + 3z = 6 \\ x + 4y \qquad = 8 \\ 2x \qquad + z = 2 \end{cases}$$

（1）证明用（相应Jacobi迭代的）Gauss-Seidel迭代格式求解上述方程组不收敛；

（2）构造一种保证收敛的Gauss-Seidel迭代格式，并说明收敛的原因。

解：Jacobi迭代对应的Gauss-Seidel迭代，其迭代矩阵为

$$\mathbf{B}_{JGS} = -(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U}$$

且 $\mathbf{D}+\mathbf{L} = \begin{bmatrix} 2 & & \\ 1 & 4 & \\ 2 & & 1 \end{bmatrix}$ $(\mathbf{D}+\mathbf{L})^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{8} & \frac{1}{4} & 0 \\ -1 & 0 & 1 \end{bmatrix}$

$$\mathbf{B}_{JGS} = -(\mathbf{D}+\mathbf{L})^{-1}\mathbf{U} = -\begin{bmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{8} & \frac{1}{4} & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{3}{2} \\ 0 & \frac{1}{8} & \frac{3}{8} \\ 0 & 1 & 3 \end{bmatrix}$$

$$\begin{cases} 2x+y+3z=6 \\ x+4y=8 \\ 2x+z=2 \end{cases}$$

由 $|\lambda\mathbf{I}-\mathbf{B}_{JGS}|=\det\begin{bmatrix}\lambda & \frac{1}{2} & \frac{3}{2}\\ 0 & \lambda-\frac{1}{8} & -\frac{3}{8}\\ 0 & -1 & \lambda-3\end{bmatrix}$

$$=\lambda(\lambda^2-\frac{25}{8}\lambda+\frac{3}{8}-\frac{3}{8})=\lambda^2(\lambda-\frac{25}{8})$$

得对应的特征值为 $\lambda_1=\lambda_2=0 \quad \lambda_3=\frac{25}{8}$

谱半径 $\rho(\mathbf{B}_{JGS})=\frac{25}{8}>1$

故（与Jacobi迭代对应的）Gauss-Seidel迭代发散。

将原方程组变形为 $\begin{cases} 2x \quad + z = 2 \\ x + 4y \quad = 8 \\ 2x + y + 3z = 6 \end{cases}$ （并非严格对角占优）

构造JGS迭代格式

$$\begin{cases} x^{(k+1)} = -\dfrac{1}{2} z^{(k)} + 1 \\ y^{(k+1)} = -\dfrac{1}{4} x^{(k+1)} + 2 \\ z^{(k+1)} = -\dfrac{2}{3} x^{(k+1)} - \dfrac{1}{3} y^{(k+1)} + 2 \end{cases} \quad k = 0, 1, \cdots \quad \mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2 = \begin{bmatrix} 0 & 0 & -\dfrac{1}{2} \\ -\dfrac{1}{4} & 0 & 0 \\ -\dfrac{2}{3} & -\dfrac{1}{3} & 0 \end{bmatrix}$$

此时有 $\|\mathbf{B}\|_1 = \max\{\dfrac{1}{4} + \dfrac{2}{3}, \dfrac{1}{3}, \dfrac{1}{2}\} = \dfrac{11}{12} < 1$

故针对上述变形的方程组，
所得的JGS迭代格式收敛。

$$\begin{cases} 2x + y + 3z = 6 \\ x + 4y \quad = 8 \\ 2x \quad + z = 2 \end{cases}$$

或将原方程组变形为 $\begin{cases} 2x \quad + z = 2 & (3) \\ x + 4y \quad = 8 & (2) \\ \quad y + 2z = 4 & (1)-(3) \end{cases}$

则系数矩阵严格对角占优，构造JGS迭代格式为

$$\begin{cases} x^{(k+1)} = -\dfrac{1}{2} z^{(k)} + 1 \\ y^{(k+1)} = -\dfrac{1}{4} x^{(k+1)} + 2 \quad k = 0, 1, \cdots \\ z^{(k+1)} = -\dfrac{1}{2} y^{(k+1)} + 2 \end{cases} \quad \mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2 = \begin{bmatrix} 0 & 0 & -\dfrac{1}{2} \\ -\dfrac{1}{4} & 0 & 0 \\ 0 & -\dfrac{1}{2} & 0 \end{bmatrix}$$

故针对上述变形的方程组，所得的JGS迭代格式收敛。

（此时，迭代矩阵有 $\|\mathbf{B}\|_1 = \dfrac{1}{2} < 1, \|\mathbf{B}\|_\infty = \dfrac{1}{2} < 1$）。

$$\begin{cases} 2x + y + 3z = 6 \\ x + 4y \quad = 8 \\ 2x \quad + z = 2 \end{cases}$$

Remarks

① 对于发散的迭代格式有时可以通过方程组变形构造收敛的迭代格式。

② 简单迭代法的收敛快慢，依赖于迭代矩阵谱半径的大小。当 $\rho(\mathbf{B})<1$ ，则迭代矩阵谱半径越小，收敛越快。当 $\rho(\mathbf{B})=0,$ 对 n 阶线性代数方程组，理论上迭代 n 步即可得到精确解（初始向量的选取对迭代步数有影响，最多迭代n 步）。

四 逐次超松弛迭代法（SOR法）

① 迭代格式

设 $a_{ii} \neq 0$，将JGS迭代格式的分量形式

$$x_i^{(k+1)} = \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}) \quad i = 1, 2, \cdots, n$$

改写为 $x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}}(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij} x_j^{(k)})$

一般 $b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij} x_j^{(k)} \neq 0$

$$\begin{cases} x_1^{(k+1)} = \frac{1}{a_{11}}(-a_{12}x_2^{(k)} - a_{13}x_3^{(k)} - \cdots - a_{1n}x_n^{(k)} + b_1) \\ x_2^{(k+1)} = \frac{1}{a_{22}}(-a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)} + b_2) \\ \cdots \\ x_n^{(k+1)} = \frac{1}{a_{nn}}(-a_{n1}x_1^{(k+1)} - a_{n2}x_2^{(k+1)} - \cdots - a_{n,n-1}x_{n-1}^{(k+1)} + b_n) \end{cases} \quad k = 0, 1, 2, \cdots$$

$$\begin{cases} b_1 - a_{11}x_1 - \cdots - a_{1n}x_n = 0 \\ b_2 - a_{21}x_1 - \cdots - a_{2n}x_n = 0 \\ \cdots\cdots\cdots \\ b_n - a_{n1}x_1 - \cdots - a_{nn}x_n = 0 \end{cases}$$

在$x_i^{(k)}$的补偿量中乘以松驰因子ω，得

SOR（Successive Over Relaxation）迭代格式的分量形式

$$x_i^{(k+1)} = x_i^{(k)} + \frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij} x_j^{(k)}\right)$$

$$i = 1, 2, \cdots, n$$

SOR迭代的计算公式也常写为：

$$x_i^{(k+1)} = (1-\omega) x_i^{(k)} + \frac{\omega}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}\right)$$

$$i = 1, 2, \cdots, n$$

ω=1 的SOR迭代即为JGS迭代。

适当选取松驰因子ω的值，可以得到比JGS方法收敛更快的格式。实际计算一般根据试算确定ω。

$$x_i^{(k+1)} = x_i^{(k)} + \frac{1}{a_{ii}}\left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij} x_j^{(k)}\right)$$

SOR迭代格式的矩阵形式

将 $x_i^{(k+1)}=(1-\omega)x_i^{(k)}+\dfrac{\omega}{a_{ii}}(b_i-\sum\limits_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum\limits_{j=i+1}^{n}a_{ij}x_j^{(k)})$

改写 $a_{ii}x_i^{(k+1)}=(1-\omega)a_{ii}x_i^{(k)}+\omega(b_i-\sum\limits_{j=1}^{i-1}a_{ij}x_j^{(k+1)}-\sum\limits_{j=i+1}^{n}a_{ij}x_j^{(k)})$

设**A**=**D**+**L**+**U**，则SOR迭代的矩阵形式为

$$\mathbf{Dx}^{(k+1)}=(1-\omega)\mathbf{Dx}^{(k)}+\omega(\mathbf{b}-\mathbf{Lx}^{(k+1)}-\mathbf{Ux}^{(k)})$$

即 $(\mathbf{D}+\omega\mathbf{L})\mathbf{x}^{(k+1)}=((1-\omega)\mathbf{D}-\omega\mathbf{U})\mathbf{x}^{(k)}+\omega\mathbf{b}$

因 $a_{ii}\neq 0$，故$\mathbf{D}+\omega\mathbf{L}$可逆，于是

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

令 $\mathbf{B}_\omega=(\mathbf{D}+\omega\mathbf{L})^{-1}((1-\omega)\mathbf{D}-\omega\mathbf{U})\quad \mathbf{g}_\omega=\omega(\mathbf{D}+\omega\mathbf{L})^{-1}\mathbf{b}$

则得 $\mathbf{x}^{(k+1)}=\mathbf{B}_\omega\mathbf{x}^{(k)}+\mathbf{g}_\omega \qquad k=0,1,2,\cdots$

$\mathbf{B}_\omega$称为SOR迭代格式的迭代矩阵。

② 收敛的充要条件

SOR迭代关于任意初始向量$\mathbf{x}^{(0)}$收敛的充要条件是 $\rho(\mathbf{B}_{\omega})<1$ 这里 $\mathbf{B}_{\omega}=(\mathbf{D}+\omega\mathbf{L})^{-1}((1-\omega)\mathbf{D}-\omega\mathbf{U})$

③ 收敛的必要条件

SOR迭代格式关于任意初始向量 $\mathbf{x}^{(0)}$ 收敛的必要条件为 $0<\omega<2$。 （证明略）

④ 收敛的充分条件

（I） 若$\|\mathbf{B}_{\omega}\|<1$，则SOR迭代关于任意初始向量$\mathbf{x}^{(0)}$收敛 。

（II ） 若**A**为实对称正定矩阵，且$0<\omega<2$，则SOR迭代关于任意初始向量 $\mathbf{x}^{(0)}$收敛 (证明略)。

（III ） 若系数矩阵**A** 严格对角占优，且$0<\omega\le 1$，则SOR迭代关于任意初始向量$\mathbf{x}^{(0)}$收敛 (证明略)。

Remarks

① SOR迭代收敛的必要条件说明：要使SOR迭代收敛，必须选取$0 < \omega < 2$；但$0 < \omega < 2$ 未必能保证SOR迭代收敛。

② 对SOR迭代首先考察系数矩阵**A**是否严格对角占优（ $0< \omega \le 1$收敛），其次考察**A**是否对称正定（ $0 < \omega<2$ 收敛）。再判断$\|\mathbf{B}_{\omega}\| < 1$是否成立。最后考察$\rho(\mathbf{B}_{\omega}) < 1$是否成立。

③ 当松弛因子 ω适当时, SOR方法收敛很快，不适当时收敛非常慢。但最佳松弛因子的选取依赖一定的理论基础与实际计算经验。

总结

1. 基本迭代格式的构造、收敛性判断以及方程组的求解

- 简单迭代格式的构造、收敛性判断以及方程组的求解
- 由简单迭代导出的Gauss—Seidel迭代格式的构造、收敛性判断以及方程组的求解

2. 三种常用迭代格式的构造、收敛性判断以及方程组的求解

- Jacobi迭代
- 基于Jacobi迭代的Gauss—Seidel迭代
- 逐次超松弛迭代

3. 矩阵三角分解法及其方程组求解

- 直接三角分解法
- 平方根法
- 追赶法

4. Gauss消去法

- Gauss主元消去法

  （列主元素消去法、总体主元素消去法）
- Gauss顺序消去法