5 解线性代数方程组的直接法

在自然科学和工程技术中,许多问题的解决常常归结为线性代数方程组的求解.例如,电学中的网络问题、船体放样中三次样条求解问题、求解非线性方程组问题、常微分方程及偏微分方程数值求解问题等,都可归结为求解线性代数方程组.有关线性代数方程组解的存在惟一性及解的结构等理论问题,线性代数已作了详细讨论.本书只介绍线性代数方程组的两类求解方法:直接法和迭代法.

直接法是指经过有限步计算后求得方程组精确解的方法.本章研究的对象是 n 阶线性代数方程组:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{5.1}$$

若用矩阵和向量的记号来表示,方程组(5.1)可写成

$$\boldsymbol{AX} = \boldsymbol{b} \tag{5.2}$$

其中,$\boldsymbol{A} = (a_{ij})$ 是 $n \times n$ 系数矩阵;$\boldsymbol{X} = \{x_i\}$ 是 n 维向量,为所求的解;$\boldsymbol{b} = \{b_i\}$,$\boldsymbol{b}$ 称为方程组(5.1)的右端项.

若 $\boldsymbol{A}$ 是非奇异矩阵,则方程组有惟一解.记 $D = \det A$,应用 Cramer 法则可得

$$x_i = \frac{D_i}{D} \quad (i = 1, 2, \cdots, n)$$

其中,D_i 是用 $\boldsymbol{b}$ 代替 $\boldsymbol{A}$ 中第 i 列而得到的相应行列式.然而在实际中,用 Cramer 法则求解(5.1)式并不可行.n 阶行列式共有 $n!$ 项,每项有 n 个因子,所以计算一个 n 阶行列式共需 $(n-1) \cdot n!$ 次乘法,应用 Cramer 法则求解方程组需要计算 $(n+1)$ 个行列式,另外要计算 x_i,还需 n 次除法.因此,用 Cramer 法则求解方程组(5.1)的乘除次数为

$$N = (n^2 - 1) \cdot n! + n$$

当 $n = 10$ 时,$N = 359\ 251\ 210$;当 $n = 20$ 时,$N = 9.707\ 3 \times 10^{20}$,可见,Cramer 法则虽然简单,但并不能直接用来求解线性代数方程组.

5.1　高斯消去法

5.1.1　顺序高斯消去法

高斯消去法是一个古老的方法,但实践证明它仍是目前计算机上一种常用的有效方法,其基本思想是通过初等变换将方程组(5.1)转化为一个等价的三角形方程组

$$\begin{cases} b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n = g_1 \\ \qquad\quad b_{22}x_2 + \cdots + b_{2n}x_n = g_2 \\ \qquad\qquad\qquad\qquad \vdots \\ \qquad\qquad\qquad\quad b_{nn}x_n = g_n \end{cases} \tag{5.3}$$

这个过程称为消元,得到方程组(5.3)后,就可逐个求出 $x_n, x_{n-1}, \cdots, x_1$. 这个过程称为回代,具体过程如下:

从方程组(5.3)中的最后一个方程直接求出 x_n

$$x_n = \frac{g_n}{b_{nn}}$$

代入方程组(5.3)中倒数第二个方程,便有

$$x_{n-1} = \frac{g_{n-1} - b_{n-1,n}x_n}{b_{n-1,n-1}}$$

一般地,有

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

5.1.1.1　顺序高斯消去法的计算过程

为了统一符号,把方程组(5.1)改写成下面形式

$$\begin{cases} a_{11}^{(1)}x_1 + a_{12}^{(1)}x_2 + \cdots + a_{1n}^{(1)}x_n = b_1^{(1)} \\ a_{21}^{(1)}x_1 + a_{22}^{(1)}x_2 + \cdots + a_{2n}^{(1)}x_n = b_2^{(1)} \\ \qquad\qquad\qquad\qquad \vdots \\ a_{n1}^{(1)}x_1 + a_{n2}^{(1)}x_2 + \cdots + a_{nn}^{(1)}x_n = b_n^{(1)} \end{cases} \tag{5.4}$$

用矩阵表示即为

$$\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{b}^{(1)} \tag{5.5}$$

其中
$$A^{(1)}=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix} \qquad b^{(1)}=\begin{bmatrix} b_1^{(1)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{bmatrix}$$

设 $a_{11}^{(1)}\neq 0$,令 $m_{i1}=\dfrac{a_{i1}^{(1)}}{a_{11}^{(1)}}$,将方程组(5.4)第1行乘 $-m_{i1}$ 后加到第 i 行($i=2,3,\cdots,n$),则方程组(5.4)化为

$$\begin{cases} a_{11}^{(1)}x_1+a_{12}^{(1)}x_2+\cdots+a_{1n}^{(1)}x_n=b_1^{(1)} \\ \qquad\quad a_{22}^{(2)}x_2+\cdots+a_{2n}^{(2)}x_n=b_2^{(2)} \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ \qquad\quad a_{n2}^{(2)}x_2+\cdots+a_{nn}^{(2)}x_n=b_n^{(2)} \end{cases} \tag{5.6}$$

其中
$$\begin{cases} a_{ij}^{(2)}=a_{ij}^{(1)}-m_{i1}a_{1j}^{(1)} & (i,j=2,3,\cdots,n) \\ b_i^{(2)}=b_i^{(1)}-m_{i1}b_1^{(1)} & (i=2,3,\cdots,n) \end{cases}$$

用矩阵符号表示为
$$A^{(2)}X=b^{(2)} \tag{5.7}$$

其中
$$A^{(2)}=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix} \qquad b^{(2)}=\begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{bmatrix}$$

类似地,若设 $a_{22}^{(2)}\neq 0$,令 $m_{i2}=\dfrac{a_{i2}^{(2)}}{a_{22}^{(2)}}$,将方程组(5.6)中第2个方程乘 $-m_{i2}$ 后加到第 i 行($i=3,4,\cdots,n$),则方程组(5.6)化为

$$\begin{cases} a_{11}^{(1)}x_1+a_{12}^{(1)}x_2+a_{13}^{(1)}x_3+\cdots+a_{1n}^{(1)}x_n=b_1^{(1)} \\ \qquad\quad a_{22}^{(2)}x_2+a_{23}^{(2)}x_3+\cdots+a_{2n}^{(2)}x_n=b_2^{(2)} \\ \qquad\qquad\qquad a_{33}^{(3)}x_3+\cdots+a_{3n}^{(3)}x_n=b_3^{(3)} \\ \qquad\qquad\qquad\qquad\qquad\qquad\vdots \\ \qquad\qquad\qquad a_{n3}^{(3)}x_3+\cdots+a_{nn}^{(3)}x_n=b_n^{(3)} \end{cases} \tag{5.8}$$

其中
$$\begin{cases} a_{ij}^{(3)}=a_{ij}^{(2)}-m_{i2}a_{2j}^{(2)} & (i,j=3,4,\cdots,n) \\ b_i^{(3)}=b_i^{(2)}-m_{i2}b_2^{(2)} & (i=3,4,\cdots,n) \end{cases}$$

用矩阵符号表示为
$$A^{(3)}X=b^{(3)} \tag{5.9}$$

其中
$$A^{(3)}=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2n}^{(2)} \\ 0 & 0 & a_{33}^{(3)} & \cdots & a_{3n}^{(3)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{n3}^{(3)} & \cdots & a_{nn}^{(3)} \end{bmatrix} \qquad b^{(3)}=\begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ b_3^{(3)} \\ \vdots \\ b_n^{(3)} \end{bmatrix}$$

若 $a_{33}^{(3)}\neq 0$,上述过程可以继续进行下去,重复 $n-1$ 次后,可得到等价的三角形方程组

$$\begin{cases} a_{11}^{(1)}x_1+a_{12}^{(1)}x_2+\cdots+a_{1n}^{(1)}x_n=b_1^{(1)} \\ \qquad\qquad a_{22}^{(2)}x_2+\cdots+a_{2n}^{(2)}x_n=b_2^{(2)} \\ \qquad\qquad\qquad\qquad\qquad\vdots \\ \qquad\qquad\qquad\qquad a_{nn}^{(n)}x_n=b_n^{(n)} \end{cases} \tag{5.10}$$

用矩阵表示为

$$A^{(n)}X=b^{(n)} \tag{5.11}$$

其中
$$A^{(n)}=\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(n)} \end{bmatrix} \qquad b^{(n)}=\begin{bmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{bmatrix}$$

消元过程中元素的计算公式可以归纳为

$$\begin{cases} a_{ij}^{(k+1)}=a_{ij}^{(k)} \qquad b_i^{(k+1)}=b_i^{(k)} & (1\leqslant i\leqslant k,i\leqslant j\leqslant n) \\ a_{ij}^{(k+1)}=a_{ij}^{(k)}-m_{ik}a_{kj}^{(k)} & (i,j=k+1,\cdots,n) \\ b_i^{(k+1)}=b_i^{(k)}-m_{ik}b_k^{(k)} & (i=k+1,\cdots,n) \\ a_{ij}^{(k+1)}=0 & (1\leqslant j\leqslant k<i\leqslant n) \end{cases}$$

从上述消元过程可以看到,顺序高斯消去法能进行下去的充要条件是 $\boldsymbol{A}$ 矩阵的所有顺序主子式不为零.

5.1.1.2 顺序高斯消去法求解算法

①输入系数矩阵 $\boldsymbol{A}$、右端项 $\boldsymbol{b}$ 及 ε.

②消元:

　　对 $k=1,2,\cdots,n-1$ 循环

　　　　若 $|a_{kk}|\leqslant\varepsilon$,则打印"求解失败",停机;否则

　　　　for　$i=k+1$　to　n　do

　　　　　$T=\dfrac{a_{ik}}{a_{kk}}$

　　　　　$b_i=b_i-T\times b_k$

for $j=k+1$ to n do

$a_{ij}=a_{ij}-T\times a_{kj}$

③回代：

若$|a_{nn}|\leqslant\varepsilon$，则打印"求解失败"，停机；否则

$x_n=\dfrac{b_n}{a_{nn}}$

for $i=n-1$ downto 1 do

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

④打印 $x_i,i=1,2,\cdots,n$

上述算法中，为了不增加存储量，消元过程直接在系数矩阵 $\boldsymbol{A}$ 中进行. ε 的引用主要是为了控制溢出.

5.1.1.3 高斯消去法的计算量

(1)消元过程的计算量

第 1 次消元：

计算 $m_{i1}(i=2,3,\cdots,n)$需要 $n-1$ 次除法；

计算 $a_{ij}^{(2)}=a_{ij}^{(1)}-m_{i1}a_{1j}^{(1)}(i,j=2,3,\cdots,n)$需要$(n-1)^2$ 次乘法；

计算 $b_i^{(2)}=b_i^{(1)}-m_{i1}b_1^{(1)}(i=2,3,\cdots,n)$需要 $n-1$ 次乘法.

因此经过一次消元后乘除次数为 n^2-1，加减法次数为$(n-1)^2+(n-1)=n(n-1)$.

第 2 次消元：

第 2 次消元等价于一个 $n-1$ 阶方程组的一次消元，所需乘除法次数为$(n-1)^2-1$，加减法次数为$(n-1)(n-2)$. 因此消元过程的计算量为：

乘除法次数

$$\sum_{i=0}^{n-1}[(n-i)^2-1]=\sum_{i=1}^{n}i^2-n=\frac{2n^3+3n^2-5n}{6}$$

加减法次数

$$\sum_{i=1}^{n}i(i-1)=\sum_{i=1}^{n}i^2-\sum_{i=1}^{n}i=\frac{n^3-n}{3}$$

(2)回代过程的计算量

乘除法次数

$$1+\sum_{i=1}^{n-1}[(n-i)+1]=\frac{n^2+n}{2}$$

加减法次数

$$\sum_{i=1}^{n-1}\left[(n-i-1)+1\right]=\frac{n^2-n}{2}$$

从而顺序高斯消去法的计算量为：

乘除法次数

$$\frac{2n^3+3n^2-5n}{6}+\frac{n^2+n}{2}=\frac{n^3}{3}+n^2-\frac{n}{3}$$

加减法次数

$$\frac{n^3-n}{3}+\frac{n^2-n}{2}=\frac{n^3}{3}+\frac{n^2}{2}-\frac{5}{6}n$$

5.1.1.4　顺序高斯消去法的矩阵解释

在上述消元过程中用到了矩阵初等变换，即将某一行减去另一行的常数倍. 在后面的列主元和全主元高斯消去法中还会用到矩阵的行交换与列交换. 现对相应初等变换的矩阵表示说明如下：

(1) 交换矩阵的两行、两列的变换

利用单位坐标向量 $\boldsymbol{e}_i\,(i=1,2,\cdots,n)$，单位矩阵 $\boldsymbol{I}$ 可以表示成

$$\boldsymbol{I}=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]$$

交换单位矩阵 $\boldsymbol{I}$ 的 i,j 两行所得矩阵记为 $\boldsymbol{P}_{ij}$

$$\boldsymbol{P}_{ij}=\begin{array}{c@{}l}\left[\begin{matrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 &&&& 1 &&& \\
&&&& 1 &&&&&& \\
&&&&& \ddots &&&&& \\
&&&&&& 1 &&&& \\
&&& 1 &&&& 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{matrix}\right] & \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \end{matrix}\end{array}$$

用 P_{ij} 左乘矩阵 $\boldsymbol{A}$，其作用就是交换 $\boldsymbol{A}$ 的第 i 行和第 j 行；用 $\boldsymbol{P}_{ij}$ 右乘矩阵 $\boldsymbol{A}$，其作用就是交换 $\boldsymbol{A}$ 的第 i 列和第 j 列.

(2) 矩阵的第 $i\,(i=2,3,\cdots,n)$ 行减去 m_{i1} 乘第 1 行的变换

这种变换相当于用矩阵

$$\boldsymbol{M}_1=\begin{bmatrix}
1 &&& \\
-m_{21} & 1 && \\
\vdots & \vdots & \ddots & \\
-m_{n1} & 0 & \cdots & 1
\end{bmatrix}$$

左乘矩阵 $\boldsymbol{A}$.

高斯消去法的消元过程,从矩阵运算的角度来看,从(5.5)式化为(5.7)式相当于系数矩阵和左端项左乘一个初等变换矩阵 $\boldsymbol{M}_1$,即

$$\boldsymbol{M}_1\boldsymbol{A}^{(1)}\boldsymbol{X}=\boldsymbol{M}_1\boldsymbol{b}^{(1)}$$

$\boldsymbol{A}^{(2)}=\boldsymbol{M}_1\boldsymbol{A}^{(1)}$,$\boldsymbol{b}^{(2)}=\boldsymbol{M}_1\boldsymbol{b}^{(1)}$,类似地,记

$$\boldsymbol{M}_i=\begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -m_{i+1,i} & 1 & & \\ & & \vdots & & \ddots & \\ & & -m_{n,i} & & & 1 \end{bmatrix}$$

则 $\boldsymbol{A}^{(i+1)}=\boldsymbol{M}_i\boldsymbol{A}^{(i)}$,$\boldsymbol{b}^{(i+1)}=\boldsymbol{M}_i\boldsymbol{b}^{(i)}$.

所以,$n-1$ 步以后得到的上三角矩阵 $\boldsymbol{A}^{(n)}$、右端项 $\boldsymbol{b}^{(n)}$ 和 $\boldsymbol{A}^{(1)}$、$\boldsymbol{b}^{(1)}$ 有下列关系式

$$\begin{cases} \boldsymbol{A}^{(n)}=\boldsymbol{M}_{n-1}\cdot\boldsymbol{M}_{n-2}\cdot\cdots\cdot\boldsymbol{M}_1\boldsymbol{A}^{(1)} \\ \boldsymbol{b}^{(n)}=\boldsymbol{M}_{n-1}\cdot\boldsymbol{M}_{n-2}\cdot\cdots\cdot\boldsymbol{M}_1\boldsymbol{b}^{(1)} \end{cases}$$

5.1.2 列主元高斯消去法

主元高斯消去法根据主元选取范围的不同,又分为列主元消去法和全主元消去法.主元高斯消去法是为了控制舍入误差而提出来的一种算法.在顺序高斯消去法的消元过程中,若出现 $a_{kk}^{(k)}=0$,则消元无法进行;即使 $a_{kk}^{(k)}\neq0$,但其值很小,用它作为除数也会导致其他元素量级的巨大增长和舍入误差的扩散,引起解的失真.先看下面的例子.

例 5-1 解方程组

$$\begin{cases} 0.003\ 000x_1+59.14x_2=59.17 \\ 5.291x_1-6.130x_2=46.78 \end{cases}$$

该方程组精确解为 $x_1^*=10.00$,$x_2^*=1.000$.

解 利用顺序高斯消去法取 4 位有效数字进行运算得

$$m_{21}=\frac{5.291}{0.003\ 000}=1\ 763.666\approx1\ 764$$

将第 1 个方程的 $-1\ 764$ 倍加到第 2 个方程,得

$$\begin{cases} 0.003\ 000x_1+59.14x_2=59.17 \\ -104\ 300x_2=-104\ 400 \end{cases}$$

通过回代得 $x_2=1.001$,$x_1=-10.00$.这与准确解相差很大.这是由于用0.003 000

作除数引起舍入误差造成的.同样对上述例题,可以交换一下方程的位置(由线性代数理论可知方程组的解并不发生改变),再利用顺序高斯消去法求解.

例 5－2 解方程组

$$\begin{cases} 5.291x_1 - 6.130x_2 = 46.78 \\ 0.003\ 000x_1 + 59.14x_2 = 59.17 \end{cases}$$

解 利用顺序高斯消去法取 4 位有效数字进行运算得

$$m_{21} = \frac{0.003\ 000}{5.291} = 0.000\ 567\ 0$$

将第 1 个方程的 $-0.000\ 576\ 0$ 倍加到第 2 个方程,得

$$\begin{cases} 5.291x_1 - 6.130x_2 = 46.78 \\ 59.14x_2 = 59.14 \end{cases}$$

通过回代,得 $x_1 = 10.00$, $x_2 = 1.000$.

小主元是不稳定的根源。因此,为使这种不稳定现象发生的可能性减至最少,必须在每次消元前选择一个相对较大的系数作为主元.

5.1.2.1 列主元高斯消去法计算过程

第一步:在系数矩阵 $\boldsymbol{A}^{(1)}$ 的第 1 列元素中选出绝对值最大的元素 $a_{i_1,1}^{(1)}$,称之为第 1 列的主元,即 $|a_{i_1,1}^{(1)}| = \max\limits_{1 \leqslant i \leqslant n} |a_{i1}^{(1)}|$. 如果 $i_1 \neq 1$,则交换第 1 行和第 i_1 行元素及相应的右端项,交换完后再进行消元得

$$\boldsymbol{A}^{(2)}\boldsymbol{X} = \boldsymbol{b}^{(2)} \tag{5.12}$$

第二步:在系数矩阵 $\boldsymbol{A}^{(2)}$ 第 2 列 $a_{22}^{(2)}$ 以下 $n-1$ 个元素中选主元 $a_{i_2,2}^{(2)}$,即 $|a_{i_2,2}^{(2)}| = \max\limits_{2 \leqslant i \leqslant n} |a_{i2}^{(2)}|$. 如果 $i_2 \neq 2$,则对调矩阵(5.12)中的第 2 行与第 i_2 行元素及相应的右端项,交换后再进行消元得

$$\boldsymbol{A}^{(3)}\boldsymbol{X} = \boldsymbol{b}^{(3)}$$

一般地,第 k 步求 i_k 使 $|a_{i_k,k}^{(k)}| = \max\limits_{k \leqslant i \leqslant n} |a_{ik}^{(k)}|$. 如果 $i_k \neq k$,则交换第 k 行与第 i_k 行及相应的右端项,经消元后得

$$\boldsymbol{A}^{(k+1)}\boldsymbol{X} = \boldsymbol{b}^{(k+1)}$$

列主元消去法除了每步需要按列选主元并可能进行矩阵行交换外,其消元过程与顺序高斯消去法的过程一样,归纳以上过程有以下算法.

5.1.2.2 列主元高斯消去算法

①输入 $\boldsymbol{A}$、$\boldsymbol{b}$ 及 ε.

②选主元及消元:

for $k=1$ to $n-1$ do

选主元: $T=\left|a_{i_k,k}\right|=\max\limits_{k\leqslant i\leqslant n}\left|a_{ik}\right|$

若 $T<\varepsilon$,则打印"求解失败",停机;否则

若 $i_k\neq k$,则交换 $\boldsymbol{A}$ 的第 i_k 行与第 k 行,交换 b_{i_k} 与 b_k;否则

消元:for $i=k+1$ to n do

$T=\dfrac{a_{ik}}{a_{kk}}$

$b_i=b_i-T\times b_k$

for $j=k+1$ to n do

$a_{ij}=a_{ij}-T\times a_{kj}$

③回代:

若 $\left|a_{nn}\right|\leqslant\varepsilon$,则打印"求解失败",停机;否则

$x_n=\dfrac{b_n}{a_{nn}}$

for $i=n-1$ downto 1 do

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

④打印 $x_i(i=1,2,\cdots,n)$

5.1.2.3 列主元高斯消去法的计算量

列主元高斯消去法的乘除及加减次数与顺序高斯消去法相同,增加了选主元时 $n(n-1)$ 次比较以及交换方程次序所需时间.

5.1.2.4 列主元高斯消去法的矩阵解释

记 $\left|a_{i_1,1}^{(1)}\right|=\max\limits_{1\leqslant i\leqslant n}\left|a_{i1}^{(1)}\right|$,则说明第1个主元在第 i_1 行,交换第1行与第 i_1 行(相当于用 $\boldsymbol{P}_{1,i_1}$ 左乘矩阵 $\boldsymbol{A}^{(1)}$),然后再进行消元,用矩阵表示即为

$$\boldsymbol{A}^{(2)}=\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\qquad \boldsymbol{b}^{(2)}=\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)}$$

同样道理,有

$$\boldsymbol{A}^{(n)}=\boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\boldsymbol{A}^{(n-1)}\qquad \boldsymbol{b}^{(n)}=\boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\boldsymbol{b}^{(n-1)}$$

由此得到等价的三角形方程组为 $\boldsymbol{A}^{(n)}\boldsymbol{X}=\boldsymbol{b}^{(n)}$. 这时

$$\begin{cases}\boldsymbol{A}^{(n)}=\boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\\ \boldsymbol{b}^{(n)}=\boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)}\end{cases}$$

例 5－3 解方程组

$$\begin{cases} 12x_1 - 3x_2 + 3x_3 = 15 \\ -18x_1 + 3x_2 - x_3 = -15 \\ x_1 + x_2 + x_3 = 6 \end{cases}$$

(它的准确解为 $x_1=1, x_2=2, x_3=3.$)

解 第 1 步:首先选取第 1 列绝对值最大的作主元. 对于本方程组,绝对值最大的主元为 -18,交换第 1 行和第 2 行得

$$\begin{cases} -18x_1 + 3x_2 - x_3 = -15 \\ 12x_1 - 3x_2 + 3x_3 = 15 \\ x_1 + x_2 + x_3 = 6 \end{cases}$$

将第 1 个方程的$\frac{2}{3}$倍加到第 2 个方程,第 1 个方程的$\frac{1}{18}$倍加到第 3 个方程得

$$\begin{cases} -18x_1 + 3x_2 - x_3 = -15 \\ -x_2 + \frac{7}{3}x_3 = 5 \\ \frac{7}{6}x_2 + \frac{17}{18}x_3 = \frac{31}{6} \end{cases}$$

第 2 步:选取第 2 列(第 2 个元素以下)绝对值最大的主元,对于本方程组,绝对值最大的主元为$\frac{7}{6}$,交换第 2 行和第 3 行,消元后得

$$\begin{cases} -18x_1 + 3x_2 - x_3 = -15 \\ \frac{7}{6}x_2 + \frac{17}{18}x_3 = \frac{31}{6} \\ \frac{22}{7}x_3 = \frac{66}{7} \end{cases}$$

把上面方程进行回代,就可逐步解出 $x_1=1, x_2=2, x_3=3.$

5.1.3 全主元高斯消去法

全主元高斯消去法与列主元高斯消去法类似,所不同的是选取主元的范围不同. 全主元高斯消去法在整个系数矩阵中找绝对值最大的元素作为主元,以控制舍入误差的增长,将舍入误差控制在一个最小的范围. 但找主元和交换行列次序要花费大量的机器时间. 因此,有时不用全主元高斯消去法而用列主元高斯消去法.

5.1.3.1 全主元高斯消去法计算过程

第 1 步:求 i_1, j_1 使

$$|a_{i_1,j_1}^{(1)}| = \max_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant n}} |a_{ij}^{(1)}|$$

如果 $i_1 \neq 1$，则交换第 1 行和第 i_1 行及相应的右端项（$\boldsymbol{P}_{1,i_1}$左乘 $\boldsymbol{A}^{(1)}$）；若 $j_1 \neq 1$，则交换第 1 列和第 j_1 列（$\boldsymbol{P}_{1,j_1}$右乘 $\boldsymbol{A}^{(1)}$），并交换 x_1 与 x_{j_1}. 交换完后再进行消元得

$$\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\boldsymbol{P}_{1,j_1}\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)} \tag{5.13}$$

记 $\boldsymbol{A}^{(2)} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\boldsymbol{P}_{1,j_1}$，$\boldsymbol{b}^{(2)} = \boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{b}^{(1)}$. 上式变为

$$\boldsymbol{A}^{(2)}\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{b}^{(2)}$$

一般地，第 k 步（$k = 1,2,\cdots,n-1$），求 i_k，j_k 使

$$|a_{i_k,j_k}^{(k)}| = \max_{\substack{k\leqslant i\leqslant n\\ k\leqslant j\leqslant n}} |a_{ij}^{(k)}|$$

如果 $i_k \neq k$，则交换 $\boldsymbol{A}^{(k)}$ 的第 k 行与第 i_k 行（$\boldsymbol{P}_{k,i_k}$左乘 $\boldsymbol{A}^{(k)}$）及相应的右端项；若 $j_k \neq k$，则交换 $\boldsymbol{A}^{(k)}$ 的第 k 列和第 j_k 列（$\boldsymbol{P}_{k,j_k}$右乘 $\boldsymbol{A}^{(k)}$），并交换 x_k 与 x_{j_k}. 交换完后进行消元得

$$\boldsymbol{M}_k\boldsymbol{P}_{k,i_k}\boldsymbol{A}^{(k)}\boldsymbol{P}_{k,j_k}\boldsymbol{P}_{k,j_k}^{-1}\cdots\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{M}_k\boldsymbol{P}_{k,i_k}\boldsymbol{b}^{(k)}$$

记 $\boldsymbol{A}^{(k+1)} = \boldsymbol{M}_k\boldsymbol{P}_{k,i_k}\boldsymbol{A}^{(k)}\boldsymbol{P}_{k,j_k}$，$\boldsymbol{b}^{(k+1)} = \boldsymbol{M}_k\boldsymbol{P}_{k,i_k}\boldsymbol{b}^{(k)}$. 上式变为

$$\boldsymbol{A}^{(k+1)}\boldsymbol{P}_{k,j_k}^{-1}\cdots\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{b}^{(k+1)}$$

归纳上述过程有以下算法：

5.1.3.2　全主元高斯消去法算法

①输入 $\boldsymbol{A}$、$\boldsymbol{b}$ 及 ε.

②for　$i=1$　to　n　do　$d_i=i$　//记录未知量位置的变化//

③选主元消元:

　　for　$k=1$　to　$n-1$　do

　　　求 i_k,j_k 使 $T=\left|a_{i_k,j_k}^{(k)}\right|=\max\limits_{\substack{k\leqslant i\leqslant n\\ k\leqslant j\leqslant n}}\left|a_{ij}^{(k)}\right|$

　　　　若 $T<\varepsilon$,则打印"求解失败",停机;否则

　　　　若 $i_k\neq k$,则交换 $\boldsymbol{A}$ 的第 i_k 行与第 k 行,交换 b_{i_k} 与 b_k

　　　　若 $j_k\neq k$,则交换 $\boldsymbol{A}$ 的第 j_k 列与第 k 列,交换 d_{j_k} 与 d_k

　　for　$i=k+1$　To　n　do

　　　$T=\dfrac{a_{ik}}{a_{kk}}$

　　　$b_i=b_i-T\times b_k$

　　for　$j=k+1$　to　n　do

　　　$a_{ij}=a_{ij}-T\times a_{kj}$

④回代:

　　$Z_n=\dfrac{b_n}{a_{nn}}$

　　for　$i=n-1$　downto　1　do

$$Z_i=\frac{b_i-\sum\limits_{j=i+1}^{n}a_{ij}Z_j}{a_{ii}}$$

⑤for　$j=1$　to　n　do

　　$x_{d_j}=Z_j$

⑥打印 x_i,$i=1,2,\cdots,n$

5.1.3.3　全主元高斯消去法的计算量

全主元高斯消去法乘除及加减次数与顺序高斯消去法相同,增加了选主元时

$$\sum_{i=1}^{n}(i^2-1)=\frac{n^3}{6}-\frac{n}{6}$$

次比较以及交换方程次序所需的时间.

5.1.3.4　全主元高斯消去法的矩阵解释

记 $\left|a_{i_1,j_1}^{(1)}\right|=\max\limits_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant n}}\left|a_{ij}^{(1)}\right|$,由上述过程知,主元在第 i_1 行、第 j_1 列,交换第 1 行与第 i_1 行、第 1 列与第 j_1 列,然后再消元即可得到 $\boldsymbol{A}^{(2)}$,用矩阵表示即为

$$\begin{cases} \boldsymbol{A}^{(2)} = \boldsymbol{M}_1 \boldsymbol{P}_{1,i_1} \boldsymbol{A}^{(1)} \boldsymbol{P}_{1,j_1} \\ \boldsymbol{b}^{(2)} = \boldsymbol{M}_1 \boldsymbol{P}_{1,i_1} \boldsymbol{b}^{(1)} \end{cases}$$

同理,得

$$\begin{cases} \boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1} \boldsymbol{P}_{n-1,i_{n-1}} \boldsymbol{A}^{(n-1)} \boldsymbol{P}_{n-1,j_{n-1}} \\ \boldsymbol{b}^{(n)} = \boldsymbol{M}_{n-1} \boldsymbol{P}_{n-1,i_{n-1}} \boldsymbol{b}^{(n-1)} \end{cases}$$

注意到 $\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{b}^{(1)}$,两边同乘 $\boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}$,得

$$\boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{b}^{(n)} \tag{5.14}$$

由于 $\boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{P}_{n-1,i_{n-1}}\cdots\boldsymbol{M}_1\boldsymbol{P}_{1,i_1}\boldsymbol{A}^{(1)}\boldsymbol{P}_{1,j_1}\cdots\boldsymbol{P}_{n-1,j_{n-1}}$,由(5.14)式知

$$\boldsymbol{A}^{(n)}\boldsymbol{P}_{n-1,j_{n-1}}^{-1}\cdots\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} = \boldsymbol{b}^{(n)} \tag{5.15}$$

记

$$\boldsymbol{Z} = \boldsymbol{P}_{n-1,j_{n-1}}^{-1}\cdots\boldsymbol{P}_{1,j_1}^{-1}\boldsymbol{X} \tag{5.16}$$

则

$$\boldsymbol{A}^{(n)}\boldsymbol{Z} = \boldsymbol{b}^{(n)} \tag{5.17}$$

通过(5.17)式可求得 $\boldsymbol{Z}$. 由(5.16)式最终求得

$$\boldsymbol{X} = \boldsymbol{P}_{1,j_1}\boldsymbol{P}_{2,j_2}\cdots\boldsymbol{P}_{n-1,j_{n-1}}Z$$

5.2 LU 分解法

前面介绍的顺序高斯消去法可以把方程组 $\boldsymbol{A}^{(1)}\boldsymbol{X} = \boldsymbol{b}^{(1)}$ 等价转化为一个上三角形方程组 $\boldsymbol{A}^{(n)}\boldsymbol{X} = \boldsymbol{b}^{(n)}$. 而从消元过程可以看到,从 $\boldsymbol{A}^{(1)}$ 转化到 $\boldsymbol{A}^{(n)}$ 的过程实际是进行若干次初等变换的过程,由顺序高斯消去法的矩阵解释得到

$$\boldsymbol{A}^{(n)} = \boldsymbol{M}_{n-1}\boldsymbol{M}_{n-2}\cdots\boldsymbol{M}_1\boldsymbol{A}^{(1)}$$

从而

$$\boldsymbol{A}^{(1)} = \boldsymbol{M}_1^{-1}\boldsymbol{M}_2^{-1}\cdots\boldsymbol{M}_{n-1}^{-1}\boldsymbol{A}^{(n)}$$

注意到 $\boldsymbol{M}_i^{-1}(i = 1, 2, \cdots, n-1)$ 均为下三角矩阵且对角线元素为 1,因此,$\boldsymbol{M}_1^{-1}\boldsymbol{M}_2^{-1}\cdots\boldsymbol{M}_{n-1}^{-1}$ 为对角线元素为 1 的下三角矩阵. 记

$$\boldsymbol{L} = \boldsymbol{M}_1^{-1}\boldsymbol{M}_2^{-1}\cdots\boldsymbol{M}_{n-1}^{-1}$$

$$\boldsymbol{U} = \boldsymbol{A}^{(n)}$$

由顺序高斯消去法看到,只要方程

$$\boldsymbol{AX} = \boldsymbol{b}$$

的系数矩阵 $\boldsymbol{A}$ 的所有顺序主子式不为零,则 $\boldsymbol{A}$ 一定可以分解成

$$\boldsymbol{A} = \boldsymbol{LU}$$

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

$\boldsymbol{L}$ 为对角线元素为 1 的单位下三角矩阵；$\boldsymbol{U}$ 为上三角矩阵.

定理 5.1 如果 n 阶矩阵 $\boldsymbol{A}$ 的所有顺序主子式不为零，则 $\boldsymbol{A}$ 有惟一的 LU 分解.

证明 设

$$\boldsymbol{A}=\boldsymbol{LU}=\boldsymbol{L}_1\boldsymbol{U}_1$$

其中，$\boldsymbol{L},\boldsymbol{L}_1$ 为单位下三角矩阵；$\boldsymbol{U},\boldsymbol{U}_1$ 为上三角矩阵. 因为 $\boldsymbol{A}$ 非奇异，故 $\boldsymbol{L}_1^{-1}\boldsymbol{U}_1^{-1}$ 存在. 于是有

$$\boldsymbol{L}^{-1}\boldsymbol{L}_1=\boldsymbol{U}\boldsymbol{U}_1^{-1}$$

上式左端为单位下三角矩阵，右端为上三角矩阵，从而上式两边均为单位矩阵，从而证得 $\boldsymbol{U}=\boldsymbol{U}_1$，$\boldsymbol{L}=\boldsymbol{L}_1$. $\boldsymbol{A}$ 的三角形矩阵分解的惟一性获证.

如果能实现这种分解，则求解方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 就相当简单. 注意到 $\boldsymbol{A}=\boldsymbol{LU}$，因此 $\boldsymbol{LUX}=\boldsymbol{b}$. 记 $\boldsymbol{UX}=\boldsymbol{Y}$，则

$$\begin{cases}\boldsymbol{LY}=\boldsymbol{b}\\ \boldsymbol{UX}=\boldsymbol{Y}\end{cases}$$

通过 $\boldsymbol{LY}=\boldsymbol{b}$，得

$$\begin{cases}y_1=b_1\\ y_i=b_i-\sum\limits_{j=1}^{i-1}l_{ij}y_j \quad (i=2,\cdots,n)\end{cases}$$

通过 $\boldsymbol{UX}=\boldsymbol{Y}$，最终求得

$$\begin{cases}x_n=\dfrac{y_n}{u_{nn}}\\ x_i=\left[y_i-\sum\limits_{j=i+1}^{n}u_{ij}x_j\right]\Big/u_{ii} \quad (i=n-1,\cdots,1)\end{cases}$$

5.2.1 直接 LU 分解法

下面详细讨论矩阵 $\boldsymbol{A}$ 的 LU 分解. 设 $\boldsymbol{A}=\boldsymbol{LU}$，即

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

两个矩阵相等，则其对应元素相等. 利用矩阵乘法，由矩阵的第 1 行对应相等得

$$u_{1j}=a_{1j}\quad (j=1,2,\cdots,n)$$

由矩阵的第 1 列对应相等得

$$l_{i1}=\frac{a_{i1}}{u_{11}}\quad (i=2,3,\cdots,n)$$

这就求出了 $\boldsymbol{U}$ 矩阵的第 1 行和 $\boldsymbol{L}$ 矩阵的第 1 列元素.

一般地,设 $\boldsymbol{U}$ 矩阵的前 $k-1$ 行和 $\boldsymbol{L}$ 矩阵的前 $k-1$ 列已经求出,则由

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

得

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

又由

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

得

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

综上所述,$\boldsymbol{A}$ 的 LU 分解公式如下

$$\begin{cases} u_{kj}=a_{kj}-\sum\limits_{m=1}^{k-1}l_{km}u_{mj} & (j=k,k+1,\cdots,n) \\ l_{ik}=\dfrac{a_{ik}-\sum\limits_{m=1}^{k-1}l_{im}u_{mk}}{u_{kk}} & (i=k+1,\cdots,n) \end{cases}$$

其中,$k=1,2,\cdots,n$.

上述计算公式有如下特点:$\boldsymbol{U}$ 的元素按行逐行求,$\boldsymbol{L}$ 的元素按列逐列求;先求 $\boldsymbol{U}$ 的第 k 行元素,然后求 $\boldsymbol{L}$ 的第 k 列元素,$\boldsymbol{U}$ 和 $\boldsymbol{L}$ 的元素一行一列交叉进行,如图 5-1 所示.

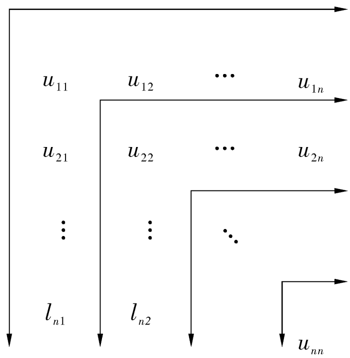

图 5－1

5.2.1.1　直接 LU 分解法算法

①输入 $\boldsymbol{A},\boldsymbol{b}$ 及 ε.

②对 $k=1,2,\cdots,n$ 循环

　　对 $j=k,k+1,\cdots,n$ 循环

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

　　如果 $|u_{kk}|<\varepsilon$，则打印"求解失败"，停机；否则

　　　对 $i=k+1,\cdots,n$，计算

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

③求解 $\boldsymbol{LY}=\boldsymbol{b}$

$$y_1=b_1$$

　　对　$i=2,3,\cdots,k$，计算

$$y_i=b_i-\sum_{j=1}^{i-1}l_{ij}y_j$$

④求解 $\boldsymbol{UX}=\boldsymbol{Y}$

$$x_n=y_n/u_{nn}$$

　　对 $i=n-1,n-2,\cdots,1$，计算

$$x_i=\frac{y_i-\sum_{j=i+1}^{n}u_{ij}x_j}{u_{ii}}$$

⑤打印 $x_i(i=1,2,\cdots,n)$

每次计算结果 l_{ik} 和 u_{kj} 仍可存放在矩阵 $\boldsymbol{A}$ 的相应元素 a_{ik} 和 a_{kj} 所占的单元内，不必占用新的单元，只要将上述算法中出现的 l_{ik} 和 u_{kj} 相应变换为 a_{ik} 和 a_{kj} 即可.

5.2.1.2 直接 LU 分解法的计算量(乘除次数)

LU 分解计算 $\boldsymbol{U}$ 的计算量：

$$\sum_{i=1}^{n}(n-i)\cdot i=\frac{n^3-n}{6}$$

计算 $\boldsymbol{L}$ 的计算量：

$$\sum_{i=1}^{n-1}(n-i)\cdot i=\frac{n^3-n}{6}$$

求解 $\boldsymbol{LY}=\boldsymbol{b}$ 的计算量：

$$\sum_{i=1}^{n-1} i=\frac{n(n-1)}{2}$$

求解 $\boldsymbol{UX}=\boldsymbol{Y}$ 的计算量：

$$\sum_{i=1}^{n} i=\frac{n(n+1)}{2}$$

总的计算量为：

$$\frac{n^3+3n^2-n}{3}$$

例 5-4 用 LU 分解法求解方程组

$$\begin{bmatrix}1 & 2 & 1\\ 2 & 2 & 3\\ -1 & -3 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 3\\ 2\end{bmatrix}$$

解 先把系数矩阵进行 LU 分解得

$$\boldsymbol{L}=\begin{bmatrix}1 & & \\ 2 & 1 & \\ -1 & 0.5 & 1\end{bmatrix}\qquad \boldsymbol{U}=\begin{bmatrix}1 & 2 & 1\\ & -2 & 1\\ & & 0.5\end{bmatrix}$$

求解 $\boldsymbol{LY}=\boldsymbol{b}$：

$$\begin{bmatrix}1 & & \\ 2 & 1 & \\ -1 & 0.5 & 1\end{bmatrix}\begin{bmatrix}y_1\\ y_2\\ y_3\end{bmatrix}=\begin{bmatrix}0\\ 3\\ 2\end{bmatrix}$$

得 $y_1=0, y_2=3, y_3=0.5$.

求解 $\boldsymbol{UX}=\boldsymbol{Y}$：

$$\begin{bmatrix}1 & 2 & 1\\ & -2 & 1\\ & & 0.5\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 3\\ 0.5\end{bmatrix}$$

解得 $x_1=1, x_2=-1, x_3=1$. 故 $\boldsymbol{AX}=\boldsymbol{b}$ 的解为 $\boldsymbol{X}=(1,-1,1)^{\mathrm{T}}$.

用直接 LU 分解法求解方程组所需要的计算量仍为$\frac{n^3+3n^2-n}{3}$,和高斯消去法

乘除法次数相同. 但 LU 分解法把对系数矩阵的计算和对右端项的计算分开了,这就使得求解系数矩阵相同而右端项不同的一系列方程组变得特别方便.

5.2.2 列主元 LU 分解法

直接进行 LU 分解时,u_{kk}可能为零,或者绝对值接近零,这样会出现溢出,或者因以一个小的数作除数,引起舍入误差的积累,最后引起解的失真. 为了避免这些问题,如果 $\boldsymbol{A}$ 非奇异,可以采用列主元的方法,叫做列主元 LU 分解.

假若第 $k-1$ 步的分解已完成,在进行第 k 步分解时,为避免出现小的 u_{kk}作除数,计算

$$S_i = a_{ik} - \sum_{m=1}^{k-1} l_{im} u_{mk}$$

选取行号 i_k,使

$$|S_{i_k}| = \max_{k \leqslant i \leqslant n} |S_i|$$

若 $i_k \neq k$,则对调 $\boldsymbol{A}$ 的第 k 行与第 i_k 行,再求 $u_{kj}(j=k,\cdots,n)$和 $l_{ik}(i=k+1,\cdots,n)$,从而完成了 $\boldsymbol{U}$ 的第 k 行和 $\boldsymbol{L}$ 的第 k 列的计算. 算法如下.

列主元 LU 分解算法

①输入 $\boldsymbol{A},\boldsymbol{b}$ 及 ε

②对 $k=1,2,\cdots,n$ 循环

计算 $S_i = a_{ik} - \sum\limits_{m=1}^{k-1} l_{im} u_{mk} \quad (i=k,k+1,\cdots,n)$

选主元 $|S_{i_k}| = \max\limits_{k \leqslant i \leqslant n} |S_i|$,并记录 i_k

若 $|S_{i_k}| < \varepsilon$,则打印"求解失败",停机;否则

交换 $\boldsymbol{A}$ 的第 k 行与第 i_k 行元素

计算 $\boldsymbol{U}$ 的第 k 行元素和 $\boldsymbol{L}$ 的第 k 列元素

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

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

③将 $\boldsymbol{b}$ 向量的第 k 个元素与第 i_k 个元素交换,$k=1,2,\cdots,n-1$.

④求解 $\boldsymbol{LY}=\boldsymbol{b},\boldsymbol{UX}=\boldsymbol{Y}$(过程与直接 LU 分解法相同)

5.3 对称正定矩阵的平方根法和 LDL$^{\mathrm{T}}$ 分解法

当 $\boldsymbol{A}$ 是对称正定矩阵时,存在一个实的非奇异下三角矩阵 $\boldsymbol{L}_1$,使

$$\boldsymbol{A}=\boldsymbol{L}_1\boldsymbol{L}_1^{\mathrm{T}}$$

且当限定 $\boldsymbol{L}_1$ 的对角线为正时,这种分解惟一. 这种分解称为矩阵的 Cholesky 分解.

定理 5.2 设 $\boldsymbol{A}$ 是对称正定矩阵,则 $\boldsymbol{A}$ 有如下分解:

$$\boldsymbol{A}=\boldsymbol{LDL}^{\mathrm{T}}$$

其中,$\boldsymbol{L}$ 是单位下三角阵;$\boldsymbol{D}$ 为对角阵,且这种分解是惟一的.

证明 因为 $\boldsymbol{A}$ 对称正定,从而 $\boldsymbol{A}$ 有惟一的 LU 分解

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

因为 $\Delta_k=u_{11}\cdots u_{kk}\neq0(k=1,2,\cdots,n)$,故上式可化为

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

将右端三角矩阵分别记为 $\boldsymbol{L}$,$\boldsymbol{D}$ 和 $\boldsymbol{R}$,因 $\boldsymbol{A}$ 对称,有

$$\boldsymbol{A}=\boldsymbol{LDR}=\boldsymbol{R}^{\mathrm{T}}\boldsymbol{DL}^{\mathrm{T}}$$

或者

$$\boldsymbol{A}=\boldsymbol{LDR}=\boldsymbol{R}^{\mathrm{T}}(\boldsymbol{DL}^{\mathrm{T}})$$

因为 $\boldsymbol{A}$ 的 LU 分解惟一,故 $\boldsymbol{L}=\boldsymbol{R}^{\mathrm{T}}$,$\boldsymbol{L}^{\mathrm{T}}=\boldsymbol{R}$,$\boldsymbol{A}=\boldsymbol{LDL}^{\mathrm{T}}$,此分解显然是惟一的.

定理 5.3 n 阶对称正定矩阵 $\boldsymbol{A}$ 一定有 Cholesky 分解 $\boldsymbol{A}=\boldsymbol{L}_1\boldsymbol{L}_1^{\mathrm{T}}$,当限定 $\boldsymbol{L}_1$ 的对角线为正时,矩阵的 Cholesky 分解惟一.

证明 由定理 5.2 知,存在惟一的单位下三角矩阵 $\boldsymbol{L}$,对角元全为正的对角阵

$$\boldsymbol{D}=\mathrm{diag}(d_1,d_2,\cdots,d_n)$$

使

$$\boldsymbol{A}=\boldsymbol{LD}^{\frac{1}{2}}(\boldsymbol{LD}^{\frac{1}{2}})^{\mathrm{T}}$$

又记 $\boldsymbol{L}_1=\boldsymbol{LD}^{\frac{1}{2}}$,它是一个非奇异下三角阵,一对角线元素为正,故 $\boldsymbol{A}$ 有 Cholesky 分解

$$\boldsymbol{A}=\boldsymbol{L}_1\boldsymbol{L}_1^{\mathrm{T}}$$

分解的惟一性显而易见。

下边推导 Cholesky 分解的计算公式

设 $\boldsymbol{A}=\boldsymbol{LL}^{\mathrm{T}}$,即

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

其中 $a_{ij}=a_{ji}(i,j=1,2,\cdots,n)$，$l_{ii}>0(i=1,2,\cdots,n)$.

第 1 步：由矩阵乘法有 $a_{11}=l_{11}^2$，且 $a_{i1}=l_{i1}l_{11}$，故求得

$$l_{11}=\sqrt{a_{11}} \qquad l_{i1}=\frac{a_{i1}}{l_{11}} \quad (i=2,3,\cdots,n)$$

一般地，设 $\boldsymbol{L}$ 矩阵的前 $k-1$ 列元素已求出，则

第 k 步：由矩阵乘法得

$$\sum_{m=1}^{k-1} l_{km}^2 + l_{kk}^2 = a_{kk} \qquad \sum_{m=1}^{k-1} l_{im}l_{km} + l_{ik}l_{kk} = a_{ik}$$

于是

$$\begin{cases} l_{kk} = \sqrt{a_{kk} - \sum\limits_{m=1}^{k-1} l_{im}^2} \\ l_{ik} = \left(a_{ik} - \sum\limits_{m=1}^{k-1} l_{im}l_{km} \right) \Big/ l_{kk} \end{cases} \quad (i=k+1,\cdots,n;k=2,3,\cdots,n) \tag{5.18}$$

由于分解公式(5.18)中的每一步都有开方运算，故又称 Cholesky 方法为平方根法.

注意到 $a_{kk}=\sum\limits_{m=1}^{k-1} l_{km}^2$，所以 $l_{km}^2 \leqslant a_{kk} \leqslant \max\limits_{1\leqslant k\leqslant n} a_{kk}$.

这说明在分解过程中元素 l_{km} 的平方不会超过 $\boldsymbol{A}$ 的最大对角元，因而舍入误差的放大受到限制. 所以，平方根法求解对称正定方程组时可以不考虑选主元的问题.

可以证明，若用顺序高斯消去法求解对称正定方程组 $\boldsymbol{AX}=\boldsymbol{b}$，则有

$$\max_{1\leqslant i,j\leqslant n} \left| a_{ij}^{(k)} \right| \leqslant \max_{1\leqslant i\leqslant n} a_{ii} \quad (k=1,2,\cdots,n)$$

其中 $a_{ij}^{(k)}$ 是第 k 步顺序高斯消去法过程所得到的元素. 这说明高斯消去法求解对称正定方程组也可以不用选主元.

从运算的角度看，平方根法是有利的，用平方根法求解 $\boldsymbol{AX}=\boldsymbol{b}$ 所需的乘除次数为$\frac{1}{6}(n^3+9n^2+2n)$，另外还有 n 次开平方运算，乘除次数只是高斯消去法的一半左右。

例 5－5　用平方根法解方程组

$$\begin{bmatrix} 3 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 12 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 7 \end{bmatrix}$$

解 由平方根分解法可得

$$\begin{bmatrix} 3 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 12 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & & \\ \frac{2}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & \\ \sqrt{3} & -\sqrt{6} & \sqrt{3} \end{bmatrix}\begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \sqrt{3} \\ & \frac{\sqrt{2}}{\sqrt{3}} & -\sqrt{6} \\ & & \sqrt{3} \end{bmatrix}$$

求解
$$\begin{bmatrix} \sqrt{3} & & \\ \frac{2}{\sqrt{3}} & \frac{\sqrt{2}}{\sqrt{3}} & \\ \sqrt{3} & -\sqrt{6} & \sqrt{3} \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 7 \end{bmatrix}$$

得 $\boldsymbol{Y} = (5/\sqrt{3}, -1/\sqrt{6}, 1/\sqrt{3})^{\mathrm{T}}$.

求解
$$\begin{bmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \sqrt{3} \\ & \frac{\sqrt{2}}{\sqrt{3}} & -\sqrt{6} \\ & & \sqrt{3} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

得 $\boldsymbol{X} = (1, 1/2, 1/3)^{\mathrm{T}}$.

平方根求解算法

①输入 $\boldsymbol{A}$,$\boldsymbol{b}$ 及 ε.

②对 $k = 1, 2, \cdots, n$,循环

计算 $S_k = a_{kk} - \sum\limits_{m=1}^{k-1} l_{km}^2$

若 $S_k < \varepsilon$,则打印"求解失败",停机;否则

计算

$$l_{kk} = \sqrt{a_{kk} - \sum_{m=1}^{k-1} l_{im}^2}$$

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

③求解 $\boldsymbol{LY} = \boldsymbol{b}$

$y_1 = b_1 / l_{11}$

对 $i=2,3,\cdots,n$,计算

$$y_i=\left(b_i-\sum_{j=1}^{i-1}l_{ij}y_j\right)\Big/l_{ii}$$

④求解 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{Y}$

$x_n=y_n/l_{nn}$

对 $i=n-1,n-2,\cdots,1$,计算

$$x_i=\left(y_i-\sum_{j=i+1}^{n}l_{ji}x_j\right)\Big/l_{ii}$$

⑤打印 $x_i(i=1,2,\cdots,n)$

为避免平方根法的开方运算,利用定理 5.2,也可以对 $\boldsymbol{A}$ 作 $\mathrm{LDL}^{\mathrm{T}}$ 分解,其分解方法可借助于 LU 分解法.

从定理 5.1 可以看到,当 $\boldsymbol{A}$ 的所有顺序主子式不为零且对称时,$\boldsymbol{A}$ 的 LU 分解存在且惟一,并且

$$\boldsymbol{U}=\boldsymbol{D}\boldsymbol{L}^{\mathrm{T}}$$

其中,$\boldsymbol{D}=\mathrm{diag}(u_{11},u_{22},\cdots,u_{nn})$,$l_{ij}=u_{ji}/u_{ii}$.

利用 LU 分解法先计算出 $\boldsymbol{U}$ 的第 k 行,由 $\boldsymbol{U}$ 的第 k 行得到 $\boldsymbol{D}$ 的第 k 个对角元素及 L 的第 k 列.

$$d_i=u_{ii}\qquad l_{ik}=u_{ki}/u_{kk}$$

因此 $\mathrm{LDL}^{\mathrm{T}}$ 分解的乘除法计算次数为

$$\frac{n^3}{6}-\frac{n}{6}+\frac{n^2-n}{2}=\frac{1}{6}n^3+\frac{1}{2}n^2-\frac{2}{3}n$$

大约只是高斯消去法的一半,与平方根法大致相同,但避免了开方运算.

例 5-6　用 $\mathrm{LDL}^{\mathrm{T}}$ 分解法求方程组

$$\begin{bmatrix}5&-4&1\\-4&6&-4\\1&-4&6\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}2\\-1\\1\end{bmatrix}$$

解　由 $\mathrm{LDL}^{\mathrm{T}}$ 分解法得

$$\begin{bmatrix}5&-4&1\\-4&6&-4\\1&-4&6\end{bmatrix}=\begin{bmatrix}1&&\\-\frac{4}{5}&1&\\\frac{1}{5}&-\frac{8}{7}&1\end{bmatrix}\begin{bmatrix}5&&\\&\frac{14}{5}&\\&&\frac{15}{7}\end{bmatrix}\begin{bmatrix}1&-\frac{4}{5}&\frac{1}{5}\\&1&-\frac{7}{8}\\&&1\end{bmatrix}$$

这样就得到

$$L=\begin{bmatrix}1 & & \\ -\frac{4}{5} & 1 & \\ \frac{1}{5} & -\frac{8}{7} & 1\end{bmatrix}\qquad D=\begin{bmatrix}5 & & \\ & \frac{14}{5} & \\ & & \frac{15}{7}\end{bmatrix}$$

解 $\boldsymbol{LY}=\boldsymbol{b}$ 得

$$\boldsymbol{Y}=\left(2,\frac{3}{5},-\frac{5}{7}\right)^{\mathrm{T}}.$$

解 $\boldsymbol{DZ}=\boldsymbol{Y}$ 得 $\boldsymbol{Z}=\left(\frac{2}{5},\frac{3}{14},-\frac{1}{3}\right)^{\mathrm{T}}$.

最后解 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{Z}$ 得 $\boldsymbol{X}=\left(\frac{1}{3},-\frac{1}{6},-\frac{1}{3}\right)^{\mathrm{T}}$.

LDL$^{\mathrm{T}}$ 分解算法

①输入 $\boldsymbol{A},\boldsymbol{b}$ 及 ε

②对 $k=1,2,\cdots,n$,循环

a)对 $j=k,k+1,\cdots,n$,计算

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

若 $|u_{kk}|<\varepsilon$,则打印"求解失败",停机;否则

b) $d_k=u_{kk}$

c)对 $i=k+1,\cdots,n$,计算

$$l_{ik}=\frac{u_{ki}}{u_{kk}}$$

③求解 $\boldsymbol{LY}=\boldsymbol{b}$

④求解 $\boldsymbol{DZ}=\boldsymbol{Y}$

⑤求解 $\boldsymbol{L}^{\mathrm{T}}\boldsymbol{X}=\boldsymbol{Z}$

⑥打印 $x_i(i=1,2,\cdots,n)$

为了更好地理解算法,引入 l_{ik} 和 u_{kj}. 实际上,l_{ik} 和 u_{kj} 仍将存放在矩阵 $\boldsymbol{A}$ 的相应元素 a_{ik} 和 a_{kj} 所占的单元内,不必占用新的单元. 要做到这一点,只要将上述算法中出现的 l_{ik} 和 u_{kj} 相应变为 a_{ik} 和 a_{kj} 即可.

5.4 向量与矩阵范数

用直接法求解线性方程组 $\boldsymbol{AX}=\boldsymbol{b}$ 时,由于有舍入误差,只能得到近似解. 为了进行解的误差分析和迭代法的收敛性分析(后面介绍),首先介绍向量范数和矩阵

范数的概念.

5.4.1 向量范数

向量范数是 n 维欧几里德空间中长度概念的推广.

定义 5.1 设向量 $\boldsymbol{X}\in\mathbf{R}^n$,若与 $\boldsymbol{X}$ 对应的非负实数 $\|\boldsymbol{X}\|$ 满足下面三个条件:

①非负数:$\forall\boldsymbol{X}\in\mathbf{R}^n$,有 $\|\boldsymbol{X}\|\geqslant 0$,且 $\|\boldsymbol{X}\|=0$ 当且仅当 $\boldsymbol{X}=0$;

②齐次性:$\forall\boldsymbol{X}\in\mathbf{R}^n,\alpha\in\mathbf{R}$ 有 $\|\alpha\boldsymbol{X}\|\leqslant|\alpha|\cdot\|\boldsymbol{X}\|$;

③三角不等式 $\forall\boldsymbol{X},\boldsymbol{Y}\in\mathbf{R}^n$,有 $\|\boldsymbol{X}+\boldsymbol{Y}\|\leqslant\|\boldsymbol{X}\|+\|\boldsymbol{Y}\|$.

则称 $\|\boldsymbol{X}\|$ 为向量 $\boldsymbol{X}$ 的范数.

常用范数有:

(1)向量的 1 范数

$$\|\boldsymbol{X}\|_1=\sum_{i=1}^{n}|x_i|$$

(2)向量的 2 范数

$$\|\boldsymbol{X}\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{\frac{1}{2}}$$

(3)向量的∞范数

$$\|\boldsymbol{X}\|_\infty=\max_{1\leqslant i\leqslant n}|x_i|$$

(4)更一般的 p 范数为

$$\|\boldsymbol{X}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}$$

容易验证,上面几种范数的定义满足范数的三个条件.

定义 5.2 设 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 是 $\mathbf{R}^n$ 上两种范数,如果存在与 $\boldsymbol{X}$ 无关的两个正常数 C_1 和 C_2,使不等式

$$C_1\|\boldsymbol{X}\|_\alpha\leqslant\|\boldsymbol{X}\|_\beta\leqslant C_2\|\boldsymbol{X}\|_\alpha\quad(\forall\boldsymbol{X}\in\mathbf{R}^n)$$

成立,则称范数 $\|\cdot\|_\alpha$ 和 $\|\cdot\|_\beta$ 等价.

定理 5.4 有限维空间上任何两种范数均等价.

定义 5.3 设 $\{\boldsymbol{X}^{(k)}\}$ 是 $\mathbf{R}^n$ 中的向量序列,若有向量 $\boldsymbol{X}^*\in\mathbf{R}^n$,使

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

则称 $\{\boldsymbol{X}^{(k)}\}$ 收敛于 $\boldsymbol{X}^*$,记为 $\lim\limits_{k\to\infty}\boldsymbol{X}^{(k)}=\boldsymbol{X}^*$.

该定义形式上依赖于所选择的向量范数,但由于 $\mathbf{R}^n$ 中范数等价,因此 $\{\boldsymbol{X}^{(k)}\}$ 的收敛性实际上与所选择的范数无关.

定理 5.5 设 $\{\boldsymbol{X}^{(k)}\}$ 是 $\mathbf{R}^n$ 中的向量序列,则 $\lim\limits_{k\to\infty}\|\boldsymbol{X}^{(k)}-\boldsymbol{X}^*\|=0$ 当且仅当

$$\lim x_j^{(k)}=x_j^*$$

其中,$x_j^{(k)}$ 和 x_j^* 分别表示 $\boldsymbol{X}^{(k)}$ 和 $\boldsymbol{X}^*$ 的第 j 个分量.

5.4.2 矩阵范数

定义 设矩阵 $\boldsymbol{A}\in\mathbf{R}^{n\times n}$,若 $\boldsymbol{A}$ 的非负实数 $\|\boldsymbol{A}\|$ 满足下列四个条件:

①非负数:$\forall \boldsymbol{X}\in\mathbf{R}^{n\times n}$, $\|\boldsymbol{A}\|\geqslant 0$,且 $\|\boldsymbol{A}\|=0$ 当且仅当 $\boldsymbol{A}=0$;

②齐次性:$\forall \boldsymbol{A}\in\mathbf{R}^{n\times n}$, $\forall \alpha\in\mathbf{R}$, $\|\alpha\boldsymbol{A}\|=|\alpha|\cdot\|\boldsymbol{A}\|$;

③三角不等式 $\forall \boldsymbol{A},\boldsymbol{B}\in\mathbf{R}^{n\times n}$, $\|\boldsymbol{A}+\boldsymbol{B}\|\leqslant\|\boldsymbol{A}\|+\|\boldsymbol{B}\|$;

④乘法不等式:$\forall \boldsymbol{A},\boldsymbol{B}\in\mathbf{R}^{n\times n}$, $\|\boldsymbol{AB}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{B}\|$.

则称 $\|\boldsymbol{A}\|$ 为矩阵 $\boldsymbol{A}$ 的范数.

同向量范数一样,矩阵范数也有下面定理.

定理 5.6 设 $\{\boldsymbol{A}^{(k)}\}$ 是 $\mathbf{R}^{n\times n}$ 中矩阵序列,$\boldsymbol{A}$ 是 $\mathbf{R}^{n\times n}$ 中矩阵,则 $\lim\limits_{k\to\infty}\|\boldsymbol{A}^{(k)}-\boldsymbol{A}\|=0$ 当且仅当 $\lim\limits_{k\to\infty}a_{ij}^{(k)}=a_{ij}$,其中 $a_{ij}^{(k)}$ 和 a_{ij} 分别表示 $\boldsymbol{A}^{(k)}$ 和 $\boldsymbol{A}$ 的分量.

定义 5.4 对于给定的向量范数 $\|\cdot\|$ 与矩阵范数 $\|\cdot\|$,如果 $\forall \boldsymbol{A}\in\mathbf{R}^{n\times n}$ 和 $\forall \boldsymbol{X}\in\mathbf{R}^{n}$,并满足

$$\|\boldsymbol{AX}\|\leqslant\|\boldsymbol{A}\|\cdot\|\boldsymbol{X}\|$$

则称此矩阵范数与向量范数是相容的.

定理 5.7 设 $\boldsymbol{A}$ 为 $n\times n$ 阶矩阵,$\|\cdot\|$ 是 $\mathbf{R}^n$ 中的向量范数,则

$$\|\boldsymbol{A}\|=\max_{\|\boldsymbol{X}\|=1}\|\boldsymbol{AX}\|$$

是一种矩阵范数,称其为由向量 $\|\cdot\|$ 诱导出的矩阵范数.

现在考察由向量范数 $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ 产生的矩阵范数.

定理 5.8 设 $\boldsymbol{A}=(a_{ij})\in\mathbf{R}^{n\times n}$,则

$$\|\boldsymbol{A}\|_1=\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

$$\|\boldsymbol{A}\|_\infty=\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|$$

$$\|\boldsymbol{A}\|_2=[\lambda_{\max}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})]^{\frac{1}{2}}$$

其中 $\boldsymbol{A}^{\mathrm{H}}$ 是 $\boldsymbol{A}$ 的共轭.

证明 下面只就 1 范数给出证明. 对 $\forall \boldsymbol{X}\in\mathbf{R}^n$, $\boldsymbol{X}\neq 0$,有

$$\|\boldsymbol{AX}\|_1=\sum_{i=1}^{n}\left|\sum_{j=1}^{n}a_{ij}x_j\right|\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|\cdot|x_j|$$
$$\leqslant\left(\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|\right)\sum_{j=1}^{n}|x_j|\leqslant\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|\cdot\|\boldsymbol{X}\|_1$$

故有

$$\|\boldsymbol{A}\|_1=\max_{X\neq 0}\frac{\|\boldsymbol{AX}\|_1}{\|\boldsymbol{X}\|_1}\leqslant\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}| \tag{5.19}$$

另一方面,设

$$\sum_{i=1}^{n}|a_{ij_0}| = \max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

取 $X^* = e_{j_0} = (0,\cdots,0,1,0,\cdots,0)^{\mathrm{T}}$,则

$$AX^* = (a_{1j_0}, a_{2j_0}, \cdots, a_{nj_0})^{\mathrm{T}}$$

$$\|AX^*\|_1 = \sum_{i=1}^{n}|a_{ij_0}| = \max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

注意到 $\|X^*\|_1 = 1$,

$$\|A\|_1 \geqslant \frac{\|AX^*\|_1}{\|X^*\|_1} = \max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}| \tag{5.20}$$

综合(5.19)式、(5.20)式得

$$\|A\|_1 = \max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|$$

在这3种范数中,1范数与∞范数的计算比较简单,2范数要计算矩阵的特征值,比较复杂.但2范数有一些好的性质,常用于理论分析.

此外,在 $\mathbf{R}^{n\times n}$ 上的一个常用且易于计算的矩阵范数为

$$\|A\|_{\mathrm{F}} = \left(\sum_{i,j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}}$$

通常称作 Frobenius 范数,它是 $\mathbf{R}^{n^2}$ 上的2范数的自然推广.

5.4.3　谱半径

定义 5.5　$A\in\mathbf{R}^{n\times n}$, A 的特征值为 $\lambda_i\ (i=1,2,\cdots,n)$,称 $\rho(A)=\max\limits_{1\leqslant i\leqslant n}|\lambda_i|$ 为 A 的谱半径.

定理 5.9　A 的谱半径不超过 A 的任何一个矩阵范数,即

$$\rho(A)\leqslant\|A\|$$

证明　设 λ 是 A 的任一特征值,X 是相应的特征向量,则有

$$AX=\lambda X$$

于是有

$$|\lambda|\cdot\|X\| = \|AX\|\leqslant\|A\|\cdot\|X\|$$

从而 $|\lambda|\leqslant\|A\|$. 于是

$$\rho(A)\leqslant\|A\|$$

定理 5.10　设 A 是任意 $n\times n$ 阶矩阵,由 A 的各次幂所组成的矩阵序列

$$I, A, A^2, \cdots, A^k, \cdots$$

收敛于零,即 $\lim\limits_{k\to\infty}A^k=0$ 的充要条件是 $\rho(A)<1$.

证明从略.读者可参阅冯康等编的《数值计算方法》.

5.4.4　方程右端误差对解的影响

考虑线性代数方程组

$$AX = b$$

设右端项有误差 δb 时，相应的方程组的解为 $X + \delta X$，即

$$A(X + \delta X) = b + \delta b$$

将 $AX = b$ 代入得

$$A\delta X = \delta b \quad 或 \quad \delta X = A^{-1}\delta b$$

这里假定 A 非奇异，从而有不等式

$$\|\delta X\| = \|A^{-1}\delta b\| \leqslant \|A^{-1}\| \cdot \|\delta b\|$$

而

$$\|AX\| = \|b\| \leqslant \|A\| \cdot \|X\|$$

所以

$$\frac{\|\delta X\|}{\|X\|} \leqslant \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\delta b\|}{\|b\|}$$

在这种情形下，数 $\|A\| \cdot \|A^{-1}\|$ 就是误差的放大率，即解的相对误差不超过右端项相对误差的 $\|A\| \cdot \|A^{-1}\|$ 倍，称 $\|A\| \cdot \|A^{-1}\|$ 为矩阵 A 的条件数，记作

$$\mathrm{Cond}(A) = \|A\| \cdot \|A^{-1}\|$$

5.4.5　系数矩阵误差对解的影响

设系数矩阵 A 有误差 δA，$X + \delta X$ 是系数矩阵为 $A + \delta A$ 时的准确解，即

$$(A + \delta A)(X + \delta X) = b \tag{5.21}$$

首先一个问题是，即使矩阵 A 非奇异，其元素有扰动 δA 后，矩阵 $A + \delta A$ 有可能变为奇异. 但是可以想像，若 δA 很小，矩阵 $A + \delta A$ 仍然可以保持它的非奇异性. 为了准确给出保持这种非奇异的条件，先证明下述引理.

引理　设矩阵 F 的范数小于 1，即 $\|F\| < 1$，则矩阵 $(I + F)$ 非奇异，且

$$\|(I + F)^{-1}\| \leqslant \frac{1}{1 - \|F\|}$$

证明　$\forall X \in \mathbf{R}^n$，恒有

$$\|(I + F)X\| = \|X + FX\| \leqslant \|X\| - \|FX\| \geqslant (1 - \|F\|) \cdot \|X\|$$

因为 $1 - \|F\| > 0$，所以，对 $X \neq 0$，$\|(I + F)X\| > 0$.

从而可知，方程组 $(I + F)X = 0$ 只有零解，因此矩阵 $(I + F)$ 非奇异.

又

$$\begin{aligned} 1 &= \|I\| = \|(I + F)(I + F)^{-1}\| \\ &= \|(I + F)^{-1} + F(I + F)^{-1}\| \\ &\geqslant \|(I + F)^{-1}\| - \|(I + F)^{-1}\| \cdot \|F\| \\ &= \|(I + F)^{-1}\|(1 - \|F\|) > 0 \end{aligned}$$

从而必有

$$\|(I+F)^{-1}\| \leqslant \frac{1}{1-\|F\|}$$

再来考虑矩阵 $A+\delta A$ 的非奇异性条件. 由于

$$A+\delta A = A(I+A^{-1}\delta A)$$

据前述引理即可知,若

$$\|A^{-1}\delta A\| < 1 \tag{5.22}$$

则矩阵 $A+\delta A$ 非奇异,且

$$\|(I+A^{-1}\delta A)^{-1}\| \leqslant \frac{1}{1-\|A^{-1}\delta A\|}$$

在条件(5.21)式成立时,据(5.21)式得

$$\begin{aligned}\delta X &= (A+\delta A)^{-1} - A^{-1}b = (A+\delta A)^{-1}[A-(A+\delta A)]A^{-1}b \\ &= (A+\delta A)^{-1}(-\delta A)A^{-1}b = (A+\delta A)^{-1}(-\delta A)X\end{aligned}$$

所以

$$\begin{aligned}\frac{\|\delta X\|}{\|X\|} &\leqslant \|(A+\delta A)^{-1}\delta A\| = \|(I+A^{-1}\delta A)^{-1}A^{-1}\delta A\| \\ &\leqslant \|(I+A^{-1}\delta A)^{-1}\| \cdot \|A^{-1}\delta A\| \leqslant \frac{\|A^{-1}\delta A\|}{1-\|A^{-1}\delta A\|}\end{aligned}$$

由于

$$\begin{aligned}\frac{\|A^{-1}\delta A\|}{1-\|A^{-1}\delta A\|} &\leqslant \frac{\|A^{-1}\| \cdot \|\delta A\|}{1-\|A^{-1}\| \cdot \|\delta A\|} \\ &= \frac{\|A\| \cdot \|A^{-1}\|}{1-\|A\| \cdot \|A^{-1}\| \dfrac{\|\delta A\|}{\|A\|}} \cdot \frac{\|\delta A\|}{\|A\|}\end{aligned}$$

从而得到不等式

$$\frac{\|\delta X\|}{\|X\|} \leqslant \frac{\mathrm{Cond}(A)}{1-\mathrm{Cond}(A)\dfrac{\|\delta A\|}{\|A\|}} \cdot \frac{\|\delta A\|}{\|A\|}$$

若 $\mathrm{Cond}(A)\dfrac{\|\delta A\|}{\|A\|}$ 很小,则 $\mathrm{Cond}(A)$ 仍表示相对误差的近似放大率.

根据以上讨论可以看出,$\mathrm{Cond}(A)$ 反映线性方程组 $AX=b$ 的解对初始数据误差的灵敏度,其值愈大,这种灵敏度愈高. 即对很小的初始误差 δb 或 δA,解的相对误差就有可能很大,从而大大破坏了解的精确度.

练习与思考

1. 分别用顺序高斯消去法、列主元高斯消去法、全主元高斯消去法求解下列方程组

$$\begin{bmatrix} 2 & 3 & 5 \\ 3 & 4 & 7 \\ 1 & 3 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 6 \\ 5 \end{bmatrix}$$

2. 用 LU 分解法求解下列方程组

$$\begin{bmatrix} 2 & -3 & 2 \\ -4 & 2 & -6 \\ 2 & 2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 14 \\ 8 \end{bmatrix}$$

3. 用平方根法求解下列方程组

$$\begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 4 \end{bmatrix}$$

4. 设 $\boldsymbol{A} = (a_{ij}^{(1)})_{n\times n}$,经顺序高斯消去法一步后 $\boldsymbol{A}$ 变为

$$\boldsymbol{A}^{(2)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix} = \begin{bmatrix} a_{11}^{(1)} & \alpha \\ 0 & \boldsymbol{A}_2 \end{bmatrix}$$

其中 $\alpha = [a_{12}^{(1)} \cdots a_{1n}^{(1)}]$,$\boldsymbol{A}_2$ 为 $(n-1)\times(n-1)$ 方阵. 证明:

(1)若 $\boldsymbol{A}$ 按行强对角占优,则矩阵 $\boldsymbol{A}_2$ 按行强对角占优;

(2)若 $\boldsymbol{A}$ 对称正定,则矩阵 $\boldsymbol{A}_2$ 对称正定.

5. 比较下面两个方程组的解

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1.00 & 0.50 & 0.33 \\ 0.50 & 0.33 & 0.25 \\ 0.33 & 0.25 & 0.20 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

6. 设有

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

其系数矩阵强对角占优. 试写出用 LU 分解法求其解的计算公式.

7. 对矩阵

$$A=\begin{bmatrix}-2&1&0&0\\1&-2&1&0\\0&1&-2&1\\0&0&1&-2\end{bmatrix}$$

求 $\|A\|_\infty$, $\|A\|_2$, $\|A\|_1$, $\|A\|_F$ 和 $\mathrm{Cond}(A)_2$.

8. 试证明 $\|A\|_2\leqslant\|A\|_1\cdot\|A\|_\infty$.

9. 画出 $\mathbf{R}^2$ 中满足下列不等式的集合

(1) $\|X\|_1\leqslant 1$;

(2) $\|X\|_2\leqslant 1$;

(3) $\|X\|_\infty\leqslant 1$.

10. 设 $A,B\in\mathbf{R}^{n\times n}$ 均非奇异，$\|\cdot\|$是任何矩阵范数. 证明：

(1) $\|A^{-1}\|\geqslant\dfrac{1}{\|A\|}$;

(2) $\|A^{-1}-B^{-1}\|\leqslant\|A^{-1}\|\cdot\|B^{-1}\|\cdot\|A-B\|$.

11. 设有方程组 $AX=b$，其中

$$A=\begin{bmatrix}1&0&-1\\2&2&1\\0&2&2\end{bmatrix}\qquad b=\begin{bmatrix}\dfrac{1}{2}\\\dfrac{1}{3}\\-\dfrac{1}{3}\end{bmatrix}$$

如果右端 b 有小扰动 $\|\delta b\|_\infty=5\times10^{-7}$，试估计由此引起的解的相对误差.

12. 设有方程组 $AX=b$.

$$\begin{bmatrix}240&-319\\-179&240\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3\\4\end{bmatrix}$$

及扰动方程组 $(A+\delta A)(X+\delta A)=b$：

$$\begin{bmatrix}240&-319.5\\-179.5&240\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}=\begin{bmatrix}3\\4\end{bmatrix}$$

(1) 试分别解出 X 和 $X+\delta X$；

(2) 利用扰动误差分析理论估计$\dfrac{\|\delta X\|_\infty}{\|X\|_\infty}$，并将理论结果与实际求解结果比较.