第2章　矩阵变换和计算

2.1　矩阵的三角分解及其应用

2.2　特殊矩阵的特征系统

2.4　矩阵的奇异值分解

2.1 矩阵的三角分解及其应用

2.1.1 Gauss消去法与矩阵的*LU*分解

2.1.2 **Gauss**列主元消去法与带列主元的*LU*分解

2.1.3 对称矩阵的**Cholesky**分解

2.1.4 三对角矩阵的三角分解

2.1.5 条件数与方程组的性态

2.1.6 矩阵的*QR*分解

2.1.1 Gauss消去法与矩阵的LU分解

例1 Gauss消去法求解线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个实例。

$$
\begin{cases}
2x_1 + x_2 + x_3 = 4 & r_1^{(0)} \\
4x_1 + 3x_2 + 3x_3 + x_4 = 11 & r_2^{(0)} \\
8x_1 + 7x_2 + 9x_3 + 5x_4 = 29 & r_3^{(0)} \\
6x_1 + 7x_2 + 9x_3 + 8x_4 = 30 & r_4^{(0)}
\end{cases}
$$

第一步，消去 $r_2^{(0)}$ 、$r_3^{(0)}$ 和 $r_4^{(0)}$ 中的 x_1,即用 $\left(-\frac{4}{2}\right)\times r_1^{(0)}+r_2^{(0)}$、$\left(-\frac{8}{2}\right)\times r_1^{(0)}+r_3^{(0)}$ 和 $\left(-\frac{6}{2}\right)\times r_1^{(0)}+r_4^{(0)}$ 得

$$\begin{cases} 2x_1 + x_2 + x_3 = 4 & r_1^{(0)} \\ x_2 + x_3 + x_4 = 3 & r_2^{(1)} \\ 3x_2 + 5x_3 + 5x_4 = 13 & r_3^{(1)} \\ 4x_2 + 6x_3 + 8x_4 = 18 & r_4^{(1)} \end{cases}$$

第二步，消去 $r_3^{(1)}$ 和 $r_4^{(1)}$ 中的 x_2，即用

$\left(-\frac{3}{1}\right)\times r_2^{(1)}+r_3^{(1)}$ 和 $\left(-\frac{4}{1}\right)\times r_2^{(1)}+r_4^{(1)}$ 得

$$
\begin{cases}
2x_1+x_2+x_3 = 4 & r_1^{(0)} \\
x_2+x_3+x_4=3 & r_2^{(1)} \\
2x_3+2x_4=4 & r_3^{(2)} \\
2x_3+4x_4=18 & r_4^{(2)}
\end{cases}
$$

第三步，消去 $r_4^{(2)}$ 中的 x_3 ，即用 $\left(-\frac{2}{2}\right)\times r_3^{(2)}+r_4^{(1)}$ 得

$$\begin{cases} 2x_1 + x_2 + x_3 \qquad\qquad = 4 & r_1^{(0)} \\ \qquad\quad x_2 + x_3 + \ x_4 = 3 & r_2^{(1)} \\ \qquad\qquad 2x_3 + 2x_4 = 4 & r_3^{(2)} \\ \qquad\qquad\qquad 2x_4 = 2 & r_4^{(3)} \end{cases}$$

$$
\begin{cases}
2x_1 + x_2 + x_3 = 4 & \Rightarrow \; 2x_1 + 1 + 1 = 4 \\
x_2 + x_3 + x_4 = 3 & \Rightarrow \; x_2 + 1 + 1 = 3 \\
2x_3 + 2x_4 = 4 & \Rightarrow \; 2x_3 + 2 \times 1 = 4 \\
2x_4 = 2 & \Rightarrow \; 2x_4 = 2 = 1
\end{cases}
$$

上述为回代求解过程，得解。$x = (1, 1, 1, 1)^T$。

Gauss 消去法的实质是首先通过一系列的初等行变换将增广矩阵 $(\boldsymbol{A}\,|\,\boldsymbol{b})$ 化成上三角增广矩阵 $(\boldsymbol{U}\,|\,\boldsymbol{c})$，然后通过回代求与 $\boldsymbol{Ax=b}$ 三角方程组 $\boldsymbol{Ux=c}$ 的解。

我们来观察Gauss消去法求 $Ax=b$ 的解，增广矩阵 $(A\,|\,b)$ 化成上三角矩阵 $(U\,|\,c)$ 的过程，如何通过矩阵的变换来实现的。首先，注意

$$A=\begin{pmatrix}2&1&1&0\\4&3&3&1\\8&7&9&5\\6&7&9&8\end{pmatrix},\qquad b=\begin{pmatrix}4\\11\\29\\30\end{pmatrix}$$

返回

解：

$$(A \mid b) \xrightarrow{\text{第三次消元}} \begin{pmatrix} 2 & 1 & 1 & 0 & 4 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 2 & 2 \end{pmatrix} = (U \mid c)$$

三次消元过程写成矩阵的形式分别为：

$$\boldsymbol{L}_1\boldsymbol{A}=\begin{pmatrix}1&&&\\-2&1&&\\-4&&1&\\-3&&&1\end{pmatrix}\begin{pmatrix}2&1&1&0\\4&3&3&1\\8&7&9&5\\6&7&9&8\end{pmatrix}=\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&3&5&5\\0&4&6&8\end{pmatrix}$$

$$\boldsymbol{L}_2(\boldsymbol{L}_1\boldsymbol{A})=\begin{pmatrix}1&&&\\&1&&\\&-3&1&\\&-4&&1\end{pmatrix}\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&3&5&5\\0&4&6&8\end{pmatrix}=\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&0&2&2\\0&0&2&4\end{pmatrix}$$

$$\boldsymbol{L}_3(\boldsymbol{L}_2\boldsymbol{L}_1\boldsymbol{A}) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix} = \boldsymbol{U}$$

注意单位下三角矩阵

$$L_k = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & -l_{n,k} & & & 1 \end{pmatrix}$$

有令人惊奇，而平凡的性质：

(1) L_k 的逆恰好是 L_k 本身的每一个对角线以下的元素都取相反数；即

$$L_k^{-1} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{n,k} & & & 1 \end{pmatrix}$$

事实上，我们定义 $\boldsymbol{l}_k = \left(0 \cdots 0 \; l_{k+1k} \cdots l_{nk}\right)^T$

则$\boldsymbol{L}_k$可写成

$$\boldsymbol{L}_k = \boldsymbol{I} - \boldsymbol{l}_k \boldsymbol{e}_k^T \text{ ,}$$

其中 $\boldsymbol{e}_k = \left(0 \cdots 0 \;\; 1 \;\; 0 \cdots 0\right)^T$，$\boldsymbol{e}_k^T \, \boldsymbol{l}_k = 0$。 而

$$\left(I - l_k e_k^T\right)\left(I + l_k e_k^T\right) = I - l_k e_k^T + l_k e_k^T - l_k e_k^T l_k e_k^T$$

$$= I - l_k e_k^T l_k e_k^T = I$$

故 L_k 的逆为：

$$
\boldsymbol{L}_k^{-1}=\left(I+l_k e_k^T\right)=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & \ddots & \\ & & & & & 1\end{pmatrix}+\begin{pmatrix}0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & l_{k+1,k} & 0 & & \\ & & \vdots & & \ddots & \\ & & l_{n,k} & & & 0\end{pmatrix}
$$

$$
=\begin{pmatrix}1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{k+1,k} & 1 & & \\ & & \vdots & & \ddots & \\ & & l_{n,k} & & & 1\end{pmatrix}
$$

则对于例题中的单位下三角阵而言，就有：

$$\boldsymbol{L}_1=\begin{pmatrix}1&&&\\-2&1&&\\-4&&1&\\-3&&&1\end{pmatrix}\quad \boldsymbol{L}_2=\begin{pmatrix}1&&&\\&1&&\\&-3&1&\\&-4&&1\end{pmatrix}\quad \boldsymbol{L}_3=\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&-1&1\end{pmatrix}$$

$$\Updownarrow\qquad\qquad\Updownarrow\qquad\qquad\Updownarrow$$

$$\boldsymbol{L}_1^{-1}=\begin{pmatrix}1&&&\\2&1&&\\4&&1&\\3&&&1\end{pmatrix}\quad \boldsymbol{L}_2^{-1}=\begin{pmatrix}1&&&\\&1&&\\&3&1&\\&4&&1\end{pmatrix}\quad \boldsymbol{L}_3^{-1}=\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&1&1\end{pmatrix}$$

（2）乘积矩阵$\boldsymbol{L}$恰好是它们具有的非零对角线以下元素嵌入到相应位置的单位下三角矩阵。

考虑矩阵乘积 $\boldsymbol{L}_k^{-1}\ \boldsymbol{L}_{k+1}^{-1}$

$$\boldsymbol{L}_k^{-1}\boldsymbol{L}_{k+1}^{-1} = \left(I + l_k e_k^T\right)\left(I + l_{k+1} e_{k+1}^T\right)$$

$$= I + l_k e_k^T + l_{k+1} e_{k+1}^T - l_k e_k^T l_{k+1} e_{k+1}^T = I + l_k e_k^T + l_{k+1} e_{k+1}^T$$

$$= \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & l_{k+1,k} & 1 & & & \\ & & l_{k+2,k} & l_{k+2,k+1} & \ddots & & \\ & & \vdots & \vdots & & 1 & \\ & & l_{n,k} & l_{n,k+1} & & & 1 \end{pmatrix}$$

当我们取所有这些矩阵乘积$\boldsymbol{L}$时，对角线下面的每处都有同样方便的性质：

$$\boldsymbol{L}=\boldsymbol{L}_1^{-1}\boldsymbol{L}_2^{-1}\cdots\boldsymbol{L}_{n-1}^{-1}=\begin{pmatrix} 1 & & & & \\ l_{21} & 1 & & & \\ l_{31} & l_{32} & 1 & & \\ \vdots & \vdots & \ddots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn-1} & 1 \end{pmatrix}$$

这样一来，例题中的计算过程可以表示为

$$L_3 L_2 L_1 A = U$$

$$A = L_1^{-1} L_2^{-1} L_3^{-1} U$$

令

$$L = L_1^{-1} L_2^{-1} L_3^{-1}$$

则由性质（2），可得出L的表达式，即

$$L_1^{-1}=\begin{pmatrix}1&&&\\2&1&&\\4&&1&\\3&&&1\end{pmatrix}\quad L_2^{-1}=\begin{pmatrix}1&&&\\&1&&\\&3&1&\\&4&&1\end{pmatrix}\quad L_3^{-1}=\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&1&1\end{pmatrix}$$

$$L=L_1^{-1}L_2^{-1}L_3^{-1}=\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{pmatrix}$$

$$L_1^{-1}=\begin{pmatrix}1 & & & \\ 2 & 1 & & \\ 4 & & 1 & \\ 3 & & & 1\end{pmatrix}\quad L_2^{-1}=\begin{pmatrix}1 & & & \\ & 1 & & \\ & 3 & 1 & \\ & 4 & & 1\end{pmatrix}\quad L_3^{-1}=\begin{pmatrix}1 & & & \\ & 1 & & \\ & & 1 & \\ & & 1 & 1\end{pmatrix}$$

$$L=L_1^{-1}L_2^{-1}L_3^{-1}=\begin{pmatrix}1 & 0 & 0 & 0\\ 2 & 1 & 0 & 0\\ 4 & 3 & 1 & 0\\ 3 & 4 & 1 & 1\end{pmatrix}$$

从而有

$$
=\begin{pmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{pmatrix}
$$

$$A$$

即得到矩阵A的LU分解

定义2.1 对于n阶方阵A，如果存在n阶单位下三角矩阵L和n阶上三角矩阵U，使得

$$A = LU$$

则称其为矩阵A的LU分解，也称 Doolittle分解。

Doolittle方法求解线性方程组：

$$AX = B \Leftrightarrow (LU)X = B$$

$$\Updownarrow$$

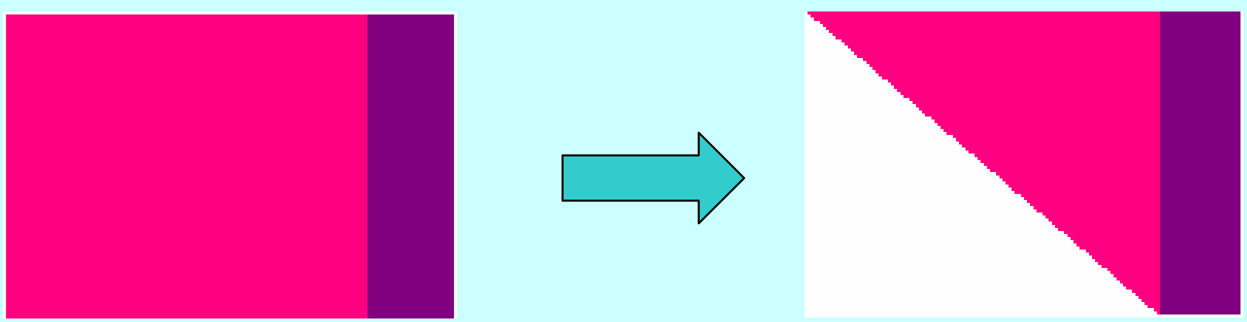
$$\begin{cases} LY = B \\ UX = Y \end{cases}$$

其中A，X，B，Y均为矩阵

下面对一般n阶方阵 A 进行 LU 分解。通过前例我们可以想到

思路

首先将A化为上三角阵，再回代求解。



步骤如下：

$$\text{第一步，第1行}\times\left(-\frac{a_{i1}}{a_{11}}\right)+\text{第}i\text{行},\quad i=2,\cdots,n$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix}$$

运算量：$(n-1)\times(1+n)$

第二步： 第2行$\times\left(\dfrac{a_{i2}^{(1)}}{a_{22}^{(1)}}\right)+$第$i$行，$i=3,\cdots,n$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix} \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & a_{n3}^{(2)} & \cdots & a_{nn}^{(2)} & b_n^{(2)} \end{pmatrix}$$

运算量： $(n\text{-}2)\times(1+n-1)=(n\text{-}2)n$

类似的做下去，我们有：

第k步： 第k行$\times\left(-\dfrac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}\right)+$第$i$行，$i=k+1,\cdots,n$

运算量： $(n-k)\times(1+n-k+1)=(n-k)(n-k+2)$

$n-1$步以后，我们可以得到变换后的矩阵为：

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ 0 & 0 & a_{33}^{(2)} & \cdots & a_{3n}^{(2)} & b_3^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn}^{(n-1)} & b_n^{(n-1)} \end{pmatrix}$$

因此，总的运算量为：

$$\sum_{k=1}^{n-1}(n-k)(n-k+2)$$

加上解上述上三角阵的运算量 $(n+1)n/2$，

总共为：

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

当n较大时，它和　　　　同阶的。

注意到，计算过程中$a_{kk}^{(k-1)}$处在被除的位置，因此整个计算过程要保证它不为0。所以，**Gauss消元法的可行条件**为：

$$a_{kk}^{(k-1)} \neq 0$$

而这一条件的等价条件是要求A的各阶顺序主子式均不为0，即

$$\det\left(\begin{bmatrix} a_{11} & \cdots & a_{1i} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ii} \end{bmatrix}\right) \neq 0, \quad i = 1, \cdots, n.$$

因此，有些有解的问题，不能用Gauss消元求解。

另外，如果某个$a_{kk}^{(k-1)}$很小的话，会引入大的误差。

于是便有了——

2.1.2

Gauss列主元消去法
与
带列主元的LU分解

1. Gauss列主元消去法

例2　在一台八位十进制的计算机上，用Gauss消去法解线性方程组

$$\begin{pmatrix} 10^{-8} & 2 & 3 \\ -1 & 3.712 & 4.623 \\ -2 & 1.072 & 5.643 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

解：在八位十进制的计算机上，经过两次消元有

$$(\boldsymbol{A}\,|\,\boldsymbol{b}) \xrightarrow{\text{第三次消元}} \begin{pmatrix} 10^{-8} & 2 & 3 & 1 \\ 0 & 0.2\times10^{9} & 0.3\times10^{9} & 0.1\times10^{9} \\ 0 & 0.4\times10^{9} & 0.6\times10^{9} & 0.2\times10^{9} \end{pmatrix}$$

$$= (\boldsymbol{U}\,|\,\boldsymbol{c})$$

显然 $(U\,|\,c)$有无穷多解. 但实际上，$\det(A)\neq 0$，线性方程组有唯一解。

因此在计算过程中的舍入误差使解的面目全非了，这些均是由于小主元作除数所致.

Gauss列主元消去法:

为避免小主元作除数、或0作分母，在**Gauss**消去法中增加选主元的过程，即在第k步 $(k=1,2,\cdots,n-1)$ 消元时，首先在第k列主对角元以下（含主对角元）元素中挑选绝对值最大的数，并通过初等行交换，使得该数位于主对角线上，然后再继续消元。

称该绝对值最大的数为列主元。

将在消元过程中，每一步都按列选主元的Gauss消去法称之为Gauss列主元消去法。

例3　用**Gauss**列主元消去法解例2中的方程组。

解　$$(\boldsymbol{A}\,|\,\boldsymbol{b})=\begin{pmatrix}10^{-8} & 2 & 3 & 1\\ -1 & 3.712 & 4.623 & 2\\ -2 & 1.072 & 5.643 & 3\end{pmatrix}$$

$$\xrightarrow{\text{选列主元，交换第一和第三行}}\begin{pmatrix}-2 & 1.072 & 5.643 & 3\\ -1 & 3.712 & 4.623 & 2\\ 10^{-8} & 2 & 3 & 1\end{pmatrix}\xrightarrow{\text{第一次消元}}\begin{pmatrix}-2 & 1.072 & 5.643 & 3\\ 0 & 0.3176\times 10 & 0.1801500\times 10 & 0.5\\ 0 & 0.2\times 10 & 0.3\times 10 & 0.1\times 10\end{pmatrix}$$

$$\xrightarrow{\text{第二次消元}}\begin{pmatrix}-2 & 1.072 & 5.643 & 3\\ 0 & 0.3176\times 10 & 0.1801500\times 10 & 0.5\\ 0 & 0 & 0.18655541\times 10 & 0.6851854\end{pmatrix}=(\boldsymbol{U}\,|\,\boldsymbol{c})$$

用回代法求 $(\boldsymbol{U}\,|\,\boldsymbol{c})$ 的解得

$$\tilde{\boldsymbol{x}}=(-0.49105820,-0.05088607,0.36725739)^T$$

方程组的精确解为:

$$\boldsymbol{x}=(-0.491058227,\ -0.050886075,\ 0.367257384)^T$$

例3　用**Gauss**列主元消去法解例2中的方程组。

解： $(\boldsymbol{A}\,|\,\boldsymbol{b}) = \begin{pmatrix} 10^{-8} & 2 & 3 & 1 \\ -1 & 3.712 & 4.623 & 2 \\ -2 & 1.072 & 5.643 & 3 \end{pmatrix}$

$$\xrightarrow[\text{第一次消元}\;\; r_1 \leftrightarrow r_2]{\text{选列主元}} \begin{pmatrix} -2 & 1.072 & 5.643 & 3 \\ 0 & 0.3176\times 10 & 0.1801\times 10 & 0.5 \\ 0 & 0 & 0.18655541\times 10 & 0.6851854 \end{pmatrix} = (\boldsymbol{U}\,|\,\boldsymbol{c})$$

用回代法求 $(\boldsymbol{U}\,|\,\boldsymbol{c})$ 得数值解为：

$$\tilde{\boldsymbol{x}} = (-0.49105820, -0.05088607, 0.36725739)^T$$

方程组的精确解为:

$$\boldsymbol{x} = (-0.491058227,\ -0.050886075,\ 0.367257384)^T$$

2. 带列主元的LU分解

由上述**Gauss**列主元消去过程可以得到矩阵的带有列选主元的LU分解，还是以例1中的系数矩阵A为例来说明。

回忆

L_3:

$= U$

实际上，上述过程可以表示为

$$L_3P_3L_2P_2L_1P_1A = U$$

显然，$L_3P_3L_2P_2L_1P_1$ 似乎并不是一个单位下三角矩阵．我们将上式改写为

$$L_3(P_3L_2P_3^{-1})(P_3P_2L_1P_2^{-1}P_3^{-1})(P_3P_2P_1)A = U$$

由 $\boldsymbol{P}_i$ 的定义知 $\boldsymbol{P}_i^{-1}=\boldsymbol{P}_i$，即

$$\boldsymbol{P}_1=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}=\boldsymbol{P}_1^{-1}\qquad \boldsymbol{P}_2=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{pmatrix}=\boldsymbol{P}_2^{-1}$$

$$\boldsymbol{P}_3=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}=\boldsymbol{P}_3^{-1}$$

从而，记

$$\tilde{L}_2 = P_3 L_2 P_3 = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & \frac{2}{7} & 1 & \\ & \frac{3}{7} & & 1 \end{pmatrix}$$

$$\tilde{L}_1 = P_3 P_2 L_1 P_2 P_3 = \begin{pmatrix} 1 & & & \\ -\frac{3}{4} & 1 & & \\ -\frac{1}{2} & & 1 & \\ -\frac{1}{4} & & & 1 \end{pmatrix}$$

显然，$\tilde{L}_2$和 $\tilde{L}_1$ 分别与 L_2和 L_3结构相同，只是下三角部分的元素进行相应的对调。从而有

$$L_3\,(P_3L_2P_3^{-1})\,(P_3P_2L_1P_2^{-1}P_3^{-1})\,(P_3P_2P_1)A = U$$

$$\Updownarrow$$

$$L_3\,\tilde{L}_2\,\tilde{L}_1\,(P_3P_2P_1)A = U \quad L_3^{-1}\tilde{L}_2^{-1}\tilde{L}_1^{-1}$$

令

$$P = P_3P_2P_1, \quad \tilde{L} = \tilde{L}_1^{-1}\tilde{L}_2^{-1}L_3^{-1}$$

则有

$$P = P_3 P_2 P_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$L = \tilde{L}_1^{-1} \tilde{L}_2^{-1} L_3^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{4} & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{2}{7} & 1 & 0 \\ \frac{1}{4} & -\frac{3}{7} & \frac{1}{3} & 1 \end{pmatrix}$$

这样，我们得到另一种形式的矩阵分解：

$$PA = LU$$

$$= $$

P A L U

一般地, 如果A为n阶方阵，进行**Gauss**列主元消去过程为：

$$L_{n-1}P_{n-1}\cdots L_2P_2L_1P_1A=U$$

类似的，可以改写成：

$$(L_{n-1}\tilde{L}_{n-2}\cdots\tilde{L}_2\tilde{L}_1)(P_{n-1}\cdots P_2P_1)A=U$$

其中，$\tilde{L}_k = P_{n-1} \cdots P_{k+1} L_k P_{k+1} \cdots P_{n-1}$（$k$=1,2,…,$n$−2）与$L_k$的结构相同，只是下三角部分元素经过了对调。因此，令

$$L = (L_{n-1}\tilde{L}_{n-2} \cdots \tilde{L}_2 \tilde{L}_1)^{-1} \quad P = P_{n-1} \cdots P_2 P_1$$

则

$$PA = LU$$

定理 对任意n阶矩阵A，均存在置换矩阵P、单位下三角矩阵L和上三角矩阵U，使得

$$PA = LU$$

例　用Gauss列主元消去法解如下方程组并给出*PA=LU*分解。

解:

$$
(A\,|\,b)=\begin{pmatrix} 0 & -6 & -1 & -2 \\ 1 & 2 & 2 & 4 \\ 2 & -2 & 1 & 1 \end{pmatrix}
$$

$$
\xrightarrow{\text{选列主元，交换第1和第三行}}
\begin{pmatrix} 2 & -2 & 1 & 1 \\ 1 & 2 & 2 & 4 \\ 0 & -6 & -1 & -2 \end{pmatrix}
\rightarrow \cdots = (U\,|\,c)
$$

$$\xrightarrow{\text{第一次消元}} \begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & 3 & \frac{3}{2} & \frac{7}{2} \\ 0 & -6 & -1 & -2 \end{pmatrix}$$

$$\xrightarrow{\text{选列主元，}r_2 \leftrightarrow r_3} \begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & -6 & -1 & -2 \\ 0 & 3 & \frac{3}{2} & \frac{7}{2} \end{pmatrix}$$

用回代法求的解得：

$$x_3=\frac{5}{2}\quad x_2=\frac{-2+\frac{5}{2}}{-6}=-\frac{1}{12}\quad x_1=-\frac{5}{6}\quad \text{即 } \boldsymbol{x}=\left(-\frac{5}{6},-\frac{1}{12},\frac{5}{2}\right)^T。$$

下面求相应的***PA*=*LU***分解

第一次选列主元，交换第1行和第3行，左乘置换矩阵$\boldsymbol{P}_1$

$$\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}0&-6&-1\\1&2&2\\2&-2&1\end{pmatrix}=\begin{pmatrix}2&-2&1\\1&2&2\\0&-6&-1\end{pmatrix}$$

第一次消元，用L_1左乘(P_1A)，即

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 0 & -6 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \\ 0 & 3 & \frac{3}{2} \\ 0 & -6 & -1 \end{pmatrix}$$

第二次选列主元，交换第2行和第3行，即左乘置换矩阵P_2

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 0 & 3 & \frac{3}{2} \\ 0 & -6 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 1 \\ 0 & -6 & -1 \\ 0 & 3 & \frac{3}{2} \end{pmatrix}$$

第二次消元，用L_2左乘$(P_2L_1P_1A)$，即

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 0 & -6 & -1 \\ 0 & 3 & \frac{3}{2} \end{pmatrix}=\begin{pmatrix} 2 & -2 & 1 \\ 0 & -6 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}=U$$

注意：

$$\tilde{L}_1=P_2L_1P_2=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{pmatrix}$$

则分解应为：

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & -6 & -1 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{pmatrix}$$

$$=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 0 & -6 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

即有: $$PA=LU$$

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & -6 & -1 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 1 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 0 & -6 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

练习题 用列主元**Gauss**消去法解如下方程组，并利用得到的上三角矩阵求出 $\det(A)$

$$\begin{pmatrix} -3 & 2 & 6 \\ 10 & -7 & 0 \\ 5 & -1 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 7 \\ 6 \end{pmatrix}$$

解:

$$\begin{pmatrix} -3 & 2 & 6 & 4 \\ 10 & -7 & 0 & 7 \\ 5 & -1 & 5 & 6 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 10 & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 5 & 6 \end{pmatrix} \xrightarrow{\text{消元}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & -\dfrac{1}{10} & 6 & \dfrac{61}{10} \\ 0 & \dfrac{5}{2} & 5 & \dfrac{5}{2} \end{pmatrix}$$

$$\begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & -\frac{1}{10} & 6 & \frac{61}{10} \end{pmatrix} \xrightarrow{\text{消元}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & 0 & \frac{31}{5} & \frac{31}{5} \end{pmatrix}$$

从而求得方程组解：$x_1 = 0 \quad x_2 = -1 \quad x_3 = 1$

又，

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad \det(\boldsymbol{P}) = 1$$

则

$$\det(\boldsymbol{PA}) = \det(\boldsymbol{LU}) = 10 \times \frac{5}{2} \times \frac{31}{5} = 155, \quad \det(\boldsymbol{A}) = 155$$

2.1.3 对称矩阵的Cholesky分解

将对称正定阵 A 做 LU 分解，得到L和U，进一步

$$U=\begin{pmatrix} & u_{ij} \\ & \end{pmatrix}=\begin{pmatrix} u_{11} & & \\ & u_{22} & \\ & & u_{nn} \end{pmatrix}\begin{pmatrix} 1 & & * \\ & 1 & \\ & & 1 \end{pmatrix} \overset{\text{记为}}{=\!=} D\tilde{U}$$

即 $A=L(D\tilde{U})$，由A 对称，得 $L(D\tilde{U})=\tilde{U}^T(DL^T)$

由A 的LU分解的唯一性 $\longrightarrow$ $L=\tilde{U}^T$ 即 $A=LDL^T$

记 $D^{1/2}=\begin{pmatrix} \sqrt{u_{11}} & & \\ & \sqrt{u_{22}} & \\ & & \sqrt{u_{nn}} \end{pmatrix}$ 则 $\tilde{L}=LD^{1/2}$ 是下三角矩阵

对称正定阵的分解为：

$$\boxed{A=\tilde{L}\tilde{L}^T}$$

定理：（Cholesky分解）

对任意n阶对称正定矩阵 A，均存在下三角矩阵L 使 $A=LL^{\mathrm{T}}$ 成立，称其为对称正定矩阵A的Cholesky分解. 进一步地, 如果规定 L 的对角元为正数，则 L 是唯一确定的。

下面研究如何进行对称正定矩阵的**Cholesky**分解。当然，上述的证明过程提供一种计算**Cholesky**分解的方法。我们还可以使用下面将要介绍的直接分解方法。

设

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

利用矩阵乘法规则和利用的下三角结构得到:

$$a_{ij} = \sum_{k=1}^{j-1} l_{ik} l_{jk} + l_{ij} l_{jj}, \quad i = j, j+1, \cdots, n$$

用平方根法解线性代数方程组的算法

（1）对矩阵A进行**Cholesky**分解，即$A=LL^T$，由矩阵乘法：

对于 $j=1, 2, \cdots, n$ 计算

$$l_{jj}=\left(a_{jj}-\sum_{k=1}^{j-1}l_{jk}^2\right)^{\frac{1}{2}},\quad l_{ij}=\left(a_{ij}-\sum_{k=1}^{j-1}l_{ik}l_{jk}\right)/l_{jj}\ ,$$

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

计算次序为：

$$l_{11}, l_{21}, \cdots, l_{n1},\ l_{22}, l_{32}, \cdots, l_{n2}, \cdots, l_{nn}\text{。}$$

（2）求解下三角形方程组

$$y_1 = b_1 / l_{11}, \quad y_i = \left(b_i - \sum_{k=1}^{i-1} l_{ik} y_k \right) / l_{ii}, \quad i = 2, 3, \cdots, n$$

（3）求解$\boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{y}$

$$x_n = y_n / l_{nn}, \quad x_i = \left(y_i - \sum_{k=i+1}^{n} l_{ki} x_k \right) / l_{ii},$$

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

由

$$l_{jj}=\left(a_{jj}-\sum_{k=1}^{j-1}l_{jk}^{2}\right)^{\frac{1}{2}}$$

得 $a_{jj}=\sum_{k=1}^{j}l_{jk}^{2}$，由此推出 $\left|l_{jk}\right|\le\sqrt{a_{jj}}$，$\boldsymbol{k}$=1, 2, ⋯, $\boldsymbol{j}$。

因此在分解过程中$\boldsymbol{L}$元素的数量级不会增长，

故平方根法通常是数值稳定的，不必选主元。

例用Cholesky方法解线性方程组*Ax=b*，其中

$$A=\begin{pmatrix}4 & -1 & 1\\ -1 & 4.25 & 2.75\\ 1 & 2.75 & 3.5\end{pmatrix}\qquad b=\begin{pmatrix}4\\ 6\\ 7.25\end{pmatrix}$$

解： 显然$A^T=A$，且 $D_1=4>0$，$D_2=16>0$，$D_3=16>0$，因此，为对称正定矩阵，故存在 $A=LL^T$。 由分解公式（2-15）和(2-16)次计算出L的诸元素：

$$l_{11}=\sqrt{a_{11}}=\sqrt{4}=2,\qquad l_{21}=\frac{a_{21}}{l_{11}}=-0.5,\qquad l_{31}=\frac{a_{31}}{l_{11}}=0.5,$$

$$l_{22}=\sqrt{a_{22}-l_{21}^2}=\sqrt{4.25-0.5^2}=2,\qquad l_{31}=\frac{a_{12}-l_{31}l_{21}}{l_{11}}=0.5\times(2.75+0.5^2)=1.5,$$

$$l_{33}=\sqrt{a_{33}-l_{31}^2-l_{32}^2}=\sqrt{3.5-0.5^2-1.5^2}=1$$

从而得

$$L=\begin{pmatrix}2 & & \\ -0.5 & 2 & \\ 0.5 & 1.5 & 1\end{pmatrix}$$

再利用（**2-18**）求下三角方程组$\boldsymbol{Ly=b}$的解，即得

$$y_1=\frac{b_1}{l_{11}}=\frac{4}{2}=2,\quad y_2=\frac{b_2-l_{21}\cdot y_1}{l_{22}}=\frac{6+1}{2}=3.5,$$

$$y_3=\frac{b_3-l_{31}\cdot y_1-l_{32}\cdot y_2}{l_{33}}=7.25-0.5\times 2-1.5\times 3.5=1,\quad \boldsymbol{y}=(2,,3.5,1)^T。$$

再利用（**2-19**）求上三角方程组$\boldsymbol{L^T x=y}$的解，即得

$$x_3=\frac{y_3}{l_{33}}=\frac{1}{1}=1,\quad x_2=\frac{y_2-l_{32}\cdot x_2}{l_{22}}=\frac{3.5-1.5}{2}=1,$$

$$x_1=\frac{y_1-l_{21}\cdot x_2-l_{31}\cdot x_3}{l_{11}}=0.5\times(2+0.5-0.5)=1,\quad \boldsymbol{x}=(1,1,1)^T。$$

2.1.4 三对角矩阵的三角分解

设三对角矩阵

$$\boldsymbol{A}=\begin{pmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ & & & a_n & b_n \end{pmatrix}$$

如果矩阵$\boldsymbol{A}$可以进行$\boldsymbol{LU}$分解$\boldsymbol{A}=\boldsymbol{LU}$, 其中

$$\boldsymbol{L}=\begin{pmatrix} 1 & & & & \\ l_2 & 1 & & & \\ & l_3 & \ddots & & \\ & & \ddots & 1 & \\ & & & l_n & 1 \end{pmatrix},\quad \boldsymbol{U}=\begin{pmatrix} u_1 & d_1 & & & \\ & u_2 & d_2 & & \\ & & \ddots & \ddots & \\ & & & u_{n-1} & d_{n-1} \\ & & & & u_n \end{pmatrix}$$

用追赶法解三对角形方程组的算法

（1）对矩阵*A*进行*LU*分解，公式如下：

$$
\begin{cases}
d_i = c_i, & i = 1,2,\cdots,n-1 \\
u_1 = b_1 & \\
l_i = a_i / u_{i-1}, & i = 2,3,\cdots,n \\
u_i = b_i - l_i c_{i-1}, & i = 2,3,\cdots,n
\end{cases}
$$

计算次序是：

$$u_1 \to l_2 \to u_2 \to l_3 \to u_3 \to \cdots \to l_n \to u_n$$

（2）求解下三角形方程组

$$y_1 = f_1, \quad y_i = f_i - l_i y_{i-1}, \quad i = 2, 3, \cdots, n$$

（3）求解$\boldsymbol{Ux} = \boldsymbol{y}$

$$x_n = y_n / u_n, \quad x_i = (y_i - c_i x_{i+1}) / u_i,$$

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

定理　设具有三对角形式的矩阵$\boldsymbol{A}$，满足条件

(1)　$|b_1| > |c_1| > 0$

(2)　$|b_n| > |a_n| > 0$

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

则方程组$\boldsymbol{Ax} = \boldsymbol{f}$可用追赶法，且解存在唯一。

证 由（**2-22**）和条件（**1**）知， $u_1 = b_1 \neq 0$且有 $0 < \left|\dfrac{c_1}{u_1}\right| < 1$。

下面用归纳法证明 $u_i \neq 0$且有 $0 < \left|\dfrac{c_i}{u_i}\right| < 1,\ i = 2,3,\cdots,n-1$。

假设 $u_{i-1} \neq 0,\ 0 < \left|\dfrac{c_{i-1}}{u_{i-1}}\right| < 1$从（**2-22**）和条件（**3**），知

$$|u_i| = |b_i - l_i c_{i-1}| > |b_i| - |a_i| \geq |c_i| \geq |b_i| - |a_i|\left|\frac{c_{i-1}}{u_{i-1}}\right| \quad i = 2,3,\cdots,n-1$$

故 $u_i \neq 0,\ 0 < \left|\dfrac{c_i}{u_i}\right| < 1,\ i = 2,3,\cdots,n-1$。

再应用条件（**2**），得

$$|u_n| = |b_n - l_n c_{n-1}| \geq |b_n| - |a_n|\left|\frac{c_{n-1}}{u_{n-1}}\right| > |b_n| - |a_n| > 0。$$

从而可得 $\det(\boldsymbol{A}) = \det(\boldsymbol{L})\det(\boldsymbol{U}) = u_1 u_2 \cdots u_n \neq 0,$

故方程组 ***Ax=f*** 的解存在唯一。又因为

$$|u_i| = |b_i - l_i c_{i-1}| > |b_i| - |a_i| \geq |c_i| \geq |b_i| - |a_i| \left|\frac{c_{i-1}}{u_{i-1}}\right| \quad i = 2, 3, \cdots, n$$

于是有

$$|b_i| - |a_i| < |u_i| < |b_i| + |a_i|, \text{ 且 } d_i = c_i, \; |l_i| = \left|\frac{a_i}{u_{i-1}}\right|, \; i = 2, 3, \ldots, n$$

即追赶法计算过程中的中间数有界，不会产生大的变化，从而说明它通常是数值稳定的。

定理条件中有 $a_i c_i \neq 0$ ，如果有某个 $a_i = 0$ 或 $c_i = 0$，则可化成低阶方程组求解。

追赶法公式简单，计算量和存储量都小。整个求解过程仅须5$\boldsymbol{n}$-4次乘除和3($\boldsymbol{n}$-1)次加减运算，总共8$\boldsymbol{n}$-7次运算。仅需4个一维数组存储向量$\boldsymbol{c}, \boldsymbol{a}, \boldsymbol{b}$和 $\boldsymbol{f}$

其中$\boldsymbol{d}_i, \boldsymbol{l}_i, \boldsymbol{u}_i$和$\boldsymbol{x}_i$分别存在数组$\boldsymbol{c}, \boldsymbol{a}, \boldsymbol{b}$和 $\boldsymbol{f}$ 中。当A对角占优时，追赶法通常数值稳定。

例 用追赶法解线性方程组$Ax=b$，其中

$$A=\begin{pmatrix}4 & -1 & 0\\ -1 & 4 & -1\\ 0 & -1 & 4\end{pmatrix} \qquad b=\begin{pmatrix}1\\2\\3\end{pmatrix}$$

解 利用公式（**2-22**），$d_i=c_i=-1$依次计算出 u_1,l_2,u_2,l_2,u_3,l_3 诸元素：

$$b_1=u_1=4,\ l_2=\frac{a_2}{u_1}=0.25,\ u_2=b_2-l_2c_1=4-(-0.25)\times(-1)=3.75$$

$$l_3=\frac{a_3}{u_2}=\frac{-1}{3.75}=-0.2667，\ u_3=b_3-l_3c_2=4-0.2667=3.7331$$

$$\boldsymbol{L}=\begin{pmatrix}1 & 0 & 0\\ -0.25 & 1 & 0\\ 0 & -0.2667 & 1\end{pmatrix}\qquad \boldsymbol{U}=\begin{pmatrix}4 & -1 & 0\\ 0 & 3.75 & -1\\ 0 & 0 & 3.7333\end{pmatrix}$$

再利用（**2-23**），求下三角线性方程组$\boldsymbol{Ly=b}$的解，即得

$$y_1=1,\ y_2=f_2-l_2\cdot y_1=3+0.25=3.25,$$

$$y_3=f_3-l_3\cdot y_2=2+0.2667\times 3.25=2.8667,$$

$$\boldsymbol{y}=(1,3.25,\ 2.8667)^T。$$

再利用（**2-24**）求上三角线性方程组$\boldsymbol{Ux=y}$的解，即得

$$x_3=\frac{y_3}{u_3}=0.7679,\quad x_2=\frac{y_2-c_2\cdot x_3}{u_2}=1.0714,$$

$$x_1=\frac{y_1-c_1\cdot x_2}{u_1}=0.5179,\quad \boldsymbol{x}=(0.7679,1.0714,\ 0.5179)^T。$$

2.1.5 条件数与方程组的性态

考虑线性方程组

$$\begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8.00001 \end{pmatrix}$$

它有准确解为： $x=(1,1)^T$ 。

如果方程组的系数矩阵以及右端项发生微小的变化，得

$$\begin{pmatrix} 2 & 6 \\ 2 & 5.99999 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 8 \\ 8.00002 \end{pmatrix}$$

它有准确解： $x=(10,-2)^T$ ，可以看出，方程组的解变化非常大。

定义 如果线性方程组$Ax=b$中，A或b的元素的微小变化，就会引起方程组解的巨大变化，则称方程组为“病态”方程组，矩阵A称为“病态”矩阵. 否则称方程组为“良态”方程组，矩阵A称为“良态”矩阵。

我们需要一种能刻画矩阵和方程组“病态”标准的量。

求解$Ax=b$时，A 和b的误差对解x有何影响？

设A精确，b有误差 δb ，得到的解为 $x+\delta x$，即

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

绝对误差放大因子

$$\Rightarrow \quad \delta x = A^{-1}\delta b \quad \Rightarrow \quad \|\delta x\| \le \|A^{-1}\|\cdot\|\delta b\|$$

又 $\|b\| = \|Ax\| \le \|A\|\cdot\|x\|$

相对误差放大因子

相对误差

$$\frac{\|\delta x\|}{\|x\|} \le \|A\|\cdot\|A^{-1}\|\cdot\frac{\|\delta b\|}{\|b\|}$$

定义 设A为非奇异矩阵，$\|\cdot\|$为矩阵的算子范数，则称 $\mathrm{cond}(A)=\|A\|\|A^{-1}\|$为矩阵$A$的条件数。

常用的条件数为：

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

$$\mathrm{cond}_1(A)=\|A\|_1\|A^{-1}\|_1$$

$$\mathrm{cond}_2(A)=\|A\|_2\|A^{-1}\|_2=\sqrt{\frac{\lambda_{\max}(A^HA)}{\lambda_{\min}(A^HA)}}$$

分别称为矩阵A的∞−条件数、1−条件数和2−条件数。

注意，由 $\boldsymbol{A}^H\boldsymbol{A}=\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{A}^H\boldsymbol{A}=\boldsymbol{A}^{-1}(\boldsymbol{A}\boldsymbol{A}^H)\boldsymbol{A}$

$$\begin{aligned}\det(\lambda\boldsymbol{I}-\boldsymbol{A}^{-1}(\boldsymbol{A}\boldsymbol{A}^H)\boldsymbol{A}) &= \det(\boldsymbol{A}^{-1}(\lambda\boldsymbol{I}-(\boldsymbol{A}\boldsymbol{A}^H))\boldsymbol{A})\\ &= \det(\boldsymbol{A}^{-1})\cdot\det(\lambda\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^H)\cdot\det(\boldsymbol{A})\\ &= \det(\lambda\boldsymbol{I}-\boldsymbol{A}\boldsymbol{A}^H)\end{aligned}$$

则 $\lambda_{\max}(\boldsymbol{A}^H\boldsymbol{A})=\lambda_{\max}(\boldsymbol{A}\boldsymbol{A}^H)$，$\quad \|\boldsymbol{A}\|_2^2=\lambda_{\max}(\boldsymbol{A}^H\boldsymbol{A})$

$$\begin{aligned}\|\boldsymbol{A}^{-1}\|_2^2 &= \lambda_{\max}((\boldsymbol{A}^{-1})^H\boldsymbol{A}^{-1})=\lambda_{\max}((\boldsymbol{A}^{\mathrm{H}})^{-1}\boldsymbol{A}^{-1})\\ &= \lambda_{\max}((\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1})=\lambda_{\max}((\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})^{-1})=\lambda_{\min}^{-1}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\end{aligned}$$

故

$$\operatorname{cond}_2(\boldsymbol{A})=\|\boldsymbol{A}\|_2\|\boldsymbol{A}^{-1}\|_2=\sqrt{\frac{\lambda_{\max}(\boldsymbol{A}^H\boldsymbol{A})}{\lambda_{\min}(\boldsymbol{A}^H\boldsymbol{A})}}$$

矩阵的条件数具有如下的性质：

(1) $\mathrm{cond}(\boldsymbol{A}) \geq 1$

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

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

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

(3) $\mathrm{cond}(\alpha A)=\mathrm{cond}(A)$， $\alpha \neq 0$ ， $\alpha \in \mathbf{R}$

$$\mathrm{cond}(\alpha A)=\|\alpha A\|\cdot\|(\alpha A)^{-1}\|=|\alpha|\cdot\|A\|\cdot\frac{1}{|\alpha|}\cdot\|A^{-1}\|$$

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

(4) 如果为U酉（正交）矩阵，则

$$\mathrm{cond}_2(U)=1$$

$$\mathrm{cond}_2(UA)=\mathrm{cond}_2(AU)=\mathrm{cond}_2(A)$$

$\mathrm{cond}(A)$ 越大，解的相对误差界可能越大，A对求解线性方程组来说就越可能呈现病态。

但 $\mathrm{cond}(A)$多大A算病态，通常没有具体的定量标准； $\mathrm{cond}(A)$ 越小，解的相对误差界越小，反之，呈现良态。

下面给出两个与条件数有关的定理

n阶Hilbert矩阵

$$H_n=(h_{ij})_{n\times n}=\left(\frac{1}{i+j-1}\right)_{n\times n}=\begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \cdots & \frac{1}{2n-1} \end{pmatrix} \quad i,j=1,2,\cdots,n$$

$$\mathrm{cond}(\boldsymbol{H}_4)=1.5514\times 10^4$$

$$\mathrm{cond}(\boldsymbol{H}_6)=1.4951\times 10^7$$

$$\mathrm{cond}(\boldsymbol{H}_8)=1.525\times 10^{10}$$

Hilbert矩阵常常出现在数据拟合和函数逼近的研究中。

与条件数有关的

一个数值例子、两个定理

一个数值例子：

在前面的例子中取 $\delta b=(0,0.00001)^T,\delta A=O_{2\times2}$。

我们观察 δb 对$\boldsymbol{x}$的影响， 由

$$A=\begin{pmatrix}2 & 6\\ 2 & 6.00001\end{pmatrix}\quad \text{易求，}\quad A^{-1}=\begin{pmatrix}300000.5 & -300000\\ -100000 & 100000\end{pmatrix}$$

则$\boldsymbol{A}$的条件数为：

$$\begin{aligned}\operatorname{cond}_\infty(\boldsymbol{A})&=\|A\|_\infty\|A^{-1}\|_\infty\\ &=8.00001\times600000.5\\ &\approx4800010\approx4.8\times10^6\end{aligned}$$

则线性方程组的相对误差界为：

$$\frac{\|\boldsymbol{\delta x}\|_\infty}{\|\boldsymbol{x}\|_\infty} \le \mathrm{cond}_\infty(A)\frac{\|\boldsymbol{\delta b}\|_\infty}{\|\boldsymbol{b}\|_\infty} \approx 4.8\times10^6\times\frac{0.00001}{8}$$

$$\approx 4.8\times10^6\times0.125\times10^{-4}$$

$$\approx 6 \approx 600\%$$

可见，右端向量$\boldsymbol{b}$其分量百分之一的变化，可能引起解向量$\boldsymbol{x}$百分之六百的变化。

这说明矩阵A是严重病态矩阵，相应的线性方程组是病态方程组。

系数矩阵和右端项的扰动对解的影响

定理1 设 $Ax=b$，A为非奇异矩阵，b为非零向量且 A 和 b 均有扰动。若A的扰动 δA 非常小，使得 $\|A^{-1}\|\|\delta A\|<1$，则

$$\frac{\|\delta x\|}{\|x\|}\le\frac{\operatorname{cond}(A)}{1-\operatorname{cond}(A)\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right)$$

注：当 $\delta A=O$ 时，上述不等式为： $\dfrac{\|\delta x\|}{\|x\|}\le\operatorname{cond}(A)\dfrac{\|\delta b\|}{\|b\|}$

近似解的余量与它的相对误差间的关系

定理2 设$Ax=b$，A为非奇异矩阵，b为非零向量，则方程组近似解$\tilde{x}$的事后估计式为

$$\frac{1}{\mathrm{cond}(A)}\frac{\|b-A\tilde{x}\|}{\|b\|}\le\frac{\|\tilde{x}-x\|}{\|x\|}\le\mathrm{cond}(A)\frac{\|b-A\tilde{x}\|}{\|b\|}$$

其中称$\|b-A\tilde{x}\|$为近似解$\tilde{x}$的余量，简称余量。

2.1.5 矩阵的*QR*分解

如何利用直接法求解一些病态方程组？

Gauss消去过程实际上是用一系列具有特定结构的单位下三角矩阵将A逐步上三角化的过程．由矩阵的条件数定义可以看出，正交矩阵是性态最好的矩阵，如果我们能用正交矩阵代替**Gauss**消去过程中的单位下三角矩阵，即

$$A = \begin{pmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{pmatrix} \xrightarrow{\text{用正交变换}\ Q_1} \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & \times & \times \end{pmatrix}$$

$$\xrightarrow{\text{用正交变换}\ Q_2} \begin{pmatrix} \times & \times & \times \\ 0 & \times & \times \\ 0 & 0 & \times \end{pmatrix} = U$$

则 $Q_1Q_2A = U$，计算知 $\text{cond}_2(A) = \text{cond}_2(U)$，因此变换后所得的矩阵$U$的条件数不变，故该计算过程具有数值稳定性。

为实现矩阵一般的QR分解，我们引入矩阵

Householder变换矩阵

定义2.4 设$\omega\in\mathbf{R}^n,\omega\neq 0$，称初等矩阵

$$H(\omega)=I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}$$

为**Householder矩阵**（简称H阵），或称其为

Householder变换。

一般的初等矩阵表示为：

$$\boldsymbol{E}(\boldsymbol{u},\boldsymbol{v};\alpha)=\boldsymbol{I}-\alpha\,\boldsymbol{u}\boldsymbol{v}^{H},\ \ \boldsymbol{u},\boldsymbol{v}\in\mathbf{C}^{\mathbf{n}},\ \alpha\in\mathbf{C}$$

此时， $\boldsymbol{u}\boldsymbol{v}^{H}$ 是一个秩1矩阵，即

$$1\le\ \mathbf{rank}(\boldsymbol{u}\boldsymbol{v}^{H})\le\mathbf{rank}(\boldsymbol{u})\ =1$$

其特征值为： $\boldsymbol{v}^{H}u,\ \underbrace{0,\cdots,0}_{n-1}$

事实上， $\left(\boldsymbol{u}\boldsymbol{v}^{H}\right)\boldsymbol{u}=\boldsymbol{u}\left(\boldsymbol{v}^{H}\boldsymbol{u}\right)=\left(\boldsymbol{v}^{H}\boldsymbol{u}\right)\boldsymbol{u}$

即 $\lambda=\boldsymbol{v}^{H}\boldsymbol{u}\neq0$ 是矩阵 $\boldsymbol{u}\boldsymbol{v}^{H}$ 的唯一非零特征值。

显然Householder矩阵矩阵具有如下性质：

(1) $H(\omega)^{\mathrm{T}}=H(\omega)$, 即 H 阵为对称阵；

$$H(\omega)^{T}=\left(I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}\right)^{T}=I-\frac{2}{\omega^{\mathrm{T}}\omega}\left(\omega\omega^{\mathrm{T}}\right)^{T}=I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}$$

(2) $H(\omega)^{\mathrm{T}}H(\omega)=I_n$， 即$H$ 阵为正交阵；

$$H(\omega)^{T}H(\omega)=\left(I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}\right)^{2}$$

$$=I-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}-\frac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}+\left(\frac{2}{\omega^{\mathrm{T}}\omega}\right)^{2}\left(\omega\omega^{\mathrm{T}}\right)\left(\omega\omega^{\mathrm{T}}\right)$$

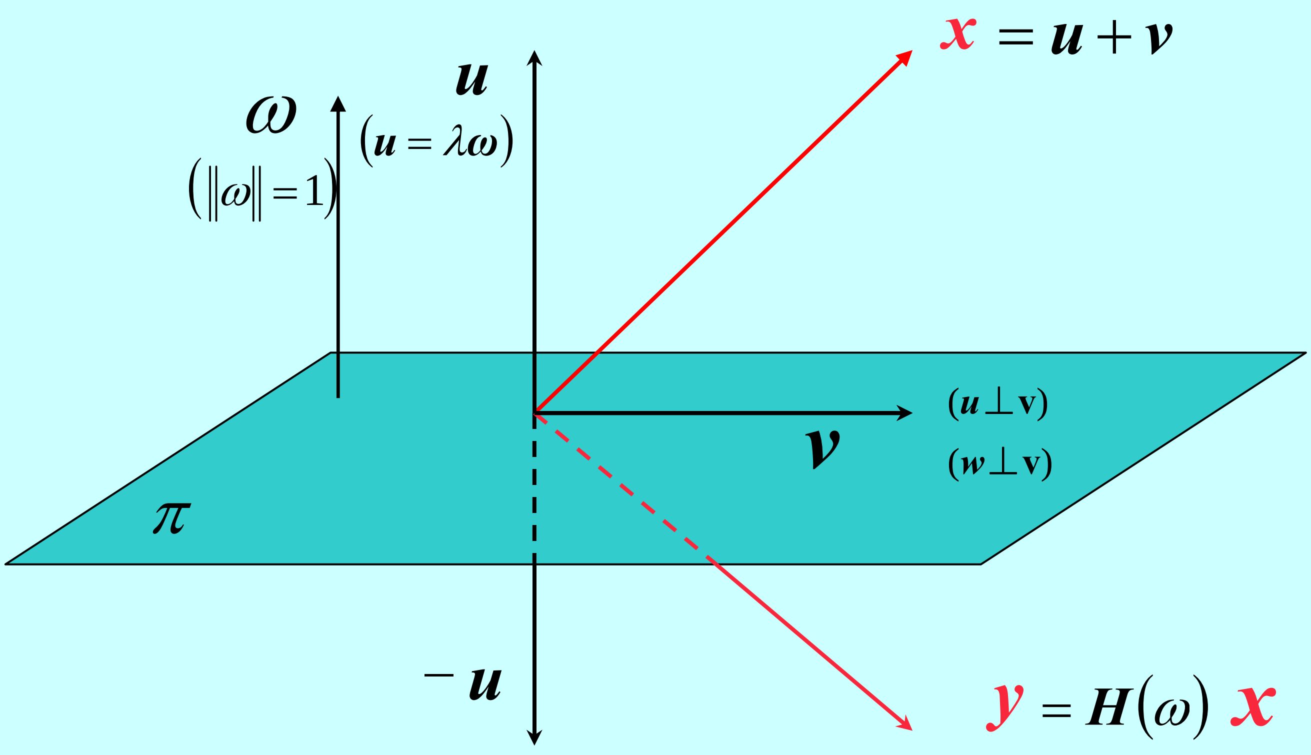
$$=I-\frac{4}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}+\frac{4}{\left(\omega^{\mathrm{T}}\omega\right)^{2}}\left(\omega^{\mathrm{T}}\omega\right)\left(\omega\omega^{\mathrm{T}}\right)=I$$

(3) 如果 $H(\omega)x = y$，则 $\|y\|_2 = \|x\|_2$; 事实上，

$$
\begin{aligned}
\|y\|_2^2 = y^T y &= (H(\omega)x)^T (H(\omega)x) \\
&= x^T \left(H(\omega)^T H(\omega)\right) x \\
&= x^T x = \|x\|_2^2
\end{aligned}
$$

(亦称Householder矩阵为镜面反射变换);

性质3亦称镜面反射变换，其几何意义如下：



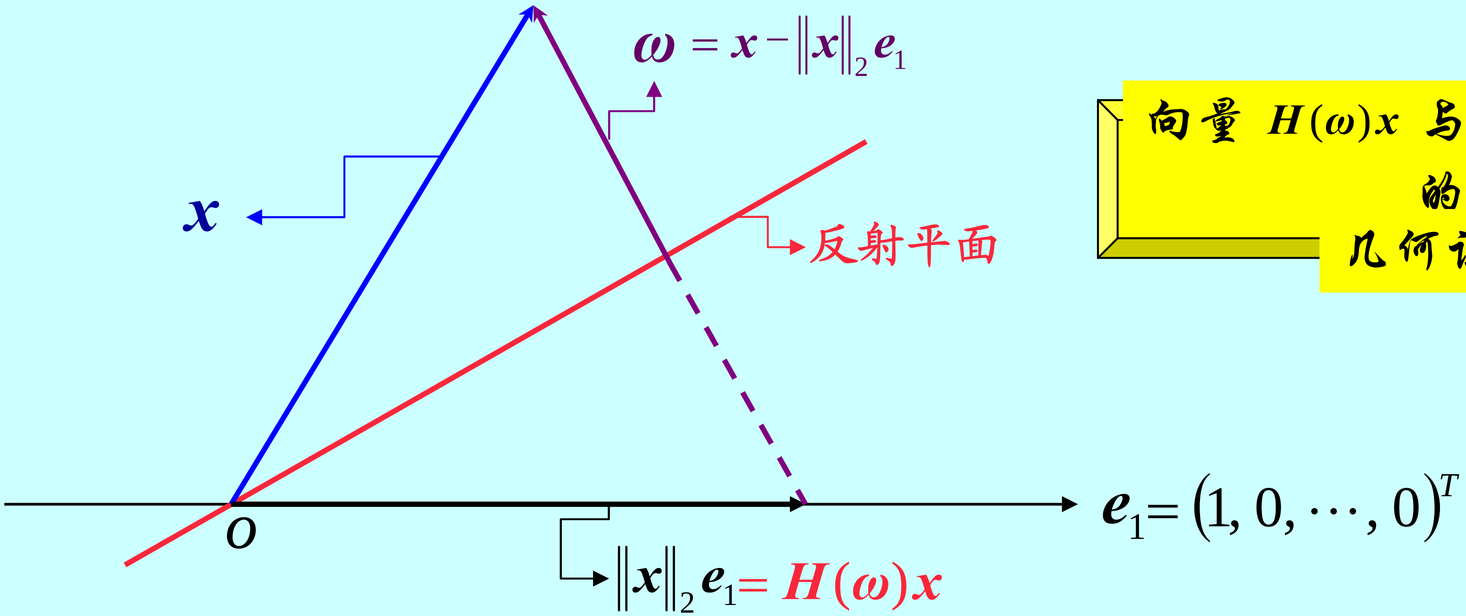
$$H(\omega)x = \left(I - 2\omega\omega^T\right)x = x - 2\omega\omega^T(u+v) = x - 2\omega\left(\omega^T u + \omega^T v\right)$$

$$= u + v - 2\lambda\omega\left(\omega^T\omega\right) = u + v - 2u \quad = -u + v$$

(4) 设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}\in\mathbf{R}^n$ 且 $\boldsymbol{x}\neq\boldsymbol{0}$, 取 $\boldsymbol{\omega}=\boldsymbol{x}-\|\boldsymbol{x}\|_2\boldsymbol{e}_1$

则

$$\boldsymbol{H}(\boldsymbol{\omega})\boldsymbol{x}=\boldsymbol{H}\left(\boldsymbol{x}-\|\boldsymbol{x}\|_2\boldsymbol{e}_1\right)\boldsymbol{x}=\left(\|\boldsymbol{x}\|_2,0,\cdots,0\right)^T=\|\boldsymbol{x}\|_2\boldsymbol{e}_1$$



向量 $H(\omega)x$ 与向量 e_1 共线的几何证明

注释:

$$Ax = b \Leftrightarrow (QR)x = b \Leftrightarrow Rx = Q^T b$$

这种直接法的数值稳定性要比LU分解好，但是QR分解的计算量远远大于LU分解，因此，QR分解只适用于求解病态线性方程组。

下面我们利用一系列H阵将A分解成QR形式。

例 利用Householder变换求A的QR分解，其中

$$A=\begin{pmatrix}0 & 3 & 1\\ 0 & 4 & -2\\ 2 & 1 & 2\end{pmatrix}.$$

解： 将A按列分块为 $A=(\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3)$,其中 $\boldsymbol{a}_1=(0,0,2)^{\mathrm{T}}$,

$\|\boldsymbol{a}_1\|_2=2$, 取 $\omega_1=\boldsymbol{a}_1-\|\boldsymbol{a}_1\|_2\boldsymbol{e}_1=\begin{pmatrix}-2\\0\\2\end{pmatrix}$，则

$$\omega_1^T\omega_1=\begin{pmatrix}-2 & 0 & 2\end{pmatrix}\begin{pmatrix}-2\\0\\2\end{pmatrix}=8,$$ 令

$$\boldsymbol{Q}_1=\boldsymbol{H}(\boldsymbol{\omega}_1)=\boldsymbol{I}-\frac{2}{\boldsymbol{\omega}_1^{\mathrm{T}}\boldsymbol{\omega}_1}\boldsymbol{\omega}_1\boldsymbol{\omega}_1^{\mathrm{T}}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\frac{1}{4}\begin{pmatrix}-2\\0\\2\end{pmatrix}\begin{pmatrix}-2&0&2\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$$

$$\boldsymbol{Q}_1\boldsymbol{A}=\boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{A}=(\boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{a}_1,\boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{a}_2,\boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{a}_3)$$

$$=\begin{pmatrix}2&1&2\\0&4&-2\\0&3&1\end{pmatrix}=\begin{pmatrix}2&\boldsymbol{b}^{\mathrm{T}}\\\boldsymbol{0}&\boldsymbol{A}_2\end{pmatrix}$$

其中

$$\boldsymbol{b}^{\mathrm{T}}=(1,2),\quad \boldsymbol{A}_2=\begin{pmatrix}4 & -2\\ 3 & 1\end{pmatrix}=(\tilde{\boldsymbol{a}}_1,\tilde{\boldsymbol{a}}_2)$$

$$\tilde{\boldsymbol{a}}_1=(4,3)^{\mathrm{T}},\ \tilde{\boldsymbol{a}}_2=(-2,1)^{\mathrm{T}},\ \|\tilde{\boldsymbol{a}}_1\|_2=5,$$ 取

$$\boldsymbol{\omega}_2=\tilde{\boldsymbol{a}}_1-\|\tilde{\boldsymbol{a}}_1\|_2\boldsymbol{e}_1=\begin{pmatrix}4\\3\end{pmatrix}-\begin{pmatrix}5\\0\end{pmatrix}=\begin{pmatrix}-1\\3\end{pmatrix}$$

$$\boldsymbol{\omega}_2^T\boldsymbol{\omega}_2=\|\boldsymbol{\omega}_2\|_2^2=10$$

$$H(\omega_2)=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}-1\\ 3\end{pmatrix}\begin{pmatrix}-1 & 3\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\frac{1}{5}\begin{pmatrix}1 & -3\\ -3 & 9\end{pmatrix}=\frac{1}{5}\begin{pmatrix}4 & 3\\ 3 & -4\end{pmatrix}$$

$$H(\omega_2)A_2=\left(H(\omega_2)\tilde{a}_1,H(\omega_2)\tilde{a}_2\right)$$

$$=\left[\frac{1}{5}\begin{pmatrix}4 & 3\\ 3 & -4\end{pmatrix}\begin{pmatrix}4\\ 3\end{pmatrix},\frac{1}{5}\begin{pmatrix}4 & 3\\ 3 & -4\end{pmatrix}\begin{pmatrix}-2\\ 1\end{pmatrix}\right]=\begin{pmatrix}5 & -1\\ 0 & -2\end{pmatrix}。$$

令

$$\boldsymbol{Q}_2=\begin{pmatrix}1 & \boldsymbol{0}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{H}(\omega_2)\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & \dfrac{4}{5} & \dfrac{3}{5}\\ 0 & \dfrac{3}{5} & -\dfrac{4}{5}\end{pmatrix}$$

则

$$(\boldsymbol{Q}_2\boldsymbol{Q}_1)\boldsymbol{A}=\begin{pmatrix}1 & \boldsymbol{0}^{T}\\ \boldsymbol{0} & \boldsymbol{H}(\omega_2)\end{pmatrix}\begin{pmatrix}2 & \boldsymbol{b}^{T}\\ \boldsymbol{0} & \boldsymbol{A}_2\end{pmatrix}=\begin{pmatrix}2 & \boldsymbol{b}^{T}\\ \boldsymbol{0} & \end{pmatrix}$$

$$=\begin{pmatrix}2 & 1 & 2\\ 0 & 5 & -1\\ 0 & 0 & -2\end{pmatrix}=\boldsymbol{R}$$

令

$$Q^{\mathrm{T}} = Q_2 Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & \frac{3}{5} & -\frac{4}{5} \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ \frac{3}{5} & \frac{4}{5} & 0 \\ -\frac{4}{5} & \frac{3}{5} & 0 \end{pmatrix}$$

从而得到Q和A的QR分解如下：

$$Q = \begin{pmatrix} 0 & \frac{3}{5} & -\frac{4}{5} \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 1 & 0 & 0 \end{pmatrix} \qquad A = QR = \begin{pmatrix} 0 & \frac{3}{5} & -\frac{4}{5} \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 2 \\ 0 & 5 & -1 \\ 0 & 0 & -2 \end{pmatrix}。$$

例　利用**Householder**变换求A的QR分解，其中

$$A=\begin{pmatrix}0&4\\0&0\\5&2\end{pmatrix}.$$

解法之一：

解 首先将A按列分块为 $A=(a_1,a_2)$，其中 $a_1=(0,0,5)^{\mathrm{T}}$，则 $\|a_1\|_2=5$

取

$$\omega_1=a_1-\|a_1\|_2 e_1=\begin{pmatrix}-5\\0\\5\end{pmatrix}，则\ \omega_1^T\omega_1=\begin{pmatrix}-5&0&5\end{pmatrix}\begin{pmatrix}-5\\0\\5\end{pmatrix}=5,$$

$$Q_1=H(\omega_1)=I-\frac{2}{\omega_1^{\mathrm{T}}\omega_1}\omega_1\omega_1^{\mathrm{T}}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\frac{1}{25}\begin{pmatrix}-5\\0\\5\end{pmatrix}\begin{pmatrix}-5&0&5\end{pmatrix}$$

$$=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}-\frac{1}{25}\begin{pmatrix}25&0&-25\\0&0&0\\-25&0&25\end{pmatrix}$$

$$Q_1A=H(\omega_1)A=(H(\omega_1)a_1,H(\omega_1)a_2)=\begin{pmatrix}5&2\\0&0\\0&4\end{pmatrix}=\begin{pmatrix}5&2\\ \boldsymbol{0}&A_2\end{pmatrix},$$ 其中

$$A_2=\begin{pmatrix}0\\4\end{pmatrix}=(\tilde{a}_1),$$ 则 $\|\tilde{a}_1\|_2=4$，取 $\omega_2=\tilde{a}_1-\|\tilde{a}_1\|_2e_1=\begin{pmatrix}0\\4\end{pmatrix}-\begin{pmatrix}4\\0\end{pmatrix}=\begin{pmatrix}-4\\4\end{pmatrix}$

$$\omega_2^T\omega_2=\|\omega_{21}\|_2^2=32,\ H(\omega_2)=\begin{pmatrix}0&1\\1&0\end{pmatrix}=\frac{1}{16}\begin{pmatrix}-4\\4\end{pmatrix}(-4\ \ 4)=\frac{1}{16}\begin{pmatrix}-16&-16\\-16&16\end{pmatrix}$$ 那么有

$$H(\omega_2)A_2=H(\omega_2)\tilde{a}_1=\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}0\\4\end{pmatrix}=\begin{pmatrix}4\\0\end{pmatrix},$$ 令

$$Q_2=\begin{pmatrix}1&\mathbf{0}^{\mathrm{T}}\\ \boldsymbol{0}&H(\omega_2)\end{pmatrix}=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$$ ，则

$$(\boldsymbol{Q}_2\boldsymbol{Q}_1)\boldsymbol{A}=\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{H}(\omega_2)\end{pmatrix}\begin{pmatrix}5 & 0\\ \boldsymbol{0} & \boldsymbol{A}_2\end{pmatrix}=\begin{pmatrix}5 & 0\\ 0 & \end{pmatrix}=\begin{pmatrix}5 & 2\\ 0 & 4\\ 0 & 0\end{pmatrix}=\begin{pmatrix}\bar{\boldsymbol{R}}\\ \boldsymbol{0}\end{pmatrix}$$

$$\boldsymbol{Q}^{\mathrm{T}}=\boldsymbol{Q}_2\boldsymbol{Q}_1=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix}$$，则

$$\boldsymbol{Q}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$$，从而得 $\boldsymbol{A}=QR$ ，即

$$\begin{pmatrix}0&4\\0&0\\5&2\end{pmatrix}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}\begin{pmatrix}5&2\\0&4\\0&0\end{pmatrix}$$

解法之二：

解： 首先，取正交矩阵 $Q_1=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$，左乘$A$，得

$$Q_1A=\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}\begin{pmatrix}0&4\\0&0\\5&2\end{pmatrix}=\begin{pmatrix}5&2\\0&0\\0&4\end{pmatrix}$$，再取正交矩阵 $Q_2=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}$

左乘$\boldsymbol{Q_1A}$，得

$$Q_2(Q_1A)=\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}5&2\\0&0\\0&4\end{pmatrix}=\begin{pmatrix}5&2\\0&4\\0&0\end{pmatrix}=\begin{pmatrix}\bar{\boldsymbol{R}}\\\boldsymbol{0}\end{pmatrix}$$

令

$$Q^{\mathrm{T}} = Q_2 Q_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\text{，则}$$

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\text{，从而得 } A = QR\text{，即}$$

$$\begin{pmatrix} 0 & 4 \\ 0 & 0 \\ 5 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 5 & 2 \\ 0 & 4 \\ 0 & 0 \end{pmatrix}$$

2.2 特殊矩阵的特征系统

本节将介绍理论上和特征系统计算上非常重要的矩阵分解，即**Schur**分解.

定理 2.7 (Schur定理) **设** $A \in \mathbf{C}^{n\times n}$，则存在酉阵 $U \in \mathbf{C}^{n\times n}$ 使得

$$A = URU^H$$

其中 $R \in \mathbf{C}^{n\times n}$ 为上三角矩阵。

$A = URU^H$ 也称为**矩阵的Schur分解**

证：对矩阵的阶数$\boldsymbol{n}$用数学归纳法证明。

n=1 时，定理显然成立。即 $\boldsymbol{A}=(a)=\boldsymbol{R}$，$\boldsymbol{U}=(1)$

设 $n=k$ 时，定理成立，即 对于$\boldsymbol{n}$阶方阵$\boldsymbol{A}$，存在$\boldsymbol{n}$阶酉阵$\boldsymbol{U}$，使得 $\boldsymbol{A}=\boldsymbol{URU}^{H}$ 现在证明，当 $n=k+1$ 时，定理仍成立。

记 λ_1 为方阵 $A \in C^{(k+1)\times(k+1)}$的一个特征值，于是存在 $\boldsymbol{u}_1 \in \mathbf{C}^{k+1}$ 为其对应于 λ_1 的特征向量。我们可以特别选取 $\boldsymbol{u}_1$，使得 $\|\boldsymbol{u}_1\| = 1$ 。再于 $\mathrm{Span}\{\boldsymbol{u}_1\} \subset \mathbf{C}^{k+1}$ 的正交补空间上选取标准正交基 $\boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1} \in \mathbf{C}^{k+1}$ ，使

$\boldsymbol{u}_1$ 构成一组标准正交基。

注

若 U 是 C^n 中的一子空间，则存在 U 的唯一的正交补空间 $V \subset C^n$，$C^n = U \oplus V$，且 $U \perp V$ 即 $\forall \alpha \in U, \forall \beta \in V$，都有 $(\alpha, \beta) = 0$。

例如，R³ 上有 $U = \mathbf{span}\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ 和 $\mathbf{V} = \mathbf{span}\left\{ e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

则 $U \perp V$，$R^3 = U \oplus V$。 U 与 V 互为R³上的正交补空间，且唯一。e_1,和 e_2, e_3 构成R³上的一组标准正交基。

回到证明

记 $U_1=(u_1,u_2,\cdots,u_{k+1})$，显然 $U_1\in \mathbf{C}^{(k+1)\times(k+1)}$ 为酉阵，

由于 $Au_1=\lambda_1 u_1$，则

$$U_1^H A u_1=\begin{pmatrix} u_1^H \\ u_2^H \\ \vdots \\ u_{k+1}^H \end{pmatrix} A u_1=\begin{pmatrix} u_1^H \\ u_2^H \\ \vdots \\ u_{k+1}^H \end{pmatrix}\lambda_1 u_1=\lambda_1\begin{pmatrix} (u_1,u_1) \\ (u_2,u_1) \\ \vdots \\ (u_{k+1},u_1) \end{pmatrix}=\lambda_1\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

由归纳法假设， $n=k$ 时结论成立， 即存在酉阵 $\tilde{\boldsymbol{U}}_2 \in \mathbf{C}^{k\times k}$，使得

$$\tilde{\boldsymbol{U}}_2^H \boldsymbol{A}_1 \tilde{\boldsymbol{U}}_2 = \begin{pmatrix} \lambda_2 & r_{23} & \cdots & r_{2k+1} \\ 0 & \lambda_3 & \cdots & r_{3k+1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_{k+1} \end{pmatrix} = R_1$$

上三角矩阵

$$\boldsymbol{A}_1 = \tilde{\boldsymbol{U}}_2 \boldsymbol{R}_1 \tilde{\boldsymbol{U}}_2^H \qquad (2\text{-}37)$$

其中 $R_1 \in \mathbf{C}^{k\times k}$ 为上三角矩阵。

现取 $U_2 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{U}_2 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & \boldsymbol{0}^T \\ \boldsymbol{0} & \tilde{U}_2 \end{pmatrix}$，则

$$\boldsymbol{U}_2^H\left(\boldsymbol{U}_1^H\boldsymbol{A}\,\boldsymbol{U}_1\right)\boldsymbol{U}_2=\boldsymbol{U}_2^H\begin{pmatrix}\lambda_1 & r_{12} & \cdots & r_{1k+1}\\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}\boldsymbol{U}_2$$

$$=\begin{pmatrix}\lambda_1 & r_{12} & \cdots & r_{1k+1}\\ 0 & & & \\ \vdots & & \boldsymbol{U}_2^H A_1\boldsymbol{U}_2^H & \\ 0 & & & \end{pmatrix}=\begin{pmatrix}\lambda_1 & r_{12} & \cdots & r_{1k+1}\\ 0 & \lambda_2 & \cdots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ 0 & \cdots & 0 & \lambda_{k+1}\end{pmatrix}=\begin{pmatrix}\lambda_1 & \tilde{\boldsymbol{c}}^T\\ \boldsymbol{0} & \boldsymbol{R}_1\end{pmatrix}=R$$

取 $U = U_1U_2$，显然 $U \in \mathbf{C}^{(k+1)\times(k+1)}$ 为酉阵，且

$U^H AU = R$，$R \in \mathbf{C}^{(k+1)\times(k+1)}$ 为上三角矩阵，则

$$A = URU^H \qquad (2\text{-}35)$$

成立，即证明了 $n = k+1$ 时定理成立。 **证毕**

称（**2-35**）为矩阵的**Schur分解**。

下面给出Schur定理的几点注记

1. 在矩阵的Schur分解中，由于$\boldsymbol{A}$和$\boldsymbol{R}$是酉相似的（$\boldsymbol{A}=\boldsymbol{URU}^H$, $U^H AU=R$），因此具有相同的特征值，而上三角矩阵的特征值即为其对角元，因此，**Schur**定理还可以表示为：

任意$\boldsymbol{n}$阶方阵酉相似于一个以其特征值为对角元的上三角矩阵$\boldsymbol{R}$

2. 通常称$\boldsymbol{R}$为$\boldsymbol{A}$的**Schur**标准型，在理论上我们得到了矩阵的特征值。但是，特征值的计算一般必须采用迭代法，通常我们无法在有限步内，准确地得到。

3. 由于实矩阵A的特征值可能是一个复数，如

$$A=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},\quad \det(\lambda I-A)=\begin{vmatrix}\lambda & 1\\ -1 & \lambda\end{vmatrix}=\lambda^2+1=0,\ \Rightarrow \begin{cases}\lambda_1=i\\ \lambda_2=-i\end{cases}$$

因此即使矩阵A是实矩阵，一般地， **Schur**分解中的 U，R有可能是复的。(因为当 A为实矩阵, U为正交矩阵时，那么, $U^T AU=R$ 也为实矩阵, 若R为上三角矩阵, 则R的对角元素(实的)为A的特征值, 这与 A的特征值可能是一个复数相矛盾)。

4. U的列向量未必都是A的特征向量, 尽管其第一列时A的特征向量。

矩阵的Schur分解还有许多应用, 在范数的性质的研究中, 用它可以证明如下定理.

定理2.8　设A为n阶方阵，$\varepsilon > 0$, 则在$\mathbf{C}^{n\times n}$中存在一种矩阵范数$\|\cdot\|_M$（依赖矩阵A和常数ε），满足$\|I_n\|_M=1$，并且

$$\|A\|_M \le \rho(A) + \varepsilon. \qquad \text{(2-41)}$$

证明:(构造性证明)根据Schur定理, 存在n阶酉阵U, 使得

$$R = U^H A U$$

为上三角矩阵, 其中对角元即为矩阵A的特征值。

记矩阵R的上三角元素为$r_{ij}\ (j \geq i)$, 对任意的$\varepsilon > 0$

取

$$\delta = \min\left\{1, \frac{\varepsilon}{(n+1)\max\limits_{1\leq i<j\leq n}|r_{ij}|}\right\}, \quad D = \begin{pmatrix} 1 & & & & \\ & \delta & & & \\ & & \delta^2 & & \\ & & & \ddots & \\ & & & & \delta^{n-1} \end{pmatrix}$$

$$\boldsymbol{R}=\begin{pmatrix} r_{11} & r_{12} & \cdots & \cdots & r_{1n} \\ & r_{22} & r_{23} & \cdots & r_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & r_{n-1n} \\ & & & & r_{nn} \end{pmatrix},\ \boldsymbol{RD}=\begin{pmatrix} r_{11} & \delta r_{12} & \delta^2 r_{13} & \cdots & \delta^{n-1} r_{1n} \\ & \delta r_{22} & \delta^2 r_{23} & \cdots & \delta^{n-1} r_{2n} \\ & & \delta^2 r_{33} & \cdots & \delta^{n-1} r_{3n} \\ & & & \ddots & \vdots \\ & & & & \delta^{n-1} r_{nn} \end{pmatrix}$$

$$\boldsymbol{D}^{-1}\boldsymbol{RD}=\begin{pmatrix} 1 & & & & \\ & \frac{1}{\delta} & & & \\ & & \frac{1}{\delta^2} & & \\ & & & \ddots & \\ & & & & \frac{1}{\delta^{n-1}} \end{pmatrix} \boldsymbol{RD} =\begin{pmatrix} r_{11} & r_{12}\delta & r_{13}\delta^2 & \cdots & r_{1n}\delta^{n-1} \\ & r_{22} & r_{23}\delta & \cdots & r_{2n}\delta^{n-2} \\ & & r_{33} & \cdots & \vdots \\ & & & \ddots & r_{n-1n}\delta \\ & & & & r_{nn} \end{pmatrix}$$

于是，

$$\left\|D^{-1}U^{H}AUD\right\|_1=\left\|D^{-1}RD\right\|_1=\max_{1\le i\le n}\sum_{i\le j}^{n}\left|r_{ij}\delta^{i-1}\right| \le \max_{1\le i\le n}\left(\left|r_{ii}\right|+\sum_{i<j}^{n}\delta^{i-1}\left|r_{ij}\right|\right)$$

$$\le \max_{1\le i\le n}\left|\lambda_i\right|+\max_{1\le i<j\le n}\left|r_{ij}\right|(1+\delta+\ldots+\delta^{n-2})\delta$$

$$\le \rho(A)+(n-1)\max_{1\le i<j\le n}\left|r_{ij}\right|\delta \le \rho(A)+\varepsilon$$

现记 $\left\|A\right\|_M=\left\|D^{-1}U^{H}AUD\right\|_1=\left\|(UD)^{-1}A(UD)\right\|_1=\left\|MAM^{-1}\right\|_1$

其中 $M=(UD)^{-1}$，A为任意n阶方阵, 且满足

$$\left\|I\right\|_M=\left\|M\,I\,M^{-1}\right\|_1=\left\|M\,M^{-1}\right\|_1=\left\|I\right\|_1=1 \quad 。$$

我们已经验证 $\|\cdot\|_M$ 为n阶方阵的一种矩阵范数（见习题）

Schur定理还可以表示为：任意n阶方阵酉相似于一个以其特征值为对角元的上三角矩阵$\boldsymbol{R}$。$\boldsymbol{R}$ 通常称为$\boldsymbol{A}$的Schur**标准型**。

定义2.5 设 $A \in \mathbf{C}^{n\times n}$，若 $A^H A = AA^H$

则称矩阵A为**正规矩阵**。

常见正规矩阵的有：

Hermite阵：$A^H = A$ 实对称矩阵：$A^T = A$

斜Hermite阵：$A^H = -A$ 实反对称矩阵：$A^T = -A$

酉阵：$A^H A = AA^H = I$

正交矩阵：$A^T A = AA^T = I$

以上矩阵均为正规矩阵。

推论 2.1 设 A为n阶方阵， 则A为正规矩阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中$D \in \mathbf{C}^{n\times n}$ 为对角矩阵。

证明：充分性　$(\Leftarrow)$

由于$A=UDU^H$，则

$$A^H A = (UDU^H)^H (UDU^H)$$

$$= UD^H U^H UDU^H = U(D^H D)U^H$$

$$AA^H = (UDU^H)(UDU^H)^H$$

$$= UDU^H UD^H U^H = U(DD^H)U^H$$

而

$$\boldsymbol{D}^H\boldsymbol{D}=\begin{pmatrix}\bar{d}_1 & & \\ & \ddots & \\ & & \bar{d}_n\end{pmatrix}\begin{pmatrix}d_1 & & \\ & \ddots & \\ & & d_n\end{pmatrix}=\begin{pmatrix}|d_1|^2 & & \\ & \ddots & \\ & & |d_n|^2\end{pmatrix}$$

$$=\begin{pmatrix}d_1 & & \\ & \ddots & \\ & & d_n\end{pmatrix}\begin{pmatrix}\bar{d}_1 & & \\ & \ddots & \\ & & \bar{d}_n\end{pmatrix}=\boldsymbol{D}\boldsymbol{D}^H$$

故

$$\boldsymbol{A}^H\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^H$$

即$\boldsymbol{A}$为正规矩阵.

必要性 $(\Rightarrow)$ 由Schur分解定理知，$\boldsymbol{A}=\boldsymbol{URU}^{H}$

$\boldsymbol{U}\in\mathbf{C}^{n\times n}$ 为酉阵, $\boldsymbol{R}$为上三角阵。那么，由假设知$\boldsymbol{A}$为正规矩阵，即 $\boldsymbol{A}^{H}\boldsymbol{A}=\boldsymbol{A}\boldsymbol{A}^{H} \Rightarrow \boldsymbol{R}^{H}\boldsymbol{R}=\boldsymbol{R}\boldsymbol{R}^{H}$，即

$\boldsymbol{R}$为正规矩阵。而上三角阵$\boldsymbol{R}$正规矩阵$\Leftrightarrow$ $\boldsymbol{R}$为对角矩阵。

事实上, 设

$$\boldsymbol{R}=\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix} \quad \boldsymbol{R}^{H}=\begin{pmatrix} \bar{r}_{11} & & & \\ \bar{r}_{12} & \bar{r}_{22} & & \\ \vdots & \vdots & \ddots & \\ \bar{r}_{1n} & \bar{r}_{2n} & \cdots & \bar{r}_{nn} \end{pmatrix}$$

再注意到,

$$R^H R=\begin{pmatrix} |r_{11}|^2 & * & \cdots & * \\ * & |r_{22}|^2+|r_{12}|^2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & \sum_{i=1}^{n}|r_{in}|^2 \end{pmatrix} = RR^H=\begin{pmatrix} \sum_{i=1}^{n}|r_{1i}|^2 & * & \cdots & * \\ * & \sum_{i=2}^{n}|r_{2i}|^2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & |r_{nn}|^2 \end{pmatrix}$$

从而可得:

$$|r_{11}|^2=|r_{11}|^2+|r_{12}|^2+\cdots+|r_{1n}|^2 \Rightarrow r_{1j}=\bar{r}_{1j}=0,\quad j=2,\cdots,n$$

$$|r_{12}|^2+|r_{22}|^2=|r_{22}|^2+|r_{23}|^2+\cdots+|r_{2n}|^2 \Rightarrow r_{2j}=\bar{r}_{2j}=0,\ j=3,\cdots,n$$

总之有: $r_{ij}=\bar{r}_{ij}=0, 1\le i<j\le n$。 即$\boldsymbol{R}$为对角矩阵。

推论 2.2 设 A为n阶方阵，则A为Hermite矩阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中$D \in \mathrm{R}^{n\times n}$ 为对角矩阵。

证明：由推论2.1，存在 n 阶酉阵 U，使得

$A=UDU^H$ 其中D为n阶对角阵。 而 $A^H=A$，则可得

$D^H=D$，即D的对角元素均为实数。注意到，由

$$d_i = a + ib = a - ib = \overline{d_i} \Rightarrow b = 0,$$

即

$$d_i = \overline{d_i} = a$$

从而D为n阶实对角阵。

推论 2.2’ 设 A为n阶方阵，则A为斜Hermite矩阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中D是对角矩阵，且对角元素为纯虚数。

证明：由推论2.1,存在 n 阶酉阵 U，使得 $A=UDU^H$ 其中D为n阶对角阵。而 $A^H=-A$，则可得 $D^H=-D$，即D的对角元素均为实数。注意到，由

$$d_i=a+ib=-(a-ib)=-\overline{d_i}\Rightarrow a=0,$$

即

$$d_i=\overline{d_i}=ib$$

从而D为n阶纯虚对角阵。

推论 2.3 设 A为n阶方阵， 则A为酉阵的充分必要条件是存在n阶酉阵U，使得

$$A = UDU^H$$

其中$D \in \mathbf{C}^{n\times n}$ 为对角矩阵，其对角元的模均为1。

证明：由推论2.1，存在n阶酉阵U使得 $A=UDU^H$

而 A为酉阵，则有 $A^H A = AA^H = I$

$$\Rightarrow U\left(DD^H\right)U^H = U\left(D^H D\right)U^H = I$$

$$\Rightarrow \quad DD^H = D^H D = I$$

即

$$D^H D = DD^H = \begin{pmatrix} |d_{11}|^2 & & \\ & \ddots & \\ & & |d_{nn}|^2 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \Rightarrow |d_{ii}| = 1$$

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

即D为n阶对角阵，其对角元的模均为1.

矩阵的基本分类

Ⅰ 在正规矩阵的集合中，特征值均为实数的子集为Hermite矩阵的集合；矩阵的特征值的模均为1的子集为酉阵的集合

Ⅱ 一般矩阵 $\supset$ 可对角化矩阵

$\supset$ 正规矩阵

$$\supset \left\{ \begin{array}{r} \text{Hermite矩阵} \supset \text{实对称矩阵} \\ \text{酉矩阵} \supset \text{实正交矩阵} \end{array} \right\}$$

定理 设A为n阶方阵，则存在n阶酉阵U和V，使得

$$\|UA\|_F=\|AV\|_F=\|UAV\|_F=\|A\|_F$$

称之为F-范数的酉不变性。

$$\|UA\|_F^2=\mathrm{tr}\left((UA)^H(UA)\right)=\mathrm{tr}\left(A^HU^HUA\right)=\mathrm{tr}\left(A^HA\right)=\|A\|_F^2$$

$$\|AV\|_F^2=\mathrm{tr}\left((AV)^H(AV)\right)=\mathrm{tr}\left((VA)(AV)^H\right)$$

$$=\mathrm{tr}\left(AV^HVA^H\right)=\mathrm{tr}\left(AA^H\right)=\|A\|_F^2$$

最后，

$$\|UAV\|_F=\|AV\|_F=\|A\|_F$$

证明 Schur不等式:

$$\sum_{i=1}^{n}\left|\lambda_i\right| \le \sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2$$

其中 λ_i 为A的特征值, 并且Schur不等式等号成立的充分必要条件是A为正规矩阵。

证：根据**Schur**定理, 存在n阶酉阵U使得 $A = URU^H$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}\right|^2 = \left\|A\right\|_F^2 = \left\|URU^H\right\|_F^2 = \left\|R\right\|_F^2 \ge \sum_{i=1}^{n}\left|r_{ii}\right|^2 = \sum_{i=1}^{n}\left|\lambda_i\right|^2 = \sum_{i=1}^{n}\sum_{j=i+1}^{n}\left|r_{ij}\right|^2$$

要使得等号成立, 只需 $r_{ij} = 0, 1 \le i < j \le n$ 即D为阶对角阵,

则由推论2. 1, 可知其充分必要条件是A为正规矩阵。

2.4 矩阵的奇异值分解

对于方阵, 利用其特征值和特征向量可以刻画矩阵的结构。对长方阵情形, 这些方法已经不适用. 而推广的特征值--矩阵的奇异值分解理论能改善这种情况。利用奇异值和奇异向量不仅可以刻画矩阵的本身结构, 而且还可以进一步刻画线代数方程组的解的结构, 是构造性的研究线代数问题的有利的工具。

定义2.10 设 $A \in \mathbf{C}^{m\times n}$, $k = \min(m, n)$

Hermite半正定矩阵 $A^H A$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$$

称非负实数

$$\sigma_i(A) = \sqrt{\lambda_i} \quad i = 1, 2, \cdots k$$

为矩阵A的奇异值。

矩阵A的奇异值满足如下性质：

定理 2.13　设 A、$B \in \mathbf{C}^{m\times n}$, 如果存在$m$阶、$n$阶酉阵$U$、$V$, 使得 $A = UBV^H$, 则矩阵A、B的奇异值相同。

证： 由　$U^H AV = B$，则有

$$B^H B = \left(U^H AV\right)^H \left(U^H AV\right) = V^H A^H \left(UU^H\right) AV$$

$$= V^H \left(A^H A\right) V$$

即 $B^H B$ 与 $A^H A$相似, 故它们具有相同的特征值, 进而命题得证。

定理 2.14 设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$, 且其秩rank($\boldsymbol{A}$)=$r$, 则存在$\boldsymbol{m}$阶、$\boldsymbol{n}$阶酉阵$\boldsymbol{U}$、$\boldsymbol{V}$使得

$$\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{V}^{H} \qquad (2\text{-}47)$$

其中

$$\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r\end{pmatrix}$$

$\sigma_i\,(i=1,2,\cdots,r)$ 为矩阵$\boldsymbol{A}$的非零奇异值。

（2-47）称为矩阵A 的奇异值分解，亦称为矩阵A的满的奇异值分解。定理2.14简称SVD定理。

关系式亦可写为

$$A = U_1 \Sigma V_1^H$$

其中 $U_1 \in \mathbf{C}^{m\times r}, V_1 \in \mathbf{C}^{n\times r}$。

并称它为矩阵A 约化的奇异值分解。

例1 求矩阵 $A=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}$ 的奇异值分解。

解 求解次序为：Σ, V, V_1, U_1, U。 计算矩阵

$$A^HA=\begin{pmatrix}1&0&0\\0&1&0\\1&1&0\end{pmatrix}\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}=\begin{pmatrix}1&0&1\\0&1&1\\1&1&2\end{pmatrix}$$

令

$$\det(\lambda I - A^HA)=\begin{vmatrix}\lambda-1&0&-1\\0&\lambda-1&-1\\-1&-1&\lambda-2\end{vmatrix}=(\lambda-1)^2(\lambda-2)-2(\lambda-1)$$

$$=(\lambda-1)[\lambda^2-3\lambda+2-2]=\lambda(\lambda-1)(\lambda-3)=0$$

则 $A^H A$ 的特征值和A的奇异值分别为

$$\lambda_1 = 3,\ \lambda_2 = 1,\ \lambda_3 = 0;\ \sigma_1 = \sqrt{3},\ \sigma_2 = 1,\ \sigma_3 = 0$$

所以

$$\Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix}$$

再求出V，注意到V满足： $V^H(A^H A)V = \begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$ 故可知V的列是$A^H A$的特征值所对应的特征向量，所以只需求解如下的方程组：

$$\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{aligned} 2x_1 \quad\ - x_3 &= 0 \\ 2x_2 - x_3 &= 0 \\ x_1 + x_2 - x_3 &= 0 \end{aligned} \Rightarrow p_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{aligned} x_3 &= 0 \\ x_3 &= 0 \\ x_1+x_2+x_3 &= 0 \end{aligned} \Rightarrow p_2=\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{aligned} x_1 \quad\quad + x_3 &= 0 \\ x_2 + x_3 &= 0 \\ x_1+x_2+2x_3 &= 0 \end{aligned} \Rightarrow p_3=\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

将它们标准化，得到酉阵V的列:

$$v_1=\frac{p_1}{\|p_1\|_2}=\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad v_2=\frac{p_2}{\|p_2\|_2}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad v_3=\frac{p_3}{\|p_3\|_2}=\frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

即得 $V=\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$，因**rank($A$)=2**，故有$(V_1)_{3\times 2}=\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix}$

进一步计算得出，

$$(\boldsymbol{U}_1)_{3\times 2}=\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1}=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}$$

得约化的奇异值分解

$$\boldsymbol{A}=\boldsymbol{U}_1\boldsymbol{\Sigma}\boldsymbol{V}_1^H=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 0 & 0\end{pmatrix}\begin{pmatrix}\sqrt{3} & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\end{pmatrix}$$

$$=\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}$$

计算U_2, 使其与U_1构成$\mathbf{R}^3$的一组标准正交基，可取 $U_2=\begin{pmatrix}0\\0\\1\end{pmatrix}$

则

$$U=\begin{pmatrix}U_1 & U_2\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}$$

是酉阵, 故矩阵A的奇异值分解（满的奇异值分解）为

$$A=\begin{pmatrix}\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{3} & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}}\\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0\\ \frac{-1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}}\end{pmatrix}$$

一阶顺序主子阵

二次阶顺序主子阵

三阶顺序主子阵

k阶顺序主子阵

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2k} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3k} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \cdots & a_{kk} & \cdots & a_{kn} \\ \vdots & \cdots & \vdots & \cdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n1} & \cdots & a_{ni} & \cdots & a_{nn} \end{pmatrix}$$

酉阵: $A \in \mathbf{C}^{n\times n}$, $A^H A = AA^H = I$

(复正交矩阵, $A^H = A^{-1}$), 例如,

$$A = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \quad A^H = \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}$$

$$A^H A = \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = AA^H$$

高斯（C.F.Gauss,1777.4.30 ~ 1855.2.23）

高斯是德国数学家、物理学家和天文学家，出生于德国布伦兹维克的一个贫苦家庭。父亲先后当过护堤工、泥瓦匠和园丁。

在全世界广为流传的一则故事说，高斯10岁时，老师当时给孩子们出了一道较难的加法题：

81297+81495+81693+…+100899

这是一个等差数列的求和问题（公差为198，项数为100）。当老师刚一写完时，高斯也算完并把写有答案的小石板交了上去，并算出了正确答案。

高斯有“数学王子”、“数学家之王”的美称、被认为是人类有史以来“最伟大的四位数学家之一”（阿基米德、牛顿、高斯和欧拉）。

高斯的研究领域，遍及纯粹数学和应用数学的各个领域，并且开辟了许多新的数学领域。人们评价到：若把18世纪的数学家想象为一系列的高山峻岭，那么最后一个令人肃然起敬的巅峰就是高斯；如果把19世纪的数学家想象为一条条江河，那么其源头就是高斯。