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

2.3 矩阵的Jordan分解

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

以下系数矩阵对应的线性方程组哪个容易求解？或者哪个容易计算行列式和特征值？

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

对角矩阵　　上(下)三角矩阵　　满矩阵

转化

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岁时算出老师给学生们出的将1到100的所有整数加起来的算术题，老师刚叙述完题目，高斯就算出了正确答案。不过，据考证老师当时给孩子们出的是一道更难的加法题： 81297+81495+81693+…+10089 。当然，这也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。

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。

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

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

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

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

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

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

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

上述为回代求解过程，得 $x = (1, 1, 1, 1)^T$。

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 $(\boldsymbol{A} \mid \boldsymbol{b})$ 化成上三角矩阵 $(\boldsymbol{U} \mid \boldsymbol{c})$，然后通过回代求与 $\boldsymbol{Ax} = \boldsymbol{b}$ 三角方程组 $\boldsymbol{Ux} = \boldsymbol{c}$ 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，其中

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

解 $(A\,|\,b)$

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

三次消元过程写成矩阵的形式分别为:

$$L_1A=\begin{pmatrix}1&&&\\-2&1&&\\-4&&1&\\-3&&&1\end{pmatrix}\begin{pmatrix}2&1&1&0\\4&3&3&1\\8&7&9&5\\6&7&9&8\end{pmatrix}=\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&3&5&5\\0&4&6&8\end{pmatrix}$$

$$L_2(L_1A)=\begin{pmatrix}1&&&\\&1&&\\&-3&1&\\&-4&&1\end{pmatrix}\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&3&5&5\\0&4&6&8\end{pmatrix}=\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&0&2&2\\0&0&2&4\end{pmatrix}$$

$$L_3(L_2L_1A)=\begin{pmatrix}1&&&\\&1&&\\&&1&\\&&-1&1\end{pmatrix}\begin{pmatrix}2&1&1&0\\0&1&1&1\\0&0&2&2\\0&0&2&4\end{pmatrix}$$

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

再注意到:

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

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

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

(单位) 下 (上) 三角矩阵的逆也是 (单位) 下 (上) 三角矩阵

所以，刚才的计算过程可以表示为

$$L_3L_2L_1A = U \qquad A = L_1^{-1}L_2^{-1}L_3^{-1}U$$

令

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

从而有

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

即得到矩阵 A 的 LU 分解。

定义2.1 对于 n 阶矩阵 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：

$$Ax = b_n \Leftrightarrow (LU)x = b_n \qquad n = 1, 2, \cdots$$

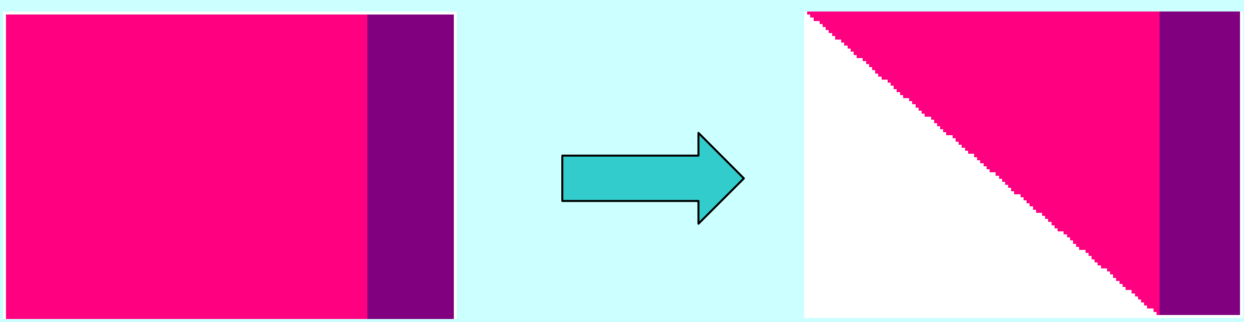
$$\Updownarrow$$

$$\begin{cases} Ly = b_n & n = 1, 2, \cdots \\ Ux = y \end{cases}$$

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:

第一步 $\quad$ 第i行$-$第1行$\times\dfrac{a_{i1}}{a_{11}}, i=2,\cdots,n$

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

运算量(乘除法): $(n\text{-}1)+(n\text{-}1)*n=(n\text{-}1)*(n+1)$

行乘子(数)

$$L_1 = \begin{pmatrix} 1 & & & \\ -l_{21} & 1 & & \\ \vdots & & \ddots & \\ -l_{n1} & & & 1 \end{pmatrix}$$

$$l_{i1} = \frac{a_{i1}^{(0)}}{a_{11}^{(0)}}, i = 2, \cdots, n$$

$$L_1(A \mid b) = \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} & b_1^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix}$$

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

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

运算量: $(n-2)+(n-2)*(n-1)=(n-2)n$

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

返回

$$L_k \cdots L_2 L_1 (A \mid b) =$$

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

$$\begin{pmatrix} a_{1,1}^{(0)} & \cdots & a_{1,k}^{(0)} & a_{1,k+1}^{(0)} & \cdots & a_{1,n}^{(0)} & b_1^{(0)} \\ & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{k,n}^{(k-1)} & b_k^{(k-1)} \\ & & & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} & b_{k+1}^{(k)} \\ & & & \vdots & \ddots & \vdots & \vdots \\ & & & a_{n,k+1}^{(k)} & \cdots & a_{n,n}^{(k)} & b_n^{(k)} \end{pmatrix}$$

运算量: $(n-k)*(1+n-k+1)=(n-k)(n-k+2)$

$n-1$步以后，我们可以得到变换后的矩阵为:

$$L_{n-1}L_{n-2}\cdots L_2L_1(A\,|\,b)=(U\,|\,c)$$

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

$$\boldsymbol{L}=(\boldsymbol{L}_{n-1}\boldsymbol{L}_{n-2}\cdots\boldsymbol{L}_2\boldsymbol{L}_1)^{-1}$$

$$\begin{pmatrix} 1 & & & & \\ l_{2,1} & 1 & & & \\ \vdots & \vdots & \ddots & & \\ l_{n-1,1} & l_{n-1,2} & \cdots & 1 & \\ l_{n,1} & l_{n,2} & \cdots & l_{n,n-1} & 1 \end{pmatrix}$$

$$\boldsymbol{L}^{-1}(\boldsymbol{A}\,|\,\boldsymbol{b})=(\boldsymbol{U}\,|\,\boldsymbol{c})$$

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

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\iff\boldsymbol{L}(\boldsymbol{U}\boldsymbol{x})=\boldsymbol{b}\iff\begin{cases}\boldsymbol{L}\boldsymbol{c}=\boldsymbol{b}\\ \boldsymbol{U}\boldsymbol{x}=\boldsymbol{c}\end{cases}$$

第k步运算量:

$$(n-k)*(1+n-k+1)=(n-k)(n-k+2)$$

因此，$n-1$步的总运算量为:

$$\sum_{k=1}^{n-1}(n-k)(n-k+2)=\frac{n^3}{3}+\frac{n^2}{2}-\frac{5n}{6}$$

$$\begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} \\ & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ & & \ddots & \vdots \\ & & & a_{nn}^{(n-1)} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\begin{pmatrix} b_1^{(0)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(n-1)} \end{pmatrix}$$

解$Ux=c$计算量

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

Gauss消去法求解n阶线性方程组的总计算量(乘除法次数)为:

$$\left(\frac{n^3}{3}+\frac{n^2}{2}-\frac{5n}{6}\right)+\left(\frac{n^2}{2}+\frac{n}{2}\right)=$$

$$\frac{n^3}{3}+n^2-\frac{n}{3} \qquad ＝3060（n=20）$$

当n较大时，它和　　　同阶的。

Gauss消去法可执行的条件？

回忆

A 的k 阶顺序主子式 主元 $a_{kk}^{(k-1)} \neq 0$

$$A = \begin{pmatrix} \boldsymbol{L}_1^* & \boldsymbol{O} \\ * & \boldsymbol{L}_2^* \end{pmatrix} \begin{pmatrix} \boldsymbol{U}_1 & \boldsymbol{U}_2 \\ \boldsymbol{O} & \boldsymbol{U}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{L}_1^*\boldsymbol{U}_1 & * \\ * & * \end{pmatrix}$$

$$\boldsymbol{D}_k = \det(\boldsymbol{L}_1^*\boldsymbol{U}_1) = \det(\boldsymbol{L}_1^*)\det(\boldsymbol{U}_1) = \det(\boldsymbol{U}_1)$$

$$= a_{11}^{(0)} a_{22}^{(1)} \cdots a_{kk}^{(k-1)} \quad = \boldsymbol{D}_{k-1} a_{kk}^{(k-1)}$$

令 $\boldsymbol{D}_0 = 1$ 则 $a_{kk}^{(k-1)} = \dfrac{\boldsymbol{D}_k}{\boldsymbol{D}_{k-1}}, k = 1,2,\ldots,n-1$

定理2.1（矩阵LU分解的存在和唯一性） 充分条件

如果n阶矩阵A的各阶顺序主子式D_k(k=1,…,n)均不为零，则必有单位下三角矩阵L和上三角矩阵U，使得 $A=LU$，而且L和U是唯一存在的。

证明唯一性，设 $L_1U_1=L_2U_2$

$$L_2^{-1}L_1=U_2U_1^{-1}=I \qquad L_2=L_1, U_2=U_1$$

单位下三角阵

上三角阵

P50 习题5 下述矩阵能否LU分解，是否唯一？

$$A=\begin{pmatrix}1&2&3\\2&4&1\\4&6&7\end{pmatrix}\xrightarrow{\begin{pmatrix}1&&\\-2&1&\\-4&&1\end{pmatrix}}\begin{pmatrix}1&2&3\\0&0&-5\\0&-2&5\end{pmatrix}$$

不能做LU分解！

事实上，A的各阶顺序主子式为 1, 0, -10

$$B=\begin{pmatrix}1&1&1\\2&2&1\\3&3&1\end{pmatrix}\xrightarrow{L_1^{-1}=\begin{pmatrix}1&&\\-2&1&\\-3&&1\end{pmatrix}}\begin{pmatrix}1&2&3\\0&0&-1\\0&0&-2\end{pmatrix}=U_1$$

$$\xrightarrow{L_2^{-1}=\begin{pmatrix}1&&\\&1&\\&-2&1\end{pmatrix}}\begin{pmatrix}1&2&3\\0&0&-1\\0&0&0\end{pmatrix}=U_2$$

可以LU分解，但不唯一！ $B=L_1U_1=(L_1L_2)U_2$

事实上，B的各阶顺序主子式为 1, 0, 0

$$C=\begin{pmatrix}1&2&6\\2&5&15\\6&15&46\end{pmatrix}\xrightarrow{L_1^{-1}=\begin{pmatrix}1&&\\-2&1&\\-6&&1\end{pmatrix}}\begin{pmatrix}1&2&6\\0&1&3\\0&3&10\end{pmatrix}$$

$$\xrightarrow{L_2^{-1}=\begin{pmatrix}1&&\\&1&\\&-3&1\end{pmatrix}}\begin{pmatrix}1&2&6\\0&1&3\\0&0&1\end{pmatrix}=U$$

可以LU分解，且唯一！ $C=LU=(L_1L_2)U$

事实上，C的各阶顺序主子式为 1, 1, 1

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：

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

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:

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

大数吃小数

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

显然 $(U\,|\,c)$有无穷多解. 但实际上, $\det(A)\neq 0$ 原线性方程组有唯一解。小主元作除数导致舍入误差使解面目全非!

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

$$\xrightarrow{\text{选列主元,交换第一和第三行}}\xrightarrow{\text{消元}}\xrightarrow{\text{选列主元}}\xrightarrow{\text{消元}}\begin{pmatrix} -2 & 1.072 & 5.643 & 3\\ 0 & 3.176 & 1.8015 & 0.5\\ 0 & 0 & 1.86554 & 0.68514 \end{pmatrix}=(U\,|\,c)$$

用回代法求$(U\,|\,c)$的解得

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

精确解 $x=(-0.491058227,-0.050886075,0.367257384)^T$.

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

$$= U$$

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 P_i 的定义知 $P_i^{-1}=P_i$, 即

$$P_1=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}=P_1^{-1}\qquad P_2=\begin{pmatrix}1&0&0&0\\0&0&0&1\\0&0&1&0\\0&1&0&0\end{pmatrix}=P_2^{-1}$$

$$P_3=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}=P_3^{-1}$$

从而，记

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

L的下标比P的下标小

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

显然，$\tilde{L}_2$和$\tilde{L}_1$分别与L_2和L_1结构相同，只是下三角部分的元素进行相应的对调。

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

$$\Updownarrow$$

$$L_3\tilde{L}_2\tilde{L}_1(P_3P_2P_1)A=U$$

进一步，得

$$P_3P_2P_1A=\tilde{L}_1^{-1}\tilde{L}_2^{-1}L_3^{-1}U$$

令

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

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

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

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 $PA=LU$。 (P可以不同，分解不唯一)

$$A = LU \qquad \det(A) = \det(L)\det(U) = \det(U)$$

$$Ax = b \quad \Leftrightarrow L(Ux) = b \quad \Leftrightarrow \begin{cases} Lc = b \\ Ux = c \end{cases}$$

$$PA = LU \qquad \det(P)\det(A) = \det(PA) = \det(L)\det(U) = \det(U)$$

$$\det(P) = (-1)^s, s\text{为换行次数}$$

$$Ax = b \quad \Leftrightarrow PAx = Pb$$

$$\Leftrightarrow L(Ux) = Pb \quad \Leftrightarrow \begin{cases} Lc = Pb \\ Ux = c \end{cases}$$

习题

例 用Gauss列主元消去法解如下方程组并给出 ***PA=LU***分解。

解:

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

$$\xrightarrow{\text{选列主元，交换第一和第三行}} \begin{pmatrix} 2 & -2 & 1 & 1 \\ 1 & 2 & 2 & 4 \\ 0 & -6 & -1 & -2 \end{pmatrix}$$

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

$$\xrightarrow{\text{选列主元，第二次消元}} \begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & -6 & -1 & -2 \\ 0 & 3 & \frac{3}{2} & \frac{7}{2} \end{pmatrix}$$

$$\begin{pmatrix} 2 & -2 & 1 & 1 \\ 0 & -6 & -1 & -2 \\ 0 & 0 & 1 & \frac{5}{2} \end{pmatrix} = (\boldsymbol{U} \mid \boldsymbol{c})$$

用回代法求的解得：

$$x_3 = \frac{5}{2} \qquad x_2 = \frac{-2+\frac{5}{2}}{-6} = -\frac{1}{12} \qquad x_2 = -\frac{5}{6}$$

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，消去第一列主对角元以下的非零元，左乘 $\boldsymbol{L}_1$

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

第二次选列主元，交换第2行和第3行，左乘置换矩阵 $\boldsymbol{P}_2$

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

第二次消元，消去第二列主对角元以下的非零元，左乘 $\boldsymbol{L}_2$

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

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，得到带列主元的LU分解，并求出 $\det(A)$

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

解:

$$\begin{pmatrix} -3 & 2 & 6 & 4 \\ 10 & -7 & 0 & 7 \\ 5 & -1 & 5 & 6 \end{pmatrix} \xrightarrow{\boldsymbol{P}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 5 & 6 \end{pmatrix}$$

$$\xrightarrow{\boldsymbol{L}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 3/10 & 1 & 0 \\ -1/2 & 0 & 1 \end{pmatrix}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & -\frac{1}{10} & 6 & \frac{61}{10} \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \end{pmatrix}$$

$$\begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & -\frac{1}{10} & 6 & \frac{61}{10} \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \end{pmatrix} \xrightarrow{P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & -\frac{1}{10} & 6 & \frac{61}{10} \end{pmatrix}$$

$$\xrightarrow{L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1/25 & 1 \end{pmatrix}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \\ 0 & 0 & \frac{31}{5} & \frac{31}{5} \end{pmatrix} = (\boldsymbol{U} \mid \boldsymbol{c})$$

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

$$L_2 P_2 L_1 P_1 A = U \qquad L_2 (P_2 L_1 P_2^{-1})(P_2 P_1) A = U$$

$$P = P_2 P_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \det(P) = 1$$

$$L = (L_2 \tilde{L}_1)^{-1} = (L_2 P_2 L_1 P_2^{-1})^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ -3/10 & -1/25 & 1 \end{pmatrix}$$

则 $PA = LU$

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

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

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

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

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

记 $D^{1/2}=\begin{pmatrix} \sqrt{u_{11}} & & & \\ & \sqrt{u_{22}} & & \\ & & \ddots & \\ & & & \sqrt{u_{nn}}\end{pmatrix}$ 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。

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

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

$$a_{21} = l_{21} l_{11} \qquad \Rightarrow l_{21} = a_{21} / l_{11}$$

$$\cdots a_{n1} = l_{n1} l_{11} \qquad \Rightarrow l_{n1} = a_{n1} / l_{11}$$

$$a_{22} = l_{21}^2 + l_{22}^2 \qquad \Rightarrow l_{22} = \sqrt{a_{22} - l_{21}^2}$$

利用矩阵乘法规则和利用的下三角结构得到

$$a_{jj}=\sum_{k=1}^{j-1}l_{jk}^2+l_{jj}^2,\quad j=1,2,\ldots n$$

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

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

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

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

$$\det(\boldsymbol{A}) = \det(\boldsymbol{L})\det(\boldsymbol{L}^T) = \prod_{j=1}^{n} l_{jj}^2$$

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \iff \boldsymbol{L}(\boldsymbol{L}^T\boldsymbol{x}) = \boldsymbol{b} \iff \begin{cases} \boldsymbol{L}\boldsymbol{y} = \boldsymbol{b} \\ \boldsymbol{L}^T\boldsymbol{x} = \boldsymbol{y} \end{cases}$$

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 $j=1, 2,\ldots, n$ 计算

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

第j列有（n−j+1）个元素，每个元素需j次乘除法

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

计算次序为$l_{11},l_{21},\cdots,l_{n1},l_{22},l_{32},\cdots,l_{n2},\cdots,l_{nn}$

计算量（乘除法次数）$\sum_{j=1}^{n}j(n-j+1)=\frac{n^3}{6}+\frac{n^2}{2}-\frac{n}{3}$

（2）求解下三角形方程组

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

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

$$x_n = y_n / l_{nn}, x_i = \left(y_i - \sum_{k=i+1}^{n} l_{ki} x_k \right) / l_{ii} \quad , \quad i = n-1, n-2, \cdots, 1$$

平方根法的数值稳定性

由 $a_{jj}=\sum_{k=1}^{j} l_{jk}^2$ ，推出 $|l_{jk}| \le \sqrt{a_{jj}}$ ，$k=1, 2, \cdots, j$.

因此在分解过程中L的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。

P28 例4 用cholesky方法求解线性方程组

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

解 (1)计算$L=\begin{pmatrix}2 & & \\ -0.5 & 2 & \\ 0.5 & 1.5 & 1\end{pmatrix}$

(2) 求解$Ly=b$，得$y=\{2,3.5,1\}^{T}$

(3) 求解$L^{T}x=y$，得$x=\{1,1,1\}^{T}$

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

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

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

$$c_i = d_i \qquad \Rightarrow d_i = c_i$$

$$b_1 = u_1, \qquad \Rightarrow u_1 = b_1,$$

$$b_i = l_i \cdot d_{i-1} + u_i \Rightarrow u_i = b_i - l_i \cdot c_{i-1}$$

$$a_i = l_i \cdot u_{i-1} \qquad \Rightarrow l_i = a_i / u_{i-1}$$

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

$$\boldsymbol{Ax}=\boldsymbol{f} \quad \Leftrightarrow \boldsymbol{L}(\boldsymbol{Ux})=\boldsymbol{f} \quad \Leftrightarrow \begin{cases} \boldsymbol{Ly}=\boldsymbol{f} \\ \boldsymbol{Ux}=\boldsymbol{y} \end{cases}$$

$$\begin{pmatrix} 1 & & & & \\ l_2 & 1 & & & \\ & l_3 & \ddots & & \\ & & \ddots & 1 & \\ & & & l_n & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix}$$

$$y_1 = f_1$$

$$l_i \cdot y_{i-1} + y_i = f_i$$

$$y_i = f_i - l_i \cdot y_{i-1}$$

$$\begin{pmatrix} u_1 & d_1 & & & \\ & u_2 & d_2 & & \\ & & \ddots & \ddots & \\ & & & u_{n-1} & d_{n-1} \\ & & & & u_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}$$

$$u_n \cdot x_n = y_n \qquad x_n = y_n / u_n$$

$$u_i \cdot x_i + d_i x_{i+1} = y_i$$

$$x_i = (y_i - d_i x_{i+1}) / u_i$$

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

LU分解计算量：乘除法2(n-1)，加减法n-1

（2）求解下三角形方程组

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

（3）求解$\boldsymbol{Ux = y}$

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

解方程计算量：乘除法3(n−1)+1,加减法2(n−1)

追赶法总计算量：乘除法5n−4,加减法3(n−1)

定理　设具有三对角形式的矩阵A，满足条件

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

(2) $|b_n| > |a_n| > 0$

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

对角占优

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

注：定理条件中 $a_i c_i \neq 0$，如果某个 $a_i = 0$ 或 $c_i = 0$ ，则可化成低阶方程组求解.

证明　已知 $u_1 = b_1, l_i = a_i / u_{i-1}, u_i = b_i - l_i \cdot c_{i-1}, i = 2,3,\ldots,n$

由 $|b_1| > |c_1| > 0 \quad \Rightarrow u_1 = b_1 \neq 0, \quad 0 < |c_1 / u_1| < 1$

归纳法证明 $u_i \neq 0, \quad 0 < |c_i / u_i| < 1, \quad i = 2,3,\ldots,n-1$

假设 $u_{i-1} \neq 0, 0 < |c_{i-1} / u_{i-1}| < 1,$ 于是

$$|u_i| = |b_i - l_i \cdot c_{i-1}| = |b_i - a_i \cdot c_{i-1} / u_{i-1}|$$

$$\geq |b_i| - |a_i| \cdot |c_{i-1} / u_{i-1}| \quad > |b_i| - |a_i| \quad \geq |c_i|$$

而且 $|u_n| = |b_n - l_n \cdot c_{n-1}| \geq |b_n| - |a_n||c_{n-1} / u_{n-1}| > |b_n| - |a_n| > 0$

于是 $\det(A) = \det(L)\det(U) = u_1 u_2 \cdots u_n \neq 0$ 方程组的解存在唯一

追赶法的优点:

- 计算量小，共8n-7次四则运算
- 存储量小，仅需要4个一维数组*a,b,c,f*,其中*d,l,u,x*分别存在*c,a,b,f*中。
- 当A为对角占优时，数值稳定（中间数有界）

$$\begin{aligned}|u_i| &= |b_i - l_i \cdot c_{i-1}| = |b_i - a_i \cdot c_{i-1}/u_{i-1}| \\ &\le |b_i| + |a_i| \cdot |c_{i-1}/u_{i-1}| < |b_i| + |a_i|\end{aligned}$$

$$|b_i| - |a_i| < |u_i| < |b_i| + |a_i|, \text{且 } d_i = c_i, |l_i| = |a_i/u_{i-1}|$$

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

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

又 $\|b\|=\|Ax\|\le\|A\|\cdot\|x\| \qquad \Rightarrow \quad \dfrac{1}{\|x\|}\le\dfrac{\|A\|}{\|b\|}$

相对误差放大因子

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

相对误差

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 $A^H A = A^{-1}AA^H A = A^{-1}(AA^H)A$ 正定

$$\det(\lambda I - A^H A) = \det(\lambda I - A^{-1}(AA^H)A)$$

$$= \det(A^{-1}(\lambda I - (AA^H))A)$$

$$= \det(A^{-1})\cdot\det(\lambda I - AA^H)\cdot\det(A) = \det(\lambda I - AA^H)$$

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

$$\|A^{-1}\|_2^2 = \lambda_{\max}((A^{-1})^H A^{-1}) = \lambda_{\max}((A^{H})^{-1}A^{-1})$$

$$= \lambda_{\max}((AA^{H})^{-1}) = \lambda_{\min}^{-1}(AA^{H}) = \lambda_{\min}^{-1}(A^{H}A) > 0$$

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

矩阵的条件数具有如下的性质：

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

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

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

求逆和数乘不改变矩阵的条件数

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

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

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

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

(4) 如果为U酉（正交）矩阵，则

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

酉变换不改变矩阵的2-范数和2-条件数

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

由第一章习题4酉矩阵与谱范数的性质可得

（5）**A，B**可逆 $\mathrm{cond}(AB)\le\mathrm{cond}(A)\cdot\mathrm{cond}(B)$

$$\mathrm{cond}(AB)=\|AB\|\left\|(AB)^{-1}\right\|$$

$$\le\|A\|\cdot\|B\|\cdot\left\|A^{-1}\right\|\cdot\left\|B^{-1}\right\|=\mathrm{cond}(A)\cdot\mathrm{cond}(B)$$

$$\frac{\|\delta \boldsymbol{x}\|}{\|\boldsymbol{x}\|} \le cond(\boldsymbol{A}) \frac{\|\delta \boldsymbol{b}\|}{\|\boldsymbol{b}\|}$$

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

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

***n*阶Hilbert矩阵**

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

$$\mathrm{cond}_2(\boldsymbol{H}_4)=1.5514\times10^4$$

$$\mathrm{cond}_2(\boldsymbol{H}_6)=1.4951\times10^7$$

$$\mathrm{cond}_2(\boldsymbol{H}_8)=1.525\times10^{10}$$

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 $\delta b = (0,\ 0.00001)^T,\ \delta A = O_{2\times 2}$。

我们观察 δb 对 x 的影响， 由

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

则 A 的条件数为：

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

$$= 8.00001 \times 600000.5$$

$$\approx 4800010 \approx 4.8 \times 10^6$$

则线性方程组的相对误差界为：

$$\frac{\|\boldsymbol{\delta x}\|_\infty}{\|\boldsymbol{x}\|_\infty} \le \mathrm{cond}_\infty(A)\frac{\|\boldsymbol{\delta b}\|_\infty}{\|\boldsymbol{b}\|_\infty} \approx 4.8\times10^6\times\frac{0.00001}{8.00001}$$

$$\approx 4.8\times10^6\times0.125\times10^{-5}$$

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

可见，右端向量$\boldsymbol{b}$其分量八十万分之一的变化，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2.5 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

$$\left(\frac{\|\delta A\|}{\|A\|}\le\|A^{-1}\|\|\delta A\|\right)$$

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

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

证明　由 $(A+\delta A)(x+\delta x)=b+\delta b \ = Ax+\delta Ax+A\delta x+\delta A\delta x$

$$\Rightarrow A\delta x=\delta b-\delta Ax-\delta A\delta x \quad \Rightarrow \delta x = A^{-1}\delta b - A^{-1}\delta Ax - A^{-1}\delta A\delta x$$

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

$$(1-\|A^{-1}\|\|\delta A\|)\|\delta x\| \le \|A^{-1}\|(\|\delta b\| + \|\delta A\|\|x\|)$$

$$\|\delta x\| \le \frac{\|A^{-1}\|(\|\delta b\| + \|\delta A\|\|x\|)}{1-\|A^{-1}\|\|\delta A\|} \qquad \frac{\|\delta x\|}{\|x\|} \le \frac{\|A^{-1}\|(\|\delta b\|/\|x\| + \|\delta A\|)}{1-\|A^{-1}\|\|\delta A\|}$$

$$\|b\| \le \|A\|\|x\|$$

$$\Rightarrow \frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}$$

$$\le \frac{\|A^{-1}\|(\|\delta b\|\|A\|/\|b\| + \|\delta A\|)}{1-\|A^{-1}\|\|\delta A\|} \le \frac{\|A\|\|A^{-1}\|(\frac{\|\delta b\|}{\|b\|} + \frac{\|\delta A\|}{\|A\|})}{1-\|A^{-1}\|\|A\|\frac{\|\delta A\|}{\|A\|}}$$

近似解的余量与它的相对误差间的关系

定理2.6 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

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

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

若cond(A) ≈ 1时，余量的相对误差可作为解的相对误差的一个好的度量，对于病态方程组，虽然余量的相对误差已经很小，但解的相对误差仍然很大。

证明　由 $Ax=b$　　$b-A\tilde{x}=Ax-A\tilde{x}=A(x-\tilde{x})$

$$x-\tilde{x}=A^{-1}(b-A\tilde{x}) \qquad \|x-\tilde{x}\|\le\|A^{-1}\|\|b-A\tilde{x}\|$$

$$\|b\|\le\|A\|\|x\| \Rightarrow \frac{1}{\|x\|}\le\frac{\|A\|}{\|b\|}$$

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

另一方面

$$x=A^{-1}b\Rightarrow\|x\|\le\|A^{-1}\|\|b\|\Rightarrow\frac{1}{\|x\|}\ge\frac{1}{\|A^{-1}\|\|b\|}$$

$$\|b-A\tilde{x}\|\le\|A\|\|x-\tilde{x}\|$$

$$\Rightarrow\|x-\tilde{x}\|\ge\frac{\|b-A\tilde{x}\|}{\|A\|} \qquad \Rightarrow\frac{\|x-\tilde{x}\|}{\|x\|}\ge\frac{\|b-A\tilde{x}\|}{\|A\|\|A^{-1}\|\|b\|}=\frac{1}{\mathrm{cond}(A)}\frac{\|b-A\tilde{x}\|}{\|b\|}$$

关于条件数的补充：有定理表明，当矩阵A十分病态时，就说明A已十分接近一个奇异矩阵。

$$\boldsymbol{A}=\begin{pmatrix}2 & 6\\ 2 & 6.00001\end{pmatrix}\qquad \det(\boldsymbol{A})=0.00002$$

思考：是否可以用行列式刻划矩阵的病态？

注意到 $\operatorname{cond}(\boldsymbol{A})=\operatorname{cond}(\boldsymbol{A}^{-1})$

$$\boldsymbol{A}^{-1}=\begin{pmatrix}300000.5 & -300000\\ -100000 & 100000\end{pmatrix}\qquad \det(\boldsymbol{A}^{-1})=50000=\det(\boldsymbol{A})^{-1}$$

2.1.6 矩阵的*QR*分解

- 回忆：求解线性方程组与矩阵的三角分解

$$\boldsymbol{Ax} = \boldsymbol{b} \qquad\qquad \boldsymbol{A} = \boldsymbol{LU}$$

$$\begin{cases} \boldsymbol{Ax} = \boldsymbol{b} \\ \boldsymbol{A} = \boldsymbol{LU} \end{cases} \Leftrightarrow \boldsymbol{LUx} = \boldsymbol{b} \Leftrightarrow \begin{cases} \boldsymbol{Ly} = \boldsymbol{b} \\ \boldsymbol{Ux} = \boldsymbol{y} \end{cases}$$

- 问题：条件数与方程组的性态

$$cond(\boldsymbol{A}) = cond(\boldsymbol{LU}) \le cond(\boldsymbol{L}) \cdot cond(\boldsymbol{U})$$

LU分解是否能保持条件数？

例

$$A=\begin{pmatrix}1&1&3\\2&4&1\\2&0&2\end{pmatrix}=LU=\begin{pmatrix}1&0&0\\2&1&0\\2&-1&1\end{pmatrix}\begin{pmatrix}1&1&3\\0&2&-5\\0&0&-9\end{pmatrix}$$

$$cond_2(A)\approx 4.89894, cond_2(L)\approx 14.9224, cond_2(U)\approx 14.2208.$$

良态方程组 $Ax=b$ 变为病态方程组 $\begin{cases}Ly=b\\Ux=y\end{cases}$

矩阵的LU分解不能保证条件数！

能否构造保持条件数的矩阵三角分解？

解决方法：正交变换保持2-条件数，即若Q为正交矩阵（$Q^{-1}=Q^T$），则

$$cond_2(Q)=\sqrt{\frac{\lambda_{\max}(Q^TQ)}{\lambda_{\min}(Q^TQ)}}=1,$$

$$cond_2(QA)=cond_2(AQ)=cond_2(A)$$

若$A=QR$，Q为正交阵，R为上三角阵

$$\begin{cases} Ax=b \\ A=QR \end{cases} \Leftrightarrow QRx=b \Leftrightarrow \begin{cases} Qy=b \\ Rx=y \end{cases} \Leftrightarrow Rx=Q^Tb$$

一般矩阵的QR分解的定义

定义：如果 $A \in \mathrm{R}^{m\times n}(m \geq n), r(A) = n,$

$$A = Q\begin{pmatrix} R_1 \\ O \end{pmatrix} = QR$$

其中 Q为正交阵，R_1为对角元大于零的上三角矩阵。上面的矩阵分解式称为矩阵的**QR**分解.

注：对角元大于零的条件不是必须的。如果小于零，只要再乘以初等（正交）矩阵 $P(i(-1))$ 即可。

以下考虑方阵的情形。

矩阵消元的几何观点

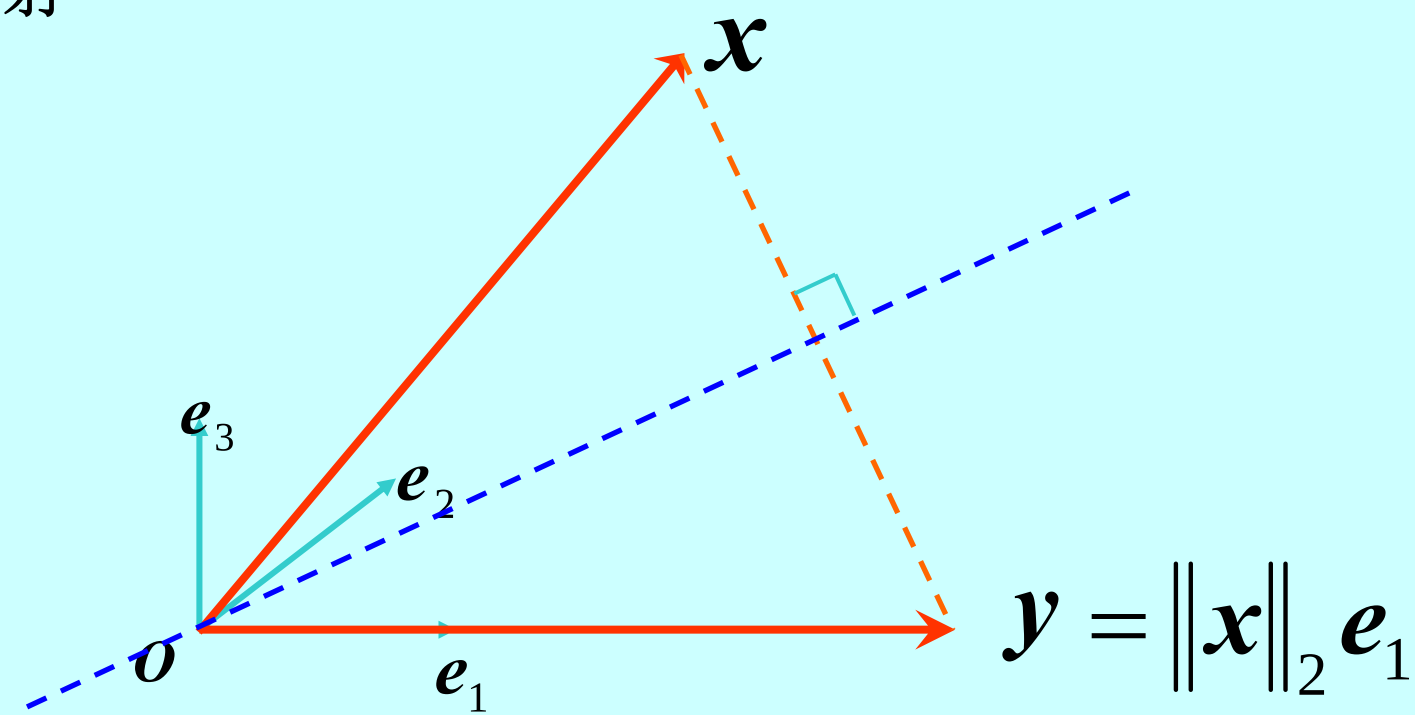
$$A=\begin{pmatrix}\times & \times & \times\\ \times & \times & \times\\ \times & \times & \times\end{pmatrix}\xrightarrow{\text{用正交变换}\boldsymbol{Q}_1}\begin{pmatrix}\times & \times & \times\\ 0 & \times & \times\\ 0 & \times & \times\end{pmatrix}\xrightarrow{\text{用正交变换}\boldsymbol{Q}_2}\begin{pmatrix}\times & \times & \times\\ 0 & \times & \times\\ 0 & 0 & \times\end{pmatrix}=\boldsymbol{R}$$

$$\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_n),\quad \boldsymbol{a}_1\rightarrow\boldsymbol{Q}_1\boldsymbol{a}_1=\begin{pmatrix}\times\\ 0\\ \vdots\\ 0\end{pmatrix}$$

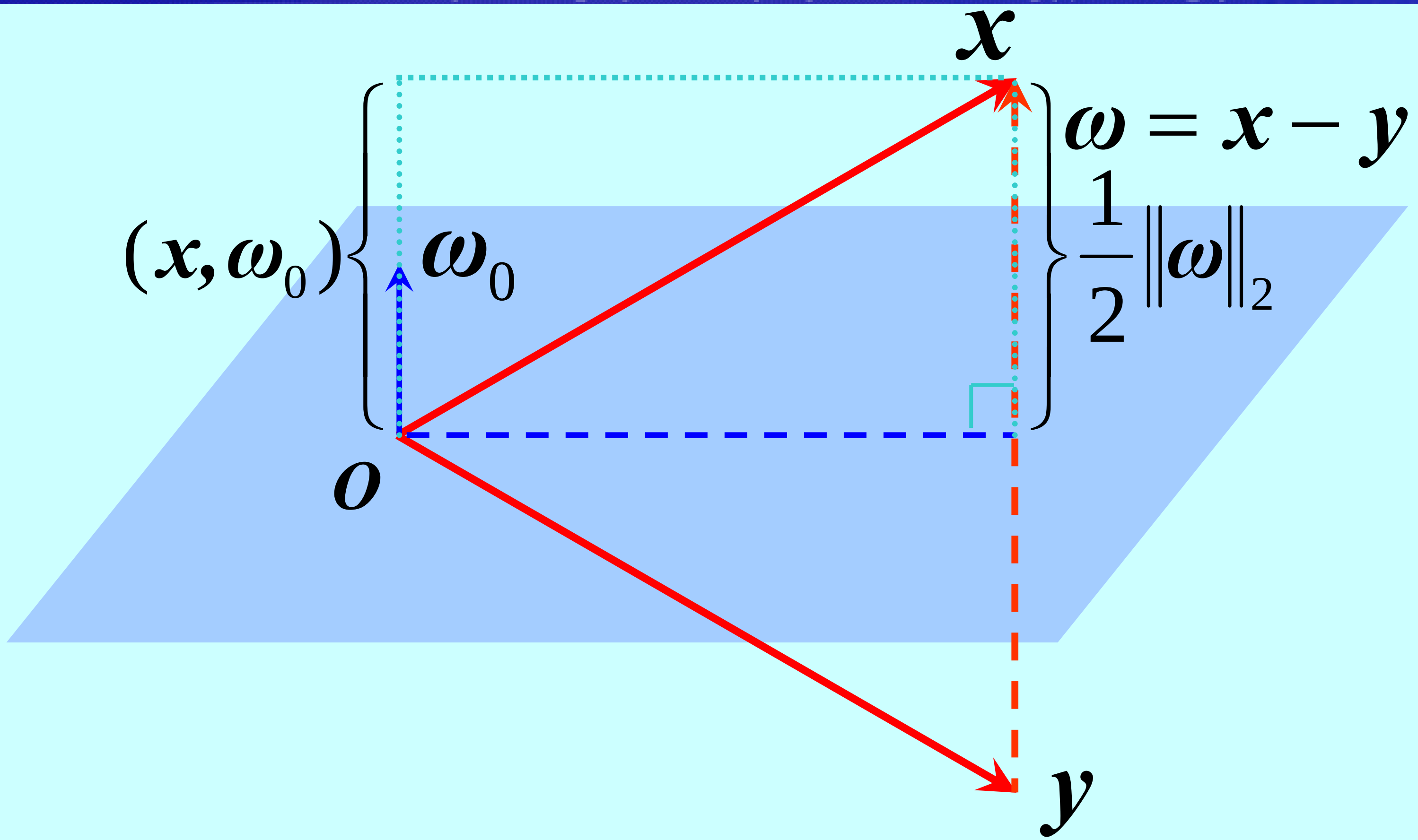
几何上看，就是把空间中的一个向量通过正交变换，变为落在第一个坐标轴上的向量。

正交变换：旋转和镜面反射，特点是保持向量的内积和长度（**2**-范数）不变。

镜面反射



如何将任意非零向量 $\boldsymbol{x}$变为落在第一个坐标轴 $\boldsymbol{e}_1$上的向量 $\boldsymbol{y}=\|\boldsymbol{x}\|_2\boldsymbol{e}_1$ ？



$$\omega_0 = \frac{\omega}{\|\omega\|_2}, \quad \omega = x - y = \|\omega\|_2 \omega_0, \ (x, \omega_0) = \frac{1}{2}\|\omega\|_2$$

$$\omega_0 = \frac{\omega}{\|\omega\|_2}, \quad \omega = x - y = \|\omega\|_2 \omega_0, \quad \|\omega\|_2 = 2(x, \omega_0) = 2\omega_0^T x$$

$$x - y = \omega_0 \cdot 2(x, \omega_0) = 2\omega_0(\omega_0^T x) = 2\frac{\omega(\omega^T x)}{\|\omega\|_2^2} = 2\frac{(\omega\omega^T)x}{\omega^T\omega}$$

$$y = x - 2\frac{\omega\omega^T}{\omega^T\omega}x = (I - \frac{2}{\omega^T\omega}\omega\omega^T)x := H(\omega)x$$

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

$$H(\omega) = I - \frac{2}{\omega^T\omega}\omega\omega^T \tag{2-33}$$

为**Householder矩阵**（简称H阵），或称**Householder变换矩阵**.

Householder矩阵的性质

1. 对称性： $H(\omega)^T = H(\omega)$

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

2. 正交性： $H(\omega)^T H(\omega) = I$

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

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

3. 如果 $H(\omega)x = y$, 则 $\|y\|_2 = \|x\|_2$ (长度不变)

$$\|y\|_2^2 = y^T y = (H(\omega)x)^T (H(\omega)x) = x^T \left(H(\omega)^T H(\omega)\right) x = x^T x = \|x\|_2^2$$

4. 设 $x = (x_1, x_2, \cdots, x_n)^T \in \mathbf{R}^n$ 且 $x \neq 0$, 取 $\omega = x - \|x\|_2 e_1$ 则

$$H(\omega)x = H(x - \|x\|_2 e_1)x = \begin{pmatrix} \|x\|_2 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \|x\|_2 e_1.$$

利用一系列H阵进行矩阵的QR分解

例5 利用Householder变换求A的分解，其中

$$A=\begin{pmatrix}1&1&1\\2&3&1\\2&1&-5\end{pmatrix}.$$

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

$$\boldsymbol{a}_1=\begin{pmatrix}1\\2\\2\end{pmatrix},\ \ \|\boldsymbol{a}_1\|_2=3,\ \ \omega_1=\boldsymbol{a}_1-\|\boldsymbol{a}_1\|_2\boldsymbol{e}_1=\begin{pmatrix}1\\2\\2\end{pmatrix}-\begin{pmatrix}3\\0\\0\end{pmatrix}=\begin{pmatrix}-2\\2\\2\end{pmatrix}$$

$$\boldsymbol{Q}_1 = \boldsymbol{H}(\boldsymbol{\omega}_1) = \boldsymbol{I} - \frac{2}{\boldsymbol{\omega}_1^{\mathrm{T}}\boldsymbol{\omega}_1}\boldsymbol{\omega}_1\boldsymbol{\omega}_1^{\mathrm{T}} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & \frac{-2}{3} \\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3} \end{pmatrix},$$

$$\boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{A} = (\boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{a}_1, \boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{a}_2, \boldsymbol{H}(\boldsymbol{\omega}_1)\boldsymbol{a}_3) = \begin{pmatrix} 3 & 3 & -\frac{7}{3} \\ 0 & 1 & \frac{13}{3} \\ 0 & -1 & -\frac{5}{3} \end{pmatrix} = \begin{pmatrix} 3 & \boldsymbol{b}^{\mathrm{T}} \\ \boldsymbol{0} & \boldsymbol{A}_2 \end{pmatrix},$$

$$A_2=\begin{pmatrix}1 & \frac{13}{3}\\ -1 & -\frac{5}{3}\end{pmatrix}=(\tilde{a}_1,\tilde{a}_2),\ \ \tilde{a}_1=\begin{pmatrix}1\\-1\end{pmatrix},\ \ \|\tilde{a}_1\|_2=\sqrt{2}$$

$$\omega_2=\tilde{a}_1-\|\tilde{a}_1\|_2 e_1=\begin{pmatrix}1\\-1\end{pmatrix}-\begin{pmatrix}\sqrt{2}\\0\end{pmatrix}=\begin{pmatrix}1-\sqrt{2}\\-1\end{pmatrix}$$

$$H(\omega_2)=\begin{pmatrix}1&0\\0&1\end{pmatrix}-\frac{1}{2-\sqrt{2}}\begin{pmatrix}3-2\sqrt{2} & -1+\sqrt{2}\\ -1+\sqrt{2} & 1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}\sqrt{2} & -\sqrt{2}\\ -\sqrt{2} & -\sqrt{2}\end{pmatrix},$$

$$H(\omega_2)A_2=\left(H(\omega_2)\tilde{a}_1,H(\omega_2)\tilde{a}_2\right)=\begin{pmatrix}\sqrt{2} & 3\sqrt{2}\\ 0 & -\frac{4}{3}\sqrt{2}\end{pmatrix}.$$

$$\boldsymbol{Q}_2=\begin{pmatrix}1 & \boldsymbol{0}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{H}(\boldsymbol{\omega}_2)\end{pmatrix}_{3\times 3}$$

$$\boldsymbol{Q}_2\boldsymbol{Q}_1\boldsymbol{A}=\begin{pmatrix}1 & \boldsymbol{0}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{H}(\boldsymbol{\omega}_2)\end{pmatrix}\begin{pmatrix}3 & \boldsymbol{b}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{A}_2\end{pmatrix}=\begin{pmatrix}3 & \boldsymbol{b}^{\mathrm{T}}\\ \boldsymbol{0} & \boldsymbol{H}(\boldsymbol{\omega}_2)\boldsymbol{A}_2\end{pmatrix}=\begin{pmatrix}3 & 3 & -\frac{7}{3}\\ 0 & \sqrt{2} & 3\sqrt{2}\\ 0 & 0 & -\frac{4}{3}\sqrt{2}\end{pmatrix}=\boldsymbol{R}$$

$$\boldsymbol{Q}=(\boldsymbol{Q}_2\boldsymbol{Q}_1)^{T}=\boldsymbol{Q}_1\boldsymbol{Q}_2=\begin{pmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\\ \frac{2}{3} & \frac{1}{3} & \frac{-2}{3}\\ \frac{2}{3} & \frac{-2}{3} & \frac{1}{3}\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\end{pmatrix}=\begin{pmatrix}\frac{1}{3} & 0 & -\frac{2\sqrt{2}}{3}\\ \frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6}\\ \frac{2}{3} & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{6}\end{pmatrix}$$

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}$$

$$R=\begin{pmatrix}3 & 3 & -\dfrac{7}{3}\\ 0 & \sqrt{2} & 3\sqrt{2}\\ 0 & 0 & -\dfrac{4}{3}\sqrt{2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}\begin{pmatrix}3 & 3 & -\dfrac{7}{3}\\ 0 & \sqrt{2} & 3\sqrt{2}\\ 0 & 0 & \dfrac{4}{3}\sqrt{2}\end{pmatrix}=Q_3\tilde{R},$$

$$A=QR=(QQ_3)\tilde{R}=\tilde{Q}\tilde{R}$$

小结

利用**Householder**矩阵对矩阵 A 进行**QR**分解，保持矩阵的条件数，数值稳定，但计算量比**LU**分解大。

1. $H(\omega)=I-\dfrac{2}{\omega^{\mathrm{T}}\omega}\omega\omega^{\mathrm{T}}$ 的定义，性质和几何意义。

2. **QR**分解的算法过程，
“降阶—变换—镶边—升阶—合并”。

- 思考题：

1. 能不能利用其它正交变换（如旋转）进行**QR**分解？
2. 除了**QR**分解，是否有别的分解（如带列主元的**LU**分解）能够保持矩阵的条件数？ 或者如何修正**LU**分解，使其保持矩阵条件数？

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

$$A = URU^H$$

酉相似

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**

证明　对矩阵阶数n用数学归纳法

$n=1$时，定理显然成立

设$n=k$时定理成立，证明$n=k+1$时定理仍成立

记 λ 为$k+1$阶方阵A的一个特征值，于是存在

$$\boldsymbol{u}_1 \in \boldsymbol{C}^{k+1}, \|\boldsymbol{u}_1\| = 1 \qquad \boldsymbol{A}\boldsymbol{u}_1 = \lambda \boldsymbol{u}_1$$

将$\boldsymbol{u}_1$扩充为$\boldsymbol{C}^{k+1}$的一组标准正交基，记为

$$\boldsymbol{U}_1 = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1}) \in \boldsymbol{C}^{(k+1)\times(k+1)}$$

$\boldsymbol{U}_1 = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{k+1}) \in \boldsymbol{C}^{(k+1)\times(k+1)}$为酉阵

$$\boldsymbol{U}_1^{\ H} \boldsymbol{A} \boldsymbol{U}_1 = \begin{pmatrix} \boldsymbol{u}_1^H \\ \boldsymbol{u}_2^H \\ \vdots \\ \boldsymbol{u}_{k+1}^H \end{pmatrix} (\boldsymbol{A}\boldsymbol{u}_1, \boldsymbol{A}\boldsymbol{u}_2, \cdots, \boldsymbol{A}\boldsymbol{u}_{k+1}) = \begin{pmatrix} \lambda & \boldsymbol{c}^T \\ \boldsymbol{0} & \boldsymbol{A}_1 \end{pmatrix}$$

其中$\boldsymbol{c} \in \boldsymbol{C}^k, \boldsymbol{A}_1 \in \boldsymbol{C}^{k\times k}$

由归纳假设，存在酉阵 $\boldsymbol{U}_2 \in \boldsymbol{C}^{k\times k}$，上三角阵 $\boldsymbol{R}_1 \in \boldsymbol{C}^{k\times k}$

$$\boldsymbol{A}_1 = \boldsymbol{U}_2 \boldsymbol{R}_1 \boldsymbol{U}_2^H$$

$$\boldsymbol{A}=\boldsymbol{U}_1\begin{pmatrix}\lambda & \boldsymbol{c}^T\\ \boldsymbol{0} & \boldsymbol{A}_1\end{pmatrix}\boldsymbol{U}_1^H=\boldsymbol{U}_1\begin{pmatrix}\lambda & \boldsymbol{c}^T\\ \boldsymbol{0} & \boldsymbol{U}_2\boldsymbol{R}_1\boldsymbol{U}_2^H\end{pmatrix}\boldsymbol{U}_1^H$$

$$=\boldsymbol{U}_1\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{U}_2\end{pmatrix}\begin{pmatrix}\lambda & \tilde{\boldsymbol{c}}^T\\ \boldsymbol{0} & \boldsymbol{R}_1\end{pmatrix}\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{U}_2^H\end{pmatrix}\boldsymbol{U}_1^H$$

其中$\tilde{\boldsymbol{c}}^T\boldsymbol{U}_2^H=\boldsymbol{c}^T\Rightarrow\tilde{\boldsymbol{c}}^T=\boldsymbol{c}^T\boldsymbol{U}_2\Rightarrow\tilde{\boldsymbol{c}}=\boldsymbol{U}_2^T\boldsymbol{c}\ \in\boldsymbol{C}^k$

记$\boldsymbol{U}=\boldsymbol{U}_1\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{U}_2\end{pmatrix},\boldsymbol{R}=\begin{pmatrix}\lambda & \tilde{\boldsymbol{c}}^T\\ \boldsymbol{0} & \boldsymbol{R}_1\end{pmatrix}$

$$\boxed{\boldsymbol{A}=\boldsymbol{U}\boldsymbol{R}\boldsymbol{U}^H}$$

$$\boldsymbol{U}^H\boldsymbol{U}=\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{U}_2^H\end{pmatrix}\boldsymbol{U}_1^H\boldsymbol{U}_1\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{U}_2\end{pmatrix}=\begin{pmatrix}1 & \boldsymbol{0}^T\\ \boldsymbol{0} & \boldsymbol{U}_2^H\boldsymbol{U}_2\end{pmatrix}=\boldsymbol{I}$$

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A的Schur**标准型**。

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

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：$A^HA=AA^H=I$

正交矩阵：$A^TA=AA^T=I$

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

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

$$=UD^H U^H UDU^H=U\left(D^H D\right)U^H$$

$$AA^H=\left(UDU^H\right)\left(UDU^H\right)^H$$

$$=UDU^H UD^H U^H=U\left(DD^H\right)U^H$$

而

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

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

故

$$A^H A = AA^H$$

即A为正规矩阵.

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

$U \in \mathbf{C}^{n\times n}$ 为酉阵, R为上三角阵。那么，由假设知A为

正规矩阵，即 $A^H A = AA^H \Rightarrow R^H R = RR^H$，即

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

事实上, 设

习题10

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

再注意到,

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

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$D \in \mathrm{R}^{n\times n}$ 为实对角矩阵。

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

$$d_k=a_k+ib_k=\;a_k-ib_k=\overline{d}_k\Rightarrow b_k=0,$$

即

$$d_k=\overline{d}_k=a_k$$

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或0。

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的对角元素为纯虚数或0。注意到，由

$$d_k=a_k+ib_k=-(a_k-ib_k)=-\overline{d}_k\Rightarrow a_k=0,$$

即

$$d_k=-\overline{d}_k=ib_k$$

从而D为n阶纯虚或0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

$$\supset \begin{Bmatrix} \text{Hermite矩阵} \supset \text{实对称矩阵} \\ \text{酉矩阵} \supset \text{实正交矩阵} \end{Bmatrix}$$

定理 设A为n阶方阵，则任意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。

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

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

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

最后，

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

习题11 证明 Schur不等式:

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

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 $A = URU^H$

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

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

定理2.8 设A为n阶方阵，ε >0，则存在 $C^{n\times n}$ 上的一种算子范数 $\|\cdot\|_M$（依赖矩阵 A 和常数 ε），使得 $\|I\|_M = 1$

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

证明　由Schur定理，存在n阶酉阵U，使得上三角矩阵

$$\begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{pmatrix} = R = U^H A U$$

取$\delta=\min\left\{1,\dfrac{\varepsilon}{(n-1)\max\limits_{1\le i<j\le n}|r_{ij}|}\right\}$, $\quad \boldsymbol{D}=\operatorname{diag}(1,\delta,\delta^2,\cdots,\delta^{n-1})$

$$\boldsymbol{D}^{-1}\boldsymbol{R}\boldsymbol{D}=\begin{pmatrix} r_{1,1} & r_{1,2}\delta & r_{1,3}\delta^2 & \cdots & r_{1,n}\delta^{n-1} \\ & r_{2,2} & r_{2,3}\delta & \ddots & \vdots \\ & & \ddots & \ddots & r_{n-2,n}\delta^2 \\ & & & \ddots & r_{n-1,n}\delta \\ & & & & r_{n,n} \end{pmatrix}$$

$$\left\|\boldsymbol{D}^{-1}\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U}\boldsymbol{D}\right\|_1=\left\|\boldsymbol{D}^{-1}\boldsymbol{R}\boldsymbol{D}\right\|_1$$

$$\le\max_{1\le i\le n}|\lambda_i|+\max_{1\le i<j\le n}|r_{i,j}|(1+\delta+\cdots+\delta^{n-2})\delta$$

$$\le\rho(\boldsymbol{A})+(n-1)\max_{1\le i<j\le n}|r_{i,j}|\delta\quad\le\rho(\boldsymbol{A})+\varepsilon$$

由第一章习题3

记$\|\boldsymbol{A}\|_M=\left\|\boldsymbol{D}^{-1}\boldsymbol{U}^H\boldsymbol{A}\boldsymbol{U}\boldsymbol{D}\right\|_1$ $\quad$ 则$\|\boldsymbol{I}\|_M=1$, $\quad \|\boldsymbol{A}\|_M\le\rho(\boldsymbol{A})+\varepsilon$

2.3 矩阵的Jordan分解

定义 2.6 设A为n阶方阵，A的特征多项式为

$$\det(\lambda \boldsymbol{I}_n - \boldsymbol{A}) = (\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_s)^{m_s} \qquad (2\text{-}42)$$

其中 $m_i(i=1,2,\ldots,s)$ 均为正整数，$\sum_{i=1}^{s} m_i = n$，$\lambda_1,\lambda_2,\ldots,\lambda_s$

为A的不同特征值，称m_i为 λ_i 的代数重数;

把与λ_i对应的线性无关的特征向量的个数，即子空间 $N(\lambda_i \boldsymbol{I}_n - \boldsymbol{A})$

(即 $(\lambda_i \boldsymbol{I}_n - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 的解空间，称为 $\lambda_i \boldsymbol{I}_n - \boldsymbol{A}$的零空间) 的维数，

称为λ_i的几何重数，记为 α_i，$\alpha_i = n - \operatorname{rank}(\lambda_i \boldsymbol{I}_n - \boldsymbol{A})$.

代数重数 $\geq$ 几何重数 $m_i \geq \alpha_i$

取特征子空间 $N(\lambda_i \boldsymbol{I}_n - \boldsymbol{A})$ 的一组基 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\alpha_i}$

扩充为$\mathbf{R}^n$ 的基 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\alpha_i}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n-\alpha_i}$

令 $\boldsymbol{U} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{\alpha_i}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{n-\alpha_i})$

$$\boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{U} = \boldsymbol{U}^{-1}(\boldsymbol{A}\boldsymbol{x}_1, \ldots, \boldsymbol{A}\boldsymbol{x}_{\alpha_i}, \boldsymbol{A}\boldsymbol{y}_1, \ldots, \boldsymbol{A}\boldsymbol{y}_{n-\alpha_i})$$

$$= (\lambda_i \boldsymbol{U}^{-1}\boldsymbol{x}_1, \ldots, \lambda_i \boldsymbol{U}^{-1}\boldsymbol{x}_{\alpha_i}, \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{y}_1, \ldots, \boldsymbol{U}^{-1}\boldsymbol{A}\boldsymbol{y}_{n-\alpha_i}) = \begin{pmatrix} \lambda_i \boldsymbol{I}_{\alpha_i} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{C} \end{pmatrix}$$

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \det(\lambda \boldsymbol{I}_{\alpha_i} - \lambda_i \boldsymbol{I}_{\alpha_i}) \cdot \det(\lambda \boldsymbol{I}_{\alpha_i} - \boldsymbol{C})$$

$$= (\lambda - \lambda_i)^{\alpha_i} \cdot \det(\lambda \boldsymbol{I}_{\alpha_i} - \boldsymbol{C}) \Rightarrow m_i \geq \alpha_i$$

定义2.7 设A为n阶方阵，λ_i为其特征值，m_i和α_i分别为其代数重数和几何重数. 如果$m_i=\alpha_i$，则称特征值λ_i为半单的；如果$m_i>\alpha_i$，则称特征值λ_i为亏损的.

- 代数重数为1的特征值一定是半单的.
- 不同特征值对应的特征向量是线性无关的.

定理2.9

- 每个特征值都是半单的矩阵称为单纯矩阵(有完备的特征向量系) $\Leftrightarrow$ 可对角化.
- 存在亏损的特征值的矩阵称为亏损矩阵 $\Leftrightarrow$ 不可对角化.

例1 下列矩阵是否可以对角化?

$$\boldsymbol{A}=\begin{pmatrix}-5 & 6 & -3\\ -1 & 0 & 1\\ 1 & 2 & -1\end{pmatrix}\qquad \boldsymbol{B}=\begin{pmatrix}1 & -1 & 2\\ 3 & -3 & 6\\ 2 & -2 & 4\end{pmatrix}\qquad \boldsymbol{C}=\begin{pmatrix}17 & 0 & -25\\ 0 & 3 & 0\\ 9 & 0 & -13\end{pmatrix}$$

(1) $\det(\lambda \boldsymbol{I}_3-\boldsymbol{A})=(\lambda-2)(\lambda^2-2\lambda-2)$

$$\lambda_1=2, \lambda_2=1+\sqrt{3}, \lambda=1-\sqrt{3}$$

$\boldsymbol{A}$为单纯矩阵,可对角化

(2) $\det(\lambda \boldsymbol{I}_3 - \boldsymbol{B}) = \lambda^2(\lambda - 2)$

$\lambda_1 = 0, m_1 = 2, \lambda_2 = 2, m_2 = 1 \qquad rank(\lambda_1 \boldsymbol{I}_3 - \boldsymbol{B}) = rank(\boldsymbol{B}) = 1,$

几何重数$\alpha_1 = 3 - 1 = 2 \Rightarrow \lambda_1$是半单的

$\boldsymbol{B}$为单纯矩阵,可对角化.

(3) $\det(\lambda \boldsymbol{I}_3 - \boldsymbol{C}) = (\lambda - 2)^2(\lambda - 3)$

$\lambda_1 = 2, m_1 = 2, \lambda_2 = 3, m_2 = 1 \qquad rank(\lambda_1 \boldsymbol{I}_3 - \boldsymbol{C}) = 2,$

几何重数$\alpha_1 = 3 - 2 = 1 \Rightarrow \lambda_1$是亏损的

$\boldsymbol{C}$为亏损矩阵,不可对角化.

定义2.8 称下面的$k\times k$阶方阵为Jordan块

$$\boldsymbol{J}_k(\lambda)=\begin{pmatrix}\lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda\end{pmatrix}$$

$$\boldsymbol{J}_4(0)=\begin{pmatrix}0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0\end{pmatrix} \quad \boldsymbol{J}_3(2)=\begin{pmatrix}2 & 1 & \\ & 2 & 1 \\ & & 2\end{pmatrix} \quad \boldsymbol{J}_2(1)=\begin{pmatrix}1 & 1 \\ & 1\end{pmatrix}$$

定义2.9 由若干个Jordan块排成的块对角矩阵为Jordan阵.

$$\boldsymbol{J}=\begin{pmatrix} \boldsymbol{J}_3(2) & & \\ & \boldsymbol{J}_4(0) & \\ & & \boldsymbol{J}_2(1) \end{pmatrix}=\operatorname{diag}(\boldsymbol{J}_3(2),\boldsymbol{J}_4(0),\boldsymbol{J}_2(1))$$

定理2.10 设$\boldsymbol{A}$为$\boldsymbol{n}$阶方阵, 则存在$\boldsymbol{n}$阶可逆矩阵$\boldsymbol{T}$使得

$$\boldsymbol{A}=\boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1} \qquad \textbf{(2-43)}$$

其中$\boldsymbol{J}=\operatorname{diag}(\boldsymbol{J}_{n_1}(\lambda_1),\boldsymbol{J}_{n_2}(\lambda_2),\ldots,\boldsymbol{J}_{n_k}(\lambda_k)), n_1+n_2+\cdots n_k=n.$

称(2-43)为$\boldsymbol{A}$的Jordan分解, $\boldsymbol{J}$称为$\boldsymbol{A}$的Jordan标准型, $\boldsymbol{T}$称为变换矩阵. 若不计Jordan块的次序, 则Jordan标准型唯一.

(一) 关于Jordan标准型*J*

Jordan标准型是一个块对角矩阵, 对角元是矩阵*J*的特征值.

对于特征值 λ_i，它的代数重数是**Jordan**标准型中以λ_i 为特征值的**Jordan**块的阶数之和. 不同**Jordan**块的特征值可能相同.

对于特征值 λ_i，它的几何重数, 即与λ_i 对应的线性无关的特征向量的个数, 恰为以λ_i为特征值的**Jordan**块的个数.

例2 求矩阵A的Jordan标准型J, 其中 $$A=\begin{pmatrix}3&0&8\\3&-1&6\\-2&0&-5\end{pmatrix}$$

解 $\det(\lambda I-A)=(\lambda+1)^3$

$$\lambda_1=\lambda_2=\lambda_3=-1 \quad 3-\mathrm{rank}(\lambda_1 I-A)=3-\mathrm{rank}\begin{pmatrix}-4&0&-8\\-3&0&-6\\2&0&4\end{pmatrix}=2$$

代数重数为3，以−1为特征值的Jordan块的阶数之和为3.

几何重数为2，以−1为特征值的Jordan块的个数为2.

$$J=\begin{pmatrix}-1&&\\&-1&1\\&&-1\end{pmatrix} \quad 或J=\begin{pmatrix}-1&1&\\&-1&\\&&-1\end{pmatrix}$$

返回

定理2.11 设$\boldsymbol{A}$为$\boldsymbol{n}$阶方阵, λ_i为其特征值, 则$\boldsymbol{A}$的**Jordan**标准型**J**中以λ_i为特征值, 阶数为$\boldsymbol{l}$ 的**Jordan**块的个数为

$$r_{l+1}+r_{l-1}-2r_l,$$

其中$r_l=\operatorname{rank}(\lambda_i \boldsymbol{I}-\boldsymbol{A})^l$.

$$r_0=\operatorname{rank}(\lambda_i \boldsymbol{I}-\boldsymbol{A})^0=\operatorname{rank}(\boldsymbol{I})=n$$

例3 求矩阵$\boldsymbol{A}$的Jordan标准型$\boldsymbol{J}$，其中

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

解 $\det(\lambda \boldsymbol{I}-\boldsymbol{A})=(\lambda-2)^4$

$$\lambda_1=\lambda_2=\lambda_3=\lambda_4=2, \qquad 4-\operatorname{rank}(\lambda_1\boldsymbol{I}-\boldsymbol{A})=2$$

代数重数为4，以2为特征值的Jordan块的阶数之和为4.

几何重数为2，以2为特征值的Jordan块的个数为2.

$$\boldsymbol{J}=\begin{pmatrix} 2 & & & \\ & 2 & 1 & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix} \quad 或\boldsymbol{J}=\begin{pmatrix} 2 & 1 & & \\ & 2 & & \\ & & 2 & 1 \\ & & & 2 \end{pmatrix}$$

(1) $l=1$ $\quad r_1=\text{rank}(\lambda_1\boldsymbol{I}-\boldsymbol{A})=\text{rank}(2\boldsymbol{I}-\boldsymbol{A})=2$

$$r_2=\text{rank}(\lambda_1\boldsymbol{I}-\boldsymbol{A})^2=\text{rank}(2\boldsymbol{I}-\boldsymbol{A})^2=0$$

以2为特征值, 阶数为1 的Jordan块的个数为

$$r_2+r_0-2r_1=0+4-2\times 2=0$$

(2) $l=2$

$$r_3=\text{rank}(\lambda_1\boldsymbol{I}-\boldsymbol{A})^3=\text{rank}(2\boldsymbol{I}-\boldsymbol{A})^3=0$$

以2为特征值, 阶数为2 的Jordan块的个数为

$$r_3+r_1-2r_2=0+2-2\times 0=2 \quad 故\boldsymbol{J}=\begin{pmatrix}2&1&&\\&2&&\\&&2&1\\&&&2\end{pmatrix}$$

(二) 关于变换矩阵T

$$\boldsymbol{A}=\boldsymbol{TJT}^{-1}\Rightarrow \boldsymbol{AT}=\boldsymbol{TJ} \qquad \boldsymbol{T}=(\boldsymbol{T}_1,\boldsymbol{T}_2,\ldots,\boldsymbol{T}_k),\ \boldsymbol{T}_i\text{为}n\times n_i\text{阶矩阵.}$$

$$\boldsymbol{AT}_i=\boldsymbol{T}_i\boldsymbol{J}_{n_i}(\lambda_i),\quad \boldsymbol{T}_i=(\boldsymbol{t}_1^i,\boldsymbol{t}_2^i,\ldots,\boldsymbol{t}_{n_i}^i),$$

$$\boldsymbol{A}(\boldsymbol{t}_1^i,\boldsymbol{t}_2^i,\ldots,\boldsymbol{t}_{n_i}^i)=(\boldsymbol{t}_1^i,\boldsymbol{t}_2^i,\ldots,\boldsymbol{t}_{n_i}^i)\begin{pmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i\end{pmatrix} \quad \begin{cases}\boldsymbol{At}_1^i=\lambda_i\boldsymbol{t}_1^i,\\ \boldsymbol{At}_2^i=\lambda_i\boldsymbol{t}_2^i+\boldsymbol{t}_1^i,\\ \quad\vdots\\ \boldsymbol{At}_{n_i}^i=\lambda_i\boldsymbol{t}_{n_i}^i+\boldsymbol{t}_{n_i-1}^i.\end{cases}$$

$\boldsymbol{t}_1^i,\boldsymbol{t}_2^i,\ldots,\boldsymbol{t}_{n_i}^i$ **构成一条关于特征值λ_i的长度为n_i的Jordan链.**

$$(\boldsymbol{A}-\lambda_i\boldsymbol{I}_n)\boldsymbol{t}_1^i=0,\quad (\boldsymbol{A}-\lambda_i\boldsymbol{I}_n)\boldsymbol{t}_j^i=\boldsymbol{t}_{j-1}^i,\ j=2,3,\ldots,n_i$$

$$(\boldsymbol{A}-\lambda_i\boldsymbol{I}_n)\boldsymbol{t}_1^i = 0, \quad (\boldsymbol{A}-\lambda_i\boldsymbol{I}_n)\boldsymbol{t}_j^i = \boldsymbol{t}_{j-1}^i,\ j=2,3,\ldots,n_i \qquad \textbf{(2-45)}$$

$\boldsymbol{t}_1^i$ 是矩阵$\boldsymbol{A}$的关于特征值λ_i的一个特征向量, 称为链首.

注意: 并不是任何一个特征向量都可以做链首, 还要求可以利用**(2-45)**求出**Jordan**链中的其余向量, 因此需要从λ_i的特征子空间中选取适当的向量作为链首, 使得方程组(2-45)可解.

例4 计算例2中矩阵A化Jordan标准型的变换矩阵T.

回忆

解

由A的Jordan标准型 $$J=\begin{pmatrix} -1 & & \\ & -1 & 1 \\ & & -1 \end{pmatrix}$$

λ_1对应两个Jordan块, 即有两条Jordan链, 长度为1和2.

求出λ_1所对应的线性无关的特征向量

$$x_1=(2,0,-1)^T, x_2=(0,1,0)^T.$$

但以x_1 或 x_2为链首时都无法求出下一个Jordan链向量.

需要$y\in span\{x_1,x_2\}$, 使得$(A-\lambda_1 I)z=y$可解

令$\boldsymbol{y}=k_1\boldsymbol{x}_1+k_2\boldsymbol{x}_2$

$$(\boldsymbol{A}-\lambda_1\boldsymbol{I}\mid\boldsymbol{y})=\begin{pmatrix}4&0&8&2k_1\\3&0&6&k_2\\-2&0&-4&-k_1\end{pmatrix}\to\begin{pmatrix}4&0&8&2k_1\\0&0&0&k_2-3k_1/2\\0&0&0&0\end{pmatrix}$$

需$2k_2-3k_1=0$ 取$k_1=2, k_2=3, \boldsymbol{y}=(4,3,-2)^T$

由$(\boldsymbol{A}-\lambda_1\boldsymbol{I})\boldsymbol{z}=\boldsymbol{y}$解出$\boldsymbol{z}=(1,0,0)^T$

由$\boldsymbol{J}=\begin{pmatrix}-1&&\\&-1&1\\&&-1\end{pmatrix}$ $\boldsymbol{T}=\begin{pmatrix}2&4&1\\0&3&0\\-1&-2&0\end{pmatrix}$ 或$\boldsymbol{T}=\begin{pmatrix}0&4&1\\1&3&0\\0&-2&0\end{pmatrix}$

定理2.12 (Hamilton-Caylay)

$$\text{设}\boldsymbol{A}\in\boldsymbol{C}^{n\times n},\psi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A}),\text{则}\quad\psi(\boldsymbol{A})=\boldsymbol{O}$$

证明　　存在$\boldsymbol{P}\in\boldsymbol{C}^{n\times n}$，使得$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$

$$\boldsymbol{J}=\begin{pmatrix}\lambda_1 & \delta & & \\ & \lambda_2 & \ddots & \\ & & \ddots & \delta \\ & & & \lambda_n\end{pmatrix},\delta=0\text{或者}1.$$

$$\psi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

$$\begin{aligned}\psi(\boldsymbol{A}) &= (\boldsymbol{A}-\lambda_1\boldsymbol{I})(\boldsymbol{A}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{A}-\lambda_n\boldsymbol{I})\\ &= (\boldsymbol{PJP}^{-1}-\lambda_1\boldsymbol{I})(\boldsymbol{PJP}^{-1}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{PJP}^{-1}-\lambda_n\boldsymbol{I})\\ &= \boldsymbol{P}(\boldsymbol{J}-\lambda_1\boldsymbol{I})\boldsymbol{P}^{-1}\boldsymbol{P}(\boldsymbol{J}-\lambda_2\boldsymbol{I})\boldsymbol{P}^{-1}\cdots\boldsymbol{P}(\boldsymbol{J}-\lambda_n\boldsymbol{I})\boldsymbol{P}^{-1}\\ &= \boldsymbol{P}(\boldsymbol{J}-\lambda_1\boldsymbol{I})(\boldsymbol{J}-\lambda_2\boldsymbol{I})\cdots(\boldsymbol{J}-\lambda_n\boldsymbol{I})\boldsymbol{P}^{-1}\end{aligned}$$

$$= \boldsymbol{P}\begin{pmatrix} 0 & \delta & & \\ & \lambda_2-\lambda_1 & \ddots & \\ & & \ddots & \delta \\ & & & \lambda_n-\lambda_1 \end{pmatrix}\begin{pmatrix} \lambda_1-\lambda_2 & \delta & & \\ & 0 & \ddots & \\ & & \ddots & \delta \\ & & & \lambda_n-\lambda_2 \end{pmatrix}\cdots\begin{pmatrix} \lambda_1-\lambda_n & \delta & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1}-\lambda_n & \delta \\ & & & 0 \end{pmatrix}\boldsymbol{P}^{-1}$$

$$= \boldsymbol{P}\begin{pmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{pmatrix}\cdots\begin{pmatrix} \lambda_1-\lambda_n & \delta & & \\ & \ddots & \ddots & \\ & & \lambda_{n-1}-\lambda_n & \delta \\ & & & 0 \end{pmatrix}\boldsymbol{P}^{-1}$$

$$= \cdots = \boldsymbol{O}_{n\times n}$$

例5 已知矩阵 $\boldsymbol{A}=\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$, 计算

(1) $\boldsymbol{A}^7-\boldsymbol{A}^5-19\boldsymbol{A}^4+28\boldsymbol{A}^3+6\boldsymbol{A}-4\boldsymbol{I};$

(2) $\boldsymbol{A}^{-1};$ (3) $\boldsymbol{A}^{100}.$

解

$$\psi(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix} \lambda+1 & -1 & 0 \\ 4 & \lambda-3 & 0 \\ -1 & 0 & \lambda-2 \end{vmatrix}=\lambda^3-4\lambda^2+5\lambda-2$$

(1)令$f(\lambda)=\lambda^7-\lambda^5-19\lambda^4+28\lambda^3+6\lambda-4$

$$=(\lambda^4+4\lambda^3+10\lambda^2+3\lambda-2)\psi(\lambda)-3\lambda^2+22\lambda-8$$

$$f(\boldsymbol{A})=-3\boldsymbol{A}^2+22\boldsymbol{A}-8\boldsymbol{I}=\begin{pmatrix} -19 & 6 & 0 \\ -64 & 43 & 0 \\ 19 & -3 & 24 \end{pmatrix}$$

(2)由$\psi(\boldsymbol{A}) = \boldsymbol{A}^3 - 4\boldsymbol{A}^2 + 5\boldsymbol{A} - 2\boldsymbol{I} = \boldsymbol{O}$得

$$\boldsymbol{A}\left(\frac{1}{2}(\boldsymbol{A}^2 - 4\boldsymbol{A} + 5\boldsymbol{I})\right) = \boldsymbol{I}$$

$$\boldsymbol{A}^{-1} = \frac{1}{2}(\boldsymbol{A}^2 - 4\boldsymbol{A} + 5\boldsymbol{I}) = \begin{pmatrix} 3 & -1 & 0 \\ 4 & -1 & 0 \\ -\frac{3}{4} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

(3)设$\lambda^{100}=g(\lambda)\psi(\lambda)+a\lambda^2+b\lambda+c$

由$\psi(\lambda)=(\lambda-2)(\lambda-1)^2$，有$\psi(2)=\psi(1)=\psi'(1)=0$

$$\begin{cases} 2^{100}=4a+2b+c \\ 1=a+b+c \\ 100=2a+b \end{cases} \Rightarrow \begin{cases} a=2^{100}-101 \\ b=-2^{101}+302 \\ c=2^{100}-200 \end{cases}$$

$$\boldsymbol{A}^{100}=g(\boldsymbol{A})\psi(\boldsymbol{A})+a\boldsymbol{A}^2+b\boldsymbol{A}+c\boldsymbol{I}=a\boldsymbol{A}^2+b\boldsymbol{A}+c\boldsymbol{I}$$

$$=\begin{pmatrix} -199 & 100 & 0 \\ -400 & 201 & 0 \\ 201-2^{100} & 2^{100}-101 & 2^{100} \end{pmatrix}.$$

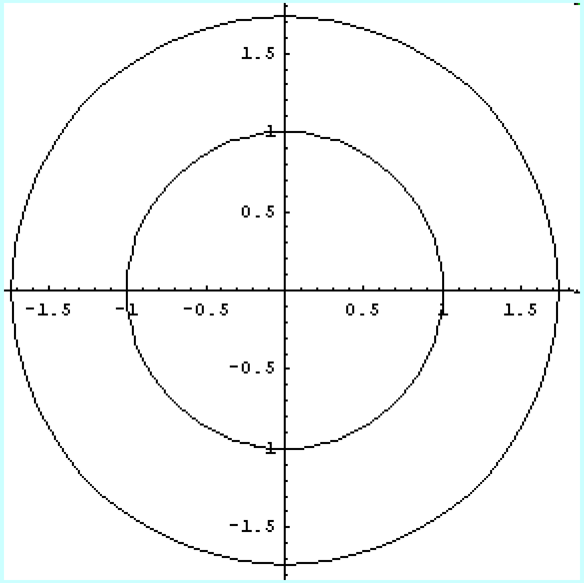
2.4 矩阵的奇异值分解

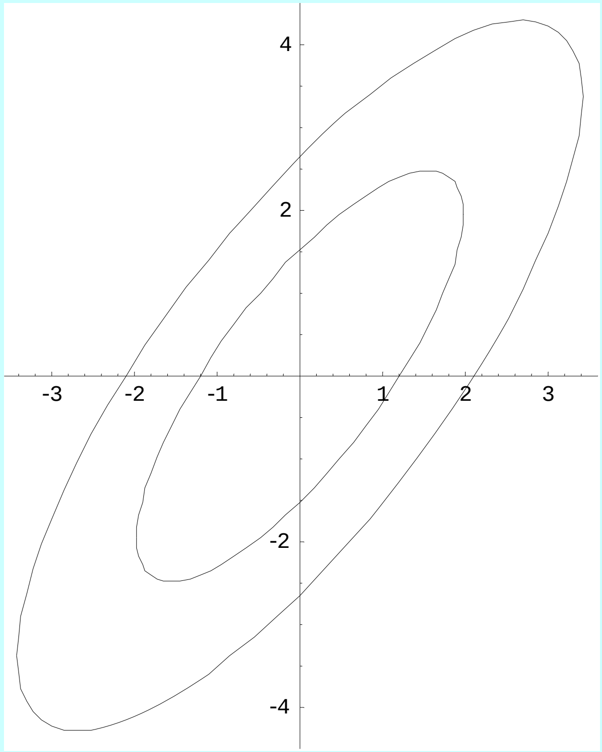
对于方阵, 利用其特征值和特征向量可以刻画矩阵的结构。对长方阵情形, 这些方法已经不适用. 而推广的特征值--矩阵的奇异值分解理论能改善这种情况。利用奇异值和奇异向量不仅可以刻画矩阵的本身结构, 而且还可以进一步刻画线代数方程组的解的结构, 是构造性的研究线代数问题的有利的工具。

2.4.1 矩阵奇异值分解的几何意义

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.96 & 1.72 \\ 2.28 & 0.96 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \boldsymbol{A}^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$x_1^2 + x_2^2 = 1, x_1^2 + x_2^2 = 3$$





$$A=\begin{pmatrix}0.96 & 1.72\\ 2.28 & 0.96\end{pmatrix}=\begin{pmatrix}0.6 & -0.8\\ 0.8 & 0.6\end{pmatrix}\begin{pmatrix}3 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}0.8 & 0.6\\ 0.6 & -0.8\end{pmatrix}^T=\boldsymbol{UDV}^T$$

$$\boldsymbol{AV}=\boldsymbol{UD}$$

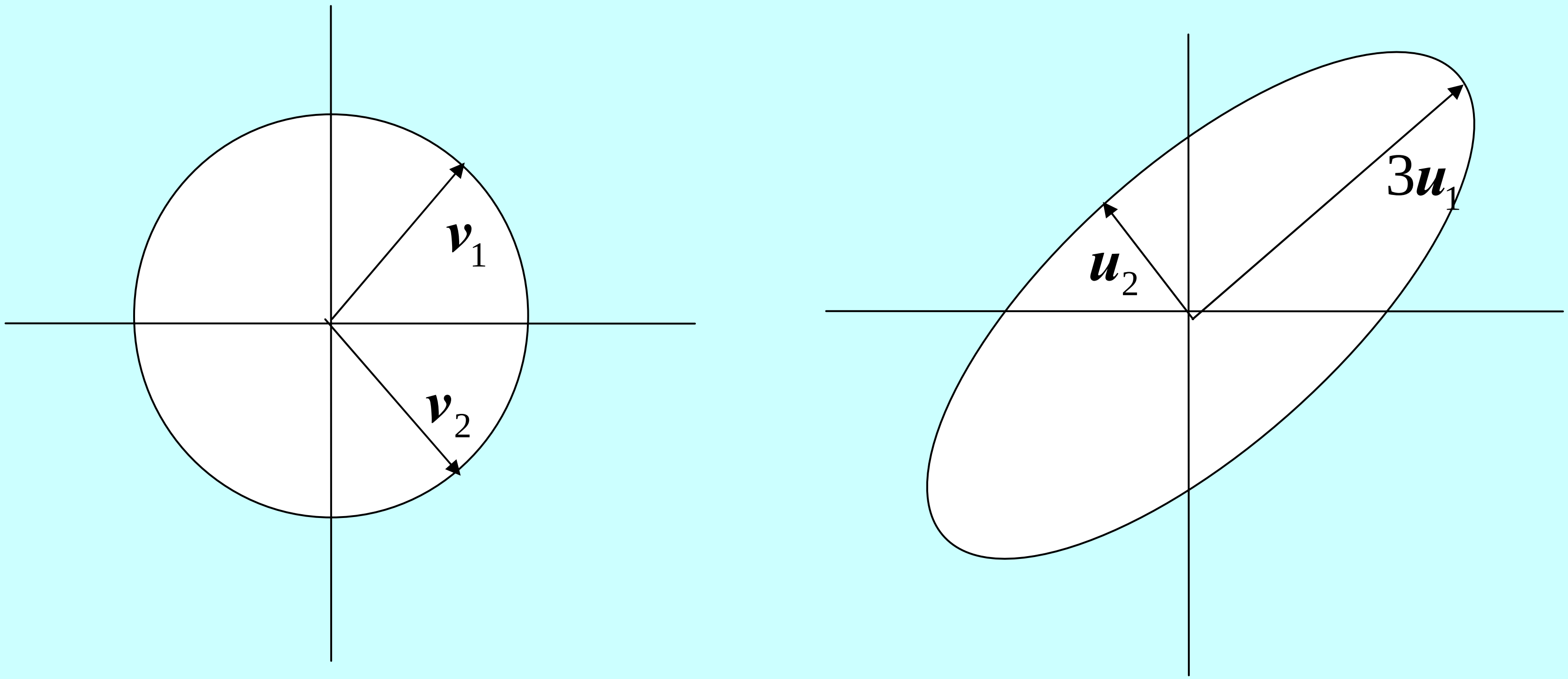
$$\boldsymbol{V}=(\boldsymbol{v}_1,\boldsymbol{v}_2),\boldsymbol{U}=(\boldsymbol{u}_1,\boldsymbol{u}_2),\boldsymbol{D}=\mathrm{diag}(\sigma_1,\sigma_2),$$

$\boldsymbol{v}_1=(0.8,0.6)^T,\boldsymbol{v}_2=(0.6,-0.8)^T,\boldsymbol{u}_1=(0.6,0.8)^T,\boldsymbol{u}_2=(-0.8,0.6)^T,$
$\sigma_1=3,\sigma_2=1$

$$\boldsymbol{v}_1^T\boldsymbol{v}_2=0,\boldsymbol{u}_1^T\boldsymbol{u}_2=0,$$

$$(\boldsymbol{Av}_1,\boldsymbol{Av}_2)=(3\boldsymbol{u}_1,\boldsymbol{u}_2)$$

$$(A\boldsymbol{v}_1, A\boldsymbol{v}_2) = (3\boldsymbol{u}_1, \boldsymbol{u}_2)$$



与特征值和特征向量对比：设A有两个线性无关(未必正交)的特征向量x_1, x_2对应的特征值为 λ_1, λ_2

$$(A\boldsymbol{x}_1, A\boldsymbol{x}_2) = (\lambda_1\boldsymbol{x}_1, \lambda_2\boldsymbol{x}_2)$$

对一般的 $\boldsymbol{A}\in\boldsymbol{R}^{m\times n}$，不妨设 $m\geq n,\mathrm{rank}(\boldsymbol{A})=r$，将其分解为

$$\boldsymbol{A}=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_m)\begin{pmatrix}\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_n)\\ \boldsymbol{O}_{(m-n)\times n}\end{pmatrix}(\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n)^T$$

$$=\boldsymbol{U}\Sigma\boldsymbol{V}^{\mathrm{T}}$$

其中 $\boldsymbol{U}$ 和 $\boldsymbol{V}$ 分别是 $\boldsymbol{m}$ 阶和 $\boldsymbol{n}$ 阶正交阵.

$$\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>0,\ \sigma_{r+1}=\sigma_{r+2}=\cdots=\sigma_n=0$$

$$AV = U\Sigma$$

则y=Ax 是将$\mathbf{R}^n$ 中的单位球 $S^n = \{x \in \mathbf{R}^n \big| \|x\|_2 = 1\}$ 变成了$\mathbf{R}^m$ 中的"超椭球" $\mathbf{E}^m = \{y \in \mathbf{R}^m \big| y = Ax, x \in \mathbf{R}^n, \|x\|_2 = 1\}$. $\mathbf{R}^m$中的"超椭球"就是将 $\mathbf{R}^m$ 中的单位球沿某些正交方向$u_1, u_2, \cdots, u_m$ 分别以拉伸因子 $\sigma_1, \sigma_2, \cdots, \sigma_n, \underbrace{0, \cdots, 0}_{m-n}$ 拉伸而成的曲面, $\{\sigma_i u_i\}$ 为 $\mathbf{E}^m$ 的主半轴, $\sigma_1, \sigma_2, \cdots, \sigma_r$ 为主半轴的长度, 它恰好是矩阵A的奇异值.

定义2.10 设 $A \in \mathbf{C}^{m\times n}$,

Hermite半正定矩阵 $A^H A$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$$

称非负实数

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

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　$U^HAV=B$, 则有

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

即 B^HB 与 A^HA相似, 故它们具有相同的特征值, 进而命题得证。

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

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

其中

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

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

证明 $rank(A^H A) = rank(AA^H) = rank(A^H) = rank(A)$

$$\boldsymbol{Ax} = 0 \Rightarrow \boldsymbol{A}^H \boldsymbol{Ax} = 0$$

$$\boldsymbol{A}^H \boldsymbol{Ax} = 0 \Rightarrow \boldsymbol{x}^H \boldsymbol{A}^H \boldsymbol{Ax} = 0 \Rightarrow (\boldsymbol{Ax})^H \boldsymbol{Ax} = 0 \Rightarrow \boldsymbol{Ax} = 0$$

方程组同解，则系数矩阵的秩相等.

$$\boldsymbol{A}^H \boldsymbol{Ax} = \lambda \boldsymbol{x} \Rightarrow \boldsymbol{AA}^H \boldsymbol{Ax} = \boldsymbol{A}\lambda \boldsymbol{x} \Rightarrow \boldsymbol{AA}^H (\boldsymbol{Ax}) = \lambda (\boldsymbol{Ax})$$

说明$\boldsymbol{A}^H\boldsymbol{A}$ 和 $\boldsymbol{AA}^H$ 的非零特征值相等, 也即秩相等.

$\boldsymbol{A}^H\boldsymbol{A}$ 是秩为$\boldsymbol{r}$ 的$\boldsymbol{n}$ 阶**Hermite**半正定矩阵, 由**Schur**定理的推论**2.2**, 必存在$\boldsymbol{n}$ 阶酉阵, 使得

$$\boldsymbol{V}^H(\boldsymbol{A}^H \boldsymbol{A})\boldsymbol{V} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}^2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}\begin{pmatrix}\boldsymbol{\Sigma}^2 & 0\\ 0 & 0\end{pmatrix}$$

$$\boldsymbol{V}=\left(\boldsymbol{V}_1 \quad | \quad \boldsymbol{V}_2\right),\quad \boldsymbol{V}_1\in\mathbf{C}^{n\times r},\qquad \boldsymbol{V}_2\in\mathbf{C}^{n\times(n-r)}$$

$$\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_1=\boldsymbol{V}_1\boldsymbol{\Sigma}^2,\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_2=0.$$

(1) $\boldsymbol{V}_1^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_1=\boldsymbol{\Sigma}^2 \Rightarrow (\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1})^H(\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1})=\boldsymbol{I}_r$

(2) $\boldsymbol{V}_2^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_2=\boldsymbol{O} \Rightarrow \boldsymbol{A}\boldsymbol{V}_2=\boldsymbol{O}$

令$\boldsymbol{U}_1=\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1}\in\boldsymbol{C}^{m\times r}$ $\Rightarrow \boldsymbol{U}_1\boldsymbol{\Sigma}=\boldsymbol{A}\boldsymbol{V}_1 \qquad \boldsymbol{U}_1^H\boldsymbol{U}_1=\boldsymbol{I}_r$

因此矩阵 $U_1=(u_1,u_2,\cdots,u_r)$ 的列是C^m 中的一个标准正交向量组, 将其扩充为C^m 的一组标准正交基

$$u_1,u_2,\cdots,u_r,u_{r+1},u_{r+2},\cdots,u_m$$

令$U_2=(u_{r+1},u_{r+2},\cdots,u_m)$ 则$U=(U_1,U_2)$是一个m阶酉阵

且$U_1^H U_1=I_r,\quad U_2^H U_1=O$

于是$U^H AV=U^H(AV_1,AV_2)$

$$U^H AV=\begin{pmatrix}U_1^H\\U_2^H\end{pmatrix}\begin{pmatrix}U_1\Sigma & O\end{pmatrix}=\begin{pmatrix}U_1^H U_1\Sigma & O\\U_2^H U_1\Sigma & O\end{pmatrix}=\begin{pmatrix}\Sigma & O\\O & O\end{pmatrix}$$

（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。

$$A = U_1 \Sigma V_1^H \quad \Rightarrow AV_1 = U_1 \Sigma, \quad U_1^H A = \Sigma V_1^H$$

$$Av_i = \sigma_i u_i, \quad u_i^H A = \sigma_i v_i^H, \quad i = 1,2,\cdots,r$$

U与V的列向量 $u_1, u_2, \cdots, u_m$和 $v_1, v_2, \cdots, v_n$ 分别称为矩阵A的与奇异值σ_i对应的左奇异向量和右奇异向量.

$$AA^H U = UU^H A(VV^H)A^H U = U(U^H AV)(U^H AV)^H = U\begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$A^H AV = VV^H A^H (UU^H)AV = V(U^H AV)^H (U^H AV) = V\begin{pmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$$

左奇异向量 $u_1, u_2, \cdots, u_m$为AA^H的单位正交特征向量,

右奇异向量 $v_1, v_2, \cdots, v_n$ 为$A^H A$的单位正交特征向量.

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

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

令

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

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

则 $\boldsymbol{A}^H\boldsymbol{A}$ 的特征值和$\boldsymbol{A}$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

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

再求出$\boldsymbol{V}$，由

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

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

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

对应的标准正交的特征向量为:

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

事实上, 实对称矩阵或**Hermite**矩阵的特征值为实数, 对应不同特征值的特征向量必正交.

$$\boldsymbol{Ax}=\lambda\boldsymbol{x} \Rightarrow \boldsymbol{x}^H\boldsymbol{A}^H=\overline{\lambda}\boldsymbol{x}^H \Rightarrow \boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{x}=\overline{\lambda}\boldsymbol{x}^H\boldsymbol{x}=\overline{\lambda}\|\boldsymbol{x}\|_2^2$$

$$\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{x}=\boldsymbol{x}^H\boldsymbol{Ax}=\boldsymbol{x}^H\lambda\boldsymbol{x}=\lambda\|\boldsymbol{x}\|_2^2 \Rightarrow \lambda=\overline{\lambda}$$

$$\boldsymbol{Ay}=\mu\boldsymbol{y}(\mu\neq\lambda) \quad (\boldsymbol{Ax},\boldsymbol{y})=\boldsymbol{y}^H\boldsymbol{Ax}=\lambda\boldsymbol{y}^H\boldsymbol{x}=\lambda(\boldsymbol{x},\boldsymbol{y})$$

$$=\boldsymbol{y}^H\boldsymbol{A}^H\boldsymbol{x}=(\boldsymbol{Ay})^H\boldsymbol{x}=(\mu\boldsymbol{y})^H\boldsymbol{x}=\overline{\mu}\boldsymbol{y}^H\boldsymbol{x}=\overline{\mu}(\boldsymbol{x},\boldsymbol{y}) \Rightarrow (\boldsymbol{x},\boldsymbol{y})=0$$

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为实对角矩阵.

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，

$$(U_1)_{3\times 2}=AV_1\Sigma^{-1}=\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{3}} & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix}$$

得约化的奇异值分解

$$A = U_1 \Sigma V_1^H = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \end{pmatrix}$$

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

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

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

是酉阵, 故矩阵A的奇异值分解（满的奇异值分解）为

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

2.4.3 利用矩阵的奇异值分解讨论矩阵的性质

定理2.15 矩阵A的非零奇异值的个数恰为矩阵A的秩.

$$rank(A^H A) = rank(AA^H) = rank(A^H) = rank(A)$$

定理2.16 $\mathbf{R}(A) = \text{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_r\},\ N(A) = \text{span}\{\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \cdots, \boldsymbol{v}_n\}$

其中 $\boldsymbol{u}_i$ 和 $\boldsymbol{v}_i$ 分别为矩阵U和V的正交向量; $\mathbf{R}(A)$为由A的列向量生成的子空间, 称为A的值域或像空间. 即

$$R(A) = \text{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$$

$\mathbf{N}(A)$称为A的零空间或核空间, 即 $\mathbf{N}(A) = \{\boldsymbol{x} \mid A\boldsymbol{x} = \boldsymbol{0}\}$

$$\boldsymbol{AV}_1 = \boldsymbol{U}_1\boldsymbol{\Sigma} \qquad \boldsymbol{AV}_2 = \boldsymbol{O}$$

定理2.17 设 $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ 则

$$\|A\|_2 = \sigma_1 \qquad \|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2}$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)} \qquad \|A\|_F = \sqrt{\mathrm{trace}(A^H A)}$$

定理2.18 如果A为Hermite矩阵, 则A的奇异值即为A的特征值的绝对值.

$$A^H A = A^2$$

$$\sigma(A) = \sqrt{\lambda(A^H A)} = \sqrt{\lambda(A^2)} = \sqrt{\lambda(A)^2} = |\lambda(A)|$$

定理2.19 如果$\boldsymbol{A}$为 $\boldsymbol{n}$ 阶方阵, 则 $$\left|\det(\boldsymbol{A})\right| = \prod_{i=1}^{n} \sigma_i$$

$$\sqrt{\det(\boldsymbol{A}^H \boldsymbol{A})} = \sqrt{\prod_{i=1}^{n} \sigma_i^2} = \prod_{i=1}^{n} \sigma_i$$

$$= \sqrt{\det(\boldsymbol{A}^H)\det(\boldsymbol{A})} = \sqrt{\overline{\det(\boldsymbol{A})}\det(\boldsymbol{A})} = \sqrt{|\det(\boldsymbol{A})|^2}$$

定理2.20 秩为r 的 $m\times n$ 阶矩阵A可以表示为 r 个秩为1的矩阵的和

$$A=\sigma_1 \boldsymbol{u}_1\boldsymbol{v}_1^H+\sigma_2 \boldsymbol{u}_2\boldsymbol{v}_2^H+\cdots+\sigma_r \boldsymbol{u}_r\boldsymbol{v}_r^H$$

$$A=U_1\Sigma V_1^H=(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_r)\begin{pmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{pmatrix}\begin{pmatrix}\boldsymbol{v}_1^H\\ \vdots \\ \boldsymbol{v}_r^H\end{pmatrix}$$

$$=(\sigma_1\boldsymbol{u}_1,\ldots,\sigma_r\boldsymbol{u}_r)\begin{pmatrix}\boldsymbol{v}_1^H\\ \vdots \\ \boldsymbol{v}_r^H\end{pmatrix}=\sigma_1 \boldsymbol{u}_1\boldsymbol{v}_1^H+\cdots+\sigma_r \boldsymbol{u}_r\boldsymbol{v}_r^H$$

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

此章结束

作者姓名：张宏伟、金光日、李崇君

工作单位：大连理工大学数学科学学院

联系方式：E-mail: chongjun@dlut.edu.cn