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

Gauss消去法

2.1.1 与

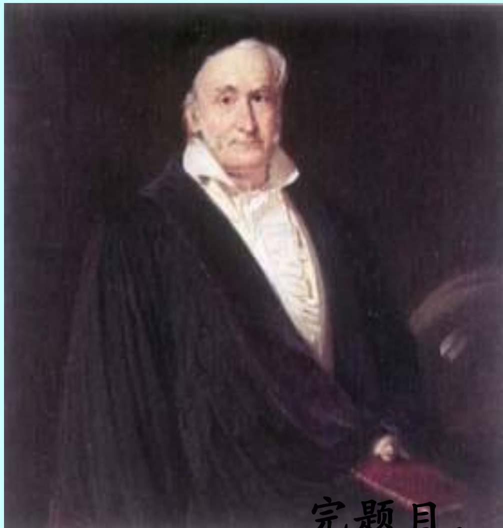
矩阵的LU分解

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。

从方程组角度考虑Gauss消去法

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

$\left(-\dfrac{4}{2}\right)\times r_1^{(0)} + r_2^{(0)}$、$\left(-\dfrac{8}{2}\right)\times r_1^{(0)} + r_3^{(0)}$**和**$\left(-\dfrac{6}{2}\right)\times r_1^{(0)} + r_4^{(0)}$**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

$$\left(-\frac{2}{2}\right) \times r_3^{(2)} + r_4^{(1)} \qquad 得$$

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

$$2x_1 + x_2 + x_3 = 4 \Rightarrow 2x_1 + 1 + 1 = 4 - 2$$

$$x_2 + x_3 + x_4 = 3 \Rightarrow x_2 + 1 + 1 = 3 - 2$$

$$2x_3 + 2x_4 = 4 \Rightarrow 2x_3 + 2 \times 1 = 4 - 2$$

$$2x_4 = 2 \Rightarrow x_4 = \frac{2}{2} = 1$$

上述为回代求解过程，得 $x = (1,1,1,1)^T$。

从增广矩阵角度考虑Gauss消去法

Gauss消去法:

增广矩阵$(A\,|\,b)$ $\xrightarrow{\text{初等行变换}}$ 上三角矩阵$(U\,|\,c)$

$\xrightarrow{\text{回代}}$ $Ux=c$ 的解 $\Longleftrightarrow$ $Ax=b$ 的解

增广矩阵 $(A|b)$ 化成上三角矩阵 $(U|c)$

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

解 $(A\,|\,b)$

$$\xrightarrow{\text{第三次消元}} \begin{pmatrix} 2 & 1 & 1 & 0 & 4 \\ 0 & 1 & 1 & 1 & 3 \\ 0 & 0 & 2 & 2 & 4 \\ 0 & 0 & 0 & 2 & 6 \end{pmatrix} (U\,|\,c)$$

矩阵变换角度考虑Gauss消去法

三次消元过程写成矩阵的形式分别为:

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

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

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

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

$$
L_i = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & -l_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & -l_{ni} & & & 1 \end{pmatrix} \Longrightarrow L_i^{-1} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & l_{i+1,i} & \ddots & & \\ & & \vdots & & \ddots & \\ & & l_{ni} & & & 1 \end{pmatrix}
$$

注意到

$$
L_i L_k = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & -l_{i+1,i} & \ddots & & & & \\ & & \vdots & & 1 & & & \\ & & \vdots & & -l_{k+1,k} & \ddots & & \\ & & \vdots & & \vdots & & \ddots & \\ & & -l_{ni} & & -l_{nk} & & & 1 \end{pmatrix}
$$

从而

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

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

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

刚才的计算过程可以表示为

$$L_3 L_2 L_1 A = U$$

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

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

$$\boldsymbol{A}$$

即得到矩阵 $\boldsymbol{A}$ 的 $\boldsymbol{LU}$ 分解。

定义2.1 对于 n 阶矩阵 $\boldsymbol{A}$，如果存在 n 阶单位下三角矩阵 $\boldsymbol{L}$ 和 n 阶上三角矩阵 $\boldsymbol{U}$，使得

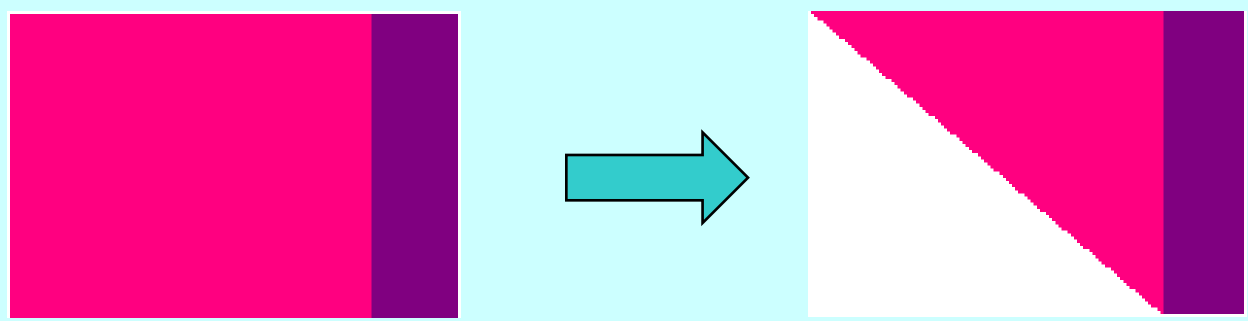
$$\boldsymbol{A} = \boldsymbol{LU}$$

则称其为矩阵 $\boldsymbol{A}$ 的 $\boldsymbol{LU}$ 分解，也称Doolittle分解。

Doolittle方法求解线性方程组：

思路

将A化为上三角阵，再回代求解。



$$Ax=b \Leftrightarrow (LU)x=b \Leftrightarrow \begin{cases} Ly=b \\ Ux=y \end{cases}$$

LU分解方法及其计算量

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

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

运算量(乘除法)： $(n-1)*(n+1)$

行乘子(数)

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

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

$$\boldsymbol{L}_1(\boldsymbol{A} \mid \boldsymbol{b}) = \begin{pmatrix} a_{11}^{(0)} & a_{12}^{(0)} & \cdots & a_{1n}^{(0)} & b_1^{(0)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} & b_2^{(1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} & b_n^{(1)} \end{pmatrix}$$

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

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

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

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

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

$$\boldsymbol{L}_k \cdots \boldsymbol{L}_2 \boldsymbol{L}_1 (\boldsymbol{A} \,|\, \boldsymbol{b}) = \begin{pmatrix} a_{1,1}^{(0)} & \cdots & a_{1,k}^{(0)} & a_{1,k+1}^{(0)} & \cdots & a_{1,n}^{(0)} & b_1^{(0)} \\ & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & a_{kk}^{(k-1)} & a_{k,k+1}^{(k-1)} & \cdots & a_{k,n}^{(k-1)} & b_k^{(k-1)} \\ & & & a_{k+1,k+1}^{(k)} & \cdots & a_{k+1,n}^{(k)} & b_{k+1}^{(k)} \\ & & & \vdots & \ddots & \vdots & \vdots \\ & & & a_{n,k+1}^{(k)} & \cdots & a_{n,n}^{(k)} & b_n^{(k)} \end{pmatrix}$$

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：

$$\boldsymbol{L}_{n-1}\boldsymbol{L}_{n-2}\cdots\boldsymbol{L}_2\boldsymbol{L}_1(\boldsymbol{A}\,|\,\boldsymbol{b})=(\boldsymbol{U}\,|\,\boldsymbol{c})$$

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

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

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

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

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

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

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

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 VS LU分解方法

对于单个线性方程组 $Ax = b$

$$(A,\ b) \longrightarrow (U,\ c)$$

$$Ux = c$$

$$A = LU$$

$$Ly = b$$

$$Ux = y$$

计算量相同

Gauss消去法 VS LU分解方法

$Ax=b_1$ $(A,b_1)\Rightarrow(U,c_1)$ $Ux=c_1$

$\cdots$ $\Downarrow$ $\cdots$

$Ax=b_m$ $(A,b_m)\Rightarrow(U,c_m)$ $Ux=c_m$

m次LU分解

对于多个线性方程组 $Ax=b_1,\dots,b_n$

$Ly_1=b_1$ $Ux=y_1$

$A=LU$

$Ly_m=b_m$ $Ux=y_m$

1次LU分解

Gauss消去法 VS LU分解方法

对于迭代公式 $Ax^{(k+1)} = x^{(k)}$， $x^{(0)}$ 给定，求 $x^{(m)}$

$$\left(A, x^{(0)}\right) \Rightarrow \left(U, y^{(0)}\right)$$
$$Ux^{(1)} = y^{(0)}$$
$$\vdots$$
$$\left(A, x^{(m-1)}\right) \Rightarrow \left(U, y^{(m-1)}\right)$$
$$Ux^{(m)} = y^{(m-1)}$$

m次LU分解

$$A = LU$$

1次LU分解

$$Ly^{(0)} = x^{(0)}$$
$$Ux^{(1)} = y^{(0)}$$
$$\vdots$$
$$Ly^{(m)} = x^{(m-1)}$$
$$Ux^{(m)} = y^{(m)}$$

右端项无法直接列出，LU分解方法节省计算量

Gauss消去法 VS LU分解方法

对于线性方程组 $A^2B^2x=b$

1、首先计算系数矩阵，再LU分解方法

$\frac{10}{3}n^3$ $C=A^2B^2 \quad Cx=b$

V 2、单个矩阵递进计算

$\frac{4}{3}n^3$ $AABBx=b \quad Az_1=b, Az_2=z_1, Bz_3=z_2, Bx=z_3$

V 3、LU分解算法

$\frac{2}{3}n^3$ $A=L_1U_1, B=L_2U_2 \quad L_1U_1L_1U_1L_2U_2L_2U_2x=b$

Gauss消去法可执行的条件?

回忆

A 的k阶顺序主子式 主元 $a_{kk}^{(k-1)} \neq 0$

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

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

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

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

充分条件

定理2.1（矩阵LU分解的存在和唯一性）

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 $\boldsymbol{L}_1\boldsymbol{U}_1=\boldsymbol{L}_2\boldsymbol{U}_2$

$$\boldsymbol{L}_2^{-1}\boldsymbol{L}_1=\boldsymbol{U}_2\boldsymbol{U}_1^{-1}=\boldsymbol{I} \qquad \boldsymbol{L}_2=\boldsymbol{L}_1,\boldsymbol{U}_2=\boldsymbol{U}_1$$

单位下三角角阵

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

DUT 大连理工大学 DALIAN UNIVERSITY OF TECHNOLOGY

$$\boldsymbol{B}=\begin{pmatrix}1&1&1\\2&2&1\\3&3&1\end{pmatrix}\xrightarrow{\boldsymbol{L}_1^{-1}=\begin{pmatrix}1&&\\-2&1&\\-3&&1\end{pmatrix}}\begin{pmatrix}1&2&3\\0&0&-1\\0&0&-2\end{pmatrix}=\boldsymbol{U}_1$$

$$\xrightarrow{\boldsymbol{L}_2^{-1}=\begin{pmatrix}1&&\\&1&\\&-2&1\end{pmatrix}}\begin{pmatrix}1&2&3\\0&0&-1\\0&0&0\end{pmatrix}=\boldsymbol{U}_2$$

可以LU分解，但不唯一!

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!

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：

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

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

显然 $(U\,|\,c)$有无穷多解．但实际上， $\det(\boldsymbol{A})\neq 0$ 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　$$(\boldsymbol{A}\,|\,\boldsymbol{b})=\begin{pmatrix}10^{-8} & 2 & 3 & 1\\ -1 & 3.712 & 4.623 & 2\\ -2 & 1.072 & 5.643 & 3\end{pmatrix}$$

$$\xrightarrow{\text{选列主元，交换第一和第三行}}\begin{pmatrix}-2 & 1.072 & 5.643 & 3\\ -1 & 3.712 & 4.623 & 2\\ 10^{-8} & 2 & 3 & 1\end{pmatrix}\xrightarrow{\text{第一次消元}}\cdots\xrightarrow{\text{第二次消元}}\begin{pmatrix}-2 & 1.072 & 5.643 & 3\\ 0 & 0.3176\times 10 & 0.18015\times 10 & 0.5\\ 0 & 0 & 0.1865\times 10 & 0.6851\end{pmatrix}$$

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

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

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

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

[illegible]

$$= U$$

上述过程可以表示为

$$L_3P_3L_2P_2L_1P_1A=U$$

显然，$L_3P_3L_2P_2L_1P_1$ 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_3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

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

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

得到另一种形式的矩阵分解：

$$PA = LU$$

P A L U

如果A为n阶方阵，进行Gauss列主元消去过程为

$$\boldsymbol{L}_{n-1}\boldsymbol{P}_{n-1}\cdots\boldsymbol{L}_2\boldsymbol{P}_2\boldsymbol{L}_1\boldsymbol{P}_1\boldsymbol{A}=\boldsymbol{U}$$

类似的，可以改写成

$$(\boldsymbol{L}_{n-1}\tilde{\boldsymbol{L}}_{n-2}\cdots\tilde{\boldsymbol{L}}_2\tilde{\boldsymbol{L}}_1)(\boldsymbol{P}_{n-1}\cdots\boldsymbol{P}_2\boldsymbol{P}_1)\boldsymbol{A}=\boldsymbol{U}$$

其中 $\tilde{\boldsymbol{L}}_k=\boldsymbol{P}_{n-1}\cdots\boldsymbol{P}_{k+1}\boldsymbol{L}_k\boldsymbol{P}_{k+1}\cdots\boldsymbol{P}_{n-1}$ （$k=1,2,\cdots,n-2$）

与L_k的结构相同，仅下三角部分元素经过了对调。

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

则

$$\boldsymbol{PA}=\boldsymbol{LU}$$

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

$$PA = LU \qquad (P,\ L\text{可以不同，分解不唯一})$$

LU分解不一定存在；若存在则唯一。列主元LU分解一定存在，但不一定唯一（P、L不唯一）。

$$\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U} \qquad \det(\boldsymbol{A})=\det(\boldsymbol{L})\det(\boldsymbol{U})=\det(\boldsymbol{U})$$

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

$$\boldsymbol{P}\boldsymbol{A}=\boldsymbol{L}\boldsymbol{U} \qquad \det(\boldsymbol{P})\det(\boldsymbol{A})=\det(\boldsymbol{P}\boldsymbol{A})=\det(\boldsymbol{L})\det(\boldsymbol{U})=\det(\boldsymbol{U})$$

$$\det(\boldsymbol{P})=(-1)^s, s\text{为换行次数}$$

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \iff \boldsymbol{P}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{P}\boldsymbol{b}$$

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

例 用Gauss列主元消去法解如下方程组并给出 $PA=LU$分解。

解：

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

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

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

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

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

用回代法求的解得:

$$x_3 = \frac{5}{2} \qquad x_2 = \frac{-2+\dfrac{5}{2}}{-6} = -\frac{1}{12} \qquad x_1 = -\frac{5}{6}$$

即 $\boldsymbol{x} = \left(-\dfrac{5}{6}, -\dfrac{1}{12}, \dfrac{5}{2}\right)^T$。

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 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 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: $$\boldsymbol{PA}=\boldsymbol{LU}$$

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

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：

$$\begin{pmatrix} -3 & 2 & 6 & 4 \\ 10 & -7 & 0 & 7 \\ 5 & -1 & 5 & 6 \end{pmatrix} \xrightarrow{P_1=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ -3 & 2 & 6 & 4 \\ 5 & -1 & 5 & 6 \end{pmatrix}$$

$$\xrightarrow{L_1=\begin{pmatrix} 1 & 0 & 0 \\ 3/10 & 1 & 0 \\ -1/2 & 0 & 1 \end{pmatrix}} \begin{pmatrix} 10 & -7 & 0 & 7 \\ 0 & -\frac{1}{10} & 6 & \frac{61}{10} \\ 0 & \frac{5}{2} & 5 & \frac{5}{2} \end{pmatrix}$$

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

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

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

$$\det(\boldsymbol{A}) = (-1)^2 \det(\boldsymbol{U}) = 10 \times \frac{5}{2} \times \frac{31}{5} = 155$$

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

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

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。

对称正定矩阵的Cholesky分解：

- 证明过程
- 直接分解方法

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

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

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

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

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

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

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

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

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

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

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

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$j=1,2,\cdots,n$ 计算

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

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

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

（2）求解下三角形方程组 $Ly=b$

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

（3）求解上三角线性方程组$L^T x = y$

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

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

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

解

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

(2) 求解$Ly=b$，得$y=\{2,3.5,1\}^{\mathrm{T}}$

(3) 求解$L^{\mathrm{T}}x=y$，得$x=\{1,1,1\}^{\mathrm{T}}$

2.1.4 三对角矩阵的三角分解

设三对角矩阵

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

如果矩阵A可以进行LU分解$A=LU$,其中

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

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

$$d_i = c_i$$

$$u_1 = b_1, \quad u_i = b_i - l_i \cdot c_{i-1}$$

$$l_i = a_i / u_{i-1}$$

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

计算次序 $u_1 \to l_2 \to u_2 \to l_3 \to u_3 \to \cdots \to l_n \to u_n$

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：

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

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

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

$$y_1 = f_1$$

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

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

计算量：乘除法n-1,加减法n-1

(3) 求解上三角形方程组

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

$$x_n = y_n / u_n$$

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

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

解方程计算量：乘除法2(n-1)+1,加减法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) $|b_i| \geq |a_i| + |c_i|, \quad a_i c_i \neq 0, i = 2,3,\cdots n-1$

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.

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

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

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

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

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

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

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

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）

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

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

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：$x=(10,-2)^{\mathrm{T}}$，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。

寻找刻画矩阵和方程组“病态”标准的量

求解 $Ax=b$ 时，A 和 b 的误差对解 x 有何影响？

$$\frac{\|\delta x\|}{\|x\|} ?????? \frac{\|\delta b\|}{\|b\|}$$

设 A 精确，b 有误差 δb，得到的解为 $x+\delta x$，即

$$\boxed{A(x+\delta x)=b+\delta b}$$

绝对误差放大因子

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

相对误差放大因子

又 $\|b\| = \|Ax\| \le \|A\| \cdot \|x\| \Rightarrow \frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}$

相对误差

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

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

常用的条件数为:

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

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

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

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

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

故

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

矩阵的条件数具有如下的性质:

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

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

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

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

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

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

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

(4) 如果U为酉（正交）矩阵，则

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

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

由第一章习题4酉矩阵与谱范数的性质可得

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

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

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

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

- $\mathrm{cond}(\boldsymbol{A})$ 越大，解的相对误差界可能越大，对求解线性方程组来说就越可能呈现病态。
- $\mathrm{cond}(\boldsymbol{A})$ 多大A算病态，通常没有具体的定量标准；
- $\mathrm{cond}(\boldsymbol{A})$ 越小，解的相对误差界越小，呈现良态。

n阶Hilbert矩阵

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

$$\mathrm{cond}_2(\boldsymbol{H}_4) = 1.5514 \times 10^4$$

$$\mathrm{cond}_2(\boldsymbol{H}_6) = 1.4951 \times 10^7$$

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

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

数值例子：

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

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

则A的条件数为：

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

解的相对误差界为：

$$\frac{\|\boldsymbol{\delta x}\|_\infty}{\|\boldsymbol{x}\|_\infty}\le\operatorname{cond}_\infty(\boldsymbol{A})\frac{\|\boldsymbol{\delta b}\|_\infty}{\|\boldsymbol{b}\|_\infty}$$
$$\approx4.8\times10^6\times\frac{0.00001}{8}$$
$$\approx6\approx600\%$$

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

原方程组 $Ax=b$ ➡ 扰动后的方程组 $(A+\delta A)(x+\delta x)=b+\delta b$

$$\frac{\|\delta x\|}{\|x\|}=\frac{\|(A+\delta A)^{-1}(\delta b-\delta Ax)\|}{\|x\|}$$

$$\leq\|(A+\delta A)^{-1}\|\left(\frac{\|\delta b\|}{\|x\|}+\|\delta A\|\right)$$

$$\leq\|A^{-1}\|\|(I+\delta A\bullet A^{-1})^{-1}\|\left(\frac{\|\delta b\|}{\|x\|}+\|\delta A\|\right)$$

$$\leq\frac{\|A^{-1}\|}{1-\|\delta A\|\|A^{-1}\|}\left(\frac{\|A\|\|\delta b\|}{\|b\|}+\|\delta A\|\right)$$

$$\leq\frac{\|A^{-1}\|\|A\|}{1-\|\delta A\|\|A^{-1}\|}\left(\frac{\|\delta b\|}{\|b\|}+\frac{\|\delta A\|}{\|A\|}\right).$$

定理： 对于线性方程组 $Ax=b$, A 为非奇异矩阵，δA 和 δb 分别为 A 和 b 的扰动。若扰动 δA 非常小,使得

$$\|A^{-1}\|\|\delta A\|<1,$$

则

$$\frac{\|\delta x\|}{\|x\|}\leq\frac{\|A\|\|A^{-1}\|}{1-\|A^{-1}\|\|\delta A\|}\left(\frac{\|\delta A\|}{\|A\|}+\frac{\|\delta b\|}{\|b\|}\right)$$

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

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

其中称$\|\boldsymbol{b}-\boldsymbol{A}\tilde{\boldsymbol{x}}\|$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 $Ax=b$ $\quad b-A\tilde{x}=Ax-A\tilde{x}=A(x-\tilde{x})$

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

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

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

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。

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

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

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

$$\mathbf{Ax}=\mathbf{b} \Leftrightarrow \mathbf{LUx}=\mathbf{b} \Leftrightarrow \begin{cases} \mathbf{Ly}=\mathbf{b} \\ \mathbf{Ux}=\mathbf{y} \end{cases}$$

- 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

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

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

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

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（$\mathbf{Q}^{-1}=\mathbf{Q}^{T}$），则

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

$$cond_2(\boldsymbol{QA})=cond_2(\boldsymbol{AQ})=cond_2(\boldsymbol{A})$$

若$A=QR$，Q为正交阵，R为上三角阵

$$\mathbf{Ax}=\mathbf{b}\Leftrightarrow\mathbf{QRx}=\mathbf{b}\Leftrightarrow\begin{cases}\mathbf{Qy}=\mathbf{b}\\ \mathbf{Rx}=\mathbf{y}\end{cases}\Leftrightarrow\mathbf{Rx}=\mathbf{Q}^T\mathbf{b}$$

利用正交变换实现矩阵的上三角化

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 即可。

矩阵消元的几何观点

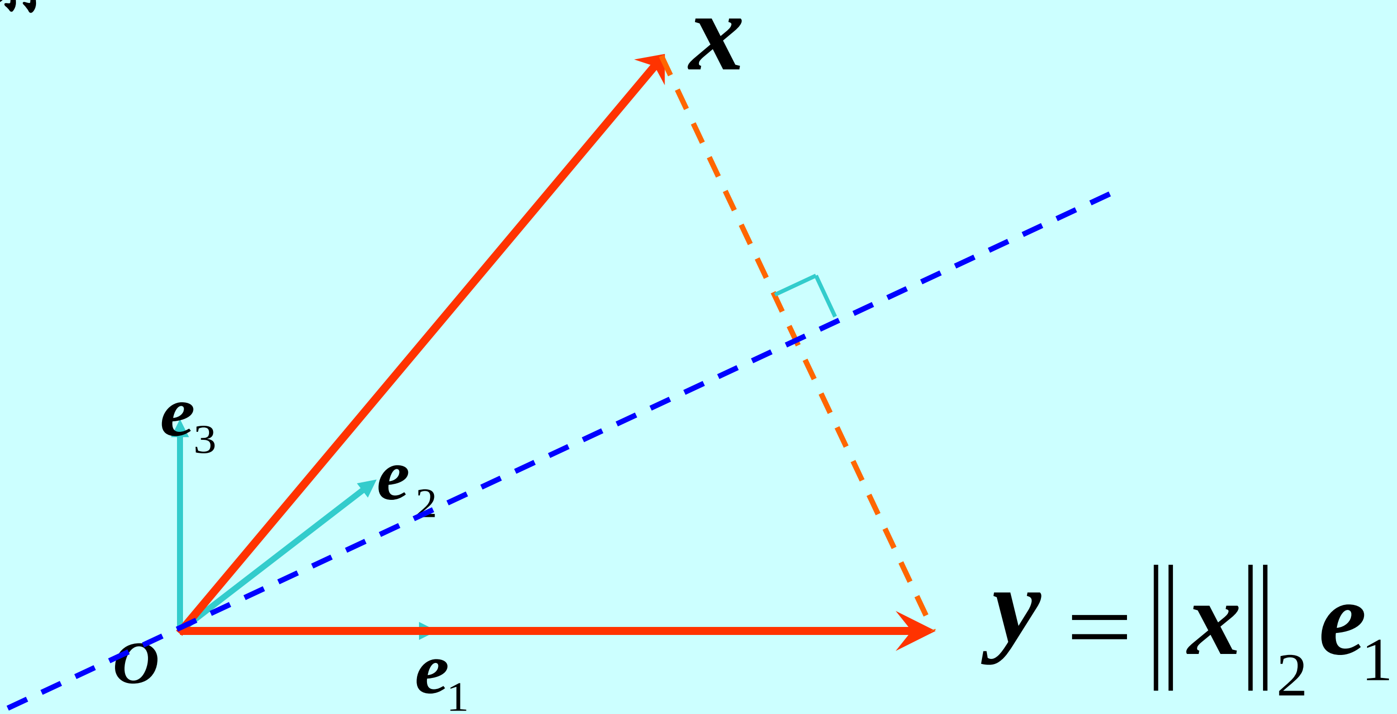
$$A=\begin{pmatrix}\times&\times&\times\\\times&\times&\times\\\times&\times&\times\end{pmatrix}\xrightarrow{\text{用正交变换 }\boldsymbol{Q}_1}\begin{pmatrix}\times&\times&\times\\0&\times&\times\\0&\times&\times\end{pmatrix}\xrightarrow{\text{用正交变换 }\boldsymbol{Q}_2}\begin{pmatrix}\times&\times&\times\\0&\times&\times\\0&0&\times\end{pmatrix}=\boldsymbol{R}$$

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

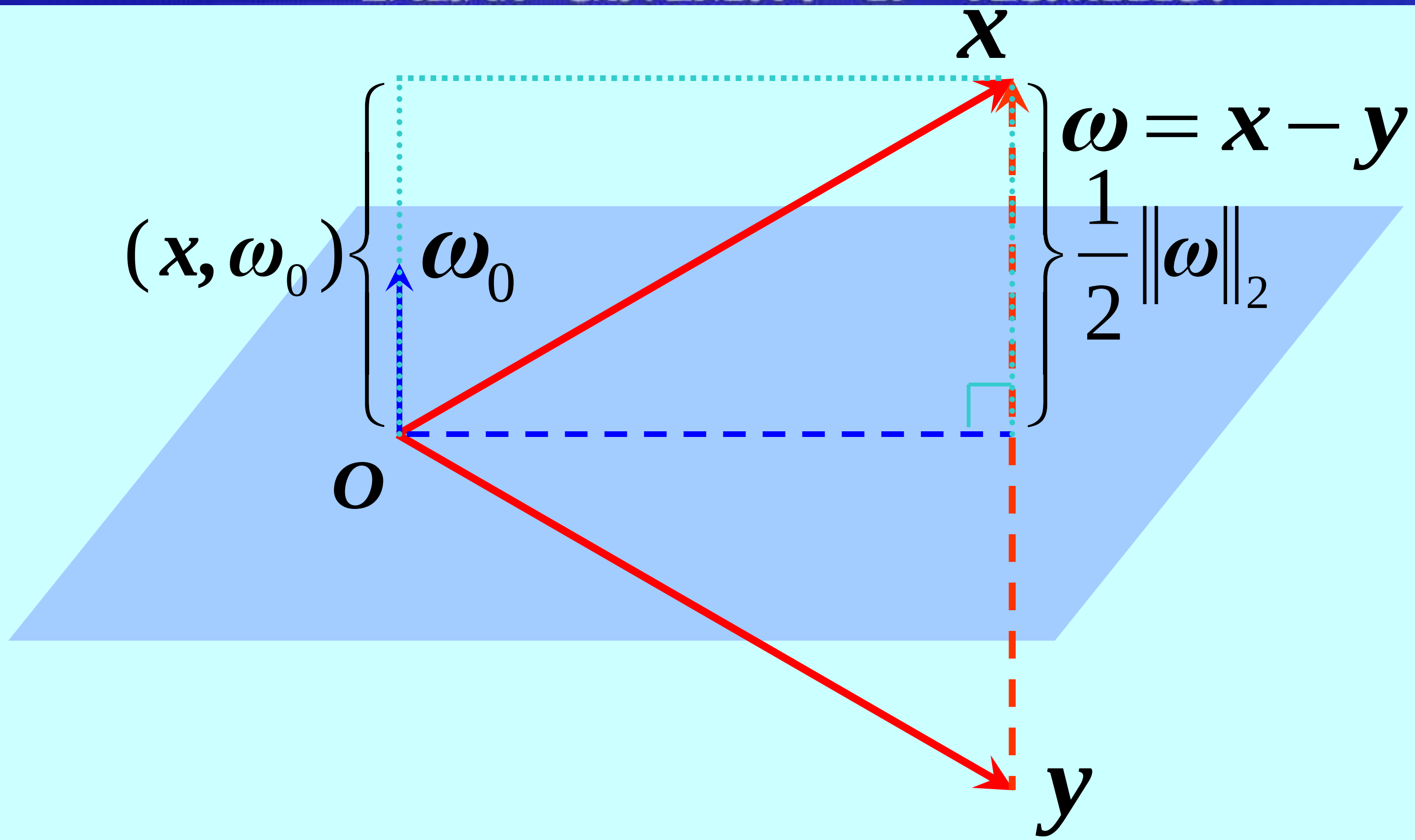
几何上看，就是把空间中的一个向量通过正交变换，变为落在第一个坐标轴上的向量。

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 x变为落在第一个坐标轴 e_1上的向量 $y=\|x\|_2 e_1$ ？



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

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

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

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

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

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

为Householder矩阵（简称H阵），或称Householder变换矩阵。单位矩阵的秩一修正。

Householder矩阵的性质

1. 对称性： $\boldsymbol{H}(\omega)^T = \boldsymbol{H}(\omega)$

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

2. 正交性： $\boldsymbol{H}(\omega)^T\boldsymbol{H}(\omega) = \boldsymbol{I}$

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

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

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

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

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

$$\boldsymbol{H}(\omega)x=\boldsymbol{H}(\boldsymbol{x}-\|\boldsymbol{x}\|_2\boldsymbol{e}_1)\boldsymbol{x}=\begin{pmatrix}\|\boldsymbol{x}\|_2\\0\\\vdots\\0\end{pmatrix}=\|\boldsymbol{x}\|_2\boldsymbol{e}_1.$$

5. $\boldsymbol{H}(\omega)$ 的特征值为 $n-1$个1和一个 -1。$|\boldsymbol{H}(\omega)|=-1$

事实上，变换后的向量可以为 $\pm\|x\|_2 e_1$

1、正负号的选取

2、正负号强制选取

3、向量x的数量级较大

例5 利用Householder变换求A的分解，其中

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

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

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

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

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

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

$$\boldsymbol{\omega}_2 = \tilde{\boldsymbol{a}}_1 - \|\tilde{\boldsymbol{a}}_1\|_2 \boldsymbol{e}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \begin{pmatrix} \sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 1-\sqrt{2} \\ -1 \end{pmatrix}$$

$$\boldsymbol{H}(\boldsymbol{\omega}_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{2-\sqrt{2}} \begin{pmatrix} 3-2\sqrt{2} & -1+\sqrt{2} \\ -1+\sqrt{2} & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & -\sqrt{2} \end{pmatrix},$$

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

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

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

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

$$
\boldsymbol{A}=\boldsymbol{QR}
$$

$$\boldsymbol{R}=\begin{pmatrix}3 & 3 & -\frac{7}{3}\\ 0 & \sqrt{2} & 3\sqrt{2}\\ 0 & 0 & -\frac{4}{3}\sqrt{2}\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}\begin{pmatrix}3 & 3 & -\frac{7}{3}\\ 0 & \sqrt{2} & 3\sqrt{2}\\ 0 & 0 & \frac{4}{3}\sqrt{2}\end{pmatrix}=\boldsymbol{Q}_3\tilde{\boldsymbol{R}},$$

$$\boldsymbol{A}=\boldsymbol{Q}\boldsymbol{R}=(\boldsymbol{Q}\boldsymbol{Q}_3)\tilde{\boldsymbol{R}}=\tilde{\boldsymbol{Q}}\tilde{\boldsymbol{R}}$$

2.2 特殊矩阵的特征系统

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

$$A = URU^H$$

酉相似

其中 $R \in C^{n\times n}$ 为上三角矩阵。

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R。

R 的对角元可以按任意给定的顺序排列，R 通常称为A的Schur标准型。

定义

设$A \in \mathbf{C}^{n\times n}$，若$A^H A = AA^H$，则称$A$为正规矩阵

复情形	实情形
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：$A^T A = AA^T = I$

正规矩阵的Schur标准型

对角矩阵

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。

A为正规矩阵, 即

$$A^H A = AA^H \Rightarrow R^H R = RR^H$$

R为正规矩阵。

上三角阵R正规矩阵 $\Leftrightarrow$ R为对角矩阵。

习题10

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

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

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

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

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

Hermite矩阵的Schur标准型

Hermite矩阵的Schur标准型为n阶对角阵

A为Hermite矩阵

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

$$\boldsymbol{UDU}^H=\left(\boldsymbol{UDU}^H\right)^H=\boldsymbol{UD}^H\boldsymbol{U}^H$$

$\boldsymbol{D}^H=\boldsymbol{D}$,D的对角元均为实数

推论

n阶方阵A为Hermite矩阵 ⟺ 存在n阶酉阵U，使得

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

其中$D\in\mathrm{R}^{n\times n}$为实对角矩阵

斜Hermite矩阵的Schur标准型

斜Hermite矩阵的Schur标准型为n阶对角阵

A为斜Hermite矩阵

$$\Downarrow$$

$$A=-A^H$$

$$\Downarrow$$

$$UDU^H=-\left(UDU^H\right)^H=-UD^HU^H$$

$$\Downarrow$$

$D=-D^H$，D的对角元为纯虚数

推论

n阶方阵A为斜Hermite矩阵

$\Longleftrightarrow$ 存在n阶酉阵U，使得

$$A=UDU^H$$

其中$D\in \mathrm{R}^{n\times n}$为纯虚对角阵

酉矩阵的Schur标准型

酉矩阵的Schur标准型为n阶对角阵

A为酉矩阵

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

$$DD^H = D^H D = I$$

$|d_i| = 1$,D的对角元模为1

推论

n阶方阵A为酉矩阵

⟺ 存在n阶酉阵U，使得

$$A = UDU^H$$

其中$D \in \mathrm{R}^{n\times n}$为对角阵

对角元模为1

定理 设 A 为 n 阶方阵，则

A 为正规矩阵 $\Longleftrightarrow$ 存在 n 阶酉阵 U，使得 $A=UDU^H$，其中 D 为对角矩阵。

A 为Hermite矩阵 $\Longleftrightarrow$ 存在 n 阶酉阵 U，使得 $A=UDU^H$ 其中 D 为实对角矩阵。

A 为斜Hermite矩阵 $\Longleftrightarrow$ 存在 n 阶酉阵 U，使得 $A=UDU^H$ 其中 D 为对角阵，且对角元为纯虚数

A 为酉阵 $\Longleftrightarrow$ 存在 n 阶酉阵 U，使得 $A=UDU^H$，其中 D 为对角阵，其对角元的模为1

矩阵的基本分类

$$
\begin{aligned}
\text{一般矩阵} \;&\supset\; \text{可对角化矩阵} \\
&\supset\; \text{正规矩阵} \\
&\supset \left\{ \begin{array}{lll} \text{Hermite阵} & \supset & \text{实对称矩阵} \\ \text{斜Hermite阵} & \supset & \text{实反对称矩阵} \\ \text{酉阵} & \supset & \text{正交矩阵} \end{array} \right\}
\end{aligned}
$$

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。

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。

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

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

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

最后，

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

习题11 证明 Schur不等式:

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

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

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

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,可知其充分必要条件是A为正规矩阵。

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

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

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　几何重数 $m_i \ge \alpha_i$

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

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

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

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

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

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

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(有完备的特征向量系) 等价于可对角化.
- 存在亏损的特征值的矩阵称为亏损矩阵等价于不可对角化.

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

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

A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

B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

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

$$J_k(\lambda)=\begin{pmatrix}\lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda\end{pmatrix}$$

$$J_4(0)=\begin{pmatrix}0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0\end{pmatrix}\qquad J_3(2)=\begin{pmatrix}2 & 1 & \\ & 2 & 1 \\ & & 2\end{pmatrix}\qquad J_2(1)=\begin{pmatrix}1 & 1 \\ & 1\end{pmatrix}$$

定义2.9 由若干个Jordan块排成的块对角矩阵为Jordan阵.

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

定理2.10 设A为n阶方阵, 则存在n阶可逆矩阵T使得

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{J}\boldsymbol{T}^{-1} \tag{2-43}$$

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

称(2-43)为A的Jordan分解, J称为A的Jordan标准型, 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 $\boldsymbol{A}=\begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}$

解 $\det(\lambda \boldsymbol{I}-\boldsymbol{A})=(\lambda+1)^3$

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

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.

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

返回

定理2.11　设A为n阶方阵，λ_i 为其特征值, 则A的Jordan标准型J中以 λ_i 为特征值, 阶数为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A的Jordan标准型J，其中

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

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.

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

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

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

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) $l=2$

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

以2为特征值, 阶数为2 的Jordan块的个数为

$$r_3+r_1-2r_2=0+2-2\times 0=2$$

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

(二) 关于变换矩阵T

$$\boldsymbol{T}=(\boldsymbol{T}_1,\boldsymbol{T}_2,\ldots,\boldsymbol{T}_k),$$

T_i 为 $n\times n_i$ 阶矩阵

$$\boldsymbol{A}=\boldsymbol{TJT}^{-1}\Rightarrow\boldsymbol{AT}=\boldsymbol{TJ}$$

$$A\,(T_1,T_2,\ldots,T_k)\;=\;(T_1,T_2,\ldots,T_k)\begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_k}(\lambda_k)\end{pmatrix}$$

$$AT_i=T_iJ_{n_i}(\lambda_i)$$

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

t_k^i 为 $n\times 1$ 阶矩阵

$$\boldsymbol{A}(\boldsymbol{t}_1^i,\boldsymbol{t}_2^i,\ldots,\boldsymbol{t}_{n_i}^i)=(\boldsymbol{t}_1^i,\boldsymbol{t}_2^i,\ldots,\boldsymbol{t}_{n_i}^i)\begin{pmatrix}\lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i\end{pmatrix}$$

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

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

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

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

$\boldsymbol{t}_1^i$是矩阵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
- 链首要求：特征向量、方程组(2-45)可解
- 选取：对应特征向量空间中所有特征向量的某种线性组合

例4 计算例2中矩阵A化Jordan标准型的变换矩阵T.

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

$$\mathbf{x}_1=(2,0,-1)^T,\ \mathbf{x}_2=(0,1,0)^T.$$

但以x_1或x_2为链首时都无法求出下一个Jordan链向量.

构造 $\mathbf{y}\in span\{\mathbf{x}_1,\mathbf{x}_2\}$, 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$\boldsymbol{T} = \begin{pmatrix} 0 & 4 & 1 \\ 1 & 3 & 0 \\ 0 & -2 & 0 \end{pmatrix}$

计算矩阵的Jordan分解

Jordan标准型J

1、计算矩阵的全部特征值

2、计算特征值的代数重数（确定对角元）

3、计算特征值的几何重数（Jordan块个数）

4、利用定理2.11确定每个k阶块的个数（为节省计算量从小到大计算）

变换矩阵T

1、求得Jordan标准型

2、计算每个Jordan块对应的Jordan链

- 若Jordan块阶数为1，直接计算特征向量
- 若阶数大于1，则先计算特征向量，利用特征向量的线性组合得到链首（保证线性方程组2.45有解）

Jordan分解的应用：

计算初等函数在某个矩阵处的值（矩阵）

最简单的情形：

多项式函数（高次多项式）

定理2.12 (Hamilton-Caylay)

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

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

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

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

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

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

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

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

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。

Jordan分解：$A = TJT^{-1}$

T可逆阵，　J若当标准型

Schur分解：$A = URU^{H}$

U酉矩阵，　R上三角矩阵

？？分解：$A = UDV^{H}$

U,V 酉矩阵，D对角矩阵

变换矩阵：

可逆 →酉矩阵

宽松 →严格

标准型：

双对角 →上三角

严格→宽松

假定矩阵 $A \in C^{m\times n}$ $\operatorname{rank}(A) = r$

$\boldsymbol{A}^H\boldsymbol{A}$ 为Hermite**半正定矩阵，特征值均为实数且非负**

$$\operatorname{rank}(A^H A) = \operatorname{rank}(AA^H) = \operatorname{rank}(A^H) = \operatorname{rank}(A)$$

由Schur定理的推论，存在 n **阶酉阵** V**，使得**

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

$$=$$

矩阵划分:

$$V=\begin{pmatrix}V_1 & V_2\end{pmatrix},\quad V_1\in C^{n\times r}, V_2\in C^{n\times(n-r)}$$

$$U=\begin{pmatrix}U_1 & U_2\end{pmatrix},\quad U_1\in C^{m\times r}, U_2\in C^{m\times(n-r)}$$

目标:寻找矩阵 U_1,U_2使得

$$\boldsymbol{U}^H\boldsymbol{AV}=\begin{pmatrix}U_1^H\\ U_2^H\end{pmatrix}A\begin{pmatrix}V_1 & V_2\end{pmatrix}=\begin{pmatrix}U_1^HAV_1 & U_1^HAV_2\\ U_2^HAV_1 & U_2^HAV_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}$$

$$U_1^H AV_1 = \boldsymbol{\Sigma} \quad U_1^H AV_2 = 0 \quad U_2^H AV_1 = 0 \quad U_2^H AV_2 = 0$$

由于

$$\boldsymbol{V}^H \boldsymbol{A^H} \boldsymbol{AV} = \begin{pmatrix} V_1^H \\ V_2^H \end{pmatrix} \boldsymbol{A^H} A \begin{pmatrix} V_1 & V_2 \end{pmatrix} = \begin{pmatrix} V_1^H \boldsymbol{A^H} AV_1 & V_1^H \boldsymbol{A^H} AV_2 \\ V_2^H \boldsymbol{A^H} AV_1 & V_2^H \boldsymbol{A^H} AV_2 \end{pmatrix}$$

$$= \begin{pmatrix} (AV_1)^H AV_1 & (AV_1)^H AV_2 \\ (AV_2)^H AV_1 & (AV_2)^H AV_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma^2} & 0 \\ 0 & 0 \end{pmatrix}$$

故

$$(AV_1)^H AV_1 = \boldsymbol{\Sigma^2} \qquad (AV_2)^H AV_2 = \boldsymbol{0}$$

从而

$U_1 = AV_1\boldsymbol{\Sigma}^{-1} \quad U_1^H U_1 = I_r \quad U_2$为$U_1$的标准正交补 $\quad U^H U = I$

定理 2.14 假定$\boldsymbol{A}\in\boldsymbol{C}^{m\times n}$，且$\operatorname{rank}(A)=r$，则存在$m$阶、$n$阶酉阵$U$、$V$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A的非零奇异值。

$\boldsymbol{A}^H\boldsymbol{A}$ 为Hermite半正定矩阵，特征值均为实数且非负

定义

设$\boldsymbol{A}\in\mathbf{C}^{m\times n}$, $k=\min(m,n)$ $\boldsymbol{A}^H\boldsymbol{A}$的特征值为

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

称非负实数

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

为矩阵A的奇异值。

矩阵A的奇异值满足性质：酉变换不改变矩阵的奇异值

定理 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，命题得证。

称（2-47）式为矩阵A的奇异值分解，亦称为矩阵A的满的奇异值分解。

$$\boldsymbol{A}=\boldsymbol{U}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\boldsymbol{V}^H=\begin{pmatrix}U_1 & U_2\end{pmatrix}\begin{pmatrix}\boldsymbol{\Sigma} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}V_1^H\\ V_2^H\end{pmatrix}=U_1\boldsymbol{\Sigma}V_1^H$$

关系式（2-47）亦可写为

$$\boldsymbol{A}=\boldsymbol{U}_1\Sigma\boldsymbol{V}_1^H$$

称为矩阵A约化的奇异值分解。

$$\boldsymbol{A}=\boldsymbol{U}_1\Sigma\boldsymbol{V}_1^H \Rightarrow \boldsymbol{A}\boldsymbol{V}_1=\boldsymbol{U}_1\Sigma,\quad \boldsymbol{U}_1^H\boldsymbol{A}=\Sigma\boldsymbol{V}_1^H$$

$$\boldsymbol{A}\boldsymbol{v}_i=\sigma_i\boldsymbol{u}_i,\quad \boldsymbol{u}_i^H\boldsymbol{A}=\sigma_i\boldsymbol{v}_i^H,\quad i=1,2,\cdots,r$$

U与V的列向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_m$和 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$分别称为矩阵$A$的与奇异值$\sigma_i$对应的左奇异向量和右奇异向量.

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

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

左奇异向量 $\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_m$ 为AA^H的单位正交特征向量

右奇异向量 $\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_n$ 为A^HA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 $A^H A$ 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

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

再求出V，由

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

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

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

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.

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

$$\boldsymbol{x}^H\boldsymbol{A}^H\boldsymbol{x}=\boldsymbol{x}^H\boldsymbol{A}\boldsymbol{x}=\boldsymbol{x}^H\lambda\boldsymbol{x}=\lambda\|\boldsymbol{x}\|_2^2 \quad \Rightarrow \lambda=\bar{\lambda}$$

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

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

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

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

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

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.

$$rank(\boldsymbol{A}^H\boldsymbol{A}) = rank(\boldsymbol{A}\boldsymbol{A}^H) = rank(\boldsymbol{A}^H) = rank(\boldsymbol{A})$$

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

其中 u_i 和 v_i 分别为矩阵U和V的正交向量;

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

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

N(A)称为A的零空间或核空间,即

$$\mathbf{N}(\boldsymbol{A}) = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$$

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

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

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

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

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.

$$\boldsymbol{A}^H \boldsymbol{A} = \boldsymbol{A}^2$$

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

定理2.19　如果A为 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

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

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

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

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

矩阵奇异值分解的几何意义

$$y = Ax = \beta_1 u_1 + \beta_2 u_2$$

$$A = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} v_1^H \\ v_2^H \end{pmatrix}$$

$$x = \alpha_1 v_1 + \alpha_2 v_2$$

$$y = Ax = \alpha_1 A v_1 + \alpha_2 A v_2 = \alpha_1 \sigma_1 u_1 + \alpha_2 \sigma_2 u_2$$

$$Av_1 = \sigma_1 u_1, Av_2 = \sigma_2 u_2$$

$$\beta_1 = \alpha_1 \sigma_1, \beta_2 = \alpha_2 \sigma_2$$

向量 x 坐标满足

$$\alpha_1^2 + \alpha_2^2 = 1$$ 单位圆

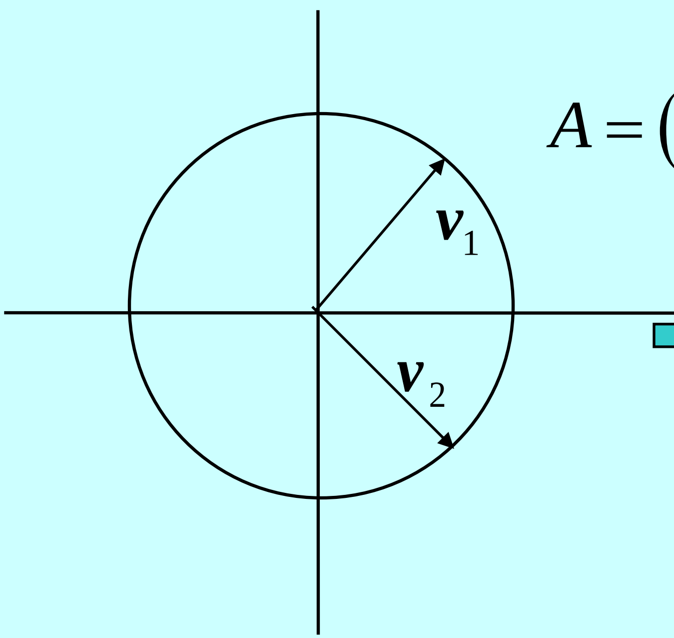
对应向量 y 坐标满足

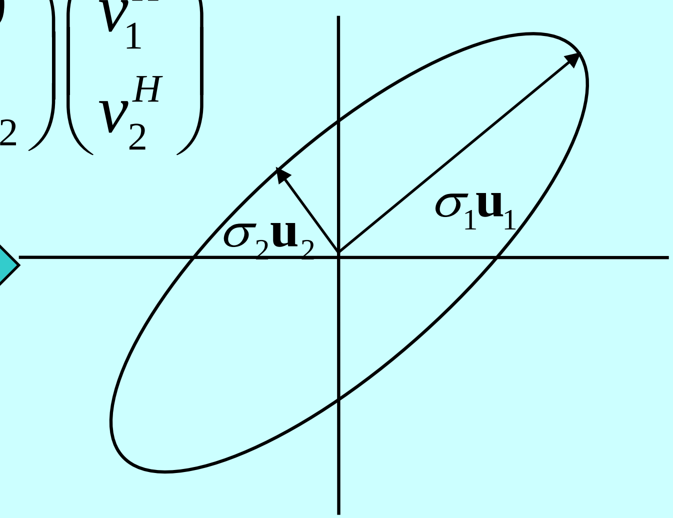
$$\left(\frac{\beta_1}{\sigma_1}\right)^2 + \left(\frac{\beta_2}{\sigma_2}\right)^2 = 1$$ 椭圆

矩阵奇异值分解的几何意义

$$x=\begin{pmatrix}x_1\\x_2\end{pmatrix}\qquad\qquad y=\begin{pmatrix}y_1\\y_2\end{pmatrix}=A\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

$$A=\begin{pmatrix}u_1 & u_2\end{pmatrix}\begin{pmatrix}\sigma_1 & 0\\0 & \sigma_2\end{pmatrix}\begin{pmatrix}v_1^H\\v_2^H\end{pmatrix}$$





高维情形

向量 $y = Ax$

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n,\ \ y = \beta_1 u_1 + \cdots + \beta_n u_n$$

仍有 $Av_i = \sigma_i u_i$ $\qquad \beta_1 = \alpha_1\sigma_1, \cdots, \beta_n = \alpha_n\sigma_n$

单位球

$$\alpha_1^2 + \cdots + \alpha_n^2 = 1$$

超椭球：将单位球沿某些正交方向分别以拉伸因子 σ_i 拉伸而成的曲面, u_i 为主半轴, σ_i 为主半轴的长度, 恰好是矩阵的奇异值.

$$\left(\frac{\beta_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{\beta_n}{\sigma_n}\right)^2 = 1$$