第六章

矩阵分解

6.1 矩阵的LU分解

许多分解源自十九世纪对二次型和双线性型的研究。**LU**分解源自**Gauss**处理表示最小二乘问题的对称正定系统时所使用的消元法，确切地说，**Gauss**使用的是分解 $A = LDL^T$ 。对于双线性型，则归功于**Jacobi**（**1857**）。**Dwyer**（**1944**）最早注意到消元法与其矩阵表示间的联系。

一、从Gauss消元法说起

例 1 求解线性方程组

$$\begin{cases} x_1 + 2x_2 - x_3 & = & 0 & ① \\ 3x_1 + x_2 & = & -1 & ② \\ -x_1 - x_2 - 2x_3 & = & 1 & ③ \end{cases} \qquad (I)$$

解:

$$(I) \xrightarrow[1\times①+③]{(-3)\times①+②} \begin{cases} x_1 + 2x_2 - x_3 & = & 0 & ① \\ -5x_2 + 3x_3 & = & -1 & ② \\ x_2 - 3x_3 & = & 1 & ③ \end{cases}$$

$$\xrightarrow{\frac{1}{5}\times②+③}\begin{cases} x_1+2x_2-x_3 = 0 & ① \\ -5x_2+3x_3 = -1 & ② \\ -\frac{12}{5}x_3 = \frac{4}{5} & ③ \end{cases}$$

$$\xrightarrow[\frac{3}{5}\times③+②]{(-\frac{5}{12})\times③}\begin{cases} x_1+2x_2 = -1/3 & ① \\ x_2 = 0 & ② \\ x_3 = -1/3 & ③ \end{cases}$$

$$\xrightarrow{(-2)\times②+①}\begin{cases} x_1 = -1/3 & ① \\ x_2 = 0 & ② \\ x_3 = -1/3 & ③ \end{cases} \qquad (II)$$

用矩阵形式表示，系数矩阵

$$A=\begin{bmatrix}1 & 2 & -1\\ 3 & 1 & 0\\ -1 & -1 & -2\end{bmatrix}\xrightarrow[r_{13}(1)]{r_{12}(-3)}\begin{bmatrix}1 & 2 & -1\\ 0 & -5 & 3\\ 0 & 1 & -3\end{bmatrix}$$

$$\xrightarrow{r_{23}(\frac{1}{5})}\begin{bmatrix}1 & 2 & -1\\ 0 & -5 & 3\\ 0 & 0 & -12/5\end{bmatrix}\equiv U$$

$$R_{23}(\tfrac{1}{5})R_{13}(1)R_{12}(-3)A = U.$$

$$A = [R_{23}(\tfrac{1}{5})R_{13}(1)R_{12}(-3)]^{-1}U$$

$$= [R_{12}(-3)]^{-1}[R_{13}(1)]^{-1}[R_{23}(\tfrac{1}{5})]^{-1}U$$

$$= R_{12}(3)R_{13}(-1)R_{23}(\tfrac{-1}{5})U \quad \boxed{\equiv LU}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & -1/5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \\ 0 & 0 & -12/5 \end{bmatrix}$$

这就是Gauss提出消元法100多年后才被Dwyer注意到的 LU 分解：

$$\boxed{A = LU}$$

据此，有

$$\boxed{Ax = LUx = L(Ux) = b}$$

因此可通过求解两个特殊的三角方程组

$$\boxed{Ly = b, Ux = y}$$

来求解线性方程组 $\boxed{Ax = b}$ ，这就是数值软件中采用的方法。

二、矩阵的LU分解

定义6.1.1 如果方阵 A 可以分解成一个下三角矩阵 L 与一个上三角矩阵 U 的乘积

$$A = LU$$

则称其为 A 的 LU 分解或三角分解。若L为单位下三角阵，则称其为矩阵A的Doolittle分解。若U为单位上三角阵，则称其为A的Crout分解。

什么样的矩阵才有LU 分解呢？

定理6.1.1（LU定理）设A是n阶非奇异矩阵，则存在唯一的单位下三角矩阵L与非奇异上三角矩阵U，使得

$$A = LU$$

当且仅当A的所有顺序主子式均非零，即

$$\Delta_k \neq 0 \quad (k = 1, 2, \cdots, n)$$

必要性：

设有 $A=LU$，将 $A=LU$ 分块为

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} L_{11} & O \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ O & U_{22} \end{pmatrix}$$

$$= \begin{pmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12}+L_{22}U_{22} \end{pmatrix}$$

因此

$$
\begin{aligned}
\det A_{11} &= \det(L_{11}U_{11}) \\
&= \det L_{11} \det U_{11} \\
&= 1 \cdot \det U_{11} = \det U_{11} \neq 0
\end{aligned}
$$

考虑到分块矩阵 A_{11} 阶数的任意性，因此上述结论对矩阵 A 的任意顺序主子式都成立。

充分性：利用数学归纳法。当n=1时，结论显然成立，$a_{11}=1\cdot a_{11}$. 设对n-1阶矩阵，定理结论成立，下证对n阶矩阵结论成立. 记

$$A=\begin{pmatrix} A_{n-1} & \beta \\ \alpha & a_{nn} \end{pmatrix}$$

其中 A_{n-1} 为 A 的n-1阶顺序主子式，根据定理条件，A_{n-1} 为非奇异矩阵，则有

$$\begin{pmatrix} I_{n-1} & 0 \\ -\alpha A_{n-1}^{-1} & 1 \end{pmatrix} A$$

$$= \begin{pmatrix} I_{n-1} & 0 \\ -\alpha A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \beta \\ \alpha & a_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} A_{n-1} & \beta \\ 0 & a_{nn} - \alpha A_{n-1}^{-1} \beta \end{pmatrix}$$

从而

$$A = \begin{pmatrix} I_{n-1} & 0 \\ \alpha A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \beta \\ 0 & a_{nn} - \alpha A_{n-1}^{-1} \beta \end{pmatrix}$$

由归纳假设，存在n-1阶单位下三角矩阵 L_{n-1} 和上三角矩阵 U_{n-1} 使得 $A_{n-1}=L_{n-1}U_{n-1}$，于是

$$A=\begin{pmatrix} I_{n-1} & 0 \\ \alpha A_{n-1}^{-1} & 1 \end{pmatrix}\begin{pmatrix} A_{n-1} & \beta \\ 0 & a_{nn}-\alpha A_{n-1}^{-1}\beta \end{pmatrix}$$

$$=\begin{pmatrix} I_{n-1} & 0 \\ \alpha A_{n-1}^{-1} & 1 \end{pmatrix}\begin{pmatrix} L_{n-1} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} U_{n-1} & L_{n-1}^{-1}\beta \\ 0 & a_{nn}-\alpha A_{n-1}^{-1}\beta \end{pmatrix}$$

$$=LU$$

$$L=\begin{pmatrix} I_{n-1} & 0 \\ \alpha A_{n-1}^{-1} & 1 \end{pmatrix}\begin{pmatrix} L_{n-1} & 0 \\ 0 & 1 \end{pmatrix},$$

$$U=\begin{pmatrix} U_{n-1} & L_{n-1}^{-1}\beta \\ 0 & a_{nn}-\alpha A_{n-1}^{-1}\beta \end{pmatrix}$$

即得 $A=LU$，其中L为单位下三角矩阵，U是上三角矩阵。

唯一性：

设有 $$A = LU = \tilde{L}\tilde{U}$$

其中，L，$\tilde{L}$ 为n阶单位下三角矩阵，U,$\tilde{U}$ 为n阶可逆上三角矩阵，则

$$\tilde{L}^{-1}L = \tilde{U}U^{-1}$$

上式左边的矩阵为单位下三角矩阵，右边的矩阵为上三角矩阵，因此，

$$\tilde{L}^{-1}L = \tilde{U}U^{-1} = I$$

于是 $L = \tilde{L}, U = \tilde{U}$.

根据 LU 定理，如果存在分解 $A=LU$，那么矩阵 L 可逆，即存在可逆矩阵 P，使得

$$PL=I$$

从而

$$PA=PLU=U$$

$$P(A,I)=(PA,P)=(U,P)$$

这说明，通过行初等变换求出 U 和 P 后，就可求出下三角阵 P 的逆矩阵 L。

注： 进行行初等变换不能进行行的交换。

例 1 求下列矩阵的Doolittle分解：

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

解:

$$(A,I)=\begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 3 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & -2 & 0 & 0 & 1 \end{bmatrix}$$

$$\Longrightarrow \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -5 & 3 & -3 & 1 & 0 \\ 0 & 1 & -3 & 1 & 0 & 1 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -5 & 3 & -3 & 1 & 0 \\ 0 & 0 & -12/5 & 2/5 & 1/5 & 1 \end{bmatrix}$$

从而得 $PA = U$，这里

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2/5 & 1/5 & 1 \end{bmatrix}, U = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \\ 0 & 0 & -12/5 \end{bmatrix}$$

因为

$$L = P^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & -1/5 & 1 \end{bmatrix}$$

所以 $A = LU$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & -1/5 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \\ 0 & 0 & -12/5 \end{bmatrix}$$

定理（LDU分解定理）设 A 为n阶非奇异矩阵，则存在唯一的单位下三角矩阵 L 与单位上三角矩阵 U，以及对角矩阵 D，使得

$$A = LDU$$

当且仅当 A 的所有阶顺序主子式均非零，即

$$\Delta_k \neq 0 \quad (k = 1, 2, \cdots, n)$$

定理6.1.3（Cholesky分解） 设 A 为n阶对称正定阵，则存在唯一的单位下三角矩阵 L 与对角矩阵 D，使得

$$\boxed{A = LDL^T}$$

定理6.1.4（Cholesky分解） 设 A 为n阶对称正定阵，则存在唯一的对角元全是正数的矩阵 G，使得

$$\boxed{A = GG^T}$$

6.2 矩阵的QR分解

QR分解在矩阵计算中占据相当重要的地位。利用QR分解，可以解决各种应用中（例如工程力学、流体力学、图像压缩处理、结构分析等）出现的最小二乘问题、特征值问题等矩阵计算中的核心问题。尤其是基于QR分解的QR算法，是求解小型稠密矩阵特征值问题的最主要的、数值稳定的算法。

一、再谈Gram-Schmidt方法

可逆阵 A 的列向量组 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关。

可令

$$\beta_1=\alpha_1$$

$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1$$

$$\vdots$$

$$\beta_n=\alpha_n-\frac{(\alpha_n,\beta_1)}{(\beta_1,\beta_1)}\beta_1-\frac{(\alpha_n,\beta_2)}{(\beta_2,\beta_2)}\beta_2-\cdots-\frac{(\alpha_n,\beta_{n-1})}{(\beta_{n-1},\beta_{n-1})}\beta_{n-1}$$

即

$$\alpha_1 = \beta_1$$

$$\alpha_2 = \frac{(\alpha_2, \beta_1)}{(\beta_1, \beta_1)} \beta_1 + \beta_2$$

$$\vdots$$

$$\alpha_n = \frac{(\alpha_n, \beta_1)}{(\beta_1, \beta_1)} \beta_1 + \frac{(\alpha_n, \beta_2)}{(\beta_2, \beta_2)} \beta_2 + \cdots + \frac{(\alpha_n, \beta_{n-1})}{(\beta_{n-1}, \beta_{n-1})} \beta_{n-1} + \beta_n$$

再由

$$\gamma_i = \frac{\beta_i}{\|\beta_i\|}$$ 可得 $$\beta_i = \|\beta_i\| \gamma_i$$

代入，有

$$\alpha_1 = \|\beta_1\|\gamma_1$$

$$\alpha_2 = (\alpha_2,\gamma_1)\gamma_1 + \|\beta_2\|\gamma_2$$

$$\vdots$$

$$\alpha_n = (\alpha_n,\gamma_1)\gamma_1 + (\alpha_n,\gamma_2)\gamma_2 + \cdots + (\alpha_n,\gamma_{n-1})\gamma_{n-1} + \|\beta_n\|\gamma_n$$

故

$$(\alpha_1,\alpha_2,\cdots,\alpha_n) = (\gamma_1,\gamma_2,\cdots,\gamma_n)\begin{pmatrix} \|\beta_1\| & (\alpha_2,\gamma_1) & \cdots & (\alpha_n,\gamma_1) \\ & \|\beta_2\| & \cdots & (\alpha_n,\gamma_2) \\ & & \ddots & \vdots \\ & & & \|\beta_n\| \end{pmatrix}$$

因此

$$\boxed{A = QR.}$$

QR
分解

其中

$$Q \equiv [\gamma_1, \gamma_2, \cdots, \gamma_n], \quad Q^H Q = E_n$$

$$R = \begin{pmatrix} \|\beta_1\| & (\alpha_2, \gamma_1) & \cdots & (\alpha_n, \gamma_1) \\ & \|\beta_2\| & \cdots & (\alpha_n, \gamma_2) \\ & & \ddots & \vdots \\ & & & \|\beta_n\| \end{pmatrix}$$

是上三角阵。

使用矩阵语言，就是QR分解：

$$\boxed{A = QR}$$

这里，Q 是酉矩阵或正交矩阵，R 是上三角阵

此时线性方程组

$$\boxed{Ax = b}$$

变成

$$\boxed{QRx = b}$$

即三角方程组

$$\boxed{Rx = Q^{-1}b = Q^{H}b}$$

二、矩阵的QR分解

定理6.2.1（QR分解） 设 A 为任一 n 阶可逆矩阵，则必存在 n 阶酉矩阵 Q 和 n 阶上三角方阵 R，使得

$$\boxed{A = QR}$$

例 1 将下列矩阵进行QR分解：

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

解: $\alpha_1 = (1,1,0,)^T$, $\beta_1 = \alpha_1 = (1,1,0)^T$

$$\gamma_1 = \frac{\beta_1}{\|\beta_1\|} = \frac{1}{\sqrt{2}}(1,1,0)^T$$

$$\beta_2 = \alpha_2 - (\alpha_2, \gamma_1)\gamma_1 = (1,-1,1)^T$$

$$\gamma_2 = \frac{\beta_2}{\|\beta_2\|} = \frac{1}{\sqrt{3}}(1,-1,1)^T$$

$$\beta_3 = \alpha_3 - (\alpha_3, \gamma_1)\gamma_1 - (\alpha_3, \gamma_2)\gamma_2 = \frac{1}{3}(-1,1,2)^T$$

$$\gamma_3 = \frac{\beta_3}{\|\beta_3\|} = \frac{1}{\sqrt{6}}(-1,1,2)^T$$

所以 A 的QR分解为：

$$Q=(\gamma_1,\gamma_2,\gamma_3)=\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix}$$

$$R = Q^T A = \begin{pmatrix} \sqrt{2} & \sqrt{2} & 2\sqrt{2} \\ 0 & \sqrt{3} & \dfrac{1}{\sqrt{3}} \\ 0 & 0 & \dfrac{\sqrt{6}}{3} \end{pmatrix}$$

例 2 利用QR分解解线性方程组 $Ax = b$，其中

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 1 & 2 \\ 2 & -5 & 8 \\ -1 & 3 & -7 \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 1 \\ -11 \\ 9 \end{pmatrix}$$

定理（**长方形矩阵QR分解**）设 $A \in C^{m\times n}$，$m \geq n$ 且 $rank(A)=n$（列满秩），则必存在 $m\times n$ 阶（列正交）矩阵 Q $(Q^HQ=E_n)$ 和非奇异上三角方阵 R，使得

$$\boxed{A = QR}$$

当 $m=n$ 时，Q 就是酉矩阵或正交矩阵。

6.3 矩阵的满秩分解

本节讨论将一个非零矩阵（长方形）分解成一个列满秩矩阵与一个行满秩矩阵的乘积问题.

主要内容:

矩阵的Hermite标准形

利用Hermite标准形进行矩阵的满秩分解

一、Hermite标准形

为了说明矩阵满秩分解定理以及满秩分解方法，先介绍Hermite 标准形（或行最简形）。

定义 $A \in C_r^{m\times n}$ 的Hermite的标准形为 H

1）它的非零行出现在前r行，且每一个非零行第一个非零元为1，而后$m-r$行全为0；

2）每一个非零行的第一个非零元出现的位置必须在前一行的第一个非零元出现位置的后面。

3）每一个非零行的第一个非零元所在的列的其它元素为0。

矩阵论 MATRIX THEORY

$$H=\begin{array}{c}\begin{array}{ccccccccccccccc} & & & j_1 & & & & j_2 & & & & j_r & & & \end{array}\\ \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}\end{array}\begin{array}{l} \\ \\ \\ \\ \text{第}r\text{行} \\ \\ \\ \end{array}$$

任何一个非零矩阵都可通过初等行变换化为Hermite标准形H，且H的前r行线性无关。采用矩阵的说法就是，存在$P \in C_m^{m\times m}$，使得

$$PA = H.$$

二、满秩分解

定理6.3.1 设 $A \in C^{m\times n}$，$rank(A) = r > 0$,

则必存在秩为 r 的两个矩阵

$$B \in C^{m\times r},\ C \in C^{r\times n},$$

使得

$$A = BC \qquad (1)$$

（1）式称为矩阵A的满秩分解.

证明： 假设矩阵 A 的前 r 个列向量是线性无关的，对矩阵A只实施初等行变换可以将其化成

$$\begin{bmatrix} E_r & D \\ 0 & 0 \end{bmatrix}$$

即存在 $P \in C_m^{m\times m}$ 使得

$$PA = \begin{bmatrix} E_r & D \\ 0 & 0 \end{bmatrix}$$

于是有

$$A = P^{-1}\begin{bmatrix} E_r \\ 0 \end{bmatrix}\begin{bmatrix} E_r & D \end{bmatrix} = BC$$

其中

$$B = P^{-1}\begin{bmatrix} E_r \\ 0 \end{bmatrix} \in C_r^{m\times r}, C = \begin{bmatrix} E_r & D \end{bmatrix} \in C_r^{r\times n}$$

如果A的前r列线性相关，那么只需对A作初等列变换使得前r个列是线性无关的。然后重复上面的过程即可。这样存在

$$P \in C_m^{m\times m}, \quad Q \in C_n^{n\times n}$$

且满足

$$PAQ = \begin{bmatrix} E_r & D \\ 0 & 0 \end{bmatrix}$$

从而

$$A = P^{-1}\begin{bmatrix} E_r & D \\ 0 & 0 \end{bmatrix} Q^{-1}$$

$$= P^{-1}\begin{bmatrix} E_r \\ 0 \end{bmatrix}\begin{bmatrix} E_r & D \end{bmatrix} Q^{-1}$$

$$= BC$$

定理 （**满秩分解定理**）设 $A\in C_r^{m\times n}(r>0)$,

且A的Hermite 标准形H为

$$H=\begin{array}{cc}
\begin{array}{ccccccccccccccc} & & & j_1 & & & & j_2 & & & & j_r & & & \end{array} & \\
\begin{pmatrix}
0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * & 0 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * & 0 & * & \cdots & * \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0
\end{pmatrix} & \begin{array}{l} \\ \\ \\ \text{第}r\text{ 行} \\ \\ \\ \\ \end{array}
\end{array}$$

则取A的第 $j_1, j_2, \cdots, j_r$ 列构成矩阵B，取H的前r行构成矩阵C，则A=BC即为矩阵A的**满秩分解**。

证明： 由于A经过初等行变换化成H，则存在m阶可逆矩阵P，使得$PA=H$，将A,H按列分块

$$A=(\alpha_1,\alpha_2,\cdots,\alpha_n)$$

$$H=P(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\widetilde{\alpha_1},\widetilde{\alpha_2},\cdots,\widetilde{\alpha_n}).$$

由行最简形矩阵的定义不难看出

$$P\alpha_{j_i}=\widetilde{\alpha_{j_i}}=e_{j_i}=(0,\cdots,0,\overset{j_i}{1},0,\cdots,0)^T\in C^n$$

其中，$i=1,2,\cdots,r;\ \widetilde{\alpha_{j_1}},\widetilde{\alpha_{j_2}},\cdots,\widetilde{\alpha_{jr}}$ 是H的列向量组的极大线性无关组。

将C按列分块

$$C=(\beta_1,\beta_2,\cdots,\beta_n),\ \beta_k\in C^r;k=1,2,\cdots,n.$$

易知，$C\in C^{r\times n}$，且rank$C$=r，且 $\beta_1,\beta_2,\cdots,\beta_n$

是向量 $\widetilde{\alpha_1},\widetilde{\alpha_2},\cdots,\widetilde{\alpha_n}$ 分别由极大线性无关组

$\widetilde{\alpha_{j_1}},\widetilde{\alpha_{j_2}},\cdots,\widetilde{\alpha_{jr}}$ 线性表示的向量，即

$$\left(\widetilde{\alpha_1},\widetilde{\alpha_2},\cdots,\widetilde{\alpha_n}\right)=\left(\widetilde{\alpha_{j_1}},\widetilde{\alpha_{j_2}},\cdots,\widetilde{\alpha_{jr}}\right)\left(\beta_1,\beta_2,\cdots,\beta_n\right)$$

令 $B=\left(\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{jr}\right)\in C^{m\times r}$ ，则有

$$
\begin{aligned}
A &= P^{-1}\left(\widetilde{\alpha_1}, \widetilde{\alpha_2}, \cdots, \widetilde{\alpha_n}\right) \\
&= P^{-1}\left(\widetilde{\alpha_{j_1}}, \widetilde{\alpha_{j_2}}, \cdots, \widetilde{\alpha_{jr}}\right)\left(\beta_1, \beta_2, \cdots, \beta_n\right) \\
&= \left(\alpha_{j_1}, \alpha_{j_2}, \cdots, \alpha_{jr}\right)\left(\beta_1, \beta_2, \cdots, \beta_n\right) \\
&= BC
\end{aligned}
$$

满秩分解的步骤

1) 求矩阵A的Hermite 标准形H;

2) 取矩阵C为H的前r个非0行;

3) 取矩阵B为A的对应于H的r个单位向量的列;

则A=BC

例1 求矩阵 $A=\begin{pmatrix} 2 & 4 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ -1 & -2 & -2 & 1 \end{pmatrix}$ 的满秩分解

首先利用行初等变换求A的Hermite 标准形H：

$$A=\begin{pmatrix} 2 & 4 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ -1 & -2 & -2 & 1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 2 & -1 & 2 \\ 2 & 4 & 1 & 1 \\ -1 & -2 & -2 & 1 \end{pmatrix} \xrightarrow[r_3+r_1]{r_2-2r_1}$$

$$\begin{pmatrix} 1 & 2 & -1 & 2 \\ 0 & 0 & 3 & -3 \\ 0 & 0 & -3 & 3 \end{pmatrix} \xrightarrow[r_2\times\frac{1}{3}]{r_3+r_2} \begin{pmatrix} 1 & 2 & -1 & 2 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = H$$

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

可见 $j_1 = 1, \quad j_2 = 3,$ 故A的满秩分解为

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

练习 求矩阵 $A=\begin{pmatrix}1 & 0 & 1\\ 2 & 1 & 3\\ 0 & 1 & 1\end{pmatrix}$ 的满秩分解

解答: $$A=\begin{pmatrix}1 & 0\\ 2 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}=BC$$

例2 **求矩阵** $A=\begin{pmatrix}0&0&1&2&3\\0&0&2&4&6\end{pmatrix}$ **的满秩分解**

首先利用行初等变换求A的Hermite 标准形H:

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

$$\xrightarrow{r_2-2r_1}\begin{pmatrix}0&0&1&2&3\\0&0&0&0&0\end{pmatrix}$$

选取

$$B=\begin{bmatrix}1\\2\end{bmatrix},\quad C=\begin{bmatrix}0 & 0 & 1 & 2 & 3\end{bmatrix}$$

事实上，也可以选取

$$B=\begin{bmatrix}2\\4\end{bmatrix}, C=\begin{bmatrix}0 & 0 & 1/2 & 1 & 3/2\end{bmatrix}$$

注1 矩阵A的满秩分解是不唯一的。

设 $A = BC, B \in C_r^{m\times r}, C \in C_r^{r\times n}$

则 $\forall D \in C_r^{r\times r}$,

$$A = (BD)(D^{-1}C) = B_1 C_1$$

注2 矩阵A的满秩分解虽然不唯一的，但对不同的分解：A=BC，乘积 $C^H(CC^H)^{-1}(B^H B)^{-1}B^H$ 保持不变。

6.4 矩阵的奇异值分解

从**Beltrami**（**1873**）和**Jordan**（**1874**）提出奇异值分解（**SVD**）至今，**SVD**及其推广已经成为矩阵计算最有用和最有效的工具之一，在最小二乘问题、最优化、统计分析、信号与图像处理、系统理论与控制等领域被广泛使用。

设矩阵 $A\in C^{m\times n}$ ，有以下结论成立。

(1) AA^H和$A^H A$都是半正定矩阵，

(2) AA^H和$A^H A$有相同的非零特征值，

(3) $r(A)=r(AA^H)=r(A^H A), r(A)$表示A的秩。

定义6.4.1 设$A \in C_r^{m\times n}$，$A^H A$ 的特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n$$

称$d_i = \sqrt{\lambda_i}$为矩阵A的奇异值，

称$d_1 \geq \ldots \geq d_r > 0$为正奇异值。

例1 求$A=\begin{bmatrix}1 & 2\\ 0 & 0\\ 0 & 0\end{bmatrix}$的奇异值。

解:

$$A^T A=\begin{bmatrix}1 & 0 & 0\\ 2 & 0 & 0\end{bmatrix}\begin{bmatrix}1 & 2\\ 0 & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}1 & 2\\ 2 & 4\end{bmatrix}$$

特征值为5, 0, 所以奇异值为$\sqrt{5},0$

例2 求$A=\begin{bmatrix}0 & -1 & 1\\ 2 & 0 & 0\end{bmatrix}$的奇异值。

解 $$AA^T=\begin{bmatrix}0 & -1 & 1\\ 2 & 0 & 0\end{bmatrix}\begin{bmatrix}0 & 2\\ -1 & 0\\ 1 & 0\end{bmatrix}=\begin{bmatrix}2 & 0\\ 0 & 4\end{bmatrix}$$

所以，A的奇异值为$2,\sqrt{2},0$

定义6.4.2 设$A, B \in C^{m\times n}$，如果存在 m 阶酉矩阵 U 和 n 阶酉矩阵V，使得

$$U^H AV = B,$$

称A与B的酉等价.

定理6.4.1 酉等价矩阵具有相同的奇异值.

$$U^H AV = B \Rightarrow B^H B = V^H (A^H A)V$$

定理6.4.2 设 $A \in C_r^{m\times n}$，那么存在酉矩阵

$U \in C^{m\times m}$ 和酉矩阵 $V \in C^{n\times n}$ 使得

$$A = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}V^H$$

奇异值分解

其中，$\Sigma = diag(\sigma_1, \sigma_2, \cdots \sigma_r)$

$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$为矩阵$A$的正奇异值。

矩阵论 MATRIX THEORY

证明 设半正定Hermite矩阵$A^H A$的特征值为

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n,$$

对应的正交单位特征向量为

$$\gamma_1, \cdots, \gamma_r, \gamma_{r+1}, \cdots, \gamma_n.$$

令

$$\Sigma = diag(\sigma_1, \cdots, \sigma_2, \cdots \sigma_r), \quad \sigma_i = \sqrt{\lambda_i}, i = 1, \cdots, r,$$

$$V = (V_1, V_2), \quad V_1 = (\gamma_1, \cdots, \gamma_r), \ V_2 = (\gamma_{r+1}, \cdots, \gamma_n)$$

则 V 为酉矩阵，且

$$V^H A^H A V = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix} \qquad (1)$$

即

$$\begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix} A^H A \begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} V_1^H A^H A V_1 & V_1^H A^H A V_2 \\ V_2^H A^H A V_1 & V_2^H A^H A V_2 \end{bmatrix} = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$$

比较上面等式，有

$$V_1^H A^H A V_1 = (AV_1)^H A V_1 = \Sigma^2$$

$$V_2^H A^H A V_2 = (AV_2)^H A V_2 = 0$$

从而

$$\Sigma^{-1} V_1^H A^H A V_1 \Sigma^{-1} = I_r$$

$$A V_2 = 0$$

令 $U_1 = AV_1\Sigma^{-1} \in C^{m\times r}$，则有

$$U_1^H U_1 = (AV_1\Sigma^{-1})^H AV_1\Sigma^{-1} = I_r$$

即 U_1 的 r 个列是两两正交的单位向量。

另外，有 $U_1^H AV_1 = (AV_1\Sigma^{-1})^H AV_1 = \Sigma$

选取 $U_2 \in C^{m\times(m-r)}$使得 $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$是酉

矩阵，即有

即有

$$U_2^H U_2 = I_{m-r},$$

$$U_2^H U_1 = U_2^H A V_1 \Sigma^{-1} = 0$$

$$\Rightarrow \quad U_2^H A V_1 = 0$$

由上述式子可得

$$U^H AV = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} A \begin{bmatrix} V_1 & V_2 \end{bmatrix}$$

$$= \begin{bmatrix} U_1^H AV_1 & U_1^H AV_2 \\ U_2^H AV_1 & U_2^H AV_2 \end{bmatrix} = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}$$

这里，要注意 $AV_2 = 0$ 。

计算矩阵A 的奇异值分解的方法：

(1)计算$A^H A$的特征值

$$\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n;$$

从而确定A的非零奇异值$\sigma_i \ (i = 1,2,\cdots,r)$

(2)求出$A^H A$对应的正交单位特征向量

$$\gamma_1,\cdots,\gamma_r,\gamma_{r+1},\cdots,\gamma_n.$$

令$V = (V_1,V_2)$, $V_1 = (\gamma_1,\cdots,\gamma_r)$, $V_2 = (\gamma_{r+1},\cdots,\gamma_n)$

(3) 计算$U_1 = AV_1\Sigma^{-1}$，将其扩充为C^m的一组标准正交基得U_2.

(4) $U=(U_1, U_2)$, $A = U\begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^H$为$A$的奇异值分解

例 3　求下列矩阵的SVD分解：

$$(1)\ A=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix};\qquad (2)\ A=\begin{pmatrix}1&0\\0&1\\1&0\end{pmatrix}.$$

解： （1）

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

矩阵 $A^T A$ 的特征值为 $3,1,0$ ，对应的特征向量为 $\beta_1 = (1,1,2)^T, \beta_2 = (1,-1,0)^T, \beta_3 = (1,1,-1)^T$

从而

$$V = \left(\begin{array}{cc:c} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \end{array}\right) \equiv (V_1, V_2), \quad \Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{pmatrix}$$

计算

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

选取

$$U_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

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

因此所求SVD为

$$A=U\begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}V^T$$

解： （2）

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

矩阵 A^TA 的特征值为 $2,1$ ，对应的特征向量为 $\beta_1=(1,0)^T,\beta_2=(0,1)^T$

从而

$$V=\begin{pmatrix}1&0\\0&1\end{pmatrix}\equiv V_1\quad,\Sigma=\begin{pmatrix}\sqrt{2}&0\\0&1\end{pmatrix}$$

计算

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

选取

$$U_2 = \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix},$$

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

因此所求SVD为

$$A=U\left(\begin{array}{cc}\sqrt{2} & 0\\ 0 & 1\\ \hline 0 & 0\end{array}\right)V^T$$