第十一讲 满秩分解与奇异值分解

一、矩阵的满秩分解

1. 定义： 设 $A \in C_r^{m\times n}(r>0)$，若存在矩阵 $F \in C_r^{m\times r}$ 及 $G \in C_r^{r\times n}$，使得 $A = FG$，则称其为 A 的一个满秩分解。

说明：（1）F 为列满秩矩阵，即列数等于秩；G 为行满秩矩阵，即行数等于秩。

（2）满秩分解不唯一。$\forall D \in C_r^{r\times r}$（$r$ 阶可逆方阵），则

$$A = FG = F(DD^{-1})G = (FD)(D^{-1}G) = F_1G_1, \quad 且\ F_1 \in C_r^{m\times r}, G_1 \in C_r^{r\times n}$$

2. 存在性定理： 任何非零矩阵均存在满秩分解。

证明：采用构造性证明方法。设 $A \in C_r^{m\times n}$，则存在初等变换矩阵 $E \in C_m^{m\times m}$，使

$$EA = B = \begin{bmatrix} G \\ \cdots\cdots \\ O \end{bmatrix} \begin{array}{l} r行 \\ \\ (m-r)行 \end{array}, \quad 其中\ G \in C_r^{r\times n}$$

将 A 写成 $A=E^{-1}B$，并把 E^{-1} 分块成 $E^{-1}=[\underset{r列}{F}\ \mid\ \underset{(m-r)列}{S}]$，其中 $F\in C_r^{m\times r}$

$$\therefore A=\begin{bmatrix} & \cdot & \\ F & \cdot & S \\ & \cdot & \end{bmatrix}\begin{bmatrix} G \\ \cdots \\ O \end{bmatrix}=FG$$

是满秩分解。

3. Hermite 标准形（行阶梯标准形）

设 $B\in C_r^{m\times n}(r>0)$，且满足

（1）B 的前 r 行中每一行至少含一个非零元素（称为非零行），且第一个非零元素为 1，而后 $(m-r)$ 行的元素全为零（称为零行）;

（2）若 B 中第 i 行的第一个非零元素（即 1）在第 j_i 列 $(i=1,2,\ldots,r)$，则 $j_1<j_2<\ldots<j_r$;

（3）矩阵 B 的第 j_1 列，第 j_2 列，…,第 j_r 列合起来恰为 m 阶单位方阵 I_m 的前 r 列（即 $j_1,j_2,\ldots,j_r$ 列上除了前述的 1 外全为 0）

则称 B 为 Hermite 标准形。

例 1. $B_1 = \begin{bmatrix} 1 & 2 & 0 & 0 & -1 & 3 \\ 0 & 0 & 1 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{5\times 6} \in C_3^{5\times 6}$ 为 **Hermite** 标准形

$B_2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}_{4\times 5} \in C_2^{4\times 5}$ 也是 **Hermite** 标准形

4. 满秩分解的一种求法

设 $A \in C_r^{m\times n}$，

（1） 采用行初等变换将 A 化成 **Hermite** 标准形，其矩阵形式为

$EA = B$，其中 B 为 **Hermite** 标准形；

（2） 选取 n 阶置换矩阵

1° **P** 的第 *i* 列为 e_{j_i}，即该列向量除第 j_i 个元素为 **1** 外，其余元素全为零 $(i=1,2,\ldots,r)$,其中 j_i 为 **Hermite** 标准形中每行第一个非零元素（即 **1**）所在的列数；

2° 其它 $(n-r)$ 列只需确保 **P** 为置换矩阵即可（**P** 的每一行，每一列均只有一个非零元素，且为 **1**）；

3° 用 **P** 右乘任何矩阵（可乘性得到满足时），即可将该矩阵的第 j_i 列置换到新矩阵（即乘积矩阵）的第 *i* 列

4° 令 $P=[\underset{r列}{P_1} \quad | \quad \underset{(n-r)列}{*}]$，即 $P_1=\begin{bmatrix} e_{j_1} & e_{j_2} \ldots e_{j_r} \end{bmatrix}_{n\times r} \in C_r^{n\times r}$

（3）令 $G=B$ 的前 r 行 $\in C_r^{r\times n}$， $F=AP_1\in C_r^{m\times r}$ 则 $A=FG$

证明： $EA=B=\begin{bmatrix} G \\ O \end{bmatrix}$， $A=E^{-1}B=[F \quad | \quad S]\begin{bmatrix} G \\ O \end{bmatrix}=FG$ 则 $F\in C_r^{m\times r}$，

$G\in C_r^{r\times n}$， G 已知， 但 $F=?$， 注意到 $BP_1=\begin{bmatrix} I_r \\ O \end{bmatrix}$， 则

$AP_1=E^{-1}BP_1=[F \quad | \quad S]\begin{bmatrix} I_r \\ O \end{bmatrix}=F$ [证毕]

例 2. $A=\begin{bmatrix}1 & 2 & 3 & 0\\ 0 & 2 & 1 & -1\\ 1 & 0 & 2 & 1\end{bmatrix}$求其满秩分解

解：（1）首先求出A的秩。显然，前两行互相独立，而第三行可由第一行减去第二行得到，故$r=2$。

(2)进行初等变换将A化为 Hermite 标准型。

$$[A \mid I_3]=\begin{bmatrix}1 & 2 & 3 & 0 & . & 1 & 0 & 0\\ 0 & 2 & 1 & -1 & . & 0 & 1 & 0\\ 1 & 0 & 2 & 1 & . & 0 & 0 & 1\end{bmatrix}\xrightarrow{(3)-(1)+(2)}\begin{bmatrix}1 & 2 & 3 & 0 & . & 1 & 0 & 0\\ 0 & 2 & 1 & -1 & . & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & . & -1 & 1 & 1\end{bmatrix}$$

$$\xrightarrow{(1)-(2)}\begin{bmatrix}1 & 0 & 2 & 1 & . & 1 & -1 & 0\\ 0 & 2 & 1 & -1 & . & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & . & -1 & -1 & 1\end{bmatrix}\xrightarrow{(2)/2}\begin{bmatrix}1 & 0 & 2 & 1 & . & 1 & -1 & 0\\ 0 & 1 & 1/2 & -1/2 & . & 0 & 1/2 & 0\\ 0 & 0 & 0 & 0 & . & -1 & 1 & 1\end{bmatrix}$$

即

$$E=\begin{bmatrix}1 & -1 & 0\\ 0 & 1/2 & 0\\ -1 & 1 & 1\end{bmatrix},\quad B=\begin{bmatrix}1 & 0 & 2 & 1\\ 0 & 1 & 1/2 & -1/2\\ 0 & 0 & 0 & 0\end{bmatrix},G=\begin{bmatrix}1 & 0 & 2 & 1\\ 0 & 1 & 1/2 & -1/2\end{bmatrix}$$

(3)求出$\mathbf{P}_1$及$\mathbf{AP}_1$

由B可见，$j_1=1, j_2=2$故$P_1=\begin{bmatrix}1&0\\0&1\\0&0\\0&0\end{bmatrix}$，$F=AP_1=\begin{bmatrix}1&2\\0&2\\1&0\end{bmatrix}$

可以验证：$FG=\begin{bmatrix}1&2\\0&2\\1&0\end{bmatrix}\begin{bmatrix}1&0&2&1\\0&1&1/2&-1/2\end{bmatrix}=\begin{bmatrix}1&2&3&0\\0&2&1&-1\\1&0&2&1\end{bmatrix}$

而$E^{-1}=\begin{bmatrix}1&2&0\\0&2&0\\1&0&1\end{bmatrix}$

二、酉对角分解与奇异值分解

1. 厄米矩阵的谱分解

A为厄米矩阵，则存在酉矩阵U，使

$$\mathbf{U}^H\mathbf{A}\mathbf{U}=\begin{bmatrix}\lambda_1 & & & \mathbf{O}\\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{O} & & & \lambda_n\end{bmatrix}=\Lambda$$

将U写成列向量形式，即$\mathbf{U}=[\mathbf{u}_1\quad \mathbf{u}_2\quad \dots\quad \mathbf{u}_n]$，则

$$\mathbf{A}=\mathbf{U}\Lambda\mathbf{U}^H=[\mathbf{u}_1\quad \mathbf{u}_2\quad \dots\quad \mathbf{u}_n]\begin{bmatrix}\lambda_1 & & & & \mathbf{O}\\ & \lambda_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ \mathbf{O} & & & & \lambda_n\end{bmatrix}\begin{bmatrix}\mathbf{u}_1^H\\ \mathbf{u}_2^H\\ \cdot\\ \cdot\\ \mathbf{u}_n^H\end{bmatrix}=\sum_{i=1}^{n}\lambda_i\mathbf{u}_i\mathbf{u}_i^H$$

2. 非奇异矩阵的酉对角分解

定理：设 **A** 为 n 阶非奇异矩阵，则存在 n 阶酉矩阵 **U** 及 **V**，使得

$$\mathbf{U}^{H}\mathbf{A}\mathbf{V}=\begin{bmatrix}\sigma_1 & & & & O\\ & \sigma_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ O & & & & \sigma_n\end{bmatrix},\quad \sigma_i>0(i=1,2,\ldots,n)$$

（若将 **U**， **V** 写成 $\mathbf{U}=[\mathbf{u}_1\quad \mathbf{u}_2\quad \ldots\quad \mathbf{u}_n],\mathbf{V}=[\mathbf{v}_1\quad \mathbf{v}_2\quad \ldots\quad \mathbf{v}_n]$，则 $A=\sum_{i=1}^{n}\sigma_i u_i v_i^H$ ）

证明：$\mathbf{A}^{H}\mathbf{A}$ 为 n 阶非奇异矩阵，而且是厄米、正定矩阵，故存在 n 阶酉矩阵 **V**，使

$$\mathbf{V}^{\mathrm{H}}(\mathbf{A}^{\mathrm{H}}\mathbf{A})\mathbf{V}=\begin{bmatrix}\sigma_1^2 & & & & \mathbf{O}\\ & \sigma_2^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ \mathbf{O} & & & & \sigma_n^2\end{bmatrix}$$

σ_i^2 为 $\mathbf{A}^{\mathrm{H}}\mathbf{A}$ 的特征值。

$$\Lambda=\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n\end{bmatrix}$$，则 $V^H A^H AV=\Lambda^2$

$\Lambda^{-1}V^H A^H AV=(AV\Lambda^{-1})^H AV=\Lambda$ 令 $U=AV\Lambda^{-1}$，则

$U^H U=(AV\Lambda^{-1})^H(AV\Lambda^{-1})=\Lambda^{-1}V^H A^H AV\Lambda^{-1}=I_n$

所以，**U** 为酉矩阵，且 $U^H AV=\Lambda$

[证毕]

称：$A=U\Lambda V^H$ 为 **A** 的酉对角分解。

3. 一般矩阵的奇异值分解

（1）定义： 设 $A\in C_r^{m\times n}$（r>0），$A^H A$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，其中 $\lambda_1,\lambda_2,\cdots,\lambda_r>0,\lambda_{r+1},\cdots,\lambda_n=0$，则称 $\sigma_i=\sqrt{\lambda_i}\quad (i=1,2,\cdots,n)$ 为 A 的奇异值。

（2）定理： 设 $A\in C_r^{m\times n}$，则存在 *m* 阶酉矩阵 U 及 *n* 阶酉矩阵 V，使

$$U^H AV=\begin{array}{cc} \left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & O \\ & & \ddots & & \\ 0 & & & \sigma_r & \\ \hline & & O & & O \end{array}\right] & \begin{array}{l} \\ \\ r\text{行} \\ \\ (m-r)\text{行} \end{array} \\ \begin{array}{cc} r\text{列} & (n-r)\text{列} \end{array} & \end{array}$$

即

$$A=U\begin{bmatrix} \sigma_1 & & & & O \\ & \sigma_2 & & & \\ & & \cdot & & \\ & & & \sigma_r & \\ O & & & & O \end{bmatrix}V^H$$

证明：首先考虑 $A^H A$。因为 $\mathrm{rank}(A^H A)=\mathrm{rank}(AA^H)=\mathrm{rank}A$，故 $A^H A \in C_r^{n\times n}$，

而且是厄米、半正定的，存在 *n* 阶酉矩阵v，使

$$V^H(A^H A)V=\begin{bmatrix}\sigma_1^2 & & & & O\\ & \sigma_2^2 & & & \\ & & \cdot & & \\ & & & \sigma_r^2 & \\ O & & & & O\end{bmatrix}_{n\times n}$$

令 $\sum=\begin{bmatrix}\sigma_1 & & & O\\ & \sigma_2 & & \\ & & \cdot & \\ & & & \cdot & \\ O & & & \sigma_r\end{bmatrix}$， $V=\begin{array}{cc}[V_1 & | \quad V_2]\\ r列 & (n-r)列\end{array}$ 则

$$V^H(A^H A)V=\begin{bmatrix}V_1^H(A^H A)V_1 & V_1^H(A^H A)V_2\\ V_2^H(A^H A)V_1 & V_2^H(A^H A)V_2\end{bmatrix}=\begin{bmatrix}\sum^2 & O\\ O & O\end{bmatrix}$$

$$V_1^H(A^H A)V_1=\sum{}^2 \qquad V_1^H(A^H A)V_2=O_{r\times(n-r)} \qquad V_2^H(A^H A)V_2=O_{(n-r)\times(n-r)}$$

令$U_1 = AV_1\sum^{-1}$ 则$U_1^H AV_1 = \sum$ ， 又$(AV_2)^H(AV_2) = 0 \to AV_2 = 0$

在U_1的基础上构造酉矩阵$U = [U_1 \quad | \quad U_2]$，即$U^H U = I$

这由前面基扩充定理可知是可行的，

$$U_1^H U_1 = I_r, U_1^H U_2 = O_{r\times(n-r)}, U_2^H U_2 = I_{n-r}$$

故

$$U^H AV = \begin{bmatrix} U_1^H \\ U_2^H \end{bmatrix} A\begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} U_1^H AV_1 & U_1^H AV_2 \\ U_2^H AV_1 & U_2^H AV_2 \end{bmatrix}$$

其中已知

$$U_1^H AV_1 = \sum \quad 而 \quad U_1^H AV_2 = 0, U_2^H AV_2 = 0 \qquad (\because AV_2 = 0)$$

$$U_2^H AV_1 = U_2^H(U_1\sum\ ) = (U_2^H U_1)\sum = 0$$

[证毕]

例2 已知矩阵$A=\begin{bmatrix}1&0&1\\0&1&1\\0&0&0\end{bmatrix}$，求其奇异值分解。

解：

$$A^TA=\begin{bmatrix}1&0&1\\0&1&1\\1&1&2\end{bmatrix}$$ 其特征值为3、1、0

它们对应的特征向量依次为：

$$\zeta_1=\begin{bmatrix}1\\1\\2\end{bmatrix},\quad \zeta_2=\begin{bmatrix}1\\-1\\0\end{bmatrix},\quad \zeta_3=\begin{bmatrix}1\\1\\-1\end{bmatrix}$$

因此，$rankA=2$，$\Sigma=\begin{bmatrix}\sqrt{3}&0\\0&1\end{bmatrix}$

则，正交矩阵 $V=\begin{bmatrix}\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}}\\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}}\\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}}\end{bmatrix}$

$$U_1=AV_1\Sigma^{-1}=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ 0 & 0\end{bmatrix}$$

构造 $U=\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0\\ 0 & 0 & 1\end{bmatrix}$

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

作业： **P225 1(2)， 2， 5；**
P233 1