§4　奇异值分解

一、奇异值分解

先给出如下引理（其中一些结论前面已证过）：

引理　设 $\boldsymbol{A}\in\mathbf{C}^{m\times n}$，则：

（1） $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的特征值均为非负实数；

（2） $\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\operatorname{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})=\operatorname{rank}\boldsymbol{A}$；

（3） $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的非零特征值相同。

证（1）设 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$，则由 $0\le\|\boldsymbol{A}\boldsymbol{x}\|_2^2=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\lambda\|\boldsymbol{x}\|_2^2$ 知 λ 为非负实数。

（2）由 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 得 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$。反之，有

$$\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\Rightarrow\|\boldsymbol{A}\boldsymbol{x}\|_2^2=\boldsymbol{x}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=0\Rightarrow\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$$

从而 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 与 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 同解，故它们的基础解系所含线性无关的向量个数相同，于是

$$n-\operatorname{rank}\boldsymbol{A}=n-\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})\quad\text{即}\quad\operatorname{rank}(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\operatorname{rank}\boldsymbol{A}$$

而 $\operatorname{rank}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})=\operatorname{rank}\boldsymbol{A}^{\mathrm{H}}=\operatorname{rank}\boldsymbol{A}$。

（3）由公式 $\det(\boldsymbol{I}_m+\boldsymbol{A}\boldsymbol{B})=\det(\boldsymbol{I}_n+\boldsymbol{B}\boldsymbol{A})$ 得，当 $\lambda\ne0$ 时，

$$\det(\lambda\boldsymbol{I}_n-\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\lambda^n\det(\boldsymbol{I}_n-\frac{1}{\lambda}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A})=\lambda^n\det(\boldsymbol{I}_m-\frac{1}{\lambda}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})=\lambda^{n-m}\det(\lambda\boldsymbol{I}_m-\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})$$

故 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 与 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ 的非零特征值相同。**证毕**

定义　设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$（$r>0$），且记 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的特征值为

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

则称 $\sigma_i=\sqrt{\lambda_i}$（$i=1,2,\cdots,n$）为 $\boldsymbol{A}$ 的**奇异值**。

由引理知，矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{A}^{\mathrm{H}}$ 有相同的非零奇异值，且非零奇异值的个数等于 $\operatorname{rank}\boldsymbol{A}$。

定理　设 $\boldsymbol{A}\in\mathbf{C}_r^{m\times n}$（$r>0$），则存在 m 阶酉矩阵 $\boldsymbol{U}$ 和 n 阶酉矩阵 $\boldsymbol{V}$，使得

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

其中 $\Sigma=\mathrm{diag}(\sigma_1,\sigma_2,\cdots,\sigma_r)$，而$\sigma_i$（$i=1,2,\cdots,r$）为$A$的非零奇异值。称

$A=U\begin{pmatrix}\Sigma & O\\ O & O\end{pmatrix}V^{\mathrm{H}}$为$A$的**奇异值分解**。

证 设A的奇异值为$\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>\sigma_{r+1}=\cdots=\sigma_n=0$，则存在$n$阶酉矩阵$V$，

使得 $$V^{\mathrm{H}}A^{\mathrm{H}}AV=\mathrm{diag}(\sigma_1^2,\cdots,\sigma_r^2,0,\cdots0)=\begin{pmatrix}\Sigma^2 & O\\ O & O\end{pmatrix}$$

令$V=[V_1,V_2]$，其中$V_1\in\mathbf{C}^{n\times r}$，$V_2\in\mathbf{C}^{n\times(n-r)}$，代入上式并比较得

$$V_1^{\mathrm{H}}A^{\mathrm{H}}AV_1=\Sigma^2,\quad V_2^{\mathrm{H}}A^{\mathrm{H}}AV_2=O$$

于是 $$\Sigma^{-1}V_1^{\mathrm{H}}A^{\mathrm{H}}AV_1\Sigma^{-1}=I_r,\quad AV_2=O$$

令$U_1=AV_1\Sigma^{-1}$，由上式知 $U_1^{\mathrm{H}}U_1=I_r$，即U_1的列由正交的单位向量构成，将U_1扩充成m阶酉矩阵 $U=(U_1,U_2)$，注意到$U_2^{\mathrm{H}}U_1=O$得

$$U^{\mathrm{H}}AV=\begin{pmatrix}U_1^{\mathrm{H}}AV_1 & U_1^{\mathrm{H}}AV_2\\ U_2^{\mathrm{H}}AV_1 & U_2^{\mathrm{H}}AV_2\end{pmatrix}=\begin{pmatrix}U_1^{\mathrm{H}}U_1\Sigma & O\\ U_2^{\mathrm{H}}U_1\Sigma & O\end{pmatrix}=\begin{pmatrix}\Sigma & O\\ O & O\end{pmatrix}$$ **证毕**

推论 设$A\in\mathbf{C}_n^{n\times n}$，即$A$是$n$阶可逆矩阵，则存在$n$阶酉矩阵$U$和$V$使

$U^{\mathrm{H}}AV=\mathrm{diag}(\sigma_1,\cdots,\sigma_n)$，其中$\sigma_1,\cdots,\sigma_n$是$A$的非零奇异值

推论 设$A\in\mathbf{R}_n^{n\times n}$，则存在$n$阶正交矩阵$P$和$Q$，使

$$P^{\mathrm{H}}AQ=\mathrm{diag}(\sigma_1,\cdots,\sigma_n)$$

推论 在$A\in\mathbf{C}_r^{m\times n}$的奇异值分解$A=U\begin{pmatrix}\Sigma & O\\ O & O\end{pmatrix}V^{\mathrm{H}}$中，$V$的列向量是$A^{\mathrm{H}}A$的特征向量，而$U$的列向量是$AA^{\mathrm{H}}$的特征向量。

证 V的列向量是$A^{\mathrm{H}}A$的特征向量，由定理的证明过程即可知。而由

$$AA^{\mathrm{H}}=U\begin{pmatrix}\boldsymbol{\Sigma} & O\\ O & O\end{pmatrix}V^{\mathrm{H}}V\begin{pmatrix}\boldsymbol{\Sigma} & O\\ O & O\end{pmatrix}U^{\mathrm{H}}=U\begin{pmatrix}\boldsymbol{\Sigma}^2 & O\\ O & O\end{pmatrix}_{m\times m}U^{\mathrm{H}}$$

得
$$AA^{\mathrm{H}}U=U\ \mathrm{diag}(\sigma_1^2,\cdots,\sigma_r^2,0,\cdots0)$$

可见U的列向量是AA^{H}的特征向量。**证毕**

例 求矩阵$A=\begin{pmatrix}1&0&0\\2&0&0\end{pmatrix}$的奇异值分解。

解 $A^{\mathrm{H}}A=\begin{pmatrix}5&0&0\\0&0&0\\0&0&0\end{pmatrix}$的特征值为$\lambda_1=5,\lambda_2=\lambda_3=0$，相应的特征向量为

$x_1=e_1$，$x_2=e_2$，$x_3=e_3$，从而 $V=(e_1,e_2,e_3)=I_3$。

法 1. 取$V_1=e_1$，$\boldsymbol{\Sigma}=(\sqrt{5})$，于是$U_1=AV_1\boldsymbol{\Sigma}^{-1}=\begin{pmatrix}\frac{1}{\sqrt5}\\ \frac{2}{\sqrt5}\end{pmatrix}$，取$U_2=\begin{pmatrix}-\frac{2}{\sqrt5}\\ \frac{1}{\sqrt5}\end{pmatrix}$，则

$U=(U_1,U_2)$为酉矩阵，故A的奇异值分解为

$$A=\begin{pmatrix}\frac{1}{\sqrt5}&-\frac{2}{\sqrt5}\\ \frac{2}{\sqrt5}&\frac{1}{\sqrt5}\end{pmatrix}\begin{pmatrix}\sqrt5&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

法 2. 因为$AA^{\mathrm{T}}=\begin{pmatrix}1&2\\2&4\end{pmatrix}$，其特征值为$\lambda_1=5,\lambda_2=0$，对应的特征向量分别为$\begin{pmatrix}1\\2\end{pmatrix}$和$\begin{pmatrix}-2\\1\end{pmatrix}$，故酉矩阵$U=\frac{1}{\sqrt5}\begin{pmatrix}1&-2\\2&1\end{pmatrix}$，且$A$的奇异值分解为

$A=U\begin{pmatrix}\sqrt5&0&0\\0&0&0\end{pmatrix}V^{\mathrm{T}}$。

二、酉相抵

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

$$U^{\mathrm{H}}AV=B$$

则称A与B**酉相抵**（或**酉等价**）。

酉等价矩阵有如下一些性质：

性质 1 A与A酉等价； （因为$A=I_mAI_n$）

性质 2　若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉等价，则 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 酉等价；

性质 3　若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 酉等价， $\boldsymbol{B}$ 与 $\boldsymbol{C}$ 酉等价，则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ 酉等价；

性质 4　酉等价矩阵有相同的奇异值。

证　设 $\boldsymbol{B}=\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}$,则 $\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}=\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}=\boldsymbol{V}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{V}$,即 $\boldsymbol{B}^{\mathrm{H}}\boldsymbol{B}$ 与 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 酉相似，它们有相同的特征值，从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的奇异值。**证毕**