2.4 矩阵的奇异值分解

对于方阵, 利用其特征值和特征向量可以刻画矩阵的结构。对长方阵情形, 这些方法已经不适用。

而推广的特征值--矩阵的奇异值分解理论能改善这种情况。利用奇异值和奇异向量不仅可以刻画矩阵的本身结构, 而且还可以进一步刻画线代数方程组的解的结构, 是构造性的研究线代数问题的有利的工具。

定义2.10 设 $A\in \mathbf{C}^{m\times n}$, $k=\min(m,n)$

Hermite半正定矩阵 $A^H A$ 的特征值为

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_k \ge 0$$

称非负实数

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

为矩阵A的奇异值。

矩阵A的奇异值满足如下性质：

定理 2.13 设 A、$B \in \mathbf{C}^{m\times n}$，如果存在$m$阶、$n$阶酉阵$U$、$V$，使得 $A=UBV^H$，则矩阵A、B的奇异值相同。

证：由 $U^H AV = B$，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，进而命题得证。

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

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

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$\boldsymbol{A}$的非零奇异值。

证 由于 rank$(\boldsymbol{A}^H\boldsymbol{A})$=rank$(\boldsymbol{A})$，因此$\boldsymbol{A}^H\boldsymbol{A}$是秩为$r$的$n$阶**Hermite**半正定矩阵，由**Hermite**半正定矩阵的性质，设其特征值为：$\sigma_1^2,\sigma_2^2,\cdots,\sigma_r^2$，且

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

由推论2.2，必存在n阶酉阵$\boldsymbol{V}$，使得

$$\boldsymbol{V}^H(\boldsymbol{A}^H\boldsymbol{A})\boldsymbol{V} = \begin{pmatrix} \sigma_1^2 & & & & & & \\ & \sigma_2^2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_r^2 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}^2 & 0 \\ 0 & 0 \end{pmatrix}$$

记 $V_1=(v_1,v_2,\cdots,v_r)\in\mathbf{C}^{n\times r}$， $V_2=(v_{r+1},v_{r+2},\cdots,v_n)\in\mathbf{C}^{n\times(n-r)}$,

则可将$\boldsymbol{V}$分块成$\boldsymbol{V}=(\boldsymbol{V}_1\ \ \boldsymbol{V}_2)$，这样有分块形式:

$$\begin{pmatrix}\boldsymbol{V}_1^H\\ \boldsymbol{V}_2^H\end{pmatrix}\boldsymbol{A}^H\boldsymbol{A}\begin{pmatrix}\boldsymbol{V}_1 & \boldsymbol{V}_2\end{pmatrix}$$

即

$$\begin{pmatrix}\boldsymbol{V}_1^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_1 & \boldsymbol{V}_1^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_2\\ \boldsymbol{V}_2^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_1 & \boldsymbol{V}_2^H\boldsymbol{A}^H\boldsymbol{A}\boldsymbol{V}_2\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Sigma}^2 & 0\\ 0 & \mathbf{0}\end{pmatrix}$$

其中 $\boldsymbol{\Sigma}^2=\mathrm{diag}(\sigma_1^2,\sigma_2^2,\cdots,\sigma_r^2)$。由此得出:

$\boldsymbol{\Sigma}^{-1}$ $\boldsymbol{\Sigma}^{-1}=\boldsymbol{I}_r$ 或 $\left(\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1}\right)^H\left(\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1}\right)=\boldsymbol{I}_r$

$=$ $\Leftrightarrow \left(\boldsymbol{A}\boldsymbol{V}_2\right)^H\left(\boldsymbol{A}\boldsymbol{V}_2\right)=0 \Leftrightarrow \boldsymbol{A}\boldsymbol{V}_2=0$

现取 $\boldsymbol{U}_1=\boldsymbol{A}\boldsymbol{V}_1\boldsymbol{\Sigma}^{-1}\in\mathbf{C}^{m\times r}$, 于是

$$U_1^{\mathrm{H}}U_1 = \Sigma^{-1}V_1^H A^H AV_1\Sigma^{-1} = I_r$$ 因此矩阵

$U_1 = (u_1, u_2, \cdots, u_r)$ 的列是$\mathbf{C}^m$中的一个标准正交向量组。

又有

$$A = U_1\ \Sigma\ V_1^H \in \mathbf{C}^{m\times n},\quad \left((U_1)_{m\times r}\ \Sigma_{\mathrm{r\times r}}\ \left(V_1^H\right)_{r\times n} \in \mathbf{C}^{m\times n}\right)$$

或 $U_1\Sigma = AV_1$。再将U_1扩充为$\mathbf{C}^m$的一标准正交基

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

令 $U_2 = (u_{r+1}, u_{r+2}, \cdots, u_m)$, $U_2 \in \mathbf{C}^{m\times(m-r)}$。则得到

$U = \begin{pmatrix} U_1 & U_2 \end{pmatrix} \in \mathbf{C}^{m\times n}$ 为酉阵。 则我们有

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

$$\begin{pmatrix} U_1\Sigma & 0 \end{pmatrix},\quad \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$$

如上关系式称为矩阵A 的奇异值分解，亦称为矩阵A 的满的奇异值分解。简称SVD定理。

关系式亦可写为

$$A = U_1 \Sigma V_1^H$$

并称它为矩阵A 约化的奇异值分解。

由 $AV_1 = U_1\Sigma$ 和 $U_1^H A = \Sigma V_1^H$ 可得

$$Av_i = \sigma_i u_i \,,\ \ u_i^H A = \sigma_i v_i^H, \quad i = 1, 2, \cdots, r \qquad \textbf{(2-53)}$$

分别称 u_i^H 和 v_i 为矩阵A的与奇异值 σ_i 对应的左奇异向量和右奇异向量。

U与V的列向量 $u_1, u_2, \cdots, u_m$ 和 $v_1, v_2, \cdots, v_n$ 从（**2-47**）可得

$$AA^HU = UU^HAVV^HA^HU = U\left(U^HAV\right)(U^HAV)^H$$

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

$$A^HAV = VV^HA^HUU^HAV = V\left(U^HAV\right)^H(U^HAV)$$

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

从而

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，注意到V满足： $V^H(A^H A)V=\begin{pmatrix}\boldsymbol{\Sigma}^2 & 0\\ 0 & 0\end{pmatrix}$ 故可知V的列是$A^H A$的特征值所对应的特征向量，所以只需求解如下的方程组：

$$\begin{pmatrix}2 & 0 & -1\\ 0 & 2 & -1\\ -1 & -1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}\Leftrightarrow\begin{aligned}2x_1\qquad -x_3&=0\\ 2x_2-x_3&=0\\ x_1+x_2-x_3&=0\end{aligned}\Rightarrow\boldsymbol{p}_1=\begin{pmatrix}1\\ 1\\ 2\end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{aligned} x_3 &= 0 \\ x_3 &= 0 \\ x_1+x_2+x_3 &= 0 \end{aligned} \Rightarrow \boldsymbol{p}_2=\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{aligned} x_1 \qquad + x_3 &= 0 \\ x_2 + x_3 &= 0 \\ x_1+x_2+2x_3 &= 0 \end{aligned} \Rightarrow \boldsymbol{p}_3=\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

将它们标准化，得到酉阵$\boldsymbol{V}$的列:

$$\boldsymbol{v}_1=\frac{\boldsymbol{p}_1}{\|\boldsymbol{p}_1\|_2}=\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \boldsymbol{v}_2=\frac{\boldsymbol{p}_2}{\|\boldsymbol{p}_2\|_2}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \boldsymbol{v}_3=\frac{\boldsymbol{p}_3}{\|\boldsymbol{p}_3\|_2}=\frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

即得 $V=\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{1}{\sqrt{3}} \end{pmatrix}$，因**rank(*A*)=2**，故有$(V_1)_{3\times 2}=\begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & 0 \end{pmatrix}$

进一步计算得出，

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

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$\mathbf{R}^3$的一组标准正交基，可取 $U_2=\begin{pmatrix}0\\0\\1\end{pmatrix}$

则

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

是酉阵, 故矩阵A的奇异值分解（满的奇异值分解）为

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

2.4.3 用矩阵的奇异值讨论矩阵的性质

下面均假定可以通过某种可靠的数值方法计算出矩阵的奇异值分解，据此讨论矩阵的一些性质。

定理2.15 矩阵A的非零奇异值的个数恰为矩阵A的秩。

证：注意，$A^H A$的非零特征值的个数应为$\mathrm{rank}(A^H A)$，又由于$\mathbf{rank}(A^H A)=\mathbf{rank}(A)$，从而$A$的非零奇异值个数恰等于$\mathbf{rank}(A)$。

该定理表明，借助矩阵的奇异值分解，可以得到计算矩阵A的秩的数值方法，同时它也是判断一个向量组是否线性相关的数值方法。

定理2.16

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

其中 $\forall \boldsymbol{A}\in\mathbf{R}^{m\times n}, \mathbf{R}(A)=\{\boldsymbol{y}\in\mathbf{R}^m \mid \boldsymbol{Ax}=\boldsymbol{y}, \forall \boldsymbol{x}\in \boldsymbol{R}^n\}$为由$A$的列向量生成的子空间，称为$A$的值域或像空间，即

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

$\mathbf{N}\left(A\right)=\{\boldsymbol{x}\in \boldsymbol{R}^n \mid \boldsymbol{Ax}=\boldsymbol{0}\}$ 称为A的零空间或核，即$\mathbf{N}(A)=\{\boldsymbol{x}\,|\,\boldsymbol{Ax}=\boldsymbol{0}\}$

证：由 $AV_1=U_1\Sigma=\left(\sigma_1\boldsymbol{u}_1,\sigma_2\boldsymbol{u}_2,\cdots,\sigma_r\boldsymbol{u}_r\right)=\left(\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_r\right)$；

$AV_2=\mathbf{0}$。

该定理表明，借助矩阵的奇异值(**SVD**)分解，可以确定子空间**R(*A*)**和**N(A)**的一组标准正交基。

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

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

证: $\|A\|_2^2 = \lambda_{\max}(A^H A) = \sigma_1^2 \quad \Longrightarrow \quad \|A\|_2 = \sigma_1$。

$$\|A\|_F^2 = \left\| U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^H \right\|_F^2 = \left\| \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \right\|_F^2$$

$$= \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2$$

该定理表明，借助矩阵的奇异值(**SVD**)分解，我们可以确定矩阵 A的2-范数和F-范数。

定理2.18 如果A为Hermite矩阵，则A的奇异值即为$A^H=A$的特征值的绝对值。

证： 由$A^HA=A^2$，则 $\sigma_i=\sqrt{\lambda_i^2}=|\lambda_i|$。

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

证： $|\det(A)|=\sqrt{\det(A)^2}=\sqrt{\det(A^HA)}$

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

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

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

证：由矩阵A 约化的奇异值分解：$\boldsymbol{A}=\boldsymbol{U}_1\Sigma\boldsymbol{V}_1^H$ 可知,

$$\boldsymbol{U}_1=(\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_r)\qquad \boldsymbol{V}_1^H=\left(v_1^H,\ v_2^H,\cdots,v_r^H\right)^T$$

$$\Sigma=\begin{pmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r\end{pmatrix}\qquad \boldsymbol{U}_1\Sigma=\left(\sigma_1u_1,\sigma_2u_2,\cdots,\sigma_ru_r\right)$$

$$\boldsymbol{A}=\boldsymbol{U}_1\Sigma V_1^H=\left(\sigma_1u_1,\sigma_2u_2,\cdots,\sigma_ru_r\right)\begin{pmatrix}v_1^H\\ v_2^H\\ \vdots\\ v_r^H\end{pmatrix}$$

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