第二章 矩阵的分解

2.4 奇异值分解

引理1: 设$A \in C^{m\times n}$, 则$r(A^H A) = r(AA^H) = r(A)$.

证明: 如果$x \in C^n$是齐次线性方程组$AX = 0$的解, 则x也是$A^H AX = 0$的解; 反之, 如果x是$A^H AX = 0$的解, 则$x^H A^H Ax = 0$, 即$(Ax)^H Ax = 0$, 于是$Ax = 0$, 这表明x也是$AX = 0$的解, 因此, $AX = 0$与$A^H AX = 0$同解, 从而$r(A^H A) = r(A)$.

同理, $r(AA^H) = r(A^H) = r(A)$.

引理2: 设$A \in C^{m\times n}$, 则

1) A^HA与AA^H的特征值均为非负实数.

2) A^HA与AA^H的非零特征值相同且非零特征值的个数为$r(A)$.

证明: 1) 设λ为A^HA的任一特征值, $x \neq 0$为相应的特征向量, 则$A^HAx = \lambda x$. 因A^HA为Hermite阵, 所以λ为实数, 且$\lambda x^Hx = x^HA^HAx = (Ax)^HAx \geq 0$. 由于$x^Hx > 0$, 所以$\lambda \geq 0$.

同理, AA^H的特征值均为非负实数.

设$r(A)=r$, λ_i为AA^H的特征值, μ_i为A^HA的特征值, 都为实数, 设

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

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_r > \mu_{r+1} = \mu_{r+2} = \dots = \mu_n = 0.$$

下证$\lambda_i = \mu_i > 0.\ (i=1,\dots,r)$

由$AA^Hx_i = \lambda_i x_i$可得(左乘A^H) $A^HAA^Hx_i = \lambda_i A^Hx_i$, 这表明$\lambda_i$既是$AA^H$的特征值, 也是$A^HA$的特征值, 同理$\mu_i$也是$AA^H$的特征值. (但不能认为$\lambda_i = \mu_i$)

设$x_1, \ldots, x_p$为AA^H对应于$\lambda_i \neq 0$的线性无关特征向量，由上述讨论知$A^H x_1, \ldots, A^H x_p$为$A^H A$属于$\lambda_i \neq 0$的特征向量，只要证明$A^H x_1, \ldots, A^H x_p$线性无关，就证明了AA^H的p重特征值也是$A^H A$的p重特征值.

下证$A^H x_1, \dots, A^H x_p$线性无关.

设$k_1 A^H x_1 + \dots + k_p A^H x_p = 0$,

则$(A^H x_1, \dots, A^H x_p)\begin{pmatrix} k_1 \\ \vdots \\ k_p \end{pmatrix} = 0$. 即

$$A^H(x_1, \dots, x_p)\begin{pmatrix} k_1 \\ \vdots \\ k_p \end{pmatrix} = 0,$$

于是$AA^H(x_1,\dots,x_p)\begin{pmatrix}k_1\\ \vdots\\ k_p\end{pmatrix}=0\Rightarrow$

$$(\lambda_i x_1,\dots,\lambda_i x_p)\begin{pmatrix}k_1\\ \vdots\\ k_p\end{pmatrix}=0\xrightarrow{\lambda_i\neq 0}(x_1,\dots,x_p)\begin{pmatrix}k_1\\ \vdots\\ k_p\end{pmatrix}=0.$$

$$(x_1, \ldots, x_p)\begin{pmatrix} k_1 \\ \vdots \\ k_p \end{pmatrix} = 0 \xrightarrow{x_i\text{线性无关}} \begin{pmatrix} k_1 \\ \vdots \\ k_p \end{pmatrix} = 0.$$

引理3: 设$A \in C_r^{m\times n}$, 则$A^H A$与AA^H都是半正定的Hermite阵.

证明: 显然$A^H A$与AA^H都是Hermite阵,

$\forall x \in C^n, y \in C^m$有

$x^H A^H A x = (Ax)^H Ax \geq 0,$

$y^H A A^H y = (A^H y)^H (A^H y) \geq 0.$

定义: 设$A \in C_r^{m\times n}$, $A^H A$的特征值$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > 0, \lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_n = 0$, 称$\sigma_i = \sqrt{\lambda_i}$为$A$的奇异值, $\sigma_i > 0$时称为A的正奇异值.

例: $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{pmatrix}$, 则$A^H A = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}$, 奇异值为$\sqrt{5}$, 1, 0.

而$AA^H = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 4 \end{pmatrix}$, $|\lambda I - AA^H| = \lambda(\lambda - 1)(\lambda - 5)$.

定理1: 若A与B酉相抵, 则A与B有相同的奇异值.

证明: 因A与B酉相抵, 所以存在酉阵U与V, 使$B=UAV$. 所以$B^HB=V^HA^HU^HUAV=V^HA^HAV$,

所以B^HB与A^HA 相似,

所以它们的特征值相同,

所以A与B有相同的奇异值.

定理2(奇异值分解)设$A \in C_r^{m\times n}$, 则存在酉阵$U = C^{m\times m}$与酉阵$V = C^{n\times n}$使

$$A = U \begin{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} V^H.$$

证明: 因$A^H A$是半正定的Hermite阵, 所以为单纯矩阵, 设其特征值为$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_n = 0$, 其对应的标准正交特征向量为$x_1, \cdots, x_r, x_{r+1}, \ldots x_n$,

记$V = (x_1, \cdots, x_r, x_{r+1}, \ldots x_n)$, 则$V$是酉矩阵.

由内积定义知$\left(Ax_i, Ax_j\right) = (Ax_j)^H Ax_i = x_j^H A^H A x_i = \lambda_i x_j^H x_i = \begin{cases} \lambda_i, \ i = j \\ 0, \ i \neq j \end{cases}$

所以$Ax_1, \ldots, Ax_r$为彼此正交的非零向量, 进而

$\frac{Ax_1}{\sigma_1}, \ldots, \frac{Ax_r}{\sigma_r}$为标基, 将其扩为$C^m$的一组标基

$\frac{Ax_1}{\sigma_1}, \ldots, \frac{Ax_r}{\sigma_r}, y_{r+1}, \ldots, y_m$.(注: 扩充不唯一)

令$U = (\frac{Ax_1}{\sigma_1}, \ldots, \frac{Ax_r}{\sigma_r}, y_{r+1}, \ldots, y_m)$, 显然$U$是酉矩阵,

且$U^H AV = U^H(Ax_1, \ldots, Ax_n) =$

$$\left(\frac{Ax_1}{\sigma_1}, \ldots, \frac{Ax_r}{\sigma_r}, y_{r+1}, \ldots, y_m\right)^H (Ax_1, \ldots, Ax_n)$$

$$= \begin{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix},$$

从而$\mathrm{A} = \mathrm{U} \begin{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} V^H.$

注: 奇异值分解不唯一.

奇异值分解的步骤:

1.计算A^HA的特征值$\lambda_1,\cdots,\lambda_r,\lambda_{r+1}=\lambda_{r+2}=\ldots=\lambda_n=0$, 得出$A$的正奇异值$\sqrt{\lambda_1},\ldots,\sqrt{\lambda_r}$.

2.求出A^HA的特征向量$x_1,\cdots,x_n$. 将其标准化, 构成矩阵V.

3.计算$Ax_1,\cdots,Ax_r$($x_1,\cdots,x_r$为非零特征值的特征向量, 实际上$Ax_{r+1}=\cdots=Ax_n=0$), 将其扩为C^m的一组基, 并将其标准化, 构成矩阵U.

则$A = U\begin{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_r} \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} V^H.$

例: $A=\begin{pmatrix}2 & 3\\ 0 & 0\\ 0 & 0\end{pmatrix}$的奇异值分解.

解: $A^HA=\begin{pmatrix}2 & 0 & 0\\ 3 & 0 & 0\end{pmatrix}\begin{pmatrix}2 & 3\\ 0 & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}4 & 6\\ 6 & 9\end{pmatrix}$.

$|\lambda I-A^HA|=\begin{vmatrix}\lambda-4 & -6\\ -6 & \lambda-9\end{vmatrix}=\lambda(\lambda-13)$, $\lambda=$ 13, 0.

故A的奇异值为$\sqrt{13}$.

$A^H A$对应$\lambda = 13$的特征向量为$x_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, 对应$\lambda = 0$的特征向量为$x_2 = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$.标准化为$z_1 = \begin{pmatrix} \frac{2}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \end{pmatrix}, z_2 = \begin{pmatrix} \frac{3}{\sqrt{13}} \\ -\frac{2}{\sqrt{13}} \end{pmatrix}$, 则$V = (z_1 \quad z_2)$.

$Ax_1=\begin{pmatrix}2&3\\0&0\\0&0\end{pmatrix}\begin{pmatrix}2\\3\end{pmatrix}=\begin{pmatrix}13\\0\\0\end{pmatrix}$将其扩为$C^3$的一组基:

$\begin{pmatrix}13\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}$, 标准化为$\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}$, 则$U=I_3$.

$$故A = U\begin{pmatrix} \sqrt{13} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} V^H =$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{13} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} & \frac{-2}{\sqrt{13}} \end{pmatrix}.$$

例: 求$A=\begin{pmatrix}-1 & 0 & 1\\ 2 & 0 & -2\end{pmatrix}$的奇异值分解. (课本例题)

例: 求$A=\begin{pmatrix}-1 & 0 & 1\\ 2 & 0 & -2\end{pmatrix}$的奇异值分解. (课本例题)

解: 令$B=A^H=\begin{pmatrix}-1 & 2\\ 0 & 0\\ 1 & -2\end{pmatrix}$, 则$B^HB=\begin{pmatrix}2 & -4\\ -4 & 8\end{pmatrix}$,

$|\lambda I-B^HB|=\begin{vmatrix}\lambda-2 & 4\\ 4 & \lambda-8\end{vmatrix}=\lambda(\lambda-10)$, $\lambda=$ 10, 0. 故B的奇异值为$\sqrt{10}$,

B^HB对应$\lambda = 10$的特征向量为$x_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$, 对应$\lambda =$

0的特征向量为$x_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, 所以$\mathrm{V} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}$,

$Bx_1 = \begin{pmatrix} -1 & 2 \\ 0 & 0 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \\ 5 \end{pmatrix}$. 将其扩为$C^3$一组

基:

$$\begin{pmatrix}-5\\0\\5\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix}, \text{标准化为} \begin{pmatrix}\frac{-1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix},\begin{pmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},$$

$$\text{则U} = \begin{pmatrix}\frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\\0 & 1 & 1\\\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0\end{pmatrix}, \text{所以} B = U\begin{pmatrix}\sqrt{10} & 0\\0 & 0\\0 & 0\end{pmatrix}V^H,$$

从而$A = V\begin{pmatrix} \sqrt{10} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} U^H =$

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{10} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix}.$$

简化奇异值分解: 设$A \in C_r^{m\times n}$, 则

$$A = U_1 \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} {V_1}^H,$$

其中$U_1 = (y_1, \ldots, y_r)_{m\times r}, V_1 = (x_1, \ldots, x_r)_{n\times r}$是列酉阵.

证明: 由奇异值分解$U\begin{pmatrix}\begin{pmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{pmatrix} & 0 \\ 0 & 0\end{pmatrix}V^H$, 令

$U=(U_1,U_2), V=(V_1,V_2)$,

$$A=(U_1,U_2)\begin{pmatrix}\Delta & 0 \\ 0 & 0\end{pmatrix}\begin{pmatrix}V_1^H \\ V_2^H\end{pmatrix}=U_1\Delta V_1^H.$$

例: 求$A=\begin{pmatrix}1&0\\0&1\\2&0\end{pmatrix}$的简化奇异值分解.

解: $A^HA=\begin{pmatrix}1&0&2\\0&1&0\end{pmatrix}\begin{pmatrix}1&0\\0&1\\2&0\end{pmatrix}=\begin{pmatrix}5&0\\0&1\end{pmatrix}\Rightarrow\lambda=$ 5, 1.故A的奇异值为$\sqrt{5}$, 1. A^HA对应$\lambda=5$的特征向量为$x_1=\begin{pmatrix}1\\0\end{pmatrix}$, 对应$\lambda=1$的特征向量为$x_2=\begin{pmatrix}0\\1\end{pmatrix}$,

令$V=(x_1,x_2)=I$.

$$Ax_1=\begin{pmatrix}1&0\\0&1\\2&0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\0\\2\end{pmatrix},Ax_2=\begin{pmatrix}1&0\\0&1\\2&0\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=$$

$\begin{pmatrix}0\\1\\0\end{pmatrix}$, 标准化为$\begin{pmatrix}\frac{1}{\sqrt{5}}\\0\\\frac{2}{\sqrt{5}}\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix}$, 则$U_1=\begin{pmatrix}\frac{1}{\sqrt{5}}&0\\0&1\\\frac{2}{\sqrt{5}}&0\end{pmatrix}$, 则

$$A=U_1\begin{pmatrix}\sqrt{5}&0\\0&1\end{pmatrix}V^H.$$

极分解: 设$A \in C_r^{n\times n}$, 则A有以下分解, $A = GU$, G为半正定Hermite矩阵, U为酉阵, 特别地, 当A满秩时, G为正定Hermite矩阵, 且分解唯一.

证明: 由奇异值分解:

$$A = U_1 \begin{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} V_1{}^H =$$

$$\left(U_1 \begin{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} U_1{}^H \right) U_1 V_1{}^H.$$

令$G = U_1 \begin{pmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{pmatrix} U_1{}^H, U = U_1 U_1{}^H.$

则G为半正定Hermite阵U为酉阵.

多选题 1分

设置

设$A \in C_r^{n\times n}$，下列说法正确的是（ ）

A $r(A^H A) = r(AA^H) = r$

B $A^H A$，AA^H都是半正定Hermite矩阵.

C A的奇异值的乘积等于A的行列式的绝对值

D A的正奇异值与A^H的正奇异值相同

提交