习题课材料（十）

注: 带♡号的习题有一定的难度、比较耗时, 请量力为之.

一 求以下矩阵的奇异值分解：

(1) $\begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$；

(2) $A = uv^T$, 其中$u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$, $\|u\| = \|v\| = 1$；

(3) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$；

(4) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$；

(5) $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

注记：(1)

$$\begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{bmatrix}$$

.

(3)这是一个列标准正交阵;

(4)这是一个实对称阵；注意特征值和奇异值的关系(特征值的绝对值是奇异值). 设A是n阶实对称阵，满足$Q^TAQ = diag(\lambda_1, \cdots, \lambda_s, \lambda_{s+1}, \cdots, \lambda_r, 0, \cdots, 0)$, 其中$r = rank(A)$, $Q = (q_1, \cdots, q_n)$是正交阵，$\lambda_1, \cdots, \lambda_s > 0$, $\lambda_{s+1}, \cdots, \lambda_r < 0$, 则$A = U\Sigma V^T$, 其中

$$U = Q, \Sigma = diag(|\lambda_1|, \cdots, |\lambda_r|, 0, \cdots, 0), V = (q_1, \cdots, q_s, -q_{s+1}, \cdots, -q_r, q_{r+1}, \cdots, q_n).$$

(5)这是一个反对称阵，注意特征值和奇异值的关系(特征值的模长是奇异值).

二 设$A \in M_{m\times n}(\mathbb{R})$的奇异值分解是$A = U\Sigma V^T$, 满足$U = (u_1,\cdots,u_m)$是$m$阶正交矩阵，$V = (v_1,\cdots,v_n)$是$n$阶正交矩阵，$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}_{m\times n}$，$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, 其中$r = rank(A)$.

(1) 证明$\|Av_1\| = \sigma_1 = max_{\|v\|=1}\|Av\|$.

(2) 证明$\|Av_2\| = \sigma_2 = max_{v\perp v_1,\|v\|=1}\|Av\|$.

证明：(1) $\forall v \in \mathbb{R}^n$, $v = c_1v_1 + \cdots + c_nv_n$. 因此，

$$\|Av\|^2 = v^TA^TAv = c_1^2\lambda_1 + \cdots + c_r^2\lambda_r \leq (c_1^2 + \cdots + c_r^2)\lambda_1 \leq \lambda_1\|v\|^2 = (\sigma_1\|v\|)^2,$$

其中$\lambda_i = \sigma_i^2, i = 1,\cdots,r$. 当$v = v_1$时，等式成立.

(2)设$v \in \mathbb{R}^n$且$v \perp v_1$，则$v = a_2v_2 + \cdots + a_nv_n$.

$$\|Av\|^2 = v^TA^TAv = a_2^2\lambda_2 + \cdots + a_r^2\lambda_r \leq (a_2^2 + \cdots + a_r^2)\lambda_2 \leq \lambda_2\|v\|^2 = (\sigma_2\|v\|)^2.$$

若$v = v_2$，等式成立.

三 讨论特征值与奇异值的差异：

(1) 特征值适用的矩阵?

(2) 奇异值适用的矩阵?

(3) 对给定的一个n阶方阵，其对应的特征值和奇异值相同吗?

(4) 对给定的一个实对称矩阵, 其对应的特征值和奇异值相同吗?

(5) 对给定的一个n阶方阵，对应的特征值和奇异值的共性?

提示: (1) 方阵, (2) 任何矩阵, (3) 不相同(书中扰动例子), (4) 特征值的绝对值, (5) 秩相同.

四 (1) 设$A, B \in M_{m\times n}(\mathbb{R})$, $A, B, A+B$的最大奇异值分别是r, s, t.证明$r + s \geq t$..

(2) 设$A \in M_{m\times n}(\mathbb{R}), B \in M_{n\times p}(\mathbb{R})$, A, B, AB的最大奇异值分别是r, s, t.证明$rs \geq t$.

(3) 设$A \in M_n(\mathbb{R})$,σ_1是A的最大奇异值，σ_r是最小(正)奇异值，λ是A的任意实特征值。证明$\sigma_r \leq |\lambda| \leq \sigma_1$.

题目有问题

提示：(1)由习题二，存在单位向量$v \in \mathbb{R}^n$, 使得$\|(A+B)v\| = t$，则$\|(A+B)v\| \le \|Av\| + \|Bv\| \le r+s$.

(2)若$AB \neq 0$, 存在单位向量$w \in \mathbb{R}^n$, 使得$\|ABw\| = t \neq 0$. 则$\|ABw\| = \|A(\frac{Bw}{\|Bw\|})\|\|Bw\| \le rs$.

五 设$A \in M_{m\times n}(\mathbb{R})$的奇异值分解是$A = U\Sigma V^T$, 满足$U = (u_1, \cdots, u_m)$是$m$阶正交矩阵，$V = (v_1, \cdots, v_n)$ 是n阶正交矩阵，$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}_{m\times n}$, $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$, 其中$r = rank(A)$. 求$A - \sigma_1 u_1 v_1^T$的奇异值分解.

解：$A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$. $A - \sigma_1 u_1 v_1^T =$

$$(u_2, u_3, \cdots, u_m, u_1) \begin{bmatrix} \sigma_2 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}_{m\times n} \begin{bmatrix} v_2^T \\ v_3^T \\ \vdots \\ v_n^T \\ v_1^T \end{bmatrix}$$

六 [♡] 设A是n阶非零实矩阵. 证明：A的奇异值与A的特征值相同当且仅当A是正定阵或半正定阵.

证明框架：假设A是正定阵或半正定阵. 设A的非零特征值是$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > 0$, 则$A^TA = A^2$的非零特征值是$\lambda_1^2 \ge \lambda_2^2 \ge \cdots \ge \lambda_r^2 > 0$. 从而$A$的(正)奇异值是$\sqrt{\lambda_1^2} \ge \sqrt{\lambda_2^2} \ge \cdots \ge \sqrt{\lambda_r^2} > 0$. 而且，$A$的零特征值和零奇异值均为$n-r$重.

$Au = u\cdot(v^Tu)$

反之，我们关于A的秩做归纳. 若A的秩是1，则存在非零向量$u, v \in \mathbb{R}^n$, $A = uv^T$. A的唯一正奇异值是$\|u\|\|v\|$,若它是A的特征值, 则它等于v^Tu. 应用内积不等式$|v^Tu| \le \|u\|\|v\|$, 等式成立当且仅当$u = cv$，$c > 0$. 此时$A = cv^Tv$是一个半正定阵(若$n > 1$). 假设题目结论对于秩小于r的实矩阵均成立. 现在假设A的秩等于r且它的奇异值分解是$A = U\Sigma V^T$, 满足$U = (u_1, \cdots, u_m)$ 是m 阶正交矩阵，$V = (v_1, \cdots, v_n)$ 是n 阶正交矩阵，$\Sigma = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ & & & \end{bmatrix}_{m\times n}$, $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$, 其中$r = rank(A)$. 所以，$A$的非零特征值就

是$\sigma_1,\cdots,\sigma_r$. 设$\sigma_1=\cdots=\sigma_k>\sigma_{k+1}\geq\cdots\geq\sigma_r>0,k\geq 1$. 令$A\eta_1=\sigma_1\eta_1$且$\|\eta_1\|=1$. 则$\eta_1=c_1v_1+\cdots+c_kv_k+c_{k+1}v_{k+1}+\cdots+c_nv_n$. 从而

$$\sigma_1^2=\|A\eta_1\|^2=\eta_1^TA^TA\eta_1=(c_1v_1+\cdots+c_nv_n)^T(c_1\sigma_1^2v_1+\cdots+c_r\sigma_r^2v_r)=c_1^2\sigma_1^2+\cdots+c_r^2\sigma_r^2.$$

$\leq\sigma_1^2(c_1^2+c_2^2+\cdots+c_n^2)=\sigma_1^2$

因为$\|\eta_1\|=1$, 所以$c_1^2+\cdots+c_n^2=1$. 因此, $\eta_1=c_1v_1+\cdots+c_kv_k$. 因为$v_1,\cdots,v_k$均是$A^TA$的特征向量，所以$\eta_1$是$A^TA$的单位特征向量. 因为$A^TA\eta_1=\lambda_1\eta_1$且$A\eta_1=\sigma_1\eta_1$, 我们有$A^T\eta_1=\sigma_1\eta_1$. 存在$n$阶正交阵$V_1$,满足$V_1$的第一列是$\eta_1$, 且

$\Rightarrow A^TA\eta_1=A^T\sigma_1\eta_1=\lambda_1\eta_1$

$\Rightarrow A^T\eta_1=\sigma_1\eta_1$

$$V_1^T(A^TA)V_1=\begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_r^2 \\ & & \end{bmatrix}_{n\times n}.$$

且因为$A^T\eta_1=A\eta_1=\sigma_1\eta_1$,

$$V_1^TAV_1=\begin{bmatrix}\sigma_1 & 0^T\\ 0 & A_1\end{bmatrix}.$$

$V_1^TAV_1=\begin{pmatrix}\eta_1^T\\ V_2^T\end{pmatrix}A(\eta_1\ V_2)=\begin{pmatrix}\sigma_1 & \eta_1^TAV_2\\ V_2^TA\eta_1 & V_2^TAV_2\end{pmatrix}$

$\sigma_1\eta_1^TV_2=0$

$\sigma_1V_2^T\eta_1=0$

注记$V_1^TAV_1$和A有相同的奇异值和特征值. 这暗示A_1的奇异值和特征值相同，且$rank(A_1)<r$. 我们应用关于A 的秩的归纳假设就可以证明结论.

七 设A是一个$m\times n$阶实数矩阵. 令$\tilde{A}=\begin{bmatrix}0 & A\\ A^T & 0\end{bmatrix}$. 证明：

考虑 $\frac{\tilde A}{\lambda}=\begin{pmatrix}0 & \frac{A}{\lambda}\\ \frac{A^T}{\lambda} & 0\end{pmatrix}$

$\Rightarrow (I-\frac{\tilde A}{\lambda})=\begin{pmatrix}I & -\frac{A}{\lambda}\\ -\frac{A^T}{\lambda} & I\end{pmatrix}$

$\to\begin{pmatrix}I & -\frac{A}{\lambda}\\ 0 & -\frac{A^TA}{\lambda^2}+I\end{pmatrix}$

$\Rightarrow |I-\frac{\tilde A}{\lambda}|=|I-\frac{A^TA}{\lambda^2}|$

$\Rightarrow \frac{|\lambda I-\tilde A|}{\lambda^{m+n}}=\frac{|\lambda^2-A^TA|}{\lambda^{2n}}$

(1) $|\lambda I_{m+n}-\tilde{A}|=\lambda^{m-n}|\lambda^2I_n-A^TA|$.

(2) 设A的奇异值分解是$A=U\Sigma V^T$, 满足$U=(u_1,\cdots,u_m)$是m阶正交矩阵，$V=(v_1,\cdots,v_n)$是n阶正交矩阵，$\Sigma=\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r \\ & & \end{bmatrix}_{m\times n}$, $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_r>0$, 其中$r=rank(A)$. 令$U_1,V_1$分别是$U,V$的前$r$列构成矩阵, U_2是U的后$m-r$列构成矩阵，V_2是V的后$n-r$列构成矩阵令

$$W=\begin{bmatrix}\frac{1}{\sqrt{2}}U_1 & -\frac{1}{\sqrt{2}}U_1 & U_2 & 0\\ \frac{1}{\sqrt{2}}V_1 & \frac{1}{\sqrt{2}}V_1 & 0 & V_2\end{bmatrix}.$$

则

$$W^T\tilde{A}W = \begin{bmatrix} \sigma_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \sigma_r & & & & & & \\ & & & -\sigma_1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -\sigma_r & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{bmatrix}$$

注：本题展示一般实对称阵特征值拥有的性质可以转换到任何实矩阵的奇异值上，例如：两实对称阵和的特征值的范围（见习题课九第八题）

证明：(1) 分块矩阵行列式.

(2) 直接验证.

八 设A是一个$m\times n$阶实矩阵，$b\in\mathbb{R}^m$. 证明：$Ax=b$的全部最小二乘解的集合是

$$\{A^+b+(I_n-A^+A)y \mid y\in\mathbb{R}^n\}.$$

$x\in N(A)$
$\Rightarrow (I-A^+A)x=x$
$\Rightarrow x\in (I_n-A^+A)y$

展示A^+b是长度最小的最小二乘解.

证明：因为$A[(I_n-A^+A)y]=0$, 所以$(I_n-A^+A)y\in N(A)$且垂直于A^+b. 我们有

$$\|A^+b+(I_n-A^+A)y\|^2=\|A^+b\|^2+\|(I_n-A^+A)y\|^2\geq\|A^+b\|^2.$$

九 [♡] 设A是$m\times n$阶阵，B是$p\times q$阶阵，C是$m\times q$阶阵. 考虑矩阵方程$AXB=C$，其中X是$n\times p$阶未知矩阵. 假设$rank(A,C)=rank(A)$且$rank\begin{bmatrix}B\\C\end{bmatrix}=rank(B)$. 证明：给定任意$n\times p$阶矩阵$W$，$A^+CB^++W-A^+AWBB^+$是$AXB=C$的解.

证明：因为题目秩的关系, 存在矩阵Y,Z，使得$AY=C,ZB=C$. 因而

$$C=AY=AA^+AY=AA^+C.$$

同理，

$$C=ZB=ZBB^+B=CB^+B.$$

(注记：以上讨论可以简化：C的列属于$C(A)$和$C(B^T)$且AA^+,B^+B分别是这两个空间上投影矩阵.) 我们有$C=AA^+CB^+B$，即A^+CB^+是$AXB=C$的解. 因为I_n-A^+A是$N(A)$上的投影矩阵，所以

$$A(W-A^+AWBB^+)B=A(WB-A^+AWB)=A(I_n-A^+A)WB=0.$$

实际上，$A^+CB^++W-A^+AWBB^+$给出了$AXB=C$的全部解. 设M是$AXB=C$的任意一个解, 则$A(M-A^+CB^+)C=0$. 令$W=M-A^+CB^+$, 则$M=A^+CB^++N=A^+CB^++W-A^+AWBB^+$.

十 Moore-Penrose(M-P)广义逆的定义: 设$A\in\mathscr{M}_{m,n}(\mathbb{R})$, 若存在$M\in\mathscr{M}_{n,m}(\mathbb{R})$满足(1) $AMA=A$, (2) $MAM=M$, (3) AM 与MA均为对称矩阵, 则称M为A的M-P广义逆, 记成A^+.

(1) M-P广义逆存在性. 对奇异值分解$AV=U\Sigma$, $V=(V_1,V_2)$和$U=(U_1,U_2)$为正交矩阵, $\Sigma=\begin{pmatrix}\Sigma_r & 0\\ 0 & 0\end{pmatrix}$, Σ_r所有正奇异值, 且$A=U_1\Sigma_rV_1^T$, 证明: $V_1\Sigma_r^{-1}U_1^T$是一个M-P广义逆.

(2) ($\heartsuit$) 证明: A^+唯一.

(3) 证明: $x=V_1\Sigma_r^{-1}U_1^Tb$为$Ax=b$的一个解, 由此得到满足M-P的广义逆A^+对应的$x=A^+b$为$Ax=b$的一个解.

(4) 一般的广义逆A^-定义为满足$AA^-A=A$, 谈谈你的认识.

证明： (1)直接验证, (2) 设广义逆A_1^+和A_2^+.

$$A_1^+=A_1^+AA_1^+=A_1^+(AA_2^+A)A_1^+=(A_1^+A)^T(A_2^+A)^TA_1^+=(A_2^+AA_1^+A)^TA_1^+=(A_2^+A)^TA_1^+=A_2^+AA_1^+,$$

同理

$$A_2^+AA_1^+=A_2^+(AA_2^+A)A_1^+=A_2^+(AA_1^+AA_2^+)^T=A_2^+(AA_2^+)^T=A_2^+AA_2^+=A_2^+.$$

由此得到$A_1^+=A_2^+$.

(3)直接验证并利用(2)得唯一性.

(4)广义逆A^-一般不唯一，例如A列满秩，A的左逆一般不唯一. 若$Ax=b$有解，则A^-b是一个特解. AA^- 是$C(A)$ 上投影矩阵.