强化练习六（不是作业）

1. 给定以下矩阵

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 3 & 2 \end{bmatrix}$$

(i) 试求方程组$Ax = b$的解集,其中 $b = [3, 3, 6]^T$.

(ii) 试求A的行空间、零空间。

(iii) 试求最小二乘问题$\inf_{x\in\mathbb{R}^3} |Ax - e_1|$的解集。试求解集中长度最小的向量。

解. *(i)* 我们用增广矩阵来求解线性方程组$Ax = b$的解集。

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 2 & 6 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -4 & -3 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

因此解集为

$$\begin{bmatrix} -3 \\ 3 \\ 0 \end{bmatrix} + \{x : x = \begin{bmatrix} 4x_3 \\ -2x_3 \\ x_3 \end{bmatrix}, x_3 \in \mathbb{R}\} = \begin{bmatrix} -3 \\ 3 \\ 0 \end{bmatrix} + span(\begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}).$$

(ii) 由第一小问的计算结果，我们知道零空间为

$$N(A) = span(\begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}).$$

行空间为

$$R(A^T) = span(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}) = span(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}).$$

(iii) 求解方法有多种。为了顺便帮助大家复习奇异值分解，此处我们用奇异值分解来做。我们先计算A^TA的谱分解.

$$A^TA = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 1 & 3 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 2 \\ 5 & 14 & 8 \\ 2 & 8 & 8 \end{bmatrix}$$

因此A^TA的特征多项式为

$$p_{A^TA}(\lambda) = det(\begin{bmatrix} \lambda-2 & -5 & -2 \\ -5 & \lambda-14 & -8 \\ -2 & -8 & \lambda-8 \end{bmatrix})$$

$$= (\lambda-2)\big((\lambda-14)(\lambda-8)-64\big)+5\big(-5(\lambda-8)-16\big)-2\big(40+2(\lambda-14)\big)$$

$$= (\lambda-2)\big(\lambda^2-22\lambda+48\big)+5\big(-5\lambda+24\big)-2\big(2\lambda+12\big)$$

$$= \lambda^3-24\lambda^2+92\lambda-96-25\lambda+120-4\lambda-24$$

$$= \lambda^3-24\lambda^2+63\lambda = \lambda(\lambda-3)(\lambda-21).$$

因此特征值为$0, 3, 21$. 相应地我们求解A^TA特征值为0的特征向量

$$N(A^TA) = N\Big(\begin{bmatrix} 2 & 5 & 2 \\ 5 & 14 & 8 \\ 2 & 8 & 8 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} 2 & 5 & 2 \\ 5 & 14 & 8 \\ 0 & 3 & 6 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} 1 & 5/2 & 1 \\ 0 & 3/2 & 3 \\ 0 & 1 & 2 \end{bmatrix}\Big)$$

$$= N\Big(\begin{bmatrix} 1 & 5/2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}\Big) = span(\begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix})$$

接下来我们求解A^TA特征值为3的特征向量,

$$N(A^TA-3I_3) = N\Big(\begin{bmatrix} -1 & 5 & 2 \\ 5 & 11 & 8 \\ 2 & 8 & 5 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} -1 & 5 & 2 \\ 0 & 36 & 18 \\ 0 & 18 & 9 \end{bmatrix}\Big)$$

$$= N\Big(\begin{bmatrix} -1 & 5 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} -1 & 0 & -1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}\Big)$$

$$= span(\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix})$$

接下来我们求解A^TA特征值为21的特征向量,

$$N(A^TA-21I_3) = N\Big(\begin{bmatrix} -19 & 5 & 2 \\ 5 & -7 & 8 \\ 2 & 8 & -13 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} 1 & 4 & -13/2 \\ -19 & 5 & 2 \\ 5 & -7 & 8 \end{bmatrix}\Big)$$

$$= N\Big(\begin{bmatrix} 1 & 4 & -13/2 \\ 0 & 81 & -243/2 \\ 0 & -27 & 81/2 \end{bmatrix}\Big) = N\Big(\begin{bmatrix} 1 & 4 & -13/2 \\ 0 & 1 & -3/2 \\ 0 & 0 & 0 \end{bmatrix}\Big)$$

$$= N\Big(\begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -3/2 \\ 0 & 0 & 0 \end{bmatrix}\Big) = span(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix})$$

因此A^TA的谱分解为

$$A^TA = Q\begin{bmatrix}21 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 0\end{bmatrix}Q^T,$$

其中

$$Q := [q_1, \cdots, q_n] := \begin{bmatrix}1/\sqrt{14} & -1//\sqrt{6} & 4/\sqrt{21}\\ 3/\sqrt{14} & -1/\sqrt{6} & -2/\sqrt{21}\\ 2/\sqrt{14} & 2/\sqrt{6} & 1/\sqrt{21}\end{bmatrix}$$

令

$$w_1 := Aq_1 = \frac{1}{\sqrt{14}}\begin{bmatrix}1 & 2 & 0\\ 0 & 1 & 2\\ 1 & 3 & 2\end{bmatrix}\begin{bmatrix}1\\ 3\\ 2\end{bmatrix} = \frac{1}{\sqrt{14}}\begin{bmatrix}7\\ 7\\ 14\end{bmatrix}, \quad \tilde{w}_1 := \frac{w_1}{|w_1|} = \begin{bmatrix}1/\sqrt{6}\\ 1/\sqrt{6}\\ 2/\sqrt{6}\end{bmatrix}$$

$$w_2 := Aq_2 = \frac{1}{\sqrt{6}}\begin{bmatrix}1 & 2 & 0\\ 0 & 1 & 2\\ 1 & 3 & 2\end{bmatrix}\begin{bmatrix}-1\\ -1\\ 2\end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix}-3\\ 3\\ 0\end{bmatrix}, \quad \tilde{w}_2 := \frac{w_2}{|w_2|} = \begin{bmatrix}-1/\sqrt{2}\\ 1/\sqrt{2}\\ 0\end{bmatrix}$$

因此A的简化奇异值分解为

$$A = U_2\Sigma V_2^T, \quad U_2 := [\tilde{w}_1, \tilde{w}_2], \quad V_2 := [q_1, q_2], \quad \Sigma := \begin{bmatrix}\sqrt{21} & 0\\ 0 & \sqrt{3}\end{bmatrix}.$$

因此最小二乘问题$\inf_{x\in\mathbb{R}^3}|Ax - e_1|$的解集中长度最小的向量为

$$x = V_2\Sigma^{-1}U_2^Te_1$$

$$= \begin{bmatrix}1/\sqrt{14} & -1//\sqrt{6}\\ 3/\sqrt{14} & -1/\sqrt{6}\\ 2/\sqrt{14} & 2/\sqrt{6}\end{bmatrix}\begin{bmatrix}1/\sqrt{21} & 0\\ 0 & 1/\sqrt{3}\end{bmatrix}\begin{bmatrix}1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6}\\ -1/\sqrt{2} & 1/\sqrt{2} & 0\end{bmatrix}e_1$$

$$= \begin{bmatrix}1/\sqrt{14} & -1//\sqrt{6}\\ 3/\sqrt{14} & -1/\sqrt{6}\\ 2/\sqrt{14} & 2/\sqrt{6}\end{bmatrix}\begin{bmatrix}1/\sqrt{126}\\ -1/\sqrt{6}\end{bmatrix} = \begin{bmatrix}\frac{1}{14\sqrt{9}} + \frac{1}{6}\\ \frac{3}{14\sqrt{9}} + \frac{1}{6}\\ \frac{2}{14\sqrt{9}} - \frac{1}{3}\end{bmatrix} = \begin{bmatrix}\frac{1}{42} + \frac{1}{6}\\ \frac{1}{14} + \frac{1}{6}\\ \frac{1}{21} - \frac{1}{3}\end{bmatrix}$$

$$= \begin{bmatrix}\frac{4}{21}\\ \frac{5}{21}\\ \frac{-2}{7}\end{bmatrix}$$

□

2. 任意给定$\mathbb{R}^n$中的子空间$\mathcal{M}_1, \mathcal{M}_2$, $\mathbb{R}^m$中的子空间$\mathcal{N}_1, \mathcal{N}_2$.
   (i) 是否存在矩阵A使得，$N(A) = \mathcal{M}_1, R(A^T) = \mathcal{M}_2, R(A) = \mathcal{N}_1, N(A^T) = \mathcal{N}_2$.
   (ii) 试给出存在满足上述性质的矩阵A的充分必要条件。

解. 因为$R(A^T) = (N(A))^{\perp}$, $R(A) = (N(A^T))^{\perp}$， 所以若$\mathcal{M}_2 \neq \mathcal{M}_1^{\perp}$则不一定存在。

- $N(A) = \mathcal{M}_1, R(A^T) = \mathcal{M}_2, R(A) = \mathcal{N}_1, N(A^T) = \mathcal{N}_2, \Longrightarrow \mathcal{M}_2 = \mathcal{M}_1^{\perp}, \mathcal{N}_2 = \mathcal{N}_1^{\perp}$, 且$dim(\mathcal{M}_1) + dim(\mathcal{N}_1) = n$.

因为$R(A^T) = (N(A))^{\perp}$, $R(A) = (N(A^T))^{\perp}$，并且$rank(A) = rank(A^T)$所以结论成立。

- $\mathcal{M}_2 = \mathcal{M}_1^{\perp}, \mathcal{N}_2 = \mathcal{N}_1^{\perp}$, 且$dim(\mathcal{M}_1) + dim(\mathcal{N}_1) = n$. $\Longrightarrow$ 存在矩阵A使得$N(A) = \mathcal{M}_1, R(A^T) = \mathcal{M}_2, R(A) = \mathcal{N}_1, N(A^T) = \mathcal{N}_2$.

由于$\mathcal{M}_2 = \mathcal{M}_1^{\perp}, \mathcal{N}_2 = \mathcal{N}_1^{\perp}$, 且$dim(\mathcal{M}_1) + dim(\mathcal{N}_1) = n$， 所以

$$dim(\mathcal{M}_2) = dim(\mathcal{N}_1).$$

记$k = dim(\mathcal{M}_2) = dim(\mathcal{N}_1)$. 令$\{u_1, \cdots, u_k\}$是子空间$\mathcal{M}_2$的一组基，$\{v_1, \cdots, v_k\}$是子空间$\mathcal{N}_1$的一组基. 同时我们分别将这两组子空间的基扩张至全空间的基$\{u_1, \cdots, u_n\}$, $\{v_1, \cdots, v_n\}$. 则

$$\mathcal{M}_1 := span(u_{k+1}, \cdots, u_n), \quad \mathcal{N}_2 := span(v_{k+1}, \cdots, v_n).$$

构造线性映射$\mathcal{A} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$使得

$$\forall i \in \{1, \cdots, k\}, \quad \mathcal{A}(u_i) = v_i, \quad i \in \{k+1, \cdots, m\}, \quad \mathcal{A}(u_i) = 0.$$

记线性映射$\mathcal{A} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$的表示矩阵为$A$. 则

$$N(A) = \mathcal{M}_1, \quad R(A) = \mathcal{N}_1, \Longrightarrow R(A^T) = \big(N(A)\big)^{\perp} = \mathcal{M}_2.$$

$$N(A^T) = (R(A))^{\perp} = \mathcal{N}_2.$$

□

3. 试求以下矩阵的行列式

$$A = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1(n-1)} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2(n-1)} & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & \ddots & a_{(n-1)(n-1)} & 0 \\ 1 & 0 & \cdots & 0 & a_{nn} \end{bmatrix}$$

解. 我们按照第一列展开，得到

$$det(A) = a_{11} det(A_{11}) + (-1)^{n+1} det(A_{n1})$$

注意到A_{11}是上三角矩阵因此

$$det(A_{11}) = \prod_{i=2}^{n} a_{ii}$$

注意到A_{n1}是如下矩阵。

$$A_{n1} = \begin{bmatrix} a_{12} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{22} & \cdots & a_{2(n-1)} & 0 \\ \ddots & \ddots & \vdots & \vdots \\ \cdots & \ddots & a_{(n-1)(n-1)} & 0 \end{bmatrix}$$

$$\det(A_{n1}) = (-1)^n a_{1n} \prod_{i=2}^{n-1} a_{ii}.$$

综上所述，

$$det(A) = \prod_{i=1}^{n} a_{ii} - a_{1n} \prod_{i=2}^{n-1} a_{ii}.$$

□

4. (i) 判断以下矩阵是否可以对角化

$$A := \begin{bmatrix} 10 & 1 & 0 \\ 0 & 10 & 1 \\ 0.001 & 0 & 10 \end{bmatrix}$$

(ii) 判断以下矩阵是否可以同时被对角化

$$A := \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 1 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

解. *(i)* 注意到

$$p_A(\lambda) = det(\begin{bmatrix} \lambda - 10 & -1 & 0 \\ 0 & \lambda - 10 & -1 \\ -0.001 & 0 & \lambda - 10 \end{bmatrix}) = (\lambda - 10)^3 - 0.001$$

因此A有三个特征值分别为$10 + 0.1, 10 + 0.1e^{i2\pi/3}, 10 + 0.1e^{i4\pi/3}$. 由于$A$有三个不同的特征值，所以$A$的特征值的代数重数和几何重数都是1, 因此可以对角化。

(ii) 注意到 A, B均可对角化。是否可以同时被对角化等价于AB是否等于BA. 注意到

$$AB = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 12 \\ 0 & 8 & 0 \\ 0 & 0 & 15 \end{bmatrix}$$

$$BA = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 8 \\ 0 & 8 & 0 \\ 0 & 0 & 15 \end{bmatrix}$$

因为$AB \neq BA$, 所以A, B不可以同时被对角化。

$\square$

5. 求以下极值问题的解

$$\sup_{x_1^2+x_2^2+4x_3^2=1} 2x_1^2 - x_2^2 - 4x_3^2 - 2x_1x_2 - 4x_1x_3.$$

解. 令

$$u_1 = x_1, \quad u_2 = x_2, \quad u_3 = 2x_3.$$

则

$$\begin{aligned} &\sup_{x_1^2+x_2^2+4x_3^2=1} 2x_1^2 - x_2^2 - 4x_3^2 - 2x_1x_2 - 4x_1x_3 \\ &= \sup_{u\in\mathbb{R}^3, |u|=1} 2u_1^2 - u_2^2 - u_3^2 - 2u_1u_2 - 2u_1u_3. \end{aligned}$$

注意到

$$2u_1^2 - u_2^2 - u_3^2 - 2u_1u_2 - 2u_1u_3 = u^T A u, \quad \Longrightarrow A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & -1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

接下来我们求解A的特征值。按照第三行展开，我们有

$$\begin{aligned} p_A(\lambda) = det(\lambda I_3 - A) = det(&\begin{bmatrix} \lambda-2 & 1 & 1 \\ 1 & \lambda+1 & 0 \\ 1 & 0 & \lambda+1 \end{bmatrix}) \\ &= (\lambda+1)\big((\lambda-2)(\lambda+1)-1\big) - (\lambda+1) \\ &= (\lambda+1)\big(\lambda^2-\lambda-4\big). \end{aligned}$$

因此A第特征值为$-1, (1+\sqrt{17})/2, (1-\sqrt{17})/2$. 由于$A$是实对称矩阵，所以$A$正交相似于一个对角矩阵

$$\begin{aligned} A = Q\Lambda Q^T, \quad \Lambda = diag((1+\sqrt{17})/2, (1-\sqrt{17})/2, -1) \\ \Longrightarrow \sup_{u\in\mathbb{R}^3, |u|=1} u^T A u = \sup_{u\in\mathbb{R}^3, |u|=1} u^T Q\Lambda Q^T u \\ = \sup_{v\in\mathbb{R}^3, |v|=1} v^T \Lambda v = (1+\sqrt{17})/2. \end{aligned}$$

注意到，在上述等式中，我们用到了正交矩阵保模长的性质。

$\square$

6. 给定任意n阶实矩阵A. 若$A^TA = A^2$，试证明$A = A^T$.
注：此题难度较大，非基础题。

证明. 记$\lambda_1, \cdots, \lambda_k \in \mathbb{C}$为所有$A$的不同特征值。首先我们证明特征值均为实数， **若存在非实数的特征值**$\lambda \in \mathbb{C}/\mathbb{R}$, 则$\lambda^2$是$A^2$的特征值，从而也是$A^TA$的特征值。 因为实对称矩阵$A^TA$的特征值为非负的实数，所以$\lambda$一定为实数（$\lambda^2$为实数，则$\lambda$为实数或者纯虚数。又因为$\lambda^2$非负，所以$\lambda$为实数）。 **矛盾**。

接下来我们证明不同特征值所对应的特征子空间$N(A-\lambda_i I_n)$相互正交. 对

$$\forall i, j \in \{1, \cdots, k\}, i \neq j, u \in N(A-\lambda_i I_n), v \in N(A-\lambda_j I_n),$$

我们有

$$\langle Au, Av\rangle = \lambda_i\lambda_j\langle u, v\rangle, \quad \langle Au, Av\rangle = \langle u, A^TAv\rangle = \langle u, A^2v\rangle = {\lambda_j}^2\langle u, v\rangle.$$
$$\langle Av, Au\rangle = {\lambda_i}^2\langle v, u\rangle, \quad \Longrightarrow \langle Au, Av\rangle = \lambda_i^2\langle u, v\rangle.$$

若$\langle u, v\rangle \neq 0$, 则从上面的等式我们可以推导出

$$\Longrightarrow \lambda_i\lambda_j = {\lambda_j}^2 = \lambda_i^2, \quad \Longrightarrow \lambda_i = \lambda_j.$$

矛盾。 因此$\langle u, v\rangle = 0$。

因为$N(A) = N(A^TA) = N(A^2)$, 所以A的特征值为零的$Jordan$链长度均为1。 接下来我们证明任何A的非零的特征值λ的 $Jordan$链长度也为1。若$Jordan$链长度不为1, 则存在一个非零的向量$u \in \mathbb{C}^n$使得

$$(A-\lambda I_n)u \neq 0, \quad (A-\lambda I_n)^2u = 0.$$

因此

$$\langle (A-\lambda I_n)^2u, u\rangle + \langle u, (A-\lambda I_n)^2u\rangle = 0$$

又因为

$$\langle (A-\lambda I_n)^2u, u\rangle = \langle A^2u, u\rangle - 2\lambda\langle Au, u\rangle + \lambda^2\langle u, u\rangle$$
$$= \langle A^TAu, u\rangle - 2\lambda\langle Au, u\rangle + \lambda^2\langle u, u\rangle = \langle Au, Au\rangle - 2\lambda\langle Au, u\rangle + \lambda^2\langle u, u\rangle$$
$$\Longrightarrow 0 = \langle (A-\lambda I_n)^2u, u\rangle + \langle u, (A-\lambda I_n)^2u\rangle$$
$$= 2\big(\langle Au, Au\rangle - \lambda\langle Au, u\rangle - \lambda\langle u, Au\rangle + \lambda^2\langle u, u\rangle\big)$$
$$= 2\langle (A-\lambda I_n)u, (A-\lambda I_n)u\rangle, \quad \Longrightarrow (A-\lambda I_n)u = 0.$$

矛盾。 所以我们知道A对应的所有$Jordan$链长度均为1.

最后我们证明$\mathbb{C}^n = N(A-\lambda_1 I_n) + \cdots + N(A-\lambda_k I_n)$. 此处重复讲义177页引理12.5的证明即可.

综上所述，选取所有子空间$N(A-\lambda_i I_n)$的标准正交基后取并集生成了全空间的标准正交基。因此存在正交矩阵Q和对角矩阵Λ使得

$$A = Q\Lambda Q^T.$$

因此

$$A^T = (Q\Lambda Q^T)^T = Q\Lambda^T Q^T = Q\Lambda Q^T = A.$$

证毕。

□