

线性代数 (I) 期末练习题参考解答

1. 给定方阵 $A = \begin{pmatrix} 2 & 0 & 0 \\ a & 2 & 0 \\ b & c & -1 \end{pmatrix}$. 证明: A 可对角化当且仅当 $a = 0$.

证明: 首先, A 是下三角矩阵, 它的特征值是 -1 (一重), 2 (二重). 因此 A 可对角化当且仅当 $\dim \ker(A - 2E) = 2$, 这等价于 $\text{rank}(A - 2E) = 1$. 注意到

$$A - 2E = \begin{pmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ b & c & -3 \end{pmatrix},$$

从而 $\text{rank}(A - 2E) = 1$ 当且仅当 $a = 0$. 因此 A 可对角化当且仅当 $a = 0$.

2. 设 $A = \begin{pmatrix} -2 & 2 & -2 \\ -1 & 1 & -1 \\ -2 & 2 & -2 \end{pmatrix}$, 请把 A^5 写成 I, A, A^2 的线性组合.

解: 首先计算 A 的特征多项式 $\chi_A(t) = t^2(t + 3)$. 注意到 $\text{rank} A = 1$, 因此 A 的极小多项式是 $\mu_A(t) = t(t + 3)$. 用 t^5 对 $t^2 + 3t$ 作带余除法, 得到

$$t^5 = (t^2 + 3t)q(t) + at + b.$$

令 $t = 0$, 则 $b = 0$. 令 $t = -3$, 得到 $a = 81$. 于是我们得到 $A^5 = 81A$.

注意: 由 A 的极小多项式是 $\mu_A(t) = t(t + 3)$ 知道 $A^2 = -3A$, 故此题的答案并不唯一, 如 $A^5 = -27A^2$.

3. 给定平面上三个点 $(1, 7)$, $(3, 3)$, $(6, 1)$, 用最小二乘法求出拟合这三个点的曲线 $y = a + \frac{b}{x}$.

解: 代入三个点的数据, 知道

$$a + b = 7, \quad a + \frac{b}{3} = 3, \quad a + \frac{b}{6} = 1.$$

令 $A = \begin{pmatrix} 1 & 1 \\ 1 & \frac{1}{3} \\ 1 & \frac{1}{6} \end{pmatrix}$, $B = [7, 3, 1]$. 上述方程三个方程可以记成 $AX = B$, 其中 $X = [a, b]$. 考虑相应的正规方程 ${}^tAAX = {}^tAB$, 得到

$$\begin{pmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & \frac{41}{36} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 11 \\ \frac{49}{6} \end{pmatrix}.$$

解得 $a = \frac{5}{21}$, $b = \frac{48}{7}$. 从而拟合出的曲线是 $y = \frac{5}{21} + \frac{48}{7x}$.

4. 设复方阵 A 的特征多项式是 $\chi_A(t) = t^5(t-1)^4(t-2)^4(t-3)^3$, 极小多项式是 $\mu_A(t) = t^3(t-1)^2(t-2)^3(t-3)$, 求 A 的约当标准型.

解: 稍加分析, 由 A 的极小多项式就知道 A 的约当标准型中特征值为 0 的约当块阶数最大是 3, 特征值为 1 的约当块阶数最大是 2, 特征值为 2 的约当块阶数最大是 3, 特征值为 3 的约当块阶数最大是 1. 再结合特征多项式就知道 A 的约当标准型有如下的可能:

$$J(A) = 2J_1(0) \dot{+} J_3(0) \dot{+} 2J_1(1) \dot{+} J_2(1) \dot{+} J_1(2) \dot{+} J_3(2) \dot{+} 3J_1(3);$$

$$J(A) = 2J_1(0) \dot{+} J_3(0) \dot{+} 2J_2(1) \dot{+} J_1(2) \dot{+} J_3(2) \dot{+} 3J_1(3);$$

$$J(A) = J_2(0) \dot{+} J_3(0) \dot{+} 2J_1(1) \dot{+} J_2(1) \dot{+} J_1(2) \dot{+} J_3(2) \dot{+} 3J_1(3);$$

$$J(A) = J_2(0) \dot{+} J_3(0) \dot{+} 2J_2(1) \dot{+} J_1(2) \dot{+} J_3(2) \dot{+} 3J_1(3).$$

5. 设 $A = \frac{1}{2} \begin{pmatrix} 1 & -\mathbf{i} & -1+\mathbf{i} \\ \mathbf{i} & 1 & 1+\mathbf{i} \\ 1+\mathbf{i} & -1+\mathbf{i} & 0 \end{pmatrix}$ 是一个酉矩阵, 求酉矩阵 B 使得 $B^{-1}AB$ 是对角阵.

解: 此题的计算过程见教材 243 页的例 6.29.

6. (1) 把 \mathbb{R}^4 中如下的正交向量组扩充为 \mathbb{R}^4 的标准正交基

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$$

(2) 设 U, V 是 \mathbb{R}^n 的子空间满足 $\mathbb{R}^n = U \oplus V$. 证明:

$$\mathbb{R}^n = U^\perp \oplus V^\perp.$$

解: (1) 简单计算, 很容易看出添加向量组 $(1, -1, 1, -1)$ 和 $(1, -1, -1, 1)$ 得到 \mathbb{R}^4 的一个正交基. 对这些向量进行单位化, 得到 \mathbb{R}^4 的标准正交基为

$$\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}\right), \left(\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right).$$

(2) 由 $\mathbb{R}^n = U \oplus V$ 知道 $\dim U + \dim V = n$. 因为 $\dim U^\perp = n - \dim U$, $\dim V^\perp = n - \dim V$. 因此 $\dim U^\perp + \dim V^\perp = n$. 从而要证明 $\mathbb{R}^n = U^\perp \oplus V^\perp$, 我们只需要证明 $U^\perp \cap V^\perp = 0$ 即可. 任取 $\mathbf{w} \in U^\perp \cap V^\perp$, 则对于任意的 $\mathbf{u} \in U, \mathbf{v} \in V$, 有 $(\mathbf{w}|\mathbf{u}) = (\mathbf{w}|\mathbf{v}) = 0$. 因为 $\mathbb{R}^n = U \oplus V$, \mathbb{R}^n 中的向量都是 $\mathbf{u} + \mathbf{v}$ 的形式, 从而 $\mathbf{w} \in (\mathbb{R}^n)^\perp = 0$, 即有 $\mathbf{w} = 0$. 这就说明 $U^\perp \cap V^\perp = 0$, 命题得证.

7. 设 $A = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & 2 & 2 & 1 \\ -1 & 0 & 0 & 0 \\ 1 & -1 & -1 & -1 \end{pmatrix}$.

(1) 求 A 的约当标准型;

(2) 求可逆矩阵 C , 使得 $C^{-1}AC$ 是 A 的约当标准型.

解: (1) 首先计算 A 的特征多项式是 $\chi_A(t) = t^2(t-1)^2$. 通过计算 $\text{rank} A = 3$ 以及 $\text{rank}(A - E) = 3$. 我们可以知道 A 的约当标准型只能是

$$J(A) = J_2(0) \dot{+} J_2(1) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(2) 为了求出 C , 首先我们求出向量 $X \in \ker A^2 \setminus \ker A$, 通过计算, 可以选取 $X = [1, -1, 0, 2]$, 此时 $AX = [0, 1, -1, 0]$. 另外, 我们求出 $Y \in \ker(A - E)^2 \setminus \ker(A - E)$, 可以选取 $Y = [1, -2, 0, 1]$, 此时 $(A - E)Y = [1, 0, -1, 1]$. 因此, 我们求出的矩阵

$$C = (AX, X, (A - E)Y, Y) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & -2 \\ -1 & 0 & -1 & 0 \\ 0 & 2 & 1 & 1 \end{pmatrix}$$

满足 $C^{-1}AC = J(A)$.

8. (1) 记 $M_n(\mathbb{R})$ 是 n 阶实方阵形成的向量空间. 对于 $A = (a_{ij}) \in M_n(\mathbb{R})$, 它的迹定义为 $\text{tr}(A) = \sum_{i=1}^n a_{ii}$. 证明: $\text{tr} : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ 是一个线性映射, 即, 对于任意的 $\lambda \in \mathbb{R}$, $A, B \in M_n(\mathbb{R})$, 有

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B), \quad \text{tr}(\lambda A) = \lambda \text{tr}(A).$$

(2) 设 A, B 是两个同阶的实方阵, 设 A 的极小多项式次数是 k , 设 B 的极小多项式次数是 l . 如果对于任意的 $i \leq k + l - 1$, 有 $\text{tr}(A^i) = \text{tr}(B^i)$. 证明 $\text{tr}(A^i) = \text{tr}(B^i), \forall i \in \mathbb{N}$.

证明: (1) 根据定义直接验证是容易的。

(2) 记 A 的极小多项式是 $\mu_A(t)$, B 的极小多项式 $\mu_B(t)$. 考虑多项式 $f(t) = \mu_A(t)\mu_B(t)$, 则多项式 $f(t)$ 的次数是 $m = k + l$, 且满足 $f(A) = f(B) = 0$. 若记 $f(t) = t^m + c_{m-1}t^{m-1} + \cdots + c_1t + c_0$. 则有

$$A^m = -(c_{m-1}A^{m-1} + \cdots + c_1A + c_0E),$$

$$B^m = -(c_{m-1}B^{m-1} + \cdots + c_1B + c_0E).$$

因此不难看出, 当 $i \geq m$ 时, A^i 和 B^i 可以分别表示成相同系数的关于 $E, A, A^2, \dots, A^{m-1}$ 以及 $E, B, B^2, \dots, B^{m-1}$ 的线性组合. 由条件, 当 $i \leq k + l - 1$ 时, 有 $\text{tr}(A^i) = \text{tr}(B^i)$. 根据第一问的结论, 我们得到 $\text{tr}(A^i) = \text{tr}(B^i), \forall i \in \mathbb{N}$.

9. 设 A 和 B 是 n 阶的实对称方阵, 且 A 正定.

(1) 证明: 存在可逆方阵 P 使得 ${}^tPAP = E$ 且 ${}^tPBP = D$, 其中 D 是对角方阵.

(2) 设 X 是 \mathbb{R}^n 中的列向量, 证明:

$$\max_{X \neq 0} \frac{{}^tXBX}{{}^tXAX} = \lambda_{\max}(A^{-1}B),$$

其中 $\lambda_{\max}(A^{-1}B)$ 表示 $A^{-1}B$ 的最大特征值.

证明: (1) 由 A 正定, 我们知道存在可逆矩阵 Q , 使得 ${}^tQAQ = E$. 此时, 我们考虑 tQBQ , 这是一个对称矩阵. 因此, 存在一个正交矩阵 T , 使得 ${}^tT{}^tQBQT = D$ 是对角矩阵, 我们令 $P = QT$. 则 ${}^tPAP = E$ 且 ${}^tPBP = D$.

(2) 令 $X = PY$, 则

$$\max_{X \neq 0} \frac{{}^tXBX}{{}^tXAX} = \max_{Y \neq 0} \frac{{}^tYDY}{{}^tYY} = \max_{\|Y\|=1} {}^tYDY.$$

简单计算得到 $A^{-1}B = D$. 记 $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 其中 $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = \lambda_{\max}(A^{-1}B)$. 于是

$${}^tYDY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2 \leq \lambda_n (y_1^2 + \dots + y_n^2) = \lambda_n.$$

令 $y_1 = y_2 = \dots = y_{n-1} = 0$, $y_n = 1$, 则 ${}^tYDY = \lambda_n$, 恰好取得最大值 $\lambda_{\max}(A^{-1}B)$. 综上, (2) 得证。