

0.1 有理标准型的几何与应用

回顾有理标准型和循环子空间相关理论.

命题 0.1

设 n 阶矩阵 A 有 n 个不同的特征值, 求证: A 的特征多项式和极小多项式相等.

证明 证法一: 设 A 的 n 个不同的特征值为 $\lambda_1, \lambda_2, \dots, \lambda_n$, 则由推论??可知, 特征多项式 $f(\lambda)$ 和极小多项式 $m(\lambda)$ 有相同的根 (不计重数), 因此

$$f(\lambda) = m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

证法二: 由于 A 有 n 个不同的特征值, 故 A 相似于对角矩阵. 又因为相似矩阵有相同的特征多项式和极小多项式, 所以只要对对角矩阵证明此结论即可. 设 $A = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 则 $\lambda I_n - A = \text{diag}\{\lambda - \lambda_1, \lambda - \lambda_2, \dots, \lambda - \lambda_n\}$, 这是一个主对角元素两两互素的对角矩阵, 由 λ -矩阵和初等因子的基本性质 (1) 以及数学归纳法可知其法式为 $\text{diag}\{1, \dots, 1, (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)\}$. 因此, A 的特征多项式和极小多项式相等. \square

命题 0.2

设 n 阶矩阵 A 有 n 个不同的特征值, 且特征值 λ_i 对应的特征向量为 α_i , 由推论??可知 $\{\alpha_1, \dots, \alpha_n\}$ 为 \mathbb{C}^n 的一组基. 则 $\alpha = \alpha_1 + \dots + \alpha_n$ 是 A 的循环空间 \mathbb{C}^n 的循环向量, 即 $\mathbb{C}^n = L(\alpha, A\alpha, \dots, A^{n-1}\alpha) = C(A, \alpha)$ 为循环空间, α 是循环向量.

证明 事实上, 由 $A^k \alpha = \lambda_1^k \alpha_1 + \dots + \lambda_n^k \alpha_n$, 利用 Vandermonde 行列式容易证明 $\{\alpha, A\alpha, \dots, A^{n-1}\alpha\}$ 是 \mathbb{C}^n 的一组基, 从而 $\mathbb{C}^n = L(\alpha, A\alpha, \dots, A^{n-1}\alpha) = C(A, \alpha)$ 为循环空间, α 是循环向量. \square

命题 0.3

设数域 \mathbb{K} 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda) \cdots P_k(\lambda)$, 其中 $P_i(\lambda) (1 \leq i \leq k)$ 是 \mathbb{K} 上互异的首一不可约多项式. 求证: A 的有理标准型只有一个 Frobenius 块, 并且 A 在复数域上可对角化.

注 我们也可以利用定理??和初等因子证明这个命题. 若利用不变因子在基域扩张下的不变性, 则这个命题也可由命题 0.1 得到.

证明 设 A 的不变因子组为 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$, 其中 $d_i(\lambda) \mid d_{i+1}(\lambda), i = 1, 2, \dots, n-1$, 则有

$$f(\lambda) = P_1(\lambda)P_2(\lambda) \cdots P_k(\lambda) = d_1(\lambda)d_2(\lambda) \cdots d_n(\lambda) \quad (1)$$

由于 $P_i(\lambda)$ 是不可约多项式, 故存在某个 j , 使得 $P_i(\lambda) \mid d_j(\lambda)$, 否则, 由不可约多项式的基本性质 (1) 可知 $(P_i(\lambda), d_j(\lambda)) = 1, j = 1, 2, \dots, n$. 再由互素多项式和最大公因式的基本性质 (5) 可知 $(P_i(\lambda), d_1(\lambda)d_2(\lambda) \cdots d_n(\lambda)) = 1$, 这与 (1) 矛盾! 从而 $P_i(\lambda) \mid d_n(\lambda) (1 \leq i \leq k)$. 由互素多项式和最大公因式的基本性质 (1) 可知, $P_1(\lambda)P_2(\lambda) \cdots P_k(\lambda) \mid d_n(\lambda)$, 因此只能是 $d_1(\lambda) = \dots = d_{n-1}(\lambda) = 1, d_n(\lambda) = f(\lambda)$, 从而 A 的有理标准型只有一个 Frobenius 块. 由于特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda) \cdots P_k(\lambda)$ 在 \mathbb{K} 上无重因式, 故 $(f(\lambda), f'(\lambda)) = 1$, 从而 $f(\lambda)$ 在复数域上无重根, 即 A 有 n 个不同的特征值, 于是 A 在复数域上可对角化. \square

推论 0.1

设数域 \mathbb{K} 上的 n 阶矩阵 A 的特征多项式 $f(\lambda) = P_1(\lambda)P_2(\lambda) \cdots P_k(\lambda)$, 其中 $P_i(\lambda) (1 \leq i \leq k)$ 是 \mathbb{K} 上互异的首一不可约多项式. 并且 α_i 为线性方程组 $P_i(A)x = 0$ 的非零解, 则 $\alpha = \alpha_1 + \dots + \alpha_k$ 是 A 的循环空间 \mathbb{K}^n 的循环向量.

证明 由命题 0.3 及定理??可知 \mathbb{K}^n 就是一个循环空间. (未完成证明) \square

命题 0.4

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的特征多项式为 $f(\lambda)$, 证明以下 3 个结论等价:

- (1) V 只有平凡的 φ -不变子空间;
- (2) V 中任一非零向量都是循环向量, 使 V 成为循环空间;
- (3) $f(\lambda)$ 是 \mathbb{K} 上的不可约多项式.

证明 (1) \Rightarrow (2): 任取 V 中非零向量 α , 则循环子空间 $C(\varphi, \alpha)$ 是非零 φ -不变子空间. 由于 V 只有平凡的 φ -不变子空间, 故 $C(\varphi, \alpha) = V$, 即 V 中任一非零向量都是循环向量, 使 V 成为循环空间.

(2) \Rightarrow (3): 用反证法, 假设 $f(\lambda) = g(\lambda)h(\lambda)$, 其中 $g(\lambda), h(\lambda)$ 是 \mathbb{K} 上次数小于 n 的首一多项式. 由 Cayley - Hamilton 定理可知 $\mathbf{0} = f(\varphi) = g(\varphi)h(\varphi)$, 故由命题??(1) 的逆否命题可知 $g(\varphi), h(\varphi)$ 中至少有一个是奇异 (不可逆/非双射) 线性变换, 不妨设为 $g(\varphi)$, 由推论??可知 $\text{Ker} g(\varphi) \neq \mathbf{0}$. 任取 $\text{Ker} g(\varphi)$ 中的非零向量 α , 设 $\deg g(\lambda) = r$, 则不妨设

$$g(\varphi) = a_r \varphi^r + a_{r-1} \varphi^{r-1} + \cdots + a_1, \quad \text{其中 } a_r \neq 0.$$

由 $\alpha \in \text{Ker} g(\varphi)$ 可知

$$g(\varphi)(\alpha) = a_r \varphi^r(\alpha) + a_{r-1} \varphi^{r-1}(\alpha) + \cdots + a_1 \alpha = 0.$$

于是

$$\varphi^r(\alpha) = -\frac{a_{r-1}}{a_r} \varphi^{r-1}(\alpha) - \cdots - \frac{a_1}{a_r} \alpha. \quad (2)$$

假设对 $k \geq r$ 且 $k \in \mathbb{N}$, 成立 $\varphi^k(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 线性表示, 则对(2)式两边同时作用 φ^{k-r+1} 可得

$$\varphi^{k+1}(\alpha) = -\frac{a_{r-1}}{a_r} \varphi^k(\alpha) - \cdots - \frac{a_1}{a_r} \varphi^{k-r+1}(\alpha).$$

于是由归纳假设可知, $\varphi^{k+1}(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 线性表示. 故由数学归纳法可得, 对 $\forall k \geq r$ 且 $k \in \mathbb{N}$, 都有 $\varphi^k(\alpha)$ 可由 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 线性表示. 因此 $C(\varphi, \alpha) = L(\alpha, \varphi(\alpha), \cdots) = L(\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha))$, 其维数 $\leq r < n$, 故 $C(\varphi, \alpha) \neq V$, 这与 V 中任一非零向量都是循环向量矛盾!

(3) \Rightarrow (1): 用反证法, 假设存在非平凡的 φ -不变子空间 $U, \dim U = r$, 则 φ 在一组基下的表示矩阵为分块上三角矩阵 $M = \begin{pmatrix} A & C \\ O & B \end{pmatrix}$, 其中 A 是 $\varphi|_U$ 的表示矩阵. 于是特征多项式

$$f(\lambda) = |\lambda I_V - \varphi| = |\lambda I_n - M| = |\lambda I_r - A| \cdot |\lambda I_{n-r} - B|.$$

是两个低次多项式的乘积, 这与 $f(\lambda)$ 的不可约性矛盾! □

命题 0.5

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的极小多项式为 $m(\lambda)$. 证明: $m(\lambda)$ 是 \mathbb{K} 上的不可约多项式的充要条件是 V 的任一非零 φ -不变子空间 U 必为如下形式:

$$U = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k)$$

并且 $\varphi|_{C(\varphi, \alpha_i)}$ 的极小多项式都是 $m(\lambda)$. 此时, $\varphi|_U$ 的极小多项式也是 $m(\lambda)$.

证明 必要性: 设 $\varphi|_U$ 的极小多项式为 $n(\lambda)$, 则 $m(\varphi|_U) = m(\varphi)|_U = \mathbf{0}$, 从而 $n(\lambda) \mid m(\lambda)$. 因为 $m(\lambda)$ 不可约, 所以 $n(\lambda) = m(\lambda)$. 又由于 $\varphi|_U$ 的所有不变因子都要整除 $m(\lambda)$ 且 $m(\lambda)$ 不可约, 故所有的非常数不变因子都等于 $m(\lambda)$. 最后, 由有理标准型的几何意义即得

$$U = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k).$$

并且 $\varphi|_{C(\varphi, \alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \cdots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(m(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. 于是 $\varphi|_{C(\varphi, \alpha_i)}$ 的极小多项式就是其表示矩阵 $C(m(\lambda))$ 的极小多项式. 又由引理??可知, $C(m(\lambda))$ 的极小多项式就是 $m(\lambda)$, 并且 $n(\lambda) = m(\lambda)$, 故结论得证.

充分性: 用反证法, 设 $m(\lambda) = g(\lambda)h(\lambda)$, 其中 $g(\lambda), h(\lambda)$ 是 \mathbb{K} 上次数小于 $m(\lambda)$ 次数的首一多项式, 则 $\mathbf{0} = m(\varphi) = g(\varphi)h(\varphi)$, 故由命题??(1) 的逆否命题可知 $g(\varphi), h(\varphi)$ 中至少有一个是奇异线性变换, 不妨设为 $g(\varphi)$, 于是由推论??可知 $\text{Kerg}(\varphi) \neq \mathbf{0}$. 任取 $\text{Kerg}(\varphi)$ 中的非零向量 α , 得到循环子空间 $U = C(\varphi, \alpha)$,

由 $g(\varphi)(\alpha) = \mathbf{0}$ 可知, 对 $\forall k \in \mathbb{N}$, 都有 $\varphi^k(g(\varphi)(\alpha)) = \mathbf{0}$. 从而对 $\forall \beta \in U = C(\varphi, \alpha)$, 存在不全为零的 a_i 使得

$$\begin{aligned} g(\varphi)(\beta) &= g(a_1\alpha + a_2\varphi(\alpha) + a_3\varphi^2(\alpha) + \cdots) \\ &= a_1g(\alpha) + a_2g(\varphi(\alpha)) + a_3g(\varphi^2(\alpha)) + \cdots \\ &= a_1g(\alpha) + a_2\varphi(g(\alpha)) + a_3\varphi^2(g(\alpha)) + \cdots \\ &= \mathbf{0} + \mathbf{0} + \mathbf{0} + \cdots = \mathbf{0}. \end{aligned}$$

因此 $g(\varphi|_U) = g(\varphi)|_U = \mathbf{0}$, 于是 $\varphi|_U$ 的极小多项式 $m(\lambda)$ 整除 $g(\lambda)$, 从而其次数 $\leq \deg g(\lambda) < \deg m(\lambda)$, 这与条件矛盾! \square

定理 0.1 (基于初等因子组的有理标准型)

设数域 \mathbb{K} 上的 n 阶矩阵 A 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \dots, P_k(\lambda)^{r_k}$, 证明: A 相似于分块对角矩阵

$$\begin{aligned} \tilde{F} &= \text{diag}\{F(P_1(\lambda)^{r_1}), F(P_2(\lambda)^{r_2}), \dots, F(P_k(\lambda)^{r_k})\} \\ \tilde{C} &= \text{diag}\{C(P_1(\lambda)^{r_1}), C(P_2(\lambda)^{r_2}), \dots, C(P_k(\lambda)^{r_k})\} \end{aligned}$$

称为 A 的基于初等因子组的有理标准型.

证明 由 Frobenius 块和友阵的性质可知, $\lambda I_n - \tilde{F}$ 和 $\lambda I_n - \tilde{C}$ 都相抵于

$$\text{diag}\{1, \dots, 1, P_1(\lambda)^{r_1}; 1, \dots, 1, P_2(\lambda)^{r_2}; \dots; 1, \dots, 1, P_k(\lambda)^{r_k}\}$$

再由 λ -矩阵和初等因子的基本性质 (2) 可知, \tilde{F}, \tilde{C} 与 A 有相同的初等因子组, 从而它们相似. \square

定理 0.2

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \dots, P_k(\lambda)^{r_k}$. 证明: 存在 $\alpha_1, \alpha_2, \dots, \alpha_k \in V$, 使得

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k)$$

证明 由基于初等因子组的有理标准型和定理??即得. \square

例题 0.1 求证: 存在 n 阶实方阵 A , 满足 $A^2 + 2A + 5I_n = O$ 的充要条件是 n 为偶数. 当 $n \geq 4$ 时, 验证满足上述条件的矩阵 A 有无限个不变子空间.

证明 必要性: 注意到 A 适合多项式 $g(\lambda) = \lambda^2 + 2\lambda + 5$, 故 A 的极小多项式 $m(\lambda) \mid g(\lambda)$, 又因为 $g(\lambda)$ 在实数域上不可约, 故只能是 $m(\lambda) = g(\lambda)$. 同理可证 A 所有的非常数不变因子都等于 $g(\lambda)$, 从而 A 的不变因子组为 $1, \dots, 1, g(\lambda), \dots, g(\lambda)$ (k 个 $g(\lambda)$). 因此 A 的特征多项式 $f(\lambda) = g(\lambda)^k$, 于是 $n = \deg f(\lambda) = 2k$ 为偶数.

充分性: 设 $n = 2k$ 为偶数, 则由必要性的证明可知, A 的不变因子组为 $1, \dots, 1, g(\lambda), \dots, g(\lambda)$ (k 个 $g(\lambda)$). 可用有理标准型构造满足条件的矩阵:

$$A = \text{diag} \left\{ \begin{pmatrix} 0 & -5 \\ 1 & -2 \end{pmatrix}, \dots, \begin{pmatrix} 0 & -5 \\ 1 & -2 \end{pmatrix} \right\} \quad (k \text{ 个二阶方阵}) \quad (3)$$

当 $n \geq 4$ 时, 设 $\{e_1, e_2, e_3, e_4\}$ 是前 4 个标准单位列向量, 则容易验证(3)式中的矩阵 A 满足 $Ae_1 = e_2, Ae_3 = e_4$, 于是构造循环子空间 $\{C_l := C(A, e_1 + le_3) = L(e_1 + le_3, e_2 + le_4), l \in \mathbb{R}\}$, 进一步容易验证循环子空间 $\{C_l := C(A, e_1 + le_3) = L(e_1 + le_3, e_2 + le_4), l \in \mathbb{R}\}$ 是两两互异的 A -不变子空间, 故 A 有无限个不变子空间. \square

命题 0.6

设 A 是数域 \mathbb{K} 上的 n 阶方阵, 求证: A 的极小多项式的次数小于等于 $r(A) + 1$.

证明 证法一: 设 A 的不变因子组为 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 则极小多项式 $m(\lambda) = d_k(\lambda)$, 并且由定理??可知 A

相似于 $F = \text{diag}\{F(d_1(\lambda)), \dots, F(d_k(\lambda))\}$. 设 $\deg d_k(\lambda) = r$, 若 $d_k(0) \neq 0$, 则由 Frobenius 块的基本性质 (1) 可知 $F(d_k(\lambda))$ 非异; 若 $d_k(0) = 0$, 则由 Frobenius 块的基本性质 (1) 可知 $F(d_k(\lambda))$ 奇异且右上角的 $r-1$ 阶子式非零, 从而秩为 $r-1$. 因此, $r(A) = r(F) \geq r(F(d_k(\lambda))) \geq r-1 = \deg d_k(\lambda) - 1$.

证法二: 从 A 的极小多项式 $m(\lambda)$ 分离出来的初等因子中, 形如 λ^r 的初等因子至多只有 1 个, 对应于零特征值的 Jordan 块 $J_r(0)$, 其余的初等因子对应于非零特征值的 Jordan 块. 因此 $r(A)$ 大于等于这些 Jordan 块秩的和, 后者等于 $\deg m(\lambda) - 1$ 或 $\deg m(\lambda)$. \square

命题 0.7

设数域 \mathbb{K} 上的 n 阶矩阵 A 的不变因子组是 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 其中 $d_i(\lambda)$ 是非常数首一多项式, $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq k-1)$. 求证: 对 A 的任一特征值 λ_0 ,

$$r(\lambda_0 I_n - A) = n - \sum_{i=1}^k \delta_{d_i(\lambda_0), 0}$$

其中记号 $\delta_{a,b}$ 表示: 若 $a = b$, 取值为 1; 若 $a \neq b$, 取值为 0.

证明 证法一: 设 $\deg d_i(\lambda) = r_i$, 则由定理??可知 A 相似于 $F = \text{diag}\{F(d_1(\lambda)), \dots, F(d_k(\lambda))\}$, 而相似矩阵有相同的特征多项式, 故 $\lambda_0 I_n - A = \lambda_0 I_n - F$. 由 Frobenius 块的基本性质 (2) 可知 $|\lambda_0 I_{r_i} - F(d_i(\lambda))| = d_i(\lambda_0)$. 若 $d_i(\lambda_0) \neq 0$, 则 $\lambda_0 I_{r_i} - F(d_i(\lambda))$ 非异; 若 $d_i(\lambda_0) = 0$, 则 $\lambda_0 I_{r_i} - F(d_i(\lambda))$ 奇异且右上角的 r_i-1 阶子式非零, 从而秩为 r_i-1 . 因此,

$$\begin{aligned} r(\lambda_0 I_n - A) &= r(\lambda_0 I_n - F) = \sum_{i=1}^k r(\lambda_0 I_{r_i} - F(d_i(\lambda))) \\ &= \sum_{i=1}^k (r_i - \delta_{d_i(\lambda_0), 0}) = n - \sum_{i=1}^k \delta_{d_i(\lambda_0), 0} \end{aligned}$$

证法二: 由定理??可知存在可逆 λ -矩阵 $P(\lambda), Q(\lambda)$, 使得

$$P(\lambda)(\lambda I_n - A)Q(\lambda) = \text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\}$$

在上式中令 $\lambda = \lambda_0$, 注意到 $P(\lambda_0), Q(\lambda_0)$ 是 \mathbb{K} 上的可逆矩阵, 故 $\lambda_0 I_n - A$ 相抵于 $\text{diag}\{1, \dots, 1, d_1(\lambda_0), \dots, d_k(\lambda_0)\}$, 于是 $r(\lambda_0 I_n - A)$ 等于 n 减去等于零的 $d_i(\lambda_0)$ 的个数, 从而结论得证. \square

命题 0.8

- (1) 设 A 是数域 \mathbb{K} 上的 n 阶矩阵, 求证: A 相似于一个 \mathbb{K} 上主对角元全为零的矩阵的充要条件是 $\text{tr}(A) = 0$.
- (2) 设 A 是数域 \mathbb{R} 上的 n 阶矩阵, 求证: A 正交相似于一个 \mathbb{R} 上主对角元全为零的矩阵的充要条件是 $\text{tr}(A) = 0$.
- (3) 设 C 是数域 \mathbb{K} 上的 n 阶矩阵, 求证: 存在 \mathbb{K} 上的 n 阶矩阵 A, B , 使得 $AB - BA = C$ 的充要条件是 $\text{tr}(C) = 0$.

证明

(1) 必要性: 由矩阵的迹是矩阵的相似不变量, 故 $\text{tr}(A) = 0$.

充分性: 对阶数进行归纳. 当 $n = 1$ 时, $A = O$, 结论显然成立. 设阶数小于 n 时结论成立, 现证 n 阶的情形. 由于题目的条件和结论在相似关系下不改变, 故不妨从一开始就假设 A 是有理标准型

$$F = \text{diag}\{F(d_1(\lambda)), \dots, F(d_k(\lambda))\}$$

其中 $d_i(\lambda)$ 是 A 的非常数不变因子, $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq k-1)$, $\deg d_i(\lambda) = r_i$. 若 r_i 都为 1, 则 $d_1(\lambda) = \dots = d_n(\lambda) = \lambda - c$, 从而 $A = cI_n$. 又 $\text{tr}(A) = 0$, 故 $c = 0$, 从而 $A = O$, 结论成立. 以下假设存在某个 $r_i > 1$, 将第 $(1, 1)$ 分块与第 (i, i) 分块对换, 这是一个相似变换, 此时矩阵的第 $(1, 1)$ 元为零, 故不妨设 A 的第 $(1, 1)$ 元为零. 注意到矩阵 $A = \begin{pmatrix} 0 & \alpha' \\ \beta & B \end{pmatrix}$, 其中 $\alpha, \beta \in \mathbb{K}^{n-1}, B \in M_{n-1}(\mathbb{K}), \text{tr}(B) = 0$. 由归纳假设, 存在 \mathbb{K} 上的 $n-1$ 阶非异阵

Q , 使得 $Q^{-1}BQ$ 的主对角元全为零, 令 $P = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}$ 为 \mathbb{K} 上的 n 阶非异阵, 则 $P^{-1}AP = \begin{pmatrix} 0 & \alpha'Q \\ Q^{-1}\beta & Q^{-1}BQ \end{pmatrix}$ 的主对角元全为零, 结论得证.

注 (1) 的充分性也可以利用矩阵的初等变换进行归纳证明 (见清疏讲义).

(2) 必要性: 若矩阵 A 正交相似于主对角元全为零的矩阵, 则由相似变换不改变迹可知, $\text{tr}(A) = 0$.

充分性: 对阶数 n 进行归纳. 当 $n = 1$ 时, 结论显然成立. 假设对任意阶数小于 n 的实迹零矩阵结论成立. 考虑 n 阶实迹零矩阵 A , 令 $T = \frac{A+A'}{2}, S = \frac{A-A'}{2}$, 则 $A = T + S$, 并且 T 为实对称阵, S 为实反称阵.

(i) 若 $T = O$, 则 $A = S$ 是实反称矩阵, 而实反称矩阵主对角元均为 0, 结论显然成立.

(ii) 若 $T \neq O$, 则存在正交矩阵 P , 使得

$$P^{-1}TP = \begin{pmatrix} I_p & & \\ & -I_q & \\ & & O \end{pmatrix},$$

其中 p, q 分别为 T 的正、负惯性指数. 由 $T \neq O$ 可知 $r(T) = p + q > 0$, 从而

$$p - q = \text{tr}(P^{-1}TP) = \text{tr}(T) = \text{tr}(A) = 0 \Rightarrow p = q = \frac{r(T)}{2} > 0.$$

于是 $p, q \geq 1$. 故

$$P_{p+1,p}^{-1} P_{p+1,2}^{-1} (P^{-1}TP) P_{p+1,2} P_{p+1,p} = \begin{pmatrix} T_1 & & & \\ & I_{p-1} & & \\ & & I_{q-1} & \\ & & & O \end{pmatrix},$$

其中 $T_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, 记 $T_2 = \begin{pmatrix} I_{p-1} & \\ & I_{q-1} \\ & & O \end{pmatrix}$, $P_0 = PP_{p+1,2}P_{p+1,p}$, 则 P_0 为正交矩阵, 并且

$$P_0^{-1}TP_0 = \begin{pmatrix} T_1 & \\ & T_2 \end{pmatrix}.$$

注意到 T_1 为 2 阶迹零阵, T_2 为 $n-2$ 阶迹零阵. 从而由归纳假设可得, 存在正交矩阵 P_1, P_2 , 使得

$$P_1^{-1}T_1P_1 = T'_1, \quad P_2^{-1}T_2P_2 = T'_2,$$

其中 T'_1, T'_2 都是主对角元全为 0 的矩阵. 故令 $Q = P_0 \begin{pmatrix} P_1 & \\ & P_2 \end{pmatrix}$, 则 Q 也是正交矩阵, 并且

$$Q^{-1}TQ = \begin{pmatrix} T'_1 & \\ & T'_2 \end{pmatrix}.$$

因此 $Q^{-1}TQ$ 的主对角元全都为 0. 又由 S 为反称矩阵可知 $(Q^{-1}SQ)' = -Q^{-1}SQ$, 故 $Q^{-1}SQ$ 也为反称矩阵, 从而 $Q^{-1}SQ$ 的主对角元也全都为 0. 于是 $Q^{-1}AQ = Q^{-1}TQ + Q^{-1}SQ$ 的主对角元也全都为 0.

(3) 必要性由矩阵迹的线性和交换性即得, 下证充分性.

由于题目的条件和结论在同时相似变换 $A \mapsto P^{-1}AP, B \mapsto P^{-1}BP, C \mapsto P^{-1}CP$ 下不改变, 故由命题 0.8(1) 不妨从一开始就假设 $C = (c_{ij})$ 的主对角元 $c_{ii} = 0 (1 \leq i \leq n)$. 取定 $A = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ 为 \mathbb{K} 上的主对角元互异的对角矩阵. 设 $B = (x_{ij})$, 则 $AB - BA = C$ 等价于方程 $\lambda_i x_{ij} - \lambda_j x_{ji} = c_{ij}$. 当 $i = j$ 时, 上式恒成立, 故 x_{ii} 可任取. 当 $i \neq j$ 时, $x_{ij} = \frac{c_{ij}}{\lambda_i - \lambda_j}$ 被唯一确定. 因此, 一定存在 \mathbb{K} 上的矩阵 A, B , 使得 $AB - BA = C$ 成立. \square

推论 0.2

$$\dim \{AB - BA | A, B \in \mathbb{C}^{n \times n}\} = \dim \{A | \text{tr}(A) = 0\} = n^2 - 1.$$

证明 由命题 0.8(3) 可知, 只须证 $\dim \{A | \operatorname{tr}(A) = 0\} = n^2 - 1$.

证法一: 注意到 $\{A | \operatorname{tr}(A) = 0\}$ 的一组基为 $E_{ij} (i \neq j), E_{ii} - E_{i+1, i+1} (i = 1, 2, \dots, n-1)$. 这组基中共含有 $n^2 - 1$ 个线性无关的矩阵.

证法二: 将矩阵的迹看作 $\mathbb{C}^{n \times n} \rightarrow \mathbb{C}$ 的一个线性映射, 记作 tr . 于是 $\dim \{A | \operatorname{tr}(A) = 0\} = \dim \operatorname{Ker} \operatorname{tr} = \dim \mathbb{C}^{n \times n} - \dim \operatorname{Im} \operatorname{tr} = n^2 - \dim \mathbb{C} = n^2 - 1$. \square

命题 0.9

设 $\operatorname{tr}(A) = 0, f(x)$ 是多项式且 $\deg f(x) \geq 1$, 证明: 存在可逆方阵 X, Y 使得 $A = f(XY) - f(YX)$.

证明 由命题 0.2(1) 可知, 存在可逆矩阵 P , 使得

$$A = P\Lambda P^{-1},$$

其中 $\Lambda = (b_{ij})_{n \times n}$ 且 $b_{ii} = 0 (i = 1, 2, \dots, n)$. 注意到条件与结论在线性变换 $A \mapsto P\Lambda P^{-1}$ 下不改变, 因此不妨设

$$A = \Lambda, \text{ 则取 } Y = \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & \ddots & \\ & & & n \end{pmatrix}, \text{ 设 } X = (x_{ij}), \text{ 记矩阵 } U = XY, A_1 = AY, \text{ 则 } A_1 \text{ 也是主对角元全为零的矩阵,}$$

设 $A_1 = (a_{ij})_{n \times n}$, 其中 $a_{ii} = 0 (i = 1, 2, \dots, n)$. 从而

$$\begin{aligned} A &= f(XY) - f(YX) \iff A = f(U) - f(YUY^{-1}) = f(U) - Yf(U)Y^{-1} \\ &\iff AY = f(U)Y - Yf(U) \iff A_1 = f(U)Y - Yf(U). \end{aligned} \quad (4)$$

由命题 0.2(3) 可知, 存在 $M_0 \in \mathbb{C}^{n \times n}$, 使得 $A_1 = M_0Y - YM_0$. 令 $M_1 = M_0 + \lambda Y, \forall \lambda \in \mathbb{C}$, 则

$$A_1 = M_1Y - YM_1. \quad (5)$$

$$\text{设 } M_0 = (x_{ij})_{n \times n}, \text{ 则 } M_1 = \begin{pmatrix} x_{11} + \lambda & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} + 2\lambda & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} + n\lambda \end{pmatrix}. \text{ 从而可设 } D_k : |z - x_{kk} - k\lambda| \leq \sum_{j \neq k} x_{kj} \quad (1 \leq$$

$k \leq n)$ 为 M_1 的 n 个戈氏圆盘. 取 λ 充分大, 使得

$$|x_{kk} - k\lambda - f(0)| > \sum_{j \neq k} x_{kj} \quad (1 \leq k \leq n), \quad (6)$$

并且 D_1, \dots, D_n 互不相交, 从而此时 $D_i (1 \leq i \leq n)$ 各自构成一个连通分支. 于是由第二圆盘定理可知, 每个连通分支 D_i 中有且仅有一个特征值, 进而 M_1 此时有 n 个不同的特征值, 故此时 M_1 可对角化. 于是存在可逆阵 G , 使得

$$M_1 = G \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} G^{-1},$$

其中 $\lambda_1, \dots, \lambda_n$ 为 M_1 的所有特征值. 取 $u_i \in \mathbb{C}$ 为 $f(x) = \lambda_i (1 \leq i \leq n)$ 的根, 又由(6)式可知, $f(0)$ 落在 $\bigcup_{i=1}^n D_i$ 之

外, 故 M_1 的所有特征值 $\lambda_1, \dots, \lambda_n \neq f(0)$. 从而 $u_i \neq 0$ ($1 \leq i \leq n$). 于是令 $U = G \begin{pmatrix} u_1 & & \\ & u_2 & \\ & & \ddots \\ & & & u_n \end{pmatrix} G^{-1}$, 则

$$f(U) = G \begin{pmatrix} f(u_1) & & \\ & f(u_2) & \\ & & \ddots \\ & & & f(u_n) \end{pmatrix} G^{-1} = G \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{pmatrix} G^{-1} = M_1.$$

再结合(5)式可知 $A_1 = f(U)Y - Yf(U)$. 令 $X = UY^{-1}$, 则再根据(4)式就可知

$$A = f(XY) - f(YX).$$

故结论得证. □