

0.1 Jordan 标准型的应用

0.1.1 利用 Jordan 标准型研究矩阵的性质

命题 0.1

设 A 是 n 阶复矩阵, 求证: A 相似于分块对角矩阵 $\text{diag}\{B, C\}$, 其中 B 是幂零矩阵, C 是可逆矩阵.

注 这个命题告诉我们: 在相似的意义下, 对复方阵的研究可归结为对幂零矩阵和可逆矩阵这两类特殊矩阵的研究, 它们的刻画分别是: 特征值全为零以及特征值全不为零. 这也是前面很多例题都处理这两类矩阵的深层次原因.

证明 我们发现 A 的初等因子分离开了零特征值和非零特征值, 从而 A 的 Jordan 标准型满足题目要求. 此时, 可将零特征值的 Jordan 块 $J_r(0)$ (幂零矩阵) 放入 B 中, 将非零特征值的 Jordan 块 $J_r(\lambda_0)$ (可逆矩阵) 放入 C 中, 即得结论. □

定理 0.1

设 λ_0 是 n 阶矩阵 A 的特征值, 其代数重数为 m . 设属于特征值 λ_0 的最大 Jordan 块的阶数为 k , 求证:

$$r(A - \lambda_0 I_n) > \cdots > r((A - \lambda_0 I_n)^k) = r((A - \lambda_0 I_n)^{k+1}) = \cdots = n - m$$

证明 设 P 为可逆矩阵, 使 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \cdots, J_{r_s}(\lambda_s)\}$ 为 Jordan 标准型, 则对任意的正整数 j ,

$$r((A - \lambda_0 I_n)^j) = r(P^{-1}(A - \lambda_0 I_n)^j P) = r((J - \lambda_0 I_n)^j) = \sum_{i=1}^s r(J_{r_i}(\lambda_i - \lambda_0)^j)$$

若 $\lambda_i \neq \lambda_0$, 则 $r(J_{r_i}(\lambda_i - \lambda_0)^j) = r_i$. 若 $\lambda_i = \lambda_0$, 则当 $1 \leq j \leq r_i$ 时, $r(J_{r_i}(0)^j) = r_i - j$; 当 $j \geq r_i$ 时, $r(J_{r_i}(0)^j) = 0$. 注意到 A 至少有一个 Jordan 块 $J_k(\lambda_0)$, 并且属于特征值 λ_0 的所有 Jordan 块阶数之和等于 m , 故当 $1 \leq j \leq k$ 时, $r((A - \lambda_0 I_n)^j)$ 严格递减; 当 $j \geq k$ 时, $r((A - \lambda_0 I_n)^j) = n - m$. □

命题 0.2

设 λ_0 是 n 阶复矩阵 A 的特征值, 并且属于 λ_0 的初等因子都是次数大于等于 2 的多项式. 求证: 特征值 λ_0 的任一特征向量 α 均可表示为 $A - \lambda_0 I_n$ 的列向量的线性组合.

注 若特征值 λ_0 有一个初等因子为一次多项式, 则必存在特征向量 α , 它不能表示为 $A - \lambda_0 I_n$ 的列向量的线性组合. 证明的细节留给读者完成. 一个极端的例子就是 $A = I_n$, 其特征值 1 的初等因子都是一次的, 并且任一特征向量都不是 $A - I_n = O$ 的列向量的线性组合. 命题 0.2 与命题 ?? (可对角化的判定) 有着密切的联系, 请读者思考两者之间的关系.

证明 因为属于特征值 λ_0 的每个 Jordan 块的线性无关的特征向量均只有一个, 并且特征值 λ_0 的任一特征向量都是这些特征向量的线性组合, 因此我们只要证明 A 的 Jordan 标准型只含一个 Jordan 块 $J_n(\lambda_0)$ 的情形即可. 设 $P = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ 为非异阵, 使得 $P^{-1}AP = J_n(\lambda_0)$, 即 $AP = PJ_n(\lambda_0)$, 利用分块矩阵的乘法可得

$$\begin{aligned} A\alpha_1 &= \lambda_0 \alpha_1, \\ A\alpha_2 &= \alpha_1 + \lambda_0 \alpha_2, \\ &\vdots \\ A\alpha_n &= \alpha_{n-1} + \lambda_0 \alpha_n \end{aligned}$$

注意到 $n \geq 2$, 故有 $\alpha_1 = (A - \lambda_0 I_n)\alpha_2$, 从而特征向量 α_1 可表示为 $A - \lambda_0 I_n$ 的列向量的线性组合. □

0.1.2 运用 Jordan 标准型进行相似问题的化简

命题 0.3

设 A, B 为 n 阶矩阵, 满足 $AB = BA = O, r(A) = r(A^2)$, 求证: $r(A + B) = r(A) + r(B)$.



证明 注意到问题的条件和结论在同时相似变换: $A \mapsto P^{-1}AP, B \mapsto P^{-1}BP$ 下不改变, 故不妨从一开始就假设 A 为 Jordan 标准型. 设 $A = \text{diag}\{A_0, A_1\}$, 其中 A_0 由零特征值的 Jordan 块构成, A_1 由非零特征值的 Jordan 块构成. 由 $r(A) = r(A^2)$ 可知, 零特征值的 Jordan 块都是一阶的, 否则, $r(A_0^2) < r(A_0)$. 即 $A_0 = O$. 将 B 进行对应的分块 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$, 则由 $AB = BA = O$ 以及 A_1 非异可知, B_{12}, B_{21} 和 B_{22} 都是零矩阵. 于是

$$r(A + B) = r \begin{pmatrix} B_{11} & O \\ O & A_1 \end{pmatrix} = r(B_{11}) + r(A_1) = r(B) + r(A).$$

□

命题 0.4 ($AX = BX$ 经典结论)

- (1) 设 $A \in \mathbb{C}^{n \times n}, B \in \mathbb{C}^{m \times m}, m, n \in \mathbb{N}$, 求证: 矩阵方程 $AX = XB$ 只有零解的充要条件是 A, B 无公共的特征值.
- (2) 设 $A \in \mathbb{C}^{n \times n}, B \in \mathbb{C}^{m \times m}, m, n \in \mathbb{N}$, 求证: 矩阵方程 $AX = XB$ 有非 0 解 $X \in \mathbb{C}^{n \times m}$ 的充分必要条件是 A, B 至少有一个相同的特征值.



证明 上述两个命题等价, 因此我们只证 (2).

证法一: 先做两步化简. 注意到问题的条件和结论在矩阵变换: $B \mapsto P^{-1}BP, X \mapsto XP$ 下不改变, 故不妨从一开始就假设 B 为 Jordan 标准型. 设 $X = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 为列分块, 则有

$$AX = (A\alpha_1, A\alpha_2, \dots, A\alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)B = XB. \quad (1)$$

若 B 有 k 个 Jordan 块, 则方程组 (1) 可分解为 k 个独立方程组. 注意到:

- (i) 方程组 (1) 只有零解当且仅当这 k 个独立方程组都只有零解;
- (ii) 方程组 (1) 有非零解当且仅当这 k 个独立方程组中至少有一个有非零解.

因此, 不妨进一步假设 $B = J_n(\lambda_0)$ 为 Jordan 块. 此时, 方程组 (1) 等价于下列方程组:

$$A\alpha_1 = \lambda_0\alpha_1, A\alpha_2 = \alpha_1 + \lambda_0\alpha_2, \dots, A\alpha_n = \alpha_{n-1} + \lambda_0\alpha_n.$$

充分性: 假设 A, B 没有公共的特征值, 则 λ_0 不是 A 的特征值, 从而由 $A\alpha_1 = \lambda_0\alpha_1$ 只能得到 $\alpha_1 = \mathbf{0}$. 代入第二个方程可得 $A\alpha_2 = \lambda_0\alpha_2$, 相同的理由可推出 $\alpha_2 = \mathbf{0}$. 不断这样做下去, 最后可得 $\alpha_i = \mathbf{0} (1 \leq i \leq n)$, 即 $X = O$, 从而矩阵方程 $AX = XB$ 只有零解.

必要性: 假设 A 和 $B = J_n(\lambda_0)$ 有公共的特征值 λ_0 , 在上述方程组中令 $\alpha_1 = \dots = \alpha_{n-1} = \mathbf{0}$. 因为 λ_0 也是 A 的特征值, 所以 $A\alpha_n = \lambda_0\alpha_n$ 有非零解 $\alpha_n = \alpha$, 于是 $X_0 = (\mathbf{0}, \dots, \mathbf{0}, \alpha)$ 是上述方程组的非零解, 从而矩阵方程 $AX = XB$ 有非零解, 矛盾!

证法二: 必要性: 若 $AX = XB$ 有非零解 X , 记 $f(x)$ 是 B 的特征多项式, 则 $f(A)X = Xf(B)$ 也有非零解. 由 Cayley-Hamilton 定理可知 $f(B) = O$, 进而 $f(A)X = O$ 有非零解, 因此 $f(A)$ 是奇异阵, 从而 $f(A)$ 有 0 特征值. 又由命题 ?? 可知 $f(A)$ 的特征值都形如

$$f(\lambda_i), i = 1, 2, \dots, n.$$

其中 λ_i 都是 A 的特征值. 因此存在 $k \in \{1, 2, \dots, n\}$, 使得 $f(\lambda_k) = 0$. 而 $f(x)$ 是 B 的特征多项式, 故 λ_k 也是 B 的特征值. 因此 λ_k 既是 A 的特征值, 也是 B 的特征值.

充分性: 若 A, B 有相同特征值, 则由矩阵与其自身的转置相似可知, A, B, B^T 都有相同特征值 $\lambda \in \mathbb{C}$. 于是取相应的特征向量 $\alpha \in \mathbb{C}^n, \beta \in \mathbb{C}^m, \alpha, \beta \neq 0$, 使

$$A\alpha = \lambda\alpha, \quad B^T\beta = \lambda\beta \iff A\alpha = \lambda\alpha, \quad \beta^T B = \lambda\beta^T.$$

从而

$$A\alpha\beta^T = \lambda\alpha\beta^T = \alpha \cdot \lambda\beta^T = \alpha\beta^T B.$$

记 $X = \alpha\beta^T \neq 0$, 则由上式可知 $AX = XB$, 故 $X = \alpha\beta^T$ 就是 $AX = XB$ 的非零解.

□

命题 0.5

设 A, B 分别是 m, n 阶矩阵, C 是 $m \times n$ 矩阵, 求证: 矩阵方程 $AX - XB = C$ 存在唯一解的充要条件是 A, B 无公共的特征值.

▲

证明 先做两步化简. 注意到问题的条件和结论在矩阵变换 $B \mapsto P^{-1}BP, C \mapsto CP, X \mapsto XP$ 下不改变, 故不妨从一开始就假设 B 为 Jordan 标准型. 设 $X = (\alpha_1, \alpha_2, \dots, \alpha_n), C = (\beta_1, \beta_2, \dots, \beta_n)$ 为列分块, 则 $AX - XB = C$ 即为:

$$(A\alpha_1, A\alpha_2, \dots, A\alpha_n) - (\alpha_1, \alpha_2, \dots, \alpha_n)B = (\beta_1, \beta_2, \dots, \beta_n). \quad (2)$$

若 B 有 k 个 Jordan 块, 则方程组 (2) 可分解为 k 个独立方程组. 注意到:

- (i) 方程组 (2) 无解当且仅当这 k 个独立方程组中至少有一个无解;
- (ii) 方程组 (2) 有唯一解当且仅当这 k 个独立方程组都只有唯一解;
- (iii) 方程组 (2) 有无穷个解当且仅当这 k 个独立方程组都有解, 且至少有一个有无穷个解.

进一步, 若假设 $B = J_n(\lambda_0)$ 为 Jordan 块, 则方程组 (2) 等价于下列方程组: $(A - \lambda_0 I_n)\alpha_1 = \beta_1, (A - \lambda_0 I_n)\alpha_2 = \alpha_1 + \beta_2, \dots, (A - \lambda_0 I_n)\alpha_n = \alpha_{n-1} + \beta_n$.

充分性: 假设 A, B 没有公共的特征值, 则 λ_0 不是 A 的特征值, 从而 $A - \lambda_0 I_n$ 是可逆矩阵. 从第一个方程可解得 $\alpha_1 = (A - \lambda_0 I_n)^{-1}\beta_1$, 代入第二个方程可解得 $\alpha_2 = (A - \lambda_0 I_n)^{-1}(\alpha_1 + \beta_2), \dots$, 代入最后一个方程可解得 $\alpha_n = (A - \lambda_0 I_n)^{-1}(\alpha_{n-1} + \beta_n)$, 从而上述方程组有唯一解, 因此矩阵方程 $AX - XB = C$ 也有唯一解.

必要性: 假设 A, B 有公共的特征值 λ_0 , 若这 k 个独立方程组中有一个无解, 则矩阵方程 $AX - XB = C$ 无解, 从而结论成立. 若这 k 个独立方程组都有解, 则不妨设 $B = J_n(\lambda_0)$ 为 Jordan 块. 由于 λ_0 是 A 的特征值, 故 $(A - \lambda_0 I_n)x = 0$ 有无穷个解. 注意到, 若 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 是上述方程组的一个解, 则对 $(A - \lambda_0 I_n)x = 0$ 的任一解 $\alpha_0, (\alpha_1, \alpha_2, \dots, \alpha_n + \alpha_0)$ 也是上述方程组的解, 因此矩阵方程 $AX - XB = C$ 有无穷个解, 矛盾!

□

命题 0.6

设 A, B 分别是 m, n 阶矩阵, $M_{m \times n}(\mathbb{C})$ 上的线性变换 φ 定义为 $\varphi(X) = AX - XB$, 则下列 3 个结论等价:

- (1) φ 是单映射;
- (2) φ 是自同构;
- (3) 对某个给定的 $m \times n$ 矩阵 C , 存在唯一的 X_0 , 使得 $\varphi(X_0) = C$.

▲

注 由这个命题可知, 命题 0.4 和命题 0.5 都等价于命题 ??, 并且命题 0.4 和命题 0.5 都给出了它们的 Jordan 标准型证法.

证明 事实上, (1) \Rightarrow (2) 由推论 ?? 立得. (2) \Rightarrow (3) 显然成立. 用反证法来证明 (3) \Rightarrow (1): 若 $\text{Ker} \varphi \neq 0$, 则 $\text{Ker} \varphi$ 中任一非零元 X_1 都满足 $\varphi(X_0 + X_1) = C$, 这与唯一性矛盾.

□

定理 0.2 (交结数)

设 \mathbb{K} 为数域, 对 $A \in \mathbb{K}^{m \times m}, B \in \mathbb{K}^{n \times n}$, 计算矩阵方程 $AX = XB$ 在 $X \in \mathbb{K}^{m \times n}$ 的解空间维数.

♥

注 因为 $AX = XB$ 本质上是一个线性方程组, 线性方程组的解空间维数不随域扩张而改变. 因此我们只需在代数闭域 (即复数域 \mathbb{C}) 研究即可.

证明 不妨设 $\mathbb{K} = \mathbb{C}$, 设

$$P^{-1}AP = \text{diag}\{J_1, J_2, \dots, J_s\}, Q^{-1}BQ = \text{diag}\{T_1, T_2, \dots, T_t\}, P \in \mathbb{C}^{m \times m}, Q \in \mathbb{C}^{n \times n},$$

这里 $J_i, T_j, 1 \leq i \leq s, 1 \leq j \leq t$ 都是相同特征值对应所有 Jordan 块构成的分块对角矩阵. 注意到

$$P^{-1}APP^{-1}XQ = P^{-1}XQQ^{-1}BQ,$$

于是定理的条件和结论在同构 $A \rightarrow P^{-1}AP, B \rightarrow Q^{-1}BQ, X \rightarrow P^{-1}XQ$ 下不改变. 因此可不妨设

$$A = \text{diag}\{J_1, J_2, \dots, J_s\}, B = \text{diag}\{T_1, T_2, \dots, T_t\},$$

这里 $J_i, T_j, 1 \leq i \leq s, 1 \leq j \leq t$ 都是相同特征值对应所有 Jordan 块构成的分块对角矩阵. 不妨设 $0 \leq s_0 \leq \min\{s, t\}$ 使得 J_1, J_2, \dots, J_{s_0} 分别和 T_1, T_2, \dots, T_{s_0} 特征值对应相同, 且 A, B 剩下的每一个块对应特征值都不相同.

为了解 $AX = XB$, 把 X 对应分块为 $(X_{ij}), 1 \leq i \leq s, 1 \leq j \leq t$, 于是 $AX = XB$ 等价于

$$J_i X_{ij} = X_{ij} T_j, 1 \leq i \leq s, 1 \leq j \leq t. \quad (3)$$

注意到由命题 0.4, 我们知道(3)中只有 $1 \leq i = j \leq s_0$ 时 X_{ij} 才能非 0. 于是记 $AX = XB$ 解空间为 W , 我们有

$$\dim W = \sum_{i=1}^{s_0} \dim\{Y : J_i Y = Y T_i\}. \quad (4)$$

对固定的 $1 \leq i \leq s_0$, 为了计算 $\dim\{Y : J_i Y = Y T_i\}$, 我们设

$$J_i = \text{diag}\{J_{n_1}(\lambda), J_{n_2}(\lambda), \dots, J_{n_{s'}}(\lambda)\}, T_i = \text{diag}\{J_{n'_1}(\lambda), J_{n'_2}(\lambda), \dots, J_{n'_{t'}}(\lambda)\}.$$

做对应分块

$$Y = (Y_{i'j'}), 1 \leq i' \leq s', 1 \leq j' \leq t'.$$

同样的 $J_i Y = Y T_i$ 等价于

$$J_{n_{i'}}(\lambda) Y_{i'j'} = Y_{i'j'} J_{n'_{j'}}(\lambda), \forall 1 \leq i' \leq s', 1 \leq j' \leq t'. \quad (5)$$

而上述方程只能通过矩阵乘法定义计算. 由推论??, 我们知道

$$\dim\{Y_{i'j'} : J_{n_{i'}}(\lambda) Y_{i'j'} = Y_{i'j'} J_{n'_{j'}}(\lambda)\} = \min\{n_{i'}, n'_{j'}\},$$

于是由(4)可计算出结果.

综上所述, 设 s_0 是 A, B 公共特征值数, J_i 由 $n_{i1}, n_{i2}, \dots, n_{ip_i}$ 阶 Jordan 块构成, T_i 由 $t_{i1}, t_{i2}, \dots, t_{iq_i}$ 阶 Jordan 块构成, 其中 $i = 1, 2, \dots, s_0$. 则

$$\dim W = \sum_{i=1}^{s_0} \sum_{r=1}^{p_i} \sum_{\ell=1}^{q_i} \min\{n_{ir}, t_{i\ell}\}. \quad (6)$$

□

命题 0.7

假定 \mathbb{F} 是一个域.

- (1) 设 $B \in \mathbb{F}^{n \times n}, A \in \mathbb{F}^{m \times m}$. 若 A, B 在域 \mathbb{F} 至少有 r 个相同特征值, 则矩阵方程 $AX = XB$ 在 $X \in \mathbb{F}^{m \times n}$ 中有秩 r 的解.
- (2) 设 $B \in \mathbb{F}^{n \times n}, A \in \mathbb{F}^{m \times m}$, 则矩阵方程 $AX = XB$ 在 $X \in \mathbb{F}^{m \times n}$ 中有秩 r 的解, 则 A, B 在包含 \mathbb{F} 的代数域上至少有 r 个相同特征值 (计重数).

▲

证明

- (1) 设 A, B 有公共特征值 $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{F}$, 于是可设

$$Au_i = \lambda_i u_i, B^T v_i = \lambda_i v_i, u_i, v_i \in \mathbb{F}^n \setminus \{0\}, i = 1, 2, \dots, r.$$

考虑

$$X \triangleq \sum_{i=1}^r u_i v_i^T \in \mathbb{F}^{n \times n},$$

我们有

$$AX = A \left(\sum_{i=1}^r u_i v_i^T \right) = \sum_{i=1}^r u_i \lambda_i v_i^T = \left(\sum_{i=1}^r u_i v_i^T \right) B = XB.$$

又

$$\sum_{i=1}^r u_i v_i^T = (u_1, u_2, \dots, u_r) \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{pmatrix},$$

即 X 是列满秩矩阵和行满秩矩阵之积, 从而由 Sylvester 不等式得

$$r = r(u_1, u_2, \dots, u_r) + r \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{pmatrix} - r \leq r(X) \leq r \begin{pmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{pmatrix} = r \implies r(X) = r.$$

这就完成了证明.

(2) 若矩阵方程 $AX = XB$ 在 $X \in \mathbb{F}^{m \times n}$ 中有秩 r 的解 X_0 , 考虑 $PAP^{-1}PXQ = PXQQ^{-1}BQ$, 则可不妨设

$$X = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}, A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}, A_1, B_1 \in \mathbb{F}^{r \times r}.$$

于是

$$AX = XB \Leftrightarrow \begin{pmatrix} A_1 & 0 \\ A_3 & 0 \end{pmatrix} = \begin{pmatrix} B_1 & B_2 \\ 0 & 0 \end{pmatrix} \Leftrightarrow A_3 = 0, B_2 = 0, A_1 = B_1,$$

即我们有

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_4 \end{pmatrix}, B = \begin{pmatrix} A_1 & 0 \\ B_3 & B_4 \end{pmatrix}.$$

于是 A_1 的特征值是 A, B 的 r 个公共特征值.

□

推论 0.1

设 $B \in \mathbb{C}^{n \times n}, A \in \mathbb{C}^{m \times m}$, 则矩阵方程 $AX = XB$ 在 $X \in \mathbb{C}^{m \times n}$ 中有秩 r 的解的充要条件是 A, B 在 \mathbb{C} 至少有 r 个相同特征值 (计重数).



证明 充分性: 由命题 0.7(1) 立得.

必要性: 从定理 0.2 我们可看到, 不妨设

$$A = \text{diag}\{J_1, J_2, \dots, J_s\}, B = \text{diag}\{T_1, T_2, \dots, T_t\},$$

这里 $J_i, T_j, 1 \leq i \leq s, 1 \leq j \leq t$ 都是相同特征值对应所有 Jordan 块构成的分块对角矩阵. 再不妨设 $m \leq n, s \leq t$ 且存在 $s_0 \in [0, s] \cap \mathbb{N}$ 使得 J_1, J_2, \dots, J_{s_0} 分别和 T_1, T_2, \dots, T_{s_0} 特征值对应相同, 且 A, B 剩下的每一个块对应特征值都不相同. J_i 由 $n_{i1}, n_{i2}, \dots, n_{ip_i}$ 阶 Jordan 块构成, T_i 由 $t_{i1}, t_{i2}, \dots, t_{iq_i}$ 阶 Jordan 块构成. 再设 X 为对应分块矩阵, 则由 $AX = XB$ 知

$$\begin{pmatrix} J_1 X_{11} & J_1 X_{12} & \cdots & J_1 X_{1t} \\ J_2 X_{21} & J_2 X_{22} & \cdots & J_2 X_{2t} \\ \vdots & \vdots & & \vdots \\ J_s X_{s1} & J_s X_{s2} & \cdots & J_s X_{st} \end{pmatrix} = \begin{pmatrix} X_{11} T_1 & X_{12} T_2 & \cdots & X_{1t} T_t \\ X_{21} T_1 & X_{22} T_2 & \cdots & X_{2t} T_t \\ \vdots & \vdots & & \vdots \\ X_{s1} T_1 & X_{s2} T_2 & \cdots & X_{st} T_t \end{pmatrix}$$

于是

$$J_i X_{ij} = X_{ij} T_j, \quad i = 1, 2, \dots, s; j = 1, 2, \dots, t.$$

再由命题??知, 当 $i \neq j$ 或 $i, j > s_0$ 时, $X_{ij} = O$. 故

$$X = \begin{pmatrix} X_{11} & & & & \\ & X_{22} & & & \\ & & \ddots & & \\ & & & X_{s_0 s_0} & \\ & & & & O \\ & & & & & \ddots \\ & & & & & & O \end{pmatrix}$$

从而

$$r(X) = \sum_{i=1}^{s_0} r(X_{ii}).$$

因为 X_{ii} 的秩小于等于其行数和列数, 所以每个 X_{ii} 至多可以提供 T_i, J_i 中较小的阶数的秩, 这个数字也恰好是属于 T_i, J_i 的特征值的公共个数 (计重数). 因此 $r(X)$ 必小于等于 A, B 的公共特征值的个数. 这就证明了 $AX = XB$ 在 $X \in \mathbb{C}^{n \times m}$ 中有秩 r 的解, 则 A, B 至少有 r 个相同特征值 (计重数).

□

推论 0.2

计算与数域上 n 阶矩阵 A 可交换的全体矩阵空间构成的维数.

♥



笔记 $C(A) \triangleq \{X : AX = XA\}$ 本质上是一个线性方程组的解, 因此不随域扩张而改变, 所以不妨只在复数域上考虑.

证明 不妨只在复数域上考虑. 设 s_0 是 A 不同特征值个数, $A = \text{diag}\{J_1, J_2, \dots, J_{s_0}\}$, 其中 J_i 是 $n_{i,1}, n_{i,2}, \dots, n_{i,m_i}$ 阶相同特征值对应 Jordan 块组成的分块对角矩阵. 于是由定理 0.2, 可得

$$\dim C(A) = \sum_{i=1}^{s_0} \sum_{r=1}^{m_i} \sum_{\ell=1}^{m_i} \min\{n_{i,r}, n_{i,\ell}\}.$$

□

命题 0.8

设数域上 n 阶矩阵 A 且 $F(A)$ 是 A 的多项式空间, 则 $F(A) \subset C(A) \triangleq \{X : AX = XA\}$ 且 $\dim F(A) = \deg m_A$, 这里 m_A 是 A 的极小多项式.

♠

证明 显然有 $F(A) \subset C(A)$. 设 $\deg m_A = r, \sum_{i=0}^{r-1} c_i A^i = O$, 则 $\sum_{i=0}^{r-1} c_i x^i$ 就是 A 的零化多项式. 由极小多项式的性质 (1) 可知 $m_A(x) \mid \sum_{i=0}^{r-1} c_i x^i$, 而 $\deg m_A = r > r-1 = \deg \sum_{i=0}^{r-1} c_i x^i$, 故 $\sum_{i=0}^{r-1} c_i x^i = O$, 即 $c_i = 0$. 因此 I_n, A, \dots, A^{r-1} 线性无关, 故

$$\dim F(A) = \dim L(I_n, A, \dots, A^{r-1}) = r.$$

□

定理 0.3 ($F(A)=C(A)$ 的充要条件和不变子空间刻画)

设数域 \mathbb{K} 上 n 阶矩阵 A , 记 $C(A) \triangleq \{X : AX = XA\}$, $F(A) \triangleq \{f(A) : f \in \mathbb{K}[x]\}$, 则有下列等价刻画:

(1) 存在 n 维向量 v 使得 $v, Av, \dots, A^{n-1}v$ 是线性无关的;

- (2) A 相似于一个友矩阵;
 (3) $F(A) = C(A)$;
 (4) $\dim F(A) = \dim C(A)$;
 (5) $\sum_{i=1}^{s_0} \max\{n_{i,1}, n_{i,2}, \dots, n_{i,m_i}\} = \sum_{i=1}^{s_0} \sum_{r=1}^{m_i} \sum_{\ell=1}^{m_i} \min\{n_{i,r}, n_{i,\ell}\}$;
 (6) $m_i = 1, i = 1, 2, \dots, s$;
 (7) 每种复特征值的几何重数都是 1;
 (8) 对每个特征值 $\lambda \in \mathbb{C}$, 都有 $r(\lambda E - A) = n - 1$;
 (9) 特征多项式 = 极小多项式;
 (10) A 的不变子空间个数有限.
 并当 A 的不变子空间个数有限时, 计算 A 的所有不变子空间.



证明 由课本定理、命题??、定理 0.2、推论 0.2 知除第一个, 第二个和最后一个等价外, 其余都是显然的.

(9) \iff (3) 见定理??.

- (1) 我们来证明第 1 个和第 2 个等价. 直接在基 $v, Av, \dots, A^{n-1}v$ 下计算 A 的表示矩阵并利用 $A^n = c_0 I_n + c_1 A + \dots + c_{n-1} A^{n-1}$ 得 A 相似于一个友矩阵. 反过来若 A 相似于一个友矩阵, 不失一般性, 假设

$$A = \begin{pmatrix} 0 & \cdots & 0 & c_0 \\ 1 & \cdots & 0 & c_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & c_{n-1} \end{pmatrix}, v = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

我们直接演算有

$$Av = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, A^2v = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, A^{n-1}v = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

即 $v, Av, \dots, A^{n-1}v$ 是线性无关的, 我们完成了证明.

- (2) 我们来证明第 2 个和第 9 个等价. 友矩阵的特征多项式和极小多项式相等, 所以第 2 条推第 9 条成立. 反过来若第 9 条成立, 即 A 特征多项式和极小多项式相等, 设为 f . 于是 A 的不变因子为 $\underbrace{1, 1, \dots, 1}_{n-1}, f$. 而 f 的友矩阵的不变因子也为 $\underbrace{1, 1, \dots, 1}_{n-1}, f$. 因此我们有第 2 条成立, 我们完成了证明.

- (3) 我们来证明最后一个和上面每一条等价并完整刻画 A 的不变子空间.

Step1 我们先做一些讨论, 使得不变子空间问题归结到 A 的特征多项式是个不可约多项式的次方. 设 A 是数域 \mathbb{F} 上线性空间 V 的线性变换, 其特征多项式有标准分解

$$f(\lambda) = p_1^{r_1}(\lambda) p_2^{r_2}(\lambda) \cdots p_s^{r_s}(\lambda),$$

这里 $r_i \in \mathbb{N}, p_i(\lambda), i = 1, 2, \dots, s$ 两两互素且不可约. 由定理??, 我们知道有直和分解

$$V = \ker f(A) = \bigoplus_{i=1}^s \ker p_i^{r_i}(A). \quad (7)$$

任取 A 的不变子空间 U . 下证明

$$U = \bigoplus_{i=1}^s \left(U \cap \ker p_i^{r_i}(A) \right). \quad (8)$$

直和由零的分解唯一易证. 对 $x \in U$, 利用分解(7), 我们设

$$x = \sum_{i=1}^s u_i, u_i \in \ker p_i^{r_i}(A), i = 1, 2, \dots, s.$$

只需证明 $u_i \in U, i \in \{1, 2, \dots, s\}$. 证明的想法通过中国剩余定理使得 u_i 暴露出来并让其余 $u_j, j \neq i$ 变成 0. 由中国剩余定理, 取多项式 $g_i \in \mathbb{F}[\lambda]$ 使得

$$\begin{cases} g_i \equiv 0 \pmod{p_j^{r_j}}, j \neq i \\ g_i \equiv 1 \pmod{p_i^{r_i}}, j = i \end{cases},$$

则

$$g_i(A)x = \sum_{j=1}^s g_i(A)u_j = u_i \in U, i = 1, 2, \dots, s,$$

这就证明了(8). 注意到 $U \cap \ker p_i^{r_i}(A)$ 是 $A|_{\ker p_i^{r_i}(A)}$ 的不变子空间且 $A|_{\ker p_i^{r_i}(A)}$ 的特征多项式是 $p_i^{r_i}(\lambda)$, 这里前者是显然的, 后者注意到写成矩阵有

$$A = \text{diag}\{A|_{\ker p_1^{r_1}(A)}, A|_{\ker p_2^{r_2}(A)}, \dots, A|_{\ker p_s^{r_s}(A)}\}.$$

若特征多项式不是 $p_i^{r_i}(\lambda)$, 则这些特征多项式的乘积就是 A 的特征多项式不为 $f(\lambda)$ 矛盾! 因此得到我们需要的. 现在任取 $A|_{\ker p_i^{r_i}(A)}$ 不变子空间 $W_i, i = 1, 2, \dots, s$, 容易验证 $\bigoplus_{i=1}^s W_i$ 是 A 不变子空间, 因此我们的

Step1 的目标完成.

Step2 设 A 是数域 \mathbb{F} 上线性空间 V 的线性变换, 其特征多项式极小多项式都为 $f(\lambda) = p^r(\lambda)$, 这里 p 不可约且 $r \in \mathbb{N}, \deg p = d > 0$. 我们将完整计算出 A 不变子空间. 我们断言

$$\dim \ker p^s(A) = ds, 0 \leq s \leq r. \quad (9)$$

一方面由维数公式, 我们有

$$\dim \ker p^s(A) = \dim V - r(p^s(A)) = rd - r(p^s(A)).$$

另外一方面因为秩不随域扩张而改变, 我们在复数域由这个定理 0.3(7)(9)知 A 的所有特征值几何重数都是 1, 可设

$$A \sim \text{diag}\{J_r(\lambda_1), J_r(\lambda_2), \dots, J_r(\lambda_d)\}, \lambda_i \text{ 两两互不相同}, i = 1, 2, \dots, d.$$

由

$$\begin{aligned} r(p^s(A)) &= r(\text{diag}\{p^s(J_r(\lambda_1)), p^s(J_r(\lambda_2)), \dots, p^s(J_r(\lambda_d))\}) \\ &= \sum_{i=1}^d r(p^s(J_r(\lambda_i))) = \sum_{i=1}^d r[(J_r(\lambda_i) - \lambda_1 E)^s (J_r(\lambda_i) - \lambda_2 E)^s \cdots (J_r(\lambda_i) - \lambda_d E)^s] \\ &\quad \xrightarrow{i \neq j \text{ 时 } (J_r(\lambda_i) - \lambda_j E)^s \text{ 可逆}} \sum_{i=1}^d r((J_r(\lambda_i) - \lambda_i E)^s) = \sum_{i=1}^d (r - s) = d(r - s), \end{aligned}$$

我们知(9)成立. 自然的, 期望 A 不变子空间一定形如 $U = \ker p^s(A), 0 \leq s \leq r$. 事实上任取 A 不变子空间 U . 注意到 $A|_U$ 特征多项式必然是 $p^s(\lambda), 0 \leq s \leq r$. 因为 $p^s(A|_U) = 0$, 所以 $U \subset \ker p^s(A)$. 由(9)的维数等式知 $U = \ker p^s(A)$. 这就完整刻画清楚了 A 不变子空间.

Step3 下证 A 的极小多项式不等于特征多项式, 不变子空间有无穷多个.

由 **Step1**, 不妨设 A 特征多项式 $f(\lambda) = p^r(\lambda)$, 这里 p 不可约且 $r \in \mathbb{N}, \deg p = d > 0$. 再假设 f 极小多项式为 $p^s(\lambda), 1 \leq s < r$. 任取向量 $a \neq 0$ 并考虑 a 生成的最小 A 不变子空间 (见定理??和定理??) $V_a = \text{span}\{a, Aa, A^2a, \dots, A^{ds-1}a\}$. 因为 $s < r$, 所以必有 $\dim V_a < \dim V = rd$, 因此可以取向量 $a_2 \notin V_a$, 继续考虑 a_2 生成的最小 A 不变子空间 V_{a_2} . 由覆盖定理, 非平凡真子空间的并不可能为全空间, 即 $V \neq V_a \cup V_{a_2}$, 于是又可取 $a_3 \in V \setminus (V_a \cup V_{a_2})$. 因此这样下去就得到无穷多个不变子空间.

□

例题 0.1 证明: 存在 $A \in \mathbb{R}^{n \times n}$ 满足 $A^2 + 2A + 5I_n = 0$ 的充要条件是 n 为偶数. 当 $n \geq 4$, 验证满足上面矩阵 A 有无限个不变子空间.

证明 由条件知 $x^2 + 2x + 5$ 是 A 的零化多项式, 故 A 的极小多项式整除 $x^2 + 2x + 5$. 又注意到 $x^2 + 2x + 5$ 在 \mathbb{R} 上不可约, 故 $x^2 + 2x + 5$ 是 A 的极小多项式. 从而 A 的不变因子组可设为

$$1, \dots, 1, x^2 + 2x + 5, \dots, x^2 + 2x + 5.$$

于是由有理标准型知

$$A \sim \begin{pmatrix} A_1 & & \\ & A_2 & \\ & & \ddots \\ & & & A_s \end{pmatrix},$$

其中 A_1, A_2, \dots, A_s 都是二阶友矩阵, 故 n 为偶数. 显然, 当 $n \geq 4$, A 的极小多项式一定不等于特征多项式. 故由定理 0.3(10)知 A 有无限个不变子空间.

□

例题 0.2 设 V 是复数域上的 n 维线性空间, A 是 V 上的一个线性变换, 证明存在 $\alpha \in V$, 使得 $\{\alpha, A\alpha, \dots, A^{n-1}\alpha\}$ 为 V 一组基当且仅当对 A 的任一特征值 λ , λ 的几何重数是 1.

证明 由定理 0.3 即得.

□

例题 0.3 假设 T 是数域上的线性空间 V 上的线性变换且存在 $v \in V$ 使得 $v, Tv, \dots, T^{n-1}v$ 是 V 的一组基, T 的极小多项式为 $m(\lambda) = \prod_{i=1}^s (\lambda - \lambda_i)^{v_i}$, 这里 $\{\lambda_i\}$ 互不相同, 证明: 该数域上的多项式 p 满足存在 $u \in V$ 使得 $u, p(T)u, \dots, p^{n-1}(T)u$ 也是 V 的一组基当且仅当 $\lambda_i \neq \lambda_j$ 时 $p(\lambda_i) \neq p(\lambda_j)$ 且 $v_i > 1$ 时 $p'(\lambda_i) \neq 0$.

证明 Step1 由 Jordan 块的性质 (5), 我们知道

$$p(J_m(\mu)) = \begin{pmatrix} p(\mu) & p'(\mu) & \cdots & \frac{p^{(m-1)}(\mu)}{(m-1)!} \\ & p(\mu) & \ddots & \vdots \\ & & \ddots & p'(\mu) \\ & & & p(\mu) \end{pmatrix}.$$

即当且仅当 $p'(\mu) \neq 0$ 时成立

$$r(p(J_m(\mu))) = r(J_m(p(\mu))) = n \text{ 或 } n-1 \implies p(J_m(\mu)) \sim J_m(p(\mu)).$$

所以条件 $\lambda_i \neq \lambda_j$ 时 $p(\lambda_i) \neq p(\lambda_j)$ 且 $v_i > 1$ 时 $p'(\lambda_i) \neq 0$ 的充要条件是若 T 有 Jordan 标准型

$$\text{diag}\{J_{m_1}(\mu_1), J_{m_2}(\mu_2), \dots, J_{m_t}(\mu_t)\},$$

则 $p(T)$ 有 Jordan 标准型

$$\text{diag}\{J_{m_1}(p(\mu_1)), J_{m_2}(p(\mu_2)), \dots, J_{m_t}(p(\mu_t))\}.$$

Step2 注意到 T 在基 $v, Tv, \dots, T^{n-1}v$ 下的矩阵是一个友矩阵, 所以特征多项式和极小多项式相等, 所以每种特征值只有 1 个 Jordan 块. 由定理 0.3 知存在 $u \in V$ 使得 $u, p(T)u, \dots, p^{n-1}(T)u$ 也是 V 的一组基当且仅当 $p(T)$ 每种特征值只有 1 个 Jordan 块. 由 Step 1 知这当且仅当 $\lambda_i \neq \lambda_j$ 时 $p(\lambda_i) \neq p(\lambda_j)$ 且 $v_i > 1$ 时 $p'(\lambda_i) \neq 0$.

□

0.1.3 应用 Jordan 标准型的三段论法

如果矩阵问题的条件和结论在相似关系下不改变, 则可以先证明结论对 Jordan 块成立, 再证明对 Jordan 标准型成立, 最后证明对一般的矩阵也成立, 这就是所谓的“三段论法”. 事实上, 我们已经利用三段论法证明过例

题??和例题??, 下面再来看一些典型的例题.

首先, 我们来看计算矩阵乘幂的问题. 设 A 为 n 阶矩阵, P 为 n 阶可逆矩阵, 使得

$$P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_k}(\lambda_k)\}$$

为 Jordan 标准型. 注意任一 Jordan 块 $J_{r_i}(\lambda_i)$ 都有分解 $J_{r_i}(\lambda_i) = \lambda_i I_{r_i} + N$, 其中 $N = J_{r_i}(0)$ 是特征值为零的 r_i 阶 Jordan 块, 故对任意的正整数 m ,

$$J_{r_i}(\lambda_i)^m = (\lambda_i I_{r_i} + N)^m = \lambda_i^m I_{r_i} + C_m^1 \lambda_i^{m-1} N + \dots + C_m^{m-1} \lambda_i N^{m-1} + N^m.$$

于是 $J^m = \text{diag}\{J_{r_1}(\lambda_1)^m, J_{r_2}(\lambda_2)^m, \dots, J_{r_k}(\lambda_k)^m\}$, 从而 $A^m = (PJP^{-1})^m = PJ^mP^{-1}$ 便可计算出来了.

例题 0.4 设 $A = \begin{pmatrix} 2 & 6 & -15 \\ 1 & 1 & -5 \\ 1 & 2 & -6 \end{pmatrix}$, 求 $A^m (m \geq 1)$.

注 本题是例题??的延拓.

解 由例题??, 我们已经计算出过渡矩阵 P , 使得 $P^{-1}AP = J = \text{diag}\{-1, J_2(-1)\}$, 于是可进一步计算出

$$\begin{aligned} A^m &= PJ^mP^{-1} = \begin{pmatrix} -2 & 3 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} (-1)^m & 0 & 0 \\ 0 & (-1)^m & (-1)^{m-1}m \\ 0 & 0 & (-1)^m \end{pmatrix} \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ 1 & 2 & -5 \end{pmatrix} \\ &= (-1)^{m-1} \begin{pmatrix} 3m-1 & 6m & -15m \\ m & 2m-1 & -5m \\ m & 2m & -5m-1 \end{pmatrix}. \end{aligned}$$

□

例题 0.5 求矩阵 B , 使得 $A = B^2$, 其中 $A = \begin{pmatrix} 3 & 1 \\ -1 & 5 \end{pmatrix}$.

解 利用过渡矩阵的求法的方法, 可求出过渡矩阵 $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, 使得 $P^{-1}AP = J = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$ 为 Jordan 标准型. 用

待定元素法不难求得 $C = \pm \begin{pmatrix} 2 & \frac{1}{4} \\ 0 & 2 \end{pmatrix}$, 使得 $C^2 = J$.

注意到 $(PCP^{-1})^2 = PC^2P^{-1} = PJP^{-1} = A$, 故可取 $B = PCP^{-1}$. 经计算可得

$$B = \pm \begin{pmatrix} \frac{7}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{9}{4} \end{pmatrix}.$$

□

定义 0.1 (非异阵的 k 次方根)

设 A 为 n 阶非异复矩阵, 称 B 为 A 的 k 次方根, 若 B 满足 $A = B^k$, 其中 $k \in \mathbb{N}_1$.

♣

定理 0.4 (非异阵存在任意次的方根)

设 A 为 n 阶非异复矩阵, 证明: 对任一正整数 m , 存在 n 阶复矩阵 B , 使得 $A = B^m$.

♡

注 定理 0.4 的结论对奇异矩阵一般并不成立. 例如, 设 $A = J_n(0)^{m-1}$, 其中 $n = mq - r$, $m \geq 2$ 且 $0 \leq r < m$, 则不存在 B , 使得 $A = B^m$. 我们用反证法来证明这个结论. 若存在满足条件的 B , 则 B 的特征值全为零, 从而 B 也是幂零矩阵, 即有 $B^n = O$. 于是 $O = B^{n+r} = (B^m)^q = A^q = J_n(0)^{(m-1)q} \neq O$, 这就导出了矛盾.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 由于 A 非异, 故 A 的所有特征值都非零. 对 A 的任一 Jordan 块 $J_{r_i}(\lambda_i)$, 取定 λ_i 的某个 m 次方根 μ_i , 即 $\mu_i^m = \lambda_i$, 则由命题??可知, $J_{r_i}(\mu_i)^m$ 相似于 $J_{r_i}(\lambda_i)$, 即存在非异阵 Q_i , 使得 $J_{r_i}(\lambda_i) = Q_i^{-1}J_{r_i}(\mu_i)^mQ_i = (Q_i^{-1}J_{r_i}(\mu_i)Q_i)^m$, 即结论对 Jordan 块

成立. 令

$$C = \text{diag}\{Q_1^{-1}J_{r_1}(\mu_1)Q_1, Q_2^{-1}J_{r_2}(\mu_2)Q_2, \dots, Q_k^{-1}J_{r_k}(\mu_k)Q_k\},$$

则 $J = C^m$, 即结论对 Jordan 标准型也成立. 最后,

$$A = PJP^{-1} = PC^mP^{-1} = (PCP^{-1})^m.$$

令 $B = PCP^{-1}$, 则有 $A = B^m$, 即结论对一般的矩阵也成立. □

例题 0.6 设 n 阶复方阵 A 的特征多项式为 $f(\lambda)$, 若复系数多项式 $g(\lambda)$ 满足

$$\gcd(f(g(\lambda)), g'(\lambda)) = 1.$$

求证: 存在 n 阶复方阵 B 使得 $g(B) = A$.



笔记 本题是定理 0.4 的推广.

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), \dots, J_{r_k}(\lambda_k)\}$ 为 Jordan 标准型, 我们先对 Jordan 块来证明结论. 任取多项式方程 $g(\lambda) - \lambda_i = 0$ 的根 μ_i , 即 $g(\mu_i) = \lambda_i$, 从而 $f(g(\mu_i)) = f(\lambda_i) = 0$. 由 $f(g(\lambda))$ 与 $g'(\lambda)$ 互素可知它们无公共根, 从而 $g'(\mu_i) \neq 0$. 由命题 ??(5) 知

$$g(J_{r_i}(\mu_i)) = \begin{pmatrix} g(\mu_i) & g'(\mu_i) & \cdots & * \\ & g(\mu_i) & \ddots & \vdots \\ & & \ddots & g'(\mu_i) \\ & & & g(\mu_i) \end{pmatrix},$$

于是 $g(J_{r_i}(\mu_i))$ 的特征值全为 $g(\mu_i) = \lambda_i$, 其几何重数等于 $r_i - r(g(J_{r_i}(\mu_i)) - \lambda_i I) = r_i - (r_i - 1) = 1$. 因此 $g(J_{r_i}(\mu_i))$ 的 Jordan 标准型中只有一个 Jordan 块, 即 $g(J_{r_i}(\mu_i))$ 相似于 $J_{r_i}(\lambda_i)$. 设 Q_i 为非异阵, 使得

$$J_{r_i}(\lambda_i) = Q_i g(J_{r_i}(\mu_i)) Q_i^{-1} = g(Q_i J_{r_i}(\mu_i) Q_i^{-1}),$$

故结论对 Jordan 块成立. 令 $Q = \text{diag}\{Q_1, \dots, Q_k\}, C = \text{diag}\{J_{r_1}(\mu_1), \dots, J_{r_k}(\mu_k)\}$, 则

$$J = \text{diag}\{J_{r_1}(\lambda_1), \dots, J_{r_k}(\lambda_k)\} = Qg(C)Q^{-1} = g(QCQ^{-1}),$$

故结论对 Jordan 标准型也成立. 最后我们有 $A = PJP^{-1} = Pg(QCQ^{-1})P^{-1} = g(PQCQ^{-1}P^{-1})$, 令 $B = PQCQ^{-1}P^{-1}$, 则 $A = g(B)$, 故结论对一般矩阵也成立. □

引理 0.1

(1) 记 $H = \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & \ddots & & \\ 1 & & & \end{pmatrix}$, 则 $H = H^T = H^{-1}$, $HJ_n(a)H = J_n^T(a)$.

(2)

$$J_n(a) = \begin{pmatrix} a & 1 & & \\ & a & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & a \end{pmatrix} = \begin{pmatrix} & & 1 & a \\ & \ddots & a & \\ & & \ddots & \\ 1 & a & & \\ a & & & \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & 1 & & \\ & & \ddots & \\ 1 & & & \end{pmatrix}$$

$$= \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & & 1 & \\ 1 & & & \end{pmatrix} \begin{pmatrix} & & a \\ & a & 1 \\ a & 1 & \\ & & \end{pmatrix}$$

证明 证明都是显然的.

□

命题 0.9

设 A 为 n 阶复矩阵, 证明: 存在 n 阶复对称矩阵 B, C , 使得 $A = BC$, 并且可以指定 B, C 中任何一个为可逆矩阵.

◆

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 考虑 Jordan 块 $J_{r_i}(\lambda_i)$ 的如下两种分解 (列倒排和行倒排):

$$\begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & \ddots \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix} = \begin{pmatrix} & & 1 & \lambda_i \\ & & \ddots & \lambda_i \\ & 1 & \ddots & \\ 1 & \lambda_i & & \\ \lambda_i & & & \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & & 1 & \\ 1 & & & \end{pmatrix} \quad (10)$$

$$= \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & & 1 & \\ 1 & & & \end{pmatrix} \begin{pmatrix} & & \lambda_i \\ & & \lambda_i & 1 \\ & \ddots & \ddots & \\ \lambda_i & 1 & & \end{pmatrix} \quad (11)$$

我们将 (10) 式的分解记为 $J_{r_i}(\lambda_i) = R_i S_i$, (11) 式的分解记为 $J_{r_i}(\lambda_i) = S_i T_i$, 注意到 R_i, S_i, T_i 都是对称矩阵, 并且 S_i 是可逆矩阵. 如果一开始选定 B 为可逆矩阵, 则利用 (11) 式的分解; 如果一开始选定 C 为可逆矩阵, 则利用 (10) 式的分解. 以下不妨设定 B 为可逆矩阵, 令

$$S = \text{diag}\{S_1, S_2, \dots, S_k\}, \quad T = \text{diag}\{T_1, T_2, \dots, T_k\}$$

则有 $J = ST$, 其中 S, T 都是对称矩阵, 并且 S 是可逆矩阵. 因此, 我们有

$$A = PJP^{-1} = PSTP^{-1} = (PSP')((P^{-1})'TP^{-1})$$

令 $B = PSP', C = (P^{-1})'TP^{-1}$, 则 $A = BC$ 即为所求分解.

□

命题 0.10

设 A 为 n 阶复矩阵, 证明: 存在 n 阶非异复对称矩阵 Q , 使得 $Q^{-1}AQ = A'$.

◆

证明 设 P 为非异阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(\lambda_1), J_{r_2}(\lambda_2), \dots, J_{r_k}(\lambda_k)\}$ 为 A 的 Jordan 标准型. 记 r_i 阶矩阵

$$S_i = \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & & 1 & \\ 1 & & & \end{pmatrix},$$

注意到 S_i 是非异对称矩阵, 且 $S_i^2 = I$, 即 $S_i^{-1} = S_i$, 我们来考虑 Jordan 块 $J_{r_i}(\lambda_i)$ 的如下相似关系:

$$J_{r_i}(\lambda_i) = \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & & 1 & \\ 1 & & & \end{pmatrix} \begin{pmatrix} \lambda_i & & & \\ 1 & \lambda_i & & \\ & 1 & \ddots & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda_i \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & \ddots & \\ & & 1 & \\ 1 & & & \end{pmatrix}$$

即有 $J_{r_i}(\lambda_i) = S_i J_{r_i}(\lambda_i)' S_i$. 令 $S = \text{diag}\{S_1, S_2, \dots, S_k\}$, 则 S 是非异对称矩阵, $S^2 = I_n$, 且 $J = S J' S$. 因此, 我们有

$$A = P J P^{-1} = P S J' S P^{-1} = P S P' A' (P^{-1})' S P^{-1} = (P S P') A' (P S P')^{-1}$$

令 $Q = P S P'$, 则 Q 为非异复对称矩阵, 使得 $Q^{-1} A Q = A'$.

□

命题 0.11

设 A 为 n 阶幂零矩阵, 证明: e^A 与 $I_n + A$ 相似.

◆

证明 证法一: 由 A 是幂零矩阵及命题??可知, A 的特征值全为零. 设 P 为非异阵, 使得 $P^{-1}AP = J = \text{diag}\{J_{r_1}(0), J_{r_2}(0), \dots, J_{r_k}(0)\}$ 为 A 的 Jordan 标准型. 先对 Jordan 块 $J_{r_i}(0)$ 进行证明. 注意到

$$\begin{aligned} e^{J_{r_i}(0)} &= I_{r_i} + \frac{1}{1!} J_{r_i}(0) + \frac{1}{2!} J_{r_i}(0)^2 + \dots + \frac{1}{(r_i-1)!} J_{r_i}(0)^{r_i-1} \\ &= \begin{pmatrix} 1 & 1 & \cdots & * \\ & 1 & \ddots & \vdots \\ & & \ddots & 1 \\ & & & 1 \end{pmatrix} \end{aligned}$$

故 $e^{J_{r_i}(0)}$ 的特征值全为 1, 其几何重数等于 $r_i - r(e^{J_{r_i}(0)} - I_{r_i}) = r_i - (r_i - 1) = 1$. 因此 $e^{J_{r_i}(0)}$ 只有一个 Jordan 块, 其 Jordan 标准型为 $J_{r_i}(1) = I_{r_i} + J_{r_i}(0)$, 即存在非异阵 Q_i , 使得 $e^{J_{r_i}(0)} = Q_i(I_{r_i} + J_{r_i}(0))Q_i^{-1}$ ($1 \leq i \leq k$). 再对 Jordan 标准型 J 进行证明. 令 $Q = \text{diag}\{Q_1, Q_2, \dots, Q_k\}$, 则 Q 为非异阵, 满足

$$e^J = \text{diag}\{e^{J_{r_1}(0)}, e^{J_{r_2}(0)}, \dots, e^{J_{r_k}(0)}\} = Q(I_n + J)Q^{-1}$$

最后对一般的矩阵 A 进行证明. 由前两步可得

$$\begin{aligned} e^A &= e^{PJP^{-1}} = P e^J P^{-1} = P Q (I_n + J) Q^{-1} P^{-1} \\ &= P Q (I_n + P^{-1} A P) Q^{-1} P^{-1} = (P Q P^{-1}) (I_n + A) (P Q P^{-1})^{-1} \end{aligned}$$

即 e^A 与 $I_n + A$ 相似.

证法二: 由 A 是幂零矩阵及命题??可知, A 的特征值全为零, 从而 $I_n + A$ 和 $e^A = I_n + A + \dots + \frac{1}{(n-1)!} A^{n-1}$ 的特征值全为 1. 容易验证 $r((e^A - I_n)^k) = r(A^k)$ ($k \geq 1$) 成立, 故由命题??即得结论.

□

0.1.4 采用 Jordan 块作为测试矩阵

在矩阵问题中, 如果需要构造满足某种性质的矩阵, 则可以采用 Jordan 块作为测试矩阵进行探索和讨论. 比如在定理 0.4 的证明中, 为了构造 Jordan 块 $J_{r_i}(\lambda_i)$ 的 m 次方根, 我们采用了 Jordan 块 $J_{r_i}(\mu_i)$ 作为测试矩阵, 并最终得到了正确的答案.

例题 0.7 证明: 存在 71 阶实方阵 A , 使得

$$A^{70} + A^{69} + \cdots + A + I_{71} = \begin{pmatrix} 2019 & 2018 & \cdots & 1949 \\ & 2019 & \ddots & \vdots \\ & & \ddots & 2018 \\ & & & 2019 \end{pmatrix}$$

证明 记 $f(x) = x^{70} + x^{69} + \cdots + x + 1$, 上述等式右边的矩阵为 B . 注意到 $f(1) < 2019$ 和 $f(2) > 2019$, 故由连续函数的性质可知, $f(x) = 2019$ 在开区间 $(1, 2)$ 中必有一实根 λ_0 . 将 Jordan 块 $J_{71}(\lambda_0)$ 代入 $f(x)$ 中, 由命题 ??(5) 计算可得

$$f(J_{71}(\lambda_0)) = \begin{pmatrix} f(\lambda_0) & f'(\lambda_0) & \cdots & * \\ & f(\lambda_0) & \ddots & \vdots \\ & & \ddots & f'(\lambda_0) \\ & & & f(\lambda_0) \end{pmatrix}$$

这是一个上三角矩阵, 主对角元全为 $f(\lambda_0) = 2019$, 上次对角元全为 $f'(\lambda_0) > 0$, 从而 $f(J_{71}(\lambda_0))$ 的特征值全为 2019, 其几何重数为 $71 - r(f(J_{71}(\lambda_0)) - 2019I_{71}) = 1$. 因此, $f(J_{71}(\lambda_0))$ 的 Jordan 标准型中只有一个 Jordan 块 $J_{71}(2019)$, 即 $f(J_{71}(\lambda_0))$ 相似于 $J_{71}(2019)$. 另一方面, 矩阵 B 也是一个上三角矩阵, 主对角元全为 2019, 上次对角元全为 2018, 从而 B 的特征值全为 2019, 其几何重数为 $71 - r(B - 2019I_{71}) = 1$. 因此, B 的 Jordan 标准型中只有一个 Jordan 块 $J_{71}(2019)$, 即 B 也相似于 $J_{71}(2019)$. 由于矩阵的相似在基域扩张下不改变, 故 $f(J_{71}(\lambda_0))$ 和 B 在实数域上相似, 即存在非异实矩阵 P , 使得 $B = P^{-1}f(J_{71}(\lambda_0))P = f(P^{-1}J_{71}(\lambda_0)P)$. 令 $A = P^{-1}J_{71}(\lambda_0)P$, 则 A 是实矩阵, 且满足 $f(A) = B$. □

例题 0.8 设 a, b 都是实数, 其中 $b \neq 0$, 证明: 对任意的正整数 m , 存在四阶实方阵 A , 使得

$$A^m = B = \begin{pmatrix} a & b & 2 & 0 \\ -b & a & 2 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & -b & a \end{pmatrix}$$

注 这个例题采用了广义 Jordan 块作为测试矩阵.

证明 显然, B 的特征多项式 $f(\lambda) = ((\lambda - a)^2 + b^2)^2$. 我们可用 3 种方法求出 B 的 Jordan 标准型 (section: Jordan 标准型的求法). 第一种方法是计算行列式因子:

$$\lambda I_4 - B = \begin{pmatrix} \lambda - a & -b & -2 & 0 \\ b & \lambda - a & -2 & 0 \\ 0 & 0 & \lambda - a & -b \\ 0 & 0 & b & \lambda - a \end{pmatrix}$$

经计算可知

$$(\lambda I_4 - B) \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{pmatrix} = -b((\lambda - a)^2 + b^2), \quad (\lambda I_4 - B) \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} = -2b(\lambda - a + b)$$

显然这两个三阶子式互素, 故三阶行列式因子 $D_3(\lambda) = 1$, 于是 B 的行列式因子组和不变因子组均为 $1, 1, 1, ((\lambda - a)^2 + b^2)^2$, 从而初等因子组为 $(\lambda - a - bi)^2, (\lambda - a + bi)^2$, 因此 B 的 Jordan 标准型 $J = \text{diag}\{J_2(a + bi), J_2(a - bi)\}$.

第二种方法是计算极小多项式: 由于 B 是实方阵, 故其极小多项式 $m(\lambda)$ 是实系数多项式, 又 $m(\lambda)$ 整除 $f(\lambda)$, 从而只能是 $m(\lambda) = (\lambda - a)^2 + b^2$ 或 $m(\lambda) = ((\lambda - a)^2 + b^2)^2$. 通过简单的计算可知 B 不适合多项式 $(\lambda - a)^2 + b^2$, 于

是 $m(\lambda) = f(\lambda) = ((\lambda - a)^2 + b^2)^2$, 剩余的讨论同第一种方法.

第三种方法是计算特征值的几何重数: B 的全体特征值为 $a + bi$ (2 重), $a - bi$ (2 重), 通过简单的计算可知 $r(B - (a + bi)I_4) = 3$ 以及 $r(B - (a - bi)I_4) = 3$, 于是 $a \pm bi$ 的几何重数都等于 1, 从而分别只有一个二阶 Jordan 块, 因此 B 的 Jordan 标准型 $J = \text{diag}\{J_2(a + bi), J_2(a - bi)\}$.

取 $a + bi$ 的 m 次方根 $c + di$ ($c, d \in \mathbb{R}$), 即满足 $(c + di)^m = a + bi$ (取定一个即可). 构造实方阵 (取法不唯一):

$$C = \begin{pmatrix} c & d & 2 & 0 \\ -d & c & 2 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{pmatrix}, \text{ 或 } C = \begin{pmatrix} c & d & 1 & 0 \\ -d & c & 0 & 1 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{pmatrix}, \text{ 或 } C = \begin{pmatrix} c & d & 0 & 0 \\ -d & c & 1 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{pmatrix}$$

注意到 $d \neq 0$, 故由开始处完全类似的讨论可知, C 的 Jordan 标准型为 $\text{diag}\{J_2(c + di), J_2(c - di)\}$. 由命题??可知, $J_2(c \pm di)^m$ 的 Jordan 标准型为 $J_2(a \pm bi)$, 从而 C^m 的 Jordan 标准型为 $\text{diag}\{J_2(a + bi), J_2(a - bi)\}$, 于是 B 与 C^m 有相同的 Jordan 标准型, 故它们在复数域上相似. 注意到 B 与 C^m 都是实矩阵, 故由矩阵的相似在基域扩张下不改变可知, 它们在实数域上也相似, 即存在非异阵 $P \in M_4(\mathbb{R})$, 使得 $B = P^{-1}C^mP = (P^{-1}CP)^m$. 令 $A = P^{-1}CP$, 则 A 为实方阵, 满足 $A^m = B$.

□