

线性变换与相似标准型

秦加俊

2026 年 5 月 7 日、5 月 21 日

1 线性变换

例题 1.1: 可对角化矩阵的稠密性 证明: 复方阵空间 $\mathbb{C}^{n \times n}$ 中的可对角化矩阵全体在通常的欧氏度量下是稠密的。即任意给定矩阵 $A \in \mathbb{C}^{n \times n}$, 都存在可对角化矩阵序列 $\{A_k\}$ 使得 $\lim_{k \rightarrow \infty} A_k = A$ 。

解答. 任取 $A \in \mathbb{C}^{n \times n}$ 。由 Schur 定理 (或复数域上矩阵的三角化定理), 存在可逆矩阵 P 使得 $P^{-1}AP = T$, 其中 T 为上三角矩阵, 且 T 的对角元 $\lambda_1, \dots, \lambda_n$ 为 A 的特征值。

构造上三角矩阵序列 $T_k = T + D_k$, 其中 $D_k = \text{diag}(\delta_1^{(k)}, \dots, \delta_n^{(k)})$ 。由于复数域是无限且稠密的, 我们总可以选取充分小的一组扰动量 $\delta_i^{(k)}$, 使得:

1. 当 $k \rightarrow \infty$ 时, 对于所有的 i , 有 $\delta_i^{(k)} \rightarrow 0$ (即 $D_k \rightarrow O$);
2. 对于每一个固定的 k , 对角元 $\lambda_1 + \delta_1^{(k)}, \dots, \lambda_n + \delta_n^{(k)}$ 互不相同。

因为 T_k 是上三角阵, 其对角元即为它的特征值。由条件 2 可知 T_k 有 n 个互不相同的特征值, 故 T_k 是可对角化的。

令 $A_k = PT_kP^{-1}$ 。由于相似变换不改变可对角化性, 序列 $\{A_k\}$ 中的每一个矩阵都是可对角化的。

最后, 在有限维空间中线性映射具有连续性 (或矩阵乘法关于各项元素是连续的), 当 $k \rightarrow \infty$ 时, $T_k \rightarrow T$, 因此:

$$\lim_{k \rightarrow \infty} A_k = \lim_{k \rightarrow \infty} (PT_kP^{-1}) = P \left(\lim_{k \rightarrow \infty} T_k \right) P^{-1} = PTP^{-1} = A$$

这说明任意 $A \in \mathbb{C}^{n \times n}$ 均可被可对角化矩阵序列逼近, 结论得证。 $\square$

例题 1.2: 对偶空间与子空间覆盖 设 V 是线性空间, $\alpha_1, \alpha_2, \dots, \alpha_n, \beta_1, \beta_2, \dots, \beta_n \in V$ 。若对任意 $f \in V^*$, 有

$$(f(\alpha_1), f(\alpha_2), \dots, f(\alpha_n)) \text{ 和 } (f(\beta_1), f(\beta_2), \dots, f(\beta_n))$$

在不计分量顺序意义下相同。证明: $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 和 $(\beta_1, \beta_2, \dots, \beta_n)$ 在不计分量顺序意义下相同。

解答. 设 S_n 为 n 元对称群 (即所有 n 阶排列的集合)。对于任意一个排列 $\sigma \in S_n$, 定义 V^* 中的子集:

$$V_\sigma = \{f \in V^* \mid f(\alpha_k) = f(\beta_{\sigma(k)}), k = 1, 2, \dots, n\}$$

由于线性函数的加法和数乘保持线性, 极易验证每个 V_σ 都是 V^* 的子空间。

由题设, 对于给定的任意 $f \in V^*$, 由于两个序列在不计顺序时相同, 必定存在一个排列 $\sigma \in S_n$, 使得对所有 $k = 1, 2, \dots, n$, 都有 $f(\alpha_k) = f(\beta_{\sigma(k)})$ 。这等价于说明, 任意 $f \in V^*$ 都至少属于某一个 V_σ 。因此我们将对偶空间写成了有限个子空间的并:

$$V^* = \bigcup_{\sigma \in S_n} V_\sigma$$

由于 V 是定义在数域 (无限域) 上的线性空间, 其对偶空间 V^* 同样是无限域上的线性空间。根据 ** 有限个真子空间不能覆盖全空间 ** 的性质, 这 $n!$ 个子空间中, 至少有一个子空间必须等于全空间 V^* 。

设存在某个固定的排列 $\sigma_0 \in S_n$ 使得 $V_{\sigma_0} = V^*$ 。这意味着对所有的线性函数 $f \in V^*$, 都成立:

$$f(\alpha_k) = f(\beta_{\sigma_0(k)}), \quad k = 1, 2, \dots, n$$

即 $f(\alpha_k - \beta_{\sigma_0(k)}) = 0, \quad \forall f \in V^*$ 。

由于对偶空间 V^* 能分离 V 中的点 (即若向量 $v \neq 0$, 必存在 $f \in V^*$ 使 $f(v) \neq 0$), 从而必须有:

$$\alpha_k - \beta_{\sigma_0(k)} = 0 \implies \alpha_k = \beta_{\sigma_0(k)}, \quad k = 1, 2, \dots, n$$

因此, 向量组 $(\alpha_1, \dots, \alpha_n)$ 和 $(\beta_1, \dots, \beta_n)$ 在相差一个排列 σ_0 的意义下完全相同, 结论得证。□

例题 1.3: 迹为零的矩阵与交换子 设 V 是 n 维复线性空间, $\mathcal{A}: V \rightarrow V$ 是线性变换。如果 $\text{Tr}(\mathcal{A}) = 0$, 证明:

1. 存在 V 的一组基, 使得 $\mathcal{A}$ 在这组基下的矩阵的主对角元都是零;
2. 存在 V 上的线性变换 $\mathcal{B}$ 和 $\mathcal{C}$, 使得 $\mathcal{A} = [\mathcal{B}, \mathcal{C}] = \mathcal{B}\mathcal{C} - \mathcal{C}\mathcal{B}$ 。

解答. 证明 (1): 我们对空间的维数 n 采用数学归纳法。

当 $n = 1$ 时, 由于 $\text{Tr}(\mathcal{A}) = 0$, 其唯一的矩阵元素必然为 0, 结论显然成立。

假设对于 $n - 1$ 维空间结论成立。现考察 n 维空间 V 上的线性变换 $\mathcal{A}$, 且 $\text{Tr}(\mathcal{A}) = 0$ 。

- 若 $\mathcal{A}$ 为纯量变换 (数量乘法), 即存在常数 c 使得 $\mathcal{A} = c\mathcal{I}$ 。则 $\text{Tr}(\mathcal{A}) = nc = 0 \implies c = 0$ 。即 $\mathcal{A} = \mathcal{O}$, 它在任何基下的矩阵均为零矩阵, 对角元全为 0, 结论成立。
- 若 $\mathcal{A}$ 不是纯量变换, 由基本引理可知, 必定存在一个非零向量 $\alpha_1 \in V$, 使得 α_1 与 $\mathcal{A}(\alpha_1)$ 线性无关。

令 $\alpha_2 = \mathcal{A}(\alpha_1)$, 由于 α_1, α_2 线性无关, 我们可以将其扩充为 V 的一组基 $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$ 。在这组基下, 由于 $\mathcal{A}(\alpha_1) = 0\alpha_1 + 1\alpha_2 + 0\alpha_3 + \dots + 0\alpha_n$, 线性变换 $\mathcal{A}$ 的矩阵表示形如:

$$A = \begin{pmatrix} 0 & * \\ \beta & A_1 \end{pmatrix}$$

其中 A_1 是一个 $(n-1) \times (n-1)$ 的矩阵, β 是列向量, $*$ 是行向量。

因为矩阵的迹在基底变换下保持不变, 所以有:

$$\operatorname{Tr}(A) = 0 + \operatorname{Tr}(A_1) = 0 \implies \operatorname{Tr}(A_1) = 0$$

根据归纳假设, 对于 $(n-1)$ 阶迹为零的矩阵 A_1 , 存在可逆矩阵 P_1 , 使得 $B_1 = P_1^{-1}A_1P_1$ 的对角元全为零。

构造 n 阶可逆矩阵 $P = \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix}$, 进行相似变换:

$$P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & P_1^{-1} \end{pmatrix} \begin{pmatrix} 0 & * \\ \beta & A_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P_1 \end{pmatrix} = \begin{pmatrix} 0 & *P_1 \\ P_1^{-1}\beta & P_1^{-1}A_1P_1 \end{pmatrix} = \begin{pmatrix} 0 & *' \\ \beta' & B_1 \end{pmatrix}$$

由于 B_1 的对角元全为零, 上述分块矩阵的主对角元也全为零。这就证明了存在一组新的基, 使得 $\mathcal{A}$ 的矩阵表示的主对角元全是零。

证明 (2): 由第 (1) 问可知, 我们可以选取 V 的一组基, 使得 $\mathcal{A}$ 在此基下的矩阵表示 $A = (a_{ij})_{n \times n}$ 满足对于所有的 $i = 1, 2, \dots, n$, 都有 $a_{ii} = 0$ 。

我们试图在这组基下寻找矩阵 B 和 C , 使得 $A = BC - CB$ 。为了简化计算, 我们构造 B 为对角阵, 即 $B = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, 并且要求所有的 λ_i 互不相同 (即当 $i \neq j$ 时, $\lambda_i \neq \lambda_j$)。

设 $C = (c_{ij})_{n \times n}$ 。考察矩阵 $BC - CB$ 的第 (i, j) 个元素: 由于 B 是对角阵, 左乘 B 相当于用 λ_i 乘 C 的第 i 行, 右乘 B 相当于用 λ_j 乘 C 的第 j 列。因此:

$$(BC - CB)_{ij} = \lambda_i c_{ij} - c_{ij} \lambda_j = (\lambda_i - \lambda_j) c_{ij}$$

我们需要让 $BC - CB = A$, 即对所有 i, j , 满足:

$$(\lambda_i - \lambda_j) c_{ij} = a_{ij}$$

- 当 $i = j$ 时, 等式左边为 $0 \cdot c_{ii} = 0$; 由于 A 的对角元 $a_{ii} = 0$, 等式自然成立, 我们可以任意取 c_{ii} (例如取 $c_{ii} = 0$)。
- 当 $i \neq j$ 时, 由于我们已经假定 $\lambda_i \neq \lambda_j$, 我们可以直接解出 c_{ij} :

$$c_{ij} = \frac{a_{ij}}{\lambda_i - \lambda_j}$$

通过上述构造, 我们找到了满足要求的矩阵 B 和 C 。将 B 和 C 对应回线性空间 V 上的线性变换 $\mathcal{B}$ 和 $\mathcal{C}$, 即满足 $\mathcal{A} = \mathcal{B}\mathcal{C} - \mathcal{C}\mathcal{B}$, 结论得证。□

例题 1.4: 矩阵秩的可加性 设矩阵 $A, B \in \mathbb{F}^{m \times n}$, 证明: $R(A+B) = R(A) + R(B)$ 的充分必要条件是, 存在可逆方阵 $P \in \mathbb{F}^{m \times m}$ 和可逆方阵 $Q \in \mathbb{F}^{n \times n}$, 使得

$$PAQ = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}, \quad PBQ = \begin{pmatrix} O & O \\ O & E_s \end{pmatrix},$$

其中 $r = R(A)$, $s = R(B)$, 且 $r + s \leq \min\{m, n\}$ 。

解答. 充分性 ($\Leftarrow$): 若 A, B 可同时相抵化为给定的分块形式, 则

$$P(A+B)Q = PAQ + PBQ = \begin{pmatrix} E_r & O \\ O & E_s \end{pmatrix}.$$

由于 P, Q 是可逆方阵, 矩阵乘法不改变秩, 因此

$$R(A+B) = R(P(A+B)Q) = r+s = R(A) + R(B).$$

必要性 ($\Rightarrow$): 将 A, B 视为从线性空间 $V = \mathbb{F}^n$ 到 $U = \mathbb{F}^m$ 的线性映射 $\mathcal{A}, \mathcal{B}$. 已知 $R(A+B) = R(A) + R(B)$, 即 $\dim(\operatorname{Im}(\mathcal{A} + \mathcal{B})) = \dim(\operatorname{Im}\mathcal{A}) + \dim(\operatorname{Im}\mathcal{B})$. 由于 $\operatorname{Im}(\mathcal{A} + \mathcal{B}) \subseteq \operatorname{Im}\mathcal{A} + \operatorname{Im}\mathcal{B}$, 空间维数相加直接意味着像空间成直和:

$$\operatorname{Im}(\mathcal{A} + \mathcal{B}) = \operatorname{Im}\mathcal{A} \oplus \operatorname{Im}\mathcal{B} \implies \operatorname{Im}\mathcal{A} \cap \operatorname{Im}\mathcal{B} = \{0\}.$$

考察核空间: 显然 $\operatorname{Ker}\mathcal{A} \cap \operatorname{Ker}\mathcal{B} \subseteq \operatorname{Ker}(\mathcal{A} + \mathcal{B})$. 由秩-零度定理, $n - R(A+B) = n - (r+s)$. 计算交空间的维数:

$$\begin{aligned} \dim(\operatorname{Ker}\mathcal{A} \cap \operatorname{Ker}\mathcal{B}) &= \dim(\operatorname{Ker}\mathcal{A}) + \dim(\operatorname{Ker}\mathcal{B}) - \dim(\operatorname{Ker}\mathcal{A} + \operatorname{Ker}\mathcal{B}) \\ &\geq (n-r) + (n-s) - n = n-r-s. \end{aligned}$$

这表明 $\dim(\operatorname{Ker}\mathcal{A} \cap \operatorname{Ker}\mathcal{B}) \geq \dim(\operatorname{Ker}(\mathcal{A} + \mathcal{B}))$, 故必定有 $\operatorname{Ker}\mathcal{A} \cap \operatorname{Ker}\mathcal{B} = \operatorname{Ker}(\mathcal{A} + \mathcal{B})$ 且 $\operatorname{Ker}\mathcal{A} + \operatorname{Ker}\mathcal{B} = V$.

令 $V_0 = \operatorname{Ker}\mathcal{A} \cap \operatorname{Ker}\mathcal{B}$. 我们可以选取子空间 $V_1 \subseteq \operatorname{Ker}\mathcal{B}$ 使得 $\operatorname{Ker}\mathcal{B} = V_1 \oplus V_0$ (此时 $\dim V_1 = r$); 选取子空间 $V_2 \subseteq \operatorname{Ker}\mathcal{A}$ 使得 $\operatorname{Ker}\mathcal{A} = V_2 \oplus V_0$ (此时 $\dim V_2 = s$). 从而定义域 V 具有直和分解: $V = V_1 \oplus V_2 \oplus V_0$.

在 V 中取顺应上述直和分解的基: $\alpha_1, \dots, \alpha_r$ (对应 V_1), $\beta_1, \dots, \beta_s$ (对应 V_2), $\gamma_1, \dots, \gamma_{n-r-s}$ (对应 V_0). 在像空间 U 中, 由于 $\mathcal{A}$ 在 V_1 上是单射, $\{\mathcal{A}(\alpha_i)\}$ 是 $\operatorname{Im}\mathcal{A}$ 的基; 同理 $\{\mathcal{B}(\beta_j)\}$ 是 $\operatorname{Im}\mathcal{B}$ 的基. 因像空间直和, $\{\mathcal{A}(\alpha_i)\} \cup \{\mathcal{B}(\beta_j)\}$ 线性无关. 将其扩充为 U 的一组基.

在这两组基下, 映射 $\mathcal{A}$ 将基底 V_1 映为 $\operatorname{Im}\mathcal{A}$, 将 V_2, V_0 映为 0; 映射 $\mathcal{B}$ 将基底 V_2 映为 $\operatorname{Im}\mathcal{B}$, 将 V_1, V_0 映为 0. 设 Q 为 V 中标准基到新基的过渡矩阵, P^{-1} 为 U 中标准基到新基的过渡矩阵, 则 PAQ 和 PBQ 恰为题目所求的分块矩阵形式. $\square$

例题 1.5: AB 与 BA 相似的秩条件 设方阵 $A, B \in \mathbb{F}^{n \times n}$ 满足 $R(ABA) = R(B)$. 证明: AB 和 BA 相似.

解答. 首先, 根据矩阵乘法秩的不等式, 有:

$$R(B) \geq R(AB) \geq R(ABA)$$

$$R(B) \geq R(BA) \geq R(ABA)$$

由题设 $R(ABA) = R(B)$ 夹逼可得:

$$R(AB) = R(BA) = R(B)$$

为了证明 $AB \sim BA$ ，我们需要证明它们关于零特征值的初等因子相同（由前置定理，非零特征值的初等因子本就相同）。这等价于证明对于任意正整数 $k \geq 1$ ，它们的幂次秩相等： $R((AB)^k) = R((BA)^k)$ 。

将矩阵视为有限维空间 $V = \mathbb{F}^n$ 上的线性变换。我们深入挖掘秩相等的几何意义：

1. 由 $R(AB) = R(B)$ ，即 $\dim(A(\text{Im}B)) = \dim(\text{Im}B)$ 。这说明线性变换 A 在子空间 $\text{Im}B$ 上是**单射**（即 $\text{Ker}A \cap \text{Im}B = \{0\}$ ）。这意味着 A 作用在 $\text{Im}B$ 的任何子空间上，都不改变该子空间的维数。
2. 由 $R(BA) = R(B)$ ，即 $\dim(\text{Im}(BA)) = \dim(\text{Im}B)$ 。因为 $\text{Im}(BA) = B(\text{Im}A) \subseteq B(V) = \text{Im}B$ ，且两者维数相同，故必须有 $\text{Im}(BA) = \text{Im}B$ 。

现在考察 $k \geq 2$ 时的像空间：对于 $(BA)^k$ ，其像空间为：

$$\text{Im}((BA)^k) = (BA)^{k-1}(BA(V)) = (BA)^{k-1}(\text{Im}(BA))$$

代入刚才得到的 $\text{Im}(BA) = \text{Im}B$ ，得到：

$$\text{Im}((BA)^k) = (BA)^{k-1}(\text{Im}B)$$

记这个子空间为 $S_k = (BA)^{k-1}(\text{Im}B)$ 。由于 B 的像必然在 $\text{Im}B$ 中，显然 $S_k \subseteq \text{Im}B$ 。此时 $R((BA)^k) = \dim(S_k)$ 。

对于 $(AB)^k$ ，我们提取最左侧的 A ，其像空间为：

$$\text{Im}((AB)^k) = A \left(B(AB)^{k-1}(V) \right) = A \left((BA)^{k-1}B(V) \right) = A \left((BA)^{k-1}(\text{Im}B) \right) = A(S_k)$$

核心在此：因为 $S_k \subseteq \text{Im}B$ ，而 A 限制在 $\text{Im}B$ 上是**单射**，所以 A 作用在 S_k 上保持维数不变，即：

$$R((AB)^k) = \dim(A(S_k)) = \dim(S_k) = R((BA)^k)$$

综上所述，对于任意正整数 $k \geq 1$ ，均有 $R((AB)^k) = R((BA)^k)$ 。因此， AB 与 BA 的秩退化序列完全一致，它们关于特征值 0 的 Jordan 块分布相同。故 $AB \sim BA$ 成立。□

例题 1.6: 可交换与幂零矩阵的行列式 设 n 阶复方阵 A 和 B 可交换，且满足 $B^n = O$ 。
证明： $|A+B| = |A|$ 。

解答. 情形一：当 A 可逆时

因为 A 可逆，我们可以将 $A+B$ 提取公因式写为：

$$A+B = A(I + A^{-1}B)$$

由于 A 与 B 可交换 ($AB = BA$)，两边同乘 A^{-1} 易得 A^{-1} 与 B 也是可交换的 ($A^{-1}B = BA^{-1}$)。因此，矩阵 $A^{-1}B$ 的幂次可以自由展开：

$$(A^{-1}B)^n = (A^{-1})^n B^n = O$$

这说明 $A^{-1}B$ 也是一个幂零矩阵。

幂零矩阵的所有特征值均为 0。因此, 矩阵 $I + A^{-1}B$ 的特征值全为 $1 + 0 = 1$ 。根据行列式等于特征值之积, 有 $|I + A^{-1}B| = 1$ 。取行列式得:

$$|A + B| = |A(I + A^{-1}B)| = |A| \cdot |I + A^{-1}B| = |A| \cdot 1 = |A|$$

情形二: 当 A 不可逆时

考虑引入辅助多项式矩阵 $A(t) = A + tI$ 。设 $f(t) = |A(t) + B| - |A(t)|$, 显然 $f(t)$ 是关于 t 的多项式。

由于 A 的特征值至多只有 n 个, 因此存在无穷多个 $t \in \mathbb{C}$ 使得 $A + tI$ 是可逆矩阵。又因为 $(A + tI)B = AB + tB = BA + tB = B(A + tI)$, 即 $A(t)$ 与 B 可交换。

对于这无穷多个使 $A(t)$ 可逆的 t , 由情形一的结论可知:

$$|A(t) + B| = |A(t)| \implies f(t) = 0$$

一个次数不超过 n 的多项式 $f(t)$ 在无穷多个点处取值为 0, 则 $f(t)$ 必为恒零多项式, 即对于任意复数 t 均有 $f(t) \equiv 0$ 。

特别地, 令 $t = 0$, 即得:

$$|A + B| = |A|$$

本题得证。 □

例题 1.7: 正特征值矩阵的平方等同性 设 n 阶方阵 A 和 B 的特征值全部大于零且 $A^2 = B^2$, 证明: $A = B$ 。

解答. 令 $C = A^2 = B^2$ 。设 C 的极小多项式为 $m(x)$ 。因为 A 的所有特征值大于零, 设 A 的特征值为 $\lambda_i > 0$, 则 C 的对应特征值为 $\mu_i = \lambda_i^2 > 0$ 。这说明 $\lambda_i = \sqrt{\mu_i}$ 是由 μ_i 唯一确定的正平方根。

我们可以利用拉格朗日-西尔维斯特 (Lagrange-Sylvester) 矩阵多项式插值理论来构造一个特殊的二次连续可微函数在矩阵上的多项式表示。具体而言, 由于极小多项式 $m(x)$ 的根全部大于 0, 存在一个多项式 $f(x)$, 使得在 C 的每个特征值 μ_i 处, $f(x)$ 满足:

$$f(\mu_i) = \sqrt{\mu_i}$$

且如果 μ_i 在 $m(x)$ 中的重数 $k > 1$, 则 $f(x)$ 在 μ_i 处的前 $k - 1$ 阶导数值与函数 $g(x) = \sqrt{x}$ 的对应导数值严格相等。

根据矩阵多项式的性质, 这样的多项式 $f(x)$ 就是矩阵 C 的主平方根, 它使得 $[f(C)]^2 = C$, 并且 $f(C)$ 的所有特征值均为正 (即 $\sqrt{\mu_i}$)。又因为 $A^2 = C$, 且 A 的特征值全部为正, 由标准型理论可知 A 本身就是这个唯一的主平方根, 即 A 可以被表示为 C 的多项式:

$$A = f(C)$$

同理, 因为 B 的特征值也全部大于零, 且 $B^2 = C$, 其对应的谱结构和特征值与 A 完全相同。应用同一个插值多项式 $f(x)$, 必然也有:

$$B = f(C)$$

综上所述, $A = f(C) = B$, 结论得证。 □

例题 1.8: 有理数域上的周期变换与维数 设 V 是有理数域 $\mathbb{Q}$ 上的有限维线性空间, p 是一个素数, $\mathcal{A}$ 是 V 上的线性变换, 满足 $\mathcal{A}^p = \mathcal{E}$, 且对任意非零的 $\alpha \in V$, α 和 $\mathcal{A}(\alpha)$ 线性无关。证明: $(p-1) \mid \dim V$ 。

解答. 由 $\mathcal{A}^p = \mathcal{E}$ 可知, 线性变换 $\mathcal{A}$ 满足多项式 $g(x) = x^p - 1 = 0$ 。在有理数域 $\mathbb{Q}$ 上, $x^p - 1$ 可以分解为:

$$x^p - 1 = (x - 1)(x^{p-1} + x^{p-2} + \cdots + x + 1)$$

设 $\Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + 1$ 。因为 p 是素数, 由艾森斯坦因 (Eisenstein) 判别法可知, $\Phi_p(x)$ 是 $\mathbb{Q}$ 上的分圆多项式, 它是不可约的。

由题设, 对任意非零向量 $\alpha \in V$, α 与 $\mathcal{A}(\alpha)$ 线性无关。这说明对于任意 $\alpha \neq 0$, $\mathcal{A}(\alpha) \neq \alpha$, 即 $(\mathcal{A} - \mathcal{E})\alpha \neq 0$ 。因此 $\lambda = 1$ 不是 $\mathcal{A}$ 的特征值, 一次多项式 $(x - 1)$ 不能整除 $\mathcal{A}$ 的极小多项式 $m(x)$ 。

因为 $m(x)$ 必须整除化零多项式 $g(x) = (x - 1)\Phi_p(x)$, 且不包含因子 $(x - 1)$, 所以 $m(x)$ 必须整除 $\Phi_p(x)$ 。又由于 $\Phi_p(x)$ 在 $\mathbb{Q}$ 上不可约, 且 $\mathcal{A} \neq \mathcal{E}$, 故极小多项式只能是:

$$m(x) = \Phi_p(x) = x^{p-1} + x^{p-2} + \cdots + 1$$

设 $\mathcal{A}$ 的特征多项式为 $f(x)$ 。根据线性代数基本定理, 特征多项式的不可约因式集合与极小多项式的不可约因式集合完全相同。在 $\mathbb{Q}$ 上, 由于 $m(x)$ 唯一包含的不可约因式是 $\Phi_p(x)$, 所以特征多项式 $f(x)$ 必定是 $\Phi_p(x)$ 的正整数次幂:

$$f(x) = [\Phi_p(x)]^k, \quad k \in \mathbb{N}^*$$

比较多项式两边的次数: 特征多项式的次数等于空间的维数 $\dim V$, 而 $[\Phi_p(x)]^k$ 的次数为 $k(p-1)$ 。因此有:

$$\dim V = k(p-1)$$

即 $(p-1) \mid \dim V$, 命题得证。 □

例题 1.9: 矩阵等式 $A^2 = A^T$ 的特征值探索 求集合 $S = \{\lambda \in \mathbb{C} \mid \text{存在 } n \in \mathbb{N}, \text{ 存在 } n \text{ 阶实矩阵 } A, \text{ 使得 } A^2 = A^T, \text{ 且 } \lambda \text{ 为 } A \text{ 的特征值}\}$ 。

解答. 设 A 是满足 $A^2 = A^T$ 的实矩阵。对等式两边取转置, 得 $(A^2)^T = (A^T)^T \implies (A^T)^2 = A$ 。将原式 $A^T = A^2$ 整体代入上式, 可得:

$$(A^2)^2 = A \implies A^4 = A \implies A(A^3 - I) = O$$

这说明矩阵 A 满足多项式 $f(x) = x^4 - x = x(x-1)(x^2 + x + 1) = 0$ 。矩阵的特征值必须是其化零多项式的根, 因此解得候选特征值为:

$$\lambda \in \left\{ 0, 1, -\frac{1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i \right\}$$

接下来我们需要分别构造实矩阵, 证明这些候选特征值确实都能被取得:

- 取 $A = (0)_{1 \times 1}$ (1 阶零矩阵), 满足 $A^2 = A^T = 0$, 此时特征值为 0。
- 取 $A = (1)_{1 \times 1}$ (1 阶单位阵), 满足 $A^2 = A^T = 1$, 此时特征值为 1。
- 为了实现共轭复数特征值, 我们需要在 $\mathbb{R}^{2 \times 2}$ 中寻找。注意到复根是 $e^{\pm i \frac{2\pi}{3}}$, 我们自然联想到旋转矩阵。取旋转角 $\theta = \frac{2\pi}{3}$ 的旋转矩阵:

$$A = \begin{pmatrix} \cos \frac{2\pi}{3} & -\sin \frac{2\pi}{3} \\ \sin \frac{2\pi}{3} & \cos \frac{2\pi}{3} \end{pmatrix} = \begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$$

它的转置为 $A^T = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix}$ 。同时, A^2 即为旋转 $2\theta = \frac{4\pi}{3}$ 的矩阵:

$$A^2 = \begin{pmatrix} \cos \frac{4\pi}{3} & -\sin \frac{4\pi}{3} \\ \sin \frac{4\pi}{3} & \cos \frac{4\pi}{3} \end{pmatrix} = \begin{pmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{pmatrix}$$

显然 $A^2 = A^T$ 成立, 且该旋转矩阵的特征值恰好为 $-\frac{1}{2} \pm \frac{\sqrt{3}}{2}i$ 。

综上所述, 所有候选根均能被实矩阵实现。故所求集合为:

$$S = \left\{ 0, 1, -\frac{1}{2} + \frac{\sqrt{3}}{2}i, -\frac{1}{2} - \frac{\sqrt{3}}{2}i \right\}$$

□

2 Jordan 标准型

知识 2.1: 矩阵相似的判定与证明方法总结

在高等代数中, 证明两个 n 阶方阵 A 与 B 相似 ($A \sim B$) 通常有以下五种核心维度:

1. **定义法 (构造过渡矩阵):** 直接寻找或构造可逆矩阵 P , 验证 $P^{-1}AP = B$; 或者将 A, B 视为同一线性变换在不同基底下的矩阵表示。
2. **不变量判定 ($D_{n-1}(\lambda) = 1$ 的特殊情形):** 若方阵的第 $n-1$ 阶行列式因子 $D_{n-1}(\lambda) = 1$ (这等价于极小多项式等于特征多项式, 即矩阵为不可逆减/循环矩阵), 则它必然相似于其特征多项式对应的友矩阵。若 A, B 特征多项式相同且均满足此条件, 则 $A \sim B$ 。
3. **极小多项式与对角化:** 利用极小多项式判断矩阵的可对角化性。若 A 与 B 的极小多项式均无重根 (即均可对角化), 且两者的特征多项式 (或各特征值的代数重数) 完全相同, 则 $A \sim B$ 。
4. **秩的二阶差分与 Jordan 块计算 (Weyl 准则):** 对于特征值 λ , 其对应阶数恰好为 k 的 Jordan 块的数量为:

$$m_k = R((\lambda I - A)^{k-1}) - 2R((\lambda I - A)^k) + R((\lambda I - A)^{k+1})$$

若 A 与 B 对于所有特征值的上述秩序列完全相同, 则相似。(特例速判: 若 $n - R(\lambda I - A) = 1$, 则说明特征值 λ 对应的几何重数为 1, 即只有一个 Jordan 块。)

5. **初等因子体系与分块观察法:** $A \sim B$ 的充分必要条件是它们具有完全相同的初等因子体系。在实战中, 若矩阵可划分为对角分块矩阵, 则其整体的初等因子集合即为各个对角子块初等因子的并集。

例题 2.1: 反对角矩阵的 Jordan 标准形 求下列复矩阵的 Jordan 标准形:

$$A = \begin{pmatrix} & & a_1 \\ & & a_2 \\ & \ddots & \\ a_n & & \end{pmatrix}_{n \times n}, \quad (a_1 a_2 \cdots a_n \neq 0).$$

解答. 考察矩阵 A 对 $\mathbb{C}^n$ 的标准基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 的作用。根据矩阵乘法, 易知:

$$A\varepsilon_i = a_{n+1-i}\varepsilon_{n+1-i}, \quad i = 1, 2, \dots, n$$

这启示我们将全空间 $V = \mathbb{C}^n$ 分解为若干个低维的 A -不变子空间的直和:

- 当 $i \neq n+1-i$ 时 (即 $1 \leq i \leq \lfloor \frac{n}{2} \rfloor$), 子空间 $W_i = \text{span}\{\varepsilon_i, \varepsilon_{n+1-i}\}$ 是一个 2 维的 A -不变

子空间。\$A\$ 限制在 \$W_i\$ 上的矩阵表示为：

$$A_i = \begin{pmatrix} 0 & a_i \\ a_{n+1-i} & 0 \end{pmatrix}$$

其特征多项式为 \$\lambda^2 - a_i a_{n+1-i}\$。由于已知 \$a_1 a_2 \cdots a_n \neq 0\$，故常数项不为零。因此 \$A_i\$ 有两个互异的非零复特征值 \$\pm \sqrt{a_i a_{n+1-i}}\$。这说明 \$A_i\$ 在复数域上可对角化，其 Jordan 标准形为对角阵 \$\text{diag}(\sqrt{a_i a_{n+1-i}}, -\sqrt{a_i a_{n+1-i}})\$。

- 如果 \$n\$ 为奇数，还存在中间的基向量，此时 \$i = \frac{n+1}{2}\$。子空间 \$W_i = \text{span}\{\varepsilon_{\frac{n+1}{2}}\}\$ 是一个 1 维不变子空间。此时 \$A\varepsilon_i = a_{\frac{n+1}{2}}\varepsilon_i\$，即 \$a_{\frac{n+1}{2}}\$ 本身就是一个特征值。

由于 \$A\$ 在全空间上可以分解为这些互不相交的 1 维或 2 维不变子空间的直和，且限制在每个子空间上 \$A\$ 都可对角化，所以 **整体矩阵 \$A\$ 必然可对角化**（即不存在阶数大于 1 的 Jordan 块）。

综上所述，\$A\$ 的 Jordan 标准形即为其特征值构成的对角矩阵（排列顺序不限）：

- 若 \$n = 2m\$ 为偶数：

$$J = \text{diag}(\sqrt{a_1 a_n}, -\sqrt{a_1 a_n}, \dots, \sqrt{a_m a_{m+1}}, -\sqrt{a_m a_{m+1}})$$

- 若 \$n = 2m + 1\$ 为奇数：

$$J = \text{diag}(\sqrt{a_1 a_n}, -\sqrt{a_1 a_n}, \dots, \sqrt{a_m a_{m+2}}, -\sqrt{a_m a_{m+2}}, a_{m+1})$$

□

例题 2.2: 特殊上三角矩阵的 **Jordan 标准型** 求下列 \$n\$ 阶矩阵的 Jordan 标准型：

$$A = \begin{pmatrix} n & n-1 & n-2 & \cdots & 1 \\ 0 & n & n-1 & \cdots & 2 \\ 0 & 0 & n & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \end{pmatrix}$$

解答. 由于 \$A\$ 是上三角矩阵，其主对角线上的元素即为其特征值。显然，矩阵 \$A\$ 只有一个 \$n\$ 重特征值 \$\lambda_0 = n\$。

为了确定 \$\lambda_0 = n\$ 对应的 Jordan 块的数量与结构，我们需要考察特征矩阵的秩退化情况，即计算秩 \$r(A - nI)\$。将 \$A\$ 的主对角线元素减去 \$n\$，得到：

$$A - nI = \begin{pmatrix} 0 & n-1 & n-2 & \cdots & 1 \\ 0 & 0 & n-1 & \cdots & 2 \\ 0 & 0 & 0 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

观察矩阵 $A - nI$ 的结构: 除了第一列全为 0 之外, 去掉第一列和最后一行后, 剩下的右上角部分是一个 $(n-1) \times (n-1)$ 的上三角子矩阵。该子矩阵的主对角线恰好是原矩阵的次对角线 (即超对角线), 其元素全为 $n-1$ 。

因为对于 n 阶方阵 (当 $n \geq 2$ 时), $n-1 \neq 0$ 。所以这个 $(n-1) \times (n-1)$ 的上三角子矩阵的对角元均不为零, 其行列式不为零, 即该子矩阵满秩。

因此, 矩阵 $A - nI$ 的秩为:

$$r(A - nI) = n - 1$$

根据特征值对应 Jordan 块的数量公式 (即特征子空间的维数/几何重数):

$$\text{Jordan 块的数量} = n - r(A - nI) = n - (n - 1) = 1$$

这说明特征值 $\lambda_0 = n$ 只对应唯一的一个 Jordan 块。既然只有一个块, 那么这个块的阶数必定等于特征值的代数重数 n 。

综上所述, 矩阵 A 的 Jordan 标准型是一个完整的 n 阶 Jordan 块:

$$J = J_n(n) = \begin{pmatrix} n & 1 & 0 & \cdots & 0 \\ 0 & n & 1 & \cdots & 0 \\ 0 & 0 & n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \end{pmatrix}_{n \times n}$$

(注: 若 $n = 1$, 则 $A = (1)$, 其 Jordan 标准型即为其本身, 结论依然成立)。□

例题 2.3: AB 与 BA 的 Jordan 块分布 设 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{n \times m}$ 。证明: 矩阵 AB 与 BA 的非零特征值完全相同; 且对于任何 $\lambda \neq 0$, 其在 AB 与 BA 的 Jordan 标准型中对应的 Jordan 块分布完全相同。

解答. 构造 $(m+n)$ 阶分块矩阵 $M_1 = \begin{pmatrix} AB & 0 \\ B & 0 \end{pmatrix}$ 与 $M_2 = \begin{pmatrix} 0 & 0 \\ B & BA \end{pmatrix}$ 。

定义可逆分块矩阵 $P = \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix}$, 易证 $P^{-1}M_1P = M_2$ 。由于 $M_1 \sim M_2$, 它们的初等因子体系完全相同。

考察 M_1 的特征矩阵 $\lambda I - M_1 = \begin{pmatrix} \lambda I_m - AB & 0 \\ -B & \lambda I_n \end{pmatrix}$ 。对于多项式环 $\mathbb{F}[\lambda]$ 中任意不等于 λ 的不可约多项式 $p(\lambda)$, 由于右下角矩阵块的行列式 $\det(\lambda I_n) = \lambda^n$ 与 $p(\lambda)$ 互素, 故右下角分块不贡献任何关于 $p(\lambda)$ 的初等因子。因此, M_1 中关于 $p(\lambda)$ 的初等因子完全等同于左上角分块 $\lambda I_m - AB$ 的初等因子。

同理, 考察 M_2 的特征矩阵 $\lambda I - M_2 = \begin{pmatrix} \lambda I_m & 0 \\ -B & \lambda I_n - BA \end{pmatrix}$, 其关于 $p(\lambda)$ 的初等因子完全等同于右下角分块 $\lambda I_n - BA$ 的初等因子。

综上所述, AB 与 BA 关于任何不可约多项式 $p(\lambda) \neq \lambda$ 的初等因子完全一致。这就意味着, 对于任何非零特征值, 它们在 Jordan 标准型中所对应的 Jordan 块分布完全相同。□

例题 2.4: Weyl 矩阵相似准则 (Weyl) 证明: n 阶复方阵 A 和 B 相似的充分必要条件是, 对于每个复数 a 和每个正整数 k , $R(aE_n - A)^k = R(aE_n - B)^k$ 。

解答. 必要性 ($\Rightarrow$): 若 A 与 B 相似, 则存在可逆方阵 P 使得 $A = PBP^{-1}$ 。对于任意复数 a 和正整数 k , 有:

$$(aE_n - A)^k = (aE_n - PBP^{-1})^k = [P(aE_n - B)P^{-1}]^k = P(aE_n - B)^k P^{-1}.$$

由于 P 及其逆矩阵均可逆, 矩阵乘以可逆阵不改变其秩, 故直接有:

$$R(aE_n - A)^k = R(aE_n - B)^k.$$

充分性 ($\Leftarrow$): 在复数域 $\mathbb{C}$ 上, 两个矩阵相似当且仅当它们具有完全相同的 Jordan 标准型。而矩阵的 Jordan 标准型由其特征值以及对应的各阶 Jordan 块的数量唯一确定。

对于任意固定的复数 $a = \lambda_0$, 考察矩阵 $N_A = \lambda_0 E_n - A$ 的幂次的序列 $\{R(N_A^k)\}_{k=0}^{\infty}$ (约定 $R(N_A^0) = n$)。根据 Jordan 标准型的结构, 特征值 λ_0 对应的阶数大于等于 k 的 Jordan 块的数量为相邻秩的差值:

$$d_k = R(N_A^{k-1}) - R(N_A^k).$$

进而, 特征值 λ_0 对应的阶数恰好为 k 的 Jordan 块的数量为二阶差分:

$$m_k(A) = d_k - d_{k+1} = R(N_A^{k-1}) - 2R(N_A^k) + R(N_A^{k+1}).$$

根据题设条件, 对于任意复数 a 和任意正整数 k , 都有 $R(aE_n - A)^k = R(aE_n - B)^k$ 。(特别地, 当 $k = n$ 时, 通过检验 $R(aE_n - A)^n < n$ 是否成立, 可知 A 和 B 具有完全相同的特征值集合)。

因此, 对于任意特征值 λ_0 , 计算上述二阶差分必定有:

$$m_k(A) = m_k(B), \quad \forall k \in \mathbb{Z}^+.$$

这说明 A 与 B 的所有特征值及其对应的各阶 Jordan 块的数量完全相同。故 A 与 B 的 Jordan 标准型相同, 从而 $A \sim B$ 。□

例题 2.5: Jordan 块个数的秩公式 设 λ_0 是 n 阶矩阵 A 的特征值, 证明: 对任意的正整数 k , 特征值为 λ_0 的 k 阶 Jordan 块 $J_k(\lambda_0)$ 在 A 的 Jordan 标准型 J 中出现的个数为

$$r((A - \lambda_0 I_n)^{k-1}) + r((A - \lambda_0 I_n)^{k+1}) - 2r((A - \lambda_0 I_n)^k),$$

其中约定 $r((A - \lambda_0 I_n)^0) = n$ 。

解答. 因为矩阵的秩在相似变换下保持不变, 且 A 相似于其 Jordan 标准型 J , 故对于任意非负整数 m , 有:

$$r((A - \lambda_0 I_n)^m) = r((J - \lambda_0 I_n)^m)$$

考察 $J - \lambda_0 I_n$ 的对角块结构。对于特征值 $\lambda_i \neq \lambda_0$ 的 Jordan 块 $J_s(\lambda_i)$, 由于 $\lambda_i - \lambda_0 \neq 0$, 块 $J_s(\lambda_i) - \lambda_0 I_s$ 及其任意次幂均是可逆矩阵, 其秩恒等于该块的阶数 s 。因此, 对于这种块, 相邻次幂的秩的差值为 $s - s = 0$, 它对秩的衰减没有贡献。

真正对秩的衰减产生贡献的, 只有对应于特征值 λ_0 的 Jordan 块。对于单个 s 阶幂零 Jordan 块 $N_s = J_s(\lambda_0) - \lambda_0 I_s$, 其幂次的秩满足极有规律的阶梯式下降: 当 $m \leq s$ 时, 每自乘一次, 对角线上非零的 1 就向右上角平移一格, 秩恰好下降 1; 当 $m \geq s$ 时, 矩阵变为零矩阵, 秩保持为 0。即:

$$r(N_s^{m-1}) - r(N_s^m) = \begin{cases} 1, & m \leq s \\ 0, & m > s \end{cases}$$

令 $d_m = r((A - \lambda_0 I_n)^{m-1}) - r((A - \lambda_0 I_n)^m)$ 。根据分块对角矩阵的秩等于各对角块秩之和的性质, 对全矩阵求 d_m 时, 只有阶数大于等于 m 的 λ_0 -Jordan 块会贡献 1 的差值。所以, d_m 的确切代数意义是: 在 J 中, 阶数大于或等于 m 的 λ_0 -Jordan 块的数量。

因此, 阶数恰好为 k 的 λ_0 -Jordan 块的数量, 等于“阶数大于等于 k 的块数”减去“阶数大于等于 $k+1$ 的块数”。即:

$$\begin{aligned} m_k &= d_k - d_{k+1} \\ &= [r((A - \lambda_0 I_n)^{k-1}) - r((A - \lambda_0 I_n)^k)] - [r((A - \lambda_0 I_n)^k) - r((A - \lambda_0 I_n)^{k+1})] \\ &= r((A - \lambda_0 I_n)^{k-1}) + r((A - \lambda_0 I_n)^{k+1}) - 2r((A - \lambda_0 I_n)^k) \end{aligned}$$

命题得证。 □

例题 2.6: 特征值全为 1 的矩阵的幂相似性 设 n 阶矩阵 A 的特征值全为 1, 求证: 对任意的正整数 k , A^k 与 A 相似。

解答. 显然 A^k 的特征值也全为 $1^k = 1$ 。

我们需要证明 A^k 与 A 关于特征值 1 的各阶 Jordan 块分布相同。由 Weyl 相似准则, 只需证明对于任意正整数 l , 都有:

$$r((A^k - I_n)^l) = r((A - I_n)^l)$$

注意到利用多项式的因式分解, 有:

$$(A^k - I_n)^l = [(A - I_n)(A^{k-1} + A^{k-2} + \cdots + A + I_n)]^l = (A - I_n)^l (A^{k-1} + A^{k-2} + \cdots + I_n)^l$$

令 $M = A^{k-1} + A^{k-2} + \cdots + I_n$ 。因为 A 的特征值全为 1, 根据多项式特征值映射定理, 矩阵 M 的特征值全为:

$$1^{k-1} + 1^{k-2} + \cdots + 1 = k$$

由于 k 是正整数, 故 M 的所有特征值 $k \neq 0$, 这意味着矩阵 M 及其幂次 M^l 都是可逆矩阵。

矩阵乘以可逆矩阵不改变其秩, 因此:

$$r((A^k - I_n)^l) = r((A - I_n)^l M^l) = r((A - I_n)^l)$$

对任意的正整数 l 都成立。由相似的秩准则可知, A^k 与 A 相似。 □

例题 2.7: 对合类矩阵与其逆矩阵相似 设 n 阶矩阵 A 的特征值全为 1 或 -1, 求证: A^{-1} 与 A 相似。

解答. 因为 A 的特征值为 ± 1 , 均不为 0, 所以 A 是可逆矩阵。显然 A^{-1} 的特征值也全为 $1^{-1} = 1$ 或 $(-1)^{-1} = -1$, 即 A 与 A^{-1} 具有完全相同的特征值集合。

设 $\lambda_0 = \pm 1$ 为其任意一个特征值。我们考察平移幂次的秩。注意到 $\lambda_0^2 = 1 \implies \lambda_0^{-1} = \lambda_0$ 。

$$\begin{aligned}(A^{-1} - \lambda_0 I_n)^l &= [A^{-1}(I_n - \lambda_0 A)]^l \\ &= (A^{-1})^l (-\lambda_0)^l (A - \lambda_0^{-1} I_n)^l \\ &= (-\lambda_0)^l A^{-l} (A - \lambda_0 I_n)^l\end{aligned}$$

由于 A 可逆, 故 $(-\lambda_0)^l A^{-l}$ 也是一个可逆矩阵。左乘可逆矩阵不改变原矩阵的秩, 因此:

$$r((A^{-1} - \lambda_0 I_n)^l) = r((A - \lambda_0 I_n)^l)$$

上式对任意的特征值 $\lambda_0 = \pm 1$ 以及任意的正整数 l 都成立。由相似的秩准则可知, A^{-1} 与 A 的 Jordan 标准型完全相同, 从而 A^{-1} 与 A 相似。□

例题 2.8: 非零特征值 Jordan 块的幂次标准型 设 $J = J_n(a)$ 是特征值为 $a \neq 0$ 的 n 阶 Jordan 块, 求 J^m 的 Jordan 标准型, 其中 m 为非零整数。

解答. 结论: J^m 的 Jordan 标准型为 $J_n(a^m)$ (即仍然是一个完整的 n 阶 Jordan 块)。

证明: 若 $m < 0$, 由于 $a \neq 0$, J 可逆。可先研究 $m > 0$ 的情形, 再结合前一题逆矩阵相似的思路推广到负整数。这里我们统一给出普适的代数证明:

记 n 阶标准幂零矩阵为 N (超对角线全为 1, 其余为 0)。则 $J = aI_n + N$ 。由于 aI_n 与 N 可交换, 利用二项式定理 (或者矩阵的泰勒展开) 计算 $J^m = (aI_n + N)^m$:

$$J^m = a^m I_n + ma^{m-1}N + \frac{m(m-1)}{2}a^{m-2}N^2 + \dots$$

由于 N^k ($k \geq n$) 均为零矩阵, 上式是一个有限项求和。显然 J^m 的特征值为 a^m 。考察其特征矩阵的退化情况:

$$J^m - a^m I_n = N \left(ma^{m-1}I_n + \frac{m(m-1)}{2}a^{m-2}N + \dots \right)$$

令括号内的矩阵为 $M = ma^{m-1}I_n + \frac{m(m-1)}{2}a^{m-2}N + \dots$ 。由于 N 的严格上三角性, M 是一个上三角矩阵, 且其主对角线元素全为 ma^{m-1} 。因为 $a \neq 0$ 且 $m \neq 0$, 所以主对角元 $ma^{m-1} \neq 0$ 。这意味着矩阵 M 的特征值均不为零, 即 M 是一个可逆矩阵。

从而对于任意的正整数 l , 有:

$$r((J^m - a^m I_n)^l) = r(N^l M^l)$$

由于 M^l 依然可逆, 故:

$$r((J^m - a^m I_n)^l) = r(N^l)$$

我们知道, 纯幂零块 N 的秩序列 $r(N^l)$ 恰好对应一个完整的 n 阶 Jordan 块。既然 $J^m - a^m I_n$ 的秩退化序列与 N 完全一致, 说明 J^m 也只有唯一的一个大小为 n 的对应于特征值 a^m 的 Jordan 块。

故 J^m 的 Jordan 标准型为 $J_n(a^m)$ 。 $\square$

例题 2.9: 矩阵指数与幂零矩阵的相似性 设 A 为 n 阶幂零矩阵, 证明: e^A 与 $I_n + A$ 相似。

解答. 因为 A 为幂零矩阵, 存在正整数 $k \leq n$ 使得 $A^k = O$ 。此时 A 的特征值全为 0。

首先考察 e^A 和 $I_n + A$ 的特征值。根据矩阵指数函数的泰勒展开, 因为 $A^k = O$, 级数成为有限项和:

$$e^A = I_n + A + \frac{1}{2!}A^2 + \cdots + \frac{1}{(k-1)!}A^{k-1}$$

由矩阵的多项式谱映射理论, e^A 的特征值全为 $e^0 = 1$; 同理 $I_n + A$ 的特征值全为 $1 + 0 = 1$ 。两者具有完全相同的特征值集合。

根据 Weyl 相似的秩准则, 只需证明对于任意正整数 l , 它们关于特征值 1 的特征矩阵的幂次秩相等, 即证明:

$$r((e^A - I_n)^l) = r(((I_n + A) - I_n)^l)$$

等式右侧显然为 $r(A^l)$ 。

对于等式左侧, 我们对多项式提取公因式 A :

$$e^A - I_n = A \left(I_n + \frac{1}{2!}A + \cdots + \frac{1}{(k-1)!}A^{k-2} \right)$$

记括号内的辅助多项式矩阵为 $M = I_n + \frac{1}{2!}A + \cdots + \frac{1}{(k-1)!}A^{k-2}$ 。由于 A 的特征值全为 0, M 的特征值全为 1。特征值全不为零意味着 M 是一个可逆矩阵。

又因为 M 是 A 的多项式, 故 A 与 M 可交换 (即 $AM = MA$)。因此, 对其进行 l 次方运算:

$$(e^A - I_n)^l = (AM)^l = A^l M^l$$

由于 M 可逆, 其幂次 M^l 必定可逆。矩阵乘以可逆矩阵不改变其秩, 故有:

$$r((e^A - I_n)^l) = r(A^l M^l) = r(A^l)$$

由此可见, 对任意的正整数 l , $r((e^A - I_n)^l) = r(A^l)$ 恒成立。这说明 e^A 与 $I_n + A$ 关于唯一特征值 1 的各阶 Jordan 块分布完全相同, 故两者相似。 $\square$

3 矩阵函数

例题 3.1: 矩阵指数的基本性质 将 $\mathbb{C}^{n \times n}$ 等同于 $\mathbb{C}^{n^2}$, 则对 $A \in \mathbb{C}^{n \times n}$, 可以定义绝对收

敛的矩阵指数:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathbb{C}^{n \times n}.$$

证明:

- (1) $\exp(O) = E_n$;
- (2) 如果 $A, B \in \mathbb{C}^{n \times n}$ 可交换, 则有 $\exp(A)\exp(B) = \exp(A+B)$;
- (3) 对任意 $A \in \mathbb{C}^{n \times n}$, $\exp(A)$ 可逆, 且 $(\exp(A))^{-1} = \exp(-A)$;
- (4) 如果 $A, B \in \mathbb{C}^{n \times n}$ 相似, 则 $\exp(A)$ 和 $\exp(B)$ 相似;
- (5) 对任意 $A \in \mathbb{C}^{n \times n}$, 有 $\det(\exp(A)) = e^{\text{Tr}(A)}$.

解答. 证明 (1): 根据矩阵指数的定义, 约定 $A^0 = E_n$. 代入零矩阵 O :

$$\exp(O) = \frac{O^0}{0!} + \frac{O^1}{1!} + \frac{O^2}{2!} + \cdots = E_n + O + O + \cdots = E_n.$$

证明 (2): 由于矩阵指数级数在 $\mathbb{C}^{n \times n}$ 中绝对收敛, 我们可以对其进行柯西乘积 (Cauchy product) 展开并重排各项:

$$\exp(A)\exp(B) = \left(\sum_{i=0}^{\infty} \frac{A^i}{i!} \right) \left(\sum_{j=0}^{\infty} \frac{B^j}{j!} \right) = \sum_{k=0}^{\infty} \sum_{i=0}^k \frac{A^i B^{k-i}}{i!(k-i)!}.$$

因为 A 和 B 可交换 ($AB = BA$), 二项式定理对矩阵同样成立:

$$(A+B)^k = \sum_{i=0}^k \frac{k!}{i!(k-i)!} A^i B^{k-i}.$$

将其代入上式, 提出 $\frac{1}{k!}$, 得:

$$\exp(A)\exp(B) = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=0}^k \frac{k!}{i!(k-i)!} A^i B^{k-i} = \sum_{k=0}^{\infty} \frac{(A+B)^k}{k!} = \exp(A+B).$$

证明 (3): 显然矩阵 A 与 $-A$ 总是可交换的 ($A(-A) = -A^2 = (-A)A$). 利用第 (1) 问和第 (2) 问的结论:

$$\exp(A)\exp(-A) = \exp(A+(-A)) = \exp(O) = E_n.$$

同理 $\exp(-A)\exp(A) = E_n$. 根据逆矩阵的定义, $\exp(A)$ 必定可逆, 并且其逆矩阵即为 $\exp(-A)$.

证明 (4): 因为 A 和 B 相似, 存在可逆矩阵 P 使得 $A = PBP^{-1}$. 注意到 $A^k = (PBP^{-1})^k = PB^kP^{-1}$. 代入级数定义中:

$$\exp(A) = \sum_{k=0}^{\infty} \frac{(PBP^{-1})^k}{k!} = \sum_{k=0}^{\infty} \frac{PB^kP^{-1}}{k!} = P \left(\sum_{k=0}^{\infty} \frac{B^k}{k!} \right) P^{-1} = P \exp(B) P^{-1}.$$

故 $\exp(A)$ 与 $\exp(B)$ 相似。

证明 (5): 在复数域上, 根据 Schur 定理, 任意矩阵 A 都酉相似于一个上三角矩阵 T 。即存在可逆阵 P 使得 $A = PTP^{-1}$, 其中 T 的主对角线元素 $\lambda_1, \lambda_2, \dots, \lambda_n$ 恰好为 A 的全部特征值。

显然矩阵的迹为特征值之和: $\text{Tr}(A) = \text{Tr}(T) = \sum_{i=1}^n \lambda_i$ 。

由第 (4) 问结论, $\exp(A)$ 相似于 $\exp(T)$, 相似矩阵行列式相等, 故:

$$\det(\exp(A)) = \det(\exp(T)).$$

考察 T 的指数 $\exp(T) = \sum \frac{T^k}{k!}$ 。由于 T 是上三角矩阵, 其任意次幂 T^k 依然是上三角矩阵, 且 T^k 的主对角线元素分别为 $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ 。因此, 当所有幂次求和后, $\exp(T)$ 也是一个上三角矩阵, 其主对角线元素就是每个特征值的标量泰勒展开, 即分别为 $e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n}$ 。

由于上三角矩阵的行列式等于主对角线元素之积:

$$\det(\exp(T)) = \prod_{i=1}^n e^{\lambda_i} = e^{\lambda_1 + \lambda_2 + \dots + \lambda_n} = e^{\text{Tr}(A)}.$$

命题得证。 □

例题 3.2: 多项式函数在 Jordan 块上的作用 设 $J_n(a)$ 是特征值为 a 的 n 阶 Jordan 块, 又 $f(x)$ 是一个多项式, 求证:

$$f(J_n(a)) = \begin{pmatrix} f(a) & \frac{f'(a)}{1!} & \frac{f^{(2)}(a)}{2!} & \dots & \frac{f^{(n-1)}(a)}{(n-1)!} \\ & f(a) & \frac{f'(a)}{1!} & \dots & \frac{f^{(n-2)}(a)}{(n-2)!} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \frac{f'(a)}{1!} \\ & & & & f(a) \end{pmatrix}.$$

解答. 设 $f(x)$ 的次数为 m 。将多项式 $f(x)$ 在 $x = a$ 处进行泰勒展开, 因为 $f(x)$ 是多项式, 该展开是精确的有限和:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \dots + \frac{f^{(m)}(a)}{m!}(x-a)^m$$

对于矩阵 $J_n(a)$, 我们可以将其拆分为数量矩阵与幂零矩阵之和: $J_n(a) = aI_n + N$ 。其中 N 是主对角线上方第一条次对角线全为 1, 其余全为 0 的 n 阶标准幂零矩阵。

将矩阵 $J_n(a)$ 代入多项式 $f(x)$ 中, 此时 $(x-a)$ 对应于 $(J_n(a) - aI_n) = N$ 。由于 $N^k = O$ 对所有 $k \geq n$ 均成立, 上述多项式的矩阵求值在第 n 项之后自然截断为零矩阵:

$$f(J_n(a)) = f(a)I_n + \frac{f'(a)}{1!}N + \frac{f^{(2)}(a)}{2!}N^2 + \dots + \frac{f^{(n-1)}(a)}{(n-1)!}N^{n-1}$$

观察右侧各项的矩阵结构:

- I_n 是主对角线为 1 的矩阵;

- N 是第一条超对角线为 1 的矩阵;
- N^2 是第二条超对角线为 1 的矩阵;
- 依此类推, N^k 是第 k 条超对角线为 1 的矩阵。

因此, 将这些项按对应系数相加后, 主对角线全为 $f(a)$, 第一超对角线全为 $\frac{f'(a)}{1!}$, 第二超对角线全为 $\frac{f^{(2)}(a)}{2!}$, 一直到右上角元素为 $\frac{f^{(n-1)}(a)}{(n-1)!}$ 。即得到了原命题中的上三角矩阵形式, 结论得证。□

例题 3.3: 矩阵指数与三阶常系数线性微分方程组 利用矩阵指数函数求解下列三阶常系数线性微分方程组的初值问题:

$$\begin{cases} \frac{dx}{dt} = x + y \\ \frac{dy}{dt} = -x + 2y + z \\ \frac{dz}{dt} = -x + 3z \end{cases}, \quad \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

解答. 将微分方程组写成矩阵形式 $\frac{dX(t)}{dt} = AX(t)$, 其中 $X(t) = (x(t), y(t), z(t))^T$, 系数矩阵为:

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

该方程组的通解可由矩阵指数函数直接给出: $X(t) = e^{At}X(0)$ 。

步骤 1: 求特征值与构造幂零矩阵

计算 A 的特征多项式 (略去行列式展开过程): $|\lambda I - A| = (\lambda - 2)^3 = 0$ 。因此 A 有一个三重特征值 $\lambda_0 = 2$ 。

我们提取特征值, 将 A 分解为纯量阵与剩余部分的和: $A = 2I + N$, 其中:

$$N = A - 2I = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

计算 N 的幂次:

$$N^2 = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

$$N^3 = N^2 \cdot N = \begin{pmatrix} 0 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

这表明 N 是一个幂零指数为 3 的矩阵。

步骤 2: 展开矩阵指数

由于 $2I$ 与 N 显然可交换, 由矩阵指数性质可得:

$$e^{At} = e^{(2I+N)t} = e^{2It}e^{Nt} = e^{2t}I \cdot e^{Nt}$$

因为 $N^3 = O$, 级数 e^{Nt} 在含 t^2 的项之后直接截断为零, 无需计算无限级数或求过渡矩阵:

$$\begin{aligned} e^{Nt} &= I + Nt + \frac{1}{2}N^2t^2 \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} -t & t & 0 \\ -t & 0 & t \\ -t & 0 & t \end{pmatrix} + \begin{pmatrix} 0 & -\frac{t^2}{2} & \frac{t^2}{2} \\ 0 & -\frac{t^2}{2} & \frac{t^2}{2} \\ 0 & -\frac{t^2}{2} & \frac{t^2}{2} \end{pmatrix} \\ &= \begin{pmatrix} 1-t & t-\frac{t^2}{2} & \frac{t^2}{2} \\ -t & 1-\frac{t^2}{2} & t+\frac{t^2}{2} \\ -t & -\frac{t^2}{2} & 1+t+\frac{t^2}{2} \end{pmatrix} \end{aligned}$$

步骤 3: 得出最终解析解

将 $e^{At} = e^{2t}e^{Nt}$ 与初始条件相乘, 即得最终解:

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = e^{2t} \begin{pmatrix} 1-t & t-\frac{t^2}{2} & \frac{t^2}{2} \\ -t & 1-\frac{t^2}{2} & t+\frac{t^2}{2} \\ -t & -\frac{t^2}{2} & 1+t+\frac{t^2}{2} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

(若需要具体到分量形式, 可按行展开, 此处矩阵形式即为最优的紧凑表达)。□

例题 3.4: 矩阵序列高次幂的收敛性与极限矩阵 设 A 是 n 阶方阵, 求 $\lim_{k \rightarrow \infty} A^k$ 存在的充要条件以及极限矩阵。

解答. 充要条件分析:

设 A 的 Jordan 标准型为 J , 则存在可逆矩阵 P 使得 $A = PJP^{-1}$ 。对于任意正整数 k , 有 $A^k = PJ^kP^{-1}$ 。因此, $\lim_{k \rightarrow \infty} A^k$ 存在当且仅当 $\lim_{k \rightarrow \infty} J^k$ 存在。

设 $J = \text{diag}(J_1, J_2, \dots, J_s)$, 其中 J_i 是对应于特征值 λ_i 的 m_i 阶 Jordan 块。由本节第一题的结论 (或者直接展开), J_i^k 的主对角线元素为 λ_i^k , 第一超对角线元素为 $\binom{k}{1}\lambda_i^{k-1}$, 一般地, 第 t 条超对角线元素为 $\binom{k}{t}\lambda_i^{k-t}$ 。

为了使 $\lim_{k \rightarrow \infty} J_i^k$ 存在, 该块中所有元素的极限都必须存在。我们分情况讨论:

- 若 $|\lambda_i| < 1$: 因为指数衰减快于多项式增长, $\lim_{k \rightarrow \infty} \binom{k}{t}\lambda_i^{k-t} = 0$, 此时 $\lim_{k \rightarrow \infty} J_i^k = O$ 。
- 若 $|\lambda_i| > 1$: λ_i^k 发散, 极限不存在。
- 若 $|\lambda_i| = 1$ 但 $\lambda_i \neq 1$: λ_i^k 在复平面 (或实轴) 上振荡, 极限不存在。
- 若 $\lambda_i = 1$: 主对角线元素 $1^k = 1$ 收敛。但如果阶数 $m_i \geq 2$, 其第一超对角线元素为 $\binom{k}{1}1^{k-1} = k$, 当 $k \rightarrow \infty$ 时发散。因此, 必须要求 $m_i = 1$ (即 $\lambda_i = 1$ 对应的 Jordan 块全为 1×1 的一阶块)。

综上所述, $\lim_{k \rightarrow \infty} A^k$ 存在的充要条件是: A 的所有特征值 λ 均满足 $|\lambda| < 1$; 或者, 如果 $\lambda = 1$ 是特征值, 则要求它是半单纯的 (即代数重数等于几何重数, 对应的初等因子全为一次), 且除 1 以外的其他特征值均满足 $|\lambda| < 1$ 。

极限矩阵的求解:

在满足上述充要条件的前提下, 我们可以适当地排列 Jordan 块, 使得可逆矩阵 P 满足:

$$A = P \begin{pmatrix} I_r & O \\ O & \tilde{J} \end{pmatrix} P^{-1}$$

其中 I_r 是对应于特征值 $\lambda = 1$ 的 r 阶单位阵 (r 为其重数, 若无特征值 1 则 $r = 0$), $\tilde{J}$ 是包含了所有满足 $|\lambda| < 1$ 的 Jordan 块的分块对角阵。

当 $k \rightarrow \infty$ 时, 由于 $\tilde{J}$ 中所有特征值的模均小于 1, 必定有 $\tilde{J}^k \rightarrow O$ 。因此, 极限矩阵为:

$$A^\infty = \lim_{k \rightarrow \infty} A^k = P \begin{pmatrix} I_r & O \\ O & O \end{pmatrix} P^{-1}$$

(注: 这个极限矩阵 A^∞ 具有性质 $(A^\infty)^2 = A^\infty$, 它是一个幂等矩阵, 本质上是沿着其他广义特征子空间, 向特征值 $\lambda = 1$ 对应的特征子空间作平行投影的投影算子。) $\square$

例题 3.5: 矩阵指数化为自身多项式的证明 设 $A \in \mathbb{C}^{n \times n}$, 证明: 矩阵指数 e^A 一定可以表示为 A 的一个多项式, 且该多项式的次数不超过 $n-1$ 。

解答. 证法一: 有限维空间的拓扑闭包法 (结合 Hamilton-Cayley 定理)

由 Hamilton-Cayley 定理可知, 矩阵 A 满足其特征多项式 $g(\lambda) = \det(\lambda I - A) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0 = 0$ 。即有矩阵等式:

$$A^n = -c_{n-1}A^{n-1} - \cdots - c_1A - c_0I$$

利用这一递推关系, 对于任何 $k \geq n$, A^k 都可以不断降次, 最终表示为 $I, A, A^2, \dots, A^{n-1}$ 的线性组合。

考虑复矩阵空间 $\mathbb{C}^{n \times n}$ 中的子空间 $V = \text{span}\{I, A, A^2, \dots, A^{n-1}\}$ 。显然, V 是一个有限维向量空间 (维数至多为 n)。

矩阵指数的定义为部分和序列的极限:

$$e^A = \lim_{m \rightarrow \infty} S_m = \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{1}{k!} A^k$$

由于每一个 $A^k \in V$, 其有限线性组合 S_m 也必定始终属于 V 。在有限维空间中, 任何子空间在拓扑上必定是闭集。因此, 柯西序列 $\{S_m\}$ 在空间中的收敛极限也必然落在这个子空间 V 内。这说明 $e^A \in V$, 即存在复数 $a_0, a_1, \dots, a_{n-1}$ 使得:

$$e^A = a_0I + a_1A + \cdots + a_{n-1}A^{n-1}$$

证法二: 多项式插值法 (纯代数构造)

设矩阵 A 的极小多项式为 $m(x) = \prod_{i=1}^s (x - \lambda_i)^{r_i}$, 其中根的重数之和 $\sum r_i = d \leq n$ 。根据矩阵函数的谱映射定义, 对于解析函数 $f(x) = e^x$, 如果我们要寻找一个多项式 $p(x)$ 使得 $p(A) = f(A) = e^A$, 只需让 $p(x)$ 在 A 的“谱”上与 $f(x)$ 具有相同的函数值与各阶导数值。即满足如下的 Hermite 插值条件:

$$p^{(j)}(\lambda_i) = f^{(j)}(\lambda_i) = e^{\lambda_i}, \quad \text{对于 } i = 1, \dots, s; j = 0, 1, \dots, r_i - 1$$

这是一个包含 d 个方程的线性插值条件组。根据多项式插值理论 (或多项式环的中国剩余定理), 必定存在一个次数严格小于 d (从而其次数至多为 $n-1$) 的唯一多项式 $p(x)$ 满足上述所有条件。

既然多项式 $p(x)$ 与 e^x 在极小多项式 $m(x)$ 的根空间上行为完全一致, 这就意味着两者之差可以被 $m(x)$ 整除, 即存在函数 $h(x)$ 使得:

$$e^x - p(x) = m(x)h(x)$$

将矩阵 A 代入上式, 由于极小多项式化零 $m(A) = O$, 直接得到:

$$e^A - p(A) = O \implies e^A = p(A)$$

结论得证。 □

4 张量积

知识 4.1: 矩阵的张量积 (Kronecker 积)

设 $A = (a_{ij})$ 是一 $m \times n$ 矩阵, B 是一 $p \times q$ 矩阵。矩阵 A 与 B 的张量积 (又称 Kronecker 积), 记作 $A \otimes B$, 是一个 $mp \times nq$ 的分块矩阵, 其定义为:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}$$

例题 4.1: 张量积的混合乘积性质 设 $A \in \mathbb{F}^{m \times n}, B \in \mathbb{F}^{p \times q}, C \in \mathbb{F}^{n \times k}, D \in \mathbb{F}^{q \times l}$ 。证明:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

解答. 直接根据张量积的分块矩阵定义和矩阵的分块乘法规则进行计算。设 $A = (a_{ij})_{m \times n}$, $C = (c_{ij})_{n \times k}$ 。矩阵 $A \otimes B$ 的第 (i, s) 个子块为 $a_{is}B$, 矩阵 $C \otimes D$ 的第 (s, j) 个子块为 $c_{sj}D$ 。

在乘积矩阵 $(A \otimes B)(C \otimes D)$ 中, 其第 (i, j) 个子块为对应子块的内积:

$$\sum_{s=1}^n (a_{is}B)(c_{sj}D)$$

由于 a_{is} 和 c_{sj} 均为标量, 可以提取并结合:

$$\sum_{s=1}^n a_{is}c_{sj}(BD) = \left(\sum_{s=1}^n a_{is}c_{sj} \right) BD = (AC)_{ij}BD$$

这恰好是矩阵 $(AC) \otimes (BD)$ 的第 (i, j) 个子块。故 $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ 成立。 $\square$

例题 4.2: 张量积与矩阵的秩 设矩阵 $A \in \mathbb{F}^{m \times n}$, 矩阵 $B \in \mathbb{F}^{p \times q}$ 。证明:

$$\text{rank}(A \otimes B) = \text{rank}(A) \cdot \text{rank}(B)$$

解答. 设 $\text{rank}(A) = r, \text{rank}(B) = s$ 。根据矩阵的相抵标准型, 存在可逆矩阵 P_1, Q_1, P_2, Q_2 使得:

$$P_1 A Q_1 = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m \times n}, \quad P_2 B Q_2 = \begin{pmatrix} E_s & O \\ O & O \end{pmatrix}_{p \times q}$$

利用混合乘积性质 (例题中的结论), 有:

$$(P_1 \otimes P_2)(A \otimes B)(Q_1 \otimes Q_2) = (P_1 A Q_1) \otimes (P_2 B Q_2) = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} \otimes \begin{pmatrix} E_s & O \\ O & O \end{pmatrix}$$

注意到 $P_1 \otimes P_2$ 和 $Q_1 \otimes Q_2$ 分别具有逆矩阵 $P_1^{-1} \otimes P_2^{-1}$ 和 $Q_1^{-1} \otimes Q_2^{-1}$, 故它们都是可逆矩阵。乘以可逆矩阵不改变矩阵的秩, 因此:

$$\text{rank}(A \otimes B) = \text{rank} \left[\begin{pmatrix} E_r & O \\ O & O \end{pmatrix} \otimes \begin{pmatrix} E_s & O \\ O & O \end{pmatrix} \right]$$

右侧的张量积是一个对角分块矩阵, 其对角线上仅包含由 $E_r \otimes E_s$ 产生的 $r \times s$ 个 1, 其余元素均为 0。故其秩恰好为 $rs = \text{rank}(A) \cdot \text{rank}(B)$ 。 $\square$

例题 4.3: 张量积与特征值 设方阵 $A \in \mathbb{C}^{m \times m}$ 的特征值为 $\lambda_1, \dots, \lambda_m$, 方阵 $B \in \mathbb{C}^{n \times n}$ 的特征值为 $\mu_1, \dots, \mu_n$ 。证明: 矩阵 $A \otimes B$ 的 mn 个特征值为 $\lambda_i \mu_j$ ($1 \leq i \leq m, 1 \leq j \leq n$)。

解答. 在复数域上, 由 Schur 定理可知, 存在可逆 (乃至酉) 矩阵 P 和 Q , 使得 $P^{-1}AP = T_A$ 且 $Q^{-1}BQ = T_B$, 其中 T_A 和 T_B 分别是主对角线元素为 λ_i 和 μ_j 的上三角矩阵。

考虑对 $A \otimes B$ 进行相似变换。利用混合乘积性质:

$$(P \otimes Q)^{-1}(A \otimes B)(P \otimes Q) = (P^{-1} \otimes Q^{-1})(A \otimes B)(P \otimes Q) = (P^{-1}AP) \otimes (Q^{-1}BQ) = T_A \otimes T_B$$

观察 $T_A \otimes T_B$ 的结构：由于 T_A 是上三角阵， $T_A \otimes T_B$ 必定是一个分块上三角矩阵，且其主对角线上的各个分块矩阵为 $\lambda_i T_B$ ($i = 1, \dots, m$)。而 $\lambda_i T_B$ 本身又是上三角矩阵，其主对角线元素依次为 $\lambda_i \mu_1, \lambda_i \mu_2, \dots, \lambda_i \mu_n$ 。

因为上三角矩阵（或分块上三角内部且对角块也是上三角）的特征值即为其主对角线上的元素，所以 $T_A \otimes T_B$ 的特征值恰为 $\lambda_i \mu_j$ 。又因相似变换不改变特征值，故 $A \otimes B$ 的特征值为 $\lambda_i \mu_j$ 。□

例题 4.4: 矩阵张量和 (Kronecker Sum) 的特征值 设方阵 $A \in \mathbb{C}^{m \times m}$ 的特征值为 $\lambda_1, \dots, \lambda_m$ ，方阵 $B \in \mathbb{C}^{n \times n}$ 的特征值为 $\mu_1, \dots, \mu_n$ 。证明：矩阵 $A \otimes I_n + I_m \otimes B$ 的 mn 个特征值为 $\lambda_i + \mu_j$ ($1 \leq i \leq m, 1 \leq j \leq n$)。

解答. 在复数域上，由 Schur 定理可知，存在可逆矩阵 P 和 Q ，使得 $P^{-1}AP = T_A$ 且 $Q^{-1}BQ = T_B$ ，其中 T_A 和 T_B 分别是主对角线元素依次为 λ_i 和 μ_j 的上三角矩阵。

对目标矩阵 $A \otimes I_n + I_m \otimes B$ 进行相似变换。利用张量积的分配律和混合乘积性质：

$$\begin{aligned} (P \otimes Q)^{-1}(A \otimes I_n + I_m \otimes B)(P \otimes Q) &= (P^{-1} \otimes Q^{-1})(A \otimes I_n)(P \otimes Q) + (P^{-1} \otimes Q^{-1})(I_m \otimes B)(P \otimes Q) \\ &= (P^{-1}AP \otimes Q^{-1}I_nQ) + (P^{-1}I_mP \otimes Q^{-1}BQ) \\ &= (T_A \otimes I_n) + (I_m \otimes T_B) \end{aligned}$$

我们来考察分块矩阵 $(T_A \otimes I_n) + (I_m \otimes T_B)$ 的结构。因为 T_A 是上三角矩阵， $T_A \otimes I_n$ 是一个分块上三角矩阵，其第 i 个主对角块恰好为纯量矩阵 $\lambda_i I_n$ 。同时， $I_m \otimes T_B$ 是一个分块对角矩阵，其每一个主对角块均为 T_B 。

由于两个矩阵的分块方式完全相同，它们的和 $(T_A \otimes I_n) + (I_m \otimes T_B)$ 依然是一个分块上三角矩阵，且其第 i 个主对角块为：

$$\lambda_i I_n + T_B$$

对于这个主对角块而言，由于 T_B 是主对角线元素为 $\mu_1, \dots, \mu_n$ 的上三角矩阵， $\lambda_i I_n + T_B$ 显然也是上三角矩阵，且其主对角线上的元素依次为 $\lambda_i + \mu_1, \lambda_i + \mu_2, \dots, \lambda_i + \mu_n$ 。

根据分块上三角矩阵的性质，其特征值由各主对角块的特征值共同构成。因此，整个大矩阵的主对角线元素（即特征值）恰为所有的 $\lambda_i + \mu_j$ 。又因相似变换不改变特征值，故原矩阵 $A \otimes I_n + I_m \otimes B$ 的特征值为 $\lambda_i + \mu_j$ 。□

例题 4.5: 线性变换 $X \mapsto AXB$ 的矩阵表示 (拉直算子) 设 $V = \mathbb{F}^{m \times n}$ 是 $m \times n$ 矩阵构成的线性空间。定义矩阵的拉直算子 (vec operator)，将矩阵 $X = (x_1, x_2, \dots, x_n)$ (其中

x_i 为列向量) 按列顺序堆叠成一个 mn 维列向量:

$$\text{vec}(X) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}_{mn \times 1}$$

设 $A \in \mathbb{F}^{m \times m}, B \in \mathbb{F}^{n \times n}$ 。证明: 对于 V 上的线性变换 $f(X) = AXB$, 在其标准基 (对应于拉直算子) 下的表示矩阵恰为 $B^T \otimes A$, 即有恒等式:

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$$

解答. 设 $B = (b_{ij})_{n \times n}$ 。我们先考察乘积矩阵 XB 。 XB 的第 j 列是由 X 的各列的线性组合构成的, 具体表达为:

$$(XB)_j = \sum_{k=1}^n b_{kj} x_k$$

接着考察 $Y = AXB = A(XB)$ 。矩阵左乘 A 相当于对 XB 的每一列独立左乘 A , 因此 Y 的第 j 列为:

$$y_j = A \left(\sum_{k=1}^n b_{kj} x_k \right) = \sum_{k=1}^n b_{kj} Ax_k$$

现在我们将 Y 进行拉直操作, 即写出 $\text{vec}(Y)$ 的分块形式。其第 j 个子块向量即为 y_j :

$$\text{vec}(AXB) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_{11}Ax_1 + b_{21}Ax_2 + \cdots + b_{n1}Ax_n \\ b_{12}Ax_1 + b_{22}Ax_2 + \cdots + b_{n2}Ax_n \\ \vdots \\ b_{1n}Ax_1 + b_{2n}Ax_2 + \cdots + b_{nn}Ax_n \end{pmatrix}$$

将上式写成矩阵与向量的乘积形式, 并提取系数 b_{ij} 与变换 A :

$$\text{vec}(AXB) = \begin{pmatrix} b_{11}A & b_{21}A & \cdots & b_{n1}A \\ b_{12}A & b_{22}A & \cdots & b_{n2}A \\ \vdots & \vdots & \ddots & \vdots \\ b_{1n}A & b_{2n}A & \cdots & b_{nn}A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

观察左侧的系数矩阵, 其第 (i, j) 个块正好是 $b_{ji}A$ 。这正是转置矩阵 B^T 与矩阵 A 的张量积定义。因此:

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$$

命题得证。 □