第二章方阵的相似化简

§2.1 特征多项式和最小多项式

对于复数域上 n 阶方阵 $A=[a_{ij}]$, 它的特征多项式

$$\det(\lambda I_n - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \ldots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \ldots & -a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ -a_{n1} & -a_{n2} & \ldots & \lambda - a_{nn} \end{vmatrix}$$

是 λ 的 n 次多项式

$$f(\lambda) := \det(\lambda I_n - A) = \lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0$$

这个 n 次多项式在复数域有 n 个根 $\lambda_1, \lambda_2, ..., \lambda_n$, (其中重根进行重复计数, 例如 $(\lambda-1)(\lambda-2)^2$ 有 3 个根, $1, 2, 2$)

$$\begin{aligned} \det(\lambda I_n - A) &= \lambda^n + b_{n-1}\lambda^{n-1} + ... + b_1\lambda + b_0 \\ &= (\lambda-\lambda_1)(\lambda-\lambda_2)...(\lambda-\lambda_n) \end{aligned}$$

(1). 方阵的迹 tr$A$(对角元之和) 等于特征值之和.

$$\begin{aligned} -b_{n-1} &= \lambda_1 + \lambda_2 + ...\lambda_n \\ &= a_{11} + a_{22} + ... + a_{nn} := \mathrm{tr}A \end{aligned}$$

(2). 方阵的行列式的值等于所有特征值的乘积.

$$(-1)^n b_0 = \lambda_1 \cdot \lambda_2 \cdot ... \cdot \lambda_n = \det A$$

方阵的多项式

定义: 若

$$g(t) = a_m t^m + a_{m-1} t^{m-1} + \ldots + a_1 t + a_0$$

是 t 的多项式. 记 $g(A)$ 为把 $g(t)$ 中 t 替换为 A 的结果

$$g(A) = a_m A^m + a_{m-1} A^{m-1} + \ldots + a_1 A + a_0 I$$

则 $g(A)$ 称为方阵 A 的多项式. 它是与 A 同阶的方阵.

例: 设 $g(t)=t^2-3t+4$, $h(t)=t^2-3t+2$,

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

则有

$$g(A)=A^2-3A+4I=\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

$$h(A)=A^2-3A+2I=\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

由矩阵乘法, 一般来说 AB 未必等于 BA, 如

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

设 $p(t), q(t)$ 是两个多项式, 它们的乘积

$$h(t) = p(t)q(t)$$

仍是一个多项式. 由于 $p(t)q(t) = q(t)p(t)$, 则 $h(A) = q(A)p(A) = p(A)q(A)$, 即$p(A)$ 与 $q(A)$ 关于矩阵乘法可交换.

定理: 设 n 阶方阵 A 的 n 个特征值是 $\lambda_1, \lambda_2, \ldots, \lambda_n$, 相应的特征向量分别为 $x_1, x_2, \ldots, x_n$, 若 $g(t)$ 是一个多项式, 则 $g(A)$ 的 n 个特征值是 $g(\lambda_1), g(\lambda_2), g(\lambda_n)$, 且 $x_1, x_2, \ldots, x_n$ 分别是相应的特征向量.

证明: 设 $g(t) = a_m t^m + a_{m-1} t^{m-1} + \ldots + a_0$,
由于 $Ax_i = \lambda x_i, A^2 x_i = \lambda_i^2 x_i, \ldots, A^m x_i = \lambda_i^m x_i$, 故

$$\begin{aligned} g(A)x_i &= (a_m A^m + a_{m-1} A^{m-1} + \ldots + a_1 A + a_0 I)x_i \\ &= (a_m \lambda_i^m + a_{m-1} \lambda_i^{m-1} + \ldots + a_1 \lambda_i + a_0)x_i \\ &= g(\lambda_i)x_i, \quad 1 \leq i \leq n \end{aligned}$$

故 $g(\lambda_i)$ 是 $g(A)$ 的特征值, x_i 是相应的特征向量.

例 若

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

求 $g(A)=2A^5-3A^4-A^3+2A-I$ 的特征值与特征向量.

解: 思路一: 先求出 $g(A)$, 再计算其特征值与特征向量.

思路二: A 的特征值和特征向量为 $\{1,2,3\}$ 和 $\{(1,0,0)^T,(0,1,0)^T,(0,0,1)^T\}$.

由前定理知, $g(A)$ 的特征值为 $\{g(1),g(2),g(3)\}=\{-1,11,221\}$, 对应的特征向量仍为 $\{(1,0,0)^T,(0,1,0)^T,(0,0,1)^T\}$.

定义: 若方阵 $A \neq O$, 且存在正整数 k, 使 $A^k = O$, 则称 A 为幂零矩阵.

例: $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $A^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $A^3 = O$.

由定义, 幂零矩阵的特征值必是 0.
(事实上, 若 $Ax_0 = \lambda_0 x_0$, 且 $A^k = O$, 则

$$0 = A^k x_0 = \lambda_0^k x_0, \quad x_0 \neq 0.$$

故 $\lambda_0^k = 0$, 即 $\lambda_0 = 0$.)

定理: 设 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵 $(m \geq n)$, 方阵 AB, BA 的特征多项式分别为 $f_{AB}(\lambda), f_{BA}(\lambda)$, 则有

$$f_{AB}(\lambda) = \lambda^{m-n} f_{BA}(\lambda)$$

证明: 设 $\mathrm{rank} A = r$, 则存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q, 使

$$PAQ = \begin{pmatrix} I_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{pmatrix}$$

证明 (续): 因而

$$\begin{aligned} PABP^{-1} &= PAQQ^{-1}BP^{-1} \\ &= \begin{pmatrix} I_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{pmatrix} \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \\ &= \begin{pmatrix} G_{11} & G_{12} \\ O_{(m-r)\times r} & O_{(m-r)\times(m-r)} \end{pmatrix}, \end{aligned}$$

其中 $Q^{-1}BP^{-1} := \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}$. 同理可得

$$\begin{aligned} Q^{-1}BAQ &= Q^{-1}BP^{-1}PAQ \\ &= \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \begin{pmatrix} I_r & O_{r\times(n-r)} \\ O_{(m-r)\times r} & O_{(m-r)\times(n-r)} \end{pmatrix} \\ &= \begin{pmatrix} G_{11} & O_{r\times(n-r)} \\ G_{21} & O_{(n-r)\times(n-r)} \end{pmatrix}, \end{aligned}$$

证明 (续): 于是,

$$\begin{aligned} f_{AB}(\lambda) &= \det(\lambda I_m - AB) \\ &= (\lambda I_m - PABP^{-1}) = \lambda^{m-r}\det(\lambda I_r - G_{11}), \\ f_{BA}(\lambda) &= \det(\lambda I_m - BA) \\ &= (\lambda I_m - Q^{-1}BAQ) = \lambda^{n-r}\det(\lambda I_r - G_{11}). \end{aligned}$$

比较以上两式, 得 $f_{AB}(\lambda) = \lambda^{m-n} f_{BA}(\lambda)$.

这个定理表明, m 阶方阵 AB 与 n 阶方阵 BA 的非零特征值是相同的, 特别的, 若 $m = n$, 则 AB 与 BA 的特征值相同.

例 1. 求镜像变换的 *Householder* 矩阵

$$H = I_n - 2\omega\omega^T$$

的特征值及它的迹和行列式.

解: 令 $A = 2\omega$, $B = \omega^T$, 则
$AB = 2\omega\omega^T$, $BA = 2\omega^T\omega = 2$, 故

$$\begin{aligned} \det(\lambda I_n - H) &= \det((\lambda - 1)I_n + 2\omega\omega^T) \\ = & \ (\lambda - 1)^{n-1}\det(\lambda - 1 + 2\omega^T\omega) = (\lambda - 1)^{n-1}(\lambda + 1) \end{aligned}$$

因而 $\lambda = 1$ 是 H 的 $n - 1$ 重特征值, $\lambda = -1$ 是单重特征值. 利用之前的结果可知,

$$\mathrm{tr}H = n - 1 + (-1) = n - 2, \quad \det H = -1.$$

例 2. 若方阵 A 的所有特征值的模都小于 1, 则方阵 $I - A$ 是可逆的.

证: 设 $\lambda_1, \lambda_2, ..., \lambda_n$ 是 A 的所有特征值, 且 $|\lambda_i| < 1 \ (1 \le i \le n)$, 则 $1 - \lambda_1, 1 - \lambda_2, ..., 1 - \lambda_n$ 是 $I - A$ 的所有特征值. 由于

$$|1 - \lambda_i| \ge 1 - |\lambda_i| > 0, \quad i = 1, 2, ..., n$$

所以 $1 - \lambda_i \ne 0$, 故

$$\det(I - A) = (1 - \lambda_1)(1 - \lambda_2)...(1 - \lambda_n) \ne 0.$$

从而 $I - A$ 是可逆的.

定义: 设 A 是一个 n 阶方阵, $g(t)$ 是一个多项式, 如果 $g(A) = O$, 则称 $g(t)$ 是 A 的零化多项式.

如 $h(t) = t^2 - 3t + 2$ 就是 $A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ 的一个零化多项式.

如果 $g(t)$ 是 A 的一个零化多项式, 而 $h(t)$ 是任一多项式, 则 $h(t)g(t)$ 也是 A 的零化多项式, 因此方阵的零化多项式不是唯一的.

问题: 那么一个方阵是否一定存在零化多项式呢?

定理 (Cayley-Hamilton): 设 n 阶方阵 A 的特征多项式为

$$f(\lambda) = \det(\lambda I - A) = \lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_1\lambda + b_0$$

则 $f(A) = O$, 即 A 的特征多项式为其零化多项式.

证: 设 $B(\lambda)$ 是 $\lambda I - A$ 的伴随矩阵, 则有

$$B(\lambda)(\lambda I - A) = f(\lambda)I = \lambda^n I + b_{n-1}\lambda^{n-1} I + \ldots + b_0 I.$$

由于方阵 $B(\lambda)$ 的元是 $\lambda I - A$ 的代数余子式, 故均是 λ 的次数不大于 $n-1$ 的多项式, 从而由矩阵的运算性质, $B(\lambda)$ 可以表示成

$$B(\lambda) = \lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \ldots + B_0,$$

证 (续): 其中 $B_0, B_1, \dots, B_{n-1}$ 都是 n 阶数矩阵. 故

$$\begin{aligned} & B(\lambda)(\lambda I - A) \\ = & (\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \dots + B_0)(\lambda I - A) \\ = & \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - B_{n-1}A) + \dots + \lambda(B_0 - B_1 A) \end{aligned}$$

又有 $B(\lambda)(\lambda I - A) = \lambda^n I + b_{n-1}\lambda^{n-1} I + \dots + b_0 I.$
比较以上两式中 λ 各次幂的系数, 得

$$\begin{aligned} \lambda^n &: B_{n-1} = I \\ \lambda^{n-1} &: B_{n-2} - B_{n-1}A = b_{n-1} I \\ \vdots \; &\quad \vdots \\ \lambda &: B_0 - B_1 A = b_1 I \\ \lambda^0 &: -B_0 A = b_0 I \end{aligned}$$

证 (续): 以 $A^n, A^{n-1}, \dots, A, I$ 依此从右边乘上面方程组中得第一式, 第二式,..., 第 n 式, 第 $n+1$ 式,

$$
\begin{array}{rcl}
\lambda^n & : & B_{n-1}A^n = A^n \\
\lambda^{n-1} & : & B_{n-2}A^{n-1} - B_{n-1}A^n = b_{n-1}A^{n-1} \\
\vdots & & \vdots \\
\lambda & : & B_0A - B_1A^2 = b_1A \\
\lambda^0 & : & -B_0A = b_0I
\end{array}
$$

再把这 $n+1$ 个式子加起来, 则左端为零矩阵, 右端即为 $f(A)$, 因此 $f(A) = O$.

即 A 的特征多项式为其零化多项式.

例 3. 已知

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

求 $g(A)=2A^5-3A^4-A^3+2A-I$.

解: A 的特征多项式

$$f(\lambda)=\begin{vmatrix} \lambda-1 & -1 & 1 \\ -1 & \lambda-1 & -1 \\ 0 & 1 & \lambda-2 \end{vmatrix}=(\lambda-1)^2(\lambda-2),$$

解 (续)：又 $g(\lambda)=2\lambda^5-3\lambda^4-\lambda^3+2\lambda-1$. 用 $f(\lambda)$ 除 $g(\lambda)$ 得

$$g(\lambda)=(2\lambda^2+5\lambda+9)f(\lambda)+15\lambda^2-33\lambda+17.$$

由于 $f(A)=O$, 故 $g(A)=15A^2-33A+17I$

$$=15\begin{pmatrix}2&3&-2\\2&1&2\\-1&-3&3\end{pmatrix}-33\begin{pmatrix}1&1&-1\\1&1&1\\0&-1&2\end{pmatrix}+17\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=\begin{pmatrix}14&12&3\\-3&-1&-3\\-15&-12&-4\end{pmatrix}.$$

如果不做多项式的除法, 可采用待定系数法, 用 $f(\lambda)$ 除 $g(\lambda)$ 一般得到下述表达式

$$g(\lambda) = h(\lambda)f(\lambda) + r(\lambda),$$

其中余式 $r(\lambda)$ 是一多项式, 其次数 $\deg r(\lambda) < \deg f(\lambda) = 3$, 因而可以设

$$r(\lambda) = a\lambda^2 + b\lambda + c,$$

代入 $g(\lambda)$ 式得

$$g(\lambda) = h(\lambda)f(\lambda) + a\lambda^2 + b\lambda + c.$$

现有三个未知数 a, b, c, 我们需要三个条件.
已知 $\lambda_1 = 1$ 和 $\lambda_2 = 2$ 分别是 $f(\lambda)$ 的二重根和一重根. 故有 $f(1) = f(2) = 0$ 和 $f'(1) = 0$.

利用 $f(1)=f(2)=0$ 和 $f'(1)=0$, 我们得到:

$$\begin{aligned} a+b+c &= g(1)=-1, \\ 4a+2b+c &= g(2)=11. \\ 2a+b &= g'(1)=-3 \end{aligned}$$

解之得

$$a=15, b=-33, c=17.$$

从而也可得到 $g(A)=r(A)=15A^2-33A+17I$.

例 4. 对于例 3 中的 A, 求逆矩阵 A^{-1}.

解: 由于 A 的特征多项式
$f(\lambda) = \lambda^3 - 4\lambda^2 + 5\lambda - 2$ 中常数项不为零, 所以
$\det A = f(0) \neq 0$, 故 A 可逆.
又 $f(A) = A^3 - 4A^2 + 5A - 2I = O$, 用 A^{-1} 乘之,
便得 A^{-1} 的表示式

$$A^{-1} = \frac{1}{2}(A^2 - 4A + 5I) = \frac{1}{2}\begin{pmatrix} 3 & -1 & 2 \\ -2 & 2 & -2 \\ -1 & 1 & 0 \end{pmatrix}.$$

定义: A 的零化多项式中, 次数最低的首一多项式称为 A 的**最小多项式**, 记为 $m_A(\lambda)$.

定理: A 的最小多项式 $m_A(\lambda)$ 可整除 A 的任何零化多项式 $g(\lambda)$, 且 $m_A(\lambda)$ 是唯一的.

证明: 反证: 若 $m_A(\lambda)$ 不能整除 $g(\lambda)$, 则有 $g(\lambda) = p(\lambda)m_A(\lambda) + r(\lambda)$, $\deg r(\lambda) < \deg m_A(\lambda)$.
由于 $O = g(A) = p(A)m_A(A) + r(A) = r(A)$, 所以 $r(\lambda)$ 是 A 的一个零化多项式, 且次数低于 $m_A(\lambda)$ 的次数. 这与 $m_A(\lambda)$ 的定义矛盾.
因而 $m_A(\lambda)$ 必能整除 $g(\lambda)$, 记为 $m_A(\lambda)|g(\lambda)$.

证明 (续): 再证唯一性:
设 $m_1(\lambda), m_2(\lambda)$ 是 A 的两个最小多项式,
则由最小多项式的定义, $m_1(\lambda)|m_2(\lambda)$, 且
$m_2(\lambda)|m_1(\lambda)$. 因而, $m_1(\lambda) = bm_2(\lambda)$, $b \neq 0$ 是常
数. 但 $m_1(\lambda), m_2(\lambda)$ 都是首一多项式, 故有 $b = 1$,
从而 $m_1(\lambda) = m_2(\lambda)$. 则定理证毕.

推论: A 的最小多项式 $m_A(\lambda)$ 必能整除 A 的特征
多项式 $f(\lambda)$, 即 $m_A(\lambda)|f(\lambda)$

定理: λ_0 是 A 的特征值, 其充分必要条件是 λ_0 是 A 的最小多项式 $m_A(\lambda)$ 的根.

证明: 设 λ_0 是 A 的特征值, x_0 是相应的特征向量, 则有

$$0 = m_A(A)x_0 = m_A(\lambda_0)x_0, \quad x_0 \neq 0,$$

故 $m_A(\lambda_0) = 0$, 即 λ_0 是 $m_A(\lambda)$ 的根.

反之, 若 λ_0 是 $m_A(\lambda)$ 的根, 那么由于 $m_A(\lambda)$ 可整除 A 的特征多项式 $f(\lambda)$, 故 λ_0 必是特征多项式的根, 即 λ_0 是 A 的特征值.

这个定理可用来通过特征多项式检验最小多项式.

事实上, 设 A 的所有不同的特征值是 $\lambda_1, \lambda_2, ..., \lambda_s$, 则 A 的特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}...(\lambda - \lambda_s)^{n_s}$$

且 $n_1 + n_2 + ... + n_s = n$, 则 A 的最小多项式一定有如下形式:

$$m_A(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2}...(\lambda - \lambda_s)^{k_s}$$

并且 $1 \leq k_i \leq n_i$ $(i = 1, 2, ..., s)$ 若 A 的特征值的代数重数都为 1, 那么 $m_A(\lambda) = f(\lambda)$.

例 5. 求

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix},$$

的最小多项式.

解: 由于 A 的特征多项式 $\det(\lambda I - A) = (\lambda - 3)(\lambda - 2)^3$, 所以 A 的最小多项式只能有一下三种可能

$$(\lambda - 3)(\lambda - 2); \quad (\lambda - 3)(\lambda - 2)^2; \quad (\lambda - 3)(\lambda - 2)^3$$

解 (续): 但 $(A-3I)(A-2I)$

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

$$= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq O$$

而 $(A-3I)(A-2I)^2 = O$,
故 $m_A(\lambda) = (\lambda-3)(\lambda-2)^2$.

例 6. 求

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

的最小多项式.

解: A 的特征多项式是 $(\lambda+2)^5(\lambda-3)^3$, 故 A 的最小多项式的形式是 $(\lambda+2)^k(\lambda-3)^l$. 由于

解 (续):

$$
\begin{aligned}
A+2I &= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix} \\
&= \operatorname{diag}\{A_1, A_2\}
\end{aligned}
$$

为对角块矩阵, 且其中 A_1 为 5 阶幂零矩阵, $A_1^3 = O$.

解 (续):

$$
\begin{aligned}
A - 3I &= \begin{pmatrix} -5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\
&= \mathrm{diag}\{A_3, A_4\}
\end{aligned}
$$

为对角块矩阵, 且其中 A_4 为 3 阶幂零矩阵, $A_4^2 = O$.

由

$$O = (A+2I)^k (A-3I)^l = \mathrm{diag}\{A_1^k A_3^l, A_2^k A_4^l\},$$

且 A_1 和 A_4 为幂零矩阵满足 $A_1^3 = A_4^2 = O$, 而 A_2 和 A_3 为满秩矩阵.

则只需通过幂零矩阵的部分乘方得到零矩阵, A_1 需自乘 3 次, 而 A_4 只需 2 次. 因此, A 的最小多项式是

$$m_A(\lambda) = (\lambda+2)^3(\lambda-3)^2$$

一般来说, 若

$$A = \operatorname{diag}(A_1, A_2, \dots, A_m),$$

且方阵 A_i 的最小多项式为

$$m_{A_i}(\lambda), \quad 1 \le i \le m,$$

则 A 的最小多项式 $m_A(\lambda)$ 是 $m_{A_1}(\lambda), m_{A_2}(\lambda), \dots, m_{A_m}(\lambda)$ 的最小公倍式. 这是因为 A_i 的最小多项式 $m_{A_i}(\lambda)$ $(1 \le i \le m)$ 必可整除 A 的最小多项式 $m_A(\lambda)$, 即

$$m_{A_i}(\lambda) | m_A(\lambda)$$

例如, 例 5 中的 $A = \begin{pmatrix} A_1 & O \\ O & A_2 \end{pmatrix}$. 且 $A_1 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.$A = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$. A_1 的最小多项式是 $(\lambda-2)^2$, A_2 的最小多项式是 $(\lambda-3)(\lambda-2)$. 而这两个多项式的最小公倍式是 $(\lambda-3)(\lambda-2)^2$, 故

$$m_A(\lambda) = (\lambda-3)(\lambda-2)^2.$$

之后会讨论如何由方阵 A 的最小多项式判断 A 是否可对角化. 在此之前, 我们先给出有关矩阵乘积的秩的一个重要不等式.

定理: 设 A 是 $m\times n$ 矩阵, B 是 $n\times s$ 矩阵, 则

$\mathrm{rank}A+\mathrm{rank}B-n\le \mathrm{rank}(AB)\le \min\{\mathrm{rank}A,\mathrm{rank}B\}$

证明: 只证前一个不等式. 设

$$B=(b_1,b_2,\ldots,b_r;b_{r+1},\ldots,b_s)=(B_1;B_2),$$

其中 $\mathrm{rank}B=\mathrm{rank}B_1=r$, 那么 B_2 中的任意一列 b_j $(r+1\le j\le s)$ 都可由 $b_1,b_2,\ldots,b_r$ 线性表出, 从而存在 $r\times(s-r)$ 矩阵 G, 使 $B_2=B_1G$, 即 $B_2-B_1G=O$. 令

$$Q=\begin{pmatrix} I_r & -G \\ O & I_{s-r}\end{pmatrix}.$$

证明 (续): 显然, Q 是 s 阶可逆方阵, 且有

$$\begin{aligned} BQ &= (B_1, B_2)\begin{pmatrix} I_r & -G \\ O & I_{s-r} \end{pmatrix} \\ &= (B_1, B_2 - B_1 G) = (b_1, b_2, \dots, b_r; O), \end{aligned}$$

其中向量组 $b_1, b_2, \dots, b_r$ 是线性无关的. 于是

$$ABQ = (Ab_1, Ab_2, \dots, Ab_r; O),$$

且 $Ab_1, Ab_2, \dots, Ab_r$ 中有 $\mathrm{rank}(AB)$ 个向量是线性无关的, 因而在这些向量中至多有 $r - \mathrm{rank}(AB)$ 个零向量. 但齐次线性方程组 $AX = 0$ 有 $n - \mathrm{rank}A$ 个线性无关的解, 因此

$$r - \mathrm{rank}(AB) = \mathrm{rank}B - \mathrm{rank}(AB) \le n - \mathrm{rank}A,$$

即前一个不等式成立,

例 7. 设 A_i $(i=1,2,...,k)$ 都是 n 阶方阵. 证明当 $A_1A_2...A_k=O$ 时有下述不等式:

$$\sum_{i=1}^{k}(n-\mathrm{rank}A_i)\geq n.$$

即

$$\sum_{i=1}^{k}\mathrm{rank}A_i\leq(k-1)n,$$

证明: 由于这两个不等式是等价的, 所以只需证明其中的一个即可.

证明 (续): 根据前定理, 有

$$
\begin{aligned}
0 &= \mathrm{rank}(A_1A_2...A_k) \\
&\geq \mathrm{rank}(A_1A_2...A_{k-1}) + \mathrm{rank}A_k - n \\
&\geq \mathrm{rank}(A_1A_2...A_k) + \mathrm{rank}A_{k-1} + \mathrm{rank}A_k - 2n \\
&\geq ... \geq \sum_{i=1}^{k} \mathrm{rank}A_i - (k-1)n,
\end{aligned}
$$

即

$$\sum_{i=1}^{k} \mathrm{rank}A_i \leq (k-1)n.$$

方阵的特征子空间

(在 §1.2 中定义了线性变换的特征子空间)
定义: 设 λ_0 是 n 阶方阵 A 的特征值, 则集合

$$E_{\lambda_0} := \{x | (A-\lambda_0 I)x = 0\} = N(A-\lambda_0 I)$$

称为 A 关于 λ_0 的特征子空间. 显然, 特征子空间 E_{λ_0} 是由 A 关于 λ_0 的所有特征向量, 再添加零向量所组成的.
$\dim E_{\lambda_0} = n - \mathrm{rank}(A-\lambda_0 I)$ 称为 λ_0 的几何重数, 就是 A 关于 λ_0 的线性无关特征向量的最大个数, 并且 λ_0 的几何重数不大于其代数重数.

定理: n 阶方阵 A 可对角化 (即相似于对角矩阵) 的充分必要条件是, A 的最小多项式没有重根.

证明: 必要性: 由于对角矩阵可以看作是 1 阶方阵所组成的对角块矩阵, 而 1 阶方阵的最小多项式是一次多项式, 所以 A 的最小多项式没有重根.

充分性: 设 $m_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_s)$,
则有 $(A - \lambda_1 I)(A - \lambda_2 I)...(A - \lambda_s I) = O$
故由例 7 知 $\sum_{i=1}^{k}(n - \text{rank}(A - \lambda_i I) \geq n.$, 即

$$\text{dim}E_{\lambda_1} + \text{dim}E_{\lambda_2} + ... + \text{dim}E_{\lambda_s} \geq n.$$

证明 (续): 但另一方面 $\dim E_{\lambda_i} \le n_i \ (1 \le i \le s)$, 故

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} + \ldots + \dim E_{\lambda_s} \le n_1 + n_2 + \ldots + n_s = n.$$

这里 n_i 是 λ_i 的代数重数. 因此

$$\dim E_{\lambda_i} = n_i, \quad 1 \le i \le s,$$

即 A 的任一特征值 λ_i 的几何重数等于 λ_i 得代数重数, 故 A 可对角化.

例 8. 设 n 阶方阵 A 满足关系式

$$A^3 - 3A^2 - A + 3I = O$$

证明 A 必可对角化.

证明: 由题设可知

$$g(\lambda) = \lambda^3 - 3\lambda^2 - \lambda + 3$$

是 A 的零化多项式, 又因

$$g(\lambda) = (\lambda + 1)(\lambda - 1)(\lambda - 3)$$

没有重根, 且 $m_A(\lambda) | g(\lambda)$, 故 $m_A(\lambda)$ 没有重根. 因此, A 必可对角化.