1 矩阵的相似与相抵

例题 1.1 设矩阵

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & -1\\ x & -2 & 0\\ -1 & -1 & y\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1 & 0 & z\\ 0 & -1 & 0\\ z & 0 & 0\end{bmatrix}$$

相似.

(1) 求 x,y,z.

(2) 求可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$.

(3) 对 $m\in\mathbb{N}^*$, 求 $\boldsymbol{A}^m$.

解.

(1) 相似的矩阵应当具有相同的特征多项式. 为此, 有

$$\begin{aligned}\det(\lambda\boldsymbol{I}-\boldsymbol{A}) &= \begin{vmatrix}\lambda-1 & -1 & 1\\ -x & \lambda+2 & 0\\ 1 & 1 & \lambda-y\end{vmatrix} = \begin{vmatrix}\lambda-1 & -1 & 1\\ -x & \lambda+2 & 0\\ 1-(\lambda-y)(\lambda-1) & 1+\lambda-y & 0\end{vmatrix}\\ &= -\begin{vmatrix}x & \lambda+2\\ \lambda^2-(1+y)\lambda+(y-1) & \lambda+(1-y)\end{vmatrix}\\ &= \lambda^3-(y-1)\lambda^2-(x+y+3)\lambda+(y-1)(2-x)\end{aligned}$$

$$\det(\lambda\boldsymbol{I}-\boldsymbol{B})=\begin{vmatrix}\lambda-1 & 0 & -z\\ 0 & \lambda+1 & 0\\ -z & 0 & \lambda\end{vmatrix}=(\lambda+1)\begin{vmatrix}\lambda-1 & -z\\ -z & \lambda\end{vmatrix}=(\lambda+1)(\lambda^2-\lambda-z^2)=\lambda^3-(1+z^2)\lambda-z^2$$

于是

$$\begin{cases}y-1=0\\ x+y+3=1+z^2\\ (y-1)(2-x)=-z^2\end{cases}$$

解得

$$x=-3,\quad y=1,\quad z=0$$

(2) 此时有

$$\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\lambda^3-\lambda$$

于是 $\boldsymbol{A}$ 的特征值为 $-1,0,1$.

考虑方程 $(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$, 对其系数矩阵做初等行变换可得

$$\begin{bmatrix}0 & -1 & 1\\3 & 3 & 0\\1 & 1 & 0\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & 1\\0 & -1 & 1\\0 & 0 & 0\end{bmatrix}$$

于是对应特征值 1 的一个特征向量为

$$\boldsymbol{\eta}_1=\begin{bmatrix}-1 & 1 & 1\end{bmatrix}^{\mathrm{t}}$$

考虑方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$, 对其系数矩阵做初等行变换可得

$$\begin{bmatrix}-1 & -1 & 1\\3 & 2 & 0\\1 & 1 & -1\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & 2\\0 & -1 & 3\\0 & 0 & 0\end{bmatrix}$$

于是对应特征值 0 的一个特征向量为

$$\boldsymbol{\eta}_2=\begin{bmatrix}-2 & 3 & 1\end{bmatrix}^{\mathrm{t}}$$

考虑方程 $(\boldsymbol{I}+\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$, 对其系数矩阵做初等行变换可得

$$\begin{bmatrix}2 & 1 & -1\\-3 & -1 & 0\\-1 & -1 & 2\end{bmatrix}\longrightarrow\begin{bmatrix}-1 & -1 & 2\\0 & -1 & 3\\0 & 2 & -6\end{bmatrix}\longrightarrow\begin{bmatrix}1 & 0 & 1\\0 & 1 & -3\\0 & 0 & 0\end{bmatrix}$$

于是对应特征值 0 的一个特征向量为

$$\boldsymbol{\eta}_3=\begin{bmatrix}-1 & 3 & 1\end{bmatrix}^{\mathrm{t}}$$

于是令

$$\boldsymbol{P}=\begin{bmatrix}\boldsymbol{\eta}_1 & \boldsymbol{\eta}_3 & \boldsymbol{\eta}_2\end{bmatrix}=\begin{bmatrix}-1 & -1 & -2\\1 & 3 & 3\\1 & 1 & 1\end{bmatrix}$$

即有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{1,-1,0\}=\boldsymbol{B}$$

(3) 由 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ 有

$$\boldsymbol{B}^m=(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^m=\boldsymbol{P}^{-1}\boldsymbol{A}^m\boldsymbol{P}$$

于是

$$\boldsymbol{A}^m=\boldsymbol{P}\boldsymbol{B}^m\boldsymbol{P}^{-1}$$

而

$$\boldsymbol{B}^m=\begin{bmatrix}1 & 0 & 0\\0 & (-1)^m & 0\\0 & 0 & 0\end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 & -2 & 1 & 0 & 0 \\ 1 & 3 & 3 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 2 & -1 & 0 & 0 \\ 0 & 2 & 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 2 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 \end{bmatrix}$$

于是

$$\boldsymbol{P}^{-1} = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{bmatrix}$$

于是

$$\boldsymbol{A}^m = \begin{bmatrix} -1 & -1\cdot(-1)^m & 0 \\ 1 & 3\cdot(-1)^m & 0 \\ 1 & 1\cdot(-1)^m & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 2 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} (-1)^{m+1} & 1+(-1)^{m+1} & -2 \\ 3\cdot(-1)^m & -1+3\cdot(-1)^m & 2 \\ (-1)^m & -1+(-1)^m & 2 \end{bmatrix}$$

例题 1.2 已知数域 $\mathbb{K}$ 上含参数 a 的矩阵

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 4 & -3 \\ 1 & a & 5 \end{bmatrix}$$

的特征多项式有一个二重根, 求 a 的值, 并判断 $\boldsymbol{A}$ 是否可以对角化.

解. 有

$$\begin{aligned} \det(\lambda \boldsymbol{I} - \boldsymbol{A}) &= \begin{bmatrix} \lambda-1 & -2 & 3 \\ 1 & \lambda-4 & 3 \\ -1 & -a & \lambda-5 \end{bmatrix} = \begin{bmatrix} \lambda-1 & -2 & 3 \\ 0 & \lambda-4-a & \lambda-2 \\ -1 & -a & \lambda-5 \end{bmatrix} \\ &= (\lambda-1)\begin{vmatrix} \lambda-(a+4) & \lambda-2 \\ -a & \lambda-5 \end{vmatrix} - \begin{vmatrix} -2 & 3 \\ \lambda-(a+4) & \lambda-2 \end{vmatrix} \\ &= (\lambda-1)(\lambda^2-9\lambda+3a+20) - (-5\lambda+3a+16) \\ &= \lambda^3 - 10\lambda^2 + (3a+34)\lambda - (6a+36) \end{aligned}$$

记上式为 $f(\lambda)$. 由于 $f(\lambda)=0$ 有一个二重根, 于是 $f(\lambda)=0$ 与 $f'(\lambda)=0$ 有相同的根, 即

$$\begin{cases} \lambda^3 - 10\lambda^2 + (3a+34)\lambda - (6a+36) = 0 \\ 3\lambda^2 - 20\lambda + (3a+34) = 0 \end{cases}$$

解得

$$\lambda = 2, \quad a = -2$$

此时

$$f(\lambda) = \lambda^3 - 10\lambda^2 + 28\lambda - 24 = (\lambda-2)^2(\lambda-6)$$

对于特征值 $\lambda = 2$, 对线性方程组 $(2\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 的系数矩阵做行变换可得

$$\begin{bmatrix} 1 & -2 & 3 \\ 1 & -2 & 3 \\ -1 & 2 & -3 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & -2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

故其几何重数为 2.

对于特征值 6, 对线性方程组 $(6\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ 的系数矩阵做行变换可得

$$\begin{bmatrix} 5 & -2 & 3 \\ 1 & 2 & 3 \\ -1 & 2 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & -12 & -12 \\ 0 & 4 & 4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

故其几何重数为 1. 于是可知 $\boldsymbol{A}$ 可以对角化.

例题 1.3 设 n 级矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 有相同的特征值.

(1) 证明: 如果 $\boldsymbol{A}$ 有 n 个互不相同的特征值, 则存在可逆矩阵 $\boldsymbol{P}$ 和矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{A} = \boldsymbol{PQ}, \boldsymbol{B} = \boldsymbol{QP}$.

(2) 如果前一问中 $\boldsymbol{A}$ 有 n 个互不相同的特征值这一条件删去, 能否得到相同的结论?

解.

(1) 由于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同有 n 个互不相同的特征值, 不妨设为 $\lambda_1, \cdots, \lambda_n$. 由于 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的阶数为 n, 因此它们的特征多项式一定为

$$f(\lambda) = \prod_{i=1}^{n} (\lambda - \lambda_i)$$

因而两者具有相同的特征多项式, 于是两者相似. 因此, 存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP}$. 于是令 $\boldsymbol{Q} = \boldsymbol{BP}^{-1}$ 就有

$$\boldsymbol{A} = \boldsymbol{PBP}^{-1} = \boldsymbol{PQ}$$

$$\boldsymbol{B} = \boldsymbol{BP}^{-1}\boldsymbol{P} = \boldsymbol{QP}$$

于是命题得证.

(2) 不能. 事实上, 将 $\boldsymbol{B} = \boldsymbol{QP}$ 变形为 $\boldsymbol{Q} = \boldsymbol{BP}^{-1}$ 代入 $\boldsymbol{A} = \boldsymbol{PQ}$ 可得

$$\boldsymbol{A} = \boldsymbol{PBP}^{-1}$$

这要求 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似. 当特征值的数目小于 n 时, 两者的特征多项式的某一项的次数可能不同, 仍然可以保证特征值相同, 但不再相似, 结论也就不一定成立了.

例题 1.4 设 n 级方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 = 2\boldsymbol{A}$. 证明: $\boldsymbol{A}$ 可以对角化.

证明. 考虑任一非零的 $\boldsymbol{x} \in \mathbb{K}^n$ 使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

两边左乘 $\boldsymbol{A}$ 可得

$$\boldsymbol{A}^2\boldsymbol{x} = \lambda \boldsymbol{A}\boldsymbol{x}$$

由 $\boldsymbol{A}^2 = 2\boldsymbol{A}$ 可得

$$2\boldsymbol{A}\boldsymbol{x} = \lambda(\boldsymbol{A}\boldsymbol{x})$$

又 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, 于是有

$$\lambda(\lambda - 2)\boldsymbol{x} = \mathbf{0}$$

由于 $\boldsymbol{x} \neq \mathbf{0}$, 从而 $\lambda = 0$ 或 $\lambda = 2$. 于是 $\boldsymbol{A}$ 仅有特征值 0 和 2.

由于 $\boldsymbol{A}(\boldsymbol{A} - 2\boldsymbol{I}) = \mathbf{0}$, 于是根据 Sylvester 秩不等式可知

$$\operatorname{rank} \boldsymbol{A} + \operatorname{rank}(\boldsymbol{A} - 2\boldsymbol{I}) \leqslant \operatorname{rank} \mathbf{0} + n = n$$

另一方面 $\boldsymbol{A} + (2\boldsymbol{I} - \boldsymbol{A}) = 2\boldsymbol{I}$. 对于任意的 n 级方阵 $\boldsymbol{P}$ 和 $\boldsymbol{Q}$ 总有

$$\operatorname{rank}(\boldsymbol{P} + \boldsymbol{Q}) \leqslant \operatorname{rank} \begin{bmatrix} \boldsymbol{P} + \boldsymbol{Q} \\ \boldsymbol{Q} \end{bmatrix} = \operatorname{rank} \begin{bmatrix} \boldsymbol{P} \\ \boldsymbol{Q} \end{bmatrix} \leqslant \operatorname{rank} \boldsymbol{P} + \operatorname{rank} \boldsymbol{Q}$$

于是

$$\operatorname{rank} \boldsymbol{A} + \operatorname{rank}(2\boldsymbol{I} - \boldsymbol{A}) \geqslant \operatorname{rank}(2\boldsymbol{I}) = n$$

从而

$$\operatorname{rank} \boldsymbol{A} + \operatorname{rank}(2\boldsymbol{I} - \boldsymbol{A}) = n$$

也即 $\boldsymbol{A}$ 的属于特征值 0 的特征空间和属于特征值 2 的特征空间的维数之和为 n, 因此 $\boldsymbol{A}$ 可对角化. □

例题 1.5 设 $\boldsymbol{A}, \boldsymbol{B}$ 是数域 $\mathbb{K}$ 上可对角化的 n 级方阵. 称 $\boldsymbol{A}, \boldsymbol{B}$ 可同时对角化, 如果存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 与 $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ 均为对角矩阵. 证明:$\boldsymbol{A}, \boldsymbol{B}$ 可同时对角化, 当且仅当它们可交换.

证明. $\Rightarrow$: 设 $\boldsymbol{C} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$, 依题意它们都是对角矩阵. 由于对角矩阵均可交换, 因此

$$\boldsymbol{C}\boldsymbol{D} = \boldsymbol{D}\boldsymbol{C}$$

于是

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$$

从而

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{A}\boldsymbol{P}$$

从而

$$\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$$

于是两者可交换.

$\Leftarrow$: 由于 $\boldsymbol{A}$ 可对角化, 因此对 $\mathbb{K}^n$ 做 $\boldsymbol{A}$ 的特征空间的直和分解:

$$\mathbb{K}^n = \bigoplus_{\lambda} E_\lambda(\boldsymbol{A})$$

任取 $\boldsymbol{A}$ 的特征值 λ 和 $\boldsymbol{x} \in E_\lambda(\boldsymbol{A})$, 则有 $\boldsymbol{Ax} = \lambda\boldsymbol{x}$. 又因为 $\boldsymbol{A}, \boldsymbol{B}$ 可交换, 于是

$$\boldsymbol{ABx} = \boldsymbol{BAx} = \boldsymbol{B}\lambda\boldsymbol{x} = \lambda(\boldsymbol{Bx})$$

从而 $\boldsymbol{Bx} \in E_\lambda(\boldsymbol{A})$, 因此 $E_\lambda(\boldsymbol{A})$ 是 $\boldsymbol{B}$ 的不变子空间. 由于 $\boldsymbol{B}$ 可对角化, 因此 $\boldsymbol{B}|_{E_\lambda(\boldsymbol{A})}$ 也可以对角化, 于是在 $E_\lambda(\boldsymbol{A})$ 中可以选择一组 $\boldsymbol{B}$ 的特征向量作为这空间的基. 依照定义, 这组特征向量也是 $\boldsymbol{A}$ 的特征向量.

于是对每个 $E_\lambda(\boldsymbol{A})$ 重复相同的操作, 就能找到 n 个线性无关的向量, 它们同时是 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的特征向量, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可同时对角化. □

例题 1.6 设 $\mathbb{K}$ 上的 n 级方阵 $\boldsymbol{T}$ 在 $\mathbb{K}$ 中有 n 个互不相同的特征值. 证明: 与 $\boldsymbol{T}$ 可交换的矩阵 $\boldsymbol{S}$ 一定可以对角化.

证明. 考虑 $\boldsymbol{T}$ 的特征值 $\lambda_1, \cdots, \lambda_n$ 和对应的特征空间 $E_1, \cdots, E_n$. 考虑 $\boldsymbol{x}_i \in E_i$ 作为该特征空间的基, 则有 $\boldsymbol{Tx}_i = \lambda_i\boldsymbol{x}_i$.

由于 $\boldsymbol{S}, \boldsymbol{T}$ 可交换, 于是

$$\boldsymbol{T}(\boldsymbol{Sx}_i) = \boldsymbol{TSx}_i = \boldsymbol{STx}_i = \boldsymbol{S}\lambda\boldsymbol{x}_i = \lambda(\boldsymbol{Sx}_i)$$

于是 $\boldsymbol{Sx}_i \in E_i$. 由于 $\dim E_i = 1$, 因此总存在 $\mu_i \in \mathbb{K}$ 使得

$$\boldsymbol{Sx}_i = \mu_i\boldsymbol{x}_i$$

即 $\boldsymbol{x}_i$ 也是 $\boldsymbol{S}$ 的特征向量. 重复上述操作, 取出的向量 $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ 是 $\boldsymbol{S}$ 的 n 个线性无关的特征向量, 因此 $\boldsymbol{S}$ 可对角化. □

例题 1.7 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 都是 $\mathbb{K}$ 上的 n 级方阵, 且满足 $\boldsymbol{AB} - \boldsymbol{BA} = \boldsymbol{C}$, $\boldsymbol{AC} = \boldsymbol{CA}$. 证明: 对任意 $k \in \mathbb{N}^+$ 都有 $\operatorname{tr}(\boldsymbol{C}^k) = 0$.

证明. 当 $k = 1$ 时有

$$\operatorname{tr}(\boldsymbol{C}) = \operatorname{tr}(\boldsymbol{AB}) - \operatorname{tr}(\boldsymbol{BA}) = 0$$

当 $k > 1$ 时有

$$\boldsymbol{C}^k = \boldsymbol{CC}^{k-1} = \boldsymbol{ABC}^{k-1} - \boldsymbol{BAC}^{k-1} = \boldsymbol{A}(\boldsymbol{BC}^{k-1}) - (\boldsymbol{BC}^{k-1})\boldsymbol{A}$$

记 $\boldsymbol{BC}^{k-1} = \boldsymbol{P}$, 则有

$$\operatorname{tr}(\boldsymbol{C}^k) = \operatorname{tr}(\boldsymbol{AP}) - \operatorname{tr}(\boldsymbol{PA}) = 0$$

□

例题 1.8 设 $\boldsymbol{A}$ 是 $\mathbb{K}$ 上的 n 级方阵, 满足对数域 $\mathbb{K}$ 上的任一 n 级方阵 $\boldsymbol{X}$ 都有 $\mathrm{tr}(\boldsymbol{AX})=0$. 证明: $\boldsymbol{A}=\boldsymbol{0}$.

证明. 设 $\boldsymbol{A}=(a_{ij})$, 则

$$\boldsymbol{A}\boldsymbol{E}_{ji}=\sum_{k=1}^{n}a_{kj}\boldsymbol{E}_{ki}$$

从而 $\boldsymbol{A}\boldsymbol{E}_{ij}$ 的对角线元素依次为 $0,\cdots,0,a_{ik},0,\cdots,0$. 由于 $\mathrm{tr}(\boldsymbol{A}\boldsymbol{E}_{ji})=0$, 因此 $a_{ij}=0$. 上述式子对任意 i,j 都成立, 因此 $\boldsymbol{A}=\boldsymbol{0}$. □

例题 1.9 证明: 一个 n 级复方阵 $\boldsymbol{B}$ 可对角化当且仅当存在一个次数不高于 $n-1$ 的复系数多项式 g 和一个有 n 个不同的复特征值的 n 级复方阵 $\boldsymbol{A}$, 满足 $\boldsymbol{B}$ 与 $g(\boldsymbol{A})$ 相似.

证明. $\Rightarrow$: 由于 $\boldsymbol{B}$ 可对角化, 因此设其相似于对角矩阵 $\boldsymbol{D}=\mathrm{diag}\{\mu_1,\cdots,\mu_n\}$, 则存在可逆矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{B}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$$

任取 n 个互不相同的 $\lambda_i\in\mathbb{C}$, 令

$$\boldsymbol{A}=\boldsymbol{P}\,\mathrm{diag}\{\lambda_1,\cdots,\lambda_i\}\boldsymbol{P}^{-1}$$

根据拉格朗日插值定理, 存在次数不高于 $n-1$ 的多项式函数 g 使得

$$g(\lambda_i)=\mu_i$$

从而

$$g(\boldsymbol{A})=\boldsymbol{P}\,\mathrm{diag}\{g(\lambda_1),\cdots,g(\lambda_n)\}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}=\boldsymbol{B}$$

从而 $\boldsymbol{B}\sim g(\boldsymbol{A})$.

$\Leftarrow$: 由于 $\boldsymbol{A}$ 有 n 个不同的特征值, 因此 $\boldsymbol{A}$ 可对角化. 于是存在可逆矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{A}=\boldsymbol{P}\,\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}\boldsymbol{P}^{-1}$$

而对任意 $m\in\mathbb{N}^+$ 有

$$\boldsymbol{A}^m=(\boldsymbol{P}\,\mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}\boldsymbol{P}^{-1})^m=\boldsymbol{P}\,\mathrm{diag}\{\lambda_1^m,\cdots,\lambda_n^m\}\boldsymbol{P}^{-1}$$

于是

$$g(\boldsymbol{A})=\boldsymbol{P}\,\mathrm{diag}\{g(\lambda_1),\cdots,g(\lambda_n)\}\boldsymbol{P}^{-1}$$

由于 $\boldsymbol{B}\sim g(\boldsymbol{A})$, 又 $g(\boldsymbol{A})\sim\mathrm{diag}\{g(\lambda_1),\cdots,g(\lambda_n)\}$, 于是 $\boldsymbol{B}$ 可对角化. □

例题 1.10 定义 n 级矩阵

$$\boldsymbol{C}_n=\begin{bmatrix}\boldsymbol{0}_{(n-1)\times 1} & \boldsymbol{I}_{n-1}\\ 1 & \boldsymbol{0}_{1\times(n-1)}\end{bmatrix}$$

n 级矩阵 $\boldsymbol{S}$ 被称作循环的, 如果存在次数不超过 $n-1$ 的多项式 p 使得 $\boldsymbol{S}=p(\boldsymbol{C}_n)$. 证明: 一个 n 级矩阵 $\boldsymbol{A}$ 可对角化当且仅当它相似于某个循环矩阵 $\boldsymbol{S}$.

证明. $\Rightarrow$: 由于 $\boldsymbol{A}$ 可对角化, 于是设 $\boldsymbol{A}\sim \mathrm{diag}\{\lambda_1,\cdots,\lambda_n\}=\boldsymbol{D}$. 由于

$$\det(\lambda\boldsymbol{I}-\boldsymbol{C}_n)=\lambda^n-1$$

于是 $\boldsymbol{C}_n$ 的全部特征值为所有 n 次单位根 $\omega_k=\mathrm{e}^{\frac{2k\pi\mathrm{i}}{n}},k=1,\cdots,n$, 从而 $\boldsymbol{C}_n$ 可对角化, 于是设可逆矩阵 $\boldsymbol{P}$ 使得

$$\boldsymbol{C}_n=\boldsymbol{P}\,\mathrm{diag}\{\omega_1,\cdots,\omega_n\}\boldsymbol{P}^{-1}$$

根据拉格朗日插值定理, 总存在次数不高于 $n-1$ 的多项式 p 使得

$$p(\omega_i)=\lambda_i$$

从而

$$\boldsymbol{S}p(\boldsymbol{C}_n)=\boldsymbol{P}\,\mathrm{diag}\{p(\omega_1),\cdots,p(\omega_n)\}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$$

从而 $\boldsymbol{S}\sim\boldsymbol{D}$, 因而 $\boldsymbol{A}\sim\boldsymbol{S}$.

$\Leftarrow$: 由前面的推导可得

$$\boldsymbol{S}=\boldsymbol{P}\,\mathrm{diag}\{p(\omega_1),\cdots,p(\omega_n)\}\boldsymbol{P}^{-1}$$

即 $\boldsymbol{S}$ 可对角化. 又 $\boldsymbol{A}\sim\boldsymbol{S}$, 因此 $\boldsymbol{A}$ 也可对角化. □

例题 1.11 设 n 级方阵 $\boldsymbol{A}$ 的特征多项式为 f, 证明: $f(\boldsymbol{A})=\boldsymbol{0}$.

证明. 根据伴随矩阵的性质可知对任意方阵 $\boldsymbol{S}$ 总有 $\boldsymbol{S}\boldsymbol{S}^*=(\det\boldsymbol{S})\boldsymbol{I}$. 现在令 $\boldsymbol{S}=\lambda\boldsymbol{I}-\boldsymbol{A}$, 则有

$$(\lambda\boldsymbol{I}-\boldsymbol{A})(\lambda\boldsymbol{I}-\boldsymbol{A})^*=\det(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{I}$$

即

$$(\lambda\boldsymbol{I}-\boldsymbol{A})(\lambda\boldsymbol{I}-\boldsymbol{A})^*=f(\lambda)\boldsymbol{I}$$

而 $(\lambda\boldsymbol{I}-\boldsymbol{A})^*$ 的每个矩阵元都是 λ 的不超过 $n-1$ 次的多项式. 于是存在矩阵 $\boldsymbol{S}_0,\cdots,\boldsymbol{S}_{n-1}$ 使得

$$(\lambda\boldsymbol{I}-\boldsymbol{A})^*=\boldsymbol{S}_0+\lambda\boldsymbol{S}_1+\cdots+\lambda^{n-1}\boldsymbol{S}^{n-1}$$

于是

$$\lambda\sum_{i=0}^{n-1}\lambda^i\boldsymbol{S}_i-\boldsymbol{A}\sum_{i=0}^{n-1}\lambda^i\boldsymbol{S}_i=f(\lambda)\boldsymbol{I}$$

即

$$\lambda^n\boldsymbol{S}_{n-1}+\lambda^{n-1}(\boldsymbol{S}_{n-2}-\boldsymbol{A}\boldsymbol{S}_{n-1})+\cdots+\lambda(\boldsymbol{S}_0-\boldsymbol{A}\boldsymbol{S}_1)-\boldsymbol{A}\boldsymbol{S}_0=f(\lambda)\boldsymbol{I}$$

设 $f(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0$, 则有

$$\lambda^n\boldsymbol{S}_{n-1}+\lambda^{n-1}(\boldsymbol{S}_{n-2}-\boldsymbol{A}\boldsymbol{S}_{n-1})+\cdots+\lambda(\boldsymbol{S}_0-\boldsymbol{A}\boldsymbol{S}_1)-\boldsymbol{A}\boldsymbol{S}_0=\lambda^n\boldsymbol{I}+a_{n-1}\lambda^{n-1}\boldsymbol{I}+\cdots+a_0\boldsymbol{I}$$

从而

$$\boldsymbol{S}_{n-1} = \boldsymbol{I}, \quad \boldsymbol{S}_{n-2} - \boldsymbol{A}\boldsymbol{S}_{n-1} = a_{n-1}\boldsymbol{I}, \cdots, \quad -\boldsymbol{A}\boldsymbol{S}_0 = a_0\boldsymbol{I}$$

于是

$$\begin{aligned} f(\boldsymbol{A}) &= \boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_1\boldsymbol{A} + a_0\boldsymbol{I} \\ &= \boldsymbol{A}^n\boldsymbol{S}_{n-1} + \boldsymbol{A}^{n-1}(\boldsymbol{S}_{n-2} - \boldsymbol{A}\boldsymbol{S}_{n-1}) + \cdots - \boldsymbol{A}\boldsymbol{S}_0 \\ &= \boldsymbol{0} \end{aligned}$$

命题得证. □

例题 1.12 设 $\boldsymbol{A}$ 是 n 级可逆方阵, 证明: 存在一个不超过 $n-1$ 次的多项式 g 使得 $g(\boldsymbol{A}) = \boldsymbol{A}^{-1}$.

证明. 考虑 $\boldsymbol{A}$ 的特征多项式 $f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$. 由前一题的结论可知

$$f(\boldsymbol{A}) = \boldsymbol{A}^n + a_{n-1}\boldsymbol{A}^{n-1} + \cdots + a_1\boldsymbol{A} + a_0\boldsymbol{I} = \boldsymbol{0}$$

由于 $\boldsymbol{A}$ 可逆, 因此 $a_0 \neq 0$. 于是则有

$$-\frac{1}{a_0}(\boldsymbol{A}^{n-1} + a_{n-1}\boldsymbol{A}^{n-2} + \cdots + a_1\boldsymbol{I})\boldsymbol{A} = \boldsymbol{I}$$

于是令 $g(x) = -\dfrac{1}{a_0}(x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_1)$ 即可. □

例题 1.13 设 n 级方阵 $\boldsymbol{A}, \boldsymbol{B}$ 满足 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{A} + \boldsymbol{B}^{2025}$. 证明: $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$.

证明. 先证明 $\boldsymbol{B} - \boldsymbol{I}$ 可逆. 考虑 n 级方阵 $\boldsymbol{X}$ 使得 $(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$, 即 $\boldsymbol{X} = \boldsymbol{B}\boldsymbol{X}$, 则有

$$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{X} = \cdots = \boldsymbol{B}^{2025}\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{B} - \boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$$

于是 $\boldsymbol{B} - \boldsymbol{I}$ 可逆. 由前一题的结论可知存在多项式 f 使得

$$f(\boldsymbol{B} - \boldsymbol{I}) = (\boldsymbol{B} - \boldsymbol{I})^{-1}$$

这也是一个仅与 $\boldsymbol{B}$ 相关的多项式, 不妨记 $g(\boldsymbol{B}) = f(\boldsymbol{B} - \boldsymbol{I})$. 从而

$$\boldsymbol{A} = \boldsymbol{B}^{2025}g(\boldsymbol{B})$$

这是一个关于 $\boldsymbol{B}$ 的多项式, 于是它与 $\boldsymbol{B}$ 可交换. □

例题 1.14 设 $\boldsymbol{A}, \boldsymbol{B}$ 分别是 $\mathbb{C}$ 上的 m, n 级方阵, 证明: 矩阵方程 $\boldsymbol{A}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{B} = \boldsymbol{0}$ 有非零解 $\boldsymbol{X}$ 当且仅当 $\boldsymbol{A}, \boldsymbol{B}$ 有共同的特征值.

证明. $\Leftarrow$: 假定 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值 λ, 那么 $\boldsymbol{B}^{\mathrm{t}}$ 也有特征值 λ. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}^{\mathrm{t}}$ 对应的特征向量分别为 $\boldsymbol{\alpha}\in\mathbb{K}^m$ 和 $\boldsymbol{\beta}\in\mathbb{K}^n$, 则有

$$\boldsymbol{A}\boldsymbol{\alpha}=\lambda\boldsymbol{\alpha},\quad \boldsymbol{B}^{\mathrm{t}}\boldsymbol{\beta}=\lambda\boldsymbol{\beta}$$

从而

$$\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{t}}-\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{t}}\boldsymbol{B}=\boldsymbol{A}\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{t}}-\boldsymbol{\alpha}\boldsymbol{B}^{\mathrm{t}}\boldsymbol{\beta}=\lambda\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{t}}-\lambda\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{t}}=\boldsymbol{0}$$

从而 $\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{t}}$ 是该矩阵方程的解.

$\Rightarrow$: 假定矩阵方程 $\boldsymbol{A}\boldsymbol{X}-\boldsymbol{X}\boldsymbol{B}=\boldsymbol{0}$ 有非零解 $\boldsymbol{X}$. 设 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 没有公共的特征值. 考虑 $\boldsymbol{A}$ 的特征多项式, 将其分解为 $\mathbb{C}$ 中的一次多项式的乘积:

$$f_{\boldsymbol{A}}(\lambda)=\prod_{j=1}^{m}(\lambda-\lambda_j)$$

于是 $f_{\boldsymbol{A}}(\boldsymbol{B})=\prod\limits_{j=1}^{m}(\boldsymbol{B}-\lambda_j\boldsymbol{I}_n)$. 由于 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 没有共同的特征值, 从而 $\det(\boldsymbol{B}-\lambda_j\boldsymbol{I}_n)\neq 0$, 从而 $\det f_{\boldsymbol{A}}(\boldsymbol{B})\neq 0$, 因而 $f_{\boldsymbol{A}}(\boldsymbol{B})$ 可逆. 由 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{X}\boldsymbol{B}$ 可得

$$f_{\boldsymbol{A}}(\boldsymbol{A})\boldsymbol{X}=\boldsymbol{X}f_{\boldsymbol{A}}(\boldsymbol{B})$$

由于 $f_{\boldsymbol{A}}(\boldsymbol{A})=\boldsymbol{0}$, 而 $f_{\boldsymbol{A}}(\boldsymbol{B})\neq\boldsymbol{0}$, 于是只能 $\boldsymbol{X}=\boldsymbol{0}$, 与假设矛盾. 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值. □

例题 1.15 证明: 幂等矩阵的秩和迹相等.

证明. 设 $\boldsymbol{A}$ 是 $\mathbb{K}$ 上的 n 级幂等矩阵. 如果 λ 是 $\boldsymbol{A}$ 的特征值, 则存在非零的 $\boldsymbol{x}\in\mathbb{K}^n$ 使得 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$. 两边左乘 $\boldsymbol{A}$ 可得

$$\boldsymbol{A}^2\boldsymbol{x}=\lambda\boldsymbol{A}\boldsymbol{x}=\lambda^2\boldsymbol{x}$$

又 $\boldsymbol{A}^2=\boldsymbol{A}$, 于是

$$\boldsymbol{A}^2\boldsymbol{x}=\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$$

从而 $(\lambda^2-\lambda)\boldsymbol{x}=\boldsymbol{0}$. 由于 $\boldsymbol{x}\neq\boldsymbol{0}$, 于是 $\lambda^2-\lambda=0$, 即 $\lambda=0$ 或 1.

设 $\operatorname{rank}\boldsymbol{A}=r$. 当 $r=0$ 时 $\boldsymbol{A}=\boldsymbol{0}$, 当 $r=n$ 时 $\boldsymbol{A}=\boldsymbol{I}$, 命题都成立. 现在设 $0<r<n$.

对于特征值 0, 齐次线性方程组 $(0\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的解空间 W_0 满足

$$\dim W_0=n-\operatorname{rank}(-\boldsymbol{A})=n-r$$

由于 $\boldsymbol{A}$ 是幂等矩阵, 因此 $\dim\boldsymbol{A}+\dim(\boldsymbol{I}-\boldsymbol{A})=n$.

对于特征值 1, 齐次线性方程组 $(1\boldsymbol{I}-\boldsymbol{A})\boldsymbol{x}=\boldsymbol{0}$ 的解空间 W_1 满足

$$\dim W_1=n-\operatorname{rank}(\boldsymbol{I}-\boldsymbol{A})=n-(n-r)=r$$

从而

$$\dim W_0+\dim W_1=n$$

从而 $\boldsymbol{A}$ 可对角化. 于是

$$\boldsymbol{A} \sim \operatorname{diag}\{1, \cdots, 1, 0, \cdots, 0\}$$

右边的对角矩阵的迹和秩都等于 r. 由于相似的矩阵具有相同的迹和秩, 因此 $\boldsymbol{A}$ 的迹和秩相等, 得证. $\square$