第一章 特征值

注 代数基本定理保证了任一 $n(n \geq 1)$ 阶复矩阵 A 或 n 维复线性空间 V 上的线性变换 φ 至少有一个复特征值 λ_0, 线性方程组的求解理论保证了 λ_0 至少有一个复特征向量。如果是在数域 $\mathbb{F}$ 上, 则需要 A 或 φ 的特征值 λ_0 属于 $\mathbb{F}$, 然后线性方程组的求解理论才能保证 λ_0 在 $\mathbb{F}^n$ 或 V 中有对应的特征向量。因此, **后面如无特殊说明, 总是假设在复数域 $\mathbb{C}$ 上考虑问题**.

1.1 特征值与特征向量

定义 1.1 (线性变换的特征值和特征向量)

设 φ 是数域 $\mathbb{K}$ 上线性空间 V 上的线性变换, 若 $\lambda_0 \in \mathbb{K}, x \in V$ 且 $x \neq 0$, 使

$$\varphi(x) = \lambda_0 x,$$

则称 λ_0 是线性变换 φ 的一个**特征值**, 向量 x 称为 φ 关于特征值 λ_0 的**特征向量**. ♣

笔记 显然 φ 的关于特征值 λ_0 的全体特征向量加上零向量构成 V 的子空间.

定义 1.2 (线性变换的特征子空间)

设 λ_0 是线性空间 V 上的线性变换 φ 的特征值, 令

$$V_{\lambda_0} = \{\alpha \in V \mid \varphi(\alpha) = \lambda_0 \alpha\} = \{\alpha \in V \mid \alpha\text{是}\varphi\text{的属于}\lambda_0\text{的特征向量}\} \cup \{0\},$$

则显然 V_{λ_0} 是 V 的子空间, 称为 φ 的属于特征值 λ_0 的**特征子空间**. ♣

笔记 显然 V_{λ_0} 是 φ 的不变子空间.

定义 1.3 (矩阵的特征值和特征向量)

设 A 是数域 $\mathbb{F}$ 上的 n 阶方阵, 若存在 $\lambda_0 \in \mathbb{F}$ 及 n 维非零列向量 α, 使

$$A\alpha = \lambda_0 \alpha$$

式成立, 则称 λ_0 为矩阵 A 的一个**特征值**,α 为 A 关于特征值 λ_0 的**特征向量**. ♣

定义 1.4 (矩阵的特征子空间)

设 λ_0 是 $\mathbb{F}$ 上的 n 阶矩阵 A 的特征值, 令

$$V_{\lambda_0} = \{x \in \mathbb{F}^n \mid Ax = \lambda_0 x\} = \{x \in \mathbb{F}^n \mid x\text{是}A\text{的属于}\lambda_0\text{的特征向量}\} \cup \{0\},$$

则 V_{λ_0} 是线性方程组 $(\lambda_0 I_n - A)x = 0$ 的解空间, 从而是 $\mathbb{F}^n$ 的子空间, 称为 A 的属于特征值 λ_0 的**特征子空间。** ♣

定义 1.5 (特征多项式)

设 A 是 n 阶方阵, 称 $|\lambda I_n - A|$ 为 A 的**特征多项式**. ♣

定理 1.1 (特征值的和与积)

矩阵 A 的 n 个特征值的和与积分别为

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(A),$$
$$\lambda_1 \lambda_2 \cdots \lambda_n = |A|.$$

♡

证明 设

$$|\lambda I_n - A| = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$
$$= (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

由 Vieta 定理知 $\lambda_1+\lambda_2+\cdots+\lambda_n = -a_1$,$\lambda_1\lambda_2\cdots\lambda_n = (-1)^n a_n$. 由例题**??**可知 $a_1 = -(a_{11}+a_{22}+\cdots+a_{nn}) = -\mathrm{tr}(A)$,$a_n = (-1)^n|A|$. 因此矩阵 A 的 n 个特征值的和与积分别为

$$\lambda_1+\lambda_2+\cdots+\lambda_n = \mathrm{tr}(A),$$
$$\lambda_1\lambda_2\cdots\lambda_n = |A|.$$

□

定义 1.6 (特征多项式)

设 φ 是线性空间 V 上的线性变换,φ 在 V 的某组基下的表示矩阵为 A, 由相似矩阵有相同特征值知 $|\lambda I_n - A|$ 与基或表示矩阵的选取无关, 称 $|\lambda I_n - A|$ 为 φ 的**特征多项式**, 记为 $|\lambda I_V - \varphi|$.

定理 1.2 (复方阵必相似于上三角阵)

任何复方阵必相似于一个上三角阵, 并且对角元素都是其特征值.

注 一般数域 $\mathbb{K}$ 上的矩阵未必相似于上三角阵.

证明 设 A 是 n 阶复方阵, 现对 n 用数学归纳法. 当 $n=1$ 时结论显然成立. 假设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵 A 来证明. 设 λ_1 是 A 的一个特征值, 则存在非零列向量 α_1, 使

$$A\alpha_1 = \lambda_1\alpha_1.$$

将 α_1 作为 C_n 的一个基向量, 并扩展为 C_n 的一组基 $\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$. 将这些基向量按照列分块方式拼成矩阵 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 则 P 为 n 阶非异阵, 且

$$AP = A(\alpha_1,\alpha_2,\cdots,\alpha_n) = (A\alpha_1,A\alpha_2,\cdots,A\alpha_n)$$
$$= (\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix},$$

其中 A_1 是一个 $n-1$ 阶方阵. 注意到 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 非异, 上式即为

$$P^{-1}AP = \begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix}.$$

因为 A_1 是一个 $n-1$ 阶方阵, 所以由归纳假设可知, 存在 $n-1$ 阶非异阵 Q, 使 $Q^{-1}A_1Q$ 是一个上三角阵. 令

$$R = \begin{pmatrix}1 & O \\ O & Q\end{pmatrix},$$

则 R 是 n 阶非异阵, 且

$$R^{-1}P^{-1}APR = \begin{pmatrix}1 & O \\ O & Q\end{pmatrix}^{-1}\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix}\begin{pmatrix}1 & O \\ O & Q\end{pmatrix}$$
$$= \begin{pmatrix}1 & O \\ O & Q^{-1}\end{pmatrix}\begin{pmatrix}\lambda_1 & * \\ O & A_1\end{pmatrix}\begin{pmatrix}1 & O \\ O & Q\end{pmatrix}$$
$$= \begin{pmatrix}\lambda_1 & * \\ O & Q^{-1}A_1Q\end{pmatrix}.$$

这是一个上三角阵, 它与 A 相似, 并且对角元素都是其特征值. □

推论 1.1

若数域 $\mathbb{K}$ 上的 n 阶方阵 A 的特征值全在 $\mathbb{K}$ 中, 则存在 $\mathbb{K}$ 上的非异阵 P, 使 $P^{-1}AP$ 是一个上三角阵.

证明 由复方阵必相似于上三角阵的证明类似可得. □

命题 1.1

1. 设 φ 为 n 维线性空间 V 上的线性变换, 则 φ 在 V 上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in V$.
2. 设 A 为 n 阶复矩阵, 则 A 在复数域上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in \mathbb{C}^n$.

证明

1. 任取 V 的一组基 $\{e_1, e_2, \cdots, e_n\}$, 设 φ 在这组基下的表示矩阵为 A, 由代数学基本定理可知, 特征多项式 $|\lambda I_V - \varphi| = |\lambda I_n - A|$ 在复数域上至少有一个根 $\lambda_0 \in \mathbb{C}$. 又由线性方程组理论可知, $(\lambda_0 I_n - A)x = 0$ 一定有非零解 $(x_1, x_2, \cdots, x_n)^T$, 即 $\lambda_0 I_n \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. 记 $\alpha_0 = (e_1, e_2, \cdots, e_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, 则
$$\begin{aligned} \varphi(\alpha_0) &= \varphi(e_1, e_2, \cdots, e_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = (e_1, e_2, \cdots, e_n) A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= (e_1, e_2, \cdots, e_n) \cdot \lambda_0 I_n \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda_0 (e_1, e_2, \cdots, e_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda_0 \alpha_0. \end{aligned}$$
故 φ 在 V 上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in V$.
2. 由代数学基本定理可知, 特征多项式 $|\lambda I_n - A|$ 在复数域上至少有一个根 $\lambda_0 \in \mathbb{C}$. 又由线性方程组理论可知, $(\lambda_0 I_n - A)x = 0$ 一定有非零解 $\alpha_0 \in \mathbb{C}^n$. 故 A 在复数域上至少存在一个特征值 $\lambda_0 \in \mathbb{C}$ 及其特征向量 $\alpha_0 \in \mathbb{C}^n$.

□

1.1.1 直接利用定义计算和证明

例题 1.1 设 V 是 n 阶矩阵全体组成的线性空间, φ 是 V 上的线性变换: $\varphi(X) = AX$, 其中 A 是一个 n 阶矩阵。求证: φ 和 A 具有相同的特征值 (重数可能不同)。

证明 设 λ_0 是 A 的特征值, x_0 是对应的特征向量, 即 $Ax_0 = \lambda_0 x_0$。令 $X = (x_0, 0, \ldots, 0)$, 则 $\varphi(X) = AX = \lambda_0 X$ 且 $X \neq 0$, 因此 λ_0 也是 φ 的特征值.

反之, 设 λ_0 是 φ 的特征值, X 是对应的特征向量, 即 $\varphi(X) = AX = \lambda_0 X$。令 $X = (x_1, x_2, \ldots, x_n)$ 为列分块, 设第 i 个列向量 $x_i \neq 0$, 则 $Ax_i = \lambda_0 x_i$, 因此 λ_0 也是 A 的特征值. □

例题 1.2 设 λ_1, λ_2 是矩阵 A 的两个不同的特征值, α_1, α_2 分别是 λ_1, λ_2 的特征向量, 求证: $\alpha_1 + \alpha_2$ 必不是 A 的特征向量。

证明 用反证法, 设 $A(\alpha_1 + \alpha_2) = \mu(\alpha_1 + \alpha_2)$, 又
$$A(\alpha_1 + \alpha_2) = A\alpha_1 + A\alpha_2 = \lambda_1 \alpha_1 + \lambda_2 \alpha_2,$$
于是 $(\lambda_1 - \mu)\alpha_1 + (\lambda_2 - \mu)\alpha_2 = 0$。由于属于不同特征值的特征向量线性无关, 故有 $\lambda_1 = \mu, \lambda_2 = \mu$, 从而 $\lambda_1 = \lambda_2$, 引出矛盾. □

命题 1.2

设 φ 是线性空间 V 上的线性变换, V 有一个直和分解:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_m,$$

其中 V_i 都是 φ-不变子空间。

(1) 设 φ 限制在 V_i 上的特征多项式为 $f_i(\lambda)$, 求证: φ 的特征多项式

$$f(\lambda) = f_1(\lambda)f_2(\lambda)\cdots f_m(\lambda).$$

(2) 设 λ_0 是 φ 的特征值, $V_0 = \{v \in V \mid \varphi(v) = \lambda_0 v\}$ 为特征子空间, $V_{i,0} = V_i \cap V_0 = \{v \in V_i \mid \varphi(v) = \lambda_0 v\}$, 求证:

$$V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}.$$

证明

(1) 取 V_i 的一组基, 将它们拼成 V 的一组基。记 A_i 是 φ 在 V_i 上的限制在 V_i 所取基下的表示矩阵, 则由定理**??**可知 φ 在 V 的这组基下的表示矩阵为分块对角矩阵 $A = \mathrm{diag}(A_1, A_2, \cdots, A_m)$, 于是

$$f(\lambda) = |\lambda I_n - A| = |\lambda I - A_1||\lambda I - A_2|\cdots|\lambda I - A_m|,$$

即 $f(\lambda) = f_1(\lambda)f_2(\lambda)\cdots f_m(\lambda)$.

(2) 任取 $\alpha \in V_0$, 设 $\alpha = \alpha_1 + \alpha_2 + \cdots + \alpha_m$, 其中 $\alpha_i \in V_i$, 则

$$\varphi(\alpha_1) + \varphi(\alpha_2) + \cdots + \varphi(\alpha_m) = \varphi(\alpha) = \lambda_0\alpha = \lambda_0\alpha_1 + \lambda_0\alpha_2 + \cdots + \lambda_0\alpha_m.$$

注意到 $\varphi(\alpha_i) \in V_i, \varphi(\alpha) = \lambda_0\alpha \in V$, 故由直和的等价条件 (5) 可得 $\varphi(\alpha_i) = \lambda_0\alpha_i$, 即 $\alpha_i \in V_{i,0}$, 从而 $V_0 = V_{1,0} \oplus V_{2,0} \oplus \cdots \oplus V_{m,0}$。注意到 $V_{i,0} \subseteq V_i$, 故

$$V_{i,0} \cap (V_{1,0} + \cdots + V_{i-1,0}) \subseteq V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}, \quad 2 \le i \le m,$$

于是由直和的等价条件 (2) 可知上述为直和。

□

推论 1.2

对分块对角矩阵 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_m\}$ 的任一特征值 λ_0, 其代数重数等于每个分块的代数重数之和, 其几何重数等于每个分块的几何重数之和.

证明 将命题 1.2的条件和结论代数化之后, 即可得到结论. □

命题 1.3 (特征向量的延拓)

设 n 阶分块对角矩阵 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_m\}$, 其中 A_i 是 n_i 阶矩阵。

(1) 任取 A_i 的特征值 λ_i 及其特征向量 $x_i \in \mathbb{C}^{n_i}$, 求证: 可在 x_i 的上下添加适当多的零, 得到非零向量 $\tilde{x}_i \in \mathbb{C}^n$, 使得 $A\tilde{x}_i = \lambda_i\tilde{x}_i$, 即 $\tilde{x}_i$ 是 A 关于特征值 λ_i 的特征向量, 称为 x_i 的**延拓**.

(2) 任取 A 的特征值 λ_0, 并设 λ_0 是 $A_{i_1}, \cdots, A_{i_r}$ 的特征值, 但不是其他 $A_j\ (1 \le j \le m, j \ne i_1, \cdots, i_r)$ 的特征值, 求证: A 关于特征值 λ_0 的特征子空间的一组基可取为 $A_{i_1}(1 \le k \le r)$ 关于特征值 λ_0 的特征子空间的一组基的延拓的并集.

证明

(1) 令 $\tilde{x}_i = (0, \cdots, 0, x_i, 0, \cdots, 0)'$, 即 $\tilde{x}_i$ 的第 i 块为 x_i, 其余块均为 0, 显然 $\tilde{x}_i \ne 0$。容易验证 $A\tilde{x}_i = \lambda_i\tilde{x}_i$, 故结论成立。

(2) 由命题 1.2(2)以及直和的等价条件 (5) 即得.

□

例题 1.3 设 A 是 n 阶整数矩阵, p, q 为互素的整数且 $q > 1$。求证：矩阵方程 $Ax = \dfrac{p}{q}x$ 必无非零解。

证明 用反证法。设上述矩阵方程有非零解，则 $\frac{p}{q}$ 为 A 的特征值，即为特征多项式 $f(\lambda)=\lambda^n+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$ 的根。由于 A 是整数矩阵，故 $f(\lambda)$ 为整数系数多项式。由整数系数多项式有有理根的必要条件可知 $q\mid 1$，从而 $q=\pm1$，于是 $q\mid p$，这与 p,q 互素矛盾。 □

例题 1.4 求下列 n 阶矩阵的特征值：

$$A=\begin{pmatrix}0&a&\cdots&a&a\\b&0&\cdots&a&a\\\vdots&\vdots&\ddots&\vdots&\vdots\\b&b&\cdots&0&a\\b&b&\cdots&b&0\end{pmatrix}.$$

解 若 $a=0$ 或 $b=0$，则 A 是主对角元全为零的下三角或上三角矩阵，故 A 的特征值全为零. 设 $a\neq0,b\neq0$，则由命题??可知：若 $a\neq b$，则 $|\lambda I_n-A|=\dfrac{a(\lambda+b)^n-b(\lambda+a)^n}{a-b}$. 设 $\dfrac{b}{a}$ 的 n 次方根为 $\omega_i(1\le i\le n)$，则

$$|\lambda I_n-A|=\frac{a(\lambda+b)^n-b(\lambda+a)^n}{a-b}=0\Rightarrow\left(\frac{\lambda+b}{\lambda+a}\right)^n=\frac{b}{a}$$
$$\Rightarrow\frac{\lambda+b}{\lambda+a}=w_i\ (1\le i\le n)\Rightarrow\lambda=\frac{a\omega_i-b}{1-\omega_i}(1\le i\le n).$$

从而 A 的特征值为 $\dfrac{a\omega_i-b}{1-\omega_i}(1\le i\le n)$.

若 $a=b$，则 $|\lambda I_n-A|=(\lambda-(n-1)a)(\lambda+a)^{n-1}$，从而 A 的特征值为 $(n-1)a$ (1 重)，$-a$ ($n-1$ 重).

综上，容易验证当 $a=b=0$ 或 $ab\neq0$ 时,A 有完全的特征向量系或有 n 个不同的特征值，从而此时 A 可对角化. 若 A 可对角化也不难得到 $a=b=0$ 或 $ab\neq0$. 故 A 可对角化的充分必要条件是 $a=b=0$ 或 $ab\neq0$. □

1.1.2 正向利用矩阵的多项式

定义 1.7 (矩阵多项式)

若 A 是一个 n 阶矩阵,$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$ 是一个多项式，记

$$f(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0I_n.$$

♣

命题 1.4

设 n 阶矩阵 A 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，$f(x)$ 是一个多项式，则 $f(A)$ 的全部特征值为 $f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)$.

♠

注 这个命题告诉我们：如果能够将一个复杂矩阵写成一个简单矩阵的多项式，那么就可以由简单矩阵的特征值得到复杂矩阵的特征值.

证明 因为任一 n 阶矩阵均复相似于上三角阵，可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1&*&\cdots&*\\0&\lambda_2&\cdots&*\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\lambda_n\end{pmatrix}.$$

因为上三角阵的和、数乘及乘方仍是上三角阵，经计算不难得到

$$P^{-1}f(A)P=f(P^{-1}AP)=\begin{pmatrix}f(\lambda_1)&*&\cdots&*\\0&f(\lambda_2)&\cdots&*\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&f(\lambda_n)\end{pmatrix}.$$

因此 $f(A)$ 的全部特征值为 $f(\lambda_1), f(\lambda_2), \cdots, f(\lambda_n)$. □

例题 1.5 设 n 阶矩阵 A 的全体特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 求 $2n$ 阶矩阵

$$\begin{pmatrix} A^2 & A \\ A^2 & A \end{pmatrix}$$

的全体特征值。

证明 由命题??可知

$$\left|\lambda I_{2n} - \begin{pmatrix} A^2 & A \\ A^2 & A \end{pmatrix}\right| = \left|\begin{pmatrix} \lambda I_n - A & -A^2 \\ -A^2 & \lambda I_n - A \end{pmatrix}\right| = |\lambda I_n - A - A^2||\lambda I_n - A + A^2|.$$

由命题 1.4可知 $A + A^2$ 的全体特征值为 $\lambda_i + \lambda_i^2 (1 \le i \le n)$, $A - A^2$ 的全体特征值为 $\lambda_i - \lambda_i^2 (1 \le i \le n)$, 因此所求矩阵的全体特征值为

$$\lambda_1 + \lambda_1^2, \lambda_1 - \lambda_1^2, \lambda_2 + \lambda_2^2, \lambda_2 - \lambda_2^2, \cdots, \lambda_n + \lambda_n^2, \lambda_n - \lambda_n^2.$$

□

命题 1.5 (循环矩阵的特征值)

求下列循环矩阵的特征值:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}.$$

♠

解 设 $J = \begin{pmatrix} O & I_{n-1} \\ 1 & O \end{pmatrix}$, $f(x) = a_1 + a_2 x + a_3 x^2 + \cdots + a_n x^{n-1}$, 则由循环矩阵的性质 2 可知 $A = f(J)$。经简单计算可得

$$|\lambda I_n - J| = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & -1 \\ -1 & 0 & 0 & \cdots & 0 & \lambda \end{vmatrix} \xlongequal{\text{按第一列展开}} \lambda \begin{vmatrix} \lambda & -1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \\ 0 & 0 & \cdots & 0 & \lambda \end{vmatrix} + (-1)^{n+1}(-1) \begin{vmatrix} -1 & 0 & \cdots & 0 & 0 \\ \lambda & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & \lambda & -1 \end{vmatrix}$$

$$= \lambda^n + (-1)^{n+2}(-1)^{n-1} = \lambda^n - 1,$$

于是 J 的特征值为

$$\omega_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \quad 0 \le k \le n-1$$

因此 A 的特征值为 $f(1), f(\omega_1), \cdots, f(\omega_{n-1})$。 □

定义 1.8 (友矩阵)

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -a_{n-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}$$

称为多项式 $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ 的 **友矩阵**.

♣

命题 1.6 (友矩阵的特征多项式及特征值)

设首一多项式 $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$, $f(x)$ 的友矩阵

$$C=\begin{pmatrix}0&0&\cdots&0&-a_0\\1&0&\cdots&0&-a_1\\0&1&\cdots&0&-a_2\\\vdots&\vdots&\ddots&\vdots&\vdots\\0&0&\cdots&1&-a_{n-1}\end{pmatrix}.$$

(1) 求证：矩阵 C 的特征多项式就是 $f(\lambda)$。

(2) 设 $f(x)$ 的根为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，$g(x)$ 为任一多项式，求以 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$ 为根的 n 次多项式。

证明

(1)

$$|xE-\boldsymbol{A}|=\begin{vmatrix}x&0&0&\cdots&0&a_0\\-1&x&0&\cdots&0&a_1\\0&-1&x&\cdots&0&a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&x&a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}\xlongequal[i=n,n-1,\cdots,2]{xr_i+r_{i-1}}\begin{vmatrix}0&0&0&\cdots&0&x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\\-1&0&0&\cdots&0&x^{n-1}+a_{n-1}x^{n-2}+\cdots+a_2x+a_1\\0&-1&0&\cdots&0&x^{n-2}+a_{n-1}x^{n-3}+\cdots+a_3x+a_2\\\vdots&\vdots&\vdots&&\vdots&\vdots\\0&0&0&\cdots&0&x^2+a_{n-1}x+a_{n-2}\\0&0&0&\cdots&-1&x+a_{n-1}\end{vmatrix}$$

$$\xlongequal{\text{按第一行展开}}\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a\right)(-1)^{n+1}\begin{vmatrix}-1&0&0&\cdots&0\\0&-1&0&\cdots&0\\\vdots&\vdots&\vdots&&\vdots\\0&0&0&\cdots&0\\0&0&0&\cdots&-1\end{vmatrix}$$

$$=\left(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a\right)(-1)^{n+1}(-1)^{n-1}$$

$$=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a=f(x).$$

(2) 由假设及 (1) 的结论可知 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 C 的全体特征值，故由命题 1.4可知 $g(\lambda_1),g(\lambda_2),\cdots,g(\lambda_n)$ 是 $g(C)$ 的全体特征值，从而 $h(x)=|xI_n-g(C)|$ 即为所求的多项式。

□

1.1.3 反向利用矩阵的多项式

命题 1.7

设 n 阶矩阵 A 适合一个多项式 $g(x)$, 即 $g(A)=O$, 则 A 的任一特征值 λ_0 也必适合 $g(x)$, 即 $g(\lambda_0)=0$.

证明 设 α 是 A 关于特征值 λ_0 的特征向量, 经简单计算得

$$g(\lambda_0)\alpha=g(A)\alpha=0.$$

而 $\alpha\neq 0$, 因此 $g(\lambda_0)=0$. □

命题 1.8 (幂零矩阵关于特征值的充要条件)

求证：n 阶矩阵 A 为幂零矩阵的充要条件是 A 的特征值全为零。

证明 若 A 为幂零矩阵, 即存在正整数 k, 使得 $A^k=O$, 则由命题 1.7可知 A 的任一特征值 λ_0 也适合 x^k, 于是 $\lambda_0=0$。

反之, 证法一:若 A 的特征值全为零, 则存在可逆矩阵 P, 使得 $P^{-1}AP=B$ 为上三角矩阵且主对角元素全为零。由上三角阵性质 (1) 可知 $B^n=O$, 于是 $A^n=(PBP^{-1})^n=PB^nP^{-1}=O$, 即 A 为幂零矩阵。

证法二:也可以利用 Cayley-Hamilton 定理来证明, 由于 A 的特征值全为零, 故其特征多项式为 λ^n, 从而 $A^n=O$. □

例题 1.6 设 V 是数域 $\mathbb{F}$ 上的 n 阶方阵全体构成的线性空间, n 阶方阵

$$P=\begin{pmatrix}0&\cdots&0&1\\0&\cdots&1&0\\\vdots&&\vdots&\vdots\\1&0&\cdots&0\end{pmatrix},$$

V 上的线性变换 η 定义为 $\eta(X)=PX'P$. 试求 η 的全体特征值及其特征向量.

笔记 任意 n 阶矩阵 A 左乘 P 相当于行倒排, 右乘 P 矩阵相当于列倒排.

解 由 $P=P', P^2=I_n$ 容易验证 $\eta^2(X)=P(PX'P)P=X$, 即 $\eta^2=I_V$, 于是 η 的特征值也适合多项式 x^2-1, 从而特征值只能是 ±1.

设 $\eta(X_0)=PX_0'P=\pm X_0$, 这等价于 $(PX_0')P=\pm PX_0$, 即 PX_0 为对称矩阵或反对称矩阵.

令 $PX_0=E_{ii}, E_{ij}+E_{ji}$ (对称矩阵空间的基向量), 容易证明 η 关于特征值 1 的线性无关的特征向量为 $X_0=PE_{ii}\ (1\le i\le n), P(E_{ij}+E_{ji})\ (1\le i<j\le n)$.

令 $PX_0=E_{ij}-E_{ji}$ (反对称矩阵空间的基向量), 容易证明 η 关于特征值 -1 的线性无关的特征向量为 $X_0=P(E_{ij}-E_{ji})\ (1\le i<j\le n)$. 注意到这些特征向量恰好构成 V 的一组基, 故 η 的特征值为 1 $\dfrac{n(n+1)}{2}$ 重, -1 $\dfrac{n(n-1)}{2}$ 重. □

例题 1.7 设 n 阶方阵 A 的每行每列只有一个元素非零, 并且那些非零元素为 1 或 -1, 证明: A 的特征值都是单位根.

证明 设 S 为由每行每列只有一个元素非零, 并且那些非零元素为 1 或 -1 的所有 n 阶方阵构成的集合, 由排列组合可得 $\overline{\overline{S}}=2^n n!$, 即 S 是一个有限集合. 注意到矩阵 $M\in S$ 当且仅当 $M=P_1P_2\cdots P_r$, 其中 P_k 是初等矩阵 P_{ij} 或 $P_i(-1)$, 因此对任意的 $M,N\in S$, $MN\in S$. 特别地, 由 $A\in S$ 可知 $A^k\in S\ (k\ge1)$, 即 $\{A,A^2,A^3,\cdots\}\subseteq S$, 于是存在正整数 $k>l$, 使得 $A^k=A^l$. 注意到 $|A|=\pm1$, 故 A 可逆, 于是 $A^{k-l}=I_n$, 从而 A 的特征值适合多项式 $x^{k-l}-1$, 即为单位根. □

例题 1.8 设 A 是 n 阶实方阵, 又 I_n-A 的特征值的模长都小于 1, 求证: $0<|A|<2^n$.

证明 设 A 的特征值为 $\lambda_1,\cdots,\lambda_n$, 则 I_n-A 的特征值为 $1-\lambda_1,\cdots,1-\lambda_n$. 由假设 $|1-\lambda_i|<1$, 若 λ_i 是实数, 则 $0<\lambda_i<2$; 若 λ_i 是虚数, 则 $\overline{\lambda_i}$ 也是 A 的特征值, 此时 $1-\overline{\lambda_i}$ 也是 I_n-A 的特征值. 从而 $|1-\lambda_i|<1, |1-\overline{\lambda_i}|<1$, 于是

$$|1-\lambda_i^2|=|(1-\lambda_i)(1-\overline{\lambda_i})|=|1-\lambda_i||1-\overline{\lambda_i}|<1.$$

因此 $0<\lambda_i^2<2$, 故此时 $0<\lambda_i<\sqrt{2}$.

综上, 无论 λ_i 是实数还是虚数, 都有 $0<|\lambda_i|<2$. 由于 $|A|$ 等于所有特征值之积, 故 $0<|A|<2^n$. □

命题 1.9 (逆矩阵的特征值)

设 n 阶矩阵 A 是可逆矩阵, 且 A 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则 A^{-1} 的全部特征值为 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$。

证明 首先注意到 A 是可逆矩阵, $\lambda_1\lambda_2\cdots\lambda_n=|A|\ne0$, 因此每个 $\lambda_i\ne0$ (事实上, A 可逆的充分必要条件是它的特征值全不为零)。由复方阵必相似于上三角阵可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1&*&\cdots&*\\0&\lambda_2&\cdots&*\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\lambda_n\end{pmatrix}.$$

因为上三角矩阵的逆矩阵仍然是上三角矩阵, 经过计算不难得到

$$P^{-1}A^{-1}P=(P^{-1}AP)^{-1}=\begin{pmatrix}\lambda_1^{-1} & * & \cdots & * \\ 0 & \lambda_2^{-1} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^{-1}\end{pmatrix}.$$

因此 A^{-1} 的全部特征值为 $\lambda_1^{-1},\lambda_2^{-1},\cdots,\lambda_n^{-1}$. □

命题 1.10 (伴随矩阵的特征值)

设 n 阶矩阵 A 的全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$，求证：$A^*$ 的全体特征值为

$$\prod_{i\neq 1}\lambda_i,\prod_{i\neq 2}\lambda_i,\cdots,\prod_{i\neq n}\lambda_i.$$

证明 因为任一 n 阶矩阵均复相似于上三角矩阵，故可设

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}.$$

注意到上三角矩阵的伴随矩阵仍是上三角矩阵，经计算可得

$$P^{-1}A^*P=P^*A^*(P^{-1})^*=(P^{-1}AP)^*=\begin{pmatrix}\prod\limits_{i\neq 1}\lambda_i & * & \cdots & * \\ 0 & \prod\limits_{i\neq 2}\lambda_i & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \prod\limits_{i\neq n}\lambda_i\end{pmatrix}.$$

因此 A^* 的全部特征值为

$$\prod_{i\neq 1}\lambda_i,\prod_{i\neq 2}\lambda_i,\cdots,\prod_{i\neq n}\lambda_i.$$

□

1.1.4 特征值的降价公式

定理 1.3 (特征值的降价公式)

设 A 是 $m\times n$ 矩阵，B 是 $n\times m$ 矩阵，且 $m\geq n$。求证：

$$|\lambda I_m-AB|=\lambda^{m-n}|\lambda I_n-BA|.$$

特别地，若 A,B 都是 n 阶矩阵，则 AB 与 BA 有相同的特征多项式.

笔记 本质上就是打洞原理.

证明 证法一 (打洞原理): 当 $\lambda\neq 0$ 时, 考虑下列分块矩阵:

$$\begin{pmatrix}\lambda I_m & A \\ B & I_n\end{pmatrix},$$

因为 $\lambda I_m, I_n$ 都是可逆矩阵, 故由行列式的降阶公式可得

$$|I_n|\cdot|\lambda I_m-A(I_n)^{-1}B|=|\lambda I_m|\cdot|I_n-B(\lambda I_m)^{-1}A|,$$

即有

$$|\lambda I_m - AB| = \lambda^{m-n}|\lambda I_n - BA|$$

成立.

当 $\lambda = 0$ 时, 若 $m > n$, 则 $r(AB) \le \min\{r(A), r(B)\} \le \min\{m, n\} = n < m$, 故 $|-AB| = 0$, 结论成立; 若 $m = n$, 则 $|-AB| = (-1)^n|A||B| = |-BA|$, 结论也成立。

事实上, $\lambda = 0$ 的情形也可以用 Cauchy-Binet 公式来处理, 还可以通过摄动法由 $\lambda \ne 0$ 的情形来得到.

证法二 (相抵标准型): 设 A 的秩等于 r, 则存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q, 使得

$$PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

令

$$Q^{-1}BP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix},$$

其中 B_{11} 是 $r \times r$ 矩阵, 则

$$PABP^{-1} = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}, \quad Q^{-1}BAQ = \begin{pmatrix} B_{11} & O \\ B_{21} & O \end{pmatrix}.$$

因此

$$|\lambda I_m - AB| = \left|\begin{pmatrix} \lambda I_r - B_{11} & -B_{12} \\ O & \lambda I_{m-r} \end{pmatrix}\right| = \lambda^{m-r}|\lambda I_r - B_{11}|,$$

同理

$$|\lambda I_n - BA| = \left|\begin{pmatrix} \lambda I_r - B_{11} & O \\ -B_{21} & \lambda I_{n-r} \end{pmatrix}\right| = \lambda^{n-r}|\lambda I_r - B_{11}|.$$

比较上面两个式子即可得出结论。

证法三 (摄动法): 先证明 $m = n$ 的情形。若 A 可逆, 则 $BA = A^{-1}(AB)A$, 即 AB 和 BA 相似, 因此它们的特征多项式相等。对于一般的方阵 A, 可取到一列有理数 $t_k \to 0$, 使得 $t_kI_n + A$ 是可逆矩阵。由可逆情形的证明可得

$$|\lambda I_n - (t_kI_n + A)B| = |\lambda I_n - B(t_kI_n + A)|.$$

注意到上述两边的行列式都是 t_k 的多项式, 从而关于 t_k 连续. 上式两边同时取极限, 令 $t_k \to 0$, 即有 $|\lambda I_n - AB| = |\lambda I_n - BA|$ 成立。

再证明 $m > n$ 的情形。令

$$C = \begin{pmatrix} A & O \end{pmatrix}, \quad D = \begin{pmatrix} B \\ O \end{pmatrix},$$

其中 C, D 均为 $m \times m$ 分块矩阵, 则

$$CD = AB, \quad DC = \begin{pmatrix} BA & O \\ O & O \end{pmatrix}.$$

因此由方阵的情形可得

$$|\lambda I_m - AB| = |\lambda I_m - CD| = |\lambda I_m - DC| = \lambda^{m-n}|\lambda I_n - BA|.$$

□

例题 1.9 设 α 是 n 维实列向量且 $\alpha'\alpha = 1$，试求矩阵 $I_n - 2\alpha\alpha'$ 的特征值。

解 设 $A = I_n - 2\alpha\alpha'$，则由特征值的降价公式可得

$$|\lambda I_n - A| = |(\lambda-1)I_n + 2\alpha\alpha'| = (\lambda-1)^{n-1}(\lambda - 1 + 2\alpha'\alpha) = (\lambda-1)^{n-1}(\lambda+1).$$

因此, 矩阵 A 的特征值为 $1(n-1)$ 重,-1(1 重). 进一步, 容易验证 A 有完全的特征向量系 ($|-I_n - A|$ 为零, 但其 $n-1$ 阶子式不为零), 于是 A 可对角化. □

例题 1.10 设 A 为 n 阶方阵, α,β 为 n 维列向量, 试求矩阵 $A\alpha\beta'$ 的特征值。

解 设 $B = A\alpha\beta'$, 则由特征值的降价公式可得

$$|\lambda I_n - B| = |I_n - (A\alpha)\beta'| = \lambda^{n-1}(\lambda - \beta' A\alpha).$$

若 $\beta' A\alpha \neq 0$, 则 B 的特征值为 $0\,(n-1$ 重$)$, $\beta' A\alpha\,(1$ 重$)$. 进一步, 容易验证此时 B 有完全的特征向量系, 从而可对角化. 若 $\beta' A\alpha = 0$, 则 B 的特征值为 $0\,(n$ 重$)$.

综上, 容易验证 $B = \boldsymbol{A\alpha\beta'}$ 可对角化的充要条件是 $\boldsymbol{\beta' A\alpha} \neq 0$ 或 $\boldsymbol{A\alpha\beta'} = \boldsymbol{O}$. □

例题 1.11 设 $a_i\ (1 \leq i \leq n)$ 都是实数, 且 $a_1 + a_2 + \cdots + a_n = 0$, 试求下列矩阵的特征值:

$$A = \begin{pmatrix} a_1^2 & a_1a_2+1 & \cdots & a_1a_n+1 \\ a_2a_1+1 & a_2^2 & \cdots & a_2a_n+1 \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1+1 & a_na_2+1 & \cdots & a_n^2 \end{pmatrix}.$$

解 矩阵 A 可以分解为 $A = -I_n + BC$, 其中

$$B = \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}, \quad C = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

由特征值的降价公式得

$$\begin{aligned} |\lambda I_n - A| &= |(\lambda+1)I_n - BC| \\ &= (\lambda+1)^{n-2}|(\lambda+1)I_2 - CB|. \end{aligned}$$

注意到 $a_1 + a_2 + \cdots + a_n = 0$, 故有

$$CB = \begin{pmatrix} a_1^2 + a_2^2 + \cdots + a_n^2 & 0 \\ 0 & n \end{pmatrix}.$$

因此 A 的特征值为 $-1\,(n-2$ 重$)$, $n-1$, $a_1^2+a_2^2+\cdots+a_n^2-1$. 进一步, 若 a_i 全部为零, 则特征值 -1 和 $n-1$ 都有完全的特征向量系. 若 $\sum\limits_{i=1}^{n} a_i^2 = n$, 利用秩的降阶公式可得特征值 -1 和 $n-1$ 都有完全的特征向量系. 在剩余情况, 利用秩的降阶公式可得 3 个特征值都有完全的特征向量系. 因此, $\boldsymbol{A}$ 可对角化. 事实上, 即使去掉 $a_1+a_2+\cdots+a_n = 0$ 的条件, 也可以计算出 $\boldsymbol{A}$ 的全体特征值的代数重数和几何重数, 从而得到 $\boldsymbol{A}$ 可对角化. 这一结论的深层次背景是: $\boldsymbol{A}$ 是实对称矩阵, 从而可正交对角化. □

例题 1.12 设 A,B,C 分别是 $m\times m, n\times n, m\times n$ 矩阵, 满足: $AC = CB$, $\mathrm{r}(C) = r$. 求证: A 和 B 至少有 r 个相同的特征值。

注 不妨设 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的原因: 假设当 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 时, 结论已经成立, 则对于一般的满足条件的矩阵 C, 由条件我们有

$$r(C) = r, \quad AC = BC.$$

由 $r(C) = r$ 可知, 存在可逆矩阵 P,Q, 使得

$$PCQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

从而对 $AC = BC$ 两边同时左乘 P, 右乘 Q 得到

$$(PAP^{-1})(PCQ) = (PCQ)(Q^{-1}BQ).$$

于是由假设可知 PAP^{-1} 和 $Q^{-1}BQ$ 都至少有 r 个相同的特征值. 又因为相似矩阵有相同的特征值, 所以 A,B 也至少有 r 个相同的特征值. 故不妨设成立.

证明 设 P 为 m 阶非异阵, Q 为 n 阶非异阵, 使得

$$PCQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}.$$

注意到问题的条件和结论在相抵变换 $C \mapsto PCQ, A \mapsto PAP^{-1}, B \mapsto Q^{-1}BQ$ 下保持不变, 故不妨从一开始就假设 $C = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

为对应的分块, 则

$$AC = \begin{pmatrix} A_{11} & O \\ A_{21} & O \end{pmatrix}, \quad CB = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}.$$

由 $AC = CB$ 可得 $A_{11} = B_{11}, A_{21} = 0, B_{12} = 0$. 于是

$$|\lambda I_m - A| = |\lambda I_r - A_{11}| \cdot |\lambda I_{m-r} - A_{22}|,$$
$$|\lambda I_n - B| = |\lambda I_r - B_{11}| \cdot |\lambda I_{n-r} - B_{22}|.$$

从而 A, B 至少有 r 个相同的特征值 (即 $A_{11} = B_{11}$ 的特征值). □

1.1.5 特征值与特征多项式系数的关系

命题 1.11 (特征值与特征多项式系数的关系)

设 n 阶矩阵 A 的特征多项式为

$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n.$$

求证: a_r 等于 $(-1)^r$ 乘以 A 的所有 r 阶主子式之和, 即

$$a_r = (-1)^r \sum_{1\le i_1<i_2<\cdots<i_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}, \quad 1 \le r \le n.$$

进一步, 若设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r} = \sum_{1\le i_1<i_2<\cdots<i_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}, \quad 1 \le r \le n.$$

♠

注 上述结论中最常用的是 $r = 1$ 和 $r = n$ 的情形:

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{tr}(A), \quad \lambda_1\lambda_2\cdots\lambda_n = |A|.$$

特别地, **A 是非异阵的充要条件是 A 的特征值全不为零.** 因此, 特征值的计算是判断矩阵是否非异阵的重要依据.

证明 第一种结论是推论**??**.

由 Vieta 定理可得

$$\sum_{1\le i_1<i_2<\cdots<i_r\le n} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r} = (-1)^r a_r = (-1)^r (-1)^r \sum_{1\le i_1<i_2<\cdots<i_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}$$
$$= \sum_{1\le i_1<i_2<\cdots<i_r\le n} A\begin{pmatrix} i_1 & i_2 & \cdots & i_r \\ i_1 & i_2 & \cdots & i_r \end{pmatrix}, 1 \le r \le n.$$

因此第二种结论也成立. □

例题 1.13 设 n 阶方阵 A 满足

$$A^2 - A - 3I_n = O,$$

求证：$A - 2I_n$ 是非奇异阵。

笔记 用特征值判断矩阵非异性.

证明 用反证法。设 $A-2I_n$ 为奇异阵, 则 2 是 A 的特征值。注意到 A 适合

$$f(x)=x^2-x-3,$$

但特征值 2 却不适合 $f(x)$, 这与命题 1.7矛盾。 □

例题 1.14 设 P 是可逆矩阵, $B=PAP^{-1}-P^{-1}AP$, 求证: B 的特征值之和为零.

证明 由特征值与特征多项式系数的关系可知, 只要证 $\mathrm{tr}(B)=0$ 即可. 由迹的线性和交换性即得

$$\mathrm{tr}(B)=\mathrm{tr}(PAP^{-1})-\mathrm{tr}(P^{-1}AP)=\mathrm{tr}(A)-\mathrm{tr}(A)=0.$$

□

例题 1.15 设 n 阶实方阵 A 的特征值全是实数, 且 A 的一阶主子式之和与二阶主子式之和都等于零. 求证: A 是零矩阵.

证明 设 A 的特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_n$, 由条件和特征值与特征多项式系数的关系可知

$$\sum_{i=1}^{n}\lambda_i=\lambda_1+\lambda_2+\cdots+\lambda_n=0,$$
$$\sum_{1\le i<j\le n}\lambda_i\lambda_j=\lambda_1\lambda_2+\lambda_1\lambda_3+\cdots+\lambda_{n-1}\lambda_n=0.$$

则

$$\sum_{i=1}^{n}\lambda_i^2=\left(\sum_{i=1}^{n}\lambda_i\right)^2-2\sum_{1\le i<j\le n}\lambda_i\lambda_j=0.$$

由于 λ_i 都是实数, 故 $\lambda_i=0\ (1\le i\le n)$ 成立, 再由命题 1.8可知 A 为零矩阵. □

例题 1.16 设 $n(n\ge 3)$ 阶非异实方阵 A 的特征值都是实数, 且 A 的 $n-1$ 阶主子式之和等于零. 证明: 存在 A 的一个 $n-2$ 阶主子式, 其符号与 $|A|$ 的符号相反.

证明 设 A 的特征值为 $\lambda_1,\lambda_2,\ldots,\lambda_n$, 由 A 非异可知它们都是非零实数. 再由条件和例 6.24 可知

$$\sum_{1\le i_1<i_2<\cdots<i_{n-1}\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_{n-1}}=0. \tag{1.1}$$

将(1.1)式左边除以 $|A|=\lambda_1\lambda_2\cdots\lambda_n$ 可得

$$\sum_{i=1}^{n}\frac{1}{\lambda_i}=0, \tag{1.2}$$

将(1.2) 式左边平方, 并将平方项移到等式的右边可得

$$\sum_{1\le i<j\le n}\frac{1}{\lambda_i\lambda_j}=-\frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{\lambda_i}\right)^2<0, \tag{1.3}$$

将(1.3)式两边同时乘以 $|A|=\lambda_1\lambda_2\cdots\lambda_n$ 可得

$$\sum_{1\le i_1<i_2<\cdots<i_{n-2}\le n}\lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_{n-2}}=-\frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{\lambda_i^2}\right)|A|. \tag{1.4}$$

由 (1.4)式和特征值与特征多项式系数的关系可得

$$\sum_{1\le i_1<i_2<\cdots<i_{n-2}\le n}A\begin{pmatrix}i_1 & i_2 & \cdots & i_{n-2}\\ i_1 & i_2 & \cdots & i_{n-2}\end{pmatrix}=-\frac{1}{2}\left(\sum_{i=1}^{n}\frac{1}{\lambda_i^2}\right)|A|,$$

于是 A 的 $n-2$ 阶主子式之和与 $|A|$ 的符号相反, 从而至少存在 A 的一个 $n-2$ 阶主子式, 其符号与 $|A|$ 的符号相反. □

结论 设 n 阶方阵 A 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则对任意的正整数 k, A^k 的特征值为 $\lambda_1^k,\lambda_2^k,\cdots,\lambda_n^k$, 于是特征值的 k 次幂和

$$s_k=\lambda_1^k+\lambda_2^k+\cdots+\lambda_n^k=\mathrm{tr}(A^k),\quad k\ge 1.$$

若已知 n 阶方阵 A 的迹 $\operatorname{tr}(A^k)(1 \le k \le n)$, 则由 Newton 公式可以计算出特征值的初等对称多项式

$$\sigma_r = \sum_{1\le i_1<i_2<\cdots<i_r\le n} \lambda_{i_1}\lambda_{i_2}\cdots\lambda_{i_r}, \quad 1 \le r \le n,$$

从而可以确定特征多项式的系数, 最后便可计算出 A 的所有特征值.

例题 1.17 设 A 是 n 阶对合矩阵, 即 $A^2 = I_n$, 证明: $n - \operatorname{tr}(A)$ 为偶数, 并且 $\operatorname{tr}(A) = n$ 的充要条件是 $A = I_n$.

证明 由 $A^2 = I_n$ 可知 A 的特征值也适合 $x^2 - 1$, 从而只能是 ± 1. 设 A 的特征值为 $1(p$ 重$)$, $-1(q$ 重$)$, 则 $p + q = n$. 且 $\operatorname{tr}(A) = p - q$, 于是 $n - \operatorname{tr}(A) = 2q$ 为偶数. 若 $A = I_n$, 则 $\operatorname{tr}(A) = n$. 反之, 若 $\operatorname{tr}(A) = n$, 则由上述讨论可知 $p = n$, $q = 0$, 从而 -1 不是 A 的特征值, 即 $A + I_n$ 是非奇阵. 最后由 $A^2 = I_n$ 可得

$$(A - I_n)(A + I_n) = O \Rightarrow A - I_n = O \Rightarrow A = I_n.$$

□

例题 1.18 设 4 阶方阵 A 满足: $\operatorname{tr}(A^k) = k\ (1 \le k \le 4)$, 试求 A 的行列式.

证明 题目条件即为 $s_k = k\ (1 \le k \le 4)$, 要求 $|A| = \sigma_4$. 根据 Newton 公式 (白皮书这一部分还没看)

$$s_k - s_{k-1}\sigma_1 + \cdots + (-1)^k k\sigma_k = 0 \quad (1 \le k \le 4)$$

可依次算出 $\sigma_1 = 1, \sigma_2 = -\dfrac{1}{2}, \sigma_3 = \dfrac{1}{6}, \sigma_4 = \dfrac{1}{24}$. 故 $|A| = \dfrac{1}{24}$. 也可以直接利用例 5.64(白皮书这一部分还没看) 来计算 σ_4. □

命题 1.12 (幂零矩阵关于迹的充要条件)

求证: n 阶矩阵 A 是零矩阵的充要条件是 $\operatorname{tr}(A^k) = 0\ (1 \le k \le n)$.

证明 设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 若 A 是零矩阵, 则 A 的特征值全为零, 从而 $\operatorname{tr}(A^k) = s_k = 0\ (k \ge 1)$. 若 $s_k = \operatorname{tr}(A^k) = 0\ (1 \le k \le n)$, 则由 Newton 公式 (白皮书这一部分还没看) 或直接利用例 5.64(白皮书这一部分还没看) 可计算出 $\sigma_r = 0\ (1 \le r \le n)$, 于是 A 的特征多项式为 λ^n, 从而 A 的特征值全为零, 再由幂零矩阵关于特征值的充要条件可知 A 为零矩阵. □

定义 1.9 (线性变换的迹)

线性变换的迹定义为它在任一组基下的表示矩阵的迹.

笔记 因为矩阵的迹在相似变换下保持不变, 并且同一线性变换在不同基下的表示矩阵必相似, 所以同一线性变换在任意一组基下的表示矩阵的迹都相同, 故线性变换的迹是良定义的.

命题 1.13

设 φ 为 n 维线性空间,λ_0 为 φ 的一个特征值,V_0 为特征值 λ_0 的特征子空间, 则存在 V_0 上的一组基, 使得 φ 在 V_0 上的限制 $\varphi|_{V_0}$ 在这组基下的表示矩阵为 $\dim V_0$ 阶的对角阵 $\operatorname{diag}\{\lambda_0, \lambda_0, \cdots, \lambda_0\}$, 从而 $\operatorname{tr}(\varphi|_{V_0}) = \lambda_0 \dim V_0$.

注 因为线性变换的特征子空间一定是不变子空间, 所以线性变换在其特征子空间上做限制后仍是线性变换, 因此线性变换在其特征子空间上的限制是良定义的.

证明 设 x_1 是 φ 属于 λ_0 的特征向量, 将其扩充成 V_0 的一组基 $\{x_1, x_2, \cdots, x_r\}$, 则 $r = \dim V_0$. 注意到 $x_1, x_2, \cdots, x_r$ 也是 $\varphi|_{V_0}$ 属于 λ_0 的特征向量, 从而

$$\varphi|_{V_0}(x_1, x_2, \cdots, x_r) = (\lambda_0 x_1, \lambda_0 x_2, \cdots, \lambda_0 x_r) = (x_1, x_2, \cdots, x_r)\begin{pmatrix} \lambda_0 & & & \\ & \lambda_0 & & \\ & & \ddots & \\ & & & \lambda_0 \end{pmatrix}.$$

故 $\varphi|_{V_0}$ 在 $\{x_1, x_2, \cdots, x_r\}$ 下的表示矩阵为 $\dim V_0$ 阶对角阵 $\operatorname{diag}\{\lambda_0, \lambda_0, \cdots, \lambda_0\}$. □

命题 1.14

设 A, B, C 是 n 阶矩阵, 其中 $C = AB - BA$. 若它们满足条件 $AC = CA, BC = CB$, 求证: C 的特征值全为零.

注 若将条件减弱为 $ABC = CAB, BAC = CBA$, 则上述结论不再成立. 原因如下:

如将条件减弱为如题所述, 则结论不再成立, 可参考下面的反例:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

由计算可得 $C = AB - BA, ABC = CAB, CBA = BAC$, 但 C 的特征值为 1 和 -1。

笔记 从下述的证法一不难看出: 只需要 $AC = CA$ 和 $BC = CB$ 这两个条件中的一个就能证明本题的结论.

证明 证法一: 由 $AC = CA$ 可知, 对任意的正整数 k,

$$C^k = C^{k-1}AB - C^{k-1}BA = A(C^{k-1}B) - (C^{k-1}B)A.$$

由迹的线性和交换性可得 $\mathrm{tr}(C^k) = 0\ (k \geq 1)$, 再由幂零矩阵关于迹的充要条件可知 C 为幂零矩阵, 从而 C 的特征值全为零。

证法二:将 A, B, C 看成是 n 维复列向量空间 V 上的线性变换. 任取 C 的特征值 λ_0 及其特征子空间 V_0, 由 $AC = CA, BC = CB$ 以及命题 1.16可知, V_0 是 $A-$ 不变子空间, 也是 $B-$ 不变子空间. 将等式 $C = AB - BA$ 两边的线性变换同时限制在 V_0 上, 可得 V_0 上线性变换的等式 $C|_{V_0} = A|_{V_0}B|_{V_0} - B|_{V_0}A|_{V_0}$. 两边同时取迹, 由迹的线性和交换性及命题 1.13可知

$$\lambda_0 \dim V_0 = \mathrm{tr}(C|_{V_0}) = \mathrm{tr}(A|_{V_0}B|_{V_0}) - \mathrm{tr}(B|_{V_0}A|_{V_0}) = 0,$$

从而 $\lambda_0 = 0$, 结论得证. □

注 上述证法二中,A, B, C 在不变子空间 V_0 上的限制只能理解成线性变换在不变子空间上的限制, 而不是矩阵在不变子空间上的限制.

推论 1.3

设 A, B, C 是 n 阶矩阵, 其中 $C = AB - BA$. 若它们满足条件 $AC = CA, BC = CB$, 求证: A, B, C 可同时上三角化.

证明 对阶数进行归纳. 由命题 1.14 证法二可知, C 的特征值全为 0, 其特征子空间 V_0 满足

$$A|_{V_0}B|_{V_0} - B|_{V_0}A|_{V_0} = C|_{V_0} = 0,$$

即 $A|_{V_0}, B|_{V_0}$ 乘法可交换. 由命题 1.19可知 $A|_{V_0}, B|_{V_0}$ 有公共的特征向量, 即存在 $0 \neq e_1 \in V_0$, 使得

$$Ae_1 = A|_{V_0}(e_1) = \lambda_1 e_1, \ Be_1 = B|_{V_0}(e_1) = \mu_1 e_1, \ Ce_1 = 0.$$

余下的证明完全类似于命题 1.20的证明, 请读者自行补充相关的细节. □

1.1.6 特征值的估计

定理 1.4 (第一圆盘定理)

设 $A = (a_{ij})$ 是 n 阶矩阵, 则 A 的特征值在复平面的下列圆盘中:

$$|z - a_{ii}| \leq R_i, \ 1 \leq i \leq n,$$

其中 $R_i = |a_{i1}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{in}|$.

注 该定理又称为 Gerschgorin 圆盘第一定理, 即戈氏圆盘第一定理. 上述圆盘称为戈氏圆盘.

证明

□

定理 1.5 (第二圆盘定理)

若 n 阶矩阵 A 的 n 个戈氏圆盘分成若干个连通区域, 其中某个连通区域恰含 k 个戈氏圆盘, 则有且仅有 k 个特征值落在该连通区域内 (若两个圆盘重合应计算重数, 若特征值为重根也要计算重数).

证明

□

例题 1.19 如果圆盘定理中有一个连通分支由两个圆盘外切组成, 证明: 每个圆盘除去切点的区域不可能同时包含两个特征值.

证明 设 $A = (a_{ij})$ 为 n 阶矩阵, $D_i : |z - a_{ii}| \le R_i (1 \le i \le n)$ 是 A 的 n 个戈氏圆盘. 不妨设 A 的两个戈氏圆盘 D_1, D_2 外切并组成一个连通分支. 令

$$A(t) = \begin{pmatrix} a_{11} & ta_{12} & \cdots & ta_{1n} \\ ta_{21} & a_{22} & \cdots & ta_{2n} \\ \vdots & \vdots & & \vdots \\ ta_{n1} & ta_{n2} & \cdots & a_{nn} \end{pmatrix},$$

由第一圆盘定理, $A(t)$ 的特征值落在下列圆盘中:

$$tD_i : |z - a_{ii}| \le tR_i,\ 1 \le i \le n.$$

由于当 $0 \le t < 1$ 时, $A(t)$ 的特征值是关于 t 的连续函数, 故 $A(t)$ 的特征值 $\lambda_i(t)$ 从 D_i 的圆心开始, 始终在圆盘 $tD_i (1 \le i \le n)$ 中连续变动. 注意此时 tD_1, tD_2 不相交, 它们是两个连通分支, 于是特征值 $\lambda_i(t)$ 落在 $tD_i (i = 1, 2)$ 中. 最后当 $t = 1$ 时, A 的特征值 $\lambda_1 = \lambda_1(1)$ 落在 D_1 中, 特征值 $\lambda_2 = \lambda_2(1)$ 落在 D_2 中. 因此, λ_1, λ_2 不可能同时落在 D_1 或 D_2 除去切点的区域中. □

例题 1.20 设 $A = (a_{ij})$ 为 n 阶矩阵, 证明: 存在正数 δ, 使得对任意的 $s \in (0, \delta)$, 下列矩阵均有 n 个不同的特征值:

$$A(s) = \begin{pmatrix} a_{11} + s & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} + s^2 & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} + s^n \end{pmatrix}.$$

证明 先证当 s 充分大时, $A(s)$ 有 n 个不同的特征值. 由第一圆盘定理, $A(s)$ 的特征值落在下列戈氏圆盘中:

$$D_i : |z - a_{ii} - s^i| \le R_i = \sum_{j=1, j \ne i}^{n} |a_{ij}|,\ 1 \le i \le n.$$

取 s 充分大, 使得 $s^n \gg s^{n-1} \gg \cdots \gg s$. 注意到 R_i 的值固定, 故 D_i 的圆心之间的距离大于半径 R_i, 从而 D_i 互不相交, 各自构成了一个连通分支. 再由第二圆盘定理, 每个连通分支 D_i 中有且仅有一个特征值, 于是 $A(s)$ 有 n 个不同的特征值.

设 $f_s(\lambda) = |\lambda I_n - A(s)|$ 是 $A(s)$ 的特征多项式, 则其判别式 $\Delta(f_s(\lambda))$ 是关于 s 的多项式. 由前面的讨论可知, 当 s 充分大时, $f_s(\lambda)$ 无重根, 从而 $\Delta(f_s(\lambda)) \ne 0$, 即 $\Delta(f_s(\lambda))$ 是关于 s 的非零多项式. 若 $\Delta(f_s(\lambda))$ 的所有复根都是零, 则任取一个正数 δ; 若 $\Delta(f_s(\lambda))$ 的复根不全为零, 则可取 δ 为 $\Delta(f_s(\lambda))$ 的非零复根的模长的最小值. 于是对任意的 $s \in (0, \delta)$, s 都不是 $\Delta(f_s(\lambda))$ 的根, 即 $\Delta(f_s(\lambda)) \ne 0$, 从而 $f_s(\lambda)$ 都无重根, 即 $A(s)$ 都有 n 个不同的特征值. □

1.2 由乘法交换性诱导的同时性质

命题 1.15 (矩阵乘法可交换的基本性质)

若两个矩阵或线性变换 A, B 乘法可交换, 即 $AB = BA$, 则有 $(AB)^m = A^m B^m$, $f(A)g(B) = g(B)f(A)$ 以及二项式定理

$$(A+B)^m = A^m + \mathrm{C}_m^1 A^{m-1}B + \cdots + \mathrm{C}_m^{m-1}AB^{m-1} + B^m$$

等成立, 其中 $m \geq 1$, $f(x), g(x)$ 为多项式.

特别地, 一个矩阵或线性变换 A 一定与其自身可交换, 从而也满足 $f(A)g(A) = g(A)f(A)$, 其中 $f(x), g(x)$ 为多项式.

证明 证明是显然的. □

1.2.1 特征子空间互为不变子空间

命题 1.16 (特征子空间互为不变子空间)

1. 设 φ, ψ 是复线性空间 V 上乘法可交换的线性变换, 即 $\varphi\psi = \psi\varphi$, 求证: φ 的特征子空间是 ψ 的不变子空间, ψ 的特征子空间是 φ 的不变子空间.
2. 若 n 阶复矩阵 A, B 乘法可交换, 即 $AB = BA$, 则 A, B 的特征子空间互为不变子空间.

注 这个命题的结论对一般的数域是不成立的. 例如, $A = I_2$, $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 显然 A, B 乘法可交换, 但它们在有理数域或实数域上没有公共的特征向量. 事实上, B 在有理数域或实数域上都没有特征值 (它的特征值是 $\pm\mathrm{i}$), 从而也没有特征向量, 所以更谈不上公共的特征向量了. 为了这个命题的结论推广到数域 $\mathbb{F}$ 上, 我们必须假设 A, B 的特征值都在 $\mathbb{F}$ 中.

证明

1. 由代数基本定理以及线性方程组的求解理论可知, $n(n \geq 1)$ 维复线性空间上的线性变换或 n 阶复矩阵至少有一个特征值和特征向量. 任取线性变换 φ 的一个特征值 λ_0, 设 V_0 是特征值 λ_0 的特征子空间, 则对任意的 $\alpha \in V_0$, 有
$$\varphi\psi(\alpha) = \psi\varphi(\alpha) = \psi(\lambda_0\alpha) = \lambda_0\psi(\alpha),$$
即 $\psi(\alpha) \in V_0$, 因此 V_0 是 ψ 的不变子空间. 同理可证 ψ 的特征子空间是 φ 的不变子空间.
2.

□

命题 1.17

设 V 为 n 维复线性空间, S 是 $\mathcal{L}(V)$ 的非空子集, 满足: S 中的全体线性变换没有非平凡的公共不变子空间. 设线性变换 φ 与 S 中任一线性变换乘法均可交换, 证明: φ 是纯量变换.

证明 任取 φ 的特征值 λ_0 及其特征子空间 V_0. 任取 $\psi \in S$, 则 $\varphi\psi = \psi\varphi$, 由命题 1.16可知 V_0 是 $\psi-$ 不变子空间, 从而是 S 中全体线性变换的公共不变子空间. 又 $V_0 \neq 0$(特征向量均非零), 故 $V_0 = V$, 从而 $\varphi = \lambda_0 I_V$ 为纯量变换. □

1.2.2 有公共的特征向量

命题 1.18

1. 设 φ, ψ 是复线性空间 V 上乘法可交换的线性变换, 求证: φ, ψ 至少有一个公共的 (复) 特征向量.
2. 若 n 阶复矩阵 A, B 乘法可交换, 即 $AB = BA$, 则 A, B 至少有一个公共的 (复) 特征向量.

注 这个命题的结论对一般的数域是不成立的. 例如, $A = I_2, B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, 显然 A, B 乘法可交换, 但它们在有理数域或实数域上没有公共的特征向量. 事实上, B 在有理数域或实数域上都没有特征值 (它的特征值是 $\pm \mathrm{i}$), 从而也没有特征向量, 所以更谈不上公共的特征向量了. 为了这个命题的结论推广到数域 $\mathbb{F}$ 上, 我们必须假设 A, B 的特征值都在 $\mathbb{F}$ 中.

证明

1. 任取 φ 的特征值 λ_0 及其特征子空间 V_0, 由命题 1.16可知, V_0 是 $\psi-$ 不变子空间. 将线性变换 ψ 限制在 V_0 上, 由于 V_0 是维数大于零的复线性空间, 故由命题 1.1可知 $\psi|_{V_0}$ 至少有一个特征值 μ_0 及其特征向量 $\alpha \in V_0$, 从而 $\varphi(\alpha) = \lambda_0 \alpha, \psi(\alpha) = \mu_0 \alpha$, 于是 α 就是 φ, ψ 的公共特征向量.
2.

□

命题 1.19

1. 设 φ, ψ 是数域 $\mathbb{F}$ 上线性空间 V 上的乘法可交换的线性变换, 且 φ, ψ 的特征值都在 $\mathbb{F}$ 中, 求证: φ, ψ 的特征子空间互为不变子空间, 并且 φ, ψ 至少有一个公共的特征向量.
2. 若数域 $\mathbb{F}$ 上的 n 阶矩阵 A, B 乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 则 A, B 的特征子空间互为不变子空间, 并且 A, B 在 $\mathbb{F}^n$ 中至少有一个公共的特征向量.

证明

1. 由线性方程组的求解理论可知, 若数域 $\mathbb{F}$ 上的线性变换或 $\mathbb{F}$ 上的矩阵在 $\mathbb{F}$ 中有一个特征值, 则在 $\mathbb{F}$ 上的线性空间或 $\mathbb{F}$ 上的列向量空间中必存在对应的特征向量. 任取线性变换 φ 的一个特征值 $\lambda_0 \in \mathbb{F}$, 设 V_0 是特征值 λ_0 的特征子空间, 则对任意的 $\alpha \in V_0$, 有
$$\varphi\psi(\alpha) = \psi\varphi(\alpha) = \psi(\lambda_0 \alpha) = \lambda_0 \psi(\alpha),$$
即 $\psi(\alpha) \in V_0$, 因此 V_0 是 $\psi-$ 不变子空间. 取 V_0 的一组基并扩张为 V 的一组基, 则 ψ 在这组基下的表示矩阵为分块对角矩阵 $\begin{pmatrix} A & C \\ O & B \end{pmatrix}$, 其中 A 是 $\psi|_{V_0}$ 在给定基下的表示矩阵, 于是 $|\lambda I_V - \psi| = |\lambda I - A||\lambda I - B|$. 因为 ψ 的特征值都在 $\mathbb{F}$ 中, 故 A 的特征值都在 $\mathbb{F}$ 中, 于是 $\psi|_{V_0}$ 的特征值都在 $\mathbb{F}$ 中. 任取 $\psi|_{V_0}$ 的一个特征值 $\mu_0 \in \mathbb{F}$ 及其特征向量 $\alpha \in V_0$, 则 $\varphi(\alpha) = \lambda_0 \alpha, \psi(\alpha) = \mu_0 \alpha$, 于是 α 就是 φ, ψ 的公共特征向量.
2.

□

1.2.3 可同时上三角化

命题 1.20

1. 设数域 $\mathbb{F}$ 上的 n 阶矩阵 A 的特征值都在 $\mathbb{F}$ 中, 求证: A 在 $\mathbb{F}$ 上可上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 是上三角矩阵.
2. 设数域 $\mathbb{F}$ 上线性空间 V 上的线性变换 φ 的特征值都在 $\mathbb{F}$ 中, 则存在 V 的一组基, 使得 φ 在这组基下的表示矩阵是上三角矩阵.

证明

1. 对阶数进行归纳. 当 $n=1$ 时结论显然成立, 设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵 A 进行证明. 设 $\lambda_1\in\mathbb{F}$ 是 A 的一个特征值, 则由线性方程组的求解理论可知, 存在特征向量 $e_1\in\mathbb{F}^n$, 使得 $Ae_1=\lambda_1e_1$. 由基扩张定理, 可将 e_1 扩张为 $\mathbb{F}^n$ 的一组基 $\{e_1,e_2,\cdots,e_n\}$, 于是
$$(Ae_1,Ae_2,\cdots,Ae_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix},$$
其中 A_1 是 $\mathbb{F}$ 上的 $n-1$ 阶矩阵. 令 $P=(e_1,e_2,\cdots,e_n)$, 则 P 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且由上式可得 $AP=P\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}$, 即 $P^{-1}AP=\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}$. 由此可得 $|\lambda I_n-A|=(\lambda-\lambda_1)|\lambda I_{n-1}-A_1|$, 又 A 的特征值全在 $\mathbb{F}$ 中, 从而 A_1 的特征值也全在 $\mathbb{F}$ 中, 故由归纳假设, 存在 $\mathbb{F}$ 上的 $n-1$ 阶可逆矩阵 Q, 使得 $Q^{-1}A_1Q$ 是上三角矩阵. 令
$$R=P\begin{pmatrix}1 & O\\ O & Q\end{pmatrix},$$
则 R 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且
$$R^{-1}AR=\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}^{-1}\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\begin{pmatrix}1 & O\\ O & Q\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & Q^{-1}A_1Q\end{pmatrix}$$
是上三角矩阵.
2.

□

命题 1.21

1. 设 A,B 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 满足: $AB=BA$ 且 A,B 的特征值都在 $\mathbb{F}$ 中, 求证: A,B 在 $\mathbb{F}$ 上可同时上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵.
2. 设数域 $\mathbb{F}$ 上线性空间 V 上的线性变换 φ,ψ 乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 则存在 V 的一组基, 使得 φ,ψ 在这组基下的表示矩阵都是上三角矩阵.

证明

1. 对阶数进行归纳. 当 $n=1$ 时结论显然成立, 设对 $n-1$ 阶矩阵结论成立, 现对 n 阶矩阵进行证明. 因为 $AB=BA$ 且 A,B 的特征值都在 $\mathbb{F}$ 中, 故由命题 1.19可知, A,B 有公共的特征向量 $e_1\in\mathbb{F}^n$, 不妨设
$$Ae_1=\lambda_1e_1,\ Be_1=\mu_1e_1,$$
其中 $\lambda_1,\mu_1\in\mathbb{F}$ 分别是 A,B 的特征值. 由基扩张定理, 可将 e_1 扩张为 $\mathbb{F}^n$ 的一组基 $\{e_1,e_2,\cdots,e_n\}$. 令 $P=(e_1,e_2,\cdots,e_n)$, 则 P 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 从而有
$$\begin{aligned}A(e_1,e_2,\cdots,e_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\Leftrightarrow AP=P\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\Leftrightarrow P^{-1}AP=\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix},\\ B(e_1,e_2,\cdots,e_n)=(e_1,e_2,\cdots,e_n)\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}\Leftrightarrow BP=P\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}\Leftrightarrow P^{-1}BP=\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}.\end{aligned}\tag{1.5}$$
其中 A_1,B_1 是 $\mathbb{F}$ 上的 $n-1$ 阶矩阵. 由 $AB=BA$ 及(1.5)式可得到
$$\begin{aligned}&\left(P^{-1}AP\right)\left(P^{-1}BP\right)=P^{-1}ABP=P^{-1}BAP=\left(P^{-1}BP\right)\left(P^{-1}AP\right)\\ \Leftrightarrow&\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & B_1\end{pmatrix}\begin{pmatrix}\lambda_1 & *\\ O & A_1\end{pmatrix}\\ \Leftrightarrow&\begin{pmatrix}\lambda_1 & *\\ O & A_1B_1\end{pmatrix}=\begin{pmatrix}\lambda_1 & *\\ O & B_1A_1\end{pmatrix}\end{aligned}$$
从而 $A_1B_1=B_1A_1$. 又由(1.5)式可得
$$|\lambda I_n-A|=|\lambda-\lambda_1|\,|\lambda I_{n-1}-A_1|,\quad |\lambda I_n-B|=|\lambda-\lambda_1|\,|\lambda I_{n-1}-B_1|.$$

因此 A_1, B_1 的特征值也是 A, B 的特征值. 又由于 A, B 的特征值都在 $\mathbb{F}$ 中, 故 A_1, B_1 的特征值都在 $\mathbb{F}$ 中. 故由归纳假设, 存在 $\mathbb{F}$ 上的 $n-1$ 阶可逆矩阵 Q, 使得 $Q^{-1}A_1Q$ 和 $Q^{-1}B_1Q$ 都是上三角矩阵. 令

$$R = P\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix},$$

则 R 是 $\mathbb{F}$ 上的 n 阶可逆矩阵, 且

$$R^{-1}AR = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}^{-1}\begin{pmatrix} \lambda_1 & * \\ O & A_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ O & Q^{-1}A_1Q \end{pmatrix},$$

$$R^{-1}BR = \begin{pmatrix} 1 & O \\ O & Q \end{pmatrix}^{-1}\begin{pmatrix} \mu_1 & * \\ O & B_1 \end{pmatrix}\begin{pmatrix} 1 & O \\ O & Q \end{pmatrix} = \begin{pmatrix} \mu_1 & * \\ O & Q^{-1}B_1Q \end{pmatrix}$$

都是上三角矩阵.

2.

□

1.2.4 可同时对角化

命题 1.22

1. 设 φ, ψ 是数域 $\mathbb{F}$ 上 n 维线性空间 V 上的线性变换, 满足: $\varphi\psi = \psi\varphi$ 且 φ, ψ 都可对角化, 求证: φ, ψ 可同时对角化, 即存在 V 的一组基, 使得 φ, ψ 在这组基下的表示矩阵都是对角矩阵.
2. 设 A, B 是数域 $\mathbb{F}$ 上的 n 阶矩阵, 满足: $AB = BA$ 且 A, B 都在 $\mathbb{F}$ 上可对角化, 则 A, B 在 $\mathbb{F}$ 上可同时对角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是对角矩阵.

证明

1. 对空间维数进行归纳. 当 $n = 1$ 时结论显然成立, 设对维数小于 n 的线性空间结论成立, 现对 n 维线性空间进行证明. 设 φ 的全体不同特征值为 $\lambda_1, \cdots, \lambda_s \in \mathbb{F}$, 对应的特征子空间分别为 $V_1, \cdots, V_s$, 则由 φ 可对角化可知

$$V = V_1 \oplus \cdots \oplus V_s.$$

若 $s = 1$, 则 $\varphi = \lambda_1 I_V$ 为纯量变换, 此时只要取 V 的一组基, 使得 ψ 在这组基下的表示矩阵为对角矩阵, 则 φ 在这组基下的表示矩阵为 $\lambda_1 I_n$, 结论成立. 若 $s > 1$, 则 $\dim V_i < n$. 注意到 $\varphi\psi = \psi\varphi$ 且 φ, ψ 的特征值都在 $\mathbb{F}$ 中, 由命题 1.16可知 V_i 都是 $\psi-$ 不变子空间. 考虑线性变换的限制 $\varphi|_{V_i}, \psi|_{V_i}$: 它们乘法可交换, 且由可对角化线性变换的性质可知它们都可对角化, 故由归纳假设可知, $\varphi|_{V_i}, \psi|_{V_i}$ 可同时对角化, 即存在 V_i 的一组基, 使得 $\varphi|_{V_i}, \psi|_{V_i}$ 在这组基下的表示矩阵都是对角矩阵. 将 V_i 的基拼成 V 的一组基, 则 φ, ψ 在这组基下的表示矩阵都是对角矩阵, 即 φ, ψ 可同时对角化.

2.

□

1.2.5 个数的推广

命题 1.23

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $A_1, A_2, \cdots, A_m$ 两两乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 求证: 它们在 $\mathbb{F}^n$ 中至少有一个公共的特征向量.

证明 对 m 进行归纳, $m = 2$ 时就是命题 1.18. 设矩阵个数小于 m 时结论成立, 现证 m 个矩阵的情形. 将所有的 A_i 都看成是列向量空间 $\mathbb{F}^n$ 上的线性变换, 任取 A_1 的一个特征值 $\lambda_1 \in \mathbb{F}$ 及其特征子空间 $V_1 \subseteq \mathbb{F}^n$. 注意到 $A_1A_i = A_iA_1$, 故由命题 1.16可知, V_1 是 $A_2, \cdots, A_m$ 的不变子空间. 将 $A_2, \cdots, A_m$ 限制在 V_1 上, 它们仍然两两乘

法可交换且特征值都在 $\mathbb{F}$ 中, 故由归纳假设可得 $A_2|_{V_1}, \cdots, A_m|_{V_1}$ 有公共的特征向量 $\alpha \in V_1$. 注意到 α 也是 A_1 的特征向量, 于是 α 是 $A_1, A_2, \cdots, A_m$ 的公共特征向量. □

命题 1.24

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $A_1, A_2, \cdots, A_m$ 两两乘法可交换, 且它们的特征值都在 $\mathbb{F}$ 中, 求证: 它们在 $\mathbb{F}$ 上可同时上三角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}A_iP(1 \le i \le m)$ 都是上三角矩阵.

证明 完全类似于命题 1.21的证明, 其中利用命题 1.23得到 $A_1, A_2, \cdots, A_m$ 的公共特征向量, 请读者自行补充相关的细节. □

命题 1.25

设数域 $\mathbb{F}$ 上的 n 阶矩阵 $A_1, A_2, \cdots, A_m$ 两两乘法可交换, 且它们都在 $\mathbb{F}$ 上可对角化, 求证: 它们在 $\mathbb{F}$ 上可同时对角化, 即存在 $\mathbb{F}$ 上的可逆矩阵 P, 使得 $P^{-1}A_iP(1 \le i \le m)$ 都是对角矩阵.

证明 若 A_i 都是纯量矩阵, 则结论显然成立. 以下不妨设 A_1 不是纯量矩阵, 余下的证明完全类似于命题 1.22的证明, 请读者自行补充相关的细节. □

例题 1.21 设 A, B 都是 n 阶矩阵且 $AB = BA$. 若 A 是幂零矩阵, 求证: $|A + B| = |B|$.

证明 证法一: 由命题 1.21可知, A, B 可同时上三角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP$ 和 $P^{-1}BP$ 都是上三角矩阵. 因为上三角矩阵的主对角元是矩阵的特征值, 而幂零矩阵的特征值全为零, 所以 $|P^{-1}AP + P^{-1}BP| = |P^{-1}BP|$, 即有 $|A + B| = |B|$.

证法二: 先假设 B 是可逆矩阵, 则 $|A + B| = |I_n + AB^{-1}||B|$, 只要证明 $|I_n + AB^{-1}| = 1$ 即可. 由 $AB = BA$ 可知 $AB^{-1} = B^{-1}A$, 再由 A 是幂零矩阵容易验证 AB^{-1} 也是幂零矩阵, 从而其特征值全为零. 因此 $I_n + AB^{-1}$ 的特征值全为 1, 故 $|I_n + AB^{-1}| = 1$.

对于一般的矩阵 B, 可取到一列有理数 $t_k \to 0$, 使得 t_kI_n+B 是可逆矩阵. 由可逆情形的证明可得 $|A+t_kI_n+B| = |t_kI_n + B|$. 注意到上式两边都是 t_k 的多项式, 从而关于 t_k 连续. 将上式两边同时取极限, 令 $t_k \to 0$, 即得结论. □

1.3 矩阵相似和可对角化的计算

1.3.1 相似初等变换及其应用

命题 1.26 (相似初等变换)

设 $\boldsymbol{A}$ 为 n 阶矩阵, 容易验证以下 3 种变换都是相似变换, 称为**相似初等变换**:

1. 对换 $\boldsymbol{A}$ 的第 i 行与第 j 行, 再对换第 i 列与第 j 列;
2. 将 $\boldsymbol{A}$ 的第 i 行乘以非零常数 c, 再将第 i 列乘以 c^{-1};
3. 将 $\boldsymbol{A}$ 的第 i 行乘以常数 c 加到第 j 行上, 再将第 j 列乘以 $-c$ 加到第 i 列上.

设 $\boldsymbol{A}$ 是具有相同行列分块方式的分块矩阵, 容易验证以下 3 种变换都是相似变换, 称为**相似分块初等变换**:

1. 对换 $\boldsymbol{A}$ 的第 i 分块行与第 j 分块行, 再对换第 i 分块列与第 j 分块列;
2. 将 $\boldsymbol{A}$ 的第 i 分块行左乘非异阵 $\boldsymbol{M}$, 再将第 i 分块列右乘 $\boldsymbol{M}^{-1}$;
3. 将 $\boldsymbol{A}$ 的第 i 分块行左乘矩阵 $\boldsymbol{M}$ 加到第 j 分块行上, 再将第 j 分块列右乘 $-\boldsymbol{M}$ 加到第 i 分块列上.

容易验证: **任一相似变换都是若干次相似初等变换的复合**.

例题 1.22 设 $\boldsymbol{A} = \mathrm{diag}\{\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m\}$ 是分块对角矩阵, 其中 $\boldsymbol{A}_i$ 都是方阵, 求证: $\mathrm{diag}\{\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m\}$ 相似于 $\mathrm{diag}\{\boldsymbol{A}_{i_1}, \boldsymbol{A}_{i_2}, \cdots, \boldsymbol{A}_{i_m}\}$, 其中 $\boldsymbol{A}_{i_1}, \boldsymbol{A}_{i_2}, \cdots, \boldsymbol{A}_{i_m}$ 是 $\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m$ 的一个排列.

证明 对换 $\boldsymbol{A}$ 的第 i 分块行与第 j 分块行, 再对换第 i 分块列与第 j 分块列. 这是一个相似变换, 变换的结果是将 $\boldsymbol{A}$ 的 (i,i) 分块和第 (j,j) 分块对换了位置. 又任一排列都可以通过若干次对换来实现, 因此 $\mathrm{diag}\{\boldsymbol{A}_1, \boldsymbol{A}_2, \cdots, \boldsymbol{A}_m\}$ 和 $\mathrm{diag}\{\boldsymbol{A}_{i_1}, \boldsymbol{A}_{i_2}, \cdots, \boldsymbol{A}_{i_m}\}$ 相似. □

命题 1.27

若分块矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 满足 $\mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathrm{r}(A)$, 则 $\mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\begin{pmatrix} A \\ C \end{pmatrix} = \mathrm{r}(A)$.

证明 由条件可得

$$\mathrm{r}(A) \leqslant \mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\begin{pmatrix} A \\ C \end{pmatrix} \leqslant \mathrm{r}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mathrm{r}(A).$$

故 $\mathrm{r}\begin{pmatrix} A & B \end{pmatrix} = \mathrm{r}\begin{pmatrix} A \\ C \end{pmatrix} = \mathrm{r}(A)$. □

例题 1.23 设 n 阶方阵 A, B 满足 $\mathrm{r}(ABA) = \mathrm{r}(B)$, 求证: AB 与 BA 相似.

注 不妨设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 的原因: 假设当 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 时, 结论已经成立. 则对于一般的满足条件的矩阵 A, 设 $\mathrm{r}(A) = r$, 则存在可逆矩阵 P, Q, 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$. 再根据条件可知

$$\mathrm{r}((PAQ)(Q^{-1}BP^{-1})(PAQ)) = \mathrm{r}(PABAQ) = \mathrm{r}(ABA) = \mathrm{r}(B) = \mathrm{r}(Q^{-1}BP^{-1}).$$

从而由假设可知,$(PAQ)(Q^{-1}BP^{-1})$ 与 $(Q^{-1}BP^{-1})(PAQ)$ 相似, 即 $PABP^{-1}$ 与 $Q^{-1}BAQ$ 相似. 又注意到 $PABP^{-1}$ 与 AB 相似,$Q^{-1}BAQ$ 与 BA 相似, 因此 AB 与 BA 也相似. 故不妨设成立.

证明 设 P, Q 为 n 阶非异阵, 使得 $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, 其中 $r = \mathrm{r}(A)$. 注意到问题的条件和结论在相抵变换: $A \mapsto PAQ, B \mapsto Q^{-1}BP^{-1}$ 下保持不变, 故不妨从一开始就假设 $A = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ 是相抵标准型. 设 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 为对应的分块, 则由 $\mathrm{r}(ABA) = \mathrm{r}(B)$ 可得 $\mathrm{r}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \mathrm{r}(B_{11})$. 从而由命题 1.27可得 $\mathrm{r}\begin{pmatrix} B_{11} & B_{12} \end{pmatrix} = \mathrm{r}(B_{11})$ 以及 $\mathrm{r}\begin{pmatrix} B_{11} \\ B_{21} \end{pmatrix} = \mathrm{r}(B_{11})$, 再由定理??可知存在矩阵 M, N, 使得 $B_{11}N = B_{12}, MB_{11} = B_{21}$.

将 $AB = \begin{pmatrix} B_{11} & B_{12} \\ O & O \end{pmatrix}$ 的第二分块行左乘 N 加到第一分块行, 再将第一分块列右乘 $-N$ 加到第二分块列, 于是 AB 相似于 $\begin{pmatrix} B_{11} & O \\ O & O \end{pmatrix}$. 将 $BA = \begin{pmatrix} B_{11} & O \\ B_{21} & O \end{pmatrix}$ 的第一分块行左乘 $-M$ 加到第二分块行, 再将第二分块列右乘 M 加到第一分块列, 于是 BA 相似于 $\begin{pmatrix} B_{11} & O \\ O & O \end{pmatrix}$. 因此, AB 与 BA 相似. □

1.3.2 利用相似不变量来判定矩阵不相似

定义 1.10 (相似不变量)

相似的矩阵具有相同的**迹**、**行列式**、**特征多项式**和**极小多项式**等, 故它们被称为矩阵相似关系下的不变量, 也称为**相似不变量**.

注 若两个矩阵的相似不变量不相同, 则它们必不相似. 利用这种方法来判断两个矩阵不相似是很简便的.

笔记 由矩阵迹、行列式、特征多项式的性质易知, 相似的矩阵具有相同的迹、行列式、特征多项式. 由命题 1.41可知, 相似的矩阵具有相同的极小多项式. 因此上述定义是良定义的.

例题 1.24 设 A, B 为 n 阶方阵, 求证: $AB - BA$ 必不相似于 kI_n, 其中 k 是非零常数.

证明 注意到 $\mathrm{tr}(AB - BA) = 0, \mathrm{tr}(kI_n) = nk \neq 0$, 又矩阵的迹是相似不变量, 因此 $AB - BA$ 和 kI_n 必不相似. □

例题 1.25 设 A, B 为 n 阶正交矩阵, 且线性方程组 $(A + B)x = 0$ 的解空间维数是奇数, 求证: A 和 B 必不相似.

证明 由假设可知 $n-\mathrm{r}(\boldsymbol{A}+\boldsymbol{B})$ 为奇数, 再由命题??可知 $|\boldsymbol{A}|=-|\boldsymbol{B}|\neq 0$, 又矩阵的行列式是相似不变量, 因此 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 必不相似. □

1.3.3 过渡矩阵 P 的计算

首先, 我们介绍一下当矩阵相似于对角矩阵时求过渡矩阵的方法. 设 n 阶矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 可逆矩阵 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)$ 为其列分块, 且

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},$$

则

$$\boldsymbol{A}\boldsymbol{P}=\boldsymbol{P}\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},$$

即

$$(\boldsymbol{A}\boldsymbol{\alpha}_1,\boldsymbol{A}\boldsymbol{\alpha}_2,\cdots,\boldsymbol{A}\boldsymbol{\alpha}_n)=(\lambda_1\boldsymbol{\alpha}_1,\lambda_2\boldsymbol{\alpha}_2,\cdots,\lambda_n\boldsymbol{\alpha}_n),$$

于是 $\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i$, 这表明 $\boldsymbol{\alpha}_i$ 就是属于特征值 λ_i 的特征向量. 因此 $\boldsymbol{P}$ 的 n 个列向量就是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量. 注意: 因为特征向量不唯一, 所以过渡矩阵 $\boldsymbol{P}$ 也不唯一. 另外, $\boldsymbol{P}$ 的第 i 个列向量对应于 $\boldsymbol{A}$ 的第 i 个特征值.

例题 1.26 设三阶矩阵 $\boldsymbol{A}$ 的特征值为 $1,1,4$, 对应的特征向量依次为 $(2,1,0)'$, $(-1,0,1)'$, $(0,1,1)'$, 试求矩阵 $\boldsymbol{A}$.

解 容易验证 $\boldsymbol{A}$ 的这 3 个特征向量线性无关, 故 $\boldsymbol{A}$ 必相似于对角矩阵, 即有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix}.$$

根据上面的分析, 有

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

于是

$$\boldsymbol{A}=\begin{pmatrix}1&0&0\\-3&7&-3\\-3&6&-2\end{pmatrix}.$$

□

1.3.4 可对角化判定的计算

例题 1.27 已知矩阵 $\boldsymbol{A}=\begin{pmatrix}1&-1&1\\2&x&-2\\-3&-3&y\end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix}2&0&0\\0&2&0\\0&0&z\end{pmatrix}$ 相似.

(1) 求 x,y,z 的值;

(2) 求一个满足 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$ 的可逆矩阵 $\boldsymbol{P}$.

解 (1) 显然 $z\neq 2$, 否则由 $\boldsymbol{A}$ 相似于 $2\boldsymbol{I}_3$ 可知 $\boldsymbol{A}=2\boldsymbol{I}_3$, 矛盾. 于是 $\boldsymbol{A}$ 的特征值为 2 (2 重), z (1 重). 因为 $\boldsymbol{A}$ 可对角化, 所以特征值 2 的几何重数也等于 2, 即线性方程组 $(A-2I_3)x=O$ 的解空间维数也是 2, 也即 $3-\mathrm{r}(\boldsymbol{A}-2\boldsymbol{I}_3)=2$. 故有

$$\mathrm{r}(\boldsymbol{A}-2\boldsymbol{I}_3)=\mathrm{r}\begin{pmatrix}-1&-1&1\\2&x-2&-2\\-3&-3&y-2\end{pmatrix}=1,$$

由此可得 $x=4,y=5$. 再由矩阵的迹等于特征值之和可得 $10=\mathrm{tr}(\boldsymbol{A})=4+z$, 故 $z=6$.

(2) 通过求解线性方程组 $(\lambda I_3 - A)x = O(\lambda = 2,6)$ 计算可得: 特征值 2 的两个线性无关的特征向量为 $\boldsymbol{\alpha}_1 = (-1,1,0)', \boldsymbol{\alpha}_2 = (1,0,1)'$; 特征值 6 的特征向量为 $\boldsymbol{\alpha}_3 = (1,-2,3)'$. 因此

$$\boldsymbol{P} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

□

例题 1.28 设 $\boldsymbol{A} = \begin{pmatrix} 3 & 2 & -2 \\ -k & -1 & k \\ 4 & 2 & -3 \end{pmatrix}$, 当 k 为何值时, 存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 是对角矩阵? 求出 $\boldsymbol{P}$ 和对角矩阵.

解 经计算可得 $|\lambda \boldsymbol{I}_3 - \boldsymbol{A}| = (\lambda-1)(\lambda+1)^2$, 因此 $\boldsymbol{A}$ 的特征值为 1 (1 重), −1 (2 重). 对单特征值 1, 其几何重数与代数重数必相等 (因为几何重数总小于代数重数且一个特征值必存在一个对应的特征向量); 因此要使 $\boldsymbol{A}$ 可对角化, 特征值 −1 的几何重数必须等于 2 才行, 故有

$$\mathrm{r}(\boldsymbol{A} + \boldsymbol{I}_3) = \mathrm{r}\begin{pmatrix} 4 & 2 & -2 \\ -k & 0 & k \\ 4 & 2 & -2 \end{pmatrix} = 1,$$

于是 $k = 0$. 通过计算可得: 特征值 1 的特征向量为 $(1,0,1)'$; 特征值 −1 的两个线性无关的特征向量为 $(-1,2,0)', (1,0,2)'$. 因此

$$\boldsymbol{P} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

□

1.3.5 可对角化矩阵的应用

例题 1.29 设矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 4 & -2 \\ -3 & -3 & 5 \end{pmatrix}$, 求 $\boldsymbol{A}^n$.

笔记 本题中的矩阵是一个可对角化矩阵, 因此可以使用下列方法: 先求出可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{B}$ 是对角矩阵. 因为对角矩阵的幂很容易求出, 故由 $\boldsymbol{A}^n = \boldsymbol{PB}^n\boldsymbol{P}^{-1}$ 即可得到结果.

解 经计算可得 $|\lambda \boldsymbol{I}_3 - \boldsymbol{A}| = (\lambda-2)^2(\lambda-6)$, 因此 $\boldsymbol{A}$ 的特征值为 2 (2 重), 6 (1 重). 通过计算可得: 特征值 2 有两个线性无关的特征向量 $(-1,1,0)', (1,0,1)'$; 特征值 6 有特征向量 $(1,-2,3)'$, 于是 $\boldsymbol{A}$ 有完全的特征向量系, 从而可对角化. 注意到

$$\boldsymbol{P} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}, \boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{AP} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix},$$

故

$$\boldsymbol{A}^n = \boldsymbol{PB}^n\boldsymbol{P}^{-1} = \frac{1}{4}\begin{pmatrix} 5\cdot 2^n - 6^n & 2^n - 6^n & 6^n - 2^n \\ 2\cdot 6^n - 2^{n+1} & 2\cdot 6^n + 2^{n+1} & 2^{n+1} - 2\cdot 6^n \\ 3\cdot 2^n - 3\cdot 6^n & 3\cdot 2^n - 3\cdot 6^n & 2^n + 3\cdot 6^n \end{pmatrix}.$$

□

例题 1.30 下列数列称为 Fibonacci 数列: $a_0 = 0, a_1 = 1, a_2 = 1, a_3 = 2, a_4 = 3, a_5 = 5, a_6 = 8, \cdots$, 通用递推式来表示为 $a_{n+2} = a_{n+1} + a_n$, 试求 Fibonacci 数列通项的显式表达式.

注 本例为用递推式定义的数列的通项提供了一个一般性的方法. 在三对角行列式中, 我们用技巧性相当高的方法

求出了用下列递推式定义的数列的通项:

$$D_n = aD_{n-1} - bcD_{n-2}, D_1 = a, D_2 = a^2 - bc.$$

现在请读者自己用上面的方法来求出通项表达式.

证明 首先用矩阵来表示递推式:

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix}.$$

令 $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, 则

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \boldsymbol{A}^2 \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} = \cdots = \boldsymbol{A}^n \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = \boldsymbol{A}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

只要求出 $\boldsymbol{A}^n$ 就可以算出 a_n 来. $\boldsymbol{A}$ 的特征多项式为 $|\lambda \boldsymbol{I}_2 - \boldsymbol{A}| = \lambda^2 - \lambda - 1$, 解得

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

对应于特征值 λ_1, λ_2 的特征向量分别是:

$$\boldsymbol{\alpha}_1 = \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix}, \boldsymbol{\alpha}_2 = \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}.$$

若记

$$\boldsymbol{P} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix},$$

则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}.$$

因此

$$\boldsymbol{A} = \boldsymbol{P} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \boldsymbol{P}^{-1}, \boldsymbol{A}^n = \boldsymbol{P} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}^n \boldsymbol{P}^{-1}.$$

经计算可得

$$\boldsymbol{A}^n = \frac{1}{\sqrt{5}} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} \\ \left(\frac{1+\sqrt{5}}{2}\right)^{n} - \left(\frac{1-\sqrt{5}}{2}\right)^{n} & \left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1} \end{pmatrix},$$

由此可得

$$a_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \right).$$

□

1.4 可对角化的判定

1.4.1 可对角化的基本知识

定义 1.11 (可对角化线性变换)

若 n 维线性空间 V 上的线性变换 φ 在某组基 $\{e_1,e_2,\cdots,e_n\}$ 下的表示矩阵为对角阵:

$$\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

则称 φ 为**可对角化线性变换**.

定理 1.6 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换，则 φ 可对角化的充分必要条件是 φ 有 n 个线性无关的特征向量。

证明 若 φ 是 V 上可对角化线性变换, 则可设 φ 在某组基 $\{e_1,e_2,\cdots,e_n\}$ 下的表示矩阵为对角阵:

$$\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

此时 $\varphi(e_i)=\lambda_i e_i$，即 $e_1,e_2,\cdots,e_n$ 是 φ 的特征向量，于是 φ 有 n 个线性无关的特征向量。

反过来，若 n 维线性空间 V 上的线性变换 φ 有 n 个线性无关的特征向量 $e_1,e_2,\cdots,e_n$，则这组向量构成了 V 的一组基，且 φ 在这组基下的表示矩阵显然是一个对角阵。 □

定义 1.12 (可对角化矩阵)

设 A 是 n 阶矩阵, 若 A 相似于对角阵, 即存在可逆阵 P, 使 $P^{-1}AP$ 为对角阵, 则称 A 为**可对角化矩阵**.

引理 1.1

设 A 是 n 阶矩阵,φ 是线性空间 V 上由矩阵 A 乘法诱导的线性变换，即 $\varphi(\alpha)=A\alpha,\forall\alpha\in V$。设 $\{e_1,e_2,\cdots,e_n\}$ 是 V 的一组标准基，则 φ 在这组基下的矩阵就是 A。证明:

(1) 矩阵 A 与线性变换 φ 的特征值相同;

(2) 矩阵 A 可对角化等价于线性变换 φ 可对角化。

证明

(1) 若 λ 为矩阵 A 的特征值，则存在 $\xi\in V$，使得 $\varphi(\xi)=A\xi=\lambda\xi$，因此矩阵 A 的特征值也是线性变换 φ 的特征值。

若 λ 为线性变换 φ 的特征值，则存在 $\eta\in V$，使得 $\varphi(\eta)=A\eta=\lambda\eta$，因此线性变换 φ 的特征值也是矩阵 A 的特征值。

故矩阵 A 与线性变换 φ 的特征值相同。

(2) 若矩阵 A 可对角化，则存在可逆矩阵 P，使得 $P^{-1}AP$ 为对角矩阵。

从而 $(e_1,e_2,\cdots,e_n)P$ 的列向量也是 V 的一组基，于是由命题**??** 可知 φ 在这组基下的矩阵为 $P^{-1}AP$ 是对角矩阵，故 φ 也可对角化。

若线性变换 φ 可对角化，则存在 V 的一组基 $\{f_1,f_2,\cdots,f_n\}$，使得 φ 在这组基下的矩阵 B 为对角矩阵。设基 $\{e_1,e_2,\cdots,e_n\}$ 到基 $\{f_1,f_2,\cdots,f_n\}$ 的过渡矩阵为 G，则由命题**??**可知 $B=G^{-1}AG$。因此矩阵 A 也可

对角化。

故矩阵 A 可对角化等价于线性变换 φ 可对角化。

$\square$

定理 1.7 (矩阵可对角化的充要条件)

设 A 是 n 阶矩阵, 则 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量.

证明 设 φ 是线性空间 V 上由矩阵 A 乘法诱导的线性变换.

若矩阵 A 有 n 个线性无关的特征值，则由引理 1.1(1)可知线性变换 φ 也有相同的 n 个线性无关的特征值，于是由定理 1.6可知线性变换 φ 可对角化，从而再由引理 1.1(2) 可知矩阵 A 也可对角化.

若矩阵 A 可对角化，则由引理 1.1(2)可知线性变换 φ 也可对角化，从而由定理 1.6可知 φ 有 n 个线性无关的特征值，于是由引理 1.1(1)可知矩阵 A 也有相同的 n 个线性无关的特征值。 $\square$

定理 1.8

若 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 为 n 维线性空间 V 上的线性变换 φ 的不同的特征值, 记 λ_i 的特征子空间为 $V_i(1\le i\le k)$, 则

$$V_1+V_2+\cdots+V_k=V_1\oplus V_2\oplus\cdots\oplus V_k.$$

证明 对 k 用数学归纳法. 若 $k=1$, 结论显然成立. 现设对 $k-1$ 个不同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_{k-1}$, 它们相应的特征子空间 $V_1,V_2,\cdots,V_{k-1}$ 之和是直和. 我们要证明 $V_1,V_2,\cdots,V_{k-1},V_k$ 之和为直和, 这只需证明:

$$V_k\cap(V_1+V_2+\cdots+V_{k-1})=0. \tag{1.6}$$

即可. 设 $v\in V_k\cap(V_1+V_2+\cdots+V_{k-1})$, 则

$$v=v_1+v_2+\cdots+v_{k-1}, \tag{1.7}$$

其中 $v_i\in V_i(i=1,2,\cdots,k-1)$. 在(1.7)式两边作用 φ, 得

$$\varphi(v)=\varphi(v_1)+\varphi(v_2)+\cdots+\varphi(v_{k-1}). \tag{1.8}$$

但 $v,v_1,v_2,\cdots,v_{k-1}$ 都是 φ 的特征向量或零向量, 因此

$$\lambda_k v=\lambda_1v_1+\lambda_2v_2+\cdots+\lambda_{k-1}v_{k-1}. \tag{1.9}$$

在(1.8)式两边乘以 λ_k 减去(1.9)式得

$$0=(\lambda_k-\lambda_1)v_1+(\lambda_k-\lambda_2)v_2+\cdots+(\lambda_k-\lambda_{k-1})v_{k-1}.$$

由于 $v_1,v_2,\cdots,v_{k-1}$ 是直和, 因此 $(\lambda_k-\lambda_i)v_i=0$, 而 $\lambda_k-\lambda_i\ne 0$, 从而 $v_i=0(i=1,2,\cdots,k-1)$. 这就证明了(1.6)式.
$\square$

推论 1.4

线性变换 φ 属于不同特征值的特征向量必线性无关.

证明 设 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是线性变换 φ 的 k 个不同特征值，由定理 1.8可知 $V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}$。于是任取 $\alpha_i\in V_{\lambda_i}\ (1\leqslant i\leqslant k)$ 且 $\alpha_i\ne 0$，假设 $\alpha_1,\alpha_2,\cdots,\alpha_k$ 线性无关，则存在一组不全为零的数 $b_1,b_2,\cdots,b_k$，使得

$$b_1\alpha_1+b_2\alpha_2+\cdots+b_k\alpha_k=0.$$

不妨设 $b_1\ne 0$，则

$$\alpha_1=\frac{b_2}{b_1}\alpha_2+\frac{b_3}{b_1}\alpha_3+\cdots+\frac{b_k}{b_1}\alpha_k\in V_{\lambda_1}\cap\left(V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}\right).$$

又由 $V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}$ 及直和的等价条件可知，

$$V_{\lambda_1}\cap\left(V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_k}\right)=\{0\},$$

从而 $\alpha_1=0$，这与 $\alpha_i\ne 0\ (1\leqslant i\leqslant k)$ 矛盾！ $\square$

推论 1.5

若 n 维线性空间 V 上的线性变换 φ 有 n 个不同的特征值, 则 φ 必可对角化.

笔记 注意这个推论只是可对角化的充分条件而非必要条件, 比如说纯量变换 $\varphi = cI_V$ 当然可对角化, 但 φ 的 n 个特征值都是 c.

证明 设 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是线性变换 φ 的 n 个不同特征值，则任取 $\alpha_i \in V_{\lambda_i}$ $(1 \leqslant i \leqslant n)$，由推论 1.4可知 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 一定线性无关。从而由定理 1.6可知，φ 一定可对角化。 □

定理 1.9 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是 φ 的全部不同的特征值, $V_i (i = 1, 2, \cdots, k)$ 是特征值 λ_i 的特征子空间, 则 φ 可对角化的充要条件是

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

证明 先证充分性. 设

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k,$$

分别取 V_i 的一组基 $\{e_{i1}, e_{i2}, \cdots, e_{it_i}\}(i = 1, 2, \cdots, k)$, 则由直和的等价条件 (4) 知这些向量拼成了 V 的一组基, 并且它们都是 φ 的特征向量. 因此 φ 有 n 个线性无关的特征向量, 从而定理 1.6可知 φ 可对角化.

再证必要性. 设 φ 可对角化, 则由定理 1.6可知 φ 有 n 个线性无关的特征向量 $\{e_1, e_2, \cdots, e_n\}$, 它们构成了 V 的一组基. 不失一般性, 可设这组基中前 t_1 个是关于特征值 λ_1 的特征向量; 接下去的 t_2 个是关于特征值 λ_2 的特征向量; $\cdots$; 最后 t_k 个是关于特征值 λ_k 的特征向量. 对任一 $\alpha \in V$, 设 $\alpha = a_1e_1 + a_2e_2 + \cdots + a_ne_n$, 则 α 可写成 $V_1, V_2, \cdots, V_k$ 中向量之和, 因此由定理 1.8可知

$$V = V_1 + V_2 + \cdots + V_k = V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

□

定义 1.13 (线性变换的几何重数与代数重数)

设 λ_0 是 n 维线性空间 V 上的线性变换 φ 的一个特征值, V_0 是属于 λ_0 的特征子空间, 称 $\dim V_0$ 为 λ_0 的**度数或几何重数**. λ_0 作为 φ 的特征多项式根的重数称为 λ_0 的**重数或代数重数**.

笔记 由线性映射的维数公式可知, 特征值 λ_0 的度数 $\dim V_0 = \dim \mathrm{Ker}(\lambda_0 I_V - \varphi) = n - r(\lambda_0 I_V - \varphi)$, 而特征值 λ_0 的重数则由特征多项式 $|\lambda I_V - \varphi|$ 的因式分解决定.

定义 1.14 (矩阵的几何重数与代数重数)

设 λ_0 是 n 阶方阵的 A 的一个特征值, V_0 是属于 λ_0 的特征子空间, 称 $\dim V_0$ 为 λ_0 的**度数或几何重数**.λ_0 作为 A 的特征多项式根的重数称为 λ_0 的**重数或代数重数**.

笔记 由线性方程组的理论可知, 特征值 λ_0 的度数 $\dim V_0 = n - r(\lambda_0 I_n - A)$, 若将 A 看作由矩阵 A 乘法诱导的 V 上的线性变换, 则由线性变换的维数公式可知 $\dim V_0 = \dim \mathrm{Ker}(\lambda_0 I_V - A) = n - r(\lambda_0 I_V - A)$. 而特征值 λ_0 的重数则由特征多项式 $|\lambda I_n - A|$ 的因式分解决定.

引理 1.2 (特征值的几何重数数总小于代数重数)

设 φ 是 n 维线性空间 V 上的线性变换, λ_0 是 φ 的一个特征值, 则 λ_0 的度数总是小于等于 λ_0 的重数.

证明 设特征值 λ_0 的重数为 m, 度数为 t, 又 V_0 是属于 λ_0 的特征子空间, 则 $\dim V_0 = t$. 设 $\{e_1, \cdots, e_t\}$ 是 V_0 的一组基. 由于 V_0 中的非零向量都是 φ 关于 λ_0 的特征向量, 故

$$\varphi(e_i) = \lambda_0 e_i, \quad i = 1, \cdots, t.$$

将 $\{e_1,\cdots,e_t\}$ 扩充为 V 的一组基, 记为 $\{e_1,\cdots,e_t,e_{t+1},\cdots,e_n\}$, 则 φ 在这组基下的表示矩阵为

$$A=\begin{pmatrix}\lambda_0 I_t & * \\ O & B\end{pmatrix},$$

其中 B 是一个 $n-t$ 阶方阵. 因此, 线性变换 φ 的特征多项式具有如下形式:

$$|\lambda I_V-\varphi|=|\lambda I_n-A|=(\lambda-\lambda_0)^t|\lambda I_{n-t}-B|,$$

这表明 λ_0 的重数至少为 t, 即 $t\le m$. □

定义 1.15 (完全的特征向量系)

设 λ_0 是 φ (或 A) 的 m 重特征值, 即它是 φ (或 A) 的特征多项式的 m 重根. 此时若有 $m=\dim V_{\lambda_0}$, 即 λ_0 的代数重数和几何重数相等, 则称 λ_0 **有完全的特征向量系**。若对 φ (或 A) 的任一特征值, 其代数重数和几何重数都相等, 则称 **φ (或 A) 有完全的特征向量系**. ♣

定理 1.10 (线性变换可对角化的充要条件)

设 φ 是 n 维线性空间 V 上的线性变换, 则 φ 可对角化的充分必要条件是 φ 有完全的特征向量系. ♡

证明 设 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是 φ 的全部不同的特征值, 它们对应的特征子空间、重数和度数分别记为 $V_i,m_i,t_i(i=1,2,\cdots,k)$. 由重数的定义以及引理 1.2可知 $m_1+m_2+\cdots+m_k=n, t_i\le m_i, i=1,2,\cdots,k$.

由定理 1.9可知, 我们只要证明 φ 有完全的特征向量系当且仅当 $V=V_1\oplus V_2\oplus\cdots\oplus V_k$.

若 $V=V_1\oplus V_2\oplus\cdots\oplus V_k$, 则

$$\begin{aligned}n=\dim V&=\dim(V_1\oplus V_2\oplus\cdots\oplus V_k)\\&=\dim V_1+\dim V_2+\cdots+\dim V_k\\&=\sum_{i=1}^k t_i\le\sum_{i=1}^k m_i=n,\end{aligned}$$

因此 $t_i=m_i(i=1,2,\cdots,k)$, 即 φ 有完全的特征向量系. 反过来, 若 φ 有完全的特征向量系, 则

$$\dim(V_1\oplus V_2\oplus\cdots\oplus V_k)=\sum_{i=1}^k t_i=\sum_{i=1}^k m_i=n=\dim V,$$

又 $V_1\oplus V_2\oplus\cdots\oplus V_k\subset V$, 故 $V=V_1\oplus V_2\oplus\cdots\oplus V_k$ 成立. □

定理 1.11

设 A 为 n 阶复矩阵, 其全体特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_r$, 并且 $\lambda_i\,(1\le i\le r)$ 的代数重数为 n_i, 则 $\sum\limits_{i=1}^r n_i=n$. 若 A 可对角化, 则 A 一定相似于 $\operatorname{diag}\{A_1,A_2,\cdots,A_r\}$, 其中 $A_i=\operatorname{diag}\{\lambda_i,\lambda_i,\cdots,\lambda_i\}\,(1\le i\le r)$ 并且阶数为 n_i. ♡

证明 由于 A 可对角化, 因此其特征值的代数重数等于几何重数. 记 V_i 为 λ_i 的特征子空间, 则任取 V_i 中一组基 $\{e_i1,e_i2,\cdots,e_{i,n_i}\}$. 由可对角化的判定条件 (3)及直和的等价条件可知,$\{e_i1,e_i2,\cdots,e_{i,n_i}\}\,(1\le i\le r)$ 可以拼成 $\mathbb{C}^n$ 的一组基. 于是记 $P=(e_{11},\cdots,e_{1,n_1},\cdots,e_{r1},\cdots,e_{r,n_r})$, 则 P 可逆, 并且

$$\begin{aligned}AP&=A\left(e_{11},\cdots,e_{1,n_1},\cdots,e_{r1},\cdots,e_{r,n_r}\right)=\left(\lambda_1 e_{11},\cdots,\lambda_1 e_{1,n_1},\cdots,\lambda_r e_{r1},\cdots,\lambda_r e_{r,n_r}\right)\\&=\left(e_{11},\cdots,e_{1,n_1},\cdots,e_{r1},\cdots,e_{r,n_r}\right)\operatorname{diag}\{A_1,A_2,\cdots,A_r\}=P\operatorname{diag}\{A_1,A_2,\cdots,A_r\}.\end{aligned}$$

故 $P^{-1}AP=\operatorname{diag}\{A_1,A_2,\cdots,A_r\}$. 结论得证. □

定理 1.12 (可对角化的判定条件)

判定 n 阶复矩阵 $\boldsymbol{A}$ (或 n 维复线性空间 V 上的线性变换 φ) 是否可对角化, 通常有以下 7 种方法:

(1) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量;

(2) 若 $\boldsymbol{A}$ 有 n 个不同的特征值, 则 $\boldsymbol{A}$ 可对角化;

(3) $\boldsymbol{A}$ 可对角化的充要条件是 $\mathbb{C}^n$ 是 $\boldsymbol{A}$ 的特征子空间的直和;

(4) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 有完全的特征向量系, 即对 $\boldsymbol{A}$ 的任一特征值, 其几何重数等于其代数重数;

(5) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 的极小多项式无重根;

(6) $\boldsymbol{A}$ 可对角化的充要条件是 $\boldsymbol{A}$ 的 Jordan 块都是一阶的 (或 $\boldsymbol{A}$ 的初等因子都是一次多项式);

(7) 若 $\boldsymbol{A}$ 相似于实对称矩阵或复正规矩阵, 则 $\boldsymbol{A}$ 可对角化.

注 上述第五、第六种方法将放在 §7.5 进行探讨, 另外命题**??**也是可对角化判定准则的一个补充; 第七种方法将放在 §9.7.4 进行探讨; 本节主要阐述可对角化判定的前 4 种方法.

证明 **几何形式**:(即 n 维复线性空间 V 上的线性变换 φ 可对角化的条件)

(1) 证明见定理 1.6.

(2) 证明见定理 1.9.

(3)

(4) 证明见定理 1.10.

(5)

(6)

(7)

代数形式:(即 n 阶复矩阵可对角化的条件) 由上述几何形式的结论及引理 1.1立即得到证明. □

注 若要考虑数域 $\mathbb{F}$ 上的 n 阶矩阵 $\boldsymbol{A}$ (或 $\mathbb{F}$ 上 n 维线性空间 V 上的线性变换 φ) 在 $\mathbb{F}$ 上的可对角化问题, 那么首先需要验证 $\boldsymbol{A}$ (或 φ) 的特征值都在 $\mathbb{F}$ 中, 否则由可对角化的定义可知, $\boldsymbol{A}$ (或 φ) 在 $\mathbb{F}$ 上必不可对角化. 若假设 $\boldsymbol{A}$ (或 φ) 的特征值都在 $\mathbb{F}$ 中, 则 $\boldsymbol{A}$ (或 φ) 在 $\mathbb{F}$ 上的可对角化判定准则也是上述前 6 种方法. 因此, 为了突出重点, 本节总是在复数域 $\mathbb{C}$ 上考虑可对角化问题. 请读者自行将某些例题推广到数域 $\mathbb{F}$ 的情形.

1.4.2 有 n 个线性无关的特征向量

寻找 $\boldsymbol{A}$ 的 n 个线性无关的特征向量, 等价于寻找 n 阶可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP}$ 为对角矩阵.

命题 1.28 (循环矩阵一定可对角化)

求证: 复数域上 n 阶循环矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix}$$

可对角化, 并求出它相似的对角矩阵及过渡矩阵.

笔记 这个命题实际上就是命题**??**.

证明 设 $f(x) = a_1 + a_2 x + \cdots + a_n x^{n-1}, \omega_k = \cos\dfrac{2k\pi}{n} + \mathrm{i}\sin\dfrac{2k\pi}{n} (0 \le k \le n-1)$, 则

$$\begin{pmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{pmatrix} \begin{pmatrix} 1 \\ \omega_k \\ \vdots \\ \omega_k^{n-1} \end{pmatrix} = f(\omega_k) \begin{pmatrix} 1 \\ \omega_k \\ \vdots \\ \omega_k^{n-1} \end{pmatrix}.$$

这表明 $(1,\omega_k,\cdots,\omega_k^{n-1})'$ 是 $\boldsymbol{A}$ 的属于特征值 $f(\omega_k)$ 的特征向量. 令

$$\boldsymbol{P}=\begin{pmatrix}1 & 1 & \cdots & 1\\ 1 & \omega_1 & \cdots & \omega_{n-1}\\ \vdots & \vdots & & \vdots\\ 1 & \omega_1^{n-1} & \cdots & \omega_{n-1}^{n-1}\end{pmatrix},$$

由 Vandermonde 行列式可知 $|\boldsymbol{P}|\neq 0$, 从而这 n 个特征向量线性无关, 因此 $\boldsymbol{A}$ 可对角化, 且有

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}\{f(1),f(\omega_1),\cdots,f(\omega_{n-1})\}.$$

□

例题 1.31 设 n 阶复矩阵 $\boldsymbol{A}$ 可对角化, 证明: 矩阵 $\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}^2\\ \boldsymbol{A}^2 & \boldsymbol{A}\end{pmatrix}$ 也可对角化.

证明 证法一: 因为 $\boldsymbol{A}$ 可对角化, 故可设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量, 满足 $\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i(1\leq i\leq n)$. 注意到

$$\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}^2\\ \boldsymbol{A}^2 & \boldsymbol{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_i\\ \boldsymbol{\alpha}_i\end{pmatrix}=(\lambda_i+\lambda_i^2)\begin{pmatrix}\boldsymbol{\alpha}_i\\ \boldsymbol{\alpha}_i\end{pmatrix},\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}^2\\ \boldsymbol{A}^2 & \boldsymbol{A}\end{pmatrix}\begin{pmatrix}\boldsymbol{\alpha}_i\\ -\boldsymbol{\alpha}_i\end{pmatrix}=(\lambda_i-\lambda_i^2)\begin{pmatrix}\boldsymbol{\alpha}_i\\ -\boldsymbol{\alpha}_i\end{pmatrix}.$$

通过定义不难验证 $\begin{pmatrix}\boldsymbol{\alpha}_i\\ \boldsymbol{\alpha}_i\end{pmatrix},\begin{pmatrix}\boldsymbol{\alpha}_i\\ -\boldsymbol{\alpha}_i\end{pmatrix}(1\leq i\leq n)$ 是线性无关的, 因此 $\begin{pmatrix}\boldsymbol{A} & \boldsymbol{A}^2\\ \boldsymbol{A}^2 & \boldsymbol{A}\end{pmatrix}$ 有 $2n$ 个线性无关的特征向量, 从而可对角化.

证法二: 容易验证 $\begin{pmatrix}I_n & I_n\\ I_n & -I_n\end{pmatrix}$ 的逆矩阵为 $\frac{1}{2}\begin{pmatrix}I_n & I_n\\ I_n & -I_n\end{pmatrix}$. 考虑如下相似变换:

$$\frac{1}{2}\begin{pmatrix}I_n & I_n\\ I_n & -I_n\end{pmatrix}\begin{pmatrix}A & A^2\\ A^2 & A\end{pmatrix}\begin{pmatrix}I_n & I_n\\ I_n & -I_n\end{pmatrix}=\begin{pmatrix}A+A^2 & O\\ O & A-A^2\end{pmatrix}.$$

由命题 1.31可知,$A+A^2,A-A^2$ 作为 A 的多项式也可对角化, 故原矩阵可对角化. 具体地, 设 P 为可逆矩阵, 使得 $P^{-1}AP=\Lambda$ 为对角矩阵, 则

$$\begin{pmatrix}P^{-1} & O\\ O & P^{-1}\end{pmatrix}\begin{pmatrix}A+A^2 & O\\ O & A-A^2\end{pmatrix}\begin{pmatrix}P & O\\ O & P\end{pmatrix}=\begin{pmatrix}\Lambda+\Lambda^2 & O\\ O & \Lambda-\Lambda^2\end{pmatrix}$$

为对角矩阵, 因此原矩阵可对角化. □

例题 1.32 设 V 为 n 阶矩阵全体构成的线性空间, V 上的线性变换 φ 定义为 $\varphi(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}$, 其中 $\boldsymbol{A}\in V$. 证明: 若 $\boldsymbol{A}$ 可对角化, 则 φ 也可对角化.

证明 设 $\boldsymbol{P}$ 为 n 阶可逆矩阵, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$, 则 $\boldsymbol{P}'\boldsymbol{A}'(\boldsymbol{P}')^{-1}=\boldsymbol{\Lambda}$, 即 $\boldsymbol{A}'$ 也可对角化. 设

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),(\boldsymbol{P}')^{-1}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$$

分别为两个矩阵的列分块, 则

$$\boldsymbol{A}\boldsymbol{\alpha}_i=\lambda_i\boldsymbol{\alpha}_i,\boldsymbol{A}'\boldsymbol{\beta}_j=\lambda_j\boldsymbol{\beta}_j,1\leq i,j\leq n,$$

且 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n$ 线性无关, $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n$ 线性无关. 由命题**??**可知, $\{\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',1\leq i,j\leq n\}$ 是 V 中 n^2 个线性无关的矩阵. 注意到

$$\varphi(\boldsymbol{\alpha}_i\boldsymbol{\beta}_j')=\boldsymbol{A}\boldsymbol{\alpha}_i\boldsymbol{\beta}_j'\boldsymbol{A}=(\boldsymbol{A}\boldsymbol{\alpha}_i)(\boldsymbol{A}'\boldsymbol{\beta}_j)'=\lambda_i\lambda_j\boldsymbol{\alpha}_i\boldsymbol{\beta}_j',$$

故 φ 有 n^2 个线性无关的特征向量, 从而可对角化. □

例题 1.33 设 V 为 n 阶矩阵全体构成的线性空间, V 上的线性变换 φ 定义为 $\varphi(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}-\boldsymbol{X}\boldsymbol{A}$, 其中 $\boldsymbol{A}\in V$. 证明: 若 $\boldsymbol{A}$ 可对角化, 则 φ 也可对角化.

证明 设 $\boldsymbol{P}$ 为 n 阶可逆矩阵, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$, 则 $\boldsymbol{P}'\boldsymbol{A}'(\boldsymbol{P}')^{-1}=\boldsymbol{\Lambda}$, 即 $\boldsymbol{A}'$ 也可对角化. 设

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n),(\boldsymbol{P}')^{-1}=(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_n)$$

分别为两个矩阵的列分块, 则

$$A\alpha_i = \lambda_i\alpha_i, A'\beta_j = \lambda_j\beta_j, 1 \le i, j \le n,$$

且 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 线性无关, $\beta_1, \beta_2, \cdots, \beta_n$ 线性无关. 由命题**??**可知, $\{\alpha_i\beta_j', 1 \le i, j \le n\}$ 是 V 中 n^2 个线性无关的矩阵. 注意到

$$\varphi(\alpha_i\beta_j') = A\alpha_i\beta_j' - \alpha_i\beta_j'A = (A\alpha_i)\beta_j' - \alpha_i(A'\beta_j)' = (\lambda_i - \lambda_j)\alpha_i\beta_j',$$

故 φ 有 n^2 个线性无关的特征向量, 从而可对角化. □

1.4.3 有 n 个不同特征值

由于属于不同特征值的特征向量线性无关, 故若 A 有 n 个不同的特征值, 则 A 必有 n 个线性无关的特征向量, 从而可对角化. 请注意 A 有 n 个不同的特征值只是可对角化的充分条件, 而并非必要条件.

例题 1.34 设 A 是实二阶矩阵且 $|A| < 0$, 求证: A 实相似于对角矩阵.

证明 设

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

由 $|A| < 0$ 可得 $ad - bc < 0$. 又 A 的特征多项式

$$|\lambda I_2 - A| = \lambda^2 - (a+d)\lambda + (ad - bc),$$

上述关于 λ 的二次方程其判别式大于零, 从而 A 有两个不相等的实特征值, 因此 A 实相似于对角矩阵. □

例题 1.35 设 A, B, C 都是 n 阶矩阵, A, B 各有 n 个不同的特征值, 又 $f(\lambda)$ 是 A 的特征多项式, 且 $f(B)$ 是可逆矩阵. 求证: 矩阵

$$M = \begin{pmatrix} A & C \\ O & B \end{pmatrix}$$

相似于对角矩阵.

证明 任取 B 的一个特征值 μ_0, 则 $f(\mu_0)$ 是 $f(B)$ 的特征值. 由于 $f(B)$ 可逆, 故 $f(B)$ 的特征值非零, 从而 $f(\mu_0) \ne 0$, 即 μ_0 不是 A 的特征值, 于是 A 和 B 的特征值互不相同. 注意到

$$|\lambda I_{2n} - M| = \begin{vmatrix} \lambda I_n - A & -C \\ O & \lambda I_n - B \end{vmatrix} = |\lambda I_n - A||\lambda I_n - B|,$$

故矩阵 M 有 $2n$ 个不同的特征值, 从而相似于对角矩阵. □

例题 1.36 设 n 阶矩阵 A, B 有相同的特征值, 且这 n 个特征值互不相等. 求证: 存在 n 阶矩阵 P, Q, 使得 $A = PQ, B = QP$.

证明 由假设以及定理 1.11和例题 1.22可知, 矩阵 A, B 相似于同一个对角矩阵, 因此 A 和 B 相似. 不妨设 $B = P^{-1}AP$, 令 $Q = P^{-1}A$, 则 $PQ = A, QP = B$. □

命题 1.29

设 A, B 是 n 阶矩阵,A 有 n 个不同的特征值, 并且 $AB = BA$, 求证:B 相似于对角矩阵, 并且 A 与 B 可同时对角化.

证明 证法一 (几何方法): 因为 A 有 n 个不同的特征值, 故 A 可对角化. 令 V 是 n 维复列向量空间, 将 A, B 看成是 V 上的线性变换. 又设 A 的特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 对应的特征向量为 $\alpha_1, \alpha_2, \cdots, \alpha_n$, 则 λ_i 的特征子空间 $V_i = L(\alpha_i)(1 \le i \le n)$, 且

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_n.$$

注意到 $AB = BA$, 故由命题 1.16可知,A 的特征子空间 V_i 是 B 的不变子空间. 将 B 限制在 V_i 上, 这是一维线性空间 V_i 上的线性变换, 从而只能是纯量变换, 即存在 μ_i, 使得 $B\alpha_i = \mu_i\alpha_i(1 \le i \le n)$, 于是 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 也是 B 的特征

向量. 因此,B 有 n 个线性无关的特征向量, 从而 B 可对角化. 事实上, 我们得到了一个更强的结果:A 和 B 可同时对角化, 即存在可逆矩阵 $P=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 使得 $P^{-1}AP=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 和 $P^{-1}BP=\mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_n\}$ 都是对角矩阵.

证法二 (代数方法): 因为 A 有 n 个不同的特征值, 故 A 可对角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$. 注意到问题的条件和结论在同时相似变换:$A\mapsto P^{-1}AP, B\mapsto P^{-1}BP$ 下保持不变, 故不妨从一开始就假设 $A=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 为对角矩阵. 设 $B=(b_{ij})$, 则

$$AB=\begin{pmatrix}\lambda_1b_{11}&\lambda_1b_{12}&\cdots&\lambda_1b_{1n}\\\lambda_2b_{21}&\lambda_2b_{22}&\cdots&\lambda_2b_{2n}\\\vdots&\vdots&&\vdots\\\lambda_nb_{n1}&\lambda_nb_{n2}&\cdots&\lambda_nb_{nn}\end{pmatrix}=\begin{pmatrix}\lambda_1b_{11}&\lambda_2b_{12}&\cdots&\lambda_nb_{1n}\\\lambda_1b_{21}&\lambda_2b_{22}&\cdots&\lambda_nb_{2n}\\\vdots&\vdots&&\vdots\\\lambda_1b_{n1}&\lambda_2b_{n2}&\cdots&\lambda_nb_{nn}\end{pmatrix}=BA,$$

比较元素可得 $\lambda_ib_{ij}=\lambda_jb_{ij}$. 注意到 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 互不相同, 故 $b_{ij}=0(i\neq j)$, 即 B 为对角矩阵. □

命题 1.30

设 A,B 是 n 阶矩阵,A 有 n 个不同的特征值, 并且 $AB=BA$, 求证: 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $B=f(A)$.

证明 由命题 1.29可知 A 和 B 可以同时对角化, 即存在可逆矩阵 P, 使得

$$P^{-1}AP=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},P^{-1}BP=\mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_n\},$$

其中 λ_i,μ_i 分别是 A,B 的特征值. 因为 λ_i 互不相同, 故由 Lagrange 插值公式可知, 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $f(\lambda_i)=\mu_i(1\le i\le n)$. 于是

$$P^{-1}BP=\mathrm{diag}\{f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)\}=f(P^{-1}AP)=P^{-1}f(A)P,$$

从而 $B=f(A)$. □

命题 1.31

若 A 可对角化, 则对任意的多项式 $f(x)$,$f(A)$ 也可对角化.

笔记 这一结论提醒我们: 在处理可对角化问题时, 如能将矩阵写成可对角化矩阵的多项式, 则往往讨论起来更加方便.

证明 事实上, 设 P 为可逆矩阵, 使得 $P^{-1}AP=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$ 为对角矩阵, 则 $P^{-1}f(A)P=f(P^{-1}AP)=\mathrm{diag}\{f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)\}$ 也为对角矩阵. □

命题 1.32

设 A 是 n 阶复矩阵且有 n 个不同的特征值, 求证:n 阶复矩阵 B 可对角化的充要条件是存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 B 相似于 $f(A)$.

证明 先证充分性. 由于 A 有 n 个不同的特征值, 故 A 可对角化, 从而由命题 1.31$f(A)$ 也可对角化, 又 B 相似于 $f(A)$, 于是 B 也可对角化.

再证必要性. 设 P,Q 为可逆矩阵, 使得

$$P^{-1}AP=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},Q^{-1}BQ=\mathrm{diag}\{\mu_1,\mu_2,\cdots,\mu_n\},$$

其中 λ_i,μ_i 分别是 A,B 的特征值. 因为 λ_i 互不相同, 故由 Lagrange 插值公式可知, 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $f(\lambda_i)=\mu_i(1\le i\le n)$. 于是

$$Q^{-1}BQ=\mathrm{diag}\{f(\lambda_1),f(\lambda_2),\cdots,f(\lambda_n)\}=f(P^{-1}AP)=P^{-1}f(A)P,$$

即有 $B=(PQ^{-1})^{-1}f(A)(PQ^{-1})$, 从而 B 相似于 $f(A)$. □

推论 1.6

n 阶复方阵 B 可对角化的充要条件是 B 相似于某个循环矩阵.

证明 设 $J=\begin{pmatrix} O & I_{n-1} \\ 1 & O \end{pmatrix}$, 经简单计算可得 $|\lambda I_n - J| = \lambda^n - 1$, 于是 J 有 n 个不同的特征值. 对任一循环矩阵 C, 由循环矩阵的性质 2 可知, 存在次数不超过 $n-1$ 的多项式 $f(x)$, 使得 $C = f(J)$, 故由命题 1.32即得本推论. □

命题 1.33

设 a, b, c 为复数且 $bc \neq 0$, 证明下列 n 阶矩阵 A 可对角化:

$$A = \begin{pmatrix} a & b & & & & \\ c & a & b & & & \\ & c & a & b & & \\ & & \ddots & \ddots & \ddots & \\ & & & c & a & b \\ & & & & c & a \end{pmatrix}$$

笔记 当 $(\lambda-a)^2 = 4bc$ 时, 利用摄动法, 设 $t > 0$, 则当 $(\lambda-a)^2 - 4bc = 0$ 时,$(\lambda+t-a)^2 - 4bc > 0$, 由下述证明可知,$\lambda + t$ 有 n 个不同取值, 从而令 $t \to 0$, 则此时 $(\lambda-a)^2 - 4bc = 0$, 并且 λ 仍有 n 个不同取值.

证明 我们先来计算 A 的特征多项式 $|\lambda I_n - A|$. 设 x_1, x_2 是二次方程 $x^2 - (\lambda-a)x + bc = 0$ 的两个根, 则当 $(\lambda-a)^2 \neq 4bc$ 时, 由推论**??**可得

$$|\lambda I_n - A| = \frac{x_1^{n+1} - x_2^{n+1}}{x_1 - x_2}.$$

注意到 x_1, x_2 都是关于 λ 的连续函数, 要求 A 的特征值 λ, 即是求 λ 的值, 使得 $|\lambda I_n - A| = 0$, 而这也等价于 $x_1^{n+1} = x_2^{n+1}$. 令 $\omega = \cos\frac{2\pi}{n+1} + \mathrm{i}\sin\frac{2\pi}{n+1}$ 为 1 的 $n+1$ 次方根, 则由 $x_1^{n+1} = x_2^{n+1}$ 可得 $x_1 = x_2\omega^k (1 \leq k \leq n)$. 由 Vieta 定理可得 $x_1x_2 = bc$, 在选定 bc 的某一平方根 $\sqrt{bc}$ 之后, 可解出

$$x_1 = \sqrt{bc}\left(\cos\frac{k\pi}{n+1} + \mathrm{i}\sin\frac{k\pi}{n+1}\right), x_2 = \sqrt{bc}\left(\cos\frac{k\pi}{n+1} - \mathrm{i}\sin\frac{k\pi}{n+1}\right), 1 \leq k \leq n.$$

再次由 Vieta 定理可得 $\lambda - a = x_1 + x_2 = 2\sqrt{bc}\cos\frac{k\pi}{n+1}$, 即

$$\lambda = a + 2\sqrt{bc}\cos\frac{k\pi}{n+1}, 1 \leq k \leq n.$$

容易验证上述 n 个数的确是 A 的 n 个不同的特征值, 从而 A 可对角化. □

1.4.4 全空间等于特征子空间的直和

矩阵或线性变换可对角化当且仅当全空间等于特征子空间的直和这一判定准则, 不仅给了我们很多几何想象的空间, 而且与矩阵或线性变换适合的多项式密切相关.

命题 1.34

设 n 阶矩阵 $\boldsymbol{A}$ 适合首一多项式 $g(x)$, 并且 $g(x)$ 在复数域中无重根, 证明:$\boldsymbol{A}$ 可对角化.

证明 设 $g(x) = (x-a_1)(x-a_2)\cdots(x-a_m)$ 是复数域上的因式分解, 其中 $a_1, a_2, \cdots, a_m$ 是互异的复数. 我们先来证明:

$$\mathbb{C}^n = \mathrm{Ker}(\boldsymbol{A} - a_1\boldsymbol{I}_n) \oplus \mathrm{Ker}(\boldsymbol{A} - a_2\boldsymbol{I}_n) \oplus \cdots \oplus \mathrm{Ker}(\boldsymbol{A} - a_m\boldsymbol{I}_n). \tag{1.10}$$

设 $g_i(x)=\prod_{j\neq i}(x-a_j)$, 则 $(g_1(x),g_2(x),\cdots,g_m(x))=1$, 故存在 $u_i(x)(1\leq i\leq m)$, 使得

$$g_1(x)u_1(x)+g_2(x)u_2(x)+\cdots+g_m(x)u_m(x)=1.$$

代入 $x=\boldsymbol{A}$, 可得恒等式

$$g_1(\boldsymbol{A})u_1(\boldsymbol{A})+g_2(\boldsymbol{A})u_2(\boldsymbol{A})+\cdots+g_m(\boldsymbol{A})u_m(\boldsymbol{A})=\boldsymbol{I}_n. \tag{1.11}$$

对任一 $\boldsymbol{\alpha}\in\mathbb{C}^n$, 由上式可知 $\boldsymbol{\alpha}=\sum_{i=1}^m g_i(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha}$. 注意到 $(\boldsymbol{A}-a_i\boldsymbol{I}_n)g_i(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha}=g(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha}=\boldsymbol{0}$, 故 $g_i(\boldsymbol{A})u_i(\boldsymbol{A})\boldsymbol{\alpha}\in\mathrm{Ker}(\boldsymbol{A}-a_i\boldsymbol{I}_n)$, 于是

$$\mathbb{C}^n=\mathrm{Ker}(\boldsymbol{A}-a_1\boldsymbol{I}_n)+\mathrm{Ker}(\boldsymbol{A}-a_2\boldsymbol{I}_n)+\cdots+\mathrm{Ker}(\boldsymbol{A}-a_m\boldsymbol{I}_n). \tag{1.12}$$

任取 $\boldsymbol{\alpha}\in\mathrm{Ker}(\boldsymbol{A}-a_1\boldsymbol{I}_n)\cap(\mathrm{Ker}(\boldsymbol{A}-a_2\boldsymbol{I}_n)+\cdots+\mathrm{Ker}(\boldsymbol{A}-a_m\boldsymbol{I}_n))$, 则 $\boldsymbol{\alpha}=\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_m$, 其中 $\boldsymbol{\alpha}_i\in\mathrm{Ker}(\boldsymbol{A}-a_i\boldsymbol{I}_n)(i\geq 2)$. 由(1.11)式可知

$$\boldsymbol{\alpha}=u_1(\boldsymbol{A})g_1(\boldsymbol{A})(\boldsymbol{\alpha}_2+\cdots+\boldsymbol{\alpha}_m)+u_2(\boldsymbol{A})g_2(\boldsymbol{A})\boldsymbol{\alpha}+\cdots+u_m(\boldsymbol{A})g_m(\boldsymbol{A})\boldsymbol{\alpha}=\boldsymbol{0}.$$

注意到下指标可任意选, 故(1.12)式是直和.

由于 $\boldsymbol{A}$ 适合 $g(x)$, 故 $\boldsymbol{A}$ 的特征值也适合 $g(x)$, 从而只可能是 $a_1,a_2,\cdots,a_m$ 中的一部分. 不妨设 a_1 不是 $\boldsymbol{A}$ 的特征值, 则对 $\forall\alpha\neq 0$, 都有 $(\boldsymbol{A}-a_1\boldsymbol{I}_n)\alpha\neq 0$, 故 $\mathrm{Ker}(\boldsymbol{A}-a_1\boldsymbol{I}_n)=0$. 于是在(1.10)式中剔除等于零的直和分量, 这就证明了全空间等于特征子空间的直和, 从而 $\boldsymbol{A}$ 可对角化. □

例题 1.37 求证:

(1) 若 n 阶矩阵 $\boldsymbol{A}$ 适合 $\boldsymbol{A}^2=\boldsymbol{I}_n$, 则 $\boldsymbol{A}$ 必可对角化;

(2) 若 n 阶矩阵 $\boldsymbol{A}$ 适合 $\boldsymbol{A}^2=\boldsymbol{A}$, 则 $\boldsymbol{A}$ 必可对角化.

(3) $M_n(\mathbb{F})$ 上的线性变换 $\boldsymbol{\eta}$ 满足 $\boldsymbol{\eta}^2=\boldsymbol{I}_V$, 则 $\boldsymbol{\eta}$ 可对角化.

证明

(1) 对合矩阵 $\boldsymbol{A}$ 适合多项式 x^2-1, 它在复数域中无重根, 故由命题 1.34即得结论.

(2) 幂等矩阵 $\boldsymbol{A}$ 适合多项式 x^2-x, 它在复数域中无重根, 故由命题 1.34即得结论.

(3) 线性变换 $\boldsymbol{\eta}$ 适合多项式 x^2-1, 它在复数域中无重根, 故由命题 1.34即得结论.

□

1.4.5 有完全的特征向量系

矩阵或线性变换有完全的特征向量系, 即任一特征值的代数重数等于其几何重数, 也就是特征值与线性无关的特征向量完全一一对应.

例题 1.38 若矩阵 $\boldsymbol{A},\boldsymbol{B}$ 有完全的特征向量系, 求证: $\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}$ 也有完全的特征向量系.

证明 因为 $\boldsymbol{A},\boldsymbol{B}$ 有完全的特征向量系, 故相似于对角矩阵. 设 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 和 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}$ 是对角矩阵, 则

$$\begin{pmatrix}\boldsymbol{P}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{Q}\end{pmatrix}^{-1}\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{P}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{Q}\end{pmatrix}=\begin{pmatrix}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q}\end{pmatrix}$$

是对角矩阵. 因此 $\begin{pmatrix}\boldsymbol{A}&\boldsymbol{O}\\\boldsymbol{O}&\boldsymbol{B}\end{pmatrix}$ 有完全的特征向量系. □

例题 1.39 设 n 阶矩阵 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}_r&\boldsymbol{B}\\\boldsymbol{O}&-\boldsymbol{I}_{n-r}\end{pmatrix}$, 求证: $\boldsymbol{A}$ 可对角化.

证明 证法一: 显然 $\boldsymbol{A}$ 有特征值 1 (r 重) 与 -1 ($n-r$ 重). 注意到矩阵 $\boldsymbol{I}_n-\boldsymbol{A}=\begin{pmatrix}\boldsymbol{O}&-\boldsymbol{B}\\\boldsymbol{O}&2\boldsymbol{I}_{n-r}\end{pmatrix}$ 的秩等于 $n-r$, 因此特征值 1 的几何重数等于 $n-\mathrm{r}(\boldsymbol{I}_n-\boldsymbol{A})=r$, 与其代数重数相等. 同理可证特征值 -1 的几何重数为 $n-r$, 与其代数重数相同. 因此 $\boldsymbol{A}$ 可对角化, 且相似于对角矩阵 $\mathrm{diag}\{\boldsymbol{I}_r,-\boldsymbol{I}_{n-r}\}$.

证法二: 容易算出 $\boldsymbol{A}^2=\boldsymbol{I}_n$, 由例题 1.37(1)可知 $\boldsymbol{A}$ 可对角化.

证法三: 做第三种初等相似变换, 由 $\begin{pmatrix}\boldsymbol{I}_r & \frac{1}{2}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}_{n-r}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{B}\\ \boldsymbol{O} & -\boldsymbol{I}_{n-r}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_r & -\frac{1}{2}\boldsymbol{B}\\ \boldsymbol{O} & \boldsymbol{I}_{n-r}\end{pmatrix}=\begin{pmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & -\boldsymbol{I}_{n-r}\end{pmatrix}$ 即得. □

命题 1.35

设 m 阶矩阵 $\boldsymbol{A}$ 与 n 阶矩阵 $\boldsymbol{B}$ 没有公共的特征值, 且 $\boldsymbol{A},\boldsymbol{B}$ 均可对角化, 又 $\boldsymbol{C}$ 为 $m\times n$ 矩阵, 求证: $\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$ 也可对角化.

证明 任取 $\boldsymbol{A}$ 的特征值 λ_0, 记其代数重数为 $m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为 $t_{\boldsymbol{A}}(\lambda_0)$. 首先注意到 $\boldsymbol{A},\boldsymbol{B}$ 没有公共的特征值, 故 λ_0 不是 $\boldsymbol{B}$ 的特征值, 又 $|\lambda\boldsymbol{I}-\boldsymbol{M}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$, 从而 $m_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)$. 由于 $\lambda_0\boldsymbol{I}-\boldsymbol{B}$ 是非异阵, 故有如下分块矩阵的初等变换:

$$\lambda_0\boldsymbol{I}-\boldsymbol{M}=\begin{pmatrix}\lambda_0\boldsymbol{I}-\boldsymbol{A} & -\boldsymbol{C}\\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B}\end{pmatrix}\to\begin{pmatrix}\lambda_0\boldsymbol{I}-\boldsymbol{A} & \boldsymbol{O}\\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B}\end{pmatrix}.$$

因为矩阵的秩在分块初等变换下不变, 故由矩阵秩的等式可得

$$\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})=\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B})=\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+n,$$

于是 $t_{\boldsymbol{M}}(\lambda_0)=(m+n)-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})=m-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})=t_{\boldsymbol{A}}(\lambda_0)$. 因为 $\boldsymbol{A}$ 可对角化, 所以 $\boldsymbol{A}$ 有完全的特征向量系, 从而 $m_{\boldsymbol{A}}(\lambda_0)=t_{\boldsymbol{A}}(\lambda_0)$, 于是 $m_{\boldsymbol{M}}(\lambda_0)=t_{\boldsymbol{M}}(\lambda_0)$. 同理可证, 对 $\boldsymbol{B}$ 的任一特征值 μ_0, 成立 $m_{\boldsymbol{M}}(\mu_0)=t_{\boldsymbol{M}}(\mu_0)$. 因此 $\boldsymbol{M}$ 有完全的特征向量系, 从而可对角化. □

命题 1.36

设 $\boldsymbol{A}$ 为 m 阶矩阵, $\boldsymbol{B}$ 为 n 阶矩阵, $\boldsymbol{C}$ 为 $m\times n$ 矩阵, $\boldsymbol{M}=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{C}\\ \boldsymbol{O} & \boldsymbol{B}\end{pmatrix}$, 求证: 若 $\boldsymbol{M}$ 可对角化, 则 $\boldsymbol{A},\boldsymbol{B}$ 均可对角化.

注 这个命题的几何版本是: 设 φ 是复线性空间 V 上的线性变换, U 是 φ-不变子空间, 若 φ 可对角化, 则 φ 在不变子空间 U 上的限制变换 $\varphi|_U$ 以及 φ 在商空间 V/U 上的诱导变换 $\overline{\varphi}$ 均可对角化.

证明 任取 $\boldsymbol{M}$ 的特征值 λ_0, 记其代数重数为 $m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为 $t_{\boldsymbol{A}}(\lambda_0)$. 由 $|\lambda\boldsymbol{I}-\boldsymbol{M}|=|\lambda\boldsymbol{I}-\boldsymbol{A}||\lambda\boldsymbol{I}-\boldsymbol{B}|$ 可得 $m_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)+m_{\boldsymbol{B}}(\lambda_0)$. 考虑如下分块矩阵:

$$\lambda_0\boldsymbol{I}-\boldsymbol{M}=\begin{pmatrix}\lambda_0\boldsymbol{I}-\boldsymbol{A} & -\boldsymbol{C}\\ \boldsymbol{O} & \lambda_0\boldsymbol{I}-\boldsymbol{B}\end{pmatrix},$$

由矩阵秩的基本公式 (4) 可得

$$\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})\geq\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A})+\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B}),$$

于是 $t_{\boldsymbol{M}}(\lambda_0)=(m+n)-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{M})\leq(m-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{A}))+(n-\mathrm{r}(\lambda_0\boldsymbol{I}-\boldsymbol{B}))=t_{\boldsymbol{A}}(\lambda_0)+t_{\boldsymbol{B}}(\lambda_0)$. 由于几何重数总是小于等于代数重数, 故有

$$t_{\boldsymbol{M}}(\lambda_0)\leq t_{\boldsymbol{A}}(\lambda_0)+t_{\boldsymbol{B}}(\lambda_0)\leq m_{\boldsymbol{A}}(\lambda_0)+m_{\boldsymbol{B}}(\lambda_0)=m_{\boldsymbol{M}}(\lambda_0).$$

因为 $\boldsymbol{M}$ 可对角化, 所以 $\boldsymbol{M}$ 有完全的特征向量系, 从而 $t_{\boldsymbol{M}}(\lambda_0)=m_{\boldsymbol{M}}(\lambda_0)$, 再由上述不等式可得 $t_{\boldsymbol{A}}(\lambda_0)=m_{\boldsymbol{A}}(\lambda_0)$, $t_{\boldsymbol{B}}(\lambda_0)=m_{\boldsymbol{B}}(\lambda_0)$. 由 λ_0 的任意性即知, $\boldsymbol{A},\boldsymbol{B}$ 均有完全的特征向量系, 从而均可对角化. □

命题 1.37

设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵, $\boldsymbol{B}$ 为 $n\times m$ 矩阵, 又 $|\boldsymbol{BA}|\neq 0$, 求证: $\boldsymbol{AB}$ 可对角化的充要条件是 $\boldsymbol{BA}$ 可对角化.

证明 记其代数重数为 $m_{\boldsymbol{A}}(\lambda_0)$, 几何重数为 $t_{\boldsymbol{A}}(\lambda_0)$(其他记号同理). 由特征值的降价公式可得 $|\lambda\boldsymbol{I}_m-\boldsymbol{AB}|=\lambda^{m-n}|\lambda\boldsymbol{I}_n-\boldsymbol{BA}|$, 因此 $\boldsymbol{AB}$ 的特征值为 $\boldsymbol{BA}$ 的特征值以及 0. 由于 $\boldsymbol{BA}$ 非异, 故其特征值全部非零, 从而 0 作为

$\boldsymbol{AB}$ 的特征值, 其代数重数为 $m-n$. 另一方面, 我们有

$$n = \mathrm{r}(\boldsymbol{BA}) \le \min\{\mathrm{r}(\boldsymbol{A}), \mathrm{r}(\boldsymbol{B})\} \le \max\{\mathrm{r}(\boldsymbol{A}), \mathrm{r}(\boldsymbol{B})\} \le \min\{m, n\} = n,$$

从而 $\mathrm{r}(\boldsymbol{A}) = \mathrm{r}(\boldsymbol{B}) = n$. 再由 Sylvester 不等式可得

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

从而 $\mathrm{r}(\boldsymbol{AB}) = n$. 因此 0 作为 $\boldsymbol{AB}$ 的特征值, 其几何重数为 $m - \mathrm{r}(\boldsymbol{AB}) = m - n$, 即特征值 0 的代数重数等于几何重数. 任取 $\boldsymbol{BA}$ 的特征值 λ_0, 它也是 $\boldsymbol{AB}$ 的非零特征值, 显然 $m_{\boldsymbol{AB}}(\lambda_0) = m_{\boldsymbol{BA}}(\lambda_0)$. 考虑分块矩阵 $\begin{pmatrix} \boldsymbol{I}_m & \boldsymbol{A} \\ \boldsymbol{B} & \lambda_0 \boldsymbol{I}_n \end{pmatrix}$, 由秩的降阶公式可得

$$m + \mathrm{r}(\lambda_0 \boldsymbol{I}_n - \boldsymbol{BA}) = n + \mathrm{r}(\boldsymbol{I}_m - \frac{1}{\lambda_0}\boldsymbol{AB}) = n + \mathrm{r}(\lambda_0 \boldsymbol{I}_m - \boldsymbol{AB}),$$

于是 $t_{\boldsymbol{AB}}(\lambda_0) = m - \mathrm{r}(\lambda_0 \boldsymbol{I}_m - \boldsymbol{AB}) = n - \mathrm{r}(\lambda_0 \boldsymbol{I}_n - \boldsymbol{BA}) = t_{\boldsymbol{BA}}(\lambda_0)$. 由 λ_0 的任意性即知, $\boldsymbol{AB}$ 有完全的特征向量系当且仅当 $\boldsymbol{BA}$ 有完全的特征向量系, 从而 $\boldsymbol{AB}$ 可对角化当且仅当 $\boldsymbol{BA}$ 可对角化. □

1.4.6 利用反证法证明不可对角化

例题 1.40 求证:

(1) 若 n 阶矩阵 $\boldsymbol{A}$ 的特征值都是 λ_0, 但 $\boldsymbol{A}$ 不是纯量矩阵, 则 $\boldsymbol{A}$ 不可对角化. 特别地, 非零的幂零矩阵不可对角化.

(2) 若 n 阶实矩阵 $\boldsymbol{A}$ 适合 $\boldsymbol{A}^2 + \boldsymbol{A} + \boldsymbol{I}_n = \boldsymbol{O}$, 则 $\boldsymbol{A}$ 在实数域上不可对角化.

证明

(1) 用反证法, 设 $\boldsymbol{A}$ 可对角化, 则存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{\Lambda}$ 为对角矩阵. 由假设 $\boldsymbol{\Lambda}$ 的主对角元素全为 λ_0, 故 $\boldsymbol{\Lambda} = \lambda_0 \boldsymbol{I}_n$, 于是 $\boldsymbol{A} = \boldsymbol{P}(\lambda_0 \boldsymbol{I}_n)\boldsymbol{P}^{-1} = \lambda_0 \boldsymbol{I}_n$, 这与假设矛盾. 特别地, 幂零矩阵的特征值都是零 (特征值适合 $x^k = 0$), 因此也不可对角化.

(2) 用反证法, 设 $\boldsymbol{A}$ 在实数域上可对角化, 则 $\boldsymbol{A}$ 的特征值都是实数. 因为 $\boldsymbol{A}$ 适合多项式 x^2+x+1, 故由命题 1.7可知, $\boldsymbol{A}$ 的特征值也适合 x^2+x+1, 从而不可能是实数, 矛盾.

□

引理 1.3 (秩 1 矩阵的列向量分解)

设 $n(n>1)$ 阶矩阵 $\boldsymbol{A}$ 的秩为 1, 则存在非零列向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$, 使得 $\boldsymbol{A} = \boldsymbol{\alpha}\boldsymbol{\beta}'$.

证明 由于 $\mathrm{r}(\boldsymbol{A}) = 1$, 因此 $\boldsymbol{A}$ 的列向量成比例, 从而存在非零列向量 $\boldsymbol{\alpha}$, 使得 $\boldsymbol{A}$ 的列分块为

$$\boldsymbol{A} = (\boldsymbol{k_1\alpha}, \boldsymbol{k_2\alpha}, \cdots, \boldsymbol{k_n\alpha}).$$

其中 $\boldsymbol{k_1}, \boldsymbol{k_2}, \cdots, \boldsymbol{k_n}$ 为不全为零的实数. 否则,$\boldsymbol{A} = \boldsymbol{O}$ 矛盾! 于是令 $\boldsymbol{\beta} = (\boldsymbol{k_1}, \boldsymbol{k_2}, \cdots, \boldsymbol{k_n})' \ne \boldsymbol{0}$, 则

$$\boldsymbol{A} = (k_1\boldsymbol{\alpha}, k_2\boldsymbol{\alpha}, \cdots, \boldsymbol{k_n}\boldsymbol{\alpha}) = \boldsymbol{\alpha}\boldsymbol{\beta}'.$$

□

命题 1.38

设 $n(n>1)$ 阶矩阵 $\boldsymbol{A}$ 的秩为 1, 求证: $\boldsymbol{A}$ 可对角化的充要条件是 $\mathrm{tr}(\boldsymbol{A}) \ne 0$.

笔记 这个命题告诉我们命题 1.37的条件 $|BA| \ne 0$ 是必要的.

证明 由 $\mathrm{r}(\boldsymbol{A}) = 1$ 及秩 1 矩阵的列向量分解可知, 存在非零列向量 $\boldsymbol{\alpha}, \boldsymbol{\beta}$, 使得 $\boldsymbol{A} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, 于是由迹的交换性可得 $\mathrm{tr}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{\alpha}\boldsymbol{\beta}') = \mathrm{tr}(\boldsymbol{\beta}'\boldsymbol{\alpha}) = \boldsymbol{\beta}'\boldsymbol{\alpha}$.

证法一: 由例题 1.10及其可对角化的讨论可知, 令例题 1.10条件中的 $\boldsymbol{A} = \boldsymbol{I_n}$ 即可得到 A 可对角化的充要条件是 $\mathrm{tr}(\boldsymbol{A}) = \boldsymbol{\beta}'\boldsymbol{\alpha} \ne 0$ 或 $A = O$, 而 $\boldsymbol{A} = \boldsymbol{O}$ 与 $\mathrm{r}(\boldsymbol{A}) = 1$ 矛盾, 故 A 可对角化的充要条件是 $\mathrm{tr}(\boldsymbol{A}) =\ne 0$.

证法二: 注意到 $\boldsymbol{A}^2=(\boldsymbol{\alpha\beta}')(\boldsymbol{\alpha\beta}')=\boldsymbol{\alpha}(\boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{\beta}'=(\boldsymbol{\beta}'\boldsymbol{\alpha})\boldsymbol{\alpha\beta}'=\mathrm{tr}(\boldsymbol{A})\boldsymbol{A}$, 故 $\boldsymbol{A}$ 适合多项式 $x^2-\mathrm{tr}(\boldsymbol{A})x$. 若 $\mathrm{tr}(\boldsymbol{A})\neq 0$, 则由命题 1.34可知 $\boldsymbol{A}$ 可对角化; 若 $\mathrm{tr}(\boldsymbol{A})=0$, 则 $\boldsymbol{A}$ 是幂零矩阵, 又 $\boldsymbol{A}\neq\boldsymbol{O}$, 故由例题 1.40(1)可知 $\boldsymbol{A}$ 不可对角化. □

1.5 极小多项式与 Cayley-Hamilton 定理

命题 1.39

数域 $\mathbb{K}$ 上的 n 阶矩阵 A 一定适合数域 $\mathbb{K}$ 上的一个非零多项式.

证明 我们已经知道, 数域 $\mathbb{K}$ 上的 n 阶矩阵全体组成了 $\mathbb{K}$ 上的线性空间, 其维数等于 n^2. 因此对任一 n 阶矩阵 A, 下列 n^2+1 个矩阵必线性相关: $A^{n^2},A^{n^2-1},\cdots,A,I_n$.

也就是说, 存在 $\mathbb{K}$ 中不全为零的数 $c_i(i=0,1,2,\cdots,c_{n^2})$, 使

$$c_{n^2}A^{n^2}+c_{n^2-1}A^{n^2-1}+\cdots+c_1A+c_0I_n=O.$$

这表明矩阵 A 适合数域 $\mathbb{K}$ 上的一个非零多项式. □

定义 1.16 (矩阵的极小多项式)

若 n 阶矩阵 A (或 n 维线性空间 V 上的线性变换 φ) 适合一个非零首一多项式 $m(x)$, 且 $m(x)$ 是 A (或 φ) 所适合的非零多项式中次数最小者, 则称 $m(x)$ 是 A (或 φ) 的一个极小多项式或最小多项式.

注 由命题 1.39可知矩阵 A 的极小多项式 $m(x)$ 一定存在, 故极小多项式是良定义的.

引理 1.4 (矩阵极小多项式的基本性质)

若 $f(x)$ 是 A 适合的一个多项式, 则 A 的极小多项式 $m(x)$ 整除 $f(x)$.

证明 由多项式的带余除法知道

$$f(x)=m(x)q(x)+r(x),$$

且 $\deg r(x)<\deg m(x)$. 将 $x=A$ 代入上式得 $r(A)=O$, 若 $r(x)\neq 0$, 则 A 适合一个比 $m(x)$ 次数更小的非零多项式, 矛盾. 故 $r(x)=0$, 即 $m(x)\mid f(x)$. □

命题 1.40 (矩阵的极小多项式必唯一)

任一 n 阶矩阵的极小多项式必唯一.

证明 若 $m(x),g(x)$ 都是矩阵 A 的极小多项式, 则由矩阵极小多项式的基本性质知道 $m(x)$ 能够整除 $g(x)$, $g(x)$ 也能够整除 $m(x)$. 因此 $m(x)$ 与 $g(x)$ 只差一个常数因子, 又极小多项式必须首项系数为 1, 故 $g(x)=m(x)$. □

命题 1.41 (相似的矩阵具有相同的极小多项式)

相似的矩阵具有相同的极小多项式.

证明 设矩阵 A 和 B 相似, 即存在可逆矩阵 P, 使 $B=P^{-1}AP$. 设 A,B 的极小多项式分别为 $m(x),g(x)$, 注意到

$$m(B)=m(P^{-1}AP)=P^{-1}m(A)P=O,$$

因此 $g(x)\mid m(x)$. 同理, $m(x)\mid g(x)$, 故 $m(x)=g(x)$. □

命题 1.42

设 A 是一个分块对角阵

$$A=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_k \end{pmatrix},$$

其中 A_i 都是方阵, 则 A 的极小多项式等于诸 A_i 的极小多项式之最小公倍式.

证明 设 A 的极小多项式为 $m(x)$, A_i 的极小多项式为 $m_i(x)$, 诸 $m_i(x)$ 的最小公倍式为 $g(x)$, 则 $g(A_i)=O$, 于是

$$g(A)=\begin{pmatrix} g(A_1) & & & \\ & g(A_2) & & \\ & & \ddots & \\ & & & g(A_k) \end{pmatrix}=O,$$

从而 $m(x)\mid g(x)$. 又因为

$$m(A)=\begin{pmatrix} m(A_1) & & & \\ & m(A_2) & & \\ & & \ddots & \\ & & & m(A_k) \end{pmatrix}=O,$$

从而 $m(x)\mid g(x)$. 又因为

$$m(A)=\begin{pmatrix} m(A_1) & & & \\ & m(A_2) & & \\ & & \ddots & \\ & & & m(A_k) \end{pmatrix}=O,$$

所以对每个 i 有 $m(A_i)=O$, 从而 $m_i(x)\mid m(x)$, 即 $m(x)$ 是 $m_i(x)$ 的公倍式. 又 $g(x)$ 是诸 $m_i(x)$ 的最小公倍式, 故 $g(x)\mid m(x)$. 综上所述,$m(x)=g(x)$. □

命题 1.43

设 $m(x)$ 是 n 阶矩阵 A 的极小多项式, λ_0 是 A 的特征值, 则 $(x-\lambda_0)\mid m(x)$.

笔记 这个命题告诉我们: 矩阵的特征值一定是其极小多项式的根.

证明 由 $m(A)=O$ 及 命题 1.7可得 $m(\lambda_0)=0$, 故结论成立. □

定理 1.13 (Cayley-Hamilton 定理)

1. **代数形式:** 设 A 是数域 $\mathbb{K}$ 上的 n 阶矩阵, $f(x)$ 是 A 的特征多项式, 则 $f(A)=O$.
2. **几何形式:** 设 φ 是 n 维线性空间 V 上的线性变换, $f(x)$ 是 φ 的特征多项式, 则 $f(\varphi)=O$.

证明

1. **代数形式:** 因为复数域是最大数域, 所以可将 A 看作一个复矩阵. 由复方阵必相似于上三角阵知 A 复相似于一个上三角阵, 也就是说存在的可逆矩阵 P, 使 $P^{-1}AP=B$ 是一个上三角阵, 其中 P 与 B 都是复矩阵, 由相似矩阵有相同特征多项式可知 A 与 B 有相同的特征多项式 $f(x)$. 记

$$f(x)=x^n+a_1x^{n-1}+\cdots+a_n,$$

则 $f(B)=O$. 而

$$f(A)=A^n+a_1A^{n-1}+\cdots+a_nI_n$$

$$
\begin{aligned}
&= (PBP^{-1})^n + a_1(PBP^{-1})^{n-1} + \cdots + a_n I_n \\
&= PB^n P^{-1} + a_1 PB^{n-1}P^{-1} + \cdots + a_n I_n \\
&= P(B^n + a_1 B^{n-1} + \cdots + a_n I_n)P^{-1} \\
&= Pf(B)P^{-1} = O.
\end{aligned}
$$

2. **几何形式**: 设 $\{e_1, e_2, \cdots, e_n\}$ 是 V 的一组标准基, φ 在这组基下的矩阵为 A, 则由 $f(x)$ 是 φ 的特征多项式可知, $f(x)$ 也是 A 的特征多项式。从而由代数形式的结论可知 $f(A) = 0$。于是对 $\forall \alpha \in V$, 都存在 $k_1, k_2, \cdots, k_n$, 使得

$$
\alpha = k_1 e_1 + k_2 e_2 + \cdots + k_n e_n = (e_1, e_2, \cdots, e_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}.
$$

两边同时作用 φ 得到

$$
\varphi(\alpha) = k_1\varphi(e_1) + k_2\varphi(e_2) + \cdots + k_n\varphi(e_n) = (\varphi(e_1), \varphi(e_2), \cdots, \varphi(e_n))\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}
$$

$$
= (e_1, e_2, \cdots, e_n)A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A(e_1, e_2, \cdots, e_n)\begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} = A\alpha.
$$

因此 $f(\varphi)(\alpha) = f(A)(\alpha) = 0$。故由 α 的任意性可知 $f(\varphi) = O$.

□

推论 1.7

n 阶矩阵 A 的极小多项式是其特征多项式的因式. 特别, A 的极小多项式的次数不超过 n.

证明 由Cayley-Hamilton 定理及矩阵极小多项式的基本性质即得结论. □

推论 1.8

n 阶矩阵 A 的极小多项式和特征多项式有相同的根 (不计重数).

证明 由命题 1.43和推论 1.7即得结论. □