0.1 可对角化的判定

0.1.1 可对角化的基本知识

定义 0.1 (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 0.1 (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 0.2 (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 0.1

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。

□

定理 0.2 (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 0.1(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 0.1可知线性变换 φ 可对角化，从而再由引理 0.1(2) 可知矩阵 A 也可对角化.

若矩阵 A 可对角化，则由引理 0.1(2)可知线性变换 φ 也可对角化，从而由定理 0.1可知 φ 有 n 个线性无关的特征值，于是由引理 0.1(1)可知矩阵 A 也有相同的 n 个线性无关的特征值。 □

定理 0.3

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:

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

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

$$v=v_1+v_2+\cdots+v_{k-1}, \tag{2}$$

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

$$\varphi(v)=\varphi(v_1)+\varphi(v_2)+\cdots+\varphi(v_{k-1}). \tag{3}$$

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

$$\lambda_k v=\lambda_1v_1+\lambda_2v_2+\cdots+\lambda_{k-1}v_{k-1}. \tag{4}$$

在(3)式两边乘以 λ_k 减去(4)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 $\lambda_k-\lambda_i\ne0$, 从而 $v_i=0(i=1,2,\cdots,k-1)$. 这就证明了(1)式.

□

推论 0.1

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 0.3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 $\alpha_i\ne0$，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 $b_1\ne0$，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 $\alpha_i\ne0\ (1\leqslant i\leqslant k)$ 矛盾！ □

推论 0.2

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 0.1可知 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 一定线性无关。从而由定理 0.1可知, φ 一定可对角化。 □

定理 0.4 (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 0.1可知 φ 可对角化.

再证必要性. 设 φ 可对角化, 则由定理 0.1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 $\alpha = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$, 则 α 可写成 $V_1, V_2, \cdots, V_k$ 中向量之和, 因此由定理 0.3可知

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

□

定义 0.3 (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 0.4 (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 0.2 (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 0.5 (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 0.5 (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 0.2可知 $m_1+m_2+\cdots+m_k=n,t_i\le m_i,i=1,2,\cdots,k$.

由定理 0.4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 0.6

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

$$\begin{aligned}AP&=A\left(e_{11},\cdots,e_{1,n_1},\cdots,e_{r1},\cdots,e_{r,n_r}\right)=\left(\lambda_1e_{11},\cdots,\lambda_1e_{1,n_1},\cdots,\lambda_re_{r1},\cdots,\lambda_re_{r,n_r}\right)\\&=\left(e_{11},\cdots,e_{1,n_1},\cdots,e_{r1},\cdots,e_{r,n_r}\right)\operatorname{diag}\{A_1,A_2,\cdots,A_r\}=P\operatorname{diag}\{A_1,A_2,\cdots,A_r\}.\end{aligned}$$

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

定理 0.7 (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 0.1.
(2) 证明见定理 0.4.
(3)
(4) 证明见定理 0.5.
(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 0.1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.

0.1.2 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 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角矩阵.

命题 0.1 (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 $f(x)=a_1+a_2x+\cdots+a_nx^{n-1}, \omega_k=\cos\dfrac{2k\pi}{n}+\mathrm{i}\sin\dfrac{2k\pi}{n}(0\leq k\leq n-1)$, 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 0.1 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 0.4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 0.2 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 0.3 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. □

0.1.3 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 0.4 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 0.5 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 $f(\mu_0) \neq 0$, 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 0.6 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 0.6和例题**??**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 0.2

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**??**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 0.3

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 0.2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 $f(\lambda_i)=\mu_i(1\leq i\leq n)$. 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 0.4

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 0.5

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 0.4$f(A)$ 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 $f(\lambda_i)=\mu_i(1\leq i\leq n)$. 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 0.3

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 0.5即得本推论. □

命题 0.6

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 $\omega = \cos\dfrac{2\pi}{n+1} + \mathrm{i}\sin\dfrac{2\pi}{n+1}$ 为 1 的 $n+1$ 次方根, 则由 $x_1^{n+1} = x_2^{n+1}$ 可得 $x_1 = x_2\omega^k (1 \le k \le n)$. 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

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

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

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

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. □