0.1 可对角化的判断 (二)

0.1.1 极小多项式无重根

命题 0.1

证明:

(1) 若 $A^2=A$, 则 A 可对角化, 并且 A 的 Jordan 标准型为 $\mathrm{diag}\{1,\cdots,1,0\cdots,0\}$;

(2) 若 $A^k=I_n$, 则 A 可对角化, 并且 A 的 Jordan 标准型为 $\mathrm{diag}\{\omega_1,\omega_2,\cdots,\omega_n\}$, 其中 $\omega_i^k=1(1\leqslant i\leqslant n)$.

证明 (1) 矩阵 A 适合 $g(x)=x^2-x$ 且 $g(x)$ 无重根, 故由命题**??**可知 A 可对角化, 并且由命题**??**可知 A 的特征值也适合 $g(x)$, 故只能是 0, 1. 因此,A 的 Jordan 标准型为 $\mathrm{diag}\{1,\cdots,1,0\cdots,0\}$, 其中有 $\mathrm{r}(A)$ 个 1.

(2) 矩阵 A 适合 $g(x)=x^k-1$ 且 $g(x)$ 无重根, 故由命题**??**可知 A 可对角化, 并且由命题**??**可知 A 的特征值也适合 $g(x)$, 故只能是 1 的 k 次方根. 因此,A 的 Jordan 标准型为 $\mathrm{diag}\{\omega_1,\omega_2,\cdots,\omega_n\}$, 其中 $\omega_i^k=1(1\leqslant i\leqslant n)$. □

例题 0.1 设 A 是有理数域上的 n 阶矩阵, 其特征多项式的所有不可约因式为 $\lambda^2+\lambda+1,\lambda^2-2$. 又 A 的极小多项式是四次多项式, 求证:A 在复数域上可对角化.

证明 因为 A 的极小多项式 $m(\lambda)$ 和特征多项式 $f(\lambda)$ 有相同的根 (不计重数), 且 $\deg m(\lambda)=4$, 所以 $m(\lambda)=(\lambda^2+\lambda+1)(\lambda^2-2)$. 注意到 $m(\lambda)$ 在复数域内无重根, 故 A 在复数域上可对角化. □

命题 0.2

设 φ 是复线性空间 V 上的线性变换,V_0 是 φ 的不变子空间. 求证: 若 φ 可对角化, 则 φ 在 V_0 上的限制变换和 φ 在 V/V_0 上的诱导变换都可对角化.

证明 证法一:由命题**??**的几何版本可知, 限制变换 $\varphi|_{V_0}$ 和诱导变换 $\overline{\varphi}$ 都有完全的特征向量系, 从而可对角化.

证法二:设线性变换 φ、限制变换 $\varphi|_{V_0}$ 和诱导变换 $\overline{\varphi}$ 的极小多项式分别为 $m(\lambda),g(\lambda)$ 和 $h(\lambda)$, 则由由命题**??**容易验证 $\varphi|_{V_0}$ 和 $\overline{\varphi}$ 的表示矩阵都适合多项式 $m(x)$, 于是 $\varphi|_{V_0}$ 和 $\overline{\varphi}$ 都适合多项式 $m(\lambda)$, 从而 $g(\lambda)\mid m(\lambda)$ 且 $h(\lambda)\mid m(\lambda)$. 由于 φ 可对角化, 故 $m(\lambda)$ 无重根, 从而 $g(\lambda),h(\lambda)$ 也无重根, 于是 $\varphi|_{V_0}$ 和 $\overline{\varphi}$ 都可对角化. □

命题 0.3

设 φ 是 n 维复线性空间 V 上的线性变换, 求证:φ 可对角化的充要条件是对任一 φ-不变子空间 U, 均存在 φ-不变子空间 W, 使得 $V=U\oplus W$. 这样的 W 称为 U 的 **φ-不变补空间**.

证明 证法一: **先证充分性**: 假设 φ 不能对角化, 则 φ 只有 m 个线性无关的特征向量, 其中 $1\leqslant m<n$. 设由这些特征向量张成的子空间为 U, 由条件可知,U 存在非零的 φ-不变补空间 W. 考虑限制变换 $\varphi|_W$, 它在 W 上必存在特征值和特征向量, 这些也是 φ 的特征值和特征向量, 于是 φ 有多于 m 个线性无关的特征向量, 矛盾!

再证必要性: 设 φ 可对角化,U 是 φ-不变子空间, 则由命题 0.2可知,$\varphi|_U$ 仍可对角化, 故存在 U 的一组基 $\alpha_1,\cdots,\alpha_r$, 它们是 $\varphi|_U$, 也是 φ 的线性无关的特征向量. 又因为 φ 可对角化, 故存在 n 个线性无关的特征向量 $\{e_1,e_2,\cdots,e_n\}$, 再由基扩张定理可知, 可从这组基中取出 $n-r$ 个向量和 $\alpha_1,\cdots,\alpha_r$ 一起组成 V 的一组基. 设这 $n-r$ 个向量张成的子空间为 W, 则 W 是 U 的 φ-不变补空间.

证法二: **必要性**: 设线性变换 $\varphi:V\to V$ 可对角化, 则

$$V=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s},\tag{1}$$

这里 $V_{\lambda_i},i=1,2,\cdots,s$ 是互不相同的属于 φ 的特征值 λ_i 的特征子空间.

对 φ 的任一不变子空间 $W\subset V$, 我们断言

$$W=(W\bigcap V_{\lambda_1})\oplus(W\bigcap V_{\lambda_2})\oplus\cdots\oplus(W\bigcap V_{\lambda_s}).\tag{2}$$

事实上, 显然有上式右边是直和 (零的分解唯一), 并且

$$W\supset(W\bigcap V_{\lambda_1})\oplus(W\bigcap V_{\lambda_2})\oplus\cdots\oplus(W\bigcap V_{\lambda_s}).\tag{3}$$

对 $x \in W$, 由(1)得 $x_i \in V_{\lambda_i}, i=1,2,\cdots,s$ 使得 $x=\sum_{i=1}^{s} x_i$. 由中国剩余定理, 我们知道存在多项式 $p_j \in \mathbb{C}[x]$ 使得

$$p_j(x) \equiv \delta_{ij}\ (\bmod\ (x-\lambda_i)), i,j=1,2,\cdots,n. \tag{4}$$

这里 δ_{ij} 是 Kronecker 符号. 由(4)得

$$x_j=\sum_{i=1}^{s}\left[k_j\left(\varphi\left(x_i\right)-\lambda_i x_i\right)+\delta_{ij} I\left(x_i\right)\right]=\sum_{i=1}^{s} p_j(\varphi) x_i=p_j(\varphi) x \in W, j=1,2,\cdots,s.$$

现在我们由 $x_j \in W \bigcap V_{\lambda_j}, j=1,2,\cdots,s$ 知

$$W \subset (W \bigcap V_{\lambda_1}) \oplus (W \bigcap V_{\lambda_2}) \oplus \cdots \oplus (W \bigcap V_{\lambda_s}). \tag{5}$$

结合(3)和(5)得(2).

对每一个 $i=1,2,\cdots,s$, 取 $W \bigcap V_{\lambda_i}$ 在 V_{λ_i} 的补空间 U_i. 注意到 $\varphi|_{V_{\lambda_i}}$ 是数量矩阵, 于是 U_i 是 φ 不变子空间. 考虑

$$U \triangleq U_1 \oplus U_2 \oplus \cdots \oplus U_s.$$

显然 U 是 φ 不变子空间且 $V=W \oplus U$.

充分性: 若 φ 的任一不变子空间 $W \subset V$, 都有 φ 的不变补空间 U, 我们来对矩阵阶数归纳证明 φ 可对角化.

当 $n=1$, 这没什么好证的, 假设命题对 $n-1$ 成立, 则对 n 时, 首先设 V_λ 是 φ 的属于特征值 $\lambda \in \mathbb{C}$ 的特征子空间, 则若 $V=V_\lambda$, 已经有 φ 是数量矩阵.

若 $V \neq V_\lambda$, 则存在 φ 的不变补空间 U_λ 使得 $V=V_\lambda \oplus U_\lambda$. 现在对任何 φ 的不变子空间 $W \subset U_\lambda$, 取 φ 的不变补空间 U_W 使得 $V=W \oplus U_W$.

注意到 $U_W \bigcap U_\lambda$ 是 φ 的不变子空间. 显然 $W+U_W \bigcap U_\lambda$ 是直和 (零的分解唯一), 且 $U_\lambda \supset W \oplus (U_W \bigcap U_\lambda)$.

又注意到 $\varphi|_{U_\lambda}$ 没有特征值 λ, 所以其极小多项式 $m_1(x)$ 和 $\varphi|_{V_\lambda}$ 极小多项式 $m_2(x)$ 互素. 由中国剩余定理知, 存在多项式 $p_j \in \mathbb{C}[x]$, 使得

$$p_j(x) \equiv \delta_{ij}\ (\bmod\ m_i(x)), i,j=1,2,$$

其中 δ_{ij} 是 Kronecker 符号. 对 $\forall x \in U_\lambda$, 由 $V=W \oplus U_W$ 知, 存在 $x_1 \in W, x_2 \in U_W$, 使得 $x=x_1+x_2$. 于是

$$x_j=\sum_{i=1}^{s}\left[k_j m_i(\varphi)\left(x_i\right)+\delta_{ij} I\left(x_i\right)\right]=\sum_{i=1}^{s} p_j(\varphi) x_i=p_j(\varphi) x \in U_\lambda,\ \ j=1,2.$$

因此 $U_\lambda \subset W \oplus (U_W \bigcap U_\lambda)$. 故

$$U_\lambda=W \oplus (U_W \bigcap U_\lambda). \tag{6}$$

由(6)得 $\varphi|_{U_\lambda}$ 满足归纳假设, 因此 $\varphi|_{U_\lambda}$ 可对角化, 又 φ 在 V_λ 上可对角化, 故 φ 在 V 上可对角化. 我们完成了证明. □

命题 0.4

设 n 阶矩阵 A 的极小多项式 $m(\lambda)$ 的次数为 s,$B=(b_{ij})$ 为 s 阶矩阵, 其中 $b_{ij}=\operatorname{tr}(A^{i+j-2})$(约定 $b_{11}=n$), 求证:A 可对角化的充要条件是 B 为可逆矩阵.

注 本题主要利用的方法是: 设矩阵 A 的全体不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_k$, 定义

$$g(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$$

若 A 可对角化, 则 A 的极小多项式就是 $g(\lambda)$(参考常见矩阵的极小多项式 (2)). 反之, 若 A 适合多项式 $g(\lambda)$, 则由极小多项式的性质可知,$g(\lambda)$ 就是 A 的极小多项式. 特别地, 由于 $g(\lambda)$ 无重根, 故 A 可对角化.

证明 设 A 的全体不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_k$, 其代数重数分别为 $m_1,m_2,\cdots,m_k$, 则由命题??可知 A^i 的全体特征值为 $\lambda_1^i,\lambda_2^i,\cdots,\lambda_k^i$, 其代数重数分别为 $m_1,m_2,\cdots,m_k$. 从而 $\operatorname{tr}(A^i)=m_1\lambda_1^i+m_2\lambda_2^i+\cdots+m_k\lambda_k^i$. 定义 $g(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$, 则 $g(\lambda) \mid m(\lambda)$, 从而 $s \geqslant k$. 若 A 可对角化, 则 $m(\lambda)=g(\lambda)$, 从而 $s=k$. 若 A 不可对角

化, 则 $m(\lambda)$ 有重根, 从而 $s>k$. 考虑矩阵 B 的如下分解:

$$B=\begin{pmatrix} m_1 & m_2 & \cdots & m_k \\ m_1\lambda_1 & m_2\lambda_2 & \cdots & m_k\lambda_k \\ \vdots & \vdots & & \vdots \\ m_1\lambda_1^{s-1} & m_2\lambda_2^{s-1} & \cdots & m_k\lambda_k^{s-1} \end{pmatrix}\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{s-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{s-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{s-1} \end{pmatrix}$$

其中上式右边第一个矩阵是 $s\times k$ 矩阵, 第二个矩阵是 $k\times s$ 矩阵.

必要性: 若 A 可对角化, 则由上述分析可知 $s=k$, 则由 Vandermonde 行列式可知

$$|B|=m_1m_2\cdots m_k\prod_{1\leqslant i<j\leqslant k}(\lambda_i-\lambda_j)^2\neq 0$$

即 B 是可逆矩阵.

充分性: 若 B 可逆, 反证, 设 A 不可对角化, 则由上述分析可知 $s>k$, 则由 Cauchy - Binet 公式可得 $|B|=0$, 即 B 不可逆, 矛盾! □

0.1.2 初等因子都是一次多项式, 或 Jordan 块都是一阶矩阵

回顾推论**??**中可对角化的充要条件.

命题 0.5

设 n 阶复方阵 A 的特征多项式为 $f(\lambda)$, 复系数多项式 $g(\lambda)$ 满足 $(f(\lambda),g'(\lambda))=1$. 证明:$A$ 可对角化的充要条件是 $g(A)$ 可对角化.

证明 必要性显然成立, 下证充分性. 用反证法, 设 A 不可对角化, 则存在可逆矩阵 P, 使得

$$P^{-1}AP=J=\mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$$

为 Jordan 标准型, 其中 $r_1>1$. 注意到

$$P^{-1}g(A)P=g(P^{-1}AP)=g(J)=\mathrm{diag}\{g(J_{r_1}(\lambda_1)),\cdots,g(J_{r_k}(\lambda_k))\}$$

由命题**??**(5) 可知

$$g(J_{r_1}(\lambda_1))=\begin{pmatrix} g(\lambda_1) & g'(\lambda_1) & \cdots & * \\ & g(\lambda_1) & \ddots & \vdots \\ & & \ddots & g'(\lambda_1) \\ & & & g(\lambda_1) \end{pmatrix}$$

由 $(f(\lambda),g'(\lambda))=1$ 可知 $g'(\lambda_1)\neq 0$, 于是 $g(J_{r_1}(\lambda_1))$ 的特征值全为 $g(\lambda_1)$, 其几何重数为 $r_1-\mathrm{r}(g(J_{r_1}(\lambda_1))-g(\lambda_1)I_{r_1})=1$, 因此 $g(J_{r_1}(\lambda_1))$ 的 Jordan 标准型为 $J_{r_1}(g(\lambda_1))$, 其阶数 $r_1>1$. 由于 $J_{r_1}(g(\lambda_1))$ 也是 $g(A)$ 的一个 Jordan 块, 故 $g(A)$ 不可对角化, 矛盾! □

命题 0.6

设 φ 是 n 维复线性空间 V 上的线性变换, 求证:φ 可对角化的充要条件是对 φ 的任一特征值 λ_0, 总有 $\mathrm{Ker}(\varphi-\lambda_0I_V)\cap\mathrm{Im}(\varphi-\lambda_0I_V)=0$.

注 这个命题 0.6是这个命题 0.8的特例.

证明 **先证必要性**: 若 φ 可对角化, 则存在一组基 $\{e_1,e_2,\cdots,e_n\}$, 使得 φ 在这组基下的表示矩阵为 $\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$. 适当调整基向量的顺序, 不妨设 $\lambda_0=\lambda_1=\cdots=\lambda_r,\lambda_0\neq\lambda_j(j>r)$, 则 $\varphi-\lambda_0I_V$ 在基 $\{e_1,e_2,\cdots,e_n\}$ 下的表示矩阵为 $\mathrm{diag}\{0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n\}$.

于是对 $\forall v\in V$, 都存在非零列向量 $(a_1,a_2,\cdots,a_n)'$, 使得

$$v=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(e_1,e_2,\cdots,e_n).$$

从而

$$(\varphi-\lambda_0I_V)(v)=\mathrm{diag}\{0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n\}\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}(e_1,e_2,\cdots,e_n)=\sum_{i=r+1}^{n}a_ie_i\in L(e_{r+1},\cdots,e_n).$$

故 $\mathrm{Im}(\varphi-\lambda_0I_V)\subset L(e_{r+1},\cdots,e_n)$. 再任取 $\alpha\in L(e_{r+1},\cdots,e_n)$, 存在非零列向量 $(0,\cdots,0,b_{r+1},\cdots,b_n)'$, 使得

$$\alpha=\begin{pmatrix}0\\\vdots\\0\\b_{r+1}\\\vdots\\b_n\end{pmatrix}(e_1,\cdots,e_r,e_{r+1},\cdots,e_n).$$

从而

$$(\varphi-\lambda_0I_V)(\alpha)=\mathrm{diag}\{0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n\}\begin{pmatrix}0\\\vdots\\0\\b_{r+1}\\\vdots\\b_n\end{pmatrix}(e_1,\cdots,e_r,e_{r+1},\cdots,e_n)=\sum_{i=r+1}^{n}b_ie_i=\alpha.$$

故 $L(e_{r+1},\cdots,e_n)\subset\mathrm{Im}(\varphi-\lambda_0I_V)$. 综上,$\mathrm{Im}(\varphi-\lambda_0I_V)=L(e_{r+1},\cdots,e_n)$.

任取 $u\in L(e_1,\cdots,e_r)$, 则存在非零列向量 $(u_1,\cdots,u_r,0,\cdots,0)'$, 使得

$$u=\begin{pmatrix}u_1\\\vdots\\u_r\\0\\\vdots\\0\end{pmatrix}(e_1,\cdots,e_r,e_{r+1},\cdots,e_n).$$

从而

$$(\varphi-\lambda_0I_V)(u)=\mathrm{diag}\{0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n\}\begin{pmatrix}u_1\\\vdots\\u_r\\0\\\vdots\\0\end{pmatrix}(e_1,\cdots,e_r,e_{r+1},\cdots,e_n)=0.$$

于是 $L(e_1,\cdots,e_r)\subset \mathrm{Ker}(\varphi-\lambda_0 I_V)$. 再任取 $y\in\mathrm{Ker}(\varphi-\lambda_0 I_V)\subset V$, 则存在非零列向量 $(y_1,y_2,\cdots,y_n)'$, 使得

$$y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}(e_1,e_2,\cdots,e_n)=\sum_{i=1}^n y_ie_i.$$

从而

$$(\varphi-\lambda_0 I_V)(y)=\mathrm{diag}\{0,\cdots,0,\lambda_{r+1},\cdots,\lambda_n\}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}(e_1,e_2,\cdots,e_n)=\sum_{i=r+1}^n y_ie_i=0.$$

因此 $y_{r+1}=y_{r+2}=\cdots=y_n=0$, 故 $y=\sum\limits_{i=1}^n y_ie_i=\sum\limits_{i=1}^r y_ie_i\in L(e_1,\cdots,e_r)$. 进而 $\mathrm{Ker}(\varphi-\lambda_0 I_V)\subset L(e_1,\cdots,e_r)$. 综上,$\mathrm{Ker}(\varphi-\lambda_0 I_V)=L(e_1,\cdots,e_r)$. 由此可知 $\mathrm{Ker}(\varphi-\lambda_0 I_V)=L(e_1,\cdots,e_r)$,$\mathrm{Im}(\varphi-\lambda_0 I_V)=L(e_{r+1},\cdots,e_n)$, 从而 $\mathrm{Ker}(\varphi-\lambda_0 I_V)\cap\mathrm{Im}(\varphi-\lambda_0 I_V)=0$.

再证充分性: 用反证法, 设 φ 不可对角化, 则存在 V 的一组基 $\{e_1,e_2\cdots,e_n\}$, 使得 φ 在这组基下的表示矩阵为 Jordan 标准型 $J=\mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$, 其中 $r_1>1$. 由表示矩阵的定义可得 $\varphi(e_1)=\lambda_1e_1$,$\varphi(e_2)=e_1+\lambda_1e_2$, 于是 $(\varphi-\lambda_1 I_V)(e_1)=0$,$(\varphi-\lambda_1 I_V)(e_2)=e_1$, 从而 $\mathbf{0}\neq e_1\in\mathrm{Ker}(\varphi-\lambda_1 I_V)\cap\mathrm{Im}(\varphi-\lambda_1 I_V)$, 这与假设矛盾. □

命题 0.7

求证:n 阶复矩阵 A 可对角化的充要条件是对 A 的任一特征值 λ_0,$(\lambda_0 I_n-A)^2$ 和 $\lambda_0 I_n-A$ 的秩相同.

注 这个命题 0.7是这个命题 0.8的特例.

证明 **先证必要性**: 若 A 可对角化, 则存在可逆矩阵 P, 使得 $P^{-1}AP=\Lambda=\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$. 适当调整 P 的列向量的顺序, 不妨设 $\lambda_0=\lambda_1=\cdots=\lambda_r$,$\lambda_0\neq\lambda_j(j>r)$, 则由于相似矩阵的特征多项式相同可知,$\lambda_0 I_n-A=\lambda_0 I_n-\Lambda$. 从而 $\mathrm{r}(\lambda_0 I_n-A)=\mathrm{r}(\lambda_0 I_n-\Lambda)=n-r$,$\mathrm{r}((\lambda_0 I_n-A)^2)=\mathrm{r}((\lambda_0 I_n-\Lambda)^2)=n-r$, 于是结论成立.

再证充分性: 用反证法, 若 A 不可对角化, 则存在可逆矩阵 P, 使得 $P^{-1}AP=J=\mathrm{diag}\{J_{r_1}(\lambda_1),\cdots,J_{r_k}(\lambda_k)\}$ 为 Jordan 标准型, 其中不妨设 $r_1>1$. 由于相似矩阵的特征多项式相等, 因此 $\lambda_1 I_n-A=\lambda_0 I_n-J$. 从而注意到

$$\mathrm{r}((\lambda_1 I_n-A)^j)=\mathrm{r}((\lambda_1 I_n-J)^j)=\sum_{i=1}^k \mathrm{r}((\lambda_1 I_{r_i}-J_{r_i}(\lambda_i))^j),\ j\geqslant 1$$

又注意到 $\lambda_1 I_{r_1}-J_{r_1}(\lambda_1)$ 是 r_1 阶基础幂零阵, 故 $\mathrm{r}(\lambda_1 I_{r_1}-J_{r_1}(\lambda_1))=r_1-1$,$\mathrm{r}((\lambda_1 I_{r_1}-J_{r_1}(\lambda_1))^2)=r_1-2$, 因此 $\mathrm{r}((\lambda_1 I_n-A)^2)<\mathrm{r}(\lambda_1 I_n-A)$, 这与假设矛盾. □

命题 0.8

设 φ 是 n 维复线性空间 V 上的线性变换, 求证: φ 可对角化的充要条件是对 φ 的任一特征值 λ_0, 下列条件之一成立:

(1) $V=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)+\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)$;

(2) $V=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)\oplus\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)$;

(3) $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)\cap\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)=0$;

(4) $\dim\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\dim\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2$;

(5) $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^3=\cdots$;

(6) $\mathrm{r}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{r}((\varphi-\lambda_0\boldsymbol{I}_V)^2)$;

(7) $\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)^3=\cdots$;

(8) $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)$ 存在 φ-不变补空间, 即存在 φ-不变子空间 U, 使得 $V=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)\oplus U$;

(9) $\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)$ 存在 φ-不变补空间, 即存在 φ-不变子空间 W, 使得 $V=\mathrm{Im}(\varphi-\lambda_0\boldsymbol{I}_V)\oplus W$.

注 命题 0.6与命题 0.7都是这个命题 0.8的特例.

笔记 由命题??可知条件 (1)~(9) 是相互等价的, 因此本题的结论由命题 0.6(与条件 (3) 对应) 或命题 0.7(与条件 (6) 对应) 即得. 事实上, 对充分性而言, 我们还可以从其他条件出发来证明 φ 可对角化, 下面是 3 种证法.

证明 证法一:对任一特征值 λ_0, 由 $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\cdots=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^n$, 取维数之后可得特征值 λ_0 的几何重数等于代数重数, 从而 φ 有完全的特征向量系, 于是 φ 可对角化.

证法二: 对任一特征值 λ_0, 由 $\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^2=\cdots=\mathrm{Ker}(\varphi-\lambda_0\boldsymbol{I}_V)^n$ 可知, 特征子空间等于根子空间, 再由根子空间的直和分解可知, 全空间等于特征子空间的直和, 从而 φ 可对角化.

证法三: 设 φ 的全体不同特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_k$, 特征多项式 $f(\lambda)=(\lambda-\lambda_1)^{m_1}(\lambda-\lambda_2)^{m_2}\cdots(\lambda-\lambda_k)^{m_k}$, 则对任意的 $\boldsymbol{\alpha}\in V$, 由 Cayley - Hamilton 定理可得

$$(\varphi-\lambda_1\boldsymbol{I}_V)^{m_1}(\varphi-\lambda_2\boldsymbol{I}_V)^{m_2}\cdots(\varphi-\lambda_k\boldsymbol{I}_V)^{m_k}(\boldsymbol{\alpha})=\boldsymbol{0},$$

即有 $(\varphi-\lambda_2\boldsymbol{I}_V)^{m_2}\cdots(\varphi-\lambda_k\boldsymbol{I}_V)^{m_k}(\boldsymbol{\alpha})\in\mathrm{Ker}(\varphi-\lambda_1\boldsymbol{I}_V)^{m_1}=\mathrm{Ker}(\varphi-\lambda_1\boldsymbol{I}_V)$, 从而

$$(\varphi-\lambda_1\boldsymbol{I}_V)(\varphi-\lambda_2\boldsymbol{I}_V)^{m_2}\cdots(\varphi-\lambda_k\boldsymbol{I}_V)^{m_k}(\boldsymbol{\alpha})=\boldsymbol{0}.$$

不断这样做下去, 最终可得对任意的 $\boldsymbol{\alpha}\in V$, 总有

$$(\varphi-\lambda_1\boldsymbol{I}_V)(\varphi-\lambda_2\boldsymbol{I}_V)\cdots(\varphi-\lambda_k\boldsymbol{I}_V)(\boldsymbol{\alpha})=\boldsymbol{0},$$

即 φ 适合多项式 $g(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_k)$, 从而 φ 可对角化. □

例题 0.2 若 $n(n\geqslant 2)$ 阶矩阵 B 相似于 $R=\mathrm{diag}\left\{\begin{pmatrix}0&1\\1&0\end{pmatrix},I_{n-2}\right\}$, 则称 B 为反射矩阵. 证明: 任一对合矩阵 A(即 $A^2=I_n$) 均可分解为至多 n 个两两乘法可交换的反射矩阵的乘积.

证明 由命题 0.1(2)可知, 对合矩阵 A 可对角化, 即存在可逆矩阵 P, 使得 $P^{-1}AP=\mathrm{diag}\{-I_r,I_{n-r}\}$, 其中 $0\leqslant r\leqslant n$. 当 $r=0$ 时,$A=I_n=R^2$, 结论成立. 当 $r\geqslant 1$ 时, 设 $B_i=P\mathrm{diag}\{1,\cdots,1,-1,1,\cdots,1\}P^{-1}$, 其中 -1 在主对角线上的第 i 个位置, 则 $B_i(1\leqslant i\leqslant r)$ 两两乘法可交换, 并且 $A=B_1B_2\cdots B_r$. 由于 $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ 的特征值是 $-1,1$, 故其相似于 $\mathrm{diag}\{-1,1\}$, 因此矩阵 B 是反射矩阵当且仅当 B 相似于 $\mathrm{diag}\{-1,1,\cdots,1\}$. 因为对角矩阵的两个主对角元素对换是一个相似变换, 所以上述 B_i 都是反射矩阵, 于是 A 可以分解为 r 个两两乘法可交换的反射矩阵的乘积. □