

0.1 有理标准型

命题 0.1

设矩阵 A 的特征矩阵 $\lambda I - A$ 的法式为

$$\text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\},$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, k-1$), 则 A 的不变因子就是

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda).$$



证明 由推论??可知, 矩阵 A 的特征矩阵 $\lambda I - A$ 的法式为

$$\text{diag}\{1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)\},$$

其中 $d_i(\lambda)$ 为非常数首一多项式且 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, 2, \dots, k-1$), 则根据不变因子的定义可知, A 的不变因子就是

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda).$$



引理 0.1 (Frobenius 块的基本性质)

设 r 阶矩阵

$$F = F(f(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_r & -a_{r-1} & -a_{r-2} & \cdots & -a_1 \end{pmatrix},$$

其中 $f(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r$, 则

- (1) $|F(f(\lambda))| = (-1)^{r+2}a_r = (-1)^{r+2}f(0)$.
- (2) $F = F(f(\lambda))$ 的特征多项式等于极小多项式等于 $f(\lambda)$.
- (3) F 的行列式因子为

$$1, \dots, 1, f(\lambda), \quad (1)$$

其中共有 $r-1$ 个 1, $f(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r$, F 的不变因子也由(1)式给出, F 的不变因子分别为:

$$1, \dots, 1, f(\lambda).$$

进而, $\lambda I - F$ 相抵于 $\text{diag}\{1, \dots, 1, f(\lambda)\}$.



注 $f(\lambda)$ 的友矩阵 $C(f(\lambda))$ (即 $F(f(\lambda))$ 的转置) 的性质与 $F(f(\lambda))$ 相同.

证明

- (1) 注意到 $f(0) = a_r$, 于是就有

$$|F(f(\lambda))| = \begin{vmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_r & -a_{r-1} & -a_{r-2} & \cdots & -a_1 \end{vmatrix} \xrightarrow{\text{按第一列展开}} (-1)^{r+2}a_r = (-1)^{r+2}f(0).$$

- (2) 由命题??(1) 同理可得

$$|\lambda I - F| = \lambda^r + a_1\lambda^{r-1} + \cdots + a_r = f(\lambda).$$

因为 F 的特征多项式为 $f(\lambda)$, 所以 F 适合多项式 $f(\lambda)$. 设 $e_i (i = 1, 2, \dots, r)$ 是 r 维标准单位行向量, 则不难算出:

$$e_1 F = e_2, \quad e_1 F^2 = e_2 F = e_3, \quad \dots, \quad e_1 F^{r-1} = e_{r-1} F = e_r.$$

显然, $e_1, e_1 F, \dots, e_1 F^{r-1}$ 是一组线性无关的向量, 从而任取 $g(x) \in P_{r-1}[x]$ 且 $g(x)$ 非零, 则存在一组不全为零的数 a_1, a_2, \dots, a_r , 使得

$$g(x) = a_1 x^{r-1} + a_2 x^{r-2} + \dots + a_r.$$

于是将 F 代入上式, 再在等式两边同乘 e_1 得到

$$e_1 g(F) = a_1 e_1 F^{r-1} + a_2 e_1 F^{r-2} + \dots + a_r e_1 F.$$

又因为 $e_1, e_1 F, \dots, e_1 F^{r-1}$ 是一组线性无关的向量, 且 a_1, a_2, \dots, a_r 不全为零, 所以 $e_1 g(F) \neq 0$. 即 $g(F)$ 的第一行不为零, 故 $g(F) \neq O$. 因此 F 不可能适合一个次数不超过 $r-1$ 的非零多项式, 从而 F 的极小多项式就是 $f(\lambda)$.

(3) F 的 r 阶行列式因子就是它的特征多项式, 由命题??(1) 同理可得

$$|\lambda I - F| = \lambda^r + a_1 \lambda^{r-1} + \dots + a_r = f(\lambda).$$

对任一 $k < r$, $\lambda I - F$ 总有一个 k 阶子式其值等于 $(-1)^k$, 故 $D_k(\lambda) = 1$. 又由推论??可知, $\lambda I - F$ 的法式为 $\text{diag}\{1, \dots, 1, f(\lambda)\}$. 故 $\lambda I - F$ 相抵于 $\text{diag}\{1, \dots, 1, f(\lambda)\}$.

□

引理 0.2

设 λ -矩阵 $A(\lambda)$ 相抵于对角 λ -矩阵

$$\text{diag}\{d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)\}, \quad (2)$$

λ -矩阵 $B(\lambda)$ 相抵于对角 λ -矩阵

$$\text{diag}\{d'_1(\lambda), d'_2(\lambda), \dots, d'_n(\lambda)\}, \quad (3)$$

且 $d'_1(\lambda), d'_2(\lambda), \dots, d'_n(\lambda)$ 是 $d_1(\lambda), d_2(\lambda), \dots, d_n(\lambda)$ 的一个置换 (即若不计次序, 这两组多项式完全相同), 则 $A(\lambda)$ 相抵于 $B(\lambda)$.

♡

证明 利用初等行对换及初等列对换即可将(2)式变成(3)式, 因此(2)式所示的矩阵与(3)式所示的矩阵相抵, 从而 $A(\lambda)$ 与 $B(\lambda)$ 相抵. □

定理 0.1 (有理标准型/Frobenius 标准型)

设 A 是数域 \mathbb{F} 上的 n 阶方阵, A 的不变因子组为

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda),$$

其中 $\deg d_i(\lambda) = m_i \geq 1$, 则 A 相似于下列分块对角阵:

$$F = \begin{pmatrix} F_1 & & & \\ & F_2 & & \\ & & \ddots & \\ & & & F_k \end{pmatrix}, \quad (4)$$

其中 F_i 的阶等于 m_i , 且 F_i 是形如引理 0.1 中的矩阵, F_i 的最后一行由 $d_i(\lambda)$ 的系数 (除首项系数之外) 的负

值组成. 此即, 设 $d_i = \lambda^{m_i} + a_{1i}\lambda^{m_i-1} + \cdots + a_{m_i,i}$, 则

$$F_i = F(d_i(\lambda)) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{m_i,i} & -a_{m_i-1,i} & -a_{m_i-2,i} & \cdots & -a_{1i} \end{pmatrix}.$$

(4)式称为矩阵 A 的**有理标准型**或**Frobenius 标准型**, 每个 F_i 称为**Frobenius 块**.

进而, A 也相似于下列分块对角阵:

$$C = \begin{pmatrix} C_1 & & \\ & C_2 & \\ & & \ddots \\ & & & C_k \end{pmatrix},$$

其中 C_i 的阶等于 m_i , C_i 就是上述 F_i 的转置, 即

$$C_i = C(d_i(\lambda)) = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_{m_i,i} \\ 1 & 0 & \cdots & 0 & -a_{m_i-1,i} \\ 0 & 1 & \cdots & 0 & -a_{m_i-2,i} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{1i} \end{pmatrix}.$$



证明 注意到 $\lambda I - A$ 的第 n 个行列式因子就是 A 的特征多项式 $|\lambda I - A|$, 再由不变因子的定义可知:

$$|\lambda I - A| = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda).$$

而 $|\lambda I - A|$ 是一个 n 次多项式, 因此 $m_1 + m_2 + \cdots + m_k = n$. 一方面, $\lambda I - A$ 的法式为

$$\text{diag}\{1, \cdots, 1, d_1(\lambda), d_2(\lambda), \cdots, d_k(\lambda)\},$$

其中有 $n - k$ 个 1. 另一方面, 对 $\lambda I - F$ 的每个分块都施以 λ -矩阵的初等变换, 由引理 0.1 可知, $\lambda I - F$ 相抵于如下对角阵:

$$\text{diag}\{1, \cdots, 1, d_1(\lambda); 1, \cdots, 1, d_2(\lambda); \cdots; 1, \cdots, 1, d_k(\lambda)\}, \quad (5)$$

其中每个 $d_i(\lambda)$ 前各有 $m_i - 1$ 个 1, 从而共有 $\sum_{i=1}^k (m_i - 1) = n - k$ 个 1. 因此 (5) 式所示的矩阵与 $\lambda I - A$ 的法式只相差主对角线上元素的置换, 由引理 0.2 可得 $\lambda I - A$ 与 $\lambda I - F$ 相抵, 从而 A 与 F 相似.

又因为矩阵与其自身的转置相似, 所以矩阵 A 也相似于 F 的转置, 即 C . □

例题 0.1 设 6 阶矩阵 A 的不变因子为

$$1, 1, 1, \lambda - 1, (\lambda - 1)^2, (\lambda - 1)^2(\lambda + 1),$$


则 A 的有理标准型为

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

定理 0.2 (极小多项式就是幂次最大的不变因子)

设数域 \mathbb{K} 上的 n 阶矩阵 A 的不变因子为

$$1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda),$$

其中 $d_i(\lambda) \mid d_{i+1}(\lambda)$ ($i = 1, \dots, k-1$), 则 A 的特征多项式为 $d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda)$, 极小多项式为 $m(\lambda) = d_k(\lambda)$. 

证明 首先证明特征多项式, 根据不变因子和行列式因子的定义可知


$$1 \cdot 1 \cdot d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda) = D_n(\lambda).$$

其中 $D_n(\lambda)$ 为 $\lambda I_n - A$ 的 n 阶行列式因子, 即 A 的特征多项式 $|\lambda I_n - A|$. 因此

$$|\lambda I_n - A| = d_1(\lambda)d_2(\lambda)\cdots d_k(\lambda).$$

然后证明极小多项式, 设 A 的有理标准型为

$$F = \begin{pmatrix} F_1 & & \\ & F_2 & \\ & & \ddots \\ & & & F_k \end{pmatrix}.$$


因为相似矩阵有相同的极小多项式, 故只需证明 F 的极小多项式是 $d_k(\lambda)$ 即可. 但 F 是分块对角阵, 由极小多项式的性质 (6) 知 F 的极小多项式是诸 F_i 极小多项式的最小公倍式. 又由引理 0.1 知 F_i 的极小多项式为 $d_i(\lambda)$. 因为 $d_i(\lambda) \mid d_{i+1}(\lambda)$, 故诸 $d_i(\lambda)$ 的最小公倍式等于 $d_k(\lambda)$. 

例题 0.2 下面两个 4 阶矩阵


$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

的不变因子分别为 $A: 1, \lambda, \lambda, \lambda^2$ 和 $B: 1, 1, \lambda^2, \lambda^2$. 它们的特征多项式和极小多项式分别相等, 但它们不相似.

定义 0.1 (循环子空间)

设 V 是数域 \mathbb{K} 上的 n 维线性空间, φ 是 V 上的线性变换. 设 $0 \neq \alpha \in V$, 则 $U = L(\alpha, \varphi(\alpha), \varphi^2(\alpha), \dots)$ 称为 V 的**循环子空间**, 记为 $U = C(\varphi, \alpha)$, α 称为 U 的**循环向量**. 若 $U = V$, 则称 V 为**循环空间**. 

定理 0.3 (循环子空间的基本性质)

设 V 是数域 \mathbb{K} 上的 n 维线性空间, φ 是 V 上的线性变换, $0 \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 为循环子空间, 则循环子空间 U 是 V 的 φ -不变子空间, 并且是包含 α 的最小 φ -不变子空间. 

证明 U 是 V 的 φ -不变子空间是显然的. 下证 U 是包含 α 的最小 φ -不变子空间.

设 $\alpha \in W$, 且 W 为 φ -不变子空间, 则由数学归纳法易知


$$\alpha, \varphi^k(\alpha) \in W, \forall k \in \mathbb{N}_1.$$

于是

$$U = L(\alpha, \varphi(\alpha), \dots) \in W.$$



定理 0.4

设 V 是数域 \mathbb{K} 上的 n 维线性空间, φ 是 V 上的线性变换, $0 \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 为循环子空间, 若 $\dim U = r$, 求证: $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基. 

证明 设 $m = \max\{k \in \mathbb{Z}^+ \mid \alpha, \varphi(\alpha), \dots, \varphi^{k-1}(\alpha) \text{ 线性无关}\}$, 则显然 $1 \leq m \leq r$, 故 m 是良定义的. 于是由命题??和数学归纳法容易验证: 对任意的 $k \geq m, \varphi^k(\alpha)$ 都是 $\alpha, \varphi(\alpha), \dots, \varphi^{m-1}(\alpha)$ 的线性组合, 于是 $\{\alpha, \varphi(\alpha), \dots, \varphi^{m-1}(\alpha)\}$ 是 U 的一组基, 从而 $m = \dim U = r$. \square

定理 0.5

设 U 是 V 的 φ -不变子空间, 求证: U 为循环子空间的充要条件是 $\varphi|_U$ 在 U 的某组基下的表示矩阵为某个首一多项式的友阵.

♡

证明 充分性: 设 $\varphi|_U$ 在 U 的一组基 $\{e_1, e_2, \dots, e_r\}$ 下的表示矩阵是友阵 $C(d(\lambda))$, 其中 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \dots + a_{r-1}\lambda + a_r$, 则由友阵的定义可知 $\varphi(e_i) = e_{i+1} (1 \leq i \leq r-1), \varphi(e_r) = -\sum_{i=1}^r a_{r-i+1}e_i$. 因此 $e_i = \varphi^{i-1}(e_1) (2 \leq i \leq r), U = L(e_1, e_2, \dots, e_r) = C(\varphi, e_1)$ 为循环子空间.

必要性: 设 $U = C(\varphi, \alpha)$ 是 r 维循环子空间, 则由定理 0.4 可知, $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基. 设

$$\varphi^r(\alpha) = -a_r\alpha - a_{r-1}\varphi(\alpha) - \dots - a_1\varphi^{r-1}(\alpha)$$


令 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \dots + a_{r-1}\lambda + a_r$, 容易验证: $\varphi|_U$ 在基 $\{\alpha, \varphi(\alpha), \dots, \varphi^{r-1}(\alpha)\}$ 下的表示矩阵就是友阵 $C(d(\lambda))$. \square

定理 0.6 (有理标准型的几何意义)

设 V 是数域 \mathbb{K} 上的 n 维线性空间, φ 是 V 上的线性变换, 且 φ 的不变因子组是 $1, \dots, 1, d_1(\lambda), \dots, d_k(\lambda)$, 其中 $d_i(\lambda)$ 是非常数首一多项式, $d_i(\lambda) \mid d_{i+1}(\lambda) (1 \leq i \leq k-1)$, 则 V 存在一个循环子空间的直和分解:

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \dots \oplus C(\varphi, \alpha_k) \quad (6)$$

使得 $\varphi|_{C(\varphi, \alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(d_i(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. \square

 **笔记** 线性变换 φ 的有理标准型诱导的 V 的上述循环子空间直和分解 (6) 就是有理标准型的几何意义.

证明 由定理 0.1 可知, 存在 V 的一组基 $\{e_1, e_2, \dots, e_n\}$, 使得 φ 在这组基下的表示矩阵为

$$C = \text{diag}\{C(d_1(\lambda)), C(d_2(\lambda)), \dots, C(d_k(\lambda))\}$$

其中 $\varphi|_{L(e_{i1}, \dots, e_{ir_i})}$ 的表示矩阵就是友阵 $C(d_i(\lambda)), i = 1, 2, \dots, k$. 再结合定理 0.5 的讨论可知, $L(e_{i1}, \dots, e_{ir_i})$ 就是一个循环子空间. 于是任取 $\alpha_i \in L(e_{i1}, \dots, e_{ir_i})$ 作为循环向量, 则

$$C(\varphi, \alpha_i) = L(e_{i1}, \dots, e_{ir_i}) = L(\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i))$$

其中 $\dim C(\varphi, \alpha_i) = r_i$.

综上所述, 此时 V 存在一个循环子空间的直和分解:

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \dots \oplus C(\varphi, \alpha_k)$$

使得 $\varphi|_{C(\varphi, \alpha_i)}$ 在基 $\{\alpha_i, \varphi(\alpha_i), \dots, \varphi^{r_i-1}(\alpha_i)\}$ 下的表示矩阵就是友阵 $C(d_i(\lambda))$, 其中 $r_i = \dim C(\varphi, \alpha_i)$. \square

定理 0.7 (循环子空间的刻画)

设 φ 是数域 \mathbb{K} 上 n 维线性空间 V 上的线性变换, φ 的特征多项式和极小多项式分别为 $f(\lambda)$ 和 $m(\lambda)$, 证明以下几个结论等价:

- (1) φ 的行列式因子组或不变因子组为 $1, \dots, 1, f(\lambda)$;
- (2) φ 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \dots, P_k(\lambda)^{r_k}$, 其中 $P_i(\lambda)$ 是 \mathbb{K} 上互异的首一不可约多项式, $r_i \geq 1, 1 \leq i \leq k$;
- (3) φ 的极小多项式 $m(\lambda)$ 等于特征多项式 $f(\lambda)$;
- (4) V 是关于线性变换 φ 的循环空间;
- (5) 存在 $\alpha \in V$, 使得 $\{\alpha, A\alpha, \dots, A^{n-1}\alpha\}$ 为 V 一组基;
- (6) φ 的任一特征值 λ 的几何重数都是 1.

♡

证明 (1) \Leftrightarrow (2): 由不变因子和初等因子之间的相互转换即得.

(1) \Leftrightarrow (3): 由极小多项式等于最大的不变因子, 以及所有不变因子的乘积等于特征多项式即得.

(1) \Leftrightarrow (4): 若 V 是循环空间, 则由 **定理 0.5** 可知, φ 在某组基下的表示矩阵是友阵 $C(g(\lambda))$, 再由友阵的性质 (**引理 0.1**) 可知, φ 的行列式因子组和不变因子组均为 $1, \dots, 1, g(\lambda) = f(\lambda)$. 若 φ 的不变因子组为 $1, \dots, 1, f(\lambda)$, 则由有理标准型的几何意义 (**定理 0.6**) 可知, V 是循环空间.

(4) \Leftrightarrow (5): 由 **定理 0.4** 知必要性成立. 充分性由 **循环子空间的定义** 立得.

(3) \Leftrightarrow (6): 记 $\varphi = A$.

充分性: 设 A 的特征多项式等于极小多项式为 $f(x)$, 不妨设

$$f(x) = \prod_{i=1}^s (x - \lambda_i)^{n_i}, 1 \leq n_i \leq n.$$

则由命题??可知, n_i 是 A 的 Jordan 标准型中所有特征值为 λ_i 的 Jordan 块的最大阶数. 而 λ_i 的代数重数也是 n_i , 即所有特征值为 λ_i 的 Jordan 块的阶数之和为 n_i . 因此特征值为 λ_i 的 Jordan 块只有一个, 即 $\lambda_i (1 \leq i \leq s)$ 的几何重数为 1.

必要性: 设 A 的特征多项式为 $f(x)$, 极小多项式为 $m(x)$, 不妨设

$$f(x) = \prod_{i=1}^s (x - \lambda_i)^{n_i}, m(x) = \prod_{i=1}^s (x - \lambda_i)^{c_i}, 1 \leq n_i \leq n.$$

则由条件可知 A 的 Jordan 标准型中特征值为 λ_i 的 Jordan 块只有一个, 且阶数为 n_i . 又由命题??可知, c_i 是 A 的 Jordan 标准型中所有特征值为 λ_i 的 Jordan 块的最大阶数, 但是特征值为 λ_i 的 Jordan 块只有一个, 故 $c_i = n_i (1 \leq i \leq s)$. 因此 $f(x) = m(x)$.

(5) \Leftrightarrow (6): 记 $\varphi = A$.

充分性: 若 A 的所有几何重数都是 1, 则由 (3) \Leftrightarrow (6) 知, A 的特征多项式等于极小多项式, 从而由 **定理 0.7** 知, $\lambda E - A$ 相抵于

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & f(\lambda) \end{pmatrix}.$$

这里 f 是 A 的极小多项式. 不妨设为 $f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$. 而由 **定理 0.1** 知 f 的友矩阵 $C(f(\lambda))$ 也满足其特征多项式等于极小多项式等于 f , 同理由 **定理 0.7** 知, $\lambda E - C(f(\lambda))$ 也与上述矩阵相抵, 进而 $\lambda E - C(f(\lambda))$ 与 $\lambda E - A$ 相抵. 于是由定理??可知

$$A \sim \begin{pmatrix} & & -a_0 \\ 1 & & -a_1 \\ & \ddots & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix}.$$

于是结合 Cayley-Hamilton 定理知, 存在可逆矩阵 $P = (\alpha, \alpha_2, \dots, \alpha_n)$, 使得

$$AP = P \begin{pmatrix} & & -a_0 \\ 1 & & -a_1 \\ & \ddots & \vdots \\ & & 1 & -a_{n-1} \end{pmatrix} \iff (A\alpha, A\alpha_2, \dots, A\alpha_n) = (\alpha_2, \alpha_3, \dots, -a_{n-1}\alpha_n - \dots - a_1\alpha_2 - a_0\alpha)$$

$$\iff A\alpha = \alpha_2, A\alpha_i = \alpha_{i+1}, i = 2, 3, \dots, n-2, A\alpha_n = -a_{n-1}\alpha_n - \dots - a_1\alpha_2 - a_0\alpha$$

$$\iff \alpha = A\alpha, \alpha_{i+1} = A\alpha_i = A^2\alpha_{i-1} = \dots = A^i\alpha, i = 2, 3, \dots, n-2;$$

$$A\alpha_n = -a_{n-1}\alpha_n - \dots - a_1\alpha_2 - a_0\alpha = -f(A)\alpha + A^n\alpha + a_{n-1}A^{n-1}\alpha - a_{n-1}\alpha_n = A^n\alpha + a_{n-1}(A^{n-1}\alpha - \alpha_n).$$

$$\iff \alpha = A\alpha, \alpha_2 = A\alpha \dots, \alpha_{n-1} = A^{n-2}\alpha, (\alpha_n - A^{n-1}\alpha)(A - a_{n-1}E) = 0.$$

显然 $A \neq a_{n-1}E$, 故 $\alpha_n = A^{n-1}\alpha$. 因此 $\alpha_i = A^{i-1}\alpha, i = 2, 3, \dots, n$. P 可逆知如此构造的 $\{\alpha, A\alpha, \dots, A^{n-1}\alpha\}$ 是一组基. 这就证明了充分性.

必要性: 直接设 $f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0$ 是 A 的极小多项式, 由 Cayley-Hamilton 定理知 $f(A)\alpha = 0$, 进而

$$A^n\alpha = -a_{n-1}A^{n-1}\alpha - \dots - a_1A\alpha - a_0\alpha.$$

于是

$$A(\alpha, A\alpha, \dots, A^{n-1}\alpha) = (\alpha, A\alpha, \dots, A^{n-1}\alpha) \begin{pmatrix} & & & -a_0 \\ & & & -a_1 \\ & & \ddots & \vdots \\ & & & 1 & -a_{n-1} \\ 1 & & & & \end{pmatrix}.$$

而友矩阵特征多项式等于极小多项式, 因此由 (3) \Leftrightarrow (6) 知 A 的所有几何重数都是 1. 这就证明了必要性. \square