0.1 Jordan 标准型的进一步讨论和应用

定理 0.1

线性变换 φ 的特征值 λ_1 的度数等于 φ 的 Jordan 标准型中属于特征值 λ_1 的 Jordan 块的个数,λ_1 的重数等于所有属于特征值 λ_1 的 Jordan 块的阶数之和.

证明 设 V 是 n 维复线性空间,φ 是 V 上的线性变换. 设 φ 的初等因子组为

$$(\lambda-\lambda_1)^{r_1},(\lambda-\lambda_2)^{r_2},\cdots,(\lambda-\lambda_k)^{r_k}, \tag{1}$$

定理??告诉我们, 存在 V 的一组基 $\{e_{11},e_{12},\cdots,e_{1r_1};e_{21},e_{22},\cdots,e_{2r_2};\cdots;e_{k1},e_{k2},\cdots,e_{kr_k}\}$, 使得 φ 在这组基下的表示矩阵为

$$\boldsymbol{J}=\begin{pmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_k\end{pmatrix}.$$

上式中每个 $\boldsymbol{J}_i$ 是相应于初等因子 $(\lambda-\lambda_i)^{r_i}$ 的 Jordan 块, 其阶正好为 r_i. 令 V_i 是由基向量 $e_{i1},e_{i2},\cdots,e_{ir_i}$ 生成的子空间, 则

$$\begin{aligned}\varphi(e_{i1})&=\lambda_i e_{i1},\\ \varphi(e_{i2})&=e_{i1}+\lambda_i e_{i2},\\ &\cdots\cdots\cdots\cdots\\ \varphi(e_{ir_i})&=e_{i,r_i-1}+\lambda_i e_{ir_i}.\end{aligned} \tag{2}$$

这表明 $\varphi(V_i)\subseteq V_i$, 即 $V_i(i=1,2,\cdots,k)$ 是 φ 的不变子空间. 显然我们有

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

线性变换 φ 限制在 V_1 上 (仍记为 φ) 便成为 V_1 上的线性变换. 这个线性变换在基 $\{e_{11},e_{12},\cdots,e_{1r_1}\}$ 下的表示矩阵为

$$\boldsymbol{J}_1=\begin{pmatrix}\lambda_1 & 1 & & & \\ & \lambda_1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1\\ & & & & \lambda_1\end{pmatrix}.$$

注意到 $\boldsymbol{J}_1$ 的特征值全为 λ_1, 并且 $\lambda_1\boldsymbol{I}-\boldsymbol{J}_1$ 的秩等于 r_1-1, 故 $\boldsymbol{J}_1$ 只有一个线性无关的特征向量, 不妨选为 e_{11}. 显然 e_{11} 也是 φ 作为 V 上线性变换关于特征值 λ_1 的特征向量. 不失一般性, 不妨设在 φ 的初等因子组即(1) 式中

$$\lambda_1=\lambda_2=\cdots=\lambda_s,\quad \lambda_i\neq\lambda_1(i=s+1,\cdots,k),$$

则 $\boldsymbol{J}_1,\cdots,\boldsymbol{J}_s$ 都以 λ_1 为特征值, 且相应于每一块有且只有一个线性无关的特征向量. 相应的特征向量可取为

$$e_{11},e_{21},\cdots,e_{s1}, \tag{3}$$

显然这是 s 个线性无关的特征向量. 如果 $\lambda_i\neq\lambda_1$, 则容易看出 $\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J}_i)=r_i$, 于是

$$\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J})=\sum_{i=1}^{k}\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J}_i)=(r_1-1)+\cdots+(r_s-1)+r_{s+1}+\cdots+r_k=n-s.$$

因此 φ 关于特征值 λ_1 的特征子空间 V_{λ_1} 的维数等于 $n-\mathrm{r}(\lambda_1\boldsymbol{I}-\boldsymbol{J})=s$, 从而特征子空间 V_{λ_1} 以(3)式中的向量为一组基. 又 λ_1 是 φ 的 $r_1+r_2+\cdots+r_s$ 重特征值, 因此 λ_1 的重数与度数之差等于

$$(r_1+r_2+\cdots+r_s)-s.$$

□

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 $\mathbf{0} \neq \alpha \in V$, 则 $U = L(\alpha, \varphi(\alpha), \varphi^2(\alpha), \cdots)$ 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 0.2 (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,$\mathbf{0} \neq \alpha \in V$, $U = C(\varphi, \alpha)$ 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

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

□

定理 0.3

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

证明 设 $m = \max\{k \in \mathbb{Z}^+ \mid \alpha, \varphi(\alpha), \cdots, \varphi^{k-1}(\alpha)$ 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 $\alpha, \varphi(\alpha), \cdots, \varphi^{m-1}(\alpha)$ 的线性组合, 于是 $\{\alpha, \varphi(\alpha), \cdots, \varphi^{m-1}(\alpha)\}$ 是 U 的一组基, 从而 $m = \dim U = r$. □

定理 0.4

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 $\{e_1, e_2, \cdots, e_r\}$ 下的表示矩阵是友阵 $C(d(\lambda))$, 其中 $d(\lambda) = \lambda^r + a_1\lambda^{r-1} + \cdots + a_{r-1}\lambda + a_r$, 则由友阵的定义可知 $\varphi(e_i) = e_{i+1}(1 \leq i \leq r-1)$,$\varphi(e_r) = -\sum\limits_{i=1}^{r} a_{r-i+1}e_i$. 因此 $e_i = \varphi^{i-1}(e_1)(2 \leq i \leq r)$,$U = L(e_1, e_2, \cdots, e_r) = C(\varphi, e_1)$ 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 0.3可知,$\{\alpha, \varphi(\alpha), \cdots, \varphi^{r-1}(\alpha)\}$ 是 U 的一组基. 设

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

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

定理 0.5

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 $1, \cdots, 1, d_1(\lambda), \cdots, d_k(\lambda)$, 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:

$$V = C(\varphi, \alpha_1) \oplus C(\varphi, \alpha_2) \oplus \cdots \oplus C(\varphi, \alpha_k) \tag{4}$$

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

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 (4)就是有理标准型的几何意义.

证明 由定理??可知, 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

$$C = \mathrm{diag}\{C(d_1(\lambda)), C(d_2(\lambda)), \cdots, C(d_k(\lambda))\}$$

其中 $\varphi|_{L(e_{i1},\cdots,e_{ir_i})}$ 的表示矩阵就是友阵 $C(d_i(\lambda)), i = 1, 2, \cdots, k$. 再结合定理 0.4的讨论可知,$L(e_{i1}, \cdots, e_{ir_i})$ 就是一个循环子空间. 于是任取 $\alpha_i \in L(e_{i1}, \cdots, e_{ir_i})$ 作为循环向量, 则

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

其中 $\dim C(\varphi, \alpha_i) = r_i$.

综上可知, 此时 V 存在一个循环子空间的直和分解:

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

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

定理 0.6 (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 4 个结论等价:

(1) φ 的行列式因子组或不变因子组为 $1, \cdots, 1, f(\lambda)$;

(2) φ 的初等因子组为 $P_1(\lambda)^{r_1}, P_2(\lambda)^{r_2}, \cdots, P_k(\lambda)^{r_k}$, 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.

♡

证明 (1) ⇔ (2): 由不变因子和初等因子之间的相互转换即得.

(1) ⇔ (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) ⇔ (4): 若 V 是循环空间, 则由定理 0.4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??) 可知,φ 的行列式因子组和不变因子组均为 $1, \cdots, 1, g(\lambda) = f(\lambda)$. 若 φ 的不变因子组为 $1, \cdots, 1, f(\lambda)$, 则由有理标准型的几何意义 (定理 0.5) 可知,V 是循环空间. □

定义 0.2 (根子空间)

设 λ_0 是 n 维复线性空间 V 上线性变换 φ 的特征值, 则

$$R(\lambda_0) = \{\boldsymbol{v} \in V \mid (\varphi - \lambda_0 \boldsymbol{I})^n(\boldsymbol{v}) = \boldsymbol{0}\}$$

构成了 V 的一个子空间, 称为属于特征值 λ_0 的**根子空间**.

♣

定理 0.7

设 φ 是 n 维复线性空间 V 上的线性变换.

(1) 若 φ 的初等因子组为

$$(\lambda - \lambda_1)^{r_1}, (\lambda - \lambda_2)^{r_2}, \cdots, (\lambda - \lambda_k)^{r_k},$$

则 V 可分解为 k 个不变子空间的直和:

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_k, \tag{5}$$

其中 V_i 是维数等于 r_i 的关于 $\varphi - \lambda_i \boldsymbol{I}$ 的循环子空间;

(2) 若 $\lambda_1, \cdots, \lambda_s$ 是 φ 的全体不同特征值, 则 V 可分解为 s 个不变子空间的直和:

$$V = R(\lambda_1) \oplus R(\lambda_2) \oplus \cdots \oplus R(\lambda_s),$$

其中 $R(\lambda_i)$ 是 λ_i 的根子空间,$R(\lambda_i)$ 的维数等于 λ_i 的重数, 且每个 $R(\lambda_i)$ 又可分解为(5) 式中若干个 V_j 的直和.

♡

证明 在定理 0.1的证明的基础上, 现在再来看 $\boldsymbol{J}_1$ 所对应的子空间 V_1, 由 (2)式中诸等式可知

$$(\varphi-\lambda_1\boldsymbol{I})(e_{1r_1})=e_{1,r_1-1},\cdots,(\varphi-\lambda_1\boldsymbol{I})(e_{12})=e_{11},(\varphi-\lambda_1\boldsymbol{I})(e_{11})=\boldsymbol{0},$$

因此, 若记 $\boldsymbol{\alpha}=e_{1r_1},\psi=\varphi-\lambda_1\boldsymbol{I}$, 则

$$\psi(\boldsymbol{\alpha})=e_{1,r_1-1},\psi^2(\boldsymbol{\alpha})=e_{1,r_1-2},\cdots,\psi^{r_1-1}(\boldsymbol{\alpha})=e_{11},\psi^{r_1}(\boldsymbol{\alpha})=\boldsymbol{0}.$$

也就是说

$$\{\boldsymbol{\alpha},\psi(\boldsymbol{\alpha}),\psi^2(\boldsymbol{\alpha}),\cdots,\psi^{r_1-1}(\boldsymbol{\alpha})\}$$

构成了 V_1 的一组基.

上面的事实说明, 每个 Jordan 块 $\boldsymbol{J}_i$ 对应的子空间 V_i 是一个循环子空间. 把属于同一个特征值, 比如属于 λ_1 的所有循环子空间加起来构成 V 的一个子空间:

$$R(\lambda_1)=V_1\oplus\cdots\oplus V_s.$$

若 $\boldsymbol{v}\in R(\lambda_1)$, 则不难算出 $(\varphi-\lambda_1\boldsymbol{I})^s(\boldsymbol{v})=\boldsymbol{0}$, 其中

$$s=\dim R(\lambda_1)=r_1+\cdots+r_s.$$

事实上, 我们可以证明

$$R(\lambda_1)=\{\boldsymbol{v}\in V\mid(\varphi-\lambda_1\boldsymbol{I})^n(\boldsymbol{v})=\boldsymbol{0}\}.\tag{6}$$

为证明 (6)式成立, 设 $U=\{\boldsymbol{v}\in V\mid(\varphi-\lambda_1\boldsymbol{I})^n(\boldsymbol{v})=\boldsymbol{0}\}$, 则由上面的分析知道,$R(\lambda_1)\subseteq U$. 另一方面, 任取 $\boldsymbol{v}\in U$, 设 $\boldsymbol{v}=\boldsymbol{v}_1+\boldsymbol{v}_2$, 其中 $\boldsymbol{v}_1\in R(\lambda_1)$,$\boldsymbol{v}_2\in V_{s+1}\oplus\cdots\oplus V_k$. 因为 $(\lambda-\lambda_1)^n$ 与 $(\lambda-\lambda_{s+1})^n\cdots(\lambda-\lambda_k)^n$ 互素, 故存在多项式 $p(\lambda),q(\lambda)$, 使

$$(\lambda-\lambda_1)^n p(\lambda)+(\lambda-\lambda_{s+1})^n\cdots(\lambda-\lambda_k)^n q(\lambda)=1.$$

将 $\lambda=\varphi$ 代入上式并作用在 $\boldsymbol{v}$ 上可得

$$\begin{aligned}\boldsymbol{v}&=p(\varphi)(\varphi-\lambda_1\boldsymbol{I})^n(\boldsymbol{v})+q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v})\\&=q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v}_1)+q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v}_2)\\&=q(\varphi)(\varphi-\lambda_{s+1}\boldsymbol{I})^n\cdots(\varphi-\lambda_k\boldsymbol{I})^n(\boldsymbol{v}_1)\in R(\lambda_1).\end{aligned}$$

这就证明了 (6)式.

上面的结果表明: 特征值 λ_0 的根子空间可表示为若干个循环子空间的直和, 每个循环子空间对应于一个 Jordan 块. 虽然我们前面的讨论是对特征值 λ_1 进行的, 其实对任一特征值 λ_i 均适用. □

命题 0.1

证明: 复数域上的方阵 $\boldsymbol{A}$ 必可分解为两个对称阵的乘积.

证明 设 $\boldsymbol{P}$ 是非异阵且使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{J}$ 为 $\boldsymbol{A}$ 的 Jordan 标准型, 于是 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}$. 设 $\boldsymbol{J}_i$ 是 $\boldsymbol{J}$ 的第 i 个 Jordan 块, 则

$$\boldsymbol{J}_i=\begin{pmatrix}\lambda_i&1&&&\\&\lambda_i&1&&\\&&\ddots&\ddots&\\&&&\ddots&1\\&&&&\lambda_i\end{pmatrix}=\begin{pmatrix}&&&1&\lambda_i\\&&1&\lambda_i&\\&\iddots&\iddots&&\\1&\iddots&&&\\\lambda_i&&&&\end{pmatrix}\begin{pmatrix}&&&&1\\&&&1&\\&&\iddots&&\\&1&&&\\1&&&&\end{pmatrix},$$

即 $\boldsymbol{J}_i$ 可分解为两个对称阵之积. 因此 $\boldsymbol{J}$ 也可以分解为两个对称阵之积, 记为 $\boldsymbol{S}_1,\boldsymbol{S}_2$, 于是

$$\boldsymbol{A}=\boldsymbol{P}\boldsymbol{J}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{S}_1\boldsymbol{S}_2\boldsymbol{P}^{-1}=(\boldsymbol{P}\boldsymbol{S}_1\boldsymbol{P}')(\boldsymbol{P}^{-1})'\boldsymbol{S}_2\boldsymbol{P}^{-1}.$$

显然 $\boldsymbol{P}\boldsymbol{S}_1\boldsymbol{P}'$ 和 $(\boldsymbol{P}^{-1})'\boldsymbol{S}_2\boldsymbol{P}^{-1}$ 都是对称矩阵, 故 $\boldsymbol{A}$ 必可分解为两个对称阵的乘积. □

例题 0.1 已知

$$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{pmatrix},$$

计算 A^k.

解 用初等变换把 $\lambda\boldsymbol{I} - \boldsymbol{A}$ 化为对角 λ-矩阵并求出它的初等因子组为

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

因此,$\boldsymbol{A}$ 的 Jordan 标准型为

$$\boldsymbol{J} = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}.$$

因为

$$\boldsymbol{P}^{-1}\boldsymbol{A}^k\boldsymbol{P} = (\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})^k = \boldsymbol{J}^k,$$

故先计算 $\boldsymbol{J}^k$. 注意 $\boldsymbol{J}$ 是分块对角阵, 它的 k 次方等于将各对角块 k 次方, 因此

$$\boldsymbol{J}^k = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & k & & \\ 0 & 1 & & \\ & & 1 & k \\ & & 0 & 1 \end{pmatrix},$$

$$\begin{aligned} \boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{J}^k\boldsymbol{P}^{-1} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -5 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & k & & \\ 0 & 1 & & \\ & & 1 & k \\ & & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ -5 & 0 & -1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 2k+1 & k & 0 & 0 \\ -4k & -2k+1 & 0 & 0 \\ 6k & k & k+1 & k \\ -14k & -5k & -k & -k+1 \end{pmatrix}. \end{aligned}$$

□

定理 0.8 (Jordan-Chevalley 分解)

设 $\boldsymbol{A}$ 是 n 阶复矩阵, 则 $\boldsymbol{A}$ 可分解为 $\boldsymbol{A} = \boldsymbol{B} + \boldsymbol{C}$, 其中 $\boldsymbol{B}, \boldsymbol{C}$ 适合下面条件:

(1) $\boldsymbol{B}$ 是一个可对角化矩阵;

(2) $\boldsymbol{C}$ 是一个幂零阵;

(3) $\boldsymbol{B}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{B}$;

(4) $\boldsymbol{B}, \boldsymbol{C}$ 均可表示为 $\boldsymbol{A}$ 的多项式.

不仅如此, 上述满足条件 (1) (3) 的分解是唯一的.

证明 先对 $\boldsymbol{A}$ 的 Jordan 标准型 $\boldsymbol{J}$ 证明结论. 设 $\boldsymbol{A}$ 的全体不同特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 且

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s \end{pmatrix},$$

其中 $\boldsymbol{J}_i$ 是属于特征值 λ_i 的根子空间对应的块, 其阶设为 m_i. 显然对每个 i 均有 $\boldsymbol{J}_i = \boldsymbol{M}_i + \boldsymbol{N}_i$, 其中 $\boldsymbol{M}_i = \lambda_i \boldsymbol{I}$ 是对角阵,$\boldsymbol{N}_i$ 是幂零阵且 $\boldsymbol{M}_i\boldsymbol{N}_i = \boldsymbol{N}_i\boldsymbol{M}_i$. 令

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{M}_1 & & & \\ & \boldsymbol{M}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{M}_s \end{pmatrix}, \quad \boldsymbol{N} = \begin{pmatrix} \boldsymbol{N}_1 & & & \\ & \boldsymbol{N}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{N}_s \end{pmatrix},$$

则 $\boldsymbol{J} = \boldsymbol{M} + \boldsymbol{N}, \boldsymbol{MN} = \boldsymbol{NM}, \boldsymbol{M}$ 是对角阵,$\boldsymbol{N}$ 是幂零阵.

因为 $(\boldsymbol{J}_i - \lambda_i \boldsymbol{I})^{m_i} = \boldsymbol{O}$, 所以 $\boldsymbol{J}_i$ 适合多项式 $(\lambda-\lambda_i)^{m_i}$. 而 λ_i 互不相同, 因此多项式 $(\lambda-\lambda_1)^{m_1}, (\lambda-\lambda_2)^{m_2}, \cdots, (\lambda-\lambda_s)^{m_s}$ 两两互素. 由中国剩余定理, 存在多项式 $g(\lambda)$ 满足条件

$$g(\lambda) = h_i(\lambda)(\lambda-\lambda_i)^{m_i} + \lambda_i,$$

对所有 $i = 1, 2, \cdots, s$ 成立 (这里 $h_i(\lambda)$ 也是多项式). 代入 $\boldsymbol{J}_i$ 得到

$$g(\boldsymbol{J}_i) = h_i(\boldsymbol{J}_i)(\boldsymbol{J}_i - \lambda_i \boldsymbol{I})^{m_i} + \lambda_i \boldsymbol{I} = \lambda_i \boldsymbol{I} = \boldsymbol{M}_i.$$

于是

$$g(\boldsymbol{J}) = \begin{pmatrix} g(\boldsymbol{J}_1) & & & \\ & g(\boldsymbol{J}_2) & & \\ & & \ddots & \\ & & & g(\boldsymbol{J}_s) \end{pmatrix} = \begin{pmatrix} \boldsymbol{M}_1 & & & \\ & \boldsymbol{M}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{M}_s \end{pmatrix} = \boldsymbol{M}.$$

又因为 $\boldsymbol{N} = \boldsymbol{J} - \boldsymbol{M} = \boldsymbol{J} - g(\boldsymbol{J})$, 所以 $\boldsymbol{N}$ 也是 $\boldsymbol{J}$ 的多项式.

现考虑一般情形, 设 $\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{J}$, 则 $\boldsymbol{A} = \boldsymbol{PJP}^{-1} = \boldsymbol{P}(\boldsymbol{M}+\boldsymbol{N})\boldsymbol{P}^{-1}$. 令 $\boldsymbol{B} = \boldsymbol{PMP}^{-1}, \boldsymbol{C} = \boldsymbol{PNP}^{-1}$, 则 $\boldsymbol{B}$ 是可对角化矩阵,$\boldsymbol{C}$ 是幂零阵,$\boldsymbol{BC} = \boldsymbol{CB}$ 并且

$$g(\boldsymbol{A}) = g(\boldsymbol{PJP}^{-1}) = \boldsymbol{P}g(\boldsymbol{J})\boldsymbol{P}^{-1} = \boldsymbol{PMP}^{-1} = \boldsymbol{B},$$

从而 $\boldsymbol{C} = \boldsymbol{A} - g(\boldsymbol{A})$.

最后证明唯一性. 假设 $\boldsymbol{A}$ 有另一满足条件 (1) (3) 的分解 $\boldsymbol{A} = \boldsymbol{B}_1 + \boldsymbol{C}_1$, 则 $\boldsymbol{B} - \boldsymbol{B}_1 = \boldsymbol{C}_1 - \boldsymbol{C}$. 由 $\boldsymbol{B}_1\boldsymbol{C}_1 = \boldsymbol{C}_1\boldsymbol{B}_1$ 不难验证 $\boldsymbol{AB}_1 = \boldsymbol{B}_1\boldsymbol{A}, \boldsymbol{AC}_1 = \boldsymbol{C}_1\boldsymbol{A}$. 因为 $\boldsymbol{B} = g(\boldsymbol{A})$, 故 $\boldsymbol{BB}_1 = \boldsymbol{B}_1\boldsymbol{B}$. 同理 $\boldsymbol{CC}_1 = \boldsymbol{C}_1\boldsymbol{C}$. 设 $\boldsymbol{C}^r = \boldsymbol{O}, \boldsymbol{C}_1^t = \boldsymbol{O}$, 用二项式定理即知 $(\boldsymbol{C}_1 - \boldsymbol{C})^{r+t} = \boldsymbol{O}$. 于是

$$(\boldsymbol{B} - \boldsymbol{B}_1)^{r+t} = (\boldsymbol{C}_1 - \boldsymbol{C})^{r+t} = \boldsymbol{O}.$$

因为 $\boldsymbol{BB}_1 = \boldsymbol{B}_1\boldsymbol{B}$, 它们都是可对角化矩阵, 由命题**??**知它们可同时对角化, 即存在可逆阵 $\boldsymbol{Q}$, 使 $\boldsymbol{Q}^{-1}\boldsymbol{BQ}$ 和 $\boldsymbol{Q}^{-1}\boldsymbol{B}_1\boldsymbol{Q}$ 都是对角阵. 注意到

$$(\boldsymbol{Q}^{-1}\boldsymbol{BQ} - \boldsymbol{Q}^{-1}\boldsymbol{B}_1\boldsymbol{Q})^{r+t} = \left(\boldsymbol{Q}^{-1}(\boldsymbol{B} - \boldsymbol{B}_1)\boldsymbol{Q}\right)^{r+t} = \boldsymbol{Q}^{-1}(\boldsymbol{B} - \boldsymbol{B}_1)^{r+t}\boldsymbol{Q} = \boldsymbol{O},$$

两个对角阵之差仍是一个对角阵, 这个差的幂要等于零矩阵, 则这两个矩阵必相等, 由此即得 $\boldsymbol{B} = \boldsymbol{B}_1$, 从而 $\boldsymbol{C} = \boldsymbol{C}_1$. □