证明. 考虑线性映射 $\mathscr{A}: F^n \to F^m$, $X \mapsto AX$ 及 $\mathscr{B}: F^p \to F^n$, $Y \mapsto BY$. 令 $\mathscr{A}_1 = \mathscr{A}|_{\mathrm{Im}(\mathscr{B})}$. 首先由 $\ker(\mathscr{A}_1) = \ker(\mathscr{A}) \cap \mathrm{Im}(\mathscr{B})$ 知 $\dim\ker(\mathscr{A}_1) \le \dim\ker(\mathscr{A})$. 另一方面, 由维数公式,

$$\dim\ker(\mathscr{A}) = n - \mathrm{rank}(A),$$

$$\dim\ker(\mathscr{A}_1) = \dim\mathrm{Im}(\mathscr{B}) - \dim\mathrm{Im}(\mathscr{A}\mathscr{B}) = \mathrm{rank}(B) - \mathrm{rank}(AB).$$

综合这三个式子即得欲证. □

习　　题

习题 3.4.1. 证明: 模 W 同余关系是 V 上的等价关系, 且若 $\alpha_1 \equiv \beta_1 \mod W$, $\alpha_2 \equiv \beta_2 \mod W$, 则对任意 $\lambda, \mu \in F$, 均有

$$\lambda\alpha_1 + \mu\alpha_2 \equiv \lambda\beta_1 + \mu\beta_2 \mod W.$$

习题 3.4.2. 设 $(\alpha_1, \cdots, \alpha_r)$ 是 W 的一组基, 将它扩充为 V 的一组基 $(\alpha_1, \cdots, \alpha_n)$. 证明: $(\bar{\alpha}_{r+1}, \cdots, (\bar{\alpha}_n)$ 是商空间 W 的一组基.

习题 3.4.3. 设 $A \in F^{m\times n}$, $B \in F^{n\times p}$, $C \in F^{p\times r}$. 证明

$$\mathrm{rank}(AB) + \mathrm{rank}(BC) \le \mathrm{rank}(ABC) + \mathrm{rank}(B).$$

习题 3.4.4. 设 V_1, V_2 是线性空间 V 的子空间. 证明: 商空间 $V_1 + V_1/V_1$ 与 $V_2/V_1 \cap V_2$ 同构.

习题 3.4.5. 设 $V_0, \cdots, V_{n+1}$ 为 F 上有限维线性空间. $V_0 = V_{n+1} = \{0\}$. $\mathscr{A}_i: V_i \longrightarrow V_{i+1}$ 线性变换, $i = 0, 1, \cdots, n$. 且 $\ker(\mathscr{A}_{i+1}) = \mathrm{Im}(\mathscr{A}_i), i = 0, 1, \cdots, n-1$. 证明

$$\sum_{i=1}^{n} (-1)^i \dim V_i = 0.$$

3.5　线性变换与方阵

3.5.1　线性变换与方阵的一一对应

我们知道线性变换是线性空间到自身的线性映射, 相对于线性映射的一般理论, 线性变换有一些特殊性. 对于一般线性映射, 我们可以分别取定定义域和值域的基, 而对于线性变换, 定义域和值域是同一线性空间, 当然应该取同一组基.

设 V 是 n 维线性空间. 取定 V 的一组基 $(\alpha_1, \cdots, \alpha_n)$. 设 $\mathscr{A}: V \to V$ 是 V 上的线性变换. 则

$$\mathscr{A}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A,$$

其中 $A = (a_{ij})_{1\le i,j\le n}$ 是 F 上的 n 阶方阵. 我们称 A 是 $\mathscr{A}$ 在基 $(\alpha_1, \cdots, \alpha_n)$ 下的矩阵. 根据线性映射的理论, 在取定基 $(\alpha_1, \cdots, \alpha_n)$ 之后, $\mathscr{A}$ 与 A 一一对应.

记

$$\mathrm{End}(V) = \mathscr{L}(V, V) = \{V\text{上线性变换全体}\}.$$

我们已经知道 $\mathrm{End}(V)$ 是 n^2 维线性空间, 并且根据线性映射的复合, 我们可以在 $\mathrm{End}(V)$ 上定义线性变换的乘积. 这样通过 § 3.3.3 中的两个定理, 我们得到线性变换与方阵的一一对应:

定理 3.5.1. *设 V 是 n 维线性空间. 设 $(\alpha_1, \cdots, \alpha_n)$ 是 V 的一组基. 设 A 是线性变换 $\mathscr{A}$ 在该基下的矩阵. 则映射*

$$\sigma : \mathrm{End}(V) \to F^{n\times n}, \quad \mathscr{A} \mapsto A$$

是保持加法、数乘和乘法的一一对应, 其中恒等变换对应单位阵, 可逆线性变换对应为可逆方阵.

证明. 只需证明可逆线性变换和可逆方阵的对应. 若 $\mathscr{A}$ 可逆, $\mathscr{B}$ 是它的逆, 则 $\mathscr{A}\mathscr{B} = \mathscr{I}$, 从而 $\sigma(\mathscr{A})\sigma(\mathscr{B}) = I_n$ 即 $\sigma(\mathscr{A})$ 可逆. 反之, 若 A 可逆, 对应线性变换 $\mathscr{A}$. 令 $\mathscr{B}$ 是在基 $(\alpha_1, \cdots, \alpha_n)$ 下矩阵为 A^{-1} 的线性变换, 则 $\mathscr{A}\mathscr{B} = \mathscr{I}$. □

注记. 定理 3.5.1 就是在说 σ 是环 $\mathrm{End}(V)$ 和环 $F^{n\times n}$ 之间的同构. 根据这个定理, 就可以将线性变换和方阵的结果互相翻译. 这是本章和下一章的主要指导思想.

对于任意变换 $\mathscr{A} \in \mathrm{End}(V)$ (包括零变换!), 记 $\mathscr{A}^0 = \mathscr{I}$. 同理对 n 阶方阵 A (包括零方阵!), 记 $A^0 = I_n$. 对于多项式 $f(x) = \sum_{i=0}^n a_i x^i = a_0 + a_1 x + \cdots + a_n x^n \in F[x]$, 根据变换和方阵的加法、数乘和乘法, 就有

$$f(\mathscr{A}) = \sum_{i=0}^n a_i \mathscr{A}^i = a_0 \mathscr{I} + a_1 \mathscr{A} + \cdots + a_n \mathscr{A}^n,$$

$$f(A) = \sum_{i=0}^n a_i A^i = a_0 I_n + a_1 A + \cdots + a_n A^n.$$

定义 3.5.1. $F[\mathscr{A}] := \{f(\mathscr{A}) \mid f(x) \in F[x]\}$, $F[A] := \{f(A) \mid f(x) \in F[x]\}$.

命题 3.5.2. *$F[\mathscr{A}]$ 是 $\mathrm{End}(V)$ 的子空间, $F[A]$ 是 $F^{n\times n}$ 的子空间. 定理 3.5.1 的对应将 $F[\mathscr{A}]$ 对应到 $F[A]$, 将 $f(\mathscr{A})$ 对应到 $f(A)$.*

对于可逆线性变换, 定理 3.4.9 给出下述结果:

定理 3.5.3. *设 $\mathscr{A} : V \to V$ 是线性变换. 则下列条件等价:*

(1) *$\mathscr{A}$ 是可逆线性变换.*

(2) *$\mathscr{A}$ 是单射, 即 $\ker(\mathscr{A}) = 0$.*

(3) *$\mathscr{A}$ 是满射, 即 $\mathrm{Im}(\mathscr{A}) = V$.*

(4) *$\mathscr{A}$ 在任意基下的矩阵是可逆阵.*

(5) *$\mathscr{A}$ 将 V 的一组基变为另外一组基.*

证明. (1), (2), (3) 的等价是定理 3.4.9 的直接推论.

(1) 与 (4) 的等价是定理 3.5.1 的结果.

(1) ⇒ (5) 是同构映射的性质. 反过来, 若 $\mathscr{A}$ 将基 $(\alpha_1,\cdots,\alpha_n)$ 映到基 $(\beta_1,\cdots,\beta_n)$. 设 P 是 $(\alpha_1,\cdots,\alpha_n)$ 到 $(\beta_1,\cdots,\beta_n)$ 的过渡矩阵, 则 P 可逆, 且

$$\mathscr{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)P,$$

故 $\mathscr{A}$ 是可逆线性变换. □

现在假设 $(\beta_1,\cdots,\beta_n)$ 是 V 的另外一组基. 则存在可逆阵 P, 使得

$$(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)P.$$

所以

$$\mathscr{A}(\beta_1,\cdots,\beta_n)=\mathscr{A}(\alpha_1,\cdots,\alpha_n)P=(\alpha_1,\cdots,\alpha_n)AP=(\beta_1,\cdots,\beta_n)P^{-1}AP.$$

我们得到

定理 3.5.4. *设 $\mathscr{A}$ 是线性变换, 它在基 $(\alpha_1,\cdots,\alpha_n)$ 下的矩阵是 A. 如果从基 $(\alpha_1,\cdots,\alpha_n)$ 到基 $(\beta_1,\cdots,\beta_n)$ 的过渡矩阵是 P, 则 $\mathscr{A}$ 在基 $(\beta_1,\cdots,\beta_n)$ 下的矩阵是 $P^{-1}AP$. 反之, 如果 P 是可逆阵, 则 $\mathscr{A}$ 在基 $(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)P$ 下的矩阵是 $P^{-1}AP$.*

3.5.2 特征多项式、特征值和特征向量

定义 3.5.2. 设 A,B 是 n 阶方阵. 若存在可逆阵 P 使得 $B=P^{-1}AP$, 则称 A 与 B **相似** (similar).

命题 3.5.5. *相似关系是等价关系.*

证明. 这由可逆阵的性质立得. 由单位阵是可逆阵知自反性, 由 P 可逆则 P^{-1} 可逆且 $(P^{-1})^{-1}=P$ 知对称性, 由 $(PQ)^{-1}=Q^{-1}P^{-1}$ 得到传递性. □

由方阵相似的定义, 定理 3.5.4 就是说线性变换 $\mathscr{A}$ 在不同基下的矩阵彼此相似, 而相似的矩阵是同一线性变换在不同基下的矩阵. 如何寻找恰当的基使得线性变换在此基下有简单的形式, 或者等价地说在方阵的相似等价类里面找到最简单的矩阵, 就是相似标准形理论的主要内容. 为了这个目标, 我们需要引入一些概念.

首先相似矩阵有如下性质.

定理 3.5.6. *设方阵 A 与 B 相似, 即 $B=P^{-1}AP$. 则*

(1) $\mathrm{rank}(A)=\mathrm{rank}(B)$, $\det(A)=\det(B)$, $\mathrm{tr}(A)=\mathrm{tr}(B)$.

(2) *对任意* $f(x)\in F[x]$, $f(B)=P^{-1}f(A)P$.

(3) *多项式* $\det(xI_n-A)=\det(xI_n-B)$.

证明. (1) 根据 $B=P^{-1}AP$, 故 B 与 A 相抵, 所以

$$\mathrm{rank}(A)=\mathrm{rank}(B).$$

由行列式的积性,

$$\det(B) = \det(P^{-1})\det(A)\det(P) = \det(A).$$

根据 $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, 故

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

(2) 这是由于对每个 n 均有 $B^n = (P^{-1}AP)^n = P^{-1}A^nP$, 再由矩阵加法即得.

(3) 由 $xI_n - B = P^{-1}(xI_n - A)P$ 立得. □

根据上述定理, 我们可以定义

定义 3.5.3. 方阵 A 的特征多项式 $\varphi_A(x) = \det(xI_n - A)$.

线性变换 $\mathscr{A}$ 的行列式 $\det(\mathscr{A})$, 迹 $\mathrm{tr}(\mathscr{A})$ 与特征多项式 $\varphi_{\mathscr{A}}(x)$ 定义为它在某组基下矩阵 A 的行列式 $\det(A)$, 迹 $\mathrm{tr}(A)$ 与特征多项式 $\varphi_A(x)$.

在学习数乘变换 $\lambda\mathscr{I}$ 的时候, 我们曾指出它将每个向量 α 变为 $\lambda\alpha$, 即拉伸或者压缩 λ 倍. 如果考虑对某一个向量出现这种情况, 就有

定义 3.5.4. 设 $\mathscr{A}$ 是 V 上的线性变换. 若存在 $\alpha \neq 0$ 及 $\lambda \in F$ 使得 $\mathscr{A}\alpha = \lambda\alpha$, 则称 λ 为 $\mathscr{A}$ 的**特征值** (eigenvalue), α 称为 $\mathscr{A}$ 关于 λ 的**特征向量** (eigenvector).

同理, 对于 n 阶方阵 A, 若存在 $0 \neq X \in F^{n\times 1}$ 及 $\lambda \in F$ 使得 $AX = \lambda X$, 则称 λ 为 A 的 **特征值**, X 称为 λ 的**特征向量**.

命题 3.5.7. *设 $\mathscr{A}$ 是线性变换, 它在基 $(\alpha_1, \cdots, \alpha_n)$ 下的矩阵是 A. 若 α 是 $\mathscr{A}$ 的特征值 λ 的特征向量, 令 X 是 α 的坐标, 则 X 是 A 的特征值 λ 的特征向量. 反之亦然.*

证明. 一方面, 若 $0 \neq \alpha$ 且 $\mathscr{A}(\alpha) = \lambda\alpha$, 则 $X \neq 0$ 且

$$\mathscr{A}(\alpha) = \mathscr{A}(\alpha_1, \cdots, \alpha_n)X = (\alpha_1, \cdots, \alpha_n)(AX) = \lambda\alpha = (\alpha_1, \cdots, \alpha_n)(\lambda X),$$

故 $AX = \lambda X$. 反之同理可得. □

下述定理将特征多项式与特征值联系起来:

定理 3.5.8. *λ 是线性变换 $\mathscr{A}$(或者方阵 A) 的特征值当且仅当它是特征多项式 $\varphi_{\mathscr{A}}(x)$ (或者 $\varphi_A(x)$) 在 F 中的根.*

证明. 只需证明矩阵情形, 线性变换情形由上述命题即得. 我们有

$$\begin{aligned}
\lambda\text{是 } A \text{ 的特征值} &\iff \text{存在 } 0 \neq X, AX = \lambda X \\
&\iff \text{齐次方程组 } (\lambda I_n - A)X = 0 \text{ 有非零解} \\
&\iff \mathrm{rank}(\lambda I_n - A) < n \iff \det(\lambda I_n - A) = 0 \\
&\iff \lambda\text{是 } \varphi_A(x) \text{ 的根.}
\end{aligned}$$

□

注记. 注意: 若 $\lambda \notin F$ 是 $\varphi_A(x)$ 的根, λ 不是 A 在 F 上的特征值.

注意到对于方阵 $A=(a_{ij})$, 特征多项式

$$\varphi_A(x)=\det(xI_n-A)=x^n+a_{n-1}x^{n-1}+\cdots+a_0$$

是 F 上的 n 次首一多项式. 设它的 n 个根 (可以是在 F 的某个扩域里面, 比如复数域 $\mathbb{C}$) 为 $\lambda_1,\cdots,\lambda_n$. 则

$$\varphi_A(x)=(x-\lambda_1)\cdots(x-\lambda_n).$$

另一方面, $a_0=\varphi_A(0)=(-1)^n\det(A)$, $a_{n-1}=-a_{11}-\cdots-a_{nn}=\mathrm{tr}(A)$. 根据韦达定理, 则有

定理 3.5.9. *设特征多项式 $\varphi_A(x)$ 的 n 个根为 $\lambda_1,\cdots,\lambda_n$. 则*

$$\mathrm{tr}(A)=\lambda_1+\cdots+\lambda_n,\quad \det(A)=\lambda_1\cdots\lambda_n. \tag{3.12}$$

3.5.3 最小多项式与零化多项式

设 A 是 n 阶方阵, 我们已经注意到 $F[A]$ 是有限维线性空间 $F^{n\times n}$ 的子空间, 故它也是有限维线性空间. 设 $\dim F[A]=d$.

定理 3.5.10. *$(I_n,A,\cdots,A^{d-1})$ 是 $F[A]$ 的一组基.*

证明. 只需证明 $\{I_n,A,\cdots,A^{d-1}\}$ 线性无关即可. 否则, 设 λ_i 不全为 0 使得 $\sum_{i=0}^{d-1}\lambda_iA^i=0$. 令 $g(x)=\sum_{i=0}^{d-1}\lambda_ix^i$, 则 $g(x)$ 是次数 $\le d-1$ 的多项式. 对于任意 $f(x)\in F[x]$, 根据多项式带余除法, $f(x)=q(x)g(x)+r(x)$, 其中 $\deg(r)<\deg(g)\le d-1$. 故 $f(A)=q(A)g(A)+r(A)=r(A)$. 这说明 $f(A)$ 都是 $\{I_n,\cdots,A^{d-2}\}$ 的线性组合, 与 $\dim F[A]=d$ 矛盾. □

根据上面的定理, 存在唯一的 $c_i\in F$, 使得 $A^d=\sum_{i=0}^{d-1}c_iA^i$. 令

$$d_A(x)=x^d-\sum_{i=0}^{d-1}c_ix^i. \tag{3.13}$$

定义 3.5.5. 多项式 $d_A(x)$ 称为 A 的**最小多项式** (minimal polynomial). 线性变换 $\mathscr{A}$ 的最小多项式即它在某组基下矩阵的最小多项式.

注意到 $F[A]$ 与 $F[P^{-1}AP]$ 同构, 且

$$A^d=\sum_{i=0}^{d-1}c_iA^i \iff (P^{-1}AP)^d=\sum_{i=0}^{d-1}c_i(P^{-1}AP)^i,$$

这就说明

定理 3.5.11. *相似方阵有着相同的最小多项式, 从而线性变换的最小多项式是良定义的.*

根据定义, $d_A(A)=0$ ($d_{\mathscr{A}}(\mathscr{A})=\mathscr{O}$). 将这条性质推广, 就有

定义 3.5.6. 若 $0\ne f(x)\in F[x]$ 使得 $f(A)=0$ (或者 $f(\mathscr{A})=\mathscr{O}$), 称 $fx)$ 是 A (或者 $\mathscr{A}$) 的一个**零化多项式** (annihilator).

定理 3.5.12. *所有零化多项式都是最小多项式的倍式, 故最小多项式是唯一的次数最低的首一零化多项式.*

证明. 由于 $(I_n, A, \cdots, A^{d-1})$ 线性无关, 故不存在次数 $\leq d-1$ 的零化多项式. 设 $f(x)$ 是 A 的零化多项式, 对 $d_A(x)$ 做带余除法, 则 $f(x) = q(x)d_A(x) + r(x)$ $(\deg r \leq d-1)$ 且 $r(A) = f(A) - q(A)d_A(A) = 0$. 这就说明 $r(x) = 0$ 即 $d_A(x) \mid f(x)$. 线性变换情形同理可得. □

注记. 我们可以直接定义最小多项式为唯一的次数最低的首一零化多项式.

设 F 是 K 的子域. 我们可以将 A 看作是 F 上的矩阵, 也可以看作是 K 上的矩阵. 定理 3.5.10 说明 $d_A(x)$ 与域的变化无关, 它总是由 A 的元素生成的最小子域上的多项式.

定理 3.5.13. *A 的特征多项式的根均为最小多项式 $d_A(x)$ 的根, 且 $d_A(x)$ 没有其他根.*

证明. 我们假设 F 足够大 (例如 $F = \mathbb{C}$) 使得 A 的特征多项式的根均在 F 中, 即

$$\varphi_A(x) = (x-\lambda_1)\cdots(x-\lambda_n),\ \lambda_i \in F.$$

对于 $\lambda = \lambda_i$, 设 X 是它的特征值. 则对于任意 $f(x) \in F[x]$, $f(A)X = f(\lambda)X$. 所以

$$0 = d_A(A)X = d_A(\lambda)X.$$

由 $X \neq 0$ 即得 $d_A(\lambda) = 0$.

另一方面, 若 $\lambda \neq \lambda_i$ 是 $d_A(x)$ 的根, 则 $d_A(x) = (x-\lambda)d'(x)$. 所以 $0 = (A-\lambda I)d'(A)$. 但由于 $\det(A-\lambda I) \neq 0$, 故 $A - \lambda I$ 可逆, 所以 $d'(A) = 0$, 这与 $d_A(x)$ 是最小多项式矛盾. □

命题 3.5.14. *若 $f(x)$ 与 $d_A(x)$ 互素, 则 $f(A)$ 可逆. 反之亦然.*

证明. 由 Bezout 等式, 设 $u(x)f(x) + v(x)d_A(x) = 1$. 两边代入 $x = A$, 则有 $u(A)f(A) = I_n$, 即 $f(A)$ 可逆.

反之, 若 $f(A)$ 可逆. 若 $\gcd(f(x), d_A(x)) = w(x) \neq 1$. 由 $f(A)$ 可逆则 $w(A)$ 可逆. 令 $d'(x) = d_A(x)/w(x)$. 则 $d'(A) = 0$, 与 $d_A(x)$ 是最小多项式矛盾. □

例 3.5.1. *设 $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ 求 A 的最小多项式及 A^{-1}.*

证明. 首先计算 A 的特征多项式, 得

$$\varphi_A(x) = (x-1)^2(x-4).$$

这说明 $d_A(x)$ 的两个根为 1 和 4, 从而 $d_A(x) = (x-1)(x-4)$ 或者 $(x-1)^2(x-4)$. 经计算,

$$(A-I)(A-4I) = A^2 - 5A + 4I = 0,$$

故 $d_A(x) = (x-1)(x-4)$. 由 $A(5I_3 - A) = 4I$ 知

$$A^{-1} = \frac{1}{4}(5I_3 - A) = \frac{1}{4}\begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix}.$$ □

例 3.5.2. 设 $A \in \mathbb{C}^{m\times n}, B \in \mathbb{C}^{n\times m}$, 则 AB 与 BA 有相同的非零特征值.

证明. 若 $\lambda \neq 0$, $X \neq 0$ 且 $(AB)X = \lambda X$, 则 $BX \neq 0$ 且 $(BA)(BX) = \lambda(BX)$, 故 λ 也是 BA 的特征值. □

例 3.5.3. 方阵 A 称为**幂零方阵** *(nilpotent)* 是指存在正整数 k 使得 $A^k = 0$. 证明: A 为幂零方阵当且仅当它的唯一特征值为 0.

证明. 若 A 幂零, 则 x^k 是 A 零化多项式, 所以 $d_A(x) = x^l$ 对某个 $l \leq k$ 成立, 它只要一个根 0. 反之, 若 A 的唯一特征值为 0, 则 $d_A(x) = x^l$ 所以 $d_A(A) = A^l = 0$, 即 A 为幂零方阵. □

3.5.4 特征子空间

将特征多项式的相同根合并, 我们有

$$\varphi_A(x) = (x-\lambda_1)^{n_1}\cdots(x-\lambda_s)^{n_s}, \quad \sum n_i = n, \quad \lambda_i \text{ 两两不同}. \tag{3.14}$$

定义 3.5.7. n_i 称为 λ_i 的**代数重数**.

定义 3.5.8. λ_i 的**特征子空间** V_{λ_i}, 即齐次线性方程组 $(\lambda_i I - A)X = 0$ 的解空间 $\ker(\lambda_i I - A)$, 它的维数 $m_i = \dim V_{\lambda_i}$ 称为 λ_i 的**几何重数**.

定理 3.5.15. 下列条件成立:

(1) $V_{\lambda_i} \neq 0$ 或者说 $m_i \geq 0$ 当且仅当 λ_i 是 A 在 F 上的特征值.

(2) $W = V_{\lambda_1} + \cdots + V_{\lambda_s}$ 是直和.

(3) 几何重数不大于代数重数: $m_i \leq n_i$.

证明. (1) 即特征值的定义.

(2) 我们要证明 W 中零向量的分解式唯一. 若不然, 不妨设 $X_1 + \cdots + X_t = 0$, 其中 $0 \neq X_i \in V_{\lambda_i}$ $(1 \leq i \leq t)$. 则 $AX_i = \lambda_i X_i$. 将 A^j $(0 \leq j \leq t-1)$ 作用到 $X_1 + \cdots + X_t = 0$, 就有

$$\lambda_1^j X_1 + \cdots + \lambda_t^j X_t = 0$$

用矩阵语言, 就是

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_t \\ & & & \\ \lambda_1^{t-1} & \lambda_2^{t-1} & \cdots & \lambda_t^{t-1} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ \cdots \\ X_t \end{pmatrix} = 0.$$

左边的矩阵是 Vandermonde 矩阵, 它可逆, 故 $X_1 = \cdots = X_t = 0$, 矛盾.

(3) 取 V_{λ_i} 的基 M_i, 将 W 的基 $M_1 \cup \cdots \cup M_s$ 扩充为 $V = F^{n\times 1}$ 的基 $(X_1, \cdots, X_n)$. 则

$$A(X_1, \cdots, X_n) = (X_1, \cdots, X_n) \begin{pmatrix} \lambda_1 I_{m_1} & & & & * \\ & \lambda_2 I_{m_2} & & & * \\ & & \ddots & & * \\ & & & \lambda_s I_{m_s} & * \\ & & & & B \end{pmatrix}.$$

这就说明 $\varphi_A(x) = (x-\lambda_1)^{m_1} \cdots (x-\lambda_s)^{m_s} \varphi_B(x)$. 比较等式两边的项就得到 $m_i \le n_i$. □

习 题

习题 3.5.1. 设 m 阶方阵 A_1 与 B_1 相似, n 阶方阵 A_2 与 B_2 相似. 证明: $\mathrm{diag}(A_1, A_2)$ 与 $\mathrm{diag}(A_2, A_1)$ 相似, 也与 $\mathrm{diag}(B_1, B_2)$ 相似.

习题 3.5.2. 设 A 是域 F 上 n 阶方阵. 设 $f(x) \in F[x]$ 是非零多项式且满足条件 $f(A) = 0$. 设多项式 $g(x)$ 与 $f(x)$ 互素. 证明: $g(A)$ 一定是可逆矩阵.

习题 3.5.3. 设方阵 A 满足条件 $A^2 - 2A = 0$. 试用 A 的多项式表示 $A + I_n$ 的逆.

习题 3.5.4. 若方阵 A, B 都是幂零阵且 $AB = BA$, 证明: $A + B$ 也是幂零阵.

习题 3.5.5. 设 n 阶方阵 A 的秩为 $n-1$. 证明: 对任意正整数 k, $\mathrm{rank}(A) \ge n-k$.

习题 3.5.6. 给出 4 个互不相似的 4 阶方阵 A, 满足条件 $(A-I)^3 = 0$.

习题 3.5.7. 设实数方阵 $A = \begin{pmatrix} B & I & I \\ I & B & I \\ I & I & B \end{pmatrix}$, 其中 $B = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. 试求 A 的特征多项式与最小多项式.

习题 3.5.8. 设复方阵 A 满足 $\det(A+I) = 1$. 证明: 存在矩阵 B, 使得 $B^2 + 2B = A$.

3.6 对角化和准对角化

定义 3.6.1. 若存在 V 的基使得 $\mathscr{A}$ 在该基下的矩阵是对角阵 (上三角阵, 准对角阵), 则称 $\mathscr{A}$ 可对角化 (上三角化, 准对角化). 换言之, 若存在可逆阵 P 使得方阵 $P^{-1}AP$ 是对角阵 (上三角阵, 准对角阵), 则称 A 可对角化 (上三角化, 准对角化).

3.6.1 方阵的对角化

首先考虑对角化情形. 若方阵 A 在 F 中相似于对角阵 $\mathrm{diag}(\lambda_1, \cdots, \lambda_n)$, 则特征多项式 $\varphi_A(x) = (x-\lambda_1)\cdots(x-\lambda_n)$ 的根全在 F 中, 即全部是特征值. 这个条件是 A 可对角化的必要条件.

定理 3.6.1. 设 A 的特征多项式的根全在 F 中. 则 A 可对角化等价于下列条件之一:

(1) $F^{n\times 1}$ 有一组以特征向量组成的基.

(2) $F^{n\times 1} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$.

(3) 几何重数等于代数重数, 即 $m_i = n_i, i = 1\cdots s$.

(4) $d_A(x) = (x-\lambda_1)\cdots(x-\lambda_s)$ 没有重根.

证明. (1) 设 A 有一组以特征向量组成的基 $(X_1,\cdots,X_n)$, 对应的特征值为 $\lambda_1,\cdots,\lambda_n$. 则

$$A(X_1,\cdots,X_n) = (X_1,\cdots,X_n)\begin{pmatrix}\lambda_1 & & \\ & \vdots & \\ & & \lambda_n\end{pmatrix}.$$

记 $P = (X_1,\cdots,X_n)$, 那么 $A = P\Lambda P^{-1}$. 反之亦然.

(2) 由 (1), 若 A 可对角化, $F^{n\times 1} \subseteq V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$ 故二者必相等. 反之, 若 $F^{n\times 1} \subseteq V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 取 M_i 为 V_{λ_i} 的基, 则 $M_1 \cup \cdots \cup M_s$ 是 $F^{n\times 1}$ 的一组以特征向量组成的基.

(3) 由 $m_i \le n_i$, 故 $m_i = n_i, i = 1\cdots s$ 就等价于 $F^{n\times 1} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$.

(4) 一方面, 若 A 相似于对角阵 $D = \mathrm{diag}(\lambda_1,\cdots,\lambda_n)$, 由于对任意 $f(x)$, 均有 $f(D) = \mathrm{diag}(f(\lambda_1),\cdots,f(\lambda_n))$, 因此 $f(x)$ 是 D 的零化多项式当且仅当 $f(x)$ 的根包括所有对角元 λ_i. 这里面次数最小的首一多项式, 即 $d_A(x)$, 就是 $(x-\lambda_1)\cdots(x-\lambda_s)$.

反过来, 设 $d_A(x) = (x-\lambda_1)\cdots(x-\lambda_s)$. 令 $f_i(x) = d_A(x)/(x-\lambda_i)$, 则 $f_1(x),\cdots,f_s(x)$ 的最大公因式是 1. 由 Bezout 定理, 存在 $u_i(x) \in F[x]$ 使得

$$u_1 f_1 + \cdots + u_s f_s = 1.$$

故 $u_1(A)f_1(A) + \cdots + u_s(A)f_s(A) = I_n$. 对任意 $X \in F^{n\times 1}$, 令 $X_i = u_i(A)f_i(A)X$. 则 $X = X_1 + \cdots + X_s$, 并且

$$(\lambda_i I - A)X_i = u_i(A)d_A(A)X = 0,$$

即 $X_i \in V_{\lambda_i}$, 所以 $F^{n\times 1} = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$. □

推论 3.6.2. 设 A 的特征多项式的根全在 F 中. 若 A 的几何重数均为 1 重, 则 A 可对角化.

证明. 此时几何重数和代数重数都等于 1. □

例 3.6.1. 若 $A^2 = I$, 证明 A 相似于对角阵.

证明. 这是由于 $x^2 - 1$ 是 A 的零化多项式没有重根, 故最小多项式也没有重根. □

例 3.6.2. 设方阵 A, B 均可对角化且 $AB = BA$. 证明: A, B 可同时对角化.

证明. 设

$$A_1 = P^{-1}AP = \begin{pmatrix} \lambda_1 I_{n_1} & & \\ & \ddots & \\ & & \lambda_s I_{n_s} \end{pmatrix}$$

并将 $B_1 = P^{-1}BP = (B_{ij})$ 与左边矩阵同样分块. 由 $AB = BA$ 知 $A_1B_1 = B_1A_1$, 比较等式两边就得 $B_{ij} = 0$ 对所有 $i \neq j$ 成立, 即

$$B_1 = P^{-1}BP = \begin{pmatrix} B_{11} & & \\ & \ddots & \\ & & B_{ss} \end{pmatrix}.$$

由 B 可对角化, 故 B_1 可对角化所以 B_1 的最小多项式无重根, 从而得知 B_{ii} 可对角化. 因而

$$B_{ii} = P_{ii}\Lambda_i P_{ii}^{-1}.$$

令

$$Q = P \begin{pmatrix} P_{11} & & \\ & \ddots & \\ & & P_{ss} \end{pmatrix},$$

则 $Q^{-1}AQ = \mathrm{diag}(\lambda_1 I_{n_1}, \cdots, \lambda_s I_{n_s})$, $Q^{-1}BQ = \mathrm{diag}(\Lambda_1, \cdots, \Lambda_s)$, 它们同为对角阵. □

3.6.2 三角化和准对角化

定理 3.6.3. *设 $\varphi_A(x)$ 的根全在 F 中 (例如 $F = \mathbb{C}$), 则 A 在 F 中相似于上三角阵.*

证明. 首先取 X_1 为 λ_1 的特征值, 将它扩充为 $F^{n\times 1}$ 的一组基 $(X_1, \cdots, X_n)$, 则

$$A(X_1, \cdots, X_n) = (X_1, \cdots, X_n) \begin{pmatrix} \lambda_1 & * \\ 0 & B \end{pmatrix}.$$

将 $(X_1, \cdots, X_n)$ 看作可逆矩阵 P, 则 $P^{-1}AP = \begin{pmatrix} \lambda_1 & * \\ 0 & B \end{pmatrix}$. 此时 $\varphi_B(x)$ 的根全在 F 中. 对方阵的阶 n 做归纳, 设 $Q^{-1}BQ = U$ 是上三角阵, 则

$$\begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix}^{-1} P^{-1}AP \begin{pmatrix} 1 & 0 \\ 0 & Q \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & U \end{pmatrix}$$

是上三角阵. □

定理 3.6.4 (Cayley-Hamilton 定理)**.** *对任意 $A \in F^{n\times n}$, 均有 $\varphi_A(A) = 0$, 即 $d_A(x) \mid \varphi_A(x)$.*

证明. 设 $xI - A$ 的伴随方阵为 $B(x) = B_{n-1}x^{n-1} + \cdots + B_1x + B_0$. 则

$$(xI - A)B(x) = \det(xI - A) \cdot I_n = \varphi_A(x) \cdot I_n.$$

设 $\varphi_A(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$. 则上式就是说

$$B_{n-1}x^n+(B_{n-2}-AB_{n-1})x^{n-1}+\cdots+(B_0-AB_1)x-AB_0=(x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0)I_n.$$

比较两边的系数, 就得到

$$\begin{cases} B_{n-1}=I_n, \\ B_{n-2}-AB_{n-1}=a_{n-1}I_n, \\ \quad\vdots \\ B_0-AB_1=a_1I_n, \\ -AB_0=a_0I_n. \end{cases}$$

所以 $\varphi_A(A)=A^n+a_{n-1}A^{n-1}+\cdots+a_1A+a_0I_n$ 就等于

$$A^nB_{n-1}+A^{n-1}(B_{n-2}-AB_{n-1})+\cdots+A(B_0-AB_1)-AB_0=0. \qquad \square$$

证 *2.* 由于等式左边与域没有关系, 我们可以扩充域 F 使得它包含所有的 λ_i. 则 A 相似于上三角阵. 故只需要对

$$A=\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}=\begin{pmatrix} \lambda_1 & * \\ 0 & B \end{pmatrix}$$

证明即可, 其中 B 是 $n-1$ 阶上三角阵, 且

$$\varphi_A(x)=(x-\lambda_1)\cdots(x-\lambda_n), \quad \varphi_B(x)=(x-\lambda_2)\cdots(x-\lambda_n).$$

我们对 n 做归纳证明. 若 $n=1$ 显然成立. 设定理对 $n-1$ 阶上三角阵成立, 则 $\varphi_B(B)=0$. 故

$$\varphi_A(A)=(A-\lambda_1I)\varphi_B(A)=\begin{pmatrix} 0 & * \\ 0 & * \end{pmatrix}\begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix}=0. \qquad \square$$

例 3.6.3. 设 $A\in F^{m\times m}, B\in F^{n\times n}$, 且 $\varphi_A(x)$ 与 $\varphi_B(x)$ 互素. 则线性变换

$$\mathscr{A}: F^{m\times n}\longrightarrow F^{m\times n}, \quad X\longrightarrow AX-XB$$

是可逆线性变换.

证明. 只要证明 $\ker(\mathscr{A})=0$ 即可. 也就是说若 $AX=XB$, 则 $X=0$.

由 $AX=XB$, 故 $A^2X=AXB=XB^2$, 从而 $A^kX=XB^k$, $f(A)X=Xf(B)$ 对任意 $f(x)\in F[x]$ 成立. 取 $f(x)=\varphi_A(x)$. 一方面 $\varphi_A(A)X=0$. 另一方面, $\varphi_A(x)$ 与 $\varphi_B(x)$ 互素, 则 $\varphi_A(B)$ 是可逆阵. 故 $X=0$. $\square$

例 3.6.4. 设

$$\varphi_A(x)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_s)^{n_s}, \qquad \lambda_i\in F, \text{ 两两不同}.$$

则 A 在 F 上相似于准对角阵

$$D=\begin{pmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{ss} \end{pmatrix}, \quad \text{其中} \quad A_{ii}=\begin{pmatrix} \lambda_i & & * \\ & \ddots & \\ 0 & & \lambda_i \end{pmatrix}.$$

证明. 我们不妨假设

$$A=\begin{pmatrix} A_{11} & B \\ 0 & A_1 \end{pmatrix}$$

的形式, 其中 $A_{11}=\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_1 \end{pmatrix}$ 为 n_1 阶方阵. 由 $\varphi_{A_{11}}(x)=(x-\lambda_1)^{n_1}$ 知 $\varphi_A(x)=(x-\lambda_2)^{n_2}\cdots(x-\lambda_s)^{n_s}$, 它们两者互素. 根据例 3.6.3, 存在 X 使得 $A_{11}X-XA_1=B$. 故

$$\begin{pmatrix} I & X \\ 0 & I \end{pmatrix}\begin{pmatrix} A_{11} & B \\ 0 & A_1 \end{pmatrix}\begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}=\begin{pmatrix} A_{11} & 0 \\ 0 & A_1 \end{pmatrix}.$$

这样只需要证明 A_1 可以如定理所示准对角化即可. 这由对 s 的归纳即得. □

习 题

习题 3.6.1. 证明: 方阵 A 是幂零阵当且仅当 A 只有一个特征值 0.

习题 3.6.2. 设 W 是 $F^{n\times n}$ 的子空间, 且对于任意 $A,B\in W$ 均有 $AB=BA$. 试证明: $\dim W\le n$.

习题 3.6.3. 设 $\mathscr{A}$ 是有限维复空间的线性变换, $\lambda_1,\cdots,\lambda_t$ 是它所有的特征值. 证明: $\mathscr{A}$ 可对角化当且仅当对所有的特征值 λ_i, $\ker(\mathscr{A}-\lambda_i\mathscr{I})^2$ 与 λ_i 的特征子空间是同一个空间.

第四章　相似标准形

本章主要目的是研究相似标准形理论. 给定方阵, 相似标准形理论就是寻找尽可能简单的矩阵与它相似, 或者从线性变换的观点来说, 就是对于一个线性变换, 寻找一组基, 使得它在该基下的矩阵足够简单. 所谓简单, 一个判断标准就是要有尽可能多的 0, 这样运算的时候会方便许多, 这是所谓的**打洞**.

标准形理论是线性代数重要组成部分. 在矩阵部分的学习中, 我们知道矩阵是否相抵是由秩唯一确定的, 一个秩为 r 的 $m \times n$ 矩阵 A 与 $\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}_{m\times n}$ 相抵, 后者是一个很简单的矩阵, 它只有 r 个位置是 1, 而其他位置都是 0, 而由于相抵矩阵的秩不变性, 我们也不能找到更简单的矩阵与 A 相抵. 这样得到的相抵标准形, 在矩阵运算中起到很大的作用.

对于方阵的相似标准形, 当然我们希望它也足够简单, 比如里面有尽可能多的 0, 实在不行 1 也可以, 并且最好是对角阵或者准对角阵, 或者退而求其次是上三角阵. 但必须注意到, 相较于相抵的情况, 相似要复杂得多.

首先, 由于相似的方阵有相同的特征多项式和特征值, 如果域 F 上的方阵与上三角阵在域 F 上相似, 那么它的特征值就是上三角阵的主对角元, 因此必须在域 F 中, 即特征多项式的根都在 F 中. 当然, 通过数学归纳法, 这个结论反过来也是对的, 如果方阵 A 的特征值全在 F 中, 则 A 一定 F 相似于上三角阵. 这就说明实矩阵 $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ 永远不可能实相似于对角阵, 因为它的特征多项式 x^2+1 没有实根.

进一步说, 如果方阵特征值都在域中, 比如说在 $F=\mathbb{C}$ 的情形 (代数基本定理), 这时候方阵相似于上三角阵, 但这个上三角阵主对角元是确定的, 就是方阵特征值 (包含重数) 的一个排列, 也就是说一般而言这些元素不是 0 或者 1. 这个时候大家的美好愿望是这个上三角阵可以进一步相似于对角阵, 但这个愿望也是达不到的. 我们举个例子. 设 $n \geq 2$. 通过计算可得 $N_n = \begin{pmatrix} 0 & I_{n-1} \\ 0 & 0 \end{pmatrix}_{n\times n}$ 的最小多项式是 x^n, 它有重根. 根据上一章学到的线性变换和方阵的对角化理论, 我们知道可对角化的方阵的最小多项式没有重根, 故 N_n 一定不可以对角化.

基于这些原因, 本章得到的相似标准形是准对角阵. 若方阵或线性变换的特征值都在域中, 我们将给出 **Jordan 标准形理论**, 此时准对角块是 Jordan 块, 而在一般情况则给出 **有理标准形理论**, 此时准对角块是特征多项式因子的友阵. 我们将首先使用空间分解定理给出 Jordan 标准形的证明, 然后发展整系数和多项式矩阵的 Smith(相抵) 标准形理论, 给出有理标准形理论的证明, 并给出 Jordan 标准形第二个证明. 概括而言, Jordan 标准形在计算上有更多应用, 而 Smith 标准形

和有理标准形则在理论上有更多应用.

4.1 空间分解定理

如不做特别说明, 本节总假定 $\mathbf{A}$ 是有限维 F-线性空间 V 上的线性变换.

4.1.1 不变子空间

定义 4.1.1. 设 $\mathbf{A}$ 是 (有限维) F-线性空间 V 上的线性变换. 若 W 是 V 的子空间且 $\mathbf{A}(W) \subseteq W$, 则称 W 是 $\mathbf{A}$ 的**不变子空间** (invariant subspace), 并记 $\mathbf{A}|_W : W \to W, w \mapsto \mathbf{A}(w)$ 为变换 $\mathbf{A}$ 在 W 上的**限制变换**.

注记. 对于任意线性映射 $\mathbf{A} : V \to U$ 以及 V 的任意子空间 W , 总可以定义限制映射 $\mathbf{A}_W : W \to U, w \mapsto \mathbf{A}(w)$. 当 $U = V$ 且 W 是 $\mathbf{A}$ 的不变子空间时, 限制映射 $\mathbf{A}_W$ 是从 W 到 V 的线性映射, 而限制变换 $\mathbf{A}|_W$ 则是从 W 到自身的线性映射 (变换).

例 4.1.1. 设 $\mathbf{A} : V \to V$ 是线性变换.

(1) 全空间 V 和零空间 $\{0\}$ 显然都是 $\mathbf{A}$ 的不变子空间.

(2) 对 $\lambda \in F$, 特征子空间 $V_\lambda = \{\alpha \in V \mid \mathbf{A}(\alpha) = \lambda\alpha\}$ 是 $\mathbf{A}$ 的不变子空间, 且 $V_\lambda \neq \{0\}$ 当且仅当 λ 是 $\mathbf{A}$ 的特征值.

(3) 核空间 $\ker \mathbf{A}$ 与像空间 $\mathrm{im}\mathbf{A}$ 是 $\mathbf{A}$-不变子空间.

(4) 不变子空间的和和交都是不变子空间.

例 4.1.2. 设 α 是 $\mathbf{A}$ 在 F 上的特征值 λ 的特征向量, 则 $F\alpha$ 是 $\mathbf{A}$ 的 1 维不变子空间. 反之, 若 W 是 $\mathbf{A}$ 的 1 维不变子空间, 则对于任意非零向量 $\alpha \in W$, 存在同一个 $\lambda \in F$, 使得 α 是 $\mathbf{A}$ 关于特征值 λ 的的特征向量.

命题 4.1.1. 设 $\mathbf{A}$ 是 V 上的线性变换. 设 $\{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一组基, 且

$$\mathbf{A}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A = (\alpha_1, \cdots, \alpha_n)\begin{pmatrix} A_1 & B \\ C & A_2 \end{pmatrix},$$

其中 A_1 是 m 阶方阵 $(m < n)$. 则下列结论成立:

(1) $W = \langle \alpha_1, \cdots, \alpha_m \rangle$ 是 $\mathbf{A}$ 的的不变子空间当且仅当 $C = 0$, 此时限制变换 $\mathbf{A}|_W$ 在基 $\{\alpha_1, \cdots, \alpha_m\}$ 下的矩阵为 A_1.

(2) $U = \langle \alpha_{m+1}, \cdots, \alpha_n \rangle$ 是 $\mathbf{A}$ 的的不变子空间当且仅当 $B = 0$, 此时限制变换 $\mathbf{A}|_U$ 在基 $\{\alpha_{m+1}, \cdots, \alpha_n\}$ 下的矩阵为 A_2.

(3) $W = \langle \alpha_1, \cdots, \alpha_m \rangle$ 与 $U = \langle \alpha_{n+1}, \cdots, \alpha_n \rangle$ 都是 $\mathbf{A}$ 的不变子空间当且仅当 B 和 C 均是零矩阵.

证明. (3) 由 (1) 和 (2) 立得, (2) 与 (1) 实际上是一回事, 我们只需证明 (1).

根据定义, W 是 $\mathbf{A}$ 的不变子空间是说 $\mathbf{A}(W) \subseteq W$, 也就是说对所有 $1 \leq i \leq m$, $\mathbf{A}(\alpha_i) = \sum_{j=1}^{m} a_{ij}\alpha_j \in W$, 这等价于 $C = 0$. □

将上述命题推广一下, 就得到

命题 4.1.2. 设 $\mathbf{A} : V \to V$ 为线性变换.

(1) 若 $V = V_1 \oplus \cdots \oplus V_s$ 且 V_i 是 $\mathbf{A}$ 的不变子空间, 设 B_i 为 V_i 的一组基且 $\mathbf{A}|_{V_i}$ 在此基下的矩阵为 A_i, 则 $\mathbf{A}$ 在基 $B = B_1 \cup \cdots \cup B_s$ 下的矩阵是准对角阵 $\mathrm{diag}(A_1, \cdots, A_s)$.

(2) 反之, 若 $\mathbf{A}$ 在基 $B = \{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是准对角阵 $\mathrm{diag}(A_1, \cdots, A_s)$, 其中 A_i 是 n_i 阶方阵, 令

$$B_i = \{\alpha_{n_1+\cdots+n_{i-1}+1}, \cdots, \alpha_{n_1+\cdots+n_i}\},$$

则 $V_i = \langle B_i \rangle$ 是 $\mathbf{A}$ 的不变子空间, $\mathbf{A}|_{V_i}$ 在基 B_i 下的矩阵为 A_i, 且 $V = V_1 \oplus \cdots \oplus V_s$.

根据上述命题, 线性变换 $\mathbf{A}$ 能否准对角化的关键在于全空间 V 能否分解为 $\mathbf{A}$ 的不变子空间的直和.

4.1.2 循环子空间

定义 4.1.2. 设 $\alpha \in V$. 包含 α 的最小的 $\mathbf{A}$-不变子空间称为 α 生成的**循环子空间** (cyclic subspace).

由例 4.1.2可知, 1 维不变子空间是特征向量生成的循环子空间, 反之特征向量生成的循环子空间一定是 1 维的. 更进一步地, 我们有

引理 4.1.3. 向量 α 生成的循环子空间即 $F[\mathbf{A}]\alpha := \{f(\mathbf{A})(\alpha) \mid f(x) \in F[x]\}$.

证明. 容易验证 $F[\mathbf{A}]\alpha$ 是包含 α 的 $\mathbf{A}$-不变子空间, 我们只需证明 $F[\mathbf{A}]\alpha \subseteq W$ 对所有包含 α 的不变子空间 W 成立即可.

由于 $\mathbf{A}(W) \subseteq W$, 故 $\mathbf{A}(\alpha) \in W$, 对 $i \geq 0$ 做归纳即得 $\mathbf{A}^i(\alpha) \in W$. 再由于 W 是线性子空间, 故对于任意 $f(x) \in F[x]$, 均有 $f(\mathbf{A})(\alpha) \in W$, 因此 $F[\mathbf{A}]\alpha \subseteq W$. □

定义 4.1.3. 设 $0 \neq \alpha \in V$. 若多项式 $0 \neq f(x) \in F[x]$ 满足 $f(\mathbf{A})(\alpha) = 0$, 则称 $f(x)$ 是 α 的**零化多项式**. 称次数最低的首一零化多项式为 α 的**最小多项式** (minimal polynomial), 记为 $d_{\mathbf{A},\alpha}(x)$ 或 $d_\alpha(x)$.

命题 4.1.4. 多项式 $f(x)$ 是向量 α 的零化多项式当且仅当它是最小多项式 $d_\alpha(x)$ 的非零倍式. 特别地,

(1) $\mathbf{A}$ 的零化多项式均是 $d_\alpha(x)$ 的非零倍式, 因而 $d_\alpha(x) \mid d_{\mathbf{A}}(x) \mid \varphi_{\mathbf{A}}(x)$.

(2) $\mathbf{A}$ 的最小多项式 $d_{\mathbf{A}}(x)$ 是所有 $d_\alpha(x)$ $(0 \neq \alpha \in V)$ 的最小公倍式.

证明. 显然. □

例 4.1.3. 设 A 是 $f(x)$ 的友阵, $\mathbf{A}: F^{n\times 1} \to F^{n\times 1}$, $X \mapsto ax$. 则向量 $e_1 = (1, 0, \cdots, 0)^T$ 的最小多项式等于 $f(x)$, 因此 $d_A(x) = f(x)$.

命题 4.1.5. 设 α 的最小多项式是 $d_\alpha(x)$, 它的次数等于 n_α. 则循环子空间 $F[\mathbf{A}]\alpha$ 是以 $\{\alpha, \mathbf{A}(\alpha), \cdots, \mathbf{A}^{n_\alpha-1}(\alpha)\}$ 为基的 n_α 维线性子空间, 且 $\mathbf{A}|_{F[\mathbf{A}]\alpha}$ 在此基下的矩阵是 $d_\alpha(x)$ 的友阵.

证明. 若 $g(x) \in F[x]$, 做带余除法 $g(x) = q(x)d_\alpha(x) + r(x)$, 其中 $\deg r(x) < n_\alpha$. 则 $g(\mathbf{A})(\alpha) = r(\mathbf{A})(\alpha)$, 故 $F[\mathbf{A}]\alpha$ 由 $\{\alpha, \mathbf{A}(\alpha), \cdots, \mathbf{A}^{n_\alpha-1}(\alpha)\}$ 生成. 另一方面, 若 $\sum_{i=0}^{n_\alpha-1} \lambda_i \mathbf{A}^i(\alpha) = 0$ 且 λ_i 不全为 0, 则 $\sum_{i=0}^{n_\alpha-1} \lambda_i x^i$ 是 α 的次数 $< n_\alpha$ 的零化多项式, 这不可能. 这说明 $\{\alpha, \mathbf{A}(\alpha), \cdots, \mathbf{A}^{n_\alpha-1}(\alpha)\}$ 线性无关, 它是 $F[\mathbf{A}]\alpha$ 的一组基.

记 $d_\alpha(x) = x^{n_\alpha} + a_{n_\alpha-1}x^{n_\alpha-1} + \cdots + a_0$. 则

$$\mathbf{A}(\alpha, \mathbf{A}(\alpha), \cdots, \mathbf{A}^{n_\alpha-1}(\alpha)) = (\alpha, \mathbf{A}(\alpha), \cdots, \mathbf{A}^{n_\alpha-1}(\alpha)) \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ & \ddots & \ddots & & \\ & & \ddots & 0 & -a_{n_\alpha-2} \\ & & & 1 & -a_{n_\alpha-1} \end{pmatrix}.$$

即 $\mathbf{A}|_{F[\mathbf{A}]\alpha}$ 在此基下的矩阵是 $d_\alpha(x)$ 的友阵. □

为 Jordan 标准形理论做准备, 我们先引入 Jordan 形矩阵的概念.

定义 4.1.4. 设 F 是域, $\lambda \in F$. 以 λ 为中心的 n 阶 **Jordan 块** $J_n(\lambda)$ 是指域 F 上的 n 阶方阵 $\lambda I_n + N_n$, 即 $J_1(\lambda) = (\lambda)$, 而当 $n \geq 2$ 时,

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}_{n\times n}. \tag{4.1}$$

若方阵 J 是对角块为 Jordan 块的准对角阵, 也就是说

$$J = \mathrm{diag}(J_{n_1}(\lambda_1), \cdots, J_{n_s}(\lambda_s)),$$

则称 J 是 **Jordan 形矩阵**. 我们称 Jordan 形矩阵中心是 λ 的块为 **λ-块**.

命题 4.1.6. 设 $\mathbf{A}: V \to V$ 是线性变换.

(1) 若 $\alpha \in V$ 的最小多项式是 $(x-\lambda)^n$, 则 $\mathbf{A}|_{F[x]\alpha}$ 在基 $\{(\mathbf{A}-\lambda\mathbf{I}_n)^{n-1}\alpha, (\mathbf{A}-\lambda\mathbf{I}_n)^{n-2}\alpha, \cdots, \alpha\}$ 下的矩阵是 $J_n(\lambda)$.

(2) 反之, 若 $\mathbf{A}$ 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 *Jordan* 块 $J_n(\lambda)$, 则 α_n 的最小多项式是 $(x-\lambda)^n$, $V = F[\mathbf{A}]\alpha$ 且

$$\alpha_i = (\mathbf{A}-\lambda\mathbf{I}_n)^{n-i}(\alpha_n) \ (1 \leq i \leq n).$$

证明. (1) 的证明与命题 4.1.5的证明类似, 只要证明 $\{(\mathbf{A}-\lambda\mathbf{I}_n)^{n-1}(\alpha),(\mathbf{A}-\lambda\mathbf{I}_n)^{n-2}(\alpha),\ \cdots,\alpha\}$ 是 $F[x]\alpha$ 的一组基即得.

对于 (2), 由 $\mathbf{A}(\alpha_1,\cdots,\alpha_n)=(\alpha_1,\cdots,\alpha_n)J_n(\lambda)$, 故得

$$\mathbf{A}(\alpha_1)=\lambda\alpha_1,\ \mathbf{A}(\alpha_2)=\alpha_1+\lambda\alpha_2,\ \cdots,\mathbf{A}(\alpha_n)=\alpha_{n-1}+\lambda\alpha_n.$$

从而 $\alpha_i=(\mathbf{A}-\lambda\mathbf{I}_n)^{n-i}(\alpha_n)$ $(1\le i\le n)$ 且 $(\mathbf{A}-\lambda\mathbf{I}_n)^n\alpha_n=0$. □

命题 4.1.7. *设 $\alpha\in V$ 的最小多项式 $d_\alpha(x)=\prod_{i=1}^s d_i(x)$, 其中 $d_i(x)$ 是两两互素的首一多项式. 则存在唯一的 $\alpha_i\in V$ $(i=1,\cdots s)$, 使得 $d_{\alpha_i}(x)=d_i(x)$ 且 $\alpha=\alpha_1+\cdots+\alpha_s$, 此时循环子空间 $F[\mathbf{A}]\alpha$ 是循环子空间 $F[\mathbf{A}]\alpha_i$ $(i=1,\cdots s)$ 的直和.*

证明. 先证唯一性. 若 α_i' 也满足条件, 则由等式 $\alpha_i-\alpha_i'=\sum_{j\ne i}(\alpha_j'-\alpha_j)$, 考虑等式左边它被 $d_i(x)$ 零化, 考虑右边它被 $\prod_{j\ne i}d_j(x)$ 零化, 但二者互素, 故 $\alpha_i-\alpha_i'=0$, 即 $\alpha_i=\alpha_i'$.

记 $D_i(x)=\prod_{j\ne i}d_j(x)$. 则 $D_i(x)$ $(i=1,\cdots s)$ 的最大公因子是 1. 由 Bezout 定理, 存在 $u_i(x)\in F[x]$ 使得

$$u_1(x)D_1(x)+\cdots+u_s(x)D_s(x)=1.$$

令 $\alpha_i=u_i(\mathbf{A})D_i(\mathbf{A})\alpha$, 则 $\alpha=\alpha_1+\cdots+\alpha_s$, 且 $d_i(\mathbf{A})\alpha_i=0$. 这说明 $F[\mathbf{A}]\alpha$ 是 $F[\mathbf{A}]\alpha_1+\cdots+F[\mathbf{A}]\alpha_s$ 的子空间, 且 $\dim_F F[\mathbf{A}]\alpha_i\le\deg(d_i(x))$ 对每个 i 成立. 从而

$$\deg(d_\alpha(x))=\dim F[\mathbf{A}]\alpha\le\sum_{i=1}^s\dim F[\mathbf{A}]\alpha_i\le\sum_{i=1}^s\deg(d_i(x))=\deg(d_\alpha(x)).$$

故两不等号均是等号, 这就说明 $d_i(x)$ 是 α_i 的最小多项式, 且 $F[\mathbf{A}]\alpha$ 是 $F[\mathbf{A}]\alpha_i$ 的直和. □

4.1.3 根子空间与空间分解定理

要使 $\mathbf{A}$ 在某组基下的矩阵是 Jordan 形矩阵, 命题 4.1.6 说明这组基的向量需要是如下定义的根向量:

定义 4.1.5. 设 λ 是 $\mathbf{A}$ 的特征值. 非零向量 α 称为 λ 的 **根向量** (或者**广义特征向量** (generalized eigenvector) 是指存在正整数 k 使得 $(\mathbf{A}-\lambda\mathbf{I}_n)^k(\alpha)=0$. 更进一步地, α 称为 λ 的 k **次根向量**是指

$$(\mathbf{A}-\lambda\mathbf{I}_n)^k(\alpha)=0,\ (\mathbf{A}-\lambda\mathbf{I}_n)^{k-1}(\alpha)\ne 0. \tag{4.2}$$

为方便起见, 称零向量是 0 次根向量.

记 λ 的所有根向量与零向量构成的集合为 W_λ, 故

$$W_\lambda=\bigcup_{k\ge 0}\ker(\mathbf{A}-\lambda\mathbf{I}_n)^k \tag{4.3}$$

是 V 的子空间, 称为 λ 的**根子空间** (root space). 记所有 $\le k$ 次的根向量构成的子空间为 W_λ^k, 故

$$W_\lambda^k=\ker(\mathbf{A}-\lambda\mathbf{I}_n)^k. \tag{4.4}$$

引理 4.1.8. *下列结果成立:*

(1) W_λ 和 W_λ^k 都是 $\mathbf{A}$-不变子空间, 即 $\mathbf{A}(W_\lambda) \subset W_\lambda$, $\mathbf{A}(W_\lambda^k) \subset W_\lambda^k$.

(2) 存在 k 使得 $W_\lambda^k = W_\lambda^{k+1}$, 从而 $W_\lambda^k = W_\lambda^{k'} = W_\lambda$ 对所有 $k' \geq k$ 成立.

(3) 设 $d_\lambda = \dim W_\lambda$ 并设 k_λ 是最小的 k 使得 $W_\lambda^k = W_\lambda^{k+1}$, 则 $\mathbf{A}|_{W_\lambda}$ 的特征多项式是 $(x-\lambda)^{d_\lambda}$, 最小多项式是 $(x-\lambda)^{k_\lambda}$.

(4) 设 α 是 k 次根向量, 对于 $j \leq k$, 令 $\alpha_j = (\mathbf{A}-\lambda\mathbf{I}_n)^{k-j}(\alpha)$. 则 α_j 是 j 次根向量, $\{\alpha_1, \cdots, \alpha_k\}$ 是循环子空间 V_α 的一组基, 且限制变换 $\mathbf{A}|_{V_\alpha}$ 在此基下的矩阵是 $J_k(\lambda)$.

证明. 显然. □

定理 4.1.9 (根子空间分解定理, Root space decomposition)**.** 设 $\mathbf{A}$ 是 n 维 F-线性空间 V 上的线性变换, 并且它的特征值 $\lambda_1, \cdots, \lambda_s$ 全在 F 中. 设 n_i 是 λ_i 的代数重数. 则 W_{λ_i} 是 $\mathbf{A}$ 的 n_i 维不变子空间, $\mathbf{A}|_{W_{\lambda_i}}$ 的特征多项式为 $(x-\lambda_i)^{n_i}$ 且 $V = \bigoplus\limits_{i=1}^{s} W_{\lambda_i}$.

证明. 简记 W_{λ_i} 为 W_i. 首先证明 $V = W_1 + \cdots + W_s$. 设 $f_i(x) = \frac{\varphi_{\mathbf{A}}(x)}{(x-\lambda_i)^{n_i}}$. 则 $f_1, \cdots, f_s$ 的最大公因子是 1, 故由 Bezout 等式, 存在 F 上的多项式 $u_1(x), \cdots, u_s(x)$ 使得

$$u_1(x)f_1(x) + \cdots + u_s(x)f_s(x) = 1.$$

故对于任意 $\alpha \in V$, 我们有

$$u_1(\mathbf{A})f_1(\mathbf{A})(\alpha) + \cdots + u_s(\mathbf{A})f_s(\mathbf{A})(\alpha) = \alpha.$$

令 $u_i(\mathbf{A})f_i(\mathbf{A})(\alpha) = \alpha_i$. 由 Cayley-Hamilton 定理, $\varphi_{\mathbf{A}}(\mathbf{A}) = 0$, 故

$$(\mathbf{A} - \lambda_i\mathbf{I}_n)^{n_i}(\alpha_i) = u_i(\mathbf{A})\varphi_{\mathbf{A}}(\mathbf{A})(\alpha) = 0.$$

这说明 $\alpha_i \in W_i$, 从而 $V = W_1 + \cdots + W_s$.

我们暂时设 $d_i = \dim W_i$. 要证明 $V = \bigoplus\limits_{i=1}^{s} W_i$ 是直和, 设 $\alpha \in W_i \cap \sum\limits_{j\neq i} W_j$, 只要证明 $\alpha = 0$. 根据引理 4.1.8, 由 $\alpha \in W_i$ 知 $(\mathbf{A} - \lambda_i\mathbf{I}_n)^{d_i}(\alpha) = 0$, 由 $\alpha \in \sum\limits_{j\neq i} W_j$ 知 $\prod\limits_{j\neq i}(\mathbf{A} - \lambda_j\mathbf{I}_n)^{d_j}(\alpha) = 0$. 由于 $(x-\lambda_i)^{d_i}$ 与 $\prod_{j\neq i}(x-\lambda_j)^{d_j}$ 互素, 故存在 F 上的多项式 $v_1(x)$ 与 $v_2(x)$, 使得

$$v_1(x)(x-\lambda_i)^{d_i} + v_2(x)\prod_{j\neq i}(x-\lambda_j)^{d_j} = 1.$$

从而

$$\alpha = v_1(\mathbf{A})(\mathbf{A} - \lambda_i\mathbf{I}_n)^{d_i}(\alpha) + v_2(\mathbf{A})\prod_{j\neq i}(\mathbf{A} - \lambda_j\mathbf{I}_n)^{d_j}(\alpha) = 0.$$

我们最后要证明 $n_i = d_i$. 这是由于 $\mathbf{A}|_{W_i}$ 的特征多项式是 $(x-\lambda_i)^{d_i}$, 故

$$\varphi_{\mathbf{A}}(x) = \prod_{i=1}^{s}(x-\lambda)^{n_i} = \prod_{i=1}^{s}\varphi_{\mathbf{A}|_{W_i}}(x) = \prod_{i=1}^{s}(x-\lambda)^{d_i},$$

这说明 $n_i = d_i$. □

4.1.4 方阵的根向量和空间分解定理

对于域 F 上的 n 阶方阵 A, 令

$$\mathbf{A}: F^{n\times 1} \to F^{n\times 1},\ X \mapsto AX.$$

我们称 A 的不变子空间、循环子空间、根向量和根子空间为对应的线性变换 $\mathbf{A}$ 的不变子空间、循环子空间、根向量和根子空间. 具体地说,

定义 4.1.6. 设 A 是域 F 上的 n 阶方阵.

(1) 设 W 是列向量空间 $F^{n\times 1}$ 的子空间, 若 $A(W) = \{AX \mid X \in W\} \subseteq W$, 则称 W 为 A 的**不变子空间**.

(2) 设 $X \in F^{n\times 1}$, 线性空间 $F[A]X = \{f(A)X \mid f(x) \in F[x]\}$ 称为 X 生成的**循环子空间**.

(3) 设 $0 \neq X \in F^{n\times 1}$. 若存在正整数 k 使得 $(\lambda I_n - A)^k X = 0$, 则称 X 是关于特征向量 λ 的**根向量**. 所有关于 λ 的根向量集合 W_λ 构成的线性子空间称为 λ 的**根子空间**.

令 V_B 表示齐次线性方程组 $BX = 0$ 的解空间. 我们同样有

$$W_\lambda = \bigcup_{k\geq 1} V_{(\lambda I_n - A)^k} = V_{(\lambda I_n - A)^{n_\lambda}}. \tag{4.5}$$

定理 4.1.10 (空间分解定理, 矩阵形式)**.** 设 n 阶方阵 A 的特征值 $\lambda_1, \cdots, \lambda_s$ 全在 F 中. 设 n_i 是 λ_i 的代数重数. 则 W_{λ_i} 是 A 的 n_i 维不变子空间, $\mathbf{A}|_{W_{\lambda_i}}$ 的特征多项式为 $(x-\lambda_i)^{n_i}$ 且 $F^{n\times 1} = \bigoplus_{i=1}^{s} W_{\lambda_i}$.

习　　题

习题 4.1.1. 设 U 是 $\mathbf{A}: V \to V$ 的不变子空间. 证明: $\mathbf{A}(U) = \{\mathbf{A}(u) \mid u \in U\}$ 和 $\mathbf{A}^{-1}(U) = \{v \mid \mathbf{A}(v) \in U\}$ 都是 $\mathbf{A}$ 的不变子空间.

习题 4.1.2. 设线性变换 $\mathbf{A}$ 与 $\mathbf{B}$ 满足条件 $\mathbf{AB} - \mathbf{BA} = \mathbf{A}$. 证明对任意正整数 k, $\ker \mathbf{A}^k$ 和 $\mathrm{im}\mathbf{A}^k$ 都是 $\mathbf{B}$-不变子空间.

习题 4.1.3. 证明: 若 V 的任意子空间都是 $\mathbf{A}: V \to V$ 的不变子空间, 则 $\mathbf{A}$ 必为数乘变换.

习题 4.1.4. 设 $V = F^{3\times 3}$, $\mathbf{A}: V \to V, X \mapsto X + X^T$. 求 V 中所有 $\mathbf{A}$-不变子空间.

习题 4.1.5. 设 $\alpha_i \in V$ $(i = 1, \cdots s)$ 且 $d_{\mathbf{A},\alpha_i}(x)$ 两两两互素. 证明: $\alpha = \alpha_1 + \cdots + \alpha_s$ 的最小多项式 $d_{\mathbf{A},\alpha}(x) = \prod_{i=1}^{s} d_{\mathbf{A},\alpha_i}(x)$, 且循环子空间 $F[\mathbf{A}]\alpha$ 是 $F[\mathbf{A}]\alpha_i$ $(i = 1, \cdots s)$ 的直和.

习题 4.1.6. 设复线性空间 V 上的线性变换 $\mathbf{A}$ 与 $\mathbf{B}$ 可交换. 证明

(1) $\mathbf{A}$ 的特征子空间和根子空间都是 $\mathbf{B}$ 的不变子空间;

(2) $\mathbf{A}$ 与 $\mathbf{B}$ 有公共特征向量.

附录: 代数基本定理的线性代数证明

我们使用线性代数方法证明代数基本定理, 这个证明由美国数学家 K. Conrad 首先给出, 只用到特征值、特征向量和不变子空间的概念, 以及奇数次实多项式有实根这个数学分析熟知结果. 代数基本定理是说

定理 4.1.11. *所有次数大于 0 的复系数多项式都有复根. 即若 $n \geq 1$, $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{C}[x]$, 则存在 $z \in \mathbb{C}$, 使得 $f(z) = 0$.*

它等价于下面的定理:

定理 4.1.12. *任意 n 阶复方阵 A 必有复特征值和复特征向量. 等价地说, 若 V 是 n 维复空间, $\mathbf{A}$ 是 V 上的线性变换, 则 $\mathbf{A}$ 有复特征值和复特征向量.*

等价性的证明. 由定理 4.1.11推出定理 4.1.12是显然的. 另一方面, 取

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

是 $f(x)$ 的友阵, 即 A 的特征多项式 $\varphi_A(x) = f(x)$. 由定理 4.1.12, 它有复特征值, 这就说明 $f(x)$ 有复根. □

引理 4.1.13. *奇数次实多项式均有实根.*

证明. 由连续函数的介值定理立得. □

引理 4.1.14. *代数基本定理对于二次复多项式成立.*

证明. 这就是一元二次方程的求根公式. □

引理 4.1.15. *设 m 是大于 1 的整数, F 是域. 假设如下条件成立:*

> *任何维数 $\notin m\mathbb{Z}$ 的线性空间上的线性变换均有 F-特征值, 等价地说, F 上的任意阶 $\notin m\mathbb{Z}$ 的方阵均有 F-特征值.*

那么对于任意维数 $d \notin m\mathbb{Z}$ 的 F-线性空间 V, V 上任意两个互相交换的线性变换 $\mathbf{A}_1$ 和 $\mathbf{A}_2$ 在 V 中有公共特征向量.

证明. 我们对 d 做归纳来证明. 当 $d = 1$ 时, V 上的任意非零向量都是 $\mathbf{A}_1$ 和 $\mathbf{A}_2$ 的公共特征向量. 现在假设对于所有 $1 \leq d' < d$ 且 $m \nmid d'$, d' 维线性空间上交换的两个线性变换有公共特征向量.

由引理的假设条件, 设 $\lambda \in F$ 是 $\mathbf{A}_1$ 的一个特征值. 令

$$U = \mathrm{im}(\mathbf{A}_1 - \lambda \mathrm{id}_V), \quad W = \ker(\mathbf{A}_1 - \lambda \mathrm{id}_V).$$

则 U 和 W 都是 $\mathbf{A}_1$ 的不变子空间, 其中 $W \neq 0$ 是 λ 的特征子空间, 且

$$\dim U + \dim W = \dim V = d.$$

若 $W = V$, 则 $\mathbf{A}_1$ 是数乘变换, $\mathbf{A}_2$ 的任意特征向量都是 $\mathbf{A}_1$ 和 $\mathbf{A}_2$ 的公共特征向量. 故总可假设 $W \neq V$, 因此 $\dim U$ 和 $\dim W$ 均小于 d, 且至少有一个不被 m 整除.

对于 $w \in W$, $\mathbf{A}_1(w) = \lambda w$, 故 $\mathbf{A}_1(\mathbf{A}_2(w)) = \mathbf{A}_2(\lambda w) = \lambda \mathbf{A}_2(w)$, 所以 $\mathbf{A}_2(w) \in W$, 这说明 W 是 $\mathbf{A}_2$ 的不变子空间. 对于 $u = \mathbf{A}_1(v) - \lambda v \in U$, 则 $\mathbf{A}_2(u) = \mathbf{A}_1(\mathbf{A}_2(v)) - \lambda \mathbf{A}_2(v) \in U$, 故 U 是 $\mathbf{A}_2$ 的不变子空间.

现在考虑 $\mathbf{A}_1$ 和 $\mathbf{A}_2$ 在 U 和 W 上的限制变换. 由归纳假设, 它们在 U 或者 W 上有公共特征向量, 故在 V 上也有公共特征向量. □

根据引理 4.1.13, 引理 4.1.15的条件在 $F = \mathbb{R}$, $m = 2$ 时成立, 从而我们有

推论 4.1.16. *奇数维实线性空间上两个互相交换的线性变换有公共特征向量.*

定理 4.1.12的证明. 设 $n = 2^k n'$, 其中 n' 是奇数. 我们对 k 做归纳来证明任意 n 阶复方阵 A 都有复特征向量.

首先考虑 $k = 0$ 即 n 是奇数的情况. 记 $\mathcal{H}_n$ 为 (Hermite) 矩阵集合

$$\mathcal{H}_n = \{M \in M_n(\mathbb{C}) : M = M^H\},$$

此处 $M^H = \overline{M}^T$ 表示 M 的共轭转置. 则 $\mathcal{H}_n$ 是 n^2 维实线性空间, 它的一组基是 $\{E_{ii}, E_{ij} + E_{ji}, \sqrt{-1}(E_{ij} - E_{ji}) \mid 1 \leq i \leq n,\ j < i\}$. 定义 $\mathcal{H}_n$ 上的实线性变换 $\mathscr{L}_1$ 和 $\mathscr{L}_2$ 如下:

$$\mathscr{L}_1(B) = \frac{AB + BA^H}{2}, \qquad \mathscr{L}_2(B) = \frac{AB - BA^H}{2\sqrt{-1}}.$$

验证得 $\mathscr{L}_1\mathscr{L}_2 = \mathscr{L}_2\mathscr{L}_1$. 由于 n^2 是奇数, 由推论 4.1.16, 存在非零方阵 $B \in \mathcal{H}_n$ 和实数 λ_1, λ_2 使得

$$\mathscr{L}_1(B) = \lambda_1 B, \qquad \mathscr{L}_2(B) = \lambda_2 B.$$

因此 $AB = \mathscr{L}_1(B) + \sqrt{-1}\mathscr{L}_2(B) = (\lambda_1 + \sqrt{-1}\lambda_2)B$. 这说明 B 的任意非零列向量都是 A 关于特征值 $\lambda_1 + \sqrt{-1}\lambda_2$ 的特征向量. 这就证明了 $k = 0$ 时的情形.

假设任意 $2^l m$ $(0 \leq l < k,\ 2 \nmid m)$ 阶方阵均有复特征向量, 我们要证明 $n = 2^k n'$ $(2 \nmid n')$ 阶方阵 A 也有复特征向量. 令 $\mathcal{S}_n$ 表示 n 阶复对称方阵的集合 $\{M \in M_n(\mathbb{C}) : M = M^T\}$, 它是 $\frac{n(n+1)}{2}$ 维复线性空间. 考虑 $\mathcal{S}_n$ 上的线性变换 $\mathscr{M}_1$ 和 $\mathscr{M}_2$:

$$\mathscr{M}_1(B) = AB + BA^T, \qquad \mathscr{M}_2(B) = ABA^T.$$

同样验证得 $\mathscr{M}_1\mathscr{M}_2 = \mathscr{M}_2\mathscr{M}_1$. 注意到由 $2^k \| n$ 即得 $2^{k-1} \| \frac{n(n+1)}{2}$. 归纳假设说明引理 4.1.15的假设条件对于 $m = 2^k$, $F = \mathbb{C}$ 的情形成立. 由于 $m \nmid \frac{n(n+1)}{2}$, 故 $\mathscr{M}_1$ 和 $\mathscr{M}_2$ 有公共特征向量, 即存在方阵 $0 \neq B \in \mathcal{S}_n$ 和 $\lambda, \mu \in \mathbb{C}$, 使得

$$AB + BA^T = \lambda B, \qquad ABA^T = \mu B.$$

因此 $A^2B+ABA^T=A^2B+\mu B=\lambda AB$, 即 $(A^2-\lambda A+\mu I_n)B=0$. 由引理 4.1.14, 令 a 和 b 是 $x^2-\lambda x+\mu$ 的两个复根, 故得

$$(A-aI_n)(A-bI_n)B=0.$$

若 $(A-bI_n)B=0$, 则 B 的任意非零列向量是 A 关于特征值 b 的特征向量. 若 $B'=(A-bI_n)B\neq 0$, 则 B' 的任意非零列向量是 A 关于特征值 a 的特征向量. 这就完成了定理的证明. □

4.2 Jordan 标准形

4.2.1 Jordan 标准形主要定理

对任意方阵 A (包括零方阵!), 我们规定 A^0 是单位阵.

定义 4.2.1. 设 A 是域 F 上的 n 阶方阵, $k\in\mathbb{Z}+$, $\lambda\in F$. 记

$$r_k(A)=\mathrm{rank}A^k,\qquad r_k(\lambda,A)=r_k(A-\lambda I_n), \tag{4.6}$$

$$\delta_k(\lambda,A)=r_{k-1}(\lambda,A)+r_{k+1}(\lambda,A)-2r_k(\lambda,A). \tag{4.7}$$

这里规定 $r_0(\lambda,A)=n$.

定理 4.2.1 (Jordan 标准形, 矩阵形式)**.** *若域 F 上方阵 A 的特征值都在 F 中, 则 A 相似于 Jordan 形矩阵 J, 且在不计 Jordan 块次序的情况下 J 是唯一的. 更确切地说, J 中形如 $J_k(\lambda)$ 的块的个数等于 $\delta_k(\lambda,A)$, 总 λ-块的个数等于 λ 的几何重数.*

注记. (1) 根据代数基本定理, 所有复方阵都满足定理的条件, 从而每个复方阵都复相似于 Jordan 形矩阵. 这是我们在其他教科书上最常见到的 Jordan 标准形定理.

(2) 总可以将域 F 扩大让它包含方阵的所有特征值, 这样 Jordan 标准形定理在扩大的域里面成立, 例如将实矩阵看成复矩阵. 这时候需要注意运算是在扩域里进行的.

(3) 讨论相似时, 准对角块的次序是不重要的: 设 A_1 与 A_2 分别是 m 和 n 阶方阵, 则 $\mathrm{diag}(A_1,A_2)$ 与 $\mathrm{diag}(A_2,A_1)$ 通过 $\begin{pmatrix} & I_m \\ I_n & \end{pmatrix}$ 相似.

(4) 我们可以这样理解 J 中 λ-块的个数等于 λ 的几何重数这个性质: 根据之后命题 4.2.4的结果, 每一个 λ-块恰好贡献特征子空间 V_λ 中的一个 1 维子空间, 而这些 1 维子空间的直和就是 V_λ.

用线性变换的语言来表述, 定理 4.2.1 就等价于下述定理:

定理 4.2.2 (Jordan 标准形, 变换形式)**.** *设 $\mathbf{A}$ 是 F-线性空间 V 上的线性变换且它的特征值都在 F 中, 则存在 V 的一组基, 使得 $\mathbf{A}$ 在此基下的矩阵是 Jordan 形矩阵 J, 且 J 中形如 $J_k(\lambda)$ 的块的个数等于 $\delta_k(\lambda,\mathbf{A})$, λ-块的总个数等于 λ 的几何重数..*

定义 4.2.2. 上述两定理中的 Jordan 形矩阵 J 称为方阵 A 或者线性变换 $\mathbf{A}$ 的 **Jordan 标准形** (Jordan canonical form).

推论 4.2.3. 设方阵 A 的 *Jordan* 标准形中 λ_i-块中最大块的大小是 d_i. 则 A 的最小多项式 $d_A(x)=\prod_{i=1}^{s}(x-\lambda_i)^{d_i}$.

证明. 根据命题 4.2.4, $J_d(\lambda)$ 的最小多项式是 $(x-\lambda)^d$, 再由 Jordan 标准形定理立得. □

4.2.2 唯一性的证明

我们证明矩阵形式 Jordan 标准形的唯一性. 首先讨论 Jordan 块 $J_n(\lambda)=\lambda I_n+N_n$ 的性质:

命题 4.2.4. *Jordan* 块矩阵 $J_n(\lambda)$ 的特征多项式和最小多项式都是 $(x-\lambda)^n$, 它有唯一特征值 λ, 其特征子空间维数是 1, 且

$$r_k(\lambda,J_n(\lambda))=\begin{cases}n-k, & 若\ 1\le k<n,\\ 0, & 若\ k\ge n;\end{cases}\tag{4.8}$$

$$r_k(\lambda,J_n(\lambda'))=n\ (若\ \lambda\ne\lambda').\tag{4.9}$$

另外, $J_n(\lambda)$ 与自己的转置相似, 且对于 *Jordan* 块 $J_{k'}(\lambda')$,

$$\delta_k(\lambda,J_{k'}(\lambda'))=\begin{cases}1, & 若\ (k,\lambda)=(k',\lambda'),\\ 0, & 若\ (k,\lambda)\ne(k',\lambda').\end{cases}\tag{4.10}$$

证明. 等式 (4.10)由命题中关于秩的结果即得, 而这些结果只需要对 $N_n=J_n(0)$ 情形证明即可. 经计算立得

$$N_n^k=\begin{cases}\begin{pmatrix}0 & I_{n-k}\\ 0 & 0\end{pmatrix}, & 若\ 1\le k<n,\\ 0, & 若\ k\ge n.\end{cases}\tag{4.11}$$

故 N_n 的特征多项式和最小多项式都是 x^n, 它有唯一特征值 0, 其几何重数是 n, 代数重数即特征子空间的维数是 1, 且

$$r_k(N_n)=\begin{cases}n-k, & 若\ 1\le k<n,\\ 0, & 若\ k\ge n.\end{cases}\tag{4.12}$$

另外, 由等式

$$N_n^T=\begin{pmatrix}&&1\\&\iddots&\\1&&\end{pmatrix}N_n\begin{pmatrix}&&1\\&\iddots&\\1&&\end{pmatrix}=\begin{pmatrix}&&1\\&\iddots&\\1&&\end{pmatrix}^{-1}N_n\begin{pmatrix}&&1\\&\iddots&\\1&&\end{pmatrix}$$

即得 N_n^T 与 N_n 相似. □

注记. 当然, N_n 与 N_n^T 相似还可以用线性变换的语言来证明. 设 V 是 n 维线性空间, $\{e_1,\cdots,e_n\}$ 是它的一组基. 令

$$\mathbf{N}_n:V\to V,\ e_1\mapsto 0,\ e_2\mapsto 1,\ \cdots,\ e_n\mapsto e_{n-1}.$$

故

$$\mathbf{N}_n(e_1,\cdots,e_n)=(e_1,\cdots,e_n)N_n,$$

$$\mathbf{N}_n(e_n,\cdots,e_1)=(e_n,\cdots,e_1)N_n^T,$$

所以 N_n 与 N_n^T 相似. 事实上上面证明出现的矩阵就是基 $\{e_1,\cdots,e_n\}$ 到 $\{e_n,\cdots,e_1\}$ 的过渡矩阵.

由 Jordan 标准形定理, 命题立刻说明:

推论 4.2.5. *复方阵 A 与 A^T 总相似.*

命题 4.2.6. *关于 $r_k(\lambda,A)$ 与 $\delta_k(\lambda,A)$, 下列性质成立:*

(1) *若 A 与 B 相似, 则 $r_k(\lambda,A)=r_k(\lambda,B)$, 从而 $\delta_k(\lambda,A)=\delta_k(\lambda,B)$.*

(2) *若 λ 不是 A 的特征值或者 $k>n$, 则 $\delta_k(\lambda,A)=0$.*

(3) *若 λ 是 A 的特征值, 则 $\sum\limits_{k=1}^{n}\delta_k(\lambda,A)=\dim V_\lambda$ 即 λ 的几何重数.*

(4) *Jordan 形矩阵 J 中 $J_k(\lambda)$-块的个数等于 $\delta_k(\lambda,J)$, 总 λ-块的个数等于 λ 的几何重数. 故 Jordan 形矩阵 J_1 与 J_2 相似当且仅当 $\delta_k(\lambda,J_1)=\delta_k(\lambda,J_2)$ 对于任意 $k\in\mathbb{Z}_+$ 与 $\lambda\in F$ 均成立.*

证明. (1) 由定义立得.

(2) 若 λ 不是 A 的特征值, 则 $A-\lambda I_n$ 可逆, 故 $r_k(\lambda,A)$ 恒等于 n, 因此 $\delta_k(\lambda,A)=0$.

设 $k>n$. 我们知道对任意方阵 B, 若 $r_i(B)=r_{i+1}(B)$, 则 $r_j(B)=r_i(B)$ 对所有 $j\geq i$ 成立. 由于 $r_i(\lambda,A)$ $(0\leq i\leq n+1)$ 不能一直严格递降, 故总存在 $i\leq n<k$, 使得 $r_i(\lambda,A)=r_{i+1}(\lambda,A)$, 这说明

$$r_i(\lambda,A)=r_{k-1}(\lambda,A)=r_k(\lambda,A)=r_{k+1}(\lambda,A),$$

因此 $\delta_k(\lambda,A)=0$.

(3) 根据 (2) 的证明, $r_n(\lambda,A)=r_{n+1}(\lambda,A)$ 恒成立. 故

$$\begin{aligned}\sum_{k=1}^{n}\delta_k(\lambda,A)=&n-r_1(\lambda,A)-r_n(\lambda,A)+r_{n+1}(\lambda,A)\\=&n-\operatorname{rank}(A-\lambda I_n)=\dim V_\lambda.\end{aligned}$$

(4) 由于 $J=\operatorname{diag}(J_{n_1}(\lambda_1),\cdots,J_{n_s}(\lambda_s))$ 是准对角阵, 故

$$\delta_k(\lambda,J)=\sum_{i=1}^{s}\delta_k(\lambda,J_{n_i}(\lambda_i)).$$

因此 (4) 由(4.10) 立得. □

注记. 事实上总有 $\delta_k(\lambda,A)\geq 0$. 这可由不等式

$$\operatorname{rank}B^k-\operatorname{rank}B^{k+1}\leq\operatorname{rank}B^{k-1}-\operatorname{rank}B^k$$

得到. 如假设 A 相似于 Jordan 形矩阵, 它也可由 (1) 和 (4) 导出.

注记. 根据线性变换与方阵的对应关系, 对于任意线性变换 $\mathbf{A}$, 类似定义 $r_k(\lambda, \mathbf{A})$ 和 $\delta_k(\lambda, \mathbf{A})$. 类似上述命题, 容易给出它们所满足的性质.

Jordan 标准形唯一性的证明. 这是命题 4.2.6 的直接推论. □

4.2.3 存在性的证明

我们证明线性变换形式 Jordan 标准形的存在性. 首先证明幂零变换情形, 这需要如下引理:

引理 4.2.7. *设 $\mathbf{A}$ 是复线性空间 V 上的幂零变换, 最小多项式是 x^k 且 $k \geq 2$. 记 $V_i = \ker \mathbf{A}^i$. 设 B_0 为空集, 设 B_{k-2}, $B_{k-2} \cup S_{k-1}$ 和 $B_{k-2} \cup S_{k-1} \cup S_k$ 分别是 V_{k-2}, V_{k-1} 和 V_k 的基. 则 $\mathbf{A}(S_k) \subset V_{k-1}$ 且它与 B_{k-2} 线性无关.*

证明. $k = 2$ 的情形是平凡情形. 现在设 $k \geq 3$. 考虑 $\alpha \in \langle S_k \rangle$. 则 $0 = \mathbf{A}^k(\alpha) = \mathbf{A}^{k-1}(\mathbf{A}(\alpha))$, 这说明 $\mathbf{A}(\alpha) \in V_{k-1}$, 故 $\mathbf{A}(S_k) \subset V_{k-1}$. 若 $\mathbf{A}(\alpha) \in \langle B_{k-2} \rangle = V_{k-2}$, 则 $\mathbf{A}^{k-2}(\mathbf{A}(\alpha)) = \mathbf{A}^{k-1}(\alpha) = 0$, 故 $\alpha \in V_{k-1} = \langle B_{k-2}, S_{k-1} \rangle$. 但 $B_{k-2} \cup S_{k-1} \cup S_k$ 是 V_k 的基, 这说明 $\alpha \in \langle S_k \rangle \cap V_{k-1} = \{0\}$. □

定理 4.2.8. *设 $\mathbf{A}$ 是 F-线性空间 V 上的幂零变换, 则存在 V 的一组基使得 $\mathbf{A}$ 在此基下的矩阵是 $\mathrm{diag}(N_{k_1}, \cdots, N_{k_t})$ 的形式.*

证明. 令 $V_i = \ker \mathbf{A}^i$. 不妨设 k 是最小的正整数使得 $V_k = V$. 由于 $V_i \subset V_{i+1}$ 且若 $V_i = V_{i+1}$, 则必有 $V_i = V_{i+1} = \cdots = V$, 故我们有 $V_1 \subsetneq V_2 \cdots \subsetneq V_k = V$, 这说明 x^i $(i \leq k)$ 是 $\mathbf{A}|_{V_i}$ 的最小多项式. 如果 $k = 1$, $\mathbf{A}$ 是零变换, 定理显然成立, 故不妨设 $k \geq 2$.

设 $B_1 = S_1$ 是 V_1 的一组基, 依次扩充, 使得 $B_i = S_1 \cup \cdots \cup S_i (i \leq k)$ 是 V_i 的一组基. 设 B_0 为空集合. 我们将 B_k 改造成 V_k 的一组新的基.

首先令 $\tilde{S}_k = S_k = S_k^*$. 由引理 4.2.7 知 $\mathbf{A}(\tilde{S}_k) \subset V_{k-1}$ 与 B_{k-2} 线性无关, 我们扩充 $B_{k-2} \cup \mathbf{A}(\tilde{S}_k)$ 为 V_{k-1} 的一组基 $B_{k-2} \cup \mathbf{A}(\tilde{S}_k) \cup S_{k-1}^*$ 并令 $\tilde{S}_{k-1} = \mathbf{A}(\tilde{S}_k) \cup S_{k-1}^*$.

从 $i = k$ 开始递降到 $i = 2$ 为止, 假设已经构造了 $\tilde{S}_i$ 使得对于 $j \geq i$, $B_{i-1} \cup \tilde{S}_i \cup \cdots \cup \tilde{S}_j$ 是 V_j 的基. 由引理 4.2.7 则 $\mathbf{A}(\tilde{S}_i) \subset V_{i-1}$ 与 B_{i-2} 线性无关, 我们扩充 $B_{i-2} \cup \mathbf{A}(\tilde{S}_i)$ 为 V_{i-1} 的一组基 $B_{i-2} \cup \mathbf{A}(\tilde{S}_i) \cup S_{i-1}^*$ 并令 $\tilde{S}_{i-1} = \mathbf{A}(\tilde{S}_i) \cup S_{i-1}^*$.

根据构造, 我们得到 V 的一组新基 $\tilde{S}_1 \cup \cdots \cup \tilde{S}_k$, 其中

$$\tilde{S}_i = \mathbf{A}^{k-i}(S_k^*) \cup \mathbf{A}^{k-1-i}(S_{k-1}^*) \cup \cdots \cup S_i^*,$$

记 $S_i^* = \{\alpha_{ij} \mid 1 \leq j \leq j_i\}$, 并令

$$T_i = \{\mathbf{A}^{i-1}(\alpha_{i1}), \mathbf{A}^{i-2}(\alpha_{i1}), \cdots, \alpha_{i1}, \cdots, \mathbf{A}^{i-1}(\alpha_{ij_i}), \mathbf{A}^{i-2}(\alpha_{ij_i}), \cdots, \alpha_{ij_i}\}.$$

那么

$$T_k \cup T_{k-1} \cup \cdots \cup T_1$$

还是 V 的基, 且在此基下, $\mathbf{A}$ 的矩阵就是定理所要求的形式. □

注记. 读者还可以通过商空间来构造向量集合 S_i^*. 考虑线性映射 $f_i: V_{i+1}/V_i \to V_i/V_{i-1}$, $\overline{\alpha} \mapsto \overline{\mathbf{A}(\alpha)}$, 它是单射. 将 $\mathrm{im}(f_i)$ 的一组基扩充为 V_i/V_{i-1} 的基, 扩充进来新向量组记为 $\overline{S_i^*}$, 那么只要取 $S_i^* \subseteq V_i$ 为 $\overline{S_i^*}$ 的一组逆元即可.

Jordan 标准形存在性的证明. 设 $\mathbf{A}$ 是 V 上的线性变换且特征值均在 F 中. 记 $W_i = W_{\lambda_i}$. 根据空间分解定理, $V = \bigoplus_{i=1}^{s} W_i$ 且 $\mathbf{A}|_{W_i}$ 的特征多项式为 $(x-\lambda_i)^{n_i}$. 故 $\mathbf{A}|_{W_i} - \lambda_i \mathrm{id}_{W_i}$ 是 W_i 上的幂零变换. 根据定理 4.2.8, 存在 W_i 的一组基 B_i, $\mathbf{A}|_{W_i} - \lambda_i \mathrm{id}_{W_i}$ 在此基下的矩阵为 $\mathrm{diag}(N_{k_1}, \cdots, N_{k_t})$ 的形式, 从而 $\mathbf{A}|_{W_i}$ 在此基下的矩阵 A_i 是 $\mathrm{diag}(J_{k_1}(\lambda_i), \cdots, J_{k_t}(\lambda_i))$ 的形式. 故 $\mathbf{A}$ 在基 $B = B_1 \cup \cdots \cup B_s$ 下的矩阵为 $\mathrm{diag}(A_1, \cdots, A_s)$, 它是 Jordan 形矩阵. □

4.2.4 Jordan 标准形的计算

我们根据前面的理论证明来分析一下 Jordan 标准形的计算问题, 也就是说, 给定域 F 上的 n 阶方阵 A, 如何通过具体计算, 求得它的 Jordan 标准形 J (不计 Jordan 块的次序), 并求得可逆矩阵 P, 使得 $P^{-1}AP = J$.

首先, Jordan 标准形 J 的存在性基于如下条件: **A 的特征值即特征多项式 $\varphi_A(x)$ 的所有根都需要在 F 中**. 对于 F 是复数域的情形, 由代数基本定理, 这个条件自然满足. 对于其他情形, 这就需要通过全面确定特征多项式 $\varphi_A(x)$ 根的位置来判定. 若假设已知 J 的存在, 要求出它的具体形式, 就必须知道 A 的所有特征值, 也就是知道特征多项式 $\varphi_A(x)$ 的所有根 (从而也知道了根的重数即特征值 λ 的代数重数 n_λ). 这是 Jordan 标准形的计算必经的第一步骤. 由于伽罗瓦的著名定理 (次数 ≥ 5 的一般复系数多项式没有通用求根公式), 实际计算中, 要么 n 较小 (比如 $n \leq 4$), 要么 $\varphi_A(x)$ 的根比较容易求出.

其次, 如果已经知道特征值 λ, 就可以计算 $r_k(\lambda, A) = \mathrm{rank}(\lambda I_n - A)^k$ $(1 \leq k \leq n_\lambda - 1)$, 从而得到 $\delta_k(\lambda, A)$, 这是 J 中 $J_k(\lambda)$-块的个数. 遍历所有的特征值 λ 就得到 Jordan 标准形 J. 其实, 如果 n_λ 不大 (例如 ≤ 5), 通过计算得到 λ 的几何重数 $m_\lambda = n - r_1(\lambda, A)$, 即 J 中 λ-块的个数, 而 n_λ 等于各 λ-块的阶之和. 这样的表达式 (不计次序) 只有很少几种情形, 往往只需要再计算 $r_2(\lambda, A)$ 就可以确定具体情形, 即得到每个 λ-块的阶的大小.

最后考虑如何求 P. 将 P 写成有序列向量的形式 $P = (X_1, \cdots, X_n)$. 由 $A(X_1, \cdots, X_n) = (X_1, \cdots, X_n)J$ 知 X_i 必须是 A 的根向量, 且 $(X_1, \cdots, X_n)$ 是按 J 中 Jordan 块次序相应的有序向量组的有序并, 其中块 $J_k(\lambda)$ 对应的有序列向量组是 $((A-\lambda I_n)^{k-1}X, (A-\lambda I_n)^{k-2}X, \cdots, X)$ 且 X 是 λ 的 k 次根向量. 这里需要好选择好来保证不同的 λ-块对应的向量组线性无关, 这正是定理 4.2.8的证明中细心处理的内容. 同样, 在 n_λ 不大时, 有序向量组的选择是容易处理的.

将上面的叙述总结一下, 我们得到如下算法:

算法 4.2.9 (Jordan 标准形算法)**.** 输入: 域 F 上 n 阶方阵 A.

(1) 求特征多项式 $\varphi_A(x)$, 以及所有特征值 λ_i 和对应的代数重数 n_i $(i = 1, \cdots, s)$. 若特征值不全在 F 中, 则 A 在 F 中没有 Jordan 标准形.

(2) 对每个 i, 计算 $r_k(\lambda, A) = \mathrm{rank}(\lambda I_n - A)^k$ $(1 \leq k \leq n_i)$ 并求得 $\delta_k(\lambda_i, A)$. 则 J 是由 $\delta_k(\lambda_i, A)$ 个 Jordan 块 $J_k(\lambda_i)$ $(i = 1, \cdots, s)$ 组成的 Jordan 形矩阵.

(3) 对每个 i, 求根子空间 W_i 的一组基, 它是形如 $((A-\lambda_i I_n)^{k-1}X, (A-\lambda_i I_n)^{k-2}X, \cdots, X)$ (其中 X 是 λ_i 的 k 阶根向量) 的有序列向量组之并. 将所有这些列向量依次横排, 即得到矩阵 P 使得 $P^{-1}AP = J$.

输出: Jordan 标准形 J 和可逆矩阵 P 使得 $P^{-1}AP = J$.

4.2.5 一些例子

例 4.2.1. 求 $A = \begin{pmatrix} 3 & -2 & 1 \\ 2 & -2 & 2 \\ 3 & -6 & 5 \end{pmatrix}$ 的 *Jordan* 标准形及最小多项式.

解. 经计算得 $\varphi_A(x) = (x-2)^3$. 故 A 只有唯一的特征值 2, 它的 Jordan 标准形 J 的可能形式为

$$J_1 = 2I_3, \quad J_2 = \begin{pmatrix} J_2(2) & \\ & 2 \end{pmatrix} \quad 和 \quad J_3 = J_3(2).$$

由于只有 J_2 满足

$$\operatorname{rank}(A-2I) = \operatorname{rank}(J-2I) = 1,$$

故 A 的 Jordan 标准形 $J = J_2 = \operatorname{diag}(J_2(2), 2)$, 最小多项式 $d_A(x) = (x-2)^2$. □

例 4.2.2. 求 $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ 的 *Jordan* 标准形.

解. 容易看出 $\varphi_A(x) = (x-1)^4$. 故 A 只有唯一的特征值 1, 它的 Jordan 标准形 J 只能是下面五者之一:

$$I_4, \begin{pmatrix} J_2(1) & \\ & I_2 \end{pmatrix}, \begin{pmatrix} J_2(1) & \\ & J_2(1) \end{pmatrix}, \begin{pmatrix} J_3(1) & \\ & 1 \end{pmatrix} \text{ 和 } J_4(1).$$

由于只有 $\operatorname{diag}(J_2(1), J_2(1))$ 满足

$$\operatorname{rank}(J-I) = \operatorname{rank}(A-I) = 2 \text{ 且 } \operatorname{rank}(J-I)^2 = \operatorname{rank}(A-I)^2 = 0,$$

故 $J = \operatorname{diag}(J_2(1), J_2(1))$. □

例 4.2.3. 设 $1 \le m < n$. 试求 n 阶方阵 $A = \begin{pmatrix} O & I_{n-m} \\ O & O \end{pmatrix}$ 的 *Jordan* 标准形.

解法 1. 注意到矩阵 $A = N_n^m$. 设 $n = (q+1)m - r$ $(0 \le r < m)$, 则 $A^{q+1} = 0$, 而对于 $1 \le k \le q$, $A^k = N_n^{km}$, 它的秩是 $n - km > 0$. 故 A 的最小多项式是 x^{q+1}, 唯一的特征值 $\lambda = 0$, 且

$$\delta_k(0, A) = \begin{cases} 0, & 如\ k < q\ 或\ k > q+1; \\ r, & 如\ k = q; \\ m-r, & 如\ k = q+1. \end{cases}$$

故 A 的 Jordan 标准形为 $J=\mathrm{diag}(\underbrace{N_{q+1},\cdots,N_{q+1}}_{m-r},\underbrace{N_q,\cdots,N_q}_{r})$. □

解法 2. 令 $\mathbf{A}:F^{n\times 1}\to F^{n\times 1}$, $X\mapsto AX$. 则 $\mathbf{A}$ 在标准基 $\{e_1,e_2,\cdots,e_n\}$ 下的矩阵是 A, 而在基 $\{e_1,e_{m+1},\cdots,e_{qm+1},\cdots e_{m-r},e_{2m-r},\cdots,e_n,e_{m-r+1},\cdots\}$ 下的矩阵就是 $J=\mathrm{diag}(\underbrace{N_{q+1},\cdots,N_{q+1}}_{m-r},\underbrace{N_q,\cdots,N_q}_{r})$. 注意到这个解法实际上给出了 A 到 J 的过渡矩阵 P. □

例 4.2.4. 设 $A=\begin{pmatrix}4&3&-4\\-1&0&2\\1&1&0\end{pmatrix}$. *求可逆矩阵 P, 使得 $P^{-1}AP$ 是 Jordan 形矩阵.*

解. 首先计算 A 的特征多项式, 得 $\varphi_A(x)=(x-1)^2(x-2)$. 这说明 A 有两个特征值, 1 和 2, 其中特征值 1 的代数重数是 2, 而特征值 2 的代数重数和几何重数都是 1, 因此以 2 为中心的 Jordan 块只有 1 块.

求解齐次线性方程组 $(A-2I)X=0$, 即

$$\begin{pmatrix}2&3&-4\\-1&-2&2\\1&1&-2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix},$$

它的解空间是 1 维的, 由 $(2,0,1)^T$ 生成.

求解线性方程组 $(A-I)X=0$, 它的解空间也是 1 维的, 由 $(1,-1,0)^T$ 生成. 这样 1 的几何重数也是 1. 这说明 1 为中心的 Jordan 块只有 1 块.

根据 Jordan 标准形定理, A 就相似于 $J=\mathrm{diag}(2,J_2(1))$. 如果 $P=(X_1,X_2,X_3)$ 满足 $P^{-1}AP=J$, 即 $AP=PJ$, 展开, 即

$$AX_1=2X_1,\quad AX_2=X_2,\quad AX_3=X_2+X_3.$$

这说明 X_1 是 2 的特征向量, 可取为 $X_1=(2,0,1)^T$; X_2 是 1 的特征向量, 可取为 $X_2=(1,-1,0)^T$. 最后只要求线性方程组 $(A-I)X=X_2$, 即

$$\begin{pmatrix}3&3&-4\\-1&-1&2\\1&1&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}1\\-1\\0\end{pmatrix},$$

得到一个解 $X_3=(-1,0,-1)^T$. 这样

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

例 4.2.5. 求

$$A=\begin{pmatrix}1&1&1&1&1&1\\0&2&0&1&2&3\\0&0&2&0&1&-1\\0&0&0&2&1&2\\0&0&0&0&2&0\\0&0&0&0&0&2\end{pmatrix}$$

的 *Jordan* 标准形 J, 并求可逆矩阵 P, 使得 $P^{-1}AP=J$.

解. 首先 $\varphi_A(x)=(x-1)(x-2)^5$, 故 A 的特征值为 1 (1 重) 和 2 (5 重).

经计算可得

$$A-2I=\begin{pmatrix}-1&1&1&1&1&1\\0&0&0&1&2&3\\0&0&0&0&1&-1\\0&0&0&0&1&2\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix},\quad (A-2I)^2=\begin{pmatrix}1&-1&-1&0&3&3\\0&0&0&0&1&2\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix},$$

$$(A-2I)^3=\begin{pmatrix}-1&1&1&0&-2&-1\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\\0&0&0&0&0&0\end{pmatrix}.$$

由于 $\operatorname{rank}(A-2I)=4$, 故 J 中 2-块的个数是 $6-4=2$, 所以 2-块的大小应该是 $\{1,4\}$ 或者 $\{2,3\}$. 再由于 $\operatorname{rank}(A-2I)^2=2$, 故 $\{1,4\}$ 不可能. 因此 $J=\operatorname{diag}(1,J_3(2),J_2(2))$.

现在需要求 $P=(X_1,X_2,X_3,X_4,X_5,X_6)$, P 可逆且 $AP=PJ$.

由 $AX_1=X_1$, 知 X_1 是 1 的特征向量, 只要求解 $(A-I)X=0$, 可取 $X_1=(1,0,0,0,0,0)^T$.

根据 Jordan 标准形定理存在性的证明, 其他 X_i 可以如下选择. 注意到 $\ker(A-2I)$, $\ker(A-2I)^2$ 和 $\ker(A-2I)^3$ 分别是 $F^{6\times1}$ 的 2, 4 和 5 维子空间. 首先求解 $(A-2I)^3X=0$, 得到它的一组基础解系, 其中一定有 (可能多个) 向量满足 $(A-2I)^2X\neq0$. 任取其中一个, 记为 X_4, 比如

$$X_4=(0,0,0,0,1,-2)^T.$$

令 $X_2=(A-2I)^2X_4$, $X_3=(A-2I)X_4$. 将 $\{X_2,X_3\}$ 扩充为 $(A-2I)^2X=0$ 的基础解系. 扩充进来的两个向量中必有一个满足 $(A-2I)X\neq0$, 取它为 X_6, 比如

$$X_6=(0,0,3,0,2,-1)^T.$$

令 $X_5=(A-2I)X_6$. 则 P 满足要求. 在我们的选择中,

$$P=\begin{pmatrix}1&-3&-1&0&4&0\\0&-3&-4&0&1&0\\0&0&3&0&3&3\\0&0&-3&0&0&0\\0&0&0&1&0&2\\0&0&0&-2&0&-1\end{pmatrix}. \qquad \square$$

例 4.2.6. 证明任意复方阵 $A\in\mathbb{C}^{n\times n}$ 均相似于上三角矩阵 $B=(b_{ij})$, 其中 $|b_{ij}|<\varepsilon,\ i\neq j$.

证明. 根据 Jordan 标准形定理, 我们不妨假设 $A=J_n(\lambda)$ 是 Jordan 块矩阵, 只需证明对任意 $\varepsilon>0$, $J_n(\lambda)$ 相似于 $\lambda_n+\varepsilon N_n$ 即可. 令 $P=\mathrm{diag}(1,\varepsilon,\cdots,\varepsilon^{n-1})$, 则 $P^{-1}J_n(\lambda)P=\lambda I_n+\varepsilon N_n$. $\square$

例 4.2.7. 设 6 方阵 A 的最小多项式 $d_A(x)=x^3$. 求 A 的所有可能的 *Jordan* 标准形.

解. 条件说明 A 只有一个特征值 0 且最大 Jordan 块的大小是 3. 我们只要看 6 表示为正整数之和且和项里中最大项为 3 的写法:

$$6=3+3=3+2+1=3+1+1+1.$$

故 A 的 Jordan 标准形有三种可能, 即 $\mathrm{diag}(N_3,N_3)$, $\mathrm{diag}(N_3,N_2,0)$ 和 $\mathrm{diag}(N_3,0,0,0)$. $\square$

注记. 同样, 若 n 阶方阵 A 只有一个特征值 $\lambda\in F$, 则 A 的可能 Jordan 标准形数就等于 n 表为正整数之和的个数, 也就是 n 的分拆数 $p(n)$. 更进一步, 若 A 的最小多项式是 $(x-\lambda)^k$, 则所有可能的 Jordan 标准形数等于 n 表为最大项为 k 的正整数之和的个数.

习 题

习题 4.2.1. 已知 5 阶方阵只有两个特征值, 且满足

$$\mathrm{rank}A=3,\ \mathrm{rank}A^2=2,\ \mathrm{rank}(A+I)=4,\ \mathrm{rank}(A+I)^2=3.$$

试问它是否存在 *Jordan* 标准形? 如存在, 求出它的 *Jordan* 标准形.

习题 4.2.2. 求复可逆矩阵 P 与 Jordan 标准形 J, 将下列方阵相似为 J:

(1) $\begin{pmatrix} 6 & 5 & -2 \\ -2 & 0 & 1 \\ -1 & -1 & 3 \end{pmatrix}$, (2) $\begin{pmatrix} -2 & 1 & 3 \\ -22 & 11 & 33 \\ 4 & -2 & -6 \end{pmatrix}$,

(3) $\begin{pmatrix} 1 & 3 & 4 & 7 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, (4) $\begin{pmatrix} 4 & 3 & 0 & 0 \\ -3 & -2 & 0 & 0 \\ 1 & 2 & -3 & -2 \\ 3 & 4 & 8 & 5 \end{pmatrix}$,

(5) $\begin{pmatrix} & & & a_1 \\ & & a_2 & \\ & \iddots & & \\ a_n & & & \end{pmatrix}$, (6) $\begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{pmatrix}$.

习题 4.2.3. 设方阵 A 的特征多项式 $\varphi_A(x) = (x-1)^4(x+1)^3x^2$, 试求 A 所有可能的 Jordan 标准形, 并求每种情形时 A 的最小多项式.

习题 4.2.4. 设 9 阶方阵 A 的最小多项式 $d_A(x) = (x-1)^2(x+1)^2x^2$, 试求 A 所有可能的 Jordan 标准形, 并求每种情形时 A 的特征多项式.

习题 4.2.5. 已知 n 阶方阵 A 相似于 Jordan 形矩阵 J, 且 $A^{n-2} \neq 0$ 但 $A^{n-1} = 0$, 试求 J.

习题 4.2.6 (Jordan 分解)**.** 证明: 任意复方阵 A 可以写为 $A = B + C$, 其中 B 相似于对角阵, C 是幂零阵且 $BC = CB$.

习题 4.2.7. 证明若 A 的特征值都是 1, 则 A 与 A^k 相似, 其中 k 是任意非零整数.

习题 4.2.8. 若方阵 A 相似于 $\operatorname{diag}\left(\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), I_{n-2}\right)$, 则称 A 为反射方阵. 证明: 若 $A^2 = I$, 则 A 可分解为有限个反射方阵之积.

习题 4.2.9. 给出 4 个互不相似的 4 阶方阵 A, 满足条件 $(A-I)^3 = 0$.

习题 4.2.10. 证明若对所有非零整数 k, A 与 A^k 相似, 则 A 的特征值只能是 0 或者 1.

习题 4.2.11. 求所有与 A^2 相似的复方阵 A.

4.3 Smith 标准形理论

4.3.1 整环上的矩阵

设 R 是整环, 为简单起见读者可以假设 R 是整数环 $\mathbb{Z}$ 或者域上的多项式环 $F[x]$. 环 R 上的 $m \times n$ 矩阵集合记为 $R^{m\times n}$.

与域上的矩阵一样, 我们定义 R 上矩阵的加法、数乘和乘法运算, 以及方阵的行列式和余子式等. 在研究域上矩阵的性质时, 只要性质的叙述和证明中**只用到域的加减法和乘法运算, 而不涉及除法运算**, 那么这些性质就可以平行地推广到整环的矩阵上来, 特别地, 我们有:

定理 4.3.1. 对于整环上的矩阵, 下列性质成立.

(1) 矩阵乘法满足结合律, 即 $(AB)C = A(BC)$.

(2) 行列式是积性函数, 即 $\det(AB) = \det(A)\det(B)$.

(3) 行列式的 *Laplace* 展开公式成立.

(4) *Binet-Cauchy* 公式成立.

(5) 设 A 是 n 阶方阵, A^* 是它的伴随阵. 则 $AA^* = A^*A = \det(A)I_n$.

定义 4.3.1. 矩阵 A 的**秩**即它的极大非零子式的阶.

方阵 A **可逆**是指存在 R 上的方阵 B 使得 $AB = BA = I_n$, 此时 B 是唯一的, 称为 A 的**逆**, 记为 A^{-1}.

引理 4.3.2. 方阵 A 可逆当且仅当 $\det(A) \in R^\times$, 即 $\det(A)$ 是乘法可逆元.

证明. 若 A 可逆, $AA^{-1} = I_n$, 故 $\det(A)\det(A^{-1}) = 1$, 因此 $\det(A) \in R^\times$. 反之, 若 $\det(A) \in R^\times$, 令 $B = \frac{1}{\det(A)}A^*$, 则 $AB = BA = I_n$, 即 A 可逆. □

例 4.3.1. 考虑 $R = \mathbb{Z}$ 和 $F[x]$ 这两种常见情况.

(1) 当 $R = \mathbb{Z}$ 时, 整系数方阵 A 可逆当且仅当 $\det(A) \in \mathbb{Z}^\times = \{\pm 1\}$. 需要注意 A 作为整系数矩阵可逆和作为有理系数矩阵可逆是不一样的, 例如 $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ 作为有理系数矩阵可逆, 但作为整系数矩阵不可逆.

(2) 当 $R = F[x]$ 时, 多项式方阵 A 可逆当且仅当 $\det(A) \in (F[x])^\times = F^\times$.

显然, 方阵可逆则必满秩.

例 4.3.2. 设 A 是 F 上的 n 阶方阵. 则 $xI_n - A$ 是 $F[x]$ 上的方阵, 它的行列式就是 A 的特征多项式 $\varphi_A(x) \neq 0$, 故 $xI_n - A$ 的秩是 n, 即它是满秩方阵, 但它不是可逆阵. 这说明对于多项式方阵而言, 满秩不保证可逆. 更一般地, 只要整环 R 不是域, 满秩与可逆就是两个不同的概念.

与域的情形一样, 整环上矩阵的初等变换和它对应的初等矩阵 (即对单位阵做此初等变换得到的矩阵) 如下给出:

(I) 第一类初等变换: 互换两行或者两列. 此变换的逆变换是它自身, 对应的初等矩阵为 P_{ij}.

(II) 第二类初等变换: 对某行或列乘以 λ 倍, 其中 $\lambda \in R^\times$. 此变换的逆变换就是对该行或列乘以 λ^{-1} 倍, 对应的初等矩阵为 $D_i(\lambda)$.

(III) 第三类初等变换: 将某行或列的 a 倍 $(a \in R)$ 加到另外一行或者列. 此变换的逆变换就是对第一行或列乘以 $-a$ 倍加到后面一行或者列, 此时的初等矩阵为 $T_{ij}(a)$.

注意到矩阵的初等变换是可逆变换, 且逆变换是相同类型的变换. 可以看出第一类和第三类初等变换与域上矩阵情形并无不同, 但对于第二类初等变换, 为保证它是可逆变换, 我们需要 λ 是 R 上的乘法可逆元这个条件.

命题 4.3.3. (1) *初等矩阵都是可逆阵:*

$$P_{ij}^{-1} = P_{ij}, \quad D_i(\lambda)^{-1} = D_i(\lambda^{-1}), \quad T_{ij}(a)^{-1} = T_{ij}(-a).$$

(2) *对矩阵做初等行变换相当于对该矩阵左乘相应的初等矩阵, 对矩阵做初等列变换相当于对该矩阵右乘相应的初等矩阵.*

4.3.2 整系数和多项式系数矩阵的相抵

本节自此起设 $R = \mathbb{Z}$ 或者 $F[x]$, 即矩阵是整系数或多项式系数矩阵.

定义 4.3.2. 设 $a \in R$. 称

$$h(a) = \begin{cases} |a|, & \text{若 } R = \mathbb{Z}; \\ \deg(a), & \text{若 } R = F[x] \end{cases}$$

为 a 的**高度**. 若 $a > 0$ ($R = \mathbb{Z}$ 时) 或者 a 首一 ($R = F[x]$ 时), 称 a **恒正**.

R 上的带余除法就是说, 对于任意 $a, b \in R$ 且 $a \neq 0$, 存在唯一的商 q 和余数 (式) r, 使得

$$b = qa + r, \qquad h(r) < h(b). \tag{4.13}$$

定义 4.3.3. 设 A 是 R 上 $m \times n$ 矩阵. 对于 $1 \leq k \leq r = \operatorname{rank}(A)$,

(1) A 的第 k 个**行列式因子** $D_k(A)$ 是指它所有 k 阶子式的最大公因子;

(2) A 的**不变因子** $d_k(A) = D_k(A)/D_{k-1}(A)$, 此处记 $D_0(A) = 1$. 多重集 $\{d_1(A), \cdots, d_r(A)\}$ 称为 A 的**不变因子组**.

命题 4.3.4. *设 A 是 R 上的矩阵, B 是 A 经初等变换得到的矩阵. 则对任意 $1 \leq k \leq \min\{m, n\}$, B 的 k 阶子式都是 A 的 k 阶子式的线性组合. 故*

(1) $\operatorname{rank}(A) - \operatorname{rank}(B)$;

(2) *对于 $1 \leq k \leq \operatorname{rank}(A)$, $D_k(A) = D_k(B)$.*

证明. 若 B 的 k 阶子式都是 A 的 k 阶子式的线性组合, 则 $\operatorname{rank}(B) \leq \operatorname{rank}(A)$ 且 $D_k(A) \mid D_k(B)$ (若 $D_k(A) \neq 0$), 但由于初等变换是可逆变换, 另一个方向就是显然的. 故只需证明此结论.

不妨设初等变换是行变换. 设讨论的 B 的子式是 $B\left(\begin{smallmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{smallmatrix}\right)$.

(i) 若初等变换涉及到的行都在或者都不在集合 $\{i_1, \cdots, i_k\}$ 中, 则

$$B\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = cA\begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix},$$

其中 $c = \pm 1$ 或者 λ.

(ii) 若初等变换为互换第 i 行与第 j 行, 其中 $i \in \{i_1, \cdots, i_k\}$ 而 $j \notin \{i_1, \cdots, i_k\}$, 将集合 $\{i_1, \cdots, i_k, j\} \backslash \{i\}$ 中元素排序为 $i'_1 < \cdots < i'_k$, 则

$$B \begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = \pm A \begin{pmatrix} i'_1 & \cdots & i'_k \\ j_1 & \cdots & j_k \end{pmatrix}.$$

(iii) 若初等变换将第 j 行的 a 倍加到第 i 行, 其中 $i \in \{i_1, \cdots, i_k\}$ 而 $j \notin \{i_1, \cdots, i_k\}$, 与 (ii) 一样排序集合 $\{i_1, \cdots, i_k, j\} \backslash \{i\}$ 中的元素, 则

$$B \begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} = A \begin{pmatrix} i_1 & \cdots & i_k \\ j_1 & \cdots & j_k \end{pmatrix} \pm aA \begin{pmatrix} i'_1 & \cdots & i'_k \\ j_1 & \cdots & j_k \end{pmatrix}. \qquad \square$$

引理 4.3.5. *设 $(a, *, \cdots, *)$ 与 $(b, *, \cdots, *)$ 是矩阵 A 的两行且 $a \neq 0$. 则经过有限次初等行变换后可以将这两条行向量变为 $(d, *, \cdots, *)$ 与 $(0, *, \cdots, *)$, 其余行不变, 这里 d 是 a 和 b 的最大公因子.*

证明. 这个结果实际上就是欧几里得算法的实现. 不妨设 $(a, *, \cdots, *)$ 是 A 的第 1 行, $(b, *, \cdots, *)$ 是第 2 行. 设 $b = qb + r$, 则第 1 行的 $-q$ 倍加到第 2 行, 第 2 行就变为 $(r, *, \cdots, *)$. 若 $r \neq 0$, 令 $a = q_1 r + r_1$. 则第 2 行的 $-q_1$ 倍加到第 1 行, 第 1 行就变为 $(r_1, *, \cdots, *)$). 继续这个过程, 经过有限次变换后, 我们就将 a 与 b 中一个变为 0, 另外一个与最大公因子 d 相差一个乘法可逆元. 再做一次第二类变换 (将非零元变为 d) 和最多一次第一类变换 (使得 $a = d$ 而 $b = 0$), 就得到引理所要求的形式. $\square$

定理 4.3.6. *任意非零阵 A 经过有限次初等变换可以变为矩阵 $S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ 的形式, 其中 $D = \mathrm{diag}(d_1, \cdots, d_r)$, d_k 恒正, 且 $d_1 \mid d_2 \mid \cdots \mid d_r$, 而且这样形式的 S 是唯一的: $r = \mathrm{rank}(A)$ 和 $d_k = d_k(A)$ 由 A 决定.*

证明. (1) 首先证明存在性.

(i) 由于 $A \neq 0$, 设 $a_{ij} \neq 0$, 互换第 i 行与第 1 行, 再互换第 j 列与第 1 列, 最后对第 1 行乘以 a_{11} 的符号或者首项系数的逆, 这样得到的新矩阵 A 就满足 a_{11} 恒正.

(ii) 由于高度不能无限递降, 我们可以假设经过有限次初等变换后 $h(a_{11})$ 达到最小. 我们断言此时 $a_{11} \mid a_{i1}$. 否则, 若 $a_{11} \nmid a_{i1}$, 根据引理 4.3.5, 我们可以通过初等行变换将 a_{11} 替换为 $\gcd(a_{11}, a_{i1})$, 它恒正且高度比 a_{11} 小, 与 $h(a_{11})$ 的最小性矛盾. 同理我们有 $a_{11} \mid a_{1j}$. 这样我们可以通过第三类初等变换使得 $a_{i1} = a_{1j} = 0 (i > 1, j > 1)$ 而 a_{11} 不变.

接下来我们断言此时 $a_{11} \mid a_{ij}$ 对所有 i, j 成立. 否则, 设 $a_{11} \nmid a_{ij}$. 将第 j 列加到第 1 列, 那么新的第 1 行是 $(a_{11}, \cdots)$ 而第 i 行是 $(a_{ij}, \cdots)$. 根据引理 4.3.5, 我们可以通过有限次初等行变换将 a_{11} 替换为 $\gcd(a_{11}, a_{ij})$, 它恒正且高度比 a_{11} 小, 与 $h(a_{11})$ 的最小性矛盾. 故经过有限次初等变换后, 可以得到新的 A 为 $\begin{pmatrix} d_1 & 0 \\ 0 & B \end{pmatrix}$ 的形式, 其中 d_1 恒正且 $B = (b_{ij})$ 中每个元素均被 d_1 整除.

(iii) 根据归纳, B 经有限次初等变换后就得到定理要求的形式 $S_1 = \begin{pmatrix} D_1 & 0 \\ 0 & 0 \end{pmatrix}$ 的形式, 其中 $D_1 = \mathrm{diag}(d_2, \cdots, d_r)$, 且 $d_2 \mid \cdots \mid d_r$. 由于 d_2 是 B 中元素的线性组合, 所以 $d_1 \mid d_2$. 这样我们就经过有限次初等变换将 A 变为 S 的形式.

(2) 我们再证唯一性. 由于初等变换不改变矩阵的秩, 故 $r = \mathrm{rank}(S) = \mathrm{rank}(A)$. 由于初等变换不改变行列式因子, 故 $D_k(A) = D_k(S)$, 但根据定义, $D_k(S) = d_1 \cdots d_k$, 故 $d_k = d_k(A) = D_k(A)/D_{k-1}(A)$. □

命题 4.3.7. *方阵 A 是可逆矩阵当且仅当它是初等矩阵的有限乘积.*

证明. 一方面, 由于初等阵是可逆阵, 它们的有限乘积也是可逆阵. 另一方面, 若 A 是可逆阵, 根据上面的定理, 则 A 等于 $P_1 \cdots P_s S Q_1 \cdots Q_t$, 其中 P_i, Q_j 是初等阵而 $S = \mathrm{diag}(d_1, \cdots, d_n)$ 且 d_i 恒正. 由于 $\det(A)$, $\det(P_i)$ 及 $\det(Q_j)$ 都是 R 的乘法可逆元, 故 $\det(S) = d_1 \cdots d_n$ 也是, 由于 d_i 恒正, 这只可能 $d_1 = \cdots = d_n = 1$, 从而 A 是有限多个初等阵的乘积. □

定义 4.3.4. 矩阵 $S = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, 其中 $D = \mathrm{diag}(d_1(A), \cdots, d_r(A))$, 称为 A 的 **Smith 标准形**.

显然, 给出 A 的 Smith 标准形等价于给出 A 的秩和不变因子组. 定理 4.3.6 还告诉我们不变因子间的互相整除关系: $d_1(A) \mid \cdots \mid d_r(A)$. 将 A 的不变因子做因式分解:

$$d_k = \prod_{i=1}^{s} p_i^{e_{ik}}, \tag{4.14}$$

这里 p_i 为互不相同的恒正素因子 (不可约因子), 且 $e_{ik} \geq 0$. 根据不变因子间的整除关系, 则对每一个 i 均有

$$e_{i1} \leq e_{i2} \leq \cdots \leq e_{ir}.$$

定义 4.3.5. 设元素 $0 \neq a \in R$ 的标准分解为 $a = u \prod_{i \in S} p_i^{e_i}$, 其中 u 是单位, p_i 恒正不可约且两两不同, 则称集合 $\{p_i^{e_i} \mid i \in S\}$ 是 a 的 **初等因子组**.

方阵 A 的**初等因子组**指由它的所有不变因子 $d_k(A)$ $(1 \leq k \leq \mathrm{rank}(A))$ 的初等因子组并起来得到的多重集, 即 $\{p_i^{e_{ij}} \mid 1 \leq i \leq s,\ 1 \leq j \leq r\}$. 这里我们规定初等因子组中可以任意增减 $1 = p_i^0$ 出现的次数.

命题 4.3.8. *矩阵的 Smith 标准形与它的秩和初等因子组互相决定.*

证明. 只需证明矩阵的秩和初等因子组决定它的不变因子组即可.

对于 A 的初等因子组出现的所有素 (不可约) 因子 p_i $(1 \leq i \leq s)$, p_i-幂次的元素最多有 $r = \mathrm{rank}(A)$ 个. 我们用 p_i^0 补全, 保证这样的元素恰好有 r 个, 并按升序排列, 即初等因子组中 p_i 的幂次元素为 $p_i^{f_{i1}}, \cdots, p_i^{f_{ir}}$, 且 $0 \leq f_{i1} \leq \cdots \leq f_{ir}$.

另一方面, 设 $p_i^{e_{ik}} \| d_k(A)$, 故 $\{e_{jk} \mid 1 \leq k \leq r\} = \{f_{ik} \mid 1 \leq k \leq r\}$. 又由于 $d_k(A) \mid d_{k+1}(A)$, 故 $e_{ik} \leq e_{i,k+1}$, 即 $0 \leq e_{i1} \leq \cdots \leq e_{ir}$. 两者综合起来就说明 $e_{ik} = f_{ik}$, 所以 $d_k = \prod_{i=1}^{s} p_i^{f_{ik}}$ 对所有 $1 \leq k \leq r$ 成立. □

注记. 对于多项式矩阵 A, 设 F_0 是包含 A 中所有元素的系数的最小的域, 这样 A 就是 $F_0[x]$ 上的多项式矩阵. 因此 A 的所有子式都是 F_0 上的多项式, 从而 (作为子式的最大公因子的) 行列式因子和不变因子也都是 F_0 上的多项式. 这说明 A 的不变因子不随着域的扩大而改变, 这就是我们称它是不变因子的原因.

这与初等因子组的情况不同. 例如, 设 A 是 $\mathbb{Q}[x]$ 上的多项式矩阵, 不变因子 $d_k(A)$ 在 $\mathbb{Q}[x]$, $\mathbb{R}[x]$ 和 $\mathbb{C}[x]$ 中有不同的因式分解, 因此得到的初等因子组也是不一样的.

定义 4.3.6. 两个矩阵 A 与 B 相抵是指存在可逆阵 P 与 Q 使得 $B = PAQ$.

定理 4.3.6 说明 A 与它的 Smith 标准形相抵.

定理 4.3.9. *相抵关系是等价关系, 且两个矩阵相抵当且仅当它们有相同的 Smith 标准形, 即它们有相同的秩和不变因子组, 亦即它们有相同的秩和初等因子组.*

证明. 相抵关系是等价关系这是显然的, 因此有相同 Smith 标准形的矩阵也一定相抵. 反之, 若矩阵 $B = PAQ$ 与 A 相抵, 由命题 4.3.7, P 和 Q 都是有限个初等矩阵的乘积, 再由命题 4.3.4, 每一次左乘或者右乘初等矩阵不改变 Smith 标准形, 故 A 与 B 有相同的 Smith 标准形. □

定理 4.3.10. *设 $A = \mathrm{diag}(a_1, \cdots, a_n)$ 是对角阵. 则 A 的初等因子组就是它的非零对角元的初等因子组之并.*

证明. 不妨设 $a_1, \cdots, a_n$ 恒正. 记 $a_j = \prod_{i=1}^{s} p_i^{f_{ij}}$ 为 a_j 的因式分解, 这里 $f_{ij} \geq 0$. 将 $\{f_{ij} \mid j = 1, \cdots, n\}$ 从小到大排序:

$$e_{i1} \leq e_{i2} \leq \cdots \leq e_{in}.$$

主要到 A 的 k 阶子式要么等于 0, 要么等于 $a_{j_1} \cdots a_{j_k}$. 设 $p_i^{t_{ik}} \| D_k(A)$. 根据定义, $D_k(A)$ 是 k 阶子式的最大公因子, 故

$$t_{ik} = \min\{f_{ij_1} + \cdots + f_{ij_k} \mid 1 \leq j_1 < \cdots < j_k \leq n\} = e_{i1} + \cdots + e_{ik}.$$

故 $d_k(A) = D_k(A)/D_{k-1}(A) = \prod_{i=1}^{s} p_i^{e_{ik}}$. 所以 A 的初等因子组就是 a_j $(1 \leq j \leq n)$ 的初等因子组之并. □

推论 4.3.11. *设 $A = \mathrm{diag}(A_1, \cdots, A_s)$ 是准对角阵. 则 A 的初等因子组是 A_i $(1 \leq i \leq s)$ 的初等因子组之并.*

4.3.3 一些例子

例 4.3.3. *设* $A = \begin{pmatrix} 4 & 3 & -4 \\ -1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}$. *通过初等变换求矩阵 $xI - A$ 的 Smith 标准形.*

解. 我们用 r_i 表示第 i 行, c_i 表示第 i 列, 且变换从上而下, 则有

$$\begin{pmatrix} x-4 & -3 & 4 \\ 1 & x & -2 \\ -1 & -1 & x \end{pmatrix} \xrightarrow[r_2+r_3\to r_2, r_1\leftrightarrow -r_3]{r_1+(x-4)r_3\to r_1} \begin{pmatrix} 1 & 1 & -x \\ 0 & x-1 & x-2 \\ 0 & -x+1 & x^2-4x+4 \end{pmatrix}$$

$$\xrightarrow[c_3+xc_1\to c_3]{c_2-c_1\to c_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & x-2 \\ 0 & -x+1 & x^2-4x+4 \end{pmatrix} \xrightarrow[r_3+r_2\to r_3]{c_3-c_2\to c_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & -1 \\ 0 & 0 & x^2-3x+2 \end{pmatrix}$$

$$\xrightarrow[r_3+(x^2-3x+2)r_2\to r_3]{c_2+(x-1)c_3\to c_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & (x-1)^2(x-2) & 0 \end{pmatrix} \xrightarrow[-r_2\to r_2]{r_2\leftrightarrow r_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x-1)^2(x-2) \end{pmatrix}.$$

□

例 4.3.4. 设 $a_1,\cdots,a_n$ 是两两互素的正整数. 证明: 方阵 $A=\mathrm{diag}(a_1,\cdots,a_n)$ 的 *Smith* 标准形为 $\mathrm{diag}(1,\cdots,1,a_1\cdots a_n)$.

证明. 首先 $\det(A)=a_1\cdots a_n\neq 0$, 故它的秩是 a. 其次, 由于 a_i 两两互素, A 的初等因子也两两互素, 这样就有 $d_1=\cdots=d_{n-1}=1$, 而 $d_n=a_1\cdots a_n$. □

例 4.3.5. 考虑多项式矩阵 xI_n-J, 其中 J 是 *Jordan* 形矩阵.

(1) 多项式矩阵 $xI_n-J_n(\lambda)$ 的不变因子组为 $\{1,\cdots,1,(x-\lambda)^n\}$, 初等因子组为 $\{(x-\lambda)^n\}$.

(2) 更一般地, 若 $J=\mathrm{diag}(J_{n_1}(\lambda_1),\cdots,J_{n_s}(\lambda_s))$, 则 xI_n-J 的初等因子组为 $\{(x-\lambda_i)^{n_i}\mid 1\leq i\leq s\}$.

解. (2) 由 (1) 和推论 4.3.11 立得. 至于 (1), 注意到

$$xI_n-J_n(\lambda)=\begin{pmatrix} x-\lambda & -1 & & \\ & x-\lambda & \ddots & \\ & & \ddots & -1 \\ & & & x-\lambda \end{pmatrix}_{n\times n}.$$

只需考虑它去掉第 1 列和第 n 行的子矩阵就知 $D_1(xI_n-J_n(\lambda))=\cdots=D_{n-1}(xI_n-J_n(\lambda))=1$, 而 $D_n(xI_n-J_n(\lambda))=(x-\lambda)^n$ 是显然的. □

注记. 设 A 是 $f(x)$ 的友阵. 与上面例子一样计算 $xI-A$ 的行列式因子, 就得到 xI_n-A 的 Smith 标准形是 $\mathrm{diag}(1,\cdots,1,f(x))$.

例 4.3.6. 试求多项式矩阵

$$A=\begin{pmatrix} x^2(x-1)^3(x+1) & & \\ & x^4(x-1)(x+1)^3 & \\ & & x(x+1)^2(x-1)^3 \end{pmatrix}$$

的不变因子, 初等因子组和 *Smith* 标准形.

解. A 的初等因子组是所有对角元的初等因子组之并, 即 $\{x, x^2, x^4, x+1, (x+1)^2, (x+1)^3, x-1, (x-1)^3, (x-1)^3\}$, 这样它的不变因子就是 $\{x(x-1)(x+1), x^2(x-1)^3(x+1)^2, x^4(x-1)^3(x+1)^3\}$, Smith 标准形等于

$$\begin{pmatrix} x(x-1)(x+1) & & \\ & x^2(x-1)^3(x+1)^2 & \\ & & x^4(x-1)^3(x+1)^3 \end{pmatrix}. \qquad \square$$

习题 4.3.1. *设多项式矩阵 A 为 4 阶方阵, 初等因子组为 $\{x^2, x^4, (x-1)^2, (x-1)^3, x+2\}$. 求它的 Smith 标准形 S.*

解. 这里没有说 A 的秩, 所以我们需要分情况讨论. 由于拥有同一个素因子底的初等因子有 2 个, 故 A 的秩 ≥ 2.

当 $\operatorname{rank}(A) = 2$ 时, $S = \operatorname{diag}(x^2(x-1)^2, x^4(x-1)^3(x+1), 0, 0)$; 当 $\operatorname{rank}(A) = 3$ 时, $S = \operatorname{diag}(1, x^2(x-1)^2, x^4(x-1)^3(x+1), 0)$; 当 $\operatorname{rank}(A) = 4$ 时, $S = \operatorname{diag}(1, 1, x^2(x-1)^2, x^4(x-1)^3(x+1))$. $\square$

习 题

习题 4.3.2. *求下列整数或多项式矩阵的 Smith 标准形.*

(1) $\begin{pmatrix} 30 & & \\ & 12 & \\ & & 72 \end{pmatrix}$, (2) $\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$,

(3) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$, (4) $\begin{pmatrix} 1 & -3 & 0 & 0 & 2 \\ 1 & 0 & -5 & 0 & 3 \\ 1 & 0 & 0 & -7 & 2 \end{pmatrix}$,

(5) $\begin{pmatrix} x^2+x & & \\ & x^4-1 & \\ & & x^3+1 \end{pmatrix}$, (6) $\begin{pmatrix} 2x & x+x^2 & x+x^3 \\ x+x^2 & 2x^2 & x^2+x^3 \\ x+x^3 & x^2+x^3 & 2x^3 \end{pmatrix}$.

习题 4.3.3. *试求下述多项式矩阵的 Smith 标准形, 不变因子和初等因子组:*

(1) $\begin{pmatrix} x & 1 & \cdots & 1 & 1 \\ 1 & x & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \cdots & x & 1 \\ 1 & 1 & \cdots & 1 & x \end{pmatrix}$, (2) $\begin{pmatrix} x & 1 & \cdots & 1 & 1 \\ 0 & x & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & x & 1 \\ 0 & 0 & \cdots & 0 & x \end{pmatrix}$,

(3) $\begin{pmatrix} x^2(x-1)^3(x+1) & & \\ & x^4(x-1)(x+1)^3 & \\ & & x(x+1)^2(x-1)^3 \end{pmatrix}$.

习题 4.3.4. 设 A 是 n 次首一多项式 $f(x)$ 的友阵. 试求 $xI_n - A$ 的 *Smith* 标准形.

习题 4.3.5. 设 $A = \begin{pmatrix} x^3-1 & \\ & x^5-1 \end{pmatrix}$. 试求 A 的不变因子, 并求它在有理数域上的初等因子组.

习题 4.3.6. 设 $R = \mathbb{Z}$ 或者 $F[x]$. 证明 R 上的方阵 A 与它的转置 A^T 相抵.

习题 4.3.7. 证明: 整系数线性方程组 $AX = b$ 有整数解的充分必要条件是 (A, b) 与 $(A, 0)$ 相抵.

4.4 有理标准形理论

4.4.1 $F[x]$ 上的模

回忆一下域 F 上线性空间 V 的定义, 其关键之处在于存在数乘映射 $F \times V \to V$, $(\lambda, v) \mapsto \lambda v$, 它满足结合律, 与向量加法间的分配律且 $1v = v$ 对所有 $v \in V$ 成立. 将这个概念推广, 就得到多项式环 $F[x]$ 上模的概念.

定义 4.4.1. 设 V 是 F-线性空间. 若存在数乘映射

$$F[x] \times V \to V,\ (f(x), v) = f(x)v, \tag{4.15}$$

它延拓域上的数乘映射 $F \times V \to V$, 且满足

(1) 加法与数乘间的分配律成立, 即

$$(f(x) + g(x))v = f(x)v + g(x)v,\ f(x)(v + w) = f(x)v + f(x)w;$$

(2) 结合律成立, 即

$$(f(x)g(x))v = f(x)(g(x)v);$$

则称 V 是 $F[x]$-**模** (module). 若 W 是 V 的子空间且对于任意 $w \in W$, $f(x) \in F[x]$, 均有 $f(x)w \in W$, 则称 W 是 V 的**子模** (sub-module). 若 $V = F[x]v$, 即它由一个元素 v 生成. 则称之为**循环模** (cyclic module).

模的概念可以做进一步推广, 模论是代数理论的重要组成部分, 读者在后续代数课程将会学到. 这里我们只需要用到下面这个简单事实:

引理 4.4.1. 设 V 是 $F[x]$-模. 对于 $(\alpha_1, \cdots, \alpha_m) \in V^m$, $A = (a_{ij}) \in F[x]^{m\times n}$, 定义

$$(\alpha_1, \cdots, \alpha_m)A := \left(\sum_{i=1}^m a_{i1}\alpha_i, \cdots, \sum_{i=1}^m a_{in}\alpha_i\right) \in V^n. \tag{4.16}$$

则对于任意 $A \in F[x]^{m\times n}$, $B \in F[x]^{n\times p}$, 均有

$$\Big[(\alpha_1, \cdots, \alpha_m)A\Big]B = (\alpha_1, \cdots, \alpha_m)(AB). \tag{4.17}$$

证明. 设 $A = (a_{ij})$, $B = (b_{jk})$. 由计算立得. □

4.4.2 有理标准形理论

线性空间 V 上每一个线性变换 $\mathbf{A}$, 都自然诱导出 V 的一个 $F[x]$-模的结构, 这个事实让我们可以从更高观点来研究线性变换和矩阵的性质, 特别是得到相似标准形的理论. 我们首先给出这个模结构.

设 V 是 F-线性空间, $\mathbf{A}$ 是 V 上的线性变换. 我们如下定义多项式环 $F[x]$ 与 V 的数乘:

$$f(x)\,v := f(\mathbf{A})(v). \tag{4.18}$$

根据这个定义, V 就成为 $F[x]$-模, 称为 $\mathbf{A}$ 决定的 $F[x]$-模. 此时 W 是 V 的子模当且仅当 W 是 $\mathbf{A}$ 的不变子空间. 元素 $\alpha \in V$ 生成的循环子空间 $F[\mathbf{A}](\alpha)$ 就是循环模 $F[x]\alpha$.

设 $\mathbf{A}$ 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 A, 即

$$\mathbf{A}(\alpha_1, \cdots, \alpha_n) = (\alpha_1, \cdots, \alpha_n)A. \tag{4.19}$$

利用 (4.16) 的定义, (4.19) 就变成

$$(\alpha_1, \cdots, \alpha_n)(xI_n) = (\alpha_1, \cdots, \alpha_n)A.$$

也就是说,

$$(\alpha_1, \cdots, \alpha_n)(xI_n - A) = 0. \tag{4.20}$$

我们将经常使用这个公式.

我们首先用模结构来再次证明 Cayley-Hamilton 定理.

定理 4.4.2 (Cayley-Hamilton)**.** *$\varphi_{\mathbf{A}}(x)$ 是 $\mathbf{A}$ 的零化多项式, 即 $\varphi_{\mathbf{A}}(\mathbf{A}) = 0$.*

证明. 在 (4.20) 两边同乘以 $xI_n - A$ 的伴随矩阵 $(xI_n - A)^*$, 注意到 $(xI_n - A)(xI_n - A)^* = \varphi_A(x)I_n$, 并利用引理 4.4.1, 就得到

$$(\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} \varphi_{\mathbf{A}}(x) & & & \\ & \varphi_{\mathbf{A}}(x) & & \\ & & \ddots & \\ & & & \varphi_{\mathbf{A}}(x) \end{pmatrix} = 0.$$

这也就是说对于 $i = 1, \cdots, n$, $\varphi_{\mathbf{A}}(\mathbf{A})(\alpha_i) = 0$. 因为 $\{\alpha_1, \cdots, \alpha_n\}$ 是 V 的一组基, 所以变换 $\varphi_{\mathbf{A}}(\mathbf{A})$ 就是零变换. □

我们接下来用模结构来讨论相似标准形理论. 这里的关键是下述定理:

定理 4.4.3. *设线性变换 $\mathbf{A} : V \to V$ 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 A, 并设*

$$P(x)(xI_n - A)Q(x) = D(x) := \mathrm{diag}(1, \cdots, 1, f_{s+1}, \cdots, f_n),$$

其中 $P(x)$, $Q(x)$ 是可逆多项式矩阵, $f_i(x)$ 是首一多项式. 令

$$(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)P(x)^{-1}.$$

则下列结果成立:

(1) 对于 $i \le s$, $\beta_i = 0$;

(2) 对于 $i > s$, β_i 的最小多项式是 $f_i(x)$;

(3) V 是循环子空间 $F[x]\beta_i$ $(s+1 \le i \le n)$ 的直和.

证明. 由 $(\alpha_1, \cdots, \alpha_n)(xI_n - A) = 0$ 即得

$$(\beta_1, \cdots, \beta_n)D(x) = 0.$$

这说明 $\beta_1 = \cdots = \beta_s = 0$, 即 (1) 成立; 且对于 $i > s$, $f_i(x)\beta_i = f_i(\mathbf{A})(\beta_i) = 0$, 即 $f_i(x)$ 是 β_i 的零化多项式, 从而

$$\sum_{i=s+1}^{n} \dim_F F[x]\beta_i \le \sum_{i=s+1}^{n} \deg(f_i(x)) = \deg \varphi_{\mathbf{A}}(x) = \dim_F V. \tag{4.21}$$

另一方面, 由 $(\alpha_1, \cdots, \alpha_n) = (\beta_1, \cdots, \beta_n)P(x)$, 我们知道

$$\alpha_i \in F[x]\beta_{s+1} + \cdots + F[x]\beta_n.$$

又由于 $\{\alpha_i\}$ 是 V 的基, 这说明

$$V = F[x]\beta_{s+1} + \cdots + F[x]\beta_n. \tag{4.22}$$

综合 (4.21) 和 (4.22) 就说明 $V = F[x]\beta_{s+1} \oplus \cdots \oplus F[x]\beta_n$, 且对于每个 $i \ge s+1$, 均有 $\dim F[x]\beta_i = \deg(f_i)$, 即 $f_i(x)$ 是 β_i 的最小多项式. □

定理 4.4.4 (有理标准形)**.** 设 A 是域 F 上的 n 阶方阵, 则多项式方阵 $xI_n - A$ 的 *Smith* 标准形为 $S(x) = \mathrm{diag}(1, \cdots, 1, d_{s+1}(x), \cdots, d_n(x))$, 且

(1) A 的特征多项式 $\varphi_A(x) = d_{s+1}(x) \cdots d_n(x)$, 最小多项式即 $d_n(x)$.

(2) 设 B_i $(i = s+1, \cdots, n)$ 是不变因子 $d_i(x)$ 的友阵, 则 A 相似于准对角阵 $\mathrm{diag}(B_{s+1}, \cdots, B_n)$.

(3) 更进一步, 设 $d_i(x)$ $(i = s+1, \cdots, n)$ 的初等因子组是 $\{p_{ij}^{e_{ij}} \mid 1 \le j \le j_i\}$, 并设 B_{ij} 是 $p_{ij}^{e_{ij}}$ 的友阵, 则 A 相似于以 B_{ij} $(s+1 \le i \le n,\ 1 \le j \le j_i)$ 为准对角阵块的准对角阵.

证明. 首先 $xI_n - A$ 的行列式 $\varphi_A(x) \ne 0$, 故 $\mathrm{rank}(xI_n - A) = n$. 其次, $\varphi_A(x) = \det(xI_n - A)$ 与 $\det(S) = d_{s+1}(x) \cdots d_n(x)$ 只相差乘法可逆元, 而它们都是首一的, 故一定相等.

取线性变换 $\mathbf{A} : V \to V$ 使得它在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵为 A. 令定理 4.4.3 中的 $D(x)$ 等于 $S(x)$, 则 $V = \bigoplus_{i=1}^{s} F[x]\beta_i$ 是循环子空间的直和且 β_i 的最小多项式为 $d_i(x)$. 这说明 $d_n(x)$ 一方面零化 β_i, 故零化 V 中所有向量, 另一方面, $d_n(x)$ 的真因子不能零化 β_n. 因此 $d_n(x)$ 是零化 V 中所有向量的次数最低的首一多项式, 也就是说它是 $\mathbf{A}$ 和 A 的最小多项式. 故 (1) 得证.

对于循环子空间 $F[x]\beta_i$, 取它的一组基为 $\{\beta_i, \mathbf{A}(\beta_i), \cdots, \mathbf{A}^{n_i-1}(\beta_i)\}$, 其中 $n_i = \deg(d_i(x))$, 则 $\mathbf{A}|_{F[x]\beta_i}$ 在此基下的矩阵是 $d_i(x)$ 的友阵 B_i. 故 A 相似于 $\mathrm{diag}(B_{s+1}, \cdots, B_n)$. 这就证明了 (2).

最后, 根据命题 4.1.7, 将 $F[x]\beta_i$ 分解为 $F[x]\beta_{ij}$ 的直和即得 (3). □

定义 4.4.2. 对于方阵 $A \in F^{n\times n}$, 多项式方阵 $xI_n - A$ 称为 A 的**特征方阵**. 定理 4.4.4 中的准对角阵称为 A 的**有理标准形**.

定理 4.4.5. 域 F 上的方阵 A 与 B 相似当且仅当它们的特征方阵 $xI_n - A$ 与 $xI_n - B$ 相抵.

证明. 若 $B = P^{-1}AP$ 与 A 相似, 则 $xI_n - B = P^{-1}(xI_n - A)P$ 相似, 因此它们自然相抵.

反过来, 设 $xI_n - A$ 与 $xI_n - B$ 相抵. 我们给出两个方法来证明 A 与 B 相似:

证法 1: 此时两个特征方阵有相同的 Smith 标准形, 根据定理 4.4.4, A 和 B 有相同的有理标准形, 故它们相似.

证法 2: 设 $xI_n - B = P(x)(xI_n - A)Q(x)$, 其中 $P(x)$ 与 $Q(x)$ 是多项式可逆阵. 设 V 是 F-线性空间, $\mathbf{A}$ 是 V 上的线性变换, 并设 $\mathbf{A}$ 在基 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 A. 将 V 通过 $\mathbf{A}$ 成为 $F[x]$-模, 则

$$(\alpha_1, \cdots, \alpha_n)(xI_n - A) = 0.$$

所以 $(\alpha_1, \cdots, \alpha_n)P^{-1}(x)(xI_n - B) = 0$. 令 $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)P(x)^{-1}$, 则

$$(\alpha_1, \cdots, \alpha_n) = (\beta_1, \cdots, \beta_n)P(x). \tag{4.23}$$

由 $(\beta_1, \cdots, \beta_n)(xI_n - B) = 0$, 故

$$\mathbf{A}(\beta_1, \cdots, \beta_n) = (\beta_1, \cdots, \beta_n)(xI_n) = (\beta_1, \cdots, \beta_n)B. \tag{4.24}$$

从而对非负整数 i, 均有

$$(\beta_1, \cdots, \beta_n)(x^n I_n) = (\beta_1, \cdots, \beta_n)B^i.$$

记 $P(x)$ 中元素 $p_{ij}(x) = \sum_{k=0}^{s} p_{ij}^{(k)} x^k$, 其中 $p_{ij}^{(k)} \in F$. 则

$$P(x) = \sum_{k=0}^{s} x^k P_k = \sum_{k=0}^{s} (x^k I_n) P_k, \text{ 其中 } P_k = (p_{ij}^{(k)}) \in F^{n\times n}.$$

故

$$\begin{aligned}(\alpha_1, \cdots, \alpha_n) =& \sum_{k=0}^{s} (\beta_1, \cdots, \beta_n)(x^k I_n \cdot P_k) \\ =& \sum_{k=0}^{s} [(\beta_1, \cdots, \beta_n) x^k I_n] \cdot P_k \\ =& (\beta_1, \cdots, \beta_n)(P_0 + BP_1 + \cdots + B^s P_s).\end{aligned}$$

这说明 $\{\alpha_1, \cdots, \alpha_n\}$ 由 $\{\beta_1, \cdots, \beta_n\}$ 线性表出, 因此 $\{\beta_1, \cdots, \beta_n\}$ 也是 V 的一组基且 B 是 $\mathbf{A}$ 在这组基下的矩阵, 故 A 与 B 相似. □

注记. 根据证法 2, $T = P(B) = P_0 + BP_1 + \cdots + B^s P_s$ 是 F 上的可逆矩阵且 $A = T^{-1}BT$. 这给出了求两个相似矩阵间的过渡矩阵的一种方法.

定理 4.4.5 和推论 4.3.11 是方阵准对角化的重要工具. 如果对于 $xI_n - A$ 的每一个初等因子 $1 \neq p_i^{e_i}$, 都能找到 A_i 使得它的特征方阵只有唯一的初等因子 $p_i^{e_i}$, 那么 A 就相似于对角块是 A_i 的准对角阵. 作为一个例子, 我们给出 Jordan 标准形定理存在性部分的新证明:

推论 4.4.6. *若方阵 A 的特征值都在 F 中, 则 A 相似于 Jordan 形矩阵.*

证明. 此时 $xI_n - A$ 的不变因子的根都在 F 中, 故可以假设 $xI_n - A$ 的初等因子组为 $\{(x-\lambda_i)^{n_i} \mid 1 \le i \le s\}$. 根据例 4.3.5, 令 $J = \mathrm{diag}(J_{n_1}(\lambda_1), \cdots, J_{n_s}(\lambda_s))$, 则 $xI_n - J$ 的初等因子组也是 $\{(x-\lambda_i)^{n_i} \mid 1 \le i \le s\}$. 故 $xI_n - A$ 与 $xI_n - J$ 相抵, 根据上面的定理, A 与 J 相似. □

定理 4.4.7. *设 A 与 B 是定义在域 F 上的矩阵, K 是 F 的扩域. 则 A 与 B 在 F 上相似当且仅当它在 K 上相似.*

证明. 这是由于特征方阵 $xI_n - A$ 的 Smith 标准形不随域的变化而变化. □

我们再举几个例子来说明有理标准形理论的应用.

例 4.4.1. *设 A 是 F 上的方阵. 由于转置不改变行列式, 所以 $xI_n - A$ 与 $xI_n - A^T$ 有相同的秩与行列式因子, 因此它们相抵, 从而 A 与 A^T 在 F 上相似.*

例 4.4.2. *根据例 4.3.3, $A = \begin{pmatrix} 4 & 3 & -4 \\ -1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}$ 的特征方阵的不变因子是 $(x-1)^2(x-2)$, 故 A 的 Jordan 标准形就是 $\mathrm{diag}(2, J_2(1))$, 这与例 4.2.5 得到的结果一样.*

例 4.4.3. *设方阵 A 的最小多项式等于特征多项式, 方阵 B 与 A 可交换. 证明: 存在 $f(x) \in F[x]$, 使得 $B = f(A)$.*

证明. 设 $\mathbf{A} : X \mapsto AX$ 和 $\mathbf{B} : B \mapsto BX$ 是 $F^{n\times 1}$ 上的线性变换, 则 $\mathbf{AB} = \mathbf{BA}$.

根据有理标准形定理, 我们知道存在 $\alpha \in F^{n\times 1}$ 使得 $F^{n\times 1} = F[\mathbf{A}]\alpha$. 则 $\mathbf{B}\alpha = f(\mathbf{A})\alpha$ 对于某个 $f(x) \in F[x]$ 成立. 对于任意 $v \in F^{n\times 1}$, 设 $v = h(\mathbf{A})\alpha$. 则

$$\mathbf{B}(v) = \mathbf{B}(h(\mathbf{A})\alpha) = h(\mathbf{A})f(\mathbf{A})\alpha = f(\mathbf{A})(h(\mathbf{A})\alpha) = f(\mathbf{A})(v).$$

这说明 $\mathbf{B} = f(\mathbf{A})$, 故 $B = f(A)$. □

$F = \mathbb{C}$ 时的矩阵证法. 不妨设 $A = J$ 是 Jordan 形矩阵. 由于 A 的最小多项式等于特征多项式, 对于每个特征值, 只有一个 Jordan 块, 故可设

$$A = \mathrm{diag}(J_1, \cdots, J_s),$$

其中 $J_i = J_{n_i}(\lambda_i)$ 且 λ_i 两两不同. 将 B 做对应分块, 再利用条件 $AB = BA$, 即得

$$B = \mathrm{diag}(B_1, \cdots, B_s), \text{ 且 } B_iJ_i = J_iB_i.$$

首先考虑 $A = J_n(\lambda)$ 的情形. 此时 $AB = BA$ 说明 $N_nB = BN_n$. 经计算即得

$$B = \sum_{i=0}^{n-1} b_i N_n^i = \sum_{i=0}^{n-1} b_i (A - \lambda I_n)^i = f(A).$$

对于一般情况, 我们已经得到 $f_i(x)$, 使得 $B_i = f_i(J_i)$. 根据中国剩余定理, 令 $f(x)$ 是同余方程组 $f(x) \equiv f_i(x) \mod (x-\lambda_i)^{n_i}$ $(i = 1, \cdots s)$ 的解. 则 $B_i = f_i(J_i) = f(J_i)$ 从而 $B = f(A)$. □

4.4.3 实方阵的实相似

这是定理 4.4.5 和推论 4.3.11 的另一个应用. 由于实系数多项式的不可约因子为 1 次或者 2 次, 实多项式矩阵的初等因子只能是 $(x-\lambda)^n$ $(\lambda\in\mathbb{R})$ 或者 $(x^2-2ax+a^2+b^2)^n$ $(a,b\in\mathbb{R},b\neq 0)$ 的形式. 我们已经知道 $J_n(\lambda)$ 的特征方阵的唯一初等因子是 $(x-\lambda)^n$. 现在只需要找到一个实方阵, 它的特征方阵只有唯一的实初等因子 $(x^2-2ax+a^2+b^2)^n$ 即可.

引理 4.4.8. 设 $a,b\in\mathbb{R}$, $b\neq 0$, $\lambda=a+bi$. 令 $K_1=K_1(a,b)=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$. 则对于任意正整数 n, 矩阵 $(xI_2-K_1)^n$ 的不变因子组为 $\{1,(x^2-2ax+a^2+b^2)^n\}$, 它作为实系数矩阵的初等因子组是 $\{(x^2-2ax+a^2+b^2)^n\}$, 它作为复系数矩阵的初等因子组是 $\{(x-\lambda)^n,(x-\bar{\lambda})^n\}$.

证明. 因为 K_1 有两个不同的特征值 λ 和 $\bar{\lambda}$, 所以存在复可逆阵 P, 使得 $K_1=P^{-1}\mathrm{diag}(\lambda,\bar{\lambda})P$, 从而 $xI_2-K_1=P^{-1}\mathrm{diag}(x-\lambda,x-\bar{\lambda})P$, 我们有

$$(xI_2-K_1)^n=P^{-1}\mathrm{diag}((x-\lambda)^n,(x-\bar{\lambda})^n)P.$$

故 $(xI_2-K_1)^n$ 作为复系数矩阵的初等因子组是 $\{(x-\lambda)^n,(x-\bar{\lambda})^n\}$, 这说明它的不变因子组是 $\{1,(x^2-2ax+a^2+b^2)^n\}$, 以及它作为实系数矩阵的初等因子组是 $\{1,(x^2-2ax+a^2+b^2)^n\}$. □

引理 4.4.9. 设 $a,b\in\mathbb{R}$, $b\neq 0$, 令 $K_n=K_n(a,b)$ 是 $2n$ 阶实矩阵

$$K_n(a,b)=\begin{pmatrix} K_1 & I_2 & & & \\ & K_1 & I_2 & & \\ & & \ddots & \ddots & \\ & & & K_1 & I_2 \\ & & & & K_1 \end{pmatrix}.$$

则 $xI_{2n}-K_n$ 的不变因子组为 $\{1,\ (2n-1\text{重}),(x^2-2ax+a^2+b^2)^n\}$, 它作为实系数矩阵的初等因子组是 $\{(x^2-2ax+a^2+b^2)^n\}$, 它作为复系数矩阵的初等因子组是 $\{(x-\lambda)^n,(x-\bar{\lambda})^n\}$.

证明. 我们将特征方阵 $xI_{2n}-K_n$ 看作是矩阵元是 2 阶方阵的 n 阶分块方阵, 做分块初等变换. 首先将块的第 2 列乘以 xI_2-K_1 加到第 1 列, 然后 xI_2-K_1 乘以第 1 行加到第 2 行, 就得到矩阵

$$\begin{pmatrix} 0 & -I_2 & & & \\ (xI_2-K_1)^2 & 0 & -I_2 & & \\ & & \ddots & \ddots & \\ & & & xI_2-K_1 & -I_2 \\ & & & & xI_2-K_1 \end{pmatrix}.$$

继续这个操作, 就得到

$$\begin{pmatrix} 0 & -I_2 & & & \\ & 0 & -I_2 & & \\ & & \ddots & \ddots & \\ & & & 0 & -I_2 \\ (xI_2-K_1)^n & & & & 0 \end{pmatrix} \Longrightarrow \begin{pmatrix} I_2 & & & & \\ & I_2 & & & \\ & & \ddots & & \\ & & & I_2 & \\ & & & & (xI_2-K_1)^n \end{pmatrix}.$$

根据前面引理就得到欲证. □

定理 4.4.10. 设 A 是 n 阶实方阵, 它的特征方阵 $xI_n - A$ 的实初等因子组是 $\{(x-\lambda_i)^{n_i}, (x^2-2a_jx+a_j^2+b_j^2)^{m_j}, i \le s,\ j \le t\}$, 则 A 实相似于准对角阵 D, 其对角块为 $J_{n_i}(\lambda_i)$ $(i \le s)$ 和 $K_{m_j}(a_j, b_j)$ $(j \le t)$.

证明. 这是因为 A 和 D 的特征方阵有相同的初等因子组. □

例 4.4.4. 设 A 为 $2n$ 阶实方阵且 $A^2+I=O$. 证明: A 实相似于 $\begin{pmatrix} O & I \\ -I & O \end{pmatrix}$.

证明. 由于 x^2+1 是 A 的化零多项式, 它没有实根, 故 A 的最小多项式 $d_A(x)=x^2+1$, A 的特征多项式 $\varphi_A(x)=(x^2+1)^n$.

证法 1 由于 $d_A(x)$ 只有单复根 $\pm i$, 故 A 可复相似于对角阵 $\mathrm{diag}(iI_n, -iI_n)$, 而此对角阵又复相似于 $\begin{pmatrix} O & I \\ -I & O \end{pmatrix}$, 因此 A 也复相似从而实相似于此矩阵.

证法 2 此时 $xI-A$ 的不变因子组就是 $\{x^2+1, \cdots, x^2+1(n重)\}$, 故 A 的实有理标准形为 $\mathrm{diag}\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)$, 它相似于 $\begin{pmatrix} O & I \\ -I & O \end{pmatrix}$. □

习　　题

习题 4.4.1. 求下述方阵在有理数域 $\mathbb{Q}$ 上的有理标准形:

$$(1)\quad \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad (2)\quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

习题 4.4.2. 称线性变换 $\mathbf{A}: V \to V$ 为**循环变换**, 若 $\mathbf{A}$ 决定的 V 的 $F[x]$-模是循环模. 证明下列条件等价:

(1) $\mathbf{A}$ 是循环变换;

(2) $\mathbf{A}$ 的特征多项式和最小多项式相等;

(3) $\mathbf{A}$ 在任意基下的矩阵 A 的特征方阵 $xI_n = A$ 的第 $n-1$ 个行列式因子等于 1.

习题 4.4.3. 设 $\lambda \in F^{\times}$. 证明: $J_n(\lambda)$ 与 $\lambda J_n(1)$ 相似.

习题 4.4.4. 设 $a, b \in \mathbb{R}$ 且 $b \neq 0$, 记 $\tau = a + bi$.

(1) 求复矩阵 P 使得 $P\begin{pmatrix} \tau & 0 \\ 0 & \bar{\tau} \end{pmatrix} P^{-1} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$.

(2) 证明 $\begin{pmatrix} J_n(\tau) & 0 \\ 0 & J_n(\bar{\tau}) \end{pmatrix}$ 复相似于 $L_n(a,b) = \begin{pmatrix} aJ_n(1) & bJ_n(1) \\ -bJ_n(1) & aJ_n(1) \end{pmatrix}$.

(3) 证明在定理 *4.4.10*中将 $K_{m_j}(a_j, b_j)$ 用 $L_{m_j}(a_j, b_j)$ 替换, 定理仍然成立.

习题 4.4.5. 设 $X_1 + iX_2$ 是 n 阶实方阵 A 的属于虚特征值 $a+bi$ 的特征向量, 其中 $X_1, X_2 \in \mathbb{R}^{n\times 1}$. 证明: $\langle X_1, X_2 \rangle$ 是 A 的 2 维不变子空间.

习题 4.4.6. 设实方阵 A 的特征多项式 $\varphi_A(x) = x(x-1)^3(x^2+1)^2$. 试讨论 A 的最小多项式、不变因子、*Jordan* 标准形及实相似标准形.

习题 4.4.7. 设 A 可逆, 其特征多项式与最小多项式相等. 证明: $d_{A^{-1}}(x) = d_A(0)^{-1} x^n d_A(x^{-1})$.

4.5 更多例子

4.5.1 常微分方程组的求解.

给定实方阵 $A \in \mathbb{R}^{n\times n}$, 考虑常微分方程组

$$\frac{d}{dt}X = AX, \text{ 其中 } X = (x_1, \cdots, x_n)^T.$$

如果做变量替换 $X = PY$, 其中 P 是可逆矩阵, 那么方程组就归结为求解

$$\frac{d}{dt}Y = P^{-1}APY.$$

如果可以找到合适的可逆阵 P, 使得 $P^{-1}AP = J$ 是 Jordan 阵时, 方程组就可以求解了. 事实上, 由于 J 是准对角阵, 不妨设 $J = J_n(\lambda)$, 则只需要求解

$$\begin{cases} \frac{dy_1}{dt} &= \lambda y_1 + y_2, \\ & \cdots \\ \frac{dy_{n-1}}{dt} &= \lambda y_{n-1} + y_n, \\ \frac{dy_n}{dt} &= \lambda y_n. \end{cases} \tag{4.25}$$

这可以从下而上依次求解, 得到

$$\begin{cases} y_1 &= \left(c_n \frac{t^{n-1}}{(n-1)!} + \cdots + c_3 \frac{t^2}{2!} + c_2 t + c_1\right) e^{\lambda t}, \\ & \cdots \\ y_{n-1} &= (c_n t + c_{n-1}) e^{\lambda t}, \\ y_n &= c_n e^{\lambda t}. \end{cases} \tag{4.26}$$

代回 $X = PY$ 就得到原常微分方程组的解.

例 4.5.1. 试求下述常微分方程组的解:

$$\begin{cases} \frac{dx_1}{dt} = 4x_1 + 3x_2 - 4x_3, \\ \frac{dx_2}{dt} = -x_1 + 2x_3, \\ \frac{dx_3}{dt} = x_1 + y_1. \end{cases}$$

解. 此时

$$A = \begin{pmatrix} 4 & 3 & -4 \\ -1 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix}.$$

根据例 4.2.5 的结果, 令

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

则 $P^{-1}AP = \mathrm{diag}(2, J_2(1))$. 根据 (4.26) 给出的公式, 得

$$(y_1, y_2, y_3) = (c_1e^{2t}, (c_3t + c_2)e^t, c_3e^t).$$

代回 $X = PY$, 就得到

$$\begin{cases} x_1 = 2c_1e^{2t} + (c_3t - c_3 + c_2)e^t, \\ x_2 = -(c_3t + c_2)e^t, \\ x_3 = c_1e^{2t} - c_3e^t. \end{cases}$$

□

4.5.2 方阵的求幂和递归数列通项的求解

Jordan 标准形的另一个应用是求矩阵的幂. 这个很好理解. 如果 $P^{-1}AP = J$, 则 $P^{-1}A^kP = J^k$, 从而 $A^k = PJ^kP^{-1}$. 对于每个 Jordan 块, 根据牛顿二项式公式, 我们有

$$J_n(\lambda)^k = (\lambda I_n + N_n)^k = \sum_{i=0}^{\min\{n-1,k\}} \binom{k}{i} \lambda^{k-i} N_n^i. \tag{4.27}$$

注意到 N_n^i 由(4.11)给出, 这样每个 Jordan 形矩阵的幂次就很容易求出.

设数列 $\{a_n\}$ 满足递归关系 $a_n = c_1a_{n-1} + \cdots + c_ka_{n-k}$, 其中 $c_1, \cdots, c_k$ 是常数, 且假设数列的前 k 项已知, 如何求它的通项? 对于这个问题, 只要设

$$A = \begin{pmatrix} c_1 & \cdots & c_{k-1} & c_k \\ & I_{k-1} & & 0 \end{pmatrix}$$

则

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

因此, 如果我们能够求出 A 的幂次, 也就得到了递归数列的通项公式.

例 4.5.2. 求实矩阵 $A=\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ 的幂次 A^k.

简解. 若 $b=0$, 则 $A^k=a^kI_2$. 现在设 $b\neq 0$, 则 A 有两个虚特征值 $a\pm bi$. 注意到 $(1,i)^T$ 是 $a+bi$ 的一个特征向量, $(1,-i)^T$ 是 $a-bi$ 的一个特征向量. 令 $P=\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$, 则 $P^{-1}=\begin{pmatrix} \frac{1}{2} & -\frac{i}{2} \\ \frac{1}{2} & \frac{i}{2} \end{pmatrix}$ 且

$$AP=P\begin{pmatrix} a+bi & 0 \\ 0 & a-bi \end{pmatrix} \Rightarrow A^k=P\begin{pmatrix} (a+bi)^k & 0 \\ 0 & (a-bi)^k \end{pmatrix}P^{-1}. \qquad \square$$

例 4.5.3. 设数列 $\{F_n\}$ 满足条件 $F_n=aF_{n-1}+bF_{n-2}$ $(n\geq 3)$, 且已知 F_1 与 F_2, 求 F_n 的通项公式.

简解. 这个时候 $A=\begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix}$, 它的特征多项式为 x^2-ax-b.

若它有重根, 则两个根都等于 $\lambda=\frac{a}{2}$, 此时 $b=-\frac{a^2}{4}$. 令 X_2 满足 $(A-\lambda I_2)^2X_2=0$ 且 $(A-\lambda I_2)X_2\neq 0$, 令 $X_1=(A-\lambda I_2)X_2$, $P=(X_1,X_2)$, 则 $P^{-1}AP=J_2(\lambda)$.

若它有两个根 λ_1, λ_2, 令 X_1 与 X_2 分别是它们的特征向量. 则 $P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2)$.

两种情况都容易求得 F_n 的值. $\square$

4.5.3 矩阵级数

本小节设 $F=\mathbb{C}$, 即讨论的矩阵都是复矩阵.

定义 4.5.1. 方阵 A 的**谱半径**, 记为 $\rho(A)$, 是它的所有特征值模的最大值.

例 4.5.4. 若幂级数 $\sum_k c_kx^k$ 的收敛半径 $r>\rho(A)$, 则 $\sum_k c_kA^k$ 收敛.

证明. 不妨设 $A=J=J_n(\lambda)$. 记 $A^k=(a_{ij}^{(k)})$. 根据牛顿二项式定理, 当 $k\geq n$ 时, 若 $i\leq j$, 则

$$a_{ij}^{(k)}=\binom{k}{j-i}\lambda^{k-j+i}.$$

此时

$$\sum_{k\geq n}c_ka_{ij}^{(k)}=\sum_{k\geq n}\binom{k}{j-i}\lambda^{k-j+i}$$

收敛. 若 $i>j$, 则 $a_{ij}^{(k)}=0$. 故 $\sum_k c_kA^k$ 收敛. $\square$

例 4.5.5. 对于方阵 A, 指数函数

$$e^A=\sum_{k=0}^{+\infty}\frac{A^k}{k!} \tag{4.28}$$

总存在, 且它满足条件 $e^{A+B}=e^A\cdot e^B$.

例 4.5.6. 设 A 是可逆复方阵, k 是大于 1 的正整数. 试求 B 使得 $B^k=A$.

解. 注意到如果 $A' = P^{-1}AP$, $B^k = A$, 那么 $(P^{-1}BP)^k = A'$, 故我们不妨假设 $A = J_n(\lambda)$ 是 Jordan 块. 再由于 λ 总可以开 k 次根号, 我们不妨设 $A = I_n + aN_n$.

函数 $(1+t)^{\frac{1}{k}}$ 的 Taylor 展开为

$$(1+t)^{\frac{1}{k}} = 1 + \sum_{i=1}^{+\infty} \frac{\frac{1}{k}(\frac{1}{k}-1)\cdots(\frac{1}{k}-i+1)}{i!} t^i.$$

将 $t = aN_n$ 代入, 注意到 $(aN_n)^n = 0$, 就得到

$$(I_n + aN_n)^{\frac{1}{k}} = I_n + \sum_{i=1}^{n-1} \frac{\frac{1}{k}(\frac{1}{k}-1)\cdots(\frac{1}{k}-i+1)}{i!} a^i N_n^i. \qquad \square$$

习　　题

习题 4.5.1. 求解常微分方程组 $\frac{d}{dt}X = AX$, 其中

$$A = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 4 & 0 \\ 2 & 3 & 0 & 4 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}.$$

习题 4.5.2. 设数列 $\{a_n\}$ 满足条件 $a_n = 3a_{n-1} - 3a_{n-2} + a_{n-3}$ $(n \geq 3)$, 且已知 $a_1 = a_2 = a_3 = 1$, 求 a_n 的通项公式.

习题 4.5.3. 设 A, B 为方阵. 证明: 若 $\mathrm{diag}(A, A)$ 与 $\mathrm{diag}(B, B)$ 相似, 则 A 与 B 相似.

习题 4.5.4. 设 A 为复方阵 A. 证明: 若对所有正整数 k 均有 $\mathrm{tr}(A^k) = 0$, 则 A 为幂零阵.

习题 4.5.5. 证明: $\lim\limits_{k\to+\infty} A^k = 0$ 当且仅当 $\rho(A) < 1$.

习题 4.5.6. 求所有实矩阵 A, 使得 $e^A = I$.