

第六章 初等标准形

§6.1 对角化

除非做出特别说明, 我们总是假定 F 是任意一个域, V 是 F 上的有限维线性空间. 我们考虑下面的问题:

- 给定线性变换 $T \in L(V)$, 寻找 V 的有序基 $\mathcal{B}$, 使得矩阵 $[T]_{\mathcal{B}}$ 尽量简单.
- 给定矩阵 $A \in F^{n \times n}$, 寻找可逆矩阵 $P \in GL_n(F)$, 使得矩阵 $P^{-1}AP$ 尽量简单.

矩阵 $[T]_{\mathcal{B}}$ 和 $P^{-1}AP$ 越简单, 关于 T 和 A 的很多问题(例如求秩、行列式、核、像集等)就越容易回答.

这两个问题在本质上是同一个问题. 事实上, T 在 V 的所有有序基下的矩阵构成一个相似等价类. 我们选择对线性变换讨论问题. 对于矩阵 A 的情况, 我们把对线性变换的结果应用到 L_A 来回答.

我们先把“简单”的矩阵取为对角矩阵.

定义 6.1. (1) 如果 $T \in L(V)$ 在 V 的某个有序基下的矩阵是对角矩阵, 则称 T 是可对角化的.

(2) 如果 $A \in F^{n \times n}$ 相似于对角矩阵, 则称 A 是可对角化的.

容易看出, A 可对角化 $\iff L_A$ 可对角化.

是否每个 T 总可对角化? 如果不是, 如何判断 T 是否可对角化? 我们引入下面的概念.

定义 6.2. 设 $T \in L(V)$. 对于 $c \in F$, 我们记

$$V_c := \text{Ker}(c\text{id}_V - T) = \{\alpha \in V \mid T\alpha = c\alpha\}.$$

如果 $V_c \neq \{0\}$, 则称 c 是 T 的特征值(eigenvalue), V_c 是 T 关于特征值 c 的特征子空间(eigenspace), V_c 中的非零向量是 T 关于特征值 c 的特征向量(eigenvector). T 的特征值的集合称为 T 的谱(spectrum), 记为 $\sigma(T)$.

命题 6.1. 设 $T \in L(V)$, $\mathcal{B}$ 是 V 的有序基. 则 $[T]_{\mathcal{B}}$ 是对角矩阵 $\iff \mathcal{B}$ 中的向量都是 T 的特征向量. 因此, T 可对角化 $\iff$ 存在由 T 的特征向量构成的基.

证明. 设 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, $A = [T]_{\mathcal{B}}$. 则

$$T\alpha_j = \sum_{i=1}^n A_{ij}\alpha_i.$$

因此, $[T]_{\mathcal{B}}$ 是对角矩阵 $\iff$ 存在 $c_j \in F$ 使得 $T\alpha_j = c_j\alpha_j \iff$ 每个 α_j 是 T 的特征向量. $\square$

对于矩阵, 我们引入类似的概念.

定义 6.3. 设 $A \in F^{n \times n}$. 对于 $c \in F$, 考虑 $F^{n \times 1}$ 的子空间

$$\text{Ker}(cI - A) = \{X \in F^{n \times 1} \mid AX = cX\}.$$

如果 $\text{Ker}(cI - A) \neq \{0\}$, 则称 c 是 A 的特征值, $\text{Ker}(cI - A)$ 是 A 关于特征值 c 的特征子空间, 其中的非零向量是 A 关于特征值 c 的特征向量. A 的特征值的集合称为 A 的谱, 记为 $\sigma(A)$.

下面讨论寻找特征值的方法.

引理 6.2. (1) 设 $T \in L(V)$. 则

$$\sigma(T) = \{c \in F \mid \det(c \text{id}_V - T) = 0\}.$$

(2) 设 $A \in F^{n \times n}$. 则

$$\sigma(A) = \{c \in F \mid \det(cI - A) = 0\}.$$

证明. (1) $c \in \sigma(T) \iff \text{Ker}(c \text{id}_V - T) \neq 0 \iff c \text{id}_V - T \text{不可逆} \iff \det(c \text{id}_V - T) = 0$.

(2) 在(1)中取 $T = L_A$. □

容易看出, 映射 $F \rightarrow F, c \mapsto \det(c \text{id}_V - T)$ 是多项式函数, 特征值即为它的“根”. 这启发我们考虑“多项式” $f = \det(x \text{id}_V - T)$. 但是, 根据我们对多项式的定义, x 是未定元, 并不是域 F 中的元素. 这时, 表达式 $x \text{id}_V - T$ 是没有意义的. 事实上, 如果 F 是有限域, 满足对任意 $c \in F$ 有 $f(c) = \det(c \text{id}_V - T)$ 的多项式 f 并不唯一. 我们利用矩阵来解决这一问题¹.

定义 6.4. 设 $A \in F^{n \times n}$. 考虑 $xI - A \in F(x)^{n \times n}$ (这里 $F(x)$ 见附录). 我们称

$$f_A := \det(xI - A) \in F[x]$$

为 A 的**特征多项式**(characteristic polynomial).

容易看出, 矩阵的特征多项式具有下面的性质:

- f_A 的首项系数是 1 (这是我们写 $xI - A$ 而不是 $A - xI$ 的原因), $\deg f_A = n$, 常数项为 $(-1)^n \det(A)$, x^{n-1} 的系数为 $-\text{tr}(A)$.
- A 的特征值即为 f_A 的根. 因此 $|\sigma(A)| \leq n$.
- 相似的矩阵有相同的特征多项式. 事实上, 若 $B = P^{-1}AP$, 则 $xI - B = P^{-1}(xI - A)P$, 因此 $xI - B$ 与 $xI - A$ 的行列式相等.
- 可以证明 AB 与 BA 有相同的特征多项式:

$$\begin{aligned} \begin{bmatrix} xI & A \\ B & I \end{bmatrix} \begin{bmatrix} I & 0 \\ -B & xI \end{bmatrix} &= \begin{bmatrix} xI - AB & xA \\ 0 & xI \end{bmatrix}, \\ \begin{bmatrix} I & 0 \\ -B & xI \end{bmatrix} \begin{bmatrix} xI & A \\ B & I \end{bmatrix} &= \begin{bmatrix} xI & A \\ 0 & xI - BA \end{bmatrix}, \end{aligned}$$

因此 $x^n \det(xI - AB) = x^n \det(xI - BA)$. 还可以这样证明: 设 $A = P \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} Q$, P, Q 可逆.

设 $B = Q^{-1} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} P^{-1}$. 则

$$AB = P \begin{bmatrix} B_{11} & B_{12} \\ 0 & 0 \end{bmatrix} P^{-1}, \quad BA = Q^{-1} \begin{bmatrix} B_{11} & 0 \\ B_{21} & 0 \end{bmatrix} Q.$$

于是 $f_{AB} = f_{BA} = x^{n-r} f_{B_{11}}$.

定义 6.5. 设 $T \in L(V)$, $\mathcal{B}$ 是 V 的有序基. 定义 T 的**特征多项式**为 $f_T = f_{[T]_{\mathcal{B}}}$.

对于不同的 $\mathcal{B}$, $[T]_{\mathcal{B}}$ 是相似的, 从而有相同的 $f_{[T]_{\mathcal{B}}}$. 因此 f_T 的定义不依赖于 $\mathcal{B}$ 的选取. 我们有:

- f_T 的首项系数是 1, $\deg f_T = \dim V$.

¹另一解决途径如下: 考虑张量积 $V \otimes_F F(x)$, 它是域 $F(x)$ 上的线性空间, $x \text{id} - T$ 可以视为 $V \otimes_F F(x)$ 上的线性变换, 于是可以定义 T 的特征多项式为 $f_T = \det(x \text{id} - T)$.

- 由于 $\sigma(T) = \sigma([T]_{\mathcal{B}})$, 所以 T 的特征值就是 f_T 的根, 因此 $|\sigma(T)| \leq \dim V$.
- 如果 T 可对角化, $[T]_{\mathcal{B}} = \text{diag}(c_1, \dots, c_n)$, 则 $f_T = (x - c_1) \cdots (x - c_n)$. 因此对角矩阵 $[T]_{\mathcal{B}}$ 的对角元在差一个排列的意义下被 T 决定. 另一方面, 通过调整有序基 $\mathcal{B}$ 中向量的排列, 可以得到对角矩阵 $[T]_{\mathcal{B}}$ 的对角元的任意排列.

为了讨论特征子空间的性质, 我们插入关于直和的部分内容(见教材§6.6). 设 $V_1, \dots, V_k$ 是域 F 上的线性空间(可以是无穷维的). 考虑集合

$$V_1 \times \cdots \times V_k := \{(\alpha_1, \dots, \alpha_k) \mid \alpha_i \in V_i\}.$$

这里 $(\alpha_1, \dots, \alpha_k)$ 为向量的 n 元有序组. 严格地说, 它指映射 $\{1, \dots, k\} \rightarrow \bigcup_{i=1}^k V_i, i \mapsto \alpha_i$. 在 $V_1 \times \cdots \times V_k$ 上定义加法和纯量乘法为

$$\begin{aligned} (\alpha_1, \dots, \alpha_k) + (\beta_1, \dots, \beta_k) &= (\alpha_1 + \beta_1, \dots, \alpha_k + \beta_k), \\ c(\alpha_1, \dots, \alpha_k) &= (c\alpha_1, \dots, c\alpha_k). \end{aligned}$$

则 $V_1 \times \cdots \times V_k$ 成为域 F 上的线性空间, 称为 $V_1, \dots, V_k$ 的**外直和**(external direct sum), 记为 $V_1 \oplus \cdots \oplus V_k$ 或 $\bigoplus_{i=1}^k V_i$. 对任意 i , 考虑单线性映射

$$\tau_i: V_i \rightarrow \bigoplus_{i=1}^k V_i, \quad \tau_i(\alpha_i) = (0, \dots, \alpha_i, \dots, 0). \quad (6.1)$$

引理 6.3. 如果 $\mathcal{B}_i$ 是 V_i 的基, 则 $\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$ 是 $\bigoplus_{i=1}^k V_i$ 的基. 从而如果 $\dim V_i < \infty$, 则 $\dim \bigoplus_{i=1}^k V_i < \infty$, 并且

$$\dim \bigoplus_{i=1}^k V_i = \sum_{i=1}^k \dim V_i.$$

证明. 首先, 对任意 $(\alpha_1, \dots, \alpha_k) \in \bigoplus_{i=1}^k V_i$, 由于 $\alpha_i \in \text{span } \mathcal{B}_i$, 所以 $\tau_i(\alpha_i) \in \text{span } \tau_i(\mathcal{B}_i)$, 从而

$$(\alpha_1, \dots, \alpha_k) = \sum_{i=1}^k \tau_i(\alpha_i) \in \text{span } \bigcup_{i=1}^k \tau_i(\mathcal{B}_i).$$

因此 $\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$ 生成全空间. 其次, 如果 $\alpha_{ij} \in \mathcal{B}_i$, 并且有限线性组合 $\sum_{i=1}^k \sum_j c_{ij} \tau_i(\alpha_{ij}) = 0$, 由于

$$\sum_{i=1}^k \sum_j c_{ij} \tau_i(\alpha_{ij}) = \sum_{i=1}^k \tau_i \left(\sum_j c_{ij} \alpha_{ij} \right) = \left(\sum_j c_{1j} \alpha_{1j}, \dots, \sum_j c_{kj} \alpha_{kj} \right),$$

所以 $\sum_j c_{ij} \alpha_{ij} = 0$, 从而 $c_{ij} = 0$. 因此 $\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$ 线性无关. $\square$

现在设 $V_1, \dots, V_k$ 是域 F 上的线性空间 V 的子空间. 考虑从外直和 $\bigoplus_{i=1}^k V_i$ 到 V 的线性映射

$$\Phi: \bigoplus_{i=1}^k V_i \rightarrow V, \quad (\alpha_1, \dots, \alpha_k) \mapsto \sum_{i=1}^k \alpha_i.$$

容易看出, $\text{Im}(\Phi) = \sum_{i=1}^k V_i$. 如果 Φ 是单映射, 则称子空间 $V_1, \dots, V_k$ 是**无关**的, 并称它们的和 $\sum_{i=1}^k V_i$ 为 $V_1, \dots, V_k$ 的**内直和**(internal direct sum). 此时, Φ 诱导了外直和与内直和之间的线性同构

$$\bigoplus_{i=1}^k V_i \cong \sum_{i=1}^k V_i. \quad (6.2)$$

引理 6.4. 设 $V_1, \dots, V_k \subset V$ 是子空间. 以下断言等价:

- (1) $V_1, \dots, V_k$ 无关.
- (2) $\alpha_i \in V_i, \sum_{i=1}^k \alpha_i = 0 \implies \alpha_i = 0$.

(3) 对任意 $2 \leq i \leq k$ 有 $V_i \cap (V_1 + \cdots + V_{i-1}) = \{0\}$.

(4) 任给 V_i 的基 $\mathcal{B}_i$, 总有 $\mathcal{B}_1, \dots, \mathcal{B}_k$ 互不相交, 并且 $\bigcup_{i=1}^k \mathcal{B}_i$ 是 $\sum_{i=1}^k V_i$ 的基.

如果 $\dim V < \infty$, 则它们还等价于

(5) $\dim \sum_{i=1}^k V_i = \sum_{i=1}^k \dim V_i$.

证明. “(1) $\iff$ (2)”. 显然.

“(1) $\implies$ (3)”. 设 $\alpha \in V_i \cap (V_1 + \cdots + V_{i-1})$. 则存在 $\alpha_1 \in V_1, \dots, \alpha_{i-1} \in V_{i-1}$ 使 $\alpha = \alpha_1 + \cdots + \alpha_{i-1}$, 即 $\Phi(\alpha_1, \dots, \alpha_{i-1}, -\alpha, 0, \dots, 0) = 0$. 由于 Φ 是单射, 所以 $\alpha = 0$.

“(3) $\implies$ (2)”. 设 $\alpha_i \in V_i, \sum_{i=1}^k \alpha_i = 0$. 如果某个 $\alpha_i \neq 0$, 取最大指标 i_0 使 $\alpha_{i_0} \neq 0$. 则 $i_0 \geq 2$, 并且

$$\alpha_{i_0} = -\sum_{i=1}^{i_0-1} \alpha_i \in V_{i_0} \cap (V_1 + \cdots + V_{i_0-1}) = \{0\},$$

矛盾. 因此所有 $\alpha_i = 0$.

“(1) $\iff$ (4)”. 注意 Φ 把 $\bigcup_{i=1}^k \tau_i(\mathcal{B}_i)$ 映为 $\bigcup_{i=1}^k \mathcal{B}_i$. 所以 Φ 单 $\iff \mathcal{B}_1, \dots, \mathcal{B}_k$ 互不相交并且 $\bigcup_{i=1}^k \mathcal{B}_i$ 线性无关.

“(1) $\iff$ (5)”. 我们总有

$$\dim \sum_{i=1}^k V_i = \dim \operatorname{Im}(\Phi) = \dim \bigoplus_{i=1}^k V_i - \dim \operatorname{Ker}(\Phi) = \sum_{i=1}^k \dim V_i - \dim \operatorname{Ker}(\Phi).$$

因此 Φ 单 $\iff$ (5) 成立. □

注意当 $k = 2$ 时, V_1 与 V_2 无关 $\iff V_1 \cap V_2 = \{0\}$; 当 $k > 2$ 时, $V_1, \dots, V_k$ 无关 $\implies \bigcap_{i=1}^k V_i = \{0\}$, 但反过来不对.

内直和与外直和有非常紧密的联系. 首先, 给定线性空间 $V_1, \dots, V_k$, 我们可以利用 (6.1) 中定义的单映射把 $\alpha_i \in V_i$ 与 $(0, \dots, \alpha_i, \dots, 0) \in \bigoplus_{i=1}^k V_i$ 等同起来, 从而把 V_i 视为 $\bigoplus_{i=1}^k V_i$ 的子空间. 这样, 我们有 $(\alpha_1, \dots, \alpha_k) = \sum_{i=1}^k \alpha_i$. 注意到 $\sum_{i=1}^k \alpha_i = 0 \implies \alpha_i = 0$, 即子空间 $V_1, \dots, V_k$ 是无关的, 并且 $\bigoplus_{i=1}^k V_i = \sum_{i=1}^k V_i$. 因此, 在这一观点下, 外直和 $\bigoplus_{i=1}^k V_i$ 成为它的子空间 $V_1, \dots, V_k$ 的内直和. 另一方面, 给定线性空间 V 的无关的子空间 $V_1, \dots, V_k$, 同构 (6.2) 说明内直和 $\sum_{i=1}^k V_i$ 中的任意向量 α 可以唯一分解成 $\alpha = \sum_{i=1}^k \alpha_i, \alpha_i \in V_i$. 我们利用同构 (6.2) 把 α 与 $(\alpha_1, \dots, \alpha_k)$ 等同起来, 从而把内直和 $\sum_{i=1}^k V_i$ 等同为外直和 $\bigoplus_{i=1}^k V_i$. 基于上述原因, 我们把内直和 $\sum_{i=1}^k V_i$ 也记为 $\bigoplus_{i=1}^k V_i$, 并且在没有歧义的情况下, 把内直和与外直和都简称为直和.

例 6.1. 设 $\operatorname{char} F \neq 2, V = F^{n \times n}$,

$$V_1 = \{A \in F^{n \times n} \mid A^t = A\}, \quad (\text{对称矩阵})$$

$$V_2 = \{A \in F^{n \times n} \mid A^t + A = 0\}. \quad (\text{反对称矩阵})$$

则 $V_1 \cap V_2 = \{0\}$, 从而 $V_1 + V_2$ 是直和. 我们断言 $V = V_1 \oplus V_2$. 事实上, 对任意 $A \in V$ 有

$$A_1 = \frac{1}{2}(A + A^t) \in V_1, \quad A_2 = \frac{1}{2}(A - A^t) \in V_2,$$

并且 $A = A_1 + A_2$. □

例 6.2. 设 $\{\alpha_1, \dots, \alpha_n\}$ 是 V 的基. 记 $V_i = \operatorname{span}\{\alpha_i\}$. 则 $V = \bigoplus_{i=1}^n V_i$. □

下面我们继续假设 $\dim V < \infty$. 设 $T \in L(V)$. V 的子空间 W 称为是 T 的不变子空间, 如果 $T(W) \subset W$. 例如, $\operatorname{Ker}(T)$ 和 $\operatorname{Im}(T)$ 是不变子空间.

命题 6.5. 设 $T \in L(V)$. 则 T 可对角化 $\iff V$ 可以分解成一维不变子空间的直和.

证明. “ $\implies$ ”. 设 T 可对角化. 则存在由 T 的特征向量构成的基 $\{\alpha_1, \dots, \alpha_n\}$. 于是 $V_i = \text{span}\{\alpha_i\}$ 是一维不变子空间, 并且 $V = \bigoplus_{i=1}^n V_i$.

“ $\impliedby$ ”. 设 V 可以分解成一维不变子空间的直和 $V = \bigoplus_{i=1}^n V_i$. 取 $\alpha_i \in V_i \setminus \{0\}$. 则 α_i 是特征向量, 并且 $\{\alpha_1, \dots, \alpha_n\}$ 是 V 的基. 因此 T 可对角化. $\square$

特征子空间的一个基本性质如下.

命题 6.6. 设 $T \in L(V)$. 则特征子空间 $\{V_c \mid c \in \sigma(T)\}$ 是无关的. 于是, 可以考虑直和 $\bigoplus_{c \in \sigma(T)} V_c$.

证明. 设 $\sigma(T) = \{c_1, \dots, c_k\}$. 由 Lagrange 插值公式, 存在 $f_1, \dots, f_k \in F[x]$ 满足 $f_j(c_i) = \delta_{ij}$. 设 $\alpha_i \in V_{c_i}$, $\sum_{i=1}^k \alpha_i = 0$. 两边取 $f_j(T)$ 得

$$0 = \sum_{i=1}^k f_j(T) \alpha_i = \sum_{i=1}^k f_j(c_i) \alpha_i = \alpha_j.$$

$\square$

对于 $f \in F[x]$, 我们把 f 的根 c 的重数记为 $m(c, f)$. 下面是特征子空间的另一个基本性质.

命题 6.7. 设 $T \in L(V)$, $c \in \sigma(T)$. 则

$$1 \leq \dim V_c \leq m(c, f_T).$$

证明. 显然 $\dim V_c \geq 1$. 取 V_c 的有序基 $\{\alpha_1, \dots, \alpha_d\}$, 并扩充为 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$. 由于当 $1 \leq j \leq d$ 时有 $T\alpha_j = c\alpha_j$, 所以 $[T]_{\mathcal{B}}$ 形如 $\begin{bmatrix} cI_d & B \\ 0 & D \end{bmatrix}$, 这里 $B \in F^{d \times (n-d)}$, $D \in F^{(n-d) \times (n-d)}$. 因此

$$\begin{aligned} f_T &= f_{[T]_{\mathcal{B}}} = \det \left(xI_n - \begin{bmatrix} cI_d & B \\ 0 & D \end{bmatrix} \right) = \det \begin{bmatrix} (x-c)I_d & -B \\ 0 & xI_{n-d} - D \end{bmatrix} \\ &= \det((x-c)I_d) \det(xI_{n-d} - D) = (x-c)^d f_D. \end{aligned}$$

因此 $m(c, f_T) \geq d$. $\square$

$m(c, f_T)$ 和 $\dim V_c$ 分别称为 T 的特征值 c 的代数重数和几何重数.

下面我们讨论线性映射可对角化的条件.

定理 6.8. 设 $T \in L(V)$. 以下论断等价:

- (1) T 可对角化.
- (2) $V = \bigoplus_{c \in \sigma(T)} V_c$.
- (3) $\dim V = \sum_{c \in \sigma(T)} \dim V_c$.
- (4) $f_T = \prod_{c \in \sigma(T)} (x-c)^{\dim V_c}$.

证明. “(1) $\iff$ (2)”.

$$T \text{ 可对角化} \iff \text{存在由 } T \text{ 的特征向量构成的基} \iff \bigoplus_{c \in \sigma(T)} V_c = V.$$

“(2) $\implies$ (4)”. 设 $\sigma(T) = \{c_1, \dots, c_k\}$, 并记 $d_i = \dim V_{c_i}$. 设 $\mathcal{B}_i = \{\alpha_{i1}, \dots, \alpha_{id_i}\}$ 是 V_{c_i} 的有序基. 则 $\mathcal{B} = \{\alpha_{11}, \dots, \alpha_{1d_1}, \dots, \alpha_{k1}, \dots, \alpha_{kd_k}\}$ 是 V 的有序基. 由于 $T\alpha_{ij} = c_i\alpha_{ij}$, 所以 $[T]_{\mathcal{B}} = \text{diag}(c_1I_{d_1}, \dots, c_kI_{d_k})$. 因此 $f_T = f_{[T]_{\mathcal{B}}} = (x-c_1)^{d_1} \cdots (x-c_k)^{d_k}$.

“(4) $\implies$ (3)”. $\dim V = \deg f_T = \sum_{c \in \sigma(T)} \dim V_c$.

“(3) $\implies$ (2)”. $\dim \bigoplus_{c \in \sigma(T)} V_c = \sum_{c \in \sigma(T)} \dim V_c = \dim V$. 因此 $V = \bigoplus_{c \in \sigma(T)} V_c$. $\square$

推论 6.9. 设 $T \in L(V)$. 以下论断等价:

- (1) T 可对角化.
 - (2) f_T 可以分解为一次式的乘积, 并且任意特征值的代数重数等于几何重数.
 - (3) f_T 可以分解为一次式的乘积, 并且任意代数重数大于1的特征值的代数重数等于几何重数.
- 特别地, 如果 $|\sigma(T)| = \dim V$, 则 T 可对角化.

证明. “(1) $\iff$ (2)”和“(2) $\implies$ (3)”. 显然.

“(3) $\implies$ (2)”. 只需注意到当 $m(c, f_T) = 1$ 时, 由命题6.7总有 $\dim V_c = 1$.

如果 $|\sigma(T)| = \dim V$, 则 f_T 可以分解为一次式的乘积并且无重根, 从而 T 可对角化. $\square$

推论 6.10. 设 $A \in F^{n \times n}$.

- (1) A 可对角化 $\iff f_A$ 可以分解为一次式的乘积, 并且任意代数重数大于1的特征值 c 的代数重数等于其几何重数 $\dim \text{Ker}(cI - A)$. 特别地, 如果 $|\sigma(A)| = n$, 则 A 可对角化.
- (2) 当 A 可对角化时, 如果 $f_A = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$, $\{\alpha_{i1}, \dots, \alpha_{id_i}\}$ 是 $\text{Ker}(c_i I - A)$ 的基, 则矩阵 $P = [\alpha_{11}, \dots, \alpha_{1d_1}, \dots, \alpha_{k1}, \dots, \alpha_{kd_k}] \in F^{n \times n}$ 可逆, 并且

$$P^{-1}AP = \text{diag}(c_1 I_{d_1}, \dots, c_k I_{d_k}).$$

证明. 在推论6.9中取 $T = L_A$ 即得(1). 当(2)的条件成立时, 由于 $F^{n \times 1} = \bigoplus_{i=1}^k \text{Ker}(c_i I - A)$, 所以 P 的列向量构成 $F^{n \times 1}$ 的基, 因此 P 可逆. 由于 $A\alpha_{ij} = c_i \alpha_{ij}$, 所以

$$\begin{aligned} AP &= [A\alpha_{11}, \dots, A\alpha_{1d_1}, \dots, A\alpha_{k1}, \dots, A\alpha_{kd_k}] \\ &= [c_1 \alpha_{11}, \dots, c_1 \alpha_{1d_1}, \dots, c_k \alpha_{k1}, \dots, c_k \alpha_{kd_k}] \\ &= P \text{diag}(c_1 I_{d_1}, \dots, c_k I_{d_k}), \end{aligned}$$

即 $P^{-1}AP = \text{diag}(c_1 I_{d_1}, \dots, c_k I_{d_k})$. $\square$

注意 A 不可对角化可能由两个原因导致:

- 如果 f_A 不能分解为一次式的乘积, 则 A 不可对角化. (可以通过扩大 F 来解决)
- 如果某特征值的几何重数小于代数重数, 即特征向量不足以构成空间的基, 则 A 不可对角化. (不能通过扩大 F 来解决, 更本质)

如果 F 是代数闭域, 则第一个原因不会出现.

例 6.3. 设 $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, 其中 $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. 则

$$f_A = \det(xI - A) = \det \begin{bmatrix} x - \cos \theta & \sin \theta \\ -\sin \theta & x - \cos \theta \end{bmatrix} = x^2 - (2 \cos \theta)x + 1.$$

容易看出, f_A 无实根. 所以 A (在 $\mathbb{R}$ 上) 不可对角化.

如果视 $A \in \mathbb{C}^{2 \times 2}$, 则 $\sigma(A) = \{e^{i\theta}, e^{-i\theta}\}$, $|\sigma(A)| = 2$, 所以 A 在 $\mathbb{C}$ 上可对角化. 关于特征值 $e^{i\theta}$ 和 $e^{-i\theta}$ 的特征子空间

$$\text{Ker}(e^{i\theta} I - A) = \text{Ker} \begin{bmatrix} i \sin \theta & \sin \theta \\ -\sin \theta & i \sin \theta \end{bmatrix}$$

和

$$\text{Ker}(e^{-i\theta}I - A) = \text{Ker} \begin{bmatrix} -i \sin \theta & \sin \theta \\ -\sin \theta & -i \sin \theta \end{bmatrix}$$

分别有基

$$\left\{ \alpha_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix} \right\}, \quad \left\{ \alpha_2 = \begin{bmatrix} 1 \\ i \end{bmatrix} \right\}.$$

设

$$P = [\alpha_1, \alpha_2] = \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix}.$$

则

$$P^{-1}AP = \text{diag}(e^{i\theta}, e^{-i\theta}).$$

注意 P 与 θ 无关.

□

例 6.4. 设 $A = \begin{bmatrix} \lambda & a & b \\ 0 & \lambda & c \\ 0 & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{3 \times 3}$, 这里 $\lambda, a, b, c \in \mathbb{R}$ 并且 a, b, c 不全为 0. 则

$$f_A = \det(xI - A) = \det \begin{bmatrix} x - \lambda & -a & -b \\ 0 & x - \lambda & -c \\ 0 & 0 & x - \lambda \end{bmatrix} = (x - \lambda)^3.$$

因此 $\sigma(A) = \{\lambda\}$. 特征值 λ 的代数重数是 3, 几何重数是

$$\dim \text{Ker}(\lambda I - A) = \dim \text{Ker} \begin{bmatrix} 0 & -a & -b \\ 0 & 0 & -c \\ 0 & 0 & 0 \end{bmatrix} = 3 - \text{rank} \begin{bmatrix} 0 & -a & -b \\ 0 & 0 & -c \\ 0 & 0 & 0 \end{bmatrix} < 3.$$

所以 A 不可对角化. 并且即使视 $A \in \mathbb{C}^{3 \times 3}$, 它也不可对角化.

□

例 6.5. 设 $A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$. 则

$$f_A = \det(xI - A) = \det \begin{bmatrix} x - 5 & 6 & 6 \\ 1 & x - 4 & -2 \\ -3 & 6 & x + 4 \end{bmatrix} = (x - 1)(x - 2)^2.$$

因此 $\sigma(A) = \{1, 2\}$. 特征值 2 的代数重数是 2, 几何重数是

$$\dim \text{Ker}(2I - A) = \dim \text{Ker} \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix} = 3 - \text{rank} \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix} = 2.$$

所以 A 可对角化. 关于特征值 1 和 2 的特征子空间

$$\text{Ker}(I - A) = \text{Ker} \begin{bmatrix} -4 & 6 & 6 \\ 1 & -3 & -2 \\ -3 & 6 & 5 \end{bmatrix}$$

和

$$\text{Ker}(2I - A) = \text{Ker} \begin{bmatrix} -3 & 6 & 6 \\ 1 & -2 & -2 \\ -3 & 6 & 6 \end{bmatrix}$$

分别有基

$$\left\{ \alpha_1 = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix} \right\}, \quad \left\{ \alpha_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \alpha_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

设

$$P = [\alpha_1, \alpha_2, \alpha_3] = \begin{bmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}.$$

则

$$P^{-1}AP = \text{diag}(1, 2, 2).$$

□

附录: 有理函数域 $F(x)$.

考虑集合

$$F[x] \times (F[x] \setminus \{0\}) = \{(f, g) \mid f, g \in F[x], g \neq 0\}$$

上的关系: $(f_1, g_1) \sim (f_2, g_2) \iff f_1 g_2 = f_2 g_1$. 容易验证它是等价关系:

- $(f, g) \sim (f, g)$,
- $(f_1, g_1) \sim (f_2, g_2) \implies (f_2, g_2) \sim (f_1, g_1)$,
- $(f_1, g_1) \sim (f_2, g_2), (f_2, g_2) \sim (f_3, g_3) \implies (f_1, g_1) \sim (f_3, g_3)$.

记 $F(x)$ 为等价类的集合, 并把 (f, g) 所在的等价类记为 $\frac{f}{g}$. 注意到 $\frac{f_1}{g_1} = \frac{f_2}{g_2} \iff f_1 g_2 = f_2 g_1$. 在 $F(x)$ 中定义加法和乘法运算为

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} = \frac{f_1 g_2 + f_2 g_1}{g_1 g_2}, \quad \frac{f_1}{g_1} \cdot \frac{f_2}{g_2} = \frac{f_1 f_2}{g_1 g_2}.$$

则 $F(x)$ 构成一个域, 称为 F 系数的**有理函数域**. 其中的零元素为 $\frac{0}{1}$, 壹元素为 $\frac{1}{1}$, $\frac{f}{g}$ 的负元素为 $-\frac{f}{g}$, $\frac{f}{g} (f \neq 0)$ 的逆元素为 $\frac{g}{f}$. 虽然 $\frac{f}{g}$ 名为“有理函数”, 我们并不把它看成函数. 通过把 $f \in F[x]$ 等同于 $\frac{f}{1} \in F(x)$, 我们把 $F[x]$ 看成 $F(x)$ 的子集. 这一等同是保持加法和乘法运算的, 因此 $F[x]$ 是 $F(x)$ 的子环, F 是 $F(x)$ 的子域.

§6.2 Cayley-Hamilton定理

定理 6.11(Cayley-Hamilton). 设 V 是域 F 上的有限维线性空间, $T \in L(V)$. 则 $f_T(T) = 0$.

推论 6.12. 设 $A \in F^{n \times n}$. 则 $f_A(A) = 0$

定理6.11 $\implies$ 推论6.12的证明. $L_{f_A(A)} = f_A(L_A) = f_{L_A}(L_A) = 0 \implies f_A(A) = 0$. □

注 6.1. • 容易看出, 推论6.12 $\implies$ 定理6.11.

- 推论的错误证明: $f_A(A) = \det(AI - A) = \det(0) = 0$.

为了证明定理6.11, 我们利用环上的模做工具.

定义 6.6. 设集合 R 上给定了两个运算, 称为加法和乘法, 分别对 R 中任意两个元素 x, y 给出 R 中的元素 $x + y$ 和 xy , 并且满足如下性质:

(1) R 对加法运算构成交换群, 即:

- 对任意 $x, y \in R$ 有 $x + y = y + x$.
- 对任意 $x, y, z \in R$ 有 $x + (y + z) = (x + y) + z$.
- 在 R 中存在唯一的元素, 称为**零元素**并记为 0 , 满足 $0 + x = x$ 对任意 $x \in R$ 成立.
- 对任意 $x \in R$, 存在唯一的 $(-x) \in R$, 称为 x 的**负元素**, 满足 $x + (-x) = 0$.

(2) 对任意 $x, y, z \in R$ 有 $x(yz) = (xy)z$, $x(y + z) = xy + xz$, $(y + z)x = yx + zx$.

则称集合 R (连同它上面的加法和乘法运算) 为**环**. 如果环 R 中的乘法运算是交换的, 即对任意 $x, y \in R$ 有 $xy = yx$, 则称 R 为**交换环**. 如果在环 R 中存在唯一的元素, 称为**壹元素**并记为 1 , 满足 $1x = x1 = x$ 对任意 $x \in R$ 成立, 则称 R 为**幺环**.

如果 A 是域 F 上的代数 (例如 $F^{n \times n}$, $L(V)$, $F[x]$), 则 A 是环.

我们主要考虑交换幺环. 注意我们允许 $1 = 0$, 即 $R = \{0\}$ 的情况 (这主要是因为我们希望交换幺环模掉所有理想 (包括 R 自身) 还是交换幺环, 见??). 除此之外, 交换幺环与域的唯一区别在于非零元素在乘法运算下的逆元素不一定存在. 也就是说:

- 域是交换幺环.
- 在交换幺环 R 中, 如果 $1 \neq 0$, 并且对任意 $x \in R \setminus \{0\}$, 存在唯一的 $x^{-1} \in R$ 满足 $xx^{-1} = 1$, 则 R 是域.

例如, 整数环 $\mathbb{Z}$ 是交换幺环但不是域. 域 F 上的多项式环 $F[x]$ 是交换幺环但不是域.

设 R 是交换幺环. 记 $R^{m \times n}$ 为矩阵元在 R 中的 $m \times n$ 矩阵的集合. 对 $A, B \in R^{m \times n}$, 定义 $A + B \in R^{m \times n}$ 为

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

对 $A \in R^{m \times n}$, $B \in R^{n \times p}$, 定义 $AB \in R^{m \times p}$ 为

$$(AB)_{ij} = \sum_{k=1}^n A_{ik} B_{kj}.$$

则

$$A(B + C) = AB + AC, \quad A(BC) = (AB)C, \quad \dots$$

定义 6.7. 给定交换幺环 R . 设对于集合 V , 给定了如下两个运算:

- 加法运算 $V \times V \rightarrow V$, $(\alpha_1, \alpha_2) \mapsto \alpha_1 + \alpha_2$, 使 V 成为交换群.
- 乘法运算 $R \times V \rightarrow V$, $(c, \alpha) \mapsto c\alpha$, 满足

$$(c_1 + c_2)\alpha = c_1\alpha + c_2\alpha, \quad c(\alpha_1 + \alpha_2) = c\alpha_1 + c\alpha_2, \quad (c_1 c_2)\alpha = c_1(c_2\alpha), \quad 1\alpha = \alpha$$

对任意 $c, c_1, c_2 \in R$ 和 $\alpha, \alpha_1, \alpha_2 \in V$ 成立.

则称集合 V (连同这两个运算) 为 R 上的**模** (或 R -**模**).

例如 R^n , $R^{n \times 1}$, $R^{m \times n}$ 在自然定义的加法和乘法运算下是 R -模. 注意如果 F 是域, 则 F -模 $\iff F$ -线性空间.

例 6.6. 设 G 是交换群, 群运算为加法. 则 G 自然成为 $\mathbb{Z}$ -模. 这里 $nx = x + \dots + x$ (n 个). □

下面的例子对我们是最重要的.

例 6.7. 设 V 是域 F 上的线性空间, $T \in L(V)$. 定义乘法 $F[x] \times V \rightarrow V$ 为 $(f, \alpha) \mapsto f\alpha = f(T)\alpha$. 则 V 成为 $F[x]$ -模. 以后我们在考虑线性空间到自身的线性映射时, 同时考虑这个模结构是有好处的. $\square$

注 6.2. 事实上 V 是 $L(V)$ -模. V 上的 $F[x]$ -模结构由环同态 $F[x] \rightarrow L(V)$, $f \mapsto f(T)$ 诱导.

设 V 为交换幺环 R 上的模. 我们定义形式矩阵乘法 $V^n \times R^{n \times p} \rightarrow V^p$ 为

$$(\alpha_1, \dots, \alpha_n)A = (\beta_1, \dots, \beta_p),$$

其中

$$\beta_j = \sum_{i=1}^n A_{ij}\alpha_i.$$

容易验证下面的乘法结合律:

$$[(\alpha_1, \dots, \alpha_n)A]B = (\beta_1, \dots, \beta_p)(AB), \quad \forall (\alpha_1, \dots, \alpha_n) \in V^n, A \in R^{n \times p}, B \in R^{p \times q}.$$

定理6.11的证明. 取 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, 并设 $A = [T]_{\mathcal{B}}$. 视 V 为 $F[x]$ -模, 即对 $f \in F[x]$, $\alpha \in V$ 定义 $f\alpha = f(T)\alpha$. 则

$$(\alpha_1, \dots, \alpha_n)A = (T\alpha_1, \dots, T\alpha_n) = (x\alpha_1, \dots, x\alpha_n) = (\alpha_1, \dots, \alpha_n)xI_n,$$

即

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

另一方面, 存在矩阵 $B \in F[x]^{n \times n}$ 使得

$$(xI_n - A)B = \det(xI_n - A)I_n = f_T I_n.$$

事实上, 取 $B = \text{adj}(xI_n - A)$ (伴随矩阵, 即余因子矩阵的转置)即可. 于是,

$$[(\alpha_1, \dots, \alpha_n)(xI_n - A)]B = (0, \dots, 0)$$

||

$$(\alpha_1, \dots, \alpha_n)[(xI_n - A)B] = (\alpha_1, \dots, \alpha_n)f_T I_n = (f_T \alpha_1, \dots, f_T \alpha_n) = (f_T(T)\alpha_1, \dots, f_T(T)\alpha_n).$$

因此 $f_T(T)\alpha_1 = \dots = f_T(T)\alpha_n = 0$. 这推出 $f_T(T) = 0$. $\square$

§6.3 最小多项式

设 V 是有限维 F -线性空间, $T \in L(V)$. 如果 $f \in F[x]$ 满足 $f(T) = 0$, 则称 f 是 T 的**零化多项式**(annihilating polynomial). 由Cayley-Hamilton定理, f_T 是 T 的零化多项式. 我们把 T 的所有零化多项式的集合记为 M_T .

引理 6.13. M_T 是 $F[x]$ 理想.

证明. 显然 M_T 非空.

M_T 是子空间: $f, g \in M_T, c \in F \implies f + cg \in M_T$.

M_T 是理想: $f \in M_T, g \in F[x] \implies fg \in M_T$. $\square$

注 6.3. 由于 $f_T \in M_T$, 所以 M_T 是非零理想. 该事实也可以不用Cayley-Hamilton定理证明: 设 $\dim V = n$, 则 $\dim L(V) = n^2$, 从而 $I, T, \dots, T^{n^2}$ 线性相关, 即存在不全为0的 $c_0, \dots, c_{n^2}$ 使得

$$c_0 I + c_1 T + \dots + c_{n^2} T^{n^2} = 0.$$

令 $f = \sum_{i=0}^{n^2} c_i x^i \neq 0$. 则 $f \in M$. $\square$

回忆 $F[x]$ 的非零理想都是主理想, 并且存在唯一的首项系数是1的生成元 p_T , 即

$$f(T) = 0 \iff p_T \mid f.$$

p_T 称为 T 的**最小多项式**. 当 $V = \{0\}$ 时, 我们约定 $f_T = p_T = 1$. 注意当 $V \neq \{0\}$ 时, 总有 $\deg p_T \geq 1$ 并且 $p_T \mid f_T$. 由理想生成元的性质, $p \in F[x]$ 是 T 的最小多项式 $\iff$ 以下两条成立:

- p 是 T 的零化多项式并且首项系数是1;
- 任何 T 的非零零化多项式的次数 $\geq \deg p$.

类似地, 如果 $A \in F^{n \times n}$, 则

$$M_A := \{f \in F[x] \mid f(A) = 0\}$$

是 $F[x]$ 的非零理想. 它的首项系数是1的生成元 p_A 称为 A 的**最小多项式**. 性质:

- $p_A \mid f_A$.
- $p_A = p_{L_A}$, $p_T = p_{[T]_{\mathcal{B}}}$.
- A 与 B 相似 $\implies p_A = p_B$: $f(P^{-1}AP) = P^{-1}f(A)P$.
- 设有两个域 $F_1 \subset F_2$, $A \in F_1^{n \times n}$, 则 A 作为 $F_1^{n \times n}$ 和 $F_2^{n \times n}$ 中的矩阵有相同的最小多项式: 设两种看法的最小多项式分别为 $p_1 \in F_1[x]$ 和 $p_2 \in F_2[x]$. 显然有 $p_2 \mid p_1$. 设 $\deg p_2 = k$, $p_2 = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$. 则

$$A^k + a_{k-1}A^{k-1} + \cdots + a_1A + a_0I = 0.$$

把这一等式视为关于未知量 $a_{k-1}, \dots, a_0$ 的线性方程组, 它有 n^2 个方程, 系数和常数项为 F_1 中的元素, 方程组在 F_2 中有解, 因此在 F_1 中也有解, 即存在 $b_{k-1}, \dots, b_0 \in F_1$ 满足 $A^k + b_{k-1}A^{k-1} + \cdots + b_1A + b_0I = 0$. 于是 $q := x^k + b_{k-1}x^{k-1} + \cdots + b_1x + b_0 \in F_1[x]$ 满足 $q(A) = 0$. 因此 $p_1 \mid q$. 再结合 $\deg p_2 = \deg q$ 即得 $p_1 = p_2$.

命题 6.14. p_T 与 f_T 有相同的根.

证明. 由于 $p_T \mid f_T$, 所以 p 的根总是 f 的根. 反过来, 设 $f(c) = 0$. 则 c 是 T 的特征值, 既存在 $\alpha \in V \setminus \{0\}$ 满足 $T\alpha = c\alpha$. 设 $p = x^k + a_{k-1}x^{k-1} + \cdots + a_1x + a_0$. 则

$$\begin{aligned} p(T)\alpha &= (T^k + a_{k-1}T^{k-1} + \cdots + a_1T + a_0I)\alpha \\ &= T^k\alpha + a_{k-1}T^{k-1}\alpha + \cdots + a_1T\alpha + a_0\alpha \\ &= c^k\alpha + a_{k-1}c^{k-1}\alpha + \cdots + a_1c\alpha + a_0\alpha \\ &= (c^k + a_{k-1}c^{k-1} + \cdots + a_1c + a_0)\alpha \\ &= p(c)\alpha. \end{aligned}$$

由于 $p(T) = 0$, 所以 $p(c)\alpha = 0$. 但 $\alpha \neq 0$. 所以 $p(c) = 0$. □

在命题中取 $T = L_A$, 即得 p_A 与 f_A 有相同的根.

由Cayley-Hamilton定理和命题6.14, 如果 $\sigma(T) = \{c_1, \dots, c_k\}$, $f_T = (x - c_1)^{d_1} \cdots (x - c_k)^{d_k}$, 则 $p_T = (x - c_1)^{r_1} \cdots (x - c_k)^{r_k}$, 其中 $1 \leq r_i \leq d_i$. 实际上, 任何在这个范围中的 r_i 都可以出现.

定理 6.15. 对于 $T \in L(V)$, TFAE:

- (1) T 可对角化.
- (2) p_T 可以分解为互不相同的monic一次式的乘积.
- (3) $p_T = \prod_{c \in \sigma(T)} (x - c)$.

为了证明定理6.15, 我们先证明:

引理 6.16. 设 $\dim V < \infty$, $T_1, \dots, T_k \in L(V)$. 则

$$\dim \operatorname{Ker}(T_1 \cdots T_k) \leq \sum_{i=1}^k \dim \operatorname{Ker}(T_i).$$

证明. 当 $k = 1$ 时无需证明. 设 $k = 2$. 考虑 $T_2|_{\operatorname{Ker}(T_1 T_2)} : \operatorname{Ker}(T_1 T_2) \rightarrow V$. 注意到

$$\operatorname{Ker}(T_2|_{\operatorname{Ker}(T_1 T_2)}) = \operatorname{Ker}(T_1 T_2) \cap \operatorname{Ker}(T_2) = \operatorname{Ker}(T_2)$$

并且 $\operatorname{Im}(T_2|_{\operatorname{Ker}(T_1 T_2)}) \subset \operatorname{Ker}(T_1)$. 因此,

$$\begin{aligned} \dim \operatorname{Ker}(T_1 T_2) &= \dim \operatorname{Ker}(T_2|_{\operatorname{Ker}(T_1 T_2)}) + \dim \operatorname{Im}(T_2|_{\operatorname{Ker}(T_1 T_2)}) \\ &\leq \dim \operatorname{Ker}(T_2) + \dim \operatorname{Ker}(T_1). \end{aligned}$$

假设 $k > 2$, 并且引理当 $< k$ 时成立. 则

$$\begin{aligned} \dim \operatorname{Ker}(T_1 \cdots T_k) &\leq \dim \operatorname{Ker}(T_1 \cdots T_{k-1}) + \dim \operatorname{Ker}(T_k) \\ &\leq \sum_{i=1}^{k-1} \dim \operatorname{Ker}(T_i) + \dim \operatorname{Ker}(T_k) = \sum_{i=1}^k \dim \operatorname{Ker}(T_i). \end{aligned}$$

□

定理6.15的证明. 设 $\sigma(T) = \{c_1, \dots, c_k\}$.

“(1) $\implies$ (2)”. 记 $g = \prod_{i=1}^k (x - c_i)$. 取 $\mathcal{B}$ 使 $[T]_{\mathcal{B}} = \operatorname{diag}(c_1 I_{d_1}, \dots, c_k I_{d_k})$. 容易验证 $g([T]_{\mathcal{B}}) = 0$. 从而 $g(T) = 0$. 因此 $p_T | g$.

“(2) $\implies$ (3)”. 由命题6.14容易看出.

“(3) $\implies$ (1)”. 设(3)成立. 则 $\prod_{i=1}^k (T - c_i I) = 0$. 由引理, 有

$$\dim V = \dim \operatorname{Ker} \left(\prod_{i=1}^k (T - c_i I) \right) \leq \sum_{i=1}^k \dim \operatorname{Ker}(T - c_i I) = \sum_{i=1}^k \dim V_{c_i} = \dim \bigoplus_{i=1}^k V_{c_i}.$$

所以只能有 $V = \bigoplus_{i=1}^k V_{c_i}$. 因此 T 可对角化.

□

在定理中取 $T = L_A$, 即得

推论 6.17. 对于 $A \in F^{n \times n}$, TFAE:

- (1) A 可对角化.
- (2) p_A 可以分解为互不相同的monic一次式的乘积.
- (3) $p_A = \prod_{c \in \sigma(A)} (x - c)$.

□

推论6.17可以用来判断 A 是否可对角化. 假设 f_A 可以分解为一次式的乘积

$$f_A = \prod_{i=1}^k (x - c_i)^{d_i}$$

(否则不能对角化). 我们知道,

$$A \text{ 可对角化} \iff \text{当 } d_i > 1 \text{ 时总有 } \dim \operatorname{Ker}(A - c_i I) = d_i.$$

推论6.17给出了另一种判断方法:

$$A \text{ 可对角化} \iff \prod_{i=1}^k (A - c_i I) = 0.$$

反过来, 推论6.17也对求最小多项式有帮助.

例 6.8. 设 $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \in \mathbb{R}^{2 \times 2}$, 其中 $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. 我们知道,

$$f_A = x^2 - (2\cos \theta)x + 1.$$

由于 f_A 无实根, 所以在 $\mathbb{R}[x]$ 中不可约. 由 Cayley-Hamilton 定理, $p_A | f_A$. 由于 $\deg p_A \geq 1$, 所以 $p_A = f_A$.

也可以这样求 p_A : 视 $A \in \mathbb{C}^{2 \times 2}$, 则 $\sigma(A) = \{e^{i\theta}, e^{-i\theta}\}$, $|\sigma(A)| = 2$, 所以 A 在 $\mathbb{C}$ 上可对角化. 由推论 6.17,

$$p_A = (x - e^{i\theta})(x - e^{-i\theta}) = f_A.$$

这里用到了最小多项式与域无关这一事实. □

例 6.9. 设 $A = \begin{bmatrix} \lambda & 0 & 1 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \in \mathbb{R}^{3 \times 3}$, 这里 $\lambda \in \mathbb{R}$. 我们知道, $f_A = (x - \lambda)^3$, 由命题 6.14, $p_A = (x - \lambda)^d$, $d \geq 1$. 容易验证, $A - \lambda I \neq 0$, $(A - \lambda I)^2 = 0$. 所以 $p_A = (x - \lambda)^2$. 这也推出 A 不可对角化. □

例 6.10. 设 $A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & 2 & -1 \\ 2 & 2 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$. 则

$$f_A = \det(xI - A) = \det \begin{bmatrix} x-3 & -1 & 1 \\ -2 & x-2 & 1 \\ -2 & -2 & x \end{bmatrix} = (x-1)(x-2)^2.$$

由 Cayley-Hamilton 定理和命题 6.14, $p_A = (x-1)(x-2)^d$, $d = 1$ 或 2 . 容易验证,

$$(A - I)(A - 2I) = \begin{bmatrix} 2 & 1 & -1 \\ 2 & 1 & -1 \\ 2 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1 \\ 2 & 0 & -1 \\ 2 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 2 & * & * \\ * & * & * \\ * & * & * \end{bmatrix} \neq 0.$$

所以 $d = 2$, $p_A = f_A$. 这也推出 A 不可对角化. □

例 6.11. 设 $A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \in \mathbb{Q}^{4 \times 4}$. 容易得到:

$$A^2 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \\ 0 & 4 & 0 & 4 \\ 4 & 0 & 4 & 0 \end{bmatrix}.$$

从而 $A^3 = 4A$. 令

$$p = x^3 - 4x = x(x+2)(x-2).$$

则 $p(A) = 0$. 这说明 $p_A | p$, 从而 A 可对角化. 容易看出, $\deg p_A \neq 1$. 我们断言, $\deg p_A \neq 2$. 事实上, 如果 $\deg p_A = 2$, 则 $p_A = x(x+2)$, $x(x-2)$ 或 $(x+2)(x-2)$. 但是 $A^2 \neq \pm 2A, 4I$. 因此 $\deg p_A \geq 3$. 只能有 $p_A = p$.

我们进一步求 f_A . 注意到 $\deg f_A = 4$. 由命题6.14, 只能有

$$f_A = x^2(x+2)(x-2), \quad x(x+2)^2(x-2) \quad \text{或} \quad x(x+2)(x-2)^2.$$

注意到特征值0的几何重数是

$$\dim \operatorname{Ker}(A) = 4 - \operatorname{rank}(A) = 2.$$

因此特征值0的代数重数 ≥ 2 . 从而只能有 $f_A = x^2(x+2)(x-2)$. 这推出 A 相似于 $\operatorname{diag}(0, 0, 2, -2)$.

□

§6.4 不变子空间

回忆: 给定 $T \in L(V)$, V 的子空间 W 称为是 T 的不变子空间, 如果 $T(W) \subset W$. 例如, $\{0\}$, V , $\operatorname{Ker}(T)$ 和 $\operatorname{Im}(T)$ 是不变子空间. 性质:

- 设 $W \subset V$ 是一维子空间. 则 W 不变 $\iff W$ 中的非零向量都是特征向量. 例如, 设

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \in \mathbb{R}^{2 \times 2},$$

其中 $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. 则 L_A 只有平凡不变子空间. 事实上, 如果 $W \subset V$ 是非平凡不变子空间, 则 $\dim W = 1$, 从而 W 中的非零向量都是 A 的特征向量. 但 A 在 $\mathbb{R}$ 中无特征值, 从而无特征向量.

- 若 $W \subset V$ 是 T 的不变子空间, 则对任意 $f \in F[x]$, W 是 $f(T)$ 的不变子空间: 设 $f = \sum_{i=0}^k a_i x^i$. 则对 $\alpha \in W$ 有

$$f(T)\alpha = \sum_{i=0}^k a_i T^i \alpha \in W.$$

- 设 $T, U \in L(V)$, $TU = UT$. 则 $\operatorname{Ker}(U)$ 和 $\operatorname{Im}(U)$ 是 T 的不变子空间:

$$\alpha \in \operatorname{Ker}(U) \implies UT\alpha = TU\alpha = 0 \implies T\alpha \in \operatorname{Ker}(U),$$

$$\alpha \in \operatorname{Im}(U) \implies \exists \beta \in V, \alpha = U\beta \implies T\alpha = TU\beta = UT\beta \in \operatorname{Im}(U).$$

注意到 U 与 T 可交换 $\implies U$ 与 $f(T)$ 可交换 $\implies \operatorname{Ker}(f(T))$ 是 U 的不变子空间. 特别地, T 的特征子空间 $V_c = \operatorname{Ker}(cI - T)$ 是 U (特别地, T)的不变子空间.

设 $W \subset V$ 是 T 的不变子空间, 则 T 诱导了 $T_W \in L(W, W)$ 和 $T_{V/W} \in L(V/W, V/W)$ 如下:

- 对 $\alpha \in W$, 定义 $T_W \alpha = T\alpha$.
- 对 $\alpha + W \in V/W$, 定义 $T_{V/W}(\alpha + W) = T\alpha + W$. 良定.

设 $\dim V < \infty$, $\mathcal{B}' = \{\alpha_1, \dots, \alpha_r\}$ 是 W 的有序基, 并扩充为 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$. 则 $\mathcal{B}'' = \{\alpha_{r+1} + W, \dots, \alpha_n + W\}$ 是 V/W 的有序基. 设 $A = [T]_{\mathcal{B}}$. 则

$$T\alpha_j = \sum_{i=1}^n A_{ij}\alpha_i.$$

当 $j \leq r$ 时有 $T\alpha_j = \sum_{i=1}^n A_{ij}\alpha_i \in W$. 所以当 $i > r$ 时有 $A_{ij} = 0$. 因此 A 形如

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

这里 $B \in F^{r \times r}$, $C \in F^{r \times (n-r)}$, $D \in F^{(n-r) \times (n-r)}$. 我们有:

- $[T_W]_{\mathcal{B}'} = B$. 这是因为

$$T_W \alpha_j = T \alpha_j = \sum_{i=1}^r A_{ij} \alpha_i = \sum_{i=1}^r B_{ij} \alpha_i, \quad j \in \{1, \dots, r\}.$$

- $[T_{V/W}]_{\mathcal{B}''} = D$. 这是因为

$$\begin{aligned} T_{V/W}(\alpha_{r+j} + W) &= T \alpha_{r+j} + W = \sum_{i=1}^r A_{i,r+j} \alpha_i + \sum_{i=r+1}^n A_{i,r+j} \alpha_i + W = \sum_{i=r+1}^n A_{i,r+j} \alpha_i + W \\ &= \sum_{i=r+1}^n A_{i,r+j}(\alpha_i + W) = \sum_{i=1}^{n-r} D_{ij}(\alpha_{r+i} + W), \quad j \in \{1, \dots, n-r\}. \end{aligned}$$

这推出:

命题 6.18. $\det(T) = \det(T_W) \det(T_{V/W})$.

证明. $\det(T) = \det(A) = \det(B) \det(D) = \det(T_W) \det(T_{V/W})$. □

类似地,

命题 6.19. $f_T = f_{T_W} f_{T_{V/W}}$.

证明.

$$f_T = f_A = \det(xI - A) = \det \begin{bmatrix} xI - B & -C \\ 0 & xI - D \end{bmatrix} = \det(xI - B) \det(xI - D) = f_B f_D = f_{T_W} f_{T_{V/W}}.$$

□

对于 $f_1, \dots, f_k \in F[x]$, 定义它们的最小公倍式(least common multiple)为理想 $\bigcap_{i=1}^k (f_i)$ 的唯一的首项系数是1的生成元, 记为 $\text{lcm}(f_1, \dots, f_k)$.

命题 6.20. $\text{lcm}(p_{T_W}, p_{T_{V/W}}) | p_T | p_{T_W} p_{T_{V/W}}$. 特别地, 如果 p_{T_W} 与 $p_{T_{V/W}}$ 互素, 则 $p_T = p_{T_W} p_{T_{V/W}}$.

证明. 容易看出, $p_T(T_W) = 0$, $p_T(T_{V/W}) = 0$. 因此 $p_{T_W} | p_T$, $p_{T_{V/W}} | p_T$. 这说明 $\text{lcm}(p_{T_W}, p_{T_{V/W}}) | p_T$. 另一方面, 对任意 $\alpha \in V$ 有 $p_{T_{V/W}} \alpha \in W$, 从而 $p_{T_W} p_{T_{V/W}} \alpha = 0$. 这说明 $p_T | p_{T_W} p_{T_{V/W}}$. □

注 6.4. 一般 $p_T = p_{T_W} p_{T_{V/W}}$ 不成立.

推论 6.21. 设 T 可对角化, 则 T_W 和 $T_{V/W}$ 可对角化.

证明. 由于 T 可对角化, 由定理6.15, p_T 为互不相同的一次式的乘积. 上一命题推出 p_{T_W} 和 $p_{T_{V/W}}$ 都是互不相同的一次式的乘积. 再由定理6.15, 即得 T_W 和 $T_{V/W}$ 可对角化. □

我们利用不变子空间讨论下面的性质.

定义 6.8. • 设 V 是有限维 F -线性空间, $\mathcal{F}$ 是 $L(V)$ 的子集. 如果存在 V 的有序基 $\mathcal{B}$, 使得所有 $T \in \mathcal{F}$ 的矩阵 $[T]_{\mathcal{B}}$ 都是对角矩阵, 则称 $\mathcal{F}$ 中的线性变换可同时对角化.

- 设 $\mathcal{M}$ 是 $F^{n \times n}$ 的子集. 如果存在可逆矩阵 $P \in \text{GL}_n(F)$, 使得对所有 $A \in \mathcal{M}$, 总有 $P^{-1}AP$ 是对角矩阵, 则称 $\mathcal{M}$ 中的矩阵可同时对角化.

定理 6.22. 对于 $\mathcal{F} \subset L(V)$, TFAE:

- (1) $\mathcal{F}$ 中的线性变换可同时对角化.
- (2) $\mathcal{F}$ 中的线性变换都是可对角化的并且两两可交换.

证明. “(1) $\implies$ (2)”. 只需证明 $\mathcal{F}$ 中的线性变换两两可交换. 取 V 的有序基 $\mathcal{B}$ 使所有 $T \in \mathcal{F}$ 的矩阵 $[T]_{\mathcal{B}}$ 是对角矩阵. 则对任意 $T, U \in \mathcal{F}$ 有 $[TU]_{\mathcal{B}} = [T]_{\mathcal{B}}[U]_{\mathcal{B}} = [U]_{\mathcal{B}}[T]_{\mathcal{B}} = [UT]_{\mathcal{B}}$. 因此 $TU =$

UT .

“(2) $\implies$ (1)”. 对 $\dim V$ 做归纳. 当 $\dim V = 1$ 时无需证明. 设 $\dim V = n \geq 2$, 并且“(2) $\implies$ (1)”当 $\dim V < n$ 时成立. 如果 $\mathcal{F}$ 中的线性变换都是恒同映射的常数倍, 则(1)自动成立. 因此假设存在 $T \in \mathcal{F}$ 不是恒同映射的常数倍. 则对任意 $c \in \sigma(T)$ 有 $\dim V_c < \dim V$. 由于 $\mathcal{F}$ 中的线性变换都与 T 可交换, 所以 V_c 是 $\mathcal{F}$ 中所有线性变换的不变子空间. 考虑 $L(V_c, V_c)$ 的子集

$$\mathcal{F}_c = \{U_{V_c} \mid U \in \mathcal{F}\}.$$

其中的线性变换都是可对角化的并且两两可交换. 由归纳假设, $\mathcal{F}_c$ 中的线性变换可同时对角化, 即存在 V_c 的基 $\mathcal{B}_c$, 其中的每个向量是 $\mathcal{F}_c$ 中所有线性变换的特征向量, 从而是 $\mathcal{F}$ 中所有线性变换的特征向量. 由于 T 可对角化, 所以 $V = \bigoplus_{c \in \sigma(T)} V_c$. 因此 $\mathcal{B} = \bigcup_{c \in \sigma(T)} \mathcal{B}_c$ 是 V 的基. 由于 $\mathcal{B}$ 中每个向量是 $\mathcal{F}$ 中所有线性变换的特征向量, 所以 $\mathcal{F}$ 中的线性变换可同时对角化. $\square$

推论 6.23. 对于 $\mathcal{M} \subset F^{n \times n}$, TFAE:

- (1) $\mathcal{M}$ 中的矩阵可同时对角化.
- (2) $\mathcal{M}$ 中的矩阵都是可对角化的并且两两可交换.

我们还考虑:

定义 6.9. (1) 如果 $T \in L(V)$ 在 V 的某个有序基下的矩阵是上三角矩阵, 则称 T 是**可三角化**的.
(2) 如果 $A \in F^{n \times n}$ 相似于上三角矩阵, 则称 A 是**可三角化**的.

注 6.5. 记

$$J_n = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \in F^{n \times n}.$$

容易验证 $J_n^{-1} = J_n$, 并且对任意 $A \in F^{n \times n}$, $J_n A J_n$ 是 A 的“中心对称”矩阵. 特别地,

$$J_n \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{bmatrix} J_n = \begin{bmatrix} a_{nn} & 0 & \cdots & 0 & 0 \\ a_{n-1,n} & a_{n-1,n-1} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{2n} & a_{2,n-1} & \cdots & a_{22} & 0 \\ a_{1n} & a_{1,n-1} & \cdots & a_{12} & a_{11} \end{bmatrix}.$$

因此在可三角化的定义中, 利用上三角矩阵和下三角矩阵是等价的. $\square$

定理 6.24. 设 $\dim V = n$, $T \in L(V)$. TFAE:

- (1) T 可三角化.
- (2) f_T 可以分解为一次式的乘积.
- (3) p_T 可以分解为一次式的乘积.
- (4) 存在 T 的不变子空间序列 $\{0\} = W_0 \subsetneq \cdots \subsetneq W_n = V$ 满足 $\dim W_i = i$.

推论 6.25. 如果 F 是代数闭域, 则任意 $A \in F^{n \times n}$ 可三角化.

定理 6.24 的证明. “(1) $\implies$ (2)”. 设 $[T]_{\mathcal{B}} = A$ 是上三角矩阵. 则 $f_T = \prod_{i=1}^n (x - A_{ii})$.

“(2) $\implies$ (3)”. 只需注意到 $p_T | f_T$.

“(3) $\implies$ (4)”. 用归纳法构造 W_k . 当 $k = 0$ 时, 只需取 $W_0 = \{0\}$. 假设 $1 \leq k \leq n$, 并且已构造出 T 的不变子空间序列 $\{0\} = W_0 \subsetneq \cdots \subsetneq W_{k-1} \subset V$ 满足 $\dim W_i = i$. 考虑 $T_{V/W_{k-1}}$. 由于 $p_{T_{V/W_{k-1}}} | p_T$, $p_{T_{V/W_{k-1}}}$ 可以分解为一次式的乘积. 特别地, $\sigma(T_{V/W_{k-1}}) \neq \emptyset$. 取 $c \in \sigma(T_{V/W_{k-1}})$ 和关于特征值 c 的特征向量 $\alpha + W_{k-1} \in V/W_{k-1}$, 这里 $\alpha \in V \setminus W_{k-1}$. 则

$$T\alpha + W_{k-1} = T_{V/W_{k-1}}(\alpha + W_{k-1}) = c(\alpha + W_{k-1}) = c\alpha + W_{k-1},$$

即 $T\alpha - c\alpha \in W_{k-1}$. 设 $W_k = W_{k-1} \oplus F\alpha$. 则 $T\alpha \in W_k$, 从而 W_k 是不变子空间. 这就完成了构造.

“(4) $\implies$ (1)”. 取 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 使得 $\{\alpha_1, \dots, \alpha_i\}$ 是 W_i 的基, $1 \leq i \leq n$. 设 $A = [T]_{\mathcal{B}}$. 则

$$T\alpha_j = \sum_{i=1}^n A_{ij}\alpha_i \in W_j = \text{span}\{\alpha_1, \dots, \alpha_j\}.$$

因此当 $i > j$ 时有 $A_{ij} = 0$, 即 A 上三角. □

注 6.6. • V 的一个子空间序列 $\{0\} = W_0 \subsetneq \cdots \subsetneq W_k = V$ 称为 V 的一个**旗(flag)**. 如果进一步还有 $\dim W_i = i$ (即 $k = \dim V$), 则称该子空间序列为 V 的一个**全旗(full flag)**或**完备旗(complete flag)**. 因此(4)就是说, 存在由 T 的不变子空间构成的全旗(即 T -不变的全旗).

- 如果不用分裂域或代数闭包等工具, “(3) $\implies$ (2)”并不显然.
- 推论6.25是较弱的结论. 一方面, $P^{-1}AP$ 可以更好(Jordan标准型). 另一方面, 如果只希望 $P^{-1}AP$ 上三角, 则 P 可以取得更特殊(例如复矩阵的Schur三角化定理).

下面的结论是可解李代数的Lie定理的特殊情况. 为了简单起见, 我们只考虑代数闭域的情况.

定理 6.26. 假设 F 是代数闭域, 并且集合 $\mathcal{F} \subset L(V)$ 中的线性变换两两可交换. 则 $\mathcal{F}$ 中的线性变换可同时三角化, 即存在 V 的有序基 $\mathcal{B}$, 使得所有 $T \in \mathcal{F}$ 的矩阵 $[T]_{\mathcal{B}}$ 都是上三角矩阵.

推论 6.27. 假设 F 是代数闭域, 并且 $\mathcal{M} \subset F^{n \times n}$ 中的矩阵两两可交换. 则 $\mathcal{M}$ 中的矩阵可同时三角化, 即存在可逆矩阵 $P \in \text{GL}_n(F)$, 使得对所有 $A \in \mathcal{M}$, 总有 $P^{-1}AP$ 是上三角矩阵.

先证明:

引理 6.28. 假设 F 是代数闭域, 并且 $\mathcal{F} \subset L(V)$ 中的线性变换两两可交换. 则 $\mathcal{F}$ 中的线性变换存在公共特征向量.

证明. 首先, 不妨假设 $\mathcal{F}$ 是有限集. 事实上, 对于一般的 $\mathcal{F}$, 取 $L(V)$ 的子空间 $\text{span}\mathcal{F}$ 的基 $\mathcal{F}_0$, 则 $\mathcal{F}_0$ 是有限集, 并且 $\mathcal{F}$ 中的线性变换两两可交换 $\iff \mathcal{F}_0$ 中的线性变换两两可交换. 假设引理对 $\mathcal{F}_0$ 成立, 即存在 $\alpha \in V \setminus \{0\}$ 使得对任意 $T \in \mathcal{F}_0$ 有 $T\alpha \in F\alpha$. 则对任意 $U \in \mathcal{F}$, 存在 $T_1, \dots, T_k \in \mathcal{F}_0$ 使得 $U = \sum_{i=1}^k c_i T_i$. 于是 $U\alpha = \sum_{i=1}^k c_i T_i \alpha \in F\alpha$. 因此 α 是 $\mathcal{F}$ 中的线性变换的公共特征向量.

假设 $\mathcal{F}$ 是有限集. 对 $|\mathcal{F}|$ 做归纳. 当 $|\mathcal{F}| = 0$ 时无需证明. 假设 $|\mathcal{F}| = k \geq 1$, 并且引理当 $|\mathcal{F}| = k - 1$ 时成立. 设 $\mathcal{F} = \{T_1, \dots, T_k\}$. 由归纳假设, $T_1, \dots, T_{k-1}$ 存在公共特征向量 α , 即存在 $c_1, \dots, c_{k-1}$ 使得 $(T_i - c_i I)\alpha = 0$, $1 \leq i \leq k-1$. 于是 $W := \bigcap_{i=1}^{k-1} \text{Ker}(T_i - c_i I) \neq \{0\}$. 由于 T_k 与 $T_1, \dots, T_{k-1}$ 可交换, 所以 W 是 T_k 的不变子空间. 由于 F 是代数闭域, $(T_k)_W$ 有特征值, 从而有特征向量 $\beta \in W \setminus \{0\}$. β 即为 $T_1, \dots, T_k$ 的公共特征向量. □

定理6.26的证明. 先证明存在子空间序列 $\{0\} = W_0 \subsetneq \cdots \subsetneq W_n = V$ 满足 $\dim W_i = i$, 并且每个 W_i 是所有 $T \in \mathcal{F}$ 的不变子空间. 用归纳法构造 W_k . 当 $k = 0$ 时, 只需取 $W_0 = \{0\}$. 假设 $1 \leq k \leq n$, 并且已构造出子空间序列 $\{0\} = W_0 \subsetneq \cdots \subsetneq W_{k-1} \subset V$, 满足当 $0 \leq i \leq k-1$ 时,

$\dim W_i = i$, 并且 W_i 是所有 $T \in \mathcal{F}$ 的不变子空间. 考虑 $L(V/W_{k-1}, V/W_{k-1})$ 的子集

$$\mathcal{F}_{k-1} := \{T_{V/W_{k-1}} \mid T \in \mathcal{F}\}.$$

它是可交换的. 对 $\mathcal{F}_{k-1}$ 用引理, 可知存在 $\alpha \in V \setminus W_{k-1}$ 使得 $\alpha + W_{k-1} \in V/W_{k-1}$ 是所有 $T_{V/W_{k-1}} (T \in \mathcal{F})$ 的公共特征向量. 所以对任意 $T \in \mathcal{F}$, 存在 $c \in F$ 使得

$$T\alpha + W_{k-1} = T_{V/W_{k-1}}(\alpha + W_{k-1}) = c(\alpha + W_{k-1}) = c\alpha + W_{k-1}.$$

设 $W_k = W_{k-1} \oplus F\alpha$. 则 $T\alpha \in W_k$, 从而 W_k 是 T 的不变子空间. 这就完成了满足要求的子空间序列的构造.

现在, 取 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 使得 $\{\alpha_1, \dots, \alpha_i\}$ 是 W_i 的基, $1 \leq i \leq n$. 则对任意 $T \in \mathcal{F}$, $[T]_{\mathcal{B}}$ 是上三角矩阵. $\square$

注 6.7. 从证明过程容易看出, 如果 F 不是代数闭域, 但是 $\mathcal{F}$ 中的线性变换的特征多项式都可以分解为一次式的乘积(这等价于 $\mathcal{F}$ 中的线性变换都是可三角化的), 则定理 6.26 的结论也成立.

§6.5 投影映射

给定 $T \in L(V)$. 假设 V 可以分解为 T 的不变子空间的直和 $V = \bigoplus_{i=1}^k W_i$. 记 $T_i = T_{W_i}$. 则对 $\alpha = \sum_{i=1}^k \alpha_i$, 有 $T\alpha = \sum_{i=1}^k T_i \alpha_i$. 我们称 T 为 $T_1, \dots, T_k$ 的直和. T 的性质被这些 T_i 的性质完全决定. 因此, 如果我们理解了每个 T_i , 我们就理解了 T . 例如:

- 设 $\mathcal{B}_i = \{\alpha_{i1}, \dots, \alpha_{id_i}\}$ 是 W_i 的有序基, 则 $\mathcal{B} = \{\alpha_{11}, \dots, \alpha_{1d_1}, \dots, \alpha_{k1}, \dots, \alpha_{kd_k}\}$ 是 V 的有序基, 并且

$$[T]_{\mathcal{B}} = \text{diag}([T_1]_{\mathcal{B}_1}, \dots, [T_k]_{\mathcal{B}_k}).$$

特别地, 如果我们可以取 $\mathcal{B}_i$ 使 $[T_i]_{\mathcal{B}_i}$ 简单, 我们就认为 $[T]_{\mathcal{B}}$ 简单.

- $\det(T) = \prod_{i=1}^k \det(T_i)$. 更一般地, $f_T = \prod_{i=1}^k f_{T_i}$.
- $p_T = \text{lcm}(p_{T_1}, \dots, p_{T_k})$.
- $\text{Ker}(T) = \bigoplus_{i=1}^k \text{Ker}(T_i)$, $\text{Im}(T) = \bigoplus_{i=1}^k \text{Im}(T_i)$.

我们后面的主要目的之一就是选择适当的直和分解, 使每个 T_i 具有比较基本的形式.

注 6.8. • 反过来, 如果给定 $T_i = L(W_i, W_i)$, 可以定义它们的直和 $\bigoplus_{i=1}^k T_i \in L(V)$ 为

$$\bigoplus_{i=1}^k T_i \left(\sum_{j=1}^k \alpha_j \right) = \sum_{i=1}^k T_i \alpha_i.$$

这也可以对外直和来定义.

- 不变直和分解就是 $F[x]$ -子模的直和分解.

§6.6 准素分解

我们准备对 T 做分解: 准素分解和循环分解. 加在一起, 得到准素循环分解. 其中循环分解推出有理标准形, 准素循环分解推出 Jordan 标准形.

设 V 是有限维 F -线性空间. 称 $T \in L(V)$ 是准素的(primary), 如果 p_T 是素多项式的幂. 此时, 我们也称 V 在 T 的作用下是准素的, 或称 V 作为 $F[x]$ -模是准素的.

注 6.9. • 以后将看到, p_T 是素多项式的幂等价于 f_T 是素多项式的幂.

- 若 p 素, 则 $F[x]$ 的理想 (p^r) 称为**准素理想**(也可以称 p^r 为准素多项式, 但这么叫的不多).

定理 6.29(准素分解). 给定 $T \in L(V)$. 将 p_T 分解为互不相同的首项系数是1的素多项式的幂的乘积 $p_T = \prod_{i=1}^k p_i^{r_i}$, 其中 $r_i \geq 1$. 则

$$V = \bigoplus_{i=1}^k \text{Ker}(p_i^{r_i}(T)),$$

其中 $W_i := \text{Ker}(p_i^{r_i}(T)) \neq \{0\}$, 并且 $p_{T_{W_i}} = p_i^{r_i}$.

证明. 记 $f_i = \prod_{j \neq i} p_j^{r_j}$. 注意到 $p_i^{r_i}(T)f_i(T) = p_T(T) = 0$ 但 $f_i(T) \neq 0$. 所以 $p_i^{r_i}(T)$ 不可逆. 因此 $W_i \neq \{0\}$.

W_i 无关: 设 $\alpha_i \in W_i$, $\sum_{j=1}^k \alpha_j = 0$. 两边取 f_i 得 $f_i \alpha_i = 0$. 注意到还有 $p_i^{r_i} \alpha_i = 0$ 并且 f_i 与 $p_i^{r_i}$ 互素. 所以存在 $a, b \in F[x]$ 使 $1 = af_i + bp_i^{r_i}$. 因此 $\alpha_i = af_i \alpha_i + bp_i^{r_i} \alpha_i = 0$.

$V = \sum_{i=1}^k W_i$: 注意到 $f_1, \dots, f_k$ 互素. 于是存在 $g_1, \dots, g_k$ 使得 $\sum_{i=1}^k g_i f_i = 1$. 设 $\alpha \in V$. 则 $f_i \alpha \in W_i$. 因此 $\alpha = \sum_{i=1}^k g_i f_i \alpha \in \sum_{i=1}^k W_i$.

$p_{T_{W_i}} = p_i^{r_i}$: 首先, 对任意 $\alpha_i \in W_i$ 有 $p_i^{r_i} \alpha_i = 0$. 因此 $p_{T_{W_i}}|p_i^{r_i}$. 另一方面, 对 $\alpha \in V$, 如果 $\alpha = \sum_{j=1}^k \alpha_j$, $\alpha_j \in W_j$, 则

$$\prod_{i=1}^k p_{T_{W_i}} \alpha = \sum_{j=1}^k \prod_{i=1}^k p_{T_{W_i}} \alpha_j = 0.$$

所以 $p_T | \prod_{i=1}^k p_{T_{W_i}}$. 这推出 $p_{T_{W_i}} = p_i^{r_i}$. □

注意到每个 W_i 是 T -不变子空间, 称为 V 的 p_i -**准素分量**(primary component).

第七章 有理标准形和Jordan标准形

§7.1 循环子空间和零化子

给定域 F , 记 $R = F[x]$. 设 V 是有限维 F -线性空间. 则 $T \in L(V)$ 诱导了 V 上的 R -模结构:
 $f\alpha = f(T)\alpha$. 对于 $\alpha \in V$, V 的子空间

$$R\alpha := \{f\alpha \mid f \in R\} = \text{span}\{\alpha, T\alpha, T^2\alpha, \dots\}$$

称为由 α 生成的循环子空间. 注意它是 V 的子模, 也称为由 α 生成的循环子模. 容易看出:

- $R\alpha$ 是包含 α 的最小的 T -不变子空间.
- $\dim R\alpha = 1 \iff \alpha$ 是特征向量.

如果 $V = R\alpha$, 则称 α 是循环向量. 如果 V 中存在循环向量, 则称 T 是循环的, 并称 V 是循环模.

例 7.1. 设 $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, $T = L_A$. 设 $\epsilon_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\epsilon_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. 注意到 $T\epsilon_1 = \epsilon_2$, $T\epsilon_2 = 0$. 因此
 $R\epsilon_1 = V$, $R\epsilon_2 = F\epsilon_2$.

因此 ϵ_1 是循环向量, ϵ_2 不是循环向量.

由这个例子可以看出, 当 T 循环时, 不一定所有非零向量都是循环向量. 为了鉴别一个向量是否是循环向量, 考虑函数

$$\Delta = \Delta_T : V \rightarrow \mathbb{Z}, \quad \Delta(\alpha) = \dim R\alpha.$$

容易看出:

- T 循环 $\iff \Delta$ 的最大值为 $\dim V$. 此时, $\alpha \in V$ 是循环向量 $\iff \Delta$ 在 α 处取得最大值.

对于 $\alpha \in V$, 为了考察 $\Delta(\alpha)$, 记

$$M(\alpha) := \{f \in R \mid f\alpha = 0\}.$$

引理 7.1. $M(\alpha)$ 是理想, 并且 $p_T \in M(\alpha)$.

证明. $p_T \in M(\alpha)$; $f, g \in M(\alpha), c \in F \implies f + cg \in M(\alpha)$; $f \in M(\alpha), g \in R \implies fg \in M(\alpha)$. $\square$

我们称 $M(\alpha)$ 为 α 的零化理想, 并称 $M(\alpha)$ 的唯一的monic生成元 p_α 为 α 的零化子(annihilator). 注意 $p_\alpha \mid p_T$, 并且 $\alpha = 0 \iff M(\alpha) = R \iff p_\alpha = 1$. p_α 与 $\Delta(\alpha)$ 的联系如下:

引理 7.2. 记 $d = \deg p_\alpha$, 则 $\{\alpha, T\alpha, \dots, T^{d-1}\alpha\}$ 是 $R\alpha$ 的基. 特别地, $\Delta(\alpha) = \deg p_\alpha$.

证明. 线性无关: 设 $\sum_{i=0}^{d-1} c_i T^i \alpha = 0$. 记 $g = \sum_{i=0}^{d-1} c_i x^i$. 则 $g\alpha = 0$. 因此 $p_\alpha \mid g$. 由于 $\deg g < \deg p_\alpha$, 只能有 $g = 0$, 即 $c_i = 0$.

$\text{span} = R\alpha$: “ $\subset$ ”显然. “ $\supset$ ”: 对任意 $\beta \in R\alpha$, 存在 $f \in R$ 使得 $\beta = f\alpha$. 设 $f = qp_\alpha + r$, 这里 $q, r \in R$, $\deg r < d$. 则 $\beta = f\alpha = r\alpha \in \text{span}$. $\square$

由这个引理, 考察 $\Delta(\alpha)$ 的最大值转化为考察 $\deg p_\alpha$ 的最大值. 由于 $p_\alpha \mid p_T$, 总有 $\Delta(\alpha) \leq \deg p_T$. 下面说明等号是可以成立的.

命题 7.3. 存在 $\alpha \in V$ 满足 $p_\alpha = p_T$.

证明. 设 $p_T = \prod_{i=1}^k p_i^{r_i}$, 其中 p_i 为互不相同的首一素多项式, $r_i \geq 1$. 记 $W_i := \text{Ker}(p_i^{r_i}(T))$. 则由准素分解, 有 $V = \bigoplus_{i=1}^k W_i$, 并且 $p_{TW_i} = p_i^{r_i}$. 注意到 $\text{Ker}(p_i^{r_i-1}(T_{W_i}))$ 为 W_i 的真子空间. 取 $\alpha_i \in$

$W_i \setminus \text{Ker}(p_i^{r_i-1}(T_{W_i}))$. 则 $p_{\alpha_i} | p_i^{r_i}$ 且 $p_{\alpha_i} \nmid p_i^{r_i-1}$, 所以 $p_{\alpha_i} = p_i^{r_i}$. 取 $\alpha = \sum_{i=1}^k \alpha_i$. 则由 $p_\alpha \alpha = 0$ 推出 $p_\alpha \alpha_i = 0$. 从而 $p_i^{r_i} | p_\alpha$. 这推出 $p_T | p_\alpha$. 只能有 $p_\alpha = p_T$. $\square$

推论 7.4. 函数 Δ 的最大值为 $\deg p_T$, 并且 $\Delta(\alpha) = \deg p_T \iff p_\alpha = p_T$.

证明. 由于总有 $p_\alpha | p_T$, 所以 $\Delta(\alpha) = \deg p_\alpha \leq \deg p_T$. 上面命题说明存在 α 使等号成立. 因此 Δ 的最大值为 $\deg p_T$. 进一步地, $\Delta(\alpha) = \deg p_T \iff \deg p_\alpha = \deg p_T \iff p_\alpha = p_T$. $\square$

由此, 容易得到:

命题 7.5. T 循环 $\iff \deg p_T = \dim V \iff p_T = f_T$. 此时, $\alpha \in V$ 是循环向量 $\iff p_\alpha = p_T$.

证明.

$$T \text{ 循环} \iff \Delta \text{ 的最大值等于 } \dim V \iff \deg p_T = \dim V \iff p_T = f_T.$$

此时, $\alpha \in V$ 是循环向量 $\iff \Delta(\alpha) = \deg p_T \iff p_\alpha = p_T$. $\square$

现在我们讨论当 T 循环时它的矩阵的形式. 对于首项系数是1的 n 次多项式

$$f = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

定义它的相伴矩阵或友阵(companion matrix) $C_f \in F^{n \times n}$ 为

$$C_f = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix}.$$

下面说明 T 循环 $\iff$ 它在某个基下的矩阵是某多项式的相伴矩阵.

命题 7.6. 设 $\dim V = n$, $T \in L(V)$.

- (1) 如果 T 循环, α 是循环向量, 则 T 在有序基 $\{\alpha, T\alpha, \dots, T^{n-1}\alpha\}$ 下的矩阵为 $C_{p_\alpha} = C_{p_T}$.
- (2) 如果 T 在某有序基下的矩阵为 C_f , 其中 f 首项系数是1并且次数是 n , 则 T 循环并且 $p_T = f$.

证明. 注意到对于 $f = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ 和 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, $[T]_{\mathcal{B}} = C_f$ 意味着

$$T\alpha_i = \alpha_{i+1}, \quad 1 \leq i \leq n-1, \quad (7.1)$$

$$T\alpha_n = -a_0\alpha_1 - a_1\alpha_2 - \cdots - a_{n-1}\alpha_n. \quad (7.2)$$

(1) 记 $\alpha_i = T^{i-1}\alpha$, $1 \leq i \leq n$. 则(7.1)和(7.2)对 $f = p_\alpha$ 成立. 这说明 T 的矩阵为 C_{p_α} .

(2) 设对于上面的 f 和 $\mathcal{B}$ 有 $[T]_{\mathcal{B}} = C_f$, 则(7.1)和(7.2)成立. (7.1)说明 α_1 是循环向量, 从而 T 循环并且 $p_{\alpha_1} = p_T$. (7.2)说明 $f\alpha_1 = 0$, 即 $p_{\alpha_1} | f$. 两者结合, 得 $p_T | f$. 但 $\deg p_T = \deg f = n$. 所以只能有 $p_T = f$. $\square$

推论 7.7. 设 f 是首项系数是1的多项式. 则 $f_{C_f} = p_{C_f} = f$.

证明. 对 T_{C_f} 和 $F^{n \times 1}$ 的标准基应用(2). $\square$

注 7.1. $f_{C_f} = f$ 也可以直接计算得到. 这一计算实际上给出了 Cayley-Hamilton 定理的另一证明: 为证明 $f_T(T) = 0$, 只需证对任意 $\alpha \in V$ 有 $f_T\alpha = 0$. 考虑循环子空间 $R\alpha$. 这一计算给出 $f_{T_{R\alpha}} = f_{C_{p_\alpha}} = p_\alpha$, 从而 $p_\alpha | f_T$, 即 $f_T\alpha = 0$.

§7.2 循环分解和有理标准形

定理 7.8(循环分解). 设 V 是有限维 F -线性空间, 设 $T \in L(V)$. 则存在 $\alpha_1, \dots, \alpha_r \in V \setminus \{0\}$ 满足:

- $V = \bigoplus_{i=1}^r R\alpha_i$;
- 设 $p_i = p_{\alpha_i}$. 则 $p_r | p_{r-1} | \dots | p_1$.

整数 r 和序列 $p_1, \dots, p_r$ 被 T 唯一决定, 并且

$$p_T = p_1, \quad f_T = \prod_{i=1}^r p_i.$$

定理中的 $p_1, \dots, p_r$ 称为 T 的不变因子(invariant factors). 对 $A \in F^{n \times n}$, L_A 的不变因子也称为 A 的不变因子.

注 7.2. • p_i 也等于 $T_{R\alpha_i}$ 的特征多项式和最小多项式.

- 有的文献中用 $p_1 | p_2 | \dots | p_r$.

引理 7.9. 设 $\alpha \in V$ 满足 $p_\alpha = p_T$. 则对任意 $L \in V/R\alpha$, 存在 $\beta \in L$ 满足 $p_\beta = p_L$.

证明. 显然对任意 $\beta \in L$ 总有 $p_\beta L = 0$, 即 $p_L | p_\beta$. 只需证存在 $\beta \in L$ 满足 $p_\beta | p_L$, 即 $p_L \beta = 0$. 任取 $\beta_0 \in L$. 则 $p_L \beta_0 \in R\alpha$. 设 $p_L \beta_0 = f\alpha$, $p_T = gp_L$. 则 $gf\alpha = gp_L \beta_0 = 0$. 于是 $gp_L = p_T = p_\alpha | gf$, 因此 $p_L | f$. 设 $f = p_L h$, $\beta = \beta_0 - h\alpha \in L$. 则 $p_L \beta = 0$. $\square$

定理7.8的证明. 对 $\dim V$ 归纳证明分解的存在性. $\dim V = 1$ 时显然. 一般情况, 取 $\alpha_1 \in V$ 使 $p_{\alpha_1} = p_T$. 由于 $\dim V/R\alpha_1 < \dim V$, 由归纳假设, 存在 $L_2, \dots, L_r \in (V/R\alpha_1) \setminus \{0\}$ 使得 $V/R\alpha_1 = \bigoplus_{i=2}^r RL_i$, 并且 $p_{L_r} | \dots | p_{L_2}$. 由引理7.9, 对 $2 \leq i \leq r$, 存在 $\alpha_i \in L_i$ 满足 $p_{\alpha_i} = p_{L_i}$. 显然 $p_{\alpha_r} | \dots | p_{\alpha_1}$. 我们验证 $V = \bigoplus_{i=1}^r R\alpha_i$. 设 $\sum_{i=1}^r g_i \alpha_i = 0$. 投到 $V/R\alpha_1$ 上, 有 $\sum_{i=2}^r g_i L_i = 0$. 因此 $g_i L_i = 0$. 这推出 $p_{\alpha_i} = p_{L_i} | g_i$, 从而 $g_i \alpha_i = 0$. 进而 $g_1 \alpha_1 = 0$. 因此 $R\alpha_1, \dots, R\alpha_r$ 无关. 另一方面, 对任意 $\gamma \in V$, 把 $\gamma + R\alpha_1 \in V/R\alpha_1$ 分解为 $\gamma + R\alpha_1 = \sum_{i=2}^r g_i L_i$, 则 $\gamma - \sum_{i=2}^r g_i \alpha_i \in R\alpha_1$. 因此 $V = \bigoplus_{i=1}^r R\alpha_i$.

下面证明唯一性和关于 p_T, f_T 的部分. 设 $\alpha_1, \dots, \alpha_r \in V \setminus \{0\}$ 满足 $V = \bigoplus_{i=1}^r R\alpha_i$, 并且 $p_i := p_{\alpha_i}$ 满足 $p_r | p_{r-1} | \dots | p_1$. 则有

$$p_T = \text{lcm}(p_{T_{R\alpha_1}}, \dots, p_{T_{R\alpha_r}}) = \text{lcm}(p_1, \dots, p_r) = p_1,$$

$$f_T = \prod_{i=1}^r f_{T_{R\alpha_i}} = \prod_{i=1}^r p_i.$$

如果还有 $\beta_1, \dots, \beta_s \in V \setminus \{0\}$ 满足 $V = \bigoplus_{i=1}^s R\beta_i$, 并且 $q_i := p_{\beta_i}$ 满足 $q_s | q_{s-1} | \dots | q_1$, 则类似地有 $p_T = q_1, f_T = \prod_{i=1}^s q_i$. 因此 $p_1 = q_1, \prod_{i=1}^r p_i = \prod_{i=1}^s q_i$. 我们需要证明 $r = s$ 并且 $p_i = q_i$. 如果这不成立, 则存在 $2 \leq t \leq \min\{r, s\}$ 满足 $p_t \neq q_t$, 并且当 $i < t$ 时有 $p_i = q_i$. 不妨设 $q_t \nmid p_t$. 在 $\bigoplus_{i=1}^r R\alpha_i = V = \bigoplus_{i=1}^s R\beta_i$ 两边乘 p_t , 得

$$\bigoplus_{i=1}^{t-1} Rp_t \alpha_i = p_t V = \bigoplus_{i=1}^{t-1} Rp_t \beta_i \oplus \bigoplus_{i=t}^s Rp_t \beta_i.$$

容易看出:

- f, g monic, $p_\alpha = fg \implies p_{f\alpha} = g$: 事实上, $p_{f\alpha} | h \iff h(f\alpha) = 0 \iff p_\alpha = fg | fh \iff g | h$.

因此, 当 $i < t$ 时有

$$\dim Rp_t \alpha_i = \deg p_{p_t \alpha_i} = \deg(p_i/p_t) = \deg(q_i/p_t) = \deg p_{p_t \beta_i} = \dim Rp_t \beta_i.$$

这推出 $\bigoplus_{i=t}^s Rp_t \beta_i = \{0\}$. 特别地, $p_t \beta_t = 0$. 因此 $q_t | p_t$. 矛盾. $\square$

推论 7.10. p_T 与 f_T 有相同的素因子. □

在定理7.8的结论下, $\mathcal{B}_i := \{\alpha_i, T\alpha_i, \dots, T^{d_i-1}\alpha_i\}$ 是 $R\alpha_i$ 的有序基, 这里 $d_i = \deg p_i = \dim R\alpha_i$. 于是, $[T_{R\alpha_i}]_{\mathcal{B}_i} = C_{p_i}$. 因此, 在 V 的有序基 $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$ 下,

$$[T]_{\mathcal{B}} = \text{diag}(C_{p_1}, \dots, C_{p_r}).$$

定义 7.1. 矩阵 $A \in F^{n \times n}$ 称为**有理形矩阵**, 如果存在非常数的首一多项式 $p_1, \dots, p_r \in F[x]$, 满足 $\sum_{i=1}^r \deg p_i = n$ 并且 $p_r | \dots | p_1$, 使得

$$A = \text{diag}(C_{p_1}, \dots, C_{p_r}).$$

定理 7.11. 存在 V 的有序基 $\mathcal{B}$ 使 $[T]_{\mathcal{B}}$ 是有理形矩阵, 并且这样的有理形矩阵是唯一的.

证明. 只需证明唯一性. 设 T 在某个有序基 $\mathcal{B}'$ 下的矩阵是有理形矩阵 $\text{diag}(C_{q_1}, \dots, C_{q_s})$, 其中 $q_s | \dots | q_1$. 设 $\mathcal{B}' = (\mathcal{B}'_1, \dots, \mathcal{B}'_s)$, $|\mathcal{B}'_i| = \deg q_i$. 则 $\mathcal{B}'_i$ 中的第一个向量 β_i 满足 $R\beta_i = \text{span } \mathcal{B}'_i$ 并且 $p_{\beta_i} = q_i$. 由定理7.8的唯一性部分, $q_1, \dots, q_s$ 只能是 T 的不变因子序列, 因此 C_{q_i} 被 T 决定. □

推论 7.12. 任意矩阵 $A \in F^{n \times n}$ 相似于唯一的有理形矩阵, 称为 A 的**有理标准形**(rational canonical form或rational normal form).

证明. 取 $T = L_A$. □

叫有理标准形的原因: 若 $A \in \mathbb{Q}^{n \times n}$ 则有理标准形也 $\in \mathbb{Q}^{n \times n}$.

推论 7.13. 两个矩阵相似 $\iff$ 有相同的有理标准形 $\iff$ 有相同的不变因子序列. □

下面说明矩阵的有理标准形和不变因子不依赖于域的选取.

推论 7.14. 设 $A \in F^{n \times n}$ 的有理标准形为 A' , 不变因子序列为 $p_1, \dots, p_r$. 如果对某个子域 $K \subset F$ 有 $A \in K^{n \times n}$, 则 $A' \in K^{n \times n}$ 并且 $A \stackrel{K}{\sim} A'$, $p_1, \dots, p_r \in K[x]$. 特别地, $p_A \in K[x]$.

证明. 设 A 在 $K^{n \times n}$ 中的有理标准形为 $A'' \in K^{n \times n}$. A' 和 A'' 都是 $F^{n \times n}$ 中的有理形矩阵并且都与 A 相似, 由推论7.12的唯一性, $A' = A''$. 因此 $A' \in K^{n \times n}$ 并且 $A \stackrel{K}{\sim} A'$. 由 $A' \in K^{n \times n}$ 即得到 $p_1, \dots, p_r \in K[x]$. □

下面看有理标准形和不变因子的两个应用.

命题 7.15. 设 $A, B \in F^{n \times n}$ 相似. 如果对子域 $K \subset F$ 有 $A, B \in K^{n \times n}$, 则 A, B 在 K 上也相似.

证明. A, B 相似推出它们在 $F^{n \times n}$ 中的有理标准形相同, 设为 C . 由 $A, B \in K^{n \times n}$ 推出 $C \in K^{n \times n}$ 并且 $A \stackrel{K}{\sim} C, B \stackrel{K}{\sim} C$. 因此 $A \stackrel{K}{\sim} B$. □

命题 7.16. 设 $A \in F^{n \times n}$. 则 A 与 A^t 相似.

证明. 先假设 A 是某个首项系数是1的多项式 f 的相伴矩阵. 此时有 $f_A = p_A = f$. 因此 A 只有一个不变因子 f . 而

$$f_{A^t} = \det(xI - A^t) = \det((xI - A)^t) = \det(xI - A) = f_A = f,$$

$$g(A^t) = g(A)^t \implies "g(A) = 0 \iff g(A^t) = 0" \implies p_{A^t} = p_A = f.$$

因此 A^t 也只有一个不变因子 f . 因此 A^t 与 A 相似.

一般情况, 设 A 的有理标准形为 $\text{diag}(C_{p_1}, \dots, C_{p_r})$. 则 $A^t \sim \text{diag}(C_{p_1}^t, \dots, C_{p_r}^t)$. 我们已经证明 $C_{p_i}^t \sim C_{p_i}$. 所以 $\text{diag}(C_{p_1}^t, \dots, C_{p_r}^t) \sim \text{diag}(C_{p_1}, \dots, C_{p_r})$. 因此 $A^t \sim A$. □

例 7.2. 2×2 的有理形矩阵只能是 cI_2 或 $\begin{bmatrix} 0 & * \\ 1 & * \end{bmatrix}$.

给定 $A \in F^{n \times n}$, 如果知道了 A 的不变因子为 $p_1, \dots, p_r$, 则 A 的有理标准形为 $\text{diag}(C_{p_1}, \dots, C_{p_r})$. 如何求可逆矩阵 P 使得 $P^{-1}AP = \text{diag}(C_{p_1}, \dots, C_{p_r})$?

命题 7.17. 设 $A \in F^{n \times n}$, $\mathcal{B}$ 是 $F^{n \times 1}$ 的有序基, P 是由 $\mathcal{B}$ 的列向量按次序排列得到的 $n \times n$ 可逆矩阵. 则 $P^{-1}AP = [L_A]_{\mathcal{B}}$.

证明. 设 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$. 则 $P = [\alpha_1, \dots, \alpha_n]$. 记 $B = [L_A]_{\mathcal{B}}$. 则

$$A\alpha_j = \sum_{i=1}^n B_{ij}\alpha_i.$$

这意味着

$$AP = [A\alpha_1, \dots, A\alpha_n] = \left[\sum_{i=1}^n B_{i1}\alpha_i, \dots, \sum_{i=1}^n B_{in}\alpha_i \right] = PB,$$

即 $P^{-1}AP = B$. □

因此, 如果 $\alpha_1, \dots, \alpha_r \in F^{n \times 1} \setminus \{0\}$ 满足 $F^{n \times 1} = \bigoplus_{i=1}^r R\alpha_i$ 并且 $p_{\alpha_i} = p_i$, 则可逆矩阵

$$P = [\alpha_1, A\alpha_1, \dots, A^{d_1-1}\alpha_1, \dots, \alpha_r, A\alpha_r, \dots, A^{d_r-1}\alpha_r], \quad d_i = \deg p_i = \dim R\alpha_i,$$

满足

$$P^{-1}AP = \text{diag}(C_{p_1}, \dots, C_{p_r}).$$

例 7.3. 设 $A = \begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix} \in \mathbb{R}^{3 \times 3}$. 习题课中求过(见P188例3):

$$f_A = (x-1)(x-2)^2.$$

容易验证 $(A-I)(A-2I) = 0$. 因此

$$p_A = (x-1)(x-2).$$

所以不变因子为

$$p_1 = p_A = (x-1)(x-2) = x^2 - 3x + 2,$$

$$p_2 = f_A/p_A = x-2.$$

因此 A 的有理标准形为

$$\text{diag}(C_{p_1}, C_{p_2}) = \text{diag}\left(\begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}, 2\right) = \begin{bmatrix} 0 & -2 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

为了寻找 P 使 $P^{-1}AP = B$, 由上面的分析, 只需寻找 $\alpha_1, \alpha_2 \in F^{3 \times 1} \setminus \{0\}$ 满足 $F^{3 \times 1} = R\alpha_1 \oplus R\alpha_2$ 并且 $p_{\alpha_i} = p_i$. 此时可逆矩阵 $P = [\alpha_1, A\alpha_1, \alpha_2]$ 就满足 $P^{-1}AP = \text{diag}(C_{p_1}, C_{p_2})$. 注意到:

- $p_{\alpha_1} = p_1 \iff \alpha_1$ 不是特征向量.
- $p_{\alpha_2} = p_2 \iff \alpha_2$ 是特征值 2 的特征向量.

此时,

- $F^{3 \times 1} = R\alpha_1 \oplus R\alpha_2 \iff \alpha_2 \notin R\alpha_1$.

尝试 $\alpha_1 = \epsilon_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. 由于 $A\alpha_1 = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$, 所以 α_1 不是特征向量. 因此可选这个 α_1 . 习题课求

过(见P188):

$$V_2 = \{(x_1, x_2, x_3)^t \mid x_1 = 2x_2 + 2x_3\}.$$

我们需要找 $\alpha_2 \in V_2$ 使得 $\alpha_2 \notin R\alpha_1$. 注意到 $R\alpha_1 = \text{span}\{\alpha_1, A\alpha_1\}$, 其中的向量总有 x_2, x_3 同时

为0或同时不为0. 取 $\alpha_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$. 则 $\alpha_2 \in V_2 \setminus R\alpha_1$. 此时, $P = \begin{bmatrix} 1 & 5 & 2 \\ 0 & -1 & 1 \\ 0 & 3 & 0 \end{bmatrix}$.

事实上, 因为 A 可对角化, 有 $F^{3 \times 1} = V_1 \oplus V_2$. 取 V_1 的基 $\{\beta_1\}$ 和 V_2 的基 $\{\gamma_1, \gamma_2\}$. 则取 $\alpha_1 = \beta_1 + \gamma_1$, $\alpha_2 = \gamma_2$ 即可. $\square$

命题 7.18. 设 T 可对角化, $\sigma(T) = \{c_1, \dots, c_k\}$.

(1) 设 $\alpha = \sum_{i=1}^k \beta_i$, 其中 $\beta_i \in V_{c_i}$, 则 $R\alpha = \text{span}\{\beta_1, \dots, \beta_k\}$, $p_\alpha = \prod_{\beta_i \neq 0} (x - c_i)$.

(2) 设 $d_i = \dim V_{c_i}$, 则不变因子为 $p_j := \prod_{d_i \geq j} (x - c_i)$, 其个数为 $r := \max d_i$.

证明. (1) 注意到 $V = \bigoplus_{i=1}^k V_{c_i}$. 由于 $R\alpha = \{f\alpha \mid f \in F[x]\}$, 而

$$f\alpha = \sum_{i=1}^k f\beta_i = \sum_{i=1}^k f(c_i)\beta_i,$$

所以 $R\alpha \subset \text{span}\{\beta_1, \dots, \beta_k\}$. 另一方面, 对任意 $\sum_{i=1}^k t_i \beta_i \in \text{span}\{\beta_1, \dots, \beta_k\}$, 存在 f 满足 $f(c_i) = t_i$, 此时 $\sum_{i=1}^k t_i \beta_i = f\alpha \in R\alpha$. 因此 $R\alpha \supset \text{span}\{\beta_1, \dots, \beta_k\}$. 我们还有

$$f\alpha = 0 \iff f(c_i)\beta_i = 0 \iff \beta_i \neq 0 \text{ 时有 } f(c_i) = 0 \iff \prod_{\beta_i \neq 0} (x - c_i) \mid f.$$

因此 $p_\alpha = \prod_{\beta_i \neq 0} (x - c_i)$.

(2) 取 V_{c_i} 的有序基 $\mathcal{B}_i = \{\beta_{i1}, \dots, \beta_{id_i}\}$. 对于 $1 \leq j \leq r$, 考虑 $\alpha_j = \sum_{d_i \geq j} \beta_{ij}$. 由(1),

$$R\alpha_j = \text{span}\{\beta_{ij} \mid d_i \geq j\}, \quad p_{\alpha_j} = p_j.$$

只需注意到 $V = \bigoplus_{j=1}^r R\alpha_j$ 并且 $p_r \mid p_{r-1} \mid \dots \mid p_1$. $\square$

§7.3 准素循环分解和Jordan标准形

称 T 为不可分解的 (indecomposable), 如果 V 不能分解成两个非零不变子空间的直和.

命题 7.19. 任何 T 总是有限个不可分解变换的直和.

证明. 对 $\dim V$ 归纳. $\dim V = 1$ 时显然. 一般情况, 如果 T 不可分解, 则无需证明. 否则, V 是两个非零不变子空间的直和, 于是可以对每个直和项用归纳假设. $\square$

定理 7.20. 设 V 是有限维 F -线性空间, $T \in L(V, V)$. 则 T 不可分解 $\iff T$ 是准素循环的.

证明. “ $\implies$ ”. 设 T 不可分解. 则 T 的准素分解和循环分解中都只有一项. 因此 T 是准素循环的.

“ $\impliedby$ ”. 设 T 是准素循环的. 则 $f_T = p_T = p^r$, 其中 p 素. 假设有不变直和分解 $V = V_1 \oplus V_2$. 则 $f_{T_{V_1}} f_{T_{V_2}} = f_T = p^r$, 从而 $f_i = p^{r_i}$, $r_i \geq 0$, $r_1 + r_2 = r$. 另一方面, 设 $p_i = p^{s_i}$, $s_i \leq r_i$. 则 $p^r = p_T = \text{lcm}(p_1, p_2) = p^{\max\{s_1, s_2\}}$. 从而 $r_1 + r_2 = r = \max\{s_1, s_2\} \leq \max\{r_1, r_2\}$. 只能有 $r_1 = 0$ 或 $r_2 = 0$. 因此 $V_1 = \{0\}$ 或 $V_2 = \{0\}$. $\square$

因此, 我们得到:

定理 7.21(准素循环分解). 设 V 是有限维 F -线性空间, 设 $T \in L(V)$. 则 V 可以分解为一些准素循

环的非零不变子空间的直和 $V = \bigoplus_{i=1}^s V_i$. 设 $q_i = p_{T_i}$. 准素多项式序列 $q_1, \dots, q_s$ 在不计排列次序的意义下被 T 唯一决定, 称为 T 的**基本因子**或**初等因子**(elementary divisors).

证明. 存在性已证明. 只需证唯一性. 设 $q_1, \dots, q_s$ 的所有素因子为 $p_1, \dots, p_k$. 也就是说, $p_1, \dots, p_k$ 是互不相同的素多项式, 每个 p_i 是某个 q_j 的素因子, 并且每个 q_j 都是某个 p_i 的幂. 将序列 $(V_1, q_1), \dots, (V_s, q_s)$ 重新排列为下表:

$$\begin{array}{ccc} (V_{11}, p_1^{r_{11}}), & \dots, & (V_{1,d_1}, p_1^{r_{1,d_1}}), \\ \vdots & & \vdots \\ (V_{k1}, p_k^{r_{k1}}), & \dots, & (V_{k,d_k}, p_k^{r_{k,d_k}}), \end{array}$$

满足对任意 $1 \leq i \leq k$ 有 $r_{i1} \geq \dots \geq r_{i,d_i}$. 设 $d = \max\{d_1, \dots, d_k\}$. 对 $1 \leq j \leq d$, 设 $W_j = \bigoplus_{i:d_i \geq j} V_{ij}$. 则:

- $V = \bigoplus_{j=1}^d W_j$.
- W_j 是循环的:

$$\begin{aligned} f_{TW_j} &= \prod_{i:d_i \geq j} f_{TV_{ij}} = \prod_{i:d_i \geq j} p_{TV_{ij}} = \prod_{i:d_i \geq j} p_i^{r_{ij}}, \\ p_{TW_j} &= \text{lcm}\{p_{TV_{ij}} \mid i : d_i \geq j\} = \text{lcm}\{p_i^{r_{ij}} \mid i : d_i \geq j\} = \prod_{i:d_i \geq j} p_i^{r_{ij}}. \end{aligned}$$

因此 $f_{TW_j} = p_{TW_j}$.

- p_{TW_j} 的公式也说明 $p_{TW_d} \mid \dots \mid p_{TW_1}$.

因此 $p_{TW_1}, \dots, p_{TW_d}$ 就是 T 的不变因子序列. 它是被 T 唯一决定的. 因此上表中的第 j 列是 p_{TW_j} 的所有准素因子, 也是被 T 决定的. $\square$

我们可以从四个角度理解准素循环分解:

- 直接把 V 分解为不可分解的不变子空间的直和.
- 先对 V 做准素分解, 再对每个直和项做循环分解, 此时每个小块都是准素循环的. 注意准素的不变子空间还是准素的.
- 先对 V 做循环分解, 再对每个直和项做准素分解, 此时每个小块都是准素循环的. 注意循环子空间的不变子空间还是循环子空间. 证明如下: 设 $W \subset R\alpha$ 是不变子空间. 考虑 $\beta := p_{\alpha+W}\alpha \in W$. 我们验证 $W = R\beta$. 显然 $R\beta \subset W$. 另一方面, 对任意 $\gamma \in W$, 存在 $f \in R$ 满足 $\gamma = f\alpha$. $\gamma \in W$ 推出 $p_{\alpha+W} \mid f$. 设 $f = gp_{\alpha+W}$. 则 $\gamma = gp_{\alpha+W}\alpha = g\beta \in R\beta$. 因此 $W \subset R\beta$.
- 同时对 V 做准素分解 $V = \bigoplus_i V_i$ 和循环分解 $V = \bigoplus_j W_j$. 则 $V = \bigoplus_{i,j} V_i \cap W_j$ 是准素循环分解. (容易验证 V 确实等于 $\sum_{i,j} V_i \cap W_j$.)

设 $V = \bigoplus_{i=1}^s V_i$ 是准素循环分解, 相应的基本因子序列为 $q_1, \dots, q_s$. 则在适当有序基下, T_{V_i} 的矩阵为 C_{q_i} , 从而 T 的矩阵为 $\text{diag}(C_{q_1}, \dots, C_{q_s})$. 这样得到的“准素有理标准形”与域 F 有关, 因此比有理标准形应用少. 此时更常用的是取另一种有序基得到的 Jordan 标准形. 为此, 我们引入下面的概念.

定义 7.2. 设 $T \in L(V)$. 如果存在正整数 r 满足 $T^r = 0$, 则称 T 是**幂零**的.

命题 7.22. 设 $\dim V = n$, $T \in L(V)$. TFAE:

- (1) T 幂零.
- (2) p_T 是 x 的幂.

(3) $f_T = x^n$.

(4) $T^n = 0$.

证明. “(1) $\implies$ (2)”. 设 $T^r = 0$. 则 $p_T | x^r$. 因此 p_T 是 x 的幂.

“(2) $\implies$ (3)”. 由推论 7.10, f_T 的素因子只有 x . 但 $\dim f_T = n$. 所以只能有 $f_T = x^n$.

“(3) $\implies$ (4)”. $T^n = f_T(T) = 0$.

“(4) $\implies$ (1)”. 显然. □

现在假设 T 准素循环, 并且 $f_T = p_T = (x - c)^n$. 我们的做法是考虑幂零变换 $N = T - cI$ 的有理标准形. 我们称形如

$$J_n(c) := cI_n + C_{x^n} = \begin{bmatrix} c & 0 & 0 & \cdots & 0 & 0 \\ 1 & c & 0 & \cdots & 0 & 0 \\ 0 & 1 & c & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c & 0 \\ 0 & 0 & 0 & \cdots & 1 & c \end{bmatrix}$$

的矩阵为**基本Jordan矩阵**或特征值为 c 的 n 阶**Jordan块**(Jordan block).

命题 7.23. 设 $\dim V = n$, $T \in L(V)$, $c \in F$.

(1) 如果 T 准素循环, 并且 $f_T = p_T = (x - c)^n$, 则 T 在某有序基下的矩阵为 $J_n(c)$.

(2) 如果 T 在某有序基下的矩阵为 $J_n(c)$, 则 T 准素循环, 并且 $f_T = p_T = (x - c)^n$.

特别地, 如果 F 是代数闭域, 则 T 准素循环 $\iff$ 它在某个基下的矩阵是 Jordan 块.

证明. 设 $N = T - cI$. 则对任意有序基 $\mathcal{B}$ 有

$$[T]_{\mathcal{B}} = [cI + N]_{\mathcal{B}} = c[I]_{\mathcal{B}} + [N]_{\mathcal{B}} = cI_n + [N]_{\mathcal{B}}.$$

因此,

$$[T]_{\mathcal{B}} = J_n(c) \iff [N]_{\mathcal{B}} = J_n(0) = C_{x^n}.$$

此外,

$$f_T = p_T = (x - c)^n \iff f_N = p_N = x^n.$$

(1) 若 T 准素循环并且 $f_T = p_T = (x - c)^n$, 则 $f_N = p_N = x^n$. 因此 N 循环并且存在 $\mathcal{B}$ 使得 $[N]_{\mathcal{B}} = C_{x^n}$. 从而 $[T]_{\mathcal{B}} = J_n(c)$.

(2) 如果存在 $\mathcal{B}$ 使 $[T]_{\mathcal{B}} = J_n(c)$, 则 $[N]_{\mathcal{B}} = C_{x^n}$, 从而 $f_N = p_N = x^n$, 因此 $f_T = p_T = (x - c)^n$. 这推出 T 准素循环. □

对于一般情况, 假设 f_T 可以分解为一次式的乘积. 取准素循环分解 $V = \bigoplus_{i=1}^s V_i$. 则基本因子形如 $q_i = p_{T|_{V_i}} = (x - c_i)^{n_i}$. 于是 $T|_{V_i}$ 在 V_i 的某有序基下的矩阵为 $J_{n_i}(c_i)$. 因此, T 在 V 的某有序基下的矩阵为

$$\text{diag}(J_{n_1}(c_1), \dots, J_{n_m}(c_m)).$$

这样的矩阵称为**Jordan形矩阵**.

定理 7.24. 假设 $T \in L(V)$ 的特征多项式可以分解为一次式的乘积. 则存在 V 的有序基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}}$ 是 Jordan 形矩阵, 并且 Jordan 块在相差次序的意义下是唯一的, 称为 T 的**Jordan标准形**(Jordan canonical form或Jordan normal form). 特别地, 如果 F 是代数闭域, 则任意 T 存在 Jordan 标准形.

证明. 只需证明唯一性. 设 T 在某个有序基 $\mathcal{B}$ 下的矩阵是Jordan形矩阵 $\text{diag}(J_{n_1}(c_1), \dots, J_{n_m}(c_m))$. 设 $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_s)$, $|\mathcal{B}_i| = n_i$. 则 $V_i := \text{span} \mathcal{B}_i$ 是非零不变子空间, 并且 $[T_{V_i}]_{\mathcal{B}_i} = J_{n_i}(c_i)$. 因此 T_{V_i} 准素循环, 并且 $p_{T_{V_i}} = (x - c_i)^{n_i}$. 由定理7.21的唯一性部分, $(x - c_1)^{n_1}, \dots, (x - c_m)^{n_m}$ 只能是 T 的基本因子序列, 它们在相差次序的意义下被 T 决定. 因此 $J_{n_i}(c_i)$ 相差次序的意义下被 T 决定. $\square$

推论 7.25. 假设 $A \in F^{n \times n}$ 的特征多项式可以分解为一次式的乘积. 则 A 相似于某个Jordan形矩阵, 并且Jordan块在相差次序的意义下是唯一的, 称为 A 的**Jordan标准形**. 特别地, 如果 F 是代数闭域, 则任意方阵存在Jordan标准形.

注 7.3. • 这里与书上稍有区别: 书上要求相同特征值在一起, 并且从大到小排列.

- 更多文献中把“1”写在主对角线上方. 我们为了方便, 与教材保持一致. 注意到两者是相似的. 事实上, 如果 T 在有序基 $\{\alpha_1, \dots, \alpha_n\}$ 下的矩阵是Jordan块 J , 则 T 在有序基 $\{\alpha_n, \dots, \alpha_1\}$ 下的矩阵是 J^t .

从 T 的Jordan标准形 $[T]_{\mathcal{B}}$ 可以直接看出很多 T 的性质:

- $[T]_{\mathcal{B}}$ 的对角元集合为 $\sigma(T)$. 设为 $\{c_1, \dots, c_k\}$. 如果 c_i 在 $[T]_{\mathcal{B}}$ 的对角元中出现 d_i 次, 则 $d_i = \dim \text{Ker}((T - c_i I)^n)$, 其中 $n = \dim V$, 并且 $f_T = \prod_{i=1}^k (x - c_i)^{d_i}$.
- 设特征值为 c_i 的Jordan块的最大阶数为 r_i . 则 $p_T = \prod_{i=1}^k (x - c_i)^{r_i}$. 特别地, T 可对角化 $\iff$ 所有Jordan块都是1阶的.
- 特征值为 c_i 的Jordan块的个数为 $\dim V_{c_i}$.

例 7.4. 2×2 的Jordan形矩阵只能是 $\begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$ 或 $\begin{bmatrix} c & 0 \\ 1 & c \end{bmatrix}$.

例 7.5. $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ 和 $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ 都是Jordan形矩阵, 特征多项式都是 $(x - 2)^4$, 最小多项式

都是 $(x - 2)^2$, 但它们不相似. (考察 $\dim \text{Ker}(A - 2I)$.)

例 7.6. 设 $A = \begin{bmatrix} 2 & 0 & 0 \\ a & 2 & 0 \\ b & c & -1 \end{bmatrix} \in \mathbb{C}^{3 \times 3}$. 则 $f_A = (x - 2)^2(x + 1)$. 如果 $p_A = f_A$, 则 A 的Jordan标准形

为 $\begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. 如果 $p_A = (x - 2)(x + 1)$, 则 A 的Jordan标准形为 $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. 后者发生

$$\iff (A - 2I)(A + I) = \begin{bmatrix} 0 & 0 & 0 \\ a & 0 & 0 \\ b & c & -3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ a & 3 & 0 \\ b & c & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 3a & 0 & 0 \\ ac & 0 & 0 \end{bmatrix} = 0 \iff a = 0.$$

现在考察与Jordan标准形密切相关的一个问题. 为了简单起见, 假设 F 是代数闭域. 设 $A \in F^{n \times n}$ 的Jordan标准形为 $J = P^{-1}AP$, 其中 $P \in GL_n(F)$. 记 J_d 是把 J 中不在主对角线上的矩阵元“1”都替换为“0”后得到的矩阵, J_n 是把 J 的对角元都替换为“0”后得到的矩阵. 则 J_d 是对角矩阵, J_n 幂零, J_d 与 J_n 可交换, 并且 $J = J_d + J_n$. 记 $A_d = PJ_dP^{-1}$, $A_n = PJ_nP^{-1}$. 则 A_d 可对角化,

A_n 幂零, A_d 与 A_n 可交换, 并且 $A = A_d + A_n$. 下面说明这样得到的 A_d 和 A_n 与 J 和 P 的选取无关.

定理 7.26. 设 V 是代数闭域 F 上的有限维线性空间, $T \in L(V)$.

- (1) 存在唯一的 $D, N \in L(V)$ 满足 $T = D + N$, D 可对角化, N 幂零, 并且 D 与 N 可交换.
- (2) D 和 N 都是 T 的多项式.

证明. 设 $p_T = \prod_{i=1}^k (x - c_i)^{r_i}$, 其中 c_i 互不相同并且 $r_i \geq 1$. 记 $q_i = (x - c_i)^{r_i}$. 我们断言存在 $f \in F[x]$ 满足

$$f \equiv c_i \pmod{q_i}, \quad i = 1, \dots, k.$$

这由多项式的中国剩余定理是显然的, 还可以直接证明如下: 由于 $h_i := \prod_{j \neq i} q_j$ 与 q_i 互素, 存在 g_i 满足 $g_i h_i \equiv 1 \pmod{q_i}$. 取 $f = \sum_{i=1}^k c_i g_i h_i$ 即可. 设 $D = f(T)$, $N = T - D$. 则 D, N 都是 T 的多项式, 从而可交换. 我们验证 D 可对角化, N 幂零. 记 $V_i = \text{Ker}(q_i(T))$. 由准素分解, $V = \bigoplus_{i=1}^k V_i$ 并且 $p_{T_{V_i}} = q_i$. 由 f 的性质有 $D_{V_i} = f(T_{V_i}) = c_i \text{id}_{V_i}$. 这说明 c_i 是 D 的特征值, 并且相应的特征子空间是 V_i . 因此 D 可对角化. 另一方面, $p_{T_{V_i}} = q_i$ 推出 $N_{V_i} = T_{V_i} - c_i \text{id}_{V_i}$ 的最小多项式为 x^{r_i} . 所以 N_{V_i} 幂零. 因此 N 幂零.

设 D', N' 也满足 (1). 由于上面的 D, N 都是 T 的多项式, 所以 D, N, D', N' 两两可交换. 注意到

$$D - D' = N' - N.$$

由于 D, D' 可对角化并且可交换, 它们可同时对角化, 从而 $D - D'$ 可对角化. 另一方面, 设 $N^r = (N')^r = 0$, 则

$$(N' - N)^{2r} = \sum_{i=0}^{2r} C_{2r}^i (N')^i (-N)^{2r-i} = 0.$$

因此 $N' - N$ 幂零. 而幂零的可对角化变换只能是 0. 因此 $D - D' = N' - N = 0$. $\square$

这里的 D 和 N 分别称为 T 的**可对角化部分**(或**半单部分**)和**幂零部分**. 分解 $T = D + N$ 称为 T 的**Jordan 分解**.

注 7.4. $\sigma(D) = \sigma(T)$, $f_D = f_T$. 如果 $p_T = \prod_{i=1}^k p_i^{r_i}$, 其中 p_i 是互不相同的首项系数是 1 的素多项式并且 $r_i \geq 1$, 则 $p_D = \prod_{i=1}^k p_i$.

§7.4 半单变换

对应教材 7.5 节.

定义 7.3. 设 V 是域 F 上的非零有限维线性空间, $T \in L(V)$.

- (1) 如果 V 没有非平凡的 T -不变子空间, 则称 T 为**单**(simple)的或**不可约**(irreducible)的.
- (2) 如果对任意 T -不变子空间 $W \subset V$, 存在 T -不变子空间 $Z \subset V$ 满足 $V = W \oplus Z$, 则称 T 为**半单**(semisimple)的或**完全可约**(completely reducible)的.

容易看出, T 单 $\iff V$ 中所有非零向量是循环向量, T 单 $\implies T$ 不可分解.

引理 7.27. 设 $T \in L(V)$.

- (1) T 单 $\implies T$ 半单.
- (2) 设 T 半单, $V' \subset V$ 是 T -不变子空间. 则 $T_{V'}$ 半单.
- (3) 设 $V = \bigoplus_{i=1}^k V_i$, 其中每个 V_i 是 T -不变子空间. 假设每个 T_{V_i} 半单. 则 T 半单.

证明. (1) 显然.

(2) 设 $W \subset V'$ 是不变子空间. 取 V 的不变子空间 Z 满足 $V = W \oplus Z$. 我们断言 $V' = W \oplus (V' \cap Z)$. 显然 $W \cap (V' \cap Z) = \{0\}$. 另一方面, 对 $v \in V'$, 设 $v = w + z$, 其中 $w \in W, z \in Z$. 则 $z = v - w \in V'$. 所以 $z \in V' \cap Z$. 因此 $v = w + z \in W + (V' \cap Z)$.

(3) 只需证明 $k = 2$ 的情况. 设 $W \subset V$ 是不变子空间. 由于 T_{V_i} 半单, 可以取不变子空间 $Z_1 \subset V_1$ 和 $Z_2 \subset V_2$ 满足 $V_1 = (W \cap V_1) \oplus Z_1, V_2 = ((W + V_1) \cap V_2) \oplus Z_2$. 我们验证 $V = W \oplus (Z_1 + Z_2)$. 首先, 如果 $w = z_1 + z_2 \in W \cap (Z_1 + Z_2)$, 其中 $z_i \in Z_i$, 则 $z_2 = w - z_1 \in (W + V_1) \cap Z_2 = \{0\}$, 从而 $z_2 = 0$. 这说明 $w = z_1 \in W \cap Z_1 = \{0\}$, 从而 $w = 0$. 因此 $W \cap (Z_1 + Z_2) = \{0\}$. 另一方面,

$$\begin{aligned} V &= V_1 + V_2 = V_1 + ((W + V_1) \cap V_2) + Z_2 \subset V_1 + (W + V_1) + Z_2 = V_1 + W + Z_2 \\ &= ((W \cap V_1) + Z_1) + W + Z_2 \subset (W + Z_1) + W + Z_2 = W + (Z_1 + Z_2). \end{aligned}$$

因此 $V = W \oplus (Z_1 + Z_2)$. □

命题 7.28. T 半单 $\iff$ 存在 T -不变直和分解 $V = \bigoplus_{i=1}^k V_i$ 使得每个 T_{V_i} 单.

证明. “ $\implies$ ”: 由(2). “ $\impliedby$ ”: 由(1)和(3). □

命题 7.29. (1) T 单 $\iff f_T$ 素.

(2) T 半单 $\iff p_T$ 无平方因子.

证明. (1) 设 T 单. 则 T 循环, 因此 $p_T = f_T$. 如果 f_T 不素, 设 $f_T = gh, \deg g, \deg h \geq 1$. 则 $g(T)h(T) = 0$. 因此 $g(T)$ 或 $h(T)$ 不可逆. 不妨设 $g(T)$ 不可逆. 则 $\text{Ker}(g(T))$ 是非零不变子空间, 因此为全空间. 这说明 $g(T) = 0$. 因此 $f_T = p_T | g$, 矛盾. 反过来, 设 T 不单, 即存在非平凡不变子空间 $W \subset V$. 则 $f_T = f_{T_W} f_{T_{V/W}}$. 因此 f_T 不素.

(2) 设 T 半单. 则 $T = \bigoplus_{i=1}^k T_i$, 其中 T_i 单. 于是 p_{T_i} 素. 因此 $p_T = \text{lcm}(p_{T_1}, \dots, p_{T_k})$ 无平方因子. 反过来, 设 $p_T = \prod_{i=1}^k p_i, p_i$ 素并且互不相同. 则在准素循环分解中, 对每个直和项有 $p = f$ 素. 因此每个直和项单. 因此 T 半单. □

推论 7.30. 设 f_T 可以分解为 $F[x]$ 中一次式的乘积(例如 F 是代数闭域).

(1) T 单 $\iff \dim V = 1$.

(2) T 半单 $\iff T$ 可对角化.

证明. 对(2), 只需注意到 T 可对角化 $\iff p_T$ 可以分解为 $F[x]$ 中互不相伴的一次式的乘积. □

下面承认 F 的代数闭包 $\bar{F}$ 的存在性. 很多重要的问题涉及到 $f \in F[x] \setminus F$ 在 $\bar{F}$ 中是否有重根. f 在 $\bar{F}$ 内有重根可以有两个原因造成: 如果 f 在 $F[x]$ 内有相同的素因子(即有平方因子), 则显然 f 在 $\bar{F}$ 内有重根. 另一个原因是由域 F 本身的性质造成的.

定义 7.4. 如果域 F 上的任意无平方因子的多项式 $f \in F[x] \setminus F$ 在 $\bar{F}$ 内无重根, 则称 F 是**完全域**.

可以证明:

- F 是完全域 $\iff F$ 上的任意素多项式在 $\bar{F}$ 内无重根.
- 代数闭域、特征是 0 的域和有限域都是完全域.

但 $F = \mathbb{F}_p(t)$ 不是完全域: $f = x^p - t$ 是 $F[x]$ 中的素多项式, 但如果 $a \in \bar{F}$ 是 f 的根则 $f = (x - a)^p$. 完全域的一个重要性质为: 设 F 是完全域, $a \in \bar{F}$. 则 $a \in F \iff$ 对所有 $\sigma \in \text{Gal}(\bar{F}/F)$ 有 $\sigma(a) = a$.

推论 7.31. 设 $A \in F^{n \times n}$.

(1) A 在 $\overline{F}$ 上可对角化 $\implies L_A$ 半单.

(2) 如果 F 是完全域, 则 A 在 $\overline{F}$ 上可对角化 $\iff L_A$ 半单.

证明. 只需注意到: A 在 $\overline{F}$ 上可对角化 $\iff p_A$ 在 $\overline{F}$ 中无重根; L_A 半单 $\iff p_A$ 无平方因子. $\square$

第八章 内积空间

§8.1 内积

本章我们总是假设 $F = \mathbb{R}$ 或 $\mathbb{C}$.

定义 8.1. 设 V 是 F -线性空间. V 上的一个内积指一个函数 $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$, $(\alpha, \beta) \mapsto \langle \alpha, \beta \rangle$ 满足:

- (a) $\langle \alpha + \beta, \gamma \rangle = \langle \alpha, \gamma \rangle + \langle \beta, \gamma \rangle$,
- (b) $\langle c\alpha, \beta \rangle = c\langle \alpha, \beta \rangle$,
- (c) $\langle \beta, \alpha \rangle = \overline{\langle \alpha, \beta \rangle}$, ($\implies \langle \alpha, \beta + \gamma \rangle = \langle \alpha, \beta \rangle + \langle \alpha, \gamma \rangle$, $\langle \alpha, c\beta \rangle = \bar{c}\langle \alpha, \beta \rangle$)
- (d) $\alpha \neq 0 \implies \langle \alpha, \alpha \rangle > 0$.

取定了内积的 F -线性空间称为**内积空间**. 有限维实内积空间也叫**Euclid空间**(Euclidean space). 复内积空间也叫**酉空间**.

注 8.1. 当 $F = \mathbb{C}$ 时, 只能要求 $\langle \alpha, c\beta \rangle = \bar{c}\langle \alpha, \beta \rangle$ ($1\frac{1}{2}$ -线性). 如果要求双线性, 则与(d)不相容:

$$\langle \alpha, \alpha \rangle > 0 \implies \langle i\alpha, i\alpha \rangle = -\langle \alpha, \alpha \rangle < 0.$$

例 8.1. 设 $V = F^{n \times 1}$. 对 $\alpha = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\beta = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$,

$$\langle \alpha, \beta \rangle = \sum_{j=1}^n x_j \bar{y}_j = \alpha^t \bar{\beta} = \beta^* \alpha$$

定义了 $F^{n \times 1}$ 上的内积, 称为**标准内积**. 其中对 $A \in F^{m \times n}$, $A^* := \overline{A^t} \in F^{n \times m}$ 为 A 的转置共轭. 类似地, 可以定义 F^n 上的标准内积.

例 8.2. 设 $V = F^{m \times n}$.

$$\langle A, B \rangle = \text{tr}(AB^*) = \text{tr}(B^*A) = \sum_{j,k} A_{jk} \overline{B_{jk}}$$

定义了 $F^{m \times n}$ 上的内积.

定义 8.2. 矩阵 $A \in F^{n \times n}$ 称为**Hermite矩阵**(Hermitian matrix) ($F = \mathbb{R}$ 时称为**对称矩阵**), 如果 $A^* = A$. 如果进一步还有

$$X^*AX > 0, \quad \forall X \in F^{n \times 1} \setminus \{0\},$$

则称 A 是**正定的**.

例 8.3. 设 $Q \in \text{GL}_n(F)$. 则 $A = Q^*Q$ 正定 Hermite.

引理 8.1. 设 $\dim V = n$, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是 V 的有序基. 对于 $\alpha, \beta \in V$, 记 $X = [\alpha]_{\mathcal{B}}$, $Y = [\beta]_{\mathcal{B}}$.

(1) 若 $A \in F^{n \times n}$ 正定 Hermite, 则

$$\langle \alpha, \beta \rangle = Y^*AX = \sum_{j,k} A_{kj} x_j \bar{y}_k \quad (8.1)$$

定义了 V 上的内积.

(2) 对任意 V 上的内积 $\langle \cdot, \cdot \rangle$, 存在唯一的正定 Hermite 矩阵 $A \in F^{n \times n}$ 使得 (8.1) 成立, 称为该**内积**

在有序基 $\mathcal{B}$ 下的矩阵.

证明. (1) 容易验证(a)-(d)成立.

(2) 唯一性: 如果矩阵 $A \in F^{n \times n}$ 满足(8.1), 取 $\alpha = \alpha_j, \beta = \alpha_k$ 得

$$A_{kj} = \langle \alpha_j, \alpha_k \rangle.$$

存在性: 由此式定义矩阵 A . 则

$$\langle \alpha, \beta \rangle = \left\langle \sum_j x_j \alpha_j, \sum_k y_k \alpha_k \right\rangle = \sum_{j,k} x_j \bar{y}_k \langle \alpha_j, \alpha_k \rangle = \sum_{j,k} x_j \bar{y}_k A_{kj} = Y^* A X,$$

即(8.1)成立. 由内积定义即得 A 正定Hermite. □

例 8.4. 设 $A = \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \in \mathbb{R}^{2 \times 2}$. 显然 A 对称. 由于对 $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0$ 有

$$X^t A X = (x_1, x_2) \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 - 2x_1 x_2 + 4x_2^2 = (x_1 - x_2)^2 + 3x_2^2 > 0,$$

所以 A 正定. 因此,

$$\langle \alpha, \beta \rangle = \beta^t A \alpha = \alpha^t A \beta = (x_1, x_2) \begin{bmatrix} 1 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1 y_1 - x_1 y_2 - x_2 y_1 + 4x_2 y_2$$

定义了 $\mathbb{R}^{2 \times 1}$ 上的内积, 它在标准基下的矩阵是 A .

例 8.5. 设 $T: V \rightarrow W$ 是单线性映射, $\langle \cdot, \cdot \rangle_0$ 是 W 上的内积. 则

$$\langle \alpha, \beta \rangle := \langle T\alpha, T\beta \rangle_0$$

定义了 V 上的内积. 例如:

- 设 $\langle \cdot, \cdot \rangle_0$ 是 $F^{n \times 1}$ 上的标准内积, $Q \in \text{GL}_n(F), T = L_Q$. 则

$$\langle \alpha, \beta \rangle = \langle Q\alpha, Q\beta \rangle_0 = (Q\beta)^* Q\alpha = \beta^* (Q^* Q) \alpha$$

也是 $F^{n \times 1}$ 上的内积. 它在标准基下的矩阵为 $Q^* Q$.

- 设 $\dim V = n, \mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是 V 的有序基, $T = \Gamma_{\mathcal{B}}: V \rightarrow F^{n \times 1}$ 是坐标映射, $\langle \cdot, \cdot \rangle_0$ 是 $F^{n \times 1}$ 上的标准内积. 则 $\langle \alpha, \beta \rangle = \langle T\alpha, T\beta \rangle_0$, 即

$$\left\langle \sum_j x_j \alpha_j, \sum_k y_k \alpha_k \right\rangle = \sum_j x_j \bar{y}_j$$

是 V 上的内积. 它在有序基 $\mathcal{B}$ 下的矩阵为 I_n . 特别地, $F^{n \times 1}$ 上的标准内积在标准基下的矩阵是 I_n .

设 V 是内积空间. 我们定义 $\alpha \in V$ 的**长度**为 $\|\alpha\| := \langle \alpha, \alpha \rangle^{\frac{1}{2}}$. 实际上, 内积被长度决定:

引理 8.2(极化恒等式). (1) 若 $F = \mathbb{R}$, 则

$$\langle \alpha, \beta \rangle = \frac{1}{4}(\|\alpha + \beta\|^2 - \|\alpha - \beta\|^2).$$

(2) 若 $F = \mathbb{C}$, 则

$$\langle \alpha, \beta \rangle = \frac{1}{4} \sum_{k=1}^4 i^k \|\alpha + i^k \beta\|^2.$$

证明. 略. □

§8.2 内积空间

长度有下面的性质:

引理 8.3. (1) $\|c\alpha\| = |c|\|\alpha\|$.

(2) $\alpha \neq 0 \implies \|\alpha\| > 0$.

(3) (Cauchy-Schwarz不等式) $|\langle \alpha, \beta \rangle| \leq \|\alpha\|\|\beta\|$, 并且“=”成立 $\iff \alpha$ 与 β 线性相关.

(4) (三角不等式) $\|\alpha + \beta\| \leq \|\alpha\| + \|\beta\|$.

证明. (1), (2)显然.

(3) 不妨设 $\alpha \neq 0$. 设 $\gamma = \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2} \alpha$ (是 β 在 $\alpha^\perp$ 上的正交投影). 则 $\langle \gamma, \alpha \rangle = 0$. 从而

$$0 \leq \|\gamma\|^2 = \left\langle \gamma, \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2} \alpha \right\rangle = \langle \gamma, \beta \rangle = \left\langle \beta - \frac{\langle \beta, \alpha \rangle}{\|\alpha\|^2} \alpha, \beta \right\rangle = \langle \beta, \beta \rangle - \frac{\langle \beta, \alpha \rangle \langle \alpha, \beta \rangle}{\|\alpha\|^2} = \|\beta\|^2 - \frac{|\langle \alpha, \beta \rangle|^2}{\|\alpha\|^2}.$$

这就推出想要的 $\leq$ 不等式. “=”成立 $\iff \gamma = 0 \iff \alpha$ 与 β 线性相关.

(4) 由(3)有

$$\begin{aligned} \|\alpha + \beta\|^2 &= \|\alpha\|^2 + \|\beta\|^2 + \langle \alpha, \beta \rangle + \langle \beta, \alpha \rangle = \|\alpha\|^2 + \|\beta\|^2 + 2\operatorname{Re}\langle \alpha, \beta \rangle \\ &\leq \|\alpha\|^2 + \|\beta\|^2 + 2|\langle \alpha, \beta \rangle| \leq \|\alpha\|^2 + \|\beta\|^2 + 2\|\alpha\|\|\beta\| = (\|\alpha\| + \|\beta\|)^2. \end{aligned}$$

□

例 8.6. 对 $\mathbb{R}^n$ 上的标准内积应用Cauchy-Schwarz不等式, 得

$$\left(\sum_{j=1}^n x_j y_j \right)^2 \leq \left(\sum_{j=1}^n x_j^2 \right) \left(\sum_{j=1}^n y_j^2 \right).$$

定义 8.3. 设 V 是内积空间.

- 对于 $\alpha, \beta \in V$, 如果 $\langle \alpha, \beta \rangle = 0$, 则称 α 与 β 垂直或正交, 记为 $\alpha \perp \beta$.
- 如果 V 的子集 S 中的向量两两正交, 则称 S 为正交集(orthogonal set). 如果进一步还有 $\|\alpha\| = 1, \forall \alpha \in S$, 则称 S 为标准正交集(orthonormal set).
- 是正交集的基称为正交基, 是标准正交集的基称为标准正交基.

例 8.7. 注意0与任何向量垂直. $F^{n \times 1}$ 的标准基在标准内积下是标准正交基.

当 $F = \mathbb{R}$ 时, 对于 $\alpha, \beta \in V \setminus \{0\}$, 由Cauchy-Schwarz不等式, 可以定义 α 与 β 之间的角度为

$$\angle(\alpha, \beta) := \arccos \frac{\langle \alpha, \beta \rangle}{\|\alpha\|\|\beta\|}.$$

则 $\alpha \perp \beta \iff \angle(\alpha, \beta) = \pi/2$. 注意有

$$\langle \alpha, \beta \rangle = \|\alpha\|\|\beta\| \cos \angle(\alpha, \beta).$$

引理 8.4. 如果 $\alpha \perp \beta$, 则

$$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2.$$

更一般地, 如果 $\{\alpha_1, \dots, \alpha_m\}$ 是正交集, 则

$$\left\| \sum_{j=1}^m \alpha_j \right\|^2 = \sum_{j=1}^m \|\alpha_j\|^2.$$

证明. 由引理8.3(4)的证明显然.

□

注意如果 $\{\alpha_1, \dots, \alpha_n\}$ 是标准正交基, 则

$$\left\langle \sum_{j=1}^n x_j \alpha_j, \sum_{k=1}^n y_k \alpha_k \right\rangle = \sum_{j=1}^n x_j \bar{y}_j.$$

特别地,

$$\left\| \sum_{j=1}^n x_j \alpha_j \right\|^2 = \sum_{j=1}^n |x_j|^2.$$

命题 8.5. 不含零向量的正交集总是线性无关的.

证明. 设 S 是不含零向量的正交集, $\alpha_1, \dots, \alpha_m \in S$. 假设 $\sum_{j=1}^m c_j \alpha_j = 0$. 则

$$0 = \left\langle \sum_{j=1}^m c_j \alpha_j, \alpha_k \right\rangle = c_k \|\alpha_k\|^2 \implies c_k = 0.$$

□

因此, 为了说明一个不含零向量的正交集是正交基, 只需说明它生成全空间.

命题 8.6. 设 V 是有限维内积空间, $\{\beta_1, \dots, \beta_n\}$ 是 V 的有序基.

(1) 存在唯一的 V 的正交基 $\{\alpha_1, \dots, \alpha_n\}$ 和对角元为1的上三角矩阵 N 满足

$$[\beta_1, \dots, \beta_n] = [\alpha_1, \dots, \alpha_n]N. \quad (8.2)$$

(2) 存在唯一的 V 的标准正交基 $\{\alpha'_1, \dots, \alpha'_n\}$, 对角元为1的上三角矩阵 N 和对角元为正数的对角矩阵 A 满足

$$[\beta_1, \dots, \beta_n] = [\alpha'_1, \dots, \alpha'_n]AN. \quad (8.3)$$

此时有

$$\text{span}\{\alpha_1, \dots, \alpha_k\} = \text{span}\{\alpha'_1, \dots, \alpha'_k\} = \text{span}\{\beta_1, \dots, \beta_k\}, \quad k = 1, \dots, n.$$

证明. (1) 将(8.2)式重写为

$$\begin{aligned} \beta_1 &= \alpha_1, \\ \beta_2 &= N_{12}\alpha_1 + \alpha_2, \\ &\vdots \\ \beta_n &= N_{1n}\alpha_1 + \dots + N_{n-1,n}\alpha_{n-1} + \alpha_n. \end{aligned}$$

注意前 m 个等式推出

$$\text{span}\{\alpha_1, \dots, \alpha_m\} = \text{span}\{\beta_1, \dots, \beta_m\}.$$

我们归纳证明: 对任意 $1 \leq m \leq n$, 存在唯一的 $\alpha_1, \dots, \alpha_m \in V \setminus \{0\}$ 和 $N_{jk} \in F (1 \leq j < k \leq m)$ 使前 m 个等式成立, 并且 $\{\alpha_1, \dots, \alpha_m\}$ 是正交集. $m = 1$ 时无需证明. 假设 $1 < m \leq n$, 并且结论

对 $m-1$ 成立. 则

α_m 和 $N_{1m}, \dots, N_{m-1,m}$ 使得前 m 个等式成立并且 $\{\alpha_1, \dots, \alpha_m\}$ 是正交集

$$\iff \alpha_m \text{ 和 } N_{1m}, \dots, N_{m-1,m} \text{ 使得第 } m \text{ 个等式成立 (即 } \beta_m = \alpha_m + \sum_{j=1}^{m-1} N_{jm} \alpha_j \text{) 并且 } \alpha_m \perp \alpha_k (1 \leq k \leq m-1)$$

$$\iff \alpha_m = \beta_m - \sum_{j=1}^{m-1} N_{jm} \alpha_j \text{ 并且 } 0 = \langle \alpha_m, \alpha_k \rangle = \langle \beta_m - \sum_{j=1}^{m-1} N_{jm} \alpha_j, \alpha_k \rangle = \langle \beta_m, \alpha_k \rangle - N_{km} \|\alpha_k\|^2 (1 \leq k \leq m-1)$$

$$\iff N_{km} = \frac{\langle \beta_m, \alpha_k \rangle}{\|\alpha_k\|^2} (1 \leq k \leq m-1) \text{ 并且 } \alpha_m = \beta_m - \sum_{j=1}^{m-1} N_{jm} \alpha_j.$$

因此 α_m 和 $N_{1m}, \dots, N_{m-1,m}$ 被唯一决定, 即结论的唯一性部分对 m 成立. 另外, 注意到这样决定出的 $\alpha_m \neq 0$. (否则, 将有

$$\beta_m = \sum_{j=1}^{m-1} N_{jm} \alpha_j \in \text{span}\{\alpha_1, \dots, \alpha_{m-1}\} = \text{span}\{\beta_1, \dots, \beta_{m-1}\},$$

与 $\{\beta_1, \dots, \beta_n\}$ 是基矛盾. 这里最后一个等号由归纳假设推出.) 因此, 结论的存在性部分对 m 也成立. 在结论中取 $m = n$ 即得 (1).

(2) 存在性: 由 (1) 的存在性部分, 可取正交基 $\{\alpha_1, \dots, \alpha_n\}$ 和对角元为 1 的上三角矩阵 N 满足 (8.2). 取 $\alpha'_j = \frac{\alpha_j}{\|\alpha_j\|}$, $A = \text{diag}(\|\alpha_1\|, \dots, \|\alpha_n\|)$. 则 (8.3) 成立. 注意到

$$\text{span}\{\alpha_1, \dots, \alpha_k\} = \text{span}\{\alpha'_1, \dots, \alpha'_k\}.$$

唯一性: 设 (8.3) 成立. 设 $A = \text{diag}(a_1, \dots, a_n)$, $\alpha_j = a_j \alpha'_j$. 则 (8.2) 成立. 由 (1) 的唯一性部分, α_j 和 N 被 β_j 唯一决定. 另一方面, α'_j 和 a_j 被 α_j 唯一决定, 从而也被 β_j 唯一决定. $\square$

在 (1) 的证明中, 利用

$$N_{km} = \frac{\langle \beta_m, \alpha_k \rangle}{\|\alpha_k\|^2}, \quad \alpha_m = \beta_m - \sum_{j=1}^{m-1} N_{jm} \alpha_j$$

可以由 β_j 归纳构造 α_j . 这一过程称为 **Gram-Schmidt 正交化**. 由任意 m 个向量 $\beta_1, \dots, \beta_m$ 出发, 应用 Gram-Schmidt 正交化过程, 如果在某步有 $\alpha_j = 0$, 则 $\beta_1, \dots, \beta_m$ 线性相关. 如果每个 $\alpha_j \neq 0$, 则 $\beta_1, \dots, \beta_m$ 线性无关, 并且对任意 $k \in \{1, \dots, m\}$, $\{\alpha_1, \dots, \alpha_k\}$ 是 $\text{span}\{\beta_1, \dots, \beta_k\}$ 的正交基.

推论 8.7. 设 V 是有限维内积空间. 则 V 中任意不含零向量的正交集可以扩充为 V 的正交基, V 中的任意标准正交集可以扩充为 V 的标准正交基. 特别地, V 存在标准正交基. $\square$

命题 8.8. 设 $S = \{\alpha_1, \dots, \alpha_m\}$ 是不含零向量的正交集, $\beta \in \text{span} S$. 则

$$\beta = \sum_{k=1}^m \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$

特别地, 如果 $\{\alpha_1, \dots, \alpha_n\}$ 是 V 的正交基, 则

$$\beta = \sum_{k=1}^n \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$

如果 $\{\alpha_1, \dots, \alpha_n\}$ 还是 V 的标准正交基, 则

$$\beta = \sum_{k=1}^n \langle \beta, \alpha_k \rangle \alpha_k.$$

证明. 设 $\beta = \sum_{j=1}^m c_j \alpha_j$. 则

$$\langle \beta, \alpha_k \rangle = \left\langle \sum_{j=1}^m c_j \alpha_j, \alpha_k \right\rangle = c_k \|\alpha_k\|^2 \implies c_k = \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2}.$$

□

命题 8.9 (Bessel不等式). 设 $S = \{\alpha_1, \dots, \alpha_m\}$ 是不含零向量的正交集, $\beta \in V$. 则

$$\sum_{k=1}^m \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2} \leq \|\beta\|^2.$$

并且 “=” 成立 $\iff \beta \in \text{span} S$.

证明. 将 S 扩充为 V 的正交基 $\{\alpha_1, \dots, \alpha_n\}$. 则

$$\beta = \sum_{k=1}^n \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k.$$

因此

$$\|\beta\|^2 = \sum_{k=1}^n \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2} \geq \sum_{k=1}^m \frac{|\langle \beta, \alpha_k \rangle|^2}{\|\alpha_k\|^2}.$$

等号成立 $\iff \langle \beta, \alpha_{m+1} \rangle = \dots = \langle \beta, \alpha_n \rangle = 0 \iff \beta = \sum_{k=1}^m \frac{\langle \beta, \alpha_k \rangle}{\|\alpha_k\|^2} \alpha_k \iff \beta \in \text{span} S$. □

这推出: 如果 $S = \{\alpha_1, \dots, \alpha_m\}$ 是标准正交集, $\beta \in V$, 则

$$\sum_{k=1}^m |\langle \beta, \alpha_k \rangle|^2 \leq \|\beta\|^2,$$

并且 “=” 成立 $\iff \beta \in \text{span} S$.

例 8.8. P282-283例12,13. 略.

定义 8.4. 设 S 是内积空间 V 的子集. 定义 S 在 V 中的正交补为

$$S^\perp := \{\alpha \in V \mid \alpha \perp \beta, \forall \beta \in S\}.$$

容易验证 $S^\perp$ 总是子空间, 并且 $S^\perp = (\text{span} S)^\perp$.

命题 8.10. 设 V 是有限维内积空间, $W \subset V$ 是子空间. 则 $\dim W + \dim W^\perp = \dim V$.

证明. 取 W 的标准正交基 $\{\alpha_1, \dots, \alpha_m\}$, 并扩充为 V 的标准正交基 $\{\alpha_1, \dots, \alpha_n\}$. 对于 $\beta \in V$, 有

$$\beta = \sum_{k=1}^n \langle \beta, \alpha_k \rangle \alpha_k.$$

因此

$$\begin{aligned} \beta \in W^\perp &\iff \langle \beta, \alpha_1 \rangle = \dots = \langle \beta, \alpha_m \rangle = 0 \\ &\iff \beta = \sum_{k=m+1}^n \langle \beta, \alpha_k \rangle \alpha_k \\ &\iff \beta \in \text{span}\{\alpha_{m+1}, \dots, \alpha_n\}. \end{aligned}$$

从而 $W^\perp = \text{span}\{\alpha_{m+1}, \dots, \alpha_n\}$, 因此 $\{\alpha_{m+1}, \dots, \alpha_n\}$ 是 $W^\perp$ 的基. □

注 8.2. 我们以前曾证明: 设 W 是有限维线性空间 V 的子空间, 并定义 $W^0 = \{f \in V^* \mid f|_W = 0\}$, 则 $\dim W + \dim W^0 = \dim V$. 上面的命题可以视为这一结论的特殊情况. 实际上, 当 $F = \mathbb{R}$ 或 $\mathbb{C}$ 时, V 上的一个内积 $\langle \cdot, \cdot \rangle$ 决定了一个映射 $\Phi: V \rightarrow V^*$, $\Phi(\beta)(\alpha) = \langle \alpha, \beta \rangle$. (书上记 $\Phi(\beta) = f_\beta$.) 容易看出, Φ 即单又满, 并且是实线性的. (注意当 $F = \mathbb{C}$ 时有 $\Phi(c\beta) = \bar{c}\Phi(\beta)$, 因此 Φ 不是复线性的.) 此外, 如果 $W \subset V$ 是子空间, 则 $\Phi(W^\perp) = W^0$. 因此 $\dim W^\perp = \dim W^0$.

推论 8.11. 设 V 是有限维内积空间, $W \subset V$ 是子空间. 则 $(W^\perp)^\perp = W$.

证明. 显然 $W \subset (W^\perp)^\perp$. 而 $\dim W = \dim (W^\perp)^\perp$. 所以 $(W^\perp)^\perp = W$. $\square$

推论 8.12. 设 V 是有限维内积空间, $W \subset V$ 是子空间. 则 $V = W \oplus W^\perp$.

证明. 显然 $W \cap W^\perp = \{0\}$. 而

$$\dim(W \oplus W^\perp) = \dim W + \dim W^\perp = \dim V.$$

所以 $W \oplus W^\perp = V$. $\square$

我们把沿 $W^\perp$ 到 W 上的投影称为到 W 上的正交投影, 记为 P_W . 对于 $\beta \in V$, 我们也称 $P_W\beta$ 为 β 在 W 上的正交投影. 容易看出, 如果 $\{\alpha_1, \dots, \alpha_m\}$ 是 W 的正交基, 则

$$P_W\beta = \sum_{j=1}^m \frac{\langle \beta, \alpha_j \rangle}{\|\alpha_j\|^2} \alpha_j.$$

如果 $\{\alpha_1, \dots, \alpha_m\}$ 还是 W 的标准正交基, 则

$$P_W\beta = \sum_{j=1}^m \langle \beta, \alpha_j \rangle \alpha_j.$$

注 8.3. • 我们知道, 如果 $V = R \oplus N$, E 是沿 N 到 R 上的投影, 则 $I - E$ 是沿 R 到 N 上的投影. 因此, $P_{W^\perp} = I - P_W$.

• 在 Gram-Schmidt 正交化过程中, 记 $W_k = \text{span}\{\beta_1, \dots, \beta_k\}$, 则 $\alpha_k = P_{W_{k-1}^\perp} \beta_k$.

命题 8.13. $P_W\beta$ 是 β 在 W 中的最佳逼近, 即函数 $W \rightarrow \mathbb{R}, \alpha \mapsto \|\beta - \alpha\|$ 在且只在 $P_W\beta$ 处达到最小值.

证明. 对 $\alpha \in W$ 有 $\beta - \alpha = (\beta - P_W\beta) + (P_W\beta - \alpha)$. 由于 $\beta - P_W\beta \in W^\perp$, $P_W\beta - \alpha \in W$, 所以 $(\beta - P_W\beta) \perp (P_W\beta - \alpha)$. 因此

$$\|\beta - \alpha\|^2 = \|\beta - P_W\beta\|^2 + \|P_W\beta - \alpha\|^2.$$

因此 $\|\beta - \alpha\|$ 达到最小值 $\iff \|P_W\beta - \alpha\|$ 达到最小值 $\iff \alpha = P_W\beta$. $\square$

§8.3 线性函数和伴随变换

设 V 是有限维内积空间. 考虑上一节定义的映射 $\Phi: V \rightarrow V^*, \Phi(\beta)(\alpha) = \langle \alpha, \beta \rangle$.

引理 8.14. Φ 即单又满. 并且共轭线性, 即 $\Phi(c\beta_1 + \beta_2) = \bar{c}\Phi(\beta_1) + \Phi(\beta_2)$.

证明. 为证即单又满, 只需证明对任意 $f \in V^*$, 存在唯一的 β 使 $\Phi(\beta) = f$. 取 V 的标准正交基 $\{\alpha_1, \dots, \alpha_n\}$. 对于 $\beta = \sum_{j=1}^n c_j \alpha_j$,

$$\Phi(\beta) = f \iff \Phi(\beta)(\alpha_k) = f(\alpha_k), \forall k.$$

而

$$\Phi(\beta)(\alpha_k) = \langle \alpha_k, \beta \rangle = \langle \alpha_k, \sum_{j=1}^n c_j \alpha_j \rangle = \sum_{j=1}^n \bar{c}_j \langle \alpha_k, \alpha_j \rangle = \bar{c}_k.$$

所以

$$\Phi(\beta) = f \iff c_k = \overline{f(\alpha_k)}, \forall k \iff \beta = \sum_{j=1}^n \overline{f(\alpha_j)} \alpha_j.$$

因此满足 $\Phi(\beta) = f$ 的 β 存在唯一.

共轭线性: 对任意 α 有

$$\Phi(c\beta_1 + \beta_2)(\alpha) = \langle \alpha, c\beta_1 + \beta_2 \rangle = \bar{c}\langle \alpha, \beta_1 \rangle + \langle \alpha, \beta_2 \rangle = \bar{c}\Phi(\beta_1)(\alpha) + \Phi(\beta_2)(\alpha) = (\bar{c}\Phi(\beta_1) + \Phi(\beta_2))(\alpha).$$

因此 $\Phi(c\beta_1 + \beta_2) = \bar{c}\Phi(\beta_1) + \Phi(\beta_2)$. $\square$

注 8.4. • 在共轭线性的基础上, 为证即单又满, 只需证明 $\text{Ker}(\Phi) = 0$ (注意到 $\dim_{\mathbb{R}} V = \dim_{\mathbb{R}} V^*$). 而这是很容易的.

- $\{\Phi(\beta)\}^0 = \text{Ker}(\Phi(\beta)) = \beta^\perp$:

$$\text{Ker}(\Phi(\beta)) = \{\alpha \in V \mid \Phi(\beta)(\alpha) = 0\} = \{\alpha \in V \mid \langle \alpha, \beta \rangle = 0\} = \beta^\perp.$$

这推出对任意 $S \subset V$ 有

$$\Phi(S)^0 = \bigcap_{\beta \in S} \text{Ker}(\Phi(\beta)) = \bigcap_{\beta \in S} \beta^\perp = S^\perp.$$

还有

$$\Phi(S^\perp) = S^0.$$

实际上就是同一个空间有两个同构的对偶.

引理 8.15. 设 V 是有限维内积空间. 则对任意 $T \in L(V)$, 存在唯一的 $T^* \in L(V)$ 满足

$$\langle T\alpha, \beta \rangle = \langle \alpha, T^*\beta \rangle, \quad \forall \alpha, \beta \in V.$$

如果 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是有序标准正交基, 则 $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*$.

证明.

$$\begin{aligned} \langle T\alpha, \beta \rangle = \langle \alpha, T^*\beta \rangle &\iff \Phi(\beta)(T\alpha) = \Phi(T^*\beta)(\alpha) \iff \Phi(\beta) \circ T = \Phi(T^*\beta) \iff T^t(\Phi(\beta)) = \Phi(T^*\beta) \\ &\iff T^t \circ \Phi = \Phi \circ T^* \iff T^* = \Phi^{-1} \circ T^t \circ \Phi. \end{aligned}$$

容易验证 $\Phi^{-1} \circ T^t \circ \Phi$ 线性. 注意 T^* 是使得图表

$$\begin{array}{ccc} V^* & \xrightarrow{T^t} & V^* \\ \Phi \uparrow \cong & & \uparrow \cong \Phi \\ V & \xrightarrow{T^*} & V \end{array}$$

可交换的唯一映射.

如果 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是有序标准正交基, 则 $\Phi(\mathcal{B}) := \{\Phi\alpha_1, \dots, \Phi\alpha_n\}$ 是对偶基. 所以 $[T^t]_{\Phi(\mathcal{B})} = [T]_{\mathcal{B}}^t$. 因此, 只需证明 $[T^*]_{\mathcal{B}} = \overline{[T^t]_{\Phi(\mathcal{B})}}$. 当 $F = \mathbb{R}$ 时, 由于 Φ 是线性的并且图表可交换, 这是显然的. 当 $F = \mathbb{C}$ 时, 证明如下. 设 $[T^*]_{\mathcal{B}} = A$. 则

$$T^*\alpha_k = \sum_{j=1}^n A_{jk}\alpha_j.$$

两边取 Φ , 得

$$T^t\Phi(\alpha_k) = \Phi(T^*\alpha_k) = \sum_{j=1}^n \overline{A_{jk}}\Phi(\alpha_j),$$

即 $[T^t]_{\Phi(\mathcal{B})} = \overline{A}$. $\square$

T^* 称为 T 的(关于内积的)伴随变换.

注: 可以对一般的 $T \in L(V, W)$ 定义 T^* .

例 8.9. 在 $F^{n \times 1}$ 上的标准内积下, $L_A^* = L_{A^*}$. 只需验证 $\langle L_A X, Y \rangle = \langle X, L_{A^*} Y \rangle$, 即 $\langle AX, Y \rangle =$

$\langle X, A^*Y \rangle$. 而 $\langle X, Y \rangle = Y^*X$. 所以后者等价于 $Y^*(AX) = (A^*Y)^*X$. 这是显然的.

伴随变换有下面的性质:

引理 8.16. (1) $(T + U)^* = T^* + U^*$.

(2) $(cT)^* = \bar{c}T^*$.

(3) $(TU)^* = U^*T^*$.

(4) $(T^*)^* = T$.

证明. 可以利用 $T^* = \Phi^{-1} \circ T^t \circ \Phi$. 也可以直接验证. 例如, 为了直接验证(4), 只需验证 $\langle T^*\alpha, \beta \rangle = \langle \alpha, T\beta \rangle$. 这是显然的. $\square$

这些性质说明: $F = \mathbb{C}$ 时, 映射 $T \mapsto T^*$ 是 $\mathbb{C}$ -代数的 $L(V)$ 的共轭线性反自同构.

如果 $T^* = T$, 即

$$\langle T\alpha, \beta \rangle = \langle \alpha, T\beta \rangle, \quad \forall \alpha, \beta \in V,$$

则称 T 是**自伴的**.

引理 8.17. 设 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是内积空间 V 的标准正交基. 则 $T \in L(V)$ 自伴 $\iff [T]_{\mathcal{B}}$ Hermite.

证明. 由于 $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*$, 所以 $T^* = T \iff [T^*]_{\mathcal{B}} = [T]_{\mathcal{B}} \iff [T]_{\mathcal{B}}^* = [T]_{\mathcal{B}} \iff [T]_{\mathcal{B}}$ Hermite. $\square$

例 8.10. 设 W 是有限维内积空间 V 的子空间. 则 P_W 自伴. 事实上, 取 W 的有序标准正交集 $\{\alpha_1, \dots, \alpha_m\}$, 并扩充为 V 的有序标准正交基 $\{\alpha_1, \dots, \alpha_n\}$. 则 $[P]_{\mathcal{B}} = \text{diag}(I_m, 0)$ 是 Hermite 矩阵.

注意对任意 $T \in L(V)$, 存在唯一的自伴变换 $T_1, T_2 \in L(V)$ 满足 $T = T_1 + iT_2$. 事实上, $T_1 = \frac{1}{2}(T + T^*)$, $T_2 = \frac{1}{2i}(T - T^*)$.

§8.4 正交变换和酉变换

定义 8.5. 设 V, W 是内积空间. 如果 $T \in L(V, W)$ 满足 $\langle T\alpha, T\beta \rangle = \langle \alpha, \beta \rangle$, 则称 T **保内积**. 如果 T 保内积并且是线性空间的同构, 则称 T 是**内积空间的同构**. 此时 T^{-1} 也是内积空间的同构. 如果在 V 与 W 间存在内积空间的同构, 则称 V 与 W 作为内积空间**同构**.

命题 8.18. T 保内积 $\iff T$ 保长度, 即 $\|T\alpha\| = \|\alpha\|$, $\forall \alpha \in V$. 特别地, 保内积 $\implies$ 单.

证明. “ $\implies$ ” 显然.

“ $\impliedby$ ” 只需应用极化恒等式.

如果 T 保长度, 则 $\text{Ker}(T) = \{0\}$, 从而 T 单. $\square$

因此, 保内积变换也叫**等距变换**(isometry).

命题 8.19. 设 V, W 是有限维内积空间, $\dim V = \dim W$, $T \in L(V, W)$. TFAE:

(1) T 保内积.

(2) T 是内积空间的同构.

(3) T 把每个 V 的标准正交基映为 W 的标准正交基.

(4) T 把某个 V 的标准正交基映为 W 的标准正交基.

证明. “(1) $\implies$ (2)” T 保内积 $\implies T$ 单. 而 $\dim V = \dim W$, 所以 T 满.

“(2) $\implies$ (3)” 设 $\{\alpha_1, \dots, \alpha_n\}$ 是 V 的标准正交基. 则 $\{T\alpha_1, \dots, T\alpha_n\}$ 是 W 的基, 并且 $\langle T\alpha_j, T\alpha_k \rangle =$

$\langle \alpha_j, \alpha_k \rangle = \delta_{jk}$. 因此 $\{T\alpha_1, \dots, T\alpha_n\}$ 是标准正交基.

“(3) $\implies$ (4)” 显然.

“(4) $\implies$ (1)” 设 T 把 V 的标准正交基 $\{\alpha_1, \dots, \alpha_n\}$ 映为 W 的标准正交基 $\{T\alpha_1, \dots, T\alpha_n\}$. 则对 $\alpha, \beta \in V$, 设 $\alpha = \sum_{j=1}^n x_j \alpha_j$, $\beta = \sum_{k=1}^n y_k \alpha_k$. 则

$$\langle T\alpha, T\beta \rangle = \left\langle \sum_{j=1}^n x_j T\alpha_j, \sum_{k=1}^n y_k T\alpha_k \right\rangle = \sum_{j=1}^n x_j \bar{y}_j = \left\langle \sum_{j=1}^n x_j \alpha_j, \sum_{k=1}^n y_k \alpha_k \right\rangle = \langle \alpha, \beta \rangle.$$

□

推论 8.20. 设 V, W 是有限维内积空间. 则 V 与 W 作为内积空间同构 $\iff \dim V = \dim W$.

证明. “ $\implies$ ” 显然.

“ $\impliedby$ ” 取 V 的标准正交基 $\{\alpha_1, \dots, \alpha_n\}$ 和 W 的标准正交基 $\{\beta_1, \dots, \beta_n\}$. 取 $T \in L(V, W)$ 满足 $T\alpha_j = \beta_j$. 则 T 是内积空间的同构. □

例 8.11. 设 V 是有限维内积空间, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是有序标准正交基. 则坐标映射 $\Gamma_{\mathcal{B}} : V \rightarrow F^{n \times 1}$, $\alpha \mapsto [\alpha]_{\mathcal{B}}$ 是内积空间的同构(关于 $F^{n \times 1}$ 上的标准内积).

定义 8.6. • 实内积空间 V 作为内积空间的自同构称为**正交变换**. 记 V 上所有正交变换在映射复合下构成的群为 $O(V)$, 称为 V 上的**正交群**.

• 复内积空间 V 作为内积空间的自同构称为**酉变换**. 记 V 上所有酉变换在映射复合下构成的群为 $U(V)$, 称为 V 上的**酉群**.

命题 8.21. $\dim V < \infty$. 则对于 $T \in L(V)$, TFAE:

- (1) T 是内积空间的自同构.
- (2) T 保内积.
- (3) $T^*T = I$.

证明. “(1) $\iff$ (2)” 由上一命题显然.

“(2) $\iff$ (3)” 注意到总有 $\langle T\alpha, T\beta \rangle = \langle \alpha, T^*T\beta \rangle$. 因此

$$(2) \iff \langle \alpha, T^*T\beta \rangle = \langle \alpha, \beta \rangle \iff T^*T = I.$$

□

定义 8.7. 设 $A \in \mathbb{C}^{n \times n}$.

- 如果 $A^t A = I$, 则称 A 称为**正交矩阵**. 记所有 n 阶实正交矩阵在矩阵乘法下构成的群为 $O(n)$, 称为 n 阶**正交群**. 记所有 n 阶复正交矩阵在矩阵乘法下构成的群为 $O(n, \mathbb{C})$, 称为 n 阶**复正交群**.
- 如果 $A^* A = I$, 则称 A 称为**酉矩阵**. 记所有 n 阶酉矩阵在矩阵乘法下构成的群为 $U(n)$, 称为 n 阶**酉群**.

显然, 如果 A 实, 则 A 正交 $\iff A$ 酉.

命题 8.22. 设 V 是域 F 上的有限维内积空间, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是有序标准正交基, $T \in L(V)$.

- (1) 如果 $F = \mathbb{R}$, 则 T 正交 $\iff [T]_{\mathcal{B}}$ 正交.
- (2) 如果 $F = \mathbb{C}$, 则 T 酉 $\iff [T]_{\mathcal{B}}$ 酉.

证明. 只需证明 $T^*T = I \iff [T]_{\mathcal{B}}^* [T]_{\mathcal{B}} = I$. 而总有 $[T^*T]_{\mathcal{B}} = [T]_{\mathcal{B}}^* [T]_{\mathcal{B}}$. 因此命题成立. □

命题 8.23. 设 $A \in \mathbb{R}^{n \times n}$. TFAE:

- (1) A 正交.
- (2) A^t 正交.
- (3) A 的行向量构成 $\mathbb{R}^n$ 的标准正交基(关于标准内积).
- (4) A 的列向量构成 $\mathbb{R}^{n \times 1}$ 的标准正交基(关于标准内积).

证明. “(1) $\iff$ (4)”. 由于 $\mathbb{R}^{n \times 1}$ 的标准基关于标准内积是标准正交基, 并且 L_A 在标准基下的矩阵为 A , 所以 A 正交 $\iff L_A$ 正交. 而 L_A 把标准基映为 A 的列向量集合. 所以 L_A 正交 $\iff$ (4).

“(2) $\iff$ (3)”. 对 A^t 应用“(1) $\iff$ (4)”.

“(1) $\iff$ (2)”. 由定义, A 正交 $\iff A^t A = I$, A^t 正交 $\iff A A^t = I$. 因此两者都等价于: A 可逆并且 $A^{-1} = A^t$.

注 8.5. 直接证明“(3) $\iff$ (4)”并不容易!

类似地, 可以证明:

命题 8.24. 设 $A \in \mathbb{C}^{n \times n}$. TFAE:

- (1) A 酉.
- (2) A^* 酉.
- (3) A 的行向量构成 $\mathbb{C}^n$ 的标准正交基(关于标准内积).
- (4) A 的列向量构成 $\mathbb{C}^{n \times 1}$ 的标准正交基(关于标准内积).

注 8.6. 还有 A 酉 $\iff A^t$ 酉 $\iff \overline{A}$ 酉: $A^* A = I \iff \overline{A^*} \overline{A} = \overline{A^*} \overline{A} = (\overline{A})^* \overline{A} = I \iff \overline{A}$ 酉 $\iff (\overline{A})^* = A^t$ 酉.

容易看出, A 正交 $\implies \det(A) = \pm 1$, A 酉 $\implies |\det(A)| = 1$. 记

$$SO(n) = \{A \in O(n) \mid \det(A) = 1\},$$

$$SU(n) = \{A \in U(n) \mid \det(A) = 1\},$$

分别称为 n 阶特殊正交群和 n 阶特殊酉群. 对任意 $A_0 \in O(n) \setminus SO(n)$, 有 $O(n) = SO(n) \sqcup A_0 SO(n)$.

例如, 取 $A_0 = \text{diag}(1, \dots, 1, -1)$. 还有

$$U(n) = \{zA \mid A \in SU(n), z \in \mathbb{C}, |z| = 1\}.$$

例 8.12.

$$O(1) = \{1, -1\}.$$

$$U(1) = \{z \in \mathbb{C} \mid |z| = 1\} = \{e^{i\theta} \mid \theta \in \mathbb{R}\}.$$

$$SO(1) = SU(1) = \{1\}.$$

$$SO(2) = \left\{ \begin{bmatrix} x & -y \\ y & x \end{bmatrix} \mid x, y \in \mathbb{R}, x^2 + y^2 = 1 \right\} = \left\{ \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \mid \theta \in \mathbb{R} \right\}.$$

$$SU(2) = \left\{ \begin{bmatrix} z & -\bar{w} \\ w & \bar{z} \end{bmatrix} \mid z, w \in \mathbb{C}, |z|^2 + |w|^2 = 1 \right\}.$$

定理 8.25(QR分解或Iwasawa分解). 设 $F = \mathbb{R}$ 或 $\mathbb{C}$. 则对任意 $A \in GL_n(F)$, 存在唯一的 $A_k \in O(n)$ (若 $F = \mathbb{R}$) 或 $U(n)$ (若 $F = \mathbb{C}$), 对角元为正数的对角矩阵 A_a , 对角元为1的上三角矩阵 A_n 满足 $A = A_k A_a A_n$.

证明. 设 $A = [\beta_1, \dots, \beta_n]$. 由命题8.6(2), 存在唯一的 $F^{n \times 1}$ 的有序标准正交基 $\{\alpha_1, \dots, \alpha_n\}$, 对角元为正数的对角矩阵 A_a , 对角元为1的上三角矩阵 A_n 满足 $A = [\alpha_1, \dots, \alpha_n] A_a A_n$. 只需再注意到 $\{\alpha_1, \dots, \alpha_n\}$ 是标准正交基 $\iff A_k := [\alpha_1, \dots, \alpha_n] \in O(n)$ 或 $U(n)$. $\square$

现在考虑内积空间上的坐标变换. 设 V 是有限维内积空间, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 和 $\mathcal{B}' = \{\alpha'_1, \dots, \alpha'_n\}$ 是有序标准正交基. 则坐标映射 $\Gamma_{\mathcal{B}} : V \rightarrow F^{n \times 1}, \alpha \mapsto [\alpha]_{\mathcal{B}}$ 和 $\Gamma_{\mathcal{B}'} : V \rightarrow F^{n \times 1}, \alpha \mapsto [\alpha]_{\mathcal{B}'}$ 是内积空间的同构(关于 $F^{n \times 1}$ 上的标准内积). 我们知道, 存在唯一的 $P \in GL_n(F)$ 满足 $\Gamma_{\mathcal{B}} = L_P \circ \Gamma_{\mathcal{B}'}$, 即 $[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$. 则 L_P 是内积空间 $F^{n \times 1}$ 的同构, 即 P (作为 L_P 在标准基下的矩阵) 是正交矩阵(当 $F = \mathbb{R}$ 时)或酉矩阵(当 $F = \mathbb{C}$ 时). 我们还知道, 对任意 $T \in L(V)$, 有

$$[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P = P^*[T]_{\mathcal{B}}P.$$

定义 8.8. • 称 $A, B \in \mathbb{R}^{n \times n}$ **正交相似**, 如果存在 $P \in O(n)$ 满足 $B = P^{-1}AP$.

• 称 $A, B \in \mathbb{C}^{n \times n}$ **酉相似**, 如果存在 $P \in U(n)$ 满足 $B = P^{-1}AP$.

因此, 在 $F = \mathbb{C}$ 的情况, 对于 $T \in L(V)$, 取有序标准正交基 $\mathcal{B}$ 得到的矩阵 $[T]_{\mathcal{B}}$ 是酉相似的, 而取所有有序标准正交基 $\mathcal{B}$ 得到的矩阵 $[T]_{\mathcal{B}}$ 构成一个酉相似等价类. 对于 $F = \mathbb{R}$ 时情况类似. 因此, 我们可以提下面两个等价的问题:

- 对有限维内积空间 V 和 $T \in L(V)$, 寻找 V 的有序标准正交基 $\mathcal{B}$, 使得矩阵 $[T]_{\mathcal{B}}$ 尽量简单.
- 给定矩阵 $A \in \mathbb{C}^{n \times n}$ (或 $\mathbb{R}^{n \times n}$), 寻找 $P \in U(n)$ (或 $O(n)$), 使得矩阵 $P^{-1}AP$ 尽量简单.

最后, 为了做习题, 我们补充下面的概念:

定义 8.9. 设 V 是有限维内积空间. 称 $T \in L(V)$ 为**正定的**(positive definite)或**正的**(positive), 如果 T 自伴, 并且 $\langle T\alpha, \alpha \rangle > 0, \forall \alpha \in V \setminus \{0\}$.

容易看出:

- T 正 $\iff$ 函数 $(\alpha, \beta) \mapsto \langle T\alpha, \beta \rangle$ 是内积.
- 如果 $\mathcal{B}$ 是 V 的有序标准正交基, 则 T 正 $\iff [T]_{\mathcal{B}}$ 正定 Hermite.

§8.5 正规变换

本节回答上一节最后提出的问题. 我们先考虑在正交相似或酉相似意义下的可对角化问题.

定义 8.10. • 有限维内积空间 V 上的线性变换 T 如果与 T^* 可交换, 则称 T 为**正规变换**(normal).

• 矩阵 $A \in \mathbb{C}^{n \times n}$ 如果与 A^* 可交换, 则称 A 是**正规的**.

注 8.7. 若 T 正规, 设 $T_1 = \frac{T+T^*}{2}, T_2 = \frac{T-T^*}{2i}$. 则 T_1, T_2 自伴可交换, 并且 $T = T_1 + iT_2$. 因此 T_1, T_2 可酉对角化 $\implies T$ 可酉对角化.

自伴变换、“反自伴”变换(即 $T^* + T = 0$)和酉变换(正交变换)都是正规的. 容易看出, 如果 $\mathcal{B}$ 是 V 的有序标准正交基, 则 T 正规 $\iff [T]_{\mathcal{B}}$ 正规. 我们证明:

定理 8.26. 设 V 是 F 上的有限维内积空间, $T \in L(V)$.

- (1) 若 $F = \mathbb{R}$, 则存在 V 的有序标准正交基 $\mathcal{B}$ 使 $[T]_{\mathcal{B}}$ 对角 $\iff T$ 自伴.
- (2) 若 $F = \mathbb{C}$, 则存在 V 的有序标准正交基 $\mathcal{B}$ 使 $[T]_{\mathcal{B}}$ 对角 $\iff T$ 正规.

下面的推论是显然的.

推论 8.27. (1) $A \in \mathbb{R}^{n \times n}$ 正交相似于对角矩阵 $\iff A$ 对称.

(2) $A \in \mathbb{C}^{n \times n}$ 酉相似于对角矩阵 $\iff A$ 正规.

注 8.8. 实正规矩阵(例如正交矩阵)一般不相似于实对角矩阵. 可以证明: 任意实正规矩阵正交相似于分块对角矩阵, 其中每块的阶数至多为2.

定理8.26的“ $\implies$ ”部分的证明. 设 V 的有序标准正交基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}} = \text{diag}(c_1, \dots, c_n)$. 则 $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^* = \text{diag}(\bar{c}_1, \dots, \bar{c}_n)$. 如果 $F = \mathbb{R}$, 则每个 $c_j \in \mathbb{R}$, 于是 $[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}$, 从而 $T^* = T$, 即 T 自伴. 如果 $F = \mathbb{C}$, 则 $[T]_{\mathcal{B}}$ 与 $[T^*]_{\mathcal{B}}$ 可交换. 这推出 T 与 T^* 可交换, 即 T 正规. $\square$

为了证明定理8.26的“ $\impliedby$ ”部分, 先证明几个引理.

引理 8.28. 设 V 是有限维内积空间, $T \in L(V)$ 自伴. 则 $f_T \in \mathbb{R}[x]$, 并且在 $\mathbb{R}[x]$ 中完全分解为一次式的乘积. 特别地, $\sigma(T) \subset \mathbb{R}$.

证明. 设 $f_T = \prod_{j=1}^k (x - c_j)^{r_j}$, $c_j \in \mathbb{C}$. 只需证明 $c_j \in \mathbb{R}$. 取 V 的有序标准正交基 $\mathcal{B}$. 则 $A := [T]_{\mathcal{B}}$ Hermite. 取 $X \in \mathbb{C}^{n \times 1} \setminus \{0\}$ 满足 $AX = c_j X$. 则

$$c_j X^* X = X^* (AX) = (X^* A) X = (AX)^* X = (c_j X)^* X = \bar{c}_j X^* X.$$

由于 $X \neq 0$, 所以 $X^* X > 0$. 因此 $c_j = \bar{c}_j$, 即 $c_j \in \mathbb{R}$. $\square$

引理 8.29. 设 V 是有限维内积空间, $T \in L(V)$. 如果 $W \subset V$ 是 T -不变子空间, 则 $W^\perp$ 是 T^* -不变子空间.

证明. 设 $\alpha \in W^\perp$. 则对任意 $\beta \in W$, 有 $\langle T^* \alpha, \beta \rangle = \langle \alpha, T \beta \rangle = 0$. 因此 $T^* \alpha \in W^\perp$. 这就证明了 $W^\perp$ 是 T^* -不变子空间. $\square$

引理 8.30. 设 V 是有限维内积空间, $T \in L(V)$ 正规.

(1) 如果 $W \subset V$ 是 T -不变子空间, 则 $W^\perp$ 是 T -不变子空间, W 是 T^* -不变子空间.

(2) T 的特征子空间两两正交.

证明. (1) 分别取 W 的标准正交基 $\mathcal{B}_1$ 和 $W^\perp$ 的标准正交基 $\mathcal{B}_2$, 则 $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ 是 V 的标准正交基, 从而 $A := [T]_{\mathcal{B}}$ 正规. 由于 W 是 T -不变的, 所以 A 形如

$$A = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}.$$

计算得

$$AA^* = \begin{bmatrix} BB^* + CC^* & * \\ * & * \end{bmatrix}, \quad A^*A = \begin{bmatrix} B^*B & * \\ * & * \end{bmatrix}.$$

A 正规推出

$$BB^* + CC^* = B^*B.$$

两边取矩阵的迹, 得 $\text{tr}(CC^*) = 0$, 从而 $C = 0$. 因此 $W^\perp$ 是 T -不变子空间. 再由引理8.29, 即知 $W = (W^\perp)^\perp$ 是 T^* -不变子空间.

(2) 设 $c_1, c_2 \in \sigma(T)$, $c_1 \neq c_2$. 需要证明 $V_{c_1} \perp V_{c_2}$. 设 $\alpha \in V_{c_1}$, $\beta \in V_{c_2}$. 由(1), $F\beta$ 是 T^* -不变子空间, 即 $T^* \beta \in F\beta$, 并且

$$\langle T^* \beta, \beta \rangle = \langle \beta, T \beta \rangle = \bar{c}_2 \langle \beta, \beta \rangle.$$

因此 $T^* \beta = \bar{c}_2 \beta$. 这推出

$$c_1 \langle \alpha, \beta \rangle = \langle c_1 \alpha, \beta \rangle = \langle T \alpha, \beta \rangle = \langle \alpha, T^* \beta \rangle = \langle \alpha, \bar{c}_2 \beta \rangle = \bar{c}_2 \langle \alpha, \beta \rangle.$$

由于 $c_1 \neq c_2$, 所以 $\langle \alpha, \beta \rangle = 0$. $\square$

注 8.9. 对于自伴变换, 引理8.30(1)是引理8.29的直接推论, 引理8.30(2)的证明也更加简单(结合引理8.28).

定理8.26的“ $\Leftarrow$ ”部分的证明. 由引理8.30(1), T 半单. 再由推论7.30和引理8.28, 可知 T 可对角化. 设 $\sigma(T) = \{c_1, \dots, c_k\}$. 则有直和分解 $V = \bigoplus_{j=1}^k V_{c_j}$. 由引理8.30(2), 该直和分解中的直和项两两正交. 取 V_{c_j} 的有序标准正交基 $\mathcal{B}_j$. 则 $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$ 是 V 的有序标准正交基. 由于 $\mathcal{B}$ 中的向量都是 T 的特征向量, 所以 $[T]_{\mathcal{B}}$ 对角. $\square$

注 8.10. 为了得到直和分解 $V = \bigoplus_{j=1}^k V_{c_j}$, 也可以不用“半单 $\Rightarrow$ 可对角化”: 假设 $W := \bigoplus_{j=1}^k V_{c_j} \neq V$. 为了得到矛盾, 只需说明 $W^\perp$ 中存在 T 的特征向量. 首先, 由引理8.30(1), $W^\perp$ 是非零 T -不变子空间. 当 $F = \mathbb{C}$ 时, $W^\perp$ 中显然有 T 的特征向量. 当 $F = \mathbb{R}$ 时并且 T 自伴时, 容易看出 $T_{W^\perp}$ 也自伴, 于是由引理8.28, $W^\perp$ 中也存在 T 的特征向量. $\square$

推论 8.31. 设 V 是 F 上的有限维内积空间, $T \in L(V)$. 假设或者 $F = \mathbb{R}$ 并且 T 自伴, 或者 $F = \mathbb{C}$ 并且 T 正规. 则有正交直和分解 $V = \bigoplus_{c \in \sigma(T)} V_c$. $\square$

可以利用讲义中命题7.17来求正交矩阵或酉矩阵 P 使得 $P^{-1}AP$ 对角. 此时需要找每个特征子空间 V_c 的标准正交基.

在复内积空间的情况, 自伴变换、反自伴变换和酉变换是三种重要的正规变换. 它们可以利用谱集来刻画.

推论 8.32. 设 V 是有限维复内积空间, $T \in L(V)$ 正规. 则:

- (1) T 自伴 $\iff \sigma(T) \subset \mathbb{R}$.
- (2) T 反自伴 $\iff \sigma(T) \subset i\mathbb{R}$.
- (3) T 酉 $\iff \sigma(T) \subset \{z \in \mathbb{C} \mid |z| = 1\}$.

证明. 取 V 的有序标准正交基 $\mathcal{B}$ 使得 $A := [T]_{\mathcal{B}}$ 对角. 则 $\sigma(T)$ 是 A 的对角元的集合.

- (1) T 自伴 $\iff A$ Hermite $\iff A$ 实对角 $\iff \sigma(T) \subset \mathbb{R}$.
- (2) 可以类似(1)来证明, 也可以注意 T 自伴 $\iff iT$ 反自伴并利用(1).
- (3) T 酉 $\iff A$ 酉 $\iff A$ 的对角元的模都是1 $\iff \sigma(T) \subset \{z \in \mathbb{C} \mid |z| = 1\}$. $\square$

推论 8.33. 设 $A \in \mathbb{C}^{n \times n}$.

- (1) A Hermite $\iff A$ 酉相似于实对角矩阵.
- (2) A 反Hermite $\iff A$ 酉相似于纯虚对角矩阵.
- (3) A 酉 $\iff A$ 酉相似于对角元的模是1的对角矩阵.

接下来讨论一般矩阵在正交相似或酉相似下的标准形.

定理 8.34(Schur三角化定理). 设 V 是有限维复内积空间, $T \in L(V)$. 则存在 V 的有序标准正交基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}}$ 上三角.

证明. 由讲义的定理6.24, 存在 T -不变的全旗 $\{0\} = W_0 \subsetneq \dots \subsetneq W_n = V$. 取 V 的有序标准正交基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 使得 $W_j = \text{span}\{\alpha_1, \dots, \alpha_j\}$. 则 $[T]_{\mathcal{B}}$ 上三角(见定理6.24的证明中的“(4) $\implies$ (1)”). $\square$

推论 8.35. 任意 $A \in \mathbb{C}^{n \times n}$ 酉相似于上三角矩阵.

注 8.11. 对于实矩阵, 与Schur三角化定理直接类似的结果不成立. 可以证明: 任意 $A \in \mathbb{R}^{n \times n}$ 正

交相似于分块上三角矩阵, 其中每块的行数和列数至多为2.

定理8.26(2)的“ $\Leftarrow$ ”部分还可以从Schur三角化定理的角度重新理解(证明).

引理 8.36. 设 $A \in \mathbb{C}^{n \times n}$ 是上三角矩阵. 则 A 正规 $\iff A$ 对角.

证明. “ $\Leftarrow$ ”: 若 A 对角, 则 $A^* = \bar{A}$ 对角, 从而与 A 可交换. 因此 A 正规.

“ $\Rightarrow$ ”: A 正规 $\implies L_A$ 正规(关于 $\mathbb{C}^{n \times 1}$ 上的标准内积). 设 $\{\epsilon_1, \dots, \epsilon_n\}$ 是标准基. 由引理8.30(1),

A 上三角 $\implies \text{span}\{\epsilon_1, \dots, \epsilon_k\}$ 是 L_A -不变的, $1 \leq k \leq n-1$

$\implies \text{span}\{\epsilon_{k+1}, \dots, \epsilon_n\} = \text{span}\{\epsilon_1, \dots, \epsilon_k\}^\perp$ 是 L_A -不变的, $1 \leq k \leq n-1$

$\implies A$ 下三角.

因此 A 对角. □

定理8.26(2)的“ $\Leftarrow$ ”部分的另一证明. 假设 T 正规. 取 V 的有序标准正交基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}}$ 上三角. 则 $[T]_{\mathcal{B}}$ 正规, 从而对角. □

第九章 内积空间上的线性变换

§9.1 线性空间上的形式

设 $F = \mathbb{R}$ 或 $\mathbb{C}$, V 是 F -线性空间. 我们考虑:

- 当 $F = \mathbb{R}$ 时 V 上的双线性形式;
- 当 $F = \mathbb{C}$ 时 V 上的 $1\frac{1}{2}$ -线性形式.

后者的定义如下:

定义 9.1. 设 V 是复线性空间. V 上的函数 $f: V \times V \rightarrow \mathbb{C}$ 如果满足:

- (a) $f(c\alpha + \beta, \gamma) = cf(\alpha, \gamma) + f(\beta, \gamma)$,
- (b) $f(\alpha, c\beta + \gamma) = \bar{c}f(\alpha, \beta) + f(\alpha, \gamma)$,

则称 f 是 V 上的 $1\frac{1}{2}$ -线性函数 (sesqui-linear function) 或 $1\frac{1}{2}$ -线性形式 (sesqui-linear form).

我们把这两者都简称为 V 上的形式, 并把 V 上的所有形式的集合记为 $\text{Form}(V)$, 它在自然定义的运算下构成线性空间.

设 $\dim V < \infty$, $f \in \text{Form}(V)$, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是 V 的有序基. 称 (j, k) -元为 $f(\alpha_k, \alpha_j)$ 的 n 阶方阵为 f 在 $\mathcal{B}$ 下的矩阵, 记为 $[f]_{\mathcal{B}}$. 对于 $\alpha = \sum_k x_k \alpha_k, \beta = \sum_j y_j \alpha_j \in V$, 有

$$f(\alpha, \beta) = f\left(\sum_k x_k \alpha_k, \sum_j y_j \alpha_j\right) = \sum_{j,k} x_k \bar{y}_j f(\alpha_k, \alpha_j) = [\beta]_{\mathcal{B}}^* [f]_{\mathcal{B}} [\alpha]_{\mathcal{B}}.$$

特别地, $[f]_{\mathcal{B}}$ 的定义与内积的矩阵的定义是一致的.

例 9.1. 设 $A \in F^{n \times n}$. 定义 $f \in \text{Form}(F^{n \times 1})$ 为

$$f(X, Y) = Y^* A X.$$

设 $\mathcal{B}$ 是 $F^{n \times 1}$ 的标准有序基. 则 $[f]_{\mathcal{B}} = A$.

命题 9.1. 如果 $\mathcal{B}' = \{\alpha'_1, \dots, \alpha'_n\}$ 是另一组有序基, $[\alpha'_1, \dots, \alpha'_n] = [\alpha_1, \dots, \alpha_n]P$, 则

$$[f]_{\mathcal{B}'} = P^* [f]_{\mathcal{B}} P.$$

证明. 此时有 $[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$. 所以

$$[\beta]_{\mathcal{B}'}^* [f]_{\mathcal{B}'} [\alpha]_{\mathcal{B}'} = f(\alpha, \beta) = [\beta]_{\mathcal{B}}^* [f]_{\mathcal{B}} [\alpha]_{\mathcal{B}} = [\beta]_{\mathcal{B}'}^* P^* [f]_{\mathcal{B}} P [\alpha]_{\mathcal{B}'}, \quad \forall \alpha, \beta \in V.$$

因此 $[f]_{\mathcal{B}'} = P^* [f]_{\mathcal{B}} P$. □

定义 9.2. 设 $f \in \text{Form}(V)$.

- (1) 如果 $F = \mathbb{R}$ 并且 $f(\alpha, \beta) = f(\beta, \alpha)$, 则称 f 是对称的.
- (2) 如果 $F = \mathbb{C}$ 并且 $f(\alpha, \beta) = \overline{f(\beta, \alpha)}$, 则称 f 是 Hermite 的.

为了与书上统一, 我们把两种情况都称为 Hermite 的.

引理 9.2. 如果 $F = \mathbb{C}$, 则 f Hermite $\iff f(\alpha, \alpha) \in \mathbb{R}, \forall \alpha \in V$.

证明. “ $\implies$ ”: 显然. “ $\impliedby$ ”: 把 $f(\alpha + \beta, \alpha + \beta)$ 展开, 条件推出

$$f(\alpha, \beta) + f(\beta, \alpha) \in \mathbb{R}, \quad \forall \alpha, \beta \in V,$$

从而

$$f(\alpha, \beta) + f(\beta, \alpha) = \overline{f(\alpha, \beta)} + \overline{f(\beta, \alpha)}, \quad \forall \alpha, \beta \in V.$$

将此式中的 β 替换为 $i\beta$, 得

$$-if(\alpha, \beta) + if(\beta, \alpha) = i\overline{f(\alpha, \beta)} - i\overline{f(\beta, \alpha)}, \quad \forall \alpha, \beta \in V.$$

由此得

$$f(\alpha, \beta) = \overline{f(\beta, \alpha)}, \quad \forall \alpha, \beta \in V.$$

□

定义 9.3. 设 $f \in \text{Form}(V)$ 是Hermite的.

- 如果 $f(\alpha, \alpha) > 0, \forall \alpha \in V \setminus \{0\}$, 则称 f 是**正定的**(positive definite)或**正的**(positive).
- 如果 $f(\alpha, \alpha) \geq 0, \forall \alpha \in V$, 则称 f 是**半正定的**(positive semi-definite)或**非负的**(non-negative).

注意正定Hermite形式就是内积.

定义 9.4. 设 $A \in F^{n \times n}$ 是Hermite矩阵.

- 如果 $X^*AX > 0, \forall X \in F^{n \times 1} \setminus \{0\}$, 则称 A 是**正定的**或**正的**.
- 如果 $X^*AX \geq 0, \forall X \in F^{n \times 1}$, 则称 A 是**半正定的**或**非负的**.

命题 9.3. 设 $\dim V < \infty$, $\mathcal{B}$ 是有序基, 则 f Hermite $\iff [f]_{\mathcal{B}}$ Hermite. 此时, f 正定 $\iff [f]_{\mathcal{B}}$ 正定, f 半正定 $\iff [f]_{\mathcal{B}}$ 半正定.

证明. “ $\implies$ ”: 显然. “ $\impliedby$ ”: 只需注意到

$$\begin{aligned} f(\alpha, \beta) &= [\beta]_{\mathcal{B}}^* [f]_{\mathcal{B}} [\alpha]_{\mathcal{B}}, \\ \overline{f(\beta, \alpha)} &= ([\alpha]_{\mathcal{B}}^* [f]_{\mathcal{B}} [\beta]_{\mathcal{B}})^* = [\beta]_{\mathcal{B}}^* [f]_{\mathcal{B}}^* [\alpha]_{\mathcal{B}}. \end{aligned}$$

其余结论由

$$f(\alpha, \alpha) = [\alpha]_{\mathcal{B}}^* [f]_{\mathcal{B}} [\alpha]_{\mathcal{B}}$$

是显然的.

□

下面讨论正定矩阵. 容易看出:

- A 正定 $\implies A$ 可逆.
- A 正定, $P \in GL_n(F) \implies P^*AP$ 正定. 特别地, $P \in GL_n(F) \implies P^*P$ 正定.

命题 9.4(Cholesky分解). 设 $A \in F^{n \times n}$ 正定. 则存在唯一的对角元为正数的上三角矩阵 R 满足 $A = R^*R$.

证明. 考虑 $F^{n \times 1}$ 上的内积 $f(X, Y) = Y^*AX$ 和标准内积 $f_0(X, Y) = Y^*X$. 则对于 $R \in GL_n(F)$,

$$\begin{aligned} &\text{映射 } L_R: (F^{n \times 1}, f) \rightarrow (F^{n \times 1}, f_0) \text{ 是内积空间的同构} \\ &\iff f_0(RX, RY) = f(X, Y), \quad \forall X, Y \in F^{n \times 1} \\ &\iff Y^*R^*RX = Y^*AX, \quad \forall X, Y \in F^{n \times 1} \\ &\iff A = R^*R. \end{aligned}$$

取定 $P \in GL_n(F)$ 使得 $L_P: (F^{n \times 1}, f) \rightarrow (F^{n \times 1}, f_0)$ 是内积空间的同构. 则对于 $R \in GL_n(F)$,

$$\begin{aligned} &L_R: (F^{n \times 1}, f) \rightarrow (F^{n \times 1}, f_0) \text{ 是内积空间的同构} \\ &\iff L_R \circ L_P^{-1}: (F^{n \times 1}, f_0) \rightarrow (F^{n \times 1}, f_0) \text{ 是内积空间的同构} \\ &\iff RP^{-1} \in O(n) \text{ (或 } U(n)) \\ &\iff PR^{-1} \in O(n) \text{ (或 } U(n)). \end{aligned}$$

由QR分解, P 可唯一分解为 $O(n)$ (或 $U(n)$)中矩阵与对角元为正数的上三角矩阵的乘积, 即存在唯一的对角元为正数的上三角矩阵 R 使得 $PR^{-1} \in O(n)$ (或 $U(n)$). $\square$

推论 9.5. A 正定 $\implies \det(A) > 0$.

证明. 设 $A = P^*P$, P 可逆. 则 $\det(A) = |\det(P)|^2 > 0$. $\square$

下面给出一个判断Hermite矩阵何时正定的方法.

定义 9.5. 设 $A \in F^{n \times n}$, $1 \leq k \leq n$. 行列式

$$\Delta_k(A) := \det \begin{bmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{bmatrix}$$

称为 A 的第 k 个顺序主子式(leading principal minor).

定理 9.6. 设 $A \in F^{n \times n}$ Hermite. 则 A 正定 $\iff \Delta_k(A) > 0, k = 1, \dots, n$.

引理 9.7(LU分解). 设 F 是任意域. 则对 $A \in GL_n(F)$, TFAE:

- (1) 存在 $L, U \in GL_n(F)$, L 下三角, U 上三角, 满足 $A = LU$.
- (2) $\Delta_k(A) \neq 0, k = 1, \dots, n-1$.

此时, 如果还要求 L (或 U)的对角元都是1, 则满足(1)的 L, U 还是唯一的.

证明. “(1) $\implies$ (2)”: 把 L, U 写成分块矩阵 $L = \begin{bmatrix} L_k & 0 \\ * & * \end{bmatrix}, U = \begin{bmatrix} U_k & * \\ 0 & * \end{bmatrix}$, 其中 $L_k, U_k \in F^{k \times k}$.

则 $A = LU = \begin{bmatrix} L_k U_k & * \\ * & * \end{bmatrix}$. 因此

$$\Delta_k(A) = \det(L_k U_k) = \det(L_k) \det(U_k) \neq 0.$$

“(2) $\implies$ (1)和分解的唯一性”: 只需证明存在唯一的严格上三角矩阵(即对角元都为0) $N \in F^{n \times n}$ 使得 $A(N+I)$ 下三角, 即 $A(N+I)$ 的第 $k+1$ 列的前 k 行为0, $1 \leq k \leq n-1$. 设 $A = \begin{bmatrix} A_k & * \\ * & * \end{bmatrix}$, 其中 $A_k \in F^{k \times k}$. 设 A, N 的第 $k+1$ 列分别为 α_{k+1}, N_{k+1} , 它们的前 k 行分别为 $\alpha'_{k+1}, N'_{k+1} \in F^{k \times 1}$. 则 $N_{k+1} = \begin{bmatrix} N'_{k+1} \\ 0 \end{bmatrix}$. 于是 $A(N+I)$ 的第 $k+1$ 列为

$$AN_{k+1} + \alpha_{k+1} = \begin{bmatrix} A_k N'_{k+1} + \alpha'_{k+1} \\ * \end{bmatrix},$$

从而其前 k 行为 $A_k N'_{k+1} + \alpha'_{k+1}$. 因此只需证明: 对任意 $1 \leq k \leq n-1$, 存在唯一的 $N'_{k+1} \in F^{k \times 1}$ 满足

$$A_k N'_{k+1} + \alpha'_{k+1} = 0.$$

这在(2)的假设下是显然的. (实际上可求得 $N'_{k+1} = -A_k^{-1} \alpha'_{k+1}$.) $\square$

注 9.1. 集合

$$\{LU \mid L, U \in GL_n(F), L \text{ 下三角}, U \text{ 上三角}\}$$

称为 $GL_n(F)$ 中的一个Bruhat big cell. 它与群 $GL_n(F)$ 的Bruhat分解有关.

定理9.6的证明. “ $\implies$ ”: 设 A 正定. 我们已经知道 $\Delta_n(A) = \det(A) > 0$. 设 $1 \leq k \leq n-1$,

$A = \begin{bmatrix} A_k & * \\ * & * \end{bmatrix}$. 则 A_k Hermite, 并且对任意 $X \in F^{k \times 1} \setminus \{0\}$ 有

$$X^* A_k X = \begin{bmatrix} X \\ 0 \end{bmatrix}^* A \begin{bmatrix} X \\ 0 \end{bmatrix} > 0.$$

所以 A_k 正定, 因此 $\Delta_k(A) = \det(A_k) > 0$.

“ $\Leftarrow$ ”: 假设 $A \in F^{n \times n}$ Hermite 并且 $\Delta_k(A) > 0, k = 1, \dots, n$. 我们证明 A 正定. 由 LU 分解, 存在 $L, U \in GL_n(F)$, L 下三角, U 上三角且对角元都为1, 满足 $A = LU$. 记 $D = (U^*)^{-1}L$. 则 $A = U^*DU$, 从而 D Hermite. 注意到 D 还是下三角的, 从而 D 对角. 由

$$\begin{bmatrix} A_k & * \\ * & * \end{bmatrix} = \begin{bmatrix} U_k^* & 0 \\ * & * \end{bmatrix} \begin{bmatrix} D_k & 0 \\ 0 & * \end{bmatrix} \begin{bmatrix} U_k & * \\ 0 & * \end{bmatrix}$$

得 $A_k = U_k^* D_k U_k$. 因此

$$\Delta_k(A) = \det(A_k) = \det(U_k^* D_k U_k) = \det(D_k) = \Delta_k(D) > 0.$$

这推出 D 的对角元都 > 0 , 从而正定. 因此 $A = U^*DU$ 正定. □

§9.2 内积空间上的形式

本节假设 V 是 $F = \mathbb{R}$ 或 $\mathbb{C}$ 上的有限维内积空间. 此时, 取定的内积给出了 $L(V)$ 与 $\text{Form}(V)$ 的一一对应. 首先, 对任意 $T \in L(V)$, $f(\alpha, \beta) = \langle T\alpha, \beta \rangle$ 是形式. 这给出了线性映射

$$\mathcal{F} : L(V) \rightarrow \text{Form}(V), \quad \mathcal{F}(T)(\alpha, \beta) = \langle T\alpha, \beta \rangle.$$

另一方面, 对任意 $f \in \text{Form}(V)$, 考虑共轭线性映射

$$\Phi_f : V \rightarrow V^*, \quad \Phi_f(\alpha)(\beta) = \overline{f(\alpha, \beta)}.$$

回忆我们还有共轭线性同构

$$\Phi : V \rightarrow V^*, \quad \Phi(\alpha)(\beta) = \langle \beta, \alpha \rangle = \overline{\langle \alpha, \beta \rangle}.$$

则

$$T_f := \Phi^{-1} \circ \Phi_f$$

线性. 这给出了线性映射

$$\mathcal{T} : \text{Form}(V) \rightarrow L(V), \quad \mathcal{T}(f) = T_f = \Phi^{-1} \circ \Phi_f.$$

注意 T_f 满足性质

$$\langle T_f \alpha, \beta \rangle = \overline{\Phi(T_f \alpha)(\beta)} = \overline{\Phi_f(\alpha)(\beta)} = f(\alpha, \beta).$$

引理 9.8. $\mathcal{F}$ 与 $\mathcal{T}$ 互为逆映射.

证明. 首先,

$$(\mathcal{F} \circ \mathcal{T}(f))(\alpha, \beta) = \mathcal{F}(T_f)(\alpha, \beta) = \langle T_f \alpha, \beta \rangle = f(\alpha, \beta),$$

即 $\mathcal{F} \circ \mathcal{T}(f) = f$. 另一方面,

$$\langle (\mathcal{T} \circ \mathcal{F}(T))\alpha, \beta \rangle = \langle T_{\mathcal{F}(T)}\alpha, \beta \rangle = \mathcal{F}(T)(\alpha, \beta) = \langle T\alpha, \beta \rangle.$$

因此 $\mathcal{T} \circ \mathcal{F}(T) = T$. □

命题 9.9. 设 $f \in \text{Form}(V)$. 则 f Hermite $\iff T_f$ 自伴. 此时, f 正定 $\iff T_f$ 正定, f 半正定 $\iff T_f$ 半正定.

这里 T_f 正定或半正定指:

定义 9.6. 设 $T \in L(V)$.

- 如果 T 自伴并且 $\langle T\alpha, \alpha \rangle > 0, \forall \alpha \in V \setminus \{0\}$, 则称 T 是正定的或正的.
- 如果 T 自伴并且 $\langle T\alpha, \alpha \rangle \geq 0, \forall \alpha \in V$, 则称 T 是半正定的或非负的.

先证明:

引理 9.10. 设 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是 V 的有序标准正交基.

- (1) 设 $T \in L(V)$, $A = [T]_{\mathcal{B}}$. 则 $A_{kj} = \langle T\alpha_j, \alpha_k \rangle$.
- (2) 设 $f \in \text{Form}(V)$. 则 $[T_f]_{\mathcal{B}} = [f]_{\mathcal{B}}$.

证明. (1) 对任意 $\alpha \in V$ 有 $\alpha = \sum_k \langle \alpha, \alpha_k \rangle \alpha_k$. 因此

$$T\alpha_j = \sum_k \langle T\alpha_j, \alpha_k \rangle \alpha_k.$$

而 A 由

$$T\alpha_j = \sum_k A_{kj} \alpha_k$$

决定. 比较系数即得.

(2)

$$[f]_{\mathcal{B}} \text{ 的 } (j, k)\text{-元} = f(\alpha_k, \alpha_j) = \langle T_f \alpha_k, \alpha_j \rangle = [T_f]_{\mathcal{B}} \text{ 的 } (j, k)\text{-元}.$$

□

注 9.2. $\langle f, g \rangle = \text{tr}(T_f T_g^*)$ 定义了 $\text{Form}(V)$ 上的内积, 并且对任意 V 的标准正交基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$, 有

$$\langle f, g \rangle = \sum_{j,k} f(\alpha_j, \alpha_k) \overline{g(\alpha_j, \alpha_k)}.$$

事实上, 只需注意到 $(T, U) \mapsto \text{tr}(TU^*)$ 是 $L(V)$ 上的内积, 并且, 记 $A = [f]_{\mathcal{B}}$, $B = [g]_{\mathcal{B}}$, 则

$$\text{tr}(T_f T_g^*) = \text{tr}([T_f T_g^*]_{\mathcal{B}}) = \text{tr}(AB^*) = \sum_{j,k} A_{jk} \overline{B_{jk}} = \sum_{j,k} f(\alpha_j, \alpha_k) \overline{g(\alpha_j, \alpha_k)}.$$

命题9.9的证明. 取 V 的有序标准正交基 $\mathcal{B}$. 由引理,

$$f \text{ Hermite} \iff [f]_{\mathcal{B}} \text{ Hermite} \iff [T_f]_{\mathcal{B}} \text{ Hermite} \iff T_f \text{ 自伴}.$$

关于正定性和半正定性的结论由定义是显然的.

□

推论 9.11. 设 $F = \mathbb{C}$. 则 $T \in L(V)$ 自伴 $\iff \langle T\alpha, \alpha \rangle \in \mathbb{R}, \forall \alpha \in V$.

证明. 取 $f \in \text{Form}(V)$ 使 $T_f = T$. 则 T_f 自伴 $\iff f$ Hermite $\iff \langle T_f \alpha, \alpha \rangle = f(\alpha, \alpha) \in \mathbb{R}$.

□

推论 9.12. 设 $f \in \text{Form}(V)$.

- (1) (主轴定理) 如果 f Hermite, 则存在有序标准正交基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}}$ 实对角.
- (2) 如果 $F = \mathbb{C}$, 则存在有序标准正交基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}}$ 上三角.

证明. (1) f Hermite $\implies T_f$ 自伴. 因此可取 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}} = [T_f]_{\mathcal{B}}$ 实对角.

(2) 由Schur三角化定理, 可取 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}} = [T_f]_{\mathcal{B}}$ 上三角.

□

注 9.3. 在主轴定理中, 如果 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 并且 $[f]_{\mathcal{B}} = \text{diag}(c_1, \dots, c_n)$, 则

$$f\left(\sum_k x_k \alpha_k, \sum_j y_j \alpha_j\right) = \sum_j c_j x_j \bar{y}_j.$$

§9.3 谱分解

设 $F = \mathbb{R}$ 或 $\mathbb{C}$, V 是 F 上的有限维内积空间. 回忆:

- 对于子空间 $W \subset V$ 有 $V = W \oplus W^\perp$, 从而任意 $\alpha \in V$ 可唯一分解为 $\alpha = \beta + \gamma$, 其中 $\beta \in W$, $\gamma \in W^\perp$. 映射 $P_W \in L(V)$, $P_W(\alpha) = \beta$ 称为到 W 上的正交投影.
- 如果直和分解 $V = \bigoplus_{i=1}^k W_i$ 中的直和项 W_i 两两正交, 则称它为正交直和分解. 如果 $F = \mathbb{R}$ 并且 T 自伴, 或者 $F = \mathbb{C}$ 并且 T 正规, 则有正交直和分解 $V = \bigoplus_{c \in \sigma(T)} V_c$.

定理 9.13(正规变换的谱分解). 设 $T \in L(V)$. 假设当 $F = \mathbb{R}$ 时假设 T 自伴, 当 $F = \mathbb{C}$ 时假设 T 正规. 设 $\sigma(T) = \{c_1, \dots, c_k\}$, P_i 为到 V_{c_i} 上的正交投影. 则

$$T = \sum_{i=1}^k c_i P_i,$$

并且对任意 $f \in F[x]$ 有

$$f(T) = \sum_{i=1}^k f(c_i) P_i.$$

证明. 取 V_{c_i} 的有序标准正交基 $\mathcal{B}_i$. 则 $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$ 是 V 的有序标准正交基, 并且 $[T]_{\mathcal{B}} = \text{diag}(c_1 I_{d_1}, \dots, c_k I_{d_k})$, 其中 $d_i = \dim V_{c_i}$. 另一方面, 任意 $\alpha \in V$ 可唯一分解为 $\alpha = \sum_{i=1}^k \alpha_i$, 其中 $\alpha_i \in V_{c_i}$. 注意到 $V_{c_i}^\perp = \bigoplus_{j \neq i} V_{c_j}$, 从而 $P_i \alpha = \alpha_i$. 这说明 $[P_i]_{\mathcal{B}} = \text{diag}(\delta_{i1} I_{d_1}, \dots, \delta_{ik} I_{d_k})$. 因此

$$[f(T)]_{\mathcal{B}} = \text{diag}(f(c_1) I_{d_1}, \dots, f(c_k) I_{d_k}) = \sum_{i=1}^k f(c_i) [P_i]_{\mathcal{B}} = \left[\sum_{i=1}^k f(c_i) P_i \right]_{\mathcal{B}},$$

即 $f(T) = \sum_{i=1}^k f(c_i) P_i$. 取 $f = x$ 即得 $T = \sum_{i=1}^k c_i P_i$. □

推论 9.14. 每个 P_i 为 T 的多项式.

证明. 取 $f_i \in F[x]$ 使 $f_i(c_j) = \delta_{ij}$. 则 $P_i = f_i(T)$. □

谱分解定理还可以用来定义 T 的其他函数: 如果 F 的子集 S 包含 $\sigma(T)$, $\phi: S \rightarrow F$ 是任意函数, 则可以定义

$$\phi(T) = \sum_{i=1}^k \phi(c_i) P_i.$$

例如, 如果 T 半正定, 则 $\sigma(T) \subset [0, +\infty)$, 从而可以定义

$$\sqrt{T} = \sum_{i=1}^k \sqrt{c_i} P_i.$$

命题 9.15. (1) $\phi(T)$ 可对角化, 并且 $\sigma(\phi(T)) = \phi(\sigma(T))$.

(2) 当 $F = \mathbb{R}$ 时, T 自伴 $\implies \phi(T)$ 自伴. 当 $F = \mathbb{C}$ 时, T 正规 $\implies \phi(T)$ 正规.

证明. 利用上面定理证明中的记号, 显然有 $[\phi(T)]_{\mathcal{B}} = \text{diag}(\phi(c_1) I_{d_1}, \dots, \phi(c_k) I_{d_k})$. □

对于矩阵, 假设 $A \in F^{n \times n}$ 可对角化, $A = P \text{diag}(a_1, \dots, a_n) P^{-1}$, 则可以定义

$$\phi(A) = P \text{diag}(\phi(a_1), \dots, \phi(a_n)) P^{-1}.$$

注意定义与 $a_1, \dots, a_n$ 的排序和 P 的选取无关: 取多项式 f 满足 $f(a_i) = \phi(a_i)$, 则 $\phi(A) = f(A)$.

作为应用, 我们证明:

定理 9.16. 设 $T \in L(V)$ 半正定. 则存在唯一的半正定 $\sqrt{T} \in L(V)$ 满足 $(\sqrt{T})^2 = T$. 此时, T 正定 $\iff \sqrt{T}$ 正定.

证明. 存在性: 设 $T = \sum_{i=1}^k c_i P_i$ 是谱分解. 则 $c_i \geq 0$. 定义

$$\sqrt{T} = \sum_{i=1}^k \sqrt{c_i} P_i.$$

则 $\sqrt{T}$ 半正定, 并且 $(\sqrt{T})^2 = \sum_{i=1}^k (\sqrt{c_i})^2 P_i = T$ (注意 $P_i^2 = P_i$ 并且当 $i \neq j$ 时有 $P_i P_j = 0$). 容易看出 T 正定 $\iff \sqrt{T}$ 正定.

唯一性: 设 N, N' 半正定并且 $N^2 = N'^2 = T$. 考虑谱分解 $N = \sum_{i=1}^k c_i P_i$, $N' = \sum_{i=1}^{k'} c'_i P'_i$, 其中 $c_i, c'_i \geq 0$. 则 $\sum_{i=1}^k c_i^2 P_i = T = \sum_{i=1}^{k'} (c'_i)^2 P'_i$. 从而 $\{c_1^2, \dots, c_k^2\} = \sigma(T) = \{(c'_1)^2, \dots, (c'_{k'})^2\}$. 这推出 $k' = k$ 并且重新排序后有 $c_i = c'_i$. 现在在 $\sum_{i=1}^k c_i^2 P_i = T = \sum_{i=1}^k (c_i)^2 P'_i$. 这推出 $\text{Im}(P_i) = \text{Ker}(T - c_i^2 \text{id}) = \text{Im}(P'_i)$. 从而 $P_i = P'_i$. 因此 $N = N'$. $\square$

注意到对任意 $T \in L(V)$, T^*T 半正定. 因此可以考虑 $N := \sqrt{T^*T}$. N 的特征值称为 T 的奇异值(singular value), 在很多方面有重要的应用.

引理 9.17. 设 $T \in L(V)$, $N = \sqrt{T^*T}$. 则

$$\|T\alpha\| = \|N\alpha\|, \quad \forall \alpha \in V.$$

特别地, $\text{Ker}(T) = \text{Ker}(N)$.

证明.

$$\|N\alpha\|^2 = \langle N\alpha, N\alpha \rangle = \langle N^2\alpha, \alpha \rangle = \langle T^*T\alpha, \alpha \rangle = \langle T\alpha, T\alpha \rangle = \|T\alpha\|^2.$$

$\square$

定理 9.18(极分解或Cartan分解). 设 $T \in L(V)$. 则存在酉变换 U (当 $F = \mathbb{R}$ 时为正交变换) 和唯一的半正定变换 N 满足 $T = UN$. T 可逆 $\iff N$ 正定, 此时 U 也唯一.

证明. 唯一性: 设 $T = UN$, U 酉, N 半正定. 则 $T^* = N^*U^* = NU^*$. 于是 $T^*T = (NU^*)(UN) = N^2$. 注意到 T^*T 半正定. 所以 $N = \sqrt{T^*T}$ 唯一. 显然有 T 可逆 $\iff N$ 可逆 $\iff N$ 正定. 此时 $U = TN^{-1}$ 也唯一.

存在性: 设 $N = \sqrt{T^*T}$. 由于 $\text{Ker}(T) = \text{Ker}(N)$, 所以存在 $U_1 \in L(\text{Im}(N), \text{Im}(T))$ 满足 $T = U_1 N$. 由于 $\|U_1 N\alpha\| = \|T\alpha\| = \|N\alpha\|$, 所以对任意 $\beta \in \text{Im}(N)$ 有 $\|U_1 \beta\| = \|\beta\|$, 即 U_1 保长度, 从而保内积. 再注意到 $\dim \text{Im}(N) = \dim \text{Im}(T)$, 所以 U_1 是内积空间的同构. 再任取内积空间的同构 $U_2 : \text{Im}(N)^\perp \rightarrow \text{Im}(T)^\perp$. 则 $U := U_1 \oplus U_2 : V \rightarrow V$ 也是内积空间的同构(可以取 $\text{Im}(N)$ 和 $\text{Im}(N)^\perp$ 的标准正交基来看), 从而酉. $\square$

推论 9.19(奇异值分解). 设 $A \in F^{n \times n}$. 则存在非负对角矩阵 D 和酉矩阵(当 $F = \mathbb{R}$ 时为正交矩阵) U_1, U_2 满足 $A = U_1 D U_2$.

证明. 由极分解, 存在酉矩阵 U 和半正定矩阵 N 满足 $A = UN$. 设 $N = V D V^{-1}$, 其中 V 酉, D 非负对角. 则 $A = U V D V^{-1}$. 取 $U_1 = UV$, $U_2 = V^{-1}$ 即可. $\square$

推论 9.20. 设 $T \in L(V)$ 可逆. 则 T 把 V 的某个正交基映为正交基, 即存在正交基 $\{\alpha_1, \dots, \alpha_n\}$ 使得 $\{T\alpha_1, \dots, T\alpha_n\}$ 也是正交基.

证明. 设 $T = UN$, U 酉, N 正定. 取有序标准正交基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 使得 $[N]_{\mathcal{B}}$ 正对角, 即 $N\alpha_i = c_i \alpha_i$, $c_i > 0$. 则 $T\alpha_i = UN\alpha_i = c_i U\alpha_i$. 显然 $\{U\alpha_1, \dots, U\alpha_n\}$ 是标准正交基. 所以 $\{c_1 U\alpha_1, \dots, c_n U\alpha_n\}$ 是正交基. $\square$

§9.4 正规变换的进一步性质

本节的主要目的之一是讨论实正规变换的性质. 我们将证明:

定理 9.21. 设 V 是有限维实内积空间, $T \in L(V)$ 正规. 则存在 V 的有序标准正交基 $\mathcal{B}$ 使得

$$[T]_{\mathcal{B}} = \text{diag} \left(a_1, \dots, a_l, r_1 \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{bmatrix}, \dots, r_m \begin{bmatrix} \cos \theta_m & -\sin \theta_m \\ \sin \theta_m & \cos \theta_m \end{bmatrix} \right),$$

其中 $a_i \in \mathbb{R}$, $r_j > 0$, $\theta_j \in (0, \pi)$.

容易看出, T 正交 $\iff a_i = \pm 1$ 并且 $r_j = 1$. 因此得到:

推论 9.22. 设 V 是有限维实内积空间, $T \in L(V)$ 正交. 则存在 V 的有序标准正交基 $\mathcal{B}$ 使得

$$[T]_{\mathcal{B}} = \text{diag} \left(I_{l_1}, -I_{l_2}, \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{bmatrix}, \dots, \begin{bmatrix} \cos \theta_m & -\sin \theta_m \\ \sin \theta_m & \cos \theta_m \end{bmatrix} \right),$$

其中 $\theta_j \in (0, \pi)$.

下面的推论也是显然的.

推论 9.23. 设 $A \in \mathbb{R}^{n \times n}$.

(1) 如果 A 正规, 则 A 正交相似于定理9.21中的矩阵.

(2) 如果 A 正交, 则 A 正交相似于推论9.22中的矩阵. □

我们给出定理9.21的三个证明, 分别利用正规变换的半单性、准素循环分解和比较特征多项式. 前两个证明本质上是等价的. 但是, 为了介绍证明过程中一些有意思的中间结果, 我们给出每个证明的细节. 实际上, 我们不只满足于中间结果对证明定理有用, 而是给出更一般的形式. 在本节中, 如果未做其他说明, 我们总是假设 $F = \mathbb{R}$ 或 $\mathbb{C}$, V 是 F 上的有限维内积空间.

第一个证明利用正规变换的半单性. 我们先证明:

引理 9.24. 设 $T \in L(V)$ 正规, $W \subset V$ 是不变子空间. 则 T_W 正规.

证明. 我们知道, W 也是 T^* -不变子空间. 对任意 $\alpha, \beta \in W$, 有

$$\langle (T_W)^* \alpha, \beta \rangle = \langle \alpha, T_W \beta \rangle = \langle \alpha, T \beta \rangle = \langle T^* \alpha, \beta \rangle.$$

由于 $T^* \alpha \in W$, 所以 $(T_W)^* \alpha = T^* \alpha$ 对任意 $\alpha \in W$ 成立, 即 $(T_W)^* = (T^*)_W$. 这推出

$$T_W (T_W)^* = T_W (T^*)_W = (T T^*)_W = (T^* T)_W = (T^*)_W T_W = (T_W)^* T_W,$$

所以 T_W 正规. □

引理 9.25. 设 $T \in L(V)$ 正规. 则 V 可分解为不变子空间的正交直和 $V = \bigoplus_{i=1}^k V_i$ 使得每个 T_{V_i} 单正规.

证明. 不妨假设 T 不单, 即存在非平凡不变子空间 $W \subset V$. 由于 T 正规, $W^\perp$ 也是不变子空间. 由引理9.24, T_W 和 $T_{W^\perp}$ 正规. 由于 $\dim W, \dim W^\perp < \dim V$, 对 $\dim V$ 用归纳法, 可以假设引理对 T_W 和 $T_{W^\perp}$ 成立. 这推出引理对 T 也成立. □

由引理9.25, 我们只需考察单正规变换.

引理 9.26. 设 V 是有限维实内积空间, $T \in L(V)$ 单正规. 假设 $\dim V > 1$. 则 $\dim V = 2$, 并且存在 $r > 0$, $\theta \in (0, \pi)$ 和 V 的有序标准正交基 $\mathcal{B}$ 使得

$$[T]_{\mathcal{B}} = r \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

其中 r 和 θ 被 T 唯一决定.

证明. 由 T 单可知 f_T 是 $\mathbb{R}[x]$ 中的素多项式. 而 $\dim V > 1$ 推出 $\deg f_T > 1$. 因此 $f_T = (x - c)(x - \bar{c})$, 其中 $c \in \mathbb{C} \setminus \mathbb{R}$. 特别地, $\dim V = \deg f_T = 2$. 取 V 的有序标准正交基 $\mathcal{B} = \{\alpha_1, \alpha_2\}$. 则 $[T]_{\mathcal{B}}$ 正规, 从而酉相似于 $\text{diag}(c, \bar{c})$. 记 $r = |c| > 0$. 则 $A := r^{-1}[T]_{\mathcal{B}}$ 酉相似于酉矩阵 $\text{diag}(r^{-1}c, r^{-1}\bar{c})$, 从而 A 也是酉矩阵. 而 A 是实矩阵, 所以 A 正交. 此外还有 $\det(A) = (r^{-1}c)(r^{-1}\bar{c}) = 1$. 因此只能有 $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$, 其中 $\theta \in [0, 2\pi)$. 注意到 T 单 $\implies T$ 无特征向量 $\implies \langle T\alpha_1, \alpha_2 \rangle \neq 0$. 必要时用 $-\alpha_2$ 替换 α_2 , 不妨设 $\langle T\alpha_1, \alpha_2 \rangle > 0$. 而 $r \sin \theta = [T]_{\mathcal{B}}$ 的 $(2, 1)$ -元 $= \langle T\alpha_1, \alpha_2 \rangle$. 所以 $\sin \theta > 0$, 即 $\theta \in (0, \pi)$. 另外, $r > 0$ 和 $\theta \in (0, \pi)$ 在各自的取值范围内被矩阵 rA 的特征多项式唯一决定, 因此被 T 唯一决定. $\square$

定理9.21的第一个证明. 由引理9.25, 存在正交直和分解 $V = \bigoplus_{i=1}^k V_i$ 使得每个 $T|_{V_i}$ 单正规. 由引理9.26, 不妨设当 $1 \leq i \leq l$ 时 $\dim V_i = 1$, 当 $1 \leq j \leq m$ 时 $\dim V_{l+j} = 2$, 其中 $l + m = k$. 对 $1 \leq i \leq l$, 任取 V_i 的有序标准正交基 $\mathcal{B}_i$. 对 $1 \leq j \leq m$, 取 V_{l+j} 的有序标准正交基 $\mathcal{B}_{l+j}$ 使得

$$[T|_{V_{l+j}}]_{\mathcal{B}_{l+j}} = r_j \begin{bmatrix} \cos \theta_j & -\sin \theta_j \\ \sin \theta_j & \cos \theta_j \end{bmatrix},$$

其中 $r_j > 0$, $\theta_j \in (0, \pi)$. 则 V 的有序标准正交基 $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$ 满足定理的要求. $\square$

第二个证明利用准素循环分解. 我们先证明一般线性变换的几个性质.

命题 9.27. 设 $T \in L(V)$. 则 $\text{Ker}(T)^\perp = \text{Im}(T^*)$, $\text{Im}(T)^\perp = \text{Ker}(T^*)$.

证明.

$$\begin{aligned} \alpha \in \text{Ker}(T) &\iff \langle T\alpha, \beta \rangle = 0, \forall \beta \in V \\ &\iff \langle \alpha, T^*\beta \rangle = 0, \forall \beta \in V \\ &\iff \alpha \in \text{Im}(T^*)^\perp. \end{aligned}$$

所以 $\text{Ker}(T) = \text{Im}(T^*)^\perp$. 两边取正交补即得第一个结论, 用 T^* 替换 T 即得第二个结论. $\square$

注 9.4. 上一命题可以与一般域上有限维线性空间的结论 $\text{Ker}(T)^0 = \text{Im}(T^t)$, $\text{Im}(T)^0 = \text{Ker}(T^t)$ 作比较.

命题 9.28. 设 $T \in L(V)$. 则 $\sigma(T^*) = \overline{\sigma(T)}$, 并且对 $c \in \sigma(T)$ 有 $\dim \text{Ker}(T^* - \bar{c}I) = \dim \text{Ker}(T - cI)$.

证明. 对 $c \in F$ 有 $(T - cI)^* = T^* - \bar{c}I$. 所以 $\text{Ker}(T - cI)^\perp = \text{Im}(T^* - \bar{c}I)$. 因此

$$\dim \text{Ker}(T - cI) = n - \dim \text{Ker}(T - cI)^\perp = n - \dim \text{Im}(T^* - \bar{c}I) = \dim \text{Ker}(T^* - \bar{c}I).$$

这推出

$$c \in \sigma(T) \iff \dim \text{Ker}(T - cI) \neq 0 \iff \dim \text{Ker}(T^* - \bar{c}I) \neq 0 \iff \bar{c} \in \sigma(T^*).$$

因此 $\sigma(T^*) = \overline{\sigma(T)}$. $\square$

下面讨论正规变换.

命题 9.29. 设 $T \in L(V)$ 正规, 则对任意 $c \in F$ 有 $\text{Ker}(T - cI) = \text{Ker}(T^* - \bar{c}I)$. 特别地, $\text{Ker}(T^*) = \text{Ker}(T)$, 从而 $\text{Ker}(T) = \text{Im}(T)^\perp$.

证明. 记 $W = \text{Ker}(T - cI)$. 则 $(T^*)_W = (T_W)^* = (c \text{id}_W)^* = \bar{c} \text{id}_W$. 这说明 $\text{Ker}(T - cI) \subset \text{Ker}(T^* - \bar{c}I)$. 由维数相等即得等式. (也可以把 T 和 c 分别替换为 T^* 和 $\bar{c}$ 看出反过来的包含关系.)

$\square$

注 9.5. 更直接的证法是利用 $\|T\alpha\| = \|T^*\alpha\|$:

$$\|T\alpha\|^2 = \langle T\alpha, T\alpha \rangle = \langle \alpha, T^*T\alpha \rangle = \langle \alpha, TT^*\alpha \rangle = \langle T^*\alpha, T^*\alpha \rangle = \|T^*\alpha\|^2.$$

注 9.6. 书上有一引理说: 设 $T \in L(V)$ 正规, 则 $T^2\alpha = 0 \implies T\alpha = 0$. 证明如下: 由 $\text{Ker}(T) \perp \text{Im}(T)$ 知 $T\alpha \in \text{Ker}(T) \cap \text{Im}(T) = \{0\}$. 事实上这对任何半单的 T 都成立: $T^2\alpha = 0 \implies p_\alpha | x^2$. 但 T 半单推出 p_T 无平方因子, 从而 p_α 无平方因子. 因此 $p_\alpha | x$, 即 $T\alpha = 0$.

我们知道正规变换的特征子空间互相正交. 这可以推广如下:

命题 9.30. 设 $T \in L(V)$ 正规, $f, g \in F[x]$ 互素. 则 $\text{Ker}(f(T))$ 与 $\text{Ker}(g(T))$ 正交.

证明. 记 $W = \text{Ker}(f(T))$, 它是 $g(T)$ 的不变子空间. 由于存在 $a, b \in F[x]$ 满足 $af + bg = 1$, 所以 $\text{id}_W = a(T)_W f(T)_W + b(T)_W g(T)_W = b(T)_W g(T)_W$, 于是 $g(T)_W$ 可逆. 因此 $W \subset \text{Im}(g(T))$. 容易看出 $g(T)$ 正规 (利用 $g(T)^* = \bar{g}(T^*)$). 所以 $\text{Im}(g(T)) \perp \text{Ker}(g(T))$. 综合起来即得 $W \perp \text{Ker}(g(T))$. $\square$

注 9.7. 类似可以证明, 对任意 F , 任意 $T \in L(V)$, 如果 $f, g \in F[x]$ 互素, 则 $\text{Ker}(f(T)) \cap \text{Ker}(g(T)) = \{0\}$.

命题 9.31. 设 $T \in L(V)$ 正规.

- (1) T 的准素分解是正交直和分解.
- (2) T 的循环分解可以取为正交直和分解.
- (3) T 的准素循环分解可以取为正交直和分解.

证明. (1) 由上一命题显然.

(2) 对 $\dim V$ 作归纳. 取 $\alpha_1 \in V$ 使 $p_{\alpha_1} = p_T$. 则 $(R\alpha_1)^\perp$ 是不变子空间. 对 $T_{(R\alpha_1)^\perp}$ 用归纳假设即可.

(3) 由(1)和(2), 显然. (先做准素分解, 再对每个准素分量做正交循环分解.) $\square$

引理 9.32. 设 V 是有限维实内积空间, $T \in L(V)$ 正规准素循环. 假设 $\dim V > 1$. 则 $\dim V = 2$, 并且存在 $r > 0, \theta \in (0, \pi)$ 和 V 的有序标准正交基 $\mathcal{B}$ 使得

$$[T]_{\mathcal{B}} = r \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

证明. 由引理 9.26, 只需证明 T 单. 由 T 准素循环可知 $f_T = p_T$ 为 $\mathbb{R}[x]$ 中素多项式的幂. 而 T 正规 $\implies T$ 半单, 所以 p_T 在 $\mathbb{R}[x]$ 中无平方因子. 因此 f_T 是 $\mathbb{R}[x]$ 中的素多项式, 从而 T 单. $\square$

定理 9.21 的第二个证明. 做正交的准素循环分解 $V = \bigoplus_{i=1}^k V_i$. 则每个 T_{V_i} 正规. 其余类似于第一个证明. $\square$

下面为定理 9.21 的第三个证明做准备. 回忆两个实方阵如果在 $\mathbb{C}$ 上相似, 则它们在 $\mathbb{R}$ 上相似. 酉相似与正交相似也有类似的关系.

定理 9.33. 如果 $A, B \in \mathbb{R}^{n \times n}$ 酉相似, 则它们正交相似.

先证明:

引理 9.34 (酉矩阵的 QS 分解). 对任意 $U \in U(n)$, 存在 $Q \in O(n)$ 和 $S \in U(n)$ 满足:

- $U = QS$,
- 存在 $f \in \mathbb{C}[x]$ 满足 $S = f(U^t U)$.

注意 S 是对称酉矩阵. 引理的结果和证明可以与极分解做比较.

证明. 设 $\sigma(U^t U) = \{c_1, \dots, c_k\}$. 取 $f \in \mathbb{C}[x]$ 满足 $f(c_i)^2 = c_i$. 设 $S = f(U^t U)$. 我们断言 $S \in U(n)$, 并且 $S^2 = U^t U$. 事实上, 由 $|c_i| = 1$ 推出 $|f(c_i)| = 1$. 如果

$$U^t U = P \operatorname{diag}(c_1 I_{d_1}, \dots, c_k I_{d_k}) P^{-1},$$

其中 $P \in U(n)$, 则

$$S = P \operatorname{diag}(f(c_1) I_{d_1}, \dots, f(c_k) I_{d_k}) P^{-1} \in U(n),$$

并且

$$S^2 = P \operatorname{diag}(f(c_1)^2 I_{d_1}, \dots, f(c_k)^2 I_{d_k}) P^{-1} = U^t U.$$

设 $Q = US^{-1} \in U(n)$. 注意到 $S = f(U^t U)$ 推出 $S^t = S$. 所以

$$\overline{Q} Q^{-1} = \overline{U} S^t S U^{-1} = (U^t)^{-1} S^2 U^{-1} = (U^t)^{-1} (U^t U) U^{-1} = I,$$

即 $\overline{Q} = Q$, 从而 $Q \in O(n)$. □

定理9.33的证明. 设 $B = UAU^{-1}$, $U \in U(n)$. 两边取复共轭, 得 $B = \overline{U} A U^t$. 由 $UAU^{-1} = \overline{U} A U^t$ 得 A 与 $U^t U$ 可交换. 由引理, 可设 $U = QS$, 其中 $Q \in O(n)$, $S = f(U^t U) \in U(n)$, $f \in \mathbb{C}[x]$. 则 A 与 S 可交换. 从而

$$B = UAU^{-1} = Q S A S^{-1} Q^{-1} = Q A Q^{-1}.$$

因此 A 与 B 正交相似. □

注意书上P356的推论(即正规矩阵酉相似于其有理标准形)是错的(这将推出Hermite矩阵的有理标准形是Hermite的, 这明显不对). 但它下面的两个结论是对的(即有理标准形相同的复(或实)正规矩阵酉(或正交)相似), 还可加强如下:

命题 9.35. 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $A, B \in F^{n \times n}$ 正规. TFAE:

- (1) A 与 B 酉相似($F = \mathbb{C}$ 时)或正交相似($F = \mathbb{R}$ 时).
- (2) A 与 B 在 $\mathbb{F}$ 上相似.
- (3) $f_A = f_B$.

证明. “(1) $\implies$ (2)”和“(2) $\implies$ (3)”显然. 只需证明“(3) $\implies$ (1)”. 先假设 $F = \mathbb{C}$. 则 A, B 分别酉相似于对角矩阵 D_1, D_2 . $f_A = f_B$ 推出 $f_{D_1} = f_{D_2}$. 因此 D_1 与 D_2 在重排对角元后相等, 从而酉相似. 这推出 A 与 B 酉相似. 再假设 $F = \mathbb{R}$. 由已经证明的 $F = \mathbb{C}$ 的情况, A 与 B 酉相似. 再由定理9.33, A 与 B 正交相似. □

推论 9.36. 设 $F = \mathbb{R}$ 或 $\mathbb{C}$, $A \in F^{n \times n}$ 正规. 则 A 与 A^T 酉相似($F = \mathbb{C}$ 时)或正交相似($F = \mathbb{R}$ 时).

证明. 只需注意到 A^T 正规并且 $f_A = f_{A^T}$. □

推论 9.37. 设 V 是 F 上的有限维内积空间, $T \in L(V)$ 正规. 设 $A \in F^{n \times n}$ 正规. 如果 $f_A = f_T$, 则存在 V 的有序标准正交基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}} = A$.

证明. 只证明 $F = \mathbb{R}$ 的情况. $F = \mathbb{C}$ 的情况类似. 任取 V 的有序标准正交基 $\mathcal{B}_0$. 设 $A_0 = [T]_{\mathcal{B}_0}$. 则 A_0 正规, 并且 $f_{A_0} = f_A$. 因此 A_0 与 A 正交相似. 而

$$\{[T]_{\mathcal{B}} \mid \mathcal{B} \text{ 是 } V \text{ 的有序标准正交基}\} = \{P^{-1} A_0 P \mid P \text{ 正交}\}.$$

A 属于后一集合, 所以也属于前一集合. □

定理9.21的另一证明. f_T 在 $\mathbb{R}[x]$ 中可以分解为

$$f_T = \prod_{i=1}^k (x - a_i) \prod_{j=1}^m (x^2 - (2r_j \cos \theta_j)x + r_j^2),$$

其中 $a_i \in \mathbb{R}$, $r_j > 0$, $\theta_j \in (0, \pi)$. 取

$$A = \text{diag} \left(a_1, \dots, a_k, r_1 \begin{bmatrix} \cos \theta_1 & -\sin \theta_1 \\ \sin \theta_1 & \cos \theta_1 \end{bmatrix}, \dots, r_m \begin{bmatrix} \cos \theta_m & -\sin \theta_m \\ \sin \theta_m & \cos \theta_m \end{bmatrix} \right).$$

则 $f_A = f_T$. 由上一推论, 存在 V 的有序标准正交基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}} = A$. □

最后再证明一个正规变换的性质.

命题 9.38. 设 $T \in L(V)$ 正规, $S \in L(V)$ 与 T 可交换. 则 S 与 T^* 可交换.

证明. 先假设 $F = \mathbb{C}$. 设 $\sigma(T) = \{c_1, \dots, c_k\}$. 取 $f \in \mathbb{C}[x]$ 满足 $f(c_i) = \bar{c}_i$. 取 V 的有序标准正交基 $\mathcal{B}$ 使得 $[T]_{\mathcal{B}} = D$ 对角. 则

$$[f(T)]_{\mathcal{B}} = f([T]_{\mathcal{B}}) = [T]_{\mathcal{B}}^* = [T^*]_{\mathcal{B}}.$$

这推出 $f(T) = T^*$. 由此即得结论.

再假设 $F = \mathbb{R}$. 任取 V 的有序标准正交基 $\mathcal{B}$. 则 $A := [T]_{\mathcal{B}}$ 正规, 并且 $B := [S]_{\mathcal{B}}$ 与 A 可交换. 只需证明 B 与 $A^* = [T^*]_{\mathcal{B}}$ 可交换. 视这些矩阵为复矩阵. 在 $\mathbb{C}^n$ 上赋予标准内积. 则 $L_A \in L(\mathbb{C}^n)$ 正规, 并且与 $L_B \in L(\mathbb{C}^n)$ 可交换. 由已经证明的 $F = \mathbb{C}$ 的情况, L_B 与 $(L_A)^* = L_{A^*}$ 可交换, 即 B 与 A^* 可交换. □

下面的结果在书上出现了三次.

命题 9.39. 设 V 是任意域上的有限维线性空间, $T \in L(V)$. 设 $V = \bigoplus_{i=1}^k V_i$ 是准素分解, $W \subset V$ 是不变子空间. 则

$$W = \bigoplus_{i=1}^k (W \cap V_i).$$

特别地, 如果 $W \cap V_1 = \{0\}$, 则 $W \subset \bigoplus_{i=2}^k V_i$.

证明. 考虑 T_W 的准素分解并注意 $p_{T_W}|p_T$ 即可. □

第十章 双线性形式

§10.1 双线性形式

参考讲义5.2和5.6节.

设 F 是任意域, V 是有限维 F -线性空间, $f \in M^2(V)$, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是 V 的有序基. 称 (i, j) -元为 $f(\alpha_i, \alpha_j)$ 的 n 阶方阵为 f 在 $\mathcal{B}$ 下的矩阵, 记为 $[f]_{\mathcal{B}}$ (注意与前面不一致, 这里的 $[f]_{\mathcal{B}}$ 是前面的 $[f]_{\mathcal{B}}^t$). 对于

$$\alpha = \sum_i x_i \alpha_i, \beta = \sum_j y_j \alpha_j \in V,$$

有

$$f(\alpha, \beta) = f\left(\sum_i x_i \alpha_i, \sum_j y_j \alpha_j\right) = \sum_{i,j} x_i y_j f(\alpha_i, \alpha_j) = [\alpha]_{\mathcal{B}}^t [f]_{\mathcal{B}} [\beta]_{\mathcal{B}}.$$

(这与前面的 $f(\alpha, \beta) = [\beta]_{\mathcal{B}}^* [f]_{\mathcal{B}} [\alpha]_{\mathcal{B}}$ 也不一致.)

引理 10.1. 映射 $M^2(V) \rightarrow F^{n \times n}$, $f \mapsto [f]_{\mathcal{B}}$ 是线性同构.

证明. 线性显然.

单: 如果 $[f]_{\mathcal{B}} = 0$, 则由 $f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t [f]_{\mathcal{B}} [\beta]_{\mathcal{B}}$ 知 $f = 0$.

满: 对任意 $A \in F^{n \times n}$, $f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}}$ 满足 $[f]_{\mathcal{B}} = A$. □

命题 10.2. 如果 $\mathcal{B}' = \{\alpha'_1, \dots, \alpha'_n\}$ 是另一组有序基, $[\alpha'_1, \dots, \alpha'_n] = [\alpha_1, \dots, \alpha_n]P$, 则

$$[f]_{\mathcal{B}'} = P^t [f]_{\mathcal{B}} P.$$

证明. 此时有 $[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}$. 所以

$$[\alpha]_{\mathcal{B}}^t [f]_{\mathcal{B}'} [\beta]_{\mathcal{B}'} = f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t [f]_{\mathcal{B}} [\beta]_{\mathcal{B}} = [\alpha]_{\mathcal{B}}^t P^t [f]_{\mathcal{B}} P [\beta]_{\mathcal{B}'}, \quad \forall \alpha, \beta \in V.$$

因此 $[f]_{\mathcal{B}'} = P^t [f]_{\mathcal{B}} P$. □

定义 10.1. $A, B \in F^{n \times n}$ 称为是合同的(congruent), 如果存在 $P \in GL_n(F)$ 满足 $B = P^t A P$.

容易看出, 对于 $f \in M^2(V)$, 取 V 的所有有序基得到矩阵的一个合同等价类, 即对 $A_0 = [f]_{\mathcal{B}_0}$,

$$\{[f]_{\mathcal{B}} \mid \mathcal{B} \text{ 是 } V \text{ 的有序基}\} = \{P^t A_0 P \mid P \in GL_n(F)\}.$$

命题 10.3. 设 $f \in M^2(V)$, $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 是 V 的有序基. 定义线性映射

$$L_f, R_f : V \rightarrow V^*, \quad L_f(\alpha)(\beta) = R_f(\beta)(\alpha) = f(\alpha, \beta).$$

则 $\text{rank}(L_f) = \text{rank}(R_f) = \text{rank}([f]_{\mathcal{B}})$.

证明. 记 $A = [f]_{\mathcal{B}}$. 则 $f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}}$. 于是

$$\alpha \in \text{Ker}(L_f) \iff [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}} = 0 \forall \beta \iff [\alpha]_{\mathcal{B}}^t A = 0 \iff [\alpha]_{\mathcal{B}} \in \text{Ker}(A^t).$$

所以 $\dim \text{Ker}(L_f) = \dim \text{Ker}(A^t)$. 这推出 $\text{rank}(L_f) = \text{rank}(A)$. 类似地,

$$\beta \in \text{Ker}(R_f) \iff [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}} = 0 \forall \alpha \iff A [\beta]_{\mathcal{B}} = 0 \iff [\beta]_{\mathcal{B}} \in \text{Ker}(A).$$

所以 $\dim \text{Ker}(R_f) = \dim \text{Ker}(A)$. 这推出 $\text{rank}(R_f) = \text{rank}(A)$. □

$\text{rank}(L_f) = \text{rank}(R_f) = \text{rank}([f]_{\mathcal{B}})$ 称为 f 的秩, 记为 $\text{rank}(f)$.

推论 10.4. 对于 $f \in M^2(V)$, TFAE:

- (1) $\text{rank}(f) = \dim V$.
- (2) 对任意 $\alpha \in V \setminus \{0\}$, 存在 $\beta \in V$ 使得 $f(\alpha, \beta) \neq 0$.
- (3) 对任意 $\beta \in V \setminus \{0\}$, 存在 $\alpha \in V$ 使得 $f(\alpha, \beta) \neq 0$.
- (4) 对某个 V 的有序基 $\mathcal{B}$, $[f]_{\mathcal{B}}$ 可逆.
- (5) 对任意 V 的有序基 $\mathcal{B}$, $[f]_{\mathcal{B}}$ 可逆.

证明. (2) $\iff \text{Ker}(L_f) = \{0\} \iff \text{rank}(L_f) = \dim V \iff (1)$.

(3) $\iff \text{Ker}(R_f) = \{0\} \iff \text{rank}(R_f) = \dim V \iff (1)$.

(4)(或(5)) $\iff \text{rank}([f]_{\mathcal{B}}) = \dim V \iff (1)$. □

满足这几个等价条件的 f 称为**非退化的**(non-degenerate).

§10.2 对称、反对称双线性形式和二次型

定义 10.2. 设 $f \in M^2(V)$.

- (1) 如果对任意 $\alpha, \beta \in V$ 有 $f(\alpha, \beta) = f(\beta, \alpha)$, 则称 f 是**对称的**(symmetric).
- (2) 如果对任意 $\alpha, \beta \in V$ 有 $f(\alpha, \beta) + f(\beta, \alpha) = 0$, 则称 f 是**反对称的**(anti-symmetric或shew-symmetric).
- (3) 如果对任意 $\alpha \in V$ 有 $f(\alpha, \alpha) = 0$, 则称 f 是**交错的**(alternating).

我们把 V 上的所有对称、反对称和交错双线性形式的集合分别记为 $S^2(V)$, $A^2(V)$ 和 $\Lambda^2(V)$. 显然它们都是 $M^2(V)$ 的子空间.

注 10.1. 如果 $\text{char } F = 2$, 则 $S^2(V) = A^2(V)$. 这种情况在这里不是讨论的重点. 所以必要时假设 $\text{char } F \neq 2$.

命题 10.5. (1) $\Lambda^2(V) \subset A^2(V)$.

(2) 如果 $\text{char } F \neq 2$, 则 $\Lambda^2(V) = A^2(V)$, 并且 $M^2(V) = S^2(V) \oplus A^2(V)$.

证明. (1) (以前已证) 若 f 交错, 则对任意 $\alpha, \beta \in V$ 有

$$0 = f(\alpha + \beta, \alpha + \beta) = f(\alpha, \alpha) + f(\beta, \beta) + f(\alpha, \beta) + f(\beta, \alpha) = f(\alpha, \beta) + f(\beta, \alpha).$$

因此 f 反对称.

(2) 若 f 反对称, 则对任意 $\alpha \in V$ 有

$$2f(\alpha, \alpha) = f(\alpha, \alpha) + f(\alpha, \alpha) = 0.$$

由于 $\text{char } F \neq 2$, 这推出 $f(\alpha, \alpha) = 0$, 即 f 交错. 此外, 若 $f \in S^2(V) \cap A^2(V)$, 则 $f(\alpha, \beta) = f(\beta, \alpha) = -f(\alpha, \beta)$, 从而 $f = 0$. 另一方面, 对任意 $f \in M^2(V)$, 取

$$f_1(\alpha, \beta) = \frac{1}{2}(f(\alpha, \beta) + f(\beta, \alpha)), \quad f_2(\alpha, \beta) = \frac{1}{2}(f(\alpha, \beta) - f(\beta, \alpha)).$$

则 $f_1 \in S^2(V)$, $f_2 \in A^2(V)$, 并且 $f = f_1 + f_2$. 因此 $M^2(V) = S^2(V) \oplus A^2(V)$. □

命题 10.6. 设 $\mathcal{B}$ 是有序基.

- (1) f 对称 $\iff [f]_{\mathcal{B}}$ 对称.
- (2) f 反对称 $\iff [f]_{\mathcal{B}}$ 反对称.
- (3) f 交错 $\iff [f]_{\mathcal{B}}$ 反对称并且对角元为0(这种矩阵称为是**交错的**).

证明. “ $\implies$ ”对于三种情况都是显然的. 下面证明“ $\impliedby$ ”. 记 $A = [f]_{\mathcal{B}}$.

(1) 假设 A 对称. 则

$$f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}} = ([\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}})^t = [\beta]_{\mathcal{B}}^t A [\alpha]_{\mathcal{B}} = f(\beta, \alpha).$$

(2) 假设 A 反对称. 则

$$f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}} = ([\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}})^t = -[\beta]_{\mathcal{B}}^t A [\alpha]_{\mathcal{B}} = -f(\beta, \alpha).$$

(3) 假设 A 反对称并且对角元为0. 记 $[\alpha]_{\mathcal{B}} = X = (x_1, \dots, x_n)^t$. 则

$$f(\alpha, \alpha) = X^t A X = \sum_{i=1}^n A_{ii} x_i^2 + \sum_{1 \leq i < j \leq n} (A_{ij} + A_{ji}) x_i x_j = 0.$$

□

定义 10.3. 设 $q: V \rightarrow F$. 如果存在 $f \in M^2(V)$ 满足 $q(\alpha) = f(\alpha, \alpha)$, 则称 q 是 V 上的二次型.

注意 $F^{n \times 1}$ 上的二次型就是二次齐次多项式函数.

记 V 上所有二次型的集合为 $Q(V)$. 它在明显的运算下构成线性空间. 我们有满线性映射 $\Phi: M^2(V) \rightarrow Q(V)$, $\Phi(f)(\alpha) = f(\alpha, \alpha)$.

命题 10.7. 设 $\text{char } F \neq 2$.

(1) $\Phi|_{S^2(V)}: S^2(V) \rightarrow Q(V)$ 是线性同构.

(2) (极化恒等式) 设 $q \in Q(V)$. 如果 $f \in S^2(V)$ 满足 $q(\alpha) = f(\alpha, \alpha)$, 则

$$f(\alpha, \beta) = \frac{1}{4}(q(\alpha + \beta) - q(\alpha - \beta)). \quad (10.1)$$

证明. (1) 显然有 $\text{Ker}(\Phi) = \Lambda^2(V)$. 只需注意到 $\text{char } F \neq 2$ 时有 $M^2(V) = S^2(V) \oplus \Lambda^2(V)$.

(2) 我们有

$$\begin{aligned} q(\alpha + \beta) &= f(\alpha + \beta, \alpha + \beta) = q(\alpha) + q(\beta) + 2f(\alpha, \beta), \\ q(\alpha - \beta) &= f(\alpha - \beta, \alpha - \beta) = q(\alpha) + q(\beta) - 2f(\alpha, \beta). \end{aligned}$$

由此即得(10.1). □

这说明, 对任意 $q \in Q(V)$, 则存在唯一的 $f \in S^2(V)$ 满足 $\Phi(f) = q$, f 的表达式由(10.1)给出. 对于 V 的有序基 $\mathcal{B}$, f 在 $\mathcal{B}$ 下的矩阵称为 q 在 $\mathcal{B}$ 下的矩阵, 记为 $[q]_{\mathcal{B}}$. 注意 $[q]_{\mathcal{B}}$ 对称, 并且

$$q(\alpha) = [\alpha]_{\mathcal{B}}^t [q]_{\mathcal{B}} [\alpha]_{\mathcal{B}}.$$

设 $f \in M^2(V)$ 对称或反对称. 我们希望寻找 V 的有序基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}}$ 充分简单.

定理 10.8. (1) 假设 $\text{char } F \neq 2$. 如果 $f \in S^2(V)$, 则存在有序基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}}$ 对角.

(2) 如果 $f \in \Lambda^2(V)$, 则存在有序基 $\mathcal{B}$ 使得

$$[f]_{\mathcal{B}} = \text{diag} \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, 0, \dots, 0 \right).$$

注 10.2. 如果存在有序基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}}$ 对角, 则 $[f]_{\mathcal{B}}$ 对称, 从而 f 对称.

先证明:

引理 10.9. 设 $f \in M^2(V)$ 对称或反对称, $W \subset V$ 是子空间,

$$W^\perp := \{\beta \in V \mid f(\alpha, \beta) = 0, \forall \alpha \in W\}.$$

假设 $f|_W$ 非退化. 则 $V = W \oplus W^\perp$. 并且如果 $\mathcal{B}_1, \mathcal{B}_2$ 分别是 W 和 $W^\perp$ 的有序基, $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$, 则

$$[f]_{\mathcal{B}} = \text{diag}([f|_W]_{\mathcal{B}_1}, [f|_{W^\perp}]_{\mathcal{B}_2}).$$

证明. 我们有

$$W^\perp = \{\beta \in V \mid L_f(\alpha)(\beta) = 0, \forall \alpha \in W\} = \bigcap_{\alpha \in W} \text{Ker}(L_f(\alpha)) = L_f(W)^0.$$

因此

$$\dim W^\perp = \dim L_f(W)^0 = \dim V - \dim L_f(W) \geq \dim V - \dim W.$$

另一方面, $f|_W$ 非退化 $\implies W \cap W^\perp = \{0\}$. 两者结合起来, 得 $V = W \oplus W^\perp$. 为了得到 $[f]_{\mathcal{B}}$ 的表达式, 只需说明对任意 $\alpha \in \mathcal{B}_1$ 和 $\beta \in \mathcal{B}_2$ 有 $f(\alpha, \beta) = f(\beta, \alpha) = 0$. $f(\alpha, \beta) = 0$ 由 $W^\perp$ 的定义是显然的. 由此和 f 的对称性或反对称性, 即得 $f(\beta, \alpha) = 0$. $\square$

定理10.8的证明. 对 $\dim V$ 用归纳法. $\dim V = 1$ 时显然. 假设 $\dim V = n > 1$, 并且定理当 $\dim V < n$ 成立. 不妨设 $f \neq 0$. 下面对两种情况分别证明.

(1) 由于 $\text{char } F \neq 2$, 由命题10.7(1), $\Phi(f)$ 非零, 即存在 $\alpha \in V$ 满足 $f(\alpha, \alpha) \neq 0$. 设 $W = F\alpha$. 则 $f|_W$ 非退化. 由引理10.9, $\dim W^\perp = n-1$. 从而由归纳假设, 存在 $W^\perp$ 的有序基 $\mathcal{B}_2$ 使 $[f|_{W^\perp}]_{\mathcal{B}_2}$ 对角. 取 W 的有序基 $\mathcal{B}_1 = \{\alpha\}$ 和 V 的有序基 $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$, 则 $[f]_{\mathcal{B}} = \text{diag}([f|_W]_{\mathcal{B}_1}, [f|_{W^\perp}]_{\mathcal{B}_2})$ 对角.

(2) 由 $f \neq 0$ 推出存在线性无关的 $\alpha, \beta \in V$ 满足 $f(\alpha, \beta) = 1$. 取 $W = \text{span}\{\alpha, \beta\}$, $\mathcal{B}_1 = \{\alpha, \beta\}$. 则 $[f|_W]_{\mathcal{B}_1} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. 从而 $f|_W$ 非退化. 由引理10.9, $\dim W^\perp = n-2$. 从而由归纳假设, 存在 $W^\perp$ 的有序基 $\mathcal{B}_2$ 使 $[f|_{W^\perp}]_{\mathcal{B}_2}$ 为定理中的形式. 设 $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$. 则 $[f]_{\mathcal{B}} = \text{diag}([f|_W]_{\mathcal{B}_1}, [f|_{W^\perp}]_{\mathcal{B}_2})$ 也为定理中的形式. $\square$

非退化的交错双线性形式称为**辛形式**.

推论 10.10. 设 V 上存在辛形式. 则 $\dim V$ 为偶数, 并且存在有序基 $\mathcal{B}$ 使得

$$[f]_{\mathcal{B}} = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}.$$

证明. 对定理10.8(2)中的有序基重新排序即可. $\square$

推论 10.11. 设 $\text{char } F \neq 2$, $A \in F^{n \times n}$.

- (1) 如果 A 对称, 则 A 合同于对角矩阵.
- (2) 如果 A 反对称, 则 A 合同于形如

$$\text{diag} \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, 0, \dots, 0 \right)$$

的矩阵.

注 10.3. 当 $\text{char } F = 2$ 时, (1) 不成立. 例如此时 $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 不合同于对角矩阵. 而此时, (2) 对反对称并且对角元为0的矩阵成立.

注 10.4. 一般矩阵(既不对称也不反对称)的合同标准形问题比较复杂. 见 Horn & Johnson.

接下来我们对特殊的域 F 进一步讨论对称双线性形式的矩阵.

定理 10.12. 设 $f \in S^2(V)$.

- (1) 如果 F 是代数闭域并且 $\text{char } F \neq 2$ (例如 $F = \mathbb{C}$), 则存在 V 的有序基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}} = \text{diag}(I_r, 0)$, 其中 $r = \text{rank}(f)$.
- (2) 如果 $F = \mathbb{R}$, 则存在 V 的有序基 $\mathcal{B}$ 使得 $[f]_{\mathcal{B}} = \text{diag}(I_{r_1}, -I_{r_2}, 0)$, 并且 r_1, r_2 被 f 决定.

证明. (1) 由于 $\text{char } F \neq 2$, 由定理10.8(1), 存在有序基 $\mathcal{B}_0 = \{\alpha_1, \dots, \alpha_n\}$ 使得

$$[f]_{\mathcal{B}_0} = \text{diag}(f(\alpha_1, \alpha_1), \dots, f(\alpha_n, \alpha_n)).$$

对 $\alpha_1, \dots, \alpha_n$ 重新排序, 不妨设 $f(\alpha_i, \alpha_i) \neq 0 \iff 1 \leq i \leq r$. 由于 F 是代数闭域, 当 $1 \leq i \leq r$ 时, 可以取 $c_i \in F$ 满足 $c_i^2 = f(\alpha_i, \alpha_i)$. 设 $\mathcal{B} = \{c_1^{-1}\alpha_1, \dots, c_r^{-1}\alpha_r, \alpha_{r+1}, \dots, \alpha_n\}$. 则 $[f]_{\mathcal{B}} = \text{diag}(I_r, 0)$.

(2) 与(1)的证明类似, 存在有序基 $\mathcal{B}_0 = \{\alpha_1, \dots, \alpha_n\}$ 使得

$$[f]_{\mathcal{B}_0} = \text{diag}(f(\alpha_1, \alpha_1), \dots, f(\alpha_r, \alpha_r), 0, \dots, 0),$$

并且对某个 $0 \leq r_1 \leq r$ 有

$$1 \leq i \leq r_1 \implies f(\alpha_i, \alpha_i) > 0,$$

$$r_1 + 1 \leq i \leq r \implies f(\alpha_i, \alpha_i) < 0.$$

当 $1 \leq i \leq r$ 时, 定义 $c_i = \sqrt{|f(\alpha_i, \alpha_i)|}$. 设 $\mathcal{B} = \{c_1^{-1}\alpha_1, \dots, c_r^{-1}\alpha_r, \alpha_{r+1}, \dots, \alpha_n\}$. 则 $[f]_{\mathcal{B}} = \text{diag}(I_{r_1}, -I_{r_2}, 0)$, 其中 $r_2 = r - r_1$.

现在设 V 的有序基 $\mathcal{B} = \{\alpha_1, \dots, \alpha_n\}$ 满足 $[f]_{\mathcal{B}} = \text{diag}(I_{r_1}, -I_{r_2}, 0)$. 我们证明 r_1, r_2 被 f 决定. 取使得 $f|_W$ 正定的维数最大的子空间 $W \subset V$. 记 $V_1 = \text{span}\{\alpha_1, \dots, \alpha_{r_1}\}$, $V_2 = \text{span}\{\alpha_{r_1+1}, \dots, \alpha_n\}$. 注意到:

$$f|_{V_1} \text{ 关于 } V_1 \text{ 的基 } \{\alpha_1, \dots, \alpha_{r_1}\} \text{ 的矩阵为 } I_{r_1} \implies f|_{V_1} \text{ 正定} \implies r_1 = \dim V_1 \leq \dim W,$$

$$f|_{V_2} \text{ 关于 } V_2 \text{ 的基 } \{\alpha_{r_1+1}, \dots, \alpha_n\} \text{ 的矩阵为 } \text{diag}(-I_{r_2}, 0) \implies f|_{V_2} \text{ 半负定}$$

$$\implies V_2 \cap W = \{0\}$$

$$\implies r_1 = n - \dim V_2 \geq \dim W.$$

因此 $r_1 = \dim W$ 被 f 决定. 类似可证明 r_2 被 f 决定. $\square$

定理10.12(2)称为Sylvester惯性定理(Sylvester's law of inertia). 这里“惯性”一词指 r_1, r_2 被 f 决定, 与物理中的“惯性”没有直接关系. r_1 与 r_2 分别称为 f 的**正惯性指数**(positive index of inertia)和**负惯性指数**(negative index of inertia). 数对 (r_1, r_2) 称为 f 的**符号**(signature)¹. 可以证明, 对任何 V 的有序基 $\mathcal{B}$, r_1 (或 r_2) 等于对称矩阵 $[f]_{\mathcal{B}}$ 的正 (或负) 特征值的重数之和. 注意这些定义也对 f 诱导的二次型适用.

注 10.5. 由主轴定理, 定理10.12(2)中的 $\mathcal{B}$ 可以取为 (关于任何内积的) 正交基.

推论 10.13. 设 $A \in F^{n \times n}$ 对称.

(1) 如果 $F = \mathbb{C}$, 则 A 合同于 $\text{diag}(I_r, 0)$. 特别地, 任何 $GL_n(\mathbb{C})$ 中的对称矩阵形如 $P^t P$.

(2) 如果 $F = \mathbb{R}$, 则 A 合同于 $\text{diag}(I_{r_1}, -I_{r_2}, 0)$, 其中 r_1, r_2 被 A 决定.

注 10.6. • 对于推论10.13(2), 由正交相似的结果, 存在 P 为正交矩阵乘正对角矩阵, 使得 $P^t A P = \text{diag}(I_{r_1}, -I_{r_2}, 0)$.

§10.3 双线性形式的自同构群

设 V 是 n 维 F -线性空间, $f \in M^2(V)$. 记

$$\text{Aut}(V, f) = \{T \in GL(V) \mid f(T\alpha, T\beta) = f(\alpha, \beta), \forall \alpha, \beta \in V\}.$$

显然 $\text{Aut}(V, f)$ 是 $GL(V)$ 的子群, 称为 f 或 (V, f) 的**自同构群**. 对于矩阵的情况, 设 $A \in F^{n \times n}$. 容易

¹在有的文献中, f 的符号定义为 $r_1 - r_2$.

验证,

$$G_A := \{M \in \text{GL}_n(F) \mid M^t A M = A\}$$

是 $\text{GL}_n(F)$ 的子群.

引理 10.14. 设 $f \in M^2(V)$, $\mathcal{B}$ 是 V 的有序基, $A = [f]_{\mathcal{B}}$. 则对于 $T \in \text{GL}(V)$, 有 $T \in \text{Aut}(V, f) \iff [T]_{\mathcal{B}} \in G_A$. 特别地, 有群同构 $\text{Aut}(V, f) \cong G_A$.

证明. 记 $M = [T]_{\mathcal{B}}$. 由于 $f(\alpha, \beta) = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}}$, $[T\alpha]_{\mathcal{B}} = M[\alpha]_{\mathcal{B}}$, 所以

$$\begin{aligned} T \in \text{Aut}(V, f) &\iff f(T\alpha, T\beta) = f(\alpha, \beta), \forall \alpha, \beta \in V \\ &\iff [T\alpha]_{\mathcal{B}}^t A [T\beta]_{\mathcal{B}} = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}}, \forall \alpha, \beta \in V \\ &\iff [\alpha]_{\mathcal{B}}^t M^t A M [\beta]_{\mathcal{B}} = [\alpha]_{\mathcal{B}}^t A [\beta]_{\mathcal{B}}, \forall \alpha, \beta \in V \\ &\iff M^t A M = A \\ &\iff M \in G_A. \end{aligned}$$

□

例 10.1. 假设 f 对称非退化. 记 $O(V, f) := \text{Aut}(V, f)$, 称为 f 或 (V, f) 的**正交群**. 如果 $\text{char } F \neq 2$, 则 T 是 f 的自同构 $\iff T$ 是 f 诱导的二次型的自同构. 记

$$O_n(F) := G_I = \{M \in \text{GL}_n(F) \mid M^t M = I\},$$

称为域 F 上的 **n 阶正交群**. 如果 F 是代数闭域并且 $\text{char } F \neq 2$ (例如 $F = \mathbb{C}$), 则存在 $\mathcal{B}$ 满足 $[f]_{\mathcal{B}} = I$. 此时 $\text{Aut}(V, f) \cong O_n(F)$. 如果 $F = \mathbb{R}$, 则存在 $\mathcal{B}$ 满足 $[f]_{\mathcal{B}} = I_{p,q} := \text{diag}(I_p, -I_q)$, 并且符号 (p, q) 被 f 决定. 此时 $\text{Aut}(V, f)$ 同构于

$$O(p, q) := G_{I_{p,q}} = \{M \in \text{GL}_n(\mathbb{R}) \mid M^t I_{p,q} M = I_{p,q}\}.$$

容易看出, $O(p, q) \cong O(q, p)$. 当 p, q 都非零时, $O(p, q)$ 称为**不定正交群**或**伪正交群**. $O(3, 1)$ 称为**Lorentz群**. □

例 10.2. 假设 f 交错非退化. 记 $\text{Sp}(V, f) := \text{Aut}(V, f)$, 称为 f 或 (V, f) 的**辛群**. 由于总存在 $\mathcal{B}$ 满足 $[f]_{\mathcal{B}} = J_{2m} := \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$ (其中 $2m = n$), 所以 $\text{Aut}(V, f)$ 同构于域 F 上的 **$2m$ 阶辛群**

$$\text{Sp}_{2m}(F) := G_{J_{2m}} = \{M \in \text{GL}_{2m}(F) \mid M^t J_{2m} M = J_{2m}\}.$$

□