# 第二章 线性算子

## §1 不同基底下线性映射的矩阵表示

## §2 线性算子代数和矩阵相似

## §3 单个算子生成的子环

## §4 算子和矩阵的极小多项式

## §5 $F[t]$ 中的最小公倍式*(复习与加细)*

## §6 商算子

设 $\mathcal{A} \in \mathcal{L}(V)$, $U$ 是 $\mathcal{A}$-子空间. 定义

$$\begin{array}{rccc} \bar{\mathcal{A}}: & V/U & \longrightarrow & V/U \\ & \mathbf{x}+U & \mapsto & \mathcal{A}(\mathbf{x})+U \end{array}.$$

我们来验证 $\bar{\mathcal{A}}$ 是商空间 $V/U$ 的线性算子.

首先, 我们来验证 $\bar{\mathcal{A}}$ 是良定义的. 设 $\mathbf{x}+U=\mathbf{y}+U$. 则 $\mathbf{x}-\mathbf{y}\in U$. 我们计算

$$\begin{aligned} \bar{\mathcal{A}}(\mathbf{x}+U)-\bar{\mathcal{A}}(\mathbf{y}+U) &= (\mathcal{A}(\mathbf{x})+U)-(\mathcal{A}(\mathbf{y})+U) && (\bar{\mathcal{A}}\text{ 的定义}) \\ &= (\mathcal{A}(\mathbf{x})-\mathcal{A}(\mathbf{y}))+U && (\text{商空间里的运算}) \\ &= \mathcal{A}(\mathbf{x}-\mathbf{y})+U && (\mathcal{A}\text{ 是线性的}) \\ &= U && (\mathbf{x}-\mathbf{y}\in U\text{ 且 }U\text{ 是 }\mathcal{A}\text{-不变的}). \end{aligned}$$

再验证 $\bar{\mathcal{A}}$ 是线性的. 对任意 $\alpha,\beta\in F$ 和 $\mathbf{x},\mathbf{y}\in V$,

$$\begin{aligned} \bar{\mathcal{A}}(\alpha(\mathbf{x}+U)+\beta(\mathbf{y}+U)) &= \bar{\mathcal{A}}((\alpha\mathbf{x}+\beta\mathbf{y})+U) && (\text{商空间里的运算}) \\ &= \mathcal{A}(\alpha\mathbf{x}+\beta\mathbf{y})+U && (\bar{\mathcal{A}}\text{ 的定义}) \\ &= (\alpha\mathcal{A}(\mathbf{x})+\beta\mathcal{A}(\mathbf{y}))+U && (\mathcal{A}\text{ 是线性的}) \\ &= \alpha(\mathcal{A}(\mathbf{x})+U)+\beta(\mathcal{A}(\mathbf{y})+U) && (\text{商空间里的运算}) \\ &= \alpha\bar{\mathcal{A}}(\mathbf{x}+U)+\beta\bar{\mathcal{A}}(\mathbf{y}+U) && (\bar{\mathcal{A}}\text{ 的定义}). \end{aligned}$$

于是, $\bar{\mathcal{A}}$ 是 $V/U$ 的线性算子, 称为 $\mathcal{A}$ 关于 $U$ 的商算子.

**命题 6.1** 设 $\mathcal{A} \in \mathcal{L}(V)$, $U$ 是 $\mathcal{A}$-子空间, $\bar{\mathcal{A}}$ 是 $\mathcal{A}$ 关于 $U$ 的商算子. 设 $\mathbf{e}_1, \ldots, \mathbf{e}_d$ 是 $U$ 的基, $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 是 $V$ 的基. 令 $\mathcal{A}$ 在上述基底下的矩阵是

$$A = \begin{pmatrix} B & * \\ O & C \end{pmatrix},$$

其中 $B \in \mathrm{M}_d(F)$ 和 $C \in \mathrm{M}_{n-d}(F)$. 则 $C$ 是 $\bar{\mathcal{A}}$ 在 $\mathbf{e}_{d+1} + U, \ldots, \mathbf{e}_n + U$ 下的矩阵.

**证明.** 注意到第二章第二讲命题 6,3 指出 $\mathcal{A}$ 在 $\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n$ 下的矩阵必然是 $A$ 具有的分块上三角的形式. 由第一章第二讲引理 4.14 的证明可知 $\mathbf{e}_{d+1} + U, \ldots, \mathbf{e}_n + U$ 是 $V/U$ 的基底. 于是, 命题中的假设都是合理的.

对任意 $j \in \{d+1, d+2, \ldots, n\}$,

$$\begin{aligned}
\bar{\mathcal{A}}(\mathbf{e}_j + U) &= \mathcal{A}(\mathbf{e}_j) + U && (\bar{\mathcal{A}} \text{ 的定义}) \\
&= (\mathbf{e}_1, \ldots, \mathbf{e}_d, \mathbf{e}_{d+1}, \ldots, \mathbf{e}_n)\vec{A}^{(j)} + U && (A \text{ 的定义}) \\
&= (\mathbf{e}_{d+1}, \ldots, \mathbf{e}_n)\vec{C}^{(j-d)} + U && (\mathbf{e}_1, \ldots, \mathbf{e}_d \text{ 的线性组合在 } U \text{ 中}) \\
&= (\mathbf{e}_{d+1} + U, \ldots, \mathbf{e}_n + U)\vec{C}^{(j-d)} && (\text{商空间里的运算}).
\end{aligned}$$

于是 $C \in \mathrm{M}_{n-d}(F)$. 则 $C$ 是 $\bar{\mathcal{A}}$ 在 $\mathbf{e}_{d+1} + U, \ldots, \mathbf{e}_n + U$ 下的矩阵. □

**推论 6.2** 设 $\mathcal{A} \in \mathcal{L}(V)$, $U$ 是 $\mathcal{A}$-子空间, $\bar{\mathcal{A}}$ 是 $\mathcal{A}$ 关于 $U$ 的商算子. 则 $\mu_{\bar{\mathcal{A}}} | \mu_{\mathcal{A}}$.

**证明.** 由第二章第二讲命题 6,3 可知 $\mu_C | \mu_A$, 其中矩阵 $C$ 和 $A$ 在上述命题中定义. 上述命题蕴含 $\mu_C = \mu_{\bar{\mathcal{A}}}$. 于是, $\mu_{\bar{\mathcal{A}}} | \mu_A = \mu_{\mathcal{A}}$. □

**例 6.3** 设 $\mathcal{D} : \mathbb{R}[x]_5 \mapsto \mathbb{R}[x]_5$ 是导数算子. 则 $\mathbb{R}[x]_3$ 是 $\mathcal{D}$-不变的. 而

$$\mathbb{R}[x]_5/\mathbb{R}[x]_3 = \langle x^3 + \mathbb{R}[x]_3, x^4 + \mathbb{R}[x]_3 \rangle.$$

算子 $\mathcal{D}$ 在 $1, x, x^2, x^3, x^4$ 下的矩阵

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

则 $\mathcal{D}$ 关于 $\mathbb{R}[x]_3$ 的商算子的矩阵是红色的子矩阵.

# §7 不可分子空间

**定义 7.1** 设 $\mathcal{A} \in \mathcal{L}(V)$, $U$ 是 $\mathcal{A}$-子空间. 如果 $U$ 不能写成两个非零的 $\mathcal{A}$-子空间的直和, 则称 $U$ 是 $\mathcal{A}$-不可分的.

**命题 7.2** 设 $\mathcal{A} \in \mathcal{L}(V)$. 则 $V$ 是有限个 $\mathcal{A}$-不可分子空间的直和.

**证明.** 设 $n = \dim(V)$. 我们对 $n$ 归纳. 当 $n = 1$ 时, $V$ 本身是 $\mathcal{A}$ 不可分的. 定理显然成立. 设 $n > 1$ 且 当空间维数小于 $n$ 时定理成立. 如果 $V$ 是 $\mathcal{A}$ 不可分的, 则定理成立. 否则存在两个非零 $\mathcal{A}$ 子空间 $U, W$ 使得 $V = U \oplus W$. 则 $\dim(U)$ 和 $\dim(W)$ 的维数都小于 $n$. 由归纳假设, $U = U_1 \oplus \cdots \oplus U_k$, 其中 $U_i$ 是 $\mathcal{A}_U$ 不可分的, 从而也是 $\mathcal{A}$ 不可分的. 同样, $W = W_1 \oplus \cdots \oplus W_\ell$, 其中 $W_j$ 是 $\mathcal{A}_W$ 不可分的, 从而也是 $\mathcal{A}$ 不可分的. 于是

$$V = U \oplus W = U_1 \oplus \cdots \oplus U_k \oplus W_1 \oplus \cdots \oplus W_\ell. \quad \Box$$

# §8 特征向量和特征多项式

在本节中 $V$ 是域 $F$ 上的有限维线性空间, 维数大于零.

## §8.1 特征向量

**定义 8.1** 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 如果 $\langle \mathbf{v} \rangle$ 是 $\mathcal{A}$ 子空间, 则称 $\mathbf{v}$ 是 $\mathcal{A}$ 的一个特征向量 *(eigenvector)*.

**命题 8.2** 设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. 则下列结论等价:

*(i)* $\mathbf{v}$ 是 $\mathcal{A}$ 的特征向量;

*(ii)* $\mathcal{A}(\mathbf{v}) \in \langle \mathbf{v} \rangle$;

*(iii)* 存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$.

**证明.** (i) $\Longrightarrow$ (ii) 显然.

(ii) $\Longrightarrow$ (iii) 因为 $\mathcal{A}(\mathbf{v}) \in \langle \mathbf{v} \rangle$, 所以存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$.

(iii) $\Longrightarrow$ (i) 设 $\mathbf{x} \in \langle \mathbf{v} \rangle$. 则存在 $\alpha \in F$ 使得 $\mathbf{x} = \alpha \mathbf{v}$. 于是,

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}(\alpha \mathbf{v}) = \alpha \mathcal{A}(\mathbf{v}) = \alpha \lambda \mathbf{v} \in \langle \mathbf{v} \rangle \implies \langle \mathbf{v} \rangle \text{ 是 } \mathcal{A} \text{ 不变的.} \quad \Box$$

从上述命题可知 $\mathbf{v}$ 是 $\mathcal{A}$ 特征向量当且仅当存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{v}) = \lambda \mathbf{v}$. 我们称 $\lambda$ 是 关于特征向量 $\mathbf{v}$ 的特征值 (eigenvalue). 简称 $\mathcal{A}$ 的特征根. 反之, 设 $\lambda \in F$ 是 $\mathcal{A}$ 的特征值. 令

$$V^\lambda = \{\mathbf{x} \in V \mid \mathcal{A}(\mathbf{x}) = \lambda \mathbf{x}\}$$

称为 $\mathcal{A}$ 关于 $\lambda$ 的特征子空间(eigenspace). 下面我们来验证 $V^\lambda$ 是 $\mathcal{A}$-子空间.

设 $\alpha, \beta \in F$, $\mathbf{x}, \mathbf{y} \in V^\lambda$. 则

$$\mathcal{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathcal{A}(\mathbf{x}) + \beta\mathcal{A}(\mathbf{y}) = \alpha\lambda\mathbf{x} + \beta\lambda\mathbf{y} = \lambda(\alpha\mathbf{x} + \beta\mathbf{y}).$$

由此可知 $\alpha\mathbf{x}+\beta\mathbf{y} \in V^\lambda$. 即 $V^\lambda$ 是子空间. 因为 $\mathcal{A}(\mathbf{x}) = \lambda\mathbf{x} \in V^\lambda$, 所以 $V^\lambda$ 是 $\mathcal{A}$ 不变的.

**例 8.3** 设 $\mathcal{D}$ 是 $\mathbb{R}[x]_n$ 中的导数算子. 求 $\mathcal{D}$ 所有特征值和特征向量.

**解.** 设 $f = f_{n-1}x^{n-1} + \cdots + f_1x + f_0$, 其中 $f_{n-1}, \ldots, f_1, f_0 \in \mathbb{R}$. 如果 $\mathcal{D}(f) = \lambda f$, 其中 $\lambda \in \mathbb{R}$. 则

$$(n-1)f_{n-1}x^{n-2} + \cdots + f_1 = \lambda(f_{n-1}x^{n-1} + \cdots + f_1x + f_0).$$

上式成立当且仅当 $\lambda = 0$ 且 $f_{n-1} = \cdots = f_1 = 0$. 注意到对任意 $r \in \mathbb{R}$, $\mathcal{D}(r) = 0 = 0r$. 于是, $f$ 是 $\mathcal{D}$ 得特征向量当且仅当 $f \in \mathbb{R} \setminus \{0\}$, 这些特征向量对应得特征值是 0. 而 $V^0 = \mathbb{R}$. $\square$

当我们把矩阵 $A \in \mathrm{M}_n(F)$ 看成 $\mathcal{L}(F^n)$ 中由 $A(\mathbf{x}) = A\mathbf{x}$ 定义得线性算子时, 我们同样有矩阵 $A$ 的特征向量, 特征值和特征子空间的概念.

**例 8.4** 求数乘矩阵的特征向量和特征值.

**解.** 设 $A = \lambda E$, 其中 $\lambda \in F$. 则对任意 $\mathbf{x} \in F^n$, $A\mathbf{x} = \lambda E\mathbf{x} = \lambda\mathbf{x}$. 于是, 任何 $F^n$ 中的非零向量都是 $A$ 的特征向量, 它们对应的特征值都是 $\lambda$. 进而, $V^\lambda = F^n$. $\square$

**例 8.5** 设 $\mathcal{A} \in \mathcal{L}(V)$ 且 $\ker(\mathcal{A}) \neq \{\mathbf{0}\}$. 证明: 0 是 $\mathcal{A}$ 的特征值且 $V^0 = \ker(\mathcal{A})$.

**证明.** 设 $\mathbf{v} \in \ker(\mathcal{A}) \setminus \{\mathbf{0}\}$. 则 $\mathcal{A}(\mathbf{v}) = \mathbf{0} = 0\mathbf{v}$. 于是, 0 是 $\mathcal{A}$ 的特征值. 设 $\mathbf{x} \in \ker(\mathcal{A})$. 则 $\mathcal{A}(\mathbf{x}) = 0\mathbf{x}$. 于是, $\mathbf{x} \in \ker(\mathcal{A})$. 反之, 设 $\mathbf{y} \in V^0$. 则 $\mathcal{A}(\mathbf{y}) = 0\mathbf{y} = \mathbf{0}$. 于是 $\mathbf{y} \in \ker(\mathcal{A})$. 由此得出 $V^0 = \ker(\mathcal{A})$. $\square$

## §8.2 特征多项式

设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 $V$ 的一组基, $\mathcal{A}$ 在该基下的矩阵等于 $A$. 设 $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$, 其中 $x_1, \ldots, x_n \in F$ 不全等于零. 则 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量当且仅当存在 $\lambda \in F$ 使得 $\mathcal{A}(\mathbf{x}) = \lambda\mathbf{x}$, 即

$$\lambda \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \iff (\lambda E - A) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

由此推出 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量蕴含 $\det(\lambda E - A) = 0$.

设 $\chi_A(t) = \det(tE - A) \in F[t]$. 则 $\mathbf{x}$ 是 $\mathcal{A}$ 的特征向量蕴含着它对应的特征值 $\lambda$ 是 $\chi_A(t)$ 的根. 反之, 设 $\lambda \in F$ 是 $\chi_A(t)$ 的根. 则方程组

$$(\lambda E - A)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

由非零解 $(\alpha_1, \ldots, \alpha_n)$. 于是, $\mathbf{v} = \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n$ 满足 $\mathcal{A}(\mathbf{v}) = \lambda\mathbf{v}$. 由此推出 $\lambda \in F$ 是 $\chi_A(t)$ 的根当且仅当 $\lambda$ 是 $\mathcal{A}$ 的特征值.

**定义 8.6** *设 $\mathcal{A} \in \mathcal{L}(V)$, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 是 $V$ 的一组基, $\mathcal{A}$ 在该基下的矩阵等于 $A$. 则 $\det(tE - A)$ 称为 $\mathcal{A}$ 的*特征多项式*(characteristic polynomial), 记为 $\chi_{\mathcal{A}}$. 特征多项式 $\chi_{\mathcal{A}}(t)$ 在 $F$ 中所有根的集合记为 $\mathrm{spec}_F(\mathcal{A})$, 称为 $\mathcal{A}$ 的在 $F$ 中的*谱*(spectrum)*

注意到矩阵 $A$ 与基底选取有关. 为了验证上述定义的合理性, 我们设 $B$ 是 $\mathcal{A}$ 在基底 $(\epsilon_1, \ldots, \epsilon_n) = (\mathbf{e}_1, \ldots, \mathbf{e}_n)P$ 下的矩阵, 其中 $P \in \mathrm{GL}_n(F)$. 则 $B = P^{-1}AP$ (见第二章第一讲第 2.2 节第一段). 我们有

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

这样就验证了上述定义的合理性.

类似地, 设 $A \in \mathrm{M}_n(F)$. 我们称 $\chi_A = \det(tE - A)$ 为矩阵 $A$ 的特征多项式. 上述合理性验证也说明 $\chi_A$ 是相似不变量. 从而, $\mathrm{spec}_A$ 中的元素都是相似不变量.

**注解 8.7** *根据上述讨论, $\mathcal{A}$ 的特征值也称为 $\mathcal{A}$ 的*特征根*(eigenroot).*

**注解 8.8** *设 $n = \dim(V)$, $\mathcal{A} \in \mathcal{L}(V)$. 则 $\chi_{\mathcal{A}}(t)$ 次数等于 $n$ 且首一(见第一章第二次习题课讲义第 1 页)*

下面我们演示通过特征多项式求特征向量的方法.

**例 8.9** *设 $\mathbf{e}_1, \mathbf{e}_2$ 是 $\mathbb{R}^2$ 的标准基, $\mathcal{A}, \mathcal{B} \in \mathcal{L}(\mathbb{R}^2)$ 分别由公式 $\mathcal{A}(\mathbf{e}_1) = \mathbf{e}_2, \mathcal{A}(\mathbf{e}_2) = \mathbf{e}_1$ 和 $\mathcal{B}(\mathbf{e}_1) = \mathbf{e}_2$, $\mathcal{B}(\mathbf{e}_2) = -\mathbf{e}_1$ 给出. 计算 $\mathcal{A}$ 和 $\mathcal{B}$ 的特征子空间.*

**解.** 算子 $\mathcal{A}$ 和 $\mathcal{B}$ 在 $\mathbf{e}_1, \mathbf{e}_2$ 下的矩阵分别是

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

于是

$$\chi_{\mathcal{A}}(t) = \det(tE-A) = \det\begin{pmatrix} t & -1 \\ -1 & t \end{pmatrix} = t^2-1, \quad \chi_{\mathcal{B}}(t) = \det(tE-B) = \det\begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2+1.$$

于是, $\mathrm{spec}_{\mathbb{R}}(\mathcal{A}) = \{1, -1\}$, $\mathrm{spec}_{\mathcal{B}} = \emptyset$. 从而 $\mathcal{B}$ 没有特征根, 从而没有特征向量和特征子空间.

对于 $\mathcal{A}$ 特征根 $\lambda_1 = 1$, 它对应的特征子空间是方程组

$$(\lambda_1 E - A)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 即方程组

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 解方程组得 $V^{\lambda_1} = \langle (1,1)^t \rangle$. 类似地, 特征根 $\lambda_2 = -1$ 对应的特征子空间是 $V^{\lambda_2} = \langle (1,-1)^t \rangle$.

**例 8.10** 设 $\mathbf{e}_1, \mathbf{e}_2$ 是 $\mathbb{C}^2$ 的标准基, $\mathcal{B} \in \mathcal{L}(\mathbb{C}^2)$ 由公式 $\mathcal{B}(\mathbf{e}_1) = \mathbf{e}_2$, $\mathcal{B}(\mathbf{e}_2) = -\mathbf{e}_1$ 给出. 计算 $\mathcal{B}$ 的特征子空间.

**解.** 算子 $\mathcal{B}$ 在 $\mathbf{e}_1, \mathbf{e}_2$ 下的矩阵是

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

于是

$$\chi_{\mathcal{B}}(t) = \det(tE - B) = \det\begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1.$$

于是, $\mathrm{spec}_{\mathbb{C}}(\mathcal{B}) = \{\sqrt{-1}, -\sqrt{-1}\}$. 特征根 $\lambda_1 = \sqrt{-1}$, 它对应的特征子空间是方程组

$$(\lambda_1 E - A)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 即方程组

$$\begin{pmatrix} \sqrt{-1} & 1 \\ -1 & \sqrt{-1} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的解空间. 解方程组得 $V^{\lambda_1} = \langle (1, -\sqrt{-1})^t \rangle$. 类似地, 特征根 $\lambda_2 = -\sqrt{-1}$ 对应的特征子空间是 $V^{\lambda_2} = \langle (1, \sqrt{-1})^t \rangle$.

**命题 8.11** 设 $V$ 是复数域 $\mathbb{C}$ 上的线性空间, $\mathcal{A} \in \mathcal{L}(V)$. 则 $\mathcal{A}$ 一定有特征向量.

**证明.** 因为 $\chi_{\mathcal{A}} \in \mathbb{C}[t] \setminus \mathbb{C}$, 所以 $\chi_{\mathcal{A}}$ 在 $\mathbb{C}$ 中至少有一个根 $\lambda$ (代数学基本定理). 即 $\mathcal{A}$ 有特征根. 于是有特征向量. □

**例 8.12** 设 $A \in \mathrm{M}_n(\mathbb{C})$. 证明: $A$ 相似于一个上三角矩阵.

**证明.** 对 $n$ 归纳. 当 $n=1$ 时, 结论显然成立. 设 $n>1$ 且 $n-1$ 时结论成立.

考虑 $A \in \mathrm{M}_n(\mathbb{C})$. 把 $A$ 看成 $\mathbb{C}^n$ 上在标准基 $\mathbf{e}_1, \ldots, \mathbf{e}_n$ 下矩阵等于 $A$ 的线性算子. 有上例可知, $A$ 有一个 1 维 $\mathcal{A}$ 子空间 $\langle \mathbf{u} \rangle$. 根据第二章第二讲命题 6.3,

$$A \sim_s \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix},$$

其中 $\lambda \in \mathbb{C}$, $B \in \mathrm{M}_{n-1}(\mathbb{C})$. 根据归纳假设. 存在 $P \in \mathrm{GL}_{n-1}(\mathbb{C})$ 使得 $P^{-1}BP$ 是上三角的. 令

$$Q = \begin{pmatrix} 1 & O_{1\times n-1} \\ O_{(n-1)\times 1} & P \end{pmatrix}.$$

则 $P$ 可逆且

$$\begin{aligned} Q^{-1} \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix} Q &= \begin{pmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & P^{-1} \end{pmatrix} \begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & B \end{pmatrix} \begin{pmatrix} 1 & O_{1\times(n-1)} \\ O_{(n-1)\times 1} & P \end{pmatrix} \\ &= \underbrace{\begin{pmatrix} \lambda & * \\ O_{(n-1)\times 1} & P^{-1}BP \end{pmatrix}}_{T}. \end{aligned}$$

因为 $P^{-1}BP$ 已经是上三角矩阵, 所以 $T$ 也是是上三角矩阵. 显然, $A \sim_s T$. □

**命题 8.13** 设 $A \in \mathrm{M}_n(F)$,

$$\chi_A = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0, \quad a_i \in F.$$

则 $a_{n-1} = -\mathrm{tr}(A)$ 和 $a_n = (-1)^n \det(A)$. 特别地, $A$ 可逆当且仅当 0 不是 $A$ 的特征根.

**证明.** 由第一章第二次习题课讲义第 1 页, 可知 $a_{n-1} = -\mathrm{tr}(A)$ 和 $a_0 = (-1)^n \det(A)$. 注意到 $A$ 可逆当且仅当 $\det(A) \neq 0$ 当且仅当 $a_0 \neq 0$ 当且仅当 $\chi_A(0) \neq 0$. □

因为矩阵的迹和行列式都是相似不变量, 所以我们可以定义算子的迹和行列式如下.

**定义 8.14** 设 $\mathcal{A} \in \mathcal{L}(V)$, $A$ 是 $\mathcal{A}$ 在 $V$ 的某组基下的矩阵. 则 $A$ 的迹称为 $\mathcal{A}$ 的迹, 记为 $\mathrm{tr}(\mathcal{A})$. 同样地, $A$ 的行列式称为 $\mathcal{A}$ 的行列式, 记为 $\det(\mathcal{A})$.

因为 $\chi_A(t)$ 是相似不变量, 所以它的系数也都是相似不变量. 由第一章第二次习题课讲义命题 5.4 可知, $A$ 的 $k$ 阶主子式之和都是相似不变量, $k = 1, 2, \ldots, n$.

**例 8.15** 设 $A \in \mathrm{M}_n(F)$ 是如下分块上三角形

$$\begin{pmatrix} A_1 & * & * & \cdots & * \\ O & A_2 & * & \cdots & * \\ O & O & A_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ O & O & O & \cdots & A_k \end{pmatrix}.$$

证明: $\chi_A = \chi_{A_1} \cdots \chi_{A_k}$.

**证明.**

$$\chi_A(t) = \det(tE - A) = \begin{pmatrix} tE_{n_1} - A_1 & * & * & \cdots & * \\ O & tE_{n_2} - A_2 & * & \cdots & * \\ O & O & tE_{n_3} - A_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & * \\ O & O & O & \cdots & tE_{n_k} - A_k \end{pmatrix},$$

其中 $A_i$ 是 $n_i$ 阶方阵, $i = 1, 2, \ldots, k$. 于是,

$$\chi_A(t) = \prod_{i=1}^{k} \det(tE_{n_i} - A_i) = \prod_{i=1}^{k} \chi_{A_i}(t). \quad \square$$

## §9 对角化

**定义 9.1** 设 $\mathcal{A} \in \mathcal{L}(V)$. 如果 $\mathcal{A}$ 在 $V$ 的某组基下的矩阵是对角矩阵, 则称 $\mathcal{A}$ 是可对角化的. 如果 $A \in \mathrm{M}_n(\mathbb{R})$ 相似于一个对角矩阵, 则称 $A$ 是可对角化的.

**定理 9.2** *(可对角化判别法I)* 设 $n = \dim(V)$, $\mathcal{A} \in \mathcal{L}(V)$. 则 $\mathcal{A}$ 可对角化当且仅当 $\mathcal{A}$ 有 $n$ 个线性无关得特征向量.

**证明.** 设 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 $\mathcal{A}$ 的 $n$ 个线性无关的特征向量. 设 $\mathcal{A}(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$, $i = 1, 2, \ldots, n$. 注意到 $\lambda_1, \ldots, \lambda_n \in F$ 不一定两两不同. 此时, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 是 $V$ 的一组基, 且

$$(\mathcal{A}(\mathbf{v}_1), \mathcal{A}(\mathbf{v}_2), \ldots, \mathcal{A}(\mathbf{v}_n)) = (\lambda_1 \mathbf{v}_1, \lambda_2 \mathbf{v}_2, \ldots, \lambda_n \mathbf{v}_n) = (\mathbf{v}_1, \ldots, \mathbf{v}_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

反之, 设 $\mathcal{A}$ 在基底 $\epsilon_1, \ldots, \epsilon_n$ 下的矩阵是 $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. 则 $\mathcal{A}(\epsilon_i) = \lambda_i \epsilon_i$. 于是, $\epsilon_i$ 是 $\mathcal{A}$ 的特征向量且 $\epsilon_1, \ldots, \mathbf{e}_n$ 线性无关. $\square$

**推论 9.3** 设 $A \in \mathrm{M}_n(F)$. 则 $A$ 可对角化当且仅当 $A$ 有 $n$ 个线性无关的特征向量. 设 $A$ 有 $n$ 个线性无关的特征向量 $\mathbf{v}_1, \ldots, \mathbf{v}_n$. 令 $P = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$. 则 $P^{-1}AP$ 是对角矩阵.

**证明.** 把 $A$ 看成 $\mathcal{L}(F^n)$ 上的线性算子满足 $\mathbf{x} \mapsto A\mathbf{x}$. 则由上述定理可知矩阵 $A$ 相似于对角阵当且仅当 $A$ 有 $n$ 个线性无关的特征向量. 此时, $P$ 是从标准基到基底 $\mathbf{v}_1, \ldots, \mathbf{v}_n$ 的转换矩阵. 于是, $P^{-1}AP$ 是对角阵. □

**注解 9.4** 线性算子 $\mathcal{A} \in \mathcal{L}(V)$ 可对角化当且仅当 $\mathcal{A}$ 在 $V$ 任何一组基下的矩阵可对角化.

**例 9.5** 如例 *8.3* 所示, $\mathbb{R}[x]_n$ 中关于导数算子 $\mathcal{D}$ 的特征向量是 $f \in \mathbb{R} \setminus \{0\}$. 不存在两个线性无关的特征向量. 于是, 当 $n > 1$ 是 $\mathcal{D}$ 不能对角化.

**例 9.6** *(科斯特利金第一卷第 72 页)* 设

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}).$$

判断 $A$ 是否能对角化. 如果可以, 求 $P \in \mathrm{M}_2(\mathbb{R})$ 使得 $P^{-1}\mathrm{M}_nP$ 是对角矩阵.

**解.** 计算

$$\chi_A(t) = \det \begin{pmatrix} t & -1 \\ -1 & t-1 \end{pmatrix} = t^2 - t - 1.$$

解方程得

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

下面计算特征向量. 特征值 $\lambda_1$ 对应得特征向量是方程组

$$\begin{pmatrix} \lambda_1 & -1 \\ -1 & \lambda_1 - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

的非零解. 因为 $\dim(V^{\lambda_1}) = 1$, 所以取 $(1, \lambda_1)^t$ 即可. 类似取 $\lambda_2$ 对应的特征向量 $(1, \lambda_2)^2$. 因为 $\lambda_1 \neq \lambda_2$, 所以这两个特征向量线性无关. 则

$$P = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix}.$$

于是

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

**引理 9.7** 设 $\mathcal{A} \in \mathcal{L}(V)$, $\lambda_1, \ldots, \lambda_k \in_F (\mathcal{A})$ 两两不同. 则

$$V^{\lambda_1} + \cdots + V^{\lambda_k}$$

是直和.

**证明.** 对 $k$ 归纳. 当 $k = 1$ 时结论显然成立. 设 $k > 1$ 且当 $k - 1$ 时结论成立.

设 $\mathbf{v}_i \in V^{\lambda_i}$, $i = 1, 2, \ldots, k$, 满足

$$\mathbf{0} = \mathbf{v}_1 + \cdots + \mathbf{v}_k.$$

则

$$\mathbf{0} = \mathcal{A}(\mathbf{v}_1 + \cdots + \mathbf{v}_{k-1} + \mathbf{v}_k) = \mathcal{A}(\mathbf{v}_1) + \cdots + \mathcal{A}(\mathbf{v}_{k-1}) + \mathcal{A}(\mathbf{v}_k) = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_{k-1} \mathbf{v}_{k-1} + \lambda_k \mathbf{v}_k.$$

第一式通乘 $\lambda_k$ 与第二式相减得

$$\mathbf{0} = (\lambda_k - \lambda_1)\mathbf{v}_1 + \cdots + (\lambda_k - \lambda_{k-1})\mathbf{v}_{k-1}.$$

由归纳假设可知, $(\lambda_k - \lambda_1) = \cdots = (\lambda_k - \lambda_{k-1})\mathbf{v}_{k-1} = \mathbf{0}$. 因为 $(\lambda_k - \lambda_1), \ldots, (\lambda_k - \lambda_{k-1})$ 都非零, 所以 $\mathbf{v}_1 = \cdots = \mathbf{v}_{k-1} = \mathbf{0}$. 进而 $\mathbf{v}_k = \mathbf{0}$. 由第一章第一讲定理 1.12, $V^{\lambda_1} + \cdots + V^{\lambda_k}$ 是直和. □