

## 第二章 线性算子

### 7 特征向量和特征值

在本节中,  $V$  是域  $F$  上的有限维线性空间且  $\dim(V) > 0$ .

#### 7.1 特征向量

**定义 7.1** 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 7.2** 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)  $\implies$  (ii) 显然.

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

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

**注解 7.3** 可直接验证:  $\mathbf{v} \in V$  是线性算子  $\mathcal{A} \in \mathcal{L}(V)$  的特征向量当且仅当  $\mu_{\mathcal{A}, \mathbf{v}}$  的次数等于 1.

根据命题 7.2, 非零  $\mathbf{v}$  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  $\mathcal{A}$  的特征值. 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 7.4** 设  $\mathcal{A} = \lambda\mathcal{E}$ ,  $\lambda \in F$ . 则  $\mathcal{A}$  的唯一特征值是  $\lambda$  且  $V^\lambda = V$ .

设  $\mathcal{A} = \mathcal{A}^2$  且  $\mathcal{A}$  既不等于  $\mathcal{E}$  也不等于  $\mathcal{O}$ . 则  $\mathcal{A}$  不可逆. 故 0 是它的特征值且  $V^0 = \ker(\mathcal{A})$ .

再设  $\lambda \in F \setminus \{0\}$  是  $\mathcal{A}$  的特征值和  $\mathbf{v}$  是其对应的特征向量. 则  $\mathcal{A}(\mathbf{v}) = \lambda\mathbf{v}$  故  $\mathcal{A}^2(\mathbf{v}) = \lambda\mathcal{A}(\mathbf{v})$ . 因为  $\mathcal{A} = \mathcal{A}^2$ ,

所以  $\lambda \mathbf{v} = \lambda^2 \mathbf{v}$ . 故  $\lambda = 1$ . 可直接验证  $\text{im}(\mathcal{A})$  中关于的非零向量都是关于 1 的特征向量. 因为  $\mathcal{A} \neq \mathcal{O}$ , 所以  $\text{im}(\mathcal{A}) \neq \{\mathbf{0}\}$ . 故 1 是  $\mathcal{A}$  的特征值且  $\text{im}(\mathcal{A}) \subset V^1$ . 而  $V^1 \subset \text{im}(\mathcal{A})$  是显然的. 故  $V^1 = \text{im}(\mathcal{A})$ .

设  $k \in \mathbb{Z}^+$  使得  $\mathcal{A}^k = \mathcal{O}$  且  $\mathcal{A} \neq \mathcal{O}$ . 因为  $\mathcal{A}$  不可逆, 所以 0 是  $\mathcal{A}$  的特征值且  $V^0 = \ker(\mathcal{A})$ . 再设  $\lambda \in F \setminus \{0\}$  是  $\mathcal{A}$  的特征值和  $\mathbf{v}$  是其对应的特征向量. 则  $\mathcal{A}^k(\mathbf{v}) = \lambda^k \mathbf{v} = \mathbf{0}$ . 故  $\lambda = 0$ , 矛盾. 由此可知, 0 是  $\mathcal{A}$  唯一的特征值且对应的特征子空间是  $V^0 = \ker(\mathcal{A})$ .

当我们把矩阵  $A \in M_n(F)$  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.

## 7.2 特征多项式

设  $A \in M_n(F)$ ,  $\mathbf{x} = (x_1, \dots, x_n)^t \in F^n \setminus \{\mathbf{0}\}$ . 则存在  $\lambda \in F$  使得  $A\mathbf{x} = \lambda\mathbf{x}$ , 即  $\mathbf{x}$  是  $A$  的特征向量, 当且仅当

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

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

**定义 7.5** 设  $A \in M_n(F)$ ,  $t$  是  $F$  上的未定元. 多项式

$$\det(tE - A) \in F[t]$$

称为  $A$  的特征多项式, 记为  $\chi_A(t)$ .

**定义 7.6** 设  $A \in M_n(F)$ . 特征多项式  $\chi_A$  在  $F$  中的根称为  $A$  的特征根 (*eigenroots*). 这些特征根的集合记为  $\text{spec}_F(A)$ , 称为  $A$  在  $F$  中的谱 (*spectrum*).

矩阵的特征根就是矩阵的特征值.

**例 7.7** 设实二阶矩阵,  $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . 求  $A$  和  $B$  的所有特征根和特征向量.

解. 直接计算得

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

和

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

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

设特征根  $\lambda_1 = 1$ . 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 7.8** 设上例中矩阵  $B$  是复矩阵. 求它的特征值和特征向量.

解. 由上例可知,  $\chi_B(t) = t^2 + 1$ . 于是,

$$\text{spec}_{\mathbb{C}}(B) = \{\sqrt{-1}, -\sqrt{-1}\}.$$

设特征根  $\lambda_1 = \sqrt{-1}$ . 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  $V^{\lambda_2} = \langle (1, \sqrt{-1})^t \rangle$ .

设  $\chi_A(t) = \det(tE - A) \in F[t]$ . 则  $\mathbf{x}$  是  $A$  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  $(\alpha_1, \dots, \alpha_n)$ . 于是,  $\mathbf{v} = \alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n$  满足  $A(\mathbf{v}) = \lambda \mathbf{v}$ . 由此推出  $\lambda \in F$  是  $\chi_A(t)$  的根当且仅当  $\lambda$  是  $A$  的特征值.

**命题 7.9** 矩阵特征多项式是相似不变量.

证明. 设  $A, B \in M_n(F)$  且  $A \sim_s B$ . 则存在  $P \in GL_n(F)$  使得  $B = P^{-1}AP$ . 直接计算得

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

即  $\chi_A = \chi_B$ .  $\square$

**定义 7.10** 设  $A \in M_n(F)$ . 特征多项式  $\chi_A$  在  $F$  中的根称为  $A$  的特征根 (*eigenroots*). 这些特征根的集合记为  $\text{spec}_F(A)$ , 称为  $A$  在  $F$  中的谱 (*spectrum*).

矩阵的特征根就是它的特征值,  $\text{spec}_F(A)$  也是相似不变量.

**命题 7.11** 设  $A \in M_n(F)$ ,

$$\chi_A = t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0, \quad a_i \in F.$$

则  $a_{n-1} = -\text{tr}(A)$  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. 设  $A = (a_{i,j})_{n \times n}$ . Note that

$$\chi_A = \det \begin{pmatrix} t - a_{1,1} & -a_{1,2} & \cdots & -a_{1,n} \\ -a_{2,1} & t - a_{2,2} & \cdots & -a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n,1} & -a_{n,2} & \cdots & t - a_{n,n} \end{pmatrix}.$$

故  $\chi_A = (t - a_{1,1})(t - a_{2,2}) \cdots (t - a_{n,n}) + p$ , 其中  $\deg(p) < n - 1$ . 于是  $\chi_A = t^n - \text{tr}(A)t^{n-1} + q$ , 其中  $\deg(q) < n - 1$ . 因为  $\chi_A(0) = \det(-A)$ , 所以  $a_n = (-1)^n \det(A)$ . 进而,  $\chi_A(0) \neq 0$  当且仅当  $\det(A) \neq 0$ , 即  $A$  可逆.  $\square$

**定义 7.12** 设  $\mathcal{A} \in \mathcal{L}(V)$ ,  $\mathbf{e}_1, \dots, \mathbf{e}_n$  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  $\text{spec}_F(\mathcal{A})$ , 称为  $\mathcal{A}$  的在  $F$  中的谱 (*spectrum*)

本节中关于矩阵特征向量和特征值的结果可以翻译成算子的语言, 表述如下设  $\mathcal{A} \in \mathcal{L}(V)$ .

1.  $\chi_{\mathcal{A}}$  是良定义的 (命题 7.9).
2.  $\chi_{\mathcal{A}}$  是次数为  $\dim(V)$  的首一多项式,  $\mathcal{A}$  可逆当且仅当  $\chi_{\mathcal{A}}(0) \neq 0$ . (命题 7.11)

下面我们描述定义 7.12 中  $\mathcal{A}$  与  $A$  特征向量之间的关系. 设  $\mathbf{v} \in V \setminus \mathbf{0}$  且

$$\mathbf{v} = (\mathbf{e}_1, \dots, \mathbf{e}_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

则  $\mathbf{v}$  是关于特征值  $\lambda$  的特征向量当且仅当

$$\begin{aligned} \mathcal{A}(\mathbf{v}) = \lambda \mathbf{v} &\iff (\mathbf{e}_1, \dots, \mathbf{e}_n) A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \lambda (\mathbf{e}_1, \dots, \mathbf{e}_n) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \\ &\iff A \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}. \end{aligned}$$

故  $\mathbf{v}$  是  $\mathcal{A}$  关于  $\lambda$  的特征向量当且仅当其坐标是  $A$  关于  $\lambda$  的特征向量. 由此可知, 计算线性算子的特征值和特征向量的问题可以转化为计算其矩阵表示的特征值和特征向量的问题.