第一章　线性变换的表示与分解

1.1　线性变换的特征值和特征向量

1.1.1　特征值和特征向量的定义

定义 1.1.1: 线性变换的特征值、特征向量、特征子空间

设 V 是一个 F 上的线性空间，$A \in \hom(V, V)$

那么如果 $\exists \lambda \in F, \alpha \in V, \alpha \neq \mathbf{0}$ 使得

$$A(\alpha) = \lambda\alpha \tag{1.1}$$

成立

那么我们称 λ 是 A 的一个特征值，α 是 A 的隶属于 λ 的一个特征向量

同时我们定义：

$$V_\lambda := \{\alpha | A(\alpha) = \lambda\alpha\} \tag{1.2}$$

称为 A 的属于特征值 λ 的特征子空间

特征值代表着一个线性映射对一个向量的伸缩程度，这里我们要求至少能找到一个向量不为零是因为：如果一个特征值只对于零向量成立，那么它过于平凡，而且会扰乱后续我们一些关于特征值数量的命题

类似地，我们自然可以定义矩阵的特征值

定义 1.1.2: 矩阵的特征值、特征向量

设 $\boldsymbol{A} \in M_n(F)$

那么如果 $\exists \lambda \in F, \alpha \in F^n$，使得

$$\boldsymbol{A}\alpha = \lambda\alpha \tag{1.3}$$

成立。

那么我们称 λ 是 $\boldsymbol{A}$ 的一个特征值，α 是 $\boldsymbol{A}$ 的隶属于 λ 的一个特征向量

1.1.2 特征值和特征向量的性质

我们探讨一些相关的基本性质

1. 我们首先验证，特征子空间的确是子空间

   命题 1.1.1. $\forall A \in \hom(V,V)$，V_λ 是 V 的一个线性子空间

   证明: 取 $\forall \alpha, \beta \in V_\lambda, k \in F$

   $A(\alpha+\beta) = \lambda\alpha + \lambda\beta = \lambda(\alpha+\beta) \Rightarrow \alpha+\beta \in V_\lambda$

   $A(k\alpha) = k\lambda\alpha = \lambda(k\alpha) \Rightarrow k\alpha \in V_\lambda$

   于是命题得证 □

   请注意：特征子空间并非全体特征向量的集合，而是全体特征向量和零向量的集合

2. 从特征值的定义可以看出，特征值就像是把线性变换的一部分转换为纯量乘法。我们猜想：每一个特征值体现了映射的不同的“方向”，不同特征子空间的线性无关向量组的并也应该线性无关

   以下命题证实了这个猜想

   命题 1.1.2. 设 $A \in \hom(V,V)$，λ_1, λ_2 是 A 的两个特征值，$\lambda_1 \neq \lambda_2$

   如果 $\alpha_1, \cdots, \alpha_m \in V_{\lambda_1}$ 线性无关，$\beta_1, \cdots, \beta_n \in V_{\lambda_2}$ 线性无关

   那么 $\alpha_1, \cdots, \alpha_m, \beta_1, \cdots, \beta_n$ 也线性无关

   证明: 取线性组合并设其为零:

   $k_1\alpha_1 + \cdots + k_m\alpha_m + l_1\beta_1 + \cdots + l_n\beta_n = \mathbf{0}$

   将 A 在其上进行变换得：$k_1A(\alpha_1) + \cdots + k_mA(\alpha_m) + l_1A(\beta_1) + \cdots + l_nA(\beta_n) = \mathbf{0}$

   $\Rightarrow k_1\lambda_1\alpha_1 + \cdots + k_m\lambda_1\alpha_m + l_1\lambda_2\beta_1 + \cdots + l_n\lambda_2\beta_n = \mathbf{0}$

但是，如果我们把最初的线性组合乘以 λ_1，得：

$k_1\lambda_1\alpha_1+\cdots+k_m\lambda_1\alpha_m+l_1\lambda_1\beta_1+\cdots+l_n\lambda_1\beta_n=\mathbf{0}$

将两式相减，得：

$l_1(\lambda_2-\lambda_1)\beta_1+\cdots+l_n(\lambda_2-\lambda_1)\beta_n=\mathbf{0}$

$\Rightarrow \forall i, l_i(\lambda_2-\lambda_1)=0 \Rightarrow l_i=0$

$\Rightarrow k_1\alpha_1+\cdots+k_m\alpha_m=\mathbf{0} \Rightarrow \forall i, k_i=0$

于是命题得证 □

一个显然的推论是此结论的 n 个子空间的版本：

推论 1.1.1

$A\in\hom(V,V)$，$\lambda_1,\cdots,\lambda_n$ 是 A 的 n 个互不相同的特征值

如果 $\forall i, \alpha_{ir_1},\cdots,\alpha_{ir_m}\in V_{\lambda_i}$ 线性无关

那么向量组 $\{\alpha_{jr_k}\}$ 也线性无关

证明： 由数学归纳法易证 □

3. 容易注意到，矩阵的特征值和线性映射的特征值其实是一样的，正如下面的命题：

命题 1.1.3. 设 V 是一个 F 上的线性空间，$\dim V=n<\aleph_0$，$\{\alpha_1,\cdots,\alpha_n\}$ 是 V 的一个基

$A\in\hom(V,V)$ 在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$

那么，λ 是 A 的一个特征值，α 是 A 的一个隶属于 λ 的一个特征向量 $\Leftrightarrow$

λ 是 $\boldsymbol{A}$ 的一个特征值，并且 α 在 $\{\alpha_1,\cdots,\alpha_n\}$ 下的坐标 $\mathbf{x}$ 是 $\boldsymbol{A}$ 的一个隶属于 λ 的特征向量

证明： $A(\alpha)=\lambda\alpha$

$\Leftrightarrow \boldsymbol{A}\mathbf{x}=\lambda\mathbf{x}$（参见命题 2.3.23）

于是命题得证。 □

1.1.3 特征矩阵与特征多项式

前面我们研究了特征值的性质，接下来我们想知道：是否可以直接去寻找计算特征值的直接方法？事实上，是可以的，我们指出：特征值是特征多项式的一个根，而特征向量是对应映射的核的一个元素

先给出定义：

定义 1.1.3: 线性变换的特征多项式

设 $A \in \hom(V,V)$，我们称

$\chi_A(\lambda) = \det(\lambda I - A)$

为 A 的特征多项式

类似地，我们可以定义矩阵的特征矩阵和特征多项式：

定义 1.1.4: 矩阵的特征矩阵和特征多项式

设 $\boldsymbol{A} \in M_n(F)$，我们称 $\lambda \boldsymbol{I} - \boldsymbol{A}$ 是 $\boldsymbol{A}$ 的特征矩阵

并称 $\chi_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A})$ 是 $\boldsymbol{A}$ 的特征多项式

下面的性质指出了我们想要的结果：

1.

命题 1.1.4. 设 V 是一个 F 上的线性空间，$\dim V = n < \aleph_0$，$\{\alpha_1, \cdots, \alpha_n\}$

如果 $A \in \hom(V,V)$ 在 $\{\alpha_1, \cdots, \alpha_n\}$ 下的矩阵是 $\boldsymbol{A}$

那么我们断言：$\chi_A(\lambda) = \chi_{\boldsymbol{A}}(\lambda)$

证明： 这是显然的，因为矩阵的行列式的定义就是其矩阵的行列式 □

2.

命题 1.1.5. $\forall \boldsymbol{A} \in M_n(F), \chi_{\boldsymbol{A}} \in F[\lambda]$

证明： 设 $\boldsymbol{A} = (a_{ij}) \in M_n(F)$

那么，$\lambda \boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix}$

由行列式的置换展开可知，$\chi_{\boldsymbol{A}}(\lambda) \in F[\lambda]$ □

推论 1.1.2

$\forall A \in \hom(V,V), \dim V = n < \aleph_0, \chi_A(\lambda) \in F[\lambda]$

证明： 这是显然的，因为我们只需要任取一个基 $\{\alpha_1,\cdots,\alpha_n\}$，并利用 A 在此基下的矩阵 $\boldsymbol{A}$

利用前面线性映射与其矩阵的特征多项式相同的命题，即可得证。 □

3.

定理 1.1.3: 线性变换的特征值即是特征多项式在域内的根

设 V 是一个 F 上的 n 维线性空间，$A\in\mathrm{hom}(V,V)$

那么：λ 是 A 的一个特征值，$\alpha\in V_\lambda\Leftrightarrow\chi_A(\lambda)=0,\alpha\in\ker(\lambda I-A)$

证明： λ 是 A 的一个特征值，$\alpha\in V_\lambda$

$\Leftrightarrow A(\alpha)=\lambda\alpha,V_\lambda\neq\{\mathbf{0}\}$

$\Leftrightarrow(\lambda I-A)(\alpha)=\mathbf{0},\ker(\lambda I-A)\neq\{\mathbf{0}\}$

$\Leftrightarrow\mathrm{rank}(\lambda I-A)<n,\alpha\in\ker(\lambda I-A)$

$\Leftrightarrow\chi_A(\lambda)=0,\alpha\in\ker(\lambda I-A)$ □

其自然推论是其矩阵版本：

推论 1.1.4

设 $\boldsymbol{A}\in M_n(F)$

那么：λ 是 $\boldsymbol{A}$ 的一个特征值，$\mathbf{x}\in F^n$ 是 $\boldsymbol{A}$ 从属于 λ 的特征向量 $\Leftrightarrow\chi_{\boldsymbol{A}}(\lambda)=0,(\lambda\boldsymbol{I}-\boldsymbol{A})\mathbf{x}=\mathbf{0},\mathbf{x}\neq\mathbf{0}$

这个推论也指出了一下结果：

推论 1.1.5

设 $\boldsymbol{A},\boldsymbol{B}\in M_n(F),\boldsymbol{A}\sim\boldsymbol{B}$

那么 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 具有相同的特征值，并且 $\chi_{\boldsymbol{A}}(\lambda)=\chi_{\boldsymbol{B}}(\lambda)$

我们之后不再完全讨论矩阵版本的特征值相关定理，因为上面的结果已经说明了：相似矩阵，以及相互对应的矩阵和映射在特征值理论中并无区别

4.

命题 1.1.6. 设 V 是一个 F 上的 n 维线性空间，$A\in\mathrm{hom}(V,V)$

那么：

$$\chi_A(\lambda) = \lambda^n - \operatorname{tr}(A)\lambda^{n-1} + \cdots + (-1)^{n-k}\left(\sum_{j_1<\cdots<j_{n-k}} A^{j_1,\cdots,j_{n-k}}_{j_1,\cdots,j_{n-k}}\right)\lambda^k + \cdots + (-1)^n \det(A) \tag{1.4}$$

证明： 事实上，λ^k 项即是下面的行列式求和：$\sum\limits_{j_1'<\cdots<j_k'} \begin{vmatrix} -a_{11} & \cdots & 0 & \cdots & 0 & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ -1 & \cdots & \lambda & \cdots & 0 & \cdots & 1 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ -1 & \cdots & 0 & \cdots & \lambda & \cdots & 1 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & 0 & \cdots & 0 & \cdots & -a_{nn} \end{vmatrix}$ 其中行列式的 $j_1',\cdots,j_k'$ 列为仅在第 j_i 个元素为 λ，其他位置都是 0 的列。

这是因为，如果想要产生 λ^k，那么求和中的每一项必须选取对角线上的 k 个元素；同时，如果在进一步展开中选取了 $-a_{ii}$ 而不是 λ，那么会导致次数降低

注意到：$\begin{vmatrix} -a_{11} & \cdots & 0 & \cdots & 0 & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ -1 & \cdots & \lambda & \cdots & 0 & \cdots & 1 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ -1 & \cdots & 0 & \cdots & \lambda & \cdots & 1 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & 0 & \cdots & 0 & \cdots & -a_{nn} \end{vmatrix}$

$= (-1)^{(j_1'+\cdots+j_k')+(j_1'+\cdots+j_k')} \det(diag\{\lambda,\cdots,\lambda\})(-1)^{n-k} A^{j_1,\cdots,j_{n-k}}_{j_1,\cdots,j_{n-k}}$

这个结果只需要对 $j_1',\cdots,j_k'$ 列展开即可得出，其中 $(-1)^{n-k}$ 是将 $-a_{ij}$ 前面的负号提出去得到的，而 $j_1,\cdots,j_{n-k}$ 是与 $j_1',\cdots,j_k'$ 互补的有限序列

于是命题得证。 □

其推论是以下结论：

推论 1.1.6

设 F 是一个代数闭域，V 是一个 F 上的 n 维线性空间，$\mathcal{A} \in \hom(V, V)$

设 $\lambda_1, \cdots, \lambda_n$ 是 $\mathcal{A}$ 的特征值（$\chi_{\mathcal{A}}(\lambda)$ 中的重根按重数计算）

那么有：

$$\prod_{i=1}^{n} \lambda_i = \det(\mathcal{A}) \tag{1.5}$$

$$\sum_{i=1}^{n} \lambda_i = \operatorname{tr}(\mathcal{A}) \tag{1.6}$$

证明： 对 $\chi_{\mathcal{A}}(\lambda)$ 作唯一分解：

$\chi_{\mathcal{A}}(\lambda) = \prod\limits_{i=1}^{n} (\lambda - \lambda_i)$

$= \lambda^n - (\lambda_1 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n \prod_{i=1}^{n} \lambda_i$

结合前面的命题既得。 □

请注意：这个命题必须要求 F 代数闭，否则可能 $\chi_{\mathcal{A}}(\lambda)$ 的根并非全体特征值，而导致论证失效。

1.2 线性变换的对角标准型

1.2.1 对角标准型的定义

定义 1.2.1: 线性映射的对角标准型

设 V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

如果存在一个基 $\{\alpha_1, \cdots, \alpha_n\}$，使 $\mathcal{A}$ 在其上的矩阵 $\boldsymbol{A}$ 是一个对角矩阵

那么我们称 $\boldsymbol{A}$ 为 $\mathcal{A}$ 的对角标准型，此时称 $\mathcal{A}$ 可对角化。

类似地，我们可以定义矩阵的对角标准型

定义 1.2.2: 矩阵的对角标准型

设 $\boldsymbol{A} \in M_n(F)$，如果存在 $\boldsymbol{P} \in GL_n(F)$ 和对角矩阵 $\boldsymbol{D} \in M_n(F)$，使得

$\boldsymbol{A} = \boldsymbol{P}^{-1}\boldsymbol{D}\boldsymbol{P}$

那么我们称 $\boldsymbol{D}$ 为 $\boldsymbol{A}$ 的对角标准型，此时称 $\boldsymbol{A}$ 可对角化。

容易看出，线性映射的对角标准型与矩阵的对角标准型是等价的。特别地，如果认为 $\boldsymbol{A}$

是 $A \in \hom(F^n, F^n)$ 在 $\{\mathbf{e}_i\}$ 下的矩阵，那么容易发现：其实 $\boldsymbol{P}^{-1}$ 的列向量就是 A 的特征向量。

1.2.2 线性变换可对角化的条件

观察对角矩阵的结构，既得以下显然的条件：

定理 1.2.1: 线性映射可对角化的条件 (1)

设 V 是一个 F 上的一个线性空间，$\dim V = n < \aleph_0$

那么：$A \in \hom(V, V)$ 可对角化 $\Leftrightarrow$ 存在一个由 A 的特征向量组成的 V 的基

证明： $A \in \hom(V, V)$ 可对角化

$\Leftrightarrow$ 存在一个基 $\{\alpha_1, \cdots, \alpha_n\}$，使得 A 在其上的矩阵为 $diag\{\lambda_1, \cdots, \lambda_n\}$

$\Leftrightarrow A(\alpha_1, \cdots, \alpha_n) = diag\{\lambda_1, \cdots, \lambda_n\}(\alpha_1, \cdots, \alpha_n)$

$\Leftrightarrow A(\alpha_1, \cdots, \alpha_n) = (\lambda_1\alpha_1, \cdots, \lambda_n\alpha_n)$

$\Leftrightarrow \lambda_i$ 是 A 的特征值，$\alpha_i \in V_{\lambda_i}$

$\Leftrightarrow$ 存在一个由 A 的特征向量组成的 V 的基 □

推论 1.2.2: 线性映射可对角化的条件 (2)

1.3 线性变换的不变子空间

1.4 线性变换的最小多项式

1.5 幂零变换的 Jordan 标准型

1.6 线性变换的 Jordan 标准型