

§3 特征值和特征向量

程光辉

2020 年 3 月 14 日

1 线性变换

定义 1 若令 V 和 W 分别是 C^m 和 C^n 的子空间, 则

$$T: V \rightarrow W$$

称为子空间 V 到子空间 W 的变换 (或函数、映射), 它表示将子空间 V 的每一个向量变成子空间 W 的一个对应向量的一种规则. 即, 对任一 $v \in V$, 则存在 $w \in W$, 有

$$w = T(v),$$

并称子空间 V 是变换 T 的始集或定义域, 称 W 是变换的终集或上域.

定义 2 若令 V 和 W 分别是 C^m 和 C^n 的子空间, 并且 $T: V \rightarrow W$ 是一种变换. 若对于 $\forall v_1, v_2 \in V$ 和任意标量 $c \in C$, 变换 T 满足叠加性

$$T(v_1 + v_2) = T(v_1) + T(v_2)$$

和齐次性

$$T(cv_1) = cT(v_1),$$

则称 T 是线性变换或线性映射.

例 1 验证如下映射都是 R^3 上的线性变换:

(1) 坐标面上投影: $\sigma_{xy}: R^3 \rightarrow R^3, (x, y, z) \rightarrow (x, y, 0)$;

(2) 关于坐标面的镜面反射: $\tau_{xy}: R^3 \rightarrow R^3, (x, y, z) \rightarrow (x, y, -z)$.

例 2 下面几个变换是否是线性变换?

$$(1) T \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = \begin{bmatrix} y - z \\ x \\ y \end{bmatrix},$$

$$(2) \quad T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} x+1 \\ z \\ y \end{bmatrix},$$

$$(3) \quad T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} xy \\ -y \\ z \end{bmatrix},$$

$$(4) \quad T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{bmatrix} x+y \\ z \end{bmatrix}.$$

若 T 是线性空间 $V_n(\mathbb{C})$ 的一种线性变换, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是线性空间 $V_n(\mathbb{C})$ 的一组基底, 则

$$T(\varepsilon_j) = \sum_{i=1}^n a_{ij} \varepsilon_i = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) a_j, \quad j = 1, 2, \dots, n,$$

其中 $a_j = (a_{1j}, a_{2j}, \dots, a_{nj})^T \in \mathbb{C}^n$, 则

$$\left(T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_n) \right) = \left(\sum_{i=1}^n a_{i1} \varepsilon_i, \sum_{i=1}^n a_{i2} \varepsilon_i, \dots, \sum_{i=1}^n a_{in} \varepsilon_i \right) = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) A,$$

其中 $A = (a_1, a_2, \dots, a_n) \in \mathbb{C}^{n \times n}$, 称 A 是线性变换 T 在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下对应的矩阵. 显然, 线性变换 T 在基 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 下和对应的矩阵 A 之间是一种 $1 \leftrightarrow 1$ 关系.

2 特征值和特征向量的概念

定义 3 设 T 是线性空间 $V_n(\mathbb{C})$ 的一种线性变换, 如果存在 $\lambda \in \mathbb{C}$ 和非零向量 $u \in V_n(\mathbb{C})$, 使 $T(u) = \lambda u$, 则 λ 叫做 T 的特征值, u 叫做 T 的属于特征值 λ 的特征向量.

若 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ 是线性空间 $V_n(\mathbb{C})$ 的一组基底, 对线性变换 T 的特征值 λ , 有

$$u = \sum_{i=1}^n x_i \varepsilon_i = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) x,$$

其中 $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{C}^n$, 则

$$T(u) = \sum_{i=1}^n x_i T(\varepsilon_i) = \left(T(\varepsilon_1), T(\varepsilon_2), \dots, T(\varepsilon_n) \right) x = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) A x,$$

和

$$\lambda u = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) \lambda x,$$

于是有

$$A x = \lambda x.$$

定义 4 设 $A \in \mathbb{C}^{n \times n}$, 如果存在 $\lambda \in \mathbb{C}$ 和非零向量 $u \in \mathbb{C}^n$, 使 $Au = \lambda u$, 则 λ 叫做 A 的特征值, u 叫做 A 的属于特征值 λ 的特征向量.

从定义可以看出, 使用矩阵 A 对向量 u 所作的线性变换 Au 不改变向量 u 的方向. 因此, 线性变换是一种“保持方向不变”的变换. 为了确定向量 u , 可将定义式改写为

$$(\lambda E - A)u = 0.$$

由于要求 $u \neq 0$, 则矩阵 $\lambda E - A$ 奇异, 即行列式

$$\det(\lambda E - A) = 0,$$

称矩阵 $(\lambda E - A)$ 为 A 的特征矩阵, $\det(\lambda E - A) = 0$ 为 A 的特征方程, $\det(\lambda E - A)$ 为 A 的特征多项式.

特征方程的根即为特征值.

定义 5 若 λ 是特征多项式 $\det(xE - A) = 0$ 的 μ 重根, 则称特征值 λ 的代数重复度为 μ ; 若与特征值 λ 对应的线性无关特征向量个数为 γ , 则称特征值 λ 的几何重复度为 γ .

特征值和特征向量的几个性质:

- (1) $\det(A) = \prod_{i=1}^n \lambda_i$;
- (2) $\text{tr}(A) = \sum_{i=1}^n \lambda_i$;
- (3) 同一特征值的几何重复度小于等于代数重复度;
- (4) A 和 A^T 的特征值相同;
- (5) 矩阵 $A + \sigma E$ 的特征值为 $\lambda + \sigma$;
- (6) 不同特征值对应的特征向量线性无关.

3 矩阵相似

定义 6 设 $A, B \in \mathbb{C}^{n \times n}$, 如果存在可逆矩阵 $P \in \mathbb{C}^{n \times n}$ 使得 $P^{-1}BP = A$, 则称矩阵 A 相似于矩阵 B .

定理 1 若 A 与 B 相似, 则 A 与 B 有相同的特征值.

定义 7 *Jordan* 标准型

$$J = \begin{bmatrix} J_{n_1} & O & \cdots & O \\ O & J_{n_2} & \cdots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & J_{n_r} \end{bmatrix} = \text{diag}(J_{n_1}, J_{n_2}, \cdots, J_{n_r}),$$

其中 J_{n_i} 为 n_i 阶对角元素都为 λ_i , λ_i 正上方紧邻元素为 1, 其余都为 0 的方阵.

定理 2 设 $A \in \mathbb{C}^{n \times n}$ 有 r 个不同的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_r$, 其代数重数分别为 n_1, n_2, \cdots, n_r , 则必存在可逆矩阵 $P \in \mathbb{C}^{n \times n}$, 使得

$$P^{-1}AP = J = \text{diag}(J_{n_1}(\lambda_1), \cdots, J_{n_r}(\lambda_r)),$$

矩阵 J 叫做 A 的 *Jordan* 标准型.