

线性映射表示的是两个线性空间之间的映射，而线性变换则是线性空间到自身的映射，它的矩阵表示具有更多性质。线性变换对应的矩阵表示为方阵，它最重要的性质就是在不同基下的表示矩阵的相似关系，以及在相似变换下的最简矩阵表示。本章详细分析了矩阵的特征值和特征向量的相关理论，给出了矩阵的若尔当标准形的基本理论和计算问题。

3.1 线性变换及表示

3.1.1 线性变换的定义

线性空间 V 到自身的映射称为 V 的一个变换。

定义 3.1.1 线性空间 V 的一个变换 T 称为**线性变换**，若对于 V 中任意的元素 $\boldsymbol{x}, \boldsymbol{y}$ 和数域 $\mathcal{F}$ 中任意数 k, l ，都有

$$T(k\boldsymbol{x} + l\boldsymbol{y}) = k \cdot T(\boldsymbol{x}) + l \cdot T(\boldsymbol{y}) \quad (3.1)$$

$T(\boldsymbol{x})$ 或 $T\boldsymbol{x}$ 代表元素 $\boldsymbol{x}$ 在变换 T 下的像。式 (3.1) 所表示的性质，有时也说成线性变换保持向量的加法与数量乘法。从线性变换的定义可以知道，它也是一个线性映射，线性映射具有的性质，线性变换都具有，因此这里就不重复了。

例题 3.1 平面上的向量构成实数域上的二维线性空间。把平面围绕坐标原点按反时针方向旋转 θ 角，就是一个线性变换，用 g_θ 表示。如果平面上一个向量 $\boldsymbol{x}$ 在直角坐标系下的坐标是 (x, y) ，那么像 $g_\theta(\boldsymbol{x})$ 的坐标，即 $\boldsymbol{x}$ 旋转 θ 角之后的坐标 $(x', y')^T$ 是

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

例题 3.2 设 $\boldsymbol{x}$ 是几何空间 $\mathbb{C}^n$ 中一固定非零列向量，把每个列向量 $\boldsymbol{\xi}$ 变到它在 $\boldsymbol{x}$ 上的内射影的变换也是一个线性变换，以 $P_{\boldsymbol{x}}$ 表示它。用公式表示就是

$$P_{\boldsymbol{x}}(\boldsymbol{\xi}) = \frac{\boldsymbol{x}^H \boldsymbol{\xi}}{\boldsymbol{x}^H \boldsymbol{x}} \boldsymbol{x}$$

这表示 $P_{\boldsymbol{x}}$ 投影变换。

例题 3.3 线性空间 V 中的恒等变换或称单位变换 I , 即

$$I(\mathbf{x}) = \mathbf{x}, \quad \mathbf{x} \in V$$

例题 3.4 在多项式线性空间 $P[x]$ 或者 $P_n[x]$ 中, 求导数是一个线性变换, 通常用 D 代表, 即

$$D(f(x)) = f'(x)$$

例题 3.5 定义在闭区间 $[a, b]$ 上的全体连续函数组成实数域上一线性空间, 以 $C(a, b)$ 代表。在这个空间中变换

$$\mathcal{J}(f(x)) = \int_a^x f(t)dt$$

是一线性变换。

若向量组 $\{T(\mathbf{x}_1), T(\mathbf{x}_2), \dots, T(\mathbf{x}_r)\}$ 线性无关, 那么向量组 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$ 线性无关, 反之未必。

3.1.2 线性变换的运算

设 T, S 是线性空间 $V(\mathcal{F})$ 的两个线性变换, 定义加法和数乘如下:

$$(T + S)(\mathbf{x}) = T(\mathbf{x}) + S(\mathbf{x})$$

$$(kT)(\mathbf{x}) = k(T(\mathbf{x}))$$

其中 $\mathbf{x} \in V, k \in \mathcal{F}$ 。

显然, 这样定义的线性变换加法与数乘的结果仍是线性变换。因此它们的全体按照这样的加法和数乘。

定理 3.1.1 线性空间 V 上全体线性变换, 对于如上定义的加法与数量乘法, 也构成数域 $\mathcal{F}$ 上一个线性空间。

定义 3.1.2 设 T, S 是线性空间 V 的两个线性变换, 定义它们乘积为

$$(TS)(\mathbf{x}) = T(S(\mathbf{x})), \quad \mathbf{x} \in V$$

显然线性变换的乘积也是线性变换。线性变换的乘法适合结合律, 即

$$(TS)P = T(SP)$$

但线性变换的乘法不适合交换律。例如, 在实数域上的线性空间中, 线性变换

$$D(f(x)) = f'(x), \quad \mathcal{J}(f(x)) = \int_a^x f(t)dt$$

的乘积 $D\mathcal{J} = I$, 但一般 $\mathcal{J}D \neq I$ 。

定义 3.1.3 V 的变换 T 称为可逆的, 如果有 V 的变换 S 存在, 使

$$TS = ST = I$$

这时, 变换 S 称为 T 的逆变换, 记为 T^{-1} 。如果线性变换 T 是可逆的, 那么它的逆变换 T^{-1} 也是线性变换。

既然线性变换的乘法满足结合律, 当若干线性变换 T 重复相乘时, 其最终结果是完全确定的, 与乘法的结合方法无关。因此当 n 个 (n 是正整数) 线性变换 T 相乘时, 就可以用

$$\overbrace{TT \cdots T}^n$$

来表示, 称为 T 的 n 次幂, 简记为 T^n . 作为定义, 令 $T^0 = I$.

例题 3.6 设 $f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0$ 是 $P[x]$ 中一多项式, T 是 V 的一个线性变换, 定义

$$f(T) = a_m T^m + a_{m-1} T^{m-1} + \cdots + a_0 I$$

显然 $f(T)$ 是一线性变换. 它称为线性变换 T 的多项式.

例题 3.7 在线性空间 $P_{n-1}[\lambda]$ 中, 求微商是一个线性变换, 用 D 表示. 同样平移变换 $f(\lambda) \rightarrow f(\lambda + a)$, $a \in P$ 也是一个线性变换, 用 T_a 表示. 根据泰勒展开式 $f(\lambda + a) = f(\lambda) + af'(\lambda) + \frac{a^2}{2!} f''(\lambda) + \cdots + \frac{a^{n-1}}{(n-1)!} f^{(n-1)}(\lambda)$, T_a 实质上是 D 的多项式:

$$T_a = I + aD + \frac{a^2}{2!} D^2 + \cdots + \frac{a^{n-1}}{(n-1)!} D^{n-1}$$

3.1.3 线性变换的矩阵表示

定义 3.1.4 设 $[x_1, x_2, \cdots, x_n]$ 是数域 $\mathcal{F}$ 上 n 维线性空间 V 的一个基, T 是 V 中的一个线性变换. 基向量的像可以被基线性表出:

$$\begin{cases} Tx_1 = a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n \\ Tx_2 = a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n \\ \vdots \\ Tx_n = a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n \end{cases}$$

用矩阵表示就是

$$TX = XA$$

其中 $X = [x_1, x_2, \cdots, x_n]$, $TX = [Tx_1, Tx_2, \cdots, Tx_n]$,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

方阵 A 称为线性变换 T 在基 $[x_1, x_2, \cdots, x_n]$ 下的表示矩阵.

从上式可以看出线性变换完全被它在一个基上的作用所决定. 在取定一个基之后, 就建立了由数域 $\mathcal{F}$ 上的 n 维线性空间 V 的线性变换到数域 $\mathcal{F}$ 上的 $n \times n$ 矩阵的一个映射, 并且这是一个双射. 这个对应的重要性表现在它保持运算, 可归纳为如下结论.

定理 3.1.2 设 $X = [x_1, x_2, \cdots, x_n]$ 是数域 $\mathcal{F}$ 上 n 维线性空间 V 的一个基. 在这个基下, V 中线性变换 T 按对应一个数域 $\mathcal{F}$ 上的 $n \times n$ 矩阵 A , 即 $TX = XA$, 且这个对应具有以下性质:

- (1) 线性变换的和对应于矩阵的和;
- (2) 线性变换的复合对应于矩阵的乘积;
- (3) 线性变换的数量乘积对应于矩阵的数量乘积;
- (4) 可逆的线性变换与可逆矩阵对应, 且逆变换对应于逆矩阵。

此定理说明数域 $\mathcal{F}$ 上 n 维线性空间 V 的全体线性变换组成的集合 $L(V)$ 对于线性变换的加法与数量乘法构成 $\mathcal{F}$ 上一个线性空间, 与数域 $\mathcal{F}$ 上 n 级方阵构成的线性空间 $\mathcal{F}^{n \times n}$ 同构。

例题 3.8 设线性空间 V 中的线性变换 T 在基 X 下的表示矩阵为 A , 那么线性变换多项式

$$f(T) = a_0 T_m + a_1 T_{m-1} + \cdots + a_{m-1} T + a_m I$$

(此处 I 表示恒等变换) 在基 X 下的矩阵表示为矩阵多项式:

$$f(A) = a_0 A^m + a_1 A^{m-1} + \cdots + a_{m-1} A + a_m I_n$$

其中 I_n 表示单位矩阵, n 为线性空间 V 的维数。

定理 3.1.3 设线性空间 V 中的线性变换 T 在基 $X = [x_1, x_2, \cdots, x_n]$ 下的表示矩阵是 A , 向量 x 在基 X 下的坐标是 $\xi = (\xi_1, \xi_2, \cdots, \xi_n)^T$, 则 $T(x)$ 在基 X 下的表示坐标 $\eta = (\eta_1, \eta_2, \cdots, \eta_n)^T$ 为 $\eta = A\xi$ 。

证明 由题设 $x = X\xi$, 由此可得 $TX = XA$ 。 $X\eta = T(x) = T(X\xi) = TX(\xi) = XA(\xi) = X(A\xi)$, 因此 $\eta = A\xi$ 。

线性变换的表示矩阵与空间中基的选择有关, 随着基的改变, 同一个线性变换就有不同的矩阵表示。为了利用矩阵来研究线性变换, 有必要弄清楚线性变换的表示矩阵是如何随着基的改变而改变的。

定理 3.1.4 设线性空间 V 中线性变换 T 在两个基 $X = [x_1, x_2, \cdots, x_n]$ 和 $Y = [y_1, y_2, \cdots, y_n]$ 的表示矩阵分别为 A 和 B 。若基 X 和 Y 之间的过渡矩阵为 C , 即 $Y = XC$, 那么 $B = C^{-1}AC$ 。

证明 由线性变换的矩阵表示理论可知, $TY = YB$, 同时 $TY = T(XC) = (TX)C = (XA)C = X(AC) = (YC^{-1})(AC) = Y(C^{-1}AC)$ 因此, 我们有 $B = C^{-1}AC$ 成立。

定义 3.1.5 设 A, B 为数域 $\mathcal{F}$ 上两个 n 阶方阵, 如果存在数域 $\mathcal{F}$ 上的 n 阶可逆方阵 C , 使得 $B = C^{-1}AC$, 就说 A 相似于 B , 记作 $A \sim B$ 。

相似是矩阵之间的一种等价关系, 这种关系具有下面三个性质:

- (1) 反身性: $A \sim A$ 。
- (2) 对称性: 如果 $A \sim B$, 那么 $B \sim A$ 。
- (3) 传递性: 如果 $A \sim B, B \sim C$, 那么 $A \sim C$ 。

所以, 有时也称相似关系为相似等价关系。它和 1.4 节定义的矩阵相抵等价关系有区别。对于线性空间 $V(\mathcal{F})$, 可以定义几个特殊的变换:

零变换 T_0 : $T_0(x) = 0$ 。

恒等变换 I : $I(x) = x$ 。

数乘变换 T_k : $T_k(x) = k \cdot x$ 。

例题 3.9 设 $f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_0$ 是 $\mathcal{F}[x]$ 中的多项式, 数域 $\mathcal{F}$ 上的

两个矩阵 A 与 B 相似, 即 $B = C^{-1}AC$ 。那么 $f(B) = C^{-1}f(A)C$ 。

这个公式是利用相似变换计算矩阵多项式的基础。有时直接计算 $f(B)$ 比较困难, 可以化为相似的对角或块对角矩阵 A , 从而计算 $f(A)$ 。

从这儿我们可以看出, 在相似等价意义下具有的性质有时也称为线性变换的性质。例如, 相似的矩阵具有相同的行列式, 所以我们可以认为线性变换的对应矩阵的行列式为线性变换对原空间的单位超立方体经变换后的多面体的体积。这正是多元积分的变量替换的雅可比行列式。又如我们知道, 相似的矩阵具有相同的特征多项式, 所以我们可以定义线性变换的特征多项式。同样可以导出线性变换的特征值和线性变换的迹 (定义为线性变换对应矩阵所有特征值的和)。

3.2 特征值与特征向量的基础

3.2.1 线性变换的特征值与特征向量

由于线性变换在不同基下的矩阵表示不相同, 那么一个很自然的问题就是, 怎样选择特殊的基使得给定线性变换在该基下的表示矩阵最简单。为此引入下列概念。

定义 3.2.1 设 T 是线性空间 $V(\mathcal{F})$ 上线性变换, 若存在非零向量 $x \in V$ 和数 $\lambda \in \mathcal{F}$ 满足 $Tx = \lambda x$, 则称 λ 为 T 的特征值, x 为相应的特征向量。

注意: 由定义, 特征向量非零。

若线性变换 T 在线性空间的一个基 X 下的矩阵表示为 A , 即

$$TX = [Tx_1, Tx_2, \dots, Tx_n] = XA$$

设 T 的特征值 λ_0 对应的特征向量 x 在基 X 的坐标表示为 ξ , 即

$$x = \xi_1 x_1 + \xi_2 x_2 + \dots + \xi_n x_n = X(\xi_1, \xi_2, \dots, \xi_n)^T = X\xi$$

那么有 $X(\lambda_0 \cdot \xi) = \lambda_0 x = Tx = T(X\xi) = (TX)\xi = X(A\xi)$ 。从而 $\lambda_0 \cdot \xi = A\xi$ 。由于上述等式推导可以倒推, 因此求线性变换 T 的特征值 λ_0 和特征向量 x 等价于在一个基 X 下求 T 的表示矩阵 A 的特征值 λ_0 和特征向量 ξ 。这时 ξ 为 x 在基 X 下的坐标, 即 $x = X\xi$ 。因此, 这个关系式将线性变换的特征值和特征向量的计算问题归结为求矩阵特征值与特征向量的计算。

定义 3.2.2 设 A 是数域 $\mathcal{F}$ 上一个 n 阶矩阵, λ 是一个数字。那么矩阵 $\lambda I_n - A$ 称为特征矩阵, 它的行列式

$$|\lambda I_n - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \quad (3.2)$$

叫作方阵 A 的特征多项式 $\varphi_A(\lambda)$, 这是数域 $\mathcal{F}$ 上的一个 n 次多项式。

上面的分析说明, 如果 λ_0 是线性变换 T 的特征值, 那么 λ_0 一定是其表示矩阵 A 的特

征多项式的一个根；反过来，如果 λ_0 是其表示矩阵 $\mathbf{A}$ 的特征多项式在数域 $\mathcal{F}$ 中的一个根，即 $|\lambda_0 \mathbf{E}_n - \mathbf{A}| = 0$ ，那么齐次线性方程组 $(\lambda_0 \mathbf{I}_n - \mathbf{A})\boldsymbol{\xi} = \mathbf{0}$ 就有非零解。这时， $\mathbf{x} = \xi_1 \mathbf{x}_1 + \xi_2 \mathbf{x}_2 + \cdots + \xi_n \mathbf{x}_n$ 为特征值 λ_0 的特征向量，即 $T\mathbf{x} = \lambda_0 \mathbf{x}$ 。因此线性变换 T 的一个特征值与特征向量的方法可以分成以下几步：

(1) 在线性空间 V 中取一个基 $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ ，写出 T 在这个基下的表示矩阵 $\mathbf{A}$ ；

(2) 求出 $\mathbf{A}$ 的特征多项式 $|\lambda \mathbf{I} - \mathbf{A}|$ 在数域 $\mathcal{F}$ 中全部的根，它们也就是线性变换 T 的全部特征值；

(3) 把所求得特征值 λ_0 逐个代入方程组 $((\lambda_0 \mathbf{I}_n - \mathbf{A})\boldsymbol{\xi} = \mathbf{0})$ ，对于每一个特征值，解方程组 $((\lambda_0 \mathbf{I}_n - \mathbf{A})\boldsymbol{\xi} = \mathbf{0})$ ，求出一个基础解系，它们就是属于这个特征值的几个线性无关的特征向量在基 $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ 下的坐标，这样，也就求出了属于每个特征值的全部线性无关的特征向量。

例题 3.10 求 $\mathbf{A} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ 的特征值和特征向量。

解： $\mathbf{A}$ 的特征多项式为

$$|\mathbf{A} - \lambda \mathbf{I}_2| = \begin{vmatrix} 3 - \lambda & -1 \\ -1 & 3 - \lambda \end{vmatrix} = (3 - \lambda)^2 - 1 = 8 - 6\lambda + \lambda^2 = (2 - \lambda)(4 - \lambda)$$

于是解得 $\mathbf{A}$ 的特征值为 $\lambda_1 = 2, \lambda_2 = 4$ 。下面分别求特征向量：

对于 $\lambda_1 = 2$ ，解齐次线性方程组

$$(\mathbf{A} - 2\mathbf{I}_2)\mathbf{x} = \mathbf{0}, \begin{pmatrix} 3 - 2 & -1 \\ -1 & 3 - 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得基础解系 $\mathbf{p}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ，因此属于 $\lambda_1 = 2$ 的全部特征向量为 $k_1 \mathbf{p}_1$ ($k_1 \neq 0$)。

对于 $\lambda_2 = 4$ ，解齐次线性方程组

$$(\mathbf{A} - 4\mathbf{I}_2)\mathbf{x} = \mathbf{0}, \begin{pmatrix} 3 - 4 & -1 \\ -1 & 3 - 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得基础解系 $\mathbf{p}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ，因此属于 $\lambda_2 = 4$ 的全部特征向量为 $k_2 \mathbf{p}_2$ ($k_2 \neq 0$)。

例题 3.11 求 $\mathbf{A} = \begin{pmatrix} -2 & 1 & 1 \\ 0 & 2 & 0 \\ -4 & 1 & 3 \end{pmatrix}$ 的特征值和特征向量。

解： $\mathbf{A}$ 的特征多项式为

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} -2-\lambda & 1 & 1 \\ 0 & 2-\lambda & 0 \\ -4 & 1 & 3-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} -2-\lambda & 1 \\ -4 & 3-\lambda \end{vmatrix} = -(\lambda+1)(\lambda-2)^2$$

所以 $\mathbf{A}$ 的特征值为 $\lambda_1 = -1, \lambda_2 = \lambda_3 = 2$ (二重根)。

$$\text{当 } \lambda_1 = -1 \text{ 时, 解 } (\mathbf{A} + \mathbf{I})\mathbf{x} = \mathbf{0}, \text{ 由 } \mathbf{A} + \mathbf{I} = \begin{pmatrix} -1 & 1 & 1 \\ 0 & 3 & 0 \\ -4 & 1 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \text{ 得}$$

$$\mathbf{p}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{ 为基础解系, 则属于 } \lambda_1 = -1 \text{ 的所有特征向量为 } k_1 \mathbf{p}_1 \text{ } (k_1 \neq 0)。$$

$$\text{当 } \lambda_2 = \lambda_3 = 2 \text{ 时, 解 } (\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}, \text{ 由 } \mathbf{A} - 2\mathbf{I} = \begin{pmatrix} -4 & 1 & 1 \\ 0 & 0 & 0 \\ -4 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} -4 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{得基础解系 } \mathbf{p}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \mathbf{p}_3 = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}, \text{ 则属于 } \lambda_2 = \lambda_3 = 2 \text{ 的全部特征向量为 } k_2 \mathbf{p}_2 + k_3 \mathbf{p}_3 \text{ } (k_2, k_3 \text{ 不全为零})。$$

例题 3.12 在空间 $P[x]_n$ 中, 线性变换 $Df(x) = f'(x)$ 在基 $1, x, \frac{x^2}{2!}, \dots, \frac{x^{n-1}}{(n-1)!}$ 下的矩阵是

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

$$D \text{ 的特征多项式是 } |\lambda \mathbf{I} - \mathbf{D}| = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & -1 \\ 0 & 0 & 0 & \cdots & \lambda \end{vmatrix} = \lambda^n. \text{ 因此, } D \text{ 的特征值只有 } 0. \text{ 通}$$

过解相应的齐次线性方程组知道, 属于特征值 0 的线性无关的特征向量组只能是任一非零常数。这表明导数为零的多项式只能是零或非零的常数。

例题 3.13 平面上全体向量构成实数域上一个二维线性空间, 旋转 g_θ 在直角坐标系下的矩阵为 $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ 。它的特征多项式为 $\begin{vmatrix} \lambda - \cos \theta & \sin \theta \\ -\sin \theta & \lambda - \cos \theta \end{vmatrix} = \lambda^2 - 2\lambda \cos \theta + 1$ 。

当 $\theta \neq k\pi$ 时, 特征多项式无实根。因之, 当 $\theta \neq k\pi$ 时, g_θ 无实特征值。从几何上看, 这个结论是明显的。

3.2.2 线性变换的特征子空间

定义 3.2.3 对于线性变换 T 的任一个特征值 λ_0 , 特征子空间 V_{λ_0} 记为

$$V_{\lambda_0} = \{\mathbf{x} \mid T\mathbf{x} = \lambda_0\mathbf{x}, \mathbf{x} \in V^n\}$$

它的维数就是属于 λ_0 的线性无关的特征向量的最大个数, 称为 λ_0 的几何重数。相应地, λ_0 为 T 的特征多项式根的重数称为 λ_0 的代数重数。

在线性变换的研究中, 矩阵的特征多项式是重要的。下面先来看一下它的系数。在

$$|\lambda \mathbf{I}_n - \mathbf{A}| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

的展开式中, 有一项是主对角线上元素的连乘积

$$(\lambda - a_{11})(\lambda - a_{22}) \cdots (\lambda - a_{nn})$$

展开式中的其余项, 至多包含 $n-2$ 个主对角线上的元素, 它对 λ 的次数最多是 $n-2$ 。因此特征多项式中含 λ 的 n 次与 $n-1$ 次的项只能在主对角线上元素的连乘积中出现, 它们是

$$\lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1}$$

在特征多项式中令 $\lambda = 0$, 即得常数项 $|\mathbf{A}| = (-1)^n |\mathbf{A}|$ 因此, 如果只写特征多项式的前两项与常数项, 就有

$$|\lambda \mathbf{I} - \mathbf{A}| = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (-1)^n |\mathbf{A}| \quad (3.3)$$

由根与系数的关系, $\mathbf{A}$ 的全体特征值的和为 $a_{11} + a_{22} + \cdots + a_{nn}$ (称为 $\mathbf{A}$ 的迹), 而 $\mathbf{A}$ 的全体特征值的积为 $|\mathbf{A}|$ 。

特征值自然是被线性变换所决定的。但是在有限维空间中, 任取一个基后, 特征值就是线性变换在这个基下矩阵的特征多项式的根。随着基的不同, 线性变换的矩阵一般是不同的, 但是这些矩阵是相似的。

定理 3.2.1 相似方阵有相同的特征多项式。

证明 设 $\mathbf{B} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}$, 那么 $\varphi_{\mathbf{B}}(\lambda) = |\lambda \mathbf{I}_n - \mathbf{B}| = |\lambda \mathbf{I}_n - \mathbf{C}^{-1}\mathbf{A}\mathbf{C}| = |\mathbf{C}^{-1}(\lambda \mathbf{I}_n - \mathbf{A})\mathbf{C}| = |\mathbf{C}^{-1}| \cdot |\lambda \mathbf{I}_n - \mathbf{A}| \cdot |\mathbf{C}| = |\lambda \mathbf{I}_n - \mathbf{A}| = \varphi_{\mathbf{A}}(\lambda)$ 。

定理3.2.1说明, 线性变换的表示矩阵的特征多项式与基的选取无关, 它直接被线性变换所决定。因此, 可以讨论线性变换的特征多项式。

既然相似的矩阵有相同的特征多项式, 特征多项式的各项系数对于相似的矩阵来说都是相同的。考虑特征多项式的常数项, 得到相似矩阵有相同的行列式。因此, 以后可以说线性变换的行列式。

定理 3.2.2 设 T 是 n 维线性空间 V 的一个线性变换, T 的表示矩阵可以在某一基下为对角矩阵的充要条件是 T 有 n 个线性无关的特征向量。

证明 (1) 必要性: 设 T 在基 $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ 下的表示矩阵 $\mathbf{A} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 为对角矩阵, 那么 $T\mathbf{x}_i = \lambda_i\mathbf{x}_i$, 因此 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 为 n 个线性无关的特征向量。

(2) 充分性: 设 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 为 n 个线性无关的特征向量, 则 $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ 为 V 的一个基, 此时存在 n 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 使得 $T\mathbf{x}_i = \lambda_i\mathbf{x}_i$ 成立, 即 $TX = X\mathbf{A}$, 其中 $\mathbf{A} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ 为对角矩阵。

定理 3.2.3 设 λ 是 $\mathbf{A}$ 的任一特征值, 若 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s$ 都是属于 λ 的特征向量, 则 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s$ 的任意非零线性组合仍是属于 λ 的特征向量。

定理 3.2.4 设 λ 是 n 阶方阵 $\mathbf{A}$ 的任一特征值, $\mathbf{x}$ 是所属的任一特征向量, 则有如下结论:

- (1) $\forall k \in \mathbb{C}, k\lambda$ 是 $k\mathbf{A}$ 的特征值, $\mathbf{x}$ 是 $k\mathbf{A}$ 的属于 $k\lambda$ 的特征向量;
- (2) $\forall k \in \mathbb{N}, \lambda^k$ 是 $\mathbf{A}^k$ 的特征值, $\mathbf{x}$ 是 $\mathbf{A}^k$ 的属于 λ^k 的特征向量;
- (3) 若 $f(\mathbf{A})$ 是 $\mathbf{A}$ 的多项式, 则 $f(\lambda)$ 是 $f(\mathbf{A})$ 的特征值, $\mathbf{x}$ 是 $f(\mathbf{A})$ 的属于 $f(\lambda)$ 的特征向量;
- (4) 若 $\mathbf{A}$ 可逆, 则 $\lambda \neq 0$, 且 $\frac{1}{\lambda}$ 是 $\mathbf{A}^{-1}$ 的特征值, $\mathbf{x}$ 是 $\mathbf{A}^{-1}$ 的属于 $\frac{1}{\lambda}$ 的特征向量;
- (5) 若 $\mathbf{A}$ 可逆, 则 $\frac{|\mathbf{A}|}{\lambda}$ 是伴随矩阵 $\mathbf{A}^*$ 的特征值, $\mathbf{x}$ 是 $\mathbf{A}^*$ 的属于 $\frac{|\mathbf{A}|}{\lambda}$ 的特征向量;
- (6) λ 也是 $\mathbf{A}^T$ 的特征值。

证明 我们证明 (2)、(4)、(6), 而将 (1)、(3)、(5) 留给读者作为练习。

(2) 由 $\mathbf{A}\mathbf{x} = \lambda \cdot \mathbf{x}$, 有 $\mathbf{A}^2\mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) = \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$, 知 $k = 2$ 时结论成立, 假设对于 $k - 1$ 有 $\mathbf{A}^{k-1}\mathbf{x} = \lambda^{k-1}\mathbf{x}$ 成立, 则由归纳原理推得对 k 有 $\mathbf{A}^k\mathbf{x} = \mathbf{A}(\mathbf{A}^{k-1}\mathbf{x}) = \mathbf{A}(\lambda^{k-1}\mathbf{x}) = \lambda^{k-1}(\mathbf{A}\mathbf{x}) = \lambda^k\mathbf{x}$ 。

(4) 设 $\mathbf{A}$ 可逆, 由 $\prod_{i=1}^n \lambda_i = |\mathbf{A}| \neq 0$ 可知所有特征值非零。进而在 $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ 两端左乘 $\mathbf{A}^{-1}$ 得 $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{A}\mathbf{x}) = \mathbf{A}^{-1}(\lambda\mathbf{x}) = \lambda(\mathbf{A}^{-1}\mathbf{x})$, 由 $\lambda \neq 0$ 便得 $\mathbf{A}^{-1}\mathbf{x} = \lambda^{-1}\mathbf{x}$, 可知命题为真。

(6) 由于 $|\mathbf{A}^T - \lambda\mathbf{I}_n| = |(\mathbf{A} - \lambda\mathbf{I}_n)^T| = |\mathbf{A} - \lambda\mathbf{I}_n|$, 可见 $\mathbf{A}^T$ 与 $\mathbf{A}$ 有相同的特征多项式, 因而有相同的特征值。

注意: 因为 $\mathbf{A}^T$ 的特征向量是齐次线性方程组 $(\mathbf{A}^T - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$ 的非零解, 与 $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}$ 的非零解一般不同, 故 $\mathbf{x}$ 未必还是 $\mathbf{A}^T$ 的特征向量。

例题 3.14 设 $\mathbf{A}$ 的特征值为 1、2、3, $\mathbf{B} = \mathbf{A}^3 - 3\mathbf{A}^2 + 3\mathbf{A} - \mathbf{I}$, 则 $\mathbf{B}$ 不可逆。

证明 易见 $\mathbf{B} = (\mathbf{A} - \mathbf{I})^3$, $\mathbf{B}$ 的特征值为 $(\lambda - 1)^3$ 。以 λ 为 1、2、3 分别代入, 得 $\mathbf{B}$ 的特征值为 0、1、2, 因此得 $|\mathbf{B}| = 0$, 从而知 $\mathbf{B}$ 不可逆。

定理 3.2.5 属于不同特征值的特征向量是线性无关的。

证明 对特征值的个数做归纳证明。显然个数为1时显然成立, 设个数为 $s-1$ 时结论成立。若 $\mathbf{x}_i$ 属于特征向量 λ_i 的特征向量, $i=1, 2, \dots, s$ 且 $i \neq j$ 时 $\lambda_i \neq \lambda_j$ 。由 $\sum_{i=1}^s c_i \mathbf{x}_i = 0$

两边作线性变换可得 $\sum_{i=1}^s c_i \cdot T\mathbf{x}_i = 0$, 即 $\sum_{i=1}^s c_i \lambda_i \cdot \mathbf{x}_i = 0$ 。

从而 $\sum_{i=1}^s c_i \lambda_i \cdot \mathbf{x}_i - \lambda_s \sum_{i=1}^s c_i \mathbf{x}_s = 0$, 即 $\sum_{i=1}^{s-1} c_i (\lambda_i - \lambda_s) \cdot \mathbf{x}_i = 0$ 由归纳假定可得 $c_i = 0, i=1, 2, \dots, s-1$, 再根据 $\sum_{i=1}^s c_i \mathbf{x}_i = 0$ 可得 $c_s \mathbf{x}_s = 0$, 即 $c_s = 0$, 从而 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s$ 线性无关。

推论 3.2.1 如果在 n 维线性空间 V 中, 线性变换 T 的特征多项式在数域 $\mathcal{F}$ 中有 n 个不同的根, 即 T 有 n 个不同的特征值, 那么 T 在某组基下的表示矩阵是对角形的。

证明 根据定理可取相应的特征向量作为基即可。

推论 3.2.2 在复线性空间中, 如果线性变换 T 的特征多项式没有重根, 那么 T 在某个基下的矩阵是对角形的。

对线性变换具有相同特征值的情形, 要判断这个线性变换的矩阵能不能相似于对角形, 问题就要复杂些。

定理 3.2.6 如果 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是线性变换 T 的不同的特征值, 而 $\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,r_i}$ 是属于特征值 λ_i 的线性无关的特征向量, $i=1, 2, \dots, k$, 那么向量组 $\{\mathbf{x}_{1,1}, \dots, \mathbf{x}_{1,r_1}, \dots, \mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,r_k}\}$ 也线性无关。

证明 设 $\sum_{i=1}^k \sum_{j=1}^{r_i} c_{ij} \mathbf{x}_{i,j} = 0$, 令 $\mathbf{x}_i = \sum_{j=1}^{r_i} c_{ij} \mathbf{x}_{i,j}$, 此时 $\sum_{i=1}^s \mathbf{x}_i = 0$ 。若存在 $\mathbf{x}_i \neq 0$, 那么 $\mathbf{x}_i$ 为特征值 λ_i 的特征向量, 这些非零的特征向量就线性相关, 这和定理3.2.5矛盾。因此对任意 $i, \mathbf{x}_i = 0$, 那么由 $\sum_{j=1}^{r_i} c_{ij} \mathbf{x}_{i,j} = 0$ 根据题设 $c_{ij} = 0, j=1, 2, \dots, r_i$ 。因此定理得证。

根据这个定理, 对于一个线性变换 T , 求出属于每个不同特征值的线性无关的特征向量, 把它们合在一起还是线性无关的。如果它们的个数等于空间的维数, 那么这个线性变换在合适的基下的表示矩阵是对角矩阵; 如果它们的个数少于空间的维数, 那么这个线性变换在任何一个基下的矩阵都不能是对角形。换句话说, 设 T 全部不同的特征值是 $\lambda_1, \lambda_2, \dots, \lambda_r$, 于是 T 在某一个基下的矩阵成对角形的充要条件是 T 的特征子空间 $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_r}$ 的维数之和等于空间的维数。

3.2.3 线性变换的值域与核

定义 3.2.4 设 T 是线性空间 V 的一个线性变换, T 的全体像组成的集合称为 T 的值域, 用 TV 表示。所有被 T 变成零向量的向量组成的集合称为 T 的核, 用 $N(T)$ 或 $T^{-1}(0)$ 表示。若用集合的记号, 则

$$TV = \{T\mathbf{x} \mid \mathbf{x} \in V\}, \quad T^{-1}(0) = \{\mathbf{x} \mid T\mathbf{x} = 0, \mathbf{x} \in V\}$$

显然, 线性变换的值域与核都是 V 的线性子空间。 TV 的维数称为 T 的秩, $T^{-1}(0)$ 的维数称为 T 的零度。

定理 3.2.7 设 T 是 n 维线性空间 V 的线性变换, $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r]$ 是 $N(T)$ 的基, 扩充这个向量组, 得到 $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ 是 V 的一个基, 在基 X 下 T 的表示矩阵是 A , 则