

§ 1.2 线性变换及其矩阵表示

定义 V^n 到 V^m 的变换 T 称为线性的, 如果对任意的数 k 及 V^n 中的任意向量 α, β , 恒有

$$T(\alpha + \beta) = T\alpha + T\beta, T(k\alpha) = kT\alpha.$$

记 $\xi = T\alpha \in V^m$, 则称 ξ 为 α 在 T 下的像, α 称为 ξ 的原像。
特别, 当 T 是 V^n 到自身的一个线性变换, 则称 T 是 V^n 的线性变换。

例1 给定 $A \in F^{m \times n}$, 定义 V^n 到 V^m 的变换 A 为

$$x \in F^n \rightarrow y = Ax \in F^m.$$

容易验证 A 是一个线性变换。

例2 给定 $P \in F^{m \times m}, Q \in F^{n \times n}$, 不难验证变换

$$T: X \in F^{m \times n} \rightarrow PXQ \in F^{m \times n}$$

是 $F^{m \times n}$ 的一个线性变换。

例3 对 $P_n(t)$ 中的多项式求导 $\frac{d}{dt}$, 很容易验证它是 $P_n(t)$ 的线性变换, 记为 D , 即

$$Dp(t) = \frac{d}{dt} p(t), \forall p(t) \in P_n(t)$$

V 的恒等变换 $I: I\alpha = \alpha, \forall \alpha \in V$, 和零变换 O : 都是 V 的线性变换。

线性变换具有下列简单性质:

(1) $O\alpha = 0, \forall \alpha \in V. T0 = 0; T(-\alpha) = -T\alpha, \forall \alpha \in V$

(2) $T(\sum_{i=1}^r k_i \alpha_i) = \sum_{i=1}^r k_i T\alpha_i$, 即任意一组向量的线性组合取像, 等于分别取像再线性组合;

(3) 一组线性相关的向量 $\alpha_1, \alpha_2, \dots, \alpha_r$, 它们在 T 下像的 $T\alpha_1, T\alpha_2, \dots, T\alpha_r$ 也是线性相关的。

但是, 线性无关的向量在 T 下的像可能是线性相关的, 例如零变换把线性无关的向量都映射为零向量。

设 T 是 V^n 到 V^m 的线性变换，在 V^n 和 V^m 中分别取基
 $B_\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 和 $B_\beta = \{\beta_1, \beta_2, \dots, \beta_m\}$ ，则 α_j 的像
 $T\alpha_j (1 \leq j \leq n)$ 可由基 B_β 唯一地线性表出：

$$T\alpha_j = \sum_{i=1}^m a_{ij}\beta_i = [\beta_1 \beta_2 \cdots \beta_m] \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

如果把 $T\alpha_j, 1 \leq j \leq n$ 按顺序排列，并使用矩阵记号，
则有

$$[T\alpha_1 T\alpha_2 \cdots T\alpha_n] = [\beta_1 \beta_2 \cdots \beta_m] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

为了简化记法和便于运算，令 $TB_\alpha \triangleq [T\alpha_1 T\alpha_2 \cdots T\alpha_n]$ ，
那么上式可简写为

$$TB_\alpha = B_\beta A, \quad (1.2-1)$$

其中 $m \times n$ 矩阵

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

(1.2-1) 式叫做 T 的矩阵表示，称 A 为 T 在基偶 $\{B_\alpha, B_\beta\}$ 下的矩阵。

特别，若 T 是 V^n 到自身的线性变换，这时
 $V^m = V^n$ ，并规定 B_β 取为 B_α ，则 (1.2-1) 式为

$$TB_\alpha = B_\alpha A,$$

称 n 阶方阵 A 为 T 在基 B_α 下的矩阵。

例6 求 $P_n(t)$ 的线性变换 $D = \frac{d}{dt}$ 在基 $B = \{1, t, \dots, t^n\}$ 下的矩阵。

解: $[D1 \ Dt \ \dots \ Dt^n] = [0 \ 1 \ \dots \ nt^{n-1}]$

$$= [1 \ t \ \dots \ t^n] \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

因此, D 在 B 下的矩阵是 $n+1$ 阶方阵:

$$A = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & n \\ & & & & 0 \end{bmatrix},$$

A 中未写出的元素皆为零。

例7 求 $P_2(t)$ 到 $P_3(t)$ 的线性变换 J :

$$J[p(t)] = \int_0^t p(t)dt$$

在基偶 $\left\{ B_1 = \{1, t, t^2\}, B_2 = \{1, t, \frac{t^2}{2}, \frac{t^3}{3}\} \right\}$ 下的矩阵。

解

$$\begin{bmatrix} J1 & Jt & Jt^2 \end{bmatrix} = \begin{bmatrix} t & \frac{t^2}{2} & \frac{t^3}{3} \end{bmatrix} = \begin{bmatrix} 1 & t & \frac{t^2}{2} & \frac{t^3}{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

有了 T 的矩阵表示, 那么 V^n 中任一向量的像就可以确定了, 事实上, $\alpha \in V^n$ 可由 B_α 线性表示出: $\alpha = \sum_{i=1}^n x_i \alpha_i$, 从而 $T\alpha = \sum_{i=1}^n x_i T\alpha_i$, 即若 $\alpha = B_\alpha x$, 则 $T\alpha = TB_\alpha x$.

由(1.2-1)式得

$$T\alpha = TB_\alpha x = B_\beta Ax. \quad (1.2-2)$$

这就是说, 若 x 是 α 在 B_α 下的坐标向量, 那么其像 $T\alpha$ 在 B_β 下的坐标向量是 Ax .

从上述讨论可知，取定基偶 $\{B_\alpha, B_\beta\}$ 后，线性变换 T 有唯一的矩阵表示，即得到 T 在此基偶下的矩阵 A ，反过来，给定 $m \times n$ 矩阵 A ，是否存在唯一的线性变换 T ，它在基偶 $\{B_\alpha, B_\beta\}$ 下的矩阵是 A 呢？我们用下述定理给出肯定的回答。

定理 设 $B_\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$, $B_\beta = \{\beta_1, \beta_2, \dots, \beta_m\}$ 分别

是 V^n 和 V^m 的基, 对于给定的 $m \times n$ 矩阵 $A = [a_{ij}]$,
则存在 V^n 到 V^m 的唯一线性变换 T , 它在 $\{B_\alpha, B_\beta\}$
下的矩阵是 A 。

证 只证存在性。

任取 $\alpha \in V^n$, 设它在 B_α 下的坐标向量为 x ,

即 $\alpha = B_\alpha x$, 令 $\beta = B_\beta Ax \in V^m$, 定义变换 T 为

$$T\alpha = \beta$$

则 T 就是定理中所要求的线性变换。

首先 T 是线性的. 设 $\xi, \eta \in V^n$, 它们在 B_α 下的坐

标向量分别为 y, z , 则

$$\xi + \eta = B_\alpha(y + z); \quad k\xi = B_\alpha(ky), \quad \forall k \in F$$

按 T 的定义, 有

$$T\xi = B_\beta Ay, \quad T\eta = B_\beta Az,$$

从而

$$\begin{aligned} T(\xi + \eta) &= B_\beta A(y + z) \\ &= B_\beta Ay + B_\beta Az = T\xi + T\eta, \end{aligned}$$

$$T(k\xi) = B_\beta A(ky) = kB_\beta Ay = kT\xi.$$

再证 $TB_{\alpha} = B_{\beta}A$. 显然, α_i 在 B_{α} 下的坐标向量是 $e_i = [0 \dots 0 \underset{i}{1} 0 \dots 0]^T$, 即第 i 个分量为1, 其余分量为0 的 n 维列向量. 从而按 T 的定义有 $T\alpha_i = B_{\beta}Ae_i$. 因而

$$\begin{aligned} TB_{\alpha} &= [T\alpha_1 \ T\alpha_2 \ \dots \ T\alpha_n] = [B_{\beta}Ae_1 \ B_{\beta}Ae_2 \ \dots \ B_{\beta}Ae_n] \\ &= B_{\beta}A[e_1 \ e_2 \ \dots \ e_n] = B_{\beta}AI = B_{\beta}A. \end{aligned}$$

由于线性空间的基不是唯一的，同一个线性变换 T 在不同基偶下的矩阵一般不会相同，因此就提出下述两个问题：

(1) 如果另取一对基 $B_{\alpha'}, B_{\beta'}$ ，又可得到 T 在基偶 $\{B_{\alpha'}, B_{\beta'}\}$ 下的矩阵 B ，那么 A 与 B 之间有什么关系呢？

(2) 如何选取 V^n, V^m 的基，才能使 T 的矩阵表示最简呢？

设 n 阶方阵 P 是基 B_α 到 $B_{\alpha'}$ 的变换矩阵, 而 m 阶方阵 Q 是基 B_β 到 $B_{\beta'}$ 的变换矩阵, $m \times n$ 矩阵 A, B 分别是 T 在基偶 $\{B_\alpha, B_\beta\}, \{B_{\alpha'}, B_{\beta'}\}$ 下的矩阵, 那么由关系式

$$B_{\alpha'} = B_\alpha P, \quad B_{\beta'} = B_\beta Q,$$

$$TB_\alpha = B_\beta A, \quad TB_{\alpha'} = B_{\beta'} B$$

可以推出,

$$B_\beta AP = B_\beta QB, \quad B_\beta (AP - QB) = O.$$

由于 B_β 是基, 所以有

$$AP = QB, \quad A = QBP^{-1}, \quad B = Q^{-1}AP.$$

这就是说矩阵 B, A 是相抵 (或等价) 的。

如果 T 是 V^n 到自身的线性变换，则在上述推导过程中，令 $V^m = V^n, B_\beta = B_\alpha, B_{\beta'} = B_{\alpha'}$ ，使得

$$AP = PB, \quad A = PBP^{-1}, \quad B = P^{-1}AP.$$

这表明方阵 B, A 是相似的。

于是， V^n 到 V^m 的一个线性变换 T 在不同基偶下的矩阵是相抵关系的，而 V^n 的线性变换 T 在不同基下的矩阵是相似关系。

从而，第二个问题也就等价于：与 A 相抵的所有矩阵中最简矩阵是什么？与方阵 A 相似的所有方阵中最简矩阵是什么。

定义 设 T 是 V^n 到 V^m 的线性变换, 集

$$N(T) \triangleq \{\alpha \in V^n \mid T\alpha = 0\}$$

$$R(T) \triangleq \{\beta \in V^m \mid \beta = T\alpha, \alpha \in V^n\}$$

分别称为 T 的**核**和 T 的**值域**。

不难证明, $N(T)$ 是 V^n 的一个子空间, 故也称它为 T 的零空间, 其维数叫做 T 的**零度**, 记为 $\text{null } T$; $R(T)$ 是 V^m 的一个子空间, 也称为值空间, 其维数叫做 T 的**秩**, 记为 $\text{rank } T$.

定理 设 T 是 V^n 到 V^m 的线性变换, 则

$$\text{null}T + \text{rank}T = n. \quad (1.2-3)$$

证 令 $\text{null}T = k$, 并设 $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ 是 $N(T)$ 的一个基. 因为 $N(T) \subset V^n$, 所以可把这个 k 个向量扩充为 V^n 的基 $\{\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n\}$.

我们证明 $\{T\alpha_{k+1}, T\alpha_{k+2}, \dots, T\alpha_n\}$ 是 $R(T)$ 的一个基, 从而 (1.2-3) 式子成立.

任取 $\alpha \in V^n$, 则 α 可由 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 线性表出:

$$\alpha = \sum_{i=1}^n b_i \alpha_i. \text{ 由于 } \alpha_i \in N(T), i=1, 2, \dots, k, \text{ 故}$$

$T\alpha_i = 0, i = 1, 2, \dots, k,$ 从而有

$$T\alpha = \sum_{i=1}^n b_i T\alpha_i = \sum_{i=k+1}^n b_i T\alpha_i .$$

这表明 $R(T)$ 中的向量 $T\alpha$ 可由 $\{T\alpha_{k+1}, T\alpha_{k+2}, \dots, T\alpha_n\}$ 线性表出. 再则, 假设有等式

$$\sum_{i=k+1}^n b_i T\alpha_i = 0,$$

则得

$$T\left(\sum_{i=k+1}^n b_i \alpha_i\right) = 0,$$

故 $\sum_{i=k+1}^n b_i \alpha_i \in N(T)$. 于是, 存在数 $c_j, j = 1, 2, \dots, k,$ 使

$$\sum_{j=1}^k c_j \alpha_j = \sum_{i=k+1}^n b_i \alpha_i ,$$

即

$$\sum_{j=1}^k (-c_j) \alpha_j + \sum_{i=k+1}^n b_i \alpha_i = 0.$$

由于 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 是线性无关的, 所以有

$$c_1 = c_2 = \dots = c_k = 0, b_{k+1} = b_{k+2} = \dots = b_n = 0.$$

这就证明了 $\{T\alpha_{k+1}, T\alpha_{k+2}, \dots, T\alpha_n\}$ 线性无关.

综合上述两点, 证明了 $\{T\alpha_{k+1}, T\alpha_{k+2}, \dots, T\alpha_n\}$
是 $R(T)$ 的基.

事实上若 $\text{rank } T = r$, 则(1.2-3)式给出 $\text{null } T = n - r$.

在 V^n 中选取这样的基 $B_\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_r, \alpha_{r+1}, \alpha_n\}$, 使 $\{\alpha_{r+1}, \alpha_{r+2}, \dots, \alpha_n\}$ 是 $N(T)$ 的基。这是做得到的, 实际上只需先选取 $N(T)$ 的一个基, 再把它扩充为 V^n 的基即可。由定理知, $\{T\alpha_1, T\alpha_2, \dots, T\alpha_r\}$ 是线性无关的, 且是 $R(T)$ 的一个基。再把它扩充为 V^m 的基 $B_\beta = \{T\alpha_1, T\alpha_2, \dots, T\alpha_r; \beta_{r+1}, \dots, \beta_m\}$. 那么, 容易验证 T 在基偶 $\{B_\alpha, B_\beta\}$ 下的矩阵是

$$\begin{matrix} & \overbrace{r} & \overbrace{n-r} \\ r\{ & \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \\ m-r\{ & \end{matrix}$$

其中 I_r 表示 r 阶单位阵, O 表示零矩阵。

下面讨论 V^n 的线性变换 T 的最简表示

注: $N(T) + R(T)$ 却不一定是 V^n , 例如, 对于 $P_{n-1}(t)$ 的线性变换 D 来说, $N(D) = P_0(t), R(D) = P_{n-2}(t)$, 故

$$\text{null } D + \text{rank } D = 1 + (n-1) = n,$$

但

$$N(D) + R(D) = P_0(t) + P_{n-2}(t) = P_{n-2}(t) \neq P_{n-1}(t).$$

定义 设 T 是 V 的线性变换, W 是 V 的子空间, 如果对于任意的 $\alpha \in W$, 都有

$$T\alpha \in W,$$

则称 W 为 T 的不变子空间。

注： T 的零空间 $N(T)$ 和 T 的值域 $R(T)$ 都是 T 的不变子空间，并且 T 的不变子空间的交空间及和空间也是 T 的不变子空间。

利用 T 的不变子空间可以简化 T 的矩阵表示。

如果 $W_i, 1 \leq i \leq s$ 都是 T 的不变子空间，且有

$$V^n = W_1 \oplus W_2 \oplus \cdots \oplus W_s,$$

则在每个 W_i 中取一个基 $\{\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{in_i}\} (i = 1, 2, \cdots, s)$,

这里 $n_i = \dim W_i$, 并把它们顺序排列为 V^n 的基

$$B = \{\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1n_1}; \alpha_{21}, \alpha_{22}, \cdots, \alpha_{2n_2}; \cdots; \alpha_{s1}, \alpha_{s2}, \cdots, \alpha_{sn_s}\},$$

那么 T 在 B 下的矩阵是对角块矩阵:

$$\begin{bmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{ss} \end{bmatrix} \triangleq \text{diag}\{A_{11}, A_{22}, \dots, A_{ss}\},$$

其中 A_{ii} ($1 \leq i \leq s$) 是 n_i 阶方阵。反之, 如果 T 在 B 下的矩阵是上述的对角块矩阵, 则

$$W_i = \text{span}\{a_{i1}, a_{i2}, \dots, a_{in_i}\} \quad (i = 1, 2, \dots, s)$$

是 T 的不变子空间, 且 V^n 是这些不变子空间的直和。

众所周知，对角矩阵是方阵中最简单的，它有许多特殊的性质。让我们看一下，若线性变换 T 在基 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 下的矩阵是对角矩阵：

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \triangleq \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

那么 T 应满足什么要求呢？

由 (1.2-1') 式, 这时有

$$[T\alpha_1 T\alpha_2 \cdots T\alpha_n] = [\alpha_1 \alpha_2 \cdots \alpha_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

即,

$$T\alpha_i = \lambda_i \alpha_i, \quad 1 \leq i \leq n, \quad (1.2-4)$$

这就引出讨论 T 的特征值和特征向量的问题。

定义 设 T 是 $V^n(F)$ 的一个线性变换, 如果存在

$\lambda_0 \in F, \xi \in V^n(F)$ 且 $\xi \neq 0$, 使

$$T\xi = \lambda_0\xi, \quad (1.2-5)$$

则称 λ_0 是 T 的一个**特征值**, ξ 称为 T 关于 λ_0 的

特征向量。

为了求出 T 的特征值和特征向量，在 V^n 中取一个基
 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ ，且设 T 在 B 下的矩阵是 A 。如果 ξ
是 T 的一个特征向量， λ_0 是相应的特征值，即
 $T\xi = \lambda_0\xi, \xi \neq 0$ ，那么 ξ 可由 B 的线性表出：

$$\xi = \sum_{i=1}^n x_i \alpha_i = Bx, x = [x_1, x_2 \cdots x_n]^T,$$

因而有

$$T\xi = TBx = BAx, \lambda_0\xi = B(\lambda_0x).$$

代入（1.2-5）式得 $BAx = B(\lambda_0x)$ ，即 $B(Ax - \lambda_0x) = 0$ 。

由此可得 λ_0 和 ξ 在 B 下的坐标向量 x 所满足的方程：

$$Ax = \lambda_0x.$$

T 的特征值问题与 A 的特征值问题是一一对应的。
由于相似矩阵有相同的特征多项式，所以我们可以把 A 的特征多项式

$$f(\lambda) \triangleq \det(\lambda I_n - A) = \lambda^n + b_1 \lambda^{n-1} + \cdots + b_{n-1} \lambda + b_n$$

称为 T 的**特征多项式**，于是 T 的特征值就是 T 的特征多项式的根。

例8 $P_2(t)$ 的线性变换 T 的定义为

$$Tp(t) = p(t) + (t+1) \frac{d}{dt} p(t),$$

求 T 的特征值和特征向量。

解 取 $P_2(t)$ 的一个基 $B = \{1, t, t^2\}$, 则 T 在 B 下的

矩阵是

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

A 的特征值是 $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$, 相应的特征向量分别为 $k_1[1 \ 0 \ 0]^T, k_2[1 \ 1 \ 0]^T, k_3[1 \ 2 \ 1]^T$. 因此, T 的特征值是 $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3$, T 关于 $\lambda_1, \lambda_2, \lambda_3$ 的特征向量分别是多项式 $k_1, k_2(1+t), k_3(1+2t+t^2)$, 上述的 k_1, k_2 和 k_3 可为任意非零实数。

定理 T 关于不同特征值的特征向量是线性无关的。

证 用数学归纳法证明. 由于特征向量是非零向量, 所以单个的特征向量线性无关. 假设 T 关于 k 个互异特征值的特征向量是线性无关的, 要证 T 关于 $k+1$ 个互异特征值 $\lambda_1, \lambda_2, \dots, \lambda_{k+1}$ 的特征向量 $\alpha_1, \alpha_2, \dots, \alpha_{k+1}$ 也线性无关.

设有等式

$$b_1\alpha_1 + b_2\alpha_2 + \dots + b_{k+1}\alpha_{k+1} = 0, \quad (1.2-6)$$

一方面在 (1.2-6) 式两端同乘以 λ_{k+1} ，得

$$b_1 \lambda_{k+1} \alpha_1 + b_2 \lambda_{k+1} \alpha_2 + \dots + b_{k+1} \lambda_{k+1} \alpha_{k+1} = 0, \quad (1.2-7)$$

另一方面，对 (1.2-6) 式两端同施行线性变换 T ，得

$$b_1 \lambda_1 \alpha_1 + b_2 \lambda_2 \alpha_2 + \dots + b_{k+1} \lambda_{k+1} \alpha_{k+1} = 0, \quad (1.2-8)$$

将 (1.2-7) 式与 (1.2-8) 式相减，得

$$b_1 (\lambda_{k+1} - \lambda_1) \alpha_1 + b_2 (\lambda_{k+1} - \lambda_2) \alpha_2 + \dots + b_k (\lambda_{k+1} - \lambda_k) \alpha_k = 0.$$

根据归纳法假设， $\alpha_1, \alpha_2, \dots, \alpha_k$ 是线性无关的，故

$$b_i(\lambda_{k+1} - \lambda_i) = 0 \quad (i \leq k),$$

但 $\lambda_{k+1} \neq \lambda_i$, 所以 $b_i = 0 (i = 1, 2, \dots, k)$, 从而 (1.2-6) 式为

$$b_{k+1} \alpha_{k+1} = 0.$$

又因 $\alpha_{k+1} \neq 0$, 故有 $b_{k+1} = 0$. 这就证明了 $\alpha_1, \alpha_2, \dots, \alpha_{k+1}$ 是线性无关的.

类似地可以证明下述定理.

定理 设 $\lambda_1, \lambda_2, \dots, \lambda_k$ 是 T 的不同特征值, $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ir_i}$ ($1 \leq i \leq k$) 是 T 关于 λ_i 的 r_i 个线性无关特征向量, 则向量组

$$\{\alpha_{11}, \alpha_{12}, \dots, \alpha_{1r_1}; \alpha_{21}, \alpha_{22}, \dots, \alpha_{2r_2}; \dots; \alpha_{k1}, \alpha_{k2}, \dots, \alpha_{kr_k}\},$$

线性无关 (证明略) .

对于 T 的任一特征值 λ_0 , T 关于 λ_0 的所有特征向量, 再添上零向量组成的集:

$$V_{\lambda_0} \triangleq \{\alpha \in V^n \mid T\alpha = \lambda_0 \alpha\}, \quad (1.2-9)$$

是 V^n 的一个子空间。事实上, 若

事实上, 若 $\alpha \in V_{\lambda_0}$, 则 $k\alpha \in V_{\lambda_0}$ 又若 $\beta \in V_{\lambda_0}$,

则有

$$T(\alpha + \beta) = T\alpha + T\beta = \lambda_0\alpha + \lambda_0\beta = \lambda_0(\alpha + \beta),$$

从而 $\alpha + \beta \in V_{\lambda_0}$.

定义 (1.2-9) 式所定义的 V_{λ_0} 称为 T 关于 λ_0 的特征子空间。 $\dim V_{\lambda_0}$ 称为 λ_0 的**几何重数**。

显然, V_{λ_0} 是 T 的不变子空间, 且 λ_0 的几何重数就是 T 关于 λ_0 的线性无关特征向量的最大个数。

我们知道，如果 $\lambda_1, \lambda_2, \dots, \lambda_s$ 是 T 的所有不同的特征值，则 T 的特征多项式 $f(\lambda)$ 可以表示为

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s},$$

且

$$n_1 + n_2 + \cdots + n_s = n,$$

而 $n_i (1 \leq i \leq s)$ 称为特征值 λ_i 的代数重数。

定理 设 $\lambda_1, \lambda_2, \dots, \lambda_s$ 是 T 的所有不同特征值, 则对

任一 $\lambda_i (1 \leq i \leq s)$, 都有

$$\dim V_{\lambda_i} \leq n_i, \quad (1.2-10)$$

即任何特征值的几何重数不大于其代数重数。

证 不失一般性, 就 λ_1 来证明. 设

$$\dim V_{\lambda_1} = k,$$

则 T 关于 λ_1 有 k 个线性无关的特征向量 $\alpha_1, \alpha_2, \dots, \alpha_k$. 从

而 $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ 是 T 的不变子空间 V_{λ_1} 的一个基, 把

它扩充为 V^n 的基 $B = \{\alpha_1, \alpha_2, \dots, \alpha_k, \alpha_{k+1}, \dots, \alpha_n\}$, 那么 T 在 B 下的矩阵为

$$A = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & O & & \\ & & & \lambda_1 & \\ & & & & \end{bmatrix} \begin{matrix} \left. \vphantom{\begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & O & & \\ & & & \lambda_1 & \\ & & & & \end{bmatrix}} \right\} k \\ \left. \vphantom{\begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & O & & \\ & & & \lambda_1 & \\ & & & & \end{bmatrix}} \right\} n-k \end{matrix}$$

从而

$$\det(\lambda I_n - A) = (\lambda - \lambda_1)^k \cdot \det(\lambda I_{n-k} - A_{22}),$$

故 λ_1 的几何重数 $k \leq n_1$.

由上面讨论可知, 若 $\lambda_1, \lambda_2, \dots, \lambda_s$ 是 T 的所有不同的特征值, 则有

$$V_{\lambda_1} + V_{\lambda_2} + \dots + V_{\lambda_s} = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_s},$$

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \dots + \dim V_{\lambda_s} \leq n_1 + n_2 + \dots + n_s = n$$

定义 T 称为是可对角化的, 如果存在 V^n 的基 B , 使 T 在 B 下的矩阵是对角矩阵。

定理 T 是可对角化的充分必要条件是下列等价条件之一成立:

(1) T 有 n 个线性无关的特征向量;

(2) $\dim V_{\lambda_i} = n_i, 1 \leq i \leq s$.

(3) $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s} = V^n$;

证 只证 (1) 设 T 在基 $B = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 下的矩阵是对角矩阵: $\text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, 即

$$[Ta_1 Ta_2 \cdots Ta_n] = [a_1, a_2, \dots, a_n] \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\},$$

则 $Ta_i = \lambda_i a_i, 1 \leq i \leq n.$ (1.2-11)

这就是说, α_i 是 T 的特征向量, λ_i 是相应的特征值, 从而 T 有 n 个线性无关的特征向量。

反之，若 T 有 n 个线性无关的特征向量，则 (1.2-11) 式成立，因此 T 在基 B 下的矩阵是对角矩阵。

这时，对角矩阵的对角线上的数 $\lambda_1, \lambda_2, \dots, \lambda_s$ （它们之中可能相同）都是 T 的特征值。

推论 若 T 有 n 个不同的特征值，则 T 必可对角化。

于是，对于 $A \in C^{n \times n}$ ，若 A 的特征多项式没有重根，则 A 必可对角化，亦即存在可逆矩阵 P ，使 $P^{-1}AP$ 为对角矩阵，且 P 的每个列向量都是 A 的特征向量。

例9 证明 $P_2(t)$ 的线性变换 D 是不可对角化的。

证 取 $P_2(t)$ 的一个基 $B = \{1, t, t^2\}$ ， D 在 B 下的矩

阵是

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

A 的特征多项式是 $\det(\lambda I_3 - A) = \lambda^3$ ， $\lambda = 0$ 是它的三重根，但齐次线性方程组

$$(A - 0I_3)x = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0,$$

的基础解系只含一个解向量，所以特征值 0 的几何重数（为 1）小于代数重数（是 3），故 D 不可对角化。

例10 R^3 的线性变换 T 定义为

$$T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

问 T 是否可对角化？

解 取 R^3 的标准基 $B = \{e_1, e_2, e_3\}$, 则 T 在 B 下的矩阵

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

A 的特征多项式为 $\det(\lambda I_3 - A) = (\lambda - 5)(\lambda - 3)^2$,

故 A 有两个不同的特征值: $\lambda_1 = 5, \lambda_2 = 3$ (二重特征值)。

关于 $\lambda_1 = 5$, 由于方阵 $A - 5 \cdot I_3$ 的秩为2, 所以只有一个线性无关的特征向量, 例如取为

$$x_1 = [1 \ 2 \ 1]^T.$$

关于 $\lambda_2 = 3$, 由于 $A - 3 \cdot I_3$ 的秩为1, 故齐次线性方程组

$$(A - 3I_3)x = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

有两个线性无关的解向量, 例如取为

$$x_2 = [1 \ 0 \ -1]^T, \quad x_3 = [0 \ 1 \ 0]^T.$$

于是 A 有三个线性无关的特征向量, 所以 A 可对角化, 即 T 是可对角化的。

若在 R^3 取基

$$B = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\},$$

则 T 在 B 下的矩阵是对角矩阵 $\text{diag}\{5, 3, 3\}$ ，即有

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & & \\ & 3 & \\ & & 3 \end{bmatrix}.$$

例11 证明矩阵

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

在实数域上是不可对角化的，但在复数域上是可对角化的。

证 A 的特征多项式

$$\det(\lambda I_2 - A) = \lambda^2 - 2\lambda + 2,$$

它在实数域上没有根，从而不存在特征向量，故 A 在实数域上不可对角化。但在复数域上，这个多项式有两个不同的根：

$$\lambda_1 = 1 + i, \quad \lambda_2 = 1 - i,$$

因而 A 可对角化。

不难求出与 λ_1, λ_2 相应的特征向量，例如分别取为

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

则

$$\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} = \begin{bmatrix} 1+i & \\ & 1-i \end{bmatrix}.$$

注：不是每个线性变换都是可对角化的。