§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

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

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

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

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 $\dfrac{d}{dt}$，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） $T(\sum_{i=1}^{r}k_i\alpha_i)=\sum_{i=1}^{r}k_iT\alpha_i,$ 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 $\alpha_1,\alpha_2,\cdots,\alpha_r$,它们在 T 下像的 $T\alpha_1,T\alpha_2,\cdots T\alpha_r$ 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 $B_\alpha=\{\alpha_1,\alpha_2,\cdots\alpha_n\}$和$B_\beta=\{\beta_1,\beta_2,\cdots\beta_m\}$，则$a_j$的像 $Ta_j(1\le j\le n)$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$T\alpha_j,1\le j\le n$按顺序排列，并使用矩阵记号，则有

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

为了简化记法和便于运算，令 $TB_a \triangleq [T\alpha_1 T\alpha_2 \cdots T\alpha_n]$,

那么上式可简写为

$$TB_\alpha = B_\beta A, \qquad (1.2\text{-}1)$$

其中 $m\times n$矩阵

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

（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 $D=\dfrac{d}{dt}$ 在基 $B=\{1,t,\cdots,t^n\}$ 下的矩阵。

解：

$$[D1\ Dt\cdots Dt^n]=[0\ 1\cdots nt^{n-1}]$$

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

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

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

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

$$\alpha = B_\alpha x, \text{则} T\alpha = TB_\alpha x.$$

由（1.2-1）式得

$$T\alpha = TB_a x = B_\beta Ax. \qquad (1.2\text{-}2)$$

这就是说，若x是a在B_a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 $B_a=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}, B_\beta=\{\beta_1,\beta_2,\cdots,\beta_m\}$ 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 $\{B_a,B_\beta\}$ 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

$$T(\xi+\eta)=B_\beta A(y+z)$$

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

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

再证 $TB_\alpha = B_\beta A$. 显然, α_i 在 B_α 下的坐标向量是

$e_i = [0\ldots0\underset{i}{1}0\ldots0]^{\mathrm{T}}$，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 $T\alpha_i = B_\beta A e_i$.

因而

$$TB_\alpha = [T\alpha_1\ T\alpha_2 \ldots T\alpha_n] = [B_\beta A e_1\ B_\beta A e_2 \ldots B_\beta A e_n]$$

$$= B_\beta A[e_1\ e_2 \ldots e_{\mathrm{n}}] = B_\beta A I = B_\beta A\ .$$

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，

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

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 $V^m = V^n, B_\beta = B_a, B_{\beta'} = B_{\alpha'}$, 便得

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

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 $\mathrm{null}\, T$; $R(T)$ 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 $\mathrm{rank}\, T$.

定理 设T是V^n到V^m的线性变换，则

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

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

我们证明$\{T\alpha_{k+1}, T\alpha_{k+2}, \ldots, T\alpha_n\}$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$\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$线性表出：$\alpha = \sum_{i=1}^{n} b_i \alpha_i$.由于 $\alpha_i \in N(T), i = 1,2, \ldots, k,$ 故

$T\alpha_i = 0, i = 1,2,\ldots,k,$ 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 $\{T\alpha_{k+1}, T\alpha_{k+2}, \ldots, T\alpha_n\}$ 线性表出.

再则，假设有等式 $\sum_{i=k+1}^{n} b_i T\alpha_i = 0,$

则得

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

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

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

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

在 V^n 中选取这样的基 $B_a = \{\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n\}$, 使 $\{\alpha_{r+1}, \alpha_{r+2}, \cdots, \alpha_n\}$ 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, $\{T\alpha_1, T\alpha_2, \cdots, T\alpha_r\}$ 是线性无关的，且是 $R(T)$ 的一个基。再把它扩充为 V^m 的基 $B_\beta = \{T\alpha_1, T\alpha_2, \cdots, T\alpha_r; \beta_{r+1}, \cdots, \beta_m\}$. 那么,容易验证 T 在基偶 $\{B_\alpha, B_\beta\}$ 下的矩阵是

$$\begin{array}{r c} & \begin{array}{cc} \overbrace{\quad}^{r} & \overbrace{\quad}^{n-r} \end{array} \\ \begin{array}{r} r\{ \\ m-r\{ \end{array} & \begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \end{array}$$

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 $W_i, 1 \le i \le s$ 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：

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

其中$A_{ii}(1 \le i \le s)$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

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

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 $B=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$ 下的矩阵是对角矩阵：

$$\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \triangleq \mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},$$

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:

$$T\alpha_i = \lambda_i \alpha_i,\ 1 \le i \le n, \qquad (1.2\text{-}4)$$

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

$$T\xi = \lambda_0 \xi, \qquad (1.2\text{-}5)$$

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 $B=\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$，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

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

代入（1.2-5）式得 $BAx=B(\lambda_0 x)$,即$B(Ax-\lambda_0 x)=0.$

由此可得 λ_0和ξ在B下的坐标向量 x 所满足的方程：

$$Ax=\lambda_0 x.$$

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$\lambda_1,\lambda_2,\ldots,\lambda_{k+1}$的特征向量$\alpha_1,\alpha_2,\ldots,\alpha_{k+1}$也线性无关.

设有等式

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

一方面在（1.2-6）式两端同乘以λ_{k+1}，得

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

另一方面，对（1.2-6）式两端同施行线性变换T，得

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

将（1.2-7）式与（1.2-8）式相减，得

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

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

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

但 $\lambda_{k+1}\neq\lambda_i$, 所以 $b_i=0(i=1,2,\ldots,k)$, 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 $\alpha_1,\alpha_2,\ldots,\alpha_{k+1}$ 是线性无关的.

类似地可以证明下述定理.

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

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

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：

$$V_{\lambda_0}\underline{\underline{\Delta}}\{\alpha\in V^n\,|\,T\alpha=\lambda_0\alpha\}, \qquad (1.2\text{-}9)$$

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 $a+\beta \in V_{\lambda_0}$.

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 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 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 $n_i (1 \le i \le s)$ 称为特征值 λ_i 的**代数重数**。

定理 设 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是T的所有不同特征值，则对任一 $\lambda_i(1\le i\le s)$都有

$$\dim V_{\lambda_i}\le n_i,\qquad （1.2\text{-}10）$$

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 $\alpha_1,\alpha_2,\cdots,\alpha_k$.从而 $\{\alpha_1,\alpha_2,\cdots,\alpha_k\}$是T的不变子空间 V_{λ_1} 的一个基，

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

$$A=\begin{array}{c}\left[\begin{array}{ccc|c}\lambda_1 & & & \\ & \ddots & & A_{12} \\ & & \lambda_1 & \\ & O & & A_{22}\end{array}\right]\begin{array}{l}\left.\vphantom{\begin{array}{c}a\\a\end{array}}\right\}k \\ \left.\vphantom{\begin{array}{c}a\\a\end{array}}\right\}n-k\end{array} \\ \underbrace{\qquad\qquad\qquad}_{k}\ \underbrace{\qquad}_{n-k}\end{array}$$

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$k\le n_1$.

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

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

$$\dim V_{\lambda_1}+\dim V_{\lambda_2}+\cdots+\dim V_{\lambda_s}\le n_1+n_2+\cdots+n_s=n$$

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）$\dim V_{\lambda_i} = n_i, 1 \le i \le s.$

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

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

$$[Ta_1Ta_2\cdots Ta_n]=[a_1,a_2,\cdots,a_n]\mathrm{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\},$$

则

$$Ta_i=\lambda_i a_i,1\le i\le n. \tag{1.2-11}$$

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 $\lambda_1,\lambda_2,\cdots,\lambda_s$（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

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

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 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：

$$\lambda_1=1+\mathrm{i},\quad \lambda_2=1-\mathrm{i}\ ,$$

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

$$x_1 = [1 \quad i] \quad x_2 = [1 \quad -i]$$

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。