§7.3 线性变换的矩阵

从前面的讨论看到线性变换的运算及其运算律与矩阵非常相似，由此启发我们进一步探讨二者的关系.

基本内容

线性变换的矩阵

线性变换的运算与矩阵运算的关系

相似矩阵(线性变换在不同基下的矩阵的关系)

1. 线性变换的矩阵

A的第j列为$\sigma(\varepsilon_j)$在基$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$下的坐标，故A是唯一确定的！

定义：**设$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$是数域P上n维线性空间V 的一组基, σ是V的一个线性变换.基向量的像$\sigma(\varepsilon_i)(i=1,2,\ldots,n)$被基$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$线性表示为**

$$\begin{cases} \sigma(\varepsilon_1) = a_{11}\varepsilon_1 + a_{21}\varepsilon_2 + \cdots a_{n1}\varepsilon_n \\ \sigma(\varepsilon_2) = a_{12}\varepsilon_1 + a_{22}\varepsilon_2 + \cdots a_{n2}\varepsilon_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \sigma(\varepsilon_n) = a_{1n}\varepsilon_1 + a_{2n}\varepsilon_2 + \cdots a_{nn}\varepsilon_n \end{cases} \quad (*) \quad 记 \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

则(*)可形式地写为

$$\sigma(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (\sigma(\varepsilon_1), \sigma(\varepsilon_2), \cdots, \sigma(\varepsilon_n)) = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A$$

称A为线性变换σ在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的矩阵表示.

特别地，恒等变换在任意基下的矩阵为单位矩阵；
零变换任意基下的矩阵为零矩阵；
数乘变换在任意基下的矩阵为数量矩阵.

问题：数域P上n维线性空间V 的线性变换σ,在取定基后有唯一确定的n阶矩阵A与之对应; 反过来, 对数域P上的n 阶矩阵A, 是否存在唯一的线性变换σ，使σ在某基下的矩阵恰好是A? 结论是肯定的.

线性变换σ被基向量的像唯一确定！

定理1: 设$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$是数域P上$n$维线性空间V 的一组基, $\alpha_1, \alpha_2, \cdots, \alpha_n$是V中任意n个向量,则存在唯一的线性变换σ使

$$\sigma(\varepsilon_j)= \alpha_j, \quad j=1,2,\cdots,n.$$

证明：(i)存在性

$$\forall \xi = \sum_{i=1}^{n} x_i \varepsilon_i \in V,\text{定义}V\text{的变换}\sigma\text{：}\ \xi \to \sum_{i=1}^{n} x_i \alpha_i\ .$$

则σ是V的线性变换,且 $\sigma(\varepsilon_j)=\alpha_j$， $j=1,2,\cdots,n.$

事实上， $\forall \beta = \sum_{i=1}^{n} b_i \varepsilon_i, \gamma = \sum_{i=1}^{n} c_i \varepsilon_i \in V$ 及 $\forall k \in P$,

$$\beta + \gamma = \sum_{i=1}^{n} (b_i + c_i)\varepsilon_i\ , k\beta = \sum_{i=1}^{n} k b_i \varepsilon_i\ ,\text{由}\sigma\text{的定义，有}$$

$$\sigma(\beta+\gamma) = \sum_{i=1}^{n} (b_i + c_i)\alpha_i = \sum_{i=1}^{n} b_i \alpha_i + \sum_{i=1}^{n} c_i \alpha_i = \sigma(\beta) + \sigma(\gamma)$$

$$\sigma(k\beta) = \sum_{i=1}^{n} k b_i \alpha_i = k \sum_{i=1}^{n} b_i \alpha_i = k\sigma(\beta)$$

因此,σ是V的线性变换. 下证 $\sigma(\varepsilon_j)=\alpha_j$， $j=1,2,\ldots,n.$

由于 $\varepsilon_j = 0\varepsilon_1 + \cdots + 0\varepsilon_{j-1} + 1\varepsilon_j + 0\varepsilon_{j+1} + \cdots + 0\varepsilon_n$, $j=1,2,\cdots,n.$

所以$\sigma(\varepsilon_j) = 0\alpha_1 + \cdots + 0\alpha_{j-1} + 1\alpha_j + 0\alpha_{j+1} + \cdots + 0\alpha_n$

$= \alpha_j$, $j=1,2,\cdots, n.$

(ii)唯一性 若还有 $\tau \in L(V)$,使 $\tau(\varepsilon_j) = \alpha_j$, $j=1,2,\cdots,n.$

$$\forall \xi = \sum_{i=1}^{n} x_i \varepsilon_i \in V\text{,由于}$$

$$\tau(\xi) = \sum_{i=1}^{n} x_i \tau(\varepsilon_i) = \sum_{i=1}^{n} x_i \alpha_i = \sum_{i=1}^{n} x_i \sigma(\varepsilon_i) = \sigma\left(\sum_{i=1}^{n} x_i \varepsilon_i\right) = \sigma(\xi)$$

因此, $\tau = \sigma$. ||

线性变换σ被基向量的像唯一确定!

有了以上的讨论，我们就可以来建立线性变换与矩阵的联系.

设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是数域$P$上的$n$维线性空间$V$的一组基，$\sigma$是$V$的一个线性变换.基向量的像在这一组基底下线性表示为

$$\begin{cases} \sigma(\varepsilon_1)=a_{11}\varepsilon_1+a_{21}\varepsilon_2+\cdots+a_{n1}\varepsilon_n \\ \sigma(\varepsilon_2)=a_{12}\varepsilon_1+a_{22}\varepsilon_2+\cdots+a_{n2}\varepsilon_n \\ \qquad\qquad\vdots \\ \sigma(\varepsilon_n)=a_{1n}\varepsilon_1+a_{2n}\varepsilon_2+\cdots+a_{nn}\varepsilon_n \end{cases}$$

用矩阵来表示就是

$$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_n))$$

$$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$$

其中

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

矩阵A称为σ在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵表示.

定理:设V是数域P上n维线性空间, σ是V的线性变换，且σ在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的矩阵为A,令

$$\varphi : L(V) \to P^{n\times n}, \sigma \xrightarrow{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n} A$$

则φ是L(V)到$P^{n\times n}$ 的双射.

证明: 由线性变换矩阵的定义, φ 是映射.

∀ σ, τ∈L(V),若 φ(σ)=φ(τ),则由线性变换矩阵的定义,

$$\sigma(\varepsilon_j)= \tau(\varepsilon_j), \quad j=1,2,\cdots,n.$$

由定理1的唯一性知σ=τ ,即φ是单射.又∀A∈ $P^{n\times n}$,令

$$\alpha_j=a_{1j}\varepsilon_1+a_{2j}\varepsilon_2+\cdots+a_{nj}\varepsilon_n \quad j=1,2,\cdots,n.$$

则由定理1知，存在V的线性变换σ使

$$\sigma(\varepsilon_j)= \alpha_j, \quad j=1,2,\cdots,n.$$

显然φ (σ)=A，故φ是满射.综上所述， φ是单射又是满射，故是双射. ||该定理也可述为

定理 在L(V)与$P^{n\times n}$之间建立了一个双射φ，它可以将线性变换转化为矩阵问题，也可以将矩阵转化为线性变换问题. 这种对应的重要性还表现如下.

定理2 设$\varepsilon_1, \varepsilon_2,\ldots,\varepsilon_n$是数域P上n维线性空间V 的一组基, V的线性变换σ, τ在该基下的矩阵分别为A、B,则

(1) σ+ τ在基$\varepsilon_1, \varepsilon_2,\ldots,\varepsilon_n$下的矩阵A+B;

L(V)与$P^{n\times n}$ 同构.

(2) στ在基$\varepsilon_1, \varepsilon_2,\ldots,\varepsilon_n$下的矩阵AB ;

(3) kσ在基$\varepsilon_1, \varepsilon_2,\ldots,\varepsilon_n$下的矩阵kA;

(4)σ可逆⇔A可逆,且σ^{-1}在基$\varepsilon_1, \varepsilon_2,\ldots,\varepsilon_n$下的矩阵为$A^{-1}$;

(5)若ξ在基$\varepsilon_1, \varepsilon_2,\ldots,\varepsilon_n$下的坐标为$(x_1,x_2,\ldots,x_n)$,则σ(ξ)在该基下的坐标$(y_1,y_2,\ldots,y_n)$满足

可计算向量的像.

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

证明: 由假设 $\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$

$\tau(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B$

(1) $(\sigma+\tau)(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)+\tau(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)$

$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A+(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B$

$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(A+B)$

(2) $(\sigma\tau)(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=\sigma(\tau(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n))=\sigma((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)B)$

$=(\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n))B=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AB$

(3) $k\sigma=(kI)\sigma$,而$(kI)(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)kE$;由(2)

得$k\sigma=(kI)\sigma$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵$kA$

④ 由于单位变换(恒等变换) E对应于单位矩阵**E**.

所以， $\sigma\tau = \tau\sigma = E$

与 **AB＝BA＝E**

相对应.

因此，可逆线性变换 σ与可逆矩阵**A**对应，且

逆变换 σ^{-1}对应于逆矩阵 A^{-1}.

(5) 由假设 $$\xi=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

于是

$$\sigma(\xi)=(\sigma(\varepsilon_1),\sigma(\varepsilon_2),\cdots,\sigma(\varepsilon_n))\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

另一方面，由假设

$$\sigma(\xi)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$$

由向量坐标的唯一性，得 $$\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$$

注： $L(V) \cong P^{n\times n}$ ； $\dim L(V) = n^2$.

事实上，任意取定V的一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 后，

对任意 $\sigma \in L(V)$，定义 φ：

$$\varphi: \; L(V) \to P^{n\times n}, \qquad \varphi(\sigma) = A,$$

这里A为σ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵.

则 φ 就是$L(V)$到 $P^{n\times n}$的一个同构映射.

3．线性变换矩阵与向量在线性变换下的象

定理3 设线性变换 σ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为A，

$\xi\in V$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标为 $(x_1,x_2,\cdots,x_n)$，

$\sigma(\xi)$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标为 $(y_1,y_2,\cdots,y_n)$，

则有

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

证：由已知有

$$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

$$\xi=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},$$

$$\sigma(\xi)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}.$$

又

$$\sigma(\xi)=\left(\sigma\varepsilon_1,\sigma\varepsilon_2,\cdots,\sigma\varepsilon_n\right)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

$$\therefore\quad\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=\left(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n\right)A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

由于$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 线性无关，所以 $\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}=A\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.$

4. 同一线性变换在不同基下矩阵之间的关系

定理4 设线性空间V的线性变换 σ在两组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n \qquad \text{(I)}$$

$$\eta_1,\eta_2,\cdots,\eta_n \qquad \text{(II)}$$

下的矩阵分别为A、B，且从基(I) 到基(II)的过渡

矩阵矩阵是X，则

$$B = X^{-1}AX.$$

证：由已知，有

$$\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$$

$$\sigma(\eta_1,\eta_2,\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)B,$$

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X.$$

于是，$\sigma(\eta_1,\eta_2,\cdots,\eta_n)=\sigma(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)X$

$$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX=(\eta_1,\eta_2,\cdots,\eta_n)X^{-1}AX.$$

由此即得 $B=X^{-1}AX.$

三、相似矩阵

1. 定义

设A、B为数域P上的两个n级矩阵，若存在可逆

矩阵 $X \in P^{n\times n}$, 使得

$$B = X^{-1}AX$$

则称矩阵A相似于B，记为 $A \sim B.$

2．基本性质

（1）相似是一个等价关系，即满足如下三条性质：

① 反身性：$A \sim A.$

$$\left(\because \quad A = E^{-1}AE.\right)$$

② 对称性：$A \sim B \Rightarrow B \sim A.$

$$\left(\because \quad B = X^{-1}AX \Rightarrow A = Y^{-1}BY, Y = X^{-1}.\right)$$

③ 传递性：$A \sim B, B \sim C \Rightarrow A \sim C.$

$$\Big(\because \quad B = X^{-1}AX,\ C = Y^{-1}BY$$

$$\Rightarrow C = Y^{-1}BY = Y^{-1}(X^{-1}AX)Y = (XY)^{-1}A(XY).\Big)$$

定理5　数域P上线性空间V的线性变换在不同组基下的矩阵是相似的；反之，相似矩阵可以看作同一个线性变换在不同组基下的矩阵.

证明：前一部分为定理4，仅需证定理的后一部分.

设数域P上n级矩阵A相似于B, 则存在P上的n级可逆矩阵X,

使　$B=X^{-1}AX$.

设线性变换σ在基$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$下的矩阵为$A$,则令

$$(\eta_1, \eta_2, \cdots, \eta_n)=(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X$$

知$\eta_1, \eta_2, \cdots, \eta_n$也是V的一组基，且

$$\begin{aligned}\sigma(\eta_1,\eta_2,\cdots,\eta_n)&=\sigma[(\varepsilon_1, \varepsilon_2, \cdots,\varepsilon_n)X]\\&=[\sigma(\varepsilon_1,\varepsilon_2,\cdots, \varepsilon_n)]X\\&=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AX\\&=(\eta_1, \eta_2,\cdots,\eta_n)X^{-1}AX\\&=(\eta_1, \eta_2,\cdots,\eta_n)B.\end{aligned}$$

即σ在基$\eta_1,\eta_2,\cdots,\eta_n$下的矩阵为$B$. //

定理6（相似矩阵的性质）设$A\sim B$，则

(i) $|A|=|B|$

(ii) $A^m\sim B^m$;

(iii) $kA\sim kB$;

(iv)对 $f(x)\in P[x]$,有$f(A)\sim f(B)$;

(v) 若A可逆，则$A^{-1}\sim B^{-1}$;

(vi) $r(A)=r(B)$;

(vii) $B_1=X^{-1}A_1X$, $B_2=X^{-1}A_2X$, 则

$$B_1+B_2=X^{-1}(A_1+A_2)X, \quad B_1B_2=X^{-1}A_1A_2X.$$

证明：由 $A\sim B$，有可逆矩阵X，使

$$B=X^{-1}AX$$

(i) $|B| = |X^{-1}AX| = |X^{-1}||A||X| = |A|$;

(ii) $B^m = (X^{-1}AX)^m = (X^{-1}AX)(X^{-1}AX)\cdots(X^{-1}AX)$

$= X^{-1}A^mX$, 即$A^m \sim B^m$;

(iii) $X^{-1}(kA)X = k(X^{-1}AX) = kB$，即 $kA \sim kB$;

(iv) 由(ii)、(iii)及矩阵的运算性质,得$f(A) \sim f(B)$;

(v) $B^{-1} = (X^{-1}AX)^{-1} = X^{-1}A^{-1}(X^{-1})^{-1} = X^{-1}A^{-1}X$,

即$A^{-1} \sim B^{-1}$;

(vi) $A \sim B$时，A与B等价，从而$r(A) = r(B)$;

(vii) $B_1 = X^{-1}A_1X$, $B_2 = X^{-1}A_2X$, 则

$$B_1 + B_2 = X^{-1}A_1X + X^{-1}A_2X = X^{-1}(A_1 + A_2)X,$$

$$B_1B_2 = (X^{-1}A_1X)(X^{-1}A_2X) = X^{-1}(A_1A_2)X. \qquad \|$$

例1 已知$\varepsilon_1, \varepsilon_2$及$\eta_1, \eta_2$是数域P上二维线性空间V的两组基，σ是V的一个线性变换.已知σ在基$\varepsilon_1, \varepsilon_2$下的矩阵为A，且 $(\eta_1, \eta_2)=(\varepsilon_1, \varepsilon_2)X$
(1)求σ在基下η_1, η_2的矩阵B; (2)求A^k. 这里

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

解: (1)依题意，有 $B=X^{-1}AX$，即

$$B=\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}^{-1}\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}=\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

(2) $\mathrm{A}=\mathrm{XBX}^{-1} \Rightarrow \mathrm{A}^k=\mathrm{XB}^k\mathrm{X}^{-1}$,即

$$A^k=\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^k\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}=\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}=\begin{bmatrix} k+1 & k \\ -k & -k+1 \end{bmatrix}.$$