矩阵分析与应用

第四讲　　线性变换之二

信息工程学院
吕旋阳

本讲主要内容

- 线性变换在给定基下的矩阵
- 线性变换的矩阵表示
- 象与原象坐标间的关系
- 线性变换在不同基下矩阵之间的关系
- 线性变换的特征值与特征向量

- 设$x_1,x_2,\cdots,x_n$是线性空间V^n的一组基，T为V^n的线性变换. 则对任意 $x\in V^n$ 存在唯一的一组数$k_1,k_2,\cdots,k_n\in K$，使 $x=k_1x_1+k_2x_2+\cdots+k_nx_n$

从而，$T(x)=k_1T(x_1)+k_2T(x_2)+\cdots+k_nT(x_n)$.

- 由此知，$T(x)$ 由$T(x_1),T(x_2),\cdots,T(x_n)$完全确定.

设 $x_1, x_2, \cdots, x_n$ 是线性空间V的一组基，T_1, T_2 为V的线性变换，若 $T_1(x_i) = T_2(x_i)$，$i = 1, 2, \cdots, n.$

则 $T_1 = T_2$

证：对 $\forall x \in V$，$x = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$

$$T_1(x) = k_1 T_1(x_1) + k_2 T_1(x_2) + \cdots + k_n T_1(x_n)$$

$$T_2(x) = k_1 T_2(x_1) + k_2 T_2(x_2) + \cdots + k_n T_2(x_n)$$

由已知，即得 $T_1(x) = T_2(x)$ $\quad \therefore \; T_1 = T_2$

- 由此知，一个线性变换完全由它在一组基上的作用所决定.

设 $x_1, x_2, \cdots, x_n$是线性空间V的一组基，对V中

任意n个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$，都存在线性变换$T$使

$$T(x_i) = \alpha_i, \qquad i = 1, 2, \cdots, n$$

证：$\forall y \in V$， $y = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$

定义 $T: V \to V$， $T(y) = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$,

易知 T 为V的一个变换，下证它是线性的.

任取 y, $z \in V$，设 $y = \sum_{i=1}^{n} b_i x_i$， $z = \sum_{i=1}^{n} c_i x_i$

则 $y+z=\sum_{i=1}^{n}(b_i+c_i)\ x_i,\qquad ky=\sum_{i=1}^{n}(kb_i)x_i$

于是 $T\left(y+z\right)=\sum_{i=1}^{n}(b_i+c_i)\ \alpha_i=\sum_{i=1}^{n}b_i\alpha_i+\sum_{i=1}^{n}c_i\alpha_i$

$$=T\left(y\right)+T\left(z\right)$$

$$T\left(kx\right)=\sum_{i=1}^{n}(kb_i)\alpha_i=k\sum_{i=1}^{n}b_i\alpha_i=kT\left(y\right)$$

$\therefore$ T 为V的线性变换.

又 $x_i=0x_1+\cdots+0x_{i-1}+x_i+0x_{i+1}+\cdots+0x_n$

$\therefore$ $T(x_i)=\alpha_i,\qquad i=1,2,\cdots,n$

由此即得

定理 设 $x_1, x_2, \cdots, x_n$ 为线性空间V的一组基，对V中任意n个向量 $\alpha_1, \alpha_2, \cdots, \alpha_n$，存在唯一的线性变换$T$使

$$T\left(x_i\right) = \alpha_i, \qquad i = 1, 2, \cdots, n.$$

1．线性变换的矩阵

设 $x_1, x_2, \cdots, x_n$为数域P上线性空间V的一组基，T为V的线性变换．基向量的象可以被基线性表出,设

$$\begin{cases} T(x_1) = a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n \\ T(x_2) = a_{12}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n \\ \qquad\qquad \cdots\cdots \\ T(x_n) = a_{1n}x_1 + a_{2n}x_2 + \cdots + a_{nn}x_n \end{cases}$$

用矩阵表示即为

$$T\left(x_1, x_2, \cdots, x_n\right) = \left(Tx_1, Tx_2, \cdots, Tx_n\right) = \left(x_1, x_2, \cdots, x_n\right)A$$

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

矩阵A称为线性变换 T 在基 $x_1,x_2,\cdots,x_n$下的矩阵.

注： ① 给定V^n的基 $x_1,x_2,\cdots,x_n$ 和线性变换T,

矩阵A是唯一的.

② 单位变换在任意一组基下的矩阵皆为单位矩阵；

零变换在任意一组基下的矩阵皆为零矩阵；

数乘变换在任意一组基下的矩阵皆为数量矩阵；

例1. 设线性空间V^3的线性变换T为

$$T(x_1, x_2, x_3) = (x_1, x_2, x_1 + x_2)$$

求T在标准基 e_1, e_2, e_3 下的矩阵.

解： $\because \quad T(e_1) = T(1,0,0) = (1,0,1)$

$$T(e_2) = T(0,1,0) = (0,1,1)$$

$$T(e_3) = T(0,0,1) = (0,0,0)$$

$$\therefore \quad T(x_1, x_2, x_3) = (x_1, x_2, x_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$$

例2. 设线性空间$P_n[t]$的线性变换T为 $Tf(t)=f'(t)$

基Ⅰ为 $f_0=1, f_1=t, f_2=\dfrac{t^2}{2!},\cdots, f_n=\dfrac{t^n}{n!}$

基Ⅱ为 $g_0=1, g_1=t, g_2=t^2,\cdots, g_n=t^n$

记T在基Ⅰ下的矩阵为A_1，在基Ⅱ下的矩阵为A_2

$$Tf_0=0, Tf_1=f_0, Tf_2=f_1,\cdots, Tf_n=f_{n-1}$$

$$Tg_0=0, Tg_1=g_0, Tg_2=2g_1,\cdots, Tg_n=ng_{n-1}$$

$$\therefore\quad A_1=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 1 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}, A_2=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ & & & 0 & n \\ & & & & 0 \end{bmatrix}$$

线性变换T的值域： $R(T)=\{y \mid y=Tx, x\in V\}$

线性变换T的核： $N(T)=\{x \mid Tx=0, x\in V\}$

定理：线性空间V的线性变换T的值域和核都是V的线性子空间

$$V\neq\varnothing\Rightarrow R(T)\neq\varnothing,$$

$$\forall y_1\in R(T)\Rightarrow \exists x_1\in V, \text{st } y_1\in Tx_1$$

$$\forall y_2\in R(T)\Rightarrow \exists x_2\in V, \text{st } y_2\in Tx_2$$

$$y_1+y_2=Tx_1+Tx_2=T\left(x_1+x_2\right)\in R(T)$$

$$ky_1=k\left(Tx_1\right)=T\left(kx_1\right)\in R(T)$$

所以，$R(T)$ 是V的线性子空间

定义：线性变换T的值域 $R(T)$的维数称为T的秩

T的核 $N(T)$ 的维数称为T的亏（零度）

例: 在线性空间 $P_n[x]$ 中，令

$$T\big(f(x)\big)=f'(x)$$

则 $$T\big(P_n[x]\big)=P_{n-1}[x]$$

$$N(T)=P$$

所以T的秩为$n-1$，T的零度为1.

定理: 设T是n 维线性空间V的线性变换，$x_1, x_2, \cdots, x_n$

是V的一组基，T在这组基下的矩阵是A，则

1）T 的值域 $R(T)$ 是由基象组生成的子空间，即

$$R(T) = L\big(T(x_1), T(x_2), \cdots, T(x_n)\big)$$

2）T 的秩＝A的秩.

证：1） $\forall y \in V$，设 $y = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$，

于是 $$T(y) = k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n)$$
$$\in L\big(T(x_1), T(x_2), \cdots, T(x_n)\big)$$

即 $$R(T) \subseteq L\big(T(x_1), T(x_2), \cdots, T(x_n)\big)$$

又对 $\forall k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n)$

有 $$k_1 T(x_1) + k_2 T(x_2) + \cdots + k_n T(x_n)$$
$$= T(k_1 x_1 + k_2 x_2 + \ldots + k_n x_n) \in R(T)$$

$$\therefore \quad L\big(T(x_1), T(x_2), \cdots, T(x_n)\big) \subseteq R(T).$$

因此，$R(T) = L\big(T(x_1), T(x_2), \cdots, T(x_n)\big)$.

2）由1），T 的秩等于基象组 $T(x_1),T(x_2),\cdots,T(x_n)$ 的秩，又

$$\big(T(x_1),T(x_2),\cdots,T(x_n)\big)=(x_1,x_2,\cdots,x_n)A$$

$T(x_1),T(x_2),\cdots,T(x_n)$ 的秩等于矩阵A的秩

$\therefore$ 秩(T)＝秩(A).

设T为n维线性空间V的线性变换，则

$$T\text{的秩}+T\text{的零度}=n$$

即 $\dim R(T)+\dim N(T)=n.$

证明：设T的零度等于r，在核$N(T)$中取一组基

$x_1, x_2, \cdots, x_r$

并把它扩充为V的一组基：$x_1, x_2, \cdots, x_r, \cdots, x_n$

$R(T)$ 是由基象组 $T(x_1), T(x_2), \cdots, T(x_n)$ 生成的.

但 $T(x_i)=0,\qquad i=1,2,\cdots,r.$

$\therefore\quad R(T)=L\left(T(x_{r+1}),\cdots,T(x_n)\right)$

下证 $T(x_{r+1}),\cdots,T(x_n)$ 为 $R(T)$ 的一组基，即证它们线性无关.

设 $\quad k_{r+1}T(x_{r+1})+\cdots+k_nT(x_n)=0$

则有 $T\left(k_{r+1}x_{r+1}+\cdots+k_nx_n\right)=0$

$\therefore\quad y=k_{r+1}x_{r+1}+\cdots+k_nx_n\in N(T)$

即 y 可被 $x_1,x_2,\cdots,x_r$ 线性表出.

设 $y = k_1 x_1 + k_2 x_2 + \cdots + k_r x_r$

于是有 $k_1 x_1 + k_2 x_2 + \cdots + k_r x_{r,} - k_{r+1} x_{r+1} - \cdots - k_n x_n = 0$

由于 $x_1, x_2, \cdots, x_n$ 为V的基.

$\therefore k_1 = k_2 = \cdots = k_n = 0$

故$T(x_{r+1}),\cdots,T(x_n)$线 性无关，即它为$R(T)$的一组基.

$\therefore$ T 的秩＝$n-r$.

因此，T 的秩＋ T 的零度＝n.

注意：

虽然 $R(T)$ 与 $N(T)$ 的维数之和等于n，但是

$R(T)+N(T)$ 未必等于V.

例: 在线性空间 $P_n[x]$ 中，令 $T\big(f(x)\big)=f'(x)$

则 $R\big(T\big)=P_{n-1}[x]$ $N(T)=P$

$$R(T)+N(T)=P_{n-1}[x]$$

定理：设$x_1, x_2, \cdots, x_n$为数域K上线性空间V的一组基，在这组基下，V的每一个线性变换都与 $P^{n\times n}$ 中的唯一一个矩阵对应，且具有以下性质：

① 线性变换的和对应于矩阵的和；

②线性变换的数量乘积对应于矩阵的数量乘积；

③线性变换的乘积对应于矩阵的乘积；

④ 可逆线性变换与可逆矩阵对应，且逆变换对应于逆矩阵.

证：设T_1,T_2为两个线性变换，它们在基$x_1,x_2,\cdots,x_n$下的矩阵分别为A、B，即

$$T_1\left(x_1,x_2,\cdots,x_n\right)=\left(x_1,x_2,\cdots,x_n\right)A$$

$$T_2\left(x_1,x_2,\cdots,x_n\right)=\left(x_1,x_2,\cdots,x_n\right)B$$

① $\because\ \left(T_1+T_2\right)\left(x_1,x_2,\cdots,x_n\right)$

$$=T_1\left(x_1,x_2,\cdots,x_n\right)+T_2\left(x_1,x_2,\cdots,x_n\right)$$

$$=\left(x_1,x_2,\cdots,x_n\right)A+\left(x_1,x_2,\cdots,x_n\right)B$$

$$=\left(x_1,x_2,\cdots,x_n\right)\left(A+B\right)$$

$\therefore\ T_1+T_2$在基$x_1,x_2,\cdots,x_n$下的矩阵为A＋B.

② $\because \ (T_1T_2)(x_1,x_2,\cdots,x_n)=T_1(T_2(x_1,x_2,\cdots,x_n))$

$$=T_1((x_1,x_2,\cdots,x_n)B)=T_1(x_1,x_2,\cdots,x_n)B$$

$$=(x_1,x_2,\cdots,x_n)(AB)$$

$\therefore \ T_1T_2$在基 $x_1,x_2,\cdots,x_n$下的矩阵为AB.

③ $\because \ (kT_1)(x_1,x_2,\cdots,x_n)=((kT_1)(x_1),\cdots,(kT_1)(x_n))$

$$=(kT_1(x_1),\cdots,kT_1(x_n))=k(T_1(x_1),\cdots,T_1(x_n))$$

$$=kT_1(x_1,x_2,\cdots,x_n)=k(x_1,x_2,\cdots,x_n)A$$

$$=(x_1,x_2,\cdots,x_n)(kA)$$

$\therefore \quad kT_1$ 在基$x_1,x_2,\cdots,x_n$下的矩阵为 kA.

④ 由于单位变换(恒等变换) T_e对应于单位矩阵E.

所以， $T_1T_2 = E$

与 AB＝BA＝E

相对应.

因此，可逆线性变换 T_1与可逆矩阵A对应，且逆变换 T_1^{-1} 对应于逆矩阵 A^{-1}.

推论： $f(t)=a_m t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0, t \in K$

T为线性空间V的线性变换，且对V的基 $x_1, x_2, \cdots, x_n$

有 $T\left(x_1, x_2, \cdots, x_n\right) = \left(x_1, x_2, \cdots, x_n\right) A$

则V的线性变换 $f(T)$ 在基 $x_1, x_2, \cdots, x_n$ 下的矩阵是

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

定理：设线性变换 T 在基 $x_1, x_2, \cdots, x_n$ 下的矩阵为A,

$x \in V$ 在基 $x_1, x_2, \cdots, x_n$ 下的坐标为 $(\xi_1, \xi_2, \cdots, \xi_n)^T$,

$T(x)$ 在基 $x_1, x_2, \cdots, x_n$ 下的坐标为 $(\eta_1, \eta_2, \cdots, \eta_n)^T$,

则有

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = A \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}.$$

证：由已知有

$$x = \xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_n x_n = (x_1, x_2, \cdots, x_n)\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

$$T(x) = T(x_1, x_2, \cdots, x_n)\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

又 $T(x) = (x_1, x_2, \cdots, x_n)(\eta_1, \eta_2, \cdots, \eta_n)^T$

所以 $$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} = A\begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

定理： 设线性空间V的线性变换T在两组基

$$x_1, x_2, \cdots, x_n \qquad (\mathrm{I})$$

$$y_1, y_2, \cdots, y_n \qquad (\mathrm{II})$$

下的矩阵分别为A、B，且从基(Ⅰ)到基(Ⅱ)的过渡矩阵矩阵是C，则

$$B = C^{-1}AC$$

证：由已知，有

$$T(x_1,x_2,\cdots,x_n)=(x_1,x_2,\cdots,x_n)A,$$

$$T(y_1,y_2,\cdots,y_n)=(y_1,y_2,\cdots,y_n)B,$$

$$(y_1,y_2,\cdots,y_n)=(x_1,x_2,\cdots,x_n)C$$

于是，$T(y_1,y_2,\cdots,y_n)=T(x_1,x_2,\cdots,x_n)C$

$$=(x_1,x_2,\cdots,x_n)AC=(y_1,y_2,\cdots,y_n)C^{-1}AC$$

由此即得 $B=C^{-1}AC$

设A、B为数域K上的两个n阶矩阵，若存在可逆矩阵 $P \in K^{n\times n}$，使得 $B = P^{-1}AP$

则称矩阵A相似于B，记为 $A \sim B.$

相似是一个等价关系，即满足如下三条性质：

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

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

$\left(\because\ B = P^{-1}AP \Rightarrow A = Q^{-1}BQ,\ Q = P^{-1}.\right)$

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

$\left(\because\ B=P^{-1}AP,\ C=Q^{-1}BQ\right.$

$\left.\Rightarrow C = Q^{-1}BQ = Q^{-1}(P^{-1}AP)Q = (PQ)^{-1}A(PQ)\right)$

定理： 线性变换在不同基下的矩阵是相似的；反过来，如果两个矩阵相似，那么它们可以看作同一线性变换在两组基下所对应的矩阵.

证：前一部分显然成立. 下证后一部分.

设 $A \sim B$,且A是线性变换 T 在基 $x_1, x_2, \cdots, x_n$ 下的矩阵.

$\because B=C^{-1}AC$, 令 $\left(y_1, y_2, \cdots, y_n\right)=\left(x_1, x_2, \cdots, x_n\right)C$

显然, $y_1, y_2, \cdots, y_n$ 也是一组基，且 T 在这组基下的矩阵就是B.

相似矩阵的运算性质

① 若 $B_1 = C^{-1}A_1C$，$B_2 = C^{-1}A_2C$， 则

$$B_1 + B_2 = C^{-1}(A_1 + A_2)C,$$

$$B_1B_2 = C^{-1}(A_1A_2)C.$$

即， $A_1 + A_2 \sim B_1 + B_2$，$A_1A_2 \sim B_1B_2$.

② 若 $B = C^{-1}AC$，$f(x) \in P[x]$，则

$$f(B) = C^{-1}f(A)C.$$

特别地， $B^m = C^{-1}A^mC.$

例：设 x_1, x_2为线性空间V一组基， 线性变换T 在这组基下的矩阵为 $A=\begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}$，

y_1, y_2 为V的另一组基，且

$$(y_1, y_2)=(x_1, x_2)\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

（1）求T在y_1, y_2下的矩阵B.

（2）求 A^k.

解：（1）T在基 y_1, y_2 下的矩阵

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

$$=\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}\begin{pmatrix}2 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.$$

（2）由 $B=C^{-1}AC$， 有 $A=CBC^{-1}$，

于是 $A^k=CB^kC^{-1}$.

$$\therefore\ A^k=\begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}^k\begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}^{-1}$$

$$=\begin{pmatrix}1 & -1\\ -1 & 2\end{pmatrix}\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}\begin{pmatrix}2 & 1\\ 1 & 1\end{pmatrix}=\begin{pmatrix}k+1 & k\\ -k & -k+1\end{pmatrix}.$$

例：在线性空间P^3中，线性变换T定义如下：

$$\begin{cases} T(y_1)=(-5,0,3) \\ T(y_2)=(0,-1,6) \\ T(y_3)=(-5,-1,9) \end{cases}$$

其中，$\begin{cases} y_1=(-1,0,2) \\ y_2=(0,1,1) \\ y_3=(3,-1,0) \end{cases}$

（1）求T在标准基 x_1,x_2,x_3 下的矩阵.

（2）求T 在 y_1,y_2,y_3 下的矩阵.

解：（1）由已知，有

$$(y_1, y_2, y_3) = (x_1, x_2, x_3)\begin{pmatrix} -1 & 0 & 3 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{pmatrix} \triangleq (x_1, x_2, x_3)C,$$

$$T(y_1, y_2, y_3) = (x_1, x_2, x_3)\begin{pmatrix} -5 & 0 & -5 \\ 0 & -1 & -1 \\ 3 & 6 & 9 \end{pmatrix},$$

设T 在标准基 x_1, x_2, x_3下的矩阵为A，即

$$T(x_1, x_2, x_3) = (x_1, x_2, x_3)A$$

$$\therefore T(y_1, y_2, y_3) = T((x_1, x_2, x_3)C) = T(x_1, x_2, x_3)C$$

$$= (x_1, x_2, x_3)AC$$

因而，$$AC=\begin{pmatrix}-5 & 0 & -5\\ 0 & -1 & -1\\ 3 & 6 & 9\end{pmatrix},$$

$$\therefore\quad A=\begin{pmatrix}-5 & 0 & -5\\ 0 & -1 & -1\\ 3 & 6 & 9\end{pmatrix}C^{-1}=\begin{pmatrix}-5 & 0 & -5\\ 0 & -1 & -1\\ 3 & 6 & 9\end{pmatrix}\begin{pmatrix}-1 & 0 & 3\\ 0 & 1 & -1\\ 2 & 1 & 0\end{pmatrix}^{-1}$$

$$=\frac{1}{7}\begin{pmatrix}-5 & 20 & -20\\ -4 & -5 & -2\\ 27 & 18 & 24\end{pmatrix}$$

（2）设 T 在 y_1,y_2,y_3下的矩阵为B，则A与B相似，且

$$B = C^{-1}AC$$

$$\therefore \quad B = \begin{pmatrix} -1 & 0 & 3 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} -5 & 0 & -5 \\ 0 & -1 & -1 \\ 3 & 6 & 9 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 3 & 5 \\ -1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

本讲主要内容

- 线性变换在给定基下的矩阵
- 线性变换的矩阵表示
- 象与原象坐标间的关系
- 线性变换在不同基下矩阵之间的关系
- 线性变换的特征值与特征向量

引入

有限维线性空间V中取定一组基后，V的任一线性变换都可以用矩阵来表示．为了研究线性变换性质，希望这个矩阵越简单越好，如对角矩阵．

从本节开始，我们主要讨论，如何选择一组适当的基，使V的某个线性变换在这组基下的矩阵就是一个对角矩阵?

一、特征值与特征向量

定义：设 T 是数域K上线性空间V的一个线性变换，

若对于K中的一个数 λ_0，存在一个V的非零向量 x，

使得

$$T(x)=\lambda_0 x \quad ,$$

则称 λ_0 为 T 的一个特征值，称 x 为 T 的属于特征值

λ_0 的特征向量.

注：① 几何意义：特征向量经线性变换后方向保持相同 $(\lambda_0 > 0)$ 或相反 $(\lambda_0 < 0)$. $\lambda_0 = 0$ 时，$T(x) = 0$.

② 若 x 是 T 的属于特征值 λ_0 的特征向量，则 $kx \quad (k \in K, k \neq 0)$ 也是 T 的属于 λ_0 的特征向量.

$$\left(\because \quad T(kx) = kT(x) = k(\lambda_0 x) = \lambda_0 (kx) \right)$$

由此知，特征向量不是被特征值所唯一确定的，但是特征值却是被特征向量所唯一确定的，即若 $T(x) = \lambda x$ 且 $T(x) = \mu x$，则 $\lambda = \mu$.

二、特征值与特征向量的求法

分析：设 $\dim V = n$， $x_1, x_2, \cdots, x_n$ 是V的一组基，

线性变换 T 在这组基下的矩阵为A.

设 λ_0是 T 的特征值，它的一个特征向量 x在基

$x_1, x_2, \cdots, x_n$下的坐标记为 $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

则 $T(x)$在基 $x_1, x_2, \cdots, x_n$下的坐标为 $A\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$,

而$\lambda_0 x$的坐标是 $\lambda_0\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$, 又 $T(x)=\lambda_0 x$

于是 $A\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=\lambda_0\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$, 从而 $(\lambda_0 I-A)\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}=0.$

即 $\begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}$ 是线性方程组 $(\lambda_0 I-A)X=0$ 的解,

又 $\because x\neq 0$, $\therefore \begin{pmatrix}\xi_1\\ \vdots\\ \xi_n\end{pmatrix}\neq 0$, $\therefore$ $(\lambda_0 I-A)X=0$ 有非零解.

所以它的系数行列式 $|\lambda_0 I-A|=0.$

以上分析说明：

若λ_0是T的特征值，则 $\left|\lambda_0 I - A\right| = 0.$

反之，若$\lambda_0 \in K$满足 $\left|\lambda_0 I - A\right| = 0,$

则齐次线性方程组 $(\lambda_0 I - A)X = 0$ 有非零解.

若 $(\xi_1, \xi_2, \cdots, \xi_n)^T$ 是 $(\lambda_0 I - A)X = 0$ 一个非零解，

则向量 $x = \xi_1 x_1 + \cdots + \xi_n x_n$ 就是T的属于λ_0的一个特征向量.

设 $A \in K^{n\times n}$，λ是一个参数，矩阵 $\lambda I - A$ 称为

A的特征矩阵，它的行列式

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

称为A的特征多项式.

（ $|\lambda I - A|$ 是数域K上的一个n次多项式）

注：① 若矩阵A是线性变换 T 关于V的一组基的矩阵，而 λ_0 是 T 的一个特征值，则 λ_0 是特征多项式 $|\lambda I - A|$ 的根，即 $|\lambda_0 I - A| = 0.$

反之，若 λ_0 是A的特征多项式的根，则 λ_0 就是 T 的一个特征值．（所以，特征值也称特征根.）

② 矩阵A的特征多项式的根有时也称为A的特征值，而相应的线性方程组 $(\lambda I - A)X = 0$ 的非零解也就称为A的属于这个特征值的特征向量.

求特征值与特征向量的一般步骤

i) 在V中任取一组基 $x_1, x_2, \cdots, x_n$ 写出 T 在这组基下的矩阵A .

ii) 求A的特征多项式 $|\lambda I - A|$ 在K上的全部根它们就是 T 的全部特征值.

iii) 把所求得的特征值逐个代入方程组

$$(\lambda I - A)X = 0$$

并求出它的一组基础解系.(它们就是属于这个特征值的全部线性无关的特征向量在基 $x_1, x_2, \cdots, x_n$ 下的坐标.)

例:在线性空间V中，数乘变换K在任意一组基下的矩阵都是数量矩阵kI，它的特征多项式是

$$\left|\lambda I - kI\right| = (\lambda - k)^n.$$

故数乘法变换K的特征值只有数k，且

对 $\forall x \in V \quad (x \neq 0)$，皆有 $K(x) = kx.$

所以，V中任一非零向量皆为数乘变换K的特征向量.

例: 设线性变换 T 在基 x_1, x_2, x_3下的矩阵是

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix},$$

求 T 特征值与特征向量.

解：A的特征多项式

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 & -2 \\ -2 & \lambda - 1 & -2 \\ -2 & -2 & \lambda - 1 \end{vmatrix} = (\lambda + 1)^2 (\lambda - 5)$$

故 T 的特征值为： $\lambda_1 = -1$（二重）， $\lambda_2 = 5$

把 $\lambda = -1$ 代入齐次方程组 $(\lambda I - A)X = 0$，得

$$\begin{cases} -2x_1 - 2x_2 - 2x_3 = 0 \\ -2x_1 - 2x_2 - 2x_3 = 0 \\ -2x_1 - 2x_2 - 2x_3 = 0 \end{cases} \quad 即 \quad x_1 + x_2 + x_3 = 0$$

它的一个基础解系为：$(1,0,-1)$，$(0,1,-1)$

因此，属于-1的两个线性无关的特征向量为

$$y_1 = x_1 - x_3,\quad y_2 = x_2 - x_3$$

而属于-1的全部特征向量为

$$k_1 y_1 + k_2 y_2, \qquad (k_1, k_2 \in K \text{ 不全为零})$$

把 $\lambda = 5$ 代入齐次方程组 $(\lambda I - A)X = 0$，得

$$\begin{cases} 4x_1 - 2x_2 - 2x_3 = 0 \\ -2x_1 + 4x_2 - 2x_3 = 0 \\ -2x_1 - 2x_2 + 4x_3 = 0 \end{cases}$$

解得它的一个基础解系为：(1,1,1)

因此，属于5的一个线性无关的特征向量为

$$y_3 = x_1 + x_2 + x_3$$

而属于5的全部特征向量为

$$k_3 y_3, \qquad (k_3 \in K, \ k_3 \neq 0)$$

定义：设T为n维线性空间V的线性变换，λ_0为T的一个特征值，令V_{λ_0}为T的属于λ_0的全部特征向量再添上零向量所成的集合，即 $V_{\lambda_0}=\{x \mid Tx=\lambda_0 x\}$

则V_{λ_0}是V的一个子空间, 称之为T的一个特征子空间.

$$\because \quad T(x+y)=T(x)+T(y)=\lambda_0 x+\lambda_0 y=\lambda_0(x+y)$$

$$T(kx)=kT(x)=k(\lambda_0 x)=\lambda_0(kx)$$

$$\therefore \quad x+y\in V_{\lambda_0}, \qquad kx\in V_{\lambda_0}$$

若 T 在n维线性空间V的某组基下的矩阵为A，则

$$\dim V_{\lambda_0} = n - \text{秩}(\lambda_0 I - A)$$

即特征子空间 V_{λ_0}的维数等于齐次线性方程组

$$(\lambda_0 I - A)X = 0 \qquad (*)$$

的解空间的维数，且由方程组(*)得到的属于λ_0的全部线性无关的特征向量就是 V_{λ_0}的一组基.

特征多项式的有关性质

1. 设 $A=\left(a_{ij}\right)\in K^{n\times n}$，则A的特征多项式

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

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

由多项式根与系数的关系还可得

① A的全体特征值的和＝$a_{11}+a_{22}+\cdots+a_{nn}$.

称之为A的迹,记作trA

② A的全体特征值的积＝ $|A|$.

定理：设 $A=\left(a_{ij}\right)_{n\times n}, B=\left(b_{ij}\right)_{n\times n}$ 则 $\mathrm{tr}\left(AB\right)=\mathrm{tr}\left(BA\right)$

证：令 $AB=\left(u_{ij}\right), BA=\left(v_{ij}\right)$，于是有

$$u_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj},\quad v_{ij}=\sum_{l=1}^{n}b_{il}a_{lj}$$

$$\therefore\quad \mathrm{tr}(AB)=\sum_{i=1}^{n}u_{ii}=\sum_{i=1}^{n}\left[\sum_{k=1}^{n}a_{ik}b_{ki}\right]$$

$$=\sum_{k=1}^{n}\left[\sum_{i=1}^{n}b_{ki}a_{ik}\right]=\sum_{k=1}^{n}v_{kk}=\mathrm{tr}(BA)$$

定理：相似矩阵有相似的迹　$\mathrm{tr}(AB) = \mathrm{tr}(BA)$

证：设 $A \sim B$，即 $\exists P \neq 0, \mathrm{st} \quad B = P^{-1}AP$

$$\begin{aligned}\mathrm{tr}(B) &= \mathrm{tr}(P^{-1}AP) \\ &= \mathrm{tr}(APP^{-1}) \\ &= \mathrm{tr}(A)\end{aligned}$$

定理： 相似矩阵具有相同的特征多项式.

证:设 $A \sim B$, 则存在可逆矩阵P，使得

$$B = P^{-1}AP$$

于是，

$$\begin{aligned}|\lambda I - B| &= |\lambda I - P^{-1}AP| \\ &= |\lambda P^{-1}IP - P^{-1}AP| \\ &= |P^{-1}(\lambda I - A)P| \\ &= |P^{-1}||\lambda I - A||P| \\ &= |\lambda I - A|\end{aligned}$$

注：① 由定理线性变换T的特征值与基的选择无关.

因此,矩阵A的特征多项式也说成是<u>线性变换 T 的特征多项式</u>；而线性变换 T 的特征值与特征向量有时也说成是<u>矩阵A的特征值与特征向量</u>.

② 有相同特征多项式的矩阵未必相似.

如 $A=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, B=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$

它们的特征多项式都是 $(\lambda-1)^2$; 但A、B不相似.

定理：任意n阶矩阵A与三角矩阵相似

证明:对阶数n利用数学归纳法证明. 当n=1时显然成立

设当阶数为n–1时定理成立。

设 $x_1, x_2, \cdots, x_n$ 是n个线性无关的列向量，

其中x_1 为A的特征值λ的特征向量, $Ax_1 = \lambda_1 x_1$

记 $P_1 = (x_1, x_2, \cdots, x_n)$

于是 $AP_1 = (Ax_1, Ax_2, \cdots, Ax_n) = (\lambda x_1, Ax_2, \cdots, Ax_n)$

由于$Ax_i \in C^n$,因此可以由 $x_1, x_2, \cdots, x_n$唯一地线性表示

即有 $Ax_i = b_{1i}x_1 + b_{2i}x_2 + \cdots + b_{ni}x_n$

于是 $AP_1 = \left(\lambda x_1, Ax_2, \cdots, Ax_n\right)$

$$= \left(x_1, x_2, \cdots, x_n\right)\begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & b_{nn} \end{bmatrix}$$

即

$$P_1^{-1}AP_1 = \begin{bmatrix} \lambda & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}$$

由归纳假定，对于n–1阶矩阵A_1存在矩阵Q使得

$$Q^{-1}A_1Q = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

记 $$P_2=\begin{bmatrix}1 & 0^T\\ 0 & Q\end{bmatrix},\quad P=P_1P_2$$

则有 $$P^{-1}AP=\left(P_1P_2\right)^{-1}A\left(P_1P_2\right)=P_2^{-1}\left(P_1^{-1}AP_1\right)P_2$$

$$=P_2^{-1}\begin{bmatrix}\lambda & b_{12} & \cdots & b_{1n}\\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{bmatrix}P_2$$

$$=\begin{bmatrix}\lambda_1 & * & \cdots & *\\ & \lambda_2 & \ddots & \vdots\\ & & \ddots & *\\ & & & \lambda_n\end{bmatrix}$$

Hamilton—Caylay定理 设 $A \in K^{n\times n}$ 其特征多项式

$$\varphi(\lambda)=|\lambda I-A|=\lambda^{n}+a_1\lambda^{n-1}+\cdots+a_{n-1}\lambda+a_n$$

则 $\varphi(A)=A^{n}+a_1A^{n-1}+\cdots+a_{n-1}A+a_nI=0.$

证明:A的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$ ，即

$$\varphi(\lambda)=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

必然存在可逆矩阵 $P_{n\times n}$ ，使得

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & * & \cdots & * \\ & \lambda_2 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n\end{bmatrix}$$

$$\varphi(P^{-1}AP)=(P^{-1}AP-\lambda_1 I)(P^{-1}AP-\lambda_2 I)\cdots(P^{-1}AP-\lambda_n I)$$

$$=\begin{bmatrix} 0 & * & \cdots & * \\ & \lambda_2-\lambda_1 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_1 \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_2 & * & \cdots & * \\ & 0 & \ddots & \vdots \\ & & \ddots & * \\ & & & \lambda_n-\lambda_2 \end{bmatrix}\cdots(P^{-1}AP-\lambda_n I)$$

$$=\begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \ddots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & & & * \end{bmatrix}\begin{bmatrix} \lambda_1-\lambda_3 & * & * & \cdots & * \\ & \lambda_2-\lambda_3 & * & \cdots & * \\ & & 0 & \ddots & \vdots \\ & & & \ddots & * \\ & & & & \lambda_2-\lambda_3 \end{bmatrix}\cdots(P^{-1}AP-\lambda_n I)$$

$$=0$$

即 $P^{-1}\varphi(A)P=0 \quad \Rightarrow \varphi(A)=0$

例: 设 $A=\begin{pmatrix}1&0&2\\0&-1&1\\0&1&0\end{pmatrix}$,求 $2A^8-3A^5+A^4+A^2-4I$

解：A的特征多项式 $\varphi(\lambda)=|\lambda I-A|=\lambda^3-2\lambda+1$

用$\varphi(\lambda)$去除 $2\lambda^8-3\lambda^5+\lambda^4+\lambda^2-4\lambda=g(\lambda)$,得

$$g(\lambda)=\varphi(\lambda)(2\lambda^5+4\lambda^3-5\lambda^2+9\lambda-14)+(24\lambda^2-37\lambda+10)$$

$\because\ \varphi(A)=0,$

$\therefore\ 2A^8-3A^5+A^4+A^2-4I=24A^2-37A+10I$

$$=\begin{pmatrix}-3&48&-26\\0&95&-61\\0&-61&34\end{pmatrix}$$

练习1：已知 $A \in P^{n\times n}$，λ 为A的一个特征值，则

（1） $kA \ (k \in P)$ 必有一个特征值为 $\underline{\quad k\lambda \quad}$；

（2） $A^m \ (m \in Z^+)$ 必有一个特征值为 $\underline{\quad \lambda^m \quad}$；

（3） A可逆时，A^{-1}必有一个特征值为 $\underline{\quad \lambda^{-1} \quad}$；

（4） A可逆时，A^*必有一个特征值为 $\underline{\quad \dfrac{|A|}{\lambda} \quad}$.

（5） $f(x) \in P[x]$，则 $f(A)$必有一个特征值为 $\underline{\quad f(\lambda) \quad}$.

练习2：已知3阶方阵A的特征值为：1、－1、2，

则矩阵 $\boldsymbol{B} = \boldsymbol{A}^3 - 2\boldsymbol{A}^2$ 的特征值为：$\underline{\quad -1, -3, 0 \quad}$，

行列式 $|\boldsymbol{B}| = \underline{\quad 0 \quad}$.

作业

- **P77：4，5**
- **P78：7，8**
- **P79：14，16**