3.矩阵与线性变换

一、线性变换的定义

观察线性函数 $f(x)=kx$知, f是从$R\to R$的一个映射,

$\forall x, y\in R, a\in R,$

$(T1)\,f(x+y)=k(x+y)=kx+ky\ =\ f(x)+f(y),$

$(T2)\,f(ax)=k(ax)=(ka)x=(ak)x=akx=a(kx)=af(x).$

对$Ax=b$的解集,可看成是在P^n到P^m的映射f：$x\ \to Ax$

下(注：$A_{m\times n}$),向量$b\in P^m$的原像.f 满足$(T1)(T2)$ 条件.

$(T1)\,f(x+y)=f(x)+f(y)$ （可加性）.

$(T2)\,f(ax)=af(x)$ （齐次性）.

如 (1)微积分中的求微分 积分的运算;

(2) $L[f(x)]=\int_0^{+\infty}f(x)e^{-px}\mathrm{d}x.$

返回

定义：设U与V是两个线性空间,U到V内的一个映射T满足$(T1)$和$(T2)$, 则称T是U到V内的线性变换(线性映射).

特别的$U=V$时,T是V自身的一个映射, 也称T是V自身的线性变换(线性算子).

$(T1)$和$(T2)$等价于：

$$T(k\alpha+l\beta)=kT(\alpha)+lT(\beta),\forall k,l\in P,\alpha,\beta\in V.$$

线性叠加原理：

$$T(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1T(\alpha_1)+k_2T(\alpha_2)+\cdots+k_sT(\alpha_s),$$

$$\forall k_j\in P,\alpha_j\in V.$$

返回

二、线性变换的性质

(1)$T(0)=0, T(-\alpha)=-T(\alpha)$;

$T(0)=0$的几何意义就是线性变换一定保持原点不动.

(2)若$\alpha_1,\alpha_2,\cdots\alpha_s$线性相关,则$T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_s)$线性相关.

(3)若$T(\alpha_1),T(\alpha_2),\cdots,T(\alpha_s)$线性无关,则$\alpha_1,\alpha_2,\cdots\alpha_s$线性无关.

由上面的性质知,线性变换把零向量变成零向量,把x的负向量变成像Tx的负向量,把线性相关的向量组变为线性相关的向量组.

三、线性变换的矩阵表示　如何确定一个线性变换T?

设$\alpha_1,\alpha_2,\cdots\alpha_n$是$n$维线性空间$V$的一组基, 设$T$是$V$的线性变换, $\forall\alpha\in V$,则　$\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n$

$$T\alpha=a_1(T\alpha_1)+a_2(T\alpha_2)+\cdots+a_n(T\alpha_n)$$

$$T\alpha_1,T\alpha_2,\cdots,T\alpha_n\in V,$$

$$T\alpha_1=a_{11}\alpha_1+a_{21}\alpha_2+\cdots+a_{n1}\alpha_n$$
$$T\alpha_2=a_{12}\alpha_1+a_{22}\alpha_2+\cdots+a_{n2}\alpha_n$$
$$\cdots\cdots\cdots\cdots$$
$$T\alpha_n=a_{1n}\alpha_1+a_{2n}\alpha_2+\cdots+a_{nn}\alpha_n$$

即

$$T(\alpha_1,\alpha_2,\cdots,\alpha_n)$$
$$=(T\alpha_1,T\alpha_2,\cdots,T\alpha_n)$$
$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$$

返回

即$T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(T\alpha_1,T\alpha_2,\cdots,T\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$

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

A的第i列恰是$T\alpha_i(i=1,2,\cdots n)$的坐标.

矩阵A称为T在V的基$\alpha_1,\alpha_2,\cdots,\alpha_n$下的矩阵,简称$A$为$T$的矩阵.

矩阵A与线性变换T一一对应.

在有限维线性空间中,线性变换与矩阵是一回事!!

基变换和坐标变换

设V是n维线性空间,$\alpha_1,\alpha_2,\cdots,\alpha_n$与$\beta_1,\beta_2,\cdots,\beta_n$是$V$的两组基,分别记为$\alpha-$基与$\beta-$基. 由基定义

$$
\begin{cases}
\beta_1 = p_{11}\alpha_1 + p_{21}\alpha_2 + \cdots + p_{n1}\alpha_n, \\
\beta_2 = p_{12}\alpha_1 + p_{22}\alpha_2 + \cdots + p_{n2}\alpha_n, \\
\qquad \cdots\cdots \\
\beta_n = p_{1n}\alpha_1 + p_{2n}\alpha_2 + \cdots + p_{nn}\alpha_n,
\end{cases}
$$

$$(\beta_1,\beta_2,\cdots,\beta_n) = (\alpha_1,\alpha_2,\cdots,\alpha_n)P,$$

可逆吗?

称n阶矩阵$P=(p_{ij})$称为由$\alpha-$基到$\beta-$基的过渡矩阵.

由$\beta-$基到$\alpha-$基的过渡矩阵是什么?

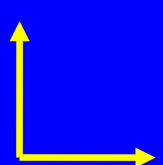
设$\gamma \in V$在α－基与β－基下的坐标分别是$x=(x_1,x_2,\cdots,x_n)^T$，$y=(y_1,y_2,\cdots,y_n)^T$，x和y有何联系？

$$\gamma = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n = (\alpha_1,\alpha_2,\cdots,\alpha_n)x,$$

$$\gamma = y_1\beta_1 + y_2\beta_2 + \cdots + y_n\beta_n = (\beta_1,\beta_2,\cdots,\beta_n)y,$$

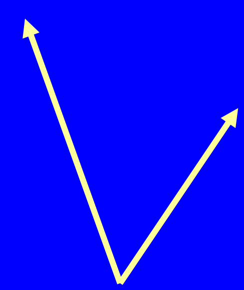
$$\gamma = (\alpha_1,\alpha_2,\cdots,\alpha_n)x = (\beta_1,\beta_2,\cdots,\beta_n)y = (\alpha_1,\alpha_2,\cdots,\alpha_n)Py$$

$\Rightarrow x = Py$或$y = P^{-1}x$. ——坐标变换公式



$\alpha_1 = (1,0), \alpha_2 = (0,1).$

$\beta_1 = (1,1), \beta_2 = (-1,2).$



$\beta_1 = \alpha_1 + \alpha_2,$

$\beta_2 = -\alpha_1 + 2\alpha_2.$

$$P = \begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$$

返回

线性空间有不同的基,在不同基下同一线性变换的矩阵有何联系?

α－基到β－基的过渡矩阵为P,

$$\because T(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$$

$$T(\beta_1,\beta_2,\cdots,\beta_n)=(\beta_1,\beta_2,\cdots,\beta_n)B=(\alpha_1,\alpha_2,\cdots,\alpha_n)PB,$$

$$\begin{aligned}T(\beta_1,\beta_2,\cdots,\beta_n)&=T[(\alpha_1,\alpha_2,\cdots,\alpha_n)P],\\&=T(\alpha_1,\alpha_2,\cdots,\alpha_n)P\\&=(\alpha_1,\alpha_2,\cdots,\alpha_n)AP\\&=(\alpha_1,\alpha_2,\cdots,\alpha_n)PB,\end{aligned}$$

$$\therefore AP=PB\Rightarrow B=P^{-1}AP.$$

返回

$$\Rightarrow B = P^{-1}AP.$$

相似矩阵的本质：同一线性变换在不同基下的矩阵！

相似矩阵具有相同的特征值！具有相同行列式和迹！

$$A\text{的迹}\,trA = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i.$$

设$A=(a_{ij})$, B为n阶方阵, k是数, 则

(1)$tr(A+B)=trA+trB$;(2)$tr(kA)=k(trA)$;(3)$tr(A^T)=trA$;

(4)$tr(AB)=tr(BA)$(此处只需$A_{m\times n}, B_{n\times m}$即可);

(5)$tr(AA^T)=\sum_{j=1}^{n}\sum_{i=1}^{n}|a_{ij}|^2$. 特别的$tr(AA^T)=0 \Leftrightarrow A=0$.

4 投影

在无线通信、雷达、时间序列分析和信号处理等领域中,许多问题的最优解都可归结为:

提取某个所希望的信号,而抑制掉其他所有干扰、杂波或者噪声. 投影是解决这类问题的一个特别重要的数学工具.

返回

一、投影算子与投影矩阵

设V_1和V_2均是V的子空间,且$V_1 \oplus V_2 = V$,则$\forall x \in V$都可以惟一的分解为　$x = y + z, y \in V_1, z \in V_2.$

称y是x沿着V_2到V_1的投影.

定义1 对$\forall x \in V$变为沿着V_2到V_1的投影变换称为沿着V_2到V_1的投影算子,记为T. 即　$Tx = y.$

由定义1知道,投影算子T将整个空间V变到子空间V_1.

即　$Tx = y$. 特别的若$\forall x \in V_1$,则$Tx = x.$

$\because T^2 = T$,所以T也叫幂等算子.

投影(幂等)算子T是线性变换.

定理1 设 T是线性空间V 上的投影,则 $V=R(T)\oplus N(T)$.

证明：设$\forall x\in V$,令$y=Tx, z=x-y$

$$\Rightarrow x=y+z, \text{其中} y\in R(T).$$

$\because Tz=T(x-y)=Tx-Ty=Tx-T^2x=Tx-Tx=0.$

$\therefore z\in N(T).\quad \Rightarrow V=R(T)+N(T).$

若$\forall x\in R(T)\cap N(T)\Rightarrow x\in R(T)$,故存在 y,使得

$$x=Ty\Rightarrow Tx=T^2y=Ty=x.$$

而$x\in N(T)\Rightarrow x=0$, 即$R(T)\cap N(T)=\{0\}$.

定理2 设$V=V_1\oplus V_2$,则$\exists$投影T,使得 $R(T)=V_1, N(T)=V_2$.

二、投影矩阵 (幂等矩阵) 的性质.

定理3 若$A \in C^{n\times n}$是幂等矩阵,即$A^2=A$,则有

(1) A^H与$I-A$也是幂等矩阵;

(2) A的特征值非0即1,且可对角化;

(3) $rank(A)=tr(A)$;

(4) $A(I-A)=(I-A)A=0$;

(5) $Ax=x \Leftrightarrow x\in R(A)$;

(6) $N(A)=R(I-A), R(A)=N(I-A)$.

证明:(2)设$Ax=\lambda x, x\neq 0 \Rightarrow A^2x=\lambda^2x \Rightarrow Ax=\lambda^2x$

$\Rightarrow \lambda x=\lambda^2x \Rightarrow \lambda=0$或者1.

设 $rank(A)=r$,考虑$Ax=0$基础解系$\xi_1,\cdots\xi_{n-r}$,

$\Rightarrow A\xi_i=0=0\xi_i, i=1,\cdots n-r.$

设$A=(\alpha_1,\cdots,\alpha_r,\cdots\alpha_n)$,

又$\because A^2=A \Rightarrow A^2=A(\alpha_1,\cdots,\alpha_r,\cdots\alpha_n)$,

$=(A\alpha_1,\cdots,A\alpha_r,\cdots A\alpha_n)$,

$=(\alpha_1,\cdots,\alpha_r,\cdots\alpha_n)$,

返回

由于$rank(A)=r$,不妨设$\alpha_1,\cdots,\alpha_r$线性无关,

$\Rightarrow A\alpha_i=\alpha_i=1\cdot\alpha_i,\quad i=1,\cdots r.$

$\Rightarrow A$有n个线性无关的向量$\Rightarrow A$可以对角化.

(3)因为$tr(A)=\sum_{i=1}^{n}\lambda_i=\sum_{i=1}^{r}1+\sum_{i=r+1}^{n}0=r$

$rank(A)$等于非零特征值的个数,1为r重根,

$\therefore rank(A)=r=tr(A).$

思考：能否说任何一个矩阵A的秩$rank(A)$均等于它的非零特征值的个数?

返回

$$(4)A(I-A)=(I-A)A=0;$$

$(5)\Rightarrow: Ax=x\Rightarrow x\in R(A);$ $\quad$ $(5)\ Ax=x\Leftrightarrow x\in R(A);$

$\Leftarrow: x\in R(A)\Rightarrow \exists y,$ 使得$Ay=x$

$$Ax=A^2y=Ay=x.$$

(6) 由(4)知, $\quad A(I-A)=0,$ $\quad$ $(6)\ N(A)=R(I-A), R(A)=N(I-A).$

$\Rightarrow \forall x\in R(I-A), \exists y, y\in C^n,$ 使得$x=(I-A)y,$

$$\Rightarrow Ax=0 \Rightarrow R(I-A)\subset N(A).$$

$$\Rightarrow \dim R(I-A)\le \dim N(A)=n-rank(A).$$

返回

$\Rightarrow$ 即 $rank(I-A) \le n - rank(A)$.

又 $\because I = A + (I-A)$,

$rank(A+B) \le rank(A) + rank(B)$.

$$\therefore n = rank I = rank[A + (I-A)]$$

$$\le rank(A) + rank(I-A)$$

$$\Rightarrow rank(I-A) = n - rank(A).$$

$\Rightarrow$ 即 $\dim R(I-A) = n - \dim R(A) = \dim N(A)$.

返回

5 特征值与特征向量

一、特征值与特征向量的概念

定义 设A是n阶方阵,若存在数λ和n维非零向量x使

$$Ax = \lambda x$$

则称λ为方阵A的一个特征值, x为方阵A对应于特征值λ的一个特征向量.

方阵A的所有特征值的集合称为A的谱,记为$\lambda(A)$.

$$Ax = \lambda x \Rightarrow (\lambda I - A)x = 0 \Rightarrow \lambda I - A\text{不可逆}$$

$\Rightarrow |\lambda I - A| = 0$, λ_i为多项式方程的根.

返回

λ_i称为A的特征值或特征根(物理等学科多称本征值).

复系数多项式$|\lambda I-A|$叫A的特征多项式.

$(\lambda I-A)x=0$的解集称为矩阵A的特征值λ的特征子空间,

记为V_λ.　　$\dim V_\lambda = n-\operatorname{rank}(\lambda I-A)=m_i$　λ的几何重数.

设$\lambda(A)=\{\lambda_1, \lambda_2,\cdots\lambda_r\}$且

$$|\lambda I-A|=\prod_{i=1}^{r}(\lambda-\lambda_i)^{n_i},\quad \sum_{i=1}^{r} n_i = n,$$

正整数n_i称为特征值λ_i的代数重数.　　$1\le m_i \le n_i \le n$.

◀ ▶ 返回

定理 (特征值的性质)

(1)矩阵A的行列式等于其所有特征值的积, i.e. $|A|=\prod_{i=1}^{r}(\lambda_i)^{n_i}$;

(2)矩阵A可逆 $\Leftrightarrow$ 0不是A的特征值;

(3)设$f(x)$是任意多项式,λ是A的一个特征值,α是属于λ的特征向量,则$f(\lambda)$是$f(A)$的一个特征值,α是属于$f(\lambda)$的特征向量;

(4) A^{-1}的特征多项式为$|\lambda I-A^{-1}|=\prod_{i=1}^{r}(\lambda-\lambda_i^{-1})^{n_i}$, 若$\alpha$是$A$的属于特征值$\lambda$的特征向量, 则$\alpha$是$A^{-1}$的属于特征值$\lambda^{-1}$的特征向量.

返回

定理 (特征向量的性质)

(1)属于不同特征值的特征向量线性无关;

(2)n阶矩阵A可以对角化 $\Leftrightarrow$ A有n个线性无关的特征向量;

$\Leftrightarrow C^n$存在由A的特征向量构成的一组基 $\Leftrightarrow m_i = n_i$.

当$m_i < n_i$时,矩阵A不能对角化,但可以块对角化.

设n阶方阵A有r个不同的特征值$\lambda_1, \lambda_2, \cdots, \lambda_r$,其代数重数是$n_i$, $(1 \le i \le r)$,则存在可逆矩阵P,使得

$$P^{-1}AP = J = diag\{J_1(\lambda_1), J_2(\lambda_2), \cdots, J_s(\lambda_s)\}. \ (r \le s)$$

$$J_i(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & 1 \\ & & \lambda_i \end{bmatrix}$$

J叫做A的$Jordan$标准型.

返回

$$P^{-1}AP = J = diag\{J_1(\lambda_1), J_2(\lambda_2), \cdots, J_s(\lambda_s)\}.$$

$$J_i(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} (i = 1, 2, \cdots, s)$$

P的计算方法.

假如$P^{-1}AP = J = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{bmatrix}$. $P = (x_1, x_2, x_3)$,于是有

返回

$$AP = PJ$$

$$i.e. A(x_1, x_2, x_3) = (x_1, x_2, x_3)\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{bmatrix}$$

$$(Ax_1, Ax_2, Ax_3) = (\lambda_1 x_1, \lambda_2 x_2, x_2 + \lambda_2 x_3)$$

$$\begin{cases} (\lambda_1 I - A)x_1 = 0 \\ (\lambda_2 I - A)x_2 = 0 \\ (\lambda_2 I - A)x_3 = -x_2 \end{cases}$$

x_1和x_2分别是λ_1和λ_2的特征向量，

x_3叫λ_2的广义特征向量.

一般情况下λ_i是A的k重特征值,则$x_1, x_2, \cdots, x_k$可由解下面方程组

$$\begin{cases} (\lambda_1 I - A)x_1 = 0 \\ (\lambda_1 I - A)x_i = -x_{i-1}(i = 2,3,\cdots,k) \end{cases}$$ 获得,

$x_1, x_2, \cdots, x_k$线性无关,故得$P = (x_1, x_2, \cdots, x_k)$,

$x_2, \cdots, x_k$为A的属于λ_1的广义特征向量.

广义的特征值和特征向量

设A、$B \subset C^{n\times n}$，如果存在$\lambda \in C$和非零向量$x \in C^{n}$，使得

$$Ax = \lambda Bx$$

则称λ为矩阵A与B确定的广义特征值，x称为与λ对应的广义特征向量.

(1)若$B = E$,则 $Ax = \lambda x.$

(2)若B可逆,则 $B^{-1}Ax = \lambda x.$

(3)若A、B都是Hermite矩阵， 即$A=A^H$、$B=B^H$ 且B正定时,有

$$B=B^H\text{且正定}\xrightarrow{\text{存在可逆矩阵}P} B=P^HP$$

$$\text{则}Ax=\lambda P^HPx \xrightarrow{\boxed{\text{记 } y=Px,\text{则}P^{-1}y=x}} (P^{-1})^HAP^{-1}y=\lambda y \xrightarrow{\boxed{Q=(P^{-1})^HAP^{-1}}}$$

$$Qy=\lambda y \xrightarrow{\boxed{Q^H=Q}} \text{广义特征值}\lambda_1,\cdots\lambda_n\text{都是实数} \xrightarrow{\boxed{\text{存在标准正交基}y_1,\cdots,y_n}}$$

$$y_i^Hy_j=\delta_{ij} \xrightarrow{y_i=Px_i} y_i^Hy_j=(Px_i)^H(Px_j)=x_i^HP^HPx_j=x_i^HBx_j$$

$$\Rightarrow x_i^HBx_j=\delta_{ij}$$

1.若$A=A^H$,则A的特征值均为实数,且属于不同特征值得特征向量彼此正交.

2.实对称矩阵可以正交对角化；Hermite矩阵可以酉对角化.

定义 2 设 A，B 为 n 阶方阵，如果存在可逆矩阵 C，使得

$$B=C^TAC,$$

则称 A 与 B 合同.

当$x_i^H B x_j = \delta_{ij}$，称 x_1，$x_2,\cdots,x_n$为 B 共轭向量系

定理　设$n\times n$矩阵 $A = A^H$，$B = B^H$，且B正定,与B共轭向量系x_1，$x_2,\cdots,x_n$具有以下性质，

（1）$x_i \neq 0 \ (i = 1, 2,\cdots,n)$；

（2）x_1，$x_2,\cdots,x_n$ 线性无关；

（3）λ_i与x_i满足方程$Ax_i = \lambda_i B x_i$；

（4）若令$X = (x_1, x_2,\cdots,x_n)$，

$$X^H B X = E,\quad X^H A X = diag(\lambda_1, \lambda_2,\cdots,\lambda_n)$$

6内积空间

定义1 在线性空间$V_n(P)$上,若映射(x, y)

$$V_n(P)\times V_n(P)\to P$$

(1) (正定性) $(x, x)\geq 0$; $(x, x)=0 \Leftrightarrow x=0$;

(2) (共轭对称性) $(x, y)=\overline{(y, x)}$;

(3) (双线性) $(ax+by, z)=\bar{a}(x, z)+\bar{b}(y, z)$,

$\forall x, y, z\in V_n(P), a, b\in P$成立.称$V$为一个内积空间.

$(x, ay+bz)=a(x, y)+b(x, z)$.(共轭双线性)

$P=R$,则内积为对称性(可交换性).

有限维的实内积空间又叫欧氏空间.

$P=C$，则内积空间称为酉空间(复内积空间).

定义2 对$\forall x\in V_n(P)$，称$\|x\|=\sqrt{(x,x)}$

为向量x的长度(或模、范数)

长度为1的向量叫单位向量或标准向量.

定理1：设(x,y)是$V_n(P)$上的内积，则

(1) $\|\lambda x\|=|\lambda|\cdot\|x\|$

(2) $\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2$

(3) $|(x,y)|\le\|x\|\cdot\|y\|$ (Cauchy不等式)

(4) $\|x+y\|\le\|x\|+\|y\|$ (三角不等式)

向量x和y的夹角

$$|(x,y)|\le\|x\|\cdot\|y\| \Rightarrow \theta = arccos\frac{|(x,y)|}{\|x\|\cdot\|y\|},\quad 0\le\theta\le\frac{\pi}{2}.$$

若 $P=R\Rightarrow \theta = arccos\dfrac{(x,y)}{\|x\|\cdot\|y\|},\qquad 0\le\theta\le\pi.$

欧氏空间通常内积的一般定义$(x, y)=y^T x = x^T y = (y, x)$,

酉空间内积的定义$(x, y)=x^H y = \overline{(y, x)} = \overline{y^H x}.$

R^2中有 $\qquad (x, y) = \|x\|\|y\| cos\theta, \theta = \angle(x, y).$

$(x-y, x-y) = \|x\|^2 - 2(x, y) + \|y\|^2.$

$c^2 = a^2 - 2ab\cos\theta + b^2.$(余弦定理)

$cos\theta = 0$,即$\theta=\dfrac{\pi}{2}$时,

即得勾股定理.

返回

定义 3 若$\forall x,y\neq 0,(x,y)=0$,则称$x$与$y$正交,记为$x\perp y$.

一组非零向量如果两两正交则称正交组.

单位向量构成的正交组称为标准正交组.

定理　正交组必是线性无关的.

定理　内积空间必存在标准正交基.

设$x_1,x_2,\cdots,x_n$是内积空间的一组标准正交基,则矩阵

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

若内积空间为欧氏空间,Q叫正交矩阵;

若内积空间为酉空间,Q叫酉矩阵.

矩阵Q是正交矩阵 $\Leftrightarrow Q^TQ=I \Leftrightarrow Q^{-1}=Q^T$.

矩阵Q是酉矩阵 $\Leftrightarrow Q^HQ=I \Leftrightarrow Q^{-1}=Q^H$.

度量矩阵(Gram矩阵)

设$\alpha_1,\alpha_2,\cdots,\alpha_n$是$n$维内积空间$V$的基,对于$V$的任意两个向量

$$x=k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n,\quad y=l_1\alpha_1+l_2\alpha_2+\cdots+l_n\alpha_n,$$

由双线性和共轭双线性知

$$(x,y)=\sum_{i,j=1}^{n}\overline{k_i}l_j(\alpha_i,\alpha_j)=\sum_{i,j=1}^{n}a_{ji}\overline{k_i}l_j.$$

其中$a_{ji}=(\alpha_i,\alpha_j)(i,j=1,2,\cdots,n)$,用矩阵表示就有

返回

其中$a_{ji}=(\alpha_i,\alpha_j)(i,j=1,2,\cdots,n)$,用矩阵表示就有

$$(x,y)=(\overline{k_1},\overline{k_2},\cdots,\overline{k_n})A\begin{pmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{pmatrix} \qquad (1)$$

$$\text{其中}A=(a_{ij})=\begin{bmatrix} (\alpha_1,\alpha_1) & (\alpha_2,\alpha_1) & \cdots & (\alpha_n,\alpha_1) \\ (\alpha_1,\alpha_2) & (\alpha_2,\alpha_2) & \cdots & (\alpha_n,\alpha_2) \\ \cdots & \cdots & \cdots & \cdots \\ (\alpha_1,\alpha_n) & (\alpha_2,\alpha_n) & \cdots & (\alpha_n,\alpha_n) \end{bmatrix}$$

称A为V对于基 $\alpha_1,\alpha_2,\cdots,\alpha_n$的度量矩阵(Gram矩阵).

返回

基向量的内积 $\Rightarrow$ 矩阵A $\Rightarrow$ 任意两个向量的内积.

$\because x \neq 0$时, $(x, x)>0$,

$\therefore$ 欧氏空间下, 度量矩阵A是对称正定矩阵.

酉空间下, 度量矩阵A是正定*Hermite*矩阵.

度量矩阵A完全确定了内积, 故可以用任意正定矩阵或正定*Hermite*矩阵作为度量矩阵来规定内积.

标准正交基对应的度量矩阵A?

返回

例5 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$与$\beta_1,\beta_2,\cdots,\beta_n$分别是欧氏空间$V$的两个基,且$V$对于两基的度量矩阵分别是 $A=(a_{ij})$与$B=(b_{ij})$,求证A与B合同.

证明：设$(\beta_1,\beta_2,\cdots\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)P$, $P=(p_{ij})$

$$\beta_i = p_{1i}\alpha_1 + p_{2i}\alpha_2+\cdots+ p_{ni}\alpha_n \quad (i=1,2,\cdots,n)$$

$$(\beta_i,\beta_j)=\sum_{s=1}^{n}\sum_{t=1}^{n} p_{si}p_{tj}(\alpha_s,\alpha_t)=(p_{1i},p_{2i},\cdots p_{ni})A\begin{bmatrix} p_{1j} \\ \vdots \\ p_{nj} \end{bmatrix},$$

$$(i=1,2,\cdots,n)$$

即 $B=P^T AP.$

度量矩阵行列式(Gram行列式)的几何意义

例 设$\alpha_1,\alpha_2\in R^2$,考察行列式

$$D=\begin{vmatrix}(\alpha_1,\alpha_1) & (\alpha_2,\alpha_1)\\(\alpha_1,\alpha_2) & (\alpha_2,\alpha_2)\end{vmatrix}$$的几何意义.

解：设α_1与α_2的夹角为θ, $\because(\alpha_1,\alpha_2)=(\alpha_2,\alpha_1)$

$$\therefore D=\|\alpha_1\|^2\|\alpha_2\|^2-(\alpha_1,\alpha_2)^2$$

$$=\|\alpha_1\|^2\|\alpha_2\|^2-\|\alpha_1\|^2\|\alpha_2\|^2\cos^2\theta$$

$$=\|\alpha_1\|^2\|\alpha_2\|^2\sin^2\theta$$

以α_1,α_2为邻边的平行四边形面积的平方.

7 内积空间的正交分解

问题：设V是n维内积空间,U是V的子空间.令

$$W=\{\alpha\in V:(\alpha,\beta)=0,\forall\beta\in U\}.$$

W是V的一个子空间吗？

$\forall\alpha,\gamma\in W\quad(\alpha+\gamma,\beta)=(\alpha,\beta)+(\gamma,\beta)=0\ \Rightarrow\alpha+\gamma\in W.$

$(k\alpha,\beta)=k(\alpha,\beta)=0\ \Rightarrow k\alpha\in W.$

故W是V的一个子空间.

定义1 设V是n维内积空间,U是V的子空间.令

$W=\{\alpha\in V:(\alpha,\beta)=0,\forall\beta\in U\}.$

则称W是U的正交补子空间. 记为$W=U^{\perp}.$

返回

定理1 设V是n维内积空间, U是V的子空间, 则$V=U\oplus U^{\perp}$.

证明 若$U=\{0\}$, 则$U^{\perp}=V$, $\Rightarrow V=\{0\}\oplus V=U\oplus U^{\perp}$.

若$U\neq\{0\}$, 设$\dim U=m(1\leq m\leq n)$, 且U的一个标准正交基为$\alpha_1,\alpha_2,\cdots\alpha_m$, 下证$V=U+U^{\perp}$.

$\forall x\in V$, 令$k_i=(x,\alpha_i)(i=1,2,\cdots,m)$.则

$$y=k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m\in U$$

再令$z=x-y$, 由于

$(z,\alpha_i)=(x-y,\alpha_i)=(x,\alpha_i)-(y,\alpha_i)=k_i-k_i=0.\Rightarrow z\in U^{\perp}$.

$\Rightarrow x=y+z, y\in U, z\in U^{\perp}$.

$\Rightarrow x = y + z, y \in U, z \in U^{\perp}$. 即$V = U + U^{\perp}$,

又$\because U \cap U^{\perp} = \{0\}$, $\therefore V = U \oplus U^{\perp}$,

推论 设V是n维内积空间, U是V的子空间, 且$\dim U = m$, 则$\dim U^{\perp} = n - m$.

$(U^{\perp})^{\perp} = U$.

正交补空间的应用： 其齐次线性方程组 $Ax = 0$.

设 $A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix}$, $Ax = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} x = \begin{pmatrix} (\alpha_1, x) \\ (\alpha_2, x) \\ \vdots \\ (\alpha_m, x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

$$Ax=\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_m\end{pmatrix}x=\begin{pmatrix}(\alpha_1,x)\\ (\alpha_2,x)\\ \vdots\\ (\alpha_m,x)\end{pmatrix}=\begin{pmatrix}0\\ 0\\ \vdots\\ 0\end{pmatrix}.$$

即 $(\alpha_1,x)=0,$
$(\alpha_2,x)=0,$
$\cdots,$
$(\alpha_m,x)=0.$

方程组的解向量就是所有与向量组 $\alpha_1,\alpha_2,\cdots,\alpha_m$正交的向量. 设$U=L(\alpha_1,\alpha_2,\cdots,\alpha_m)$

$\therefore Ax=0$的解空间就是求U的正交补空间.

最佳近似定理、广义相对论等也有重要的应用.

EX: 试利用正交分解的理论在空间中建立关于面积的勾股定理. 能否建立更高维的勾股定理?

$N(A^H)$与$R(A)$互补, $N(A)$与$R(A^H)$互补, 是正交补吗?

设$A=(a_{ij})\in C^{m\times n}$的第$j$个列向量为$\alpha_j$, $(j=1,2,\cdots,n)$,

并记$U=R(A)=L(\alpha_1,\alpha_2,\cdots\alpha_n)\subset C^m$, 于是有

$$U^{\perp}=\{y \mid y\perp(k_1\alpha_1+k_2\alpha_2+\cdots+k_n\alpha_n, k_j\in C)\}$$

$$=\{y \mid y\perp\alpha_j, j=1,2,\cdots,n\}$$

$$=\{y \mid \alpha_j^H y=0, j=1,2,\cdots,n\}$$

$$=\{y \mid A^H y=0\}=N(A^H)$$

由定理1知 $C^m=U\oplus U^{\perp}=R(A)\oplus N(A^H)$.

定理2 $R^{\perp}(A)=N(A^H)$, $R(A)\oplus N(A^H)=C^m$;

$$R^{\perp}(A^H)=N(A), R(A^H)\oplus N(A)=C^n.$$

正交投影

定义2 设T是V上的投影, $V = R(T) \oplus N(T)$,

若$R(T) \perp N(T)$, 则称T为正交投影.

正交投影算子或正交投影矩阵的性质

1 设投影T的矩阵是A,则T为正交投影 $\Leftrightarrow A^H = A$.

$\Leftarrow: \forall x \in N(A), y \in R(A)$,则$(x, y) = (x, Ay)$

$= (A^H x, y) = (Ax, y) = (0, y) = 0 \quad \Rightarrow R(A) \perp N(A)$.

$\Rightarrow: \forall x \in V, y \in N(A)$.

$\Rightarrow 0 = (y, Ax) = (A^H y, x) \Rightarrow A^H y = 0 \Rightarrow y \in N(A^H)$.

$\Rightarrow A^H = A$.(正交补的唯一性)

2　A为正交投影算子 $\Leftrightarrow I-A$也是正交投影算子.

3　若T_1,T_2均为正交投影算子,则T_1+T_2正交投影算子

$\Leftrightarrow T_1T_2=T_2T_1=0.$

证: $\Leftarrow \because T_1=T_1^2, T_2=T_2^2,$

$\therefore (T_1+T_2)^2=T_1^2+T_1T_2+T_2T_1+T_2^2=T_1^2+T_2^2=T_1+T_2.$

又 $\because T_1^H=T_1, T_2^H=T_2,$　　$\therefore (T_1+T_2)^H=T_1+T_2$

$\Rightarrow T_1+T_2$为正交投影算子.

$\Rightarrow \because (T_1+T_2)^2=T_1^2+T_1T_2+T_2T_1+T_2^2=T_1+T_2,$

$\Rightarrow T_1T_2+T_2T_1=0.$

$T_1T_2+T_2T_1=0.$

$\Rightarrow T_1T_1T_2+T_1T_2T_1=0\Rightarrow T_1T_2+T_1T_2T_1=0,$(左)

$\Rightarrow T_1T_2T_1+T_2T_1T_1=0\Rightarrow T_1T_2T_1+T_2T_1=0.$(右)

$\Rightarrow T_1T_2-T_2T_1=0 \quad \Rightarrow T_1T_2=T_2T_1=0.$

4 若T_1,T_2均为正交投影算子,则T_1-T_2正交投影算子

$\Leftrightarrow T_1T_2=T_2T_1=T_2.$

证明：T_1-T_2正交投影算子$\Leftrightarrow I-(T_1-T_2)$为正交投影算子.

$\because I-(T_1-T_2)=(I-T_1)+T_2,$

$\therefore T_1-T_2$正交投影算子$\Leftrightarrow (I-T_1)+T_2$为正交投影算子.

$\Leftrightarrow (I-T_1)T_2=T_2(I-T_1)=0,$

$\Leftrightarrow T_2-T_1T_2=0, T_2-T_2T_1=0, \Leftrightarrow T_1T_2=T_2T_1=T_2.$

8 内积空间中的线性变换

问题:在内积空间,线性变换将三角形变成什么图形?是否仍将抛物线变为抛物线?

本质问题:在内积空间中,线性变换是否保持长度、角度和距离?

一、保距变换

定义1 设V是内积空间,T是线性变换且保持向量间距离,即$\forall \alpha, \beta \in V$,均有$d(T(\alpha),T(\beta))=d(\alpha, \beta)$,则称$T$是等距变换或保距变换.

返回

定理1　设V是内积空间, T是线性变换, 则T是等距变换.

$\Leftrightarrow T$保持向量的长度$\Leftrightarrow T$保持内积.

证：$\forall \alpha \in V$, 由于$\|\alpha\| = \sqrt{(\alpha,\alpha)} = d(\alpha, 0)$, 所以保持距离等价于保持长度. 保持内积显然蕴含了保持长度.

设T保持长度,

$$\because (T(\alpha+\beta), T(\alpha+\beta))$$

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

$$=(\alpha,\alpha)+(T(\alpha),T(\beta))+\overline{(T(\alpha),T(\beta))}+(\beta,\beta)$$

$$(\alpha+\beta, \alpha+\beta)=(\alpha, \alpha)+(\alpha, \beta)+\overline{(\alpha, \beta)}+(\beta, \beta)$$

$$\therefore (T(\alpha),T(\beta))+\overline{(T(\alpha),T(\beta))} = (\alpha,\beta)+\overline{(\alpha,\beta)}$$

返回

$$(T(\alpha),T(\beta))+\overline{(T(\alpha),T(\beta))}=(\alpha,\beta)+\overline{(\alpha,\beta)}$$

$\therefore (T(\alpha),T(\beta))$与$(\alpha,\beta)$有相同的实部.

又$\because Re(i\alpha,\beta))=Im(\alpha,\beta)$有相同的虚部.

$\therefore (T(\alpha),T(\beta))$与$(\alpha,\beta)$有相同的虚部.

$\therefore (T(\alpha),T(\beta))=(\alpha,\beta)$.

定理2 设V是内积空间,T是线性变换,$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的一组标准正交基, A是T在该组基下的矩阵,则T是等距变换$\Leftrightarrow$ A是酉矩阵.

定理2 设V是内积空间, T是线性变换,$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的一组标准正交基, A是T在该组基下的矩阵,则T是等距变换 $\Leftrightarrow$ A是酉矩阵.

证: $\Leftarrow$ $\forall \alpha \in V$,设$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)x,\ x\in C^n$,则

$$\|\alpha\|=\sqrt{(\alpha,\alpha)}=\sqrt{x^H x}.$$

设$T(\alpha)=\beta=(\alpha_1,\alpha_2,\cdots,\alpha_n)Ax$,则

$$\|\beta\|=\sqrt{(Ax)^H(Ax)}=\sqrt{x^H A^H Ax}.$$

$\because A^H A=I$, 则$\|T(\alpha)\|=\|\beta\|=\sqrt{x^H x}=\|\alpha\|$.

$\therefore T$保持长度,是等距变换.

返回

$$\Rightarrow (\beta,\beta)=x^H A^H Ax=(\alpha,\alpha)=x^H x.$$

$$\therefore x^H x - x^H A^H Ax = x^H(I-A^H A)x = 0. \therefore A^H A = I.$$

例：设A是$Hermite$矩阵,如果对任意向量x均有$x^H Ax=0$,则$A=0$.

证　$\because A^H = A, \therefore$ 存在酉矩阵C,使得$A = Cdiag\{\lambda_1,\lambda_2,\cdots\lambda_n\}C^H$

令$C=(c_1,c_2,\cdots c_n)$,　由$C^H C=E$知$c_i^H c_j = 1$,当$i=j$;

$c_i^H c_j = 0$,当$i \neq j$.　$\therefore c_i^H C = (0,\cdots 0,1,0,\cdots 0)=e_i^H$.

取$x=c_i$,　$0 = x^H Ax = c_i^H Ac_i$

$$=c_i^H Cdiag\{\lambda_1,\lambda_2,\cdots\lambda_n\}C^H c_i$$

$$=e_i^H diag\{\lambda_1,\lambda_2,\cdots\lambda_n\}e_i=\lambda_i \quad (\forall i \in N).$$

$$A = Cdiag\{\lambda_1,\lambda_2,\cdots\lambda_n\}C^H = Cdiag\{0,0,\cdots,0\}C^H = 0.$$

定理2 设V是内积空间, T是线性变换,$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的一组标准正交基, A是T在该组基下的矩阵,则T是等距变换 $\Leftrightarrow$ A是酉矩阵.

思考：定理2中的条件“标准正交”可以去掉吗？

例 设T为实平面上逆时针旋转$\dfrac{\pi}{2}$,求它在基(e_1, e_1+e_2)下的矩阵.

解 因为T为实平面上的旋转变换,故

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases} \quad \text{即} T \to A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases} \text{即} T \to \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

当$\theta=\dfrac{\pi}{2}$时, $A=\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

基(e_1,e_2)到基(e_1,e_1+e_2)的过渡矩阵 $P=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

所以T在基(e_1,e_1+e_2)下的矩阵

$B=P^{-1}AP=\begin{bmatrix} -1 & -2 \\ 1 & 1 \end{bmatrix}$ B显然不是酉矩阵!

所以,定理2中的条件“标准正交”不可以去掉!!

返回

注：欧氏空间中的等距变换又叫正交变换；
复内积空间中的等距变换又叫酉变换.
等距变换的积和逆仍是等距变换.

定义2 设T是欧氏空间V的线性变换,若

$$(T(\alpha),T(\alpha))=(\alpha,\alpha),$$ 则称T是正交变换.

定理3 线性变换T是正交变换 $\Leftrightarrow \forall x, y\in V$, 均有

$$(T(x),T(y))=(x,y).$$

证: $\Leftarrow$ 令$y=x$,充分性显然成立.

$\Rightarrow$ 设T是正交变换,则

$$(x-y,x-y)=(T(x-y),T(x-y)).$$

返回

$$(x-y,x-y)=(T(x-y),T(x-y)).$$

$$(x,x)-2(x,y)+(y,y)=(Tx,Tx)-2(Tx,Ty)+(Ty,Ty)$$
$$=(x,x)-2(Tx,Ty)+(y,y)$$

即 $(x,y)=(Tx,Ty)$.

注 若空间为复空间,则线性变换T也称酉变换.

保距变换保持长度和内积,从而保持了向量的角度,因此保距变换不改变图形的形状.

二、两个著名的正交变换

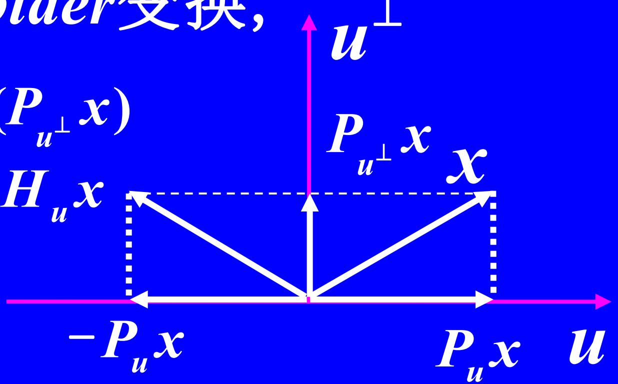
*Householder*变换 & *Givens*变换

$a)$Householder变换

设$u \in C^n$,且$u^H u = 1$,定义n阶复阵H : $H = E_n - 2uu^H$ 称为初等酉矩阵或Householder矩阵.

由矩阵H定义的线性变换H_u称为Householder变换,

即对$\forall x \in C^n$, $H_u x = x - 2uu^H x = H_u(P_u x) + H_u(P_{u^\perp} x)$



$$x = P_u x + P_{u^\perp} x$$

$$\begin{aligned} H_u(P_u x) &= (E_n - 2uu^H)(P_u x) \\ &= P_u x - 2uu^H (P_u x) \\ &= P_u x - 2uu^H (\lambda u) \\ &= P_u x - 2\lambda u(u^H u) \\ &= P_u x - 2\lambda u \\ &= P_u x - 2P_u x \\ &= -P_u x. \end{aligned}$$

$$\begin{aligned} H_u(P_{u^\perp} x) &= (E_n - 2uu^H) P_{u^\perp} x \\ &= P_{u^\perp} x - 2uu^H (P_{u^\perp} x) \\ &= P_{u^\perp} x - 2bu(u^H u^\perp) \\ &= P_{u^\perp} x - 2bu(u^\perp, u) = P_{u^\perp} x. \end{aligned}$$

H_u实际上是关于超平面$u^\perp$的镜面反射.

返回

b) *Givens*变换

正交矩阵$A=\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$决定的旋转称为*Jacobi*旋转或(二维)*Givens*旋转，

*n*阶实矩阵

$$G(i, j, \theta)=E_n-(1-\cos\theta)(E_{ii}+E_{jj})+\sin\theta(E_{ij}-E_{ji})$$

称为*Givens*旋转矩阵.

$$
G(i,j,\theta)=\begin{array}{c@{}c}
\begin{bmatrix}
1 & \cdots & 0 & & 0 & & 0 \\
\vdots & \ddots & & & & & \\
0 & & \cos\theta & \cdots & \sin\theta & & \\
\vdots & & & \ddots & & & \\
0 & & -\sin\theta & \cdots & \cos\theta & & \\
\vdots & & & & & \ddots & \\
0 & \cdots & 0 & & 0 & & 1
\end{bmatrix} &
\begin{array}{c}
\\ \\ i \\ \\ j \\ \\ \\
\end{array} \\
\begin{array}{ccccccc}
 & & i & & j & &
\end{array} &
\end{array}
$$

Householder,*Givens*变换均可以用来将矩阵中某些特定的元素变为0,是矩阵快速计算中不可或缺的方法.

三、对称变换与对称矩阵

定义3　设T是欧氏空间V的线性变换, 如果对$\forall \alpha, \beta \in V$, 均有

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

则称T是V中的一个对称变换.

若V是酉空间, 线性变换T称为V的*Hermite*变换.

*Householder*变换是?

定理4　欧式空间的线性变换T是对称变换$\Leftrightarrow$ T在一组标准正交基下的矩阵是实对称矩阵.

返回

定理4 欧式空间的线性变换T是对称变换$\Leftrightarrow$ T在一组标准正交基下的矩阵是实对称矩阵.

证明:$\Leftarrow$ 设T在一组标准正交基$\alpha_1, \alpha_2, \cdots \alpha_n$下的矩阵为$A$是实对称矩阵,设$\alpha, \beta$在该组下的坐标分别为$x, y$,

则$T(\alpha)=(\alpha_1, \alpha_2, \cdots \alpha_n)Ax$, $T(\beta)=(\alpha_1, \alpha_2, \cdots \alpha_n)Ay$,

$$\therefore (T(\alpha), \beta)=(Ax, y)=y^T Ax=(A^T y)^T x=(Ay)^T x$$

$$=(x, Ay)=(\alpha, T(\beta)).$$

$\Rightarrow$ 设T是对称变换,在标准正交基$\alpha_1, \alpha_2, \cdots \alpha_n$下的矩阵

$A=(a_{ij})$, 则$T(\alpha_j)=\sum_{i=1}^{n} a_{ij}\alpha_i$,

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

$$\begin{aligned} T\alpha_1 &= a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n \\ T\alpha_2 &= a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n \\ &\cdots\cdots\cdots\cdots \\ T\alpha_n &= a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n \end{aligned}$$

$A=(a_{ij})$,则 $T(\alpha_j)=\sum_{i=1}^{n} a_{ij}\alpha_i$,

$$a_{ij}=(T(\alpha_j),\alpha_i)=(\alpha_i,T(\alpha_j))=(T(\alpha_i),\alpha_j)=a_{ji}.$$

内积的对称性

返回