第5章 矩阵的相合与相似

5.1 欧氏空间

5.2 正交化

5.3 二次型

5.4 实对称方阵相合标准形

5.5 特征向量与相似矩阵

5.6 正交相似

5.7 更多例子

5.8 若当标准型

5.1 欧氏空间

最小二乘法

例1 在几何向量组成的3维向量空间v中，用几何方法定义了内积：

$$(\alpha,\beta)=|\alpha||\beta|\cos\theta$$

求是否存在向量 $\alpha_1,\alpha_2,\alpha_3$ 满足条件：

$$(\alpha_1,\alpha_1)=(\alpha_2,\alpha_2)=(\alpha_3,\alpha_3)=1,(\alpha_1,\alpha_2)=2,$$

$$(\alpha_1,\alpha_3)=-3\text{且}(\alpha_2,\alpha_3)=-2?$$

解： 设有 $\alpha_1,\alpha_2,\alpha_3$ 满足条件，它们在实数域R上的**任意线性组合** $\alpha = x\alpha_1 + y\alpha_2 + z\alpha_3$ 含于V，应满足条件（α,α）≥ 0. 即

$$(\alpha,\alpha) = (x\alpha_1 + y\alpha_2 + z\alpha_3, x\alpha_1 + y\alpha_2 + z\alpha_3)$$

$$= x^2 + y^2 + z^2 + 4xy - 6xz - 4yz \geq 0$$

配方得

$$(\alpha,\alpha) = x^2 + y^2 + z^2 + 4xy - 6xz - 4yz \geq 0$$

$$= (x + 2y - 3z)^2 - 3(y - \frac{4}{3}z)^2 - \frac{8}{3}z^2$$

选 $x = -2, y = 1, z = 0,$ 则（α，α）$=-3<0$

矛盾。故不存在满足所说条件的 $\alpha_1,\alpha_2,\alpha_3$.

例2（最小二乘法） 已知某种材料在生产过程中的废品率y与某种化学成分的含量百分比x有关。以下是某工厂在生产过程中实测到的*x*, *y*的一些对应数据。

X(%)	3.6	3.7	3.8	3.9	4.0	4.1	4.2
Y(%)	1.00	0.9	0.9	0.81	0.60	0.56	0.35

试给出*x*, *y*之间函数关系的一个近似公式。

解: 可以将x，y 的函数关系用一次函数$y=kx+b$来近似的表示。现在需要确定k，b使直线与所有点的总的偏差尽可能小，点(x_i, y_i)的偏差为kx_i+b-y_i，则所有偏差的平方和为：

$$d(k,b)=\sum_{i=1}^{7}(kx_i+b-y_i)^2$$

用来度量"总的偏差"，选择k, b使$d(k,b)$达到最小值。将$d(k, b)$展开得

$$d(k,b)=Ak^2+2Bkb+Cb^2-2Dk-2Eb+F$$

其中 $A=\sum_{i=1}^{7}x_i^2=106.75, B=\sum_{i=1}^{7}x_i=27.3, C=\sum_{i=1}^{7}1=7$

$$D=\sum_{i=1}^{7}x_iy_i=19.675, E=\sum_{i=1}^{7}y_i=5.12, F=\sum_{i=1}^{7}y_i^2=4.0722$$

将数值代入，$d(k,b)$ 配方得:

$$d(k,b)=106.75(k+0.255738b-0.184309)^2+$$

$$0.0183607(b-4.8125)^2+0.0206821$$

取 $b=4.8125, k=-0.255738b+0.184309=-1.04643$

即可使 $d(k,b)$ 达到最小值0.0206821。

所以所求直线方程为 $y=-1.04643x+4.8125.$

内积的推广

实数域R上的三维几何空间R^3中，向量α,β的内积可以通过模长和夹角余弦来定义

$$(\alpha,\beta)=|\alpha||\beta|\cos\theta,$$

设向量α,β的坐标为$\alpha=(x_1,y_1,z_1),\beta=(x_2,y_2,z_2)$,则有内积的坐标表达式和余弦公式：

$$(\alpha,\beta)=x_1x_2+y_1y_2+z_1z_2,$$

$$\cos\theta=\frac{(\alpha,\beta)}{|\alpha||\beta|}=\frac{x_1x_2+y_1y_2+z_1z_2}{\sqrt{x_1^2+y_1^2+z_1^2}\sqrt{x_2^2+y_2^2+z_2^2}}$$

实数域R上的n维空间R^n，设向量α,β为$\alpha=(x_1,\dots,x_n),\beta=(y_1,\dots,y_n)$，很自然地推广3维内积的坐标表达式和余弦公式：

$$(\alpha,\beta)=\sum_{i=1}^{n}x_iy_i=x_1y_1+x_2y_2+\cdots+x_ny_n$$

$$|\alpha|=\sqrt{(\alpha,\alpha)}=\sqrt{x_1^2+\cdots+x_n^2}$$

$$\cos\theta=\frac{(\alpha,\beta)}{\sqrt{(\alpha,\alpha)}\sqrt{(\beta,\beta)}}=\frac{\sum_{i=1}^{n}x_iy_i}{\sqrt{\sum_{i=1}^{n}x_i^2}\sqrt{\sum_{i=1}^{n}y_i^2}}$$

定义5.1.1 (内积)设 V是实数域R上线性空间. 如果给定了V上的2元实函数，将V中任意两个向量α, β对应到一个实数(α, β), 并且满足如下条件:

(1) (双线性)

$$(\alpha_1+\alpha_2,\beta)=(\alpha_1,\beta)+(\alpha_2,\beta),(\lambda\alpha_1,\beta)=\lambda(\alpha_1,\beta);$$

$$(\beta,\alpha_1+\alpha_2)=(\beta,\alpha_1)+(\beta,\alpha_2),(\beta,\lambda\alpha_1)=\lambda(\beta,\alpha_1)$$

对任意 $\alpha_1,\alpha_2, \beta \in V$和$\lambda\in R$成立。

(2) (对称性)$(\alpha,\beta)=(\beta,\alpha)$对任意$\alpha, \beta \in V$成立

(3) (正定性)$(\alpha, \alpha,)>0$对任意$0 \neq\alpha\in V$成立

就称(α, β)为内积，V为欧几里德空间，简称欧氏空间。

记号E(R)表示欧氏空间，$E_n(R)$表示n维欧氏空间。

例 如下的空间V是欧氏空间:

1.V是实数域R上n维数组空间R^n。对$X=(x_1,\ldots,x_n)$, $Y=(y_1,\ldots,y_n)$，定义$(X,Y)=x_1y_1+\ldots+x_ny_n$，则$(X,Y)$是内积，$R^n$成为欧氏空间。此内积称为$R^n$上的标准内积。

2.V是闭区间$[a,b]$上所有的连续实函数组成的实向量空间$C[a,b]$。对任意$f(x)$，$g(x)\in V$，定义

$$(f(x),g(x))=\int_a^b f(x)g(x)dx$$

则$(f(x),g(x))$是内积, $C[a,b]$在此内积下是欧氏空间。

由内积的定义立即得出如下性质:

- $(0,\alpha)=(00,\alpha)=0(0,\alpha)=0$。特别$(0,0)=0$。因此, 内积的正定性也可以叙述为:$(\alpha, \alpha) \geq 0$对任意$\alpha \in V$成立， 其中等号成立仅当$\alpha=0$。

- 由正定性定义任意向量α的长度 $|\alpha|=\sqrt{(\alpha,\alpha)}$,则$|\alpha| \geq 0$, 且仅当$\alpha=0$时$|\alpha|=0$。

- 对任意$\boldsymbol{\alpha} \neq 0$, $\boldsymbol{\alpha}_0=(1/|\boldsymbol{\alpha}|)\boldsymbol{\alpha}$是与$\boldsymbol{\alpha}$方向相同、长度为1的向量，称为单位向量。由$\boldsymbol{\alpha}$得到单位向量$(1/|\alpha|)\boldsymbol{\alpha}$ 的过程称为将向量$\boldsymbol{\alpha}$归一化。

- 设V是欧氏空间。如果$\alpha,\beta \in V$满足 $(\alpha, \beta)=0$, 就称α, β正交，记为$\alpha \perp \beta$。

例3 对任意$A\in R^{m\times n}$，$b\in R^{m\times 1}$，求证：

1. $A^TAX=A^Tb$的解X_0使 $|AX-b|^2$取**最小值**$|AX_0-b|^2$。
2. 齐次线性方程组$AX=0$与$A^TAX=0$同解。
   $\text{rank}(A^TA)=\text{rank}A$。
3. 方程组$A^TAX=A^Tb$总有解，且当$\text{rank}A=n$时有**唯一解**$X=(A^TA)^{-1}A^Tb$。

证明：(1)设$A^TAX_0=A^Tb$。对任意$X\in \mathrm{R}^{n\times 1}$, $AX-b=(AX_0-b)+A(X-X_0)$，且

$$(A(X-X_0))^T(AX_0-b)=(X-X_0)^TA^T(AX_0-b)$$

$$=(X-X_0)^T(A^TAX_0-A^Tb)=0$$

即$AX_0-b\perp A(X-X_0)$,由勾股定理得$|AX-b|^2\geq|AX_0-b|^2$。

(2)$AX=0$与$A^TAX=0$ 的解空间为 $V_A\subseteq V_{A^TA}$.

反之， $X\in V_{A^TA}\Longleftrightarrow A^TAX=0\Longrightarrow X^TA^TAX=0\Longrightarrow (AX)^TAX=0\Longrightarrow AX=0$即$X\in V_A$。**从而** $V_{A^TA}\subseteq V_A$.

$$\left.\begin{aligned}\dim V_{A^TA}&=n-\operatorname{rank} A^TA\\ \dim V_A&=n-\operatorname{rank} A\end{aligned}\right\}\Rightarrow \operatorname{rank} A^TA=\operatorname{rank} A.$$

(3) 由rankAB ≤ rank$A$知 rank(A^TA,A^Tb)≤rank(A^T)= rank(A)=rank(A^TA)。反之, A^TA为(A^TA,A^Tb)的子矩阵于是rank(A^TA) ≤ rank(A^TA,A^Tb)，因此方程组$A^TAX=A^Tb$ 的系数矩阵与增广矩阵的秩相等，一定有解。当rank$A$=n时，rank(A^TA)=rank$A$=n。因此A^TA为可逆矩阵，方程组有唯一解$X=(A^TA)^{-1}A^Tb$。

✎**算法5.1.1**(最小二乘法)设n元实系数线性方程组$AX=b$的系数矩阵的秩rank$A$=n。若方程组无解，则$(A^TA)X=A^Tb$的唯一解X_0是使$|AX-b|^2$最小的近似值。

定理5.1.2(Cauchy-Schwatz不等式)对欧氏空间V中任意$\alpha, \beta\in V$,$(\alpha, \beta)^2\le(\alpha, \alpha)(\beta, \beta)$成立, 其中的等号仅当$\alpha, \beta$线性相关时成立。

证明： 显然, 当$\alpha=0$时, $(\alpha, \beta)^2=0=(\alpha, \alpha)(\beta, \beta)$, C-S不等式中的等式成立。

以下设$\alpha\neq0$, $(\alpha, \alpha)>0$由内积的正定性知: 对任意实数x，有$(x\alpha+\beta, x\alpha+\beta)\ge0$即$(\alpha, \alpha)x^2+2(\alpha, \beta)x+(\beta, \beta)\ge0$。

这说明以x为自变量的二次函数$f(x)=(\alpha, \alpha)x^2+2(\alpha, \beta)x+(\beta, \beta)$的图像在$x$ 轴的上方。

判别式$\Delta = 4(\alpha, \beta)^2 - 4(\alpha, \alpha)(\beta, \beta) \le 0$，即

$(\alpha, \beta)^2 \le (\alpha, \alpha)(\beta, \beta)$，其中的等号成立当且仅当方程 $(\alpha, \alpha)x^2 + 2(\alpha, \beta)x + (\beta, \beta) = 0$有根, 即存在$x$使$(x\alpha+\beta, x\alpha+\beta)=0$，即$x\alpha+\beta=0$,也即$\alpha, \beta$线性相关。

✎应用到例题中的2个例子中, 分别得到:

(1)$(x_1y_1+\ldots+x_ny_n)^2 \le (x_1^2+\ldots+x_n^2)(y_1^2+\ldots+y_n^2)$

对任意实数x_i, y_i $(1 \le i \le n)$成立。

(2) $$\left(\int_a^b f(x)g(x)dx\right)^2 \le \int_a^b f(x)^2 dx \int_a^b g(x)^2 dx$$

对区间[*a*,*b*]上任意连续函数*f*(*x*),*g*(*x*)成立。

☞根据C-S不等式, 可以定义两个非零向量α, β的夹角

$$\theta = \arccos\frac{(\alpha,\beta)}{|\alpha||\beta|}$$

例4 (三角形不等式)对欧氏空间V中任意向量*α*, *β*，有|*α*+*β*|≤|*α*|+|*β*|。

证明: 由$(\alpha, \beta)^2 \le (\alpha, \alpha)(\beta,\beta)$有$|(\alpha, \beta)| \le |\alpha||\beta|$。因此

$$|\alpha+\beta|^2 \le |\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 = (|\alpha|+|\beta|)^2$$

$$\Rightarrow |\alpha+\beta| \le |\alpha|+|\beta|.$$

5.2 正交化

标准正交基

例1 实数域R上方程$x_1+x_2+x_3+x_4=0$的解空间W是欧氏空间R^4的子空间，$a_1=(1,-1,0,0)^T$, $a_2=(0,1,-1,0)^T$, $a_3=(0,0,1,-1)^T$,组成W的一组基S。将W中每个向量用它在S下的坐标表示，根据$a,b\in W$在S下的坐标$X=(x_1,x_2,x_3)^T$, $Y=(y_1,y_2,y_3)^T$并计算内积(a,b)。

解：$a=x_1a_1+x_2a_2+x_3a_3=(a_1,a_2,a_3)(x_1,x_2,x_3)^T=AX$,
$b=y_1a_1+y_2a_2+y_3a_3=(a_1,a_2,a_3)(y_1,y_2,y_3)^T=AY$，其中
$A=(a_1,a_2,a_3)$，于是

$$(a,b) = a^T b = (AX)^T (AY) = X^T A^T AY = X^T GY,$$

$$G = A^T A = \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T \end{pmatrix} (a_1, a_2, a_3) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

$$(a,b) = X^T GY = (x_1, x_2, x_3) \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

$$= 2x_1y_1 + 2x_2y_2 + 2x_3y_3 - x_1y_2 - x_2y_1 - x_2y_3 - x_3y_2.$$

由内积的双线性性可以推出: 对欧氏空间V中任意向量α_i, β_j和任意实数x_i, y_j, $(1\le i\le k, 1\le j\le m)$, 有

$$\left(\sum_{i=1}^{k} x_i\alpha_i, \sum_{j=1}^{m} y_j\beta_j\right) = \sum_{i=1}^{k}\sum_{j=1}^{m} x_i y_j(\alpha_i, \beta_j)$$

特别地, 取V的任意一组基 $M=\{\alpha_1,\ldots,\alpha_n\}$, 设$\alpha,\beta$在这组基$M$下的坐标分别是$X=(x_1,\ldots,x_n)^T$, $Y=(y_1,\ldots,y_n)^T$, 则

$$(\alpha,\beta) = \left(\sum_{i=1}^{n} x_i\alpha_i, \sum_{j=1}^{n} y_j\alpha_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j(\alpha_i, \alpha_j) = X^T SY$$

$$S = (s_{ij})_{n\times n}, s_{ij} = (\alpha_i, \alpha_j), \forall 1\le i, j\le n.$$

S是由基M中的向量两两的内积组成的矩阵, 称为内积(α,β)在基M下的度量矩阵(metric matrix), 也称为Gram方阵。

✎上述定义可推广到一般的向量组$T=\{\alpha_1,\ldots,\alpha_m\}$, 即$G=(t_{ij})_{m\times m}$, $t_{ij}=(\alpha_i, \alpha_j)$, $1\le i,j\le m$称为T的度量矩阵(metric matrix), 也称为Gram方阵。

☞选取3个两两正交的单位向量α_1, α_2, α_3作为基向量，则空间中的向量在这样的基向量下用坐标表示, 则由于

$$(\alpha_i,\alpha_j)=\begin{cases}1, & \text{when } i=j\\ 0, & \text{when } i\neq j\end{cases}$$

内积在这组基下的度量矩阵$S=I_{(3)}$为单位阵。坐标分别为$X=(x_1,x_2,x_3)^T$，$Y=(y_1,y_2,y_3)^T$的向量α, β的内积$(\alpha, \beta)=X^T I Y= X^T Y =x_1y_1+x_2y_2+x_3y_3$.

在n维实向量空间V中定义内积之后, 选取V的这样的基$M=\{\alpha_1,\ldots,\alpha_n\}$，使度量矩阵$S=((\alpha_i,\alpha_j))_{nxn}$是$n$阶单位阵, 也就是要求基向量之间的内积满足条件

$$(\alpha_i,\alpha_j)=\begin{cases}1, & \text{when } i=j\\ 0, & \text{when } i\neq j\end{cases}$$

从而向量α, β的坐标 $X=(x_1,\ldots,x_n)^T$，$Y=(y_1,\ldots,y_n)^T$ 计算内积的公式有最简单的形式$(\alpha, \beta)=X^TIY= X^TY =x_1y_1+\ldots+x_ny_n$。因此, 以上对基向量内积的要求就是:基 $\{\alpha_1,\ldots,\alpha_n\}$由两两正交的单位向量组成。

定义5.2.1 V中由两两正交的非零向量组成的向量组$\{\alpha_1,\ldots,\alpha_k\}$，称为正交向量组。如果$V$的基$M$是正交向量组，就称$M$为正交基。 由两两正交的单位向量组成的基称为标准正交基。

引理5.2.2 (1) 实矩阵A的列向量组是正交向量组$\Leftrightarrow$ A^TA是可逆对角矩阵;

(2)实矩阵A的列向量组是标准正交向量组$\Leftrightarrow A^TA=I$;

(3)n阶实方阵A列向量组是R^{nx1}标准正交基$\Leftrightarrow$ $A^TA=I \Leftrightarrow A^T=A^{-1} \Leftrightarrow AA^T=I \Leftrightarrow$ A行向量组是R^{1xn}标准正交基.

引理 欧氏空间V中的正交向量组$\{\alpha_1,\ldots,\alpha_k\}$线性无关。

证明：略

Gram-Schmidt 正交化方法

定理 n维欧氏空间V必然存在标准正交基。

证明(算法5.2.1)任取V的一组基$M=\{\alpha_1,\ldots,\alpha_n\}$.取$\beta_1=\alpha_1$,并依次对$2\le k\le m$取

$$\beta_k=\alpha_k-\sum_{i=1}^{k-1}\frac{(\alpha_k,\beta_i)}{(\beta_i,\beta_i)}\beta_i$$

得到正交向量组，在将之归一化即得标准正交基。

命题: 设$M_1=\{\alpha_1,\ldots,\alpha_n\}$是欧氏空间的一组基, 则存在$V$的标准正交基$M=\{\gamma_1,\ldots,\gamma_n\}$使

$$(\gamma_1,\ldots,\gamma_n)=(\alpha_1,\ldots,\alpha_n)T$$

对某个上三角矩阵

$$T=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ & b_{22} & \cdots & b_{2n} \\ & & \ddots & \vdots \\ & & & b_{nn} \end{pmatrix}$$

成立。对每个$1\le k\le n$，$\{\gamma_1,\ldots,\gamma_k\}$是$\{\alpha_1,\ldots,\alpha_k\}$生成的子空间的一组标准正交基。

推论 5.2.1 欧氏空间V中任何一组两两正交的单位向量$S=\{\alpha_1,\ldots,\alpha_k\}$都能扩充为$V$的一组标准正交基。

证明: 将S扩充为V的一组基$M_1=\{\alpha_1,\ldots\alpha_k,\ldots,\alpha_n\}$, 再利用Gram-Schmidt正交化方法, 用$\beta_{k+1},\ldots,\beta_n$替换$\alpha_{k+1},\ldots,\alpha_n$得到标准正交基 $M=\{\alpha_1,\ldots\alpha_k, \beta_{k+1},\ldots,\beta_n\}$。

练习 试求R^4中线性无关的向量组$\alpha_1=(1,0,1,0)^T$, $\alpha_2=(0,-1,1,-1)^T$, $\alpha_3=(1,1,1,1)^T$所生成的子空间的一组正交基，并扩充成为R^4的一组标准正交基。

$$\gamma_1=\frac{1}{\sqrt{2}}(1,0,1,0),\gamma_2=\frac{1}{\sqrt{10}}(-1,-2,1,-2),\gamma_3=\frac{1}{\sqrt{10}}(-2,1,2,1)$$

$$\gamma_4=\frac{1}{\sqrt{2}}(0,1,0,-1).$$