第五章 相似矩阵及二次型

一、向量的内积、长度及正交性

定义1 设有n维向量

$$x=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad y=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},$$

令 $[x,y]=x_1y_1+x_2y_2+\cdots+x_ny_n$

$[x,y]$ 称为向量 x 与 y 的内积。

矩阵记号：$[x,y]=x^Ty$

内积的性质：

(i) $[x,y]=[y,x]$

(ii) $[\lambda x,y]=\lambda[x,y]$

(iii) $[x+y,z]=[x,z]+[y,z]$

(iv) 当 $x=0$ 时，$[x,x]=0$；当 $x\neq0$ 时，$[x,x]>0$.

Schwarz 不等式：$[x,y]^2\leq[x,x][y,y]$

解析几何中引进向量的数量积，即

$$x\cdot y=\underbrace{|x|\,|y|}_{\text{大小或长度}}\cos\underbrace{\theta}_{\text{夹角}}$$

在直角坐标系中，有

$$(x_1,x_2,x_3)\cdot(y_1,y_2,y_3)=x_1y_1+x_2y_2+x_3y_3$$

定义2 令 $\|x\|=\sqrt{[x,x]}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2}$，

$\|x\|$ 称为n维向量 x 的长度(或范数).

当 $\|x\|=1$ 时，称 x 为单位向量。

向量的长度的性质：

(i) 非负性：当 $x\neq 0$ 时，$\|x\|>0$；当 $x=0$ 时，$\|x\|=0$；

(ii) 齐次性：$\|\lambda x\|=|\lambda|\|x\|$；

(iii) 三角不等式：$\|x+y\|\leq\|x\|+\|y\|$.

当 $x\neq 0, y\neq 0$ 时，

$$\theta=\arccos\frac{[x,y]}{\|x\|\|y\|}$$

称为n维向量 x 与 y 的夹角。

当 $[x,y]=0$ 时，称向量 x 与 y 正交.

若 $x=0$，则 x 与任何向量都正交.

正交向量组：一组两两正交的非零向量。

定理1 若n维向量 $a_1,a_2,\cdots,a_r$ 是一组两两正交的非零向量，则 $a_1,a_2,\cdots,a_r$ 线性无关。

证明：考察 $\lambda_1a_1+\lambda_2a_2+\cdots+\lambda_ra_r=0$

$$\left.\begin{aligned}&\Rightarrow \lambda_1a_1^Ta_1=0\\ &a_1\neq 0,\ a_1^Ta_1=\|a_1\|^2\neq 0\end{aligned}\right\}\Rightarrow\lambda_1=0$$

同理 $\lambda_2=\cdots=\lambda_r=0$

则正交向量组 $a_1,a_2,\cdots,a_r$ 线性无关.

例1 解：据题意，即有 $\begin{cases}a_1^Ta_3=0\\ a_2^Ta_3=0\end{cases}$ 即 $\begin{pmatrix}a_1^T\\ a_2^T\end{pmatrix}a_3=0$

记 $A=\begin{pmatrix}a_1^T\\ a_2^T\end{pmatrix}$，则 a_3 满足 $Ax=0$，即转化为求解齐次线性方程组

$$\begin{pmatrix}1&1&1\\1&-2&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$$

$$A \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \Rightarrow \begin{cases} x_1 = -x_3 \\ x_2 = 0 \end{cases}$$

基础解系 $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ 即可取 $a_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ 为所求.

定义3 设n维向量 $e_1, e_2, \cdots, e_r$ 是向量空间 $V \subset R^n$ 的一个基，如果 $e_1, e_2, \cdots, e_r$ 两两正交，且都是单位向量，则称 $e_1, \cdots, e_r$ 是V的一个规范正交基。

例如 $e_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}$，$e_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{pmatrix}$，$e_3 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$，$e_4 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$

是 R^4 的一个规范正交基。

若 $e_1, \cdots, e_r$ 是V的一个规范正交基，那么V中任一向量 a 应能由 $e_1, \cdots, e_r$ 线性表示，设表示式为

$$a = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_r e_r$$

则系数可由下式求解

$$\lambda_i = e_i^T a = [a, e_i]$$

即向量在规范正交基中的坐标计算公式。

设 $a_1, \cdots, a_r$ 是向量空间V的一个基，要求V的一个规范正交基，即要找一组两两正交的单位向量 $e_1, \cdots, e_r$，使 $e_1, \cdots, e_r$ 与 $a_1, \cdots, a_r$ 等价，这一问题称为把 $a_1, \cdots, a_r$ 这个基规范正交化。

*Schmidt正交化的几何解释：　　　课后思考！

$b_2=a_2-c_2$，　$c_2=[a_2,\frac{b_1}{\|b_1\|}]\frac{b_1}{\|b_1\|}$ 表明 c_2 为 a_2 在 b_1 上的投影向量.

$b_3=a_3-c_3$，　$c_3=c_{31}+c_{32}=[a_3,\frac{b_1}{\|b_1\|}]\frac{b_1}{\|b_1\|}+[a_3,\frac{b_2}{\|b_2\|}]\frac{b_2}{\|b_2\|}$

表明 c_3 为 a_3 分别在 b_1, b_2 上的投影向量之和或者说

c_3 为 a_3 在 b_1 和 b_2 确定的平面上的投影向量.

… … … … …

$b_r=a_r-c_r$，$c_r=c_{r1}+c_{r2}+\cdots+c_{r,r-1}$

$$=[a_r,\frac{b_1}{\|b_1\|}]\frac{b_1}{\|b_1\|}+[a_r,\frac{b_2}{\|b_2\|}]\frac{b_2}{\|b_2\|}+\cdots+[a_r,\frac{b_{r-1}}{\|b_{r-1}\|}]\frac{b_{r-1}}{\|b_{r-1}\|}$$

表明 c_r 为 a_r 分别在 $b_1,b_2,\cdots,b_{r-1}$ 上的投影向量之和或者说

c_r 为 a_r 在 $b_1,b_2,\cdots,b_{r-1}$ 确定的超平面上的投影向量.

三维空间中正交化的另一种方法：

取 $b_1=a_1$，$b_2=a_1\times a_2$，$b_3=b_2\times b_1$

规范正交化具体步骤：

(i) Schmidt 正交化

取 $b_1 = a_1$；

$$b_2 = a_2 - \frac{[b_1, a_2]}{[b_1, b_1]} b_1 \ ;$$

$$\cdots\ \cdots\ \cdots\ \cdots\ \cdots$$

$$b_r = a_r - \frac{[b_1, a_r]}{[b_1, b_1]} b_1 - \frac{[b_2, a_r]}{[b_2, b_2]} b_2 - \cdots - \frac{[b_{r-1}, a_r]}{[b_{r-1}, b_{r-1}]} b_{r-1} \ .$$

(ii) 单位化

取 $e_1 = \frac{1}{\|b_1\|} b_1,\ e_2 = \frac{1}{\|b_2\|} b_2,\ \cdots,\ e_r = \frac{1}{\|b_r\|} b_r$ ，

即得 V 的一个规范正交基。

不仅满足 $b_1, \cdots, b_r$ 与 $a_1, \cdots, a_r$ 等价，还满足：对任何 $k\ (1 \le k \le r)$，向量组 $b_1, \cdots, b_k$ 与 $a_1, \cdots, a_k$ 等价。

例2 解：(i) 取 $b_1 = a_1$

$$b_2 = a_2 - \frac{[a_2, b_1]}{[b_1, b_1]} b_1$$

$$= \begin{pmatrix} -1 \\ 3 \\ 1 \end{pmatrix} - \frac{4}{6} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \frac{5}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

$$b_3 = a_3 - \frac{[a_3, b_1]}{[b_1, b_1]} b_1 - \frac{[a_3, b_2]}{[b_2, b_2]} b_2$$

$$= \begin{pmatrix} 4 \\ -1 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \frac{5}{3} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

(ii) 单位化

$$e_1 = \frac{b_1}{\|b_1\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix},\ e_2 = \frac{b_2}{\|b_2\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},\ e_3 = \frac{b_3}{\|b_3\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$ 即所求.

例3 （略） 课后练习.

定义4 如果n阶矩阵A满足

$$A^TA=E \quad (\text{即} A^{-1}=A^T)$$

那么称A为正交矩阵，简称正交阵.

若用A的列向量表示，即

$$\begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix}(a_1,a_2,\cdots,a_n)=E$$

即 n^2 个关系式 $a_i^Ta_j=\delta_{ij}=\begin{cases}1, & \text{当} i=j \\ 0, & \text{当} i\neq j\end{cases} \quad (i,j=1,2,\cdots,n)$

方阵A为正交阵的充分必要条件是A的列向量都是单位向量，且两两正交.

例4 验证（略）.

正交矩阵的性质：

(i) 若A为正交阵，则 $A^{-1}=A^T$ 也是正交阵，且 $|A|=1$ 或 (-1).

(ii) 若A和B都是正交阵，则AB也是正交阵.

定义5 若P为正交矩阵，则线性变换 $y=Px$ 称为正交变换.

正交变换的特性：

设 $y=Px$ 为正交变换，则

$$\|y\|=\sqrt{y^Ty}=\sqrt{x^TP^TPx}=\sqrt{x^Tx}=\|x\|$$

即经过正交变换，向量的长度保持不变.

二、方阵的特征值与特征向量

定义6 设A是n阶矩阵，如果数λ和n维非零列向量x使关系式

$$Ax=\lambda x \tag{1}$$

成立，那么这样的数λ称为方阵A的特征值，非零向量x称为A的对应于λ的特征向量。

(1)式即

$$(A-\lambda E)x=0$$

此即含n个未知数n个方程的齐次线性方程组。具有非零解的充分必要条件是系数行列式

$$|A-\lambda E|=0$$

即

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}=0$$

以λ为未知数的一元n次方程，称为方阵A的特征方程。

左端$|A-\lambda E|$是λ的n次多项式，记作$f(\lambda)$，称为方阵A的特征多项式。

表明A的特征值就是特征方程的解。特征方程在复数范围内恒有解，其个数为方程的次数（重根按重数计算），因此，n阶矩阵A在复数范围内有n个特征值。

特征值的性质:

(i) $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$

(ii) $\lambda_1\lambda_2\cdots\lambda_n=|A|$

设 $\lambda=\lambda_i$ 是方阵A的一个特征值，则由方程

$$(A-\lambda_i E)x=0$$

可求得非零解 $x=p_i$，那么 p_i 便是A的对应于特征值 λ_i 的特征向量.
（若 λ_i 为实数，则 p_i 可取实向量；若 λ_i 为复数，则 p_i 为复向量.）

例5 求特征值和特征向量.

基本步骤：① 写出特征多项式，求特征方程的根

② 写出对应特征值的齐次线性方程组，求对应特征值的特征向量.

解：① 写出特征多项式

$$|A-\lambda E|=\begin{vmatrix}3-\lambda & -1\\ -1 & 3-\lambda\end{vmatrix}=(3-\lambda)^2-1=8-6\lambda+\lambda^2$$

$$=(4-\lambda)(2-\lambda)$$

$\therefore$ A的特征值为 $\lambda_1=4$，$\lambda_2=4$

② 求对应特征值的特征向量

(i) $\lambda_1=2$ 即求 $(A-2E)x=0$

$$A-2E=\begin{pmatrix}3-2 & -1\\ -1 & 3-2\end{pmatrix}=\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix}\overset{r_2+r_1}{\sim}\begin{pmatrix}1 & -1\\ 0 & 0\end{pmatrix}\Rightarrow x_1=x_2$$

特征向量可取 $p_1=\begin{pmatrix}1\\1\end{pmatrix}$

(ii) $\lambda_2=4$ 即求 $(A-4E)x=0$

$$A-4E=\begin{pmatrix}3-4 & -1\\ -1 & 3-4\end{pmatrix}=\begin{pmatrix}-1 & -1\\ -1 & -1\end{pmatrix}\overset{r_2-r_1}{\sim}\begin{pmatrix}-1 & -1\\ 0 & 0\end{pmatrix}\overset{r_1\times(-1)}{\sim}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\Rightarrow x_1=-x_2$$

特征向量可取 $p_2=\begin{pmatrix}-1\\ 1\end{pmatrix}$

实际上，对应于特征值 λ_i 的齐次线性方程组 $(A-\lambda_i E)x=0$ 的通解中所有非零解都是方阵A的对应于特征值 λ_i 的特征向量。

例6 解：①写出A的特征多项式

$$|A-\lambda E|=\begin{vmatrix}-1-\lambda & 1 & 0\\ -4 & 3-\lambda & 0\\ 1 & 0 & 2-\lambda\end{vmatrix}=(2-\lambda)(1-\lambda)^2$$

∴ A的特征值为 $\lambda_1=2$，$\lambda_2=\lambda_3=1$

②求对应特征值的特征向量

(i) $\lambda_1=2$ 即求 $(A-2E)x=0$

$$A-2E=\begin{pmatrix}-3 & 1 & 0\\ -4 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\Rightarrow\begin{cases}x_1=0\\ x_2=0\end{cases}$$

特征向量可取 $p_1=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$

(ii) $\lambda_2=\lambda_3=1$，即求 $(A-E)x=0$

$$A-E=\begin{pmatrix}-2 & 1 & 0\\ -4 & 2 & 0\\ 1 & 0 & 1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 2\\ 0 & 0 & 0\end{pmatrix}\Rightarrow\begin{cases}x_1=-x_3\\ x_2=-2x_3\end{cases}$$

特征向量可取 $p_2=\begin{pmatrix}-1\\ -2\\ 1\end{pmatrix}$

例7 解：①写出A的特征多项式

$$|A-\lambda E|=\begin{vmatrix}-2-\lambda & 1 & 1\\ 0 & 2-\lambda & 0\\ -4 & 1 & 3-\lambda\end{vmatrix}=(2-\lambda)\begin{vmatrix}-2-\lambda & 1\\ -4 & 3-\lambda\end{vmatrix}$$

$$=(2-\lambda)(\lambda^2-\lambda-2)=-(\lambda+1)(\lambda-2)^2$$

∴ A的特征值为 $\lambda_1=-1,\ \lambda_2=\lambda_3=2$

②求对应特征值的特征向量

(i) $\lambda_1=-1$ 即求 $(A+E)x=0$

$$A+E=\begin{pmatrix}-1 & 1 & 1\\ 0 & 3 & 0\\ -4 & 1 & 4\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\Rightarrow\begin{cases}x_1=x_3\\ x_2=0\end{cases}$$

可得基础解系 $p_1=\begin{pmatrix}1\\0\\1\end{pmatrix}$

则 $k_1p_1\ (k_1\neq 0)$ 是对应于 $\lambda_1=-1$ 的全部特征向量.

(ii) $\lambda_2=\lambda_3=2$ 即求 $(A-2E)x=0$

$$A-2E=\begin{pmatrix}-4 & 1 & 1\\ 0 & 0 & 0\\ -4 & 1 & 1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & -\frac{1}{4} & -\frac{1}{4}\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\Rightarrow x_1=\frac{1}{4}x_2+\frac{1}{4}x_3$$

可得基础解系

$$p_2=\begin{pmatrix}\frac{1}{4}\\1\\0\end{pmatrix},\quad p_3=\begin{pmatrix}\frac{1}{4}\\0\\1\end{pmatrix}$$

则 $k_2p_2+k_3p_3$ (k_2、k_3不同时为0)是对应于 $\lambda_2=\lambda_3=2$ 的全部特征向量.

例8　略. 若λ是A的特征值，则$\varphi(\lambda)$是$\varphi(A)$的特征值.

例9　解：据例8的结论.

$$\left.\begin{array}{l} A^*=|A|A^{-1}, \\ |A|=\lambda_1\lambda_2\lambda_3=-2 \end{array}\right\} \Rightarrow A^*=-2A^*$$

原式 $=-2A^{-1}+3A-2E$

记 $\varphi(A)=-2A^{-1}+3A-2E$，则 $\varphi(\lambda)=-\frac{2}{\lambda}+3\lambda-2$

$\because$ A的特征值为$1,-1,2$，$\therefore$ $\varphi(A)$的特征值为$\varphi(1),\varphi(-1),\varphi(2)$

即 $-1,-3,3$.

定理2　设$\lambda_1,\lambda_2,\cdots,\lambda_m$是方阵$A$的$m$个特征值，$p_1,p_2,\cdots,p_m$依次是与之对应的特征向量，如果$\lambda_1,\lambda_2,\cdots,\lambda_m$各不相等，则$p_1,p_2,\cdots,p_m$线性无关.

证明略.

例10. 反证法：假设p_1+p_2是A的特征向量，则应存在数λ，使

$$\left.\begin{array}{l} A(p_1+p_2)=\lambda(p_1+p_2) \\ \text{而}\ Ap_1=\lambda_1p_1 \\ \quad\ \ Ap_2=\lambda_2p_2 \end{array}\right\} \Rightarrow \lambda(p_1+p_2)=\lambda_1p_1+\lambda_2p_2$$

$$\left.\begin{array}{l} \text{即}\ (\lambda_1-\lambda)p_1+(\lambda_2-\lambda)p_2=0 \\ \lambda_1\neq\lambda_2 \Rightarrow p_1,p_2\text{线性无关} \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \lambda_1=\lambda \\ \lambda_2=\lambda \end{array}\right\} \Rightarrow \lambda_1=\lambda_2\ \text{与已知矛盾}$$

故 p_1+p_2不是A的特征向量.

（第五版）P134
课后作业：习题五 2.(2)，3，4，5，6(2)，7，

三、相似矩阵

定义7 设A、B都是n阶矩阵，若有可逆矩阵P，使

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

则称B是A的相似矩阵，或说矩阵A与B相似。对A进行的运算$P^{-1}AP$称为对A进行相似变换，可逆矩阵P称为把A变成B的相似变换矩阵。

定理3 若n阶矩阵A与B相似，则A与B的特征多项式相同，从而A与B的特征值亦相同。

证明：存在$P^{-1}AP=B$

$$\begin{aligned}\text{则 } |B-\lambda E| &= |P^{-1}AP-\lambda E| = |P^{-1}AP-P^{-1}(\lambda E)P| \\ &= |P^{-1}(A-\lambda E)P| = |P^{-1}|\,|A-\lambda E|\,|P| \\ &= |A-\lambda E|\end{aligned}$$

推论 若n阶矩阵A与对角阵

$$\Lambda=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

相似，则$\lambda_1,\lambda_2,\cdots,\lambda_n$即是A的n个特征值。

若存在$P^{-1}AP=\Lambda$为对角阵，则

$$A^k=P^{-1}\Lambda^k P,\quad \varphi(A)=P\varphi(\Lambda)P^{-1}$$

由此可方便地计算出A的多项式$\varphi(A)$。

附加的结论：$f(\lambda)$是矩阵A的特征多项式，则$f(A)=0$。证明略。

一个矩阵与某个对角阵相似在求矩阵的特征值和矩阵多项式时会带来很大的便利，因此我们很有必要研究对于n阶矩阵A，如何将其化成对角阵。寻求相似变换矩阵P，使$P^{-1}AP=\Lambda$为对角阵，这类问题称为把矩阵A对角化。

为简单起见，我们假设已经找到P，使$P^{-1}AP=\Lambda$，分析相似变换矩阵P的性质。从向量组的角度来看，把P看成是由列向量组$p_1, p_2, \cdots, p_n$构成的矩阵

$$\left.\begin{array}{l}\text{即 } P=(p_1, p_2, \cdots, p_n) \\ \text{又 } P^{-1}AP=\Lambda \Rightarrow AP=P\Lambda\end{array}\right\} \Rightarrow A(p_1, p_2, \cdots, p_n)=(p_1, p_2, \cdots, p_n)\begin{pmatrix}\lambda_1 & & \\ & \lambda_2 & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$$

即 $A(p_1, p_2, \cdots, p_n)=(\lambda_1 p_1, \lambda_2 p_2, \cdots, \lambda_n p_n)$

$\Rightarrow \quad Ap_i=\lambda_i p_i \quad (i=1,2,\cdots,n)$

上式表明λ_i是A的特征值，P的列向量p_i是A的对应于特征值λ_i的特征向量。

反过来，若知道A恰好有n个特征值，并对应地求得n个特征向量，则这n个特征向量即可构成矩阵P，只是该矩阵P仅使$AP=P\Lambda$成立，显然这样构成的矩阵P不一定可逆，只有在构成的矩阵P可逆时，才有$P^{-1}AP=\Lambda$，即A与Λ相似，此时P才是相似变换矩阵。

定理4　n阶矩阵A与对角阵相似（即A能对角化）的充分必要条件是A有n个线性无关的特征向量。

推论　如果n阶矩阵A的n个特征值互不相等，则A与对角阵相似。

当A的特征方程有重根时，A不一定能对角化。

。。。

例11 解： $|A-\lambda E| = \begin{vmatrix} -\lambda & 0 & 1 \\ 1 & 1-\lambda & x \\ 1 & 0 & -\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix}$

$$= (1-\lambda)(\lambda^2-1) = -(\lambda-1)^2(\lambda+1)$$

可得A的特征值 $\lambda_1=-1$，$\lambda_2=\lambda_3=1$

① 对单根 $\lambda_1=-1$，即对应 $(A+E)x=0$

$$A+E = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & x \\ 1 & 0 & 1 \end{pmatrix} \xrightarrow[r_2-r_1]{r_3-r_1} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & x-1 \\ 0 & 0 & 0 \end{pmatrix}$$

可见 $R(A+E)=2$ 故 $(A+E)x=0$ 的基础解系只含1个向量.

② 对重根 $\lambda_2=\lambda_3=1$，即对应 $(A-E)x=0$

$$A-E = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & x \\ 1 & 0 & -1 \end{pmatrix} \xrightarrow[r_2+r_1]{r_3+r_1} \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & x+1 \\ 0 & 0 & 0 \end{pmatrix}$$

可见 $R(A-E)$ 决定于 x 的取值，据题意，若要使A能够对角化，只须再找出2个线性无关的特征向量，故只须 $R(A-E)=1$ 即可，此时 $(A-E)x=0$ 的基础解系中就含2个向量，故 $x=-1$ 即可.

综上所述，$x=-1$ 时，A能对角化。

课后作业：P_{135}，15

上一节我们学习了矩阵的相似关系，同时引入了相似变换的定义以及描述相似变换运算的矩阵——相似变换矩阵，并进一步讨论了一类特殊的相似变换——对角化。然而对于n阶矩阵而言，对角化是比较复杂的，我们只讨论A是对称阵的情形.这一节我们来学习对称阵的对角化.

四、对称矩阵的对角化

定理5 对称阵的特征值为实数.

证明：设$\lambda \in \mathbb{Z}$为对称阵A的特征值，复向量x为对应的特征向量。

即 $Ax = \lambda x \quad (x \neq 0)$.

A为实矩阵，则 $\bar{A} = A$，有

$$\left.\begin{aligned} \overline{Ax} &= \bar{A}\bar{x} = A\bar{x} \\ \overline{Ax} &= \overline{\lambda x} = \bar{\lambda}\bar{x} \end{aligned}\right\} \Rightarrow A\bar{x} = \bar{\lambda}\bar{x}$$

构造

$$\left.\begin{aligned} \bar{x}^T A x &= \bar{x}^T (Ax) = \bar{x}^T \lambda x = \lambda \bar{x}^T x \\ \bar{x}^T A x &= (\bar{x}^T A^T) x = (A\bar{x})^T x = (\bar{\lambda}\bar{x})^T x = \bar{\lambda}\bar{x}^T x \end{aligned}\right\}$$

$$\left.\begin{aligned} &\Rightarrow (\lambda - \bar{\lambda}) \bar{x}^T x = 0 \\ &x \neq 0 \Rightarrow \bar{x}^T x = \sum_{i=1}^{n} \bar{x}_i x_i = \sum_{i=1}^{n} |x_i|^2 \neq 0 \end{aligned}\right\}$$

$$\Rightarrow \lambda = \bar{\lambda} \Rightarrow \lambda \in R$$

当$\lambda_i \in R$，$(A - \lambda_i E)x = 0$是实系数方程组，由$|A - \lambda_i E| = 0$知必有实的基础解系，即对应的特征向量可以取实向量。

定理6 设λ_1, λ_2是对称阵A的两个特征值，p_1, p_2是对应的特征向量。若$\lambda_1 \neq \lambda_2$，则p_1与p_2正交。

证明：根据已知，有 $Ap_1 = \lambda_1 p_1$，$Ap_2 = \lambda_2 p_2$，$\lambda_1 \neq \lambda_2$，$A^T = A$

由于 $A^T=A$，则

$$\lambda_1 p_1^T=(\lambda_1 p_1)^T=(Ap_1)^T=p_1^TA^T=p_1^TA$$

构造

$$\lambda_1 p_1^T p_2=p_1^TAp_2=p_1^T\lambda_2 p_2=\lambda_2 p_1^T p_2$$

$$\left.\begin{array}{r}(\lambda_1-\lambda_2)p_1^Tp_2=0\\ \lambda_1\neq\lambda_2\end{array}\right\}\Rightarrow p_1^Tp_2=0\Rightarrow p_1\text{与}p_2\text{正交}$$

定理7 设 A 为 n 阶对称阵，则必有正交阵 P，使 $P^{-1}AP=P^TAP=\Lambda$，其中 Λ 是以 A 的 n 个特征值为对角元的对角阵.

推论 设 A 为 n 阶对称阵，λ 是 A 的特征方程的 k 重根，则矩阵 $A-\lambda E$ 的秩 $R(A-\lambda E)=n-k$，从而对应特征值 λ 恰有 k 个线性无关的特征向量.

证明：A 是对称阵 $\Rightarrow A$ 与 Λ 相似 $\Longrightarrow A-\lambda E$ 与 $\Lambda-\lambda E$ 相似

$$\Updownarrow$$

$$P^{-1}AP=\Lambda\Longrightarrow P^{-1}(A-\lambda E)P=P^{-1}AP-P^{-1}\lambda EP=\Lambda-\lambda E$$

设 λ 是 A 的 k 重特征根，则 $\lambda_1,\cdots,\lambda_n$ 中有 k 个等于 λ，有 $n-k$ 个不等于 λ，从而对角阵 $\Lambda-\lambda E$ 的对角元恰有 k 个等于 0，于是 $R(\Lambda-\lambda E)=n-k$，又 $R(A-\lambda E)=R(\Lambda-\lambda E)$.

故 $R(A-\lambda E)=n-k$

对称阵A对角化的步骤：

(1) 求出A的全部互不相等的特征值 $\lambda_1, \cdots, \lambda_s$，它们的重数依次为 $k_1, \cdots, k_s$ $(k_1 + \cdots + k_s = n)$.

(2) 对每个 k_i 重特征值 λ_i，求方程 $(A - \lambda_i E)x = 0$ 的基础解系，得出 k_i 个线性无关的特征向量。再把它们正交化、单位化，得 k_i 个两两正交的单位向量。因 $k_1 + \cdots + k_s = n$，故总共可得 n 个两两正交的单位特征向量.

(3) 用这 n 个两两正交的单位特征向量构成正交阵 P，便有 $P^{-1}AP = P^{T}AP = \Lambda$.

注意：Λ 中对角元的排列次序应与 P 中列向量的排列次序相对应.

例12 解：(1) 求特征值

$$|A - \lambda E| = \begin{vmatrix} -\lambda & -1 & 1 \\ -1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} \xlongequal{r_1 - r_2} \begin{vmatrix} 1-\lambda & \lambda-1 & 0 \\ -1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} \xlongequal{c_2 + c_1} \begin{vmatrix} 1-\lambda & 0 & 0 \\ -1 & -1-\lambda & 1 \\ 1 & 2 & -\lambda \end{vmatrix}$$

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

$$= -(\lambda - 1)^2 (\lambda + 2)$$

A的特征值为 $\lambda_1 = -2$，$\lambda_2 = \lambda_3 = 1$

(2) 求对应特征值的特征向量

(i) $\lambda_1 = -2$ 即求 $(A + 2E)x = 0$

$$A + 2E = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

得基础解系 $\xi_1 = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$，将 ξ_1 单位化得 $p_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$

(ii) $\lambda_2=\lambda_3=1$ 即求 $(A-E)x=0$

$$A-E=\begin{pmatrix}-1&-1&1\\-1&-1&1\\1&1&-1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&1&-1\\0&0&0\\0&0&0\end{pmatrix}$$

得基础解系 $\xi_2=\begin{pmatrix}-1\\1\\0\end{pmatrix}$, $\xi_3=\begin{pmatrix}1\\0\\1\end{pmatrix}$

将 ξ_2 和 ξ_3 正交化：取 $\eta_2=\xi_2$

$$\eta_3=\xi_3-\frac{[\eta_2,\xi_3]}{\|\eta_2\|^2}\eta_2=\begin{pmatrix}1\\0\\1\end{pmatrix}+\frac{1}{2}\begin{pmatrix}-1\\1\\0\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1\\1\\2\end{pmatrix}$$

再将 η_2,η_3 单位化，得

$$p_2=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix},\quad p_3=\frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\2\end{pmatrix}$$

(3) 构造正交阵

$$P=(p_1,p_2,p_3)=\begin{pmatrix}-\frac{1}{\sqrt{3}}&-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}\\-\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&0&\frac{2}{\sqrt{6}}\end{pmatrix}$$

即 $P^{-1}AP=P^TAP=\Lambda=\begin{pmatrix}-2&0&0\\0&1&0\\0&0&1\end{pmatrix}$

例13 解：A 对称 $\Rightarrow$ A 可对角化，即存在可逆矩阵 P 及对角阵 Λ，使

$P^{-1}AP=\Lambda$，则 $A=P\Lambda P^{-1}$

因此，$A^n=P\Lambda^n P^{-1}$，则原问题转化为求 P 和 Λ.

(1) 求特征值

$$|A-\lambda E|=\begin{vmatrix}2-\lambda&-1\\-1&2-\lambda\end{vmatrix}=\lambda^2-4\lambda+3=(\lambda-1)(\lambda-3)$$

得 A 的特征值为 $\lambda_1=1,\lambda_2=3$. 则 $\Lambda=\begin{pmatrix}1&0\\0&3\end{pmatrix}$, $\Lambda^n=\begin{pmatrix}1&0\\0&3^n\end{pmatrix}$.

(2) 求对应特征值的特征向量

(i) $\lambda_1=1$，即求 $(A-E)x=0$

$$A-E=\begin{pmatrix}1&-1\\-1&1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&-1\\0&0\end{pmatrix}，得\ \xi_1=\begin{pmatrix}1\\1\end{pmatrix}$$

(ii) $\lambda_2=3$，即求 $(A-3E)x=0$

$$A-3E=\begin{pmatrix}-1&-1\\-1&-1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&1\\0&0\end{pmatrix}，得\ \xi_2=\begin{pmatrix}1\\-1\end{pmatrix}$$

根据定理6，ξ_1 与 ξ_2 正交

故 $P=(\xi_1,\xi_2)=\begin{pmatrix}1&1\\1&-1\end{pmatrix}$，则 $P^{-1}=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$

$$A^n=P\Lambda^nP^{-1}=\frac{1}{2}\begin{pmatrix}1&1\\1&-1\end{pmatrix}\begin{pmatrix}1&0\\0&3^n\end{pmatrix}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$$

$$=\frac{1}{2}\begin{pmatrix}1+3^n&1-3^n\\1-3^n&1+3^n\end{pmatrix}$$

课后作业：P_{135} 17, 19(1), 25(1).

五. 二次型及其标准形

解析几何中，二次曲线方程

$$ax^2+bxy+cy^2=1.$$

当我们选择适当的坐标旋转变换

$$\begin{cases}x=x'\cos\theta-y'\sin\theta\\y=x'\sin\theta+y'\cos\theta\end{cases}$$

把方程化为标准形 $mx'^2+ny'^2=1$

化标准形的过程就是通过变量的线性变换化简一个二次齐次多项式，使它只含有平方项。

这节课我们讨论一般化的 n 个变量的二次齐次多项式的化简问题。

定义8 含有 n 个变量 $x_1, x_2, \cdots, x_n$ 的二次齐次函数

$$f(x_1, x_2, \cdots, x_n) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2$$
$$+ 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{n-1,n}x_{n-1}x_n$$

称为二次型。

若取 $a_{ji} = a_{ij}$，则 $2a_{ij}x_ix_j = a_{ij}x_ix_j + a_{ji}x_jx_i$，则

$$\begin{aligned} f = & a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ & + \cdots + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2 \\ = & \sum_{i,j=1}^{n} a_{ij}x_ix_j \end{aligned}$$

对于二次型，我们要讨论的主要问题是：

寻求可逆的线性变换

$$\begin{cases} x_1 = c_{11}y_1 + c_{12}y_2 + \cdots + c_{1n}y_n \\ x_2 = c_{21}y_1 + c_{22}y_2 + \cdots + c_{2n}y_n \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \\ x_n = c_{n1}y_1 + c_{n2}y_2 + \cdots + c_{nn}y_n \end{cases} \qquad (7)$$

使得二次型只含平方项，使

$$f = k_1y_1^2 + k_2y_2^2 + \cdots + k_ny_n^2 \text{ —— 二次型的标准形}$$

若 $k_1, k_2, \cdots, k_n \in \{-1, 0, 1\}$，使

$$f = y_1^2 + \cdots + y_p^2 - y_{p+1}^2 - \cdots - y_r^2$$ —— 二次型的规范形

仅限于实系数范围：

将二次型表示成矩阵的形式

$$f = x_1(a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)$$
$$+ x_2(a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n)$$
$$+ \cdots$$
$$+ x_n(a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n)$$

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

$$= (x_1, x_2, \cdots, x_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

令 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$，$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

则二次型可表示成

$$f = x^T A x$$，其中 A 为对称阵.

例如 $f = x^2 - 3z^2 - 4xy + yz$ 的矩阵形式

$$f = (x, y, z)\begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

二次型 $\xleftrightarrow{\text{一一对应}}$ 对称阵

对称阵A的　　二次型f的

秩

若记 $C=(c_{ij})$，把可逆变换(7)记作

$$x=Cy$$

$$f=x^TAx=(Cy)^TACy=y^TC^TACy$$

定义9　设A和B是n阶矩阵，若有可逆矩阵C，使 $B=C^TAC$，则称矩阵A与B合同.

经可逆变换 $x=Cy$ 后，二次型f的矩阵由A变为与A合同的矩阵 C^TAC，且二次型的秩不变。

要使二次型f经可逆变换 $x=Cy$ 变成标准形，就是要使

$$y^TC^TACy=k_1y_1^2+k_2y_2^2+\cdots+k_ny_n^2$$

$$=(y_1,y_2,\cdots,y_n)\begin{pmatrix}k_1&&&\\&k_2&&\\&&\ddots&\\&&&k_n\end{pmatrix}\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}$$

也就是要使 C^TAC 成为对角阵.

即对于对称阵A，寻求可逆矩阵C，使 C^TAC 为对角阵. 这种二次型的标准化问题称为把对称阵A合同对角化.

定理8 任给二次型 $f=\sum_{i,j=1}^{n} a_{ij}x_i x_j \ (a_{ij}=a_{ji})$，总有正交变换 $x=Py$，

使 f 化为标准形

$$f=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2 \quad (=y^T\Lambda y)$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 f 的矩阵 $A=(a_{ij})$ 的特征值.

推论 任给 n 元二次型 $f(x)=x^TAx \ (A^T=A)$，总有

可逆变换 $x=Cz$，使 $f(Cz)$ 为规范形.

证明：据定理8，有

$$f(Py)=y^T\Lambda y=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2$$

设二次型 f 的秩为 r，则特征值 λ_i 中恰有 r 个不为0，

不妨设 $\lambda_1,\cdots,\lambda_r$ 不等于0，$\lambda_{r+1}=\cdots=\lambda_n=0$，

令 $K=\begin{pmatrix} k_1 & & & \\ & k_2 & & \\ & & \ddots & \\ & & & k_n \end{pmatrix}$，其中 $k_i=\begin{cases} \dfrac{1}{\sqrt{|\lambda_i|}}, & i\le r \\ 1, & i>r \end{cases}$

则 K 可逆，变换 $y=Kz$ 把 $f(Py)$ 化为

$$f(PKz)=z^TK^TP^TAPKz=z^TK^T\Lambda Kz$$

而 $K^T\Lambda K=\mathrm{diag}\left(\dfrac{\lambda_1}{|\lambda_1|},\cdots,\dfrac{\lambda_r}{|\lambda_r|},0,\cdots,0\right)$

记 $C=PK$，即可逆变换 $x=Cz$ 把 f 化成规范形

$$f(Cz)=\frac{\lambda_1}{|\lambda_1|}z_1^2+\cdots+\frac{\lambda_r}{|\lambda_r|}z_r^2$$

例14 解：二次型的矩阵为

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

直接使用例12的结果，可得正交阵

$$P=\begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}，使 P^TAP=\Lambda=\begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

即有正交变换

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

把二次型 f 化为标准形

$$f=-2y_1^2+y_2^2+y_3^2$$

若要把二次型 f 化成规范形，只需令

$$\begin{cases} y_1=\frac{1}{\sqrt{2}}z_1 \\ y_2=z_2 \\ y_3=z_3 \end{cases}$$

即得 f 的规范形

$$f=-z_1^2+z_2^2+z_3^2$$