第六章
第二节： 矩阵相似与矩阵的对角化
习题课

董荣
数学与统计学院

一、实对称矩阵的对角化

对称矩阵： 满足$A^T=A$或$a_{ij}=a_{ji}(\forall i,j=1,2,\cdots,n,i\neq j)$,

共扼矩阵： 称$\overline{A}=(\overline{a_{ij}})$为复矩阵$A=(a_{ij})$的共轭矩阵

矩阵的共扼运算满足：

$$\overline{A+B}=\overline{A}+\overline{B}$$

$$\overline{AB}=\overline{A}\ \overline{B}$$

$$\overline{kA}=\overline{k}\ \overline{A}$$

方阵A为实矩阵 $\Leftrightarrow \overline{A}=A$

性质6.2.1 实对称矩阵的特征值都是实数.

证明 设λ为实对称矩阵A的一个特征值且$x=(x_1,x_2,\cdots,x_n)^T$为对应的特征向量,则有 $Ax=\lambda x$

两端取共轭,得 $\overline{A}\,\overline{x}=\overline{\lambda}\,\overline{x}$, 两端取转置,得

$$\overline{x}^T A=\overline{\lambda}\,\overline{x}^T \Rightarrow \overline{x}^T A x=\overline{\lambda}\,\overline{x}^T x \Rightarrow \overline{x}^T \lambda x=\overline{\lambda}\,\overline{x}^T x$$

$$\Rightarrow (\lambda-\overline{\lambda})\overline{x}^T x=0$$

因为$\overline{x}^T x=|x_1|^2+|x_2|^2+\cdots+|x_n|^2>0$

$\Rightarrow \lambda=\overline{\lambda}$,即$\lambda$为实数.

性质6.2.2 设λ_1,λ_2是实对称矩阵的两个不同特征值.

$x_1=\begin{bmatrix}a_1\\ \vdots\\ a_n\end{bmatrix},x_2=\begin{bmatrix}b_1\\ \vdots\\ b_n\end{bmatrix}$分别为对应的特征向量.

实对称矩阵的属于互不相同特征值的特征向量**互相正交**

则x_1与x_2正交，即$\langle x_1,x_2\rangle=x_1^Tx_2=x_2^Tx_1=\sum_{i=1}^{n}a_ib_i=0$

证明 $Ax_1=\lambda_1x_1 \quad \Rightarrow x_2^TAx_1=x_2^T\lambda_1x_1=\lambda_1x_2^Tx_1=\lambda_1\langle x_1,x_2\rangle,$

$Ax_2=\lambda_2x_2 \quad \Rightarrow x_1^TAx_2=x_1^T\lambda_2x_2=\lambda_2x_1^Tx_2=\lambda_2\langle x_1,x_2\rangle,$

$x_2^TAx_1$是1×1矩阵, 故$x_2^TAx_1=(x_2^TAx_1)^T=x_1^TA^Tx_2$

因为A为对称矩阵，所以 $x_2^TAx_1=x_1^TAx_2 \quad \Rightarrow \lambda_1\langle x_1,x_2\rangle=\lambda_2\langle x_1,x_2\rangle$

$\Rightarrow(\lambda_1-\lambda_2)\langle x_1,x_2\rangle=0,$ 由于$\lambda_1\neq\lambda_2$，得$\langle x_1,x_2\rangle=0$

定理6.2.3 设A为n阶实对称矩阵，则必存在n阶正交矩阵P，使得

$$P^{-1}AP = P^{T}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)\text{为对角矩阵.}$$

$$P^{T}P = PP^{T} = I$$

证明: 用归纳法. n=1显然成立; 假设阶数为n-1时成立，下证阶数为n时成立:

设A的一个特征向量为$\alpha_1 \in R^n$，则将α_1扩充得R^n的一组标准正交基$\alpha_1, \alpha_2, \ldots, \alpha_n$，令$P_1=[\alpha_1, \alpha_2, \ldots, \alpha_n]$，则$P_1$为正交矩阵.

$AP_1 = [A\alpha_1, A\alpha_2, \cdots, A\alpha_n]$ 设向量$A\alpha_i$ (i=2,…n)在基$\alpha_1, \ldots, \alpha_n$下的坐标为$b_{1i}, \ldots, b_{ni}$

$$AP_1 = [\lambda_1\alpha_1, A\alpha_2, \cdots, A\alpha_n] = [\alpha_1, \alpha_2, \cdots, \alpha_n]\begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & b_{n2} & \cdots & b_{nn} \end{bmatrix} = P_1\begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix}$$

$$\Rightarrow A = P_1\begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix}P_1^T = A^T = P_1\begin{bmatrix} \lambda_1 & 0 \\ \beta & A_1^T \end{bmatrix}P_1^T \quad \Rightarrow P_1^T A P_1 = \begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ \beta & A_1^T \end{bmatrix}$$

定理6.2.3 设A为n阶实对称矩阵，则必存在n阶正交矩阵P，使得

$$P^{-1}AP = P^TAP = diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$$为对角矩阵.

$$\boxed{P^TP = PP^T = I}$$

$$\Rightarrow P_1^TAP_1 = \begin{bmatrix} \lambda_1 & \beta^T \\ 0 & A_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ \beta & A_1^T \end{bmatrix} \Rightarrow \beta = 0, A_1 = A_1^T \Rightarrow$$ A_1为n-1阶实对称阵

由归纳假设，存在n-1阶正交阵P_2使 $P_2^{-1}A_1P_2 = P_2^TA_1P_2 = \mathrm{diag}(\lambda_2,\cdots,\lambda_n)$

令 $P = P_1\begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix}$，则$P$为两个正交矩阵之积，故$P$为$n$阶正交阵

$$P^TAP = \begin{bmatrix} 1 & 0 \\ 0 & P_2^T \end{bmatrix} P_1^TAP_1 \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & P_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 \\ 0 & P_2^TA_1P_2 \end{bmatrix} = diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$$

定理6.2.3 设A为n阶实对称矩阵，则必存在n阶正交矩阵P，使得

$$P^{-1}AP = P^{T}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$为对角矩阵.

其中$\lambda_1, \lambda_2, \cdots, \lambda_n$为$A$的全部特征值；

矩阵P的列向量组为A的n个标准正交的特征向量.

推论6.2.3 实对称矩阵每个特征值的几何重数等于其代数重数.

利用正交矩阵将对称矩阵对角化的方法：

1. **求出A的所有特征值;**
2. 求出A的n个线性无关的特征向量;
3. 将特征向量正交化与单位化，组成P，则有

$$P^{-1}AP = P^{T}AP = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$$为对角矩阵.

例： 对于$A=\begin{bmatrix}1&2\\2&1\end{bmatrix}$,求一个正交矩阵$P$,使得$P^{-1}AP$成为对角矩阵.

解： $|\lambda I-A|=\begin{vmatrix}\lambda-1&-2\\-2&\lambda-1\end{vmatrix}=0 \Rightarrow \lambda_1=3,\lambda_2=-1.$

求$(3I-A)x=0$和$(-I-A)x=0$的基础解系,

分别求得$\lambda_1=3,\lambda_2=-1$对应的特征向量

$$\xi_1=\begin{bmatrix}1\\1\end{bmatrix},\quad \xi_2=\begin{bmatrix}1\\-1\end{bmatrix}$$

ξ_1与ξ_2已经正交,再单位化：

$$e_1=\frac{\xi_1}{\|\xi_1\|}=\begin{bmatrix}\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\end{bmatrix},e_2=\frac{\xi_2}{\|\xi_2\|}=\begin{bmatrix}\frac{1}{\sqrt{2}}\\-\frac{1}{\sqrt{2}}\end{bmatrix}$$

令$P=[e_1\quad e_2]$,

则$P^{-1}AP=\begin{bmatrix}3&\\&-1\end{bmatrix}$

二、习题课

本章基本要求

(1)理解特征值与特征向量的定义,了解其性质,会计算特征值与特征向量.

(2)了解相似矩阵的概念及性质.

(3)理解方阵可对角化的条件,掌握用相似变换化方阵为对角矩阵的方法.

(4)了解实对称矩阵的性质,掌握实对称矩阵正交相似对角化的方法.

特征值与特征向量：A为n阶方阵，λ为复数，x为n维非零列向量，若

$$Ax = \lambda x$$

则λ称为A的**特征值**，x 称为A的对应于 λ的**特征向量**.

求特征值与特征向量的一般步骤

(1)求出$|\lambda I - A| = 0$在复数范围内的全部根

$\lambda_1, \lambda_2, \cdots, \lambda_n$(重根按重数计算)，这就是$A$的全部特征值.

(2)对于A的特征值λ_i，求出齐次线性方程组

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

的一个基础解系$\xi_{i1}, \xi_{i2}, \cdots, \xi_{ik_i}$，则属于$\lambda_i$的全部特征向量为

$x = c_1\xi_{i1} + c_2\xi_{i2} + \cdots + c_{k_i}\xi_{ik_i}$，$(c_1, \cdots, c_{k_i}$为不全为零的任意常数)

例　若存在某正整数k，使得方阵A满足$A^k=O$，证明：A的特征值都为0.

幂零矩阵

证:　设λ是A的特征值，α是对应的特征向量，且$A\alpha=\lambda\alpha$，

则$A^k\alpha=\lambda^k\alpha=0$，

而特征向量$\alpha\neq 0$，

$\lambda^k=0\Rightarrow\lambda=0$.

注:　零矩阵O的特征值都为0

幂零矩阵的例子：$M=\begin{bmatrix}5&-3&2\\15&-9&6\\10&-6&4\end{bmatrix}$满足$M^2=O$

习题6.1(A)17

若矩阵 $A_{n\times n}$ 满足 $A^2=A$, 证明: A的特征值必为0或1.

幂等矩阵

证: 设λ是A的特征值, α是对应的特征向量, 且 $A\alpha=\lambda\alpha$,

则$A^2\alpha=\lambda^2\alpha$,

由 $A^2=A$得 $(\lambda^2-\lambda)\alpha=0$, 而特征向量$\alpha\neq 0$,

$\lambda^2-\lambda=0\Rightarrow\lambda=0$ 或 1.

注: 虽然 A的特征值必为0或1, 但是0,1未必都能取到.

例如$\begin{pmatrix}1&0\\0&1\end{pmatrix}$, $\begin{pmatrix}0&0\\0&0\end{pmatrix}$, $\begin{pmatrix}1&0\\0&0\end{pmatrix}$.

习题6.1(A)19

设A,B均为n阶矩阵, 证明: AB与BA有相同的特征值.

证: (1) 设$\lambda\neq 0$是AB的特征值, α是对应特征向量, 即$(AB)\alpha=\lambda\alpha$,

则$B(AB)\alpha=\lambda B\alpha$, 即$(BA)(B\alpha)=\lambda(B\alpha)$, 而$B\alpha\neq 0$

(否则$A(B\alpha)=0$,与$(AB)\alpha=\lambda\alpha\neq 0$矛盾)

故λ也是BA的特征值, 相应的特征向量为$B\alpha$.

(2) 设$\lambda=0$是AB的一个特征值, 则有$\det(AB)=0$, 从而

$\det(BA)=\det(AB)=0$, 故0也是BA的特征值.

总之, AB与BA有相同的特征值.

习题6.1(B)1： 设任何n维非零列向量都是n阶矩阵$\boldsymbol{A}$的特征向量，证明：$\boldsymbol{A}$为数量矩阵（即存在常数k，使得$\boldsymbol{A}=k\boldsymbol{I}$）

证： 由于任何n维非零列向量都是$\boldsymbol{A}$的特征向量，则$\boldsymbol{\varepsilon_1}=[1,0\cdots,0]^T$，$\boldsymbol{\varepsilon_2}=[0,1\cdots,0]^T,\cdots,\boldsymbol{\varepsilon_n}=[0,0\cdots,1]^T$是$\boldsymbol{A}$的特征向量，有$\boldsymbol{A\varepsilon_i}=\lambda_i\boldsymbol{\varepsilon_i}$

从而有 $\boldsymbol{A}[\boldsymbol{\varepsilon_1},\cdots,\boldsymbol{\varepsilon_n}]=[\boldsymbol{A\varepsilon_1},\cdots,\boldsymbol{A\varepsilon_n}]=[\lambda_1\boldsymbol{\varepsilon_1},\cdots,\lambda_n\boldsymbol{\varepsilon_n}]$

因为$[\boldsymbol{\varepsilon_1},\cdots,\boldsymbol{\varepsilon_n}]=\boldsymbol{I}$, 所以$\boldsymbol{A}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$

$\boldsymbol{\varepsilon_i}+\boldsymbol{\varepsilon_j}$也是$\boldsymbol{A}$的特征向量，设其特征值为$\lambda$，我们有

$$\lambda_i\boldsymbol{\varepsilon_i}+\lambda_j\boldsymbol{\varepsilon_j}=\boldsymbol{A}(\boldsymbol{\varepsilon_i}+\boldsymbol{\varepsilon_j})=\lambda\left(\boldsymbol{\varepsilon_i}+\boldsymbol{\varepsilon_j}\right)=\lambda\boldsymbol{\varepsilon_i}+\lambda\boldsymbol{\varepsilon_j}$$

$$\Rightarrow(\lambda_i-\lambda)\boldsymbol{\varepsilon_i}=\left(\lambda-\lambda_j\right)\boldsymbol{\varepsilon_j}$$

进而可得$\lambda_i=\lambda_j=\lambda,(i,j=1,2,\cdots,n)$，从而

$$\boldsymbol{A}=\lambda\boldsymbol{I}$$

习题6.1(B)2： 设λ为正交矩阵$\boldsymbol{A}$的特征值．证明：$1/\lambda$也是$\boldsymbol{A}$的特征值．

证： 由于λ为正交矩阵$\boldsymbol{A}$的特征值，从而存在特征向量$\boldsymbol{x}$，使得

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$$

用$\boldsymbol{A}^T$同时左乘上式两端，得

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{A}^T\boldsymbol{x}$$

从而有

$$\frac{1}{\lambda}\boldsymbol{x} = \boldsymbol{A}^T\boldsymbol{x}$$

即$1/\lambda$是$\boldsymbol{A}^T$的特征值。

由于$\det(\lambda\boldsymbol{I}-\boldsymbol{A}) = \det((\lambda\boldsymbol{I}-\boldsymbol{A})^T) = \det(\lambda\boldsymbol{I}-\boldsymbol{A}^T)$，我们得知$\boldsymbol{A}, \boldsymbol{A}^T$具有相同的特征值，从而$1/\lambda$是$\boldsymbol{A}$的特征值.

习题6.2(A)4 已知3阶矩阵A与B相似，A的特征值为$\frac{1}{2}$，$\frac{1}{3}$，$\frac{1}{4}$，则行列式$|B^{-1}-I|=?$

解: A与B相似，故B的特征值也为$\frac{1}{2}$，$\frac{1}{3}$，$\frac{1}{4}$，则B^{-1}的特征值为$2, 3, 4$

现在来看$B^{-1}-I$的特征值，其特征方程为

$$|\lambda I-(B^{-1}-I)|=0$$

$$|(\lambda+1)I-B^{-1}|=0$$

显然有当$\lambda+1=2,3,4$时，上述等式成立，故$B^{-1}-I$的特征值为$1,2,3$

因此$|B^{-1}-I|=1\times 2\times 3=6$

由此题可得结论： 若λ是A的特征值，则$\lambda+k$是$A+kI$的特征值.

$$|\lambda I-A|=|(\lambda+k)I-(A+kI)|=0$$