§2 特征值与特征向量

一、特征值和特征向量的概念

定义1 $A\in C^{n\times n}$,如果存在$\lambda\in C$和非零向量$x\in C^{n}$, 使

$$Ax=\lambda x,$$

则λ叫做A的特征值, x叫做A的属于特征值λ的特征向量.

(1) 矩阵的谱： A的所有特征值的全体, 叫做A的谱

记为$\lambda(A)$.

返回

(2) 特征多项式：

$$|\lambda E - A| = (\lambda-\lambda_1)^{n_1}\cdots(\lambda-\lambda_i)^{n_i}\cdots(\lambda-\lambda_r)^{n_r}\text{（其中}\sum_{i=1}^{r} n_i = n.)$$

称为特征多项式

(3)代数重数： n_i叫做λ_i的 代数重数

(4)几何重数：

（I） 若$\operatorname{rank}(\lambda_i E - A) = n - m_i$， m_i叫做λ_i的几何重数.

(II) $W = \{x \mid (\lambda_i E - A)x = 0\}$

$m_i = \dim W = n - \operatorname{rank}(\lambda_i E - A)$ 叫做λ_i的 几何重数.

(5) 特征值的几何重数不超过代数重数 ($m_i \le n_i$)

(6) 不同特征值对应的特征向量线性无关.

注: 1.特征值的几何重数不超过代数重数

$V_{\lambda_0}=\{x\,|\,(\lambda_0 E-A)x=0\}$, $\dim V_{\lambda_0}=k$ 且 $x_1,x_2,\cdots,x_k$ 为基

V_{λ_0} 的基: $x_1,x_2,\cdots,x_k \xrightarrow{\text{扩充}} C^n$ 的基: $x_1,x_2,\cdots,x_k,x_{k+1},\cdots,x_n$

$$C=(x_1,\cdots,x_k,x_{k+1},\cdots,x_n)\Rightarrow E_n=C^{-1}C=C^{-1}(x_1,\cdots,x_k,x_{k+1},\cdots,x_n)$$

$$\Rightarrow C^{-1}x_k=e_k=(0,\cdots,0,\underset{k}{1},0,\cdots,0)^T\ (k=1,\cdots,n)\ \Rightarrow$$

$$C^{-1}AC=C^{-1}A(x_1,\cdots,x_k,x_{k+1},\cdots,x_n)=C^{-1}(Ax_1,\cdots,Ax_k,Ax_{k+1},\cdots,Ax_n)$$

$$=C^{-1}(\lambda_0x_1,\cdots,\lambda_0x_k,Ax_{k+1},\cdots,Ax_n)$$

$$=(\lambda_0C^{-1}x_1,\cdots,\lambda_0C^{-1}x_k,C^{-1}Ax_{k+1},\cdots,C^{-1}Ax_n)$$

$$=(\lambda_0e_1,\cdots,\lambda_0e_k,C^{-1}Ax_{k+1},\cdots,C^{-1}Ax_n)$$

$$C^{-1}AC=\left(\begin{array}{ccc|c}\lambda_0 & & & * \\ & \ddots & & \\ & & \lambda_0 & \\ \hline & 0 & & A_0\end{array}\right)=B \quad \Rightarrow |\lambda E-A|=|\lambda E-B|=$$

$$\left|\begin{array}{ccc|c}\lambda-\lambda_0 & & & * \\ & \ddots & & \\ & & \lambda-\lambda_0 & \\ \hline & 0 & & \lambda E_{n-k}-A_0\end{array}\right|=(\lambda-\lambda_0)^k\,|\lambda E_{n-k}-A_0|$$

$\Rightarrow$ λ_0 至少为 k 重特征值 $\Leftrightarrow$

代数重数$\geq k=\dim V_{\lambda_0}$ ----几何重数

注：2. 不同特征值对应的特征向量线性无关

$\lambda_i (i=1,2,\cdots,r)$---------互异特征值

$x_1^{(i)}, x_2^{(i)}, \cdots, x_{s_i}^{(i)}$----------$\lambda_i$对应的线性无关特征向量$(i=1,2,\cdots,r)$

$$\Rightarrow \quad \underbrace{x_1^{(1)}, x_2^{(1)}, \cdots, x_{s_1}^{(1)}}\ \underbrace{x_1^{(2)}, x_2^{(2)}, \cdots, x_{s_2}^{(2)}} \cdots \underbrace{x_1^{(r)}, x_2^{(r)}, \cdots, x_{s_r}^{(r)}}$$ 线性无关

证：

$$\underbrace{k_{11}x_1^{(1)} + \cdots + k_{1s_1}x_{s_1}^{(1)}} + \underbrace{k_{21}x_1^{(2)} + \cdots + k_{2s_2}x_{s_2}^{(2)}} + \cdots + k_{r1}\underbrace{x_1^{(r)} + \cdots + k_{rs_r}x_{s_r}^{(r)}} = 0$$

$$\Rightarrow$$

$$(\lambda_1 E - A)(\underbrace{k_{11}x_1^{(1)} + \cdots + k_{1s_1}x_{s_1}^{(1)}} + \underbrace{k_{21}x_1^{(2)} + \cdots + k_{2s}x_{s_2}^{(2)}} + \cdots + \underbrace{k_{r1}x_1^{(r)} + \cdots + k_{rs_r}x_{s_r}^{(r)}}) = 0$$

$$(\lambda_1 - \lambda_2)(\underbrace{k_{21}x_1^{(2)} + \cdots + k_{2s}x_{s_2}^{(2)}}) + \cdots + (\lambda_1 - \lambda_r)(\underbrace{k_{r1}x_1^{(r)} + \cdots + k_{rs_r}x_{s_r}^{(r)}}) = 0$$

返回

$\Rightarrow$

$$(\lambda_2 E - A)[(\lambda_1-\lambda_2)(\underbrace{k_{21}x_1^{(2)}+\cdots+k_{2s_2}x_{s_2}^{(2)}})+\cdots+(\lambda_1-\lambda_r)(\underbrace{k_{r1}x_1^{(r)}+\cdots+k_{rs_r}x_{s_r}^{(r)}})]=0$$

$$(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)(\underbrace{k_{31}x_1^{(3)}+\cdots+k_{3s_3}x_{s_3}^{(3)}})+\cdots+(\lambda_1-\lambda_r)(\lambda_2-\lambda_r)(\underbrace{k_{r1}x_1^{(r)}+\cdots+k_{rs_r}x_{s_r}^{(r)}})]=0$$

…………

$$\Rightarrow \quad (\lambda_1-\lambda_r)(\lambda_2-\lambda_r)\cdots(\lambda_{r-1}-\lambda_r)(\underbrace{k_{r1}x_1^{(r)}+\cdots+k_{rs_r}x_{s_r}^{(r)}})=0$$

$$\Rightarrow \quad k_{r1}=\cdots=k_{rs_r}=0 \Rightarrow \quad \cdots\cdots\cdots \Rightarrow$$

$$k_{11}=\cdots=k_{1s_1}=k_{21}=\cdots=k_{2s_2}=\cdots=k_{r1}=\cdots=k_{rs_r}=0$$

$$\Rightarrow \underbrace{x_1^{(1)},x_2^{(1)},\cdots,x_{s_1}^{(1)}}\ \underbrace{x_1^{(2)},x_2^{(2)},\cdots,x_{s_2}^{(2)}}\cdots\underbrace{x_1^{(r)},x_2^{(r)},\cdots,x_{s_r}^{(r)}}$$ 线性无关

定义2 设$A\in C^{n\times n}$， 如果存在可逆矩阵 $P\in C^{n\times n}$，使得

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

则矩阵A 叫做 可对角化矩阵.

注：n阶矩阵A可对角化$\Leftrightarrow A$有n个线性无关特征向量

定理1 设矩阵 $A\in C^{n\times n}$，则存在可逆矩阵 $P\in C^{n\times n}$， 使得

$$P^{-1}AP=J=diag(J(\lambda_1),J(\lambda_2),\cdots,J(\lambda_r))$$

(1) J 称为A的Jordan标准形;

(2) 其中 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 不一定不相同

Jordan矩阵的结构与几个结论:

(1) Jordan块的个数 r是线性无关特征向量的个数;

(2) 矩阵可对角化,当且仅当$r=n$;

(3) (I): 对应于同一特征值的Jordan块的个数是该特征值的几何重数, 它是相应的特征子空间的维数.

(II):对应于同一特征值的所有Jordan块的阶数之和是该特征值的代数重数.

定理 2 设$A\in C^{n\times n}$，则下列命题等价：

(1) A 是可对角化矩阵；

(2) C^n 存在由A的特征值向量构成的一组基底。

(3) A 的Jordan标准形中的 Jordan块都是一阶的。

(4) $m_i = n_i \quad (i=1,\ 2,\cdots,r)$

二. 特征值与特征向量的几何性质

1. 变换　　　　V---线性空间

$$V \xrightarrow{T} V$$

$$\forall \alpha \in V, \alpha \xrightarrow{T} \alpha' \in V$$

（V中任一元素α，都有V中唯一确定的元素α'与之对应）.

则称 T为V的变换

2. 线性变换

T为V的变换且满足 $\begin{cases} <1> \quad \forall \ \alpha,\beta \in V, T(\alpha+\beta)=T(\alpha)+T(\beta) \\ <2> \quad \forall \ k \in P, \qquad T(k\alpha)=kT(\alpha) \end{cases}$

则称 T为V的线性变换

例:在线性空间 P_n 中 ,求微分是一线性变换 , 即

$$Df(t)=f'(t), \forall f(t) \in P_n$$

3. 线性变换的特征值

定义 $1'$ 设T是线性空间$V_n(C)$的一个线性变换，如果存在$\lambda \in C$和非零向量$\xi \in V_n(C)$，使得$T\xi=\lambda\xi$，则λ 叫做T的特征值，ξ 叫做T的属于特征值λ 的特征向量。

返回

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$$

4. 线性变换与矩阵 $(T \text{-------------}\to A)$

$$1 \leftrightarrow 1$$

V---n维线性空间, $\varepsilon_1, \varepsilon_1, \cdots, \varepsilon_n$ 为基, T---V上的线性变换

$$T\varepsilon_1 = \sum_{i=1}^{n} a_{i1}\varepsilon_i, T\varepsilon_2 = \sum_{i=1}^{n} a_{i2}\varepsilon_i, \cdots, T\varepsilon_n == \sum_{i=1}^{n} a_{in}\varepsilon_i$$

则有

$$T(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) = (T\varepsilon_1, T\varepsilon_2, \cdots, T\varepsilon_n)$$

$$= (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)A$$

A: 线性变换T在基 $\varepsilon_1, \varepsilon_1, \cdots, \varepsilon_n$ 下的矩阵.

$$\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$$

$$T \text{-----------}\rightarrow A$$

$$1 \leftrightarrow 1$$

5. 线性变换与矩阵特征值关系 $\left(T\alpha = \lambda\alpha \xleftrightarrow{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n} Ax = \lambda x\right)$

$$\alpha = \sum_{i=1}^{n} x_i \varepsilon_i, \quad x = \left(x_1, x_2, \cdots, x_n\right)^T, \quad T\alpha = \lambda\alpha$$

$$\left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)\lambda x = \sum_{i=1}^{n} \lambda x_i \varepsilon_i = \lambda \sum_{i=1}^{n} x_i \varepsilon_i = \lambda\alpha = T\alpha$$

$$= T\sum_{i=1}^{n} x_i \varepsilon_i = \sum_{i=1}^{n} x_i T\varepsilon_i = \left(T\varepsilon_1, T\varepsilon_2, \cdots, T\varepsilon_n\right)x = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)Ax$$

即得 $\lambda x = Ax$

定理3:

$$\begin{cases} \xi \xleftrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n} x=(x_1,x_2,\cdots,x_n)^T \\ \eta \xleftrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n} y=(y_1,y_2,\cdots,y_n)^T \\ T \xleftrightarrow{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n} A \end{cases} \qquad \eta=T\xi \Leftrightarrow y=Ax.$$

$$(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)y=\eta=T\xi=T(\sum_{i=1}^{n}x_i\varepsilon_i)$$

$$=\sum_{i=1}^{n}x_iT\varepsilon_i$$

$$=(T\varepsilon_1,T\varepsilon_2,\cdots,T\varepsilon_n)x$$

$$=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)Ax$$

$$\Rightarrow y=Ax.$$

6. 线性变换在不同基下矩阵之间的关系

定理4: $T \xrightarrow{(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)} A,\quad T \xrightarrow{(\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n')} B,$

$$(\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n') = (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)C$$

$$\Downarrow$$

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

证明： $T(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A,$

$$T(\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n')=(\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n')B,$$

$$(\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n')=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)C$$

$$\begin{aligned}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)CB &= (\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n')B = T(\varepsilon_1',\varepsilon_2',\cdots,\varepsilon_n')\\ &= T((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)C)\\ &= (T(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n))C\\ &= ((\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A)C\\ &= (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)AC\end{aligned}$$

$$\Rightarrow (\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)(CB-AC)=0$$

$$\Rightarrow CB-AC=0 \qquad \Rightarrow B=C^{-1}AC$$

$$\Rightarrow AC=CB$$

返回

三、广义特征值问题

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

$$Ax = \lambda Bx \quad (1\text{-}3)$$

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。

2. 广义特征值问题的简化

(1) 当B可逆时，式(1-3)可化为 $B^{-1}Ax=\lambda x$ (1-4)

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

$$Qy=\lambda y(\text{其中}Q=\left(P^{-1}\right)^H AP^{-1}, y=Px, B=P^H P)$$

证明：$B=B^H$且正定 $\xrightarrow{\text{存在可逆矩阵P}}$ $B=P^H P$ $\xrightarrow{\text{则(1-3)式化为}}$

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

$$Qy=\lambda y$$

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

$$y_i^H y_j=\delta_{ij} \xrightarrow{y_i=Px_i} y_i^H y_j=(Px_i)^H(Px_j)=x_i^H P^H Px_j=x_i^H Bx_j$$

$$\Longrightarrow \quad x_i^H Bx_j=\delta_{ij}=\begin{cases}1, & j=i\\ 0, & j\neq i\end{cases}$$

◀ ▶ 返回

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

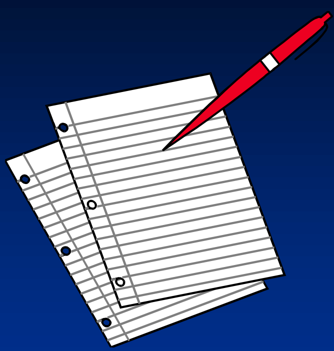
（1）$x_i \neq 0 \ (i = 1,\ 2,\cdots,n)$；

（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$X = (x_1,\ x_2,\cdots,x_n)$，

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



返回

四．二次特征值问题

对于一个震动系统，当衰减效应被引入模型时，结构的固有模式和频率为二次特征值问题的解，二次特征值问题由如下公式给出：

$$Mx'' + Cx' + Kx = 0, \qquad x = e^{\lambda t}u$$

$$(\lambda^2 M + \lambda C + K)u = 0 \quad (1) \quad 或 \quad v^H(\lambda^2 M + \lambda C + K) = 0 \quad (2)$$

$$令P = \begin{pmatrix} -C & -K \\ E & 0 \end{pmatrix}, \quad G = \begin{pmatrix} M & \\ & E \end{pmatrix}, \quad z = \begin{pmatrix} \lambda u \\ u \end{pmatrix}, 则(1)可转化为:$$

$$Pz = \lambda Gz.$$

当外力的频率和结构的固有频率非常接近时就会引起共振，从而会破坏系统的结构，使得系统不稳定．因此，确定一个系统的固有频率尤为重要．

返回

On its opening day in June 2000, the 320-meter-long Millennium footbridge over the river Thames in London (see Figure) started to wobble alarmingly under the weight of thousands of people; two days later the bridge was closed.

Fig. ***The Millennium footbridge over the river Thames.***

推荐参考书：

(1) 《矩阵论》，方保镕、周继东、李医民编著，清华大学出版社

(2) 《Matrix Analysis》，Roger A. Horn, Charles. – R. Johnson，人民邮电出版社

(3) 《矩阵分析与应用》，张贤达著，清华大学出版社

返回