第一章 线性代数基础

1.5 特征值与特征向量

定义1（特征值与特征向量） 设V是F上的线性空间，$T \in L(V)$,若存在$\lambda \in F$及V的非零向量ξ，使得$T\xi = \lambda\xi$，则称λ为T的一个**特征值**，而ξ为T的属于特征值λ的一个**特征向量**.

注：特征向量在线性变换作用下**保持方位不变**(在同一直线上).

设$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$的一组基, T在这组基下的矩阵为$A=(a_{ij})\in F^{n\times n}$,即

$$T(\alpha_1,\ldots,\alpha_n)=(\alpha_1,\ldots,\alpha_n)A$$

设λ_0是T的一个特征值，ξ为T的属于特征值λ_0的一个特征向量, 且$\xi=(\alpha_1,\ldots,\alpha_n)X,\ X\in F^n$,则

$$T\xi=\lambda_0\xi\Rightarrow T(\alpha_1,\ldots,\alpha_n)X=\lambda_0(\alpha_1,\ldots,\alpha_n)X$$

即：$(\alpha_1,\ldots,\alpha_n)AX=\lambda_0(\alpha_1,\ldots,\alpha_n)X\Rightarrow AX=\lambda_0X$

$\Rightarrow(A-\lambda_0I)X=0$（$X$非零向量）

所以行列式$|A-\lambda_0I|=0$.

定义2（特征多项式）设$A=\left(a_{ij}\right)\in F^{n\times n}$,$\lambda$是一个参量, 矩阵$\lambda I-A$称为$A$的特征矩阵. 行列式

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

称为A的特征多项式.展开式是一个n次多项式，其根为A的特征值，而相应于$(\lambda_0 I-A)X=0$的非零解向量X称为A的属于λ的特征向量.

注：(1)λ_0是T的特征值$\Leftrightarrow$ λ_0是A的特征值.

(2)ξ是T的特征向量$\Leftrightarrow$ X是A的特征向量，这里，$\xi=(\alpha_1,\ldots,\alpha_n)X$.

(3) A的属于λ_0的全部特征向量再添上零向量就构成了F^n的一个线性子空间，称为A的一个特征子空间，记为$E(\lambda_0)$，它就是齐次线性方程组（$\lambda_0 I-A$）$X=0$的解空间.

求矩阵A的全部特征值及特征向量的步骤：

1）计算行列式$|\lambda I - A|$；

2）求出多项式$f(\lambda) = |\lambda I - A|$在数域$F$中的全部根（即$A$的特征值）；

3）对A的每个特征值λ，解齐次线性方程组$(\lambda I - A)\mathrm{X} = 0$，求出它的一组基础解系$X_1, \ldots, X_t$，则$A$的属于$\lambda$的全部特征向量为$k_1X_1 + \cdots + k_tX_t$, k_i不全为零.

4）T的属于λ的特征向量为$\xi_i=(\alpha_1,\ldots,\alpha_n)X_i$,$(1\le i\le \mathrm{t})$，从属于$\lambda$而全部特征向量为$k_1\xi_1+\cdots+k_t\xi_t$, k_i不全为零.

注：(1) 特征值与特征向量是否存在**依赖于**V所在的数域F,比如$\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$的特征多项式为$f(\lambda)=\lambda^2+1$

(2)当$\mathrm{dimV}=\mathrm{n}$很大时，上述求法太繁琐，可借助于计算机.

(3)$\mathrm{E}（\lambda_i）=\{\mathrm{X}\in \mathrm{F}^n|(\lambda_i\mathrm{I}-\mathrm{A})\mathrm{X}=0\}=\mathrm{N}(\lambda_i\mathrm{I}-\mathrm{A})$ (解空间).由亏加秩定理有$\mathrm{r}(\lambda_i\mathrm{I}-\mathrm{A})+\mathrm{dimN}(\lambda_i\mathrm{I}-\mathrm{A})=n$，所以**$\mathrm{E}（\lambda_i）$的维数**为$\mathrm{dimE}(\lambda_i)=n-\mathrm{r}(\lambda_i\mathrm{I}-\mathrm{A})$称为**$\lambda_i$的几何重数**.

例1 设$T\in L(R^3)$在基$\alpha_1,\alpha_2,\alpha_3$下矩阵为

$$A=\begin{bmatrix}2&-2&2\\-2&-1&4\\2&4&-1\end{bmatrix}$$,求T的全部特征值与全部特征向量.

解:$f(\lambda)=|\lambda I-A|=(\lambda-3)^2(\lambda+6)$,解得

$$\lambda_1=3,(\text{二重})\lambda_2=-6.$$

对于$\lambda_1=3$,求解齐次方程组$(3I-A)X=0$得基础解系$X_1=(-2,1,0)^T$,$X_2=(2,0,1)^T$.所以T的属于$\lambda_1=3$的特征向量为$\xi_1=(\alpha_1,\alpha_2,\alpha_3)X_1=-2\alpha_1+\alpha_2$.

$\xi_2 = (\alpha_1, \alpha_2,\ \alpha_3)X_2 = 2\alpha_1 + \alpha_3$.所以$T$的属于$\lambda_1 = 3$的全部特征向量为$k_1\xi_1 + k_2\xi_2$, k_i不同时为零.

对于$\lambda_2 =- 6$，求解齐次方程组$(-6I - A)\mathrm{X} = 0$得基础解系X_3=(1，2，$-$2)^T.所以T的属于$\lambda_2 =- 6$的特征向量为$\xi_3 = (\alpha_1, \alpha_2,\ \alpha_3)X_3 = \alpha_1 + 2\alpha_1 - 2\alpha_3$,所以$T$的属于$\lambda_2 =- 6$的全部特征向量为$k_3\xi_3$, k_3不为零.

定理3 相似矩阵有相同的特征多项式及特征值，反之不然.

证明：设$P^{-1}AP=B$，则

$$|\lambda\mathrm{I}-\mathrm{B}|=|\lambda\mathrm{I}-P^{-1}AP|=|P^{-1}||\lambda\mathrm{I}-A||P|$$
$$=|\lambda\mathrm{I}-A|.$$

反例：$A=\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$,$I_2=\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$,$A$与$I_2$的特征多项式均为$(\lambda-1)^2$，但它们不相似.

注： 定理表明，线性变换的矩阵的特征多项式与**基的选取无关**，而直接由线性变换决定，故可称之为**线性变换的特征多项式**.

A的特征多项式$f(\lambda)=|\lambda I-A|$是一个首1的多项式，其$n-1$次系数是$-(\sum_{i=1}^{n}\lambda_i)=-\sum_{i}^{n}a_{ii}=-trA$(A的迹);常数项为$-1^n|A|=-1^n\prod_{i=1}^{n}\lambda_i$.

定理4设$A=(a_{ij})\in F^{m\times n}$, $B=(b_{ij})\in F^{n\times m}$,

则$\mathrm{tr}(AB)=\mathrm{tr}(BA)$.

证明：令$AB=(c_{ij})\in F^{m\times m}$, $BA=(d_{ij})\in F^{n\times n}$.

所以$c_{ij}=\sum_{k=1}^{n}a_{ik}b_{kj}$, $d_{ij}=\sum_{l=1}^{m}b_{il}a_{lj}$.

$$\mathrm{tr}(AB)=\sum_{i=1}^{m}c_{ii}=\sum_{i=1}^{m}\sum_{k=1}^{n}a_{ik}b_{ki}$$

$$\mathrm{tr}(BA)=\sum_{j=1}^{n}d_{jj}=\sum_{j=1}^{n}\sum_{l=1}^{m}b_{jl}a_{lj}=$$

$$\sum_{l=1}^{m}\sum_{j=1}^{n}a_{lj}b_{jl}.$$ 结论成立.

定理5 相似的矩阵有相同的迹.

例2（换位公式）设$A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$,

则$\lambda^n|\lambda I_m - AB| = \lambda^m|\lambda I_n - BA|$.

例2（换位公式）设$A \in \mathbb{C}^{m\times n}$, $B \in \mathbb{C}^{n\times m}$,

则$\lambda^n|\lambda I_m - AB| = \lambda^m|\lambda I_n - BA|$.

证明：任意验证

$$\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix}\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix}$$

因为$\begin{bmatrix} I_m & -A \\ 0 & I_n \end{bmatrix}$可逆,所以$\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}$与$\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$相似.

故 $\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}$与$\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$特征多项式相等.

则：$\left|\lambda I_{m+n} - \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}\right| = \left|\lambda I_{m+n} - \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}\right|$,

即：$\begin{vmatrix} \lambda I_m - AB & 0 \\ -B & \lambda I_n \end{vmatrix} = \begin{vmatrix} \lambda I_m & 0 \\ -B & \lambda I_n - AB \end{vmatrix}$

从而$\lambda^n|\lambda I_m - AB| = \lambda^m|\lambda I_n - BA|$.

注：例题表明AB与BA有相同的非零特征值,故$\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

定理5(Schur引理) 任意的$A \in \mathbb{C}^{n\times n}$,都相似于一个上三角阵，即存在满秩阵$P$，使$P^{-1}AP$为上三角阵，其对角上元为A的全部特征值.

推论：设A的n个特征值为$\lambda_1, \cdots, \lambda_n$, $\varphi(x)$为任一多项式，则矩阵多项式$\varphi(x)$的n个特征值为$\varphi(\lambda_1), \cdots, \varphi(\lambda_n)$,

特别的，kA的特征值为$k\lambda_1, \cdots, k\lambda_n$，$A^m$特征值为${\lambda_1}^n, \cdots, {\lambda_n}^m$.

定理6(Hamilton-Cayley) 任意的$A \in \mathbb{C}^{n\times n}$,其特征多项式为$f(\lambda) = |\lambda \mathrm{I} - A|$，则必有$f(A) = 0$(零矩阵).

定理6(Hamilton-Cayley) 任意的$A \in \mathbb{C}^{n\times n}$,其特征多项式为$f(\lambda) = |\lambda \mathrm{I} - A|$，则必有$f(A) = 0$(零矩阵).

证明：设$f(\lambda) = (\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$,以及,

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = A_1(\text{上三角}).$$

则$P^{-1}f(A)P = f(P^{-1}AP) = (A_1 - \lambda_1 I)\cdots(A_1 - \lambda_n I) = 0$,故$f(A) = 0$.

注：(1)由于$\mathrm{L}(V,V)\cong F^{n\times n}$，故对于线性变换$T$有平行的结果：$T\in \mathrm{L}(V,V)$，且$f(\lambda)$为$T$的特征多项式，则$f(T)$为零变换.

(2)：Cayley定理对于一般数域F的矩阵也成立.

由**Hamilton-Cayley定理**可知，任取矩阵A，必有可使其零化的多项式，引入:

定义3(最小多项式) 设$A \in F^{n\times n}$, 使A零化的最小次数的首1多项式称为A的最小多项式，记为 $m_A(\lambda)$.

注：$m_A(\lambda)$是唯一的，且可整除任一A的零化多项式特别地，有$m_A(\lambda)|f(\lambda)$.

定理7 A的特征多项式$f(\lambda)$与最小多项式$m_A(\lambda)$有相同的根(不计重数).

证明：由 $m_A(\lambda)|f(\lambda)$知最小多项式$m_A(\lambda)$的根是$f(\lambda)$的特征根.

反之，若$f(\lambda_0)=0$，设ξ是属于λ_0的特征向量，即$A\xi=\lambda_0\xi$，从而$0=m_A(A)\xi=m_A(\lambda_0)\xi$,因$\xi\neq 0$，故$m_A(\lambda_0)=0$.

例3求矩阵$A=\begin{bmatrix} 7 & 4 & -1 \\ 4 & 7 & -1 \\ -4 & -4 & 4 \end{bmatrix}$的最小多项式.

例3求矩阵$A=\begin{bmatrix}7 & 4 & -1\\ 4 & 7 & -1\\ -4 & -4 & 4\end{bmatrix}$的最小多项式.

解： $f(\lambda)=|\lambda \mathrm{I}-A|=(\lambda-3)^2(\lambda-12)$,

由定理7,$m_A(\lambda)$与$f(\lambda)$有相同的根,所以

$m_A(\lambda)=f(\lambda)$或$m_A(\lambda)=(\lambda-3)(\lambda-12)$,

经验证$(A-3I)(A-12I)=0$,所以

$m_A(\lambda)=(\lambda-3)(\lambda-12)$.

定理8 矩阵A的属于不同特征值的特征向量线性无关.

证明：设$\lambda_1,\cdots,\lambda_k$是$A$的$k$个互异特征值, $\xi_1,\cdots,\xi_k$为它们对应的特征向量, 即$A\xi_i=\lambda_i\xi_i,(\xi_i\neq 0)$.令

$$c_1\xi_1+c_2\xi_2+\cdots+c_k\xi_k=0,$$

将A左乘$k-1$次，得

$$\begin{cases} c_1\xi_1+c_2\xi_2+\cdots+c_k\xi_k=0 \\ \lambda_1c_1\xi_1+\lambda_2c_2\xi_2+\cdots+\lambda_kc_k\xi_k=0 \\ \vdots \\ {\lambda_1}^{k-1}c_1\xi_1+{\lambda_2}^{k-1}c_2\xi_2+\cdots+{\lambda_k}^{k-1}c_k\xi_k=0 \end{cases}$$

把$c_i\xi_i$看成未知变量，系数行列式为

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \vdots & \vdots & \ddots & \vdots \\ {\lambda_1}^{k-1} & {\lambda_2}^{k-1} & \cdots & {\lambda_k}^{k-1} \end{vmatrix} = \prod_{1\le i<j\le k}\left(\lambda_j-\lambda_i\right) \neq 0.$$

所以方程组只有零解$c_i\xi_i=0$,而$\xi_i\neq 0 \Rightarrow c_i=0$,其中$1\le i\le k$.