5.5 特征向量与相似矩阵

定义5.5.1 设A,B是数域F上的两个n阶方阵。如果存在F上的n阶可逆方阵P使$B=P^{-1}AP$，就称A,B在F上相似。

定理5.5.1 设$A,B\in F^{n\mathrm{x}n}$相似当且仅当它们是F上的同一n维空间V的同一线性变换在两组基下的矩阵。

引理5.5.2 方阵之间的相似关系满足下列性质:

(1)自反性　任意$A,B\in F^{n\mathrm{x}n}$与自身相似;

(2)对称性　如果$F^{n\mathrm{x}n}$中A与B相似,则B与A相似;

(3)传递性　设$A,B,C\in F^{n\mathrm{x}n}$,且$A$与$B$相似, B与C相似，则A与C相似。

定义5.5.2：若在线性空间V中选一组基$\{\beta_1,\ldots,\beta_n\}$使线性变换$\underline{A}$在这组基下的矩阵$B$是对角矩阵，就称$\underline{A}$可对角化。如果方阵$A$相似于某个对角矩阵$B$就称$A$可相似对角化。

定义5.5.3 设$\underline{A}:V\to V$是线性变换，如果非零向量$\beta\in V$被$\underline{A}$映到它的某个倍向量,即$\underline{A}(\beta)=\lambda\beta$对某个$\lambda\in$F成立，就称$\lambda$是$\underline{A}$的特征值，$\beta$是$\underline{A}$的属于特征值$\lambda$的特征向量。

设$A\in F^{n\times n}$，则$V=F^{n\times 1}$上的线性变换$\underline{A}:X\to AX$的特征值λ和特征向量X称为A的特征值和特征向量。即如果$\lambda\in F$，$0\neq X\in F^{n\times 1}$满足$AX=\lambda X$，就称λ是A的特征值，X是属于特征值λ的特征向量。

定理5.5.3 线性变换 A:$V\rightarrow V$可对角化的充要条件是：存在A的一组特征向量$\beta_1,\ldots,\beta_n$ 组成V的一组基。

n阶复方阵A相似于对角矩阵当且仅当A有n个线性无关的特征向量$X_1,\ldots,X_n$。

从几何上看，矩阵A的一个特征向量X 经过作用后得到的向量AX与特征向量X是共线的，而比例系数 λ 就是特征向量X所属的特征值。

对于数域$\boldsymbol{P}$上给定的n阶方阵$\boldsymbol{A}$，它可能有多个特征值，也可能没有特征值. 如果A有特征值λ，那么$\boldsymbol{A}$的属于λ的特征向量有多少呢?

定理5.5.4 若$\alpha_1, \alpha_2, \ldots, \alpha_s$是$A$的属于$\lambda$的特征向量，则$\alpha_1, \alpha_2, \ldots, \alpha_s$的任何非零线性组合$\beta=k_1\alpha_1+k_2\alpha_2+\ldots+k_s\alpha_s$也是$A$的属于$\lambda$的特征向量。

证明: 由条件 $A\alpha_i=\lambda\alpha_i, i=1, 2,\ldots, s$。

从而 $\boldsymbol{A}\beta=\boldsymbol{A}(k_1\alpha_1+\ldots+k_s\alpha_s)=k_1\boldsymbol{A}\alpha_1+\ldots+k_s\boldsymbol{A}\alpha_s=k_1\lambda\alpha_1+\ldots+k_s\lambda\alpha_s$

$$=\lambda(k_1\alpha_1+\ldots+k_s\alpha_s)=\lambda\beta。$$

因此，β是A的属于λ的特征向量，证毕。

由定理5.5.4可知，若A有特征值λ，则$\boldsymbol{A}$的属于λ的特征向量有无穷多个。相反，若已知$\boldsymbol{A}$有特征向量α，则α只能属于$\boldsymbol{A}$的一个特征值。

事实上，若 α 属于A的特征值 λ_1， λ_2，则

$\boldsymbol{A}\alpha=\lambda_1\alpha$，$A\alpha=\lambda_2\alpha$，从而$\lambda_1\alpha=\lambda_2\alpha$，

得$(\lambda_1-\lambda_2)\alpha=0$，由于特征向量$\alpha\neq 0$，故$\lambda_1-$

$\lambda 2=0$，即$\lambda_1=\lambda_2$.

下面给出寻找特征值与特征向量的方法。

设$\boldsymbol{A}=(a_{ij})_{nn}$是数域$P$上的$n$阶方阵，若 λ 是

$\boldsymbol{A}$的特征值，

$$\alpha=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

是A的属于 λ 的特征向量，

$$\boldsymbol{A}\alpha = \lambda\alpha,$$

得

$$\lambda\alpha - \boldsymbol{A}\alpha = 0,$$

即

$$(\lambda I - \boldsymbol{A})\,\alpha = 0$$

注意$\lambda I-\boldsymbol{A}$是一个n阶矩阵，把 α 看作未知向量，上式就是一个齐次线性方程组

$$\left.\begin{aligned}(\lambda-a_{11})x_1-a_{12}x_2-\cdots-a_{1n}x_n&=0,\\-a_{21}x_1+(\lambda-a_{22})x_2-\cdots-a_{2n}x_n&=0,\\\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots&\\-a_{n1}x_1-a_{n2}x_2-\cdots+(\lambda-a_{nn})x_n&=0.\end{aligned}\right\}\quad(5.5.1)$$

由于α≠0，故$x_1,\ldots,x_n$不全为零，即$x_1,\ldots,x_n$是(5.5.1)的非零解。而齐次线性方程组(5.5.1)有非零解的充分必要条件是它的系数行列式为零，即

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

定义5.5.4 设A是数域P上的n阶方阵，λ是在P上取值的变量。矩阵λI一A称为A的**特征矩阵**。

$$|\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} \quad (5.5.2)$$

称为A的特征多项式。它是数域P上以λ为变元的一个n次多项式。

上面分析说明，如果 λ 是方阵A的特征值，则 λ 必是A的特征多项式的一个根；反之，如果λ是A的特征多项式在数域P中的一个根，则齐次线性方程组(5.5.1)必有非零解。这样，λ 就是A的一个特征值，而式(5.5.1)的非零解$\alpha=(x_1,\ldots,x_n)^{\mathrm{T}}$就是$A$的属于 λ 的特征向量。

综上所述，确定方阵A的特征值与特征向量的方法分为以下几步：

（1）写出A的特征多项式$|\lambda I-A|$，并求出它在数域P中全部的根（称为A的特征根），这些根也就是A的全部特征值；

（2）把所求得的特征值逐个地代入方程组(5.5.1)，对每个特征值解方程组(5.5.1)，求出它的基础解系，它们就是属于这个特征值的线性无关特征向量。

例5.5.1　求n阶数量矩阵kI的特征值与特征向量.

解： kI的特征多项式为

$$|\lambda I-kI|=\begin{vmatrix}\lambda-k & & & \\ & \lambda-k & & \\ & & \ddots & \\ & & & \lambda-k\end{vmatrix}=(\lambda-k)^n$$

特征多项式的根为$\lambda=k$，即kE的特征值只有k，它是一个n重特征根。

把$\lambda=k$代入$(\lambda I-kI)\alpha=0$，得

$$0\,\boldsymbol{\alpha}=0.$$

这说明任何非0向量都是kI的特征向量。直接由特征向量的定义也可知，数量矩阵kI左乘任何向量α后得到$k\alpha$。

例5.5.2 设

$$A=\begin{pmatrix} 2 & 3 & 2 \\ 1 & 4 & 2 \\ 1 & -3 & 1 \end{pmatrix}$$

为实数域R上的矩阵，求A的特征值与特征向量。

解　A的特征多项式为 $|\lambda I-A|=(\lambda-1)(\lambda-3)^2$

故A的特征值是3（二重特征根）和1.

对于特征值解 $\lambda=1$，齐次线性方程组 $(\lambda I-A)X=0$，得属于特征值1的特征向量

$$\alpha_1=\begin{pmatrix}-3\\-1\\3\end{pmatrix}$$

从而属于1的全部特征值为 $k_1\alpha_1$，$k_1\neq 0$ 。

对于二重特征值 $\lambda=3$ ，解齐次线性

方程组 $(3I-A)X=0$，得属于特征值3的特征向量 $\alpha_1=(-1,-1,1)^T$，从而属于特征值3的全部特征向量为 $k_1\alpha_1$，$k_1\neq 0$。

例5. 5. 3 设

$$A=\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$$

为实数域R上的矩阵，求A的特征值与特征向量。

解：　　A的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda-1 & -2 & -2 \\ -2 & \lambda-1 & -2 \\ -2 & -2 & \lambda-1 \end{vmatrix} = (\lambda+1)^2(\lambda-5)$$

故A的特征值是－1（二重特征根）和5.

把特征值－1代入 $(\lambda I - A)X = 0$ 得齐次线性方程组

$$\begin{cases} -2x_1 - 2x_2 - 2x_3 = 0, \\ -2x_1 - 2x_2 - 2x_3 = 0, \\ -2x_1 - 2x_2 - 2x_3 = 0. \end{cases}$$

它的基础解系是

$$\alpha_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \qquad \alpha_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

故属于−1的两个线性无关特征向量就是α_1，α_2，而属于−1的全部特征向量是$k_1\alpha_1 + k_2\alpha_2$，其中$k_1, k_2$为不同时为零的所有实数。

再把特征值5代入｜$\lambda I - A$｜＝ 0得齐次线性方程组

$$\begin{cases} 4x_1 - 2x_2 - 2x_3 = 0, \\ -2x_1 + 4x_2 - 2x_3 = 0, \\ -2x_1 - 2x_2 + 4x_3 = 0. \end{cases}$$

它的基础解系是

$$\alpha_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

它就是属于5的一个线性无关特征向量. 属于5的全部特征向量就是$k\alpha_3$，$k \in R$, $k \neq 0$。

由上述两个例子看出，如果 λ_0 是特征方程 $|\lambda I - A| = 0$ 的单根，那么属于 λ_0 的线性无关特征向量的个数只有一个；如果 λ_0 是特征方程 $|\lambda I - A| = 0$ 的重根，那么属于 λ_0的线性无关特征向量的个数可能等于λ_0 的重数，也可能小于 λ_0 的重数。

定理5.5.5 设 λ_0 是n阶方阵A的k重特征值，则A的属于特征值 λ_0 的线性无关的特征向量的个数不超过k。

证明：反证法，设属于特征值 λ_0 的线性无关的特征向量的个数为 l（$l>k$）个，分别用 $\alpha_1,\alpha_2,\cdots,\alpha_l$ 表示，由之前知识可知，可找到n−l个n维向量 $\alpha_{l+1},\alpha_{l+2},\cdots,\alpha_n$，使得 $\alpha_1,\alpha_2,\cdots,\alpha_l,\ \alpha_{l+1},\alpha_{l+2},\cdots,\alpha_n$

构成n维向量空间的一组基。以它们作列向量，得到n阶满秩矩阵

$$P=(\ \alpha_1,\alpha_2,\cdots,\alpha_l,\ \alpha_{l+1},\cdots,\alpha_n)$$

由于 $A\alpha_i(i=l+1,\cdots,n)$ 为n维向量，故可由

$\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出，且表达式唯一。设

$$A\alpha_i = p_{1i}\alpha_1 + p_{2i}\alpha_2 + p_{ni}\alpha_n \quad (l+1 \le i \le n)$$

于是 $AP = (A\alpha_1, A\alpha_2, \cdots, A\alpha_l, A\alpha_{l+1}, \cdots, A\alpha_n)$

$$= (\lambda_0\alpha_1, \lambda_0\alpha_2, \cdots, \lambda_0\alpha_l, A\alpha_{l+1}, \cdots, A\alpha_n)$$

$$= (\alpha_1,\alpha_2,\cdots,\alpha_l,\alpha_{l+1},\cdots,\alpha_n)\begin{pmatrix} \lambda_0 & 0 & \cdots & 0 & p_{1\,l+1} & \cdots & p_{1n} \\ 0 & \lambda_0 & \cdots & 0 & p_{2\,l+1} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_0 & p_{l\,l+1} & \cdots & p_{l\,n} \\ 0 & 0 & \cdots & 0 & p_{l+1\,l+1} & \cdots & p_{l+1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & p_{\mathrm{n}\,l+1} & \cdots & p_{nn} \end{pmatrix}$$

$$= P\begin{pmatrix} \lambda_0 I_l & P_1 \\ 0 & P_2 \end{pmatrix}$$

即

$$P^{-1}AP=\begin{pmatrix}\lambda_0 I_l & P_1\\ 0 & P_2\end{pmatrix}$$

故

$$|\lambda I-A|=\left|\lambda I-P^{-1}AP\right|=(\lambda-\lambda_0)^l\,|\lambda I-P_2|$$

这意味着λ_0至少是A的 /重特征值，而 />k，
这与 λ_0为A的k重特征值矛盾，证毕。

特征值与特征向量的性质

先看矩阵A的特征多项式

$$f(\lambda)=|\lambda I-A|$$

的形式. 由于

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

由行列式定义，展开式中有一项是主对角线元素的连乘积:

$$f(\lambda)=(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})+\cdots+C$$

而其余各项中至多包含n-2个主对角线的元素，它对λ的次数最多是 n-2，因此特征多项式中含λ的n次与n-1次的项只能在主对角线上元素的连乘积中出现，所以

$$f(\lambda)=\lambda^{n}-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+C$$

把λ=0代入上式得:

$$C=f(0)=\left|-A\right|=(-1)^{n}\left|A\right|$$

从而，

$$f(\lambda)=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n|A| \tag{5.5.3}$$

定义5.5.5　方阵A的主对角线元素之和$(a_{11}+a_{22}+\ldots+a_{nn})$称 为$A$的迹，记为$tr(A)$。

定理5.5.6　若n阶方阵A在数域P上有n个特征值（重根按重数计），则A的全体特征值之和等于A的迹$tr(A)$。A的全体特征值之积等于A的行列式｜A｜。

证明： 设A的特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$, 则

$$f(\lambda)=\left|\lambda I-A\right|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

即

$$f(\lambda)=\lambda^n-(\lambda_1+\lambda_2+\cdots+\lambda_n)\lambda^{n-1}+\cdots+(-1)^n\lambda_1\lambda_2\cdots\lambda_n$$

与(5.5.3)式比较即得

$$\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn},$$

$$\lambda_1\lambda_2\cdots\lambda_n=\left|A\right|.$$

证毕。

推论 复数域方阵A可逆的充分必要条件是A的特征值全不为零。

定理5.5.7 若n阶可逆阵A的特征值为λ_1, $\lambda_2,\cdots,\lambda_n$，则$A^{-1}$的特征值恰为 $1/\lambda_1, 1/\lambda_2,\ldots, 1/\lambda_n$。

证明： 由于A可逆，由定理5.5.6知 $\lambda_i\neq0$(i=1, 2, …, n)，因此 $1/\lambda_1, 1/\lambda_2,\ldots, 1/\lambda_n$有意义.

设α_i是A的属于特征值λ_i的特征向量，则

$$A\alpha_i=\lambda_i\alpha_i,\ (i=1, 2, \ldots,n),$$

此式左乘A^{-1}得 $\alpha_i = \lambda_i (A^{-1}\alpha_i)$,

即

$$A^{-1}\alpha_i = (1/\lambda_i)\alpha_i$$

从而$1/\lambda_i$是A^{-1}的特征值。故A^{-1}的全部特征值恰为 $1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n$，证毕。

例5.5.4　设A是准对角阵

$$A = \begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}$$

则$A_1, A_2,\ldots, A_s$的所有特征值就是A的全部特征值.

证明：令$I_i\,(i=1, 2,\ldots, s)$是与$A_i\,(i=1, 2,\ldots, s)$同阶的单位阵，则有

$$|\lambda I-A|=\begin{vmatrix}\lambda I_1-A_1 & & & \\ & \lambda I_2-A_2 & & \\ & & \ddots & \\ & & & \lambda I_s-A_s\end{vmatrix}$$

$$=|\lambda I_1-A_1||\lambda I_2-A_2|\cdots|\lambda I_s-A_s|$$

从而A的特征多项式是所有$A_i\,(i=1, 2,\ldots, s)$的特征多项式之积. 故$A_i\,(i=1, 2,\ldots,$

s)的所有特征值就是A的全部特征值。

下面给出特征向量的一个重要性质。

定理5. 5. 8 若$\lambda_1, \lambda_2, \ldots, \lambda_m$是$A$的$m$个不同特征值，$\alpha_1, \alpha_2, \ldots, \alpha_m$是分别属于它们的特征向量，则$\alpha_1,\ldots, \alpha_m$是线性无关的。

证明： 对不同特征值的个数m作归纳

当m=1时，单个非零的特征向量总是线性无关的，定理成立。

假设对$m-1$个属于不同特征值的特征向量定理成立，考察m个属于不同特征值的特征向量$\alpha_1, \alpha_2, \ldots, \alpha_m$，令

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0 \qquad (5.5.4)$$

用A左乘(5.5.4)式得

$$k_1\lambda_1\alpha_1 + k_2\lambda_2\alpha_2 + \cdots + k_m\lambda_m\alpha_m = 0 \qquad (5.5.5)$$

再用λ_m乘(5. 5. 4)式并与(5. 5. 5)式相减得

$$k_1(\lambda_m - \lambda_1)\alpha_1 + k_2(\lambda_m - \lambda_2)\alpha_2 + \cdots + k_{m-1}(\lambda_m - \lambda_{m-1})\alpha_{m-1} = 0$$

由归纳假设$\alpha_1, \ldots, \alpha_{m-1}$线性无关，且

$$\lambda_m - \lambda_i \neq 0\,(i=1,2,\cdots,m-1)$$

故$k_1=k_2=\ldots=k_{m-1}=0$，这时，式(5.5.4)变为

$$k_m\alpha_m = 0$$

由于$\alpha_m\neq 0$，又有$k_m=0$. 这样证明了$\alpha_1, \alpha_2, \ldots, \alpha_m$线性无关. 证毕。

更一般地有如下定理

定理5.5.9 设n阶方阵有m个互不相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_m$ ，而属于$\lambda_i\,(i=1,2,\cdots,m)$的所有线性无关特征向量有 r_i个： $\alpha_{i1},\alpha_{i2},\cdots\alpha_{ir_i}$， 那么

由这些特征向量组成的向量组

$\alpha_{11},\alpha_{12},\cdots,\alpha_{1r_1}$， $\alpha_{21},\alpha_{22},\cdots,\alpha_{2r_2}$ ， ...， $\alpha_{m1},\alpha_{m2},\cdots,\alpha_{mr_m}$

也线性无关.

证明： 设

$$k_{11}\alpha_{11}+\cdots+k_{1r_1}\alpha_{1r_1}+k_{21}\alpha_{21}+\cdots+k_{2r_2}\alpha_{2r_2}+\cdots+k_{m1}\alpha_{m1}+\cdots+k_{mr_m}\alpha_{mr_m}=0. \quad (5.5.6)$$

记 $\alpha_i=k_{i1}\alpha_{i1}+\cdots+k_{ir_i}\alpha_{ir_i}=\sum_{j=1}^{r_i}k_{ij}\alpha_{ij}$， $i=1,2,\cdots,m$

则(5.5.6)式可以写成

$$\alpha_1+\alpha_2+\cdots+\alpha_m=0 \qquad (5.5.7)$$

显然这m个向量全为零向量. 若有某些$\alpha_i \neq 0$，由定理5.5.1可知仍是属于 λ_i的特征向量，而(5.5.7)式说明，这些属于不同特征值的特征向量线性相关，矛盾。

故必 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 全为0。 即

$$\alpha_i=k_{i1}\alpha_{i1}+\cdots+k_{ir_i}\alpha_{ir_i}=0,(i=1,2,\cdots,n)$$

又 $\alpha_{i1},\cdots,\alpha_{ir_i}$ 线性无关，得k_{ij}=0， i=1，…， m; j=1，…，r_i。 从而向量组 $\alpha_{11},\cdots,\alpha_{1r_1}$, $\alpha_{21},\cdots,\alpha_{2r_2}$，…， $\alpha_{m1},\cdots,\alpha_{mr_m}$ 线性无关. 证毕.