第五章 矩阵的相似对角化

第一节　矩阵的特征值和特征向量 相似矩阵

❖ 本节进一步讨论方阵的内在性质，加深对矩阵的认识和理解，以便更好地使用矩阵解决线性代数中的问题。

一.矩阵的特征值和特征向量

定义**1** 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，则称数λ为矩阵**A**的特征值**(eigenvalue)**，非零向量**x**称为**A**的对应于特征值λ的特征向量**(eigenvector)**。

例1 试验证

$$\alpha=\begin{bmatrix}1\\1\end{bmatrix},\quad \beta=\begin{bmatrix}2\\-1\end{bmatrix}$$

是矩阵

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

分别是属于特征值λ**1 = 1**和λ**2= –5**的特征向量

证 只需验证**A**α **= 1·**α，**A**β **= –5·**β：

$$A\alpha=\begin{bmatrix}-3 & 4\\ 2 & -1\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}=\begin{bmatrix}1\\ 1\end{bmatrix}=1\cdot\alpha$$

$$A\beta=\begin{bmatrix}-3 & 4\\ 2 & -1\end{bmatrix}\begin{bmatrix}2\\ -1\end{bmatrix}=\begin{bmatrix}-10\\ 5\end{bmatrix}=-5\begin{bmatrix}2\\ -1\end{bmatrix}=-5\cdot\beta \qquad \textbf{(1)}$$

式也可以写成，

$$(A-\lambda E)x = 0 \qquad \textbf{(2)}$$

这是**n**个未知数**n**个方程的齐次线性方程组，我们知道，它有非零解的充分必要条件是系数行列式

$$|A-\lambda E| = 0 \qquad \textbf{(3)}$$

即

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

上式是以λ为未知数的一元**n**次方程**,**称为方阵**A**的特征方程。其左边是λ 的**n**次多项式，记作**f (λ)**，称为**A**的特征多项式。显然， **A**的特征值就是特征方程的根或特征多项式的零点。这个**n**次的特征方程在计算根的重数时应共有**n**个实根或复根。因此**,n**阶矩阵有**n**个特征值。

读者注意：（1）式也可改写成（λE－A）x＝0，从而（3）式变成| λE –A | = 0，

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

即有些教材把上式定义为方阵A的特征方程，其左边是λ 的n次首一多项式，亦记作f (λ)，并称为A的特征多项式。事实上，为叙述的方便，在本章的最后一节中，作者采用的就是这种定义。

设n阶矩阵$A=(a_{ij})$的特征值为$\lambda_1,\lambda_2,\ldots,\lambda_n$，由多项式根与系数之间的关系可得：

1. $\lambda_1 + \lambda_2 + \ldots + \lambda_n = a_{11} + a_{22} + \ldots + a_{nn}$
2. $\lambda_1\lambda_2\ldots\lambda_n = |A|$

由(2)知，方阵A可逆的充要条件是A有n个非零的特征值。

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就是A的对应于特征值λ_i的特征向量。

（注意：若为λ_i是实数，则P_i可取实向量；若λ_i为复数，则P_i为复向量。）

例2　求

$$A=\begin{bmatrix}3 & -1\\ -1 & 3\end{bmatrix}$$

的特征值和特征向量。

解　A的特征多项式为

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

所以A的特征值为$\lambda_1=4$，$\lambda_2=2$。

当$\lambda_1=2$时，由$(A-\lambda_1 E)x=0$，即

$$\begin{bmatrix}3-2 & -1\\ -1 & 3-2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}$$

即

$$\begin{cases} x_1 - x_2 = 0 \\ -x_1 + x_2 = 0 \end{cases}$$

解得x1 = x2，所以对应的特征向量可取为

$$\mathbf{p}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

当λ2 = 4时，由(A – λ2E) x = 0，即

$$\begin{bmatrix} 3-4 & -1 \\ -1 & 3-4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

解得x1 = – x2，所以对应的特征向量可取为

$$\mathbf{p}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

显然，若**Pi**为**A**的对应于特征值λ**i**的特征向量，则**kPi**（**k≠0**）也是对应于λ**i**的特征向量

例**3**　求矩阵

$$A = \begin{bmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

的特征值和特征向量。

解　**A**的特征多项式为

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

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

当$\lambda_1 = 2$时，解方程$(A - 2E)\,x = 0$。由

$$A - 2E = \begin{bmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

得基础解系

$$\mathbf{p}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

所以$k_1 P_1 (k_1 \neq 0)$是对应于$\lambda_1 = 2$的全部特征值。

当$\lambda_2 = \lambda_3 = 1$时,解方程$(A - E)\,x = 0$。由

$$A - E = \begin{bmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

得基础解系

$$\mathbf{p}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

所以**k_2P_2 ($k_2 \neq 0$)**是对应于**$\lambda_2 = \lambda_3 = 1$**的全部特征向量。

例4　求矩阵

$$A=\begin{bmatrix}-2 & 1 & 1\\ 0 & 2 & 0\\ -4 & 1 & 3\end{bmatrix}$$

的特征值和特征向量

解

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

$$=(2-\lambda)\begin{bmatrix}-2-\lambda & 1\\ -4 & 3-\lambda\end{bmatrix}$$

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

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

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

当$\lambda_1 = -1$时，解方程$(A + E)x = 0$。由

$$A + E = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 3 & 0 \\ -4 & 1 & 4 \end{bmatrix} \to \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

得基础解系

$$\mathbf{p}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

所以对应于$\lambda_1 = -1$的全部特征向量为k_1P_1
$(k_1 \neq 0)$。

当$\lambda_2 = \lambda_3 = 2$时，解方程$(A - 2E)x = 0$。由

$$A - 2E = \begin{bmatrix} -4 & 1 & 1 \\ 0 & 0 & 0 \\ -4 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} -4 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

得基础解系

$$\mathbf{p}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \qquad \mathbf{p}_3 = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix},$$

所以对应于$\lambda_2 = \lambda_3 = 2$的全部特征向量为$k_2P_2 + k_3P_3$　$(k_2, k_3$不同时为零$)$

例5 设λ为方阵A的特征值，证明λ^2是A^2的特征值。

证 因λ为A的特征值，故有$P \neq 0$使$AP=\lambda P$。于是

$A^2 P = A(AP) = A(\lambda P) = \lambda(AP) = \lambda^2 P$,

所以，λ^2是A^2的特征值 。

由此例类推，不难证明：若λ为方阵A的特征值，则λ^k是A^k的特征值；

$\varphi(\lambda)$是$\varphi(A)$的特征值（其中，$\varphi(\lambda)$为λ的m次多项式）。若A为可逆矩阵，则还有λ^{-1}是A^{-1}的特征值。

定理1 设$\lambda_1, \lambda_2, \ldots, \lambda_m$是方阵A的m个特征值，$p_1, p_2, \ldots, p_m$依次是与之对应的特征向量。如果$\lambda_1, \lambda_2, \ldots, \lambda_m$各不相等，则$p_1, p_2, \ldots, p_m$线性无关。

证　设有常数x1，x2，…，xm使

x1p1 + x2p2 + … + xmpm= 0

则　A(x1p1 + x2p2 + … + xmpm)= 0，即

λ1x1p1 +λ2 x2p2 + … + λmxmpm= 0

又 A (λ1x1p1 +λ2 x2p2 + … + λmxmpm) = 0，即

类推之，有

$$\lambda_1^k x_1 \mathbf{p}_1 + \lambda_2^k x_2 \mathbf{p}_2 + \cdots + \lambda_m^k x_m \mathbf{p}_m = \mathbf{0} \qquad (k = 1, 2, \cdots, m-1)$$

把上述各式合写成矩阵形式，得

$$(x_1 \mathbf{p}_1, x_2 \mathbf{p}_2, \cdots, x_m \mathbf{p}_m) \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & \lambda_m & \cdots & \lambda_m^{m-1} \end{bmatrix} = (\mathbf{0}, \mathbf{0}, \cdots, \mathbf{0}).$$

上式等号左边第二个矩阵的行列式为范德蒙行列式，当λI各不相同时该行列式不等于零，从而该矩阵可逆，于是有

$(x_1p_1, x_2p_2, \ldots, x_mp_m) = (0, 0, \ldots, 0)$

即$x_i p_i = 0$，（$i = 1, 2, \ldots, m$）。但$p_i \neq 0$，故$x_i = 0$（$i = 1, 2, \ldots, m$）。

所以，向量组$p_1, p_2, \ldots, p_m$线性无关。

二、 相似矩阵及其性质

定义2 设A，B都是n阶方阵，若有可逆方阵P，使 $P^{-1}AP = B$ 则称B是A的相似矩阵，或说A与B相似。对A进行运算 $P^{-1}AP$ 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。

容易证明，方阵之间的相似关系是等价关系。

定理2 若n阶方阵A，B相似，则A，B的特征多项式相同，从而A与B的特征值相同 。

证　因为A，B相似，所以存在可逆矩阵P，使 $P^{-1}AP = B$。故

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

推论　若n阶方阵A与对角阵。

$$
\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}
$$

相似，则$\lambda_1,\lambda_2,\ldots,\lambda_n$就是A的n个特征值。

定理3 若n阶方阵A，B相似，相似变换矩阵为P，则

1. A^k与B^k相似，且相似变换矩阵为P，即$A^k = P^{-1}B^k P$ (k为正整数)；
2. $\varphi(A)$与$\varphi(B)$相似，且相似变换矩阵为P，即$\varphi(A) = P^{-1}\varphi(B)P$。

其中$\varphi(A)$为A的m次多项式。

证 (1) 因为$A = P^{-1}BP$，所以

$$A^2 = (P^{-1}BP)(P^{-1}BP) = P^{-1}B^2P,$$

$$A^3 = (P^{-1}B^2P)(P^{-1}BP) = P^{-1}B^3P,$$

$$\cdots$$

$$A^k = (P^{-1}B^{k-1}P)(P^{-1}BP) = P^{-1}B^kP,$$

(2) 设

$\varphi(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mE$,

由（1）知,

$$\varphi(A) = a_0A^m + a_1A^{m-1} + \ldots + a_{m-1}A + a_mE$$
$$= a_0P^{-1}B^mP + a_1P^{-1}B^{m-1}P + \ldots + a_{m-1}P^{-1}BP + a_mP^{-1}EP$$
$$= P^{-1}(a_0B^m + a_1B^{m-1} + \ldots + a_{m-1}B + a_mE)P$$
$$= P^{-1}\varphi(B)P$$

三　一般矩阵的对角化

- 这一段我们要讨论的主要问题是：对**n**阶方阵**A**，寻求相似变换矩阵**P**，使$P^{-1}AP=\Lambda$为对角阵，该手续称为把方阵**A**对角化，并称可与对角阵相似的矩阵为可对角化矩阵。

定理4 **n**阶方阵**A**可对角化的充要条件是**A**具有**n**个线性无关的特征向量。

证　必要性　设**A**与对角阵$D=\mathrm{diag}(d_1,d_2,\ldots,d_n)$相似，故有满秩矩阵**P**，成立

$$A = PDP^{-1}$$

即

$$AP = PD$$

对P按列分块，

$$P = [p1，p2，…，pn]$$

则可将（4）式写成n个向量等式

$$Api = dipi \qquad （i=1，2，…，n）$$

显然Pi是A的对应于特征值di的特征向量。又因P可逆，故p1，p2，…，pn线性无关。

充分性

若A有n个线性无关的特征向量p1，p2，…，pn，它们所对应的特征值分别为λ1，λ2，…，λn。则有

$$Api = \lambda ipi \qquad （i=1，2，…，n）$$

将这n个向量等式合并成矩阵等式，得

$$A[p1，p2，…，pn] = [\lambda 1p1，\lambda 2p2，…，\lambda npn]$$

联系定理1，可得如下推论：

推论　如果n阶方阵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，就不一定有n个线性无关的特征向量，从而就不一定可对角化。例如在例3中A的特征方程有重根，确实找不到3个线性无关的特征向量，因此，例3中的A不能对角化；而在例4中虽然A的特征方程也有重根，但却能找到3个线性无关的特征向量，因此例4中的A可对角化。

　一个方阵具备什么条件才能对角化是一个较复杂的问题。我们对此不作一般性的讨论，而仅讨论当A为实对称阵的情形。

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -5 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{bmatrix}$$