第5章　特征值

郑州大学数学与统计学院 线性代数教研室

目录

定义 5.1 设P为一个数域，$\boldsymbol{A}$为P上的n阶方阵，若存在数$\lambda \in P$及n维非零(列)向量$\boldsymbol{X} \in P^n$，使

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

则称λ是$\boldsymbol{A}$的一个**特征值**，$\boldsymbol{X}$是属于λ的一个**特征向量**。

定义 5.1 设P为一个数域，$\boldsymbol{A}$为P上的n阶方阵，若存在数$\lambda \in P$及n维非零(列)向量$\boldsymbol{X} \in P^n$，使

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

则称λ是$\boldsymbol{A}$的一个**特征值**，$\boldsymbol{X}$是属于λ的一个**特征向量**。

易知：(1) 若$\boldsymbol{X}$是$\boldsymbol{A}$属于λ的特征向量，则对任意非零数$c \in P$，$c\boldsymbol{X}$也是$\boldsymbol{A}$属于λ的特征向量；

(2) 对应于一个特征向量的特征值是唯一的，而对应于一个特征值的特征向量有无穷多个；

(3) 若$\boldsymbol{X}_1, \boldsymbol{X}_2$都是$\boldsymbol{A}$属于特征值$\lambda$的特征向量，且$\boldsymbol{X}_1 + \boldsymbol{X}_2 \neq 0$，则$\boldsymbol{X}_1 + \boldsymbol{X}_2$也是$\boldsymbol{A}$属于特征值$\lambda$的特征向量。

综合(1),(2),(3), 不难得到

命题 设$\lambda \in P$是P上n阶方阵$\boldsymbol{A}$的一个特征值，则$\boldsymbol{A}$属于λ的全体特征向量添上零向量按向量的加法和数乘运算作成P^n的一个子空间，称为$\boldsymbol{A}$关于特征值λ 的**特征子空间**。

下面讨论如何求$\boldsymbol{A}$的特征值与特征向量。由定义知

$$(\lambda \boldsymbol{E} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0},$$

即

$$\begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

下面讨论如何求$\boldsymbol{A}$的特征值与特征向量。由定义知

$$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\mathbf{0},$$

即

$$\begin{bmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

可以看出，λ为$\boldsymbol{A}$的特征值，且$\boldsymbol{X}$是$\boldsymbol{A}$属于λ的特征向量的充要条件是$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\mathbf{0}$有非零解，即$|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$.

下面讨论如何求$\boldsymbol{A}$的特征值与特征向量。由定义知

$$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0},$$

即

$$\begin{bmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

可以看出，λ为$\boldsymbol{A}$的特征值，且$\boldsymbol{X}$是$\boldsymbol{A}$属于λ的特征向量的充要条件是$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$有非零解，即$|\lambda \boldsymbol{E}-\boldsymbol{A}|=0$.

显然，$\boldsymbol{A}$关于λ的特征子空间就是$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$的解空间，其维数为$n-r(\lambda \boldsymbol{E}-\boldsymbol{A})$。

定义 5.2 设$\boldsymbol{A}$为P上的n阶方阵，λ为未定元。则称$\lambda\boldsymbol{E}-\boldsymbol{A}$为$\boldsymbol{A}$的**特征矩阵**，称行列式$|\lambda\boldsymbol{E}-\boldsymbol{A}|$为$\boldsymbol{A}$的**特征多项式**，又称$n$次代数方程$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$为$\boldsymbol{A}$的**特征方程**。

定义 5.2 设$\boldsymbol{A}$为P上的n阶方阵，λ为未定元。则称$\lambda\boldsymbol{E}-\boldsymbol{A}$为$\boldsymbol{A}$的**特征矩阵**，称行列式$|\lambda\boldsymbol{E}-\boldsymbol{A}|$为$\boldsymbol{A}$的**特征多项式**，又称$n$次代数方程$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$为$\boldsymbol{A}$的**特征方程**。

因此，数λ为$\boldsymbol{A}$的特征值的充要条件是λ为特征方程$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$的根。所以，特征值又称为**特征根**。它由$\boldsymbol{A}$唯一确定，而属于$\lambda$的特征向量$\boldsymbol{X}$是$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$的非零解，当然也是由$\boldsymbol{A}$所确定的。

定义 5.2 设$\boldsymbol{A}$为P上的n阶方阵，λ为未定元。则称$\lambda\boldsymbol{E}-\boldsymbol{A}$为$\boldsymbol{A}$的**特征矩阵**，称行列式$|\lambda\boldsymbol{E}-\boldsymbol{A}|$为$\boldsymbol{A}$的**特征多项式**，又称$n$次代数方程$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$为$\boldsymbol{A}$的**特征方程**。

因此，数λ为$\boldsymbol{A}$的特征值的充要条件是λ为特征方程$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$的根。所以，特征值又称为**特征根**。它由$\boldsymbol{A}$唯一确定，而属于$\lambda$的特征向量$\boldsymbol{X}$是$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$的非零解，当然也是由$\boldsymbol{A}$所确定的。

当然，对于给定的数域P，不是所有的方阵都有特征值！

下面的定理揭示了特征值与原矩阵的关系。

定理 5.1 设$\boldsymbol{A}$为P上的n阶方阵，且$\boldsymbol{A} = (a_{ij})_{nn}$在$P$上有$n$个特征值$\lambda_1, \lambda_2, \cdots, \lambda_n$(重根按重数计算)，则

(1) $|\boldsymbol{A}| = \lambda_1\lambda_2\cdots\lambda_n$;

(2) $\boldsymbol{A}$的主对角线元素之和称为$\boldsymbol{A}$的**迹**，记为tr$\boldsymbol{A}$,

$$\mathrm{tr}\boldsymbol{A} = a_{11} + a_{22} + \cdots + a_{nn} = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

下面的定理揭示了特征值与原矩阵的关系。

定理 5.1 设$\boldsymbol{A}$为P上的n阶方阵，且$\boldsymbol{A}=(a_{ij})_{nn}$在$P$上有$n$个特征值$\lambda_1,\lambda_2,\cdots,\lambda_n$(重根按重数计算)，则

(1) $|\boldsymbol{A}|=\lambda_1\lambda_2\cdots\lambda_n$;

(2) $\boldsymbol{A}$的主对角线元素之和称为$\boldsymbol{A}$的**迹**，记为tr$\boldsymbol{A}$，

$$\mathrm{tr}\boldsymbol{A}=a_{11}+a_{22}+\cdots+a_{nn}=\lambda_1+\lambda_2+\cdots+\lambda_n.$$

证明 (1) 因为$\lambda_1,\cdots,\lambda_n$为特征方程$|\lambda\boldsymbol{E}-\boldsymbol{A}|=0$的根，所以特征多项式

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

令$\lambda=0$，得$|-\boldsymbol{A}|=(-1)^n\lambda_1\lambda_2\cdots\lambda_n$，从而结论(1)成立。

(2) 先计算$|\lambda \boldsymbol{E}-\boldsymbol{A}|$中$\lambda^{n-1}$的系数。

可以看出，$|\lambda \boldsymbol{E}-\boldsymbol{A}|$中$\lambda^{n-1}$的系数与$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$中$\lambda^{n-1}$ 的系数相同，都是$-(a_{11}+a_{22}+\cdots+a_{nn})$. 而$(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$中$\lambda^{n-1}$的系数为$-(\lambda_1+\lambda_2+\cdots+\lambda_n)$. 由(1)的证明知

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

例 1 在有理数域Q上求方阵$\boldsymbol{A}=\begin{bmatrix}3 & 1\\ 5 & -1\end{bmatrix}$的特征值与特征向量。

例 1 在有理数域$\boldsymbol{Q}$上求方阵$\boldsymbol{A}=\begin{bmatrix}3 & 1\\ 5 & -1\end{bmatrix}$的特征值与特征向量。

解 由$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-3 & -1\\ -5 & \lambda+1\end{vmatrix}=(\lambda-4)(\lambda+2)$，得$\boldsymbol{A}$的两个特征值$4,-2$.

例 1 在有理数域$\boldsymbol{Q}$上求方阵$\boldsymbol{A}=\begin{bmatrix}3 & 1\\ 5 & -1\end{bmatrix}$的特征值与特征向量。

解 由$|\lambda \boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-3 & -1\\ -5 & \lambda+1\end{vmatrix}=(\lambda-4)(\lambda+2)$，得$\boldsymbol{A}$的两个特征值$4,-2$.

将$\lambda=4$代入$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得

$$\begin{cases} x_1-x_2=0,\\ -5x_1+5x_2=0.\end{cases}$$

其基础解系是$\boldsymbol{X}_1=\begin{bmatrix}1\\ 1\end{bmatrix}$，所以属于$\lambda=4$的全部特征向量是$k\begin{bmatrix}1\\ 1\end{bmatrix}$，$k\neq 0$是任意常数。

将$\lambda=-2$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得

$$\begin{cases}-5x_1-x_2=0,\\-5x_1-x_2=0.\end{cases}$$

其基础解系是$\boldsymbol{X}_2=\begin{bmatrix}1\\-5\end{bmatrix}$，所以属于$\lambda=-2$的全部特征向量是$l\begin{bmatrix}1\\-5\end{bmatrix}$，$l\neq0$是任意常数。

例 2 在有理数域Q上求方阵$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$的特征值与特征向量。

例 2 在有理数域$\boldsymbol{Q}$上求方阵$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$的特征值与特征向量。

解 由

$$|\lambda \boldsymbol{E} - \boldsymbol{A}| = \begin{vmatrix} \lambda - 1 & -1 & 0 \\ 0 & \lambda - 1 & 0 \\ 0 & 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)^3,$$

得$\boldsymbol{A}$的特征值是$\lambda_1 = \lambda_2 = \lambda_2 = 1.$

将$\lambda=1$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得$x_2=0$，其基础解系是

$$\boldsymbol{X}_1=\begin{bmatrix}1\\0\\0\end{bmatrix},\quad \boldsymbol{X}_2=\begin{bmatrix}0\\0\\1\end{bmatrix}.$$

所以属于特征值1的全部特征向量是

$$k_1\begin{bmatrix}1\\0\\0\end{bmatrix}+k_2\begin{bmatrix}0\\0\\1\end{bmatrix},\quad k_1,k_2\text{不全为零}.$$

例 3 在复数域C上求复矩阵$\boldsymbol{A}=\begin{bmatrix} 1 & 2 & 2 \\ 1 & -1 & 1 \\ -4 & 12 & 1 \end{bmatrix}$的特征值与特征向量。

例 3 在复数域C上求复矩阵$\boldsymbol{A}=\begin{bmatrix}1 & 2 & 2\\ 1 & -1 & 1\\ -4 & 12 & 1\end{bmatrix}$的特征值与特征向量。

解 由 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1 & -2 & -2\\ -1 & \lambda+1 & -1\\ -4 & 12 & \lambda-1\end{vmatrix}=(\lambda-1)(\lambda+\mathrm{i})(\lambda-\mathrm{i})$,

得$\boldsymbol{A}$的特征值是$1, \mathrm{i}, -\mathrm{i}$.

例 3 在复数域C上求复矩阵$\boldsymbol{A}=\begin{bmatrix}1 & 2 & 2\\ 1 & -1 & 1\\ -4 & 12 & 1\end{bmatrix}$的特征值与特征向量。

解 由 $|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-1 & -2 & -2\\ -1 & \lambda+1 & -1\\ -4 & 12 & \lambda-1\end{vmatrix}=(\lambda-1)(\lambda+\mathrm{i})(\lambda-\mathrm{i})$,
得$\boldsymbol{A}$的特征值是$1,\mathrm{i},-\mathrm{i}$. 分别代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得

属于特征值1的全部特征向量是$k_1\begin{bmatrix}3\\ 1\\ -1\end{bmatrix},\quad k_1\neq 0;$

属于特征值i的全部特征向量是$k_2\begin{bmatrix}4+2\mathrm{i}\\ 1+\mathrm{i}\\ -4\end{bmatrix},\quad k_2\neq 0;$

属于特征值$-$i的全部特征向量是$k_3\begin{bmatrix}4-2\mathrm{i}\\ 1-\mathrm{i}\\ -4\end{bmatrix},\quad k_3\neq 0;$

定义 5.3 设$\boldsymbol{A},\boldsymbol{B}$为数域$P$上$n$阶方阵，若存在$P$上$n$阶可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}$，则称$\boldsymbol{A}$与$\boldsymbol{B}$相似，记作$\boldsymbol{A}\overset{C}{\sim}\boldsymbol{B}$ 或$\boldsymbol{A}\sim\boldsymbol{B}$或$\boldsymbol{A}\overset{P}{\sim}\boldsymbol{B}$.

定义 5.3 设$\boldsymbol{A},\boldsymbol{B}$为数域$P$上$n$阶方阵，若存在$P$上$n$阶可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\boldsymbol{B}$，则称$\boldsymbol{A}$与$\boldsymbol{B}$**相似**，记作$\boldsymbol{A}\overset{C}{\sim}\boldsymbol{B}$ 或$\boldsymbol{A}\sim\boldsymbol{B}$或$\boldsymbol{A}\overset{P}{\sim}\boldsymbol{B}$.

矩阵相似是矩阵等价的特殊情况，它当然也保持矩阵的秩不变。此外，它还有类似于矩阵等价的三个性质：

(1) 自反性：对任意方阵$\boldsymbol{A}$，有$\boldsymbol{A}\overset{E}{\sim}\boldsymbol{A}$;

(2) 对称性：若$\boldsymbol{A}\overset{C}{\sim}\boldsymbol{B}$，则$\boldsymbol{B}\overset{C^{-1}}{\sim}\boldsymbol{A}$;

(3) 传递性：若$\boldsymbol{A}\overset{C_1}{\sim}\boldsymbol{B},\boldsymbol{B}\overset{C_2}{\sim}\boldsymbol{C}$，则$\boldsymbol{A}\overset{C_1C_2}{\sim}\boldsymbol{C}$.

矩阵相似还有下面重要的性质：

定理 5.2 (1) 相似矩阵有相同的特征多项式，当然也有相同的特征值(重根按重数计算)；

(2) 相似矩阵的行列式相等；

(3) 相似矩阵有相同的迹，即两个相似矩阵的主对角线上元素之和相等。

定理 5.3 数域P上n阶方阵$\boldsymbol{A}$相似于对角矩阵的充要条件是$\boldsymbol{A}$在数域P上有n个线性无关的特征向量。

定理 5.3 数域P上n阶方阵$\boldsymbol{A}$相似于对角矩阵的充要条件是$\boldsymbol{A}$在数域P上有n个线性无关的特征向量。

证明 必要性. 设存在数域P上可逆矩阵$\boldsymbol{C}$，使

$$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}, \quad \text{则} \quad \boldsymbol{A}\boldsymbol{C}=\boldsymbol{C}\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}$$

把$\boldsymbol{C}$按列分块，得$\boldsymbol{C}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)$. 于是有

$$\boldsymbol{A}(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

即 $(\boldsymbol{A}\boldsymbol{X}_1,\boldsymbol{A}\boldsymbol{X}_2,\cdots,\boldsymbol{A}\boldsymbol{X}_n)=(\lambda_1\boldsymbol{X}_1,\lambda_2\boldsymbol{X}_2,\cdots,\lambda_n\boldsymbol{X}_n)$，故$\boldsymbol{A}\boldsymbol{X}_i=\lambda_i\boldsymbol{X}_i$

把$\boldsymbol{C}$按列分块，得$\boldsymbol{C}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)$. 于是有

$$\boldsymbol{A}(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

即$(\boldsymbol{A}\boldsymbol{X}_1,\boldsymbol{A}\boldsymbol{X}_2,\cdots,\boldsymbol{A}\boldsymbol{X}_n)=(\lambda_1\boldsymbol{X}_1,\lambda_2\boldsymbol{X}_2,\cdots,\lambda_n\boldsymbol{X}_n)$，故$\boldsymbol{A}\boldsymbol{X}_i=\lambda_i\boldsymbol{X}_i$

由于$\boldsymbol{C}$可逆，所以$\boldsymbol{C}$的列向量组$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$线性无关。注意到$\boldsymbol{X}_i\neq 0$，所以$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$就是矩阵$\boldsymbol{A}$在数域$P$上$n$个线性无关的特征向量。

充分性. 设n阶方阵$\boldsymbol{A}$在数域P上有n个线性无关的特征向量$\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$，它们对应的特征值分别为$\lambda_1, \lambda_2, \cdots, \lambda_n \in P$.

令$\boldsymbol{C} = (\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n)$，则$\boldsymbol{C}$为数域$P$上可逆矩阵，且

$$\begin{aligned}
\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} &= \boldsymbol{C}^{-1}\boldsymbol{A}(\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n) = \boldsymbol{C}^{-1}(\boldsymbol{A}\boldsymbol{X}_1, \boldsymbol{A}\boldsymbol{X}_2, \cdots, \boldsymbol{A}\boldsymbol{X}_n) \\
&= \boldsymbol{C}^{-1}(\lambda_1\boldsymbol{X}_1, \lambda_2\boldsymbol{X}_2, \cdots, \lambda_n\boldsymbol{X}_n) \\
&= \boldsymbol{C}^{-1}(\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n)\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.
\end{aligned}$$

下面定理表明，属于不同特征值的特征向量线性无关。

定理 5.4 设$\lambda_1, \lambda_2, \cdots, \lambda_s \in P$是矩阵$\boldsymbol{A}$在$P$中的$s$个互异的特征值。若对每个特征值$\lambda_k$，$\boldsymbol{A}$有$m_k$个线性无关的特征向量$\boldsymbol{X}_{k1}, \boldsymbol{X}_{k2}, \cdots, \boldsymbol{X}_{km_k}$，$k = 1, 2, \cdots, s$，则向量组$\boldsymbol{X}_{11}, \cdots, \boldsymbol{X}_{1m_1}, \boldsymbol{X}_{21}, \cdots, \boldsymbol{X}_{2m_2}, \cdots, \boldsymbol{X}_{s1}, \cdots, \boldsymbol{X}_{sm_s}$也线性无关。

下面定理表明，属于不同特征值的特征向量线性无关。

定理 5.4 设$\lambda_1, \lambda_2, \cdots, \lambda_s \in P$是矩阵$\boldsymbol{A}$在$P$中的$s$个互异的特征值。若对每个特征值$\lambda_k$，$\boldsymbol{A}$有$m_k$个线性无关的特征向量$\boldsymbol{X}_{k1}, \boldsymbol{X}_{k2}, \cdots, \boldsymbol{X}_{km_k}$，$k = 1, 2, \cdots, s$，则向量组$\boldsymbol{X}_{11}, \cdots, \boldsymbol{X}_{1m_1}, \boldsymbol{X}_{21}, \cdots, \boldsymbol{X}_{2m_2}, \cdots, \boldsymbol{X}_{s1}, \cdots, \boldsymbol{X}_{sm_s}$也线性无关。

证明 对s用数学归纳法。当$s = 1$时显然成立。假设定理5.4对$s-1$成立，即已知向量组$\boldsymbol{X}_{11}, \cdots, \boldsymbol{X}_{1m_1}, \cdots, \boldsymbol{X}_{s-1,1}, \cdots, \boldsymbol{X}_{s-1,m_{s-1}}$线性无关。令

$$\sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\boldsymbol{X}_{ij} = \boldsymbol{0}, \quad k_{ij} \in P, \tag{5}$$

则有

$$\boldsymbol{A}\Big(\sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\boldsymbol{X}_{ij}\Big) = \sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\lambda_i\boldsymbol{X}_{ij} = \boldsymbol{0},$$

$$\lambda_s\sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\boldsymbol{X}_{ij} = \sum_{i=1}^{s}\sum_{j=1}^{m_i} k_{ij}\lambda_s\boldsymbol{X}_{ij} = \boldsymbol{0}.$$

两式相减，得

$$\sum_{i=1}^{s-1}\sum_{j=1}^{m_i} k_{ij}(\lambda_i-\lambda_s)\boldsymbol{X}_{ij}=\boldsymbol{0}.$$

由归纳假设得$k_{ij}(\lambda_i-\lambda_s)=0$，$i=1,\cdots,s-1$，$j=1,\cdots,m_i$. 由于$\lambda_i\neq\lambda_s$，故$k_{ij}=0$，$i=1,\cdots,s-1$，$j=1,\cdots,m_i$. 代入(5)式，得

$$\sum_{j=1}^{m_s} k_{sj}\boldsymbol{X}_{sj}=\boldsymbol{0}.$$

又因为$\boldsymbol{X}_{s1},\cdots,\boldsymbol{X}_{sm_s}$线性无关，所以$k_{sj}=0$，$j=1,\cdots,m_s$. 故向量组$\boldsymbol{X}_{11},\cdots,\boldsymbol{X}_{1m_1},\boldsymbol{X}_{21},\cdots,\boldsymbol{X}_{2m_2},\cdots,\boldsymbol{X}_{s1},\cdots,\boldsymbol{X}_{sm_s}$线性无关。 □

设数域P上n阶矩阵$\boldsymbol{A}$有特征值$\lambda \in P$，相应的特征子空间记为V_λ，我们称V_λ的维数$\dim V_\lambda$为$\boldsymbol{A}$的特征值λ在P上的**几何重数**，

设数域P上n阶矩阵$\boldsymbol{A}$有特征值$\lambda \in P$，相应的特征子空间记为V_λ，我们称V_λ的维数$\dim V_\lambda$为$\boldsymbol{A}$的特征值λ在P上的**几何重数**，称特征值(根)作为$\boldsymbol{A}$的特征多项式的根的重数为$\boldsymbol{A}$的特征值λ在P上的**代数重数**。实际上，$\dim V_\lambda = n - r(\lambda \boldsymbol{E} - \boldsymbol{A})$.

设数域P上n阶矩阵$\boldsymbol{A}$有特征值$\lambda \in P$，相应的特征子空间记为V_λ，我们称V_λ的维数$\dim V_\lambda$为$\boldsymbol{A}$的特征值λ在P上的**几何重数**，称特征值(根)作为$\boldsymbol{A}$的特征多项式的根的重数为$\boldsymbol{A}$的特征值λ在P上的**代数重数**。实际上，$\dim V_\lambda = n - r(\lambda \boldsymbol{E} - \boldsymbol{A})$.

我们不加证明地给出

定理 5.5 设$\boldsymbol{A}$为数域P上n阶矩阵，则存在P上可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$为对角矩阵的充要条件是下列各式都成立：

(1) $\boldsymbol{A}$在P上有n个特征值(重根按重数计算);

(2) 对于$\boldsymbol{A}$在P上的每个特征根λ，其代数重数都等于几何重数。

例 1 设$\boldsymbol{A}=\begin{bmatrix}1 & 2 & 2\\ 2 & 1 & 2\\ 2 & 2 & 1\end{bmatrix}$，求可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$为对角阵。

例 1 设$\boldsymbol{A}=\begin{bmatrix}1&2&2\\2&1&2\\2&2&1\end{bmatrix}$，求可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$为对角阵。

解

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

所以$\boldsymbol{A}$在任意数域P上有特征值 5, -1(2重根).

例 1 设$\boldsymbol{A}=\begin{bmatrix}1&2&2\\2&1&2\\2&2&1\end{bmatrix}$，求可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$为对角阵。

解

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

所以$\boldsymbol{A}$在任意数域P上有特征值 5, -1(2重根).

将$\lambda=5$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得

$$\begin{bmatrix}4&-2&-2\\-2&4&-2\\-2&-2&4\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

由此可得属于$\lambda=5$的一个线性无关的特征向量$\begin{bmatrix}1\\1\\1\end{bmatrix}$.

将$\lambda=-1$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得

$$\begin{bmatrix}-2 & -2 & -2\\ -2 & -2 & -2\\ -2 & -2 & -2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

由此可得属于$\lambda=-1$的两个线性无关的特征向量$\begin{bmatrix}1\\ -1\\ 0\end{bmatrix},\ \begin{bmatrix}1\\ 0\\ -1\end{bmatrix}.$

将$\lambda=-1$代入$(\lambda \boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得

$$\begin{bmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

由此可得属于$\lambda=-1$的两个线性无关的特征向量$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$，$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$.

由定理5.4知上面三个特征向量线性无关。令$\boldsymbol{C}=\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}$，

则$\boldsymbol{C}$可逆，且

$$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}=\begin{bmatrix} 5 & & \\ & -1 & \\ & & -1 \end{bmatrix}.$$

□

例 2 设$\boldsymbol{A}=\begin{bmatrix}3 & 1 & 0\\ -4 & -1 & 0\\ 4 & -8 & -2\end{bmatrix}$，则$\boldsymbol{A}$不与对角矩阵相似。

例 2 设$\boldsymbol{A}=\begin{bmatrix}3 & 1 & 0\\ -4 & -1 & 0\\ 4 & -8 & -2\end{bmatrix}$，则$\boldsymbol{A}$不与对角矩阵相似。

证明 由于$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-3 & -1 & 0\\ 4 & \lambda+1 & 0\\ -4 & 8 & \lambda+2\end{vmatrix}=(\lambda+2)(\lambda-1)^2$,
则$\boldsymbol{A}$任意数域P上有特征值 -2，1 (2重根).

将$\lambda=-2$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得

$$\begin{bmatrix}-5 & -1 & 0\\ 4 & -1 & 0\\ -4 & 8 & 0\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

由此可得属于$\lambda=-2$的一个线性无关的特征向量$\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}$.

将$\lambda=1$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得

$$\begin{bmatrix}-2&-1&0\\4&2&0\\-4&8&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

由此可得属于$\lambda=1$的一个线性无关的特征向量$\begin{bmatrix}3\\-6\\20\end{bmatrix}$.

将$\lambda=1$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=0$，得

$$\begin{bmatrix}-2 & -1 & 0\\ 4 & 2 & 0\\ -4 & 8 & 3\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ x_3\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\end{bmatrix}.$$

由此可得属于$\lambda=1$的一个线性无关的特征向量$\begin{bmatrix}3\\ -6\\ 20\end{bmatrix}$.

由于$\boldsymbol{A}$在任意数域P上没有三个线性无关的特征向量，所以$\boldsymbol{A}$不与对角矩阵相似。

利用定理5.3和定理5.4，不难得到下列关于矩阵对角化(相似于对角矩阵)的充分条件。

推论 1 若数域P上n阶矩阵$\boldsymbol{A}$在P上有n个互异的特征值，则存在数域P上可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$为对角矩阵。

利用定理5.3和定理5.4，不难得到下列关于矩阵对角化(相似于对角矩阵)的充分条件。

推论 1 若数域P上n阶矩阵$\boldsymbol{A}$在P上有n个互异的特征值，则存在数域P上可逆矩阵$\boldsymbol{C}$，使$\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$为对角矩阵。

推论 2 若n阶复矩阵$\boldsymbol{A}$的特征多项式没有重根，则$\boldsymbol{A}$在复数域上与对角矩阵相似。

上述两个推论的逆命题不成立。

正交向量组

定义 5.4 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$是$n$维欧氏空间$\mathbf{R}^n$中非零向量。如果它们两两正交，则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为一个**正交向量组**，并规定由一个非零向量所组成的向量组是正交的。若每个$\boldsymbol{\alpha}_i$都是单位向量，则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为一个**单位正交向量组**。

正交向量组

定义 5.4 设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$是$n$维欧氏空间$\mathbf{R}^n$中非零向量。如果它们两两正交，则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为一个**正交向量组**，并规定由一个非零向量所组成的向量组是正交的。若每个$\boldsymbol{\alpha}_i$都是单位向量，则称$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为一个**单位正交向量组**。

特别地，如果$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$是$\mathbf{R}^n$的一个基，且为正交向量组，则称$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n\}$为$\mathbf{R}^n$的一个**正交基**。

正交向量组

定义 5.4 设$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$是$n$维欧氏空间$\mathbf{R}^n$中非零向量。如果它们两两正交，则称$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$为一个**正交向量组**，并规定由一个非零向量所组成的向量组是正交的。若每个$\boldsymbol{\alpha}_i$都是单位向量，则称$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$为一个**单位正交向量组**。

特别地，如果$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$是$\mathbf{R}^n$的一个基，且为正交向量组，则称$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$为$\mathbf{R}^n$的一个**正交基**。

又若$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$是一个正交基，且每个$\boldsymbol{\alpha}_i$都是单位向量，则称$\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n\}$是$\mathbf{R}^n$的一个**标准正交基(**或**规范正交基)**。

例 1 $\mathbf{R}^n$中基本向量组$\boldsymbol{\varepsilon}_1=(1,0,\cdots,0),\boldsymbol{\varepsilon}_2=(0,1,\cdots,0),\cdots,$
$\boldsymbol{\varepsilon}_n=(0,0,\cdots,1)$是$\mathbf{R}^n$的一个标准正交基。

例 1 $\mathbf{R}^n$中基本向量组$\boldsymbol{\varepsilon}_1 = (1, 0, \cdots, 0), \boldsymbol{\varepsilon}_2 = (0, 1, \cdots, 0), \cdots,$ $\boldsymbol{\varepsilon}_n = (0, 0, \cdots, 1)$是$\mathbf{R}^n$的一个标准正交基。

定理 5.6 正交向量组必是线性无关的。

正交矩阵

定义 5.5 设$\boldsymbol{A}$为n阶实矩阵，如果$\boldsymbol{A}$满足

$$\boldsymbol{A}'\boldsymbol{A}=\boldsymbol{E},$$

则称$\boldsymbol{A}$为**正交矩阵**。

正交矩阵

定义 5.5 设$\boldsymbol{A}$为n阶实矩阵，如果$\boldsymbol{A}$满足

$$\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{E},$$

则称$\boldsymbol{A}$为**正交矩阵**。

例如，单位矩阵$\boldsymbol{E}$是正交矩阵；平面解析几何中直角坐标变换公式

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$

对应的矩阵$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$也是正交矩阵。

定理 5.7 实n阶矩阵$\boldsymbol{A}$是正交矩阵的充要条件是它的列(行)向量组是n维欧氏空间$\mathbf{R}^n$的一个单位正交向量组。

定理 5.7 实n阶矩阵$\boldsymbol{A}$是正交矩阵的充要条件是它的列(行)向量组是n维欧氏空间$\mathbf{R}^n$的一个单位正交向量组。

由定理5.7不难看出，实n阶矩阵$\boldsymbol{A}$为正交矩阵的充要条件是它的列(行)向量组是n维欧氏空间$\mathbf{R}^n$的一个标准正交基。

施密特(Schmidt)正交化方法

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为线性无关的向量组，令

$$
\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_1,\\
\boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1,\\
\boldsymbol{\beta}_3 &= \boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2,\\
&\vdots\\
\boldsymbol{\beta}_m &= \boldsymbol{\alpha}_m-\frac{(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2-\cdots-\frac{(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_{m-1})}{(\boldsymbol{\beta}_{m-1},\boldsymbol{\beta}_{m-1})}\boldsymbol{\beta}_{m-1},
\end{aligned}
$$

则$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$是与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$等价的正交向量组。

施密特(Schmidt)正交化方法

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$为线性无关的向量组，令

$$\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_1,\\
\boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1,\\
\boldsymbol{\beta}_3 &= \boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2,\\
&\vdots\\
\boldsymbol{\beta}_m &= \boldsymbol{\alpha}_m-\frac{(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2-\cdots-\frac{(\boldsymbol{\alpha}_m,\boldsymbol{\beta}_{m-1})}{(\boldsymbol{\beta}_{m-1},\boldsymbol{\beta}_{m-1})}\boldsymbol{\beta}_{m-1},
\end{aligned}$$

则$\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_m$是与$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$等价的正交向量组。再单位化即得单位正交向量组

$$\gamma_1=\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|},\quad \gamma_2=\frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|},\quad \cdots\quad \gamma_m=\frac{\boldsymbol{\beta}_m}{|\boldsymbol{\beta}_m|}.$$

注 实际上，上述正交化方法是从非零向量$\boldsymbol{\alpha}_1$出发，令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\alpha}_2$出发找$\boldsymbol{\beta}_2$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$. 令$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + k\boldsymbol{\beta}_1$，则由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$可得$k = -\frac{(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}$, 故$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - \frac{(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}\boldsymbol{\beta}_1$.

注 实际上，上述正交化方法是从非零向量$\boldsymbol{\alpha}_1$出发，令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\alpha}_2$出发找$\boldsymbol{\beta}_2$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$. 令$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + k\boldsymbol{\beta}_1$，则由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$可得$k = -\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}$, 故$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - \frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\alpha}_3$出发找$\boldsymbol{\beta}_3$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_3) = 0, (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3) = 0$. 令$\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_2 + k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2$，由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_3) = (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3) = 0$得$k_1 = -\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}$, $k_2 = -\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}$.

注 实际上，上述正交化方法是从非零向量$\boldsymbol{\alpha}_1$出发，令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\alpha}_2$出发找$\boldsymbol{\beta}_2$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$. 令$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + k\boldsymbol{\beta}_1$，则由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$可得$k = -\frac{(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}$，故$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - \frac{(\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}\boldsymbol{\beta}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\alpha}_3$出发找$\boldsymbol{\beta}_3$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_3) = 0, (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3) = 0$. 令$\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_2 + k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2$，由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_3) = (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3) = 0$得$k_1 = -\frac{(\boldsymbol{\alpha}_3, \boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1, \boldsymbol{\beta}_1)}$，$k_2 = -\frac{(\boldsymbol{\alpha}_3, \boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2, \boldsymbol{\beta}_2)}$.

如此进行下去，后面逐个求得的每个向量$\boldsymbol{\beta}_i$和前面的正交向量组$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_{i-1}$中的每个向量都正交。

注 实际上，上述正交化方法是从非零向量$\boldsymbol{\alpha}_1$出发，令$\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\alpha}_2$出发找$\boldsymbol{\beta}_2$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$. 令$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + k\boldsymbol{\beta}_1$，则由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = 0$可得$k = -\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}$, 故$\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 - \frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1$.

再由$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\alpha}_3$出发找$\boldsymbol{\beta}_3$，使$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_3) = 0, (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3) = 0$. 令$\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_2 + k_1\boldsymbol{\beta}_1 + k_2\boldsymbol{\beta}_2$，由$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_3) = (\boldsymbol{\beta}_2, \boldsymbol{\beta}_3) = 0$得$k_1 = -\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}$, $k_2 = -\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}$.

如此进行下去，后面逐个求得的每个向量$\boldsymbol{\beta}_i$和前面的正交向量组$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_{i-1}$中的每个向量都正交。

例 2 把向量组$\boldsymbol{\alpha}_1 = (0, 1, 1)$，$\boldsymbol{\alpha}_2 = (1, 0, 1)$，$\boldsymbol{\alpha}_3 = (1, 1, 0)$化为单位正交向量组。

复n维向量空间的内积

在复n维向量空间$\mathbf{C}^n$中，考虑如下二元函数

$$(\boldsymbol{X},\boldsymbol{Y})=\boldsymbol{X}'\overline{\boldsymbol{Y}}=\overline{\boldsymbol{Y}}'\boldsymbol{X},\quad \boldsymbol{X},\boldsymbol{Y}\in\mathbf{C}^n.$$

利用矩阵的运算法则，不难证明二元函数$(\cdot,\cdot)$满足

(1) $(\boldsymbol{X},\boldsymbol{Y})=\overline{(\boldsymbol{Y},\boldsymbol{X})}$;

(2) $(\boldsymbol{X}_1+\boldsymbol{X}_2,\boldsymbol{Y})=(\boldsymbol{X}_1,\boldsymbol{Y})+(\boldsymbol{X}_2,\boldsymbol{Y})$,
$(\boldsymbol{X},\boldsymbol{Y}_1+\boldsymbol{Y}_2)=(\boldsymbol{X},\boldsymbol{Y}_1)+(\boldsymbol{X},\boldsymbol{Y}_2)$;

(3) $(\lambda\boldsymbol{X},\boldsymbol{Y})=\lambda(\boldsymbol{X},\boldsymbol{Y}),\quad (\boldsymbol{X},\lambda\boldsymbol{Y})=\bar{\lambda}(\boldsymbol{X},\boldsymbol{Y})$;

(4) 对任意向量$\boldsymbol{X}\in\mathbf{C}^n$, $(\boldsymbol{X},\boldsymbol{X})$为非负实数，且$(\boldsymbol{X},\boldsymbol{X})=0$当且仅当$\boldsymbol{X}=0$.

复n维向量空间的内积

在复n维向量空间$\mathbf{C}^n$中，考虑如下二元函数

$$(\boldsymbol{X},\boldsymbol{Y})=\boldsymbol{X}'\overline{\boldsymbol{Y}}=\overline{\boldsymbol{Y}}'\boldsymbol{X},\quad \boldsymbol{X},\boldsymbol{Y}\in\mathbf{C}^n.$$

利用矩阵的运算法则，不难证明二元函数$(\cdot,\cdot)$满足

(1) $(\boldsymbol{X},\boldsymbol{Y})=\overline{(\boldsymbol{Y},\boldsymbol{X})}$;

(2) $(\boldsymbol{X}_1+\boldsymbol{X}_2,\boldsymbol{Y})=(\boldsymbol{X}_1,\boldsymbol{Y})+(\boldsymbol{X}_2,\boldsymbol{Y})$,
$(\boldsymbol{X},\boldsymbol{Y}_1+\boldsymbol{Y}_2)=(\boldsymbol{X},\boldsymbol{Y}_1)+(\boldsymbol{X},\boldsymbol{Y}_2)$;

(3) $(\lambda\boldsymbol{X},\boldsymbol{Y})=\lambda(\boldsymbol{X},\boldsymbol{Y}),\quad (\boldsymbol{X},\lambda\boldsymbol{Y})=\bar{\lambda}(\boldsymbol{X},\boldsymbol{Y})$;

(4) 对任意向量$\boldsymbol{X}\in\mathbf{C}^n$, $(\boldsymbol{X},\boldsymbol{X})$为非负实数，且$(\boldsymbol{X},\boldsymbol{X})=0$当且仅当$\boldsymbol{X}=0$.

我们把上述二元函数称为$\mathbf{C}^n$中向量的**内积**，又把引入内积运算之后的复n维向量空间$\mathbf{C}^n$称为**酉空间**。

可以证明，对任意$\boldsymbol{X},\boldsymbol{Y}\in\mathbf{C}^n$和$n$阶实矩阵$\boldsymbol{A}$，都有

(1) $(\boldsymbol{AX},\boldsymbol{Y})=(\boldsymbol{X},\boldsymbol{A}'\boldsymbol{Y})$.

特别地，若$\boldsymbol{A}$为实对称矩阵，则有$(\boldsymbol{AX},\boldsymbol{Y})=(\boldsymbol{X},\boldsymbol{AY})$;

(2) 若$\boldsymbol{A}$为正交矩阵，则有$(\boldsymbol{AX},\boldsymbol{AY})=(\boldsymbol{X},\boldsymbol{Y})$.

可以证明，对任意$\boldsymbol{X},\boldsymbol{Y}\in\mathbf{C}^n$和$n$阶实矩阵$\boldsymbol{A}$，都有

(1) $(\boldsymbol{AX},\boldsymbol{Y})=(\boldsymbol{X},\boldsymbol{A}'\boldsymbol{Y})$.
特别地，若$\boldsymbol{A}$为实对称矩阵，则有$(\boldsymbol{AX},\boldsymbol{Y})=(\boldsymbol{X},\boldsymbol{AY})$;

(2) 若$\boldsymbol{A}$为正交矩阵，则有$(\boldsymbol{AX},\boldsymbol{AY})=(\boldsymbol{X},\boldsymbol{Y})$.

对于实对称矩阵有下面两个引理:

引理 1 实对称矩阵$\boldsymbol{A}$的特征值必为实数，并且对于它的每一个实特征值，$\boldsymbol{A}$都有实特征向量。

可以证明，对任意$\boldsymbol{X},\boldsymbol{Y}\in\mathbf{C}^n$和$n$阶实矩阵$\boldsymbol{A}$，都有

(1) $(\boldsymbol{AX},\boldsymbol{Y})=(\boldsymbol{X},\boldsymbol{A}'\boldsymbol{Y})$.
特别地，若$\boldsymbol{A}$为实对称矩阵，则有$(\boldsymbol{AX},\boldsymbol{Y})=(\boldsymbol{X},\boldsymbol{AY})$;

(2) 若$\boldsymbol{A}$为正交矩阵，则有$(\boldsymbol{AX},\boldsymbol{AY})=(\boldsymbol{X},\boldsymbol{Y})$.

对于实对称矩阵有下面两个引理：

引理 1 实对称矩阵$\boldsymbol{A}$的特征值必为实数，并且对于它的每一个实特征值，$\boldsymbol{A}$都有实特征向量。

引理 2 实对称矩阵$\boldsymbol{A}$属于不同特征值的(实)特征向量正交。

下面引理表明，任意一个单位正交向量组都可以扩充为$\mathbf{R}^n$的一个标准正交基。

引理 3 若$\boldsymbol{\eta}_1, \boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_s$ $(s \leq n)$是任意一个单位正交向量组，则可以用这s个向量为前s列作出一个n阶正交矩阵。

下面引理表明，任意一个单位正交向量组都可以扩充为$\mathbf{R}^n$的一个标准正交基。

引理 3 若$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_s$ $(s\le n)$是任意一个单位正交向量组，则可以用这s个向量为前s列作出一个n阶正交矩阵。

证明* 令矩阵$\boldsymbol{A}=(\boldsymbol{\eta}_1',\boldsymbol{\eta}_2',\cdots,\boldsymbol{\eta}_s')^\top$, 则$r(\boldsymbol{A})=s$.

不妨设$s<n$，则齐次线性方程组$\boldsymbol{AX}=0$有基础解系，且基础解系令有$n-s$个解向量，记为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-s}$. 把$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-s}$正交化再单位化，得到与之等价的单位正交向量组$\boldsymbol{\eta}_{s+1},\boldsymbol{\eta}_{s+2},\cdots,\boldsymbol{\eta}_n$. 又因为$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{n-s}$中每个向量都与$\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\cdots,\boldsymbol{\eta}_s$正交，所以$\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_s,\boldsymbol{\eta}_{s+1},\cdots,\boldsymbol{\eta}_n$为单位正交向量组，从而得到一个$n$阶正交矩阵$(\boldsymbol{\eta}_1,\cdots,\boldsymbol{\eta}_s,\boldsymbol{\eta}_{s+1},\cdots,\boldsymbol{\eta}_n)$. □

定理 5.8 对任意n阶实对称矩阵$\boldsymbol{A}$，总可以找到n阶正交矩阵$\boldsymbol{T}$，使$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}$为对角形，即

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}'\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

其中$\lambda_1,\lambda_2,\cdots,\lambda_n$为$\boldsymbol{A}$的全体特征值。

例 3 设$\boldsymbol{A}=\begin{bmatrix} 2 & -1 & -1 & 1 \\ -1 & 2 & 1 & -1 \\ -1 & 1 & 2 & -1 \\ 1 & -1 & -1 & 2 \end{bmatrix}$，求正交矩阵$\boldsymbol{T}$，使$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$为对角矩阵。

例 3 设 $\boldsymbol{A}=\begin{bmatrix} 2 & -1 & -1 & 1 \\ -1 & 2 & 1 & -1 \\ -1 & 1 & 2 & -1 \\ 1 & -1 & -1 & 2 \end{bmatrix}$，求正交矩阵$\boldsymbol{T}$，使$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$为对角矩阵。

解 由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix} \lambda-2 & 1 & 1 & -1 \\ 1 & \lambda-2 & -1 & 1 \\ 1 & -1 & \lambda-2 & 1 \\ -1 & 1 & 1 & \lambda-2 \end{vmatrix}=(\lambda-1)^3(\lambda-5),$$

得$\boldsymbol{A}$的特征值为1(三重), 5.

例 3 设$\boldsymbol{A}=\begin{bmatrix}2&-1&-1&1\\-1&2&1&-1\\-1&1&2&-1\\1&-1&-1&2\end{bmatrix}$，求正交矩阵$\boldsymbol{T}$，使$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$为对角矩阵。

解 由

$$|\lambda\boldsymbol{E}-\boldsymbol{A}|=\begin{vmatrix}\lambda-2&1&1&-1\\1&\lambda-2&-1&1\\1&-1&\lambda-2&1\\-1&1&1&\lambda-2\end{vmatrix}=(\lambda-1)^3(\lambda-5),$$

得$\boldsymbol{A}$的特征值为1(三重), 5.

将$\lambda=1$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得

$$\begin{bmatrix}-1&1&1&-1\\1&-1&-1&1\\1&-1&-1&1\\-1&1&1&-1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix},$$

它的一个基础解系为

$$\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{\alpha}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}.$$

它的一个基础解系为

$$\boldsymbol{\alpha}_1=\begin{bmatrix}1\\1\\0\\0\end{bmatrix},\quad \boldsymbol{\alpha}_2=\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\quad \boldsymbol{\alpha}_3=\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}.$$

正交化，得

$$\begin{aligned}
\boldsymbol{\beta}_1 &= \boldsymbol{\alpha}_1=\begin{bmatrix}1\\1\\0\\0\end{bmatrix},\\
\boldsymbol{\beta}_2 &= \boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\frac{1}{2}\boldsymbol{\beta}_1=\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\1\\0\end{bmatrix},\\
\boldsymbol{\beta}_3 &= \boldsymbol{\alpha}_3-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2=\boldsymbol{\alpha}_3-\frac{1}{2}\boldsymbol{\beta}_1-\frac{1}{3}\boldsymbol{\beta}_2=\begin{bmatrix}\frac{1}{3}\\-\frac{1}{3}\\-\frac{1}{3}\\-1\end{bmatrix}.
\end{aligned}$$

再单位化，得

$$\gamma_1=\frac{\beta_1}{|\beta_1|}=\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\\ 0\\ 0\end{bmatrix},\quad \gamma_2=\frac{\beta_2}{|\beta_2|}=\begin{bmatrix}\frac{\sqrt{6}}{6}\\ -\frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\\ 0\end{bmatrix},\quad \gamma_3=\frac{\beta_3}{|\beta_3|}=\begin{bmatrix}\frac{\sqrt{3}}{6}\\ -\frac{\sqrt{3}}{6}\\ -\frac{\sqrt{3}}{6}\\ -\frac{\sqrt{3}}{2}\end{bmatrix}.$$

再单位化，得

$$\boldsymbol{\gamma}_1=\frac{\boldsymbol{\beta}_1}{|\boldsymbol{\beta}_1|}=\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}\\ 0\\ 0\end{bmatrix},\quad \boldsymbol{\gamma}_2=\frac{\boldsymbol{\beta}_2}{|\boldsymbol{\beta}_2|}=\begin{bmatrix}\frac{\sqrt{6}}{6}\\ -\frac{\sqrt{6}}{6}\\ \frac{\sqrt{6}}{3}\\ 0\end{bmatrix},\quad \boldsymbol{\gamma}_3=\frac{\boldsymbol{\beta}_3}{|\boldsymbol{\beta}_3|}=\begin{bmatrix}\frac{\sqrt{3}}{6}\\ -\frac{\sqrt{3}}{6}\\ -\frac{\sqrt{3}}{6}\\ -\frac{\sqrt{3}}{2}\end{bmatrix}.$$

将$\lambda=5$代入$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$，得

$$\begin{bmatrix}3&1&1&-1\\1&3&-1&1\\1&-1&3&1\\-1&1&1&3\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

它的一个基础解系为$\boldsymbol{\alpha}_4=\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix}$. 单位化得$\boldsymbol{\gamma}_4=\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$.

令

$$\boldsymbol{T}=(\gamma_1,\gamma_2,\gamma_3,\gamma_4)=\begin{bmatrix}\frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} & \frac{\sqrt{3}}{6} & \frac{1}{2}\\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} & -\frac{\sqrt{3}}{6} & -\frac{1}{2}\\ 0 & \frac{\sqrt{6}}{3} & -\frac{\sqrt{3}}{6} & -\frac{1}{2}\\ 0 & 0 & -\frac{\sqrt{3}}{2} & \frac{1}{2}\end{bmatrix},$$

则$\boldsymbol{T}$为正交矩阵，且

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 5\end{bmatrix}.$$

□

注 对于特征值$\lambda = 1$，通过观察可知正交向量组

$$\boldsymbol{\eta}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{\eta}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad \boldsymbol{\eta}_3 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

是$(\lambda \boldsymbol{E} - \boldsymbol{A})\boldsymbol{X} = \mathbf{0}$的一个基础解系，将它们单位化，得$\frac{1}{2}\boldsymbol{\eta}_1, \frac{1}{2}\boldsymbol{\eta}_2, \frac{1}{2}\boldsymbol{\eta}_3$.

注 对于特征值$\lambda=1$，通过观察可知正交向量组

$$\boldsymbol{\eta}_1=\begin{bmatrix}1\\1\\1\\1\end{bmatrix},\quad \boldsymbol{\eta}_2=\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix},\quad \boldsymbol{\eta}_3=\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}$$

是$(\lambda\boldsymbol{E}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$的一个基础解系，将它们单位化，得$\frac{1}{2}\boldsymbol{\eta}_1,\frac{1}{2}\boldsymbol{\eta}_2,\frac{1}{2}\boldsymbol{\eta}_3$.

再结合特征值$\lambda=5$的单位特征向量$\frac{1}{2}\boldsymbol{\eta}_4$ $(\boldsymbol{\eta}_4=\boldsymbol{\alpha}_4)$，可得到正交矩阵

$$\boldsymbol{Q}=\frac{1}{2}(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3,\boldsymbol{\eta}_4)=\frac{1}{2}\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix},$$

使

$$\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\begin{bmatrix}1&&&\\&1&&\\&&1&\\&&&5\end{bmatrix}.$$

小结

求正交矩阵$\boldsymbol{T}$，将实对称矩阵$\boldsymbol{A}$对角化的一般步骤:

1. 首先求$\boldsymbol{A}$的特征值;
2. 求每个特征值对应的一组线性无关的特征向量，即$(\lambda \boldsymbol{E} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$的基础解系;
3. 将每个特征值对应的特征向量组正交化并单位化，便得到对应的单位正交向量组;
4. 将每个特征值对应的单位正交向量组按列组成一个矩阵，即为正交矩阵$\boldsymbol{T}$;
5. 利用正交矩阵$\boldsymbol{T}$将$\boldsymbol{A}$对角化，即

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \boldsymbol{T}'\boldsymbol{A}\boldsymbol{T} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

其中$\lambda_1, \lambda_2, \cdots, \lambda_n$为$\boldsymbol{A}$的全体特征值。

若尔当(Jordan)矩阵

定义 5.6 设λ_i为复数，形如

$$\boldsymbol{J}_i=\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{bmatrix}$$

的m_i阶矩阵($m_i \geq 1$)叫做一个**若尔当块**。由若尔当块$\boldsymbol{J}_i$组成的准对角矩阵

$$\begin{bmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s\end{bmatrix}$$

称为一个**若尔当矩阵**。

若尔当(Jordan)矩阵

定义 5.6 设λ_i为复数，形如

$$\boldsymbol{J}_i=\begin{bmatrix}\lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i\end{bmatrix}$$

的m_i阶矩阵($m_i \geq 1$)叫做一个**若尔当块**。由若尔当块$\boldsymbol{J}_i$组成的准对角矩阵

$$\begin{bmatrix}\boldsymbol{J}_1 & & & \\ & \boldsymbol{J}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{J}_s\end{bmatrix}$$

称为一个**若尔当矩阵**。

显然，对角矩阵是若尔当矩阵的特殊情况。

例 1 $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$，$\begin{bmatrix} \sqrt{2} & 1 & 0 \\ 0 & \sqrt{2} & 1 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$，$\begin{bmatrix} \mathrm{i} & 1 \\ 0 & \mathrm{i} \end{bmatrix}$都是若尔当块，而$\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$，$\begin{bmatrix} \sqrt{2} & 1 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$则不是若尔当块。

例 1 $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} \sqrt{2} & 1 & 0 \\ 0 & \sqrt{2} & 1 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$, $\begin{bmatrix} \mathrm{i} & 1 \\ 0 & \mathrm{i} \end{bmatrix}$都是若尔当块，而$\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} \sqrt{2} & 1 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$则不是若尔当块。

例 2 $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} \sqrt{2} & 1 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & \mathrm{i} & 1 \\ 0 & 0 & 0 & 0 & \mathrm{i} \end{bmatrix}$都是若尔当矩阵，

而 $\begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & \mathrm{i} & 1 & 0 \\ 0 & 0 & 0 & \mathrm{i} & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$则不是若尔当矩阵。

下面的定理通常称为若尔当定理。

定理 5.9 任意一个复方阵$\boldsymbol{A}$都相似于一个若尔当矩阵，且该若尔当矩阵在各个若尔当块可以相差一个次序的意义下由$\boldsymbol{A}$唯一确定。因此，我们把这个若尔当矩阵$\boldsymbol{J}$称为复矩阵$\boldsymbol{A}$的**若尔当标准形**。

下面的定理通常称为若尔当定理。

定理 5.9 任意一个复方阵$\boldsymbol{A}$都相似于一个若尔当矩阵，且该若尔当矩阵在各个若尔当块可以相差一个次序的意义下由$\boldsymbol{A}$唯一确定。因此，我们把这个若尔当矩阵$\boldsymbol{J}$称为复矩阵$\boldsymbol{A}$的**若尔当标准形**。

显然，复矩阵$\boldsymbol{A}$相似于对角矩阵的充要条件是$\boldsymbol{A}$的若尔当标准形为对角矩阵。

例 3 我们在5.2节例2中已证明$\boldsymbol{A}=\begin{bmatrix}3 & 1 & 0\\ -4 & -1 & 0\\ 4 & -8 & -2\end{bmatrix}$不相似于对角矩阵。取

$$\boldsymbol{T}=\begin{bmatrix}0 & 3 & \frac{3}{2}\\ 0 & -6 & 0\\ 1 & 20 & -\frac{14}{3}\end{bmatrix},$$

则$\boldsymbol{T}$可逆，且$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}-2 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$.

例 3 我们在5.2节例2中已证明$\boldsymbol{A}=\begin{bmatrix}3 & 1 & 0\\ -4 & -1 & 0\\ 4 & -8 & -2\end{bmatrix}$不相似于对角矩阵。取

$$\boldsymbol{T}=\begin{bmatrix}0 & 3 & \frac{3}{2}\\ 0 & -6 & 0\\ 1 & 20 & -\frac{14}{3}\end{bmatrix},$$

则$\boldsymbol{T}$可逆，且$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}-2 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}$.

例 4 令$\boldsymbol{A}=\begin{bmatrix}-2 & 1 & 0\\ -1 & 0 & 0\\ -1 & -1 & 1\end{bmatrix}$, $\boldsymbol{T}=\begin{bmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$，则$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$为$\boldsymbol{A}$的若尔当标准形。

Thanks for your attention!