第5讲、特征值与特征向量

May 31, 2017

1 矩阵的特征值与特征向量 相似矩阵

1.1 特征值与特征向量

定义 1 (特征值与特征向量)**.** 设$\mathbf{A}$为复数域$\mathbb{C}$上的n阶矩阵，如果存在数$\lambda \in \mathbb{C}$和非零的n维向量x使得

$$\mathbf{A}x = \lambda x$$

则称λ为矩阵$\mathbf{A}$的特征值，x为$\mathbf{A}$的对应于特征值λ的特征向量。

(1) 特征向量$x \neq 0$；

(2) 特征值问题是对方针而言的。

由定义，n阶矩阵$\mathbf{A}$的特征值，就是使齐次线性方程组

$$(\mathbf{A} - \lambda \mathbf{I})x = 0$$

有非零解的λ值，即满足方程

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

的λ都是矩阵$\mathbf{A}$的特征值。

结论 1. 特征值λ是关于λ的多项式$\det(\mathbf{A} - \lambda \mathbf{I})$的根。

定义 2 (特征多项式、特征矩阵、特征方程)**.** 设n阶矩阵$\mathbf{A} = (a_{ij})$，则

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

称为矩阵$\mathbf{A}$的特征多项式，$\mathbf{A} - \lambda \mathbf{I}$称为$\mathbf{A}$的特征矩阵，$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$称为$\mathbf{A}$的特征方程。

(1) n阶矩阵$\mathbf{A}$的特征多项式是λ的n次多项式。

(2) 特征多项式的k重根称为k重特征值。

例 1. 求矩阵

$$\mathbf{A} = \begin{pmatrix} 5 & -1 & -1 \\ 3 & 1 & -1 \\ 4 & -2 & 1 \end{pmatrix}$$

的特征值与特征向量。

解.

$$\begin{aligned}
\det(\mathbf{A}-\lambda\mathbf{I}) &= \begin{vmatrix} 5-\lambda & -1 & -1 \\ 3 & 1-\lambda & -1 \\ 4 & -2 & 1-\lambda \end{vmatrix} = \begin{vmatrix} 3-\lambda & -1 & -1 \\ 3-\lambda & 1-\lambda & -1 \\ 3-\lambda & -2 & 1-\lambda \end{vmatrix} \\
&= (3-\lambda)\begin{vmatrix} 1 & -1 & -1 \\ 1 & 1-\lambda & -1 \\ 1 & -2 & 1-\lambda \end{vmatrix} = (3-\lambda)\begin{vmatrix} 1 & -1 & -1 \\ 0 & 2-\lambda & 0 \\ 0 & -1 & 2-\lambda \end{vmatrix} \\
&= (3-\lambda)(\lambda-2)^2 = 0
\end{aligned}$$

故$\mathbf{A}$的特征值为$\lambda_1 = 3,\ \lambda_2 = 2$（二重特征值）。

当$\lambda_1 = 3$时，由$(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = \mathbf{0}$，即

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

得其基础解系为$\mathbf{x}_1 = (1,1,1)^{\mathsf{T}}$，因此$k_1\mathbf{x}_1$（$k_1$为非零任意常数）是$\mathbf{A}$对应于$\lambda_1 = 3$的全部特征向量。

当$\lambda_1 = 2$时，由$(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = \mathbf{0}$，即

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

得其基础解系为$\mathbf{x}_2 = (1,1,2)^{\mathsf{T}}$，因此$k_2\mathbf{x}_2$（$k_1$为非零任意常数）是$\mathbf{A}$对应于$\lambda_1 = 2$的全部特征向量。

例 2.

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}, \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

的特征多项式为

$$(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$$

故其n个特征值为n个对角元。

1.2 特征值与特征值的性质

定理 1. 若$\mathbf{x}_1$和$\mathbf{x}_2$都是$\mathbf{A}$的对应于特征值λ_0的特征向量，则$k_1\mathbf{x}_1 + k_2\mathbf{x}_2$也是$\mathbf{A}$的对应于特征值$\lambda_0$的特征向量（其中$k_1, k_2$为任意常数，但$k_1\mathbf{x}_1 + k_2\mathbf{x}_2 \neq 0$）。

证明. 由于$\mathbf{x}_1$和$\mathbf{x}_2$是齐次线性方程组

$$(\mathbf{A}-\lambda_0\mathbf{I})\mathbf{x} = 0$$

的解，因此$k_1\mathbf{x}_1 + k_2\mathbf{x}_2$也是上式的解，故当$k_1\mathbf{x}_1 + k_2\mathbf{x}_2 \neq 0$时，是$\mathbf{A}$的属于$\lambda_0$的特征向量。 □

在$(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = 0$的解空间中，除零向量以外的全体解向量就是$\mathbf{A}$的属于特征值$\lambda$的全体特征向量。因此，$(\mathbf{A}-\lambda\mathbf{I})\mathbf{x} = 0$的解空间也称为$\mathbf{A}$关于特征值$\lambda$的特征子空间，记作$V_\lambda$。$n$阶矩阵$\mathbf{A}$的特征子空间就是$n$维向量空间的子空间，其维数为

$$\mathrm{dim} V_\lambda = n - \mathrm{rank}(\mathbf{A}-\lambda\mathbf{I}).$$

需要注意的是，n维实矩阵的特征值可能是复数，所以特征子空间一般是n维复向量空间$\mathbb{C}^n$的子空间。

上例中，矩阵$\mathbf{A} = \begin{pmatrix} 5 & -1 & -1 \\ 3 & 1 & -1 \\ 4 & -2 & 1 \end{pmatrix}$的两个特征子空间为

$$\begin{aligned}
V_{\lambda_1} &= \{k\mathbf{x} | \mathbf{x} = (1,1,1)^{\mathsf{T}}, k \in \mathbb{C}\}, \\
V_{\lambda_2} &= \{k\mathbf{x} | \mathbf{x} = (1,1,2)^{\mathsf{T}}, k \in \mathbb{C}\}.
\end{aligned}$$

定理 2. 设n阶矩阵$\mathbf{A}=(a_{ij})$的n个特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，则

(1) $\sum_{i=1}^{n}\lambda_i=\sum_{i=1}^{n}a_{ii}$;

(2) $\prod_{i=1}^{n}\lambda_i=\det(\mathbf{A})$,

其中$\sum_{i=1}^{n}a_{ii}$是$\mathbf{A}$的主对角元之和，称为$\mathbf{A}$的迹*(trace)*，记为$\mathrm{tr}(\mathbf{A})$。

证明. 设

$$\det(\mathbf{A}-\lambda\mathbf{I})=\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}=\lambda^n+c_1\lambda^{n-1}+c_2\lambda^{n-2}+\cdots+c_{n-1}\lambda+c_n,$$

展开后含λ^{n-1}项的行列式有下面n个

$$\begin{vmatrix} a_{11} & & & & \\ a_{21} & -\lambda & & & \\ a_{31} & & -\lambda & & \\ \vdots & & & \ddots & \\ a_{n1} & & & & -\lambda \end{vmatrix},\begin{vmatrix} -\lambda & a_{12} & & & \\ & a_{22} & & & \\ & a_{32} & -\lambda & & \\ & \vdots & & \ddots & \\ & a_{n2} & & & -\lambda \end{vmatrix},\cdots,\begin{vmatrix} -\lambda & & & & a_{1n} \\ & -\lambda & & & a_{2n} \\ & & -\lambda & & a_{3n} \\ & & & \ddots & \vdots \\ & & & & a_{nn} \end{vmatrix}$$

它们之和等于

$$(a_{11}+a_{22}+\cdots+a_{nn})(-\lambda)^{n-1}=(-1)^{n-1}\sum_{i=1}^{n}a_{ii}\lambda^{n-1},$$

即$c_1=(-1)^{n-1}\sum_{i=1}^{n}a_{ii}$。

展开后常数项为

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}=\det(\mathbf{A}),$$

即$c_n=\det(\mathbf{A})$。

假设$\mathbf{A}$的n个特征值为$\lambda_1,\lambda_2,\cdots,\lambda_n$，根据$n$次多项式的根与系数的关系，得

$$(-1)^{n-1}\sum_{i=1}^{n}\lambda_i=c_1=(-1)^{n-1}\sum_{i=1}^{n}a_{ii},$$
$$\prod_{i=1}^{n}\lambda_i=c_n=\det(\mathbf{A}),$$

故

$$\det(\mathbf{A})=\prod_{i=1}^{n}\lambda_i.$$

□

注 1.
- 当$\det(\mathbf{A})\neq 0$，即$\mathbf{A}$为可逆矩阵时，其特征值全为非零数；
- 奇异矩阵$\mathbf{A}$至少有一个零特征值。

定理 3. 一个特征向量不能属于不同的特征值。

证明. 若$\mathbf{x}$是$\mathbf{A}$的属于特征值$\lambda_1,\lambda_2(\lambda_1\neq\lambda_2)$的特征向量， 即有

$$\mathbf{A}\mathbf{x}=\lambda_1\mathbf{x},\quad \mathbf{A}\mathbf{x}=\lambda_2\mathbf{x}\quad\Rightarrow\quad(\lambda_1-\lambda_2)\mathbf{x}=0\quad\Rightarrow\quad \mathbf{x}=0$$

这与$\mathbf{x}\neq 0$矛盾。 □

性质 1. 若λ是矩阵$\mathbf{A}$的特征值，$\mathbf{x}$是$\mathbf{A}$属于λ的特征向量，则

(i) $k\lambda$是$k\mathbf{A}$的特征值；

(ii) λ^m是$\mathbf{A}^m$的特征值；

(iii) 当$\mathbf{A}$可逆时，λ^{-1}是$\mathbf{A}^{-1}$的特征值；

且$\mathbf{x}$仍是矩阵$k\mathbf{A},\mathbf{A}^m,\mathbf{A}^{-1}$分别对应于$k\lambda,\lambda^m,\lambda^{-1}$的特征向量。

证明. (i)

(ii)

(iii) 当$\mathbf{A}$可逆时，$\lambda\neq 0$，由$\mathbf{A}\mathbf{x}=\lambda\mathbf{x}$可得

$$\mathbf{A}^{-1}(\mathbf{A}\mathbf{x})=\mathbf{A}^{-1}(\lambda\mathbf{x})=\lambda\mathbf{A}^{-1}\mathbf{x},$$

因此

$$\mathbf{A}^{-1}\mathbf{x}=\lambda^{-1}\mathbf{x},$$

故λ^{-1}是$\mathbf{A}^{-1}$的特征值，且$\mathbf{x}$也是$\mathbf{A}^{-1}$对应于λ^{-1}的特征向量。

□

注. 若λ是$\mathbf{A}$的特征值，则$\varphi(\lambda)$是$\varphi(\mathbf{A})$的特征值，其中$\varphi(\lambda)=a_0+a_1\lambda+\cdots+a_m\lambda^m,\varphi(\mathbf{A})=a_0\mathbf{I}+a_1\mathbf{A}+\cdots+a_m\mathbf{A}^m$.

例 3. 设3阶矩阵$\mathbf{A}$的特征值为$1,-1,2$，求$|\mathbf{A}^*+3\mathbf{A}-2\mathbf{I}|$.

解. 因$\mathbf{A}$的特征值全不为零，故$\mathbf{A}$可逆，从而$\mathbf{A}^*=|\mathbf{A}|\mathbf{A}^{-1}$。又因$|\mathbf{A}|=\lambda_1\lambda_2\lambda_3=-2$，故

$$\mathbf{A}^*+3\mathbf{A}-2\mathbf{I}=-2\mathbf{A}^{-1}+3\mathbf{A}-2\mathbf{I}.$$

令$\varphi(\lambda)=-\frac{2}{\lambda}+3\lambda-2$，则$\varphi(\lambda)$为上述矩阵的特征值，分别为$\varphi(1)=-1,\varphi(-1)=-3,\varphi(2)=3$，于是

$$|\mathbf{A}^*+3\mathbf{A}-2\mathbf{I}|=(-1)\cdot(-3)\cdot 3=9.$$

性质 2. 矩阵$\mathbf{A}$与$\mathbf{A}^\mathsf{T}$的特征值相同。

证明. 因$(\mathbf{A}-\lambda\mathbf{I})^\mathsf{T}=\mathbf{A}^\mathsf{T}-(\lambda\mathbf{I})^\mathsf{T}=\mathbf{A}^\mathsf{T}-\lambda\mathbf{I}$，故

$$\det(\mathbf{A}-\lambda\mathbf{I})=\det(\mathbf{A}^\mathsf{T}-\lambda\mathbf{I}),$$

故$\mathbf{A}$与$\mathbf{A}^\mathsf{T}$有完全相同的特征值。 □

例 4. 设$\mathbf{A}=\begin{pmatrix}1&-1&1\\2&-2&2\\-1&1&-1\end{pmatrix}$

(i) 求$\mathbf{A}$的特征值与特征向量

(ii) 求可逆矩阵$\mathbf{P}$，使得$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$为对角阵。

解. 由

$$|\mathbf{A}-\lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & -1 & 1 \\ 2 & -2-\lambda & 2 \\ -1 & 1 & -1-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 2 & -\lambda & 2 \\ -1 & -\lambda & -1-\lambda \end{vmatrix}$$

$$= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 2 & -\lambda & 2 \\ -3 & 0 & -3-\lambda \end{vmatrix} = -\lambda[(\lambda-1)(\lambda+3)+3] = -\lambda^2(\lambda+2),$$

知$\mathbf{A}$的特征值为$\lambda_1=\lambda_2=0$和$\lambda_3=-2$.

当$\lambda_{1,2}=0$时，由$(\mathbf{A}-0\mathbf{I})\mathbf{x}=0$，即$\mathbf{A}\mathbf{x}=0$得基础解系

$$\mathbf{x}_1=(1,1,0)^{\mathsf{T}}, \quad \mathbf{x}_2=(-1,0,1)^{\mathsf{T}},$$

故$\mathbf{A}$对应于$\lambda_{1,2}=0$的全体特征向量为$k_1\mathbf{x}_1+k_2\mathbf{x}_2=k_1(1,1,0)^{\mathsf{T}}+k_2(-1,0,1)^{\mathsf{T}}$，其中$k_1,k_2$为不全为零的任意常数。

当$\lambda_3=-2$时，由$(\mathbf{A}-\lambda_3\mathbf{I})\mathbf{x}=0$，即

$$\begin{pmatrix} 3 & -1 & 1 \\ 2 & 0 & 2 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

得基础解系

$$\mathbf{x}_3=(-1,-2,1)^{\mathsf{T}},$$

故$\mathbf{A}$对应于$\lambda_3=-2$的全体特征向量为$k_3\mathbf{x}_3=k_3(-1,-2,1)^{\mathsf{T}}$，其中$k_3$为非零的任意常数。

将$\mathbf{A}\mathbf{x}_i=\lambda_i\mathbf{x}_i(i=1,2,3)$表示成

$$\mathbf{A}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)=(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)\begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$$

取

$$\mathbf{P}=(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3)=\begin{pmatrix} 1 & -1 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{\Lambda}=\begin{pmatrix} 0 & & \\ & 0 & \\ & & -2 \end{pmatrix},$$

则$\mathbf{AP}=\mathbf{P\Lambda}$，且$|\mathbf{P}|=2\neq 0$，故得

$$\mathbf{P}^{-1}\mathbf{AP}=\mathbf{\Lambda}$$

为对角阵。

定理 4. 设$\lambda_1,\lambda_2,\cdots,\lambda_m$是方阵$\mathbf{A}$的$m$个特征值，$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_m$依次是与之对应的特征向量，若$\lambda_1,\lambda_2,\cdots,\lambda_m$互不相等，则$\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_m$线性无关。

证明. 设有常数$k_1,k_2,\cdots,k_m$使得

$$k_1\mathbf{x}_1+k_2\mathbf{x}_2+\cdots+k_m\mathbf{x}_m=0,$$

则$\mathbf{A}(k_1\mathbf{x}_1+k_2\mathbf{x}_2+\cdots+k_m\mathbf{x}_m)=0$，即

$$\lambda_1k_1\mathbf{x}_1+\lambda_2k_2\mathbf{x}_2+\cdots+\lambda_mk_m\mathbf{x}_m=0,$$

以此类推，有

$$\lambda_1^lk_1\mathbf{x}_1+\lambda_2^lk_2\mathbf{x}_2+\cdots+\lambda_m^lk_m\mathbf{x}_m=0, \quad l=1,2,\cdots,m-1.$$

写成矩阵形式即为

$$(k_1\mathbf{x}_1,k_2\mathbf{x}_2,\cdots,k_m\mathbf{x}_m)\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{m-1} \end{pmatrix}=(0,0,\cdots,0).$$

上式左边的第二个矩阵的行列式为范德蒙德行列式，当λ_i互不相等时该行列式不为零，从而该矩阵可逆。于是有

$$(k_1\boldsymbol{x}_1, k_2\boldsymbol{x}_2, \cdots, k_m\boldsymbol{x}_m) = (0, 0, \cdots, 0),$$

即$k_j\boldsymbol{x}_j = 0,\quad j = 1, 2, \cdots, m$。但$\boldsymbol{x}_j \neq 0$，故$k_j = 0,\quad j = 1, 2, \cdots, m$，从而$\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$线性无关。 □

例 5. 设λ_1和λ_2是矩阵$\boldsymbol{A}$的两个不同特征值，对应的特征向量依次为$\boldsymbol{x}_1, \boldsymbol{x}_2$，证明$\boldsymbol{x}_1 + \boldsymbol{x}_2$不是$\boldsymbol{A}$的特征向量。

证明. 按题设，有$\boldsymbol{A}\boldsymbol{x}_1 = k_1\boldsymbol{x}_1$，$\boldsymbol{A}\boldsymbol{x}_2 = k_2\boldsymbol{x}_2$，故

$$A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2.$$

（反证法）假设$\boldsymbol{x}_1 + \boldsymbol{x}_2$是$\boldsymbol{A}$的特征向量，则应存在$\lambda$使得$\boldsymbol{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2)$，于是

$$\lambda(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2,$$

即

$$(\lambda_1 - \lambda)\boldsymbol{x}_1 + (\lambda_2 - \lambda)\boldsymbol{x}_2 = 0.$$

因$\lambda_1 \neq \lambda_2$，由上述定理知$\boldsymbol{x}_1, \boldsymbol{x}_2$线性无关，从而有

$$\lambda_1 - \lambda = \lambda_2 - \lambda,$$

即$\lambda_1 = \lambda_2$，与题设矛盾，从而$\boldsymbol{x}_1 + \boldsymbol{x}_2$不是$\boldsymbol{A}$的特征向量。 □

2 相似矩阵

定义 3. 设$\boldsymbol{A}, \boldsymbol{B}$为$n$阶矩阵，若存在可逆矩阵$\boldsymbol{P}$，使得

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B},$$

则称$\boldsymbol{B}$是$\boldsymbol{A}$的相似矩阵，或$\boldsymbol{A}$与$\boldsymbol{B}$相似，记为$\boldsymbol{A} \sim \boldsymbol{B}$。对$\boldsymbol{A}$进行运算$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$称为对$\boldsymbol{A}$进行相似变化，可逆矩阵$\boldsymbol{P}$称为把$\boldsymbol{A}$变成$\boldsymbol{B}$的相似变换矩阵。

定理 5. 若$\boldsymbol{A} \sim \boldsymbol{B}$，则$\boldsymbol{A}$与$\boldsymbol{B}$的$u$特征多项式相同，从而$\boldsymbol{A}$与$\boldsymbol{B}$的特征值相同。

证明. 因$\boldsymbol{A} \sim \boldsymbol{B}$，即有可逆阵$\boldsymbol{P}$使得$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$。故

$$|B - \lambda\boldsymbol{I}| = |\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} - \boldsymbol{P}^{-1}(\lambda\boldsymbol{I})\boldsymbol{P}| = |\boldsymbol{P}^{-1}(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{P}| = |\boldsymbol{P}^{-1}||\boldsymbol{A} - \lambda\boldsymbol{I}||\boldsymbol{P}| = |\boldsymbol{A} - \lambda\boldsymbol{I}|.$$

□

推论 1. 若$\boldsymbol{A}$与对角阵

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

相似，则$\lambda_1, \lambda_2, \cdots, \lambda_n$为$\boldsymbol{A}$的$n$个特征值。

3 矩阵可对角化的条件

矩阵可对角化，即矩阵与对角阵相似。

定理 6. 矩阵可对角化 $\iff$ n阶矩阵有n个线性无关的特征向量

证明. (⇒) 设

$$P^{-1}AP = \Lambda \implies AP = P\Lambda$$

将P按列分块，即

$$P = (x_1,\ x_2,\ \cdots,\ x_n),$$

则

$$A(x_1,\ x_2,\ \cdots,\ x_n) = (x_1,\ x_2,\ \cdots,\ x_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

于是

$$Ax_i = \lambda_i x_i \quad (i = 1, 2, \cdots, n).$$

故x_1，x_2，$\cdots$，x_n是A分别对应于λ_1，λ_2，$\cdots$，λ_n的特征向量。由于P可逆，所以它们是线性无关的。

(⇐) 上述步骤显然可逆，故充分性也成立。

□

若A与Λ相似，则Λ的主对角元都是A的特征值。 若不计λ_k的排列次序，则Λ是唯一的，称Λ为A的相似标准型。

定理 7. A的属于不同特征值的特征向量是线性无关的。

证明. 设A的m个互不相同的特征值为$\lambda_1, \lambda_2, \cdots, \lambda_m$，其相应的特征向量为$x_1$，$x_2$，$\cdots$，$x_m$. 对$m$做数学归纳法。

1° 当$m = 1$时，结论显然成立。

2° 设k个不同特征值$\lambda_1, \lambda_2, \cdots, \lambda_k$的特征向量$x_1$，$x_2$，$\cdots$，$x_k$。下面考虑$k+1$个不同特征值的特征向量。设

$$\begin{aligned}
& a_1x_1 + a_2x_2 + \cdots + a_kx_k + a_{k+1}x_{k+1} = 0 \qquad (1) \\
\implies\ & A(a_1x_1 + a_2x_2 + \cdots + a_kx_k + a_{k+1}x_{k+1}) = 0 \\
\implies\ & a_1\lambda_1x_1 + a_2\lambda_2x_2 + \cdots + a_k\lambda_kx_k + a_{k+1}\lambda_{k+1}x_{k+1} = 0 \quad (2) \\
\xrightarrow{(2)-\lambda_{k+1}(1)}\ & a_1(\lambda_{k+1} - \lambda_1)x_1 + a_2(\lambda_{k+1} - \lambda_2)x_2 + \cdots + a_k(\lambda_{k+1} - \lambda_k)x_k = 0 \\
\implies\ & a_i(\lambda_{k+1} - \lambda_i) = 0, \quad i = 1, 2, \cdots, k \\
\implies\ & a_i = 0, \quad i = 1, 2, \cdots, k \\
\implies\ & a_{k+1}x_{k+1} = 0 \\
\implies\ & a_{k+1} = 0 \\
\implies\ & x_1,\ x_2,\ \cdots,\ x_k,\ \ x_{k+1}\text{线性无关}
\end{aligned}$$

□

推论 2. 若A有n个互不相同的特征值，则A与对角阵相似。

例 6. 设实对称矩阵

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

问A是否可对角化？若可对角化，求对角阵Λ及可逆矩阵P使得$P^{-1}AP = \Lambda$，再求A^k。

解. 由

$$
\begin{aligned}
|\mathbf{A}-\lambda\mathbf{I}| &= \begin{vmatrix} 1-\lambda & -1 & -1 & -1 \\ -1 & 1-\lambda & -1 & -1 \\ -1 & -1 & 1-\lambda & -1 \\ -1 & -1 & -1 & 1-\lambda \end{vmatrix} \\
&= -(\lambda+2)\begin{vmatrix} 1 & -1 & -1 & -1 \\ 1 & 1-\lambda & -1 & -1 \\ 1 & -1 & 1-\lambda & -1 \\ 1 & -1 & -1 & 1-\lambda \end{vmatrix} \\
&= -(\lambda+2)\begin{vmatrix} 1 & -1 & -1 & -1 \\ 0 & 2-\lambda & 0 & 0 \\ 0 & 0 & 2-\lambda & 0 \\ 0 & 0 & 0 & 2-\lambda \end{vmatrix} = (\lambda+2)(\lambda-2)^3,
\end{aligned}
$$

故$\mathbf{A}$的特征值为$\lambda_1=-2$（单根），$\lambda_2=2$（三重根）。

由$(\mathbf{A}-\lambda_1\mathbf{I})\mathbf{x}=0$，即

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

得λ_1对应的特征向量为$\{k_1\mathbf{x}_1|\mathbf{x}_1=(1,1,1,1)^T, k_1\neq 0\}$。

由$(\mathbf{A}-\lambda_2\mathbf{I})\mathbf{x}=0$，即

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

得基础解系：

$$
\mathbf{x}_{21}=(1,-1,0,0)^T, \quad \mathbf{x}_{22}=(1,0,-1,0)^T, \quad \mathbf{x}_{23}=(1,0,0,-1)^T.
$$

因$\mathbf{A}$有4个线性无关的特征向量，故$\mathbf{A}\sim\mathbf{\Lambda}$。

取

$$
\mathbf{P}=(\mathbf{x}_1,\mathbf{x}_{21},\mathbf{x}_{22},\mathbf{x}_{23})=\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix},
$$

则

$$
\mathbf{P}^{-1}\mathbf{A}\mathbf{P}=\begin{pmatrix} -2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix}=\mathbf{\Lambda}.
$$

再由$\mathbf{A}=\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$得

$$
\begin{aligned}
\mathbf{A}^k &= (\mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1})^k = \mathbf{P}\mathbf{\Lambda}^k\mathbf{P}^{-1} \\
&= \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} (-2)^k & & & \\ & 2^k & & \\ & & 2^k & \\ & & & 2^k \end{pmatrix}\frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{pmatrix} \\
&= \begin{cases} 2^k\mathbf{I}_4, & k \textit{ even}, \\ 2^{k-1}\mathbf{A}, & k \textit{ odd}. \end{cases}
\end{aligned}
$$

例 7. 设

$$
\mathbf{A}=\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & x \\ 1 & 0 & 0 \end{pmatrix}
$$

问x为何值时，矩阵$\mathbf{A}$能对角化?

解. 由

$$|\mathbf{A}-\lambda\mathbf{I}| = \begin{vmatrix} -\lambda & 0 & 1 \\ 1 & 1-\lambda & x \\ 1 & 0 & -\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = -(\lambda-1)^2(\lambda+1),$$

即$\lambda_1=-1, \lambda_2=\lambda_3=1$。

对应于单根$\lambda_1=-1$，可求得线性无关的特征向量恰有1个，故$\mathbf{A}$可对角化的充分必要条件是对应重根$\lambda_2=\lambda_3=1$，有2个线性无关的特征向量，即$(\mathbf{A}-\mathbf{I})\mathbf{x}=0$有两个线性无关的解，亦即$\mathbf{A}-\mathbf{I}$的秩$\mathrm{R}(\mathbf{A}-\mathbf{I})=1$。由

$$\mathbf{A}-\mathbf{I} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 0 & x \\ 1 & 0 & -1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & x+1 \\ 0 & 0 & 0 \end{pmatrix}$$

欲使$\mathrm{R}(\mathbf{A}-\mathbf{I})=1$，须有$x+1=0$，即$x=-1$。因此当$x=-1$时，矩阵$\mathbf{A}$能对角化。

例 8. 设$\mathbf{A}=(a_{ij})_{n\times n}$是主对角元全为2的上三角矩阵，且存在$a_{ij}\neq 0(i<j)$，问$\mathbf{A}$是否可对角化?

解. 设

$$A = \begin{pmatrix} 2 & * & \cdots & * \\ & 2 & \cdots & * \\ & & \ddots & \vdots \\ & & & 2 \end{pmatrix}$$

其中$*$为不全为零的任意常数，则

$$|\mathbf{A}-\lambda\mathbf{I}| = (2-\lambda)^n,$$

即$\lambda=2$为$\mathbf{A}$的n重特征根，而$\mathrm{R}(\mathbf{A}-2\mathbf{I})\geqslant 1$，故$(\mathbf{A}-2\mathbf{I})\mathbf{x}=0$的基础解系所含向量个数$\leqslant n-1$个，即$\mathbf{A}$的线性无关的特征向量的个数$\leqslant n-1$个，因此$\mathbf{A}$不与对角阵相似。