# 4.3 矩阵可对角化的条件

研究特征值特征向量的其中一个原因是

我们希望将一个矩阵相似变换成一个对角矩阵.

但是并不是所有的矩阵都能相似于一个对角矩阵的.

$A_n$可对角化 $\Leftrightarrow$ $A_n$有$n$个无关特征向量

**定义4.3.1(可对角化)** 若方阵$A$相似于一个对角矩阵，则称$A$可对角化.

**例4.3.1** 说明矩阵 $A=\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$ 不可对角化.

**解** 由 $|\lambda E-A|=\begin{vmatrix}\lambda-1 & -2\\ 0 & \lambda-1\end{vmatrix}=(\lambda-1)^2=0$

可得特征值 =1(二重)，再假设矩阵$A$可对角化，即存在可逆矩阵$P$，有

$$P^{-1}AP=\begin{pmatrix}s_1 & 0\\ 0 & s_2\end{pmatrix}$$

因为相似矩阵有相同的特征值，故$s_1=s_2=1$，于是$P^{-1}AP=E$，从而得$A=PEP^{-1}=E$，与假设矛盾. 故$A$不可对角化.

# 可对角化的一般性条件

**定理4.3.1** $n$阶矩阵可对角化的充要条件是有$n$个线性无关的特征向量；且对角矩阵的主对角线由特征值(可按任意次序)构成，相似变换矩阵由属于相应特征值的特征向量构成．

说明：

可对角化：$(\xi_1, \dots ,\xi_n)^{-1}A(\xi_1, \dots ,\xi_n)=\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}$

$\Updownarrow$

$$\begin{cases} (A\xi_1, \dots, A\xi_n)=(\xi_1, \dots ,\xi_n)\begin{pmatrix} s_1 & & \\ & \ddots & \\ & & s_n \end{pmatrix}=(s_1\xi_1, \dots ,s_n\xi_n) \\ \text{矩阵}(\xi_1, \dots , \xi_n)\text{可逆} \end{cases}$$

$\Updownarrow$

$A\xi_i=s_i\xi_i$ , $i=1,2,\dots,n$ ； $\xi_1, \dots , \xi_n$ 线性无关

$$(\xi_{i1}, \dots ,\xi_{in})^{-1}A(\xi_{i1}, \dots ,\xi_{in})=(\xi_{i1}, \dots ,\xi_{in})^{-1} (A\xi_{i1}, \dots , A\xi_{in})$$

$$=(\xi_{i1}, \dots ,\xi_{in})^{-1} (s_{i1}\xi_{i1}, \dots , s_{in}\xi_{in})=(\xi_{i1}, \dots ,\xi_{in})^{-1} (\xi_{i1}, \dots , \xi_{in})\begin{pmatrix} s_{i_1} & & \\ & \ddots & \\ & & s_{i_n} \end{pmatrix}=\begin{pmatrix} s_{i_1} & & \\ & \ddots & \\ & & s_{i_n} \end{pmatrix}$$

## 特征向量的关系

**定理4.3.2 属于不同特征值的特征向量线性无关.**

说明: $y=k_1\xi_1+k_2\xi_2+\ldots+k_m\xi_m=\theta$,

$$\lambda_m y = k_1\lambda_m\xi_1 + \ldots + k_{m-1}\lambda_m\xi_{m-1} + k_m\lambda_m\xi_m = \theta$$

$$-)\ Ay = k_1\lambda_1\xi_1 + \ldots + k_{m-1}\lambda_{m-1}\xi_{m-1} + k_m\lambda_m\xi_m = \theta$$

$$k_1(\lambda_m-\lambda_1)\xi_1 + \ldots + k_{m-1}(\lambda_m-\lambda_{m-1})\xi_{m-1} + \theta = \theta$$

利用归纳假设

结论: $k_1(\lambda_m-\lambda_1)=k_2(\lambda_m-\lambda_2)=\ldots=k_{m-1}(\lambda_m-\lambda_{m-1})=0$

$$k_1=\ldots=k_{m-1}=0 \Rightarrow k_m=0$$

**推论4.3.3 若$n$阶矩阵有$n$个互不相同的特征值，则矩阵可对角化.**

**例4.3.2 将矩阵** $A=\begin{pmatrix}1&1&1\\-2&1&2\\2&1&0\end{pmatrix}$ **对角化.**

**解 在例4.2.5中求得*A*的特征值为：** $\lambda=1,-1,2$ **.**

**对于**$\lambda=1$**，**

**解齐次方程组 $(E-A)x=\theta$，由** $\begin{pmatrix}0&-1&-1\\2&0&-2\\-2&-1&1\end{pmatrix}\to\begin{pmatrix}0&-1&-1\\1&0&-1\\0&-1&-1\end{pmatrix}\to\begin{pmatrix}1&0&-1\\0&1&1\\0&0&0\end{pmatrix}$,

**求得该齐次方程组的一个基础解系为：**$\alpha_1=(1,-1,1)^T$ **.**

**对于**$\lambda=-1$**，**

**解齐次方程组 $(-E-A)x=\theta$，由** $\begin{pmatrix}-2&-1&-1\\2&-2&-2\\-2&-1&-1\end{pmatrix}\to\begin{pmatrix}0&-3&-3\\1&-1&-1\\0&-3&-3\end{pmatrix}\to\begin{pmatrix}1&0&0\\0&1&1\\0&0&0\end{pmatrix}$,

**求得该齐次方程组的一个基础解系为：**$\alpha_2=(0,-1,1)^T$ **.**

**对于**$\lambda=2$**，**

**解齐次方程组 $(2E-A)x=\theta$，由** $\begin{pmatrix}1&-1&-1\\2&1&-2\\-2&-1&2\end{pmatrix}\to\begin{pmatrix}1&-1&-1\\0&3&0\\0&-3&0\end{pmatrix}\to\begin{pmatrix}1&0&-1\\0&1&0\\0&0&0\end{pmatrix}$,

**求得该齐次方程组的一个基础解系为：**$\alpha_3=(1,0,1)^T$ **.**

**令** $P=\begin{pmatrix}1&0&1\\-1&-1&0\\1&1&1\end{pmatrix}$, **则有** $P^{-1}AP=\mathrm{diag}(1,-1,2)$ **.**

**补充例4G** 已知矩阵 $A=\begin{pmatrix}1&1&1\\-2&1&2\\2&1&0\end{pmatrix}$，计算 $A^n$.

**解** $$|\lambda E-A|=\begin{vmatrix}\lambda-1&-1&-1\\2&\lambda-1&-2\\-2&-1&\lambda\end{vmatrix}=\begin{vmatrix}\lambda-2&-1&-1\\0&\lambda-1&-2\\\lambda-2&-1&\lambda\end{vmatrix}=(\lambda-1)(\lambda+1)(\lambda-2)$$

故$A$的特征值为：$\lambda=1,-1,2$.

当$\lambda=1$时，解得一个无关特征向量为：$\alpha_1=(1,-1,1)^T$.

当$\lambda=-1$时，解得一个无关特征向量为：$\alpha_2=(0,-1,1)^T$.

当$\lambda=2$时，解得一个无关特征向量为：$\alpha_3=(1,0,1)^T$.

令 $P=(\alpha_1,\alpha_2,\alpha_3)$，则有 $P^{-1}AP=\mathrm{diag}(1,-1,2)=D$，即 $A=PDP^{-1}$，故

$$A^n=(PDP^{-1})^n=PD^nP^{-1}=\begin{pmatrix}1&0&1\\-1&-1&0\\1&1&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&(-1)^n&0\\0&0&2^n\end{pmatrix}\begin{pmatrix}1&-1&-1\\-1&0&1\\0&1&1\end{pmatrix}$$

$$=\begin{pmatrix}1&-1+2^n&-1+2^n\\-1+(-1)^n&1&1-(-1)^n\\1-(-1)^n&-1+2^n&-1+(-1)^n+2^n\end{pmatrix}.$$

**补充例4H** 已知$P^{-1}AP=D$，其中 $P=\begin{pmatrix}1&0&-2\\-1&2&0\\1&1&1\end{pmatrix}, D=\begin{pmatrix}3&0&0\\0&1&0\\0&0&0\end{pmatrix}$.

计算矩阵 $B=A^2+(2E-A)^{-1}$的特征值和特征向量.

解 令 $P=(\xi_1,\xi_2,\xi_3)=\begin{pmatrix}1&0&-2\\-1&2&0\\1&1&1\end{pmatrix}$.

易知 $A=PDP^{-1}$, 则 $B=A^2+(2E-A)^{-1}=PD^2P^{-1}+(P(2E-D)P^{-1})^{-1}=P(D^2+(2E-D)^{-1})P^{-1}$.

而 $D^2+(2E-D)^{-1}$ =diag(9,1,0)+(diag(-1,1,2))$^{-1}$=diag(8,2,0.5)，

故 $B(\xi_1,\xi_2,\xi_3)=BP=P$diag(8,2,0.5)=$(8\xi_1,2\xi_2,0.5\xi_3)$.

故 $B=A^2+(2E-A)^{-1}$ 特征值为8,2,0.5,对应特征向量 $k_1\xi_1$, $k_2\xi_2$, $k_2\xi_2$， $k_1,k_2,k_3\in$R.

解法二 令$\lambda_1=3$, $\lambda_2=1$, $\lambda_3=0$，$\xi_1=\begin{pmatrix}1\\-1\\1\end{pmatrix}, \xi_2=\begin{pmatrix}0\\2\\1\end{pmatrix}, \xi_3=\begin{pmatrix}-2\\0\\1\end{pmatrix}$.

则有 $D=\begin{pmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{pmatrix}, P=(\xi_1,\xi_2,\xi_3)$. 于是有 $AP=PD$， 即 $A\xi_i=\lambda_i\xi_i$ , $i$=1,2,3，

故 $(A^2+(2E-A)^{-1})\xi_i=(\lambda_i^2+(2-\lambda_i)^{-1})\xi_i$ , $i$=1,2,3.

可得 $B=A^2+(2E-A)^{-1}$ 的特征值$\lambda_i^2+(2-\lambda_i)^{-1}$ 为8,2,0.5，对应特征向量为

$k_1\xi_1$, $k_2\xi_2$, $k_2\xi_2$， $k_1,k_2,k_3$为非零实数.

**补充例4I** 设$\xi_1$和$\xi_2$为矩阵$A$的属于不同特征值的特征向量，证明$a\xi_1+b\xi_2$当$ab\neq 0$时不是$A$的特征向量.

解 设$A$的特征向量$\xi_1$和$\xi_2$对应特征值为$\lambda_1$、$\lambda_2$，且有$\lambda_1\neq\lambda_2$ .

假设$\xi_3=a\xi_1+b\xi_2$也是$A$的特征向量，对应特征值为$\lambda_3$，

则有 $A\xi_3=\lambda_3\xi_3$，

又有 $A\xi_3=A(a\xi_1+b\xi_2)=aA\xi_1+bA\xi_2=a\lambda_1\xi_1+b\lambda_2\xi_2$，

故有 $a\lambda_1\xi_1+b\lambda_2\xi_2=\lambda_3(a\xi_1+b\xi_2)$，即 $a(\lambda_1-\lambda_3)\xi_1+b(\lambda_2-\lambda_3)\xi_2=\theta$ .

由于$\xi_1$和$\xi_2$对应于不同的特征值，故$\xi_1$和$\xi_2$线性无关，

于是由 $a(\lambda_1-\lambda_3)\xi_1+b(\lambda_2-\lambda_3)\xi_2=\theta$ 可得$a(\lambda_1-\lambda_3)=b(\lambda_2-\lambda_3)=0$，

再由$ab\neq 0$，得$\lambda_1=\lambda_3=\lambda_2$，与$\lambda_1\neq\lambda_2$矛盾，故$a\xi_1+b\xi_2$不是特征向量.

**补充例4J** 设$A\in\mathbf{R}^{n\times n}$，$A$可对角化，且 $f(\lambda)=|\lambda E-A|$，证明 $f(A)=O$.

证明 因为$A$可对角化，故存在可逆矩阵$P$使得 $P^{-1}AP=\mathrm{diag}(\lambda_1,\ldots\lambda_n)=D$，

易知$\lambda_1,\ldots,\lambda_n$是$A$的$n$个特征值. 由 $f(\lambda)=|\lambda E-A|$得 $f(\lambda_i)=0,i=1,2,\ldots,n$.

由 $P^{-1}AP=D$可得 $P^{-1}f(A)P=f(P^{-1}AP)=f(D)=\mathrm{diag}(f(\lambda_1),\ldots,f(\lambda_n))=O$，

于是 $f(A)=POP^{-1}=O$.

注 由哈密顿-凯莱定理，任意方阵$A$都满足$f(A)=O$，其中$f(\lambda)=|\lambda E-A|$.

## 重特征值条件下的对角化

**定理4.3.4** 若$\lambda_1, \lambda_2, \dots, \lambda_m$是矩阵$A$的不同特征值，而$A$的属于$\lambda_i$的线性无关的特征向量为$\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{is_i}$ $(i=1,2,\dots,m)$，则向量组

$$\alpha_{11}, \dots, \alpha_{1s_1}, \alpha_{21}, \dots, \alpha_{2s_2}, \dots, \alpha_{m1}, \dots, \alpha_{ms_m},$$

线性无关 .

说明：

$$\underbrace{k_{11}\alpha_{11}+\dots+k_{1s_1}\alpha_{1s_1}}_{\beta_1\ \ \lambda_1}+\underbrace{k_{21}\alpha_{21}+\dots+k_{2s_2}\alpha_{2s_2}}_{\beta_2\ \ \lambda_2}+\dots+\underbrace{k_{m1}\alpha_{m1}+\dots+k_{ms_m}\alpha_{ms_m}}_{\beta_m\ \ \lambda_m}=\theta$$

$\beta_1$为$\theta$或$\lambda_1$的特征向量

$A\beta_1=\lambda_1\beta_1$　　$A\beta_2=\lambda_2\beta_2$　　$A\beta_m=\lambda_m\beta_m$

$$\beta_1+\beta_2+\dots+\beta_s+\beta_{s+1}+\dots+\beta_k+\dots+\beta_m=\theta$$ (红色为0向量)

$\beta_2,\beta_s,\dots,\beta_k$无关, 矛盾

$$k_{i1}\alpha_{i1}+\dots+k_{is_i}\alpha_{is_i}=\beta_i=\theta \quad \Rightarrow k_{i1}=k_{i2}=\dots=k_{is_i}=0$$

$$k_{11}=k_{12}=\dots=k_{1s_1}=k_{21}=\dots=k_{m1}=\dots=k_{ms_m}=0$$

**定理4.3.5** 设$\lambda_0$是$n$阶方阵$A$的$k$重特征值，则$A$的属于特征值$\lambda_0$的线性无关的特征向量个数不超过 $k$ .

说明：

$$\underbrace{\xi_1, \xi_2, \cdots, \xi_s}_{\lambda_0}, \eta_{s+1}, \cdots, \eta_n \quad \text{线性无关}$$

- 将向量 $e_1, e_2, \cdots, e_n$依次添加到向量组$\xi_1, \xi_2, \cdots, \xi_s$中
- 当添加的向量使得添加后的向量组线性相关时，删去该添加的向量
- 最后得扩充的线性无关组$\xi_1, \xi_2, \cdots, \xi_s, \eta_{s+1}, \cdots, \eta_n$ .

$$A(\xi_1, \cdots, \xi_s, \eta_{s+1}, \cdots, \eta_n)=(\lambda_0\xi_1, \cdots, \lambda_0\xi_s, \varphi_{s+1}, \cdots, \varphi_n)$$

$$=\underbrace{(\xi_1, \cdots, \xi_s, \eta_{s+1}, \cdots, \eta_n)}_{P}\underbrace{\left(\begin{array}{ccc|cc} \lambda_0 & & & * & * \\ & \ddots & & * & * \\ & & \lambda_0 & * & * \\ \hline & & & * & * \\ & & & * & * \end{array}\right)}_{B}$$

结论：$P^{-1}AP=B$

$$|\lambda E-A|=|\lambda E-B|=(\lambda-\lambda_0)^s g(\lambda)=(\lambda-\lambda_0)^k f(\lambda)\text{，其中 } f(\lambda_0)\neq 0$$

定理4.3.6 $n$阶方阵$A$可对角化的充要条件是每个$k_i$重特征值$\lambda_i$对应的特征矩阵$\lambda_i E\text{-}A$的秩为$n\text{-}k_i$ .

无关特征向量

$\lambda_1$ ($k_1$重) $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1s1}$ $s_1 \leqslant k_1$

$\lambda_2$ ($k_2$重) $\alpha_{21}, \alpha_{22}, \dots, \alpha_{2s2}$ $s_2 \leqslant k_2$

…… …… ……

$\lambda_m$ ($k_m$重) $\alpha_{m1}, \alpha_{m2}, \dots, \alpha_{msm}$ $s_m \leqslant k_m$ ( +

$\sum s_i \leqslant \sum k_i = n$

$$n=\sum s_i \leqslant \sum k_i = n \Leftrightarrow s_i = k_i \Leftrightarrow \mathrm{r}(\lambda_i E\text{-}A) = n\text{-}k_i$$

$n$个无关特征向量

# 矩阵对角化的具体步骤

将矩阵$A$对角化的具体步骤：

(1) 解特征方程 $|\lambda E-A|=0$ 得特征值 $\lambda=\lambda_1(s_1$重$), \dots, \lambda_m(s_m$重$)$；

(2) 对每个特征值 $\lambda_i$, $i=1,2,\dots,m$，解齐次方程组

$$(\lambda_i E-A)x=\theta,$$

得一个基础解系 $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{ir_i}$ .

若有某个 $i$ 使得 $r_i<s_i$，则矩阵 $A$ 不可对角化；

(3) 当所有的 $r_i=s_i$, $i=1,2,\dots,m$，则令

$$P=(\alpha_{11}, \dots, \alpha_{1s_1}, \alpha_{21}, \dots, \alpha_{2s_2}, \dots, \alpha_{m1}, \dots, \alpha_{ms_m}),$$

即得 $P^{-1}AP=\mathrm{diag}(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_m, \dots, \lambda_m)$ .

例4.3.3 问 $A=\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix}$ 是否可对角化，为什么？

解 由

$$|\lambda E-A|=\begin{vmatrix} \lambda-2 & 1 & -2 \\ -5 & \lambda+3 & -3 \\ 1 & 0 & \lambda+2 \end{vmatrix}=(\lambda+1)^3=0$$

得三重特征值 $\lambda=-1$，显然 $-E-A\neq O$，因而其秩 $r\geq 1$，但 $r\neq n-3=0$，故$A$不可能与对角矩阵相似.

**例4.3.4** 将上节例4.2.1的矩阵 $A=\begin{pmatrix}5 & -2 & 1\\0 & 4 & 0\\1 & -2 & 5\end{pmatrix}$ 对角化.

**解** 在上节例4.2.1已求得$A$的特征值为：$\lambda=6,4$(二重)；对$\lambda=6$，已求得一个特征向量为：$\alpha_1=(1,0,1)^T$；对 $\lambda=4$，已求得二个线性无关的特征向量为：$\alpha_2=(2,1,0)^T$ , $\alpha_3=(-1,0,1)^T$ .

令 $P=(\alpha_1,\alpha_2,\alpha_3)=\begin{pmatrix}1 & 2 & -1\\0 & 1 & 0\\1 & 0 & 1\end{pmatrix}$，则有 $P^{-1}AP=\mathrm{diag}(6,4,4)$ .

**例4.3.5** 证明：若方阵$A$满足关系 $A^2=E$，则$A$可对角化 .

**证明** 设$A\in \mathbf{R}^{n\times n}$，由 $A^2=E$ 可得

$$(E-A)(-E-A)=(-E-A)(E-A)=O\ .$$

由于 $-E-A$ 的非零列是 $(E-A)x=\theta$ 的非零解，所以$A$至少有 $r(-E-A)$ 个属于特征值 $\lambda=1$的线性无关特征向量，同理$A$至少有 $r(E-A)$个属于特征值 $\lambda=-1$的线性无关特征向量, 且$r(-E-A)+r(E-A)\le n$. 又由
$r(E-A)+r(-E-A)=r(E-A)+r(E+A)\ge r((E-A)+(E+A))=r(2E)=n$
可得$A$有$n$个线性无关的特征向量，可以对角化.

**补充例4K** 已知 $A=\begin{pmatrix}3&1&-3\\4&3&-6\\4&2&-5\end{pmatrix}, B=\begin{pmatrix}1&0&0\\0&0&-1\\0&-1&0\end{pmatrix}$，$A\sim B$，求$P$使得$P^{-1}BP=A$.

解 由 $|\lambda E-A|=\begin{vmatrix}\lambda-3&-1&3\\-4&\lambda-3&6\\-4&-2&\lambda+5\end{vmatrix}=(\lambda-1)^2(\lambda+1),$

得$A$的特征值 $\lambda=1$(二重), -1，对应特征向量 $\begin{pmatrix}-1\\2\\0\end{pmatrix},\begin{pmatrix}3\\0\\2\end{pmatrix},\begin{pmatrix}1\\2\\2\end{pmatrix}.$

$B$有相同特征值$\lambda=1$(二重), -1，对应特征向量 $\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\-1\\1\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix}.$

令 $Q_1=\begin{pmatrix}-1&3&1\\2&0&2\\0&2&2\end{pmatrix}, Q_2=\begin{pmatrix}1&0&0\\0&-1&1\\0&1&1\end{pmatrix}$. 则 $Q_1^{-1}AQ_1=Q_2^{-1}BQ_2=\mathrm{diag}(1,1,-1)$.

故有 $A=Q_1Q_2^{-1}BQ_2Q_1^{-1}=P^{-1}BP$，则 $P=Q_2Q_1^{-1}=\begin{pmatrix}1&1&-3/2\\-2&-1&5/2\\0&0&1/2\end{pmatrix}.$

**补充例4L** 设矩阵 $A=\begin{pmatrix}1 & -1 & 1\\ x & 4 & y\\ -3 & -3 & 5\end{pmatrix}$，已知$A$有3个线性无关的特征向量，$\lambda=2$是$A$的二重特征值，试求矩阵$A$.

解 $A$有3个线性无关的特征向量，$\lambda=2$是$A$的二重特征值，故$A$的属于$\lambda=2$的无关特征向量有两个，于是有 $r(2E-A)=1$. 对$2E-A$做初等行变换如下

$$2E-A=\begin{pmatrix}1 & 1 & -1\\ -x & -2 & -y\\ 3 & 3 & -3\end{pmatrix}\to\begin{pmatrix}1 & 1 & -1\\ 0 & x-2 & -x-y\\ 0 & 0 & 0\end{pmatrix},$$

故有 $x=2,y=-2$，于是 $A=\begin{pmatrix}1 & -1 & 1\\ 2 & 4 & -2\\ 3 & 3 & 5\end{pmatrix}$.