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：必要性：A可对角化，有可逆矩阵 $P=(\xi_1, \xi_2, \dots , \xi_n)$使得

$$P^{-1}AP=\mathrm{diag}(s_1,s_2, \dots ,s_n)=\Lambda，\text{即}AP=P\Lambda，$$

则 $(A\xi_1, \dots,A\xi_n)=(s_1\xi_1, \dots ,s_n\xi_n)$, 即$A\xi_i=s_i \xi_i$,且$\xi_1, \dots ,\xi_n$无关($P$可逆).

充分性：设$\xi_1, \dots , \xi_n$为n个无关特征向量, 对应特征值$s_1,\dots ,s_n$. 则$P=(\xi_1, \dots , \xi_n)$可逆. 且$AP=(A\xi_1, \dots,A\xi_n)=(s_1\xi_1,\dots, s_n\xi_n)=P\Lambda$, 其中 $\Lambda=\mathrm{diag}(s_1, \dots,s_n)$，故有$P^{-1}AP=\mathrm{diag}(s_1,\dots,s_n)$，矩阵可对角化.

进一步令 $C=(e_{i1},e_{i2}, \dots ,e_{in})$, $D=PC$, 其中 $(i_1,i_2,\dots,i_n)$为$(1,2,\dots,n)$的一个排列, 则$C^TC=(d_{kj})_{n\times n}=(e_{ik}{}^Te_{ij})_{n\times n}=(\delta_{kj})_{n\times n}=E$, 故$C^T=C^{-1}$, 于是

$$D^{-1}AD=C^{-1}P^{-1}APC=C^T\mathrm{diag}(s_1,s_2, \dots ,s_n)C$$

$$= C^T(s_{i1}e_{i1}, s_{i2}e_{i2}, \dots , s_{in}e_{in})=\mathrm{diag}(s_{i1},s_{i2}, \dots ,s_{in}) ,$$

即对角线的特征值可按任意次序排列.

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 .

证明：设A有互不相同的特征值 $\lambda_1, \lambda_2, \ldots, \lambda_r$及对应的特征向量$\xi_1, \xi_2, \ldots, \xi_r$.
现在用数学归纳法证明线性无关.

当 $r=2$ 时，设有 $k_1\xi_1+k_2\xi_2=\theta$，左乘A得 $k_1\lambda_1\xi_1+k_2\lambda_2\xi_2=\theta$，即

$$\begin{cases} k_1\xi_1+k_2\xi_2=\theta, \\ k_1\lambda_1\xi_1+k_2\lambda_2\xi_2=\theta. \end{cases}$$

上述第一式乘以λ_2再减去第二式得到：$k_1(\lambda_2-\lambda_1)\xi_1=\theta$，因为$\xi_1\neq\theta$，$\lambda_1\neq\lambda_2$，故得到$k_1=0$，从而也有$k_2=0$，故$\xi_1, \xi_2$线性无关 .

假设 $r=m-1$ 时, 结论成立，当$r=m$时, 设$y=k_1\xi_1+k_2\xi_2+\ldots+k_m\xi_m=\theta$，则有

$$\begin{cases} y=k_1\xi_1+k_2\xi_2+\cdots+k_m\xi_m=\theta, \\ Ay=k_1\lambda_1\xi_1+k_2\lambda_2\xi_2+\cdots+k_m\lambda_m\xi_m=\theta. \end{cases}$$

第一式乘以k_m再减去第二式得到：

$$k_1(\lambda_m-\lambda_1)\xi_1+k_2(\lambda_m-\lambda_2)\xi_2+\ldots+k_{m-1}(\lambda_m-\lambda_{m-1})\xi_{m-1}=\theta .$$

由假设知 $k_1(\lambda_m-\lambda_1)=k_2(\lambda_m-\lambda_2)=\ldots=k_{m-1}(\lambda_m-\lambda_{m-1})=0$，因为$\lambda_1, \ldots, \lambda_m$互不相同，故有 $k_1=\ldots=k_{m-1}=0$，从而$k_m=0$，即线性无关. 故$r=m$时结论成立 .

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 $\begin{pmatrix}0&-1&-1\\2&0&-2\\-2&-1&1\end{pmatrix}\xrightarrow[r_2\div 2]{r_3+r_2}\begin{pmatrix}0&-1&-1\\1&0&-1\\0&-1&-1\end{pmatrix}\xrightarrow[r_1\leftrightarrow r_2]{\substack{r_3-r_1\\(-1)r_1}}\begin{pmatrix}1&0&-1\\0&1&1\\0&0&0\end{pmatrix}$,

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 $\begin{pmatrix}-2&-1&-1\\2&-2&-2\\-2&-1&-2\end{pmatrix}\xrightarrow[r_2\div 2]{\substack{r_1+r_2\\r_3+r_2}}\begin{pmatrix}0&-3&-3\\1&-1&-1\\0&-3&-3\end{pmatrix}\xrightarrow[\substack{(-\frac{1}{3})r_1\\r_1\leftrightarrow r_2}]{\substack{r_3-r_1\\r_2-\frac{1}{3}r_1}}\begin{pmatrix}1&0&0\\0&1&1\\0&0&0\end{pmatrix}$,

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 $\begin{pmatrix}1&-1&-1\\2&1&-2\\-2&-1&2\end{pmatrix}\xrightarrow[r_3+2r_1]{r_2-2r_1}\begin{pmatrix}1&-1&-1\\0&3&0\\0&-3&0\end{pmatrix}\xrightarrow[r_2\div 3]{\substack{r_3+r_2\\r_1+\frac{1}{3}r_2}}\begin{pmatrix}1&0&-1\\0&1&0\\0&0&0\end{pmatrix}$,

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 $P^{-1}AP=\mathrm{diag}(1,-1,2)$.

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$\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{isi}$ ($i=1,2,\dots,m$)，则向量组

$$\alpha_{11}, \dots, \alpha_{1s1}, \alpha_{21}, \dots, \alpha_{2s2}, \dots, \alpha_{m1}, \dots, \alpha_{msm},$$

线性无关 .

证明 考虑如下式子

$$k_{11}\alpha_{11}+\dots+k_{1s1}\alpha_{1s1}+k_{21}\alpha_{21}+\dots+k_{2s2}\alpha_{2s2}+\dots+k_{m1}\alpha_{m1}+\dots+k_{msm}\alpha_{msm}=\theta .$$

设

$$\begin{cases} \beta_1 = k_{11}\alpha_{11} + k_{12}\alpha_{12} + \cdots + k_{1s_1}\alpha_{1s_1}, \\ \beta_2 = k_{21}\alpha_{21} + k_{22}\alpha_{22} + \cdots + k_{2s_2}\alpha_{2s_2}, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \beta_m = k_{m1}\alpha_{m1} + k_{m2}\alpha_{m2} + \cdots + k_{ms_m}\alpha_{ms_m}, \end{cases}$$

则有 $\beta_1+\beta_2+\dots+\beta_m=\theta$,

且由定理4.2.1知 $A\beta_1=\lambda_1\beta_1$， $A\beta_2=\lambda_2\beta_2$， … ，$A\beta_m=\lambda_m\beta_m$.

若$\beta_{i1}, \beta_{i2}, \dots, \beta_{ir}$为非零向量，其余为零向量，则有 $\beta_{i1}+\beta_{i2}+\dots+\beta_{ir}=\theta$，即这些向量线性相关. 但由定理4.3.2可知$\beta_{i1}, \beta_{i2}, \dots, \beta_{ir}$线性无关，产生矛盾. 故必有$\beta_1=\beta_2=\dots=\beta_m=\theta$.

对每个 $i\in\{1,2,\dots,m\}$，由于$\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{isi}$是线性无关的特征向量，有$\beta_i=k_{i1}\alpha_{i1}+k_{i2}\alpha_{i2}+\dots+k_{isi}\alpha_{isi}=\theta$， 可得 $k_{i1}=k_{i2}=\dots=k_{isi}=0$, $i=1,2,\dots,m$.
这就证得向量组$\alpha_{11}, \dots, \alpha_{1s1}, \alpha_{21}, \dots, \alpha_{2s2}, \dots, \alpha_{m1}, \dots, \alpha_{msm}$线性无关 .

定理4.3.5 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个数不超过 k .

证明 设有A的属于λ_0的特征向量的极大无关组$\xi_1, \xi_2, \dots, \xi_s$. 将向量$e_1, e_2, \dots, e_n$依次添加到向量组$\xi_1, \xi_2, \dots, \xi_s$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. 最后可得扩充的线性无关组$\xi_1, \xi_2, \dots, \xi_n$. 令 $P=(\xi_1, \xi_2, \dots, \xi_n)$，则 $r(P)=n$，于是P可逆. 因为 $Pe_i=\xi_i$，所以有 $P^{-1}\xi_i=e_i$, $(i=1,2,\dots,n)$.

令$B=P^{-1}AP$，则

$$B = P^{-1}(\lambda_0\xi_1,\cdots,\lambda_0\xi_s,\alpha_{s+1},\cdots,\alpha_n) = (\lambda_0 e_1,\cdots,\lambda_0 e_s,\beta_{s+1},\cdots,\beta_n) = \begin{pmatrix} \lambda_0 E_s & * \\ O & * \end{pmatrix}.$$

于是据定理4.2.3， $|\lambda E-A|=|\lambda E-B|=(\lambda-\lambda_0)^s g(\lambda)$.

其中 $g(\lambda)$是λ的多项式. 上式说明特征值λ_0的重数 $k\geq s$. 由λ_0的任意性，可知线性无关的特征向量个数不超过对应特征值的重数.

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λ_i对应的特征矩阵$\lambda_i E-A$的秩为$n-k_i$.

证明 设$\lambda_1, \dots, \lambda_s$为$A$的全部不同特征值， $(\lambda_i E-A)x=\theta$ 的基础解系含 $n-(n-k_i)=k_i$个线性无关的特征向量，由定理4.3.4可知，A有$\Sigma k_i=n$个线性无关的特征向量，则A与对角矩阵相似.

反之，若A相似于对角矩阵，则A有n个线性无关的特征向量，而$k_i\geq 1$且$\Sigma k_i=n$，由定理4.3.5，A的属于λ_i的特征向量的极大无关组恰好有k_i个特征向量. 这个极大无关组又是 $(\lambda_i E-A)x=\theta$ 的基础解系，从而$\lambda_i E-A$的秩为$n-k_i$.

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$), \ldots, \lambda_m(s_m$重$)$；

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

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

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

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 $r_i=s_i$, $i=1,2,\ldots,m$， 则令

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

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

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 $r\geq1$，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$r(-E-A)+r(E-A)\leq n$. 又由 $r(E-A)+r(-E-A)=r(E-A)+r(E+A)\geq r((E-A)+(E+A))=r(2E)=n$ 可得A有n个线性无关的特征向量，可以对角化.