第五章 矩阵相似对角化

本章介绍矩阵特征值和特征向量概念和求法，并给出两个矩阵相似的概念，介绍矩阵与对角矩阵相似的条件和方法. 这些方法在工程技术中有着广泛的应用.

§1 矩阵的特征值和特征向量

一. 特征值和特征向量的概念

定义5.1 设**A**是n阶方阵，如果数λ_0和n维非零列向量ξ满足关系式

$$\mathbf{A}\xi=\lambda_0\xi,$$

则称λ_0为**A**的特征值，ξ为**A**的属于λ_0的一个特征向量.

注：特征向量一定是非零向量.

对于n阶零矩阵$\mathbf{O}$，由于对于任意n维非零向量ξ都有

$$\mathbf{O}\xi = \mathbf{0} = 0\xi.$$

所以，数0是零矩阵$\mathbf{O}$的特征值，任意n维非零向量都是n阶零矩阵$\mathbf{O}$的属于特征值0的特征向量.

对于n阶单位矩阵$\mathbf{E}$，由于对于任意n维非零向量ξ都有

$$\mathbf{E}\xi = \xi = 1\xi.$$

所以，数1是单位矩阵$\mathbf{E}$的特征值，任意n维非零向量都是n阶单位矩阵$\mathbf{E}$的属于特征值1的特征向量.

如果$\mathbf{A}$是奇异矩阵($|\mathbf{A}|=0$)，则齐次线性方程组$\mathbf{Ax=0}$有非零解. 若记ξ为$\mathbf{Ax=0}$的非零解，则有

$$\mathbf{A}\xi = \mathbf{0} = 0\xi.$$

可见，$\lambda_0=0$为奇异矩阵$\mathbf{A}$的(一个)特征值；方程组$\mathbf{Ax=0}$的所有非零解向量都是$\mathbf{A}$的属于特征值$\lambda_0=0$的特征向量.

对于一般的n阶方阵**A**，如何确定其全部的特征值和特征向量?

一般地，由$\mathbf{A}\xi=\lambda_0\xi$可得

$$(\lambda_0\mathbf{E}-\mathbf{A})\xi=\mathbf{0}.$$

可见:

ξ是n元齐次线性方程组$(\lambda_0\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$的非零解.

这等价于以下条件:

$$|\lambda_0\mathbf{E}-\mathbf{A}|=0.$$

思考: 这意味着……?

定义5.2 设**A**是n阶方阵，λ是参数，记

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

称$f(\lambda)$为方阵**A**的**特征多项式**. 称$f(\lambda)=0$, 也即

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

为方阵**A**的**特征方程**.

1. 可能会有重特征值;
2. 若有复特征值，会共轭成对.

注1: **A**的特征值就是其特征方程的解.

注2: $f(\lambda)$是n次多项式，因此**A**必有n个特征值.

注3: **A**的属于特征值λ_i的特征向量就是齐次线性方程组

$$(\lambda_i\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$$

的所有非零解.

例1 求矩阵

$$\mathbf{A}=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}$$

的特征值和特征向量.

解 **A**的特征多项式为

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

所以**A**的特征值为 $\lambda_1=\lambda_2=1,\ \lambda_3=3.$

对$\lambda_1=\lambda_2=1$，解方程组$(\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$，由于

$$\mathbf{E}-\mathbf{A}=\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ -1 & -3 & 0 \end{pmatrix} \overset{\text{行变换}}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

得同解方程: $\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$, 基础解系可取为$\xi_1 = (0, 0, 1)^T$.

所以 $k\xi_1\,(k\in R$ 且 $k\neq 0)$ 是属于$\lambda_1=\lambda_2=1$的全部特征向量.

对$\lambda_3=3$, 解方程$(3\mathbf{E}\text{-}\mathbf{A})\mathbf{x}=\mathbf{0}$, 由于

$$3\mathbf{E}-\mathbf{A}=\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ -1 & -3 & 2 \end{pmatrix} \overset{\text{行变换}}{\sim} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

得同解方程: $\begin{cases} x_1 = -x_3 \\ x_2 = x_3 \end{cases}$, 基础解系可取为$\xi_2 = (-1, 1, 1)^T$.

所以 $k\xi_2\,(k\in R$ 且 $k\neq 0)$ 是属于$\lambda_3=3$的全部特征向量.

例2 求矩阵

$$\mathbf{A}=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}$$

的特征值和特征向量.

解 **A**的特征多项式为

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

所以**A**的特征值为 $\lambda_1=\lambda_2=1$, $\lambda_3=3$.

对$\lambda_1=\lambda_2=1$, 解方程组$(\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$, 由于

$$\mathbf{E}-\mathbf{A}=\begin{pmatrix}-1 & 1 & 0\\ 1 & -1 & 0\\ -1 & 1 & 0\end{pmatrix}\overset{\text{行变换}}{\sim}\begin{pmatrix}1 & -1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

得同解方程：$x_1=x_2$，基础解系为$\xi_1=(1, 1, 0)^T$, $\xi_2=(0, 0, 1)^T$.

所以属于$\lambda_1=\lambda_2=1$的全部特征向量为

$k_1\xi_1+k_2\xi_2$，($k_1, k_2\in R$, 且 k_1, k_2 不同时为0)

对$\lambda_3=3$, 解方程组$(3\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$, 由于

$$3\mathbf{E}-\mathbf{A}=\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ -1 & 1 & 2\end{pmatrix}\overset{\text{行变换}}{\sim}\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}$$

得同解方程：$\begin{cases}x_1=x_3\\ x_2=-x_3\end{cases}$，基础解系为$\xi_3=(1, -1, 1)^T$.

所以 $k\xi_3$ ($k\in R$, 且 $k\neq0$) 是属于$\lambda_3=3$的全部特征向量.

例3 设方阵$\mathbf{A}$可逆，且λ是$\mathbf{A}$的特征值，求证：

$1/\lambda$是$\mathbf{A}^{-1}$的特征值.

证 首先证明$\lambda\neq 0$. 用反证法：假设$\lambda=0$是$\mathbf{A}$的特征值，则

$$|0\mathbf{E}-\mathbf{A}|=|-\mathbf{A}|=0,$$

这与$\mathbf{A}$可逆矛盾，故$\lambda\neq 0$.

再设ξ是$\mathbf{A}$对应特征值λ的特征向量，则

$$\mathbf{A}\xi=\lambda\xi \Longrightarrow \mathbf{A}^{-1}\xi=(1/\lambda)\xi$$

所以$(1/\lambda)$是$\mathbf{A}^{-1}$的特征值，而且与$\mathbf{A}$有相同的特征向量.

证毕.

例5.3 设方阵A满足条件$A^2=A$ (此时称A为幂等矩阵). 求证:

A的特征值只能是1或0.

思考: 哪些常见矩阵是幂等的?

证 设λ是A的特征值, ξ是属于λ的特征向量, 则

$$A\xi=\lambda\xi.$$

于是

$$A^2\xi = A(A\xi)= A(\lambda \xi)= \lambda A\xi= \lambda^2\xi.$$

又因为$A^2=A$, 所以

$$A^2\xi= A\xi= \lambda\xi.$$

综上可知

$$\lambda^2\xi= \lambda\xi,$$

也即

$$(\lambda^2- \lambda)\xi= \mathbf{0}.$$

由于$\xi\neq\mathbf{0}$, 故必有$\lambda^2- \lambda=0$, 因此λ可能的取值只能为1或0.

证毕.

二. 特征值和特征向量的性质

由于

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

$$=\lambda^n-(a_{11}+a_{22}+\ldots+a_{nn})\lambda^{n-1}+\ldots+(-1)^n|\mathbf{A}|, \quad (5.3)$$

设这个多项式的n个根(也即矩阵**A**的全部n个特征值)分别为$\lambda_1, \lambda_2,\ldots, \lambda_n$, 则此多项式又可以写为:

$$\det(\lambda\mathbf{E}-\mathbf{A})=(\lambda-\lambda_1)(\lambda-\lambda_2)\ldots(\lambda-\lambda_n)$$

$$=\lambda^n-(\lambda_1+\lambda_2+\ldots+\lambda_n)\lambda^{n-1}+\ldots+(-1)^n \lambda_1\lambda_2\ldots \lambda_n, \qquad (5.4)$$

比较式(5.3)与(5.4)的各项系数, 可得重要而有趣的性质:

性质5.1(特征值的性质) 设λ_1, λ_2,…, λ_n是n阶方阵**A**的全部特征值，则

$$\lambda_1+\lambda_2+\ldots+\lambda_n = a_{11}+a_{22}+\ldots+a_{nn},$$

$$\lambda_1\lambda_2\ldots\lambda_n = \det(\mathbf{A}).$$

注：通常称矩阵**A**的对角线之和$a_{11}+a_{22}+\ldots+a_{nn}$为矩阵**A**的迹，记作tr(**A**).

推论 n阶方阵**A**可逆的充要条件是：

A的n个特征值均不为零.

性质5.2 设λ是方阵$\mathbf{A}$的特征值，ξ是$\mathbf{A}$的属于λ的特征向量，$p(\lambda)$是关于λ的多项式，也即

$$p(\lambda)=a_0+a_1\lambda+\ldots+a_m\lambda^m\ ,$$

则$p(\lambda)$一定是矩阵$p(\mathbf{A})$的特征值，且ξ是矩阵$p(\mathbf{A})$的属于特征值$p(\lambda)$的特征向量．其中

$$p(\mathbf{A})=a_0\mathbf{E}+a_1\mathbf{A}+\ldots+a_m\mathbf{A}^m.$$

证明 由已知条件，有$\mathbf{A}\xi=\lambda\xi$，所以

$$\begin{aligned}p(\mathbf{A})\xi &= (a_0\mathbf{E}+a_1\mathbf{A}+\ldots+a_m\mathbf{A}^m)\xi\\ &=a_0\mathbf{E}\xi+a_1\mathbf{A}\xi+\ldots+a_m\mathbf{A}^m\xi\\ &= a_0\xi+a_1\lambda\xi+\ldots+a_m\lambda^m\xi\\ &= (a_0+a_1\lambda+\ldots+a_m\lambda^m)\xi= p(\lambda)\xi,\end{aligned}$$

所以$p(\lambda)$是的矩阵$p(\mathbf{A})$的特征值，ξ是$p(\mathbf{A})$的属于$p(\lambda)$的特征向量．证毕．

推论 设λ是方阵A的特征值，ξ是A的属于λ的特征向量，则λ^k (k为正整数)一定是矩阵A^k的特征值，且ξ是矩阵A^k的属于特征值λ^k的特征向量.

注1: 设A为可逆矩阵，λ是方阵A的特征值，ξ是A的属于λ的特征向量，则λ^k (k为整数)一定是矩阵A^k的特征值，且ξ是矩阵A^k的属于特征值λ^k的特征向量.

此时k可取负整数. 其中$A^{-k}=(A^{-1})^k$.

注2: 当矩阵A可逆时，性质5.2中的多项式$p(\lambda)$也可以含有负幂指数的项.

性质5.3 (特征向量性质I) 设$\lambda_1, \lambda_2, \ldots, \lambda_s$是方阵**A**的互异特征值，$\xi_1, \xi_2, \ldots, \xi_s$是分别属于$\lambda_1, \lambda_2, \ldots, \lambda_s$的特征向量，则$\xi_1, \xi_2, \ldots, \xi_s$线性无关.

证明 对(序号) s用数学归纳法. 因$\xi_1 \neq \mathbf{0}$，故ξ_1线性无关.

假设$\xi_1, \xi_2, \ldots, \xi_{s-1}$线性无关，以下往证$\xi_1, \xi_2, \ldots, \xi_{s-1}, \xi_s$也线性无关.

采用反证法. 若$\xi_1, \xi_2, \ldots, \xi_{s-1}, \xi_s$线性相关，则$\xi_s$可由$\xi_1, \xi_2, \ldots, \xi_{s-1}$线性表示，于是有

$$\xi_s = k_1\xi_1 + k_2\xi_2 + \ldots + k_{s-1}\xi_{s-1}. \qquad (*1)$$

用**A**左乘上式可得： $\mathbf{A}\xi_s = \mathbf{A}(k_1\xi_1 + k_2\xi_2 + \ldots + k_{s-1}\xi_{s-1}).$

即： $$\lambda_s \xi_s = k_1\lambda_1\xi_1 + k_2\lambda_2\xi_2 + \ldots + k_{s-1}\lambda_{s-1}\xi_{s-1}. \qquad (*2)$$

(*1)式两端左乘λ_s可得：$\lambda_s\xi_s = k_1\lambda_s\xi_1 + k_2\lambda_s\xi_2 + \ldots + k_{s-1}\lambda_s\xi_{s-1}.$ (*3)

(*2)与(*3)式相减可得……

(*2)与(*3)式相减可得

$$\mathbf{0}=k_1(\lambda_1-\lambda_s)\xi_1+k_2(\lambda_2-\lambda_s)\xi_2+\ldots+k_{s-1}(\lambda_{s-1}-\lambda_s)\xi_{s-1}.$$

由于$\xi_1, \xi_2,\ldots, \xi_{s-1}$线性无关, 所以

$$k_1(\lambda_1-\lambda_s)=k_2(\lambda_2-\lambda_s)=\ldots= k_{s-1}(\lambda_{s-1}-\lambda_s)=0.$$

又由于$\lambda_1, \lambda_2,\ldots, \lambda_s$互异, 所以可知

$$k_1= k_2=\ldots= k_{s-1}=0,$$

代入(*1)式会得到

$$\xi_s =\mathbf{0},$$

这与$\xi_s \neq \mathbf{0}$矛盾(特征向量不能为零向量).

因此, 假设错误. 所以$\xi_1, \xi_2,\ldots, \xi_s$必然线性无关. 证毕.

命题含义:

矩阵**A**的属于不同特征值的特征向量必然线性无关.

性质5.4(特征向量性质II) 设λ_1, λ_2是A的两个互异特征值, $\xi_1,\ldots,\xi_s$是属于λ_1的线性无关的特征向量, $\eta_1,\ldots,\eta_t$是属于λ_2的线性无关的特征向量, 则$\xi_1,\ldots,\xi_s,\eta_1,\ldots,\eta_t$线性无关.

证明 设有常数$k_1,\ldots,k_s,l_1,\ldots,l_t$, 使

$$k_1\xi_1+\ldots+k_s\xi_s+l_1\eta_1+l_2\eta_2+\ldots+l_t\eta_t=\mathbf{0}. \qquad (*)$$

记: $\xi=k_1\xi_1+\ldots+k_s\xi_s$, $\eta=l_1\eta_1+\ldots+l_t\eta_t$.

假设$\xi\neq\mathbf{0}$, 那么$\eta\neq\mathbf{0}$ (否则若$\eta=\mathbf{0}$, 则由式(*)即知$\xi+\eta=\mathbf{0}$, 可得$\xi=\mathbf{0}$).

在上述假设下, 易见ξ是属于λ_1的特征向量, 这是因为

$$\mathbf{A}\xi=\mathbf{A}(k_1\xi_1+\ldots+k_s\xi_s)=k_1\mathbf{A}\xi_1+\ldots+k_s\mathbf{A}\xi_s$$

$$=k_1\lambda_1\xi_1+\ldots+k_s\lambda_1\xi_s=\lambda_1(k_1\xi_1+\ldots+k_s\xi_s)$$

$$=\lambda_1\xi.$$

同理, 在此假设下, η也是属于λ_2的特征向量. 但是… (见下页)

但是由(*)式知$\xi+\eta=\mathbf{0}$, 所以得到结论"ξ, η线性相关", 这与性质5.3矛盾. 于是前面假设"$\xi\neq\mathbf{0}$, $\eta\neq\mathbf{0}$"错误, 故只能有

$$\xi=\eta=\mathbf{0},$$

也即

$$k_1\xi_1+\ldots+k_s\xi_s=\mathbf{0},$$

$$l_1\eta_1+\ldots+l_t\eta_t=\mathbf{0}.$$

又由$\xi_1, \ldots, \xi_s$是线性无关的, 可得

$$k_1=\ldots=k_s=0,$$

再由$\eta_1,\ldots, \eta_t$是线性无关的, 可得

$$l_1=\ldots=l_t=0,$$

观察(*)式知: $\xi_1,\ldots, \xi_s, \eta_1,\ldots, \eta_t$是线性无关的. 证毕.

注: 以上性质5.4对于矩阵A有任意多个互异的特征值的情形也是成立的.

例4 设3阶方阵$\mathbf{A}$的特征值为1, -1, 2, 求$|\mathbf{A}^*+3\mathbf{A}-2\mathbf{E}|$.

解 由于$\mathbf{A}$的特征值都不为0, 故$\mathbf{A}$可逆. 而$|\mathbf{A}|= -2$,

于是 $\mathbf{A}^*=|\mathbf{A}|\mathbf{A}^{-1}= -2\mathbf{A}^{-1}$.

根据性质5.1

因此 $$\mathbf{A}^*+3\mathbf{A}-2\mathbf{E}= -2\mathbf{A}^{-1}+3\mathbf{A}-2\mathbf{E} = p(\mathbf{A}),$$

其中$p(\lambda)$为:

$$p(\lambda)= -2\lambda^{-1}+3\lambda -2.$$

由于方阵$\mathbf{A}$的特征值为1, -1, 2, 根据性质5.2的推论可知

$p(\mathbf{A})$的3个特征值为:

$$p(1)= -1, p(-1)= -3, p(2)= 3.$$

于是

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

根据性质5.1

§2 矩阵相似对角化

一. 相似矩阵

定义5.3 设A, B都是n阶方阵. 若存在可逆矩阵P, 使

$$P^{-1}AP=B,$$

则称B是A的**相似矩阵**, 或说矩阵 A与B**相似**. 对A进行运算 $P^{-1}AP=B$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**.

注: 若两个矩阵相似, 它们必等价. 但等价的矩阵未必相似.

矩阵的相似关系具有下述性质:

(i) 反身性: A必然与A自身相似;

(ii) 对称性: 若A与B相似, 则B必然与A也相似;

(iii) 传递性: 若A与B相似, B与C相似, 则A必然与C也相似.

定理5.1 相似矩阵有相同的特征多项式，因此也有相同的特征值.

证 若矩阵A与B相似，则存在矩阵P，使$P^{-1}AP=B$，故

$$|\lambda E-B|=|P^{-1}(\lambda E)P-P^{-1}AP|=|P^{-1}(\lambda E-A)P|$$

$$=|P^{-1}|\,|\lambda E-A|\,|P|=|\lambda E-A|.$$ 证毕.

注：定理5.1的逆命题不成立. 例如矩阵

$$\begin{pmatrix}1&0\\0&1\end{pmatrix} \text{和} \begin{pmatrix}1&1\\0&1\end{pmatrix}$$

的特征多项式都是$(\lambda-1)^2$，

但以上两个矩阵不相似(具体理由以后会说明).

以上反例表明：即使矩阵A的特征值与B的特征值完全相同，A与B也未必相似.

推论 若n阶方阵**A**与对角矩阵

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

相似，则$\lambda_1, \lambda_2, \ldots, \lambda_n$是矩阵**A**的n个特征值.

注1: 由以上可见，如果不计主对角线上元素的顺序，则与**A**相似的对角矩阵是唯一的.

注2: 以上推论中，把对角矩阵**Λ**改为上三角阵(或下三角阵)，结论也是成立的.

例5 设矩阵 $A=\begin{pmatrix}2 & -1 & 0\\ a & 2 & 0\\ 1 & -1 & 1\end{pmatrix}$ 与矩阵 $B=\begin{pmatrix}1 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & b\end{pmatrix}$ 相似，求a, b的值.

解 由于相似矩阵有相同的特征值，也就有相同的迹，也有相同的行列式，所以可得:

$$2+2+1=\mathrm{tr}(\mathbf{A})=\mathrm{tr}(\mathbf{B})=1+3+b,$$

$$4+a=\det(\mathbf{A})=\det(\mathbf{B})=3b,$$

于是得解:

$$a=-1,\ b=1\ .$$

如果$\mathbf{A}$与$\mathbf{B}$相似，即存在可逆矩阵$\mathbf{Q}$，使得$\mathbf{A}=\mathbf{QBQ}^{-1}$，则

$$\mathbf{A}^k=(\mathbf{QBQ}^{-1})^k$$

$$=\mathbf{QBQ}^{-1}\mathbf{QBQ}^{-1}\mathbf{QBQ}^{-1}\cdots\cdots\cdots\ \mathbf{QBQ}^{-1}\mathbf{QBQ}^{-1}$$

$$=\mathbf{QB}^k\mathbf{Q}^{-1},$$

而对于多项式$p(\lambda)=a_0+a_1\lambda+\ldots+a_m\lambda^m$，矩阵$p(\mathbf{A})$此时可化为

$$p(\mathbf{A})=p(\mathbf{QBQ}^{-1})$$

$$=a_0\mathbf{E}+a_1\mathbf{QBQ}^{-1}+\ldots+a_m(\mathbf{QBQ}^{-1})^m$$

$$=a_0\mathbf{QEQ}^{-1}+a_1\mathbf{QBQ}^{-1}+\ldots+a_m\mathbf{QB}^m\mathbf{Q}^{-1}$$

$$=\mathbf{Q}(a_0\mathbf{E}+a_1\mathbf{B}+\ldots+a_m\mathbf{B}^m)\mathbf{Q}^{-1}$$

$$=\mathbf{Q}p(\mathbf{B})\mathbf{Q}^{-1}.$$

也就是说，若矩阵$\mathbf{A}$与$\mathbf{B}$相似，则$\mathbf{A}^k$与$\mathbf{B}^k$相似，且$p(\mathbf{A})$与$p(\mathbf{B})$相似，其中$p(.)$代表某个给定的多项式.

特别地，若矩阵**A**与对角矩阵**Λ**相似，其中对角阵**Λ**形如：

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

则存在可逆矩阵**Q**，使得$\mathbf{A}=\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$，则有

$$\mathbf{A}^{k}=\mathbf{Q}\boldsymbol{\Lambda}^{k}\mathbf{Q}^{-1},$$

由对角矩阵**Λ**的形状可知：

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

而对于多项式$p(\lambda)=a_0+a_1\lambda+\ldots+a_m\lambda^m$, 矩阵$p(\mathbf{A})$必可化为

$$p(\mathbf{A})=\mathbf{Q}p(\mathbf{\Lambda})\mathbf{Q}^{-1}.$$

其中矩阵$p(\mathbf{\Lambda})=a_0\mathbf{E}+a_1\mathbf{\Lambda}+\ldots+a_m\mathbf{\Lambda}^m$, 可知其为对角阵, 形如:

$$p(\Lambda)=\begin{pmatrix} p(\lambda_1) & & & \\ & p(\lambda_2) & & \\ & & \ddots & \\ & & & p(\lambda_n) \end{pmatrix}$$

以上给出了当矩阵$\mathbf{A}$相似于某对角矩阵时求$\mathbf{A}$的幂$\mathbf{A}^k$以及多项式$p(\mathbf{A})$的有效方法.

问题1: 是否任意方阵$\mathbf{A}$都能与对角矩阵相似?

问题2: 若$\mathbf{A}$与对角阵$\mathbf{\Lambda}$相似, 如何求$\mathbf{\Lambda}$与相似变换矩阵$\mathbf{Q}$?

这就是矩阵的相似对角化问题.

二. 与对角矩阵相似的条件

假设n阶方阵A与对角矩阵

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

相似. 也就是存在可逆矩阵P, 使得

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

这等价于

$$\mathbf{A}\mathbf{P}=\mathbf{P}\Lambda. \qquad (***)$$

若记$\mathbf{P}=(\xi_1, \xi_2,\dots, \xi_n)$, 其中$\xi_1, \xi_2,\dots, \xi_n$是矩阵P的列向量组, 则(***)式可具体写为:

$$\mathbf{A}(\xi_1,\xi_2,\ldots,\xi_n)=(\xi_1,\xi_2,\ldots,\xi_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n\end{pmatrix},$$

即

$$(\mathbf{A}\xi_1, \mathbf{A}\xi_2,\ldots, \mathbf{A}\xi_n)=(\lambda_1\xi_1, \lambda_2\xi_2,\ldots, \lambda_n\xi_n),$$

也即

$$\mathbf{A}\xi_i= \lambda_i\xi_i,\ i=1, 2,\ldots, n.$$

因为矩阵$\mathbf{P}$可逆，所以$\mathbf{P}$的n个列向量$\xi_1, \xi_2,\ldots, \xi_n$线性无关.因而可见：若矩阵$\mathbf{A}$与对角矩阵相似，则$\mathbf{A}$必有n个线性无关的特征向量.

反之，设矩阵$\mathbf{A}$有n个线性无关的特征向量$\xi_1, \xi_2,\ldots, \xi_n$,且

$$\mathbf{A}\xi_i=\lambda_i\xi_i,\ i=1, 2,\ldots, n,$$

令$P=(\xi_1, \xi_2,\dots, \xi_n)$，则$P$可逆，且有

$$AP=(A\xi_1, A\xi_2,\dots, A\xi_n)=(\lambda_1\xi_1, \lambda_2\xi_2,\dots, \lambda_n\xi_n)$$

$$=(\xi_1,\xi_2,\dots,\xi_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n\end{pmatrix}$$

$$=P\Lambda.$$

于是$AP=P\Lambda$，也即

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

也就是说此时矩阵A与对角矩阵Λ相似.

定理5.2 n阶矩阵A与对角矩阵相似的充分必要条件是：矩阵A有n个线性无关的特征向量.

推论 若n阶矩阵A有n个互异的特征值，则A与对角矩阵相似

可见，前面的分析不但证明了定理5.2，还给出了相似变换矩阵$\mathbf{P}$和对角矩阵$\mathbf{\Lambda}$的求法.

例如例1中的矩阵$\mathbf{A}=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}$

没有3个线性无关的特征向量，故$\mathbf{A}$不能与对角矩阵相似.

而例2中的矩阵$\mathbf{A}=\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}$

由于其3个特征值为$\lambda_1=\lambda_2=1$，$\lambda_3=3$，对应的特征向量：

$$\xi_1=(1, 1, 0)^T, \xi_2=(0, 0, 1)^T, \xi_3=(1, -1, 1)^T$$

线性无关，所以...

取相似变换矩阵$\mathbf{P}=(\xi_1, \xi_2, \xi_3)=\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$

可求得$\mathbf{P}$的逆矩阵为

$$\mathbf{P}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}$$

与$\mathbf{A}$相似的对角矩阵为

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

$$= \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 3 \end{pmatrix}$$

定义 方程$f(x)=0$的解α称为它的**根**，或称为函数$f(x)$的**零点**.

定义 若

$$f(x)=(x-\alpha)^m h(x),$$

其中$h(x)$在$x=\alpha$处连续且$h(\alpha)\neq 0$，则称α是方程$f(x)=0$的**m重根**.

注： 由以上定义可知：

λ_0是n阶矩阵**A**的k重特征值，当且仅当其特征多项式满足：

$$f(\lambda)=det(\lambda\boldsymbol{E}-\boldsymbol{A})=(\lambda-\lambda_0)^k p(\lambda),$$

其中$p(\lambda)$在$\lambda=\lambda_0$处连续且$p(\lambda_0)\neq 0$.

可以理解为：多项式$p(\lambda)$不再含有因式$(\lambda-\lambda_0)$.

为了进一步研究矩阵相似对角化问题，再给出矩阵特征值、特征向量的如下性质.

定理5.3 (特征向量性质III) 设λ_0是n阶矩阵**A**的k重特征值，则属于λ_0的线性无关的特征向量的个数不大于k.

证明： 设$\xi_1, \ldots, \xi_t$是属于λ_0的线性无关的特征向量. 则存在向量$\xi_{t+1},\ldots, \xi_n$使$\xi_1, \xi_2,\ldots, \xi_n$线性无关(可将线性无关组$\xi_1,\ldots, \xi_t$ 不断地添加新成员扩为

$$\xi_1, \ldots, \xi_t, \xi_{t+1}, \ldots, \xi_n,$$

使之成为极大线性无关组).

$\xi_{t+1}, \xi_{t+2},\ldots, \xi_n$ 未必是**A**的特征向量

由于$\xi_1,\ldots, \xi_n$是R^n的一个基，故$\mathbf{A}\xi_{t+1},\ldots, \mathbf{A}\xi_n$都能被$\xi_1,\ldots, \xi_n$线性表示，具体可表示为：

$$\mathbf{A}\xi_j = c_{1j}\xi_1 + \ldots + c_{tj}\xi_t + c_{t+1,j}\xi_{t+1} + \ldots + c_{nj}\xi_n,$$

$$j = t+1, t+2, \ldots, n.$$

令矩阵$P=(\xi_1,\xi_2,\ldots,\xi_n)$，则$P$是可逆矩阵，且可得

$$\begin{aligned}AP&=(A\xi_1,\ldots,A\xi_t,A\xi_{t+1},\ldots,A\xi_n)\\&=(\lambda_0\xi_1,\ldots,\lambda_0\xi_t,A\xi_{t+1},\ldots,A\xi_n)\end{aligned}$$

$$=(\xi_1,\ldots,\xi_t,\xi_{t+1},\ldots,\xi_n)\begin{pmatrix}\lambda_0 & & & c_{1,t+1} & \cdots & c_{1,n}\\ & \cdots & & \cdots & & \cdots\\ & & \lambda_0 & c_{t,t+1} & \cdots & c_{t,n}\\ & & & c_{t+1,t+1} & \cdots & c_{t+1,n}\\ & & & \cdots & & \cdots\\ & & & c_{n,t+1} & \cdots & c_{n,n}\end{pmatrix}$$

$$=P\begin{pmatrix}\lambda_0E_t & C_1\\ O & C_2\end{pmatrix}.$$

其中E_t是t阶单位矩阵

因此可得 $P^{-1}AP=\begin{pmatrix}\lambda_0E_t & C_1\\ O & C_2\end{pmatrix}$. 由此可知

矩阵A与$B=\begin{pmatrix}\lambda_0E_t & C_1\\ O & C_2\end{pmatrix}$相似.

所以，A与B有相同的特征多项式，即

$$|\lambda E-A|=|\lambda E-B|=\begin{vmatrix}\lambda E_t-\lambda_0 E_t & -C_1 \\ O & \lambda E_{n-t}-C_2\end{vmatrix}$$

$$=|\lambda E_t-\lambda_0 E_t||\lambda E_{n-t}-C_2|$$

$$=(\lambda-\lambda_0)^t|\lambda E_{n-t}-C_2|.$$

因此，矩阵A的特征值λ_0的重数$k\geq t$. (其中t代表矩阵A的属于λ_0的线性无关的特征向量的个数)

证毕.

定义 某方阵的特征方程的重数为1的根称为此矩阵的**单特征值**.

推论1 属于矩阵的单特征值的线性无关的特征向量有且仅有一个.

注1: 特征值λ_i作为特征方程的根的重数也称为λ_i的**代数重数**, 属于特征值λ_i的线性无关的特征向量的个数也称为λ_i的**几何重数**.

注2: 定理5.3的大意:

$$\text{特征值}\lambda_i\text{的代数重数} \geq \text{特征值}\lambda_i\text{的几何重数}.$$

任意n阶方阵A的特征值与特征向量的关系的示意图

$A \in R^{n\times n}$

设A有s个互异的特征值

特征值的重数	k_1	k_2	……	k_s	$\sum_{j=1}^{s} k_j = n$
特征值	λ_1	λ_2	……	λ_s	
特征向量	属于λ_1的线性无关的特征向量 $\xi_1^{(1)} \ \cdots \ \xi_{m_1}^{(1)}$	属于λ_2的线性无关的特征向量 $\xi_1^{(2)} \ \cdots \ \xi_{m_2}^{(2)}$	……	属于λ_s的线性无关的特征向量 $\xi_1^{(s)} \ \cdots \ \xi_{m_s}^{(s)}$	
特征向量重数	m_1	m_2	……	m_s	$\sum_{j=1}^{s} m_j \le n$
	$m_1 \le k_1$	$m_2 \le k_2$	……	$m_s \le k_s$	

推论2 设$\lambda_1,\ldots,\lambda_s$是n阶矩阵**A**的全部互异特征值，其(代数)重数依次为$k_1,\ldots,k_s$ (且$k_1+k_2+\ldots+k_s=n$)，则**A**与对角矩阵相似的充分必要条件是:

$$R(\lambda_i \mathbf{E}-\mathbf{A})=n-k_i,\ i=1,2,\ldots,s.$$

对所有特征值都要成立

证明： **A**与对角矩阵相似. $\Longleftrightarrow$ **A**有n个线性无关的特征向量.

$\Longleftrightarrow$ 属于**A**的每个特征值λ_i $(i=1,2,\ldots,s)$的线性无关的特征向量个数恰为k_i个. (否则，假设某特征值λ_t的线性无关的特征向量个数$<k_t$，由于属于矩阵**A**的每个特征值λ_i，$(i=1,2,\ldots,s,)$的线性无关的特征向量不多于k_i个，会得到**A**的全体线性无关的特征向量个数$<k_1+k_2+\ldots+k_s=n$. 矛盾.)

$\Longleftrightarrow$ 每个齐次线性方程组$(\lambda_i \mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$的基础解系恰有$k_i$个解向量.

$\Longleftrightarrow$ $R(\lambda_i \mathbf{E}-\mathbf{A})=n-k_i,\ i=1,2,\ldots,s.$

证毕.

例6 a, b为何值时，矩阵$A=\begin{pmatrix} 4 & 2 & b \\ -1 & 1 & 3 \\ 0 & 0 & a \end{pmatrix}$与对角矩阵相似？

解 矩阵**A**的特征多项式为

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

$$=(\lambda-a)(\lambda^2-5\lambda+6)=(\lambda-a)(\lambda-2)(\lambda-3).$$

可见，**A**的特征值是$\lambda_1=a, \lambda_2=2, \lambda_3=3$.

所以，当$a\neq 2$且$a\neq 3$时，矩阵**A**必与对角矩阵相似.

当$a=2$时，$\lambda=2$是**A**的二重特征值，由于

$$2E-A=\begin{pmatrix}-2 & -2 & -b \\ 1 & 1 & -3 \\ 0 & 0 & 0\end{pmatrix} \sim \begin{pmatrix}1 & 1 & -3 \\ 0 & 0 & -b-6 \\ 0 & 0 & 0\end{pmatrix}$$

所以, b=-6时, R(2**E**-**A**)=1=3-特征值重数, 此时**A**与对角矩阵相似.

当a=3时, λ=3是**A**的二重特征值, 由于

$$3E-A=\begin{pmatrix}-1 & -2 & -b \\ 1 & 2 & -3 \\ 0 & 0 & 0\end{pmatrix} \sim \begin{pmatrix}1 & 2 & -3 \\ 0 & 0 & -b-3 \\ 0 & 0 & 0\end{pmatrix}$$

所以, b=-3时, R(3**E**-**A**)=1=3-特征值重数, 此时**A**与对角矩阵相似.

综上, 当$a\neq2$且$a\neq3$, 或$a=2$, $b=-6$, 或$a=3$, $b=-3$时, 矩阵**A**与对角矩阵相似.

对n阶方阵$\mathbf{A}$相似对角化(通过相似变换将$\mathbf{A}$化为对角矩阵)，可按如下步骤进行：

(1) 解特征方程$\det(\lambda\mathbf{E}-\mathbf{A})=0$，求出$\mathbf{A}$的全部互异的特征值$\lambda_1$, $\lambda_2,\ldots,\lambda_s$;

(2) 对每个特征值λ_i，(设其重数为k_i)，观察齐次线性方程组$(\lambda_i\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$系数矩阵的秩$R(\lambda_i\mathbf{E}-\mathbf{A})$，如果$R(\lambda_i\mathbf{E}-\mathbf{A})=n-k_i$， 则求出这个齐次线性方程组的一个基础解系$\boldsymbol{\alpha}_{i1}$，$\boldsymbol{\alpha}_{i2},\ldots$，$\boldsymbol{\alpha}_{i,k_i}$，它们是属于特征值$\lambda_i$的$k_i$个线性无关的特征向量；

以上过程中，若有某个$R(\lambda_i\mathbf{E}-\mathbf{A})>n-k_i$，则判断$\mathbf{A}$不能相似于对角矩阵，终止计算.

(3) 写出相似变换矩阵$\mathbf{P}=(\boldsymbol{\alpha}_{11}, \boldsymbol{\alpha}_{12},\ldots \boldsymbol{\alpha}_{1,k_1},\ldots, \boldsymbol{\alpha}_{s1}, \boldsymbol{\alpha}_{s2},\ldots \boldsymbol{\alpha}_{s,k_s})$以及相似变换的结果：

$$P^{-1}AP=\begin{pmatrix}\lambda_1 E_{k_1} & & & \\ & \lambda_2 E_{k_2} & & \\ & & \cdots & \\ & & & \lambda_s E_{k_s}\end{pmatrix},$$

其中, E_{k_1}, E_{k_2} …, E_{k_s}分别是k_1, k_2, … k_s阶单位矩阵.

注意: 以上对角阵中每个特征值λ_i的位置必须与相似变换矩阵$P=(\alpha_{11}, \alpha_{12},\dots \alpha_{1,k_1},\dots, \alpha_{s1}, \alpha_{s2},\dots \alpha_{s,k_s})$中属于$\lambda_i$的特征向量的位置互相对应.

§3　实对称矩阵的相似对角化

一. 实对称矩阵的特征值和特征向量

设复矩阵$A=(a_{ij})$，用 $\overline{a_{ij}}$ 表示a_{ij}的共轭复数，记

$$\overline{A}=(\overline{a_{ij}}).$$

称 $\overline{A}$ 为A的共轭矩阵. 显然，当A为实矩阵时，$\overline{A}=A$.

共轭矩阵具有下列性质:

(i) $\overline{A+B}=\overline{A}+\overline{B}$;

(ii) $\overline{\lambda A}=\overline{\lambda}\,\overline{A}$;

其中λ是常数

(iii) $\overline{AB}=\overline{A}\,\overline{B}$;

(iv) $\overline{A}^{T}=\overline{A^{T}}$.

当A为实对称矩阵时，有 $\overline{A}=A=A^{T}$.

需注意的是：对于一般的n阶**实方阵A**，其特征值**未必是实数**，但是……

定理5.4 实对称矩阵的特征值都是实数.

证 设λ为实对称矩阵$\mathbf{A}$的特征值,ξ是属于λ的特征向量,即

$$\mathbf{A}\xi = \lambda\xi.$$

对上式两端左乘 $\overline{\xi}^{\mathrm{T}}$, 可得

$$\overline{\xi}^{\mathrm{T}}\mathbf{A}\xi = \overline{\xi}^{\mathrm{T}}\lambda\xi = \lambda\,\overline{\xi}^{\mathrm{T}}\xi, \tag{*1}$$

对上式两端同时取共轭转置, 得

$$\overline{\xi}^{\mathrm{T}}\,\overline{\mathbf{A}}^{\mathrm{T}}\xi = \overline{\lambda}\,\overline{\xi}^{\mathrm{T}}\xi, \tag{*2}$$

由于$\mathbf{A}$是实对称矩阵, 因此 $\overline{\mathbf{A}}^{\mathrm{T}}=\mathbf{A}$, 故(*2)式可化为

$$\overline{\xi}^{\mathrm{T}}\mathbf{A}\xi = \overline{\lambda}\,\overline{\xi}^{\mathrm{T}}\xi, \tag{*3}$$

(*1)与(*3)相减可得

$$(\lambda-\overline{\lambda})\overline{\xi}^{\mathrm{T}}\xi = 0.$$

由于$\xi\neq\mathbf{0}$, 所以 $\overline{\xi}^{\mathrm{T}}\xi\neq 0$, 因此 $\lambda=\overline{\lambda}$, 故$\lambda$必为实数. 证毕.

注: 实对称矩阵的特征向量都可以取为实向量. (见P104)

性质5.3与性质5.4已表明:

对于一般的方阵A, 属于它的不同特征值的特征向量之间必然是线性无关的.

对于实对称矩阵A, 还有进一步的结果......

定理5.5 设λ_1，λ_2是实对称矩阵A的不同特征值，ξ_1，ξ_2分别是属于λ_1，λ_2的特征向量，则ξ_1，ξ_2是正交的.

证明： 由

$$\begin{cases} A\xi_1 = \lambda_1\xi_1, \\ A\xi_2 = \lambda_2\xi_2, \end{cases} \xrightarrow{\text{可得}} \begin{cases} \xi_2^T A\xi_1 = \xi_2^T(\lambda_1\xi_1) = \lambda_1(\xi_2^T\xi_1), & (1) \\ \xi_1^T A\xi_2 = \xi_1^T(\lambda_2\xi_2) = \lambda_2(\xi_1^T\xi_2), & (2) \end{cases}$$

ξ_1与ξ_2均可取为实向量，故$\xi_2^T\xi_1$、$\xi_1^T\xi_2$、$\xi_2^T A\xi_1$以及$\xi_1^T A\xi_2$均为实数，且显然有$\xi_2^T\xi_1 = (\xi_2^T\xi_1)^T = \xi_1^T\xi_2$，

且由$\xi_2^T A\xi_1$为实数，也可得

A为实对称

$$\xi_2^T A\xi_1 = (\xi_2^T A\xi_1)^T = \xi_1^T A^T\xi_2 = \xi_1^T A\xi_2$$

观察式(1)、(2)的左右两端可知：$\lambda_1\xi_2^T\xi_1 = \lambda_2\xi_1^T\xi_2 = \lambda_2\xi_2^T\xi_1$.

于是成立：$(\lambda_1 - \lambda_2)\xi_2^T\xi_1 = 0$.

由于$\lambda_1 \neq \lambda_2$，所以$\xi_2^T\xi_1 = 0$，即ξ_1，ξ_2正交. 证毕.

关于定理5.5证明的另一表述

证 由于

$$\mathbf{A}\xi_1 = \lambda_1\xi_1,\ \mathbf{A}\xi_2 = \lambda_2\xi_2,$$

因此

$$\lambda_1\xi_2^{\mathrm{T}}\xi_1 = \xi_2^{\mathrm{T}}(\lambda_1\xi_1) = \xi_2^{\mathrm{T}}(\mathbf{A}\xi_1) = (\xi_2^{\mathrm{T}}\mathbf{A})\xi_1$$
$$= (\mathbf{A}^{\mathrm{T}}\xi_2)^{\mathrm{T}}\xi_1 = (\mathbf{A}\xi_2)^{\mathrm{T}}\xi_1 = (\lambda_2\xi_2)^{\mathrm{T}}\xi_1 = \lambda_2\xi_2^{\mathrm{T}}\xi_1,$$

A是实对称矩阵

也即: $$\lambda_1\xi_2^{\mathrm{T}}\xi_1 = \lambda_2\xi_2^{\mathrm{T}}\xi_1.$$

于是

$$(\lambda_1 - \lambda_2)\,\xi_2^{\mathrm{T}}\xi_1 = 0.$$

由于$\lambda_1 \neq \lambda_2$, 所以$\xi_2^{\mathrm{T}}\xi_1 = 0$, 即$\xi_1$, ξ_2正交. 证毕.

例7 三阶实对称矩阵A满足$|\mathbf{E}+\mathbf{A}|=0$，且$\alpha_1=(1, 1, 1)^T$，$\alpha_2=(1, -1, 0)^T$都是齐次线性方程组$\mathbf{Ax=0}$的解，求矩阵A的所有特征值和特征向量.

解 由$|\mathbf{E}+\mathbf{A}|=0$知：$\lambda_1=-1$是矩阵A的一个特征值.

又由于$\mathbf{Ax=0}$的非零解都是A的属于的特征值0的特征向量，又由于代数重数不小于几何重数，因此特征值0的代数重数至少为2；而三阶方阵只有3个特征值，所以，$\lambda_2=\lambda_3=0$是矩阵A的二重特征值.

所以，矩阵A的全体特征值为：$\lambda_1=-1, \lambda_2=\lambda_3=0$.

特征值0的所有特征向量为：$k_1\alpha_1+k_2\alpha_2$，k_1, k_2取不全为零的任意常数.

记特征值-1的特征向量为$\alpha_3=(x_1, x_2, x_3)^T$，由于$\alpha_3$与$\alpha_1$，$\alpha_2$都正交，可得：

$$\begin{cases} x_1+x_2+x_3=0, \\ x_1-x_2=0, \end{cases}$$

其通解为

$$\begin{cases} x_1=c, \\ x_2=c, \\ x_3=-2c, \end{cases}$$

于是, 可取

$$\boldsymbol{\alpha}_3=(1, 1, -2)^T,$$

特征值-1的所有特征向量为: $k\boldsymbol{\alpha}_3$, k是任意非零常数.

二. 实对称矩阵正交相似对角化

定理5.6 设A是实对称矩阵，则必存在正交矩阵Q，使得$Q^{-1}AQ$ $(=Q^TAQ)$为对角矩阵.

下一个推论

证 对矩阵A的阶数n用数学归纳法. n=1时定理结论显然成立. 以下假设对n-1阶实对称矩阵，定理结论成立.

对于n阶实对称矩阵A，任取特征值λ_1和属于λ_1的特征向量ξ_1，且令ξ_1为单位向量. (若ξ_1原先不是单位向量，总可以进行单位化，且单位化后的向量仍为λ_1的特征向量.)

则有

$$A\xi_1 = \lambda_1\xi_1,$$

且在R^n中必有n-1个向量$\xi_2, \xi_3,\ldots, \xi_n$，使$\xi_1, \xi_2,\ldots, \xi_n$为$R^n$的一个规范正交基. (可以先将$\xi_1$扩充为含n个向量的某极大线性无关组，再对其做施密特正交化和单位化而得到.)

由于$\mathbf{A}\xi_2,\ldots,\mathbf{A}\xi_n$都可以由$\xi_1,\xi_2,\ldots,\xi_n$线性表示，即有

$$\mathbf{A}\xi_j = c_{1j}\xi_1 + c_{2j}\xi_2 + \ldots + c_{nj}\xi_n,\quad j=2,3,\ldots,n.$$

令矩阵$\mathbf{P}_1=(\xi_1,\xi_2,\ldots,\xi_n)$，则$\mathbf{P}_1$是正交矩阵，且有

$$\begin{aligned}
\mathbf{A}\mathbf{P}_1 &= \mathbf{A}(\xi_1,\xi_2,\ldots,\xi_n)\\
&= (\lambda_1\xi_1, \mathbf{A}\xi_2,\ldots,\mathbf{A}\xi_n),\\
&= \left(\lambda_1\xi_1, \sum_{k=1}^{n} c_{k2}\xi_k, \ldots, \sum_{k=1}^{n} c_{kn}\xi_k\right)\\
&= (\xi_1,\xi_2,\ldots,\xi_n)\begin{pmatrix} \lambda_1 & c_{12} & \cdots & c_{1n}\\ 0 & c_{22} & \cdots & c_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ 0 & c_{n2} & \cdots & c_{nn}\end{pmatrix},
\end{aligned}$$

这是矩阵$\mathbf{P}_1$

因此得到

$$\mathbf{P}_1^T\mathbf{A}\mathbf{P}_1=\mathbf{P}_1^{-1}\mathbf{A}\mathbf{P}_1=\begin{pmatrix}\lambda_1 & c_{12} & \cdots & c_{1n}\\ 0 & c_{22} & \cdots & c_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ 0 & c_{n2} & \cdots & c_{nn}\end{pmatrix}.$$

$\mathbf{P}_1$是正交矩阵

由于**A**是实对称矩阵，可知$\mathbf{P}_1^T\mathbf{A}\mathbf{P}_1$也必为实对称矩阵，因此还可知以上矩阵$\mathbf{P}_1^T\mathbf{A}\mathbf{P}_1$中的第一行中，$c_{12}=c_{13}=\ldots=c_{1n}=0$，且(右下角分块)矩阵

$$\mathbf{B}=\begin{pmatrix}c_{22} & \cdots & c_{2n}\\ \cdots & \cdots & \cdots\\ c_{n2} & \cdots & c_{nn}\end{pmatrix}$$

是n-1阶的实对称矩阵．也即

$$\mathbf{P}_1^T\mathbf{A}\mathbf{P}_1=\begin{pmatrix}\lambda_1 & \mathbf{0}^T\\ \mathbf{0} & \mathbf{B}\end{pmatrix},$$

其中**B**是n-1阶的实对称矩阵．由归纳假设，存在n-1阶正交矩阵**P**，使$\mathbf{P}^T\mathbf{B}\mathbf{P}$为对角矩阵，也即

$$\mathbf{P}^{-1}\mathbf{B}\mathbf{P}=\mathbf{P}^{\mathrm{T}}\mathbf{B}\mathbf{P}=\begin{pmatrix}\lambda_2 & & & \\ & \lambda_3 & & \\ & & \cdots & \\ & & & \lambda_n\end{pmatrix}.$$

取

$$\mathbf{P}_2=\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{P}\end{pmatrix},$$

则$\mathbf{P}_2$是n阶正交矩阵，且

$$\mathbf{P}_2^{\mathrm{T}}\begin{pmatrix}\lambda_1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{B}\end{pmatrix}\mathbf{P}_2=\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{P}^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\lambda_1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{B}\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{P}\end{pmatrix}$$

$$=\begin{pmatrix}\lambda_1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \mathbf{P}^{\mathrm{T}}\mathbf{B}\mathbf{P}\end{pmatrix}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n\end{pmatrix}\overset{\text{def}}{=}\boldsymbol{\Lambda}.$$

记为$\boldsymbol{\Lambda}$

观察、整理上述过程，可知

$$\mathbf{P}_2^{\mathrm{T}}\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1\mathbf{P}_2 = \mathbf{P}_2^{\mathrm{T}}(\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1)\mathbf{P}_2 = \mathbf{\Lambda}.$$

记$\mathbf{Q}=\mathbf{P}_1\mathbf{P}_2$，则$\mathbf{Q}$是正交矩阵，且有

$$\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q} = \mathbf{P}_2^{\mathrm{T}}\mathbf{P}_1^{\mathrm{T}}\mathbf{A}\mathbf{P}_1\mathbf{P}_2 = \mathbf{\Lambda}$$

$$= \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \cdots & \\ & & & \lambda_n \end{pmatrix}$$

因此，对任意n阶实对称矩阵$\mathbf{A}$，定理结论成立. 证毕.

返回定理5.6

推论 设λ_0是n阶实对称矩阵$\mathbf{A}$的k重特征值，则

$$R(\lambda_0\mathbf{E}-\mathbf{A}) = n-k,$$

也即：属于λ_0的线性无关的特征向量恰有k个.

证明思路： 由于实对称矩阵必与对角矩阵相似，根据定理5.3的推论2，k重特征值λ_0的线性无关的特征向量恰有k个，也即对应的特征方程(齐次线性方程组)的基础解系恰含k个向量，因此$R(\lambda_0\mathbf{E}-\mathbf{A})=n-k$.

注：以上结论的另一等价说法是：

实对称矩阵$\mathbf{A}$的每个特征值的代数重数都等于几何重数.

三. 实对称矩阵正交相似对角化的方法

将实对称矩阵**A**正交相似对角化(化为对角矩阵)，可按如下步骤进行:

(1) 解特征方程$\det(\lambda E-A)=0$，求出**A**的全部互异的特征值λ_1, $\lambda_2,\ldots,\lambda_s$;

(2) 对每个特征值λ_i，(设其重数为k_i)，求出齐次线性方程组$(\lambda_i \mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$的一个基础解系$\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2},\ldots, \boldsymbol{\alpha}_{i,k_i}$，它们是属于特征值$\lambda_i$的$k_i$个线性无关的特征向量;

(3) 对于上一步中每个特征值λ_i, $(i=1, 2,\ldots, s.)$ 将其对应的向量组$\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2},\ldots, \boldsymbol{\alpha}_{i,k_i}$规范正交化，得到$\xi_{i1}, \xi_{i2},\ldots, \xi_{i,k_i}$，它们是属于特征值$\lambda_i$的$k_i$个两两正交的单位特征向量;

(4) 写出正交矩阵$\mathbf{Q}=(\xi_{11}, \xi_{12},\ldots\xi_{1,k_1},\ldots, \xi_{s1}, \xi_{s2},\ldots\xi_{s,k_s})$以及正交变换的结果:

$$\mathbf{Q}^T\mathbf{A}\mathbf{Q}=\begin{pmatrix}\lambda_1\mathbf{E}_{k_1} & & & \\ & \lambda_2\mathbf{E}_{k_2} & & \\ & & \cdots & \\ & & & \lambda_s\mathbf{E}_{k_s}\end{pmatrix},$$

其中, $\mathbf{E}_{k_1}, \mathbf{E}_{k_2}, \ldots, \mathbf{E}_{k_s}$ 分别是 $k_1, k_2, \ldots k_s$ 阶单位矩阵.

注意: 以上对角阵中每个特征值 λ_i 的位置必须与正交矩阵 $\mathbf{Q}=(\xi_{11}, \xi_{12}, \ldots \xi_{1,k_1}, \ldots, \xi_{s1}, \xi_{s2}, \ldots \xi_{s,k_s})$ 中属于 λ_i 的特征向量的位置互相对应.

例8 设

$$\mathbf{A}=\begin{pmatrix}1&2&4\\2&1&4\\4&4&7\end{pmatrix}$$

求一个正交矩阵Q, 使$\mathbf{Q}^{\mathrm{T}}\mathbf{AQ}$为对角矩阵.

解 先求A的所有特征值. 由于矩阵A的特征多项式为:

$\det(\lambda\mathbf{E}-\mathbf{A})$

$$=\begin{vmatrix}\lambda-1&-2&-4\\-2&\lambda-1&-4\\-4&-4&\lambda-7\end{vmatrix}=\begin{vmatrix}\lambda-1&-1-\lambda&-4\\-2&\lambda+1&-4\\-4&0&\lambda-7\end{vmatrix}=\begin{vmatrix}\lambda-3&0&-8\\-2&\lambda+1&-4\\-4&0&\lambda-7\end{vmatrix}$$

$=(\lambda+1)(\lambda^2-10\lambda-11)=(\lambda+1)^2(\lambda-11)$

得特征值$\lambda_1=\lambda_2=-1$, $\lambda_3=11$.

然后再求相应的特征向量, 并对其规范正交化...

对$\lambda_1=\lambda_2=-1$，由于

$$-\mathbf{E}-\mathbf{A}=\begin{pmatrix}-2 & -2 & -4\\ -2 & -2 & -4\\ -4 & -4 & -8\end{pmatrix}\to\begin{pmatrix}1 & 1 & 2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

所以方程组$(-\mathbf{E}-\mathbf{A})\mathbf{x}=\mathbf{0}$等价于$x_1+x_2+2x_3=0$，其一基础解系为

$\boldsymbol{\alpha}_1=(-1, 1, 0)^T$，$\boldsymbol{\alpha}_2=(-2, 0, 1)^T$，将其正交化得：

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=(-1, 1, 0)^T,$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\frac{\boldsymbol{\alpha}_2^T\boldsymbol{\beta}_1}{\boldsymbol{\beta}_1^T\boldsymbol{\beta}_1}\boldsymbol{\beta}_1=\boldsymbol{\alpha}_2-\boldsymbol{\beta}_1=(-1, -1, 1)^T,$$

再单位化得：

$$\xi_1=\frac{1}{|\boldsymbol{\beta}_1|}\boldsymbol{\beta}_1=\left(\frac{-1}{\sqrt{2}},\ \frac{1}{\sqrt{2}},\ 0\right)^T,$$

$$\xi_2=\frac{1}{|\boldsymbol{\beta}_2|}\boldsymbol{\beta}_2=\left(\frac{-1}{\sqrt{3}},\ \frac{-1}{\sqrt{3}},\ \frac{1}{\sqrt{3}}\right)^T.$$

对$\lambda_3=11$，由于

$$\mathbf{11E}-\mathbf{A}=\begin{pmatrix}10 & -2 & -4\\ -2 & 10 & -4\\ -4 & -4 & 4\end{pmatrix}\to\begin{pmatrix}1 & 0 & -1/2\\ 0 & 1 & -1/2\\ 0 & 0 & 0\end{pmatrix},$$

所以方程组$(11\mathbf{E}\text{-}\mathbf{A})\mathbf{x}=\mathbf{0}$的一个基础解系为$\boldsymbol{\beta}_3=(1, 1, 2)^T$，

将其单位化得：$\xi_3=\dfrac{1}{|\boldsymbol{\beta}_3|}\boldsymbol{\beta}_3=\left(\dfrac{1}{\sqrt{6}}, \dfrac{1}{\sqrt{6}}, \dfrac{2}{\sqrt{6}}\right)^T$，

所以得正交矩阵：$\mathbf{Q}=(\xi_1, \xi_2, \xi_3)=\begin{pmatrix}\dfrac{-1}{\sqrt{2}} & \dfrac{-1}{\sqrt{3}} & \dfrac{1}{\sqrt{6}}\\ \dfrac{1}{\sqrt{2}} & \dfrac{-1}{\sqrt{3}} & \dfrac{1}{\sqrt{6}}\\ 0 & \dfrac{1}{\sqrt{3}} & \dfrac{2}{\sqrt{6}}\end{pmatrix}$，

而且……

而且使

$$\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 11 \end{pmatrix} = \mathrm{diag}(-1,\ -1,\ 11).$$

diag(-1, -1, 11)即对角线元素为-1, -1, 11的对角矩阵.

注：理论上，方程组$x_1+x_2+2x_3=0$的基础解系也可直接取为：

$$\boldsymbol{\xi}_1=\begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix},\ \boldsymbol{\xi}_2=\begin{pmatrix}-1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{pmatrix}.$$

这样，就不需要再进行规范正交化了.

此类操作的前提是：你得能“猜对”这些规范正交的特征向量.

例9 设三阶实对称矩阵**A**的秩等于2, $\alpha_1=(1, -1, 0)^T$和$\alpha_2=(1, 1, 2)^T$分别是**A**的属于特征值1和-1的特征向量, 求**A**.

解 由R(**A**)=2知$\lambda=0$是**A**的特征值. 所以, 矩阵**A**的三个特征值为: $\lambda_1=1, \lambda_2=-1, \lambda_3=0$.

又由于$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}_3$是分别属于三个不同的特征值的特征向量, 因而向量组$\boldsymbol{\alpha}_1$, $\boldsymbol{\alpha}_2$, $\boldsymbol{\alpha}_3$是正交向量组. 所以, 特征值$\lambda_3=0$的特征向量$\boldsymbol{\alpha}_3=(x_1, x_2, x_3)^T$应满足:

$$x_1-x_2=0,\ x_1+x_2+2x_3=0.$$

所以, 可取$\boldsymbol{\alpha}_3=(1, 1, -1)^T$. 将$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$单位化得:

$$\xi_1=\frac{1}{|\alpha_1|}\alpha_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix},\quad \xi_2=\frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\2\end{pmatrix},\quad \xi_3=\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\-1\end{pmatrix}$$

构造正交矩阵

$$Q=(\xi_1,\xi_2,\xi_3)=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \end{pmatrix}$$

则有 $\mathbf{Q}^{\mathrm{T}}\mathbf{A}\mathbf{Q}=\mathbf{\Lambda}$，所以有

$$\mathbf{A}=\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{T}}=\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \end{pmatrix}\begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix}$$

故得

$$\mathbf{A}=\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{T}}=\begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{2}{3} \end{pmatrix}$$

$$=\frac{1}{3}\begin{pmatrix} 1 & -2 & -1 \\ -2 & 1 & -1 \\ -1 & -1 & -2 \end{pmatrix}$$

建议自修题目：

P108: 例5.9.

P97: 例5.3.