## 6．2　相似矩阵与矩阵的对角化

在上节例 6．1 中，矩阵 $A=\begin{pmatrix}2 & 1\\ 1 & 2\end{pmatrix}$ 有特征值 1，3 ，相应的特征向量为 $p_1=\begin{pmatrix}-1\\ 1\end{pmatrix}$，$p_2=\begin{pmatrix}1\\ 1\end{pmatrix}$，令 $P=\left(p_1,p_2\right)=\begin{pmatrix}-1 & 1\\ 1 & 1\end{pmatrix}$，则 $AP=\left(Ap_1,Ap_2\right)=\left(p_1,3p_2\right)=\left(p_1,p_2\right)\begin{pmatrix}1 & 0\\ 0 & 3\end{pmatrix}=P\begin{pmatrix}1 & 0\\ 0 & 3\end{pmatrix}$，所以 $P^{-1}AP=\begin{pmatrix}1 & 0\\ 0 & 3\end{pmatrix}$．这样我们通过可逆矩阵 $P$，将矩阵 $A$ 化为对角矩阵，这个过程称为相似变换.

**定义 6．2**　设 $A,B$ 都是 $n$ 阶矩阵，若存在可逆矩阵 $P$，使得

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

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

“相似”是矩阵之间的一种关系，它具有下述性质（请学生课后证明）：

（1）**反身性**　对任意的方阵 $A$，$A$ 与 $A$ 相似；

（2）**对称性**　若 $A$ 与 $B$ 相似，则 $B$ 与 $A$ 相似；

（3）**传递性**　若 $A$ 与 $B$ 相似，$B$ 与 $C$ 相似，则 $A$ 与 $C$ 相似.

矩阵的相似关系是一种等价关系，我们可以将同阶的矩阵进行等价分类，即把所有相互相似的矩阵归为一类，我们将探讨同类的相似矩阵有什么样的共性，相似变换的不变量是什么？

相似矩阵还具有以下性质：

（1）若 $A$ 与 $B$ 相似，则 $R(A)=R(B)$.

**证**　因为 $A$ 与 $B$ 相似，即存在可逆矩阵 $P$，使 $B=P^{-1}AP$，所以 $R(A)=R(B)$.

（2）若 $A$ 与 $B$ 相似，则 $\left|A\right|=\left|B\right|$.

**证**　因为 $A$ 与 $B$ 相似，即存在可逆矩阵 $P$，使 $B=P^{-1}AP$，因此

$$\left|B\right|=\left|P^{-1}AP\right|=\left|P^{-1}\right|\left|A\right|\left|P\right|=\left|A\right|.$$

（3）若 $A$ 与 $B$ 相似，则 $A$ 与 $B$ 有相同的可逆性，且可逆时其逆也相似.

**证**　由（2）有 $\left|A\right|=\left|B\right|$，所以 $A$ 与 $B$ 有相同的可逆性.

设 $A$ 与 $B$ 相似，且 $A$ 与 $B$ 均可逆．因为 $A$ 与 $B$ 相似，即存在可逆矩阵 $P$，使 $B=P^{-1}AP$，故

$$B^{-1}=(P^{-1}AP)^{-1}=P^{-1}A^{-1}P,$$

即 $A^{-1}$ 与 $B^{-1}$ 相似.

（4）若 $A$ 与 $B$ 相似，则 $A^T$ 与 $B^T$ 也相似.

**证** 因为 $A$ 与 $B$ 相似，即存在可逆矩阵 $P$，使 $B=P^{-1}AP$，故

$$B^T=(P^{-1}AP)^T=P^TA^T(P^{-1})^T=P^TA^T(P^T)^{-1}.$$

令 $Q=(P^T)^{-1}$，则 $B^T=Q^{-1}A^TQ$，即 $A^T$ 与 $B^T$ 相似.

（5）若 $A$ 与 $B$ 相似，则矩阵多项式

$$f(A)=a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E$$

与

$$f(B)=a_mB^m+a_{m-1}B^{m-1}+\cdots+a_1B+a_0E \quad （m\geq 1为整数）$$

也相似.

**证** 因为 $A$ 与 $B$ 相似，即存在可逆矩阵 $P$，使 $B=P^{-1}AP$，故

$$\begin{aligned}f(B)&=a_m(P^{-1}AP)^m+a_{m-1}(P^{-1}AP)^{m-1}+\cdots+a_1(P^{-1}AP)+a_0E\\&=P^{-1}(a_mA^m)P+P^{-1}(a_{m-1}A^{m-1})P+\cdots+P^{-1}(a_1A)P+P^{-1}(a_0E)P\\&=P^{-1}(a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E)P\\&=P^{-1}f(A)P,\end{aligned}$$

即 $f(A)$ 与 $f(B)$ 相似.

**定理 6．1** 若 $n$ 阶矩阵 $A$ 与 $B$ 相似，则 $A$ 与 $B$ 有相同的特征多项式，从而 $A$ 与 $B$ 有相同的特征值.

**证** 因为 $A$ 与 $B$ 相似，即存在可逆矩阵 $P$，使 $B=P^{-1}AP$，因此

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

即 $A$ 与 $B$ 有相同的特征多项式，从而 $A$ 与 $B$ 有相同的特征值.

**推论** 若 $n$ 阶矩阵 $A$ 与对角矩阵 $\Lambda=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$ 相似，则 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 为 $A$ 的 $n$ 个特征值.

如果矩阵 $A$ 与对角矩阵 $\Lambda=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix}$ 相似，即存在可逆矩阵 $P$，使

$$P^{-1}AP=\Lambda=\begin{pmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{pmatrix},$$

则 $A=P\Lambda P^{-1}$，于是有

$$A^k=(P\Lambda P^{-1})(P\Lambda P^{-1})\cdots(P\Lambda P^{-1})=P\Lambda^k P^{-1}=P\begin{pmatrix}\lambda_1^k & 0 & \cdots & 0\\ 0 & \lambda_2^k & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n^k\end{pmatrix}P^{-1}.$$

若 $\varphi(\lambda)$ 为 $\lambda$ 的多项式，矩阵多项式 $\varphi(A)$ 可由下式求得

$$\varphi(A)=P\begin{pmatrix}\varphi(\lambda_1) & & & \\ & \varphi(\lambda_2) & & \\ & & \cdots & \\ & & & \varphi(\lambda_n)\end{pmatrix}P^{-1}.$$

由此可方便地计算 $A$ 的多项式 $\varphi(A)$ .

由此可见，一个能和对角矩阵相似的矩阵具有良好的性质，但遗憾的是，并不是每一个 $n$ 阶矩阵都能和一个对角矩阵相似．寻求一个相似变换矩阵 $P$ ，使得 $P^{-1}AP=\Lambda$ 为对角矩阵，称作把方阵 **$A$ 对角化**．究竟什么样的方阵能对角化？相似变换矩阵 $P$ 有什么样的特点？

假设 $n$ 阶方阵 $A$ 可对角化，即存在 $n$ 阶可逆矩阵 $P$ ，使 $P^{-1}AP=\Lambda$ 为对角阵，我们来讨论 $P$ 应满足的条件.

把 $P$ 用其列向量组表示为

$$P=(p_1,p_2,\cdots,p_n),$$

则 $p_1,p_2,\cdots,p_n$ 线性无关．由 $P^{-1}AP=\Lambda$，得 $AP=P\Lambda$，即

$$A(p_1,p_2,\cdots,p_n)=(p_1,p_2,\cdots,p_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}=(\lambda_1 p_1,\lambda_2 p_2,\cdots,\lambda_n p_n),$$

于是有 $Ap_i=\lambda_i p_i\ (i=1,2,\cdots,n)$ .

可见相似变换矩阵 $P$ 的列向量组 $p_1,p_2,\cdots,p_n$ 是 $A$ 的 $n$ 个线性无关的特征向量.

以上分析告诉我们：若 $n$ 阶方阵 $A$ 可对角化，则 $A$ 有 $n$ 个线性无关的特征向量.

反过来，若 $A$ 有 $n$ 个线性无关的特征向量 $p_1,p_2,\cdots,p_n$ ，它们对应的特征值分别为 $\lambda_1,\lambda_2,\cdots,\lambda_n$ ，即 $Ap_i=\lambda_i p_i\ (i=1,2,\cdots,n)$，把这 $n$ 个等式写成矩阵形式

$$(Ap_1\,,\,Ap_2\,,\cdots,\,Ap_n)=(\lambda_1 p_1\,,\,\lambda_2 p_2\,,\cdots,\,\lambda_n p_n)\,,$$

$$A(p_1\,,\,p_2\,,\cdots,\,p_n)=(p_1\,,\,p_2\,,\cdots,\,p_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}.$$

令 $P=(p_1\,,\,p_2\,,\cdots,\,p_n)$，由于 $p_1\,,\,p_2\,,\cdots,\,p_n$ 线性无关，故 $P$ 可逆．上式两端左乘 $P^{-1}$，得

$$P^{-1}AP=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

所以 $A$ 可对角化．

以上分析又告诉我们：若 $n$ 阶方阵 $A$ 有 $n$ 个线性无关的特征向量，则 $A$ 可对角化．

综合以上分析，可得下面定理：

**定理 6．2**　$n$ 阶方阵 $A$ 可对角化的充分必要条件是 $A$ 有 $n$ 个线性无关的特征向量．

结合特征向量的性质 1，可得以下推论：

**推论**　如果 $n$ 阶矩阵 $A$ 的 $n$ 个特征值互不相等，则 $A$ 可对角化．

说明："特征值互不相等"是可对角化的充分条件，但不是必要条件．

例 1．1、例 1．3、例 1．4 中的矩阵都可以对角化．

当 $A$ 的特征方程有重根时，就不一定有 $n$ 个线性无关的特征向量，从而不一定能对角化．例如例 1．5 中 $A$ 的特征方程有重根，却找不到 3 个线性无关的特征向量，因此例 1．5 中的矩阵 $A$ 不能对角化；而在例 1．4 中 $A$ 的特征方程也有重根，但能找到 3 个线性无关的特征向量，因此该矩阵 $A$ 能对角化．

补充一个定理：

**定理 6.3**　设 $\lambda_1\,,\,\lambda_2\,,\cdots,\,\lambda_s$ 为 $n$ 阶方阵 $A$ 的 $s$ 个互异特征根，它们的重数分别为 $n_1\,,\,n_2\,,\cdots,\,n_s$（$n_1+n_2+\cdots+n_s=n$），则 $A$ 可对角化的充要条件是 $R(A-\lambda_i E)=n-n_i\ (i=1,2,\cdots,s)$．

**证　充分性**

若 $R(A-\lambda_i E)=n-n_i\ (i=1,2,\cdots,s)$，则齐次线性方程组 $(A-\lambda_i E)x=0\ (i=1,2,\cdots,s)$ 的基础解系中分别含有 $n_1\,,\,n_2\,,\cdots,\,n_s$ 个解向量，从而 $A$ 有 $n_1+n_2+\cdots+n_s=n$ 个线性无关的解向量，所以 $A$ 可对角化．

**必要性**

若 $A$ 可对角化，即 $A$ 与对角阵 $\Lambda=diag\,(\lambda_1\,,\lambda_2\,,\cdots,\lambda_n)$ 相似，其中 $\lambda_1\,,\,\lambda_2\,,\cdots,\,\lambda_n$ 为 $A$ 的 $n$ 个特征值，故 $A-\lambda_i E$ 与 $\Lambda-\lambda_i E=diag\,(\lambda_1-\lambda_i\,,\,\lambda_2-\lambda_i\,,\cdots,\,\lambda_n-\lambda_i)$ 相似．因为 $\lambda_i$ 是 $A$ 的 $n_i$ 重特征

值，故对角阵$\Lambda-\lambda_i E$的主对角元中恰有$n_i$个等于 0，于是$R(A-\lambda_i E)=R(\Lambda-\lambda_i E)=n-n_i$.

**例 2．1** 设矩阵

$$A=\begin{bmatrix}3&2&2\\2&3&2\\2&2&3\end{bmatrix},\quad P=\begin{bmatrix}0&1&0\\1&0&1\\0&0&1\end{bmatrix},$$

$B=P^{-1}A^*P$，求$B+2E$的特征值.

**解** $A$的特征多项式为

$$|\lambda E-A|=\begin{vmatrix}\lambda-3&-2&-2\\-2&\lambda-3&-2\\-2&-2&\lambda-3\end{vmatrix}=(\lambda-1)^2(\lambda-7),$$

解特征方程$|\lambda E-A|=0$，求得$A$的 3 个特征值是 1，1，7.

因为$|A|=1\times1\times7=7$，所以$A^*$的 3 个特征值是 7，7，1.

因为$B$与$A^*$相似，故$B$的特征值为 7，7，1，从而$B+2E$的特征值为 9，9，3.

**例 2．2** 设矩阵$B=\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}$，已知矩阵$A$相似于矩阵$B$，则秩$(A-2E)$与秩$(A-E)$之和等于（　　）.

（A）2；　　（B）3；　　（C）4；　　（D）5.

**解** 因为$A$相似于$B$，故$A-2E$相似于$B-2E$，$A-E$相似于$B-E$，从而

$$R(A-2E)+R(A-E)=R(B-2E)+R(B-E)=3+1=4.$$

故选（C）.

**例 2．3** 已知$A=\begin{bmatrix}2&a&2\\5&b&3\\-1&1&-1\end{bmatrix}$有特征值$1$和$-1$，问$A$是否能对角化？

**解** 由已知得

$$\begin{cases}|E-A|=0,\\|-E-A|=0,\end{cases}$$

即

$$\begin{cases}-7(a+1)=0,\\3a-2b-3=0,\end{cases}$$

解得$a=-1$，$b=-3$.

由特征值的性质“所有特征值的和等于矩阵的迹”得$A$的第3个特征值为$-2$．因为$A$的所有特征值互异，故$A$可对角化.

**例 2．4** 设$A=\begin{pmatrix}0&0&1\\1&1&x\\1&0&0\end{pmatrix}$，问$x$为何值时，矩阵$A$能对角化?

**解** 解特征方程

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

得$A$的特征值为$\lambda_1=-1$，$\lambda_2=\lambda_3=1$.

对应单根$\lambda_1=-1$，必有$R(A+E)=3-1=2$，故矩阵$A$可对角化$\Leftrightarrow R(A-E)=3-2=1$．由

$$A-E=\begin{pmatrix}-1&0&1\\1&0&x\\1&0&-1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&0&-1\\0&0&x+1\\0&0&0\end{pmatrix},$$

得$x+1=0$，即$x=-1$.

因此，当$x=-1$时，矩阵$A$能对角化.

矩阵对角化的步骤（讲授算法之前，先分析算法的依据）:

（i）解特征方程$|\lambda E-A|=0$，求出$n$阶方阵$A$的特征值$\lambda_1,\lambda_2,\cdots,\lambda_n$；

（ii）对每个互异特征值$\lambda_1,\lambda_2,\cdots,\lambda_s$，解齐次线性方程组$(\lambda_i E-A)x=0\ (i=1,2,\cdots,s)$，分别求出它们的基础解系，再将这些基础解系合并为一个向量组，从而得到$A$的线性无关的特征向量组$p_1,p_2,\cdots,p_m$.

（iii）若$m<n$，则$A$不可对角化；若$m=n$，则$A$可对角化．令$P=(p_1,p_2,\cdots,p_n)$，则$P$可逆，且

$$P^{-1}AP=\begin{pmatrix}\lambda_1&0&\cdots&0\\0&\lambda_2&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&\lambda_n\end{pmatrix}.$$

**例 2．5** 设$A=\begin{pmatrix}3&1\\5&-1\end{pmatrix}$，求$A^n$.

**解**　矩阵 $A$ 的特征方程为

$$|A-\lambda E|=\begin{vmatrix}3-\lambda & 1\\ 5 & -1-\lambda\end{vmatrix}=(\lambda+2)(\lambda-4)=0,$$

解得 $A$ 的特征值为 $\lambda_1=4$，$\lambda_2=-2$.

当 $\lambda_1=4$ 时，解齐次线性方程组 $(A-4E)x=0$，得基础解系 $p_1=(1,1)^T$.

当 $\lambda_2=-2$ 时，解齐次线性方程组 $(A+2E)x=0$，得基础解系 $p_2=(1,-5)^T$.

令 $P=(p_1,p_2)=\begin{pmatrix}1 & 1\\ 1 & -5\end{pmatrix}$，则 $P^{-1}=\begin{pmatrix}\frac{5}{6} & \frac{1}{6}\\ \frac{1}{6} & -\frac{1}{6}\end{pmatrix}$，且 $P^{-1}AP=\begin{pmatrix}4 & 0\\ 0 & -2\end{pmatrix}$，所以

$$A^n=P\begin{pmatrix}4^n & 0\\ 0 & (-2)^n\end{pmatrix}P^{-1}=\begin{pmatrix}1 & 1\\ 1 & -5\end{pmatrix}\begin{pmatrix}4^n & 0\\ 0 & (-2)^n\end{pmatrix}\begin{pmatrix}\frac{5}{6} & \frac{1}{6}\\ \frac{1}{6} & -\frac{1}{6}\end{pmatrix}$$

$$=\frac{1}{6}\begin{pmatrix}5\times 4^n+(-2)^n & 4^n-(-2)^n\\ 5\times 4^n-5\times(-2)^n & 4^n+5\times(-2)^n\end{pmatrix}.$$