第 2 节 矩阵的相似对角化

安徽财经大学

统计与应用数学学院

目录

矩阵的相似

在习题 1.3 的第 13 题中，我们已经知道，如果已知可逆矩阵 $\boldsymbol{P}$，且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ (对角矩阵) $=\begin{pmatrix}\lambda_1 & \\ & \lambda_2\end{pmatrix}$，则 $\boldsymbol{A}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，且

$$\boldsymbol{A}^k=\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}=\boldsymbol{P}\begin{pmatrix}\lambda_1^k & \\ & \lambda_2^k\end{pmatrix}\boldsymbol{P}^{-1}.$$

对于同阶方阵 $\boldsymbol{A},\boldsymbol{B}$，如果存在可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$，那么对于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之间的这种关系，我们给出如下定义：

定义 (4.2.1)

对于 n 阶矩阵 $\boldsymbol{A},\boldsymbol{B}$，若存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，记为 $\boldsymbol{A}\sim\boldsymbol{B}$.

矩阵的相似

在习题 1.3 的第 13 题中，我们已经知道，如果已知可逆矩阵 $\boldsymbol{P}$，且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$ (对角矩阵) $= \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$，则 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，且

$$\boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix} \lambda_1^k & \\ & \lambda_2^k \end{pmatrix}\boldsymbol{P}^{-1}.$$

对于同阶方阵 $\boldsymbol{A}, \boldsymbol{B}$，如果存在可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$，那么对于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之间的这种关系，我们给出如下定义:

定义 (4.2.1)

对于 n 阶矩阵 $\boldsymbol{A}, \boldsymbol{B}$，若存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，记为 $\boldsymbol{A} \sim \boldsymbol{B}$.

矩阵的相似

在习题 1.3 的第 13 题中，我们已经知道，如果已知可逆矩阵 $\boldsymbol{P}$，且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$ (对角矩阵) $= \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}$，则 $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$，且

$$\boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix} \lambda_1^k & \\ & \lambda_2^k \end{pmatrix}\boldsymbol{P}^{-1}.$$

对于同阶方阵 $\boldsymbol{A}, \boldsymbol{B}$，如果存在可逆矩阵 $\boldsymbol{P}$，使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$，那么对于 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 之间的这种关系，我们给出如下定义:

定义 (4.2.1)

对于 n 阶矩阵 $\boldsymbol{A}, \boldsymbol{B}$，若存在可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$，则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相似**，记为 $\boldsymbol{A} \sim \boldsymbol{B}$.

矩阵之间的相似关系具有以下性质:

1° **反身性**: $\boldsymbol{A} \sim \boldsymbol{A}$;

2° **对称性**: 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $\boldsymbol{B} \sim \boldsymbol{A}$;

3° **传递性**: 若 $\boldsymbol{A} \sim \boldsymbol{B}$ 且 $\boldsymbol{B} \sim \boldsymbol{C}$, 则 $\boldsymbol{A} \sim \boldsymbol{C}$.

1° 和 2° 的证明是很显然的. 3° 的证明如下:

证明.

设 $\boldsymbol{A} \sim \boldsymbol{B}$, 则存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$, 又 $\boldsymbol{B} \sim \boldsymbol{C}$, 则存在可逆矩阵 $\boldsymbol{Q}$, 使 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} = \boldsymbol{C}$. 所以

$$\boldsymbol{Q}^{-1}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Q} = (\boldsymbol{P}\boldsymbol{Q})^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{Q}) = \boldsymbol{C},$$

记 $\boldsymbol{R} = \boldsymbol{P}\boldsymbol{Q}$, 则 $\boldsymbol{R}$ 可逆, 且 $\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R} = \boldsymbol{C}$. 故 $\boldsymbol{A} \sim \boldsymbol{C}$. □

矩阵之间的相似关系具有以下性质:

1° **反身性**: $A \sim A$;

2° **对称性**: 若 $A \sim B$, 则 $B \sim A$;

3° **传递性**: 若 $A \sim B$ 且 $B \sim C$, 则 $A \sim C$.

1° 和 2° 的证明是很显然的. 3° 的证明如下:

证明.

设 $A \sim B$, 则存在可逆矩阵 P, 使得 $P^{-1}AP = B$, 又 $B \sim C$, 则存在可逆矩阵 Q, 使 $Q^{-1}BQ = C$. 所以

$$Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ) = C,$$

记 $R = PQ$, 则 R 可逆, 且 $R^{-1}AR = C$. 故 $A \sim C$. $\square$

矩阵之间的相似关系具有以下性质:

1° **反身性**: $\boldsymbol{A} \sim \boldsymbol{A}$;

2° **对称性**: 若 $\boldsymbol{A} \sim \boldsymbol{B}$, 则 $\boldsymbol{B} \sim \boldsymbol{A}$;

3° **传递性**: 若 $\boldsymbol{A} \sim \boldsymbol{B}$ 且 $\boldsymbol{B} \sim \boldsymbol{C}$, 则 $\boldsymbol{A} \sim \boldsymbol{C}$.

1° 和 2° 的证明是很显然的. 3° 的证明如下:

证明.

设 $\boldsymbol{A} \sim \boldsymbol{B}$, 则存在可逆矩阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$, 又 $\boldsymbol{B} \sim \boldsymbol{C}$, 则存在可逆矩阵 $\boldsymbol{Q}$, 使 $\boldsymbol{Q}^{-1}\boldsymbol{B}\boldsymbol{Q} = \boldsymbol{C}$. 所以

$$\boldsymbol{Q}^{-1}\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{Q} = (\boldsymbol{P}\boldsymbol{Q})^{-1}\boldsymbol{A}(\boldsymbol{P}\boldsymbol{Q}) = \boldsymbol{C},$$

记 $\boldsymbol{R} = \boldsymbol{P}\boldsymbol{Q}$, 则 $\boldsymbol{R}$ 可逆, 且 $\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{R} = \boldsymbol{C}$. 故 $\boldsymbol{A} \sim \boldsymbol{C}$. □

矩阵之间的相似关系具有以下性质:

1° **反身性**: $A \sim A$;

2° **对称性**: 若 $A \sim B$, 则 $B \sim A$;

3° **传递性**: 若 $A \sim B$ 且 $B \sim C$, 则 $A \sim C$.

1° 和 2° 的证明是很显然的. 3° 的证明如下:

证明.

设 $A \sim B$, 则存在可逆矩阵 P, 使得 $P^{-1}AP = B$, 又 $B \sim C$, 则存在可逆矩阵 Q, 使 $Q^{-1}BQ = C$. 所以

$$Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ) = C,$$

记 $R = PQ$, 则 R 可逆, 且 $R^{-1}AR = C$. 故 $A \sim C$. □

定理 (4.2.1)

相似矩阵的特征值相同.

证明.

设 $A \sim B$, 则存在可逆矩阵 P, 使

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

$$\begin{aligned}\det(\lambda I - B) &= \det\left(\lambda I - P^{-1}AP\right) = \det\left[P^{-1}(\lambda I - A)P\right] \\ &= \det P^{-1}\det(\lambda I - A)\det P = \det(\lambda I - A),\end{aligned}$$

A 与 B 的特征多项式相同, 因此 A 与 B 的特征值相同. □

定理 (4.2.1)

相似矩阵的特征值相同.

证明.

设 $A \sim B$, 则存在可逆矩阵 P, 使

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

$$\begin{aligned}\det(\lambda I - B) &= \det\left(\lambda I - P^{-1}AP\right) = \det\left[P^{-1}(\lambda I - A)P\right] \\ &= \det P^{-1}\det(\lambda I - A)\det P = \det(\lambda I - A),\end{aligned}$$

A 与 B 的特征多项式相同, 因此 A 与 B 的特征值相同. □

定理 (4.2.1)

相似矩阵的特征值相同.

证明.

设 $\boldsymbol{A} \sim \boldsymbol{B}$, 则存在可逆矩阵 $\boldsymbol{P}$, 使

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

$$\begin{aligned}\det(\lambda\boldsymbol{I} - \boldsymbol{B}) &= \det\left(\lambda\boldsymbol{I} - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right) = \det\left[\boldsymbol{P}^{-1}(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{P}\right] \\ &= \det \boldsymbol{P}^{-1} \det(\lambda\boldsymbol{I} - \boldsymbol{A}) \det \boldsymbol{P} = \det(\lambda\boldsymbol{I} - \boldsymbol{A}),\end{aligned}$$

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征值相同. □

定理 (4.2.1)

相似矩阵的特征值相同.

证明.

设 $\boldsymbol{A} \sim \boldsymbol{B}$, 则存在可逆矩阵 $\boldsymbol{P}$, 使

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

$$\begin{aligned}\det(\lambda\boldsymbol{I}-\boldsymbol{B}) &= \det\left(\lambda\boldsymbol{I}-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right) = \det\left[\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}\right] \\ &= \det\boldsymbol{P}^{-1}\det(\lambda\boldsymbol{I}-\boldsymbol{A})\det\boldsymbol{P} = \det(\lambda\boldsymbol{I}-\boldsymbol{A}),\end{aligned}$$

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征值相同. □

定理 (4.2.1)

相似矩阵的特征值相同.

证明.

设 $\boldsymbol{A} \sim \boldsymbol{B}$, 则存在可逆矩阵 $\boldsymbol{P}$, 使

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

$$\begin{aligned}\det(\lambda\boldsymbol{I} - \boldsymbol{B}) &= \det\left(\lambda\boldsymbol{I} - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}\right) = \det\left[\boldsymbol{P}^{-1}(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{P}\right] \\ &= \det\boldsymbol{P}^{-1}\det(\lambda\boldsymbol{I} - \boldsymbol{A})\det\boldsymbol{P} = \det(\lambda\boldsymbol{I} - \boldsymbol{A}),\end{aligned}$$

$\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征多项式相同, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的特征值相同. □

相似矩阵有以下性质

定理（相似矩阵的性质）

若矩阵 A 与 B 相似，则

(1) A^k 与 B^k 相似，其中 k 为正整数.

(2) A 与 B 的行列式相等.

(3) A 与 B 有相同的可逆性，当 A，B 可逆时，A^{-1} 与 B^{-1} 也相似.

(4) A 与 B 有相同的特征多项式和特征值.

(5) A 与 B 的秩相等.

(6) A 与 B 的迹相同.

证明.

由 $A \sim B$ 可知, 存在可逆矩阵 P, 使得 $P^{-1}AP = B$.

(1) 由于

$$
\begin{aligned}
B^k &= (P^{-1}AP)^k \\
&= \underbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)(P^{-1}AP)}_{k\text{个}} \\
&= P^{-1}A(PP^{-1})A(PP^{-1})\cdots(PP^{-1})A(PP^{-1})AP \\
&= P^{-1}A^kP,
\end{aligned}
$$

故 $A^k \sim B^k$.

(2) $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A|$, 即 A 与 B 的行列式相等. □

证明.

由 $A \sim B$ 可知, 存在可逆矩阵 P, 使得 $P^{-1}AP = B$.

(1) 由于

$$\begin{aligned} B^k &= (P^{-1}AP)^k \\ &= \underbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)(P^{-1}AP)}_{k\text{个}} \\ &= P^{-1}A(PP^{-1})A(PP^{-1})\cdots(PP^{-1})A(PP^{-1})AP \\ &= P^{-1}A^kP, \end{aligned}$$

故 $A^k \sim B^k$.

(2) $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A|$, 即 A 与 B 的行列式相等. □

证明.

由 $A \sim B$ 可知, 存在可逆矩阵 P, 使得 $P^{-1}AP = B$.
(1) 由于

$$\begin{aligned}
B^k &= (P^{-1}AP)^k \\
&= \underbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)(P^{-1}AP)}_{k\text{个}} \\
&= P^{-1}A(PP^{-1})A(PP^{-1})\cdots(PP^{-1})A(PP^{-1})AP \\
&= P^{-1}A^kP,
\end{aligned}$$

故 $A^k \sim B^k$.
(2) $|B| = |P^{-1}AP| = |P^{-1}||A||P| = |A|$, 即 A 与 B 的行列式相等. □

证明.

(3) 由 (2) 可知 $|\boldsymbol{A}|=|\boldsymbol{B}|$, 所以 $|\boldsymbol{A}|$ 与 $|\boldsymbol{B}|$ 同时为零或不为零, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同时可逆或不可逆, 即有相同的可逆性.
当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可逆时, 由于 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 则

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

即 $\boldsymbol{A}^{-1}\sim\boldsymbol{B}^{-1}$.

(4) 由于

$$\begin{aligned}|\lambda\boldsymbol{I}-\boldsymbol{B}|&=|\lambda\boldsymbol{I}-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}|=|\boldsymbol{P}^{-1}(\lambda\boldsymbol{I})\boldsymbol{P}-\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}|\\&=|\boldsymbol{P}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{P}|=|\boldsymbol{P}^{-1}|\cdot|\lambda\boldsymbol{I}-\boldsymbol{A}|\cdot|\boldsymbol{P}|\\&=|\lambda\boldsymbol{I}-\boldsymbol{A}|.\end{aligned}$$

所以, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征多项式, 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值.
(5) 由于 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价, 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的秩相等.
(6) 由 (4) 可推出 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的迹相同, 即 $\mathrm{tr}(\boldsymbol{A})=\mathrm{tr}(\boldsymbol{B})$.

证明.

(3) 由 (2) 可知 $|A| = |B|$, 所以 $|A|$ 与 $|B|$ 同时为零或不为零, 因此 A 与 B 同时可逆或不可逆, 即有相同的可逆性.
当 A 与 B 可逆时, 由于 $P^{-1}AP = B$, 则

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

即 $A^{-1} \sim B^{-1}$.
(4) 由于

$$\begin{aligned}|\lambda I - B| &= |\lambda I - P^{-1}AP| = |P^{-1}(\lambda I)P - P^{-1}AP| \\ &= |P^{-1}(\lambda I - A)P| = |P^{-1}| \cdot |\lambda I - A| \cdot |P| \\ &= |\lambda I - A|.\end{aligned}$$

所以, A 与 B 有相同的特征多项式, 从而 A 与 B 有相同的特征值.
(5) 由于 $P^{-1}AP = B$, 则 A 与 B 等价, 从而 A 与 B 的秩相等.
(6) 由 (4) 可推出 A 与 B 的迹相同, 即 $\mathrm{tr}(A) = \mathrm{tr}(B)$.

证明.

(3) 由 (2) 可知 $|\boldsymbol{A}| = |\boldsymbol{B}|$, 所以 $|\boldsymbol{A}|$ 与 $|\boldsymbol{B}|$ 同时为零或不为零, 因此 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 同时可逆或不可逆, 即有相同的可逆性.
当 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 可逆时, 由于 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$, 则

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

即 $\boldsymbol{A}^{-1} \sim \boldsymbol{B}^{-1}$.
(4) 由于

$$\begin{aligned}|\lambda\boldsymbol{I} - \boldsymbol{B}| &= |\lambda\boldsymbol{I} - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}| = |\boldsymbol{P}^{-1}(\lambda\boldsymbol{I})\boldsymbol{P} - \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}| \\ &= |\boldsymbol{P}^{-1}(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{P}| = |\boldsymbol{P}^{-1}| \cdot |\lambda\boldsymbol{I} - \boldsymbol{A}| \cdot |\boldsymbol{P}| \\ &= |\lambda\boldsymbol{I} - \boldsymbol{A}|.\end{aligned}$$

所以, $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征多项式, 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征值.
(5) 由于 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{B}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价, 从而 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的秩相等.
(6) 由 (4) 可推出 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的迹相同, 即 $\mathrm{tr}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{B})$.

证明.

(3) 由 (2) 可知 $|A| = |B|$, 所以 $|A|$ 与 $|B|$ 同时为零或不为零, 因此 A 与 B 同时可逆或不可逆, 即有相同的可逆性.
当 A 与 B 可逆时, 由于 $P^{-1}AP = B$, 则

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

即 $A^{-1} \sim B^{-1}$.
(4) 由于

$$\begin{aligned}|\lambda I - B| &= |\lambda I - P^{-1}AP| = |P^{-1}(\lambda I)P - P^{-1}AP| \\ &= |P^{-1}(\lambda I - A)P| = |P^{-1}| \cdot |\lambda I - A| \cdot |P| \\ &= |\lambda I - A|.\end{aligned}$$

所以, A 与 B 有相同的特征多项式, 从而 A 与 B 有相同的特征值.
(5) 由于 $P^{-1}AP = B$, 则 A 与 B 等价, 从而 A 与 B 的秩相等.
(6) 由 (4) 可推出 A 与 B 的迹相同, 即 $\mathrm{tr}(A) = \mathrm{tr}(B)$. □

例 (4.2.1)

设 n 阶方阵 $\boldsymbol{A} \sim \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$, 求 $\boldsymbol{A}^k$ (k 为正整数).

解 因为 $\boldsymbol{A} \sim \boldsymbol{\Lambda}$, 所以存在可逆方阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$, $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$, 故

$$\boldsymbol{A}^k = \left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right)\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right)\cdots\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right) = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}$$

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

只需求出 $\boldsymbol{P}^{-1}$, 再计算出 $\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}$ 就行了. 当 k 比较大时, 这比直接计算 $\boldsymbol{A}^k$ 要方便得多.

例 (4.2.1)

设 n 阶方阵 $\boldsymbol{A} \sim \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$, 求 $\boldsymbol{A}^k$ (k 为正整数).

解 因为 $\boldsymbol{A} \sim \boldsymbol{\Lambda}$, 所以存在可逆方阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$, $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$, 故

$$\begin{aligned} \boldsymbol{A}^k &= \left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right)\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right)\cdots\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right) = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} \\ &= \boldsymbol{P}\begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix}\boldsymbol{P}^{-1}, \end{aligned}$$

只需求出 $\boldsymbol{P}^{-1}$, 再计算出 $\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}$ 就行了. 当 k 比较大时, 这比直接计算 $\boldsymbol{A}^k$ 要方便得多.

例 (4.2.1)

设 n 阶方阵 $\boldsymbol{A} \sim \boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$, 求 $\boldsymbol{A}^k$ (k 为正整数).

解　因为 $\boldsymbol{A} \sim \boldsymbol{\Lambda}$, 所以存在可逆方阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$, $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}$, 故

$$\begin{aligned} \boldsymbol{A}^k &= \left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right)\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right)\cdots\left(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\right) = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} \\ &= \boldsymbol{P}\begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix}\boldsymbol{P}^{-1}, \end{aligned}$$

只需求出 $\boldsymbol{P}^{-1}$, 再计算出 $\boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1}$ 就行了. 当 k 比较大时, 这比直接计算 $\boldsymbol{A}^k$ 要方便得多.

我们自然要提出的问题是:

- 什么样的矩阵 $\boldsymbol{A}$ 可以与对角矩阵相似?
- 或者说, 对于给定的矩阵 $\boldsymbol{A}$, 在什么条件下存在对角矩阵 $\boldsymbol{\Lambda}$ 与可逆矩阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$?
- 如果这样的矩阵 $\boldsymbol{\Lambda}$ 与 $\boldsymbol{P}$ 存在, 那么又应该怎样求出?

下面就讨论这些问题.

定理 (4.2.2)

若 n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$ 相似，则 $\lambda_1,\lambda_2,\cdots\ \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征值.

证明.

因为 $\boldsymbol{A}\sim\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$，所以 $\boldsymbol{A}$ 与 $\boldsymbol{\Lambda}$ 的特征值相同，又

$$\det(\lambda\boldsymbol{I}-\boldsymbol{\Lambda})=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

所以 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{\Lambda}$ 的全部特征值，也就是 $\boldsymbol{A}$ 的全部特征值. □

定理 (4.2.2)

若 n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$ 相似，则 $\lambda_1,\lambda_2,\cdots\ \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征值.

证明.

因为 $\boldsymbol{A}\sim\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$，所以 $\boldsymbol{A}$ 与 $\boldsymbol{\Lambda}$ 的特征值相同，又

$$\det(\lambda\boldsymbol{I}-\boldsymbol{\Lambda})=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

所以 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{\Lambda}$ 的全部特征值，也就是 $\boldsymbol{A}$ 的全部特征值. □

定理 (4.2.2)

若 n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$ 相似，则 $\lambda_1,\lambda_2,\cdots\ \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征值.

证明.

因为 $\boldsymbol{A}\sim\boldsymbol{\Lambda}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$，所以 $\boldsymbol{A}$ 与 $\boldsymbol{\Lambda}$ 的特征值相同，又

$$\det(\lambda\boldsymbol{I}-\boldsymbol{\Lambda})=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n),$$

所以 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 是 $\boldsymbol{\Lambda}$ 的全部特征值，也就是 $\boldsymbol{A}$ 的全部特征值. □

定理 2 指出, 若 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似, 则 $\boldsymbol{\Lambda}$ 的主对角线上的元就是 $\boldsymbol{A}$ 的全部特征值, 那么, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 的矩阵 $\boldsymbol{P}$ 又是怎样构成的呢?

设 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$, $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{P}$ 的列向量组, 则

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

$$\text{即 } \boldsymbol{A}(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

$$(\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\cdots,\boldsymbol{A}\boldsymbol{p}_n)=(\lambda_1\boldsymbol{p}_1,\lambda_2\boldsymbol{p}_2,\cdots,\lambda_n\boldsymbol{p}_n),$$

$$\boldsymbol{A}\boldsymbol{p}_i=\lambda_i\boldsymbol{p}_i\quad(i=1,2,\cdots,n).$$

因为 $\boldsymbol{P}$ 可逆, 所以 $\boldsymbol{p}_i\neq\boldsymbol{0}$ $(i=1,2,\cdots,n)$, 于是, $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量.

定理 2 指出，若 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似，则 $\boldsymbol{\Lambda}$ 的主对角线上的元就是 $\boldsymbol{A}$ 的全部特征值，那么，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 的矩阵 $\boldsymbol{P}$ 又是怎样构成的呢？
设 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{P}$ 的列向量组，则

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

即 $\boldsymbol{A}(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix}$，

$$(\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\cdots,\boldsymbol{A}\boldsymbol{p}_n)=(\lambda_1\boldsymbol{p}_1,\lambda_2\boldsymbol{p}_2,\cdots,\lambda_n\boldsymbol{p}_n),$$

$$\boldsymbol{A}\boldsymbol{p}_i=\lambda_i\boldsymbol{p}_i\quad(i=1,2,\cdots,n).$$

因为 $\boldsymbol{P}$ 可逆，所以 $\boldsymbol{p}_i\neq\boldsymbol{0}\ (i=1,2,\cdots,n)$，于是，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量.

定理 2 指出，若 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似，则 $\boldsymbol{\Lambda}$ 的主对角线上的元就是 $\boldsymbol{A}$ 的全部特征值，那么，使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{\Lambda}$ 的矩阵 $\boldsymbol{P}$ 又是怎样构成的呢？
设 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{P}$ 的列向量组，则

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

即 $$\boldsymbol{A}(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

$$(\boldsymbol{A}\boldsymbol{p}_1,\boldsymbol{A}\boldsymbol{p}_2,\cdots,\boldsymbol{A}\boldsymbol{p}_n)=(\lambda_1\boldsymbol{p}_1,\lambda_2\boldsymbol{p}_2,\cdots,\lambda_n\boldsymbol{p}_n),$$

$$\boldsymbol{A}\boldsymbol{p}_i=\lambda_i\boldsymbol{p}_i\quad(i=1,2,\cdots,n).$$

因为 $\boldsymbol{P}$ 可逆，所以 $\boldsymbol{p}_i\neq\boldsymbol{0}\ (i=1,2,\cdots,n)$，于是，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量.

定理 2 指出，若 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似，则 $\boldsymbol{\Lambda}$ 的主对角线上的元就是 $\boldsymbol{A}$ 的全部特征值，那么，使 $\boldsymbol{P}^{-1}\boldsymbol{AP}=\boldsymbol{\Lambda}$ 的矩阵 $\boldsymbol{P}$ 又是怎样构成的呢？

设 $\boldsymbol{P}=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)$，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{P}$ 的列向量组，则

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

即 $$\boldsymbol{A}(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)=(\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n)\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{pmatrix},$$

$$(\boldsymbol{Ap}_1,\boldsymbol{Ap}_2,\cdots,\boldsymbol{Ap}_n)=(\lambda_1\boldsymbol{p}_1,\lambda_2\boldsymbol{p}_2,\cdots,\lambda_n\boldsymbol{p}_n),$$

$$\boldsymbol{Ap}_i=\lambda_i\boldsymbol{p}_i\quad(i=1,2,\cdots,n).$$

因为 $\boldsymbol{P}$ 可逆，所以 $\boldsymbol{p}_i\neq\boldsymbol{0}\ (i=1,2,\cdots,n)$，于是，$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是 $\boldsymbol{A}$ 的 n 个线性无关的特征向量.

反之，若 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$，即 $\boldsymbol{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i (i = 1, 2, \cdots, n)$，设 $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$，则 $\boldsymbol{P}$ 可逆，且

$$\boldsymbol{A}\boldsymbol{P} = (\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \cdots, \boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1\boldsymbol{p}_1, \lambda_2\boldsymbol{p}_2, \cdots, \lambda_n\boldsymbol{p}_n)$$

$$= (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = \boldsymbol{P}\boldsymbol{\Lambda},$$

所以 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$，即 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似.

由以上讨论可得

定理 (4.2.3)

n 阶矩阵 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda}$ 相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

反之，若 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$，即 $\boldsymbol{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i (i = 1, 2, \cdots, n)$，设 $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$，则 $\boldsymbol{P}$ 可逆，且

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{P} &= (\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \cdots, \boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1 \boldsymbol{p}_1, \lambda_2 \boldsymbol{p}_2, \cdots, \lambda_n \boldsymbol{p}_n) \\
&= (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n) \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = \boldsymbol{P}\boldsymbol{\Lambda},
\end{aligned}
$$

所以 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$，即 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似.
由以上讨论可得

定理 (4.2.3)

n 阶矩阵 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda}$ 相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

反之，若 $\boldsymbol{A}$ 有 n 个线性无关的特征向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$，即 $\boldsymbol{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i (i = 1, 2, \cdots, n)$，设 $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$，则 $\boldsymbol{P}$ 可逆，且

$$\begin{aligned}\boldsymbol{A}\boldsymbol{P} &= (\boldsymbol{A}\boldsymbol{p}_1, \boldsymbol{A}\boldsymbol{p}_2, \cdots, \boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1 \boldsymbol{p}_1, \lambda_2 \boldsymbol{p}_2, \cdots, \lambda_n \boldsymbol{p}_n) \\ &= (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = \boldsymbol{P}\boldsymbol{\Lambda},\end{aligned}$$

所以 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$，即 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似.

由以上讨论可得

定理 (4.2.3)

n 阶矩阵 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda}$ 相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

由此定理可知, 若 n 阶矩阵 $\boldsymbol{A}$ 有 n 个线性无关的特征向量

$$\boldsymbol{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i \quad (i = 1, 2, \cdots, n),$$

令 $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$, 则

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

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征值.

注: $\boldsymbol{P}$ 中列向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$ 的顺序要与 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的顺序一致. 由于 $\boldsymbol{p}_i$ 是 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$ 的基础解系中的解向量, 故 $\boldsymbol{p}_i$ 的取法不是惟一的, 因此 $\boldsymbol{P}$ 也不是惟一的. 而 $f_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ 的根只有 n 个 (重根按重数计算), 所以, 若不计 λ_i 的排列顺序, 则 Λ 是惟一确定

由此定理可知, 若 n 阶矩阵 $\boldsymbol{A}$ 有 n 个线性无关的特征向量

$$\boldsymbol{A}\boldsymbol{p}_i = \lambda_i \boldsymbol{p}_i \quad (i = 1, 2, \cdots, n),$$

令 $\boldsymbol{P} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$, 则

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

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征值.

注: $\boldsymbol{P}$ 中列向量 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$ 的顺序要与 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的顺序一致. 由于 $\boldsymbol{p}_i$ 是 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$ 的基础解系中的解向量, 故 $\boldsymbol{p}_i$ 的取法不是惟一的, 因此 $\boldsymbol{P}$ 也不是惟一的. 而 $f_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = 0$ 的根只有 n 个 (重根按重数计算), 所以, 若不计 λ_i 的排列顺序, 则 $\boldsymbol{\Lambda}$ 是惟一确定的.

例 (4.2.2)

设 $\boldsymbol{A}=\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$, 求 $\boldsymbol{A}^{10}$.

解

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix} \lambda-4 & -6 & 0 \\ 3 & \lambda+5 & 0 \\ 3 & 6 & \lambda-1 \end{vmatrix}=(\lambda+2)(\lambda-1)^2,$$

$\boldsymbol{A}$ 的特征值为 $\lambda_1=-2, \lambda_2=1(2$ 重$)$.

例 (4.2.2)

设 $A = \begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$, 求 A^{10}.

解

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 4 & -6 & 0 \\ 3 & \lambda + 5 & 0 \\ 3 & 6 & \lambda - 1 \end{vmatrix} = (\lambda + 2)(\lambda - 1)^2,$$

A 的特征值为 $\lambda_1 = -2, \lambda_2 = 1$(2 重).

解

对于 $\lambda_1=-2$, $(\lambda_1\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的系数矩阵为

$$\begin{pmatrix} -6 & -6 & 0 \\ 3 & 3 & 0 \\ 3 & 6 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},$$

对应的齐次线性方程组为 $\begin{cases} x_1=-x_3, \\ x_2=x_3, \end{cases}$ 其基础解系为 $\boldsymbol{\alpha}_1=(-1,1,1)^{\mathrm{T}}$.

对于 $\lambda_2=1$, $(\lambda_2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的系数矩阵为

$$\begin{pmatrix} -3 & -6 & 0 \\ 3 & 6 & 0 \\ 3 & 6 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

对应的齐次线性方程组为 $x_1=-2x_2+0x_3$, 其基础解系为

$$\boldsymbol{\alpha}_2=\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{\alpha}_3=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

解

对于 $\lambda_1=-2$, $(\lambda_1\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的系数矩阵为

$$\begin{pmatrix}-6&-6&0\\3&3&0\\3&6&-3\end{pmatrix}\to\begin{pmatrix}1&1&0\\0&0&0\\0&1&-1\end{pmatrix}\to\begin{pmatrix}1&0&1\\0&1&-1\\0&0&0\end{pmatrix},$$

对应的齐次线性方程组为 $\begin{cases}x_1=-x_3,\\x_2=x_3,\end{cases}$ 其基础解系为 $\boldsymbol{\alpha}_1=(-1,1,1)^{\mathrm{T}}$.

对于 $\lambda_2=1$, $(\lambda_2\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的系数矩阵为

$$\begin{pmatrix}-3&-6&0\\3&6&0\\3&6&0\end{pmatrix}\to\begin{pmatrix}1&2&0\\0&0&0\\0&0&0\end{pmatrix},$$

对应的齐次线性方程组为 $x_1=-2x_2+0x_3$, 其基础解系为

$$\boldsymbol{\alpha}_2=\begin{pmatrix}-2\\1\\0\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

解

令

$$\boldsymbol{P} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3) = \begin{pmatrix} -1 & -2 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

易见, $\boldsymbol{P}$ 是可逆矩阵, 且

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} -2 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \quad \boldsymbol{A} = \boldsymbol{P}\begin{pmatrix} -2 & & \\ & 1 & \\ & & 1 \end{pmatrix}\boldsymbol{P}^{-1},$$

$$\boldsymbol{P}^{-1} = \begin{pmatrix} 1 & 2 & 0 \\ -1 & -1 & 0 \\ -1 & -2 & 1 \end{pmatrix},$$

解

所以

$$\boldsymbol{A}^{10}=\boldsymbol{P}\begin{pmatrix}(-2)^{10} & & \\ & 1 & \\ & & 1\end{pmatrix}\boldsymbol{P}^{-1}$$

$$=\begin{pmatrix}-1 & -2 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}\begin{pmatrix}1024 & & \\ & 1 & \\ & & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 0\\ -1 & -1 & 0\\ -1 & -2 & 1\end{pmatrix}$$

$$=\begin{pmatrix}-1022 & -2046 & 0\\ 1023 & 2047 & 0\\ 1023 & 2046 & 1\end{pmatrix}.$$

例 (4.2.3)

设 $A \neq O, A^k = O$ (k 为正整数), 证明 A 不能与对角矩阵相似.

证明.

设 A 能与对角矩阵 $\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则存在可逆矩阵 P, 使

$$A = P\Lambda P^{-1}, \tag{4.1}$$

$$A^k = P\Lambda^k P^{-1} = P\begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix} P^{-1} = O,$$

于是可得 $\operatorname{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k) = O$. 从而, $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, 即有 $\Lambda = O$. 由式 (4.1) 可得 $A = O$, 与题设 $A \neq O$ 矛盾. 故 A 不能与对角矩阵相似. □

例 (4.2.3)

设 $\boldsymbol{A} \neq \boldsymbol{O}, \boldsymbol{A}^k = \boldsymbol{O}$ (k 为正整数), 证明 $\boldsymbol{A}$ 不能与对角矩阵相似.

证明.

设 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则存在可逆矩阵 $\boldsymbol{P}$, 使

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}, \tag{4.1}$$

$$\boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix}\boldsymbol{P}^{-1} = \boldsymbol{O},$$

于是可得 $\operatorname{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k) = \boldsymbol{O}$. 从而, $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, 即有 $\boldsymbol{\Lambda} = \boldsymbol{O}$. 由式 (4.1) 可得 $\boldsymbol{A} = \boldsymbol{O}$, 与题设 $\boldsymbol{A} \neq \boldsymbol{O}$ 矛盾. 故 $\boldsymbol{A}$ 不能与对角矩阵相似. □

例 (4.2.3)

设 $\boldsymbol{A} \neq \boldsymbol{O}, \boldsymbol{A}^k = \boldsymbol{O}$ (k 为正整数), 证明 $\boldsymbol{A}$ 不能与对角矩阵相似.

证明.

设 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则存在可逆矩阵 $\boldsymbol{P}$, 使

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}, \tag{4.1}$$

$$\boldsymbol{A}^k = \boldsymbol{P}\boldsymbol{\Lambda}^k\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix}\boldsymbol{P}^{-1} = \boldsymbol{O},$$

于是可得 $\operatorname{diag}(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k) = \boldsymbol{O}$. 从而, $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, 即有 $\boldsymbol{\Lambda} = \boldsymbol{O}$. 由式 (4.1) 可得 $\boldsymbol{A} = \boldsymbol{O}$, 与题设 $\boldsymbol{A} \neq \boldsymbol{O}$ 矛盾. 故 $\boldsymbol{A}$ 不能与对角矩阵相似.

例 (4.2.3)

设 $\boldsymbol{A} \neq \boldsymbol{O}, \boldsymbol{A}^k = \boldsymbol{O}$ (k 为正整数), 证明 $\boldsymbol{A}$ 不能与对角矩阵相似.

证明.

设 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则存在可逆矩阵 $\boldsymbol{P}$, 使

$$\boldsymbol{A} = \boldsymbol{P\Lambda P}^{-1}, \tag{4.1}$$

$$\boldsymbol{A}^k = \boldsymbol{P\Lambda}^k\boldsymbol{P}^{-1} = \boldsymbol{P}\begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & \\ & & \ddots & \\ & & & \lambda_n^k \end{pmatrix}\boldsymbol{P}^{-1} = \boldsymbol{O},$$

于是可得 $\operatorname{diag}\left(\lambda_1^k, \lambda_2^k, \cdots, \lambda_n^k\right) = \boldsymbol{O}$. 从而, $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, 即有 $\boldsymbol{\Lambda} = \boldsymbol{O}$. 由式 (4.1) 可得 $\boldsymbol{A} = \boldsymbol{O}$, 与题设 $\boldsymbol{A} \neq \boldsymbol{O}$ 矛盾. 故 $\boldsymbol{A}$ 不能与对角矩阵相似. □

定理 3 给出了 n 阶矩阵与对角矩阵相似的充要条件, 但是对于一个具体的 n 阶矩阵, 要直接判断它是否有 n 个线性无关的特征向量一般是很困难的, 下面我们进一步讨论什么样的 n 阶矩阵能与对角矩阵相似.

定理 (4.2.4)

设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是矩阵 $\boldsymbol{A}$ 的互异特征值, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 $\boldsymbol{A}$ 分别对应于这些特征值的特征向量, 则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关.

证明.

用数学归纳法证明.

当 $m=1$ 时, 结论显然成立. 因为特征向量 $\boldsymbol{\alpha}_1 \neq \mathbf{0}$, 所以一个非零向量是线性无关的.

假设对 $m-1$ 个互异特征值结论成立. □

定理 3 给出了 n 阶矩阵与对角矩阵相似的充要条件，但是对于一个具体的 n 阶矩阵，要直接判断它是否有 n 个线性无关的特征向量一般是很困难的，下面我们进一步讨论什么样的 n 阶矩阵能与对角矩阵相似.

定理 (4.2.4)

设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是矩阵 $\boldsymbol{A}$ 的互异特征值，$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 $\boldsymbol{A}$ 分别对应于这些特征值的特征向量，则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关.

证明.

用数学归纳法证明.
当 $m = 1$ 时，结论显然成立. 因为特征向量 $\boldsymbol{\alpha}_1 \neq \mathbf{0}$，所以一个非零向量是线性无关的.
假设对 $m-1$ 个互异特征值结论成立. □

证明.

对 m 个互异特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 以及它们所对应的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$，设

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}, \tag{4.2}$$

用 $\boldsymbol{A}$ 左乘式 (4.2) 两端得

$$k_1(\boldsymbol{A}\boldsymbol{\alpha}_1) + k_2(\boldsymbol{A}\boldsymbol{\alpha}_2) + \cdots + k_m(\boldsymbol{A}\boldsymbol{\alpha}_m) = \mathbf{0},$$

$$k_1(\lambda_1\boldsymbol{\alpha}_1) + k_2(\lambda_2\boldsymbol{\alpha}_2) + \cdots + k_m(\lambda_m\boldsymbol{\alpha}_m) = \mathbf{0}. \tag{4.3}$$

用 λ_m 乘式 (4.2) 两端得

$$k_1(\lambda_m\boldsymbol{\alpha}_1) + k_2(\lambda_m\boldsymbol{\alpha}_2) + \cdots + k_m(\lambda_m\boldsymbol{\alpha}_m) = \mathbf{0}, \tag{4.4}$$

式 (4.3) 与式 (4.4) 两端相减可得

$$k_1(\lambda_1 - \lambda_m)\boldsymbol{\alpha}_1 + k_2(\lambda_2 - \lambda_m)\boldsymbol{\alpha}_2 + \cdots + k_{m-1}(\lambda_{m-1} - \lambda_m)\boldsymbol{\alpha}_{m-1} = \mathbf{0},$$

证明.

对 m 个互异特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 以及它们所对应的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$, 设

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}, \tag{4.2}$$

用 $\boldsymbol{A}$ 左乘式 (4.2) 两端得

$$k_1\left(\boldsymbol{A}\boldsymbol{\alpha}_1\right) + k_2\left(\boldsymbol{A}\boldsymbol{\alpha}_2\right) + \cdots + k_m\left(\boldsymbol{A}\boldsymbol{\alpha}_m\right) = \mathbf{0},$$

$$k_1\left(\lambda_1\boldsymbol{\alpha}_1\right) + k_2\left(\lambda_2\boldsymbol{\alpha}_2\right) + \cdots + k_m\left(\lambda_m\boldsymbol{\alpha}_m\right) = \mathbf{0}. \tag{4.3}$$

用 λ_m 乘式 (4.2) 两端得

$$k_1\left(\lambda_m\boldsymbol{\alpha}_1\right) + k_2\left(\lambda_m\boldsymbol{\alpha}_2\right) + \cdots + k_m\left(\lambda_m\boldsymbol{\alpha}_m\right) = \mathbf{0}, \tag{4.4}$$

式 (4.3) 与式 (4.4) 两端相减可得

$$k_1\left(\lambda_1 - \lambda_m\right)\boldsymbol{\alpha}_1 + k_2\left(\lambda_2 - \lambda_m\right)\boldsymbol{\alpha}_2 + \cdots + k_{m-1}\left(\lambda_{m-1} - \lambda_m\right)\boldsymbol{\alpha}_{m-1} = \mathbf{0},$$

证明.

对 m 个互异特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 以及它们所对应的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$, 设

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}, \tag{4.2}$$

用 $\boldsymbol{A}$ 左乘式 (4.2) 两端得

$$k_1(\boldsymbol{A}\boldsymbol{\alpha}_1) + k_2(\boldsymbol{A}\boldsymbol{\alpha}_2) + \cdots + k_m(\boldsymbol{A}\boldsymbol{\alpha}_m) = \mathbf{0},$$

$$k_1(\lambda_1\boldsymbol{\alpha}_1) + k_2(\lambda_2\boldsymbol{\alpha}_2) + \cdots + k_m(\lambda_m\boldsymbol{\alpha}_m) = \mathbf{0}. \tag{4.3}$$

用 λ_m 乘式 (4.2) 两端得

$$k_1(\lambda_m\boldsymbol{\alpha}_1) + k_2(\lambda_m\boldsymbol{\alpha}_2) + \cdots + k_m(\lambda_m\boldsymbol{\alpha}_m) = \mathbf{0}, \tag{4.4}$$

式 (4.3) 与式 (4.4) 两端相减可得

$$k_1(\lambda_1 - \lambda_m)\boldsymbol{\alpha}_1 + k_2(\lambda_2 - \lambda_m)\boldsymbol{\alpha}_2 + \cdots + k_{m-1}(\lambda_{m-1} - \lambda_m)\boldsymbol{\alpha}_{m-1} = \mathbf{0},$$

证明.

对 m 个互异特征值 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 以及它们所对应的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$, 设

$$k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}, \tag{4.2}$$

用 $\boldsymbol{A}$ 左乘式 (4.2) 两端得

$$k_1(\boldsymbol{A}\boldsymbol{\alpha}_1) + k_2(\boldsymbol{A}\boldsymbol{\alpha}_2) + \cdots + k_m(\boldsymbol{A}\boldsymbol{\alpha}_m) = \mathbf{0},$$

$$k_1(\lambda_1\boldsymbol{\alpha}_1) + k_2(\lambda_2\boldsymbol{\alpha}_2) + \cdots + k_m(\lambda_m\boldsymbol{\alpha}_m) = \mathbf{0}. \tag{4.3}$$

用 λ_m 乘式 (4.2) 两端得

$$k_1(\lambda_m\boldsymbol{\alpha}_1) + k_2(\lambda_m\boldsymbol{\alpha}_2) + \cdots + k_m(\lambda_m\boldsymbol{\alpha}_m) = \mathbf{0}, \tag{4.4}$$

式 (4.3) 与式 (4.4) 两端相减可得

$$k_1(\lambda_1 - \lambda_m)\boldsymbol{\alpha}_1 + k_2(\lambda_2 - \lambda_m)\boldsymbol{\alpha}_2 + \cdots + k_{m-1}(\lambda_{m-1} - \lambda_m)\boldsymbol{\alpha}_{m-1} = \mathbf{0},$$

证明.

式 (4.3) 与式 (4.4) 两端相减可得

$$k_1(\lambda_1-\lambda_m)\boldsymbol{\alpha}_1+k_2(\lambda_2-\lambda_m)\boldsymbol{\alpha}_2+\cdots+k_{m-1}(\lambda_{m-1}-\lambda_m)\boldsymbol{\alpha}_{m-1}=\mathbf{0},$$

由归纳假设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 线性无关, 故

$$k_1(\lambda_1-\lambda_m)=k_2(\lambda_2-\lambda_m)=\cdots=k_{m-1}(\lambda_{m-1}-\lambda_m)=0.$$

又因为 $\lambda_i-\lambda_m\neq 0\ (i=1,2,\cdots,m-1)$, 所以只有 $k_1=k_2=\cdots=k_{m-1}=0$, 代入式 (4.2) 得 $k_m\boldsymbol{\alpha}_m=\mathbf{0}$, 由 $\boldsymbol{\alpha}_m\neq\mathbf{0}$ 可得 $k_m=0$, 于是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关. □

证明.

式 (4.3) 与式 (4.4) 两端相减可得

$$k_1(\lambda_1-\lambda_m)\boldsymbol{\alpha}_1+k_2(\lambda_2-\lambda_m)\boldsymbol{\alpha}_2+\cdots+k_{m-1}(\lambda_{m-1}-\lambda_m)\boldsymbol{\alpha}_{m-1}=\mathbf{0},$$

由归纳假设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_{m-1}$ 线性无关, 故

$$k_1(\lambda_1-\lambda_m)=k_2(\lambda_2-\lambda_m)=\cdots=k_{m-1}(\lambda_{m-1}-\lambda_m)=0.$$

又因为 $\lambda_i-\lambda_m\neq 0\ (i=1,2,\cdots,m-1)$, 所以只有 $k_1=k_2=\cdots=k_{m-1}=0$, 代入式 (4.2) 得 $k_m\boldsymbol{\alpha}_m=\mathbf{0}$, 由 $\boldsymbol{\alpha}_m\neq\mathbf{0}$ 可得 $k_m=0$, 于是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关. □

推论 (4.2.1)

设 n 阶矩阵 $\boldsymbol{A}$ 的特征值都是**单特征根**, 则 $\boldsymbol{A}$ 能与对角矩阵相似.

证明.

因为 A 的特征值都是 $\det(\lambda I - A) = 0$ 的单根, 所以 A 有 n 个互异特征值. 互异特征值对应的特征向量是线性无关的, 故 A 有 n 个线性无关的特征向量, 因而 A 能与对角矩阵相似. □

与定理 4 的证明类似, 我们可以得到下面的推论:

推论 (4.2.1)

设 n 阶矩阵 $\boldsymbol{A}$ 的特征值都是单特征根，则 $\boldsymbol{A}$ 能与对角矩阵相似.

证明.

因为 $\boldsymbol{A}$ 的特征值都是 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$ 的单根，所以 $\boldsymbol{A}$ 有 n 个互异特征值. 互异特征值对应的特征向量是线性无关的，故 $\boldsymbol{A}$ 有 n 个线性无关的特征向量，因而 $\boldsymbol{A}$ 能与对角矩阵相似. $\square$

与定理 4 的证明类似，我们可以得到下面的推论:

推论 (4.2.2)

设 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是矩阵 $\boldsymbol{A}$ 的互异特征值, $\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, \cdots, \boldsymbol{\alpha}_{ir_i}$ 是对应于特征值 λ_i 的线性无关的特征向量, 则 $\boldsymbol{\alpha}_{11}, \cdots, \boldsymbol{\alpha}_{1r_1}, \cdots, \boldsymbol{\alpha}_{k1}, \cdots, \boldsymbol{\alpha}_{kr_k}$ 也线性无关.

设 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 是 n 阶矩阵 $\boldsymbol{A}$ 的全部互异特征值, λ_i 是 $\boldsymbol{A}$ 的 k_i 重特征值 $(k_i \geqslant 1)$, 则 $k_1 + k_2 + \cdots + k_r = n$. 若对每一个特征值 $\lambda_i (i = 1, 2, \cdots, r)$, $(\lambda_i \boldsymbol{I} - \boldsymbol{A}) \boldsymbol{X} = \boldsymbol{0}$ 的基础解系由 k_i 个解向量组成, 即 λ_i 恰有 k_i 个线性无关的特征向量, 则由推论 2 可知, $\boldsymbol{A}$ 有 n 个线性无关的特征向量. 而 $(\lambda_i \boldsymbol{I} - \boldsymbol{A}) \boldsymbol{X} = \boldsymbol{0}$ 的基础解系所含解向量个数不大于 k_i, 故可得下面的定理:

推论 (4.2.2)

设 $\lambda_1,\lambda_2,\cdots,\lambda_k$ 是矩阵 $\boldsymbol{A}$ 的互异特征值，$\boldsymbol{\alpha}_{i1},\boldsymbol{\alpha}_{i2},\cdots,\boldsymbol{\alpha}_{ir_i}$ 是对应于特征值 λ_i 的线性无关的特征向量，则 $\boldsymbol{\alpha}_{11},\cdots,\boldsymbol{\alpha}_{1r_1},\cdots,\boldsymbol{\alpha}_{k1},\cdots,\boldsymbol{\alpha}_{kr_k}$ 也线性无关.

设 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 是 n 阶矩阵 $\boldsymbol{A}$ 的全部互异特征值，λ_i 是 $\boldsymbol{A}$ 的 k_i 重特征值 $(k_i \geqslant 1)$，则 $k_1+k_2+\cdots+k_r=n$. 若对每一个特征值 $\lambda_i(i=1,2,\cdots,r)$，$(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系由 k_i 个解向量组成，即 λ_i 恰有 k_i 个线性无关的特征向量，则由推论 2 可知，$\boldsymbol{A}$ 有 n 个线性无关的特征向量. 而 $(\lambda_i\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{0}$ 的基础解系所含解向量个数不大于 k_i，故可得下面的定理:

定理 (4.2.5)

n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是对于 $\boldsymbol{A}$ 的每一个 k_i 重特征根 λ_i，齐次线性方程组 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{X} = \mathbf{0}$ 的基础解系由 k_i 个解向量组成.

$\lambda_i \boldsymbol{I} - \boldsymbol{A}$ 是齐次线性方程组 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{X} = \mathbf{0}$ 的系数矩阵，由系数矩阵的秩与基础解系所含解向量的个数的关系可以得到定理 5 的一个推论.

推论 (4.2.3)

n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是对于每一个 k_i 重特征根 λ_i，$R(\lambda_i \boldsymbol{I} - \boldsymbol{A}) = n - k_i$.

定理 (4.2.5)

n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是对于 $\boldsymbol{A}$ 的每一个 k_i 重特征根 λ_i, 齐次线性方程组$(\lambda_i \boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{X} = \boldsymbol{0}$ 的基础解系由 k_i 个解向量组成.

$\lambda_i \boldsymbol{I} - \boldsymbol{A}$ 是齐次线性方程组 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{X} = \boldsymbol{0}$ 的系数矩阵, 由系数矩阵的秩与基础解系所含解向量的个数的关系可以得到定理 5 的一个推论.

推论 (4.2.3)

n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是对于每一个 k_i 重特征根 λ_i, $R\,(\lambda_i \boldsymbol{I} - \boldsymbol{A}) = n - k_i$.

例 (4.2.4)

下列矩阵能否与对角矩阵相似?

$$\boldsymbol{A}=\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & -2 & 1 \end{pmatrix}, \boldsymbol{B}=\begin{pmatrix} 3 & -1 & -2 \\ 2 & 0 & -2 \\ 2 & -1 & -1 \end{pmatrix}, \boldsymbol{C}=\begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}.$$

解

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & 2\\ 2 & 2 & \lambda-1\end{vmatrix}=(\lambda-1)(\lambda+1)(\lambda-3),$$

$\boldsymbol{A}$ 的特征值都是单根，故 $\boldsymbol{A}$ 能与对角矩阵相似.

$$\det(\lambda \boldsymbol{I}-\boldsymbol{B})=\begin{vmatrix}\lambda-3 & 1 & 2\\ -2 & \lambda & 2\\ -2 & 1 & \lambda+1\end{vmatrix}=\lambda(\lambda-1)^2,$$

对于 2 重特征根 $\lambda_2=1$,

$$\lambda_2\boldsymbol{I}-\boldsymbol{B}=\begin{pmatrix}-2 & 1 & 2\\ -2 & 1 & 2\\ -2 & 1 & 2\end{pmatrix}, R(\lambda_2\boldsymbol{I}-\boldsymbol{B})=1,$$

所以 $\boldsymbol{B}$ 能与对角矩阵相似.

解

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & 2\\ 2 & 2 & \lambda-1\end{vmatrix}=(\lambda-1)(\lambda+1)(\lambda-3),$$

$\boldsymbol{A}$ 的特征值都是单根，故 $\boldsymbol{A}$ 能与对角矩阵相似.

$$\det(\lambda \boldsymbol{I}-\boldsymbol{B})=\begin{vmatrix}\lambda-3 & 1 & 2\\ -2 & \lambda & 2\\ -2 & 1 & \lambda+1\end{vmatrix}\equiv\lambda(\lambda-1)^2,$$

对于 2 重特征根 $\lambda_2=1$,

$$\lambda_2\boldsymbol{I}-\boldsymbol{B}=\begin{pmatrix}-2 & 1 & 2\\ -2 & 1 & 2\\ -2 & 1 & 2\end{pmatrix}, R\left(\lambda_2\boldsymbol{I}-\boldsymbol{B}\right)=1,$$

所以 $\boldsymbol{B}$ 能与对角矩阵相似.

解

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix}\lambda-1 & -2 & -2\\ -2 & \lambda-1 & 2\\ 2 & 2 & \lambda-1\end{vmatrix}=(\lambda-1)(\lambda+1)(\lambda-3),$$

$\boldsymbol{A}$ 的特征值都是单根，故 $\boldsymbol{A}$ 能与对角矩阵相似.

$$\det(\lambda \boldsymbol{I}-\boldsymbol{B})=\begin{vmatrix}\lambda-3 & 1 & 2\\ -2 & \lambda & 2\\ -2 & 1 & \lambda+1\end{vmatrix}\equiv\lambda(\lambda-1)^2,$$

对于 2 重特征根 $\lambda_2=1$,

$$\lambda_2\boldsymbol{I}-\boldsymbol{B}=\begin{pmatrix}-2 & 1 & 2\\ -2 & 1 & 2\\ -2 & 1 & 2\end{pmatrix}, R\left(\lambda_2\boldsymbol{I}-\boldsymbol{B}\right)=1,$$

所以 $\boldsymbol{B}$ 能与对角矩阵相似.

解

$$|\lambda \boldsymbol{I}-\boldsymbol{C}|=\begin{vmatrix}\lambda-3 & -1 & 0\\ 4 & \lambda+1 & 0\\ -4 & 8 & \lambda+2\end{vmatrix}=(\lambda-1)^2(\lambda+2),$$

对于 2 重特征根 $\lambda_1=1$,

$$\lambda_1\boldsymbol{I}-\boldsymbol{C}=\begin{pmatrix}-2 & -1 & 0\\ 4 & 2 & 0\\ -4 & 8 & 3\end{pmatrix},R(\lambda_1\boldsymbol{I}-\boldsymbol{C})=2,$$

所以 $\boldsymbol{C}$ 不能与对角矩阵相似.

解

$$|\lambda \boldsymbol{I}-\boldsymbol{C}|=\begin{vmatrix}\lambda-3 & -1 & 0\\ 4 & \lambda+1 & 0\\ -4 & 8 & \lambda+2\end{vmatrix}=(\lambda-1)^2(\lambda+2),$$

对于 2 重特征根 $\lambda_1=1$,

$$\lambda_1\boldsymbol{I}-\boldsymbol{C}=\begin{pmatrix}-2 & -1 & 0\\ 4 & 2 & 0\\ -4 & 8 & 3\end{pmatrix}, R\left(\lambda_1\boldsymbol{I}-\boldsymbol{C}\right)=2,$$

所以 C 不能与对角矩阵相似.

例 (4.2.5)

已知 $\boldsymbol{\alpha}=\begin{pmatrix}1\\1\\-1\end{pmatrix}$ 是矩阵 $\boldsymbol{A}=\begin{pmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{pmatrix}$ 的特征向量，试确定 a,b 的值与 $\boldsymbol{\alpha}$ 所对应的特征值，并讨论 $\boldsymbol{A}$ 能否与对角矩阵相似.

解

设 α 所对应的特征值为 λ，则

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{\alpha}=\begin{pmatrix}\lambda-2&1&-2\\-5&\lambda-a&-3\\1&-b&\lambda+2\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\boldsymbol{0},$$

解得 $a=-3,\ b=0,\ \lambda=-1$. 于是

例 (4.2.5)

已知 $\boldsymbol{\alpha}=\begin{pmatrix}1\\1\\-1\end{pmatrix}$ 是矩阵 $\boldsymbol{A}=\begin{pmatrix}2&-1&2\\5&a&3\\-1&b&-2\end{pmatrix}$ 的特征向量，试确定 a,b 的值与 $\boldsymbol{\alpha}$ 所对应的特征值，并讨论 $\boldsymbol{A}$ 能否与对角矩阵相似.

解

设 $\boldsymbol{\alpha}$ 所对应的特征值为 λ，则

$$(\lambda\boldsymbol{I}-\boldsymbol{A})\boldsymbol{\alpha}=\begin{pmatrix}\lambda-2&1&-2\\-5&\lambda-a&-3\\1&-b&\lambda+2\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\mathbf{0},$$

解得 $a=-3,\ b=0,\ \lambda=-1$. 于是

解

于是

$$\boldsymbol{A}=\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix},$$

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix} \lambda-2 & 1 & -2 \\ -5 & \lambda+3 & -3 \\ 1 & 0 & \lambda+2 \end{vmatrix}=(\lambda+1)^3.$$

故 $\lambda=-1$ 是 $\boldsymbol{A}$ 的 3 重特征根.

$$R(-\boldsymbol{I}-\boldsymbol{A})=R\left(\begin{pmatrix} -3 & 1 & -2 \\ -5 & 2 & -3 \\ 1 & 0 & 1 \end{pmatrix}\right)=2,$$

所以 $\boldsymbol{A}$ 不能与对角矩阵相似.

解

于是

$$\boldsymbol{A}=\begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix},$$

$$\det(\lambda \boldsymbol{I}-\boldsymbol{A})=\begin{vmatrix} \lambda-2 & 1 & -2 \\ -5 & \lambda+3 & -3 \\ 1 & 0 & \lambda+2 \end{vmatrix}=(\lambda+1)^3.$$

故 $\lambda=-1$ 是 $\boldsymbol{A}$ 的 3 重特征根.

$$R(-\boldsymbol{I}-\boldsymbol{A})=R\left(\begin{pmatrix} -3 & 1 & -2 \\ -5 & 2 & -3 \\ 1 & 0 & 1 \end{pmatrix}\right)=2,$$

所以 $\boldsymbol{A}$ 不能与对角矩阵相似.

小结 (I)

- 相似矩阵的基本概念: 对于 n 阶矩阵 $\boldsymbol{A}, \boldsymbol{B}$, 若存在可逆矩阵 $\boldsymbol{P}$, 使 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$, 则称 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 记为 $\boldsymbol{A} \sim \boldsymbol{B}$.
- 矩阵之间的相似关系是一种等价关系, 具有反身性, 对称性, 传递性.
- 相似矩阵的性质: 若矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似, 则
  (1) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的特征多项式和特征值.
  (2) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行列式相等.
  (3) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的迹相同.
  (4) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的秩相等.
  (5) $\boldsymbol{A}^k$ 与 $\boldsymbol{B}^k$ 相似, 其中 k 为正整数. $f(\boldsymbol{A}) \sim f(\boldsymbol{B})$.
  (6) $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 有相同的可逆性, 当 $\boldsymbol{A}$, $\boldsymbol{B}$ 可逆时, $\boldsymbol{A}^{-1}$ 与 $\boldsymbol{B}^{-1}$ 也相似.
- 矩阵的相似对角化: 若 n 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{\Lambda}=\operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 相似, 则 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是 $\boldsymbol{A}$ 的全部特征值.
- 若 $\boldsymbol{A}$ 与对角矩阵 $\boldsymbol{\Lambda}$ 相似, 则 $\boldsymbol{\Lambda}$ 的主对元就是 $\boldsymbol{A}$ 的全部特征值.

小结 (II)

- 矩阵的相似对角化: n 阶矩阵 $\boldsymbol{A}$ 能与对角矩阵 $\boldsymbol{\Lambda}$ 相似的充要条件是 $\boldsymbol{A}$ 有 n 个线性无关的特征向量.
- 设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是矩阵 $\boldsymbol{A}$ 的互异特征值, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 $\boldsymbol{A}$ 分别对应于这些特征值的特征向量, 则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关.
- 设 n 阶矩阵 $\boldsymbol{A}$ 的特征值都是单特征根, 则 $\boldsymbol{A}$ 能与对角矩阵相似.
- 设 $\lambda_1, \lambda_2, \cdots, \lambda_k$ 是矩阵 $\boldsymbol{A}$ 的互异特征值, $\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, \cdots, \boldsymbol{\alpha}_{ir_i}$ 是对应于特征值 λ_i 的线性无关的特征向量, 则 $\boldsymbol{\alpha}_{11}, \cdots, \boldsymbol{\alpha}_{1r_1}, \cdots, \boldsymbol{\alpha}_{k1}, \cdots, \boldsymbol{\alpha}_{kr_k}$ 也线性无关.
- n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是对于 $\boldsymbol{A}$ 的每一个 k_i 重特征根 λ_i, $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$ 的基础解系由 k_i 个解向量组成.
- n 阶矩阵 $\boldsymbol{A}$ 与对角矩阵相似的充要条件是对于每一个 k_i 重特征根 λ_i, $R(\lambda_i \boldsymbol{I} - \boldsymbol{A}) = n - k_i$.

小结 (III)

- n 阶矩阵 $\boldsymbol{A}$ 对角化的一般步骤:

(1) 写出 n 阶矩阵 $\boldsymbol{A}$ 的特征多项式 $|\lambda \boldsymbol{I} - \boldsymbol{A}|$, 求出 $\boldsymbol{A}$ 的全部特征值.

(2) 若 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是 $\boldsymbol{A}$ 的全部不同的特征值. 对每个 $\lambda_i\,(i = 1, 2, \cdots, m)$ 设其重数为 $k_i\,(k_1 + k_2 + \cdots + k_m = n)$, 求解齐次方程组 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{0}$, 得到一个基础解系为 $\boldsymbol{\alpha}_{i1}, \boldsymbol{\alpha}_{i2}, \cdots, \boldsymbol{\alpha}_{is_i}$. 若 $s_i = k_i\,(i = 1, 2, \cdots, m)$, 则 $\boldsymbol{A}$ 可对角化, 否则 $\boldsymbol{A}$ 不可以对角化.

(3) 若 $\boldsymbol{A}$ 可以对角化, 用已求出的全部基础解系的解向量 (所有解向量的个数必为 n) 作为矩阵 $\boldsymbol{P}$ 的列向量. 则 $\boldsymbol{P}$ 可逆, 且 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}$, 其中 $\boldsymbol{\Lambda}$ 是对角矩阵, 其主对角线上的元素为 $\boldsymbol{A}$ 的全部特征值.