线性代数

leip@lzu.edu.cn
兰州大学 数学与统计学院

2017 年 10 月 25 日

第五章

特征值与特征向量、矩阵的对角化

目录

由上一章的讨论, 我们知道, 在数域 F 上的 n 维线性空间 V 中取定一组基后, V 上的所有线性变换的集合 $L(V)$ 与数域 F 上的所有 n 阶矩阵的集合 $F^{n\times n}$ 之间存在一一对应. 且 V 上的每一个线性变换都可以用它在这组基下的矩阵 A 在 F^n 上定义的线性变换

$$\begin{aligned} \mathcal{A}: \quad F^n &\to F^n \\ X &\mapsto AX \end{aligned}$$

表示. 这样, 给定一个线性变换我们希望能找到一组基使得线性变换在这组基下的矩阵最为简单.

对角矩阵被认为是一类简单的矩阵.

矩阵的对角化问题

- ☞ 满足什么条件的线性变换, 可以找到一组基使得线性变换在这组基下的矩阵是对角矩阵?
- ☞ 或等价地：满足什么条件的矩阵可以相似于对角矩阵?

目录

定义5.1.1

设 $\mathcal{A}$ 是数域 F 上线性空间 V 的一个线性变换. 如果存在数域 F 中的数 λ_0 以及 V 中的非零向量 ξ 使得

$$\mathcal{A}\xi = \lambda_0\xi \tag{5.1}$$

那么 λ_0 称为 $\mathcal{A}$ 的一个特征值, ξ 称为 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量.

定义5.1.1

设 $\mathcal{A}$ 是数域 F 上线性空间 V 的一个线性变换. 如果存在数域 F 中的数 λ_0 以及 V 中的非零向量 ξ 使得

$$\mathcal{A}\xi = \lambda_0\xi \tag{5.1}$$

那么 λ_0 称为 $\mathcal{A}$ 的一个特征值, ξ 称为 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量.

特征向量的几何意义在于：特征向量的经过线性变换后, 与原向量是共线的, 或者方向不变($\lambda_0 > 0$), 或者方向相反($\lambda_0 < 0$), 而当 $\lambda_0 = 0$时, 特征向量被线性变换变成零向量 0.

定义5.1.1

设 $\mathcal{A}$ 是数域 F 上线性空间 V 的一个线性变换. 如果存在数域 F 中的数 λ_0 以及 V 中的非零向量 ξ 使得

$$\mathcal{A}\xi = \lambda_0\xi \tag{5.1}$$

那么 λ_0 称为 $\mathcal{A}$ 的一个特征值, ξ 称为 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量.

特征向量的几何意义在于：特征向量的经过线性变换后, 与原向量是共线的, 或者方向不变($\lambda_0 > 0$), 或者方向相反($\lambda_0 < 0$), 而当 $\lambda_0 = 0$时, 特征向量被线性变换变成零向量 0.

推论5.1.2

设 $\mathcal{A}$ 是数域 F 上线性空间 V 的一个线性变换.

(1) 若 ξ 是 $\mathcal{A}$ 的特征向量, 则存在 $\mathcal{A}$ 的唯一的特征值 λ_0 使得 ξ 是 $\mathcal{A}$ 的属于 λ_0 的特征向量, 即一个特征向量属于唯一的一个特征值.

(2) 若 λ_0 是 $\mathcal{A}$ 的特征值, 则

$$V_{\lambda_0} = \{\xi \in V | \mathcal{A}\xi = \lambda_0\xi\}$$

是 V 的一个非零子空间, 称为 $\mathcal{A}$ 的属于特征值 λ_0 的特征子空间.

证明 (1) 设 ξ 是 $\mathcal{A}$ 的同时属于特征值 λ_1和 λ_2 的特征向量, 即有

$$\mathcal{A}\xi = \lambda_1\xi, \quad \mathcal{A}\xi = \lambda_2\xi.$$

则 $(\lambda_1 - \lambda_2)\xi = 0$. 由于 $\xi \neq 0$, 因此 $\lambda_1 - \lambda_2 = 0$, 即 $\lambda_1 = \lambda_2$.

证明 (1) 设 ξ 是 $\mathcal{A}$ 的同时属于特征值 λ_1和 λ_2 的特征向量, 即有

$$\mathcal{A}\xi = \lambda_1\xi, \quad \mathcal{A}\xi = \lambda_2\xi.$$

则 $(\lambda_1 - \lambda_2)\xi = 0$. 由于 $\xi \neq 0$, 因此 $\lambda_1 - \lambda_2 = 0$, 即 $\lambda_1 = \lambda_2$.
(2) 若 λ_0 是 $\mathcal{A}$ 的特征值, 则存在非零向量 $\xi \in V$ 使得 $\mathcal{A}\xi = \lambda_0\xi$, 因而 $V_{\lambda_0} \neq \{0\}$. 对任意 $\eta_1, \eta_2 \in V_{\lambda_0}$ 以及 $k \in F$, 有

$$\mathcal{A}(\eta_1 + \eta_2) = \mathcal{A}\eta_1 + \mathcal{A}\eta_2 = \lambda_0(\eta_1 + \eta_2)$$
$$\mathcal{A}(k\eta_1) = k\mathcal{A}\eta_1 = k\lambda_0\eta_1 = \lambda_0(k\eta_1).$$

所以 V_{λ_0} 是 V 的非零子空间.

证明 (1) 设 ξ 是 $\mathcal{A}$ 的同时属于特征值 λ_1 和 λ_2 的特征向量, 即有

$$\mathcal{A}\xi = \lambda_1\xi, \quad \mathcal{A}\xi = \lambda_2\xi.$$

则 $(\lambda_1 - \lambda_2)\xi = 0$. 由于 $\xi \neq 0$, 因此 $\lambda_1 - \lambda_2 = 0$, 即 $\lambda_1 = \lambda_2$.
(2) 若 λ_0 是 $\mathcal{A}$ 的特征值, 则存在非零向量 $\xi \in V$ 使得 $\mathcal{A}\xi = \lambda_0\xi$, 因而 $V_{\lambda_0} \neq \{0\}$. 对任意 $\eta_1, \eta_2 \in V_{\lambda_0}$ 以及 $k \in F$, 有

$$\mathcal{A}(\eta_1 + \eta_2) = \mathcal{A}\eta_1 + \mathcal{A}\eta_2 = \lambda_0(\eta_1 + \eta_2)$$
$$\mathcal{A}(k\eta_1) = k\mathcal{A}\eta_1 = k\lambda_0\eta_1 = \lambda_0(k\eta_1).$$

所以 V_{λ_0} 是 V 的非零子空间.

1. 若 ξ 是 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量,则 $k\xi\,(k \neq 0)$ 也是 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量.
2. 特征向量一定是非零向量. 特征子空间 V_{λ_0} 中除零向量外的所有向量都是 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量.
3. 线性变换的属于一个特征值的特征向量有无穷多个.

例5.1.1

1. 设 $\mathcal{R}_\pi$ 是平面 R^2 上的绕坐标原点逆时针旋转 π 弧度的旋转变换.则

$$\mathcal{R}_\pi\alpha = -\alpha, \quad \forall\alpha \in R^2.$$

因此 -1 是 $\mathcal{R}_\pi$ 的特征值, 而 R^2 中任意非零向量都是 $\mathcal{R}_\pi$ 的属于特征值 -1 的特征向量.

2. 如果 $\mathcal{T}_x$ 是平面 R^2 上关于 x-轴的镜面反射, 那么

$$\mathcal{T}_x\alpha = \alpha, \quad \forall\alpha = (x, 0)^T \in R^2.$$

因此 1 是 $\mathcal{T}_x$ 的特征值, $(x, 0)^T$, $x \neq 0$, 都是 $\mathcal{T}_x$ 的属于特征值 1 的特征向量;又

$$\mathcal{T}_x\beta = -\beta, \quad \forall\beta = (0, y)^T \in R^2.$$

因此 -1 也是 $\mathcal{T}_x$ 的特征值, $(0, y)^T$, $y \neq 0$, 都是 $\mathcal{T}_x$ 的属于特征值 -1 的特征向量.

例5.1.2

在线性空间 V 中

1. 1 是单位变换 $\mathcal{E}$ 的特征值, 任意非零向量都是 $\mathcal{E}$ 的属于特征值 1 的特征向量.
2. 0 是零变换的特征值, 任意非零向量都是零变换的属于特征值 0 的特征向量.

例5.1.2

在线性空间 V 中

1. 1 是单位变换 $\mathcal{E}$ 的特征值, 任意非零向量都是 $\mathcal{E}$ 的属于特征值 1 的特征向量.
2. 0 是零变换的特征值, 任意非零向量都是零变换的属于特征值 0 的特征向量.

补充例题 5.1.1

如果 $\mathcal{T}_{xoy}$ 是几何空间 R^3 上关于 xoy-平面的镜面反射,那么

例5.1.2

在线性空间 V 中

1. 1 是单位变换 $\mathcal{E}$ 的特征值, 任意非零向量都是 $\mathcal{E}$ 的属于特征值 1 的特征向量.
2. 0 是零变换的特征值, 任意非零向量都是零变换的属于特征值 0 的特征向量.

补充例题 5.1.1

如果 $\mathcal{T}_{xoy}$ 是几何空间 R^3 上关于 xoy-平面的镜面反射,那么

1. $\mathcal{T}_{xoy}\alpha = \alpha, \quad \forall \alpha = (x, y, 0)^T \in R^3.$
   因此 1 是 $\mathcal{T}_{xoy}$ 的特征值, $(x, y, 0)^T$, $xy \neq 0$, 都是 $\mathcal{T}_{xoy}$ 的属于特征值 1 的特征向量;

例5.1.2

在线性空间 V 中

1. 1 是单位变换 $\mathcal{E}$ 的特征值, 任意非零向量都是 $\mathcal{E}$ 的属于特征值 1 的特征向量.
2. 0 是零变换的特征值, 任意非零向量都是零变换的属于特征值 0 的特征向量.

补充例题 5.1.1

如果 $\mathcal{T}_{xoy}$ 是几何空间 R^3 上关于 xoy-平面的镜面反射,那么

1. $\mathcal{T}_{xoy}\alpha = \alpha, \quad \forall \alpha = (x, y, 0)^T \in R^3.$
   因此 1 是 $\mathcal{T}_{xoy}$ 的特征值, $(x, y, 0)^T$, $xy \neq 0$, 都是 $\mathcal{T}_{xoy}$ 的属于特征值 1 的特征向量;
2. $\mathcal{T}_{xoy}\beta = -\beta, \quad \forall \beta = (0, 0, z)^T \in R^3.$
   因此 -1 也是 $\mathcal{T}_{xoy}$ 的特征值, $(0, 0, z)^T$, $z \neq 0$, 都是 $\mathcal{T}_{xoy}$ 的属于特征值 -1 的特征向量.

补充例题 5.1.2

1. k 是线性空间 V 中数乘变换 $\mathcal{K}$ 的特征值, 任意非零向量都是 $\mathcal{K}$ 的属于特征值 k 的特征向量.
2. 如果 $\mathcal{R}_z$ 是几何空间 R^3 中绕 z-轴逆时针旋转 $\theta\ (\neq k\pi,\ k \in Z)$ 弧度的变换,那么

补充例题 5.1.2

1. k 是线性空间 V 中数乘变换 $\mathcal{K}$ 的特征值, 任意非零向量都是 $\mathcal{K}$ 的属于特征值 k 的特征向量.
2. 如果 $\mathcal{R}_z$ 是几何空间 R^3 中绕 z-轴逆时针旋转 θ $(\neq k\pi,\ k \in Z)$ 弧度的变换,那么 $\mathcal{R}_z\beta = \beta,\quad \forall\beta = (0,0,z)^T \in R^3$. 因此 1 是 $\mathcal{R}_z$ 的特征值, $(0,0,z)^T$, $z \neq 0$, 都是 $\mathcal{R}_z$ 的属于特征值 1 的特征向量.

补充例题 5.1.2

1. k 是线性空间 V 中数乘变换 $\mathcal{K}$ 的特征值, 任意非零向量都是 $\mathcal{K}$ 的属于特征值 k 的特征向量.
2. 如果 $\mathcal{R}_z$ 是几何空间 R^3 中绕 z-轴逆时针旋转 θ $(\neq k\pi,\ k \in Z)$ 弧度的变换,那么 $\mathcal{R}_z\beta = \beta,\quad \forall\beta = (0,0,z)^T \in R^3$. 因此 1 是 $\mathcal{R}_z$ 的特征值, $(0,0,z)^T$, $z \neq 0$, 都是 $\mathcal{R}_z$ 的属于特征值 1 的特征向量.

补充例题 5.1.3

如果 $\mathcal{S}_{xoy}$ 是几何空间 R^3 上关于 xoy-平面的投影变换,那么

补充例题 5.1.2

1. k 是线性空间 V 中数乘变换 $\mathcal{K}$ 的特征值, 任意非零向量都是 $\mathcal{K}$ 的属于特征值 k 的特征向量.
2. 如果 $\mathcal{R}_z$ 是几何空间 R^3 中绕 z-轴逆时针旋转 $\theta\ (\neq k\pi,\ k\in Z)$ 弧度的变换,那么 $\mathcal{R}_z\beta=\beta,\quad \forall\beta=(0,0,z)^T\in R^3$. 因此 1 是 $\mathcal{R}_z$ 的特征值, $(0,0,z)^T$, $z\neq 0$, 都是 $\mathcal{R}_z$ 的属于特征值 1 的特征向量.

补充例题 5.1.3

如果 $\mathcal{S}_{xoy}$ 是几何空间 R^3 上关于 xoy-平面的投影变换,那么

1. $\mathcal{S}_{xoy}\alpha=\alpha,\quad \forall\alpha=(x,y,0)^T\in R^3.$
   因此 1 是 $\mathcal{S}_{xoy}$ 的特征值, $(x,y,0)^T$, $xy\neq 0$, 都是 $\mathcal{S}_{xoy}$ 的属于特征值 1 的特征向量;

补充例题 5.1.2

1. k 是线性空间 V 中数乘变换 $\mathcal{K}$ 的特征值, 任意非零向量都是 $\mathcal{K}$ 的属于特征值 k 的特征向量.
2. 如果 $\mathcal{R}_z$ 是几何空间 R^3 中绕 z-轴逆时针旋转 θ $(\neq k\pi,\ k \in Z)$ 弧度的变换,那么 $\mathcal{R}_z\beta = \beta,\quad \forall\beta = (0,0,z)^T \in R^3$. 因此 1 是 $\mathcal{R}_z$ 的特征值, $(0,0,z)^T$, $z \neq 0$, 都是 $\mathcal{R}_z$ 的属于特征值 1 的特征向量.

补充例题 5.1.3

如果 $\mathcal{S}_{xoy}$ 是几何空间 R^3 上关于 xoy-平面的投影变换,那么

1. $\mathcal{S}_{xoy}\alpha = \alpha,\quad \forall\alpha = (x,y,0)^T \in R^3.$
   因此 1 是 $\mathcal{S}_{xoy}$ 的特征值, $(x,y,0)^T$, $xy \neq 0$, 都是 $\mathcal{S}_{xoy}$ 的属于特征值 1 的特征向量;
2. $\mathcal{S}_{xoy}\beta = 0 = 0\beta,\quad \forall\beta = (0,0,z)^T \in R^3.$
   因此 0 也是 $\mathcal{S}_{xoy}$ 的特征值, $(0,0,z)^T$, $z \neq 0$, 都是 $\mathcal{S}_{xoy}$ 的属于特征值 0 的特征向量.

现在我们讨论如何求特征值和特征向量.
设 V 是数域 F 上的 n 维线性空间, $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 是它的一组基, V 上的线性变换 $\mathcal{A}$ 在这组基下的矩阵是 A. 又设 λ_0 是 $\mathcal{A}$ 的一个特征值, ξ 是 $\mathcal{A}$ 的属于特征值 λ_0 的一个特征向量, 它在基 $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 下的坐标是 $X_0 = (x_1^0, x_2^0, \cdots, x_n^0)^T$. 则 $\mathcal{A}\xi$ 的坐标是 AX_0. 此时, (5.1) 等价于

$$AX_0 = \lambda_0 X_0$$

或

$$(\lambda_0 E - A)X_0 = 0.$$

这说明特征向量 ξ 的坐标 X_0 是齐次线性方程组

$$(\lambda_0 E - A)X = 0 \tag{5.2}$$

的解. 因为 $\xi \neq 0$, 所以 $X_0 \neq 0$, 方程组(5.2)有非零解, 从而方程组系数矩阵的行列式等于零, 即 $|\lambda_0 E - A| = 0$.

反过来, 如果数域 F 中的数 λ_0 使得 $|\lambda_0 E - A| = 0$, 那么齐次线性方程组 (5.2) 有非零解. 对于 (5.2) 的任一非零解 X_0, 令

$$\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) X_0.$$

则

$$\mathcal{A}\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) A X_0 = \lambda_0 (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n) X_0 = \lambda_0 \xi,$$

因而 ξ 就是 $\mathcal{A}$ 的属于特征值 λ_0 的特征向量.

定义5.1.3

设 $A=(a_{ij})$ 是复数域 C 上的 n 阶矩阵. 如果存在复数 λ_0 以及 C^n 中的非零向量 η 使得

$$A\eta=\lambda_0\eta, \tag{5.3}$$

那么称 λ_0 为矩阵 A 的一个特征值, 称非零向量 η 为矩阵 A 的属于特征值 λ_0 的特征向量. 设 λ 是一文字, 称多项式

$$|\lambda E-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{1n}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix} \tag{5.4}$$

为 A 的特征多项式, 它是复数域 C 上 λ 的一个首项系数为 1 的 n 次多项式.

由定义 5.1.3 不难得到

推论5.1.4

1. λ_0 是矩阵 A 的特征值当且仅当 λ_0 是 A 的特征多项式的根, 即 $|\lambda_0 E - A| = 0$.
2. 若 λ_0 是 A 的一个特征值, 则齐次线性方程组 $(\lambda_0 E - A)X = 0$ 的所有非零解就是 A 的属于特征值 λ_0 的全部特征向量.

由定义 5.1.3 不难得到

推论5.1.4

1. λ_0 是矩阵 A 的特征值当且仅当 λ_0 是 A 的特征多项式的根, 即 $|\lambda_0 E - A| = 0$.
2. 若 λ_0 是 A 的一个特征值, 则齐次线性方程组 $(\lambda_0 E - A)X = 0$ 的所有非零解就是 A 的属于特征值 λ_0 的全部特征向量.

设 A 是复数域 C 上的矩阵, $\lambda_0 \in C$ 是 A 的一个特征值. 则

$$V_{\lambda_0} = \{\eta \in F^n | A\eta = \lambda_0 \eta\}$$

是 C^n 的子空间($(\lambda_0 E - A)X = 0$ 的解空间), 亦称为 A 属于(对应于) λ_0 的特征子空间.

☞ $|A - \lambda E| =$

由定义 5.1.3 不难得到

推论5.1.4

1. λ_0 是矩阵 A 的特征值当且仅当 λ_0 是 A 的特征多项式的根, 即 $|\lambda_0 E - A| = 0$.
2. 若 λ_0 是 A 的一个特征值, 则齐次线性方程组 $(\lambda_0 E - A)X = 0$ 的所有非零解就是 A 的属于特征值 λ_0 的全部特征向量.

设 A 是复数域 C 上的矩阵, $\lambda_0 \in C$ 是 A 的一个特征值. 则

$$V_{\lambda_0} = \{\eta \in F^n | A\eta = \lambda_0 \eta\}$$

是 C^n 的子空间($(\lambda_0 E - A)X = 0$ 的解空间), 亦称为 A 属于(对应于) λ_0 的特征子空间.

☞ $|A - \lambda E| = (-1)^n |\lambda E - A|$.

定理5.1.5

相似矩阵有相同的特征多项式, 进而有相同的特征值.

定理5.1.5

相似矩阵有相同的特征多项式, 进而有相同的特征值.

证明 设 A, B 是相似矩阵, 即存在可逆矩阵 T 使得 $B = T^{-1}AT$. 于是

$$|\lambda E - B| = |\lambda E - T^{-1}AT| = |T^{-1}||\lambda E - A||T| = |\lambda E - A|.$$

定理5.1.5

相似矩阵有相同的特征多项式, 进而有相同的特征值.

证明 设 A, B 是相似矩阵, 即存在可逆矩阵 T 使得 $B=T^{-1}AT$. 于是

$$|\lambda E-B|=|\lambda E-T^{-1}AT|=|T^{-1}||\lambda E-A||T|=|\lambda E-A|.$$

定理 5.1.5 说明, 线性变换的矩阵的特征多项式与基的选取无关, 它是被线性变换唯一决定的, 因此我们也称它是线性变换的特征多项式.

☞ 定理 5.1.5 的逆不成立, 即特征多项式相同的矩阵不一定相似. 如

$$A=\begin{pmatrix}1&0\\0&1\end{pmatrix},\quad B=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

A 和 B 有相同的特征多项式 $(\lambda-1)^2$, 但是 A 和 B 不相似.

定理5.1.5

相似矩阵有相同的特征多项式, 进而有相同的特征值.

证明 设 A, B 是相似矩阵, 即存在可逆矩阵 T 使得 $B = T^{-1}AT$. 于是

$$|\lambda E - B| = |\lambda E - T^{-1}AT| = |T^{-1}||\lambda E - A||T| = |\lambda E - A|.$$

定理 5.1.5 说明, 线性变换的矩阵的特征多项式与基的选取无关, 它是被线性变换唯一决定的, 因此我们也称它是线性变换的特征多项式.

☞ 定理 5.1.5 的逆不成立, 即特征多项式相同的矩阵不一定相似. 如

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

A 和 B 有相同的特征多项式 $(\lambda - 1)^2$, 但是 A 和 B 不相似. 因为和单位矩阵相似的矩阵只能是单位矩阵本身.

综上所述,我们有

定理5.1.6

设 V 是数域 F 上的 n 维线性空间, $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 是它的一组基, V 上的线性变换 $\mathcal{A}$ 在这组基下的矩阵是 A.

1. 如果 λ_0 是 $\mathcal{A}$ 的特征值, ξ 是 $\mathcal{A}$ 的属于 λ_0 的一个特征向量, 那么 λ_0 是 A 的特征值, ξ 在基 $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$ 下的坐标 X_0 是 A 的属于 λ_0 的一个特征向量.
2. 反之, 如果 $\lambda_0 \in F$ 是 A 的特征值, X_0 是 A 的属于 λ_0 的特征向量, 那么 λ_0 是 $\mathcal{A}$ 的特征值, 且
$$\xi = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)X_0$$
是 $\mathcal{A}$ 的属于 λ_0 的特征向量.

根据定理 5.1.5, 确定数域 F 上线性空间 V 上的线性变换 $\mathcal{A}$ 的特征值和特征向量的过程可以归结为如下步骤:
(1) 在 V 中取一组基 $(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n)$, 求出 $\mathcal{A}$ 在这组基下的矩阵 A;
(2) 求出 A 的特征多项式 $|\lambda E - A|$ 在数域 F 中的全部根 $\lambda_1, \lambda_2, \cdots, \lambda_s$, 这就是线性变换 $\mathcal{A}$ 的全部特征值;
(3) 对于 $i = 1, 2, \cdots, s$, 求出齐次线性方程组

$$(\lambda_i E - A)X = 0$$

的一个基础解系 $(\eta_{i1}, \eta_{i2}, \cdots, \eta_{ir_i})$, 这就是矩阵 A 的特征子空间 V_{λ_i} 的一组基, 即 A 的属于 λ_i 的全部特征向量组成的向量组的一个极大线性无关组. 因此

$$\eta = k_1\eta_{i1} + k_2\eta_{i2} + \cdots + k_{r_i}\eta_{ir_i},$$

就给出了 A 属于 λ_i 的全部特征向量, 其中 $k_1, k_2, \cdots, k_{r_i}$ 是 F 中任意一组不全为零的常数.

(4) 对于 $i=1,2,\cdots,s$, 令

$$\xi_{ij}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\eta_{ij},\quad j=1,2,\cdots,r_i.$$

则 $(\xi_{i1},\xi_{i2},\cdots,\xi_{ir_i})$ 就是 $\mathcal{A}$ 的特征子空间 V_{λ_i} 的一组基,

$$\xi=k_1\xi_{i1}+k_2\xi_{i2}+\cdots+k_{r_i}\xi_{ir_i},$$

其中 $k_1,k_2,\cdots,k_{r_i}$ 是 F 中任意一组不全为零的常数, 就给出了 $\mathcal{A}$ 属于 λ_i 的全部特征向量.

其中 (2), (3) 两步就是求矩阵 A 的特征值和特征向量的过程.

(4) 对于 $i=1,2,\cdots,s$, 令

$$\xi_{ij}=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)\eta_{ij},\quad j=1,2,\cdots,r_i.$$

则 $(\xi_{i1},\xi_{i2},\cdots,\xi_{ir_i})$ 就是 $\mathcal{A}$ 的特征子空间 V_{λ_i} 的一组基,

$$\xi=k_1\xi_{i1}+k_2\xi_{i2}+\cdots+k_{r_i}\xi_{ir_i},$$

其中 $k_1,k_2,\cdots,k_{r_i}$ 是 F 中任意一组不全为零的常数, 就给出了 $\mathcal{A}$ 属于 λ_i 的全部特征向量.
其中 (2), (3) 两步就是求矩阵 A 的特征值和特征向量的过程.

例5.1.3

设 $(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ 是 3 维空间 V 的一组基, V 上的线性变换 $\mathcal{A}$ 在基 $(\varepsilon_1,\varepsilon_2,\varepsilon_3)$ 下的矩阵是

$$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}.$$

求 $\mathcal{A}$ 的特征值与特征向量.

解 因为 A 的特征多项式为

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

所以 A 的特征值为$-1, -1$, 5.
对于特征值 -1, 对应的齐次方程组为

$$\begin{cases} -2x_1 & -2x_2 & -2x_3 & = & 0 \\ -2x_1 & -2x_2 & -2x_3 & = & 0 \\ -2x_1 & -2x_2 & -2x_3 & = & 0 \end{cases}$$

解得它的一个基础解系

$$\eta_1 = (1, 0, -1)^T, \quad \eta_2 = (0, 1, -1)^T,$$

这是 A 的属于特征值 -1 的特征向量组的一个极大线性无关组,

因此 $\mathcal{A}$ 的属于特征值 -1 的全部特征向量为

$$\eta = k_1\eta_1 + k_2\eta_2,$$

其中 k_1, k_2 是不全为零的任意常数. 而 $\mathcal{A}$ 属于特征值 -1 的两个线性无关的特征向量为

$$\xi_1 = \varepsilon_1 - \varepsilon_3, \quad \xi_2 = \varepsilon_2 - \varepsilon_3,$$

而属于 -1 的全部特征向量就是

$$k_1\xi_1 + k_2\xi_2,$$

其中 k_1, k_2 为不全为零的任意常数.
对应特征值 5 的齐次线性方程组为

$$\begin{cases} 4x_1 & -2x_2 & -2x_3 & = & 0 \\ -2x_1 & +4x_2 & -2x_3 & = & 0 \\ -2x_1 & -2x_2 & +4x_3 & = & 0 \end{cases}$$

解得它的一个基础解系

$$\eta_3 = (1, 1, 1)^T,$$

这是 A 的属于特征值 5 的特征向量组的一个极大无关组, 因此 A 的属于 5 的全部特征向量为

$$k_3\eta_3 = k_3(1, 1, 1)^T,$$

其中 k_3 是任意不等于零的数. 而线性变换 $\mathcal{A}$ 的属于特征值 5 的特征向量组的一个极大线性无关组为

$$\xi_3 = \varepsilon_1 + \varepsilon_2 + \varepsilon_3,$$

$\mathcal{A}$ 的属于特征值 5 的全部特征向量就是

$$k_3\xi_3 = k_3\varepsilon_1 + k_3\varepsilon_2 + k_3\varepsilon_3,$$

其中 k_3 是任意不等于零的数.

矩阵的特征多项式有重要的应用，下面我们研究它的一些基本性质.
设 $A=(a_{ij})$ 是复数域 C 上的 n 阶矩阵. 在 A 的特征多项式

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

的展开式中，有一项是主对角线上 n 个元素的乘积

$$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn}),$$

展开式中的其余各项，至多包含 $n-2$ 个主对角线上的元素，λ 的次数最多是 $n-2$，因此特征多项式的 n 次项和 $n-1$ 次项只能包含在对角线上元素的乘积中，它们是

$$\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}.$$

另一方面, 在特征多项式中令 $\lambda = 0$, 即得常数项 $|-A| = (-1)^n|A|$. 因此, 如果只写出特征多项式的前两项与常数项, 就有

$$\begin{aligned}|\lambda E - A| &= \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (-1)^n|A| \\ &= \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n\lambda_1\lambda_2\cdots\lambda_n \\ &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)\end{aligned}$$

根据多项式根与系数的关系可知, A 的全体特征值的和为 $a_{11} + a_{22} + \cdots + a_{nn}$, 称为 A 的迹, 记作 $\operatorname{tr} A$. 而 A 的全体特征值的乘积等于 $|A|$. 因此我们有

另一方面, 在特征多项式中令 $\lambda = 0$, 即得常数项 $|-A| = (-1)^n|A|$. 因此, 如果只写出特征多项式的前两项与常数项, 就有

$$\begin{aligned}|\lambda E - A| =&\lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (-1)^n|A| \\ =&\lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \cdots + (-1)^n\lambda_1\lambda_2\cdots\lambda_n \\ =&(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)\end{aligned}$$

根据多项式根与系数的关系可知, A 的全体特征值的和为 $a_{11} + a_{22} + \cdots + a_{nn}$, 称为 A 的迹, 记作 $\mathrm{tr}\ A$. 而 A 的全体特征值的乘积等于 $|A|$. 因此我们有

定理5.1.7

设 $A = (a_{ij})$ 是 n 阶矩阵, $\lambda_1, \lambda_2, \cdots, \lambda_n$ 是它的全部 n 个特征值(包括重特征值). 则

1. $\mathrm{tr}\ A = \lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}$,
2. $|A| = \lambda_1\lambda_2\cdots\lambda_n$.

最后我们给出特征多项式的一个重要性质, 我们略去它的证明, 有兴趣的读者可以参考有关的文献给出它的证明.

设 A 是数域 F 上的 n 阶矩阵,

$$f(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0$$

是数域 F 上的一个 m 次多项式. 则

$$f(A) = a_m A^m + a_{m-1} A^{m-1} + \cdots + a_1 A + a_0 E,$$

也是数域 F 上的 n 阶矩阵, 称为**矩阵 A 的多项式**.

定理5.1.8 (哈密顿—凯莱(Hamilton-Cayley)定理)

设 $A = (a_{ij})$ 是数域 F 上的一个 n 阶矩阵, $f(\lambda) = |\lambda E - A|$ 是 A 的特征多项式. 则

$$f(A) = A^n - (a_{11} + a_{22} + \cdots + a_{nn}) A^{n-1} + \cdots + (-1)^n |A| E = 0.$$

最后我们给出特征多项式的一个重要性质, 我们略去它的证明, 有兴趣的读者可以参考有关的文献给出它的证明.

设 A 是数域 F 上的 n 阶矩阵,

$$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$$

是数域 F 上的一个 m 次多项式. 则

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

也是数域 F 上的 n 阶矩阵, 称为矩阵 A 的多项式.

定理5.1.8 (哈密顿一凯莱(Hamilton-Cayley)定理)

设 $A=(a_{ij})$ 是数域 F 上的一个 n 阶矩阵, $f(\lambda)=|\lambda E-A|$ 是 A 的特征多项式. 则

$$f(A)=A^n-(a_{11}+a_{22}+\cdots+a_{nn})A^{n-1}+\cdots+(-1)^n|A|E=0.$$

使得 $f(A)=0$ 的多项式 $f(x)$ 称为矩阵 A 的零化多项式. 据定理 5.1.8 知一个矩阵的特征多项式是它的零化多项式.

补充例题 5.1.4 (补充: 矩阵的多项式)

设 A 是一个 n 阶方阵, $f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ 是一个关于变量 x 的多项式, 则

$$f(A) \xlongequal{\text{def}} a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 E,$$

是一个跟 A 有关的 n 阶方阵,称为矩阵 A 的多项式.

补充例题 5.1.4 (补充: 矩阵的多项式)

设 A 是一个 n 阶方阵, $f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0$ 是一个关于变量 x 的多项式, 则

$$f(A) \xlongequal{\text{def}} a_m A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 E,$$

是一个跟 A 有关的 n 阶方阵,称为矩阵 A 的多项式.

例如 设 $f(x) = x^2 - (a+d)x + ad - bc$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,则

$$f(A) = A^2 - (a+d)A + (ad-bc)E$$

补充例题 5.1.4 (补充: 矩阵的多项式)

设 A 是一个 n 阶方阵, $f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$ 是一个关于变量 x 的多项式, 则

$$f(A)\xlongequal{\text{def}}a_mA^m+a_{m-1}A^{m-1}+\cdots+a_1A+a_0E,$$

是一个跟 A 有关的 n 阶方阵,称为矩阵 A 的多项式.

例如 设 $f(x)=x^2-(a+d)x+ad-bc$, $A=\begin{pmatrix}a&b\\c&d\end{pmatrix}$,则

$$\begin{aligned}f(A)=&A^2-(a+d)A+(ad-bc)E\\=&\begin{pmatrix}a&b\\c&d\end{pmatrix}^2-(a+d)\begin{pmatrix}a&b\\c&d\end{pmatrix}+(ad-bc)\begin{pmatrix}1&0\\0&1\end{pmatrix}\\=&\begin{pmatrix}a^2+bc&ab+db\\ac+dc&d^2+bc\end{pmatrix}-\begin{pmatrix}a^2+ad&ab+bd\\ac+cd&ad+d^2\end{pmatrix}+\begin{pmatrix}ad-bc&0\\0&ad-bc\end{pmatrix}\\=&\begin{pmatrix}0&0\\0&0\end{pmatrix}=0.\end{aligned}$$

补充例题 5.1.5

证明可逆阵 A 的特征值 $\lambda \neq 0$, 且 $\frac{1}{\lambda}, \frac{|A|}{\lambda}$ 分别是 A^{-1}, A^{*} 的特征值.

补充例题 5.1.5

证明可逆阵 A 的特征值 $\lambda \neq 0$, 且 $\frac{1}{\lambda}, \frac{|A|}{\lambda}$ 分别是 A^{-1}, A^* 的特征值.

解: 先证 $\lambda \neq 0$.

- 证法一. 因可逆阵 A 的列是线性无关的,故对任意非零 n 维列向量 X, $AX \neq 0$. 特别取 X 为 A 的属于特征值 λ 的特征向量,则知 $\lambda X = AX \neq 0$. 故 $\lambda \neq 0$.
- 证法二. 反证法. 假设 $\lambda = 0$. 取 X 为 A 的属于特征值 λ 的特征向量,则知 $AX = \lambda X = 0$. 因特征向量 $X \neq 0$, 即齐次线性方程组 $AX = 0$ 有非零解. 这与 A 可逆矛盾.

再证 $\frac{1}{\lambda}$ 是 A^{-1} 的特征值. 取 X 为 A 的属于特征值 λ 的特征向量,则有 $AX = \lambda X$, 两边同时左乘以 $\frac{1}{\lambda}A^{-1}$ 得,

$$A^{-1}X = \frac{1}{\lambda}X,$$

即 X 是 A^{-1} 的属于特征值 $\frac{1}{\lambda}$ 的特征向量.

在上式两端同时乘以 $|A|$ 即得 X 是 A^* 的属于特征值 $\frac{|A|}{\lambda}$ 的特征向量.

补充例题 5.1.6

设 A 为 n 阶方阵. 证明 A 可逆当且仅当 A 的所有特征值都不为零.

补充例题 5.1.6

设 A 为 n 阶方阵. 证明 A 可逆当且仅当 A 的所有特征值都不为零.

证法一： 设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$.

补充例题 5.1.6

设 A 为 n 阶方阵. 证明 A 可逆当且仅当 A 的所有特征值都不为零.

- **证法一：** 设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 利用结论 A 可逆当且仅当 $|A| \neq 0$ 和等式

$$|A| = \lambda_1 \lambda_2 \cdots \lambda_n.$$

补充例题 5.1.6

设 A 为 n 阶方阵. 证明 A 可逆当且仅当 A 的所有特征值都不为零.

- **证法一：** 设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 利用结论 A 可逆当且仅当 $|A| \neq 0$ 和等式

$$|A| = \lambda_1 \lambda_2 \cdots \lambda_n.$$

- **证法二：** 利用结论 A 不可逆当且仅当 $|A| = 0$, 当且仅当

$$|0E - A| = (-1)^n |A| = 0.$$

补充例题 5.1.6

设 A 为 n 阶方阵. 证明 A 可逆当且仅当 A 的所有特征值都不为零.

- **证法一：** 设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 利用结论 A 可逆当且仅当 $|A| \neq 0$ 和等式

$$|A| = \lambda_1 \lambda_2 \cdots \lambda_n.$$

- **证法二：** 利用结论 A 不可逆当且仅当 $|A| = 0$, 当且仅当

$$|0E - A| = (-1)^n |A| = 0.$$

补充例题 5.1.7

证明幂等阵(即满足 $A^2 = A$ 的方阵)的特征值只能是 0 或 1.

补充例题 5.1.6

设 A 为 n 阶方阵. 证明 A 可逆当且仅当 A 的所有特征值都不为零.

- **证法一：** 设 A 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$. 利用结论 A 可逆当且仅当 $|A| \neq 0$ 和等式

$$|A| = \lambda_1 \lambda_2 \cdots \lambda_n.$$

- **证法二：** 利用结论 A 不可逆当且仅当 $|A| = 0$, 当且仅当

$$|0E - A| = (-1)^n |A| = 0.$$

补充例题 5.1.7

证明幂等阵(即满足 $A^2 = A$ 的方阵)的特征值只能是 0 或 1.

证明： 取 X 为 A 的属于特征值 λ 的特征向量,则有

$$\lambda X = AX = A^2 X = A(AX) = A(\lambda X) = \lambda(AX) = \lambda^2 X.$$

即 $(\lambda^2 - \lambda)X = 0$, 又因 $X \neq 0$, 故 $\lambda^2 - \lambda = 0$, 由此得 $\lambda = 0$ 或 $\lambda = 1$.

补充例题 5.1.8

证明对合阵(即满足 $A^2 = E$ 的方阵)的特征值只能是 ± 1.

补充例题 5.1.8

证明对合阵(即满足 $A^2=E$ 的方阵)的特征值只能是 ± 1.

补充例题 5.1.9 (与 A 相关矩阵的特征值与特征向量)

设 n 阶方阵 A 有一特征向量 α，与其对应的特征值 λ. 则

1. α 也为 kA(k 为常数)的特征向量，与其对应的特征值 $k\lambda$;
2. α 也为 A^m(m 为正整数)的特征向量，与其对应的特征值 λ^m;
3. α 也为 $A+\mu E$(μ 为常数)的特征向量，与其对应的特征值 $\lambda+\mu$;
4. 设 $f(x)=c_m x^m+c_{m-1}x^{m-1}+\cdots+c_1 x+c_0$ 是一个关于 x 的多项式，则 α 也为矩阵多项式 $f(A)$ 的特征向量，与其对应的特征值 $f(\lambda)$;
5. λ 也为 A^T 的特征值,

补充例题 5.1.8

证明对合阵(即满足 $A^2=E$ 的方阵)的特征值只能是 ± 1.

补充例题 5.1.9 (与 A 相关矩阵的特征值与特征向量)

设 n 阶方阵 A 有一特征向量 α，与其对应的特征值 λ. 则

1. α 也为 kA(k 为常数)的特征向量，与其对应的特征值 $k\lambda$;
2. α 也为 A^m(m 为正整数)的特征向量，与其对应的特征值 λ^m;
3. α 也为 $A+\mu E$(μ 为常数)的特征向量，与其对应的特征值 $\lambda+\mu$;
4. 设 $f(x)=c_m x^m+c_{m-1}x^{m-1}+\cdots+c_1 x+c_0$ 是一个关于 x 的多项式，则 α 也为矩阵多项式 $f(A)$ 的特征向量，与其对应的特征值 $f(\lambda)$;
5. λ 也为 A^T 的特征值，$|\lambda E-A^T|=|(\lambda E-A^T)^T|=|\lambda E-A|$.

目录

本节讨论矩阵的对角化问题, 即当矩阵 A 满足什么条件时, A 相似于对角矩阵, 也就是说, 存在可逆矩阵 T 使得 $T^{-1}AT$ 是对角矩阵? 以及如何确定满足条件的可逆矩阵 T. 下面我们着手解决这些问题.
我们知道,

$$\begin{aligned} & n\text{阶矩阵 } A \text{ 相似于对角矩阵 } diag\{\lambda_1,\lambda_2,\cdots,\lambda_n\} \\ \Leftrightarrow\quad & \text{存在可逆矩阵 } T \text{ 使得 } T^{-1}AT = diag\{\lambda_1,\lambda_2,\cdots,\lambda_n\} \\ \Leftrightarrow\quad & \text{存在可逆矩阵 } T \text{ 使得 } AT = Tdiag\{\lambda_1,\lambda_2,\cdots,\lambda_n\}. \end{aligned}$$

对矩阵 T 按列向量进行分块, 即令 $T=(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 则有

$$\begin{aligned} & AT = Tdiag\{\lambda_1,\lambda_2,\cdots,\lambda_n\} \\ \Leftrightarrow\quad & (A\alpha_1,A\alpha_2,\cdots,A\alpha_n) = (\lambda_1\alpha_1,\lambda_2\alpha_2,\cdots,\lambda_n\alpha_n) \\ \Leftrightarrow\quad & A\alpha_i = \lambda_i\alpha_i,\ i=1,2,\cdots,n. \end{aligned}$$

由于 T 可逆当且仅当 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关, 因此我们就证明了

定理5.2.1

设 A 是 n 阶矩阵. 则 A 相似于对角矩阵的充分必要条件是 A 有 n 个线性无关的特征向量. 此时, 若 $\eta_1, \eta_2, \cdots, \eta_n$ 是 A 的 n 个线性无关的特征向量, 令

$$T = (\eta_1 \ \eta_2 \ \cdots \ \eta_n).$$

则 T 是可逆矩阵, 且

$$T^{-1}AT = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 分别为 A 的对应于特征向量 $\eta_1, \eta_2, \cdots, \eta_n$ 的特征值.

定理5.2.1

设 A 是 n 阶矩阵. 则 A 相似于对角矩阵的充分必要条件是 A 有 n 个线性无关的特征向量. 此时, 若 $\eta_1, \eta_2, \cdots, \eta_n$ 是 A 的 n 个线性无关的特征向量, 令

$$T = (\eta_1 \ \eta_2 \ \cdots \ \eta_n).$$

则 T 是可逆矩阵, 且

$$T^{-1}AT = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 分别为 A 的对应于特征向量 $\eta_1, \eta_2, \cdots, \eta_n$ 的特征值.

1. 如果矩阵 A 相似于对角矩阵, 那么这个对角阵的主对角线上的元素就是 A 的全部特征值.
2. n 阶矩阵 A 能否相似于对角矩阵就看 A 是否有 n 个线性无关的特征向量.

因此讨论矩阵的线性无关的特征向量是必要的. 为此, 我们有

定理5.2.2

方阵 A 的属于不同特征值的特征向量是线性无关的.

定理5.2.2

方阵 A 的属于不同特征值的特征向量是线性无关的.

证明 设 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是方阵 A 的互不相同的特征值, $\eta_1, \eta_2, \cdots, \eta_s$ 分别为它们对应的特征向量. 下面对 s 用数学归纳法证明定理结论成立.

当 $s = 1$ 时, 因为特征向量非零, 所以 η_1 线性无关.

假设结论对 $s-1$ 成立, 即 A 的属于 $s-1$ 个不同特征值的特征向量是线性无关的, 我们考察 s 个不同特征值的情况. 设有

$$k_1\eta_1 + k_2\eta_2 + \cdots + k_{s-1}\eta_{s-1} + k_s\eta_s = 0. \tag{5.1}$$

用矩阵 A 左乘上式两端并注意到 $A\eta_i = \lambda_i\eta_i$, 得

$$k_1\lambda_1\eta_1 + k_2\lambda_2\eta_2 + \cdots + k_{s-1}\lambda_{s-1}\eta_{s-1} + k_s\lambda_s\eta_s = 0. \tag{5.2}$$

用 λ_s 数乘 (5.1) 式两端得

$$k_1\lambda_s\eta_1 + k_2\lambda_s\eta_2 + \cdots + k_{s-1}\lambda_s\eta_{s-1} + k_s\lambda_s\eta_s = 0. \tag{5.3}$$

用 (5.3) 式减去 (5.2) 式得

$$k_1(\lambda_s - \lambda_1)\eta_1 + k_2(\lambda_s - \lambda_2)\eta_2 + \cdots + k_{s-1}(\lambda_s - \lambda_{s-1})\eta_{s-1} = 0.$$

用 (5.3) 式减去 (5.2) 式得

$$k_1(\lambda_s-\lambda_1)\eta_1+k_2(\lambda_s-\lambda_2)\eta_2+\cdots+k_{s-1}(\lambda_s-\lambda_{s-1})\eta_{s-1}=0.$$

根据归纳假设, $\eta_1,\eta_2,\cdots,\eta_{s-1}$ 线性无关, 因此

$$k_i(\lambda_s-\lambda_i)=0,\ i=1,2,\cdots,s-1.$$

由于 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相同, 所以 $\lambda_s-\lambda_i\neq 0$, $i=1,2,\cdots,s-1$,
故 $k_1=k_2=\cdots=k_{s-1}=0$. 带入到 (5.2) 式得 $k_s\eta_s=0$. 由 $\eta_s\neq 0$ 得 $k_s=0$.
于是 $\eta_1,\eta_2,\cdots,\eta_n$ 线性无关. 根据归纳法原理, 定理成立.

用 (5.3) 式减去 (5.2) 式得

$$k_1(\lambda_s-\lambda_1)\eta_1+k_2(\lambda_s-\lambda_2)\eta_2+\cdots+k_{s-1}(\lambda_s-\lambda_{s-1})\eta_{s-1}=0.$$

根据归纳假设, $\eta_1,\eta_2,\cdots,\eta_{s-1}$ 线性无关, 因此

$$k_i(\lambda_s-\lambda_i)=0,\ i=1,2,\cdots,s-1.$$

由于 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相同, 所以 $\lambda_s-\lambda_i\neq 0$, $i=1,2,\cdots,s-1$,
故 $k_1=k_2=\cdots=k_{s-1}=0$. 带入到 (5.2) 式得 $k_s\eta_s=0$. 由 $\eta_s\neq 0$ 得 $k_s=0$.
于是 $\eta_1,\eta_2,\cdots,\eta_n$ 线性无关. 根据归纳法原理, 定理成立.

推论5.2.3

若 n 阶方阵 A 有 n 个互不相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则 A 相似于对角矩阵 $diag\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$.

用 (5.3) 式减去 (5.2) 式得

$$k_1(\lambda_s-\lambda_1)\eta_1+k_2(\lambda_s-\lambda_2)\eta_2+\cdots+k_{s-1}(\lambda_s-\lambda_{s-1})\eta_{s-1}=0.$$

根据归纳假设, $\eta_1,\eta_2,\cdots,\eta_{s-1}$ 线性无关, 因此

$$k_i(\lambda_s-\lambda_i)=0,\ i=1,2,\cdots,s-1.$$

由于 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 互不相同, 所以 $\lambda_s-\lambda_i\neq 0, i=1,2,\cdots,s-1$, 故 $k_1=k_2=\cdots=k_{s-1}=0$. 带入到 (5.2) 式得 $k_s\eta_s=0$. 由 $\eta_s\neq 0$ 得 $k_s=0$. 于是 $\eta_1,\eta_2,\cdots,\eta_n$ 线性无关. 根据归纳法原理, 定理成立.

推论5.2.3

若 n 阶方阵 A 有 n 个互不相同的特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$, 则 A 相似于对角矩阵 $diag\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$.

证明 设 $\eta_1,\eta_2,\cdots,\eta_n$ 分别为 A 的对应于特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 的特征向量. 则由定理 5.2.2 知 $\eta_1,\eta_2,\cdots,\eta_n$ 线性无关. 据定理 5.2.1 知 A 相似于对角矩阵, 若令 $T=(\eta_1,\eta_2,\cdots,\eta_n)$, 则 T 可逆, 且有

$$T^{-1}AT=diag\{\lambda_1,\lambda_2,\cdots,\lambda_n\}.$$

例5.2.1

设

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

问 A 能否相似于对角矩阵？若能，求出使得 $T^{-1}AT$ 为对角阵的可逆矩阵 T.

例5.2.1

设

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

问 A 能否相似于对角矩阵？若能，求出使得 $T^{-1}AT$ 为对角阵的可逆矩阵 T.

解 因为

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

所以 A 有 3 个互不相同的特征值 $\lambda_1=1$, $\lambda_2=2$, $\lambda_3=3$. 据推论 5.2.3 知 A 相似于对角阵.

把 $\lambda_1 = 1$ 带入到 $(\lambda_1 E - A)X = 0$, 得齐次线性方程组

$$\begin{cases} \quad\quad -x_2 \quad\quad = 0 \\ \quad\quad -x_2 \quad\quad = 0 \\ -2x_1 + x_2 - 2x_3 = 0 \end{cases}$$

解得它得一个基础解系 $\eta_1 = (1, 0, -1)^T$. η_1 是 A 的属于特征值 1 的特征向量.
把 $\lambda_2 = 2$ 带入到 $(\lambda_2 E - A)X = 0$, 解得 A 的属于特征值 2 的一个特征向量 $\eta_2 = (1, 1, -1)^T$.
把 $\lambda_3 = 3$ 带入到 $(\lambda_3 E - A)X = 0$, 解得 A 的属于特征值 3 的一个特征向量 $\eta_3 = (0, 0, 1)^T$.
据定理 5.2.2 知 η_1, η_2, η_3 是 A 的 3 个线性无关的特征向量.

令

$$T=(\eta_2,\eta_1,\eta_3)=\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 0\\ -1 & -1 & 1\end{pmatrix}.$$

则 T 是可逆矩阵, 且 $T^{-1}AT=diag\{\lambda_2,\lambda_1,\lambda_3\}=diag\{2,1,3\}$.

令
$$T=(\eta_2,\eta_1,\eta_3)=\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 0\\ -1 & -1 & 1\end{pmatrix}.$$
则 T 是可逆矩阵, 且 $T^{-1}AT=diag\{\lambda_2,\lambda_1,\lambda_3\}=diag\{2,1,3\}$.
注意推论 5.2.3 只是一个矩阵相似于对角阵的充分条件, 而不是必要条件, 这可以由例 5.2.2 说明. 为了给出进一步的判别条件, 我们需要推广定理 5.2.2.

令

$$T=(\eta_2,\eta_1,\eta_3)=\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 0\\ -1 & -1 & 1\end{pmatrix}.$$

则 T 是可逆矩阵, 且 $T^{-1}AT=diag\{\lambda_2,\lambda_1,\lambda_3\}=diag\{2,1,3\}$.
注意推论 5.2.3 只是一个矩阵相似于对角阵的充分条件, 而不是必要条件, 这可以由例 5.2.2 说明. 为了给出进一步的判别条件, 我们需要推广定理 5.2.2.

定理5.2.4

若 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是矩阵 A 的互不相同的特征值, $\eta_{i1},\eta_{i2},\cdots,\eta_{it_i}$ 是 A 的属于特征值 λ_i 的线性无关的特征向量, $i=1,2,\cdots,s$, 则 $\eta_{11},\eta_{12},\cdots,\eta_{1t_1},\eta_{21},\cdots,\eta_{s1},\eta_{s2},\cdots,\eta_{st_s}$ 线性无关.

令

$$T=(\eta_2,\eta_1,\eta_3)=\begin{pmatrix}1 & 1 & 0\\ 1 & 0 & 0\\ -1 & -1 & 1\end{pmatrix}.$$

则 T 是可逆矩阵, 且 $T^{-1}AT=diag\{\lambda_2,\lambda_1,\lambda_3\}=diag\{2,1,3\}$.
注意推论 5.2.3 只是一个矩阵相似于对角阵的充分条件, 而不是必要条件, 这可以由例 5.2.2 说明. 为了给出进一步的判别条件, 我们需要推广定理 5.2.2.

定理5.2.4

若 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是矩阵 A 的互不相同的特征值, $\eta_{i1},\eta_{i2},\cdots,\eta_{it_i}$ 是 A 的属于特征值 λ_i 的线性无关的特征向量, $i=1,2,\cdots,s$, 则 $\eta_{11},\eta_{12},\cdots,\eta_{1t_1},\eta_{21},\cdots,\eta_{s1},\eta_{s2},\cdots,\eta_{st_s}$ 线性无关.

证明 类似于定理 5.2.2 的证明, 对 s 用数学归纳法.(略)

设 λ_0 是 n 阶矩阵 A 的特征值, $\eta_1, \eta_2, \cdots, \eta_t$ 是 A 的属于 λ_0 的线性无关的特征向量. 把 $\eta_1, \eta_2, \cdots, \eta_t$ 扩充为 C^n 的一组基 $(\eta_1, \eta_2, \cdots, \eta_t, \eta_{t+1}, \cdots, \eta_n)$, 并令

$$T = (\eta_1, \eta_2, \cdots, \eta_t, \eta_{t+1}, \cdots, \eta_n).$$

则 T 是可逆矩阵, 且有

$$T^{-1}AT = T^{-1}(\lambda_0\eta_1, \lambda_0\eta_2, \cdots, \lambda_0\eta_t, A\eta_{t+1}, \cdots, A\eta_n) = \begin{pmatrix} \lambda_0 E_t & B_1 \\ 0 & A_1 \end{pmatrix}.$$

因此

$$|\lambda E - A| = |\lambda E - T^{-1}AT| = (\lambda - \lambda_0)^t |\lambda E_{n-t} - A_1|,$$

从而 λ_0 至少是 A 的 t 重特征根.

为方便起见, 我们引入

定义5.2.5

设 λ_0 是方阵 A 的特征值. 称 λ_0 在 A 特征多项式 $|\lambda E - A|$ 中的重数为 λ_0 的代数重数; 称特征子空间 V_{λ_0} 的维数, 即线性方程组 $(\lambda_0 E - A)X = 0$ 的解空间 $N(\lambda_0 E - A)$ 的维数为 λ_0 的几何重数.

定义5.2.5

设 λ_0 是方阵 A 的特征值. 称 λ_0 在 A 特征多项式 $|\lambda E - A|$ 中的重数为 λ_0 的代数重数; 称特征子空间 V_{λ_0} 的维数, 即线性方程组 $(\lambda_0 E - A)X = 0$ 的解空间 $N(\lambda_0 E - A)$ 的维数为 λ_0 的几何重数.

显然单特征根的代数重数等于 1. 在例 5.1.3 中, -1 的代数重数等于 2, 几何重数也等于 2.
这样我们就证明了

定理5.2.6

特征值的几何重数不超过它的代数重数.

现在我们可以给出矩阵对角化的新的判定定理.

定理5.2.7

设 A 是 n 阶矩阵, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是它的全部互不相同的特征值, 它们的代数重数分别为 $n_1, n_2, \cdots, n_s$, 它们的几何重数分别为 $m_1, m_2, \cdots, m_s$. 则

$$\begin{aligned} A\text{ 相似于对角阵} \quad &\Leftrightarrow \quad m_i = n_i, i = 1, 2, \cdots, s, \\ &\Leftrightarrow \quad r(\lambda_i E - A) = n - n_i, i = 1, 2, \cdots, s. \end{aligned}$$

现在我们可以给出矩阵对角化的新的判定定理.

定理5.2.7

设 A 是 n 阶矩阵, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是它的全部互不相同的特征值, 它们的代数重数分别为 $n_1, n_2, \cdots, n_s$, 它们的几何重数分别为 $m_1, m_2, \cdots, m_s$. 则

$$\begin{aligned} A\text{ 相似于对角阵} \quad & \Leftrightarrow \quad m_i = n_i, i = 1, 2, \cdots, s, \\ & \Leftrightarrow \quad r(\lambda_i E - A) = n - n_i, i = 1, 2, \cdots, s. \end{aligned}$$

证明 注意到 $n_1 + n_2 + \cdots + n_s = n$, 根据定理 3.3.3 和 5.2.1, 我们有

$$\begin{aligned} A\text{ 相似于对角阵} \quad & \Leftrightarrow \quad A\text{ 有 } n \text{ 个线性无关的特征向量} \\ & \Leftrightarrow \quad m_1 + m_2 + \cdots + m_s = n \\ & \Leftrightarrow \quad m_i = n_i, i = 1, 2, \cdots, s, \\ & \Leftrightarrow \quad r(\lambda_i E - A) = n - n_i, i = 1, 2, \cdots, s. \end{aligned}$$

现在我们可以给出矩阵对角化的新的判定定理.

定理5.2.7

设 A 是 n 阶矩阵, $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是它的全部互不相同的特征值, 它们的代数重数分别为 $n_1, n_2, \cdots, n_s$, 它们的几何重数分别为 $m_1, m_2, \cdots, m_s$. 则

$$
\begin{aligned}
A\text{ 相似于对角阵} \quad &\Leftrightarrow \quad m_i = n_i, i = 1, 2, \cdots, s, \\
&\Leftrightarrow \quad r(\lambda_i E - A) = n - n_i, i = 1, 2, \cdots, s.
\end{aligned}
$$

证明 注意到 $n_1 + n_2 + \cdots + n_s = n$, 根据定理 3.3.3 和 5.2.1, 我们有

$$
\begin{aligned}
A\text{ 相似于对角阵} \quad &\Leftrightarrow \quad A\text{ 有 } n \text{ 个线性无关的特征向量} \\
&\Leftrightarrow \quad m_1 + m_2 + \cdots + m_s = n \\
&\Leftrightarrow \quad m_i = n_i, i = 1, 2, \cdots, s, \\
&\Leftrightarrow \quad r(\lambda_i E - A) = n - n_i, i = 1, 2, \cdots, s.
\end{aligned}
$$

☞ 定理 5.2.7 表明, 如果矩阵 A 的所有特征值的几何重数等于代数重数, 那么 A 相似于对角阵. 否则 A 不相似于对角阵.

例5.2.2

判断矩阵

$$(1)\quad A=\begin{pmatrix}2 & -1 & 2\\ 5 & -3 & 3\\ -1 & 0 & -2\end{pmatrix},\qquad (2)\quad A=\begin{pmatrix}1 & -1 & 1\\ 2 & 4 & -2\\ -3 & -3 & 5\end{pmatrix}$$

能否对角化？如能对角化，求出相似变换矩阵 T 使得 $T^{-1}AT$ 为对角阵.

例5.2.2

判断矩阵

$$(1)\quad A=\begin{pmatrix}2&-1&2\\5&-3&3\\-1&0&-2\end{pmatrix},\qquad (2)\quad A=\begin{pmatrix}1&-1&1\\2&4&-2\\-3&-3&5\end{pmatrix}$$

能否对角化？如能对角化，求出相似变换矩阵 T 使得 $T^{-1}AT$ 为对角阵.

解 (1) 由

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

得 -1 是 A 的三重特征值. 因为

$$r(-E-A)=r\begin{pmatrix}-3&1&-2\\-5&2&-3\\1&0&1\end{pmatrix}=2,$$

所以 $3-r(-E-A)=1<3$, 据定理 5.2.7 知 A 不能相似于对角阵.

(2) 由

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

解得 A 的特征值 $\lambda_1 = \lambda_2 = 2$ (二重), $\lambda_3 = 6$.
解齐次线性方程组

$$(2E - A)X = \begin{pmatrix} 1 & 1 & -1 \\ -2 & -2 & 2 \\ 3 & 3 & -3 \end{pmatrix} X = 0$$

得 A 属于特征值 2 的线性无关的特征向量为

$$\eta_1 = (1, -1, 0)^T, \qquad \eta_2 = (1, 0, 1)^T.$$

解齐次线性方程组

$$(6E-A)X=\begin{pmatrix} 5 & 1 & -1 \\ -2 & 2 & 2 \\ 3 & 3 & 1 \end{pmatrix} X=0$$

得 A 属于特征值 6 的线性无关的特征向量为

$$\eta_3=(1,-2,3)^T.$$

综上可知矩阵 A 的所有特征值的几何重数等于代数重数, 因而 A 相似于对角阵. 令

$$T=(\eta_1,\eta_2,\eta_3)=\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & -2 \\ 0 & 1 & 3 \end{pmatrix}.$$

则 $T^{-1}AT=diag\{2,2,6\}$.

目录

上一节我们已经看到并不是所有的矩阵都能与对角矩阵相似. 本节我们讨论一类能与对角矩阵相似的矩阵, 即实对称矩阵.
我们首先介绍实对称矩阵的特征值和特征向量的一些基本性质.

定理5.3.1

实对称矩阵的特征值都是实数.

上一节我们已经看到并不是所有的矩阵都能与对角矩阵相似. 本节我们讨论一类能与对角矩阵相似的矩阵, 即实对称矩阵.

我们首先介绍实对称矩阵的特征值和特征向量的一些基本性质.

定理5.3.1

实对称矩阵的特征值都是实数.

据定理 5.3.1 知, n 阶实对称矩阵在实数域内有 n 个特征值(重特征值按重数计算). 而 A 的属于特征值 λ 的特征向量就是实系数齐次线性方程组

$$(\lambda E - A)X = 0$$

的非零解解, 方程组存在实向量构成的基础解系, 它构成了 A 的属于特征值 λ 的所有特征向量组的一个极大线性无关组. 因此我们约定, 实对称矩阵的特征向量都是欧氏空间 R^n 中的向量.

证明 设 λ 是 n 阶实对称矩阵 A 的任意一个特征值, $\eta=(x_1,x_2,\cdots,x_n)^T$ 是 A 的属于 λ 的一个特征向量. 则有

$$A\eta=\lambda\eta. \tag{5.1}$$

令

$$\overline{\eta}=(\overline{x}_1,\overline{x}_2,\cdots,\overline{x}_n)^T,$$

其中$\overline{x}_i$, $i=1,2,\cdots,n$, 表示 x_i 的共轭复数. 则

$$\overline{\eta}^T\eta=\overline{x}_1x_1+\overline{x}_2x_2+\cdots+\overline{x}_nx_n\neq 0.$$

注意到 A 是实对称矩阵. 据 (5.1) 式得

$$\lambda\eta^T=(\lambda\eta)^T=(A\eta)^T=\eta^TA^T=\eta^TA,\quad A\overline{\eta}=\overline{A}\overline{\eta}=\overline{A\eta}=\overline{\lambda\eta}=\overline{\lambda}\overline{\eta}.$$

于是

$$\lambda\eta^T\overline{\eta}=(\lambda\eta^T)\overline{\eta}=\eta^TA\overline{\eta}=\eta^T(A\overline{\eta})=\eta^T(\overline{\lambda}\overline{\eta})=\overline{\lambda}\eta^T\overline{\eta}.$$

因为 $\eta^T\overline{\eta}\neq 0$, 所以 $\lambda=\overline{\lambda}$, 故 λ 是实数.

定理5.3.2

设 A 是实对称矩阵. 则 A 的属于不同特征值的特征向量是正交的.

定理5.3.2

设 A 是实对称矩阵. 则 A 的属于不同特征值的特征向量是正交的.

证明 设 η_1, η_2 分别是 A 的属于不同特征值 λ_1, λ_2 的特征向量. 则有

$$A\eta_1 = \lambda_1\eta_1, \quad A\eta_2 = \lambda_2\eta_2.$$

于是在 R^n 中,

$$\begin{aligned}\lambda_1(\eta_1,\eta_2) &= (\lambda_1\eta_1,\eta_2) = (A\eta_1,\eta_2) = (A\eta_1)^T\eta_2 = \eta_1^T A^T\eta_2 \\ &= \eta_1^T A\eta_2 = (\eta_1, A\eta_2) = (\eta_1,\lambda_2\eta_2) = \lambda_2(\eta_1,\eta_2),\end{aligned}$$

因此有

$$(\lambda_1-\lambda_2)(\eta_1,\eta_2) = 0.$$

因为$\lambda_1 \neq \lambda_2$, 所以

$$(\eta_1,\eta_2) = 0,$$

即η_1 与 η_2 正交.

现在可以证明本节的主要定理.

定理5.3.3

对于任意 n 阶实对称矩阵 A, 都存在 n 正交矩阵 T, 使得

$$T^{-1}AT = T'AT = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 的全部特征值.

现在可以证明本节的主要定理.

定理5.3.3

对于任意 n 阶实对称矩阵 A, 都存在 n 正交矩阵 T, 使得

$$T^{-1}AT = T'AT = diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 为 A 的全部特征值.

证明 对矩阵的阶数 n 进行归纳证明.
当 $n = 1$ 时, 结论显然成立.
假设 $n-1$ 时定理结论成立. 现考察 n 时的情况. 设 λ_1 是 n 阶实对称矩阵 A 的一个特征值, η_1 是 A 的属于 λ_1 的一个特征向量. 由于 $(1/|\eta_1|)\eta_1$ 也是 A 的属于 λ_1 的特征向量, 故不妨设 η_1 是单位向量. 把 η_1 扩充为 R^n 的一组标准正交基

$$(\eta_1, \eta_2, \cdots, \eta_n).$$

令 $T_1 = (\eta_1, \eta_2, \cdots, \eta_n)$. 则 T_1 是正交矩阵, 且有

$$T_1^{-1}AT_1 = T_1'AT_1 = \begin{pmatrix} \lambda_1 & C \\ 0 & B \end{pmatrix}, \tag{5.2}$$

其中 B 是一个 $n-1$ 阶矩阵, C 是一个 $1\times(n-1)$ 矩阵. 因为 T_1 是正交矩阵, A 是实对称矩阵, 所以矩阵 $T_1^{-1}AT_1$, 即

$$\begin{pmatrix} \lambda_1 & C \\ 0 & B \end{pmatrix}$$

是实对称矩阵. 于是 $C=0$, B 为 $n-1$ 阶实对称矩阵.
根据归纳法假设,存在 $n-1$ 阶正交矩阵 S, 使得

$$S^{-1}BS = S'BS = diag\{\lambda_2, \cdots, \lambda_n\},$$

其中 $\lambda_2, \cdots .\lambda_n$ 为 B 的全部特征值.

令
$$T_2 = \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}.$$
则 T_2 是 n 阶正交矩阵, 从而 $T = T_1 T_2$ 也是 n 阶正交矩阵, 且有
$$\begin{aligned} T^{-1}AT &= T_2^{-1}(T_1^{-1}AT_1)T_2 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & S^{-1} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & S^{-1}BS \end{pmatrix} \\ &= diag\{\lambda_1, \lambda_2, \cdots, \lambda_n\}. \end{aligned}$$
据 (5.2) 式知 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 就是 A 的全部特征值. 根据归纳法原理, 定理得证.

正交矩阵 T 的求法可以按如下步骤进行:
(1) 求出实对称矩阵 A 的全部互异的特征值 $\lambda_1, \lambda_2, \cdots, \lambda_s$.
(2) 对每个特征值 λ_i, $i = 1, 2, \cdots, s$, 求得 V_{λ_i} 的一组基, 即齐次线性方程组

$$(\lambda_i E - A)X = 0$$

的一个基础解系. 对这组基施行施密特正交化得到 V_{λ_i} 的一组标准正交基

$$(\eta_{i1}, \eta_{i2}, \cdots, \eta_{ir_i}).$$

(3) 据定理 5.2.1, 5.3.2 及 5.3.3 知

$$(\eta_{11}, \eta_{12}, \cdots, \eta_{1r_1}, \cdots, \eta_{s1}, \eta_{s2}, \cdots, \eta_{sr_s}),$$

就构成 R^n 的一组标准正交基, 并且都是 A 的特征向量. 令

$$T = (\eta_{11}, \eta_{12}, \cdots, \eta_{1r_1}, \cdots, \eta_{s1}, \eta_{s2}, \cdots, \eta_{sr_s}).$$

则 T 是 n 阶正交矩阵, 且有

$$T^{-1}AT = diag\{\overbrace{\lambda_1, \cdots, \lambda_1}^{r_1}, \cdots, \overbrace{\lambda_s, \cdots, \lambda_s}^{r_s}\}.$$

例5.3.1

设

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

求正交矩阵 T, 使得 $T^{-1}AT = T'AT$ 为对角矩阵.

例5.3.1

设
$$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}.$$
求正交矩阵 T, 使得 $T^{-1}AT=T'AT$ 为对角矩阵.

解 据例 5.1.3 知, A 的特征值为 -1 (二重), 5.
V_{-1} 的一组基为
$$(\xi_1=(1,0,-1)^T,\quad \xi_2=(0,1,-1)^T).$$
用施密特正交化方法把 (ξ_1,ξ_2) 化为标准正交基:
$$\xi_1'=\xi_1=(1,0,-1)^T,$$
$$\xi_2'=\xi_2-\frac{(\xi_2,\xi_1')}{(\xi_1',\xi_1')}\xi_1'=(0,1,-1)^T-\frac{1}{2}(1,0,-1)^T=\frac{1}{2}(-1,2,-1)^T.$$

把 ξ_1', ξ_2' 单位化得

$$\eta_1 = \frac{1}{\sqrt{2}}(1, 0, -1)^T, \quad \eta_2 = \frac{1}{\sqrt{6}}(-1, 2, -1)^T.$$

V_5 的一组基为 $(\xi_3 = (1, 1, 1)^T)$, 单位化得

$$(\eta_3 = \frac{1}{\sqrt{3}}(1, 1, 1)^T).$$

则 (η_1, η_2, η_3) 就是 R^3 的一组标准正交基. 令

$$T = (\eta_1\ \eta_2\ \eta_3) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$

则 T 是正交矩阵, 且 $T^{-1}AT = T'AT = diag\{-1, -1, 5\}$.