线性代数与解析几何

第五章 特征值与特征向量

- 特征值(eigenvalue) 和特征向量(eigenvector) 是线性代数的主要研究对象之一,
- 它综合了矩阵、行列式和方程组等诸多知识点, 与第七章将要讲的线性变换关系甚为密切.
- 作为工程计算中不可或缺的组成部分, 特征值和特征向量被广泛用于物理学中的各种临界值问题、机械及电磁振动中的固有值问题等等.
- 本章将对特征值及特征向量进行深入的讨论, 并由此探讨相似意义下矩阵的可对角化问题.

§ 5.1 矩阵的特征值和特征向量

若 $A=\begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix}$, $\alpha=\begin{pmatrix}0\\1\end{pmatrix}$, $\beta=\begin{pmatrix}2\\1\end{pmatrix}$, $\gamma=\begin{pmatrix}1\\1\end{pmatrix}$, 则有

$$A\alpha=\begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}0\\-1\end{pmatrix}=-\alpha,$$

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

$$A\gamma=\begin{pmatrix}1 & 0\\ 1 & -1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}\neq k\gamma,\quad \text{对任意的}k.$$

定义 1.1

设 A 是一个 n 阶矩阵, λ 是一个数. 如果存在非零列向量α(即 $n \times 1$ 矩阵) 使得

$$A\alpha = \lambda\alpha,$$

就称 λ 是 A 的一个特征值, α 是 A 的对应特征值 λ 的特征向量, 简称特征向量.

- 在上面的例子中, -1和1就是矩阵 A 的两个特征值, α 和 β 分别是对应于它们的特征向量; 而 γ 则是一个不属于A 的任意一个特征值的向量.

设 $A=(a_{ij})_{n\times n}$, $\alpha=\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}\neq 0$, 且 $A\alpha=\lambda\alpha$, 则 $(A-\lambda E)\alpha=0$,

即 $(\lambda E-A)\alpha=0$. 于是 α 是齐次线性方程组

$$\begin{pmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=0.$$

的一组非零解. 故由齐次线性方程组的解的理论知道, 齐次方程组的系数行列式等于零. 于是有

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0,$$

即

$$|\lambda E - A| = 0.$$

定义 1.2

设 $A = (a_{ij})_{n\times n}$, 我们称关于 λ 的一元 n 次方程 $f(\lambda) = |\lambda E - A| = 0$ 为矩阵 A 的特征方程或特征多项式; 称 $\lambda E - A$ 为矩阵 A 的特征矩阵.

- 从定义我们可以看出, A 的特征方程或特征多项式的根就是 A 的特征值, 所以特征值又叫作特征根.
- 若在复数域$\mathbb{C}$ 中考虑问题, 由代数基本定理知, 在计算根的重数的前提下, A 的特征值的总个数为 n.

这样我们就找到了寻求矩阵特征值和特征向量的方法:

- (1) 求出矩阵 A 的特征方程$|\lambda E - A| = 0$;
- (2) 求出特征方程的所有根 (计重), 就是 A 的全部特征值;
- (3) 对于每一个特征值 λ_i, 求出线性方程组 $(\lambda E - A)X = 0$ 的非零解, 就是对应于 λ_i 的全部特征向量.

例题 1.1

求矩阵 A 的特征值和特征向量, 这里

$$A = \begin{pmatrix} -3 & 4 & 0 \\ -1 & 1 & 0 \\ 7 & 5 & 1 \end{pmatrix}.$$

例题 1.1

求矩阵 A 的特征值和特征向量, 这里

$$A = \begin{pmatrix} -3 & 4 & 0 \\ -1 & 1 & 0 \\ 7 & 5 & 1 \end{pmatrix}.$$

解 A 的特征方程为:

$$|\lambda E - A| = \begin{vmatrix} \lambda+3 & -4 & 0 \\ 1 & \lambda-1 & 0 \\ -7 & -5 & \lambda-1 \end{vmatrix} = (\lambda-1)(\lambda+1)^2 = 0.$$

所以 A 的特征值为 1, -1(二重).

将 $\lambda=1$ 代入齐次线性方程组 $(\lambda E-A)X=0$ 中, 得

$$\begin{cases} 4x_1-4x_2=0, \\ x_1=0, \\ -7x_1-5x_2=0. \end{cases}$$

它的一个基础解系是

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

将 $\lambda=-1$ 代入齐次线性方程组 $(\lambda E-A)X=0$ 中, 得

$$\begin{cases} 2x_1-4x_2=0, \\ x_1-2x_2=0, \\ -7x_1-5x_2-2x_3=0. \end{cases}$$

它的一个基础解系是

$$\begin{pmatrix} 4 \\ 2 \\ -19 \end{pmatrix}.$$

因此, A 的特征值为 1,-1. 对应于 1 的特征向量是

$$k_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad (k_1 \neq 0);$$

对应于 -1 的特征向量是

$$k_2 \begin{pmatrix} 4 \\ 2 \\ -19 \end{pmatrix} \quad (k_2 \neq 0).$$

例题 1.2

平面上旋转变换 T_θ 在单位向量组成的基 e_1, e_2 下的矩阵为

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

此矩阵在复数域内肯定有特征值，试问在实数范围内，此矩阵在 θ 取何值时有特征值？求出此时对应于特征值的特征向量.

例题 1.2

平面上旋转变换 T_θ 在单位向量组成的基 e_1, e_2 下的矩阵为

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

此矩阵在复数域内肯定有特征值，试问在实数范围内，此矩阵在 θ 取何值时有特征值？求出此时对应于特征值的特征向量.

解 A 的特征方程为：

$$|\lambda E - A| = \begin{vmatrix} \lambda - \cos\theta & \sin\theta \\ -\sin\theta & \lambda - \cos\theta \end{vmatrix} = \lambda^2 - 2\lambda\cos\theta + 1 = 0.$$

- 因为当 $\theta \neq k\pi$ 时, $\cos\theta \neq \pm 1$, 判别式 $\Delta = 4\cos^2\theta - 4 < 0$, 方程在实数范围内无解. 于是此矩阵要想在实数范围内存在特征值, θ 只能取 $k\pi$. 此时, 特征方程为

$$|\lambda E - A| = \lambda^2 \pm 2\lambda + 1 = (\lambda \pm 1)^2 = 0.$$

- 当 $\theta = 2k\pi$ 时, 特征值为 1(二重). 将 $\lambda = 1$ 代入齐次线性方程组 $(\lambda E - A)X = 0$ 中, 求得其基础解系为

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

对应于 1 的特征向量是

$$k_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (k_1, k_2 \text{ 不全为零 }).$$

- 当 $\theta = (2k+1)\pi$ 时，特征值为 -1(二重). 将 $\lambda = -1$ 代入齐次线性方程组 $(\lambda E - A)X = 0$ 中，求得其基础解系仍为

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

对应于 1 的特征向量是

$$k_3 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + k_4 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (k_3, k_4 \text{ 不全为零 }).$$

矩阵的特征值和特征向量有很多优良的性质:

性质 1.1

n 阶矩阵 A 与其转置 A^T 有相同的特征值.

矩阵的特征值和特征向量有很多优良的性质:

性质 1.1

n 阶矩阵 A 与其转置 A^T 有相同的特征值.

证明 因为 $(\lambda E - A)^T = \lambda E - A^T$, 又 $|\lambda E - A| = |(\lambda E - A)^T| = |\lambda E - A^T|$. 因此 A 与 A^T 有相同的特征值多项式, 所以有相同的特征值.

矩阵的特征值和特征向量有很多优良的性质:

性质 1.1

n 阶矩阵 A 与其转置 A^T 有相同的特征值.

证明 因为 $(\lambda E - A)^T = \lambda E - A^T$, 又 $|\lambda E - A| = |(\lambda E - A)^T| = |\lambda E - A^T|$. 因此 A 与 A^T 有相同的特征值多项式, 所以有相同的特征值.

性质 1.2

设 n 阶矩阵 $A = (a_{ij})_{n\times n}$ 的 n 个特征值为 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 则有

(1) A 的 n 个特征值之和等于矩阵 A 的迹, 即

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}.$$

(2) A 的 n 个特征值之积等于矩阵 A 的行列式的值, 即

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

证明

首先由条件, 矩阵 A 的特征多项式

$$f(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n). \tag{1}$$

其次, 利用行列式的定义将矩阵 A 的特征多项式

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

展开, 含 λ^n, λ^{n-1} 的项只能出现在展开式的 $(\lambda - a_{11})(\lambda - a_{22})\cdots(\lambda - a_{nn})$ 这一项中, 从而 $f(\lambda)$ 的最高项的系数为 1, $n-1$ 次项的系数为 $-(a_{11} + a_{22} + \cdots + a_{nn})$. 在特征多项式 $|\lambda E - A|$ 中令 $\lambda = 0$ 得常数项 $|-A| = (-1)^n|A|$. 这样若只写出其前两项和常数项, 特征多项式又可以写成

$$f(\lambda) = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \cdots + (-1)^n|A|. \qquad (2)$$

比较(1)和(2)两边的 λ^{n-1} 的系数和常数项 (或由方程的根与系数的关系)，即得

$$\lambda_1 + \lambda_2 + \cdots + \lambda_n = a_{11} + a_{22} + \cdots + a_{nn}, \quad \lambda_1\lambda_2\cdots\lambda_n = |A|.$$

例题 1.3

设 λ 是 n 阶矩阵 A 的一个特征值, α 是 A 的对应于特征值 λ 的一个特征向量. 证明 λ^2 是 A^2 的一个特征值, α 也是 A^2 的对应于特征值 λ^2 的一个特征向量.

例题 1.3

设 λ 是 n 阶矩阵 A 的一个特征值, α 是 A 的对应于特征值 λ 的一个特征向量. 证明 λ^2 是 A^2 的一个特征值, α 也是 A^2 的对应于特征值 λ^2 的一个特征向量.

证明　由条件知 $A\alpha = \lambda\alpha$, 且 $\alpha \neq 0$. 于是

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

所以 λ^2 是 A^2 的一个特征值, α 也是 A^2 对应于特征值 λ^2 的一个特征向量.

此例题不难推广到一般的情形: 若 λ 是 n 阶矩阵 A 的一个特征值, α 是 A 的对应于特征值 λ 的一个特征向量, 对任意的正整数 k, λ^k 是 A^k 的特征值, 且 α 也是 A^k 的对应于特征值 λ^k 的一个特征向量.

更进一步地, 设 $f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0$ 为 m 次多项式, 则矩阵 A 的多项式为

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

$$\begin{aligned} f(A)\alpha &= a_mA^m\alpha+a_{m-1}A^{m-1}\alpha+\cdots+a_1A\alpha+a_0E\alpha \\ &= a_m\lambda^m\alpha+a_{m-1}\lambda^{m-1}\alpha+\cdots+a_1\lambda\alpha+a_0\alpha \\ &= f(\lambda)\alpha, \end{aligned}$$

所以 $f(\lambda)=a_m\lambda^m+a_{m-1}\lambda^{m-1}+\cdots+a_1\lambda+a_0$ 是 $f(A)$ 的一个特征值, 且 α 也是 $f(A)$ 的对应于特征值 $f(\lambda)$ 的一个特征向量.特别地, 当 A 是可逆矩阵时, 一定有 $\lambda\neq 0$, 而且由 $A\alpha=\lambda\alpha$ 得 $A^{-1}A\alpha=A^{-1}\lambda\alpha$, 即 $\alpha=\lambda A^{-1}\alpha$, 故有 $A^{-1}\alpha=\frac{1}{\lambda}\alpha$. 所以 $\frac{1}{\lambda}$ 是 A^{-1} 的特征值, 且 α 也是 A^{-1} 的对应于特征值 $\frac{1}{\lambda}$ 的一个特征向量.

性质 1.3

设 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是 n 阶矩阵 A 的两两不同的特征值, $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ir_i}$ 是对应于 λ_i 的一些线性无关的特征向量$(i = 1, 2, \cdots, t)$. 那么 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}, \alpha_{21}, \cdots, \alpha_{t1}, \cdots, \alpha_{tr_t}$ 也是线性无关的.

性质 1.3

设 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是 n 阶矩阵 A 的两两不同的特征值, $\alpha_{i1}, \alpha_{i2}, \cdots, \alpha_{ir_i}$ 是对应于 λ_i 的一些线性无关的特征向量$(i = 1, 2, \cdots, t)$. 那么 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}, \alpha_{21}, \cdots, \alpha_{t1}, \cdots, \alpha_{tr_t}$ 也是线性无关的.

证明 对特征值的个数 t 用数学归纳法.

当 $t = 1$ 时, 由假设知 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}$ 是线性无关的.

假设定理对 $t = k$ 时成立, 即 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}, \alpha_{21}, \cdots, \alpha_{k1}, \cdots, \alpha_{kr_k}$ 是线性无关的, 下证当 $t = k + 1$ 时结论同样成立. 设

$$\sum_{j=1}^{r_1} a_{1j}\alpha_{1j} + \cdots + \sum_{j=1}^{r_k} a_{kj}\alpha_{kj} + \sum_{j=1}^{r_{k+1}} a_{k+1,j}\alpha_{k+1,j} = 0. \tag{3}$$

在等式两边同乘以 λ_{k+1}, 得

$$\sum_{j=1}^{r_1} a_{1j}\lambda_{k+1}\alpha_{1j} + \cdots + \sum_{j=1}^{r_k} a_{kj}\lambda_{k+1}\alpha_{kj} + \sum_{j=1}^{r_{k+1}} a_{k+1,j}\lambda_{k+1}\alpha_{k+1,j} = 0.$$

在等式 (1.3) 的两端同时用矩阵 A 左乘, 得

$$\sum_{j=1}^{r_1} a_{1j}\lambda_1\alpha_{1j} + \cdots + \sum_{j=1}^{r_k} a_{kj}\lambda_k\alpha_{kj} + \sum_{j=1}^{r_{k+1}} a_{k+1,j}\lambda_{k+1}\alpha_{k+1,j} = 0.$$

新得到的两个等式相减得

$$\sum_{j=1}^{r_1} a_{1j}(\lambda_{k+1} - \lambda_1)\alpha_{1j} + \cdots + \sum_{j=1}^{r_k} a_{kj}(\lambda_{k+1} - \lambda_k)\alpha_{kj} = 0.$$

由归纳假设知 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}, \alpha_{21}, \cdots, \alpha_{k1}, \cdots, \alpha_{kr_k}$ 是线性无关的, 于是

$$a_{ij}(\lambda_{k+1} - \lambda_i) = 0 \quad (i = 1, 2, \cdots, k; j = 1, 2, \cdots, r_i).$$

由题意知 $\lambda_1, \lambda_2, \cdots, \lambda_t$ 是两两不同的特征值, 从而有

$$a_{ij} = 0 \quad (i = 1, 2, \cdots, k; j = 1, 2, \cdots, r_i).$$

将其代入 (1.3) 式有

$$\sum_{j=1}^{r_{k+1}} a_{k+1,j}\alpha_{k+1,j} = 0.$$

又因为 $\alpha_{k1}, \alpha_{k2}, \cdots, \alpha_{kr_k}$ 是对应于 λ_k 的一些线性无关的特征向量,

$$a_{k+1,j} = 0 \quad (j = 1, 2, \cdots, r_{k+1}).$$

这样就得到了 $\alpha_{11}, \alpha_{12}, \cdots, \alpha_{1r_1}, \alpha_{21}, \cdots, \alpha_{t1}, \cdots, \alpha_{tr_t}$ 也是线性无关的.

下面的一个重要的性质是性质 1.3 的特殊情形.

性质 1.4

对应于不同特征值的特征向量线性无关.

本章作业 P149

1(2)(4); 2(2); 3; 9(2)(4); 14(1)(3); 15; 17; 19; 21; 23