问题: 如何在 $\mathbb{C}$ 上求一个 n 阶方阵 $\boldsymbol{A}$ 的特征值和相应的特征向量?

(1) 首先, 计算矩阵 $\boldsymbol{A}$ 的特征多项式 $p_{\boldsymbol{A}}(\lambda)$. 由**代数学基本定理**可知, $p_{\boldsymbol{A}}(\lambda)$ 恰好有 n 个复根 (可能有重根), 从而可以写成

$$p_{\boldsymbol{A}}(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}\cdots(\lambda-\lambda_s)^{n_s},$$

其中 $\lambda_1,\ldots,\lambda_s \in \mathbb{C}$ 是 $\boldsymbol{A}$ 的全部不同的特征值, 相应的重数 $n_i \geq 1$ 被称为 λ_i 的**代数重数**, 满足

$$n_1 + \cdots + n_s = n.$$

(2) 对于每个特征值 λ_i, 求解齐次方程组 $(\lambda_i \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$. 相应的解空间 $V_{\boldsymbol{A}}(\lambda_i)$ 不是零空间, 不妨设 $\boldsymbol{x}_{i1}, \boldsymbol{x}_{i2}, \ldots, \boldsymbol{x}_{im_i}$ 是一个基础解系 (m_i 被称为 λ_i 的**几何重数**). 则 $\boldsymbol{A}$ 的属于 λ_i 的所有特征向量为 $V_{\boldsymbol{A}}(\lambda_i)$ 中的所有非零向量, 即 $\boldsymbol{x}_{i1},\ldots,\boldsymbol{x}_{im_i}$ 的非零线性组合的全体.

例 6.3.10. 求矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ 的特征值及特征向量.

证明. (1) 求解特征多项式

$$|\lambda \boldsymbol{I} - \boldsymbol{A}| = \begin{vmatrix} \lambda-1 & -2 & -2 \\ -2 & \lambda-1 & -2 \\ -2 & -2 & \lambda-1 \end{vmatrix} = \lambda^3 - 3\lambda^2 - 9\lambda - 5 = (\lambda-5)(\lambda+1)^2.$$

正确计算简单的三阶方阵的特征多项式, 并将其分解因式, 是非常重要的

故矩阵 $\boldsymbol{A}$ 的特征值为 $\lambda_1 = 5$ 和 $\lambda_2 = -1$ (两重).

(2) 对于 $\lambda_1 = 5$,

$$\lambda_1 \boldsymbol{I} - \boldsymbol{A} = \begin{pmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{pmatrix} \xrightarrow{\text{初等行变换}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

解相应的方程组, 我们可以以 x_3 为自由元, 解得 $x_1 = x_2 = x_3$. 这说明解空间有基础解系 $(1,1,1)^{\mathsf{T}}$.

(3) 对于 $\lambda_2 = -1$,

$$\lambda_2 - \boldsymbol{A} = \begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix} \xrightarrow{\text{初等行变换}} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

我们以 x_2, x_3 为自由元, 解得 $x_1 = -x_2 - x_3$. 即

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 - x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

从而解空间有基础解系 $(-1,1,0)^{\mathsf{T}}, (-1,0,1)^{\mathsf{T}}$.

综上, $\boldsymbol{A}$ 有两个特征值 $\lambda_1 = 5$ 和 $\lambda_2 = -1$. 属于 $\lambda_1 = 5$ 的特征向量为 $c_1(1,1,1)^{\mathsf{T}}$, 其中 $c_1 \neq 0$. 属于 $\lambda_2 = -1$ 的特征向量为 $c_2(-1,1,0)^{\mathsf{T}} + c_3(-1,0,1)^{\mathsf{T}}$, 其中 c_2, c_3 不全为 0. □

例 6.3.11. 设 3 阶矩阵 $\boldsymbol{A}$ 的 3 个特征值为 $1,1,2$, 对应地分别有特征向量

$$\boldsymbol{\xi}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad \boldsymbol{\xi}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{\xi}_3 = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

求矩阵 $\boldsymbol{A}$.

解. 由定义, $\boldsymbol{A}\boldsymbol{\xi}_1 = \boldsymbol{\xi}_1$, $\boldsymbol{A}\boldsymbol{\xi}_2 = \boldsymbol{\xi}_2$, $\boldsymbol{A}\boldsymbol{\xi}_3 = 2\boldsymbol{\xi}_3$. 将这些由矩阵表示, 我们得到

$$\boldsymbol{A}\underbrace{(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3)}_{\boldsymbol{T}} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3)\underbrace{\begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix}}_{\boldsymbol{D}},$$

可以验证, 我们有 $\det(\boldsymbol{T}) \neq 0$, 这说明 $\boldsymbol{T}$ 可逆, 从而有

$$\begin{aligned} \boldsymbol{A} &= \boldsymbol{T}\boldsymbol{D}\boldsymbol{T}^{-1} \\ &= \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 2 \\ -2 & 3 & -4 \\ 1 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -2 & 2 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}. \end{aligned}$$

□

注 6.3.12. 由多项式的理论可知, 实多项式的虚根都是成对出现的. 这说明实方阵的虚复特征值也是成对出现的.

事实上, 若 λ 是某个 n 阶实方阵 $\boldsymbol{A}$ 的复特征值, $\boldsymbol{x} \in \mathbb{C}^n$ 是对应的一个特征向量, 则

$$\boldsymbol{A}\overline{\boldsymbol{x}} = \overline{\boldsymbol{A}}\overline{\boldsymbol{x}} = \overline{\boldsymbol{A}\boldsymbol{x}} = \overline{\lambda\boldsymbol{x}} = \overline{\lambda}\overline{\boldsymbol{x}}.$$

由于 $\overline{\boldsymbol{x}} \neq \boldsymbol{0}$, 这说明 $\overline{\boldsymbol{x}}$ 是实矩阵 $\boldsymbol{A}$ 关于特征值 $\overline{\lambda}$ 的特征向量. 特别地, 这说明实方阵的复特征向量也是依照复共轭运算成对出现的. (当然, 这儿需要简要说明一下, 当 $\lambda \neq \overline{\lambda}$, 则必有 $\boldsymbol{x} \neq \overline{\boldsymbol{x}}$, 从而 $\boldsymbol{x}$ 不是实向量. 为此, 我们用反证法. 假设 $\boldsymbol{x} = \overline{\boldsymbol{x}}$, 则我们会得到

$$\overline{\lambda}\boldsymbol{x} = \overline{\lambda}\overline{\boldsymbol{x}} = \overline{\boldsymbol{A}\boldsymbol{x}} \xlongequal[\text{从而 } \boldsymbol{A}\boldsymbol{x} \text{ 也是实的}]{\boldsymbol{A}, \boldsymbol{x} \text{ 都是实的}} \boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x},$$

这说明 $(\lambda - \overline{\lambda})\boldsymbol{x} = \boldsymbol{0}$. 但是 $\lambda \neq \overline{\lambda}$, 从而可以推出 $\boldsymbol{x} = \boldsymbol{0}$. 而这与 $\boldsymbol{x}$ 为特征值相矛盾.)

另一方面, 如果实矩阵 $\boldsymbol{A}$ 的特征值是实数, 那么为了讨论的方便, 作为解空间的基本解系的特征向量我们一般可以取为实向量.

若 $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathsf{T}} \in \mathbb{C}^n$, 则

$$\|\boldsymbol{x}\| := \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2} = \sqrt{\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{x}}$$

称为 $\boldsymbol{x}$ 在标准内积下的**模长**. 显然, $\boldsymbol{x} = \boldsymbol{0}$ 当且仅当 $\|\boldsymbol{x}\| = 0$.

例 6.3.13. 设 λ 为 n 阶方阵 $\boldsymbol{A}$ 的一个特征值, $\boldsymbol{x}$ 为对应的一个特征向量, 则 教材例题 6.3.3

(1) λ^k 是 $\boldsymbol{A}^k$ 的特征值, 其中 k 为正整数, 并且更一般地, 若 $f(t)$ 是一个一元多项式, 则 $f(\lambda)$ 是 $f(\boldsymbol{A})$ 的特征值;

(2) λ 为 $\boldsymbol{A}^{\mathsf{T}}$ 的特征值;

(3) 若 $\lambda \neq 0$, 则 $\frac{\det(\boldsymbol{A})}{\lambda}$ 为 $\boldsymbol{A}$ 的伴随矩阵 $\boldsymbol{A}^*$ 的特征值;

(4) 若 $\boldsymbol{A}$ 为实方阵且满足 $\boldsymbol{A}\boldsymbol{A}^{\mathsf{T}} = \boldsymbol{I}$ (即 $\boldsymbol{A}$ 为正交矩阵), 则 $|\lambda| = 1$. (此时 $\lambda \in \mathbb{C}$)

证明. (1) 可以推出,

$$\boldsymbol{A}^k\boldsymbol{x} = \boldsymbol{A}^{k-1}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}^{k-1}(\lambda\boldsymbol{x}) = \lambda\boldsymbol{A}^{k-1}\boldsymbol{x} = \cdots = \lambda^k\boldsymbol{x},$$

以及 $f(\boldsymbol{A})\boldsymbol{x} = f(\lambda)\boldsymbol{x}$. 从而 $\boldsymbol{x}$ 是 $\boldsymbol{A}^k$ 关于特征值 λ^k 的一个特征向量, 也是 $f(\boldsymbol{A})$ 关于特征值 $f(\lambda)$ 的一个特征值.

(2) 此时 $\det(\lambda\boldsymbol{I}-\boldsymbol{A})=0$, 但是转置运算不改变方阵的行列式, 从而 $\det(\lambda\boldsymbol{I}-\boldsymbol{A}^{\mathsf{T}})=0$. 这说明 λ 也是 $\boldsymbol{A}^{\mathsf{T}}$ 的特征值.

(3) 我们有 $\boldsymbol{A}^*\boldsymbol{A}=\det(\boldsymbol{A})\boldsymbol{I}$, 从而可以推出 $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}=\det(\boldsymbol{A})\boldsymbol{x}$. 可是 $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}^*(\lambda\boldsymbol{x})=\lambda\boldsymbol{A}^*\boldsymbol{x}$, 这说明 $\boldsymbol{A}^*\boldsymbol{x}=\frac{\det(\boldsymbol{A})}{\lambda}\boldsymbol{x}$, 从而 $\boldsymbol{x}$ 是 $\boldsymbol{A}^*$ 的属于 $\frac{\det(\boldsymbol{A})}{\lambda}$ 的特征向量.

(4) 由于 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$, 取转置运算和复共轭后, 我们有 $\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}=\overline{\lambda}\overline{\boldsymbol{x}}^{\mathsf{T}}$, 再分别右乘等式 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$ 的两边, 我们有 $\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\boldsymbol{x}=\lambda\overline{\lambda}\overline{\boldsymbol{x}}^{\mathsf{T}}\boldsymbol{x}$, 即 $\|\boldsymbol{x}\|^2=|\lambda|^2\cdot\|\boldsymbol{x}\|^2$. 由于 $0\neq\|\boldsymbol{x}\|\in\mathbb{C}$, 我们推出 $|\lambda|=1$. □

注 6.3.14. (1) 设 n 阶方阵 $\boldsymbol{A}$ 的每行元素之和都等于 s. 容易看出, 对于全为 1 的列向量 $\boldsymbol{x}=(1,1,\ldots,1)^{\mathsf{T}}$ 有 $\boldsymbol{A}\boldsymbol{x}=s\boldsymbol{x}$. 从而, $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的关于特征值 s 的特征向量.

(2) 设 n 阶方阵 $\boldsymbol{B}$ 的每列元素之和都等于 t. 由于转置运算不改变矩阵的特征值, 利用上面一条可以看出, t 是 $\boldsymbol{B}$ 的特征值. 不过, 此时的特征向量就不一定容易描述了.

命题 6.3.15. 若 $\boldsymbol{A}=\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}$, 则 $p_{\boldsymbol{A}}(\lambda)=p_{\boldsymbol{B}}(\lambda)$, 即它们的特征多项式相等. 特别地, 相似的矩阵的特征值 (重根按重数计算) 相同, 特征多项式的各个系数 (如这些矩阵的迹、行列式等等) 也对应相等.

证明. 直接计算, 有

$$p_{\boldsymbol{A}}(\lambda)=\det(\lambda\boldsymbol{I}-\boldsymbol{A})=\det(\lambda\boldsymbol{I}-\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T})=\det(\boldsymbol{T}^{-1}(\lambda\boldsymbol{I}-\boldsymbol{B})\boldsymbol{T})=\det(\lambda\boldsymbol{I}-\boldsymbol{B})=p_{\boldsymbol{B}}(\lambda).$$

□

注 6.3.16. 对于相似的矩阵, 虽然其特征值相同, 对应的特征向量还是会改变的. 事实上, 若 $\boldsymbol{B}=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$, 则 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$ 当且仅当 $\boldsymbol{T}\boldsymbol{B}\boldsymbol{T}^{-1}\boldsymbol{x}=\lambda\boldsymbol{x}$, 当且仅当 $\boldsymbol{B}(\boldsymbol{T}^{-1}\boldsymbol{x})=\lambda(\boldsymbol{T}^{-1}\boldsymbol{x})$. 这说明 $\boldsymbol{x}$ 是 $\boldsymbol{A}$ 的属于 λ 的特征向量, 当且仅当 $\boldsymbol{T}^{-1}\boldsymbol{x}$ 是 $\boldsymbol{B}$ 的属于 λ 的特征向量.

命题 6.3.17. 设 $\boldsymbol{A},\boldsymbol{B}$ 为同阶方阵, 则 $\boldsymbol{A}\boldsymbol{B}$ 和 $\boldsymbol{B}\boldsymbol{A}$ 具有相同的特征多项式.

证明. 我们只需说明 $\det(\lambda\boldsymbol{I}-\boldsymbol{A}\boldsymbol{B})=\det(\lambda\boldsymbol{I}-\boldsymbol{B}\boldsymbol{A})$. 当 $\lambda=0$ 时, 这等价于证明 $\det(-\boldsymbol{A}\boldsymbol{B})=\det(-\boldsymbol{B}\boldsymbol{A})$, 而这是显然的. 当 $\lambda\neq 0$ 时, 我们将教材第四章习题 #25 用于 $\frac{1}{\lambda}\boldsymbol{A}$ 和 $\boldsymbol{B}$, 从而有 $\det(\boldsymbol{I}-\frac{1}{\lambda}\boldsymbol{A}\boldsymbol{B})=\det(\boldsymbol{I}-\frac{1}{\lambda}\boldsymbol{B}\boldsymbol{A})$. 对其稍加变形即可. □

注 6.3.18. 设 $\lambda_1, \ldots, \lambda_n$ 是 n 阶方阵 $\boldsymbol{A}$ 的所有特征值, $f(t)$ 为一元多项式, 则 $f(\lambda_1), \ldots, f(\lambda_n)$ 皆为 $f(\boldsymbol{A})$ 的特征值. 由于可能出现 $i \neq j$ 但是 $f(\lambda_i) = f(\lambda_j)$ 的情形, 目前这并没有表明 $f(\lambda_1), \ldots, f(\lambda_n)$ 为 $f(\boldsymbol{A})$ 全部的特征值. 为此, 我们利用下一节要提到的 *Schur* 定理. 不妨设所考虑的矩阵为复方阵, 由该定理可知, 存在可逆方阵 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 为上三角矩阵, 其对角线上的元素依次为 $\lambda_1, \ldots, \lambda_n$. 可以直接验证, $\boldsymbol{T}^{-1}f(\boldsymbol{A})\boldsymbol{T}$ 仍然是上三角矩阵, 其对角线上的元素依次为 $f(\lambda_1), \ldots, f(\lambda_n)$. 这足以说明, $f(\lambda_1), \ldots, f(\lambda_n)$ 为 $f(\boldsymbol{A})$ 全部的特征值.

我们下面开始具体研究复方阵的特征多项式. 设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{C}^{n\times n}$, 则

$$p_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & \lambda - a_{nn} \end{vmatrix}.$$

很明显, $p_{\boldsymbol{A}}(\lambda)$ 是关于 λ 的多项式. 由于 λ 仅出现在主对角线上, 不难看出 $p_{\boldsymbol{A}}(\lambda)$ 是关于 λ 的最高次系数为 1 的 n 次多项式, 从而

$$\begin{aligned} p_{\boldsymbol{A}}(\lambda) &\xlongequal{\text{记作}} \lambda^n + \sigma_1\lambda^{n-1} + \cdots + \sigma_{n-1}\lambda + \sigma_n \\ &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n). \end{aligned}$$

其中 $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ 为 $p_{\boldsymbol{A}}(\lambda)$ 的根, 即 $\boldsymbol{A}$ 的特征值. 由**韦达定理** (**多项式系数与根的关系**) 可知,

$$\sigma_i = (-1)^i \sum_{1\le j_1<\cdots<j_i\le n} \lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_i},$$

例如, 当 $n = 4$ 而 $i = 2$, 则

$$\sigma_2 = (-1)^2(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_1\lambda_4 + \lambda_2\lambda_3 + \lambda_2\lambda_4 + \lambda_3\lambda_4).$$

需要特别关注的是,

$$\sigma_1 = -\sum_{i=1}^n \lambda_i, \qquad \sigma_n = (-1)^n\lambda_1\cdots\lambda_n.$$

另一方面, 对于给定 n 阶矩阵 $\boldsymbol{A}$, 形如 $\left|\boldsymbol{A}\left(\begin{smallmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{smallmatrix}\right)\right|$ 的 k 阶子式称为行列式 $|\boldsymbol{A}|$ 的 k 阶**主子式**, 其中 $1 \le i_1 < i_2 < \cdots < i_k \le n$. 在注 6.3.19 中, 我们可以证明: 上面的 $(-1)^k\sigma_k$ 是所有 k 阶主子式的和. 特别地, 我们有

$$-\sigma_1 = \sum_{i=1}^n a_{ii}, \qquad (-1)^n\sigma_n = \det(\boldsymbol{A}).$$

这说明

$$\det(\boldsymbol{A}) = \lambda_1 \cdots \lambda_n, \tag{6.2}$$

以及

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i. \tag{6.3}$$

注 6.3.19. 这儿, 我们证明: $(-1)^k\sigma_k$ 是所有 k 阶主子式的和. 设 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 是 $\boldsymbol{A}$ 的 n 个列向量, 于是 $p_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{e}_1 - \boldsymbol{\alpha}_1, \lambda \boldsymbol{e}_2 - \boldsymbol{\alpha}_2, \ldots, \lambda \boldsymbol{e}_n - \boldsymbol{\alpha}_n)$. 由行列式函数的多重线性性, 该式等于 感兴趣的同学可以了解一下

$$\sum_{k=0}^{n} \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} D_{i_1, i_2, \ldots, i_k},$$

其中, 若设 $\{i_{k+1} < i_{k+2} < \cdots < i_n\}$ 为 $\{i_1 < i_2 < \cdots < i_k\}$ 在 $\{1, 2, \ldots, n\}$ 中的补集, 则 $D_{i_1, i_2, \ldots, i_k}$ 为行列式函数, 其 $i_1, i_2, \ldots, i_k$ 分量分别为 $-\boldsymbol{\alpha}_{i_1}, -\boldsymbol{\alpha}_{i_2}, \ldots, -\boldsymbol{\alpha}_{i_k}$, 而 $i_{k+1}, i_{k+2}, \ldots, i_n$ 分量分别为 $\lambda \boldsymbol{e}_{i_{k+1}}, \lambda \boldsymbol{e}_{i_{k+2}}, \ldots, \lambda \boldsymbol{e}_{i_n}$. 不难验证,

$$D_{i_1, i_2, \ldots, i_n} = (-1)^k \lambda^{n-k} \left| \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix} \right|,$$

从而完成了证明.

习题 6.3.20. 设 $\boldsymbol{A} = (a_{ij})_{3\times 3}$. 若 $\boldsymbol{A}$ 的特征多项式为 $|\lambda \boldsymbol{I}_3 - \boldsymbol{A}| = \lambda^3 + \sigma_1 \lambda^2 + \sigma_2 \lambda + \sigma_3$, 直接验证 σ_2 是 $\boldsymbol{A}$ 的所有 2 阶主子式的和.

由公式 (6.2) 出发, 我们立刻得到

推论 6.3.21. n 阶方阵可逆当且仅当零不是它的特征值.

例 6.3.22. 学生课堂上自学教材 *P174-175* 的例 *6.3.4, 6.3.5*.

例 6.3.23. 设 $\boldsymbol{A}$ 为 3 阶方阵, 有特征值 $1, 2, 3$. 设 $\boldsymbol{B} = \boldsymbol{A}^3 + 2\boldsymbol{A}^2 + \boldsymbol{A} + 2\boldsymbol{I}$. 若以 $f(x) = x^3 + 2x^2 + x + 2$, 则 $\boldsymbol{B} = f(\boldsymbol{A})$, 从而以 $f(1) = 6$, $f(2) = 20$ 和 $f(3) = 50$ 为特征值. 特别地, $|\boldsymbol{B}| = 6 \cdot 20 \cdot 50 = 6000$, $\operatorname{tr}(\boldsymbol{B}) = 6 + 20 + 50 = 76$.

习题 6.3.24. 证明: 若 $\boldsymbol{A}^{2021} = \boldsymbol{O}$, 则 $\det(\boldsymbol{I} - 1958\boldsymbol{A}) = 1$.

例 6.3.25. 若 $\boldsymbol{A} = (a_{ij})_{4\times 4}$, 以 1 为二重特征值, 以 -2 为一重特征值, 求 $\boldsymbol{A}$ 的特征多项式.

解. 设 $\lambda_{1,2} = 1$, $\lambda_3 = -2$, 则 $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = \operatorname{tr}(\boldsymbol{A}) = a_{11} + a_{22} + a_{33} + a_{44}$. 故, $\lambda_4 = a_{11} + a_{22} + a_{33} + a_{44}$. 这说明

$$p_{\boldsymbol{A}}(\lambda) = (\lambda - 1)^2(\lambda + 2)(\lambda - (a_{11} + a_{22} + a_{33} + a_{44})). \qquad \square$$

线性变换的特征值与特征向量

定义 6.3.26. 设 F 为数域, V 为 F 上的线性空间, $\mathscr{A}$ 为 V 上的线性变换. 若存在 $\lambda \in F$ 以及非零向量 $\boldsymbol{x} \in V$, 满足 $\mathscr{A}(\boldsymbol{x}) = \lambda\boldsymbol{x}$, 则称 λ 为 $\mathscr{A}$ **的一个特征值**, 并称 $\boldsymbol{x}$ 为**属于 λ 的一个特征向量**.

类似地, 我们可以考察集合 $V_{\mathscr{A}}(\lambda) \coloneqq \{\, \boldsymbol{x} \in V \mid \mathscr{A}(\boldsymbol{x}) = \lambda\boldsymbol{x} \,\}$. 由于 $\mathscr{A}$ 为线性变换, $\mathbf{0} \in V_\lambda$, 从而该集合非空. 容易用定义验证, $V_{\mathscr{A}}(\lambda)$ 对于向量的加法与数乘运算封闭, 从而构成 V 的一个子空间. 由定义 6.3.26 知, λ 是 $\mathscr{A}$ 的一个特征值当且仅当 $V_{\mathscr{A}}(\lambda)$ 存在非零向量. 若该条件满足, 我们称 $V_{\mathscr{A}}(\lambda)$ 为**属于 λ 的特征子空间**. 一个向量是属于 λ 的特征向量当且仅当它是 $V_{\mathscr{A}}(\lambda)$ 中的一个非零向量.

例 6.3.27. 若 $\mathscr{A}$ 为恒等变换, 则 $\mathscr{A}$ 仅有特征值 1, 特征子空间 $V_{\mathscr{A}}(1) = V$ 为全空间, 即, V 中任意的非零向量都是属于 1 的特征向量.

注 6.3.28. *线性变换的特征值与特征向量的求解问题, 一般都是利用下面的命题 6.3.29, 化成在某组基下的矩阵的相应问题的求解. 关于这方面具体的例子, 见教材 P175 的例 6.3.6. 学生课后自习.*

命题 6.3.29. *若 $\mathscr{A}$ 是 n 维线性空间 V 上的一个线性变换, 在基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 下的矩阵为 $\boldsymbol{A}$. 那么, 我们有以下几条.* 教材命题 6.3.1

(1) $\boldsymbol{A}$ 与 $\mathscr{A}$ 具有相同的特征值.

(2) 设 λ 是 $\mathscr{A}$ 的一个特征值, 则 $\boldsymbol{x}$ 是 $\mathscr{A}$ 的关于 λ 的特征向量的充要条件是 $\boldsymbol{x}$ 在基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 下的坐标列向量 $\boldsymbol{X}$ 是 $\boldsymbol{A}$ 关于 λ 的特征向量. 从而,

$$V_{\mathscr{A}}(\lambda) = \left\{\, \boldsymbol{x} \in V \;\middle|\; \boldsymbol{x} \text{ 在基 } \boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n \text{ 下的坐标 } \boldsymbol{X} \text{ 属于 } V_{\boldsymbol{A}}(\lambda) \,\right\}.$$

特别地, $\mathscr{A}$ 至多有 n 个不同的特征值.

证明. 设 $\lambda \in F$, $\boldsymbol{x} \in V$ 在 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 下的坐标为 $\boldsymbol{X} = (x_1, \dots, x_n)^\mathsf{T}$, 即, $\boldsymbol{x} = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n)\boldsymbol{X} = \sum_i x_i\boldsymbol{\alpha}_i$. 此时, 我们有

$$\mathscr{A}(\boldsymbol{x}) = \mathscr{A}\left(\sum_i x_i\boldsymbol{\alpha}_i\right) = \sum_i x_i\mathscr{A}(\boldsymbol{\alpha}_i) = (\mathscr{A}(\boldsymbol{\alpha}_1), \dots, \mathscr{A}(\boldsymbol{\alpha}_n))\boldsymbol{X} = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n)\boldsymbol{A}\boldsymbol{X},$$

以及

$$\lambda\boldsymbol{x} = \lambda\left((\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n)\boldsymbol{X}\right) = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n)\lambda\boldsymbol{X}.$$

由坐标表示的唯一性可知 $\boldsymbol{A}\boldsymbol{X} = \lambda\boldsymbol{X}$ 的充要条件是 $\mathscr{A}(\boldsymbol{x}) = \lambda\boldsymbol{x}$. 另一方面, $\boldsymbol{x}$ 是零向量当且仅当 $\boldsymbol{X}$ 是零向量. 从而, 命题的各条容易验证. □

请先自行认真阅读教材上 P175 的例 6.3.6, 然后完成下面习题.

习题 6.3.30. 设 $\boldsymbol{A}=\begin{pmatrix}1&2\\0&4\end{pmatrix}$, $V=F^{2\times 2}$ 为线性空间, 定义 V 上的线性变换 $\mathscr{A}:\boldsymbol{M}\mapsto\boldsymbol{AM}$, $\boldsymbol{M}\in V$.

(1) 求 $\mathscr{A}$ 的特征值和特征向量.

(2) 求 $\mathscr{A}$ 在基 $\begin{pmatrix}1&1\\1&1\end{pmatrix}$, $\begin{pmatrix}-1&1\\1&-1\end{pmatrix}$, $\begin{pmatrix}1&-1\\1&-1\end{pmatrix}$, $\begin{pmatrix}1&1\\-1&-1\end{pmatrix}$ 下的矩阵.

习题 6.3.31. 在以下题目中, 设 $\boldsymbol{A}$ 是线性变换 $\mathscr{A}$ 在自然基下的矩阵. 不写出 $\boldsymbol{A}$, 直接描述 (求出) $\boldsymbol{A}$ 的特征值与特征子空间.

(1) $\mathscr{A}$ 是 $\mathbb{R}^2$ 的相对于某条过原点的直线的反射变换 (镜像映射).

(2) $\mathscr{A}$ 是 $\mathbb{R}^3$ 的相对于某条过原点的直线的旋转变换.

例 6.3.32. 对于实数域上的线性空间 $\mathbb{R}^{n\times n}$ 上的线性变换 $\mathscr{A}:\boldsymbol{A}\mapsto\boldsymbol{A}^{\mathsf{T}}$, 讨论其特征值和特征向量.

解. 设 $\boldsymbol{X}=(x_{ij})$ 是 $\mathscr{A}$ 属于特征值 $\lambda\in\mathbb{R}$ 的一个特征向量, 从而 $\boldsymbol{X}\neq\boldsymbol{O}$, 且 $\boldsymbol{X}^{\mathsf{T}}=\lambda\boldsymbol{X}$, 即 $x_{ij}=\lambda x_{ji}$ 对任意的 $1\le i,j\le n$ 都成立. 由于 $\boldsymbol{X}\neq\boldsymbol{O}$, 存在 $x_{ij}\neq 0$. 由于 $x_{ij}=\lambda x_{ji}$ 且 $x_{ji}=\lambda x_{ij}$, 故 $x_{ij}=\lambda^2x_{ij}$. 这说明 $\lambda^2=1$, 即 $\lambda=\pm1$. 若 $\lambda=1$, 则 $\boldsymbol{X}^{\mathsf{T}}=\boldsymbol{X}$, 即 $\boldsymbol{X}$ 为非零的对称矩阵. 若 $\lambda=-1$, 则 $\boldsymbol{X}^{\mathsf{T}}=-\boldsymbol{X}$, 即 $\boldsymbol{X}$ 为非零的反对称矩阵. □

定义 6.3.33. 在数域 F 上的有限维线性空间 V 上的线性变换 $\mathscr{A}$, 在 V 的一组基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下有矩阵 $\boldsymbol{A}$. 在命题 *6.3.29* 中我们已经见到了 $\boldsymbol{A}$ 的特征向量与 $\mathscr{A}$ 的特征向量之间的关系. 由此出发, 我们定义如下:

$\mathscr{A}$ 的**特征多项式** $p_{\mathscr{A}}(\lambda):=p_{\boldsymbol{A}}(\lambda)$, $\mathscr{A}$ 的**秩** $\operatorname{rank}(\mathscr{A}):=\operatorname{rank}(\boldsymbol{A})$,

$\mathscr{A}$ 的**行列式** $\det(\mathscr{A}):=\det(\boldsymbol{A})$, $\mathscr{A}$ 的**迹** $\operatorname{tr}(\mathscr{A}):=\operatorname{tr}(\boldsymbol{A})$.

由于特征多项式、秩、行列式、迹是方阵的相似不变量, 故上面的定义不依赖于特定的基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ (从而 $\boldsymbol{A}$) 的选取.

习题 6.3.34. 设 $\boldsymbol{A}$ 是 $V=F^{n\times n}$ 中取定的矩阵, 并据此定义 V 到自身的映射 $\ell_{\boldsymbol{A}}:V\to V$, $\boldsymbol{X}\mapsto\boldsymbol{AX}$, 以及 $r_{\boldsymbol{A}}:V\to V$, $\boldsymbol{X}\mapsto\boldsymbol{XA}$. 不难看出, $\ell_{\boldsymbol{A}}$ 和 $r_{\boldsymbol{A}}$ 都是 V 上的线性变换. 证明: $\operatorname{tr}(\ell_A)=\operatorname{tr}(r_{\boldsymbol{A}})=n\operatorname{tr}(\boldsymbol{A})$.

定义 6.3.35. 对于线性空间 V 上的一个线性变换 $\mathscr{A}$, 若存在 V 的一组基, 使得 $\mathscr{A}$ 在这组基下的矩阵为对角阵, 那么我们称 $\mathscr{A}$ 为**可对角化的**.

注 6.3.36. 在上面的定义中, 任取 V 的一组基, 并设 $\mathscr{A}$ 在这组基下的矩阵为 $\boldsymbol{A}$. 则由定理 *6.2.7* 容易验证, $\boldsymbol{A}$ 可对角化当且仅当 $\mathscr{A}$ 可对角化.

6.4 矩阵的相似对角化

首先我们需要指出, 并不是所有的方阵都可以相似对角化.

例 6.4.1. 考察矩阵 $\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$. 由于该矩阵为上三角矩阵, 在例 *6.3.8* 中我们已经看到, $\boldsymbol{A}$ 的特征值为 $2, 2, 2$, 故若 $\boldsymbol{B}$ 与 $\boldsymbol{A}$ 相似, 则 $\boldsymbol{B}$ 的特征值也同样地为 $2, 2, 2$. 注意: $\boldsymbol{B}$ 不能为对角阵, 否则 $\boldsymbol{B}$ 必为 $2\boldsymbol{I}_3$. 但是 $2\boldsymbol{I}_3$ 的相似等价类中仅有一个元素 $2\boldsymbol{I}_3$, 并不包含 $\boldsymbol{A}$. 故 $\boldsymbol{A}$ 与 $2\boldsymbol{I}_3$ 并不相似等价, 即, $\boldsymbol{A}$ 不可对角化.