特征值与特征向量的定义

定义 1

设 $\boldsymbol{A}$ 为数域 F 上的 n 阶方阵. 如果存在 $\lambda \in F$ 和非零列向量 $\boldsymbol{x} \in F^n$ 使得 $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, 则称 λ 为 $\boldsymbol{A}$ 的一个特征值 (*eigenvalue, characteristic value*), 并称 $\boldsymbol{x}$ 为属于 λ 的一个特征向量 (*eigenvector, characteristic vector*).

设 $\boldsymbol{A} = (a_{ij})$ 为数域 F 上的一个 n 阶方阵, 则

$$\lambda \in F \text{ 为 } \boldsymbol{A} \text{ 的特征值} \Leftrightarrow V_A(\lambda) \text{ 有非零向量} \Leftrightarrow (\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0} \text{ 有非零解}$$

$$\Leftrightarrow \lambda \boldsymbol{I} - \boldsymbol{A} \text{ 不可逆} \Leftrightarrow \text{行列式 } |\lambda \boldsymbol{I} - \boldsymbol{A}| = 0.$$

我们可以观察到

$$
\begin{aligned}
|\lambda \boldsymbol{I}-\boldsymbol{A}| &= \begin{vmatrix} \lambda-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a_{n-1,n} \\ -a_{n1} & \cdots & -a_{n,n-1} & \lambda-a_{nn} \end{vmatrix} \\
&= \lambda^n + \text{中间次数的复杂项} + (-1)^n \det(\boldsymbol{A}).
\end{aligned}
$$

最后的这个多项式是关于未知元 λ 的 F 上的 n 次多项式，其中间次数的系数较为复杂，我们稍后会作进一步解释. 该多项式将被记作 $p_{\boldsymbol{A}}(\lambda)$，称为 $\boldsymbol{A}$ 的特征多项式 (characteristic polynomial). 相应地，方程 $p_{\boldsymbol{A}}(\lambda)=0$ 称为矩阵 $\boldsymbol{A}$ 的特征方程 (characteristic equation). 具体取值 $\lambda_0 \in F$ 为 $\boldsymbol{A}$ 的特征值，当且仅当 λ_0 是 $p_{\boldsymbol{A}}(\lambda)=0$ 的根. 由域上的多项式的理论，我们知 $\boldsymbol{A}$ 的特征值最多有 n 个. 特别地，$\boldsymbol{A}$ 的特征值的个数有限.

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 2

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.

例 3

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 $\boldsymbol{A}$.

注 4

1. 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．
2. 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 $\overline{\boldsymbol{x}} \neq \mathbf{0}$，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 $(\lambda - \overline{\lambda})\boldsymbol{x} = \mathbf{0}$．但是 $\lambda \neq \overline{\lambda}$，从而可以推出 $\boldsymbol{x} = \mathbf{0}$．而这与 $\boldsymbol{x}$ 为特征值相矛盾．)

3. 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．

定义 5

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, $\boldsymbol{x} = \mathbf{0}$ 当且仅当 $\|\boldsymbol{x}\| = 0$.

例 6

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

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 $\lambda \in \mathbb{C}$)

注 7

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 8

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.

注 9

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 10

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.

注 11

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.

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 $\left|\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \cdots & i_k \\ i_1 & i_2 & \cdots & i_k \end{pmatrix}\right|$ 的 k 阶子式称为行列式 $|\boldsymbol{A}|$ 的 k 阶主子式, 其中 $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. 在注 12 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

$$\det(\boldsymbol{A}) = \lambda_1 \cdots \lambda_n, \tag{1}$$

以及

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i. \tag{2}$$

注 12 ($(-1)^k\sigma_k$ 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

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

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.

由公式 (1) 出发, 我们立刻得到

推论 13

n 阶方阵可逆当且仅当零不是它的特征值.

例 14

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

例 15

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, $|\boldsymbol{B}|=6\cdot 20\cdot 50=6000$, $\mathrm{tr}(\boldsymbol{B})=6+20+50=76$.

例 16

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.

定义 17

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.

例 18

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

定义 19

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 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 下有矩阵 $\boldsymbol{A}$. 在之前, 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:

$$\mathscr{A} \text{ 的特征多项式 } p_{\mathscr{A}}(\lambda) := p_{\boldsymbol{A}}(\lambda), \qquad \mathscr{A} \text{ 的秩 } \operatorname{rank}(\mathscr{A}) := \operatorname{rank}(\boldsymbol{A}),$$
$$\mathscr{A} \text{ 的行列式 } \det(\mathscr{A}) := \det(\boldsymbol{A}), \qquad \mathscr{A} \text{ 的迹 } \operatorname{tr}(\mathscr{A}) := \operatorname{tr}(\boldsymbol{A}).$$

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 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ (从而 $\boldsymbol{A}$) 的选取.

定义 20

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 21

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 $\boldsymbol{A}$. 则容易验证, $\boldsymbol{A}$ 可对角化当且仅当 $\mathscr{A}$ 可对角化.

矩阵的相似对角化

并不是所有的方阵都可以相似对角化

例 22

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, $\boldsymbol{A}$ 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.