《线性代数》

第七章 线性运算与线性变换

§7.2 特征值与特征向量

2019秋

杨晶 主讲

(零) 问题引入

例1 (人口流动问题) 设某城市的总人口固定不变 (记为c)，开始时市区人口数为x_0，郊区人口数为y_0 (满足$x_0+y_0=c$). 如果今后每年都有5%的市区居民搬到郊区，而有15%的郊区居民搬到市区.

(1) 请问: 20年后市区和郊区人口各为多少?

(2) 请对该市长期的人口分布做出预测.

分析： 首先找出第n年和第$n+1$年市区和郊区人口数的关系；然后根据递推关系得到第20年与开始时(第0年)的人口数的关系.

进而，对充分大的n，根据第n年与开始时(第0年)的人口数的关系，预测长期的人口分布规律.

解： 设第n年市区人口数和郊区人口数分别为x_n与y_n，则根据题意，x_{n+1}, y_{n+1}与x_n, y_n有如下关系式：

$$\begin{cases} x_{n+1} = 0.95x_n + 0.15y_n, \\ y_{n+1} = 0.05x_n + 0.85y_n. \end{cases}$$

表为矩阵形式为：

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \overbrace{\boxed{\frac{1}{20}\begin{pmatrix} 19 & 3 \\ 1 & 17 \end{pmatrix}}}^{A} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = A\begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

于是有递推关系式：

求A^n为关键

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = A\begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = A^2\begin{pmatrix} x_{n-2} \\ y_{n-2} \end{pmatrix} = \cdots = A^n\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

然而，手动计算A^{20}并不是容易的事情，于是我们借助计算机，计算得

$$A^{20} \approx \begin{pmatrix} 0.7525 & 0.7425 \\ 0.2475 & 0.2575 \end{pmatrix} \approx \begin{pmatrix} 3/4 & 3/4 \\ 1/4 & 1/4 \end{pmatrix}.$$

因此，$$\begin{pmatrix} x_{20} \\ y_{20} \end{pmatrix} = A^{20}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \approx \begin{pmatrix} 3/4 & 3/4 \\ 1/4 & 1/4 \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 3(x_0+y_0)/4 \\ (x_0+y_0)/4 \end{pmatrix} = \begin{pmatrix} (3/4)c \\ (1/4)c \end{pmatrix}.$$

借助计算机，对更大的n，进行数值计算可知：$\begin{pmatrix} x_n \\ y_n \end{pmatrix} \to c\begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix} \quad (n \to \infty)$

问题：是否有方法，不需要依赖计算机，也可以计算一般的A^n，并对长期的人口分布进行预测？

回答：有！但需要学习本章即将介绍的内容.

问题回顾

(1) 对于给定的线性变换σ, 如果存在一组基 $\vec{\beta}_1,\vec{\beta}_2,\ldots,\vec{\beta}_n$ 使得σ在这组基下的矩阵D是一个对角矩阵, 则称线性变换σ可对角化. 中心问题: 线性变换σ可对角化的条件?

(2) 对于给定方阵A, 如果存在可逆矩阵P使得$P^{-1}AP$为对角矩阵, 则称方阵A可对角化. 中心问题: 方阵A可对角化的条件?

分析: 仅讨论n阶矩阵$A=(a_{ij})$可对角化的条件. 设存在n阶可逆矩阵P使得

$$P^{-1}AP=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$$

$$\Rightarrow\ AP=P\,\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$$

进一步, 令 $P=(\vec{X}_1,\vec{X}_2,\ldots,\vec{X}_n)$, 代入得

$$A(\vec{X}_1,\vec{X}_2,\ldots,\vec{X}_n)=(\vec{X}_1,\vec{X}_2,\ldots,\vec{X}_n)\mathrm{diag}(\lambda_1,\ldots,\lambda_n)$$

$$\Rightarrow\ (A\vec{X}_1,\ldots,A\vec{X}_n)=(\lambda_1\vec{X}_1,\ldots,\lambda_n\vec{X}_n)$$

$$\Rightarrow\ A\vec{X}_i=\lambda_i\vec{X}_i,\ (i=1,2,\ldots,n).$$

- 说明我们需要专门研究满足下列条件的$\vec{X}$和λ:

$$\boxed{A\vec{X}=\lambda\vec{X}.}$$

内容提要

- 特征值与特征向量的概念
- 特征值与特征向量的计算
- 特征值与特征多项式的性质
- 特征向量的性质与特征子空间

(一) 特征值与特征向量的概念

定义1 设A为n阶方阵，若存在数λ及n维非零(列)向量 $\vec{X}$，使得

$$A\vec{X}=\lambda\vec{X}, \tag{1}$$

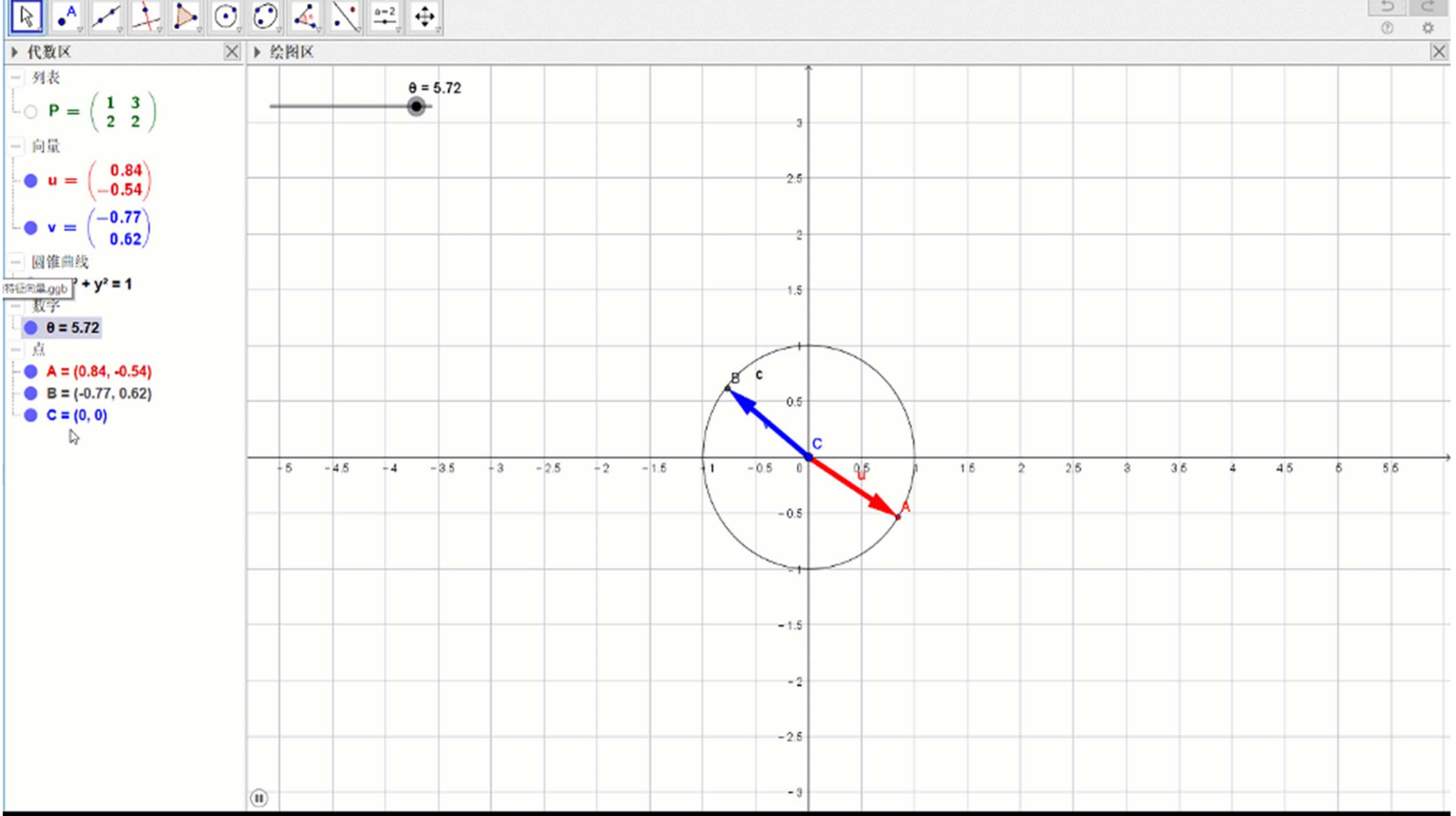
则称 λ 是 A 的特征值(eigenvalue)，$\vec{X}$ 是 A 的属于特征值 λ 的特征向量(eigenvector).

例如：

- 对于零方阵O，0为其特征值，所有非零(列)向量都是属于特征值0的特征向量.
- 对于单位阵I_n，1为其特征值，所有非零(列)向量都是属于特征值1的特征向量.

注: (i) 特征值与特征向量是成对绑定出现的，但有从属关系.

(ii) 特征向量必非零(向量)，但特征值可以为零.


代数区
列表
P = (1 3; 2 2)
向量
u = (0.84, −0.54)
v = (−0.77, 0.62)
圆锥曲线
特征向量.ggb
+ y² = 1
数字
θ = 5.72
点
A = (0.84, -0.54)
B = (-0.77, 0.62)
C = (0, 0)
绘图区
θ = 5.72
B
c
C
u
A

特征值与特征向量的定义并不是凭空想象出来的，而是有具体的几何意义与实际意义，例如：

- 当今网络搜索中，输入“电脑”和“计算机”二词，搜出的相关结果几乎是相同的. 究其原因，其实是后台算法将这两个词出现的语境列为两个很大的矩阵，进而计算这两个矩阵的特征值，发现它们的特征值几乎相同. 故认为它们表示同样的含义, 从而列出几乎相同的搜索结果. ——说明特征值反映了一些事物的本质特征.
- 几何方面，特征向量对应了映射作用下的某种不变量(构成不变子空间)，而特征值对应了这个不变集合上的共同特征量. ——这将在下学期<线性代数2>中详细讨论.
- 引出矩阵对角化问题，从而给出计算A^k的理论方法.
- 在数学的其他分支，以及其他学科, 如: 微分方程数值解、矩阵函数、物体与波形的固有频率等等, 实际问题中有非常重要的应用。

➢ 特征值与特征向量的基本问题

对于给定的n阶方阵A，有如下问题有待讨论：

- **存在问题：** A是否存在特征值与特征向量？
- **个数问题：** A若存在特征值，共有多少个？对于A的一个特征值，属于它的特征向量共有多少个？
- **计算问题：** 给出方法计算出A所有的特征值与特征向量.
- **关系问题：** A的各个特征值之间有何关系？属于A的不同特征值的特征向量有何关系？
- **结构问题：** 属于A的同一个特征值的特征向量组成的集合结构如何?

(二) 特征值与特征向量的计算方法

首先来试算一个简单的例子，这是二阶方阵的情形.

例2 求 $A=\begin{bmatrix}1 & -2\\ 1 & 4\end{bmatrix}$ 的所有特征值和特征向量.

解 设 A 的特征值是 λ, 属于特征值 λ 的特征向量为 $\vec{X}$, 则 $A\vec{X}=\lambda\vec{X}$,

即 $\begin{bmatrix}1 & -2\\ 1 & 4\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\lambda\begin{bmatrix}x_1\\ x_2\end{bmatrix}$ 亦即 $\begin{cases}(\lambda-1)x_1+2x_2=0\\ -x_1+(\lambda-4)x_2=0\end{cases}$

这是两个未知数两个方程的齐次线性方程组, 按定义, 特征向量是非零向量, 于是要求上述齐次线性方程组有非零解, 这等价于

$$\begin{vmatrix}\lambda-1 & 2\\ -1 & \lambda-4\end{vmatrix}=(\lambda-1)(\lambda-4)+2=(\lambda-2)(\lambda-3)=0,$$

解得 $\lambda_1=2$, $\lambda_2=3$，这就是 A 的两个特征值.

下面，分别求解属于特征值 $\lambda_1 = 2$与$\lambda_2 = 3$的特征向量.

对于特征值2, 代入得到齐次线性方程组: $\begin{cases} x_1 + 2x_2 = 0 \\ -x_1 - 2x_2 = 0 \end{cases}$

解得属于2的特征向量是 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k\begin{bmatrix} 2 \\ -1 \end{bmatrix}$, 其中$k$是任意非零常数.

对于特征值3, 代入得到齐次线性方程组: $\begin{cases} 2x_1 + 2x_2 = 0 \\ -x_1 - x_2 = 0 \end{cases}$

解得属于3的特征向量是 $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, 其中$k$是任意非零常数.

问题：对在本例中，我们发现一些规律. 但是，对于一般矩阵是不是也有类似的方法和规律呢?

- 给定 n 阶方阵 $A=(a_{ij})_{n\times n}$，设 A 的特征值λ与特征向量$\vec{X}$存在，则有

$$A\vec{X}=\lambda\vec{X}, \tag{1}$$

$$\Leftrightarrow\ (\lambda I_n-A)\vec{X}=\vec{0},\ \text{即}\quad \begin{cases}(\lambda-a_{11})x_1-a_{12}x_2-\cdots-a_{1n}x_n=0\\ -a_{21}x_1+(\lambda-a_{22})x_2-\cdots-a_{2n}x_n=0\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ -a_{n1}x_1-a_{n2}x_2-\cdots+(\lambda-a_{nn})x_n=0\end{cases} \tag{2}$$

由特征向量$\vec{X}$存在, 齐次线性方程组(2)有非零解, 这又等价于(2)的系数矩阵的行列式为0，即

$$\Leftrightarrow\quad |\lambda I-A|=\begin{vmatrix}\lambda-a_{11} & -a_{12} & \cdots & -a_{n1}\\ -a_{21} & \lambda-a_{22} & \cdots & -a_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn}\end{vmatrix}=0 \tag{3}$$

上述推导说明: A 的所有特征值一定满足(3). 反之, 满足(3)的一个数λ也一定是 A 的一个特征值. 求属于这个特征值的所有特征向量，只需将λ代入(2)式，求出齐次线性方程组的所有非零解即可.

特别地，对(3)式左边的含参行列式，我们有如下定义.

定义2 下述关于λ的n次多项式称为n次方阵A的**特征多项式(eigenpolynomial)**

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

- 求给定矩阵A的特征值相当于求 A 的特征多项式的根，故特征值也叫做**特征根(eigenroot)**.
- λ作为 $f_A(\lambda)$的根的重数，称为特征值λ的**代数重数(Algebraic multiplicity)**.

● 根据以上讨论，对方阵 A 的特征值与特征向量的**计算步骤**如下：

(1) 计算 A 的特征多项式 $f_A(\lambda)$.

(2) 求出特征多项式 $f_A(\lambda)$ 的所有根 λ_i，它们就是 A 的全部特征值.

(3) 分别把每个特征值 λ_i 代入方程组 $(\lambda I-A)\vec{X}=\vec{0}$，并求出它的基础解系并确定 $N(\lambda I-A)$，其中的所有非零向量就是 A 的属于 λ 的全部特征向量.

含参行列式计算

因式分解，一元高次方程求根

齐次线性方程组的求解

- 关于存在问题：由于特征多项式 $f_A(\lambda)$ 在复数范围内总是有根的，所以A的特征值总是存在的，于是特征向量也总是存在的.

- 关于个数问题：由特征多项式 $f_A(\lambda)$ 在复数范围内共有n个根 (累计重根)，故A的特征值共有n个 (累计重数). 而特征向量$\vec{X}$是齐次线性方程组的非零解，必有无穷多解，故属于A的同一个特征值的特征向量有无穷多个. 实际上，$A\vec{X}=\lambda\vec{X} \Rightarrow A(k\vec{X})=\lambda\left(k\vec{X}\right), \forall k \neq 0$.

- 本门课程中我们只考虑实矩阵A，由特征向量$\vec{X} \in N(\lambda I-A)$知，当特征值$\lambda$为实数时，对应的特征向量必为实向量；而由定义式$A\vec{X}=\lambda\vec{X}$知，当特征值$\lambda$为复数时，对应的特征向量必为复向量 (此情况超出了本课程的讨论范围)，因此，我们把问题集中到：

  求出A的所有实特征值以及属于它们的实特征向量.

再来看一个特殊的例子.

例3 考虑对角阵 $A=\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix}$，其中 $a_1,a_2,\cdots,a_n$ 为互不相同的实数.

求A的特征值与特征向量.

解: (1) 计算特征多项式: $|\lambda I-A|=(\lambda-a_1)(\lambda-a_2)\ldots(\lambda-a_n)$，显然得到全部特征值为

$$\lambda_1=a_1,\lambda_2=a_2,\cdots,\lambda_n=a_n.$$

(2) 对每个特征值 $\lambda = a_i$, 计算属于它的特征向量满足的线性方程组

$$(a_i I - A)\vec{X} = \vec{0}\ ,\ 即\ \begin{pmatrix} a_i - a_1 & & & & \\ & \ddots & & & \\ & & 0 & & \\ & & & \ddots & \\ & & & & a_i - a_n \end{pmatrix} \vec{X} = \vec{0}$$

由于$a_i \neq a_j\ (i \neq j)$，故上述方程组系数矩阵主对角线上只有一个位置为0, 故它的基础解系为 $\vec{X}_i = \vec{e}_i = (0,\cdots,1,\cdots,0)^T$.

故属于特征值 $\lambda = a_i$ 的特征向量为 $k_i\vec{e}_i\ (k_i \neq 0)$, 对所有的 $i=1,2,\ldots,n$.

问题：在本例中，属于不同特征值的特征向量显然是线性无关的. 这是偶然的吗？还是有一般的规律？

(三) 特征值与特征多项式的性质

下面讨论方阵 $A=(a_{ij})_{n\times n}$ 的特征多项式$f_A(\lambda)$的性质. 设

$$f_A(\lambda)=\lambda^n+c_1\lambda^{n-1}+c_2\lambda^{n-2}+\cdots+c_s\lambda^{n-s}+\cdots+c_{n-1}\lambda+c_n \quad (c_i\in\mathbb{R}).$$

我们用两种方式表示特征多项式的系数: 用矩阵元素、用矩阵特征根.

(1) 特征多项式的部分系数与矩阵A的元素的关系.

记 $A=(\vec{\alpha}_1,\vec{\alpha}_2,\ldots,\vec{\alpha}_n)$, $I_n=(\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_n)$, 则

$$f_A(\lambda)=|\lambda I-A|=\begin{vmatrix}\lambda\vec{e}_1+(-\vec{\alpha}_1) & \lambda\vec{e}_2+(-\vec{\alpha}_2) & \cdots & \lambda\vec{e}_n+(-\vec{\alpha}_n)\end{vmatrix}$$

根据行列式的列加法的拆分性质，可得两个如下两个系数与A中元素的关系:

$$c_n=(-1)^n|A|,\quad c_1=-\sum_{i=1}^{n}a_{ii}.$$

主对角线上元素之和称为A的迹, 记为$\mathrm{tr}(A)$

(2) 特征多项式的部分系数与矩阵A的特征值的关系.

设矩阵 A 的特征多项式的全部(复)根为: $\lambda_1,\lambda_2,\cdots,\lambda_n$，则

$$f_A(\lambda)=\lambda^n+c_1\lambda^{n-1}+c_2\lambda^{n-2}+\cdots+c_s\lambda^{n-s}+\cdots+c_{n-1}\lambda+c_n$$
$$=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n).$$

将此式展开, 利用韦达公式 (根与系数的关系)可知:

$$c_n=(-1)^n\prod_{i=1}^{n}\lambda_i,\quad c_1=-\sum_{i=1}^{n}\lambda_i.$$

比较上述两方面，即可得到如下结论:

定理1 设$A=(a_{ij})_{n\times n}$为n阶方阵，又设$\lambda_1,\lambda_2,\cdots,\lambda_n$为$A$的全部(复)特征值，则 (1) $\sum_{i=1}^{n}\lambda_i=\sum_{i=1}^{n}a_{ii}=trA$; (2) $|A|=\lambda_1\lambda_2\ldots\lambda_n=\prod_{i=1}^{n}\lambda_i$.

由此我们可得到一个利用特征值判断 A 是否可逆的结论.

推论 n 阶方阵 A 可逆 $\Leftrightarrow$ $|A|\neq 0$ $\Leftrightarrow$ A的特征值全不为零.

(等价命题) A 不可逆 $\Leftrightarrow$ $|A|=0$ $\Leftrightarrow$ 0是A的一个特征值.

扩展问题： 除了常数项c_n与次高项系数c_1之外，特征多项式的其他项系数c_k也可由韦达公式表为特征值$\lambda_1,\ldots,\lambda_n$的表达式；那么，$c_k$与$A$的分量元素$a_{ij}$有何关系？

例4 (1) 设 $A=\begin{pmatrix} a_{11} & & & * \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix}$，则 $f_A(\lambda)=\prod_{i=1}^{n}(\lambda-a_{ii})$.

从而，三角阵A的特征值就为主对角线上的全体元素$a_{11}, a_{22}, \dots, a_{nn}$.

(2) 设 $A=\begin{pmatrix} A_{11} & & & * \\ & A_{22} & & \\ & & \ddots & \\ 0 & & & A_{ss} \end{pmatrix}$，其中$A_{ii}$为$n_i$阶方阵，则

$$f_A(\lambda)=f_{A_{11}}(\lambda)f_{A_{22}}(\lambda)\cdots f_{A_{ss}}(\lambda).$$

从而，准三角阵A的特征值就为主对角线上的每个子块$A_{11}, A_{22}, \dots, A_{nn}$的特征值的并集.

性质： 设 λ_0为A 的特征值, $\vec{X}_0$为对应的特征向量，则

(i) $k\lambda_0$为kA 的特征值, $\vec{X}_0$为对应的特征向量;

(ii) λ_0^h为A^h 的特征值, $\vec{X}_0$为对应的特征向量;
(其中h为任意正整数)

(iii) $g(\lambda_0)$为$g(A)$ 的特征值, $\vec{X}_0$为对应的特征向量;
(其中$g(x)$是任意多项式)

进一步，若A可逆，还有如下两条性质：

(iv) λ_0^{-1}为A^{-1} 的特征值, $\vec{X}_0$为对应的特征向量;

(v) $\lambda_0^{-1}|A|$为A^* 的特征值,$\vec{X}_0$为对应的特征向量;

(iv) $A\vec{X}_0 = \lambda_0\vec{X}_0 \Rightarrow A^{-1}(A\vec{X}_0) = A^{-1}(\lambda_0\vec{X}_0) \Rightarrow \vec{X}_0 = \lambda_0 A^{-1}\vec{X}_0 \Rightarrow A^{-1}\vec{X}_0 = \lambda_0^{-1}\vec{X}_0$.

(v) $A^*A = |A|I \Rightarrow A^*A\vec{X}_0 = |A|\vec{X}_0 \Rightarrow A^*(\lambda_0\vec{X}_0) = |A|\vec{X}_0 \Rightarrow A^*\vec{X}_0 = \lambda_0^{-1}|A|\vec{X}_0$.

设置

例5 设 A 是三阶方阵, 它的特征值为 1, 2, −1, $B = A^3 - 5A^2$,

求 $|B|$=[填空1] .

正常使用填空题需3.0以上版本雨课堂

作答

例6 已知n阶方阵A的n个特征值为$\lambda_1, \ldots, \lambda_n$, 求$2I-A$的特征值及$\det(2I-A)$.

解 由题设知 $|\lambda I - A| = \prod_{i=1}^{n}(\lambda - \lambda_i)$. 下考虑$(2I-A)$的特征多项式

$$|\mu I - (2I - A)| = |(\mu - 2)I + A| = (-1)^n |(2-\mu)I - A|$$

$$= (-1)^n \prod_{i=1}^{n}((2-\mu) - \lambda_i) = \prod_{i=1}^{n}(\mu - (2 - \lambda_i)).$$

所以$(2I-A)$的n个特征值为 $2-\lambda_i$, $i = 1,2, \ldots, n$. 从而

$$\det(2I - A) = \prod_{i=1}^{n}(2 - \lambda_i).$$

注: 本例结论也可由上述性质(iii)直接得到.

例6 已知 $A^2 = A$, 求矩阵A的特征值.

解 设 λ 是 A 的任意一个特征值, $\vec{X}$ 是 λ 所属的特征向量, 则$A\vec{X} = \lambda\vec{X}$,

所以 $A^2\vec{X} = A(\lambda\vec{X}) = \lambda A\vec{X} = \lambda^2\vec{X}$,

利用已知条件 $A^2 = A$与上式, 有

$$\lambda\vec{X} = A\vec{X} = A^2\vec{X} = \lambda^2\vec{X},$$

所以 $(\lambda^2 - \lambda)\vec{X} = \vec{0}$, 因为 $\vec{X}$ 为特征向量，所以$\vec{X} \neq \vec{0}$, 故

$$\lambda^2 = \lambda, \quad \Rightarrow \quad \lambda = 0 \text{ 或 } \lambda = 1.$$

注： 本例可推广为一般情况, 若矩阵A满足某个多项式, 即$g(A)=O$, A的特征值λ也满足该多项式, 即$g(\lambda)=0$, 即性质(iii)

(四) 特征向量的性质与特征子空间

下面讨论特征向量的性质. 取定n阶方阵 A 的一个特征值 λ, 则

1. 设 $\vec{X}$ 是 A 属于特征值 λ 的特征向量, 即 $A\vec{X}=\lambda\vec{X}$, 又设 $k\in\mathbb{R}$, 则

$$A(k\vec{X})=kA\vec{X}=k\lambda\vec{X}=\lambda(k\vec{X}),$$

若 $k\neq 0$, 则 $k\vec{X}$ 是A 属于特征值λ的特征向量.

2. 设 $\vec{X}_1, \vec{X}_2$ 是 A 属于特征值 λ 的特征向量, 即 $A\vec{X}_1=\lambda\vec{X}_1$, $A\vec{X}_2=\lambda\vec{X}_2$,

则 $A(\vec{X}_1+\vec{X}_2)=A\vec{X}_1+A\vec{X}_2=\lambda\vec{X}_1+\lambda\vec{X}_2=\lambda(\vec{X}_1+\vec{X}_2)$, 若 $\vec{X}_1+\vec{X}_2\neq\vec{0}$, 则它为 A 的属于特征值 λ 的特征向量.

由这两条性质, A 属于特征值 λ 的特征向量的任意非零线性组合仍是属于 λ 的特征向量, 加上零向量就构成 $\mathbb{R}^n$ 的一个子空间.

把上述讨论总结一下，我们可如下定义.

定义3 设 A为n阶方阵, λ是 A 的一个特征值, 则称

$$V_\lambda := \{\vec{X}\in \mathbb{R}^n \mid A\vec{X} = \lambda\vec{X}\}$$

为 A 的关于特征值 λ的**特征子空间(eigen-subspace)**, 其维数称为特征值 λ 的**几何重数(geometric multiplicity)**.

注：(i) 特征子空间V_λ中并不全都是特征向量，零向量除外.

(ii) 几何重数 = dim V_λ =dim N$(\lambda I-A)$ $= n -$ r$(\lambda I-A) \geq 1$.

至此，我们回到了本章初提出的几个基本问题中的结构问题，即属于A的同一个特征值λ的特征向量再加上零向量构成$\mathbb{R}^n$的一个子空间；

自然地，属于A的不同特征值的特征向量有何关系呢？

特征向量的相互关系

定理2 属于A的不同特征值的特征向量是线性无关的.

证明 设$\lambda_1, \lambda_2, \ldots, \lambda_s$为$A$的$s$个互不相同的特征值 $(s \le n)$, $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_s$为相应的特征向量，即 $A\vec{X}_i = \lambda_i \vec{X}_i$, $i = 1, 2, \ldots, s$.

对不同特征值的个数s做归纳法.

当s=1时，由$\vec{X}_1 \neq \vec{0}$, 结论成立. 假设$s-1$个特征值时, 结论成立, 下考虑s的情形, 设

$$k_1\vec{X}_1 + k_2\vec{X}_2 + \ldots + k_s\vec{X}_s = \vec{0}. \tag{1}$$

用A左乘(1)式两边, 得

$$k_1A\vec{X}_1 + k_2A\vec{X}_2 + \ldots + k_sA\vec{X}_s = \vec{0}.$$

$$\Rightarrow\ k_1\lambda_1\vec{X}_1 + k_2\lambda_2\vec{X}_2 + \ldots + \underline{k_s\lambda_s\vec{X}_s} = \vec{0}. \tag{2}$$

用λ_s左乘(1)式两边, 得

$$k_1\lambda_s\vec{X}_1 + k_2\lambda_s\vec{X}_2 + \ldots + \underline{k_s\lambda_s\vec{X}_s} = \vec{0}. \tag{3}$$

(2), (3)两式相减, 得

$$k_1(\lambda_1-\lambda_s)\vec{X}_1+k_2(\lambda_2-\lambda_s)\vec{X}_2+\ldots+k_{s-1}(\lambda_{s-1}-\lambda_s)\vec{X}_{s-1}=\vec{0}.$$

由归纳假设知, $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_{s-1}$ 是线性无关的, 于是有

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

由已知条件$\lambda_i\neq\lambda_s$, $i=1, 2, \ldots, s-1$, 得到

$$k_i=0\ \ (i=1, 2, \ldots, s-1).$$

再代回到(1)式, 得

$$k_s\vec{X}_s=\vec{0}$$

又因$\vec{X}_s$为特征向量, $\vec{X}_s\neq\vec{0}$, 所以有 $k_s=0$.

由于(1)式中所有组合系数均为0, 故 $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_s$ 线性无关, 定理2结论成立. ■

把几何重数m_i考虑上, 定理2可以推广为以下结论:

定理3 设$\lambda_1, \lambda_2, \dots, \lambda_s$为$A$的$s$个互异的特征值, 且

$\underbrace{\vec{X}_{11}, \vec{X}_{12}, \dots, \vec{X}_{1m_1}}_{\lambda_1}$ (m_1个线性无关)

$\underbrace{\vec{X}_{21}, \vec{X}_{22}, \dots, \vec{X}_{2m_2}}_{\lambda_2}$ (m_2个线性无关)

……

$\underbrace{\vec{X}_{s1}, \vec{X}_{s2}, \dots, \vec{X}_{sm_s}}_{\lambda_s}$ (m_s个线性无关)

⇓

$\sum_{i=1}^{s} m_i$个特征向量线性无关

本讲小结1

- 从矩阵高次幂的计算说起
- 特征值与特征向量的概念 —— 绑定出现，从属关系
- 特征值与特征向量的计算—— 特征值计算即特征多项式求根，特征向量计算即求解对应线性方程组

$A\vec{X}=\lambda\vec{X}\ (\vec{X}\neq\vec{0})$ ⟺ $(\lambda I-A)\vec{X}=\vec{0}$ 有非零解 ⟺ 特征多项式 $f_A(\lambda)=|\lambda I-A|=0$

(1) (2) (3)

本讲小结2

- 特征多项式系数的性质 —— $g(A)$的特征值为$g(\lambda)$.

  $\mathrm{tr}A=\sum_{i=1}^{n}\lambda_i;\quad \det A=\prod_{i=1}^{n}\lambda_i.\quad \lambda_i$的代数重数$n_i$

- 特征子空间 —— $V_{\lambda_i}:=\{\vec{X}\in\mathbb{R}^n \mid A\vec{X}=\lambda_i\vec{X}\};$

  $\dim V_{\lambda_i}=\lambda_i$的几何重数$m_i$

- 特征向量的相互关系 ——不同特征值下的特征向量线性无关