线性代数

张神星 (合肥工业大学)

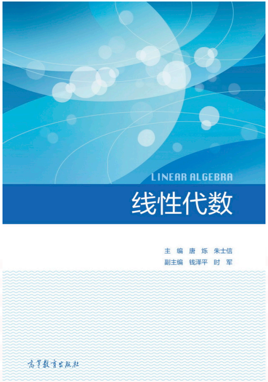
办公室: 翡翠科教楼 *B1810* 东

Email: zhangshenxing@hfut.edu.cn

课件地址: https://zhangshenxing.github.io

课程信息

- 课时:
  - 10 周 40 课时;
  - 2024-09-18 $\sim$ 2024-11-26
- 课程 QQ 群 (入群答案 1400071B)
  - 003 班 (自动化) **973042523**
  - 004 班 (电气) **980820998**
- 教材: 唐烁 朱士信《线性代数》
- 本课程讲授顺序与教材有所不同

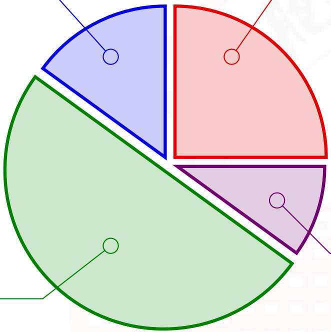


作业 15 分

作业为配套练习册，每两周交一次．作业不允许迟交．没带的请当天联系助教补交，迟一天交 −50% 当次作业分，迟两天或以上 *0* 分．请假需提前交给我请假条．

课堂测验 25 分

课堂测验共 *3* 次，取最高的两次平均．测验范围和时间会提前通知．测验时在教室内作答，否则按未考处理．



期末考试 50 分

期末卷面需要达到 *45* 分才计算总评分数，*45* 分以下直接不及格．

期末报告 10 分

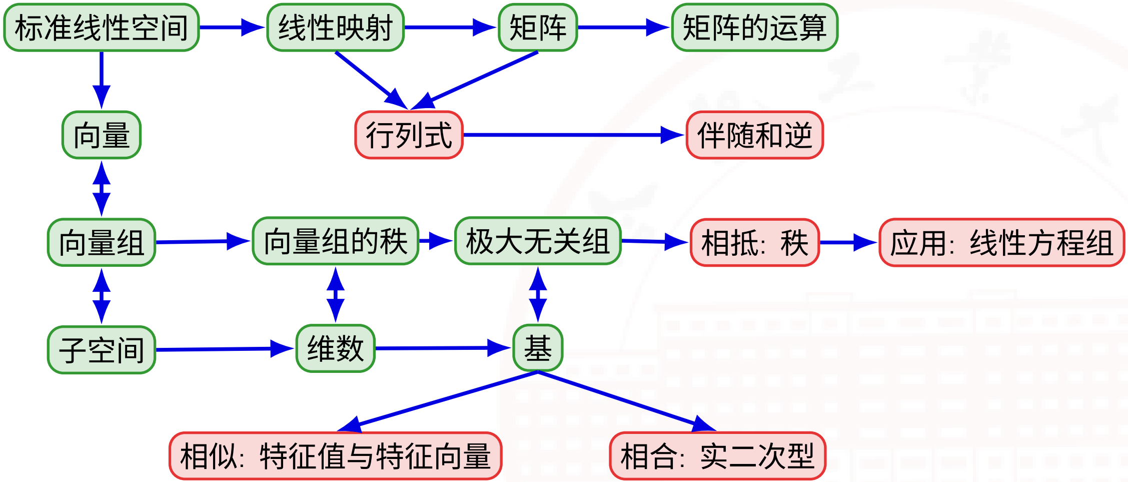
期末之前会告知主题．请交手写纸质版，并自行留存电子版本以免意外丢失．

线性代数是一门利用代数方法研究线性方程、线性空间、线性变换等线性结构的课程.

线性代数通过从具体的、几何化的观念出发, 抽象出一套代数化的方法, 从而避免了高维情形缺乏几何直观的问题. 如同微积分中“以直代曲”思想引出导数、切线、积分等一系列概念, 线性代数利用“以直代曲”思想将许多非线性问题的处理转化为线性问题, 非线性模型近似为线性模型等.

这些内容在统计学、密码学、运筹学、物理学、工程学、管理学、信息学、计算机科学等很多领域有着广泛的应用. 我们不在此处逐一列举, 在之后的授课中我们会见到它的各种应用.

课程内容关系


标准线性空间
线性映射
矩阵
矩阵的运算
向量
行列式
伴随和逆
向量组
向量组的秩
极大无关组
相抵：秩
应用：线性方程组
子空间
维数
基
相似：特征值与特征向量
相合：实二次型

课程学习方法

课前
预习课本

课上
认真听课
记好笔记

课后
过一遍教材
与课上知识点

作业
检测学
习效果

核心
理解抽象概念
掌握计算方法

第一章　向量和矩阵

第一节　向量和矩阵的定义

- 向量和向量空间
- 线性映射
- 矩阵

我们在高中学习过向量的概念. 在平面上建立一个直角坐标系. 对于平面上的点 A, 连接 OA 的有向线段就是一个**向量**. 它可以用 $\boldsymbol{u}=(x,y)$ 来表示.

由于向量和平面上的点是一一对应的, 因此我们可以用 $\mathbb{R}^2$ 来表示平面上所有向量形成的集合. 在这个集合中有一个特殊的元素, 叫作**零向量**:

$$\mathbf{0}=(0,0),$$

而且我们可以定义加法和数乘:

$$(x_1,y_1)+(x_2,y_2)=(x_1+x_2,y_1+y_2),\quad \lambda(x,y)=(\lambda x,\lambda y),\lambda\in\mathbb{R}.$$

类似地, 立体空间中的所有向量形成集合 $\mathbb{R}^3$. 在这个集合中也有零向量、加法和数乘.

我们总使用粗小写字母 $\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{x}, \boldsymbol{y}, \ldots$ 表示向量.

$\mathbb{R}^2$ 或 $\mathbb{R}^3$ 上的零向量、加法和数乘满足:

(V1) $\boldsymbol{\alpha} + \boldsymbol{\beta} = \boldsymbol{\beta} + \boldsymbol{\alpha}$;

(V2) $\boldsymbol{\alpha} + (\boldsymbol{\beta} + \boldsymbol{\gamma}) = (\boldsymbol{\alpha} + \boldsymbol{\beta}) + \boldsymbol{\gamma}$;

(V3) $\mathbf{0} + \boldsymbol{\alpha} = \boldsymbol{\alpha}$;

(V4) 对任意 $\boldsymbol{\alpha}$, 存在 $\boldsymbol{\beta}$ 使得 $\boldsymbol{\alpha} + \boldsymbol{\beta} = \mathbf{0}$. 称 $\boldsymbol{\beta}$ 为 $\boldsymbol{\alpha}$ 的**负向量**;

(V5) $(\lambda\mu)\boldsymbol{\alpha} = \lambda(\mu\boldsymbol{\alpha})$;

(V6) $(\lambda + \mu)\boldsymbol{\alpha} = \lambda\boldsymbol{\alpha} + \mu\boldsymbol{\alpha}$;

(V7) $\lambda(\boldsymbol{\alpha} + \boldsymbol{\beta}) = \lambda\boldsymbol{\alpha} + \lambda\boldsymbol{\beta}$;

(V8) $1 \cdot \boldsymbol{\alpha} = \boldsymbol{\alpha}$.

将上述概念稍作推广，我们可以得到具有 n 个**分量**的向量

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n),$$

称为 n 维**行向量**. 出于后续描述的统一性，我们将向量的分量写成一列而不是一行，从而得到 n 维**列向量**:

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

之后凡是提到向量，均是指列向量. 为了书写方便，也可以把列向量写成

$$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}.$$

其中 T 表示**转置**(Transpose)，它是将行列关系对换的一种操作.

将全体 n 维列向量记为 $\mathbb{R}^n$. 在这个集合中可类似定义零向量、加法和数乘:

(1) $\mathbf{0}=(0,\dots,0)^{\mathrm{T}}$;

(2) $(a_1,\dots,a_n)^{\mathrm{T}}+(b_1,\dots,b_n)^{\mathrm{T}}=(a_1+b_1,\dots,a_n+b_n)^{\mathrm{T}}$;

(3) $\lambda(a_1,\dots,a_n)^{\mathrm{T}}=(\lambda a_1,\dots,\lambda a_n)^{\mathrm{T}},\lambda\in\mathbb{R}$.

而且它们也满足(V1)–(V8).

定义

若集合 V 带有零向量、加法和实数的数乘, 且满足(V1)–(V8), 则称 V 是一个实线性空间.

这样, $\mathbb{R}^n$ 成为了一个实线性空间.

例

(1) 平面上经过原点的直线 $V=\{(x,y)\in\mathbb{R}^2 \mid ax+by=0\}$ 是实线性空间.

(2) $\mathbb{R}^3$ 中经过原点的平面 $V=\{(x,y,z)\in\mathbb{R}^3 \mid ax+by+cz=0\}$ 是实线性空间.

(3) 如果 V 是 $\mathbb{R}^n$ 的子集, 且它和继承自 $\mathbb{R}^n$ 的零向量、加法和数乘构成线性空间, 则称 V 是**向量空间**.

(4) 实系数多项式全体 $\mathbb{R}[x]$ 是实线性空间, 其中的零向量是指零多项式.

(5) 次数不超过 n 的多项式 (含零多项式) 全体是实线性空间.

若数乘的系数是复数而不是实数, 称 V 为**复线性空间**. 我们可以类似地定义复向量, 且 n 维复向量全体形成一个复线性空间 $\mathbb{C}^n$.

考虑如下映射

$$
\begin{aligned}
\mathscr{A}:\mathbb{R}^n &\longrightarrow \mathbb{R}^m\\
\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n \end{pmatrix} &\longmapsto \begin{pmatrix} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n\\ \vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n \end{pmatrix}.
\end{aligned}
$$

不难验证，对任意 $\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^n,\lambda\in\mathbb{R}$,

(L1) $\mathscr{A}(\boldsymbol{u}+\boldsymbol{v})=\mathscr{A}(\boldsymbol{u})+\mathscr{A}(\boldsymbol{v})$;

(L2) $\mathscr{A}(\lambda\boldsymbol{u})=\lambda\mathscr{A}(\boldsymbol{u})$.

定义

若映射 $\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^m$ 满足(L1),(L2), 称 $\mathscr{A}$ 是一个**线性变换**或**线性映射**.

反过来, 是不是所有的线性变换 $\mathbb{R}^n\to\mathbb{R}^m$ 都可以表达为前述形式呢?

以线性映射 $\mathscr{A}:\mathbb{R}^3\to\mathbb{R}^4$ 为例, 记

$$\mathscr{A}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\\a_{41}\end{pmatrix},\qquad \mathscr{A}\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}a_{12}\\a_{22}\\a_{32}\\a_{42}\end{pmatrix},\qquad \mathscr{A}\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}a_{13}\\a_{23}\\a_{33}\\a_{43}\end{pmatrix}.$$

根据线性变换的性质,

$$\mathscr{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \mathscr{A}\left(x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = x_1 \mathscr{A}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \mathscr{A}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \mathscr{A}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ a_{42} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \\ a_{43} \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 \end{pmatrix}.$$

因此每个线性变换 $\mathbb{R}^4 \to \mathbb{R}^3$ 都具有这种形式, 它完全由 4×3 个系数

$$(a_{ij})_{1 \leqslant i \leqslant 4, 1 \leqslant j \leqslant 3}$$

所确定. 由此引出矩阵的定义.

定义

将 mn 个数按照每行 n 个元素，每列 m 个元素，排成的数表称为 m 行 n 列矩阵，或简称为 $m\times n$ 矩阵：

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}_{m\times n},$$

其中 a_{ij} 表示 A 的第 i 行 j 列元素，并记 $A=(a_{ij})_{m\times n}$.

我们总使用粗大写字母 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{\Lambda},\ldots$ 表示矩阵. 不强调矩阵的阶时，也可省略右下角 $m\times n$.

矩阵的圆括号也可用方括号来代替.

例

(1) $\boldsymbol{A}=\begin{pmatrix}0 & -1 & -2 & -3\\ 1 & 0 & -1 & -2\\ 2 & 1 & 0 & -1\\ 3 & 2 & 1 & 0\end{pmatrix}=(i-j)_{4\times 4}.$

(2) $\boldsymbol{A}=\begin{pmatrix}1 & 1 & \cdots & 1\\ 1 & 2 & \cdots & n\\ \vdots & \vdots & \ddots & \vdots\\ 1 & 2^{n-1} & 3^{n-1} & n^{n-1}\end{pmatrix}_n=(j^{i-1})_n.$ $m=n$ 时可用 n 来表示右下角的 $n\times n$.

线性映射与矩阵一一对应

- 令 $M_{m\times n}$ 表示 m 行 n 列的矩阵 (Matrix) 全体.
- 若 $m=n$, 称相应矩阵为 n **阶方阵**, 并用 M_n 来表示 n 阶方阵全体.
- 若要强调矩阵元素都是实数, 我们称相应的矩阵为**实矩阵**, 并用 $M_{m\times n}(\mathbb{R}), M_n(\mathbb{R})$ 来表示相应集合.
- 全体线性变换 $\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^m$ 和 $\boldsymbol{A}\in M_{m\times n}(\mathbb{R})$ 是一一对应的.

类似可定义**复矩阵**, $M_{m\times n}(\mathbb{C})$, $M_n(\mathbb{C})$, 且全体线性变换 $\mathbb{C}^n\to\mathbb{C}^m$ 和 $M_{m\times n}(\mathbb{C})$ 一一对应. 之后此种类比不再赘述.

- 称方阵的对角线上的元素 $a_{11},\dots,a_{nn}$ 为方阵的**对角元**.

元素全为零的矩阵为**零矩阵** $\boldsymbol{O} = \boldsymbol{O}_{m\times n} \in M_{m\times n}$. 方阵中

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ & a_{22} & \cdots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \quad \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} \in M_n$$

分别为**上三角阵**, **下三角阵**和**对角阵** (空白部分表示元素都是零). 为书写方便, 对角阵也可记作 (对角: diagonal)

$$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

称方阵

$$\boldsymbol{E} = \boldsymbol{E}_n = \mathrm{diag}(1, 1, \ldots, 1) \in M_n$$

为**单位阵**.

将只有一行的矩阵

$$(a_1, a_2, \ldots, a_n) \in M_{1\times n}$$

称为 n 维行矩阵. 只有一列的矩阵

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in M_{n\times 1}$$

称为 n 维列矩阵.

我们将在后面定义矩阵的加法和数乘, 它与将行 (列) 矩阵看作行 (列) 向量的加法和数乘是相同的. 于是可将行 (列) 矩阵和对应的行 (列) 向量视为同一对象.

如何用矩阵表示平面 $\mathbb{R}^2$ 上的旋转？设 $A(x_1, x_2)$ 是平面上的一个点，沿着原点逆时针旋转角度 θ 变成 $B(y_1, y_2)$。利用极坐标将 A 表示为

$$\begin{cases} x_1 = \rho\cos\alpha, \\ x_2 = \rho\sin\alpha, \end{cases}$$

则

$$\begin{cases} y_1 = \rho\cos(\alpha+\theta) = \rho(\cos\alpha\cos\theta - \sin\alpha\sin\theta) = (\cos\theta)x_1 - (\sin\theta)x_2, \\ y_2 = \rho\sin(\alpha+\theta) = \rho(\cos\alpha\sin\theta + \sin\alpha\cos\theta) = (\sin\theta)x_1 + (\cos\theta)x_2. \end{cases}$$

因此上述**旋转变换** $\mathscr{R}_\theta$ 对应的矩阵为

$$\boldsymbol{R}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in M_2(\mathbb{R}).$$

例：线性变换

- $\boldsymbol{A}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$ 表示各个分量分别放大为 λ_i 倍的线性变换.
- $\boldsymbol{A}=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$ 表示平面中沿着直线 $x_1=x_2$ 翻转.
- $\boldsymbol{A}=\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}$ 表示三维空间中沿着直线 $x_1=x_2=x_3$ 旋转 $\dfrac{2\pi}{3}$.
- 想一想：$\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1\end{pmatrix}$ 表示什么线性变换?

第二节　矩阵的线性运算、乘法和转置

- 矩阵的线性运算
- 矩阵的乘法
- 矩阵的幂
- 矩阵的转置

给定线性变换 $\mathscr{A}, \mathscr{B} : \mathbb{R}^n \to \mathbb{R}^m$ 和数 $\lambda \in \mathbb{R}$, 定义

$$(\mathscr{A} + \mathscr{B})(\boldsymbol{x}) = \mathscr{A}(\boldsymbol{x}) + \mathscr{B}(\boldsymbol{x}), \qquad (\lambda f)(\boldsymbol{x}) = \lambda(\mathscr{A}(\boldsymbol{x})).$$

对任意 $\boldsymbol{x} \in \mathbb{R}^n$,

$$(\mathscr{A} + \mathscr{B})(\boldsymbol{x}) = \mathscr{A}(\boldsymbol{x}) + \mathscr{B}(\boldsymbol{x}) \xlongequal{\text{(V1)}} \mathscr{B}(\boldsymbol{x}) + \mathscr{A}(\boldsymbol{x}) = (\mathscr{B} + \mathscr{A})(\boldsymbol{x}).$$

故 (1) $\mathscr{A} + \mathscr{B} = \mathscr{B} + \mathscr{A}$. 同理

(2) $(\mathscr{A} + \mathscr{B}) + \mathscr{C} = \mathscr{A} + (\mathscr{B} + \mathscr{C})$;

(3) $\mathscr{A} + \mathscr{O} = \mathscr{A}$, 这里 $\mathscr{O}$ 表示零映射: $\mathscr{O}(\boldsymbol{x}) = \mathbf{0}$;

(4) 对任意线性映射 $\mathscr{A}$, 存在线性映射 $\mathscr{B} = (-1)\mathscr{A}$ 使得 $\mathscr{A} + \mathscr{B} = \mathscr{O}$;

(5) $(\lambda\mu)\mathscr{A} = \lambda(\mu\mathscr{A}) = \mu(\lambda\mathscr{A})$;

(6) $(\lambda + \mu)\mathscr{A} = \lambda\mathscr{A} + \mu\mathscr{A}$;

(7) $\lambda(\mathscr{A} + \mathscr{B}) = \lambda\mathscr{A} + \lambda\mathscr{B}$;

(8) $1 \cdot \mathscr{A} = \mathscr{A}$, $0 \cdot \mathscr{A} = \mathscr{O}$,$\lambda \cdot \mathscr{O} = \mathscr{O}$.

由此得到对应的矩阵的加法和数乘:

定义

设 $\boldsymbol{A}=(a_{ij})_{m\times n}, \boldsymbol{B}=(b_{ij})_{m\times n}$, $\lambda\in\mathbb{R}$. **定义**

$$\boldsymbol{A}+\boldsymbol{B}=(a_{ij}+b_{ij})_{m\times n}, \qquad (\lambda f)(\boldsymbol{x})=\lambda(f(\boldsymbol{x})).$$

只有同型矩阵(行列数都相同的矩阵) 才能相加. 行 (列) 矩阵的加法和数乘就是其对应的行 (列) 向量的加法和数乘.

自然地, 矩阵的减法为

$$\boldsymbol{A}-\boldsymbol{B}=(a_{ij}-b_{ij})_{m\times n}.$$

矩阵加法的性质

由线性映射的相应性质可知:

(1) $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{B}+\boldsymbol{A}$;

(2) $(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C}=\boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C})$;

(3) $\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$;

(4) 对于任意矩阵 $\boldsymbol{A}$, 存在矩阵 $\boldsymbol{B}=(-1)\boldsymbol{A}$ 使得 $\boldsymbol{A}+\boldsymbol{B}=\boldsymbol{O}$. 称 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的负矩阵, 记作 $-\boldsymbol{A}$.

(5) $(\lambda\mu)\boldsymbol{A}=\lambda(\mu\boldsymbol{A})=\mu(\lambda\boldsymbol{A})$;

(6) $(\lambda+\mu)\boldsymbol{A}=\lambda\boldsymbol{A}+\mu\boldsymbol{A}$;

(7) $\lambda(\boldsymbol{A}+\boldsymbol{B})=\lambda\boldsymbol{A}+\lambda\boldsymbol{B}$;

(8) $1\cdot\boldsymbol{A}=\boldsymbol{A}, 0\cdot\boldsymbol{A}=\boldsymbol{O}, \lambda\boldsymbol{O}=\boldsymbol{O}$.

一张图片由一些像素构成, 上图包含 341×512 个像素. 在 RGB (红绿蓝) 颜色模式下, 每个像素包含红绿蓝三个通道, 每个通道为一个 $0 \sim 255$ 之间的整数, 数值越高对应颜色越亮. 例如

- $R = 0$, $G = 0$, $B = 0$ 表示纯黑色;
- $R = 255$, $G = 255$, $B = 255$ 表示纯白色;
- $R = 255$, $G = 0$, $B = 255$ 表示纯紫色.

若三个通道相同, 图片就是一张灰色的图. 此时图片对应一个 341×512 的矩阵 A.

矩阵线性运算的应用：图像处理

想一想：如何将这个图像变亮？

$\Longrightarrow$

我们只需要增加每个元素的值，例如 (超过 255 的需要修正为 255)

$$\boldsymbol{A}+\begin{pmatrix}50 & \cdots & 50\\ \vdots & \ddots & \vdots\\ 50 & \cdots & 50\end{pmatrix}, \qquad 1.5\boldsymbol{A}, \qquad \frac{1}{2}\left(\boldsymbol{A}+\begin{pmatrix}255 & \cdots & 255\\ \vdots & \ddots & \vdots\\ 255 & \cdots & 255\end{pmatrix}\right).$$

如何让图像反色?

$\Longrightarrow$

$$\begin{pmatrix} 255 & \cdots & 255 \\ \vdots & \ddots & \vdots \\ 255 & \cdots & 255 \end{pmatrix} - \boldsymbol{A}.$$

设线性映射

$$\mathscr{A}:\mathbb{R}^n \to \mathbb{R}^m,\quad \mathscr{B}:\mathbb{R}^p \to \mathbb{R}^n,\quad \mathscr{C}=\mathscr{A}\circ\mathscr{B}:\mathbb{R}^p \to \mathbb{R}^m$$

对应的矩阵为 $\boldsymbol{A}=(a_{ij})_{m\times n}, \boldsymbol{B}=(b_{ij})_{n\times p}, \boldsymbol{C}=(c_{ij})_{m\times p}$. 如何用 $\boldsymbol{A},\boldsymbol{B}$ 来表示 $\boldsymbol{C}$ 呢?

设 $\boldsymbol{x}=(x_1,\ldots,x_p)^{\mathrm{T}}\in\mathbb{R}^p$, 则

$$\mathscr{B}(\boldsymbol{x})=(y_1,\ldots,y_n)^{\mathrm{T}},\quad y_k=\sum_{j=1}^{p} b_{kj}x_j.$$

$$\mathscr{A}(\mathscr{B}(\boldsymbol{x}))=(z_1,\ldots,z_m)^{\mathrm{T}},\quad z_i=\sum_{k=1}^{n} a_{ik}y_k=\sum_{k=1}^{n} a_{ik}\sum_{j=1}^{p} b_{kj}x_j=\sum_{j=1}^{p}\Big(\sum_{k=1}^{n} a_{ik}b_{kj}\Big)x_j.$$

$$\Longrightarrow\ \boldsymbol{C}=(c_{ij})_{m\times p},\qquad c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}.$$

我们把它定义为矩阵的乘法 $\boldsymbol{C}=\boldsymbol{AB}$.

矩阵乘法的定义

定义

设 $\boldsymbol{A}=(a_{ij})_{m\times n}, \boldsymbol{B}=(b_{ij})_{n\times p}$. 定义矩阵的乘法为 $\boldsymbol{C}=\boldsymbol{AB}=(c_{ij})_{m\times p}$, 其中

$$c_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}.$$

只有第一个矩阵的列数等于第二个矩阵的行数才能相乘.

$$\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} =?$$

不能相乘!

简单来说, $\boldsymbol{AB}$ 的 (i,j) 元就是 $\boldsymbol{A}$ 的 i 行和 $\boldsymbol{B}$ 的 j 列对应分量相乘后求和得到的.

行向量与列向量的乘法

设 $\boldsymbol{A}=(a_1,\ldots,a_n)$ 是 n 维行向量, $\boldsymbol{B}=(b_1,\ldots,b_n)^{\mathrm{T}}$ 是 n 维列向量. $\boldsymbol{AB},\boldsymbol{BA}=?$

$$\boldsymbol{AB}=\sum_{i=1}^{n}a_ib_i,\qquad \boldsymbol{BA}=(b_ia_j)_{n\times n}\in M_n.$$

对于矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n},\boldsymbol{B}=(b_{ij})_{n\times p}$. $\boldsymbol{AB}$ 的 (i,j) 元其实就是 $\boldsymbol{A}$ 第 i 行对应的行向量和 $\boldsymbol{B}$ 第 j 列对应的列向量相乘得到的数 (1 阶方阵包含的唯一元素):

$$\begin{pmatrix}\boldsymbol{\alpha}_1\\ \boldsymbol{\alpha}_2\\ \vdots\\ \boldsymbol{\alpha}_m\end{pmatrix}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_p)=\begin{pmatrix}\boldsymbol{\alpha}_1\boldsymbol{\beta}_1 & \boldsymbol{\alpha}_1\boldsymbol{\beta}_2 & \cdots & \boldsymbol{\alpha}_1\boldsymbol{\beta}_p\\ \boldsymbol{\alpha}_2\boldsymbol{\beta}_1 & \boldsymbol{\alpha}_2\boldsymbol{\beta}_2 & \cdots & \boldsymbol{\alpha}_2\boldsymbol{\beta}_p\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{\alpha}_m\boldsymbol{\beta}_1 & \boldsymbol{\alpha}_m\boldsymbol{\beta}_2 & \cdots & \boldsymbol{\alpha}_m\boldsymbol{\beta}_p\end{pmatrix}.$$

例：矩阵乘法的计算

例

求矩阵 $\boldsymbol{A}=\begin{pmatrix}1 & 2 & 0 & -1\\ 2 & 1 & 4 & 0\end{pmatrix}$ 与 $\boldsymbol{B}=\begin{pmatrix}2 & 0 & 1\\ -2 & 3 & 1\\ 1 & 5 & 0\\ 1 & -3 & 4\end{pmatrix}$ 的乘积 $\boldsymbol{AB}$.

解

$$\begin{pmatrix}1 & 2 & 0 & -1\\ 2 & 1 & 4 & 0\end{pmatrix}\begin{pmatrix}2 & 0 & 1\\ -2 & 3 & 1\\ 1 & 5 & 0\\ 1 & -3 & 4\end{pmatrix}=\begin{pmatrix}-3 & 9 & -1\\ 6 & 23 & 3\end{pmatrix}.$$

矩阵乘法的性质

线性映射满足如下性质:

(1) $(\mathscr{A}\circ\mathscr{B})\circ\mathscr{C}=\mathscr{A}\circ(\mathscr{B}\circ\mathscr{C})$;

(2) $\lambda(\mathscr{A}\circ\mathscr{B})=(\lambda\mathscr{A})\circ\mathscr{B}=\mathscr{A}\circ(\lambda\mathscr{B})$;

(3) $\mathscr{A}\circ(\mathscr{B}+\mathscr{C})=\mathscr{A}\circ\mathscr{B}+\mathscr{A}\circ\mathscr{C}$;

(4) $\mathscr{E}_m\circ\mathscr{A}=\mathscr{A}\circ\mathscr{E}_n=\mathscr{A}$, 这里 $\mathscr{E}$ 表示恒等映射: $\mathscr{E}(\boldsymbol{x})=\boldsymbol{x}$;

(5) $\mathscr{O}\circ\mathscr{A}=\mathscr{A}\circ\mathscr{O}=\mathscr{O}$.

由此可知矩阵乘法满足如下性质:

(1) $(\boldsymbol{AB})\boldsymbol{C}=\boldsymbol{A}(\boldsymbol{BC})$;

(2) $\lambda(\boldsymbol{AB})=(\lambda\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(\lambda\boldsymbol{B})$;

(3) $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}$;

(4) 若 $\boldsymbol{A}\in M_{m\times n}$, 则 $\boldsymbol{E}_m\boldsymbol{A}=\boldsymbol{A}\boldsymbol{E}_n=\boldsymbol{A}$.

(5) 若 $\boldsymbol{A}\in M_{m\times n}$, 则 $\boldsymbol{O}_{p\times m}\boldsymbol{A}=\boldsymbol{O}_{p\times n}$, $\boldsymbol{A}\boldsymbol{O}_{n\times p}=\boldsymbol{O}_{m\times p}$.

矩阵 $\boldsymbol{A}$ 对应的线性变换就是

$$\mathscr{A}: \mathbb{R}^n \to \mathbb{R}^m,\ \boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}.$$

因此解线性方程组

$$\begin{cases} a_{11}\ x_1 + a_{12}\ x_2 + \cdots + a_{1n}\ x_n = b_1 \\ a_{21}\ x_1 + a_{22}\ x_2 + \cdots + a_{2n}\ x_n = b_2 \\ \qquad\qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

等价于解矩阵方阵 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, 其中

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

矩阵乘法无交换律和消去律

矩阵的乘法不能随意交换顺序. 一般称 AB 为 A 左乘 B, 或者 B 右乘 A.

若 $AB = BA$, 则称 A, B 是可交换的. 此时 A, B 必为同阶方阵. 例如

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

方阵总和同阶单位阵交换.

矩阵乘法也没有消去律: $AB = O$ 推不出 $A = O$ 或 $B = O$. 例如

$$\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix} = O_2.$$

由此可知: $AC = BC$ 推不出 $A = B$.

练习

设 A, B 为 $n > 1$ 阶方阵, 则 $A + AB =$(C).

(A) $A(1 + B)$ (B) $(E + B)A$ (C) $A(E + B)$ (D) 以上都不对

例：与给定矩阵可交换

例

求与矩阵 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ & 0 & 1 \\ & & 0 \end{pmatrix}$ 可交换的所有矩阵.

解

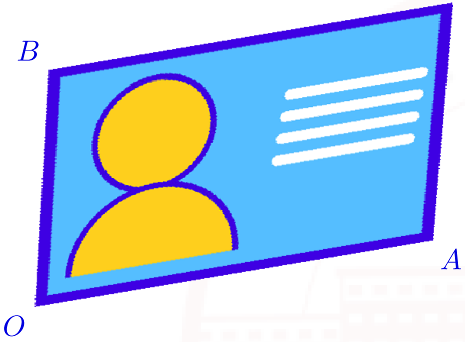
设 $\boldsymbol{B} = (a_{ij})_{3\times 3}$ 与 $\boldsymbol{A}$ 可交换，则

$$\boldsymbol{AB} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ 0 & 0 & 0 \end{pmatrix} = \boldsymbol{BA} = \begin{pmatrix} 0 & a_{11} & a_{12} \\ 0 & a_{21} & a_{22} \\ 0 & a_{31} & a_{32} \end{pmatrix},$$

$$a_{21} = a_{31} = a_{32} = 0, \quad a_{11} = a_{22} = a_{33}, \quad a_{23} = a_{12},$$

即 $\boldsymbol{B} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ & a_{11} & a_{12} \\ & & a_{11} \end{pmatrix}$.

某位同学拍身份证照片拍成了下图的样子，如何才能修复好呢？



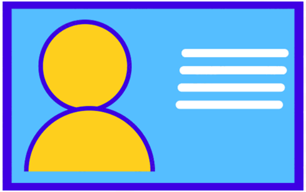
以左下角为原点，通过测量发现 A 坐标为 $(521, 88)$，B 坐标为 $(19, 311)$.

经过查询知道身份证长宽比为 $42.7 : 27$. 令 $A' = (427, 0), B' = (0, 270)$. 我们希望找到一个线性变换，将 A, B 变为 A', B'.

设该线性变换对应的矩阵为 $\boldsymbol{A}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$，则

$$\boldsymbol{A}\begin{pmatrix} 521 & 19 \\ 88 & 311 \end{pmatrix}=\begin{pmatrix} 427 & 0 \\ 0 & 270 \end{pmatrix},\qquad \text{即}\quad \begin{cases} 521a+88b=427 \\ 19a+311b=0 \\ 521c+88d=0 \\ 19c+311d=270 \end{cases}$$

解得 $\boldsymbol{A}=\begin{pmatrix} 0.828 & -0.051 \\ -0.148 & 0.877 \end{pmatrix}$. 通过应用该变换，图片被修复成如下效果：



定义

设 A 为 n 阶方阵, 定义 A 的幂

$$A^0 = E_n, \quad A^k = \underbrace{A \cdot A \cdot \cdots \cdot A}_{k\text{ 个}}.$$

矩阵幂满足如下性质 (k, ℓ 为非负整数):

(1) $A^{k+\ell} = A^k \cdot A^\ell$;

(2) $A^{k\ell} = (A^k)^\ell$.

注意 $(AB)^k$ 一般不等于 $A^k \cdot B^k$. 想一想: 下面的等式成立吗?

$$(A-B)(A+B) = A^2 - B^2? \qquad (A+B)^2 = A^2 + 2AB + B^2?$$

当且仅当 $AB = BA$ 时成立.

例

设 $\boldsymbol{A}=\operatorname{diag}(\lambda_1,\cdots,\lambda_n)$. 求 $\boldsymbol{A}^k$.

解

$$\boldsymbol{A}^2=\boldsymbol{A}\cdot\boldsymbol{A}=\operatorname{diag}(\lambda_1^2,\cdots,\lambda_n^2),$$
$$\boldsymbol{A}^3=\boldsymbol{A}\cdot\boldsymbol{A}^2=\operatorname{diag}(\lambda_1^3,\cdots,\lambda_n^3),$$

递推下去可知

$$\boldsymbol{A}^k=\operatorname{diag}(\lambda_1^k,\cdots,\lambda_n^k).$$

例

设 $\boldsymbol{A}=\begin{pmatrix}\lambda & 1 & 0\\ & \lambda & 1\\ & & \lambda\end{pmatrix}$. 求 $\boldsymbol{A}^k$.

解

$$\boldsymbol{A}^2=\begin{pmatrix}\lambda & 1 & 0\\ & \lambda & 1\\ & & \lambda\end{pmatrix}\begin{pmatrix}\lambda & 1 & 0\\ & \lambda & 1\\ & & \lambda\end{pmatrix}=\begin{pmatrix}\lambda^2 & 2\lambda & 1\\ & \lambda^2 & 2\lambda\\ & & \lambda^2\end{pmatrix},$$

$$\boldsymbol{A}^3=\begin{pmatrix}\lambda & 1 & 0\\ & \lambda & 1\\ & & \lambda\end{pmatrix}\begin{pmatrix}\lambda^2 & 2\lambda & 1\\ & \lambda^2 & 2\lambda\\ & & \lambda^2\end{pmatrix}=\begin{pmatrix}\lambda^3 & 3\lambda^2 & 3\lambda\\ & \lambda^3 & 3\lambda^2\\ & & \lambda^3\end{pmatrix}.$$

归纳可知 $\boldsymbol{A}^k=\begin{pmatrix}\lambda^k & k\lambda^{k-1} & \frac{1}{2}k(k-1)\lambda^{k-2}\\ & \lambda^k & k\lambda^{k-1}\\ & & \lambda^k\end{pmatrix}$.

例：矩阵幂的计算

另解

设 $\boldsymbol{N} = \begin{pmatrix} 0 & 1 & 0 \\ & 0 & 1 \\ & & 0 \end{pmatrix}$，则 $\boldsymbol{N}^2 = \begin{pmatrix} 0 & 0 & 1 \\ & 0 & 0 \\ & & 0 \end{pmatrix}$，$\boldsymbol{N}^3 = \boldsymbol{O}$. 由于 $\boldsymbol{A} = \lambda\boldsymbol{E} + \boldsymbol{N}$ 且 $\boldsymbol{E}$ 和 $\boldsymbol{N}$ **可交换**，因此

$$\begin{aligned} \boldsymbol{A}^k &= (\lambda\boldsymbol{E} + \boldsymbol{N})^k = \lambda^k\boldsymbol{E} + \mathrm{C}_k^1\lambda^{k-1}\boldsymbol{N} + \mathrm{C}_k^2\lambda^{k-2}\boldsymbol{N}^2 \\ &= \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \frac{1}{2}k(k-1)\lambda^{k-2} \\ & \lambda^k & k\lambda^{k-1} \\ & & \lambda^k \end{pmatrix}. \end{aligned}$$

例：矩阵幂的计算

例

设 $\boldsymbol{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$. 求 $\boldsymbol{A}^k$.

解

注意到 $\boldsymbol{A}$ 对应平面 $\mathbb{R}^2$ 上的线性变换是逆时针旋转 θ，所以 $\boldsymbol{A}^k$ 就是逆时针旋转 $n\theta$，对应的矩阵为

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

例

设 $\boldsymbol{A}=(1,2,3), \boldsymbol{B}=\begin{pmatrix}-1\\2\\0\end{pmatrix}$. 求 $(\boldsymbol{BA})^k$.

解

注意到 $\boldsymbol{AB}=3$, 因此

$$(\boldsymbol{BA})^k=\boldsymbol{B}(\boldsymbol{AB})^{k-1}\boldsymbol{A}=\boldsymbol{B}\cdot 3^{k-1}\cdot\boldsymbol{A}=3^{k-1}\boldsymbol{BA}=\begin{pmatrix}-3^{k-1} & -2\cdot 3^{k-1} & -3^k\\ 2\cdot 3^{k-1} & 4\cdot 3^{k-1} & 2\cdot 3^k\\ 0 & 0 & 0\end{pmatrix}.$$

把数 λ 看成一阶方阵时，它右乘一列向量、或左乘一行向量和数乘效果一样.

该方法可用于计算每行之间成比例的方阵的幂.

练习

设 $\boldsymbol{A}=\begin{pmatrix}1&2&3\\2&4&6\\3&6&9\end{pmatrix}$. 求 $\boldsymbol{A}^k$.

答案

注意到 $\boldsymbol{A}=\begin{pmatrix}1\\2\\3\end{pmatrix}(1,2,3)$, 因此 $\boldsymbol{A}^k=14^{k-1}\boldsymbol{A}=\begin{pmatrix}14^{k-1}&2\cdot 14^{k-1}&3\cdot 14^{k-1}\\2\cdot 14^{k-1}&4\cdot 14^{k-1}&6\cdot 14^{k-1}\\3\cdot 14^{k-1}&6\cdot 14^{k-1}&9\cdot 14^{k-1}\end{pmatrix}$.

想一想：$\boldsymbol{A}^2=\boldsymbol{E}$ 能推出 $\boldsymbol{A}=\boldsymbol{E}$ 或 $-\boldsymbol{E}$ 吗？

方阵的多项式

设

$$f(x) = a_m x^m + \cdots + a_1 x + a_0$$

是一个多项式. 对于 n 阶方阵 $\boldsymbol{A}$, 定义

$$f(\boldsymbol{A}) = a_m \boldsymbol{A}^m + \cdots + a_1 \boldsymbol{A} + a_0 \boldsymbol{E}.$$

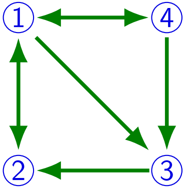
和幂的计算类似, 我们有

(1) $f\big(\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\big) = \mathrm{diag}\big(f(\lambda_1), \ldots, f(\lambda_n)\big)$.

(2) $f(\lambda \boldsymbol{E} + \boldsymbol{N}) = f(\lambda)\boldsymbol{E} + f'(\lambda)\boldsymbol{N} + \dfrac{1}{2!} f''(\lambda)\boldsymbol{N}^2 + \cdots + \dfrac{1}{(n-1)!} f^{(n-1)} \boldsymbol{N}^{n-1}$.

(3) 若 $a_0 = 0$, 则 $f(\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}) = \dfrac{f(\lambda)}{\lambda}\boldsymbol{A}$, 其中 $\lambda = \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}$.

网上订票系统里记录了所有能直飞的航班线路. 对于不能直达的城市, 该怎么确定是否有换乘方案呢? 例如 4 个城市之间的航线如图所示:



$\Longleftrightarrow$ 邻接矩阵 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$

邻接矩阵中 $a_{ij} = 1$ 表示从 i 到 j 有直飞航线.

于是 $\boldsymbol{A}^2$ 的 (i,j) 元

$$b_{ij}=\sum_{k=1}^{4}a_{ik}a_{kj}$$

就是从 i 到 j 换乘一次的方案数. 例如从②$\implies$③:

$$\boldsymbol{A}^2=\begin{pmatrix}0&1&1&1\\1&0&0&0\\0&1&0&0\\1&0&1&0\end{pmatrix}\begin{pmatrix}0&1&1&1\\1&0&0&0\\0&1&0&0\\1&0&1&0\end{pmatrix}=\begin{pmatrix}2&1&1&0\\0&1&\boxed{1}&1\\1&0&0&0\\0&2&1&1\end{pmatrix}.$$

由于 $b_{23}=1$, 因此可通过② $\implies$ ① $\implies$ ③换乘一次到达.

想一想: 如何从③到达④? 考虑 $\boldsymbol{A}^3$, 即换乘两次即可.

矩阵转置的定义

若 $\boldsymbol{A}=(a_{ij})_{m\times n}$, 称

$$\boldsymbol{A}^{\mathrm{T}}=(a_{ji})_{n\times m}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}\in M_{n\times m}$$

为矩阵 $\boldsymbol{A}$ 的**转置**. 如同我们之前使用过的写法, 行向量的转置就是列向量

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}=(a_1,\ldots,a_n)^{\mathrm{T}}.$$

方阵的转置还是方阵, 上三角阵的转置是下三角阵.

$\boldsymbol{A}^{\mathrm{T}}$ 对应的线性变换是 $\boldsymbol{A}$ 对应的线性变换诱导的对偶空间上的线性变换, 感兴趣的可自行阅读有关材料.

矩阵的转置满足如下性质:

(1) $(\boldsymbol{A}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}$;

(2) $(\boldsymbol{A}+\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}} + \boldsymbol{B}^{\mathrm{T}}$;

(3) $(\lambda\boldsymbol{A})^{\mathrm{T}} = \lambda\boldsymbol{A}^{\mathrm{T}}$;

(4) $(\boldsymbol{A}\boldsymbol{B})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}$.

例如

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix},$$

两边取转置得到

$$(b_1, b_2, b_3) \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = (c_1, c_2, c_3).$$

对称阵和反对称阵

定义

- 若方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$，称 $\boldsymbol{A}$ 为对称阵；
- 若 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$，称 $\boldsymbol{A}$ 为反对称阵.

例如 $\begin{pmatrix} 12 & 6 & 1 \\ 6 & 8 & 0 \\ 1 & 0 & 6 \end{pmatrix}$ 是对称阵. 对角矩阵都是对称阵.

例如 $\begin{pmatrix} 0 & 6 & 1 \\ -6 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$ 是反对称阵. 反对称阵的对角线均为 0.

例

证明：若 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{AB}$ 都是对称阵，则 $\boldsymbol{AB} = \boldsymbol{BA}$.

证明

由题设可知 $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}, \boldsymbol{B}^{\mathrm{T}} = \boldsymbol{B}$,

$$\boldsymbol{AB} = (\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{BA}.$$

□

想一想：若 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{AB}$ 中有一个对称阵和两个反对称阵呢？同理 $\boldsymbol{AB} = \boldsymbol{BA}$.

练习

设 $\boldsymbol{A}$ 是 n 阶方阵，(A)一定是对称阵？

(A) $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ (B) $\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}}$ (C) $\boldsymbol{A}^2$ (D) $\boldsymbol{A}^{\mathrm{T}} - \boldsymbol{A}$

一般地，若 $\boldsymbol{A} \in M_{m\times n}$, $\boldsymbol{AA}^{\mathrm{T}}$ 是 m 阶对称阵，$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是 n 阶对称阵.

任一方阵可表为对称阵与反对称阵之和

例

证明：任一方阵均可写成一对称阵和一反对称阵之和.

证明

$$A = \frac{A + A^{\mathrm{T}}}{2} + \frac{A - A^{\mathrm{T}}}{2}.$$

□

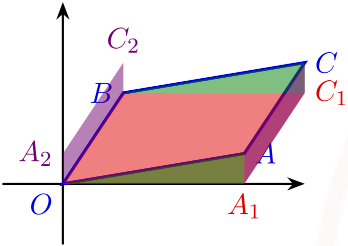
想一想:

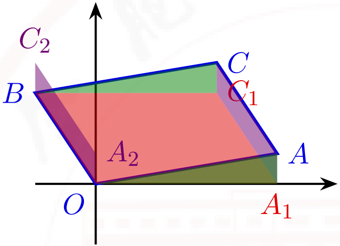
- 若函数 $f(x)$ 的定义域关于原点对称，则 $f(x)$ 可以表示成一个偶函数和一个奇函数之和.
- 复数 z 可以写成 $z_1 + z_2$，其中 $\overline{z_1} = z_1, \overline{z_2} = -z_2$.

第三节　方阵的行列式

- 行列式的定义
- 行列式的性质
- 拉普拉斯展开
- 行列式的计算举例
- 三对角和范德蒙型行列式

设平面上有 $\square OACB$，其中 A,B 坐标分别为 $\boldsymbol{u}=(a,b)^{\mathrm{T}},\boldsymbol{v}=(c,d)^{\mathrm{T}}$。如果 $\boldsymbol{u}=(1,0)^{\mathrm{T}},\boldsymbol{v}=(0,1)^{\mathrm{T}}$，那么面积为 1。如果将 $\boldsymbol{u}$ 换成 $k\boldsymbol{u}$，那么面积变为 $|k|$ 倍。如果将 $\boldsymbol{u}$ 拆分为 $(a,0)^{\mathrm{T}}+(0,b)^{\mathrm{T}}$，那么得到的三个平行四边形的面积有什么联系呢？





令 $A_1(a,0)$，并作 $\square OA_1C_1B$；令 $A_2(0,b)$，并作 $\square OA_2C_2B$。那么 $\square OACB$ 的面积是这两个相加还是相减？使用割补法可知为**二者相减**。

如果 A 在第一象限，B 在第二象限。那么 $\square OACB$ 的面积是这两个相加还是相减？使用割补法可知为**二者相加**。

为何第一种情形是相减而第二种情形是相加呢? 观察发现: 这些平行四边形中只有第一种情形的 $\varnothing OA_2C_2B$, 从 OA_2 到 OB 是顺时针方向. 如果定义**有向面积**并记为

$$|\boldsymbol{u}, \boldsymbol{v}| = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = \begin{cases} S_{\varnothing OACB}, & \text{从 } OA \text{ 到 } OB \text{ 是逆时针;} \\ -S_{\varnothing OACB}, & \text{从 } OA \text{ 到 } OB \text{ 是顺时针.} \end{cases}$$

那么

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = \begin{vmatrix} a & c \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & c \\ b & d \end{vmatrix} = a \begin{vmatrix} 1 & c \\ 0 & d \end{vmatrix} + b \begin{vmatrix} 0 & c \\ 1 & d \end{vmatrix}.$$

换言之, $\boldsymbol{v}$ 固定时, 则 $|\boldsymbol{u}, \boldsymbol{v}|$ 关于 $\boldsymbol{u}$ 是线性的. 同理, $\boldsymbol{u}$ 固定时, $|\boldsymbol{u}, \boldsymbol{v}|$ 关于 $\boldsymbol{v}$ 也是线性的.

将上述概念推广到 n 维情形. 考虑 n 维空间中由从原点出发的向量 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ 张成的平行多面体. 如果 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ 就是按顺序各个分量上的单位向量

$$\boldsymbol{e}_1 = (1, 0, \ldots, 0)^{\mathrm{T}}, \quad \boldsymbol{e}_2 = (0, 1, \ldots, 0)^{\mathrm{T}}, \quad \cdots, \quad \boldsymbol{e}_n = (0, 0, \ldots, 1)^{\mathrm{T}},$$

则有向面积为 1. 如果交换 $\boldsymbol{v}_i, \boldsymbol{v}_j$ 的位置, 有向面积相差 -1 倍. 于是得到 n 维情形的有向面积 $|\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n|$ 应当满足:

(1) $|\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n| = 1$;

(2) 反对称性: $|\cdots, \boldsymbol{v}_i, \cdots, \boldsymbol{v}_j, \cdots| = -|\cdots, \boldsymbol{v}_j, \cdots, \boldsymbol{v}_i, \cdots|$;

(3) $|\cdots, k\boldsymbol{\alpha}, \cdots| = k|\cdots, \boldsymbol{\alpha}, \cdots|$;

(4) $|\cdots, \boldsymbol{\alpha} + \boldsymbol{\beta}, \cdots| = |\cdots, \boldsymbol{\alpha}, \cdots| + |\cdots, \boldsymbol{\beta}, \cdots|$.

行列式的定义

设 n 阶方阵 $\boldsymbol{A}=(a_{ij})$ 的各列形成的向量为 $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n$, 称有向面积 $|\boldsymbol{v}_1,\cdots,\boldsymbol{v}_n|$ 就是方阵 $\boldsymbol{A}$ 的行列式. 注意到

$$\boldsymbol{v}_j=a_{1j}\boldsymbol{e}_1+a_{2j}\boldsymbol{e}_2+\cdots+a_{nj}\boldsymbol{e}_n,$$

利用线性性质将行列式展开将会得到 n^n 项

$$\sum_{k_1,k_2,\ldots,k_n=1}^{n}|\boldsymbol{e}_{k_1},\cdots,\boldsymbol{e}_{k_n}|a_{k_1 1}a_{k_2 2}\cdots a_{k_n n}.$$

如果 $k_i=k_j$, 则交换 $\boldsymbol{e}_{k_i},\boldsymbol{e}_{k_j}$ 可知 $|\boldsymbol{e}_{k_1},\cdots,\boldsymbol{e}_{k_n}|=0$. 从而只剩下 $k_1,k_2,\ldots,k_n$ 是 $1,2,\ldots,n$ 的排列时的那些项. 例如:

$$\begin{aligned}\begin{vmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{vmatrix}&=|a_{11}\boldsymbol{e}_1,a_{12}\boldsymbol{e}_1+a_{22}\boldsymbol{e}_2|+|a_{21}\boldsymbol{e}_2,a_{12}\boldsymbol{e}_1+a_{22}\boldsymbol{e}_2|\\&=a_{11}a_{12}|\boldsymbol{e}_1,\boldsymbol{e}_1|+a_{11}a_{22}|\boldsymbol{e}_1,\boldsymbol{e}_2|+a_{21}a_{12}|\boldsymbol{e}_2,\boldsymbol{e}_1|+a_{21}a_{22}|\boldsymbol{e}_2,\boldsymbol{e}_2|\\&=a_{11}a_{22}-a_{21}a_{12}.\end{aligned}$$

设 $k_1, k_2, \ldots, k_n$ 是 $1, 2, \ldots, n$ 的排列. 如果排列 $k_1, k_2, \ldots, k_n$ 需要奇数次对换变成 $1, 2, \ldots, n$, 记 $\operatorname{sgn}(k_1, \ldots, k_n) = -1$; 否则 $\operatorname{sgn}(k_1, \ldots, k_n) = +1$. 根据反对称性,

$$|\boldsymbol{e}_{k_1}, \cdots, \boldsymbol{e}_{k_n}| = \operatorname{sgn}(k_1, \ldots, k_n)|\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n| = \operatorname{sgn}(k_1, \ldots, k_n).$$

定义

设 $\boldsymbol{A} = (a_{ij})$ 是 n 阶方阵. 定义 $\boldsymbol{A}$ 的行列式为

$$|\boldsymbol{A}| = \sum \operatorname{sgn}(k_1, \ldots, k_n) a_{k_1 1} a_{k_2 2} \cdots a_{k_n n},$$

其中 $k_1, k_2, \ldots, k_n$ 取遍 $1, 2, \ldots, n$ 的全体排列.

当 $n=2$ 时, $\mathrm{sgn}(12)=1, \mathrm{sgn}(21)=-1$, 于是

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}.$$

当 $n=3$ 时,

$$\mathrm{sgn}(123)=\mathrm{sgn}(231)=\mathrm{sgn}(312)=1,$$

$$\mathrm{sgn}(132)=\mathrm{sgn}(213)=\mathrm{sgn}(321)=-1,$$

于是

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} := a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31}.$$

例: $2,3$ 阶行列式的计算

例

$$\begin{vmatrix}1&3&2\\3&-5&1\\2&1&4\end{vmatrix}=1\cdot(-5)\cdot4+3\cdot1\cdot2+2\cdot3\cdot1-1\cdot1\cdot1-3\cdot3\cdot4-2\cdot(-5)\cdot2$$
$$=-20+6+6-1-36+20=-25.$$

练习

若 $k>0$ 且 $\begin{vmatrix}k&2&1\\2&k&1\\k&1&2\end{vmatrix}=0$, 则 $k=$ __2__.

注记

(1) 行列式将一个方阵映射到一个数.

(2) 1 阶行列式就是方阵里面唯一的那个元素, 尽管也记作 $|\cdot|$, 但注意和绝对值区分.

(3) 2, 3 阶行列式可以用对角线法直接得到展开式, 但是更高阶的没有这种表示方法.

(4) 对角阵的行列式

$$\left|\operatorname{diag}(a_1, a_2, \ldots, a_n)\right| = a_1 a_2 \cdots a_n,$$

特别地 $|\boldsymbol{E}_n| = 1, |\boldsymbol{O}_n| = 0$.

(5) $|\boldsymbol{A}|$ 是由一些 $\pm a_{k_1 1} a_{k_2 2} \cdots a_{k_n n}$ 相加得到, 其中 $k_1, k_2, \ldots, k_n$ 取遍 $1, 2, \ldots, n$ 的所有排列, 一共有 $n!$ 个这样的项, 其中一半取 $+$, 一半取 $-$ $(n \geqslant 2)$.

(6) $|\boldsymbol{A}|$ 对应的线性变换 (常数倍) 是 $\boldsymbol{A}$ 对应的线性变换 $\mathbb{R}^n \to \mathbb{R}^n$ 诱导的 n 次外代数上的线性变换 $\bigwedge^n \mathbb{R}^n \to \bigwedge^n \mathbb{R}^n$, 感兴趣的可自行阅读有关材料.

(1) $|\boldsymbol{AB}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$.

证明

设 $f(\boldsymbol{X}) := |\boldsymbol{AX}|$, 即 $f(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) = |\boldsymbol{Av}_1, \ldots, \boldsymbol{Av}_n|$. 容易知道 f 也满足反对称性, 且对任意 $\boldsymbol{v}_i$ 是线性的. 类似于行列式展开可知, 对于 $\boldsymbol{X} = (a_{ij})$,

$$\begin{aligned} f(\boldsymbol{X}) &= \sum f(\boldsymbol{e}_{k_1}, \ldots, \boldsymbol{e}_{k_n}) a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} \\ &= \sum f(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) \operatorname{sgn}(k_1, \ldots, k_n) a_{k_1 1} a_{k_2 2} \cdots a_{k_n n} \\ &= f(\boldsymbol{E}) |\boldsymbol{X}| = |\boldsymbol{A}| \cdot |\boldsymbol{X}|, \end{aligned}$$

其中 $k_1, k_2, \ldots, k_n$ 取遍 $1, 2, \ldots, n$ 的全体排列. 故 $|\boldsymbol{AB}| = f(\boldsymbol{B}) = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$. □

由此可知, 对于平行多面体 $V \subset \mathbb{R}^n$, $\boldsymbol{A}$ 对应的线性映射将其面积变为 $|\boldsymbol{A}|$ 倍.

例

证明：
$$\begin{vmatrix} a_1+b_1 & b_1+c_1 & c_1+a_1 \\ a_2+b_2 & b_2+c_2 & c_2+a_2 \\ a_3+b_3 & b_3+c_3 & c_3+a_3 \end{vmatrix} = 2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

证明

$$\begin{vmatrix} a_1+b_1 & b_1+c_1 & c_1+a_1 \\ a_2+b_2 & b_2+c_2 & c_2+a_2 \\ a_3+b_3 & b_3+c_3 & c_3+a_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \cdot \begin{vmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = 2\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}. \quad \square$$

(2) 转置不改变行列式: $|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|$.

一个排列 $k_1,\dots,k_n$ 可以看成是集合 $\{1,2,\dots,n\}$ 到自身的双射 $i\mapsto k_i$. 设它的逆映射对应的排列是 $\ell_1,\dots,\ell_n$, 则 $\ell_{k_i}=i$. 由于

$$|\boldsymbol{A}|=\sum\operatorname{sgn}(k_1,\dots,k_n)a_{k_1 1}\cdots a_{k_n n}=\sum\operatorname{sgn}(k_1,\dots,k_n)a_{1\ell_1}\cdots a_{n\ell_n},$$
$$|\boldsymbol{A}^{\mathrm{T}}|=\sum\operatorname{sgn}(\ell_1,\dots,\ell_n)a_{1\ell_1}\cdots a_{n\ell_n},$$

我们只需说明 $\operatorname{sgn}(k_1,\dots,k_n)=\operatorname{sgn}(\ell_1,\dots,\ell_n)$. 设

$$\boldsymbol{P}=(\boldsymbol{e}_{k_1},\boldsymbol{e}_{k_2},\dots,\boldsymbol{e}_{k_n})$$

的 k_i 行 i 列为 1, 其余项为零. 那么 $|\boldsymbol{P}|=\operatorname{sgn}(k_1,\dots,k_n)$, $|\boldsymbol{P}^{\mathrm{T}}|=\operatorname{sgn}(\ell_1,\dots,\ell_n)$. 由于 $\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{E}$, 因此 $|\boldsymbol{P}|\cdot|\boldsymbol{P}^{\mathrm{T}}|=|\boldsymbol{E}|=1$. 而 $|\boldsymbol{P}|=\pm1$, 因此 $|\boldsymbol{P}|=|\boldsymbol{P}^{\mathrm{T}}|$.

例

设 $\boldsymbol{A}=\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}$，求 $|\boldsymbol{A}|$.

解

这题可以直接硬算，不过我们可以利用一点小技巧：

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}=\begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix}\begin{pmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{pmatrix}=(a^2+b^2+c^2+d^2)\boldsymbol{E}.$$

因此 $|\boldsymbol{A}|=\pm(a^2+b^2+c^2+d^2)^2$. 因为 $|\boldsymbol{A}|$ 一定有 a^4 项，所以 $|\boldsymbol{A}|=(a^2+b^2+c^2+d^2)^2$.

再根据行列式关于每个列向量的线性性和反对称性有:

(3) 互换两行 (列) 后, 方阵的行列式变为 -1 倍.

(4) 方阵的某一行 (列) 乘 k 后, 方阵的行列式变为 k 倍.

(5) 将方阵某一行 (列) 对应向量写成两个向量之和, 则行列式也可对应拆成两个行列式之和.

推论

(1) 具有相同的两行 (列) 的方阵的行列式为零: $|\cdots, \boldsymbol{v}, \cdots, \boldsymbol{v}, \cdots| = 0$.

(2) 若方阵有一行 (列) 全为零, 则行列式为零: $|\cdots, \mathbf{0}, \cdots| = 0$.

(3) 若方阵有两行 (列) 成比例, 则行列式为零: $|\cdots, \boldsymbol{v}, \cdots, k\boldsymbol{v}, \cdots| = 0$.

(4) 行列式中某一行 (列) 的公因子可以提到行列式外面.

计算行列式可以通过实施下列变换来化简:

初等变换

(1) 互换两行 (列): $\boldsymbol{r_i \leftrightarrow r_j, c_i \leftrightarrow c_j}$, 行列式变号;

(2) 一行 (列) 乘**非零常数** k: $\boldsymbol{kr_i, kc_i}$, 行列式变为 k 倍;

(3) j 行 (列) 乘 k 加到 i 行 (列): $\boldsymbol{r_i + kr_j, c_i + kc_j}$, 行列式不变.

实施第三类初等变换 $c_i + kc_j$ 时, 第 j 列不变, 改变的是第 i 列. 由于

$$\begin{aligned}|\cdots, \boldsymbol{v}_i, \cdots, \boldsymbol{v}_j, \cdots| &= |\cdots, \boldsymbol{v}_i, \cdots, \boldsymbol{v}_j, \cdots| + |\cdots, k\boldsymbol{v}_j, \cdots, \boldsymbol{v}_j, \cdots| \\ &= |\cdots, \boldsymbol{v}_i + k\boldsymbol{v}_j, \cdots, \boldsymbol{v}_j, \cdots|,\end{aligned}$$

因此第三类初等变换不改变行列式的值.

练习

(1) 判断题：$|\lambda \boldsymbol{A}| = \lambda|\boldsymbol{A}|$. $|\lambda \boldsymbol{A}| = \lambda^n|\boldsymbol{A}|$

(2) 判断题：$\begin{vmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{vmatrix} = -\begin{vmatrix} & & & 1 \\ & & 2 & \\ & 3 & & \\ 4 & & & \end{vmatrix}$. $|\boldsymbol{e}_4, \boldsymbol{e}_3, \boldsymbol{e}_2, \boldsymbol{e}_1| = 1$

(3) 计算 $\begin{vmatrix} a_1+b_1 & b_1+c_1 & c_1+d_1 & d_1+a_1 \\ a_2+b_2 & b_2+c_2 & c_2+d_2 & d_2+a_2 \\ a_3+b_3 & b_3+c_3 & c_3+d_3 & d_3+a_3 \\ a_4+b_4 & b_4+c_4 & c_4+d_4 & d_4+a_4 \end{vmatrix} = \underline{\quad 0 \quad}$.

(4) 设 $\boldsymbol{A}$ 为 5 阶方阵，$|\boldsymbol{A}| = -1$，则 $|2\boldsymbol{A}| = \underline{\quad -32 \quad}$，$||\boldsymbol{A}|\boldsymbol{A}| = \underline{\quad 1 \quad}$.

例

计算
$$\begin{vmatrix} 2\sin a\cos a & \sin a\cos b+\cos a\sin b & \sin a\cos c+\cos a\sin c \\ \sin b\cos a+\cos b\sin a & 2\sin b\cos b & \sin b\cos c+\cos b\sin c \\ \sin c\cos a+\cos c\sin a & \sin c\cos b+\cos c\sin b & 2\sin c\cos c \end{vmatrix}.$$

容易看出该方阵可写成两个方阵之和

$$\begin{pmatrix} \sin a\cos a & \sin a\cos b & \sin a\cos c \\ \sin b\cos a & \sin b\cos b & \sin b\cos c \\ \sin c\cos a & \sin c\cos b & \sin c\cos c \end{pmatrix} + \begin{pmatrix} \cos a\sin a & \cos a\sin b & \cos a\sin c \\ \cos b\sin a & \cos b\sin b & \cos b\sin c \\ \cos c\sin a & \cos c\sin b & \cos c\sin c \end{pmatrix}.$$

这两个方阵各自满足各行成比例，因此可分别写成

$$\begin{pmatrix} \sin a \\ \sin b \\ \sin c \end{pmatrix}(\cos a, \cos b, \cos c), \qquad \begin{pmatrix} \cos a \\ \cos b \\ \cos c \end{pmatrix}(\sin a, \sin b, \sin c).$$

因此原方阵为

$$\begin{pmatrix} \sin a & \cos a \\ \sin b & \cos b \\ \sin c & \cos c \end{pmatrix} \cdot \begin{pmatrix} \cos a & \cos b & \cos c \\ \sin a & \sin b & \sin c \end{pmatrix}.$$

解

$$原式 = \begin{vmatrix} \sin a & \cos a & 0 \\ \sin b & \cos b & 0 \\ \sin c & \cos c & 0 \end{vmatrix} \cdot \begin{vmatrix} \cos a & \cos b & \cos c \\ \sin a & \sin b & \sin c \\ 0 & 0 & 0 \end{vmatrix} = 0.$$

设 $\boldsymbol{A} \in M_{m\times n}, \boldsymbol{B} \in M_{n\times m}$. 若 $m > n$, 则

$$|\boldsymbol{AB}| = \left| (\boldsymbol{A}, \boldsymbol{O}_{m\times(m-n)}) \begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{O}_{(m-n)\times m} \end{pmatrix} \right| = 0.$$

例

设 n 阶方阵 $\boldsymbol{A}$ 是反对称阵. 若 n 是奇数, 则 $|\boldsymbol{A}|=0$.

证明

由于 $\boldsymbol{A}^{\mathrm{T}}=-\boldsymbol{A}$, 于是

$$|\boldsymbol{A}|=|\boldsymbol{A}^{\mathrm{T}}|=|-\boldsymbol{A}|=(-1)^n|\boldsymbol{A}|=-|\boldsymbol{A}|.$$

故 $|\boldsymbol{A}|=0$. □

例：使用初等变换计算行列式

例

若 $abcd=1$，证明 $\boldsymbol{A}=\begin{pmatrix} a^2+a^{-2} & a & a^{-1} & 1 \\ b^2+b^{-2} & b & b^{-1} & 1 \\ c^2+c^{-2} & c & c^{-1} & 1 \\ d^2+d^{-2} & d & d^{-1} & 1 \end{pmatrix}$ 行列式为零.

证明

$$|\boldsymbol{A}|=\begin{vmatrix} a^2 & a & a^{-1} & 1 \\ b^2 & b & b^{-1} & 1 \\ c^2 & c & c^{-1} & 1 \\ d^2 & d & d^{-1} & 1 \end{vmatrix}+\begin{vmatrix} a^{-2} & a & a^{-1} & 1 \\ b^{-2} & b & b^{-1} & 1 \\ c^{-2} & c & c^{-1} & 1 \\ d^{-2} & d & d^{-1} & 1 \end{vmatrix}=abcd\begin{vmatrix} a & 1 & a^{-2} & a^{-1} \\ b & 1 & b^{-2} & b^{-1} \\ c & 1 & c^{-2} & c^{-1} \\ d & 1 & d^{-2} & d^{-1} \end{vmatrix}+\begin{vmatrix} a & a^{-2} & 1 & a^{-1} \\ b & b^{-2} & 1 & b^{-1} \\ c & c^{-2} & 1 & c^{-1} \\ d & d^{-2} & 1 & d^{-1} \end{vmatrix}.$$

由于 $abcd=1$，且等式右侧两个行列式相差 -1 倍，因此 $|\boldsymbol{A}|=0$. □

余子式和代数余子式

我们来介绍行列式与方阵子式的联系.

定义

设 $\boldsymbol{A} = (a_{ij})$ 是 $n \geqslant 2$ 阶方阵.

(1) $\boldsymbol{A}$ 去掉第 i 行和 j 列得到的 $n-1$ 阶方阵的行列式称为 $\boldsymbol{A}$ 在 (i,j) 处的**余子式** (Minor), 记为 M_{ij}.

(2) 称 $A_{ij} = (-1)^{i+j} M_{ij}$ 为 $\boldsymbol{A}$ 在 (i,j) 处的**代数余子式** (Algebraic Minor).

注意余子式和代数余子式是数而不是矩阵.

行列式与余子式的联系

假设 $\boldsymbol{A}$ 的第 n 列除了 a_{nn} 都是零. 若 $k_n \neq n$, 则 $a_{k_1 1} \cdots a_{k_n n} = 0$; 若 $k_n = n$, 则 $\operatorname{sgn}(k_1, \dots, k_{n-1}, n) = \operatorname{sgn}(k_1, \dots, k_{n-1})$. 因此

$$|\boldsymbol{A}| = \sum \operatorname{sgn}(k_1, \dots, k_{n-1}, n) a_{k_1,1} \cdots a_{k_{n-1},n-1} a_{n,n} = a_{nn} M_{nn} = a_{nn} A_{nn}.$$

假设 $\boldsymbol{A}$ 的第 j 列除了 a_{ij} 都是零. 依次对 $\boldsymbol{A}$ 实施

$$r_i \leftrightarrow r_{i+1}, \quad r_{i+1} \leftrightarrow r_{i+2}, \quad \dots, \quad r_{n-1} \leftrightarrow r_n,$$

得到的方阵 $\boldsymbol{B}$ 就是将 $\boldsymbol{A}$ 的第 i 行移动到第 n 行的后面得到的方阵. 由于一共 $n-i$ 次列互换, 因此 $|\boldsymbol{B}| = (-1)^{n-i}|\boldsymbol{A}|$.

同理, 将 $\boldsymbol{B}$ 的第 j 列移动到第 n 列的后面得到的方阵记为 $\boldsymbol{C}$, 则

$$|\boldsymbol{C}| = (-1)^{n-j}|\boldsymbol{B}| = (-1)^{i+j}|\boldsymbol{A}|.$$

注意到 $\boldsymbol{C}$ 在 (n, n) 处元素是 a_{ij}, 余子式是 M_{ij}, 因此

$$|\boldsymbol{C}| = a_{ij} M_{ij}, \qquad |\boldsymbol{A}| = (-1)^{i+j} a_{ij} M_{ij} = a_{ij} A_{ij}.$$

将 $\boldsymbol{A}$ 的第 j 列写成

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a_{2j} \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_{nj} \end{pmatrix},$$

根据行列式的线性性质，我们得到

$$|\boldsymbol{A}| = a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}.$$

由于转置不改变方阵的行列式，于是得到

$$|\boldsymbol{A}| = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}.$$

行列式沿任一行或列展开

方阵的行列式等于任一行 (列) 的元素与其对应的代数余子式乘积的和:

$$\begin{aligned}|\boldsymbol{A}| &= a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in}\\ &= a_{1j}A_{1j} + a_{2j}A_{2j} + \cdots + a_{nj}A_{nj}.\end{aligned}$$

由此也可以看出 $i \neq k$ 时,

$$a_{i1}A_{k1} + a_{i2}A_{k2} + \cdots + a_{in}A_{kn} = 0,$$

因为它是第 i, k 行相同的方阵的行列式.

例

$$\begin{vmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & & \\ \vdots & \ddots & \\ a_{n2} & \cdots & a_{nn} \end{vmatrix} = a_{11}a_{22} \begin{vmatrix} a_{33} & & \\ \vdots & \ddots & \\ a_{n3} & \cdots & a_{nn} \end{vmatrix}$$

$$= \cdots = a_{11}a_{22}\cdots a_{nn}.$$

由于转置不改变行列式，因此上三角阵行列式也等于对角元乘积.

例

计算 $|\boldsymbol{A}|$, 其中 $\boldsymbol{A}=\begin{pmatrix} & & & a_1 \\ & & a_2 & \\ & \iddots & & \\ a_n & & & \end{pmatrix}$.

解

$$|\boldsymbol{A}|=(-1)^{n+1}a_1\begin{vmatrix} & & a_2 \\ & \iddots & \\ a_n & & \end{vmatrix}=(-1)^{n+1}a_1\cdot(-1)^{n}a_2\begin{vmatrix} & & a_3 \\ & \iddots & \\ a_n & & \end{vmatrix}$$

$$=\cdots=\prod_{i=1}^{n}(-1)^{n-i}a_i=(-1)^{\frac{n(n-1)}{2}}a_1a_2\cdots a_n.$$

对于具体的方阵，我们可以利用初等变换将其化为三角阵来计算行列式. 也可以在某一行或一列只有少数非零元时用拉普拉斯展开来降阶.

例

$$\begin{vmatrix} 2 & 3 & 1 & -1 \\ -4 & -5 & 1 & 3 \\ -3 & 1 & -5 & 3 \\ 1 & -2 & 0 & -1 \end{vmatrix} \xlongequal[r_1 - 2r_4]{\substack{r_2+4r_4 \\ r_3+3r_4}} \begin{vmatrix} 0 & 7 & 1 & 1 \\ 0 & -13 & 1 & -1 \\ 0 & -5 & -5 & 0 \\ 1 & -2 & 0 & -1 \end{vmatrix} = (-1)^{4+1} \begin{vmatrix} -13 & 1 & -1 \\ -5 & -5 & 0 \\ 7 & 1 & 1 \end{vmatrix}$$

$$\xlongequal{r_3+r_1} - \begin{vmatrix} -13 & 1 & -1 \\ -5 & -5 & 0 \\ -6 & 2 & 0 \end{vmatrix} = (-1)^{1+3} \begin{vmatrix} -5 & -5 \\ -6 & 2 \end{vmatrix} = -40.$$

例：利用初等变换计算行列式

练习

(1) $\begin{vmatrix} -2 & 0 & 1 \\ 501 & 200 & 299 \\ 500 & 200 & 300 \end{vmatrix} = \underline{\quad -200 \quad}$.

(2) $\begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ b+c & c+a & a+b \end{vmatrix} = \underline{\quad 0 \quad}$.

(3) 设 $\boldsymbol{\alpha} = (1, 0, -1), \boldsymbol{A} = \boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}$, 则 $|5\boldsymbol{E} - \boldsymbol{A}^3| = \underline{\quad -75 \quad}$.

回忆：若 $\boldsymbol{A} = \boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}}$, 则 $\boldsymbol{A}^k = \lambda^{k-1}\boldsymbol{A}$, 其中 $k = \boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}$.

例

设

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix},$$

$$\boldsymbol{C} = \left(\begin{array}{ccc:ccc} a_{11} & \cdots & a_{1m} & & & \\ \vdots & \ddots & \vdots & & 0 & \\ a_{m1} & \cdots & a_{mm} & & & \\ \hdashline * & \cdots & * & b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ * & \cdots & * & b_{n1} & \cdots & b_{nn} \end{array}\right).$$

证明 $|\boldsymbol{C}| = |\boldsymbol{A}| \cdot |\boldsymbol{B}|$.

例：分块矩阵行列式

证明

对 m 归纳．当 $m=1$ 时将 $|\boldsymbol{C}|$ 沿第一行展开可知成立．

假设命题对于 $m-1$ 成立．设 $\boldsymbol{A}$ 在 $(1,j)$ 处的余子式为 M_{1j}, $\boldsymbol{C}$ 在 $(1,j)$ 处的余子式为 N_{1j}．则由归纳假设 $N_{1j}=M_{1j}|\boldsymbol{B}|$．因此

$$\begin{aligned}|\boldsymbol{C}| &= \sum_{j=1}^{m}(-1)^{1+j}a_{1j}N_{1j}\\ &= \sum_{j=1}^{m}(-1)^{1+j}a_{1j}M_{1j}|\boldsymbol{B}| = |\boldsymbol{A}|\cdot|\boldsymbol{B}|.\end{aligned}$$

□

例：拉普拉斯展开的应用

例

设 $\boldsymbol{A} = \begin{pmatrix} 3 & 0 & 4 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & -7 & 0 & 0 \\ 5 & 3 & -2 & 2 \end{pmatrix}$. 计算 $A_{41}+A_{42}+A_{43}+A_{44}$ 和 $M_{41}+M_{42}+M_{43}+M_{44}$.

解

由拉普拉斯展开可知

$$A_{41}+A_{42}+A_{43}+A_{44} = \begin{vmatrix} 3 & 0 & 4 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & -7 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0.$$

续解

$$M_{41}+M_{42}+M_{43}+M_{44}=-A_{41}+A_{42}-A_{43}+A_{44}=\begin{vmatrix} 3 & 0 & 4 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & -7 & 0 & 0 \\ -1 & 1 & -1 & 1 \end{vmatrix}$$

$$=7\begin{vmatrix} 3 & 4 & 0 \\ 2 & 2 & 2 \\ -1 & -1 & 1 \end{vmatrix}=-28.$$

练习

若 $\boldsymbol{A}=\begin{pmatrix} a_1 & a_2 & a_3 & f \\ b_1 & b_2 & b_3 & f \\ c_1 & c_2 & c_3 & f \\ d_1 & d_2 & d_3 & f \end{pmatrix}$，则 $A_{11}+A_{21}+A_{31}+A_{41}=\underline{\quad 0 \quad}$.

计算 n 阶矩阵的行列式可以使用初等变换将其变为三角型，也可以使用拉普拉斯展开来对其实施降阶.

例

$$\begin{vmatrix} a & 1 & \cdots & 1 \\ 1 & a & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & a \end{vmatrix} \xlongequal[i \geqslant 2]{c_1 + c_i} \begin{vmatrix} a+n-1 & 1 & \cdots & 1 \\ a+n-1 & a & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a+n-1 & 1 & \cdots & a \end{vmatrix} = (a+n-1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & a & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & a \end{vmatrix}$$

$$\xlongequal[i \geqslant 2]{r_i - r_1} (a+n-1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & a-1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a-1 \end{vmatrix} = (a+n-1)(a-1)^{n-1}.$$

例：行和为常数的行列式

若方阵的每行 (列) 之和为常数, 可用此法化简.

练习

计算 n 阶行列式 $\begin{vmatrix} 1+a_1 & a_2 & \cdots & a_n \\ a_1 & 1+a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \cdots & 1+a_n \end{vmatrix} = \underline{\quad 1+a_1+\cdots+a_n \quad}$.

例：箭形行列式

例

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & n \end{vmatrix} \xlongequal[i \geqslant 2]{c_1 - \frac{1}{i} c_i} \begin{vmatrix} 1 - \frac{1}{2} - \cdots - \frac{1}{n} & 1 & \cdots & 1 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{vmatrix} = \left(1 - \frac{1}{2} - \cdots - \frac{1}{n}\right) n!.$$

一般的箭形行列式均可用此法处理.

例：特殊形状行列式

练习

计算 n 阶行列式 $\begin{vmatrix} 1 & 2 & 3 & \cdots & n-1 & n \\ -1 & 0 & 3 & \cdots & n-1 & n \\ -1 & -2 & 0 & \cdots & n-1 & n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -2 & -3 & \cdots & 0 & n \\ -1 & -2 & -3 & \cdots & -(n-1) & 0 \end{vmatrix} = \underline{\quad n! \quad}.$

答案

对该方阵实施 $r_i + r_1, i \geqslant 2$ 即可化为上三角阵.

例

计算矩阵 $\boldsymbol{A}_n = \begin{pmatrix} x & -1 & 0 & \cdots & 0 & 0 \\ 0 & x & -1 & \cdots & 0 & 0 \\ 0 & 0 & x & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x & -1 \\ a_n & a_{n-1} & a_{n-2} & \cdots & a_2 & x+a_1 \end{pmatrix}$ 的行列式.

解

沿着第一列展开得到

$$|\boldsymbol{A}_n| = x|\boldsymbol{A}_{n-1}| + (-1)^{1+n}a_n(-1)^{n-1} = x|\boldsymbol{A}_{n-1}| + a_n,$$

递推或归纳可知

$$|\boldsymbol{A}_n| = x(x|\boldsymbol{A}_{n-2}| + a_{n-1}) + a_n = \cdots = x^n + a_1x^{n-1} + a_2x^{n-2} + \cdots + a_n.$$

例：降阶法计算三对角矩阵行列式

例

计算矩阵 $\boldsymbol{A}_n = \begin{pmatrix} 2 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2 \end{pmatrix}$ 的行列式.

解

设 $D_n = |\boldsymbol{A}_n|$. 沿着第一行展开得到

$$|\boldsymbol{A}_n| = 2|\boldsymbol{A}_{n-1}| - \begin{vmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 2 & 1 & \cdots & 0 & 0 \\ 0 & 1 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 2 \end{vmatrix}_{n-1} = 2|\boldsymbol{A}_{n-1}| - |\boldsymbol{A}_{n-2}|,$$

因此

$$|\boldsymbol{A}_n| - |\boldsymbol{A}_{n-1}| = |\boldsymbol{A}_{n-1}| - |\boldsymbol{A}_{n-2}| = \cdots = |\boldsymbol{A}_2| - |\boldsymbol{A}_1| = 1,$$

从而 $|\boldsymbol{A}_n| = n - 1 + |\boldsymbol{A}_1| = n + 1.$

例：降阶法计算三对角矩阵行列式

若主对角线元素均为 a，上下副对角线元素均为 b 和 c，则

$$|\boldsymbol{A}_n| - a|\boldsymbol{A}_{n-1}| + bc|\boldsymbol{A}_{n-2}| = 0.$$

设 $\lambda^2 - a\lambda + bc = 0$ 的两个根为 λ_1, λ_2，则归纳可知

$$\begin{aligned}|\boldsymbol{A}_n| &= \lambda_1^n + \lambda_1^{n-1}\lambda_2 + \cdots + \lambda_1\lambda_2^{n-1} + \lambda_2^n \\ &= \begin{cases} \dfrac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1 - \lambda_2}, & \text{若 } \lambda_1 \neq \lambda_2; \\ (n+1)\left(\dfrac{a}{2}\right)^n, & \text{若 } \lambda_1 = \lambda_2 = \dfrac{a}{2}. \end{cases}\end{aligned}$$

范德蒙行列式

设 $\boldsymbol{A}_n = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$. 证明 $|\boldsymbol{A}_n| = \prod_{1\leqslant i<j\leqslant n}(x_j - x_i)$.

证明

归纳证明. 当 $n=1,2$ 时显然成立. 设 $n\geqslant 3$, 由 $r_n - x_1 r_{n-1}, \ldots, r_2 - x_1 r_1$ 得到

$$|\boldsymbol{A}_n| = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2 - x_1 & x_3 - x_1 & \cdots & x_n - x_1 \\ 0 & x_2(x_2 - x_1) & x_3(x_3 - x_1) & \cdots & x_n(x_n - x_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_2^{n-2}(x_2 - x_1) & x_3^{n-2}(x_3 - x_1) & \cdots & x_n^{n-2}(x_n - x_1) \end{vmatrix}.$$

续证

沿着第一列展开，然后提取每一列的公因式 $(x_j - x_1)$ 得到

$$|\boldsymbol{A}_n| = \prod_{j=2}^{n}(x_j - x_1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}.$$

由归纳假设可知

$$|\boldsymbol{A}_n| = \prod_{j=2}^{n}(x_j - x_1) \cdot \prod_{2\leqslant i<j\leqslant n}(x_j - x_i) = \prod_{1\leqslant i<j\leqslant n}(x_j - x_i).$$

□

例：范德蒙行列式的应用

练习

(1) $\begin{vmatrix} x_1^{-3} & x_2^{-3} & x_3^{-3} & x_4^{-3} \\ x_1^{-1} & x_2^{-1} & x_3^{-1} & x_4^{-1} \\ x_1 & x_2 & x_3 & x_4 \\ x_1^3 & x_2^3 & x_3^3 & x_4^3 \end{vmatrix} = \underline{x_1^{-3}x_2^{-3}x_3^{-3}x_4^{-3} \prod_{1\leqslant i<j\leqslant 4}(x_j^2 - x_i^2)}.$

(2) $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 65 \end{vmatrix} = \underline{\quad 14 \quad}. = \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 1 & 8 & 27 & 64 \end{vmatrix} + \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \\ 0 & 0 & 0 & 1 \end{vmatrix}$

(3) 设 a, b, c 两两不等，且 $\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ b+c & c+a & a+b \end{vmatrix} = 0$，则 $a+b+c = \underline{\quad 0 \quad}$.

例：范德蒙行列式

范德蒙行列式还有另一种证明方式，这种思路对于其它行列式的计算也有帮助.

证明

$f = |\boldsymbol{A}_n|$ 是 $x_1, \ldots, x_n$ 的多项式，且次数不超过 $1+2+\cdots+(n-1)$. 由于当 $x_i = x_j$ 时 $f = 0$，因此 f 包含因式 $x_i - x_j$，从而

$$f = g \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i).$$

比较两边次数可知 g 是常数. 注意到 $\prod_{i=1}^{n} x_i^{i-1}$ 只出现在范德蒙行列式对角元的乘积中，且它在 $\prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)$ 中的系数是 1. 因此 $g = 1$. □

例

计算 $\begin{vmatrix} 1^{50} & 2^{50} & \cdots & 100^{50} \\ 2^{50} & 3^{50} & \cdots & 101^{50} \\ \vdots & \vdots & \ddots & \vdots \\ 100^{50} & 101^{50} & \cdots & 199^{50} \end{vmatrix}$.

证明

将第一行换成 $(x+1)^{50}, \dots, (x+100)^{50}$, 并将行列式记为 $f(x)$. 那么 $f(1) = \cdots = f(99) = 0$. 注意到 f 的次数不超过 50, 因此 $f \equiv 0$. $\square$

同理若 $k < n-1$, $\left|\left((a_i + b_j)^k\right)_{1 \leqslant i,j \leqslant n}\right| = 0$.

(1) $2,3$ 阶行列式可用对角线法直接展开.

(2) 三角阵行列式等于对角元的乘积, 分块三角阵行列式等于对角阵行列式乘积.

(3) 行列式的计算一般需要用到三类初等变换, 创造出足够多的零.

(4) 行列式沿一行 (列) 的展开往往是降阶法的必要手段.

(5) 范德蒙型行列式可处理方阵为元素幂次递增的情形.

第四节　逆矩阵

- 方阵的伴随矩阵
- 逆矩阵的定义和形式
- 逆矩阵的性质
- 克拉默法则
- 逆矩阵的应用

定义

设 $\boldsymbol{A}=(a_{ij})$ 为 $n \geqslant 2$ 阶方阵. 由 $\boldsymbol{A}$ 的代数余子式形成的 n 阶方阵

$$\boldsymbol{A}^*=(A_{ji})=\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

称为 $\boldsymbol{A}$ 的伴随矩阵.

注意, 伴随矩阵的 (i,j) 元是代数余子式 A_{ji} 而不是 A_{ij}.

$\boldsymbol{A}^*$ 对应的线性变换是 $\boldsymbol{A}$ 对应的线性变换 $\mathbb{R}^n \to \mathbb{R}^n$ 诱导的 $n-1$ 次外代数的对偶空间上的线性变换 $\left(\wedge^{n-1}\mathbb{R}^n\right)^* \to \left(\wedge^{n-1}\mathbb{R}^n\right)^*$, 感兴趣的可自行阅读有关材料.

例

若 $\boldsymbol{A}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, 则 $\boldsymbol{A}^*=\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

伴随矩阵满足如下性质:

(1) $\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{E}_n$.

这是因为

$$\boldsymbol{A}\boldsymbol{A}^*=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix}$$

的 (i,j) 元是 $a_{i1}A_{j1}+\cdots+a_{in}A_{jn}=\begin{cases} |\boldsymbol{A}|, & i=j; \\ 0, & i\neq j. \end{cases}$

(2) $(k\boldsymbol{A})^* = k^{n-1}\boldsymbol{A}^*$.

(3) $(\boldsymbol{A}^{\mathrm{T}})^* = (\boldsymbol{A}^*)^{\mathrm{T}}$.

(4) $|\boldsymbol{A}^*| = |\boldsymbol{A}|^{n-1}$.

若 $|\boldsymbol{A}| \neq 0$, 由 $|\boldsymbol{A}^*| \cdot |\boldsymbol{A}| = |\boldsymbol{A}^*\boldsymbol{A}| = ||\boldsymbol{A}|\boldsymbol{E}_n| = |\boldsymbol{A}|^n$ 可得.

若 $|\boldsymbol{A}| = 0$, 则

$$\boldsymbol{A}\boldsymbol{A}^*(\boldsymbol{A}^*)^* = |\boldsymbol{A}|(\boldsymbol{A}^*)^* = \boldsymbol{O} = |\boldsymbol{A}^*|\boldsymbol{A},$$

于是 $|\boldsymbol{A}^*| = 0$ 或 $\boldsymbol{A} = \boldsymbol{O}$, $\boldsymbol{A}^* = \boldsymbol{O}$. 故 $|\boldsymbol{A}^*| = 0$.

(5) $(\boldsymbol{A}^*)^* = \begin{cases} \boldsymbol{A}, & n=2; \\ |\boldsymbol{A}|^{n-2}\boldsymbol{A}, & n \geqslant 3. \end{cases}$

容易知道, 2 阶方阵满足 $(\boldsymbol{A}^*)^* = \boldsymbol{A}$. 若 $\boldsymbol{A}$ 是 $n \geqslant 3$ 阶方阵且 $|\boldsymbol{A}| \neq 0$, 则由

$$\boldsymbol{A}\boldsymbol{A}^*(\boldsymbol{A}^*)^* = |\boldsymbol{A}|(\boldsymbol{A}^*)^* = |\boldsymbol{A}^*|\boldsymbol{A}$$

可知

$$(\boldsymbol{A}^*)^* = \frac{|\boldsymbol{A}^*|}{|\boldsymbol{A}|}\boldsymbol{A} = |\boldsymbol{A}|^{n-2}\boldsymbol{A}.$$

若 $\boldsymbol{A}$ 是 $n \geqslant 3$ 阶方阵且 $|\boldsymbol{A}| = 0$, 我们会在 §2.4 证明 $(\boldsymbol{A}^*)^* = \boldsymbol{O}$.

例

设非零 $n \geqslant 3$ 实方阵 $\boldsymbol{A}$ 满足对任意 i, j, $a_{ij} = A_{ij}$. 求 $|\boldsymbol{A}|$.

解

由题设可知 $\boldsymbol{A}^* = \boldsymbol{A}^{\mathrm{T}}$. 因此 $|\boldsymbol{A}| = |\boldsymbol{A}^*| = |\boldsymbol{A}|^{n-1}$, $|\boldsymbol{A}| = 0, 1$ 或 -1.
注意 $\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{E}$ 的第 i 个对角元

$$\sum_{k=1}^{n} a_{ik}^2 \geqslant 0.$$

因此 $|\boldsymbol{A}| \geqslant 0$. 若 $|\boldsymbol{A}| = 0$, 则所有的 $a_{ik} = 0, \boldsymbol{A} = \boldsymbol{O}$, 矛盾! 因此 $|\boldsymbol{A}| = 1$.

事实上, $\boldsymbol{A}$ 是正交阵, 即满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = \boldsymbol{E}$ 的实方阵.

给定一个线性变换 $f:\mathbb{R}^n\to\mathbb{R}^m$, 若线性变换 $g:\mathbb{R}^m\to\mathbb{R}^n$ 满足

$$(gf)(\boldsymbol{x})=\boldsymbol{x},\forall\boldsymbol{x}\in\mathbb{R}^n,\quad(fg)(\boldsymbol{y})=\boldsymbol{y},\forall\boldsymbol{y}\in\mathbb{R}^m,$$

则称 g 是 f 的逆.

设 f,g 对应的矩阵分别是 $\boldsymbol{A},\boldsymbol{B}$, 则

$$\boldsymbol{A}\boldsymbol{B}=\boldsymbol{E}_m,\quad\boldsymbol{B}\boldsymbol{A}=\boldsymbol{E}_n.$$

由于 $m>n$ 时 $|\boldsymbol{A}\boldsymbol{B}|=0$; $m<n$ 时 $|\boldsymbol{B}\boldsymbol{A}|=0$. 因此线性变换的逆只可能在 $m=n$ 时存在.

由此得到对应的矩阵的逆的定义:

定义

设 A 是 n 阶方阵. 若存在 n 阶方阵 B 使得

$$AB = BA = E_n,$$

则称 A 是**可逆矩阵**, B 是 A 的**逆矩阵**.

可逆矩阵的逆矩阵唯一吗? 设 B, B' 都是 A 的逆矩阵, 则

$$AB = E_n, \quad B'A = E_n.$$

于是

$$B = (B'A)B = B'(AB) = B'.$$

因此**若逆矩阵存在必唯一, 记为 A^{-1}**.

若 $\boldsymbol{A}$ 可逆, 从 $\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{E}$ 可知 $|\boldsymbol{A}|\cdot|\boldsymbol{A}^{-1}|=1$. 从而 $|\boldsymbol{A}|\neq 0$.

反之, 若 $|\boldsymbol{A}|\neq 0$, 则

$$\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{E},\qquad \boldsymbol{A}\cdot\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*\cdot\boldsymbol{A}=\boldsymbol{E},\qquad \boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*.$$

逆矩阵的存在性和形式

n **阶方阵** $\boldsymbol{A}$ **可逆当且仅当** $|\boldsymbol{A}|\neq 0$. **此时** $\boldsymbol{A}^{-1}=\dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*$.

推论

设 $\boldsymbol{A}, \boldsymbol{B}$ 为 n 阶方阵. 若 $\boldsymbol{AB}=\boldsymbol{E}$ (或 $\boldsymbol{BA}=\boldsymbol{E}$), 则 $\boldsymbol{B}=\boldsymbol{A}^{-1}$.

证明

若 $\boldsymbol{AB}=\boldsymbol{E}$, 则 $|\boldsymbol{A}|\cdot|\boldsymbol{B}|=1$, $|\boldsymbol{A}|\neq 0$. 因此 $\boldsymbol{A}$ 可逆.

$$\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}(\boldsymbol{AB})=\boldsymbol{B}.$$

□

行列式为零的方阵也叫**退化方阵**.

例

证明 $\boldsymbol{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ 可逆并求其逆矩阵.

解

由于 $|\boldsymbol{A}|=\cos^2\theta+\sin^2\theta=1$, 因此 $\boldsymbol{A}$ 可逆. 由于 $\boldsymbol{A}^*=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$, 因此

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}.$$

注意到 $\boldsymbol{A}$ 对应的是平面上沿原点逆时针旋转 θ, 因此 $\boldsymbol{A}^{-1}$ 对应的是平面上沿原点逆时针旋转 $-\theta$.

例：计算逆矩阵

例

$\boldsymbol{A}=\begin{pmatrix}1&2&1\\0&1&-1\\-1&3&4\end{pmatrix}$，$\boldsymbol{B}=\begin{pmatrix}2&7&-2\\-1&-3&5\\1&5&11\end{pmatrix}$ 是否可逆？若可逆求其逆矩阵.

解

由于 $|\boldsymbol{A}| \xlongequal{r_3+r_1} \begin{vmatrix}1&2&1\\0&1&-1\\0&5&5\end{vmatrix}=10$，因此 $\boldsymbol{A}$ 可逆，

$$\boldsymbol{A}^{-1}=\frac{1}{10}\boldsymbol{A}^*=\frac{1}{10}\begin{pmatrix}A_{11}&A_{21}&A_{31}\\A_{12}&A_{22}&A_{32}\\A_{13}&A_{23}&A_{33}\end{pmatrix}=\frac{1}{10}\begin{pmatrix}7&-5&-3\\1&5&1\\1&-5&1\end{pmatrix}.$$

由于 $|\boldsymbol{B}| \xlongequal[c_2-5c_1]{c_3-11c_1} \begin{vmatrix}2&-3&-24\\-1&2&16\\1&0&0\end{vmatrix}=\begin{vmatrix}-3&-24\\2&16\end{vmatrix}=0$，因此 $\boldsymbol{B}$ 不可逆.

例

设 $\lambda_1,\ldots,\lambda_n \neq 0$,

$$\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n),$$

则

$$\boldsymbol{\Lambda}^{-1} = \operatorname{diag}(\lambda_1^{-1},\lambda_2^{-1},\ldots,\lambda_n^{-1}).$$

通过计算伴随矩阵还可知道, 行列式非零的上 (下) 三角阵的逆也是上 (下) 三角的,且二者对角元对应元素互为逆.

逆矩阵通常采用下述方法计算:

(1) 利用公式 $A^{-1}=\frac{1}{|A|}A^*$, 适用于 $2,3$ 阶方阵, 或用于抽象分析.

(2) 寻找方阵 B 使得 $AB=E$, 适用于抽象矩阵求逆.

(3) 利用矩阵的初等变换求逆矩阵, 该方法我们会在下一节中学习.

例

设 A 为 n 阶方阵且满足 $A^2+3A-2E=O$. 求 A^{-1} 和 $(A-E)^{-1}$.

解

(1) 由于 $A^2+3A=2E$, 因此 $A(A+3E)=2E$, $A^{-1}=\frac{1}{2}(A+3E)$.

(2) 由于 $(A-E)(A+4E)=-2E$, 因此 $(A-E)^{-1}=-\frac{1}{2}(A+4E)$.

设 $f(x)$ 是一多项式, $f(\boldsymbol{A})=\boldsymbol{O}$, 我们想求 $(\boldsymbol{A}-a\boldsymbol{E})^{-1}$. 注意到

$$\frac{x^k-a^k}{x-a}=x^{k-1}+ax^{k-2}+\cdots+a^{k-1}$$

是一个多项式, 因此存在多项式 $g(x)=\dfrac{f(x)-f(a)}{x-a}$. 于是

$$f(x)-f(a)=(x-a)g(x),$$

$$f(\boldsymbol{A})-f(a)\boldsymbol{E}=(\boldsymbol{A}-a\boldsymbol{E})g(\boldsymbol{A}).$$

从而当 $f(a)\neq 0$ 时, $(\boldsymbol{A}-a\boldsymbol{E})^{-1}=-\dfrac{1}{f(a)}g(\boldsymbol{A})$.

当 $f(a)=0$ 时, $\boldsymbol{A}-a\boldsymbol{E}$ 未必可逆.

逆矩阵的性质

逆矩阵满足如下性质:

(1) 设 $\boldsymbol{A}$ 可逆, $\lambda \neq 0$.

- $|\boldsymbol{A}^{-1}| = |\boldsymbol{A}|^{-1}$;
- $\boldsymbol{A}^{-1}$ 可逆, 且 $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$;
- $\lambda\boldsymbol{A}$ 可逆, 且 $(\lambda\boldsymbol{A})^{-1} = \dfrac{1}{\lambda}\boldsymbol{A}^{-1}$;
- $\boldsymbol{A}^{\mathrm{T}}$ 可逆, 且 $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$;
- $\boldsymbol{A}^*$ 可逆, 且 $(\boldsymbol{A}^*)^{-1} = (\boldsymbol{A}^{-1})^* = \dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}$.

由于 $\boldsymbol{A}\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{E}$, 因此 $\boldsymbol{A}^* = |\boldsymbol{A}|\boldsymbol{A}^{-1}$. 于是

$$(\boldsymbol{A}^{-1})^* = |\boldsymbol{A}^{-1}|(\boldsymbol{A}^{-1})^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}, \quad (\boldsymbol{A}^*)^{-1} = (|\boldsymbol{A}|\boldsymbol{A}^{-1})^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}.$$

(2) 若 A, B 为同阶可逆矩阵，则 AB 也可逆，且

$$(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}.$$

一般地

$$(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n)^{-1} = \boldsymbol{A}_n^{-1}\cdots\boldsymbol{A}_2^{-1}\boldsymbol{A}_1^{-1}.$$

注意矩阵不能相除 $\dfrac{\boldsymbol{A}}{\boldsymbol{B}}$，因为一般 $\boldsymbol{B}^{-1}\boldsymbol{A} \neq \boldsymbol{A}\boldsymbol{B}^{-1}$.

一般地，$(\boldsymbol{A}+\boldsymbol{B})^{-1} \neq \boldsymbol{A}^{-1}+\boldsymbol{B}^{-1}$. 例如 $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ 均可逆，但 $\boldsymbol{A}+\boldsymbol{B} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$ 不可逆.

例

多选题：若 A, B, C 为同阶方阵，且 A 可逆，则(AC).

(A) 若 $AB = AC$，则 $B = C$　　(B) 若 $AB = CB$，则 $A = C$

(C) 若 $AB = O$，则 $B = O$　　(D) 若 $BC = O$，则 $B = O$

练习

(1) 设 $A = \begin{pmatrix} 4 & 1 & 3 \\ 4 & a & 7 \\ 3 & -1 & 4 \end{pmatrix}$，且存在矩阵 $B \neq C$ 使得 $AB = AC$，则 $a =$ $\underline{\ -3\ }$.

(2) 若 A 为 n 阶方阵，则下面命题正确的有 $\underline{\ 1\ }$ 个.

(i) $A^{-1} = \dfrac{1}{|A|}A^*$;　　(ii) $A^* = |A|A^{-1}$;　　(iii) $|A^*| = |A|^{n-1}$.

例

设 $\boldsymbol{A}$ 是 n 阶方阵，若(D)，则 $\boldsymbol{A}-\boldsymbol{E}$ 可逆.

(A) $\boldsymbol{A}$ 可逆

(B) $|\boldsymbol{A}|=0$

(C) $\boldsymbol{A}$ 的主对角线元素均为 0

(D) 存在某个正整数 m 使得 $\boldsymbol{A}^m=\boldsymbol{O}$

练习

(1) 设 3 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^3-2\boldsymbol{A}+\boldsymbol{E}=\boldsymbol{O}$，且 $|\boldsymbol{A}|=2$，则 $|(\boldsymbol{A}^2-2\boldsymbol{E})^{-1}|=$ $\underline{\quad -2 \quad}$.

(2) 设 n 阶方阵 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}$ 满足 $\boldsymbol{ABC}=\boldsymbol{E}$，则(D).

(A) $\boldsymbol{ACB}=\boldsymbol{E}$ (B) $\boldsymbol{CBA}=\boldsymbol{E}$ (C) $\boldsymbol{BAC}=\boldsymbol{E}$ (D) $\boldsymbol{CAB}=\boldsymbol{E}$

想一想 $\boldsymbol{B}^{-1}=?$

例：逆矩阵的性质

例

设 $\boldsymbol{A}$ 是 3 阶方阵，$|\boldsymbol{A}| = \dfrac{1}{2}$. 求 $|(2\boldsymbol{A})^{-1} - (2\boldsymbol{A})^*|$.

解

$$(2\boldsymbol{A})^{-1} - (2\boldsymbol{A})^* = \frac{1}{2}\boldsymbol{A}^{-1} - 2^2\boldsymbol{A}^* = \boldsymbol{A}^* - 4\boldsymbol{A}^* = -3\boldsymbol{A}^*,$$

因此

$$\left|(2\boldsymbol{A})^{-1} - (2\boldsymbol{A})^*\right| = -27|\boldsymbol{A}^*| = -27|\boldsymbol{A}|^2 = -\frac{27}{4}.$$

例：逆矩阵计算方阵的幂

例

设 $\boldsymbol{P}=\begin{pmatrix}1&2\\1&4\end{pmatrix},\boldsymbol{\Lambda}=\begin{pmatrix}1&\\&2\end{pmatrix}$，求 $(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1})^n$.

解

$$|\boldsymbol{P}|=2,\boldsymbol{P}^{-1}=\frac{1}{2}\begin{pmatrix}4&-2\\-1&1\end{pmatrix},$$

$$\begin{aligned}(\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1})^n&=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\cdot\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\cdots\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}=\boldsymbol{P}\boldsymbol{\Lambda}^n\boldsymbol{P}^{-1}\\&=\frac{1}{2}\begin{pmatrix}1&2\\1&4\end{pmatrix}\begin{pmatrix}1&\\&2^n\end{pmatrix}\begin{pmatrix}4&-2\\-1&1\end{pmatrix}=\begin{pmatrix}2-2^n&2^n-1\\2-2^{n+1}&2^{n+1}-1\end{pmatrix}.\end{aligned}$$

由此可知，对于多项式 $f(x)$，$f(\boldsymbol{P}\boldsymbol{A}\boldsymbol{P}^{-1})=\boldsymbol{P}f(\boldsymbol{A})\boldsymbol{P}^{-1}$.

例

证明 $(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*$.

证明

若 $\boldsymbol{A}, \boldsymbol{B}$ 均可逆，则

$$(\boldsymbol{AB})^* = |\boldsymbol{AB}|(\boldsymbol{AB})^{-1} = |\boldsymbol{A}| \cdot |\boldsymbol{B}|\boldsymbol{B}^{-1}\boldsymbol{A}^{-1} = \boldsymbol{A}^*\boldsymbol{B}^*.$$

一般情形下，注意到 $|x\boldsymbol{E} + \boldsymbol{A}|$ 展开式包含 x^n 项，所以它是 x 的多项式. 由此可知存在无穷多 x 使得 $x\boldsymbol{E} + \boldsymbol{A}, x\boldsymbol{E} + \boldsymbol{B}$ 均可逆.

设 $f(x), g(x)$ 分别为 $((x\boldsymbol{E}+\boldsymbol{A})(x\boldsymbol{E}+\boldsymbol{B}))^*$ 和 $(x\boldsymbol{E}+\boldsymbol{B})^*(x\boldsymbol{E}+\boldsymbol{A})^*$ 的 (i,j) 元，那么 $f(x) = g(x)$ 对无穷多 x 成立. 而 f, g 均是 x 的多项式，这迫使 $f = g, f(0) = g(0)$. 故 $(\boldsymbol{AB})^* = \boldsymbol{B}^*\boldsymbol{A}^*$. □

伴随、逆、转置相互可交换，且均满足此种性质.

克拉默法则

设 A 是 n 阶方阵，$\boldsymbol{b}=(b_1,\dots,b_n)^{\mathrm{T}}$ 是 n 维列向量，将 A 的第 j 列换成 $\boldsymbol{b}$ 得到的方阵记为 $\boldsymbol{A}_j$. 当 $|\boldsymbol{A}|\neq 0$，线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有唯一解

$$\boldsymbol{x}=\left(\frac{|\boldsymbol{A}_1|}{|\boldsymbol{A}|},\frac{|\boldsymbol{A}_2|}{|\boldsymbol{A}|},\dots,\frac{|\boldsymbol{A}_n|}{|\boldsymbol{A}|}\right).$$

证明

显然唯一解就是 $\boldsymbol{x}=\boldsymbol{A}^{-1}\boldsymbol{b}=\dfrac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*\boldsymbol{b}$. 由于方阵 $\boldsymbol{A}_j$ 沿着第 j 列展开得到

$$|\boldsymbol{A}_j|=b_1A_{1j}+b_2A_{2j}+\cdots+b_nA_{nj}=(A_{1j},A_{2j},\dots,A_{nj})\boldsymbol{b}.$$

因此 $\boldsymbol{A}^*\boldsymbol{b}=(|\boldsymbol{A}_1|,\dots,|\boldsymbol{A}_n|)^{\mathrm{T}}$，从而 $\boldsymbol{x}$ 具有题述形式. □

定理

设 $\boldsymbol{A}$ 是 n 阶方阵，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有唯一解当且仅当 $|\boldsymbol{A}|\neq 0$.

因此行列式起到了线性方程组的“判别式”的作用. 我们将在 §1.6 证明该结论.

若线性方程组的常数都是零，即 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 或

$$\begin{cases} a_{11}\ x_1+a_{12}x_2+\cdots+a_{1n}\ x_n=0 \\ a_{21}\ x_1+a_{22}x_2+\cdots+a_{2n}\ x_n=0 \\ \qquad\qquad\qquad\vdots \\ a_{m1}x_1+a_{n2}x_2+\cdots+a_{mn}x_n=0 \end{cases}$$

称之为齐次线性方程组. 否则称之为非齐次线性方程组.

对于齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$, 令

$$V=\{\boldsymbol{x}\mid \boldsymbol{A}\boldsymbol{x}=\mathbf{0}\}$$

表示该方程的所有解形成的集合. 显然 $\mathbf{0}\in V$. 若 $\boldsymbol{u},\boldsymbol{v}\in V$, 则 $\boldsymbol{A}\boldsymbol{u}=\boldsymbol{A}\boldsymbol{v}=\mathbf{0}$. 于是

$$\boldsymbol{A}(\boldsymbol{u}+\boldsymbol{v})=\boldsymbol{A}(\lambda\boldsymbol{v})=\mathbf{0},\quad \forall\lambda\in\mathbb{R}.$$

换言之, 从而 V 构成一个线性空间. 对于一般情形, 若 $\boldsymbol{A}\boldsymbol{\xi}=\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, 则 $\boldsymbol{A}(\boldsymbol{\xi}-\boldsymbol{x}_0)=\mathbf{0}$, 从而它所有的解就是 $\boldsymbol{\xi}+\boldsymbol{x},\boldsymbol{x}\in V$.

设 $\boldsymbol{A}$ 是方阵. 显然零解 $\boldsymbol{x}=\mathbf{0}$ 是 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的解, 其它解被称为非零解. 由于非零解的倍数还是它的解, 因此 $|\boldsymbol{A}|=0 \iff \boldsymbol{A}\boldsymbol{x}=0$ **有无穷多 (非零) 解**. 再根据解的特点可知 $|\boldsymbol{A}|=0 \iff \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ **无解或有无穷多解**.

例：克拉默法则的应用

例

已知 $\begin{cases} \lambda x_1 + x_2 + x_3 = 1 \\ x_1 + \lambda x_2 + x_3 = 1 \\ x_1 + x_2 + \lambda x_3 = -2 \end{cases}$ 有无穷多解，求 λ.

解

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

因此 $\lambda = 1$ 或 -2. 显然 $\lambda = 1$ 时无解. $\lambda = -2$ 时, $(t, t, t+1)$ 是方程的解. 因此 $\lambda = -2$.

例：克拉默法则的应用

练习

若 $\begin{cases} x_1 + \ \ x_2 + \qquad\quad\ \ x_3 = 0 \\ x_1 + \lambda x_2 + \qquad\quad\ \ x_3 = 0 \\ x_1 + \ \ x_2 + (\lambda+1)x_3 = 0 \end{cases}$ 有非零解，则 $\lambda =$ <u>$0, 1$</u>.

例

证明：若三条不同的直线

$$\begin{aligned} ax + by + c &= 0 \\ bx + cy + a &= 0 \\ cx + ay + b &= 0 \end{aligned}$$

相交于一点，则 $a + b + c = 0$.

例：克拉默法则的应用

证明

线性方程组 $\begin{cases} ax_1+bx_2+cx_3=0 \\ bx_1+cx_2+ax_3=0 \\ cx_1+ax_2+bx_3=0 \end{cases}$ 有非零解 $(x,y,1)$. 因此 $\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix}=0$.

$$\begin{aligned}\begin{vmatrix} a & b & c \\ b & c & a \\ c & a & b \end{vmatrix} &= (a+b+c)\begin{vmatrix} 1 & 1 & 1 \\ b & c & a \\ c & a & b \end{vmatrix} \\ &= (a+b+c)(bc+ac+ab-a^2-b^2-c^2) \\ &= -\frac{1}{2}(a+b+c)\big[(a-b)^2+(b-c)^2+(c-a)^2\big].\end{aligned}$$

由于这是三条不同直线，因此 a,b,c 不可能全部相等，从而 $a+b+c=0$. □

想一想：为什么 $a+b+c=0$ 时，三条直线一定相交于一点？

使用克拉默法则解线性方程组需要:

(1) 方程组的数量和未知数数量相同;
(2) 系数矩阵行列式非零.

但由于使用克拉默法则计算量较大, 一般不使用该方法解方程, 通常仅用于理论研究.

练习

何时下述线性方程组有非零解? $b=0$ 或 $-a_1-\cdots-a_n$.

$$\begin{cases} (a_1+b)x_1 + a_2x_2 + \cdots + a_nx_n = 0 \\ a_1x_1 + (a_2+b)x_2 + \cdots + a_nx_n = 0 \\ \vdots \\ a_1x_1 + a_2x_2 + \cdots + (a_n+b)x_n = 0 \end{cases}$$

例

设 $a_1, \ldots, a_n$ 两两不同. 解方程组 $\begin{cases} x_1 + a_1 x_2 + \cdots + a_1^{n-1} x_n = 1 \\ x_1 + a_2 x_2 + \cdots + a_2^{n-1} x_n = 1 \\ \qquad\qquad \vdots \\ x_1 + a_n x_2 + \cdots + a_n^{n-1} x_n = 1 \end{cases}$

解

由于系数矩阵行列式为范德蒙行列式

$$\prod_{1 \leqslant i < j \leqslant n} (a_j - a_i) \neq 0,$$

因此方程有唯一解 $x_1 = 1, x_2 = \cdots = x_n = 0.$

左图是一张夏天的风景图，我们希望把它修改成秋天的景色. Photoshop 提供了将颜色重新搭配的通道混合器. 用取色工具选取树叶、蓝天、地面的颜色，分别得到 RGB 值为

$$(59,181,19),\quad (90,185,249),\quad (210,205,186).$$

我们希望将树叶变成金黄色 RGB$(234,228,70)$ 而保持蓝天和地面的颜色不变. 则我们需要的矩阵 $\boldsymbol{A}$ 满足

$$\boldsymbol{A}\begin{pmatrix} 59 & 90 & 210 \\ 181 & 185 & 205 \\ 19 & 249 & 186 \end{pmatrix} = \begin{pmatrix} 234 & 90 & 210 \\ 228 & 185 & 205 \\ 70 & 249 & 186 \end{pmatrix}.$$

逆矩阵的应用：图像处理

非考试内容

解得

$$A = \begin{pmatrix} 234 & 90 & 210 \\ 228 & 185 & 205 \\ 70 & 249 & 186 \end{pmatrix} \begin{pmatrix} 59 & 90 & 210 \\ 181 & 185 & 205 \\ 19 & 249 & 186 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0.43 & 1.23 & -0.70 \\ -0.15 & 1.33 & -0.19 \\ -0.17 & 0.36 & 0.79 \end{pmatrix}.$$

分别在通道混合器的红绿蓝通道输入上面三行即可.

$\Longrightarrow$

第五节　分块矩阵

- 分块矩阵的定义和运算
- 特殊分块矩阵

有时为了研究矩阵和其部分元素形成的矩阵的联系，需要使用分块法将其进行拆分:

定义

用若干条横线和竖线将矩阵 A 分成许多小矩阵，每个小矩阵成为 A 的子块，以子块为元素的矩阵称为分块矩阵.

例如

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{O}_{m\times n} & \boldsymbol{E}_m\\ \boldsymbol{E}_n & \boldsymbol{O}_{n\times m}\end{pmatrix}$$

就是一个分块矩阵. 事实上我们已经不加定义地用过分块矩阵了.

分块矩阵的运算: 加法和数乘

若分块矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 同型, 且每个对应分块也同型, 则 $\boldsymbol{A}+\boldsymbol{B}$ 就是对应分块相加形成的分块矩阵:

$$\begin{pmatrix} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1r} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{s1} & \cdots & \boldsymbol{A}_{sr} \end{pmatrix} + \begin{pmatrix} \boldsymbol{B}_{11} & \cdots & \boldsymbol{B}_{1r} \\ \vdots & \ddots & \vdots \\ \boldsymbol{B}_{s1} & \cdots & \boldsymbol{B}_{sr} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_{11}+\boldsymbol{B}_{11} & \cdots & \boldsymbol{A}_{1r}+\boldsymbol{B}_{1r} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{s1}+\boldsymbol{B}_{s1} & \cdots & \boldsymbol{A}_{sr}+\boldsymbol{B}_{sr} \end{pmatrix}.$$

数 λ 和分块矩阵的数乘, 就是 λ 和对应分块数乘形成的分块矩阵:

$$\lambda \begin{pmatrix} \boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1r} \\ \vdots & \ddots & \vdots \\ \boldsymbol{A}_{s1} & \cdots & \boldsymbol{A}_{sr} \end{pmatrix} = \begin{pmatrix} \lambda\boldsymbol{A}_{11} & \cdots & \lambda\boldsymbol{A}_{1r} \\ \vdots & \ddots & \vdots \\ \lambda\boldsymbol{A}_{s1} & \cdots & \lambda\boldsymbol{A}_{sr} \end{pmatrix}.$$

设

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \cdots & \boldsymbol{A}_{1r}\\ \vdots & \ddots & \vdots\\ \boldsymbol{A}_{s1} & \cdots & \boldsymbol{A}_{sr}\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \cdots & \boldsymbol{B}_{1t}\\ \vdots & \ddots & \vdots\\ \boldsymbol{B}_{r1} & \cdots & \boldsymbol{B}_{rt},\end{pmatrix}$$

且 $\boldsymbol{A}_{ik}$ 的列数和 $\boldsymbol{B}_{kj}$ 的行数相同，则

$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{C}_{11} & \cdots & \boldsymbol{C}_{1t}\\ \vdots & \ddots & \vdots\\ \boldsymbol{C}_{s1} & \cdots & \boldsymbol{C}_{st}\end{pmatrix},\quad \boldsymbol{C}_{ij}=\sum_{k=1}^{r}\boldsymbol{A}_{ik}\boldsymbol{B}_{kj}.$$

分块矩阵能相乘，需要 $\boldsymbol{A}$ 的列块数和 $\boldsymbol{B}$ 的行块数要相等．然后得到乘积的每个分块表达式后，相应的分块乘法也要能相乘．这样分块矩阵的运算就如同把这些分块视作数一样运算．

分块矩阵的运算：转置

设

$$A=\begin{pmatrix} A_{11} & \cdots & A_{1r} \\ \vdots & \ddots & \vdots \\ A_{s1} & \cdots & A_{sr} \end{pmatrix},$$

则

$$A^{\mathrm{T}}=\begin{pmatrix} A_{11}^{\mathrm{T}} & \cdots & A_{s1}^{\mathrm{T}} \\ \vdots & \ddots & \vdots \\ A_{1r}^{\mathrm{T}} & \cdots & A_{sr}^{\mathrm{T}} \end{pmatrix}.$$

注意小块的位置需要转置，每个小块也需要转置.

若方阵

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_m \end{pmatrix},$$

其中 $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ 都是方阵, 称 $\boldsymbol{A}$ 为分块对角阵. 记作 $\boldsymbol{A} = \mathrm{diag}(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m)$.

分块对角阵具有如下性质:

(1) $|\boldsymbol{A}| = |\boldsymbol{A}_1| \cdots |\boldsymbol{A}_m|$;

(2) $\boldsymbol{A}$ 可逆当且仅当 $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ 均可逆, 此时 $\boldsymbol{A}^{-1} = \mathrm{diag}(\boldsymbol{A}_1^{-1}, \ldots, \boldsymbol{A}_m^{-1})$.

(3) $\boldsymbol{A}^k = \mathrm{diag}(\boldsymbol{A}_1^k, \ldots, \boldsymbol{A}_m^k)$.

例

求 $\boldsymbol{A}=\begin{pmatrix}2&1&0&0\\1&1&0&0\\0&0&2&0\\0&0&-1&3\end{pmatrix}$ 的逆矩阵.

解

设 $\boldsymbol{A}_1=\begin{pmatrix}2&1\\1&1\end{pmatrix}$, $\boldsymbol{A}_2=\begin{pmatrix}2&0\\-1&3\end{pmatrix}$, 则 $\boldsymbol{A}_1^{-1}=\begin{pmatrix}1&-1\\-1&2\end{pmatrix}$, $\boldsymbol{A}_2^{-1}=\frac{1}{6}\begin{pmatrix}3&0\\1&2\end{pmatrix}$.

故 $\boldsymbol{A}^{-1}=\mathrm{diag}(\boldsymbol{A}_1^{-1},\boldsymbol{A}_2^{-1})=\begin{pmatrix}1&-1&&\\-1&2&&\\&&1/2&0\\&&1/6&1/3\end{pmatrix}$.

例: 分块三角阵的逆

例

设 $\boldsymbol{A}, \boldsymbol{B}$ 均可逆, 求 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$ 的逆矩阵.

解

由 $\begin{vmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}| \cdot |\boldsymbol{B}| \neq 0$ 可知该方阵可逆. 设 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix} \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{pmatrix} = \boldsymbol{E}$.
则 $\boldsymbol{A}\boldsymbol{A}_1 = \boldsymbol{E}, \boldsymbol{A}\boldsymbol{A}_2 = \boldsymbol{O}, \boldsymbol{C}\boldsymbol{A}_2 + \boldsymbol{B}\boldsymbol{A}_4 = \boldsymbol{E}$. 于是 $\boldsymbol{A}_1 = \boldsymbol{A}^{-1}, \boldsymbol{A}_2 = \boldsymbol{O}, \boldsymbol{A}_4 = \boldsymbol{B}^{-1}$.
再由 $\boldsymbol{C}\boldsymbol{A}_1 + \boldsymbol{B}\boldsymbol{A}_3 = \boldsymbol{O}$ 可得

$$\boldsymbol{A}_3 = -\boldsymbol{B}^{-1}\boldsymbol{C}\boldsymbol{A}_1 = -\boldsymbol{B}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1}.$$

故 $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} & \boldsymbol{O} \\ -\boldsymbol{B}^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & \boldsymbol{B}^{-1} \end{pmatrix}$.

由此可知, (分块) 上/下三角阵的逆还是 (分块) 上/下三角阵.

例：分块方阵的伴随

练习

设 A, B 为同阶方阵，$C = \begin{pmatrix} A & \\ & B \end{pmatrix}$，则 $C^* =$(C).

(A) $\begin{pmatrix} A^* & \\ & B^* \end{pmatrix}$

(B) $\begin{pmatrix} B^* & \\ & A^* \end{pmatrix}$

(C) $\begin{pmatrix} |B|A^* & \\ & |A|B^* \end{pmatrix}$

(D) $\begin{pmatrix} |A|B^* & \\ & |B|A^* \end{pmatrix}$

第六节　矩阵的初等变换

- 初等矩阵
- 矩阵等价
- 初等变换解矩阵方程

我们曾利用如下三种初等变换来帮助计算行列式:

初等变换

(1) 互换两行 (列): $\boldsymbol{r_i \leftrightarrow r_j, c_i \leftrightarrow c_j}$, 行列式变号;

(2) 一行 (列) 乘**非零常数** k: $\boldsymbol{kr_i, kc_i}$, 行列式变为 k 倍;

(3) j 行 (列) 乘 k 加到 i 行 (列): $\boldsymbol{r_i + kr_j, c_i + kc_j}$, 行列式不变.

实际上我们也可以对矩阵实施初等变换, 而且这三类变换过程都是可逆的, 且其逆变换是同一类变换. 以行变换为例:

(1) $r_i \leftrightarrow r_j$ 的逆是 $r_i \leftrightarrow r_j$;

(2) kr_i 的逆是 $\frac{1}{k} r_i$;

(3) $r_i + kr_j$ 的逆是 $r_i - kr_j$.

我们使用矩阵来表示上述变换.

单位阵 $\boldsymbol{E}$ 经过一次初等变换得到的方阵称为**初等矩阵**.

(1) $r_i \leftrightarrow r_j$ 和 $c_i \leftrightarrow c_j$ 都对应初等矩阵

$$
\boldsymbol{E}(i,j) = \left(\begin{array}{ccccccccccc}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & \cdots & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & \vdots & & \ddots & & \vdots & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & \cdots & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{array}\right)
\begin{array}{l}
 \\
 \\
 \\
\leftarrow \text{第 } i \text{ 行} \\
 \\
 \\
 \\
\leftarrow \text{第 } j \text{ 行} \\
 \\
 \\
 \\
\end{array}
$$

(2) kr_i, kc_i 都对应初等矩阵

$$
\boldsymbol{E}(i(k)) = \begin{array}{c} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & k & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \\ \\ \end{array} \\ \uparrow \\ \text{第 } i \text{ 列} \end{array}
$$

(3) $r_i + kr_j, c_j + kc_i$ 都对应初等矩阵

$$\boldsymbol{E}(i,j(k)) = \begin{array}{cl} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & k & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow \text{第 } i \text{ 行} \\ \\ \leftarrow \text{第 } j \text{ 行} \\ \\ \\ \end{array} \\ \begin{array}{ccccccc} & & \uparrow & & \uparrow & & \\ & & \text{第 } i \text{ 列} & & \text{第 } j \text{ 列} & & \end{array} & \end{array}$$

需要注意的是 $c_i + kc_j$ 对应的初等矩阵不是 $E(i,j(k))$ 而是 $E(j,i(k))$.

我们来看

$$\boldsymbol{E}(1,3)\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}.$$

$\boldsymbol{E}(i,j)$ 左乘在矩阵 $\boldsymbol{A}$ 上, 即对 $\boldsymbol{A}$ 实施 $r_i \leftrightarrow r_j$.

从

$$\boldsymbol{E}(i,j)\boldsymbol{A} = \boldsymbol{E}(i,j) \begin{pmatrix} \vdots \\ \boldsymbol{\alpha}_i \\ \vdots \\ \boldsymbol{\alpha}_j \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ \boldsymbol{\alpha}_j \\ \vdots \\ \boldsymbol{\alpha}_i \\ \vdots \end{pmatrix}$$

可以看出确实如此.

$$\boldsymbol{E}(2(k))\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4k & 5k & 6k \\ 7 & 8 & 9 \end{pmatrix}.$$

$\boldsymbol{E}(i(k))$ 左乘在矩阵 $\boldsymbol{A}$ 上, 即对 $\boldsymbol{A}$ 实施 kr_i.

$$\boldsymbol{E}(3,1(k))\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ k & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7+k & 8+2k & 9+3k \end{pmatrix}.$$

$\boldsymbol{E}(i,j(k))$ 左乘在矩阵 $\boldsymbol{A}$ 上, 即对 $\boldsymbol{A}$ 实施 $r_i + kr_j$. 即, 初等矩阵左乘矩阵 $\boldsymbol{A}$ 等同于对 $\boldsymbol{A}$ 实施对应的初等行变换.

同理, 初等矩阵右乘矩阵 $\boldsymbol{A}$ 等同于对 $\boldsymbol{A}$ 实施对应的初等列变换.

定理

设 $\boldsymbol{A} \in M_{m\times n}$.

(1) 对 $\boldsymbol{A}$ 实施一次初等行变换, 相当于在$\boldsymbol{A}$ 的左边乘对应的 m 阶初等矩阵.

(2) 对 $\boldsymbol{A}$ 实施一次初等列变换, 相当于在$\boldsymbol{A}$ 的右边乘对应的 n 阶初等矩阵.

对应的初等矩阵就是对单位阵 $\boldsymbol{E}$ 实施相应的初等变换得到的矩阵.

即左行右列.

例

设 $\boldsymbol{A}$ 为 3 阶可逆方阵，将 $\boldsymbol{A}$ 的第 1 列与第 2 列交换得 $\boldsymbol{B}$，再把 $\boldsymbol{B}$ 的第 2 列加到第 3 列得到 $\boldsymbol{C}$．求 $\boldsymbol{Q}=\boldsymbol{A}^{-1}\boldsymbol{C}$.

解

$$\boldsymbol{B}=\boldsymbol{A}\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix},\boldsymbol{C}=\boldsymbol{B}\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}=\boldsymbol{A}\boldsymbol{Q}.$$ 因此

$$\boldsymbol{Q}=\begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}=\begin{pmatrix}0&1&1\\1&0&0\\0&0&1\end{pmatrix}.$$

练习

设 $\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \boldsymbol{B}=\begin{pmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31}+2a_{11} & a_{33}+2a_{13} & a_{32}+2a_{12} \end{pmatrix},$

$\boldsymbol{P}_1=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \boldsymbol{P}_2=\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \boldsymbol{P}_3=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$. 则 $\boldsymbol{B}=$(C).

(A) $\boldsymbol{P}_3\boldsymbol{A}\boldsymbol{P}_2$ (B) $\boldsymbol{P}_2\boldsymbol{A}\boldsymbol{P}_3$ (C) $\boldsymbol{P}_3\boldsymbol{A}\boldsymbol{P}_1$ (D) $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{A}$

由于初等变换都是可逆，因此初等矩阵也都是可逆的：

(1) $\boldsymbol{E}(i,j)\boldsymbol{E}(i,j)=\boldsymbol{E} \implies \boldsymbol{E}(i,j)^{-1}=\boldsymbol{E}(i,j)$;

(2) $\boldsymbol{E}(i(k))\boldsymbol{E}(i(\frac{1}{k}))=\boldsymbol{E} \implies \boldsymbol{E}(i(k))^{-1}=\boldsymbol{E}(i(\frac{1}{k}))$;

(3) $\boldsymbol{E}(i,j(k))\boldsymbol{E}(i,j(-k))=\boldsymbol{E} \implies \boldsymbol{E}(i,j(k))^{-1}=\boldsymbol{E}(i,j(-k))$.

例

将可逆方阵 $\boldsymbol{A}$ 的第 1 行的 2 倍加到第 2 行得到 $\boldsymbol{B}$，则对 $\boldsymbol{A}^{-1}$ 实施初等变换(D)可得到 $\boldsymbol{B}^{-1}$.

(A) r_2+2r_1 (B) r_2-2r_1 (C) c_1+2c_2 (D) c_1-2c_2

练习

设 $\boldsymbol{A}$ 是 n 阶可逆矩阵，将 $\boldsymbol{A}$ 的第 i 行与第 j 行对换后得到的矩阵记为 $\boldsymbol{B}$，则 $\boldsymbol{A}\boldsymbol{B}^{-1}=$ $\underline{\quad \boldsymbol{E}(i,j) \quad}$.

例：初等矩阵

例

设 $\boldsymbol{P}_1=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$, $\boldsymbol{P}_2=\begin{pmatrix}1&&&\\0&1&&\\0&0&1&\\a&0&0&1\end{pmatrix}$, $\boldsymbol{P}_3=\begin{pmatrix}1&&&\\&k&&\\&&1&\\&&&1\end{pmatrix}$, 求 $\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3$ 及逆.

解

$$\boldsymbol{P}_1 \xrightarrow{c_1+ac_4} \boldsymbol{P}_1\boldsymbol{P}_2=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\a&0&0&1\end{pmatrix} \xrightarrow{kc_2} \boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3=\begin{pmatrix}0&0&1&0\\0&k&0&0\\1&0&0&0\\a&0&0&1\end{pmatrix},$$

$$\boldsymbol{P}_1^{-1}=\boldsymbol{P}_1 \xrightarrow{r_4-ar_1} (\boldsymbol{P}_1\boldsymbol{P}_2)^{-1}=\begin{pmatrix}0&0&1&0\\0&1&0&0\\1&0&0&0\\0&0&-a&1\end{pmatrix} \xrightarrow{\frac{1}{k}r_2} (\boldsymbol{P}_1\boldsymbol{P}_2\boldsymbol{P}_3)^{-1}=\begin{pmatrix}0&0&1&0\\0&1/k&0&0\\1&0&0&0\\0&0&-a&1\end{pmatrix}.$$

练习

设

$$A=\begin{pmatrix}a_{11}&a_{12}&a_{13}&a_{14}\\a_{21}&a_{22}&a_{23}&a_{24}\\a_{31}&a_{32}&a_{33}&a_{34}\\a_{41}&a_{42}&a_{43}&a_{44}\end{pmatrix},\qquad B=\begin{pmatrix}a_{14}&a_{13}&a_{12}&a_{11}\\a_{24}&a_{23}&a_{22}&a_{21}\\a_{34}&a_{33}&a_{32}&a_{31}\\a_{44}&a_{43}&a_{42}&a_{41}\end{pmatrix},$$

$$P_1=\begin{pmatrix}0&0&0&1\\0&1&0&0\\0&0&1&0\\1&0&0&0\end{pmatrix},\qquad P_2=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}.$$

若 A 可逆，则 $B^{-1}=$(C).

(A) $A^{-1}P_1P_2$ (B) $P_1A^{-1}P_2$ (C) $P_1P_2A^{-1}$ (D) $P_2A^{-1}P_1$

我们来看初等行变换能够将矩阵化简成何种形式.

$$\begin{pmatrix}1 & 3 & -2 & 4\\3 & 6 & -2 & 11\\2 & 1 & 1 & 3\end{pmatrix}\xrightarrow[r_3-2r_1]{r_2-3r_1}\begin{pmatrix}1 & 3 & -2 & 4\\0 & -3 & 4 & -1\\0 & -5 & 5 & -5\end{pmatrix}\xrightarrow{-\frac{1}{5}r_3}\begin{pmatrix}1 & 3 & -2 & 4\\0 & -3 & 4 & -1\\0 & 1 & -1 & 1\end{pmatrix}$$

$$\xrightarrow{r_2\leftrightarrow r_3}\begin{pmatrix}1 & 3 & -2 & 4\\0 & 1 & -1 & 1\\0 & -3 & 4 & -1\end{pmatrix}\xrightarrow{r_3+3r_2}\begin{pmatrix}1 & 3 & -2 & 4\\0 & 1 & -1 & 1\\0 & 0 & 1 & 2\end{pmatrix}$$

经过若干次初等行变换, 矩阵变为行阶梯形矩阵.

定义

满足下述条件的矩阵称为行阶梯形矩阵:

(1) 每个非零行的第一个非零元只出现在上一行第一个非零元的右边;

(2) 零行只可能出现在最下方.

换言之, 若 $\boldsymbol{A} \in M_{m\times n}$, 存在正整数

$$1 \leqslant k_1 < k_2 < \cdots < k_\ell, \qquad \ell \leqslant m$$

使得 $a_{1,k_1}, \ldots, a_{\ell,k_\ell}$ 均非零; $j < k_i$ 或 $i > \ell$ 时 $a_{ij} = 0$.

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 0 & 3 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 2 & 3 & 4 & -1 \\ 1 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 3 & -1 \end{pmatrix} \times$$

任何矩阵都可通过初等行变换化为行阶梯形.

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} \underset{r_1+2r_3}{\overset{r_2+r_3}{\sim}} \begin{pmatrix} 1 & 3 & 0 & 8 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix} \overset{r_1-3r_2}{\sim} \begin{pmatrix} \boxed{1} & 0 & 0 & -1 \\ 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & \boxed{1} & 2 \end{pmatrix}$$

再经过若干次初等变换，矩阵变为行最简形矩阵.

定义

满足下述条件的行阶梯形矩阵称为行最简形矩阵:

(1) 每个非零行的第一个非零元是 1;

(2) 每个非零行的第一个非零元所在列其它元素均为 0.

例如 $\begin{pmatrix} 0 & \boxed{1} & 0 & 2 & 0 \\ 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$. 任何矩阵都可通过初等行变换化为行最简形.

例

用初等行变换将 $\boldsymbol{A}=\begin{pmatrix}1 & 3 & -9 & 3\\ 0 & 1 & -3 & 4\\ -2 & -3 & 9 & 6\end{pmatrix}$ 化为行最简形矩阵.

解

$$\begin{pmatrix}1 & 3 & -9 & 3\\ 0 & 1 & -3 & 4\\ -2 & -3 & 9 & 6\end{pmatrix}\underset{\sim}{r_3+2r_1}\begin{pmatrix}1 & 3 & -9 & 3\\ 0 & 1 & -3 & 4\\ 0 & 3 & -9 & 12\end{pmatrix}\underset{\sim}{r_3-3r_2}\begin{pmatrix}1 & 3 & -9 & 3\\ 0 & 1 & -3 & 4\\ 0 & 0 & 0 & 0\end{pmatrix}$$

$$\underset{\sim}{r_1-3r_2}\begin{pmatrix}1 & 0 & 0 & -9\\ 0 & 1 & -3 & 4\\ 0 & 0 & 0 & 0\end{pmatrix}.$$

命题

n 阶方阵 $\boldsymbol{A}$ 可逆 $\iff$ 它的行最简形为 $\boldsymbol{E}_n$.

注意到初等行变换不改变行列式的非零性. 若 $\boldsymbol{A}$ 可逆, 则它的行最简形 $\boldsymbol{B}$ 不能有零行, 从而 $\boldsymbol{B}$ 有 n 个非零行. 而每个非零行的首个非零元所在列数递增, 因此第 i 行的首个非零元一定在第 i 列, 从而 $\boldsymbol{B} = \boldsymbol{E}$.

由此可知, 存在一系列初等矩阵 $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_k$ 使得 $\boldsymbol{P}_1 \cdots \boldsymbol{P}_k \boldsymbol{A} = \boldsymbol{E}$. 从而 $\boldsymbol{A} = \boldsymbol{P}_k^{-1} \cdots \boldsymbol{P}_1^{-1}$, 可逆方阵可以写成有限个初等矩阵的乘积.

命题

设 $\boldsymbol{A}$ 是 n 阶方阵. 若 $|\boldsymbol{A}|=0$, 则 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 有非零解.

证明

设 $\boldsymbol{A}^{\mathrm{T}}$ 的行最简形为 $\boldsymbol{B}=\boldsymbol{P}\boldsymbol{A}^{\mathrm{T}}$, 其中 $\boldsymbol{P}$ 是一些初等矩阵的乘积, 从而可逆. 由于 $|\boldsymbol{A}^{\mathrm{T}}|=|\boldsymbol{A}|=0$, 因此 $\boldsymbol{B}\neq\boldsymbol{E}$ 有零行. 设 $\boldsymbol{v}=(0,\ldots,0,1)^{\mathrm{T}}$, 那么 $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{v}=\mathbf{0}$. 从而 $\boldsymbol{A}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{v}=\mathbf{0}$. 由于 $\boldsymbol{P}^{\mathrm{T}}$ 可逆, 它的最后一列 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{v}$ 非零. □.

推论

设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵. 若 $m<n$, 则 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 有非零解.

证明

将 $\boldsymbol{A}$ 补充 $n-m$ 个零行得到 $\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{O}\end{pmatrix}$, 则 $|\boldsymbol{B}|=0$, 存在非零向量 $\boldsymbol{x}$ 使得 $\boldsymbol{B}\boldsymbol{x}=\mathbf{0}$. 从而 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$. □

定义

(1) 若 A 经过有限次初等行变换变为 B, 则称 A 和 B 行等价, 记作 $A \overset{r}{\sim} B$.

(2) 若 A 经过有限次初等列变换变为 B, 则称 A 和 B 列等价, 记作 $A \overset{c}{\sim} B$.

(3) 若 A 经过有限次初等行变换和初等列变换变为 B, 则称 A 和 B 等价, 记作 $A \sim B$.

由于可逆方阵可以写成有限个初等矩阵的乘积, 因此:

定理

(1) $A \overset{r}{\sim} B$ 当且仅当存在可逆矩阵 P 使得 $B = PA$.

(2) $A \overset{c}{\sim} B$ 当且仅当存在可逆矩阵 Q 使得 $B = AQ$.

(3) $A \sim B$ 当且仅当存在可逆矩阵 P, Q 使得 $B = PAQ$.

由此可知

命题

矩阵的行等价、列等价、等价均满足

(1) 自反性: $A \sim A$;

(2) 对称性: $A \sim B \implies B \sim A$;

(3) 传递性: $A \sim B, B \sim C \implies A \sim C$.

任一矩阵通过有限次初等行变换变为行最简形后, 可通过初等列变换将其变为标准型 $\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$. 例如:

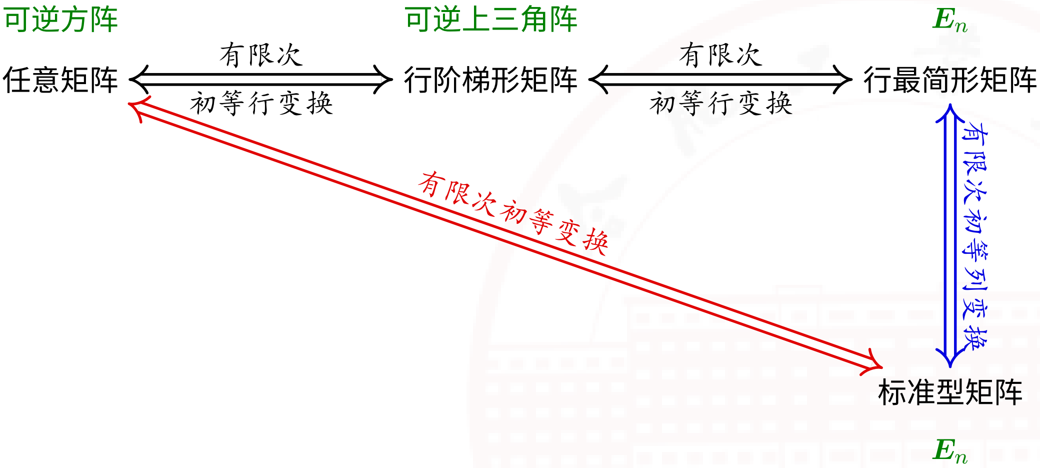
$$\begin{pmatrix} 1 & 0 & 0 & -9 \\ 0 & 1 & -3 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \underset{}{\overset{c_4+9c_1}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -3 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \underset{c_4-4c_2}{\overset{c_3+3c_2}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

矩阵的等价也叫做相抵, 上述标准型也叫作相抵标准型. 我们会看到不同的 r 对应的相抵标准型不等价. 所以相抵标准型相当于在每一个等价类中找到了一个具有代表性的矩阵.

命题

n 阶方阵 $\boldsymbol{A}$ 可逆当且仅当它的标准型为 $\boldsymbol{E}_n$.

矩阵的变换关系



可逆方阵
可逆上三角阵
E_n
任意矩阵
有限次
初等行变换
行阶梯形矩阵
有限次
初等行变换
行最简形矩阵
有限次初等变换
有限次初等列变换
标准型矩阵
E_n

例

将矩阵 $\boldsymbol{A}=\begin{pmatrix}1&0&0\\2&0&-1\\0&-1&0\end{pmatrix}$ 表示成有限个初等阵的乘积.

解

$$\boldsymbol{A}\overset{r_2\leftrightarrow r_3}{\sim}\begin{pmatrix}1&0&0\\0&-1&0\\2&0&-1\end{pmatrix}\overset{r_3-2r_1}{\sim}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}\underset{-r_3}{\overset{-r_2}{\sim}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}.$$

因此 $\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&1&0\\-2&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&0&1\\0&1&0\end{pmatrix}\boldsymbol{A}=\boldsymbol{E}$,

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

若 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{AX}=\boldsymbol{B} \iff \boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$. 若 $(\boldsymbol{A},\boldsymbol{B}) \overset{r}{\sim} (\boldsymbol{E},\boldsymbol{X})$, 则存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}(\boldsymbol{A},\boldsymbol{B})=(\boldsymbol{E},\boldsymbol{X})$. 即 $\boldsymbol{P}=\boldsymbol{A}^{-1}, \boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$. 所以

$$\boldsymbol{AX}=\boldsymbol{B} \iff \boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B} \iff (\boldsymbol{A},\boldsymbol{B}) \overset{r}{\sim} (\boldsymbol{E},\boldsymbol{X}).$$

同理

$$\boldsymbol{XA}=\boldsymbol{C} \iff \boldsymbol{X}=\boldsymbol{CA}^{-1} \iff \begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{B}\end{pmatrix} \overset{c}{\sim} \begin{pmatrix}\boldsymbol{E}\\ \boldsymbol{X}\end{pmatrix}.$$

所以我们可使用初等变换解该类型矩阵方程.

特别地, $(\boldsymbol{A},\boldsymbol{E}) \overset{r}{\sim} (\boldsymbol{E},\boldsymbol{A}^{-1})$ 可用来帮助计算矩阵的逆.

例

求 $\boldsymbol{A}=\begin{pmatrix}1&1&2\\1&2&3\\2&1&4\end{pmatrix}$ 的逆.

例：初等变换解矩阵方程

解

$$(\boldsymbol{A},\boldsymbol{E}) = \left(\begin{array}{ccc:ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 2 & 1 & 4 & 0 & 0 & 1 \end{array}\right) \underset{r_3-2r_1}{\overset{r_2-r_1}{\sim}} \left(\begin{array}{ccc:ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & -1 & 0 & -2 & 0 & 1 \end{array}\right)$$

$$\overset{r_3+r_2}{\sim} \left(\begin{array}{ccc:ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -3 & 1 & 1 \end{array}\right) \underset{r_2-r_3}{\overset{r_1-2r_3}{\sim}} \left(\begin{array}{ccc:ccc} 1 & 1 & 0 & 7 & -2 & -2 \\ 0 & 1 & 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & -3 & 1 & 1 \end{array}\right)$$

$$\overset{r_1-r_2}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 5 & -2 & -1 \\ 0 & 1 & 0 & 2 & 0 & -1 \\ 0 & 0 & 1 & -3 & 1 & 1 \end{array}\right).$$

故 $\boldsymbol{A}^{-1} = \begin{pmatrix} 5 & -2 & -1 \\ 2 & 0 & -1 \\ -3 & 1 & 1 \end{pmatrix}.$

例

若 $\boldsymbol{A}=\begin{pmatrix}2&2&0\\2&1&3\\0&1&0\end{pmatrix}$，$\boldsymbol{AX}=\boldsymbol{A}+\boldsymbol{X}$，求 $\boldsymbol{X}$.

解

由题设知 $(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{X}=\boldsymbol{A}$，$\boldsymbol{X}=(\boldsymbol{A}-\boldsymbol{E})^{-1}\boldsymbol{A}$.

$$(\boldsymbol{A}-\boldsymbol{E},\boldsymbol{A})=\left(\begin{array}{ccc|ccc}1&2&0&2&2&0\\2&0&3&2&1&3\\0&1&-1&0&1&0\end{array}\right)\xrightarrow[r_2\leftrightarrow r_3]{r_2-2r_1}\left(\begin{array}{ccc|ccc}1&2&0&2&2&0\\0&1&-1&0&1&0\\0&-4&3&-2&-3&3\end{array}\right)$$

例：初等变换解矩阵方程

续解

$$(\boldsymbol{A}-\boldsymbol{E},\boldsymbol{A}) \overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 2 & 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & -4 & 3 & -2 & -3 & 3 \end{array}\right) \underset{-r_3}{\overset{r_3+4r_2}{\sim}} \left(\begin{array}{ccc:ccc} 1 & 2 & 0 & 2 & 2 & 0 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & -1 & -3 \end{array}\right)$$

$$\overset{r_2+r_3}{\sim} \left(\begin{array}{ccc:ccc} 1 & 2 & 0 & 2 & 2 & 0 \\ 0 & 1 & 0 & 2 & 0 & -3 \\ 0 & 0 & 1 & 2 & -1 & -3 \end{array}\right) \overset{r_1-2r_2}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & -2 & 2 & 6 \\ 0 & 1 & 0 & 2 & 0 & -3 \\ 0 & 0 & 1 & 2 & -1 & -3 \end{array}\right).$$

故 $\boldsymbol{X} = \begin{pmatrix} -2 & 2 & 6 \\ 2 & 0 & -3 \\ 2 & -1 & -3 \end{pmatrix}.$

例

解矩阵方程 $\boldsymbol{A}^*\boldsymbol{X} = 2\boldsymbol{E} + 2\boldsymbol{X}$，其中 $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$.

解

注意到 $|\boldsymbol{A}| = 4$. 两边同时左乘 $\boldsymbol{A}$ 得到 $4\boldsymbol{X} = 2\boldsymbol{A} + 2\boldsymbol{A}\boldsymbol{X}$, $(2\boldsymbol{E} - \boldsymbol{A})\boldsymbol{X} = \boldsymbol{A}$.

$$(2\boldsymbol{E} - \boldsymbol{A}, \boldsymbol{A}) = \left(\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 & -1 & 1 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{array}\right).$$

故 $\boldsymbol{X} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$.

练习

解矩阵方程 $A^*XA = 2XA - 8E$，其中 $A = \begin{pmatrix} 1 & 2 & -2 \\ 0 & -2 & 4 \\ 0 & 0 & 1 \end{pmatrix}$.

答案

两边同时左乘 A 右乘 A^{-1} 得到 $-2X = 2AX - 8E$，$(A+E)X = 4E$，

$$(A+E, 4E) = \left(\begin{array}{ccc:ccc} 2 & 2 & -2 & 4 & 0 & 0 \\ 0 & -1 & 4 & 0 & 4 & 0 \\ 0 & 0 & 2 & 0 & 0 & 4 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc:ccc} 1 & 0 & 0 & 2 & 4 & -6 \\ 0 & 1 & 0 & 0 & -4 & 8 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{array}\right).$$

练习

(1) 设 $\boldsymbol{A}$ 是 3 阶方阵, 存在可逆阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}1&&\\&2&\\&&3\end{pmatrix}$, 则

$\boldsymbol{P}^{-1}\boldsymbol{A}^*\boldsymbol{P}=$ ____diag(6, 3, 2)____.

(2) 设 $\boldsymbol{A}$ 是 3 阶方阵, 存在可逆阵 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\begin{pmatrix}1&&\\&2&\\&&3\end{pmatrix}$.

若 $\boldsymbol{Q}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2)$, 则 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=$ ____diag(1, 3, 2)____.

(3) 设 n 阶方阵 $\boldsymbol{A},\boldsymbol{B}$ 满足 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{E}$, 则以下说法正确的有____4____个.

(i) $\boldsymbol{A}$ 等价于 $\boldsymbol{E}$;

(ii) $\boldsymbol{A}$ 等价于 $\boldsymbol{B}$;

(iii) $\boldsymbol{A}$ 可经过有限次初等行变换化为 $\boldsymbol{B}$;

(iv) $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}$.

第二章　等价和秩

1. 向量组
2. 矩阵的秩
3. 标准正交基
4. 线性方程组

第一节　向量组

- 线性组合
- 线性相关和线性无关
- 线性相关和线性无关的性质
- 维数和秩
- 极大线性无关组

向量组

我们将一些具有相同维数的向量放在一起称之为**向量组**(可以有重复的)：例如:

- $\boldsymbol{\alpha}_1 = (1,1,-1)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (2,1,2)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (3,2,1)^{\mathrm{T}}$;
- $\boldsymbol{\alpha}_1^{\mathrm{T}} = (1,1,-1), \boldsymbol{\alpha}_2^{\mathrm{T}} = (2,1,2), \boldsymbol{\alpha}_3^{\mathrm{T}} = (3,2,1)$;
- $m \times n$ 矩阵 $\boldsymbol{A}$ 的 m 行可以看成 m 个行向量，它们构成一个向量组，叫做 $\boldsymbol{A}$ 的**行向量组**;
- 类似地，$\boldsymbol{A}$ 的列向量构成它的**列向量组**.
- $\boldsymbol{e}_1 = (1,0,\ldots,0)^{\mathrm{T}}, \boldsymbol{e}_2 = (0,1,\ldots,0)^{\mathrm{T}}, \ldots, \boldsymbol{e}_n = (0,0,\ldots,1)^{\mathrm{T}}$.

对于标准向量空间 $V=\mathbb{R}^n$, 我们可以将任一 $\boldsymbol{v}\in V$ **唯一**地表达为形式

$$\boldsymbol{v}=\lambda_1\boldsymbol{e}_1+\lambda_2\boldsymbol{e}_2+\cdots+\lambda_n\boldsymbol{e}_n,$$

其中

$$\boldsymbol{e}_1=(1,0,\ldots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,\ldots,0)^{\mathrm{T}},\ldots,\boldsymbol{e}_n=(0,0,\ldots,1)^{\mathrm{T}}.$$

定义

设 V 是线性空间. 若 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m\in V$ 满足: 对任意 $\boldsymbol{v}\in V$, 存在唯一的一组数 $\lambda_1,\ldots,\lambda_m$ 使得

$$\boldsymbol{v}=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m,$$

则称 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m$ 是 V 的一组基.

如何判断一组向量是不是基呢? 这需要线性组合和线性无关的概念.

向量组的线性组合

定义

设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$ 为 n 维向量. 若存在一组数 $\lambda_1,\ldots,\lambda_m$ 使得

$$\boldsymbol{\beta}=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m,$$

则称 $\boldsymbol{\beta}$ 可以被向量组 $\{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m\}$ **线性表示**, 或称 $\boldsymbol{\beta}$ 是向量组 $\{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m\}$ 的**线性组合**.

例

(1) n 维零向量是任一 n 维向量组的线性组合.

(2) 任意 n 维向量是 $\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n$ 的线性组合.

(3) $v\in V$ 是它的一组基的线性组合.

(4) 空间中两条不共线的向量的线性组合全体就是过二者的平面.

向量 β 能被向量组 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性表示, 当且仅当存在 $\lambda_1, \dots, \lambda_m$ 使得

$$\beta = \lambda_1\boldsymbol{\alpha}_1 + \cdots + \lambda_m\boldsymbol{\alpha}_m = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m)\begin{pmatrix}\lambda_1\\ \vdots \\ \lambda_m\end{pmatrix},$$

即 $\boldsymbol{A}\boldsymbol{x} = \beta$ 有解, 其中 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m)$.

定理

向量 β 能被 $\boldsymbol{A}$ 的列向量组线性表示, 当且仅当 $\boldsymbol{A}\boldsymbol{x} = \beta$ 有解.

设向量组 S 是 $\boldsymbol{A}$ 的列向量组. 记 V 为向量组 S 能线性表示的向量全体, 则

$$V = \{\boldsymbol{A}\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^m\} \subseteq \mathbb{R}^n$$

是一个线性空间, 称为 S 生成的空间. 它是包含 S 中所有向量的最小的线性空间.

这样, β 能被 S 线性表示 $\iff \beta \in V$.

定义

(1) 设有两个向量组 $S=\{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m\}, T=\{\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_k\}$. 若 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_k$ 均可以被 S 线性表示, 则称向量组 T 可以被向量组 S 线性表示.

(2) 若向量组 S 和 T 能相互线性表示, 则称 S,T **向量组等价**.

设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m), \boldsymbol{B}=(\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_k)$. T 能被 S 线性表示, 当且仅当存在 $\boldsymbol{x}_1,\dots,\boldsymbol{x}_k$ 使得

$$\boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{\beta}_1,\quad \dots,\quad \boldsymbol{A}\boldsymbol{x}_k=\boldsymbol{\beta}_k,$$

即存在矩阵 $\boldsymbol{X}$ 使得 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$.

定理

设 S,T 分别为 $\boldsymbol{A},\boldsymbol{B}$ 的列向量组, 且分别生成空间 V,W.

(1) T 可以被 S 线性表示 $\iff W\subseteq V \iff \exists\boldsymbol{X}$ 使得 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{B}$.

(2) S,T 向量组等价 $\iff W=V \iff \exists\boldsymbol{X},\boldsymbol{Y}$ 使得 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{X}, \boldsymbol{A}=\boldsymbol{B}\boldsymbol{Y}$.

命题

向量组的等价满足如下性质:

(1) 自反性: $S \sim S$;

(2) 对称性: $S \sim T \implies T \sim S$;

(3) 传递性: $S \sim T, T \sim R \implies S \sim R$.

若矩阵 $\boldsymbol{A} \overset{c}{\sim} \boldsymbol{B}$ 列等价, 则存在可逆矩阵 $\boldsymbol{Q}$ 使得 $\boldsymbol{B} = \boldsymbol{A}\boldsymbol{Q}$, 于是二者的列向量组作为向量组等价. 但是反过来不成立. 这是因为列等价的矩阵一定是同型矩阵, 但等价的向量组并不要求向量数量相同. 不过, 同型矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 列向量组等价 $\iff \boldsymbol{A} \overset{c}{\sim} \boldsymbol{B}$. 我们稍后证明.

定义

对于 n 维向量组 $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m\}$, 若存在一组不全为零的数 $\lambda_1, \dots, \lambda_m$ 使得

$$\lambda_1\boldsymbol{\alpha}_1 + \cdots + \lambda_m\boldsymbol{\alpha}_m = \mathbf{0},$$

则称该向量组线性相关. 否则称该向量组线性无关.

例

(1) $\boldsymbol{\alpha}_1 = (1,2,3)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2 = (2,3,4)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3 = (0,0,0)^{\mathrm{T}}$ 线性相关. 包含零向量的向量组总是线性相关的.

(2) $\boldsymbol{e}_1 = (1,0,0)^{\mathrm{T}}, \boldsymbol{e}_2 = (0,1,0)^{\mathrm{T}}, \boldsymbol{e}_3 = (0,0,1)^{\mathrm{T}}$ 线性无关. 一般地, n 维基本向量组 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ 线性无关.

(3) $\boldsymbol{\alpha}$ 线性相关 $\Longleftrightarrow$ $\boldsymbol{\alpha} = \mathbf{0}$.

(4) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 线性相关 $\Longleftrightarrow$ $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 对应分量成比例 (共线).

向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 线性无关当且仅当

$$\lambda_1\boldsymbol{\alpha}_1 + \cdots + \lambda_m\boldsymbol{\alpha}_m = \mathbf{0} \implies \lambda_1 = \cdots = \lambda_m = 0,$$

即 $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 只有零解, 其中 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m)$.

若 $\boldsymbol{\beta} = \boldsymbol{A}\boldsymbol{\xi}$, 则 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{\beta} \iff \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{\xi}) = \mathbf{0}$. 因此:

定理

(1) 设 V 是 $\boldsymbol{A}$ 列向量生成的空间, 则以下结论等价:

- $\boldsymbol{A}$ 的列向量组线性无关;
- $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 只有零解;
- $\exists \boldsymbol{v} \in V$, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{v}$ 只有唯一解;
- $\forall \boldsymbol{v} \in V$, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{v}$ 只有唯一解.

(2) 向量组 S 是线性空间 V 的一组基 $\iff$ S 线性无关且生成 V.

例：线性无关和线性相关

练习

设 $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ 3 & 0 & 4 \end{pmatrix}, \boldsymbol{\alpha} = \begin{pmatrix} k \\ 1 \\ 1 \end{pmatrix}$. 若 $\boldsymbol{A}\boldsymbol{\alpha}$ 和 $\boldsymbol{\alpha}$ 线性相关，则 $k = \underline{\ -1\ }$.

练习

已知向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2\}$ 线性无关，请问向量组 $\{\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_1\}$ 是否线性无关？

答案

线性相关，因为 $(\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2) + (\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2) - 2\boldsymbol{\alpha}_1 = \boldsymbol{0}$.

例

已知向量组 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 线性无关，证明向量组 $\{\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2, \boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3, \boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1\}$ 线性无关.

证明

我们可以用定义来直接证明. 设

$$\lambda_1(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)+\lambda_2(\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3)+\lambda_3(\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1)=\mathbf{0}.$$

那么

$$(\lambda_1+\lambda_3)\boldsymbol{\alpha}_1+(\lambda_1+\lambda_2)\boldsymbol{\alpha}_2+(\lambda_2+\lambda_3)\boldsymbol{\alpha}_3=\mathbf{0}.$$

由于 $\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ 线性无关，因此

$$\lambda_1+\lambda_3=\lambda_1+\lambda_2=\lambda_2+\lambda_3=0,$$

解得 $\lambda_1=\lambda_2=\lambda_3=0$. 证毕. □

例：判断线性无关

我们来看另一种证法.

另证

我们有

$$(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{A},$$

其中 $\boldsymbol{A}=\begin{pmatrix}1&0&1\\1&1&0\\0&1&1\end{pmatrix}$. 若 $(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_1)\boldsymbol{x}=\boldsymbol{0}$, 则

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\implies \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$$

由于 $|\boldsymbol{A}|=2$, $\boldsymbol{A}$ 可逆, 因此 $\boldsymbol{x}=\boldsymbol{0}$. □

例：判断线性无关

命题

(1) 设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m$ 线性无关，$(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n)=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m)\boldsymbol{C}$. $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_n$ 线性无关 $\iff$ $\boldsymbol{C}\boldsymbol{x}=\mathbf{0}$ 只有零解.

(2) 设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m$ 线性无关，$(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_m)=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m)\boldsymbol{C}$. $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_m$ 线性无关 $\iff$ $|\boldsymbol{C}|\neq 0$.

(3) n 维向量组 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 线性无关 $\iff$ $|\boldsymbol{\alpha}_1,\cdots,\boldsymbol{\alpha}_n|\neq 0$.

例

$\boldsymbol{\alpha}_1=(1,2,3)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(2,3,4)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3=(3,5,7)^{\mathrm{T}}$ 线性相关. 因为它们构成的 3 阶矩阵行列式为零.

练习

(1) 若向量组 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(1,2,3)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3=(1,3,t)^{\mathrm{T}}$ 线性相关, 则 $t=$ __5__.

(2) 若任一 3 维向量都可由向量组 $\boldsymbol{\alpha}_1=(a,3,2)^{\mathrm{T}}$, $\boldsymbol{\alpha}_2=(2,-1,3)^{\mathrm{T}}$, $\boldsymbol{\alpha}_3=(3,2,1)^{\mathrm{T}}$ 线性表示, 则 $a\neq$ __5__.

定理

向量组 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m$ 线性相关 $\Longleftrightarrow$ 其中至少有一个向量可以由其它向量线性表示.

证明

若该向量组线性相关, 则存在不全为零的数 $\lambda_1,\dots,\lambda_m$ 使得

$$\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m=\mathbf{0}.$$

设 $\lambda_i\neq 0$, 则 $\boldsymbol{\alpha}_i=-\dfrac{1}{\lambda_i}\sum\limits_{\substack{j=1\\j\neq i}}^{m}\lambda_j\boldsymbol{\alpha}_j$ 可由其它向量线性表示.

反之, 若 $\boldsymbol{\alpha}_i=\sum\limits_{\substack{j=1\\j\neq i}}^{m}\lambda_j\boldsymbol{\alpha}_j$ 可由其它向量线性表示. 则 $-\boldsymbol{\alpha}_i+\sum\limits_{\substack{j=1\\j\neq i}}^{m}\lambda_j\boldsymbol{\alpha}_j=\mathbf{0}$, 向量组 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m$ 线性相关. □

线性相关和线性无关的等价刻画

向量组 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性无关 $\Longleftrightarrow$ 其中任一向量不可以由其它向量线性表示.

注意, 向量组 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性相关 $\not\Longrightarrow$ 其中任一向量可以由其它向量线性表示.

练习

设 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 是 m 个 n 维向量, 则下列结论正确的有 __1__ 个.

(1) 若 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性相关, 则其中任一向量均可由其余向量线性表示

(2) 若 $\boldsymbol{\alpha}_m$ 不能由 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{m-1}$ 线性表示, 则向量组 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性无关

(3) 若 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性相关, 且存在不全为零的 $\lambda_1, \dots, \lambda_{m-1}$ 使得 $\lambda_1\boldsymbol{\alpha}_1 + \cdots + \lambda_{m-1}\boldsymbol{\alpha}_{m-1} = \mathbf{0}$, 则 $\boldsymbol{\alpha}_m$ 不能由 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{m-1}$ 线性表示

(4) 若 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性相关, 且 $\boldsymbol{\alpha}_m$ 不能由 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_m$ 线性表示, 则 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_{m-1}$ 线性相关

定理

若向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 线性无关，向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}$ 线性相关，则 $\boldsymbol{\beta}$ 可以由 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 线性表示，且表达形式唯一.

证明

由于向量组 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}$ 线性相关，因此存在不全为零的数 $\lambda_1, \ldots, \lambda_m, k$ 使得

$$\lambda_1 \boldsymbol{\alpha}_1 + \cdots + \lambda_m \boldsymbol{\alpha}_m + k\boldsymbol{\beta} = \mathbf{0}.$$

若 $k = 0$，则由 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 线性无关可知 $\lambda_1 = \cdots = \lambda_k = 0$. 这与 $\lambda_1, \ldots, \lambda_m, k$ 不全为零矛盾. 因此 $k \neq 0$，

$$\boldsymbol{\beta} = -\frac{1}{k}(\lambda_1 \boldsymbol{\alpha}_1 + \cdots + \lambda_m \boldsymbol{\alpha}_m).$$

由 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 线性无关可知线性组合表达方式唯一. □

定理

设向量组 $S=\{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m\}, T=\{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m,\boldsymbol{\alpha}_{m+1},\dots,\boldsymbol{\alpha}_s\}$.

(1) 若向量组 S 线性相关，则 T 也线性相关.

(2) 若向量组 T 线性无关，则 S 也线性无关.

即部分相关 $\Longrightarrow$ 整体相关，整体无关 $\Longrightarrow$ 部分无关.

例

n 维向量组 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s (3\leqslant s\leqslant n)$ 线性无关 $\Longleftrightarrow$ (D).

(A) $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s$ 中存在一个向量不能由其余向量线性表示 必要

(B) $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s$ 中任两个向量都线性无关 必要

(C) $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s$ 中不含零向量 必要

(D) $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s$ 中任一个向量都不能由其余向量线性表示

练习

若向量组 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性无关，$\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\delta}$ 线性相关，则(C).

(A) $\boldsymbol{\alpha}$ 一定能由 $\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta}$ 线性表示

(B) $\boldsymbol{\beta}$ 一定不能由 $\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\delta}$ 线性表示

(C) $\boldsymbol{\delta}$ 一定能由 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性表示

(D) $\boldsymbol{\delta}$ 一定不能由 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 线性表示

例

设向量 $\boldsymbol{\beta}$ 可由 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m$ 线性表示，但不能由向量组 $S=\{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{m-1}\}$ 线性表示．记 $T=\{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{m-1},\boldsymbol{\beta}\}$，则(B). $\boldsymbol{\beta}=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m, \lambda_m\neq 0$

(A) $\boldsymbol{\alpha}_m$ 不能由 S 线性表示，也不能由 T 线性表示

(B) $\boldsymbol{\alpha}_m$ 不能由 S 线性表示，但能由 T 线性表示

(C) $\boldsymbol{\alpha}_m$ 能由 S 线性表示，也能由 T 线性表示

(D) $\boldsymbol{\alpha}_m$ 能由 S 线性表示，但不能由 T 线性表示

例

设向量组 $\alpha_1, \alpha_2, \alpha_3$ 线性相关，$\alpha_2, \alpha_3, \alpha_4$ 线性无关，证明

(1) α_1 能由 α_2, α_3 线性表示；

(2) α_4 不能由 $\alpha_1, \alpha_2, \alpha_3$ 线性表示.

证明

(1) 由 $\alpha_2, \alpha_3, \alpha_4$ 线性无关可知 α_2, α_3 线性无关.但是 $\alpha_1, \alpha_2, \alpha_3$ 线性相关，所以 α_1 能由 α_2, α_3 线性表示.

(2) 若 α_4 能由 $\alpha_1, \alpha_2, \alpha_3$ 线性表示，由于 α_1 能由 α_2, α_3 线性表示，于是 α_4 也能由 α_2, α_3 线性表示. 这与 $\alpha_2, \alpha_3, \alpha_4$ 线性无关矛盾. □

定理

设 $\boldsymbol{\alpha}_j=(a_{1j},\dots,a_{nj})^{\mathrm{T}},\boldsymbol{\beta}_j=(a_{1j},\dots,a_{nj},a_{n+1,j})^{\mathrm{T}}$.

(1) 若向量组 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m$ 线性无关, 则 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m$ 线性无关.

(2) 若向量组 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m$ 线性相关, 则 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m$ 线性相关.

证明

设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m),\boldsymbol{B}=(\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m)$, 则存在 m 维向量 $\boldsymbol{\gamma}$ 使得 $\boldsymbol{B}=\begin{pmatrix}\boldsymbol{A}\\ \boldsymbol{\gamma}^{\mathrm{T}}\end{pmatrix}$. 若向量组 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m$ 线性无关, 则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 只有零解. 而

$$\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}\iff\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0},\boldsymbol{\gamma}^{\mathrm{T}}\boldsymbol{x}=0,$$

因此 $\boldsymbol{x}=\boldsymbol{0}$, $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 只有零解, $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m$ 线性无关. □

即高维相关 $\Longrightarrow$ 低维相关, 低维无关 $\Longrightarrow$ 高维无关.

例：线性相关和线性无关

练习

(1) 判断下列向量组的线性相关性：

(i) $(1,2,3,4)^{\mathrm{T}},(2,3,4,5)^{\mathrm{T}},(0,0,0,0)^{\mathrm{T}}$. 相关

(ii) $(a,b,1,0,0)^{\mathrm{T}},(c,d,0,6,0)^{\mathrm{T}},(a,c,0,5,6)^{\mathrm{T}}$. 无关

(iii) $(a,1,0,b,0)^{\mathrm{T}},(c,0,6,d,0)^{\mathrm{T}},(a,0,5,c,6)^{\mathrm{T}}$. 无关

(2) 若 $(1,0,0,2)^{\mathrm{T}},(0,1,5,0)^{\mathrm{T}},(2,1,t+2,4)^{\mathrm{T}}$ 线性相关，则 $t=$___3___.

定理

设向量组 $S=\{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s\}$ 可由 $T=\{\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_t\}$ 线性表示.

(1) 若 $s>t$, 则 S 线性相关.

(2) 若 S 线性无关, 则 $s\leqslant t$.

即多的由少的表示, 多的一定线性相关.

证明

设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s), \boldsymbol{B}=(\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_t)$. 则存在矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A}=\boldsymbol{B}\boldsymbol{P}$. 由于 $\boldsymbol{P}$ 行数小于列数, 因此 $\boldsymbol{P}\boldsymbol{x}=\boldsymbol{0}$ 有非零解 $\boldsymbol{x}$. 从而 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{B}\boldsymbol{P}\boldsymbol{x}=\boldsymbol{0}$, S 线性相关. □

推论

(1) $m>n$ 个 n 维向量一定线性相关.

(2) 任意两个等价的线性无关向量组所含向量的个数相同.

(1) 向量组线性相关 $\Longleftrightarrow$ 其中至少有一个向量可以由其它向量线性表示.

(2) 若 S 线性无关, $S \cup \{\beta\}$ 线性相关, 则 β 可以由 S 唯一线性表示.

(3) 部分相关 $\Longrightarrow$ 整体相关, 整体无关 $\Longrightarrow$ 部分无关.

(4) 高维相关 $\Longrightarrow$ 低维相关, 低维无关 $\Longrightarrow$ 高维无关.

(5) 多的由少的表示, 多的一定线性相关.

定义

(1) 若 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 是线性空间 V 的一组基，则称 m 为 V 的**维数**，记作 $\dim V$.

(2) 设向量组 S 生成空间 V. 称 V 的维数为该向量组的**秩(Rank)**，记作 $\mathrm{R}(S)$.

由前面的结论可知，V 不同的基向量组的个数总是相同的，即维数是唯一的.

设 $S = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m\}$ 生成 V，T 是 V 的一组基. 由于向量组等价 $\Longleftrightarrow$ 生成同一个空间，因此 S, T 是等价向量组. 由 T 线性无关可知 $\mathrm{R}(S) \leqslant m$.

定理

设 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m)$ 的列向量组为 S.

(1) S 线性无关 $\Longleftrightarrow$ $\mathrm{R}(S) = m$ $\Longleftrightarrow$ $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 只有零解.

(2) S 线性相关 $\Longleftrightarrow$ $\mathrm{R}(S) < m$ $\Longleftrightarrow$ $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ 有非零解.

例：方阵的列向量组

推论

设方阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n)$ 的列向量组为 S.

(1) S 线性无关 $\iff \mathrm{R}(S)=m \iff \boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 只有零解 $\iff |\boldsymbol{A}|\neq 0$.

(2) S 线性相关 $\iff \mathrm{R}(S)<m \iff \boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 有非零解 $\iff |\boldsymbol{A}|=0$.

练习

设 $\boldsymbol{A}$ 是 n 阶方阵，且其行列式 $|\boldsymbol{A}|=0$. 下列说法正确的是(C).

(A) $\boldsymbol{A}$ 必有一列元素全为零

(B) $\boldsymbol{A}$ 必有两列元素对应成比例

(C) $\boldsymbol{A}$ 必有一个列向量可由其余列向量线性表示

(D) $\boldsymbol{A}$ 中任意列向量均可由其余列向量线性表示

定理

(1) 设向量组 S 可由向量组 T 线性表示，则 $\mathrm{R}(S) \leqslant \mathrm{R}(T)$.

(2) 若线性空间 $V \subseteq W$，则 $\dim V \leqslant \dim W$.

(3) 设向量组 S 可由向量组 T 线性表示，且 $\mathrm{R}(S) = \mathrm{R}(T)$，则 S, T 向量组等价.

(4) 若线性空间 $V \subseteq W$ 且 $\dim V = \dim W$，则 $V = W$.

我们来证明(4). 设 S, T 是 V, W 的一组基. 那么 S, T 大小相同，且 S 可由 T 线性表示. 设 S, T 分别是 $\boldsymbol{A}, \boldsymbol{B}$ 的列向量组，那么存在方阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A} = \boldsymbol{B}\boldsymbol{P}$. 若 $\boldsymbol{P}$ 不可逆，存在非零向量 $\boldsymbol{x}$ 使得 $\boldsymbol{P}\boldsymbol{x} = \boldsymbol{0}$. 于是 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{B}\boldsymbol{P}\boldsymbol{x} = \boldsymbol{0}$，$S$ 线性相关，矛盾!

定理

设 V 是 n 维空间，S 是由其中向量构成向量组. 那么 S 是一组基当且仅当如下任意两条满足 (剩下一条自动成立):

(1) S 大小是 n;　　(2) S 生成 V;　　(3) S 线性无关.

例

若

$$S_1=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3\},\quad S_2=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4\},\quad S_3=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5\}$$

满足 $\mathrm{R}(S_1)=\mathrm{R}(S_2)=3,\mathrm{R}(S_3)=4$. 证明向量组 $S=\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5-\boldsymbol{\alpha}_4\}$ 线性无关.

证明

由 $\mathrm{R}(S_1)=3$ 可知 S_1 线性无关. 由 $\mathrm{R}(S_2)=3$ 可知 S_2 线性相关. 从而 $\boldsymbol{\alpha}_4$ 可由 S_1 线性表示. 于是 S_3 可由 S 线性表示. 显然 S 可由 S_3 线性表示, 因此二者等价, $\mathrm{R}(S)=\mathrm{R}(S_3)=4$. □

练习

判断题: 设 S 和 T 为两个 n 维向量组, 且 $\mathrm{R}(S)=\mathrm{R}(T)$, 则 S 和 T 等价. ✗

如何从一组能生成空间 V 的向量找到 V 的一组基？我们只需要取极大线性无关组.

定义

设 S 为一个向量组. 若 S 的部分组 $S_0 = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m\}$ 满足

(1) S_0 线性无关;

(2) S_0 添加 S 中的若干向量得到的向量组均线性相关.

则称 S_0 是 S 的一个**极大线性无关组**.

根据上一节相关结论可知, S 中所有向量均可由 S_0 线性表示. 换言之, S_0 和 S 等价, 它们生成相同的子空间 V, $m = \mathrm{R}(S)$, S_0 是 V 的一组基.

极大线性无关组的定义

定理

S_0 是 S 的极大线性无关组当且仅当

(1) S_0 线性无关;

(2) S 中任意 $m+1$ 个向量线性相关.

证明

若 S_0 是 S 的极大线性无关组, 则 S 中任意 $m+1$ 个向量可由 S_0 线性表示. 从而线性相关.

反之, 若 S 中任意 $m+1$ 个向量线性相关, 则 S 中任意 $s>m$ 个向量线性相关. 于是 S_0 添加 S 中的若干向量得到的向量组均线性相关. $\square$

(1) 若 $\mathrm{R}(S) = r$, 则 S 中任意 r 个线性无关的向量构成 S 的一个极大线性无关组.

(2) 只含有零向量的向量组没有极大线性无关组 (空集), 它的秩为 0 (空集生成 0 维空间 $\{\mathbf{0}\}$).

(3) 极大线性无关组一般不是唯一的. 例如

$$\boldsymbol{\alpha}_1 = (1, 0)^{\mathrm{T}}, \quad \boldsymbol{\alpha}_2 = (0, 1)^{\mathrm{T}}, \quad \boldsymbol{\alpha}_3 = (1, 1)^{\mathrm{T}}.$$

$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 是一个极大线性无关组, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_3$ 也是一个极大线性无关组.

(4) 向量组和它的一个极大线性无关组是等价的, 于是同一向量组的任意两个极大线性无关组等价.

设 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_r),\boldsymbol{B}=(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_r)$ 的列向量组是等价的线性无关组. 我们之前证明过若 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{P}$, 则 $\boldsymbol{P}$ 可逆.

一般地, 若 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m),\boldsymbol{B}=(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_m)$ 的列向量组是等价的向量组, 秩为 r. 通过适当的列变换, 可以让 $\boldsymbol{A}$ 的前 r 列是极大无关组, 后面全是零向量. 设 $\boldsymbol{A}=(\boldsymbol{A}_r',\boldsymbol{O})$. 对 $\boldsymbol{B}$ 作类似操作, 则 $\boldsymbol{A}'\overset{c}{\sim}\boldsymbol{B}'$, 即存在可逆矩阵 $\boldsymbol{P}\in M_r$ 使得 $\boldsymbol{B}'=\boldsymbol{A}'\boldsymbol{P}$. 于是

$$\boldsymbol{B}=\boldsymbol{A}\begin{pmatrix}\boldsymbol{P} & \\ & \boldsymbol{E}\end{pmatrix}\overset{c}{\sim}\boldsymbol{A}.$$

因此**同型矩阵列 (行) 向量组等价 $\Longleftrightarrow$ 列 (行) 等价**.

例

矩阵 $\boldsymbol{A}=\begin{pmatrix}1&1&3&1\\0&1&-1&4\\0&0&0&5\\0&0&0&0\end{pmatrix}$ 的行向量组为

$$\boldsymbol{\alpha}_1^{\mathrm{T}}=(1,1,3,1),\ \boldsymbol{\alpha}_2^{\mathrm{T}}=(0,1,-1,4),\ \boldsymbol{\alpha}_3^{\mathrm{T}}=(0,0,0,5),\ \boldsymbol{\alpha}_4^{\mathrm{T}}=(0,0,0,0).$$

由于 $\boldsymbol{\alpha}_1^{\mathrm{T}},\boldsymbol{\alpha}_2^{\mathrm{T}},\boldsymbol{\alpha}_3^{\mathrm{T}}$ 的第 $1,2,4$ 个分量形成可逆矩阵 $\begin{pmatrix}1&1&1\\0&1&4\\0&0&5\end{pmatrix}$，因此它们线性无关. 它们构成一个极大线性无关组，$\boldsymbol{A}$ 的行向量组的秩是 3. 类似可知，$\boldsymbol{A}$ 的列向量组的秩也是 3.

实际上，任意矩阵的行向量组的秩等于列向量组的秩. 为了说明这一点，我们考虑矩阵的秩.

第二节　矩阵的秩

- 矩阵秩的定义
- 矩阵秩的性质
- 极大线性无关组的计算方法

我们知道, 每个矩阵 $\boldsymbol{A}$ 都等价于某个标准型 $\begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$. 称 r 为 $\boldsymbol{A}$ 的秩, 记作 $\mathrm{R}(\boldsymbol{A})$.

称 $\boldsymbol{A}$ 的行向量组的秩为行秩, 列向量组的秩为列秩.

行秩等于列秩等于秩

$\boldsymbol{A}$ 的行秩和列秩均等于秩 $\mathrm{R}(\boldsymbol{A})$.

由此可知矩阵的秩是唯一确定的.

对于行阶梯形矩阵, 再实施初等变换使其变为行最简形矩阵或标准型矩阵, 并不会改变它的非零行的个数. 换言之, 行阶梯形矩阵的秩就是非零行的个数.

设 $\boldsymbol{A}$ 通过初等行变换变为行阶梯形矩阵 $\boldsymbol{B}$, 则二者秩相等, 二者的行向量组等价, 从而行秩也相等. 对于 $\boldsymbol{B}$, 它的行秩就是非零行的个数, 也就是 $\mathrm{R}(\boldsymbol{B})$. 因此 $\boldsymbol{A}$ 的行秩等于秩. 不难知道 $\mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}^{\mathrm{T}})$, 从而 $\boldsymbol{A}$ 的列秩 $= \boldsymbol{A}^{\mathrm{T}}$ 的行秩 $= \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{R}(\boldsymbol{A})$.

例

求矩阵 $\boldsymbol{A}=\begin{pmatrix}2 & -1 & 0 & 3 & -2\\ 0 & 3 & 1 & -2 & 5\\ 0 & 0 & 0 & 4 & -3\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}, \boldsymbol{B}=\begin{pmatrix}1 & 2 & 3\\ 2 & 3 & -5\\ 4 & 7 & 1\end{pmatrix}$ 的秩.

解

$\boldsymbol{A}$ 是行阶梯形矩阵, 因此 $\mathrm{R}(\boldsymbol{A})=3$.

$$\boldsymbol{B}\overset{r_2-2r_1}{\underset{r_3-4r_1}{\sim}}\begin{pmatrix}1 & 2 & 3\\ 0 & -1 & -11\\ 0 & -1 & -11\end{pmatrix}\overset{r_3-r_2}{\underset{-r_2}{\sim}}\begin{pmatrix}1 & 2 & 3\\ 0 & 1 & 11\\ 0 & 0 & 0\end{pmatrix}\Longrightarrow \mathrm{R}(\boldsymbol{B})=2.$$

例

求矩阵 $\boldsymbol{A}=\begin{pmatrix}1&1&a\\-1&a-1&1-a\\1&1&a^2\\1&1&2a+1\end{pmatrix}$ 的秩.

解

$$\boldsymbol{A}\underset{r_4-r_1}{\overset{\substack{r_2+r_1\\r_3-r_1}}{\sim}}\begin{pmatrix}1&1&a\\0&a&1\\0&0&a^2-a\\0&0&a+1\end{pmatrix}\underset{-\frac{1}{2}r_3}{\overset{r_3-ar_4}{\sim}}\begin{pmatrix}1&1&a\\0&a&1\\0&0&a\\0&0&a+1\end{pmatrix}\sim\begin{pmatrix}1&1&a\\0&a&1\\0&0&1\\0&0&0\end{pmatrix}$$

因此 $a\neq 0$ 时，$\mathrm{R}(\boldsymbol{A})=3$; $a=0$ 时，$\mathrm{R}(\boldsymbol{A})=2$.

注意处理带未知数的矩阵时，不宜实施 $\dfrac{1}{a+1}r_2, (a-2)r_3$ 等类似操作，因为其分母或系数可能为零.

练习

求矩阵 $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & -2 & 3 \\ 2 & 1 & -6 & 4 \\ 3 & 2 & m & 7 \end{pmatrix}$ 的秩.

答案

$m \neq -8$ 时，$\mathrm{R}(\boldsymbol{A}) = 3$；$m = -8$ 时，$\mathrm{R}(\boldsymbol{A}) = 2$.

例：计算矩阵的秩

例

求矩阵 $\boldsymbol{A}=\begin{pmatrix} a & 1 & 1 & 1\\ 1 & a & 1 & 1\\ 1 & 1 & a & 1\\ 1 & 1 & 1 & a \end{pmatrix}$ 的秩.

解

$$\boldsymbol{A} \underset{\substack{r_3 - r_1\\ r_4 - a r_1}}{\overset{\substack{r_1 \leftrightarrow r_4\\ r_2 - r_1}}{\sim}} \begin{pmatrix} 1 & 1 & 1 & a\\ 0 & a-1 & 0 & 1-a\\ 0 & 0 & a-1 & 1-a\\ 0 & 1-a & 1-a & 1-a^2 \end{pmatrix} \underset{r_4 + r_3}{\overset{r_4 + r_2}{\sim}} \begin{pmatrix} 1 & 1 & 1 & a\\ 0 & a-1 & 0 & 1-a\\ 0 & 0 & a-1 & 1-a\\ 0 & 0 & 0 & -(a+3)(a-1) \end{pmatrix}$$

因此 $a \neq 1, -3$ 时, $\mathrm{R}(\boldsymbol{A})=4$; $a=-3$ 时, $\mathrm{R}(\boldsymbol{A})=3$; $a=1$ 时, $\mathrm{R}(\boldsymbol{A})=1$.

矩阵秩有另一种刻画方式. 矩阵 $\boldsymbol{A}$ 任取 k 行 k 列交叉得到的 k^2 个元素 (不改变位置次序) 形成的 k 阶方阵的行列式, 称为 $\boldsymbol{A}$ 的 k **阶子式**. 例如 n 阶方阵的余子式是 $n-1$ 阶子式.

定理

设 $\mathrm{R}(\boldsymbol{A})=r$, 则存在非零的 r 阶子式, 但所有的 $r+1$ 阶子式都是零.

根据行列式的拉普拉斯展开, 若 $\boldsymbol{A}$ 的 k 阶子式均为零, 则 $k+1$ 阶子式也都是零. 因此 $\boldsymbol{A}$ 的任意 $s>r$ 阶子式都是零.

推论

(1) $\mathrm{R}(\boldsymbol{A})\geqslant r \iff \boldsymbol{A}$ 存在非零 r 阶子式.

(2) $\mathrm{R}(\boldsymbol{A})\leqslant r \iff \boldsymbol{A}$ 所有 $r+1$ 阶子式均为零.

(3) $\mathrm{R}(\boldsymbol{A})=r \implies \boldsymbol{A}$ 存在 $1,2,\ldots,r$ 阶非零子式.

证明

设 $\boldsymbol{B}=\boldsymbol{P}\boldsymbol{A}$, 其中 $\boldsymbol{P}$ 是初等矩阵.

(1) 若 $\boldsymbol{P}=\boldsymbol{E}(i,j)$, 则 $\boldsymbol{B}$ 的 k 阶子式总等于 $\boldsymbol{A}$ 的某个 k 阶子式, 最多相差 -1.

(2) 若 $\boldsymbol{P}=\boldsymbol{E}(i(a))$, 则 $\boldsymbol{B}$ 的 k 阶子式总等于 $\boldsymbol{A}$ 的某个 k 阶子式或 a 倍.

(3) 若 $\boldsymbol{P}=\boldsymbol{E}(i,j(a))$, 则 $\boldsymbol{B}$ 的 k 阶子式总等于 $\boldsymbol{A}$ 的某个 k 阶子式.

因此若 $\boldsymbol{A}$ 的 k 阶子式都是零, 则 $\boldsymbol{B}$ 的 k 阶子式也都是零.

由于 $\boldsymbol{P}^{-1}$ 也是初等矩阵, 因此反过来也成立. 对于 $\boldsymbol{B}=\boldsymbol{A}\boldsymbol{P}$ 情形同理. 因此, 若 $\boldsymbol{A}\sim\boldsymbol{B}$, 则 $\boldsymbol{A}$ 的 k 阶子式都是零 $\Longleftrightarrow$ $\boldsymbol{B}$ 的 k 阶子式都是零.

对于标准型矩阵, 该定理显然成立. 因此该定理对任意矩阵都成立. $\square$

命题

设 $\boldsymbol{A} \in M_{m\times n}$, 则 $0 \leqslant \mathrm{R}(\boldsymbol{A}) \leqslant \min(m,n)$.

定义

(1) 若 $\mathrm{R}(\boldsymbol{A}) = m$, 称 $\boldsymbol{A}$ 行满秩;

(2) 若 $\mathrm{R}(\boldsymbol{A}) = n$, 称 $\boldsymbol{A}$ 列满秩;

(3) 若 $\mathrm{R}(\boldsymbol{A}) = m = n$, 称 $\boldsymbol{A}$ 满秩.

命题

(1) $\mathrm{R}(\boldsymbol{A}) = 0 \iff \boldsymbol{A} = \boldsymbol{O}$;

(2) n 阶方阵 $\boldsymbol{A}$ 可逆 $\iff \mathrm{R}(\boldsymbol{A}) = n$;

(3) $\mathrm{R}(k\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}), k \neq 0$;

(4) $\boldsymbol{A} \sim \boldsymbol{B} \iff \mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{B})$;

(5) $\mathrm{R}(\boldsymbol{A}\boldsymbol{B}) \leqslant \min(\mathrm{R}(\boldsymbol{A}), \mathrm{R}(\boldsymbol{B}))$;

(6) 若 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times \ell} = \boldsymbol{O}$, 则 $\mathrm{R}(\boldsymbol{A}) + \mathrm{R}(\boldsymbol{B}) \leqslant n$;

(7) $\mathrm{R}(a\boldsymbol{A} + b\boldsymbol{B}) \leqslant \mathrm{R}(\boldsymbol{A}, \boldsymbol{B}) \leqslant \mathrm{R}(\boldsymbol{A}) + \mathrm{R}(\boldsymbol{B})$.
特别地, $\max(\mathrm{R}(\boldsymbol{A}), \mathrm{R}(\boldsymbol{B})) \leqslant \mathrm{R}(\boldsymbol{A}, \boldsymbol{B})$.

命题

(5) $\mathrm{R}(\boldsymbol{AB}) \leqslant \min(\mathrm{R}(\boldsymbol{A}), \mathrm{R}(\boldsymbol{B}))$.

证明

$\boldsymbol{AB}$ 的列向量为 $\boldsymbol{A}$ 列向量组的线性组合, 从而 $\boldsymbol{AB}$ 的列秩 $\leqslant$ $\boldsymbol{A}$ 的列秩, 即 $\mathrm{R}(\boldsymbol{AB}) \leqslant \mathrm{R}(\boldsymbol{A})$. 于是

$$\mathrm{R}(\boldsymbol{AB}) = \mathrm{R}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}) \leqslant \mathrm{R}(\boldsymbol{B}^{\mathrm{T}}) = \mathrm{R}(\boldsymbol{B}).$$ □

若 $\boldsymbol{B}$ 行满秩, 则 $\boldsymbol{B}$ 有 $\mathrm{R}(\boldsymbol{B})$ 阶子式非零, 它对应的方阵右乘 $\boldsymbol{A}$ 得到的列向量组和 $\boldsymbol{A}$ 列向量组等价, 从而 $\mathrm{R}(\boldsymbol{AB}) = \mathrm{R}(\boldsymbol{A})$; 若 $\boldsymbol{B}$ 列满秩, 则 $\mathrm{R}(\boldsymbol{BA}) = \mathrm{R}(\boldsymbol{A})$;

命题

(6) 若 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times \ell}=\boldsymbol{O}$，则 $\mathrm{R}(\boldsymbol{A})+\mathrm{R}(\boldsymbol{B})\leqslant n$.

证明

我们将会 §2.4 证明空间

$$V=\{\boldsymbol{x}\in\mathbb{R}^n\mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\}$$

的维数是 $n-\mathrm{R}(\boldsymbol{A})$. 设

$$W=\{\boldsymbol{B}\boldsymbol{y}\mid \boldsymbol{y}\in\mathbb{R}^\ell\}$$

是 $\boldsymbol{B}$ 列向量生成的空间，则 W 的维数是 $\mathrm{R}(\boldsymbol{B})$. 由 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}$ 可知 $W\subseteq V$，因此 $\mathrm{R}(\boldsymbol{B})\leqslant n-\mathrm{R}(\boldsymbol{A})$. □

命题

(7) $\mathrm{R}(a\boldsymbol{A}+b\boldsymbol{B}) \leqslant \mathrm{R}(\boldsymbol{A},\boldsymbol{B}) \leqslant \mathrm{R}(\boldsymbol{A})+\mathrm{R}(\boldsymbol{B})$.
特别地，$\max(\mathrm{R}(\boldsymbol{A}),\mathrm{R}(\boldsymbol{B})) \leqslant \mathrm{R}(\boldsymbol{A},\boldsymbol{B})$.

证明

设 S, T 分别为 $\boldsymbol{A},\boldsymbol{B}$ 列向量组的极大线性无关组，则 S, T 的大小分别是 $\mathrm{R}(\boldsymbol{A})$, $\mathrm{R}(\boldsymbol{B})$. 由于 $(\boldsymbol{A},\boldsymbol{B})$ 的列向量组和 $S\cup T$ 向量组等价，因此

$$\mathrm{R}(\boldsymbol{A},\boldsymbol{B}) = \mathrm{R}(S\cup T) \leqslant \mathrm{R}(\boldsymbol{A})+\mathrm{R}(\boldsymbol{B}).$$

由于 $a\boldsymbol{A}+b\boldsymbol{B}$ 的列向量组可以由 $(\boldsymbol{A},\boldsymbol{B})$ 的列向量组线性表示，因此

$$\mathrm{R}(a\boldsymbol{A}+b\boldsymbol{B}) \leqslant \mathrm{R}(\boldsymbol{A},\boldsymbol{B}).$$

□

另证

由于添加零行或零列不改变秩，因此不妨设 $\boldsymbol{A},\boldsymbol{B}$ 都是方阵. 由于

$$a\boldsymbol{A}+b\boldsymbol{B}=(\boldsymbol{A},\boldsymbol{B})\begin{pmatrix} a\boldsymbol{E} & \\ & b\boldsymbol{E}\end{pmatrix},\quad (\boldsymbol{A},\boldsymbol{B})=(\boldsymbol{E},\boldsymbol{E})\begin{pmatrix} \boldsymbol{A} & \\ & \boldsymbol{B}\end{pmatrix}.$$

因此 $\mathrm{R}(a\boldsymbol{A}+b\boldsymbol{B})\leqslant \mathrm{R}(\boldsymbol{A},\boldsymbol{B})\leqslant \mathrm{R}(\boldsymbol{A})+\mathrm{R}(\boldsymbol{B})$. □

练习

(1) 设 $\mathrm{R}(\boldsymbol{A})=2, \boldsymbol{B}=\begin{pmatrix}1 & 0 & 2\\ 0 & 2 & 0\\ -1 & 0 & 3\end{pmatrix}$, 则 $\mathrm{R}(\boldsymbol{AB})=$ __2__.

(2) 若 $\boldsymbol{A}$ 是 n 阶方阵且 $\mathrm{R}(\boldsymbol{AB})<\mathrm{R}(\boldsymbol{B})$, 则 $|\boldsymbol{A}|=$ __0__.

(3) 若 $\boldsymbol{A}=\begin{pmatrix}1 & 0 & 2\\ 0 & 1 & 1\\ 2 & 0 & 5\end{pmatrix}, \boldsymbol{B}=\begin{pmatrix}3 & 4 & 1\\ 1 & 2 & 1\\ 4 & 6 & t\end{pmatrix}$, $\boldsymbol{AX}=\boldsymbol{B}$ 且 $\mathrm{R}(\boldsymbol{X})=2$, 则 $t=$ __2__.

(4) 若 $\boldsymbol{A}=\begin{pmatrix}t & 2 & 3\\ 2 & 1 & -1\\ 0 & 0 & 5\end{pmatrix}$ 且存在非零矩阵 $\boldsymbol{B}$ 使得 $\boldsymbol{AB}=\boldsymbol{O}$, 则 $t=$ __4__.

例：矩阵秩性质的应用

例

证明：若 n 阶方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2 = \boldsymbol{A}$，则 $\mathrm{R}(\boldsymbol{A}) + \mathrm{R}(\boldsymbol{A} - \boldsymbol{E}) = n$.

证明

由于 $\boldsymbol{A}(\boldsymbol{A} - \boldsymbol{E}) = \boldsymbol{A}^2 - \boldsymbol{A} = \boldsymbol{O}$，因此

$$\mathrm{R}(\boldsymbol{A}) + \mathrm{R}(\boldsymbol{A} - \boldsymbol{E}) \leqslant n.$$

由于 $\boldsymbol{A} + (\boldsymbol{E} - \boldsymbol{A}) = \boldsymbol{E}$，因此

$$n = \mathrm{R}(\boldsymbol{E}) \leqslant \mathrm{R}(\boldsymbol{A}) + \mathrm{R}(\boldsymbol{E} - \boldsymbol{A}).$$

故 $\mathrm{R}(\boldsymbol{A}) + \mathrm{R}(\boldsymbol{A} - \boldsymbol{E}) = n$. □

伴随矩阵的秩

设 $\boldsymbol{A}$ 是 n 阶方阵，则

$$\mathrm{R}(\boldsymbol{A}^*)=\begin{cases} n, & \mathrm{R}(\boldsymbol{A})=n; \\ 1, & \mathrm{R}(\boldsymbol{A})=n-1; \\ 0, & \mathrm{R}(\boldsymbol{A})\leqslant n-2. \end{cases}$$

证明

(1) 若 $\mathrm{R}(\boldsymbol{A})=n$，$\boldsymbol{A}$ 可逆，从而 $\boldsymbol{A}^*$ 可逆，$\mathrm{R}(\boldsymbol{A}^*)=n$.

(2) 若 $\mathrm{R}(\boldsymbol{A})=n-1$，由 $\boldsymbol{A}\boldsymbol{A}^*=|\boldsymbol{A}|\boldsymbol{E}=\boldsymbol{O}$ 可知 $\mathrm{R}(\boldsymbol{A}^*)\leqslant 1$. 由于 $\mathrm{R}(\boldsymbol{A})=n-1$，$\boldsymbol{A}$ 存在非零的 $n-1$ 子式，从而 $\boldsymbol{A}^*\neq\boldsymbol{O}$. 故 $\mathrm{R}(\boldsymbol{A}^*)=1$.

(3) 若 $\mathrm{R}(\boldsymbol{A})\leqslant n-2$，则 $\boldsymbol{A}$ 的 $n-1$ 子式均为零，从而 $\boldsymbol{A}^*=\boldsymbol{O}$. □

例：矩阵秩性质的应用

练习

(1) 设 $\boldsymbol{\alpha}=(1,0,-1,2)^{\mathrm{T}}, \boldsymbol{\beta}=(0,1,0,2)^{\mathrm{T}}$，则 $\mathrm{R}(\boldsymbol{\alpha}\boldsymbol{\beta}^{\mathrm{T}})=$ __1__.

(2) 若 $\boldsymbol{A}=\begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}$ 且 $\mathrm{R}(\boldsymbol{A}^*)=1$，则(B).

(A) $a\neq b, a+2b\neq 0$　　(B) $a\neq b, a+2b=0$

(C) $a=b, a\neq 0$　　(D) $a=b=0$

(3) 设 $\boldsymbol{A},\boldsymbol{B}$ 均为 n 阶非零矩阵，且 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}$，则 $\mathrm{R}(\boldsymbol{A})$ 与 $\mathrm{R}(\boldsymbol{B})$(B).

(A) 必有一个等于 0　　(B) 都小于 n

(C) 都等于 n　　(D) 一个小于 n，一个等于 n

例：矩阵秩性质的应用

练习

(4) 设 $\boldsymbol{P}$ 为 3 阶非零矩阵，$\boldsymbol{Q}=\begin{pmatrix}1&2&3\\2&4&t\\3&6&9\end{pmatrix}$ 且 $\boldsymbol{PQ}=\boldsymbol{O}$，则(A).

(A) $t\neq 6$ 时，$\mathrm{R}(\boldsymbol{P})=1$

(B) $t\neq 6$ 时，$\mathrm{R}(\boldsymbol{P})=2$

(C) $t=6$ 时，$\mathrm{R}(\boldsymbol{P})=1$

(D) $t=6$ 时，$\mathrm{R}(\boldsymbol{P})=2$

(5) 设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶方阵，则(A).

(A) $\mathrm{R}(\boldsymbol{A},\boldsymbol{AB})=\mathrm{R}(\boldsymbol{A})$

(B) $\mathrm{R}(\boldsymbol{A},\boldsymbol{BA})=\mathrm{R}(\boldsymbol{A})$

(C) $\mathrm{R}(\boldsymbol{A},\boldsymbol{AB})=\max\big(\mathrm{R}(\boldsymbol{A}),\mathrm{R}(\boldsymbol{B})\big)$

(D) $\mathrm{R}(\boldsymbol{AB})=\mathrm{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}})$

答案

存在 $\boldsymbol{AB}=\boldsymbol{O},\boldsymbol{BA}\neq\boldsymbol{O}$，D 错误. 令 $\boldsymbol{A}=\boldsymbol{E}$，C 错误. $(\boldsymbol{E},\boldsymbol{B})$ 行满秩，选 A.

练习

(6) 设 $\boldsymbol{A} \in M_{m\times n}, \boldsymbol{B} \in M_{n\times m}$，则(A).

(A) 当 $m > n$ 时，必有 $|\boldsymbol{AB}| = 0$
(B) 当 $m > n$ 时，必有 $|\boldsymbol{AB}| \neq 0$
(C) 当 $m < n$ 时，必有 $|\boldsymbol{AB}| = 0$
(D) 当 $m < n$ 时，必有 $|\boldsymbol{AB}| \neq 0$

(7) 设 $\boldsymbol{A} \in M_{n\times m}, \boldsymbol{B} \in M_{m\times n}, n < m$. 若 $\boldsymbol{AB} = \boldsymbol{E}$，则 $\mathrm{R}(\boldsymbol{B}) =$ __n__.

(8) 若 $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 2 & 3 & a+2 \end{pmatrix}$ 和 $\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 1 \\ a+3 & a+6 & a+4 \end{pmatrix}$ 等价，则(B).

(A) $a = -1$　(B) $a \neq -1$　(C) $a \neq 1$　(D) $a = 1$

(9) 设四阶方阵 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4)$ 满足 $\boldsymbol{\alpha}_1 + \boldsymbol{\alpha}_2 - 2\boldsymbol{\alpha}_3 = \boldsymbol{0}, \boldsymbol{\alpha}_2 + 5\boldsymbol{\alpha}_4 = \boldsymbol{0}$，则 $\mathrm{R}(\boldsymbol{A}^*) =$ __0__.

定理

若 $\boldsymbol{A}$ 经过初等行变换变为 $\boldsymbol{B}$, 则

(1) $\boldsymbol{A}$ 的行向量组与 $\boldsymbol{B}$ 的行向量组等价;

(2) $\boldsymbol{A}$ 任意 k 列和 $\boldsymbol{B}$ 对应的 k 列具有相同的线性相关性.

即初等行变换保持行向量组的等价性, 列向量组的线性组合关系.

证明

(1)我们已经证明过. 设 $\boldsymbol{B}=\boldsymbol{P}\boldsymbol{A}$, 其中 $\boldsymbol{P}$ 是可逆矩阵. 若 $\boldsymbol{B}\boldsymbol{x}=\mathbf{0}$, 则 $\boldsymbol{P}\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$, $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$. 反之亦然, 即 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0} \iff \boldsymbol{B}\boldsymbol{x}=\mathbf{0}$. 设

$$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_m),\qquad \boldsymbol{B}=(\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m).$$

则对于 $\boldsymbol{x}=(\lambda_1,\dots,\lambda_m)^{\mathrm{T}}$,

$$\boldsymbol{A}\boldsymbol{x}=\lambda_1\boldsymbol{\alpha}_1+\cdots+\lambda_m\boldsymbol{\alpha}_m=\mathbf{0} \iff \boldsymbol{B}\boldsymbol{x}=\lambda_1\boldsymbol{\beta}_1+\cdots+\lambda_m\boldsymbol{\beta}_m=\mathbf{0}.$$

□

极大线性无关组和秩的计算方法

(1) 将向量组以列向量形式组成矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_m)$.

(2) 通过初等行变换将 $\boldsymbol{A}$ 变为行阶梯形矩阵.

- 行阶梯形矩阵非零行的行数就是秩 $\mathrm{R}(\boldsymbol{A})$;
- 行阶梯形矩阵每个非零行的首个非零元对应的 $\boldsymbol{A}$ 的列向量，就是极大线性无关组.

(3) 继续化简为行最简形矩阵，则可将其余向量表示为极大线性无关组的线性组合.

例

求下述向量组的秩和一个极大无关组，并把其余向量用这个极大无关组线性表示：

$$\boldsymbol{\alpha}_1 = \begin{pmatrix} -7 \\ -2 \\ 1 \\ -11 \end{pmatrix},\ \boldsymbol{\alpha}_2 = \begin{pmatrix} 1 \\ -1 \\ 5 \\ 8 \end{pmatrix},\ \boldsymbol{\alpha}_3 = \begin{pmatrix} 3 \\ 1 \\ -1 \\ 4 \end{pmatrix},\ \boldsymbol{\alpha}_4 = \begin{pmatrix} 5 \\ 3 \\ -7 \\ 0 \end{pmatrix},\ \boldsymbol{\alpha}_5 = \begin{pmatrix} -4 \\ -2 \\ 1 \\ -11 \end{pmatrix}.$$

解

$$\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3, \boldsymbol{\alpha}_4, \boldsymbol{\alpha}_5) = \begin{pmatrix} -7 & 1 & 3 & 5 & -4 \\ -2 & -1 & 1 & 3 & -2 \\ 1 & 5 & -1 & -7 & 1 \\ -11 & 8 & 4 & 0 & -11 \end{pmatrix}$$

典型例题：求极大线性无关组

续解

$$\underset{\sim}{r_1 \leftrightarrow r_3}\begin{pmatrix} 1 & 5 & -1 & -7 & 1 \\ -2 & -1 & 1 & 3 & -2 \\ -7 & 1 & 3 & 5 & -4 \\ -11 & 8 & 4 & 0 & -11 \end{pmatrix} \sim \begin{pmatrix} 1 & 5 & -1 & -7 & 1 \\ 0 & 9 & -1 & -11 & 0 \\ 0 & 36 & -4 & -44 & 3 \\ 0 & 63 & -7 & -77 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 5 & -1 & -7 & 1 \\ 0 & 9 & -1 & -11 & 0 \\ 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} \boxed{1} & 0 & -4/9 & -8/9 & 0 \\ 0 & \boxed{1} & -1/9 & -11/9 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

因此 $\mathrm{R}(\boldsymbol{A}) = 3, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_5$ 是一个极大线性无关组，且

$$\boldsymbol{\alpha}_3 = -\frac{4}{9}\boldsymbol{\alpha}_1 - \frac{1}{9}\boldsymbol{\alpha}_2, \quad \boldsymbol{\alpha}_4 = -\frac{8}{9}\boldsymbol{\alpha}_1 - \frac{11}{9}\boldsymbol{\alpha}_2.$$

练习

求下述矩阵列向量的一个极大无关组，并把其余向量用这个极大无关组线性表示：

$$\boldsymbol{A}=\begin{pmatrix}2 & -1 & -1 & 1 & 2\\ 1 & 1 & -2 & 1 & 4\\ 4 & -6 & 2 & -2 & 4\\ 3 & 6 & -9 & 7 & 9\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & -1 & 0 & 4\\ 0 & 1 & -1 & 0 & 3\\ 0 & 0 & 0 & 1 & -3\\ 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

答案

设 $\boldsymbol{\alpha}_j$ 是 $\boldsymbol{A}$ 的第 j 列，则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$ 是一个极大线性无关组，且

$$\boldsymbol{\alpha}_3=-\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2,\quad \boldsymbol{\alpha}_5=4\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2-3\boldsymbol{\alpha}_4.$$

例

假设下述向量组线性相关

$$\boldsymbol{\alpha}_1=(1,1,1,1,2),\ \boldsymbol{\alpha}_2=(2,1,3,2,3),\ \boldsymbol{\alpha}_3=(2,3,3,2,3),\ \boldsymbol{\alpha}_4=(1,3,-1,1,a).$$

求 a，并求它的秩和一个极大无关组，并把其余向量用这个极大无关组线性表示.

解

$$\boldsymbol{A}=(\boldsymbol{\alpha}_1^{\mathrm{T}},\boldsymbol{\alpha}_2^{\mathrm{T}},\boldsymbol{\alpha}_3^{\mathrm{T}},\boldsymbol{\alpha}_4^{\mathrm{T}})\overset{r}{\sim}\begin{pmatrix}1&0&0&5\\0&1&0&-2\\0&0&1&0\\0&0&0&a-4\\0&0&0&0\end{pmatrix}.$$

因此 $a=4$，秩为 3，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是一个极大线性无关组，且 $\boldsymbol{\alpha}_4=5\boldsymbol{\alpha}_1-2\boldsymbol{\alpha}_2$.

练习

(1) 设矩阵 $\boldsymbol{A}$ 经初等行变换化为 $\boldsymbol{B}$, 则二者的(A).

(A) 行向量组等价, 列向量组同相关性

(B) 行向量组同相关性, 列向量组等价

(C) 行向量组未必等价, 列向量组同相关性

(D) 行向量组等价, 列向量组未必同相关性

(2) 设 $\boldsymbol{A} \in M_{m\times n}, \boldsymbol{B} \in M_{n\times k}, \boldsymbol{AB} = \boldsymbol{O}, \boldsymbol{B} \neq \boldsymbol{O}$, 则(A).

(A) $\boldsymbol{A}$ 的列向量组线性相关

(B) $\boldsymbol{A}$ 的行向量组线性相关

(C) $\boldsymbol{A}$ 的列向量组线性无关

(D) $\boldsymbol{A}$ 的行向量组线性无关

练习

多选题：设 $\boldsymbol{A}^*$ 是 $n>1$ 阶方阵，以下说法正确的是(ABCD).

(A) 若 $\boldsymbol{A}$ 的列向量组线性相关，则 $\boldsymbol{A}^*$ 的列向量组线性相关

(B) 若 $\boldsymbol{A}$ 的列向量组线性无关，则 $\boldsymbol{A}^*$ 的列向量组线性无关

(C) 若 $\boldsymbol{A}^*$ 的某两列向量线性相关，则 $\boldsymbol{A}$ 的列向量组线性相关

(D) 若 $\boldsymbol{A}^*$ 的某两列向量线性无关，则 $\boldsymbol{A}$ 的列向量组线性无关

第三节　标准正交基

- 向量的内积
- 正交向量组与格拉姆-施密特正交化

本节考虑的向量都是实向量.

设向量组 $S=\{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_s\}$ 的秩为 r, 则它们生成的线性空间 V 的维数就是 r. S 的极大无关组 S_0 的大小就是 r, 且 S_0 是 V 的一组基.

有时候我们想更进一步, 就像 $\mathbb{R}^n$ 的基本向量组 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ 一样, 我们希望找到 V 的一组基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 使得

(1) $\boldsymbol{\alpha}_i$ 长度都是 1;

(2) $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_r$ 两两垂直.

定义

设 $\boldsymbol{\alpha}=(a_1,\dots,a_n)^{\mathrm{T}},\boldsymbol{\beta}=(b_1,\dots,b_n)^{\mathrm{T}}\in\mathbb{R}^n$, 定义内积

$$[\boldsymbol{\alpha},\boldsymbol{\beta}]=\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\beta}=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\alpha}=a_1b_1+\cdots+a_nb_n\in\mathbb{R}.$$

内积是数量积的推广, 它满足

(1) $[\boldsymbol{\alpha},\boldsymbol{\beta}]=[\boldsymbol{\beta},\boldsymbol{\alpha}]$;

(2) $[\boldsymbol{\alpha},\boldsymbol{\alpha}]\geqslant 0$. 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时, $[\boldsymbol{\alpha},\boldsymbol{\alpha}]=0$;

(3) $[\lambda\boldsymbol{\alpha},\boldsymbol{\beta}]=[\boldsymbol{\alpha},\lambda\boldsymbol{\beta}]=\lambda[\boldsymbol{\alpha},\boldsymbol{\beta}]$;

(4) $[\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma}]=[\boldsymbol{\alpha},\boldsymbol{\gamma}]+[\boldsymbol{\beta},\boldsymbol{\gamma}]$.

这说明内积是一个对称正定双线性型.

定义

设 $\boldsymbol{x}=(x_1,\ldots,x_n)\in\mathbb{R}^n$, 定义 $\boldsymbol{x}$ 的**长度**或**模**为

$$\|\boldsymbol{x}\|=\sqrt{[\boldsymbol{x},\boldsymbol{x}]}=\sqrt{x_1^2+\cdots+x_n^2}.$$

当 $\|\boldsymbol{x}\|=1$ 时, 称 $\boldsymbol{x}$ 为**单位向量**. 对于非零向量 $\boldsymbol{x}$, $\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ 为 $\boldsymbol{x}$ 的**单位化向量**.

我们有 $\boldsymbol{x}=\mathbf{0}\iff\|\boldsymbol{x}\|=0\iff[\boldsymbol{x},\boldsymbol{x}]=0.$

定义

若 $[\boldsymbol{\alpha},\boldsymbol{\beta}]=0$, 称 $\boldsymbol{\alpha},\boldsymbol{\beta}$ **正交**(垂直).

设 $\boldsymbol{\alpha} \neq \mathbf{0}$. 那么 X 的二次多项式

$$\|X\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=[X\boldsymbol{\alpha}+\boldsymbol{\beta},X\boldsymbol{\alpha}+\boldsymbol{\beta}]=[\boldsymbol{\alpha},\boldsymbol{\alpha}]X^2+2[\boldsymbol{\alpha},\boldsymbol{\beta}]X+[\boldsymbol{\beta},\boldsymbol{\beta}]\geqslant 0$$

恒成立. 因此其判别式

$$\Delta=4([\boldsymbol{\alpha},\boldsymbol{\beta}]^2-\|\boldsymbol{\alpha}\|^2\cdot\|\boldsymbol{\beta}\|^2)\leqslant 0,$$

于是我们得到**柯西-施瓦兹不等式**

$$\pm[\boldsymbol{\alpha},\boldsymbol{\beta}]\leqslant\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|.$$

显然 $\boldsymbol{\alpha}=\mathbf{0}$ 时它也成立.

定义

设非零向量 $\boldsymbol{\alpha}=(a_1,\ldots,a_n)^{\mathrm{T}},\boldsymbol{\beta}=(b_1,\ldots,b_n)^{\mathrm{T}}\in\mathbb{R}^n$, 定义 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 的**夹角**为

$$\theta=\arccos\frac{[\boldsymbol{\alpha},\boldsymbol{\beta}]}{\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|}\in[0,\pi].$$

注意正交比夹角为 $\frac{\pi}{2}$ 略微广泛点, 因为零向量与任意向量正交.

若 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 正交, 则

$$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2+2[\boldsymbol{\alpha},\boldsymbol{\beta}]=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2,$$

此即**勾股定理**.

定义

(1) 若向量组 S 中的向量两两正交且非零，则称 S 为**正交向量组**.

(2) 若向量组 S 中的向量两两正交且均为单位向量，则称 S 为**标准正交向量组**.

例

设 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,-2,1)^{\mathrm{T}}\in\mathbb{R}^3$. 求向量 $\boldsymbol{\alpha}_3$ 使得 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是正交向量组.

解

显然 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 正交. 设 $\boldsymbol{\alpha}_3=(x_1,x_2,x_3)^{\mathrm{T}}$, 则

$$
\begin{aligned}
[\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3]&=x_1+x_2+x_3=0,\\
[\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3]&=x_1-2x_2+x_3=0.
\end{aligned}
$$

解得 $(x_1,x_2,x_3)=(k,0,-k)$. 故可取 $\boldsymbol{\alpha}_3=(1,0,-1)^{\mathrm{T}}$.

定理

正交向量组必线性无关.

证明

设 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 是正交向量组, $\lambda_1\boldsymbol{\alpha}_1 + \cdots + \lambda_r\boldsymbol{\alpha}_r = \mathbf{0}$. 对任意 $1 \leqslant i \leqslant r$,

$$\begin{aligned} 0 = [\mathbf{0}, \boldsymbol{\alpha}_i] &= [\lambda_1\boldsymbol{\alpha}_1 + \cdots + \lambda_r\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_i] \\ &= \lambda_1[\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_i] + \cdots + \lambda_r[\boldsymbol{\alpha}_r, \boldsymbol{\alpha}_i] = \lambda_i[\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i]. \end{aligned}$$

由于 $\boldsymbol{\alpha}_i$ 非零, $[\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i] \neq 0, \lambda_i = 0$. 故 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 线性无关. □

如果 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 是线性空间 V 的一组标准正交基, 则对任意 $\boldsymbol{v} \in V$, 有

$$\boldsymbol{v} = \sum_{i=1}^{r} [\boldsymbol{v}, \boldsymbol{\alpha}_i]\boldsymbol{\alpha}_i.$$

定义

若实方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$, 则称 $\boldsymbol{A}$ 为正交阵.

正交阵满足如下性质:

(1) $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)$ 是正交阵 $\iff$ $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 是标准正交向量组.

(2) $\boldsymbol{A}$ 是正交阵 $\iff$ $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}^{-1}$.

(3) $\boldsymbol{A}$ 是正交阵 $\implies$ $|\boldsymbol{A}|=\pm 1$ 且 $\boldsymbol{A}^{\mathrm{T}},\boldsymbol{A}^{-1},\boldsymbol{A}^*$ 都是正交阵.

定义

若 $\boldsymbol{P}$ 为正交阵, 称线性变换 $\boldsymbol{y}=\boldsymbol{P}\boldsymbol{x}$ 为正交变换.

例如, $\mathbb{R}^2$ 上的正交变换就是绕原点的旋转、反射, 以及它们的复合. 由于

$$[\boldsymbol{P}\boldsymbol{x},\boldsymbol{P}\boldsymbol{y}]=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{y}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}=[\boldsymbol{x},\boldsymbol{y}],$$

因此正交变换保持向量的长度和夹角.

现在我们来看如何从空间 V 的一组基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_r$ 得到一组标准正交基. 令 $\boldsymbol{\beta}_1 = \boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_2 = \boldsymbol{\alpha}_2 + \lambda\boldsymbol{\beta}_1$. 若 $\boldsymbol{\beta}_1$ 和 $\boldsymbol{\beta}_2$ 正交, 则

$$0 = [\boldsymbol{\beta}_2, \boldsymbol{\beta}_1] = [\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1] + \lambda[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1],$$

因此 $\lambda = -\dfrac{[\boldsymbol{\alpha}_2, \boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1]}$.

令 $\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_3 + \lambda_1\boldsymbol{\beta}_1 + \lambda_2\boldsymbol{\beta}_2$, 类似地, 若 $\boldsymbol{\beta}_3$ 和 $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ 均正交, 则

$$\lambda_1 = -\frac{[\boldsymbol{\alpha}_3, \boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1]}, \qquad \lambda_2 = -\frac{[\boldsymbol{\alpha}_3, \boldsymbol{\beta}_2]}{[\boldsymbol{\beta}_2, \boldsymbol{\beta}_2]},$$

$$\boldsymbol{\beta}_3 = \boldsymbol{\alpha}_3 - \frac{[\boldsymbol{\alpha}_3, \boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1, \boldsymbol{\beta}_1]}\boldsymbol{\beta}_1 - \frac{[\boldsymbol{\alpha}_3, \boldsymbol{\beta}_2]}{[\boldsymbol{\beta}_2, \boldsymbol{\beta}_2]}\boldsymbol{\beta}_2.$$

依次递推下去可得一组正交基.

格拉姆-施密特正交单位化方法

$$\beta_1 = \alpha_1$$

$$\beta_2 = \alpha_2 - \frac{[\alpha_2, \beta_1]}{[\beta_1, \beta_1]}\beta_1$$

$$\beta_3 = \alpha_3 - \frac{[\alpha_3, \beta_1]}{[\beta_1, \beta_1]}\beta_1 - \frac{[\alpha_3, \beta_2]}{[\beta_2, \beta_2]}\beta_2$$

$$\vdots$$

$$\beta_r = \alpha_r - \frac{[\alpha_r, \beta_1]}{[\beta_1, \beta_1]}\beta_1 - \cdots - \frac{[\alpha_r, \beta_{r-1}]}{[\beta_{r-1}, \beta_{r-1}]}\beta_{r-1}$$

则 $e_1 = \dfrac{\beta_1}{\|\beta_1\|}, \ldots, e_r = \dfrac{\beta_r}{\|\beta_r\|}$ 就是 V 的一组标准正交基.

典型例题：格拉姆-施密特正交化

例

将 $\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,0,1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,1,2)^{\mathrm{T}}$ 正交单位化.

解

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1=(1,1,0)^{\mathrm{T}}$$

$$\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{[\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1=(1,0,1)^{\mathrm{T}}-\frac{1}{2}(1,1,0)^{\mathrm{T}}=(\frac{1}{2},-\frac{1}{2},1)^{\mathrm{T}}$$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{[\boldsymbol{\alpha}_3,\boldsymbol{\beta}_1]}{[\boldsymbol{\beta}_1,\boldsymbol{\beta}_1]}\boldsymbol{\beta}_1-\frac{[\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2]}{[\boldsymbol{\beta}_2,\boldsymbol{\beta}_2]}\boldsymbol{\beta}_2=(1,1,2)^{\mathrm{T}}-(1,1,0)^{\mathrm{T}}-\frac{2}{3/2}(\frac{1}{2},-\frac{1}{2},1)^{\mathrm{T}}=(-\frac{2}{3},\frac{2}{3},\frac{2}{3})^{\mathrm{T}}$$

$$\boldsymbol{e}_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|}=\frac{1}{\sqrt{2}}(1,1,0)^{\mathrm{T}},\quad \boldsymbol{e}_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|}=\frac{1}{\sqrt{6}}(1,-1,2)^{\mathrm{T}},\quad \boldsymbol{e}_3=\frac{\boldsymbol{\beta}_3}{\|\boldsymbol{\beta}_3\|}=\frac{1}{\sqrt{3}}(-1,1,1)^{\mathrm{T}}.$$

第四节　线性方程组

- 齐次线性方程组解的存在性
- 齐次线性方程组解的结构
- 非齐次线性方程组
- 向量组的线性表示

线性方程组是指

$$\begin{cases} a_{11}\ x_1 + a_{12}\ x_2 + \cdots + a_{1n}\ x_n = b_1 \\ a_{21}\ x_1 + a_{22}\ x_2 + \cdots + a_{2n}\ x_n = b_2 \\ \qquad\qquad\qquad\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

它的系数形成了一个 $m \times n$ 矩阵 $\boldsymbol{A}$, 称为系数矩阵.

线性方程组等价于

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

其中

$$\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{T}}.$$

齐次线性方程组非零解的判定

当 $\boldsymbol{b}=\mathbf{0}$ 为零向量时，称该线性方程组为齐次的；否则称为非齐次的．齐次线性方程组总有解 $\boldsymbol{x}=\mathbf{0}$．$\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 有非零解 $\iff$ $\boldsymbol{A}$ 的列向量线性相关 $\iff$ $\mathrm{R}(\boldsymbol{A})<n$．

定理

(1) $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\mathbf{0}$ 有 (无穷多) 非零解 $\iff$ $\mathrm{R}(\boldsymbol{A})<n$；

(2) $\boldsymbol{A}_{m\times n}\boldsymbol{x}=\mathbf{0}$ 只有零解 $\iff$ $\mathrm{R}(\boldsymbol{A})=n$．

推论

设 $\boldsymbol{A}$ 是 n 阶方阵．

(1) $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 有 (无穷多) 非零解 $\iff$ $|\boldsymbol{A}|=0$；

(2) $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 只有零解 $\iff$ $|\boldsymbol{A}|\neq 0$．

推论

若方程个数小于未知元个数，则齐次线性方程组有非零解．

例：齐次线性方程组非零解的判定

例

假设

$$\begin{cases} x_1 + 2x_2 - 2x_3 = 0 \\ 4x_1 + ax_2 + 3x_3 = 0 \\ 3x_1 - x_2 + x_3 = 0 \end{cases}$$

有非零解，求 a.

解

此时系数矩阵行列式为零：

$$0 = \begin{vmatrix} 1 & 2 & -2 \\ 4 & a & 3 \\ 3 & -1 & 1 \end{vmatrix} = 7a + 21, \quad a = -3.$$

例

若下述方程有非零解，求 a.

$$\begin{cases} x_1 + x_2 + ax_3 = 0 \\ -x_1 + (a-1)x_2 + (1-a)x_3 = 0 \\ x_1 + x_2 + a^2x_3 = 0 \\ x_1 + x_2 + (2a+1)x_3 = 0 \end{cases}$$

解

$$\begin{pmatrix} 1 & 1 & a \\ -1 & a-1 & 1-a \\ 1 & 1 & a^2 \\ 1 & 1 & 2a+1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 1 & a \\ 0 & a & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

的秩小于 3，因此 $a = 0$.

定义

称空间 $\{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$ 的一组基为该齐次线性方程组的**基础解系**.

齐次线性方程组的解

设 $\boldsymbol{A} \in M_{m\times n}, \mathrm{R}(\boldsymbol{A}) = r$. 线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 的基础解系包含 $n - r$ 个向量.

证明

通过交换未知元的位置 (相当于交换 $\boldsymbol{A}$ 列的位置), 不妨设 $\boldsymbol{A}$ 可化为行最简形

$$\left(\begin{array}{ccc|ccc} 1 & \cdots & 0 & b_{11} & \cdots & b_{1,n-r} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & b_{r,1} & \cdots & b_{r,n-r} \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right) = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}.$$

续证

方程化为 $(\boldsymbol{E}_r, \boldsymbol{B})\boldsymbol{x} = \boldsymbol{0}$, 即

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = -\boldsymbol{B}\begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} -\boldsymbol{B} \\ \boldsymbol{E}_{n-r} \end{pmatrix}\begin{pmatrix} x_{r+1} \\ \vdots \\ x_n \end{pmatrix}.$$

于是 $\boldsymbol{C} := \begin{pmatrix} -\boldsymbol{B} \\ \boldsymbol{E}_{n-r} \end{pmatrix}$ 的 $n-r$ 个列向量生成了整个解空间. 由于 $\mathrm{R}(\boldsymbol{C}) \geqslant \mathrm{R}(\boldsymbol{E}_{n-r}) = n-r$, $\boldsymbol{C}$ 列满秩, 因此它的列向量就是一组基础解系. □

推论

$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$ 任意 $n-r$ 个线性无关的解都是一组基础解系.

求基础解系的步骤

齐次线性方程组的解法

(1) 将系数矩阵通过初等行变换化为行最简形.

(2) 将矩阵重新写成方程形式 $x_i + \cdots = 0$.

(3) 移项, 使得等式左侧只有阶梯拐角列 i 对应的 $x_i = \cdots$.

(4) 添加 $n-r$ 项 $x_j = x_j$, 使得等式左边凑成 $\boldsymbol{x}$.

(5) 等式右侧是非拐角列 j 对应的 x_j 的组合, 其系数形成的 $n-r$ 个向量就是基础解系.

典型例题：求基础解系

例

解方程 $\begin{cases} x_1+2x_2+4x_3+x_4=0 \\ 2x_1+4x_2-2x_3-x_4=0 \\ 3x_1+6x_2+2x_3 \qquad =0 \end{cases}$.

解

$$\begin{pmatrix} 1 & 2 & 4 & 1 \\ 2 & 4 & -2 & -1 \\ 3 & 6 & 2 & 0 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 2 & 4 & 1 \\ 0 & 0 & -10 & 3 \\ 0 & 0 & -10 & 3 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 2 & 0 & -1/5 \\ 0 & 0 & 1 & 3/10 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\Longrightarrow \begin{cases} x_1+2x_2-1/5x_4=0 \\ x_3+3/10x_4=0 \end{cases}$$

续解

将 x_1, x_3 保留在等式左侧，其它项移动到等式右边，并添加 $x_2 = x_2, x_4 = x_4$，得到

$$\begin{cases} x_1 = -2x_2 + 1/5x_4 \\ x_2 = x_2 \\ x_3 = -3/10x_4 \\ x_4 = x_4 \end{cases} \Longrightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_2 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1/5 \\ 0 \\ -3/10 \\ 1 \end{pmatrix},$$

通解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = k_1 \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 1/5 \\ 0 \\ -3/10 \\ 1 \end{pmatrix}, \quad k_1, k_2 \text{为任意常数.}$$

练习

解方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，其中 $\boldsymbol{A}=\begin{pmatrix}1&2&1&1&1\\2&4&3&1&1\\-1&-2&1&3&-3\\0&0&2&5&-2\end{pmatrix}$.

答案

$$\boldsymbol{A}\overset{r}{\sim}\begin{pmatrix}1&2&0&0&2\\0&0&1&0&-1\\0&0&0&1&0\\0&0&0&0&0\end{pmatrix}\Longrightarrow\begin{cases}x_1=-2x_2-2x_5\\x_2=x_2\\x_3=x_5\\x_4=x_4\\x_5=x_5\end{cases}\Longrightarrow\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}=k_1\begin{pmatrix}-2\\1\\0\\0\\0\end{pmatrix}+k_2\begin{pmatrix}-2\\0\\1\\0\\1\end{pmatrix}.$$

例

设 $\boldsymbol{A} \in M_{m\times n}, \mathrm{R}(\boldsymbol{A}) = n-3$, ξ_1, ξ_2, ξ_3 为 $\boldsymbol{Ax} = \boldsymbol{0}$ 的三个线性无关的解. 则(B)是该方程的一组基础解系.

(A) $\xi_1, -\xi_2, \xi_2-\xi_3, \xi_3-\xi_1$

(B) $\xi_1, \xi_1+\xi_2, \xi_1+\xi_2+\xi_3$

(C) ξ_1, ξ_2

(D) $\xi_1, \xi_1-\xi_2-\xi_3, \xi_1+\xi_2+\xi_3$

练习

设 ξ_1, ξ_2, ξ_3 是 $\boldsymbol{Ax} = \boldsymbol{0}$ 的一组基础解系，则(D)也是该方程的一组基础解系.

(A) 与 ξ_1, ξ_2, ξ_3 等价的一组向量

(B) 与 ξ_1, ξ_2, ξ_3 同秩的一组向量

(C) $\xi_1-\xi_2, \xi_2-\xi_3, \xi_3-\xi_1$

(D) $\xi_1+\xi_2, \xi_2+\xi_3, \xi_3+\xi_1$

例：基础解系的应用

例

设 $\boldsymbol{A}$ 是 n 阶方阵, $\mathrm{R}(\boldsymbol{A})=n-1$ 且每行元素之和为 0. 则齐次线性方程组 $\boldsymbol{Ax}=\mathbf{0}$ 的解为 $\underline{\quad k(1,1,\ldots,1)^{\mathrm{T}}, k \text{ 为任意常数} \quad}$.

例

设 n 阶方阵 $\boldsymbol{A}$ 列向量的一个极大线性无关组为 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_{n-1}$. 则 $\boldsymbol{A}^*\boldsymbol{x}=\mathbf{0}$ 的解为 $\underline{\quad k_1\boldsymbol{\alpha}_1+\cdots+k_{n-1}\boldsymbol{\alpha}_{n-1}, k_1,\ldots,k_{n-1} \text{ 为任意常数} \quad}$.

练习

设 n 阶方阵 $\boldsymbol{A}$ 满足 $\mathrm{R}(\boldsymbol{A})=n-1$, 代数余子式 $A_{11}\neq 0$. 则 $\boldsymbol{Ax}=\mathbf{0}$ 的解为 $\underline{\quad k(A_{11},\ldots,A_{1n})^{\mathrm{T}}, k \text{ 为任意常数} \quad}$.

例

设 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times s}=\boldsymbol{O}$，证明 $\mathrm{R}(\boldsymbol{A})+\mathrm{R}(\boldsymbol{B})\leqslant n$.

证明

由于 $\boldsymbol{B}$ 的列向量都是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解，因此 $\mathrm{R}(\boldsymbol{B})$ 不超过该方程解空间的维数，即 $n-\mathrm{R}(\boldsymbol{A})$. □

例

设 $\boldsymbol{A}$ 是实矩阵，证明 $\mathrm{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})=\mathrm{R}(\boldsymbol{A})$.

证明

若 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$，则 $0=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=[\boldsymbol{A}\boldsymbol{x},\boldsymbol{A}\boldsymbol{x}]$. 由内积的正定性可知 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$. 所以 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}\iff\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$. 二者解空间维数相等. 由于二者列数相同，因此二者秩相同. □

由此可知

$$\mathrm{R}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}) = \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}) = \mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}),$$

$$\mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) = \cdots$$

由于

$$\mathrm{R}(\boldsymbol{A}) \geqslant \mathrm{R}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}) \geqslant \mathrm{R}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}),$$

因此这些矩阵的秩都相等. 类似地, 任意多个 $\boldsymbol{A}, \boldsymbol{A}^{\mathrm{T}}$ 交错相乘得到的矩阵秩也都等于 $\mathrm{R}(\boldsymbol{A})$.

注意, 对于复矩阵这并不成立, 例如

$$\boldsymbol{A} = \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{A} = 0.$$

此时有 $\mathrm{R}\left(\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{A}\right) = \mathrm{R}(\boldsymbol{A})$, 其中 $\overline{\boldsymbol{A}}$ 表示所有元素取共轭.

例

若 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5)$ 且 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的解为 $k_1(1,0,-1,0,1)^{\mathrm{T}}+k_2(1,0,0,1,-1)^{\mathrm{T}}$，则 $\boldsymbol{A}$ 列向量组的一个极大无关组是(D).

(A) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5$ (B) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ (C) $\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5$ (D) $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_5$

练习

若 $\boldsymbol{A}=\begin{pmatrix} a & 1 & a^2 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$ 且存在 3 阶非零矩阵 $\boldsymbol{B}$ 使得 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{O}$，则(A).

(A) $a=1,|\boldsymbol{B}|=0$ (B) $a=-2,|\boldsymbol{B}|=0$

(C) $a=1,|\boldsymbol{B}|\neq 0$ (D) $a=-2,|\boldsymbol{B}|\neq 0$

设 $\boldsymbol{A}\in M_{m\times n}$. 对于非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, 若方程有解, 则 $\boldsymbol{b}$ 可以由 $\boldsymbol{A}$ 的列向量线性表示, 从而 $\boldsymbol{A}$ 的列向量组和 $(\boldsymbol{A},\boldsymbol{b})$ 的列向量组等价. 因此 $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})$. 我们称 $m\times(n+1)$ 矩阵 $(\boldsymbol{A},\boldsymbol{b})$ 为**增广矩阵**.

注意到 $\boldsymbol{A}$ 列向量生成的空间 V 是 $(\boldsymbol{A},\boldsymbol{b})$ 列向量生成的空间 W 的子空间. 若 $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})$, 则 $V=W$, $\boldsymbol{A}$ 列向量组的一个极大无关组 S 也是 $(\boldsymbol{A},\boldsymbol{b})$ 的极大无关组. 从而 $\boldsymbol{b}$ 是 S 的线性组合, 也是 $\boldsymbol{A}$ 列向量的线性组合.

定理

$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有解 $\iff$ $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})$.

推论

若 $\mathrm{R}(\boldsymbol{A}_{m\times n})=m$, 则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 总有解.

非齐次线性方程组解的结构

若非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有解 $\boldsymbol{x}=\boldsymbol{x}_0$，则 $\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}_0)=\mathbf{0}$. 从而 $\boldsymbol{x}-\boldsymbol{x}_0$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解. 设 $\xi_1,\ldots,\xi_{n-r}$ 为 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的一组基础解系，则 $\boldsymbol{A}\boldsymbol{x}=\mathbf{0}$ 的通解为

$$\boldsymbol{x}=\boldsymbol{x}_0+k_1\boldsymbol{\xi}_1+\cdots+k_{n-r}\boldsymbol{\xi}_{n-r},$$

$k_1,\ldots,k_{n-r}$ 为任意常数.

线性方程组解的判定准则

(1) 若 $\mathrm{R}(\boldsymbol{A})<\mathrm{R}(\boldsymbol{A},\boldsymbol{b})$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 无解;

(2) 若 $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})=n$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有唯一解;

(3) 若 $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})<n$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有无穷多解.

推论

若 $\boldsymbol{A}$ 是 n 阶方阵，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有唯一解 $\Longleftrightarrow$ $|\boldsymbol{A}|\neq 0$.

若 $|\boldsymbol{A}|=0$，则 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 无解或有无穷多解.

求解非齐次线性方程组的步骤

非齐次线性方程组的解法

(1) 写：写出方程组对应的增广矩阵 $(\boldsymbol{A}, \boldsymbol{b})$;

(2) 变：通过初等行变换将其化为行最简形;

(3) 判：通过行最简形判定方程是否有解;

(4) 解：若系数矩阵部分零行对应的常数项均为零，则方程有解.

(5) 类似于齐次情形，将矩阵重新写成方程形式、移项、添恒等式，使得等式左边凑成 $\boldsymbol{x}$.

(6) 等式右侧的常数部分是特解，其余是非拐角列 j 对应的 x_j 的组合，其系数形成的 $n-r$ 个向量就是基础解系.

例

解方程 $\begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 1 \\ 2x_1 - x_2 + 2x_3 - 2x_4 = 3 \\ 3x_1 + x_2 + 5x_3 + 2x_4 = 2 \end{cases}$

解

$$\left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 2 & -1 & 2 & -2 & 3 \\ 3 & 1 & 5 & 2 & 2 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 0 & -5 & -4 & -10 & 1 \\ 0 & -5 & -4 & -10 & -1 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 1 \\ 0 & -5 & -4 & -10 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

于是 $\mathrm{R}(\boldsymbol{A}) = 2 < \mathrm{R}(\boldsymbol{A}, \boldsymbol{b}) = 3$，无解.

典型例题：解非齐次线性方程组

例

解方程 $\begin{cases} x_1 - x_2 - x_3 + x_4 = 0 \\ x_1 - x_2 + x_3 - 3x_4 = 1 \\ x_1 - x_2 - 2x_3 + 3x_4 = -1/2 \end{cases}$

解

$$\left(\begin{array}{cccc:c} 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & 1 & -3 & 1 \\ 1 & -1 & -2 & 3 & -1/2 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{cccc:c} 1 & -1 & 0 & -1 & 1/2 \\ 0 & 0 & 1 & -2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

于是 $\mathrm{R}(\boldsymbol{A}) = 2 = \mathrm{R}(\boldsymbol{A}, \boldsymbol{b}) = 2$，有解.

$$\begin{cases} x_1 = x_2 + x_4 + 1/2 \\ x_2 = x_2 \\ x_3 = 2x_4 + 1/2 \\ x_4 = x_4 \end{cases} \Longrightarrow \boldsymbol{x} = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix} + k_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, k_1, k_2 \in \mathbb{R}.$$

例

已知

$$\boldsymbol{\alpha}_1 = (1,4,0,2)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (2,7,1,3)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (0,1,-1,a)^{\mathrm{T}}, \boldsymbol{\beta} = (3,10,b,4)^{\mathrm{T}}.$$

问 a,b 为何值时，

(1) $\boldsymbol{\beta}$ 不能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表示；

(2) $\boldsymbol{\beta}$ 能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 唯一线性表示；

(3) $\boldsymbol{\beta}$ 能由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 不唯一线性表示.

典型例题：线性方程组的性质

解

即问 $\boldsymbol{A}\boldsymbol{x} = \beta$ 的解的情况，其中 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$.

$$(\boldsymbol{A}, \beta) = \left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 4 & 7 & 1 & 10 \\ 0 & 1 & -1 & b \\ 2 & 3 & a & 4 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & -1 & 1 & -2 \\ 0 & 1 & -1 & b \\ 0 & -1 & a & -2 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & a-1 & 0 \\ 0 & 0 & 0 & b-2 \end{array}\right)$$

于是可知 $\mathrm{R}(\boldsymbol{A})$ 和 $\mathrm{R}(\boldsymbol{A}, \beta)$，故

(1) $b \neq 2$ 时，β 不能由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性表示;

(2) $a \neq 1, b = 2$ 时，β 能由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 唯一线性表示;

(3) $a = 1, b = 2$ 时，β 能由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 不唯一线性表示.

例：线性方程组解的性质

例

设 $\boldsymbol{A} \in M_{m\times n}$, 则(D).

(A) 若 $\boldsymbol{Ax}=\boldsymbol{0}$ 仅有零解, 则 $\boldsymbol{Ax}=\boldsymbol{b}$ 有唯一解

(B) 若 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解, 则 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解

(C) 若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解, 则 $\boldsymbol{Ax}=\boldsymbol{0}$ 只有零解

(D) 若 $\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解, 则 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解

练习

设 $\boldsymbol{A} \in M_{m\times n}, \mathrm{R}(\boldsymbol{A})=m<n$, 则(C).

(A) $\boldsymbol{A}$ 的任意 m 个列向量线性无关

(B) $\boldsymbol{A}$ 的任意一个 m 阶子式不等于 0

(C) $\boldsymbol{Ax}=\boldsymbol{b}$ 一定有无穷多个解

(D) $\boldsymbol{A} \overset{r}{\sim} (\boldsymbol{E}, \boldsymbol{O})$

例

a 为何值时，以下方程 (1) 有唯一解；(2) 无解；(3) 有无穷多解？并在有无穷多解时求其通解.

$$\begin{cases} (1+a)x_1 + x_2 + x_3 = 0 \\ x_1 + (1+a)x_2 + x_3 = 3 \\ x_1 + x_2 + (1+a)x_3 = a \end{cases}$$

和之前求带参数矩阵的秩类似，此处不宜实施常数不确定是否非零的第二类初等变换 $\frac{1}{a+1}r_2, (a-2)r_3$ 等.

解

$$(\boldsymbol{A}, \boldsymbol{b}) = \left(\begin{array}{ccc|c} 1+a & 1 & 1 & 0 \\ 1 & 1+a & 1 & 3 \\ 1 & 1 & 1+a & a \end{array}\right) \overset{r}{\sim} \left(\begin{array}{ccc|c} 1 & 1+a & 1 & 3 \\ 0 & -a & a & a-3 \\ 0 & a & a^2+2a & a^2+a \end{array}\right)$$

续解

$$\overset{r}{\sim}\left(\begin{array}{ccc:c}1 & 1+a & 1 & 3\\ 0 & a & -a & 3-a\\ 0 & 0 & a^2+3a & a^2+2a-3\end{array}\right).$$

(1) 若 $a \neq 0, -3$, 则 $\mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{A}, \boldsymbol{b}) = 3$, 方程有唯一解.

(2) 若 $a = 0$, 则 $(\boldsymbol{A}, \boldsymbol{b}) \overset{r}{\sim} \left(\begin{array}{ccc:c}1 & 1 & 1 & 3\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{array}\right)$, $\mathrm{R}(\boldsymbol{A}) = 1 < \mathrm{R}(\boldsymbol{A}, \boldsymbol{b}) = 2$, 方程无解.

续解

(3) 若 $a=-3$, 则 $(\boldsymbol{A},\boldsymbol{b}) \overset{r}{\sim} \left(\begin{array}{ccc|c} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right)$, $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})=2$, 方程有无穷多解. 特解为 $\begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}$, 基础解系为 $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, 通解为

$$\boldsymbol{x}=\begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}+k\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

k 为任意常数.

由于系数矩阵为 3 阶方阵, 也可以先通过 $|\boldsymbol{A}| \neq 0$ 得到唯一解情形.

练习

a, b 为何值时，以下方程

(1) 有唯一解; (2) 无解; (3) 有无穷多解? 并在有无穷多解时求其通解.

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ x_2 - x_3 + 2x_4 = 1 \\ 2x_1 + 3x_2 + (a+2)x_3 + 4x_4 = b+3 \\ 3x_1 + 5x_2 + x_3 + (a+8)x_4 = 5 \end{cases}$$

典型例题：解非齐次线性方程组

答案

$$\left(\begin{array}{cccc:c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 2 & 1 \\ 2 & 3 & a+2 & 4 & b+3 \\ 3 & 5 & 1 & a+8 & 5 \end{array}\right) \overset{r}{\sim} \left(\begin{array}{cccc:c} 1 & 0 & 2 & -1 & 0 \\ 0 & 1 & -1 & 2 & 1 \\ 0 & 0 & a+1 & 0 & b \\ 0 & 0 & 0 & a+1 & 0 \end{array}\right).$$

(1) $a \neq -1$ 时有唯一解;

(2) $a = -1, b \neq 0$ 时无解;

(3) $a = -1, b = 0$ 时有无穷多解，通解为

$$\boldsymbol{x} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + k_1 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} 1 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

例

设四元非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的系数矩阵 $\boldsymbol{A}$ 的秩为 3. 已知 $\boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\boldsymbol{\eta}_3$ 是它的三个解向量，且

$$\boldsymbol{\eta}_1=(2,3,4,5)^{\mathrm{T}},\quad \boldsymbol{\eta}_2+\boldsymbol{\eta}_3=(1,2,3,4)^{\mathrm{T}}.$$

求 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的通解.

解

由于 $\mathrm{R}(\boldsymbol{A})=3$，因此 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的基础解系只包含一个向量. 根据解的性质，

$$2\boldsymbol{\eta}_1-(\boldsymbol{\eta}_2+\boldsymbol{\eta}_3)=(3,4,5,6)^{\mathrm{T}}$$

是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的一个解，因此这是它的一个基础解系. 故 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的通解为

$$\boldsymbol{x}=\boldsymbol{\eta}_1+k(3,4,5,6)^{\mathrm{T}}=(2,3,4,5)^{\mathrm{T}}+k(3,4,5,6)^{\mathrm{T}}.$$

例

已知 4 阶方阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4)$，且 $\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 线性无关，$\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$. 若 $\boldsymbol{\beta}=\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2+\boldsymbol{\alpha}_3+\boldsymbol{\alpha}_4$，求 $\boldsymbol{Ax}=\boldsymbol{\beta}$ 的通解.

解

由题设可知 $\mathrm{R}(\boldsymbol{A})=3$，因此 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系只包含一个向量. 由 $\boldsymbol{\alpha}_1=2\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$ 可知 $(1,-2,1,0)^{\mathrm{T}}$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的一个解，因此这是它的一个基础解系. 注意到 $(1,1,1,1)^{\mathrm{T}}$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的一个特解，故通解为

$$\boldsymbol{x}=(1,1,1,1)^{\mathrm{T}}+k(1,-2,1,0)^{\mathrm{T}}.$$

例：线性方程组解的性质

例

已知 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{b}$ 的两个不同的解，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 是 $\boldsymbol{Ax}=\boldsymbol{0}$ 的基础解系，则 $\boldsymbol{Ax}=\boldsymbol{b}$ 的通解为（ B ），k_1,k_2 为任意常数.

(A) $\dfrac{\boldsymbol{\beta}_1-\boldsymbol{\beta}_2}{2}+k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2)$

(B) $2\boldsymbol{\beta}_1-\boldsymbol{\beta}_2+k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2)$

(C) $\dfrac{\boldsymbol{\beta}_1+\boldsymbol{\beta}_2}{2}+k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)$

(D) $\dfrac{\boldsymbol{\beta}_1-\boldsymbol{\beta}_2}{2}+k_1\boldsymbol{\alpha}_1+k_2(\boldsymbol{\beta}_1-\boldsymbol{\beta}_2)$

练习

已知 $\boldsymbol{\eta}_1=(0,1,0)^{\mathrm{T}},\boldsymbol{\eta}_2=(-3,2,2)^{\mathrm{T}}$ 是线性方程组 $\begin{cases} x_1-x_2+2x_3=-1\\ 3x_1+x_2+4x_3=1\\ ax_1+bx_2+cx_3=d\end{cases}$ 的两个解向量，则该方程组的通解为 $\underline{\quad(0,1,0)^{\mathrm{T}}+k(-3,1,2)^{\mathrm{T}}\quad}$.

若 $\boldsymbol{B}$ 的列向量可由 $\boldsymbol{A}$ 的列向量组线性表示, 则 $(\boldsymbol{A},\boldsymbol{B})$ 的列向量组和 $\boldsymbol{A}$ 的列向量组等价, 因此 $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{B})$.

注意到 $\boldsymbol{A}$ 列向量生成的空间 V 是 $(\boldsymbol{A},\boldsymbol{B})$ 列向量生成的空间 W 的子空间. 若 $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{b})$, 则 $\boldsymbol{A}$ 列向量组的一个极大无关组 S 也是 $(\boldsymbol{A},\boldsymbol{B})$ 的极大无关组. 从而 $\boldsymbol{B}$ 的列向量都是 S 的线性组合, 也是 $\boldsymbol{A}$ 列向量的线性组合.

定理

(1) $\boldsymbol{B}$ 的列向量组可由 $\boldsymbol{A}$ 的列向量组线性表示 $\iff$ $\boldsymbol{AX}=\boldsymbol{B}$ 有解 $\iff$ $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{B})$.

(2) $\boldsymbol{B}$ 的列向量组和 $\boldsymbol{A}$ 的列向量组等价 $\iff$ $\mathrm{R}(\boldsymbol{A})=\mathrm{R}(\boldsymbol{A},\boldsymbol{B})=\mathrm{R}(\boldsymbol{B})$.

例

证明向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 与 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3$ 等价，其中

$$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\-1\\1\\-1\end{pmatrix},\ \boldsymbol{\alpha}_2=\begin{pmatrix}3\\1\\1\\3\end{pmatrix},\ \boldsymbol{\beta}_1=\begin{pmatrix}2\\0\\1\\1\end{pmatrix},\ \boldsymbol{\beta}_2=\begin{pmatrix}1\\1\\0\\2\end{pmatrix},\ \boldsymbol{\beta}_3=\begin{pmatrix}3\\-1\\2\\0\end{pmatrix}.$$

证明

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=\left(\begin{array}{cc:ccc}1&3&2&1&3\\-1&1&-&1&-1\\1&1&1&0&2\\-1&3&1&2&0\end{array}\right)\overset{r}{\sim}\left(\begin{array}{cc:ccc}1&3&2&1&3\\0&2&1&1&1\\0&0&0&0&0\\0&0&0&0&0\end{array}\right).$$

因此 $\mathrm{R}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=\mathrm{R}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2)=\mathrm{R}(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_3)=2.$ □

在物理实验中，经常会出现实验数据与预期不符的情况. 例如变量 y 应当为变量 $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{T}}$ 的线性组合，即存在 n 维向量 β 使得 $y=\boldsymbol{x}^{\mathrm{T}}\beta$. 但从实验数据解方程却是无解. 因此我们需要寻找参数 β 使得 $y=\boldsymbol{x}^{\mathrm{T}}\beta$ 尽可能接近实验数据.

比较常见的是最小二乘法：即寻找参数 β 使得

$$\sum_{i=1}^{k}|y_i-\boldsymbol{x}_i^{\mathrm{T}}\beta|^2=\|\boldsymbol{y}-\boldsymbol{A}\beta\|^2$$

尽可能小，其中 $(\boldsymbol{x}_i,y_i)$ 是实验数据，$\boldsymbol{y}=(y_1,\dots,y_k)^{\mathrm{T}}$，$\boldsymbol{A}$ 是由行向量 $\boldsymbol{x}_i^{\mathrm{T}}$ 构成的 $k\times n$ 矩阵. 注意所有向量 $\boldsymbol{A}\beta$ 形成一个向量空间 V，也就是 $\boldsymbol{A}$ 的列向量生成的空间. $\boldsymbol{y}$ 距离这个空间的距离 $\|\boldsymbol{y}-\boldsymbol{A}\beta\|$ 达到最小时，$\boldsymbol{y}-\boldsymbol{A}\beta$ 应当和这个空间正交. 于是 $\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{A}\beta)=\boldsymbol{0}$，即 β 是方程

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\beta=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}$$

的解.

第三章　相似和合同

第一节　方阵的相似

- 特征值与特征向量
- 特征值和特征向量的性质
- 方阵的相似
- 相似矩阵的性质
- 相似对角化

设 $f:\mathbb{C}^n\to\mathbb{C}^n$ 是一个线性映射. 由于 $\mathbb{C}^n$ 中的向量由一组基

$$\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$$

唯一线性表示, 因此它完全由它在该组基下的像决定. 若 f 将每个 $\boldsymbol{\alpha}_i$ 映射为它的倍数, 则 f 将会变得很容易研究.

定义

若常数 λ 和非零向量 x 满足 $Ax=\lambda x$, 称 λ 为 A 的特征值, x 为 A 关于 λ 的特征向量.

设 $\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$, 则

$$(\boldsymbol{A} - \lambda \boldsymbol{E})\boldsymbol{x} = \boldsymbol{0}.$$

该方程有非零解当且仅当

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

注意到该行列式是 λ 的 n 次多项式, 最高项为 $(-1)^n \lambda^n$. 称之为 $\boldsymbol{A}$ 的特征多项式.

在复数域中 n 次多项式总有 n 个根 (计算重数), 也就是说 $\boldsymbol{A}$ 的特征多项式可以写成

$$f(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda).$$

所以 $\boldsymbol{A}$ 的**特征值有 n 个 (计算重数)**.

特征值和特征向量只对方阵存在, 且有如下性质:

(1) 零向量不是特征向量;

(2) 若 $\boldsymbol{x}$ 是对应 λ 的特征向量, 则它的非零倍也是;

(3) 若 $\boldsymbol{x}_1, \boldsymbol{x}_2 \neq -\boldsymbol{x}_1$ 是对应 λ 的特征向量, 则 $\boldsymbol{x}_1 + \boldsymbol{x}_2$ 也是;

(4) $|\boldsymbol{A}| = 0 \iff 0$ 是特征值; $|\boldsymbol{A}| \neq 0 \iff 0$ 不是特征值;

(5) 若 n 阶方阵 $\boldsymbol{A}$ 的各行元素之和为 k, 则 k 是 $\boldsymbol{A}$ 的一个特征值, 且特征向量为 $(1, \dots, 1)^{\mathrm{T}}$.

特征值和特征向量的计算步骤如下:

(1) 求 $\boldsymbol{A}$ 的特征多项式 $f(\lambda)=|\boldsymbol{A}-\lambda\boldsymbol{E}|$;

(2) 解 $f(\lambda)=|\boldsymbol{A}-\lambda\boldsymbol{E}|=0$ 得到特征值;

(3) 对于每一个特征值 λ_i, 解 $(\boldsymbol{A}-\lambda_i\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$, 其非零解就是对应特征向量.

例

求 $\boldsymbol{A}=\begin{pmatrix}1 & 3\\ 4 & 2\end{pmatrix}$ 的特征值和特征向量.

解

(1) 特征多项式 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix}1-\lambda & 3\\ 4 & 2-\lambda\end{vmatrix}=\lambda^2-3\lambda-10.$

(2) 由 $\lambda^2-3\lambda-10=0$ 解得特征值 $\lambda_1=5, \lambda_2=-2.$

(3) 对于 $\lambda_1=5$, $\boldsymbol{A}-5\boldsymbol{E}=\begin{pmatrix}-4 & 3\\ 4 & -3\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & -3/4\\ 0 & 0\end{pmatrix}$, 得到基础解系 $\begin{pmatrix}3/4\\ 1\end{pmatrix}$. 故对应的所有特征向量为 $k(3,4)^{\mathrm{T}}, k\neq 0.$

(4) 对于 $\lambda_2=-2$, $\boldsymbol{A}+2\boldsymbol{E}=\begin{pmatrix}3 & 3\\ 4 & 4\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}$, 得到基础解系 $\begin{pmatrix}1\\ -1\end{pmatrix}$. 故对应的所有特征向量为 $k(1,-1)^{\mathrm{T}}, k\neq 0.$

例

求上三角阵 $\boldsymbol{A}=\begin{pmatrix} a_1 & b & c \\ 0 & a_2 & d \\ 0 & 0 & a_3 \end{pmatrix}$ 的特征值.

解

特征多项式

$$|\boldsymbol{A}-\lambda \boldsymbol{E}|=\begin{vmatrix} a_1-\lambda & b & c \\ 0 & a_2-\lambda & d \\ 0 & 0 & a_3-\lambda \end{vmatrix}=(a_1-\lambda)(a_2-\lambda)(a_3-\lambda).$$

因此特征值为 a_1, a_2, a_3.

上三角阵、下三角阵、对角阵的特征值就是对角元.

例

求 $\boldsymbol{A}=\begin{pmatrix}-1 & 1 & 0\\ -4 & 3 & 0\\ 1 & 0 & 2\end{pmatrix}$ 的特征值和特征向量.

解

由特征多项式

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

得到特征值 $\lambda=1,1,2.$

续解

对于 $\lambda_1 = 1$,

$$\boldsymbol{A}-\boldsymbol{E}=\begin{pmatrix}-2 & 1 & 0\\ -4 & 2 & 0\\ 1 & 0 & 1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 2\\ 0 & 0 & 0\end{pmatrix},$$

得到基础解系 $(-1,-2,1)^{\mathrm{T}}$. 故对应的所有特征向量为 $k(-1,-2,1)^{\mathrm{T}}, k \neq 0$.

对于 $\lambda_2 = 2$,

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}-3 & 1 & 0\\ -4 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

得到基础解系 $(0,0,1)^{\mathrm{T}}$. 故对应的所有特征向量为 $k(0,0,1)^{\mathrm{T}}, k \neq 0$.

练习

求 $\boldsymbol{A}=\begin{pmatrix}-2 & 1 & 1\\ 0 & 2 & 0\\ -4 & 1 & 3\end{pmatrix}$ 的特征值和特征向量.

答案

(1) 特征值 $\lambda=-1,2,2$.

(2) -1 对应的所有特征向量为 $k(1,0,1)^{\mathrm{T}}, k\neq 0$.

(3) 2 对应的所有特征向量为 $k_1(0,1,-1)^{\mathrm{T}}+k_2(1,0,4)^{\mathrm{T}}, k_1,k_2$ 不全为零.

若 λ 是 k 重特征值, 则它对应的线性无关的特征向量最多 k 个.

通过

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

的展开可以看出, 若 $i \neq j$, 则 i 行 j 列的代数余子式中最多只会出现 λ^{n-2} 项. 所以

$$\begin{aligned} |\boldsymbol{A}-\lambda\boldsymbol{E}| &= (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) + \text{至多 } n-2 \text{ 次项} \\ &= (-1)^n(\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}) + \text{至多 } n-2 \text{ 次项}. \end{aligned}$$

根据韦达定理, 特征值之和为 $a_{11}+a_{22}+\cdots+a_{nn}$.

$\lambda=0$ 时, 特征多项式

$$f(\lambda)=(\lambda_1-\lambda)\cdots(\lambda_n-\lambda)$$

的取值 $|\boldsymbol{A}|$ 就是特征值的乘积.

特征值和特征向量的性质

定义 $\boldsymbol{A}$ 的**迹**为对角元之和:

$$\mathrm{Tr}(\boldsymbol{A}) = a_{11} + a_{22} + \cdots + a_{nn}.$$

(1) 特征值之和等于迹: $\lambda_1 + \lambda_2 + \cdots + \lambda_n = \mathrm{Tr}(\boldsymbol{A})$;

(2) 特征值之积等于行列式: $\lambda_1\lambda_2\cdots\lambda_n = |\boldsymbol{A}|$.

例

$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ 的特征值为(C).

(A) $1, 0, 1$　　(B) $1, 1, 2$　　(C) $-1, 1, 2$　　(D) $-1, 1, 1$

定理

若 λ 是 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 是对应特征向量，则下述矩阵有如下对应的特征值与特征向量：

方阵	$k\boldsymbol{A}$	$\boldsymbol{A}^m$	$\boldsymbol{A}^{-1}$	$\boldsymbol{A}^*$	$g(\boldsymbol{A})h(\boldsymbol{A})^{-1}$	$\boldsymbol{A}^{\mathrm{T}}$	$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$
特征值	$k\lambda$	λ^m	λ^{-1}	$\lvert\boldsymbol{A}\rvert/\lambda$	$g(\lambda)/h(\lambda)$	λ	λ
对应特征向量	$\boldsymbol{x}$	$\boldsymbol{x}$	$\boldsymbol{x}$	$\boldsymbol{x}$	$\boldsymbol{x}$	未必是 $\boldsymbol{x}$	$\boldsymbol{P}^{-1}\boldsymbol{x}$

这里 g, h 是多项式，且满足 $h(\boldsymbol{A})$ 可逆.

由此可知，若 $g(\boldsymbol{A}) = \boldsymbol{O}$，则 $\boldsymbol{A}$ 的所有特征值 λ 均满足 $g(\lambda) = 0$.

例：特征值和特征向量的性质

例

设 3 阶方阵 $\boldsymbol{A}$ 的特征值为 $2,2,3$，则

(1) $\boldsymbol{A}^2$ 的特征值为 $\underline{\quad 4,4,9 \quad}$;

(2) $\boldsymbol{A}^2-2\boldsymbol{A}+\boldsymbol{E}$ 的特征值为 $\underline{\quad 1,1,4 \quad}$;

(3) $|\boldsymbol{A}^2-2\boldsymbol{A}+\boldsymbol{E}|=\underline{\quad 4 \quad}$;

(4) $\boldsymbol{A}^{-1}$ 的特征值为 $\underline{\quad 1/2,1/2,1/3 \quad}$;

(5) $\boldsymbol{A}^*$ 的特征值为 $\underline{\quad 4,6,6 \quad}$;

(6) $A_{11}+A_{22}+A_{33}=\underline{\quad 16 \quad}$, 其中 A_{ij} 表示 $\boldsymbol{A}$ 的代数余子式.

例：特征值和特征向量的性质

例

设 3 阶方阵 $\boldsymbol{A}$ 的特征值为 $1,-1,2$. 求 $|\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}|$.

解

由于 $|\boldsymbol{A}|=-2$, $\boldsymbol{A}^*=-2\boldsymbol{A}^{-1}$. 于是

$$\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}=-2\boldsymbol{A}^{-1}+3\boldsymbol{A}-2\boldsymbol{E}$$

的特征值为 $-1,-3,3$, $|\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}|=(-1)\times(-3)\times 3=9$.

练习

若 4 阶方阵 $\boldsymbol{A}$ 的特征值为 $1,2,-2,0$, 则下列矩阵可逆的是(C).

(A) $\boldsymbol{A}$　　(B) $\boldsymbol{A}-2\boldsymbol{E}$　　(C) $\boldsymbol{A}+\boldsymbol{E}$　　(D) $\boldsymbol{A}-\boldsymbol{E}$

例

若 $\boldsymbol{\alpha} = (1, k, 1)^{\mathrm{T}}$ 为 $\boldsymbol{A} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ 的逆 $\boldsymbol{A}^{-1}$ 的特征向量，求 k.

解

$\boldsymbol{\alpha}$ 也是 $\boldsymbol{A}$ 的特征向量.

$$\boldsymbol{A\alpha} = \begin{pmatrix} k+3 \\ 2k+2 \\ k+3 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ k \\ 1 \end{pmatrix},$$

因此 $\lambda = k+3, 2k+2 = (k+3)k, k^2+k-2=0, k=-2$ 或 1.

例：特征值的应用

例

计算 $\boldsymbol{A}=\begin{pmatrix} a & b & \cdots & b \\ b & a & \cdots & b \\ \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a \end{pmatrix}$ 的行列式.

解

设 $\boldsymbol{B}=\boldsymbol{A}+(b-a)\boldsymbol{E}$, 则 $\boldsymbol{B}$ 所有元素为 b, 秩为 1, $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 基础解系有 $n-1$ 个向量. 从而 0 是 $\boldsymbol{B}$ 的至少 $n-1$ 重特征值. 由于 $\mathrm{Tr}(\boldsymbol{B})=nb$, 因此 $\boldsymbol{B}$ 的所有特征值为 $nb,0,\ldots,0$, $\boldsymbol{A}$ 的所有特征值为 $nb+a-b,a-b,\ldots,a-b$,

$$|\boldsymbol{A}|=(a-b)^{n-1}(nb-b+a).$$

定理

若 $\lambda_1, \ldots, \lambda_m$ 是 $\boldsymbol{A}$ 的 m 个两两不同的特征值，则其对应的特征向量 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ 线性无关.

即对应于不同特征值的特征向量线性无关.

证明

设 $k_1\boldsymbol{\alpha}_1 + \cdots + k_m\boldsymbol{\alpha}_m = \mathbf{0}$. 左乘 $\boldsymbol{A}^k$ 得到 $k_1\lambda_1^k\boldsymbol{\alpha}_1 + \cdots + k_m\lambda_m^k\boldsymbol{\alpha}_m = \mathbf{0}$. 令 $k = 1, 2, \ldots, m-1$, 我们得到

$$(k_1\boldsymbol{\alpha}_1, \cdots, k_m\boldsymbol{\alpha}_m)\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{m-1} \end{pmatrix} = \boldsymbol{O}.$$

注意到第二个方阵的行列式是范德蒙行列式，当 $\lambda_1, \ldots, \lambda_m$ 两两不同时它非零，从而

$$(k_1\boldsymbol{\alpha}_1, \cdots, k_m\boldsymbol{\alpha}_m) = \boldsymbol{O}, \qquad k_1 = \cdots = k_m = 0.$$

□

设 λ_1 对应线性无关的特征向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$, λ_2 对应线性无关的特征向量 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$. 若 $\lambda_1 \neq \lambda_2$, 则 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ 也是线性无关的. 这是因为

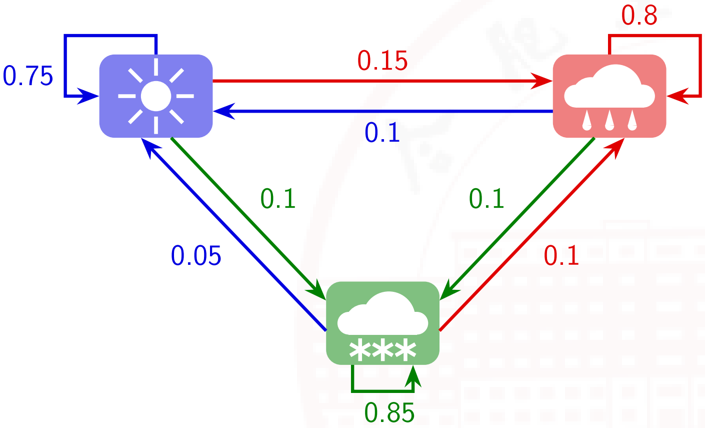
$$\boldsymbol{A}(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2) = \lambda_1(k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2),$$

$$\boldsymbol{A}(\ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2) = \lambda_2(\ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2).$$

若 $k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \ell_1\boldsymbol{\beta}_1 + \ell_2\boldsymbol{\beta}_2 = \mathbf{0}$, 同理可证明这些向量都是零向量.

由此也可以知道, 不同特征值的特征向量的线性组合不可能还是特征向量.

将天气简化为晴雨雪三种，其它天气由它们组合得到. 根据该地历史的天气信息，得到当天与第二天天气的关系:



例

某地某季节天气仅受前一天天气状态影响：设 k 天后天气为晴天、雨天、雪天概率分别为 a_k, b_k, c_k，则

$$\begin{pmatrix} a_k \\ b_k \\ c_k \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} a_{k-1} \\ b_{k-1} \\ c_{k-1} \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 0.75 & 0.1 & 0.05 \\ 0.15 & 0.8 & 0.1 \\ 0.1 & 0.1 & 0.85 \end{pmatrix}.$$

如果今天天气为晴天，请问未来七天的各种天气概率分别是多少？

解

设 $\boldsymbol{x}_k=(a_k,b_k,c_k)^{\mathrm{T}}$，则 $\boldsymbol{x}_k=\boldsymbol{A}\boldsymbol{x}_{k-1}$. 因此 $x_k=\boldsymbol{A}x_{k-1}=\boldsymbol{A}^2x_{k-2}=\cdots=\boldsymbol{A}^k\boldsymbol{x}_0$. 解特征多项式

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix}0.75-\lambda & 0.1 & 0.05\\ 0.15 & 0.8-\lambda & 0.1\\ 0.1 & 0.1 & 0.85-\lambda\end{vmatrix}=0$$

得到特征值 $\lambda_1=1,\lambda_2=0.75,\lambda_3=0.65$. 解方程 $(\boldsymbol{A}-\lambda_i\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$ 得到特征向量

$$\boldsymbol{v}_1=(8,13,14)^{\mathrm{T}},\quad \boldsymbol{v}_2=(1,1,-2)^{\mathrm{T}},\quad \boldsymbol{v}_3=(1,-1,0)^{\mathrm{T}}.$$

由于今天天气是晴天，所以 $\boldsymbol{x}_0=(1,0,0)^{\mathrm{T}}$. 注意到 $\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3$ 线性无关，因此 $\boldsymbol{x}_0$ 可以由它们线性表示. 通过计算得到 $\boldsymbol{x}_0=\frac{1}{35}\boldsymbol{v}_1+\frac{1}{5}\boldsymbol{v}_2+\frac{4}{7}\boldsymbol{v}_3$，因此

$$\boldsymbol{x}_k=\boldsymbol{A}^k\boldsymbol{x}_0=\frac{1}{35}\lambda_1^k\boldsymbol{v}_1+\frac{1}{5}\lambda_2^k\boldsymbol{v}_2+\frac{4}{7}\lambda_3^k\boldsymbol{v}_3.$$

应用：天气预报

续解

代入计算可得 $\boldsymbol{x}_k$，从而得到未来七天晴雨雪的概率：

	明天	后天	3 天后	4 天后	5 天后	6 天后	7 天后
晴	0.75	0.58	0.47	0.39	0.34	0.31	0.28
雨	0.15	0.24	0.30	0.33	0.35	0.36	0.37
雪	0.10	0.18	0.23	0.27	0.31	0.33	0.35

基的变换

设 $f:\mathbb{C}^n\to\mathbb{C}^n$ 是一个线性映射. 对于任意 $\boldsymbol{x}\in\mathbb{C}^n$, 它可以唯一表达为

$$\boldsymbol{x}=x_1\boldsymbol{e}_1+\cdots+x_n\boldsymbol{e}_n.$$

通过将 $f(\boldsymbol{e}_1),\ldots,f(\boldsymbol{e}_n)$ 表达为这组标准基的线性组合, 我们建立了线性映射 f 与一个 n 阶方阵 $\boldsymbol{A}$ 的联系.

一般的线性空间并没有这样的标准正交基, 或者 $\mathbb{C}^n$ 本身我们也可以选择其它基. 则这种对应会有什么变化呢? 设 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 是 $\mathbb{C}^n$ 中一组线性无关的向量,

$$\boldsymbol{P}=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n).$$

则

$$f(\boldsymbol{\alpha}_i)=\boldsymbol{A}\boldsymbol{\alpha}_i=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n)\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_i,$$

所以 $f(\boldsymbol{\alpha}_i)$ 表达为 $\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_n$ 线性组合的系数形成的向量是 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_i$. 它们构成矩阵

$$(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_1,\ldots,\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{\alpha}_n)=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}.$$

也就是说, 若我们将线性空间 $\mathbb{C}^n$ 换一组基表达, 线性映射对应的矩阵就会变成 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$. 我们称 f 在不同基下矩阵为相似的.

定义

若存在可逆矩阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$, 则称方阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似.

注意, 相似只对方阵有定义.

命题

方阵的相似满足

(1) 自反性: $\boldsymbol{A}$ 与自身相似;

(2) 对称性: 若 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$, 则 $\boldsymbol{B}$ 相似于 $\boldsymbol{A}$;

(3) 传递性: 若 $\boldsymbol{A}$ 相似于 $\boldsymbol{B}$, $\boldsymbol{B}$ 相似于 $\boldsymbol{C}$, 则 $\boldsymbol{A}$ 相似于 $\boldsymbol{C}$.

若 $\boldsymbol{A},\boldsymbol{B}$ 相似 $(\boldsymbol{B}=\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P})$, 则 $\boldsymbol{A},\boldsymbol{B}$ 等价 $(\boldsymbol{B}=\boldsymbol{P}\boldsymbol{A}\boldsymbol{Q})$. 反之未必成立, 例如 $\boldsymbol{A}=\begin{pmatrix}1&0\\0&1\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}1&1\\0&1\end{pmatrix}$. 这是因为和 $\boldsymbol{A}=\boldsymbol{E}$ 相似的只有它自己.

例

若 3 阶方阵 $\boldsymbol{A}\overset{r_1\leftrightarrow r_3}{\sim}\boldsymbol{B}\overset{c_1\leftrightarrow c_3}{\sim}\boldsymbol{C}$, 则 $\boldsymbol{A}$ 与 $\boldsymbol{C}$ (C).

(A) 等价但不相似　　(B) 相似但不等价

(C) 等价而且相似　　(D) 既不等价也不相似

相似矩阵的性质

若 A 与 B 相似, 则二者的特征多项式相同, 从而特征值也相同.

这是因为若 $P^{-1}AP = B$, 则

$$P^{-1}(A - \lambda E)P = P^{-1}AP - \lambda P^{-1}EP = B - \lambda E.$$

两边取行列式并利用行列式的可乘性得到 $|A - \lambda E| = |B - \lambda E|$.

注意反过来未必成立, 例如 $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. 的特征多项式相同, 但它们不相似.

相似矩阵具有相同的特征值, 但对应的特征向量未必相同.

相似矩阵的性质

若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，则

(1) $\boldsymbol{A}, \boldsymbol{B}$ 特征值相同 (包括重数);

(2) $|\boldsymbol{A}| = |\boldsymbol{B}|$, $\mathrm{Tr}(\boldsymbol{A}) = \mathrm{Tr}(\boldsymbol{B})$;

(3) $\boldsymbol{A} \sim \boldsymbol{B}$, 即 $\mathrm{R}(\boldsymbol{A}) = \mathrm{R}(\boldsymbol{B})$;

(4) 对于多项式 g, h, 若 $h(\boldsymbol{A})$ 可逆，则 $h(\boldsymbol{B})$ 也可逆，且 $g(\boldsymbol{A})/h(\boldsymbol{A})$ 与 $g(\boldsymbol{B})/h(\boldsymbol{B})$ 相似. 特别地，$\boldsymbol{A} - \lambda \boldsymbol{E}$ 与 $\boldsymbol{B} - \lambda \boldsymbol{E}$ 相似.

推论

若 $\boldsymbol{A}$ 与对角阵 $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ 相似，则 $\lambda_1, \ldots, \lambda_n$ 是 $\boldsymbol{A}$ 的 n 个特征值.

例：相似矩阵的性质

例

若 $\boldsymbol{A}=\begin{pmatrix}2&-1&4\\0&a&7\\0&0&3\end{pmatrix}$ 与 $\boldsymbol{\Lambda}=\begin{pmatrix}1&0&0\\0&2&0\\0&0&b\end{pmatrix}$ 相似，则 $a=$ <u>1</u>，$b=$ <u>3</u>.

例

若 3 阶可逆阵 $\boldsymbol{A},\boldsymbol{B}$ 相似，$\boldsymbol{A}^{-1}$ 的特征值为 $1/2,1/3,1/4$，则 $|\boldsymbol{E}-\boldsymbol{B}|=$ <u>-6</u>.

练习

若 3 阶矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相似，且存在非零矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{AC}=2\boldsymbol{C},|\boldsymbol{A}+2\boldsymbol{E}|=0,|2\boldsymbol{A}-\boldsymbol{E}|=0$，则 $|\boldsymbol{B}^{-1}-\boldsymbol{E}|=$ <u>$3/4$</u>.

若 $\boldsymbol{AB}=k\boldsymbol{B}$，则 $\boldsymbol{B}$ 的每个非零列向量均为 $\boldsymbol{A}$ 的属于特征值 k 的特征向量.

相似对角化的定义

定义

若 n 阶方阵 $\boldsymbol{A}$ 相似于某个对角阵

$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),$$

则称 $\boldsymbol{A}$ 可**(相似) 对角化**.

设 $\boldsymbol{P} = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$, 则

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda} \iff \boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{\Lambda} \iff (\boldsymbol{A}\boldsymbol{p}_1, \ldots, \boldsymbol{A}\boldsymbol{p}_n) = (\lambda_1\boldsymbol{p}_1, \ldots, \lambda_n\boldsymbol{p}_n),$$

即 $\boldsymbol{A}\boldsymbol{p}_i = \lambda_i\boldsymbol{p}_i$. 由于 $\boldsymbol{P}$ 可逆, $\boldsymbol{A}$ 拥有 n 个线性无关的特征向量 $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$. 反之, 若 $\boldsymbol{A}$ 拥有 n 个线性无关的特征向量, 则选择 $\boldsymbol{P}$ 以它们为列向量即可使 $\boldsymbol{A}$ 对角化.

相似对角化的等价刻画

n 阶矩阵 $\boldsymbol{A}$ 可对角化 $\Longleftrightarrow$ $\boldsymbol{A}$ 有 n 个线性无关的特征向量.

推论

若 $\boldsymbol{A}$ 的特征值两两不同，则 $\boldsymbol{A}$ 可对角化.

反之未必成立.

例

设 3 阶方阵 $\boldsymbol{B}$ 的特征值为 $1, 2, -2$, $\boldsymbol{A} = \boldsymbol{B}^3 - 4\boldsymbol{B} + \boldsymbol{E}$, 求 $\boldsymbol{A}$ 的相似对角阵.

解

由于 $\boldsymbol{B}$ 特征值两两不同，因此存在 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P} = \boldsymbol{\Lambda} = \operatorname{diag}(1, 2, -2)$. 于是

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{\Lambda}^3 - 4\boldsymbol{\Lambda} + \boldsymbol{E} = \operatorname{diag}(-2, 1, 1).$$

回忆 k 重特征值对应的线性无关的特征向量最多 k 个.

相似对角化的等价刻画

若 λ 是 $\boldsymbol{A}$ 的 k 重特征值, 则 $\boldsymbol{A}$ 可对角化 $\iff \forall\lambda, \mathrm{R}(\boldsymbol{A}-\lambda\boldsymbol{E}) = n-k$, 即 λ 对应的线性无关的特征向量恰有 k 个.

相似对角化的步骤如下:

(1) 求出 $\boldsymbol{A}$ 的所有特征值 λ_i 和特征向量 $\boldsymbol{p}_i$;

(2) 根据上述判定方法判断 $\boldsymbol{A}$ 是否可以相似对角化;

(3) 若能, 将 n 个对应的线性无关的特征向量 $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ 组成方阵 $\boldsymbol{P} = (\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n)$,

$$\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

例：对角化的计算

例

$\boldsymbol{A}=\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}$ 能否对角化？若能，求 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 是对角阵.

解

(1) 上三角阵 $\boldsymbol{A}$ 特征值为 $1,0,0$.

(2) 对于 $\lambda_1=1$, $\boldsymbol{A}-\boldsymbol{E}=\begin{pmatrix}0&1&1\\0&-1&0\\0&0&-1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$, 取特征向量 $\boldsymbol{p}_1=(1,0,0)^{\mathrm{T}}$.

(3) 对于 $\lambda_2=\lambda_3=0$, $\boldsymbol{A}$ 对应的基础解系可以取 $\boldsymbol{p}_2=(-1,1,0)^{\mathrm{T}}$, $\boldsymbol{p}_3=(-1,0,1)^{\mathrm{T}}$.

(4) 因此 $\boldsymbol{A}$ 可对角化，取 $\boldsymbol{P}=\begin{pmatrix}1&-1&-1\\0&1&0\\0&0&1\end{pmatrix}$, 则 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(1,0,0)$.

例

判断下列矩阵 A 能否相似对角化: (1) A 是二阶实矩阵且 $|A| < 0$;

(2) $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$; (3) $A = \begin{pmatrix} -2 & 1 & 1 \\ 0 & 2 & 0 \\ -4 & 1 & 3 \end{pmatrix}$; (4) $A = \begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$.

解

(1) 特征值一正一负, 能对角化; (2) 特征值为 $1, 3, 0$, 能对角化;

(3) $(2-\lambda)(\lambda^2-\lambda-2) \implies \lambda = -1, 2, 2$. $A - 2E = \begin{pmatrix} -4 & 1 & 1 \\ 0 & 0 & 0 \\ -4 & 1 & 1 \end{pmatrix}$ 秩 1, 能对角化;

(4) $(2-\lambda)(\lambda^2-2\lambda+2) \implies \lambda = 2, 1, 1$. $A - E = \begin{pmatrix} -2 & 1 & 0 \\ -4 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ 秩 2, 不能对角化.

练习

(1) 若 $\boldsymbol{A}=\begin{pmatrix}0&0&1\\a&1&1\\1&0&0\end{pmatrix}$ 可对角化，则 $a=$ $\underline{\quad -1 \quad}$.

(2) 设 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3),\boldsymbol{Q}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_2)$. 若 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\operatorname{diag}(1,2,3)$，则 $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q}=\underline{\operatorname{diag}(1,3,2)}$.

(3) 若 3 阶方阵 $\boldsymbol{A}$ 的特征值互不相同且 $|\boldsymbol{A}|=0$，则 $\mathrm{R}(\boldsymbol{A})=$ $\underline{\quad 2 \quad}$.

(4) 若 $\boldsymbol{A}=\begin{pmatrix}2&0&0\\0&2&1\\0&0&1\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix},\boldsymbol{C}=\begin{pmatrix}1&0&0\\0&2&0\\0&0&2\end{pmatrix}$，则(C).

(A) $\boldsymbol{A},\boldsymbol{C}$ 相似，$\boldsymbol{B},\boldsymbol{C}$ 相似

(B) $\boldsymbol{A},\boldsymbol{C}$ 不相似，$\boldsymbol{B},\boldsymbol{C}$ 相似

(C) $\boldsymbol{A},\boldsymbol{C}$ 相似，$\boldsymbol{B},\boldsymbol{C}$ 不相似

(D) $\boldsymbol{A},\boldsymbol{C}$ 不相似，$\boldsymbol{B},\boldsymbol{C}$ 不相似

例

设 $\boldsymbol{A}=\begin{pmatrix}1 & -1 & 1\\ x & 4 & y\\ -3 & -3 & 5\end{pmatrix}$ 有 3 个线性无关特征向量, $\lambda=2$ 是 $\boldsymbol{A}$ 的二重特征值. 求可逆阵 $\boldsymbol{P}$, 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ 为对角阵.

解

由 $\mathrm{Tr}(\boldsymbol{A})=10$ 可知特征值为 $2,2,6$. 由 $\mathrm{R}(\boldsymbol{A}-2\boldsymbol{E})=1$ 可知 $x=2, y=-2$.

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}-1 & -1 & 1\\ x & 2 & y\\ -3 & -3 & 3\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 1 & -1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\Longrightarrow \boldsymbol{p}_1=\begin{pmatrix}-1\\ 1\\ 0\end{pmatrix}, \boldsymbol{p}_2=\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}.$$

$$\boldsymbol{A}-6\boldsymbol{E}=\begin{pmatrix}-5 & -1 & 1\\ 2 & -2 & -2\\ -3 & -3 & -1\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1 & 0 & -1/3\\ 0 & 1 & 2/3\\ 0 & 0 & 0\end{pmatrix}\Longrightarrow \boldsymbol{p}_3=\begin{pmatrix}1\\ -2\\ 3\end{pmatrix}. \quad \boldsymbol{P}=\begin{pmatrix}-1 & 1 & 1\\ 1 & 0 & -2\\ 0 & 1 & 3\end{pmatrix}.$$

例：对角化的计算

例

若 $\boldsymbol{A}$ 的各行元素之和为 2, $\boldsymbol{A}\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, 则 $\boldsymbol{A}$ 相似于(C).

(A) $\mathrm{diag}(1,1,2)$ (B) $\mathrm{diag}(2,1,1)$ (C) $\mathrm{diag}(2,1,-1)$ (D) $\mathrm{diag}(2,-1,-1)$

练习

设 $\boldsymbol{A} = \begin{pmatrix} 2 & a & 2 \\ 5 & b & 3 \\ -1 & 1 & -1 \end{pmatrix}$ 有特征值 ± 1, 问 $\boldsymbol{A}$ 能否相似对角化?

答案

$|\boldsymbol{A} \pm \boldsymbol{E}| = 0 \implies a = -1, b = -3$. $\mathrm{Tr}(\boldsymbol{A}) = -2 \implies \lambda_3 = -2$, 可对角化.

例：对角化的计算

例

设 $\boldsymbol{A}$ 为三阶方阵, $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 是线性无关的三维列向量且

$$\boldsymbol{A\alpha}_1=2\boldsymbol{\alpha}_1,\boldsymbol{A\alpha}_2=3\boldsymbol{\alpha}_2+2\boldsymbol{\alpha}_3,\boldsymbol{A\alpha}_3=2\boldsymbol{\alpha}_2+3\boldsymbol{\alpha}_3.$$

证明 $\boldsymbol{A}$ 可对角化.

解

设 $\boldsymbol{P}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)$, 则

$$\boldsymbol{AP}=\boldsymbol{A}(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3)\boldsymbol{B}=\boldsymbol{PB},\quad \text{其中 } \boldsymbol{B}=\begin{pmatrix}2&0&0\\0&3&2\\0&2&3\end{pmatrix}.$$

由于 $\boldsymbol{B}$ 的特征值是 $2,1,5$, 因此 $\boldsymbol{B}$ 能对角化, 从而 $\boldsymbol{A}=\boldsymbol{PBP}^{-1}$ 也可以.

一般地, 实对称矩阵一定能对角化. 我们将在下一节中解释这为何成立.

任何方阵都可以相似于约当标准形

$$\operatorname{diag}(\boldsymbol{J}_{k_1}(\lambda_1),\ldots,\boldsymbol{J}_{k_t}(\lambda_t)),$$

其中 $\boldsymbol{J}_k(\lambda)=\begin{pmatrix}\lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda\end{pmatrix}$ 是 k 阶方阵. 当 $k_1=\cdots=k_t=1$ 时这就是对角阵.

特别地, 任何方阵都可以相似于一个上三角阵.

容易知道上三角阵的伴随矩阵的对角元. 因此 $\boldsymbol{A}^*$ 的所有特征值的就是 $\boldsymbol{A}$ 的 n 个特征值中 $n-1$ 个相乘得到的.

(1) 当 $|\boldsymbol{A}| \neq 0$ 时, $\boldsymbol{A}^*$ 的所有特征值为 $|\boldsymbol{A}|/\lambda$.

(2) 当 $|\boldsymbol{A}| = 0$ 且 $\lambda_1 = 0$ 是一重特征值, 则 $\boldsymbol{A}^*$ 唯一的非零特征值为 $\boldsymbol{A}$ 非零特征值之乘积.

(3) 当 $|\boldsymbol{A}| = 0$ 且 $\lambda_1 = 0$ 是 $\geqslant 2$ 重特征值, 则 $\boldsymbol{A}^* = \boldsymbol{O}$.

练习

若 4 阶实矩阵 $\boldsymbol{A}^*$ 的特征值为 $-1, 1, 2, 4$, 则下列矩阵可逆的是(D).

(A) $\boldsymbol{A}+2\boldsymbol{E}$ (B) $\boldsymbol{A}-2\boldsymbol{E}$ (C) $2\boldsymbol{A}+\boldsymbol{E}$ (D) $\boldsymbol{A}-\boldsymbol{E}$

例：方阵的相似的应用

利用约当标准形还可以计算任意矩阵的方幂：

$$J_k(\lambda)^m = (\lambda E + N)^m = \sum_{i=0}^{m} \mathrm{C}_m^i \lambda^{m-i} N^i, \quad \text{其中 } N = J_k(0).$$

例

设 $A = \begin{pmatrix} 5 & -6 \\ 1 & 0 \end{pmatrix}$, 计算 A^k.

解

A 特征值为 $2, 3$, 对应的特征向量可以取 $p_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $p_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. 设 $P = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix}$, 则

$P^{-1}AP = \Lambda = \begin{pmatrix} 2 & \\ & 3 \end{pmatrix}$, $A^k = P\Lambda^k A^{-1}$

续解

因此

$$A^k = \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2^k & \\ & 3^k \end{pmatrix} \begin{pmatrix} -1 & 3 \\ 1 & -2 \end{pmatrix} = \begin{pmatrix} 3^{k+1} - 2^{k+1} & 6(2^k - 3^k) \\ 3^k - 2^k & 6(2^{k-1} - 3^{k-1}) \end{pmatrix}.$$

线性递推数列是指满足如下递推关系的数列

$$a_{n+k} - c_{k-1}a_{n+k-1} - \cdots - c_1 a_{n+1} - c_0 a_n = 0,$$

它的通项可以用矩阵方法来计算. 例如

$$a_{n+2} = 5a_{n+1} - 6a_n, \quad a_0 = a_1 = 1.$$

设 $\boldsymbol{x}_n = \begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix}$, 则 $\boldsymbol{x}_n = \boldsymbol{A}\boldsymbol{x}_{n-1} = \boldsymbol{A}^n \boldsymbol{x}_0 = \begin{pmatrix} 2^{n+2} - 3^{n+1} \\ 2^{n+1} - 3^n \end{pmatrix}$, 故 $a_n = 2^{n+1} - 3^n$.

第二节　实对称阵的正交合同

- 实二次型

- 实对称阵和实二次型的合同

我们知道，

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1,\quad \frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$

分别表示椭圆和双曲线. 对于二次曲线

$$Ax^2+Bxy+Cy^2=1,$$

它又表示什么图形呢?

实二次型的定义

定义

若 n 元多项式 $f(x_1, \ldots, x_n)$ 满足

$$f(\lambda x_1, \ldots, \lambda x_n) = \lambda^2 f(x_1, \ldots, x_n), \quad \forall \lambda \in \mathbb{R},$$

则称 f 为二次齐次多项式或实二次型. 它可以写成

$$f(x) = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \cdots + 2a_{n-1,n}x_{n-1}x_n.$$

本课程仅讨论实二次型. 根据定义, f 不能包含一次项和常数项. 若 f 的交叉项 $x_ix_j(i < j)$ 系数均为零, 则称 f 为实二次型的标准形.

实二次型的矩阵形式

设实二次型 f 的 x_i^2 项的系数为 a_{ii}, $x_ix_j(i<j)$ 项的系数为 $2a_{ij}$. 设 $a_{ji}=a_{ij}$, 对称阵 $\boldsymbol{A}=(a_{ij})_n$, 则

$$(x_1,\dots,x_n)\boldsymbol{A}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\Big(\sum_{i=1}^n a_{i1}x_i,\dots,\sum_{i=1}^n a_{in}x_i\Big)\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$$

$$=\sum_{j=1}^n\Big(\sum_{i=1}^n a_{ij}x_i\Big)x_j=\sum_{i,j=1}^n a_{ij}x_ix_j=f(x_1,\dots,x_n),$$

即 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$, 其中 $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{T}}$, $\boldsymbol{A}$ 为实对称阵.

反过来, 任给一个实对称阵 $\boldsymbol{A}$, 多项式 $f(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 显然满足

$$f(\lambda\boldsymbol{x})=(\lambda\boldsymbol{x})^{\mathrm{T}}\boldsymbol{A}(\lambda\boldsymbol{x})=\lambda^2f(\boldsymbol{x}),$$

故 f 是实二次型. 因此实二次型 f 与对称阵 $\boldsymbol{A}$ 之间存在一一对应的关系.

例

写出实二次型 $f(x_1, x_2, x_3) = x_1^2 + 4x_1x_2 + 4x_2^2 + 2x_1x_3 + x_3^2 + 4x_2x_3$ 对应的矩阵.

解

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

若 f 是标准形, 则 f 对应矩阵 $\boldsymbol{A}$ 是对角阵.

定义

(1) 若存在可逆线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 使得实二次型 f 在变量 $\boldsymbol{x},\boldsymbol{y}$ 下的矩阵分别为 $\boldsymbol{A},\boldsymbol{B}$, 则称矩阵 $\boldsymbol{A}$ 合同或相合于 $\boldsymbol{B}$.

(2) 若 $\boldsymbol{P}$ 是正交阵, 则称矩阵 $\boldsymbol{A}$ 正交合同或正交相合于 $\boldsymbol{B}$.

若 $\boldsymbol{A}$ 是对称阵, $\boldsymbol{P}$ 可逆, 则 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 也是对称阵. 由

$$f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}\boldsymbol{y}=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P})\boldsymbol{y}$$

可知 $\boldsymbol{A}$ (正交) 合同于 $\boldsymbol{B}$ 当且仅当存在可逆 (正交) 方阵 $\boldsymbol{P}$ 使得 $\boldsymbol{B}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$.

命题

对称方阵的 (正交) 合同满足

(1) 自反性: A 与自身 (正交) 合同;

(2) 对称性: 若 A (正交) 合同于 B, 则 B (正交) 合同于 A;

(3) 传递性: 若 A (正交) 合同于 B, B (正交) 合同于 C, 则 A (正交) 合同于 C.

合同、等价、相似有如下关系:

(1) 若 A, B 合同, 则 A, B 等价, $\mathrm{R}(A) = \mathrm{R}(B)$. 反之未必.

(2) 若 A, B 正交合同, 则 A, B 相似. 反之, 若实对称阵 A, B 相似, 则二者正交合同.

定理

对于实对称阵 $\boldsymbol{A}$, 存在正交阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 是对角阵. 从而 $\boldsymbol{A}$ 对应的实二次型在线性变换 $\boldsymbol{y}=\boldsymbol{P}\boldsymbol{x}$ 下变为标准形.

命题

实对称阵的特征值一定是实数, 从而其特征向量均可取实向量.

证明

设 $\boldsymbol{A}$ 是实对称阵, 非零向量 $\boldsymbol{x}$ 满足 $\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x}$. 两边取转置和共轭并右乘 $\boldsymbol{x}$ 得到

$$\bar{\lambda}\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=\overline{\boldsymbol{x}}^{\mathrm{T}}\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{x}=\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\lambda\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}.$$

显然 $\overline{\boldsymbol{x}}^{\mathrm{T}}\boldsymbol{x}=|x_1|^2+\cdots+|x_n|^2>0$, 因此 λ 是实数. 由于特征向量是方程 $(\boldsymbol{A}-\lambda\boldsymbol{E})\boldsymbol{x}=\boldsymbol{0}$ 的解, 系数矩阵是实方阵, 因此特征向量可取实向量. $\square$

定理的证明

归纳证明 $\boldsymbol{A}$ 存在 n 个两两正交的单位特征向量. 假设我们已找到 $k<n$ 个两两正交的单位特征向量 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_k$, 分别对应特征值 $\lambda_1,\dots,\lambda_k$. 设 V 是与这些向量均正交的实向量全体, 即方程 $(\boldsymbol{e}_1,\dots,\boldsymbol{e}_k)^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 的解空间. 由于系数秩为 k, 存在基础解系 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k}\in V$. 对任意 i,j,

$$[\boldsymbol{e}_i,\boldsymbol{A}\boldsymbol{v}_j]=\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}_j=(\boldsymbol{A}\boldsymbol{e}_i)^{\mathrm{T}}\boldsymbol{v}_j=\lambda_i\boldsymbol{e}_i^{\mathrm{T}}\boldsymbol{v}_j=0\implies\boldsymbol{A}\boldsymbol{v}_j\in V.$$

设 $(n-k)$ 阶矩阵 $\boldsymbol{B}$ 的第 j 列是 $\boldsymbol{A}\boldsymbol{v}_i$ 表示为 $\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k}$ 线性组合的系数, 即

$$\boldsymbol{A}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k})=(\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k})\boldsymbol{B}.$$

设非零向量 $\boldsymbol{x}$ 满足 $\boldsymbol{B}\boldsymbol{x}=\lambda\boldsymbol{x}$, 则

$$\boldsymbol{A}(\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k})\boldsymbol{x}=\lambda(\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k})\boldsymbol{x},$$

即非零向量 $(\boldsymbol{v}_1,\dots,\boldsymbol{v}_{n-k})\boldsymbol{x}$ 是 $\boldsymbol{A}$ 关于 λ 的特征向量, 于是 $\lambda\in\mathbb{R}$ 且可选择 $\boldsymbol{x}$ 使得它是实向量, 它和 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_k$ 均正交. 令 $\boldsymbol{e}_{k+1}$ 为该向量的标准化. 归纳可知 $\boldsymbol{A}$ 存在 n 个两两正交的特征向量, 它们构成的正交阵 $\boldsymbol{P}=(\boldsymbol{e}_1,\dots,\boldsymbol{e}_n)$ 满足题述要求. □

由于特征值 λ 对应的实特征向量就是 $\boldsymbol{P}$ 中 λ 对应的那些列向量的线性组合, 因此:

推论

实对称阵的不同特征值对应的实特征向量正交.

练习

(1) 设 $\boldsymbol{\alpha}_1=(1,-3,1)^{\mathrm{T}}, \boldsymbol{\alpha}_2=(1,a,2)^{\mathrm{T}}$ 是实对称阵 $\boldsymbol{A}$ 对应特征值 $\lambda_1=1$ 和 $\lambda_2=2$ 的特征向量, 则 $a=$ __1__.

(2) 若 3 阶实对称阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2=\boldsymbol{A}, \mathrm{R}(\boldsymbol{A})=1$, 则 $\boldsymbol{A}$ 的特征值为 __0, 0, 1__.

例

设 3 阶实对称阵 $\boldsymbol{A}$ 的特征值为 $6,3,3$, 与特征值 6 对应的特征向量为 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}}$, 求 $\boldsymbol{A}$.

解

由于 $\boldsymbol{A}$ 有两个对应特征值 3 的线性无关特征向量, 因此与 $\boldsymbol{\alpha}_1$ 正交的向量都是与特征值 3 对应的特征向量. 由 $\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{x}=0$ 得 $\boldsymbol{\alpha}_2=(-1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(-1,0,1)^{\mathrm{T}}$. 故

$$\boldsymbol{A}=\boldsymbol{P}\mathrm{diag}(6,3,3)\boldsymbol{P}^{-1}=\begin{pmatrix}1&-1&-1\\1&1&0\\1&0&1\end{pmatrix}\begin{pmatrix}6&&\\&3&\\&&3\end{pmatrix}\begin{pmatrix}1&-1&-1\\1&1&0\\1&0&1\end{pmatrix}^{-1}=\begin{pmatrix}4&1&1\\1&4&1\\1&1&4\end{pmatrix}.$$

另解

根据 $\boldsymbol{A}$ 的行和、迹和对称性可设 $\boldsymbol{A}=\begin{pmatrix}a&b&6-a-b\\b&9-a-b&a\\6-a-b&a&3+b\end{pmatrix}$. 再由 $\mathrm{R}(\boldsymbol{A}-3\boldsymbol{E})=1$ 可知 $a=4,b=1$.

对称阵正交合同对角化, 或求正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 将实二次型 f 化为标准形的步骤:

(1) 写出 f 对应的对称阵 $\boldsymbol{A}$.

(2) 求出 $\boldsymbol{A}$ 的特征值.

(3) **若特征值是 k 重的, 求出 k 个特征向量后, 用格拉姆-施密特方法将其正交单位化.**

(4) 这些特征向量构成正交阵 $\boldsymbol{P}$, $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\mathrm{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n)$.

(5) 写出正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 以及对应的实二次型

$$f=\lambda_1 y_1^2+\cdots+\lambda_n y_n^2.$$

例

设 $\boldsymbol{A}=\begin{pmatrix}2&0&0\\0&0&1\\0&1&x\end{pmatrix}, \boldsymbol{B}=\begin{pmatrix}2&0&0\\0&y&0\\0&0&-1\end{pmatrix}$ 相似，求 x,y 以及正交阵 $\boldsymbol{P}$ 使得 $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}=\boldsymbol{B}$.

解

由 $\boldsymbol{A},\boldsymbol{B}$ 相似得 $|\boldsymbol{A}|=-2=|\boldsymbol{B}|=-2y$，$\mathrm{Tr}(\boldsymbol{A})=2+x=\mathrm{Tr}(\boldsymbol{B})=1+y$，故 $x=0,y=1$.

- 对于 $\lambda_1=2$，$\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}0&0&0\\0&-2&1\\0&1&-2\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}\Longrightarrow \boldsymbol{\alpha}_1=\begin{pmatrix}1\\0\\0\end{pmatrix}$.

典型例题：实二次型的对角化

续解

- 对于 $\lambda_2 = 1$, $\boldsymbol{A} - \boldsymbol{E} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \boldsymbol{\alpha}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.
- 对于 $\lambda_3 = -1$, $\boldsymbol{A} + \boldsymbol{E} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \overset{r}{\sim} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \Longrightarrow \boldsymbol{\alpha}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.
- 将特征向量单位化得到 $\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \\ 0 & \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \end{pmatrix}$.

典型例题：实二次型的对角化

例

求正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 化 $f=4x_1^2+4x_2^2+4x_3^2+4x_1x_2+4x_2x_3+4x_3x_1$ 为标准形.

解

- f 对应 $\boldsymbol{A}=\begin{pmatrix}4&2&2\\2&4&2\\2&2&4\end{pmatrix}$. 由 $|\boldsymbol{A}-\lambda\boldsymbol{E}|=-(\lambda-2)^2(\lambda-8)$ 得到特征值 $8,2,2$.
- 对于 $\lambda_1=8$, $\boldsymbol{A}-8\boldsymbol{E}=\begin{pmatrix}-4&2&2\\2&-4&2\\2&2&-4\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&0&-1\\0&1&-1\\0&0&0\end{pmatrix}\Longrightarrow\boldsymbol{\alpha}_1=\begin{pmatrix}1\\1\\1\end{pmatrix}$. 将

  其单位化得到 $\boldsymbol{e}_1=\dfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$.

续解

- 对于 $\lambda_2,\lambda_3=2$, $\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix}2&2&2\\2&2&2\\2&2&2\end{pmatrix}\overset{r}{\sim}\begin{pmatrix}1&1&1\\0&0&0\\0&0&0\end{pmatrix}\Longrightarrow \boldsymbol{\alpha}_2=\begin{pmatrix}-1\\1\\0\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}-1\\0\\1\end{pmatrix}.$

  将其正交单位化得到 $\boldsymbol{\beta}_2=\begin{pmatrix}-1\\1\\0\end{pmatrix},\boldsymbol{e}_2=\frac{1}{\sqrt{2}}\begin{pmatrix}-1\\1\\0\end{pmatrix},$

$$\boldsymbol{\beta}_3=\boldsymbol{\alpha}_3-\frac{[\boldsymbol{\alpha}_3,\boldsymbol{\beta}_2]}{[\boldsymbol{\beta}_2,\boldsymbol{\beta}_2]}\boldsymbol{\beta}_2=\boldsymbol{\alpha}_3-\frac{1}{2}\boldsymbol{\beta}_2=\begin{pmatrix}-1/2\\-1/2\\1\end{pmatrix},\quad \boldsymbol{e}_3=\frac{1}{\sqrt{6}}\begin{pmatrix}-1\\-1\\2\end{pmatrix}.$$

- 因此经过正交变换 $\boldsymbol{x}=\begin{pmatrix}\frac{1}{\sqrt{3}}&-\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&\frac{1}{\sqrt{2}}&-\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{3}}&0&\frac{2}{\sqrt{6}}\end{pmatrix}\boldsymbol{y}$, f 化为标准形 $f=8y_1^2+2y_2^2+2y_3^2$.

例

设实二次型 $f=ax_1^2+x_2^2+x_3^2+4x_1x_2+4x_2x_3+4x_3x_1$ 经过正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 化为 $f=5y_1^2-y_2^2-y_3^2$. 求常数 a 和正交阵 $\boldsymbol{P}$.

解

f 对应 $\boldsymbol{A}=\begin{pmatrix} a & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$, $\mathrm{Tr}(\boldsymbol{A})=a+2=5-1-1, a=1.$

同上例可得 $\boldsymbol{P}=\begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$.

典型例题：实二次型的对角化

例

设实二次型 $f=a(x_1^2+x_2^2+x_3^2)+4x_1x_2+4x_2x_3+4x_3x_1$ 经正交变换化成标准形 $f(y_1,y_2,y_3)=6y_1^2$，则 $a=$ __2__.

练习

设实二次型 $f=x_1^2+2x_2^2+ax_3^2-4x_1x_2-4x_2x_3$ 经过正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 化为 $f=2y_1^2+5y_2^2+by_3^2$. 求常数 a,b 和正交阵 $\boldsymbol{P}$.

答案

$$a=3,b=-1,\boldsymbol{P}=\frac{1}{3}\begin{pmatrix}-2&1&2\\1&-2&2\\2&2&1\end{pmatrix}.$$

例

设二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$. 证明：当 $\|x\|=1$ 时，$f(\boldsymbol{x})$ 的最大（小）值为实对称阵 $\boldsymbol{A}$ 的最大（小）特征值.

证明

将 $\boldsymbol{A}$ 的特征值排序为 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_n$. 存在正交变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 使得

$$\boldsymbol{\Lambda}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}=\operatorname{diag}(\lambda_1,\ldots,\lambda_n),\quad f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{y}=\lambda_1y_1^2+\cdots+\lambda_ny_n^2.$$

由于正交变换保持长度，因此 $\|\boldsymbol{x}\|=1\iff\|\boldsymbol{y}\|=1$.

$$f=\lambda_1y_1^2+\cdots+\lambda_ny_n^2\leqslant\lambda_1(y_1^2+\cdots+y_n^2)=\lambda_1,$$

且等式可在 $\boldsymbol{y}=(1,0,\ldots,0)$ 处取得.

$$f=\lambda_1y_1^2+\cdots+\lambda_ny_n^2\geqslant\lambda_n(y_1^2+\cdots+y_n^2)=\lambda_n,$$

且等式可在 $\boldsymbol{y}=(0,\ldots,0,1)$ 处取得. 故 f 的最大值为 λ_1，最小值为 λ_n. □

第三节　实对称阵的合同

- 惯性指数
- 正定二次型

设 A, B, C 是不全为零的实数. 二次曲线 $Ax^2 + Bxy + Cy^2 = 1$ 左侧的实二次型对应方阵 $\boldsymbol{A} = \begin{pmatrix} A & B/2 \\ B/2 & C \end{pmatrix}$. 由

$$|\boldsymbol{A} - \lambda \boldsymbol{E}| = \begin{vmatrix} A - \lambda & B/2 \\ B/2 & C - \lambda \end{vmatrix} = \lambda^2 - (A + C)\lambda + (AC - B^2/4)$$

可知,

(1) 当 $B^2 - 4AC > 0$ 时, $\boldsymbol{A}$ 特征值一正一负, 从而通过正交变换 $\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{P} \begin{pmatrix} s \\ t \end{pmatrix}$ 可知该曲线为双曲线.

(2) 同理, $B^2 - 4AC < 0$ 时该曲线为椭圆 (或空集);

(3) $B^2 - 4AC = 0$ 时该曲线为两条直线 (若有一次项则为抛物线).

可以看出我们有时候只关心实二次型对应的矩阵的特征值的符号. 由于合同矩阵秩相同, 因此可定义:

定义

称实二次型 f 对应矩阵的秩为 f 的**秩**.

惯性定理

(1) 若 A 和 B 为合同的对角阵, 则 A, B 对角元中正数的个数相同.

(2) 设实二次型 f 的秩为 r. 若可逆线性变换 $x = Py = Qz$ 分别将 f 变为

$$f = k_1 y_1^2 + \cdots + k_r y_r^2, \qquad = \ell_1 z_1^2 + \cdots + \ell_r z_r^2,$$

则 $k_1, \ldots, k_r$ 中正的个数和 $\ell_1, \ldots, \ell_r$ 中正的个数相同.

证明

设

$$\begin{aligned} \boldsymbol{A} &= \operatorname{diag}(k_1, \ldots, k_r, 0, \ldots, 0), \\ \boldsymbol{B} &= \operatorname{diag}(\ell_1, \ldots, \ell_r, 0, \ldots, 0) = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}, \end{aligned}$$

其中 $\boldsymbol{P} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n)$ 为可逆矩阵. 不难得到 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 的对角元:

$$\boldsymbol{B} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P} = \operatorname{diag}(k_1\boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{\alpha}_1, \ldots, k_r\boldsymbol{\alpha}_r^{\mathrm{T}}\boldsymbol{\alpha}_r, 0, \ldots, 0).$$

从而 $\ell_i = k_i\|\boldsymbol{\alpha}_i\|^2$, 二者符号相同. (2) 由(1) 得到. □

定义

把实二次型 f 标准形系数中为正数的个数称为 f 的正惯性指数 p, 为负数的个数称为 f 的负惯性指数 $q = r - q$.

推论

实二次型 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 的正 (负) 惯性指数等于实对称阵 $\boldsymbol{A}$ 的正 (负) 特征值的个数.

定理

任意 n 阶实对称矩阵 $\boldsymbol{A}$ 合同于对角矩阵

$$\begin{pmatrix} \boldsymbol{E}_p & & \\ & -\boldsymbol{E}_q & \\ & & \boldsymbol{O}_{n-p-q} \end{pmatrix},$$

其中 p,q 分别为正负特征值个数 (计算重数), $\mathrm{R}(\boldsymbol{A})=p+q$.

从而正负惯性指数相同的实对称阵是合同的.

推论

n 阶实对称阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 合同 $\iff$ $\boldsymbol{A},\boldsymbol{B}$ 的正负特征值个数均相同.

例

设 $\boldsymbol{A}=\begin{pmatrix}1&2\\2&1\end{pmatrix}$，则 $\boldsymbol{A}$ 合同于(D).

(A) $\begin{pmatrix}-2&1\\1&-2\end{pmatrix}$ (B) $\begin{pmatrix}2&-1\\-1&2\end{pmatrix}$ (C) $\begin{pmatrix}2&1\\1&2\end{pmatrix}$ (D) $\begin{pmatrix}1&-2\\-2&1\end{pmatrix}$

例

矩阵 $\begin{pmatrix}1&1\\1&1\end{pmatrix}$ 与 $\begin{pmatrix}3&0\\0&0\end{pmatrix}$(B).

(A) 合同且相似 (B) 合同但不相似

(C) 不合同但相似 (D) 既不合同也不相似

定义

若实二次型的标准形的系数只在 $-1,0,1$ 三个数中取值，则称此实二次型为**规范形**.

定理

任意一个实二次型总可经过适当的可逆线性变换化为规范形，且规范形是唯一的 (可任意交换变量顺序):

$$f = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2,$$

其中 p,q 分别为正负惯性指数.

例

若实对称矩阵 $\boldsymbol{A}$ 合同于 $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 2 & 0 \end{pmatrix}$，则通过可逆线性变换 $\boldsymbol{x} = \boldsymbol{P}\boldsymbol{y}$ 可将二次型 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 化为规范形 $\underline{\quad y_1^2 + y_2^2 - y_3^2 \quad}$.

例：二次曲面的分类

对于三元实二次型, 正负惯性指数确定了二次曲面的类别.

(1) $p=3, q=0$ 为椭球面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}+\dfrac{z^2}{c^2}=1.$

(2) $p=2, q=1$ 为单叶双曲面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1.$

(3) $p=1, q=2$ 为双叶双曲面 $\dfrac{x^2}{a^2}-\dfrac{y^2}{b^2}-\dfrac{z^2}{c^2}=1.$

(4) $p=2, q=0$ 为椭圆柱面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1.$

(5) $p=1, q=1$ 为双曲柱面 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1.$

定义

设 $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 是二次型.

(1) 若对任意 $\boldsymbol{x}\neq\mathbf{0}$, 都有 $f(\boldsymbol{x})>0$, 则称 f 为正定二次型, 并称实对称阵 $\boldsymbol{A}$ 为正定矩阵, 也记作 $\boldsymbol{A}>0$.

(2) 若对任意 $\boldsymbol{x}$, 都有 $f(\boldsymbol{x})\geqslant 0$, 则称 f 为半正定二次型, 并称实对称阵 $\boldsymbol{A}$ 为半正定矩阵, 也记作 $\boldsymbol{A}\geqslant 0$.

(3) 若 $-f$ (半) 正定, 则称 f 为(半) 负定二次型, 并称实对称阵 $\boldsymbol{A}$ 为(半) 负定矩阵, 也记作 $\boldsymbol{A}<0$ $(\boldsymbol{A}\leqslant 0)$.

(4) 除此之外, 称 f 不定.

可逆线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 不会影响正定性. $\boldsymbol{A}$ 正定 $\Longleftrightarrow$ $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ 正定.

例

(1) $f(x_1,x_2,x_3)=x_1^2+x_2^2$ 半正定.

(2) $f(x_1,x_2,x_3)=x_1^2-2x_2^2+x_3^2$ 不定.

(3) $f(x_1,x_2,x_3)=(x_1+x_2)^2+(x_2+x_3)^2+(x_3+x_1)^2$ 正定.

(4) $f(x_1,x_2,x_3)=(x_1-x_2)^2+(x_2-x_3)^2+(x_3-x_1)^2$ 半正定.

(5) 椭球面 $f(x,y,z)=1$ 对应的二次型正定.

(6) 单叶/双叶双曲面 $f(x,y,z)=1$ 对应的二次型不定.

定理

设 $\boldsymbol{A}$ 是 n 阶实对称阵, $f=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$. 如下命题等价:

(1) $\boldsymbol{A}>0$ 正定, 即 f 正定.

(2) f 的正惯性指数为 n, 即 $\boldsymbol{A}$ 特征值全为正.

(3) 存在正交阵 $\boldsymbol{P}$ 使得 $\boldsymbol{A}=\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}$.

(4) (赫尔维茨定理) $\boldsymbol{A}$ 的各阶顺序主子式都为正, 即

$$a_{11}>0,\quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}>0,\quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}>0,\quad \ldots,\quad |\boldsymbol{A}|>0.$$

将(4)中 $>$ 换成 $\geqslant$ 即可判断半正定, 这也等价于 f 的负惯性指数为 0, 即 $\boldsymbol{A}$ 特征值全非负.

推论

若 $\boldsymbol{A}$ 正定，则 $|\boldsymbol{A}|>0$ 且对角元全为正.

例

若 $f(x_1,x_2,x_3)=2x_1^2+x_2^2+x_3^2+2x_1x_2+tx_2x_3$ 正定，求 t 的取值范围.

解

f 对应 $\boldsymbol{A}=\begin{pmatrix}2&1&0\\1&1&t/2\\0&t/2&1\end{pmatrix}$，顺序主子式

$$2>0,\quad \begin{vmatrix}2&1\\1&1\end{vmatrix}=1>0,\quad |\boldsymbol{A}|=1-\frac{t^2}{2}>0$$

得到 $-\sqrt{2}<t<\sqrt{2}$.

例

求可逆线性变换 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{y}$ 二次型 $f=-5x_1^2+6x_2^2-2x_3^2+6x_1x_2+4x_1x_3$ 化为规范形.

解

$$
\begin{aligned}
f &= -2(x_3-x_1)^2-3x_1^2+6x_2^2+6x_1x_2\\
&= -2(x_3-x_1)^2-3(x_1-x_2)^2+9x_2^2
\end{aligned}
$$

因此 $\boldsymbol{x}=\boldsymbol{y}\boldsymbol{P},\boldsymbol{P}=\begin{pmatrix}-\sqrt{2} & \sqrt{3} & 0\\ 0 & -\sqrt{3} & 3\\ \sqrt{2} & 0 & 0\end{pmatrix}^{-1}=\begin{pmatrix}0 & 0 & \frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{3}}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{pmatrix}$，将 f 化为规范形

$f=-y_1^2-y_2^2+y_3^2$,

例：正定的性质和判定

例

若实对称阵 $\boldsymbol{A}$ 正定，证明 $|\boldsymbol{A}+\boldsymbol{E}|>1$.

证明

由 $\boldsymbol{A}$ 正定可知其特征值均为正，从而 $\boldsymbol{A}+\boldsymbol{E}$ 特征值都大于 1. 从而 $|\boldsymbol{A}+\boldsymbol{E}|>1$. □

例

设 3 阶实对称阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^2+2\boldsymbol{A}=\boldsymbol{O}, \mathrm{R}(\boldsymbol{A})=2$. 当 k 为何值时，矩阵 $\boldsymbol{A}+k\boldsymbol{E}$ 为正定矩阵.

解

由 $\boldsymbol{A}^2+2\boldsymbol{A}=\boldsymbol{O}$ 可知 $\boldsymbol{A}$ 的特征值满足 $\lambda^2+2\lambda=0, \lambda=0,-2$. 由 $\mathrm{R}(\boldsymbol{A})=2$ 可知 $\boldsymbol{A}$ 特征值为 $0,-2,-2$, $\boldsymbol{A}+k\boldsymbol{E}$ 特征值为 $k, k-2, k-2$. 因此 $k>2$.

例

设 $\boldsymbol{A}$ 是 $m\times n$ 实矩阵且 $\mathrm{R}(\boldsymbol{A})=n$. 证明 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 正定.

证明

显然 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是对称的. 注意到

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}=(\boldsymbol{A}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}=\|\boldsymbol{A}\boldsymbol{x}\|^2.$$

由于 $\mathrm{R}(\boldsymbol{A})=n$, $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 只有零解. 因此当 $\boldsymbol{x}\neq\boldsymbol{0}$ 时, $\boldsymbol{A}\boldsymbol{x}\neq\boldsymbol{0}$, 从而

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x}=\|\boldsymbol{A}\boldsymbol{x}\|^2>0.$$

□

例

设 $\boldsymbol{A}$ 为 n 阶实对称矩阵. 证明 $\mathrm{R}(\boldsymbol{A}) = n \iff$ 存在一个 n 阶实方阵 $\boldsymbol{B}$ 使得 $\boldsymbol{AB} + \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ 正定.

证明

显然 $\boldsymbol{AB} + \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ 是对称的, 且

$$\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{AB} + \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A})\boldsymbol{x} = (\boldsymbol{Ax})^{\mathrm{T}}\boldsymbol{Bx} + (\boldsymbol{Bx})^{\mathrm{T}}(\boldsymbol{Ax}) = 2[\boldsymbol{Ax}, \boldsymbol{Bx}].$$

若 $\mathrm{R}(\boldsymbol{A}) = n$, 令 $\boldsymbol{B} = \boldsymbol{A}$, 则当 $\boldsymbol{x} \neq \mathbf{0}$ 时, $\boldsymbol{Ax} \neq \mathbf{0}$, 从而

$$[\boldsymbol{Ax}, \boldsymbol{Bx}] = \|\boldsymbol{Ax}\|^2 > 0.$$

若 $\mathrm{R}(\boldsymbol{A}) < n$, 则存在 $\boldsymbol{x} \neq \mathbf{0}$ 使得 $\boldsymbol{Ax} = \mathbf{0}$, 从而 $[\boldsymbol{Ax}, \boldsymbol{Bx}] = 0$, $\boldsymbol{AB} + \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}$ 不正定. □

例

设 x 是实数，证明 $\begin{vmatrix} 10 & 2 & 3 & 5 \\ 2 & 10 & 1 & -1 \\ 3 & 1 & 10 & x \\ 5 & -1 & x & 10 \end{vmatrix} \leqslant 10000.$

证明

设 $\boldsymbol{A}$ 为题述方阵. 它的前三个顺序主子式

$$10 > 0, \quad \begin{vmatrix} 10 & 2 \\ 2 & 10 \end{vmatrix} = 96 > 0, \quad \begin{vmatrix} 10 & 2 & 3 \\ 2 & 10 & 1 \\ 3 & 1 & 10 \end{vmatrix} = 872 > 0,$$

若 $|\boldsymbol{A}| \leqslant 0$，命题显然成立. 若 $|\boldsymbol{A}| > 0$，则 $\boldsymbol{A}$ 正定，从而特征值全正. 因此

$$|\boldsymbol{A}| = \lambda_1\lambda_2\lambda_3\lambda_4 \leqslant \left(\frac{\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4}{4}\right)^4 = 10000. \qquad \square$$

正定的应用：函数极值

实对称阵可用于判断多元函数的极值. 设 $f(\boldsymbol{x}) = f(x_1, \ldots, x_n)$ 是一个 n 元实函数, $\boldsymbol{a}$ 是其定义域内一点, 且 f 在 $\boldsymbol{a}$ 附近具有连续的二阶偏导数. 记 $f''_{ij} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$, 则 $f''_{ij} = f''_{ji}$. 于是 $\boldsymbol{A} = (f''_{ij}(\boldsymbol{a}))$ 是 n 阶实对称阵.

定理

设 f 在 $\boldsymbol{a}$ 处各阶偏导均为零.

(1) 若 $\boldsymbol{A}$ 正定, 则 f 在 $\boldsymbol{a}$ 处取极小值;

(2) 若 $\boldsymbol{A}$ 负定, 则 f 在 $\boldsymbol{a}$ 处取极大值.

若 $\boldsymbol{A}$ 不定, 则无法判断 $\boldsymbol{a}$ 是否是极值点.

对于一般的 $m\times n$ 实矩阵 $\boldsymbol{A}$, 我们有奇异值分解

$$\boldsymbol{A}=\boldsymbol{U}_{m\times m}\boldsymbol{\Sigma}_{m\times n}\boldsymbol{V}_{n\times n}^{\mathrm{T}},$$

其中 $\boldsymbol{U},\boldsymbol{V}$ 分别是 m,n 阶正交阵, $\boldsymbol{\Sigma}$ 是 $m\times n$ 型对角阵, 对角元非负且按降序排列.

首先对 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 这一半正定对称阵做正交合同对角化

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{V}^{\mathrm{T}}\boldsymbol{\Lambda}\boldsymbol{V},\quad \boldsymbol{V}=(\boldsymbol{v}_1,\ldots,\boldsymbol{v}_n),\quad \boldsymbol{\Lambda}=\mathrm{diag}(\lambda_1,\ldots,\lambda_n),\lambda_1\geqslant\cdots\geqslant\lambda_n\geqslant 0.$$

奇异值是指 $\sigma_i=\sqrt{\lambda_i}$. 令 $\boldsymbol{\Sigma}\in M_{m\times n}$ 为对角阵, 对角元为 $\sigma_1,\ldots,\sigma_n$.

对于 $1\leqslant j\leqslant r=\mathrm{R}(\boldsymbol{A})$, 令 $\boldsymbol{u}_j=\dfrac{1}{\sigma_j}\boldsymbol{A}\boldsymbol{v}_j$, 设 $\boldsymbol{u}_{r+1},\ldots,\boldsymbol{u}_m$ 是 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}$ 的一组标准正交基础解系, 则 $\boldsymbol{U}=(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_m)$ 是正交阵, 且 $\boldsymbol{A}=\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}$.

注意到, 如果 $\boldsymbol{U}'$ 是 $\boldsymbol{U}$ 的前 r 列, $\boldsymbol{V}'$ 是 $\boldsymbol{V}$ 的前 r 行, $\boldsymbol{\Sigma}'$ 是 $\boldsymbol{\Sigma}$ 的前 r 行 r 列, 则

$$\boldsymbol{A} = \boldsymbol{U}'_{m\times r}\boldsymbol{\Sigma}'_{r\times r}\boldsymbol{V}'_{r\times n},$$

其中 $\boldsymbol{\Sigma}'$ 是 $\boldsymbol{A}$ 奇异值降序的对角阵, $\boldsymbol{U}', \boldsymbol{V}'$ 为列/行为标准向量且两两正交的矩阵.

这意味着当 r 相比 m, n 较小时, 只需存储 $(m+n+1)r$ 个元素即可还原 $\boldsymbol{A}$. 这是一种无损压缩算法. 如果我们只截取前 $k < r$ 个奇异值以及对应的 $\boldsymbol{U}, \boldsymbol{V}$ 部分, 则可以对 $\boldsymbol{A}$ 进行有损压缩到 $\boldsymbol{A}'$. 例如 $\boldsymbol{A}$ 表示一张图像的像素信息, 保留它较大的奇异值往往对它的信息影响很小. 有时候, 我们甚至需要主动舍弃较小的奇异值, 只保留较大的奇异值来实现信号降噪.

矩阵还有诸如 LU 分解, QR 分解, 科列斯基分解等. 这些分解往往都在压缩或降噪中发挥着作用.

当 $m = n = 3$ 时, 可以看出线性变换可以分解为旋转、放缩、旋转的复合.