线性代数讲义

冉诗勇

简介

本讲义用于线性代数教学，配合本校所使用的教材同济版《工程数学–线性代数》，根据本校教学实际 (32 学时) 对讲授内容进行了取舍和增补.

对使用讲义的同学的要求: 通过课堂听课或自学教材完成讲义的思考题和例题练习, 积极进行课堂演算, 及时依据讲义做好笔记 (动笔而不要用手机拍拍照) 和知识梳理, 以加深对概念和知识点的理解; 课外独立完成所布置的作业以及时巩固所学知识.

目录

第 零 章　引言

即得易见平凡, 仿照上例显然. 留作习题答案略, 读者自证不难.

反之亦然同理, 推论自然成立, 略去过程 QED, 由上可知证毕.

Mario Li (网名) «西江月·证明»

线性代数, 顾名思义, 是研究线性问题的代数理论. 线性问题遍及数学乃至科学与工程的几乎所有领域. 为什么会这样? 这是因为我们在处理非线性问题 (真实世界多是非线性的, 譬如人的听觉系统对声强的感觉是非线性的) 时, 优先考虑的是能否将之线性处理或线性近似. 为了对线性运算及其所引出的种种结构进行有效的描述, 线性代数需要引入三个基本工具: 矩阵、向量与行列式. 为描述这些结构, 需要进一步进行抽象, 引入线性空间、秩等概念.

联想一下微积分里面的以直代曲操作、物理学小角度单摆运动分析的近似处理, 就是用线性去近似非线性.

有必要对代数和线性两个术语咬文嚼字一下. 问两个问题: 什么是代数? 什么是线性?

0.1　什么是代数

代数英文是 algebra, 源于 9 世纪阿拉伯数学家阿尔·花拉子米的一本著作的书名 (书名内有 al-jabr 一词, 本意是 "重新组合", 意指方程两端的移项和同类项合并) , 该书主要涉及初等代数问题. 而后该书被译为拉丁文传入西方, "al-jabr" 译为拉丁文 "aljebra", 英文译作 "algebra". 中文 "代数"(1859 年由李善兰翻译) 意指 "以字母等符号代替数". 在小学阶段我们接触的是算术. 我们知道 2 个苹果加 1 个苹果等于 3 个苹果; 3 个梨子加 2 个梨子等于 5 个梨子. 算术进行了初步的抽象, 去除苹果和梨子的区别, 只留下数的运算. 在初中阶段, 我们将数字用字母表示, 是进一步的抽象, 这就是初等代数. 其目的在于研究更一般的数的运算规律, 如两数之和的平方的普遍表达式、一元二次方程的解等. 同样的道理, 在线性代数 (属于高等代数范畴, 因此数学系里这门课的名字叫高等代数) 里, 我们又进行了进一步的抽象, 提出了**线性空间** (又称**向量空间**, 是满足 "加法" 和 "数乘" 等八条公理的元素的集合) 这一抽象概念, 其元素称为**向量**. 针对所有的线性空间, 线性代数研究其普遍的规律和性质.

线性空间是一个非常抽象的概念. 要将结论普适化, 抽象是需要付出的代价, 相应地也会带来理解上的困难. 线性空间对于加法和数乘满足封闭性, 即对该空间的向量进行加法和数乘操作后得到的向量仍然在集合内部. 可以用简单的二维情况去理解这一概念. 所有二维平面向量组成的集合称为二维向量空间, 在该空间内, 任何向量经加法和数乘运算后得到的向量仍属于该集合, 即对这两种运算是封闭的. 如果向量是 n 维的, 则称为 n 维向量空间.

0.2 什么是线性

数学问题可分为两类: 一类线性问题, 一类非线性问题. 线性问题是研究最久、理论最完善的; 而非线性问题则可以在一定基础上转化为线性问题求解. 因此遇到一个具体的问题后, 我们通常考虑的是它是线性还是非线性的. 如果是线性问题如何处理, 如果是非线性问题则能否转化为线性问题.

有一个烧开水的段子形象地说明了这种数学思维. 问: 如何用一个空水壶、燃气灶烧一壶开水? 答: 将水壶灌满水, 放到燃气灶上, 点燃燃气灶, 把一满壶水烧开. 再问: 水壶里已有半壶水, 如何烧一壶开水? 与一般回答不同, 数学家的回答是: 把装了半壶水的水壶倒空, 就化归为刚才已解决的问题了.

在中学的初等数学里, 我们知道 $ax+by=c(a,b,c$为常数$)$ 方程在平面上的图形是一条直线, 此时 y,x 之间是**线性关系**, 该方程称为**线性方程**.

将之推广, n 个变量 $(x_1,x_2,\cdots,x_n)$ 的一次方程

$$k_1x_1+k_2x_2+\cdots+k_nx_n=b \tag{0-1}$$

称为**线性方程**$(k_1,k_2,\cdots,k_n,b$ 是常数$)$. 此时各个变量之间呈**线性关系**.

在线性代数里, 形如 $L(x)=k_1x_1+k_2x_2+\cdots+k_nx_n$ 的函数称为**线性函数**. 线性函数满足两个标准: **可加性**和**齐次性**.

1) 可加性. 即有

$$L\left(x_1+x_2+\cdots+x_n\right)=L\left(x_1\right)+L\left(x_2\right)+\cdots+L\left(x_n\right)$$

2) 齐次性: 也叫做**比例性**、**数乘性**. 即有

$$L(kx)=kL(x)\quad \text{其中}k\text{是常数.}$$

按照上述标准, 形如 $f(x)=kx+b(b\neq 0)$ 的函数就不能称为线性函数了.

思考 动手研究一下为什么不满足可加性和齐次性标准.

如果将变量 $x_1,x_2,\cdots,x_n$ 的组合视为一个向量 $\boldsymbol{x}$, 记作**列矩阵**(又称为**列向量**) 形式

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

相应的系数组合用**行矩阵**(又称为**行向量**) 形式 $\boldsymbol{A}=(k_1,k_2,\cdots,k_n)$ 表示, 则该线性函数可简化表示为 $\boldsymbol{Ax}$, 我们可表述为对向量 $\boldsymbol{x}$ 进行了一次**线性变换**, 得到了另一个向量 $\boldsymbol{Ax}$, 这一过程又称为**线性映射**, 记为 $\boldsymbol{x}\mapsto\boldsymbol{Ax}$.

本段表述所用的概念如矩阵、向量超前了. 可先放在一边, 学到相关概念之后再体会. 只要知道这里的变换称为线性变换即可.

为什么上述变换可称为**线性**变换? 原因在于其满足可加性和齐次性两个标准. 即有

$$\boldsymbol{A}(\boldsymbol{x}_1+\boldsymbol{x}_2)=\boldsymbol{A}\boldsymbol{x}_1+\boldsymbol{A}\boldsymbol{x}_2$$

和

$$\boldsymbol{A}(k\boldsymbol{x}) = k\boldsymbol{A}\boldsymbol{x}$$

形式的可加性和齐次性规律.

一句话总结, 线性代数主要研究的就是上述形式的**线性变换**, 这是线性代数的精华部分.

举一个例子说明线性变换.

考虑二维平面的一个向量 $\boldsymbol{a} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, 经过矩阵 $\boldsymbol{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ 的线性变换后, 有

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2\cos\theta - \sin\theta \\ 2\sin\theta + \cos\theta \end{pmatrix}$$

线性变换后得到的向量 $\boldsymbol{b} = \begin{pmatrix} 2\cos\theta - \sin\theta \\ 2\sin\theta + \cos\theta \end{pmatrix}$, 如图0-1所示. 由图可知, 该线性变换的结果是将向量 $\boldsymbol{a}$ 逆时针方向旋转了 θ 角度. 数学上我们称之为 $\mathbb{R}^2$ 向量空间的 $\boldsymbol{a}$ 向量经过旋转线性映射后变为了 $\mathbb{R}^2$ 向量空间的 $\boldsymbol{b}$ 向量. 变换类似于物理学中位置矢量的变化, 即向量在运动 (在物理学中相当于质点在二维平面运动).

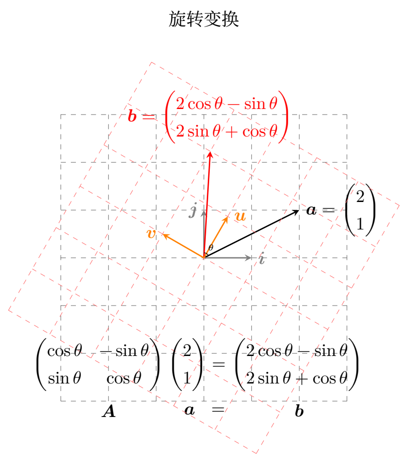


图 0-1 将向量旋转 θ 角度的线性变换.

初学者可能会觉得解线性方程组中学学过, 不就是消元法么? 并非如此简单, 中学所接触的是一种简单的线性方程组, 体现在方程个数较少 (两到三个), 解唯一. 但我们更常遇到方程组包含多个方程, 还可能有无解、无穷多解等情况.

0.3 线性代数课程的内容

对于非数学专业的学习者而言, 线性代数的核心在于求解线性方程组, 为此引入了矩阵、向量和行列式. 简述如下:

1) 求解线性方程组时所用的高斯消元法解法中, 参与计算的只有方程的系数和常数, 变量的记号不重要. 因此引入了矩阵的概念.
2) 消元过程等价于对矩阵作初等行变换. 这就涉及了矩阵的运算, 为此需要定义矩阵的加法、数乘和矩阵的乘法, 并了解其运算规律.
3) 线性方程组的求解等价于向量组的线性表示. 从向量的角度去理解线性方程组的求解, 需要研究向量的线性相关问题.
4) 对于某一类线性方程组, 行列式可简洁表述其解的规律. 为此, 需要掌握行列式的定义、性质以及计算方法.

总而言之, 围绕求解线性方程组, 本课程将学习行列式、矩阵及其运算、矩阵的初等变换以及向量组的线性相关等内容.

在学习时, 一定要将概念和知识点融汇贯通起来, 认识到同一个问题可以有不同的表述和理解方式. 举例来说, 求解三元一次方程组时, 其解可从向量角度用向量组的线性组合表示, 也可从几何角度去理解其解的规律.

0.4 如何学习线性代数

线性代数课程有以下两个特点:

1) 概念、定理多, 内容**非常抽象**, 区别于微积分等基础课. 在学习具体内容时, 不易理解和掌握, 解题时往往只会套用而不知道为什么要这样做.
2) 计算繁琐, 机械重复. 例如, 计算问题涉及到初等行变换的, 就有求逆矩阵、矩阵的秩, 向量组的秩和基础解系等等. 一般而言, 阶数大于 4 的高阶矩阵问题, 笔算非常繁琐, 极易犯错.

针对这些特点, 有如下一些建议:

1) 对概念、定理以及它们之间的联系的准确理解十分重要, 否则极有可能出现学完只掌握了一堆行列式和矩阵的计算技巧的情况. 即使考试能拿高分, 也并不意味着学会了线性代数.
2) **利用具体的问题和简单的情况去理解抽象的概念和定理**. 譬如在行列式的性质学习中, 可以以二阶行列式为例去验证各个性质, 以理解、掌握并应用这些性质. 至于各个性质的具体的证明, 如个人感兴趣可自学; 如无充分时间和兴趣, 则可忽略相关的证明, 将重点放在理解和应用上面.
3) 充分的习题训练是必不可少的一环, 原则上多多益善. 演算过程务必仔细, 宁慢勿快, 避免犯低级错误. 以行列式或矩阵的数值计算为例, 一般只要求掌握四阶及以下情况的计算, 目的在于理解、巩固所学的概念和定理. 更高阶的行列式或矩阵的数值运算, 我们一般都交给计算机完成 (不用练习); 而如果在习题中遇到高阶行列式的计算, 多以字母符号给出, 一般需要用一些技巧而不要硬算 (这违背练习的目的).
4) 做好笔记和课堂即时练习, 对课堂讲授的一些典型题型记下来并自己总结.
5) 充分利用网络资源学习, 如网课、各种概念的科普视频. 一些值得推荐的资源包括 3Blue1Brown 线性代数科普、MIT 的线性代数教学视频等.
6) 课外的预习和自学是必须的, 所花的时间至少是课堂听课时间的两倍.
7) 作业独立完成, 不要抄答案. 数学是一门诚实不会欺骗你的学科, 不会就是不会 (双关语).

第 一 章 行列式

有那么两种数学书: 第一种你看了第一页就不想看了, 第二种是你看了第一句话就不想看了.

杨振宁 (1922–)

本章关注以下几个问题.

1) 为什么要引入行列式?

2) 行列式如何定义?

3) 行列式有哪些性质?

4) 如何计算行列式, 有哪些常用的方法?

在学习时, 请注意行列式是一个辅助工具, 其相关的理论和概念并不多, 重点在于利用其性质和定义计算行列式. 初学者碰到的第一个坎是行列式奇怪的定义方式. 为什么要引入逆序数这么别扭、违背直觉的概念来定义行列式? 定义式为何如此不直观, 以至于在一般情况下根本就没有将之应用于计算行列式的价值 (一般不要通过算高阶行列式各项的逆序数来计算行列式)? 请先接受这种定义并适应它 (形式上没有更好更简洁的办法去定义行列式了). 为回避这种别扭的定义, 本章在行列式的定义一节还提供了另外一种基于代数余子式的递归定义, 应该更符合我们的直觉, 也更有应用于计算行列式的价值.

1.1 二阶和三阶行列式

1.1.1 二阶行列式

行列式的引入可以方便地**表达线性方程组解**的规律. 行列式的英文单词为 determinant, 隐含有 “决定条件” 的意思, 行列式的值不为 0 也正是决定 n 个 n 元线性方程组有唯一解的条件. 我们先来看下面的二元线性方程组:

$$\begin{cases} a_{11}x + a_{12}y = b_1 \\ a_{21}x + a_{22}y = b_2 \end{cases},$$

利用消元法, 可求得它的解为 (设分母不为零):

$$\begin{cases} x = \dfrac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}} \\ y = \dfrac{a_{11} b_2 - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}} \end{cases}.$$

定义 1.1: 二阶行列式

主对角线 副对角线 元素

行标 列标

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

$-$ $+$

其中, $a_{ij}(i, j = 1, 2)$ 称为**元素**, i 称为**行标**, j 称为**列标**.

请始终记住行列式是一个值, 对之作出如此符号约定只是一种形式. 对于一阶行列式, $|-2| = -2$.

练习 1.1 《孙子算经》有题: 今有雉兔同笼, 上有三十五头, 下有九十四足, 问雉兔各几何? 请用二阶行列式秒之.

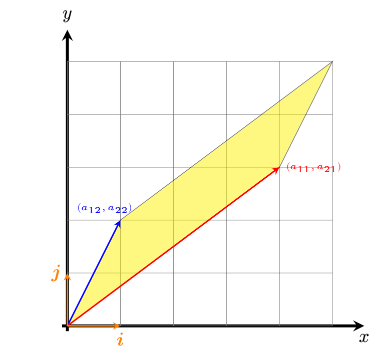


图 1-1

思考 如果将行列式里的第一列值 a_{11}, a_{21} 视为 XY 平面上的一个向量 $\boldsymbol{A}_1$ 的坐标 (a_{11}, a_{21}), 第二列值 a_{12}, a_{22} 视为另一个向量 $\boldsymbol{A}_2$ 的坐标 (a_{12}, a_{22})(图1-1), 则可以将二阶行列式的绝对值理解为由这两个向量构成的平行四边形的面积, 请验证并证明之.

练习 1.2 仿照二阶行列式用于表示二元线性方程组解的方法, 写出三元线性方程组的三阶行列式形式的表述形式.

方程的解可用行列式表示为:

$$x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{D_1}{D}, \qquad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} = \frac{D_2}{D}.$$

1.1.2 三阶行列式

类似地, 对于一般的三元线性方程组:

$$\begin{cases} a_{11}x + a_{12}y + a_{13}z = b_1 \\ a_{21}x + a_{22}y + a_{23}z = b_2 \\ a_{31}x + a_{32}y + a_{33}z = b_3 \end{cases},$$

利用消元法, 可求得它的解为 (设分母不为零):

$$\begin{cases} x = \dfrac{b_1a_{22}a_{33} + a_{12}a_{23}b_3 + a_{13}b_2a_{32} - b_1a_{23}a_{32} - a_{12}b_2a_{33} - a_{13}a_{22}b_3}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}} \\ y = \dfrac{a_{11}b_2a_{33} + b_1a_{23}a_{31} + a_{13}a_{21}b_3 - a_{11}a_{23}b_3 - b_1a_{21}a_{33} - a_{13}b_2a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}} \\ z = \dfrac{a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}} \end{cases}$$

定义 1.2: 三阶行列式

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

三阶行列式计算的辅助记忆方法: 沙路法

$$\begin{array}{ccccc} + & + & + & & \\ \left|\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}\right. & \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} \\ - & - & - & & \end{array}$$

思考　四阶及以上的行列式还能用三阶行列式那样的辅助图画法进行计算吗?

例 1.1　求下列三阶行列式的值.

$$(1)\begin{vmatrix} 1 & 2 & 3 \\ 8 & 0 & 4 \\ 7 & 6 & 5 \end{vmatrix};\quad (2)\begin{vmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \\ 2 & 7 & 8 \end{vmatrix}.$$

1.2　全排列和对换

从三阶行列式的定义式可以看到以下特点:

1) 如果把每一项元素的**行标**按 $1,2,3$ 依次排列, 则每项元素的**列标**排列分别是 $123,231,312,132,213,321$, 恰好是 $1,2,3$ 这三个数的所有可能排列;

2) 共有 $3!=6$ 项求和, 且每一项是**不同行不同列**的三个元素乘积;

3) 6 项中有三项符号为正, 三项符号为负.

以上三个特点不是偶然的, 当对此进行深入讨论后, 就可以给出一般 n 阶行列式的定义. 为此我们首先定义几个术语.

定义 1.3: 全排列

把 n 个不同的元素排成一列, 称为这 n 个元素的**全排列**(简称**排列**). n 个不同元素的所有排列的种数用 p_n 表示. 有

$$p_n = n\cdot(n-1)\cdot\cdots\cdot 3\cdot 2\cdot 1 = n!$$

定义 1.4: 逆序、排列的逆序数、奇排列、偶排列、对换

对于 n 个不同的元素, 先规定各元素之间有一个标准次序. 在 $n!$ 个不同的全排列中, 规定某一个排列为标准顺序的排列, 即在各元素中规定一个**标准顺序**(对自然数, 一般规定**从小到大**的顺序为标准顺序). 在任一排列中, 当两个元素的次序与标准顺序不一致时, 就说有一个**逆序**. 一个排列中的所有逆序的总数称为这个排列的**逆序数**.

逆序数的计算方法如下:

设有 n 个自然数, 记为 $p_i(i=1,2,\cdots,n)$, $p_1p_2\cdots p_n$ 为这 n 个数的一个排列. 则对每个 $p_i(i=1,2,\cdots,n)$, 如果比 p_i 大且排在 p_i 前

面的元素个数有 t_i 个, 就称 p_i 在这个排列中的逆序数是 t_i, 而

$$t = t_1 + t_2 + \cdots + t_n = \sum_{i=1}^{n} t_i$$

就是这个排列的**逆序数**.

逆序数为奇数的排列称为**奇排列**, 逆序数为偶数的排列称为**偶排列**.

在排列中, 将两个元素对调位置, 其余元素不动, 这种作出新排列的操作称为**对换**.

定理 1.1

一个排列中, 任意两个元素**对换**, 排列**改变奇偶性**.

推论 1.1

由于标准顺序的排列逆序数是零, 由 **定理 1.1** , 有:

奇排列调成标准顺序排列的对换次数是奇数, 偶排列调成标准顺序排列的对换次数是偶数.

再次观察三阶行列式值的表达形式, 由于排列 $123, 231, 312, 132, 213, 321$ 的逆序数分别是 $0, 2, 2, 1, 1, 3$, 可以将之用逆序数表达成如下形式:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_{p_1p_2p_3} (-1)^t a_{1p_1} a_{2p_2} a_{3p_3}.$$

其中 p_1, p_2, p_3 是 $1, 2, 3$ 三个数的一个排列, t 是这个排列的逆序数, 共 $3! = 6$ 项求和. 推广这种定义方法, 下节给出一般 n 阶行列式的定义.

1.3 n 阶行列式的定义

1.3.1 行列式的组合定义

定义 1.5: n 阶行列式的组合定义

由 n 行 n 列共 n^2 个元素 $a_{ij}(i, j = 1, \cdots, n)$ 组成的多项式, 其值可利用逆序数有如下组合定义:

主对角线 副对角线

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum (-1)^t a_{1p_1} a_{2p_2} \cdots a_{np_n},$$

其中 a_{ij} 是第 i 行第 j 列的数 (称为**元素**). 在每一行中取出一个数, 并且要求取出的 n 个数都**在不同的列上**. 把所取元素的行数按自然顺序排列, 相应的列数设为 $p_1p_2\cdots p_n$. 作出这 n 个元素的乘积 $a_{1p_1}a_{2p_2}\cdots a_{np_n}$, 然后按第二个下标排列 $(p_1p_2\cdots p_n)$ 的奇偶性, 将乘积乘以 $+1$ 或 -1 . 如 $(p_1p_2\cdots p_n)$ 是偶排列则乘以 $+1$, 否则乘以 -1. 如把 $(p_1p_2\cdots p_n)$ 的逆序数记作 t, 则各乘积所带的符号为 $(-1)^t$, 然后对所有 n 级排列 $(p_1p_2\cdots p_n)$ 求和, 这个和数称为 n 阶行列式, 记作 $\det(a_{ij})$

利用行列式的组合定义我们来研究几个特殊行列式的计算. 先定义几个特殊的行列式.

定义 1.6: 单位行列式

主对角线上元素为1其他元素为0的行列式, 其值为 1. 如: 二阶单位行列式 $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$; 三阶单位行列式 $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$.

定义 1.7: 上三角行列式、下三角行列式、对角行列式

主对角线以下的元素均为 0 (即当 $i > j$ 时, $a_{ij} = 0$), 这种行列式称为**上三角行列式**.

$$\begin{vmatrix} a_{11} & & & & \\ & a_{22} & & * & \\ & & a_{33} & & \\ & 0 & & \ddots & \\ & & & & a_{nn} \end{vmatrix}$$

主对角线以上的元素均为 0 (即当 $i < j$ 时, $a_{ij} = 0$) 的行列式称为**下三角行列式**.

$$\begin{vmatrix} a_{11} & & & & \\ & a_{22} & & 0 & \\ & & a_{33} & & \\ & * & & \ddots & \\ & & & & a_{nn} \end{vmatrix}$$

主对角线之外都为零的行列式称为**对角行列式**.

$$\begin{vmatrix} a_{11} & & & & \\ & a_{22} & & \text{\Large 0} & \\ & & a_{33} & & \\ & \text{\Large 0} & & \ddots & \\ & & & & a_{nn} \end{vmatrix}$$

以**下三角行列式**为例探究其值的计算. 根据行列式的组合定义, n 阶下三角行列式的展开式有 $n!$ 项, 但在每一项中的乘积 $a_{1p_1}a_{2p_2}\cdots a_{np_n}$ 中只要有一个元素等于 0，乘积就为 0． 所以只要计算展开式中不明显为 0 的项. 由于第 1 行除了 a_{11} 外都为 0，故只须考虑 $p_1 = 1$ 的项. 第 2 行中除了 a_{21}, a_{22} 外都是 0，现已取 $p_1 = 1$, 所以必须取 $p_2 = 2$. 依次按行推理可推知第 n 行的元素只能选取 $p_n = n$. 即行列式中不为 0 的乘积只可能是主对角线上各元素的乘积 $a_{11}a_{22}\cdots a_{nn}$, 而排列 $12\cdots n$ 的逆序数为 0，所以这一项所带的符号是正的, 因此该行列式等于 $a_{11}a_{22}\cdots a_{nn}$. 即

$$\begin{vmatrix} a_{11} & & & & \\ & a_{22} & & \text{\Large 0} & \\ & & a_{33} & & \\ & \text{\Large $*$} & & \ddots & \\ & & & & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdot\cdots\cdot a_{nn}.$$

同理可得上三角行列式和对角行列式的值, 得到如下推论.

推论 1.2

上三角、下三角和对角行列式的值等于**主对角线元素的乘积**.

1.3.2　行列式的递归定义

再次回顾三阶行列式值的表达形式, 发现可以将之写作二阶行列式的组合形式

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

其中 M_{1j} 表示在三阶行列式中**划去**元素 a_{1j} 所在的第 1 行元素与第 j 列元素后, 剩下的元素按原有的先后次序排成的 2 阶的行列式. 规定 $A_{1j}=(-1)^{1+j}M_{ij}$.

依此规律, 可写出 n 阶行列式的递归定义. 先对**余子式**和**代数余子式**作出定义.

定义 1.8: 余子式、代数余子式

假设 D 为一个 n 阶行列式:

- **余子式** M_{ij} 是将 D 的第 i 行和第 j 列去掉之后得到的 $n-1$ 阶行列式.
- **代数余子式** $A_{ij}=(-1)^{i+j}M_{ij}$.

以 5 阶行列式为例说明, 元素 a_{12} 的余子式和代数余子式如下所示.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}$$

$$\begin{vmatrix} a_{21} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{53} & a_{54} & a_{55} \end{vmatrix}$$

余子式M_{ij}

$$(-1)^{1+2}\begin{vmatrix} a_{21} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{53} & a_{54} & a_{55} \end{vmatrix}$$

代数余子式 $A_{ij}=(-1)^{i+j}M_{ij}$

定义 1.9: n 阶行列式的递归定义

由 n 行 n 列共 n^2 个元素 $a_{ij}(i,j=1,\cdots,n)$ 组成的多项式

主对角线　　副对角线

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

记为 $D=\det(a_{ij})$. 当 $n=1$ 时, 其值为 $D=|a_{11}|=a_{11}$; 当 $n\geq 2$

时, 其值为

$$D = a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n} = \sum_{j=1}^{n} a_{1j}A_{1j}$$

其中,

$$A_{1j} = (-1)^{1+j}M_{1j}$$

M_{1j} 表示在 D 中**划去**元素 a_{1j} 所在的第 1 行元素与第 j 列元素后, 剩下的元素按原有的先后次序排成的 $n-1$ 阶的余子行列式, 即

$$M_{1j} = \begin{vmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3,j-1} & a_{3,j+1} & \cdots & a_{3n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}, j = 1, 2, \cdots, n$$

行列式的递归定义可将高阶行列式的计算转化为低阶行列式的计算, 相比用组合定义计算具有明显的优越性. 在高阶行列式的计算中, 将行列式降阶后计算是最常用的方法之一.

1.4　行列式的性质

行列式有如下几种性质.

性质 1.1: 转置不变性

行列式转置之后值不变, 即

$$D^{\mathrm{T}} = D$$

该性质说明行列式的**行和列具有同等的地位**, **凡是对行成立的性质, 对列也一样成立**. 反之亦然.

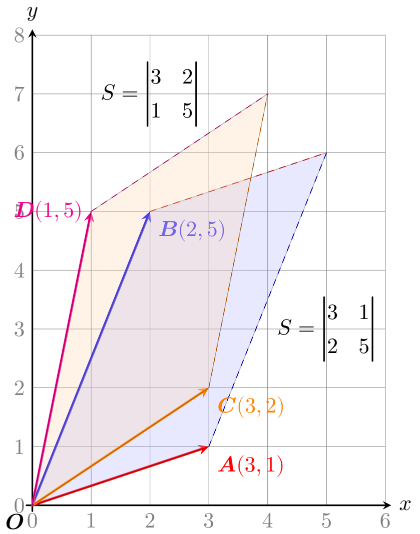


图 1-2　性质 1.1 的几何意义.

思考　如何理解互换变号性的几何意义?

定义 1.10: 转置行列式

对一个行列式 D, 将它的行和列交换, 即将所有 a_{ij} 和 a_{ji} 元素交换, 得到的新行列式, 称为行列式 D 的**转置行列式**, 记为 D^{T}.

性质 1.2: 互换变号性

交换两行 (列) 后, 值**变号**.

根据该性质可有如下推论.

推论 1.3

如果行列式中有**两行 (列) 相同**, 则此行列式的值等于零.

证明:　因为互换行列式 D 中的两个相同的行 (列), 其结果仍是

D, 但由 **性质 1.2** 可知其结果为 $-D$, 因此 $D=-D$, 所以 $D=0$.

□

≡ 性质 1.3: 数乘性

某行 (列) 乘 k 倍, 值变 k 倍.

› 推论 1.4

行列式某一行 (列) 的公因子可以提到行列式外边.

› 推论 1.5

若行列式其中某行 (列) 所有元素都为零, 则它的值为零.

≡ 性质 1.4: 成比例为零性

若行列式其中两行 (列) 对应元素成比例, 则它的值为零.

≡ 性质 1.5: 可加性

若行列式 D_1 和 D_2 在某行 (列) 之外的其他元素都相同, 则两者之和等于另一个行列式 D. 其中 D 在该行 (列) 的元素等于 D_1 和 D_2 对应元素之和, 其他位置的元素与 D_1 和 D_2 相同. 即: 两式仅一行 (列) 不同可相加.

≡ 性质 1.6: 倍加不变性

行列式的某行 (列) 的每个元素都加上另一行 (列) 对应位置元素的 k 倍, 它的值不变.

行列式有交换、数乘、相加三种初等运算, 我们约定如下的符号表示三种变换. 对于行的运算:

1) $r_i \leftrightarrow r_j$ 表示第 i 行和第 j 行交换
2) $r_i \times k$ 表示第 i 行乘以 k 倍
3) $r_i + kr_j$ 表示第 i 行加上第 j 行的 k 倍

对于列的运算:

1) $c_i \leftrightarrow c_j$ 表示第 i 列和第 j 列交换
2) $c_i \times k$ 表示第 i 列乘以 k 倍
3) $c_i + kc_j$ 表示第 i 列加上第 j 列的 k 倍

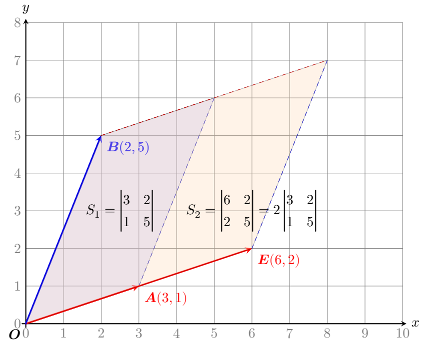


图 1-3 性质 1.3 的几何意义.

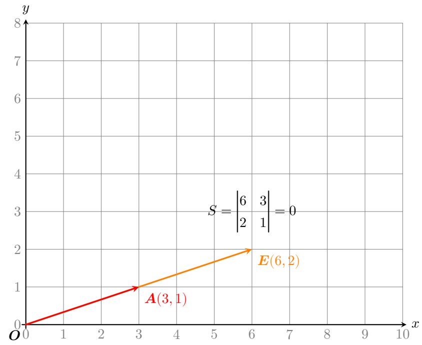


图 1-4 性质 1.4 的几何意义.

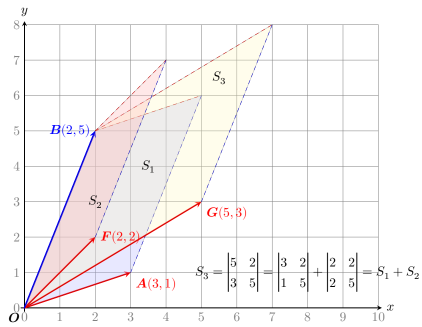


图 1-5 性质 1.5 的几何意义.

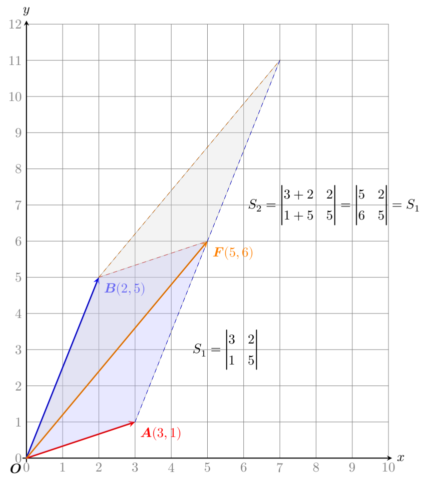


图 1-6 性质 1.6 的几何意义.

🔔 性质 **1.6** 用于计算行列式非常有用.

★ **例子**

$$\begin{vmatrix}1&2&3\\4&5&6\\7&8&9\end{vmatrix}\xlongequal{r_3-r_2}\begin{vmatrix}1&2&3\\4&5&6\\3&3&3\end{vmatrix}$$

$$\xlongequal{c_1\leftrightarrow c_2}-\begin{vmatrix}2&1&3\\5&4&6\\3&3&3\end{vmatrix}\xlongequal{r_3\div 3}-3\begin{vmatrix}2&1&3\\5&4&6\\1&1&1\end{vmatrix}$$

性质 1.7: 行列式展开性质

1) 行列式等于它的任一行 (列) 的各元素与其对应的代数余子式乘积之和.
2) 行列式某一行 (列) 的元素与另一行 (列) 的对应元素的代数余子式乘积之和等于零.

即

$$\sum_{k=1}^{n}a_{ki}A_{kj}=\begin{cases}D, & 当 i=j,\\ 0, & 当 i\neq j;\end{cases}$$

或

$$\sum_{k=1}^{n}a_{ik}A_{jk}=\begin{cases}D, & 当 i=j,\\ 0, & 当 i\neq j;\end{cases}$$

这个性质叫做行列式按行 (列) 展开法则. 利用这一法则可以将高阶行列式降阶, 简化行列式的计算, 务必掌握.

1.5　行列式的计算

行列式的计算是本章需要重点练习的一个环节. 原则上来讲, 行列式的计算可采用多种方法, 通过行列式的组合定义、行列式三角化、利用行列式的展开降阶、递推等技巧可简化行列式的计算. 下面通过具体的习题展示这些技巧和一些典型行列式的结论.

练习 1.3　计算行列式

$$\begin{vmatrix}0&0&0&1\\0&0&2&0\\0&3&0&0\\4&0&0&0\end{vmatrix}.$$

1.5.1　利用组合定义计算

例 1.2　计算行列式

$$\begin{vmatrix}a&b&c&d\\e&f&g&h\\x&y&0&0\\u&v&0&0\end{vmatrix}.$$

1.5.2　利用行列式的性质计算

例 1.3 计算

$$D=\begin{vmatrix} 3 & 1 & -1 & 2 \\ -5 & 1 & 3 & -4 \\ 2 & 0 & 1 & -1 \\ 1 & -5 & 3 & -3 \end{vmatrix}$$

1.5.3　行列式的展开–降阶法

例 1.4 计算 n 阶行列式

$$D_n=\begin{vmatrix} a & b & 0 & \cdots & 0 & 0 \\ 0 & a & b & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a & b \\ b & 0 & 0 & \cdots & 0 & a \end{vmatrix}.$$

1.5.4　递推法计算行列式

例 1.5 计算 $2n$ 阶行列式

$$D_{2n}=\begin{vmatrix} a & & & & & b \\ & \ddots & & & \iddots & \\ & & a & b & & \\ & & c & d & & \\ & \iddots & & & \ddots & \\ c & & & & & d \end{vmatrix},$$

其中未写出的元素为 0 .

1.5.5　行列式三角化

例 1.6 设

$$D=\begin{vmatrix} a_{11} & \cdots & a_{1k} & & & \\ \vdots & & \vdots & & 0 & \\ a_{k1} & \cdots & a_{kk} & & & \\ c_{11} & \cdots & c_{1k} & b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{n1} & \cdots & c_{nk} & b_{n1} & \cdots & b_{nn} \end{vmatrix},$$

$$D_1=\det\left(a_{ij}\right)=\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix}, D_2=\det\left(b_{ij}\right)=\begin{vmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{vmatrix},$$

证明 $D=D_1D_2$.

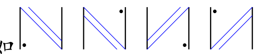
思考　判断：计算

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} \xlongequal[r_1+r_2]{r_2-r_1} \begin{vmatrix} a+c & b+d \\ c-a & d-b \end{vmatrix}$$

是否正确？

练习 1.4　计算

$$D=\begin{vmatrix} a & b & c & d \\ a & a+b & a+b+c & a+b+c+d \\ a & 2a+b & 3a+2b+c & 4a+3b+2c+d \\ a & 3a+b & 6a+3b+c & 10a+6b+3c+d \end{vmatrix}$$

形如的行列式可叫做两杠一星行列式，不难发现这类行列式都可用降阶法计算.

本例题结论可简化行列式计算，记下来.

1.5.6　副对角行列式

利用行列式的性质可证明对于 n 阶副对角行列式, 有如下结论.

$$\begin{vmatrix} & & & & a_1 \\ & * & & a_2 & \\ & & a_3 & & \\ & \iddots & & & \\ a_n & & & & \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}} a_1 a_2 \cdot \cdots \cdot a_n$$

1.5.7　同行和行列式

例 1.7　计算 n 阶行列式

$$D_n = \begin{vmatrix} x & a & \cdots & a \\ a & x & \cdots & a \\ \vdots & \vdots & & \vdots \\ a & a & \cdots & x \end{vmatrix}.$$

练习 1.5　计算

$$D = \begin{vmatrix} 3 & 1 & 1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 3 & 1 \\ 1 & 1 & 1 & 3 \end{vmatrix}$$

1.5.8　爪形行列式

例 1.8　计算行列式 (假定 $a_i \neq 0$) :

$$D_{n+1} = \begin{vmatrix} a_0 & 1 & 1 & \cdots & 1 \\ 1 & a_1 & 0 & \cdots & 0 \\ 1 & 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & a_n \end{vmatrix}.$$

形如 $\left|\nwarrow\right|$ $\left|\nearrow\right|$ $\left|\swarrow\right|$ $\left|\searrow\right|$ 的行列式可称为爪形行列式.　它们的计算方法都是类似的, 通过作业去练习.

1.5.9　升阶法 (加边法) 计算行列式

例 1.9　计算 n 阶行列式

$$D_n = \begin{vmatrix} a_1 - x & a_2 & \cdots & a_n \\ a_1 & a_2 - x & \cdots & a_n \\ \vdots & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_n - x \end{vmatrix}.$$

1.5.10　范德蒙行列式

例 1.10　证明范德蒙行列式

$$D = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \\ a_1^2 & a_2^2 & a_3^2 & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_1^{n-1} & a_2^{n-1} & a_3^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{n \geqslant i > j \geqslant 1} (a_i - a_j), \quad (n \geqslant 2).$$

式中 $\prod$ 为连乘号, $\prod\limits_{n\geqslant i>j\geqslant 1}(a_i-a_j)$ 表示 $a_1,a_2,\cdots,a_n$ 这 n 个数的所有可能 $a_i-a_j(i>j)$ 的乘积. 例如, 当 $n=4$ 时,

$$\prod_{4\geqslant i>j\geqslant 1}(a_i-a_j)=(a_2-a_1)(a_3-a_1)(a_4-a_1)(a_3-a_2)\cdot(a_4-a_2)(a_4-a_3).$$

证明: 用数学归纳法. 当 $n=2$ 时,

$$\begin{vmatrix}1 & 1\\ a_1 & a_2\end{vmatrix}=a_2-a_1$$

结论正确.

假设对于 $n-1$ 阶范德蒙行列式结论成立. 现在来看 n 阶的情形, 从第 n 行开始, 逐行减去上面相邻行的 a_1 倍, 得

$$D=\begin{vmatrix}1 & 1 & 1 & \cdots & 1\\ 0 & a_2-a_1 & a_3-a_1 & \cdots & a_n-a_1\\ 0 & a_2(a_2-a_1) & a_3(a_3-a_1) & \cdots & a_n(a_n-a_1)\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & a_2^{n-2}(a_2-a_1) & a_3^{n-2}(a_3-a_1) & \cdots & a_n^{n-2}(a_n-a_1)\end{vmatrix}$$

对第一列展开, 然后提出每一列元素的公因子, 就有

$$D=(a_2-a_1)(a_3-a_1)\cdot\cdots\cdot(a_n-a_1)\begin{vmatrix}1 & 1 & \cdots & 1\\ a_2 & a_3 & \cdots & a_n\\ a_2^2 & a_3^2 & \cdots & a_n^2\\ \vdots & \vdots & \vdots & \vdots\\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2}\end{vmatrix}$$

后面的行列式是一个 $n-1$ 阶的范德蒙行列式, 由归纳法假设, 它等于 $\prod\limits_{n\geqslant i>j\geqslant 2}(a_i-a_j)$, 代人上式, 即得

$$D=(a_2-a_1)(a_3-a_1)\cdots(a_n-a_1)\prod_{n\geqslant i>j\geqslant 2}(a_i-a_j)=\prod_{n\geqslant i>j\geqslant 1}(a_i-a_j).$$

□

本题结论在求解问题时很有用, 记下来. 口诀: 后减前连乘, 依次穷尽.

练习 1.6 对于 4 阶行列式

$$D=\begin{vmatrix}2 & -3 & 1 & 5\\ -1 & 5 & 7 & -8\\ 2 & 2 & 2 & 2\\ 0 & 1 & -1 & 0\end{vmatrix},$$

计算:

(1) $2A_{31}-3A_{32}+A_{33}+5A_{34}$;

(2) $M_{14}+M_{24}+M_{34}+M_{44}$.

1.5.11 代数余子式的和

例 1.11 设 $D_4=\begin{vmatrix}1 & 1 & 1 & 2\\ 1 & 1 & -2 & 0\\ 1 & 2 & 0 & -1\\ 2 & -3 & 4 & 3\end{vmatrix}$, 求 $A_{41}+A_{42}+A_{43}+A_{44}$.

1.6　克拉默法则

含有 n 个未知数 $x_1, x_2, \cdots, x_n$ 的 n 个线性方程的方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases} \tag{1-1}$$

与二、三元线性方程组相类似, 它的解可以用 n 阶行列式表示, 即有

克拉默法则

如果线性方程组的系数行列式不等于零, 即

$$D = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \neq 0,$$

那么, 方程组 (1-1) 有惟一解

$$x_1 = \frac{D_1}{D}, \quad x_2 = \frac{D_2}{D}, \cdots, \quad x_n = \frac{D_n}{D},$$

其中 $D_j (j = 1, 2, \cdots, n)$ 是把系数行列式 D 中第 j 列的元素用方程组右端的常数项代替后所得到的 n 阶行列式, 即

$$D_j = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}.$$

撇开上述求解公式, **克拉默法则**可叙述为下面的定理:

定理 1.2

如果线性方程组 (1-1) 的系数行列式 $D \neq 0$, 则方程组 (1-1) 一定有解, 且解是惟一的.

定理 1.3

如果线性方程组 (1-1) 无解或有两个不同的解, 则它的系数行列式必为零.

第　二　章　　矩阵及其计算

数缺形时少直观, 形缺数时难入微. 数形结合百般好, 隔离分家万事休.

华罗庚 (1910–1985)

本章主要关注以下几个问题.

1) 矩阵概念的引入. 矩阵与线性变换之间的联系.
2) 矩阵的加法、数乘和乘法的运算规则. 尤其要注意矩阵乘法为什么要如此定义以及其不满足交换律的特点.
3) 伴随矩阵、逆矩阵的概念和求解方法.
4) 矩阵分块用于矩阵运算的技巧. 该技巧能使阶数较高的矩阵运算通过分块降成阶数低的矩阵的运算, 使问题简化.

本章学习中除注意概念和矩阵运算的特点外, 需要充分地练习. 只有通过具体动手计算才能深入体会、熟练并理解本章所学的知识点.

2.1　线性方程组与矩阵

2.1.1　线性方程组

我们用线性方程组的求解说明**矩阵**概念引入的必要性. 首先对线性方程组作出定义.

定义 2.1: 线性方程组

形如

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

的具有 n 个未知数、m 个方程的线性方程组合称为**线性方程组**. 其中 a_{ij} 是第 i 个方程的第 j 个未知数的系数, b_i 是第 i 个方程的常数项 $(i=1,2,\cdots,m; j=1,2,\cdots,n)$. 当常数项 $b_1,b_2,\cdots,b_m$ **不全为 0** 时, 线性方程组称为 **n 元非齐次线性方程组**. 而当常数项**全为 0** 时,

2-1变为

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & 0, \\ \cdots\cdots\cdots & & & & & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0, \end{cases} \tag{2-2}$$

2-2称为n **元齐次线性方程组**.

在第一章中利用克拉默法则我们已经会求解一种特殊 (方程个数与未知量个数**相等**且系数行列式**不等于零**) 的线性方程组, 其解由方程组中未知量的系数及常数项构成的行列式**唯一确定**, 与未知量的记号无关. 对于一般的 n **元线性方程组** (方程个数 m 与未知量个数 n **可以不等**), 其解的判定及求解显然**不能再利用克拉默法则**(因为系数数表的行数与列数不等, 不存在行列式), 此时我们可以通过未知量的系数和常数项构成的**矩形数表**

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{matrix} \tag{2-3}$$

研究其解的规律, 如: 线性方程组是否有解? 有解时是否是唯一的解? 如果有多个解, 如何求出其所有的解? 为表明该数表的整体性, 我们通常用括号 (可用圆括号或方括号, 含义完全相同) 将之括起来, 表示如下:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}. \tag{2-4}$$

在实际问题处理中, 我们也通常将数据按照一定的顺序列成**矩形数表**以简洁表示数据信息, 并方便后续的数据处理.

2.1.2 矩阵的定义

定义 2.2: 矩阵

由 $m \times n$ 个数 $a_{ij}(i = 1, 2, \cdots, m; j = 1, 2, \cdots, n)$ 按一定顺序排成的 m 行 n 列的矩形数表, 称为 $m \times n$ 矩阵, 简称**矩阵**. 矩阵用**黑**

体的**大写英文字母**表示, 记作

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

其中数 a_{ij} 为位于矩阵 $\boldsymbol{A}$ 中第 i 行第 j 列的元素. 以数 a_{ij} 为元素的矩阵记为 (a_{ij}) 或 $(a_{ij})_{m\times n}$, $m\times n$ 矩阵也记为 $\boldsymbol{A}_{m\times n}$.
元素 a_{ij} 是实数的矩阵称为**实 (数) 矩阵**, 元素 a_{ij} 是复数的矩阵称为**复 (数) 矩阵**. 如无特殊声明, 本书中的矩阵都是指实 (数) 矩阵.
行数和列数都等于 n 的矩阵称为 n **阶矩阵**或 n **阶方阵**.
特殊情况下, **行数为 1** 、列数为 n (只有一行) 的矩阵

$$\boldsymbol{A}=\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$$

称为**行矩阵**或**行向量**. 为避免元素间的混淆, 行矩阵也记为

$$\boldsymbol{A}=(a_1,a_2,\cdots,a_n).$$

行数为 m 、**列数为 1** (只有一列) 的矩阵

$$\boldsymbol{B}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$

称为**列矩阵**或**列向量**.
若矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ 的**行数与列数均相等**, 则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 为**同型矩阵**.
如果同型矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 对应位置上的元素都相等, 则称矩阵 $\boldsymbol{A}$ 与矩阵 $\boldsymbol{B}$ **相等**, 记为 $\boldsymbol{A}=\boldsymbol{B}$.

2.1.3 一些特殊矩阵

定义 2.3: 零矩阵

元素均为零的矩阵称为**零矩阵**, 记为 $\boldsymbol{O}_{m\times n}$, 也简记为 $\boldsymbol{O}$ (大写字母). 注意, **不同型的零矩阵是不相等**的, 例如, 二阶零矩阵与 4×1

零矩阵不相等, 即

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

定义 2.4: 三角矩阵、对角矩阵、单位矩阵、数量矩阵

主对角线上方元素均为 0 的 n 阶方阵

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

称为 n **阶下三角矩阵**.　**主对角线下方元素均为 0** 的 n 阶方阵

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

称为 n **阶上三角矩阵**.　上三角矩阵与下三角矩阵统称为**三角矩阵**.
主对角线之外的元素均为 0, 而主对角线元素不全为 0 的 n 阶方阵

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

称为 n **阶对角矩阵**.　n 阶对角矩阵也常记为 $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$.
主对角线元素均为 1 的 n 阶对角矩阵

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

称为 n **阶单位矩阵**, 记为 $\boldsymbol{E}_n$ 或 $\boldsymbol{I}_n$, 简记为 $\boldsymbol{E}$ 或 $\boldsymbol{I}$.

主对角线上元素均为 a 的 n 阶对角矩阵

$$\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & a & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a \end{pmatrix}$$

称为 n **阶数量矩阵**或 n **阶标量矩阵**, 简记为 $a\boldsymbol{E}$ 或 $a\boldsymbol{I}$.

2.2 矩阵的计算

2.2.1 矩阵的加法和数乘

定义 2.5: 矩阵的加法和减法

设矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{ij})_{m\times n}$, 称矩阵 $\boldsymbol{C} = (c_{ij})_{m\times n} = (a_{ij}+ b_{ij})_{m\times n}$ 为**矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和**, 记为 $\boldsymbol{C} = \boldsymbol{A} + \boldsymbol{B}$.

设矩阵 $\boldsymbol{A} = (a_{ij})_{m\times n}$, 称 $-\boldsymbol{A} = (-a_{ij})_{m\times n}$ 为 $\boldsymbol{A}$ 的**负矩阵**, 从而定义**矩阵的减法**为

$$\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B}).$$

注意

只有**同型矩阵**才能进行加法运算.

定义 2.6: 矩阵的数乘

设 $\boldsymbol{A} = (a_{ij})_{m\times n}$, k 为任一常数, 则称矩阵

$$(ka_{ij})_{m\times n} = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}$$

为 k 与 $\boldsymbol{A}$ 的乘积, 简称**数乘**, 记为 $k\boldsymbol{A}$ 或 $\boldsymbol{A}k$.

由数乘运算定义, 数量矩阵

$$\begin{pmatrix} a & & & \\ & a & & \\ & & \ddots & \\ & & & a \end{pmatrix} = a\boldsymbol{E}.$$

矩阵的加法与矩阵的数乘运算统称为矩阵的**线性运算**. **线性运算**满足以下八条运算规律.

1) 加法交换律: $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$.
2) 加法结合律: $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$.

3) $\boldsymbol{A}+\boldsymbol{O}=\boldsymbol{A}$.

4) $\boldsymbol{A}+(-\boldsymbol{A})=\boldsymbol{A}-\boldsymbol{A}=\boldsymbol{O}$,

5) $0\boldsymbol{A}=\boldsymbol{O}, 1\boldsymbol{A}=\boldsymbol{A}$.

6) 数乘结合律: $(kl)\boldsymbol{A}=k(l\boldsymbol{A})=l(k\boldsymbol{A})$.

7) 数乘分配律: $(k+l)\boldsymbol{A}=k\boldsymbol{A}+l\boldsymbol{A}$.

8) 数乘分配律: $k(\boldsymbol{A}+\boldsymbol{B})=k\boldsymbol{A}+k\boldsymbol{B}$.

以后会定义满足八条规律的向量集合为线性空间.

其中 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 和零矩阵 $\boldsymbol{O}$ 是同型矩阵. k, l 为两个常数.

2.2.2　线性变换与矩阵的乘法

在许多实际问题中, 我们经常遇到 m 个变量 $y_1, y_2, \cdots, y_m$ 用 n 个变量 $x_1, x_2, \cdots, x_n$ 线性地表示, 即

$$\left\{\begin{array}{lllllllll} y_1 & = & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n, \\ y_2 & = & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n, \\ & & \cdots\cdots\cdots & & & & & & \\ y_m & = & a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n, \end{array}\right. \tag{2-5}$$

定义 2.7: 线性变换、系数矩阵

给定 n 个数 $x_1, x_2, \cdots, x_n$, 经过线性计算得到了 m 个数 $y_1, y_2, \cdots, y_m$, 从变量 $x_1, x_2, \cdots, x_n$ 到变量 $y_1, y_2, \cdots, y_m$ 的变换就定义为**线性变换**. 线性变换的系数 a_{ij} 构成矩阵, 称 $\boldsymbol{A}=(a_{ij})_{m\times n}$ 为**系数矩阵**.

给定了线性变换, 就确定了一个系数矩阵; 反之, 若给出一个矩阵作为线性变换的系数矩阵, 则线性变换也就确定了. 在这个意义上, 线性变换与矩阵之间存在**一一对应**的关系.

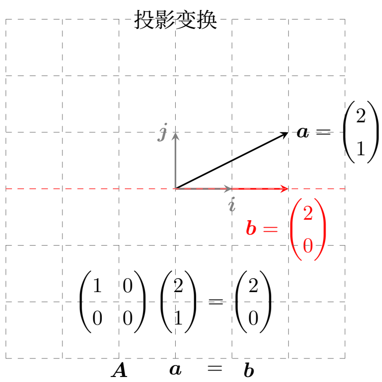


图 2-1　投影变换.

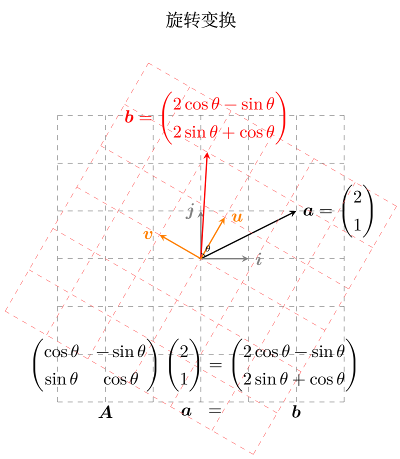


图 2-2　旋转变换

线性变换

接下来利用线性变换来解释矩阵的含义.

矩阵 $\begin{pmatrix}1&0\\0&0\end{pmatrix}$ 对应的线性变换施加给向量 $\begin{pmatrix}2\\1\end{pmatrix}$ 的结果如图2-1所示, 这是一个投影变换.

而矩阵 $\begin{pmatrix}\cos\theta&-\sin\theta\\\sin\theta&\cos\theta\end{pmatrix}$ 对应的线性变换施加给向量 $\begin{pmatrix}2\\1\end{pmatrix}$ 的结果如图2-2所示, 这是一个旋转角度 θ 的变换.

矩阵的乘法

设有两个线性变换

$$\left\{\begin{array}{l} y_1=a_{11}x_1+a_{12}x_2+a_{13}x_3, \\ y_2=a_{21}x_1+a_{22}x_2+a_{23}x_3, \end{array}\right. \tag{2-6}$$

$$\begin{cases} x_1 = b_{11}t_1 + b_{12}t_2, \\ x_2 = b_{21}t_1 + b_{22}t_2, \\ x_3 = b_{31}t_1 + b_{32}t_2, \end{cases} \tag{2-7}$$

式 (2-6) 对应的矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}, \tag{2-8}$$

式 (2-7) 对应的矩阵

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}, \tag{2-9}$$

为了求出从 t_1, t_2 到 y_1, y_2 的线性变换, 可将 式 (2-7) 代入 式 (2-6), 得

$$\begin{cases} y_1 = (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31})\, t_1 + (a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32})\, t_2 \\ y_2 = (a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31})\, t_1 + (a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32})\, t_2 \end{cases} \tag{2-10}$$

式 (2-10) 可看成是先作式 (2-7) 线性变换再作式 (2-6) 线性变换的结果. 我们把式 (2-10) 对应的矩阵记为

$$\begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}. \tag{2-11}$$

我们把式 (2-11) 称为式 (2-6) 与式 (2-7) 的乘积, 相应地, 其所对应的矩阵定义为式 (2-6) 与式 (2-7) 所对应的矩阵的乘积, 即

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}.$$

由此推广, 可得到一般矩阵乘法的定义.

定义 2.8: 矩阵的乘法

设矩阵 $\boldsymbol{A} = (a_{ij})_{m\times s}$, 矩阵 $\boldsymbol{B} = (b_{ij})_{s\times n}$, 则它们的乘积 $\boldsymbol{AB}$ 等于矩阵 $\boldsymbol{C} = (c_{ij})_{m\times n}$, 记作 $\boldsymbol{AB} = \boldsymbol{C}$, 其中

$$c_{ij} = (a_{i1}, a_{i2}, \cdots, a_{is}) \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{sj} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj}.$$

$$(i=1,2,\cdots,m;j=1,2,\cdots,n)$$

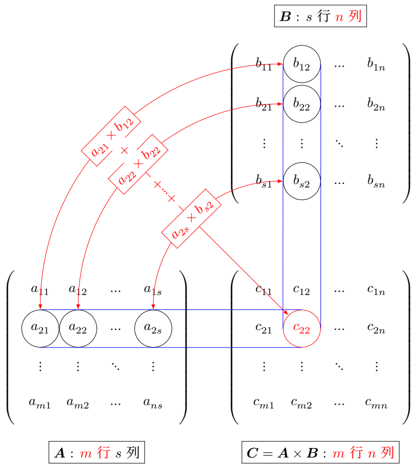


图 2-3　矩阵乘法图示

思考　为什么我们没有定义矩阵的除法?

注意

只有第 1 个矩阵的列数等于第 2 个矩阵的行数时, 两个矩阵的乘法才有意义, 即应有 $A_{m\times s}B_{s\times n}=C_{m\times n}$. 而乘积矩阵 C 的元素 c_{ij} 是把矩阵 A 中的**第 i 行元素**与矩阵 B 中的**第 j 列元素**对应相乘后再相加得到的, 即 $c_{ij}=\sum\limits_{t=1}^{s}a_{it}b_{tj}$.

例 2.1　计算矩阵乘积 AB 与 BA, 其中

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

例 2.2　设 $A=\begin{pmatrix}1&1\\-1&-1\end{pmatrix},B=\begin{pmatrix}-2&1\\2&-1\end{pmatrix},C=\begin{pmatrix}2&3\\1&-3\end{pmatrix},D=\begin{pmatrix}1&-1\\2&1\end{pmatrix}$, 计算 AB,BA,AC,AD.

注意

矩阵乘法与数的乘法在运算中有许多不同之处, 需要注意.

1) 矩阵乘法**不满足交换律**. 这是因为 AB 与 BA 不一定都有意义; 即使 AB 与 BA 都有意义, 也不一定有 $AB=BA$ 成立. 特别地, 对于方阵 A,B, 如果有 $AB=BA$, 则称矩阵 A,B **可交换**.
2) 在矩阵乘法的运算中, "若 $AB=O$, 则必有 $A=O$ 或 $B=O$" 这个结论不一定成立.
3) 矩阵乘法的消去律不成立, 即 "若 $AD=AC$ 且 $A\neq O$, 则 $D=C$" 这个结论不一定成立.

矩阵乘法的运算规律

矩阵乘法虽然不满足交换律, 但它满足以下运算规律.

1) 结合律 $(AB)C=A(BC)$.
2) 分配律 $A(B+C)=AB+A,(B+C)A=BA+CA$.
3) $\lambda AB=(\lambda A)B=A(\lambda B)$.

不难得到以下结论.

1) $E_mA_{m\times n}=A_{m\times n},A_{m\times n}E_n=A_{m\times n}$, 或者写成 $EA=AE=A$, 即单位矩阵 E 在矩阵乘法中的作用类似于数 1 .
2) 由于 n 阶方阵 A 可以自乘, 我们给出**方阵 A 幂**的运算定义: 设 A 为 n 阶方阵, k 是正整数, 规定

$$A^k = \overbrace{AA\cdots A}^{k}.$$

特别地, 当 A 为非零方阵时, 规定 $A^0 = E$. 由此易证

$$A^m A^n = A^{m+n}, (A^m)^n = A^{mn}\ (\ m, n \text{ 是正整数}).$$

2.2.3 矩阵的转置

思考 设 A, B 均为 n 阶方阵, k 是正整数, 判断 $(AB)^k = A^k B^k$ 是否正确, 并说明理由.

定义 2.9: 转置矩阵

设 $A = (a_{ij})_{m\times n}$, 把矩阵 A 的行换成同序数的列而得到的新矩阵, 叫作矩阵 A 的转置矩阵, 记为 A^{T}.

矩阵转置的运算规律

矩阵的转置也是一种运算, 满足下述运算规律 (假设运算都是可行的):

1) $(A^{\mathrm{T}})^{\mathrm{T}} = A$;
2) $(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$;
3) $(\lambda A)^{\mathrm{T}} = \lambda A^{\mathrm{T}}$;
4) $(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}$.

例 2.3 已知

$$A = \begin{pmatrix} 2 & 0 & -1 \\ 1 & 3 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 7 & -1 \\ 4 & 2 & 3 \\ 2 & 0 & 1 \end{pmatrix},$$

求 $(AB)^{\mathrm{T}}$.

定义 2.10: 对称矩阵

设 A 为 n 阶方阵, 如果满足 $A^{\mathrm{T}} = A$, 即 $a_{ij} = a_{ji} (i, j = 1, 2, \cdots, n)$, 则称为 **$n$ 阶对称矩阵**. 例如

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 7 & 5 \\ -3 & 5 & 6 \end{pmatrix}$$

是三阶对称矩阵. 对称矩阵的特点: 关于**主对角线对称的元素相等**.

2.2.4 方阵的行列式

定义 2.11: 方阵的行列式

由 n 阶方阵 A 的元素所构成的行列式 (各元素的位置不变), 称为 **方阵 A 的行列式**, 记作 $|A|$ 或 $\det A$.

注意

方阵与行列式是两个不同的概念, n 阶方阵是 n^2 个数按一定方式排成的**数表**, 而 n 阶行列式则是这些数 (也就是数表 $\boldsymbol{A}$) 按一定的运算法则所确定的**一个数**.

由 $\boldsymbol{A}$ 确定 $|\boldsymbol{A}|$ 的这个运算满足下述运算规律 (设 $\boldsymbol{A},\boldsymbol{B}$ 为 n 阶方阵, λ 为数):

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

2) $|\lambda\boldsymbol{A}|=\lambda^n|\boldsymbol{A}|$;

3) $|\boldsymbol{A}\boldsymbol{B}|=|\boldsymbol{A}||\boldsymbol{B}|$.

推论 2.1

对于 n 阶矩阵 $\boldsymbol{A},\boldsymbol{B}$, 一般来说 $\boldsymbol{A}\boldsymbol{B}\neq\boldsymbol{B}\boldsymbol{A}$, 但总有

$$|\boldsymbol{A}\boldsymbol{B}|=|\boldsymbol{B}\boldsymbol{A}|.$$

2.2.5　伴随矩阵

定义 2.12: 伴随矩阵

行列式 $|\boldsymbol{A}|$ 的各个元素的代数余子式 A_{ij} 所构成的如下的矩阵

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

称为矩阵 $\boldsymbol{A}$ 的**伴随矩阵**, 简称**伴随阵**.

例 2.4　$\boldsymbol{A}^*$ 是 n 阶方阵 $\boldsymbol{A}$ 的伴随矩阵. 试证

$$\boldsymbol{A}\boldsymbol{A}^*=\boldsymbol{A}^*\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{E}.$$

2.3　逆矩阵

在数的运算中, 当数 $a\neq 0$ 时, 有 $a\cdot a^{-1}=a^{-1}\cdot a=1$, 其中 a^{-1} 为 a 的倒数. 在矩阵的乘法运算中, 单位矩阵 $\boldsymbol{E}$ 相当于数的乘法运算中的 1, 那么对于方阵 $\boldsymbol{A}$, 是否存在一个矩阵 $\boldsymbol{A}^{-1}$, 使 $\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{E}$ 呢?

2.3.1　逆矩阵的定义、性质与求法

定义 2.13: 逆矩阵

对于 n 阶方阵 $\boldsymbol{A}$, 如果有一个 n 阶方阵 $\boldsymbol{B}$, 使 $\boldsymbol{A}\boldsymbol{B}=\boldsymbol{B}\boldsymbol{A}=\boldsymbol{E}$, 则称矩阵 $\boldsymbol{A}$ **可逆**, 矩阵 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的**逆矩阵**. 我们将方阵 $\boldsymbol{A}$ 的 (唯一

的) 逆阵记作 $\boldsymbol{A}^{-1}$, 而 $\boldsymbol{A}^{-1}$ 满足

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}.$$

注意

如果方阵 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{A}$ 的逆矩阵是**唯一**的. 这是因为, 若方阵 $\boldsymbol{B}, \boldsymbol{C}$ 都是方阵 $\boldsymbol{A}$ 的逆矩阵, 则有

$$\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{E}, \quad \boldsymbol{AC} = \boldsymbol{CA} = \boldsymbol{E},$$

可推出

$$\boldsymbol{B} = \boldsymbol{BE} = \boldsymbol{B}(\boldsymbol{AC}) = (\boldsymbol{BA})\boldsymbol{C} = \boldsymbol{EC} = \boldsymbol{C},$$

即 $\boldsymbol{B} = \boldsymbol{C}$.

定理 2.1

方阵 $\boldsymbol{A}$ 可逆等价于 $|\boldsymbol{A}| \neq 0$, 且当 $\boldsymbol{A}$ 可逆时, 有

$$\boxed{\boldsymbol{A}^{-1} = \frac{1}{|\boldsymbol{A}|}\boldsymbol{A}^*},$$

其中

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

是 $\boldsymbol{A}$ 的**伴随矩阵**, 而 A_{ij} 是 a_{ij} 的代数余子式.

该定理实际上提供了一种计算逆矩阵的方法, 但仅仅是理论上的, 一般只适用于低阶矩阵的逆矩阵的计算. 高阶矩阵情况下, 应用该定理计算明显不明智.

若 n 阶矩阵 $\boldsymbol{A}$ 的行列式 $|\boldsymbol{A}| = 0$, 则称 $\boldsymbol{A}$ 是**奇异矩阵**, 否则成**非奇异矩阵**. 因此, **矩阵可逆** $\Longleftrightarrow$ **矩阵非奇异**.

推论 2.2

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

性质 2.1: 逆矩阵的性质

1) 如果 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{A}^{-1}$ 也可逆, 而且 $(\boldsymbol{A}^{-1})^{-1} = \boldsymbol{A}$;
2) 如果 $\boldsymbol{A}$ 可逆, k 为非零常数, 则 $k\boldsymbol{A}$ 也可逆, 而且 $(k\boldsymbol{A})^{-1} = \dfrac{1}{k}\boldsymbol{A}^{-1}$;
3) 如果 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{A}^{\mathrm{T}}$ 也可逆, 而且 $(\boldsymbol{A}^{\mathrm{T}})^{-1} = (\boldsymbol{A}^{-1})^{\mathrm{T}}$;
4) 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 同阶而且都可逆, 则 $\boldsymbol{AB}$ 也可逆, 而且 $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$;
5) 如果 $\boldsymbol{A}$ 可逆, 则有 $|\boldsymbol{A}^{-1}| = |\boldsymbol{A}|^{-1}$.

例 2.5 求二阶矩阵 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的逆矩阵.

将本例题的结论记下来. 口诀: 主对调, 副变号, 再除以矩阵的行列式.

例 2.6 求 $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ 的逆矩阵.

总结

对称阵的逆矩阵仍为对称阵.

例 2.7 求 $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$ 的逆矩阵.

2.3.2 逆矩阵的应用

可逆矩阵在线性代数中非常重要, 有广泛的应用. 常见的使用场合是求解矩阵方程、计算矩阵的多项式. 下面用例题进行介绍.

例 2.8 设

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 1 \\ 3 & 4 & 3 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} 1 & 3 \\ 2 & 0 \\ 3 & 1 \end{pmatrix},$$

求矩阵 $\boldsymbol{X}$ 使其满足

$$\boldsymbol{AXB} = \boldsymbol{C}.$$

例 2.9 设 $\boldsymbol{P} = \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix}, \boldsymbol{\Lambda} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \boldsymbol{AP} = \boldsymbol{P\Lambda}$, 求 $\boldsymbol{A}^n$.

2.4 矩阵分块法

对于行数和列数较多的矩阵, 采取分块法来计算可以提高效率, 同时减少运算错误.

2.4.1 分块矩阵的定义

定义 2.14: 分块矩阵

用若干条贯穿矩阵的横线和纵线将矩阵 $\boldsymbol{A}$ 分成若干个小块, 每一小块称为矩阵 $\boldsymbol{A}$ 的**子块** (或子阵). 以子块为元素的形式上的矩阵称为**分块矩阵**.

例如, 3×4 矩阵

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix},$$

分块方式有很多, 下面列举其中的 3 种分块方法.

1) 普通分块

$$\boldsymbol{A}=\left(\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right)=\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}.$$

2) 按行分块

$$\boldsymbol{A}=\left(\begin{array}{cccc} a_{11} & a_{12} & a_{13} & a_{14} \\ \hline a_{21} & a_{22} & a_{23} & a_{24} \\ \hline a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right).$$

3) 按列分块

$$\boldsymbol{A}=\left(\begin{array}{c|c|c|c} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{array}\right).$$

在矩阵运算中, 对矩阵做分块时需要掌握两个原则: (1) 使矩阵的子块像 "数" 一样满足矩阵运算的要求, 不同的运算要采取不同的分块方法; (2) 使运算尽量**简单方便**.

2.4.2　分块矩阵的运算

分块矩阵的运算与普通矩阵的运算规则相类似, 详细说明如下.

分块矩阵的代数和运算

设矩阵 $\boldsymbol{A},\boldsymbol{B}$ 的行数和列数相同, 对 $\boldsymbol{A},\boldsymbol{B}$ 采用相同的分块法, 有

$$\boldsymbol{A}=\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st} \end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1t} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{B}_{s1} & \boldsymbol{B}_{s2} & \cdots & \boldsymbol{B}_{st} \end{pmatrix},$$

其中 $\boldsymbol{A}_{ij},\boldsymbol{B}_{ij}(i=1,2,\cdots,s;j=1,2,\cdots,t)$ 的行数和列数也相同, 则

$$\boldsymbol{A}\pm\boldsymbol{B}=\begin{pmatrix} \boldsymbol{A}_{11}\pm\boldsymbol{B}_{11} & \boldsymbol{A}_{12}\pm\boldsymbol{B}_{12} & \cdots & \boldsymbol{A}_{1t}\pm\boldsymbol{B}_{1t} \\ \boldsymbol{A}_{21}\pm\boldsymbol{B}_{21} & \boldsymbol{A}_{22}\pm\boldsymbol{B}_{22} & \cdots & \boldsymbol{A}_{2t}\pm\boldsymbol{B}_{2t} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{A}_{s1}\pm\boldsymbol{B}_{s1} & \boldsymbol{A}_{s2}\pm\boldsymbol{B}_{s2} & \cdots & \boldsymbol{A}_{st}\pm\boldsymbol{B}_{st} \end{pmatrix}.$$

分块矩阵的数乘运算

设 $\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2t}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st}\end{pmatrix}$, λ 是数, 则

$$\lambda\boldsymbol{A}=\begin{pmatrix}\lambda\boldsymbol{A}_{11} & \lambda\boldsymbol{A}_{12} & \cdots & \lambda\boldsymbol{A}_{1t}\\ \lambda\boldsymbol{A}_{21} & \lambda\boldsymbol{A}_{22} & \cdots & \lambda\boldsymbol{A}_{2t}\\ \vdots & \vdots & & \vdots\\ \lambda\boldsymbol{A}_{s1} & \lambda\boldsymbol{A}_{s2} & \cdots & \lambda\boldsymbol{A}_{st}\end{pmatrix}.$$

分块矩阵的乘法

设矩阵 $\boldsymbol{A}=(a_{ij})_{m\times s}$, $\boldsymbol{B}=(b_{ij})_{s\times n}$, 且对 $\boldsymbol{A}$ 的列分块方法与对 $\boldsymbol{B}$ 的行分块方法相同, 即分块为

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \cdots & \boldsymbol{A}_{1t}\\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \cdots & \boldsymbol{A}_{2t}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \cdots & \boldsymbol{A}_{st}\end{pmatrix},\quad \boldsymbol{B}=\begin{pmatrix}\boldsymbol{B}_{11} & \boldsymbol{B}_{12} & \cdots & \boldsymbol{B}_{1r}\\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} & \cdots & \boldsymbol{B}_{2r}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{B}_{t1} & \boldsymbol{B}_{t2} & \cdots & \boldsymbol{B}_{tr}\end{pmatrix},$$

其中矩阵 $\boldsymbol{A}$ 的第 i 行各子块 $\boldsymbol{A}_{i1},\boldsymbol{A}_{i2},\cdots,\boldsymbol{A}_{it}$ 的列数分别等于矩阵 $\boldsymbol{B}$ 的第 j 列的各子块 $\boldsymbol{B}_{1j},\boldsymbol{B}_{2j},\cdots,\boldsymbol{B}_{ij}$ 的行数, 则

$$\boldsymbol{AB}=\begin{pmatrix}\boldsymbol{C}_{11} & \boldsymbol{C}_{12} & \cdots & \boldsymbol{C}_{1r}\\ \boldsymbol{C}_{21} & \boldsymbol{C}_{22} & \cdots & \boldsymbol{C}_{2r}\\ \vdots & \vdots & & \vdots\\ \boldsymbol{C}_{s1} & \boldsymbol{C}_{s2} & \cdots & \boldsymbol{C}_{sr}\end{pmatrix}$$

其中

$$\boldsymbol{C}_{ij}=\sum_{k=1}^{t}\boldsymbol{A}_{ik}\boldsymbol{B}_{kj}=\boldsymbol{A}_{i1}\boldsymbol{B}_{1j}+\boldsymbol{A}_{i2}\boldsymbol{B}_{2j}+\cdots+\boldsymbol{A}_{it}\boldsymbol{B}_{ij}$$

$$i=1,2,\cdots,s;\quad j=1,2,\cdots,r$$

例 2.10 设

$$\boldsymbol{A}=\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ -1 & 2 & 1 & 0\\ 1 & 1 & 0 & 1\end{pmatrix},\boldsymbol{B}=\begin{pmatrix}1 & 0 & 1 & 0\\ -1 & 2 & 0 & 1\\ 1 & 0 & 4 & 1\\ -1 & -1 & 2 & 0\end{pmatrix},$$

求 $\boldsymbol{AB}$.

线性方程组与矩阵乘法的等价性

对于线性方程组

$$\left\{\begin{array}{lllllll} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m, \end{array}\right. \tag{2-12}$$

记 $\boldsymbol{A}=(a_{ij})$, $\boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\boldsymbol{b}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$, $\boldsymbol{B}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$, 其中 $\boldsymbol{A}$ 称为**系数矩阵**, $\boldsymbol{x}$ 称为**未知数向量**, $\boldsymbol{b}$ 称为**常数项向量**, $\boldsymbol{B}$ 称为**增广矩阵**. 按分块矩阵的记法, 可记

$$\boldsymbol{B}=(\boldsymbol{A}|\boldsymbol{b}), \text{ 或} \boldsymbol{B}=(\boldsymbol{A},\boldsymbol{b})=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n,\boldsymbol{b}).$$

利用矩阵的乘法, 此方程组可记作

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \tag{2-13}$$

方程 (2-13) 以向量 $\boldsymbol{x}$ 为未知元, 它的解称为方程组 (2-12) 的解向量.

如果把系数矩阵 $\boldsymbol{A}$ 按行分成 m 块, 则线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 可记作

$$\begin{pmatrix} \boldsymbol{\alpha}_1^{\mathrm{T}} \\ \boldsymbol{\alpha}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{\alpha}_m^{\mathrm{T}} \end{pmatrix}\boldsymbol{x}=\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \text{ 或} \left\{\begin{array}{c} \boldsymbol{\alpha}_1^{\mathrm{T}}\boldsymbol{x}=b_1, \\ \boldsymbol{\alpha}_2^{\mathrm{T}}\boldsymbol{x}=b_2, \\ \vdots \\ \boldsymbol{\alpha}_m^{\mathrm{T}}\boldsymbol{x}=b_m, \end{array}\right. \tag{2-14}$$

这就相当于把每个方程

$$a_{i1}x_1+a_{i2}x_2+\cdots+a_{in}x_n=b_i$$

记作

$$\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x}=b_i(i=1,2,\cdots,m).$$

如果把系数矩阵 $\boldsymbol{A}$ 按列分成 n 块, 则与 $\boldsymbol{A}$ 相乘的 $\boldsymbol{x}$ 应对应地按行分成 n 块, 从而记作

$$(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=\boldsymbol{b}, \tag{2-15}$$

即

$$x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_n\boldsymbol{\alpha}_n = \boldsymbol{b}. \tag{2-16}$$

(2-13), (2-14), (2-16) 是线性方程组 (2-12) 的各种变形. 因此, 它们可与 (2-12) 混同使用而不加区分, 并都称为线性方程组或线性方程. 解与解向量亦不加区别.

第　三　章　　矩阵的初等变换与线性方程组

今有醇酒一斗, 直钱五十; 行酒一斗, 直钱一十．今将钱三十, 得酒二斗．问醇、行酒各得几何?

《九章算术》

鸡翁一值钱五, 鸡母一值钱三, 鸡雏三值钱一. 百钱买百鸡, 问鸡翁、鸡母、鸡雏各几何?

北魏 · 张邱建《张邱建算经》

本章主要关注如下几个问题.

1) 线性方程组求解的消元法与矩阵初等变换之间的联系和等价性. 由此引入初等矩阵、初等变换、行最简型性矩阵等概念.
2) **矩阵的秩**的概念. 这是线性代数中一个**非常基本且重要并非常抽象**的概念. 利用这一概念可很好解决线性方程组是否有解和有多少解的问题. 在求解矩阵的秩时, 除利用定义求解外, 更常用的方法是通过将矩阵转化为行最简型矩阵判断.
3) 如何判断非齐次和齐次线性方程组的解.

本章学习时可从简单的包含 2–3 个方程的线性方程组求解实例出发以理解引入的概念和定理.

3.1　矩阵的初等变换

矩阵的初等变换是处理矩阵问题的一种基本方法.

3.1.1　矩阵的初等变换

定义 3.1: 矩阵的初等行 (列) 变换

下面三种变换称为**矩阵的初等行 (列) 变换**:

1) 交换矩阵中的第 i 行 (列) 与第 j 行 (列) 的元素, 记作 $r_1 \leftrightarrow r_j$ 或 $c_i \leftrightarrow c_j$;
2) 用一个非零常数 k 乘遍矩阵的第 i 行 (列), 记作 kr_i 或 kc_i;
3) 矩阵的第 j 行 (列) 元素的 k 倍加到第 i 行 (列) 对应元素上, 记作 $r_i + kr_j\,(c_i+\ kc_j\ )$. (第 j 行 (列) 的元素并没有改变)

矩阵的初等行 (列) 变换统称为**初等变换**.

显然, 三种初等变换都是可逆的, 且其逆变换是同一类型的初等变换; 变换 $r_i \leftrightarrow r_j$ 的逆变换就是其本身; 变换 $r_i \times k$ 的逆变换为

$r_i \times \left(\frac{1}{k}\right)$ (或记作 $r_i \div k$); 变换 $r_i + kr_j$ 的逆变换为 $r_i + (-k)r_j$ (或记作 $r_i - kr_j$).

定义 3.2: 矩阵与矩阵等价

如果矩阵 $\boldsymbol{A}$ 经有限次初等行变换变成矩阵 $\boldsymbol{B}$, 就称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **行等价**, 记作[a] $\boldsymbol{A} \overset{r}{\cong} \boldsymbol{B}$; 如果矩阵 $\boldsymbol{A}$ 经有限次初等列变换变成矩阵 $\boldsymbol{B}$, 就称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **列等价**, 记作 $\boldsymbol{A} \overset{c}{\cong} \boldsymbol{B}$; 如果矩阵 $\boldsymbol{A}$ 经有限次初等变换变成矩阵 $\boldsymbol{B}$, 就称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **等价**, 记作 $\boldsymbol{A} \cong \boldsymbol{B}$.

[a] 在很多书中关于矩阵等价没有统一的符号, 同济版教材里用的是 $\sim$ 符号, 讲义在这里经过权衡考虑后用 $\cong$ 符号, 以区别后面章节矩阵相似用的 $\sim$ 符号. 为避免误解, 在做题和考试时, 可以直接用文字表达.

容易证明, 矩阵等价具有以下三个性质:

1) 反身性: $\boldsymbol{A} \cong \boldsymbol{A}$;

2) 对称性: 若 $\boldsymbol{A} \cong \boldsymbol{B}$, 则 $\boldsymbol{B} \cong \boldsymbol{A}$;

3) 传递性: 若 $\boldsymbol{A} \cong \boldsymbol{B}, \boldsymbol{B} \cong \boldsymbol{C}$, 则 $\boldsymbol{A} \cong \boldsymbol{C}$.

定义 3.3: 标准型矩阵

设矩阵 $\boldsymbol{D}_{m\times n}$ 的左上角为一个单位矩阵, 其他元素都是零, 称其为**标准形矩阵**, 即

$$\boldsymbol{D}_{m\times n} = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}_{m\times n} = \begin{pmatrix} \boldsymbol{E}_r & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}_{m\times n}$$

$$(0 \leqslant r \leqslant \min\{m, n\})$$

定理 3.1

任何一个非奇异矩阵 $\boldsymbol{A}$, 都可经过有限次初等行变换变成单位矩阵 $\boldsymbol{E}$.

例 3.1 用初等行变换将 $\boldsymbol{A} = \begin{pmatrix} 0 & -2 & 1 \\ 3 & 8 & -2 \\ 1 & 3 & 0 \end{pmatrix}$ 化成单位矩阵.

定理 3.2

任意一个矩阵 $\boldsymbol{A}_{m\times n}$, 都能经过有限次初等变换变成标准形矩阵.

例 3.2 将矩阵 $\boldsymbol{A}=\begin{pmatrix}2&1&2&3\\4&1&3&5\\2&0&1&2\end{pmatrix}$ 化为标准形矩阵.

3.1.2 初等矩阵

定义 3.4: 初等矩阵

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

三种初等变换对应有三种初等矩阵.

1) 把单位阵中第 i,j 两行对调 (或第 i,j 两列对调), 得初等矩阵

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

用 m 阶初等矩阵 $\boldsymbol{E}_m(i,j)$ 左乘矩阵 $\boldsymbol{A}=(a_{ij})_{m\times n}$, 得

$$\boldsymbol{E}_m(i,j)\boldsymbol{A}=\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\\vdots&\vdots&&\vdots\\a_{j1}&a_{j2}&\cdots&a_{jn}\\\vdots&\vdots&&\vdots\\a_{i1}&a_{i2}&\cdots&a_{in}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\begin{array}{l}\\\\\text{第 i 行}\\\\\text{第 j 行}\\\\\\\end{array}$$

其结果相当于对矩阵 $\boldsymbol{A}$ 施行第一种初等行变换: 把 $\boldsymbol{A}$ 的第 i 行与第 j 行对调 $(r_i\leftrightarrow r_j)$. 类似地, 以 n 阶初等矩阵 $\boldsymbol{E}_n(i,j)$ 右乘矩阵 $\boldsymbol{A}$, 其结果相当于对矩阵 $\boldsymbol{A}$ 施行第一种初等列变换: 把 $\boldsymbol{A}$ 的第 i 列与第 j 列对调 $(c_i\leftrightarrow c_j)$.

2) 以数 $k \neq 0$ 乘单位阵的第 i 行 (或第 i 列), 得初等矩阵

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

可以验知: 以 $\boldsymbol{E}_m(i(k))$ 左乘矩阵 $\boldsymbol{A}$, 其结果相当于以数 k 乘 $\boldsymbol{A}$ 的第 i 行 $(r_i \times k)$; 以 $\boldsymbol{E}_n(i(k))$ 右乘矩阵 $\boldsymbol{A}$, 其结果相当于以数 k 乘 $\boldsymbol{A}$ 的第 i 列 $(c_i \times k)$.

3) 以 k 乘 $\boldsymbol{E}$ 的第 j 行加到第 i 行上或以 k 乘 $\boldsymbol{E}$ 的第 i 列加到第 j 列上, 得初等矩阵

$$
\boldsymbol{E}(ij(k))=\begin{array}{cl}
\begin{pmatrix}
1 &&&&&& \\
& \ddots &&&&& \\
&& 1 & \cdots & k && \\
&&& \ddots & \vdots && \\
&&&& 1 && \\
&&&&& \ddots & \\
&&&&&& 1
\end{pmatrix} &
\begin{array}{l} \\ \\ \text{第 i 行} \\ \\ \text{第 j 行} \\ \\ \\ \end{array}
\end{array}
$$

可以验知: 以 $\boldsymbol{E}_m(ij(k))$ 左乘矩阵 $\boldsymbol{A}$, 其结果相当于把 $\boldsymbol{A}$ 的第 j 行乘 k 加到第 i 行上 $(r_i + kr_j)$, 以 $\boldsymbol{E}_n(ij(k))$ 右乘矩阵 $\boldsymbol{A}$, 其结果相当于把 $\boldsymbol{A}$ 的第 i 列乘 k 加到第 j 列上 $(c_j + kc_i)$.

初等矩阵具有以下性质:

≡ 性质 3.1

初等矩阵都是可逆矩阵, 且其逆矩阵也是同类型的初等矩阵.

$$
\boldsymbol{E}(i,j)^{-1}=\boldsymbol{E}(i,j), \boldsymbol{E}(i(k))^{-1}=\boldsymbol{E}\left(i\left(\frac{1}{k}\right)\right), \boldsymbol{E}(i,j(k))^{-1}=\boldsymbol{E}(i,j(-k))
$$

≡ 性质 3.2

初等矩阵的转置仍是同类型的初等矩阵.

$$
\boldsymbol{E}^{\mathrm{T}}(i,j)=\boldsymbol{E}(i,j); \boldsymbol{E}^{\mathrm{T}}(i(k))=\boldsymbol{E}(i(k)); \boldsymbol{E}^{\mathrm{T}}\left(r_i+kr_j\right)=\boldsymbol{E}\left(r_j+kr_i\right)
$$

≡ 性质 3.3

对 $\boldsymbol{A}_{m\times n}$ 作一次初等行变换, 相当于对 $\boldsymbol{A}$ 左乘一个相应的 m 阶初等矩阵; 对 $\boldsymbol{A}_{m\times n}$ 作一次初等列变换, 相当于对 $\boldsymbol{A}$ 右乘一个相应的 n 阶初等矩阵.

口诀: 左乘行变, 右乘列变.

性质 3.3 表明

1) $\boldsymbol{E}(i,j)\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 作初等行变换 $r_i \leftrightarrow r_j$, $\boldsymbol{AE}(i,j)$ 相当于对 $\boldsymbol{A}$ 作初等列变换 $c_i \leftrightarrow c_j$;
2) $\boldsymbol{E}(i(k))\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 作初等行变换 kr_i, $\boldsymbol{AE}(i(k))$ 相当于对 $\boldsymbol{A}$ 作初等列变换 kc_i;
3) $\boldsymbol{E}(i,j(k))\boldsymbol{A}$ 相当于对 $\boldsymbol{A}$ 作初等行变换 $r_i + kr_j$, $\boldsymbol{AE}(i(k))$ 相当于对 $\boldsymbol{A}$ 作初等列变换 $c_j + kc_i$.

3.1.3 初等变换求逆矩阵

设 $\boldsymbol{A}$ 是 n 阶可逆矩阵, 由 **定理 3.1** 知, $\boldsymbol{A}$ 可经过有限次初等行变换变成单位矩阵, 根据初等矩阵的 **性质 3.3** , 说明存在一系列初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_s$, 使得 $\boldsymbol{P}_s \cdots \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A} = \boldsymbol{E}$, 两边同时右乘 $\boldsymbol{A}^{-1}$, 得 $\boldsymbol{P}_s \cdots \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{E} = \boldsymbol{A}^{-1}$.

比较这两个式子:

$$\boldsymbol{P}_s \cdots \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{A} = \boldsymbol{E}$$
$$\boldsymbol{P}_s \cdots \boldsymbol{P}_2\boldsymbol{P}_1\boldsymbol{E} = \boldsymbol{A}^{-1}$$

说明对 $\boldsymbol{A}, \boldsymbol{E}$ 作相同的初等行变换, 当 $\boldsymbol{A}$ 经过这些初等行变换化成单位矩阵 $\boldsymbol{E}$ 时, $\boldsymbol{E}$ 就变化成了 $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$. 这就是用初等行变换求逆矩阵的方法:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{E} \end{pmatrix} \xrightarrow{\text{初等行变换}} (\boldsymbol{E} \quad \boldsymbol{A}^{-1})$$

利用初等变换求逆矩阵的方法相比利用伴随矩阵求解在某些情况下更简洁, 将该方法熟练掌握, 要考.

用同样的方法可以得到用初等列变换求逆矩阵的方法:

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{E} \end{pmatrix} \xrightarrow{\text{初等列变换}} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{A}^{-1} \end{pmatrix}$$

3.1.4 用初等变换求解矩阵方程

矩阵方程的基本形式有下面三种:

1) $\boldsymbol{AX} = \boldsymbol{B}$, 当 $\boldsymbol{A}$ 可逆时, $\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{B}$;
2) $\boldsymbol{XA} = \boldsymbol{B}$, 当 $\boldsymbol{A}$ 可逆时, $\boldsymbol{X} = \boldsymbol{BA}^{-1}$;
3) $\boldsymbol{AXB} = \boldsymbol{C}$, 当 $\boldsymbol{A}, \boldsymbol{B}$ 可逆时, $\boldsymbol{X} = \boldsymbol{A}^{-1}\boldsymbol{CB}^{-1}$.

对于矩阵方程为 $\boldsymbol{AX} = \boldsymbol{B}$ 的形式, 若 $\boldsymbol{A}$ 可逆, 先求出 $\boldsymbol{A}^{-1}$, 再计算 $\boldsymbol{A}^{-1}\boldsymbol{B}$, 而计算两个矩阵乘积是很麻烦的. 下面介绍一种较简便的方法, 就是利用初等变换直接求出 $\boldsymbol{A}^{-1}\boldsymbol{B}$.

类似上面推导求逆矩阵 $\boldsymbol{A}^{-1}$ 的过程，若 $\boldsymbol{A}$ 可逆，由 **定理 3.1**，存在初等矩阵 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_s$，使得 $\boldsymbol{P}_s \cdots \boldsymbol{P}_2 \boldsymbol{P}_1 \boldsymbol{A} = \boldsymbol{E}$，两边同时右乘 $\boldsymbol{A}^{-1}\boldsymbol{B}$，得 $\boldsymbol{P}_s \cdots \boldsymbol{P}_2 \boldsymbol{P}_1 \boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{B}$. 比较这两个式子:

$$\boldsymbol{P}_s \cdots \boldsymbol{P}_2 \boldsymbol{P}_1 \boldsymbol{A} = \boldsymbol{E}$$
$$\boldsymbol{P}_s \cdots \boldsymbol{P}_2 \boldsymbol{P}_1 \boldsymbol{B} = \boldsymbol{A}^{-1}\boldsymbol{B}$$

说明对 $\boldsymbol{A}, \boldsymbol{B}$ 作相同的初等行变换，当 $\boldsymbol{A}$ 经过这些初等行变换化成单位矩阵 $\boldsymbol{E}$ 时，$\boldsymbol{B}$ 就变化成了矩阵方程的解 $\boldsymbol{A}^{-1}\boldsymbol{B}$. 由此，我们得到了一个用初等行变换求解矩阵方程的方法:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \end{pmatrix} \xrightarrow{\text{初等行变换}} (\boldsymbol{E} \quad \boldsymbol{A}^{-1}\boldsymbol{B})$$

矩阵方程与线性方程组等价. 这种方法也与高斯消元法求解线性方程组的方法等价, 多加练习掌握.

同理，利用初等列变换，也可求解矩阵方程 $\boldsymbol{X}\boldsymbol{A} = \boldsymbol{B}$. 即当 $\boldsymbol{A}$ 可逆时，

$$\begin{pmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{pmatrix} \xrightarrow{\text{初等列变换}} \begin{pmatrix} \boldsymbol{E} \\ \boldsymbol{B}\boldsymbol{A}^{-1} \end{pmatrix}$$

此时，$\boldsymbol{X} = \boldsymbol{B}\boldsymbol{A}^{-1}$ 就是矩阵方程 $\boldsymbol{X}\boldsymbol{A} = \boldsymbol{B}$ 的解.

例 3.3 已知 $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 2 \\ 0 & -1 & 0 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} -2 & 0 \\ -3 & 2 \\ 3 & -1 \end{pmatrix}$，求矩阵方程 $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{B}$ 的解.

3.2　矩阵的秩

矩阵的秩是矩阵理论中的一个重要概念，它不仅与可逆矩阵的讨论有着密切关系，而且在线性方程组的解的情况讨论中也有重要应用. 为引入该概念，我们先定义几种特殊矩阵.

3.2.1　行最简形矩阵

在矩阵 $\boldsymbol{A}_{m\times n}$ 中，如果某行元素全为 0，则称其为矩阵 $\boldsymbol{A}$ 的**零行**，否则称为**非零行**；在非零行中，从左至右数第一个不为零的元素称为**首非零元**.

定义 3.5: 行阶梯型矩阵

如果矩阵 $\boldsymbol{A}_{m\times n}$ 满足:

1) 若有零行，则零行位于矩阵的最下方;
2) 首非零元前面 0 的个数逐行严格增加. 则称矩阵 $\boldsymbol{A}_{m\times n}$ 为**行阶梯形矩阵** (简称为**阶梯形**).

★ **例子**

$$\begin{pmatrix} 3 & -2 & 0 & 2 \\ 0 & -1 & 4 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & 0 & -3 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 1 & 0 \\ 0 & 0 & -1 & 7 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

都是行阶梯形矩阵, 而

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

不是行阶梯形矩阵.

定义 3.6: 行最简形矩阵

设矩阵 $\boldsymbol{A}_{m\times n}$ 是行阶梯形矩阵, 利用初等行变换, 将 $\boldsymbol{A}$ 化成首非零元全为 1 , 且它所在列的其他元素都是 0 , 这样的矩阵称为**行最简形矩阵**或**最简形**.

★ **例子**

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

为行最简形矩阵.

行最简形矩阵是一个非常重要的概念. 在线性代数中, **很多问题的求解方法都会涉及将矩阵转化为行最简形矩阵**这一步.

3.2.2 矩阵的秩

矩阵的秩的概念

定义 3.7: 矩阵的 k 阶子式

在矩阵 $\boldsymbol{A}_{m\times n}$ 中任取 k 行 k 列交点上的 k^2 个元素, 按照原来的次序组成的 k 阶行列式, 称为**矩阵 $\boldsymbol{A}$ 的一个 k 阶子式**, 其中 $k \leqslant \min\{m, n\}$.

★ **例子**

矩阵 $\boldsymbol{A}=\begin{pmatrix}1&3&0&0&4\\0&0&5&0&-1\\0&0&0&6&2\\0&0&0&0&0\end{pmatrix}$, 取 $\boldsymbol{A}$ 的第 1、2、3 行, 第 2、3、4 列相交处的元素可组成 $\boldsymbol{A}$ 的一个 3 阶子式 $\begin{vmatrix}3&0&0\\0&5&0\\0&0&6\end{vmatrix}$, 取 $\boldsymbol{A}$ 的第 2、3 行, 第 4 、5 列相交处的元素可组成 $\boldsymbol{A}$ 的一个 2 阶子式 $\begin{vmatrix}0&-1\\6&2\end{vmatrix}$.

显然, 矩阵 $\boldsymbol{A}_{m\times n}$ 的一阶子式为 $\boldsymbol{A}_{m\times n}$ 的元素, 共有 $m\times n$ 个; k 阶子式共有 $\mathrm{C}_m^k\mathrm{C}_n^k$ 个 k 阶子式; 方阵 $\boldsymbol{A}_{n\times n}$ 的 n 阶子式只有一个, 即 $\boldsymbol{A}_{n\times n}$ 的行列式 $|\boldsymbol{A}_{n\times n}|$; $n-1$ 阶子式是 $\boldsymbol{A}_{n\times n}$ 的余子式.

定义 3.8: 矩阵的秩

矩阵 $\boldsymbol{A}$ 中的非零子式的最高阶数称为矩阵 $\boldsymbol{A}$ 的秩, 记作 $r(\boldsymbol{A})$ 或 $R(\boldsymbol{A})$. 规定零矩阵的秩为零, 即 $R\left(\boldsymbol{O}_{m\times n}\right)=0$.

定义 3.8 表明: 当 $R\left(\boldsymbol{A}_{m\times n}\right)=r$ 时, 矩阵 $\boldsymbol{A}$ 中有一个 r 阶子式不等于 0 , 而**所有的 $r+1$ 阶子式 (若存在) 均为零**.

★ **例子**

从矩阵 $\boldsymbol{A}=\begin{pmatrix}1&3&0&0&4\\0&0&5&0&-1\\0&0&0&6&2\\0&0&0&0&0\end{pmatrix}$ 中可取一阶、二阶、三阶、四阶子式. 易知, 四阶子式有 5 个, 且均为零, 而三阶子式有 $\begin{vmatrix}3&0&0\\0&5&0\\0&0&6\end{vmatrix}$ 非零, 也就是说矩阵 $\boldsymbol{A}$ 中的非零子式的最高阶数为 3 , 即 $R(\boldsymbol{A})=3$.

计算矩阵的秩

根据 **定义 3.8** , 计算矩阵的秩, 只需求出矩阵非零子式的最高阶数即可.

从解题过程能总结出如下规律: **行阶梯形矩阵的秩等于它的非零行的行数或首非零元的个数**.

对于一般矩阵, 如何找到简单的方法计算矩阵的秩呢?

为了解决这个问题, 有如下定理.

定理 3.3

矩阵的初等变换不改变矩阵的秩.

定理 3.4

n 阶方阵 $\boldsymbol{A}$ 可逆的充要条件是 $R(\boldsymbol{A})=n$.

定理 3.3 给出了求矩阵的秩的一般方法, 即**用初等变换将矩阵化为阶梯形矩阵**, 从而等到矩阵的秩.

例 3.4 设 $\boldsymbol{A}=\begin{pmatrix}1 & -1 & 0\\ 0 & 2 & 3\\ 1 & 0 & 1\end{pmatrix}$, 求 $R(\boldsymbol{A})$.

例 3.5 设 $\boldsymbol{A}=\begin{pmatrix}1 & 3 & -1 & -2\\ 2 & -1 & 2 & 3\\ 3 & 2 & 1 & 1\\ 1 & -4 & 3 & 5\end{pmatrix}$, 求 $R(\boldsymbol{A})$.

矩阵的秩的相关结论和性质

根据矩阵的秩的 **定义 3.8** , 可以得到如下结论:

1) 若矩阵 $\boldsymbol{A}$ 的所有 r 阶子式全为 0 , 则 $R(\boldsymbol{A})<r$.
2) 若矩阵 $\boldsymbol{A}$ 的所有 r 阶子式全不为 0 , 则 $R(\boldsymbol{A})\geqslant r$.

矩阵的秩具有以下性质:

性质 3.4: 矩阵的秩的性质

设矩阵 $\boldsymbol{A}_{m\times n},\boldsymbol{B}_{m\times n}$, 有

1) $0\leqslant R(\boldsymbol{A})\leqslant \min\{m,n\}$;
2) $R(\boldsymbol{A})=R\left(\boldsymbol{A}^{\mathrm{T}}\right)$;
3) 若 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价, 即 $\boldsymbol{A}\cong\boldsymbol{B}$, 则 $R(\boldsymbol{A})=R(\boldsymbol{B})$;
4) 若矩阵 $\boldsymbol{P}_{m\times m},\boldsymbol{Q}_{n\times n}$ 可逆, 则 $R(\boldsymbol{PAQ})=R(\boldsymbol{A})$;
5) $R(\boldsymbol{A}+\boldsymbol{B})\leqslant R(\boldsymbol{A})+R(\boldsymbol{B})$.

定理 3.5

设矩阵 $\boldsymbol{A}_{m\times n},\boldsymbol{B}_{n\times s}$, 则有 $R(\boldsymbol{AB})\leqslant\min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$.

3.3 线性方程组的解

3.3.1 非齐次线性方程组的判断与求解

设包含 m 个方程的 n 元线性方程组

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

其矩阵形式为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

如果其中常数项 $b_1, b_2, \cdots, b_m$ 不全为零, 此线性方程组称为**非齐次线性方程组**. 矩阵 $\boldsymbol{A}$ 为方程组的**系数矩阵**, 矩阵 $\boldsymbol{B} = (\boldsymbol{A}, \boldsymbol{b})$ 为方程组的**增广矩阵**.

非齐次线性方程组的解有三种情况: (1) 无解 (也称方程组不相容); (2) 有唯一解; (3) 有无限多解. 利用如下定理可以判断解的情况.

定理 3.6

对于 n 元非齐次线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

1) 无解当且仅当 $R(\boldsymbol{A}) < R(\boldsymbol{A}, \boldsymbol{b})$;

2) 有唯一解当且仅当 $R(\boldsymbol{A}) = R(\boldsymbol{A}, \boldsymbol{b}) = n$;

3) 有无限多解当且仅当 $R(\boldsymbol{A}) = R(\boldsymbol{A}, \boldsymbol{b}) < n$.

例 3.6 判断方程组是否有解.

$$\begin{cases} 2x_1 + 4x_2 + 6x_3 + 2x_4 = 2, \\ 3x_1 + x_2 + x_3 + 2x_4 = 0, \\ -2x_1 + x_2 + 2x_3 + 2x_4 = 3. \end{cases}$$

例 3.7 判断方程组是否有解, 如果有解求出方程组的所有解.

$$\begin{cases} 2x_1 + x_2 + 6x_4 + 6x_5 = 7, \\ x_1 - x_3 + 4x_4 + 2x_5 = 6, \\ -3x_1 + 3x_3 - 12x_4 - 2x_5 = -10. \end{cases}$$

3.3.2　齐次线性方程组的判断与求解

设包含 m 个方程的 n 元线性方程组

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & 0, \\ \cdots\cdots\cdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0, \end{cases}$$

其矩阵形式为

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}.$$

这类常数项全为零的线性方程组称为**齐次线性方程组**.

齐次线性方程组是非齐次线性方程组的特殊形式. 因为未知数全为零一定是齐次线性方程组的解, 所以, 齐次线性方程组不存在无解情况, 其解有两种情况: (1) 只有零解; (2) 有非零解 (此时一定有无限多解). 由 **定理 3.6** 可得下面的定理.

定理 3.7

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

1) 只有零解当且仅当 $R(\boldsymbol{A}) = n$;
2) 有非零解当且仅当 $R(\boldsymbol{A}) < n$.

例 3.8 问 λ 取何值时, 齐次线性方程组

$$\begin{cases} (1-\lambda)x_1 - 2x_2 + 4x_3 = 0, \\ 2x_1 + (3-\lambda)x_2 + x_3 = 0, \\ x_1 + x_2 + (1-\lambda)x_3 = 0 \end{cases}$$

有非零解?

第 四 章　向量组的线性相关性

前不见古人, 后不见来者, 念天地之悠悠, 独怆然而涕下.

唐·陈子昂 (661–702) 《登幽州台歌》

本章关注以下几个问题:

1) 从解析几何中学过的概念引入 n 维向量、n 维向量空间、向量组的线性组合、线性相关和线性无关概念.
2) 向量组线性相关、线性无关的性质和判别方法.
3) 向量组的最大无关组、向量组秩的概念以及如何求解.
4) 如何利用向量组的相关性理论得到线性方程组的解.
5) 向量空间的定义. 在向量空间的基础上引入欧氏空间的概念, 此空间定义了向量的内积运算, 由此引入向量长度、向量夹角、向量正交化的概念.

本章内容集中展示了线性代数这一学科的**高度抽象性**, 概念和定理多, 证明繁琐, 学习难度较大. 在学习时, 可以利用解析几何中学到的二维和三维几何中的知识理解引入的抽象概念, 弄清楚概念的准确内涵以及概念之间的联系.

4.1 n 维向量及其线性运算

在物理学中, 建立三维直角坐标系后, 空间任一质点的坐标可由一个三元有序数组 (x, y, z) 描述, 可与位置矢量 (向量) 一一对应. 在许多实际问题中, 所研究的对象需要用多个数构成的有序数组来描述. 如果将维度推广到 n 维, 可以引入 n 维向量、n 维向量空间的概念.

4.1.1 n 维向量的定义

定义 4.1: n 维向量、行向量、列向量

一组有序的 n 个数 $(x_1, x_2, \cdots, x_n)$ 称为一个 n **维向量**, 其中数 $x_i(i = 1, 2, \cdots, n)$ 称为该向量的第 i 个分量. n 维向量记为 $\boldsymbol{x} =$

$(x_1, x_2, \cdots, x_n)$. 有时也可写成

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

为了区别起见, 前者称为**行向量**, 后者称为**列向量**.

不难注意到 n 维行向量即为 $1 \times n$ 矩阵, 而 n 维列向量即为 $n \times 1$ 矩阵. 根据矩阵的运算, 我们可定义向量 $\boldsymbol{x}$ 的转置与矩阵的转置类似, 记为 $\boldsymbol{x}^{\mathrm{T}}$.

这样 $\boldsymbol{x}$ 等向量可用于表示列向量, $\boldsymbol{x}^{\mathrm{T}}$ 等可用于表示行向量, 以简化向量的表示.

如果两个 n 维向量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n), \boldsymbol{y} = (y_1, y_2, \cdots, y_n)$ 的对应分量相等, 即 $x_i = y_i (i = 1, 2, \cdots, n)$, 则称此两向量是**相等**的, 记作 $\boldsymbol{x} = \boldsymbol{y}$.

如果分量为实数, 则向量称为**实向量**; 如果分量为复数, 则向量称为**复向量**. 除非特别说明, 讲义所述向量均为实向量.

定义 4.2: 向量和、向量数乘、零向量、负向量、线性运算

两向量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n), \boldsymbol{y} = (y_1, y_2, \cdots, y_n)$ 的和 $\boldsymbol{x} + \boldsymbol{y}$ 规定为

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n).$$

向量 $\boldsymbol{x}$ 乘以实数 λ 的乘积 $\lambda\boldsymbol{x}$ 称为 **向量的数乘**, 规定为

$$\lambda\boldsymbol{x} = (\lambda x_1, \lambda x_2, \cdots, \lambda x_n).$$

分量全为 0 的向量 $(0, 0, \cdots, 0)$ 称为**零向量**, 记作 $\mathbf{0}$, 即

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

向量 $(-x_1, -x_2, \cdots, -x_n)$ 称为 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$ 的**负向量**, 记作 $-\boldsymbol{x}$.

向量的加法和数乘统称为**向量的线性运算**.

4.1.2 向量的线性运算性质

根据矩阵的运算规律, 向量的线性运算满足如下八条性质.

1) $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$

2) $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$;

3) $\boldsymbol{x} + 0 = \boldsymbol{x}$

4) $\boldsymbol{x} + (-\boldsymbol{x}) = 0$;

5) $1\boldsymbol{x} = \boldsymbol{x}$

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

7) $(\lambda+\mu)\boldsymbol{x}=\lambda\boldsymbol{x}+\mu\boldsymbol{x}$;

8) $\lambda(\boldsymbol{x}+\boldsymbol{y})=\lambda\boldsymbol{x}+\lambda\boldsymbol{y}$.

4.2 向量组的线性关系

定义 4.3: 向量组

由若干个同维数的行向量 (或同维数的列向量) 组成的集合称为**向量组**.

对于一个 $m\times n$ 矩阵 $\boldsymbol{A}=(a_{ij})$, 它的 n 个 m 维列向量

$$\boldsymbol{\alpha}_i=(a_{1i},a_{2i},\cdots,a_{mi})^{\mathrm{T}}\,(i=1,2,\cdots,n)$$

构成的向量组, 称为 $\boldsymbol{A}$ 的**列向量组**; 同理, m 个 n 维行向量 $\beta_i^{\mathrm{T}}=(a_{i1},a_{i2},\cdots,a_{in})\,(i=1,2,\cdots,m)$ 构成的向量组, 称为 $\boldsymbol{A}$ 的**行向量组**.

由此, 矩阵 $\boldsymbol{A}$ 可记为

$$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_n)=\begin{pmatrix}\beta_1^{\mathrm{T}}\\ \beta_2^{\mathrm{T}}\\ \vdots\\ \beta_m^{\mathrm{T}}\end{pmatrix}.$$

所以, 矩阵 $\boldsymbol{A}$ 与其列向量组或行向量组之间建立了**一一对应**关系.

4.2.1 向量组的线性组合

在三维向量空间 $\mathbb{R}^3$ 中, 在基底 $\boldsymbol{i}=(1,0,0)^{\mathrm{T}},\boldsymbol{j}=(0,1,0)^{\mathrm{T}},\boldsymbol{k}=(0,0,1)^{\mathrm{T}}$ 下, 任意一个向量 $\boldsymbol{\alpha}=(x,y,z)^{\mathrm{T}}$ 可表示为 $\boldsymbol{\alpha}=x\boldsymbol{i}+y\boldsymbol{j}+z\boldsymbol{k}$.

下面, 我们将这样的概念推广到 n 维向量空间 $\mathbb{R}^n$ 中来.

定义 4.4: 线性组合

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 是 n 维向量, $k_1,k_2,\cdots,k_s\in\mathbb{R}$, 则称

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s$$

是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的一个 **线性组合**.

如果对向量 β, 存在一组实数 $k_1,k_2,\cdots,k_s$, 使

$$\beta=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s,$$

则称 β 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的一个线性组合, 也称 β 可由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ **线性表出**.

★ **例子**

对 $\boldsymbol{\alpha}_1=(1,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,1)^{\mathrm{T}},\boldsymbol{\beta}=(-3,2)^{\mathrm{T}}$, 有 $\boldsymbol{\beta}=-5\boldsymbol{\alpha}_1+2\boldsymbol{\alpha}_2$. 即 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 的一个线性组合.

对 $\boldsymbol{\alpha}_1=(0,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,2)^{\mathrm{T}},\boldsymbol{\alpha}_3=(-2,4)^{\mathrm{T}},\boldsymbol{\beta}=(3,5)^{\mathrm{T}}$, 有 $\boldsymbol{\beta}=-\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2+0\boldsymbol{\alpha}_3$, 以及 $\boldsymbol{\beta}=7\boldsymbol{\alpha}_1+\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3$, 还有 $\boldsymbol{\beta}=11\boldsymbol{\alpha}_1+0\boldsymbol{\alpha}_2-\frac{3}{2}\boldsymbol{\alpha}_3$ 等等, 故 $\boldsymbol{\beta}$ 是 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 的线性组合, 而且线性表出的方式**不是唯一**的.

对 $\boldsymbol{\alpha}_1=(2,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(3,0)^{\mathrm{T}},\boldsymbol{\beta}=(0,1)^{\mathrm{T}}$, 无论 k_1,k_2 取什么数, $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2\neq\boldsymbol{\beta}$, 即 $\boldsymbol{\beta}$**不能**由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2$ 线性表出.

向量 $\boldsymbol{b}$ 能由向量组 A 线性表示, 也就是方程组

$$x_1\boldsymbol{a}_1+x_2\boldsymbol{a}_2+\cdots+x_m\boldsymbol{a}_m=\boldsymbol{b}$$

有解. 由上章定理, 立即可得

定理 4.1

向量 $\boldsymbol{b}$ 能由向量组 $A:\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m$ 线性表示的充分必要条件是矩阵 $\boldsymbol{A}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m)$ 的秩等于矩阵 $\boldsymbol{B}=(\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_m,\boldsymbol{b})$ 的秩.

例 4.1 已知 $\boldsymbol{\alpha}_1=(-1,0,1,2)^{\mathrm{T}},\boldsymbol{\alpha}_2=(3,4,-2,5)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,4,0,9)^{\mathrm{T}},\boldsymbol{\beta}=(5,4,-4,1)^{\mathrm{T}}$. 试问 $\boldsymbol{\beta}$ 能否由 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性表出? 如果能线性表出就写出其表达式.

通过此例题可理解向量能否由向量组线性表出等价于相应的线性方程组是否有解.

4.2.2 向量组的等价

定义 4.5: 向量组的等价

设有两个向量组: (I) $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$; (II) $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$. 若向量组 (I) 中每个向量都可由向量组 (II) 线性表示, 则称向量组 (I) 可由向量组 (II) 线性表示. 若两个向量组可相互线性表示, 则称它们**等价**.

★ **例子**

例如, $\boldsymbol{e}_1=(1,0,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,0)^{\mathrm{T}},\boldsymbol{e}_3=(0,0,1)^{\mathrm{T}}$ 和 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,1,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,0,0)^{\mathrm{T}}$ 这两个向量组是等价的.

原因如下.

$\boldsymbol{\alpha}_1=\boldsymbol{e}_1+\boldsymbol{e}_2+\boldsymbol{e}_3,\boldsymbol{\alpha}_2=\boldsymbol{e}_1+\boldsymbol{e}_2,\boldsymbol{\alpha}_3=\boldsymbol{e}_1$, 又容易得 $\boldsymbol{e}_1=\boldsymbol{\alpha}_3,\boldsymbol{e}_2=\boldsymbol{\alpha}_2-\boldsymbol{\alpha}_3,\boldsymbol{e}_3=\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2$, 即两个向量组等价.

不难证明, 向量组等价满足下列 3 条性质.

性质 4.1: 向量组等价的性质

1) 反身性: 每一个向量组都与其自身等价.
2) 对称性: 若向量组 (I) 与 (II) 等价, 则向量组 (II) 与 (I) 也等价.

3) 传递性: 若向量组 (I) 与 (II) 等价, 向量组 (II) 与 (III) 等价, 则向量组 (I) 与 (III) 等价.

定义 4.6: 表示矩阵

由 **定义 4.5** , 若向量组 (II) $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 可由向量组 (I) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表示, 则存在一组数

$$c_{1i}, c_{2i}, \cdots, c_{mi}\,(i=1,\cdots,s)\,,$$

使

$$\begin{aligned}\boldsymbol{\beta}_i &= c_{1i}\boldsymbol{\alpha}_1 + c_{2i}\boldsymbol{\alpha}_2 + \cdots + c_{mi}\boldsymbol{\alpha}_m \\ &= (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)\begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{mi}\end{pmatrix},\end{aligned}$$

从而

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)\,\boldsymbol{C},$$

其中 $m \times s$ 矩阵 $\boldsymbol{C} = (c_{ij})$ 称为这一线性表示过程的**表示矩阵**.

令 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)\,, \boldsymbol{B} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s)$, 则有 $\boldsymbol{AC} = \boldsymbol{B}$. 进而有以下定理.

定理 4.2

向量组 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 可由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表示的充要条件是矩阵方程 $\boldsymbol{AX} = \boldsymbol{B}$ 有解, 也即 $R(\boldsymbol{A}) = R(\boldsymbol{A}, \boldsymbol{B})$.

推论 4.1

向量组 $A : \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 与向量组 $B : \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s$ 等价的充分必要条件是

$$R\,(\boldsymbol{\alpha}) = R\,(\boldsymbol{\beta}) = R\,(\boldsymbol{\alpha}, \boldsymbol{\beta})\,,$$

其中 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 是向量组 A 和 B 所构成的矩阵.

定理 4.3

行向量组 $\boldsymbol{\beta}_1^{\mathrm{T}}, \boldsymbol{\beta}_2^{\mathrm{T}}, \cdots, \boldsymbol{\beta}_s^{\mathrm{T}}$ 可由行向量组 $\boldsymbol{\alpha}_1^{\mathrm{T}}, \boldsymbol{\alpha}_2^{\mathrm{T}}, \cdots, \boldsymbol{\alpha}_m{}^{\mathrm{T}}$ 线性表示的充要条件是矩阵方程 $\boldsymbol{XA}^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}$ 有解.

定理 4.4

设向量组 $B : \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_l$ 能由向量组 $A : \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表示, 则 $R\,(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_s) \leq R\,(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m)$.

总结

向量组 $B:\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 能由向量组 $A:\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性表示. (几何语言)

$\Longleftrightarrow$ 有矩阵 $\boldsymbol{K}$, 使 $\boldsymbol{\beta}=\boldsymbol{\alpha}\boldsymbol{K}$. (矩阵语言)

$\Longleftrightarrow$ 方程 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{\beta}$ 有解.

$\Longleftrightarrow$ $R(\boldsymbol{\alpha})=R(\boldsymbol{\alpha},\boldsymbol{\beta})$

4.2.3　向量组线性相关性的定义

考虑 $\mathbb{R}^2$ 的两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$, 如果两者共线, 则必然存在实数 $\lambda\in\mathbb{R}$, 使 $\boldsymbol{\alpha}=\lambda\boldsymbol{\beta}$ 或 $\boldsymbol{\beta}=\lambda\boldsymbol{\alpha}$.

经过变形移项后, 有: 若两个向量 $\boldsymbol{\alpha},\boldsymbol{\beta}$ 共线 $\Longleftrightarrow$ 存在不全为 0 的实数 k_1,k_2, 使 $k_1\boldsymbol{\alpha}+k_2\boldsymbol{\beta}=\mathbf{0}$.

类似地, 如果 $\mathbb{R}^3$ 的三个向量 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 共面, 则 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 中某一个可被另外两个线性表出 $\Longleftrightarrow$ 存在不全为 0 的实数 λ_1, λ_2,λ_3, 使得 $\lambda_1\boldsymbol{\alpha}+\lambda_2\boldsymbol{\beta}+\lambda_3\boldsymbol{\gamma}=\mathbf{0}$.

下面, 我们将共线、共面的概念推广到一般 n 维空间中. 此时, 几何意义很难想象, 所以主要从代数的角度来刻画.

定义 4.7: 线性相关、线性无关

给定向量空间 $\mathbb{R}^n$ 中 s 个向量 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$, 如果存在不全为 0 的 $k_1,k_2,\cdots,k_s\in F$, 使得 $k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_s\boldsymbol{\alpha}_s=\mathbf{0}$, 则称向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ **线性相关** . 否则称 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ **线性无关** .

根据定义, 在 $\mathbb{R}^3$ 中, 共线的向量及共面的向量都是线性相关的.

例子

$\boldsymbol{\alpha}_1=\begin{pmatrix}1\\2\\1\end{pmatrix},\boldsymbol{\alpha}_2=\begin{pmatrix}2\\4\\2\end{pmatrix},\boldsymbol{\alpha}_3=\begin{pmatrix}1\\3\\5\end{pmatrix}$ 是 $\mathbb{R}^3$ 中三个向量, 由于 $\boldsymbol{\alpha}_2=2\boldsymbol{\alpha}_1$, 因而有 $2\boldsymbol{\alpha}_1+(-1)\boldsymbol{\alpha}_2+0\boldsymbol{\alpha}_3=\mathbf{0}$, 系数 $2,-1,0$ 不全为 0 , 按定义 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3$ 线性相关.

4.2.4　向量组线性相关性的判定

定理 4.5

向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性相关的充分必要条件是它所构成的矩阵 $\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$ 的秩小于向量个数 m ; 向量组线性无关的充分必要条件是 $R(\boldsymbol{A})=m$.

推论 4.2

任意 n 个 n 维向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n$ 线性无关的充要条件是矩阵 $\boldsymbol{A} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_n)$ 的行列式不等于零.

由 **定理 4.5** 可进一步证明得到如下定理.

定理 4.6

(1) 若向量组 $A: \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$ 线性相关, 则向量组 $B: \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m$, $\boldsymbol{\alpha}_{m+1}$ 也线性相关. 反言之, 若向量组 B 线性无关, 则向量组 A 也线性无关.

(2) m 个 n 维向量组成的向量组, 当维数 n 小于向量个数 m 时一定线性相关. 特别地, $n+1$ 个 n 维向量一定线性相关.

(3) 设向量组 $A: \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关, 而向量组 $B: \boldsymbol{\alpha}_1, \cdots, \boldsymbol{\alpha}_m, \boldsymbol{\beta}$ 线性相关, 则向量 $\boldsymbol{\beta}$ 必能由向量组 A 线性表示, 且表示式是**唯一**的.

例 4.2 已知向量 $\boldsymbol{\alpha}_1 = (1, 2, -1, 3)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (2, -1, 3, 5)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (-1, a+17, a, -1)^{\mathrm{T}}$, 问: a 为何值时, 向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 线性相关、线性无关?

4.3 向量组的秩

4.3.1 极大线性无关组

定义 4.8: 极大线性无关组

在向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 中, 若存在 r 个向量 $\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i, \cdots, \boldsymbol{\alpha}_i$ 线性无关, 再加进任意一个向量 $\boldsymbol{\alpha}_j (j = 1, 2, \cdots, s)$ 就线性相关, 则称 $\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_i, \cdots, \boldsymbol{\alpha}_i$ 是向量组 $\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的一个**极大线性无关组**.

例子

向量组 $\boldsymbol{\alpha}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \boldsymbol{\alpha}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \boldsymbol{\alpha}_3 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ 中, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 是线性无关的, 而 $\boldsymbol{\alpha}_3 = \dfrac{1}{2}\boldsymbol{\alpha}_1 + \dfrac{3}{2}\boldsymbol{\alpha}_2$, 即添加 $\boldsymbol{\alpha}_3$ 后就线性相关, 所以, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ 就是一个极大线性无关组.

注意

若向量组是线性无关的, 那么极大无关组是唯一的, 就是向量组本身. 若向量组是线性相关的, 则极大无关组不一定唯一. 例如向量组 $\boldsymbol{\alpha}_1 = (1, 0)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (0, 1)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (2, 3)^{\mathrm{T}}$ 是线性相关的, 任意两个向量线性无关, 所以, 任意两个向量都是极大无关组.

由极大线性无关组的定义还可以得到, 向量组和它的任意一个极大无关组都是等价的, 进而同一向量组的任意两个极大无关组是等价的. 当向量组线性相关时, 其极大无关组不唯一, 但是由 **定理 4.6** , 极

大无关组所含向量的个数是**唯一**的.

定义 4.9: 向量组的秩

向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 的极大无关组所含向量的个数称为**向量组的秩**, 记为 $R(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$.

只含有零向量的向量组没有极大无关组, 规定**秩为零**.

由定义可知, 向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关的充要条件是秩等于 m, 线性相关的充要条件是秩小于 m .

定理 4.7

等价的向量组有相同的秩.

注意

若两个向量组的秩相同, 这两个向量组未必等价. 例如, 一个向量 $\boldsymbol{\alpha}=(1,0)^{\mathrm{T}}$ 构成的向量组的秩为 1, 一个向量 $\boldsymbol{\beta}=(0,1)^{\mathrm{T}}$ 构成的向量组的秩也为 1, 但是显然二者不等价.

4.3.2 向量组的秩和矩阵的秩的关系

对于特殊的向量组, 比如, 全体 n 维向量构成的向量组, 记为 $\mathbb{R}^n$, 则单位向量组 $\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{e}_n=(0,0,\cdots,1)^{\mathrm{T}}$ 为 $\mathbb{R}^n$ 的一个极大无关组, $\mathbb{R}^n$ 的秩为 n . 那么, 对于一般的向量组, 如何求出它的极大无关组和秩呢?

定义 4.10: 行秩、列秩

对于 $m\times n$ 矩阵 $\boldsymbol{A}$, 我们称 $\boldsymbol{A}$ 的 m 个 n 维行向量构成的向量组为 $\boldsymbol{A}$ 的行向量组, 称 n 个 m 维列向量构成的向量组为 $\boldsymbol{A}$ 的列向量组, 并分别称它们的秩为 $\boldsymbol{A}$ 的**行秩**和**列秩**.

定理 4.8

设 $\boldsymbol{A}$ 为 $m\times n$ 矩阵, 则 $\boldsymbol{A}$ 的秩等于 $\boldsymbol{A}$ 的行秩, 也等于 $\boldsymbol{A}$ 的列秩.

推论 4.3

若 D_r 是矩阵 $\boldsymbol{A}$ 的一个最高阶非零子式, 则 D_r 所在的 r 列就是 $\boldsymbol{A}$ 的列向量组的一个极大无关组, 而 D_r 所在的 r 行就是 $\boldsymbol{A}$ 的行向量组的一个极大无关组.

例 4.3 求向量组 $\boldsymbol{\alpha}_1=(1,1,4)^{\mathrm{T}},\boldsymbol{\alpha}_2=(1,0,4)^{\mathrm{T}},\boldsymbol{\alpha}_3=(1,2,4)^{\mathrm{T}},\boldsymbol{\alpha}_4=(1,3,4)^{\mathrm{T}}$ 的秩和一个极大无关组, 并将其余向量用该极大无关组线性表示.

可利用向量组的极大无关组和秩证明得到矩阵的秩的两个重要性质.

☰ 性质 4.2

设 $\boldsymbol{A},\boldsymbol{B}$ 为 $m\times n$ 矩阵, 则 $R(\boldsymbol{A}+\boldsymbol{B})\leq R(\boldsymbol{A})+R(\boldsymbol{B})$.

☰ 性质 4.3

设 $\boldsymbol{A}$ 为 $m\times s$ 矩阵, $\boldsymbol{B}$ 为 $s\times n$ 矩阵, 则 $R(\boldsymbol{AB})\leq \min\{R(\boldsymbol{A}),R(\boldsymbol{B})\}$.

4.4 线性方程组的解的结构

在上一章中, 我们已经介绍了用矩阵的初等变换解线性方程组的方法, 并建立了两个重要定理, 即

1) n 个未知数的齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{0}$ 有非零解的充分必要条件是系数矩阵的秩 $R(\boldsymbol{A})<n$.

2) n 个未知数的非齐次线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 有解的充分必要条件是系数矩阵 $\boldsymbol{A}$ 的秩等于增广矩阵 $\boldsymbol{B}$ 的秩, 且当 $R(\boldsymbol{A})=R(\boldsymbol{B})=n$ 时方程组有唯一解, 当 $R(\boldsymbol{A})=R(\boldsymbol{B})=r<n$ 时方程组有无限多个解.

下面我们用向量组线性相关性的理论来讨论线性方程组的解. 先讨论齐次线性方程组.

4.4.1 齐次线性方程组的解的结构

设有齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0, \\ \cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0, \end{cases} \tag{4-1}$$

记

$$\boldsymbol{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \boldsymbol{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

则方程组 (4-1) 可写成向量方程

$$\boldsymbol{Ax}=\boldsymbol{0}. \tag{4-2}$$

若 $x_1=\xi_{11},x_2=\xi_{21},\cdots,x_n=\xi_{n1}$ 为方程组 (4-1) 的解, 则

$$\boldsymbol{x}=\boldsymbol{\xi}_1=\begin{pmatrix} \xi_{11} \\ \xi_{21} \\ \vdots \\ \xi_{n1} \end{pmatrix}$$

称为方程组 (4-1) 的**解向量**, 它也就是**向量方程** (4-2) 的解.

根据向量方程 (4-2), 我们来讨论解向量的性质.

性质 4.4

若 $\boldsymbol{x}=\boldsymbol{\xi}_1, \boldsymbol{x}=\boldsymbol{\xi}_2$ 为 (4-2) 的解, 则 $\boldsymbol{x}=\boldsymbol{\xi}_1+\boldsymbol{\xi}_2$ 也是 (4-2) 的解.

证明: 只要验证 $\boldsymbol{x}=\boldsymbol{\xi}_1+\boldsymbol{\xi}_2$ 满足方程 (4-2):

$$\boldsymbol{A}\left(\boldsymbol{\xi}_1+\boldsymbol{\xi}_2\right)=\boldsymbol{A}\boldsymbol{\xi}_1+\boldsymbol{A}\boldsymbol{\xi}_2=\mathbf{0}+\mathbf{0}=\mathbf{0}.$$

□

性质 4.5

若 $\boldsymbol{x}=\boldsymbol{\xi}_1$ 为 (4-2) 的解, k 为实数, 则 $\boldsymbol{x}=k\boldsymbol{\xi}_1$ 也是 (4-2) 的解.

证明:

$$\boldsymbol{A}\left(k\boldsymbol{\xi}_1\right)=k\left(\boldsymbol{A}\boldsymbol{\xi}_1\right)=k\mathbf{0}=\mathbf{0}.$$

□

定义 4.11: 向量方程的通解、基础解系

把方程 (4-2) 的全体解所组成的集合记作 S, 如果能求得解集 S 的一个最大无关组 $S_0: \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_t$, 那么方程 (4-2) 的任一解都可由最大无关组 S_0 线性表示; 另一方面, 由 **性质 4.4**、 **性质 4.5** 可知, 最大无关组 S_0 的任何线性组合

$$\boldsymbol{x}=k_1\boldsymbol{\xi}_1+k_2\boldsymbol{\xi}_2+\cdots+k_t\boldsymbol{\xi}_t \quad (k_1, k_2, \cdots, k_t\text{为任意实数})$$

都是方程 (4-2) 的解, 因此上式便是方程 (4-2) 的**通解**.
齐次线性方程组的解集的**最大无关组**称为该齐次线性方程组的**基础解系**.

由上面的讨论可知, **要求齐次线性方程组的通解, 只需求出它的基础解系**.

上一章我们用初等变换的方法求线性方程组的通解, 下面我们用同一方法来求齐次线性方程组的基础解系.

设方程组 (4-1) 的系数矩阵 $\boldsymbol{A}$ 的秩为 r, 不妨设 $\boldsymbol{A}$ 的前 r 个列向量线性无关, 于是 $\boldsymbol{A}$ 的行最简形矩阵为

$$\boldsymbol{B}=\begin{pmatrix} 1 & \cdots & 0 & b_{11} & \cdots & b_{1,n-r} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & b_{r1} & \cdots & b_{r,n-r} \\ 0 & & & \cdots & & 0 \\ \vdots & & & & & \vdots \\ 0 & & & \cdots & & 0 \end{pmatrix},$$

与 $\boldsymbol{B}$ 对应, 即有方程组

$$\begin{cases} x_1 = -b_{11}x_{r+1} - \cdots - b_{1,n-r}x_n, \\ \cdots\cdots\cdots\cdots \\ x_r = -b_{r1}x_{r+1} - \cdots - b_{r,n-r}x_n, \end{cases} \tag{4-3}$$

把 $x_{r+1}, \cdots, x_n$ 作为自由未知数, 并令它们依次等于 $c_1, \cdots, c_{n-r}$, 可得方程组 (4-1) 的通解

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \\ x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = c_1 \begin{pmatrix} -b_{11} \\ \vdots \\ -b_{r1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -b_{12} \\ \vdots \\ -b_{r2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + c_{n-r} \begin{pmatrix} -b_{1,n-r} \\ \vdots \\ -b_{r,n-r} \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

把上式记作

$$\boldsymbol{x} = c_1\boldsymbol{\xi}_1 + c_2\boldsymbol{\xi}_2 + \cdots + c_{n-r}\boldsymbol{\xi}_{n-r},$$

可知解集 S 中的任一向量 $\boldsymbol{x}$ 能由 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_{n-r}$ 线性表示, 又因为矩阵 $(\boldsymbol{\xi}_1\ , \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_{n-r})$ 中有 $n-r$ 阶子式 $|\boldsymbol{E}_{n-r}| \neq 0$. 故 $R(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_{n-r}) = n - r$, 所以 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_{n-r}$ 线性无关. 根据最大无关组的等价定义, 即知 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_{n-r}$ 是解集 S 的最大无关组, 即 $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \cdots, \boldsymbol{\xi}_{n-r}$ 是方程组 (4-1) 的基础解系.

总结

在上面的讨论中, 我们先求出齐次线性方程组的通解, 再从通解求得基础解系. 其实我们也可先求基础解系, 再写出通解. 这只需在得到方程组 (4-3) 以后, 令自由未知数 $x_{r+1}, x_{r+2}, \cdots, x_n$ 取下列 $n-r$ 组数

$$\begin{pmatrix} x_{r+1} \\ x_{r+2} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

由 (4-3) 即依次可得

$$\begin{pmatrix} x_1 \\ \vdots \\ x_r \end{pmatrix} = \begin{pmatrix} -b_{11} \\ \vdots \\ -b_{r1} \end{pmatrix}, \begin{pmatrix} -b_{12} \\ \vdots \\ -b_{r2} \end{pmatrix}, \cdots, \begin{pmatrix} -b_{1,n-r} \\ \vdots \\ -b_{r,n-r} \end{pmatrix},$$

合起来便得基础解系

$$\xi_1=\begin{pmatrix}-b_{11}\\ \vdots\\ -b_{r1}\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix},\xi_2=\begin{pmatrix}-b_{12}\\ \vdots\\ -b_{r2}\\ 0\\ 1\\ \vdots\\ 0\end{pmatrix},\cdots,\xi_{n-r}=\begin{pmatrix}-b_{1,n-r}\\ \vdots\\ -b_{r,n-r}\\ 0\\ 0\\ \vdots\\ 1\end{pmatrix}.$$

依据以上的讨论, 还可推得

定理 4.9

设 $m\times n$ 矩阵 $\boldsymbol{A}$ 的秩 $R(\boldsymbol{A})=r$, 则 n 元齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解集 S 的秩 $R_S=n-r$.

1) 当 $R(\boldsymbol{A})=n$ 时, 方程 (4-1) 只有零解, 没有基础解系 (此时解集 S 只含一个零向量);
2) 当 $R(\boldsymbol{A})=r<n$ 时, 由 **定理 4.9** 可知方程组 (4-1) 的基础解系含 $n-r$ 个向量. 因此, 由最大无关组的性质可知, 方程组 (4-1) 的任何 $n-r$ 个线性无关的解都可构成它的基础解系. 并由此可知齐次线性方程组的基础解系并不是唯一的, 它的通解的形式也不是唯一的.

例 4.4 求齐次线性方程组

$$\begin{cases}x_1+x_2-x_3-x_4=0,\\ 2x_1-5x_2+3x_3+2x_4=0,\\ 7x_1-7x_2+3x_3+x_4=0\end{cases}$$

的基础解系与通解.

定理 4.9 不仅是线性方程组各种解法的理论基础, 在讨论向量组的线性相关性时也很有用.

例 4.5 设 $\boldsymbol{A}_{m\times n}\boldsymbol{B}_{n\times l}=\boldsymbol{O}$, 证明 $R(\boldsymbol{A})+R(\boldsymbol{B})\leq n$.

证明: 记 $\boldsymbol{B}=(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l)$, 则

$$\boldsymbol{A}(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_l)=(\boldsymbol{0},\boldsymbol{0},\cdots,\boldsymbol{0}),$$

即 $\boldsymbol{A}\boldsymbol{b}_i=\boldsymbol{0}(i=1,2,\cdots,l)$, 表明矩阵 $\boldsymbol{B}$ 的 l 个列向量都是齐次方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解. 记方程 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解集为 S , 由 $\boldsymbol{b}_i\in S$, 知有 $R(\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_t)\leq R_S$, 即 $R(\boldsymbol{B})\leq R_S$. 而由 **定理 4.9** 有 $R(\boldsymbol{A})+R_S=n$, 故 $R(\boldsymbol{A})+R(\boldsymbol{B})\leq n$. □

例 4.6 设 n 元齐次线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 与 $\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0}$ 同解, 证明

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

证明: 由于方程组 $\boldsymbol{Ax}=\mathbf{0}$ 与 $\boldsymbol{Bx}=\mathbf{0}$ 有相同的解集, 设为 S , 则由 **定理 4.9** 即有 $R(\boldsymbol{A})=n-R_S, R(\boldsymbol{B})=n-R_S$. 因此 $R(\boldsymbol{A})=R(\boldsymbol{B})$.

本例的结论表明, 当矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的列数相等时, 要证 $R(\boldsymbol{A})=R(\boldsymbol{B})$, 只需证明齐次方程 $\boldsymbol{Ax}=\mathbf{0}$ 与 $\boldsymbol{Bx}=\mathbf{0}$ 同解. □

4.4.2　非齐次线性方程组的解的结构

下面讨论非齐次线性方程组.

设有非齐次线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{4-4}$$

它也可写作向量方程

$$\boldsymbol{Ax}=\boldsymbol{b}, \tag{4-5}$$

向量方程 (4-5) 的解也就是方程组 (4-4) 的解向量, 它具有

≡ 性质 4.6

设 $x=\boldsymbol{\eta}_1$ 及 $x=\boldsymbol{\eta}_2$ 都是 (4-5) 的解, 则 $x=\boldsymbol{\eta}_1-\boldsymbol{\eta}_2$ 为对应的齐次线性方程组

$$\boldsymbol{Ax}=\mathbf{0} \tag{4-6}$$

的解.

证明:

$$\boldsymbol{A}(\boldsymbol{\eta}_1-\boldsymbol{\eta}_2)=\boldsymbol{A\eta}_1-\boldsymbol{A\eta}_2=\boldsymbol{b}-\boldsymbol{b}=\mathbf{0},$$

即 $\boldsymbol{x}=\boldsymbol{\eta}_1-\boldsymbol{\eta}_2$ 满足方程 (4-6). □

≡ 性质 4.7

设 $\boldsymbol{x}=\boldsymbol{\eta}$ 是方程 (4-5) 的解, $\boldsymbol{x}=\boldsymbol{\xi}$ 是方程 (4-6) 的解, 则 $\boldsymbol{x}=\boldsymbol{\xi}+\boldsymbol{\eta}$ 仍是方程 (4-5) 的解.

证明:

$$\boldsymbol{A}(\boldsymbol{\xi}+\boldsymbol{\eta})=\boldsymbol{A\xi}+\boldsymbol{A\eta}=\mathbf{0}+\boldsymbol{b}=\boldsymbol{b},$$

即 $\boldsymbol{x}=\boldsymbol{\xi}+\boldsymbol{\eta}$ 满足方程 (4-5). □

由 **性质 4.6** 可知, 若求得 (4-5) 的一个解 $\boldsymbol{\eta}^*$ (称为**特解**) , 则 (4-5) 的任一解总可表示为

$$\boldsymbol{x}=\boldsymbol{\xi}+\boldsymbol{\eta}^*,$$

其中 $\boldsymbol{x}=\boldsymbol{\xi}$ 为方程 (4-6) 的解, 又若方程 (4-6) 的通解为 $\boldsymbol{x}=k_1\boldsymbol{\xi}_1+$

$\cdots + k_{n-r}\boldsymbol{\xi}_{n-r}$ ，则方程 (4-5) 的任一解总可表示为

$$\boldsymbol{x} = k_1\boldsymbol{\xi}_1 + \cdots + k_{n-r}\boldsymbol{\xi}_{n-r} + \boldsymbol{\eta}^*.$$

而由 **性质 4.7** 可知, 对任何实数 $k_1, \cdots, k_{n-r}$ ，上式总是方程 (4-5) 的解. 于是方程 (4-5) 的通解为

$$\boldsymbol{x} = k_1\boldsymbol{\xi}_1 + \cdots + k_{n-r}\boldsymbol{\xi}_{n-r} + \boldsymbol{\eta}^* \, (k_1, \cdots, k_{n-r} \text{ 为任意实数}).$$

其中 $\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_{n-r}$ 是方程组 (4-6) 的基础解系.

例 4.7 求解方程组

$$\begin{cases} x_1 - x_2 - x_3 + x_4 = 0, \\ x_1 - x_2 + x_3 - 3x_4 = 1, \\ x_1 - x_2 - 2x_3 + 3x_4 = -\dfrac{1}{2}. \end{cases}$$

4.5　向量空间

4.5.1　向量空间的定义

定义 4.12: 向量空间

设 V 是 n 维向量的非空集合, 如果 V 对加法和数乘两种运算都封闭, 即

(1) 若 $\boldsymbol{\alpha} \in V, \boldsymbol{\beta} \in V$ ，则 $\boldsymbol{\alpha} + \boldsymbol{\beta} \in V$ ，

(2) 若 $\boldsymbol{\alpha} \in V, k \in \mathbb{R}$ ，则 $k\boldsymbol{\alpha} \in V$ ，

则称 V 是 $\mathbb{R}$ 上的**向量空间**.

全体 n 维向量的集合 $\mathbb{R}^n$ 构成的向量空间称为 n 维向量空间. 特别地, $n = 1$ 时, $\mathbb{R}^1$ 表示一维向量空间, 即数轴; $n = 2$ 时, $\mathbb{R}^2$ 表示二维向量空间, 即平面; $n = 3$ 时, $\mathbb{R}^3$ 表示三维向量空间.

例 4.8 判断下列集合是否构成向量空间.

(1) $V_1 = \left\{(0, x_2, \cdots, x_n)^{\mathrm{T}} \mid x_2, \cdots, x_n \in \mathbb{R}\right\}$.

(2) $V_2 = \left\{(1, x_2, \cdots, x_n)^{\mathrm{T}} \mid x_2, \cdots, x_n \in \mathbb{R}\right\}$.

例 4.9 设 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 为 m 个 n 维向量, 证明集合

$$V = \{\boldsymbol{\xi} \mid \boldsymbol{\xi} = k_1\boldsymbol{\alpha}_1 + k_2\boldsymbol{\alpha}_2 + \cdots + k_m\boldsymbol{\alpha}_m, k_1, k_2, \cdots, k_m \in \mathbb{R}\}$$

是一个向量空间.

该例说明 V 中的向量对向量的加法和数乘运算具有封闭性, 从而构成向量空间. 我们称这个向量空间 V 是由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ **张成**的向量空间.

定义 4.13: 子空间

设 V_1, V_2 为向量空间, 若 $V_1 \subseteq V_2$, 称 V_1 是 V_2 的**子空间**.
例如, 任意向量空间 V 都有两个子空间, 即由零向量构成的零子空间 $\{\mathbf{0}\}$ 和 V 本身, 称为 **平凡子空间**; 任意向量空间 V 都是 $\mathbb{R}^n$ 的子空间, 特别地, 由 n 维向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 生成的向量空间是 $\mathbb{R}^n$ 的子空间.

定义 4.14: 基、维数

设 V 是向量空间, 若 V 中有 m 个向量 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 满足
(1) $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性无关;
(2) V 中任意一个向量都能由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 线性表示,
则称 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 为向量空间 V 的一组**基**, 向量空间 V 的基中所含的向量个数称为向量空间 V 的**维数**, 并称 V 为 m 维向量空间.
我们规定零向量空间 $\{\mathbf{0}\}$ 的维数为 0, 称为 **0 维向量空间**, 它**没有基**.

由定义可知, 若把向量空间 V 看作向量组, **V 的一组基就是一个极大无关组, 而维数就是向量组的秩**. 而且, 若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 为向量空间 V 的一组基, 则 V 可看作由 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 张成的向量空间.

思考 向量空间的基是唯一的吗?

★ 例子

对于向量空间 $\mathbb{R}^n$, n 维基本单位向量组 $\boldsymbol{e}_1 = (1,0,\cdots,0)^{\mathrm{T}}, \boldsymbol{e}_2 = (0,1,\cdots,0)^{\mathrm{T}}, \cdots, \boldsymbol{e}_n = (0,0,\cdots,1)^{\mathrm{T}}$ 是一组基, 维数为 n , 称为 n 维向量空间.
对于由向量组 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 生成的向量空间 V, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的极大无关组就是 V 的一组基, $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 的秩就是 V 的维数.

4.5.2 过渡矩阵与坐标变换

定义 4.15: 坐标

设 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 是 m 维向量空间 V 的一组基, 则对 V 中任意向量 $\boldsymbol{\beta}$, 存在唯一的一组实数 $x_1, x_2, \cdots, x_m$, 使 $\boldsymbol{\beta} = x_1\boldsymbol{\alpha}_1 + x_2\boldsymbol{\alpha}_2 + \cdots + x_m\boldsymbol{\alpha}_m$, 称有序实数组 $(x_1, x_2, \cdots, x_m)^{\mathrm{T}}$ 为向量 $\boldsymbol{\beta}$ 在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 下的**坐标**.

由定义可知, 向量空间 $\mathbb{R}^n$ 中任意向量 $\boldsymbol{\alpha} = (a_1, a_2 \cdots, a_n)$ 在基 $\boldsymbol{e}_1 = (1,0,\cdots,0)^{\mathrm{T}}, \boldsymbol{e}_2 = (0,1,\cdots,0)^{\mathrm{T}}, \cdots, \boldsymbol{e}_n = (0,0,\cdots,1)^{\mathrm{T}}$ 下的坐标就是它的分量.

例 4.10 设向量组 $\boldsymbol{\alpha}_1 = (1,2,1)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (1,3,2)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (1,a,3)^{\mathrm{T}}$ 为 $\mathbb{R}^3$ 的一个基, $\boldsymbol{\beta} = (1,1,1)^{\mathrm{T}}$ 在这组基下的坐标为 $(b,c,1)^{\mathrm{T}}$, 求 a, b, c .

定义 4.16: 基变换公式、过渡矩阵

由向量空间的基的定义, 任意两组基是等价的. 若 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 和 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 是 m 维向量空间 V 的两组基, 则它们可以相互线性表示, 从而存在矩阵 $\boldsymbol{C}$, 使

$$(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m) = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m) \boldsymbol{C}.$$

我们称这个表达式为**基变换公式**, 称矩阵 $\boldsymbol{C}$ 为从基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 的**过渡矩阵**.

从基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 的过渡矩阵实际上就是向量 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_m$ 分别在基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \cdots, \boldsymbol{\alpha}_m$ 下的坐标组成的矩阵, 也就是前文定义的表示矩阵.

例 4.11 设 $\boldsymbol{\alpha}_1 = (1,1,0)^{\mathrm{T}}, \boldsymbol{\alpha}_2 = (0,1,1)^{\mathrm{T}}, \boldsymbol{\alpha}_3 = (0,0,1)^{\mathrm{T}}$ 和 $\boldsymbol{\beta}_1 = (1,-1,-1)^{\mathrm{T}}, \boldsymbol{\beta}_2 = (1,1,-1)^{\mathrm{T}}, \boldsymbol{\beta}_3 = (-1\ ,\ 1,0)^{\mathrm{T}}$ 是向量空间 $\mathbb{R}^3$ 的两组基.

求由基 $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ 到基 $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3$ 的过渡矩阵.

4.6 向量的内积

在向量空间中, 向量之间的运算只有加法和数乘两种, 与几何空间相比, 就会发现向量的长度、向量间的夹角等度量性质在向量空间的理论中都没有得到反映. 而这些度量性质无论是在理论上还是在实际问题中都是非常重要的, 因此有必要引入这些概念.

4.6.1 向量的内积

在三维几何空间中, 向量的长度、夹角都可用数量积 (内积) 来描述; 数量积 (内积) 定义为

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = |\boldsymbol{\alpha}||\boldsymbol{\beta}| \cos\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle,$$

其坐标表达式为

$$\boldsymbol{\alpha} \cdot \boldsymbol{\beta} = a_1 b_1 + a_2 b_2 + a_3 b_3,$$

其中 $\boldsymbol{\alpha} = (a_1, a_2, a_3), \boldsymbol{\beta} = (b_1, b_2, b_3)$. 由此可得:

向量 $\boldsymbol{\alpha}$ 的长度 $|\boldsymbol{\alpha}| = \sqrt{\boldsymbol{\alpha} \cdot \boldsymbol{\alpha}} = \sqrt{a_1^2 + a_2^2 + a_3^2}$, 向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的夹角 $\langle \boldsymbol{\alpha}, \boldsymbol{\beta} \rangle = \arccos \dfrac{\boldsymbol{\alpha} \cdot \boldsymbol{\beta}}{|\boldsymbol{\alpha}||\boldsymbol{\beta}|}$. 下面我们将这一概念推广到 n 维实向量空间 $\mathbb{R}^n$.

定义 4.17: 向量的内积

在 n 维实向量空间 $\mathbb{R}^n$ 中, 设

$$\boldsymbol{\alpha}=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix},\quad \boldsymbol{\beta}=\begin{pmatrix}b_1\\b_2\\\vdots\\b_n\end{pmatrix},$$

则

$$\boldsymbol{\alpha}^T\boldsymbol{\beta}=a_1b_1+a_2b_2+\cdots+a_nb_n=\sum_{i=1}^n a_ib_i$$

称为向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ 的**内积**, 记作 ($\boldsymbol{\alpha},\boldsymbol{\beta}$). 定义内积的实向量空间 $\mathbb{R}^n$ 称为**欧几里得空间**, 简称**欧氏空间**, 仍用 $\mathbb{R}^n$ 表示欧氏空间.

向量内积的性质

1) $(\boldsymbol{\alpha},\boldsymbol{\beta})=(\boldsymbol{\beta},\boldsymbol{\alpha})$.
2) $(k\boldsymbol{\alpha},\boldsymbol{\beta})=k(\boldsymbol{\alpha},\boldsymbol{\beta})$.
3) $(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\gamma})=(\boldsymbol{\alpha},\boldsymbol{\gamma})+(\boldsymbol{\beta},\boldsymbol{\gamma})$.
4) $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0,(\boldsymbol{\alpha},\boldsymbol{\alpha})=0$ 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$,

其中 $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ 为 n 维实向量, k 为任意实数.

由于 $(\boldsymbol{\alpha},\boldsymbol{\alpha})\geqslant 0$, 在欧氏空间 $\mathbb{R}^n$ 中可引入向量 $\boldsymbol{\alpha}$ 长度的概念.

定义 4.18: 向量长度、向量的范数、向量单位化

对任意 $\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}}\in\mathbb{R}^n$, 规定向量 $\boldsymbol{\alpha}$ 的**长度** $\|\boldsymbol{\alpha}\|$ 为

$$\|\boldsymbol{\alpha}\|=\sqrt{(\boldsymbol{\alpha},\boldsymbol{\alpha})}=\sqrt{\boldsymbol{\alpha}^{\mathrm{T}}\boldsymbol{\alpha}}=\sqrt{a_1^2+a_2^2+\cdots+a_n^2}.$$

向量的长度也称为向量的**范数** (norm). 当 $\|\boldsymbol{\alpha}\|=1$ 时, 称 $\boldsymbol{\alpha}$ 为**单位向量** (unit vector). 显然, $\|\boldsymbol{\alpha}\|=1$ 当且仅当 ($\boldsymbol{\alpha}$, $\boldsymbol{\alpha}$) $=1$. 若 $\boldsymbol{\alpha}\neq\mathbf{0}$, 则 $\|\boldsymbol{\alpha}\|\neq 0$, 那么

$$\left(\frac{1}{\|\boldsymbol{\alpha}\|}\boldsymbol{\alpha},\frac{1}{\|\boldsymbol{\alpha}\|}\boldsymbol{\alpha}\right)=\frac{1}{\|\boldsymbol{\alpha}\|^2}(\boldsymbol{\alpha},\boldsymbol{\alpha})=\frac{1}{\|\boldsymbol{\alpha}\|^2}\cdot\|\boldsymbol{\alpha}\|^2=1,$$

也就是说 $\dfrac{1}{\|\boldsymbol{\alpha}\|}\boldsymbol{\alpha}$ 是一个单位向量.

把一个非零实向量 $\boldsymbol{\alpha}$ 变成单位向量 $\dfrac{1}{\|\boldsymbol{\alpha}\|}\boldsymbol{\alpha}$, 称为把向量 $\boldsymbol{\alpha}$ **单位化**.

向量长度的性质

1) $\|\boldsymbol{\alpha}\|\geqslant 0$, 当且仅当 $\boldsymbol{\alpha}=\mathbf{0}$ 时, 有 $\|\boldsymbol{\alpha}\|=0$.
2) $\|k\boldsymbol{\alpha}\|=|k|\cdot\|\boldsymbol{\alpha}\|$ (k 为常数).

为了引入向量夹角的概念, 可证明得到如下不等式.

定理 4.10

设 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n,\boldsymbol{\alpha}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}},\boldsymbol{\beta}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}}$, 则

$$\|(\boldsymbol{\alpha},\boldsymbol{\beta})\|\leqslant\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|$$

或者

$$\left|\sum_{i=1}^{n}a_ib_i\right|\leqslant\sqrt{\sum_{i=1}^{n}a_i^2}\cdot\sqrt{\sum_{i=1}^{n}b_i^2}.$$

该式称为**柯西——施瓦茨 (Cauchy-Schwarz) 不等式**.

根据柯西——施瓦茨不等式, 对于 $\boldsymbol{\alpha}\neq\mathbf{0},\boldsymbol{\beta}\neq\mathbf{0}$, 总有

$$-1\leqslant\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\cdot\|\boldsymbol{\beta}\|}\leqslant 1.$$

由此可对欧氏空间 $\mathbb{R}^n$ 中的向量引入夹角的概念.

定义 4.19: 向量与向量的夹角、正交

设 $\alpha,\beta\in\mathbb{R}^n,\alpha\neq 0,\beta\neq 0$, 定义 α 与 β 的**夹角** $\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle$ 为

$$\theta=\langle\boldsymbol{\alpha},\boldsymbol{\beta}\rangle=\arccos\frac{(\boldsymbol{\alpha},\boldsymbol{\beta})}{\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|}.$$

显然, $0\leqslant\theta\leqslant\pi$. 当 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$ 时, $\theta=\dfrac{\pi}{2}$, 此时称向量 $\boldsymbol{\alpha}$ 与 $\boldsymbol{\beta}$ **正交**(或垂直) , 记作 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$. 显然, 若 $\boldsymbol{\alpha}=\mathbf{0}$, 则 $\boldsymbol{\alpha}$ 与任何向量正交.

定理 4.11

设 $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathbb{R}^n$, 则有

1) $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|\leqslant\|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|$ (三角不等式);
2) 当 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$ 时, $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2$ (勾股定理).

证明: (1) $\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\alpha})+2(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\beta})$

$$\begin{aligned}&\leqslant\|\boldsymbol{\alpha}\|^2+2\|\boldsymbol{\alpha}\|\|\boldsymbol{\beta}\|+\|\boldsymbol{\beta}\|^2\\&=(\|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|)^2,\end{aligned}$$

因此

$$\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|\leqslant\|\boldsymbol{\alpha}\|+\|\boldsymbol{\beta}\|.$$

(2) 因为 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$, 故 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 那么

$$\begin{aligned}\|\boldsymbol{\alpha}+\boldsymbol{\beta}\|^2&=(\boldsymbol{\alpha}+\boldsymbol{\beta},\boldsymbol{\alpha}+\boldsymbol{\beta})=(\boldsymbol{\alpha},\boldsymbol{\alpha})+2(\boldsymbol{\alpha},\boldsymbol{\beta})+(\boldsymbol{\beta},\boldsymbol{\beta})\\&=(\boldsymbol{\alpha},\boldsymbol{\alpha})+(\boldsymbol{\beta},\boldsymbol{\beta})=\|\boldsymbol{\alpha}\|^2+\|\boldsymbol{\beta}\|^2.\end{aligned}$$

□

4.6.2 标准正交向量组

定义 4.20: 正交

若向量 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 的内积为零, 即 $(\boldsymbol{\alpha},\boldsymbol{\beta})=0$, 则称 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 正交.

由定义, 零向量与任何向量都正交. 而且, 若 $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 正交, 则夹角为 $\frac{\pi}{2}$, 即正交的几何意义为垂直, 所以, $\boldsymbol{\alpha}$ 和 $\boldsymbol{\beta}$ 正交, 可以记为 $\boldsymbol{\alpha}\perp\boldsymbol{\beta}$.

定义 4.21: 正交向量组、标准正交向量组、标准正交基

若一个非零向量组的任意两个向量都是正交的, 则称该向量组为**正交向量组**. 若正交向量组的每一个向量都是**单位向量**, 则称为**标准正交向量组**.

特别地, 如果**标准正交向量组的秩等于向量空间的维数**, 则称该标准正交向量组为**标准正交基**.

例子

n 维基本单位向量组 $\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}},\boldsymbol{e}_2=(0,1,\cdots,0)^{\mathrm{T}},\cdots,\boldsymbol{e}_n=(0,0,\cdots,1)^{\mathrm{T}}$ 是 n 维向量空间 $\mathbb{R}^n$ 中的一组标准正交基.

又如, 四维向量组 $\boldsymbol{\alpha}_1=(1,0,0,0)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,1,0,0)^{\mathrm{T}},\boldsymbol{\alpha}_3=\left(0,0,\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)^{\mathrm{T}},\boldsymbol{\alpha}_4=\left(0,0,\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)^{\mathrm{T}}$ 是四维向量空间 $\mathbb{R}^4$ 的一组标准正交基.

关于正交向量组, 有如下重要性质.

定理 4.12

设 n 维向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是正交向量组, 则 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是线性无关的.

证明: 设有一组数 $k_1,k_2,\cdots,k_m$, 使

$$k_i\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0},$$

用 $\boldsymbol{\alpha}_i\,(i=1,2,\cdots,m)$ 分别和上式两端做内积, 当 $i\neq j$ 时, $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=0$, 所以有

$$k_i\,(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)=0,\ i=1,2,\cdots,m.$$

又因为 $\boldsymbol{\alpha}_i\neq\mathbf{0}\,(i=1,2,\cdots,m)$, 有 $(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_i)\neq0$ 所以 $k_i=0\,(i=1,2,\cdots,m)$, 即 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 线性无关.

□

现在的问题是, 给定线性无关的向量组, 如何变成与之等价的标准正交向量组, 或者, 给定向量空间的一组基, 如何变成与之等价的标准正交基? 这就是下面要介绍的**施密特正交化**方法.

施密特正交化

设 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 线性无关, 令

$$\boldsymbol{\beta}_1=\boldsymbol{\alpha}_1,\boldsymbol{\beta}_2=\boldsymbol{\alpha}_2-\frac{(\boldsymbol{\alpha}_2,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1,\cdots,$$

$$\boldsymbol{\beta}_s=\boldsymbol{\alpha}_s-\frac{(\boldsymbol{\alpha}_s,\boldsymbol{\beta}_1)}{(\boldsymbol{\beta}_1,\boldsymbol{\beta}_1)}\boldsymbol{\beta}_1-\frac{(\boldsymbol{\alpha}_s,\boldsymbol{\beta}_2)}{(\boldsymbol{\beta}_2,\boldsymbol{\beta}_2)}\boldsymbol{\beta}_2-\cdots-\frac{(\boldsymbol{\alpha}_s,\boldsymbol{\beta}_{s-1})}{(\boldsymbol{\beta}_{s-1},\boldsymbol{\beta}_{s-1})}\boldsymbol{\beta}_{s-1},$$

则 $\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\cdots,\boldsymbol{\beta}_s$ 是与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 等价的正交向量组.

再令

$$\gamma_1=\frac{\boldsymbol{\beta}_1}{\|\boldsymbol{\beta}_1\|},\gamma_2=\frac{\boldsymbol{\beta}_2}{\|\boldsymbol{\beta}_2\|},\cdots,\gamma_s=\frac{\boldsymbol{\beta}_s}{\|\boldsymbol{\beta}_s\|},$$

则 $\gamma_1,\gamma_2,\cdots,\gamma_s$ 是与 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 等价的标准正交向量组.

上述过程称为向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_s$ 的**正交规范化**过程.

例 4.12 设 $\boldsymbol{\alpha}_1=(1,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_2=(0,1,1)^{\mathrm{T}},\boldsymbol{\alpha}_3=(0,0,1)^{\mathrm{T}}$ ，用施密特正交化方法将该向量组正交化规范.

4.6.3　正交矩阵

向量空间 V 的一个基 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$ 是标准正交基的充分必要条件是

$$(\boldsymbol{\alpha}_i,\boldsymbol{\alpha}_j)=\begin{cases}1, & i=j,\\ 0, & i\neq j,\end{cases}\quad i,j=1,2,\cdots,m.$$

故 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$ 是 n 维向量空间 $\mathbb{R}^n$ 的标准正交基, 记 $\boldsymbol{A}=(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)$, 则有

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}\boldsymbol{e}_1^{\mathrm{T}}\\ \boldsymbol{e}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{e}_n^{\mathrm{T}}\end{pmatrix}(\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n)=\boldsymbol{E}.$$

定义 4.22: 正交矩阵

如果 n 阶实方阵 $\boldsymbol{A}$ 满足 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}=\boldsymbol{E}$, 则称 $\boldsymbol{A}$ 为正交矩阵.

由定义容易证明 n 阶正交矩阵 $\boldsymbol{A}$ 具有以下性质.

性质 4.8

设 $\boldsymbol{A}$ 为 n 阶正交矩阵, 则

1) $\boldsymbol{A}$ 的行列式为 1 或 -1 ;
2) $\boldsymbol{A}$ 为可逆矩阵, 且 $\boldsymbol{A}^{-1}=\boldsymbol{A}^{\mathrm{T}}$;
3) $\boldsymbol{A}$ 的列向量组是 $\mathbb{R}^n$ 的一个标准正交基.

第 五 章　特征向量、相似矩阵与二次型

疾风知劲草, 板荡识诚臣. 勇夫安识义, 智者必怀仁.

唐·李世民 (599–649) «赐萧瑀»

本章主要关注以下几个问题.

1) 矩阵特征值、特征向量概念的引入和计算.
2) 相似矩阵概念的引入、相似矩阵的性质以及矩阵对角化的充要条件和方法.
3) 二次型引入的目的、形式和求解方法.

5.1　方阵的特征值与特征向量

5.1.1　特征值与特征向量的定义

定义 5.1: 特征值、特征向量、特征方程

设 $\boldsymbol{A}$ 是 n 阶矩阵, 如果数 λ 和 n 维非零列向量 $\boldsymbol{x}$ 使关系式

$$\boldsymbol{A}\boldsymbol{x}=\lambda\boldsymbol{x} \tag{5-1}$$

成立, 那么, 这样的数 λ 称为矩阵 $\boldsymbol{A}$ 的**特征值**, 非零向量 $\boldsymbol{x}$ 称为 $\boldsymbol{A}$ 的对应于特征值 λ 的**特征向量**.
式 (5-1) 也可写成

$$(\boldsymbol{A}-\lambda\boldsymbol{E})\,\boldsymbol{x}=0, \tag{5-2}$$

这是 n 个未知数 n 个方程的齐次线性方程组, 它有非零解的充分必要条件是系数行列式

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=0,$$

即

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

上式是以 λ 为未知数的一元 n 次方程, 称为矩阵 $\boldsymbol{A}$ 的**特征方程**.

其左端 $|\boldsymbol{A}-\lambda\boldsymbol{E}|$ 是 λ 的 n 次多项式, 记作 $f(\lambda)$, 称为矩阵 $\boldsymbol{A}$ 的**特征多项式**. 显然, $\boldsymbol{A}$ 的特征值就是特征方程的解. 特征方程在复数范围内恒有解, 其个数为方程的次数 (重根按重数计算), 因此, n 阶矩阵 $\boldsymbol{A}$ 在复数范围内有 n 个特征值.

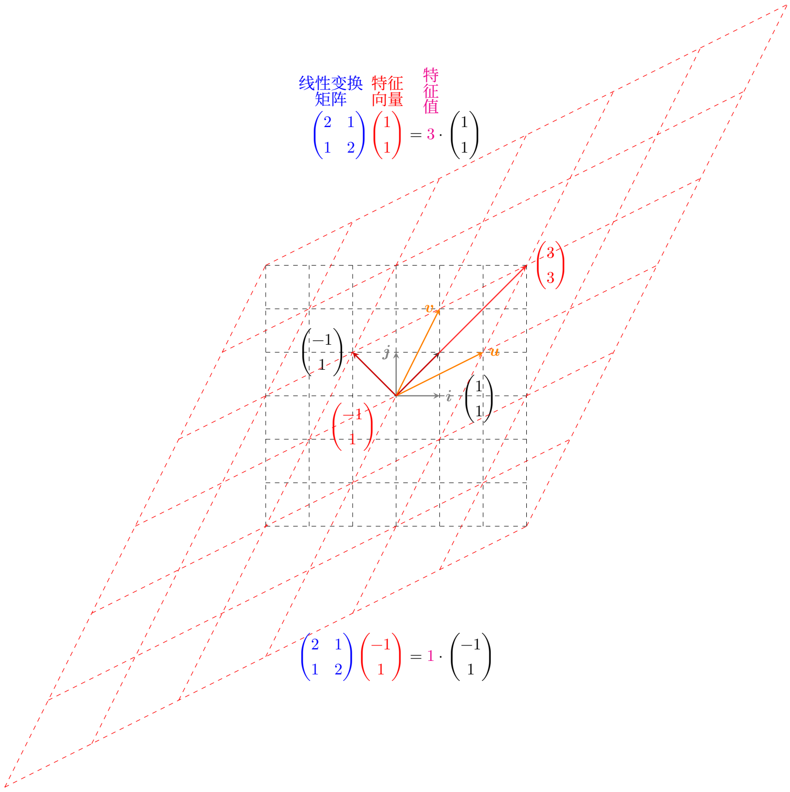


图 5-1　特征值、特征向量示例.

设 $\lambda=\lambda_i$ 为矩阵 $\boldsymbol{A}$ 的一个特征值, 则由方程

$$(\boldsymbol{A}-\lambda_i\boldsymbol{E})\,\boldsymbol{x}=0$$

可求得非零解 $\boldsymbol{x}=\boldsymbol{p}_i$, 那么 $\boldsymbol{p}_i$ 便是 $\boldsymbol{A}$ 的对应于特征值 λ_i 的**特征向量**. (若 λ_i 为实数, 则 $\boldsymbol{p}_i$ 可取实向量; 若 λ_i 为复数, 则 $\boldsymbol{p}_i$ 为复向量.)

例 5.1　求矩阵 $\boldsymbol{A}=\begin{pmatrix}3&-1\\-1&3\end{pmatrix}$ 的特征值和特征向量.

解:　$\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix}3-\lambda&-1\\-1&3-\lambda\end{vmatrix}=(3-\lambda)^2-1=8-6\lambda+\lambda^2$$
$$=(4-\lambda)(2-\lambda),$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=2,\lambda_2=4$.

当 $\lambda_1=2$ 时, 对应的特征向量应满足

$$\begin{pmatrix} 3-2 & -1 \\ -1 & 3-2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix},\text{即}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

解得 $x_1=x_2$, 所以对应的特征向量可取为

$$\boldsymbol{p}_1=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

当 $\lambda_2=4$ 时, 由

$$\begin{pmatrix} 3-4 & -1 \\ -1 & 3-4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix},\text{即}\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

解得 $x_1=-x_2$, 所以对应的特征向量可取为

$$p_2=\begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

显然, 若 $\boldsymbol{p}_i$ 是矩阵 $\boldsymbol{A}$ 的对应于特征值 λ_i 的特征向量, 则 $k\boldsymbol{p}_i(k\neq 0)$ 也是对应于 λ_i 的特征向量.

例 5.2 求矩阵

$$\boldsymbol{A}=\begin{pmatrix} -1 & 1 & 0 \\ -4 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix}$$

的特征值和特征向量.

解: $\boldsymbol{A}$ 的特征多项式为

$$|\boldsymbol{A}-\lambda\boldsymbol{E}|=\begin{vmatrix} -1-\lambda & 1 & 0 \\ -4 & 3-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix}=(2-\lambda)(1-\lambda)^2,$$

所以 $\boldsymbol{A}$ 的特征值为 $\lambda_1=2,\lambda_2=\lambda_3=1$.

当 $\lambda_1=2$ 时, 解方程 $(\boldsymbol{A}-2\boldsymbol{E})\boldsymbol{x}=0$. 由

$$\boldsymbol{A}-2\boldsymbol{E}=\begin{pmatrix} -3 & 1 & 0 \\ -4 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\overset{r}{\cong}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

得基础解系

$$p_1=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

所以 $kp_1(k\neq 0)$ 是对应于 $\lambda_1=2$ 的全部特征向量.

当 $\lambda_2=\lambda_3=1$ 时, 解方程 $(\boldsymbol{A}-\boldsymbol{E})\boldsymbol{x}=0$. 由

$$\boldsymbol{A}-\boldsymbol{E}=\begin{pmatrix}-2&1&0\\-4&2&0\\1&0&1\end{pmatrix}\overset{r}{\cong}\begin{pmatrix}1&0&1\\0&1&2\\0&0&0\end{pmatrix},$$

得基础解系

$$p_2=\begin{pmatrix}-1\\-2\\1\end{pmatrix}$$

所以 $kp_2(k\neq 0)$ 是对应于 $\lambda_2=\lambda_3=1$ 的全部特征向量.

例 5.3 设 λ 是方阵 $\boldsymbol{A}$ 的特征值, 证明

(1) λ^2 是 $\boldsymbol{A}^2$ 的特征值;

(2) 当 $\boldsymbol{A}$ 可逆时, $\dfrac{1}{\lambda}$ 是 $\boldsymbol{A}^{-1}$ 的特征值.

证明: 因 λ 是 $\boldsymbol{A}$ 的特征值, 故有 $\boldsymbol{p}\neq 0$ 使 $\boldsymbol{A}\boldsymbol{p}=\lambda\boldsymbol{p}$. 于是

(1)

$$\boldsymbol{A}^2\boldsymbol{p}=\boldsymbol{A}(\boldsymbol{A}\boldsymbol{p})=\boldsymbol{A}(\lambda\boldsymbol{p})=\lambda(\boldsymbol{A}\boldsymbol{p})=\lambda^2\boldsymbol{p},$$

所以 λ^2 是 $\boldsymbol{A}^2$ 的特征值.

(2) 当 $\boldsymbol{A}$ 可逆时, 由 $\boldsymbol{A}\boldsymbol{p}=\lambda\boldsymbol{p}$, 有 $\boldsymbol{p}=\lambda\boldsymbol{A}^{-1}\boldsymbol{p}$, 因 $\boldsymbol{p}\neq 0$, 知 $\lambda\neq 0$, 故

$$\boldsymbol{A}^{-1}\boldsymbol{p}=\frac{1}{\lambda}\boldsymbol{p},$$

所以 $\dfrac{1}{\lambda}$ 是 $\boldsymbol{A}^{-1}$ 的特征值.

按此例类推, 不难证明: 若 λ 是 $\boldsymbol{A}$ 的特征值, 则 λ^k 是 $\boldsymbol{A}^k$ 的特征值; $\varphi(\lambda)$ 是 $\varphi(\boldsymbol{A})$ 的特征值 (其中 $\varphi(\lambda)=a_0+a_1\lambda+\cdots+a_m\lambda^m$ 是 λ 的多项式, $\varphi(\boldsymbol{A})=a_0\boldsymbol{E}+a_1\boldsymbol{A}+\cdots+a_m\boldsymbol{A}^m$ 是矩阵 $\boldsymbol{A}$ 的多项式). □

例 5.4 设 3 阶矩阵 $\boldsymbol{A}$ 的特征值为 $1,-1,2$, 求 $\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}$ 的特征值.

解: 因 $\boldsymbol{A}$ 的特征值全不为 0, 知 $\boldsymbol{A}$ 可逆, 故 $\boldsymbol{A}^*=|\boldsymbol{A}|\,\boldsymbol{A}^{-1}$. 而 $|\boldsymbol{A}|=\lambda_1\lambda_2\lambda_3=-2$, 所以

$$\boldsymbol{A}^*+3\boldsymbol{A}-2\boldsymbol{E}=-2\boldsymbol{A}^{-1}+3\boldsymbol{A}-2\boldsymbol{E}.$$

把上式记作 $\varphi(\boldsymbol{A})$, 有 $\varphi(\lambda)=-\dfrac{2}{\lambda}+3\lambda-2$. 这里, $\varphi(\boldsymbol{A})$ 虽不是矩阵多项式, 但也具有矩阵多项式的特性, 从而可得 $\varphi(\boldsymbol{A})$ 的特征值为 $\varphi(1)=-1$, $\varphi(-1)=-3,\varphi(2)=3$.

5.1.2 特征值与特征向量的性质

☰ 性质 5.1: 特征值性质

设 n 阶矩阵 $\boldsymbol{A}=(a_{ij})$ 的特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$ ，可证明

1) $\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$ ；
2) $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$.

定义 5.2: 迹

设矩阵 $\boldsymbol{A}=(a_{ij})_{n\times n}$ ，称 $a_{11}+a_{22}+\cdots+a_{nn}$ 为 $\boldsymbol{A}$ 的**迹**, 记为 $\operatorname{tr}\boldsymbol{A}$.

☰ 性质 5.2

矩阵 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征值.

证明: 因为

$$\left|\boldsymbol{A}^{\mathrm{T}}-\lambda\boldsymbol{E}\right|=\left|(\boldsymbol{A}-\lambda\boldsymbol{E})^{\mathrm{T}}\right|=|\boldsymbol{A}-\lambda\boldsymbol{E}|,$$

所以 $\boldsymbol{A}$ 和 $\boldsymbol{A}^{\mathrm{T}}$ 有相同的特征多项式, 故它们有相同的特征值.

□

☰ 性质 5.3

设 $\boldsymbol{A}$ 是 n 阶可逆矩阵, 则

1) $\boldsymbol{A}$ 的特征值都不为零;
2) 若 λ 是 $\boldsymbol{A}$ 的特征值, 则 λ^{-1} 是 $\boldsymbol{A}^{-1}$ 的特征值.

证明: (1) 设 $\boldsymbol{A}$ 的全部特征值为 $\lambda_1,\lambda_2,\cdots,\lambda_n$. 由 **性质 5.1** , $\lambda_1\lambda_2\cdots\lambda_n=|\boldsymbol{A}|$, 因 $\boldsymbol{A}$ 可逆, 所以 $|\boldsymbol{A}|\neq 0$, 故 $\lambda_1,\lambda_2,\cdots,\lambda_n$ 都不为零.

(2) 设 $\boldsymbol{\xi}$ 是 $\boldsymbol{A}$ 的属于 λ 的特征向量, 则 $\boldsymbol{A\xi}=\lambda\boldsymbol{\xi}$. 根据 (1), $\lambda\neq 0$, 于是

$$\frac{1}{\lambda}\boldsymbol{A\xi}=\boldsymbol{\xi}$$

用 $\boldsymbol{A}^{-1}$ 同时左乘上式的两边得

$$\boldsymbol{A}^{-1}\boldsymbol{\xi}=\frac{1}{\lambda}\boldsymbol{\xi}$$

由 $\boldsymbol{\xi}\neq 0$ 知, λ^{-1} 是 $\boldsymbol{A}^{-1}$ 的特征值. □

☰ 性质 5.4

设 $f(x)=a_mx^m+\cdots+a_1x+a_0$ 是关于 x 的多项式, $\boldsymbol{A}$ 是 n 阶方阵, 此时

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

> 若 λ 是 $\boldsymbol{A}$ 的特征值, 则 $f(\lambda)$ 是 $f(\boldsymbol{A})$ 的特征值.

例 5.5 已知三阶方阵 $\boldsymbol{A}$ 的特征值为 $-1,1,2$, 求 $\left|\boldsymbol{A}^3-5\boldsymbol{A}^2\right|$.

解: 设 $f(x)=x^3-5x^2$, 则 $f(A)=A^3-5A^2$, 由性质 5.4 知, $f(A)$ 的全部特征值为 $f(-1)=-6, f(1)=-4, f(2)=-12$, 因此

$$\left|A^3-5A^2\right|=(-6)\times(-4)\times(-12)=-288.$$

定理 5.1

设 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 是方阵 $\boldsymbol{A}$ 的 m 个特征值, $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 依次是与之对应的特征向量, 如果 $\lambda_1,\lambda_2,\cdots,\lambda_m$ 各不相等, 则 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 线性无关.