线性代数课程讲义

June 26, 2021

Abstract

本讲义主要是梳理知识, 发现定理, 严格证明参考教科书.

Contents

1 开场白

新的数学方法和概念, 常常比解决数学问题本身更重要. 中学与大学的不同: 前者更注重题目, 后者更注重概念. 但注重概念不代表可以忽视做题. 题目使得大家更深刻理解概念. 做题目就好似与刚刚认识的一位朋友的一次聊天, 一次约会. 增进你对她的了解. 无论安排几次会面, 熟识她才是目的.

数学的美在于为不同的事物起了相同的名字. 这也是为什么很多人认为数学抽象的原因. 你要和她利用公式对话. 公式是她的语言. 让公式说话, 用公式说话. 听得懂, 说得出. 你能从梵高的向日葵中看到生命的律动, 也能用图画展示自己的情怀.

我的任务是将复杂的东西简洁地呈现给大家, 而大家用更多的精力去学习和研究更复杂的知识.

学习和研究数学, 既要有异想天开的不羁, 也要有脚踏实地的严谨. 思想不被束缚, 行动要有准则. 发现比证明更重要. 大家需要总问自己, 这个定理存在的意义是什么? 它勾勒出怎样的一个图景, 描述了什么样的关系? 有思想地去做题.

数学更像是在习武, 分为内功和外功. 内功是思维方式, 外功是计算推导能力. 那么学习一本武林秘籍, 必然是既要有心法口诀也要有武功招式. 类比来看, 数学中的公式定理就是武功招式, 而心法口诀就是串联所有公式定理的思维方式. 做题目修炼外功, 强化对武功招式的熟练. 思考修炼内功, 将心法口诀烂熟于胸. 只会思考, 如同王语嫣. 只会做题目, 如同鸠摩智. 武学的一招一式都会与心法口诀相呼应. 数学中的每个公式定理都会与思维方式相联系.

线性代数, 这个教科书就是一本武林典籍. 历史上诸多武林宗师共同创造了这门经典武功. 它有四大招式: 行列式, 矩阵, 线性空间, 线性映射. 最初武林宗师们只为解决一个问题: 线性方程组的求解. 我们将会用这四大招式分别阐述和求解线性方程组. 我也会为大家阐明这四大招式的内在联系.

线性代数的第一层内功心法是"消元法". 心法生招式: 消元法贯通行列式, 矩阵, 线性空间, 线性映射. 内功修为的增进需通过外功的修炼, 所以做题目是必不可少的, 而且要有思想地做题目, 最终无招胜有招.

对于线性空间, 初看与线性方程组求解完全没有关系. 其实它是另一种视角来描述线性方程组 (线性方程组的几何意义). 线性代数的上半部的主要内容是消元法, 行列式和线性方程组求解. 下半部的主要内容是矩阵, 线性空间和线性映射. 前半部是进行内外功的修炼, 后半部是进行佛法的修行. "武功的戾气需要用佛法来化解". 线性代数的下半部将线性方程组问题用"集合和映射"的观点完美的阐述和解决. 只修炼上半部, 纵然你武功再高, 也需要下半部统领全局, 不至于被武功所累. 若只会上半部分, 那你只是耍把式的, 街头卖艺混口饭吃. 将下半部分融汇贯通的习武之人, 才有可能成为一代宗师.

值得说明的是, 集合与映射是数学最基本的对象, 数学的研究都基于两者展开. 线性代数的集合与映射具体体现为线性空间和线性映射. 另外, 线性代数不单单只是消元法, 只能求解线性方程组. 这门基础性武学的习得, 会为大家修炼更上乘的武林绝学铺平道路. 数学分析和线性代数, 打通大家的任督二脉.

"少年, 我看你骨骼精奇, 是万中无一的武学奇才, 维护世界和平就靠你了, 我这有本秘籍《线性代数》, 见与你有缘, 就十块卖给你了!"

Remark 1.1. *每引入一个新的定义, 需要问自己三个问题:*

- *为什么它很重要?*
- *它是什么?*

- 它与其它定义有什么联系?

每证明一个新的定理, 需要问自己三个问题:

- 为什么它很重要?
- 为什么它是对的?
- 它与哪些定义有关联?

定义是“点”, 定理是“线”.

2 发展简史 – 读历史明更替

线性代数就像一串冰糖葫芦, 各个知识点就是一个个裹着冰糖的红果, 而求解线性方程组就是那个将它们串起来的竹签.

中国古代就有一个非常著名的线性方程组求解问题:

Example 2.1 (鸡兔同笼)**.** *今有雉、兔同笼, 上有三十五头, 下有九十四足. 问: 雉、兔各几何? 答曰: 雉二十三, 兔一十二.*

Remark 2.1 (代数学 – Algebra)**.** *研究数, 数量, 关系与结构的数学分支.*

- *公元 3 世纪, 丢番图, 古希腊,《算术》: 引入未知数的概念, 并用文字缩写表达未知数, 开创了简化代数.*
- *公元 820, 阿尔 · 花剌子模, 阿拉伯,《代数学》: 归位与对消的科学, 把负项移到方程另一边变成正项, 方程两边也可消去相同项或合并同类项.*
- *1519, 韦达, 法国,《分析方法入门》: 首次系统地使用了符号表示未知量的值进行运算, 提出符号运算与数的区别, 规定了代数与算术的分界. 经笛卡尔改进后成为现代的形式. 笛卡尔用小写字母 a, b, c 等表示已知量, 而用 x, y, z 代表未知量.*
- *1859, 李善兰, 中国,《代数学》: 代数之法, 无论何数, 皆可以任何记号代之.*

Remark 2.2. *代数学的思想 – 同解简化: 比如收拾屋子. 首先为整理的东西分类 (找共同点), 然后为每一类制定一种处理方式.*

符号只是一种形式, 描述问题, 表达思想和过程.

Remark 2.3 (方程 – Equation)**.** *描述事物之间关系的等式.*

- *一个桌子四条腿 $y = 4x$.*
- *道高一尺魔高一丈 $y = 10x$.*
- *唇亡齿寒, 此消彼长, 一荣俱荣一损俱存, 差之毫厘失之千里, 一失足成千古恨.*

2.1 时间线

- 1693, Gottfried Wilhelm Leibniz, 法国, 利用 Determinant 求解线性方程组.
- 1750, Gabriel Cramer, 瑞士, 发现 Cramer's Rule 求解线性方程组.
- Later, A. T. Vandermonde, 法国, 系统阐述行列式理论, 将行列式理论与线性方程组求解相分离.
- Later, Augustin Louis Cauchy, 法国, 将元素排成方阵来记行列式, 并首次采用双重下标.
- Later, Carl Friedrich Gauss, 德国, 利用 Gaussian elimination 求解线性方程组.
- 1844, Hermann Grassmann, 德国, 发表 Theory of Extension 建立线性代数雏形.
- 1848, James Joseph Sylvester, 英国, 引入 matrix (矩阵) 这个词, 它源于拉丁语, 代表一排数.

- Later, Arthur Cayley, 英国, 研究线性变换时引入矩阵乘法和求逆, 并发现 matrix 和 determinant 的联系.
- 1882, Hüseyin Tevfik Pasha, 土耳其, 出版 Linear Algebra 一书.
- 1888, Giuseppe Peano, 意大利, 精确定义 Vector space.
- ···

历史上线性代数的第一个问题是解线性方程组, 而线性方程组理论的发展又促成了作为工具的矩阵论和行列式理论的创立与发展, 最终创立了线性空间理论.

- 判别式 — 莱布尼茨的行列式.
- 一切的源头 — 凯莱的矩阵.
- 集大成 — 皮亚诺的线性空间.

2.2 内容线

秩序. 有统一的标准才有秩序, 有了标准就可以进行分类. "尊师重道", "忠、信、孝、悌、礼、义、廉、耻", 皆是依道德之理所遵循的标准原则或是行为准则. 顺应道德的准则是善，违逆道德的准则是恶. 换句话说, 只要有标准, 我们就能判断一件事情的好与坏. 善与恶. 有了所公认的准则就有了秩序, 有了分类.

线性代数是研究秩序的一门学科, 线性方程组的求解过程就是分类整理的过程. 秩, 这个概念更能体现秩序与准则, 等价关系与分类的思想. 矩阵的相抵关系, 相似关系, 合同关系都是这些思想地具体实现.

Linear Algebra

- Linear Equations (线性方程组)
  - 代数:
    - Determinant (行列式): 用一个数判别线性方程组解的存在唯一性.
    - Cramer Rule (克拉默法则): 用线性方程组的系数表达解.
    - Gaussian Elimination (高斯消元法): 一套流程计算线性方程组的解.
  - 几何:
    - Vector Space (向量空间): 求线性组合的系数.
  - 集合与映射:
    - Matrix (矩阵): 已知像, 求原象.
- 等价关系
  - 相抵关系
  - 相似关系
  - 合同关系

3 消元法 – 高斯的三板斧

高斯消元法可以完全解决线性方程组求解的问题.

- 什么是方程, 什么是线性方程?
- 为什么要研究线性方程组, 研究它的哪些性质?
- 完全解决是什么意思?
- 解决的思路是什么?
- 具体步骤是什么?
- 为什么是这样的步骤?
- 谁是高斯, 为何这个方法被称作消元法?

内功心法, 贯通全书.

3.1 什么是线性方程组?

先说说什么是线性方程组.

Definition 3.1. 所谓线性方程组 *(Linear Equations)* 是指形式为

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad \cdots\cdots \\ a_{m1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n = b_s \end{cases}$$

的方程组, 其中 $x_1, x_2, \cdots, x_n$ 代表 n 个未知数, m 个线性方程的个数, a_{ij} 称为方程组的系数, b_i 称为常数项.

思考为何采取双角标描述线性方程组系数. 思考什么是线性.

Definition 3.2. 所谓方程组的一个解就是指由 n 个数 $k_1, k_2, \cdots, k_n$ 组成有序数组, 当 $x_1, x_2, \cdots, x_n$ 分别用 $k_1, k_2, \cdots, k_n$ 代入后, 每个线性方程都变成恒等式.

方程组的解的全体称为它的解集合. 如果两个线性方程组有相同的解集合, 称这两个线性方程组是同解的.

Example 3.1. 二元一次方程组

$$\begin{cases} 3x + 1y = 4 \\ 1x + 2y = 3 \end{cases} \iff \begin{pmatrix} 3 & 1 & 4 \\ 1 & 2 & 3 \end{pmatrix} \qquad r_1 \times (-1/3) + r_2$$

$$\begin{cases} 3x + 1y = 4 \\ 0x + \dfrac{5}{3}y = \dfrac{5}{3} \end{cases} \iff \begin{pmatrix} 3 & 1 & 4 \\ 0 & \dfrac{5}{3} & \dfrac{5}{3} \end{pmatrix} \qquad r_2 \times (-3/5) + r_1$$

$$\begin{cases} 3x + 0y = 3 \\ 0x + \dfrac{5}{3}y = \dfrac{5}{3} \end{cases} \iff \begin{pmatrix} 3 & 0 & 3 \\ 0 & \dfrac{5}{3} & \dfrac{5}{3} \end{pmatrix} \qquad r_1 \times (1/3), r_2 \times (3/5)$$

$$\begin{cases}1x+0y=1\\0x+1y=1\end{cases}\iff\begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}$$

得到 $x=1$, $y=1$. 这里 r_1 和 r_2 分别表示第 1 行和第 2 行 (或者是对应的方程). 3 和 5/3 就是第一, 二个主元.

左边是完整的记号, 表达了线性方程组的同解变形. 右边是简写的记号, 略去了未知数, 只记录方程组系数和常数项的变化, 与左边完全等价.

线性方程组看作两条平面上的直线, 消元法时, 则逐渐将直线平行于坐标轴. 分别将四组线性方程组看作相交的两直线, 我们说线性方程组们同解, 其实就是两直线的交点不变.

3.2 三板斧劈到了谁?

消元法要诀: 让非零元聚集在右上角 (制造更多的零), 然后回代, 将非零元只聚集在对角元. 或者说, 尽可能一个线性方程一个未知数, 不同方程不同未知数. 总结为:

- 依次按顺序用主元将其下的系数消为零;
- 依次按逆序用主元将其上的系数消为零.

消元法的三种基本变换:

- 用一非零数乘某个线性方程;
- 把一个线性方程的倍数加到另一个线性方程;
- 互换两个线性方程的位置.

这三个变换称为线性方程组的初等变换.

Remark 3.1. 根据定义证明上述三类线性方程组的初等变换将线性方程组变成同解的线性方程组. 思考是否存在其它变换保持线性方程组的解集合不变. 特别地, 如果线性方程组无解, 即解集合是空集, 则通过线性方程组的初等变换得到的线性方程组也无解.

消元法的基本步骤:

- 利用初等变换, "自上而下, 自下而上" 消元;
- 去掉 "$0=0$";
- 若有 "$0=1$", 则无解;
- 若未知数等于方程数, 则有唯一解, 否则有无穷多解.

另外, 未知数的相对位置可以调换. 也就是可以以列调换线性方程组的系数

三元一次方程组

$$\begin{cases}x-y+z=6\\3x+y+z=4\\3x+2y=5\end{cases}\iff\begin{pmatrix}1&-1&1&6\\3&1&1&4\\3&2&0&5\end{pmatrix}$$

$$\begin{cases} x-y+z=6 \\ 3x+y+z=4 \\ 2x+z=1 \end{cases} \iff \begin{pmatrix} 1 & -1 & 1 & 6 \\ 3 & 1 & 1 & 4 \\ 2 & 0 & 1 & 1 \end{pmatrix}$$

对于未知数个数与方程组个数不同的线性方程组也可以利用消元法进行计算.

Example 3.2.

$$\begin{cases} x-y+z=6 \\ 3x+y+z=4 \end{cases} \iff \begin{pmatrix} 1 & -1 & 1 & 6 \\ 3 & 1 & 1 & 4 \end{pmatrix} \qquad r_1 \times (-3) + r_2$$

$$\begin{cases} x-y+z=6 \\ 0x+4y-2z=-14 \end{cases} \iff \begin{pmatrix} 1 & -1 & 1 & 6 \\ 0 & 4 & -2 & -14 \end{pmatrix} \qquad r_2 \times (1/4)$$

$$\begin{cases} x-y+z=6 \\ 0x+y-\frac{1}{2}z=-\frac{7}{2} \end{cases} \iff \begin{pmatrix} 1 & -1 & 1 & 6 \\ 0 & 1 & -\frac{1}{2} & -\frac{7}{2} \end{pmatrix} \qquad r_1 + r_2$$

$$\begin{cases} x+0y+\frac{1}{2}z=\frac{5}{2} \\ 0x+y-\frac{1}{2}z=-\frac{7}{2} \end{cases} \iff \begin{pmatrix} 1 & 0 & \frac{1}{2} & \frac{5}{2} \\ 0 & 1 & -\frac{1}{2} & -\frac{7}{2} \end{pmatrix}$$

所以, $x=\frac{5}{2}-\frac{1}{2}k$, $y=-\frac{7}{2}+\frac{1}{2}k$, $z=k$, 任取 $k \in \mathbb{R}$.

$$\begin{cases} x-y+z=6 \\ 2x-2y+2z=12 \end{cases} \iff \begin{pmatrix} 1 & -1 & 1 & 6 \\ 2 & -2 & 2 & 12 \end{pmatrix}$$

以 $n=m=3$ 为例, 一个线性方程代表三维空间中的一个平面. 如果三个平面相交于一点, 则方程组解唯一. 如果三个平面相交于一条线, 则方程组解有无穷多个, 都落在这条交线上. 如果三个平面相交于一个平面, 则方程组解有无穷多个, 都落在这个交平面上. 如果三个平面无交点, 则方程组无解.

Remark 3.2. 请阅读 *[3]* 第三章, 第一节 – 消元法. 通过计算题完全明白消元法的解题步骤. 消元法的思想和方法是线性代数上半册最关键的.

线性方程组 $\to$ 解, 解集合, 同解 $\to$ 利用线性方程组的初等变换将原线性方程组对角元上下的系数都消为零 $\to$ 去掉 "0=0"$\to$ 若存在 "0=1", 则无解. 否则, 未知数等于方程数, 则唯一解; 未知数大于方程数, 则多解.

Remark 3.3. 视频是青年人不可分离的生命伴侣和导师.

- 厦门大学 – 高等代数 – 第二章 – 第 *1* 节 – 消元法

  http://www.icourses.cn/sCourse/course_3077.html

- 麻省理工学院 – 线性代数 – 第 *1* 集 – 方程组的几何解释

  http://open.163.com/movie/2010/11/7/3/M6V0BQC4M_M6V29E773.html

- 麻省理工学院 – 线性代数 – 第 *2* 集 – 矩阵消元

  http://open.163.com/movie/2010/11/P/P/M6V0BQC4M_M6V29EGPP.html

4 行列式 – 隐藏在 “大公式” 背后的真相

利用行列式可以部分解决线性方程组求解的问题.

- **什么是行列式, 为何称之为行列式?**
- **为何只能部分解决, 哪部分可以解决?**
- **解决的思路是什么?**
- **具体步骤是什么?**
- **为什么是这样的步骤?**
- **是否可以改进?**
- **与消元法有什么关联?**

外家功夫, 第一招, 方形线性方程组的判定条件和解的表达.

要诀: 基于消元法构建行列式应有的性质; 基于线性方程组的几何意义构建行列式的几何意义.

行列式是一个数, 它由 $n\times n$ 个数计算得到的 1 个数. 行列式从它的英文 – Determinant 更能体现它的来源. 最初希望通过方形线性方程组的所有 $n\times n$ 个系数构造出一个数. 当这个数不为零时, 线性方程组有唯一解. 行列式真正的名字其实是判别式. 这正如一元二次方程是否有解, 有多少解, 可以通过另外一个我们更熟知的判别式确定. 另外, 我们可以通过一元二次方程的判别式将解表达出来. 同样, 我们也希望利用行列式将线性方程组的解表达出来, 这便是 Cramer 法则.

起初行列式是为了判断 n 个未知数 n 个线性方程是否有唯一解而提出的. n 个未知数 n 个线性方程对应的判别式称之为 n 阶行列式. n 的不同对应的判别式自然不同, 那么它们之间是否有联系呢? 是否有一个统一的, 递归的定义呢? 无论如何, 行列式是由 $n\times n$ 个数定义的一个数.

要想完全说清楚行列式, 就要明白 n 个未知数 n 个线性方程的线性方程组解存在唯一的充要条件. 可以从 $n=2,3$ 说起, 这个情况可以利用消元法计算得到, 但写出公式也是比较繁杂, 毕竟分别是 4 个数和 9 个数的一个公式. 如果 n 很大, 公式的长度就难以想象了. 但归根结底我们是想研究线性方程组解的存在唯一性, 而消元法恰是解决这个问题的良方. 消元法催生行列式, 虽然历史上行列式的提出早于消元法.

4.1 线性方程组的几何意义

数形结合总是一个非常好用的数学方法. 正如一元二次方程和抛物线的完美结合一样, 线性方程和平面有不解之缘.

对于一个二元线性方程

$$ax_1+bx_2=b_1$$

它无非是一条直线而已, 而且这条直线的法向量是 $\vec{n}_1=(a,b)$. 那么线性方程组

$$\begin{cases} ax_1+bx_2=b_1,\\ cx_1+dx_2=b_2. \end{cases}$$

什么时候上述线性方程组存在唯一解呢? 即, 两条直线相交于一点. 换种思维, 两条直线的法向量不平行. 再换种思维, 这**两个法向量张成的平行四边形面积不为零**. 换句话说, 线性方程组存在唯一

解的充要条件就是这两个法向量构成的平行四边形面积不为零. 就是它, 这个平行四边面积就是我们要找的那个判别线性方程组存在唯一解的数. 更重要的是, 这个数可以非常简单的推广到 n 个未知数 n 个线性方程的情形. 即, n 维空间中的 n 个超平面的 n 个法向量张成的“平行四边形” 的“面积”.

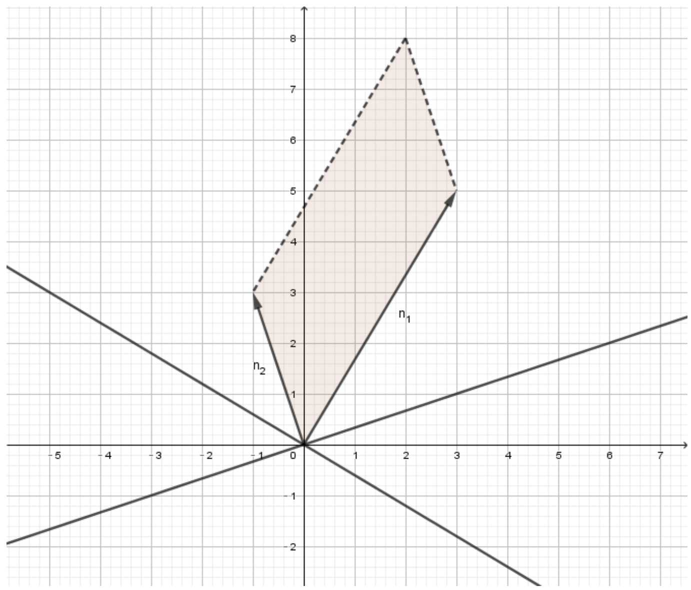


Figure 4.1: 几何意义

下面就探讨这个面积如何计算. 思考, 为什么平行四边形面积等于底乘以高. 如何用数学的语言表达你的答案.

首先我们用下述符号表示这个平行四边形的面积

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

首先, 根据图4.2, 容易得到,

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1, \ \begin{vmatrix} ka & kb \\ c & d \end{vmatrix} = k \begin{vmatrix} a & b \\ c & d \end{vmatrix},$$

$$\begin{vmatrix} a + a' & b + b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix},$$

以及

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c + ka & d + kb \end{vmatrix}.$$

这样, 我们基于上述规定, 可以得到 “交换两行, 行列式变号”. 这个规定也使得两行相同的行列

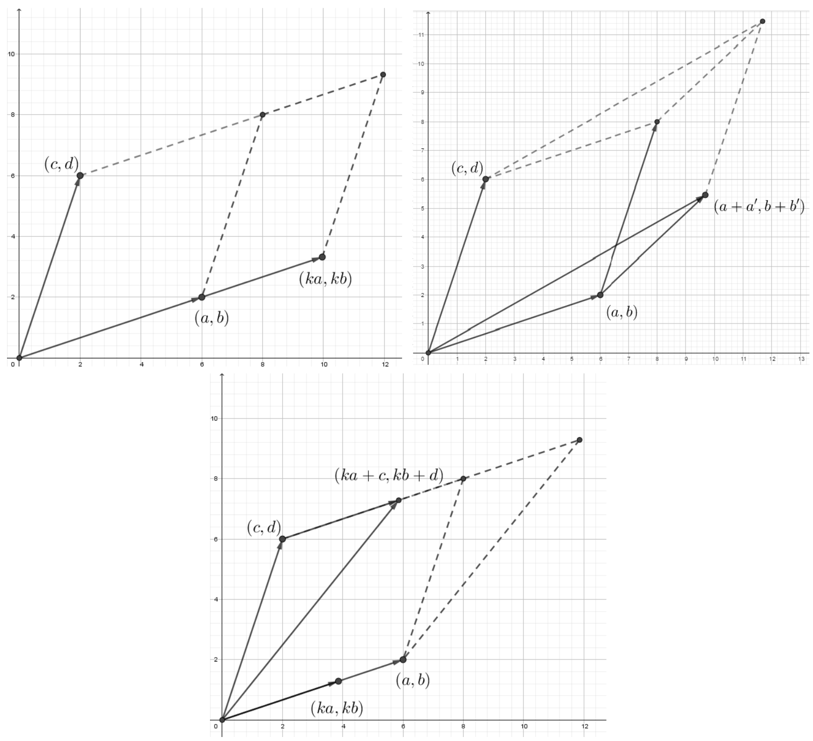


Figure 4.2: 几何性质

式为零.

$$\begin{aligned}\begin{vmatrix}a & b\\ c & d\end{vmatrix} &= \begin{vmatrix}a & b\\ c+a & d+b\end{vmatrix} = \begin{vmatrix}a+(-1)(c+a) & b+(-1)(d+b)\\ c+a & d+b\end{vmatrix}\\ &= \begin{vmatrix}-c & -d\\ c+a & d+b\end{vmatrix} = \begin{vmatrix}-c & -d\\ a & b\end{vmatrix} = -\begin{vmatrix}c & d\\ a & b\end{vmatrix}\end{aligned}$$

值得注意的是, 某行每个元素分裂成两个数的行列式等于分别两个行列式的和. 以及某行每个元素乘以一个数的行列式等于原行列式乘以这个数. 这个规定可以归结为行列式关于固定的一行有线性性.

所以,

$$\begin{aligned}\begin{vmatrix}a & b\\ c & d\end{vmatrix} &= \begin{vmatrix}a & 0\\ c & d\end{vmatrix} + \begin{vmatrix}0 & b\\ c & d\end{vmatrix} = \begin{vmatrix}a & 0\\ c & 0\end{vmatrix} + \begin{vmatrix}a & 0\\ 0 & d\end{vmatrix} + \begin{vmatrix}0 & b\\ c & 0\end{vmatrix} + \begin{vmatrix}0 & b\\ 0 & d\end{vmatrix}\\ &= ac\begin{vmatrix}1 & 0\\ 1 & 0\end{vmatrix} + ad\begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix} + bc\begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix} + bd\begin{vmatrix}0 & 1\\ 0 & 1\end{vmatrix} = ad-bc.\end{aligned}$$

这就是二阶行列式的定义啦.

Example 4.1.

$$\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} = \begin{vmatrix}a_{11} & 0 & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}0 & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}$$

$$
\begin{aligned}
&=\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&=\begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&\quad + \cdots + \cdots \\
&=\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & a_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{vmatrix} \\
&\quad + \begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix} \\
&=(-1)^{\tau(123)} a_{11}a_{22}a_{33} + (-1)^{\tau(132)} a_{11}a_{23}a_{32} \\
&\quad + (-1)^{\tau(231)} a_{12}a_{23}a_{31} + (-1)^{\tau(213)} a_{12}a_{21}a_{33} \\
&\quad + (-1)^{\tau(312)} a_{13}a_{21}a_{32} + (-1)^{\tau(321)} a_{13}a_{22}a_{31} \\
&=\sum_{j_1j_2j_3} (-1)^{\tau(j_1j_2j_3)} a_{1j_1}a_{2j_2}a_{3j_3} \ (\text{大公式}).
\end{aligned}
$$

n 阶行列式的定义就是根据二阶, 三阶行列式的计算方法归纳总结出来的. n 阶行列式是 $n!$ 个行列式的和, 其中每个行列式中每行每列只有一个元素. 再根据交换行, 行列式变号, 以及“单位正方形面积”为 1, 得到 n 阶行列式的“大公式”定义.

4.2 行列式与线性方程组的初等变换

回顾高斯消元法求解下述线性方程组的过程:

$$
\begin{cases} 3x + 1y = 4 \\ 1x + 2y = 3 \end{cases} \iff \begin{pmatrix} 3 & 1 & 4 \\ 1 & 2 & 3 \end{pmatrix}
$$

$$
\begin{cases} 3x + 1y = 4 \\ 0x + \frac{5}{3}y = \frac{5}{3} \end{cases} \iff \begin{pmatrix} 3 & 1 & 4 \\ 0 & \frac{5}{3} & \frac{5}{3} \end{pmatrix}
$$

$$
\begin{cases} 3x + 0y = 3 \\ 0x + \frac{5}{3}y = \frac{5}{3} \end{cases} \iff \begin{pmatrix} 3 & 0 & 3 \\ 0 & \frac{5}{3} & \frac{5}{3} \end{pmatrix}
$$

$$
\begin{cases} 1x + 0y = 1 \\ 0x + 1y = 1 \end{cases} \iff \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}
$$

仔细想来, 线性方程组的系数构成的行列式是这样变化的

$$
\begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ 0 & \frac{5}{3} \end{vmatrix} = \begin{vmatrix} 3 & 0 \\ 0 & \frac{5}{3} \end{vmatrix} = 3 \times \frac{5}{3} \times \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 5.
$$

这样看来, 消元法求解线性方程组和计算行列式, 看似两个不同的东西, 前者是给出一个集合 (解集合), 后者是给出一个数 (行列式). 但是计算过程是如此的相似.

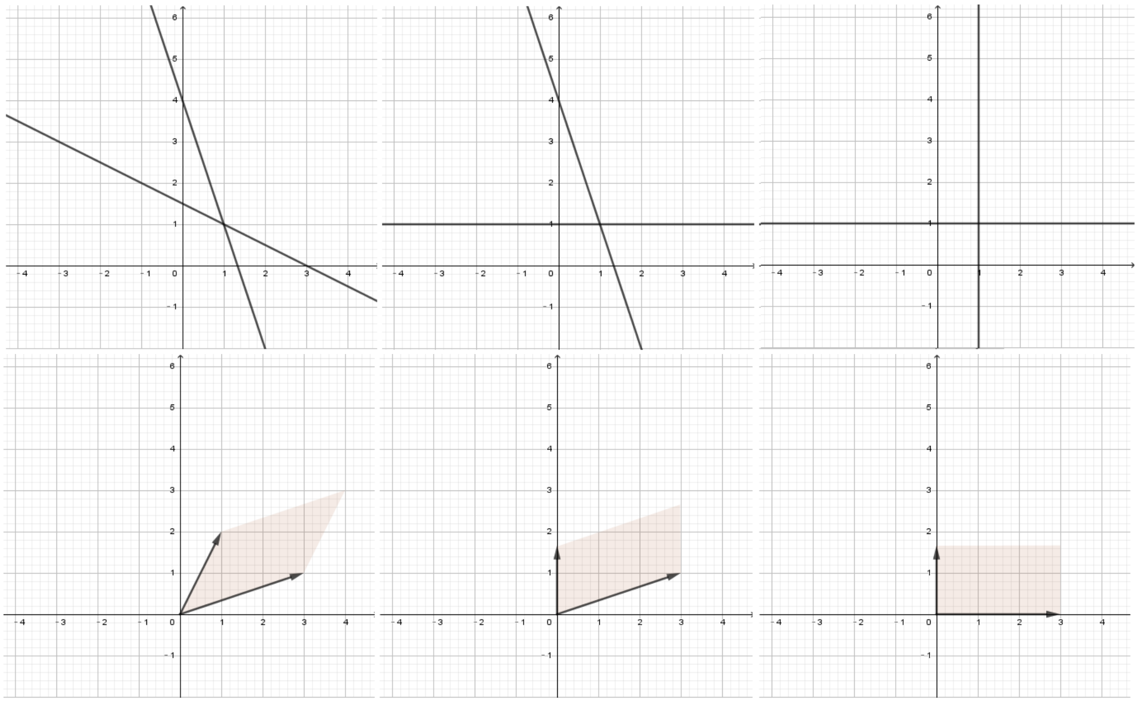


Figure 4.3: 高斯消元法与行列式计算的几何意义

平行四边形是通过直线的法向量张成的. 根据平行四边形的面积等于“底乘以高”, 从上图也能很清晰地看出上述三个平行四边形的面积相等.

另外, 高斯消元法的“三板斧”:

(1) 用一非零数乘某个线性方程;

(2) 把一个线性方程的倍数加到另一个线性方程;

(3) 互换两个线性方程的位置.

与行列式的性质一一对应:

(1) 某一行乘以实数的行列式等于该实数乘以原行列式;

(2) 将某行乘以某实数加到另一行后的行列式保持不变;

(3) 交换两行, 行列式变号.

高斯消元法保证线性方程组的解集合不变, 再根据行列式的性质, 它也保证了线性方程组对应的行列式不会从“等于零”变为“不为零”, 也不会从“不为零”变为“等于零”这样我们就更深入地认识到, 判断线性方程组解存在唯一的充要条件是线性方程组对应的行列式不为零. 需要注意的是, 这里讨论的线性方程组的方程数和未知数是相等的.

4.3 关键性质

从“大公式”出发给出下述性质. 对于 3 阶行列式, 这些性质可以具体写成数学公式. 另外, 根据前三个性质也可以等价地定义行列式.

(1) 对角线上全为 1, 其它元素全为 0 的行列式为 1.

(2) 交换两行的顺序, 行列式变号.

(3) 某一行乘以实数的行列式等于该实数乘以原行列式; 某行每个元素分裂成两个数的行列式等于分别两个行列式的和.

最后一个性质说明行列式关于行是线性的. 基于上面的三个性质可以证明:

(4) 某两行相等的行列式为 0. (交换两行, 行列式变号)

(5) 将某行乘以某实数加到另一行后的行列式保持不变.

(6) 某行全是零的行列式为 0.

(7) 上 (下) 三角形行列式等于对角线元素的乘积.

容易证明性质 (2-6) 将行换为列同样成立. 只需利用“大公式”证明性质 (2,3) 对列成立即可. 通过列的性质 (2,3) 同理证明 (4,5,6). 另外, 可以通过“大公式”证明

(8)
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$$

(HW) 利用列的性质 (2,3) 证明性质 (4,5,6) 对列也成立.

(HW) 基于上述性质给出 3 阶行列式的计算公式.

4.4 计算方法

行列式可以通过“将某行乘以某实数加到另一行”的方法计算. 也可通过“行列式按行有线性性”将高阶行列式转换为低阶行列式.

前面说了, 二阶行列式的几何意义是平行四边形的面积 (有正负之分), 三阶行列式的几何意义是平行六面体的体积 (有正负之分). 我们在中学所学的, 平行四边形的面积是“底 × 高”, 平行六面体的体积是“底面积 × 高”. 那么, 在行列式的计算中, 哪里是“底”哪里是“高”呢?

我们首先以一个例子来说明:

$$\begin{vmatrix} 2 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ -1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 2 & 1 \\ 2 & 3 \end{vmatrix} \times 2 = (6-2) \times 2 = 8,$$

这可以理解为三个向量

$$\overrightarrow{OA} = (2,0,0),\ \overrightarrow{OB} = (1,2,1),\ \overrightarrow{OC} = (-1,2,3)$$

张成的平行六面体的体积, 如下图所示.

第一个等式成立的几何意义是, $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 张成的平行四边形与 $\overrightarrow{OA}$ 与 $\overrightarrow{OB_1}$ 张成的平行四边形在同一个平面上且面积相等, 那么

$$\overrightarrow{OA},\ \overrightarrow{OB},\ \overrightarrow{OC}$$

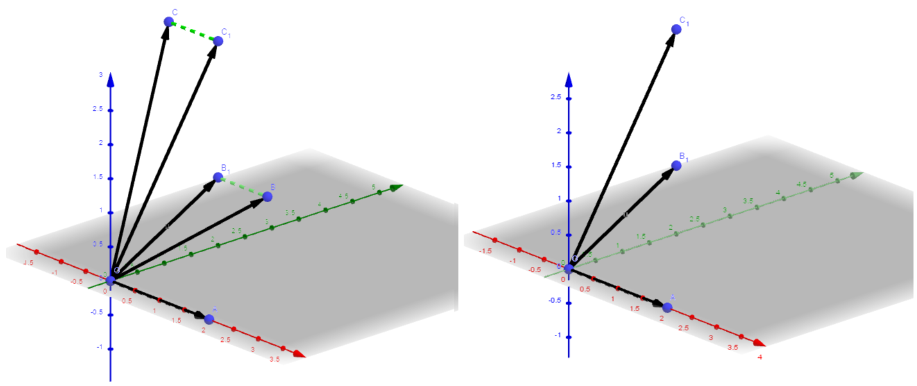


Figure 4.4: 三阶行列式

张成的平行六面体的体积与

$$\overrightarrow{OA},\ \overrightarrow{OB_1}=(0,2,1),\ \overrightarrow{OC}$$

张成的平行六面体的体积相等. 同理, 得到第二个等式成立的几何意义. 这两个等式成立, 都是性质“某行乘以某数加到另外一行去, 行列式保持不变” 的运用. 那么第三个等式为什么对呢? 那就是“底面积乘以高” 的结果, 其中

$$底面积=\begin{vmatrix}2 & 1\\ 2 & 3\end{vmatrix},\ 高=2.$$

底面积是由

$$\overrightarrow{OB_1}=(0,2,1),\ \overrightarrow{OC_1}=(0,2,3)$$

张成平行四边形的面积, 而且这个平行四边形与 $\overrightarrow{OA}$ 是垂直的, 那么 $\overrightarrow{OA}$ 的长度就是高.

说了这么多, 其实就是希望大家明白, 为什么

$$\begin{vmatrix}2 & 0 & 0\\ 1 & 2 & 1\\ 2 & 2 & 3\end{vmatrix}=\begin{vmatrix}2 & 1\\ 2 & 3\end{vmatrix}\times 2$$

成立. 无非就是平行六面体的体积等于“底面积”×“高”. 需要注意的是, 这里平行六面体已经等面积地变化过两次. 之所以要变化, 原因是让平行六面体的底面与高是垂直的.

抽象地来说, 我们就可以理解下述三个等式

$$\begin{vmatrix}a_{11} & 0 & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=a_{11}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix},$$

$$\begin{vmatrix}0 & a_{12} & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=-a_{12}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix},$$

$$\begin{vmatrix}0 & 0 & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix}=a_{13}\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix}.$$

这里需要注意符号, 三个等式的符号分别是“正负正”. 符号的出现起源于“交换两行或两列, 行列式变号”. 也可以从“大公式”中得到.

那么, 我们就可以引入代数余子式的概念, 结合“行列式按行的线性性”得到

$$\begin{aligned}\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} = \begin{vmatrix}a_{11} & 0 & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}0 & a_{12} & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}0 & 0 & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} \\ =a_{11}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix} - a_{12}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{13}\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix} \\ =\sum_{i=1}^{3} a_{1i}A_{1i} \text{ (降阶公式)}.\end{aligned}$$

基于性质 (2, 3, 5), 可以引入利用代数余子式计算行列式. 对于二阶行列式

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

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

这样的话, 就把三阶行列式的计算转化为二阶行列式的计算. 正如, 三维平行六面体的体积计算转化为一维高度和二维平行四边形面积的计算.

总结一下, 计算行列式有三种方法;

- “大公式”: $n!$ 项的加减, 每一项是 n 个数的乘积;
- “降阶公式”: n 阶行列式转化为 n 个 $n-1$ 阶行列式的计算;
- “性质 (5)”: “某行乘以某数加到另外一行去, 行列式保持不变”.

那么三阶行列式按照第一行展开的话, 可以得到三个二阶行列式求和.

$$\begin{aligned}\begin{vmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} = \begin{vmatrix}a_{11} & 0 & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}0 & a_{12} & 0\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}0 & 0 & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} \\ =a_{11}\begin{vmatrix}a_{22} & a_{23}\\ a_{32} & a_{33}\end{vmatrix} - a_{12}\begin{vmatrix}a_{21} & a_{23}\\ a_{31} & a_{33}\end{vmatrix} + a_{13}\begin{vmatrix}a_{21} & a_{22}\\ a_{31} & a_{32}\end{vmatrix} \\ =\sum_{i=1}^{3} a_{1i}A_{1i} \text{ (降阶公式)} \\ =\begin{vmatrix}a_{11} & 0 & 0\\ a_{21} & 0 & 0\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}a_{11} & 0 & 0\\ 0 & a_{22} & 0\\ a_{31} & a_{32} & a_{33}\end{vmatrix} + \begin{vmatrix}a_{11} & 0 & 0\\ 0 & 0 & a_{23}\\ a_{31} & a_{32} & a_{33}\end{vmatrix} \\ + \cdots + \cdots \\ =\begin{vmatrix}a_{11} & 0 & 0\\ 0 & a_{22} & 0\\ 0 & 0 & a_{33}\end{vmatrix} + \begin{vmatrix}a_{11} & 0 & 0\\ 0 & 0 & a_{23}\\ 0 & a_{32} & 0\end{vmatrix} + \begin{vmatrix}0 & a_{12} & 0\\ 0 & 0 & a_{23}\\ a_{31} & 0 & 0\end{vmatrix}\end{aligned}$$

$$
\begin{aligned}
&+\begin{vmatrix} 0 & a_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & a_{32} & 0 \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix} \\
=&(-1)^{\tau(123)}a_{11}a_{22}a_{33}+(-1)^{\tau(132)}a_{11}a_{23}a_{32} \\
&+(-1)^{\tau(231)}a_{12}a_{23}a_{31}+(-1)^{\tau(213)}a_{12}a_{21}a_{33} \\
&+(-1)^{\tau(312)}a_{13}a_{21}a_{32}+(-1)^{\tau(321)}a_{13}a_{22}a_{31} \\
=&\sum_{j_1j_2j_3}(-1)^{\tau(j_1j_2j_3)}a_{1j_1}a_{2j_2}a_{3j_3}\ (\text{大公式}).
\end{aligned}
$$

所以, 高阶行列式的计算可以降阶, 最后降到 1 阶行列式, 得到“大公式”. 从而也可以引入逆序数的概念 (描述行行交换的次数, 使得行列式变成对角的).

对于三阶行列式, 我们利用由行的性质推得到大公式证明行的性质 (1, 2, 3) 同样对列也成立.

$$
\begin{aligned}
\begin{vmatrix} a_{11} & a_{13} & a_{12} \\ a_{21} & a_{23} & a_{22} \\ a_{31} & a_{33} & a_{32} \end{vmatrix}=&\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & a_{23} & 0 \\ 0 & 0 & a_{32} \end{vmatrix}+\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{22} \\ 0 & a_{33} & 0 \end{vmatrix}+\begin{vmatrix} 0 & a_{13} & 0 \\ 0 & 0 & a_{22} \\ a_{31} & 0 & 0 \end{vmatrix} \\
&+\begin{vmatrix} 0 & a_{13} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{32} \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{12} \\ a_{21} & 0 & 0 \\ 0 & a_{33} & 0 \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{12} \\ 0 & a_{23} & 0 \\ a_{31} & 0 & 0 \end{vmatrix} \\
=&-\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.
\end{aligned}
$$

$$
\begin{aligned}
\begin{vmatrix} a_{11} & ka_{12} & a_{13} \\ a_{21} & ka_{22} & a_{23} \\ a_{31} & ka_{32} & a_{33} \end{vmatrix}=&\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & ka_{22} & 0 \\ 0 & 0 & a_{33} \end{vmatrix}+\begin{vmatrix} a_{11} & 0 & 0 \\ 0 & 0 & a_{23} \\ 0 & ka_{32} & 0 \end{vmatrix}+\begin{vmatrix} 0 & ka_{12} & 0 \\ 0 & 0 & a_{23} \\ a_{31} & 0 & 0 \end{vmatrix} \\
&+\begin{vmatrix} 0 & ka_{12} & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{13} \\ a_{21} & 0 & 0 \\ 0 & ka_{32} & 0 \end{vmatrix}+\begin{vmatrix} 0 & 0 & a_{13} \\ 0 & ka_{22} & 0 \\ a_{31} & 0 & 0 \end{vmatrix} \\
=&k\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.
\end{aligned}
$$

(HW) 利用大公式证明三阶行列式的性质 (3) 对于列也成立.

(HW) 利用行列式关于列的性质 (1, 2, 3), 证明行列式关于行的性质 (4, 5, 6) 对列也成立.

(HW) 思考性质 (7) 的几何意义.(底乘以高)

4.5 Cramer 法则

经典的 Cramer 法则的证明, 是非常精巧的. 每个人都可以从中学到很意思的东西. 这里我并不是想提出一个新的证明, 而是出于好奇. 既然高斯消元法可以求解线性方程组, 同时 Cramer 法则可以判别线性方程组解的存在唯一性, 且在解存在唯一时, 可以利用行列式写出解的表达式. 这里我们从一个特殊的例子出发, 去理解这两者的联系.

二元一次方程组

$$\begin{cases}3x+1y=4\\1x+2y=3\end{cases} \iff \begin{pmatrix}3&1&4\\1&2&3\end{pmatrix} \qquad r_1\times(-1/3)+r_2$$

$$\begin{cases}3x+1y=4\\0x+\dfrac{5}{3}y=\dfrac{5}{3}\end{cases} \iff \begin{pmatrix}3&1&4\\0&\dfrac{5}{3}&\dfrac{5}{3}\end{pmatrix} \qquad r_2\times(-3/5)+r_1$$

$$\begin{cases}3x+0y=3\\0x+\dfrac{5}{3}y=\dfrac{5}{3}\end{cases} \iff \begin{pmatrix}3&0&3\\0&\dfrac{5}{3}&\dfrac{5}{3}\end{pmatrix}$$

根据行列式的性质得到

$$\begin{vmatrix}3&1\\1&2\end{vmatrix}=\begin{vmatrix}3&0\\0&5/3\end{vmatrix},\ \begin{vmatrix}3&4\\1&3\end{vmatrix}=\begin{vmatrix}3&3\\0&5/3\end{vmatrix},\ \begin{vmatrix}4&1\\3&2\end{vmatrix}=\begin{vmatrix}3&0\\5/3&5/3\end{vmatrix}.$$

大家应该很认同第一个等式, 可能会对第二个和第三个的等式产生疑问. 其实第二个和第三个等式中的行列式不过是将列交换了一下而已.

显然, 根据消元法和上述三个等式得到,

$$x=\frac{\begin{vmatrix}3&3\\0&5/3\end{vmatrix}}{\begin{vmatrix}3&0\\0&5/3\end{vmatrix}}=\frac{\begin{vmatrix}3&4\\1&3\end{vmatrix}}{\begin{vmatrix}3&1\\1&2\end{vmatrix}},\ y=\frac{\begin{vmatrix}3&0\\5/3&5/3\end{vmatrix}}{\begin{vmatrix}3&0\\0&5/3\end{vmatrix}}=\frac{\begin{vmatrix}4&1\\3&2\end{vmatrix}}{\begin{vmatrix}3&1\\1&2\end{vmatrix}}.$$

这便是 Cramer 法则呀.

我们利用抽象的符号再次说明上述推导. 我们假设线性方程组进行消元法时只做“将某行乘以某实数加到另一行”(这样对应行列式保持不变), 且线性方程组解存在唯一. 那么, 通过高斯消元法将下述线性方程组

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

同解简化为

$$\begin{cases}d_1x+0y+0z=c_1\\0x+d_2y+0z=c_2\\0x+0y+d_3z=c_3\end{cases} \iff \begin{pmatrix}d_1&0&0&c_1\\0&d_2&0&c_2\\0&0&d_3&c_3\end{pmatrix}.$$

根据线性方程组存在唯一解, 则 d_1,d_2,d_3 都不为零. 同时,

$$x=\frac{c_1}{d_1}=\frac{\begin{vmatrix}c_1&0&0\\c_2&d_2&0\\c_3&0&d_3\end{vmatrix}}{\begin{vmatrix}d_1&0&0\\0&d_2&0\\0&0&d_3\end{vmatrix}},\ y=\frac{c_2}{d_2}=\frac{\begin{vmatrix}d_1&c_1&0\\0&c_2&0\\0&c_3&d_3\end{vmatrix}}{\begin{vmatrix}d_1&0&0\\0&d_2&0\\0&0&d_3\end{vmatrix}},\ z=\frac{c_3}{d_3}=\frac{\begin{vmatrix}d_1&0&c_1\\0&d_2&c_2\\0&0&c_3\end{vmatrix}}{\begin{vmatrix}d_1&0&0\\0&d_2&0\\0&0&d_3\end{vmatrix}}.$$

再根据行列式的性质, 得到

$$\begin{vmatrix} c_1 & 0 & 0 \\ c_2 & d_2 & 0 \\ c_3 & 0 & d_3 \end{vmatrix} = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad \begin{vmatrix} d_1 & c_1 & 0 \\ 0 & c_2 & 0 \\ 0 & c_3 & d_3 \end{vmatrix} = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix},$$
$$\begin{vmatrix} d_1 & 0 & c_1 \\ 0 & d_2 & c_2 \\ 0 & 0 & c_3 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}, \quad \begin{vmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

这样我们就得到 Cramer 法则:

线性方程组解存在唯一的充要条件是

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{vmatrix} \neq 0.$$

同时, 当线性方程组解存在唯一时,

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

这里还需要说明的是, 我们并不通过 Cramer 法则计算线性方程组, 它只是给出一个代数的公式来表达解的形式.

Example 4.2. 计算与两个不平行的 $\mathbb{R}^3$ 向量都垂直的向量. 根据垂直的关系列出方程组

$$\begin{cases} ax + by + cz = 0, \\ dx + ey + fz = 0. \end{cases}$$

利用消元法得到, 对任意 k

$$\begin{cases} x = k(bf - ce)/(ae - bd), \\ y = -k(af - cd)/(ae - bd), \\ z = k. \end{cases}$$

取 $k = (ae - bd)$, 得到

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & b & c \\ d & e & f \end{vmatrix}.$$

计算一个平面的法线方向 $\vec{n}$. 并且计算两个三维向量张成的平行四边形面积. 该平面的法向量, 得到平面与三个坐标平面之间的二面角, 进而解释空间平行四边形面积等于在三个坐标平面的投影平行四边形面积的平方和.

Example 4.3. 根据顶点的坐标, 计算三角形面积, 平行四边形的面积, 平行六面体的体积.

利用行列式的定义 *(三条性质)*, 定义平行四边形的面积, 它满足:

$$
\begin{aligned}
S(1,0;0,1) &= 1; \\
S(a,b;c,d) &= -S(c,d;a,b); \\
S(a,b;kc,kd) &= kS(a,b;c,d); \\
S(a,b;c+c',d+d') &= S(a,b;c,d)+S(a,b;c',d').
\end{aligned}
$$

其中 (a,b) 和 (c,d) 是形成平行四边形的两个向量.

4.6 真相只有一个

- 原名: 判别式 – 希望找到一个数判断 n 个未知数 n 个线性方程的线性方程组是否存在唯一解;
- 化名: 行列式 – 凡是关于行的性质, 列同样具备, 行与列是平等的;
- 表象: 大公式 – 它的复杂性源于行列式的逐行分裂, 且交换行列形成逆序数;
- 内心: 平行四边形的“有向”面积 – 三条性质的图像呈现出行列式的几何意义;
- 性格: 一共七条性质 – 前三条性质是本质, 后四条性质是外延;
- 父亲: 莱布尼兹;
- 兄弟: $\Delta = b^2 - 4ac$.

Remark 4.1. 请阅读 *[3]* 第二章 – 行列式. 行列式是一个非常有用的数. 用来判断方形方程组是否有唯一解.

线性方程组的几何意义 → 行列式的几何意义 →“大公式”→ 行的三个性质 → 行的四个性质 → 列的三个性质 → 列的四个性质 → 代数余子式 → 降阶公式 → 线性方程组的初等变换 → Cramer 法则.

Remark 4.2. 光阴给我们经验, 视频给我们知识.

- 厦门大学 – 高等代数 – 第一章 – 第 *4* 节 – 行列式

  http://www.icourses.cn/sCourse/course_3077.html
- 麻省理工学院 – 线性代数 – 第 *18* 集 – 行列式及其性质

  http://open.163.com/movie/2010/11/5/0/M6V0BQC4M_M6V2AP150.html
- 麻省理工学院 – 线性代数 – 第 *19* 集 – 行列式公式和代数余子式

  http://open.163.com/movie/2010/11/3/4/M6V0BQC4M_M6V2APP34.html

5 2 维向量空间与矩阵

主线: 线性方程组求解.

道理:

- 方形线性方程组, 利用行列式判别是否有解, Cramer 法则表述解的形式.
- 方形线性方程组, 利用矩阵的可逆性判别是否有解, 矩阵的逆矩阵表述解的形式.
- 一般线性方程组, 矩阵理论描述消元法; 线性空间描述解的存在性和解的形式: 四个子空间一个大图像.
- 一般线性方程组, 利用消元法实际计算线性方程组的解.
- 一般线性方程组, 看作线性映射的作用过程.

前两个是等价的, 第二个更具技巧性, 只对方形线性方程组. 第三个理论更深刻, 完全解决了线性方程组的理论问题. 第四个解决了线性方程组的计算问题. 第五个是对线性方程组的新认识. 行列式, 矩阵, 线性空间, 消元法, 线性映射 – 大家都围绕着线性方程组!

本节用三个视角看同一个问题.

一个问题: 求解线性方程组

$$\begin{cases} 3x + y = 4 \\ x + 2y = 3 \end{cases} . \tag{5.1}$$

行视角和列视角, 前者是平面几何中, 两个平面相交的问题. 后者是向量空间的线性组合问题. 另外一个视角就是线性映射, 将一个向量变成另外一个向量.

5.1 行视角

行视角可以进行计算 (消元法), 列视角的语言陈述解的存在唯一性. 线性方程组(5.1)解可能不存在或者不唯一. 从行视角, 可能没有交点后者有很多交点. 从列视角, 右端向量是否在列向量张成的向量空间中. 以线性映射的角度, 某个向量是否在它的值域里. 行视角很直接, 列视角很抽象.

线性方程组求解可以通过初等变换进行求解 (消元法), 这是行视角. 利用行列式表述线性方程组的解, 即为 Cramer 法则.

5.2 列视角

列视角就是如何确定组合系数 x 和 y 使得向量 $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ 和 $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ 形成 $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$. 即

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 2 \end{pmatrix} y = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

5.3 映射视角

集合与映射是数学最基本的要素. 线性方程组问题依然可以用集合与映射的语言描述和解决. 通过向量空间描述 “集合”, 利用矩阵描述 “映射”. 向量空间和矩阵恰是基于线性方程组提出的数学对象, 它们的本质就是集合和映射.

将矩阵看作一个函数, 它的定义域和值域都是向量空间 $\mathbb{R}^2$.

$$f:\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}3x+y\\x+2y\end{pmatrix}$$

或者

$$\begin{pmatrix}3&1\\1&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}3x+y\\x+2y\end{pmatrix}.$$

那么 $\begin{pmatrix}4\\3\end{pmatrix}$ 的原象就是线性方程组的解.

5.4 再进一步

从下面的等式看出

$$\begin{pmatrix}3&1\\1&2\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}3\\1\end{pmatrix}\text{ 和 }\begin{pmatrix}3&1\\1&2\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}.$$

它将一个正方形映射成一个平行四边形. 那么它把一个面积为 1 的正方形变成面积为 $\det A$ 的平行四边形 (有正负号). 显然, 映射 f 有线性性

$$f(\alpha\vec{a}+\beta\vec{b})=\alpha f(\vec{a})+\beta f(\vec{b}).$$

线性映射的复合又是什么意思? (矩阵的乘法)

如何描述线性映射 f 的逆映射? (矩阵的逆)

简要说明线性空间 V 的定义. “1128”:

$$\text{线性空间}=\text{集合}+\text{数域}+\text{加法和数乘}+\text{八条法则}.$$

- 1 个集合 S. 将感兴趣的所有元素都放这里, 每个元素称作向量.
- 1 个数域 $\mathbb{P}$. 一般考虑实数域或者复数域.
- 2 种运算. 加法是两个向量之间的加法, 数乘是数域中的数与集合中的向量的数乘. 这里的加法与数乘都是抽象地运算.
- 8 条法则. 使得加法和数乘相容. $k(s\vec{x}+t\vec{y})=(ks)\vec{x}+(kt)\vec{y}$.

Example 5.1. *设 $S=\mathbb{R}^n$ 即实数域上所有 n 元有序数组构成的集合, $\mathbb{P}=\mathbb{R}$, 在通常意义的加法和数乘运行定义下, 构成线性空间. 也称为 n 维向量空间.*

粗略地说, 线性空间不只是集合, 它是有“结构”的集合. “结构”体现在两种运算上, 它使得向量与向量之间有了关系. 我们将向量和数之间的加法和数乘运算统称为线性运算, 将由线性运算得到的向量称为原来向量的线性组合.

Example 5.2 (数学的美在于为不同的事物起了相同的名字)**.** *将股票市场中每种股票看作向量. 例如, 恒大地产和浦发银行这两支股票就是两个不同的向量 $\vec{x}$ 和 $\vec{y}$. 数域选择实数域 $\mathbb{R}$. 定义数乘和加法, $k\vec{x}$ 表示投资者拥有 k 份恒大地产的股票. $\vec{x}+\vec{y}$ 表示投资者分别拥有 1 份恒大地产和 1 份浦发银行的股票. 集合 S 囊括了所有由数乘和加法生成的向量, 它表示投资者在股票市场的所有投资组合. 显然满足八条性质. 这样就构成了一个线性空间. 在这个线性空间下我们可以研究一个投资者在股票市场的投资策略. 股票是向量, 线性组合是投资组合.*

Example 5.3. 集合 S 是所有的 2 维向量, 数域是实数域 $\mathbb{R}$. 加法和数乘是一般意义下 2 维向量和实数的加法和数乘. 显然满足八条性质. 这样就构成了一个线性空间, 记为 $\mathbb{R}^2$.

从线性空间的定义中我们并不需要定义一个向量的长度. 但是对于线性空间 $\mathbb{R}^n$, 向量长度的概念可以很容易的定义. 首先定义两个向量的数量积 (内积), 然后定义向量的长度和向量间的角度. 那么就可以描述两个向量是否垂直.

对于线性空间 $\mathbb{R}^2$ 中的向量 $\vec{v}=(v_1,v_2)^T$, $\vec{w}=(w_1,w_2)^T$,

$$\vec{v}\cdot\vec{w}=v_1w_1+v_2w_2.$$

$$\|\vec{v}\|=\sqrt{\vec{v}\cdot\vec{v}}.$$

(HW) 如果 $\|\vec{v}\|^2+\|\vec{w}\|^2=\|\vec{v}-\vec{w}\|^2$, 即 $\vec{v}\perp\vec{w}$. 验证, 此时 $\vec{v}\cdot\vec{w}=0$.

线性映射 (保持线性关系的映射, $f(a\vec{x}+b\vec{y})=af(\vec{x})+bf(\vec{y})$, 映射作用与线性运算可交换顺序)

$$\begin{aligned}&A:\mathbb{R}^2\to\mathbb{R}^2\\ &\begin{pmatrix}x\\y\end{pmatrix}\to\begin{pmatrix}a_{11}x+a_{12}y\\a_{21}x+a_{22}y\end{pmatrix}=A\begin{pmatrix}x\\y\end{pmatrix}.\end{aligned}$$

其中

$$A=\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}.$$

进而引出矩阵与向量的乘法运算 (映射作用在原象上得到象).

(HW) 验证所定义的映射 A 是线性映射.

矩阵 A 乘以向量 $\vec{a}$ 等于另一个向量 $\vec{b}$: 映射 A 作用在 $\vec{a}$ 等于 $\vec{b}$. 矩阵的线性运算 (加法和数乘): 映射的加法和数乘. 矩阵的乘法 $C=AB$: 映射的复合.

$$C(\begin{pmatrix}x\\y\end{pmatrix})=A(B(\begin{pmatrix}x\\y\end{pmatrix})).$$

矩阵相乘, 若将矩阵 B 的每一列都看作一个向量, 则

$$A\left(\vec{b}_1,\vec{b}_2,\cdots,\vec{b}_n\right)=\left(A\vec{b}_1,A\vec{b}_2,\cdots,A\vec{b}_n\right).$$

基于映射的复合, 可以定义映射的逆映射. 用矩阵的语言描述, 即为矩阵 A 的逆矩阵 A^{-1}.

$$\begin{pmatrix}x\\y\end{pmatrix}=I(\begin{pmatrix}x\\y\end{pmatrix})=A(A^{-1}(\begin{pmatrix}x\\y\end{pmatrix}))=A^{-1}(A(\begin{pmatrix}x\\y\end{pmatrix})).$$

其中 I 是恒等映射. 只有方阵才能求逆矩阵, 或者说, 只有定义域和值域相等的映射才求逆映射.

矩阵逆的计算与线性方程组的求解是同一回事.

$$\begin{pmatrix}1&3\\2&7\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix}\Leftrightarrow\begin{pmatrix}1&3\\2&7\end{pmatrix}\begin{pmatrix}a\\c\end{pmatrix}=\begin{pmatrix}1\\0\end{pmatrix}$$

and

$$\begin{pmatrix}1&3\\2&7\end{pmatrix}\begin{pmatrix}b\\d\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$$

那么能否利用消元法或者 Cramer 法则去计算和表达线性映射的逆映射或者矩阵的逆矩阵呢?

(HW) 利用消元法求解上述矩阵的逆矩阵. 尝试利用 Cramer 法则写出一般二阶方阵的逆矩阵的代数公式.

行列式是一个数. 几何意义: 线性变换 (定义域和值域相同的线性映射) 的面积比 (伸缩比). 两个向量生成的平行四边形的有向面积. 根据伸缩比这个几何意义给出行列式的三个性质.

向量组的秩向量组等价的线性相关与线性无关.

线性表出 $\rightarrow$ 线性相关与线性无关 $\rightarrow$ 向量组等价 $\rightarrow$ 极大线性无关组 $\rightarrow$ 向量组的秩

向量组的秩 $\rightarrow$ 矩阵的秩 (行秩, 列秩)

向量组等价 $\rightarrow$ 矩阵的 (行, 列) 初等变换

Remark 5.1. **请阅读** *[3]* **第六章 – 线性空间的定义与简单性质.**

集合 $\rightarrow$ 映射 $\rightarrow$ 线性空间 $\rightarrow$ “1128” $\rightarrow$ 线性运算 $\rightarrow$ 线性组合 $\rightarrow$ 数量积与长度 $\rightarrow$ 垂直 $\rightarrow$ 线性映射与矩阵 $\rightarrow$ 矩阵乘以向量 $\rightarrow$ 映射复合 $\rightarrow$ 矩阵乘以矩阵 $\rightarrow$ 恒等映射与恒等矩阵 $\rightarrow$ 逆映射与逆矩阵 $\rightarrow$ 行列式与面积 $\rightarrow$ 四大基本子空间 $\rightarrow$ 图解线性方程组.

6 矩阵 – 万法归一, 一归何处?

矩阵理论可以完全描述和解决线性方程组求解的问题.

- **什么是矩阵, 为何称之为矩阵?**
- **矩阵理论有哪些?**
- **如何描述, 用到哪些矩阵理论?**
- **解决的思路是什么?**
- **具体步骤是什么?**
- **为什么是这样的步骤?**
- **为何要用矩阵来研究线性方程组?**
- **与消元法有什么关联?**
- **与行列式有什么关联?**

内功心法: 变换 — 等价关系 — 分类 — 代表元 — 不变量.

要诀: 利用矩阵和它的理论精确描述消元法. 线性方程组表达为矩阵, 它的各种求解过程归于矩阵的初等变换.

6.1 概念引入

Figure 6.1: 黑客帝国 (The Matrix)

黑客帝国 (The Matrix), 是一部 1999 年的好莱坞科幻电影, 由沃卓斯基姐妹执导, 基努・里维斯、劳伦斯・菲什伯恩、凯莉・安摩丝及雨果・威文等人主演, 并由香港电影界的袁和平担任武术指导. 此片以其类似佛学的故事概论和子弹时间的特殊镜头及电脑特效著名, 在全球获取高票房,

并在 2003 年, 推出续集《黑客帝国 2: 重装上阵》及第三集《黑客帝国 3: 矩阵革命》, 在 2019 年, 因该片的 20 周年庆而推出了该片的 4K 修复版本, 将经典重返大银幕.

Matrix 的本意是子宫, 母体, 孕育生命的地方. 同时, 在数学名词中, 矩阵用来表示统计数据等方面的各种有关联的数据. 这个定义很好地解释了 Matrix 代码制造世界的数学逻辑基础. 在电影中, Matrix 不仅是一个虚拟程序, 也是一个实际存在的地方. 在这里, 人类的身体被放在一个盛满营养液的器皿中, 身上插满了各种插头以接受电脑系统的感官刺激信号. 人类就依靠这些信号, 生活在一个完全虚拟的电脑幻景中. 机器用这样的方式占领了人类的思维空间, 用人类的身体作为电池以维持自己的运行.

在电影中, Matrix 是一套复杂的模拟系统程序, 它是由具有人工智能的机器建立的, 模拟了人类以前的世界, 并用之控制人类. 在 Matrix 中出现的人物, 都可以看做是具有人类意识特征的程序. 这些程序根据所附着的载体不同有三类: 一类是附着在生物载体上的, 就是在 Matrix 中生活的普通人; 一类是附着在电脑芯片上的, 就是具有人工智能的机器, 这些载体通过硬件与 Matrix 连接; 另一类则是自由程序, 它没有载体, 诸如特工, 先知, 建筑师, 梅罗文加, 火车人等.

Matrix 是一个巨大的网络, 连接着无数人的意识, 系统分配给他们不同的角色, 就像电脑游戏中的角色扮演游戏一样, 只是他们没有选择角色的权利和意识. 人类通过这种联网的虚拟生活来维持自身的生存需要, 但 Matrix 中的智能程序, 也就是先知的角色, 发现在系统中有的人由于自主意识过强, 不能兼容系统分配的角色, 如果对他们不进行控制就会导致系统的不稳定, 进而导致系统崩溃. 因此编写 Matrix 的智能程序, 也就是建筑师就制造了“救世主”, 让他有部分自主意识, 并成为觉醒人类的领袖, 带领他们建造了锡安.

数码就是矩阵, 依照电影所言, Matrix 缔造了整个虚拟的数码时代. 真实的人类, 其肉体生活在“人造”的子宫 (matrix) 内, 其思想生活在“人造”的数码世界 (Matrix). 可见, 电影名 The Matrix 着实让人印象深刻.

在数学中, matrix 的中文译名是矩阵, 顾名思义, 矩形的阵列. 它是由数或符号组成的多行或多列的矩形阵列. 既然是矩形阵列, 也就是说组成矩阵的行或列的长度是一样的. 值得注意的是, 矩阵以一个单独的数学对象被定义其实并不是一蹴而就的. 数学家并不会无病呻吟, 随随便便提出一个定义. 他们是意识到了可以利用矩阵表达非常多的事物, 可以说, 矩阵创造世界.

矩阵这节从数学上就是研究矩阵各种运算以及它们之间的关系. 矩阵的一个非常重要的应用就是线性方程组求解的精确论述.

6.2 八种运算总述

矩阵的基本运算有八种, 即矩阵的“加法, 数乘, 转置, 行列式, 乘法, 逆, 伴随, 秩”. 那么, 不同运算之间是否有交换性呢? 比如说, 先做加法运算后做乘法运算, 是否等于, 先做乘法运算后做加法运算.

为什么交换性那么重要呢? 一句话, 合理利用交换性可以简化计算. 举个例子, 在我们小学学习加法的时候, 计算 $9+3$, 你可能会默默地扳起指头. 扳起第一个指头, 嘴里默念 10, 扳起第二个指头, 嘴里默念 11, 扳起第三个智斗, 嘴里默念 12. 那么, 我们就知道了 $9+3$ 等于 12. 按照这个想法, 我们应该如何计算 $3+9$ 呢? 扳起第一个指头, 嘴里默念 4, 扳起第二个指头, 嘴里默念 5, $\cdots$, 扳起第八个指头, 嘴里默念 11, 扳起第九个指头, 嘴里默念 12, 那么, 我们就知道了 $3+9$ 等于 12. 有人可能会说, 才不是这样呢. 算 $3+9$ 的过程, 应该和算 $9+3$ 的过程一样, 只需要扳三个指头就可以了, 哪里要这样麻烦, 扳九个指头数九次. 你要是这么想, 其实你已经用到了 $3+9=9+3$ 这样一条法则. 而且, 这样一条法则确实给你带来了方便. 试想, 如果你要计算 $8+85$ 且不知道“加法交换律”, 那扳指头的过程是相当吓人的.

下面将矩阵的“加法, 数乘, 乘法, 逆, 伴随, 转置, 行列式, 秩”的交换性进行总结如下:
加法

$$A+B=B+A,\ (A+B)+C=A+(B+C)$$

加法, 数乘

$$\begin{aligned}&k(A+B)=kA+kB\\&(k+l)A=kA+lA,\ (kl)A=k(lA)\end{aligned}$$

加法, 数乘, 转置

$$\begin{aligned}&(A+B)^T=A^T+B^T\\&(kA)^T=kA^T\\&(A^T)^T=A\end{aligned}$$

加法, 数乘, 转置, 行列式

$$\begin{aligned}&\det(A+B)\neq\det(A)+\det(B)\\&\det(kA)=k^n\det(A)\ (n\text{ 是 }A\text{ 的阶数})\\&\det(A^T)=\det(A)\end{aligned}$$

加法, 数乘, 转置, 行列式, 乘法

$$\begin{aligned}&(A+B)C=AC+BC,\ C(A+B)=CA+CB\\&(kA)B=k(AB)=A(kB)\\&(AB)^T=B^TA^T\\&\det(A)\det(B)=\det(AB)\\&AB\neq BA,\ (AB)C=A(BC)\end{aligned}$$

加法, 数乘, 转置, 行列式, 乘法, 逆

$$\begin{aligned}&(A+B)^{-1}\neq A^{-1}+B^{-1}\\&(kA)^{-1}=k^{-1}A^{-1}\ (k\neq 0)\\&(A^T)^{-1}=(A^{-1})^T=A^{-T}\\&(\det(A))^{-1}=\det(A^{-1})\\&(AB)^{-1}=B^{-1}A^{-1}\\&(A^{-1})^{-1}=A\end{aligned}$$

加法, 数乘, 转置, 行列式, 乘法, 逆, 伴随

$$\begin{aligned}&(A+B)^*\neq A^*+B^*\\&(kA)^*=k^{n-1}A^*\ (n\text{ 是 }A\text{ 的阶数},\ n\geq 2)\\&(A^T)^*=(A^*)^T\end{aligned}$$

$$\det(A^*) = (\det(A))^{n-1} \ (n \text{ 是 } A \text{ 的阶数})$$
$$(AB)^* = B^*A^*$$
$$(A^{-1})^* = (A^*)^{-1} = (\det(A))^{-1}A$$
$$(A^*)^* = (\det(A))^{n-2}A \ (n \text{ 是 } A \text{ 的阶数}, n \geq 2)$$

加法, 数乘, 转置, 行列式, 乘法, 逆, 伴随, 秩

$$\begin{aligned}
&\mathrm{rank}(A+B) \leq \mathrm{rank}(A) + \mathrm{rank}(B)\\
&\mathrm{rank}(kA) = \mathrm{rank}(A) \ (k \neq 0)\\
&\mathrm{rank}(A^T) = \mathrm{rank}(A)\\
&\mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}\\
&\mathrm{rank}(AB) \geq \mathrm{rank}(A) + \mathrm{rank}(B) - n\\
&\mathrm{rank}(A^{-1}) = \mathrm{rank}(A) = n\\
&\mathrm{rank}(A^*) = \begin{cases} n, & \text{if } \mathrm{rank}(A) = n,\\ 1, & \text{if } \mathrm{rank}(A) = n-1,\\ 0, & \text{otherwise} \end{cases}
\end{aligned}$$

6.3　矩阵的逆

矩阵之间的乘法源于线性映射的复合; 矩阵的逆矩阵源于线性映射的逆映射.

基于消元法计算矩阵的逆矩阵.

$$\begin{pmatrix} a_{11} & a_{12} & 1 & 0\\ a_{21} & a_{22} & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & b_{11} & b_{12}\\ 0 & 1 & b_{21} & b_{22} \end{pmatrix}.$$

抽象地讲,

$$\begin{pmatrix} A & I \end{pmatrix} \to \begin{pmatrix} I & A^{-1} \end{pmatrix}.$$

那么可以利用 Cramer 法则写出逆矩阵的代数表达式: 设 $A = \left(\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n\right)$ 可逆, 其逆矩阵为 $B = \left(\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n\right)$, 则

$$AB = A\left(\vec{b}_1, \vec{b}_2, \cdots, \vec{b}_n\right) = I = \left(\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\right).$$

其中

$$\vec{a}_i = \begin{pmatrix} a_{1i}\\ a_{2i}\\ \vdots\\ a_{n-1,i}\\ a_{ni} \end{pmatrix}, \ \vec{b}_i = \begin{pmatrix} b_{1i}\\ b_{2i}\\ \vdots\\ b_{n-1,i}\\ b_{ni} \end{pmatrix}, \ \vec{e}_i = \begin{pmatrix} 0\\ \vdots\\ 1\\ \vdots\\ 0 \end{pmatrix}.$$

$\vec{e}_i$ 的第 i 个元素为 1, 其它全为 0. 所以,

$$A\vec{b}_i = \vec{e}_i, \ \forall i = 1, 2, \cdots, n.$$

根据 Cramer 法则, 第 i 个线性方程的第 j 个未知数为

$$b_{ji}=(\vec{b}_i)_j=\frac{\left|(\vec{a}_1,\cdots,\vec{a}_{j-1},\vec{e}_i,\vec{a}_{j+1},\cdots,\vec{a}_n)\right|}{\left|(\vec{a}_1,\cdots,\vec{a}_{j-1},\vec{a}_j,\vec{a}_{j+1},\cdots,\vec{a}_n)\right|}=\frac{A_{ij}}{\left|A\right|}.$$

所以,

$$\vec{b}_i=\frac{1}{\left|A\right|}\begin{pmatrix}A_{i1}\\A_{i2}\\\vdots\\A_{in}\end{pmatrix},$$

即

$$A^{-1}=B=\frac{1}{\left|A\right|}\begin{pmatrix}A_{11}&A_{21}&\cdots&A_{n1}\\A_{12}&A_{22}&\cdots&A_{n2}\\\vdots&\vdots&\ddots&\vdots\\A_{1n}&A_{2n}&\cdots&A_{nn}\end{pmatrix}=\frac{1}{\left|A\right|}A^*.$$

6.4 初等变换与初等矩阵

要诀: 消元法的矩阵形式.

矩阵的初等变换:

- 交换矩阵的两行 (列), 记作 $r_i\leftrightarrow r_j$;
- 用一个非零常数去乘矩阵的某一行 (列) 的元, 记作 kr_i;
- 将矩阵某一行 (列) 的元乘数 k 后加到另一行 (列) 的对应元上去, 记作 r_i+kr_j.

初等矩阵可以精确表达初等变换, 将初等变换作用于单位矩阵得到初等矩阵.

矩阵初等变换也是基于消元法提出的, 正所谓招式一定要契合心法. 初等矩阵由初等变换对单位矩阵作用得到. 更关键的是, 一般矩阵的初等变换可以通过左 (右) 乘初等矩阵得到. “左行右列”.

初等矩阵:

- $P(i,j)$, $i\neq j$; $P(i,j)^{-1}=P(i,j)$
- $P(i(k))$, $k\neq 0$; $P(i(k))^{-1}=P(i(1/k))$
- $P(i,j(k))$, $i\neq j$; $P(i,j(k))^{-1}=P(i,j(-k))$

初等变换与初等矩阵的对应关系:

- $r_i\leftrightarrow r_j$: $P(i,j)$; $c_i\leftrightarrow c_j$: $Q(i,j)$
- kr_i: $P(i(k))$; kc_i: $Q(i(k))$
- r_i+kr_j: $P(i,j(k))$; c_i+kc_j: $Q(i,j(k))$

其实, $Q(i,j)=P(i,j)$, $Q(i(k))=P(i(k))$, $Q(i,j(k))=P(j,i(k))$.

阐述初等矩阵的行列式, 逆, 伴随.

这样我们就可以利用矩阵乘法重现消元法, 并且可以得到基于矩阵初等变换的求矩阵逆的计算方法.

$$\begin{cases}3x+y=4\\x+2y=3\end{cases}\iff\begin{pmatrix}3&1&4\\1&2&3\end{pmatrix}$$

$$\begin{cases}3x+y=4\\0x+\dfrac{5}{3}y=\dfrac{5}{3}\end{cases}\iff\begin{pmatrix}1&0\\-\dfrac{1}{3}&1\end{pmatrix}\begin{pmatrix}3&1&4\\1&2&3\end{pmatrix}=\begin{pmatrix}3&1&4\\0&\dfrac{5}{3}&\dfrac{5}{3}\end{pmatrix}$$

$$\begin{cases}3x+0y=3\\0x+\dfrac{5}{3}y=\dfrac{5}{3}\end{cases}\iff\begin{pmatrix}1&-\dfrac{3}{5}\\0&1\end{pmatrix}\begin{pmatrix}3&1&4\\0&\dfrac{5}{3}&\dfrac{5}{3}\end{pmatrix}=\begin{pmatrix}3&0&3\\0&\dfrac{5}{3}&\dfrac{5}{3}\end{pmatrix}$$

得到 $x=1$, $y=1$. 那么

$$\begin{pmatrix}1&0\\0&\dfrac{3}{5}\end{pmatrix}\begin{pmatrix}\dfrac{1}{3}&0\\0&1\end{pmatrix}\begin{pmatrix}1&-\dfrac{3}{5}\\0&1\end{pmatrix}\begin{pmatrix}1&0\\-\dfrac{1}{3}&1\end{pmatrix}\begin{pmatrix}3&1\\1&2\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix},$$

或者

$$\begin{pmatrix}\dfrac{1}{3}&-\dfrac{1}{5}\\0&\dfrac{3}{5}\end{pmatrix}\begin{pmatrix}1&0\\-\dfrac{1}{3}&1\end{pmatrix}\begin{pmatrix}3&1\\1&2\end{pmatrix}=\begin{pmatrix}1&0\\0&1\end{pmatrix},$$

或者

$$\begin{pmatrix}3&1\\1&2\end{pmatrix}=\begin{pmatrix}1&0\\\dfrac{1}{3}&1\end{pmatrix}\begin{pmatrix}3&1\\0&\dfrac{5}{3}\end{pmatrix}.$$

第一个式子是对消元法的总结, 第二个式子是矩阵逆的表达, 最后一个式子看作将矩阵分解为下三角矩阵和上三角矩阵的乘积.

$A=LU$. 利用初等行变换 (对应初等矩阵), 将矩阵 A 化成上三角矩阵 U (要求 A 的所有顺序主子式不为零). 若无乘数且无换行, 则可以轻易写出矩阵 L.

$$P(3,1(-3))P(2,1(-3))\begin{pmatrix}1&-1&1\\3&1&1\\3&2&0\end{pmatrix}=\begin{pmatrix}1&-1&1\\0&4&-2\\0&5&-3\end{pmatrix},$$

$$P(3,2(-5/4))\begin{pmatrix}1&-1&1\\0&4&-2\\0&5&-3\end{pmatrix}=\begin{pmatrix}1&-1&1\\0&4&-2\\0&0&-1/2\end{pmatrix}.$$

或者

$$P(3,2(-5/4))P(3,1(-3))P(2,1(-3))\begin{pmatrix}1&-1&1\\3&1&1\\3&2&0\end{pmatrix}=\begin{pmatrix}1&-1&1\\0&4&-2\\0&0&-1/2\end{pmatrix}.$$

$$\begin{pmatrix}1&-1&1\\3&1&1\\3&2&0\end{pmatrix}=P(2,1(3))P(3,1(3))P(3,2(5/4))\begin{pmatrix}1&-1&1\\0&4&-2\\0&0&-1/2\end{pmatrix}.$$

其中

$$P(2,1(3)) = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, P(3,1(3)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix},$$

and

$$P(3,2(5/4)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5/4 & 1 \end{pmatrix}.$$

而且

$$P(2,1(3))P(3,1(3))P(3,2(5/4)) = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 5/4 & 1 \end{pmatrix}.$$

矩阵的逆与初等矩阵: 矩阵 A 可逆的充要条件是存在一系列初等矩阵 $Q_1, Q_2, \cdots, Q_t$, 使得

$$A = Q_1 Q_2 \cdots Q_t.$$

另外, 可逆矩阵总可以通过一系列初等行变换化成单位矩阵. 即

$$P_s \cdots P_2 P_1 \begin{pmatrix} A & E \end{pmatrix} = \begin{pmatrix} E & P_s \cdots P_2 P_1 \end{pmatrix} = \begin{pmatrix} E & A^{-1} \end{pmatrix}.$$

6.5 矩阵的秩

变换 — 等价关系 — 分类 — 代表元 — 不变量.

秩, 秩序. 有统一的标准才有秩序, 有了标准就可以进行分类. 比如说, 大家去超市买东西, 排队结账. 大家都认同先来后到, 所以对插队这种行为嗤之以鼻. 换句话说, 只要有标准, 我们就能判断一件事情的好与坏. 社会价值观也是这样.

礼让行人. 红灯停绿灯行. 女士优先. 皆是准则, 有了大家所公认的准则就有了秩序, 有了分类.

要说明事物的特征, 往往从单方面不易说清楚, 可以根据形状、性质、成因、功用等属性的异同, 把事物分成若干类, 然后依照类别逐一加以说明. 这种说明技巧, 叫分类别.

分类别是将复杂的事物说清楚的重要方法, 有时事物的特征、本质需要分成几点或几个方面来说明, 也属分类别.

运用分类别方法要注意分类的标准. 一次分类只能用同一个标准, 以免产生重叠交叉的现象. 例如, 图书馆的藏书, 按国别分, 有中国的、外国的; 按时代分, 有古典的、现代的; 按性质分, 有科技的、文学的以及政治经济方面的等.

对于矩阵 $A, B \in \mathbb{R}^{n\times m}$, 如果 A 可以通过一系列初等变换变成 B, 称矩阵 A 和 B 是等价的. 用初等矩阵和矩阵乘法的语言来说, 如果矩阵 A 等价于 B, 则存在一系列初等矩阵 $P_1, P_2, \cdots, P_s \in \mathbb{R}^{n\times n}$ 和 $Q_1, Q_2, \cdots, Q_t \in \mathbb{R}^{m\times m}$, 使得

$$A = P_s \cdots P_2 P_1 B Q_1 Q_2 \cdots Q_t.$$

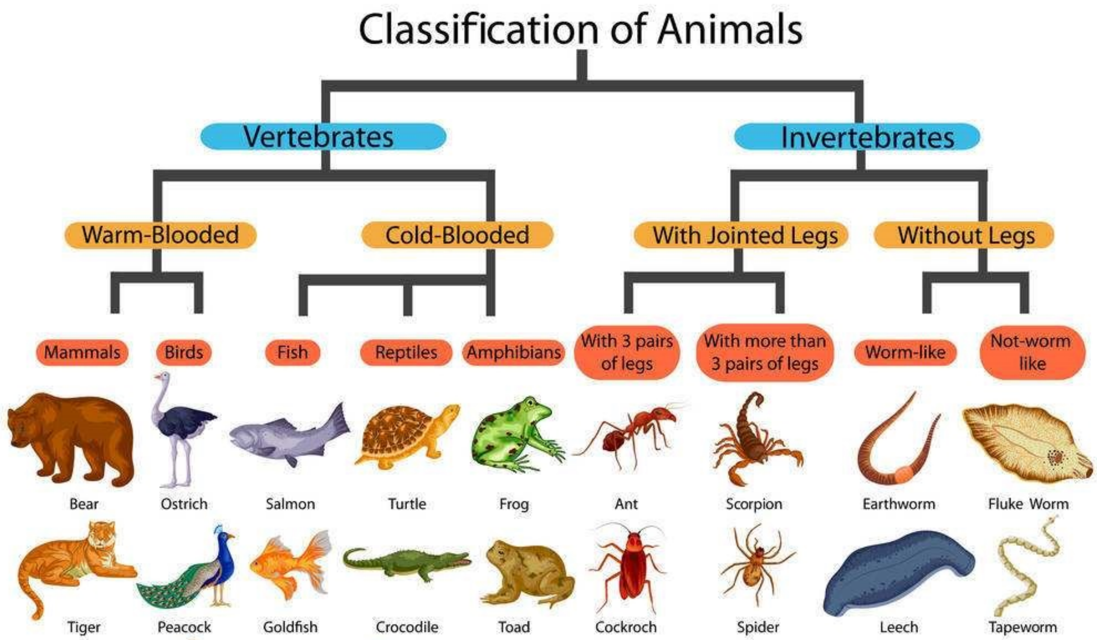


Figure 6.2: Classification of Animals

Figure 6.3: Garbage Classification

对于任一个矩阵 $A\in\mathbb{R}^{n\times m}$, 它都与一形式为

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

的矩阵等价, 称 D_r 为 A 的标准形, 下标 r 表示主对角线上 1 的个数.

基于上述等价关系, 将所有 n 行 m 列的矩阵分为 $l=\min\{n,m\}+1$ 类. 即

$$\mathbb{R}^{n\times m}=S_0\cup S_1\cup\cdots\cup S_{l-1},$$

其中 S_i 是与矩阵 D_i 等价的所有矩阵的集合. D_i 称为 S_i 的代表元.

对任一 $A\in\mathbb{R}^{n\times m}$, 则 A 会属于某个集合 S_i. 我们将 i 称作矩阵 A 的秩, 记为 $\operatorname{rank}(A)$. 显然 $\operatorname{rank}(A)\le\min\{n,m\}$.

Example 6.1. 对于 4 行 5 列的矩阵, 基于初等变换可以分成五类:

$$D_0=\begin{pmatrix}0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix},\ D_1=\begin{pmatrix}1&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix},\ D_2=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&0&0&0\\0&0&0&0&0\end{pmatrix},$$
$$D_3=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0\end{pmatrix},\ D_4=\begin{pmatrix}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\end{pmatrix}.$$

如果矩阵 A 的秩为 r, 则 A 有一个 r 级子式不为零且所有 $r+1$ 级子式都为零. 反之亦然. (矩阵的初等变换并不改变对应 k 级子式的零性)

矩阵的秩与映射的值域.

Remark 6.1. 相同行列数的矩阵基于初等变换建立等价关系, 并将矩阵分类, 同时选择代表元. 每一类不等于其它类的特征量就是矩阵的秩.

6.6 矩阵的分块

矩阵分块源于矩阵的加法. 以映射的观点就是将一个映射分成多个映射的和.

矩阵按列分块作用在向量上, 可以看作是矩阵列向量的线性组合.

$$M=\begin{pmatrix}A_{n\times s}&B_{n\times t}\\C_{m\times s}&D_{m\times t}\end{pmatrix},\ N=\begin{pmatrix}E_{s\times p}&F_{s\times q}\\G_{t\times p}&H_{t\times q}\end{pmatrix},$$

则

$$MN=\begin{pmatrix}AE+BG&AF+BH\\CE+DG&CF+DH\end{pmatrix}.$$

设 $A, B, C \in \mathbb{R}^{n\times n}$, $M = \begin{pmatrix} A & O \\ C & B \end{pmatrix}$, 则 $\det(M) = \det(A)\det(B)$ (可以对 n 进行数学归纳法证明). 若 A 和 B 都可逆, 则 M 可逆且 $M^{-1} = \begin{pmatrix} A^{-1} & O \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix}$.

$R_1 \leftrightarrow R_2$:

$$\begin{pmatrix} O & E_n \\ E_n & O \end{pmatrix} = P(n, n+n)\cdots P(2, n+2)P(1, n+1) \begin{pmatrix} E_n & O \\ O & E_n \end{pmatrix}.$$

$k \cdot R_1$:

$$\begin{pmatrix} kE_n & O \\ O & E_n \end{pmatrix} = P(n(k))\cdots P(2(k))P(1(k)) \begin{pmatrix} E_n & O \\ O & E_n \end{pmatrix}.$$

$A \cdot R_1 + R_2$:

$$\begin{aligned}\begin{pmatrix} E_n & O \\ A & E_n \end{pmatrix} = &\prod_{i=1}^{n} P(n+n, i(a_{ni}))\cdots \prod_{i=1}^{n} P(n+2, i(a_{2i})) \\ &\prod_{i=1}^{n} P(n+1, i(a_{1i})) \begin{pmatrix} E_n & O \\ O & E_n \end{pmatrix}.\end{aligned}$$

可以得到块矩阵的块初等变换, 对应行列式的变化. 设 $A, B, C, D, F \in \mathbb{R}^{n\times n}$, $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, 则

$$\begin{aligned}
\det(\begin{pmatrix} O & E_n \\ E_n & O \end{pmatrix} M) &= (-1)^n \det(M), \\
\det(\begin{pmatrix} kE_n & O \\ O & E_n \end{pmatrix} M) &= k^n \det(M), \\
\det(\begin{pmatrix} E_n & O \\ F & E_n \end{pmatrix} M) &= \det(M).
\end{aligned}$$

Example 6.2. 设 $M = \begin{pmatrix} E_n & E_n \\ E_n & E_n \end{pmatrix}$, 则 $rank(M) = n$.

6.7 关于矩阵的定理

Theorem 6.1. *若矩阵可逆, 则其逆矩阵是唯一的.*

Proof. 知 $AB = BA = E$, 若存在 C 使得 $AC = CA = E$. 那么 $C = B$. □

Theorem 6.2. *设矩阵 $A \in \mathbb{R}^{n\times n}$, 则 A 可逆的充要条件是 $det(A) \neq 0$. 当 A 可逆时,*

$$A^{-1} = \frac{1}{det(A)} A^*.$$

Proof. 若可逆, $AB = E$, 以及逆矩阵的唯一性, 利用 Cramer 法则, 方形方程组存在唯一解的充要条件为 $\det(A) \neq 0$, 并给出逆矩阵的公式.

若 $\det(A) \neq 0$, 则由 Cramer 法则, 线性方程组 $AX = E$ 存在唯一解, 且 $X = \dfrac{1}{\det(A)}A^*$. 再验证 $XA = E$, 则 A 可逆, 且 X 为 A 的逆矩阵. □

Theorem 6.3. 设 $A, B, C \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, $C \in \mathbb{R}^{m\times n}$, $M = \begin{pmatrix} A & O \\ C & B \end{pmatrix}$, 则

$$det(M) = det(A)det(B).$$

若 A 和 B 都可逆, 则 M 可逆且

$$M^{-1} = \begin{pmatrix} A^{-1} & O \\ -B^{-1}CA^{-1} & B^{-1} \end{pmatrix}.$$

Proof. 第一个等式可以对 n 进行数学归纳法证明, 也可以根据行列式的定义. 第二个等式直接利用定义验证. □

Theorem 6.4. 设矩阵 $A, B \in \mathbb{R}^{n\times n}$, 则

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

Proof. 设 $M = \begin{pmatrix} A & O \\ -E & B \end{pmatrix}$, 则

$$\begin{pmatrix} E & A \\ O & E \end{pmatrix} M = \begin{pmatrix} E & A \\ O & E \end{pmatrix} \begin{pmatrix} A & O \\ -E & B \end{pmatrix} = \begin{pmatrix} O & AB \\ -E & B \end{pmatrix}.$$

容易得到, $\det(M) = \det(A)\det(B)$, 并且上述操作不改变矩阵 M 对应的行列式. 然而 $\begin{vmatrix} O & AB \\ -E & B \end{vmatrix} = (-1)^n\det(AB)\det(-E)$. □

Theorem 6.5. 设矩阵 $A \in \mathbb{R}^{n\times n}$, 存在矩阵 $B \in \mathbb{R}^{n\times n}$ 使得

$$AB = E,$$

则 A, B 皆可逆且 A 和 B 互为逆矩阵.

Proof. $1 = \det(E) = \det(AB) = \det(A)\det(B)$, 即 $\det(A)$ 和 $\det(B)$ 都不为零. 所以 A 和 B 都可逆. 然而 $B = EB = BE = BAB \Leftrightarrow (BA - E)B = O \Leftrightarrow B^T(BA - E)^T = O$. 由 $\det(B) \neq 0$ 和 Cramer 法则, $(BA - E)^T = O \Leftrightarrow BA = E$. 即 A 和 B 互为逆矩阵. □

Theorem 6.6. 初等矩阵都是可逆矩阵, 其逆矩阵也是初等矩阵.

Proof. $P(i,j)^{-1} = P(i,j)$, $P(i(k))^{-1} = P(i(1/k))$ $(k \neq 0)$, $P(i,j(k))^{-1} = P(i,j(-k))$ $(i \neq j)$. $Q(i,j)^{-1} = Q(i,j)$, $Q(i(k))^{-1} = Q(i(1/k))$ $(k \neq 0)$, $Q(i,j(k))^{-1} = Q(i,j(-k))$ $(i \neq j)$. □

Theorem 6.7. 设矩阵 $A \in \mathbb{R}^{n\times m}$, 则 A 必等价于一形如

$$D_r = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} E_r & O_{r\times(m-r)} \\ O_{(n-r)\times r} & O_{(n-r)\times(m-r)} \end{pmatrix}.$$

的矩阵.

Proof. 不妨设 $a_{11} \neq 0$, 则

$$A \sim \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & O_{1\times(m-1)} \\ O_{(n-1)\times 1} & A_1 \end{pmatrix},$$

其中 $A_1 \in \mathbb{R}^{(n-1)\times(m-1)}$. 若 $a_{11} = 0$, 则不妨设 a_{i1} 或 a_{1j} 至少有一个不为 0, 将之换到第一行第一列. 否则第一行和第一列全为 0. □

Theorem 6.8. 设矩阵 $A \in \mathbb{R}^{n\times n}$ 可逆的充要条件是存在一系列初等矩阵 $R_1, \cdots, R_l$, 使得

$$A = R_1 \cdots R_l.$$

Proof. 若 A 可逆, 下面证明 $A \sim E$. 首先, 存在一系列初等矩阵 $P_1, \cdots, P_s$ 和 $Q_1, \cdots, Q_t$, 使得 $P_s \cdots P_1 A Q_1 \cdots Q_t = D_r$. 由于 $\det(A) \neq 0$, 则

$$0 \neq \det(P_s)\cdots\det(P_1)\det(A)\det(Q_1)\cdots\det(Q_t) = \det(D_r).$$

所以, $r = n$. 那么 $A = (P_s \cdots P_1)^{-1} E (Q_1 \cdots Q_t)^{-1}$. 即

$$A = P_1^{-1} \cdots P_s^{-1} Q_t^{-1} \cdots Q_1^{-1}.$$

反之, 直接验证. □

Theorem 6.9. 设矩阵 $A, B \in \mathbb{R}^{n\times m}$, 则下述论述互为充要条件:

- $A \sim B$;
- 存在可逆矩阵 $P \in \mathbb{R}^{n\times n}$ 和 $Q \in \mathbb{R}^{m\times m}$, 使得 $A = PBQ$;
- $rank(A) = rank(B)$.

Proof. 直接利用定义验证. 关键是初等变换与初等矩阵的对应关系, $A \sim D_r$ 和可逆矩阵可以写成一系列初等矩阵的乘积. □

Theorem 6.10. 设矩阵 $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{s\times t}$, $C \in \mathbb{R}^{n\times t}$, $M = \begin{pmatrix} A & O \\ C & B \end{pmatrix}$, 则

$$rank(M) \geq rank(A) + rank(B).$$

若 C 是零矩阵, 则等号成立.

Proof. $\begin{pmatrix} P & O \\ O & \tilde{P} \end{pmatrix}\begin{pmatrix} A & O \\ C & B \end{pmatrix}\begin{pmatrix} Q & O \\ O & \tilde{Q} \end{pmatrix} = \begin{pmatrix} PAQ & O \\ \tilde{P}CQ & \tilde{P}B\tilde{Q} \end{pmatrix} = \begin{pmatrix} D_p & O \\ \tilde{P}DQ & \tilde{D}_q \end{pmatrix}$.

$$\begin{pmatrix} D_p & O \\ \tilde{P}DQ & \tilde{D}_q \end{pmatrix} = \begin{pmatrix} E_p & O & O & O \\ O & O & O & O \\ U & V & E_q & O \\ W & Z & O & O \end{pmatrix} \to \begin{pmatrix} E_p & O & O & O \\ O & E_q & O & O \\ O & O & Z & O \\ O & O & O & O \end{pmatrix}.$$

□

Theorem 6.11. 设矩阵 $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{n\times l}$, $M = \begin{pmatrix} A & B \end{pmatrix}$, 则

$$\max\{rank(A), rank(B)\} \leq rank(M) \leq rank(A) + rank(B).$$

Proof. $P\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} Q & O \\ O & E \end{pmatrix} = \begin{pmatrix} PAQ & PB \end{pmatrix} = \begin{pmatrix} D_r & PB \end{pmatrix}$ 且

$$\begin{pmatrix} D_r & PB \end{pmatrix} = \begin{pmatrix} E_r & O & G \\ O & O & H \end{pmatrix} \to \begin{pmatrix} E_r & O & O \\ O & O & H \end{pmatrix} \to \begin{pmatrix} E_r & O & O \\ O & H & O \end{pmatrix}.$$

那么 $\mathrm{rank}(M) = \mathrm{rank}(A) + \mathrm{rank}(H) \geq \mathrm{rank}(A)$. 同理, $\mathrm{rank}(M) \geq r = \mathrm{rank}(B)$. 然而 $\mathrm{rank}(H) \leq \mathrm{rank}(PB) = \mathrm{rank}(B)$. □

Theorem 6.12. 设矩阵 $A, B \in \mathbb{R}^{n\times n}$, 则

$$\begin{aligned}
&rank(A+B) \leq rank(A) + rank(B),\\
&rank(kA) = rank(A) \ \ if\ k \neq 0,\\
&rank(AB) \leq \min\{rank(A), rank(B)\},\\
&rank(AB) \geq rank(A) + rank(B) - n,\\
&rank(A^T) = rank(A),\\
&rank(A^{-1}) = rank(A) = n,\\
&rank(A^*) = \begin{cases} n, & if\ rank(A) = n,\\ 1, & if\ rank(A) = n-1,\\ 0, & otherwise. \end{cases}
\end{aligned}$$

Proof.

- $\begin{pmatrix} A & B \end{pmatrix} \sim \begin{pmatrix} A & A+B \end{pmatrix}$.
- $(kE)A = A$.
- $PAQQ^{-1}B = D_rQ^{-1}B = \begin{pmatrix} E_r & O \\ O & O \end{pmatrix} Q^{-1}B$.

- $\begin{pmatrix} E & A \\ O & E \end{pmatrix}\begin{pmatrix} A & O \\ -E & B \end{pmatrix} = \begin{pmatrix} O & AB \\ -E & B \end{pmatrix} \to \begin{pmatrix} O & AB \\ -E & O \end{pmatrix} \to \begin{pmatrix} AB & O \\ O & -E \end{pmatrix}$. 那么 $\mathrm{rank}(A) + \mathrm{rank}(B) \le$ 等式左边矩阵的秩 = 等式右边矩阵的秩 $= \mathrm{rank}(AB) + \mathrm{rank}(E)$.
- $PAQ = D_r$, $Q^TA^TP^T = D_r^T = \tilde{D}_r$.
- A可逆 $\iff A \sim E \iff \mathrm{rank}(A) = n = \mathrm{rank}(A^{-1})$.

□

Theorem 6.13. *设矩阵 $A \in \mathbb{R}^{n\times m}$, 则 $rank(A) = r$ 的充要条件是存在 A 的 r 阶子式不等于 0, 并且所有 $r+1, \cdots, \min\{n, m\}$ 阶子式都为零. 这里 A 的 k 阶子式是从 A 中选出某 k 行某 k 列得到的 k 阶行列式, $k = 1, 2, \cdots, \min\{n, m\}$.*

Proof. 若 $\mathrm{rank}(A) = r$, 则不妨设 $(A)_{11} \neq 0$, 则

$$A \sim \begin{pmatrix} 1 & O_{1\times(m-1)} \\ O_{(n-1)\times 1} & A_1 \end{pmatrix},$$

再对 A_1 进行同样的讨论, 这样将 A 化为 D_r. 那么 A 前 r 行前 r 列对应的 r 阶子式必不为零. 因为对于 D_r, 其相同行列的 r 阶子式不为零. 另外 D_r 的更高阶的子式都为 0, 则 A 也是.

反之, 将不为 0 的 A 的 r 阶子式对应的行列通过行列互换, 换到 A 的前 r 行前 r 列, 那么此时矩阵的前 r 行前 r 列对应的矩阵与 E_r 等价. 所以 $\mathrm{rank}(A) \ge r$. 又由其它高阶子式都为 0, 得到 $\mathrm{rank}(A) = r$.

$$\begin{pmatrix} P & O \\ O & E \end{pmatrix}\begin{pmatrix} L & M \\ N & S \end{pmatrix}\begin{pmatrix} Q & O \\ O & E \end{pmatrix} = \begin{pmatrix} PLQ & PM \\ NQ & S \end{pmatrix} = \begin{pmatrix} E_r & PM \\ NQ & S \end{pmatrix}$$
$$\to \begin{pmatrix} E_r & PM \\ O & -NQPM+S \end{pmatrix} \to \begin{pmatrix} E_r & O \\ O & -NQPM+S \end{pmatrix}.$$

□

Remark 6.2. *矩阵的运算: 加法, 数乘, 乘法, 逆, 转置, 行列式, 秩, 分块. 单个运算本身或两两组合就会对应定理和性质. 本节定理主要是研究的是逆, 行列式, 秩, 分块. 有关秩的定理, 就会与初等变换和初等矩阵相关联.*

(HW) 思考块矩阵求逆和块矩阵求行列式的计算公式.

(HW) 思考矩阵相加的求行列式的计算公式.

(HW) 思考关于矩阵秩的不等式, 何时等号成立, 是否有充要条件.

(HW) 思考如果将矩阵求伴随矩阵看作一种运算, 那么它与上述八种运算之间的联系.

Remark 6.3. *关于秩, 总结如下:*

设矩阵

$$A = (a_{ij})_{n\times m} = \begin{pmatrix} \vec{\alpha}_1 & \cdots & \vec{\alpha}_m \end{pmatrix} = \begin{pmatrix} \vec{\beta}_1 \\ \vdots \\ \vec{\beta}_n \end{pmatrix},$$

- 向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_m\}$ 的极大线性无关组中向量个数为 p;

- 向量组 $\{\vec{\beta}_1, \cdots, \vec{\beta}_n\}$ 的极大线性无关组中向量个数为 q;
- A 通过初等变换化为 D_r;
- A 存在一个 s 阶子式不为零, 所有 $s+1$ 阶子式全为零.

那么,

$$p = q = r = s.$$

6.8　再论消元法

现在我们已经了解了矩阵的基本运算, 矩阵的初等变换和矩阵的秩. 接下来我们将用这些矩阵的理论, 重新认识消元法. 首先, 我们回顾对于三个未知数, 三个线性方程的消元法求解过程. 线性方程组

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

对于上述线性方程组进行消元, 就是对增广矩阵 $\bar{A}$ 进行初等行变换. 那么, 通过初等行变换和交换线性方程组系数矩阵的列必能将增广矩阵 $\bar{A}$ 化成下述三个矩阵之一:

$$\begin{pmatrix} 1 & 0 & 0 & d_1 \\ 0 & 1 & 0 & d_2 \\ 0 & 0 & 1 & d_3 \end{pmatrix}, \begin{pmatrix} 1 & 0 & c_{13} & d_1 \\ 0 & 1 & c_{23} & d_2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & c_{13} & d_1 \\ 0 & 1 & c_{23} & d_2 \\ 0 & 0 & 0 & d_3 \end{pmatrix}.$$

其中, 第三个矩阵中 $d_3 \neq 0$. 换句话说, 增广矩阵 $\bar{A}$ 必与上述某个矩阵相抵, 即两者的秩相等. 容易得到, 三个矩阵的秩分别为 $3, 2, 3$. 它们分别对应消元法的第一, 二, 三种情况. 即, 线性方程组, 存在唯一解, 存在多解和无解.

n 个线性方程, m 个未知数的线性方程组. 矩阵 A 是线性方程组的系数矩阵, $\bar{A}$ 是对应的增广矩阵. 增广矩阵通过初等行变换和列列交换 (只交换线性方程组系数矩阵的列), 可以化为 $\bar{D}$.

$$\bar{A} \sim \bar{D} = \begin{pmatrix} 1 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1m} & d_1 \\ 0 & 1 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2m} & d_2 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rm} & d_r \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

此时,

$$A \sim D = \begin{pmatrix} 1 & 0 & \cdots & 0 & c_{1,r+1} & \cdots & c_{1m} \\ 0 & 1 & \cdots & 0 & c_{2,r+1} & \cdots & c_{2m} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & c_{r,r+1} & \cdots & c_{rm} \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

容易看出, $\operatorname{rank}(A) = r$. 更重要的是 $\operatorname{rank}(A) = \operatorname{rank}(\bar{A})$ 的充要条件是 $d_{r+1} = 0$.

- 若 $\operatorname{rank}(\bar{A}) = \operatorname{rank}(A) = m$, 则 $r = m$, 即线性方程组有唯一解.
- 若 $\operatorname{rank}(\bar{A}) = \operatorname{rank}(A) = r < m$, 则 $d_{r+1} = 0$, 即线性方程组有无穷多个解.
- 若 $\operatorname{rank}(\bar{A}) \neq \operatorname{rank}(A)$, 则 $d_{r+1} \neq 0$. 此时, $\operatorname{rank}(\bar{A}) = \operatorname{rank}(A) + 1 = r + 1$, 即线性方程组无解.

反之同样成立.

Remark 6.4. *请阅读 [3] 第四章 – 矩阵, 第 3, 5, 6, 7 节.*

线性方程组的初等变换 → 矩阵的初等变换 → 初等矩阵 →"左行右列" → 等价关系 → 矩阵的秩 → 线性方程组消元法的再表达.

副产品: 矩阵各种运算交换顺序所对应的定理和性质.

Remark 6.5. 视频是在时代的波涛中航行的思想之船, 它小心翼翼地把珍贵的货物运送给一代又一代.

- 厦门大学 – 高等代数 – 第一章 – 第 2 节 – 矩阵和运算

  http://www.icourses.cn/sCourse/course_3077.html

- 厦门大学 – 高等代数 – 第一章 – 第 6 节 – 可逆矩阵

  http://www.icourses.cn/sCourse/course_3077.html

- 麻省理工学院 – 线性代数 – 第 3 集 – 乘法和逆矩阵

  http://open.163.com/movie/2010/11/H/0/M6V0BQC4M_M6V29FCH0.html

7 变换与特征量 – 分类与整理的艺术

等价简化, 分而治之.

- **等价是什么意思?**
- **如何判定是简化了, 还是繁难了?**
- **为何能分, 如何分?**
- **治谁, 怎么治?**

首先设集合 S 和变换族 $\mathcal{T}=\{P:S\to S\}$, 即 $\mathcal{T}$ 是从 S 到 S 的映射的集合 (称定义域和值域相同的映射为变换). $\mathcal{T}$ 满足:

- 恒等映射 $I\in\mathcal{T}$;
- $\mathcal{T}$ 中元素基于映射复合运算是封闭的;
- $\forall P\in\mathcal{T}$, 存在 $Q\in\mathcal{T}$, 使得 $PQ=QP=I$ (乘法表示映射复合), 即两个映射复合后变成恒等映射. 称 P 的逆是 Q, 记 Q 为 P^{-1}. 反之亦然.

接下来基于变换族 $\mathcal{T}$ 定义等价关系: 对于 $a,b\in S$, 若存在 $P,Q\in\mathcal{T}$, 使得 $P(a)=b$ 且 $Q(b)=a$, 则称 a 和 b 是等价的, 记为 $a\sim b$.

下面证明这样定义的关系确实是等价关系, 即要证明所定义的关系具有反身性, 对称性, 传递性. 对于 $a,b,c\in S$, 首先, $a\sim a$, 这是由于 $\mathcal{T}$ 中包含恒等映射; 其次, 若 $a\sim b$, 则 $b\sim a$, 这由该关系的定义以及 $\mathcal{T}$ 满足的性质得到; 最后, 若 $a\sim b$ 且 $b\sim c$, 则 $a\sim c$, 这由该关系的定义以及 $\mathcal{T}$ 满足的性质得到.

那么基于等价关系将等价的元素分别放在一起, 则设将集合 S 分成 n 类, 记为 $S_1,S_2,\cdots,S_n$. 它们都是 S 的子集, 并且

$$S_i\cap S_j=\emptyset,\ \text{if } i\neq j \text{ and } S=S_1\cup S_2\cup\cdots\cup S_n.$$

取 $s_i\in S_i$, 称 s_i 为 S_i 的一个代表元, $i=1,2,\cdots,n$. 一般地, 所取得 s_i 是可以明显地明确表达集合 S_i 中所有元素共同特征的元素. 显然, $a,b\in S_i\Leftrightarrow a\sim b$.

最后我们基于对 S 的分类, 在每个类中寻找一种共同特征, 而且这个特征对每个类都是不等的. 也就是说, 通过这个特征我们可以判断任意 $s\in S$ 属于哪一类, 或者说属于哪个 S_i. 我们称这样的特征为特征量或完全不变量. 记 $\forall s\in S$ 的特征量为 $\mathrm{ch}(s)$. 那么,

- 若 $a,b\in S_i$, 则 $\mathrm{ch}(a)=\mathrm{ch}(b)$.
- 若 $a\in S_i$ 且 $a\in S_j$, $i\neq j$, 则 $\mathrm{ch}(a)\neq\mathrm{ch}(b)$.

即每个类都有于其它类不等的特征量. 换句话说,

- 若存在 $P\in\mathcal{T}$, 使得 $b=P(a)$, 则 $\mathrm{ch}(a)=\mathrm{ch}(b)$.
- 若不存在 $P\in\mathcal{T}$, 使得 $b=P(a)$, 则 $\mathrm{ch}(a)\neq\mathrm{ch}(b)$.

这就是将特征量称为完全不变量的原因 $(\mathrm{ch}(a)=\mathrm{ch}(P(a))$, $\forall P\in\mathcal{T})$.

Remark 7.1. 将所研究的事物用集合 S 描述, 然后找到从 S 到 S 的变换族. 这样就可以定义等价关系以及对 S 分类, 进而探索出特征量 *(完全不变量)*.

Example 7.1. 以上述 "变换与特征量" 的理论重新阐述高斯消元法和矩阵理论求解线性方程组的过程.

8 线性空间 – 万丈高楼平地起

线性空间理论可以完全描述和解决线性方程组求解的问题.

- **什么是线性空间, 为何称之为线性空间?**
- **线性空间理论有哪些?**
- **如何描述, 用到哪些线性空间理论?**
- **解决的思路是什么?**
- **具体步骤是什么?**
- **为什么是这样的步骤?**
- **为何要用线性空间来研究线性方程组?**
- **与消元法是否有所关联?**
- **与行列式是否有所关联?**
- **与矩阵是否有所关联?**

内家功夫, 第三招, 消元法的全新解读.

要诀: 集合与映射, 线性空间与矩阵.

我们之前学习了主要学习了两大 “招式”, 比如行列式和矩阵. 这两个招式之间也有很多联系. 这两个招式使得我们对付线性方程组有了办法. 换句话说, 我们遇到线性方程组求解的问题时, 已经可以按照既定的 “招式” 将其制服了.

但数学讲的是 “集合与映射”. 如果线性代数只停留在行列式和矩阵, 那就只能成为 “胸口碎大石” 般的技艺. 线性空间的提出使得线性代数走向了 “集合与映射” 的数学框架下, 使线性代数一跃而成流芳百世的武学经典.

线性空间也许是你接触的第一个非常抽象的数学概念. 线性空间的习得是你的线性代数从 “招式” 到 “ 心法” 的一次巨大提升. “招式” 好练, “心法” 难懂. “心法” 是对 “招式” 的抽象和升华.

8.1 概念引入

数学的美在于为不同的事物起了相同的名字. 线性空间就是数学美的体现. 空间并不单单是一个集合, 而且里面的元素需要有一定的关联. 这种结构充分体现了线性方程组的列视角.

“1128”:

$$\text{线性空间} = \text{集合} + \text{数域} + \text{加法和数乘} + \text{八条法则}.$$

什么是集合?

最简单的数学对象, 将我们所关心的事物放在一起.

什么是数域?

首先是数的集合, 而且要定义数与数之间的加减乘除, 并且所得的数还要属于这个集合.

什么是加法?

这里的加法是抽象的, 它可能并不是数与数之间的, 像 $3+2=5$ 这样的加法. 一个简单的例子就是二维空间中向量的加法, 两个向量相加满足平行四边形法则, 得到另外一个向量.

什么是数乘?

故名思议, 首先有数, 然后用之乘. 这里是用数域中的数乘以向量空间中的向量. 数是比较具体, 或是实数或是复数或是有理数. 然后由于向量的抽象性使得乘变得很抽象. 比如一个二维向量和 3 相乘.

什么是八条法则?

规定加法和数乘满足的运算规则, 并使二者协调.

什么是向量?

向量其实是一个非常抽象的概念. 向量可以是我们中学理解的有方向和长度的线段, 也可以是二维空间中的坐标, 还可以是函数. 当我们想研究一类具体事物的线性性质时, 我们就将这个事物称作向量, 而这个事物的全体称作向量空间或线性空间.

什么是线性运算?

加法和数乘统称为线性运算.

(HW) 阐述 $\vec{a}+\vec{a}=2\vec{a}$ 中所有符号和运算的意义, 其中 $\vec{a}$ 是线性空间中的元素.

什么是线性结构?

以 $\mathbb{R}^2$ 为例, 每个向量都放在平面上. 以某一个向量进行数乘得到另外一个向量. 某一个向量加另一个向量, 利用平行四边形法则得到另外一个向量, 使得三个向量有了联系. 比如集合中的元素为 -1 和 1, 定义普通的乘法为加法, 则可绘出其结构图.

(HW) 验证以 n 行 m 列的矩阵为集合, 数域, 以矩阵的加法和数乘为加法和数乘, 构成线性空间.

8.2 基本概念和运算

线性组合, 线性表出, 向量组等价.

线性相关, 线性无关.

极大线性无关组, 向量组的秩.

上述概念在任何一本线性代数的教科书中都可以找到它们的精确定义. 这里我们需要理解的是, 这些定义之间的联系.

首先, 线性空间是一个集合, 但它不单单是一个集合. 线性空间中的向量之间可以进行两种运算, 加法和数乘. 那么, 线性组合和线性表出的定义就出现了. 例如, 若成立

$$\vec{v}=k_1\vec{v}_1+\cdots+k_m\vec{v}_m,$$

则称向量 $\vec{v}$ 是向量组 $\{\vec{v}_1,\cdots,\vec{v}_m\}$ 的一个线性组合, 或者称向量组 $\{\vec{v}_1,\cdots,\vec{v}_m\}$ 可以线性表出向量 $\vec{v}$.

然后, 要明确, 线性相关与线性无关是向量组的属性. 例如, 令上式中的 $\vec{v}$ 等于 $\vec{0}$. 若向量组 $\{\vec{v}_1,\cdots,\vec{v}_m\}$ 可以唯一线性表出零向量 $\vec{0}$, 则称向量组 $\{\vec{v}_1,\cdots,\vec{v}_m\}$ 线性无关, 否则称之为线性相关. 这里, 唯一性指的是表出系数 $k_1,\cdots,k_m$ 是否有非零数. 因为, $k_1=\cdots=k_m=0$ 必然使得

$$\vec{0}=k_1\vec{v}_1+\cdots+k_m\vec{v}_m$$

成立.

最后, 要明确, 极大线性无关组 S_1 是向量组 S_0 的子集, 即 $S_1 \subset S_0$. 极大线性无关组 S_1 要满足两条性质, 其一是线性无关, 其二是真包含 S_1 的所有其它 S_0 的子集都是线性相关. 极大线性无关组这个名字中的极大就体现在第二点中. 那么, 向量组的秩, 这个数, 就是向量组的极大线性无关组中向量的个数. 一般地, 向量组的极大线性无关组是不唯一的. 所以, 需要一个定理去说明, 向量组的秩, 这个定义的合理性.

Definition 8.1. 设向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_n\}$ 的一部分向量构成一个向量组 $\{\vec{\alpha}_{i_1}, \cdots, \vec{\alpha}_{i_s}\}$, 即

$$\{\vec{\alpha}_{i_1}, \cdots, \vec{\alpha}_{i_s}\} \subset \{\vec{\alpha}_1, \cdots, \vec{\alpha}_n\}.$$

若向量组 $\{\vec{\alpha}_{i_1}, \cdots, \vec{\alpha}_{i_s}\}$ 线性无关且取任意 $\vec{\alpha}_l \in \{\vec{\alpha}_1, \cdots, \vec{\alpha}_n\}$, 向量组 $\{\vec{\alpha}_{i_1}, \cdots, \vec{\alpha}_{i_s}, \vec{\alpha}_l\}$ 线性相关, 则称向量组 $\{\vec{\alpha}_{i_1}, \cdots, \vec{\alpha}_{i_s}\}$ 为向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_n\}$ 的**极大线性无关组**.

极大线性无关组 $\{\vec{\alpha}_{i_1}, \cdots, \vec{\alpha}_{i_s}\}$ 中向量的个数称为向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_n\}$ 的**秩**.

维数, 基和坐标.

线性无关向量组 $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n\} \subset V$, $\forall \vec{b} \in V$ 可由其线性表出, 则称之为 V 的基, 此时 n 称为 V 的维数. 那么 $\vec{x} \in \mathbb{R}^n$ 满足

$$\begin{pmatrix} \vec{\varepsilon}_1 & \cdots & \vec{\varepsilon}_n \end{pmatrix} \vec{x} = \vec{b},$$

称之为 $\vec{b} \in V$ 在基 $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n\}$ 下的坐标. 线性表出的唯一性是由于基向量组的线性无关性.

基变换

$$\begin{pmatrix} \vec{\eta}_1 & \cdots & \vec{\eta}_n \end{pmatrix} = \begin{pmatrix} \vec{\varepsilon}_1 & \cdots & \vec{\varepsilon}_n \end{pmatrix} A.$$

其中, $A = (a_{ij})_{n\times n} \in \mathbb{R}^{n\times n}$.

坐标变换

$$A\vec{y} = \vec{x}.$$

其中

$$\begin{pmatrix} \vec{\eta}_1 & \cdots & \vec{\eta}_n \end{pmatrix} \vec{y} = \vec{b}.$$

(HW) 过渡矩阵的可逆性.

(HW) 讨论 $\mathbb{R}^n$ 上过渡矩阵的求法.

8.3 向量组与矩阵

我们将矩阵看作一组列向量的排列, 则矩阵乘以向量看作这组向量的线性组合, 组合系数就是这个向量. 将矩阵乘以矩阵看作是多个线性组合形成向量组, 并又构成一个矩阵. (严格地说, 将向量在一组基下对应的坐标作为列向量. 或者将 V 直接设为 $\mathbb{R}^n$)

记由向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 和 $\{\vec{\beta}_1, \cdots, \vec{\beta}_s\} \subset V$ 形成的矩阵为

$$A = \begin{pmatrix} \vec{\alpha}_1 & \cdots & \vec{\alpha}_r \end{pmatrix},\ B = \begin{pmatrix} \vec{\beta}_1 & \cdots & \vec{\beta}_s \end{pmatrix}.$$

(严格地说, 是向量组基于基向量的坐标形成的 $\mathbb{R}^n$ 上的向量组排列形成矩阵.) 下面所说的秩, 即可以理解为矩阵的秩, 也可以理解为列向量组的秩.

如何判断一个向量 $\vec{\alpha} \in V$ 是否可以用向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 线性表出?

若可线性表出, 则线性方程组

$$A\vec{x} = \vec{\alpha}$$

有解 $\vec{x}$. 反之亦然. 所以, 可线性表出的充要条件是 $\text{rank}(A) = \text{rank}(A, \vec{\alpha})$.

如何判断向量组 $\{\vec{\beta}_1, \cdots, \vec{\beta}_s\} \subset V$ 可以经向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 线性表出?

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} E & Q \\ O & -E \end{pmatrix} = \begin{pmatrix} A & AQ - B \end{pmatrix} = \begin{pmatrix} A & O \end{pmatrix}$$

若可以线性表出, 则线性方程组 $AX = B$ 有解. 所以, 可以线性表出的充要条件是 $\text{rank}(A) = \text{rank}(A, B)$.

如何判断向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 和 $\{\vec{\beta}_1, \cdots, \vec{\beta}_s\} \subset V$ 是等价的?

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} E & Q \\ O & -E \end{pmatrix} = \begin{pmatrix} A & AQ - B \end{pmatrix} = \begin{pmatrix} A & O \end{pmatrix}$$

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} -E & O \\ \tilde{Q} & E \end{pmatrix} = \begin{pmatrix} -A + B\tilde{Q} & B \end{pmatrix} = \begin{pmatrix} O & B \end{pmatrix}$$

所以, 两向量组等价的充要条件是 $\text{rank}(A) = \text{rank}(B) = \text{rank}(A, B)$.

如何判断向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 线性相关?

若线性相关, 则

$$A\vec{x} = \vec{0}$$

有非零解. 反之亦然. 所以, 线性相关的充要条件是 $\text{rank}(A) < r$.

如何判断向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 线性无关?

若线性无关, 则

$$A\vec{x} = \vec{0}$$

只有零解. 反之亦然. 所以, 线性无关的充要条件是 $\text{rank}(A) = r$.

向量组 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\} \subset V$ 是向量组 $\{\vec{\beta}_1, \cdots, \vec{\beta}_s\} \subset V$ 的极大线性无关组的充要条件是 $r = \text{rank}(A) = \text{rank}(B) = \text{rank}(A, B)$.

Remark 8.1. 请阅读 *[3]* 第三章, 第二节 — n 维向量空间和第三节 — 线性相关性. 第六章 — 线性空间.

"1128" $\to$ 向量组 $\to$ 线性相关与线性无关 $\to$ 极大线性无关组 $\to$ 向量组的秩 $\to$ 基, 维数与坐标.

8.4 线性子空间

Definition 8.2. 设线性空间 V, 若满足

- $\emptyset \neq W \subset V$;
- W 关于 V 上的线性运算是封闭的.

则 W 是 V 的线性子空间:

Example 8.1. $W = \{\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2 : a_1^2 + a_2^2 \geq 1\} \subset \mathbb{R}^2$, 虽然 W 包含零向量和负向量, 即

$$\vec{0} \in W, \ -\vec{\alpha} \in W, \forall \alpha \in W.$$

但是 W 关于 $\mathbb{R}^2$ 上的加法和数乘不构成线性子空间. 因为 W 关于这两种运算不封闭.

由向量组生成的子空间. 任何一个有限维子空间都可以用一个向量组张成.

Theorem 8.1. 两个向量组等价 $\Leftrightarrow$ 两向量组生成相同的子空间.

Theorem 8.2. $dim(L(\vec{\alpha}_1, \cdots, \vec{\alpha}_r)) = rank(\vec{\alpha}_1, \cdots, \vec{\alpha}_r)$.

Theorem 8.3. 子空间基的扩充.

子空间的交与和与直和.

Theorem 8.4. 维数公式:

$$dim(V_1) + dim(V_2) = dim(V_1 + V_2) + dim(V_1 \cap V_2).$$

Theorem 8.5. V_1 和 V_2 都是 V 的线性子空间, 下面这些条件是等价的:

- $V_1 + V_2$ 是直和;
- 零向量的表法是唯一的;
- $V_1 \cap V_2 = \{\vec{0}\}$;
- $dim(V_1 + V_2) = dim(V_1) + dim(V_2)$.

Example 8.2. 设 V 的非平凡线性子空间为 V_1 和 V_2, 则存在 $\alpha \in V$ 满足 $\alpha \notin V_1$, $\alpha \notin V_2$.

Proof. 反证法. 假设 $\forall \alpha \in V$, 若 $\alpha \notin V_1$, 则 $\alpha \in V_2$. 同时, 若 $\alpha \notin V_2$, 则 $\alpha \in V_1$. 那么, 由于 $V_1 \subsetneqq V$, 则存在 $\beta \notin V_1$, 则 $\beta \in V_2$. 同理, 存在 $\gamma \notin V_2$ 且 $\gamma \in V_1$.

基于假设条件且 $\beta + \gamma \in V$, 若 $\beta + \gamma \notin V_1$, 则 $\beta + \gamma \in V_2$, 即 $\gamma \in V_2$. 矛盾. 若 $\beta + \gamma \notin V_2$, 则 $\beta + \gamma \in V_1$, 则 $\beta \in V_1$. 矛盾. □

Example 8.3. 设 V 的线性子空间为 V_1 和 V_2, 则

$$L(V_1 \cup V2) = V_1 + V_2.$$

8.5 线性同构

Definition 8.3. V 同构于 U, 即存在一个从 V 到 U 的双射 σ, 且保持线性关系.

Theorem 8.6. 线性空间的同构是一个等价关系.

Proof. 验证同构关系满足如下性质.

- 反身性: 构造恒等映射;
- 对称性: 同构映射的逆映射保持线性关系;
- 传递性: 两个同构映射的复合也是同构映射.

□

Theorem 8.7. 设 $\sigma : V \to U$ 的同构映射, 则 $\{\vec{e}_1, \cdots, \vec{e}_s\}$ 线性相关 (无关) 的充要条件是 $\{\sigma(\vec{e}_1), \cdots, \sigma(\vec{e}_s)\}$ 线性相关 (无关).

Proof. 存在不全为零的 $a_i \in P, \forall i = 1, \cdots, s$, 使得

$$a_1\vec{e}_1 + \cdots + a_s\vec{e}_s = \vec{0} \Leftrightarrow \sigma(a_1\vec{e}_1 + \cdots + a_s\vec{e}_s) = \vec{0}$$
$$\Leftrightarrow a_1\sigma(\vec{e}_1) + \cdots + a_s\sigma(\vec{e}_s) = \vec{0}$$

□

Theorem 8.8. 任意 n 维线性空间 V 都与 $\mathbb{R}^n$ 同构.

Proof. 构造从 V 到 $\mathbb{R}^n$ 的一一映射 σ, 并验证 σ 保持线性关系. □

Theorem 8.9. 设有限维线性空间 V 和 U, 则 V 同构于 U 的充要条件是

$$dim(V) = dim(U).$$

Remark 8.2. 对于所有有限维线性空间来说, 可以基于同构这个**等价关系**进行**分类**, 同一类的不同线性空间通过同构映射进行**变换**, 所对应的**不变量**是线性空间的维数.

8.6 再论线性方程组

在中学时, 我们非常熟悉一个题目: 求解下述关于未知数 x 的方程

$$x^2 = b.$$

即, 得到方程的可解性和解的结构. 换句话说, 对于方程,

- b 满足什么条件时, 上述方程有解;
- 如果上述方程有解, 什么条件下, 解是唯一的, 什么条件下, 解不唯一;
- 如果上述方程有解, 写出解的表达式.

答案是非常容易的

- $b \in [0, +\infty) \Leftrightarrow$ 上述方程有解;
- $b \in \{0\} \Leftrightarrow$ 解是唯一的; $b \in (0, +\infty) \Leftrightarrow$ 解不唯一;
- 当 $b \in [0, +\infty)$ 时, 解 $x = \pm\sqrt{b}$.

换个角度看这些问题, 构造函数 $f(x) = x^2$, 或者更精确的表达为

$$f : \mathbb{R} \to \mathbb{R}$$
$$x \to x^2$$

映射 f 的定义域, 值域, 映射关系分别是 $\mathbb{R}$, $\mathbb{R}$, x 映成 x^2. 那么, 我们从映射 f 的角度再次看上述三个问题: 对于映射 f,

- 值域 $\mathbb{R}$ 中的每个数 b 都有原象吗?
- 如果 b 存在原象, 那么原象是唯一的吗?

- 如何给出象 b 的原象的表达式?

这样, 我们就把方程求解的问题转化到了集合与映射的框架下处理. 我们将这样的想法应用于求解线性方程组.

设矩阵

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix},\ \vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},\ \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

我们讨论下述线性方程组

$$\begin{cases} 1x_1 + 2x_2 = b_1 \\ 3x_1 + 6x_2 = b_2 \end{cases} \Leftrightarrow A\vec{x} = \vec{b}$$

的可解性和解的结构. 首先, 构造一个映射

$$\begin{aligned} &F : \mathbb{R}^2 \to \mathbb{R}^2 \\ &\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} 1x_1 + 2x_2 \\ 3x_1 + 6x_2 \end{pmatrix} = A\vec{x} \end{aligned}$$

我们同样可以提出问题: 对于线性方程组,

- $\vec{b}$ 满足什么条件时, 上述线性方程组有解;
- 如果上述线性方程组有解, 什么条件下, 解是唯一的, 什么条件下, 解不唯一;
- 如果上述线性方程组有解, 写出解的表达式.

或者等价地, 对于映射 F,

- 值域 $\mathbb{R}^2$ 中的每个向量 $\vec{b}$ 都有原象吗?
- 如果 $\vec{b}$ 存在原象, 那么原象是唯一的吗?
- 如何给出象 $\vec{b}$ 的原象的表达式?

我们观察 F 的映射关系

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} 1x_1 + 2x_2 \\ 3x_1 + 6x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 6 \end{pmatrix}$$

所以, 映射 F 的象集合是 $\mathbb{R}^2$ 的线性子空间, 它是由向量组

$$\left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \end{pmatrix} \right\}$$

张成的. 我们记象空间为 $C(A)$, 其实就是矩阵 A 的列向量组张成的值域 $\mathbb{R}^2$ 的线性子空间. 这里 C 其实就是列 (column) 的意思.

那么, 我们就得到了

$$\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in C(A) \Leftrightarrow \text{线性方程组有解} \Leftrightarrow \vec{b}\ \text{有原象}.$$

而解的唯一性问题其实就是向量组

$$\left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}2\\6\end{pmatrix}\right\}$$

关于向量 $\vec{b}\in C(A)$ 的表出系数是否唯一的问题. 根据之前的向量组线性相关和线性无关的定义, 我们得到, 当解存在, 即 $\vec{b}\in C(A)$ 时,

$$\left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}2\\6\end{pmatrix}\right\}\text{线性无关} \Leftrightarrow \text{解唯一} \Leftrightarrow \vec{b}\text{ 的原象唯一}.$$

显然, 向量组

$$\left\{\begin{pmatrix}1\\3\end{pmatrix},\begin{pmatrix}2\\6\end{pmatrix}\right\}$$

是线性相关的, 那么这就说明, 如果解存在, 则必是多解.

最后一个问题就是, 当有解时, 如何表达解? 一图说明之.

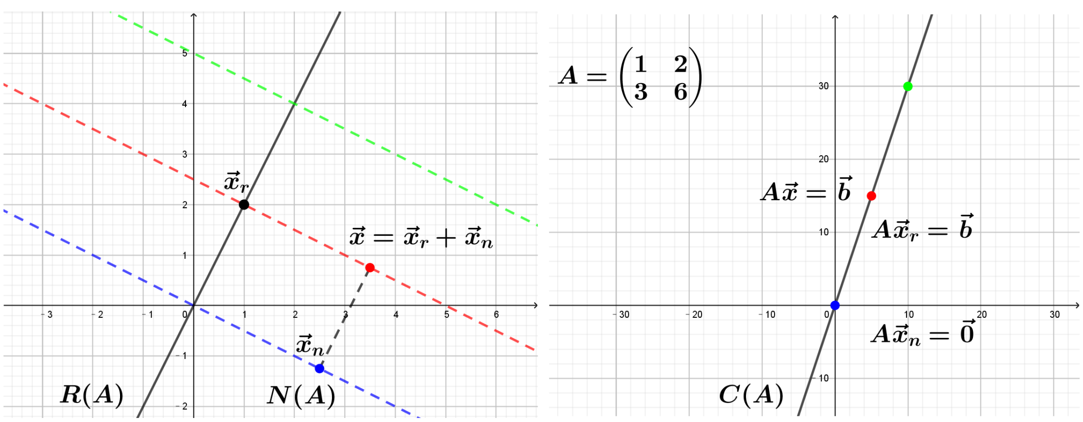


Figure 8.1: 本图参考 [1]

这里, $R(A)$ 表示矩阵 A 的行向量组张成的线性子空间. $N(A)$ 表示线性方程组 $A\vec{x}=\vec{0}$ 的解集合, 它其实是定义域 $\mathbb{R}^2$ 的线性子空间. 左图代表定义域 $\mathbb{R}^2$, 右图代表值域 $\mathbb{R}^2$. $C(A)$ 是象集合, 且绿色标记的象对应的原象集合是左图中的绿色虚线. 红色标记的象对应的原象集合是左图中的红色虚线. 蓝色标记的象对应的原象集合是左图中的蓝色虚线. 容易看出, 这三个象集合是平行的. 同时, 对于 $\vec{b}\in C(A)$ 的任一原象 $\vec{x}$, 存在唯一的 $\vec{x}_r\in R(A)$ 和 $\vec{x}_n\in N(A)$, 使得

$$\vec{x}=\vec{x}_r+\vec{x}_n$$

成立.

Remark 8.3. 显然, f 并不是一个满射, 即并不是值域中的每个数都有原象, 比如 -3 就没有原象, 即在 $\mathbb{R}$ 中找不到一个数 a 使得, $f(a)=-3$. 同时, f 也不是一个单射, 即存在一个象对应多个原象的情形, 比如 $f(3)=f(-3)=9$, 那么, 同一个象 9 对应原象 3 和 -3. 考虑 F 是单射的充要条件以及是满射的充要条件.

抽象地, 我们有下述两个定理和一张图.

Theorem 8.10. 设齐次线性方程组

$$A\vec{x} = \vec{0},$$

其中 $A \in \mathbb{R}^{m\times n}$. 那么齐次方程组

(1) 解的全体构成 $\mathbb{R}^n$ 的 $n - rank(A)$ 维子空间, 记为 $N(A)$.

(2) 矩阵 A 的所有行向量张成 $\mathbb{R}^n$ 的 $rank(A)$ 维子空间, 记为 $R(A)$.

(3) $N(A) + R(A) = N(A) \oplus R(A) = \mathbb{R}^n$.

Proof. 设 $\operatorname{rank}(A) = r$, 则存在一个 r 阶子式不为零. 不妨设所对应的 r 行和 r 列为矩阵 A 的前 r 行和前 r 列 (思考: 交换系数矩阵的行列会对方程组的解有什么影响). 那么将矩阵和向量分块

$$A = \begin{pmatrix} R & N \\ T & W \end{pmatrix}, \ \vec{x} = \begin{pmatrix} \vec{x}_R \\ \vec{x}_N \end{pmatrix}$$

其中 $\vec{x}_R \in \mathbb{R}^r$, $R \in \mathbb{R}^{r\times r}$ 为上述不为零的 r 阶子式对应的可逆矩阵矩阵. 对矩阵 A 做块初等变换, 得到

$$\begin{pmatrix} R^{-1} & O \\ -TR^{-1} & E_{n-r} \end{pmatrix} \begin{pmatrix} R & N \\ T & W \end{pmatrix} = \begin{pmatrix} E_r & R^{-1}N \\ O & W - TR^{-1}N \end{pmatrix}.$$

由于 $\operatorname{rank}(A) = r$, 则 $W - TR^{-1}S = O$. 否则存在不为零 $r+1$ 阶子式. 再根据初等变换不改变方程组的解, 则方程组 $A\vec{x} = \vec{0}$ 与方程组

$$\begin{pmatrix} E_r & R^{-1}N \\ O & O \end{pmatrix} \begin{pmatrix} \vec{x}_R \\ \vec{x}_N \end{pmatrix} = \vec{0}$$

同解. 后者的解为

$$\begin{pmatrix} \vec{x}_R \\ \vec{x}_N \end{pmatrix} = \begin{pmatrix} -R^{-1}N \\ E_{n-r} \end{pmatrix} \vec{k}.$$

其中, 任意 $\vec{k} \in \mathbb{R}^{n-r}$. 所以,

$$N(A) = \left\{ \vec{x} \in \mathbb{R}^n : \vec{x} = \begin{pmatrix} -R^{-1}N \\ E_{n-r} \end{pmatrix} \vec{k}, \forall \vec{k} \in \mathbb{R}^{n-r} \right\}.$$

容易验证 $N(A)$ 是 $\mathbb{R}^n$ 的 $n-r$ 维子空间.

容易验证, 由矩阵 $\begin{pmatrix} E_r & R^{-1}N \end{pmatrix}$ 的行向量张成了 r 维子空间, 并且 $\begin{pmatrix} E_r & R^{-1}N \end{pmatrix}$ 的行向量组与 A 的行向量组是等价的. 所以 $R(A)$ 就是 r 维子空间.

下证 $N(A) \cap R(A) = \left\{ \vec{0} \right\}$. 任取 $\vec{v} \in N(A) \cap R(A)$, 即 $\vec{v}$ 可以用 A 的行向量线性表出. 所以

$$\vec{v} = \begin{pmatrix} E_r \\ N^T R^{-T} \end{pmatrix} \vec{g}$$

其中 $\vec{g} \in \mathbb{R}^r$. 另外, $A\vec{v} = \vec{0}$, 即

$$\begin{pmatrix} E_r & R^{-1}N \end{pmatrix} \vec{v} = \begin{pmatrix} E_r & R^{-1}N \end{pmatrix} \begin{pmatrix} E_r \\ N^T R^{-T} \end{pmatrix} \vec{g} = \vec{0}.$$

那么

$$\vec{v}^T\vec{v} = \vec{g}^T \begin{pmatrix} E_r & R^{-1}N \end{pmatrix} \begin{pmatrix} E_r \\ N^T R^{-T} \end{pmatrix} \vec{g} = 0.$$

所以 $\vec{v} = \vec{0}$. 那么 $N(A) \cap R(A) = \left\{\vec{0}\right\}$, $N(A) + R(A) = N(A) \oplus R(A)$. 另外, 由

$$\dim(N(A) \oplus R(A)) = \dim(N(A)) + \dim(R(A)) = n,$$

得到 $N(A) \oplus R(A) = \mathbb{R}^n$.

□

Theorem 8.11. 设线性方程组

$$A\vec{x} = \vec{b},$$

其中 $A \in \mathbb{R}^{m\times n}$. 则

(1) A 的所有列向量张成的 $\mathbb{R}^n$ 的 $rank(A)$ 维子空间 $C(A)$.

(2) 线性方程组有解的充要条件是 $\vec{b} \in C(A)$.

(3) 若线性方程组有一个解 $\vec{x}_0$, 则所有解的集合为

$$S = \left\{\vec{x} \in \mathbb{R}^n : \vec{x} = \vec{x}_0 + \vec{v},\ \forall \vec{v} \in N(A)\right\}.$$

(4) 若线性方程组有解, 则对任一解向量 $\vec{x}$, 必存在唯一的 $\vec{x}_r \in R(A)$ 和 $\vec{x}_n \in N(A)$, 满足

$$\vec{x} = \vec{x}_r + \vec{x}_n.$$

Proof. 设 $\mathrm{rank}(A) = r$. 由向量组等价与矩阵初等列变换的关系, 容易得到 $C(A)$ 是 r 维子空间.

$$\text{方程组有解} \Leftrightarrow \mathrm{rank}(A) = \mathrm{rank}(\begin{pmatrix} A & \vec{b} \end{pmatrix}) \Leftrightarrow \vec{b}\ \text{可由}A\ \text{的列向量线性表出}.$$

显然集合 S 中的任意向量都是方程组的解, 下证方程组的解都包含在集合 S 中. 设方程组的一个解为 $\vec{y}$, 即 $A\vec{y} = \vec{b}$. 再由 $A\vec{x}_0 = \vec{b}$, 得到 $A(\vec{y} - \vec{x}_0) = \vec{0}$. 所以 $\vec{y} - \vec{x}_0 \in N(A)$, 即存在 $\vec{v} \in N(A)$, 使得 $\vec{y} - \vec{x}_0 = \vec{v}$, 即 $\vec{y} = \vec{x}_0 + \vec{v}$. 所以方程组的解都属于集合 S. □

Example 8.4. *P259, 例 2,* 对于线性方程组解空间的描述以及子空间和的描述. 另外, 讨论两个解空间的和是直和的充要条件.

设 $V = V_1 \oplus V_2 = U_1 \oplus U_2$, 且 $V_1 \perp V_2$, $U_1 \perp U_2$. 则

$$V_2 \cap U_2 = (V_1 + U_1)^{\perp},\ V_2 + U_2 = (V_1 \cap U_1)^{\perp}.$$

显然,

$$V_2 + U_2 = V_2 \oplus U_2 \Leftrightarrow V_1 + U_1 = V.$$

这里 $V_1 = R(A)$, $V_2 = N(A)$ 和 $U_1 = R(B)$, $U_2 = N(B)$. 其中, A 和 B 分别是齐次线性方程组 $A\vec{x} = \vec{0}$ 和 $B\vec{x} = \vec{0}$ 的系数矩阵.

Theorem 8.12. 设两个齐次线性方程组

$$A\vec{x} = \vec{0},\ B\vec{x} = \vec{0},$$

其中 $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{l\times n}$. 那么上述两个齐次方程组的对应的解空间的和是直和的充要条件是

$$rank(A) + rank(B) = rank(\begin{pmatrix} A \\ B \end{pmatrix}).$$

Proof.

$$n - \mathrm{rank}(\begin{pmatrix} A \\ B \end{pmatrix}) = \dim(V_2 \cap U_2)$$

□

利用矩阵和向量的语言表达线性方程组有解判别定理和解的结构, 并且表达出消元法的每个步骤. 设矩阵 $A \in \mathbb{R}^{m\times n}$, 首先保证无多余线性方程且方程组有解, 则 $m \leq n$, 也就是保证矩阵 A 行满秩. 或者直接设 A 为阶梯形矩阵. 然后将矩阵 A 分块, 使得矩阵 $R \in \mathbb{R}^{m\times m}$ 可逆.

$$A\vec{x} = \begin{pmatrix} R & N \end{pmatrix} \begin{pmatrix} x_r \\ x_n \end{pmatrix} = \vec{b}$$

$$\begin{pmatrix} x_r \\ x_n \end{pmatrix} = \begin{pmatrix} R^{-1}\vec{b} \\ 0 \end{pmatrix} + \begin{pmatrix} -R^{-1}N \\ I \end{pmatrix} \vec{k}.$$

Remark 8.4. 请阅读 *[3]* 第三章, 第五节 — 线性方程组有解判别定理和第六节 — 线性方程组的结构.

Remark 8.5. 生活里没有视频, 就好像没有阳光; 智慧里没有视频, 就好像鸟儿没有翅膀.

- 厦门大学 – 高等代数 – 第三章 – 第 *1* 节 – 线性空间

  http://www.icourses.cn/sCourse/course_3077.html

- 厦门大学 – 高等代数 – 第三章 – 第 *5* 节 – 直和分解

  http://www.icourses.cn/sCourse/course_3077.html

- 麻省理工学院 – 线性代数 – 第 *8* 集 – 求解 $Ax = b$: 可解性和解的结构

  http://open.163.com/movie/2010/11/V/8/M6V0BQC4M_M6V2ABHV8.html

9 The Big Picture

子空间与直和分解.

矩阵的行空间, 列空间, 零空间.

线性代数基本定理.

描述四个空间以及它们与线性方程组的关联.

首先基于矩阵 $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ 引入四个基本子空间.

- $C(A)$, 由矩阵 A 的所有列向量张成的子空间, 包含了由向量 $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ 进行数乘运算得到的所有向量.
- $N(A)$, 由矩阵 A 的零空间, 包含了由向量 $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ 进行数乘运算得到的所有向量, 满足使得 $A\vec{x} = 0$.
- $C(A^T)$, 由矩阵 A 的所有行向量张成的子空间, 包含了由向量 $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ 进行数乘运算得到的所有向量.
- $N(A^T)$, 由矩阵 A 的零空间, 包含了由向量 $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$ 进行数乘运算得到的所有向量, 满足使得 $A\vec{x} = 0$.

那么将线性方程组

$$\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

以映射的观点描述, 一图说明之.

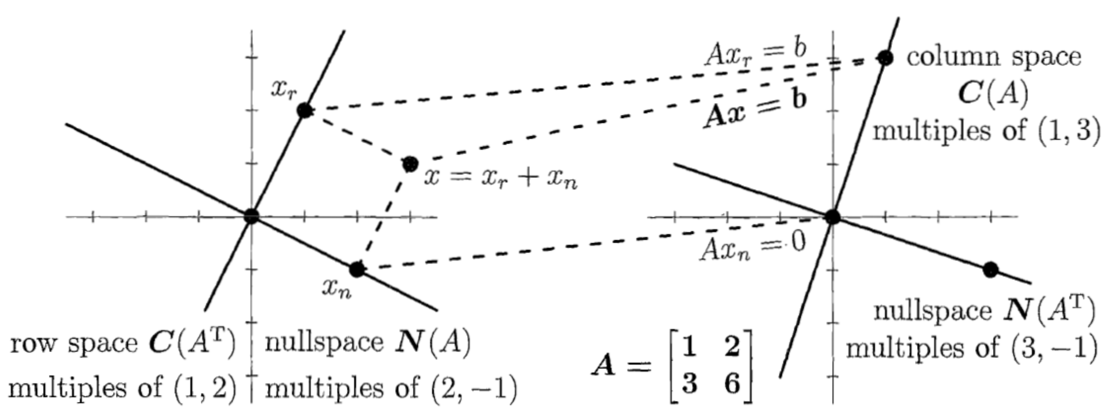


Figure 9.1: 本图摘自 [1]

此图非常形象地将线性方程组利用集合和映射表述, 并且道出了解的结构. 一图言出线性代数的核心内容.

$R(A), C(A), N(A), N(A^T)$ 分别表述矩阵 A 的行空间, 列空间, 零空间和转置的零空间, 它们都是线性空间. $R(A)$ 和 $N(A)$ 是 $\mathbb{R}^n$ 的子空间. $C(A)$ 和 $N(A^T)$ 是 $\mathbb{R}^m$ 的子空间. 设矩阵 $A \in \mathbb{R}^{m\times n}$ 且秩为 r, 则 $\dim R(A) = r$, $\dim N(A) = n - r$ 且 $R(A) \perp N(A)$. 同理, $\dim C(A) = r$, $\dim N(A^T) = m - r$ 且 $C(A) \perp N(A^T)$.

$Ax = b$ 有解的充要条件是 $b \in C(A)$, 此时 $x = x_r + x_n$, 其中 $x_r \in R(A)$ 且 $x_n \in N(A)$, 满足 $Ax_r = b$. 这样线性方程组求解的问题完全清楚了.

本节内容贯穿整个线性代数的主要内容, 深刻且简洁. 一幅图描述了线性方程组解的结构. 将线性方程组问题用集合与映射的观点描述和解决.

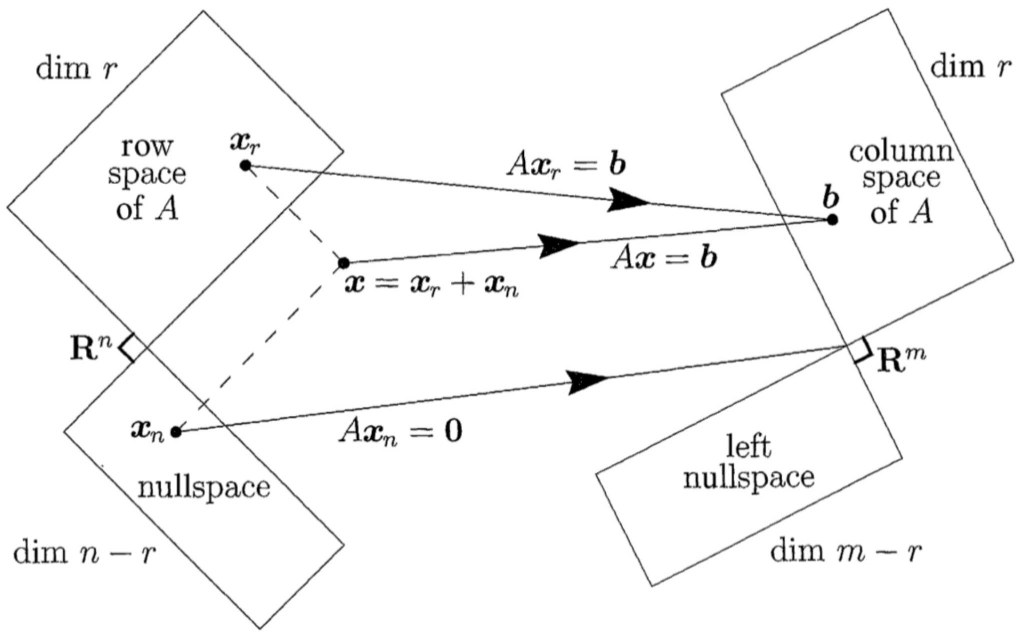


Figure 9.2: 本图摘自 [1]

不同的矩阵 A 将定义域 $\mathbb{R}^n$ 和值域 $\mathbb{R}^m$ 分解成不同的子空间. 箭头描述了映射关系, 何种情况下是“一对一”, 何种情况下是“ 多对一”, 以及值域中的哪些元素没有原象. “多对一” 时, 诸多原象的样子.

Remark 9.1. 请阅读 *[3]* 第六章 – 线性空间.

10 总结

消元法: 同解简化, 三类初等变换.

行列式: 判别式, 三条性质, 降阶公式, 大公式, Cramer 法则.

矩阵: 消元法的精确描述 (初等矩阵与矩阵的秩), 八大运算三十六个性质.

线性空间: “1128”, 线性相关与线性无关, 基与坐标, 对应概念与性质在 $\mathbb{R}^n$ 上的判别.

以行列式, 矩阵, 线性空间的观点分别描述线性方程组的初等变换. 以消元法, 行列式, 矩阵, 线性空间这四种理论论述线性方程组.

下学期主要内容: 多项式, 矩阵相抵, 矩阵相似和矩阵合同. 上述这三种等价关系对应的线性空间, 线性映射的内容. 另外就是欧式空间.

主线依然是: 变换 — 等价关系 — 不变量 — 代表元.

11 Q&A

问: 学习线性代数需要做多少题目, 什么叫做学明白了呢?

答: 学问总有根源, 找到根源就是学懂. 如何找到, 就要在做题目中思考定义, 理解逻辑. 做题目并非狭义的课后习题, 更是书本中的定理和性质. 它们更加体现数学定义的核心. 定义好比离散的点, 那么定理就是连接这些离散点的通路. 做题就是从此点到彼点的一场徒步旅行. 将所有离散点全部连接起来, 即学通. 学懂, 自然是找到一个点, 从此点出发完全覆盖所有离散点. 这也就是所谓的, 读书从薄到厚再到薄的过程. 学通的人是学者, 学懂的人是大师.

问: 一言以蔽之, 这学期高等代解决了什么问题?

答: 线性方程组求解.

问: 这学期研习线性代数, 它的本源在哪里?

答: 矩阵的初等变换.

问: 行列式与矩阵看似相似, 实则不同, 但似乎有千丝万缕的联系, 如何理解?

答: 从代数角度, n 阶方阵可以计算它的行列式. 从几何角度, 若用方阵描述正方形变作平行四边形的过程, 那么行列式就描述这个过程中面积的变化. 从线性方程组角度, 若用矩阵描述线性方程组, 那么行列式是这个线性方程组的判别式. 大体可以认为, 行列式是方阵在某个维度的投影. 直观看, 行列式是一个数, 方阵是一个数阵.

问: 矩阵的相抵关系, 能拿它做什么呢?

答: 相抵源于初等变换, 初等变换源于消元法. 相抵其实是在寻找保持矩阵秩不变的矩阵最简形式. 而消元法是在寻找线性方程组解集合不变的线性方程组最简形式. 然而, 消元法中的线性方程组的三类初等变换保持线性方程组解不变. 矩阵的初等变换源于此, 它们恰是矩阵的初等行变换. 综上, 矩阵的相抵关系源于消元法, 而高于消元法. 高于之处不只是前者更抽象, 也在于前者是对消元法同解简化思想的进一步提升. 相抵关系得到的矩阵分解, 让我们在解决一些问题时有了一个统一的套路.

问: 都说数学是哲学, 那么秩这个概念在生活中哪里有体现?

答: 秩这个概念也是由同解简化派生的. 有分类才有秩, 有分类自然有等价关系. 秩说的是它处在哪一类里, 秩的英文 rank 也体现了这个意思. 譬如, 古代中央行政机构的六部 “吏, 户, 礼, 兵, 刑, 工”. 若按这个次序编号. 那么吏部的秩就是一, 户部的秩就是二, 以此类推. 换句话说, 秩等于一的是吏部, 秩等于二的是户部. 再举一例, 兵书有云 “知己知彼百战不殆”. 战争胜败的关键点有两个 “知己” 和 “知彼”, 秩等于一的就是 “ 知己”, 秩等于二的就是 “知彼”. 这样看来秩似乎也没什么, 无非是把类型做个编号罢了. 但为何古代中央行政机构是六部而非七部或者八部, 难道战争的秘诀就有且只有两点吗? 古代中央设立行政机构无非是想处理天下事, 而天下事无非六件. 战争无非是人与人之间的大规模打斗. 研究战争是要取得战争的胜利. 而战场上是敌非友, 是友非敌. 要赢得战争的胜利, 自然就是两个关键点. 所以, 秩背后的分类更关键, 而如何分类分几类就需要仔细斟酌了. 这样看来, 向量组的秩的定义更符合这种理解.

问: 向量空间与向量的集合, 二者的差别好似前者在后者上定义了两种运算, 而且还牵扯到数域. 我们真的有必要这样做吗?

答: 空间较之集合, 它有了结构的概念. 那何为结构? 先举个例子, 我们要建校舍, 首先要搞清楚我们需要哪些材料, 将这些材料列个清单. 譬如, 水泥五吨, 沙八吨, 实木十方. 我们并没有写五个水泥一吨, 八个沙一吨, 十五个实木一方. 这里首先就有数乘概念的体现, 集合中的元素就是水泥, 沙和实木, 而数五, 八和十五就是数域中的数. 接下来, 我们就要利用水泥, 沙和实木开始修建校舍. 开工前, 堆在面前的是水泥五吨, 沙八吨, 实木十方. 竣工后, 摆在眼前的是校舍. 然而仔细想来, 校舍不就是水泥五吨, 沙八吨, 实木十方吗! 这好比, 空间是集合, 且不只是集合. 校舍区别于原材料的原因在于它是原材料遵循一定规则架构而成的. 那么, 空间就是通过加法和数乘架构而成, 所言的规则就是那八条法则. 这种基于运算的架构给了空间一种结构感. 而结构感源于关联性, 运算恰恰制造了关联.

问: 线性空间中的诸多概念最后的落脚点大都是线性方程组, 判断其有解无解, 唯一解还是多解. 那么是什么促成了这样一个结果呢?

答: 线性空间无非是线性相关与线性无关, 而二者源于线性运算 — 加法和数乘. 设定数乘中的数为未知数, 那么可以线性表出就是一个线性方程组有解. 线性相关就是一个齐次线性方程组有非零解. 线性无关就是一个齐次线性方程组只有零解. 其它概念无外乎于此.

问: 矩阵有秩, 向量组也有秩. 通过行列式可以矩阵的秩, 通过相抵关系的标准形也可以定义秩, 通过线性相关与无关也可以定义秩. 秩的这种多面性起源于哪里呢?

答: "万法归一, 一归何处", 本学期的线性代数所学概念方法归于矩阵这一个概念, 而矩阵最核心的概念是初等变换. 矩阵乃万物之源, 此源如何生万物, 初等变换是也. 秩的多面性源于初等变换在线性方程组, 行列式, 矩阵, 线性空间中的表现形式. 具体来讲, 之于线性方程组, 初等变化乃 Gauss 消元法, 秩乃有效方程个数. 之于行列式, 初等变换乃计算行列式之利器, 秩可用于判别行列式是否为零. 之于矩阵, 初等变换回归变换的本源, 秩回归不变量的本源. 之于向量组, 初等变换乃线性运算, 秩乃分类数.

数学善于为不同事物起相同的名字, 矩阵, 初等变换和秩都是抽象的名字. 在不同情景下, 它们又变成其它名字. 在线性方程组中, 矩阵变成了系数和右端项, 初等变换就是消元法的三板斧, 秩变成了有效方程数. 在线性空间中, 矩阵变成了向量组, 初等变换就是线性运算, 秩变成了基于线性相关对向量组的分类数.

集合, 变换, 不变量又是矩阵, 初等变换和秩的抽象.

问: 是否可以举几个生活中运用 "变换 — 等价关系 — 不变量 — 分类 — 代表元" 的例子.

答: 按照某个标准, 动物的分类, 垃圾的分类, 手机应用的分类. 以手机应用为例, 理财、出行、健康、系统工具、社交通讯、新闻阅读、影音娱乐. 再以数钱为例, 若以面值为分类标准, 则数钱这份工作都变成了数张数. 而且, 这种分类的不变量除了面值外, 还有尺寸. 这样, 使得三岁顽童依然可以胜任分类这个工作. 生活中, 分类无处不在.

12 多项式 – 一个萝卜一个坑

阐述多项式因式分解与多项式函数根的关系.

Definition 12.1. 数域 $\mathbb{P}$ *(加减乘除)*, 一元多项式环 $\mathbb{P}[x]$ *(加乘)*.

Theorem 12.1 (带余除法)**.** 对于 $\mathbb{P}[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$, 其中 $g(x) \neq 0$, 则在 $\mathbb{P}[x]$ 中存在唯一多项式 $q(x)$ 与 $r(x)$, 使得

$$f(x) = q(x)g(x) + r(x)$$

成立, 其中 $\partial(r(x)) < \partial(g(x))$ 或者 $r(x) = 0$.

若 $r(x) = 0$, 则称多项式 $g(x)$ 整除 $f(x)$, 记为 $g(x)|f(x)$.

整除性的常用性质: 对称等同性; 传递性; 组合性.

记两个不全为零多项式 $f(x)$ 和 $g(x)$ 的首一最大公因式为 $(f(x), g(x))$.

若 $(f(x), g(x)) = 1$ 则称 $f(x)$ 和 $g(x)$ 互素.

Theorem 12.2. 若成立

$$f(x) = q(x)g(x) + r(x),$$

则 $f(x)$, $g(x)$ 和 $g(x)$, $r(x)$ 有相同的公因子.

Theorem 12.3. 对于 $\mathbb{P}[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$, 则存在最大公因子 $d(x) \in \mathbb{P}[x]$ 且满足

$$d(x) = u(x)f(x) + v(x)g(x),$$

其中 $u(x), v(x) \in \mathbb{P}[x]$.

记 $(f(x), g(x))$ 为 $f(x)$ 和 $g(x)$ 的首一 (最高次项系数为 1) 的最大公因式.

Example 12.1. 设 $f(x) = 5(x-1)$, $g(x) = (x-1)(x+2)$, 则 $(x-1), 2(x-1), \sqrt{3}(x-1), \cdots$ 都是 $f(x)$ 和 $g(x)$ 的最大公因式. 它们之间相差非零常数倍.

因式分解 — 不可约多项式.

Theorem 12.4. 设 $f(x) \in \mathbb{P}[x]$, $\partial(f(x)) \geq 1$, 那么存在互不相同的首一不可约多项式 $p_1(x), \cdots, p_s(x) \in \mathbb{P}[x]$, 正整数 $r_1, \cdots, r_s \in \mathbb{N}^+$, $c \in \mathbb{P}$, 满足

$$f(x) = cp_1(x)^{r_1} \cdots p_s(x)^{r_s}.$$

并且这种分解是唯一的.

Theorem 12.5. 设 $p(x) \in \mathbb{P}[x]$, 则 $p(x)$ 是不可约多项式的充要条件是 $\forall f(x) \in \mathbb{P}[x]$, 满足

$$要么 p(x)|f(x),\ 要么 (p(x), f(x)) = 1.$$

Theorem 12.6 (复数域上的因式分解定理)**.** 设 $f(x) \in \mathbb{C}[x]$, $\partial(f(x)) \geq 1$, 那么存在互不相同的首一一次多项式 $p_1(x), \cdots, p_s(x) \in \mathbb{C}[x]$, 正整数 $r_1, \cdots, r_s \in \mathbb{N}^+$, $c \in \mathbb{C}$, 满足

$$f(x) = cp_1(x)^{r_1} \cdots p_s(x)^{r_s} = c(x-\alpha_1)^{r_1} \cdots (x-\alpha_s)^{r_s}.$$

Remark 12.1. 数域 — 一元多项式环 — 整除 — 带余除法 — 因式 — 公因子 — 最大公因式 — 辗转相除 — 互素 — 不可约多项式 — 因式分解定理.

13 线性映射 – 两条腿走路的巨人

线性代数中的集合与映射:

- **线性空间;**
- **线性映射.**

Definition 13.1. 保持线性运算的映射 $\mathcal{A}$ 称为线性映射, 它的定义域 V 和值域 W 都是数域 $\mathbb{K}$ 上的线性空间.

若定义域和值域相同, 则称 $\mathcal{A}$ 为 V 上的线性变换.

这里保持线性运算, 可以理解为, 映射 $\mathcal{A}$ 可以与线性空间中的加法和数乘交换顺序. 或者说, 它们是可交换的. 具体来讲,

$$\mathcal{A}(\vec{v}+\vec{u})=\mathcal{A}(\vec{v})+\mathcal{A}(\vec{u}),\ \mathcal{A}(k\vec{v})=k\mathcal{A}(\vec{v}),\ \forall\vec{u},\vec{v}\in V,\ \forall k\in\mathbb{K}.$$

线性映射之间可以定义加法, 数乘, 乘法 (复合).

Theorem 13.1. $\mathcal{L}(V,W)$ 是线性空间.

Definition 13.2. 设 $\mathcal{A}\in\mathcal{L}(V)$, 若存在 $\mathcal{B}\in\mathcal{L}(V)$ 使得

$$\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}=\mathcal{E}.$$

称 $\mathcal{B}$ 为 $\mathcal{A}$ 的逆变换.

容易验证 $\mathcal{B}\in\mathcal{L}(V)$.

Theorem 13.2. 设 $f(x),g(x)\in\mathbb{P}[x]$, 且 $h(x)=f(x)g(x)$. 那么 $\forall\mathcal{A}\in\mathcal{L}(V)$,

$$h(\mathcal{A})=f(\mathcal{A})\circ g(\mathcal{A})$$

那么可以定义线性变换 $\mathcal{A}$ 的多项式.

13.1 线性映射与矩阵

主线: 建立线性空间 $\mathcal{L}(V,W)$ 与 $\mathbb{P}^{m\times n}$ 线性同构关系.

- 建立映射

$$\begin{aligned}\sigma:\mathcal{L}(V,W)&\to\mathbb{P}^{m\times n}\\ \mathcal{A}&\to A.\end{aligned}$$

- 证明映射 σ 是单射也是满射.
- 证明映射 σ 保持线性运算.

V 和 W 分别是 n 维和 m 维线性空间, 其基分别是

$$\left\{\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n\right\}$$

和
$$\left\{\vec{\eta}_1,\cdots,\vec{\eta}_m\right\}.$$
那么线性映射 $\mathcal{A}$ 满足 $\forall i=1,\cdots,n$
$$\mathcal{A}\vec{\varepsilon}_i=\sum_{k=1}^{m}a_{ki}\vec{\eta}_k.$$
形式记号
$$\mathcal{A}(\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n)=(\vec{\eta}_1,\cdots,\vec{\eta}_m)A,$$
其中, $A=(a_{ij})\in\mathbb{R}^{m\times n}$.

任意 $\vec{\xi}=\sum_{i=1}^{n}x_i\vec{\varepsilon}_i\in V$,
$$\mathcal{A}\vec{\xi}=\mathcal{A}\sum_{i=1}^{n}x_i\vec{\varepsilon}_i=\sum_{i=1}^{n}x_i\mathcal{A}\vec{\varepsilon}_i=\sum_{i=1}^{n}x_i\sum_{k=1}^{m}a_{ki}\vec{\eta}_k=\sum_{k=1}^{m}\sum_{i=1}^{n}a_{ki}x_i\vec{\eta}_k.$$
利用形式记号表达为
$$\mathcal{A}\vec{\xi}=\mathcal{A}(\vec{\varepsilon}_1,\vec{\varepsilon}_2,\cdots,\vec{\varepsilon}_n)\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}=(\vec{\eta}_1,\vec{\eta}_2,\cdots,\vec{\eta}_m)A\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}.$$
这就是映射 σ 的映射关系.

基于以下定理, 可以证明 σ 的双射性质. 容易验证它的线性性质.

Theorem 13.3. 设 $\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n$ 是 n 维线性空间 V 上的一组基, $\vec{\alpha}_1,\cdots,\vec{\alpha}_n$ 是 m 维线性空间 W 中的任意 n 个向量, 则存在唯一的线性映射线性映射 $\mathcal{A}:V\to W$ 使得
$$\mathcal{A}\vec{\varepsilon}_i=\vec{\alpha}_i,\ i=1,\cdots,n.$$

Theorem 13.4. 线性映射复合与矩阵相乘, 逆映射与逆矩阵.

Theorem 13.5. 通过线性映射的表示矩阵计算像.

然而, 线性同构映射 σ 的映射关系与 V 和 W 的基的选取直接相关. 不同的基对应不同的映射关系. 设可逆矩阵 $P\in\mathbb{R}^{n\times n}$, $Q\in\mathbb{R}^{m\times m}$. 若基变换
$$(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)=(\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n)P,$$
$$(\vec{\psi}_1,\cdots,\vec{\psi}_n)=(\vec{\eta}_1,\cdots,\vec{\eta}_m)Q,$$
则
$$\mathcal{A}(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)P^{-1}=(\vec{\psi}_1,\cdots,\vec{\psi}_m)Q^{-1}A,$$
即
$$\mathcal{A}(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)=(\vec{\psi}_1,\cdots,\vec{\psi}_m)Q^{-1}AP.$$
这就是过渡矩阵与表示矩阵的联系.

Example 13.1. 求导运算在多项式空间的矩阵表达.

Example 13.2. 一元多项式环上的平移运算是线性变换.

13.2 基的选取与表示矩阵

我们知道, 对于线性映射

$$\mathcal{A}: V \to W,$$

选取 V 和 W 的基的不同, 会导致线性映射 $\mathcal{A}$ 对应的表示矩阵不一样.

$$\mathcal{A}(\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n) = (\vec{\eta}_1, \cdots, \vec{\eta}_m)A,$$

其中, $A = (a_{ij}) \in \mathbb{R}^{m\times n}$.

我们现在的目标是如何选择 V 和 W 的基, 使得线性映射 $\mathcal{A}$ 的表示矩阵变简单. 或者说, 表示矩阵是对角阵或接近于对角阵.

13.2.1 对称矩阵的例子

设 $V = W = \mathbb{R}^2$, 且

$$\begin{aligned}\mathcal{A}: \mathbb{R}^2 &\to \mathbb{R}^2\\ \vec{v} &\to A\vec{v}\end{aligned}$$

其中

$$A = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}.$$

其实, 若 V 和 W 取同一组基, 即

$$\vec{\varepsilon}_1 = \vec{\eta}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \vec{\varepsilon}_2 = \vec{\eta}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

那么, $\mathcal{A}$ 对应的表示矩阵为 A. 由图13.1, 向量 $\vec{\varepsilon}_1$ (左图红色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_1$ (右图红色向量); 向量 $\vec{\varepsilon}_2$ (左图蓝色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_2$ (右图蓝色向量).

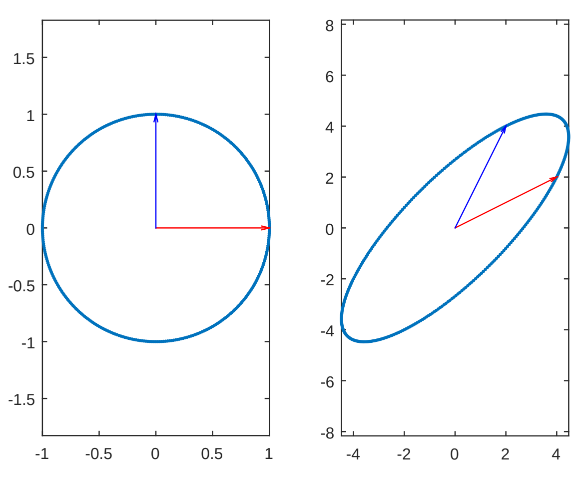


Figure 13.1: 左图是 V, 右图是 W.

若取 V 和 W 取另外同一组基, 即

$$\vec{\varepsilon}_1=\vec{\eta}_1=\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix},\ \vec{\varepsilon}_2=\vec{\eta}_2=\begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}.$$

那么, $\mathcal{A}$ 对应的表示矩阵为

$$\Lambda=\begin{pmatrix}6&0\\0&2\end{pmatrix}.$$

这是由于

$$\mathcal{A}\vec{\varepsilon}_1=6\vec{\varepsilon}_1=6\vec{\eta}_1,\ \mathcal{A}\vec{\varepsilon}_2=2\vec{\varepsilon}_2=2\vec{\eta}_2.$$

更重要的是,

$$\{\vec{\varepsilon}_1,\vec{\varepsilon}_2\},\ \{\vec{\eta}_1,\vec{\eta}_2\}$$

都是标准正交基, 而且这两组基是一样的. 由图13.2, 向量 $\vec{\varepsilon}_1$ (左图红色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_1$ (右图红色向量); 向量 $\vec{\varepsilon}_2$ (左图蓝色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_2$ (右图蓝色向量).

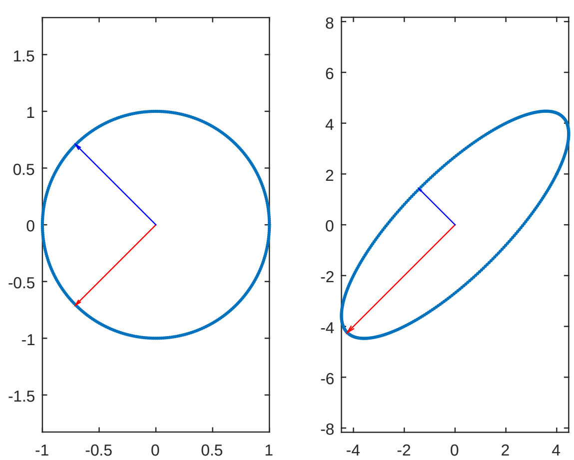


Figure 13.2: 左图是 V, 右图是 W.

这两个表示矩阵的关系为

$$A=Q\Lambda Q^{-1}=Q\Lambda Q^T,$$

其中

$$Q=\begin{pmatrix}\vec{\varepsilon}_1&\vec{\varepsilon}_2\end{pmatrix}=\begin{pmatrix}1/\sqrt{2}&-1/\sqrt{2}\\1/\sqrt{2}&1/\sqrt{2}\end{pmatrix}.$$

13.2.2 非对称矩阵的例子

设 $V=W=\mathbb{R}^2$, 且

$$\begin{aligned}\mathcal{A}:\mathbb{R}^2&\to\mathbb{R}^2\\ \vec{v}&\to A\vec{v}\end{aligned}$$

其中

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

其实, 若 V 和 W 取同一组基, 即

$$\vec{\varepsilon}_1 = \vec{\eta}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ \vec{\varepsilon}_2 = \vec{\eta}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

那么, $\mathcal{A}$ 对应的表示矩阵为 A. 由图13.3, 向量 $\vec{\varepsilon}_1$ (左图红色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_1$ (右图红色向量); 向量 $\vec{\varepsilon}_2$ (左图蓝色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_2$ (右图蓝色向量).

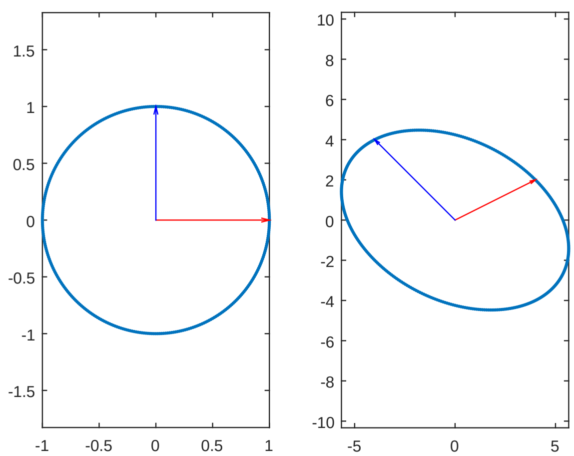


Figure 13.3: 左图是 V, 右图是 W.

若取 V 和 W 取另外同一组基, 即

$$\vec{\varepsilon}_1 = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix},\ \vec{\varepsilon}_2 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}.$$

且

$$\vec{\eta}_1 = \begin{pmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix},\ \vec{\eta}_2 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}.$$

那么, $\mathcal{A}$ 对应的表示矩阵为

$$\Sigma = \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix}.$$

这是由于

$$\mathcal{A}\vec{\varepsilon}_1 = 6\vec{\eta}_1,\ \mathcal{A}\vec{\varepsilon}_2 = 4\vec{\eta}_2.$$

更重要的是,

$$\{\vec{\varepsilon}_1, \vec{\varepsilon}_2\},\ \{\vec{\eta}_1, \vec{\eta}_2\}$$

都是标准正交基. 但是, 由于此时 A 是非对称矩阵, 所以并不能找到同一组标准正交基. 由图13.4, 向量 $\vec{\varepsilon}_1$ (左图红色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_1$ (右图红色向量); 向量 $\vec{\varepsilon}_2$ (左图蓝色向量) 作用于线性映射 $\mathcal{A}$, 得到向量 $A\vec{\varepsilon}_2$ (右图蓝色向量).

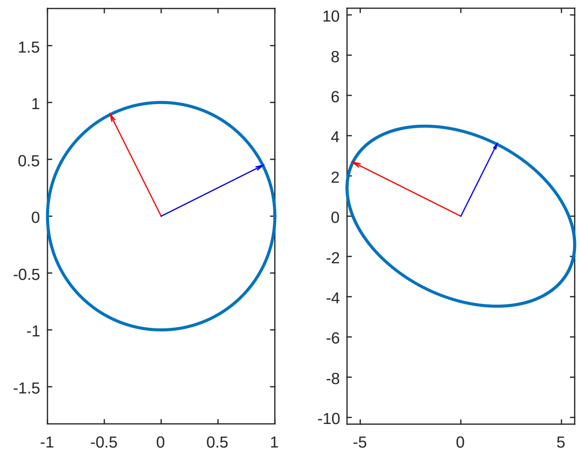


Figure 13.4: 左图是 V, 右图是 W.

这两个表示矩阵的关系为

$$A = Q\Sigma P^{-1} = Q\Sigma P^T,$$

其中

$$Q = \begin{pmatrix} \vec{\eta}_1 & \vec{\eta}_2 \end{pmatrix} = \begin{pmatrix} -2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}, \ P = \begin{pmatrix} \vec{\varepsilon}_1 & \vec{\varepsilon}_2 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix}.$$

13.3 线性映射的像空间与核空间

Definition 13.3. 下述集合称为线性映射 $\mathcal{A}: V \to W$ 的像空间和核空间, 分别记为

$$Im(\mathcal{A}) = \mathcal{A}V = \{\mathcal{A}\xi \in W | \xi \in V\},$$
$$Ker(\mathcal{A}) = \mathcal{A}^{-1}(\vec{0}) = \{\xi \in V | \mathcal{A}\xi = \vec{0}\}.$$

称 $dim(Im(\mathcal{A}))$ 和 $dim(Ker(\mathcal{A}))$ 分别为 $\mathcal{A}$ 的秩和零度.

Theorem 13.6. $Im(\mathcal{A})$ 和 $Ker(\mathcal{A})$ 分别是 W 和 V 的子空间, 且

$$Im(\mathcal{A}) = L(\mathcal{A}(\varepsilon_1, \cdots, \varepsilon_n)),$$

其中 $\{\varepsilon_1, \cdots, \varepsilon_n\}$ 是 V 中的一组基.

Theorem 13.7. 设 $\{\vec{\eta}_1, \cdots, \vec{\eta}_r\}$ 是 $Im(\mathcal{A})$ 的一组基, 且 $\mathcal{A}\vec{\varepsilon}_i = \vec{\eta}_i$, $i = 1, \cdots, r$, $\{\vec{\varepsilon}_{r+1}, \cdots, \vec{\varepsilon}_n\}$ 是 $Ker(\mathcal{A})$ 的一组基, 则 $\{\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_r, \vec{\varepsilon}_{r+1}, \cdots, \vec{\varepsilon}_n\}$ 构成 V 上的一组基. 那么

$$dim(Im(\mathcal{A})) + dim(Ker(\mathcal{A})) = dim(V).$$

Proof. 首先证明向量组线性无关; 然后证明向量组可以表出 V 中所有向量. □

Theorem 13.8. 设同构映射 $\sigma_1 : V \to \mathbb{P}^n$ 和同构映射 $\sigma_2 : W \to \mathbb{P}^m$, 满足

$$\sigma_1(\vec{\alpha}) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ \sigma_2(\vec{\beta}) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

其中

$$\vec{\alpha} = (\vec{\varepsilon}_1, \vec{\varepsilon}_2, \cdots, \vec{\varepsilon}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ \vec{\beta} = (\vec{\eta}_1, \vec{\eta}_2, \cdots, \vec{\eta}_m) \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

$\{\vec{\varepsilon}_1, \vec{\varepsilon}_2, \cdots, \vec{\varepsilon}_n\}$ 和 $\{\vec{\eta}_1, \vec{\eta}_2, \cdots, \vec{\eta}_m\}$ 分别是 V 和 W 的一组基. 基于 V 和 W 的这组基, 构造同构映射 $\sigma : L(V, W) \to \mathbb{P}^{m\times n}$, 使得 $\sigma(\mathcal{A}) = A$. 那么

$$\sigma_2 \circ \mathcal{A} = A \circ \sigma_1 : V \to \mathbb{P}^m,$$

且

$$\sigma_2(Im(\mathcal{A})) = Im(A), \ \sigma_1(Ker(\mathcal{A})) = Ker(A).$$

那么

$$Im(\mathcal{A}) \text{ 同构于} Im(A), \ Ker(\mathcal{A}) \text{ 同构于} Ker(A).$$

换句话说,

$$dim(Im(\mathcal{A})) = dim(Im(A)) = rank(A), \ dim(Ker(\mathcal{A})) = dim(Ker(A)).$$

Proof.

$$\begin{array}{ccc} V & \xrightarrow{\mathcal{A}} & W \\ \downarrow{\sigma_1} & & \downarrow{\sigma_2} \\ \mathbb{P}^n & \xrightarrow{A} & \mathbb{P}^m \end{array}$$

□

Example 13.3. 设 $\mathcal{A} \in \mathcal{L}(V, W)$, 则

- $\mathcal{A}$ 是单射 $\Leftrightarrow Ker(\mathcal{A}) = \{\vec{0}\} \Leftrightarrow$ 表示矩阵 A 列满秩;
- $\mathcal{A}$ 是满射 $\Leftrightarrow Im(\mathcal{A}) = W \Leftrightarrow$ 表示矩阵 A 行满秩;
- 若 $V = W$, 则 $\mathcal{A}$ 是单射 $\Leftrightarrow \mathcal{A}$ 是满射.

Example 13.4. 求导运算和平移运算在多项式空间的像与核. 并在某组基下写出其表示矩阵, 并验证本节中定理的结论.

Example 13.5. 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$, 且 $dim(V) = n$, W 是 V 的一个子空间. 求证:

- $\mathcal{A}(W)$ 是 V 的线性子空间.
- $dim(W) = dim(Ker\mathcal{A} \cap W) + dim(\mathcal{A}(W))$.

Example 13.6. 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$, 则

$$dim(Ker(\mathcal{A} \circ \mathcal{B})) \le dim(Ker(\mathcal{A})) + dim(Ker(\mathcal{B})).$$

Proof. **方法一:** 设 $\mathrm{Ker}(\mathcal{A}) \cap \mathrm{Ker}(\mathcal{B})$ 的一组基为

$$\varepsilon_1, \cdots, \varepsilon_i.$$

将之扩充成 $\mathrm{Ker}(\mathcal{A})$ 的基为

$$\varepsilon_1, \cdots, \varepsilon_i, \eta_1, \cdots, \eta_j.$$

将之扩充成 $\mathrm{Ker}(\mathcal{B})$ 的基为

$$\varepsilon_1, \cdots, \varepsilon_i, \gamma_1, \cdots, \gamma_k.$$

将之扩充成 V 的基为

$$\varepsilon_1, \cdots, \varepsilon_i, \eta_1, \cdots, \eta_j, \gamma_1, \cdots, \gamma_k, \beta_1, \cdots, \beta_l.$$

又由于

$$\mathrm{Im}(\mathcal{B}) \cap \mathrm{Ker}(\mathcal{A}) \subset \mathrm{Ker}(\mathcal{A}),$$

所以,

$$\dim(\mathrm{Ker}(\mathcal{A} \circ \mathcal{B})) \leq (i+k) + (i+j) = \dim(\mathrm{Ker}(\mathcal{B})) + \dim(\mathrm{Ker}(\mathcal{A})).$$

方法二: 设 $\mathrm{Im}(\mathcal{B}) \cap \mathrm{Ker}(\mathcal{A})$ 的一组基为

$$\varepsilon_1, \cdots, \varepsilon_i.$$

将之扩充成 $\mathrm{Ker}(\mathcal{A})$ 的基为

$$\varepsilon_1, \cdots, \varepsilon_i, \eta_1, \cdots, \eta_j.$$

将之扩充成 $\mathrm{Im}(\mathcal{B})$ 的基为

$$\varepsilon_1, \cdots, \varepsilon_i, \gamma_1, \cdots, \gamma_k.$$

设

$$\mathcal{B}(\alpha_1) = \varepsilon_1, \cdots, \mathcal{B}(\alpha_i) = \varepsilon_i, \mathcal{B}(\beta_1) = \gamma_1, \cdots, \mathcal{B}(\beta_k) = \gamma_k,$$

以及 $\mathrm{Ker}(\mathcal{B})$ 的基为

$$\xi_1, \cdots, \xi_l.$$

所以, V 的基为

$$\alpha_1, \cdots, \alpha_i, \beta_1, \cdots, \beta_k, \xi_1, \cdots, \xi_l,$$

而且 $\mathrm{Ker}(\mathcal{A} \circ \mathcal{B})$ 的基为

$$\xi_1, \cdots, \xi_l, \alpha_1, \cdots, \alpha_i.$$

显然,

$$\dim(\mathrm{Ker}(\mathcal{A} \circ \mathcal{B})) = l + i \leq i + j + l = \dim(\mathrm{Ker}(\mathcal{A})) + \dim(\mathrm{Ker}(\mathcal{B})).$$

□

Example 13.7. 设 $\mathcal{A}, \mathcal{B} \in \mathcal{L}(V)$, 且 $dim(V) = n$, 求证:

$$dim(Im(\mathcal{A} \circ \mathcal{B})) \geq dim(Im\mathcal{A}) + dim(Im\mathcal{B}) - n.$$

Proof. 设 $\mathrm{Im}(\mathcal{A} \circ \mathcal{B})$ 的一组基为

$$\varepsilon_1, \cdots, \varepsilon_i.$$

将之扩充成 $\mathrm{Im}(\mathcal{A})$ 的基为

$$\varepsilon_1, \cdots, \varepsilon_i, \eta_1, \cdots, \eta_j.$$

设

$$\mathcal{B}(\alpha_1) = \varepsilon_1, \cdots, \mathcal{B}(\alpha_i) = \varepsilon_i,$$

以及 $\mathrm{Ker}(\mathcal{A}|_{\mathrm{Im}\mathcal{B}})$ 的基为

$$\xi_1, \cdots, \xi_k.$$

所以, Im$\mathcal{B}$ 的基为

$$\alpha_1, \cdots, \alpha_i, \xi_1, \cdots, \xi_k.$$

显然,

$$\dim(\mathrm{Im}(\mathcal{A} \circ \mathcal{B})) = i,\ \dim(\mathrm{Im}\mathcal{A}) + \dim(\mathrm{Im}\mathcal{B}) = i + j + i + k.$$

由于

$$\mathrm{Ker}(\mathcal{A}|_{\mathrm{Im}\mathcal{B}}) \subset \mathrm{Ker}(\mathcal{A}),$$

则 $k \leq n - (i + j)$, 即

$$i \geq i + j + i + k - n.$$

□

Remark 13.1. 视频的真正目的在于诱导头脑自己去思考.

- 厦门大学 – 高等代数 – 第四章 – 第 *2* 节 – 线性映射和运算
  http://www.icourses.cn/sCourse/course_3077.html
- 厦门大学 – 高等代数 – 第四章 – 第 *4* 节 – 像与核
  http://www.icourses.cn/sCourse/course_3077.html
- 麻省理工学院 – 线性代数 – 第 *31* 集 – 线性变换及对应矩阵
  http://open.163.com/newview/movie/free?pid=M6V0BQC4M&mid=M6V2B60PJ

14 等价关系: 相抵

线性映射 $\mathcal{A}: V \to W$ 在不同基下的表示矩阵.

$$\begin{aligned}\mathcal{A}(\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n) &= (\vec{\eta}_1, \cdots, \vec{\eta}_m)A,\\ (\vec{\varphi}_1, \cdots, \vec{\varphi}_n) &= (\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n)Q,\\ (\vec{\psi}_1, \cdots, \vec{\psi}_m) &= (\vec{\eta}_1, \cdots, \vec{\eta}_m)P^{-1}.\end{aligned}$$

$$\mathcal{A}(\vec{\varphi}_1, \cdots, \vec{\varphi}_n) = (\vec{\psi}_1, \cdots, \vec{\psi}_m)PAQ = (\vec{\psi}_1, \cdots, \vec{\psi}_m)\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}.$$

Definition 14.1. 称矩阵 A 与 B 是相抵的, 若存在可逆矩阵 P 与 Q 使得 $B = PAQ$.

(HW) 验证相抵关系是等价关系.

Theorem 14.1. 矩阵的秩是矩阵相抵关系的完全不变量.

(HW) 对于具体一个线性映射, 找出基变换, 使得它的表示矩阵为 D_r.

(HW) 将上述理论应用于线性方程组的描述和求解中.

15 等价关系: 相似

斐波那契数列 (Fibonacci sequence), 又称黄金分割数列. 它因数学家列昂纳多 · 斐波那契 (Leonardoda Fibonacci) 以兔子繁殖为例子而引入, 故又称为 “兔子数列”.

一般而言, 兔子在出生两个月后就有繁殖能力. 一对兔子每个月能生出一对小兔子来. 如果所有兔子都不死, 那么一年以后可以繁殖多少对兔子?

我们不妨拿新出生的一对小兔子分析一下:

- 第一个月, 小兔子没有繁殖能力, 所以还是一对;
- 两个月后, 生下一对小兔对数共有两对;
- 三个月以后, 老兔子又生下一对, 因为小兔子还没有繁殖能力, 所以一共是三对;
- $\cdots$

设第 n 个月后, 兔子一共是 $F(n)$ 对, 则我们可以得到递推公式

$$F(n) = F(n-1) + F(n-2),$$

其中, $F(0) = F(1) = 1$.

这就是一种变化过程, 知道上个月和上上个月的兔子对数, 就可以知道这个月的兔子对数. 我们是否可以通过线性映射去描述这种变化呢?

$$\begin{pmatrix} F(n+1) \\ F(n) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F(n) \\ F(n-1) \end{pmatrix} = A \begin{pmatrix} F(n) \\ F(n-1) \end{pmatrix} = A^2 \begin{pmatrix} F(n-1) \\ F(n-2) \end{pmatrix} = \cdots = A^n \begin{pmatrix} F(1) \\ F(0) \end{pmatrix}$$

如果我们能计算出 A^n 的公式, 那么我们就可以得到斐波那契数列的通项公式.

考虑, 若存在 λ_1 和 λ_2, 满足 $\lambda_1 + \lambda_2 = 1$, $\lambda_1\lambda_2 = -1$. 那么,

$$\begin{aligned} F(n+1) - \lambda_2 F(n) &= \lambda_1(F(n) - \lambda_2 F(n-1)), \\ F(n+1) - \lambda_1 F(n) &= \lambda_2(F(n) - \lambda_1 F(n-1)), \end{aligned}$$

即,

$$\begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix} \begin{pmatrix} F(n+1) \\ F(n) \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix} \begin{pmatrix} F(n) \\ F(n-1) \end{pmatrix}$$

换句话讲,

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -\lambda_2 \\ 1 & -\lambda_1 \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -\lambda_1 & \lambda_2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -\lambda_1 & \lambda_2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

如果存在 $\lambda_1, \lambda_2 \in \mathbb{R}$, $\vec{\xi_1}, \vec{\xi_2} \in \mathbb{R}^2$, 使得

$$A\vec{\xi_1} = \lambda_1\vec{\xi_1},\ A\vec{\xi_2} = \lambda_2\vec{\xi_2},$$

且对任意 $\vec{u}\in\mathbb{R}^2$,

$$\vec{u}=k_1\vec{\xi_1}+k_2\vec{\xi_2},$$

那么

$$A^n\vec{u}=A^n(k_1\vec{\xi_1}+k_2\vec{\xi_2})=k_1\lambda_1^n\vec{\xi_1}+k_2\lambda_2^n\vec{\xi_2}.$$

$$\det(\lambda E-A)=0\Leftrightarrow\lambda^2-\lambda-1=0\Leftrightarrow\lambda=\frac{1\pm\sqrt{5}}{2}$$

$$\lambda_1=\frac{1+\sqrt{5}}{2},\ \vec{\xi_1}=\begin{pmatrix}(1+\sqrt{5})/2\\1\end{pmatrix},\ \lambda_2=\frac{1-\sqrt{5}}{2},\ \vec{\xi_2}=\begin{pmatrix}(1-\sqrt{5})/2\\1\end{pmatrix}.$$

另外,

$$\begin{pmatrix}F(1)\\F(0)\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix}=\frac{\sqrt{5}+1}{2\sqrt{5}}\vec{\xi_1}+\frac{\sqrt{5}-1}{2\sqrt{5}}\vec{\xi_2},$$

即,

$$k_1=\frac{\sqrt{5}+1}{2\sqrt{5}},\ k_2=\frac{\sqrt{5}-1}{2\sqrt{5}}.$$

那么, 得到通项

$$F(n)=\frac{1}{\sqrt{5}}\Big((\frac{1+\sqrt{5}}{2})^{n+1}-(\frac{1-\sqrt{5}}{2})^{n+1}\Big).$$

15.1 为什么特征值、特征向量很重要?

因为可以很容易计算 $A^k\vec{u}$. 为什么计算 $A^k\vec{u}$ 很重要? 因为它代表一种演化过程, 就好比

$$\vec{v}_n=A\vec{v}_{n-1},$$

我们知道 $\vec{v}_0=\vec{u}$, 要计算 $\vec{v}_k$. 而上面那个迭代公式, 就好比一种演化, 一种递推公式, 它可以用于求解离散时间下的方程, 也被应用于连续时间下的常微分方程求解.

可以考虑先从利息说起, 然后到连续复利, 以及求解过程. 之后再说, 小兔变大兔再生小兔子的事情.

设 n 维线性空间 V, 其基是

$$\Big\{\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n\Big\}.$$

V 上的线性变换

$$\mathcal{A}(\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n)=(\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n)A,$$

设可逆矩阵 $P\in\mathbb{R}^{n\times n}$, 若基变换

$$(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)=(\vec{\varepsilon}_1,\cdots,\vec{\varepsilon}_n)P,$$

则

$$\mathcal{A}(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)P^{-1}=(\vec{\psi}_1,\cdots,\vec{\psi}_n)P^{-1}A,$$

即

$$\mathcal{A}(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)=(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)P^{-1}AP=(\vec{\varphi}_1,\cdots,\vec{\varphi}_n)B.$$

Definition 15.1. 若存在可逆矩阵 P 使得

$$P^{-1}AP = B,$$

则称矩阵 A 与 B 是相似的.

(HW) 验证相似关系是等价关系.

变换 — 等价关系 — 不变量 — 代表元.

15.2 特征值与特征向量

特征值和特征向量的提出, 本质是希望找到一组基, 使得 $\mathcal{A}$ 在其上的表示矩阵 " 对角化".

Definition 15.2. 特征值与特征向量, 特征多项式, 特征子空间.

Definition 15.3. 特征值的代数重数与几何重数.

在 V 中取定一组基, 将

$$\mathcal{A}\vec{\xi} = \lambda\vec{\xi}$$

变成矩阵方程

$$A\vec{x} = \lambda\vec{x}.$$

其中 $\lambda \in \mathbb{P}$, 非零向量 $\vec{\xi} \in V$, 非零向量 $\vec{x} \in \mathbb{P}^n$.

寻找 λ 使得下述线性方程组

$$(\lambda E - A)\vec{x} = \vec{0}$$

有非零解. 根据 Cramer 法则, 上述方程组有非零解的充要条件是线性方程组系数矩阵对应的行列式为零, 即

$$\det(\lambda E - A) = 0.$$

若存在满足条件的 λ, 则线性方程组的解 $\vec{x}$ 即为特征值 λ 对应的特征向量.

Theorem 15.1. *若 A 和 B 相似, 则它们的特征多项式相同. 反之不然.*

Proof. 对于任意复数 λ, 下述等式恒成立:

$$\det(\lambda E - B) = \det(\lambda E - P^{-1}AP) = \det(P^{-1}(\lambda E - A)P) = \det(\lambda E - A).$$

□

(HW) 举例说明, 上述定理的反命题不成立.

Theorem 15.2. *若 A 和 B 相似, 则它们的特征值对应相等, 行列式相等, 迹相等.*

Definition 15.4. 线性变换的迹和行列式, 记为 $trace(\mathcal{A})$ 和 $det(\mathcal{A})$.

Theorem 15.3. *设 $\mathcal{A}$ 是 n 维线性空间 V 上的线性变换, $\lambda_1, \cdots, \lambda_n$ 是 $\mathcal{A}$ 的特征值 (含重特征值), 则*

$$trace(\mathcal{A}) = \sum_{i=1}^{n} \lambda_i, \ \ det(\mathcal{A}) = \prod_{i=1}^{n} \lambda_i.$$

Example 15.1 (矩阵的幂次)**.**

$$\begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}^{100}, \begin{pmatrix}1 & -1\\ 4 & 5\end{pmatrix}^{100}.$$

Example 15.2 (Fibonacci 数列)**.**

$$F_{k+2} = F_{k+1} + F_k$$

设 $\vec{u}_k = \begin{pmatrix}F_{k+1}\\ F_k\end{pmatrix}$, 则

$$\vec{u}_{k+1} = A\vec{u}_k = \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\vec{u}_k = A^k\vec{u}_1 = P^{-1}\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}^k P\vec{u}_1.$$

归纳出数列通过递推式求解通项的方法.

Example 15.3 (HW)**.** 设多项式空间 $\mathbb{P}_3[x]$ 上的求导运算 $\mathcal{D}$.

- 给出在基 $\{1, x, x^2, x^3\}$ 下 $\mathcal{D}$ 的表示矩阵 D_1;
- 给出在基 $\{1, x-1, (x-1)^2, (x-1)^3\}$ 下 $\mathcal{D}$ 的表示矩阵 D_2;
- 验证 D_1 和 D_2 是相似的;
- 给出 $Ker(\mathcal{D})$ 以及其上的一组基;
- 给出 $Im(\mathcal{D})$ 以及其上的一组基;
- 计算 $trace(\mathcal{D})$, $det(\mathcal{D})$ 和 $\mathcal{D}$ 的特征多项式;
- 验证 $\mathcal{D}$ 只有 0 特征值, 并给出对应的特征子空间以及其上的一组基. 基于这组基扩充成 $\mathbb{P}_3[x]$ 上的一组基, 并给出 $\mathbb{P}_3[x]$ 上这组基下 $\mathcal{D}$ 的表示矩阵.

Example 15.4 (HW)**.** 设 $\mathcal{A}$ 是 n 维线性空间 V 上的线性变换, 则 $\mathcal{A}$ 可逆的充要条件是

- 存在 V 上的线性变换 $\mathcal{B}$, 使得 $\mathcal{AB} = \mathcal{BA} = \mathcal{E}$;
- 存在 V 上的线性变换 $\mathcal{B}$, 使得 $\mathcal{AB} = \mathcal{E}$ 或 $\mathcal{BA} = \mathcal{E}$;
- $\mathcal{A}$ 的表示矩阵都是可逆矩阵;
- $\mathcal{A}$ 将 V 中的一组基仍映成 V 中的一组基;
- $\mathcal{A}$ 是双射;
- $\mathcal{A}$ 是单射;
- $\mathcal{A}$ 是满射;
- $Ker(\mathcal{A}) = \{\vec{0}\}$;
- $Im(\mathcal{A}) = V$;
- $det(\mathcal{A}) \neq 0$;
- $\mathcal{A}$ 没有零特征值.

Remark 15.1. 特征值与特征向量 — 特征多项式与特征子空间 — 特征值的代数重数与几何重数 — 矩阵相似的不变量 — 方阵可对角化的判定条件 — 线性变换可逆的判定条件.

15.3 可对角化

Theorem 15.4. 若线性变换 $\mathcal{A}: V \to V$ 有 n 个线性无关的特征向量, 则 $\mathcal{A}$ 在某组基下的表示矩阵为对角矩阵. 反之亦然.

设 $A \in \mathbb{P}^{2\times 2}$, 它的特征向量 $\vec{x}_1, \vec{x}_2$ 线性无关, 对应的特征值分别为 λ_1, λ_2, 则

$$A\begin{pmatrix}\vec{x}_1 & \vec{x}_2\end{pmatrix} = \begin{pmatrix}\vec{x}_1 & \vec{x}_2\end{pmatrix}\begin{pmatrix}\lambda_1 & 0 \\ 0 & \lambda_2\end{pmatrix}.$$

Remark 15.2. 如果线性变换有 n 个线性无关的特征向量, 那么以这些特征向量为线性空间的一组基, 线性变换的表示矩阵是对角矩阵且对角元是对应的特征值.

Theorem 15.5. 数域 $\mathbb{P}$ 上, 若矩阵 $A \in \mathbb{P}^{n\times n}$ 的特征多项式有 n 个不同的根, 则 A 相似于一个对角矩阵. 反之不然.

Theorem 15.6. 数域 $\mathbb{P}$ 上, 设矩阵 $A \in \mathbb{P}^{n\times n}$ 的特征多项式为

$$f(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_s)^{r_s},$$

其中 $\lambda_i \neq \lambda_j$, $i \neq j$. 那么, 若

$$r_i = n - rank(\lambda_i E - A),\ i = 1, \cdots, s$$

则 A 可对角化.

Remark 15.3. 上述定理的逆命题依然成立, 基于代数重数大于等于几何重数.

Theorem 15.7. 设矩阵 $A, B \in \mathbb{P}^{n\times n}$ 且皆可对角化, 则 A 相似于 B 等价于它们的特征多项式相等.

15.4 不变子空间

Definition 15.5. 若 $W \subset V$, $\mathcal{A}W \subset W$, 则称 W 为 $\mathcal{A}$ 的不变子空间.

Definition 15.6. 设

$$f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$$

是 $P[x]$ 中一多项式, $\mathcal{A} \in \mathcal{L}(V)$, 则称 V 上的线性变换 $f(\mathcal{A})$ 为线性变换 $\mathcal{A}$ 的多项式.

Theorem 15.8. 若线性变换 $\mathcal{A}$ 和 $\mathcal{B}$ 可交换, 则 $\mathcal{B}$ 的核空间和像空间都是 $\mathcal{A}$– 子空间.

Remark 15.4. 线性变换 $\mathcal{A}$ 和它的多项式 $f(\mathcal{A})$ 可交换, 则 $f(\mathcal{A})$ 的核空间和像空间都是 $\mathcal{A}$– 子空间.

不变子空间与线性变换的矩阵化简.

Theorem 15.9 (Hamilton–Caylay)**.** 设 $A \in \mathbb{P}^{n\times n}$, 其特征多项式为 $f(\lambda) = |\lambda E - A|$, 则

$$f(A) = A^n - (a_{11} + \cdots + a_{nn})A^{n-1} + \cdots + (-1)^n |A| E = O.$$

对 V 上的线性变换 $\mathcal{A}$ 有类似的结论.

特征多项式的因式分解与线性空间的不变子空间直和分解.

Theorem 15.10. 设线性变换 $\mathcal{A}\in\mathcal{L}(V)$, 它的特征多项式 $f(\lambda)$ 分解成一次因式的乘积

$$f(\lambda)=(\lambda-\lambda_1)^{r_1}\cdots(\lambda-\lambda_s)^{r_s},$$

其中 $\lambda_i\neq\lambda_j$, $i\neq j$, 则 V 可分解成不变子空间的直和

$$V=V_1\oplus\cdots\oplus V_s,$$

其中 $V_i=\{\vec{\xi}|(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}\vec{\xi}=\vec{0},\vec{\xi}\in V\}=Ker((\mathcal{A}-\lambda_i\mathcal{E})^{r_i})$.

Proof. 设 $W_i=\mathrm{Im}(f_i(\mathcal{A}))$, 其中 $f_i(\lambda)=f(\lambda)/(\lambda-\lambda_i)^{r_i}$.

(i) 先证 $W_i\subset V_i$:

设 $\vec{w}\in W_i$, 则存在 $\vec{u}\in V$, 使得 $f_i(\mathcal{A})\vec{u}=\vec{w}$. 所以,

$$(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}\vec{w}=(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}f_i(\mathcal{A})\vec{u}=f(\mathcal{A})\vec{u}=\vec{0}.$$

即, $\vec{w}\in V_i$.

(ii) 再证 $V=W_1+\cdots+W_s$:

由于多项式 $\{f_i(\lambda)\}_{i=1}^s$ 是互素的, 所以存在 $\{u_i(\lambda)\}_{i=1}^s$ 使得

$$\sum_{i=1}^s f_i(\lambda)u_i(\lambda)=1.$$

那么

$$\sum_{i=1}^s f_i(\mathcal{A})\circ u_i(\mathcal{A})=\mathcal{E}.$$

上述等式两边的线性变换分别作用于 $\forall\vec{\beta}\in V$, 得到

$$\sum_{i=1}^s f_i(\mathcal{A})(u_i(\mathcal{A})\vec{\beta})=\vec{\beta}.$$

设 $\vec{\beta}_i=f_i(\mathcal{A})(u_i(\mathcal{A})\vec{\beta})\in W_i$, $i=1,2,\cdots,s$. 即, $\forall\vec{\beta}\in V$ 存在 $\vec{\beta}_i\in W_i$ 使得

$$\vec{\beta}=\sum_{i=1}^s\vec{\beta}_i.$$

所以 $V=W_1+\cdots+W_s$. 又由于 $W_i\subset V_i$, 那么 $V=V_1+\cdots+V_s$.

(iii) 再证 $V_1+\cdots+V_s=V_1\oplus\cdots\oplus V_s$:

设 $\vec{\beta}_i\in V_i=\mathrm{Ker}((\mathcal{A}-\lambda_i\mathcal{E})^{r_i})$, 使得

$$\vec{0}=\sum_{i=1}^s\vec{\beta}_i.$$

上述等式两边同时作用 $f_k(\mathcal{A})$, 得到

$$\vec{0}=f_k(\mathcal{A})\vec{\beta}_k.$$

又因为 $(f_k(\lambda),(\lambda-\lambda_k)^{r_k})=1$, 所以存在 $u(\lambda),v(\lambda)$ 使得

$$u(\lambda)f_k(\lambda)+v(\lambda)(\lambda-\lambda_k)^{r_k}=1$$

则
$$u(\mathcal{A})\circ f_k(\mathcal{A})+v(\mathcal{A})\circ(\mathcal{A}-\lambda_k\mathcal{E})^{r_k}=\mathcal{E}$$
上述等式两边的线性变换分别作用于 $\vec{\beta}_k$, 得到
$$\vec{0}=u(\mathcal{A})(f_k(\mathcal{A})\vec{\beta}_k)+v(\mathcal{A})((\mathcal{A}-\lambda_k\mathcal{E})^{r_k}\vec{\beta}_k)=\vec{\beta}_k.$$
所以对于零向量的表法是唯一的, 即, $V_1+\cdots+V_s=V_1\oplus\cdots\oplus V_s$.

综合上述结论得到,
$$V=V_1\oplus\cdots\oplus V_s.$$

(iv) 最后证 $W_i=V_i$:

由于 $W_i\subset V_i$, 所以 $W_1+\cdots+W_s=W_1\oplus\cdots\oplus W_s$. 综上得到
$$V=V_1\oplus\cdots\oplus V_s=W_1\oplus\cdots\oplus W_s.$$
根据维数公式,
$$\dim(V)=\sum_{i=1}^{s}\dim(V_i)=\sum_{i=1}^{s}\dim(W_i),$$
且 $\dim(V_i)\ge\dim(W_i)$, 所以 $\dim(V_i)=\dim(W_i)$. 再由 $W_i\subset V_i$, 得到 $W_i=V_i$.

□

15.5 Jordan 标准形

Definition 15.7. *Jordan* 块
$$J(\lambda,t)=\begin{pmatrix}\lambda&0&\cdots&0&0\\1&\lambda&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&\lambda\end{pmatrix}_{t\times t},$$
Jordan 形矩阵
$$J=diag\{A_1,A_2,\cdots,A_s\}$$
其中 $A_i=J(\lambda_i,k_i)$.

Theorem 15.11. 设 $\mathcal{B}\in\mathcal{L}(V)$, $dim(V)=n$, $\mathcal{B}^k=\mathcal{O}$, $k\in\mathbb{N}^+$. 那么, 在 V 中必有下列形式的一组元素作为基
$$\begin{array}{cccc}\vec{\alpha}_1&\vec{\alpha}_2&\cdots&\vec{\alpha}_s\\\mathcal{B}\vec{\alpha}_1&\mathcal{B}\vec{\alpha}_2&\cdots&\mathcal{B}\vec{\alpha}_s\\\vdots&\vdots&&\vdots\\\mathcal{B}^{k_1-2}\vec{\alpha}_1&\mathcal{B}^{k_2-2}\vec{\alpha}_2&\cdots&\mathcal{B}^{k_s-2}\vec{\alpha}_s\\\mathcal{B}^{k_1-1}\vec{\alpha}_1&\mathcal{B}^{k_2-1}\vec{\alpha}_2&\cdots&\mathcal{B}^{k_s-1}\vec{\alpha}_s,\end{array}$$
满足 $\mathcal{B}^{k_1}\vec{\alpha}_1=\mathcal{B}^{k_2}\vec{\alpha}_2=\cdots\mathcal{B}^{k_s}\vec{\alpha}_s=\vec{0}$. 在这组基下 $\mathcal{B}$ 的表示矩阵为
$$J=diag\{J(0,k_1),J(0,k_2),\cdots,J(0,k_s)\}.$$

Proof. 对 V 的维数 n 作数学归纳法.

(i) 当 $\dim(V)=1$ 时成立. 设 $\mathcal{B}\vec{\alpha}_1=\lambda\vec{\alpha}_1$, 则

$$\vec{0}=\mathcal{O}\vec{\alpha}_1=\mathcal{B}^k\vec{\alpha}_1=\lambda^k\vec{\alpha}_1.$$

由于 $\vec{\alpha}_1\neq\vec{0}$, 所以 $\lambda^k=0$, 即 $\lambda=0$. 那么取 $\{\vec{\alpha}_1\}$ 为 V 的基, 且满足 $\mathcal{B}\vec{\alpha}_1=\vec{0}$. 即, 命题成立.

(ii) 假设 $\dim(V)<n$ 时命题成立, 下证 $\dim(V)=k$ 时命题成立.

(iii) 当 $\dim(V)=n$ 时, 易得 $\dim(\mathcal{B}V)<\dim(V)=n$, 否则与 $\mathcal{B}^k=\mathcal{O}$ 矛盾. 那么将 $\mathcal{B}$ 看作 $\dim(\mathcal{B}V)<\dim(V)$ 上的线性变换, 仍有 $\mathcal{B}^k=\mathcal{O}$. 那么利用 (ii) 对于线性子空间 $\mathcal{B}V=\mathrm{Im}\mathcal{B}$ 有命题所述形式的一组向量作为基,

$$\begin{array}{cccc}
\vec{\varepsilon}_1 & \vec{\varepsilon}_2 & \cdots & \vec{\varepsilon}_t\\
\mathcal{B}\vec{\varepsilon}_1 & \mathcal{B}\vec{\varepsilon}_2 & \cdots & \mathcal{B}\vec{\varepsilon}_t\\
\vdots & \vdots & & \vdots\\
\mathcal{B}^{k_1-2}\vec{\varepsilon}_1 & \mathcal{B}^{k_2-2}\vec{\varepsilon}_2 & \cdots & \mathcal{B}^{k_t-2}\vec{\varepsilon}_t\\
\mathcal{B}^{k_1-1}\vec{\varepsilon}_1 & \mathcal{B}^{k_2-1}\vec{\varepsilon}_2 & \cdots & \mathcal{B}^{k_t-1}\vec{\varepsilon}_t,
\end{array}$$

满足 $\mathcal{B}^{k_1}\vec{\varepsilon}_1=\mathcal{B}^{k_2}\vec{\varepsilon}_2=\cdots\mathcal{B}^{k_t}\vec{\varepsilon}_t=\vec{0}$.

由于 $\{\vec{\varepsilon}_i\}_{i=1}^t\subset\mathcal{B}V=\mathrm{Im}\mathcal{B}$, 则存在 $\{\vec{\alpha}_i\}_{i=1}^t$ 使得 $\mathcal{B}\vec{\alpha}_i=\vec{\varepsilon}_i$. 所以线性子空间 $\mathcal{B}V=\mathrm{Im}\mathcal{B}$ 上的基可以重写为

$$\begin{array}{cccc}
\mathcal{B}\vec{\alpha}_1 & \mathcal{B}\vec{\alpha}_2 & \cdots & \mathcal{B}\vec{\alpha}_t\\
\mathcal{B}^2\vec{\alpha}_1 & \mathcal{B}^2\vec{\alpha}_2 & \cdots & \mathcal{B}^2\vec{\alpha}_t\\
\vdots & \vdots & & \vdots\\
\mathcal{B}^{k_1-1}\vec{\alpha}_1 & \mathcal{B}^{k_2-1}\vec{\alpha}_2 & \cdots & \mathcal{B}^{k_t-1}\vec{\alpha}_t\\
\mathcal{B}^{k_1}\vec{\alpha}_1 & \mathcal{B}^{k_2}\vec{\alpha}_2 & \cdots & \mathcal{B}^{k_t}\vec{\alpha}_t.
\end{array}$$

由定理13.7, $\mathrm{Im}\mathcal{B}$ 的一组基的原象集合

$$\begin{array}{cccc}
\vec{\alpha}_1 & \vec{\alpha}_2 & \cdots & \vec{\alpha}_t\\
\mathcal{B}\vec{\alpha}_1 & \mathcal{B}\vec{\alpha}_2 & \cdots & \mathcal{B}\vec{\alpha}_t\\
\mathcal{B}^2\vec{\alpha}_1 & \mathcal{B}^2\vec{\alpha}_2 & \cdots & \mathcal{B}^2\vec{\alpha}_t\\
\vdots & \vdots & & \vdots\\
\mathcal{B}^{k_1-1}\vec{\alpha}_1 & \mathcal{B}^{k_2-1}\vec{\alpha}_2 & \cdots & \mathcal{B}^{k_t-1}\vec{\alpha}_t,
\end{array}$$

与 $\mathrm{Ker}\mathcal{B}$ 的一组基共同构成 V 的一组基. 然而基于 $\mathrm{Ker}\mathcal{B}$ 中线性无关的向量组

$$\mathcal{B}^{k_1}\vec{\alpha}_1,\mathcal{B}^{k_2}\vec{\alpha}_2,\cdots,\mathcal{B}^{k_t}\vec{\alpha}_t$$

可扩充为 $\mathrm{Ker}\mathcal{B}$ 的基, 即

$$\mathcal{B}^{k_1}\vec{\alpha}_1,\mathcal{B}^{k_2}\vec{\alpha}_2,\cdots,\mathcal{B}^{k_t}\vec{\alpha}_t,\vec{\alpha}_{t+1},\cdots,\vec{\alpha}_s.$$

综上, 得到 V 的一组基

$$\begin{array}{ccccccc}
\vec{\alpha}_1 & \vec{\alpha}_2 & \cdots & \vec{\alpha}_t & & & \\
\mathcal{B}\vec{\alpha}_1 & \mathcal{B}\vec{\alpha}_2 & \cdots & \mathcal{B}\vec{\alpha}_t & & & \\
\mathcal{B}^2\vec{\alpha}_1 & \mathcal{B}^2\vec{\alpha}_2 & \cdots & \mathcal{B}^2\vec{\alpha}_t & & & \\
\vdots & \vdots & & \vdots & & & \\
\mathcal{B}^{k_1-1}\vec{\alpha}_1 & \mathcal{B}^{k_2-1}\vec{\alpha}_2 & \cdots & \mathcal{B}^{k_t-1}\vec{\alpha}_t & & & \\
\mathcal{B}^{k_1}\vec{\alpha}_1 & \mathcal{B}^{k_2}\vec{\alpha}_2 & \cdots & \mathcal{B}^{k_t}\vec{\alpha}_t & \vec{\alpha}_{t+1} & \cdots & \vec{\alpha}_s,
\end{array}$$

满足 $\mathcal{B}^{k_1+1}\vec{\alpha}_1=\mathcal{B}^{k_2+1}\vec{\alpha}_2=\cdots\mathcal{B}^{k_t+1}\vec{\alpha}_t=\vec{0}$ 且 $\mathcal{B}\vec{\alpha}_{t+1}=\mathcal{B}\vec{\alpha}_s=\vec{0}$. 所以, 命题对 $\dim(V)=n$ 时成立.

另外, 注意到

$$\begin{aligned}&\mathcal{B}(\vec{\alpha}_i,\mathcal{B}\vec{\alpha}_i,\cdots,\mathcal{B}^{k_i-1}\vec{\alpha}_i)=(\mathcal{B}\vec{\alpha}_i,\cdots,\mathcal{B}^{k_i-1}\vec{\alpha}_i,\vec{0})\\ =&(\vec{\alpha}_i,\mathcal{B}\vec{\alpha}_1,\cdots,\mathcal{B}^{k_i-1}\vec{\alpha}_i)\begin{pmatrix}0&0&\cdots&0&0\\1&0&\cdots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\cdots&1&0\end{pmatrix}_{k_i\times k_i}\\ =&(\vec{\alpha}_i,\mathcal{B}\vec{\alpha}_1,\cdots,\mathcal{B}^{k_i-1}\vec{\alpha}_i)J(0,k_i).\end{aligned}$$

所以, $\mathcal{B}$ 在这组基下 $\mathcal{B}$ 的表示矩阵为

$$J=\operatorname{diag}\{J(0,k_1),J(0,k_2),\cdots,J(0,k_s)\}.$$

□

Remark 15.5. 验证对于 $i=1,2,\cdots,s$, 向量组 $\{\mathcal{B}^l\vec{\alpha}_i\}_{l=0}^{k_i-1}$ 是线性无关的.

Theorem 15.12. *设 $\mathcal{A}$ 是 $\mathbb{C}$ 上线性空间 V 的一个线性变换, 则在 V 中必存在一组基, 使得 $\mathcal{A}$ 在这组基下的表示矩阵是 Jordan 形矩阵.*

Proof. 设线性变换 $\mathcal{A}$ 的特征多项式 $f(\lambda)$ 分解成一次因式的乘积

$$f(\lambda)=(\lambda-\lambda_1)^{r_1}\cdots(\lambda-\lambda_s)^{r_s},$$

其中 $\lambda_i\neq\lambda_j$, $i\neq j$, 则 V 可分解成不变子空间的直和

$$V=V_1\oplus\cdots\oplus V_s,$$

其中 $V_i=\{\vec{\xi}|(\mathcal{A}-\lambda_i\mathcal{E})^{r_i}\vec{\xi}=\vec{0},\vec{\xi}\in V\}=\operatorname{Ker}((\mathcal{A}-\lambda_i\mathcal{E})^{r_i})$.

对任意 $i=1,2,\cdots,s$, 设 $\mathcal{B}_i=(\mathcal{A}-\lambda_i\mathcal{E})|_{V_i}\in\mathcal{L}(V_i)$, 则 $\mathcal{B}_i^{r_i}=\mathcal{O}$. 所以在 V_i 中存在一组基使得 $\mathcal{B}_i$ 的表示矩阵为

$$J_i(0)=\operatorname{diag}\{J(0,k_{i1}),J(0,k_{i2}),\cdots,J(0,k_{il_i})\}.$$

由于 V_i 是 $\mathcal{A}$-子空间, 则 $\mathcal{A}|V_i$ 在这组基下的表示矩阵为

$$J_i(\lambda_i)=J_i(0)-\lambda_iE=\operatorname{diag}\{J(\lambda_i,k_{i1}),J(\lambda_i,k_{i2}),\cdots,J(\lambda_i,k_{il_i})\}.$$

以上述方法将 V_i 的合并成 V 的基, 则 $\mathcal{A}$ 在这组基下的表示矩阵为

$$J=\operatorname{diag}\{J_1(\lambda_1),J_2(\lambda_2),\cdots,J_s(\lambda_s)\}.$$

□

Remark 15.6. $J_i(\lambda_i)$ 的阶数是 r_i, 即

$$r_i=k_{i1}+k_{i2}+\cdots+k_{il_i}.$$

另外,

$$l_i=dim(Ker(\mathcal{A}-\lambda_i\mathcal{E})).$$

所以, λ_i 的代数重数是 r_i, 几何重数是 l_i, 且 $r_i\geq l_i$.

换作矩阵的语言, 每个 n 级复矩阵都与一个 Jordan 形矩阵相似.

(HW) 以 $\mathbb{P}_n[x]$ 上的求导运算 $\mathcal{D}$ 为例, 验证上述定理.

Example 15.5. 设 n 维线性空间 V 上的线性变换 $\mathcal{A}$. 求证,

若 $rank(\mathcal{A}) = rank(\mathcal{A}^2)$, 则 $V = Im\mathcal{A} \oplus Ker\mathcal{A}$.

Proof. **法一**: 只需证明 $W = \mathrm{Im}\mathcal{A} \cap \mathrm{Ker}\mathcal{A} = \{\vec{0}\}$.

对任意 $\vec{v} \in W$, 则存在 $\vec{\alpha} \in V$ 使得 $\mathcal{A}\vec{\alpha} = \vec{v}$, $\mathcal{A}\vec{v} = \vec{0}$. 那么 $\vec{\alpha} \in \mathrm{Ker}\mathcal{A}^2$. 即, 只需证明 $\mathrm{Ker}\mathcal{A}^2 = \mathrm{Ker}\mathcal{A}$, 就可以得到 $\vec{v} = \vec{0}$. (可利用齐次线性方程组 $A\vec{x} = \vec{0}$ 与 $A^2\vec{x} = \vec{0}$ 同解得到)

法二: 设 V 中存在一组基使得 $\mathcal{A}$ 在这组基下的表示矩阵是 Jordan 形矩阵

$$J = \mathrm{diag}\{J_1(\lambda_1), \cdots, J_{s-1}(\lambda_{s-1}), J_s(\lambda_s)\}$$

满足 $\{\lambda_i\}_{i=1}^s$ 互不相同且 $\lambda_s = 0$. 若无零特征值, 则 $\mathrm{Ker}(\mathcal{A}) = \{\vec{0}\}$, $\mathrm{Im}(\mathcal{A}) = V$. 命题成立.

因为 $\mathrm{rank}(J) = \mathrm{rank}(\mathcal{A}) = \mathrm{rank}(\mathcal{A}^2) = \mathrm{rank}(J^2)$, 则 $J_s(\lambda_s)$ 只能是零矩阵. 设

$$J = \begin{pmatrix} K & O \\ O & O \end{pmatrix}$$

其中 $K = \mathrm{diag}\{J_1(\lambda_1), \cdots, J_{s-1}(\lambda_{s-1})\}$ 是 r 阶可逆矩阵. 另外, V 中的一组基 $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r, \vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n\}$ 使得

$$\mathcal{A}(\vec{\alpha}_1, \cdots, \vec{\alpha}_r, \vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n) = (\vec{\alpha}_1, \cdots, \vec{\alpha}_r, \vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n)J.$$

所以

$$\mathcal{A}(\vec{\alpha}_1, \cdots, \vec{\alpha}_r, \vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n) = (\vec{\alpha}_1, \cdots, \vec{\alpha}_r) \begin{pmatrix} K & O \end{pmatrix},$$

$$\mathcal{A}(\vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n) = (\vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n)O.$$

即, $\{\vec{\alpha}_1, \cdots, \vec{\alpha}_r\}$ 是 $\mathrm{Im}(\mathcal{A})$ 的基, $\{\vec{\alpha}_{r+1}, \cdots, \vec{\alpha}_n\}$ 是 $\mathrm{Ker}(\mathcal{A})$ 的基.

综上,

$$V = \mathrm{Im}(\mathcal{A}) \oplus \mathrm{Ker}(\mathcal{A}).$$

□

Example 15.6. 设 $\mathcal{A} \in \mathcal{L}(V)$, 且 $dim(V) = n$, $\mathcal{A}$ 的特征多项式为

$$f(\lambda) = \prod_{i=1}^{s} (\lambda - \lambda_i)^{r_i},\ \lambda_i \neq \lambda_j,\ for\ i \neq j.$$

已知对任意 $i = 1, 2, \cdots, s$ 成立

$$dim(Ker(\mathcal{A} - \lambda_i\mathcal{E})^k) = m_{ik},\ for\ k = 1, 2, \cdots, r_i.$$

尝试给出 φ 的 *Jordan* 形表示矩阵 J.

Proof. 设 n_{il} 表示 J 中包含 Jordan 块 $J(\lambda_i, l)$ 的个数, 则

$$\begin{aligned}
n_{i1} &= m_{i1} - (m_{i2} - m_{i1}) = 2m_{i1} - m_{i2}, \\
n_{i2} &= (m_{i2} - m_{i1}) - (m_{i3} - m_{i2}) = 2m_{i2} - m_{i1} - m_{i3}, \\
n_{i3} &= (m_{i3} - m_{i2}) - (m_{i4} - m_{i3}) = 2m_{i3} - m_{i2} - m_{i4},
\end{aligned}$$

$$\vdots$$
$$n_{il} = (m_{i,l} - m_{i,l-1}) - (m_{i,l+1} - m_{i,l}) = 2m_{i,l} - m_{i,l-1} - m_{i,l+1},$$
$$\vdots$$
$$n_{ir_i} = (m_{i,r_i} - m_{i,r_i-1}) - (m_{i,r_i+1} - m_{i,r_i}) = m_{i,r_i} - m_{i,r_i-1}.$$

□

15.6 λ–矩阵相抵关系的完全不变量

Definition 15.8. λ–矩阵的加法, 乘法, 数乘, 秩, 逆, 行列式, 初等变换, 初等矩阵.

Remark 15.7. λ–矩阵的逆是数字矩阵逆的直接推广, 但是它与行列式的关系有所不同. λ–矩阵的秩与相抵关系是数字矩阵秩与相抵关系的直接定义, 但代表元有所不同.

Definition 15.9. 称 λ–矩阵 $A(\lambda)$ 与 $B(\lambda)$ 是相抵的, 若可以经过一系列初等变换将 $A(\lambda)$ 化为 $B(\lambda)$. λ–矩阵的标准形.

Theorem 15.13. 设 $A(\lambda)$ 是一个 n 阶 λ–矩阵, 则 $A(\lambda)$ 相抵于唯一的 λ– 矩阵的标准形

$$D_r(\lambda) = diag\{d_1(\lambda), \cdots, d_r(\lambda); 0, \cdots, 0\},$$

其中 $d_i(\lambda)$ 是非零首一多项式且 $d_i(\lambda)|d_{i+1}(\lambda)$, $i = 1, \cdots, r-1$.

Definition 15.10. λ–矩阵的行列式因子和不变因子.

Theorem 15.14. λ–矩阵的行列式因子或不变因子是 λ–矩阵相抵关系的完全不变量.

15.7 矩阵相似关系的完全不变量

Theorem 15.15. 矩阵 A 与 B 相似的充要条件是 λ– 矩阵 $\lambda E - A$ 与 $\lambda E - B$ 相抵.

Theorem 15.16. λ–矩阵 $\lambda E - A$ 的行列式因子, 不变因子, 初等因子都是矩阵相似的完全不变量.

15.8 初等因子与 Jordan 标准形的计算

给定初等因子组, 可以唯一确定不变因子组. 反之亦然.

每一个初等因子 $(\lambda - \lambda_i)^{k_i}$ 对应一个 Jordan 块 $J(\lambda_i, k_i)$.

15.9 Jordan 块 – 循环子空间

- 同一特征值对应的循环子空间的个数是其特征值对应的几何重数;
- 同一特征值对应的循环子空间的直和形成根子空间, 根子空间的维数是其特征值对应的代数重数;
- 每个循环子空间对应一个初等因子, 对应一个 Jordan 块, 对应一个特征向量.

$$\mathcal{A}: V \to V$$

在一组基下的表示矩阵为 Jordan 形矩阵:

$$\mathcal{A}(\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n) = (\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_n) \begin{pmatrix} J(\lambda_1, k_1) & & & \\ & J(\lambda_2, k_2) & & \\ & & \ddots & \\ & & & J(\lambda_t, k_t) \end{pmatrix}.$$

以 λ_1 为例,

$$\mathcal{A}\vec{\varepsilon}_1 = \lambda_1\vec{\varepsilon}_1 + \vec{\varepsilon}_2,\ \cdots,\ \mathcal{A}\vec{\varepsilon}_{k_1-1} = \lambda_1\vec{\varepsilon}_{k_1-1} + \vec{\varepsilon}_{k_1},\ \mathcal{A}\vec{\varepsilon}_{k_1} = \lambda_1\vec{\varepsilon}_{k_1},\ \cdots$$

即

$$(\mathcal{A} - \lambda_1\mathcal{E})\vec{\varepsilon}_1 = \vec{\varepsilon}_2,\ \cdots,\ (\mathcal{A} - \lambda_1\mathcal{E})\vec{\varepsilon}_{k_1-1} = \vec{\varepsilon}_{k_1},\ (\mathcal{A} - \lambda_1\mathcal{E})\vec{\varepsilon}_{k_1} = \vec{0},\ \cdots$$

$\vec{\varepsilon}_1, \cdots, \vec{\varepsilon}_{k_1}$ 张成一个 λ_1 的循环子空间 U_1 (是不变子空间), 对应一个 λ_1 的特征向量 $\vec{\varepsilon}_{k_1}$, 对应一个初等因子 $(\lambda - \lambda_1)^{k_1}$, 对应一个 Jordan 块 $J(\lambda_1, k_1)$. 那么,

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_t.$$

将 λ_1 的所有循环子空间直和得到 $V_1 = \mathrm{Ker}((\lambda_1\mathcal{E} - \mathcal{A})^{r_1})$. 其中, $\mathcal{A}$ 的特征多项式的一次因式分解为

$$f(\lambda) = (\lambda - \lambda_1)^{r_1} \cdots (\lambda - \lambda_s)^{r_s}.$$

λ_1 的所有循环子空间的维数和等于 r_1.

Remark 15.8. $\lambda_1, \cdots, \lambda_s$ 是互不相同的 s 个特征值. $\lambda_1, \cdots, \lambda_t$ 是 t 个特征值, 可以有相同的特征值.

(HW) 计算 Jordan 块 $J(\lambda_i, k_i)$ 的不变因子和初等因子.

Example 15.7. $f_1(x) = xe^x,\ f_2(x) = e^x,\ f_3(x) = xe^{-x},\ f_4(x) = e^{-x}$.

$$\mathcal{D}(f_1, f_2, f_3, f_4) = (f_1, f_2, f_3, f_4) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Remark 15.9. *λ–矩阵的相抵关系 — 特征矩阵 $\lambda E - A$ 的标准形 — 数字矩阵相似关系的完全不变量— 不变因子 — 初等因子 — Jordan 块 — Jordan 形矩阵.*

16 等价关系: 合同

线性代数的核心思想是分类与整理.

如何分类, 首先要给出一个等价关系. 那么, 何来等价关系, 先从定义变换开始. 分类之后如何整理, 每一类给出一个代表元, 并给出不变量. 不变量可以理解为一个标签, 同类的标签相同, 不同类的标签互不相同. 不变量使分类变得容易. 总结起来: 变换 - 等价关系 - 分类 - 代表元 - 不变量. 我们以这样的逻辑研究二次型.

首先, 给出二次型及其相关定义.

Definition 16.1. *称数域 $\mathbb{P}$ 上的 n 元二次齐次多项式*

$$f(x_1,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\vec{x}^TA\vec{x},$$

为数域 $\mathbb{P}$ 上的 n 元二次型, 简称二次型. 其中 $\vec{x}=(x_1,\cdots,x_n)^T$, 并称 $A=(a_{ij})_{n\times n}$ 为二次型 $f(x_1,\cdots,x_n)$ 的矩阵, 称矩阵 A 的秩 $rank(A)$ 为二次型 $f(x_1,\cdots,x_n)$ 的秩.

给出了我们要研究的对象, 接下来, 我们给出作用于这些对象的变换.

Definition 16.2. *线性替换, 可逆的线性替换.*

$$f(x_1,\cdots,x_n)=\vec{x}^TA\vec{x}=\vec{y}^TB\vec{y},$$

其中 $\vec{x}=C\vec{y}$, $B=C^TAC$.

Definition 16.3. *称 n 阶对称矩阵 A 和 B 是合同的, 若存在可逆矩阵 C 使得 $C^TAC=B$.*

Theorem 16.1. *合同关系是等价关系.*

在描述剩下的概念之前, 先围观一下我们曾经见过的与二次型相关的东西. 在二维空间中,

- $x^2-y^2=1$ (红色虚线); $7/4x^2-5/2xy+3/4y^2=1$ (红色实线);
- $x^2+0y^2=1$ (绿色虚线); $4x^2-4xy+y^2=1$ (绿色实线);
- $x^2+y^2=1$ (蓝色虚线); $25/4x^2-11/2xy+5/4y^2=1$ (蓝色实线).

右图其实是左图旋转拉伸之后的结果. 对于左图, 我们可以轻易地将曲线与方程做对应. 也就是说, 我们会非常自然地写出这些几何曲线对应的代数方程. 对于右图, 事情就没有那么简单. 但是, 你也会认得右图中蓝色实线是椭圆. 那么, 会不会存在一个坐标变换, 将左图中的椭圆变成右图中的椭圆呢?

$$\begin{aligned}
x^2+y^2=1 &\Leftrightarrow \begin{pmatrix} x & y\end{pmatrix}\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=1\\
&\Leftrightarrow \begin{pmatrix} x & y\end{pmatrix}P^T\begin{pmatrix}25/4&-11/4\\-11/4&5/4\end{pmatrix}P\begin{pmatrix}x\\y\end{pmatrix}=1\\
&\Leftrightarrow \begin{pmatrix} w & z\end{pmatrix}\begin{pmatrix}25/4&-11/4\\-11/4&5/4\end{pmatrix}\begin{pmatrix}w\\z\end{pmatrix}=1\\
&\Leftrightarrow 25/4w^2-11/2wz+5/4z^2=1,
\end{aligned}$$

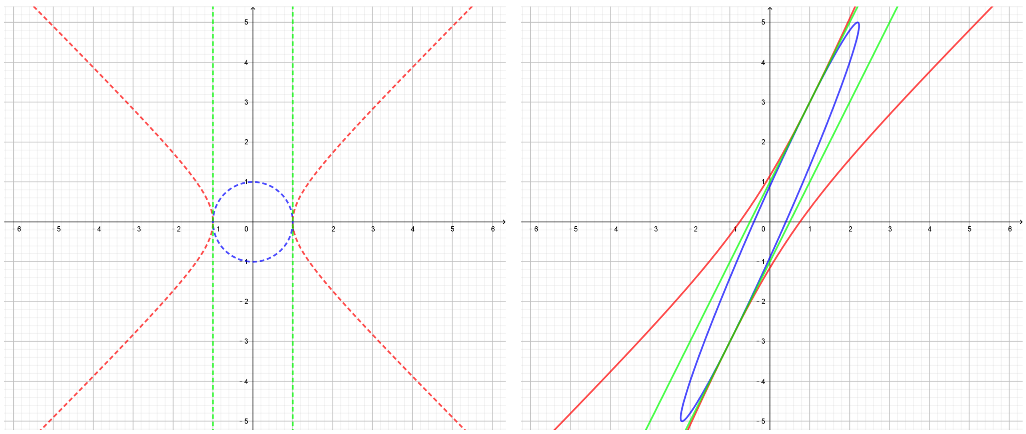


Figure 16.1: 二次曲线

其中,

$$P=\begin{pmatrix}1&2\\3&4\end{pmatrix},\ \begin{pmatrix}w\\z\end{pmatrix}=P\begin{pmatrix}x\\y\end{pmatrix}.$$

这么说来, 左右两图之间只差一个可逆的线性替换. 换句话讲, 由于坐标系的不同, 原本一个图形呈现出两种不同的表象 (无论从几何曲线还是代数方程).

那么, 二次型 $f_1(x,y)=x^2+y^2=1$ 与 $f_2(x,y)=25/4x^2-11/2xy+5/4y^2=1$ 是等价的, 或者说, 是属于同一类的. 从几何上看, 即 $f_1(x,y)=1$ 与 $f_2(x,y)=1$ 都是椭圆. 同样, 可以对其它两类曲线进行讨论.

我们已经简述了变换和等价关系, 之后来说说分类. 对于我们研究的二次型来说, 我们希望知道合同关系将二次型分成了几类. 先以 2 元二次型为例, 下图展示了二次型 $f(x,y)=x^2+ay^2$ 对应的二次曲线 $x^2+ay^2=1$ 随 a 的变化而变化的过程 (下图也可以理解为, 三维曲面 $z=f(x,y)=x^2+ay^2$ 与平面 $z=1$ 的交线).

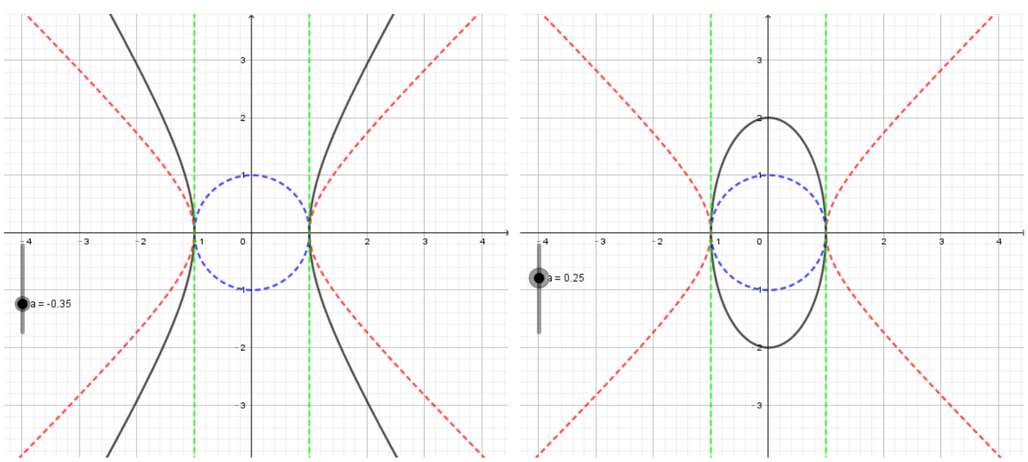


Figure 16.2: $x^2+ay^2=1$ (黑色), a 从 -1 (红色) 变化到 0 (绿色) 再到 $+1$ (蓝色).

容易发现, 当 a 从负数变为零, 再从零变为正数时, 曲线会发生巨大的变化. 当 $a<0$ 时, 黑色曲线是双曲线. 当 $a=0$ 时, 双曲线瞬间变成了两条直线. 当 $a>0$ 时, 两条直线瞬间连接在了一起

变成了椭圆. 无论 a 等于哪个正数, 黑色曲线依然是椭圆. 无论 a 等于哪个负数, 黑色曲线依然是双曲线. 这两种情况的临界情况, $a=0$, 黑色曲线是两条直线. 换句话说, a 或负或零或正, 是为二次曲线 $x^2+ay^2=1$ 分类的关键. 然而, "或负或零或正" 就是研究 n 元二次型分类的要诀.

Theorem 16.2 (惯性定理)**.** 设 n 阶矩阵 A, $rank(A)=r$.

- 若 A 是实对称矩阵, 则必唯一合同于如下形式的对角矩阵
$$diag\{1,\cdots,1;-1,\cdots,-1;0,\cdots,0\},$$
  其中有 p 个 1, q 个 -1, $n-r$ 个 0.
- 若 A 是复对称矩阵, 则必唯一合同于如下形式的对角矩阵
$$diag\{1,\cdots,1;0,\cdots,0\},$$
  其中有 r 个 1, $n-r$ 个 0.

称 p 为正惯性指数, q 为负惯性指数, $p-q$ 为符号差. 称上述对角线上只有 -1, 0, 1 的对角矩阵为标准形.

Proof. **法一**: 对二次型进行线性替换.

以三元二次型为例, 假设三元实二次型 $f(x_1,x_2,x_3)$ 通过不同的非退化线性替换转化为规范形 $z_1^2-z_2^2-z_3^2$ 和 $y_1^2+y_2^2-y_3^2$, 即
$$f(x_1,x_2,x_3)=z_1^2-z_2^2-z_3^2=y_1^2+y_2^2-y_3^2,$$
其中
$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=C\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=P\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}.$$
所以
$$\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=C^{-1}P\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=G\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}.$$
其中 $G=(g_{ij})_{3\times3}=C^{-1}P$.

另外, 下述关于 (y_1,y_2,y_3) 的线性方程组
$$\begin{pmatrix}0\\g_{21}y_1+g_{22}y_2\\g_{31}y_1+g_{32}y_2\end{pmatrix}=G\begin{pmatrix}y_1\\y_2\\0\end{pmatrix}\Leftrightarrow\begin{pmatrix}0\\0\end{pmatrix}=\begin{pmatrix}g_{11}y_1+g_{12}y_2\\y_3\end{pmatrix}$$
必然非零解 $(k_1,k_2,0)$. 那么, 取
$$\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix}=\begin{pmatrix}k_1\\k_2\\k_3\end{pmatrix},\quad\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix}=\begin{pmatrix}0\\g_{21}k_1+g_{22}k_2\\g_{31}k_1+g_{32}k_2\end{pmatrix},$$

满足

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = G \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

然而

$$0 \geq z_1^2 - z_2^2 - z_3^2 = y_1^2 + y_2^2 - y_3^2 > 0,$$

矛盾.

法二: 对二次型的矩阵进行初等变换. □

Remark 16.1. 基于惯性定理, 我们只需对标准形进行分类, 就可以得到对所有二次型的分类.

Example 16.1. 将下列二次型化为对角型:

- $f(x_1, x_2, x_3) = x_1^2 - 3x_2^2 - 2x_1x_2 + 2x_1x_3 - 6x_2x_3$;
- $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3$.

那么, 不变量就一目了然了.

Theorem 16.3. *秩 r 与符号差 $p - q = p - (r - p) = 2p - r$ 是实对称矩阵合同关系的完全不变量. 秩 r 是复对称矩阵合同关系的完全不变量.*

Example 16.2. 对于 2 元二次型 $f(x_1, x_2)$, 它的标准形必然是下面情形之一:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

换句话讲, 基于合同关系, 2 元二次型 $f(x_1, x_2)$ 被分成了 6 类.

Example 16.3. 在三维空间中:

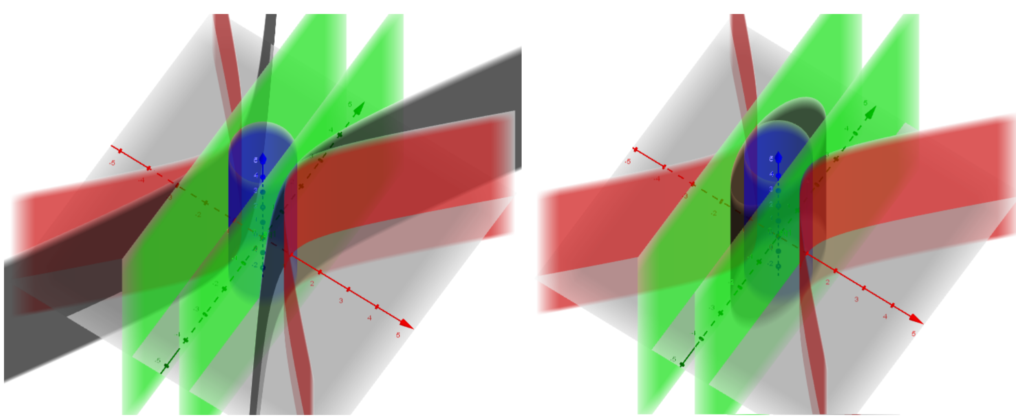


Figure 16.3: $x^2 + ay^2 + 0z^2 = 1$ (黑色), a 从 -1 (红色) 变化到 0 (绿色) 再到 $+1$ (蓝色).

Remark 16.2. 变换 *(二次型的线性替换)* — 等价关系 *(对称矩阵的合同关系)* — 不变量 *(正负惯性指数和秩)* — 代表元 *(标准形矩阵)*.

Definition 16.4. 正定矩阵, 负定矩阵, 半正定矩阵.

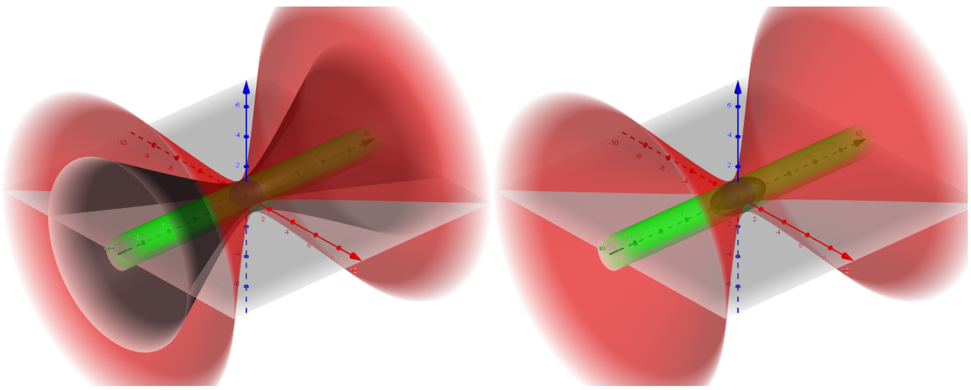


Figure 16.4: $x^2 + ay^2 + z^2 = 1$ (黑色), a 从 -1 (红色) 变化到 0 (绿色) 再到 $+1$ (蓝色).

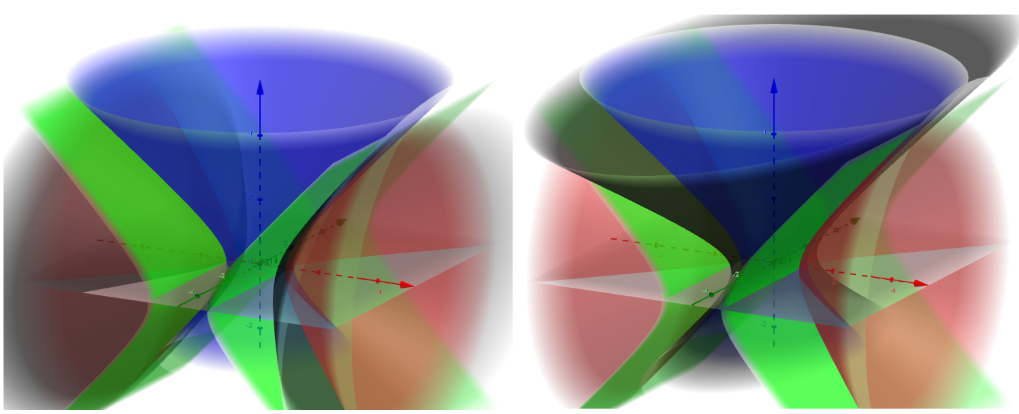


Figure 16.5: $x^2 + ay^2 - z^2 = 1$ (黑色), a 从 -1 (红色) 变化到 0 (绿色) 再到 $+1$ (蓝色).

Theorem 16.4. 设 n 阶实对称矩阵 A, 则下述命题等价:

- A 是正定矩阵;
- A 的正惯性指数 p 等于 n;
- A 的 n 个顺序主子式全大于零;
- A 合同于 n 阶单位矩阵 E_n;
- 存在 n 阶实可逆矩阵 C, 使得 $A = C^TC$.

Theorem 16.5. 设 n 阶实对称矩阵 A, 则下述命题等价:

- A 是半正定矩阵;
- A 的正惯性指数 p 等于 A 的秩 r;
- A 的 n 个顺序主子式全大于等于零;
- A 合同于 n 阶单位矩阵 $D_r = diag(E_r, O)$;
- 存在 n 阶实矩阵 C, 使得 $A = C^TC$.

Example 16.4. 设 A 是实对称正定矩阵, $\vec{\alpha}$ 是实向量, 求证: $A + \vec{\alpha}\vec{\alpha}^T$ 也是对称正定矩阵.

Proof. 构造矩阵
$$B = \begin{pmatrix} C \\ \vec{\alpha}^T \end{pmatrix}, \text{ and } B^TB = C^TC + \vec{\alpha}\vec{\alpha}^T,$$
其中 $A = C^TC$. 又因为 $\text{rank}(B) = \text{rank}(B^TB)$, 则 $A + \vec{\alpha}\vec{\alpha}^T$ 可逆. 容易证明 $A + \vec{\alpha}\vec{\alpha}^T$ 是对称半正定矩阵, 则它为对称正定矩阵. □

Example 16.5. 设实对称正定矩阵
$$M = \begin{pmatrix} A & F \\ F^T & B \end{pmatrix},$$
求证:
$$N = \begin{pmatrix} A & -F \\ -F^T & B \end{pmatrix},$$
也是实对称正定矩阵.

Proof. 法一:
$$\begin{pmatrix} C & O \\ O^T & 1 \end{pmatrix}^T \begin{pmatrix} A & F \\ F^T & B \end{pmatrix} \begin{pmatrix} C & O \\ O^T & 1 \end{pmatrix} = \begin{pmatrix} E & C^TF \\ F^TC & B \end{pmatrix},$$
以及
$$\begin{pmatrix} E & -C^TF \\ O^T & 1 \end{pmatrix}^T \begin{pmatrix} E & C^TF \\ F^TC & B \end{pmatrix} \begin{pmatrix} E & -C^TF \\ O^T & 1 \end{pmatrix} = \begin{pmatrix} E & O \\ O^T & B - F^TCC^TF \end{pmatrix}.$$

法二:
$$\begin{pmatrix} E & O \\ O^T & -E \end{pmatrix} N \begin{pmatrix} E & O \\ O^T & -E \end{pmatrix} = M.$$
□

Example 16.6. 设

$$\vec{\alpha}=\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix},\ \vec{\beta}=\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix},$$

则

$$\vec{\alpha}\vec{\beta}^T=\begin{pmatrix}a_1b_1 & a_1b_2 & a_1b_3\\a_2b_1 & a_2b_2 & a_2b_3\\a_3b_1 & a_3b_2 & a_3b_3\end{pmatrix},\ \vec{\alpha}^T\vec{\beta}=a_1b_1+a_2b_2+a_3b_3$$

Example 16.7. 设 A 是 3 阶方阵. 求证:

$$rank(A)=1\Leftrightarrow \text{存在非零向量}\vec{\alpha},\vec{\beta}\in\mathbb{R}^3,\text{使得}A=\vec{\alpha}\vec{\beta}^T.$$

Proof. "$\Rightarrow$": 设

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

则 A 的列向量组

$$\vec{\gamma}_1=\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\end{pmatrix},\ \vec{\gamma}_2=\begin{pmatrix}a_{12}\\a_{22}\\a_{32}\end{pmatrix},\ \vec{\gamma}_3=\begin{pmatrix}a_{13}\\a_{23}\\a_{33}\end{pmatrix},$$

的秩为 1. 不妨设 $\{\vec{\gamma}_1\}$ 为其极大线性无关组, 那么, 向量组 $\{\vec{\gamma}_1\}$ 可以线性表出 $\vec{\gamma}_2$ 和 $\vec{\gamma}_3$, 即, 存在 a 和 b 使得

$$\vec{\gamma}_2=a\vec{\gamma}_1,\ \vec{\gamma}_3=b\vec{\gamma}_1,$$

即,

$$A=\begin{pmatrix}\vec{\gamma}_1 & a\vec{\gamma}_1 & b\vec{\gamma}_1\end{pmatrix}=\vec{\gamma}_1\begin{pmatrix}1 & a & b\end{pmatrix}.$$

那么, 取

$$\vec{\alpha}=\begin{pmatrix}a_{11}\\a_{21}\\a_{31}\end{pmatrix},\ \vec{\beta}=\begin{pmatrix}1\\a\\b\end{pmatrix}.$$

"$\Leftarrow$": 若存在非零向量 $\vec{\alpha}$ 和 $\vec{\beta}$, 使得 $A=\vec{\alpha}\vec{\beta}^T$. 将向量 $\vec{\alpha}$ 和 $\vec{\beta}$ 看作 3 行 1 列的矩阵, 则

$$0\le \mathrm{rank}(A)\le\min\{\mathrm{rank}(\vec{\alpha}),\mathrm{rank}(\vec{\beta})\}=1.$$

同时, A 必不是零矩阵, 则 $\mathrm{rank}(A)=1$. □

Example 16.8. 设 A 是 3 阶对称矩阵且其特征值均大于等于零. 求证:

$$rank(A)=2\Leftrightarrow \text{存在非零向量}\vec{\alpha},\vec{\beta}\in\mathbb{R}^3\text{且}\vec{\alpha}^T\vec{\beta}=0,\text{使得}A=\vec{\alpha}\vec{\alpha}^T+\vec{\beta}\vec{\beta}^T.$$

Proof.

$$A=Q^T\Lambda Q,\ Q^T=Q^{-1}.$$

注意到

$$\Lambda=\begin{pmatrix}g & 0 & 0\\0 & t & 0\\0 & 0 & k\end{pmatrix}=g\begin{pmatrix}1\\0\\0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\end{pmatrix}+t\begin{pmatrix}0\\1\\0\end{pmatrix}\begin{pmatrix}0 & 1 & 0\end{pmatrix}+k\begin{pmatrix}0\\0\\1\end{pmatrix}\begin{pmatrix}0 & 0 & 1\end{pmatrix},\ g,t,k\ge 0.$$

那么

$$A = gQ^T \begin{pmatrix}1\\0\\0\end{pmatrix} \begin{pmatrix}1 & 0 & 0\end{pmatrix} Q + tQ^T \begin{pmatrix}0\\1\\0\end{pmatrix} \begin{pmatrix}0 & 1 & 0\end{pmatrix} Q + kQ^T \begin{pmatrix}0\\0\\1\end{pmatrix} \begin{pmatrix}0 & 0 & 1\end{pmatrix} Q.$$

“$\Rightarrow$”:

若 $\text{rank}(A) = 2$, 则 g, t, k 必只有一个为零. 不妨设 $g = 0$, 则取

$$\vec{\alpha} = Q^T \begin{pmatrix}0\\\sqrt{t}\\0\end{pmatrix},\ \vec{\beta} = Q^T \begin{pmatrix}0\\0\\\sqrt{k}\end{pmatrix}$$

满足条件.

“$\Leftarrow$”:

注意到

$$A\vec{\alpha} = (\vec{\alpha}^T\vec{\alpha})\vec{\alpha},\ A\vec{\beta} = (\vec{\beta}^T\vec{\beta})\vec{\beta}.$$

那么, 由于 $\vec{\alpha}$, $\vec{\beta}$ 皆为非零向量, 则 $\vec{\alpha}^T\vec{\alpha}$ 和 $\vec{\beta}^T\vec{\beta}$ 是 A 的特征值且皆不为零, 对应的特征向量是 $\vec{\alpha}$ 和 $\vec{\beta}$.

另外, 必存在 $\vec{\gamma} \in \mathbb{R}^3$ 既与 $\vec{\alpha}$ 垂直, 也与 $\vec{\beta}$ 垂直. 那么,

$$A\vec{\gamma} = 0 = 0\vec{\gamma}.$$

所以, 0 是 A 的第三个特征值. 那么, $\text{rank}(A) = 2$.

□

Example 16.9. 设 A 是 3 阶对称矩阵且其特征值均大于等于零, 求证:

$$rank(A) = 1 \Leftrightarrow 存在非零向量\vec{\alpha} \in \mathbb{R}^3, 使得A = \vec{\alpha}\vec{\alpha}^T.$$

Example 16.10. 设 A 是 3 阶对称矩阵, 且 $rank(A) = 1$. 求证: *(1)* A 相似于对角矩阵

$$\begin{pmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & k\end{pmatrix},$$

即, A 的特征值为 0, 0, k, 其中 $k \neq 0$; *(2)* 矩阵 A 的非零列向量均是特征值 k 的特征向量.

Example 16.11. 设 A, B 是 3 阶实对称矩阵, 且 $AB = BA = O$, $A^2 + A = O$, $B^2 + B = O$. 求证:

(1) A 只有两个不同的特征值, 分别是 0 和 -1;

(3) B 关于特征值 -1 的特征向量可以由 B 的列向量组线性表出;

(2) B 的每一非零列向量均是 A 关于特征值 0 的特征向量;

(4) A 关于 -1 的一个特征向量和 B 关于 -1 的一个特征向量形成的向量组线性无关.

Proof. (1) 反证法. 若 A 有其它特征值 a 且 $a \neq 0$, $a \neq -1$. 设 a 对应的特征向量为 $\vec{\alpha}$, 则

$$A\vec{\alpha} = a\vec{\alpha}.$$

那么,

$$\vec{0} = O\vec{\alpha} = (A^2 + A)\vec{\alpha} = (a^2 + a)\vec{\alpha}$$

然而 $a^2 + a \neq 0$ 且 $\vec{\alpha} \neq \vec{0}$, 与上式矛盾. 所以, A 只有两个不同的特征值, 分别是 0 和 -1;

(2) 设

$$B = \begin{pmatrix} \vec{\beta}_1 & \vec{\beta}_2 & \vec{\beta}_3 \end{pmatrix}$$

关于 -1 的特征向量为 $\vec{\xi}_2 = (s_1, s_2, s_3)^T$, 即

$$\begin{pmatrix} \vec{\beta}_1 & \vec{\beta}_2 & \vec{\beta}_3 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = B\vec{\xi}_2 = (-1)\vec{\xi}_2.$$

整理得到,

$$\vec{\xi}_2 = -s_1\vec{\beta}_1 - s_2\vec{\beta}_2 - s_3\vec{\beta}_3.$$

所以, B 关于特征值 -1 的特征向量可以由 B 的列向量组线性表出.

(3) 由于 $AB = O$, 则

$$\begin{pmatrix} A\vec{\beta}_1 & A\vec{\beta}_2 & A\vec{\beta}_3 \end{pmatrix} = A\begin{pmatrix} \vec{\beta}_1 & \vec{\beta}_2 & \vec{\beta}_3 \end{pmatrix} = AB = O = \begin{pmatrix} \vec{0} & \vec{0} & \vec{0} \end{pmatrix}.$$

那么, $i = 1, 2, 3$,

$$A\vec{\beta}_i = \vec{0} = 0\vec{\beta}_i.$$

所以, B 的每一非零列向量均是 A 关于特征值 0 的特征向量.

(4) 设

$$A\vec{\xi}_1 = (-1)\vec{\xi}_1,\ B\vec{\xi}_2 = (-1)\vec{\xi}_2,$$

且由 (2) 知

$$\vec{\xi}_2 = -s_1\vec{\beta}_1 - s_2\vec{\beta}_2 - s_3\vec{\beta}_3,$$

由 (3) 知

$$A\vec{\xi}_2 = A(-s_1\vec{\beta}_1 - s_2\vec{\beta}_2 - s_3\vec{\beta}_3) = -s_1A\vec{\beta}_1 - s_2A\vec{\beta}_2 - s_3A\vec{\beta}_3 = \vec{0},$$

即, $\vec{\xi}_2$ 是 A 关于 0 的特征向量.

又由于 A 的不同特征值对应的特征向量线性无关, 得到 $\vec{\xi}_1$ 和 $\vec{\xi}_2$ 线性无关. 那么, A 关于 -1 的一个特征向量 $\vec{\xi}_1$ 和 B 关于 -1 的一个特征向量 $\vec{\xi}_2$ 形成的向量组线性无关.

□

Remark 16.3. 二次型 — 线性替换 — 合同关系 — 标准形矩阵 — 正定矩阵.

知识能使你增加一双眼睛

17 欧式空间 – 让生活大不同

内积让世界更美好.

- **有了内积, 我们是如何拥有长度和角度的?**
- **有了内积, 对于线性空间中的基提出什么新要求, 所对应的基变换又有什么新特性?**
- **有了内积, 同构关系有什么新要求?**
- **有了内积, 基扩充定理有什么变化?**
- **有了内积, 提出了怎样新的线性变换?**
- **有了内积, 矩阵的合同和相似如何统一起来?**

Definition 17.1. *欧式空间是赋予内积的线性空间. 基于内积定义距离. 基于内积和距离定义夹角. 基于夹角定义正交 (垂直).*

内积与度量矩阵.

标准正交基与正交矩阵.

Schimidt 正交化过程.

欧式空间的同构关系.

正交变换与正交矩阵.

子空间直和与正交补.

实对称矩阵的标准形 (合同关系与相似关系).

Example 17.1. 设 $f_1(x) = \sin(x)$, $f_2(x) = \cos(x)$, $f_3(x) = \sin(2x)$, $f_4(x) = \cos(2x)$ 是 $[-\pi, \pi]$ 上的实值函数, 记由这四个函数在复数域上张成的线性空间为 V, 即

$$V = \{f : [0, 2\pi] \to \mathbb{R} | f = \sum_{i=1}^{4} k_i f_i, \forall k_i \in \mathbb{C}\},$$

对应的加法和数乘是函数之间的加法和数乘. 在 V 上赋予内积

$$(f_i, f_j) = \int_{-\pi}^{\pi} f_i(x) f_j(x)\, dx,$$

则 V 是欧式空间.

- 验证 $\{f_1, f_2, f_3, f_4\}$ 是 V 中的一组基;
- 给出 V 的一组标准正交基 $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$;
- 分别写出基于基 $\{f_1, f_2, f_3, f_4\}$ 和 $\{\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4\}$ 的度量矩阵, 并验证这两个度量矩阵是合同的.

Example 17.2. 以概率空间为例, 说明欧式空间.

设样本集合 $\Omega = \{\omega_1, \omega_2, \omega_3\}$. 且发生 ω_i 的概率为 $p_i > 0$, $i = 1, 2, 3$, 且 $p_1 + p_2 + p_3 = 1$.

设随机变量

$$X : \Omega \to \mathbb{R}$$

$$\omega_i \to x_i$$

换句话讲, 我们可以利用一个三维向量

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

表达随机变量 X, 其中 $x_i \in \mathbb{R}$ 表示发生 ω_i 时, 随机变量 X 所取的值. 同样, 设另外一个随机变量

$$\begin{aligned} Y : \Omega &\to \mathbb{R} \\ \omega_i &\to y_i \end{aligned}$$

并以一个三维向量

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

表达随机变量 Y.

我们将为包含所有这样的随机变量的集合, 赋予加法和数乘, 并构建线性空间. 首先, 定义加法和数乘这两种运算, 即,

$$\begin{aligned} X + Y : \Omega &\to \mathbb{R} \\ \omega_i &\to x_i + y_i \end{aligned}$$

以及

$$\begin{aligned} k \cdot X : \Omega &\to \mathbb{R} \\ \omega_i &\to kx_i \end{aligned}$$

显然, $X + Y$ 与 $k \cdot X$ 也是随机变量. 自然地, 三维向量

$$\vec{x} + \vec{y},\ k\vec{x}$$

与随机变量 $X + Y, k \cdot X$ 对应. 这里, 取 $k \in \mathbb{R}$.

可以验证, *8* 条法则成立, 那么我们就构造了一个线性空间, 记为 V.

在这个线性空间中有一个很特殊的向量, 我们记

$$\vec{e} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

向量 $\vec{e}$ 对应于一个没有不确定性的随机变量

$$\begin{aligned} E : \Omega &\to \mathbb{R} \\ \omega_i &\to 1 \end{aligned}$$

换句话说, 无论取哪个样本点, 这个随机变量 E 的取值始终为 1. 也就是说, E 其实就是常数 1. 这里, 我们依然将常数 1 看作是一个分量都为 1 的三维向量, 一个没有不确定性随机变量.

我们在 V 上定义内积

$$(\vec{x},\vec{y}) = \vec{x}^T P\vec{y} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}^T \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \sum_{i=1}^{3} p_i x_i y_i.$$

那么, 赋予内积的线性空间 V, 就成了欧式空间.

根据内积, 我们就可以定义距离

$$\|\vec{x}\| = \sqrt{(\vec{x},\vec{x})} = \sqrt{\sum_{i=1}^{3} p_i x_i^2}.$$

一件有趣的事情是, 向量

$$\vec{e} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

的长度为 1.

那么, 在概率论中, 我们熟知的期望, 方差, 协方差, 相关系数是什么样子呢?

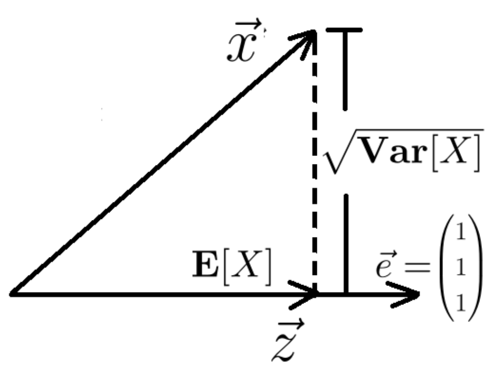
$$\begin{aligned}
\mathbf{E}[X] &= p_1x_1 + p_2x_2 + p_3x_3 = \vec{e}^T P\vec{x} = (\vec{e},\vec{x}),\\
\mathbf{Var}[X] &= \mathbf{E}[(X-\mathbf{E}[X])(X-\mathbf{E}[X])] = (\vec{x}-(\vec{e},\vec{x})\vec{e}, \vec{x}-(\vec{e},\vec{x})\vec{e}),\\
\mathbf{Cov}[X,Y] &= \mathbf{E}[(X-\mathbf{E}[X])(Y-\mathbf{E}[Y])] = (\vec{x}-(\vec{e},\vec{x})\vec{e}, \vec{y}-(\vec{e},\vec{y})\vec{e}),\\
\rho_{XY} &= \frac{\mathbf{Cov}[X,Y]}{\sqrt{\mathbf{Var}[X]\mathbf{Var}[Y]}} = \mathbf{E}[\frac{X-\mathbf{E}[X]}{\sqrt{\mathbf{Var}[X]}}\frac{Y-\mathbf{E}[Y]}{\sqrt{\mathbf{Var}[Y]}}] = (\frac{\vec{x}-(\vec{e},\vec{x})\vec{e}}{\|\vec{x}-(\vec{e},\vec{x})e\|}, \frac{\vec{y}-(\vec{e},\vec{y})\vec{e}}{\|\vec{y}-(\vec{e},\vec{y})e\|}).
\end{aligned}$$

接下来我们说说这些概念在欧式空间 V 中的几何意义.

设向量

$$\vec{z} = (\vec{e},\vec{x})\vec{e},$$

则它是向量 $\vec{x}$ 在向量 $\vec{e}$ 上的垂直投影向量.



那么, 大家熟知的方差计算公式

$$\mathbf{Var}[X] = \mathbf{E}[X^2] - (\mathbf{E}[X])^2$$

其实就是勾股定理公式. 图中, 直角三角形斜边长度的平方为

$$\|\vec{x}\|^2 = (\vec{x},\vec{x}) = \mathbf{E}[X^2],$$

一条直角边长度的平方为

$$\|\vec{z}\|^2 = (\vec{z},\vec{z}) = (\vec{e},\vec{x})^2(\vec{e},\vec{e}) = (\mathbf{E}[X])^2$$

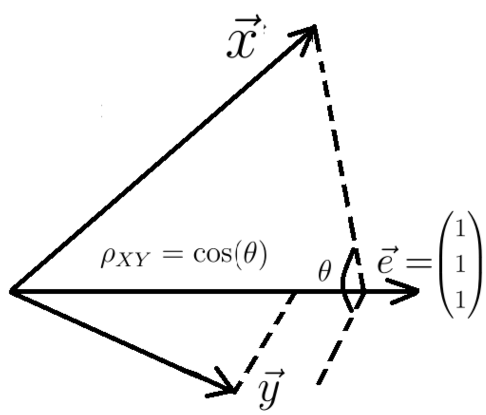
另一条直角边长度的平方为

$$\|\vec{x}-\vec{z}\|^2 = \mathbf{Var}[X].$$

我们再看看相关系数的几何意义. 根据

$$\rho_{XY} = \left(\frac{\vec{x}-(\vec{e},\vec{x})\vec{e}}{\|\vec{x}-(\vec{e},\vec{x})e\|}, \frac{\vec{y}-(\vec{e},\vec{y})\vec{e}}{\|\vec{y}-(\vec{e},\vec{y})e\|}\right).$$

以及期望和方差的几何意义, 我们可以得到随机变量 X 和 Y 的相关系数 ρ_{XY}, 就是向量 $\vec{x}$ 和 $\vec{e}$ 张成的半平面与向量 $\vec{y}$ 和 $\vec{e}$ 张成的半平面之间夹角的余弦值, $\cos(\theta)$.



如果 $\rho_{XY}=0$, 说明这两个半平面的夹角为 $\pi/2$. 如果 $\rho_{XY}=1$, 说明这两个半平面的夹角为 0, 即 $\vec{y}$ 可以由 $\vec{x}$ 和 $\vec{e}$ 线性表出,

$$\vec{y} = k\vec{x} + b\vec{e},$$

且 $k>0$. 如果 $\rho_{XY}=-1$, 说明这两个半平面的夹角为 π, 即 $\vec{y}$ 可以由 $\vec{x}$ 和 $\vec{e}$ 线性表出,

$$\vec{y} = k\vec{x} + b\vec{e},$$

且 $k<0$. 换成随机变量的记号, 即

$$Y = kX + b.$$

最后, 我们说说条件期望的几何意义. 一句话, 随机变量 X 关于随机变量 Y 的条件期望 $E[X|Y]$ 是向量 $\vec{x}$ 在某个与向量 $\vec{y}$ 相关的子空间上的垂直投影.

首先, 我们要定义特征映射 $I_A:\Omega\to\mathbb{R}$, 满足

$$I_A(\omega) = \begin{cases} 1, & \textit{if}\ \omega\in A \\ 0, & \textit{otherwise.}\end{cases}$$

其中 A 是 Ω 的子集. 那么, 随机变量 X 和 Y 分别可以写成

$$X = x_1 I_{\{\omega_1\}} + x_2 I_{\{\omega_2\}} + x_3 I_{\{\omega_3\}},\ Y = y_1 I_{\{\omega_1\}} + y_2 I_{\{\omega_2\}} + y_3 I_{\{\omega_3\}}.$$

下面讨论在一个特殊的情形下, $y_1 = y_2 = a$, $y_3 = b$, 且 $a \neq b$, 条件期望

$$\mathbf{E}[X|Y=a],\ \mathbf{E}[X|Y=b]$$

的几何意义.

此时,

$$Y = aI_{\{\omega_1,\omega_2\}} + bI_{\{\omega_3\}}.$$

$I_{\{\omega_1,\omega_2\}}$ 和 $I_{\{\omega_3\}}$ 这个两个随机变量对应的向量分别为

$$\vec{u} = \begin{pmatrix}1\\1\\0\end{pmatrix},\ \vec{v} = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

我们先算算向量 $\vec{x}$ 到向量 $\vec{u}$ 和 $\vec{v}$ 张成的平面的垂直投影是什么. 换句话说, 我们需要求解下述最小化问题:

$$\min_{s,t} \|\vec{x} - (s\vec{u} + t\vec{v})\|.$$

根据一阶最优条件, 容易得到, 上述问题等价于求解下述线性方程组

$$\begin{pmatrix}(\vec{u},\vec{u}) & (\vec{u},\vec{v})\\(\vec{v},\vec{u}) & (\vec{v},\vec{v})\end{pmatrix}\begin{pmatrix}s\\t\end{pmatrix} = \begin{pmatrix}(\vec{u},\vec{x})\\(\vec{v},\vec{x})\end{pmatrix}.$$

求解上述线性方程组, 得到

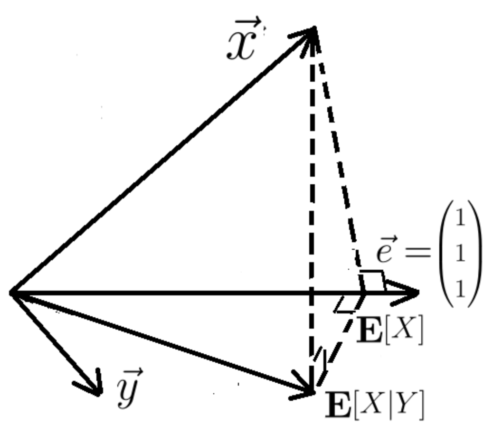
$$s = \frac{p_1x_1 + p_2x_2}{p_1 + p_2},\ t = \frac{p_3x_3}{p_3}.$$

那么, 向量 $\vec{x}$ 到向量 $\vec{u}$ 和 $\vec{v}$ 张成的平面的垂直投影为

$$\vec{w} = \frac{p_1x_1 + p_2x_2}{p_1 + p_2}\vec{u} + \frac{p_3x_3}{p_3}\vec{v}.$$

再根据条件期望的定义, 发现

$$\mathbf{E}[X|Y=a] = s,\ \mathbf{E}[X|Y=b] = t.$$



我们记向量 $\vec{x}$ 投影到向量 $\vec{u}$ 和 $\vec{v}$ 张成的平面的向量 $\vec{w}$, 对应的随机变量为

$$\mathbf{E}[X|Y] = \mathbf{E}[X|Y=a]I_{\{\omega_1,\omega_2\}} + \mathbf{E}[X|Y=b]I_{\{\omega_3\}}.$$

那么, 全期望公式

$$\mathbf{E}[\mathbf{E}[X|Y]] = \mathbf{E}[X].$$

就是几何中的三垂线定理.

随机变量 X 的期望 $E[X]$ 就是向量 $\vec{x} = (x_1, x_2, x_3)^T$ 在向量 $(1,1,1)^T$ 方向上的正交投影 $\vec{z}$,

随机变量 X 的标准差 $\sqrt{\mathrm{Var}[X]}$ 就是 $\vec{x} - \vec{z}$ 的长度

相关系数是二面角的余弦值.

随机变量 X 关于随机变量 Y 的条件期望 $E[X|Y]$ 是向量 $\vec{x}$ 在某个与 $\vec{y}$ 相关的子空间上的投影.

18 Q&A

问: 为什么研究相似呢?

答: 之所以要研究相似, 原因在于数学是一门崇尚“简单”的学科. 研究抽象的线性变换, 在线性空间上取定一组基后可以与一个矩阵唯一对应. 映射的作用转化为矩阵乘以向量. 这是表示理论 (从抽象到具体). 数学的“简单”要求我们找到最简单的矩阵与线性变换与之对应. 这就是相似的起因. 同一个线性变换在不同基下的表示矩阵是相似的.

问: 相似到底怎么理解?

答: 可以从代数和几何两方面描述. 一句话, 同一事物的不同表象. 这也是为什么相似关系是等价关系的本源.

问: 从几何层面, 循环基与 Jordan 块之间的对应关系是怎么建立起来的?

答: 首先循环基的构成就与线性变换直接关联; 其次循环基张成的线性子空间是这个线性变换的不变子空间; 然后线性变换在这个子空间下的限制在循环基下的表示矩阵就是 Jordan 块了.

问: 如何从代数层面, 理解循环基 Jordan 块的对应关系?

答: 首先, 若存在可逆矩阵 P 使得 $P^{-1}AP = J$, 称 A 相似于 Jordan 形矩阵 J. 然后, 将 P 按列分块并展开 $AP = PJ$. 那么, 可以得到 P 的每一列对应某个循环基的一个基向量.

问: 相似关系的代表元为什么选择用 Jordan 形矩阵, 它是最简单的吗?

答: 从代数层面看, Jordan 形矩阵是块对角的, 每一块是一个 Jordan 块. 而且 Jordan 块也足够简单, 对角线上是同一个数 (某个特征值), 下面的次对角线全是 1. 从几何层面看, 要使得线性变换的表示矩阵简单, 最重要的是寻找线性空间的不变子空间的直和分解. 即将线性变换分解成互不干涉的几个部分, 每一部分对应一个不变子空间. 而且在每个不变子空间上所取的基是循环基, 使得线性变换在这个子空间上的限制的表示矩阵足够接近于数量矩阵. 换句话说, 将线性空间分的越细, 则表示矩阵越像对角矩阵.

问: 维数公式 $\dim(\mathrm{Ker}(\mathcal{A})) + \dim(\mathrm{Im}(\mathcal{A})) = \dim(V)$ 从几何层面展现出怎么的图景?

答: 一句话, 空间的直和分解. 维数公式体现了将 V 基于 $\mathcal{A}$ 进行分解的过程. 可以参考图9.2.

19 后记

学了一年的线性代数, 或有收获, 也有遗憾. 现在到了总结的时候.

很多同学问我, 怎么学好数学. 不妨从上课前的五分钟开始, 上课早来五分钟, 看看今天老师准备讲什么内容, 一丢丢时间就行. 正如, 若你有一个职业规划, 那么你的人生就有了方向, 不至于太迷茫. 上课也是一样, 如果提前花一点时间预习下, 上课的效率会成倍的提升.

也有很多同学问我, 怎么做题, 学好数学是不是就是会做题. 做题很重要, 请不要忽视课本上的定理证明. 很多题目的解题思想都来源于定理证明. 当然, 数学的学习好比英语的学习, 你确实需要一定的单词量和阅读量, 但最后还是需要理解语言背后的文化.

每个数学定义的名字, 要自己想想, 体会它的内涵. 当一个数学家发现一个新的数学概念的时候, 对待这个概念一定是像对待自己刚出生的孩子一样. 所以, 他一定不会随随便便起名字的, 名字中一定包含了数学家对这个新生儿的期寄. 比如说, 行列式 – determinant (判别式), 秩 – rank (秩序). 就拿秩来说, 秩序是什么, 就是大家都以一个统一的标准, 做事情, 并且以这个标准对事情分类. 比方说, 大家去超市买东西, 都会以先来后到的标准, 去结账. 如果你这么做了, 就认为是好的, 如果你加三, 就认为是坏的. 这不正是线性代数, 变换, 等价关系, 分类的思想吗.

数学是一个色彩斑斓的超高维物体, 这也是数学抽象的原因. 老师们没有办法让大家一眼就认识她. 初学数学, 可能你会觉得很枯燥, 原因在于, 无论是自己看书还是老师讲解. 你必须按部就班, 一页一页看, 一个定理一个定理的讲. 换句话说, 学习数学的过程呢, 是一个一维的过程. 就像一个非常精妙的工艺品, 蚂蚁在上面爬, 等它一步步一维式地爬完了才能在自己脑海中构建出这个艺术品的样子, 才能品味这个艺术品. 数学也像一部, 情节紧凑, 毫不拖沓, 每个细节都是为结尾铺垫的不刷上几遍恐是不能看懂的悬疑烧脑大剧.

一门数学课, 线性代数, 数学分析, 都可以这样去想. 等到大家大三, 大四, 你会发现, 你所学的每一门课都是有关联的, 慢慢地, 你就会发觉, 数学在你的脑海里就是一棵树, 树上结了很多果实这颗果实是实变函数, 那颗果实是复变函数, 另外一颗是泛函分析它们的根都是集合与映射, 每一门数学课无非就是研究不同的集合与不同的映射原来它们的思路都是一样一样的, 只是因为不同的集合, 不同的映射, 各个果子的形状各异, 但是你要细看, 其实它们的纹路, 层次, 都是一样的.

那么大学四年的数学生涯, 给你带来了什么呢, 会影响你以后的生活吗? 我不敢说你能把这颗树抱回家, 但是你会油然而生一种审美. 你会对着一个别人压根不知道怎么念的一个公式, 欣喜若狂, 高兴的说不出话来, 因为看到了美的东西, 看到了一个艺术品.

大学四年的数学生涯, 给我了一双发现美的眼睛. 可以让我忘我地, 毫无杂念地, 静静地去观赏. 欣赏这个无比精妙绝伦的艺术品.

References

[1] Gilber Strang, *Linear Algebra and Its Application*, Fourth Edition, Thomson, 2006.

[2] David C. Lay, Steven R. Lay, Judi J. McDonald, *Linear Algebra and Its Applications*. Fifth Edition, Pearson, 2016.

[3] 北京大学数学系几何与代数教研室前代数小组, 高等代数, 第三版, 高等教育出版社, 2005.

[4] 任广千, 胡翠芳, 线性代数的几何意义 — 图解线性代数, 西安电子科技大学出版社, 2015.

[5] 史树中, 金融学中的数学, 高等教育出版社, 2006.

[6] 同济大学数学系, 线性代数, 第五版, 高等教育出版社, 2007.

[7] 姚慕生, 高等代数学, 第三版, 复旦大学出版社, 2003.

[8] 张禾瑞, 郝鈵新, 高等代数, 第五版, 高等教育出版社, 2007.