线性代数

刘博文, 余成龙

2025 年 5 月 27 日

目录

前言

这是 2025 春示范班线性代数课程的讲义草稿, 由于本人精力以及能力的限制, 在整理过程中难免出现一些疏漏, 而这都是我本人的错误. 如有您发现了任何笔误或有任何修改意见, 欢迎通过邮件与我联系: liubw22@mails.tsinghua.edu.cn.

刘博文

1 线性方程组

1.1 引言

PageRank 算法是 Google 搜索引擎早期用于计算网页链接权重的核心算法. 我们通过一个简化模型来说明其基本原理：假设有四个网页 1、2、3、4, 它们之间的超链接关系如下图所示：

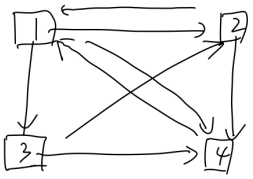


图 1: 网页链接关系示意图

其中 x_i 表示第 i 个网页的重要性分数. 对于 1 号网页, 其重要性分数（流量）被均分为三部分, 分别流向 2、3、4 号网页：

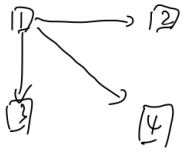


图 2: 1 号网页的流量分配

同时, 1 号网页也从 2 号和 4 号网页接收流量：

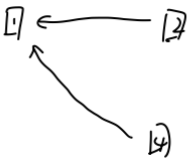


图 3: 1 号网页的流量输入

为简化模型, 我们假设流出的流量是平均分配的. 根据网页间的超链接关系, 可

以建立如下线性方程组来描述各网页的重要性分数:

$$\begin{cases} x_1 = \frac{1}{2}x_1 + x_4 & (1 \text{ 号网页的流量平衡}) \\ x_2 = \frac{1}{3}x_1 + \frac{1}{2}x_3 & (2 \text{ 号网页的流量平衡}) \\ x_3 = \frac{1}{3}x_1 & (3 \text{ 号网页的流量平衡}) \\ x_4 = \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 & (4 \text{ 号网页的流量平衡}) \end{cases}$$

通过高斯消元法求解上述方程组, 我们得到通解形式为:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = t \cdot \begin{pmatrix} \frac{4}{3} \\ \frac{2}{3} \\ \frac{4}{9} \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}^+$$

其中 t 为正实数, 表示解的比例系数. 这表明所有网页的重要性分数具有固定的比例关系.

1.2 列向量空间与线性函数

本节将从三个基本问题出发, 系统介绍列向量空间与线性函数的概念:

(1) 研究对象所在的集合（空间）是什么?

(2) 线性方程的本质特征是什么?

(3) 线性方程组的求解方法有哪些?

1.2.1 列向量空间以及加法, 数乘运算

我们记 $\mathbb{R}$ 为全体实数的集合. 在例子中我们有

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

其中 x_i 都是实数, 这样的 x 被称为一个长度为 4 的实数组. 如果我们考虑长度为 2 的实数组全体构成的集合

$$\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R}\},$$

这个集合记做 $\mathbb{R}^2$, 并且这个集合中的元素一一对应于平面上的点. 同时, $\mathbb{R}^2$ 中的点 P 也可以一一对应与从原点到这个点的向量 $\vec{OP}$, 如下图所示:

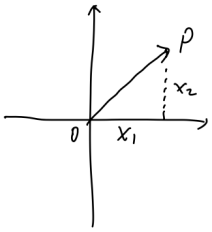


图 4: 点与向量的一一对应

类似的, 我们可以认为 $\mathbb{R}^3$ 一一对应于三维空间中的点, 也一一对应于从原点到这个点的向量.

定义 1.2.1. 列向量空间 *(column vector space)*定义为如下的集合

$$\mathbb{R}^n = \{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R} \},$$

其中的元素称为**列向量** *(column vector).*

定义 1.2.2. $\mathbb{R}^n$ 上可以定义如下两种运算*:*

(1) **加法** *(addition):*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

(2) **数乘** *(scalar product):* 任取实数 $c \in \mathbb{R}$, 定义

$$c \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} cx_1 \\ cx_2 \\ cx_3 \\ \vdots \\ cx_n \end{pmatrix}.$$

注 1.2.1. 向量的加法和数乘有其对应的几何意义, 对于 $\mathbb{R}^2$ 来说, 其中向量的加法和数乘的几何意义可以通过如下具体图形象的展示*:*

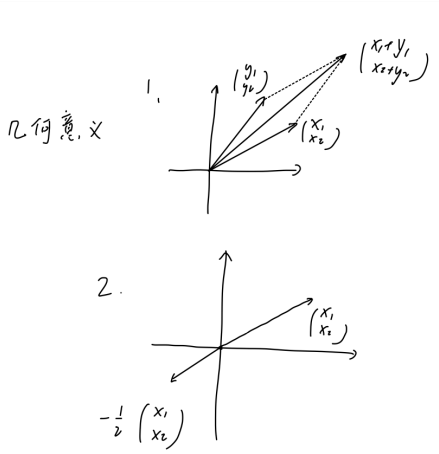


图 5: 向量加法与数乘的几何意义

定义 1.2.3. 对于 $\mathbb{R}^n$ 上的函数 $F\colon \mathbb{R}^n \to \mathbb{R}$, 如果存在 $a_1, \ldots, a_n \in \mathbb{R}$ 是常数, 使得 F 有如下表达式

$$F(x) = a_1x_1 + \cdots + a_nx_n,$$

那么称 F 是 $\mathbb{R}^n$ 上的**线性函数** *(linear function).*

例 1.2.1. 如下的 F 是 $\mathbb{R}^2$ 上的线性函数*:*

$$\begin{aligned} &F\colon \mathbb{R}^2 \to \mathbb{R} \\ &\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto (x_1+1)^2 - (x_1-1)^2 + (x_2-1)^2 - (x_2+1)^2. \end{aligned}$$

定理 1.2.1. 函数 $F\colon \mathbb{R}^n \to \mathbb{R}$ 是线性函数当且仅当 F 满足*:*

(1) 对任意 $x, y \in \mathbb{R}^n$, 有

$$F(x+y) = F(x) + F(y).$$

(2) 对任意 $c \in \mathbb{R}, x \in \mathbb{R}^n$, 有

$$F(cx) = cF(x).$$

注 1.2.2. 这个定理说明 $\mathbb{R}^n$ 上的函数 F 是线性函数当且仅当 F 与 $\mathbb{R}^n$ 上加法与数乘, 也就是与 $\mathbb{R}^n$ 上的线性结构相容, 这也是 F 为什么被称为线性函数.

证明. 如果 F 是线性函数, 可以直接验证 F 满足 (1), (2) 两条性质; 另一方面, 如果

F 满足 (1), (2), 我们记

$$a_1 = F(\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}), \quad a_2 = F(\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}), \quad \dots$$

则

$$F(\begin{pmatrix}x_1\\0\\\vdots\\0\end{pmatrix}) = a_1x_1, \quad F(\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}) = a_2x_2, \quad \dots$$

那么任取 $v_1, \dots, v_m \in \mathbb{R}^n$, 则

$$\begin{aligned} F(v_1 + v_2 + \cdots + v_m) &= F((v_1 + \cdots + v_{m-1}) + v_m) \\ &= F(v_1 + \cdots + v_{m-1}) + F(v_m) \\ &= F(v_1) + \cdots + F(v_m). \end{aligned}$$

因此,

$$F(\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}) = F(\begin{pmatrix}x_1\\0\\\vdots\\0\end{pmatrix} + \cdots + \begin{pmatrix}0\\0\\\vdots\\x_n\end{pmatrix}) = F(\begin{pmatrix}x_1\\0\\\vdots\\0\end{pmatrix}) + \cdots + F(\begin{pmatrix}0\\0\\\vdots\\x_n\end{pmatrix}) = a_1x_1 + \cdots + a_nx_n.$$

□

命题 1.2.1. *如果 F 是线性函数, 那么 $F(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = 0$.*

证明. 任意取 $0 \neq c \in \mathbb{R}$, 根据定义则有

$$F(c\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = cF(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}),$$

因此 $F(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = 0$. □

例 1.2.2. 如下的 F_1, F_2 不是 $\mathbb{R}^2$ 上的线性函数:

$$\begin{aligned} F_1 &: \mathbb{R}^2 \to \mathbb{R} \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\mapsto x_1 + x_4 + 4, \end{aligned}$$

$$\begin{aligned} F_2 &: \mathbb{R}^2 \to \mathbb{R} \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &\mapsto x_1^2 + x_2^2. \end{aligned}$$

对于函数 F_1, 通过直接的计算得到 $F_1(\begin{pmatrix} 0 \\ 0 \end{pmatrix}) \neq 0$, 从而利用命题*1.2.1*可知 F_1 不是线性函数. 对于函数 F_2, 我们可以发现

$$F_2(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = 1, \quad F_2(2\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = 4 \neq 2,$$

从而根据线性函数的定义中第二条可知 F_2 不是线性函数.

定义 1.2.4. $\mathbb{R}^n$ 上 m 个线性函数 $F_1, F_2, \ldots, F_m$ 和 m 个实数 $b_1, b_2, \ldots, b_m$ 满足的方程组

$$\begin{cases} F_1(x) = b_1 \\ F_2(x) = b_2 \\ \vdots \\ F_m(x) = b_m \end{cases}$$

称为 n 个变元的**线性方程组** *(system of linear equations)*, 带入方程组使得其成立的 x 称为**线性方程组的解** *(solution of system of linear equations)*.

注 1.2.3. 我们可以给线性方程组如下的一些几何解释:

(1) 在 $\mathbb{R}^2$ 中, 单个线性函数 $F_1(x) = a_1x_1 + a_2x_2$ 以及实数 b_1 给出的线性方程组 $a_1x_1 + a_2x_2 = b_1$ 的解是 $\mathbb{R}^2$ 中的一条直线.

(2) 在 $\mathbb{R}^2$ 中, 根据 *(1)* 的几何解释不难理解如下线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

的解是 $\mathbb{R}^2$ 中两条直线的交点. 注意, 在 $\mathbb{R}^2$ 中两条不一样的直线不一定相交, 即如上线性方程组不一定有有解, 但是如果有解一定只有唯一解.

(3) 在 $\mathbb{R}^3$ 中, 如下线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{cases}$$

的解可以看成是 $\mathbb{R}^3$ 中两个平面的交线. 注意: 在 $\mathbb{R}^3$ 中两个不一样的平面不一定相交, 即如上线性方程组不一定有有解, 并且如果相交, 也是交出一条线, 即此时解不唯一.

(4) 在更高维中也有同样的解释: 由一个线性函数给出的线性方程的解可以看成是一个低一维的超平面, 而多个线性函数给出的线性方程组的解则是这些超平面的交.

1.3 高斯消元法和最简阶梯型

根据注记1.2.3可知对于一个线性方程组其可能没有解, 并且即使有解也不一定只有唯一解, 那么该如何求解线性方程组呢? 在本节中我们将利用高斯消元法, 来求解一般的线性方程组. 我们先来看下面的一个简单的例子.

例 1.3.1.

$$\begin{cases} 4x_2 - x_3 = 7 & (r_1) \\ x_1 + 2x_2 = 5 & (r_2) \\ 2x_1 + x_3 = 3 & (r_3) \end{cases}$$

显然我们交换 r_1, r_2 不改变上述方程组的解, 因此我们得到:

$$\begin{cases} x_1 + 2x_2 = 5 & (r_1') \\ 4x_2 - x_3 = 7 & (r_2') \\ 2x_1 + x_3 = 3 & (r_3') \end{cases}$$

我们考虑如下操作: 保持 r_1', r_2' 不变, 用 r_3' 减去 $2r_1'$, 得到如下的方程组:

$$\begin{cases} x_1 + 2x_2 = 5 & (r_1'') \\ 4x_2 - x_3 = 7 & (r_2'') \\ -4x_2 + x_3 = -7 & (r_3'') \end{cases}$$

上述操作并不改变方程组的解, 因为可由 r_1'', r_2'', r_3'' 恢复出 r_1', r_2', r_3'. 类似的最后再保持 r_1'', r_2'' 不变, 用 r_3'' 加上 r_2'', 得到

$$\begin{cases} x_1 + 2x_2 = 5 \\ 4x_2 - x_3 = 7 \\ 0 = 0 \end{cases}$$

对于上述方程组我们可以用 x_3 来如下的表示 x_1, x_2, 其中 x_3 可以取任意的实数

$$\begin{aligned} x_1 &= -\frac{1}{2}x_3 + \frac{3}{2} \\ x_2 &= \frac{1}{4}x_3 + \frac{7}{4} \end{aligned}$$

因此我们可以将方程组的解写作

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{3}{2} \\ \frac{7}{4} \\ 0 \end{pmatrix}.$$

注 1.3.1. 根据上述结果可以发现该线性方程组有无穷组解, 这对应于几何解释中 $\mathbb{R}^3$ 中三个平面相交出一条线.

回顾例1.3.1, 在解方程中我们主要用到了如下三种操作:

(E1) 交换方程组的某两行.

(E2) 将某一行乘以非零常数 c.

(E3) 将某一行的非零常数 c 倍加到另一行上.

我们称如上的三种操作为**基础行变换** (elementary row operations). 不难发现基础行变换均可逆, 并且其逆也是基础行变换.

定义 1.3.1. 有限个基础行变换的复合称为**行变换** *(row operations)*.

命题 1.3.1. 行变换均可逆, 并且其逆也为行变换.

证明. 因为基础行变换可逆, 且其逆也为基础行变换, 并且操作 O_1O_2 的逆为 $O_2^{-1}O_1^{-1}$. □

推论 1.3.1. 行变换不改变线性方程组的解.

由于作行变换只关注方程的系数以及右侧常数项, 因此对于如下的线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

我们将其系数及常数项提出来记作

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

并将这个线性方程组记做 $Ax = b$, 这也引出了矩阵的概念.

定义 1.3.2. *由 $m \times n$ 个数 a_{ij} 排成的 m 行 n 列的 (实) 数表称为 m 行 n 列* **矩阵** *(matrix), 记做 $(a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{R})$. 当 $m = n$ 时, $A \in M_{n\times n}(\mathbb{R})$ 被称为 n 阶* **方阵** *(square matrix), 此时 $M_{n\times n}(\mathbb{R})$ 通常简记为 $M_n(\mathbb{R})$.*

例 1.3.2. *$\mathrm{I}_n \in M_n(\mathbb{R})$ 是只有 (i, i) 元为 1, 其余分量为零的矩阵, 称为* **单位矩阵** *(identity matrix).*

对于线性方程组 $Ax = b$, A 称为**系数矩阵** (coefficient matrix), (A, b) 称为**增广矩阵** (augmented matrix), 并将上述方程记作 $Ax = b$. 现在我们即可以通过行变换来操作我们的增广矩阵, 使其最终的形式方便于我们求解, 那么究竟该操作到什么样子为止呢?

根据例1.3.1, 我们发现如果我们的增广矩阵有如下的形式, 线性方程组可以直接求解:

(1) 所有非零行在零行的上面.

(2) 对某一非零行, 称最左边的非零元为**主元** (pivot), 第 i 行的主元严格比第 $i + 1$ 行的主元靠左.

满足上述条件的矩阵称为**行阶梯型** (row echelon form), 并且如果主元所在列的其他元素均为零, 主元本身为 1, 则称此时为**最简行阶梯型** (reduced row echelon form).

定理 1.3.1. *矩阵 A 可通过行变换变成最简行阶梯型, 并且该最简行阶梯型不依赖于行变换的选取, 记作 $\operatorname{rref} A$.*

证明. 对 $m \times n$ 矩阵的列数作归纳: 假设 $n = 1$ 时, 对于 $m \times 1$ 矩阵

$$A = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}.$$

如果 $a_{11} = \cdots = a_{m1} = 0$, 则此时已经是最简行阶梯型. 若 $a_{11} = a_{21} = \cdots = a_{(k-1)1} = 0, a_{k1} \neq 0$, 那么通过 (E1) 将 a_{k1} 换到第一行, 用 (E2) 将第一行乘以 $(a_{k1})^{-1}$ 使得主元变为 1, 再用 (E3) 将第一行以下变为零, 因此此时最简行阶梯型为

$$\operatorname{rref} A = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

并且此时容易看出不依赖于行列变换的选取, 既最简行阶梯型是唯一的.

假设对列数为 n 的时候成立, 对于 $m\times(n+1)$ 的矩阵 A, 将其写作 $A=(B,y)$, 其中 B 是 $m\times n$ 矩阵. 根据归纳假设 B 可由行变换得到最简行阶梯型, 记作 B', 将同样的变换作用在 A 上得到 $A'=(B',y')$. 如果 B' 没有非零行, 则次数 A' 已经是最简行阶梯型. 如果 B' 从 $k+1$ 行开始是零行, 则对

$$\begin{pmatrix} y'_{k+1}\\ \vdots\\ y'_m \end{pmatrix}$$

应用 $n=1$ 时的结论, 可做行变换得到最简行阶梯型, 同时也对 B' 作. 但由于行变换不改变零矩阵, 因此不改变 B', 得到的矩阵记作 A''. 考虑如下两种情况:

(1) 如果

$$\mathrm{rref}\begin{pmatrix} y'_{k+1}\\ \vdots\\ y'_m \end{pmatrix}=\begin{pmatrix} 0\\ \vdots\\ 0 \end{pmatrix},$$

则此时 A'' 已经是最简行阶梯型.

(2) 如果

$$\mathrm{rref}\begin{pmatrix} y'_{k+1}\\ \vdots\\ y'_m \end{pmatrix}=\begin{pmatrix} 1\\ \vdots\\ 0 \end{pmatrix},$$

则作 (E3) 将第 $k+1$ 行加到第 $1,2,\dots,k$ 行, 将 $y'_1,\dots,y'_k$ 变成零, 此时得到的矩阵也是最简行阶梯型. 最简行阶梯形的唯一性我们留在1.5中证明, 一个好的阅读材料是第一位女菲尔兹奖得主写的短文 “A Simple Proof of a Theorem of Schur”, https://doi.org/10.1080/00029890.1998.12004879.

□

定义 1.3.3. 对于矩阵 $A\in M_{m\times n}(\mathbb{R})$, 其主元的个数被定义为 A 的**秩** *(rank)*, 记做 $\mathrm{rank}\,A$.

注 1.3.2. 根据最简行阶梯型的唯一性, 可知矩阵的秩的定义是良好的.

1.4 线性方程组解的结构

定义 1.4.1. 对于线性方程组的系数矩阵 A, $\mathrm{rref}\,A$ 中主元所在的列对应的未知元称为**主元** *(principal unknowns)*, 其余未知元称为**自由元** *(free unknowns)*.

例 1.4.1. 例如线性方程组 $Ax = b$, 其中

$$A = \begin{pmatrix} 1 & 0 & 3 & 5 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix},$$

则 x_1, x_2, x_5 是主元, x_3, x_4, x_6 是自由元. 并且根据上述最简行阶梯型, 我们可以直接分析出方程组的解的情况:

1. 如果 b_4 或者 b_5 不是零, 则方程组 $Ax = b$ 无解.
2. 如果 $b_4 = b_5 = 0$, 则 x_3, x_4, x_6 取定任意实数后, 主元由方程组唯一确定:

$$\begin{aligned} x_1 &= b_1 - 3x_3 - 5x_4 - x_6 \\ x_2 &= b_2 - 2x_3 - x_4 - 2x_6 \\ x_5 &= b_3. \end{aligned}$$

定理 1.4.1 (线性方程组解的结构定理)**.** 对于方程 $Ax = b$, 用行变换将 (A, b) 化作最简行阶梯型 $(\overline{A}, \overline{b})$, 则

(1) 方程有解等价于 $\overline{A}$ 的零行对应的 $\overline{b}_i$ 也是零.

(2) 方程有解时自由元可以任意取值, 且自由元的每一组取值都唯一决定了一组解. 特别地, 方程有唯一解当且仅当没有自由元.

推论 1.4.1. 线性方程组 $Ax = b$

(1) 有解当且仅当 $\operatorname{rank} A = \operatorname{rank}(A, b)$.

(2) 有唯一解当且仅当 $\operatorname{rank} A$ 等于 A 的列数相同.

定义 1.4.2. 方程 $Ax = 0$ 称为**齐次线性方程组** *(system of homogeneous linear equations).*

定理 1.4.2. 齐次线性方程组 $Ax = 0$ 的解在加法和数乘下封闭.

证明. 注意到

$$\begin{aligned} A(x + y) &= Ax + Ay, \\ A(cx) &= cAx. \end{aligned}$$

□

定理 1.4.3. 对于线性方程组 $Ax = b$, 如果 $\widetilde{x}$ 是其某一解 *(特解)*, 则 $Ax = b$ 的所有解均可唯一的表达为 $x = y + \widetilde{x}$, 其中 y 是 $Ax = 0$ 的解.

证明. 只需验证如下两点:

(1) 验证 $y+\widetilde{x}$ 是解.

(2) 验证当 x 是解时, $x=(x-\widetilde{x})+\widetilde{x}$, 其中 $x-\widetilde{x}$ 满足 $Ax=0$.

□

注 1.4.1. 从几何上来看, 齐次线性方程组 $Ax=0$ 的解构成了 $\mathbb{R}^n$ 中的一个对加法数乘封闭的子集, 之后我们会用更抽象的观点去描述这种子集, 并称其为一个子空间. 而 $Ax=b$ 的解相当于是将这个子空间做了平移.

定义 1.4.3. 对于线性方程组 $Ax=b$, 我们有如下定义:

(1) 如果方程有解, 我们称这个线性方程组是**相容的** *(consistent)*.

(2) 如果方程无解, 我们称这个线性方程组是**不相容的** *(inconsistent)*.

(3) 如果方程有唯一解, 我们称这个线性方程组是**确定的** *(definite)*.

推论 1.4.2.

(1) 线性方程组 $Ax=b$ 是相容的等价于 $\operatorname{rank}A=\operatorname{rank}(A,b)$.

(2) 线性方程组 $Ax=b$ 是确定的等价于 $Ax=b$ 是相容的, 且 $\operatorname{rank}A$ 等于 A 的列数.

定理 1.4.4. 对于 n 阶方阵 A, 线性方程组 $Ax=b$ 是否有唯一解只取决于 A, 与 b 无关.

证明. 根据推论1.4.2可知, $Ax=b$ 有唯一解等价于 $\operatorname{rank}A=n$. □

例 1.4.2 (Shafarevich-Remizov)**.** 对于互不相同的实数 $c_1,\dots,c_r$, 以及任意实数 $k_1,\dots,k_r$, 存在唯一的次数小于等于 $r-1$ 的多项式 $f(x)=a_0+a_1x+\cdots+a_{r-1}x^{r-1}$ 使得对于任意 $i=1,\dots,r$ 使得

$$f(c_i)=k_i \tag{1}$$

证明. 注意到 (1) 是关于 $a_0,\dots,a_{r-1}$ 这 r 个未知元的线性方程组, 不妨记做

$$A\begin{pmatrix}a_0\\a_1\\\vdots\\a_{r-1}\end{pmatrix}=\begin{pmatrix}k_0\\k_1\\\vdots\\k_{r-1}\end{pmatrix}.$$

根据线性方程组的解的结构定理, 上述方程组有唯一解当且仅当下面的线性方程组有唯一解

$$A\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{r-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

这等价于 $c_1, \dots, c_r$ 是 $f(x)$ 的根, 而以 $c_1, \dots, c_r$ 为根的次数不超过 $r-1$ 的多项式是唯一的, 从而说明了 (1) 解的唯一性. □

1.5 选讲: 最简行阶梯形唯一性的证明

令 A 是一个 $m \times n$ 阶的矩阵, 在本节中我们介绍如何证明最简行阶梯形 $\operatorname{rref} A$ 的唯一性.

证明. 我们对 n 做归纳法来证明: 当 $n = 1$ 时, $\operatorname{rref} A$ 只有如下两种情况:

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

注意到行变换将零矩阵变成零矩阵, 并且由于行变换可逆, 从而行变换将非零矩阵变成非零矩阵. 因此在 $n = 1$ 的情况, 最简行阶梯形 $\operatorname{rref} A$ 是否为零完全由 A 是不是零来决定.

现在假设命题对列数是 n 的情况都成立, 对于 $n+1$ 列的矩阵 A, 我们将其写成 $A = (B, y)$ 的形式, 其中 B 是一个 n 列的矩阵, y 是一个列向量. 假设此时 $\operatorname{rref} A$ 有 $A' = (B', y')$ 和 $A'' = (B'', y'')$ 两种形式, 则 B' 和 B'' 都是由 B 经过行变换得到, 且都是最简行阶梯形, 从而根据归纳假设 $B' = B''$.

现在考虑线性方程组

$$B\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y. \tag{2}$$

由于行变换不改变线形方程组解的情况, 从而方程组 (2) 等价于如下两种方程组

$$B\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y', \quad B\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y''.$$

对于方程组 (2), 我们有如下两种情况:

(i) (2) 无解, 此时是不相容情形. 假设 B 的秩为 r, 那么

$$y' = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = y'',$$

其中 1 位于第 $r+1$ 行.

(ii) (2) 有解. 假设 B 的秩为 r, 并且主元分别在 $i_1, \ldots, i_r$ 列. 令自由元均取 0, 则得到关于 $x_{i_1}, \ldots, x_{i_r}$ 这 r 个未知元的线性方程组, 并且有解:

$$x_{i_1} = y'_1, \ldots, x_{i_r} = y'_r,$$

或者

$$x_{i_1} = y''_1, \ldots, x_{i_r} = y''_r.$$

但由于此时解是唯一的, 从而 $y' = y''$, 从而得到最简行阶梯形 $\operatorname{rref} A$ 是唯一的.

□

1.6 选讲: 舒伯特胞腔

我们用 $G(m,n)$ 记 $\mathbb{R}^n$ 中穿过原点的所有 $\mathbb{R}^m$ 组成的集合.

例 1.6.1. 考虑 2×3 阶矩阵 A, 满足 $\operatorname{rank} A = 2$, 则 $Ax = 0$ 这个线性方程组的解反映在 $\mathbb{R}^3$ 中的几何意义则是 $\mathbb{R}^3$ 中穿过原点的直线. 由于 $\operatorname{rank} A = 2$, 则最简行阶梯形 $\operatorname{rref} A$ 拥有的可能性如下:

$$\begin{pmatrix} 1 & 0 & a_1 \\ 0 & 1 & a_2 \end{pmatrix}, \quad \begin{pmatrix} 1 & b_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

所以 $\mathbb{R}^3$ 中通过原点的直线的集合 $G(1,3)$ 一一对应于

$$\mathbb{R}^2 \coprod \mathbb{R} \coprod \mathbb{R}^0,$$

其中 $\mathbb{R}^0$ 表示单点集.

例 1.6.2. 考虑 2×3 阶矩阵 A, 满足 $\operatorname{rank} A=1$, 则 $Ax=0$ 这个线性方程组的解反映在 $\mathbb{R}^3$ 中的几何意义则是 $\mathbb{R}^3$ 中穿过原点的二维平面. 由于 $\operatorname{rank} A=1$, 则最简行阶梯形 $\operatorname{rref} A$ 拥有的可能性如下:

$$\begin{pmatrix}1 & a_1 & a_2\\ 0 & 0 & 0\end{pmatrix},\quad \begin{pmatrix}0 & 1 & b\\ 0 & 0 & 0\end{pmatrix},\quad \begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}.$$

所以 $\mathbb{R}^3$ 中通过原点的平面的集合 $G(2,3)$ 一一对应于

$$\mathbb{R}^2\coprod\mathbb{R}\coprod\mathbb{R}^0.$$

例 1.6.3. 考虑 2×4 阶矩阵 A, 满足 $\operatorname{rank} A=2$, 则 $Ax=0$ 这个线性方程组的解反映在 $\mathbb{R}^4$ 中的几何意义则是 $\mathbb{R}^4$ 中穿过原点的二维平面. 由于 $\operatorname{rank} A=2$, 则最简行阶梯形 $\operatorname{rref} A$ 拥有的可能性如下:

$$\begin{pmatrix}1 & 0 & a_1 & a_2\\ 0 & 1 & a_3 & a_4\end{pmatrix},\quad \begin{pmatrix}1 & b_1 & 0 & b_2\\ 0 & 0 & 1 & b_3\end{pmatrix},\quad \begin{pmatrix}1 & c_1 & c_2 & 0\\ 0 & 0 & 0 & 1\end{pmatrix},\quad \begin{pmatrix}0 & 1 & 0 & d_1\\ 0 & 0 & 1 & d_2\end{pmatrix},$$

$$\begin{pmatrix}0 & 1 & e & 0\\ 0 & 0 & 0 & 1\end{pmatrix},\quad \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

所以 $\mathbb{R}^3$ 中通过原点的平面的集合 $G(2,4)$ 一一对应于

$$\mathbb{R}^4\coprod\mathbb{R}^3\coprod\mathbb{R}^2\coprod\mathbb{R}^2\coprod\mathbb{R}\coprod\mathbb{R}^0.$$

1.7 作业一

这次作业里所有矩阵不加说明都是实数矩阵.

1.7.1 基础题

本部分题必做.

习题 1.7.1. 用消元法解线性方程组

1. 关于两个变元 x_1, x_2 的线性方程组

$$\begin{cases} x_1 - 2x_2 &= 1 \\ 2x_1 - x_2 &= 2 \end{cases}$$

2. 关于四个变元 $x_1, x_2\ x_3, x_4$ 的线性方程组

$$\begin{cases} x_1 - 2x_2 + 3x_3 - 4x_4 &= 4 \\ x_2 - x_3 + x_4 &= -3 \\ x_1 + 3x_2 + x_4 &= 1 \\ -7x_2 + 3x_3 + x_4 &= -3 \end{cases}$$

3. 关于三个变元 x, y, z 的线性方程组

$$\begin{cases} 2x + y - z = 8 \\ -3x - y + 2z = -11 \\ -2x + y + 2z = -3 \end{cases}$$

习题 1.7.2. 讨论 λ 取何值时，线性方程组

$$\begin{cases} \lambda x_1 + x_2 + x_3 &= 1 \\ x_1 + \lambda x_2 + x_3 &= \lambda \\ x_1 + x_2 + \lambda x_3 &= \lambda^2 \end{cases}$$

有唯一解，无穷多解，无解，并在有解时求其解。

习题 1.7.3. 考虑包含 m 个主元的最简阶梯型的 $m \times n$ 的实矩阵，如果主元出现的位置相同的这样的矩阵视作一类，求不同的种类数有多少？

习题 1.7.4. 考虑一个连通无向无圈无多重边的有限图 G. 令 V 是顶点的集合，将其视为网页，构造如下网络. 如果两个顶点之间有边连接，则假设两个网页之间有彼此两个方向之间的超链接连接. 我们利用 *Google* 的 *PageRank* 算法得到关于每个网页重要性 $(x_i)_{i \in V}$ 的线性方程组. 证明此时 x_i 等于经过顶点 i 的边数是这个方程组的一组解.

习题 1.7.5. 构造一个 *3* 阶方阵, 其 *9* 个元素各不相同, 且行简化阶梯形有且只有一个主元.

习题 1.7.6. 设

$$A=\begin{pmatrix}1 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & -1\end{pmatrix},\quad b=\begin{pmatrix}b_1\\ b_2\\ b_3\end{pmatrix}.$$

证明

1. $Ax=b$ 有解当且仅当 $b_1+b_2+b_3=0$.
2. $Ax=0$ 的解集是 $\{kx_1 : k\in\mathbb{R}\}$, 其中 $x_1=\begin{pmatrix}1\\ -1\\ 1\end{pmatrix}$.
3. 当 $Ax=b$ 有解时, 若 x_0 是一个解, 则解集是 $\{x_0+kx_1 : k\in\mathbb{R}\}$.

1.8 作业二

这次作业里所有矩阵不加说明都是实数矩阵.

1.8.1 基础题

本部分题必做.

习题 1.8.1. 把下列矩阵化为最简行阶梯型: *(默认空格处为 0)*

1. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix}$
2. $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{pmatrix}$
4. $\begin{pmatrix} 2 & 1 & & \\ 1 & 2 & 1 & \\ & 1 & 2 & 1 \\ & & 1 & 2 \end{pmatrix}$
5. $\begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$

习题 1.8.2. 考虑一个连通无向无圈无多重边的有限图 G. 假设 V 是顶点的集合, 其中只有一条边相连的顶点称为边界点, 有多条边相连的顶点称为内部点. 假设内部点和边界点的集合都非空。对每一个顶点 i 取一个温度 $T_i \in \mathbb{R}$, 称 $T = (T_i)_{i \in V}$ 是一个图上的温度分布. 如果每一个内部点的温度等于与之相连的点的温度的平均值, 则称这一分布称为稳定的的. 证明: 对于每一组边界点的温度值, 存在唯一的内部点的温度取值, 使得这一温度分布是稳定的.

习题 1.8.3. 将下列问题转化为求解线性方程组的问题, 并求解:

1. 设 2×2 矩阵 A 满足 $A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix}$ 且 A 的第一列元素之和为 2, 求所有可能的 A.

2. 空间中有一个平面经过点 $\begin{pmatrix}1\\1\\0\end{pmatrix}$, $\begin{pmatrix}0\\1\\1\end{pmatrix}$, $\begin{pmatrix}1\\0\\1\end{pmatrix}$, 求所有与该平面垂直的向量.

3. 写出通过 5 点 $M_1(0,1), M_2(2,0), M_3(-2,0), M_4(1,-1), M_5(-1,-1)$ 的二次曲线的方程. 这里二次曲线是 xy-平面上形如 $ax^2+bxy+cy^2+dx+ey+f=0$ 的方程决定的曲线.

习题 1.8.4. 若 A, C 均为 $m\times n$ 矩阵, 如果对任何 b, 线性方程组 $Ax=b$ 与 $Cx=b$ 都有相同的解集, 是否一定有 $A=C$?

1.8.2 思考题

本部分题选做，学期中任何时间都可以交，不计成绩。

习题 1.8.5. （还没讲到，2 月 25 日课会讲）课上我们研究过 $G(m,n)$ 的分解中 $\mathbb{R}^i$ 的个数, 记为 b_i.

1. 求 b_i 的生成函数 $\sum_i b_i t^i$.

2. 验证 $b_i = b_{m(n-m)-i}$.

3. 任取一组正实数 m, n，验证 b_i 是单峰的 (先单调递增后单调递减).

习题 1.8.6. 若 A, A' 均为 $m\times n$ 矩阵, b, b' 为 m 维向量, 方程 $Ax=b$ 与 $A'x=b'$ 的解集相同且非空, 请思考 (A', b') 是否一定可由 (A, b) 经过行变换得到，你能对 $m=n=2$ 写出证明吗？

2 矩阵及其运算

2.1 矩阵乘法

在 $M_{m\times n}(\mathbb{R})$ 上有如下的运算:

(1) 加法: $A=(a_{ij})_{m\times n}, B=(b_{ij})_{m\times n}$, 则 $A+B:=(a_{ij}+b_{ij})_{m\times n}$.

(2) 数乘: $A=(a_{ij})_{m\times n}, c\in\mathbb{R}$, 则 $cA:=(ca_{ij})_{m\times n}$.

除了以上两种结构, 还满足额外的乘法结构, 在研究线性方程组的时候我们已经见到了矩阵和向量相乘的例子: 给定 $c=(c_1,\dots,c_n)\in\mathbb{R}^n$ 和 $A\in M_{m\times n}(\mathbb{R})$, 如果按照列向量排列的形式将 A 记做 $A=(v_1,\dots,v_n)$, 那么

$$Ac:=c_1v_1+\cdots+c_nv_n,$$

也即是以 $c_1,\dots,c_n$ 为系数的 A 中的列向量 $v_1,\dots,v_n$ 的线性组合.

定义 2.1.1. *对于向量 $v_1,\dots,v_n\in\mathbb{R}^m, c_1,\dots,c_n\in\mathbb{R}$, 则 $c_1v_1+\cdots+c_nv_n$ 称为以 $c_1,\dots,c_n$ 为系数的 $v_1,v_2,\dots,v_n$ 的***线性组合*** (linear combination).*

对于两个矩阵 A,B, 如果 A 的列数与 B 的行数相同, 我们则可以用上面的办法将矩阵乘法定义为:

定义 2.1.2. *对于 $A\in M_{m\times n}(\mathbb{R})$ 以及 $B=(w_1,\dots,w_\ell)\in M_{n\times\ell}(\mathbb{R})$, 则***矩阵乘法*** (matrix multiplication)定义为*

$$AB:=(Aw_1,\dots,Aw_\ell).$$

同时, 矩阵乘法还可以用如下等价的方式定义:

定义 2.1.3. *对于 $A\in M_{m\times n}(\mathbb{R}), B\in M_{n\times l}(\mathbb{R})$, 则***矩阵乘法*** (matrix multiplication) 定义为 $AB:=(c_{ij})_{m\times l}$, 其中*

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

注 2.1.1. *对于矩阵 A,B, 只有 A 的列数与 B 的行数相同时, 矩阵乘法 AB 才有意义.*

命题 2.1.1. *矩阵乘法具有结合律.*

证明. 为了方便起见我们不妨假设 $A,B,C\in M_n(\mathbb{R})$, 更一般地情况证明是类似的. 记 $C=(c_1,\dots,c_n)$, 其中 c_i 是列向量. 那么

$$(AB)C=((AB)c_1,\dots,(AB)c_n)$$
$$A(BC)=A(Bc_1,\dots,Bc_n)=(A(Bc_1),\dots,A(Bc_n))$$

因此只需对每个 c_i 验证 $(AB)c_i = A(Bc_i)$ 即可, 因此我们不妨假设 C 是 $n \times 1$ 的矩阵. 将 A 写作

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

其中 a_i 是行向量, 那么

$$(AB)C = \begin{pmatrix} a_1B \\ a_2B \\ \vdots \\ a_nB \end{pmatrix} C = \begin{pmatrix} a_1BC \\ a_2BC \\ \vdots \\ a_nBC \end{pmatrix}$$

$$A(BC) = \begin{pmatrix} a_1(BC) \\ a_2(BC) \\ \vdots \\ a_n(BC) \end{pmatrix}$$

因此只需要对每一个 a_i 验证即可, 因此我们不妨假设 A 是 $1 \times n$ 的矩阵, 那么

$$((a_1, \ldots, a_n)B) \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = \sum_{j=1}^{n} (\sum_{i=1}^{n} a_i b_{ij}) c_j = \sum_{i,j} a_i b_{ij} c_j,$$

$$(a_1, \ldots, a_n)(B \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}) = \sum_{j=1}^{n} a_i (\sum_{i=1}^{n} b_{ij} c)_j = \sum_{i,j} a_i b_{ij} c_j.$$

□

有了矩阵乘法, 我们可以重新将之前对系数矩阵做的初等行变换用矩阵的语言再解释一遍.

定义 2.1.4. 如下的三类矩阵被称为**初等矩阵** *(elementary matrix):*

(E1)

$$E[ij] = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & \ddots & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

即交换 I_n 的第 i 行与第 j 行得到的矩阵.

(E2)

$$E[i,c] = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$$

即将 I_n 的第 i 行乘以 c 得到的矩阵, 其中 $c \neq 0$.

(E3)

$$E[ij,c] = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & \vdots & \ddots & & & \\ & & c & \cdots & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

即将 I_n 的第 i 行乘以 c 加到第 j 行得到的矩阵, 其中 $c \neq 0$.

例 2.1.1. 假设系数矩阵

$$A = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix},$$

那么如果我们对其作用 *(E3)* 将第三行的 2 倍加到第一行上去, 得到的新的系数矩阵记做 A', 那么

$$A' = \begin{pmatrix} r_1 + 2r_3 \\ r_2 \\ r_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A.$$

因此我们可以看出初等行变换 *(E3)* 可以看作是初等矩阵 *(E3)* 左乘.

命题 2.1.2. 对 $A \in M_{m\times n}(\mathbb{R})$ 做初等行变换 O 等价于对左乘相应的初等矩阵 E.

证明. 根据定义验证即可. □

推论 2.1.1. 对 A 做行变换 $O_1 \dots O_k$ 等价于左乘初等矩阵 $E_k \dots E_2 E_1$.

注意到我们的初等行变换是可逆的, 用矩阵的语言来说, 对于初等矩阵 $B \in M_n(\mathbb{R})$, 总存在另一个初等矩阵 B' 使得 $BB'A = \mathrm{I}_n$, 其中 I_n 是只有对角线为 1, 其余地方全为零的 $n \times n$ 矩阵.

定义 2.1.5. 对于矩阵 $A \in M_n(\mathbb{R})$

(1) 若有 B 使得 $BA = \mathrm{I}_n$, 则称 B 为 A 的**左逆** *(left inverse)*.

(2) 若有 C 使得 $AC = \mathrm{I}_n$, 则称 C 为 A 的**右逆** *(right inverse)*.

(3) 如果左逆右逆均存在, 则称 A **可逆** *(invertible)*.

定理 2.1.1. 对于矩阵 $A \in M_n(\mathbb{R})$, 如下叙述等价:

(1) A 可逆.

(2) A 存在左逆.

(3) A 存在右逆.

(4) $\operatorname{rref} A = \mathrm{I}_n$.

(5) $Ax = b$ 有唯一解.

(6) $Ax = 0$ 有唯一解.

(7) $\operatorname{rank} A = n$.

证明. 根据线性方程组解的结构定理, 即定理1.4.1, 我们已经证明了 (4),(5),(6),(7) 的等价性.

(2) $\Longrightarrow$ (6): 如果 A 存在左逆, 那么 $A^{-1}Ax = 0$ 意味着 $x = 0$, 即 $Ax = 0$ 只有唯一的解.

(5) $\Longrightarrow$ (3): 如果 $Ax = b$ 存在唯一解, 那么我们不妨取

$$b_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots$$

那么不妨记 $Ax = b_1, Ax = b_2, \dots, Ax = b_n$ 唯一的解分别是 $w_1, \dots, w_n$. 令 $C = (w_1, \dots, w_n)$ 则 $AC = \mathrm{I}_n$, 即 C 是 A 的右逆.

注 2.1.2. 至此已经证明了如果矩阵 A 存在左逆, 那其一定存在右逆, 即 (2) $\Longrightarrow$ (3).

(3) $\Longrightarrow$ (2): 假设 A 有右逆, 存在 C 使得即 $AC = \mathrm{I}_n$, 从而 $CAC = C$. 另一方面, 由于 C 存在左逆, 从而 C 存在右逆, 不妨记为 D, 因此

$$CA = CACD = CD = \mathrm{I}_n$$

即 C 也是 A 的左逆.

注 2.1.3. 从上述证明可以看出, 如果 A 存在左逆, 那么其右逆不仅存在, 并且还和左逆相同. 类似的可以说明如果 A 存在右逆则其左逆不仅存在, 也与右逆相同.

□

命题 2.1.3. 若 A 可逆, 则左逆与右逆均唯一存在且相同, 记做 A^{-1}.

证明. 我们只需要证明如果 A 可逆, 那么其左逆右逆都唯一: 假设 C 是 A 的一个左逆, D 是 A 的一个右逆, 那么

$$C = C\,\mathrm{I}_n = CAD = \mathrm{I}_n\,D = D$$

即 A 的任何左逆与右逆都相同. 那么假设 C_1, C_2 是 A 的两个左逆, 由于 C_1 也是 A 的右逆, 从而 $C_1 = C_2$, 即 A 的左逆唯一, 类似的, 我们也可以说明 A 的右逆唯一. □

注 2.1.4. 上述结论表明, 如果 A 可逆, 那么 $\operatorname{rref} A = \mathrm{I}_n$, 而根据推论*2.1.1*可知行变换等价于左乘初等矩阵, 因此将其化为最简行阶梯型的初等矩阵的乘积就是 A^{-1}. 那么我们该如何将这些初等矩阵的乘积记录下来呢? 考虑矩阵 (A, I_n), 对其进行操作使得 A 化为最简行阶梯型则有 (I, A^{-1}), 这也给出了我们求逆的办法, 并且我们也有如下简单的推论.

推论 2.1.2. 矩阵 A 可逆当且仅当其为初等矩阵的乘积.

例 2.1.2. 考虑 $A = \left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$, 则

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

即 $A^{-1} = \left(\begin{smallmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{smallmatrix}\right)$

注 2.1.5. 更一般的, 那么我们有如下的求 2 阶可逆方阵逆的办法:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2.2 矩阵的转置

定义 2.2.1. 对于矩阵 $A \in M_{m\times n}(\mathbb{R})$, 其**转置矩阵** *(transpose matrix)*是一个 $n \times m$ 阶矩阵 $A^T = (b_{ij})_{n\times m}$, 其中 $b_{ij} = a_{ji}$.

定义 2.2.2. 矩阵 $A \in M_{m\times n}(\mathbb{R})$ 被称为**对称矩阵** *(symmetric matrix)*, 如果 $A^T = A$.

例 2.2.1. 对于列向量来说, 其转置为行向量; 对于行向量来说, 其转置为列向量.

命题 2.2.1. 对于矩阵转置来说, 我们有如下简单的性质:

(1) $(A^T)^T = A$.

(2) $(AB)^T = B^T A^T$.

(3) $AA^T = 0$ 当且仅当 $A = 0$.

证明. 直接验证即可. □

例 2.2.2. 对于方阵 $A \in M_n(\mathbb{R})$, $AA^T = 0$ 等价于 $A = 0$.

推论 2.2.1. 对矩阵 A 做列变换等价于右乘可逆矩阵.

证明. 利用转置矩阵的观点, 对矩阵 A 进行列变换, 等价于对 A^T 进行行变换再转置, 而列变换等价于左乘可逆矩阵, 因此根据命题2.2.1的 (2) 即可. □

回忆定义1.3.3, 我们定义矩阵 A 的秩为其最简行阶梯型的主元个数. 一个自然的问题就是 A^T 的秩与 A 的秩有什么关系呢?[1] 我们可以证明 $\operatorname{rank} A = \operatorname{rank} A^T$, 这主要依赖于下面的定理.

定理 2.2.1. 列变换不改变矩阵 A 的秩.

证明. 假设 $A \in M_{m\times n}(\mathbb{R})$, 根据推论2.2.1, 我们只需要对可逆矩阵 $B \in M_n(\mathbb{R})$ 证明 $\operatorname{rank} A = \operatorname{rank}(AB)$ 即可. 我们不妨记 $\operatorname{rank} A = k, \operatorname{rank}(AB) = l$. 根据线性方程组解的理论可知

1. $Ax = 0$ 有主元 $x_{i_1}, \ldots, x_{i_k}$ 以及自由元 $x_{i_{k+1}}, \ldots, x_{i_n}$.
2. $ABy = 0$ 有主元 $y_{i_1}, \ldots, y_{i_l}$ 以及自由元 $y_{i_{l+1}}, \ldots, y_{l_n}$.

由于 $Ax = 0$ 的解与 $ABy = 0$ 的解之间满足 $x = By$, 由于 B 是可逆矩阵, 根据定理 2.1.1 可知两者解之间存在一一对应, 从而主元与自由元的情况是相同的, 从而 $k = l$. □

[1] 在一些教材中我们这里定义的矩阵的秩又被称为行秩, A^T 的秩被称为 A 的列秩, 即我们要证明矩阵的行秩与列秩相同.

推论 2.2.2. *对于矩阵 $A \in M_{m\times n}(\mathbb{R})$ 来说,* $\operatorname{rank} A = \operatorname{rank} A^T$.

证明. 我们对 A 的最简行阶梯型 $\operatorname{rref} A$ 作列变换, 将其化作如下形式

$$\begin{pmatrix} \mathrm{I}_k & \mathrm{O}_{k\times n-k} \\ \mathrm{O}_{m-k\times k} & \mathrm{O}_{m-k\times n-k} \end{pmatrix}$$

其中 O 代表分量全为零的矩阵. 此时 A 与 A^T 都为最简行阶梯型, 从而 $\operatorname{rank} A = \operatorname{rank} A^T = k$. □

从上述证明过程中, 根据可逆矩阵与行列变换的关系, 我们还能看出:

推论 2.2.3. *对于矩阵 $A \in M_{m\times n}(\mathbb{R})$, 存在可逆矩阵 $P \in M_m(\mathbb{R}), Q \in M_n(\mathbb{R})$ 使得*

$$PAQ = \begin{pmatrix} \mathrm{I}_k & \mathrm{O}_{k\times n-k} \\ \mathrm{O}_{m-k\times k} & \mathrm{O}_{m-k\times n-k} \end{pmatrix}$$

其中 $k = \operatorname{rank} A$, 这被称为 A 的 **相抵标准型** *(canonical form).*

定义 2.2.3. *矩阵 $A, B \in M_{m\times n}(\mathbb{R})$ 之间被称为* **相抵** *(equivalent), 如果存可逆矩阵 $P \in M_m(\mathbb{R}), Q \in M_n(\mathbb{R})$ 使得 $PAQ = B$.*

定理 2.2.2. *$m \times n$ 阶矩阵 A, B 之间相抵当且仅当 $\operatorname{rank} A = \operatorname{rank} B$, 即相抵关系完全由矩阵的秩分类.*

证明. 由于行列变换不改变矩阵的秩, 从而相抵矩阵有相同的秩; 另一方面, 如果 A, B 有相同的秩, 它们的相抵标准型相同, 从而相抵. □

2.3 分块矩阵

一般来说, 当 n 较大时, 求解 $n \times n$ 矩阵的逆对人工操作来说是相对较麻烦的, 但如果矩阵有相对较好的形式, 此时的求解也可以化简. 下面将介绍分块矩阵的想法, 给定矩阵 A, 我们可以做适当的划分, 将其看作矩阵元素是矩阵的矩阵. 例如

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

我们可以将其看成 2×2 的矩阵 $(A_{ij})_{2\times 2}$, 其中

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad A_{12} = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} \quad A_{21} = \begin{pmatrix} a_{31} & a_{32} \end{pmatrix} \quad A_{22} = \begin{pmatrix} a_{33} \end{pmatrix}$$

如果可逆矩阵 A 可以写成分块对角的形式, 即

$$A=\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & A_3 & \\ & & & A_4 \end{pmatrix}$$

那么则有

$$A^{-1}=\begin{pmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & A_3^{-1} & \\ & & & A_4^{-1} \end{pmatrix}$$

同样的, 我们可以对分块矩矩阵进行分块行列变换, 得到相对较好的形式. 例如对于分块矩阵

$$M=\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

其中 A,B,C,D 都是方阵. 如果 A 可逆, 那么

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \mathrm{I} & -A^{-1}B \\ 0 & \mathrm{I} \end{pmatrix}=\begin{pmatrix} A & 0 \\ C & D-CA^{-1}B \end{pmatrix}$$

即通过行列变换将其下三角化, 对于 B,C,D 可逆的时候我们也可以做类似的事情. 特别地是, 如果我们采取不同的变换得到相同的等式, 这有时候可以给我们带来一些非平凡的结果.

例 2.3.1. 对于列向量 α,β, 考虑

$$\begin{pmatrix} \mathrm{I} & \alpha \\ \beta^T & 1 \end{pmatrix}^{-1}$$

一方面我们考虑

$$\begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ \beta^T & 1 & \mathrm{O} & 1 \end{pmatrix}\to\begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ 0 & 1-\beta^T\alpha & -\beta^T & 1 \end{pmatrix}\to\begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ 0 & 1 & -\beta^T(1-\beta^T\alpha)^{-1} & (1-\beta^T\alpha)^{-1} \end{pmatrix}$$
$$\to\begin{pmatrix} \mathrm{I} & 0 & \mathrm{I}+\alpha(1-\beta^T\alpha)^{-1}\beta^T & -(1-\beta^T\alpha)^{-1}\alpha \\ 0 & 1 & -(1-\beta^T\alpha)^{-1}\beta^T & (1-\beta^T\alpha)^{-1} \end{pmatrix}$$

另一方面我们有

$$\begin{pmatrix} \mathrm{I} & \alpha & \mathrm{I} & \mathrm{O} \\ \beta^T & 1 & \mathrm{O} & 1 \end{pmatrix}\to\begin{pmatrix} \mathrm{I}-\alpha\beta^T & 0 & \mathrm{I} & -\alpha \\ \beta^T & 1 & 0 & 1 \end{pmatrix}\to\begin{pmatrix} \mathrm{I} & 0 & (\mathrm{I}-\alpha\beta^T)^{-1} & -(\mathrm{I}-\alpha\beta^T)^{-1}\alpha \\ \beta^T & 1 & 0 & 1 \end{pmatrix}$$
$$\to\begin{pmatrix} \mathrm{I} & 0 & (\mathrm{I}-\alpha\beta^T)^{-1} & -(\mathrm{I}-\alpha\beta^T)^{-1}\alpha \\ 0 & 1 & -\beta^T(\mathrm{I}-\alpha\beta^T)^{-1} & 1+\beta^T(\mathrm{I}-\alpha\beta^T)^{-1}\alpha \end{pmatrix}$$

从而我们有非平凡等式

$$(\mathrm{I}-\alpha\beta^T)^{-1}=\mathrm{I}+\alpha(1-\beta^T\alpha)^{-1}\beta^T$$
$$(1-\beta^T\alpha)^{-1}=1+\beta^T(\mathrm{I}-\alpha\beta^T)^{-1}\alpha$$

引理 2.3.1.

$$\operatorname{rank}\begin{pmatrix}A & C\\ 0 & B\end{pmatrix}\geq\operatorname{rank}(A)+\operatorname{rank}(B)$$

证明. 分别对 A,B 所在的行做初等行变换将其化为其相抵标准型, 则有

$$\begin{pmatrix}A & C\\ 0 & B\end{pmatrix}\to\begin{pmatrix}\mathrm{I}_{r_1} & *\\ 0 & \mathrm{I}_{r_2}\end{pmatrix},$$

其中 $r_1=\operatorname{rank}(A),r_2=\operatorname{rank}(B)$, 从而可以得到期待的不等式. □

命题 2.3.1 (Frobenius 秩不等式)**.** 对于矩阵 $A\in M_{m\times n}(\mathbb{R}),B\in M_{n\times k}(\mathbb{R}),C\in M_{k\times\ell}$, 有

$$\operatorname{rank}(AB)+\operatorname{rank}(BC)\leq\operatorname{rank}(ABC)+\operatorname{rank}(B).$$

证明. 注意到初等行变换不改变矩阵的秩, 从而下面的矩阵有相同的秩:

$$\begin{pmatrix}AB & 0\\ B & BC\end{pmatrix}\to\begin{pmatrix}0 & -ABC\\ B & BC\end{pmatrix}\to\begin{pmatrix}0 & ABC\\ B & 0\end{pmatrix}\to\begin{pmatrix}ABC & 0\\ 0 & B\end{pmatrix}.$$

根据引理2.3.1, 我们有

$$\operatorname{rank}(AB)+\operatorname{rank}(BC)\leq\operatorname{rank}(ABC)+\operatorname{rank}(B).$$

□

例 2.3.2. 对于矩阵 $A\in M_n(\mathbb{R})$, 有 $A^2=\mathrm{I}$ 当且仅当 $\operatorname{rank}(A+\mathrm{I}_n)+\operatorname{rank}(A-\mathrm{I}_n)=n$.

证明. 注意到

$$\begin{pmatrix}\mathrm{I}_n+A & 0\\ 0 & \mathrm{I}_n-A\end{pmatrix}\to\begin{pmatrix}\mathrm{I}_n+A & \mathrm{I}_n+A\\ 0 & \mathrm{I}_n-A\end{pmatrix}\to\begin{pmatrix}\mathrm{I}_n+A & 2\,\mathrm{I}_n\\ 0 & \mathrm{I}_n-A\end{pmatrix}\to\begin{pmatrix}\mathrm{I}_n+A & 2\,\mathrm{I}_n\\ \frac{1}{2}(A^2-\mathrm{I}_n) & 0\end{pmatrix}$$

从而 $A^2=\mathrm{I}$ 当且仅当 $\operatorname{rank}(A+\mathrm{I}_n)+\operatorname{rank}(A-\mathrm{I}_n)=n$. □

注 2.3.1. 类似的技巧可以证明, 如果 $f(x)$ 和 $g(x)$ 是互素多项式, 则对于 $A\in M_n(\mathbb{R})$, 有 $f(A)g(A)=0$ 当且仅当 $\operatorname{rank}(f(A))+\operatorname{rank}(g(A))=n$.

2.4 选讲: 快速傅立叶变换

给定两个 d 次多项式

$$h_1(x) = a_0 + a_1 x + \cdots + a_d x^d$$
$$h_2(x) = b_0 + b_1 x + \cdots + b_d x^d$$

多项式乘积为

$$h(x) = h_1(x)h_2(x) := c_0 + c_1 x + \cdots + c_{2d} x^{2d},$$

其中 $c_k = \sum_{i+j=k} a_i b_j$. 因此如果直接通过定义计算 c_k, 我们需要进行 $k+1$ 次乘法运算, 因此总共需要进行

$$\sum_{k=0}^{2d}(k+1) = \frac{(1+2d+1)(2d+1)}{2} = d(2d+1)$$

次乘法运算, 即复杂度为 $O(d^2)$.

在本节中我们将介绍快速傅立叶变换 (fast Fourier transformation)与快速傅立叶逆变换快速傅立叶逆变换 (inverse fast Fourier transformation), 将计算多项式乘法的复杂度降到 $O(d\log d)$.

注意到 $h(x)$ 是一个 $2d$ 次多项式, 其可以由在 $2d+1$ 个不同的 x 处取值决定, 即我们有如下的矩阵表达式

$$\begin{pmatrix} h(x_0) \\ h(x_1) \\ \vdots \\ h(x_{2d}) \end{pmatrix} = M \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2d} \end{pmatrix} = \begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{2d} \\ 1 & x_1 & x_1^2 & \dots & x_1^{2d} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_d & x_d^2 & \dots & x_d^{2d} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{2d} \end{pmatrix}$$

其中 M 被称为**范德蒙德矩阵** (Vandermond matrix).

如果知道了 $h_1(x_i)$ 与 $h_2(x_i)$ 的这些节点值, 节点值 $h(x_i)$ 可以通过 $h(x_i) = h_1(x_i)h_2(x_i)$ 计算, 这只需要进行 $2d+1$ 次乘法运算, 其时间复杂度为 $O(d)$. 因此关键在于如何以 $O(d\log d)$ 的时间复杂度通过 $h_1(x)$ 与 $h_2(x)$ 的系数来得到节点值 $h_1(x_i), h_2(x_i)$, 以及如何以 $O(d\log d)$ 的时间复杂度通过节点值 $h(x_i)$ 来得到 $h(x)$ 的系数, 这就是快速傅立叶变换与快速傅立叶逆变换要完成的事情.

现在我们对一个一般的 $n-1$ 次多项式 $f(x) = a_0 x + a_1 x + \cdots + a_{n-1}x^{n-1}$ 来介绍快速傅立叶变换与快速傅立叶逆变换. 不失一般性的, 我们可以假设 $n = 2^s$, 因为我们总可以把一个多项式等价的看成是次数更高的多项式, 但是那些高次项的系数是零.

注意到任意给 n 个节点 $x_1, \dots, x_n$, 通过暴力计算得到所有的 $f(x_i)$ 需要 n^2 次乘法, 因此时间复杂度为 $O(n^2)$. 快速傅立叶变换并不选取任意 n 个节点, 而是首先

选取 n 个单位根 ω_n^i，然后按照奇偶性将 $f(x)$ 分成两部分

$$f(x) = (a_0 + a_2x^2 + \cdots + a_{n-2}) + (a_1x + a_3x^3 + \cdots + a_{n-1}x^{n-1}),$$

则

$$f(x) = f_1(x^2) + xf_2(x^2).$$

例 2.4.1. 例如 $h(x) = x^4 + 3x^3 + 2x^2 + x + 1$, 我们将其写作

$$\begin{aligned} f(x) &= (x^4 + 2x^2 + 1) + x(3x^2 + 1) \\ &= f_1(x^2) + xf_2(x^2). \end{aligned}$$

对于任意 $k < \frac{n}{2}$, 我们有

$$f(\omega_n^k) = f_0(\omega_n^{2k}) + \omega_n^k f_2(\omega_n^{2k}) = f_0(\omega_{\frac{n}{2}}^k) + \omega_n^k f_2(\omega_{\frac{n}{2}}^k)$$

以及

$$f(\omega_n^{k+\frac{n}{2}}) = f_0(\omega_n^{2k+n}) + \omega_n^k f_2(\omega_n^{2k+n}) = f_0(\omega_{\frac{n}{2}}^k) + \omega_n^k f_2(\omega_{\frac{n}{2}}^k).$$

因此如果我们已经知道了 $f_1(\omega_{\frac{n}{2}}^k)$ 与 $f_2(\omega_{\frac{n}{2}}^k)$ 这些节点值, 则可以通过 $O(n)$ 的复杂度得到 $f(\omega_n^k)$ 这些节点值. 而对于 $f_1(x)$ 和 $f_2(x)$ 来说这是规模缩小了一半的子问题, 所以我们不断向下递归分治, 当 $n = 1$ 时结束, 从而可以以 $O(n\log n)$ 的复杂度通过多项式系数得到节点值.

如果我们用矩阵的语言来解释快速傅立叶变换, 这相当于以 $O(n\log n)$ 的复杂度计算了如下矩阵乘法

$$\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_n & \omega_n^2 & \omega_n^3 & \cdots & \omega_n^{n-1} \\ 1 & \omega_n^2 & \omega_n^4 & \omega_n^6 & \cdots & \omega_n^{2(n-1)} \\ 1 & \omega_n^3 & \omega_n^6 & \omega_n^9 & \cdots & \omega_n^{3(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{n-1} & \omega_n^{2(n-1)} & \omega_n^{3(n-1)} & \cdots & \omega_n^{(n-1)(n-1)} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} f(1) \\ f(\omega_n) \\ f(\omega_n^2) \\ f(\omega_n^3) \\ \vdots \\ f(\omega_n^{n-1}) \end{pmatrix}$$

如果想要通过节点值得到多项式系数, 实际上就是要计算如上矩阵的逆矩阵, 但是不难发现

$$W = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega_n & \omega_n^2 & \omega_n^3 & \cdots & \omega_n^{n-1} \\ 1 & \omega_n^2 & \omega_n^4 & \omega_n^6 & \cdots & \omega_n^{2(n-1)} \\ 1 & \omega_n^3 & \omega_n^6 & \omega_n^9 & \cdots & \omega_n^{3(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n^{n-1} & \omega_n^{2(n-1)} & \omega_n^{3(n-1)} & \cdots & \omega_n^{(n-1)(n-1)} \end{pmatrix}$$

的逆矩阵为

$$W^{-1} = \frac{1}{n}\overline{W}.$$

这意味着, 如果想要通过 $f(x)$ 的节点值得到 $f(x)$ 的系数, 只需要将这些节点值作为多项式 $g(x)$ 的系数, 然后取 $\omega_n^0, \omega_n^{-1}, \ldots, \omega_n^{-n+1}$ 作为节点进行快速傅立叶变换, 得到的节点值 $g(\omega_n^{-i})$ 就是 $f(x)$ 的系数.

2.5 作业三

2.5.1 基础题

本次作业中的矩阵均为实矩阵.

习题 2.5.1. 计算矩阵乘法:

1. $\begin{bmatrix} 1 & 4 & 7 & 9 \\ -3 & 3 & 8 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 & 6 & 7 \\ 9 & -2 & 0 \\ 11 & -1 & -3 \\ 1 & 0 & 0 \end{bmatrix}$;

2. $\begin{bmatrix} X & 1 & 0 \\ X^2+X & 2 & 0 \\ 0 & X & X-1 \end{bmatrix} \begin{bmatrix} -1 & X & -X \\ 8 & -X-2 & -2 \\ 0 & 0 & 1 \end{bmatrix}$.

3. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ $(\theta, \varphi \in \mathbb{R})$.

4. $\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}^6$. (为什么?)

习题 2.5.2. 设矩阵 A, B 的行数相等. 证明: 存在矩阵 X 使得 $AX = B$ 当且仅当 $\operatorname{rank}(A) = \operatorname{rank}((A, B))$. (其中 (A, B) 表示将两个矩阵拼接得到的矩阵.)

习题 2.5.3. 设有 n 个矩阵 $A^{(1)}, \cdots, A^{(n)}$(注意, 此处上标不是乘方), 其大小未知, 但满足乘积 $P = A^{(1)}A^{(2)}\cdots A^{(n)}$ 有意义. 记 $A^{(k)}$ 的第 i 行第 j 个元素为 $a_{ij}^{(k)}$, P 的第 i 行第 j 个元素为 p_{ij}. 请用 $a_{ij}^{(k)}$ 表示 p_{ij}.

提示: 答案并不复杂. $n = 2$ 时的答案为

$$p_{ij} = \sum_k a_{ik}^{(1)} a_{kj}^{(2)}.$$

习题 2.5.4. 设 $G = (V, E)$ 是一个图, 顶点集 $V = \{1, 2, \cdots, n\}$. n 阶矩阵 A 是 G 的邻接矩阵, 即 A 的元素 a_{ij} 等于顶点 i, j 之间边的数量.

证明 A^k 的第 i 行第 j 个元素等于 i, j 之间长度为 k 的道路的数量. (所谓 i, j 之间长度为 k 的道路, 是指 V 的一列元素 $i = v_0, v_1, \cdots, v_k = j$ 和 E 的一列元素 $e_1, \cdots, e_k$, 满足 e_h 的顶点为 v_{h-1}, v_h.)

提示: 使用问题2.5.3的结果.

习题 2.5.5. 证明: 与所有 n 阶方阵均可交换的 n 阶方阵必为纯量方阵, 即形如 $\lambda I_n, \lambda \in \mathbb{R}$.

习题 2.5.6. 对 $n \times n$ 矩阵 $X = (x_{ij})_{1 \le i,j \le n}$, 定义其 "迹" 为

$$\mathrm{tr}(X) = x_{11} + x_{22} + \cdots + x_{nn}.$$

1. 设 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵, 证明 $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

2. 证明不存在 $n \times n$ 的矩阵 A, B 使得 $AB - BA = I_n$.

2.6 作业四

2.6.1 基础题

本部分题必做.

习题 2.6.1. 考虑如下 2 阶方阵的集合 $M = \{A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mid a, b \in \mathbb{R}\}$.

1. 请证明 M 在矩阵的加法，数乘和乘法下封闭.
2. 请证明 M 上的乘法满足交换律，而且 M 中的任何非零矩阵均可逆，且逆矩阵也在 M 中.

习题 2.6.2. 计算如下矩阵的逆矩阵:

1. $\begin{bmatrix} 1 & a & z \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}$;
2. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}, ad - bc \neq 0$;
3. $\begin{bmatrix} 17 & 8 & 3 \\ 2 & 3 & 1 \\ 0 & 8 & 2 \end{bmatrix}$. 利用你计算的结果解方程 $\begin{bmatrix} 17 & 8 & 3 \\ 2 & 3 & 1 \\ 0 & 8 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 18 \\ 36 \\ 0 \end{bmatrix}$.

习题 2.6.3. 回顾第一次课里介绍的 *Google* 的 *PageRank* 算法. 对任意一个有向图 G, 其对应的线性方程组是否一定有非零解?

提示: 这个方程可以表示为 $Ax = 0$, A 是某个方阵. 考虑 A^T 以及方程 $A^T y = 0$.

习题 2.6.4. 证明线性方程组 $A\mathbf{x} = \mathbf{b}$ 有解当且仅当 $\begin{bmatrix} A^T \\ \mathbf{b}^T \end{bmatrix} \mathbf{y} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$ 无解.

习题 2.6.5. 令 $A = (a_{ij})_{1 \leq i,j \leq n}$ 为 n 阶实矩阵。证明若 $\forall 1 \leq i \leq n, |a_{ii}| > \Sigma_{j \neq i} |a_{ij}|$, 则 A 可逆.

提示: 利用 $Ax = 0$ 是否有非零解的判定法则.

习题 2.6.6. 证明 $n \times m$ 的实矩阵 A 的 *rank* 小于或等于 1 等价于存在 n 维列向量 α 和 m 维列向量 β 使得 $A = \alpha \cdot \beta^T$.

2.6.2 思考题

习题 2.6.7. 设 $A = (a_{ij})_{1 \leq i,j \leq n}$ 是一个可逆方阵. A 的一个 LDU 分解指 $A = LDU$, 其中 L 是一个主对角线均为 1 的下三角矩阵, U 是一个主对角线均为 1 的上三角矩

阵, D 是一个可逆的对角矩阵. 记 $A_m = (a_{ij})_{1\le i,j\le m}$. 证明,$A$ 存在 LDU 分解当且仅当对每个 $1 \le m \le n$, A_m 都可逆, 并且当 A 存在 LDU 分解时, 其 LDU 分解是唯一的. (注: 这 A_m 称做 A 的顺序主子阵。请用这个结论说服自己, "大部分" 实矩阵都具有 LDU 分解, 思考如何来定义 "大部分".)

习题 2.6.8. 令 $\mathrm{GL}(n,\mathbb{R})$ 是域 $\mathbb{R}$ 上的 n-阶可逆矩阵全体, B 是上三角矩阵全体. 记 S_n 是每行每列有且仅有一个 1 的 n-矩阵全体. 对任意 $w \in S_n$, 定义 $\mathrm{GL}(n,\mathbb{R})$ 的子集为

$$BwB = \{A_1 \cdot w \cdot A_2 \mid A_1 \in B, A_2 \in B\}.$$

证明 $\mathrm{GL}(n,\mathbb{R})$ 是所有 BwB 的无交并.

习题 2.6.9. 称 n 阶实方阵 A 是幂零矩阵, 如果 $A^k = 0$ 对某个正整数 k 成立. 证明

1. 若 A 是 n 阶幂零矩阵, 则 $I + A$ 可逆.
2. 假设 $I - X$ 是 n 阶幂零方阵, B 是 n 阶实方阵, 且 $X^m B = BX^m$ 对某一个正整数 m 成立. 请问是否一定有 $XB = BX$ 成立? 如果是请证明, 如果不是请给出反例.

2.7 作业五

2.7.1 基础题

本部分题必做.

习题 2.7.1. 假设 A 是可逆矩阵，u, v 是列向量。证明：

1. $A + uv^T$ 可逆当且仅当 $1 + v^T A^{-1} u \neq 0$。
2. $A + uv^T$ 可逆时有 $(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1+v^T A^{-1}u}$

习题 2.7.2. 求下述 n 阶矩阵的逆：

$$\begin{pmatrix} 1+a_1 & 1 & \cdots & 1 \\ 1 & 1+a_2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1+a_n \end{pmatrix}$$

(其中 $a_i \neq 0$.)

习题 2.7.3. 假设 $A \in M_{n\times m}(\mathbb{R}), B \in M_{m\times n}(\mathbb{R})$. 证明 $I_n + AB$ 可逆当且仅当 $I_m + BA$ 可逆.

习题 2.7.4. 设 A, B 为 $2^n \times 2^n$ 矩阵，我们想要计算乘积 $C = AB$。根据维基百科，“*Strassen* 算法将 A, B 和 C 分割成大小相等的块矩阵

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}, \quad C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

其中 $A_{ij}, B_{ij}, C_{ij} \in \mathrm{M}_{2^{n-1}\times 2^{n-1}}(\mathbb{R})$。朴素算法如下：

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} A_{11}\times B_{11} + A_{12}\times B_{21} & A_{11}\times B_{12} + A_{12}\times B_{22} \\ A_{21}\times B_{11} + A_{22}\times B_{21} & A_{21}\times B_{12} + A_{22}\times B_{22} \end{bmatrix}$$

这种构造并没有减少乘法的数量：仍然需要 *8* 次矩阵块的乘法来计算 C_{ij} 矩阵，这与使用标准矩阵乘法所需的乘法数量相同。*Strassen* 算法定义了新的值：

$$\begin{aligned} M_1 &= (A_{11} + A_{22}) \times (B_{11} + B_{22}); \\ M_2 &= (A_{21} + A_{22}) \times B_{11}; \\ M_3 &= A_{11} \times (B_{12} - B_{22}); \\ M_4 &= A_{22} \times (B_{21} - B_{11}); \\ M_5 &= (A_{11} + A_{12}) \times B_{22}; \\ M_6 &= (A_{21} - A_{11}) \times (B_{11} + B_{12}); \\ M_7 &= (A_{12} - A_{22}) \times (B_{21} + B_{22}), \end{aligned}$$

使用 *7* 次乘法（每个 M_k 一次）而不是 *8* 次。现在我们可以用 M_k 来表达 C_{ij}：

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} M_1 + M_4 - M_5 + M_7 & M_3 + M_5 \\ M_2 + M_4 & M_1 - M_2 + M_3 + M_6 \end{bmatrix}$$

”递归地这样做，比较朴素算法和 *Strassen* 算法需要的两个数乘法的次数。让自己相信 *Strassen* 算法更高效。（还有一个更快的算法，由 *Coppersmith-Winograd* 提出）。

2.7.2 思考题

本部分题不用交.

习题 2.7.5. 设 A, B 是 $\mathbb{R}$ 上的 $m \times n$ 矩阵，$\mathrm{rank}(A) = r, \mathrm{rank}(B) = s$，并且 $\mathrm{rank}(A+B) = r+s$ 证明：存在 m 阶可逆矩阵 P 与 n 阶可逆矩阵 Q 使得

$$PAQ = \begin{pmatrix} I_r & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \quad PBQ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & I_s & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

3 线性空间与线性映射

3.1 $\mathbb{R}$-线性空间

3.1.1 $\mathbb{R}^n$ 的子空间

回忆对于线性方程组 $Ax = 0$ 的解构成的集合, 其满足:

(1) 加法封闭.

(2) 数乘封闭.

定义 3.1.1. 将满足 *(1)* 和 *(2)* 的 $\mathbb{R}^n$ 的非空子集称为 $\mathbb{R}^n$ 的**子空间** *(subspace)*.

例 3.1.1. $\{0\}$ 是子空间[2], 称为零空间 *(zero space)*.

命题 3.1.1. $\mathbb{R}^n$ 的任何一个子空间 W 均包含零空间.

证明. 任取 $w \in W$, 由于 W 对于数乘封闭, 那么 $0 = 0w \in W$. □

注 3.1.1. 零空间是在包含关系下最小的子空间.

定义 3.1.2. 给定矩阵 $A \in M_{m\times n}(\mathbb{R})$, 齐次线性方程组 $Ax = 0$ 的解是 $\mathbb{R}^n$ 的子空间, 被称为解空间, 也被称为 A 的**核** *(kernel)*, 记为 $\ker A$.

例 3.1.2. 如果 A 是 3×3 阶矩阵, 根据线性方程组解的结构定理, 即定理*1.4.1*, 我们可以发现 $\ker A$ 在 $\mathbb{R}^3$ 中的形式与 $\operatorname{rank} A$ 关系密切:

(1) $\operatorname{rank} A = 0$, 此时 A 是零矩阵, 从而 $\ker A = \mathbb{R}^3$.

(2) $\operatorname{rank} A = 1$, 此时 $\ker A$ 是通过原点的平面.

(3) $\operatorname{rank} A = 2$, 此时 $\ker A$ 是通过原点的直线.

(4) $\operatorname{rank} A = 3$, 此时 A 可逆, 从而 $\ker A = \{0\}$.

定义 3.1.3. 给定 $v_1, \dots, v_n \in \mathbb{R}^m$, 则

$$\operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_n\} := \{x_1v_1 + \cdots + x_nv_n \mid x_i \in \mathbb{R}\}$$

称为 $v_1, \dots, v_n$ 的**线性生成** *(linearly combination)*, 可以直接验证 $\operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$ 是子空间.

注 3.1.2. 更一般的, 我们也可以定义无穷多个向量的线性生成, 假设 $\{v_i \mid v_i \in \mathbb{R}^n\}_{i\in I}$ 是一族由指标集 I 为下标的列向量, 其中 $v_i \in \mathbb{R}^n$, 则定义

$$\operatorname{span}_{\mathbb{R}}\{v_i\}_{i\in I} := \{x_{i_1}v_{i_1} + \cdots + x_{i_k}v_{i_k} \mid x_{i_1}, \dots, x_{i_k} \in \mathbb{R}, k \in \mathbb{Z}_{\geq 0}, i_1, \dots, i_k \in I\}.$$

换言之, 即便指标集 I 是无限集, 我们依然只考虑有限线性生成.

[2]这里的 0 指代零向量, 之后可能会用 0 即代指实数零又代指零向量, 请读者注意自己仔细区分.

注 3.1.3. 现在我们来看一下线性方程组与线性生成的关系: 给定 $A=(v_1,\dots,v_n)\in M_{m\times n}(\mathbb{R})$, 那么 $Ax=b$ 有解当且仅当

$$b\in \operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_n\}.$$

这个时候也称 $\operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_n\}$ 是 A 的**列空间** *(column space)*, 记做 $\operatorname{im} A$. 类似的, 我们也可以定义其**行空间** *(row space)*, 记做 $\operatorname{im} A^T$.

注 3.1.4. 以上我们介绍了两种生成 $\mathbb{R}^n$ 的线性子空间的办法: 考虑矩阵的核以及线性生成. 现在我们来考虑一下这两种操作之间的联系: 给定矩阵 $A\in M_{m\times n}(\mathbb{R})$, 根据定理*1.4.1*, 线性方程组 $Ax=0$ 的解可以由自由元 $x_{i_1},\dots,x_{i_k}$ 以及主元给出, 其中 $k=n-\operatorname{rank}(A)$, 并且在自由元确定后, 主元被唯一确定, 从而

$$\ker A=\operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_k\},$$

其中 v_n 是只有第 i_n 位置为 1, 其余位置为 0 的列向量. 反之, 任取列向量 $\{v_1,\dots,v_k\}\subset\mathbb{R}^n$, 是否存在矩阵 $A\in M_{m\times n}(\mathbb{R})$ 使得 $\operatorname{span}_{\mathbb{R}}\{v_1,\dots,v_k\}=\ker A$? 答案也是肯定的.

例 3.1.3. 在 $\mathbb{R}^3$ 中考虑如下的线性生成

$$\operatorname{span}_{\mathbb{R}}\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix}\}.$$

则

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}\in \operatorname{span}_{\mathbb{R}}\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix}\}$$

当且仅当如下关于 y_1,y_2 的线性方程组

$$\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}\begin{pmatrix}y_1\\y_2\end{pmatrix}=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

有解. 根据定理*1.4.1*, 可知这等价于

$$\operatorname{rank}\begin{pmatrix}1&4\\2&5\\3&6\end{pmatrix}=\operatorname{rank}\begin{pmatrix}1&4&x_1\\2&5&x_2\\3&6&x_3\end{pmatrix}.$$

根据初等行变换可知上述条件等价于 $x_3-2x_2+x_1=0$, 即取 $A=(1,-2,1)$ 即可.

命题 3.1.2. *列变换不改变 A 的列空间.*

证明. 不妨将 A 写作 $A = (v_1, \dots, v_n)$, 根据推论2.2.1, 做列变换等价于右乘可逆矩阵 B, 因此不妨记 A 做列变换得到的矩阵为 $AB = (\overline{v}_1, \dots, \overline{v}_n)$. 根据矩阵乘法的定义, 我们有 $\overline{v}_1, \dots, \overline{v}_n$ 都是 $v_1, \dots, v_n$ 的线性组合, 从而

$$\mathrm{span}_{\mathbb{R}}\{\overline{v}_1, \dots, \overline{v}_n\} \subseteq \mathrm{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$$

并且由于列变换是可逆的, 即 A 也可以由 AB 做列变换得到, 从而

$$\mathrm{span}_{\mathbb{R}}\{v_1, \dots, v_n\} \subseteq \mathrm{span}_{\mathbb{R}}\{\overline{v}_1, \dots, \overline{v}_n\}$$

即列变换不改变 A 的列空间. □

推论 3.1.1. *行变换不改变 A 的行空间.*

证明. 证明同上. □

3.1.2 $\mathbb{R}$-线性空间

在实际应用中, 我们不仅会遇见线性方程组, 同时也会遇见微分方程. 例如当求解 $[0, \infty)$ 上的 $\mathbb{R}$ 值光滑函数 $f(t)$ 满足一维的自由弹簧方程

$$f''(t) + f(t) = 0$$

时, 会发现其解的结构为

$$f(t) = x_1 \sin t + x_2 \cos t, \quad x_1, x_2 \in \mathbb{R}.$$

而有外力的弹簧方程

$$f''(t) + f(t) = \sin t$$

的解的结构为

$$f(t) = -\frac{1}{2} t \cos t + x_1 \sin t + x_2 \cos t, \quad x_1, x_2 \in \mathbb{R}.$$

如果我们依然想用线性空间的语言去描述其解的结构, 我们则需要引入更一般的线性空间的定义.

定义 3.1.4. **一个$\mathbb{R}$-线性空间** *($\mathbb{R}$-vector space)*由非空集合 V 以及以下结构给出:

(1) *加法:* $V \times V \to V, (v_1, v_2) \mapsto v_1 + v_2$;

(2) *数乘:* $\mathbb{R} \times V \to V, (c, v) \mapsto cv$;

满足

(a) 加法满足交换律结合律, 并且存在零元素 0, 逆元素;

(b) 数乘满足结合律, 并且存在单位元 1;

(c) 加法和数乘满足分配律.

定义 3.1.5. 令 V 是 $\mathbb{R}$-线性空间, 子集 $W \subseteq V$ 称为 V 的一个 $\mathbb{R}$-**子空间**, 如果 W 在 V 的加法和数乘运算下封闭.

例 3.1.4. $\mathbb{R}^n$ 以及 $\mathbb{R}^n$ 中的子空间都是 $\mathbb{R}$-线性空间.

例 3.1.5. 全体 $\mathbb{R}$-系数多项式 $\mathbb{R}[x]$ 构成了 $\mathbb{R}$-线性空间.

例 3.1.6. 全体次数小于等于 n 的 $\mathbb{R}$-系数多项式, 记做 $\mathbb{R}[x]_{\leq n}$ 构成了 $\mathbb{R}$-线性空间, 但次数等于 n 的 $\mathbb{R}$ 系数多项式全体不构成 $\mathbb{R}$-线性空间.

例 3.1.7. $C^\infty([a,b]) = \{f(x) \mid f(x)$是定义在 $[a,b]$ 上的光滑函数$\}$ 构成了 $\mathbb{R}$-线性空间.

例 3.1.8. $M_{m\times n}(\mathbb{R})$ 构成了 $\mathbb{R}$-线性空间, 但 n 阶可逆方阵全体 $\mathrm{GL}_n(\mathbb{R})$ 不构成 $\mathbb{R}$-线性空间, 从而也不是 $M_n(\mathbb{R})$ 的子空间.

命题 3.1.3. 加法零元唯一.

证明. 假设 $0, 0'$ 都是加法的零元, 从而根据定义有

$$0 + 0' = 0, \quad 0' + 0 = 0,$$

从而 $0 = 0'$. □

命题 3.1.4. 加法逆元唯一.

证明. 任取 $v \in V$, 如果 $v + w_1 = v + w_2 = 0$, 那么

$$w_1 = w_1 + (v + w_2) = (w_1 + v) + w_2 = w_2.$$

□

命题 3.1.5. $0v = 0$.

证明. 根据分配律有

$$0v = (0 + 0)v = 0v + 0v.$$

两侧同时加上 $-(0v)$ 则有

$$-(0v) + (0v + 0v) = (-0v) + 0v$$

从而有 $0v = 0$. □

3.2 线性相关性

例 3.2.1. 令

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}.$$

注意到 $v_3 = 2v_2 - v_1$, 从而

$$\begin{aligned} a_1v_1 + a_2v_2 + a_3v_3 &= a_1v_1 + a_2v_2 + a_3(2v_2 - v_1) \\ &= (a_1 - a_3)v_1 + (a_2 + 2a_3)v_2. \end{aligned}$$

即

$$\mathrm{span}_{\mathbb{R}}\{v_1, v_2\} = \mathrm{span}_{\mathbb{R}}\{v_1, v_2, v_3\},$$

换言之, 在考虑 v_1, v_2, v_3 的线性生成的时候, v_3 是多余的信息.

定义 3.2.1. $v_1, \ldots, v_n \in \mathbb{R}^m$ 被称为**线性无关** *(linearly independent)*, 如果 0 表示为 $v_1, \ldots, v_n$ 的线性组合的方式唯一, 即只有

$$0 = 0v_1 + \cdots + 0v_n$$

否则 $v_1, \ldots, v_n$ 被称为**线性相关** *(linearly dependent)*.

注 3.2.1. 给定 $\mathbb{R}$-线性空间 V, 对于无穷个向量 $\{v_i\}_{i \in I} \subseteq V$, 其线性无关性定义为: 任意有限和 $a_{i_1}v_{i_1} + \cdots + a_{i_n}v_{i_n} = 0$ 当且仅当 $a_{i_1} = \cdots = a_{i_n} = 0$.

例 3.2.2.

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

是线性无关的: 根据定义, $\{v_1, v_2\}$ 线性无关当且仅当

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

有唯一解 $(0, 0)$, 这等价于

$$\mathrm{rank} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} = 2.$$

定理 3.2.1. $v_1, \ldots, v_n \in \mathbb{R}^m$ 线性无关当且仅当 $\mathrm{rank}\, A = n$, 其中 $A = (v_1, \ldots, v_n)$ 是 $m \times n$ 阶矩阵, 此时也称 A 是列满秩的.

证明. 注意到 $v_1, \dots, v_n$ 线性无关当且仅当 $Ax = 0$ 只有零解, 根据定理1.4.1可知这当且仅当 $\operatorname{rank} A = n$. □

推论 3.2.1. $\mathbb{R}^m$ 中 k 个列向量当 $k > m$ 时一定线性相关.

证明. $\mathbb{R}^m$ 中 k 个列向量组成的矩阵 A 的秩在 $k > m$ 时最大为 m. □

推论 3.2.2. 列变换不改变矩阵列向量的线性相关性.

证明. 根据定理2.2.1, 列变换不改变矩阵的秩, 从而不改变列向量的线性相关性. □

注 3.2.2. 行变换会改变矩阵列向量的线性相关性, 试着举例说明.

命题 3.2.1. 若 $v_1, \dots, v_n \in \mathbb{R}^m$ 线性无关, 则对于 $v \in \operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$, 其被写成 $v_1, \dots, v_n$ 线性组合式子的系数是唯一的.

证明. 不妨假设

$$\begin{aligned} v &= a_1 v_1 + \cdots + a_n v_n \\ &= b_1 v_1 + \cdots + b_n v_n \end{aligned}$$

则

$$0 = (a_1 - b_1)v_1 + \cdots + (a_n - b_n)v_n$$

根据线性无关的定义可知 $a_i = b_i$ 对任意的 $1 \le i \le n$ 成立. □

定理 3.2.2. 假设 $v_1, \dots, v_n$ 线性无关, 则 $v_1, \dots, v_n, v_{n+1}$ 线性相关等价于 $v_{n+1} \in \operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$.

证明. 如果 $v_{n+1} \in \operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$, 则显然 $v_1, \dots, v_{n+1}$ 线性相关. 另一方面, 假设

$$a_1 v_1 + \cdots + a_{n+1} v_{n+1} = 0$$

的系数 $a_1, \dots, a_{n+1}$ 不全为零, 那么一定有 $a_{n+1} \neq 0$, 否则有

$$a_1 v_1 + \cdots + a_n v_n = 0$$

并且 $a_1, \dots, a_n$ 不全为零, 这与线性无关相矛盾, 从而

$$v_{n+1} = -a_{n+1}^{-1}(a_1 v_1 + \cdots + a_n v_n) \in \operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_n\}$$

□

定义 3.2.2. 对于 $v_1, \dots, v_n \in \mathbb{R}^m$, 称 $v_{i_1}, \dots, v_{i_k}$ 是**极大线性无关组** *(maximal linearly independent set)*, 如果

(1) $v_{i_1}, \dots, v_{i_k}$ 线性无关.

(2) 任何包含 $v_{i_1},\dots,v_{i_k}$ 的 $\{v_1,\dots,v_n\}$ 的子集中的向量都线性相关.

注 3.2.3. 若 $v_1,\dots,v_n$ 不全为零, 则其一定存在极大线性无关组: 假设 $v_1\neq 0$, 考虑 v_1,v_2 是否线性相关, 如线性相关则剔除 v_2, 线性无关则保留 v_2. 再依次考虑 $v_3,v_4,\dots$ 即可.

定理 3.2.3. 对于 $v_1,\dots,v_n\in\mathbb{R}^m$, $v_{i_1},\dots,v_{i_k}$ 是其极大线性无关组, 则

$$\mathrm{span}_{\mathbb{R}}\{v_1,\dots,v_n\}=\mathrm{span}_{\mathbb{R}}\{v_{i_1},\dots,v_{i_k}\}$$

证明. 显然 $\mathrm{span}_{\mathbb{R}}\{v_{i_1},\dots,v_{i_k}\}\subseteq\mathrm{span}_{\mathbb{R}}\{v_1,\dots,v_n\}$, 并且根据定理以及极大线性无关组的定义可知任取 $i\in\{1,\dots,n\}\backslash\{i_1,\dots,i_k\}$ 有

$$v_i\in\mathrm{span}_{\mathbb{R}}\{v_{i_1},\dots,v_{i_k}\}$$

从而 $\mathrm{span}_{\mathbb{R}}\{v_1,\dots,v_n\}=\mathrm{span}_{\mathbb{R}}\{v_{i_1},\dots,v_{i_k}\}$. □

对于一组向量来说, 其极大线性无关组可能有很多, 但是任何两个极大线性无关组中向量的个数是一样的, 极大线性无关组不仅仅是包含意义下极大, 也是绝对数目的极大.

定理 3.2.4. 假设 $v_1,\dots,v_n$ 的某个极大线性无关组中有 k 个向量, 则对于任意 $v_{j_1},\dots,v_{j_l}$, 其中 $l>k$, 其线性相关.

证明. 假设 $v_{i_1},\dots,v_{i_k}$ 是 $v_1,\dots,v_n$ 的一个极大线性无关组, 任取 $l>k$ 以及 $v_{j_1},\dots,v_{j_l}$, 根据极大线性无关组的定义有

$$(v_{j_1},\dots,v_{j_l})=(v_{i_1},\dots,v_{i_k})A$$

其中 $A\in M_{k\times l}(\mathbb{R})$, 从而

$$x_1v_{j_1}+\cdots+x_lv_{j_l}=(v_{j_1},\dots,v_{j_l})\begin{pmatrix}x_1\\x_2\\\vdots\\x_l\end{pmatrix}=(v_{i_1},\dots,v_{i_k})A\begin{pmatrix}x_1\\x_2\\\vdots\\x_l\end{pmatrix}$$

由于 $l>k$, 从而根据线性方程组解的结构定理1.4.1可知 $Ax=0$ 有非零解, 从而 $v_{j_1},\dots,v_{j_l}$ 线性相关. □

推论 3.2.3. $v_1,\dots,v_n$ 的极大线性无关组中向量的数目是确定的.

3.3 基与维数

定义 3.3.1. 令 V 是一个 $\mathbb{R}$-线性空间, 如果 V 中的一组向量满足

(1) $\{v_i\}_{i\in I}$ 线性无关;

(2) $V = \mathrm{span}_{\mathbb{R}}\{v_i\}_{i\in I}$,

则称 $\{v_i\}_{i\in I}$ 是 V 的一组**基** *(basis)*.

注 3.3.1. $\{v_i\}_{\in I}$ 构成了 V 的一组基当且仅当任取 $v \in V$, v 是 $\{v_i\}_{i\in I}$ 的线性组合, 并且组合系数唯一.

定义 3.3.2. 对于 $\mathbb{R}$-线性空间 V, 其基中向量个数被称为 V 的实**维数** *(dimension)*, 记做 $\dim_{\mathbb{R}} V$.

注 3.3.2. $\mathbb{R}$-线性空间 V 的维数是良好定义的: 因为 V 的一组基构成了一个极大线性无关组, 根据引理*3.2.3*可知其中向量个数是不依赖于基的选取的.

例 3.3.1. 对 $\mathbb{R}$-线性空间 $\mathbb{R}^n$,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

是 $\mathbb{R}^n$ 的一组基, 从而 $\dim_{\mathbb{R}} \mathbb{R}^n = n$.

例 3.3.2. 对于 $\mathbb{R}$-线性空间 $M_{m\times n}(\mathbb{R})$, $\{E_{ij}\}$ 构成了的一组基, 其中 E_{ij} *(i,j)* 位置为 1, 其余位置为 0 的矩阵, 从而 $\dim_{\mathbb{R}} M_{m\times n}(\mathbb{R}) = mn$.

例 3.3.3. 对于全体次数小于等于 n 的 $\mathbb{R}$ 系数多项式组成的 $\mathbb{R}$-线性空间 $\mathbb{R}[x]_{\leq n}$, $\{1, x, x^2, \ldots, x^{n-1}\}$ 构成了的一组基, 从而 $\dim_{\mathbb{R}} \mathbb{R}[x]_{\leq n} = n$.

例 3.3.4. 对于一维自由弹簧方程 $f''(t)+f(t) = 0$ 的解构成的 $\mathbb{R}$-线性空间, $\{\sin t, \cos t\}$ 构成了一组基, 从而解空间的实维数为 2.

例 3.3.5. $V = \{\frac{ax^2+bx+c}{x^3-x} \mid a, b, c \in \mathbb{R}\}$ 构成了一个 $\mathbb{R}$-线性空间, 其有如下两组不同的基:

$$B = \{\frac{x^2}{x^3-x}, \frac{x}{x^3-x}, \frac{1}{x^3-x}\}$$
$$C = \{\frac{1}{x}, \frac{1}{x-1}, \frac{1}{x+1}\}$$

对于一般的 $\mathbb{R}$-线性空间 V, 其维数不一定有限, 比如下面的例子:

例 3.3.6. 对于 $\mathbb{R}$ 系数多项式全体组成的 $\mathbb{R}$-线性空间 $\mathbb{R}[x]$, $\{1, x, x^2, \dots\}$ 构成了一组基, 因此 $\dim_{\mathbb{R}} \mathbb{R}[x] = \infty$.

但是在这门课程中, 我们主要关心有限维的线性空间. 在之后, 如果不加特殊说明, 我们总假设线性空间是有限维的.

注 3.3.3. 根据推论*3.2.3*可知 W 的维数是良定义的, 并且如下三条中任意满足两条即可说明 $v_1, \dots, v_k$ 是 W 的基:

(1) $W = \operatorname{span}_{\mathbb{R}}\{v_1, \dots, v_k\}$.

(2) $v_1, \dots, v_k$ 线性无关.

(3) $\dim_{\mathbb{R}} W = k$.

命题 3.3.1. 给定 $\mathbb{R}$-向量空间 W_1, W_2, 满足 $W_1 \subseteq W_2$, 则

(1) $\dim_{\mathbb{R}} W_1 \leq \dim_{\mathbb{R}} W_2$, 并且等号成立当且仅当 $W_1 = W_2$.

(2) W_1 的基可以扩充为 W_2 的基.

证明. (1). 由于 W_1 中的线性无关组一定是 W_2 中的线性无关组, 从而$\dim W_1 \leq \dim W_2$, 等号取得是显然的.

(2). 假设 W_1 的基是 $v_1, \dots, v_k$, W_2 的基是 $w_1, \dots, w_l$. 我们在取 $v_1, \dots, v_k, w_1, \dots, w_l$ 的极大线性无关组的时候仔细一些: 即前 k 个向量取 $v_1, \dots, v_k$, 这是可以做到的, 因为 $v_1, \dots, v_k$ 本身线性无关. 这样取出的极大线性无关组就是由 $v_1, \dots, v_k$ 扩充得到的 W_2 的基. □

定理 3.3.1. 对于矩阵 $A \in M_{m\times n}(\mathbb{R})$, $\operatorname{rank} A$ 是行空间维数, $\operatorname{rank} A^T$ 是列空间维数.

证明. 根据推论3.1.1可知行变换不改变行空间, 因此我们通过行变换将其化作最简行阶梯型, 此时主元所在的行向量构成了行空间的一组基, 因此 $\operatorname{rank} A$ 是行空间维数. 类似的可以证明 $\operatorname{rank} A^T$ 是列空间维数. □

推论 3.3.1. 对于矩阵 $A \in M_{m\times n}(\mathbb{R})$, $\dim_{\mathbb{R}} \operatorname{im} A = \operatorname{rank} A$.

证明. 根据推论2.2.2即可. □

推论 3.3.2. 给定矩阵 $A \in M_{m\times n}(\mathbb{R}), B \in M_{n\times \ell}(\mathbb{R})$, 则

$$\operatorname{rank}(AB) \leq \operatorname{rank}(A).$$

证明. 只需要注意到 $\operatorname{im} AB \subseteq \operatorname{im} A$. □

3.4 向量的坐标表达

定义 3.4.1. 令 V 是 $\mathbb{R}$-线性空间, $B=\{v_1,\dots,v_n\}$ 是其一组基, 则对于 $v\in V$, 有

$$v=a_1v_1+\cdots+a_nv_n$$

其中 $a_i\in\mathbb{R}$. 列向量

$$[v]_B=\begin{pmatrix}a_1\\a_2\\\vdots\\a_n\end{pmatrix}\in\mathbb{R}^n$$

被称为 v 在基 B 下的**坐标** *(coordinate)*.

注 3.4.1. 对于 $\mathbb{R}$-线性空间 V, $v\in V$ 在不同基下的坐标往往有不同的优势: 在例*3.3.5*, 考虑

$$v=\frac{3x^2+x+2}{x^3-x}.$$

我们可以很轻易地写出其在 B 下的坐标

$$[v]_B=\begin{pmatrix}2\\1\\3\end{pmatrix}.$$

然而很难直接写出 $[v]_C$. 另一方面, 基 C 便于积分, 因此如果我们有 $[v]_C$, 则可以很轻易地将 v 的积分写出来. 因此搞清楚同一个向量在不同基下坐标的变换关系是非常有意义的问题.

对于 $\mathbb{R}$-线性空间 V, 给定其一组基 $B=\{v_1,\dots,v_n\}$, 我们可以将其视作矩阵 $B=(v_1,\dots,v_n)$, 此时我们有矩阵乘法的等式

$$v=B[v]_B$$

并且对于另一个基 $C=\{w_1,\dots,w_n\}$, 任取 $1\le i\le j$, 有

$$w_i=v_1p_{1i}+\cdots+v_np_{ni}$$

如果我们记矩阵 $P_{B\leftarrow C}=(p_{ij})_{n\times n}$, 并称其为基的**转移矩阵** (transition matrix), 则有

$$C=BP_{B\leftarrow C}$$

命题 3.4.1. 对于转移矩阵, 我们有如下简单的性质.

(1) $[v]_B=P_{B\leftarrow C}[v]_C$.

(2) $P_{B_1\leftarrow B_3}=P_{B_1\leftarrow B_2}P_{B_2\leftarrow B_3}$.

(3) $P_{B\leftarrow B}=\mathrm{I}_n$.

(4) $P_{B\leftarrow C}$ 是可逆矩阵, 并且 $P_{B\leftarrow C}^{-1}=P_{C\leftarrow B}$.

例 3.4.1. 在例*3.3.5*中, 对于

$$v=\frac{3t^2+t+2}{t^3-3}$$

在基 B 下的坐标为

$$[v]_B=\begin{pmatrix}2\\1\\3\end{pmatrix}.$$

直接计算有基的转移矩阵为

$$P_{B\leftarrow C}=\begin{pmatrix}-1&0&0\\9&1&-1\\1&1&1\end{pmatrix},\quad P_{C\leftarrow B}=\begin{pmatrix}-1&0&0\\9&1&-1\\1&1&1\end{pmatrix}^{-1}=\begin{pmatrix}-1&0&0\\\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}\end{pmatrix}.$$

从而有

$$[v]_C=P_{C\leftarrow B}[v]_B=\begin{pmatrix}-2\\2\\2\end{pmatrix}.$$

3.5 线性空间的构造

3.5.1 子空间的和与交

给定 $\mathbb{R}$-线性空间 V 以及 $\mathbb{R}$-线性子空间 W_1,W_2, 我们有以下两种得到新的子空间的方法:

(1) $W_1\cap W_2$ 是 V 的 $\mathbb{R}$-线性子空间, 称为 W_1 与 W_2 的**交** (intersection);

(2)

$$W_1+W_2:=\{w_1+w_2\mid w_1\in W_1,w_2\in W_2\},$$

是 V 的 $\mathbb{R}$-线性子空间, 称为 W_1 与 W_2 的**和** (sum).

例 3.5.1. 考虑 $\mathbb{R}^3=\{(x_1,x_2,x_3)\mid x_i\in\mathbb{R}\}$ 中的子空间 $W_1=\{(x_1,x_2,x_3)\mid x_1=0\},W_2=\{(x_1,x_2,x_3)\mid x_2=0\}$, 则

$$W_1+W_2=\mathbb{R}^3,\quad W_1\cap W_2=\{(x_1,x_2,x_3)\mid x_2=x_3=0\}$$

定理 3.5.1. 给定 $\mathbb{R}$-线性空间 V 以及 $\mathbb{R}$-线性子空间 W_1, W_2. 假设 $\dim V_i < \infty$, 则

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim W_1 \cap W_2.$$

证明. 假设 $W_1 \cap W_2$ 有一组基 $v_1, \ldots, v_k$, 并且 $v_1, \ldots, v_k$ 添加 $u_1, \ldots, u_n$ 扩充成 W_1 的一组基, 添加 $w_1, \ldots, w_m$ 扩充成 W_2 的一组基, 则我们断言

$$v_1, \ldots, v_k, u_1, \ldots, u_n, w_1, \ldots, w_m$$

构成了 $W_1 + W_2$ 的一组基.

首先根据 W_1+W_2 的定义不难发现 W_1+W_2 可以由 $v_1, \ldots, v_k, u_1, \ldots, u_n, w_1, \ldots, w_m$ 生成, 因此只需要证明其线性无关性. 假设

$$a_1v_1 + \cdots + a_kv_k + b_1u_1 + \cdots + b_nu_n + c_1w_1 + \cdots + c_mw_m = 0.$$

令

$$r_1 = a_1v_1 + \cdots + a_kv_k + b_1u_1 + \cdots + b_nu_n \in W_1$$
$$r_2 = c_1w_1 + \cdots + c_mw_m,$$

从而有 $r_1 = -r_2 \in W_1 \cap W_2$, 不妨记

$$r_1 = d_1v_1 + \cdots + d_kv_k,$$

则

$$(a_1 - d_1)v_1 + \cdots + (a_k - d_k)v_k + b_1u_1 + \cdots + b_nu_n = 0.$$

根据 $v_1, \ldots, v_k, u_1, \ldots, u_n$ 的线性无关性有 $b_1 = \cdots = b_n = 0$, 从而有

$$a_1v_1 + \cdots + a_kv_k + c_1w_1 + \cdots + c_mw_m = 0,$$

再利用 $v_1, \ldots, v_k, w_1, \ldots, w_m$ 的线性无关性有 $a_1 = \cdots = a_k = c_1 = \cdots = c_m = 0$. □

注 3.5.1. 给定 $\mathbb{R}$-线性空间 V 以及 $\mathbb{R}$-线性子空间 W_1, W_2, W_3, 以下公式不成立:

$$\dim(W_1+W_2+W_3) = \sum_{i=1}^{3} \dim W_i - \dim W_1 \cap W_2 - \dim W_1 \cap W_3 - \dim W_2 \cap W_3 + \dim W_1 \cap W_2 \cap W_3.$$

另一方面, 如下等式一般也不成立

$$(W_1 + W_2) \cap W_3 = W_1 \cap W_3 + W_2 \cap W_3.$$

考虑 $\mathbb{R}^2$ 中 $W_1 = \operatorname{span}\{\binom{1}{0}\}, W_2 = \operatorname{span}\{\binom{0}{1}\}, W_3 = \operatorname{span}\{\binom{1}{1}\}$, 则以上两条均不满足.

3.5.2 直和

定义 3.5.1. 给定 $\mathbb{R}$-线性空间 V, W, 定义线性空间的**外直和** *(external direct sum)*为集合 $V \times W$, 其上带有结构

(1) $(v_1, w_1) + (v_2, w_2) := (v_1 + v_2, w_1 + w_2)$, 其中 $v_1, v_2 \in V, w_1, w_2 \in W$.

(2) $c(v, w) := (cv, cw)$, 其中 $v \in V, w \in W, c \in \mathbb{R}$.

使得其成为一个 $\mathbb{R}$-线性空间, 记做 $V \oplus W$.

例 3.5.2. $\mathbb{R}^n$ 可以视作 n 个 $\mathbb{R}$ 的外直和.

定义 3.5.2. 给定 $\mathbb{R}$-线性空间 V 以及其子空间 V_1, V_2, 如果

(1) $V = V_1 + V_2$, 即任何 V 中的向量可以表示成 V_1, V_2 中向量的组合.

(2) $V_1 \cap V_2 = \{0\}$.

则称 V 是 V_1 和 V_2 的**内直和** *(internal direct sum)*.

命题 3.5.1. 给定 $\mathbb{R}$-线性空间 V 以及其子空间 V_1, V_2, 如果 V 是 V_1 和 V_2 的内直和, 那么

$$T : V_1 \oplus V_2 \to V_1 + V_2$$
$$(v_1, v_2) \mapsto v_1 + v_2$$

是线性同构, 即 $V \cong V_1 \oplus V_2$.

证明. 线性映射 T 显然是满射, 并且根据内直和的定义有 $\ker T = V_1 \cap V_2 = \{0\}$. □

注 3.5.2. 这意味着内直和与外直和是一体两面, 因此我们之后并不再区分内外直和, 而统称为直和.

定义 3.5.3. 给定 $\mathbb{R}$-线性空间 V 以及子空间 V_1, 如果存在子空间 V_2 满足 $V = V_1 \oplus V_2$, 则称 V_2 是 V_1 的补空间 *(complement space)*.

注 3.5.3. 根据命题*3.3.1*, 我们总可以将子空间的一组基延拓成全空间的一组基, 因此补空间总是存在的.

命题 3.5.2. $\mathbb{R}$-线性空间 $V = V_1 \oplus \cdots \oplus V_k$ 当且仅当如下两条满足:

(1) $V = V_1 + \cdots + V_k$.

(2) 对任意 $1 \le i \le k$, 有 $V_i \cap \sum_{i \ne j} V_j = \{0\}$.

命题 3.5.3. 给定 $\mathbb{R}$-线性空间 V, W, 有

$$\dim_{\mathbb{R}} V \oplus W = \dim_{\mathbb{R}} V + \dim_{\mathbb{R}} W$$

证明. 假设 $\{v_1, \ldots, v_n\}$ 是 V 的一组基, $\{w_1, \ldots, w_m\}$ 是 W 的一组基, 直接验证 $\{(v_1, 0), \ldots, (v_n, 0), (0, w_1), \ldots, (0, w_m)\}$ 构成了 $V \oplus W$ 的一组基. □

3.5.3 商空间

定义 3.5.4. 给定 $\mathbb{R}$-线性空间 V 以及其子空间 W, **商空间** *(quotient space)*定义为集合 V/W, 其上带有结构:

(1) $(v_1+W)+(v_2+W):=(v_1+v_2)+W$, 其中 $v_1,v_2\in V$.

(2) $c(v+W):=cv+W$, 其中 $c\in\mathbb{R}$.

使得其称为一个 $\mathbb{R}$-线性空间.

注 3.5.4. 注意我们在处理商集的时, 定义的良好性是我们始终要考虑的问题, 即定义不依赖于代表元的选取.

命题 3.5.4. 给定 $\mathbb{R}$-线性空间以及商空间 V/W, 自然投射 *(canonical projection)*

$$\begin{aligned}\pi: V &\to V/W\\ v &\mapsto v+W\end{aligned}$$

是满的线性映射, 并且 $\ker\pi=W$.

命题 3.5.5. 给定 $\mathbb{R}$-线性空间以及商空间 V/W, 有

$$\dim_{\mathbb{R}} V=\dim_{\mathbb{R}} V/W+\dim_{\mathbb{R}} W$$

证明. 假设 $v_1+W,\dots,v_n+W$, $w_1,\dots,w_m$ 分别构成了 V/W 和 W 的基, 直接验证 $v_1,\dots,v_n,w_1,\dots,w_m$ 构成了 V 的一组基. □

3.6 线性映射

3.6.1 定义与例子

定义 3.6.1. 对于 $\mathbb{R}$-线性空间 V,W, 映射 $T\colon V\to W$ 被称为 $\mathbb{R}$-**线性映射** *(linear map)*, 如果

(1) 对任意 $v_1,v_2\in V$ 有 $T(v_1+v_2)=Tv_1+Tv_2$.

(2) 对于任意 $c\in\mathbb{R},v\in V$ 有 $T(cv)=cv$.

例 3.6.1. $\mathbb{R}^n$ 到子空间 W 的投影映射是 $\mathbb{R}$-线性映射.

定义 3.6.2. 给定 $\mathbb{R}$-线性空间 V,W, 全体 V 到 W 的 $\mathbb{R}$ 线性映射组成的集合记做 $\mathrm{Hom}_{\mathbb{R}}(V,W)$.

注 3.6.1. $\mathrm{Hom}_{\mathbb{R}}(V,W)$ 上有自然的加法与 $\mathbb{R}$-数乘结构如下:

(1) $(T_1+T_2)v=T_1v+T_2v$, 其中 $v\in V$.

(2) $(cT)v = cTv$, 其中 $v \in V, c \in \mathbb{R}$.

并且可以直接验证上述结构使得 $\mathrm{Hom}_{\mathbb{R}}(V,W)$ 是 $\mathbb{R}$-线性空间.

命题 3.6.1. 给定 $\mathbb{R}$-线性映射 $T_1 : V_1 \to V_2, T_2 : V_2 \to V_3$, 则 $T_2 \circ T_1 \in \mathrm{Hom}_{\mathbb{R}}(V_1, V_3)$.

3.6.2 线性映射与矩阵

定理 3.6.1. 对于 n 维 $\mathbb{R}$-线性空间 V 以及 m 维 $\mathbb{R}$-线性空间 W, 有如下 $\mathbb{R}$-线性空间的同构

$$M_{m\times n}(\mathbb{R}) \longleftrightarrow \mathrm{Hom}_{\mathbb{R}}(V,W)$$

证明. 考虑映射

$$\begin{aligned} M_{m\times n}(\mathbb{R}) &\to \mathrm{Hom}_{\mathbb{R}}(V,W) \\ A &\mapsto T_A \end{aligned}$$

首先线性性是显然的. 我们固定 V 的一组基 $\{v_1,\dots,v_n\}$ 以及 W 的基 $C = \{w_1,\dots,w_m\}$. 对于矩阵 $A_1, A_2 \in M_{m\times n}(\mathbb{R})$, 注意到 $[T_{A_1}v_i]_C$ 是 A_1 的第 i 列, $[T_{A_2}v_i]_C$ 是 A_2 的第 i 列, 从而如果 $T_{A_1} = T_{A_2}$, 从而 $A_1 = A_2$; 另一方面, 对于 $T \in \mathrm{Hom}_{\mathbb{R}}(V,W)$, 我们可以得到矩阵[3] $[T]_B^C = ([Tv_1]_C,\dots,[Tv_n]_C) \in M_{m\times n}(\mathbb{R})$, 直接验证有

$$\begin{aligned} Tv &= T((v_1,\dots,v_n)[v]_B) \\ &= (Tv_1,\dots,Tv_n)[v]_B \\ &= C([Tv_1]_C,\dots,[Tv_n]_C)[v]_B \\ &= T_{[T]_B^C}v \end{aligned}$$

□

注 3.6.2. 这是一个非常重要的观点, 即当我们取定线性空间 V 的一组基之后, 根据命题3.6.4我们可以将 V 看作 $\mathbb{R}^n$, 并且根据定理3.6.1我们可以将线性映射看作矩阵, 因此线性空间上线性映射的问题可以被转化成矩阵问题来解决, 这更加的具体, 以及可计算.

注意到定理3.6.1中的对应是依赖于基的选取的, 那么一个自然的问题是给定一个 $\mathbb{R}$-线性映射, 选取不同的基得到的矩阵之间有什么关系?

命题 3.6.2. 假设 B_1, B_2 都是 $\mathbb{R}$-线性空间 V 的基, C_1, C_2 都是 W 的基, 则对于 $T \in \mathrm{Hom}_{\mathbb{R}}(V,W)$, 有

$$[T]_{B_1}^{C_1} = P_{C_1\leftarrow C_2}[T]_{B_2}^{C_2}P_{B_2\leftarrow B_1}$$

[3] 我们称矩阵 $[T]_B^C$ 为线性映射 T 在基 B, C 下的矩阵.

命题 3.6.3. 假设 $T\colon V\to W$ 是 $\mathbb{R}$-线性空间之间的线性映射, B,C 分别是 V 和 W 的一组基, 则任取 $v\in V$ 有

$$[T(v)]_C=[T]_B^C[v]_B.$$

定义 3.6.3. 矩阵 $A,B\in M_n(\mathbb{R})$ 称为**相似** *(similar)*, 如果存在可逆矩阵 $P\in M_n(\mathbb{R})$ 使得 $PAP^{-1}=B$.

因此, 对于 $\mathbb{R}$-线性映射 $T\colon V\to V$, 其在不同基下的矩阵是相似的, 这也是相似矩阵的几何解释. 如果对于矩阵来说一些量是相似不变量[4], 则我们可以对线性映射定义.

定义 3.6.4. 对于 $\mathbb{R}$-线性映射 $T\colon V\to W$, 其**行列式** *(determinant)*定义为 $\det[T]_B^C$, 其中 B,C 分别是 V,W 的基, 记做 $\det T$.

注 3.6.3. 根据命题*3.7.1*的 *(5)* 即可.

定义 3.6.5. 对于 $\mathbb{R}$-线性映射 $T\colon V\to W$, 其**秩** *(rank)*定义为 $\operatorname{rank}[T]_B^C$, 其中 B,C 分别是 V,W 的基, 记做 $\operatorname{rank}T$.

注 3.6.4. 由于相似矩阵只相差可逆矩阵左乘右乘, 并且左乘右乘可逆矩阵不改变矩阵的秩, 从而线性映射的秩定义是良好的.

例 3.6.2. 考虑 $\mathbb{R}^2$ 上由逆时针旋转 θ 给出的线性映射, 并且记 $B=\{e_1,e_2\}$ 是 $\mathbb{R}^2$ 的一组标准基, 则

$$T_\theta(e_1)=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},\quad T_\theta(e_2)=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}.$$

从而有

$$[T]_B^B=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

根据旋转的几何直观我们有 $T_{\theta_1}\circ T_{\theta_2}=T_{\theta_1+\theta_2}$, 从而有如下等式:

$$\begin{pmatrix}\cos\theta_1 & -\sin\theta_1\\ \sin\theta_1 & \cos\theta_1\end{pmatrix}\begin{pmatrix}\cos\theta_2 & -\sin\theta_2\\ \sin\theta_2 & \cos\theta_2\end{pmatrix}=\begin{pmatrix}\cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2)\\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2)\end{pmatrix}.$$

这证明了和角公式.

例 3.6.3. 考虑如下线性映射

$$\begin{aligned}T\colon \mathbb{R}[x]_{\leq n}&\to\mathbb{R}[x]_{\leq n}\\ f(x)&\mapsto\frac{\mathrm{d}f}{\mathrm{d}x}\end{aligned}$$

[4] 即指如果两个矩阵相似, 它们的这个量相同.

考虑基 $B=\{1,x,\dots,x^n\}$, 则有

$$[T]_B^B=\begin{pmatrix}0&0&\dots&0&0\\1&0&\dots&0&0\\0&2&\dots&0&0\\\vdots&\vdots&&\vdots&\vdots\\0&0&\dots&n&0\end{pmatrix}_{(n+1)\times(n+1)}$$

注 3.6.5. 在上面的例子中, $T^{n+1}=0$, 并且 $\operatorname{rank}T=n$.

3.6.3 线性同构

定义 3.6.6. $\mathbb{R}$-线性映射 T 是**线性同构** *(linear isomorphism)*, 如果其既是双射也是满射.

定义 3.6.7. $\mathbb{R}$-线性空间 V,W 称为**线性同构** *(linear isomorphism)*, 如果存在线性同构 $T\colon V\to W$.

例 3.6.4. 逆时针旋转角度 θ 是 $\mathbb{R}^2\to\mathbb{R}^2$ 的 $\mathbb{R}$-线性同构.

命题 3.6.4. 给定 n 维 $\mathbb{R}$-线性空间 V 以及其一组基 $B=\{v_1,\dots,v_n\}$, 如下映射是 $\mathbb{R}$-线性同构

$$T\colon V\to\mathbb{R}^n$$
$$\sum_{i=1}^n a_iv_i\mapsto\sum_{i=1}^n a_ie_i$$

证明. 映射 T 显然是 $\mathbb{R}$-线性的, 并且满射也是显然的. 假设

$$T(\sum_{i=1}^n a_iv_i)=\sum_{i=1}^n a_ie_i=0$$

根据 $\{e_i\}$ 的线性无关性可知 $a_i=0$, 从而 $\sum_{i=1}^n a_iv_i=0$. □

推论 3.6.1. $\mathbb{R}$-线性空间 V,W 之间同构当且仅当 $\dim_{\mathbb{R}}V=\dim_{\mathbb{R}}W$, 即维数是线性空间的完全不变量.

证明. 显然同构的线性空间有相同的维数; 另一方面, 假设 $\dim_{\mathbb{R}}V=\dim_{\mathbb{R}}W=n$, 则 $V\cong\mathbb{R}^n\cong W$. □

例 3.6.5. 对于矩阵 $A\in M_{m\times n}(\mathbb{R})$, 对于 n 维 $\mathbb{R}$-线性空间 V 以及 m 维 $\mathbb{R}$-线性空间 W

$$T_A:V\to W$$
$$v\mapsto CA[v]_B$$

是 $\mathbb{R}$-线性映射, 其中 B,C 分别是 V 和 W 的基. T_A 为线性同构当且仅当 A 可逆.

3.6.4 维数公式

定义 3.6.8. 对于 $\mathbb{R}$-线性映射 $T\colon V \to W$.

(1) **核** *(kernel)*定义为

$$\ker T\colon = \{v \in V \mid Tv = 0\}.$$

(2) **像** *(image)*定义为

$$\operatorname{im} T\colon = \{Tv \in W \mid v \in V\}.$$

命题 3.6.5. 对于 $\mathbb{R}$-线性映射 $T\colon V \to W$, $\ker T$ 是 V 的子空间, $\operatorname{im} T$ 是 W 的子空间.

定理 3.6.2 (维数公式)**.** 对于 $\mathbb{R}$-线性映射 $T\colon V \to W$, 有

$$\dim V = \dim \ker T + \dim \operatorname{im} T$$

证明. 由于 $\ker T$ 勾成了 V 的一个子空间, 我们取其一组基 $\{v_1, \dots, v_k\}$, 并将其拓展成 V 的一组基 $\{v_1, \dots, v_k, v_{k+1}, \dots, v_n\}$, 从而有

$$\operatorname{im} T = \operatorname{span}_{\mathbb{R}}\{T(v_1), \dots, T(v_n)\} = \operatorname{span}_{\mathbb{R}}\{T(v_{k+1}), \dots, T(v_n)\},$$

并且不难验证 $\{T(v_{k+1}), \dots, T(v_n)\}$ 线性无关, 从而有维数公式. □

推论 3.6.2. 对于 $A \in M_{m\times n}(\mathbb{R})$, 有如下维数公式成立

$$\dim \ker A + \operatorname{rank} A = n$$

推论 3.6.3. 对于 $A \in M_{m\times n}(\mathbb{R})$, 线性方程组 $Ax = 0$ 解空间的维数为 $n - \operatorname{rank} A$.

定理 3.6.3. 对任意 $(x_0, y_0), \dots, (x_n, y_n)$, 其中 $x_0, \dots, x_n$ 互不相同, 则存在次数小于等于 n 的多项式 $f(x)$ 使得 $f(x_i) = y_i$.

证明. 考虑 $\mathbb{R}$-线性映射 $T\colon \mathbb{R}[x]_{\le n} \to \mathbb{R}^{n+1}$, 定义为

$$T(f) = \begin{pmatrix} f(x_0) \\ f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}.$$

由于小于等于 n 次的非零多项式最多拥有 n 个零点, 从而 $\ker T = \{0\}$, 根据维数公式可知 T 是满射. □

3.7 行列式

3.7.1 定义与基本性质

回忆在注2.1.5中, 我们有

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

因此 2×2 矩阵可逆当且仅当 $ad-bc\neq 0$. 实际上, $ad-bc$ 有如下的几何含义: 考虑在 $\mathbb{R}^2$ 中由列向量 $v=\binom{a}{c}, w=\binom{b}{d}$ 围成的平行四边形, 向量 v,w 的夹角 θ 满足

$$\cos\theta = \frac{ab+cd}{\sqrt{a^2+c^2}\sqrt{b^2+d^2}}$$

从而 v,w 围成的平行四边形面积为

$$\begin{aligned} S &= \sqrt{1-\cos\theta^2}\sqrt{a^2+c^2}\sqrt{b^2+d^2} \\ &= \sqrt{(ad-bc)^2} \\ &= |ad-bc| \end{aligned}$$

注意 $ad-bc>0$ 与 <0 两种情况分别对应了 v 在 w 左侧或右侧两种情况, 因此 $ad-bc$ 可以看作是 v,w 围成的平行四边形的 "有向面积", 我们将要定义的行列式, 就是这种有向面积的高维推广.

定义 3.7.1. $\mathbb{R}^n$ 上的 n 次**多重反对称线性函数** *(multiple skew symmetric linear function)*(多重反对称线性函数, *multiple skew symmetric linear function*) 是函数

$$f\colon \underbrace{\mathbb{R}^n\times\cdots\times\mathbb{R}^n}_{n\text{个}} \to \mathbb{R}$$

满足条件

(1) $f(v_1,\ldots,cv_i,\ldots,v_n) = cf(v_1,\ldots,v_n)$, 其中 $c\in\mathbb{R}$.

(2) $f(v_1,v_2,\ldots,v_i'+v_i,\ldots,v_n) = f(v_1,\ldots,v_i',\ldots,v_n) + f(v_1,\ldots,v_i,\ldots,v_n)$.

(3) $f(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_n) = -f(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_n)$.

(4) $f(e_1,\ldots,e_n)=1$.

定理 3.7.1. $\mathbb{R}^n$ 上的 n 次多重反对称线性函数存在且唯一.

证明. 对 n 进行归纳: 当 $n=1$ 时, 由于 $f(e_1)=1$ 以及任取 $v\in\mathbb{R}$ 我们有 $v=ce_1$, 从而 $f(v)=c$, 即此时被唯一确定. 假设当 $n<k$ 时假设成立, 考虑 $n=k$ 的情形. 任取 $v=(v_1,\ldots,v_k)\in\mathbb{R}^k$, □

定义 3.7.2. 假设 f 是 $\mathbb{R}^n$ 上 n 次多重反对称线性函数, 对于 $A \in M_n(\mathbb{R})$, 记 $A = (v_1, \dots, v_n)$, 则 A 的**行列式** *(determinant)*定义为

$$\det A = |A| := f(v_1, \dots, v_n).$$

例 3.7.1. 对于三种初等矩阵[5], 其行列式分别为

(E1) $\det E[ij] = 1$.

(E2) $\det E[i, c] = c$.

(E3) $\det E[ij, c] = 1$.

特别地, 对于矩阵 A 以及初等矩阵 E, 有 $|AE| = |A||E|$.

命题 3.7.1. 行列式有如下性质:

(1) 如果 A 的某一列为零, 则 $|A| = 0$.

(2) $|A| \neq 0$ 当且仅当 $\operatorname{rank} A = n$.

(3) $|\mathrm{I}_n| = 1$.

(4) $|AB| = |A||B|$.

(5) $|A^{-1}| = |A|^{-1}, |PAP^{-1}| = |A|$.

(6) $|A| = |A^T|$.

(7) 假设 A 是分块上三角矩阵, 并且对角线分块矩阵为 $A_1, \dots, A_n$, 则 $|A| = |A_1| \dots |A_n|$. 特别地, 如果 A 是上三角矩阵, 并且对角线元素为 $a_1, \dots, a_n$, 则 $|A| = a_1 \dots a_n$.

证明. (1). 根据定义3.7.1中 (1) 即可.

(2). 根据例3.7.1可知, 如果 E 是初等矩阵, 则 $|AE| = |A||E|$, 因此因此不妨将 A 写作 $E_k \dots E_1 \operatorname{rref} A$, 其中 E_i 是初等矩阵. 而 $\operatorname{rank} A = n$ 当且仅当 $\operatorname{rref} A$ 没有零列, 并且由于 $|E_i| \neq 0$, 从而根据 (1) 可知 $|A| \neq 0$ 当且仅当 $\operatorname{rank} A = n$.

(3). 根据定义3.7.1中 (4) 即可.

(4). 假设 B 不可逆, 即 $\operatorname{rank} B < n$, 根据 (2) 则有 $|B| = 0$. 而根据命题**??**有 $\operatorname{rank} AB \leq \operatorname{rank} B < n$, 因此 $|AB| = 0$, 即 $|AB| = |A||B|$. 假设 B 可逆, 则根据推论2.1.2不妨将其写成初等矩阵的乘积, 再根据初等矩阵的性质即可.

(5). 由 (4) 即得.

[5] 见定义2.1.4

(6). 根据推论2.2.2可知 $\operatorname{rank} A = \operatorname{rank} A^T$. 因此如果 $\operatorname{rank} A < n$, 则 $\operatorname{rank} A^T < n$, 即根据 (2) 可知 $|A| = |A^T| = 0$. 假设 $\operatorname{rank} A = n$, 再根据推论2.2.2将 A 写成初等矩阵的乘积, 并注意到对于初等矩阵 E 有 $|E| = |E^T|$.

(7). 不妨假设

$$A = \begin{pmatrix} A_1 & A_2 \\ \mathrm{O} & A_3 \end{pmatrix}$$

假设 A_1, A_3 中有一个不可逆, 则此时 A 也不可逆, 因此 $|A| = |A_1||A_2| = 0$. 假设 A_1, A_2 都可逆, 则此时 $\operatorname{rref} A = \mathrm{I}_n$. 将 A_1 化作最简行阶梯型的初等矩阵记做 $E_1, \dots, E_k$, 将 A_3 化作最简行阶梯型的初等矩阵记做 $E'_1, \dots, E'_l$, 则考虑

$$\widetilde{E}_i = \begin{pmatrix} E_i & \mathrm{O} \\ \mathrm{O} & \mathrm{I} \end{pmatrix}, \quad \widetilde{E}'_j = \begin{pmatrix} \mathrm{I} & \mathrm{O} \\ \mathrm{O} & E'_j \end{pmatrix}$$

则 $\widetilde{E}_k \dots \widetilde{E}_1 \widetilde{E}'_l \dots \widetilde{E}'_1 A = \mathrm{I}_n$, 从而 $|A|^{-1} = |\widetilde{E}_k| \dots |\widetilde{E}_1||\widetilde{E}'_l| \dots |\widetilde{E}'_1|$. 注意到 $|\widetilde{E}_i| = |E_i|, |\widetilde{E}'_j| = |E'_j|$ 以及 $|A_1| = \prod_{i=1}^{k} |E_i|, |A_3| = \prod_{j=1}^{l} |E'_j|$ 即可. □

定义 3.7.3. 给定 $\mathbb{R}$-线性空间 V 以及 $\mathbb{R}$-线性映射 $T\colon V \to V$, 则 T 的**行列式** *(determinant)*定义为 $\det A$, 其中 A 是 T 在任意一组基下的矩阵.

注 3.7.1. 线性映射的行列式是良好定义的, 因为同一个线性映射在不同基下的矩阵之间是相似的, 而根据命题3.7.1中的 *(4)* 可知相似的矩阵有着相同的行列式.

3.7.2 行列式的计算方法

接下来我们介绍一些计算矩阵行列式的常用的方法. 第一种常见的办法通过行变换将矩阵变换成上三角矩阵的情形, 此时的行列式是容易计算的, 并且矩阵的行列式在行变换下如何改变也是清楚的, 这种办法有着普遍性, 但是缺点在于计算相对繁琐.

另一种常见的办法就是通过按行展开: 给定 $A = (a_{ij})_{n\times n} \in M_n(\mathbb{R})$, 我们记 $A = (v_1, \dots, v_n)$, 则根据行列式的定义有

$$\begin{aligned} \det A &= f(v_1, \dots, v_n) \\ &= f(a_{11}e_1 + \cdots + a_{n1}e_n, v_2, \dots, v_n) \\ &= a_{11} f(e_1, v_2, \dots, v_n) + \cdots + a_{n1} f(e_n, v_2, \dots, v_n). \end{aligned}$$

对于某个 $i = 1, \dots, n$, 我们记 $v_2 = a_{i2}e_i + v'_2$, 则 $v'_2 \in W := \operatorname{span}_{\mathbb{R}}\{e_1, \dots, \widehat{e_i}, \dots, e_n\} \cong \mathbb{R}^{n-1}$. 类似地, 我们可以构造 $v'_3, \dots, v'_n \in W$. 根据行列式的交错性, 我们有

$$f(e_i, v_2, \dots, v_n) = f(e_i, v'_2, \dots, v'_n).$$

考虑

$$g\colon \underbrace{W\times\cdots\times W}_{n-1\text{个}}\to\mathbb{R}$$

$$(w_1,\ldots,w_{n-1})\mapsto f(e_i,w_1,\ldots,w_n).$$

函数 g 满足 $\mathbb{R}^{n-1}$ 上多重反对称线性函数的所有性质, 除了

$$g(e_1,\ldots,\widehat{e_i},\ldots,e_n)=f(e_i,e_1,\ldots,e_n)=(-1)^{i+1}f(e_1,\ldots,e_n)=(-1)^{i+1}.$$

因此根据多重反对称线性函数的唯一性, 我们有

$$f(e_i,v_2',\ldots,v_n')=(-1)^{i+1}|M_{i1}|,$$

其中 M_{i1} 是由矩阵 A 去掉第 i 行和第 j 列得到的矩阵, 因此有

$$|A|=\sum_{i=1}^{n}(-1)^{i+1}a_{i1}|M_{i1}|$$

定义 3.7.4. 给定 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{R})$.

1. 矩阵 A 去掉第 i 行和第 j 列得到的矩阵 M_{ij} 的行列式被称为 A 的(i,j) **余子式** *(minor).*
2. 矩阵 A 的(i,j) **代数余子式** *(algebraic minor)*定义为 $(-1)^{i+j}|M_{ij}|$;

命题 3.7.2. 给定 $A=(a_{ij})_{n\times n}\in M_n(\mathbb{R})$, 则

$$|A|=\sum_{j=1}^{n}(-1)^{i+j}a_{ij}|M_{ij}|.$$

例 3.7.2. 当 $n=2$ 时

$$\det\begin{pmatrix}a&b\\c&d\end{pmatrix}=ad-bc$$

例 3.7.3. 当 $n=3$ 时

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

注 3.7.2. 上述公式又被称为对角线法则.

命题 3.7.3. 对于 $2n\times 2n$ 阶矩阵

$$M=\begin{pmatrix}A&B\\C&D\end{pmatrix}$$

其中 A,B,C,D 是 $n\times n$ 阶矩阵, 并且 $AC=CA$, 则 $|M|=|AD-CB|$.

证明. 首先我们假设 A 可逆, 则

$$\begin{aligned}|M| &= |\begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}| \\ &= |A||D - CA^{-1}B| \\ &= |A(D - CA^{-1}B)| \\ &= |AD - CB|\end{aligned}$$

而当 A 不可逆时, 我们不妨考虑 $A_\lambda = A + \lambda I$. 由于 $|A_\lambda|$ 是 λ 的 n 次多项式, 从而对有限多个 λ 之外 A_λ 总可逆, 因此我们不妨取足够小的 λ 使得 A_λ 总可逆, 根据可逆时的情形我们有

$$|\begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix}| = |A_\lambda D - CB|.$$

从而我们令 $\lambda \to 0$ 即有我们期待的结果[6]. □

3.7.3 伴随矩阵

定义 3.7.5. 对于 $A \in M_n(\mathbb{R})$, **伴随矩阵** *(adjugate matrix)*$A^* = (a_{ij})_{n\times n}$ 定义为

$$a'_{ij} = (-1)^{i+j}|A_{ij}|.$$

命题 3.7.4. 对于 $A \in M_n(\mathbb{R})$, 有

$$AA^* = A^*A = \det A\, \mathrm{I}_n$$

证明. 由命题3.7.2可得. □

推论 3.7.1. 如果 $A \in M_n(\mathbb{R})$ 可逆, 则

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

推论 3.7.2 (克拉姆法则 (cramer's rule))**.** 对于线性方程组 $Ax = b$, 如果 A 可逆, 则有唯一解

$$x = \frac{1}{\det A}A^*b.$$

3.7.4 Laplace 公式

在本节中我们介绍用于矩阵行列式计算的 Laplace 公式, 这个公式可以看作是矩阵行列式按行展开的一个自然的推广.

为了公式的陈述, 我们先引入一些记号: 对于 $A \in M_n(\mathbb{R})$ 以及指标集 $M \subset \{1, 2, \ldots, n\}$. 定义 $A_{M,L}$ 表示以 M 为行, 以 L 为列组成的 A 的子矩阵, A_{M^c,L^c} 表示以 $\{1, 2, \ldots, n\} \setminus M$ 为行, 以 $\{1, 2, \ldots, n\} \setminus L$ 为列组成的 A 的子矩阵.

[6] 这个操作称为微扰法, 是矩阵中一个非常经典的技巧.

定理 3.7.2 (Laplace 公式)**.** 对于 $A \in M_n(\mathbb{R})$,

$$|A| = \sum_{\substack{L \subset \{1,2,\ldots,n\} \\ |L|=|M|}} (-1)^{\epsilon_{M,L}} |A_{M,L}||A_{M^c,L^c}|,$$

其中

$$\epsilon_{M,L} = (\sum_{i \in M} i) + (\sum_{j \in L} j)$$

注 3.7.3. 当 $|M| = 1$ 时, *Laplace* 公式给出了矩阵行列式的按行展开公式.

定理 3.7.3 (Cauchy-Binet 公式)**.** 对于 $A \in M_{m\times n}(\mathbb{R})$ 以及 $B \in M_{n\times m}(\mathbb{R})$,

$$|AB| = \begin{cases} 0, & m > n \\ \sum_{\substack{S \subset \{1,2,\ldots,n\} \\ |S|=m}} |A_{[m],S}||B_{S,[m]}|, & m \leq n, \end{cases}$$

其中 $A_{[m],S}$ 表示所有列指标由 S 给出的 A 的 m 阶子方阵, $B_{S,[m]}$ 表示所有行指标由 S 给出的 A 的 m 阶子方阵.

3.8 域上的线性空间

3.8.1 域

定义 3.8.1. 一个集合 $\mathbb{F}$ 被称为一个**域** *(field)*, 如果其上拥有如下两种运算:

(1) $+ \colon \mathbb{F} \times \mathbb{F} \to \mathbb{F}, (a, b) \mapsto a + b$;

(2) $\times \colon \mathbb{F} \times \mathbb{F} \to \mathbb{F}, (a, b) \mapsto ab$.

并且满足:

- 存在 $\boldsymbol{0}, \boldsymbol{1} \in \mathbb{F}$, 使得任取 $m \in \mathbb{F}$ 有 $\boldsymbol{0} + m = m$, $\boldsymbol{1}m = m$;
- 任取 $m \in \mathbb{F}$, 存在 $-m \in \mathbb{F}$ 使得 $m + (-m) = \boldsymbol{0}$;
- 任取 $0 \neq m \in \mathbb{F}$, 存在 $m^{-1} \in \mathbb{F}$ 使得 $mm^{-1} = \boldsymbol{1}$;
- 任取 $a, b, c \in \mathbb{F}$, 有如下结合律, 交换律和分配律:

  (i) $a + b = b + a$, $(a + b) + c = a + (b + c)$;

  (ii) $ab = ba$, $(ab)c = a(bc)$;

  (iii) $(a + b)c = ac + bc$, $c(a + b) = ca + cb$;

例 3.8.1. 实数 $\mathbb{R}$, 有理数 $\mathbb{Q}$, 复数 $\mathbb{C}$ 都是域.

例 3.8.2. 在 $\mathbb{R}^2$ 上定义

$$(a,b)+(c,d)=(a+c,b+d)$$
$$(a,b)\times(c,d)=(ac-bd,ad+bc).$$

直接验证可知上述运算满足结合律, 交换律和分配律. 其中 $\boldsymbol{0}=(0,0)$, $\boldsymbol{1}=(1,0)$. 任取 $(a,b)\in\mathbb{R}^2$, 我们有 $-(a,b)=(-a,-b)$. 对于 $(a,b)\neq(0,0)$, 我们有

$$(a,b)^{-1}=(\frac{a}{a^2+b^2},\frac{-b}{a^2+b^2}).$$

注 3.8.1. 这实际上用 $\mathbb{R}^2$ 构造了复数 $\mathbb{C}$, 即 (a,b) 对应于 $a+\sqrt{-1}b$.

3.8.2 域上的线性空间

定义 3.8.2. 给定集合 V, 如果其上有如下结构:

(1) 加法 $V\times V\to V$, 记做 $(v_1,v_2)\mapsto v_1+v_2$.

(2) 数乘 $\mathbb{F}\times V\to V$, 记做 $(c,v)\mapsto cv$.

满足:

(3) 加法满足交换律, 结合律, 0 向量以及逆元.

(4) 数乘满足结合律以及单位元.

(5) 加法和数乘满足分配律.

则称 V 构成了一个 $\mathbb{F}$-**线性空间** *(vector space)*.

例 3.8.3. $\mathbb{F}^n$ 构成了一个 $\mathbb{F}$-线性空间. 特别地, $\mathbb{R}^n$ 是一个 $\mathbb{R}$-线性空间.

例 3.8.4. 假设有域之间的包含关系 $\mathbb{F}\subseteq\mathbb{E}$, 则任何 $\mathbb{E}$-线性空间都可以视作 $\mathbb{F}$-线性空间. 例如, 任何 $\mathbb{C}$-线性空间都可以视作 $\mathbb{R}$-线性空间.

我们之前对于 $\mathbb{R}^n$ 所发展的理论都可以对一般的 $\mathbb{F}$-线性空间发展.

定义 3.8.3. 给定 $\mathbb{F}$-线性空间 V, $W\subseteq V$ 被称为 V 的**子空间** *(subspace)*, 如果 W 对 V 上的数乘与加法都封闭.

定义 3.8.4. $\mathbb{F}$-线性空间 V 中的向量 $v_1,\dots,v_n$ 称为 $\mathbb{F}$-**线性无关** *(linearly independent)*(线性无关, *linearly independent*), 如果 0 表示为 $v_1,\dots,v_n$ 的 $\mathbb{F}$-线性组合的方式唯一, 否则 $v_1,\dots,v_n$ 称为**线性相关** *(linearly dependent)*.

定义 3.8.5. $\mathbb{F}$-线性空间 V 的一个 $\mathbb{F}$-极大线性无关组被称作 V 的 $\mathbb{F}$-**基** *(basis)*.

定义 3.8.6. $\mathbb{F}$-线性空间 V 的一个 $\mathbb{F}$-极大线性无关组中向量的个数被称作 V 的 $\mathbb{F}$-**维数** *(dimension).*

注 3.8.2. 注意, 在这里我们强调它作为某个域上线性空间的维数, 因为同一个集合作为不同线性空间的维数可能不同, 例如 $\dim_{\mathbb{C}} \mathbb{C} = 1$ 但是 $\dim_{\mathbb{R}} \mathbb{C} = 2$.

3.9 作业六

3.9.1 基础题

习题 3.9.1. 如下归纳地定义方阵

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}$$

. 求 A_n 的平方 $(A_n)^2$.

习题 3.9.2. 请判断以下向量组是否线性无关，并找出下述向量组生成的子空间的一组数目最少的生成元.

1. $a_1 = (1, 2, 3), a_2 = (3, 6, 7)$;
2. $a_1 = (2, -3, 1), a_2 = (3, -1, 5), a_3 = (1, -4, 3)$;
3. $a_1 = (4, -5, 2, 6), a_2 = (2, -2, 1, 3), a_3 = (6, -3, 3, 9), a_4 = (4, -1, 5, 6)$;
4. $a_1 = (1, 0, 0, 2, 5), a_2 = (0, 1, 0, 3, 4), a_3 = (0, 0, 1, 4, 7), a_4 = (2, -3, 4, 11, 12)$.

习题 3.9.3. 请将 $Span_{\mathbb{R}}(v_1, v_2)$ 写成某个矩阵的 *kernel*. 其中 $v_1 = (1, 2, 3, 4)^T$ *and* $v_2 = (5, 6, 7, 8)^T$.

习题 3.9.4. 令 $S = \{v \in \mathbb{R}^3 : v = (r - 2s, 3r + s, s)^T, r, s \in \mathbb{R}\}$.

1. 请验证 S 是 $\mathbb{R}^3$ 的子空间.
2. 证明 S 是平面 $3x - y + 7z = 0$.

习题 3.9.5. 判断以下集合和运算是否构成 $\mathbb{R}$-线性空间.

1. n-阶实对称矩阵 $A = A^T$ 全体，在矩阵的加法和数乘下.
2. n-阶实反对称矩阵 $A = -A^T$ 全体，在矩阵的加法和数乘下.
3. 满足 $p(1) = p(2)$ 的所有实系数多项式，在通常多项式的加法和数乘下.
4. 秩小于或等于 1 的三阶方阵全体，在矩阵的加法和数乘下.
5. 在实轴上定义的周期等于 1 的全体实值函数，在通常函数的加法和数乘下.
6. $\mathbb{R}^2$ 中满足方程 $x^2 = y^2$ 的点集，在 $\mathbb{R}^2$ 的加法和数乘下.
7. 实轴上的光滑函数，满足 $f'(t) + f(t) = \cos t$，在通常函数的加法和数乘下.

习题 3.9.6.

1. 请举出 $\mathbb{R}^2$ 中满足加法封闭，但是数乘不封闭的非空集合的例子.

2. 请举出 $\mathbb{R}^2$ 中满足数乘封闭，但是加法不封闭的非空集合的例子.

习题 3.9.7. 在 $\mathbb{R}$-线性空间 V 中，

1. 验证对任意 $v \in V$, $-1 \cdot v = -v$.

2. 验证对任意 $c \in F$, $c\mathbf{0} = \mathbf{0}$, 这里 $\mathbf{0}$ 指的是 V 中的加法单位元.

习题 3.9.8. 固定某个向量 $w \in \mathbb{R}^n$. 对于 $a \in \mathbb{R}$ *and* $u \in \mathbb{R}^n$, 定义

$$a \otimes u = a(u - w) + w.$$
$$u \oplus v = u + v - w.$$

请判断并证明 $V = \mathbb{R}^n$ 在数乘 $\otimes$ 和加法 $\oplus$ 下是否做成 $\mathbb{R}$ 线性空间. 如果是，其中零向量是什么? *(Note:* 我们用记号 $\otimes$ *and* $\oplus$ 来和 $\mathbb{R}^n$ 上的通常加法和数乘做区分*).*

3.10 作业七

3.10.1 基础题

习题 3.10.1. 决定以下向量组是否组成 $\mathbb{R}^3$ 的基.

$$\left\{\begin{bmatrix}1\\2\\1\end{bmatrix},\begin{bmatrix}3\\-1\\2\end{bmatrix},\begin{bmatrix}1\\1\\-1\end{bmatrix}\right\}.$$

习题 3.10.2. 判定以下 $M_{2\times2}(\mathbb{R})$ 中的向量组是否线性无关:

$$A_1=\begin{bmatrix}1&1\\0&1\end{bmatrix},\quad A_2=\begin{bmatrix}2&-1\\0&1\end{bmatrix},\quad A_3=\begin{bmatrix}3&6\\0&1\end{bmatrix}.$$

习题 3.10.3. $C^\infty(\mathbb{R})$ ($\mathbb{R}$ 上的光滑函数) 是 (无穷维) $\mathbb{R}$-线性空间. 证明, 函数

$$\sin(t),\sin(2t),\cdots,\sin(Nt)$$

线性无关.

习题 3.10.4. 如果矩阵 $A=A^T$, 则称 A 是对称矩阵. 如果 $A=-A^T$, 则称 A 是反对称矩阵. 分别证明对称矩阵和反对称矩阵构成 $M_n(\mathbb{R})$ 的子空间, 计算这两个子空间的维数.

习题 3.10.5. 记 $V=C^\infty(\mathbb{R})$ 是 $\mathbb{R}$ 上的光滑函数组成的 $\mathbb{R}$-线性空间. 验证两组元素 B,C 满足 $\mathrm{Span}_{\mathbb{R}}B=\mathrm{Span}_{\mathbb{R}}C=W$, 且 B,C 均为 W 的基.

1. $B=(1,\cos x,\cos 2x,\cos 3x)$, $C=(1,\cos x,\cos^2 x,\cos^3 x)$.
2. $B=(1,x,x^2,x^3)$, $C=(1,x-a,(x-a)^2,(x-a)^3)$ ($a\in\mathbb{R}$ 为常数)

习题 3.10.6. 设 V 是 $\mathbb{R}$ 上有限维向量空间, $W_1,W_2,\cdots,W_n$ 是 V 的真子空间, 证明:

$$W_1\cup W_2\cup\cdots\cup W_n\neq V.$$

习题 3.10.7. 考虑 *"Shifted Legendre polynomial"*

$$\widetilde{P}_n(x)=\frac{1}{n!}\frac{d^n}{dx^n}(x^2-x)^n$$

证明 $\widetilde{P}_0,\widetilde{P}_1,\cdots$ 构成 $\mathbb{R}[x]$ 的一组基.

3.11 作业八

3.11.1 基础题

习题 3.11.1. 记 $V = C^\infty(\mathbb{R})$ 是 $\mathbb{R}$ 上的光滑函数组成的 $\mathbb{R}$ 线性空间. 第七次作业中，我们验证了两组元素 B, C 满足 $\mathrm{Span}_{\mathbb{R}} B = \mathrm{Span}_{\mathbb{R}} C = W$, 且 B, C 均为 W 的基. 请写出 B, C 的转换矩阵 $P_{B\leftarrow C}$ 和 $P_{C\leftarrow B}$.

1. $B = (1, \cos x, \cos 2x, \cos 3x)$, $C = (1, \cos x, \cos^2 x, \cos^3 x)$.
2. $B = (1, x, x^2, x^3)$, $C = (1, x-a, (x-a)^2, (x-a)^3)$ ($a \in \mathbb{R}$ 为常数)

习题 3.11.2. 考虑线性映射 $F\colon \mathbb{R}^3 \to \mathbb{R}^2$ 使得 $(x_1, x_2, x_3)^T \mapsto (x_1 + 2x_2, x_1 - x_2)^T$. 计算 T 关于以下 $\mathbb{R}^3$ 的基 $\{\alpha_1, \alpha_2, \alpha_3\}$ of $\mathbb{R}^3$ 以及 $\mathbb{R}^2$ 的基 $\{\beta_1, \beta_2\}$ 对应的矩阵:

1. $\alpha_1 = (1,0,0)^T, \alpha_2 = (0,1,0)^T, \alpha_3 = (0,0,1)^T, \beta_1 = (1,0)^T, \beta_2 = (0,1)^T$
2. $\alpha_1 = (1,1,1)^T, \alpha_2 = (0,1,1)^T, \alpha_3 = (0,0,1)^T, \beta_1 = (1,1)^T, \beta_2 = (1,0)^T$
3. $\alpha_1 = (1,2,3)^T, \alpha_2 = (0,1,-1)^T, \alpha_3 = (-1,-2,3)^T, \beta_1 = (1,2)^T, \beta_2 = (2,1)^T$

习题 3.11.3. 令 $B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$

1. 证明和 B 交换的矩阵 A 的集合 $W = \{A \in M_{2\times 2}(\mathbb{R}) | AB = BA\}$ 是 $\mathbb{R}$-线性空间 $M_{2\times 2}(\mathbb{R})$ 的子空间.
2. 找到 W 的一组基.

习题 3.11.4. 对于 $A \in M_n(\mathbb{R})$, 证明 $A \cdot A^T$ 的列空间和 A 的列空间相同.

习题 3.11.5. 令 V 是形如 $AB - BA$ 的矩阵生成的 $M_n(\mathbb{C})$ 的子空间. 证明 $V = \{A \mid Trace(A) = 0\}$.

习题 3.11.6. 设 $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 3 \end{bmatrix}$. 考虑线性映射 $f\colon M_3(\mathbb{R}) \to M_3(\mathbb{R})$ 满足对 $B \in M_3(\mathbb{R})$ 有 $f(B) = A \cdot B$.

1. 求 A 的（行）最简阶梯型。
2. 求 $\mathrm{Ker}(f)$ 作为 $\mathbb{R}$-线性空间的一组基。
3. 求 $\mathrm{Im}(f)$ 作为 $\mathbb{R}$-线性空间的一组基。

习题 3.11.7. 假设 V 是 $\mathbb{R}$-线性空间 $\mathbb{R}[x]$, 考虑两个线性映射 $T_1, T_2\colon V \to V$, 使得 $T_1(f) = f'$ 和 $T_2(f) = xf$. 证明 $T_1 \circ T_2 - T_2 \circ T_1 = \mathrm{I}$. 请问在有限维线性空间 V 中是否存在这样的线性映射？

3.12 作业九

3.12.1 基础题

习题 3.12.1. 假设 l 是二维平面上过原点的直线，且 l 与横轴的夹角为 θ. 记 T 为以 l 为对称轴的反射. 求线性变换 T 在标准基下的矩阵. 假设另一条二维平面上过原点的直线 l' 与横轴的夹角为 α, 记 T' 为以 l' 为对称轴的反射. 利用线性变化与矩阵乘法的关系，证明 $T \circ T'$ 是绕原点旋转的线性变换.

习题 3.12.2. 设 T_i 是 $\mathbb{R}^3 = \{(x_1, x_2, x_3)^T \mid x_i \in \mathbb{R}\}$ 上的线性变换, 定义为绕 x_i 轴旋转角度 π. 请写下这三个变换在标准基下的矩阵并证明 $T_1 \circ T_2$ 仍然是某个旋转变换.

习题 3.12.3. 证明复数在通常的加法和实数的乘法下做成 $\mathbb{R}$ 线性空间，且有基 $B\colon 1, \sqrt{-1}$. 对某一固定的复数 $u + v\sqrt{-1}$. 定义 $T\colon \mathbb{C} \to \mathbb{C}$ 为 $T(z) = u \cdot z$. 证明 T 是 $\mathbb{C}$ 上的 $\mathbb{R}$ 线性变换并写出 T 在基 B 下的矩阵.

习题 3.12.4. 找到一个 $\mathbb{R}^2$ 上的线性变换将曲线 C

$$\{(x, y)^T \mid x^2 + 4xy + 10y^2 = 1\}$$

映射为半径为 1 的圆. 找到一个旋转变换将这个圆映射为长轴在 x 轴上的椭圆.

习题 3.12.5. 对于 $\mathbb{R}^2$ 上的线性变换 T, 请找出 $\mathrm{tr}(T^3)$ 和 $\mathrm{tr}(T^2), \mathrm{tr}(T)$ 之间的关系.

3.13 作业十

3.13.1 基础题

习题 3.13.1. 设 V 是有限维线性空间, $T\colon V \to V$ 是线性变换. 证明 T 是可逆的当且仅当 T 的核是 $\{0\}$. 当 V 是无限维线性空间时, 这个结论是否成立?

证明. 假设 V 是有限维线性空间, 则根据维数公式有

$$\dim V = \dim \ker T + \dim \operatorname{im} T.$$

若 $\ker T = \{0\}$, 则 $\dim \operatorname{im} T = \dim V$, 故 $\operatorname{im} T = V$, 即 T 是满射, 从而可逆. 反之若 T 可逆, 则 T 是单射, 故 $\ker T = \{0\}$ 从而 $\ker T = \{0\}$ 当且仅当 $\operatorname{im} T = V$, 这也当且仅当 T 是线性同构.

当 V 是无限维线性空间时, 这个结论不成立: 考虑 $\ell^2 = \{(x_1, x_2, \dots) \mid x_i \in \mathbb{R}\}$ 以及其上的右移位算子 $T\colon \ell^2 \to \ell^2$, 定义为:

$$T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots),$$

其核为 $\{0\}$, 但非满射, 故不可逆. □

习题 3.13.2. 定义 P_n 为关于未定元 x 次数小于或等于 n 的复系数多项式组成的复线性空间. 定义 P_n 上的线性变换 $T\colon P_n \to P_n$ 为 $T(f) = f' + f$. 请问这个线性变换是否可逆, 并写下其逆变换.

证明. 任取 $g \in P_n[x]$, 则 $T(f) = f' + f = g$ 等价于

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(e^x f(x)\right) = e^x g.$$

两侧积分有

$$f(x) = e^{-x} \int e^x g + Ce^{-x}.$$

由于 $f(x)$ 是多项式, 从而 $C = 0$. 因此当 $g = 0$ 时 $f = 0$, 即 T 是单射.

由于 $g \in P_n[x]$, 从而 g 的 $n+1$ 阶导数为零, 从而分布积分若干次有

$$\int e^x g = e^x \sum_{i=0}^{n} (-1)^i g^{(i)}(x),$$

从而有逆变换

$$T^{-1}(g) = \sum_{i=0}^{n} (-1)^n g^{(i)}(x).$$

□

习题 3.13.3. 假设 A, B 都是 n 阶实矩阵, 且 A 可逆, $B^n = 0$. 证明以下映射 F 是双射

$$F\colon M_n(\mathbb{R}) \to M_n(\mathbb{R})$$
$$X \mapsto AX + XB.$$

证明. 由于 $M_n(\mathbb{R})$ 是有限维线性空间, 所以只需证明 F 是单射即可. 设 $F(X) = 0$, 即 $AX = -XB$. 左乘 A^{-1} 得:

$$X = -A^{-1}XB.$$

重复上述操作 k 次则得到

$$X = (-1)^k A^{-k} X B^k$$

由于 $B^n = 0$, 取 $k = n$, 由 $B^n = 0$ 得 $X = 0$, 这证明了单射性. □

习题 3.13.4. 假设 V 是 $M_n(\mathbb{R})$ 中的对称矩阵组成的子空间, W 是反对称矩阵组成的子空间. 证明 $V \oplus W = M_n(\mathbb{R})$.

证明. 任意实矩阵 A 可唯一分解为:

$$A = \underbrace{\frac{A + A^T}{2}}_{\text{对称}} + \underbrace{\frac{A - A^T}{2}}_{\text{反对称}}.$$

若 A 既对称又反对称, 则 $A = 0$, 故 $V \cap W = \{0\}$, 且 $V + W = M_n(\mathbb{R})$. □

习题 3.13.5. 假设 V 是实轴上的全体实值函数组成的线性空间, 其中又子空间 W_1 是全体奇函数组成的线性空间, W_2 是全体偶函数组成的线性空间. 证明 $V = W_1 \oplus W_2$.

证明. 对任意 $f \in V$, 唯一分解:

$$f(x) = \underbrace{\frac{f(x) - f(-x)}{2}}_{\in W_1} + \underbrace{\frac{f(x) + f(-x)}{2}}_{\in W_2}.$$

若 $f \in W_1 \cap W_2$, 则 $f(x) = -f(-x) = f(-x)$, 故 $f(x) = 0$. □

习题 3.13.6. 假设 T 是实线性空间 V 上的线性变换且满足 $T^2 = \mathrm{I}$. 证明 $V = \ker(T - \mathrm{I}) \oplus \ker(T + \mathrm{I})$. 这里 I 是恒等变换. 你能用这个事实来解释前两题的结论吗?

证明. 对任意 $v \in V$, 存在分解:

$$v = \underbrace{\frac{v + T(v)}{2}}_{v_1} + \underbrace{\frac{v - T(v)}{2}}_{v_2}.$$

我们断言 $v_1 \in \ker(T-\mathrm{I})$, $v_2 \in \ker(T+\mathrm{I})$. 直接验证如下:

$$(T-\mathrm{I})(v_1) = \frac{T(v)+T^2(v)}{2} - \frac{v+T(v)}{2} = \frac{T(v)+v}{2} - \frac{v+T(v)}{2} = 0.$$

同理 $(T+\mathrm{I})(v_2) = 0$. 若 $v \in \ker(T-\mathrm{I}) \cap \ker(T+\mathrm{I})$, 则 $T(v) = v = -v$, 故 $v = 0$. 前两题中:

(1) 取 $T(A) = A^T$, 则 $\ker(T-\mathrm{I})$ 为对称矩阵空间, $\ker(T+\mathrm{I})$ 为反对称矩阵空间;

(2) 取 $T(f)(x) = f(-x)$, 则 $\ker(T-\mathrm{I})$ 为偶函数空间, $\ker(T+\mathrm{I})$ 为奇函数空间;

□

习题 3.13.7. 假设 T 是有限维线性空间 V 上的线性变换且满足 $T^2 = T$. 证明 $V = \ker T \oplus \operatorname{im} T$.

证明. 分解 $v = (v - T(v)) + T(v)$, 其中:

- $T(v - T(v)) = T(v) - T^2(v) = 0$, 故 $v - T(v) \in \ker T$;
- $T(v) \in \operatorname{im} T$.

若 $u \in \ker T \cap \operatorname{im} T$, 则存在 w 使得 $u = T(w)$ 且 $T(u) = 0$, 于是 $u = T(w) = T^2(w) = T(u) = 0$. □

习题 3.13.8. 以下假设线性空间的维数均有限. 对于子空间 $W \subset V$, 定义余维数 $\operatorname{codim}(W) = \dim V - \dim W$. 设

$$V_0 \xrightarrow{T_1} V_1 \xrightarrow{T_2} \cdots \xrightarrow{T_m} V_m$$

是一串线性映射, 证明

$$\sum_{i=1}^{m} \dim \ker T_i - \sum_{i=1}^{m} \operatorname{codim} \operatorname{Im} T_i = \dim V_0 - \dim V_m.$$

证明. 对每个线性映射 $T_i\colon V_{i-1} \to V_i$, 根据维数公式有:

$$\dim V_{i-1} = \dim \ker T_i + \dim \operatorname{im} T_i,$$

改写为:

$$\dim \ker T_i - (\dim V_i - \dim \operatorname{im} T_i) = \dim V_{i-1} - \dim V_i.$$

对 $i = 1$ 到 m 累加得:

$$\sum_{i=1}^{m} \dim \ker T_i - \sum_{i=1}^{m} \operatorname{codim} \operatorname{im} T_i = \sum_{i=1}^{m} (\dim V_{i-1} - \dim V_i) = \dim V_0 - \dim V_m.$$

□

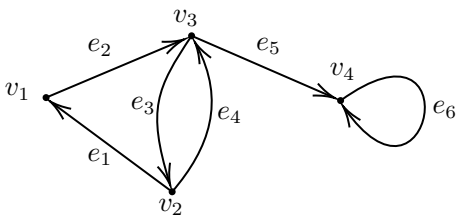
3.13.2 思考题

习题 3.13.9. 设 $G=(V,E)$ 是一个有向图, 带有映射 $s,t\colon E\to V$, 其中 $s(e),t(e)$ 分别表示有向边 e 的起点和终点. 记 $\mathbb{R}\{V\}$ 为形如 $\sum_{v\in V}x_v\cdot v\,(x_v\in\mathbb{R})$ 的表达式构成的线性空间; 类似地有 $\mathbb{R}\{E\}$.

定义线性映射

$$\begin{aligned} d\colon \mathbb{R}\{E\} &\to \mathbb{R}\{V\} \\ e &\mapsto t(e)-s(e), \end{aligned}$$

那么 $\dim\ker d, \operatorname{codim}\operatorname{im} d$ 给出了有向图中的什么信息? 尝试对如下例子计算 $\dim\ker d$.



3.14 小测一

习题 3.14.1. 找到一个 4×3 矩阵 A 使得 A 的列空间有基向量 $\begin{pmatrix}1\\2\\3\\4\end{pmatrix}$，行空间有基向量 $\begin{pmatrix}2\\3\\4\end{pmatrix}$.

习题 3.14.2. 设 A 是一个 3×3 矩阵，假设对于任意 3 维列向量 x，都有某个依赖于 x 的实数，使得 $Ax=c(x)x$. 证明 $A=cI_3$.

习题 3.14.3. 请写出一个 3×3 矩阵 A，使得 A 的 *kernel* 是 $\mathrm{span}_{\mathbb{R}}\{\begin{pmatrix}2\\3\\4\end{pmatrix}\}$.

习题 3.14.4. 设 A 是一个 $m\times n$ 矩阵，秩为 r. 假设 $Ax=b$ 对于某些右端 b 没有解，而对于另一些右端 b 有无穷多解。

1. 决定 A 的零空间是否只包含零向量，并说明原因.
2. 决定 A 的列空间是否是 $\mathbb{R}^m$，并说明原因.
3. 是否存在某个右端 b 使得 $Ax=b$ 有且只有一个解？为什么？
4. 找出 r,m 和 n 之间的大小关系.

习题 3.14.5. 对 n 阶方阵 A，证明以下秩等式

$$\mathrm{rank}(A+I)+\mathrm{rank}(A-I)=\mathrm{rank}(A^2-I)+n.$$

3.15 作业十一

3.15.1 基础题

习题 3.15.1. 计算下列行列式:

1.

$$\begin{vmatrix} 4 & -2 & 0 & 5 \\ 3 & 2 & -2 & 1 \\ -2 & 1 & 3 & -1 \\ 2 & 3 & -6 & -3 \end{vmatrix}$$

2.

$$\begin{vmatrix} x & y & z & 1 \\ y & z & x & 1 \\ z & x & y & 1 \\ \frac{z+x}{2} & \frac{x+y}{2} & \frac{y+z}{2} & 1 \end{vmatrix}$$

3.

$$\begin{vmatrix} 3 & 2 & 0 & \cdots & 0 \\ 1 & 3 & 2 & \cdots & 0 \\ 0 & 1 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 3 \end{vmatrix}_{n\times n}$$

4. 因式分解

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix}$$

5. Vandermonde 行列式

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{n-1} & X_2^{n-1} & \cdots & X_n^{n-1} \end{vmatrix}$$

6.

$$\begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{vmatrix}$$

其中 $s_k = X_1^k + X_2^k + \cdots X_n^k$.

7. 在复数域 $\mathbb{C}$ 上, 将关于 n 个变量 $a_1, a_2, \cdots, a_n$ 的多项式

$$\begin{vmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{vmatrix}$$

分解为不可约因子乘积.

8. $\det(A^*)$, 其中 A^* 是方阵 A 的伴随.

9. $(n+1) \times (n+2)$ 的矩阵

$$A = (a_{ij}) = \left(\binom{j-1}{i-1} \right), 1 \le i \le n+1, 1 \le j \le n+2,$$

A_k 为 A 去掉第 k 列得到的矩阵, 计算 $\det(A_k)$.

解. 1. -21(计算过程略).

2. 第一行加第三行等于第四行的 2 倍, 故行列式为 0.

3. 我们记这个行列式为 A_n, 这是一个依赖于 n 的函数. 若 $n \ge 3$, 对第一行展开有 $A_n = 3A_{n-1} - 2A_{n-2}$. 方程 $x^2 = 3x - 2$ 有两个单根 $x = 1, 2$, 故上述递归表达式有一般解 $A_n = a + b2^n$. 容易验证 $A_1 = 3, A_2 = 7$, 代入一般表达式解得 $a = -1, b = 2$, 故 $A_n = 2^{n+1} - 1$.

4. 我们可以假设 $a \neq 0$, 用第一列的 a 消去其他非零元, 再对称地用第一行的 a 消去其他非零元. 然后我们用第二列的 a 和第二行的 a 消去其他非零元. 结果如

下：

$$\begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{pmatrix} \xrightarrow{\text{1st column}} \begin{pmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ 0 & -d & -bd/a & f-eb/a \\ 0 & -e & -f-cd/a & -ec/a \end{pmatrix}$$

$$\xrightarrow{\text{1st row}} \begin{pmatrix} 0 & a & 0 & 0 \\ -a & 0 & d & e \\ 0 & -d & 0 & f-eb/a+cd/a \\ 0 & -e & -f-cd/a+eb/a & 0 \end{pmatrix}$$

$$\xrightarrow{\text{2nd column \& row}} \begin{pmatrix} 0 & a & 0 & 0 \\ -a & 0 & 0 & 0 \\ 0 & 0 & 0 & f-eb/a+cd/a \\ 0 & 0 & -f-cd/a+eb/a & 0 \end{pmatrix}$$

由于以上行列变换不改变行列式, 故

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = a^2(f+cd/a-eb/a)^2 = (af-be+cd)^2.$$

注：严格来讲, 我们以上的运算是在环 $\mathbb{Z}[a,b,c,d,e,f][1/a]$ 中进行的. 但是环 $\mathbb{Z}[a,b,c,d,e,f]$ 是 $\mathbb{Z}[a,b,c,d,e,f][1/a]$ 的一个子环, 故等式

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & d & e \\ -b & -d & 0 & f \\ -c & -e & -f & 0 \end{vmatrix} = (af-be+cd)^2$$

在 $\mathbb{Z}[a,b,c,d,e,f][1/a]$ 上成立, 且等式两边的元素都落入 $\mathbb{Z}[a,b,c,d,e,f]$ 中, 则等式在 $\mathbb{Z}[a,b,c,d,e,f]$ 上成立. 也可以考虑摄动法或直接计算.

5. 我们消去第一行有

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{n-1} & X_2^{n-1} & \cdots & X_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & X_2-X_1 & \cdots & X_n-X_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & X_2^{n-1}-X_1^{n-1} & \cdots & X_n^{n-1}-X_1^{n-1} \end{vmatrix}$$

对 $2 \le i \le n-1$, 我们把第 i 行乘以 $-X_1$ 加到第一行, 行列式不变, 为

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & X_2 - X_1 & \cdots & X_n - X_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & X_2^{n-2}(X_2 - X_1) & \cdots & X_n^{n-2}(X_n - X_1) \end{vmatrix}$$

$$= (X_2 - X_1)\cdots(X_n - X_1) \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & X_2^{n-2} & \cdots & X_n^{n-2} \end{vmatrix}$$

$$= (X_2 - X_1)\cdots(X_n - X_1) \begin{vmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ X_2^{n-2} & \cdots & X_n^{n-2} \end{vmatrix}$$

注意到式子后一部分是一个 $n-1$ 阶的 Vandermonde 行列式, 故可用归纳法得到 n 阶 Vandermonde 行列式为

$$\Pi_{1 \le i < j \le n}(X_j - X_i).$$

6.

$$\begin{vmatrix} s_0 & s_1 & s_2 & \cdots & s_{n-1} \\ s_1 & s_2 & s_3 & \cdots & s_n \\ s_2 & s_3 & s_4 & \cdots & s_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & s_{n+1} & \cdots & s_{2n-2} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ X_1 & X_2 & \cdots & X_n \\ \vdots & \vdots & \ddots & \vdots \\ X_1^{n-1} & X_2^{n-1} & \cdots & X_n^{n-1} \end{vmatrix} \begin{vmatrix} 1 & X_1 & \cdots & X_1^{n-1} \\ 1 & X_2 & \cdots & X_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_n & \cdots & X_n^{n-1} \end{vmatrix}$$

$$= \Pi_{1 \le i < j \le n}(X_j - X_i)^2.$$

7. 令 $J = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix} = \begin{pmatrix} 0 & I_{n-1} \\ 1 & 0 \end{pmatrix}$, 则可以验证对任意 $1 \le i \le$

$n-1$, $J^i = \begin{pmatrix} 0 & I_{n-i} \\ I_i & 0 \end{pmatrix}$. 由于 $|\lambda I_n - J| = \lambda^n - 1$(请自行验证), 故该多项式有 n 个单根, 即所有的 n 次单位根. 记 $\omega = e^{\frac{2\pi i}{n}}$, 则 J 的所有特征值为 $\omega^i, 0 \le i \le n-1$. 由于 J 的特征多项式只有单根, J 可以对角化, 于是存在可逆阵 P 使得 $P^{-1}JP = diag(1, \omega, \cdots, \omega^{n-1})$. 令 $f(x) = a_1 + a_2 x + \cdots + a_n x^{n-1}$, 注意到题目中的矩阵恰为 $a_1 + a_2 J + \cdots + a_n J^{n-1} = f(J)$, 于是 $P^{-1}f(J)P = diag(f(1), f(\omega), \cdots, f(\omega^{n-1}))$, 于是 $\det(f(J)) = \det(P^{-1}JP) = \Pi_{0 \le i \le n-1}(a_0 + a_1\omega^i + \cdots + a_n\omega^{i(n-1)})$. 我们得到了一个该行列式的因式分解, 而每个因子都是一次的, 故不可约.

8. 若 $A = 0$, 则 $A^* = 0$, 从而 $\det(A^*) = 0$. 若 $A \neq 0$ 且不可逆, 则 $AA^* = \det(A)I_n = 0$. 于是得到 A^* 也不可逆, 否则等式两边乘以 $(A^*)^{-1}$ 得到 $A = 0$, 矛盾! 故 $\det(A^*) = 0$. 若 A 可逆, 则由 $AA^* = \det(A)I_n$ 知 $\det(A)\det(A^*) = \det(\det(A)I_n) = \det(A)^n$, 而 $\det(A) \neq 0$, 故 $\det(A^*) = \det(A)^{n-1}$.

9. 我们先说明: 对于固定的 i, 函数 $j \to \binom{j}{i}$ 是一个首项系数为 $1/i!$ 的 i 次多项式. 这是因为 $\binom{j}{i} = \frac{j(j-1)\cdots j-i+1}{i!}$ (注意此式对 $0 \le j \le i-1$ 仍然成立, 此时等式两边为 0.) 我们记 $P_i(x) = \frac{x(x-1)\cdots x-i+1}{i!}$. 由归纳法我们可以证明存在系数 $c_{ij}, 0 \le j \le i-1$ 使得 $\frac{1}{i!}x^i = P_i(x) + c_{i,i-1}P_{i-1}(x) + \cdots + c_{i,0}P_0(x)$. 所求行列式为

$$\begin{vmatrix} P_0(0) & P_0(1) & \cdots & P_0(k-2) & P_0(k) & \cdots & P_0(n+1) \\ P_1(0) & P_1(1) & \cdots & P_1(k-2) & P_1(k) & \cdots & P_1(n+1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ P_n(0) & P_n(1) & \cdots & P_n(k-2) & P_n(k) & \cdots & P_1(n+1) \end{vmatrix}$$

我们从下到上把第 j 行乘以系数 c_{ij} 加到第 i 行上去, 得到行列式

$$\begin{vmatrix} 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & k-2 & k & \cdots & n+1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1/n! & \cdots & (k-2)^n/n! & k^n/n! & \cdots & (n+1)^n/n! \end{vmatrix}$$

由第 5 问知它等于

$$\frac{\Pi_{0 \le i < j \le n+1, i, j \neq k-1}(j-i)}{\Pi_{1 \le i \le n} i!}$$

由于

$$\frac{\Pi_{1 \le i \le n+1} i!}{\Pi_{0 \le i < j \le n+1}(j-i)} = 1$$

故可将它乘到上一式中并消去相同的项, 得到

$$\frac{(n+1)!}{\Pi_{0 \le i < j \le n+1, i = k-1}(j-i)\Pi_{0 \le i < j \le n+1, j = k-1}(j-i)} = \frac{(n+1)!}{(n+1-(k-1))!(k-1)!} = \binom{n+1}{k-1}.$$

□

习题 3.15.2. 设 A 是 $m \times n$ 矩阵, B 是 $n \times m$ 矩阵, 证明

$$\det(I_m + AB) = \det(I_n + BA).$$

证明. 我们考虑矩阵 $\begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix}$ 用 I_m 消去左下角的 B 知其行列式为 $\det(I_n - BA)$, 用 I_n 消去右上角的 A 知其行列式为 $\det(I_m - AB)$, 故二者相等.

□

习题 3.15.3. $M_n(\mathbb{R})$ 是实数域上的 n 阶方阵的集合. $\Phi : M_n(\mathbb{R}) \to \mathbb{R}$ 是一个映射, 满足以下条件

(1) $\Phi(AB) = \Phi(A)\Phi(B), \forall A, B \in M_n(\mathbb{R})$;

(2) 对任意上三角矩阵 $A \in M_n(\mathbb{R}), \Phi(A)$ 等于 A 的主对角线元素之积;

(3) 对任意下三角矩阵 $A \in M_n(\mathbb{R}), \Phi(A)$ 等于 A 的主对角线元素之积.

1. 证明: $\Phi(A) = |A|, \forall A \in M_n(\mathbb{R})$.

2. 如果 Φ 只满足条件 *(a)* 和 *(b)*, 结论是否成立? 请证明.

习题 3.15.4. 考虑一串线性映射

$$\cdots \xrightarrow{d_{n+1}} V_n \xrightarrow{d_n} V_{n-1} \xrightarrow{d_{n-1}} V_{n-2} \xrightarrow{d_{n-2}} \cdots$$

其中 V_k 都是有限维 $\mathbb{R}$-向量空间, 并且对任何 $k \in \mathbb{Z}$, $d_{k-1} \circ d_k = 0$.

- 记 $Z_k = \ker d_k, B_k = \operatorname{im} d_{k+1}$, 证明 B_k 是 Z_k 的子空间, 由此定义商空间 $H_k = Z_k/B_k$.
- 设 $\{f_n : V_n \to V_n\}_{n\in\mathbb{Z}}$ 是一串线性映射, 满足对任何 n, $d_n \circ f_n = f_{n-1} \circ d_n$, 则 $f_n(Z_n) \subset Z_n, f_n(B_n) \subset B_n$.
- 利用商空间的性质说明, f_n 诱导了线性映射 $f_{n*} : H_n \to H_n$, 使得如下图表交换:

$$\begin{array}{ccc} Z_n & \xrightarrow{f_n} & Z_n \\ \downarrow{\scriptstyle \pi_n} & & \downarrow{\scriptstyle \pi_n} \\ H_n & \xrightarrow{f_{n*}} & H_n \end{array}$$

  其中 $\pi_n : Z_n \to H_n = Z_n/B_n$ 是商空间的投影映射.

- *(Hopf 迹公式)* 设对某个 $N \in \mathbb{Z}_+$, 当 $|n| > N$ 时, $V_n = 0$. 证明

$$\sum_{n\in\mathbb{Z}}(-1)^n \operatorname{tr}(f_n \colon V_n \to V_n) = \sum_{n\in\mathbb{Z}}(-1)^n \operatorname{tr}(f_{n_*} \colon H_n \to H_n).$$

  注意这里操作的实际上是有限和, 不涉及级数收敛问题.

- 假设所有的 f_n 都可逆, 且对某个 $N \in \mathbb{Z}_+$, 当 $|n| > N$ 时, $V_n = 0$. 请证明

$$\Pi_{n\in\mathbb{Z}}(\det(f_n \colon V_n \to V_n))^{(-1)^n} = \Pi_{n\in\mathbb{Z}}(\det(f_{n_*} \colon H_n \to H_n))^{(-1)^n}.$$

  注意这里零维向量空间的线性变换的 det 定义为 1, 以上操作的实际上是有限乘积, 不涉及级数收敛问题.

证明. 若 $x \in B_k$ 则 $x = d_{k+1}(y)$, 又因为 $d_k \circ d_{k+1} = 0$ 我们有 $d_k(x) = d_k \circ d_{k+1}(y) = 0$, 因此 $B_k \subset Z_k$. 先证明 $f_n(Z_n) \subset Z_n$, 若 $d_n(x) = 0$ 则 $d_n(f_n(x)) = f_n(d_n(x)) = 0$; 再证明 $f_n(B_n) \subset B_n$, 若 $x = d_{n+1}(y)$ 则 $f_n(x) = d_{n+1}(f_n(y))$. 利用商的万有性质可得映射 $f_n : H_n \to H_n$ 以及交换性.

由于 $B_n \subset Z_n \subset V_n$ 都是 f_n 不变线性空间, 因此 $tr(f_n; V_n) = tr(f_n; B_n) + tr(f_n; H_n) + tr(f_n; V_n/Z_n)$. 我们断言 $tr(f_{n-1}; B_{n-1}) = tr(f_n; V_n/Z_n)$, 而这个式子不难推出 Hopf 迹公式. 事实上根据第一同构定理 $\tilde{d}_n : V_n/Z_n \to B_{n-1}$ 是线性空间的同构, 且 $\tilde{d} \circ f = f \circ \tilde{d}$ 从而得到断言. □

习题 3.15.5. 记 $w = e^{-\frac{2\pi i}{N}}$. 证明矩阵

$$W = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(N-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \omega^{3(N-1)} & \cdots & \omega^{(N-1)(N-1)} \end{bmatrix}$$

可逆, 并求 W^{-1}.

证明. 由 Van der monde 行列式的计算得矩阵可逆. $W(w)W(w^{-1}) = I$ □

3.16 作业十二

3.16.1 基础题

习题 3.16.1. 模仿四元域的构造方法，构造一个域 F，使得 F 包含 9 个元素.

习题 3.16.2. 利用域 F 的运算规律证明 $0 \cdot a = 0$ 对任意 $a \in F$ 成立.

习题 3.16.3. 证明域中不存在零因子，或者等价的说，不存在两个非零元素的乘积等于零.

习题 3.16.4. 验证 $\mathbb{Q}[\sqrt{-1}] = \{a + b\sqrt{-1} \mid a, b \in \mathbb{Q}\}$ 作为 $\mathbb{C}$ 的子集，在 $\mathbb{C}$ 的乘法和加法下是一个域.

习题 3.16.5. 设 $f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x]$，其中 $a_n \neq 0$。令 $f(x) \cdot \mathbb{R}[x]$ 为 $\mathbb{R}[x]$ 的子空间，包含所有能被 $f(x)$ 整除的多项式。

1. 求 $V = \mathbb{R}[x]/(f(x) \cdot \mathbb{R}[x])$ 的一个基，并计算其维数。
2. 考虑线性变换 $T\colon V \to V$，定义为 $\overline{g}(x) \mapsto \overline{x} \cdot \overline{g}(x)$。求 T 在所选基下的矩阵表示。

习题 3.16.6 (系数的限制)**.** 假设 F 是域 K 的子集，且在 K 的加法乘法的运算下做成一个域，则称 F 是 K 的子域. 证明

1. K 在乘法和加法下做成 F 的线性空间.
2. 假设 V 是一个 K 线性空间, 证明 V 在 F 继承 K 的数乘运算下做成 F 线性空间，并证明 $\dim_F V = \dim_F K \cdot \dim_K V$.*(你可以假设这里的维数均有限.)*
3. 以下问题中，我们假设 $F = \mathbb{R}$，$\dim_F K = 2$. 证明可将 $1 \in K$ 添加 $\alpha \in K$ 扩张成 K 的一组 F 基，且 $\alpha^2 = -1$.
4. 用以上各问中的记号和假设, 证明 α 的数乘定义了一个 V 上的 F 线性变换 $T_\alpha(v) = \alpha \cdot v$. 请找到一组基，写下这组基下的 T_α 的矩阵. 并且证明这个矩阵的平方等于 $-I$.

习题 3.16.7. 假设 $\alpha \in \mathbb{R}$. 将 $\mathbb{R}$ 视为 $\mathbb{Q}$ 上的线性空间.

1. 证明 α 是无理数等价于 $1, \alpha$ 是 $\mathbb{Q}$-线性无关.
2. 称 α 是代数数，如果存在多项式 $f(x) \in \mathbb{Q}[x]$，使得 $f(\alpha) = 0$. 否则，我们称 α 是超越数. 定义 $\mathbb{Q}[\alpha] = \{g(\alpha) \mid g(x) \in \mathbb{Q}[x]\}$. 证明 α 是代数数等价于 $\mathbb{Q}[\alpha]$ 是 $\mathbb{Q}$ 上的有限维线性空间.

3.17 小测二

习题 3.17.1. 考虑二阶实矩阵构成的实线性空间 V, 记矩阵 $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. 定义线性变换 $T\colon V \to V$ 为 $T(X) = AXB$. 求 $\operatorname{tr}(T)$.

证明. 选取 V 的一组基

$$e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad e_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad e_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

直接计算可知

$$T(e_1) = \begin{pmatrix} 5 & 6 \\ 15 & 18 \end{pmatrix}, \quad T(e_2) = \begin{pmatrix} 7 & 8 \\ 21 & 24 \end{pmatrix}, \quad T(e_3) = \begin{pmatrix} 10 & 12 \\ 20 & 24 \end{pmatrix}, \quad T(e_4) = \begin{pmatrix} 14 & 16 \\ 28 & 32 \end{pmatrix},$$

从而 T 在这组基下的表示矩阵为

$$\begin{pmatrix} 5 & 7 & 10 & 14 \\ 6 & 8 & 12 & 16 \\ 15 & 21 & 20 & 28 \\ 18 & 24 & 24 & 2 \end{pmatrix}.$$

从而有 $\operatorname{tr}(T) = 65$. □

习题 3.17.2. 如下归纳地定义方阵

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}.$$

求 $\det A_n$.

证明. 注意到 $A_1^2 = \mathrm{I}_2$, 以及

$$A_n^2 = \begin{pmatrix} A_{n-1}^2 + \mathrm{I}_{2^{n-1}} & 0 \\ 0 & A_{n-1}^2 + \mathrm{I}_{2^{n-1}} \end{pmatrix},$$

我们可以归纳地证明 $A_n^2 = n\,\mathrm{I}_{2^n}$, 特别地, A_n 是可逆矩阵. 根据降阶公式, 我们有

$$\begin{aligned} |A_n| &= |A_{n-1}|\,|-A_{n-1} - A_{n-1}^{-1}| \\ &= (-1)^{2^{n-1}} |A_{n-1}^2 + \mathrm{I}_{2^{n-1}}\,| \\ &= |(n-1)\,\mathrm{I}_{2^{n-1}} + \mathrm{I}_{2^{n-1}}\,| \\ &= n^{2^{n-1}}. \end{aligned}$$

从而有

$$|A_n| = \begin{cases} -1, & n = 1 \\ n^{2^{n-1}}, & n \geq 2 \end{cases}$$

□

习题 3.17.3. 设 V 是一个 2025 维实线性空间，$T\colon V \to V$ 是一个线性变换，且 $T^2 = 0$. 求 T 的秩的最大可能值.

证明. 由于 $T^2 = 0$, 从而 $\ker T \supseteq \operatorname{im} T$, 因此 $\dim \ker T \geq \dim \operatorname{im} T$. 根据维数公式有

$$\dim V = \dim \ker T + \dim \operatorname{im} T \geq 2 \operatorname{im} T.$$

从而 $\dim T \leq 1012$. 为了说明 1012 是 T 秩的最大可能值, 只需要构造一个秩为 1012 的矩阵使得 $T^2 = 0$ 即可. 考虑

$$J_2(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

这是一个秩为 1 的矩阵满足 $J_2(0)^2 = 0$. 考虑分块对角矩阵

$$T = \operatorname{diag}\{\underbrace{J_2(0), \ldots, J_2(0)}_{1012\text{个}}, 0\},$$

这给出了一个秩 1012 的矩阵满足 $T^2 = 0$. □

习题 3.17.4.

1. 假设 V 是一个 7 维实线性空间，是否存在线性变换 $T_1\colon V \to V$, $T_2\colon V \to V$ 使得 $T_1 \circ T_2 - T_2 \circ T_1 = \mathrm{I}_7$. 这里 I_7 是恒等变换.
2. 假设有限域 $\mathbb{F}_7 = \mathbb{Z}/7\mathbb{Z}$, V 是一个 7 维的 $\mathbb{F}_7$ 线性空间，是否存在线性变换 $T_1\colon V \to V$, $T_2\colon V \to V$ 使得 $T_1 \circ T_2 - T_2 \circ T_1 = \mathrm{I}_7$. 这里 I_7 是恒等变换.

证明. 当 V 是实线性空间时不存在, 因为如果存在 T_1, T_2 满足

$$T_1 \circ T_2 - T_2 T_1 = \mathrm{I}_7,$$

考虑两侧的迹得到 $0 = 7$ 矛盾. 当 V 是 $\mathbb{Z}/7\mathrm{Z}$ 上的线性空间时, 这样的线性变换存在, 取 V 的一组基 $\{e_0, \ldots, e_6\}$, 我们考虑如下的构造:

$$T_1(e_k) = k e_{k-1}.$$

以及

$$T_2(e_k)=\begin{cases}e_{k+1}, & k<6\\ 0, & k=6,\end{cases}$$

直接计算可知

$$T_1\circ T_2(e_k)=(k+1)e_k,\quad T_2\circ T_1(e_k)=ke_k,$$

从而有

$$T_1\circ T_2-T_2\circ T_1=\mathrm{I}_7\,.$$

□

习题 3.17.5. 已知

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

将 A 的 a_{ij} 所在的第 i 行第 j 列划掉后得一个 3×3 子矩阵，其行列式记为 M_{ij}，定义 $A_{ij}=(-1)^{i+j}M_{ij}$ ，试求

(1) $A_{11}+A_{12}+A_{13}+A_{14}$;

(2) $M_{12}+2M_{22}+3M_{32}+4M_{42}$.

证明. 利用行列式的按行展开公式构造新的行列式, 然后再用初等行列变换将其化为上三角阵即可, 结果如下:

$$A_{11}+A_{12}+A_{13}+A_{14}=-12,$$
$$M_{12}+2M_{22}+3M_{32}+4M_{42}=-18.$$

□

4 对角化

4.1 特征值与特征向量

对于 $\mathbb{F}$-线性映射 $T\colon V\to V$, 能否找到 V 的一组基 $\{v_1,\dots,v_n\}$ 满足对任意的 $1\le i\le n$ 有 $Tv_i=\lambda_i v_i$, 其中 $\lambda_i\in\mathbb{F}$, 称为线性映射的可**对角化** (diagonalizable)问题. 用矩阵的语言来说, 给定矩阵 $A\in M_{n\times n}(\mathbb{F})$, 是否存在可逆矩阵 $P\in M_{n\times n}(\mathbb{F})$ 使得 PAP^{-1} 是对角矩阵.

当 $\dim_{\mathbb{F}}V=1$ 时是显然的, 因为任何 1×1 阶矩阵都是对角矩阵. 但是对一般的维数来说, 并不是所有的线性映射都是可对角化的. 考虑 $\dim_{\mathbb{F}}=2$ 的情形: 固定 V 的一组基 B 将 T 看作矩阵

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}.$$

那么 $0\neq v\in V$ 满足 $Tv=\lambda v$, 其中 $\lambda\in\mathbb{F}$, 等价于 $(\lambda \mathrm{I}_2-A)v=0$ 有非零解, 根据定理2.1.1以及命题3.7.1的 (2) 可知这等价于 λ 满足 $|\lambda \mathrm{I}_2-A|=0$, 即

$$\lambda^2-(a+d)\lambda+ad-bc=0$$

我们考虑如下的情况:

(1) 如果 $\lambda^2-(a+d)\lambda+ad-bc=0$ 不存在根, 那么一定不存在 V 的一组基 $\{v_1,v_2\}$ 使得 $Tv_i=\lambda_i v_i$, 其中 $i=1,2$.

(2) 如果 $\lambda^2-(a+d)\lambda+ad-bc=0$ 有两个不同的根 λ_1,λ_2, 那么考虑对于每一个 λ_i 考虑 $(\lambda_i \mathrm{I}_2-A)v=0$ 的非零解 v_i, 那么我们断言 $\{v_1,v_2\}$ 构成了一组基, 因此满足我们的要求: 假设 $\{v_1,v_2\}$ 线性相关, 不妨假设 $v_1=\mu v_2$, 从而

$$\lambda_1v_1=Tv_1=T(\mu v_2)=\mu\lambda_2v_2.$$

这意味着 $\mu(\lambda_1-\lambda_2)=0$, 从而有 $\mu=0$, 从而证明了 v_1,v_2 线性无关.

(3) 如果 $\lambda^2-(a+d)\lambda+ad-bc=0$ 有重根 λ, 我们要考虑 $\lambda \mathrm{I}_2-A$ 的秩:

   (a) $\mathrm{rank}(\lambda \mathrm{I}_2-A)=0$, 此时根据推论3.6.3解空间的维数是 2, 因此可以找到两个线性无关的向量 $\{v_1,v_2\}$ 满足 $Tv_i=\lambda v_i$, 其中 $i=1,2$.

   (b) $\mathrm{rank}(\lambda \mathrm{I}_2-A)=1$, 此时根据推论3.6.3解空间的维数是 1, 则此时找不到两个线性无关的向量.

4.1.1 线性映射语言

定义 4.1.1. 对于 $\mathbb{F}$-线性映射 $T\colon V\to V$, $\lambda\in\mathbb{F}$ 被称为 T 的**特征值** *(eigenvalue)*, 如果存在 $0\neq v\in V$ 使得

$$Tv=\lambda v.$$

此时 v 称为特征值 λ 对应的**特征向量** *(eigenvector).*

定义 4.1.2. 对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 其中 $\dim_{\mathbb{F}} V = n$. 关于 λ 的 n 次多项式 $|\lambda \mathrm{I}_n - T|$ 称为 T 的**特征多项式** *(characteristic polynomial).*

注 4.1.1. 显然, $\mathbb{F}$-线性映射 T 的特征值是特征多项式在域 $\mathbb{F}$ 中的根, 因此 n 维线性空间上的线性映射至多有 n 个不同的特征值. 并且值得注意的是其可能没有特征值.

命题 4.1.1. 假设 V 是 n 维 $\mathbb{C}$-线性空间, 给定 $\mathbb{C}$-线性映射 $T\colon V \to V$, 存在 V 的一组基 B 使得 $[T]_B^B$ 是上三角矩阵.

证明. 我们对维数 n 做归纳法: 当 $n = 1$ 的时候是显然的. 假设命题对 $n < k$ 都成立, 当 $n = k$ 时, 由于 $\mathbb{C}$ 是代数闭域, 因此 T 的特征多项式总存在根, 即 T 总有特征值. 不妨假设 $v_1 \in V$ 是 T 对于特征值 λ_1 的特征向量. 将 v_1 扩充为 V 的一组基 $B = \{v_1, \dots, v_n\}$, 则

$$[T]_B^B = \begin{pmatrix} \lambda_1 & * \\ \mathrm{O} & A \end{pmatrix},$$

其中 A 是 $(k-1) \times (k-1)$ 阶矩阵. 根据归纳法存在可逆矩阵 P 使得 PAP^{-1} 是上三角矩阵, 从而考虑基 $B' = B \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix}$, 则此时 T 在这组基下的矩阵为上三角矩阵. □

推论 4.1.1. 对于 $\mathbb{C}$-线性映射 $T\colon V \to V$, 其特征值恰为 $[T]_B^B$ 对角线上的元素, 其中 B 是使得 $[T]_B^B$ 为上三角矩阵的基.

定义 4.1.3. 对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 全体特征值 λ 对应的特征向量构成了一个 $\mathbb{F}$-线性空间, 称为特征值 λ 的**特征子空间** *(eigenspace),* 记做 V_λ.

命题 4.1.2. 对于 $\mathbb{F}$-线性映射 $T\colon V \to V$ 以及其特征值 λ, 有

$$V_\lambda = \ker(\lambda \mathrm{I} - T).$$

证明. 根据定义即可. □

定理 4.1.1. $\mathbb{F}$-线性映射 $T\colon V \to V$ 可对角化当且仅当 V 有特征子空间分解, 即 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 其中 $\lambda_1, \dots, \lambda_s$ 是 T 的全体不同特征值.

证明. 假设 $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 则任取 $V_{\lambda_1}, \dots, V_{\lambda_s}$ 的基, 将其并起来得到 V 的一组基, 则在这组基下 T 对应的矩阵是对角矩阵; 另一方面, 假设 T 可对角化, 则可以找到 V 的一组由特征向量构成的基, 这组基给出了 V 的特征子空间分解. □

4.1.2 矩阵语言

定义 4.1.4. 对于矩阵 $A \in M_{m\times n}(\mathbb{F})$, $\lambda \in \mathbb{F}$ 被称为 A 的**特征值** *(eigenvalue)*, 如果存在非零列向量使得

$$Av = \lambda v.$$

此时 v 称为特征值 λ 对应的**特征向量** *(eigenvector)*.

定义 4.1.5. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 关于 λ 的 n 次多项式 $|\lambda \mathrm{I}_n - A|$ 称为 A 的**特征多项式** *(characteristic polynomial)*.

定义 4.1.6. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 全体特征值 λ 对应的特征向量构成了一个 $\mathbb{F}$-线性空间, 称为特征值 λ 的**特征子空间** *(eigenspace)*, 记做 V_λ.

命题 4.1.3. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$ 以及其特征值 λ, 有

$$V_\lambda = \ker(\lambda \mathrm{I}_n - T).$$

定理 4.1.2. 矩阵 $A \in M_{n\times n}(\mathbb{F})$ 可对角化当且仅当 $\mathbb{F}^n$ 有特征子空间分解, 即 $\mathbb{F}^n = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s}$, 其中 $\lambda_1, \ldots, \lambda_s$ 是 T 的全体不同特征值.

4.2 代数重数与几何重数

定义 4.2.1. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 特征值 λ 在特征多项式中的重数称为 λ 的**代数重数** *(algebraic multiplicity)*.

定义 4.2.2. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 特征值 λ 的特征子空间 V_λ 的维数称为 λ 的**几何重数** *(geometric multiplicity)*.

例 4.2.1. 假设矩阵 $A \in M_{n\times n}(\mathbb{F})$ 只有一个特征值 λ_1, 即其特征多项式为 $(\lambda - \lambda_1)^n$. 根据定理*4.3.1*可知 $(A - \lambda \mathrm{I}_n)^n = 0$, 因此 $\mathbb{F}^n = \ker(A - \lambda \mathrm{I}_n)^n$. 特别地, A 的特征值 λ_1 的代数重数为

$$n = \dim\{v \in \mathbb{F}^n \mid (\lambda \mathrm{I}_n - A)^n v = 0\}$$

命题 4.2.1. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$, 特征值 λ 的代数重数为 $\dim V'_\lambda$, 其中

$$V'_\lambda = \{v \in \mathbb{F}^n \mid \text{存在 } k \in \mathbb{N}_{\geq 0} \text{ 使得 } (\lambda \mathrm{I}_n - A)^k v = 0\}$$

推论 4.2.1. 对于矩阵 $A \in M_{n\times n}(\mathbb{F})$ 的特征值 λ, 其几何重数小于等于代数重数.

证明. 注意到对于特征值 λ 我们有 $V_\lambda \subseteq V'_\lambda$. □

定理 4.2.1. 矩阵 $A \in M_{n\times n}(\mathbb{F})$ 可对角化当且仅当对于每一个特征值 λ 其几何重数等于代数重数.

证明. 假设 A 的特征多项式为 $f(\lambda)=(\lambda-\lambda_1)^{n_1}\dots(\lambda-\lambda_s)^{n_s}$, 从而给出了 $\mathbb{F}^n$ 的如下直和分解

$$\mathbb{F}^n=V'_{\lambda_1}\oplus\cdots\oplus V'_{\lambda_s}$$

注意到代数重数等于几何重数当且仅当对任意 $1\leq i\leq s$ 有 $V'_{\lambda_i}=V_{\lambda_i}$, 根据定理4.1.2即可. □

4.3 Cayley-Hamilton 定理

定义 4.3.1. 对于 $\mathbb{F}$-线性映射 $T\colon V\to V$, 多项式 $f(\lambda)\in\mathbb{F}[\lambda]$ 称为其**零化多项式** *(annihilation polynomial)*, 如果 $f(T)=0$.

注 4.3.1. 类似的, 我们可以对矩阵定义其零化多项式, 即将矩阵视作线性映射. 本节之后所有的概念以及结果都可以用矩阵的语言叙述, 在此不再赘述.

引理 4.3.1. 对于任何 $\mathbb{F}$-线性变换 $T\colon V\to V$, 零化多项式总是存在.

证明. 任意取 V 的一组基, 考虑 T 在这组基下对应的矩阵 A. 假设 V 的维数为 n, 那么全体矩阵组成的线性空间的维数是 n^2, 从而 $\mathrm{I}_n, A, A^2, \dots, A^{n^2}$ 一定线性相关, 即存在不全为零的 $a_0, a_1, \dots, a_{n^2}$ 使得

$$a_0\,\mathrm{I}_n+a_1A+\cdots+a_{n^2}A^{n^2}=0,$$

从而 $f(x)=a_0+a_1x+\cdots+a_nx^{n^2}$ 是零化多项式. □

给定 $\mathbb{F}$-线性变换 $T\colon V\to V$, Cayley-Hamilton 定理改进上上述结果, 证明总是存在一个 n 次的多项式是 T 的零化多项式, 其中 n 是 V 的维数.

定理 4.3.1 (Cayley-Hamilton)**.** 对于 $\mathbb{F}$-线性映射 $T\colon V\to V$, 其特征多项式是其零化多项式.

为了证明这个结果, 我们先定义多项式环 $\mathbb{F}[\lambda]$ 在线性空间 V 上的作用:

$$\begin{aligned}\mathbb{F}[\lambda]\times V&\to V\\(h(\lambda),v)&\to h(T)v.\end{aligned}$$

不难验证上述作用有如下的性质:

(1) 对任意的 $f_1,f_2\in\mathbb{F}[\lambda], v\in V$, $(f_1f_2)v=f_1(f_2v)$;

(2) 对任意的 $f_1,f_2\in\mathbb{F}[\lambda], v\in V$, $(f_1+f_2)v=f_1v+f_2v$;

(3) 对任意的 $f\in\mathbb{F}[\lambda], c\in\mathbb{F}, v\in V$, $(cf)v=c(fv)$;

(4) 对任意的 $f\in\mathbb{F}[\lambda], v,w\in V$, $fv+fw$;

(5) 对任意的 $f \in \mathbb{F}[\lambda], c \in \mathbb{F}, v \in V$, $f(cv) = c(fv)$.

我们任意取 V 的一组基 $B = (v_1, \ldots, v_n)$, 那么

$$\lambda(v_1, \ldots, v_n) = T(v_1, \ldots, v_n) = (v_1, \ldots, v_n)A,$$

其中 A 是线性变换 T 在基 B 下的表示矩阵. 将上述等式转置则有

$$\lambda \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = A^T \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

即

$$(\lambda \mathrm{I}_n - A^T) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{3}$$

我们可以将 $\lambda \mathrm{I}_n - A^T$ 视为元素在多项式环 $\mathbb{F}[\lambda]$ 中的矩阵, 对于这样的矩阵, 我们仍可以定义其行列式 $|(\lambda \mathrm{I}_n - A^T)|$ 以及伴随矩阵 $(\lambda \mathrm{I}_n - A^T)^*$, 因为这只涉及乘法运算, 并且如下等式依然成立

$$(\lambda \mathrm{I}_n - A^T)(\lambda \mathrm{I}_n - A^T)^* = |\lambda \mathrm{I}_n - A^T| \mathrm{I}_n .$$

对于 (3), 两侧同时乘以伴随矩阵 $(\lambda \mathrm{I}_n - A^T)^*$ 则有

$$\left|\lambda \mathrm{I}_n - A^T\right| \mathrm{I}_n \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

从而对任意的 $1 \leq i \leq n$, 有 $|\lambda \mathrm{I}_n - A^T| v_i = 0$, 但是 λ 作用在 V 上的方式由线性变换 T 给出, 从而这证明了特征多项式是线性变换 T 的零化多项式.

4.4 极小多项式

定义 4.4.1. 对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 其次数最低的首一零化多项式被称为**极小多项式** *(minimal polynomial).*

命题 4.4.1. 对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 其极小多项式整除其任何一个零化多项式.

证明. 假设 $m(\lambda)$ 是 T 的极小多项式, $f(\lambda)$ 是 T 的某个零化多项式, 作带余除法则有

$$f(\lambda) = m(\lambda)q(\lambda) + r(\lambda)$$

其中 $r(\lambda)$ 显然也是 T 的零化多项式. 如果 $r(\lambda)$ 不为零, 根据带余除法的结果我们有 $\deg r(\lambda) < m(\lambda)$, 这与极小多项式的定义矛盾. □

推论 4.4.1. *对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 其极小多项式是唯一的.*

证明. 假设 $m_1(\lambda), m_2(\lambda)$ 都是 T 的极小多项式, 从而有 $m_1(\lambda) \mid m_2(\lambda)$ 以及 $m_2(\lambda) \mid m_1(\lambda)$, 因此 $m_1(\lambda), m_2(\lambda)$ 之间相差一个非零常数 c, 再利用首一性可知 $m_1(\lambda) = m_2(\lambda)$. □

推论 4.4.2. *对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 其极小多项式整除其特征多项式.*

证明. 根据定理4.3.1可知特征多项式是零化多项式. □

命题 4.4.2. *对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 如果 $f(\lambda)$ 是其零化多项式, 则其特征值 λ 也是该多项式的根. 特别地, 是其极小多项式的根.*

注 4.4.1. *根据推论4.4.2以及命题4.4.2可知, 如果 $\mathbb{F}$-线性映射 $T\colon V \to V$ 的分解为一次因式的乘积*

$$f(\lambda) = (\lambda - \lambda_1)^{n_1} \dots (\lambda - \lambda_s)^{n_s},$$

那么其极小多项式为

$$m(\lambda) = (\lambda - \lambda_1)^{k_1} \dots (\lambda - \lambda_s)^{k_s},$$

其中 $k_i \le n_i, 1 \le i \le s$.

定理 4.4.1. *对于 $\mathbb{F}$-线性映射 $T\colon V \to V$, 其在 $\mathbb{F}$ 上可对角化当且仅当极小多项式 $m(\lambda)$ 在 $\mathbb{F}[\lambda]$ 中可分解为一次因式的乘积, 且没有重根.*

证明. 如果 T 可对角化, 即其在某一组基下的矩阵 A 是对角阵, 其不同的特征值分别记为 $\lambda_1, \dots, \lambda_s$, 此时 A 的极小多项式 $m(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_s)$; 另一方面, 假设 T 的极小多项式为 $m(\lambda) = (\lambda - \lambda_1) \dots (\lambda - \lambda_s)$, 其中 $\lambda_1, \dots, \lambda_s$ 互不相同, 考虑

$$h_i(\lambda) = \frac{m(\lambda)}{(\lambda - \lambda_i)},$$

则 $\gcd(h_1, \dots, h_s) = 1$, 从而根据裴蜀定理存在 $k_1(\lambda), \dots, k_s(\lambda) \in \mathbb{F}[\lambda]$ 使得

$$k_1(\lambda)h_1(\lambda) + \dots + k_s(\lambda)h_s(\lambda) = 1.$$

即

$$k_1(T)h_1(T) + \dots + k_s(T)h_s(T) = \mathrm{I}_n,$$

其中 $n=\dim V$. 任取 $v\in V$, 我们可以将其分解为 $v=v_1+\cdots+v_s$, 其中

$$v_i=k_i(T)h_i(T)v.$$

如果我们记 $V_i=\ker(\lambda_i\mathrm{I}-T)$, 那么

(1) $v_i\in V_i$, 这是因为 $(A-\lambda_i\mathrm{I}_n)v_i=k_i(T)m(T)v=0$.

(2) 任取 $1\leq i\leq s$ 以及 $v\in V_i\cap\sum_{j\neq i}V_j$, 则有

$$0=(k_1(T)h_1(T)+\cdots+k_s(T)h_s(T))v=v.$$

从而根据命题3.5.2有

$$V=V_{\lambda_1}\oplus\cdots\oplus V_{\lambda_s}.$$

根据定理4.1.1可知 T 可对角化. □

推论 4.4.3. *n 维 $\mathbb{F}$-线性空间 V 上的 $\mathbb{F}$-线性映射如果有 n 个不同的特征值, 则其可对角化.*

证明. 假设特征多项式 $f(\lambda)=(\lambda-\lambda_1)\dots(\lambda-\lambda_n)$, 根据命题4.4.2可知此时极小多项式等于特征多项式, 并且是没有重根的一次因式乘积. □

推论 4.4.4. *$\mathbb{F}$-线性映射 $T\colon V\to V$ 可对角化, 假设 W 是 T-不变子空间, 则 $T|_W\colon W\to W$ 也可对角化.*

证明. 假设 $m_V(\lambda)$ 是 $T\colon V\to V$ 的极小多项式, 那么其也是 $T|_W\colon W\to W$ 的零化多项式, 从而 $m_W(\lambda)$ 整除 $m_V(\lambda)$. 从而如果 $m_V(\lambda)$ 可以分解为没有重根的一次因式乘积, $m_W(\lambda)$ 也可以分解为没有重根的一次因式乘积, 从而 $T|_W$ 可对角化. □

定义 4.4.2. *矩阵 $A\in M_{n\times n}(\mathbb{F})$ 称为***幂等矩阵** *(idempotent matrix), 如果 $A^2=\mathrm{I}_n$.*

命题 4.4.3. *假设 $\mathbb{F}$ 的特征不为 2, 则幂等矩阵 A 在 $\mathbb{F}$ 上可对角化.*

证明. 注意到 λ^2-1 是 A 的零化多项式, 从而其极小多项式整除 λ^2-1. 如果 $\mathbb{F}$ 的特征不为 2, λ^2-1 在 $\mathbb{F}$ 可以分解为 $(\lambda-1)(\lambda+1)$, 从而根据定理4.4.1可知 A 可对角化. □

定义 4.4.3. *矩阵 $A\in M_{n\times n}(\mathbb{F})$ 被称为***幂零矩阵** *(nilpotent matrix), 如果存在 $k\in\mathbb{N}_{\geq 0}$ 使得 $A^k=0$.*

命题 4.4.4. *幂零矩阵的特征值都为零.*

证明. 如果 A 是幂零矩阵, 则某个单项式 λ^k 是 A 的零化多项式, 特别地, 根据命题4.4.2有幂零矩阵的特征值都为零. □

推论 4.4.5. 幂零矩阵可对角化当且仅当其为零矩阵.

证明. 一方面, 如果其可对角化, 由于其特征值都为零从而有其为零矩阵; 另一方面, 零矩阵当然是可对角化的幂零矩阵. □

注 4.4.2. 上述命题也可以通过极小多项式的观点看出: 如果 A 是非零的幂零矩阵, 那么其极小多项式对于其 λ^k, $k \geq 2$, 从而不是没有重根的一次因式的乘积.

4.5 作业十三

4.5.1 基础题

习题 4.5.1 (4 选 2 做即可). 判断下列矩阵 A 是否可在复数域上对角化. 在可对角化的情形, 给出可逆矩阵 P, 使得 $P^{-1}AP$ 是对角矩阵.

1. $A = \begin{bmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 4 & 7 & -5 \\ -4 & 5 & 0 \\ 1 & 9 & -4 \end{bmatrix}$

3. $A = \begin{bmatrix} 4 & 2 & -5 \\ 6 & 4 & -9 \\ 5 & 3 & -7 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

习题 4.5.2. 已知 $A \in M_n(\mathbb{C})$ 的特征值 (按代数重数计) 为 $\lambda_1, \lambda_2, \cdots, \lambda_n$(其中可能有相同的数). 对多项式 $f = a_m X^m + \cdots a_0 \in \mathbb{C}[X]$, 定义 $f(A) = a_m A^m + \cdots + a_0 I_n$. 则 $f(A)$ 的特征值集合 (按代数重数计) 是什么?

习题 4.5.3. 固定复数 $b, c \in \mathbb{C}$. 假设有数列 (a_n) 满足递推公式 $a_{n+1} = ba_n + ca_{n-1}$. 以下你将利用矩阵对角化和上三角化来求通项公式.

1. 令向量 $x_n = (a_n, a_{n-1})^T$. 写下 x_n 的递推公式 $x_n = Ax_{n-1}$. 其中 A 是二阶复方阵.

2. 求矩阵 A 的特征值 λ_1, λ_2 并判断在复数域上是否可以对角化.

3. 如果 A 可以对角化, 利用对角化求 A^n.

4. 如果 A 不能对角化, 证明 A 相似于矩阵 $\lambda_1 I_2 + N$, 其中 $N^2 = 0$. 利用二项展开求 A^n.

习题 4.5.4. 证明复循环矩阵

$$\begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_n \\ a_n & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_2 & a_3 & a_4 & \cdots & a_1 \end{bmatrix}$$

(在复数域上) 可对角化, 并求该矩阵的特征值以及行列式. *(Hint:* 考虑循环方阵

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

和习题*4.5.2.)*

习题 4.5.5. 设 $A \in M_n(\mathbb{C})$ 满足 $\mathrm{tr}(A^k) = 0$, 对任何 $0 < k \leq n$. 证明 $A^n = 0$. *(Hint:* 考虑习题*4.5.2.)*

习题 4.5.6. 设 $A, B \in M_n(\mathbb{C})$ 满足 $AB = BA$, 证明 A 的某个特征值对应的特征子空间也是 B 的不变子空间.

习题 4.5.7. 设 $\mathcal{T} \subset M_n(\mathbb{C})$ 是一些交换矩阵构成的集合, 即对任何 $T_1, T_2 \in \mathcal{T}, T_1T_2 = T_2T_1$. 证明, $\mathcal{T}$ 可以同时上三角化, 即存在 $P \in \mathrm{GL}_n(\mathbb{C})$, 使得对任何 $T \in \mathcal{T}, P^{-1}TP$ 是上三角矩阵.

习题 4.5.8. 考虑平面上绕原点的旋转变换对应的矩阵, 求其复特征值, 并判断在复数域上是否可以对角化.

4.5.2 附加题

习题 4.5.9. 设 $A, B \in M_n(\mathbb{C})$ 满足 $AB - BA = A$, 证明 A, B 有公共的特征向量.

习题 4.5.10. 假设 $A = (a_{ij})$ 是复方阵. 定义 $D_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq \sum_{j \neq i, 1 \leq j \leq n} |a_{ij}|\}$ 为复平面上的圆盘. 证明 A 的特征值落在这些圆盘的并里 $\cup_i D_i$. (这称为 *Gershgorin circle theorem*, 有更精细的估计在哪些圆盘里有多少特征值的版本. 在估计特征值范围时还可以用对角矩阵 Λ 共轭作用于 A, 即考虑 $\Lambda^{-1}A\Lambda$ 来改变圆盘的位置, 请思考如何选取 Λ 来得到更精细的估计.)

4.6 作业十四

习题 4.6.1. 对域 F 上的二阶方阵 A，请通过直接计算验证 *Cayley-Hamilton* 定理：A 的特征多项式为 $f_A(x) = x^2 - \mathrm{tr}(A)x + \det(A)$，则 A 满足 $f_A(A) = 0$. 请思考该定理中将域上的方阵替换成任意环上是否依然成立.

习题 4.6.2. 利用 *Cayley-Hamilton* 定理，证明域 F 上的任意可逆方阵 A 的伴随都可以写成 A 的多项式. 请思考对于环 R 上的一般方阵 A*(不要求可逆)* 这个结论是否依然成立.

习题 4.6.3. 假设 T 是 F-线性空间 V 上的线性变换，W 是 T 的不变子空间. 如果 T 可对角化，请问 T 在 W 的限制是否一定可对角化? 如果是请证明，如果不是请给出反例.

习题 4.6.4. 对域 F 上的方阵 A，如果 μ 是 A 的特征值，证明 μ 一定是 A 的极小多项式的根.

习题 4.6.5. 假设复方阵 A 满足 $A^n = I$, 证明 A 可以复对角化.

习题 4.6.6. 假设两个实方阵在复数域上相似，证明其在实数域上也相似.

习题 4.6.7. 假设 n 阶实方阵 A 满足 $A^2 = -I$，证明 A 的阶数 n 为偶数，且 A 在实数域上不能对角化，在复数域上可以对角化. 请对每个偶数 n 给出一个这样的矩阵的例子.

习题 4.6.8. 找出以下矩阵 A 的极小多项式.

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

4.7 作业十五

习题 4.7.1. 在 *Google PageRank* 算法中,我们写下了一个 n 维向量 $x=(x_1,x_2,\cdots,x_n)^T$, 表示一个网页的权重。我们假设每个网页都有一个初始权重 x_i，并且第 j 个网页都有一个链到其他网页 i 的概率 p_{ij}。我们可以用一个 $n\times n$ 的矩阵 $P=(p_{ij})$ 来表示这个概率分布, 并且求解权重向量时，写下了方程 $Px=x$。请证明 P 有特征值 1, 并且所有复特征值的绝对值小于等于 1.

习题 4.7.2. 假设写出一个用极小多项式判定域 F 上的 n 阶方阵 A 是可上三角化的充分必要条件，并证明.

习题 4.7.3. 假设域 F 上的 n 阶方阵 A 是可对角化的，证明 A^T 也是可对角化的. 请思考当 A 可以上三角化时，A^T 是否也可以上三角化.

习题 4.7.4. 假设域 F 上的 n 阶方阵 A 是可对角化的，m 阶方阵 B 也是可以对角化的, 证明 F 线性空间 $M_{n\times m}(F)$ 上的线性变换 T，$T(X)=AXB$ 也是可对角化的. 请找出 T 的 *determinant* 和 *trace* 与 A 和 B 的关系. 请思考 A 和 B 可以上三角化时，T 是否也可以上三角化.

习题 4.7.5. 假设域 F 上的有限维线性空间有线性变换 T，和 T 的不变子空间 W, 则 T 诱导了商空间上 V/W 上的线性变换 $\widetilde{T}$. 证明 T 的特征多项式等于 $T|_W$ 和 $\widetilde{T}$ 的特征多项式的乘积.

习题 4.7.6. 对域 F 上的 n 阶方阵 A, 特征多项式 $\det(\lambda I-A)=\lambda^n+s_1\lambda^{n-1}+\cdots+s_n$ 的展开系数 s_k 定义了函数 $S_k\colon M_n(F)\to F, A\mapsto s_k$. 证明这些函数满足 $S_k(AB)=S_k(BA)$. 请思考将域 F 换成任意交换环 R, 这个结论是否成立.

习题 4.7.7. 假设可交换的复方阵 A 和 B, 满足 $B^n=0$. 证明 A 和 $A+B$ 有相同的特征多项式. 请思考将复数域换成任意域 F, 这个结论是否成立.

习题 4.7.8. 请验证 R-模 M 中满足以下等式

1. $0_R\cdot m=0_M,\forall m\in M$.
2. $-1\cdot m=-m,\forall m\in M$.
3. $r\cdot 0_M=0_M,\forall r\in R$.

习题 4.7.9. 请验证课上关于商模的构造的良定部分.

5 环论与模论

在 Cayley-Hamilton 定理的证明过程中, 我们看到给定有限维 $\mathbb{F}$-线性空间 V 以及其上的线性映射 T, 我们如下给 V 一个 $\mathbb{F}[\lambda]$-模结构:

$$\begin{aligned}\mathbb{F}[\lambda] \times V &\to V\\ (f(\lambda), v) &\mapsto f(T)v,\end{aligned}$$

并且满足若干性质, 将这些性质抽象出来, 就是本节中我们要研究的概念: 模, 而这个概念为我们研究线性映射提供了新的角度和思路.

线性代数中一个重要的问题是如何分类所有的线性映射, 如果用 $\mathcal{M}$ 记所有线性空间和线性映射的二元组构成的集合:

$$\{(V,T) \mid V \text{ 是 } \mathbb{F}\text{-线性空间}, T\colon V \to V \text{ 是 } \mathbb{F} \text{ 线性映射}\}.$$

但是可能上述集合中存在着很多 “相同” 的二元组, 这取决于我们在什么意义下分类. 因此在分类之前, 我们先在这个集合上定义一个等价关系, 使得我们的分类的结果更加简洁. 给定两个线性映射 $T_1\colon V_1 \to V_1$ 以及 $T_2\colon V_2 \to V_2$, 我们称 (V_1, T_1) 和 (V_2, T_2) 是等价的, 如果存在线性同构 $\psi\colon V_1 \to V_2$ 使得如下图表交换

$$\begin{array}{ccc} V_1 & \xrightarrow{T_1} & V_1 \\ {\scriptstyle\psi}\downarrow & & \downarrow{\scriptstyle\psi} \\ V_2 & \xrightarrow{T_2} & V_2. \end{array}$$

即 $T_1 = \psi^{-1} \circ T_2 \circ \psi$. 我们用 $\mathcal{M}$ 记商掉上述等价关系的集合, 即

$$\mathcal{M} = \{(V,T) \mid V \text{ 是 } \mathbb{F}\text{-线性空间}, T\colon V \to V \text{ 是 } \mathbb{F} \text{ 线性映射}\}/\sim.$$

我们可以将 $\mathcal{M}$ 分解成 $\mathcal{M} = \bigcup_{n\in\mathbb{Z}_{>0}} \mathcal{M}_n$, 其中

$$\mathcal{M}_n = \{(V,T) \mid V \text{ 是 } n \text{ 维 } \mathbb{F}\text{-线性空间}, T\colon V \to V \text{ 是 } \mathbb{F} \text{ 线性映射}\}/\sim.$$

因此为了研究 $\mathcal{M}$, 只需要对任意 $n \in \mathbb{Z}_{>0}$, 理解清楚 $\mathcal{M}_n$. 任取 $\mathcal{M}_n$ 中两个等价的线性映射 (V_1, T_1) 和 (V_2, T_2) 后, 分别取 V_1 和 V_2 的一组基 B_1 和 B_2, 则 B_1 和 B_2 分别给出了 V_1 和 V_2 到 $\mathbb{F}^n$ 的同构. 如果我们记 A_1 和 A_2 分别是 T_1 和 T_2 在基 B_1 和 B_2 下的矩阵, 那么 $T_1 = \psi^{-1} \circ T_2 \circ \psi$ 等价于 A_1 相似于 A_2, 从而我们有如下的一一对应

$$\mathcal{M}_n \overset{1\text{–}1}{\Longleftrightarrow} \{\text{域 } \mathbb{F} \text{ 上 } n \text{ 阶矩阵的相似类}\}.$$

当 (V,T) 可对角化时, 这意味着其在某一组基下的矩阵可以写成对角矩阵, 因此有一个很简单的代表元, 而当 T 不可对角化时, 此时想要找到一个” 良好” 的代表元并不是一件容易的事情,

在本节中, 我们将会说明等价的线性映射会给出 V 上同构的 $\mathbb{F}[\lambda]$-模结构, 从而通过分类 $\mathbb{F}[\lambda]$-模的结构来给不可对角化的矩阵一个好的描述.

5.1 环与理想

5.1.1 环与环同态

定义 5.1.1. 一个**环** *(ring)*, 是一个集合 R 有如下两种运算:

1. 加法运算 "$+$", 使得 R 构成一个阿贝尔群 $(R,+)$, 其中的单位元被记做 0;
2. 乘法运算 "$\times$", 使得 R 构成一个幺半群, 即对于乘法运算存在单位元 1 以及满足结合律;
3. 乘法与加法运算之间存在分配律.

注 5.1.1. 更严格的来说, 我们这里定义的是含有单位元的环, 这与一些教材上对环的定义不一样.

定义 5.1.2. 一个**交换环** *(commutative ring)*, 是指一个乘法运算交换的环.

注 5.1.2. 在此之后, 当我们提及环时, 总指的是交换环.

例 5.1.1. 一些交换环的例子:

- $(\mathbb{Z},+,\times,0,1)$;
- $(\mathbb{Q},+,\times,0,1)$;
- 零环 $R=\{0\}$, 即只有一个元素组成的环.

命题 5.1.1. 在环 R 中, 有 $0\times a=0$ 对任意 $a\in R$ 成立.

证明. 首先由于 $0+0=0$, 那么任取 $a\in A$, 根据分配律可知

$$\begin{aligned} 0\times a &= (0+0)\times a \\ &= 0\times a+0\times a \end{aligned}$$

两侧同时加上 $-(0\times a)$, 则有

$$0=0\times a$$

□

命题 5.1.2. 如果在环 R 中满足 $1=0$, 那么 R 是零环.

证明. 任取 $a\in R$, 则 $a=1\times a=0\times a=0$, 即 R 是零环. □

定义 5.1.3. 一个环 R 中的元素 a 被称为**单位** *(unit)*, 如果 a 存在一个乘法逆 b, 即存在 $b\in R$ 使得 $ab=ba=1$.

注 5.1.3. 与处理群中的逆元类似, 不难证明一个元素 a 如果存在逆那么其逆一定唯一, 我们通常记做 a^{-1}.

定义 5.1.4. 给定一个环 R, 其上的一个**形式多项式** *(formal polynomial)*, 是形如下式的元素

$$f(x) := a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0, \quad a_n \neq 0, a_i \in R, i = 0, 1, 2, \ldots, n$$

其中 x 被称为**单项式** *(monomial)*, n 被称作多项式的次数, 记做 $\deg f$. 两个形式多项式

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$$
$$g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_0$$

相等当且仅当 $m = n$, 并且 $a_i = b_i$ 对任意的 i 成立.

定义 5.1.5. 给定一个环 R, 其上的**多项式环** *(polynomial ring)*定义为

$$R[x] := \{f(x) \mid f(x)\text{是 } R \text{ 上的形式多项式}\}$$

其中给定 $f, g \in R[x]$, 加法乘法运算如下给出:

1. $f + g(x) := (a_0 + b_0) + (a_1 + b_1)x + \ldots$;
2. $fg(x) := a_0b_0 + (a_0b_1 + a_1b_0)x + (a_1b_1 + a_2b_0 + a_0b_2)x^2 + \ldots$.

习题 5.1.1. 验证 $R[x]$ 构成了一个环.

注 5.1.4. 如果在构造多项式环时取环 $R' = R[x]$, 其中 R 是一个环, 那么我们可以定义 R 上的二元多项式环 $R[x, y]$ 为 $R[x][y]$.

定义 5.1.6. 给定 $f(x), g(x) \in R[x]$, 如果存在 $q(x), r(x) \in R[x]$ 使得

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

其中 $\deg r < \deg f$, 那么我们称 $f(x)$ 带余式 $r(x)$ 除 $g(x)$, 这个过程被称为**带余除法** *(division with remainder)*

命题 5.1.3. 带余除法总可以进行, 只要除式 $f(x)$ 的首项 a_n 是 R 中的单位.

定义 5.1.7. 环 R 的一个**子环** *(subring)*指的是 R 的一个子集 R', 满足对加、减、乘运算封闭, 并且环 R 的 $0, 1$ 在 R' 中.

定义 5.1.8. 给定两个环 R_1, R_2, 一个映射 $f\colon R_1 \to R_2$ 被称为**环同态** *(ring homomorphism)*, 如果

1. f 保持加法和乘法运算;

2. $f(1_{R_1}) = 1_{R_2}$.

注 5.1.5. 注意, f 保持加法运算我们则一定有 $f(0_{R_1}) = 0_{R_2}$, 但是保持乘法运算不一定有 $f(1_{R_1}) = 1_{R_2}$(为什么？), 因此我们需要在第二条中要求这件事情.

定义 5.1.9. 一个环同态 f 被称为**环同构** *(ring isomorphism)*, 如果其作为映射是双射.

定义 5.1.10. 给定环同态 $f\colon R_1 \to R_2$, 映射的**核** *(kernel)*被定义为

$$\ker f = \{a \in R_1 \mid f(a) = 0_{R_2}\}.$$

命题 5.1.4. 对于环同态 $f\colon R_1 \to R_2$ 的核 $\ker f$, 我们有如下性质:

1. $\ker f$ 对加法构成子群;
2. 任取 $s \in \ker f, r \in R_1$, 有 $rs \in \ker f$;
3. 如果 f 是环同构, 那么 $\ker f = \{0_{R_1}\}$;
4. 如果 $1_{R_1} \in \ker f$, 那么 $\ker f = R_1$.

注 5.1.6. 上述命题表示 1_R 不一定在环同态的核中, 即一般来说环同态的核不是子环, 这与我们在群的时候情况并不一样.

5.1.2 理想

将一个环同态的核所具有的性质抽象出来, 这就得到了理想的定义.

定义 5.1.11. 环 R 的一个子集 I 被称为 R 的**理想** *(ideal)*, 如果

1. 对于加法 I 构成子群;
2. 任取 $s \in I, r \in R$, 有 $rs \in I$.

例 5.1.2. 对于环 R 来说, 其存在两个平凡理想:R 本身与 $\{0\}$, 其他的理想被称为非平凡理想.

例 5.1.3. 对环同态 $f\colon R_1 \to R_2$ 来说, $\ker f$ 是 R_1 的理想.

定义 5.1.12. 环 R 的一个理想 I 被称为**主理想** *(principal ideal)*, 如果存在 $s \in R$, 使得

$$I = (s) := \{rs \mid r \in R\}.$$

定义 5.1.13. 环 R 的一个理想 I 被称为**有限生成理想** *(finitely generated ideal)*, 如果存在 $s_1, s_2, \ldots, s_r \in R$, 使得

$$I = (s_1, \ldots, s_r) := \{\sum_{i=1}^{r} r_i s_i \mid r_i \in R\}.$$

命题 5.1.5. 给定一个环 R 以及一个理想 I, 在集合 R/I 上存在唯一的环结构, 使得如下映射是环同态:

$$\pi\colon R \to R/I$$
$$a \mapsto a + I.$$

我们称 R/I 是一个**商环** *(quotient ring)*.

证明. 首先如果只考虑 R 以及 I 上的加法结构, 显然 R/I 构成了一个商群, 因此我们只需要去定义 R/I 上的乘法结构即可. 注意到如果想要 π 是一个环同态, 我们只能如下定义我们的乘法结构: 任取 $a+I, b+I \in R/I$,

$$(a+I)(b+I) := ab + I.$$

此时我们需要验证我们的定义不依赖于代表元的选取, 即如果

$$a_1 + I = a_2 + I$$
$$b_1 + I = b_2 + I,$$

那么一定有

$$a_1 b_1 + I = a_2 b_2 + I.$$

根据理想的定义, 我们有

$$(a_1 - a_2)b_1 \in I,$$
$$a_2(b_1 - b_2) \in I.$$

上面两式相加即有 $a_1 b_1 - a_2 b_2 \in I$, 即我们的乘法运算是良好定义的. □

定理 5.1.1 (第一同构定理)**.** 如果 $f\colon R_1 \to R_2$ 是满的环同态, 那么 $R_1/I \cong R_2$, 其中 I 是 f 的核.

命题 5.1.6. 给定环同态 $f\colon R_1 \to R_2$ 以及 R_1 的一个理想 I, 我们记 $K = \ker f, \pi\colon R \to R/I$, 那么:

1. 如果 $I \subset K$, 那么存在唯一的映射 $\widetilde{f}: \overline{R} := R_1/I \to R_2$, 使得 $\widetilde{f} \circ \pi = f$.
2. $I = K$ 当且仅当上述映射 $\widetilde{f}$ 是一个同构.

例 5.1.4. 考虑赋值映射 $e_2\colon \mathbb{Q}[x] \to \mathbb{Q}$, 定义为 $e_2(p(x)) = p(2), p(x) \in \mathbb{Q}[x]$. 显然 $(x-2) \in \ker(e_2)$, 并且根据带余除法我们可知 $\mathbb{Q}[x]/(x-2) \cong \mathbb{Q}$, 从而由上述命题可知 $\ker(e_2) = (x-2)$ 是一个主理想.

定理 5.1.2 (对应定理)**.** 给定满的环同态 $f\colon R_1 \to R_2$, 并记 $K = \ker f$, 那么我们有如下的一一对应:

$$\{R_1\text{中包含 } K \text{ 的理想}\} \overset{1\text{–}1}{\Longleftrightarrow} \{R_2\text{中的理想}\},$$

并且一一对应如下给出:

1. 如果 $K \subset I$, 那么其对应到 R_2 中的理想 $f(I)$;

2. 如果 $\widetilde{I}$ 是 R_2 中的理想, 那么 $f^{-1}(\widetilde{I})$ 是 R_1 中包含 K 的理想.

5.1.3 极大理想

定义 5.1.14. 环 R 的一个真理想 I 被称为**极大理想** *(maximal ideal)*, 如果任何包含 I 的理想都是 R.

为了说明极大理想总是存在的, 我们需要 Zorn 引理, 而为了陈述这个技术性的引理, 我们需要引入偏序集这个概念.

定义 5.1.15. 给定一个集合 S, 其上的一个**偏序** *(partial order)*是一个关系 $R \subseteq S \times S$, 如果 $(a,b) \in R$, 我们记作 $a \leq b$, 并且 R 满足

1. $a \leq a$;

2. 如果 $a \leq b$ 并且 $b \leq a$, 那么 $a = b$;

3. 如果 $a \leq b$ 并且 $b \leq c$, 那么 $a \leq c$.

带有偏序关系的一个集合被称为**偏序集** *(partially ordered set)*.

定义 5.1.16. 偏序集 S 的一个子集 C 被称为一个**链** *(chain)*, 如果任取 $a, b \in C$, 我们有 $a \leq b$ 或者 $b \leq a$.

定义 5.1.17. 偏序集 S 的一个链 C 称为有上界, 如果存在 $c \in S$ 使得任意 $a \in C$ 都有 $a \leq c$.

引理 5.1.1 (Zorn 引理)**.** 如果偏序集 S 的每一条链都有上界, 那么 S 存在至少一个极大元.

命题 5.1.7. 对于环 R 来说, 极大理想总是存在的.

证明. 考虑集合 S 是由环 R 所有真理想组成的集合, 显然 S 非空, 并且包含关系是 S 上的一个偏序关系. 为了证明极大理想的存在性, 根据 Zorn 引理只需要对每一条由真理想组成的链, 其都存在上界即可. 给定链 $C = \{I_i\}_{i \in I}$, 考虑 $\widetilde{I} = \bigcup_{i \in I} I_i$, 我们有如下事实:

1. $\widetilde{I}$ 是一个理想;

2. $\widetilde{I} \neq R$, 这是因为 $1 \notin \widetilde{I}$, 以及一个理想 $I = R$ 当且仅当 $1 \in I$.

从而 $\widetilde{I}$ 是 C 的一个上界, 从而根据 Zorn 引理可知极大理想存在. □

命题 5.1.8. I 是 R 的极大理想当且仅当 R/I 是域.

例 5.1.5. 对于任意 $a \in \mathbb{C}$, 我们有 $(x-a) \subseteq \mathbb{C}[x]$ 是一个极大理想. 这个事实可以通过考虑如下映射来得到

$$\begin{aligned} e_a : \mathbb{C}[x] &\to \mathbb{C} \\ p(t) &\mapsto p(a). \end{aligned}$$

实际上不难验证 $\mathbb{C}[x]$ 的所有极大理想都是形如 $(x-a), a \in \mathbb{C}$ 的样子, 即在集合上有如下的一一对应:

$$\{\mathbb{C}[x]\text{的所有极大理想}\} \overset{1-1}{\Longleftrightarrow} \mathbb{C}.$$

例 5.1.6. 考虑实值连续函数环 $R = C((a,b),\mathbb{R})$, 其中 (a,b) 是开区间, 不难发现任取 $a \in \mathbb{R}$, 如下集合构成了 R 的一个极大理想

$$I_a = \{f \in R \mid f(a) = 0\}.$$

例 5.1.7. 更一般的, 我们有如下的一一对应

$$\{\mathbb{C}[x,y]\text{的所有极大理想}\} \overset{1-1}{\Longleftrightarrow} \mathbb{C}^2.$$

这是一个更加困难的事实, 被称为希尔伯特零点定理 *(Hilbert's Nullstellensatz)*. 如果承认这个事实, 回顾我们之前的一个例子, 环同态

$$\begin{aligned} \varphi\colon \mathbb{C}[x,y] &\to \mathbb{C}[t] \\ x &\mapsto t \\ y &\mapsto t^2 \end{aligned}$$

给出了同构 $\mathbb{C}[x,y]/(x-y^2) \cong \mathbb{C}[t]$, 从而有如下的一一对应

$$\{\mathbb{C}[x,y]/(x-y^2)\text{的所有极大理想}\} \overset{1-1}{\Longleftrightarrow} \{\mathbb{C}[t]\text{的所有极大理想}\}.$$

并且根据对应我们知道 $\mathbb{C}[t]$ 的极大理想 $(t-a)$ 会对应到 $\mathbb{C}[x,y]/(x-y^2)$ 中的极大理想 $(x-a, y-a^2)$, 从而 $\mathbb{C}[x,y]/(x-y^2)$ 的极大理想与 $\mathbb{C}^2$ 中的一条抛物线上的点一一对应. 这意味着去研究 $\mathbb{C}^2$ 中的一条曲线与研究 $\mathbb{C}[x,y]/(x-y^2)$ 的极大理想是一样的, 这也是代数几何的最初的想法.

5.2 整环

5.2.1 整环与素理想

定义 5.2.1. 环 R 被称为**整环** *(domain)*, 如果对于任意 $a,b \in R$ 满足 $ab = 0$, 那么 a 和 b 中至少有一个为 0.

定义 5.2.2. 环 R 的理想 I 被称为**素理想** *(prime ideal)*, 如果对任意 $a,b \neq I$, 一定有 $ab \neq I$.

命题 5.2.1. 对于环 R 的理想 I, R/I 是整环当且仅当 I 是素理想.

证明. 根据定义直接验证. □

命题 5.2.2. 极大理想一定是素理想.

证明. 注意到域一定是整环. □

当环 R 是整环时, 我们有着相对良好的分解性质, 并且有些时候这种分解是非常重要的, 例如下面的例子.

例 5.2.1. $\sqrt{2}$ 是无理数: 假设 $\sqrt{2}$ 不是无理数, 那么存在互素整数 p, q 使得 $\sqrt{2}q = p$, 即 $2q^2 = p^2$. 由于 p, q 互素从而根据整数的唯一分解性质有 $2 \mid p$, 不妨假设 $p = 2m$, 从而有 $q^2 = 2m^2$, 即 $2 \mid q$, 进而 $2 \mid \gcd(p, q) = 1$, 矛盾.

因此在本节接下来的内容中我们主要研究这件性质. 为了后续的需要, 我们来固定一些术语以及列出一些简单的事实, 其中涉及到的所有的元素都是在整环 R 中的:

1. u 是单位当且仅当 $(u) = R$;
2. a 整除 b[7]当且仅当 $(b) \subseteq (a)$;
3. a 是 b 的真因子[8]当且仅当 $(b) \subsetneq (a) \subsetneq R$;
4. a 和 b 相伴[9]当且仅当 $(a) = (b)$;
5. $a \neq 0$, a 不可约[10]当且仅当 $(a) \subseteq R$ 并且不存在主理想 (c) 使得 $(a) \subsetneq (c) \subsetneq R$;
6. p 是素元[11]当且仅当 (p) 是素理想.

5.2.2 主理想整环与欧几里德整环

定义 5.2.3. 整环 R 被称为**主理想整环** *(principal ideal domain)*, 如果 R 的每一个理想都是主理想.

定义 5.2.4. 整环 R 被称为**欧几里得整环** *(Euclidean domain)*, 如果存在函数 $\sigma\colon R\backslash\{0\} \to \mathbb{Z}_{\geq 0}$ 使得任取 $a, b \in R, b \neq 0$, 存在 $q, r \in R$ 使得

$$a = bq + r.$$

如果 $r \neq 0$, 则有 $\sigma(r) < \sigma(b)$.

[7] 即存在 c 使得 $b = ac$.
[8] 即 a 不是单位, 并且存在非单位的 c 使得 $b = ac$.
[9] 即存在某些单位 c 使得 $a = bc$.
[10] 即 a 不是单位并且 a 没有真因子.
[11] p 如果整除 ab, 那么 p 整除 a 或 p 整除 b

命题 5.2.3. 欧几里得整环是主理想整环.

例 5.2.2. 多项式环是欧几里得整环.

定义 5.2.5. 高斯整数环 *(Gauss integer ring)*定义为

$$\mathbb{Z}[i] := \{a + bi \mid a, b \in \mathbb{Z}\},$$

其中 i 是纯虚数.

例 5.2.3. 高斯整数环是欧几里得环, 其中 $\sigma(a+bi) = |a+bi|^2$. 考虑 $z_1 = a+bi, z_2 = c+di$, 其中 $a \neq 0, b \neq 0$. 如果想要写成

$$z_2 = z_1 q + r, \quad r \neq 0$$

的形式, 并且满足 $\sigma(r) < \sigma(z_2)$, 我们应该选 q 尽可能的接近 z_1/z_2. 我们先如下计算

$$\frac{z_1}{z_2} = (c+di)\frac{a-bi}{a^2+b^2} = m + ni, \quad m, n \in \mathbb{Q}.$$

选取 $m_0, n_0 \in \mathbb{Z}$ 使得

$$\begin{cases} |m - m_0| \leq \frac{1}{2}, \\ |n - n_0| \leq \frac{1}{2}. \end{cases}$$

我们记 $q = m_0 + n_0 i$, 那么

$$q - \frac{z_2}{z_1} = (m_0 - m) + (n_0 - n)i,$$

并且满足

$$|q - \frac{z_2}{z_1}|^2 = (m_0 - m)^2 + (n - n_0)^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2} < 1.$$

因此

$$|r|^2 = |z_2 - qz_1|^2 = |z_1(\frac{z_2}{z_1} - q)|^2 < |z_1|^2.$$

从而 $z_2 = qz_1 + r$ 满足 $\sigma(r) < \sigma(z_1)$.

5.2.3 唯一分解整环

给定整环 R, 以及非单位 $a \in R\backslash\{0\}$, 如果 a 不是不可约的, 那么 a 可以写成 $a = a_1 a_2$, 其中 a_1, a_2 都不是单位. 如果 a_1, a_2 中存在不是不可约的, 那么对其再进行类似的分解. 我们重复上述操作, 如果在有限次操作停止, 即得到的所有因子都是不可约的, 那么我们将 a 写成如下形式

$$a = a_1 \dots a_n,$$

其中 $a_1, \ldots, a_n$ 是 R 中的不可约的元素. 此时我们称对 a 的**分解终止** (factorization terminates), 并称 $a_1 \ldots a_n$ 是 a 的一个不可约分解. 如果 a 有两个不可约分解

$$\begin{aligned} a &= p_1 \ldots p_m \\ &= q_1 \ldots q_n. \end{aligned}$$

这两个不可约分解被称为相同的, 如果 $m = n$, 并且经过合适的顺序排列之后有 $(p_i) = (q_i)$.

例 5.2.4. *在高斯整数环 $\mathbb{Z}[i]$ 中, 对 5 有如下分解*

$$\begin{aligned} 5 &= (1+2i)(1-2i) \\ &= (2+i)(2-i), \end{aligned}$$

但是这两个分解是相同的, 因为 $(1+2i)i = (i-2)$, 以及 i 是 $\mathbb{Z}[i]$ 的单位.

定义 5.2.6. *整环 R 被称为***唯一分解整环** *(unique factorization domain), 如果任取非单位 $a \in R\backslash\{0\}$, 如果对 a 的分解终止, 并且得到的任何两个不可约分解是相同的.*

引理 5.2.1. *在整环 R 中, 任何素元都是不可约元.*

证明. 回忆:

1. p 是素元, 如果 $p \mid ab$ 可以推出 $p \mid a$ 或 $p \mid b$.
2. p 是不可约元, 如果 $p = ab$ 意味着 a 或 b 中一定有一个是单位.

下面我们来证明在整环中素元都是不可约元: 假设 p 是素元, 并且 $p = ab$, 那么不妨假设 $p \mid a$, 即 $a = pc$, 从而 $p = pcb$, 这意味着 $bc = 1$(这里用到了整环的性质), 从而 b 是单位, 即 p 是不可约元. □

引理 5.2.2. *在主理想整环 R 中, 任何不可约元都是素元.*

证明. 假设 p 不可约的, 从而不存在主理想 (c) 使得

$$(p) \subsetneq (c) \subsetneq (1).$$

由于 R 是主理想整环, 从而 (p) 是极大理想, 进而 (p) 是一个素理想, 这意味着 p 是素元. □

引理 5.2.3. *给定整环 R, 以及如下两条等价:*

1. 对任意非单位 $a \in R\backslash\{0\}$ 的分解终止;

2. R 不存在无穷严格上升的主理想链

$$(a_1) \subsetneq (a_2) \subsetneq \dots$$

证明. 对非单位 $a_1 \in R\backslash\{0\}$ 进行不可约分解

$$\begin{aligned} a_1 &= a_2 b_2 \\ &= a_2 a_3 b_3 \\ &= a_2 a_3 a_4 \dots \end{aligned}$$

我们可以得到一个严格上升的主理想链

$$(a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \dots$$

从这里可以观察出 (1) 与 (2) 的等价性. □

命题 5.2.4. *给定整环 R,*

1. *如果在 R 中对任何元素的分解都终止, 那么 R 是唯一分解整环当且仅当每个不可约元都是素元;*
2. *主理想整环是唯一分解整环.*

证明. (1). 我们先假设每个不可约元都是素元: 任取非单位 $a \in R\backslash\{0\}$, 有如下两个不可约分解

$$\begin{aligned} a &= p_1 \dots p_m \\ &= q_1 \dots q_n, \end{aligned}$$

那么 p_1 是素元, 并且 $p_1 \mid q_1 \dots q_n$, 从而 $p_1 \mid q_i, 1 \le i \le n$, 然而 q_i 是不可约的, 从而 p_1 和 q_i 是相伴的. 经过合适的顺序调整以及系数条件, 我们不妨假设 $p_1 = q_1$, 从而有

$$q_2 \dots q_n = p_2 \dots p_m$$

利用归纳法即可证明 $m = n$ 并且这两个分解是相同的, 从而 R 是唯一分解整环. 另一方面, 我们假设 R 是唯一分解整环, 取 a 是一个不可约元, 假设 $a \mid bc$, 不妨写作 $ad = bc$, 对 b, c, d 进行唯一分解, 即

$$\begin{aligned} b &= p_1 \dots p_m \\ c &= q_1 \dots q_n \\ d &= r_1 \dots r_s, \end{aligned}$$

那么由于唯一分解性, a 一定相伴于某个 p_i 或者 q_j, 即 $a \mid b$ 或 $a \mid c$.

(2). 由于在主理想整环中所有不可约元都是素元, 从而只需要证明任取非单位 $a \in R\backslash\{0\}$, 对 a 的分解终止即可, 根据之前的引理可知只需要证明任何严格包含的主理想链都会稳定即可, 考虑

$$(a_1) \subsetneq (a_2) \subsetneq (a_3) \subsetneq \dots$$

考虑 $I = \bigcup_{i=1}^{\infty}(a_i)$, 由于 R 是主理想整环从而 $I = (a_j)$ 对某个 j 成立, 这意味着严格升链是有限长的. □

例 5.2.5. 高斯整数环 $\mathbb{Z}[i]$ 是欧几里得整环, 从而是主理想整环, 从而是唯一分解整环.

例 5.2.6. $\mathbb{Z}[\sqrt{-5}]$ 不是唯一分解整环: 考虑 6 我们有如下分解

$$\begin{aligned} 6 &= 2 \times 3 \\ &= (1+\sqrt{-5})(1-\sqrt{-5}). \end{aligned}$$

定义 5.2.7. 给定整环 R 以及 $a, b \in R$, 如果存在 $d \in R, d \mid a, d \mid b$ 满足对任意 $c \mid a, c \mid b$ 有 $c \mid d$, 那么称 d 是 a, b 的**最大公因数** *(greatest common divisor)*, 记做 $\gcd(a, b)$.

注 5.2.1. 在一般的整环 R 中最大公因数不一定存在, 但是在唯一分解整环中最大公因数总是存在的: 将 a, b 做唯一分解

$$\begin{aligned} a &= p_1^{r_1} \dots p_s^{r_s} \\ b &= p_1^{t_1} \dots p_s^{t_s}, \end{aligned}$$

其中 $r_i, t_j \geq 0$. 不难发现最大公因子 d 为 $d = p_1^{\min\{r_1,t_1\}} \dots p_s^{\min\{r_s,t_s\}}$.

我们来考虑多项式版本的费马大定理:

命题 5.2.5. 假设 $f(x), g(x), h(x) \in \mathbb{C}[x]$, 则当 $n \geq 3$ 时

$$f^n + g^n = h^n$$

不存在解满足 $\deg f \geq 1$ 并且 $\gcd(f, g) = 1$.

证明. 假设存在解, 我们不妨假设 (f, g, h) 是满足 $\deg f + \deg g + \deg h$ 最小的解. 那么

$$\begin{aligned} f^n &= h^n - g^n \\ &= \prod_{k=0}^{n-1}(h - \xi^k g), \end{aligned}$$

其中 $\xi = e^{\frac{2\pi i}{n}}$. 那么 $\gcd(h, g) = 1$ 意味着 $\gcd(h - \xi_k g, h - \xi_l g) = 1$, 其中 $k \neq l$. 由于 $\mathbb{C}[t]$ 是唯一分解整环, 那么

$$\begin{aligned} h - g &= (x(t))^n, \\ h - \xi g &= (y(t))^n, \\ h - \xi^2 g &= (z(t))^n, \end{aligned}$$

并且由于 $n \geq 3$ 我们知道 $1, \xi, \xi^2$ 互不相同. 从而有

$$a(x(t))^n + b(y(t))^n = c(z(t))^n,$$

其中 $a, b, c \neq 0$, 从而得到了一个次数更小的解, 相矛盾. □

5.3 模

5.3.1 模的定义与例子

定义 5.3.1. 给定环 R, 一个**R-模** *(R-module)*是一个阿贝尔群 $(M, +)$ 以及一个 R 的作用 $R \times M \to M$, 记做 $(r, m) \mapsto rm$, 满足如下公理:

1. $1_R v = v$;
2. $(rs)m = r(sm)$;
3. $(r + s)m = rm + sm$;
4. $r(m_1 + m_2) = rm_1 + rm_2$.

其中 $r, s \in R$ 以及 $m_1, m_2 \in M$.

例 5.3.1. 域 $\mathbb{F}$ 上的线性空间 V 是 $\mathbb{F}$-模.

例 5.3.2. 任意阿贝尔群都是 $\mathbb{Z}$-模.

例 5.3.3. 给定域 $\mathbb{F}$, $\mathbb{F}$-线性空间 V 以及一个线性映射 $T\colon V \to V$, 此时 V 上具有一个 $\mathbb{F}[\lambda]$-模结构, 如下给出:

$$\begin{aligned} \mathbb{F}[\lambda] \times V &\to V \\ (f(\lambda), v) &\mapsto f(T)v. \end{aligned}$$

例 5.3.4. 给定环 R, 我们有如下 R-模:

1. 环 R 可以看成一个 R-模, 其中 $R \times R \to R$ 由 R 自身的乘法给出;
2. 环 R 的任何理想 I 可以看出一个 R-模;

3. $R^n = R \times \cdots \times R$*, 也可以看成是* R*-模, 被称为自由* R*-模.*

定义 5.3.2. 给定 R-模 M, M 的**子模** *(submodule)*是 M 的一个阿贝尔群子群 N, 并且对 R 的作用封闭.

命题 5.3.1. 环 R 的一个子集 I 是 R 作为 R-模的子模当且仅当 I 是 R 的理想.

证明. 直接根据定义. □

定义 5.3.3. 给定 R-模 M 以及 M 的子模 N, 商群 M/N 上有一个自然的 R-模结构, 由如下给出:

$$R \times M/N \to M/N$$
$$(r, m+N) \to rm + N.$$

如此得到的 R-模 M/N 称为**商模** *(quotient module)*

例 5.3.5. 给定多项式环 $R = F[x]$ 以及理想 $I = (f(x))$, 考虑 $M = F[x]/(f(x))$. 一方面 M 上有 $F[x]$-模结构. 另一方面, M 上有维数为 $n = \deg f(x)$ 的 F-线性空间结构, 并且 $\{1, x, x^2, \ldots, x^{n-1}\}$ 构成了 M 作为 F-线性空间的一组基. 此时 M 上 $F[x]$-模结构对应的线性变换在这组基下的表示矩阵为

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_n \\ 1 & 0 & \ldots & 0 & -a_{n-1} \\ 0 & 1 & \ldots & 0 & -a_{n-2} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_1 \end{pmatrix},$$

其中 $f(x) = x^n + a_1 x^{n-1} + \ldots a_{n-1}x + a_n$. 矩阵 A 被称为是多项式 $f(x)$ 的**友阵** *(companion matrix).*

定义 5.3.4. 给定 R-模 M_1, M_2, 阿贝尔群的外直和 $M_1 \oplus M_2$ 上有自然的 R-模结构, 由如下给出:

$$R \times M_1 \oplus M_2 \to M_1 \oplus M_2$$
$$(r, (m_1, m_2)) \to (rm_1, rm_2).$$

5.3.2 模同态与同态基本定理

定义 5.3.5.

(1) 给定 R-模 M_1, M_2, **R-模同态** *(R-module homomorphism)*是一个阿贝尔群 $\varphi\colon M_1 \to M_2$ 之间的群同态, 并且满足

$$\varphi(rm) = r\varphi(m)$$

对任意 $r \in R, m \in M_1$ 成立.

(2) R-模同态 φ 是**单射 *(满射，同构)***, 如果 φ 作为群同态是单射 *(*满射，同构*)*.

(3) R-模同态的**核 *(像)*** 定义为它作为群同态的核 *(*像*)*.

例 5.3.6. 给定 R-模 M 以及 M 的子模 N, 如果考虑 M/N 上自然的 R-模结构, 此时投影映射 $\pi\colon M \to M/N$ 是 R-模同态.

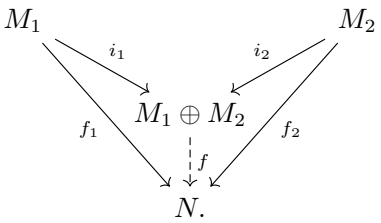
例 5.3.7. 给定 R-模 M_1, M_2 以及 R-模同态 φ, 此时 $\ker\varphi$ 是 M_1 的子模, $\operatorname{im}\varphi$ 是 M_2 的子模.

引理 5.3.1. 给定两个 $\mathbb{F}$-线性映射 $T_1\colon V_1 \to V_1$ 以及 $T_2\colon V_2 \to V_2$, 则 (V_1, T_1) 和 (V_2, T_2) 是等价的 $\mathbb{F}$-线性映射当且仅当 V_1 和 V_2 作为 $\mathbb{F}[\lambda]$-模是同构的.

证明. 如果 (V_1, T_1) 和 (V_2, T_2) 是等价的, 则存在 $\mathbb{F}$-线性同构 $\psi\colon V_1 \to V_2$ 使得 $T_1 = \psi^{-1} \circ T_2 \circ \psi$, 则 ψ 也给出了 $\mathbb{F}[\lambda]$-模之间的同构, 反之亦然. □

我们有如下与群论/环论/线性空间中平行的结果:

命题 5.3.2. 给定 R-模同态 $f_1\colon M_1 \to N, f_2\colon M_2 \to N$, 那么存在唯一的 R-模同态 $f\colon M_1 \oplus M_2 \to N$ 使得下图交换*:*



命题 5.3.3. 给定 R-模 M_1, M_2 以及 R-模同态 $\varphi\colon M_1 \to M_2$.

1. 如果 R-模 N 满足 $N \subset \ker\varphi$, 那么存在唯一的 R-模同态 $\overline{\varphi}$ 使得下图交换*:*
2. $M_1/\ker\varphi \cong \operatorname{im}\varphi$
3. *(*对应定理*)* 如果 φ 此时是满 R-模同态, 那么 M_1 包含 $\ker\varphi$ 的子模与 M_2 的子模一一对应.

定义 5.3.6. 给定 R-模 M 的子模 M_1, M_2, M 称为 M_1, M_2 的**内直和**, 如果满足

1. M_1, M_2 生成 M*;*
2. $M_1 \cap M_2 = \{0\}$.

命题 5.3.4. 如果 R-模 M 是子模 M_1, M_2 的内直和, 那么有 R-模同构 $M \cong M_1 \oplus M_2$.

定义 5.3.7. 给定 R-模 M_1, M_2, M_3 以及 R-模同态 $\varphi\colon M_1 \to M_2, \psi\colon M_2 \to M_3$, 如下序列被称为在 M_2 处**正合** *(exact)*

$$M_1 \xrightarrow{\varphi} M_2 \xrightarrow{\psi} M_3,$$

如果 $\ker\psi = \operatorname{im}\varphi$.

例 5.3.8. $\varphi\colon M_1 \to M_2$ 是满射当且仅当 $M_1 \xrightarrow{\varphi} M_2 \to 0$ 在 M_2 处正合; $\varphi\colon M_1 \to M_2$ 是单射当且仅当 $0 \to M_1 \xrightarrow{\varphi} M_2$ 在 M_1 处正合.

例 5.3.9. 下述 R-模序列被称为一个**短正合列** *(short exact sequence)*, 如果下列序列在 $M_i, i = 1, 2, 3$ 处都正合:

$$0 \to M_1 \xrightarrow{\varphi} M_2 \xrightarrow{\psi} M_3 \to 0.$$

5.3.3 自由模

定义 5.3.8. 给定集合 X 以及环 R, 定义

$$R^X := \{\sum_{x\in X} a_x x \mid a_x \in R, \text{ 只有有限多个 } a_x \text{ 非零}\}.$$

并且如下定义其上的 R-模结构:

$$\begin{cases} (\sum_{x\in X} a_x x) + (\sum_x b_x x) = \sum_{x\in X}(a_x + b_x)x, \\ r(\sum_{x\in X} a_x x) = \sum_{x\in X}(ra_x)x. \end{cases}$$

命题 5.3.5. 任何集合间的映射 $\varphi\colon X \to M$ 给出了一个 R-模映射 $\widetilde{\varphi}\colon R^X \to M$.

证明. 考虑

$$\begin{aligned} \widetilde{\varphi}\colon R^X &\to M \\ \sum_{x\in X} a_x x &\to \sum_{x\in X} a_x \varphi(x). \end{aligned}$$

□

定义 5.3.9. R-模 M 被称为**自由模** *(free module)*, 如果存在某个集合 X 使得 $M \cong R^X$ 作为 R-模同构. 特别地, M 被称为**有限生成自由模** *(finitely generated free module)*, 如果 X 是有限集.

定义 5.3.10. 给定 R-模 M, M 的子集 B 称为 M 的**基** *(basis)*, 如果:

1. M 由 B 生成, 即任取 $m \in M, m = \sum_{b\in B} r_b b$, 其中 $r_b \in R$ 并且只有有限多个 b 不是零;
2. B 是 R-线性无关的, 即如果 $0 = \sum_{b\in B} r_b b$, 则 $r_b = 0$.

命题 5.3.6. *给定 R-模 M 以及其的基 B, 那么有如下同构*

$$\begin{aligned} R^B &\to M \\ b &\mapsto b. \end{aligned}$$

特别地, 如果 $|B| = n$, 那么 $M \cong R^n$.

证明. 基的第一个条件保证了满射, 第二个条件保证了单射. □

注 5.3.1. *上面的命题告诉我们, R-模 M 是自由模当且仅当其存在一组基; R-模 M 是有限生成自由模当且仅当其存在一组有限基.*

当 R 是域的时候, 此时有限生成自由 R-模就是有限维线性空间, 并且我们知道对于有限维线性空间来说其维数就已经完全决定了它本身.

那么一个自然的问题就是, 对于有限生成自由模来说, 如果 B, C 都是它的基, 那么我们是否有 $|B| = |C|$ 呢? 为了解决这个问题, 我们需要借助取值在一般环上的矩阵, 并借助这个工具来研究我们的问题.

定义 5.3.11. *环 R 上的矩阵定义为*

$$M_{m\times n}(R) := (a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}},$$

其中 $a_{ij} \in R$. 为了方便, 我们记 $M_{n\times n}(R)$ 为 $M_n(R)$.

与域上的矩阵类似, 我们有如下的事情:

1. 矩阵乘法与之前相同, 并且此时矩阵乘法也具有结合律, 因为我们的环 R 具有结合律;
2. 我们可以同样地定义行列式映射

   $$\det\colon M_n(R) \to R,$$

   并且也满足 $\det(AB) = \det A \cdot \det B$ 对任意 $A, B \in M_n(R)$ 成立;
3. 对于 $A \in M_n(R)$, 我们可以同样地定义伴随矩阵 A^*, 并且满足 $A^*A = AA^* = \det A \cdot I_n$.

定义 5.3.12. *矩阵 $A \in M_{m\times n}(R)$ 被称为***可逆** *(invertible), 如果存在 $B \in M_{n\times m}(R)$ 使得 $AB = I_m$ 以及存在 $C \in M_{m\times n}(R)$ 使得 $CA = I_n$.*

命题 5.3.7. *如果 $A \in M_{m\times n}(R)$ 可逆, 此时一定有 $m = n$, 并且 A 可逆当且仅当 $\det A \in R^\times$, 此时 $B = C = (\det A)^{-1}A^*$.*

证明. 取 R 的一个极大理想 I, 此时 R/I 是一个域. 对 $AB = I_m, CA = I_n$ 中的分量做商, 在 R/I 中考虑等式

$$\overline{AB} = I_m, \quad \overline{CA} = I_n.$$

根据线性空间的结果我们可知 $m = n$.

现在假设 $A \in M_n(R)$ 是可逆的, 则 $1 = \det I_m = \det A \cdot \det B$ 意味着 $\det A \in R^\times$; 另一方面, 对 $B = C = (\det A)^{-1} I_n$ 直接验证 $AB = CA = I_n$. □

定义 5.3.13. 环 R 上的**一般线性群** *(general linear group)*定义为 $\mathrm{GL}_n(R) := \{R \in M_n(R) \mid R \text{ 可逆}\}$.

定义 5.3.14. 给定有限生成自由 R-模 M 以及其一组基 $B = \{b_1, \ldots, b_n\}$, 由于任意 $v \in M$ 可以被唯一写成 $v = \sum_{i=1}^n v_i b_i$, 我们称如下列向量为 v 在基 B 下的坐标:

$$[v]_B := \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

现在假设 M 的任何一个有限子集 $C = \{c_1, \ldots, c_m\}$, 以及 B 是 M 的一组基, 我们有:

命题 5.3.8. $([c_1]_B, \ldots, [c_m]_B) = P$ 可逆当且仅当 C 是 M 的一组基. 特别地, M 的任何两组基有相同的元素个数.

证明. 为了方便起见我们用 $(b_1, \ldots, b_n)[c_i]_B$ 去记:

$$c_i = \sum_{j=1}^n [c_i]_B^j b_j,$$

其中 $[c_i]_B^j$ 是 $[c_i]_B$ 的第 j-行. 利用这种记号, 我们可以记 $(c_1, \ldots, c_m) = (b_1, \ldots, b_n)P$.

假设 P 可逆, 即此时 $m = n$, 并且存在 $Q \in M_n(R)$ 使得 $PQ = I_n$, 那么:

$$\begin{aligned}(c_1, \ldots, c_n)Q &= ((b_1, \ldots, b_n)P)Q \\ &= (b_1, \ldots, b_n)(PQ) \\ &= (b_1, \ldots, b_n).\end{aligned}$$

从而:

1. 任取 $v \in M$, $v = (b_1, \ldots, b_n)[v]_B = (c_1, \ldots, c_n)(Q[v]_B)$, 即 M 可由 C 生成;
2. $0 = (b_1, \ldots, b_n)[0]_B = (c_1, \ldots, c_n)(Q[0]_B)$, 由于 $[0]_B$ 是全零组成的列向量, 从而 $Q[0]_B$ 也是.

综上所述 C 此时是 M 的一组基.

假设 B, C 都是 M 的基, 那么

$$(c_1, \ldots, c_m) = (b_1, \ldots, b_n) = P,$$
$$(b_1, \ldots, b_n) = (c_1, \ldots, c_n) = Q.$$

从而有

$$(c_1, \ldots, c_m) = (c_1, \ldots, c_m)QP,$$
$$(b_1, \ldots, b_n) = (b_1, \ldots, b_n)PQ,$$

从而 $QP = I_m, PQ = I_n$, 从而此时 P, Q 可逆, 并且 $m = n$. □

上面的结果告诉我们有限生成自由模的基的大小是良定义的, 即如果 $R^m \cong R^n$, 那么一定有 $m = n$. 我们利用的工具是 R 上的矩阵, 这与线性空间情形我们所做的事情几乎完全一致. 实际上, 我们完全可以通过一些约化将这个问题重新回归到线性代数的情形.

给定环 R 的一个理想 I 以及一个 R-模 M, 那么

$$IM := \{\sum_{i=1}^{n} r_i m_i \mid r_i \in I, m_i \in M\}.$$

是 M 的一个子模. 特别地, 如果 I 是一个主理想 (a), 此时 $IM = aM = \{am \mid m \in M\}$. 在 M/IM 上有一个自然的 R/I-模结构, 并且不难发现作为 R/I-模我们有如下同构:

$$(M_1 \oplus M_2)/I(M_1 \oplus M_2) \cong M_1/IM_1 \oplus M_2/IM_2,$$

其中 M_1, M_2 是 R-模.

推论 5.3.1. *给定环 R, 如果 $R^m \cong R^n$, 那么 $m = n$.*

证明. 取 R 的极大理想 I, 那么有作为 R/I-模我们有

$$(R/I)^n \cong R^n/IR^n \cong R^m/IR^m \cong (R/I)^m,$$

然而 R/I 是一个域, 从而

$$(R/I)^n \cong (R/I)^m$$

是作为线性空间的同构, 从而 $m = n$. □

总而言之, 有限生成自由模的结构总是简单的. 事实上, 我们更关心那些有限生成模的结构, 这也是下一节的关心内容.

5.4 有限生成模

定义 5.4.1. R-模 M 被称为是**有限生成的** *(finitely generated),* 如果存在一个有限生成模 R^n 以及如下满同态

$$R^n \xrightarrow{\varphi} M \to 0,$$

即 $M \cong R^n/\ker\varphi$.

注意到有限生成自由模的子模并不一定有限生成, 从而 $\ker\varphi$ 一般来说可能会很复杂. 现在我们期待一个比较好的条件, 使得此时的 $\ker\varphi$ 都是有限生成模.

命题 5.4.1. 对于环 R 来说, 如下条件等价:

1. 没有严格递增的理想序列

$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \dots,$$

其中 I_i 都是 R 的理想;

2. R 的任何理想都是有限生成的;

3. 任何有限生成 R-模的子模都是有限生成的.

满足上述条件之一的环被称为**诺特环** *(noetherian ring).*

证明. (1) 到 (2). 假设 I 不是有限生成的, 任取 $a_1 \in I$, 则 $(a_1) \subsetneq I$, 取 $a_2 \in I\backslash(a_1)$, 从而 $(a_1) \subsetneq (a_1, a_2) \subsetneq I$. 不断重复上述操作则可以得到一个理想的严格升链.

(2) 到 (1). 任给一个理想序列 $I_1 \subset I_2 \subset \dots$, 考虑 $I = \bigcup_{i=1}^{\infty} I_i$, 由于 I 是有限生成的, 这意味着存在 $a_1, \dots, a_n \in I$ 使得 $I = (a_1, \dots, a_n)$. 并且不难发现存在足够大的 N 使得 $a_1, \dots, a_n \in I_N$, 这意味着这个理想升链到 I_N 时已经稳定, 即不是严格上升.

(3) 到 (2). 注意到 R 作为 R-模的子模恰是 R 的理想;

(2) 到 (3). 首先我们有如下观察:

1. 对于正合列 $M_1 \xrightarrow{\varphi} M_2 \to 0$, 如果 M_1 是有限生成的, 那么 M_2 也是. 这是显然的, 因为 M_1 是有限生成的等价于存在满射 $R^n \to M_1$, 只需要复合 φ 即可得到 R^n 到 M_2 的满射;

2. 对于短正合列 $0 \to M_1 \xrightarrow{\varphi} M_2 \xrightarrow{\psi} M_3 \to 0$, 如果 M_1, M_3 都是有限生成的, 那么 M_2 也是. 假设 M_3 由 $v_1, \dots, v_n$ 生成, 那么选取 $\overline{v}_1, \dots, \overline{v}_n$ 使得 $\psi(\overline{v}_i) = v_i, 1 \le i \le n$. 假设 M_1 由 $w_1, \dots, w_m$ 生成, 并且记 $\widetilde{w}_i = \varphi(w_i)$. 任取 $\widetilde{v} \in M_2$, 记 $\psi(\widetilde{v}) = \sum_{i=1}^n r_i v_i$, 那么

$$\psi(\widetilde{v} - \sum_{i=1}^{n} r_i \widetilde{v}_i) = 0.$$

从而 $\widetilde{v} - \sum_{i=1}^n r_i\widetilde{v}_i \in \ker\psi = \operatorname{im}\varphi$, 即 $\widetilde{v} - \sum_{i=1}^n r_i\widetilde{v}_i = \sum_{j=1}^m s_j\widetilde{w}_j$. 即 M_2 可由 $\widetilde{v}_1, \dots, \widetilde{v}_n, \widetilde{w}_1, \dots, \widetilde{w}_m$ 生成.

根据上述观察, 我们做以下约化:

1. 根据对应定理以及观察 (1), 我们只需要对有限生成自由模 R^n 证明其所有的子模都是有限生成的即可.
2. 注意到我们有如下的短正合列:
$$0 \to R^{n-1} \to R^n \to R \to 0.$$
3. 利用归纳, 我们只需要证明 R 作为 R-模的子模都是有限生成的即可, 然而这恰是 (2).

□

命题 5.4.2. *如果 R 是诺特环, 则对任意理想 I 有 R/I 是诺特环.*

定理 5.4.1. *如果 R 是诺特环, 则 $R[x]$ 是诺特环.*

例 5.4.1. *R 是主理想整环, 则 R 是诺特环.*

定义 5.4.2. *给定环 R, R-模 N 被称为***诺特模** *(noetherian module), 如果任何子模的升链都稳定, 即如果有如下的子模升链*
$$0 \subseteq N_1 \subseteq N_2 \subseteq \dots,$$
则存在 m 使得 $N_m = N_{m+1} = \dots$.

命题 5.4.3. *给定环 R, 如果有如下 R-模的正合列*
$$0 \to N \to M \to M/N \to 0,$$
则 M 是诺特模当且仅当 $N, M/N$ 是诺特模.

定理 5.4.2. *假设 R 是诺特环, 则任何有限生成 R-模都是诺特模.*

现在如果 R 是诺特环, M 是有限生成 R-模, 即存在如下正合列:
$$\ker\varphi \to R^n \xrightarrow{\varphi} M \to 0,$$
并且此时 $\ker\varphi$ 是有限生成的, 从而存在满态射 $R^m \to \ker\varphi$, 即我们可以写作正合列:
$$R^m \xrightarrow{\psi} R^n \xrightarrow{\varphi} M \to 0,$$
并且 $M \cong R^n/\operatorname{im}\psi$. 如果我们能对态射 ψ 有很好的刻画, 那么我们就可以对诺特环上的有限生成模有一个较好的刻画.

给定 R^n 的一组基 $B=\{b_1,\dots,b_n\}$ 以及 R^m 的一组基 $C=\{c_1,\dots,c_m\}$, 我们可以将 ψ 写成矩阵的形式, 即:

$$(\psi(c_1),\dots,\psi(c_m))=(b_1,\dots,b_n)A,$$

其中 $A=(a_{ij})$ 由如下关系给出:

$$\psi(c_j)=\sum_{i=1}^{n}a_{ij}b_i.$$

为了表示 A 对于基 B,C 的依赖性, 我们通常也记做 $[\psi]_B^C:=A$. 如果我们能够选取 R^n,R^m 合适的基 B,C, 使得矩阵 $[\psi]_B^C$ 有尽可能简单的形式, 从而我们可以得到我们期待的结果. 幸运的是, 当 R 是主理想整环的时候, 确实有如此好的事情.

注 5.4.1. 由于 $\psi\circ\varphi=0$, 从而

$$\sum_{i=1}^{n}a_{ij}\psi(b_i)=0,$$

并且由于 $\{\psi(b_1),\dots\psi(b_n)\}$ 生成了 M, 因此我们可以将 M 视作是由 $\{\psi(b_1),\dots\psi(b_n)\}$ 生成, 并且满足上述关系.

5.4.1 主理想整环上的有限生成模

从现在开始到本节结束, 除非特殊说明, 我们总假设 R 是指主理想整环.

定理 5.4.3. 给定 R-模同态 $\psi\colon R^m\to R^n$, 我们可以选出 R^m,R^n 的基使得 ψ 在这些基下的表达式为

$$[\psi]_B^C=\left[\begin{array}{cccc|ccc}d_1&0&\cdots&0&0&\cdots&0\\0&d_2&\cdots&0&0&\cdots&0\\\vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&d_s&0&\cdots&0\\\hline0&0&\cdots&0&0&\cdots&0\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\0&0&\cdots&0&0&\cdots&0\end{array}\right]_{n\times m},$$

其中 $d_1\mid d_2\mid\cdots\mid d_s$.

推论 5.4.1. 如果 M 是有限生成 R-模, 则

$$M\cong R/(d_1)\oplus R/(d_2)\oplus\cdots\oplus R/(d_s)\oplus R^{n-s},$$

其中 $d_1\mid d_2\mid\cdots\mid d_s$.

定理 5.4.4. 给定有限生成 R-模 M, 如果 $(d_1) \neq R$, 我们有 $d_1, \ldots, d_s$ 被 M 完全决定. 此时 $(d_1, \ldots, d_s)$ 称为 M 的不变因子组.

例 5.4.2. $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \cong \mathbb{Z}/6\mathbb{Z}$, 即此时 $d_1 = 6$.

例 5.4.3. $\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/9\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/18\mathbb{Z}$, 此时 $d_1 = 3, d_2 = 18$.

定理5.4.3的证明. 首先给定 R^n, R^m 的基 B, C, 并且令 $B' = BP, C' = CQ$, 那么有

$$[\psi]_{B'}^{C'} = P^{-1}[\psi]_B^C Q$$

这是 ψ 在基变换下的变换公式. 回忆我们在线性空间的时候, 我们可以通过初等行列变换来得到矩阵的相抵标准型, 这里我们要做类似的事情. 给定矩阵非零矩阵

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

首先我们假设 R 是欧几里德整环, 其上的带有的函数记做 σ.

1. 第一步: 通过变换行与列, 我们可以选取 a_{11} 使得 $\sigma(a_{11}) = \min_{i,j}\{\sigma(a_{ij}) \mid a_{ij} \neq 0\}$
2. 第二步: 通过带余除法, 我们可以做到 $\sigma(a_{1i}) < \sigma(a_{11}), \sigma(a_{1j}) < \sigma(a_{11})$, 对任意的 $i = 2, \ldots, m, j = 2, \ldots, n$.
3. 此时再重复上述一、二两个步骤, 经过有限次行列变换之后, 我们可以得到 $a_{11} \neq 0, a_{1i} = a_{1j} = 0$, 对任意的 $i = 2, \ldots, m, j = 2, \ldots, n$ 成立. 即此时有:
$$\begin{pmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{pmatrix}$$
4. 第三步: 此时我们需要 a_{11} 可以整除右下角的 $(n-1) \times (m-1)$ 阶的子矩阵中每一个元素. 如果不然, 不妨假设 $a_{11} \nmid a_{i2}$, 我们不妨将第 i 行加到第一行得到
$$\begin{pmatrix} a_{11} & a_{i2} & \ldots & a_{in} \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$
此时再重复第一、二步即可.

5. 通过归纳假设, 我们可以得到经过有限次行列变换之后, A 的相抵标准型为我们期待的结果.

对于 R 是主理想整环的时候, 我们不能再进行辗转相除法, 但是整体上的思路是几乎一致的:

1. 第一步: 通过变换行与列, 我们可以选取 $a_{11} \neq 0$.
2. 第二步: 当 $i \geq 2$ 时, 我们考虑如下两种情况:
   (a) 如果 $(a_{1i}) \subset (a_{11})$, 那么直接通过行列变换将 a_{1i} 消去;
   (b) 如果 $(a_{1i}) \subsetneq (a_{11})$, 那么考虑 (a_{11}, a_{1i}) 生成的主理想 (d), 此时有
   $$d = ka_{11} + la_{1i}$$
   $$a_{11} = md$$
   $$a_{1i} = nd$$
   其中 $k, l, m, n \in R$. 考虑 A 右乘 B, 其中 B 是 $(1,1), (1,i), (i,1), (i,i)$ 元分别为 $k, -n, l, m$, $j \neq 2$ 的时候 (j,j) 元为 1, 其余都为零的可逆矩阵. 如果用 A' 去记这个新得到的矩阵, 即 $a'_{11} = d$, 那么我们有 $(a'_{11}) \supsetneq (a_{11})$. 如果还存在 $i \geq 2$ 使得 $(a'_{1i}) \subsetneq (a'_{11})$, 则不断重复上述操作, 由于升链总会稳定, 因此在有限步行列变换之后我们一定可以得到一个矩阵, 使得第一行除了 $(1,1)$ 元以外都是零; 同样的可以对列做同样的事情, 最终可以得到
   $$\begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & \end{pmatrix}.$$
3. 第三步: 同之前.

□

注 5.4.2. 在主理想整环证明的情形, 我们实际上也可以找一个函数来替代欧几里德整环情形下的 σ: 给定 $r \in R$, 我们用 $\sigma(r)$ 记 r 的素因子分解中的素因子个数. 此时我们需要一个类似辗转相除的操作来使得进行这个操作后得到的元素满足函数 σ 在其上的值严格小于操作前. 这实际上是我们在第二步 (b) 中的操作: 假设 $(a_{1i}) \subsetneq (a_{11})$, 我们可以找到一个可逆矩阵 B' 使得 AB' 的 $(1,i)$ 元为 a_{1i}, a_{11} 的最大公因子 d, 显然 d 的素因子分解的素因子个数严格小于 a_{1i}, 这便达成了我们的要求.

定义 5.4.3. 如果 R 是一个整环, 给定 R-模 M, 其**挠子模** *(torsion submodule)*定义为:

$$M_{tor} := \{m \in M \mid \text{存在 } r \in R\backslash\{0\} \text{ 使得 } rm = 0\}.$$

注 5.4.3. 可以根据定义直接验证 M_{tor} 是一个子模.

给定一个有限生成 R-模 M, 首先我们可以将其分解为

$$M = R/(d_1) \oplus \cdots \oplus R/(d_s) \oplus R^{n-s}.$$

我们现在的目的在于说明如果 $(d_1) \neq R$, $d_1, \ldots, d_s$ 是被 M 唯一决定的. 首先从上述同构可以看出 $M_{tor} \cong R/(d_1) \oplus \cdots \oplus R/(d_s)$, 即我们有如下短正合列:

$$0 \to M_{tor} \to M \to R^{n-s} \to 0,$$

此时我们有 s 是内蕴的, 即 s 由 M 唯一决定. 为了想要进一步说明 $d_1, \ldots, d_s$ 都由 M 决定, 我们需要考虑更精细的结构.

定义 5.4.4. 给定一般的环 R, R-模 M 被称为**循环模** *(cyclic module)*, 如果存在下述正合列:

$$R \to M \to 0.$$

命题 5.4.4. 假设 R-模 M 是循环模, 则有如下分解

$$M \cong R/(p_1^{n_1}) \oplus \cdots \oplus R/(p_k^{n_k}),$$

其中 $p_1, \ldots, p_k$ 是互不相同的素元.

证明. 由于 R-模 M 是循环模, R 是主理想整环, 从而 $M \cong R/(d), d \neq 0$, 将 d 做如下分解:

$$d = p_1^{n_1} \ldots p_k^{n_k},$$

其中 $p_1, \ldots, p_k$ 是不同的素元, 再利用中国剩余定理即可. □

现在回到我们的情况, 假设 R-模 M 有如下同构

$$M \cong R/(d_1) \oplus \cdots \oplus R/(d_s).$$

我们可以对 d_s 做如下分解:

$$d_s = p_1^{n_{1,s}} \ldots p_k^{n_{k,s}},$$

其中 $n_{i,s} \geq 1, 1 \leq i \leq k$. 由于有整除关系, 我们可知对任意的 $1 \leq r \leq s$, 我们有

$$d_r = p_1^{n_{1,r}} \ldots p_k^{n_{k,r}},$$

其中 $0 \leq n_{i,1} \leq \cdots \leq n_{i,r} \leq \cdots \leq n_{i,s}, 1 \leq i \leq k$. 从而我们有如下分解

$$\begin{aligned} M \cong \quad & R/(p_1^{n_{1,1}}) \oplus \cdots \oplus R/(p_k^{n_{k,1}}) \\ & \oplus R/(p_1^{n_{1,2}}) \oplus \cdots \oplus R/(p_k^{n_{k,2}}) \\ & \cdots \\ & \oplus R/(p_1^{n_{1,s}}) \oplus \cdots \oplus R/(p_k^{n_{k,s}}). \end{aligned} \tag{4}$$

我们称 $\{p_i^{n_{i,r}} \mid n_{i,r} \geq 1\}$ 为 M 的初等因子.

命题 5.4.5. 给定 R-模 M, 其初等因子组与不变因子组之间可以相互决定.

证明. 与线性代数的时候情形相同. □

定义 5.4.5. R 是一个整环, 给定环 R 的一个素元 p 以及 R-模 M, 其**p-挠子模** *(p-torsion submodule)*定义为:

$$M_{(p)} := \{m \in M \mid \text{存在 } n \in \mathbb{Z}_{\geq 1} \text{ 使得 } p^n m = 0\}.$$

注 5.4.4. 可以根据定义直接验证 $M_{(p)}$ 是一个子模.

命题 5.4.6. 分解4可以被写成

$$M \cong M_{(p_1)} \oplus \cdots \oplus M_{(p_s)}.$$

证明. 任取 $m = m_1 + \cdots + m_s$, 其中 $m_i \in R/(p_i^{n_{i,1}}) \oplus \cdots \oplus R/(p_i^{n_{i,s}})$. 如果 $m \in M_{(p_1)}$, 那么存在 $n \in \mathbb{Z}_{\geq 1}$ 使得

$$p_1^n m = 0,$$

这意味着 $p_1^n m_i = 0$ 对每一个 i 都成立. 现在假设 $i \geq 2$, 我们把 m_i 更详细的写成 $m_i = m_{i,1} + \cdots + m_{i,s}$, 其中 $m_{i,r} \in R/(p_i^{n_{i,r}})$. 那么有

$$p_1^n(m_{i,1} + \cdots + m_{i,s}) = 0,$$

这意味着 $p_1^n m_{i,r} = 0$ 对 $1 \leq r \leq s$ 都成立. 由于 p_1 和 $p_i, i \geq 2$ 互素可知 $m_{i,r} = 0$, 即此时有 $m \in R/(p_1^{n_{1,1}}) \oplus \cdots \oplus R/(p_1^{n_{1,s}})$, 即

$$M_{(p_1)} \subseteq R/(p_1^{n_{1,1}}) \oplus \cdots \oplus R/(p_1^{n_{1,s}}).$$

反包含关系是显然的, 从而我们有

$$M_{(p_1)} = R/(p_1^{n_{1,1}}) \oplus \cdots \oplus R/(p_1^{n_{1,s}}).$$

即

$$M \cong M_{(p_1)} \oplus \cdots \oplus M_{(p_s)}.$$

□

上述结果表明 M 的初等因子组可以由 $M_{(p_i)}$ 的分解来决定, 那么问题在于我们如何确定出每一个 $M_{(p_i)}$ 该如何分解呢?

例如我们先考虑下述简单的例子:

$$M \cong \mathbb{Z}/(p^2) \oplus \mathbb{Z}/(p) \oplus \mathbb{Z}/(p),$$

其中 p 是一个素数. 我们现在想定义一个内蕴的量来区分 p 和 p^2, 自然的想法如下:

定义 5.4.6. *R 是一个一般的环, p 是 R 的一个素元, M 是一个 R-模, 我们定义如下子模*

$$M[p] := \{m \in M \mid pm = 0\}.$$

并且不难发现 $M[p]$ 上存在 $R/(p)$-模结构.

特别地, 当 R 是主理想整环, p 是 R 的素元时, $R/(p)$ 是一个域, 即此时 $M[p]$ 是一个线性空间, 并且

$$\dim_{R/(p)} M[p]$$

完全由 M 本身决定, 例如在上面的例子中, 可以直接看出 $M[p] = \mathbb{Z}/(p) \oplus \mathbb{Z}/(p) \oplus \mathbb{Z}/(p)$, 此时 $M[p]$ 作为 $\mathbb{Z}/(p)$-线性空间的维数为 3.

命题 5.4.7. *假设 R-模 M 同构于*

$$M \cong R/(p^{n_1}) \oplus \dots R/(p^{n_s}),$$

其中 p 是 R 的素元, $n_1 \le \cdots \le n_s$, 那么 $M[p] \cong (R/(p))^{\oplus s}$.

到这里实际上我们几乎已经解决了我们的问题. 类似地我们可以定义 $M[p^2]$ 等等, 从 $M \cong \mathbb{Z}/(p^2) \oplus \mathbb{Z}/(p) \oplus \mathbb{Z}/(p)$ 我们可以看出 $\dim_{\mathbb{Z}/(p)} M$ 确定了 ≥ 1 的 n_r 的个数, 同样的 $\dim_{\mathbb{Z}/(p^2)} M$ 确定了 ≥ 2 的 n_r 的个数.

因此对于 R-模 M, 如果我们想要确定其初等因子组 $\{p_i^{n_{i,r}} \mid n_{i,r} \ge 1\}$, 我们只要对每一个可能的 p_i 不断计算 $M[p_i], M[p_i^2], \dots$ 的维数, 我们就可以从 M 确定其初等因子组. 再由于初等因子组可以唯一确定不变因子组, 至此我们证明了定理5.4.4.

定义 5.4.7. *如果 R 是一个整环, R-模 M 被称为***无挠的***, 如果 $M_{tor} = 0$.*

推论 5.4.2. *M 是有限生成 R-模, R 是主理想整环[12], 则 M 无挠等价于 M 自由.*

推论 5.4.3 (有限阿贝尔群的结构定理)**.** *M 是有限生成阿贝尔群, 则有*

$$M \cong \mathbb{Z}/d_1\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/d_s\mathbb{Z} \oplus \mathbb{Z}^{\ell},$$

其中 $2 \le d_1 \mid \cdots \mid d_s$, $\ell \ge 0$, 并且这些数由 M 唯一决定.

[12] 虽然在本节开头已经声明如无特殊声明 R 都是主理想整环, 但在这里我们还是依然强调这个事实.

5.4.2 应用：若尔当标准型

另一个主理想整环上有限生成模的结构定理的重要的应用, 就是用这个观点给出线性映射的分类. 给定有限维 $\mathbb{F}$-线性空间 V 以及其上的线性映射 T, 我们如下给 V 一个 $\mathbb{F}[\lambda]$-模结构:

$$\mathbb{F}[\lambda] \times V \to V$$
$$(f(x), v) \mapsto f(T)v$$

取定 V 的一组基 $B = \{v_1, \ldots, v_n\}$, 并用矩阵 $[T]_B$ 来记 T 在这组基下的矩阵. 由于作为 $\mathbb{F}$-线性空间 V 可以由 B 生成, 从而作为 $\mathbb{F}[\lambda]$-模, V 也可以由 B 生成.

根据注记5.4.1, 可知只要将 $(\lambda I - [T]_B^t)$ 通过行列变换化为

$$\left[\begin{array}{cccc|ccc} d_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_s & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right]_{n\times n}$$

的形式, 则有 V 作为 $\mathbb{F}[\lambda]$-模同构于

$$\mathbb{F}[\lambda]/(d_1) \oplus \cdots \oplus \mathbb{F}[\lambda]/(d_s) \oplus (\mathbb{F}[\lambda])^{n-s}$$

其中 $d_1 | \cdots | d_s$ 是首一多项式. 由于 $\mathbb{F}[\lambda]$ 作为 $\mathbb{F}$-线性空间是无穷维的, 而 V 作为 $\mathbb{F}$-线性空间是有限维的, 从而一定有 $n = s$. 即作为 $\mathbb{F}[\lambda]$-模有如下同构

$$V \cong \bigoplus_{i=1}^{n} \mathbb{F}[\lambda]/(d_i)$$

并且可以发现 T 的特征多项式 $f(\lambda)$ 等于 $\prod_{i=1}^n d_i$. 注意到任取 $v \in V$, $d_n([T]_B)v = 0$, 从而 $f([T]_B) = 0$, 这便又给出了 Cayley-Hamilton 定理的另一个证明. 这些 $\{d_i\}_{i=1}^n$ 称为 T 的**不变因子组**, 并且 d_n 就是 T 的极小多项式.

命题 5.4.8.

$$\{f \in \mathbb{F}[\lambda] \mid f(T) = 0\} = (d_n).$$

证明. 从上述分解即得. □

由于不变因子组 $\{d_n\}_{i=1}^n$ 可能存在常多项式, 不妨假设从 $i \geq k$ 开始 d_i 不是常多项式, 并且假设 $d_i = x^{n_i} + a_{n_i-1}x^{n_i-1} + \cdots + a_0$, 那么 $\mathbb{F}[\lambda]/(d_i)$ 视作 $\mathbb{F}$-线性空间有一组基 $\{1, x, \ldots, x^{n_i-1}\}$, 并且

$$T\left(1, x, \ldots, x^{n_i-1}\right) = \left(1, x, \ldots, x^{n_i-1}\right) A_i$$

其中 A_i 是 d_i 的友阵. 因此我们可以找到 V 的一组基使得 T 在这组基下的矩阵为分块对角矩阵 $\operatorname{diag}\{A_k,\dots,A_n\}$, 其中 A_i 是 d_i 的友阵, 这叫作 T 的**有理标准型** (Frobenius normal form).

如果此时我们考虑域 $\mathbb{F}$ 是复数域 $\mathbb{C}$, 那么我们不妨记 $d_i = \prod_{j=1}^{l} p_j^{n_{ij}}$, 其中 $p_j^{n_{ij}} = (\lambda-\lambda_j)^{n_{ij}}$. 那么我们则有

$$\mathbb{C}[\lambda]/(d_i) = \bigoplus_{j=1}^{\ell} \mathbb{C}[\lambda]/(\lambda-\lambda_j)^{n_{ij}},$$

对于每一个因子 $\mathbb{C}[\lambda]/(\lambda-\lambda_j)^{n_{ij}}$ 来说, 其有如下的一组 $\mathbb{C}$-基,

$$1, \lambda-\lambda_j, \dots, (\lambda-\lambda_j)^{n_{ij}-1},$$

并且不难发现

$$\lambda \cdot \begin{pmatrix} 1 \\ \lambda-\lambda_j \\ \vdots \\ (\lambda-\lambda_j)^{n_{ij}-1} \end{pmatrix} = J_{\lambda_j}(n_{ij}) \begin{pmatrix} 1 \\ \lambda-\lambda_j \\ \vdots \\ (\lambda-\lambda_j)^{n_{ij}-1} \end{pmatrix}$$

其中

$$J_{\lambda_i}(n_{ij}) = \begin{pmatrix} \lambda_j & 1 & & & \\ & \lambda_j & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_j & 1 \\ & & & & \lambda_j \end{pmatrix}_{n_{ij}\times n_{ij}}$$

被称为 Jordan 块, 并且从这里可以直接看出 λ_j 是 T 的一个特征值. 并且我们可以找到一组基, 使得 T 在这组基下的矩阵为分块对角矩阵, 并且每一个分块都是 Jordan 块, 这就是 **Jordan 分解** (Jordan decompostion).

5.5 作业十六

5.5.1 基础题

习题 5.5.1. 考虑复数域 $\mathbb{C}$ 上的 n 阶方阵组成的线性空间 $V = M_n(\mathbb{C})$, 以及 V 上的线性变换

$$T\colon V \to V, X \mapsto A^T X A.$$

1. 假设 W 是对称矩阵组成的 V 的子空间, U 是反对称矩阵组成的 V 的子空间, 证明 $V = W \oplus U$, 且 W 和 U 都是 T 的不变子空间.

2. 求线性变换 $T|_W\colon W \to W$ 和 $T|_U\colon U \to U$ 的行列式以及迹, 请用 A 的行列式以及迹表示.

习题 5.5.2. 如下归纳地定义方阵

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_n = \begin{bmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{bmatrix}.$$

求 A_n 的特征多项式.

习题 5.5.3. 请对以下 $f, g \in F[t]$ 求以下带余除法, $f = gq + r$.

1. $f = t^3 + 2t^2 + 3t + 4, g = t^2 + t + 1$.

2. $f = t^5 + t^4 + t^3 + t^2 + t + 1, g = t^3 + t^2 + t + 1$.

习题 5.5.4. 假设域 F 上的 n-阶矩阵 A 的特征多项式 $f_A(\lambda)$ 等于其极小多项式, 证明 A 相似于 f_A 的友阵.

习题 5.5.5. 请用直接计算行列式的方法计算出友阵的特征多项式.

习题 5.5.6.

1. 假设 I 是环 R 的理想, 请在商集 R/I 上定义加法和乘法, 使得 R/I 成为一个环.

2. 如果一个 F 线性空间 V 上的线性变换 T 对应的极小多项式是 $m(\lambda)$, 请证明 V 有 $F[\lambda]/(m(\lambda))$-模结构.

3. 假设 n-维实线性空间 V 上有一个线性变换 T 满足 $T^2 = -Id$. 请利用 T 给出 V 的一个复线性空间结构.

4. 在以上的条件下, 假设 A 是 V 上的所有和 T 交换的实线性变换组成的实线性空间. 请求出 $\dim_{\mathbb{R}} A$. *(提示: 构造 A 和 $\mathrm{End}_{\mathbb{C}} V$ 之间的同构.)*

习题 5.5.7. 以下是使用对角化的极小多项式判定法则来求解微分方程的例子. 假设 V 是 $\mathbb{R}$ 上无穷次可导的实值函数组成的线性空间. 线性变换 $D\colon V \to V$ 定义为 $D(f) = f'$. 令 $a_1 \cdots a_n$ 是互不相同的实数, W 是线性变换 $(D-a_1Id)\circ\cdots\circ(D-a_nId)$ 的 *kernel*. 证明

1. W 是 D 的不变子空间.
2. 模仿线性代数中对角化的极小多项式判定法则, 证明 W 是 $\ker(D - a_iId)$ 的直和, 即 $W = \oplus_{i=1}^{n} \ker(D - a_iId)$
3. 已知 $f' = f$ 的解形如 $f(x) = Ce^x$, 其中 C 是任意实数. 求解
$$f'' - 3f' + 2f = 0.$$
4. *(选做)* 请尝试推广以上结论到齐次的常系数线性微分方程的情形. 你需要求解哪些基础的微分方程来得到所有的解.

5.5.2 选做题

习题 5.5.8. 假设 R 是交换环, $m \leq n$ 是正整数. 假设 $A \in M_{m\times n}(R)$ 是一个元素取值在 R 上的 $m \times n$ 矩阵. 假设 I 是 A 的 $m \times m$ 子式生成的 R 的理想. 对任意 $f \in I$, 存在矩阵 $B \in M_{n\times m}(R)$ 使得 $AB = fI_m$. 这里 $m \times m$ 子式指的是 A 中任意取出不重复的 m 行 m 列得到的行列式.

5.6 作业十七

5.6.1 中文版

习题 5.6.1. 设 F 是一个域, $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots a_0\in F[x]$. 类似微积分, 定义 f 的导数为 $f'(x)=na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+a_1$. 对于 $f(x),g(x)\in F[x]$, 证明:

1. $(fg)'=fg'+f'g$.
2. $(f(g(x)))'=f'(g(x))\cdot g'$.
3. $\gcd(f,f')=1$ 当且仅当 f 的不可约分解中没有重因子.

习题 5.6.2. 证明 $\mathbb{Q}/\mathbb{Z}$ 不是有限生成的 $\mathbb{Z}$-模.

习题 5.6.3. 求以下 $\mathbb{Z}$ 上矩阵的 *Smith* 标准形:

$$\begin{bmatrix} 15 & 6 & 9 \\ 6 & 6 & 6 \\ -3 & -12 & -12 \end{bmatrix}.$$

习题 5.6.4.

定义 5.6.1. 设 R 是一个环, $A\in M_{m\times n}(R)$. 取 $i_1,i_2,\cdots,i_k$ 行和 $j_1,j_2,\cdots,j_k$ 列组成的子矩阵的行列式称为 A 的 $k\times k$ 子式. 所有 $k\times k$ 子式的最大公因数称为第 k 个行列式因子 a_k.

1. 类似域 F, 定义 R 上的初等矩阵. 有三类初等矩阵:
   (a) $E_{ij}(\lambda)$: 在单位矩阵上, 将第 i 行的 $\lambda\in R$ 倍加到第 j 行.
   (b) $E_{ii}(\lambda)$: 在单位矩阵上, 将第 i 行乘以 $\lambda\in R^\times$, 其中 $R^\times$ 是 R 中的可逆元集合.
   (c) $E_{ij}(\lambda)$: 在单位矩阵上, 交换第 i 行和第 j 行.

   证明: 若 R 是欧几里得环, 则任意可逆矩阵 $A\in M_n(R)$ 都可以表示为有限个初等矩阵的乘积.
2. 当 R 是欧几里得环时, 证明 A 左右乘可逆矩阵时, (a_k) 不变.
3. 证明行列式因子和不变因子可以互相确定. *(在乘以 R 中乘法可逆元的意义下)*

习题 5.6.5.

1. 假设 $f(x)\in F[x]$ 是域 F 上的不可约多项式, 证明 $F[x]/(f)$ 在商环结构下也是一个域.
2. 请对 $\mathbb{R}$ 上的不可约多项式构造域, 并且证明这些域只能同构于 $\mathbb{R}$ 和 $\mathbb{C}$.

5.6.2 英文版

习题 5.6.6. *Let F be a field and $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots a_0 \in F[x]$. Define the derivative of f similarly as calculus. $f'(x) = na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1$. For $f(x), g(x) \in F[x]$, prove*

1. $(fg)' = fg' + f'g$.
2. $(f(g(x)))' = f'(g(x)) \cdot g'$.
3. $\gcd(f, f') = 1$ *if and only if in the irreducible factorization of f, there are no factors with multiplicities.*

习题 5.6.7. *Prove that $\mathbb{Q}/\mathbb{Z}$ is not a finitely generated $\mathbb{Z}$-module.*

习题 5.6.8. *Find the Smith normal form of the following matrix over $\mathbb{Z}$:*

$$\begin{bmatrix} 15 & 6 & 9 \\ 6 & 6 & 6 \\ -3 & -12 & -12 \end{bmatrix}.$$

习题 5.6.9.

定义 5.6.2. *Let R be a ring and $A \in M_{m\times n}(R)$. The determinant of submatrix with $i_1, i_2, \cdots, i_k$th rows and $j_1, j_2, \cdots, j_k$th columns is called a $k \times k$-minor of A. The greatest common divisor of all $k \times k$-minors is called a determinant divisors a_k.*

1. *Define the elementary matrix over R similarly as field F. There are three types of elementary matrices:*

   (a) *$E_{ij}(\lambda)$: For indentiy matrix, add $\lambda \in R$ times ith row to jth row.*

   (b) *$E_{ii}(\lambda)$: For indentity matrix, multiply ith row by $\lambda \in R^\times$. Here $R^\times$ is the set of mulplicative invertible elements in R.*

   (c) *$E_{ij}(\lambda)$: For indentiy matrix, swap ith row and jth row.*

   Show that if R is a Euclidean domain, then any invertible matrix $A \in M_n(R)$ is the product of a finite number of elementary matrices.

2. *When R is Euclidean Domain, show that a_k does not change when A is multiplied by invertible matrices on the left or right.*

3. *Show that determinant divisors and invariant factors determines each other.*

习题 5.6.10.

1. *Suppose $f(x) \in F[x]$ is an irreducible polynomial over the field F. Prove that $F[x]/(f)$ is also a field under the quotient ring structure.*

2. *Construct fields from irreducible polynomials over $\mathbb{R}$ and prove that these fields can only be isomorphic to $\mathbb{R}$ or $\mathbb{C}$.*

索引