4.1 引言: 回到线性方程组

回首 §1.3 考察的线性方程组

$$\begin{cases} a_{11}X_1 + \cdots + a_{1n}X_n = b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n = b_2 \\ \qquad\qquad\qquad \vdots \qquad \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n = b_m. \end{cases} \tag{4.1.1}$$

当时我们默认系数和所求的解都在 $\mathbb{C}$ 中. 然而解方程所用的 Gauss–Jordan 消元法所需的仅是四则运算, 无关复数的一切其他性质, 所以现在可以假定 a_{ij} 和 b_i 都落在选定的域 F 中, 在 F 上求解. 我们有:

> 在任意域 F 上, 线性方程组 (4.1.1) 可以由 Gauss–Jordan 消元法判定是否有解; 若有解, 则消元法可以描述所有的解.

这是因为 §§1.3—1.4 介绍的行运算和 Gauss–Jordan 消元法理论可以一字不易地推广到任意域上. 在线性方程组的研究中, $b_1 = \ldots = b_m = 0$ 的情形占有特殊的地位.

定义 4.1.1 考虑域 F 上形如 (4.1.1) 的 n 元线性方程组. 如果 $b_1 = \cdots = b_m = 0$, 则称此方程组为**齐次**的.

消元法的基础是矩阵的初等行变换. 以下转换视角, 从映射的观点来理解域 F 上的线性方程组.

给定 $n, m \in \mathbb{Z}_{\geq 1}$ 和一族系数 $(a_{ij})_{\substack{1\leq i\leq m \\ 1\leq j\leq n}}$, 其中 $a_{ij} \in F$, 定义映射

$$\begin{aligned} T : F^n &\longrightarrow F^m \\ (x_j)_{j=1}^n &\longmapsto \left(\sum_{j=1}^n a_{1j}x_j, \ldots, \sum_{j=1}^n a_{mj}x_j \right). \end{aligned}$$

今后也将数组 $(x_1, \ldots, x_n)$ 写作 $\boldsymbol{x}$ 的形式. 对于给定的 $\boldsymbol{b} = (b_1, \ldots, b_m) \in F^m$. 解 (4.1.1) 相当于研究纤维 $T^{-1}(\boldsymbol{b})$ (定义 2.2.10). 对任何映射 $F^n \to F^m$ 都可以考虑同样的问题, 然而眼下问题的特点在于 T 是 "线性" 的. 更明确地说, 在 F^n 上定义

⋆ 两个数组的**加法**运算

$$(x_1, \ldots, x_n) + (x'_1, \ldots, x'_n) := (x_1 + x'_1, \ldots, x_n + x'_n),$$

其中 $(x_1, \ldots, x_n), (x'_1, \ldots, x'_n) \in F^n$;

- ⋆ 域 F 之于数组的**纯量乘法**运算

$$t(x_1,\ldots,x_n):=(tx_1,\ldots,tx_n);$$

其中 $t\in F$.

如此定义的加法满足结合律和交换律, 而纯量乘法对加法满足结合律和分配律:

$$(tt')\boldsymbol{x}=t(t'\boldsymbol{x}),\quad t(\boldsymbol{x}+\boldsymbol{y})=t\boldsymbol{x}+t\boldsymbol{y},\quad (t+t')\boldsymbol{x}=t\boldsymbol{x}+t'\boldsymbol{x};$$

我们不厌其烦地重申: $t,t'\in F$ 而 $\boldsymbol{x},\boldsymbol{y}\in F^n$. 此外, $\mathbf{0}:=(0,\ldots,0)\in F^n$ (必要时记为 $\mathbf{0}_n$) 对加法充当了零元的角色:

$$\boldsymbol{x}+\mathbf{0}=\boldsymbol{x}=\mathbf{0}+\boldsymbol{x}.$$

由于运算是逐分量定义的, 上述性质全都归结为域 F 的相应性质. 另记 $-\boldsymbol{x}:=(-1)\boldsymbol{x}$ 和 $\boldsymbol{x}'-\boldsymbol{x}:=\boldsymbol{x}'+(-\boldsymbol{x})$.

这对解方程的问题有何启发? 请读者验证映射 T 满足

$$T(\underbrace{\boldsymbol{x}+\boldsymbol{y}}_{F^n})=\underbrace{T(\boldsymbol{x})+T(\boldsymbol{y})}_{F^m},\quad T(t\boldsymbol{x})=t\cdot T(\boldsymbol{x}).$$

事实上, $T(\boldsymbol{x})$ 的每个分量 $\sum_{j=1}^n a_{ij}x_j\in F$ 都有这些性质 $(1\le i\le m)$. 一个立即的结论是线性方程组的解可以线性叠加. 何谓线性叠加? 先固定方程组 (4.1.1) 的系数矩阵 $(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$.

- ⋆ 若 $\boldsymbol{x}=(x_1,\ldots,x_n)$ 是方程组对 $\boldsymbol{b}=(b_1,\ldots,b_m)$ 的解, 而 $\boldsymbol{x}'=(x_1',\ldots,x_n')$ 是对 $\boldsymbol{b}'=(b_1',\ldots,b_m')$ 的解, 则 $\boldsymbol{x}+\boldsymbol{x}'$ 是方程组对 $\boldsymbol{b}+\boldsymbol{b}'$ 的解;
- ⋆ 若 $\boldsymbol{x}=(x_1,\ldots,x_n)$ 是 (4.1.1) 对 $\boldsymbol{b}=(b_1,\ldots,b_m)$ 的解, $t\in F$, 则 $t\boldsymbol{x}$ 是方程组对 $t\boldsymbol{b}$ 的解.

因此像集 $\mathrm{im}(T)$ 在 F^m 的加法和纯量乘法运算下封闭. 像集还是非空的, 因为 $\mathbf{0}_m=T(\mathbf{0}_n)\in\mathrm{im}(T)$.

作为推论, 若 $\boldsymbol{x}$ 和 $\boldsymbol{x}'$ 是方程组对同一个 $\boldsymbol{b}\in F^m$ 的解, 则 $\boldsymbol{y}:=\boldsymbol{x}-\boldsymbol{x}'$ 是对应的齐次方程组的解. 反之, 若已知 $\boldsymbol{x}$ 是原方程组的解, 而 $\boldsymbol{y}$ 是对应的齐次方程组的解, 则 $\boldsymbol{x}+\boldsymbol{y}$ 仍是原方程组的解.

综之, 只要找出特解 $\boldsymbol{x}\in T^{-1}(\boldsymbol{b})$, 方程组 (4.1.1) 的通解便由下式描述

$$\begin{aligned}T^{-1}(\boldsymbol{b})&=\boldsymbol{x}+T^{-1}(\mathbf{0})\\&:=\left\{\boldsymbol{x}+\boldsymbol{y}:\boldsymbol{y}\in F^n,\ T(\boldsymbol{y})=\mathbf{0}_m\right\}.\end{aligned}$$

这就引出了以下观察:

- ⋆ 线性方程组 (4.1.1) 有解当且仅当 $\boldsymbol{b} := (b_1, \ldots, b_m) \in \mathrm{im}(T)$;
- ⋆ 一旦它有解, 则精确到用某个特解的 "平移", 解集 $T^{-1}(\boldsymbol{b})$ 的性状完全由对应的齐次方程组确定, 或者说由系数矩阵 $(a_{ij})_{\substack{1\leq i\leq m\\1\leq j\leq n}}$ 确定.

这表明对于解方程而言, F^n 的子集 $T^{-1}(\mathbf{0})$, 亦即对应的齐次方程组的解集和 $\mathrm{im}(T)$ 一样是重要的对象. 一如 $\mathrm{im}(T)$, 子集 $T^{-1}(\mathbf{0})$ 同样对 F^n 的加法和纯量乘法运算封闭, 这是因为

$$\begin{aligned} &T(\boldsymbol{x}) = \mathbf{0}_m = T(\boldsymbol{x}') \implies \\ &\qquad\qquad T(\boldsymbol{x}+\boldsymbol{x}') = \mathbf{0}_m + \mathbf{0}_m = \mathbf{0}_m, \\ &t \in F, \quad T(\boldsymbol{x}) = \mathbf{0}_m \implies \\ &\qquad\qquad\qquad T(t\boldsymbol{x}) = tT(\boldsymbol{x}) = t\,\mathbf{0}_m = \mathbf{0}_m. \end{aligned}$$

对 $T^{-1}(\mathbf{0}_m)$ 可以提许多问题, 比如说, 解集的参数化所指究竟为何? 如何确定所需的参数个数, 即先前所谓的 "自由度"?

定义 4.1.2 设 $T : F^n \to F^m$ 按上述方式对应到一个齐次线性方程组. 若 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h \in F^n$ 都是方程组的解, 而且所有解 $\boldsymbol{x} \in F^n$ 都可以通过加法和纯量乘法表作

$$\boldsymbol{x} = \sum_{i=1}^{h} t_i \boldsymbol{v}_i, \quad t_1, \ldots, t_h \in F,$$

其中 $(t_1, \ldots, t_h)$ 由 $\boldsymbol{x}$ 唯一确定, 则称 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h$ 是该齐次方程组的一组**基础解系**.

求出基础解系 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_h$ 相当于用 h 个参数来描述齐次线性方程组的解集.

命题 4.1.3 考虑形如 (4.1.1) 的 n 元齐次线性方程组, 其中 $\boldsymbol{b} = 0$. 设消元法给出的简化行梯矩阵有 r 个主元, 则对应的齐次方程组有基础解系 $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-r}$.

证明 沿用 §1.4 的记号. 将简化行梯矩阵中不含主元的列按编号枚举为

$$1 \leq f_1 < \cdots < f_{n-r} \leq n.$$

对公式 (1.4.1) 代入 $b_k = 0$ 给出齐次线性方程组的通解, 其中变元 $X_{f_1}, \ldots, X_{f_{n-r}}$ 可以任意赋值, 其余对应到主元的变元则由这些值唯一确定. 因此对每个 $1 \leq i \leq n-r$ 都可以取 $\boldsymbol{v}_i \in F^n$ 为齐次线性方程组的解, 使得

- ⋆ 它的第 f_i 个分量为 1,
- ⋆ 当 $1 \leq j \leq n-r$ 而 $j \neq i$ 时它的第 f_j 个分量为 0,
- ⋆ 其余分量由上述资料和 (1.4.1) 唯一确定.

若 $\boldsymbol{x} \in F^n$ 是齐次方程组的任意解, 则上述讨论表明 $\boldsymbol{x} = \sum_{i=1}^{n-r} t_i \boldsymbol{v}_i$, 其中的 $t_i \in F$ 是且仅能是 $\boldsymbol{x}$ 的第 f_i 个分量, 因此它们由 $\boldsymbol{x}$ 唯一确定. □

进一步, 我们还可以问解集 $T^{-1}(\mathbf{0})$ 的参数个数 $n-r$ 能否由 T 内在地刻画. 基于几何直观, 答案理应是肯定的. 我们现在有充分的动机来研究:

- ⋆ 带有加法和来自域 F 的纯量乘法这两种运算的代数结构, 例如 F^n, 或者先前提到的 $\mathrm{im}(T)$ 和 $T^{-1}(\mathbf{0})$;
- ⋆ 保持这种代数结构的映射, 例如之前考虑的 $T: F^n \to F^m$.

这就将向量空间的抽象理论推到了聚光灯下. 稍后展开的理论将会表明, 此一观点和 Gauss–Jordan 消元法是相互交融的.

4.2 向量空间

域 F 上的向量空间又称为线性空间. 这是具有加法, 纯量乘法[1)]运算以及零元的一种代数结构; 加法是空间上的二元运算, 纯量乘法则涉及域 F 的元素在 V 上的作用.

定义 4.2.1 域 F 上的 **F-向量空间**简称向量空间, 这是指资料 $(V, +, \cdot, 0_V)$, 其中 V 是集合, $0_V \in V$, 而 $+: V \times V \to V$ 和 $\cdot: F \times V \to V$ 分别写作 $(u, v) \mapsto u + v$ 和 $(t, v) \mapsto t \cdot v$ 的形式, 使得以下条件成立.

(i) 加法满足以下条件:

- ▷ **结合律** $(u+v)+w = u+(v+w)$;
- ▷ **幺元性质** $v + 0_V = v = 0_V + v$;
- ▷ **交换律** $u + v = v + u$;
- ▷ **加法逆元** 对所有 v 皆存在 $-v$ 使得 $v + (-v) = 0_V$.

(ii) 纯量乘法 $t \cdot v$ 也简写为 tv, 它满足以下条件:

- ▷ **结合律** $s \cdot (t \cdot v) = (st) \cdot v$;
- ▷ **幺元性质** $1 \cdot v = v$, 此处 1 代表 F 的乘法幺元.

(iii) 纯量乘法对加法满足:

- ▷ **分配律之一** $(s+t) \cdot v = sv + tv$;
- ▷ **分配律之二** $s \cdot (u+v) = su + sv$.

其中 u, v, w (或 s, t) 代表 V (或 F) 中的任意元素. 不致混淆时, 我们也简记 0_V 为 0, 将 $u + (-v)$ 写作 $u - v$, 并以 V 总括资料 $(V, +, \cdot, 0)$.

向量空间 V 的元素也称为其中的**向量**. 定义中的 0_V 又称为 V 的零元, 或**零向量**. 与向量相对, 域 F 的元素则称为**纯量**.

定义的模式和环的定义 3.1.1 明显相似. 由之推出的以下几条性质也是出于同样理路.

[1)] 在一些教材中, 纯量乘法又称为数乘或标量乘法.

- ⋆ 零元 0_V 由相应的幺元性质唯一确定, 所以资料中的 0_V 其实可以略去, 要求存在满足该性质的元素即可.
- ⋆ 加法满足消去律: $u+v=u'+v$ 蕴涵 $u=u'$. 作为推论, 向量 v 的加法逆元 $-v$ 唯一.
- ⋆ 逆元的唯一性也蕴涵 $-(-v)=v$.
- ⋆ 对于任何 $t\in F$ 都有 $t\cdot 0_V=0_V$, 这是对 $t\cdot(0_V)=t\cdot(0_V+0_V)=t\cdot 0_V+t\cdot 0_V$ 应用加法消去律的结论.
- ⋆ 我们有恒等式 $0\cdot v=0$; 等号左边的 $0\in F$, 而右边的 $0=0_V\in V$. 这是缘于 $0\cdot v=(0+0)\cdot v=0\cdot v+0\cdot v$ 和消去律.
- ⋆ 我们有 $(-1)\cdot v=-v$. 这是缘于 $(-1)\cdot v+v=(-1)\cdot v+1\cdot v=(-1+1)\cdot v=0\cdot v=0$.

定义 4.2.2 设 V 为 F-向量空间. 如果 V 的子集 V_0 包含 0, 而且在加法和纯量乘法运算下封闭, 则 $(V_0,+,\cdot,0)$ 也是 F-向量空间, 称之为 V 的**子空间**.

子空间的定义中未要求 V_0 对取逆 $v\mapsto -v$ 封闭, 这是因为 $-v=(-1)\cdot v$, 所以要求纯量乘法的封闭性已经足够.

练习 4.2.3 给定 F-向量空间 V, 请验证任意多个子空间的交仍然是 V 的子空间.

例 4.2.4 (零空间) 最平凡的向量空间是 $\{0\}$, 它是所有 F-向量空间的子空间.

向量空间理论也有直观的几何实例, 关乎中学数学里熟悉的平面与空间向量.

例 4.2.5 (向量几何) 平面向量对加法和纯量乘法构成 $\mathbb{R}$-向量空间, 空间向量亦同. 这是向量空间最直观的例子. 在未取定坐标的情况下, 向量的加法是按照平行四边形法则确定的, 而向量的纯量乘法则是带方向的伸缩, 负号代表反向; 相关的几何图像在本章开头已有回顾. 为了用代数语言表达角度和长度等几何概念, 还需要引入内积的概念, 这将是第九章的主题.

例 4.2.6 (域本身作为向量空间) 域 F 相对于域的加法和乘法 (在此扮演纯量乘法的角色) 构成 F-向量空间, 这是最简单的非零 F-向量空间, 它在后续章节还会反复现身.

例 4.2.7 设 $n\in\mathbb{Z}_{\geq 0}$. 按以下方式赋予 F^n 向量空间的结构

$$\begin{aligned}(x_1,\ldots,x_n)+(y_1,\ldots,y_n) &:= (x_1+y_1,\ldots,x_n+y_n), \quad x_i,y_i\in F;\\ t(x_1,\ldots,x_n) &:= (tx_1,\ldots,tx_n), \quad t\in F.\end{aligned}$$

其中的零元取为 $\mathbf{0}:=(0,\ldots,0)$. 这是 §4.1 讨论的例子. 定义所需的性质都是明显的. 注意到 F^0 规定为零空间, 而 F^1 无非是例 4.2.6 介绍的空间 F.

如果在 F^n 的定义中具体取 $F=\mathbb{R}$ 和 $n=2$ (或 $n=3$), 则 $\mathbb{R}^2$ (或 $\mathbb{R}^3$) 可以等同于坐标化的平面 (或空间). 中学数学所出现的平面 (或空间) 向量带有始点和终点, 写作诸如 $\overrightarrow{PQ}$ 的形式. 一旦在平面 (或空间) 中选定一点作为原点, 则总是可以将所有向量的始点平移到原点, 从而化约到例 4.2.5 的情境. 若进一步建立坐标系[2], 便化约到此处讨论的 $\mathbb{R}^2$ 和 $\mathbb{R}^3$.

例 4.2.8 域 F 上的多项式环 $F[X]$ 相对于纯量乘法

$$t\cdot\sum_{n\geq 0}c_nX^n := \sum_{n\geq 0}tc_nX^n, \quad t\in F$$

和 $F[X]$ 本身的加法构成 F-向量空间. 推而广之, 任意个变元的多项式环 $F[X,Y,\ldots]$ 也都是 F-向量空间.

以下介绍两种从已有的向量空间构造新空间的抽象途径.

例 4.2.9 (向量空间的直积) 首先回忆 (2.3.1) 的定义: 对任意一族集合 V_i, 其中下标 i 遍历给定的集合 I (暂时要求非空), 积集 $\prod_{i\in I}V_i$ 的元素是以 I 为下标的元素组 $(v_i)_{i\in I}$, 其中 $v_i\in V_i$. 如果每个 V_i 都带有 F-向量空间的结构, 则 $\prod_{i\in I}V_i$ 也有自然的 F-向量空间结构, 方法是逐分量地定义运算为

$$(v_i)_i+(w_i)_i=(v_i+w_i)_i, \quad t(v_i)_i=(tv_i)_i,$$

下标 i 遍历 I 而 $t\in F$. 我们称此向量空间 $\prod_{i\in I}V_i$ 为 $(V_i)_{i\in I}$ 的直积. 所需的结合律等诸般性质全部都能化到每个 V_i 上来验证. 思路和环的直积 (例 3.1.13) 并无二致.

如果取所有 V_i 都是同一个向量空间 V, 则和积集的情形类似, 我们将对应的 $\prod_{i\in I}V$ 记为 V^I. 回忆到在关于集合论的讨论中, 我们在 $I=\varnothing$ 时不无理由地规定 V^I 为独点集. 建立在独点集上的向量空间结构能且仅能是零空间, 因此我们规定当 $I=\varnothing$ 时 $V^I:=\{0\}$. 这一切可以视为例 4.2.7 的推广: 不难看出 F^n 即是 $F^{\{0,\ldots,n-1\}}$.

例 4.2.10 (向量空间的直和) 例 4.2.9 介绍的直积 $\prod_{i\in I}V_i$ 有一个更常用的子空间, 称为 $(V_i)_{i\in I}$ 的**直和**, 定义为

$$\bigoplus_{i\in I}V_i := \left\{(v_i)_i\in\prod_{i\in I}V_i : \text{至多仅有有限个 } i\in I \text{ 使得 } v_i\neq 0\right\}.$$

留意到如果 $(v_i)_{i\in I}$ 和 $(w_i)_{i\in I}$ 都仅有至多有限个分量非零, 则 $(v_i+w_i)_{i\in I}$ 亦然, 因为两个有限集的并仍有限. 因此 $\bigoplus_{i\in I}V_i$ 对加法封闭. 此外它显然也对纯量乘法封闭, 并且包含 $0=(0)_{i\in I}$. 于是 $\bigoplus_{i\in I}V_i$ 确实是 $\prod_{i\in I}V_i$ 的子空间.

每个 V_i 都自然地嵌入为 $\bigoplus_{j\in I}V_j$ 作为子空间, 方式是映 $v\in V_i$ 为 $(v_j)_{j\in I}$, 其中 $v_i:=v$ 而 $j\neq i$ 时 $v_j:=0$. 考察各个分量, 立见在这些嵌入之下有

$$i\neq j \implies V_i\cap V_j=\{0\}.$$

[2] 几何直观中的平面或空间既无指定的原点, 又无指定的坐标轴, 所以 $\mathbb{R}^n$ 对此是个不尽合身的模型, 以后讨论仿射空间时还会回到这个问题.

如果取所有 V_i 都是同一个向量空间 V, 则对应的直和表作 $V^{\oplus I}$, 以区别于直积 V^I. 当 I 有限时, 至多有限个分量非零的条件自动成立, 这时 $\prod_{i\in I} V_i = \bigoplus_{i\in I} V_i$. 本书涉及的直和的以这类情形居多.

约定 4.2.11 规定空直和与空直积 (相当于 $I = \varnothing$ 情形) 为零空间 $\{0\}$. 这与 §2.3 的规定兼容.

直和是向量空间理论中的一则基本构造. 关于直和的进一步讨论将在 §4.10 接续.

4.3 矩阵及其运算

矩阵在 §1.3 关于消元法的讨论中已经出现. 当时我们对矩阵元在哪个数系取值含糊其词. 现在既然有了域的概念, 便可以清楚地定义一般的矩阵.

定义 4.3.1 (域上的矩阵) 设 $m, n \in \mathbb{Z}_{\geq 1}$. 域 F 上的 $m \times n$ **矩阵**是指如下的资料

$$\begin{aligned} \boldsymbol{A} &= (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} \\ &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{array}{cc} \begin{pmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{pmatrix} & \text{第 } i \text{ 行} \\ \text{第 } j \text{ 列} & \end{array} \end{aligned}$$

其中 $a_{ij} \in F$, 称为矩阵 $\boldsymbol{A}$ 的第 (i,j) 个矩阵元或 (i,j)-项, 而 $n \times n$ 矩阵也称为 n 阶**方阵**.

今后我们将 F 上的所有 $m \times n$ 矩阵 (m 行 n 列) 构成的集合记为 $\mathrm{M}_{m\times n}(F)$.

按惯例, 我们经常将矩阵 $\boldsymbol{A} = (a_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$ 简记为 $(a_{ij})_{i,j}$. 严格地说, F 上的 $m \times n$ 矩阵是集合 $F^{\{1,\ldots,m\}\times\{1,\ldots,n\}}$ 的元素; 这种原教旨主义的解读当然是庸人自扰.

现在赋予集合 $\mathrm{M}_{m\times n}(F)$ 作为向量空间所需的运算.

▷ **加法** 对 $\mathrm{M}_{m\times n}(F)$ 的任两个元素 $\boldsymbol{A} = (a_{ij})_{i,j}$ 和 $\boldsymbol{B} = (b_{ij})_{i,j}$, 定义

$$\boldsymbol{A} + \boldsymbol{B} := (a_{ij} + b_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}.$$

▷ **纯量乘法** 对任意 $t \in F$ 和 $\boldsymbol{A} = (a_{ij})_{i,j} \in \mathrm{M}_{m\times n}(F)$, 定义

$$t \cdot \boldsymbol{A} = t\boldsymbol{A} := (ta_{ij})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}.$$

给定 m 和 n, 对应尺寸的**零矩阵**定义为

$$\mathbf{0}_{m\times n} := (0)_{i,j} \in \mathrm{M}_{m\times n}(F).$$

命题 4.3.2 这些运算使得 $\mathrm{M}_{m\times n}(F)$ 成为 F-向量空间. 其中的零元为零矩阵, 而矩阵 $\boldsymbol{A}=(a_{ij})_{i,j}$ 的加法逆元 $-\boldsymbol{A}$ 为 $(-a_{ij})_{i,j}$.

证明 理应是明显的. 一切运算性质都对每个矩阵元化约到 F 中来验证. □

定义 $m\times n$ 矩阵的减法为 $\boldsymbol{A}-\boldsymbol{B}:=\boldsymbol{A}+(-\boldsymbol{B})$, 它的 (i,j) 矩阵元是 $a_{ij}-b_{ij}$.

约定 4.3.3 (行向量与列向量) 对于特例 $n=1$, 向量空间 $\mathrm{M}_{m\times 1}(F)$ 化约为例 4.2.7 讨论的 F^m; 出于直观的理由, $\mathrm{M}_{m\times 1}(F)$ 的元素也称为 m 维的**列向量**. 同样道理, 向量空间 $\mathrm{M}_{1\times n}(F)$ 可以等同于 F^n, 其元素称为 n 维的**行向量**. 当 $n=m=1$, 空间 $\mathrm{M}_{1\times 1}(F)$ 无非就是 F 自身.

作为向量空间, $\mathrm{M}_{m\times n}(F)$ 和 F^{mn} 似乎没有实质区别, 然而矩阵的特色在于除了加法和纯量乘法之外, 它们还具备乘法运算.

定义 4.3.4 (矩阵乘法) 矩阵乘法是按以下方式定义的映射

$$\begin{aligned}\mathrm{M}_{m\times n}(F)\times\mathrm{M}_{n\times r}(F) &\longrightarrow \mathrm{M}_{m\times r}(F)\\ (\boldsymbol{A},\boldsymbol{B}) &\longmapsto \boldsymbol{AB};\end{aligned}$$

若 $\boldsymbol{A}=(a_{ij})_{\substack{1\le i\le m\\ 1\le j\le n}}$, $\boldsymbol{B}=(b_{jk})_{\substack{1\le j\le n\\ 1\le k\le r}}$, 则 $\boldsymbol{AB}=(c_{ik})_{\substack{1\le i\le m\\ 1\le k\le r}}$, 其中

$$c_{ik}:=\sum_{j=1}^{n}a_{ij}b_{jk}=\underbrace{\begin{pmatrix} a_{i1} & \cdots & a_{in}\end{pmatrix}}_{\text{第 } i \text{ 行}}\underbrace{\begin{pmatrix} b_{1k}\\ \vdots \\ b_{nk}\end{pmatrix}}_{\text{第 } k \text{ 列}}.$$

注意: 只有行数和列数合乎规格的矩阵才能相乘. 按惯例, $\boldsymbol{AB}$ 有时也写成 $\boldsymbol{A}\cdot\boldsymbol{B}$.

定义 n 阶**单位矩阵**为

$$\mathbf{1}_{n\times n}:=\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix}\in\mathrm{M}_{n\times n}(F)$$

其中对角线矩阵元全为 1, 其余留白部分全为 0. 请读者循定义动手验证单位矩阵对所有 $\boldsymbol{A}\in\mathrm{M}_{m\times n}(F)$ 都有类似于乘法幺元的性质:

$$\mathbf{1}_{m\times m}\cdot\boldsymbol{A}=\boldsymbol{A}=\boldsymbol{A}\cdot 1_{n\times n}.$$

命题 4.3.5 矩阵乘法满足以下性质:

▷ 结合律 $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$;

▷ 分配律 $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC}$, $(\boldsymbol{B}+\boldsymbol{C})A = \boldsymbol{BA}+\boldsymbol{CA}$;

▷ 线性 $\boldsymbol{A}(t\boldsymbol{B}) = t(\boldsymbol{AB}) = (t\boldsymbol{A})\boldsymbol{B}$;

其中 $t \in F$ 和矩阵 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 是任意的, 前提是矩阵的行数和列数使运算有意义.

证明 对 $(\boldsymbol{AB})\boldsymbol{C}$ 和 $\boldsymbol{A}(\boldsymbol{BC})$ 的每个矩阵元进行考察, 可见结合律相当于说

$$\sum_k \left(\sum_j a_{ij} b_{jk} \right) c_{kl} = \sum_j a_{ij} \left(\sum_k b_{jk} c_{kl} \right),$$

下标 i, j, k, l 的范围取决于矩阵尺寸. 根据求和符号的基本操作, 交换 $\sum_k$ 和 $\sum_j$ 以后, 问题归结为证

$$(a_{ij} b_{jk}) c_{kl} = a_{ij} (b_{jk} c_{kl}).$$

然而这无非是 F 本身的乘法结合律. 同理, 分配律归结为 F 本身的分配律

$$a_{ij}(b_{jk} + c_{jk}) = a_{ij} b_{jk} + a_{ij} c_{jk}.$$

最后一则性质归结为 F 中的等式 $a_{ij}(tb_{jk}) = t(a_{ij}b_{jk}) = (ta_{ij})b_{jk}$, 而这又归结为 F 的乘法交换律. □

尽管推论 4.6.3 将包含下述结果, 眼下先就矩阵定义来验证是有益的, 细节也毫无困难.

练习 4.3.6 验证 $\mathrm{M}_{n\times n}(F)$ 相对于矩阵加法和乘法成环, 它以零矩阵 $\boldsymbol{0}_{n\times n}$ 为零元, 以单位矩阵 $\boldsymbol{1}_{n\times n}$ 为乘法幺元. 提示 前述诸性质的综合.

练习 4.3.7 矩阵给出非交换环的自然例子. 请动手验证以下等式以说明 $\mathrm{M}_{2\times 2}(F)$ 并非交换环.

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

注记 4.3.8 (环上的矩阵) 如果将域 F 换成一般的环 R, 则以上定义和大部分性质仍然通行无阻; 特别地, 我们有集合 $\mathrm{M}_{m\times n}(R)$ 以及矩阵之间的加法和乘法, 并且使 $\mathrm{M}_{n\times n}(R)$ 成环. 唯一成问题的是命题 4.3.5 的等式 $\boldsymbol{A}(t\boldsymbol{B}) = t(\boldsymbol{AB})$ (其中 $t \in R$): 从证明不难看出, 障碍在于 $t \in R$ 对乘法未必能和 R 中的其他元素交换. 如果我们要求 R 是交换环, 则 $\boldsymbol{A}(t\boldsymbol{B}) = t(\boldsymbol{AB}) = (t\boldsymbol{A})\boldsymbol{B}$ 仍然成立.

此外, 在非交换环上必须区别 $t \in R$ 对矩阵的左乘 $t\boldsymbol{A}$ 和右乘 $\boldsymbol{A}t$.

一如 §1.3 所述, 线性方程组的布列和消元在一般的域 F 上也能用 F 上的矩阵来表述. 矩阵乘法还给出考量线性方程组的另一种视角: 对于取在 F 中的变元 $X_1, \ldots, X_n$, 我们有

$$\begin{cases} a_{11}X_1 + \cdots + a_{1n}X_n &= b_1 \\ a_{21}X_1 + \cdots + a_{2n}X_n &= b_2 \\ \qquad\qquad\qquad\vdots & \quad\vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n &= b_m \end{cases} \iff \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

取方程组的系数矩阵 $\boldsymbol{A} := (a_{ij})_{i,j} \in \mathrm{M}_{m\times n}(F)$, 并且定义列向量

$$\boldsymbol{x} := \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}, \quad \boldsymbol{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix},$$

则原方程组等价于以 $\boldsymbol{x}$ 为变量的矩阵方程

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

这是矩阵乘法和线性方程组的直接联系, 矩阵写法显然胜在简洁, 而这也为矩阵乘法的定义提供了一部分的解释.

我们更愿意从线性映射的视角理解矩阵. 一旦在 §4.5 明确了矩阵和线性映射的关系, 则无论是命题 4.3.5 的矩阵乘法性质, 练习 4.3.6 或线性方程组的矩阵写法 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ 都会得到自然的诠释. 在此之前, 有必要先引进基和维数的抽象概念.

4.4 基和维数

对于经典的平面向量或空间向量, 我们熟知一旦选定坐标系, 所有向量都能按坐标唯一地展开. 将此思路扩及一般的向量空间, 就引向了基的概念.

向量空间的基并不唯一, 通常也没有标准的取法. 在许多应用中[3)], 适当选基, 自由换基还是解决问题的关键手段. 阐明基与基之间如何变换是稍后的重要任务.

严谨的解释需要一系列的准备工作. 设 S 为 F-向量空间 V 的子集, 或有限或无穷. 我们称形如

$$a_1 v_1 + \cdots + a_m v_m, \qquad \begin{aligned} &m \in \mathbb{Z}_{\geq 0}, \\ &a_i \in F, \\ &v_i \in S \end{aligned} \tag{4.4.1}$$

3) 例如中学物理力学中的静态平衡.