

代数笔记

Algebra

© 晨沐公[†]

[†] 成都市锦江区嘉祥外国语高级中学

临时笔记

bilibili: 晨沐公 Kasumi github: MATHhahetaDEATH

2024 年 8 月 29 日

这一切只是虚张声势的线性代数.

WENWEI LI

前言

在数学中, 对于每一个数学对象 (例如极限), 我们会例行公事般地考虑它的一些常见的性质. 比如说, 这个对象最基本的例子是什么, 这种对象是否存在, 如果存在的话它是否具有唯一性, 它的子对象和商对象 (如果有的话) 都具有什么性质 (比如说遗传了原来的对象的什么性质), 这个对象的可计算性以及它在特定映射下的行为等等.

这份讲义 cover 的内容:

- “抽象废话” 范畴论, 当然只有一点点
- 作为工具和灵感来源的矩阵理论
- 几乎同构于矩阵理论的线性映射理论, 但我们会研究一些无限维情况
- 经典的群结构 $\mathbb{Z}/n\mathbb{Z}$, 以及一般的群论
- 经典的环结构 $\mathbb{F}[x]$, 以及初步的环论
- 域理论
- 初步的模论

换句话说, 本讲义大致覆盖高等代数与基础抽象代数的内容.

抽象的代数需要直觉, 例如“自由度”“置换”等想法. 用具体详实的例子引出抽象理论.

很多想法都是读者所熟悉的, 例如数论中的中国剩余定理, 多项式理论中的因式分解等. 因此这份讲义将不会侧重解释这些初等理论, 而是试图将其推广.

目录

I	预备知识	1
1	集合论回顾	2
1.1	公理化集合论	2
1.2	映射	4
1.3	二元关系	6
1.4	序数	8
1.4.1	全序集与序数	8
1.4.2	归纳法与选择公理	10
1.5	基数	12
2	范畴论初步	15
II	线性映射与代数结构	16
3	向量空间	17
3.1	从 \mathbb{F}^n 说起	17
3.1.1	复数与复数域	17
3.1.2	\mathbb{F}^n	19
3.2	向量空间	22
3.3	子空间	25
3.3.1	子空间	25
3.3.2	子空间的和	26
3.3.3	直和	27
3.4	有限维向量空间	28
3.4.1	张成空间	28
3.4.2	有限维向量空间	31
3.5	线性无关	32
3.5.1	线性无关	32
3.5.2	线性相关性与张成	35

目录	iv
3.6 基与维数	38
3.6.1 基	38
3.6.2 维数	39
4 线性映射与矩阵	42
4.1 基本概念	42
4.1.1 向量空间的线性映射	42
4.1.2 线性映射的表示矩阵	46
4.1.3 矩阵的运算	48
4.2 零空间与值域	51
4.2.1 解线性方程组问题	51
4.2.2 可逆矩阵	52
4.2.3 零空间与值域	53
4.2.4 线性映射基本定理	55
4.3 可逆性与同构的向量空间	58
4.3.1 线性映射的可逆性	58
4.3.2 同构的向量空间	59
4.3.3 基的变换	61
4.4 向量空间的积与商	63
4.4.1 向量空间的积	63
4.4.2 向量空间的商	63
4.5 对偶与矩阵的秩	66
4.5.1 对偶	66
4.5.2 矩阵的秩	67
4.5.3 对偶映射的零空间和值域	68
5 群论	71
5.1 基本想法	71
6 多项式与环	74
III DLC1-范畴论	75
IV DLC2-线性代数	76
7 线性方程组与行列式计算	77

8 对角化	78
8.1 基本概念	79
8.2 极小多项式	81
8.2.1 本征值与极小多项式	81
8.2.2 奇数维实向量空间上的本征值	83
8.3 上三角矩阵	84
8.4 可对角化算子	86

Part I

预备知识

Chapter 1

集合论回顾

任何人都不能把我们从 Cantor 创造的乐园里赶出去.

Hilbert

1.1 公理化集合论

我们选择采用 ZFC 公理系统, 即承认以下九条关于集合论的公理 (详细解释可在任何一本集合论的书上看到, 这里略去):

1. **外延公理**: 两个集合 A 和 B 相等当且仅当它们的元素相同.

外延公理告诉我们, 描述同一群对象的任意集合都是相等的. 因此, 从等价类的角度来看, 它是唯一确定的.

2. **配对公理**: 对于任意集合 X, Y , 存在一个集合 Z 使得 X 和 Y 是它仅有的元素. 特别地, 若 $X = Y$, 则将 Z 视作只有唯一元素.

由配对公理, 存在集合 $\{X, Y\}$ 和 $\{Y, X\}$, 而由外延公理这两个集合是相等的, 于是集合是无序的¹. 另一方面, 容易说明集合 $\{X, X\}$ 就等于 $\{X\}$, 于是集合是不重复的.

3. **并集公理**: 对于一个集合族 M (即元素均为集合的集合²), 存在另一个集合 $\bigcup M$, 其元素恰包含所有属于 M 的集合的元素. 这样的集合称作 M 的并.

4. **分离公理**: 任意集合 A ³ 和性质 P 都对应另一个集合 B , 其元素恰包含那些在集合 A 中而具有性质 P 的.

这实际上是在说, $B = \{x \in A : P(x)\}$ 也是一个集合.

¹如果想要描述有序的集合 (列表), 例如 (X, Y) , 可采用 $\{\{X, Y\}, \{X\}\}$ 的形式, 之后通过嵌套可构造出更长的列表.

²前提是一些集合的确构成集合, 反例如所有集合无法构成一个更大的集合. 后文会针对性地讨论这种操作.

³同注释 2, 此处必须要求 A 是集合才能进行分离操作, 否则会导致 Russell 悖论.

结合并集公理, 马上可以定义非空⁴集合族 M 的交为:

$$\bigcap M := \{x \in \bigcup M : \forall X, X \in M \Rightarrow x \in X\}.$$

特别地, 若 $M = \{A, B\}$, 则 $\bigcap M$ 记作 $A \cap B$.

同理可以定义集合的差和补:

$$A - B := \{x \in A : x \notin B\}.$$

如果 A 是 M 的一个子集, 则定义:

$$A^c := M - A.$$

另外, 分离公理也表明, 对任意集合 X 都存在一个不包含任何元素的子集 $\emptyset_X := \{x \in X : x \neq x\}$. 由外延公理可知对任意集合 X, Y 都有 $\emptyset_X = \emptyset_Y$. 我们称该集合为空集, 记为 \emptyset .

5. **幂集公理**: 对任意集合 X , 总存在它的幂集 $\mathcal{P}(X)$, 其元素恰为 X 的所有子集.

应用幂集公理的一个例子是 Cartesian 积构造之合理性.

6. **无穷公理**: 存在包含空集和自身任何一个元素的后继的集合 (这样的集合称作归纳集). 这里, 所谓后继是指 $X^+ := X \sqcup \{X\}$.

前文已经表明, 存在唯一的空集 \emptyset . 无穷公理允许我们构造任何归纳集的一个归纳子集:

$$\{\emptyset, \emptyset^+, (\emptyset^+)^+, \dots\}.$$

同样可以验证这是唯一的. 这实际上就是可数无穷序数 ω .

7. **替换公理**: 令 $\mathcal{F}(x, y)$ 是如下命题: 对于 X 中的任意元素 x_0 , 存在唯一的 y_0 使得 $\mathcal{F}(x_0, y_0)$ 成立. 那么满足以下条件的 y 构成一个集合: 存在 $x \in X$ 使得 $\mathcal{F}(x, y)$ 成立.

我们选择用一阶逻辑来阐述替换公理而不是直接规定“映射的值域是集合”, 在于这种语言更加精确.

替换公理的一个常见应用是: 构造集合 A 和映射 f , 则 $f(A)$ 必为集合, 由此得到一些矛盾.

8. **正则公理**: 任何非空集合 X 都存在一个元素 x , 使得 $x \cap X = \emptyset$. 等价地说, X 包含一个相对 \in 关系的极小元.

正则公理的直接推论是, 不存在 $X \in X$ 的情形.

9. **选择公理**: 对于任何由互不相交且非空的集合形成的集合族 X , 存在选择函数 $f: X \rightarrow \bigcup X$ 使得对任意 $x \in X$ 都有 $f(x) \in x$.

命题 1.1 集合运算的运算律

设集合 A, B, C , 集合族 $\{B_\alpha : \alpha \in I\}$ (这里 I 是指标集).

- 交、并满足交换律, 即

$$A \cap B = B \cap A, \quad A \cup B = B \cup A.$$

⁴从定义不难看出, 当 M 为空集时 $\bigcap M$ 只能是全体集合构成的类.

- 交、并满足结合律, 即

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

$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C).$$

- 交对并、并对交满足分配律, 即

$$A \cap \left(\bigcup_{\alpha \in I} B_\alpha \right) = \bigcup_{\alpha \in I} (A \cap B_\alpha),$$

$$A \cup \left(\bigcap_{\alpha \in I} B_\alpha \right) = \bigcap_{\alpha \in I} (A \cup B_\alpha).$$

- de Morgan 定律:

$$\left(\bigcup_{\alpha \in I} B_\alpha \right)^c = \bigcap_{\alpha \in I} B_\alpha^c, \quad \left(\bigcap_{\alpha \in I} B_\alpha \right)^c = \bigcup_{\alpha \in I} B_\alpha^c.$$

证明 略.

最后对所谓类的概念作出说明: 在 ZFC 系统中, 凡是非集合的对象都称之为类, 因而其只能作为一个指代而不具有任何特殊性质. 一个例子是所有集合构成的类 \mathbf{V} . 利用分离公理, 可以证明这的确无法构成集合:

证明 下面证明对任意集合 X 均存在集合 $R_X \notin X$. 令 $P(x) = (x \notin x)$, 则对于集合 X 存在集合 $R_X = \{x \in X : P(x)\}$. 此时 $R_X \notin R_X$, 否则 $R_X \in R_X$ 将导致 $R_X \in X$ 和 $R_X \notin R_X$, 得到矛盾. 从而, 由 $R_X \notin R_X$ 可得 $R_X \notin X$.

1.2 映射

作为对应关系存在的映射强调的是元素间的相互作用, 而这一节我们会着重于研究映射之间的关系. 最为明显的例子是映射的复合运算, 这一运算使得 (B^A, \circ) 构成一个群.

定义 1.2 映射

- 设 A 和 B 为两个集合, 若对 A 中每个元素 x , 都存在 B 中唯一的元素 y 与之对应, 则称此对应关系为一个映射, 记作

$$f : A \rightarrow B, \quad x \mapsto y.$$

- x 在 B 中的对应元素 y 称为 x 在 f 下的象, x 称为 y 在 f 下的原象, 记作

$$f(x) = y, \quad x \in A.$$

- 集合 A 称作映射 f 的定义域; 集合 B 称为映射 f 的陪域; A 中所有元素在 f 下的象组成的集合称为 f 的值域, 记作 $f(A)$ 或 $\text{im } f$.
- 两个映射相等, 当且仅当它们的定义域、对应关系、陪域相同.

从集合论的视角看, 一个映射其实就是确定的三元组 (A, B, f) , 其中 A 是定义域, B 是陪域, f 是对应关系.

定义 1.3 双射

设映射 $f: A \rightarrow B$.

- 称 f 是单射, 若 $\forall x, y \in A, f(x) = f(y) \Rightarrow x = y$. 记 $f: A \hookrightarrow B$.
- 称 f 是满射, 若 $\forall y \in B, \exists x \in A (f(x) = y)$. 记 $f: A \twoheadrightarrow B$.
- 称 f 是双射, 若 f 既是单射又是满射.

定义 1.4 映射间的运算

设映射 $f: A \rightarrow B, g: B \rightarrow C$.

- 定义 f, g 的复合映射

$$gf: A \rightarrow C, x \mapsto g(f(x)).$$

注意复合运算有先后顺序. 另外, 为了强调复合运算, gf 也可记作 $g \circ f$.

- 称 f 是 A 上的一个恒等映射, 如果对任意 $x \in A$ 有 $f(x) = x$. 此时记 f 为 id_A .
- 称 f 是可逆的, 如果存在映射 $f^{-1}: B \rightarrow A$ 满足

$$ff^{-1} = \text{id}_B, \quad f^{-1}f = \text{id}_A.$$

特别地, 称 f^{-1} 为 f 的逆映射.

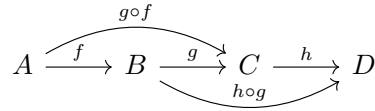
注. 另一种记号是所谓拉回和推出, 分别记作 $f^*g := g \circ f$ 和 $g_*f := g \circ f$.

容易验证恒等映射与逆映射的唯一性. 由此, 映射 id 和 \cdot^{-1} 就是良定义的.

不难得到, 复合运算满足结合律, 分配律但不满足交换律. 这就说明 (B^A, \circ) 是一个群.

为了更加方便地描述映射的复合以及映射之间何时相等, 可以作如下的交换图:

下面的命题从可逆性的角度考虑单射和满射. 我们可以发现, 单射与满射似乎只是方向不同的同类性质而已. 这一点稍后会得到验证.

图 1.1: 复合运算的结合律, 即 $(f \circ g) \circ h = f \circ (g \circ h)$ **命题 1.5**

设映射 $f : A \rightarrow B$. f 为单射, 当且仅当存在映射 g 使得 $gf = \text{id}$. f 为满射, 当且仅当存在映射 g 使得 $fg = \text{id}$. 从而 f 可逆当且仅当它是双射.

接下来我们尝试用第三种方法理解双射:

定义 1.6 单态射, 满态射

设映射 $f : A \rightarrow B$.

- 称 f 为单态射, 若对任意集合 C 和任意映射 $g, h \in C$ 都有 $fg = fh \Leftrightarrow g = h$, 即 f 满足左消去律.
- 同理, 称 f 为满态射, 若 f 满足右消去律.

命题 1.7

设映射 $f : A \rightarrow B$, 则 f 是单射当且仅当 f 是单态射, f 是满射当且仅当 f 是满态射.

1.3 二元关系

幂集公理允许我们构造两个集合的 Cartesian 积.

定义 1.8 Cartesian 积

设集合 A 和 B , 定义它们的 Cartesian 积如下:

$$A \times B := \{(a, b) : a \in A, b \in B\}.$$

注. 特别地, 记 $A^2 := A \times A$, 以及 $A^n := A^{n-1} \times A$ ($n \geq 2$).

定义 1.9 二元关系

设非空集合 S , 则称 S^2 的一个子集 \mathcal{R} 为 S 上的一个二元关系 (binary relation). 若 $(a, b) \in \mathcal{R}$, 则称 a, b 有 \mathcal{R} 关系, 记作 $a\mathcal{R}b$.

例如, 对于集合族 M , 定义在 M 上的关系

$$\Delta := \{(X, Y) \in M^2 : \forall x, (x \in X) \Leftrightarrow (x \in Y)\},$$

那么集合 A, B 相等就可以表述为 $A\Delta B$.

一类在数学中很重要的关系就是等价关系, 它为我们阐明了数学对象的相似性和一致的本质.

定义 1.10 等价关系

设集合 S 及定义在 S 上的关系 \mathcal{R} , 如果对任意 $a, b, c \in S$ 都有:

1. 自反性: $a\mathcal{R}a$;
2. 对称性: $a\mathcal{R}b \Rightarrow b\mathcal{R}a$;
3. 传递性: $a\mathcal{R}b \wedge b\mathcal{R}c \Rightarrow a\mathcal{R}c$.

则称 \mathcal{R} 是 S 上的一个等价关系 (equivalence relation), 通常记作 \sim .

把所有等价的元素放在一起, 就形成了等价类 (equivalence class). 具体地, 定义 a 在 \mathcal{R} 下的等价类

$$[a]_{\mathcal{R}} := \{x \in S : x\mathcal{R}a\},$$

其中 \mathcal{R} 是 S 上的一个等价关系.

例如, 数论中模 n 的同余关系就是一类等价关系, 而模 n 的同余类就是等价类.

等价类内元素都具有同等地位, 都能代表整个等价类, 否则它们也不会被称作是等价的.

命题 1.11 等价类相等等价于代表元素等价

设 \mathcal{R} 是 S 上的等价关系, 对于 $a, b \in S$ 有 $[a]_{\mathcal{R}} = [b]_{\mathcal{R}} \Leftrightarrow a\mathcal{R}b$.

证明 必要性显然. 充分性: 任取 $c \in [a]_{\mathcal{R}}$, 由传递性知 $c\mathcal{R}b$, 所以 $c \in [b]_{\mathcal{R}}$, 从而 $[a]_{\mathcal{R}} \subseteq [b]_{\mathcal{R}}$. 同理有 $[b]_{\mathcal{R}} \subseteq [a]_{\mathcal{R}}$, 所以 $[a]_{\mathcal{R}} = [b]_{\mathcal{R}}$.

还是以模 n 的同余类为例. 我们发现, 任何一个整数都会出现且仅会出现在一个同余类里, 换句话说, 所有的同余类构成类对整数集合的划分.

一般地, 所有的等价类都可以构成对特定集合的划分.

定义 1.12 集合的划分

对于给定集合 S , 集合族 $X = \{S_\alpha : \alpha \in I\}$. 称 X 是 S 的一个划分 (partition), 如果

$$1) S = \bigcup_{\alpha \in I} S_\alpha; \quad 2) \forall \alpha \neq \beta, S_\alpha \cap S_\beta = \emptyset.$$

命题 1.13

设 \mathcal{R} 是 S 上的一个等价关系, 则集合族 $\{[a]_{\mathcal{R}} : a \in S\}$ 构成了 S 的一个划分. 特别地, 称该集合为 S 模 \mathcal{R} 的商集, 记作 S/\mathcal{R} .

证明 首先我们证明, 所有 $[a]_{\mathcal{R}}$ 的并集恰等于 S . 注意到对任意 $a \in S$, $a \in [a]_{\mathcal{R}} \wedge [a]_{\mathcal{R}} \subseteq S$. 所以 $S \subseteq \bigcup_{a \in S} [a]_{\mathcal{R}} \subseteq S$, 从而 $S = \bigcup_{a \in S} [a]_{\mathcal{R}}$.

接着证明这些集合都是不交并. 对于 $[a]_{\mathcal{R}} \neq [b]_{\mathcal{R}}$, 假设存在 $c \in [a]_{\mathcal{R}} \cap [b]_{\mathcal{R}}$, 那么 $c \in [a]_{\mathcal{R}} \wedge c \in [b]_{\mathcal{R}}$, 由等价关系的传递性, $a \mathcal{R} b$, 与假设矛盾. 于是该集合族中任意两个元素交集为空.

借助商集, 我们可以证明一个有趣的命题:

命题 1.14 映射的典范分解

设集合 A, B . 则任一映射 $f : A \rightarrow B$ 均可被写成满射 $\varphi : A \rightarrow C$ 与单射 $\psi : C \rightarrow B$ 的复合.

$$\begin{array}{ccccc} A & \xrightarrow{\varphi} & A/\sim & \xrightarrow{\psi} & B \\ & \searrow & & \nearrow & \\ & & f & & \end{array}$$

图 1.2: 映射的典范分解

证明 考虑等价关系 $\sim := \{(x_1, x_2) \in A^2 : f(x_1) = f(x_2)\}$, 取 $C = A/\sim$. 构造 $\varphi : A \rightarrow C, x \mapsto [x]_{\sim}, \psi : C \rightarrow B, [x]_{\sim} \mapsto f(x)$. 显然 φ 是满射, ψ 是单射, 且 $f = \psi \circ \varphi$.

1.4 序数

1.4.1 全序集与序数

紧接着上一节的内容, 我们可以定义序关系. 此处 $<$ 和 \leq 的区别是自然定义的.

定义 1.15 偏序关系

设集合 S 及定义在 S 上的关系 \mathcal{R} , 如果对任意 $a, b, c \in S$ 都有:

1. 自反性: $a \mathcal{R} a$;
2. 反对称性: $a \mathcal{R} b \wedge b \mathcal{R} a \Rightarrow a = b$;

3. 传递性: $a\mathcal{R}b \wedge b\mathcal{R}c \Rightarrow a\mathcal{R}c$.

则称 \mathcal{R} 是 S 上的一个偏序关系, 记作 \leq .

注. 仅满足自反性与传递性的序结构称作预序关系. 下一章会见到一个预序关系的例子.

注. 若 (S, \leq) 满足 $a \leq b \vee b \leq a$, 则称 S 为全序集 (线序集, 链). 比较符合直觉 (但不严谨) 的看法是: 每个集合上的序结构均可视作是一系列全序子集“粘合”的结果 (这种粘合包括形成线状, 环状等等). 于是研究全序集的性质就是研究序结构最为基础而方便的途径.

我们来思考如何在不同的序结构 $(S_1, \leq_1), (S_2, \leq_2)$ 之间建立映射. 实际上这一操作的思想会在多处有所体现, 例如向量空间之间的线性映射, 拓扑空间之间的连续映射等.

这一映射当然要给出集合 S_1, S_2 元素间的关系, 因此构造映射 $\varphi: S_1 \rightarrow S_2$. 接下来, φ 应当沟通 \leq_1 和 \leq_2 , 故我们定义 φ 是保序的当且仅当对任意 $x, y \in S_1$ 都有 $x \leq_1 y \Rightarrow \varphi(x) \leq_2 \varphi(y)$. 例如, 取数列 $\{a_n\}$ 的子列 $\{a_{k_n}\}$ 实际上就构造了保序映射 k .

单向的映射无法建立一一对应, 因此还要考虑 φ^{-1} . 对 φ^{-1} 重复上述操作, 我们称 φ 是 $(S_1, \leq_1), (S_2, \leq_2)$ 间的同构, 若 φ, φ^{-1} 均为保序映射 (当然, 这个条件可以弱化为 φ 保序). 一个例子就是 Riemann 积分的分割点集 π , 若存在同构 φ , 则极限过程 $\|\pi\| \rightarrow 0$ 和 $\|\varphi(\pi)\| \rightarrow 0$ 是等价的.

下面的定义虽然没有明显动机, 但后面我们会看到这一类序结构的良好性质.

定义 1.16 良序集

称全序集 P 是良序的, 若每个 P 的非空子集均有极小元.

序结构间的同构自然可以构成等价关系, 因而就有了对应的等价类⁵(称作序型). 根据良序定理, 每个集合均可被赋予良序, 因而我们先来研究良序序型.

下面的想法是自然的: 所有良序序型能否排序? 这种想法来源于对最初几个良序序型的枚举, 即 $0, 1, 2, \dots$ (具体构造可借助 $\emptyset, \emptyset^+, (\emptyset^+)^+$). 另一个想法是, 在序型超出自然数集大小 (即可数无穷序数 ω) 时, 需要在原有序型基础上增加何种长度的良序集才能得到一个新的序型 (进而不致被视作与原序型同构).

然而, 在解决问题之前, 正如前文所述, 我们需要优化序型的定义, 否则在 ZFC 系统中考虑一个类 (同构类) 的类 (全体序型构成的类) 是风险重重的. von Neumann 的解决思路是在每个同构类中挑出一个标准的良序集:

定义 1.17 序数

称集合 α 是序数, 若 α 是传递的, 即其任一元素均为其子集, 且 α 对于 \in 构成良序集.

此种定义方式显然是针对后继集 X^+ 设计的, 因为 α 是序数蕴含 α^+ 是序数. 方便起见, 后文均用 $\alpha + 1$ 表示 α^+ , 用 $<$ 表示 \in , 记所有序数构成类 **On**.

⁵这不一定是集合.

下方的命题可借助自然数集来理解. 证明仅是简单的推演, 故略去.

命题 1.18

- 设序数 α , 则任意 $\beta \in \alpha$ 均为序数, 进而 $\alpha = \{\beta \in \mathbf{On} : \beta < \alpha\}$.
- 对于序数 α, β , 若 $\beta \subsetneq \alpha$ 则 $\beta \in \alpha$.
- 非空的序数集配上 $<$ 是良序集.

现在逐一解决先前的两个问题. 首先, 我们 (不加证明地⁶) 给出下方的定理:

定理 1.19 良序集基本定理

对于良序集 (P, \leq) , 称形如 $P_{<a} := \{x \in P : x < a\}$ 的集合配上继承自 P 的序结构为 P 的一个前段. 于是, 对于良序集 $(P_1, \leq_1), (P_2, \leq_2)$, 下列三种情况有且仅有一种成立:

- 1) P_1 与 P_2 同构; 2) P_1 与 P_2 的一个前段同构; 3) P_2 与 P_1 的一个前段同构.

自然数的三歧性就是该定理的直接推论.

第二个问题就涉及到怎样构造更大 (以及最大) 的序数. 实际上我们有:

命题 1.20

- 对任意非空序数集合 X , $\bigcap X$ 是序数, $\bigcap X = \inf X$, 且 $\inf X \in X$. 进而, $\alpha + 1 = \inf\{\beta \in \mathbf{On} : \beta > \alpha\}$.
- 对任意序数集合 X , $\bigcup X$ 是序数, $\bigcup X = \sup X$.

由此, \mathbf{On} 不是集合, 否则 $\sup \mathbf{On}$ 是一个序数, 从而 $\sup \mathbf{On} + 1$ 是序数, 易得矛盾.

根据该命题, $\alpha + 1$ 就是大于 α 的最小序数. 我们称 α 为后继序数, 若存在序数 β 使得 $\beta + 1 = \alpha$. 若 α 不是后继序数, 则称其为极限序数. 不难证明此时 $\alpha = \sup\{\beta \in \mathbf{On} : \beta < \alpha\}$.

容易验证, 将自然数集视作序数 ω 后, ω 就是最小的非零极限序数.

1.4.2 归纳法与选择公理

这一节的中心内容就是如下 (互相等价的) 三个命题: 选择公理, 良序定理, Zorn 引理. 我们将用归纳构造验证后两个命题, 并给出序数类似于自然数的一些性质.

⁶证明可参考 [?] 的定理 4.1.6. 读者亦可结合下一节内容用归纳法给出证明.

定理 1.21 超穷归纳法

设性质 $P(x)$, 若对所有序数 α 均有

$$\forall \beta < \alpha, P(\beta) \Rightarrow P(\alpha),$$

则 $P(\alpha)$ 对所有序数 α 均成立.

注. 若对后继序数和极限序数分别讨论, 则条件可等价写作: (1) $P(0)$ 成立; (2) 对所有后继序数 α , $P(\alpha) \Rightarrow P(\alpha + 1)$; (3) 对所有非零极限序数, $\forall \beta < \alpha, P(\beta) \Rightarrow P(\alpha)$.

证明 假设存在 γ 使得 $P(\gamma)$ 不成立, 考虑 $\alpha = \inf\{\beta \leq \gamma : P(\beta) \text{ 不成立}\}$, 则任意比 α 小的序数均成立 P , 从而 $P(\alpha)$ 成立, 矛盾.

接下来验证归纳构造这一操作的合理性.

根据前文的说法, 函数并不必要建立在集合基础上, 因此形如 $\mathbf{G} : \mathbf{V} \rightarrow \mathbf{V}$ 的函数是可被定义的. 在下方的定理中, 一个 θ -序列是指定义域为序数 θ 的函数.

推论 1.22 超穷递归原理

设函数 $\mathbf{G} : \mathbf{V} \rightarrow \mathbf{V}$, 则存在唯一的 θ -序列 a 使得当 $\alpha < \theta$ 时总有 $a_\alpha = \mathbf{G}(a|_\alpha)$. 从而, 存在唯一的序列 $a : \mathbf{On} \rightarrow \mathbf{V}$ 使得对任意 α 成立 $a_\alpha = \mathbf{G}(a|_\alpha)$.

注. 由于 $\alpha = \{\beta \in \mathbf{On} : \beta < \alpha\}$, 故 $\mathbf{G}(a|_\alpha)$ 实际在说将 \mathbf{G} 应用到所有比 α 小的序数上.

注. 这一条件蕴含了 $a_0 = \mathbf{G}(\emptyset) = \mathbf{G}(0)$.

证明 进行归纳证明. $\theta = 0$ 时显然. 当 $\theta > 0$ 时, 假设对任意 $\theta' < \theta$ 命题均成立.

存在性: 基本的想法是将每个 θ' 所对应的 $a[\theta']$ 粘在一起使得 $a|_{\theta'} = a[\theta']$. 唯一的问题在于这些 $a[\theta']$ 重叠的部分应当一致, 而这由归纳假设中的唯一性可得.

唯一性: 若 a, a' 均合题, 由归纳假设 $a|_\theta = a'|_\theta$, 从而 $a_\theta = a'_\theta$.

类似于自然数, 可以归纳定义序数的运算:

	$\alpha + \beta$	$\alpha \cdot \beta$	α^β
$\beta = 0$	0	0	1
β 是后继序数	$(\alpha + (\beta - 1)) + 1$	$\alpha \cdot (\beta - 1) + \alpha$	$\alpha^{\beta-1} \cdot \alpha$
β 是极限序数	$\sup\{\alpha + \theta : \theta < \beta\}$	$\sup\{\alpha \cdot \theta : \theta < \beta\}$	$\sup\{\alpha^\theta : \theta < \beta\}$

由于涉及到极限序数, 序数的加法和乘法满足结合律, 但一般不满足交换律. 例如 $n + \omega = \sup\{n + m : m < \omega\} = \omega \neq \omega + n$. 另外不难验证, 如果正确地定义了良序集的加法和乘法, 则序数的加/乘法即相当于将其视作良序集做加/乘法. 有关序数算术的进一步内容, 仍请参考 [?] 的 4.4 小节.

定理 1.23 良序定理

任一集合均可赋予良序.

证明 考虑集合 S 上的选择函数 f . 对每个序数 α , 当 $\{a_\beta : \beta < \alpha\} \neq S$ 时 (即还未取尽时), 递归定义 $a_\alpha := f(S \setminus \{a_\beta : \beta < \alpha\})$. 取最小的 θ 使得 a_θ 无定义 (若这样的 θ 不存在, 则 $a : \mathbf{On} \rightarrow S$, 然而由替换公理 $a^{-1}(a(\mathbf{On})) = \mathbf{On}$ 是集合, 矛盾), 易得 a 是 S 到 θ 的同构.

下方的命题验证了序数确是良序集等价类的代表元.

命题 1.24 Mirimanoff-von Neumann

对任意良序集 P , 存在唯一的序数 α 与之同构.

证明 唯一性: 这由序数的全序性立得.

存在性: 只需将良序定理证明中的选择函数 f 换做最小值函数 \min 即可.

由于良序集与唯一的序数等价, 我们有另一个版本的超穷递归原理:

推论 1.25 超穷递归原理

设函数 $\mathbf{G} : \mathbf{V} \rightarrow \mathbf{V}$, 良序集 P , 则存在唯一的序列 $a : P \rightarrow \mathbf{V}$ 使得对任意 $x \in P$ 成立 $a_x = \mathbf{G}(a|_x)$.

最后介绍 Zorn 引理. 该定理的一个应用是证明任意 (有限或无限维) 向量空间都存在一组基.

定理 1.26 Zorn 引理

设非空偏序集 P , 且 P 中每个链均有上界, 则 P 存在极大元.

证明

1.5 基数

高中数学中, 我们学过有限集合的元素个数. 从直观上看, 似乎无限集合不会存在元素个数这一说法, 但我们又熟知实数远比整数多, 那么这种相对的元素个数比较是怎样建立的?

来考虑这样一个问题: 给定两个有限集合 A, B , 如何比较它们的元素个数. 最一般的想法应该是在它们之间构造一个映射 $f : A \rightarrow B$, 如果 f 是双射则 A, B 元素个数相等, 如果是单射则 A 的元素个数不多于 B 的元素个数, 如果是满射则 B 的元素个数不多于 A 的元素个数 (这些用反证法容易说明).

相对应地, 既然我们只需要考虑无限集合之间的相对“元素个数”多少, 而不需要得到一个绝对数值, 就可以仿照上方的方法定义一个无限集合的“相对元素个数”. 非常直观地, 我们也将其称为“势”, 这是否让你想起电势? 在接下来的内容中, 你将看到集合的“势”的参考位置一般取用自然数集合.

定义 1.27 等势集合

对于集合 A, B , 若存在单射 $f: A \rightarrow B$, 则称 A 的势小于等于 B . 特别地, 若单射 f 同时也是一个满射, 即 f 是双射, 则称 A, B 等势 (equipollent).

很自然地, 我们可以证明集合的等势关系是一个等价关系. 为了证明势的小于等于是一个全序关系, 需要下方的定理:

定理 1.28 Schröder–Bernstein

给定集合 A, B . 若在 A, B 间存在两个单射 $f: A \rightarrow B$ 与 $g: B \rightarrow A$, 则在它们之间也存在一个双射 $h: A \rightarrow B$.

证明 证法一 (不依赖选择公理的构造性证明) 不妨考虑 A, B 非空. 由于 g 是 $B \rightarrow g(B)$ 的双射, 我们可以用 $g(B)$ 替换 B , 即不妨设 $f(A) \subseteq B \subseteq A$. 令 $A_0 = A, B_0 = B$, 递归地定义 $A_{n+1} = f(A_n), B_{n+1} = f(B_n)$. 于是得到 $A_0 \supseteq B_0 \supseteq \cdots \supseteq A_n \supseteq B_n \supseteq \cdots$. 直接给出 h 的构造:

$$h(x) = \begin{cases} f(x) & \exists n \in \mathbb{Z}_+, x \in A_n - B_n \\ x & \text{否则} \end{cases}.$$

下面验证 h 是双射: 若 $h(x) = h(y)$ 而 $x \neq y$, 只能 $x \in A_N - B_N, y \notin A_n - B_n, \forall n$. 但是 $y = h(y) = h(x) = f(x) \in A_{N+1}, \notin B_{N+1}$, 即 $y \in A_{N+1} - B_{N+1}$, 矛盾. 这说明 h 是单射. 另一方面, 任取 $y \in B$, 若 $y \notin f(A)$, 由 $B_0 \supseteq A_1$ 可知不存在 $n \geq 1$ 使得 $f(y) \in A_n$, 从而 $h(y) = y$. 这说明 h 是满射.

由上方的定理, 容易得到势的小于等于关系满足反对称性. 该关系的完全性是选择公理的推论 (这里略去). 再加上传递性 (例如, A, B 之间存在单射 f, B, C 之间存在单射 g , 则 $g|_{f(A)} \circ f$ 是 A, C 间的单射), 马上得到该关系是一个全序关系.

从而, 我们可以利用等价类的思想刻画一个无限集合的相对元素个数.

定义 1.29 集合的基数

- 设集合的等势关系 \mathcal{R} . 对于集合 X , 称 $[X]_{\mathcal{R}}$ 为其基数 (cardinal) 或势, 记作 $|X|$.
- 定义 $|X| = |Y|$, 如果 X 与 Y 等势.
- 定义 $|X| \leq |Y|$, 如果 X 与 Y 的某个子集等势.

容易证明集合基数的小于等于关系也是一个全序关系.

现在对用等价类定义的基数做一些说明: 这种定义方式其实不够好, 因为如果我们要考虑所有基数构成的“集合”, 实际上是在考虑一个集合族, 而集合族不一定是集合. 更好的方法是取等价类中的某个代表元素 (一般取的是最小序数). 利用取代表元素的定义方法, 我们可以引入定理28的第二种证明:

证明 证法二(承认选择公理的证明, 了解即可) 先不加证明地给出一个引理 (良序定理): 任何集合 S 上均存在一个序结构 \prec , 使得 (S, \prec) 是一个良序结构, 即对任意 S 的子集都存在关于 \prec 的极小元. 等价地, 存在唯一的极小序数 $|S|$ 使得 $S, |S|$ 之间存在双射. 在下方的证明中, 我们实际上将基数的反对称性处理成了序数的反对称性 (不加证明地承认).

回到原题, 不妨考虑 A, B 非空. 则存在唯一的极小序数 $|A|$ 和双射 $\varphi: A \rightarrow |A|$. 类似地定义 $|B|$ 和 $\psi: B \rightarrow |B|$. 由于 $\psi \circ f: X \rightarrow |Y|$ 是单射, 将其视作陪域等于值域的映射就是双射. 由替换公理, 其值域 $\text{range}(\psi \circ f)$ 亦是集合, 故可良序化, 即存在唯一极小序数 $|\text{range}(\psi \circ f)|$ 和双射 $\alpha: \text{range}(\psi \circ f) \rightarrow |\text{range}(\psi \circ f)|$. 于是

$$\alpha \circ \psi \circ f: X \rightarrow |\text{range} \psi \circ f| \leq |Y|$$

是双射, 而由 $|X|$ 的极小性可知 $|X| \leq |\text{range} \psi \circ f| \leq |Y|$. 同理可得 $|Y| \leq |X|$. 于是 $|X| = |Y|$.

关于无限集合, Cantor 曾证明: (这里 $|X| < |Y|$ 承自然的严格偏序定义)

定理 1.30

设集合 X , 则 $|X| < |\mathcal{P}(X)|$.

证明 若 X 是空集, 则显然成立. 从而, 只考虑 X 非空的情况.

由于 $\mathcal{P}(X)$ 涵盖所有 X 的一元子集, 故显然有 $|X| \leq |\mathcal{P}(X)|$. 假设有 $|X| = |\mathcal{P}(X)|$, 那么存在双射 $f: X \rightarrow \mathcal{P}(X)$.

根据 f , 取 $B = \{x \in X : x \notin f(x)\}$, 显然 $B \in \mathcal{P}(X)$, 从而存在 x 使得 $f(x) = B$. 此时, 若 $x \in B$, 则由 B 的定义知 $x \notin B$, 矛盾; 同理, 若 $x \notin B$, 则可得 $x \in B$, 也矛盾.

Chapter 2

范畴论初步

定义 2.1 范畴

例 2.1 Slice Categories

Part II

线性映射与代数结构

Chapter 3

向量空间

3.1 从 \mathbb{F}^n 说起

3.1.1 复数与复数域

首先来温习一下复数域 \mathbb{C} 的定义与它满足的性质：

定义 3.1 复数

记 $z = a + bi (a, b \in \mathbb{R})$ 为一个复数，其中 $i^2 = -1$. 由所有复数构成的集合记为 \mathbb{C} . \mathbb{C} 上的加法与乘法定义如下：

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

命题 3.2 复数运算的性质

命题 3.2 (1) 交换性质

$$\forall \alpha, \beta \in \mathbb{C}, \alpha + \beta = \beta + \alpha, \alpha\beta = \beta\alpha$$

(2) 结合性质

$$\forall \alpha, \beta, \lambda \in \mathbb{C}, (\alpha + \beta) + \lambda = \alpha + (\beta + \lambda), (\alpha\beta)\lambda = \alpha(\beta\lambda)$$

(3) 单位元

$$\forall \lambda \in \mathbb{C}, \lambda + 0 = \lambda, 1\lambda = \lambda$$

(4) 加法逆元

$$\forall \alpha \in \mathbb{C}, \exists! \beta \in \mathbb{C}, \alpha + \beta = 0$$

(5) 乘法逆元

$$\forall \alpha \in \mathbb{C} (\alpha \neq 0), \exists \beta \in \mathbb{C}, \alpha\beta = 1$$

(6) 分配性质

$$\forall \lambda, \alpha, \beta \in \mathbb{C}, \lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$$

证明 这里只选择部分性质证明:

(1) 加法交换性质: 设 $\alpha = a + bi, \beta = c + di$ ($a, b, c, d \in \mathbb{R}$), 则

$$\begin{aligned}\alpha + \beta &= (a + bi) + (c + di) \\ &= (a + c) + (b + d)i \\ &= (c + a) + (d + b)i \\ \beta + \alpha &= (c + di) + (a + bi) \\ &= (c + a) + (d + b)i\end{aligned}$$

因此有 $\alpha + \beta = \beta + \alpha$

(2) 乘法单位元: 设 $\lambda = a + bi$ ($a, b \in \mathbb{R}$), 那么

$$1\lambda = (1 + 0i)(a + bi) = a + bi = \lambda$$

(3) 加法逆元: 先证明存在. 设 $\alpha = a + bi$, 取 $\beta = (-a) + (-b)i$, 则 $\alpha + \beta = 0 + 0i = 0$;

再证明唯一. 假设 $\beta_1, \beta_2 \in \mathbb{C}$ 均为 α 的加法逆元, 那么

$$\beta_1 = \beta_1 + 0 = \beta_1 + \alpha + \beta_2 = 0 + \beta_2 = \beta_2$$

这与假设矛盾, 则 α 的加法逆元是唯一的.

由此可以引出域的正式定义:

定义 3.3 域

域是一个集合 \mathbb{F} , 它带有加法与乘法两种运算 (分别在加法与乘法上封闭), 且这些运算满足命题??所示所有性质.

注. 最小的域是一个集合 $\{0, 1\}$, 带有通常的加法与乘法运算, 但规定 $1 + 1 = 0$.

容易验证, \mathbb{R} 与 \mathbb{C} 都是域. 本书中用 \mathbb{F} 来表示 \mathbb{R} 或 \mathbb{C} .

总是用 β 表示 α 的逆元非常不自然, 因此定义出加/乘法逆元的表示与减/除法.

定义 3.4 加法逆元, 减法, 乘法逆元, 除法

设 $\alpha, \beta \in \mathbb{C}$.

- 令 $-\alpha$ 表示 α 的加法逆元, 即 $-\alpha$ 是使得

$$\alpha + (-\alpha) = 0$$

成立的唯一复数.

- 对于 $\alpha \neq 0$, 令 α^{-1} 表示 α 的乘法逆元, 即 α^{-1} 是使得

$$\alpha(\alpha^{-1}) = 1$$

成立的唯一复数.

- 定义 \mathbb{C} 上的减法:

$$\beta - \alpha = \beta + (-\alpha)$$

- 定义 \mathbb{C} 上的除法:

$$\beta/\alpha = \beta(1/\alpha)$$

3.1.2 \mathbb{F}^n

在中学的向量板块, 我们认识到一个向量可以表示为有序数组 (a, b) 的形式, 并且在立体几何板块利用三维下的向量进行了许多计算. 那么向量的定义能否推广到更高维度呢?

定义 3.5 \mathbb{F}^n

\mathbb{F}^n 是 \mathbb{F} 中元素组成的长度为 n 的组的集合, 即

$$\mathbb{F}^n = \{(x_1, \dots, x_n) : x_j \in \mathbb{F}, j = 1, \dots, n\}$$

特别地, 对于由无限长度序列构成的集合, 称作 \mathbb{F}^∞ , 即

$$\mathbb{F}^\infty = \{(x_1, \dots, x_n, \dots) : x_j \in \mathbb{F}, j = 1, \dots, n, \dots\}$$

对于 \mathbb{F}^n 中的某个元素 (x_1, \dots, x_n) , 称 x_j ($j = 1, \dots, n$) 为 (x_1, \dots, x_n) 的第 j 个坐标.

\mathbb{F}^n 上的加法定义为对应坐标相加, 即

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

对于 \mathbb{F}^∞

$$(x_1, \dots, x_n, \dots) + (y_1, \dots, y_n, \dots) = (x_1 + y_1, \dots, x_n + y_n, \dots)$$

\mathbb{F}^n 上的标量乘法：一个数 λ ($\lambda \in \mathbb{F}$) 与 \mathbb{F}^n 中元素的乘积这样计算：用 λ 乘以该元素的每个坐标，即

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

对于 \mathbb{F}^∞

$$\lambda(x_1, \dots, x_n, \dots) = (\lambda x_1, \dots, \lambda x_n, \dots)$$

我们暂时不讨论 \mathbb{F}^n 上元素之间的乘法.

当 \mathbb{F} 代表 \mathbb{R} 且 $n = 2, 3$ 时, \mathbb{F}^n 中的元素就相当于我们熟悉的平面向量、空间向量. 实际上, 所有在 \mathbb{F} 中的元素都被称为标量, 所有在 \mathbb{F}^n 中的元素如果被看做是一个从原点指向某定点的有向线段时, 它就是向量. 我们一般用小写字母表示标量, 用加粗的小写字母表示 \mathbb{F}^n 中的元素, 例如 \mathbb{F}^4 中的元素

$$\mathbf{x} = (x_1, x_2, x_3, x_4)$$

特别地, 用 $\mathbf{0}$ 表示所有坐标全是 0 的元素, 即

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

\mathbb{F}^n 同样也具有类似于 \mathbb{F} 的一些性质:

命题 3.6 \mathbb{F}^n 的性质

xlxkvi (1) 交换性质

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{F}^n, \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

(2) 结合性质

$$\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{F}^n, a, b \in \mathbb{F}, (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}), (ab)\mathbf{v} = a(b\mathbf{v})$$

(3) 加法单位元

$$\exists! \mathbf{0} \in \mathbb{F}^n, \forall \mathbf{v} \in \mathbb{F}^n, \mathbf{v} + \mathbf{0} = \mathbf{v}$$

(4) 加法逆元

$$\forall \mathbf{v} \in \mathbb{F}^n, \exists! \mathbf{w} \in \mathbb{F}^n, \mathbf{v} + \mathbf{w} = \mathbf{0}$$

(5) 乘法单位元

$$\forall \mathbf{v} \in \mathbb{F}^n, 1\mathbf{v} = \mathbf{v}$$

(6) 分配性质

$$\forall a, b \in \mathbb{F}, \mathbf{u}, \mathbf{v} \in \mathbb{F}^n, a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}, (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$$

证明 这里只选择部分证明:

(1) 交换性质: 设 $\mathbf{u} = (u_1, \dots, u_n), \mathbf{v} = (v_1, \dots, v_n)$, 则

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= (u_1, \dots, u_n) + (v_1, \dots, v_n) \\ &= (u_1 + v_1, \dots, u_n + v_n) \\ &= (v_1 + u_1, \dots, v_n + u_n) \\ &= (v_1, \dots, v_n) + (u_1, \dots, u_n) \\ &= \mathbf{v} + \mathbf{u} \end{aligned}$$

(2) 加法单位元: 先证明存在. 若 $\mathbf{v} = (v_1, \dots, v_n)$, 取 $-\mathbf{v} = (-v_1, \dots, -v_n)$, 容易发现 $\mathbf{v} + -\mathbf{v} = \mathbf{0}$; 再证明唯一. 假设存在两个加法单位元 $\mathbf{0}$ 与 $\mathbf{0}'$, 则

$$\mathbf{0} = \mathbf{0} + \mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}'$$

这与假设矛盾. 因此最多只有一个加法单位元.

3.2 向量空间

类似于 \mathbb{F}^n ，我们把向量空间定义为带有加法和标量乘法的集合 V ，其满足命题??中的性质. 请注意，由于不一定满足乘法交换性质，向量空间不一定是一个域.

定义 3.7 加法，标量乘法

- 集合 V 上的加法是一个函数，它把每一对 $u, v \in V$ 都对应到 V 中的一个元素 $u + v$.
- 集合 V 上的标量乘法是一个函数，它把任意 $\lambda \in \mathbb{F}$ 和 $v \in V$ 都对应到 V 中的一个元素 λv .

注. 换句话说， V 对加法和标量乘法封闭.

接下来可以正式定义向量空间：

定义 3.8 向量空间

向量空间就是带有加法和标量乘法的集合 V ，满足如下性质：

(1) 交换性质

$$\forall u, v \in V, u + v = v + u$$

(2) 结合性质

$$\forall u, v, w \in V, a, b \in \mathbb{F}, (u + v) + w = u + (v + w), (ab)v = a(bv)$$

(3) 加法单位元

$$\exists \mathbf{0} \in V, \forall v \in V, v + \mathbf{0} = v$$

(4) 加法逆元

$$\forall v \in V, \exists w \in V, v + w = \mathbf{0}$$

(5) 乘法单位元

$$\forall v \in V, 1v = v$$

(6) 分配性质

$$\forall a, b \in \mathbb{F}, u, v \in V, a(u + v) = au + av, (a + b)v = av + bv$$

向量空间中的元素被称为向量或点.

注. 因为向量空间的标量乘法依赖于 \mathbb{F} ，所以一般会说 V 是 \mathbb{F} 上的向量空间. 例如，平面点集 \mathbb{R}^2 是 \mathbb{R} 上的向量空间. 如果没有特别说明，默认 V 就表示在 \mathbb{F} 上的向量空间.

注. 最小的向量空间是 $\{0\}$ ，它带有通常的加法和乘法运算.

在向量空间的定义中并没有说明唯一性，这是因为唯一性可以通过已有的性质证明出.

现在介绍一个具体的例子:

定义 3.9 \mathbb{F}^S

设 S 是一个集合, 我们用 \mathbb{F}^S 表示 S 到 \mathbb{F} 的所有函数的集合.

对于 $f, g \in \mathbb{F}^S$, 对所有 $x \in S$, 规定 \mathbb{F}^S 上的加和 $f + g$ 满足

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

对于 $\lambda \in \mathbb{F}$ 和 $f \in \mathbb{F}^S$, 对所有 $x \in S$, 规定 \mathbb{F}^S 上的标量乘法得到的乘积 $\lambda f \in \mathbb{F}^S$ 满足

$$(\lambda f)(x) = \lambda f(x)$$

例 3.1 请证明 \mathbb{F}^S 是 \mathbb{F} 上的向量空间, 并指出它的加法单位元与加法逆元.

向量空间的定义中缺少了一些显而易见的性质, 我们现在进行补充:

命题 3.10 向量空间的性质

xlkjxkvi

- 向量空间有唯一的加法单位元.
- 向量空间中的每个元素都有唯一的加法逆元.
- 对任意 $v \in V$ 都有 $0v = 0$.
- 对任意 $a \in \mathbb{F}$ 都有 $a0 = 0$.
- 对任意 $v \in V$ 都有 $(-1)v = -v$. (等式右边的 $-v$ 表示 v 的加法逆元)

证明 设向量空间 V ,

(1) 假设 V 中有两个不同的加法单位元 $0, 0'$, 那么

$$0 = 0 + 0' = 0' + 0 = 0'$$

这与假设矛盾, 于是向量空间中只有唯一的加法单位元.

(2) 对于 $v \in V$, 假设 w, w' 都是它的加法逆元, 那么

$$w = w + 0 = w + v + w' = 0 + w' = w'$$

这与假设矛盾, 于是向量空间中每个元素都有唯一的加法逆元.

(3) 对于 $v \in V$, 由于

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

在等式两边同时加上 $0v$ 的加法逆元, 可得 $0v = 0$.

(4) 与 (3) 同理, 请读者自行证明.

(5) 对于 $v \in V$, 由于

$$0 = (1 + (-1))v = v + (-1)v$$

在等式两边同时加上 v 的加法逆元, 可得 $(-1)v = -v$.

注. 在 (3) 的证明过程中, 由于在向量空间中只有分配性质能将标量乘法与向量的加法联系在一起, 故必然会利用分配性质.

3.3 子空间

就像构造集合时要研究一个集合的子集一样, 在向量空间中, 我们也要研究它的子集. 特别地, 向量空间的子集如果也是向量空间, 我们把它称作子空间.

3.3.1 子空间

定义 3.11 子空间

设向量空间 V 和它的一个子集 U (采用与 V 相同的加法法则与标量乘法法则), 如果 U 也是一个向量空间, 则称 U 是 V 的子空间.

然而在实际应用中, 每遇到一个子集 U 都证明一遍它是向量空间是很麻烦的. 其实只需要证明以下三个关键性质:

命题 3.12 子空间的判定条件

设向量空间 V 的子集 U , U 是 V 的子空间当且仅当 U 满足下列条件:

(1) 加法单位元

$$0 \in U$$

(2) 加法封闭性

$$\forall u, v \in U, u + v \in U$$

(3) 标量乘法封闭性

$$\forall \lambda \in \mathbb{F}, v \in U, \lambda v \in U$$

证明 1° 必要性: 当 U 是 V 的子空间时, 由定义可知 U 是一个向量空间, 则它自然满足上述条件.

2° 充分性: 当 U 满足上述条件时, 由于 U 是 V 的子集并拥有相同的运算规则, 显然 U 可以满足向量空间的所有性质.

注. 该判定条件中有关加法单位元的性质等价于“ U 非空”. (取 $v \in U, 0 \in \mathbb{F}$, 由标量乘法封闭性与命题 3.12 的第三条可知 $0v = 0 \in U$)

注. 实际上子空间的判定条件就是向量空间的必要条件: 拥有加法单位元, 且对加法和标量乘法封闭.

例 3.2 请指出下列向量空间的所有子空间: (不要求证明唯一性, 我们会在下一章给出证明)

(1) 定义在 \mathbb{R} 上的向量空间 \mathbb{R}^2 ;

(2) 定义在 \mathbb{R} 上的向量空间 \mathbb{R}^3 .

解. (1) $\{0\}$ 、 \mathbb{R}^2 和 \mathbb{R}^2 中过原点的所有直线.

(2) $\{0\}$ 、 \mathbb{R}^3 和 \mathbb{R}^3 中过原点的所有平面.

例 3.3 证明下列结论:

- (1) 若 $b \in \mathbb{F}$, 则 $U = \{(x_1, x_2, x_3, x_4) \in \mathbb{F}^4 : x_3 = 5x_4 + b\}$ 是 \mathbb{F}^4 的子空间当且仅当 $b = 0$;
- (2) 区间 $[0, 1]$ 上的全体实值连续函数的集合是 $\mathbb{R}^{[0,1]}$ 的子空间;
- (3) 区间 $(0, 3)$ 上满足条件 $f'(2) = b$ 的实值可微函数的集合是 $\mathbb{R}^{(0,3)}$ 的子空间当且仅当 $b = 0$;
- (4) 极限为 0 的复数序列组成的集合是 \mathbb{C}^∞ 的子空间.

证明 (1) 1° 充分性: 当 $b = 0$ 时, 显然 $0 = (0, 0, 0, 0) \in U$. 取 U 中两个元素 $v = (v_1, v_2, 5v_4, v_4)$ 与 $u = (u_1, u_2, 5u_4, u_4)$, 取 \mathbb{F} 中标量 λ . 因为

$$v + u = (v_1 + u_1, v_2 + u_2, 5v_4 + 5u_4, v_4 + u_4) = (v_1 + u_1, v_2 + u_2, 5(v_4 + u_4), v_4 + u_4) \in U$$

$$\lambda v = (\lambda v_1, \lambda v_2, \lambda 5v_4, \lambda v_4) = (\lambda v_1, \lambda v_2, 5(\lambda v_4), \lambda v_4) \in U$$

这告诉我们 U 对加法和标量乘法封闭, 于是 U 是 \mathbb{F}^4 的子空间.

2° 必要性: 任取 U 中两个元素 $v = (v_1, v_2, 5v_4 + b, v_4)$ 与 $u = (u_1, u_2, 5u_4 + b, u_4)$, 取 \mathbb{F} 中标量 λ . 因为

$$(0, 0, 0, 0) \in U$$

$$v + u = (v_1, v_2, 5v_4 + b, v_4) + (u_1, u_2, 5u_4 + b, u_4) = (v_1 + u_1, v_2 + u_2, 5(v_4 + u_4) + 2b, v_4 + u_4) \in U$$

$$\lambda v = (\lambda v_1, \lambda v_2, 5\lambda v_4 + \lambda b, \lambda v_4) \in U$$

则 $0 = 0 + b$, $5(v_4 + u_4) + 2b = 5(v_4 + u_4) + b$, $5\lambda v_4 + \lambda b = 5\lambda v_4 + b$, 这要求 $b = 0$.

(3) 1° 充分性: 设函数 $0 : x \mapsto 0$, 容易验证 0 是该集合的加法单位元; 取函数 $f, g \in \mathbb{R}^{(0,3)}$, 由于 $(f+g)'(2) = f'(2) + g'(2) = 0$, 可知 $f+g \in \mathbb{R}^{(0,3)}$, 即该集合对加法封闭; 取函数 $f \in \mathbb{R}^{(0,3)}$, 标量 $\lambda \in \mathbb{F}$, 由于 $(\lambda f)'(2) = \lambda f'(2) = 0$, 可知 $\lambda f \in \mathbb{R}^{(0,3)}$, 即该集合对标量乘法封闭.

2° 必要性: 由例题 1.2.1 的结论, 该集合中必有加法单位元 $0 : x \mapsto 0$, 则 $0'(2) = 0 = b$; 取函数 $f, g \in \mathbb{R}^{(0,3)}$, 由于该集合对加法封闭, 可知 $(f+g)'(2) = f'(2) + g'(2) = 2b = b$, 则 $b = 0$; 取函数 $f \in \mathbb{R}^{(0,3)}$, 标量 $\lambda \in \mathbb{F}$, 由于该集合对标量乘法封闭, 有 $(\lambda f)'(2) = \lambda f'(2) = \lambda b = b$, 则 $b = 0$.

3.3.2 子空间的和

继续与集合比较. 我们发现集合间有交、并、补等运算, 向量空间中也有对应的运算, 不过我们感兴趣的通常是它们的和.(详细原因参考本节习题)

定义 3.13 子集的和

设 U_1, \dots, U_m 都是 V 的子集, 定义 U_1, \dots, U_m 的和为 U_1, \dots, U_m 中元素所有可能的和构成的集合, 记作 $U_1 + \dots + U_m$, 即

$$U_1 + \dots + U_m = \{u_1 + \dots + u_m : u_j \in U_j, j = 1, \dots, m\}$$

例 3.4 证明下列结论:

(1) 设

$$U = \{(x, 0, 0) \in \mathbb{F}^3 : x \in \mathbb{F}\}, \quad W = \{(0, y, 0) \in \mathbb{F}^3 : y \in \mathbb{F}\}$$

则

$$U + W = \{(x, y, 0) : x, y \in \mathbb{F}\}$$

(2) 设

$$U = \{(x, x, y, y) \in \mathbb{F}^4 : x, y \in \mathbb{F}\}, \quad W = \{(x, x, x, y) \in \mathbb{F}^4 : x, y \in \mathbb{F}\}$$

则

$$U + W = \{(x, x, y, z) : x, y, z \in \mathbb{F}\}$$

两个集合的并集是包含它们的最小集合. 相应地, 两个子空间的和是包含它们的最小子空间.

命题 3.14 子空间的和是包含这些子空间的最小子空间

设 U_1, \dots, U_m 都是 V 的子空间, 则 $U_1 + \dots + U_m$ 是 V 的包含 U_1, \dots, U_m 的最小子空间.

证明 记 $U = U_1 + \dots + U_m$.

1° 证明 U 是 V 的子空间: 显然 $0 = 0 + \dots + 0 \in U$; 取 $x_1 + \dots + x_m, y_1 + \dots + y_m \in U$, 其中 $x_i, y_i \in U_i (i = 1, \dots, m)$, 由于对任意 i 都有 $x_i + y_i \in U_i$, 所以 $(x_1 + y_1) + \dots + (x_m + y_m)$ 也在 U 中, 因此 U 对加法封闭; 取 $x_1 + \dots + x_m \in U$, 由于对任意 i 都有 $\lambda x_i \in U_i$, 所以 $\lambda x_1 + \dots + \lambda x_m$ 也在 U 中, 因此 U 对标量乘法封闭. 综上, U 是 V 的子空间.

2° 证明 U 包含 U_1, \dots, U_m : 取 U_j 中元素 u_j , 再取其他子空间中的元素 0 , 可知 $u_j \in U$. 因此任意一个子空间都包含于 U .

3° 证明 U 是最小的满足条件的子空间: 假设存在一个更小的 U' , 由于 U' 包含 U_1, \dots, U_m 中的所有元素, 又因为 U' 对加法封闭, 故 U' 中必有 $U_1 + \dots + U_m$ 中所有元素, 这与假设矛盾. 因此 U 是最小的满足条件的子空间.

3.3.3 直和

注意到子空间的和中的元素 u 可以用不同的 $u_1 + \dots + u_m$ 来表示. 为了尽量避免这种不确定性, 规定一种能够唯一地表示为上述形式的情形.

定义 3.15 直和

设 U_1, \dots, U_m 都是 V 的子空间. 和 $U_1 + \dots + U_m$ 称为直和, 如果 $U_1 + \dots + U_m$ 中的每个元素都能唯一地表示成 $u_1 + \dots + u_m$ 的形式, 其中每个 u_j 都属于 U_j . 特别地, 用 $U_1 \oplus \dots \oplus U_m$ 表示一个直和.

例 3.5 证明下列结论:

(1) 设

$$U = \{(x, y, 0) \in \mathbb{F}^3 : x, y \in \mathbb{F}\}, \quad W = \{(0, 0, z) \in \mathbb{F}^3 : z \in \mathbb{F}\}$$

则 $\mathbb{F}^3 = U \oplus W$.

(2) 设 U_j 是 \mathbb{F}^n 中除第 j 个坐标以外其余坐标全是 0 的向量所组成的子空间 (例如, $U_2 = \{(0, x, 0, \dots, 0) \in \mathbb{F}^n : x \in \mathbb{F}\}$), 则 $\mathbb{F}^n = U_1 \oplus \dots \oplus U_n$.

(3) 设

$$U_1 = \{(x, y, 0) \in \mathbb{F}^3 : x, y \in \mathbb{F}\}, \quad U_2 = \{(0, 0, z) \in \mathbb{F}^3 : z \in \mathbb{F}\}, \quad U_3 = \{(0, y, y) \in \mathbb{F}^3 : y \in \mathbb{F}\}$$

则 $U_1 + U_2 + U_3$ 不是直和.

每次都要构造一个反例来说明某个和不是直和过于麻烦, 实际上有一种更简易的判别方法:

命题 3.16 直和的判定条件

设 U_1, \dots, U_m 都是 V 的子空间. “ $U_1 + \dots + U_m$ 是直和” 当且仅当 “0 表示成 $u_1 + \dots + u_m$ (其中每个 u_j 都属于 U_j) 的唯一方式是每个 u_j 都等于 0”.

证明 1° 必要性: 由定义可知, 若 $U_1 + \dots + U_m$ 是直和, 则 $\mathbf{0}$ 只有一种表示. 又由 $\mathbf{0} + \dots + \mathbf{0} = \mathbf{0}$ (其中第 j 个 $\mathbf{0}$ 属于 U_j) 可知, 这是唯一的表示方法.

2° 充分性: 设 $U_1 + \dots + U_m$ 中元素 v , 若 v 可以表示为 $u_1 + \dots + u_m$ 或 $v_1 + \dots + v_m$ (其中 $u_j, v_j \in U_j$), 那么 $0 = (u_1 - v_1) + \dots + (u_m - v_m)$, 即 $u_j = v_j$ ($j = 1, \dots, m$), 于是 $U_1 + \dots + U_m$ 是直和.

命题 3.17 两个子空间的直和

设 U 和 W 都是 V 的子空间, 则 $U + W$ 是直和当且仅当 $U \cap W = \{0\}$.

证明 1° 必要性: 设 $v \in (U \cap W)$, 由于 $0 = v + -v$, 由命题??可知, $v = 0$.

2° 充分性: 假设有不为 0 的两个向量 $u \in U, v \in W$, 使得 $0 = u + v$, 那么 $u = -v$. 又因为 $-v \in W$, 可知 $u \in v \in (U \cap W)$, 于是 $u = 0$, 这与假设矛盾.

3.4 有限维向量空间

3.4.1 张成空间

首先介绍线性组合:

定义 3.18 线性组合

对于 V 中的一组向量 v_1, \dots, v_m , 取 $a_1, \dots, a_m \in \mathbb{F}$ 分别与每个元素相乘, 就得到这组向量的线性组

合, 即

$$a_1v_1 + \cdots + a_mv_m$$

容易发现, 一组向量的线性组合也是向量.

注. 在描述一组向量时, 为了避免出现歧义, 通常不用括号括起来. 这就类似于集合的表示中 “|” 与 “:” 的关系一样.

注. 线性组合, 实际上就是用来描述加法封闭性与标量乘法封闭性的. 可以说, 一个对加法、标量乘法封闭的集合中的任意元素都能被由所有元素构成的组的线性组合表示出来.

例 3.6 请判断下列向量是否是 $(2, 1, -3), (1, -2, 4)$ 的线性组合:

$$(17, -4, 2) \quad (17, -4, 5)$$

当这组向量的长度为 2 时, 联系 “平面向量基本定理”, 可知若取平面上的两个基本的不共线向量 \vec{e}_1, \vec{e}_2 , 则平面上任意一个向量都能用这两个向量的线性组合表示. 就像我们会用 “所有满足 $(x-a)^2 + (y-b)^2 = r^2$ 的点构成的集合” 表示一个圆一样, 所有能用这两个向量的线性组合表示的元素构成的集合是什么呢?

定义 3.19 张成空间

V 中的一组向量 v_1, \cdots, v_m 的所有线性组合所构成的集合称为 v_1, \cdots, v_m 的张成空间, 记为 $\text{span}(v_1, \cdots, v_m)$, 即

$$\text{span}(v_1, \cdots, v_m) = \{a_1v_1 + \cdots + a_mv_m : a_1, \cdots, a_m \in \mathbb{F}\}$$

特别地, 定义空组 $()$ 的张成空间为 $\{0\}$.

例 3.7 前面的例子表明在 \mathbb{F}^3 中,

$$(17, -4, 2) \in \text{span}((2, 1, -3), (1, -2, 4))$$

$$(17, -4, 5) \notin \text{span}((2, 1, -3), (1, -2, 4))$$

有了张成空间的定义, 可知上文所述集合就是 \mathbb{R}^2 , 表示为 $\mathbb{R}^2 = \text{span}(\vec{e}_1, \vec{e}_2)$.

命题 3.20 张成空间是包含这组向量的最小子空间

V 中一组向量的张成空间是包含这组向量的最小子空间.

证明 设 V 中向量组 v_1, \cdots, v_m 的张成空间 $\text{span}(v_1, \cdots, v_m)$, 记为 U .

1° 证明 U 是 V 的子空间: 显然 $0 = 0v_1 + \cdots + 0v_m \in U$; 任取 U 中两个元素 $u = a_1v_1 + \cdots + a_mv_m$ 与

$w = b_1v_1 + \cdots + b_mv_m$, 作 $u + w = (a_1 + b_1)v_1 + \cdots + (a_m + b_m)v_m$, 由 V 对加法封闭, 可知 U 也对加法封闭; 取 U 中一个元素 $u = a_1v_1 + \cdots + a_mv_m$ 与标量 $\lambda \in \mathbb{F}$, 由于 $\lambda u = (\lambda a_1)v_1 + \cdots + (\lambda a_m)v_m$, 由 V 对标量乘法封闭, 可知 U 也对标量乘法封闭. 综上, U 是 V 的子空间.

2° 证明 U 包含 v_1, \cdots, v_m : 取 U 中元素 u_j , 令 $u_j = 0v_1 + \cdots + 1v_j + \cdots + 0v_m = v_j$, 于是任意一个 $v_j \in U$.

3° 证明 U 是最小的满足条件的子空间: 假设存在一个更小的 U' , 由于 U' 包含 v_1, \cdots, v_m , 又因为 U' 对加法与标量乘法封闭, 故 U' 中必有 v_1, \cdots, v_m 的所有线性组合, 即 $|U'| \geq |\text{span}(v_1, \cdots, v_m)| = |U|$, 这与假设矛盾. 因此 U 是最小的满足条件的子空间.

定义 3.21 张成

若 $\text{span}(v_1, \cdots, v_m) = V$, 则称 v_1, \cdots, v_m 张成 V .

继续上文的例子. 由于 $\mathbb{R}^2 = \text{span}(\vec{e}_1, \vec{e}_2)$, 可知 \vec{e}_1, \vec{e}_2 张成 \mathbb{R}^2 . 现在, 取 $\vec{e}_1 = (1, 0), \vec{e}_2 = (0, 1)$, 则 \mathbb{R}^2 中的任意一个向量均能表示为 $a\vec{e}_1 + b\vec{e}_2 = (a, b)$ 的形式, 这是一个标准的 Cartesian 坐标系. 那如果 \vec{e}_1, \vec{e}_2 只是两个普通的向量呢? 可以构造出一种“平面非直角非单位长度坐标系”. 总的来说, 不论 \vec{e}_1, \vec{e}_2 如何选取, 它们总能作为两个“基底”张成 \mathbb{R}^2 . 更进一步, \mathbb{R}^2 中所有元素的自由度都是 2(实际上这一点会在后面讲到, 我们称能张成 V 的最小组的长度为 V 的维度).

例 3.8 请证明:

(1) \mathbb{F}^2 上的向量组 $(1, 2), (3, 5)$ 张成 \mathbb{F}^2 .

(2) \mathbb{F}^2 上的向量组 $(1, 2), (3, 5), (6, 7)$ 张成 \mathbb{F}^2 .

(3) \mathbb{F}^n 上的向量组 $(1, 0, \cdots, 0), (0, 1, \cdots, 0), \cdots, (0, 0, \cdots, 1)$ 张成 \mathbb{F}^n . (其中第 j 个向量的第 j 个坐标为 1, 其余都为 0)

(4) 设 v_1, v_2, v_3, v_4 张成 V , 则 $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ 也张成 V .

证明 只选择部分证明:

(1) 任取 \mathbb{F}^2 上的向量 (x, y) , 由于 $(3y - 5x)(1, 2) + (2x - y)(3, 5) = (x, y)$, 可知 $\text{span}((1, 2), (3, 5)) = \mathbb{F}^2$, 即 $(1, 2), (3, 5)$ 张成 \mathbb{F}^2 .

(4) 由于 v_1, v_2, v_3, v_4 张成 V , 任取 V 中元素 v , 设

$$v = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 \quad (a_1, a_2, a_3, a_4 \in \mathbb{F})$$

因为

$$v = a_1(v_1 - v_2) + (a_1 + a_2)(v_2 - v_3) + (a_1 + a_2 + a_3)(v_3 - v_4) + (a_1 + a_2 + a_3 + a_4)v_4$$

且由 \mathbb{F} 对加法封闭, $a_1, a_1 + a_2, a_1 + a_2 + a_3, a_1 + a_2 + a_3 + a_4 \in \mathbb{F}$, 可知 $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ 也张成 V .

注. 从第二个例子可以看出, 张成向量空间的组的长度不一定与 \mathbb{R}^2 的维度相等.

3.4.2 有限维向量空间

现在我们给出线性代数中的一个关键定义：

定义 3.22 有限维向量空间, 无限维向量空间

- 如果一个向量空间可以由该空间中的某个向量组张成, 则称这个向量空间是有限维的.
- 相对应地, 如果一个向量空间不是有限维的, 则称这个向量空间是无限维的. 也就是说, 如果一个向量空间不能由该空间中的任何向量组张成, 它就是无限维的.

联系上一个例子中的第三条, 由于 \mathbb{F}^n 总能被这样一个向量组张成, 它是有限维的. “维度”这个概念会在后面详细介绍, 现在只是定性分析.

现在介绍一个具体的例子:

定义 3.23 多项式, 多项式的次数

- 对于函数 $p: \mathbb{F} \rightarrow \mathbb{F}$, 若对任意 $z \in \mathbb{F}$ 均存在 $a_0, \dots, a_m \in \mathbb{F}$ 使得

$$p(z) = a_0 + a_1 z + a_2 z^2 + \dots + a_m z^m$$

则称 p 是系数属于 \mathbb{F} 的多项式.

- 特别地, 对于上式, 当要求 $a_m \neq 0$ 时, 称 p 的次数为 m , 记为 $\deg p = m$. 规定恒等于 0 的多项式的次数为 $-\infty$.
- 定义 $\mathcal{P}(\mathbb{F})$ 是系数属于 \mathbb{F} 的所有多项式构成的集合.
- 对于非负整数 m , 定义 $\mathcal{P}_m(\mathbb{F})$ 表示系数在 \mathbb{F} 中且次数不超过 m 的所有多项式构成的集合. (约定 $-\infty < m$).

例 3.9 请证明:

(1) 对每个非负整数 m , $\mathcal{P}_m(\mathbb{F})$ 是有限维向量空间.

(2) $\mathcal{P}(\mathbb{F})$ 是无限维向量空间.

证明 (1) 由于 $\mathcal{P}_m(\mathbb{F}) = \text{span}(1, z, \dots, z^m)$, 可知 $\mathcal{P}_m(\mathbb{F})$ 是有限维向量空间.

(2) 假设 $\mathcal{P}(\mathbb{F})$ 中的一组多项式可以张成 $\mathcal{P}(\mathbb{F})$, 记这组多项式中次数最高的多项式的次数为 m , 那么总能找到 z^{m+1} 不属于该张成空间, 这与假设矛盾. 故不存在任何一组多项式可以张成 $\mathcal{P}(\mathbb{F})$, 即 $\mathcal{P}(\mathbb{F})$ 是无限维向量空间.

3.5 线性无关

3.5.1 线性无关

与子空间的和一样, 我们倾向于研究那些有唯一表示形式的元素.

定义 3.24 线性无关, 线性相关

- V 中的一组向量 v_1, \dots, v_m 称为线性无关, 如果 $\text{span}(v_1, \dots, v_m)$ 中每个向量可以唯一地表示成 v_1, \dots, v_m 的线性组合. 规定空组 $()$ 是线性无关的.
- 相对应地, 如果一组向量不是线性无关的, 则称这组向量线性相关. 也就是说, 对于这组向量的张成空间, 如果其中存在向量有不唯一的表示, 它就是线性相关的.

命题 3.25 线性相关性的判定

- V 中一组向量 v_1, \dots, v_m 线性无关当且仅当使得 $a_1v_1 + \dots + a_mv_m = 0$ 成立的 $a_1, \dots, a_m \in \mathbb{F}$ 只有 $a_1 = \dots = a_m = 0$.
- 由线性相关的定义可知, V 中一组向量 v_1, \dots, v_m 线性相关当且仅当存在不全为 0 的 $a_1, \dots, a_m \in \mathbb{F}$ 使得 $a_1v_1 + \dots + a_mv_m = 0$ 成立.

证明 1° 充分性: 假设 $v \in \text{span}(v_1, \dots, v_m)$ 有两种不同的线性组合表示, 即

$$v = a_1v_1 + \dots + a_mv_m \quad v = b_1v_1 + \dots + b_mv_m$$

两式相减, 得到 $0 = (a_1 - b_1)v_1 + \dots + (a_m - b_m)v_m$. 由所给条件, 知 $a_j = b_j$ ($j = 1, \dots, m$), 这与假设矛盾, 于是 v 只有一种表示方法, 即 v_1, \dots, v_m 线性无关.

2° 必要性: 首先, 若令 a_1, \dots, a_m 全为 0, 则有 $0 = 0v_1 + \dots + 0v_m$, 这是 0 的一种表示形式; 其次, 由于 v_1, \dots, v_m 线性无关, 0 只有一种表示形式. 综上, 0 的唯一表示形式就是 $a_1 = \dots = a_m = 0$.

例 3.10 线性无关的判断 请证明:

- (1) V 中一个向量 v 所构成的向量组是线性无关的当且仅当 $v \neq 0$.
- (2) V 中两个向量构成的向量组线性无关当且仅当每个向量都不能写成另一个向量的标量倍.
- (3) 对每个非负整数 m , $\mathcal{P}(\mathbb{F})$ 中的组 $1, z, \dots, z^m$ 线性无关.
- (4) 设 v_1, v_2, v_3, v_4 在 V 中是线性无关的, 则 $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ 也是线性无关的.
- (5) 设 v_1, \dots, v_m 在 V 中线性无关, 并设 $w \in V$. 证明: 若 $v_1 + w, \dots, v_m + w$ 线性相关, 则 $w \in \text{span}(v_1, \dots, v_m)$.

证明 (1) 分别证明充分性和必要性的逆否命题成立, 即证明“ v 构成的向量组线性相关当且仅当 $v = 0$ ”. 充分性: 当 $v = 0$, 设 $av = 0$ ($a \in F$), 则存在不为 0 的 a ; 必要性: 设存在不为 0 的 $a \in F$ 使 $av = 0$, 则 $v = 0$.

(2) 设向量 $u, v \in V$. 分别证明充分性和必要性的逆否命题成立, 即证明“ u, v 线性相关当且仅当每个向量可以写成另一个向量的标量倍”. 充分性: 记 $u = \lambda v$ ($\lambda \neq 0$), 则存在不全为零的 $a_1, a_2 \in \mathbb{F}$ 满足 $a_1 + a_2\lambda = 0$ 使 $a_1v + a_2u = 0$ 成立; 必要性: 设存在不全为零的 $a_1, a_2 \in \mathbb{F}$ 使 $a_1v + a_2u = 0$ 成立, 不妨设 $a_2 \neq 0$, 则 $u = -\frac{a_1}{a_2}v$.

(3) 首先不加证明地阐释一个引理: 若一个多项式是零函数, 则其所有系数均为 0 (会在第四章进行证明). 于是, 对于 $p(z) = a_0 + a_1z + \cdots + a_mz^m = 0$, 必然有 $a_0 = a_1 = \cdots = a_m$, 即 $1, z, \cdots, z^m$ 线性无关.

(4) 由 v_1, v_2, v_3, v_4 线性无关, 设

$$a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 = 0 \quad (3.1)$$

其中 $a_1, a_2, a_3, a_4 \in \mathbb{F}$. 则必有 $a_1 = a_2 = a_3 = a_4 = 0$. 对式 3.1 进行变形, 得到

$$a_1(v_1 - v_2) + (a_1 + a_2)(v_2 - v_3) + (a_1 + a_2 + a_3)(v_3 - v_4) + (a_1 + a_2 + a_3 + a_4)v_4 = 0$$

由上可得此时 $a_1 = a_1 + a_2 = a_1 + a_2 + a_3 = a_1 + a_2 + a_3 + a_4 = 0$, 即 $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ 线性无关.

(5) 设不全为 0 的 $c_1, \cdots, c_m \in \mathbb{F}$ 满足

$$c_1(v_1 + w) + \cdots + c_m(v_m + w) = 0$$

即

$$c_1v_1 + \cdots + c_mv_m = -(c_1 + \cdots + c_m)w$$

由 v_1, \cdots, v_m 线性无关, 左式一定不为 0. 当 $w = 0$ 时, 必然可以表示为 v_1, \cdots, v_m 的线性组合形式, 命题成立; 当 $w \neq 0$ 时, $c_1 + \cdots + c_m \neq 0$, 故

$$w = \frac{-c_1}{c_1 + \cdots + c_m}v_1 + \cdots + \frac{-c_m}{c_1 + \cdots + c_m}v_m$$

则命题成立.

例 3.11 线性相关的判断 请证明:

(1) \mathbb{F}^3 中的向量组 $(2, 3, 1), (1, -1, 2), (7, 3, 8)$ 线性相关.

(2) \mathbb{F}^3 中的向量组 $(2, 3, 1), (1, -1, 2), (7, 3, c)$ 线性相关当且仅当 $c = 8$.

(3) 包含 0 向量的向量组线性相关.

证明 (1) 设 $a_1, a_2, a_3 \in \mathbb{F}$ 满足

$$a_1(2, 3, 1) + a_2(1, -1, 2) + a_3(7, 3, 8) = (0, 0, 0)$$

即

$$(2a_1 + a_2 + 7a_3, 3a_1 - a_2 + 3a_3, a_1 + 2a_2 + 8a_3) = (0, 0, 0)$$

可得

$$\begin{cases} 2a_1 + a_2 + 7a_3 = 0 \\ 3a_1 - a_2 + 3a_3 = 0 \\ a_1 + 2a_2 + 8a_3 = 0 \end{cases}$$

化简之, 得到任意满足 $a_1 = -2a_3, a_2 = -3a_3$ 的 (a_1, a_2, a_3) 均符合要求, 则存在一组不全为 0 的 (a_1, a_2, a_3) 满足上式, 即 $(2, 3, 1), (1, -1, 2), (7, 3, 8)$ 线性相关.

(实际上, 将方程组中的第一个式子乘以 $\frac{7}{5}$ 再与第二个式子乘以 $-\frac{3}{5}$ 相加, 就得到了第三个式子. 也就是说, 这三个式子中有一个式子是多余的, 自然可以解出不全为 0 的 a_1, a_2, a_3 .)

(2) 充分性同上, 下证必要性: 设 $a_1, a_2, a_3 \in \mathbb{F}$ 满足

$$a_1(2, 3, 1) + a_2(1, -1, 2) + a_3(7, 3, c) = (0, 0, 0)$$

即

$$\begin{cases} 2a_1 + a_2 + 7a_3 = 0 \\ 3a_1 - a_2 + 3a_3 = 0 \\ a_1 + 2a_2 + ca_3 = 0 \end{cases}$$

容易发现, $(a_1, a_2, a_3) = (0, 0, 0)$ 是方程组的一组解. 要得到另一组不同的解, 要求有效方程的个数严格小于变量个数, 即其中一个方程可以表示为另两个的线性组合形式, 记

$$a_1 + 2a_2 + ca_3 = x(2a_1 + a_2 + 7a_3) + y(3a_1 - a_2 + 3a_3) \quad (x, y \in \mathbb{F})$$

化简之, 即

$$(2x + 3y - 1)a_1 + (x - y - 2)a_2 + (7x + 3y - c)a_3 = 0$$

对任意 a_1, a_2, a_3 均成立, 即 $2x + 3y - 1 = x - y - 2 = 7x + 3y - c = 0$, 解得 $x = \frac{7}{5}, y = -\frac{3}{5}$, 于是 $c = 8$.

线性相关与下列定义等价:

命题 3.26 线性相关的第二定义

V 中一组向量 v_1, \dots, v_m 线性相关当且仅当其中存在一个向量能表示为其余向量的线性组合形式.

证明 1° 充分性: 设该向量 v 能表示为 v_1, \dots, v_m 的线性组合形式, 即

$$v = a_1v_1 + \dots + a_mv_m$$

那么 $0 = a_1v_1 + \dots + a_mv_m + (-1)v$. 其中 -1 显然不为 0, 因此 v_1, \dots, v_m, v 线性相关.

2° 必要性: 设 $0 = a_1v_1 + \cdots + a_mv_m$. 不妨令 $a_j \neq 0$, 那么有

$$v_j = \frac{a_1}{-a_j}v_1 + \cdots + \frac{a_m}{-a_j}v_m$$

这说明 v_j 可以表示为其余元素的线性组合.

注. 在该证明过程中, 不难发现定义里“其余”的重要性.

实际上, 利用这个定义更好理解线性相关的本质. 上一小节的例题告诉我们, 张成组 (即张成某向量空间的向量组) 的长度可以不同. 容易证明, 第一个例子中的向量组是线性无关的, 而第二个例子中的向量组是线性相关的. 实际上, 像这样既是张成组又是线性无关的组, 就称为基 (详细内容在下一小节会讲到).

3.5.2 线性相关性与张成

下面的命题为我们阐释了线性相关性与张成的一个基本关系.

命题 3.27 线性相关性引理

设 v_1, \cdots, v_m 是 V 中的一个线性相关的向量组, 则存在 $j \in \{1, 2, \cdots, m\}$ 使得:

(a) $v_j \in \text{span}(v_1, \cdots, v_{j-1})$;

(b) 若从 v_1, \cdots, v_m 中去掉第 j 项, 则剩余组的张成空间等于 $\text{span}(v_1, \cdots, v_m)$.

证明 (a) 由于 v_1, \cdots, v_m 线性相关, 存在不全为 0 的数 $a_1, \cdots, a_m \in \mathbb{F}$ 使得 $a_1v_1 + \cdots + a_mv_m = 0$. (人为地) 设该向量组的顺序满足 a_1, \cdots, a_j 均不为 0, 从而有

$$v_j = \frac{a_1}{-a_j}v_1 + \cdots + \frac{a_{j-1}}{-a_j}v_{j-1} \quad (3.2)$$

这意味着 $v_j \in \text{span}(v_1, \cdots, v_{j-1})$;

(b) 取 $\text{span}(v_1, \cdots, v_m)$ 中某一元素 u , 设 $u = b_1v_1 + \cdots + b_mv_m$, 将式 3.2 代入可得

$$u = \left(\frac{a_1b_j}{-a_j} + b_1 \right) v_1 + \cdots + \left(\frac{a_{j-1}b_j}{-a_j} + b_{j-1} \right) v_{j-1} + b_{j+1}v_{j+1} + \cdots + b_mv_m$$

这表明对于 $\text{span}(v_1, \cdots, v_m)$ 中任一元素, 它都在 $\text{span}(v_1, \cdots, v_{j-1}, v_{j+1}, \cdots, v_m)$ 中, 即原命题所述.

由线性无关与张成的几何意义, 我们能够想象: 对于任意一个有限维向量空间, 总是存在一组“基底”, 这组基底可以线性表示任何向量空间中的元素, 并且它们之间互不多余、缺一不可. 这就类似于欧氏几何中的五条公理一样. 通过这种直观的理解, 不难得出以下命题, 难的在于如何规整地证明.

命题 3.28 线性无关组与张成组长度的关系

xxwgv 在有限维向量空间 V 中, 线性无关组的长度总是小于等于向量空间的每一个张成组的长度.

证明 设 V 中一个线性无关组 u_1, \cdots, u_m 与张成组 w_1, \cdots, w_n .

1° 第 1 步: 将线性无关组中的第 1 个元素 u_1 添加在张成组的开头, 便形成组

$$u_1, w_1, \dots, w_n$$

由线性相关性引理, 我们可以去掉某个 w 使得新的组仍张成 V .

2° 第 j 步: 将线性无关组中的第 j 个元素 u_j 添加在 u_{j-1} 后, 由线性相关性引理, 又因为 u_1, \dots, u_j 是线性无关的, 我们可以去掉某个 w 使得新的组仍张成 V .

每经过一步, 都会将组中的一个 w 换成一个 u . 因为在第 m 步后把所有的 u 都换完, 可知 $n \geq m$, 即原命题所述.

利用这一“直观”的结论, 我们可以“直观”地证伪某些命题.

例 3.12 证明下列结论:

(1) 组 $(1, 2, 3), (4, 5, 8), (9, 6, 7), (-3, 2, 8)$ 在 \mathbb{R}^3 中一定不是线性无关的.

(2) 组 $(1, 2, 3, -5), (4, 5, 8, 3), (9, 6, 7, -1)$ 一定不能张成 \mathbb{R}^4 .

利用命题??的证明思路, 还可以说明更多直观的结论:

命题 3.29 向量空间中的一些结论

yixpjply

- 在向量空间 V 中, 每个线性相关的张成组都能通过去除某些元素得到一个线性无关的张成组.
- V 是无限维向量空间当且仅当 V 中存在一个向量序列 v_1, v_2, \dots 使得当 m 是任意正整数时 v_1, \dots, v_m 都是线性无关的.
- 有限维向量空间的子空间都是有限维的.

证明 (1)1° 第 1 步: 设 $\mathcal{W}_1 = v_1, \dots, v_m$ 张成 V . 若 $v_1 \notin \text{span}(v_2, \dots, v_m)$, 则保持该组不变, 并停止操作; 若 $v_1 \in \text{span}(v_2, \dots, v_m)$, 则去掉 v_1 , 并记新组 v_2, \dots, v_m 为 \mathcal{W}_2 .

2° 第 j 步: 若 $v_j \notin \text{span}(v_{j+1}, \dots, v_m)$, 则保持 \mathcal{W}_j 不变, 并停止操作; 若 $v_j \in \text{span}(v_{j+1}, \dots, v_m)$, 则去掉 v_j , 并记新组 v_{j+1}, \dots, v_m 为 \mathcal{W}_{j+1} . 在经过有限次操作后, 一定会在某一步停止并返回一个线性无关的组, 且能张成 V .

(2) 充分性显然. 下证必要性:

1° 第 1 步: 取 V 中的一个线性无关向量组 \mathcal{W}_1 , 作它的张成空间 U_1 , 取一元素 $u \in (V \setminus U_1)$ 放入该组, 得到一个新的组 \mathcal{W}_2 . 显然该组仍是线性无关的 (因为线性相关性的第二定义).

2° 第 j 步: 作 \mathcal{W}_j 的张成空间 U_j , 取一元素 $u \in (V \setminus U_j)$ 放入 \mathcal{W}_j , 得到一个新的组 \mathcal{W}_{j+1} .

由于可以不断重复该过程, 因此这样一个组 \mathcal{W}_j 会不断扩张并保持线性无关, 即符合原命题要求.

(3) 设有限维向量空间 V 及其子空间 U . 由命题??, V 中任意一个线性无关组的长度小于等于每一个 V 的张成组的长度, 故该线性无关组的长度是有限的.

1° 第 1 步: 若 $U = \{0\}$, 即 $U = \text{span}()$, 则 U 符合要求; 否则存在非零向量 $v_1 \in U$.

2° 第 j 步: 若 $U = \text{span}(v_1, \dots, v_{j-1})$, 则 U 符合要求; 否则存在 $v_j \in U$ 满足 $v_j \notin \text{span}(v_1, \dots, v_{j-1})$.

由于第 $j+1$ 步能够说明存在 $v_1, \dots, v_j \in U$ 使 v_1, \dots, v_j 线性无关, 而 U 中任意一个线性无关组的长度是有限的, 故一定会在某一步停止, 此时 U 即符合要求.

3.6 基与维数

3.6.1 基

上一节中多次出现“基底”这一关键词，现在我们来集中研究它：

定义 3.30 基

若 V 中的一个向量组既线性无关又张成 V ，则称为 V 的基。

例 3.13 基的例子 请验证：

- (1) 组 $(1, 0, \dots, 0), (0, 1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ 是 \mathbb{F}^n 的基.(实际上, 这称为 \mathbb{F}^n 的标准基)
- (2) 组 $(1, 1, 0), (0, 0, 1)$ 是 $\{(x, x, y) \in \mathbb{F}^3 : x, y \in \mathbb{F}\}$ 的基.
- (3) 组 $(1, -1, 0), (1, 0, -1)$ 是 $\{(x, y, z) \in \mathbb{F}^3 : x + y + z = 0\}$ 的基.
- (4) 组 $1, z, \dots, z^m$ 是 $\mathcal{P}_m(\mathbb{F})$ 的基.

我们发现张成和线性无关的定义十分类似：都出现了“线性组合”这一形式. 将它们综合起来，就是基的判定命题：

命题 3.31 基的判定

V 中的向量组 v_1, \dots, v_m 是 V 的基当且仅当每个 $v \in V$ 都能唯一地写成以下形式

$$v = a_1 v_1 + \dots + a_m v_m$$

其中 $a_1, \dots, a_m \in \mathbb{F}$.

证明 必要性显然. 直接来看充分性：在 V 中任取一元素 v ，设它可以唯一地表示为 $v = a_1 v_1 + \dots + a_m v_m$ 的形式.

1° 张成：由张成的定义可知， v_1, \dots, v_m 张成 V .

2° 线性无关：令 a_1, \dots, a_m 全为 0，则有 $0 = 0v_1 + \dots + 0v_m$ ，这是 0 的唯一表示形式，因此 v_1, \dots, v_m 线性无关.

回顾上一节中命题??的第一条，实际上现在我们就能将其写成基的形式.

命题 3.32 基、线性无关组、张成组 I

在有限维向量空间 V 中，

- 每个张成组都可以化简成 V 的一个基.
- 每个线性无关的向量组都可以扩充成 V 的一个基.

证明 第一条已经在命题??中证明过，这里只证明第二条.

设 V 中的线性无关组 v_1, \dots, v_m 与一个基 u_1, \dots, u_n . 作组 $\mathcal{W} = v_1, \dots, v_m, u_1, \dots, u_n$, 它显然张成 V . 由命题??可知 $m \leq n$. 利用第一条的证明过程将 \mathcal{W} 化简为 $v_1, \dots, v_m, u_1, \dots, u_j$, 可知它是 V 的一个基.

以上的命题具有很强的可操作性. 例如, 取 \mathbb{F}^3 的基的一部分 (也就是一个线性无关组) $(1, 1, 4), (5, 1, 4)$, 再取一个标准基 $(1, 0, 0), (0, 1, 0), (0, 0, 1)$. 在 $(1, 1, 4), (5, 1, 4), (1, 0, 0), (0, 1, 0), (0, 0, 1)$ 中去掉

$$(1, 0, 0) = -\frac{1}{4}(1, 1, 4) + \frac{1}{4}(5, 1, 4)$$

在 $(1, 1, 4), (5, 1, 4), (0, 1, 0), (0, 0, 1)$ 中, 由于 $(0, 1, 0), (0, 0, 1)$ 都不能被 $(1, 1, 4), (5, 1, 4)$ 线性表示, 所以最后可以保留其中任意一个, 即 $(1, 1, 4), (5, 1, 4), (0, 1, 0)$ 和 $(1, 1, 4), (5, 1, 4), (0, 0, 1)$ 都是 \mathbb{F}^3 的基.

有了基这个工具之后, 我们可以证明更多之前不能证明的结论:

命题 3.33 子空间与直和的关系

设 V 是有限维的, U 是 V 的子空间, 则存在 V 的子空间 W 使得 $V = U \oplus W$.

证明 设 U 的一个基 u_1, \dots, u_m , 按照命题??的方法将这个 V 中的线性无关组扩充为 V 的基, 记为 $u_1, \dots, u_m, w_1, \dots, w_n$. 取 $W = \text{span}(w_1, \dots, w_n)$, 下证这样的 W 就是满足题目要求的子空间: 显然 $U + W = V$. 任取 $v \in (U \cap W)$, 设

$$v = a_1 u_1 + \dots + a_m u_m \quad v = b_1 w_1 + \dots + b_n w_n$$

两式作差, 得 $0 = a_1 u_1 + \dots + a_m u_m + (-b_1) w_1 + \dots + (-b_n) w_n$, 由 $u_1, \dots, u_m, w_1, \dots, w_n$ 是 V 的基可得 $a_1 = \dots = b_n = 0$, 则 $v = 0$. 由命题??知 $V = U \oplus W$.

3.6.2 维数

继续研究向量空间的几何意义, 我们发现基已经被定义了, 但是最小的基还不太清楚. 实际上, 容易说明所有基的长度都是相等的, 而任意一个基的长度就称作维数.

命题 3.34 基的长度不依赖于基的选取

有限维向量空间的任意两个基的长度都相同.

证明 由命题??可知, 因为任意两个基 \mathcal{U}, \mathcal{V} 都同时是线性无关向量组与张成组, 所以 \mathcal{U} 的长度小于等于 \mathcal{V} 的长度、 \mathcal{U} 的长度大于等于 \mathcal{V} 的长度, 于是它们的长度相等.

定义 3.35 维数

有限维向量空间 V 的任意基的长度称为这个向量空间的维数, 记作 $\dim V$.

很明显, 一个有限维向量空间的子空间也是有限维的, 它的维数应当满足下列命题要求:

命题 3.36 子空间的维数

若 U 是有限维向量空间 V 的子空间, 则 $\dim U \leq \dim V$.

证明 取 U 的一个基 \mathcal{U} , 取 V 的一个基 \mathcal{V} . 因为 \mathcal{U} 是 V 的一个线性无关的子空间, 由命题??可知, \mathcal{U} 的长度小于等于 \mathcal{V} 的长度, 即 $\dim U \leq \dim V$.

借助维数, 可以更快捷地证明一个向量空间的基.

命题 3.37 基、线性无关组、张成组 II

设 V 是有限维向量空间.

- V 中每个长度为 $\dim V$ 的线性无关向量组都是 V 的基.
- V 中每个长度为 $\dim V$ 的张成组都是 V 的基.

证明 (1) 设 V 中的一个线性无关向量组 v_1, \dots, v_m , 取 V 中一个基 w_1, \dots, w_n , 由命题??可知, v_1, \dots, v_m 可以扩充为基, 然而在此过程中由于 $m = n$, 实际上没有发生任何扩充, 故 v_1, \dots, v_m 本来就是一个基.

(2) 证明过程同 (1), 留作习题.

例 3.14 证明以下结论:

- (1) 组 $(5, 7), (4, 3)$ 是 \mathbb{F}^2 的基.
- (2) 证明 $1, (x-5)^2, (x-5)^3$ 是 $\mathcal{P}_3(\mathbb{R})$ 的子空间 U 的一个基, 其中 U 定义为

$$U = \{p \in \mathcal{P}_3(\mathbb{R}) : p'(5) = 0\}$$

类比并集的元素个数计算公式 (即容斥原理), 子空间和的维数计算公式如下:

命题 3.38 子空间和的维数

如果 U_1 和 U_2 是有限维向量空间的两个子空间, 则

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$$

证明 设 $U_1 \cap U_2$ 的基 v_1, \dots, v_m , 则 $v_1, \dots, v_m \in U_1, U_2$; 设 U_1, U_2 的基分别为 $v_1, \dots, v_m, u_1, \dots, u_j$ 与 $v_1, \dots, v_m, u'_1, \dots, u'_k$.

取 $U_1 + U_2$ 中的 0 元素, 它能被表示为 $w_1 + w_2$ 的形式 (其中 $w_1 \in U_1, w_2 \in U_2$). 因此设

$$\begin{aligned} 0 &= w_1 + w_2 = (a_1 v_1 + \dots + a_m v_m + b_1 u_1 + \dots + b_j u_j) + (a'_1 v_1 + \dots + a'_m v_m + b'_1 u'_1 + \dots + b'_k u'_k) \\ &= (a_1 + a'_1) v_1 + \dots + (a_m + a'_m) v_m + b_1 u_1 + \dots + b_j u_j + b'_1 u'_1 + \dots + b'_k u'_k \end{aligned}$$

于是

$$(-a_1 - a'_1)v_1 + \cdots + (-a_m - a'_m)v_m + (-b_1)u_1 + \cdots + (-b_j)u_j = b'_1u'_1 + \cdots + b'_ku'_k$$

等式左边属于 U_1 , 等式右边属于 U_2 , 因此 $b'_1u'_1 + \cdots + b'_ku'_k \in U_1 \cap U_2$. 设

$$b'_1u'_1 + \cdots + b'_ku'_k = c_1v_1 + \cdots + c_mv_m$$

于是

$$b'_1u'_1 + \cdots + b'_ku'_k + (-c_1)v_1 + \cdots + (-c_m)v_m = 0$$

这个式子告诉我们 $b'_1 = \cdots = b'_k = c_1 = \cdots = c_m = 0$. 同理可得 $b_1 = \cdots = b_j = 0$. 带入上式, 可得

$$(a_1 + a'_1)v_1 + \cdots + (a_m + a'_m)v_m = 0$$

所以 $a_1 + a'_1 = \cdots = a_m + a'_m = 0$.

综上, 对于 $U_1 + U_2$ 中的 0, 它只有唯一一种线性表示形式, 即满足

$$a_1 + a'_1 = \cdots = a_m + a'_m = b'_1 = \cdots = b'_k = b_1 = \cdots = b_j$$

故 $v_1, \cdots, v_m, u_1, \cdots, u_j, u'_1, \cdots, u'_k$ 是 $U_1 + U_2$ 的基. 于是 $\dim(U_1 + U_2) = m + j + k = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2)$.

注. 注意子空间的维数计算公式不一定能推广到更多元的情况, 例如本节习题中所示.

Chapter 4

线性映射与矩阵

4.1 基本概念

4.1.1 向量空间的线性映射

在描述一种颜色时, 我们会给出几个不同方面的数据进行叠加. 以 RGB 颜色为例, 其通过对红 (R), 绿 (G), 蓝 (B) 三个通道分别赋予 $0 \sim 255$ 的值再进行组合而产生出 256^3 种颜色. 设颜色向量 c_1, c_2, c_3 . 如果想将 c_1, c_2 叠加, 我们会考虑一种加权平均, 也即新的颜色 $c_4 = tc_1 + (1-t)c_2, 0 \leq t \leq 1$. 如果考虑 c_5 为另一种方式的叠加, 即 $c_5 = sc_1 + (1-s)c_2, 0 \leq s \leq 1$, 再定义两种颜色的直接和 $\tilde{+}$ 为其算术平均, 那么会有 $c_4 \tilde{+} c_5 = (t \tilde{+} s)c_1 + ((1-t) \tilde{+} (1-s))c_2$. 将 t 与 s 视作映射 C 的自变量, 即可得到 $C(t \tilde{+} s) = C(t) \tilde{+} C(s)$. 同样地, 如果定义标量乘法 $\tilde{\times}$ 表示将白色 $(0, 0, 0)$ 与该颜色按量 λ 混合, 可以得到 $C(\lambda \tilde{\times} t) = \lambda \tilde{\times} C(t)$.

上方的例子解释了所谓叠加原理: 即和的输入所得输出等于输入所得输出的和, 标量倍输入所得输出等于输入所得输出的标量倍. 由此引入线性映射的概念.

定义 4.1 线性映射

设函数 $T: V \rightarrow W$, 若该函数满足下列性质:

1. 加性:

$$\forall u, v \in V, T(u+v) = Tu + Tv.$$

2. 齐性:

$$\forall \lambda \in \mathbb{F}, v \in V, T(\lambda v) = \lambda Tv.$$

则称 T 是线性映射.

注. 为了让你想起线性映射可以直接写成矩阵乘法 (见 3.1.3 节), 通常用 Tv 代替 $T(v)$.

规定记号 $\mathcal{L}(V, W)$, 表示所有从 V 到 W 的线性映射构成的集合.

另一种理解线性映射的方式: 记 T 的图为 $\{(v, Tv) \in V \times W : v \in V\}$, 则 T 是线性映射当且仅当 T 的图是 $V \times W$ 的子空间. (这里, \times 表示笛卡尔积)

例 4.1 线性映射的例子 (1) 零函数: 定义 $0 \in \mathcal{L}(V, W)$ 如下

$$\forall v \in V, 0v = 0.$$

(2) 恒等映射: 定义 $I \in \mathcal{L}(V, V)$ 如下

$$\forall v \in V, Iv = v.$$

(3) 微分: 定义 $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$ 如下

$$Dp = p'.$$

(4) 乘以 x^2 : 定义 $T \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$ 满足

$$\forall x \in \mathbb{R}, (Tp)(x) = x^2 p(x).$$

(5) 向后移位: 定义 $T \in \mathcal{L}(\mathbb{F}^\infty, \mathbb{F}^\infty)$ 如下

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

(6) 从 \mathbb{R}^3 到 \mathbb{R}^2 : 定义 $T \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ 如下

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z).$$

(7) 从 \mathbb{F}^n 到 \mathbb{F}^m : 定义 $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ 如下

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \dots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \dots + A_{m,n}x_n).$$

其中, $A_{j,k} \in \mathbb{F}$, $j = 1, \dots, m$, $k = 1, \dots, n$. 事实上, 从 \mathbb{F}^n 到 \mathbb{F}^m 的每个线性映射都是这种形式的. 我们稍后会进行证明.

观察线性映射所具有的加性与齐性, 似乎可以将其与线性组合联系起来. 例如, 对于 $T \in \mathcal{L}(V, W)$, 若给定 v_1, \dots, v_m 是 V 的基, 设 c_1, \dots, c_m 是 \mathbb{F} 中的任意元素, 则

$$T(c_1v_1 + \dots + c_mv_m) = c_1Tv_1 + \dots + c_mTv_m.$$

由这种想法, 以下的命题不难证明:

命题 4.2 线性映射与定义域的基

对于 $T \in \mathcal{L}(V, W)$, 设 v_1, \dots, v_m 是 V 的基, $w_1, \dots, w_m \in W$, 则存在唯一一个线性映射 $T \in \mathcal{L}(V, W)$ 使得对任意 $j = 1, \dots, m$ 都有

$$Tv_j = w_j.$$

证明 1° 存在性: 利用先前得到的想法, 构造 $T: V \rightarrow W$ 如下:

$$\forall c_j \in \mathbb{F} (j = 1, \dots, m), T(c_1 v_1 + \dots + c_m v_m) = c_1 w_1 + \dots + c_m w_m.$$

由于 v_1, \dots, v_m 是 V 的基, T 的定义域是 V . 下证 $T \in \mathcal{L}(V, W)$.

取 $u = a_1 v_1 + \dots + a_m v_m$, $v = c_1 v_1 + \dots + c_m v_m$, $\lambda \in \mathbb{F}$, 因为

$$\begin{aligned} T(u+v) &= T((a_1+c_1)v_1 + \dots + (a_m+c_m)v_m) \\ &= (a_1+c_1)w_1 + \dots + (a_m+c_m)w_m, \\ Tu+Tv &= a_1 w_1 + \dots + a_m w_m + c_1 w_1 + \dots + c_m w_m. \end{aligned}$$

所以 $T(u+v) = Tu+Tv$, 即 T 满足加性. 因为

$$\begin{aligned} T(\lambda u) &= T((\lambda a_1)v_1 + \dots + (\lambda a_m)v_m) \\ &= \lambda a_1 w_1 + \dots + \lambda a_m w_m, \\ \lambda T(u) &= \lambda(a_1 w_1 + \dots + a_m w_m). \end{aligned}$$

所以 $T(\lambda u) = \lambda T(u)$, 即 T 满足齐性. 综上, 这样的 $T \in \mathcal{L}(V, W)$.

2° 唯一性: 若 $T \in \mathcal{L}(V, W)$, 记 $Tv_j = w_j$ ($j = 1, \dots, m$), 设 c_1, \dots, c_m 是 \mathbb{F} 中的任意元素, 有

$$T(c_1 v_1 + \dots + c_m v_m) = c_1 w_1 + \dots + c_m w_m.$$

这表明任何一个满足题目要求的映射 T 都满足上述形式要求. 结合 1°, 可知映射 T 的唯一形式即为该形式.

对上方的命题稍加延伸, 不难证明:

命题 4.3

设 $v_1, \dots, v_n \in V, T \in \mathcal{L}(V, W)$.

- (1) v_1, \dots, v_n 线性相关, 则 Tv_1, \dots, Tv_n 线性相关;
- (2) v_1, \dots, v_n 线性无关且满足 $Tv = 0$ 的 v 只有 0 (即 T 的零空间为 $\{0\}$), 则 Tv_1, \dots, Tv_n 线性无关.

例 4.2 设 $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. 证明存在标量 $A_{j,k} \in \mathbb{F}$ (其中 $j = 1, \dots, m, k = 1, \dots, n$) 使得对任意 $(x_1, \dots, x_n) \in \mathbb{F}^n$ 都有

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \dots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \dots + A_{m,n}x_n).$$

证明 记 $T(e_i) = (A_{1i}, \dots, A_{mi})$, 则 $T(x_1, \dots, x_n) = x_1 Te_1 + \dots + x_n Te_n = (A_{1,1}x_1 + \dots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \dots + A_{m,n}x_n)$.

例 4.3 将二维平面 \mathbb{R}^2 上的绕原点逆时针旋转 θ 看做映射 $\varphi_\theta : (x, y) \mapsto (x_1, y_1)$. 容易验证 φ_θ 是一个线性映射. 由于

$$\varphi_\theta(1, 0) = (\cos \theta, \sin \theta), \varphi_\theta(1, 0) = (\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2})),$$

我们可以得到 $\varphi_\theta(x, y) = (x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta)$.

我们接着完善线性映射的定义. 先在 $\mathcal{L}(V, W)$ 上定义加法和标量乘法:

定义 4.4 $\mathcal{L}(V, W)$ 上的加法和标量乘法

设 $S, T \in \mathcal{L}(V, W)$, $\lambda \in \mathbb{F}$.

(1) 定义和 $S + T$ 是 V 到 W 的映射, 满足

$$\forall u \in V, (S + T)(u) = Su + Tu.$$

(2) 定义积 λT 是 V 到 W 的映射, 满足

$$\forall u \in V, (\lambda T)(u) = \lambda(Tu).$$

注. 实际上, 定义中的 $S + T, \lambda T$ 均为从 V 到 W 的线性映射, 也即上述定义的加法、标量乘法是封闭的. 更进一步, $\mathcal{L}(V, W)$ 就是一个向量空间. 这个命题易于证明.

一般来说, 向量空间中元素之间的乘法是没有意义的. 然而对于线性映射, 我们倾向于将一种特殊的运算视作乘积:

定义 4.5 线性映射的乘积

设 $T \in \mathcal{L}(U, V)$, $S \in \mathcal{L}(V, W)$, 定义乘积 ST 是 U 到 W 的映射, 满足

$$\forall u \in U, (ST)(u) = S(Tu).$$

注. 同样的, 这里的 ST 是从 U 到 W 的线性映射.

注. 此处线性映射的乘积就是所谓“函数复合” $S \circ T$.

命题 4.6 线性映射乘积的代数性质

(1) 结合性: 设线性映射 T_1, T_2, T_3 , 在运算有意义的情况下, 有

$$(T_1 T_2) T_3 = T_1 (T_2 T_3).$$

(2) 单位元: 设 $T \in \mathcal{L}(V, W)$, I 是 W 上的恒等映射, 则

$$TI = IT = T.$$

(3) 分配性质: 设 $T, T_1, T_2 \in \mathcal{L}(U, V)$, $S, S_1, S_2 \in \mathcal{L}(V, W)$, 则

$$(S_1 + S_2)T = S_1T + S_2T, \quad S(T_1 + T_2) = ST_1 + ST_2.$$

请注意, 线性映射的乘法不满足交换性, 即 $ST = TS$ 不一定成立.

4.1.2 线性映射的表示矩阵

我们知道, 对于线性映射 $T: V \rightarrow W$, 通过 V 的基的象 Tv_1, \dots, Tv_n 可以确定任意 V 中元素的象. 稍后我们会利用 W 的基在矩阵上记录这些 Tv_j 的值.

定义 4.7 矩阵

设正整数 m, n , $m \times n$ 矩阵 A 是由 \mathbb{F} 的元素构成的 m 行 n 列的矩形阵列:

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}.$$

其中, 记号 $A_{j,k}$ 表示 A 的第 j 行第 k 列的元素.

定义 4.8 线性映射的矩阵

设 $T \in \mathcal{L}(V, W)$, 并设 v_1, \dots, v_n 是 V 的基, w_1, \dots, w_m 是 W 的基. 规定 T 关于这些基的矩阵为 $m \times n$ 矩阵 $\mathcal{M}(T)$, 其中 $A_{j,k}$ 满足

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m.$$

如果这些基不是上下文自明的, 则采用记号 $\mathcal{M}(T, (v_1, \dots, v_n), (w_1, \dots, w_m))$.

构造 $\mathcal{M}(T)$ 的方法如下图所示: 把 Tv_k 写成 w_1, \dots, w_m 的线性组合形式 $A_{1,k}w_1 + \cdots + A_{m,k}w_m$, 那

么所有系数自上而下构成的矩阵的第 k 列.

$$\mathcal{M}(T) = \begin{matrix} & Tv_1 & \cdots & Tv_k & \cdots & Tv_n \\ \begin{matrix} w_1 \\ \vdots \\ w_m \end{matrix} & \begin{pmatrix} & & & & \\ & A_{1,k} & & & \\ & \vdots & & & \\ & & & A_{m,k} & \end{pmatrix} \end{matrix}.$$

例如, 对于线性映射 $T: \mathbb{F}^2 \rightarrow \mathbb{F}^3$ 定义为 $T(x, y) = (x + 3y, 2x + 5y, 7x + 9y)$, 则 T 关于 \mathbb{F}^2 与 \mathbb{F}^3 的标准基的矩阵为

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}.$$

这是因为 $T(1, 0) = (1, 2, 7) = 1(1, 0, 0) + 2(0, 1, 0) + 7(0, 0, 1)$, $T(0, 1) = (3, 5, 9) = 3(1, 0, 0) + 5(0, 1, 0) + 9(0, 0, 1)$.

对于微分映射 $D: \mathcal{P}_3(\mathbb{R}) \rightarrow \mathcal{P}_2(\mathbb{R})$ 满足 $Dp = p'$, 它关于 $\mathcal{P}_3(\mathbb{R})$ 和 $\mathcal{P}_2(\mathbb{R})$ 的标准基的矩阵为

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

如果将向量表示为如下形式:

定义 4.9 向量的矩阵

设 $v \in V$, 并设 v_1, \dots, v_n 是 V 的基. 若 $v = c_1 v_1 + \dots + c_n v_n$, 规定 v 关于这个基的矩阵是一个 $n \times 1$ 矩阵

$$\mathcal{M}(v) = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

那么对于具有表示矩阵 A 的线性映射 $T \in \mathcal{L}(V, W)$, Tv 可以表示为:

$$\begin{pmatrix} A_{11}c_1 + \dots + A_{1n}c_n \\ \vdots \\ A_{m1}c_1 + \dots + A_{mn}c_n \end{pmatrix} := \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}.$$

我们将其定义为矩阵与向量的乘积. 如此, Tv 与 Av 的结果就相等了, 因而常将线性映射的表示矩阵写作其本身的符号. 当然, 这样的运算满足线性分配律 $A(c_1 v_1 + \dots + c_n v_n) = c_1(Av_1) + \dots + c_n(Av_n)$.

前文所提到的旋转变换就可以这样写:

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix}.$$

4.1.3 矩阵的运算

矩阵的加法与标量乘法定义很符合直觉:

定义 4.10 矩阵的加法与标量乘法

(1) 规定两个同样大小的矩阵的和是将对应元素相加得到的矩阵:

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} C_{1,1} & \cdots & C_{1,n} \\ \vdots & & \vdots \\ C_{m,1} & \cdots & C_{m,n} \end{pmatrix} = \begin{pmatrix} A_{1,1} + C_{1,1} & \cdots & A_{1,n} + C_{1,n} \\ \vdots & & \vdots \\ A_{m,1} + C_{m,1} & \cdots & A_{m,n} + C_{m,n} \end{pmatrix}.$$

(2) 规定标量与矩阵的乘积是将标量乘以每个元素得到的矩阵:

$$\lambda \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}.$$

因而, 我们有

命题 4.11 线性映射与矩阵运算

(1) 设 $S, T \in \mathcal{L}(V, W)$, 则 $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$;

(2) 设 $\lambda \in \mathbb{F}$, $T \in \mathcal{L}(V, W)$, 则 $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

命题 4.12 $\mathbb{F}^{m,n}$ 是向量空间

对于正整数 m, n , 元素取自 \mathbb{F} 的所有 $m \times n$ 矩阵的集合记为 $\mathbb{F}^{m,n}$. 按照矩阵运算的定义, $\mathbb{F}^{m,n}$ 是 mn 维向量空间.

上述命题的证明是显然的.

我们注意到, 线性映射不止有加法和标量乘法, 还有元素之间的乘法. 猜测是否会有 $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$? 为了得到这个结果, 尝试倒推矩阵乘法的定义:

考虑 $T \in \mathcal{L}(U, V)$, $S \in \mathcal{L}(V, W)$, 并设 u_1, \dots, u_p 是 U 的基, v_1, \dots, v_n 是 V 的基, w_1, \dots, w_m 是

W 的基. 记 $\mathcal{M}(S) = A$, $\mathcal{M}(T) = C$. 那么对于任意的 $1 \leq k \leq p$, 有

$$(ST)u_k = S\left(\sum_{r=1}^n C_{r,k}v_r\right) = \sum_{r=1}^n C_{r,k}Sv_r = \sum_{r=1}^n C_{r,k}\sum_{j=1}^m A_{j,r}w_j = \sum_{j=1}^m \left(\sum_{r=1}^n A_{j,r}C_{r,k}\right)w_j.$$

因此 $\mathcal{M}(ST)$ 是 $m \times p$ 矩阵, 且满足

$$\mathcal{M}(ST)_{j,k} = \sum_{r=1}^n A_{j,r}C_{r,k}.$$

于是可以定义:

定义 4.13 矩阵乘法

设 A 是 $m \times n$ 矩阵, C 是 $n \times p$ 矩阵. AC 定义为 $m \times p$ 矩阵, 满足

$$(AC)_{j,k} = \sum_{r=1}^n A_{j,r}C_{r,k}.$$

也即, 将 A 的第 j 行与 C 的第 k 列元素对应相乘再相加.

这样的矩阵乘法脱胎于线性映射的乘法, 因此其代数性质也类似线性映射乘法的结合性、单位元、分配性质, 且不满足交换性.

例如, 将一个 3×2 矩阵与一个 2×4 矩阵相乘, 得到一个 3×4 矩阵:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 6 & 5 & 4 & 3 \\ 2 & 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 10 & 7 & 4 & 1 \\ 26 & 19 & 12 & 5 \\ 42 & 31 & 20 & 9 \end{pmatrix}.$$

若将第二个矩阵写作行向量的列向量形式, 可以视作矩阵对向量的乘法:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} A_1^T \\ A_2^T \end{pmatrix} = \begin{pmatrix} A_1^T + 2A_2^T \\ 3A_1^T + 4A_2^T \\ 5A_1^T + 6A_2^T \end{pmatrix}.$$

实际上是将每一列分开乘再进行特殊的相加 (这里记作 $\tilde{+}$):

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \left(\begin{pmatrix} 6 \\ 2 \end{pmatrix} \tilde{+} \begin{pmatrix} 5 \\ 1 \end{pmatrix} \tilde{+} \begin{pmatrix} 4 \\ 0 \end{pmatrix} \tilde{+} \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} 10 \\ 26 \\ 42 \end{pmatrix} \tilde{+} \begin{pmatrix} 7 \\ 19 \\ 31 \end{pmatrix} \tilde{+} \begin{pmatrix} 4 \\ 12 \\ 20 \end{pmatrix} \tilde{+} \begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix}.$$

例如将列向量转为行向量, 有时需要考虑矩阵的转置:

定义 4.14 矩阵的转置

$m \times n$ 矩阵 A 的转置定义为将其行列交换得到的 $n \times m$ 矩阵, 记作 A^T .

容易验证如下命题:

命题 4.15 转置的运算

设 $m \times n$ 矩阵 A , $n \times p$ 矩阵 B , $\lambda, \mu \in \mathbb{F}$. 我们有

$$1. (\lambda A + \mu B)^T = \lambda A^T + \mu B^T, \quad 2. (AB)^T = B^T A^T.$$

4.2 零空间与值域

4.2.1 解线性方程组问题

考虑解一个非齐次线性方程组 (即右侧常数不全为 0):

$$\begin{cases} 4x_1 + 3x_2 - 6x_3 = -17 \\ x_1 + 2x_2 + 3x_3 = 22 \\ 2x_1 + 3x_2 = 11 \end{cases}.$$

我们当然可以使用一些奇技淫巧, 例如通过后两个方程直接得到 x_1, x_2, x_3 的倍数关系, 但这毕竟不是通用方法. 最经典的方法应该是消元: 不断地利用倍乘变换 (将某个方程左右同乘常数 k)、倍加变换 (将某个方程的 k 倍加到另一个方程上) 使得最后每个方程都变成左边只含一个变量 (如果将多余的变量视作常量的话). 我们可以将方程组的系数提取出来操作:

$$\left(\begin{array}{ccc|c} 4 & 3 & -6 & -17 \\ 1 & 2 & 3 & 22 \\ 2 & 3 & 0 & 11 \end{array} \right) \xrightarrow{\text{化简}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 5 \end{array} \right).$$

左侧矩阵被称作方程组的增广矩阵. 从化简的结果来看, 我们得到了 $x_1 = 1, x_2 = 3, x_3 = 5$. 需要注意的是, 为了保证 x_1, x_2, x_3 顺次排列, 可能还会用到对换变换 (将矩阵的两行调换位置).

有些时候不一定能够得到每个变量的固定解, 而是需要用一部分变量来表示另一部分变量. 例如:

$$\begin{cases} 2x_1 + 3x_2 = 5 \\ x_1 + x_4 = 8 \\ x_1 + x_2 + x_3 = 4 \end{cases} \xrightarrow{\text{表示}} \left(\begin{array}{cccc|c} 2 & 3 & 0 & 0 & 5 \\ 1 & 0 & 0 & 1 & 8 \\ 1 & 1 & 1 & 0 & 4 \end{array} \right) \xrightarrow{\text{化简}} \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{2} & \frac{65}{8} \\ 0 & 1 & 0 & -\frac{1}{3} & -\frac{15}{4} \\ 0 & 0 & 1 & -\frac{2}{3} & -\frac{1}{3} \end{array} \right).$$

从而 $x_1 = -\frac{1}{2}x_4 + \frac{65}{8}, x_2 = \frac{1}{3}x_4 - \frac{15}{4}, x_3 = \frac{2}{3}x_4 - \frac{1}{3}$.

从上述例子中, 尝试提取一些共性: 通过三种变换 (倍乘变换, 倍加变换, 对换变换, 统称为基本变换), 我们总是能将一个增广矩阵化简为一部分列恰存在一个 1 元素的矩阵. (利用归纳法容易证明) 这样的列对应主变量, 称作主列; 而剩下的列对应自由变量, 称作自由列. 最后化简出来的矩阵称作原矩阵 A 的行简化阶梯型矩阵, 记作 $\text{rref}(A)$.

容易验证, 一个齐次线性方程组只有零解当且仅当其系数矩阵的行简化阶梯型矩阵的阶梯数等于未知数个数; 一个非齐次线性方程组有解当且仅当增广矩阵和系数矩阵的行简化阶梯型矩阵的阶梯数相等.

既然是变换, 能否找到其对应的 (线性) 映射呢? 在 \mathbb{F}^n 上考虑, 我们从最简单的恒等变换开始:

设 $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^n)$ 满足 $T: v \mapsto v$, 由 $Te_i = e_i$ 可知 (其中空白部分为 0)

$$T = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} =: I_n.$$

将列向量 v 替换为行向量的列向量即可验证对任意矩阵 A 都有 $I_n A = A I_n = A$.

在 I_n 的基础上, 可以分别得到倍乘变换、倍加变换、对换变换的表示矩阵:

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

实际上这些矩阵都是对 I_n 做一次变换得到的.

从线性映射的角度, 设初等变换的表示矩阵 R , 则 RA 表示将 A 进行初等行变换. 相对应地, AR 则表示将 A 进行初等列变换, 其法则类似于行变换.

4.2.2 可逆矩阵

定义 4.16 矩阵的逆

设 A 是 n 阶方阵, 若存在 n 阶方阵 B 满足 $AB = BA = I_n$, 则称 A 是可逆的, B 为其逆.

注. 广义的矩阵逆定义如上. 然而, 研究矩阵的“左逆”和“右逆”往往更有趣 (这是因为, $AB = I_n$ 足以推出 $BA = I_n$ 了, 这一点将在后文说明), 只不过原书中没有涉及这一部分内容, 只好忍痛删去.

容易验证, 一个矩阵若可逆, 则其逆一定唯一. 因此, 我们得以用 A^{-1} 来指代 A 的逆.

举一些可逆矩阵的例子: I_n 的逆是它本身, 旋转 θ 变换的逆是旋转 $-\theta$, 初等变换的逆就是对应操作的逆. 一个重要的例子是所谓对角矩阵:

$$\text{diag}(a_1, \dots, a_n) = \begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{pmatrix} =: D, \text{ 若 } a_1 \cdots a_n \neq 0 \text{ 则 } D \text{ 的逆存在且 } D^{-1} = \text{diag}(a_1^{-1}, \dots, a_n^{-1}).$$

容易验证如下命题.

命题 4.17 矩阵的逆的运算

设 A, B 是可逆矩阵, 则

$$1. (A^T)^{-1} = (A^{-1})^T, \quad 2. (AB)^{-1} = B^{-1}A^{-1}.$$

命题 4.18 矩阵可逆的条件

设 n 阶方阵 A , 则以下说法等价:

- a) A 可逆. b) 齐次线性方程组 $Ax = 0$ 的解唯一. c) $\text{rref } A = I_n$.

d) A 是有限个初等矩阵的乘积.

证明 采用轮换证法. $a) \Rightarrow b)$: A 可逆即其对应的线性映射可逆, 亦等价于该线性映射为双射. 满射性可以导出 $Ax = 0$ 存在解, 单射性可以导出 $Ax = 0$ 解唯一.

$b) \Rightarrow c)$: 显然.

$c) \Rightarrow d)$: 在将 A 化简为 $\text{rref } A$ 即 I_n 时, 会应用有限个可逆的初等行变换 E_1, \dots, E_n , 即 $A = E_1 \cdots E_n I_n$.

$d) \Rightarrow a)$: $A^{-1} = E_n^{-1} \cdots E_1^{-1}$.

由该命题可以得到一种计算矩阵的逆的方法. 对于 n 阶矩阵 A , 将矩阵 $\begin{pmatrix} A & I_n \end{pmatrix}$ 变换为 $\begin{pmatrix} I_n & E \end{pmatrix}$, 所需初等行变换的乘积为 E , 那么 $EA = I_n$, 进一步 $A^{-1} = EAA^{-1} = E$.

有一些特殊的矩阵的逆需要知晓.

定义上三角矩阵为对角线左下方的元素均为 0 的矩阵. 那么其可逆当且仅当其对角元素均不为 0. 逆矩阵也是上三角矩阵, 且对角线元素为原矩阵对角线元素的倒数. 下三角矩阵类似.

定义 (行) 对角占优矩阵 $A = (a_{ij})_{n \times n}$ 满足对所有 $i = 1, \dots, n$ 有 $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$. 对角占优矩阵可逆.

证明 即证 $Ax = 0$ 只有 0 解. 不然, 设 x_i 为 x 的分量中绝对值最大的. 考虑第 i 个方程

$$a_{i1}x_1 + \cdots + a_{ii}x_i + \cdots + a_{in}x_n = 0.$$

显然 $x_i \neq 0$, 则

$$|a_{ii}||x_i| = \left| \sum_{j \neq i} a_{ij}x_j \right| \leq \sum_{j \neq i} |a_{ij}||x_j| \leq |x_i| \sum_{j \neq i} |a_{ij}| < |a_{ii}||x_i|.$$

矛盾!

4.2.3 零空间与值域

定义 4.19 零空间

对于 $T \in \mathcal{L}(V, W)$, T 的零空间 (或称为“核”) 定义如下:

$$\text{null } T = \{v \in V : Tv = 0\}$$

例 4.4 零空间的例子 (1) 若 T 是 V 到 W 的零映射, 则 $\text{null } T = V$.

(2) 设 $\varphi \in \mathcal{L}(\mathbb{C}^3, \mathbb{F})$ 定义为 $\varphi(z_1, z_2, z_3) = z_1 + 2z_2 + 3z_3$. 则

$$\text{null } \varphi = \{(z_1, z_2, z_3) \in \mathbb{C}^3 : z_1 + 2z_2 + 3z_3 = 0\}$$

并且 $\text{null } \varphi$ 的一个基为 $(-2, 1, 0), (-3, 0, 1)$.

(3) 设 $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$ 是微分映射 $Dp = p'$. 只有常函数的导数才能等于零函数, 故 T 的零空间是常

函数组成的集合.

(4) 设 $T \in \mathcal{L}(\mathbb{F}^\infty, \mathbb{F}^\infty)$ 是向后移位映射

$$T(x_1, x_2, x_3, \cdots) = (x_2, x_3, \cdots)$$

为让 $Tv = 0$, 要求 $x_2 = x_3 = \cdots = 0$, 故 $\text{null } T = \{(a, 0, 0, \cdots) : a \in \mathbb{F}\}$.

自然地, 零空间是向量空间.

命题 4.20 零空间是子空间

设 $T \in \mathcal{L}(V, W)$, 则 $\text{null } T$ 是 V 的子空间.

证明 略. 为证明上述命题, 只需注意到 $T(0) = 0$ (因为 $T(0+0) = T(0) + T(0)$).

定义 4.21 单射

如果当 $Tu = Tv$ 时必有 $u = v$, 则称映射 $T: V \rightarrow W$ 是单射.

命题 4.22 单射性的判定

设 $T \in \mathcal{L}(V, W)$, 则 T 是单射当且仅当 $\text{null } T = \{0\}$.

证明 1° 充分性: 当 $\text{null } T = \{0\}$ 时, 设 $Tu = Tv$, 则 $Tu - Tv = T(u - v) = 0$, 于是 $u - v = 0$, 即 $u = v$. 这表明 T 是单射.

2° 必要性: 任取 $\text{null } T$ 中的元素 v , 则 $Tv = 0$. 因为 $T0 = 0$ 且 T 是单射, 所以必有 $v = 0$, 即 $\text{null } T = \{0\}$.

定义 4.23 值域

对于 $T \in \mathcal{L}(V, W)$, T 的值域 (或称为“像”) 定义如下:

$$\text{range } T = \{Tv : v \in V\}$$

例 4.5 值域的例子 (1) 若 T 是 V 到 W 的零映射, 则 $\text{range } T = \{0\}$.

(2) 设 $T \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^3)$ 定义为 $T(x, y) = (2x, 5y, x + y)$, 则

$$\text{range } T = \{(2x, 5y, x + y) : x, y \in \mathbb{R}\}$$

并且 $\text{range } T$ 的一个基为 $(2, 0, 1), (0, 5, 1)$.

(3) 设 $D \in \mathcal{L}(\mathcal{P}(\mathbb{R}), \mathcal{P}(\mathbb{R}))$ 是微分映射 $Dp = p'$. 由于对每个多项式 $q \in \mathcal{P}(\mathbb{R})$ 均存在多项式 p 使得 $p' = q$, 故 D 的值域为 $\mathcal{P}(\mathbb{R})$.

自然地, 值域是向量空间.

命题 4.24 值域是子空间

设 $T \in \mathcal{L}(V, W)$, 则 $\text{range } T$ 是 W 的子空间.

证明 略.

定义 4.25 满射

如果函数 $T: V \rightarrow W$ 的值域等于 W , 则称 T 为满射.

上述例子中只有微分映射是满的.

4.2.4 线性映射基本定理

命题 4.26 线性映射基本定理

设 V 是有限维的, $T \in \mathcal{L}(V, W)$. 则 $\text{range } T$ 是有限维的并且

$$\dim V = \dim \text{null } T + \dim \text{range } T.$$

证明 设 v_1, \dots, v_m 是 $\text{null } T$ 的基. 将其扩展为 V 的一个基 $v_1, \dots, v_m, u_1, \dots, u_n$. 注意到原命题等价于证明 $\dim \text{range } T = n$, 于是下证 Tu_1, \dots, Tu_n 为 $\text{range } T$ 的基:

首先, 任取 $v \in V$, 记 $v = a_1v_1 + \dots + a_mv_m + b_1u_1 + \dots + b_nu_n$, 则

$$Tv = a_1Tv_1 + \dots + a_mTv_m + b_1Tu_1 + \dots + b_nTu_n = b_1Tu_1 + \dots + b_nTu_n$$

故 Tu_1, \dots, Tu_n 张成 $\text{range } T$.

另外, 若 $b_1Tu_1 + \dots + b_nTu_n = 0$, 则

$$T(b_1u_1 + \dots + b_nu_n) = 0$$

这表明 $b_1u_1 + \dots + b_nu_n \in \text{null } T$. 记 $b_1u_1 + \dots + b_nu_n = c_1v_1 + \dots + c_mv_m$, 则由 $v_1, \dots, v_m, u_1, \dots, u_n$ 线性无关, 可得 $b_1 = \dots = b_n = c_1 = \dots = c_m = 0$, 于是 Tu_1, \dots, Tu_n 线性无关.

注. 需要注意的是, 证明线性映射基本定理的前提是 V 为有限维.

利用线性映射基本定理, 可以快速证伪某些命题.

命题 4.27

设有限维向量空间 V, W . 若 $\dim V > \dim W$, 那么 V 到 W 的线性映射 T 一定不是单射; 相反地, 若 $\dim V < \dim W$, 那么 V 到 W 的线性映射 T 一定不是满射.

证明 只证明第一部分. 由 $\dim V > \dim W \geq \dim \text{range } T$, 可知 $\dim \text{null } T = \dim V - \dim \text{range } T > 0$, 于是 T 不是单射.

例 4.6 用线性映射重述齐次线性方程组是否有非零解的问题. 即, 对给定的正整数 m, n , 设 $A_{j,k} \in \mathbb{F}$ ($j = 1, \dots, m, k = 1, \dots, n$), 考虑齐次线性方程组

$$\begin{cases} \sum_{k=1}^n A_{1,k} x_k = 0 \\ \dots\dots\dots \\ \sum_{k=1}^n A_{m,k} x_k = 0 \end{cases}$$

是否有不全为 0 的解.

解. 构造 $T: \mathbb{F}^n \rightarrow \mathbb{F}^m$ 满足:

$$T(x_1, \dots, x_n) = \left(\sum_{k=1}^n A_{1,k} x_k, \dots, \sum_{k=1}^n A_{m,k} x_k \right)$$

易于证明 T 是线性映射. 则原方程有不全为 0 的解等价于 T 不是单射. 由上述命题可知, 若 $n > m$ 则 T 一定不是单射. 故当 $n > m$ 时原方程组有不全为 0 的解.(即变量个数大于方程个数时)

例 4.7 用线性映射重述是否可以选取常数项使得非齐次线性方程组无解的问题. 即, 对给定的正整数 m, n , 设 $A_{j,k} \in \mathbb{F}$ ($j = 1, \dots, m, k = 1, \dots, n$) 及 $c_1, \dots, c_m \in \mathbb{F}$, 考虑线性方程组

$$\begin{cases} \sum_{k=1}^n A_{1,k} x_k = c_1 \\ \dots\dots\dots \\ \sum_{k=1}^n A_{m,k} x_k = c_m \end{cases}$$

是否存在某些常数 c_1, \dots, c_m 使得上述方程组无解.

解. 构造 $T: \mathbb{F}^n \rightarrow \mathbb{F}^m$ 满足:

$$T(x_1, \dots, x_n) = \left(\sum_{k=1}^n A_{1,k} x_k, \dots, \sum_{k=1}^n A_{m,k} x_k \right)$$

易于证明 T 是线性映射. 则存在这样的一组常数等价于 T 不是满的. 由上述命题可知, 若 $n < m$ 则

T 一定不是满射. 故当 $n < m$ 存在这样一组常数. (即变量个数小于方程个数时)

4.3 可逆性与同构的向量空间

4.3.1 线性映射的可逆性

类似于一般的函数，我们可以定义线性映射的可逆性：

定义 4.28 线性映射的可逆性

线性映射 $T \in \mathcal{L}(V, W)$ 称为可逆的，如果存在线性映射 $S \in \mathcal{L}(W, V)$ 使得 ST 等于 V 上的恒等映射且 TS 等于 W 上的恒等映射. 这样的 S 称作 T 的逆，记为 T^{-1} .

这里的“逆”，在线性映射的乘法意义下，即为其乘法逆元. 自然它是唯一的.

命题 4.29

可逆的线性映射有唯一的逆.

证明 设 $T \in \mathcal{L}(V, W)$ 可逆，且 S_1, S_2 均为 T 的不同的逆. 由于

$$S_1 = S_1 I = S_1 (TS_2) = (S_1 T) S_2 = I S_2 = S_2$$

这与假设矛盾. 故 T 的逆是唯一的.

以映射的观点来看，一个函数可逆当且仅当它是双射. 这一点对于线性映射也成立.

命题 4.30 线性映射可逆性的判定

一个线性映射是可逆的当且仅当它既是单的又是满的.

证明 1° 必要性：设 $T \in \mathcal{L}(V, W)$ 是可逆的. 设 $u_1, u_2 \in V$ 使得 $Tu_1 = Tu_2$ ，那么

$$u_1 = T^{-1}Tu_1 = T^{-1}Tu_2 = u_2.$$

于是 T 是单的. 另一方面，设 $w \in W$ ，则由 $w = T(T^{-1}w)$ 可知 $W \subseteq \text{range } T$ ，又 $\text{range } T \subseteq W$ ，则 $W = \text{range } T$ ，即 T 是满的.

2° 充分性：设 T 既是单的又是满的，构造映射 S 满足：对于每个 $w \in W$ ， Sw 表示使得 $T(Sw) = w$ 成立的 V 中的唯一元素 (这里的存在与唯一可以由 T 的单射与满射得到). 我们证明 S 是线性映射且 ST 是 V 上的恒等映射.

首先，设 $w_1, w_2 \in W$ ，由于

$$T(Sw_1 + Sw_2) = TS w_1 + TS w_2 = w_1 + w_2,$$

$$T(S(w_1 + w_2)) = w_1 + w_2,$$

所以 $S(w_1 + w_2) = Sw_1 + Sw_2$. 类似地可得 S 的齐性. 故 S 是线性映射.

接着, 任取 $v \in V$, 由于

$$T(STv) = (TS)Tv = ITv = Tv.$$

所以 $STv = v$, 即 ST 是 V 上的恒等映射.

其实我们可以做出更为细致的分析:

命题 4.31

- (1) 设 W 是有限维的, $T \in \mathcal{L}(V, W)$, 则 T 是单的当且仅当存在 $S \in \mathcal{L}(W, V)$ 使得 $ST = I$.
 (2) 设 V 是有限维的, $T \in \mathcal{L}(V, W)$, 则 T 是满的当且仅当存在 $S \in \mathcal{L}(W, V)$ 使得 $TS = I$.

证明 (1) " \Rightarrow ": 假设 T 是单的. 记 $S' : \text{range } T \rightarrow V$ 使得 $S'(Tv) = v$ (由于 Tv 一定对应一个 v , 这是良定义的). 将 S' 延伸到 $S : W \rightarrow V$ 即可.

" \Leftarrow ": 设 u, v 使得 $Tu = Tv$, 则 $ST(u) = ST(v)$, 从而 $u = v$.

(2) " \Rightarrow ": 假设 T 是满的. 设 Tv_1, \dots, Tv_m 是 W 的一组基 (由于任意 W 中元素都能表示为 Tv 的形式, 这是良定义的). 记 $S : W \rightarrow V$ 使得 $S(c_1Tv_1 + \dots + c_mTv_m) = c_1v_1 + \dots + c_mv_m$. 容易验证 S 是线性的.

" \Leftarrow ": 对于任意 $w \in W$, $TS(w) = w$, 即 T 是满的.

一般来说, 单射性和满射性并不相互蕴含. 然而对于从向量空间映射到自身的线性映射, 称为算子, 满射性和单射性是等价的 (如果有限维). 设 V 是有限维的, $T \in \mathcal{L}(V)$, 则 T 是单的等价于 $\dim \text{null } T = 0$, T 是满的等价于 $\dim \text{range } T = \dim V$, 由线性映射基本定理就可以得到结论. 利用这一点我们可以得到:

命题 4.32

若 V, W 都是有限维向量空间, $S \in \mathcal{L}(V, W), T \in \mathcal{L}(W, V)$, 那么 $ST = I$ 等价于 $TS = I$.

证明 只证明必要性. 由上一个命题我们知道 T 是单的, 则 T 是满的, 进而 T 可逆. 所以 $ST = I$ 就得到 $S = T^{-1}$, 进一步有 $TS = TT^{-1} = I$.

4.3.2 同构的向量空间

在高中数学中, “同构” 这个词被大量滥用, 但其也能为我们揭示同构的内涵. 例如, 我们说等式

$$x(\ln x + 1) = ye^{y-1}$$

关于 x 和 y 是同构的, 是因为若作换元 $y = \ln t + 1$, 可得 $x(\ln x + 1) = t(\ln t + 1)$.

为什么 x 与 y “同构” 呢? 因为 y 和 x 可以通过一个映射联系起来¹. 类似地, 我们正式给出两个向量空间的同构定义:

¹请注意, 这是直观而不严格的说法.

定义 4.33 向量空间的同构

同构就是可逆的线性映射. 若两个向量空间之间存在一个同构, 则称这两个向量空间是同构的.

同构 $T: V \rightarrow W$ 做了一步操作, 将 $v \in V$ 重新标记为 $Tv \in W$; T 的逆 T^{-1} 同等地将每个 $Tv \in W$ 重新标记为 $v \in V$. 于是 V 与 W 中的元素只是形式不一样, 其性质是一样的.

回想之前提到的“矩阵乘法和线性映射乘法的代数性质一致”这件事, 本质上是因为 $\mathcal{L}(V, W)$ 与 $\mathbb{F}^{m,n}$ 同构.

命题 4.34 $\mathcal{L}(V, W)$ 与 $\mathbb{F}^{m,n}$ 同构

设 v_1, \dots, v_n 是 V 的基, w_1, \dots, w_m 是 W 的基, 则 \mathcal{M} 是 $\mathcal{L}(V, W)$ 与 $\mathbb{F}^{m,n}$ 之间的一个同构.

证明 将 \mathcal{M} 视作一个映射, 那么由命题 1.1 可知它是线性的. 现在只需证明它可逆.

一方面, 若对于 $T \in \mathcal{L}(V, W)$, $\mathcal{M}(T) = 0$, 则由定义可得 $Tv_k = 0$, $k = 1, \dots, n$, 那么 $T(c_1v_1 + \dots + c_nv_n) = c_1Tv_1 + \dots + c_nTv_n = 0$, 即 $\text{null } \mathcal{M} = \{0\}$, 于是 \mathcal{M} 是单的.

另一方面, 任取 $A \in \mathbb{F}^{m,n}$, 构造线性映射 $T: V \rightarrow W$ 满足

$$Tv_k = \sum_{j=1}^m A_{j,k} w_j$$

则 $\mathcal{M}(T) = A$. 这表明 \mathcal{M} 是满射.

作为一个例子, 我们注意到为证明 $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W$ 需要一些构造, 但是将其转化为矩阵空间则是显然的. 这启示我们可以从与一个向量空间同构的另一个更简单 (或已知) 的向量空间来考察该向量空间.

注意到一个问题: 类似于高中数学中“集合的对应原理”: 若两个有限集合 A, B 之间存在一个双射 f , 则 $|A| = |B|$. 两个向量空间同构, 它们的维数应当相同. 而更进一步, 由于“维数相同”这一概念比“集合元素个数相等”更强, 上面的说法反过来也可以是对的.

命题 4.35 向量空间同构的判定

\mathbb{F} 上两个有限维向量空间同构当且仅当其维数相同.

证明 1° 必要性: 设 V 和 W 是同构的有限维向量空间, 即存在可逆的线性映射 $T: V \rightarrow W$. 于是 $\dim \text{null } T = 0$, $\dim \text{range } T = \dim W$. 又由线性映射基本定理可知

$$\dim V = \dim \text{null } T + \dim \text{range } T = \dim W$$

2° 充分性: 设 V 和 W 维数相同, v_1, \dots, v_n 是 V 的基, w_1, \dots, w_n 是 W 的基. 由命题 2 可知存在一个线性映射 $T: V \rightarrow W$ 满足

$$T(c_1v_1 + \dots + c_nv_n) = c_1w_1 + \dots + c_nw_n$$

只需证明这个 T 是可逆的. 实际上, 若 $T(c_1v_1 + \cdots + c_nv_n) = 0$, 由于 w_1, \cdots, w_n 是线性无关的, 必有 $c_1 = \cdots = c_n = 0$, 即 $\text{null } T = \{0\}$, 即 T 是单的; 另一方面, 等式右侧是 w_1, \cdots, w_n 的线性组合形式, 于是 $\text{range } T = W$, 即 T 是满的.

4.3.3 基的变换

在开始下方的命题之前, 我们先来看一个显然而有用的等式: 设 $T \in \mathcal{L}(V, W)$, V 和 W 的一组基分别为 $v_1, \cdots, v_n, w_1, \cdots, w_m$, 记 $A = \mathcal{M}(T, (v_1, \cdots, v_n), (w_1, \cdots, w_m))$, 那么

$$\begin{pmatrix} w_1 & \cdots & w_m \end{pmatrix} A = \begin{pmatrix} Tv_1 & \cdots & Tv_n \end{pmatrix}.$$

注意, 由于我们不在乎左右两侧的矩阵对应的线性映射, 请不要认为它们和基的选取相关.

命题 4.36 线性映射之积的矩阵

设 $T \in \mathcal{L}(U, V), S \in \mathcal{L}(V, W)$. 我们有

$$\begin{aligned} \mathcal{M}(ST, (u_1, \cdots, u_m), (w_1, \cdots, w_p)) = \\ \mathcal{M}(S, (v_1, \cdots, v_n), (w_1, \cdots, w_p)) \mathcal{M}(T, (u_1, \cdots, u_m), (v_1, \cdots, v_n)). \end{aligned}$$

证明 我们定义矩阵乘法就是为了这个命题.

例 4.8 矩阵

$$\mathcal{M}(I, (u_1, \cdots, u_n), (v_1, \cdots, v_n)), \quad \mathcal{M}(I, (v_1, \cdots, v_n), (u_1, \cdots, u_n))$$

是可逆的, 且互为逆矩阵.

命题 4.37 换基公式

设 $T \in \mathcal{L}(V)$, V 有两组基 v_1, \cdots, v_n 和 u_1, \cdots, u_n . 记

$$A = \mathcal{M}(T, (u_1, \cdots, u_n)), \quad B = \mathcal{M}(T, (v_1, \cdots, v_n)), \quad C = \mathcal{M}(I, (u_1, \cdots, u_n), (v_1, \cdots, v_n)).$$

那么我们有

$$A = C^{-1}BC.$$

形式化的证明直接来源于前面的命题, 这里省略. 上述命题的另一种理解方式是

$$\begin{aligned} \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} A &= \begin{pmatrix} Tu_1 & \cdots & Tu_n \end{pmatrix}, & \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} B &= \begin{pmatrix} Tv_1 & \cdots & Tv_n \end{pmatrix}, \\ \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} C &= \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}. \end{aligned}$$

只需注意到

$$\begin{pmatrix} Tv_1 & \cdots & Tv_n \end{pmatrix} = T \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix},$$

即可完成同样的证明.

4.4 向量空间的积与商

4.4.1 向量空间的积

定义 4.38 向量空间的积

设 \mathbb{F} 上的向量空间 V_1, \dots, V_m . 定义:

- (1) 积 $V_1 \times \dots \times V_m := \{(v_1, \dots, v_m) : v_j \in V_j, i = 1, \dots, m\}$.
- (2) 积上的加法 $(v_1, \dots, v_m) + (u_1, \dots, u_m) = (v_1 + u_1, \dots, v_m + u_m)$.
- (3) 积上的标量乘法 $\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m)$.

容易验证, $V_1 \times \dots \times V_m$ 是一个 \mathbb{F} 上的向量空间. 非常直观 (且容易证明) 地, 我们有:

命题 4.39 积的维数是维数之和

设有限维向量空间 V_1, \dots, V_m , 那么 $\dim(V_1 \times \dots \times V_m) = \dim V_1 + \dots + \dim V_m$.

证明 将 V_j 的一组基认作除第 j 位是基向量以外都是 0 的向量, 这构成 $V_1 \times \dots \times V_m$ 中的一组基.

另外, 我们可以将直和与积联系在一起.

命题 4.40

设 V_1, \dots, V_m 是 V 的子空间, 设

$$\Gamma : V_1 \times \dots \times V_m \rightarrow V_1 + \dots + V_m, \quad (v_1, \dots, v_m) \mapsto v_1 + \dots + v_m.$$

那么, $V_1 + \dots + V_m$ 是直和当且仅当 Γ 是单的.

证明是显然的. 从而, 当 V 是有限维向量空间时, 联系上两个命题可知 $V_1 + \dots + V_m$ 是直和当且仅当 $\dim(V_1 + \dots + V_m) = \dim V_1 + \dots + \dim V_m$.

4.4.2 向量空间的商

回顾一般的商集: 首先在集合 X 上定义一个等价关系 \sim , 称所有 X 上的等价类 $[a]_\sim$ 构成的集合是商集 X/\sim . 进一步, 我们会研究商映射 $\pi : X \rightarrow X/\sim, x \mapsto [x]_\sim$.

对于向量空间, 我们似乎可以使用平移的方式定义等价关系:

定义 4.41 仿射子集

设 $v \in V$, U 是 V 的子空间. 定义 V 的仿射子集 $v + U := \{v + u : u \in U\}$.

命题 4.42

设 U 是 V 的子空间, $v, w \in V$, 那么以下说法等价:

$$1) \ v - w \in U, \quad 2) \ v + U = w + U, \quad 3) \ (v + U) \cap (w + U) \neq \emptyset.$$

证明 $1) \Rightarrow 2)$: 任取 $u \in U$, 那么 $v + u = w + (v - w + u) \in w + U$, 从而 $v + U \subseteq w + U$, 反之亦然.

$2) \Rightarrow 3)$: 显然. $3) \Rightarrow 1)$: 即存在 $u_1, u_2 \in U$ 使得 $v + u_1 = w + u_2$, 从而 $v - w \in U$.

定义 4.43 商空间

设 U 是 V 的子空间. 定义:

- (1) 商空间 $V/U := \{v + U : v \in V\}$.
- (2) 商空间上的加法 $(v + U) + (w + U) = (v + w) + U$.
- (3) 商空间上的标量乘法 $\lambda(v + U) = \lambda v + U$.

我们需要验证 V/U 是一个向量空间, 即是要验证其加法和标量乘法的良定义性. 以前者为例, 设 $v + U = v' + U, w + U = w' + U$, 则 $v - v' \in U, w - w' \in U$, 说明 $v - v' + w - w' = (v + w) - (v' + w') \in U$, 从而 $(v + w) + U = (v' + w') + U$. 接下来只需注意到 $0 + U$ 是加法单位元, $(-v) + U$ 是加法逆元即可完成证明.

定义 4.44 商映射

设 U 是 V 的子空间, 定义商映射 $\pi : V \rightarrow V/U, v \mapsto v + U$.

注意到若 V 是有限维的, 那么 $\dim \text{null } \pi = \dim U, \dim \text{range } \pi = \dim V/U$, 所以有

$$\dim V/U = \dim V - \dim U.$$

下面来看一个和商空间相关的重要实例. 设向量空间 $V, W, T \in \mathcal{L}(V, W), w \in W$. 我们知道 $Tv = w$ 的解集就是非齐次线性方程组 $\mathcal{M}(T)v = w$ 的解集, 而 $\text{null } T$ 是齐次线性方程组 $\mathcal{M}(T)v = 0$ 的解集. 将两式相减, 我们得到:

命题 4.45 线性方程组的解空间

设线性方程组 $\mathcal{M}(T)v = w$ 的一个特解为 v_0 , 那么该线性方程组的解集就是 $v_0 + \text{null } T$.

这里我们是固定了 w 来考虑的. 那么, 如果 w 遍历 W , 解空间的集合 \mathcal{S} 应该具有怎样的结构? 对于有限维的情况, 显然 $\mathcal{S} = V / \text{null } T$, 所以 $\dim \mathcal{S} = \dim V - \dim \text{null } T = \dim \text{range } T$, 这说明 \mathcal{S} 与 $\text{range } T$ 同构. 实际上可以证明, 这一点对于无限维情况仍然适用.

定义 4.46

设 $T \in \mathcal{L}(V, W)$. 定义 $\tilde{T} : V / \text{null } T \rightarrow \text{range } T$, $v + \text{null } T \mapsto Tv$.

首先不难验证, \tilde{T} 是一个线性映射. 我们只要证明 \tilde{T} 是双射: 定义表明 \tilde{T} 是满的; 现在设 $v \in V$ 使得 $\tilde{T}(v + \text{null } T) = 0 = Tv$, 即 $v \in \text{null } T$, 从而 $v + \text{null } T = 0 + \text{null } T$, 这就是说 \tilde{T} 是单的.

4.5 对偶与矩阵的秩

4.5.1 对偶

之前我们证明了矩阵乘法就等价于其对应线性映射的复合, 也说明了可逆矩阵和可逆线性映射的关系. 实际上, 矩阵的对偶也会有对应的线性映射的形式, 但这一套理论并不显然. 现在让我们以反向探索的方式进行研究.

设线性映射 $T \in \mathcal{L}(V, W)$, V 的一组基 v_1, \dots, v_m , W 的一组基 w_1, \dots, w_n , 简记 $\mathcal{M}(T)$ 表示 $\mathcal{M}(T, (v_1, \dots, v_m), (w_1, \dots, w_n))$.

首先, 由于 $\mathcal{M}(T)$ 作用到 $v \in V$ 上而 $\mathcal{M}(T)^T$ 作用到 v^T 即 $1 \times m$ 的矩阵上, 我们需要定义这些矩阵 (或者说线性映射).

定义 4.47 线性泛函, 对偶空间

- 称 V 上的线性泛函为 $\mathcal{L}(V, \mathbb{F})$ 中的元素.
- V 上所有线性泛函组成的向量空间 $\mathcal{L}(V, \mathbb{F})$ 称为 V 的对偶空间, 记作 V' .

我们来寻找 V' 的一组基. 从矩阵的角度, 容易想到的构造是: $\varphi_k(v_j) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}, k = 1, \dots, m$. 称

这组基为 v_1, \dots, v_m 的对偶基.

现在考虑构造 $\mathcal{L}(W', V')$ 里的线性映射 T' 使得 $\mathcal{M}(T') := \mathcal{M}(T', (\psi_1, \dots, \psi_n), (\varphi_1, \dots, \varphi_m))$ 满足 $\mathcal{M}(T') = \mathcal{M}(T)^T$, 其中 ψ_1, \dots, ψ_n 是 w_1, \dots, w_n 的对偶基. 设 $Tv_i = c_{1i}w_1 + \dots + c_{ni}w_n$, 那么 $\psi_jTv_i = c_{ji}$, 进而可以排布出理想中的 $\mathcal{M}(T')$:

$$\mathcal{M}(T') = \begin{pmatrix} \psi_1Tv_1 & \cdots & \psi_nTv_1 \\ \vdots & & \vdots \\ \psi_1Tv_m & \cdots & \psi_nTv_m \end{pmatrix}.$$

也就是说, 对 $k = 1, \dots, n$, $T'(\psi_k) = (\psi_kTv_1)\varphi_1 + \dots + (\psi_kTv_m)\varphi_m$. 左右同时作用于 $j = 1, \dots, m$ 可得 $T'(\psi_k)(v_j) = \psi_kTv_j$. 因此, 我们定义 $T'(\psi) = \psi \circ T$ 即可.

定义 4.48 对偶映射

设 $T \in \mathcal{L}(V, W)$, 则 T 的对偶映射 $T' \in \mathcal{L}(W', V')$ 满足 $T'(\psi) = \psi \circ T$.

容易验证, 将映射变为对偶映射的映射 $(\cdot)'$ 满足下方代数运算性质.

命题 4.49

- 加法. 对所有 $S, T \in \mathcal{L}(V, W)$, $(S + T)' = S' + T'$.
- 标量乘法. 对所有 $\lambda \in \mathbb{F}$ 和 $T \in \mathcal{L}(V, W)$, $(\lambda T)' = \lambda T'$.
- 复合. 对所有 $T \in \mathcal{L}(U, V)$ 和 $S \in \mathcal{L}(V, W)$, $(ST)' = T'S'$.

当然, 这些性质可以直接从定义推导出来, 无需借助上方的结论. 另一方面, 上述命题也向我们阐明了 $(\cdot)'$ 和 $(\cdot)^T$ 的统一性.

4.5.2 矩阵的秩

现在我们暂时从对偶理论中跳脱出来, 看看如何衡量一个矩阵 (所对应线性映射) 的零空间和值域.

定义 4.50 矩阵的秩

设矩阵 A .

- 称 A 的列向量的张成空间为 A 的列空间, 记作 $\mathcal{R}(A)$. $\{x : Ax = 0\}$ 称为 A 的零空间, 记作 $\mathcal{N}(A)$.
- 称 $\dim \mathcal{R}(A)$ 为 A 的列秩, 记作 $\text{rank } A$. 称 $\dim \mathcal{R}(A^T)$ 为 A 的行秩.

注. 若 A 对应线性映射 T , 则 $\mathcal{R}(A) = \text{range } T$, $\mathcal{N}(A) = \text{null } T$.

注. 后面我们会说明, 矩阵的行秩和列秩其实是相等的. 这里我们先用列秩定义秩.

我们注意到, 例如, 设 A 是 $m \times n$ 矩阵, 那么 $\text{rank } A \leq n$ (因为 A 只有 n 列) 并且 $\text{rank } A \leq m$ (因为每个列向量只有 m 个坐标).

命题 4.51 行列分解

设 A 是 $m \times n$ 矩阵, $c = \text{rank } A \geq 1$, 则存在 $m \times c$ 矩阵 C 和 $c \times n$ 矩阵 R 使得 $A = CR$.

证明 设 A 的列向量 v_1, \dots, v_n , 取其极大线性无关组 v_{k_1}, \dots, v_{k_c} , 令 $C = (v_{k_1} \ \dots \ v_{k_c})$. 设 $v_j = a_{1j}v_{k_1} + \dots + a_{cj}v_{k_c}$, $j = 1, \dots, n$, 令 $R = (a_{ij})$. 容易验证 $A = CR$.

上述命题直接的推论:

命题 4.52 矩阵的行秩等于列秩

设 A 是 $m \times n$ 矩阵, 则 $\text{rank } A = \text{rank } A^T$.

证明 设 A 的秩为 c , 则存在 $m \times c$ 矩阵 C 和 $c \times n$ 矩阵 R 使得 $A = CR$. 和上个命题恰好相反地, 容易验证 A 的行都可由 R 的所有行线性表出 (其中系数取自 C). 从而 $\text{rank } A^T \leq \text{rank } A$. 对 A^T 作同样的操作可以证明 $\text{rank } A \leq \text{rank } A^T$, 所以原命题成立.

从矩阵计算的角度来看, 可以按如下的方式计算矩阵的秩以及零空间的一组基: 容易说明, $\text{rref}A$ 和 A 的列向量具有相同的特征, 即若 A 中某个列向量能用其他线性无关的列向量线性表出, 则在 $\text{rref}A$ 中对应的列向量也能被对应的列向量组线性表出, 且表示系数相同. 因此, $\text{rref}A$ 的阶梯数就是 $\text{rank } A$, 且我们能够通过 $\text{rref}A$ 简单地得到解空间的结构.

$$A = \begin{pmatrix} 1 & 1 & 4 & 2 \\ 5 & 1 & 4 & 4 \\ 2 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{化简}} \text{rref}A = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 4 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

在上例中, 设 A 的列向量为 v_1, \dots, v_4 , 则可得 $v_3 = 4v_2, v_4 = \frac{1}{2}v_1 + \frac{3}{2}v_2$.

另一方面, 作初等行变换时也不会改变行空间 \mathcal{R}^T (但是行向量之间的关系会改变), 这样我们就得到了 $\text{rank } A = \text{rank } A^T$ 的另一种证明.

定义 4.53 满秩

设 $m \times n$ 矩阵 A . 当 $\text{rank } A = m$ 时, 称 A 行满秩; 当 $\text{rank } A = n$ 时, 称 A 列满秩; 当 $\text{rank } A = m = n$ 时, 称 A 满秩.

容易验证, 设 T 是 A 所对应的线性映射, 则 T 是单射当且仅当 A 列满秩, T 是满射当且仅当 A 行满秩, T 是双射当且仅当 A 满秩. 将这一观点结合行列分解, 可以得到和命题31类似的想法.

命题 4.54

设 $m \times n$ 矩阵 A 和 $n \times p$ 矩阵 B , 则 $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$, 等号在 B 可逆时取得.

证明 将 A, B 视作线性映射, 结论是显然的.

命题 4.55 矩阵乘法和秩的关系

设 $m \times n$ 矩阵 A 和 $n \times p$ 矩阵 B , m 阶可逆矩阵 P , n 阶可逆矩阵 Q .

- $\text{rank}(AB) \leq \min\{\text{rank } A, \text{rank } B\}$.
- $\text{rank}(PAQ) = \text{rank } A$.

证明 这是上个命题的直接推论 (需要用矩阵转置说明两个方面的关系).

4.5.3 对偶映射的零空间和值域

我们先尝试来猜测对偶映射的零空间和值域的结构. 设 $T' \in \mathcal{L}(W', V')$, 则

$$T'(\psi) = 0 \Leftrightarrow \psi \circ T = 0 \Leftrightarrow \forall v \in V, \psi(Tv) = 0 \Leftrightarrow \text{range } T \subseteq \text{null } \psi.$$

即 $\text{null } T' = \{\psi \in W' : \text{range } T \subseteq \text{null } \psi\}$.

另一方面, 设 $\text{null } T$ 的一组基为 v_1, \dots, v_m , 扩充为 V 的一组基 $v_1, \dots, v_m, u_1, \dots, u_n$; $\text{range } T$ 的一组基 Tu_1, \dots, Tu_n , 扩充为 W 的一组基 $Tu_1, \dots, Tu_n, w_1, \dots, w_p$.

则 T 关于 V 的这组基的矩阵为

$$\mathcal{M}(T) = \begin{pmatrix} 0 & \cdots & 0 & \cdot & \cdots & \cdot \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdot & \cdots & \cdot \end{pmatrix}.$$

而当 ψ 遍历 W' 时,

$$\mathcal{M}(\psi) \cdot \mathcal{M}(T) = \begin{pmatrix} \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \end{pmatrix}_{1 \times (n+p)} \cdot \begin{pmatrix} 0 & \cdots & 0 & \cdot & \cdots & \cdot \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdot & \cdots & \cdot \end{pmatrix}_{(n+p) \times (m+n)} = \begin{pmatrix} 0 & \cdots & 0 & \cdot & \cdots & \cdot \end{pmatrix}_{1 \times (m+n)}$$

容易验证, $\psi \circ T$ 遍历所有使得 $\text{null } T \subseteq \text{null } \varphi$ 的 $\varphi \in V'$, 即 $\text{range } T' = \{\varphi \in V' : \text{null } T \subseteq \text{null } \varphi\}$.

将以上两个结论放在一起, 自然引出下方定义:

定义 4.56 零化子

设 U 是 V 的子空间, 那么定义 U 的零化子 U^0 如下:

$$U^0 := \{\varphi \in V' : U \subseteq \text{null } \varphi\}.$$

容易验证, U^0 是 V' 的子空间. 我们可以将上方的两个结论写成:

命题 4.57 对偶映射的零空间与值域

设有限维向量空间 V, W , $T \in \mathcal{L}(V, W)$, 则

$$\text{null } T' = (\text{range } T)^0, \quad \text{range } T' = (\text{null } T)^0.$$

注. 实际上, 第一个结论的得出并不需要有限维的假设, 参见先前的证明过程. 第二个结论需要, 这是因为在验证“遍历”的时候我们隐含使用了有限维的假设 (否则没办法把线性映射写成矩阵).

为了得到 $\text{null } T', \text{range } T'$ 具体的维数, 我们需要研究 U^0 的性质.

命题 4.58 零化子的维数

设 V 是有限维的, U 是 V 的子空间, 则 $\dim U + \dim U^0 = \dim V$.

证明 构造嵌入映射 $\iota \in \mathcal{L}(U, V)$, 使得 $\iota|_U = I_U$. 那么 $\iota' \in \mathcal{L}(V', U')$, 满足 $\text{null } \iota' = U^0$. 由线性映射基

本定理,

$$\dim \operatorname{range} \iota' + \dim U^0 = \dim V' = \dim V.$$

另外, 对任意 $\varphi \in U'$, 存在 $\varphi_1 \in V'$ 使得 $\varphi_1|_U = \varphi$, 而 $\iota'(\varphi_1) = \varphi$, 从而 $\operatorname{range} \iota' = U$. 这就完成了证明.

因此可以得到: $\dim \operatorname{null} T' = \dim \operatorname{null} T + \dim W - \dim V$, $\dim \operatorname{range} T' = \dim \operatorname{range} T$. 因此, T 是满射当且仅当 T' 是单射, T 是单射当且仅当 T' 是满射.

利用对偶, 我们可以给出 $\operatorname{rank} A = \operatorname{rank} A^T$ 的第三种证明: 设 T 是 A 对应的线性映射, 那么

$$\operatorname{rank} A = \dim \operatorname{range} T = \dim \operatorname{range} T' = \operatorname{rank} A^T.$$

Chapter 5

群论

5.1 基本想法

定义 5.1 二元运算

定义非空集合 S 上的二元运算 $\cdot : S \times S \rightarrow S$, 考虑 (S, \cdot) 的如下性质:

- S 是交换的, 如果对任意 $x, y \in S$ 有 $xy = yx$.
- S 是结合的, 如果对任意 $x, y, z \in S$ 有 $(xy)z = x(yz)$.
- S 上的一个幺元 (单位元) e , 满足对任意 $x \in S$ 都有 $xe = ex = e$.
- x 的一个逆元 x^{-1} , 满足 $xx^{-1} = x^{-1}x = e$.

注. 不造成歧义时, $x \cdot y$ 可简写为 xy , (S, \cdot) 可简写为 S . 括号表示优先的运算.

注. 不难 (并且应当) 验证逆元 x^{-1} (相对于 x) 和幺元 e (相对于 S) 的唯一性.

注. 后面会看到, 这些性质的常见程度为结合性质 \sim 存在幺元 $>$ 所有元素存在逆元 $>$ 交换性质.

我们引入自然的 x^n 定义, 其中 n 是整数.

定义 5.2 幺半群, 群

设带有二元运算 \cdot 的非空集合 S , 称 (S, \cdot) 为幺半群, 若运算 \cdot 满足结合律且 S 中存在幺元. 进一步, 称幺半群 G 为群, 如果其所有元素均可逆.

注. 关于交换性质的扩展定义: 称交换的幺半群为交换幺半群, 交换的群为 *Abel* 群.

定义 5.3 子幺半群, 子群

对于幺半群 S 和非空子集 $T \subset S$, 称 T 是一个子幺半群, 如果

- $e \in T$.
- T 对于 \cdot 封闭, 即 \cdot 在 T 上的限制映射之值域包含于 T .

对于群 G 和非空子集 $H \subset G$, 称 H 是一个子群 (记作 $H < G$), 如果

- $e \in H$.
- H 对于 \cdot 封闭.
- H 对于取逆元映射 \cdot^{-1} 封闭.

注. 子群的后两个条件可以合并为一个: 对任意 $x, y \in H$ 有 $xy^{-1} \in H$.

例 5.1 么半群中所有可逆元素构成的子么半群是一个群.

例 5.2 记 n 阶实矩阵构成集合 $M(n, \mathbb{R})$, 该集合对矩阵乘法构成么半群. 考虑其子集 $GL(n, \mathbb{R}) := \{A \in M(n, \mathbb{R}) : \det A \neq 0\}$ 和 $SL(n, \mathbb{R}) := \{A \in M(n, \mathbb{R}) : \det A = 1\}$, 它们对矩阵乘法构成群, 分别称作一般线性群和特殊线性群.

例 5.3 对于集合 X , 考虑所有 $X \rightarrow X$ 的双射所成集合 $\text{Sym } X$, 这个集合对映射复合构成群, 称为对称群 (置换群). 参考后文.

记群 G 的阶 $\text{ord } G := |G|$, 当 G 为无限集合时 $|G|$ 表示其基数.

在 G 中任取集合 E , 考虑 E 所生成的“闭包”, 即包含 E 的最小子群 $\langle E \rangle := \bigcap_{H < G, E \subset H} H$. 容易证明, 对于有限集合 $E = \{a_1, \dots, a_n\}$, $\langle E \rangle = \{a_1^{\alpha_1} \cdots a_n^{\alpha_n} : \alpha_j \in \mathbb{Z}, j = 1, \dots, n\}$. 特别地, 当 E 是单点集 $\{x\}$ 时, 简记 $\langle x \rangle = \langle \{x\} \rangle$. 此时定义 $\text{ord } x := \text{ord } \langle x \rangle$ 为 x 的阶. 亦可证明, $\text{ord } x$ 有限时即为最小的使得 $x^n = e$ 的正整数 n .

例 5.4 对于群 G , 若存在 $x \in G$ 使得 $G = \langle x \rangle$, 则称 G 为一个循环群. 下一节会研究循环群的结构.

我们采用惯常的记号 AB 表示 $\{ab : a \in A, b \in B\}$, 并简记 $\{x\}A$ 为 xA , Bx 同理.

类似于利用等价类和商集对集合进行划分的方法, 这里可以考虑一个群 G 的划分. 选取其任一子群 H , 我们希望利用 H 的某一特征划分 G . 自然的想法是作出 GH 或者 HG 并设法去除其中重复的集合. 这就是定义陪集的动机:

定义 5.4 陪集

设 H, K 为群 G 的子群.

- 定义左陪集为 G 中形如 xH 的子集, 并记全体左陪集构成集合 G/H . 类似可得右陪集和 $G \setminus H$ 的定义.
- 定义双陪集为 G 中形如 HxK 的子集, 并记全体双陪集构成集合 $H \setminus G/K$.
- 定义 H 在 G 中的指数 $(G : H) := \text{ord } G/H$.

注. 在三种陪集中, 我们也许更偏爱左陪集, 因为左乘一个元素可以自然地视作映射 $H \rightarrow G, h \mapsto xh$.

注. 左右陪集是双陪集的特例, 即 H 或 K 为 $\{e\}$ 的情况.

注. 令 $x \sim y$ 当且仅当 $HxK = HyK$, 那么这是一个等价关系.

注. 可以证明 $\text{ord } G/H = \text{ord } G \setminus H$. 因此讨论指数时无需指定左或右.

自然想到, 取映射 $\tau_K : K \rightarrow K, k \mapsto y^{-1}xk$, 则 $xK = yK$ 等价于 τ_K 是双射. 同理可得, 令 $\tau_H : H \rightarrow H, h \mapsto h y x^{-1}$, 则 $Hx = Hy$ 等价于 τ_H 是双射. 最后, $HxK = HyK$ 等价于 $\tau_H \circ \tau_K$ 是双射.

更进一步, 以 τ_K 为例: 若想要 τ_K 是双射, 由于其本身就是单射, 故只要求满射, 一个充分条件是 $y^{-1}x \in K$. 反过来, 当 $xK = yK$ 时自然有 $xe = yk$ 即 $y^{-1}x = k \in K$. 这就证明了下面的命题.

另一种证明该命题的方法是考虑 x 驱使 H “平移”的作用. 如果将 H 想象成向量空间, x 想象成向量并取加法为 \cdot , 这一点会非常直观.

命题 5.5

设 H 为群 G 的子群, 则对于 $x \in G$ 有 $xH = H \Leftrightarrow x \in H$. 进而对于 $x, y \in G$ 有 $xH = yH \Leftrightarrow y^{-1}x \in H$. 右陪集和双陪集的情况同理.

证明 必要性: 取 $e \in H$ 即得 $x = xe \in xH = H$. 充分性: 由乘法封闭性可知 $xH \subset H$. 同理有 $x^{-1}H \subset H$, 即 $H \subset xH$. 这说明 $xH = H$.

现在考虑用陪集划分群 G :

命题 5.6

设 H, K 为群 G 的子群, 则

- $G = \bigsqcup_x HxK$, 其中 x 遍历每个双陪集的代表元.
- (Lagrange 定理) $\text{ord } G = (G : H) \text{ord } H$.

证明 (1) 先证明若 HxK, HyK 有交则相等: 记 $h x k = h' y k'$, 则 $x = h^{-1} h' y k' k^{-1} \in HyK$, 由等价关系的传递性可知 $HxK = HyK$. 这就证明了无交部分. 并部分则是显然的.

(2) 将 G 进行无交并分解即 $G = \bigsqcup_x xH$, 记 $E = \{e\}$, 则 $G = \bigsqcup_x x \bigsqcup_y yE = \bigsqcup_{x,y} xyE$. 注意到对于 G, H , $(G : H)$ 就是可能的 x 的个数, 而 $(G : 1) = \text{ord } G, (H : 1) = \text{ord } H$, 结合 x, y 个数是 x 个数与 y 个数之积, 立得原式成立.

注. 在 (2) 中实际上证明了: 若 $K < H < G$, 则 $(G : K) = (G : H)(H : K)$. 当然这直接由 Lagrange 定理可得.

Chapter 6

多项式与环

Part III

DLC1-范畴论

Part IV

DLC2-线性代数

Chapter 7

线性方程组与行列式计算

Chapter 8

对角化

引子一

设 $T \in \mathcal{L}(V)$, 设 V 的一个直和分解 $V = V_1 \oplus \cdots \oplus V_m$, 则在研究 T 时我们只需要知道每个 $T|_{V_k}$ 的行为即可. 但是, $T|_{V_k}$ 不一定是 V_k 上的算子, 这样很多有效的工具就没有作用了. 因此, 我们要考虑是否存在 V 的一种直和分解, 使得对所有 k , $T|_{V_k} \in \mathcal{L}(V_k)$.

一般地, 设 U 是 V 的子空间, $T \in \mathcal{L}(V)$. 称 U 是 T 下的一个不变子空间, 如果 $T(U) \subseteq U$. 显然不变子空间的交与和都是不变子空间, 因此可以考虑最简单的不变子空间, 即由向量 v 张成的一维子空间. 此时 $Tv \in \text{span}(v)$ 说明存在 $\lambda \in \mathbb{F}$ 使得 $Tv = \lambda v$. 反过来, 若上式成立, 则 $\text{span}(v)$ 自然是 T 的一个不变子空间.

从矩阵的角度来看, 若想要让 V 分解为若干个一维不变子空间的直和, 就是要寻找一组基使得 T 的表示矩阵为对角矩阵, 此时我们称算子 T 是可对角化的.

引子二¹

我们来看这样一个例子: 甲地和乙地之间每年都会有人口的流动. 设每年甲地向乙地转移其人口的 40%, 乙地向甲地转移其人口的 10%, 那么当足够久之后两地的人口比例是否会趋于稳定? 用矩阵变换来描述这个问题, 就是说对任意的 $0 < t < 1$, 当 $n \rightarrow \infty$ 时式子

$$\begin{pmatrix} 0.6 & 0.1 \\ 0.4 & 0.9 \end{pmatrix}^n \cdot \begin{pmatrix} t \\ 1-t \end{pmatrix}$$

是否存在极限?

一方面, 由压缩映射原理, 我们可以猜测稳定状态下的结果 x 就是 $Ax = x$ 的解, 而后者解集为向量 $\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$ 的张成空间. 另外不难验证对任意 n , 最终 x 的两个坐标之和为 1. 因此我们猜出了稳定的情况 $x_0 = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$.

¹摘编自 梁鑫, 田垠, 杨一龙. 线性代数入门. 清华大学出版社, 2022

另一方面, 设 A 是上方的矩阵, 我们要考虑怎样选取一组基才能将 A 变为对角矩阵 (因为对角矩阵的乘积就是直接将对角线上元素相乘). 这就是说, 要寻找 x_1, x_2 使得 $Ax_1 = \lambda_1 x_1$ 和 $Ax_2 = \lambda_2 x_2$. 一般地, 考虑方程 $(A - \lambda I)x = 0$. 我们知道该方程有非零解当且仅当 $\text{rank } A = 1$, 从而

$$\text{rank } A = 1 \Leftrightarrow \frac{0.6 - \lambda}{0.4} = \frac{0.1}{0.9 - \lambda} \Leftrightarrow \lambda^2 - 1.5\lambda + 0.5 = 0 \Leftrightarrow \lambda_1 = 1, \lambda_2 = 0.5.$$

计算可得, λ_1, λ_2 对应的 x_1, x_2 分别为 $\begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. 记 $X = \begin{pmatrix} x_1 & x_2 \end{pmatrix}$. 那么

$$A^n X = \begin{pmatrix} A^n x_1 & A^n x_2 \end{pmatrix} = \begin{pmatrix} x_1 & 0.5^n x_2 \end{pmatrix} = X \begin{pmatrix} 1 & \\ & 0.5 \end{pmatrix}^n.$$

由 x_1, x_2 线性无关知 X 可逆, 那么

$$A^n = \begin{pmatrix} 0.2 & 1 \\ 0.8 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & \\ & 0.5^n \end{pmatrix} \cdot \begin{pmatrix} 0.2 & 1 \\ 0.8 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 0.2 + \frac{0.8}{2^n} & 0.2 - \frac{0.2}{2^n} \\ 0.8 - \frac{0.8}{2^n} & 0.8 + \frac{0.2}{2^n} \end{pmatrix}.$$

由数分的知识, 一个矩阵的极限就等于其所有分量极限的矩阵. 因此 $A^n \rightarrow \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix}$, 从而 $A^n \begin{pmatrix} t \\ 1-t \end{pmatrix} \rightarrow \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$. 这与我们的猜想是一致的.

8.1 基本概念

总结以上两个引子, 我们希望找到一组基使得 $T \in \mathcal{L}(V)$ 的表示矩阵为对角矩阵, 这直接等价于寻找 λ 和 v 使得 $Tv = \lambda v$. 于是引出如下定义:

定义 8.1 本征值, 本征向量

设 $T \in \mathcal{L}(V)$. 称 $\lambda \in \mathbb{F}$ 为 T 的一个本征值, 如果存在 $V \ni v \neq 0$ 使得 $Tv = \lambda v$, 同时这样的 v 称作 λ 的一个本征向量.

在 \mathbb{R}^n 平面上, $Tv = \lambda v$ 的几何意义就是 Tv 和 v 共线.

为了更好表示本征值的本征向量, 我们可以定义本征空间的概念: 设 $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{F}$, 则 T 对应 λ 的本征空间定义为

$$E(\lambda, T) := \text{null}(T - \lambda I).$$

显然, 若 $E(\lambda, T) \neq \{0\}$, λ 是 T 的本征值, $E(\lambda, T)$ 是所有 λ 的本征向量的集合.

例 8.1 设旋转变换 $R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. 若 $\theta \neq k\pi$, 在 \mathbb{R}^2 上显然不存在 v 使得 Rv 和 v 共线. 从矩

阵的角度, $(R - \lambda I)v = 0$ 有解等价于

$$\frac{\cos \theta - \lambda}{\sin \theta} = \frac{-\sin \theta}{\cos \theta - \lambda} \Leftrightarrow \lambda^2 - 2\cos \theta \lambda + 1 = 0 \Leftrightarrow \Delta = 4(\cos^2 \theta - 1) \geq 0.$$

而这是不可能的. 但是另一方面, 若考虑 \mathbb{C}^2 , 则方程有两解 $\lambda_{1,2} = \cos \theta \pm i \sin \theta$.

上面的例子说明, 在 \mathbb{R}^n 和 \mathbb{C}^n 上我们可能得到完全不同的结果.

显然, 若 v 是 λ 的本征向量, 则 $kv, k \in \mathbb{F}$ 也是. 因此不难得到下方结论:

命题 8.2

设 $T \in \mathcal{L}(V)$, $\lambda_1, \dots, \lambda_m$ 是 T 的本征值, v_j 分别是 λ_j 对应的本征向量 ($j = 1, \dots, m$), 则 v_1, \dots, v_m 线性无关.

证明 设若不然, 取 v_1, \dots, v_m 的一个极大线性无关组, 不妨设为 v_1, \dots, v_k . 记 $v_m = a_1 v_1 + \dots + a_k v_k$, 进而

$$Tv_m = a_1 Tv_1 + \dots + a_k Tv_k \Leftrightarrow \lambda_m v_m = a_1 \lambda_1 v_1 + \dots + a_k \lambda_k v_k.$$

消元可得 $a_1(\lambda_1 - \lambda_m)v_1 + \dots + a_k(\lambda_k - \lambda_m)v_k = 0$, 显然矛盾.

该命题的一个直接推论是: 不同本征值的个数不能超过向量空间的维数.

现在来看一些平凡的子空间. 设 $T \in \mathcal{L}(V)$, 则

$$\text{null } T, \quad \text{range } T, \quad \{0\}, \quad V,$$

都是 T 下的不变子空间. 一般地, 设多项式 $p(z) = a_0 + a_1 z + \dots + a_n z^n$, 若定义关于算子 $T \in \mathcal{L}(V)$ 的多项式 $p(T) := a_0 + a_1 T + \dots + a_n T^n$, 则有:

命题 8.3

设 $T \in \mathcal{L}(V)$, $P \in \mathcal{P}(\mathbb{F})$, 则 $\text{null } p(T)$ 和 $\text{range } p(T)$ 是 T 下的不变子空间.

证明 (1) 设 $p(T)u = 0$, 则 $p(T)(Tu) = T(p(T)u) = 0$, 即是说 $Tu \in \text{null } p(T)$.

(2) 设存在 $v \in V$ 使得 $u = p(T)v$, 则 $Tu = T(p(T)v) = p(T)(Tv)$, 即是说 $Tu \in \text{range } p(T)$.

最后来看一个有用的工具:

定义 8.4 商算子

设 $T \in \mathcal{L}(V)$, U 是 T 下的不变子空间. 定义商算子 $T/U \in \mathcal{L}(V/U)$ 使得 $(T/U)(v + U) = Tv + U$.

容易验证, 若 $v + U = w + U$, 则 $v - w \in U$, 进而 $Tv - Tw \in U$, 于是 $Tv + U = Tw + U$. 这就是说 T/U 是良定义的.

8.2 极小多项式

8.2.1 本征值与极小多项式

正如前文所述, 在 \mathbb{R} 和 \mathbb{C} 上的向量空间关于本征值的存在性可能不一样. 下文全部讨论 \mathbb{C} 上的向量空间.

命题 8.5 复向量空间上的算子总存在本征值

设 V 是 \mathbb{C} 上的有限维向量空间, 则任意的 $T \in \mathcal{L}(V)$ 都存在本征值.

证明 设 $v \neq 0$, $n = \dim V$, 则 $v, Tv, \dots, T^n v$ 线性相关, 从而存在非零多项式 p 使得 $p(T)v = 0$, 特别地我们要求 p 的次数是使得该式成立最小的 (这句话是先验的). 由代数基本定理, 存在 $\lambda \in \mathbb{C}$ 使得 $p(\lambda) = 0$, 进而存在多项式 q 使得 $p(T) = (T - \lambda I)q(T)$, 这说明 $(T - \lambda I)(q(T)v) = 0$. 由 p 次数的最小性, $q(T)v \neq 0$, 于是 λ 是 T 的一个本征值, $q(T)v$ 是 λ 的本征向量.

去掉有限维的假设是不可行的, 反例如:

例 8.2 设 $T \in \mathcal{L}(\mathcal{P}(\mathbb{C}))$ 使得 $(Tp)(z) = zp(z)$. 容易发现对于非零多项式 p , Tp 的次数会严格大于 p 的次数, 因此 Tp 不能表示为 p 的标量倍.

在上方命题的证明中, 我们发现使得 $p(T)v = 0$ 的多项式 p 起到了关键作用. 实际上, 我们可将其扩展为更强的命题:

命题 8.6 极小多项式的存在性和唯一性

设 V 是有限维向量空间, $T \in \mathcal{L}(V)$, 则存在首一多项式 $p \in \mathcal{P}(\mathbb{F})$ 使得 $p(T) = 0$. 特别地, 称满足该式的次数最小的多项式为极小多项式, 那么极小多项式唯一且次数 $\deg p \leq \dim V$.

注. 这并不是说极小多项式不存在重根, 后面会看到具体例子.

证明 (1) 用归纳法. 当 $\dim V = 0$ 时取 $p(z) = 1$ 即可; $\dim V = 1$ 时, 任取 $v \neq 0$, 由 v, Tv 线性相关可知存在首一多项式 $p \in \mathcal{P}_1(\mathbb{F})$ 使得 $p(T)v = 0$. 记 $p(z) = z + a_0$, 那么 $Tv + a_0v = 0$. 另一方面, 熟知存在 λ 使得 $Tv = \lambda v$ 对所有 $v \in V$ 成立, 因此 $\lambda = -a_0$, 即是说 $p(T)v = 0$ 对所有 $v \in V$ 成立.

设 $\dim V = n + 1$, 假设当 $\dim V \leq n$ 时命题均成立. 任取 $v \neq 0$, 由 $v, Tv, \dots, T^{n+1}v$ 线性相关可知存在最小的 m 使得 $T^m v$ 能被 $v, Tv, \dots, T^{m-1}v$ 线性表出 (也就是说 $v, Tv, \dots, T^{m-1}v$ 是极大线性无关组). 由此构造首一 m 次多项式 p 使得 $p(T)v = 0$.

我们首先注意到对任意 $k \geq 0$, $p(T)(T^k v) = T^k(p(T)v) = 0$, 所以 $\dim \text{null } p(T) \geq m$. 另外, 由之前的命题, $\text{range } p(T)$ 是 T 下的不变子空间, 故 $T|_{\text{range } p(T)} \in \mathcal{L}(\text{range } p(T))$, 进而可以应用归纳假设, 也就是说存在首一多项式 q 使得 $\deg q (\leq \dim \text{range } p(T) \leq n + 1 - m)$ 是满足 $q(T|_{\text{range } p(T)}) = 0$ 的最小数.

最后只需验证, 对任意 $v \in V$, $qp(T)(v) = q(T)(p(T)v) = 0$ 且 $\deg qp = \deg q + \deg p \leq n + 1$, 立得 qp 是满足要求的极小多项式.

(2) 假设存在不同的 n 次首一多项式 p, q 使得 $p(T) = 0, q(T) = 0$ 且 n 是满足该式的最小正整数. 设 $p - q$ 的首项系数为 $\lambda \neq 0$, 那么 $\deg(\frac{p-q}{\lambda}) < n$ 和 $(\frac{p-q}{\lambda})(T) = 0$, 与 n 的最小性矛盾. 于是 $p - q = 0$.

在实际计算极小多项式时, 一般采用如下策略: 取 $v \in V$, 考虑方程 $c_0v + c_1Tv + \cdots + c_{n-1}T^{n-1}v + T^nv = 0$, 其中 $n = \dim V$. 如果该方程存在唯一解 c_0, \cdots, c_n , 则极小多项式的系数就是这组解. 这对于表示矩阵中 0 的个数较多的线性映射是很方便的.

例 8.3 设 $T \in \mathcal{L}(\mathbb{F}^5)$, 关于标准基的矩阵是 $\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$. 求 T 的极小多项式.

解. 考虑方程 $c_0e_1 + c_1Te_1 + \cdots + T^5e_1 = 0$, 注意到

$$Te_1 = e_2, \quad T^2e_1 = Te_2 = e_3, \quad T^3e_1 = Te_3 = e_4, \quad T^4e_1 = Te_4 = e_5, \quad T^5e_1 = Te_5 = -3e_1 + 6e_2,$$

因此方程等价于 $(c_0 - 3)e_1 + (c_1 + 6)e_2 + c_2e_3 + c_3e_4 + c_4e_5 = 0$. 由 e_1, \cdots, e_5 线性无关可知方程的唯一解是 $c_0 = 3, c_1 = -6, c_2 = c_3 = c_4 = 0$. 因此 T 的极小多项式就是 $p(z) = 3 - 6z + z^5$.

根据命题5的证明过程可以得到:

命题 8.7

设有限维复向量空间 V 上的算子 $T \in \mathcal{L}(V)$, 设 T 的本征值为 $\lambda_1, \cdots, \lambda_m$, 则 T 的极小多项式是 $p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$.

注. 通过合并一些因式, 可以将该命题推广到实数的情况.

注. 从这个命题可以看出, 一些矩阵的特征值是不可求 (解析) 解的, 例如上个例子的五次方程.

命题 8.8

设有限维向量空间 V 上的算子 $T \in \mathcal{L}(V)$, 则多项式 q 满足 $q(T) = 0$ 当且仅当它能被 T 的极小多项式乘除.

证明 充分性是显然的. 必要性: 设 T 的极小多项式 p , 令 $q = ps + r$, 其中 $\deg r < \deg p$, 那么 $0 = q(T) = p(T)s(T) + r(T) = r(T)$. 若 $r \neq 0$, 不妨让 r 的首项系数为 1, 即得到矛盾. 因此 $p \mid q$.

推论 8.9

设有限维向量空间 V 上的算子 $T \in \mathcal{L}(V)$, $U \subseteq V$ 是 T 下的不变子空间. 那么 T 的极小多项式被 $T|_U$ 的极小多项式整除.

命题 8.10

设有限维向量空间 V 上的算子 $T \in \mathcal{L}(V)$, 则 T 不可逆当且仅当 T 的极小多项式常数项为 0.

8.2.2 奇数维实向量空间上的本征值**引理 8.11**

设有限维实向量空间 V 上的算子 $T \in \mathcal{L}(V)$, $b, c \in \mathbb{R}$ 满足 $b^2 < 4c$. 则 $\dim \text{null}(T^2 + bT + cI)$ 是偶数.

证明 由于 $\text{null}(T^2 + bT + cI)$ 是 T 下的不变子空间, 不妨令 $V = \text{null}(T^2 + bT + cI)$. 设 $T^2 + bT + cI = 0$, 假设 λ 是 T 的本征值, $v \neq 0$ 是 λ 的本征向量, 那么

$$0 = (T^2 + bT + cI)v = (\lambda^2 + b\lambda + c)v.$$

从而 $\lambda^2 + b\lambda + c = 0$, 矛盾. 因此 T 不存在本征值.

设 U 是最大的 V 的偶数维数子空间, 使得 U 在 T 下不变. 不妨考虑 $U \neq V$, 即存在 $w \in V - U$. 取 $W = \text{span}(w, Tw)$, 那么 $T(Tw) = -bTw - cw \in W$, 即 W 在 T 下不变. 另外显然 $\dim W = 2$. 注意到 $U \cap W = \{0\}$, 否则 $U \cap W = \text{span}(Tw)$ 是 T 的不变子空间, 从而 T 存在本征值. 于是 $U + W$ 是 T 下的偶数维不变子空间, 与 U 的最大性矛盾. 由此得到 $U = V$.

下一个命题强化了命题5的结果:

命题 8.12 奇数维实向量空间上的算子总存在本征值

设 V 是 \mathbb{R} 上的奇数维向量空间, 则任意的 $T \in \mathcal{L}(V)$ 都存在本征值.

证明 对 V 的维数归纳证明. 当 $\dim V = 1$ 时显然成立. 假设 $\dim V \leq n$ 时成立, 考虑 $\dim V = n + 2$: 设 p 是 T 的极小多项式, 不妨设 p 没有一次因式, 即存在 $b, c \in \mathbb{R}$ 满足 $b^2 < 4c$ 且 $z^2 + bz + c \mid p(z)$, 那么存在首一多项式 q 使得

$$p(z) = q(z)(z^2 + bz + c), \quad \Rightarrow \quad 0 = p(T) = q(T)(T^2 + bT + cI).$$

这就是说, $q(T)$ 在 $\text{range}(T^2 + bT + cI)$ 上为 0, 特别地有 $\text{range}(T^2 + bT + cI) \neq V$.

另一方面, 由引理可知 $\dim \text{null}(T^2 + bT + cI)$ 是偶数, 于是 $\dim \text{range } T^2 + bT + cI$ 是奇数, 应用归纳假设可得 T 在 $\text{range}(T^2 + bT + cI)$ 上存在本征值, 进一步 T 在 V 上有本征值.

8.3 上三角矩阵

下面我们给出一个方便的结果:

引理 8.13

设 $T \in \mathcal{L}(V)$ 在 V 的一组基上具有上三角矩阵, 设对角线元素为 $\lambda_1, \dots, \lambda_n$, 则有 $(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0$.

证明 设这组基为 v_1, \dots, v_n , 则 $(T - \lambda_1 I)v_1 = 0$, 进而 $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$, 其中 $m = 1, \dots, n$. 又 $(T - \lambda_2 I)(v_2) \in \text{span}(v_1)$, 所以 $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_2 = 0$, 其中 $m = 2, \dots, n$. 递推可得对任意 v_k 都有 $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_k = 0$, 即 $(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0$.

命题 8.14 上三角矩阵的本征值

设 $T \in \mathcal{L}(V)$ 在 V 的一组基上具有上三角矩阵, 则 T 的本征值恰为对角线上的所有元素.

注. 我们并不能通过将矩阵化为阶梯型矩阵来找到其本征值, 因为在操作的过程中本征值会改变.

证明 证法一 设对角线元素分别为 $\lambda_1, \dots, \lambda_n$, 这组基为 v_1, \dots, v_n . 那么 $Tv_1 = \lambda_1 v_1$. 又 $(T - \lambda_2 I)(v_2) \in \text{span}(v_1)$, 显然 $T - \lambda_2 I$ 不是满射, 因此 λ_2 是本征值. 递推可得任意 λ_k 均为本征值.

另一方面, 由上个引理, T 的极小多项式整除 $(z - \lambda_1) \cdots (z - \lambda_n)$, 因此 T 没有除了 $\lambda_1, \dots, \lambda_n$ 以外的本征值.

证法二 只需注意到上三角矩阵可逆当且仅当对角线元素均非零. 考虑 $T - \lambda I$ 的表示矩阵, 则 λ 是 T 的本征值等价于 $\mathcal{M}(T - \lambda I)$ 不可逆, 进而等价于 λ 等于某个 λ_k .

于是可以用本征值来判断 T 是否存在上三角矩阵:

命题 8.15

设有限维实向量空间 V 上的算子 $T \in \mathcal{L}(V)$. 则 T 在某组基上存在上三角矩阵当且仅当 T 的极小多项式具有 $(z - \lambda_1) \cdots (z - \lambda_n)$ 的形式. (不要求 $\lambda_1, \dots, \lambda_n$ 两两不同)

证明 必要性由前面的命题是显然的. 充分性: 用归纳法证明. 当 $n = 1$ 时, T 的极小多项式是 $z - \lambda_1$, 即是说 $T = \lambda_1 I$, 显然对任何基 T 的矩阵都是上三角矩阵. 假设对任意 $n \leq m - 1$ 命题均成立, 下证 $n = m > 1$ 时亦成立:

证法一 令 $U = \text{range}(T - \lambda_m I)$, 则 U 是 T 下的不变子空间, 从而 $T|_U \in \mathcal{L}(U)$. 对任意 $u \in U$, 若 $(T - \lambda_m I)(v) = u$, 则

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)(u) = (T - \lambda_1 I) \cdots (T - \lambda_m I)(v) = 0.$$

这就是说 $(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I) = 0$, 于是 $T|_U$ 的极小多项式能整除该多项式. 那么, 由归纳假设, 存在 U 的一组基 u_1, \dots, u_p 使得 $T|_U$ 的表示矩阵是上三角矩阵. 将 u_1, \dots, u_p 扩展为 V 的一组基

$u_1, \dots, u_p, v_1, \dots, v_t$, 而对任意 $j = 1, \dots, t$ 有

$$Tv_j = (T - \lambda_m I)v_j + \lambda_m v_j \in \text{span}(u_1, \dots, u_p, v_1, \dots, v_j).$$

即对于基 $u_1, \dots, u_p, v_1, \dots, v_t$, T 的表示矩阵是上三角矩阵.

证法二 令 u_1 是 λ_1 的本征向量, 作 $U = E(\lambda_1, T)$ 并设其一组基为 u_1, \dots, u_p . 对任意 $v \in V$, 注意到 $(T - \lambda_2 I) \cdots (T - \lambda_m I)(v) \in U$, 于是

$$(T/U - \lambda_2 I) \cdots (T/U - \lambda_m I)(v + U) = (T - \lambda_2 I) \cdots (T - \lambda_m I)(v) + U = 0 + U.$$

即是说 $(T/U - \lambda_2 I) \cdots (T/U - \lambda_m I) = 0$. 那么 T/U 的极小多项式能整除该多项式. 于是, 由归纳假设, 存在 V/U 的一组基 $v_1 + U, \dots, v_t + U$ 使得 T/U 的表示矩阵是上三角矩阵, 也即

$$(T/U)(v_j + U) \in \text{span}(v_1 + U, \dots, v_j + U) \Leftrightarrow Tv_j \in \text{span}(u_1, \dots, u_p, v_1, \dots, v_j).$$

其中 $j = 1, \dots, t$. 显然 $u_1, \dots, u_p, v_1, \dots, v_t$ 就是 V 的一组基, 由此证毕.

需要注意, 这里没有要求 $\lambda_1, \dots, \lambda_n$ 两两不同, 例如:

例 8.4 设 $T \in \mathcal{L}(\mathbb{F}^3)$, 关于标准基的矩阵是 $\mathcal{M}(T) = \begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}$. 求 T 的极小多项式.

解. 由第一个引理, 可知极小多项式整除 $(z - 6)^2(z - 7)$, 另一方面由上方的命题可知 $6, 7$ 是 T 的本征值. 因此极小多项式可能为 $(z - 6)(z - 7)$ 或 $(z - 6)^2(z - 7)$. 取标准基, 计算可得

$$\mathcal{M}(T - 6I)\mathcal{M}(T - 7I) = \begin{pmatrix} 0 & -3 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (\mathcal{M}(T - 6I))^2\mathcal{M}(T - 7I) = O.$$

于是极小多项式为 $(z - 6)^2(z - 7)$.

上面的命题结合代数基本定理, 立得:

推论 8.16

设 V 是有限维复向量空间, $T \in \mathcal{L}(V)$. 则存在 V 的一组基使得 T 的表示矩阵是上三角矩阵.

8.4 可对角化算子

根据之前的铺垫, 下方的命题是显然的:

命题 8.17 算子可对角化的条件

设 V 是有限维向量空间, $T \in \mathcal{L}(V)$. 设 $\lambda_1, \dots, \lambda_m$ 是 T 的本征值, 则下列说法等价:

- 1) T 是可对角化的.
- 2) T 的本征向量构成 V 的一组基.
- 3) $V = E(\lambda_1, T) \oplus \dots \oplus E(\lambda_m, T)$.
- 4) $\dim V = \dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T)$.

注意到 2) 部分的直接推论: 若 T 恰好有 $\dim V$ 个本征值, 则 T 是可对角化的. 反之则不一定.

上方的命题不具有实际操作性. 不过, 结合极小多项式理论, 我们可以给出如下有用的充要条件:

命题 8.18

设 V 是有限维向量空间, $T \in \mathcal{L}(V)$. 则 T 可对角化当且仅当 T 的极小多项式具有 $(z - \lambda_1) \cdots (z - \lambda_n)$ 的形式, 其中 $\lambda_1, \dots, \lambda_n$ 两两不同.

证明 必要性显然. 充分性: 用归纳法. $n = 1$ 时可知 $T = \lambda_1 I$, 显然可对角化. 假设对 $n \leq m - 1$ 命题均成立, 下证命题对 $n = m$ 亦成立:

令 $U = \text{range}(T - \lambda_m I)$, 则 U 是 T 下的不变子空间, 从而 $T|_U \in \mathcal{L}(U)$. 对任意 $u \in U$, 若 $(T - \lambda_m I)(v) = u$, 则

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)(u) = (T - \lambda_1 I) \cdots (T - \lambda_m I)(v) = 0.$$

这就是说 $(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I) = 0$, 于是 $T|_U$ 的极小多项式能整除该多项式. 那么, 由归纳假设, 存在 U 的一组基恰好为 $T|_U$ 的本征向量.

另一方面, 显然 $\text{null}(T - \lambda_m I)$ 存在一组基恰好为 T 的本征向量. 于是只要证明 $\text{range}(T - \lambda_m I) \oplus \text{null}(T - \lambda_m I) = V$ 即得原命题成立. 实际上, 任取 u 满足 $Tu = \lambda_m u$ 和 $(T - \lambda_m I)(v) = u$, 可得

$$0 = (T - \lambda_1 I) \cdots (T - \lambda_m I)(v) = (T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)(u) = (\lambda_m - \lambda_1) \cdots (\lambda_m - \lambda_{m-1})(u).$$

由于 $\lambda_1, \dots, \lambda_m$ 两两不同, 可得 $u = 0$.

推论 8.19

设 $T \in \mathcal{L}(V)$ 可对角化, U 是 T 下的不变子空间, 则 $T|_U$ 在 U 上可对角化.