

Calculus

Unsinn

2014 年 10 月 19 日

Preface

As a student in physics school, mathematics is an essential but difficult tool for the next round of learning in your university life. Of course, the famous calculus is the first and maybe the most different one because it's your first math course here. Indeed, as I will show in this book, mathematics is an interesting and powerful subject where you can find a new world which is different from the math learned in your middle school though it is also important.

I will assume that you have a good command of the math learned in the middle school which is the base of this book. Except it, you needn't know anything about the calculus, the main contents of this book. If you feel hard to read, don't worry, it is a common thing in learning mathematics. You can find a way of yourself to deal with this problem, and it may be different for different people.

I should emphasize here that I would like to show main ideas and the importance of the theory but not its applications in this brochure. Because of its small capacity, it lost a lot of tiny details which may be not essential in the context or may cost us too many pages to prove it(details?). The most important one is that we may lose a lot of applications of the theory we built in this book. So it is not enough for you to read this book only to have an entire view on calculus. It will just take you to the new world, and never tell you what is the fastest way to get the greatest utility.

Because of a lot of reasons, there may be a big number of mistakes in this small book though I have done my best to correct them. Thus, I leave this work to you as a big exercise through this book.

In preparing this book, I received helpful suggestions from A.H. and Pachi, as well as from a lot of people who read the draft and give some useful information. Many thanks to them all. I finally wish to thank Professor T.L. Peng, his linear algebra gives me the encouragement to complete this book.

Then I will give you some tips to read this book, or indiscreetly, how to use this book.

How to use this book

Contents

0	Set and Map	1
1	Vector Space	4
1.1	n 维 Euclid 空间 \mathbb{R}^n	5
1.2	矢量空间和线性映射	8
1.3	内积和对偶空间	11
1.4	矩阵	14
1.5	正交和矢量相关概念的几何直观	17
1.6	关于线性映射的进一步讨论	23
1.7	线性映射和内积、投影变换	26
1.8	多重线性映射	28
1.9	行列式	34
1.10	基的变换	38
1.11	一些补充	40
1.12	小结	42
2	Groups and Algebra*	44
2.1	群论的一些基础概念	44
2.2	群作用和群表示	45
2.3	有限群的表示	46
2.4	对称群	46
2.5	代数	46
2.6	张量代数	46
3	Topological Space and Continuous Function	53
3.1	度量空间	54

3.1.1	度量函数	54
3.1.2	开集和闭集	56
3.2	连续函数	57
3.2.1	函数的极限	57
3.2.2	空间的完备性	57
3.2.3	\mathbb{R} 的一些基本性质	57
3.2.4	连续函数	58
3.2.5	紧集和一致连续	58
3.3	\mathbb{R}^n 中的 Heine-Borel 定理	58
3.4	拓扑空间	59
3.5	线性赋范空间	59

This page intentionally left blank.

Chapter 0

Set and Map

这章带大家回顾一下集合 (*set*) 和映射 (*map*)¹的一些概念, 这在后面是常用的, 而本章前面大部分对于读者来说是熟知的, 特别是一些集合的基本命题不再重复。而后边部分涉及一些群论的基本概念。

Definition 0.1. 如果关系 $f: A \rightarrow B$ 是一个映射, 需要满足以下两个要求

- (a). 对任意一个 $x \in A$, 都有一个 $f(x) \in B$ 。即对每一个输入值, B 中都有与之对应的输出值, 反之不然。
- (b). 如果 $x, y \in A$, 且 $f(x) = a, f(y) = b$, 则 $a = b$ 。即一个输入只能有一个输出, 反之不然。

A 称为 f 的定义域, 而 B 称为 f 的值域。

Definition 0.2. 设 $f: A \rightarrow B$ 和 $g: B \rightarrow C$, 定义映射的 $g \circ f$ 为

$$g \circ f(x) = g(f(x))$$

注意 g 和 f 的复合只有当 f 的值域等于 g 的定义域的时候才有意义。此外, 有时候我们也省略 \circ , 而直接使用 gf 来表示 g 和 f 的复合。

根据 Definition 0.1 的 (b), 多个元素可以映射到同一个元素。比如 $f(x) = x^2$ 把 2 和 -2 都映到 4。为了让我们的讨论更加清楚。

Definition 0.3. 给定一个映射 $f: A \rightarrow B$, 对于一个集合 $C \subset B$, 我们定义 C 的原象为

$$f^{-1}(C) = \{x \in A | f(x) \in C\}$$

此外, 对于 $b \in B$, 我们定义它的原象为 $f^{-1}(b) = \{x \in A | f(x) = b\}$ 。

¹后面文中出现的“函数”、“算子”、“算符”、“变换”、“对应”等都是映射的同义词。

注意 f^{-1} 不一定是一个映射。

Definition 0.4. 以下三种映射最为重要：

1. 如果 B 任意元素的原象都不空，则称映射为满射。
2. 如果 $x \neq y$ 可以推出 $f(x) \neq f(y)$ ，则称映射为单射。
3. 如果一个映射既是单射又是满射，则称他是双射。

Remark. 如果一个映射不是满射，那么原象可以是空集，如果有一个原象是空集，但我们去掉原象为空集的那些元素就构成了一个新的 B' ，这就自然地让 $f: A \rightarrow B'$ 变成了满射。

如果一个映射是双射，那么他的没一个元素都有原象且原象是一个单元素集，此时我们才可以定义映射 $f^{-1}: B \rightarrow A$ ，我们称其为 f 的逆映射。所以映射有逆等价于这个映射是双射。

下面这个命题很重要。

Proposition 0.1. 映射复合符合结合律：

$$f \circ (g \circ h) = (f \circ g) \circ h$$

Proof.

$$(f \circ g) \circ h(x) = (f \circ g)(h(x)) = f \circ g(h(x)) = f(g(h(x)))$$

$$f \circ (g \circ h)(x) = f((g \circ h)(x)) = f(g \circ h(x)) = f(g(h(x)))$$

□

因为映射复合符合结合律，所以我们可以定义 $f \circ g \circ h$ ，因为括号怎么加都不影响最后结果。

为了研究多变量的映射，我们需要集合直积的定义。

Definition 0.5. 对于两个集合 $\{V_1, V_2\}$ ，那么他们两个集合的直积 $V_1 \times V_2$ 定义为所有如下有序组的集合：

$$V_1 \times V_2 = \{(v_1, v_2) | v_1 \in V_1, v_2 \in V_2\}$$

所谓的有序组就是 (v_1, v_2) 和 (v_2, v_1) 一般不相等。

因为 $(a, (b, c))$ 和 (a, b, c) 可以看作等价，所以我们也不加区分。通过 $V_3 \times (V_1 \times V_2)$ 来定义 $V_3 \times V_1 \times V_2$ ，很容易检验结合律成立，那么我们也定义了多个集合的直积。

此外，我们通常将映射 $f: V_1 \times V_2 \rightarrow A$ 不写作 $f((a, b))$ 而是 $f(a, b)$ 。

下面定义集合间的运算（或者作用）。

Definition 0.6. 设有两个集合 X 和 Y , 则称映射 $f: X \times Y \rightarrow X$ 为集合 X 上的一个右作用。称映射 $g: Y \times X \rightarrow X$ 为集合 X 上的一个左作用。称映射 $h: X \times X \rightarrow X$ 为集合 X 上的一个二元运算。同样地, 我们可以定义多元运算。

Remark. 可以看到 \mathbb{R} 上的加法 $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ 定义为 $f(a, b) = a + b$ 是一个二元运算, 同样可以检验乘法。再比如设 Y 是集合 X 上所有双射 $f: X \rightarrow X$ 的集合, 那么映射复合构成 Y 上的一个运算。

对于一个未知的运算或者作用, 我们通常称之为“乘法”。左作用称为“左乘”, 右作用称为“右乘”。对于运算 $f(a, b)$ 通常直接记作 $a * b$, 或者再干脆一些省略中间的符号 $*$, 记作 ab , 读作 a 左乘 b 、 b 右乘 a 或 a 乘以 b 。

可以清晰地看到, 对于 X 上的二元运算的 $a * b$ 其结果 ab 也是在 X 里面的。

下面是群和域的定义, 我们打算详细研究它们。

Definition 0.7. 有一个集合 E 和其上的二元运算 $*$, 或者记作 $(E, *)$, 称为一个群 (group), 如果满足:

- (1). 结合律: 对于任意 $a, b, c \in E$, 有 $(a * b) * c = a * (b * c)$;
- (2). 单位元: 对于任意的 $a \in E$, 存在一个元素 $e \in E$, 使得 $e * a = a * e = a$;
- (3). 反元素: 对于每一个 $a \in E$, 存在一个元素 $b \in E$, 使得 $b * a = a * b = e$ 。

a 的反元素通常记作 a^{-1} 。此外如果不产生歧义, 我们可以直接称呼 E 为一个群。显然 $(\mathbb{R}, +)$ 是一个群, 单位元是 0 , $a^{-1} = -a$ 。将 \mathbb{R} 去掉了 0 之后的集合记作 $\mathbb{R} - \{0\}$, 则 $(\mathbb{R} - \{0\}, \cdot)$ 构成一个群, 单位元是 1 , $a^{-1} = 1/a$ 。

Definition 0.8. 如果有群 $(E, *)$, 对于任意的 $a, b \in E$ 满足 $a * b = b * a$, 则称 $(E, *)$ 是一个交换群, 或者称为 Abel 群 (Abelian group)。

Definition 0.9. 我们称呼一个三元组 $(\mathbb{F}, +, \cdot)$ 为一个域 (field), 如果满足下列性质:

- (1). $(\mathbb{F}, +)$ 构成一个交换群, 运算称为加法, 其中的单位元记作 0 , a 的反元素记作 $-a$;
- (2). $(\mathbb{F} - \{0\}, \cdot)$ 构成一个交换群, 运算称为乘法, 其中的单位元记作 1 , a 的反元素记作 a^{-1} ;
- (3). 分配律成立: 对所有 $a, b, c \in \mathbb{F}$, $a \cdot (b + c) = a \cdot b + a \cdot c$ 。

同样地, 如果不产生歧义, 我们可以直接称呼 \mathbb{F} 为一个域。我们刚刚已经检查过了 \mathbb{R} 满足前两条性质, 第三条分配律也是成立的, 故而 \mathbb{R} 是一个域。所以我们通常称呼其为实数域十分合理。

Chapter 1

Vector Space

空间 (space), 是一个相当有吸引力且让人浮想联翩的词汇。从各色各样的科普读物中或从中学老师口中我们可以自然地发问, 我们到底是身处三维空间还是四维空间之中? 当然, 我不会就这个有趣的问题做出回答, 而是径直地来到我们的重点, 也就是数学。

Proposition 1.1. 在数学上, 空间是集合。

这个陈述似乎和刚刚所说的空间相去甚远, 而且太过模棱两可, 并没有指出到底什么是空间。所以, 另一个自然的问题也就提出了。

Question 1.1. 在数学上, 什么样的集合是空间?

在回答这个问题之前, 先让我们看一下我们身处的三维空间 (似乎我是在回答上面的问题?), 我们会说他的三维分别是长、宽、高。或者用中学数学中熟知的三根轴 (x 轴, y 轴和 z 轴) 表示吧, 那么只要指定一个参考点 (所谓的原点), 那我们就可以写出任何一个点的坐标, 也就是用三根轴上的读数 (或者说在轴上的投影距离原点的距离) 来确定位置。如果一个物体在这个空间中, 比如这时我们就可以写出他表面上每一点的坐标。

接着需要确定的是, 这个我们熟知的空间中到底有什么? 这为下面抽象出空间这个概念提供了一个足够好的模板。首先是这个空间中的对象, 也就是“点”, 空间是这些点的集合。这符合 Proposition 1.1 所提到的。然后, 显而易见地, 只有点是不够的。就如上文所说的, 这个空间里还有“参考点”, 还有“三根轴”, 还有“距离”。可以想象, 没有这些概念, 我们不能如一般直觉一样去描述一个点的位置, 去定量比较三个点两两的远近。这些概念, 我们往往称之为空间的结构。所以, 作为对 Question 1.1 的回应, 我们有

Proposition 1.2. 在数学上, 空间是指一种具有特殊性质及一些额外结构的集合。

在具体地给出一些结构之前, 还需要明确的一点是: 我们往往称空间中的元素叫做“点”, 而事实是, 这些“点”可以是其他一些东西, 因为我们并没有明确指出集合中的

元素是什么。这就是数学的抽象，这种抽象是有必要性的，因为经过这样的抽象证明的结论，我们可以放心地应用到看似完全没关系的领域中去。比如我们完全可以把空间中的轴取做动量，或者温度，或者面积，而不单单是长、宽、高。但无论如何，数学的结论一样是可以使用的。

此外，上面说过，集合的元素是对象，自然地，我们还希望建立对象间的关系。正如我们所熟知的，函数（或者映射）这个概念就建立了元素间的关系，比如 $x \in X$ 和 $y \in Y$ ，那么我们就有函数关系 $f(x) = y$ ，这也正是我们最感兴趣的一个关系，这种关系也理所当然地建立起了集合之间的关系，这在高中数学中已经有过比较精细的考察，在第 0 章也已经有过简单的回顾。有趣的是，这种关系也可以成为对象，换言之，函数也可以成为空间中的“点”。

接下来几节，将把我们熟知的 \mathbb{R}^2 推广到 \mathbb{R}^n ，继而转变到一般的矢量空间上去，并且研究上面的线性映射，以及顺便研究一下他的矩阵和行列式。

1.1 n 维 Euclid 空间 \mathbb{R}^n

回忆一下二维矢量，回忆他的一些性质。首先给出二维矢量的定义：

Definition 1.1. 二维矢量是这样的一个有序数组 (a, b) ，其中 a 和 b 都是实数。

设有两个二维矢量 \mathbf{x}, \mathbf{y} （矢量一般用粗体标识），一个标量 $c \in \mathbb{R}$ ，其中 $\mathbf{x} = (x^1, x^2), \mathbf{y} = (y^1, y^2)$ （这是的上标不代表幂），如同我们熟知的，满足以下两个运算，加法和数乘：

$$\begin{aligned}\mathbf{x} + \mathbf{y} &= (x^1, x^2) + (y^1, y^2) = (x^1 + y^1, x^2 + y^2) \\ \mathbf{x}c &= (x^1, x^2)c = (x^1c, x^2c)\end{aligned}$$

此时，所有二维矢量的集合 $\mathbb{R} \times \mathbb{R}$ 称之为二维 Euclid 空间 \mathbb{R}^2 。同样，我们可以用完全一样的定义方法，我们可以直接把上面这些运算推广到 $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ 上去，此时 \mathbb{R}^n 称为 n 维 Euclid 空间。

虽然繁琐，但还是让我们先考察一下这样定义的 n 维 Euclid 空间 \mathbb{R}^n 的元素和结构。元素，也就是 \mathbb{R}^n 中的点，显然就是有序数组 (a, b) （比如 \mathbb{R}^2 中）。而结构，也就是我们所定义的两运算，加法和数乘。没错，这就是结构。

这样定义的空间已经可以用来做一些事情了，比如 $n = 3$ 的时候，确定实体空间中的某个点的位置。

在考察一般的矢量空间之前，要对 \mathbb{R}^n 的基有些概念。

Definition 1.2. 若 \mathbb{R}^n 上存在 n 个确定的矢量 $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$, 对每一个矢量 \mathbf{x} , 都满足以下两个条件时, 称 $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ 为 \mathbb{R}^n 的一组基, 或简写作 $\{\mathbf{u}_i\}$.

(a). 存在一组实数 $\{x^1, x^2, \dots, x^n\}$ (称之为坐标), 或简写作 $\{x^i\}$, 使得,

$$\mathbf{x} = \sum_{i=1}^n \mathbf{u}_i x^i$$

这被称为矢量关于基的分解。后面样子的表达式又叫 $\{\mathbf{u}_i\}$ 的线性组合。

(b). 若上式中 $\mathbf{x} = \mathbf{0} := (0, 0, \dots, 0)$, 则

$$x^1 = x^2 = \dots = x^n = 0$$

虽然 \mathbb{R}^n 的基不是唯一的。但 \mathbb{R}^n 中的任何一个矢量对于一组确定的基的分解应当是唯一的。

证明是简单的, 令 $\{x^i\}$ 与 $\{x'^i\}$ 是 \mathbf{x} 对于 $\{\mathbf{e}_i\}$ 的两组坐标, 那么,

$$\mathbf{x} = \sum_{i=1}^n \mathbf{u}_i x^i = \sum_{i=1}^n \mathbf{u}_i x'^i$$

移项有,

$$\mathbf{0} = \sum_{i=1}^n \mathbf{u}_i (x^i - x'^i)$$

根据 (b), 就有,

$$x^i - x'^i = 0$$

即两个分解其实是相同的。

另外, 在 \mathbb{R}^n 中显然有这样一组基

$$\begin{cases} \mathbf{e}_1 := (1, 0, 0, \dots, 0) \\ \mathbf{e}_2 := (0, 1, 0, \dots, 0) \\ \dots\dots\dots \\ \mathbf{e}_n := (0, 0, 0, \dots, 1) \end{cases}$$

我们称 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ 是 \mathbb{R}^n 的标准基。

为了阐述清楚后面的内容, 我们再引入下面的定义。

Definition 1.3. 设存在一组矢量 $\{\mathbf{u}_i\}$, 如果

$$\sum_{i=1}^n \mathbf{u}_i x^i = \mathbf{0}$$

但 $\{x^i\}$ 不都等于 0. 那么就称这组矢量为线性相关的, 否则就是线性独立的。可以看到, \mathbb{R}^n 中的基是线性独立的。

下面让我们来看一个例子，也就是我们熟知的复数。这是一个给已知空间 \mathbb{R}^2 添加结构（即后面所谓的 \times 运算）的尝试。

Example 1.1. 记实数域上的 \mathbb{R}^2 中标准基 $\mathbf{e}_1 = (1, 0), \mathbf{e}_2 = (0, 1)$ 和任意三个元素 $\mathbf{a}, \mathbf{b}, \mathbf{c}$ 还有一个标量 k ，如果存在一种运算规则 \times 满足下列规则，则称 \mathbb{R}^2 为复数域 \mathbb{C} （请检查成为域的条件），其中的元素称为复数。

$$(a). \mathbf{e}_1 \times \mathbf{e}_1 = \mathbf{e}_1$$

$$(b). \mathbf{e}_2 \times \mathbf{e}_2 = -\mathbf{e}_1$$

$$(c). \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_2 \times \mathbf{e}_1 = \mathbf{e}_2$$

$$(d). \mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$$

$$(e). \mathbf{a} \times (b\mathbf{k}) = (\mathbf{a} \times \mathbf{b})k$$

为了求出一般的 \times 公式，把 $\mathbf{a} = (a^1, a^2), \mathbf{b} = (b^1, b^2)$ 按标准基展开，然后 \times 他们，有

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= (\mathbf{e}_1 a^1 + \mathbf{e}_2 a^2) \times (\mathbf{e}_1 b^1 + \mathbf{e}_2 b^2) \\ &= (\mathbf{e}_1 \times \mathbf{e}_1) a^1 b^1 + (\mathbf{e}_1 \times \mathbf{e}_2) (a^1 b^2 + a^2 b^1) + (\mathbf{e}_2 \times \mathbf{e}_2) a^2 b^2 \\ &= \mathbf{e}_1 (a^1 b^1 - a^2 b^2) + \mathbf{e}_2 (a^1 b^2 + a^2 b^1) \\ &= (a^1 b^1 - a^2 b^2, a^1 b^2 + a^2 b^1) \end{aligned}$$

如果 $\mathbf{a} = (a, 0), \mathbf{b} = (b, 0), \mathbf{c} = (c, 0)$ ，且 $ab = c$ ，则可以看到 $\mathbf{a} \times \mathbf{b} = (ab, 0) = \mathbf{c}$ ，反过来也一样可以成立，则 $ab = c$ 和 $\mathbf{a} \times \mathbf{b} = \mathbf{c}$ 等价！就是说，当所有的元素都是 $\mathbf{a} = \mathbf{e}_1 a$ 的形式的时候， \times 和我们熟知的实数乘法是一样的。从这种意义上，复数的乘法可以完全地看成实数的乘法。所以我们可以形式地等同 \mathbf{a} 和 $(a, 0)$ ，而建立一种乘法。

形式地记 $\mathbf{e}_1 = 1, \mathbf{e}_2 = i$ ，那么前三条法则也就变成了

$$(a'). 1 \times 1 = 1$$

$$(b'). i \times i = i^2 = -1$$

$$(c'). 1 \times i = i \times 1 = i$$

这里完全可以把 \times 看成（扩展的）实数的乘法了，为此我们重新证明一下一般的 \times 公式，现在我们叫他复数的乘法公式，而使用的符号也和实数的时候一样（必要时略去），也不再用粗体标识复数。

设 $a = a^1 + a^2 i, b = b^1 + b^2 i$ ，注意到 $i^2 = -1$ 我们有

$$\begin{aligned} ab &= (a^1 + a^2 i)(b^1 + b^2 i) \\ &= (a^1 b^1 - a^2 b^2) + (a^1 b^2 + a^2 b^1) i \end{aligned}$$

一般在术语上，设 $a = a^1 + a^2 i$ ，则称 a^1 为复数 a 的实部， a^2 为复数 a 的虚部。再规定一个运算共轭， $a^* = a^1 - a^2 i$ ，其中 a^* 称为 a 的共轭复数，或简略地，共轭。读者可以

自行验证，加法的共轭等于共轭的加法，乘法的共轭也等于共轭的乘法。这应当是读者在中学时候就已经熟知的结论。

1.2 矢量空间和线性映射

Definition 1.4. 给定域 \mathbb{F} 和一个交换群 $(V, +)$ 。现任选 $k \in \mathbb{F}$ 和 $v \in V$ ，如果存在右作用 $f: V \times \mathbb{F} \rightarrow V$ ，记作 vk ，满足（假设 $u, v, w \in V$ ， $a, b \in \mathbb{F}$ ）：

$$(a). (v + w)a = ua + va;$$

$$(b). u(a + b) = ua + ub;$$

$$(c). (ub)a = u(ba);$$

$$(d). \mathbb{F} \text{ 中的单位元 } 1, \text{ 称为单位标量, 对任何矢量 } u \text{ 有 } u1 = u.$$

则称 V 为 \mathbb{F} 上的右矢量空间（或向量空间、线性空间）。 V 中的元素称为矢量，而 \mathbb{F} 中的元素被称为标量，而那个右作用称为标量乘法或者标量作用。

有点长的定义，不过很容易就可以验证 \mathbb{R}^n 是 \mathbb{R} 上的一个矢量空间。这是最一般的矢量空间的定义，同时我们也放弃了粗体标识矢量的方式，而采用事先声明来表示，除非有必要，我们一律使用一般字体。

Example 1.2. 域 \mathbb{F} 上的 n 维 Euclid 空间。

可以看到 $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ ，那么我们可以类似定义 $\mathbb{F}^n = \mathbb{F} \times \cdots \times \mathbb{F}$ ，其上的加法和标量乘法定义为

$$x + y = (x^1, \cdots, x^n) + (y^1, \cdots, y^n) = (x^1 + y^1, \cdots, x^n + y^n)$$

$$xc = (x^1, \cdots, x^n)c = (x^1c, \cdots, x^nc)$$

其中 $c \in \mathbb{F}$ 而 $x, y \in \mathbb{F}^n$ 。 \mathbb{F}^n 被称为域 \mathbb{F} 上的 n 维 Euclid 空间，这样记就不用声明所处的域，因为已经用 \mathbb{F}^n 明确写出。

注意到，我们这边把标量乘法作用在矢量右边。那么就有这样的问题：

Question 1.2. 标量可以直接移到矢量左边吗？

答案是不可以，因为我们没有定义标量左乘！可以看到，前面所有的标量都是右乘的，包括在讨论 \mathbb{R}^n 的时候，也许有些奇怪，但这就是如此。乘法的定义可以不是实数乘法，或者复数乘法，在具体问题的时候可以具体讨论，体现了抽象定义的灵活性。有些时候，比如在 \mathbb{R}^n 中我们就常常定义左乘，并且让左乘和右乘的结果相同。

我们称上面这样定义的矢量空间为 \mathbb{F} 上的右矢量空间。自然也可以同样定义左矢量空间。甚至，同时定义左乘右乘。不管怎么样，值得声明的是：

Proposition 1.3. 乘法一般来说是不可交换的。如果 u 和 v 之间存在一种乘法，即 uv 一般不等于 vu 。

这在后面将会一直遇到。就具体例子来说，实数乘法和复数乘法都是可以交换的。在下面的线性映射上，我们将会再一次接触了不可交换的乘法。但需要点明的是，域 \mathbb{F} 中标量和标量的乘法和加法都是满足交换律的¹。

不过在讨论线性映射之前，让我们回忆一下 \mathbb{R}^n 的基和分解，并且已经在具体的计算上，比如 Example 1.1 上展现了威力。我们自然要问：

Question 1.3. 在一般的矢量空间上，我们是否也可以有基和分解的概念？

答案当然是可以的，但我不打算写出那么长的证明，所以我不在这里具体阐述，可以参见任何一本线性代数（比如我小节里面的参考书）。不过，我还是可以去解释一些重要而在前面的直观的概念。

Question 1.4. 或者我们可以问这样一些问题。

1. 如果我们要建立分解，那么唯一性我们是希望保留的，怎么保留？（参考 Definition 1.2 的 (b)）
2. 在 Definition 1.2 里面， \mathbb{R}^n 上的一组基有 n 个矢量，同时我们说 \mathbb{R}^n 是 n 维的。那么一般矢量空间的维数是否可以这样定义？
3. 矢量空间的维数是否一定是有限的？

这里我直接给出答案。请读者尝试自己完成证明。

Proposition 1.4. 如果知道了一组线性独立的矢量，则他的所有线性组合的集合构成一个矢量空间，或者说生成一个矢量空间，那么这样的线性独立矢量组被称为这个矢量空间的基。反过来，一个矢量空间总可以挑出无数线性独立矢量组，这个矢量空间的维数（维度）被定义为这些线性独立矢量组中矢量的个数的最大值。如果这个值有限，则称这个矢量空间是有限维的，否则无限。

既然我们已经定义了一般的矢量空间，我们接下来的工作都将在矢量空间下完成。下面讨论线性映射。

所谓线性，就是“像直线一样的”，那么对于我们最熟知的直线 $f(x) = kx$ ，我们考察他在运算上的性质，然后推广开来。

对加法有，

$$f(x^1 + x^2) = kx^1 + kx^2 = k(x^1 + x^2) = f(x^1 + x^2)$$

¹也许有些教材会让域中的乘法不满足交换律，然后让满足乘法交换律的域叫做交换域，这是有些古老的称呼，让我们放弃他。

对数乘有,

$$cf(x) = ckx = kcx = f(cx)$$

Definition 1.5. 设两个右向量空间 V 和 W , 以及非空集合 X, Y , $X \subseteq V, Y \subseteq W$, 若存在这样一个从 X 到 Y 的映射 A , 使得所有 $x_1, x_2, x \in X$ 和所有标量 c , 满足

$$A(x_1 + x_2) = A(x_1) + A(x_2), A(xc) = A(x)c$$

那么便称为一个 (左作用的) 线性映射 (或线性变换、线性函数、线性算子、线性算符等等)。如果 A 是线性的, 那么通常把 $A(x)$ 写作 Ax .

把所有 $V \rightarrow W$ 的线性映射构成的集叫做线性映射空间, 记作 $L(V, W)$, 若是 $L(V, V)$, 则将其简记作 $L(V)$, 简称为 V 上的线性映射空间。

很容易就可以看出, 标量左乘 $f(x) = kx$ 就是一个 (左作用的) 线性映射, 当然也可以我们也可以对标量右乘做相同的论断。

然后让我们看看一下任意线性映射作用在零矢量上的结果。选一个不为 0 的标量 k , 因为 $f(0) = f(0k) = f(0)k$, 所以 $f(0) = 0$, 也就是说任何线性映射作用在零矢量上都得到零矢量。这是所有线性映射都有的性质。

Definition 1.6. 设两个右向量空间 V, W, X , V 中的矢量 x , 再设 $A_1, A_2 \in L(V, W)$, c_1, c_2 是两个标量, 则这两个线性映射的加法定义为,

$$(c_1A_1 + c_2A_2)x := A_1(xc_1) + A_2(xc_2)$$

显然, $(c_1A_1 + c_2A_2) \in L(V, W)$.

设 $A \in L(V, W), B \in L(W, X)$. 那么这两个线性映射的乘法 $BA : V \rightarrow X$ 定义为 $B \circ A$. 显然, $BA \in L(V, X)$.

值得注意的是, 上面定义的线性映射空间 $L(V, W)$ 是一个 (左) 向量空间。 A 是一个函数, 但却又成为了空间中的一个“点”, 这就是引论里面的“关系也可以成为对象”的鲜活例证。这就是说, 我们对一般向量空间的研究结论也能应用到线性映射空间去。另外, 在向量空间外我们还引入了线性映射空间一个新的结构, 也就是线性映射的乘法。

因为这里线性映射的乘法就是映射的复合。所以满足 $C(BA) = (CB)A$.

一般来说, 假如 BA 存在, AB 也不一定存在。作为例外, 当 A 和 B 都是 $L(V)$ 中的元素时, BA 和 AB 都是存在的。注意到我们对乘法一般的 Proposition 1.3, 我们自然就想要知道 BA 和 AB 是否相等? 这个问题将会明确地在矩阵那里回答, 但是先让我们猜一下。由于线性映射的乘法就是映射的复合, 映射的复合一般又不满足交换律²。所以我们应该相信, 线性映射的乘法应该也不满足交换率, 如果满足, 则是线性映射一个很优良的性质。

²比如 $f(x) = x^2, g(x) = x + 1$, 则 $f(g(x)) = (x + 1)^2, g(f(x)) = x^2 + 1$.

1.3 内积和对偶空间

本节讨论的是内积和对偶空间，在此这个之前，先定义群上的一种一元运算，对偶。

Definition 1.7. 现在有一个群 G 上的一个运算 $f: X \rightarrow X$ 满足：

- (1). f 是一个双射;
- (2). $f = f^{-1}$;
- (3). $f(ab) = f(b)f(a)$.

则将 f 称为对偶。 $g \in G$ 的对偶通常记作 g' ，那么条件 (2) 就是 $(g')' = g$ ，条件 (3) 就是 $(ab)' = b'a'$ 。显然，映射 $f(g) = g^{-1}$ ，也就是逆，自然地成为一种对偶。

现在设 V 是 \mathbb{F} 上的 n 维右矢量空间，任意固定一组基 $\{v_i\}$ 考察一个特殊的线性映射 $f: V \rightarrow \mathbb{F}$ ，设有一个矢量 $x = \sum_{i=1}^n v_i x^i$ 。则，

$$f(x) = f\left(\sum_{i=1}^n v_i x^i\right) = \sum_{i=1}^n f(v_i) x^i$$

记 $f(v_i) = u_i$ ，则 $f(x)$ 就是

$$f(x) = u_1 x^1 + u_2 x^2 + \cdots + u_n x^n \quad (1.1)$$

这样的线性映射 f ，我们叫做标量线性映射，或者称之为对偶矢量（注意我没说他是 x 的对偶矢量），记作 u' ，我们把记号 u 依旧留给原空间中的矢量。

上一节中说过，线性映射空间也可以成为一个矢量空间。如果希望对偶矢量的集合也构成一个矢量空间，我们需要定义标量乘法。如果原矢量空间定义的是标量右乘，那么我们在对偶空间中定义标量左乘。为了不显得突兀，上面是引入，下面给出严格的定义。

Definition 1.8. 一个 \mathbb{F} 上的右（左）矢量空间 V 到 \mathbb{F} 的所有线性映射构成一个 \mathbb{F} 左（右）矢量空间 V' ，我们称之为对偶空间，对偶空间中的元素称为对偶矢量。

对偶空间作为矢量空间的结构我们之后考虑。然后是对偶基的定义：

Definition 1.9. 选取 V 中的一组基 $\{v_i\}$ ，那些使得 $f^j(v_i) = \delta_i^j$ 成立³的映射 $\{f^j\}$ 是对偶空间的一组基，称为 $\{v_i\}$ 的对偶基，记作 $\{v^i\}$ 。

接着是一个矢量的对偶矢量的定义：

Definition 1.10. 如果 $x \in V$ 对的基 $\{v_i\}$ 展开是 $x = \sum_{i=1}^n v_i x^i$ ，那么他的对偶矢量定义为在 V' 中展开为 $\sum_{i=1}^n x_i v^i$ 的矢量，记作 x' 。

³这里的 δ_i^j 是 δ 函数，当 $i = j$ 时 $\delta_i^j = 1$ ，否则等于 0，以后会遇到 δ_{ij} 和 δ^{ij} ，定义类似，当 $i = j$ 的时候 $\delta^{ij} = \delta_{ij} = 1$ ，否则等于 0。

简单说，对偶矢量就是把对基 $\{v_i\}$ 展开的坐标 $\{x^i\}$ 变成了对对偶基 $\{v^i\}$ 展开 $\{x_i\}$ 。其中 $\{x_i\}$ 被称为 $\{x^i\}$ 的对偶标量。一般地，标量 a 的对偶标量记作 a' 。

把对偶矢量对矢量的作用严密化的定义：

Definition 1.11. 对偶矢量 u' 对 x 的作用我们称之为内积，记作 $\langle u, x \rangle$ ，或者 $\langle u|x \rangle$ 。有时候我们也叫 u 点乘 x ，记作 $u \cdot x$ 。

可以看到 u' 对 x 作用我们定义为了原空间 V 中的两个矢量的内积，而不涉及对偶空间。对对偶基来说，内积写作 $\langle v_i, v_j \rangle = \delta_j^i$ （不记作 δ_{ij} 是为了以后的求和约定）。

现在考虑对偶空间的结构。首先考察矢量加法 $x + y$ ，

$$x + y = \sum_{i=1}^n v_i(x^i + y^i)$$

他的对偶矢量

$$(x + y)' = \sum_{i=1}^n (x^i + y^i)' v^i$$

要使其成为域上的对偶，则对偶法则 $(a + b)' = a' + b'$ 满足⁴，所以

$$(x + y)' = \sum_{i=1}^n ((x^i)' + (y^i)') v^i = x' + y' \quad (1.2)$$

然后考察标量右乘 xa ，我们寻找他的对偶矢量。将其展开

$$xa = \sum_{i=1}^n v_i(x^i a)$$

那么他的对偶矢量应该为

$$(xa)' = \sum_{i=1}^n (x^i a)' v^i$$

要使其成为域上的对偶，则对偶法则 $(ab)' = b'a'$ 满足，所以

$$(xa)' = \sum_{i=1}^n (x^i a)' v^i = a' \sum_{i=1}^n x_i v^i = a' x' \quad (1.3)$$

可以清晰地看到，这样子通过(1.2)和(1.3)我们定义了对偶空间中的矢量加法和标量乘法。可以很容易证明对偶空间是一个左矢量空间。

按照上面的一些运算规则，我们可以写出内积的运算规则

$$\langle u, x + y \rangle = \langle u, x \rangle + \langle u, y \rangle, \langle u + v, x \rangle = \langle u, x \rangle + \langle v, x \rangle$$

⁴因为域上的加法群是 Abel 群，所以 $b' + a' = a' + b'$ 。

$$\langle u, xa \rangle = \langle u, x \rangle a \quad (1.4)$$

$$\langle ub, x \rangle = b' \langle u, x \rangle \quad (1.5)$$

这里仅仅说明一下(1.5). 因为 $\langle ub, x \rangle = \sum_{i=1}^n (u_i b)' x^i = b' \sum_{i=1}^n u_i x^i = b' \langle u, x \rangle$.

需要注意的是, 内积又叫点乘, 也是一种乘法, 根据一般表述 Proposition 1.3 则 $\langle u, x \rangle$ 和 $\langle x, u \rangle$ 一般来说是不同的, 因为第一个是 $\langle u, x \rangle = u_1 x^1 + u_2 x^2 + \cdots + u_n x^n$, 另一个是 $\langle x, u \rangle = x_1 u^1 + x_2 u^2 + \cdots + x_n u^n$, 可以看到 $\langle u, x \rangle = \langle x, u \rangle'$.

可以看到, 对偶标量之间, 我们没有定义半点关系, 所以还是比较自由的。在实际使用中, 往往喜欢在对偶标量间加一点关系。比如在量子力学里, 我们的计算在复数域上进行, 对于对偶标量我们会让 $x' = x^*$ (x^* 是复数 x 的共轭), 可以看到如此的话 $\langle u|x \rangle = \langle x|u \rangle^*$.

值得一提的是可以看到 $f(x) = \langle u, x \rangle$ 只和 x 和 f 有关。更一般地, 对于线性映射来说, 不仅仅是内积, 他的定义不含基的选取, 所以线性映射和内积与基的选取无关。但是从 $\langle v_i, v_j \rangle = \delta_j^i$ 似乎可以看到内积和基的选取有关? 如果我们换一组基 $\{u_i\}$, 那么是不是一定有 $\langle u_i, u_j \rangle = \delta_j^i$?

不是这样的。为此让我们更精细地考察 $\langle v_i, v_j \rangle = \delta_j^i$ 这个关系, 这个关系写成 $v^i(v_j) = \delta_{ij}$ 可能让问题变得好回答一些。可以看到内积由 v^i 和 v_j 唯一确定, 而和基无关, 但是注意 $\langle v_i, v_j \rangle$ 中可不显含 v^i , 也就是说, 我们没有给定 v_i 是怎么变成 v^i 的。那么我们怎么给定呢? 通过 $\langle v_i, v_j \rangle = \delta_j^i$! 所以 $\langle v_i, v_j \rangle = \delta_j^i$ 是一个定义式, 由他确定了内积。如果此时换了一组基 $\{u_i\}$, $\langle v_i, v_j \rangle = \delta_j^i$ 这个关系还是成立的, 但是如果要问 $\langle u_i, u_j \rangle = \delta_j^i$ 成不成立, 那么自然是不一定的。因为成立的话, 内积的定义就可能改变了。

在形式地定义了内积之后, 让我们回顾一下引入中的式(1.1)。

$$\langle u, x \rangle = \sum_i \sum_j \langle v_i u^i, v_j x^j \rangle = \sum_i \sum_j u_i x^j \langle v_i, v_j \rangle = \sum_i \sum_j u_i x^j \delta_j^i = \sum_i u_i x^i$$

接下来是两个重要的记号方法, 在许多时候使用会十分方便。

Example 1.3. Dirac 符号体系。

Dirac 符号体系是最常用的量子力学的数学描述, 要点是采取了对偶空间这个概念, 这里将演示一些代数上的基础。或者说, 发展出一套新的符号体系。

遵从 Dirac, 我们将原空间中的矢量记作 $|x\rangle$, 而对偶空间中的向量记作 $\langle u|$. 那么内积就变成了这两个符号的自然拼接 $\langle u|x \rangle$. 这样的记号的优点是清楚。

让接着我们考虑一下线性映射 F 在矢量 $|x\rangle$ 上的作用, 一般称其为线性算子。遵从 Dirac, 我们一般将线性映射的作用直接写作 $F|x\rangle$, 他自然生成了一个矢量, 这个矢量和对偶空间中矢量的内积则记作 $\langle u|(F|x\rangle)$. 这时候我们自然需要把 F 定义到对偶矢量上。也就是说 $(\langle u|F)|x\rangle := \langle u|(F|x\rangle)$, 这样我们就可以去掉括号, 简记为 $\langle u|F|x\rangle$.

$\langle u|F|x\rangle$ 同时也是一个内积, 所以可以换一下位置 $\langle (F|x)|u\rangle'$, $\langle (F|x)|$ 是对偶空间的一个矢量, 我们把他表示成一个线性算子 F^\dagger 对 $\langle x|$ 的作用, 即 $\langle x|F^\dagger$, 这样也保持了 $\langle u|F|x\rangle = \langle x|F^\dagger|u\rangle'$ 的紧凑形式, 其中 F^\dagger 被称为对偶算子。特别地, 当 $F = F^\dagger$, 则称 F 为对称算子, 或自伴算子。

接着, 毫无疑问, $|x\rangle\langle v|y\rangle$ 是一个原空间中的矢量, 我们也可以理解做 $|x\rangle\langle v|$ 作用在原空间的矢量 $|y\rangle$ 上, 很容易检验 $|x\rangle\langle v|$ 是一个线性算子。显然, 这个线性算符也可以右作用在一个对偶矢量上。

Example 1.4. 协变和逆变, *Einstein* 求和约定。

将分量指标出现在上面的矢量叫做逆变矢量, 而其对偶空间中的矢量叫做协变矢量。

Einstein 求和约定即是, 当一个符号在上标和下标同时出现时, 则对他求和 (求和的范围一般需要给出)。比如内积

$$A = \sum_{i=1}^n u_i x^i$$

在 *Einstein* 求和约定下则记作 $u_i x^i$. 而 $u_i x^j$ 则仅代表一个数。

再比如求和约定下 $u_i^j x^i y_j$ 代表

$$\sum_{i=1}^n \sum_{j=1}^m u_i^j x^i y_j = \sum_{i=1}^n x^i \left(\sum_{j=1}^m u_i^j y_j \right)$$

以后除非特殊情况 (会事先声明), 我们会一直使用 *Einstein* 求和约定。

1.4 矩阵

若 V 的一组基为 $\{u_1, u_2, \dots, u_n\}$, W 的一组基为 $\{v_1, v_2, \dots, v_m\}$. 设 $A \in L(V, W)$, 于是将 Au_i 对 $\{v_j\}$ 展开, 用求和约定, j 从 1 加到 m .

$$Au_i = v_j a_i^j$$

将这些 a_i^j 按照 j 是行数、 i 是列数按顺序写出来就是一个 $m \times n$ 的长方阵, 我们称这个长方阵为线性映射 A 的矩阵, 或者直接称为一个 $m \times n$ 矩阵。下面写出这个矩阵,

$$(A) = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \cdots & a_n^m \end{pmatrix}$$

可以看到 Au_j 关于基 $\{v_i\}$ 的坐标出现在 (A) 的第 j 列, 所以一个线性映射 A , 确定了两组基才可以确定他的矩阵。所以改变基必然会改变 A 的矩阵。换言之, 一个线性映射, 只要选定了两组基, 矩阵则和他等价, 所以我们没有理由说矩阵不是一个线性映射。下面将要发展一种乘法, 将矩阵直接作为一个线性映射看待, 目标就是把这种乘法的具体形式写出来。

为此我们先从线性映射的乘法出发。

设 U, V, W 是三个矢量空间, 分别为 l, m, n 维, 而他们的基分别是 $\{u_i\}, \{v_j\}, \{w_k\}$, 则对于 $x \in U$, 若存在 $A \in L(U, V), B \in L(V, W), C = BA \in L(U, W)$ 有

$$Ax = A(u_i x^i) = A(u_i) x^i = v_j a_i^j x^i$$

然后在把 B 作用上去,

$$Cx = (BA)x = B(Ax) = B(v_j) a_i^j x^i = w_k b_j^k a_i^j x^i \quad (1.6)$$

另一方面

$$Cx = C(u_i x^i) = C(u_i) x^i = w_k c_i^k x^i \quad (1.7)$$

则可以从(1.6)和(1.7)直接读出, $c_i^k = b_j^k a_i^j$. 为了直观, 我给出这条式子的矩阵表示:

$$\begin{pmatrix} b_1^1 & b_2^1 & \cdots & b_m^1 \\ b_1^2 & b_2^2 & \cdots & b_m^2 \\ \vdots & \vdots & \ddots & \vdots \\ b_1^n & b_2^n & \cdots & b_m^n \end{pmatrix} \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_l^1 \\ a_1^2 & a_2^2 & \cdots & a_l^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \cdots & a_l^m \end{pmatrix} = \begin{pmatrix} b_j^1 a_1^j & b_j^1 a_2^j & \cdots & b_j^1 a_l^j \\ b_j^2 a_1^j & b_j^2 a_2^j & \cdots & b_j^2 a_l^j \\ \vdots & \vdots & \ddots & \vdots \\ b_j^n a_1^j & b_j^n a_2^j & \cdots & b_j^n a_l^j \end{pmatrix} \quad (1.8)$$

很显然, 如果 (BA) 的矩阵元为 c_{ij} , 则 c_{ij} 由第一个矩阵的第 i 行和第二个矩阵的第 j 列逐个相乘然后相加而得。矩阵乘法把 $n \times m$ 和 $m \times l$ 的矩阵变成了 $n \times l$ 的矩阵。

在(1.8)式中令 $n = l = 1$, 则变成

$$\begin{pmatrix} b_1 & b_2 & \cdots & b_m \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^m \end{pmatrix} = (b_j a^j) = (\langle b, a \rangle) \quad (1.9)$$

注意, 如果我们把(1.9)式左边的第二项看成原空间中的矢量, 而左边的第一项看成对偶空间中的矢量。我们再一次在右边得到了熟知的内积。则我们就这么干, 让左边的第二项是一个矢量, 称为列矢量, 另一个称为行矢量, 可见这样子一个列矢量的对偶矢量是

行列式。那么我们就可以改写一下(1.9)式。先写出改写成果：

$$\begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \cdots & a_l \end{pmatrix} = \begin{pmatrix} \langle b_1, a_1 \rangle & \langle b_1, a_2 \rangle & \cdots & \langle b_1, a_l \rangle \\ \langle b_2, a_1 \rangle & \langle b_2, a_2 \rangle & \cdots & \langle b_2, a_l \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle b_n, a_1 \rangle & \langle b_n, a_2 \rangle & \cdots & \langle b_n, a_l \rangle \end{pmatrix} \quad (1.10)$$

其中 b^1 等代表行矢量, a_1 等代表列矢量, 因为出现在内积里面的都是原空间的矢量, 所以出现的是 b_1 等而不是 b^1 。可以看到, (1.10)式右边, 第 (i, j) 位置的值就是左边第一项的 i 行和左边第二项的 j 列内积而来的。

既然这样用矩阵再一次给出了原空间中矢量的表示 (列矢量), 那么我们就可以去给出线性映射的表示了。

让(1.10)式中的 $l = 1$, 则

$$\begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^n \end{pmatrix} \begin{pmatrix} a \end{pmatrix} = \begin{pmatrix} \langle b_1, a \rangle \\ \langle b_2, a \rangle \\ \vdots \\ \langle b_n, a \rangle \end{pmatrix} \quad (1.11)$$

前面看到了线性映射 A 的矩阵作用在 m 行的列矢量上成为了 n 行的列矢量。对应的方程也就是 $Ba = a'$, a 和 a' 都是原空间中的矢量。在这里, 矩阵乘法表现出了很神奇的作用。前面说过 “一个线性映射, 只要选定了两组基, 矩阵则和他等价”, 那么线性映射对矢量作用的方式, 用矩阵表示, 就是一个矩阵对列矢量的矩阵乘法!

显然矩阵 $(A), (B), (C)$ 之间的乘法成立关系 $(A)[(B)(C)] = [(A)(B)](C)$, 这是线性映射 $A(BC) = (AB)C$ 的直接对应。当然, 读者可以直接计算验证。

前面问过一个问题, 线性映射的乘法, 比如 $L(V)$ 中两个线性映射 A 和 B , BA 和 AB 是否相等? 在这里选终于可以做出方便的反例。

选定基, 给出两个线性映射 A 和 B 的矩阵:

$$(A) = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}, (B) = \begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix} \quad (1.12)$$

则

$$(AB) = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & 3 \end{pmatrix} \quad (1.13)$$

$$(BA) = \begin{pmatrix} 0 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ -2 & 1 \end{pmatrix} \quad (1.14)$$

比较就可以知道, $BA \neq AB$

1.5 正交和矢量相关概念的几何直观

这里要对前面的许多概念给出一些几何图像。首先给出正交的概念。

Definition 1.12. 设矢量空间 V 和其中的两个矢量 u 和 v , 若 $\langle u, v \rangle = 0$, 则称 u 和 v 正交 (或垂直)。

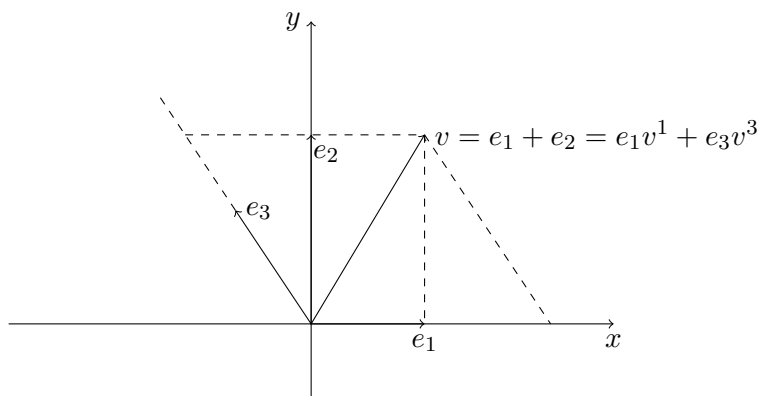
Example 1.5. \mathbb{R}^n 中的标准内积。

对于 \mathbb{R} 上的 \mathbb{R}^n 来说, 一般还附加一个很漂亮的结构。就是说他的对偶空间和其本身的结构一模一样。简单来说就是 $v' = v$. 那么 $\langle u, v \rangle = \langle u_i e^j, v_j e^j \rangle = u_i v_j \langle e^j, e^i \rangle = u_i v_j \delta^{ij} = \sum_i u_i v_i$. 这样的内积我们称为 \mathbb{R}^n 中的标准内积。并且可以看到 $\langle v, v \rangle = \sum_i v_i^2 \geq 0$.

Definition 1.13. 如果如果 V 中的一组基 $\{e_i\}$ 中的任意俩元素满足 $\langle e_i, e_j \rangle = 0$ ($i \neq j$), 而当 $i = j$ 的时候 $\langle e_i, e_j \rangle \neq 0$, 则称这样的基 $\{e_i\}$ 为正交基。若矢量 v 满足 $\langle v, v \rangle = 1$, 则称矢量 v 为归一的。而若 $\langle e_i, e_j \rangle = \delta_i^j$, 则称这样的基为正交归一基。

显然, 按照一般法则, 我们选取的对偶基都是正交归一基, 特别地, \mathbb{R}^n 中的标准基是正交归一基。

接着让我们回归 \mathbb{R}^2 中的几何。这将给出所有矢量空间的几何性质一个很直观的印象。



正如读者在中学里面熟知的, 一个矢量可以分解到 x 轴和 y 轴上去, 这就是分解 $v = e_1 + e_2$. 注意到这里 e_1 和 e_2 是垂直的, 并且 $e_1 = (a, 0)$ 且 $e_2 = (0, b)$, $\langle e_1, e_2 \rangle = 0a + 0b = 0$, 这两个矢量正交, 所以从几何上来看, 正交就是垂直!

另一方面, 矢量也可以按 e_1 和 e_3 展开, 这就是分解 $v = e_1 v^1 + e_3 v^3$. 可以简单的看到 e_1 和 e_3 不垂直, 但是这样的分解也是可以做到的。说明这是两组不同的基。

为了更好地描述甚至解决这些问题, 引入另一些几何概念。

Definition 1.14. 如果两个矢量 $\{v_1, v_2\}$ 是线性相关的, 则称这两个矢量是平行的。

Definition 1.15. n 维 Euclid 空间 \mathbb{R}^n 中的矢量 u 的长度定义为 $|u| = \sqrt{\langle u, u \rangle}$. 两个矢量 u 和 v 的距离定义为他们差的长度, 即 $|u - v|$. 显然 $|0| = 0$.

Definition 1.16. n 维 Euclid 空间 \mathbb{R}^n 中的两个矢量 u 和矢量 v 间的角度 $\theta (\theta \in [0, \pi))$ 定义为

$$|u||v| \cos \theta = \langle u, v \rangle$$

上面唯一没有明确定义的角度就是零矢量和任意矢量间的角度, 我们让他可以任意选取。

可以从 Definition 1.14 后面会有几何的例子解释为什么两矢量线性相关就是平行。Definition 1.15 则对矢量空间引入了新的结构。即所谓的长度的概念, 有了长度, 我们就可以说距离, 这些都是读者在中学就熟知的几何概念。Definition 1.16 是读者熟知的定义。当 $\langle u, v \rangle = 0$, 但 u 和 v 都不是零矢量的时候, 可以计算出 $\theta = \pi/2$, 这是一个熟知的几何事实。可以计算得, $(r \cos \phi, r \sin \phi)$ 和 x 轴上的标准基 $(1, 0)$ 的夹角就是 ϕ 。

有了长度的概念, 归一化就是理所当然的了。假设矢量为 v , 那么新的矢量 $v/|v|$ 的长度就是 1, 这就称为矢量的归一化。证明很显然, $|v/|v||^2 = \langle v/|v|, v/|v| \rangle = \langle v, v \rangle / |v|^2 = 1$ 。

接着看一下线性算子的几何, 这并不是读者那么熟悉的概念。这里的讨论限于 \mathbb{R}^2 , 但可以推广到任意矢量空间上去。这里尝试用矩阵来表示, 前面已经说过了, 选定了基, 那么也就直接确定了矩阵。首先给出一两个不那么与几何相关的美妙命题。

Proposition 1.5. 在相同的域上的任意 n 维矢量空间 V 都和 \mathbb{F}^n 具有相同的结构 (即所谓他们是同构的⁵)。

Proof. 所谓同构, 首先要在 V 中的元素和 \mathbb{F}^n 矢量空间中的元素存在双射 f , 这就是元素可以看成相同的。然后要保持结构, 也就是说保持加法 $f(v + w) = f(v) + f(w)$, 保持数乘 $f(va) = f(v)a$ 。这就是说经过映射, 加法还是加法, 数乘还是数乘, 可以看到, 这是一个线性映射。所以证明同构, 就是要证明这样的映射 (所谓的同构映射) 是存在的。

证明并不复杂, 主要依托于 Proposition 1.4. 先在 V 中选一组基 $\{v_i\}$, 然后把所有的矢量按这个展开 $a = v_i a^i$, 然后可以知道这一组 $\{a^i\}$ 唯一确定了 V 中所有的矢量。然后让 \mathbb{F}^n 中的一个矢量的坐标 (按标准基展开) 也为 $\{a^i\}$, 这样也唯一确定了 \mathbb{F}^n 中的一个矢量 $a' = e_i a^i$ (注意不是对偶矢量), 那么 V 中的一个矢量 a 一对一对应了一个 $a' \in \mathbb{F}^n$, 或者写作 $a' = f(a)$. 这样我们就找到了一个一对一的映射 f , 接着就是证明这个映射是线性的。

⁵更明确地说, 是在矢量空间意义上的同构。

设 $a + b \in V$, 然后 $a + b$ 按照基 $\{v_i\}$ 的展开是 $a + b = v_i(a^i + b^i)$, 所以 $f(a + b) = e_i(a^i + b^i) = e_i a^i + e_i b^i = f(a) + f(b)$, 同样 $ak = v_i(a^i k)$, 这就是说 $f(ak) = e_i(a^i k) = (e_i a^i)k = f(a)k$, 所以 f 是线性的。那么 f 就是一个同构映射。找到了同构映射, 我们也就可以说 V 和 \mathbb{F}^n 同构。□

可以注意到, 我并没有给出这个同构映射 f 的具体形式。此外可以有些很显然的推论。

Proposition 1.6. 同一域上的任意的 n 维矢量空间是同构的。

Proof. 注意到线性映射的复合依旧是线性映射, 双射的复合依旧是双射, 则 $A: V \rightarrow \mathbb{F}^n$ 和 $B: \mathbb{F}^n \rightarrow V'$ 的复合 $AB: V \rightarrow V'$ 依然是同构映射。□

下面这个命题也是重要的:

Proposition 1.7. 假设有同构映射 $f: V \rightarrow W$, 如果 V 是矢量空间, 那么 W 也是; 如果 W 是矢量空间, 那么 V 也是。此外, 如果 V 是 n 维的, 那么 W 也是, 反过来也一样。

Proof. 如果说留作习题你们会打我吗? 喵 ~~~ □

所以在同构意义上, 我们可以把同一域上所有 n 维矢量空间看成是等同的。Proposition 1.5 的证明思路其实在我们前面的讨论中已经出现过。我们曾经认为一个矢量和他的列矢量是等价的, 事实上还是有区别的, 列矢量是 $L(\mathbb{F}, \mathbb{F}^n)$ 里面一个元素的矩阵, 但是事实上, 矢量是 \mathbb{F}^n (或者 V) 里面的元素。 $L(\mathbb{F}, \mathbb{F}^n)$ 是一个矢量空间, 选定基之后, 他的所有元素的矩阵也构成一个矢量空间。这种等价是同构意义上的, 之所以这里要说清楚, 是因为下面要不加区分地反复使用。

有了上面一些关于代数细节的讨论, 现在我们可以放心地把线性映射作用于矢量等价于一个矩阵和列矢量相乘了。我们只讨论实数域上 $L(\mathbb{R}^2, \mathbb{R}^2)$ 上选定标准基的 2×2 的矩阵, 因为 1×2 的矩阵作用在 2×1 的矩阵上就是 \mathbb{R}^2 中的标准内积, 此外 $n \times 2 (n > 2)$ 的矩阵作用后又不在平面上。

可以给出一个著名的矩阵, 江湖人称单位矩阵

$$(E) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

此后在不至于产生混淆的情况下我们也省略掉矩阵和列矢量的括号了。可以计算有

$$Ev = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = v$$

也就是单位矩阵作用在任意列矢量上都得到这个列矢量本身。而且可以检验

$$EA = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

$$AE = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = A$$

这就是说 $AE = EA = A$, 前面我们说过矩阵乘法一般不满足交换律, 但单位矩阵却可以, 单位矩阵很神奇吧! 有木有! 有木有!

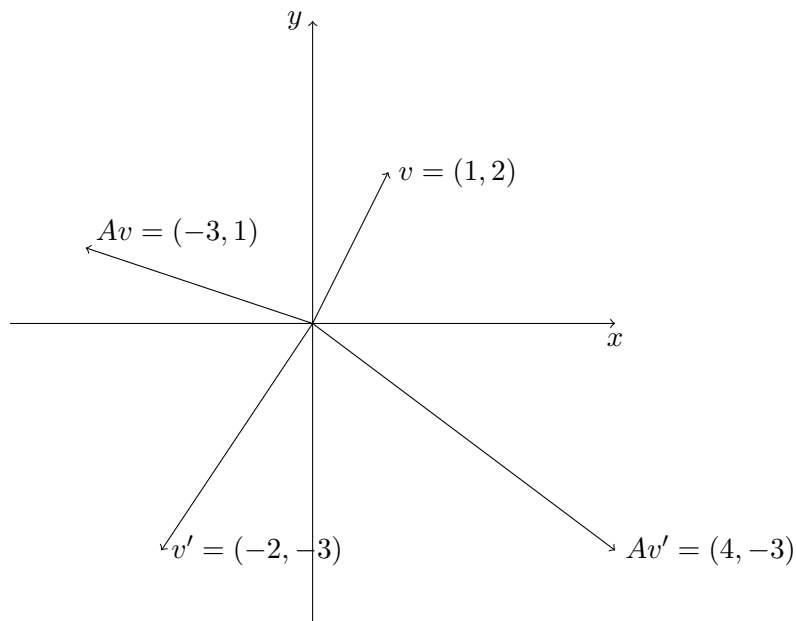
(1) 固定一个列矢量, 比如 $v = (1, 2)$, 找一个矩阵 (随意地),

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

计算一下

$$Av = \begin{pmatrix} 1 & -2 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

画出他的图 (顺便再外加一组矢量)



反正没什么规律。(2) 再比如一个矩阵正比于单位矩阵

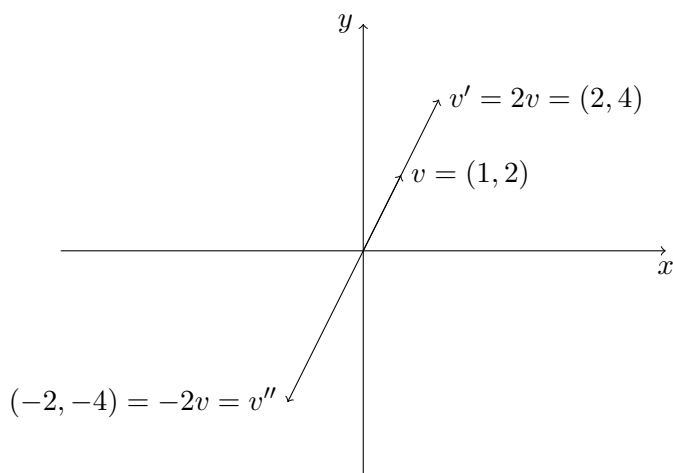
$$B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2E$$

再算一下

$$Bv = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2v$$

其实，这就是看作线性映射的数乘的矩阵，也就是 $v' = v\lambda$ ，这就是说 v' 和 v 线性相关。反过来，如果线性相关，也可以写作 $v' = v\lambda$ 的形式。所以平行和线性相关等价。

画图



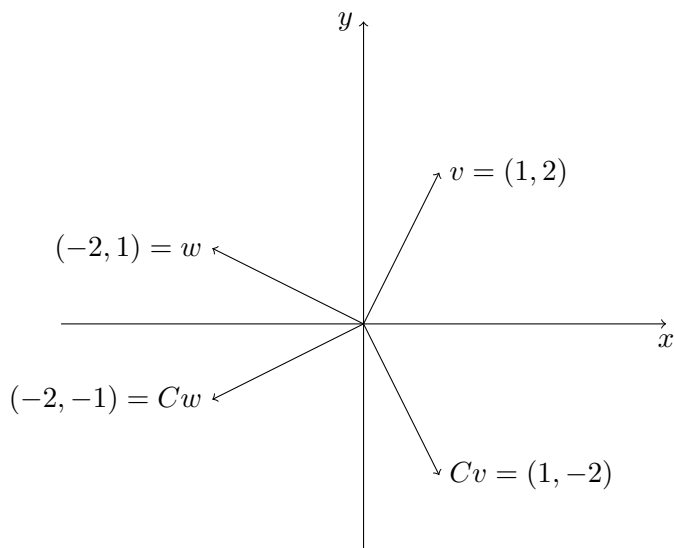
这个几何意义相当显然，如果是一个正比于单位矩阵的矩阵，那么他的作用就是拉长或者缩短。这里的放缩比例 λ 是可正可负的，如果是负 λ 的话，我们可以看到就是矢量反向的正放缩。总之，作用后的矢量和原矢量是平行的，如图。(3) 或者

$$C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.15}$$

$$Cv = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

很显然，这是关于 x 轴的对称。注意，可能对于一个矢量来说，他也可以是一个旋转，但是一旦多于一个点，则不能通过一个旋转操作得到（如下图所示），所以反射和旋转是不一样的。

画图有



(4) 既然上面提到了旋转，那么这里给出结果是旋转的矩阵。

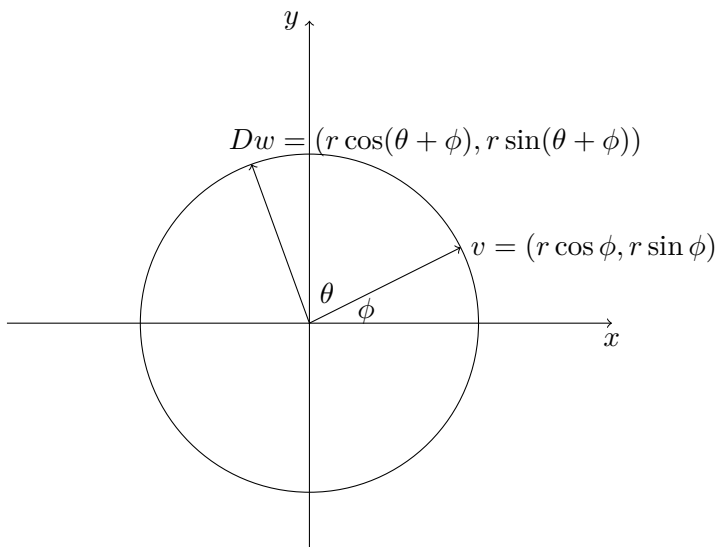
在中学的课程中已经知道了，任意二维矢量都可以写成 $v = (r \cos \phi, r \sin \phi)$ 的形式，其中 $r = |v|$ ，此即所谓的极坐标表示。现在给出一个矩阵，

$$D = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (1.16)$$

我们看一下他在 v 上的作用

$$Dv = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} r \cos \phi \\ r \sin \phi \end{pmatrix} = \begin{pmatrix} r(\cos \theta \cos \phi - \sin \theta \sin \phi) \\ r(\sin \theta \cos \phi + \cos \theta \sin \phi) \end{pmatrix} = \begin{pmatrix} r \cos(\theta + \phi) \\ r \sin(\theta + \phi) \end{pmatrix}$$

为了明晰起见，可以画出他的图



这样就给出了旋转变换的矩阵 D 了, 为了明确选择的角度, 往往加个下标 D_θ 。

(5) 线性映射的复合。这个意义是清晰的, 比如 $D_\theta C$ 就是将矢量先按 x 轴反射, 然后再把他绕原点旋转 θ . $CC = C^2$ 是将矢量先按 x 轴反射然后再反射一次, 直觉告诉我们作用后矢量并没有变, 所以 $C^2 = E$. 显然 $D_\theta D_{-\theta} = D_{-\theta} D_\theta = E$.

1.6 关于线性映射的进一步讨论

正如前面在矢量空间的同构中看到的, 一般来说同构映射仅仅是线性的是不够的, 而且还要是一个双射。这里需要讨论的就是线性映射是一个双射的条件。继而给出线性映射逆的概念。给出子矢量空间的定义, 并研究两个重要的子矢量空间。

假设我们下面考虑的都是满射, 所有映射都是线性映射。

如果线性映射 f 不是双射, 那么存在

$$f(a) = f(b) = c$$

这就是说

$$f(a) - f(b) = f(a - b) = 0$$

也就是说, 如果 f 不是双射, 那么对 f 来说 0 的原象就不止为 0 , 还有 $a - b$ 以及他的一切数乘 $(a - b)k$. ($f(0) = 0$ 这是普遍的)

反过来, 如果 0 的原象只有 0 , 那么

$$0 \neq f(a - b) = f(a) - f(b)$$

那么对于两个不同的 a 和 b , 则 $f(a) \neq f(b)$, 反过来也一样, 这就保证了单射。

在研究了 0 的原象之后, 我们再看一下任意元素的单射。设 0 的原象为 N , 考察任意一个元素 $f(a)$, 任选 $b \in N$, 则

$$f(a + b) = f(a) + f(b) = f(a) + 0 = f(a)$$

所以 $a + b$ 在 $f(a)$ 的原象里, 如果我们记 $a + N = \{v | v = a + b, b \in N\}$, 这就构成了全部的 a 的原象。全部是容易证明的, 如果 $b \notin a + N$ 但 $f(b) = f(a)$, 那么第一个条件给出 $(b - a) \notin N$, 但第二个条件给出 $(b - a) \in N$, 矛盾。

我们称呼 $a + N$ 为 N 的陪集。可以看到线性映射关于 0 的原象 N 是如此的重要, 以致于其他所有 B 中元素的原象都是 N 的陪集。

如果一个映射是单射, 且 $f(a) = b$, 那么自然地, b 的原象只有 a , 则可以定义 $a = f^{-1}(b)$, 其中 f^{-1} 显然是一个映射, 称为 f 的逆映射。逆映射的存在条件是 f 是单

射。那么在我们向量空间和线性映射的情况下，要求就是 N 的元素只有一个 0. 当然我们也可以这么探讨矩阵的逆，或者说逆矩阵。

如果向量空间 V 和 W 有着不同的维数，则那么一定不是同构的，之间也不存在同构映射（参见 Proposition 1.7）。那么对于一切线性映射 $f: V \rightarrow W$ ，我们都不能谈逆，因为此时 f 一定不是单射。因此只有当谈论俩同构的向量空间的时候，我们才可以谈论之间的映射是否一定有逆，更具体的，我们往往讨论的是 $f \in L(V)$ 是否有逆。

下面要深入讨论作为子向量空间的 N 。首先给出子向量空间的定义。

Definition 1.17. 设有一个域 \mathbb{F} 上的向量空间 V 和其中的集合 X . 对于集合 X 如果对于 V 中的加法和 V 和 \mathbb{F} 的标量乘法也构成一个 \mathbb{F} 上的向量空间，那么 X 被称为 V 的子向量空间（或线性子空间），或者在不混淆的情况下称为子空间。

一个显然的推论，根据 Proposition 1.4，从 n 维的向量空间 V ，他的一组基 $\{v_i\}$ 其中选 k 个，他们的线性组合就是 V 的一个 k 维的子空间。

如果线性映射 $f: V \rightarrow W$ 是两个向量空间间的映射（当然，假设是满射），那么此时 N 就成为 V 的一个子空间，被称为 V 中关于 f 的零空间。

Definition 1.18. 称集合 A 中的关系 \sim 为等价关系。如果满足

1. $a \sim a$ （自反性）
2. 如果 $a \sim b$, 那么 $b \sim a$ （互反性）
3. 如果 $a \sim b$ 且 $b \sim c$, 那么 $a \sim c$ （传递性）

把 A 中所有相互等价的元素放在一起，构成一个集合，我们称之为等价类。

假设有向量空间 V 以及其零空间 N ，如果定义，若 $(a - b) \in N$ ，则 $a \sim b$ ，那么可以很简单验证这是一个等价关系。所以对任意一个 $f(a) \in W$ ，他的原象构成一个等价类，记作 \bar{a} ， a 可以是等价类 $a + N$ 中任意一个元素，这里只是作为这个等价类的代表元素，注意 \bar{a} 是一个集合， V 中与 a 等价的所有元素都在里面，当然由 $a \sim a$ ， a 也在里面。

如果两个不同的等价类 \bar{a} 和 \bar{b} 有一个元素相同，那么根据传递性， b 的任意元素都和 a 中的元素等价，则 $\bar{b} \subseteq \bar{a}$ ，同理 $\bar{a} \subseteq \bar{b}$ ，那么 $\bar{a} = \bar{b}$ ，那么也就是说这两个等价类相同，矛盾。所以不同的两个等价类不相交。

所以一个向量空间 V 被肢解成了许多个不相交等价类的并。而这些等价类又都是零空间的陪集。

好的，现在开始我们在有限维向量空间下讨论。我们从每一个陪集中都挑选一个元素出来构成一个集合 H ，如果让 f 的定义域是 H ，我们称之为 f 在 H 上的限制，记作 $f|_H$ 。那么 $f|_H$ 称为了一个双射。这样的 H 显然不止一个，重要的是我们希望有一个 H 是 V 的子空间，那么 $f|_H$ 就变成了向量空间到向量空间的双射（一个同构）。这交由下面一个命题完成。

Proposition 1.8. 在众多的 H 中存在一个是 V 的子空间。

Proof. 这其实是一个构造过程。

首先, 可以验证 $\overline{u+v} = \bar{u} + \bar{v}$ 以及 $\overline{uc} = \bar{u}c$, 其中 u 和 v 是矢量, c 是一个标量 (请读者自行验证)。

现在开始构造, 先选 V 中的一组基 $\{v_i\}$, 以及他们的等价类 $\{\bar{v}_i\}$, 如果有重复, 就剔除挑那些重复的, 然后只留一个。就得到了一组新的 $\{\bar{v}_j\}$ (上标换成 j 以示区别), 然后挑出一组新的 $\{v_j\}$, 根据 Proposition 1.4, 这构成一组新的基, 把他们作为 H 中的元素。那么考虑他们线性组合 $v_j a^j = a$ 的等价类 $\bar{v}_j a^j = \bar{a}$, a 确实在 $\bar{v}_j a^j$ 中, 我们挑出这个 a 放入 H 中。这样进行下去得到 H , H 是所有 $\{v_j\}$ 的线性组合的集合, 由于 $\{v_j\}$ 是基, 所以 H 是一个矢量空间。又 H 在 V 内, 所以 H 是 V 的子空间。

下面我们所指的 H 都是我们构造的这个子空间。 □

由于 H 一个矢量空间, 所以 $0 \in H$, 但是除此之外, H 和 N 再无交集, 这个证明留给读者, 这个事实是不同等价类不相交的显然推论。

Remark. 为了构造和 W 的同构, 我们其实也可以不那么干。干脆直接把所有等价类 \bar{a} 的集合记作 V' , 定义映射 $f': V' \rightarrow W$ 满足 $f'(\bar{a}) = f(a)$, 可以验证这是一个同构映射 (交由读者完成)。这个集合 V' 由一个同构映射变成一个矢量空间, 那么 V' 也是一个矢量空间, 这由 Proposition 1.7 完成。

Definition 1.19. 如果有三个矢量空间 V, X, Y , 满足

$$1. X \cap Y = \{0\}$$

$$2. V = \{v | v = x + y, x \in X, y \in Y\}$$

那么称 X, Y 的直和为 V , 记作 $V = X + Y$, 或者有些人记作 $V = X \oplus Y$ 。

读者可以自己验证有限维矢量空间总可以写成 $V = N + H$ 。

此前说过, 内积是特殊的线性映射, 考虑矢量空间 V 和他的对偶空间 V' , 我们寻找对固定的 x 让 $\langle u, x \rangle = 0$ 的所有 u , 也就是说和 x 正交的所有 u . 让我们不使用上面的结论, 重新来解析一下这个例子。

由内积的线性性, 很容易验证, 这些 u 也构成一个矢量空间 U , 是 V 的子空间。选择 V 的一组基 $\{v_i\}$, 设 U 由前面的 k 个矢量构成的基生成。那么我们可以写出 $u = v_i u^i$, 显然这是从 0 加到 k (这里认为 $v^0 = 0$), 以及

$$\langle u, x \rangle = u^i \langle v_i, x \rangle = 0$$

由于 u_i 的任意性, 则要求 $\langle v_i, x \rangle = 0$ 其中 $0 \leq i \leq k$, 那么

$$\langle v_i, x \rangle = \langle v_i, v_j x^j \rangle = x^j \langle v_i, v_j \rangle = 0 \quad (0 \leq i \leq k)$$

所以 $x = \sum_{i=k+1}^n v_i x^i$, 那么所有这样的 x 和 U 里面的任意矢量都正交。这样的 x 也构成了一个子空间 X . 可证明的是 $V = X + U$. 这是上面的 $V = N + H$ 的一个特殊例子。

可以证明, 如果 X, Y 分别是 m, n 维的, 那么 $X + Y$ 就是 $m + n$ 维的。这从上面的例子来看是显然的, 具体证明交由小伙伴们完成。

这里谈论的, 都可以在群论里面找到更抽象的对应。比如零空间就变成了正规子群, 诸如此类。至于这些更抽象的知识, 可以参考彭解解教授的《线性代数》。

1.7 线性映射和内积、投影变换

设 A 是一个线性映射, 我们有一个 n 维矢量空间 V 以及他的对偶空间 V' , 我们先考察内积和线性映射之间的关系。

如果选定 V 中的一组基 $\{v_i\}$, 那么 $Av_j = v_k a_j^k$, 对于 $\langle v_i, Av_j \rangle$, 有

$$\langle v_i, Av_j \rangle = \langle v_i, v_k a_j^k \rangle = \delta_k^i a_j^k = a_j^i$$

如果使用 Dirac 符号, 矩阵元 (即一个矩阵任意位置上的元素, 一般是指第 i 行和第 j 列的元素) 就会显得特别清楚:

$$a_j^i = \langle v_i | A | v_j \rangle$$

其中左边的 i 代表行, 而 j 代表列。那么 A 的矩阵就可以写作

$$(A) = \begin{pmatrix} \langle v_1 | A | v_1 \rangle & \langle v_1 | A | v_2 \rangle & \cdots & \langle v_1 | A | v_n \rangle \\ \langle v_2 | A | v_1 \rangle & \langle v_2 | A | v_2 \rangle & \cdots & \langle v_2 | A | v_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle v_m | A | v_1 \rangle & \langle v_m | A | v_2 \rangle & \cdots & \langle v_m | A | v_n \rangle \end{pmatrix}$$

在介绍投影之前, 让我们用这个工具考察之前已经说过的对偶算子 (对偶映射)。 A 的对偶算子 A^\dagger 定义为 $\langle Ax | y \rangle = \langle x | A^\dagger | y \rangle$, 我们来求一下 A^\dagger 的矩阵元

$$\langle v_i | A^\dagger | v_j \rangle = \langle Av_i | v_j \rangle = \langle v_k a_i^k | v_j \rangle = (a_i^k)' \langle v_k | v_j \rangle = (a_i^k)' \delta_{kj} = (a_i^j)' = (\langle v_j | A | v_i \rangle)'$$

可以看到 A^\dagger 第 (i, j) 位置的矩阵元就是 A 第 (j, i) 位置的矩阵元的对偶。特别地, 如果 $A^\dagger = A$, 那么我们称呼其为自伴算子或对称算子。显然 $(A^\dagger)^\dagger = A$, 所以对偶算子的对偶算子是原算子, 这是 $(a')' = a$ 的自然结果。可以验证的是, 列矢量的对偶算子就是行矢量。

在现实生活中, 阳光在地面上留下各种影子。可以理想化地假设阳光都是沿着同一个方向 (比如说垂直于地面的角度) 照射而来, 大地是严格的平面, 那么, 对于任意一

个物体（比如说一只正在飞行的鸟），它的位置可以用向量 (x, y, z) 来表示，而这只鸟在阳光下对应着一个影子，也就是 $(x, y, 0)$ 。这样子的一个变换我们称为一个投影变换。它将三维空间中的向量 (x, y, z) 到映射到向量 $(x, y, 0)$ 。这是在 $x-y$ 平面上的投影。显然如果是 $(x, y, 0)$ ，他将被投影到 $(x, y, 0)$ ，也就是其本身。并且显然投影变换按照我们的直觉应该是一个线性映射。下面可以抽象出投影变换的定义。

Definition 1.20. 设 V 是矢量空间，如果算子 $P \in L(V)$ ，满足 $P^2 = P$ ，就说 P 是 V 里的一个投影变换（或者叫射影）。明确点，就是要求对任意的 $v \in V$ ， $P(Pv) = Pv$ 。换句话说， P 把每个向量固定在他的值域里面。

这里不使用求和约定。如果选定一组基 $\{v_i\}$ ，这样一组投影变换 $\{P_i\}$ 是常用的，要求 $P_i v_j = v_i \delta_{ij}$ ，我们可以计算有，

$$\begin{aligned} P_i x &= P_i \sum_j v_j x^j = \sum_j (P_i v_j) x^j = v_i \sum_j (\delta_{ij} x^j) = v_i x^i \\ P_i^2 x &= P_i (v_i x^i) = P_i v_i x^i = v_i x^i = P_i x \end{aligned}$$

所以 $\{P_i\}$ 确实是 V 中的一组投影变换。

为了消去 $P_i x = v_i x^i$ 中的坐标 x^i ，我们把他改成 $x^i = \langle v_i, x \rangle$ ，则 $P_i x = v_i \langle v_i, x \rangle$ ，或者改写成 Dirac 符号的形式

$$P_i |x\rangle = |v_i\rangle \langle v_i | x \rangle$$

所以 $P_i = |v_i\rangle \langle v_i|$ 。

此外可以证明如果 $i \neq j$ ，那么 $P_i P_j = |v_i\rangle \langle v_i | v_j \rangle \langle v_j| = 0$ 。所以 $\{P_i\}$ 表现地就像一组线性映射空间的正交基。下面这个问题是有趣的。

Question 1.5. 是否 $L(V)$ 中任一个线性映射都是 $\{P_i\}$ 的线性组合？

如果可以，即

$$A = \sum_i a^i P_i$$

那么

$$A(P_j v) = \sum_i a^i P_i P_j v = a^j (P_j v)$$

记 $x = P_j v$ ，这就归结于是否有矢量 x 和标量 λ 满足方程

$$Ax = x\lambda \tag{1.17}$$

如果有，那么问题就解决了一半了。形如(1.17)的方程，我们称之为本征方程。关于本征方程的研究，后面会阐述。

本节的最后，我们写出矩阵乘法的矩阵元。对于 $L(V)$ 中线性映射对应的矩阵的乘法 (BA) ，他矩阵元按照矩阵乘法的法则就是

$$\sum_{i=1}^n \langle v_k | B | v_i \rangle \langle v_i | A | v_j \rangle$$

同时，根据 $P_i = |v_i\rangle\langle v_i|$ ，我们有

$$\sum_{i=1}^n \langle v_k | B | v_i \rangle \langle v_i | A | v_j \rangle = \sum_{i=1}^n \langle v_k | B P_i A | v_j \rangle = \left\langle v_k \left| \sum_{i=1}^n B P_i A \right| v_j \right\rangle$$

所以

$$BA = \sum_{i=1}^n B P_i A$$

也就是说

$$B \left(E - \sum_{i=1}^n P_i \right) A = 0 \quad (1.18)$$

其中 E 是恒同映射（有时候也记作 I 甚至 1 ），作用在任何一个矢量上结果还是那个矢量。因为(1.18)对任意的 A 和 B 都是成立的，那么我们只能说

$$\sum_{i=1}^n P_i = E$$

或者

$$\sum_{i=1}^n |v_i\rangle\langle v_i| = E$$

这个等式在有限维矢量空间中往往是成立的，但是一旦到了无限维就不一定成立了。

1.8 多重线性映射

所谓多重线性映射，就是指我们的映射不再是单一变量了，而是变成了多个，并且对各个变量来说，他们都是线性的。这是线性映射（更准确说是内积）的直接推广，他当然有个大名鼎鼎的名字，张量。我们从二重开始，并且恢复使用求和约定。

设有域 \mathbb{F} 上的右矢量空间 V ，我们考虑映到 \mathbb{F} 上的双线性映射 $f: V \times V \rightarrow \mathbb{F}$ ，也就是满足下面四条法则

$$\begin{aligned} (1). f(x+y, z) &= f(x, z) + f(y, z) \\ (2). f(x, y+z) &= f(x, y) + f(x, z) \\ (3). f(xa, y) &= f(x, y)a \\ (4). f(x, yb) &= f(x, y)b \end{aligned} \quad (1.19)$$

然后假设 V 上的一组基为 $\{v_i\}$, 则

$$f(x, y) = f(v_i x^i, v_j y^j) = f(v_i, v_j) x^i y^j = t_{ij} x^i y^j$$

定义 t 为 (二阶) 张量, 记

$$\langle t, xy \rangle = t_{ij} x^i y^j = f(x, y)$$

在 $\langle t, xy \rangle = f(x, y)$ 中看出, 张量 t 只和给定的 f, x, y 有关, 和 V 中基的选取是无关的。和内积类似的, 这是张量一个很好的性质。 $t_{ij} = f(v_i, v_j)$ 称为张量的分量。

我们求一下变换基之后的张量分量, 设新的基为 $\{u_i\}$, 而对应的坐标的矩阵为 (a_j^i) .

$$t'_{ij} = f(u_i, u_j) = f(v_p a_i^p, v_q a_j^q) = t_{pq} a_i^p a_j^q$$

注意到张量和基的选取无关, 所以将 t'_{ij} 还是写作 t_{ij} , 最后得到了

$$t_{ij} = t_{pq} a_i^p a_j^q$$

实际上, 如果我们记 V' 为 V 的对偶空间, 我们也定义双线性函数 $f: V' \times V \rightarrow \mathbb{F}$, 也就是满足

$$\begin{aligned} (1). f(x + y, v) &= f(x, v) + f(y, v) \\ (2). f(x, u + v) &= f(x, u) + f(x, v) \\ (3). f(ax, v) &= af(x, v) \\ (4). f(x, vb) &= f(x, v)b \end{aligned} \tag{1.20}$$

如果对于这个多线性映射, 对应的张量的分量为 $t^i_j = f(v^i, v_j) = \delta^i_j$, 这正是内积。

同样, 我们就可以定义多线性映射 $f: V \times V \times \cdots \times V' \times V' \times \cdots \rightarrow \mathbb{F}$ 和张量 t 满足 $\langle t, xyz \cdots uvw \cdots \rangle = f(x, y, z, \cdots, u, v, w, \cdots)$. 如果前 p 个变量满足来自于 V , 后 q 个变量来自于 V' , 那么张量的分量的一般形式为

$$t^{abc \cdots}_{ABC \cdots} = f(v_a, v_b, \cdots, v_A, v_B, \cdots)$$

我们称之为 (p, q) 型张量。在 Einstein 求和约定那里说过, 一般将分量指标出现在上面的矢量叫做逆变矢量, 出现在下面的叫协变矢量。所以, 我们称呼分量指标全出现在上面 (即 $q = 0$) 的张量叫做逆变张量, 全出现在下面的 (即 $p = 0$) 叫协变张量, 上下都有的叫做混合张量。因此称呼 p 为逆变阶数, q 为协变阶数。

多重线性映射的出现为我们理所当然推广内积做了足够好的铺垫，尽管其本身也有着足够丰富的内涵。最直接的推广方式，就是把双线性映射 $f: V \times V \rightarrow \mathbb{F}$ 定义为 V 中的内积。所以单一空间中的内积的一般形式就是

$$\langle x, y \rangle = f(x, y) = t_{ij} x^i y^j$$

其余一切定义（比如正交）都类似。一般而言，给定了一组基 $\{v_i\}$ 和一个双线性映射 f ，并不能得到 $f(v_i, v_j) = \delta_{ij}$ ，那么就有这样的问题

Question 1.6. 给定双线性映射 $f: V \times V \rightarrow \mathbb{F}$ ，这就是说给定了内积的形式。那么我们是否可以从 n 维向量空间 V 中挑出一组基 $\{v_i\}$ ，使得 $f(v_i, v_j) = \delta_{ij}$ 成立？

至少在 \mathbb{R}^n 答案是可以的。为此我们演示一种思路，人称 Gram-Schmidt 正交化过程。从两个线性无关的矢量 $\{v_1, v_2\}$ （即不平行）开始，那么首先把 v_1 归一化，即

$$v_1^* = \frac{v_1}{|v_1|}$$

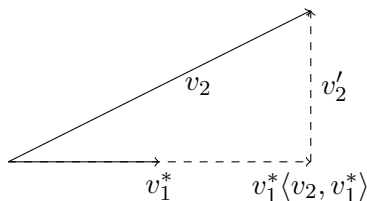
可以检验 $\langle v_1^*, v_1^* \rangle = 1$ 。然后令构造与 v_1^* 垂直的一个矢量

$$v_2' = v_2 - v_1^* \langle v_2, v_1^* \rangle$$

可以检验

$$\langle v_2', v_1^* \rangle = \langle v_2, v_1^* \rangle - \langle v_1^*, v_1^* \rangle \langle v_2, v_1^* \rangle = 0$$

最后对 v_2' 进行归一化就可以了。几何上这就是说， v_2 减去他在 v_1^* 上的投影，那么他剩余部分就自然和 v_1^* 垂直了，然后剩下的就是归一化了。如图。



类似地，如果我们正交归一化了前面 k 个矢量，记作 $\{v_k^*\}$ ，那么用

$$v_{k+1}' = v_{k+1} - \sum_{i=1}^k v_i^* \langle v_{k+1}, v_i^* \rangle$$

就可以正交化，然后对 v_{k+1}' 归一化就完成了前 $k+1$ 个矢量的正交归一化。那么这就是告诉我们，至少在 \mathbb{R}^n 上总可以通过一组基构造出一组正交基。

下面讨论两种特殊的多重线性映射，自然，对应了两种特殊的张量。

1. 反对称多重线性映射

如果 $f(x, y) = -f(y, x)$ ，那么称呼他为反对称的。类似地，称 $f(x, y, z, u, v, w \cdots)$ 为反对称的，就是当任意两个变量交换位置，则函数变号。比如这里我调换 y 和 u ， $f(x, y, z, u, v, w \cdots) = -f(x, u, z, y, v, w \cdots)$ 。然后我们考虑反对称多线性映射 $f(x, y, z)$ 的张量

$$f(x, y, z) = f(v_i, v_j, v_k \cdots) x^i y^j z^k = t_{ijk} x^i y^j z^k$$

$$f(x, z, y) = f(v_i, v_k, v_j \cdots) x^i z^k y^j = t_{ikj} x^i z^k y^j$$

注意到 $y^j z^k = z^k y^j$ （因为域中的标量乘法是可交换的），而且 $f(x, y, z) = -f(x, z, y)$ ，就可以推出张量的分量满足

$$t_{ijk} = -t_{ikj}$$

所以反对称多线性映射的张量交换脚标是会变号的，我们称呼这样的张量为反对称张量，上面只有三个变量，推广到多个并不复杂。

那么如果在一个反对称的多线性映射，出现了相同的变量多次，则这个多线性映射恒为 0。以二元为例 $f(x, x) = -f(x, x)$ ，所以 $f(x, x) = 0$ 。对于反对称张量的分量也有类似结论。一个反问题是有趣的，如果 $f(x, x) = 0$ ，也可以推出 f 反对称。这是因为 $0 = f(x + y, x + y) = f(x, y) + f(y, x)$ ，至于更高阶是否可行，这交由读者。

那么对于一个反对称张量来说，确定他全部的分量，只要确定那些不会出现重复的排列就可以了比如 $t_{123\dots}$ ，而重复的比如 $t_{113\dots}$ 的值都是 0。为此我们选定一个标准 $t_{123\dots n} = a$ ，然后希望用它来找到其他的分量。这是一个置换的问题，为此，把叙述精细化，需要下列定义。

Definition 1.21. 设集合 $N = \{1, 2, 3 \cdots, n\}$ ，对于任意一个双射 $\sigma: N \rightarrow N$ ，我们称为 N 上的一个置换。

一般而言，我们习惯将置换用矩阵写出来，比如

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

上面一排写原排列，后面一排写新的排列。然后这就是说 $\sigma(1) = 2$ 等等。那么我们是否可以通过不断的两两调换位置来从第一排换到第二排呢？这总归是可以的。具体证明细节我就不出了，比如写成矩阵的那个例子可以这样完成：3 位置不变，12 位置对调，然后 24 位置对调。接下来这个暂时不予证明的命题是有趣的，也是至关重要的，需要证明的可以去第二章。

Proposition 1.9. 我们总是可以通过两两调换来完成置换，但方式是不一定。但无论通过那种方式调换完成同一个置换置换，对应的调换的次数 k ，我们必然有 $(-1)^k$ 不变。我们将一个置换 σ 对应的 $(-1)^k$ 记作 $\text{sgn}(\sigma)$ 。

因此我们有定义

Definition 1.22. 对于每一个置换 σ ，如果 $\text{sgn}(\sigma) = -1$ 这个置换称为奇置换，如果 $\text{sgn}(\sigma) = 1$ 这个置换称为偶置换。

回到一开始的问题上，这里我们将目光锁定在一个 n 维向量空间的 n 重反对称线性映射上面。并且为了写得更形式一些，我们需要定义。

Definition 1.23. 设多重线性映射 $f: V \times \cdots \times V \rightarrow \mathbb{F}$ ，定义置换 σ 在上面的作用为 $\sigma f(v_1, v_2, \cdots, v_n) = f(v_{\sigma(1)}, v_{\sigma(2)}, \cdots, v_{\sigma(n)})$ ，且是线性的。

我们自然可以把置换定义到张量分量上

$$\sigma t_{12 \cdots n} = t_{\sigma(1)\sigma(2) \cdots \sigma(n)}$$

所有的置换 σ 对应所有的排列 $\sigma t_{12 \cdots n}$ 。如果我们需要找到他所有反对称张量的分量，其实也就是找到所有的排列 $\sigma t_{12 \cdots n}$ 。由上面的讨论，我们可以通过他最小的两两调换过程，把标准排列对应的张量的分量 $t_{12 \cdots n}$ 变成 $\sigma t_{12 \cdots n}$ ，于是 $\sigma t_{12 \cdots n} = \text{sgn}(\sigma) t_{12 \cdots n} = \text{sgn}(\sigma) a$ （这也可以作为反对称张量的定义 $\sigma t = \text{sgn}(\sigma) t$ ）。所以所有反对称 n 阶张量的分量都仅仅只由 $t_{12 \cdots n} = a$ 决定。令 $a = 1$ ，这决定了一个特殊的张量 d ，也意味着决定了一个特殊的多重线性映射 D 。因为

$$\sigma d_{12 \cdots n} = \text{sgn}(\sigma)$$

$$D(x, y, z \cdots) = \sum_{\sigma} \text{sgn}(\sigma) x_{\sigma(1)} y_{\sigma(2)} z_{\sigma(3)} \cdots$$

这里求和是指对所有置换求和。特别地，因为 $d_{12 \cdots n} = 1$ ，所以

$$D(v_1, v_2, v_3 \cdots, v_n) = d_{12 \cdots n} = 1$$

那么对任意的反对称多线性映射 F ，他的张量设为 s ，我们有

$$s_{12 \cdots n} = a$$

$$\sigma s_{12 \cdots n} = a \text{sgn}(\sigma) = s_{12 \cdots n} \sigma d_{12 \cdots n}$$

所以有

$$F(x, y \cdots) = \sum_{\sigma} \sigma s_{12 \cdots n} x_{\sigma(1)} y_{\sigma(2)} z_{\sigma(3)} \cdots = s_{12 \cdots n} \sum_{\sigma} \sigma d_{12 \cdots n} x_{\sigma(1)} y_{\sigma(2)} z_{\sigma(3)} \cdots$$

或者我们直接写作下面这个漂亮的形式

$$F(x, y \dots) = F(v_1, v_2, \dots, v_n) D(x, y \dots) \quad (1.21)$$

简单来说, 所有的反对称多线性映射 F 都是 D 乘以一个常数而已。

2. 对称多重线性映射

对称多重线性映射和上面反对称是相反的, 调换两个变量位置但是函数值不变, 在二重多线性映射情况下这就是说 $f(x, y) = f(y, x)$. 同样的, 他的张量在调换脚标时也不变 $t_{ij} = t_{ji}$ (对称张量的定义也可以写作 $\sigma t = t$). 我们来展示一种从任意的多重线性映射构造对称多重线性映射的方式. 譬如 $f(x, y)$ 不是对称的, 那么 $g(x, y) = f(x, y) + f(y, x)$ 是对称的. 对于 $n > 2$ 的情况, 把所有的排列加起来也一样可以得到. 同样的问题也可以在反对称里面问, 就是说, 如果我给定一个多重线性映射, 不对称也不反对称, 能不能用他构造出一个反对称多重线性映射? 这交由读者自行探索。

接着让我们考虑这样的一个双线性映射 $f(u, v) = \langle u, x \rangle \langle v, y \rangle$, 其中 $u, v, x, y \in V$, 后者是给定的, 那么对于这个映射的张量 t 有

$$\langle t, xy \rangle = t_{ij} x^i y^j = f(x, y)$$

这等价于

$$t^{ij} u_i v_j = u_i x^i v_j y^j$$

那么就有

$$t^{ij} = x^i y^j$$

我们将这样的张量记作 $t = x \otimes y$, 称 t 为 x 和 y 的张量积。

类似地, 对于任意俩张量 t 和 s , 也通过

$$\langle t \otimes s, abc \dots xyz \dots \rangle = \langle t, abc \dots \rangle \langle s, xyz \dots \rangle$$

来定义张量积。张量积的分量等于各个分量相乘。显然 $(t \otimes s) \otimes r = s \otimes (t \otimes r)$. 这就是说, 张量积满足交换律和结合律。

由于完成了对于各个张量的分析, 本节最后我们来看一下张量所构成的空间。以二重线性映射为例, 所有 $f(x, y)$ 构成了一个集合 W 。现在从中选出 f, g , 对应的张量为 t, s , 然后再设有一个标量 c , 我们有

$$\langle t, xy \rangle + \langle s, xy \rangle = t_{ij} x^i y^j + s_{ij} x^i y^j = (t_{ij} + s_{ij}) x^i y^j$$

如果定义 $t_{ij} + s_{ij} = s_{ij} + t_{ij}$ 是张量 $t + s = s + t$ 的分量, 那么

$$\langle t, xy \rangle + \langle s, xy \rangle = \langle t + s, xy \rangle \quad (1.22)$$

此外

$$\langle t, xy \rangle c = t_{ij} x^i y^j c = (t_{ij} c) x^i y^j$$

如果定义 $t_{ij} c$ 是张量 tc 的分量, 则

$$\langle t, xy \rangle c = \langle tc, xy \rangle \quad (1.23)$$

式(1.22)和式(1.23)给 W 赋予了线性结构, 使得 W 变成了一个矢量空间。这是一个让普通集合加上结构变成空间的鲜活例证。这里尽管只讨论了二重线性映射, 其他的也类似可以得到相同结论。关于这个矢量空间更细致地考察, 留待以后完成。

1.9 行列式

大名鼎鼎的行列式终于千呼万唤始出来啊, 我会说我其实原本根本不想写他的咩 (有种你来打我呀 233333)

还记得(1.9)式中我把一个矩阵

$$A = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & a_l^1 \\ a_1^2 & a_2^2 & \cdots & a_l^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^m & a_2^m & \cdots & a_l^m \end{pmatrix}$$

用列矢量缩写成了

$$(a_1 \ a_2 \ \cdots \ a_l)$$

了吗? (众: 鬼记得啊摔!) 其中每一个 a_i 代表一个 m 维矢量。

Definition 1.24. 对一组 n 维矢量空间 V 中的 n 个矢量 $\{x_i\}$, 定义这组矢量的行列式为 $D(x_1, x_2, \cdots, x_n)$, 其中 D 为上节中定义的特殊反对称多重线性映射。

我们自然地把 $D(a_1, a_2, \cdots, a_l)$ 定义为矩阵 A 的行列式。因为是 n 维矢量空间中挑 n 个矢量, 那么自然地, 行列式只定义在方阵上 (行数等于列数的矩阵)。因为 D 是反对称的多重线性映射, 所以如果矩阵中有两列相同, 则其行列式为 0。

对于一个 $n \times n$ 的矩阵 A , 我们通常使用 $\det A$ 或者 $|A|$ 来表达他的行列式, 并且记作

$$|A| = \det A = \begin{vmatrix} a_1^1 & a_2^1 & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^n & a_2^n & \cdots & a_n^n \end{vmatrix}$$

如果我们选定 n 维向量空间 V 中的一组基 $\{v_i\}$, 那么按照 D 的定义有

$$D(v_1, v_2, \dots, v_n) = 1$$

但是把 $\{v_i\}$ 用列向量写出来, 则变成了

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \cdots \quad v_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

这就是说, $(v_1 \ v_2 \ \cdots \ v_n)$ 对应的矩阵是单位矩阵 E ,

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

那么 $D(v_1, v_2, \dots, v_n) = 1$ 就代表

$$|E| = \det E = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 1$$

根据上面的讨论, 总结一下行列式的性质,

1. 行列式对每列都是线性的;
2. 行列式如果两列相同, 那么行列式为 0;
3. 单位矩阵的行列式为 1.

神奇的是, 如果有个式子也满足这些性质, 那么他就是行列式。这个交由读者证明。

下面是几个关于行列式运算的有名命题

Proposition 1.10. 如果列向量 a_n 是其他列向量 $\{a_1, a_2, \dots, a_{n-1}\}$ 的线性组合, 或者说列向量们线性相关, 则行列式为 0.

Proof. 请自行咨询小霸王学习机, 哪里不会点哪里, so easy. □

Proposition 1.11. 将某一行乘以某个标量后, 行列式值的乘以那个标量。显然推论是, 如果有一行全为 0, 那么行列式值为 0.

Proof. 你是我的小呀小苹果，怎么爱你都不嫌多！ □

Proposition 1.12. 将某一行乘以某个标量然后加到另一行上去，行列式值不变。

Proof. What does the fox say? □

接着，我们需要的证明的就是著名的行列式乘法定理。

Proposition 1.13. 对于 $n \times n$ 矩阵 A 和 B ，有 $\det(AB) = \det(A) \det(B) = \det(BA)$ 。

Proof. 首先由矩阵乘法，我们有

$$A \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} Ab_1 & Ab_2 & \cdots & Ab_n \end{pmatrix}$$

那么 $\det(AB) = D(Ab_1, Ab_2, \dots, Ab_n)$ 然后对于 $D(Ab_1, Ab_2, \dots, Ab_n) = F(b_1, b_2, \dots, b_n)$ 是一个反对称的多重线性映射，按照式(1.21)，可以写成

$$F(b_1, b_2, \dots, b_n) = F(v_1, v_2, \dots, v_n) D(b_1, b_2, \dots, b_n)$$

也就是说

$$D(Ab_1, Ab_2, \dots, Ab_n) = D(Av_1, Av_2, \dots, Av_n) D(b_1, b_2, \dots, b_n) \quad (1.24)$$

我们需要计算一下 Av_i 这些，写出矩阵

$$Av_1 = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_1$$

同样地可以计算得 $Av_i = a_i$ ，所以式(1.24)就变成了

$$\det(AB) = D(Ab_1, Ab_2, \dots, Ab_n) = D(a_1, a_2, \dots, a_n) D(b_1, b_2, \dots, b_n) = \det(A) \det(B)$$

同理 $\det(BA) = \det(B) \det(A)$ ，所以最后得到

$$\det(AB) = \det(BA) = \det(A) \det(B) \quad (1.25)$$

□

接着我们计算一下具体行列式，这需要应用到 Proposition 1.10, 1.11, 1.12. 比如我们在复数域上计算最基本的 2×2 行列式

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

反复应用 Proposition 1.10, 1.11, 1.12

$$k = a \begin{vmatrix} 1 & b \\ c/a & d \end{vmatrix} = a \begin{vmatrix} 1 & 0 \\ c/a & d - bc/a \end{vmatrix} = a \begin{vmatrix} 1 & 0 \\ 0 & d - bc/a \end{vmatrix} = a(d - bc/a) \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = ad - bc$$

可以看到，这是很低效的计算方法，虽然理论上可以得到所有行列式的值。作为一个大习题，读者需要自己发展一种行列式计算方法，即所谓的按列展开。这种方法可以将 $n \times n$ 阶行列式的计算化成 n 个 $(n-1) \times (n-1)$ 阶行列式的计算。大体思路如下： $n \times n$ 阶行列式 $D(a_1, a_2, \dots, a_n)$ 的第一个 a_1 展开为 $a_1 = v_j a_1^j$ ，按照 $D(a_1, a_2, \dots, a_n)$ 是线性映射，有

$$D(a_1, a_2, \dots, a_n) = D(v_j a_1^j, a_2, \dots, a_n) = D(v_j, a_2, \dots, a_n) a_1^j$$

计算 $D(v_j, a_2, \dots, a_n)$ 是什么样的 $(n-1) \times (n-1)$ 行列式？把行列式写出来找找看。

最后我们来讨论一下矩阵的逆，这在前面一直没有讨论，利用行列式，可以给出一个矩阵是否有逆的判别法则。首先，先指出只有方阵才有可能存在逆。如果 A 是 $m \times n$ 的矩阵，但 $m \neq n$ 。那么 A 对应的线性映射是 $L(V, W)$ 里面的元素，而且 V 是 n 维的， W 是 m 维的，那么 V 和 W 一定不同构，之间的一切线性映射一定不是单射。对应的，只有方阵才有可能存在逆。

Definition 1.25. 对于一个 $n \times n$ 的矩阵 A 和单位矩阵 E ，考虑方程 $AB = E$ ，如果存在解 B ，那我们称之为矩阵 A 的逆，记作 A^{-1} ，有 $AA^{-1} = E$ 。

Proposition 1.14. E 的逆为 E 。如果 A 有逆，则 $\det(A)^{-1} = \det(A^{-1}) \neq 0$ ，并且 $AA^{-1} = A^{-1}A = E$ 。所以这就是说， $(A^{-1})^{-1} = A$ 。

Proof. 分开来证明

首先，因为 $EE = E$ ，所以 E 的逆还是 E

因为 $AA^{-1} = E$ ，对两边都取行列式有 $\det(A)\det(A^{-1}) = 1$ ，所以 $\det(A) \neq 0$ 且 $\det(A)^{-1} = \det(A^{-1})$ 。

因为 $AA^{-1} = E$ ，所以 $AA^{-1}A = EA = A$ ，这就是说 $A(A^{-1}A) = A$ ，设 $X = A^{-1}A$ ，使得 $AX = A$ 。我们下面证明 $X = E$ 。为此首先 $AX - AE = A(X - E) = A - A = 0$ 。因为 $X - E = (x_1 - v_1 \ x_2 - v_2 \ \cdots \ x_n - v_n)$ ，则

$$A \begin{pmatrix} x_1 - v_1 & x_2 - v_2 & \cdots & x_n - v_n \end{pmatrix} = \begin{pmatrix} A(x_1 - v_1) & A(x_2 - v_2) & \cdots & A(x_n - v_n) \end{pmatrix} = 0$$

这等价于对任意的 i , $A(x_i - v_i) = 0$, 根据线性映射可逆的一般法则 (零空间只有 0), 如果 A 可逆, 则如果 $Av = 0$ 可以推出 $v = 0$, 那么对于所有的 i , 有 $x_i - v_i = 0$, 这就是说 $X = E$. 那么这就证明了 $A^{-1}A = E$. \square

作为上述命题的逆

Proposition 1.15. 如果一个方阵的行列式不为 0, 那么他可逆。

Proof. 原命题等价于他的逆否命题: 如果一个方阵 A 不可逆, 那么 $\det A = 0$. 而前半句等价于至少一个非零的 x 使得 $Ax = 0$ 成立 (即零空间不只有零矢量)

$$Ax = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} = a_i x^i = 0$$

如果所有的 a_i 中有一个等于 0, 那么根据 Proposition 1.11, 就得到了 $\det A = 0$, 那么就证明完毕了。现在假设他们都不等于 0.

因为 $x \neq 0$, 所以存在至少一个 $x_j \neq 0$, 不失一般性, 不妨让 $j = 1$, 那么

$$\det A = D(a_1, a_2, \cdots, a_n) = \frac{1}{x^1} D(a_1 x^1, a_2, \cdots, a_n)$$

根据 Proposition 1.12, 将其他列矢量的任意线性组合加到第一个列矢量上去, 行列式值不变, 所以 (暂时不用求和约定)

$$\det A = \frac{1}{x^1} D\left(a_1 x^1 + \sum_{i=2}^n a_i x^i, a_2, \cdots, a_n\right) = \frac{1}{x^1} D(0, a_2, \cdots, a_n) = 0$$

\square

这样, 我们就可以说一个方阵的行列式是否为 0 是其不可逆的充要条件。没有逆的方阵称为奇异矩阵, 有逆的方阵称为非奇异矩阵。

1.10 基的变换

在实际使用中, 我们往往会去选取一组好用的基, 而这样就涉及了从原来的基变成一个新的基变换, 我们希望用一个线性映射 (更明确地, 同构映射) 来完成这个变换。这就是基的变换问题。

考虑一个映射 $A \in L(V, W)$, 在选定了 \mathbb{F}^m 中的基 $\{u_i\}$ 和 \mathbb{F}^n 中的基 $\{v_i\}$ 后, 我们可以写出 A 的矩阵 (A) 。同样如果我们通过两个同构映射 $F \in L(\mathbb{F}^m, \mathbb{F}^m), G \in L(\mathbb{F}^n, \mathbb{F}^n)$

使得基变成了 $u'_i = u_j f_i^j$ 和 $v'_i = v_j g_i^j$. 那么我们去写出在 $\{u'_i\}$ 和 $\{v'_i\}$ 下的矩阵 (A') 呢? 方法其实和之前的一样。

首先, 我们可以将 $v'_i = v_j g_i^j$ 写作 $v_i = v'_j h_i^j$, 其中 h_i^j 是 $H = G^{-1}$ 的矩阵元。则

$$Au'_i = Au_k f_i^k = v_j a_k^j f_i^k = v'_l h_j^l a_k^j f_i^k$$

所以 A 在新基下的矩阵的矩阵就是 $(A') = (G^{-1})(A)(F)$, 或者 $A' = G^{-1}AF$. 特别地, $A \in L(V, W)$, 则 $A' = F^{-1}AF$. 这就是方阵的变换规则, 一般称为相似变换, 然后称这俩方阵是相似的。可以看到 $\det A' = \det(P^{-1}AP) = \det(P^{-1}PA) = \det A$, 所以一个线性映射的行列式不随基的改变而改变, 或者说行列式在相似变换下是不变的。

我们从三种方式来看相似变换

1. 固定 P , 那么设 $f_P(A) = P^{-1}AP$, f_P 成了一线性映射, 考察方程 $f_P(A) = 0$ 来寻找零空间. 由于 $\det f_P(A) = \det A$, 如果 A 可逆, 那么 $\det f_P(A) = \det A \neq 0$, 那么 f_P 的零空间将是行列式为 0 的矩阵构成集合的子集。注意, 他的零空间一般不同行列式为 0 的矩阵构成的集合, 这个考察 f_E 就可以了, $f_E(A) = A = 0$, 显然比所有行列式为 0 的矩阵构成集合小。
2. 固定 A , 能证明对于所有同构映射 P , $P^{-1}AP$ 构成一个等价类, 相似是等价关系。
3. 相似变换的几何。这个在平面上进行演示。设 $A = (a_1 \ a_2)$, 我们在平面上画出 a_1 和 a_2 , 这样就等价画出了一个矩阵, 还是采用 $f_P(A) = P^{-1}AP$ 的记法。我们来看一下 f_P 这个线性变换的作用结果, 也就是 a'_1 和 a'_2 。

设

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \quad P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad P^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

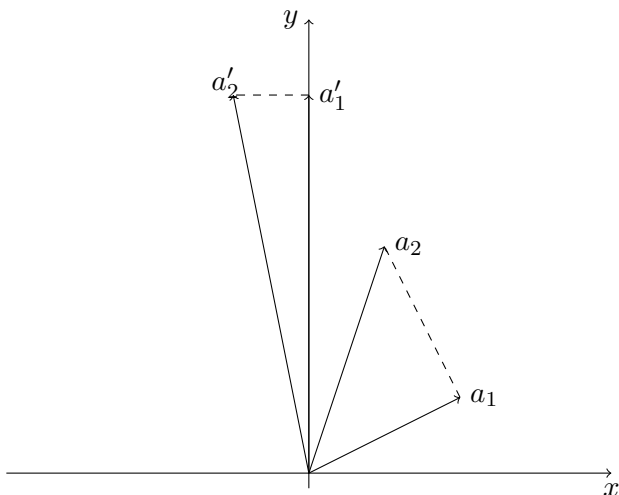
可以计算得

$$P^{-1}AP = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 5 & 5 \end{pmatrix}$$

附加地, 可以计算出 f_P 的矩阵

$$F = \begin{pmatrix} 1/5 & -2/5 \\ 2 & 1 \end{pmatrix}$$

作图,



我们计算一下 Ob_1a_2 和 $Ob'_1a'_2$ 这两个三角形的面积。

$$S(Oa'_1a'_2) = 5 \times 1/2 = 5/2$$

由于夹角 θ 满足 $\langle a_1, a_2 \rangle = |a_1||a_2| \cos \theta$, 所以使用 $S(Oa_1a_2) = (|a_1||a_2| \sin \theta)/2$ (熟知的三角形面积公式)

$$S(Oa_1a_2) = \frac{|a_1||a_2|}{2} \sqrt{1 - \frac{\langle a_1, a_2 \rangle^2}{(|a_1||a_2|)^2}} = \frac{1}{2} \sqrt{(|a_1||a_2|)^2 - \langle a_1, a_2 \rangle^2} = \frac{1}{2\sqrt{2}} \sqrt{5} \sqrt{10} = 5/2$$

所以 $S(Oa'_1a'_2) = S(Oa_1a_2)$, 似乎我们可以断言相似变换是保面积 (体积) 的, 这和行列式又有什么关系? 这个交由读者自己考虑。

1.11 一些补充

1. 线性方程组.

这是一个古老的内容, 似乎一切代数都来自于解方程。解方程一直是代数的重要领域。

下面这个方程组是线性方程组,

$$\begin{cases} a_1^1 x^1 + a_2^1 x^2 + \cdots + a_n^1 x^n = b^1 \\ a_1^2 x^1 + a_2^2 x^2 + \cdots + a_n^2 x^n = b^2 \\ \vdots \\ a_1^m x^1 + a_2^m x^2 + \cdots + a_n^m x^n = b^m \end{cases}$$

当然我们一般会让指标乱飞，避免和幂混淆，通常写作

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

其中 a_{ij} 的 i 指行数 j 指列数。这就是常见的线性方程组。如果记

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

那么原方程就变成了矩阵方程 $Ax = b$ ，其中

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

线性映射 A 的零空间 N 对应着那些 x 使得 $Ax = 0$ 。所以如果找到了一个 $Ax = b$ 的一个解 a （我们称之为特解），那么这个方程所有的解（我们称之为通解）就是所有陪集 $a + N$ 中的元素。这就是线性方程组解的一般结构。可以看到方程 $Ax = 0$ 和他的解是很重要的，这样的方程我们称为齐次方程组，而 $Ax = b$ 就自然称为非齐次方程组。

解 $Ax = b$ 最直接的方式就是 $x = A^{-1}b$ ，但是值得注意的是，对于 A 不是方阵，则不能有逆，所以 $x = A^{-1}b$ 只能对有逆的方阵来说，因为方阵有逆就等价于其行列式不为 0，所以判断方程 $Ax = b$ 有没有唯一解，只要问 A 的行列式是不是 0 就可以了。

2. 对偶算子的补充。

考虑内积 $\langle Fx | Fx \rangle$ ，由于 $\langle Fx | = \langle x | F^\dagger$ ，所以 $\langle Fx | Fx \rangle = \langle x | F^\dagger F | x \rangle$ ，如果 $F^\dagger F = E$ （单位算子）或者 $F^\dagger = F^{-1}$ ，那么我们就有 $\langle Fx | Fx \rangle = \langle x | x \rangle$ 换句话说，内积在 F 算子下不变。这样的算子称为正交算子，其对应的矩阵称为正交矩阵。可以看到，正交算子直接依赖于内积的定义，不同内积对应的正交算子一般是不同的。

Example 1.6. 转置：

首先设, $A \in L(\mathbb{R}^n)$, 所以 $a = a'$. 而以前我们求过 A 的对偶算子 A^\dagger 的矩阵元 $\langle v_i | A^\dagger | v_j \rangle = (\langle v_j | A | v_i \rangle)'$, 所以有 $\langle v_i | A^\dagger | v_j \rangle = \langle v_j | A | v_i \rangle$. 可以看到 A 的第 i 行第 j 列的元素变成了 A^\dagger 的第 j 行第 i 列的元素。

现在我们不限制于 $A \in L(\mathbb{R}^n)$, 而在任何情况下定义一个有关于 A 的新的算子 A^T , 称为 A 的转置, 就是满足

$$\langle v_i | A^T | v_j \rangle = \langle v_j | A | v_i \rangle$$

所以在 $A \in L(\mathbb{R}^n)$ 的情况下, A 的对偶算子即其转置, 也即 $A^\dagger = A^T$. 此时正交算子写作 $F^T F = E$ 或者 $F^T = F^{-1}$.

Example 1.7. 共轭:

现在我们考虑复数域上的矢量空间, 定义对偶标量为复数共轭, 即 $a' = a^*$. 则

$$\langle v_i | A^\dagger | v_j \rangle = (\langle v_j | A | v_i \rangle)^*$$

可以看到 A 的第 i 行第 j 列的元素的共轭变成了 A^\dagger 的第 j 行第 i 列的元素。此时正交算子写作 $F^\dagger = F^{-1}$.

Definition 1.26. 考虑 \mathbb{F} 上的 V 和 $L(V)$ 中算子对应的矩阵的群 $M(n, \mathbb{F})$, 定义如下几种矩阵群 (群运算为矩阵乘法, 请自行检验):

1. 一般线性群: $GL(n, \mathbb{F}) = \{A \in M(n, \mathbb{F}) | \det A \neq 0\}$;
2. 特殊线性群: $SL(n, \mathbb{F}) = \{A \in M(n, \mathbb{F}) | \det A = 1\}$;
3. 正交群: $O(n, \mathbb{F}) = \{Q \in M(n, \mathbb{F}) | Q^\dagger Q = Q Q^\dagger = E\}$;
4. 特殊正交群: $SO(n, \mathbb{F}) = O(n, \mathbb{F}) \cap SL(n, \mathbb{F})$;
5. 一般将 $SO(n, \mathbb{C})$ 记作 $SU(n)$.

3. Gauss 消元法等

可以注意到, 前面我们始终没有给出一个解线性方程的有效方法。一切都在理论上, 比如我们知道解 $Ax = b$ 只要知道 A^{-1} 就可以了, 但是我们根本不知道怎么求 A^{-1} . 这里将要给出一些应用上的方法, 并且可以看到, 他在理论上也是好用的。

1.12 小结

本章主要介绍了 (有限维) 矢量空间和线性映射的相关知识, 这其实是线性代数的内容, 如果再继续下去, 就未免有些喧宾夺主, 大概 50 页的内容已经足够丰富。在本章内, 我没有证明一个非常重要的结论, 就是 Proposition 1.4, 在后面诸多的 Proposition 和其证明中, Proposition 1.4 都起到了至关重要的作用。所以希望得到更严密基础的人,

可以参考 B.L.van der Waerden 的 *Algebra* 第一卷第四章（这也是我的主要参考文献）或者 Michael Artin 的 *Algebra* 第三章。值得一提的是，看似用坐标或者基定义矢量空间是一个很自然的方式，但是一旦涉及无穷维空间就不会是那么便利了。

一般矢量空间的定义在上世纪才完成，而他的发展在之前百年就已经开始，可以说是既古老又年轻的一个数学。矢量空间是一个很常见的数学对象，在微积分里面充斥着矢量空间，因为矢量空间是一个很重要的工具。所以我们在谈论微积分之前一定要有一定的矢量空间知识，这对阅读下面的内容是有益处的。

Chapter 2

Groups and Algebra*

这章将深入探讨一些代数上的东西，这对后面的内容没什么影响，但是本身也足够有趣，可以作为上一章许多内容的延伸和拓展。

2.1 群论的一些基础概念

这一节的许多概念都可以在向量空间里面找到对应，可以看到，这样的抽象是很自然的。

首先还是把群的定义列出来。

Definition 2.1. 有一个集合 E 和其上的二元运算 $*$ ，或者记作 $(E, *)$ ，称为一个群 (group)，如果满足：

- (1). 结合律：对于任意 $a, b, c \in E$ ，有 $(a * b) * c = a * (b * c)$;
- (2). 单位元：对于任意的 $a \in E$ ，存在一个元素 $e \in E$ ，使得 $e * a = a * e = a$;
- (3). 反元素：对于每一个 $a \in E$ ，存在一个元素 $b \in E$ ，使得 $b * a = a * b = e$.

正如同我们在向量空间研究中看到的一样，研究群和群之间的映射也是重要的。特别地，这种映射还要能保持群的结构。所以有定义：

Definition 2.2. 如果一个群 $(G, *)$ ，和另一个群 (G', \times) ，设满射 $f : G \rightarrow G'$ 满足 $f(a * b) = f(a) \times f(b)$ ，则称 f 为同态，特别地，如果 f 是单射，则称 f 为同构。

显然 $f(e_G) = e_{G'}$ 。对应着向量空间，线性映射就是一个同态。那么同态何时称为同构？这在线性映射里面我们也研究过。只有零空间为 0 的时候才是。所以我们有定义 $\ker(f) = \{a \in G | f(a) = e_{G'}\}$ ，称为同态 f 的核 (kernel)。只有核为单位元的时候，同态才成为一个同构。设 H 是一个同态 f 的核，类似于零空间的陪集，对应着每一个元素

$a \in G$ 我们可以有左陪集 $a * H = \{a * h | h \in H\}$. 类似于通解和特解, 陪集 $a * H$ 所有 $f(x) = a$ 的解的集合。同样, 可以定义右陪集。

对应着子空间, 我们有子群的概念。

Definition 2.3. 有一个群 $(G, *)$, 有一个集合 $H \in G$, 如果 $(H, *)$ 也是一个群, 则称 $(H, *)$ 为 $(G, *)$ 的子群 (subgroup), 或者再简单一些, H 是 G 的子群。

可以检验, 同态的核是一个子群。但是, 核还有这更加优秀的性质: 他的左陪集等于右陪集。证明交给读者。有这样性质的子群是特殊的, 所以我们有定义:

Definition 2.4. 左陪集等于右陪集的子群称为正规子群 (normal subgroup)。

正规子群之所以是特殊的, 这是因为我们可以使用正规子群对整个群进行划分, 分成诸不相交集的并 (或者说等价类), 这其实在矢量空间中已经做过了。

Proposition 2.1. 现在有一个正规子群 H , 可以如下定义等价关系: 如果 $b * a^{-1} \in H$, 则 $a \sim b$.

Proof. 习题。 □

Remark. 可以看到 $a \sim b$ 当且仅当 $b \in H * a = a * H$, 所以等价关系也可以这样描述, 如果 a 和 b 同时在一个正规子群的陪集里, 则两者等价。

所以, 类似于 Proposition 1.8 下面的 remark 里面的处理。我们有如下定义。

Proposition 2.2. 设 H 是 G 的一个正规子群, 商群 G/H 定义为所有陪集的集合。

下面再罗列一下定义。

Definition 2.5. 现有群 G, G_1, G_2 , 定义:

- (1). 有限群 G 的元素个数称为群的阶, 记作 $|G|$.
- (2). 群元 h 关于 g 的共轭 (conjugation) 记作 $Conj_g(h) := ghg^{-1}$.
- (3). 群 G 的核 (center) 被定义为集合 $\{g \in G | \forall h \in G, hg = gh\}$.
- (4). 两个群 G_1, G_2 的直积 (direct product) 定义为集合 $G_1 \times G_2 := \{(g_1, g_2) | g_1 \in G_1, g_2 \in G_2\}$, 上面的乘法定义为 $(g_1, g_2)(g'_1, g'_2) = (g_1g'_1, g_2g'_2)$.

2.2 群作用和群表示

先定义群作用

Definition 2.6. 设 G 是一个群, E 是一个集合。则所谓群作用就是定义了一个运算 $f: G \times E \rightarrow E$, 如果 $g, h \in G, a \in E$, 则运算记作 $g \cdot a$, 且满足 $h \cdot (g \cdot a) = (hg) \cdot a$. 而对于 G 中的单位元 e , 他的作用 $e \cdot a = a$.

简单来说, 可以说 G 是 E 上的可逆变换构成的群。

现在定义群表示的概念,

Definition 2.7. 让 V 是一个域 \mathbb{F} 上的有限维向量空间:

群 G 的一个表示 U 指存在这样的群同态 $U: G \rightarrow Gl(V)$, 使得

$$U(g)U(g') = U(gg'), U(g^{-1}) = U(g)^{-1}, U(e_G) = e_{Gl(V)}$$

成立。 $\dim V$ 称为表示 U 的维度。

2.3 有限群的表示

如果 $\langle \cdot, \cdot \rangle$ 是 V 上的一个内积, 如果对任意的 $g \in G, u, v \in V$ 有

$$\langle u, v \rangle = \langle U(g)u, U(g)v \rangle$$

我们称呼其为么正 (unitarizable) 表示。有限群有着一个足够好的性质, 即为, 其所有表示都是么正的, 只需定义新的内积 $\langle \cdot, \cdot \rangle'$ 为

$$\langle u, v \rangle' = \frac{1}{|G|} \sum_{g \in G} \langle U(g)u, U(g)v \rangle$$

即可。证明的关键如果 h 跑遍整个群, 则 gh 也跑遍整个群。而群元素有限总可以让这个新的内积是良定义的, 所以有限群所有表示都是么正的。

现在我们有一个 V 的子空间 W , 如果对所有 g 有 $U(g)W \subset W$, 则 W 称为在 U 下的不变子空间。如果 V 在 U 没有不平凡的不变子空间的话, 则称 U 为不可约表示, 否则是可约的。

2.4 对称群

2.5 代数

2.6 张量代数

这节所有向量空间都同时定义左乘右乘且相等, 这就是说, 如果 a 是标量, v 是向量, av 和 va 是同一个向量。同时这节也可以看做一个超大的习题。

在 1.10 节我们已经阐述过了, (同一型的) 张量构成一个矢量空间。很简单就可以检验, 矢量是一个张量, 那么我们就可以定义矢量空间的张量积, 方式很简单 $V \otimes W = \{x | x = v \otimes w, v \in V, w \in W\}$. 我们再一次扩充多线性映射的定义

Definition 2.8. 设有矢量空间 V_1, V_2, \dots, V_m, W , 映射 $f: V_1 \times V_2 \times \dots \times V_m \rightarrow W$ 被称为多线性映射, 如果他对每一个变元都是线性的。而所有这样多线性映射的集合记作 $L(V_1, V_2, \dots, V_m; W)$.

然后要证明下面一个有趣而重要的命题, 他可以用来证明 $L(V, W; X)$ 和 $L(V \otimes W; X)$ 是同构的。

Proposition 2.3. 考虑三个矢量空间 V, W, X 和一个双线性映射 $h: V \times W \rightarrow V \otimes W$

$$h(v, w) = v \otimes w$$

则对任意的双线性映射 $f: V \times W \rightarrow X$ 存在唯一的线性映射 $g: V \otimes W \rightarrow X$, 使得

$$f = gh: V \times W \rightarrow X$$

Proof. 设 V, W, X 都是矢量空间, 前两者基分别为 $\{v_i\}$ 和 $\{w_j\}$, 定义线性映射 $g: V \otimes W \rightarrow X$, 使他在基上的作用是 $g(v_i \otimes w_j) = f(v_i, w_j)$, 若设

$$v = v_i a^i, w = w_j b^j$$

那么

$$g(v \otimes w) = g(v_i \otimes w_j) a^i b^j = f(v_i, w_j) a^i b^j = f(v, w)$$

显然, g 唯一确定, 且 $f = gh: V \times W \rightarrow X$. □

定义 $\phi: L(V, W; X) \rightarrow L(V \otimes W; X)$ 使得 $\phi(g) = gh$, 显然 ϕ 是一个双射, 且是线性的, 所以 ϕ 是一个同构映射。

Proposition 2.4. 考虑两个矢量空间 V, W , V 和 W 分别是 m, n 维的, 那么 $V \otimes W$ 也是一个矢量空间, 且是 $m \times n$ 维的。

Proof. $V \otimes W$ 是矢量空间前面已经说过了。设 V 的基为 $\{v_i\}$, W 的基为 $\{w_j\}$, 以及 $x \in V, y \in W$, 那么

$$x \otimes y = x^i y^j v_i \otimes w_j$$

要成为一组基, 首先是他要线性独立, 很容易证明 $\{v_i \otimes w_j\}$ 线性独立。由于任意 $V \otimes W$ 中的元素都是他的线性组合, 所以他是一组 $V \otimes W$ 的基, 一共 $m \times n$ 个, 所以 $V \otimes W$ 是 $m \times n$ 维的。 □

在定义张量积的时候，我们考虑了一个双线性映射 $f(v, w) = \langle x, v \rangle \langle y, w \rangle$ ，其中 $v, w, x, y \in V$ ，当时固定了 $x, y \in V$ 。由于上面证明过 $L(V, V; \mathbb{F})$ 和 $L(V \otimes V; \mathbb{F})$ 同构，所以存在线性映射 $g(v \otimes w) = \langle x, v \rangle \langle y, w \rangle$ ， g 由 x, y 直接确定。与此同时，如果我们固定 v 和 w ，那么 g 又可以写作 $g(x, y) \in L(V \otimes V; \mathbb{F})$ ，因为 $L(V, V; \mathbb{F})$ 和 $L(V \otimes V; \mathbb{F})$ 同构，所以存在线性映射 $h(x \otimes y) = \langle x, v \rangle \langle y, w \rangle$ 。

综合上面两条，我们得到了一个双线性映射 $h(x \otimes y, v \otimes w) = \langle x, v \rangle \langle y, w \rangle$ 。我们干脆顺势将其定义成内积！记作 $\langle x \otimes y, v \otimes w \rangle = \langle x, v \rangle \langle y, w \rangle$ 。

接着我们来求对偶基。 $V \otimes V$ 的基为 $\{v_i \otimes v_j\}$ ，那么 $\langle v_k \otimes v_l, v_i \otimes v_j \rangle = \langle v_k, v_i \rangle \langle v_l, v_j \rangle = \delta_i^k \delta_j^l$ 。只有当 $i = k$ 且 $j = l$ 的时候才为 1，其他时候都为 0，这就是说对偶矢量 $(v_i \otimes v_j)'$ ，对 $\{v_i \otimes v_j\}$ 这组总共有 n^2 个元素的基来说，只有作用在 $v_i \otimes v_j$ 上才是 1，其他都是 0。这正是对偶基！没有理由不让 $(v_i \otimes v_j)' = v^i \otimes v^j$ ，这就是说 $\{v_i \otimes v_j\}$ 的对偶基为 $\{v^i \otimes v^j\}$ 。类似地，

Definition 2.9. 定义内积：

$$\langle v_1 \otimes \cdots \otimes v_k, u_1 \otimes \cdots \otimes u_k \rangle = \langle v_1, u_1 \rangle \cdots \langle v_k, u_k \rangle$$

Proposition 2.5. 人话：对偶空间的张量积是原空间张量积的对偶。下面是非人话表述：

$$(V \otimes W)' = V' \otimes W'$$

接着我们可以用矢量的张量积构造出 n^{p+q} 个 (p, q) 型张量。对于 p 个矢量空间 V ，和 q 个对偶空间 V' ，他们的基分别为 $\{v_i\}$ 和 $\{v^i\}$ 。则

$$v^{k_1} \otimes \cdots \otimes v^{k_q} \otimes v_{i_1} \otimes \cdots \otimes v_{i_p}$$

可以证明他们线性独立，所以构成 (p, q) 型张量的空间 V_q^p （特别地，约定 $V_0^0 = \mathbb{F}, V_0^1 = V, V_1^0 = V'$ ）的基。也就是说，任意的 (p, q) 型张量都可以按这个基打开

$$x = x^{i_1 \cdots i_p}_{k_1 \cdots k_q} v^{k_1} \otimes \cdots \otimes v^{k_q} \otimes v_{i_1} \otimes \cdots \otimes v_{i_p}$$

那么可以很容易证明，如果 t 是 (p_1, q_1) 型张量， s 是 (p_2, q_2) 型张量，那么 $t \otimes s$ 是 $(p_1 + p_2, q_1 + q_2)$ 型张量。

Definition 2.10. 取两指标 r, s 。假设在 V 和 V' 中分别有一组矢量 $\{v_i\}$ 和 $\{u^i\}$ （不一定为基），

$$x = u^{k_1} \otimes \cdots \otimes u^{k_q} \otimes v_{i_1} \otimes \cdots \otimes v_{i_p}$$

构成一个 (p, q) 型张量。定义

$$C_{rs}(x) = \langle u_r, v_s \rangle u^{k_1} \otimes \cdots \otimes u^{k_r} \otimes \cdots \otimes \widetilde{u^{k_q}} \otimes v_{i_1} \otimes \cdots \otimes \widetilde{v_{i_s}} \otimes \cdots \otimes v_{i_p}$$

其中记号 $\widetilde{u^{k_q}}$ 指在连续的张量积中去掉该因子, 那么映射 $C_{rs} : V_q^p \rightarrow V_{q-1}^{p-1}$ 。此外, 再让映射 C_{rs} 是线性的, 这样的映射我们称之为缩并。

Example 2.1. 作为张量的矩阵。

现在设有一个 $L(V)$ 的线性映射。在选定了 V 的基之后, 线性映射有着矩阵 (a^j_i) , 所以矩阵的分量可以看成是一个 $(1,1)$ 型张量的分量, 那么我们可以直接将这两个东西等同。对于 $(1,1)$ 型张量, 我们求一下他的缩并。

首先, $(1,1)$ 型张量 a 只能表示成 $a = v^i \otimes v_j a^j_i$, 那么

$$C_{11}(a) = \langle v_i, v_j \rangle a^j_i = \delta^i_j a^j_i = a^j_j$$

这也就是说, 缩并在矩阵上的作用就是将所有对角元素加起来。一般而言, 我们称之为矩阵的迹。请乃们自己证明, 一个线性映射对应的矩阵的迹只和线性映射有关, 而和选取的基无关。

更一般的缩并, 也请小伙伴们自己尝试。

在讲外积前, 有必要再一次接触反对称张量的概念。回忆线性映射置换的定义为 $\sigma f(v_1, \dots, v_n) = f(v_{\sigma(1)}, \dots, v_{\sigma(n)})$, 那么对于张量积 $u_1 \otimes \dots \otimes u_n$, 同样定有

$$\sigma(u_1 \otimes \dots \otimes u_n) = u_{\sigma(1)} \otimes \dots \otimes u_{\sigma(n)}$$

因为前面 σ 是定义在 $V \otimes \dots \otimes V$ 上的线性算子, 那么我们顺便求一下他在 $V \otimes \dots \otimes V$ 的对偶空间 $V' \otimes \dots \otimes V'$ 上的对偶算子 σ^\dagger , 这不失为一种有趣的尝试。因为

$$\langle u_1 \otimes \dots \otimes u_n, \sigma(u_1 \otimes \dots \otimes u_n) \rangle = \langle u_1 \otimes \dots \otimes u_n, u_{\sigma(1)} \otimes \dots \otimes u_{\sigma(n)} \rangle = \langle u_1, u_{\sigma(1)} \rangle \dots \langle u_n, u_{\sigma(n)} \rangle$$

注意到上式右边就等于

$$\langle u_1, u_{\sigma(1)} \rangle \dots \langle u_n, u_{\sigma(n)} \rangle = \langle u_{\sigma^{-1}(1)}, u_1 \rangle \dots \langle u_{\sigma^{-1}(n)}, u_n \rangle$$

同时对于对偶算子有

$$\begin{aligned} \langle (\sigma^\dagger(u^1 \otimes \dots \otimes u^n))', u_1 \otimes \dots \otimes u_n \rangle &= \langle (u^{\sigma^\dagger(1)} \otimes \dots \otimes u^{\sigma^\dagger(n)})', u_1 \otimes \dots \otimes u_n \rangle \\ &= \langle u_{\sigma^\dagger(1)} \otimes \dots \otimes u_{\sigma^\dagger(n)}, u_1 \otimes \dots \otimes u_n \rangle \\ &= \langle u_{\sigma^\dagger(1)}, u_1 \rangle \dots \langle u_{\sigma^\dagger(n)}, u_n \rangle \end{aligned}$$

按照对偶算子的定义

$$\langle (\sigma^\dagger(u^1 \otimes \dots \otimes u^n))', u_1 \otimes \dots \otimes u_n \rangle = \langle u_1 \otimes \dots \otimes u_n, \sigma(u_1 \otimes \dots \otimes u_n) \rangle$$

于是

$$\langle u_1, u_{\sigma^\dagger(1)} \rangle \cdots \langle u_n, u_{\sigma^\dagger(n)} \rangle = \langle u_1, u_{\sigma(1)} \rangle \cdots \langle u_n, u_{\sigma(n)} \rangle = \langle u_1, u_{\sigma^{-1}(1)} \rangle \cdots \langle u_n, u_{\sigma^{-1}(n)} \rangle$$

所以 $\sigma^\dagger = \sigma^{-1}$. 于是如果 $x = v^1 \otimes \cdots \otimes v^n$, 则

$$\sigma^\dagger x = v^{\sigma^{-1}(1)} \otimes \cdots \otimes v^{\sigma^{-1}(n)}$$

Definition 2.11. 记 $T^p(V) = V_0^p$ 和 $T^q(V') = V_q^0$, 分别对应逆变和协变张量的情况。将所有 p 阶反对称逆变张量的集合记作 $\Lambda^p(V)$, 所有 p 阶对称逆变张量的集合记作 $P^p(V)$.

很容易证明, $P^p(V)$ 和 $\Lambda^p(V)$ 都是 $T^p(V)$ 的子空间。后面将会引出外积的概念。外积, 在近代物理和数学都有着很多的用处, 但展示其丰富内涵并非我的主要目的, 而是作为一个大例子, 使用前面的一些工具。

Definition 2.12. p 阶反对称逆变张量又叫做 p -外矢量, 空间 $\Lambda^p(V)$ 被称为 V 上的 p -外矢量空间。还像之前一样约定 $\Lambda^1(V) = V$ 和 $\Lambda^0(V) = \mathbb{F}$ 。

Definition 2.13. 对任意的 $x \in T^p(V)$, 令

$$S_p(x) = \frac{1}{p!} \sum_{\sigma} \sigma x$$

$$A_p(x) = \frac{1}{p!} \sum_{\sigma} (-1)^{k(\sigma)} \sigma x$$

前者称为 p 阶逆变张量的对称化算子和反对称化算子, 他们都是投影算子 (所以也是线性的), 证明交给读者。他们将任意一个 p 阶逆变张量变成了一个 p 阶对称逆变张量和 p 阶反对称逆变张量。(这在前面问过)

在 q 阶协变张量上我们也可以进行类似讨论, 这时候可以使用上面的 σ^\dagger 。

Definition 2.14. 设 ξ 和 η 分别是 m -外矢量和 n -外矢量, 定义 ξ 和 η 的外积 $\xi \wedge \eta$ 为

$$\xi \wedge \eta = \frac{(m+n)!}{m!n!} A_{m+n}(\xi \otimes \eta)$$

$\xi \wedge \eta$ 是一个 $(m+n)$ -外矢量。

Proposition 2.6. 设有三个 m -外矢量 ξ, ξ_1, ξ_2 , 三个 n -外矢量 η, η_1, η_2 和一个 h -外矢量 ζ , 有

(1). 分配律:

$$(\xi_1 + \xi_2) \wedge \eta = \xi_1 \wedge \eta + \xi_2 \wedge \eta$$

$$\xi \wedge (\eta_1 + \eta_2) = \xi \wedge \eta_1 + \xi \wedge \eta_2$$

(2). 反变换律: $\xi \wedge \eta = (-1)^{mn} \eta \wedge \xi$

(3). 结合律: $(\xi \wedge \eta) \wedge \zeta = \xi \wedge (\eta \wedge \zeta)$

Proof. (1) 由反对称化算子线性和张量积线性显然。

(2) 设置换

$$\tau = \begin{pmatrix} 1 & \cdots & m & m+1 & \cdots & m+n \\ 1+n & \cdots & m+n & 1 & \cdots & n \end{pmatrix}$$

接着我们来求出一个调换次数完成了置换 τ .

首先给出反序需要的步数

$$\begin{pmatrix} 1 & 2 & \cdots & l \\ l & l-1 & \cdots & 1 \end{pmatrix}$$

如果 l 是一个偶数, 挨次从 i 为 0 到 $l/2$ 将 i 和 $l-i$ 调换, 一共调换了 $l/2$ 次, 完成了反序。如果 l 是一个奇数, 挨次从 i 为 0 到 $(l-1)/2$ 将 i 和 $l-i$ 调换, 一共调换了 $(l-1)/2$, 完成了反序。

对于 τ , 可以将前 n 个先反序, 然后将后 m 个反序, 最后整体 $m+n$ 个反序。分三步

$$\tau_1 = \begin{pmatrix} 1 & \cdots & m & m+1 & \cdots & m+n \\ m+n & \cdots & 1+n & 1 & \cdots & n \end{pmatrix}$$

$$\tau_2 = \begin{pmatrix} 1 & \cdots & m & m+1 & \cdots & m+n \\ m+n & \cdots & 1+n & n & \cdots & 1 \end{pmatrix}$$

$$\tau_3 = \begin{pmatrix} 1 & \cdots & m & m+1 & \cdots & m+n \\ 1 & \cdots & m & m+1 & \cdots & m+n \end{pmatrix}$$

如果 m, n 都是偶数, 那么一共需要 $m+n$ 步, 为偶数步。如果 m, n 都是奇数或者一奇一偶, 那么一共需要 $m+n-1$ 步, 前者为奇数步, 后者为偶数步。所以奇偶关系和 mn 是一样的。这就证明了 $(-1)^{k(\tau)} = (-1)^{mn}$ 。

接着开始正式的证明。首先按基打开直接写出

$$\xi \wedge \eta = \frac{\xi^{i_1 \cdots i_m} \eta^{j_1 \cdots j_n}}{m!n!} \sum_{\sigma} (-1)^{k(\sigma)} \sigma(v_{i_1} \otimes \cdots \otimes v_{i_m} \otimes v_{j_1} \otimes \cdots \otimes v_{j_n})$$

$$\eta \wedge \xi = \frac{\xi^{i_1 \cdots i_m} \eta^{j_1 \cdots j_n}}{m!n!} \sum_{\sigma} (-1)^{k(\sigma)} \sigma(v_{j_1} \otimes \cdots \otimes v_{j_n} \otimes v_{i_1} \otimes \cdots \otimes v_{i_m})$$

注意到 σ 跑遍所有置换, 那么 σ 和 τ 的复合 $\sigma' = \sigma\tau$ (先进行 τ , 然后进行 σ) 也跑遍

所有置换, 且 $(-1)^{k(\sigma')} = (-1)^{k(\sigma)}(-1)^{k(\tau)}$, 所以

$$\begin{aligned}
 \eta \wedge \xi &= \frac{\xi^{i_1 \dots i_m} \eta^{j_1 \dots j_n}}{m!n!} \sum_{\sigma'} (-1)^{k(\sigma')} \sigma'(v_{j_1} \otimes \dots \otimes v_{j_n} \otimes v_{i_1} \otimes \dots \otimes v_{i_m}) \\
 &= \frac{\xi^{i_1 \dots i_m} \eta^{j_1 \dots j_n}}{m!n!} \sum_{\sigma} (-1)^{k(\sigma)} (-1)^{k(\tau)} \sigma(v_{i_1} \otimes \dots \otimes v_{i_m} \otimes v_{j_1} \otimes \dots \otimes v_{j_n}) \\
 &= (-1)^{k(\tau)} \frac{\xi^{i_1 \dots i_m} \eta^{j_1 \dots j_n}}{m!n!} \sum_{\sigma} (-1)^{k(\sigma)} \sigma(v_{i_1} \otimes \dots \otimes v_{i_m} \otimes v_{j_1} \otimes \dots \otimes v_{j_n}) \\
 &= (-1)^{k(\tau)} \xi \wedge \eta
 \end{aligned}$$

代入 $(-1)^{k(\tau)} = (-1)^{mn}$ 即得证 (2).

第三个性质, 也就是结合律留习题吧, 犯晚期懒癌了. 需要证明的是:

$$(\xi \wedge \eta) \wedge \zeta = \xi \wedge (\eta \wedge \zeta) = \frac{(m+n+h)!}{m!n!h!} A_{m+n+h}(\xi \otimes \eta \otimes \zeta)$$

□

当 $v, w \in V$ 的时候, 显然 $v \wedge w = -w \wedge v$ 且 $v \wedge v = 0$. 设 $\{v_i\}$ 是 V 的一组基, 对结合律进行归纳可得 n 个的结合律, 有

$$v_{i_1} \wedge \dots \wedge v_{i_k} = k! A_k(v_{i_1} \otimes \dots \otimes v_{i_k})$$

只要在 $v_{i_1} \wedge \dots \wedge v_{i_k}$ 中有重复出现的元素, 那么这个式子就为 0. 特别地, 当 k 比 V 的维数大的时候, $v_{i_1} \wedge \dots \wedge v_{i_k}$ 中必有重复, 所以必为 0.

任何 k -外矢量 x , 都可以分解成

$$x = x^{i_1 \dots i_k} v_{i_1} \otimes \dots \otimes v_{i_k}$$

由于 $A_k x = x$, 所以

$$x = x^{i_1 \dots i_k} A_k(v_{i_1} \otimes \dots \otimes v_{i_k}) = \frac{x^{i_1 \dots i_k}}{k!} v_{i_1} \wedge \dots \wedge v_{i_k}$$

这样我们也找到了外矢量空间的基就是原矢量空间的基的外积组合而成。

Chapter 3

Topological Space and Continuous Function

在这章里，我们主要处理的是另一种空间——拓扑空间，在这个空间里面，我们加上的结构使得我们可以研究点与点之间的连接关系，用比较生活化的语言来说，研究的就是我们怎么区分这个点是不是在那个点附近等问题。但是，为了让我们的教程不至于一下子变得太过抽象，就让我们先处理一种特殊的拓扑空间，即度量空间。在度量空间中，我们定义了所谓的“距离”。而度量空间也是我们在后面章节讨论所处的基本环境。

正如在上一章中讨论的，研究一个集合（空间）是不够的，而且要研究其上的关系，最最主要的，是要研究其上的映射，比如矢量空间上的线性映射，可以注意到线性同构映射是保持线性结构的（参考 Proposition 1.7）。在拓扑空间上，我们主要研究的映射就是可能能够保持拓扑结构的映射——连续映射。

与连续映射相关的，我们会一步一步建立序列的极限，函数的极限的相关命题，这些是下面章节讨论的基本工具。此外，我们还要给出一个重要的拓扑空间，即实直线，或者说实数域。实数域的建立是一个复杂而漫长的过程，我们会采用直接定义或者按照直觉的方式来避开这个漫长的讨论。

作为引论的最后，我们来比较两个集合的个数，其本身也是一个好玩的问题。

当一个集合的个数有限时，比较两个集合的个数多少可以直接比较。但当集合元素的个数是无限的时候，个数无法得到了，因此无法比较元素的多少。但我们有时候也要去比较两个无限集谁的元素多。比如实数集和有理数集哪一个元素多？

所以人类引入势的概念¹。有限集的势等于这个集合元素的个数。而对于两个集合势的比较，则依靠于两个集合间的关系，比如一一映射。

¹可自行搜索 Hilbert 旅馆问题

我们定义, 当两个集合的势相同的时候, 则称两个集合是等势的。如果一个集合和另一个集合之间存在双射, 即 X 的元素 x 和 Y 的元素 y 同时满足 $f(x) = y$ 和 $f^{-1}(y) = x$, 那么这两个集合是等势的。而当集合 X 和 Y 的某个子集等势的时候, 则 X 的势小于等于 Y 的势。

Definition 3.1. 如果一个集合和正整数集 \mathbb{N} 等势, 则称这个集合为可数的, 叫做可数集。如果一个集合是有限的或可数的, 则称为至多可数的。

作为习题, 试证明, 至多可数个至多可数集的并是至多可数的。

3.1 度量空间

3.1.1 度量函数

度量函数还有一个比较熟悉的名字“距离”, 我们说定义了距离的空间就自然成为度量空间。因为独立空间有了距离, 所以在他上面可以清晰地明了点与点之间的连接关系。我们从在上一章中已经知道了在 \mathbb{R}^n 中的距离。如果 $a, b \in \mathbb{R}^n$, 那么将 $\|a - b\| = \sqrt{\langle a - b, a - b \rangle} = \sqrt{\sum_i (a_i - b_i)^2}$ 定义为他们之间的距离², 我们现在来提取出它的一些性质。

Proposition 3.1. 对任意的三个点 $a, b, c \in \mathbb{R}^n$,

1. $\|b - a\| = \|a - b\| \geq 0$, 当且仅当 $a = b$ 的时候不等式取等号;
2. $\|a\|^2 \|b\|^2 \geq \langle a, b \rangle^2$;
3. $\|a - b\| + \|b - c\| \geq \|a - c\|$.

Proof. 性质 1 由定义显然。性质 2 就是有名的 Cauchy 不等式, 或者写作

$$\left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n b_i^2 \right) \geq \left(\sum_{i=1}^n a_i b_i \right)^2 \quad (3.1)$$

证明方法很多, 仅举其一。

设有一个参数 t , 显然对任意的 t 我们有 $\langle a + bt, a + bt \rangle \geq 0$. 因为标准内积是双线性的, 且 $\langle a, b \rangle = \langle b, a \rangle$, 所以

$$\langle a + bt, a + bt \rangle = \langle a, a + bt \rangle + \langle b, a + bt \rangle t = \langle a, a \rangle + 2\langle a, b \rangle t + \langle b, b \rangle t^2 \geq 0$$

²我们将 $|$ 留给 \mathbb{R} 上的绝对值, 尽管可以轻松地证明, 在 \mathbb{R} 上 $\|$ 和 $|$ 表达的是同一个意思, 但我们还是区分一下符号, 在这将来是有益处的。

因为这对所有的 t 都成立，所以作为 t 的二次函数，比如有二次项系数为正且其判别式小于等于 0. 前者已经满足，后者就是不等式

$$(2\langle a, b \rangle)^2 - 4\langle a, a \rangle \langle b, b \rangle \leq 0$$

注意到 $\langle a, a \rangle = \|a\|^2$ ，我们就有

$$\|a\|^2 \|b\|^2 \geq \langle a, b \rangle^2$$

或者更精简一些

$$\|a\| \|b\| \geq |\langle a, b \rangle| \geq \langle a, b \rangle$$

接着我们证明性质 3，先证明 $\|p\| + \|q\| \geq \|p + q\|$ ，

$$(\|p\| + \|q\|)^2 - \|p + q\|^2 = (\|p\| + \|q\|)^2 - \langle p + q, p + q \rangle = 2\|p\|\|q\| - 2\langle p, q \rangle$$

由 Cauchy 不等式显然有 $(\|p\| + \|q\|)^2 \geq \|p + q\|^2$ ，由于 $\|p\|$ 等非负，所以

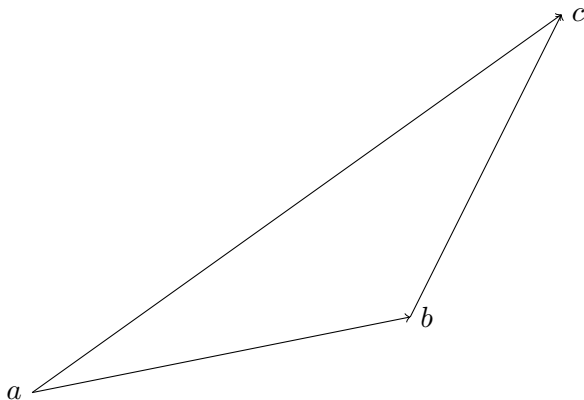
$$\|p\| + \|q\| \geq \|p + q\|$$

然后将 $p = a - b, q = b - c$ 代入

$$\|a - b\| + \|b - c\| \geq \|a - b + b - c\| = \|a - c\|$$

□

性质 1 告诉我们两个不同点之间距离总是正的，只有相同点之间的距离才是 0，在我们推广距离概念的时候，这条性质应该满足。而性质 2 依托于 \mathbb{R}^n 上标准内积的性质，重要的是，他和距离没有关系，但用其证明的性质 3 却是有关距离的不等式。其几何意义如图：



对于 a, b, c 三点, a 到 c 的距离比 a 先到 b 然后再从 b 到 c 的距离短, 简单说, 三角形两边之和大于第三边。如果我们定义抽象的距离, 这点也是希望可以满足的³, 所以有这样的定义。

Definition 3.2. 对于一个集合 E , 其中任意取三点 a, b, c , 如果有一个实值函数 $d: E \times E \rightarrow \mathbb{R}$, 满足如下两个性质:

1. $d(a, b) = d(b, a) \geq 0$, 当且仅当 $a = b$ 的时候不等式取等号;
2. $d(a, b) + d(b, c) \geq d(a, c)$.

这样的函数被称为度量函数。如果一个集合上存在度量函数, 则称这个集合为度量空间。

Remark. 一个度量空间上的度量函数不一定只有一个。比如我们可以在任意的集合上定义 $d(a, b) = 1 - \delta_{ab}$ 为度量函数⁴, 这样的度量函数意义不大。我们称这样的度量空间为平凡的度量空间。

有时候有些书也会采用 $\|a - b\|$ 来代替 $d(a, b)$, 这其实很不科学, 因为我们未必在度量空间 E 上定义了减法。

显然 \mathbb{R} 可以是一个非平凡的度量空间, 比如选取 $d(a, b) = |a - b|$. 再比如在 \mathbb{R}^n 上选取 $d(a, b) = \|a - b\|$.

3.1.2 开集和闭集

这节里面要尝试精确描述什么是“附近”, 这涉及了一个很重要的概念“开集”.

³当然, 可以不要性质 3, 定义更抽象的距离, 但是显然证明的结论就会少很多。

⁴ δ_{ab} 定义为 $a = b$ 的时候为 1, 否则为 0.

3.2 连续函数

3.2.1 函数的极限

3.2.2 空间的完备性

3.2.3 \mathbb{R} 的一些基本性质

在这节里面我们选取 $d(a, b) = |a, b|$ 为 \mathbb{R} 上的度量函数，称为 \mathbb{R} 上的标准度量。

3.2.4 连续函数

3.2.5 紧集和一致连续

Definition 3.3. 假设空间 X 有一个集合 E , 对于 E 来说, 存在 X 中的开集族, 这个开集族中所有开集的并可以覆盖 E (就是 E 包含于其中), 那么这样的开集族就称为 E 的开覆盖。如果从任意一个 E 的开覆盖中可以挑选出有限个开集, 他们的并依旧可以覆盖这个 E (有限开覆盖), 则这样的集合 E 被称为紧集。

Proposition 3.2. 紧集的性质:

- (1). 紧集必然闭。
- (2). 紧集的闭子集是紧集。

Proof. (1). 设 K 是一个紧集, 只要证明 K 的补集是一个开集就可以了。取一点 p 在 K 的补集里。再取一点 q 在 K 里面, 令 V_p 和 W_q 分别是 p 和 q 的邻域, 他们的半径小于 $\frac{1}{2}d(p, q)$. 因为 K 是紧集, 所以 K 中有有限多个点 q_1, q_2, \dots, q_n , 使得,

$$K \subset W_{q_1} \cup W_{q_2} \cup \dots \cup W_{q_n} = W$$

如果令 $V = V_{P_1} \cap V_{P_2} \cap \dots \cap V_{P_n}$, 那么 V 就是 p 的一个和 W 不相交的邻域。因此 V 在 K 的补集里面, 所以 p 是 K 的补集的内点。证毕。

(2). 设 $F \subset K \subset X$, 其中 F 是闭的, K 是紧的, 而 X 是一个度量空间。取 F 的一个开覆盖 $\{V_a\}$, 然后把 F 的补集加进去 (开的), 他构成了一个 K 的开覆盖, 由于 K 是紧集, 所以可以从中挑出有限的开集依旧覆盖 K 。如果 F 的补集不在这些有限开覆盖里面, 则这些开覆盖也覆盖 F , 则命题证毕。如果在里面, 把他去掉, 则剩下的那些有限个开集的并依旧可以覆盖 F 。命题证毕。

□

3.3 \mathbb{R}^n 中的 Heine-Borel 定理

Proposition 3.3. \mathbb{R}^n 中的 Heine-Borel 定理:

以下三个条件在 \mathbb{R}^n 中等价:

- (a). E 是有界闭集。
- (b). E 是紧集。
- (c). E 中的每个无限子集在 E 内有极限点。

Proof. 1. 我们按照 $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (a)$ 的顺序进行证明。因为 $(a) \Rightarrow (b)$ 已经由 10/17 的第二题给出了, 所以下面先证明 $(b) \Rightarrow (c)$ 。

这里不限定在 \mathbb{R}^n 中, 而是可以放在任意紧集 E 里面。假设 F 是他的一个无限子集。如果 E 里面没有 F 的极限点, 那么每一个 $q \in F$ 将有一个邻域 V_q , 他最多含有 F 的一个点 (如果 $q \in F$, 那么这个点就是 q)。显然, 没有 $\{V_q\}$ 的有限子组可以覆盖 F 。因为 $F \subset E$, 所以对于 E 也一样, 没有 $\{V_q\}$ 的有限子组可以覆盖 E 。但这与 E 是个紧集矛盾, 所以 $(b) \Rightarrow (c)$ 成立。

最后证明 $(c) \Rightarrow (a)$ 。如果 E 不是有界的, 那么挑出这样的一个子列 $|x_n| > n$, 构成了一个无界无限子集。显然在 \mathbb{R}^k 中没有极限点。所以 E 是有界的。

如果 E 不是闭集, 那么存在一点 $x_0 \in \mathbb{R}^k$, 他是 E 的极限点, 但是却不在于 E 里面。对于 $n = 1, 2, 3, \dots$, 存在点 $x_n \in E$, 使得 $|x_n - x_0| < 1/n$ 。令 S 就是这些点构成的集合, 显然这是 E 的无限子集。如果 $y \in \mathbb{R}^k$, 且是一个不同于 x_0 的 S 的极限点。那么总可以挑出 $|x_0 - y| > 2/n$, 则

$$|x_n - y| \geq |x_0 - y| - |x_n - x_0| \geq |x_0 - y| - \frac{1}{n} \geq \frac{1}{2}|x_0 - y|$$

那么 y 就不可能是 S 的一个极限点。因此 S 在 E 里面没有极限点。

所以上面证明了, 如果 E 不是闭集, 则存在一个无限子集在 E 中没有极限点。其逆否命题就是, 如果每一个 E 中的无限子集都在 E 中都有极限点, 那么 E 就是闭集。 $(c) \Rightarrow (a)$ 证毕。

命题证毕。这就给出了 \mathbb{R}^n 中的 Heine-Borel 定理的完整形式。 \square

Remark. 可以看到 $(b) \Rightarrow (c)$ 并不依托于 \mathbb{R}^n 的结构, 可以适用于任何度量空间。但是 $(a) \Rightarrow (b)$ 却用了 \mathbb{R}^n 的结构。那么似乎提示了我们, 在一般的度量空间中, (a) 是推不出 (b) 的。事实正是如此。而对于 (b) 和 (c) , 则可以证明他们是等价的。可见 (a) 是其中最弱的条件。

3.4 拓扑空间

3.5 线性赋范空间

Appendix

Backmatter