

数学物理：其基础的现代介绍

Mathematical Physics: A modern Introduction to Its Foundations

目录

第一章 数学准备	3
1.1 集合	3
1.1.1 等价关系	6
1.2 映射	9
1.3 度量空间	13
1.4 基数	16
1.5 数学归纳法	19
1.6 本章习题	20
第一部分 有限维矢量空间	23
第二章 矢量与线性映射	27
2.1 矢量空间	27
2.1.1 子空间	30
2.1.2 商空间	32
2.1.3 直和	35
2.1.4 矢量空间的张量积	39
2.2 内积	40
2.2.1 正交性	44
2.2.2 Gram-Schmidt 过程	46
2.2.3 Schwarz 不等式	48
2.2.4 矢量的长度	50
2.3 线性映射	53
2.3.1 线性映射的核	57
2.3.2 线性同构	60
2.4 复结构	64
2.5 线性泛函	69
2.6 多重线性映射	75
2.6.1 线性算符的行列式	79

2.6.2 古典伴随	79
2.7 本章习题	82
第三章 代数	87
3.1 从矢量空间到代数	87
3.1.1 一般性质	88
3.1.2 同态	96
3.2 理想	99
3.2.1 商代数	104
3.3 全矩阵代数	105
3.4 代数的导子	107
3.5 代数的分解	111
3.5.1 代数的根	112
3.5.2 半单代数	118
3.5.3 单代数的分类	122
3.6 多项式代数	125
3.7 本章习题	128
第四章 算符代数	131
4.1 自同态代数 $\text{End}(\mathcal{V})$	131

1

数学准备

本章为介绍性章节, 汇总了本书会用到的那些最为基本的工具以及概念. 同时, 我们还介绍了现代数学物理文献当中一些通用的词汇以及记号. 那些熟悉诸如集合、映射、等价关系、度量空间等概念的读者可以跳过本章.

1.1. 集合

现代数学会从集合 (set) 这一基本 (但却未经定义) 的概念开始. 我们可以将集合看作是一些对象 (object) 构成的无结构族 (structureless family), 或者无结构组 (structureless collection).¹ 比如说, 我们可以说某大学中学生的集合、某城市中男人的集合、某公司中女性劳动者的集合、空间中向量的集合、平面内点的集合、时空连续统 (continuum of space-time) 中事件的集合, 等等.² 对于一个集合 A 而言, 它里面的任意成员 a 称作这个集合的一个元素 (element). 这一关系记作 $a \in A$ (读作 “ a 是 A 的元素”, 或者 “ a 属于 A ”), 其否定形式为 $a \notin A$, 即 a 不是 A 的元素, 或者说 a 不属于 A . 有时候, 为强调其几何内涵 (geometric connotation), 我们会将 a 称作集合 A 中的点 (point).

通常我们会在一对大括号内枚举集合的元素来指定这个集合. 比如说, $\{2, 4, 6, 8\}$ 就表示前四个正偶数构成的集合;³ $\{0, \pm 1, \pm 2, \pm 3, \dots\}$ 就表示所有整数的集合; $\{1, x, x^2, x^3, \dots\}$ 为 x 所有非负次幂之集合; $\{1, i, -1, -i\}$ 是四次单位根的集合. 在很多实际场景中, 我们通过集合中所有元素满足的一条 (数学) 语句来定义这个集合. 这样的集合通常记

¹译者注: 此处涉及英文中对集合的两种常用表述方式, 一种是 a family of xx , 一种是 a collection of, 在翻译为中文时我们通常都会将其翻译为 xx 的集合, 如果这里 xx 本身就是集合, 则将其翻译为 xx 的集族. 此处为了在中文中区分这两个说法, 我将 family 翻译为族, 而将 collection 翻译为组. 但之后不会在翻译时区分这两种说法.

²译者注: 连续统 continuum 是个数学概念, 最初只是实数集 \mathbb{R} 的另一个名字, 用以强调实数的连续性. 现在则泛化为一个一般概念, 表示具备某种连续特征的数学对象, 在不同数学分支有着不同内涵.

³译者注: 自然数集 \mathbb{N} 是否包含 0 是个相当微妙的话题, 在不同的数学领域中, 数学研究者有着不同的偏好. 在本书中, 自然数与非负整数同义, 即 $0 \in \mathbb{N}$. 此外, 原文中对这个集合的说法是 even natural numbers, 即偶的自然数, 因此作者似乎默认奇数和偶数都是对正整数而言的, 但这与通常的说法不一致. 因此在翻译时改成了正偶数.

作 $\{x \mid P(x)\}$, 读作“使得 $P(x)$ 为真的所有 x 构成的集合”. 前面给出的那些例子就可以通过这种方式表述如下:

$$\begin{aligned} & \{n \mid n \text{ 为偶数且 } 1 < n < 9\}, \\ & \{\pm n \mid n \text{ 为自然数}\}, \\ & \{y \mid y = x^n, \text{ 这里 } n \text{ 为自然数}\}, \\ & \{z \mid z^4 = 1 \text{ 且 } z \text{ 为复数}\}. \end{aligned}$$

在常用的简化记号下, 后面两个集合可以缩写为 $\{x^n \mid n \geq 0 \text{ 且 } n \text{ 为自然数}\}$ 以及 $\{z \in \mathbb{C} \mid z^4 = 1\}$. 类似地, 单位圆 (unit circle) 就可以记作 $\{z \in \mathbb{C} \mid |z| = 1\}$, 闭区间 (closed interval) $[a, b]$ 就是 $\{x \mid a \leq x \leq b\}$, 开区间 (open interval) (a, b) 就是 $\{x \mid a < x < b\}$, 而 x 的非负次幂构成的集合简写为 $\{x^n\}_{n=0}^{\infty}$ 或者 $\{x^n\}_{n \in \mathbb{N}}$, 其中 \mathbb{N} 是自然数集 (即非负整数的集合). 本书会经常使用最后这个记号. 只有一个元素的集合称作单点集 (singleton). ⁴

如果 $a \in B$ 时总有 $a \in A$, 我们就称 B 是 A 的子集 (subset), 并记作 $B \subseteq A$ 或者 $A \supset B$, 读作“ B 含于 A ”或“ A 包含 B ”. 如果 $B \subseteq A$ 且 $A \subseteq B$, 则称 $A = B$. 如果 $B \subseteq A$ 但是 $A \neq B$, 则称 B 是 A 的真子集 (proper subset). 由 $\{a \mid a \neq a\}$ 定义的集合称作空集 (empty set), 记作 \emptyset . ⁵ 很明显, \emptyset 中没有元素, 并且是任意集合的子集. 给定集合 A , 其所有子集 (含空集 \emptyset) 的集族称作 A 的幂集 (power set), 记作 2^A . 采用这一记号的原因在于, 当 n 有限时, 由 n 个元素构成的集合, 其所有子集的数目为 2^n (见『习题 1.1』).

* 此处 I 称作指标集 (index set), 也称作计数集 (counting set), α 是这个集合中的一个代表元素. 在多数情况下, I 就是 (非负) 整数集. 但是, 原则上, 它可以是任意集合, 比如说实数集.

如果 A 和 B 均为集合, 由那些要么属于 A , 要么属于 B , 要么同时属于 A 和 B 的元素所构成的集合就称作 A, B 的并集 (union), 记作 $A \cup B$. 而那些同时属于 A 和 B 的元素所构成的集合就称作 A, B 的交集 (intersection), 记作 $A \cap B$. 设 $\{B_\alpha\}_{\alpha \in I}$ 是一族集合, * 我们将它们的并集记作 $\bigcup_{\alpha \in I} B_\alpha$, 将它们的交集记作 $\bigcap_{\alpha \in I} B_\alpha$.

⁴译者注: 这里给出的集合描述法 $\{x \mid P(x)\}$ 是朴素集合论 (naive set theory) 中的记号, 但这种记号有个问题: 哪些 $P(x)$ 是合规的数学语句? 大数学家兼哲学家 Russel 就给出了著名的 Russel 悖论, 具体的矛盾就是取 $P(x)$ 为 $x \notin x$. 为了解决这个悖论, 人们提出了公理集合论, 在这个体系下集合的描述法表述要通过分离公理模式给出, 具体就是写作 $\{x \in S \mid P(x)\}$, 对此感兴趣的读者可以自行检索相关内容.

⁵译者注: 作者选用了 \emptyset 这个记号, 不过 \varnothing 这个记号现在更为常见, 故翻译时采纳后者.

集合 A 的补集 (complement, 亦译作余集) 记作 $\sim A$, 其定义为⁶

$$\sim A := \{a \mid a \notin A\}.$$

集合 B 在 A 中的余集 (或者它们的差集, difference) 为

$$A \sim B := \{a \mid a \in A \text{ 且 } a \notin B\}.$$

在使用集合论时, 总会有个基本的全集 (universal set), 它的子集就是我们所研究的那些对象. 通常而言, 这个全集在具体语境下是清晰的. 比如说, 在研究整数的性质时, 全集就是所有整数的集合, 记作 \mathbb{Z} . 而在实分析中, 全集就是所有实数的集合, 即 \mathbb{R} ; 类似地, 复分析中的全集就是所有复数的集合 \mathbb{C} . 为强调全集 X 的存在, 我们可以将补集具体写作 $X \sim A$ 而非 $\sim A$.

从两个给定集合 A 和 B 出发, 我们就可以构造它们的 **Descartes 积** (Cartesian product, 或译作卡氏积, 中文常见说法为笛卡尔积), 其符号为 $A \times B$, 这是序对 (ordered pair) (a, b) 的集合, 其中 $a \in A, b \in B$. 在集合论的记号下, 它表示如下:

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

我们可以将这个概念推广到任意有限多个集合. 设 A_1, \dots, A_n 为集合, 那么它们的 Descartes 积就是

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i\},$$

它是有序 n 元组 (ordered n -tuples) 的集合. 如果 $A_1 = A_2 = \cdots = A_n = A$, 则我们通常会将 $A \times A \times \cdots \times A$ 简写作 A^n , 即

$$A^n = \{(a_1, a_2, \dots, a_n) \mid a_i \in A\}.$$

我们最熟悉的 Descartes 积的例子就是 $A = \mathbb{R}$ 的情况. 此时 \mathbb{R}^2 就是点对 (x_1, x_2) 的集合, 其中 $x_1, x_2 \in \mathbb{R}$. 这其实就是平面里面点的集合. 类似地, \mathbb{R}^3 就是三元组 (x_1, x_2, x_3) 的集合, 或者说空间里面点的集合, 而 $\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R}\}$ 就是所有实 n 元组构成的集合.

⁶译者注: 在通常的数学文献中, 补集的符号还有 A^c , $C_X A$ 以及 \bar{A} 等. 其中 $C_X A$ 强调了全集 X 的存在, 但绝大多数情况下全集都是语境自明的, 因此更多的书籍喜欢 A^c 和 \bar{A} 这组记号. 另外, 差集通常记作 $A - B$ 或者 $A \setminus B$. 本书采用的这个记号大概率是来自数理逻辑那边, 他们喜欢用 \sim 或者 \neg 表示否定. 另外, 这里用到的符号 \coloneqq 意思是将冒号左边的内容定义为右边的内容. 这就是所谓定义号, 亦写作 $\stackrel{\text{def}}{=}$, \triangleq 等, 在物理书里面经常写作 \equiv (恒等于). 原书采用的也是物理书里用恒等于表示定义的写法, 但这种写法和之前的两种符号一样都无法区分定义内容和被定义内容, 因此翻译时我选择了数学书中现在普遍采用的 \coloneqq 写法以强调定义方向.

1.1.1. 等价关系

在很多情况下, 集合中的一些元素会自然地组合在一起. 比方说, 两个磁矢势之间如果只相差一个标量函数的梯度, 那么它们就可以视作同一类, 因为它们给出相同的磁场. 类似地, 两个量子态(已归一化)如果只相差一个模为 1 的复数乘法因子, 那么它们就可以视作同一类, 因为它们表示相同的物理态. 将这些想法抽象化之后, 我们就引入了下述定义⁷.

定义 1.1.1 (关系、等价关系、等价类). 设 A 为集合, 它上面的一个关系 (*relation*), 就是对 A 中有顺序的一对元素之间进行的一种比较测试 (*comparison test*). 如果序对 $(a, b) \in A \times A$ 通过了这个测试, 我们就记 $a \triangleright b$, 读作 “ a 与 b 相关 (或 a 相关于 b)” (a is related to b). 特别地, 如果某个关系 \triangleright 满足下述性质:

$$\begin{aligned} a \triangleright a & \quad \forall a \in A, && \text{(自反性, } reflexivity) \\ a \triangleright b \Rightarrow b \triangleright a & \quad a, b \in A, && \text{(对称性, } symmetry) \\ a \triangleright b, \text{ 且 } b \triangleright c \Rightarrow a \triangleright c & \quad a, b, c \in A. && \text{(传递性, } transitivity) \end{aligned}$$

那么, 我们就称这是个等价关系 (*equivalence relation*). 此时, 若 $a \triangleright b$, 我们就称 “ a 等价于 b ” (a is equivalent to b). 并且, 这种情况下, 所有等价于 a 的元素之集合 $[\![a]\!] = \{b \in A \mid b \triangleright a\}$ 就称作 a 的等价类 (*equivalence class*).

请读者验证下面所述等价关系的性质.

命题 1.1.2. 设 \triangleright 是 A 上的等价关系, $a, b \in A$. 那么, 要么 $[\![a]\!] \cap [\![b]\!] = \emptyset$, 要么 $[\![a]\!] = [\![b]\!]$.

因此, $a' \in [\![a]\!]$ 就可以推出 $[\![a']\!] = [\![a]\!]$. 换句话说, 等价类中任意一个元素都可以选作这个等价类的代表元 (*representative*). 由于等价关系满足对称性, 有时候我们也会将其记作 \bowtie .

例 1.1.3. (a) 设 A 为人类的集合. 我们将 $a \triangleright b$ 解释为 “ a 年龄比 b 大”. 很明显, \triangleright 是 A 上的关系, 但不是等价关系. 但换个解释, 若将 $a \triangleright b$ 解释为 “ a 和 b 居住在同一座城市”, 那么 \triangleright 就是个等价关系, 读者自可验证其满足等价关系的定义. 而 a 的等价类就是 a 所在城市的总人口.

⁷译者注: 下文对关系的定义用了个模糊的说法(比较测试), 这可能把定义反而变复杂了. 实际上, 所谓的关系就是 Descartes 积 $A \times A$ 的任意子集, 给定它的一个子集, 我们就确定了一个关系. 举个例子, 我们考虑 $\mathbb{R} \times \mathbb{R}$ 的子集 $R = \{(a, b) \in \mathbb{R}^2 \mid a > b\}$. 那么只要 $(x, y) \in R$, 我们就说 $x \triangleright y$, 此时这里的 \triangleright 其实就是两个实数的大小关系.

- (b) 设 V 是磁矢势的集合. 若存在某个标量函数 f 使得 $\mathbf{A} - \mathbf{A}' = \nabla f$, 我们就记 $\mathbf{A} \triangleright \mathbf{A}'$. 读者可以验证, 这里的 \triangleright 就是个等价关系, 并且 $[\![\mathbf{A}]\!]$ 就是那些和 \mathbf{A} 给出同一磁场的磁矢势的集合.
- (c) 设底集为 $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$.⁸ 假若 $ad = bc$, 则称 “ (a, b) 相关于 (c, d) ” . 可以验证, 这个关系是个等价关系. 不仅如此, 我们可以将等价类 $[(a, b)]$ 等同于比值 a/b .⁹

定义 1.1.4 (划分). 设 A 为集合, $\{B_\alpha\}$ 是其子集族. 如果这些 B_α 互不相交 (*disjoint*), 即没有共同元素, 并且满足 $\bigcup_\alpha B_\alpha = A$, 则称 $\{B_\alpha\}$ 是 A 的一个划分 (*partition*, 也译作分划或者剖分), 也称 $\{B_\alpha\}$ 将 A 划分.

现在考察 A 中所有等价类的集合 $\{[\![a]\!] \mid a \in A\}$. 这个集合中的元素互不相交, 并且它们的并集覆盖了整个 A . 因此, 集合 A 所有等价类构成的集合就构成 A 的一个划分. 若等价关系为 \bowtie , 则这个集族就记作 A/\bowtie , 并称其为 A 在等价关系 \bowtie 下的商集 (*quotient set*, 也写作 *factor set*, 后者通常也译作商集).

例 1.1.5. (a) 设底集为 \mathbb{R}^3 . 在 \mathbb{R}^3 这样定义等价关系: 若 $P_1, P_2 \in \mathbb{R}^3$ 落在同一条过原点的直线上,¹⁰ 则称 P_1, P_2 等价. 这一定义下, \mathbb{R}^3/\bowtie 就是空间中所有过原点的直线所构成的集合. 如果我们将沿给定直线, 第三个坐标分量为正的单位矢量选作这条直线的代表元, 那么 \mathbb{R}^3/\bowtie (称作关联于 \mathbb{R}^3 的射影空间, *projective space associated with \mathbb{R}^3*) 几乎就是 (但并不完全是) 上单位半球. 二者区别在于半球边缘上落在同一直径上的两个点要等同起来, 这样才能将上单位半球转化为射影空间.

(b) 在整数集 \mathbb{Z} 上, 我们这样定义 \triangleright : 对 $m, n \in \mathbb{Z}$, 如果 $m - n$ 可以被 k 整除 ($m - n$ is divisible by k), 其中 k 是个固定的整数, 那么就称

⁸译者注: 集合 A 上如果带有额外结构 (所谓结构, 就是集合上的关系或者运算), 我们就将去除结构后的 A 称作对应结构的底集 (*underlying set*). 类似的术语还有底空间 (*underlying space*), 它和底集含义差不多. 不过使用底空间这个术语的时候通常指这个集合上带有多个结构, 我们去除某些结构后剩下的结构就叫底空间. 比如说当我们研究内积空间时, 去除内积结构后底集上还带有线性空间结构, 因此将其称作底空间. 当然, 这俩术语有合流的趋势.

⁹译者注: 等同 (*identification*) 是个常见的数学术语. 一般而言, 如果我们定义了一个等价关系 \bowtie . 那么当我们说将 a 与 b 等同时, 实际上就是在说 $a \bowtie b$. 当然, 更常见的情况是通过说 a 和 b 等同定义这样一个等价关系. 等同的另一种用法就是此处所述含义, 即将两组概念视作是一个东西的不同表示方法. 举个例子, 汉语中数字 1 写作 “一”, 记账时则写作壹, 于是我们就可以说 “一等同于壹”. 采用这一说法的原因是现代数学更关注对象之间的联系 (即所谓结构). 对于同一结构, 我们可以有不同的实现方式 (称作表示, *representation*), 这些表示虽然表面上不同, 但是如果我们从结构的角度去审视它们, 它们完全无法区分 (所谓 “鸭子论”: 如果一个动物看上去像鸭子, 叫声像鸭子, 吃起来也是鸭肉口感, 那么它就是鸭子), 因而 “等同”. 对于这个词, 可以联系物理学中的术语全同粒子 (*identical particles*), 它们的含义实际上是差不多的. 这个理念会在后续学习中不断深化, 特别是在学习完群论之后.

¹⁰译者注: 换言之, 它们的连线过原点.

$m \triangleright n$. 这样定义的 \triangleright 不仅是个关系, 还是个等价关系. 在这种情况下, 我们有

$$\mathbb{Z}/\triangleright = \{\llbracket 0 \rrbracket, \llbracket 1 \rrbracket, \dots, \llbracket k-1 \rrbracket\},$$

我们敦促读者验证这一结果.¹¹

- (c) 对于『例 1.1.3』中定义在 $\mathbb{Z} \times (\mathbb{Z} \sim \{0\})$ 上的等价关系, 商集 $(\mathbb{Z} \times (\mathbb{Z} \sim \{0\}))/\bowtie$ 就可以等同于有理数集 \mathbb{Q} .¹²

¹¹译者注: 这里定义的等价关系 \triangleright 称作同余 (congruence), 在此处其具体含义是 m 和 n 除以 k 后所得余数相等. 举个例子, 取 $k=3$, 则 $0, 3, 6, \dots$ 除以 3 余数都是 0, 于是我们称它们模 3 同余 0; $1, 4, 7, \dots$ 除以 3 余数都是 1, 于是我们称它们模 3 同余 1; $2, 5, 8, \dots$ 除以 3 余数都是 2, 称它们模 3 同余 2. 在数论中, 若 $m \in \llbracket p \rrbracket$, 则记 $m := p \pmod{k}$.

¹²译者注: 我们经常会遇到在数集中扣除 0 这个点的情况. 更进一步, 我们会遇到在某个环 R 中去掉加法零元 0_R 的情况. 方便起见, 人们约定 R^* 表示 $R \sim \{0_R\}$ 这个集合. 这样一来非零整数就可以记作 \mathbb{Z}^* . 高中时我们经常用 \mathbb{N}^* 表示正整数就是这个原因. 本书作者没有引入这个记号, 但在后面翻译时如果不至于产生符号上的混淆(毕竟右上角加星号是个相当常见的记号), 偶尔我也会采用这一写法.

1.2. 映射

为建立起集合之间的联系, 我们引入映射这一概念. 所谓从集合 X 到集合 Y 的映射 (map) f , 就是 X 中诸元素与 Y 中元素之间的对应, 并且要满足两个条件: (1) X 中所有元素都在 Y 中有对应; (2) 每个 X 只对应 Y 中的一个元素 (见图 1.1). 我们将这样的映射记作 $f : X \rightarrow Y$ 或者 $X \xrightarrow{f} f$. 如果 $x \in X$ 在映射 f 下对应了元素 $y \in Y$, 我们就记

$$y = f(x) \text{ 或 } x \mapsto f(x) \text{ 或 } x \xrightarrow{f} y,$$

并称这里的 $f(x)$ 为 x 在 f 下的像 (image). 因此, 根据映射的定义, $x \in X$ 只能有一个像. 这里的集合 X 称作映射 f 的定义域 (domain), 而 Y 称作 f 的陪域 (codomain) 或者到达域 (target space). 当我们说两个映射 $f : X \rightarrow Y$ 和 $g : X \rightarrow Y$ 相等的时候, 实际上是在说对所有的 $x \in X$ 均有 $f(x) = g(x)$. ¹³

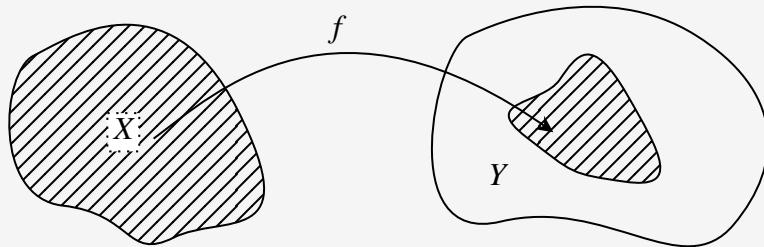


图 1.1: 映射 f 将整个集合 X 满映到 Y 的一个子集上. 图中 Y 内的阴影部分就是 f 的值域 (range), 记作 $f(X)$.

定义 1.2.1 (函数). 所谓函数 (function), 就是陪域为实数集 \mathbb{R} 或者复数集 \mathbb{C} 的映射.

对于每个集合 A , 我们都可以定义一个特殊的映射 $\text{id}_A : A \rightarrow A$, 称作 A 的恒等映射 (identity map), 其定义为

$$\text{id}_A(a) = a \quad \forall a \in A.$$

给定映射 $f : A \rightarrow B$, 它的图象 (graph) 指 $A \times B$ 的一个特定子集 Γ_f , 其定义为

$$\Gamma_f = \{(a, f(a)) \mid a \in A\} \subseteq A \times B.$$

¹³译者注: 请注意此处定义相等映射时陪域是相同的. 原则上讲, 映射有三要素: 定义域、陪域、对应关系. 这三个要素必须全部相同才能说是同一个映射. 但是在有些文献中则只要求定义域和对应关系相同, 因为他们将陪域默认为值域, 但这会导致满射成为无效概念, 本书就没有采用这种约定.

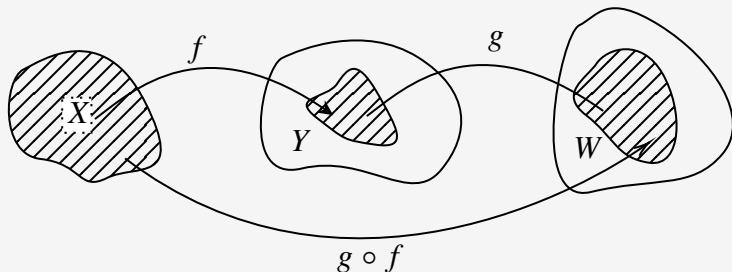


图 1.2: 两个映射的复合是另一个映射.

当 $A = B = \mathbb{R}$, 并视 $A \times B$ 为 xy 平面时, 上面这个一般定义就和我们在代数学以及微积分中遇到的那个(函数)图象意义一致了.

设 A 为 X 的子集, 我们称 $f(A) = \{f(x) \mid x \in A\}$ 为 A 的像集 (image). 类似地, 若 $B \subseteq f(X)$, 我们就称 $f^{-1}(B) = \{x \in X \mid f(x) \in B\}$ 为 B 的原像集 (preimage), 亦称作逆像集 (inverse image). ¹⁴ 换言之, $f^{-1}(B)$ 就是 X 中那些像落入 B 的元素所构成的集合. 如果 B 中只有一个元素 b , 则 $f^{-1}(\{b\}) = \{x \in X \mid f(x) = b\}$ 就是 X 中那些被映为 b 的元素所构成的集合. 在不致混淆情况下, 我们将这一特殊情况简单记作 $f^{-1}(b)$. 注意, 集合 X 内可能有多个元素对应到 Y 中的同一元素. 映射 f 域的子集 $f(X)$ 就称作它的值域 (range). (见『图 1.1』.)

设有映射 $f : X \rightarrow Y$ 以及映射 $g : Y \rightarrow W$, 由 $h(x) = g(f(x))$ 给出的映射 $h : X \rightarrow W$ 就称作 f 和 g 的复合映射 (composition mapping), 记作 $h = g \circ f$ (见『图 1.2』)*, 这里的运算 \circ 就称作复合运算 (composition). 容易验证下述结果:

$$f \circ \text{id}_X = f = \text{id}_Y \circ f.$$

如果 $f(x_1) = f(x_2)$ 就可以推出 $x_1 = x_2$, 我们就称 f 为单射 (injection), 或者说它是单的 (injective), 也称其为一对一的 (one-to-one, 记作 1-1, 也译作一一的). 对于单射而言, 给定 Y 中的一个元素, 至多只有一个 X 中的元素与之对应. 如果 $f(X) = Y$, 那么此时就称这个映射是满射 (surjection), 或者说它是满的 (surjective), 或者映满的 (onto). ¹⁵ 如果一个映射同时是单射和满射, 那么就称其为双射 (bijection), 或者说它既单且满 (bijective), 也称其为一一对应 (one-to-one correspondence). 如果两

¹⁴译者注: 之后我们不再特意区分集合的像集以及单个点的像, 将它们统称为像. 与之对应, 也不区强调原像集的集合属性, 而直接简称其为原像.

¹⁵译者注: onto 这个词通常就直接译作满的, 不过一些学者将其译作到上、上到、映上、映到、映成. 在描述映射时, 如果我们说 “ f maps X onto Y ”, 那么就是在说 $f : X \rightarrow Y$ 这个映射是满射. 但这个短语如何翻译却没有通用说法, 有人直接说 f 将 X 映到 Y 上, 用“上”这个词强调其为满射. 在这份译作中, 我将其固定译作 “ f 将 X 满映到 Y (上)”, 直接在 “映” 这个动词前用副词 “满” 修饰.

个集合之间存在一一对应, 那么根据定义, 它们就有相同数目的元素. 如果 $f : X \rightarrow Y$ 是 X 到 Y 的双射 (bijection from X onto Y)¹⁶, 那么对每个 $y \in Y$, 都有且仅有一个 $x \in X$ 使得 $f(x) = y$. 因此, 我们就可以由 $f^{-1}(y) = x$ 定义一个映射 $f^{-1} : Y \rightarrow X$, 这里 x 就是使得 $f(x) = y$ 的那个唯一的元素. 这个映射称作 f 的逆映射 (inverse). 逆映射还有另一种定义方式: 我们将其定义为同时满足 $f \circ f^{-1} = \text{id}_Y$ 和 $f^{-1} \circ f = \text{id}_X$ 的映射 f^{-1} . 比方说, 读者可以轻易验证 $\ln^{-1} = \exp$ 且 $\exp^{-1} = \ln$, 因为 $\ln e^x = x$ 且 $e^{\ln x} = x$.

给定映射 $f : X \rightarrow Y$, 我们可以在 X 上定义一个关系 \bowtie , 具体就是规定 $x_1 \bowtie x_2$ 等价于 $f(x_1) = f(x_2)$. 读者可以验证这是个等价关系. 对应的每个等价类, 其所有元素都对应到 Y 中同一个点. 事实上, $[\![x]\!] = f^{-1}(f(x))$. 对应于 f , 就有个映射 $\tilde{f} : X/\bowtie \rightarrow Y$, 其定义为 $\tilde{f}([\![x]\!]) = f(x)$, 这个映射称作商映射 (quotient map, 或作 factor map). 这个映射必然是单的, 因为若 $\tilde{f}([\![x_1]\!]) = \tilde{f}([\![x_2]\!])$, 则 $f(x_1) = f(x_2)$, 从而 x_1 和 x_2 就属于同一等价类; 进而 $[\![x_1]\!] = [\![x_2]\!]$. 由此推得下述结论:

命题 1.2.2. 映射 $\tilde{f} : X/\bowtie \rightarrow f(X)$ 既单且满.

如果 f 和 g 都是双射, 其逆映射分别为 f^{-1} 和 g^{-1} , 那么 $g \circ f$ 同样存在逆映射, 并且可以直接验证有 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

例 1.2.3. 本例用以说明集合原像. 考察正弦函数 $\sin : \mathbb{R} \rightarrow \mathbb{R}$ 以及余弦函数 $\cos : \mathbb{R} \rightarrow \mathbb{R}$. 那么应当有下述显然的结果:

$$\sin^{-1}(0) = \{n\pi\}_{n=-\infty}^{\infty}, \quad \cos^{-1}(0) = \left\{\frac{\pi}{2} + n\pi\right\}_{n=-\infty}^{\infty}.$$

类似地, 闭区间 $[0, \frac{1}{2}] \subseteq \mathbb{R}$ 的原像 $\sin^{-1}([0, \frac{1}{2}])$ ——也就是正弦值落在 0 和 $\frac{1}{2}$ 之间的所有点——就由『图 1.3』中 x 轴上加粗的直线段所对应的那些区间构成.

例 1.2.4. 设集合 X 上定义有等价关系 \bowtie , 那么就有个天然的映射 $\pi : X \rightarrow X/\bowtie$, 它的定义为 $\pi(x) = [\![x]\!]$, 这个映射称作投影映射 (projection). 它显然是满射, 但通常不是单射, 因为若 $y \bowtie x$, 则必然有 $\pi(y) = \pi(x)$. 仅当等价关系变成恒等映射时投影映射才会是单射, 即得有 $\bowtie = \text{id}_X$.¹⁷ 此时投影映射成为双射, 并且我们记 $X \cong X/\text{id}_X$.

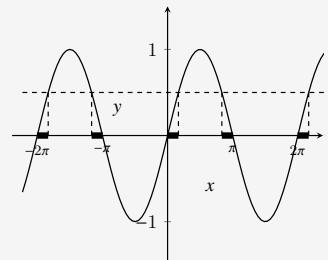


图 1.3: x 轴上由所有加粗线段标记的区间之并就是 $\sin^{-1}([0, \frac{1}{2}])$.

¹⁶译者注: 由于双射暗含满, 是故不再将 onto 明确翻译出来.

¹⁷译者注: 这里的表述可能有点让人迷惑. 根据单射的定义, 当 $\pi(y) = \pi(x)$ 时我们必须有 $y = x$. 而我们又知道此时 $y \bowtie x$, 于是 \bowtie 就必须是 $=$ 这个关系. 但是写成 $\bowtie ==$ 似乎有点让人不明所以, 是故作者采用了 id_X 来代指相等这个关系. 事实上, 在抽象理论中, 映射就是通过图象定义的, 映射 $f : A \rightarrow B$ 实际就是 Γ_f 这个集合. 若 $A = B = X$, 那么 Γ_f 本身就定义了一个关系. 容易验证, id_X 这个映射对应的关系实际就是相等关系 $=$.

例 1.2.5. 这里给出映射的更多实例.

- (a) 我们考察微积分中研究的函数 $f : \mathbb{R} \rightarrow \mathbb{R}$. 考察函数 $f : \mathbb{R} \rightarrow \mathbb{R}$ 和 $g : \mathbb{R} \rightarrow (-1, 1)$, 它们分别由 $f(x) = x^3$ 和 $g(x) = \tanh x$ 给出. 这两个函数都是双射. 顺带一提, 后面这个函数 g 表明整条实直线 (*real line*) 上的点和区间 $(-1, 1)$ 上的点一样多. 如果我们将正实数的集合记作 \mathbb{R}^+ , 将非负实数的集合记作 $\mathbb{R}_{\geq 0}$. 那么, 由 $f(x) = x^2$ 定义的函数 $f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ 就是个满射, 但不是单射 (因为 x 和 $-x$ 都映到同一点 x^2). 通过同样的对应关系 $g(x) = x^2$ 定义的函数 $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ 是单射, 但不是满射. 然而, 由同一对应关系 $h(x) = x^2$ 定义的函数 $h : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ 就是双射了. 不过, 类似定义的函数 $u : \mathbb{R} \rightarrow \mathbb{R}$ 既不是单射, 也不是满射.¹⁸
- (b) 将所有 $n \times n$ 实矩阵的集合记作 $M^{n \times n}$.¹⁹ 定义函数 $\det : M^{n \times n} \rightarrow \mathbb{R}$ 为 $\det(A) = \det A$, 即矩阵 A 的行列式. 这个函数显然是满的 (为什么?), 但不是单的. 行列式等于 1 的所有矩阵构成的集合就是 $\det^{-1}(1)$. 这样的矩阵在物理学中经常出现.
- (c) 我们感兴趣的另一个例子为 $f(z) = |z|$ 给出的函数 $f : \mathbb{C} \rightarrow \mathbb{R}$. 这个函数既不是单射, 也不是满射. 在这一情景中, $f^{-1}(1)$ 就是单位圆 (*unit circle*), 即复平面上半径为 1 的圆. 很明显, $f(\mathbb{C}) = \{0\} \cup \mathbb{R}^+$. 此外, f 诱导出 \mathbb{C} 上的一个等价关系: $z_1 \bowtie z_2$ 等价于说 z_1, z_2 落在同一个圆上. 进而, \mathbb{C}/\bowtie 就是复平面上以原点为圆心的同心圆之集合, 映射 $\tilde{f} : \mathbb{C}/\bowtie \rightarrow \{0\} \cup \mathbb{R}^+$ 为双射, 它将每个圆对应到其半径.

映射的定义域也可以是一个集合的 Descartes 积, 比如说 $f : X \times X \rightarrow Y$. 有两个具体的例子值得一提. 第一个例子就是取 $Y = \mathbb{R}$. 这一情形的例子之一就是矢量点乘. 因此, 如果 X 为 (三维) 空间中矢量的集合, 我们就可以定义 $f(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$. 第二种情形是 $Y = X$. 此时 f 就称作 X 上的二元运算 (binary operation), 此时 X 中的一个元素与 X 中的两个元素相关联. 比如说, 取 $X = \mathbb{Z}$, 即整数集; 那么, 由 $f(m, n) = mn$ 定义的函数 $f : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}$ 就是整数相乘这个二元运算. 类似地, 由 $g(x, y) = x + y$ 给出的函数 $g : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ 就是实数相加这个二元运算.

¹⁸译者注: 原书中 f 的定义为 $f : \mathbb{R} \rightarrow \mathbb{R}^+$. 但是, 根据书中定义, \mathbb{R}^+ 不包含 0, 从而 $f(0) = 0 \notin \mathbb{R}^+$, 即 $f(0)$ 没有定义! 因此, 在翻译时我引入了 $\mathbb{R}_{\geq 0}$ 这个记号.

¹⁹译者注: 这个对象的记号有很多, 常见的是 $\mathbb{R}^{n \times n}$ 和 $M_n(\mathbb{R})$, 它们都很好地强调了对应矩阵元所在数域. 本书选用的这套记号反倒比较少见, 不过具体语境下也能确定到底是实数还是复数矩阵元.

1.3. 度量空间

虽说集合是现代数学的根基, 但它们本身却只具有形式上和抽象上的意义. 要让集合变得有用起来, 我们就需要在它们上面附加一些额外的结构. 实现这些结构有两种一般流程. 这两种流程源自对数学两大分支(即代数学和分析学)的抽象化.

要在集合上附带代数结构(algebraic structure), 我们就需要在它上面引入二元运算(binary operation). 比方说, 矢量空间就由矢量加法这个二元运算以及其他附带内容构成. 群就由“群乘法”以及其他附带内容构成. 这样的代数系统还有很多, 它们构成了代数学这个丰富的数学分支.

而在数学的另一分支(也就是分析学)中使用集合这一概念进行抽象时, 就将我们引向了拓扑学(topology). 在这一主题中, 连续性这个概念起着中心作用. 此外, 拓扑学也是一门内容丰富、意义深远、应用广泛的学科. 我们不准备深入这两个数学领域. 虽说如此, 在后面的章节中我们还是会讨论一些代数系统, 并且会用到极限和连续性这些思想, 不过是以一种直观的方式实现的, 并在使用到的时候才会引入并介绍相关概念. 但另一方面, 还是有一些覆盖主体内容所需的最小预备知识. 其中之一就是度量空间.

定义 1.3.1 (度量空间). 所谓度量空间(*metric space*), 就是一个集合 X 连带着它上面的一个实值函数 $d : X \times X \rightarrow \mathbb{R}$, 并且这个函数需要满足下述条件:

- (a) $d(x, y) \geq 0$ 对所有 x, y 成立, 并且 $d(x, y) = 0$ 的充要条件是 $x = y$.
- (b) $d(x, y) = d(y, x)$ (对称性, *symmetry*).
- (c) $d(x, y) \leq d(x, z) + d(z, y)$ (三角不等式, *triangle inequality*).

值得指出的是, 这里的集合 X 完全是任意的, 它不需额外的结构. 在这方面, 『[定义 1.3.1](#)』相当宽泛, 并包含了很多不同的情形, 我们接下来的几个例子就会说明这一点. 不过, 在具体进入这些例子之前, 我们还是要注意一下, 上述定义中的函数 d 其实是距离(distance)这一概念的抽象: 条件(a)是在说任意两点之间的距离总是非负的, 并且仅当这两点重合时其距离为零; 条件(b)是在说交换两点顺序之后它们的距离不会发生改变; 条件(c)其实是我们所熟知的“三角形任意两边长度之和大于第三边的长度, 仅当三角形退化为直线时取等”.

集合中任意两点距离是个正实数, 这一事实是 **Euclid 度量空间** (Euclidean metric space) 的属性. 而在相对论中, 我们就必须面对另一种度量空间, 所谓的 **Minkowsky 度量空间** (Minkowskian metric space), 此时距离(的平方)有取负值的可能.

例 1.3.2. 此例用以给出度量空间的些许实例.

- (a) 取 $X = \mathbb{Q}$, 即有理数集, 并定义 $d(x, y) = |x - y|$.
- (b) 取 $X = \mathbb{R}$, 同样定义 $d(x, y) = |x - y|$.
- (c) 取 X 为球面上的所有点. 我们可以定义 X 上的两个距离函数. 距离函数 $d_1(P, Q)$ 定义为球面上连接 P, Q 所得弦 (*chord*) 的长度. 我们还可以定义另外一个度量 $d_2(P, Q)$, 它定义为球面上过 P, Q 的大圆 (*great circle*) 上 PQ 段的弧长. 要说服自己认为这两个函数确实满足度量函数的所有性质并非难事. 另外注意, 对于 d_2 而言, 如果三个点中有两个点是球面的极点 (*poles*), 那么三角不等式此时就变成一个等式.
- (d) 用 $\mathcal{C}^0[a, b]$ 表示闭区间 $[a, b]$ 上所有实值连续函数的集合. 对 $f, g \in \mathcal{C}^0[a, b]$, 我们可以定义 $d(f, g) = \int_a^b |f(x) - g(x)| dx$.
- (e) 用 $\mathcal{C}_B(a, b)$ 表示闭区间 $[a, b]$ 上诸有界 (*bounded*) 实值连续函数的集合. 对 $f, g \in \mathcal{C}_B(a, b)$ 我们定义

$$d(f, g) = \max_{x \in [a, b]} \{|f(x) - g(x)|\}.$$

这个记号的含义是说, 首先计算函数 f, g 在区间 $[a, b]$ 中所有点 x 处函数值之差的绝对值, 然后取这些绝对值中的最大值.

度量函数为我们提供了一种自然的工具来测试度量空间中一些点的“靠近度” (*closeness*). 靠近度这个概念在研究序列时会变得至关重要. 所谓序列 (*sequence*), 就是从正整数集 \mathbb{N}^* 映入度量空间 X 的一个映射 $s : \mathbb{N}^* \rightarrow X$. 这样的一个映射就将一个正整数 n 和度量空间中的一个点 $s(n)$ 关联起来. 习惯上, 我们会将 $s(n)$ 写作 s_n (或者写作 x_n 用以匹配 X 这个符号), 并用 $\{x_n\}_{n=1}^\infty$ 来枚举其函数值.²⁰

了解在 n 充分大时序列的行为相当重要. 特别地, 当 n 增加时, 判断序列是否会趋近于某个确定值很重要.²¹

定义 1.3.3 (收敛). 假设对某个 x 以及任意正实数 ε , 存在自然数 N 使得 $d(x_n, x) < \varepsilon$ 在 $n > N$ 时均成立. 那么, 我们就称序列 $\{x_n\}_{n=1}^\infty$ 收敛到 (*converge to*) 其极限点 x , 并记 $\lim_{n \rightarrow \infty} d(x_n, x) = 0$ 或者 $d(x_n, x) \rightarrow 0$, 亦或者简单记作 $x_n \rightarrow x$.

要直接检验一个给定序列是否收敛有时候是不可能的, 因为这需要知道具体的极限点 x 才可以进行. 不过, 我们却可以去干下面这件最好的事情——去判断这个序列中的点是否会随着 n 越来越大而越来越近.

²⁰译者注: 原书中序列定义为映射 $s : \mathbb{N} \rightarrow X$, 但是 $0 \in \mathbb{N}$, 因此最终的序列应当写作 $\{x_n\}_{n=0}^\infty$, 但作者又习惯上把序列写成 $\{x_n\}_{n=1}^\infty$. 是故这里我将所有的自然数都改成了正整数.

²¹译者注: 原文用到的是 *finite value*, 即有限值. 但考虑到只有度量空间为数集时有限这个说法才合适, 是故此处我改用确定值这一说法.

定义 1.3.4 (Cauchy 列). 所谓 **Cauchy 列 (Cauchy sequence)**, 就是满足

$$\lim_{m,n \rightarrow \infty} d(x_m, x_n) = 0$$

的序列.

『图 1.4』直观地展示了 Cauchy 列.

不同于收敛, 我们可以直接判断一个序列是否 Cauchy. 然而, 即便一个序列是 Cauchy 列, 它也未必是收敛的. 比方说, 取度量空间为有理数 \mathbb{Q} 连带度量函数 $d(x, y) = |x - y|$. 现在考察 $x_n = \sum_{k=1}^n (-1)^{k+1}/k$ 给出的序列 $\{x_n\}_{n=1}^\infty$. ²²很明显, 对任意的 n , 这里的 x_n 都是个有理数. 而『习题 1.7』则说明了如何去证明 $|x_m - x_n| \rightarrow 0$. 因此, 这个序列是个 Cauchy 列. 然而, 读者可能已经知道 $\lim_{n \rightarrow \infty} x_n = \ln 2$, 这并不是有理数.

定义 1.3.5 (完备度量空间). 如果一度量空间中所有的 Cauchy 列都收敛, 则称其为完备度量空间 (*complete metric space*).

在现代分析学中, 完备度量空间扮演关键角色. 前面的那个例子说明有理数集 \mathbb{Q} 不是完备度量空间. 然而, 如果我们将所有 Cauchy 列的极限点都添加到 \mathbb{Q} 中, 所得到的这个空间就完备了. 当然, 这个完备空间就是实数系 \mathbb{R} . 可以证明, 任意不完备的度量空间都可以“扩充成”一个完备度量空间.²³

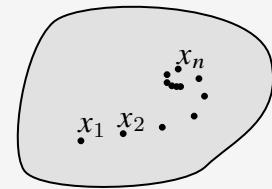


图 1.4: Cauchy 列中元素之间的距离会越来越小.

²²译者注: 学过微积分的读者会发现, 这实际上就是自然对数 $\ln(1+x)$ 的幂级数展开, 当然此处 x 已经赋值为 1 了.

²³译者注: 这一操作称作完备化 (completion), 并且可以证明任意度量空间完备化后的空间是唯一的. 请注意, 此处的唯一是指同构意义下唯一. 至于什么叫同构, 可以翻阅群论相关章节. 直观上说, 所谓同构, 就是单纯从结构的角度我们无法区分两个对象. 严格地说, 就是两个对象之间存在保结构的双射.

1.4. 基数

计数这一过程就是将一个集合和另一个集合中的元素逐一进行匹配 (one-to-one comparison). 如果两个集合一一对应, 则称它们有相同的基数 (cardinality, 也译作势). 具有相同基数的集合就有相同“数目”的元素. 集合 $F_n = \{1, 2, \dots, n\}$ 元素数目有限, 它的基数为 n . 任意一个与 F_n 之间存在双射的集合就称作 n 元有限集 (finite set with n elements).



Georg Cantor
(格奥尔格·康托尔)
1845-1918

人物志 (Georg Cantor)

虽说在 Cantor (1845-1918) 之前, 就有了一些尝试, 用以建立一套确定的集合理论, 但公认的集合论创建者就是他. Cantor 出生于沙俄的一家丹麦裔犹太人家庭, 但之后随父母移居德国. 他的父亲敦促他学习工程学, 而他也在 1863 年怀揣这一目的进入柏林大学. 在大学中, 他受到 Weierstrass 的影响, 转而踏入纯数学的学习研究. 1869 年, 他在哈雷大学获得无薪讲师资格,²⁴ 并于 1879 年成为教授. 在 29 岁那年, Cantor 在期刊 *Journal für Mathematik* 上发表了他的第一篇革命性论文, 这篇论文是关于无限集的. 虽说一些老派数学家认为文中的一些命题是错的, 但文中整体体现出的创意以及才华获得了人们的关注. 一直到 1897 年, 他都不断在发表关于集合论以及超限数的文章.²⁵ Cantor 的主要关注点之一就是通过“大小 (size)”来区分无限集. 和之前的 Bolzano 一样, Cantor 同样将一一对应确立为基本原则. 在他 1873 年与 Dedekind 的通信中, Cantor 提出了这样一个问题: 实数是否和整数一一对应? 而几周过后, 他给出了否定答案, 并给出了两个证明. 第一个证明远复杂于第二个, 而后者是我们今天通常采用的证明方式. 在 1874 年, Cantor 致力于研究直线上的点和 \mathbb{R}^n 中的点之间的等价性, 并试图证明这两个集合之间不存在一一对应. 不过, 就在三年过后, 他证明了这样的对应反而是存在的. 他在给 Dedekind 的信件中说道, “我看到了, 但我不信 (I see it but I do not believe it)”. 之后他证明了, 对任意给定集合, 都可以构造出一个新的集合, 即给定集合所有子集构成的集合,²⁶ 这个集合的基数一定比原本的集合来得大. 将自然数集 \mathbb{N} 的基数记作 \aleph_0 , 而将对应于前述子集集合的基数记作 2^{\aleph_0} .²⁷ Cantor 证明了 $2^{\aleph_0} = c$, 其中 c 是连续统 (也就是实数集) 的基数. Cantor 的研究解决了一些由来已久的问题, 但也颠覆了许多过去的观念, 因而很难期望其很快得到接受. 他关于超限序数 (transfinite ordinal number) 以及超限基数 (transfinite cardinal number) 的想法引起了大数学家 Leopold Kronecker 的敌意, 他在十多年的时间里野蛮攻讦 Cantor 的理论, 并阻止 Cantor 在柏林大学获得更重要的任命. 虽说

²⁴译者注: 无薪讲师 (privatdozent) 是德语国家的一个特殊职业, 其职业内容和大学讲师相同, 负责向大学生授课, 但是大学不向其支付薪水, 其收入完全由学生的听课费构成.

²⁵译者注: Cantor 把基数大于任意 n 的基数称作超限数 (transfinite number).

²⁶译者注: 回忆一下, 这就是我们前面定义过的幂集.

²⁷译者注: 此处的 \aleph 是希伯来字母, 读作阿列夫, 于是 \aleph_0 就读作阿列夫零.

Kronecker 于 1891 年逝世, 但他的攻击还是使得数学家们对 Cantor 的研究表示怀疑. Poincaré 就将集合论称作一种有趣的 “病态情形 (pathological case)”. 同时预言说 “后人们会将 Cantor 的集合论视作是一种已痊愈的疾病 (Later generations will regard Mengenlehre as a disease from which one has recovered)”. Cantor 曾一度精神崩溃, 但他还是在 1887 年恢复工作.

然而另一方面, 很多杰出的数学家则因这个新理论在分析学、测度论、拓扑学中的使用而印象深刻. Hilbert 在德国传播了 Cantor 的思想. 在 1926 年, 他说 “没有人可以将我们从 Cantor 为我们创造的天堂中驱逐出去 (No one shall expel us from the paradise which Cantor created for us)”. 他将 Cantor 的超限算术 (transfinite arithmetic) 称赞为 “数学思想最令人惊异的产物, 是人类活动在纯可理解领域中最漂亮的表现之一 (the most astonishing product of mathematical thought, one of the most beautiful realization of human activity in the domain of the purely intelligible)”. Bertrand Russel 则将 Cantor 的理论描述为 “这可能是这个时代最值得夸耀的 (probably the greatest of which the age can boast)”. Cantor 的理论在主要由 Hilbert 引领的 “数学形式化运动” 得到广泛应用, 但这一后续应用似乎与 Cantor 的 Plato 观点不一致: Cantor 认为其工作的重要性在于对形而上学和神学的影响. 作品可以如此无缝偏离其创作者的用意, 这是对其客观性以及精湛工艺最强有力的证明.

现在考虑正整数集 $\mathbb{N}^* = \{1, 2, 3, \dots\}$. 如果集合 A 和 \mathbb{N}^* 之间存在一个双射, 则称 A 是可数无穷的 (countably infinite). ²⁸ 可数无穷集的例子有整数集、正偶数集、正奇数集、素数集等, 氢原子束缚态能级的集合也是个可数无穷集.

一个集合的子集 (比如说所有正偶数的集合) 可能会和这个集合 (对应前文就是正整数集) 之间存在一一对应, 这似乎让人惊讶; 然而, 这是所有无穷集共有的性质. 事实上, 有时候无穷集就定义为那些和其至少一个真子集之间存在一一对应的集合. 同样令人惊讶的是, 有理数和自然数一样多. 要知道, 仅仅在区间 $(0, 1)$, 或者任意两个不同的实数之间, 就有无穷多个有理数! *

那些既不是有限集, 也不是可数无穷集的集合就是不可数的 (uncountable). 在某种意义上, 不可数集比任意可数集 “更无穷”. 不可数集的例子有区间 $(-1, 1)$ 中的点、实数、平面中的点、空间中的点. 可以证明, 这些集合都有相同的基数. 换言之, 三维空间 (也就是全宇宙) 和区间

²⁸译者注: 原书中写的是自然数集 $\mathbb{N} = \{1, 2, 3, \dots\}$. 但这个写法又和第一次引入这个符号的说法矛盾了. 这本书中经常出现自然数有时候包含零, 有时候不包含零的情况. 一个原因就是不同文献有着不同约定, 作者在写作时可能受不同参考文献的影响经常犯糊涂.

* 为证明有理数和自然数一样多, 我们可以将有理数 m/n 写成一个 $\infty \times \infty$ 矩阵中处于 (m, n) 位置的矩阵元, 然后从 $(1, 1)$ 矩阵元开始 “计数”, 然后向右数 $(1, 2)$, 接下来沿对角数 $(2, 1)$, 接着向下数 $(3, 1)$, 再沿对角向上数, 以此类推. 很明显, 这个集合是可数的, 并且穷尽了所有有理数. 事实上, 这个过程会重复计入某些有理数.

$(-1, 1)$, 或者任意其他有限区间, 有一样多的点.



图 1.5: 前四次“切除 (dissection)”后所得 Cantor 集.

基数是个非常复杂的数学概念, 它有很多令人惊奇的结果. 考虑区间 $[0, 1]$. 从它中间移去开区间 $(\frac{1}{3}, \frac{2}{3})$ (只保留端点 $\frac{1}{3}$ 和 $\frac{2}{3}$). 接下来从剩下的部分 $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ 中再去掉两段各自中间的 $\frac{1}{3}$, 这样剩下来的部分就是

$$\left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

这样无限进行下去(见『图 1.5』). 这样剩下的那个集合就称作 **Cantor 集** (Cantor set), 这个集合的基数是什么呢? 直觉上, 我们会认为几乎不剩下什么东西. 我们可能会说服自己接受这样的结果: 剩下的这个集合中点的数目至多是可数无穷的. 但令人惊讶的是, 它的基数和连续统相同. 因此, 在去除无穷多个中间的三分之一后, 剩下的集合和最开始的集合有一样多的点.

1.5. 数学归纳法

很多时候人们都希望做出一个对所有自然数都成立的数学论断。比如说，我们可能希望建立一个涉及到整数参数的公式，这个公式对所有正整数都成立。人们会在这样的情形遇到这类问题：在对前几个正整数做试验后，人们识别出一个模式，并发现了某个公式，然后希望这个公式对所有自然数都成立。为此，我们就需要用到数学归纳法 (mathematical induction)，它的内容如下所述：

命题 1.5.1. 假设对每个自然数 (正整数) n 都有个陈述 S_n 。那么，只要下面两个条件同时满足，则 S_n 对每个正整数都是成立的：

- (a) S_1 成立。
- (b) 若对某个给定正整数 m 有 S_m 成立，则 S_{m+1} 也成立。

为说明如何使用数学归纳法，我们以二项式定理 (binomial theorem) 的证明为例。

例 1.5.2. 证明二项式定理，即对任意正整数 m ，我们有

$$\begin{aligned} (a+b)^m &= \sum_{k=0}^m \binom{m}{k} a^{m-k} b^k = \sum_{k=0}^m \frac{m!}{k!(m-k)!} a^{m-k} b^k \\ &= a^m + ma^{m-1}b + \frac{m(m-1)}{2!} a^{m-2}b^2 + \cdots + mab^{m-1} + b^m, \end{aligned} \quad (1.1)$$

其中

$$\binom{m}{k} := \frac{m!}{k!(m-k)!}. \quad (1.2)$$

证。对于此例，对应的数学陈述 S_m 就是『(1.1) 式』。我们注意到 S_1 显然成立： $(a+b)^1 = a^1 + b^1$ 。现在假定 S_m 成立，我们需要证明此时 S_{m+1} 同样成立。换言之，我们要从『(1.1) 式』出发去证明

$$(a+b)^{m+1} = \sum_{k=0}^{m+1} \binom{m+1}{k} a^{m+1-k} b^k.$$

如此一来，归纳原理就保证了我们的目标论断 (方程) 对所有正整数都成立。给『(1.1) 式』两边同时乘以 $a+b$ ，如此即得

$$(a+b)^{m+1} = \sum_{k=0}^m \binom{m}{k} a^{m-k+1} b^k + \sum_{k=0}^m \binom{m}{k} a^{m-k} b^{k+1}.$$

我们将第一个和式中 $k=0$ 这一项以及第二个和式中 $k=m$ 这一项单独

分离出来:

$$\begin{aligned}
 (a+b)^{m+1} &= a^{m+1} + \sum_{k=1}^m \binom{m}{k} a^{m-k+1} b^k + \underbrace{\sum_{k=0}^{m-1} \binom{m}{k} a^{m-k} b^{k+1} + b^{m+1}}_{k=j-1} \\
 &= a^{m+1} + \sum_{k=1}^m \binom{m}{k} a^{m-k+1} b^k + \sum_{j=1}^m \binom{m}{j-1} a^{m-j+1} b^j + b^{m+1}.
 \end{aligned}$$

最后一行第二个求和中涉及到指标 j . 由于这是个哑指标 (dummy index), 我们可以将其替换为自己喜欢的任意符号. 选用 k 取替换 j 则尤为有用, 因为这样我们可以将两个求和的形式统一起来. 如此一来就有

$$(a+b)^{m+1} = a^{m+1} + \sum_{k=1}^m \left\{ \binom{m}{k} + \binom{m}{k-1} \right\} a^{m-k+1} b^k + b^{m+1}.$$

读者很容易就可以验证下述结果²⁹:

$$\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1}.$$

使用上述结果, 最后即可得到

$$\begin{aligned}
 (a+b)^{m+1} &= a^{m+1} + \sum_{k=1}^m \binom{m+1}{k} a^{m-k+1} b^k + b^{m+1} \\
 &= \sum_{k=0}^{m+1} \binom{m+1}{k} a^{m-k+1} b^k.
 \end{aligned}$$

这就完成了证明. □

数学归纳法也用于定义涉及到整数的量. 这样的定义方法通常称作归纳定义 (inductive definition). 在定义正整数次幂时我们就用到了这种方法: $a^1 = a$, 然后定义 $a^m = a^{m-1}a$.

1.6. 本章习题

习题 1.1. 求证, n 元集合所有子集的数目为 2^n .

习题 1.2. 设 A, B, C 为全集 U 中的集合. 求证:

²⁹译者注: 请读者回忆高中知识. 二项式系数 $\binom{n}{k}$ 就是从 n 个元素中选出 k 个元素的组合数. 如此一来, $\binom{m+1}{k}$ 就可以解释为从 $\{0, 1, \dots, m\}$ 这个集合中挑选出 k 个元素的方法数. 选出 k 个元素有两种互斥的选择: (1) 这 k 个元素中包含元素 0; (2) 这 k 个元素中不包含元素 0. 前者意味着我们要从剩下的 $\{1, 2, \dots, m\}$ 中选出 $k-1$ 个元素, 它的方法数为 $\binom{m}{k-1}$. 后者意味着我们要从 $\{1, 2, \dots, m\}$ 中完全选出 k 个元素, 它的方法数为 $\binom{m}{k}$. 由于两种情况互斥, 根据分类加法计数原理, 最后的总方法数就是两种方法数之和. 这样我们就从组合学的角度证明了这里的这个恒等式. 这种证明方法也是高中必须学会的基础知识之一.

- (a) 若 $A \subseteq B$ 且 $B \subseteq C$, 则 $A \subseteq C$.
- (b) $A \subseteq B$ 等价于 $A \cap B = A$, 也等价于 $A \cup B = B$.
- (c) $A \subseteq B$ 且 $B \subseteq C$ 时就有 $(A \cup B) \subseteq C$.
- (d) $A \cup B = (A \sim B) \cup (A \cap B) \cup (B \sim A)$.

提示: 要证明两个集合相等, 只需证明彼此互为对方的子集.

习题 1.3. 对每个 $n \in \mathbb{N}$, 令

$$I_n = \left\{ x \left| |x - 1| < n \text{ 且 } |x + 1| > \frac{1}{n} \right. \right\}.$$

求 $\bigcup_n I_n$ 以及 $\bigcap_n I_n$.

习题 1.4. 证明 $a' \in [\![a]\!]$ 时就有 $[\![a']\!] = [\![a]\!]$.

习题 1.5. 你可以在空间矢量的集合中定义“乘法”这个二元运算吗? 平面上的矢量呢? 对每一情形, 将乘积的分量用参与运算的两个矢量的分量表示出来.

习题 1.6. 若 f, g 均为双射, 求证 $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

习题 1.7. 此题目目标是证明由 $x_n = \sum_{k=1}^n (-1)^{k+1}/k$ 定义的序列 $\{x_n\}_{n=1}^\infty$ 为 *Cauchy* 列. 不失一般性, 可以假设 $n > m$, 并且 $n - m$ 为偶数 ($n - m$ 为奇数时也可以类似操作).

(a) 证明:

$$x_n - x_m = -(-1)^m \sum_{j=1}^{n-m} \frac{(-1)^j}{j+m}.$$

(b) 将 (a) 中所得求和按照奇数项和偶数项分离. 证明

$$x_n - x_m = -(-1)^m \left\{ - \sum_{k=1}^{(n-m)/2} \frac{1}{2k-1+m} + \sum_{k=1}^{(n-m)/2} \frac{1}{2k+m} \right\}.$$

(c) 将 (b) 最后的两个求和加起来, 使其变成单个求和. 证明

$$x_n - x_m = -(-1)^m \left\{ \sum_{k=1}^{(n-m)/2} \frac{1}{(2k+m)(2k+m-1)} \right\},$$

进而证明

$$\begin{aligned} |x_n - x_m| &\leq \sum_{k=1}^{(n-m)/2} \frac{1}{(2k+m-1)^2} \\ &= \frac{1}{(1+m)^2} + \sum_{k=2}^{(n-m)/2} \frac{1}{(2k+m-1)^2}. \end{aligned}$$

(d) 让自己相信对任意连续的递减函数³⁰ $f(x)$ 有

$$\int_1^s f(x) dx \geq \sum_{k=2}^s f(k).$$

利用该关系, 将 (c) 的结果进一步写成³¹

$$\begin{aligned} |x_n - x_m| &\leq \frac{1}{(1+m)^2} + \int_1^{(n-m)/2} \frac{1}{(2x+m-1)^2} dx \\ &= \frac{1}{(1+m)^2} - \frac{1}{2} \left(\frac{1}{n-1} - \frac{1}{m+1} \right). \end{aligned}$$

最后一行中的每一项都会在 m, n 趋于无穷时独立地趋近于零.

习题 1.8. 构造双射 $f : \mathbb{N} \rightarrow \mathbb{Z}$.

提示: 想办法让 f 将偶整数映满到非负整数上, 并将奇整数映满到负整数上.

习题 1.9. 任取两个开区间 (a, b) 和 (c, d) , 证明无论区间的间隔有多大, 这两个区间中都有相同数目的点.

提示: 在两个区间之间构造一个(线性的)代数关系.

习题 1.10. 利用数学归纳法导出函数乘积导数的 *Leibniz* 法则:

$$\frac{d^n}{dx^n}(f \cdot g) = \sum_{k=0}^n \binom{n}{k} \frac{d^k f}{dx^k} \frac{d^{n-k} g}{dx^{n-k}}.$$

习题 1.11. 利用数学归纳法导出下述结果:

$$\sum_{k=0}^n r^k = \frac{r^{n+1} - 1}{r - 1}, \quad \sum_{k=0}^n k = \frac{n(n+1)}{2}.$$

³⁰译者注: 原书说是连续函数, 但是可以找到反例: 设 $f(x) = x$, 则 $f(2) + f(3) = 5$, 而 $\int_1^3 x dx = (9 - 1)/2 = 4$. 也就是说 $\int_1^3 f(x) dx < \sum_{k=2}^3 f(x)$. 不过若是递减的函数就没问题了, 我们可以通过积分中值定理得到该结果: 根据中值定理可知存在 $x_n \in [n-1, n]$ 使得 $\int_{n-1}^n f(x) dx = f(x_n)[n - (n-1)] = f(x_n)$. 递减保证了 $f(x_n) \geq f(n)$. 于是 $\int_{n-1}^n f(x) dx \leq f(n)$, 两边对 n 求和即可得到该不等式.

³¹译者注: 此处原本最后一行有个 $1/n$, 但经校验当为笔误.



第1部分
有限维矢量空间

2

矢量与线性映射

我们在高中接触并学习过的平面矢量 (planar vector) 和空间矢量 (spatial vector) 可以很容易推广到高维情形。通过将矢量用其分量表示，我们可以想象矢量具有 N 个分量。这种方式是将平面矢量以及空间矢量推广的最快方式，这样得到的矢量就称作 N 维 **Descartes 矢量** (Cartesian vector)。Descartes 矢量受到两方面的限制：其分量是实的，并且其维数有限。在物理学中，我们有时候需要移除这两个限制中的一个或者全部。因此，我们就有必要研究去除分量为实这一要求或者维数有限这一要求的矢量。而这些属性均可视作是更一般定义的推论。尽管在这一部分我们将注意力集中在有限维的矢量空间，但是这里涉及到的很多概念以及例子同样也适用于无穷维的空间。

2.1. 矢量空间

我们从对抽象的(复)矢量空间下定义开始。^{*}

定义 2.1.1 (矢量空间). 复数集 \mathbb{C} 上的矢量空间 (*vector space*) \mathcal{V} 是一些记作形如 $|a\rangle$, $|b\rangle$, $|x\rangle$ 的对象的集合，这些对象称作矢量 (*vector*, 现代数学文献中通常译作向量)，并且需要满足下述性质：

(1) 对 \mathcal{V} 中每对矢量 $|a\rangle$ 和 $|b\rangle$ ，都对应了同样在 \mathcal{V} 中的一个矢量 $|a\rangle + |b\rangle$ ，称作这两个矢量的和 (*sum*)，并且这个和要满足下述要求：

- (a) $|a\rangle + |b\rangle = |b\rangle + |a\rangle$,
- (b) $|a\rangle + (|b\rangle + |c\rangle) = (|a\rangle + |b\rangle) + |c\rangle$,
- (c) 存在唯一的矢量 $|0\rangle \in \mathcal{V}$ ，使得对每个矢量 $|a\rangle$ 均有 $|a\rangle + |0\rangle = |a\rangle$ ，该矢量称作零矢量 (*zero vector*, 或者 *null vector*)，
- (d) 对每个矢量 $|a\rangle \in \mathcal{V}$ ，均存在唯一的矢量 $-|a\rangle \in \mathcal{V}$ 使得 $|a\rangle + (-|a\rangle) = |0\rangle$ 。

(2) 对每个复数[†] α 以及每个矢量 $|a\rangle$ ，都对应了 \mathcal{V} 中的一个矢量 $\alpha|a\rangle$ ，称作这一组合的标量乘法 (*scalar multiplication*, 数学文献中也译作纯量乘法，或者数乘)，这里的 α 就称作标量 (*scalar*, 数学文献中也译作纯量)。标量乘法要满足下述要求：

^{*} 始终牢记， \mathbb{C} 为复数集，而 \mathbb{R} 为实数集。

[†] 复数通常记作 z (特别是将其视作变量时)，我们会在『第??部分』坚持这一约定。不过，在矢量空间的讨论中，我们会看到使用小写希腊字母表示复标量会更方便。

$$(a) \alpha(\beta|a\rangle) = (\alpha\beta)|a\rangle,$$

$$(b) 1|a\rangle = |a\rangle.$$

(3) 涉及到矢量和标量的这个乘法满足分配律 (*distributive law*), 即

$$(a) \alpha(|a\rangle + |b\rangle) = \alpha|a\rangle + \alpha|b\rangle.$$

$$(b) (\alpha + \beta)|a\rangle = \alpha|a\rangle + \beta|a\rangle.$$

用 $\langle |$ 和 $| \rangle$ 表示矢量的这一记法, 是由 Dirac 发明的, 根据其箭头开口分别命名为左矢 (bra) 和右矢 (ket), 当我们处理复矢量空间时这一记法会非常有用.¹ 虽说如此, 但在涉及到范数和度规时, 这一记号却稍显笨拙, 因此在讨论这两个主题时我们会暂时舍弃这一记法.

上文定义的这个矢量空间通常也称作复矢量空间 (complex vector space). 如果将定义中的复数集 \mathbb{C} 替换为实数集 \mathbb{R} , 所得的这个空间就称作实矢量空间 (real vector space).

实数和复数是域 (field) 这一数学结构的原型. 所谓域, 就是一些对象的集合 \mathbb{F} , 在这些对象之间定义了两个二元运算: 加法 (addition) 和乘法 (multiplication). 乘法对加法满足分配律, 并且每个运算都有对应的单位元. 加法的单位元记作 0, 并称作加法单位元 (additive identity). 乘法的单位元记作 1, 并称作乘法单位元 (multiplicative identity). 不仅如此, 对每个元素 $\alpha \in \mathbb{F}$, 都有个加法逆元 $-\alpha$; 而对每个非加法单位元以外的元素, 都有个乘法逆元 α^{-1} .²

例 2.1.2 (一些矢量空间).

(1) \mathbb{R} 是实数域上的矢量空间.

(2) \mathbb{C} 是实数域上的矢量空间.

(3) \mathbb{C} 是复数域上的矢量空间.

(4) 令 $\mathcal{V} = \mathbb{R}$, 并取标量域为 \mathbb{C} . 这就不是个矢量空间, 因为『[定义 2.1.1](#)』中的性质 (2) 不成立: 复数乘以实数所得结果未必还是实数, 从而可能不属于 \mathcal{V} .

(5) 平面 (或者空间) 中“箭头” (arrows) 的集合, 在平面矢量 (或者空间矢量) 加法的平行四边形法则下构成 \mathbb{R} 上的矢量空间.

¹译者注: 这里的 bra 和 ket 实际上是 Dirac 将括号 (bracket) 这个单词拆分成两部分给出的, 我们可以将这个单词理解为某个常量 c 被 bra 和 ket 包围, 这就给出一个括号 (c). 值得一提的是, 在对 bra 和 ket 的翻译中, 还有一个译法曾流传一时: 将 bra 译作刁矢, 将 ket 译作刃矢, 实际上就是取横折钩这个笔画内下一笔到底从左往右 (刁) 还是从右往左 (刃) 来区分箭头的左右朝向. 这个想法虽然和 Dirac 命名时的本意有异曲同工之妙, 但是因为此处刁和刃和汉字本意无关, 虽有匠心, 却失意涵, 后来被左矢和右矢这一直观译法广泛取代.

²译者注: 此处的加法和乘法除了文中提到的单位元和逆元性质以外, 还必须满足结合律和交换律. 最后, 在矢量空间的定义中, \mathbb{C} 可以替换为任何一个域 \mathbb{F} , 所得的结果就称作这个域 \mathbb{F} 上的矢量空间 (vector space over \mathbb{F}).

- (6) 设 $\mathcal{P}^c[t]$ 为变量 t 的复系数多项式构成的空间. 在多项式的加法以及与复数和多项式的乘法下, $\mathcal{P}^c[t]$ 就是个矢量空间.³ 在这一例子中, 零矢量就是零多项式.
- (7) 对于给定的正整数 n , 取 $\mathcal{P}_n^c[t]$ 为次数小于等于 n 的复系数多项式构成的空间. 同样容易验证在多项式加法以及复数和多项式的乘法下, 这构成一个矢量空间. 特别地, 两个次数小于等于 n 的多项式之和, 同样是一个次数小于或者等于 n 的多项式. 此外, 复系数多项式与复数相乘给出的是另一个同类型多项式. 此处零多项式照样就是零矢量.
- (8) 所有次数小于等于 n 的实系数多项式构成的集合 $\mathcal{P}_n^r[t]$ 是实数域上的矢量空间, 但是不是复数域上的矢量空间.
- (9) 设 \mathbb{C}^n 为所有形如 $|a\rangle = (\alpha_1, \dots, \alpha_n)$ 、 $|b\rangle = (\beta_1, \dots, \beta_n)$ 这样的复 n 元组构成. 设 η 为一复数. 现在我们定义

$$\begin{aligned}|a\rangle + |b\rangle &= (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_n + \beta_n), \\ \eta|a\rangle &= (\eta\alpha_1, \eta\alpha_2, \dots, \eta\alpha_n), \\ |0\rangle &= (0, 0, \dots, 0), \\ -|a\rangle &= (-\alpha_1, -\alpha_2, \dots, -\alpha_n).\end{aligned}$$

容易验证 \mathbb{C}^n 是复数域上矢量空间. 我们称其为 n 维复坐标空间 (*complex coordinate space*).

- (10) 和 \mathbb{C}^n 类似的运算下, 所有实 n 元组的集合 \mathbb{R}^n 就是实数域上的矢量空间, 称其为 n 维实坐标空间 (*real coordinate space*), 亦或者 n 维 Descartes 空间 (*Cartesian n-space*). 这不是复数域上的矢量空间.
- (11) 所有 m 行 n 列复矩阵的集合 $\mathcal{M}^{m \times n}$ 在常规的矩阵加法以及与复数的标量乘法下构成矢量空间. 与之对应的零矢量就是所有矩阵元都为 0 的 $m \times n$ 矩阵, 即零矩阵.
- (12) 设 \mathbb{C}^∞ 为所有使得 $\sum_{i=1}^{\infty} |\alpha_i|^2 < \infty$ 的复序列 $|a\rangle = \{\alpha_i\}_{i=1}^{\infty}$ 构成的集合. 可以证明, 在逐分量定义的加法和标量乘法下, \mathbb{C}^∞ 是复数域上的矢量空间.
- (13) 实区间 (a, b) 上, 所有连续单(实)变量复值函数的集合, 是复数域上的矢量空间.
- (14) 区间 (a, b) 上, 所有 n 阶连续可导单(实)变量实值函数的集合 $\mathcal{C}^n(a, b)$, 是实数域上的矢量空间.

³译者注: 在通常的数学文献中, 我们将复系数多项式的空间记作 $\mathbb{C}[t]$, 实系数多项式的空间记作 $\mathbb{R}[t]$. 这里表征多项式自变量的字母 t 也叫作不定元. 而下面次数不超过 n 的多项式构成的集合则对应记作 $\mathbb{C}[t]_n$ 和 $\mathbb{R}[t]_n$. 这套记号可以推广到任意交换环 R 上, 环的概念可见后面代数相关章节.

(15) 区间 (a, b) 上, 所有光滑单(实)变量实值函数的集合 $\mathcal{C}^\infty(a, b)$, 是实数域上的矢量空间.

从上面的例子可以清晰地看出, 矢量空间既取决于矢量的性质, 也取决于标量的性质.⁴

定义 2.1.3 (线性无关、线性相关、线性组合). 给定一组矢量 $|a_1\rangle, |a_2\rangle, \dots, |a_n\rangle$, 如果对 $\alpha_i \in \mathbb{C}$, 由 $\sum_{i=1}^n \alpha_i |a_i\rangle = 0$ 可以推出 $\alpha_i = 0$ 对所有 i 成立, 则称这组矢量线性无关 (*linear independent*), 反之则称其线性相关 (*linear dependent*). 我们称 $\sum_{i=1}^n \alpha_i |a_i\rangle$ 为 $\{|a_i\rangle\}_{i=1}^n$ 的线性组合 (*linear combination*).

2.1.1. 子空间

给定矢量空间 \mathcal{V} , 我们可以考察 \mathcal{V} 内一些矢量的集合 \mathcal{W} , 即 \mathcal{V} 的子集. 因为 \mathcal{W} 是个子集, 它就包含了很多矢量, 但是这并不足以保证这些矢量的线性组合仍旧在里面. 我们现在来探究使其成立的条件.

定义 2.1.4 (子空间). 所谓矢量空间 \mathcal{V} 的子空间 \mathcal{W} , 就是 \mathcal{V} 的一个非空子集, 并且满足若 $|a\rangle, |b\rangle \in \mathcal{W}$, 则对任意 $\alpha, \beta \in \mathbb{C}$, 都有 $\alpha|a\rangle + \beta|b\rangle$ 属于 \mathcal{W} .

读者可以自行验证, 子空间本身也是个矢量空间. 此外, 两个子空间的交集同样是子空间.

例 2.1.5. 接下来会给出『[例 2.1.2](#)』中所述一些矢量空间的子空间. 请读者验证它们确实构成对应的子空间.

- (1) 实数这个“空间”是实数域上矢量空间 \mathbb{C} 的一个子空间.
- (2) \mathbb{R} 不是 \mathbb{C} 在复数域上的子空间, 因为如『[例 2.1.2](#)』解释的那样, \mathbb{R} 不可能是复数域上的矢量空间.
- (3) 沿给定直线, 并且过原点的所有矢量之集合, 是 \mathbb{R} 上平面矢量(或空间矢量)这个矢量空间的子空间.
- (4) $\mathcal{P}_n^r[t]$ 是 $\mathcal{P}_n^c[t]$ 的子空间.
- (5) 若将 \mathbb{C}^{n-1} 等同于最后一个分量为零的所有复 n 元组, 则 \mathbb{C}^{n-1} 就是 \mathbb{C}^n 的一个子空间. 一般地, 设 $m < n$, 如果将 \mathbb{C}^n 等同于最后 $n-m$ 个分量均为零的所有复 n 元组, 则 \mathbb{C}^{n-m} 是 \mathbb{C}^n 的子空间.
- (6) 设 $r \leq m, s \leq n$, 则 $\mathcal{M}^{r \times s}$ 就是 $\mathcal{M}^{m \times n}$ 的子空间. 此处, 我们要将 $r \times s$ 矩阵等同于其最后 $m-r$ 行和 $n-s$ 列均为零的 $m \times n$ 矩阵.

⁴译者注: 从这里的例子我们能够直接看到的是, 给定所谓矢量的集合 \mathcal{V} 和标量的集合 \mathbb{F} , 这两者的性质同时决定了 \mathcal{V} 是否是 \mathbb{F} 上的矢量空间.

- (7) $\mathcal{P}_m^c[t]$ 是 $\mathcal{P}_n^c[t]$ 的子空间, 其中 $m < n$.
- (8) $\mathcal{P}_m^r[t]$ 是 $\mathcal{P}_n^r[t]$ 的子空间, 其中 $m < n$. 注意, $P_n^r[t]$ 和 $\mathcal{P}_m^r[t]$ 都只是实数域上的矢量空间.
- (9) $m < n$ 时 \mathbb{R}^m 是 \mathbb{R}^n 的子空间. 因此, \mathbb{R} (也就是平面) 是 \mathbb{R}^3 (也就是 Euclid 空间) 的子空间. 同样地, $\mathbb{R}^1 := \mathbb{R}$ 同时是平面 \mathbb{R}^2 和 Euclid 空间 \mathbb{R}^3 的子空间.
- (10) 设 \mathbf{a} 沿着 x 轴 (这是 \mathbb{R}^2 的一个子空间), \mathbf{b} 沿着 y 轴 (这也是 \mathbb{R}^2 的一个子空间). 那么, $\mathbf{a} + \mathbf{b}$ 一般既不沿着 x 轴, 也不沿着 y 轴. 这就说明, 两个子空间的并集通常未必还是子空间.

定理 2.1.6. 若 S 为矢量空间 \mathcal{V} 内矢量构成的任意非空集合, 则它里面所有矢量的线性组合构成的集合 \mathcal{W}_S 就是 \mathcal{V} 的子空间. 我们称 \mathcal{W}_S 为 S 的张成子空间 (*span*), 也称 S 张成了 \mathcal{W}_S , 亦或者说 \mathcal{W}_S 由 S 张成. 通常也将 \mathcal{W}_S 记作 $\text{Span}\{S\}$.

我们将『定理 2.1.6』的证明留作习题 (『习题 2.6』).

定义 2.1.7 (基). 所谓矢量空间 \mathcal{V} 的一组基 (*basis*, 也称作基底),⁵ 就是一个线性无关矢量构成的集合 B , 并且这个集合张成了整个 \mathcal{V} . 这个集合 B 中的元素称作 \mathcal{V} 的基矢量, 简称基矢. 如果 B 是有限集, 则称 \mathcal{V} 是有限维的 (*finite-dimensional*); 反之, 则称其是无穷维的 (*infinite-dimensional*).

注意, 上述定义中我们只用了矢量空间的一组基就规定了它是有限维还是无限维, 这是没有问题的, 因为我们有下述不加证明的定理 (见 [Axler, 1996], p. 31):

定理 2.1.8. 对于给定的有限维矢量空间, 其所有基都具有相同数目的基矢.

定义 2.1.9 (维数). 矢量空间 \mathcal{V} 基的基数称作 \mathcal{V} 的维数 (*dimension*), 记作 $\dim \mathcal{V}$. 为强调其对标量域的依赖性, 我们也会使用 $\dim_{\mathbb{C}} \mathcal{V}$ 和 $\dim_{\mathbb{R}} \mathcal{V}$ 这对记号. 有时候也会将 N 维矢量空间记作 \mathcal{V}_N .

若 $|a\rangle$ 是 N 维矢量空间 \mathcal{V} 中的矢量, 且 $B = \{|a_i\rangle\}_{i=1}^N$ 是这个空间的一组基, 则根据基的定义可知, 存在唯一的 (见『习题 2.4』) 一组标量 $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ 使得 $|a\rangle = \sum_{i=1}^N \alpha_i |a_i\rangle$. 集合 $\{\alpha_i\}_{i=1}^N$ 就称作 $|a\rangle$ 在基 B 下的分量 (*components*).⁶

⁵译者注: 有学者主张基不应该用量词“组”, 因为我们提到基的时候实际上是个集合, 所以应当用量词“个”. 但考虑到我们同样可以将基理解为基矢, 因此一组基就是一组基矢的含义. 在这个意义上用量词“组”就是说得过去的, 是故我们这里采用这一常见说法.

⁶译者注: 原则上讲, 应该把这个集合叫做坐标 (分量), 而非分量.

例 2.1.10. 下面给出的是『例 2.1.2』中所给矢量空间的基.

- (1) 数字 1 (或者任意非零实数) 就是 \mathbb{R} 的基, 从而 \mathbb{R} 是 1 维的.
- (2) 数字 1 以及 $i = \sqrt{-1}$ (或任意一对不同的非零复数⁷) 就是实数域上的矢量空间 \mathbb{C} 的一组基. 因此, 这个空间的维数为 2.
- (3) 数字 1 (或者任意非零复数) 都是复数域上的矢量空间 \mathbb{C} 的基, 从而这个空间是 1 维的. 注意, 尽管这个例子中的矢量和上一个例子相同, 但是改变标量的性质后, 空间的维数就变了.
- (4) 沿空间中三条坐标轴的单位矢量构成的集合 $\{\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z\}$ 构成空间的一组基. 因此该空间就是 3 维的.
- (5) 多项式空间 $\mathcal{P}^c[t]$ 的一组基可由单项式 $1, t, t^2, \dots$ 构成. 很明显, 该空间是无穷维的.
- (6) 空间 \mathbb{C}^n 的一组基由 $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_n$ 构成, 其中 $\hat{\mathbf{e}}_j$ 就是第 j 个位置等于 1, 剩余位置都等于 0 的 n 元组. 这组基就是 \mathbb{C}^n 的标准基 (*standard basis*). 很明显, 该空间是 n 维的.
- (7) 空间 $\mathcal{M}^{m \times n}$ 的一组基由 $\mathbf{e}_{11}, \mathbf{e}_{12}, \dots, \mathbf{e}_{ij}, \dots, \mathbf{e}_{mn}$ 给出, 其中 \mathbf{e}_{ij} 是除了 i 行 j 列元素为 1, 其余矩阵元都是 0 的 $m \times n$ 矩阵.
- (8) 由单项式 $1, t, t^2, \dots, t^n$ 组成的集合就构成 $\mathcal{P}_n^c[t]$ 的一组基. 从而, 这个空间是 $(n+1)$ 维的.
- (9) \mathbb{C}^n 的标准基也是 \mathbb{R}^n 的一组基, 它同样称作 \mathbb{R}^n 的标准基. 因此, \mathbb{R}^n 就是 n 维的.
- (10) 如果我们假设 $a < 0 < b$, 则单项式 $1, x, x^2, \dots$ 的集合就构成 $C^\infty(a, b)$ 的一组基, 这是因为根据 *Taylor* 定理, 任意属于 $C^\infty(a, b)$ 的函数都可以在 $x = 0$ 处展开为无穷幂级数. 因此, 这个空间就是无穷维的.

注 2.1.1. 给定空间 \mathcal{V} 以及它的一组基 $B = \{|a_i\rangle\}_{i=1}^n$, 则 B 中任意 m 个矢量 ($m < n$) 的张成就是 \mathcal{V} 的一个 m 维子空间.

2.1.2. 商空间

设 \mathcal{W} 是矢量空间 \mathcal{V} 的子空间, 在 \mathcal{V} 上定义关系如下: 若 $|a\rangle \in \mathcal{V}$, $|b\rangle \in \mathcal{V}$, 则称 $|a\rangle - |b\rangle$ 属于 \mathcal{W} 时 $|a\rangle$ 和 $|b\rangle$ 相关, 记作 $|a\rangle \bowtie |b\rangle$. 容易验证, \bowtie 是个等价关系. 将 $|a\rangle$ 在该等价关系下的等价类记作 $[\![a]\!]$, 并将该等价关系下的商集 $\{[\![a]\!] \mid |a\rangle \in \mathcal{V}\}$ 记作 \mathcal{V}/\mathcal{W} . 在商集上定义加法和标量乘法如下:

$$\alpha [\![a]\!] + \beta [\![b]\!] = [\![\alpha a + \beta b]\!], \quad (2.1)$$

⁷译者注: 注意, 这对复数里面至少得有个虚数. 后面讨论基时提到多个复数时, 都应当作此理解.

其中 $\llbracket \alpha a + \beta b \rrbracket$ 是 $\alpha|a\rangle + \beta|b\rangle$ 的等价类. 这个式子要有意义, 就必须让它不依赖于等价类代表元的选择. 换言之, 若 $\llbracket a' \rrbracket = \llbracket a \rrbracket$ 且 $\llbracket b' \rrbracket = \llbracket b \rrbracket$, 我们是否有 $\llbracket \alpha a' + \beta b' \rrbracket = \llbracket \alpha a + \beta b \rrbracket$? 要让这个关系成立, 我们就需要有

$$(\alpha|a'\rangle + \beta|b'\rangle) - (\alpha|a\rangle + \beta|b\rangle) \in \mathcal{W}.$$

由于 $|a'\rangle \in \llbracket a \rrbracket$, 因此存在 $|w_1\rangle \in \mathcal{W}$ 使得 $|a'\rangle = |a\rangle + |w_1\rangle$. 类似地, 可以写下 $|b'\rangle = |b\rangle + |w_2\rangle$. 于是乎,

$$(\alpha|a'\rangle + \beta|b'\rangle) - (\alpha|a\rangle + \beta|b\rangle) = \alpha|w_1\rangle + \beta|w_2\rangle.$$

由于 \mathcal{W} 是个子空间, 因此上式的右边就属于 \mathcal{W} .

有时候也会将 $\llbracket a \rrbracket$ 写作 $|a\rangle + \mathcal{W}$. 在这一记号下, 就有如下等式:

$$|w\rangle + \mathcal{W} = \mathcal{W}, \mathcal{W} + \mathcal{W} = \mathcal{W}, \alpha\mathcal{W} = \mathcal{W}, \alpha\mathcal{W} + \beta\mathcal{W} = \mathcal{W}.$$

上述等式不过是下述事实的缩写罢了: (1) \mathcal{W} 中两个矢量的和矢量还在 \mathcal{W} 内; (2) 标量和 \mathcal{W} 中矢量的乘积还在 \mathcal{W} 内; (3) \mathcal{W} 中任意两个矢量的线性组合还在 \mathcal{W} 内.

我们如何才能得到 \mathcal{V}/\mathcal{W} 的一组基呢? 设 $\{|a_i\rangle\}$ 是 \mathcal{W} 的一组基. 现将其扩充为 \mathcal{V} 的一组基 $\{|a_i\rangle, |b_j\rangle\}$. 此时, $\{\llbracket b_j \rrbracket\}$ 就构成 \mathcal{V}/\mathcal{W} 的一组基. 事实上, 令 $\llbracket a \rrbracket \in \mathcal{V}/\mathcal{W}$. 那么, 由于 $|a\rangle \in \mathcal{V}$, 所以

$$\begin{aligned} \llbracket a \rrbracket &:= |a\rangle + \mathcal{W} = \overbrace{\sum_i^{\infty} \alpha_i |a_i\rangle} + \sum_j \beta_j |b_j\rangle + \mathcal{W} \\ &= \sum_j \beta_j |b_j\rangle + \mathcal{W}. \end{aligned}$$

由此即可推出

$$\llbracket a \rrbracket = \left\llbracket \sum_j \beta_j |b_j\rangle \right\rrbracket = \sum_j \beta_j \llbracket b_j \rrbracket.$$

而这就说明 $\{\llbracket b_j \rrbracket\}$ 张成了 \mathcal{V}/\mathcal{W} . 要证明其构成一组基, 则只需再证明它们线性无关即可. 因此, 假定 $\sum_j \beta_j \llbracket b_j \rrbracket = \llbracket 0 \rrbracket$. 而这意味着

$$\sum_j \beta_j |b_j\rangle + \mathcal{W} = |0\rangle + \mathcal{W} = \mathcal{W} \Rightarrow \sum_j \beta_j |b_j\rangle \in \mathcal{W}.$$

于是, 上面最后给出的那个和式就必然是 $\{|a_i\rangle\}$ 的线性组合:

$$\sum_j \beta_j |b_j\rangle = \sum_i \alpha_i |a_i\rangle.$$

将其改写为

$$\sum_j \beta_j |b_j\rangle - \sum_i \alpha_i |a_i\rangle = 0.$$

上式左边是 \mathcal{V} 中一组基的线性组合, 它的结果等于零. 因此, 出现的所有系数都必然等于零, 这自然包括所有的 β_j .

上述论证给出的一个直接推论就是⁸

$$\dim(\mathcal{V}/\mathcal{W}) = \dim \mathcal{V} - \dim \mathcal{W}. \quad (2.2)$$

⁸译者注: 因为 $\{a_i\}$ 的基数为 $\dim \mathcal{W}$, 而它要扩充为 \mathcal{V} 的一组基, 就必须再添加 $\dim \mathcal{V} - \dim \mathcal{W}$ 个线性无关的矢量, 而这正是 $\{|b_j\rangle\}$ 的基数, 也就是 \mathcal{V}/\mathcal{W} 基矢量的数目.

2.1.3. 直和

有时候可以将一个矢量空间分割成一些特殊的(不相交)子空间, 而且这样还会带来一些方便. 举个例子, 为研究在中心力场影响下某粒子在 \mathbb{R}^3 中的运动, 我们可以将运动分解为到角动量方向上的投影, 以及到垂直于角动量的平面上的投影. 这就对应了将任意空间矢量分解为 xy 平面上的一个矢量, 以及沿着 z 轴方向的一个矢量. 我们可以将此推广到任意矢量空间, 不过首先要给出一些要用到的记号: 用 \mathcal{U} 和 \mathcal{W} 表示矢量空间 \mathcal{V} 的两个子空间. 用 $\mathcal{U} + \mathcal{W}$ 表示 \mathcal{V} 中所有可以写成两个矢量之和(其中一个属于 \mathcal{U} , 另一个属于 \mathcal{W})的那些矢量构成的集合. 这样定义的就是所谓子空间之和⁹ (sum of subspaces). 容易证明, $\mathcal{U} + \mathcal{W}$ 是 \mathcal{V} 的子空间¹⁰.

例 2.1.11. 设 \mathcal{U} 为 xy 平面, 而 \mathcal{W} 为 yz 平面. 它们都是 \mathbb{R}^3 的子空间, 于是 $\mathcal{U} + \mathcal{W}$ 也是其子空间. 事实上, $\mathcal{U} + \mathcal{W} = \mathbb{R}^3$, 因为给定任意的 $(x, y, z) \in \mathbb{R}^3$, 我们都可以将其改写为

$$(x, y, z) = \underbrace{\left(x, \frac{1}{2}y, 0\right)}_{\in \mathcal{U}} + \underbrace{\left(0, \frac{1}{2}y, z\right)}_{\in \mathcal{W}}.$$

这个分解并不唯一: 我们也可以将其分解为

$$(x, y, z) = \left(x, \frac{1}{3}y, 0\right) + \left(0, \frac{2}{3}y, z\right),$$

当然还有很多其他的分解方式.

定义 2.1.12 (直和). 设 \mathcal{U} 和 \mathcal{W} 为矢量空间 \mathcal{V} 的子空间, 若它们满足 $\mathcal{V} = \mathcal{U} + \mathcal{W}$ 和 $\mathcal{U} \cap \mathcal{W} = \{0\}$. 那么, 我们就称 \mathcal{V} 是 \mathcal{U} 和 \mathcal{W} 的直和 (direct sum), 记作 $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$.

命题 2.1.13 (直和的唯一性). 设 \mathcal{U} 和 \mathcal{W} 是 \mathcal{V} 的子空间, 并使得 $\mathcal{V} = \mathcal{U} + \mathcal{W}$. 那么, $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$ 的充要条件是: \mathcal{V} 中任意非零矢量可以唯一地写成 \mathcal{U} 中一个矢量和 \mathcal{W} 中一个矢量之和.

⁹译者注: 用符号表示的话, 子空间之和定义为

$$\mathcal{U} + \mathcal{W} := \{|u\rangle + |w\rangle \in \mathcal{V} \mid |u\rangle \in \mathcal{U}, |w\rangle \in \mathcal{W}\}.$$

私以为用符号写比用自然语言写更容易理解.

¹⁰译者注: 要证明是子空间, 只需证明其对加法和标量乘法封闭. 由于 \mathcal{U} 和 \mathcal{W} 本身是子空间, 它们对加法和数乘封闭, 自然其元素的线性组合仍旧封闭. 进而我们看到 $\mathcal{U} + \mathcal{W}$ 中任意两个矢量的线性组合还在这个空间中(只需利用交换律将其重新组合一下即可).

证. 假设 $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, 并且设 $|v\rangle \in \mathcal{V}$ 有两种符合要求, 但又不同的分解方式:

$$|v\rangle = |u\rangle + |w\rangle = |u'\rangle + |w'\rangle.$$

它等价于说

$$|u\rangle - |u'\rangle = |w'\rangle - |w\rangle.$$

上式左边属于 \mathcal{U} , 右边属于 \mathcal{W} , 既然二者相等, 则左边也必须属于 \mathcal{W} . 因此, 左边就必须等于零.¹¹ 同样地, 右边也必须为零. 由此可知 $|u\rangle = |u'\rangle$, $|w'\rangle = |w\rangle$. 这就说明 $|v\rangle$ 的分解方式是唯一的.

反过来, 假设 \mathcal{V} 中任意矢量符合要求的分解方式唯一. 现在假设 $|a\rangle \in \mathcal{U} \cap \mathcal{W}$, 则我们就有如下分解方式:

$$|a\rangle = \underbrace{\frac{1}{3}|a\rangle}_{\in \mathcal{U}} + \underbrace{\frac{2}{3}|a\rangle}_{\in \mathcal{W}} = \underbrace{\frac{1}{4}|a\rangle}_{\in \mathcal{U}} + \underbrace{\frac{3}{4}|a\rangle}_{\in \mathcal{W}}.$$

这就表明 $|a\rangle$ 有两种不同的分解方式. 根据假设, $|a\rangle$ 不可能是非零矢量, 这就说明仅有同时属于 \mathcal{U} 和 \mathcal{W} 的矢量就是零矢量. 由此推出 $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$. \square

更一般地, 我们有如下情形:

定义 2.1.14 (一般的直和). 设 $\{\mathcal{U}_i\}_{i=1}^r$ 为 \mathcal{V} 的子空间, 若它们满足

$$\mathcal{V} = \mathcal{U}_1 + \cdots + \mathcal{U}_r, \text{ 且对所有 } i, j = 1, \dots, r \text{ 有 } \mathcal{U}_i \cap \mathcal{U}_j = \{|0\rangle\}.$$

则称 \mathcal{V} 是 $\{\mathcal{U}_i\}_{i=1}^r$ 的直和, 记作

$$\mathcal{V} = \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_r = \bigoplus_{i=1}^r \mathcal{U}_i.$$

设 $\mathcal{W} = \mathcal{U}_1 \oplus \cdots \oplus \mathcal{U}_s$ 是 s 个子空间的直和 (它们不必张成整个 \mathcal{V}). 现将 \mathcal{W} 写作 $\mathcal{W} = \mathcal{U}_1 \oplus \mathcal{W}'$, 其中 $\mathcal{W}' = \mathcal{U}_2 \oplus \cdots \oplus \mathcal{U}_s$. 设 $\{|u_i\rangle\}_{i=1}^s$ 为一组非零矢量, 其中 $|u_i\rangle \in \mathcal{U}_i$. 假若有

$$\alpha_1|u_1\rangle + \alpha_2|u_2\rangle + \cdots + \alpha_s|u_s\rangle = |0\rangle, \quad (2.3)$$

令上式左边后面的 $s-1$ 项为 $\alpha|w'\rangle \in \mathcal{W}'$, 则有

$$\alpha_1|u_1\rangle + \alpha|w'\rangle = |0\rangle,$$

此即是说

$$\alpha_1|u_1\rangle = -\alpha|w'\rangle.$$

¹¹译者注: 注意我们假设了 $\mathcal{V} = \mathcal{W} \oplus \mathcal{U}$, 因此 \mathcal{W} 和 \mathcal{U} 交集为零矢量.

通过上式左边可得 $\alpha_1|u_1\rangle \in \mathcal{U}_1$, 而通过右边则可得到 $\alpha_1|u_1\rangle \in \mathcal{W}'$, 进而我们必然就有 $\alpha_1|u_1\rangle = |0\rangle$. 然而根据假设 $|u_1\rangle \neq |0\rangle$, 是故此时必有 $\alpha_1 = 0$. 进而现在就可以将『(2.3) 式』写作

$$\alpha_2|u_2\rangle + \alpha_3|u_3\rangle + \cdots + \alpha_s|u_s\rangle = |0\rangle.$$

与前文操作类似, 将上式改写为

$$\alpha_2|u_2\rangle + \beta|w''\rangle = |0\rangle \Rightarrow \alpha_2|u_2\rangle = -\beta|w''\rangle,$$

其中 $\mathcal{W}' = \mathcal{U}_2 \oplus \mathcal{W}''$, 而 $\mathcal{W}'' = \mathcal{U}_3 \oplus \cdots \oplus \mathcal{U}_s$, 并且 $|w''\rangle \in \mathcal{W}''$. 同样方式即可论证得到 $\alpha_2 = 0$. 以此类推, 最终我们就得到了下述命题:

命题 2.1.15. 『定义 2.1.14』中来自不同子空间中的矢量是线性无关的.

命题 2.1.16 (直和分解的存在性). 设 \mathcal{U} 是 \mathcal{V} 的子空间, 则存在 \mathcal{V} 的子空间 \mathcal{W} 使得 $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$.

证. 设 $\{|u_i\rangle\}_{i=1}^m$ 是 \mathcal{U} 的一组基. 现将其扩充为 \mathcal{V} 的一组基 $\{|u_i\rangle\}_{i=1}^N$, 则 $\mathcal{W} = \text{Span}\{u_j\}_{j=m+1}^N$ 就满足要求. \square

例 2.1.17. 设 \mathcal{U} 为 xy 平面, \mathcal{W} 为 z 轴. 它们均为 \mathbb{R}^3 的子空间, 从而 $\mathcal{U} + \mathcal{W}$ 亦为其子空间. 不仅如此, 显然有 $\mathcal{U} + \mathcal{W} = \mathbb{R}^3$. 这是因为, 对任意 \mathbb{R}^3 中的矢量 (x, y, z) , 我们均可以将其写作

$$(x, y, z) = \underbrace{(x, y, 0)}_{\in \mathcal{U}} + \underbrace{(0, 0, z)}_{\in \mathcal{W}}.$$

这一分解显然是唯一的, 因此 $\mathbb{R}^3 = \mathcal{U} \oplus \mathcal{W}$.

命题 2.1.18 (直和空间的维数). 若 $\mathcal{V} = \mathcal{U} \oplus \mathcal{W}$, 则 $\dim \mathcal{V} = \dim \mathcal{U} + \dim \mathcal{W}$.

命题 2.1.19. 设 $\{|u_i\rangle\}_{i=1}^m$ 是 \mathcal{U} 的一组基, $\{|w_i\rangle\}_{i=1}^k$ 是 \mathcal{W} 的一组基. 那么容易验证 $\{|u_1\rangle, \dots, |u_m\rangle, |w_1\rangle, \dots, |w_k\rangle\}$ 就是 \mathcal{V} 的一组基. 具体验证细节留作习题¹².

设 \mathcal{U} 和 \mathcal{V} 是任意两个实矢量空间 (或复矢量空间). 现在考察它们底集的 Descartes 积 $\mathcal{W} := \mathcal{U} \times \mathcal{V}$. 在 \mathcal{W} 上定义加法以及标量乘法如下:

$$\begin{aligned} \alpha(|u\rangle, |v\rangle) &= (\alpha|u\rangle, \alpha|v\rangle), \\ (|u_1\rangle, |v_1\rangle) + (|u_2\rangle, |v_2\rangle) &= (|u_1\rangle + |u_2\rangle, |v_1\rangle + |v_2\rangle). \end{aligned} \tag{2.4}$$

¹²译者注: 这组矢量张成 \mathcal{V} 显然, 是故只需证明其线性无关. 令其线性组合等于零矢量, 移项后则可得到 $\alpha_1|u_1\rangle + \cdots + \alpha_m|u_m\rangle = -\beta_1|w_1\rangle - \cdots - \beta_k|w_k\rangle$ 这样的表达式, 左边属于 \mathcal{U} , 右边属于 \mathcal{W} , 是故必然都等于零. 而它们本身是子空间基矢量的线性组合, 因此系数必然全部为零, 由此即证它们线性无关.

再规定 $|0\rangle_{\mathcal{W}} = (|0\rangle_{\mathcal{U}}, |0\rangle_{\mathcal{V}})$, 则 \mathcal{W} 就是个矢量空间¹³. 不仅如此, 如果我们将 \mathcal{U} 和 \mathcal{V} 分别等同于形式为 $(|u\rangle, |0\rangle_{\mathcal{V}})$ 和 $(|0\rangle_{\mathcal{W}}, |v\rangle)$ 的矢量构成的空间, 那么 \mathcal{U} 和 \mathcal{V} 就成了 \mathcal{W} 的子空间. 如果 $|w\rangle \in \mathcal{W}$ 同时属于 \mathcal{U} 和 \mathcal{V} , 则它必然可以同时写成 $(|u\rangle, |0\rangle_{\mathcal{V}})$ 和 $(|0\rangle_{\mathcal{U}}, |v\rangle)$ 的形式, 亦即 $(|u\rangle, |0\rangle_{\mathcal{V}}) = (|0\rangle_{\mathcal{U}}, |v\rangle)$. 而这仅有可能在 $|u\rangle = |0\rangle_{\mathcal{U}}$ 且 $|v\rangle = |0\rangle_{\mathcal{V}}$ 的时候发生, 而这也就意味着 $|w\rangle = |0\rangle_{\mathcal{W}}$. 因此, \mathcal{U} 和 \mathcal{V} 中仅有的共同矢量就是零矢量. 职是之故, 我们得到下述命题:

命题 2.1.20. 设 \mathcal{U} 和 \mathcal{V} 为任意两个实矢量空间 (或复矢量空间). 那么, 在『(2.4) 式』定义的运算下, 它们的 Descartes 积 $\mathcal{W} := \mathcal{U} \times \mathcal{V}$ 就是个矢量空间. 不仅如此, 如果将 \mathcal{U} 和 \mathcal{V} 分别等同于那些形式为 $(|u\rangle, |0\rangle_{\mathcal{V}})$ 和 $(|0\rangle_{\mathcal{U}}, |v\rangle)$ 的矢量, 那么 $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.

设 $\{|a_i\rangle\}_{i=1}^M$ 是 \mathcal{U} 的一组基, $\{|b_j\rangle\}_{j=1}^N$ 是 \mathcal{V} 的一组基. 现于 $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$ 中定义矢量 $\{|c_k\rangle\}_{k=1}^{M+N}$ 如下:

$$\begin{aligned}|c_k\rangle &= (|a_k\rangle, |0\rangle_{\mathcal{V}}), & \text{若 } 1 \leq k \leq M \\ |c_k\rangle &= (|0\rangle_{\mathcal{U}}, |b_{k-M}\rangle), & \text{若 } M+1 \leq k \leq M+N.\end{aligned}\tag{2.5}$$

此时 $\{|c_k\rangle\}_{k=1}^{M+N}$ 就是线性无关的. 事实上,

$$\sum_{k=1}^{M+N} \gamma_k |c_k\rangle = |0\rangle_{\mathcal{W}}$$

等价于

$$\sum_{k=1}^M \gamma_k (|0\rangle_{\mathcal{U}}, |0\rangle_{\mathcal{V}}) + \sum_{j=1}^N \gamma_{M+j} (|0\rangle_{\mathcal{U}}, |b_j\rangle) = (|0\rangle_{\mathcal{U}}, |0\rangle_{\mathcal{V}}),$$

此即是说

$$\left(\sum_{k=1}^M \gamma_k |a_k\rangle, |0\rangle_{\mathcal{V}} \right) + \left(|0\rangle_{\mathcal{U}}, \sum_{j=1}^N \gamma_{M+j} |b_j\rangle \right) = (|0\rangle_{\mathcal{U}}, |0\rangle_{\mathcal{V}}),$$

亦或者说

$$\left(\sum_{k=1}^M \gamma_k |a_k\rangle, \sum_{j=1}^N \gamma_{M+j} |b_j\rangle \right) = (|0\rangle_{\mathcal{U}}, |0\rangle_{\mathcal{V}}).$$

¹³译者注: 数学上将这样定义的矢量空间 \mathcal{W} 称作 \mathcal{U} 和 \mathcal{V} 的直积 (direct product). 我们接下来的论证说明, 有限个矢量空间的直积和直和实际上是一回事 (或者用数学的语言来说, 它们同构). 它们的区别会在参与运算的矢量空间有无限个时产生. 值得一提的是, 在部分物理学书籍中, 会将我们后面会定义的张量积命名为直积, 这在一定程度上导致了两个学科同名术语的混乱, 译者不支持这种做法.

而这就意味着

$$\sum_{k=1}^M \gamma_k |a_k\rangle = |0\rangle_{\mathcal{U}} \quad \text{且} \quad \sum_{j=1}^N \gamma_{M+j} |b_j\rangle = |0\rangle_{\mathcal{V}}.$$

最后考虑到 $\{|a_i\rangle\}_{i=1}^M$ 和 $\{|b_j\rangle\}_{j=1}^N$ 各自线性无关, 于是对任意的 $1 \leq k \leq M+N$ 均有 $\gamma_k = 0$.

不难证明 $\mathcal{W} = \text{Span}\{|c_k\rangle\}_{k=1}^{M+N}$. 因此, 我们就有下述定理:

定理 2.1.21. 设 $\{|a_i\rangle\}_{i=1}^M$ 是 \mathcal{U} 的一组基, $\{|b_j\rangle\}_{j=1}^N$ 是 \mathcal{V} 的一组基. 那么, 『(2.5) 式』定义的矢量组 $\{|c_k\rangle\}_{k=1}^{M+N}$ 就构成直和空间 $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$ 的一组基. 特别地, \mathcal{W} 的维数就是 $M+N$.

2.1.4. 矢量空间的张量积

上一小节介绍的直和是通过两个矢量空间构造新空间的一种方式, 而这一小节则介绍另一种构造的流程. 同样地, 设 \mathcal{U} 和 \mathcal{V} 为任意两个矢量空间. 但是这时在它们的 Descartes 积上引入标量乘法和双线性条件如下:

$$\begin{aligned} \alpha(|u\rangle, |v\rangle) &= (\alpha|u\rangle, |v\rangle) = (|u\rangle, \alpha|v\rangle), \\ (\alpha_1|u_1\rangle + \alpha_2|u_2\rangle, |v\rangle) &= \alpha_1(|u_1\rangle, |v\rangle) + \alpha_2(|u_2\rangle, |v\rangle), \\ (|u\rangle, \beta_1|v_1\rangle + \beta_2|v_2\rangle) &= \beta_1(|u\rangle, |v_1\rangle) + \beta_2(|u\rangle, |v_2\rangle). \end{aligned} \quad (2.6)$$

在这些性质下, $\mathcal{U} \times \mathcal{V}$ 同样会变成一个矢量空间, 这就是 \mathcal{U} 和 \mathcal{V} 的张量积 (tensor product), 记作 $\mathcal{U} \otimes \mathcal{V}$.^{*} 张量积空间中的矢量则记作 $|u\rangle \otimes |v\rangle$ (偶尔也记作 $|uv\rangle$). 如果 $\{|a_i\rangle\}_{i=1}^M$ 和 $\{|b_j\rangle\}_{j=1}^N$ 分别是 \mathcal{U} 和 \mathcal{V} 的基, 并且

$$|u\rangle = \sum_{i=1}^M \alpha_i |a_i\rangle \quad \text{且} \quad |v\rangle = \sum_{j=1}^N \beta_j |b_j\rangle,$$

那么根据『(2.6) 式』就有

$$\begin{aligned} |u\rangle \otimes |v\rangle &= \left(\sum_{i=1}^M \alpha_i |a_i\rangle \right) \otimes \left(\sum_{j=1}^N \beta_j |b_j\rangle \right) \\ &= \sum_{i=1}^M \sum_{j=1}^N \alpha_i \beta_j |a_i\rangle \otimes |b_j\rangle. \end{aligned}$$

因此, $\{|a_i\rangle \otimes |b_j\rangle\}$ 就构成 $\mathcal{U} \otimes \mathcal{V}$ 的一组基, 从而 $\dim(\mathcal{U} \otimes \mathcal{V}) = \dim \mathcal{U} \dim \mathcal{V}$.

根据『(2.6) 式』, 我们有

$$|0\rangle_{\mathcal{U}} \otimes |v\rangle = (|u\rangle - |u\rangle) \otimes |v\rangle = |u\rangle \otimes |v\rangle - |u\rangle \otimes |v\rangle = |0\rangle_{\mathcal{U} \otimes \mathcal{V}}.$$

与之类似, 就有 $|u\rangle \otimes |0\rangle_{\mathcal{V}} = |0\rangle_{\mathcal{U} \otimes \mathcal{V}}$.

* 我们会在『第??章』中更为详尽地讨论张量积以及张量.

2.2. 内积

在『[定义 2.1.1](#)』中给出的矢量空间太过一般, 缺少结构, 和物理学的需求不匹配¹⁴. 因此, 我们就需要在它上面引入一些别的结构. 矢量空间上可以引入的诸多结构中, 标量积就是有用的一个. 回忆一下高中所学平面矢量以及空间矢量的标量积(或者说点积). 它是这样一个规则: 对每对矢量 \mathbf{a} 和 \mathbf{b} , 我们都赋予一个实数. 这个规则可以用符号表示为 $g : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$, 其中 $g(\mathbf{a}, \mathbf{b}) = \mathbf{a} \cdot \mathbf{b}$. 它是对称的, 即有 $g(\mathbf{a}, \mathbf{b}) = g(\mathbf{b}, \mathbf{a})$; 还对第一个因子是线性的(并且根据对称性, 对第二个因子也是如此)*:

$$g(\alpha\mathbf{a} + \beta\mathbf{b}, \mathbf{c}) = \alpha g(\mathbf{a}, \mathbf{c}) + \beta g(\mathbf{b}, \mathbf{c}),$$

也就是我们熟悉的

$$(\alpha\mathbf{a} + \beta\mathbf{b}) \cdot \mathbf{c} = \alpha\mathbf{a} \cdot \mathbf{c} + \beta\mathbf{b} \cdot \mathbf{c};$$

最后, 这个规则还为矢量赋予了其“长度”: $|\mathbf{a}|^2 = g(\mathbf{a}, \mathbf{a}) \geq 0$, 并且保证了长度为零[†]的矢量只能是零矢量, 即 $g(\mathbf{a}, \mathbf{a}) = 0$ 的充要条件是 $\mathbf{a} = \mathbf{0}$.

我们希望将上面的那些性质推广到抽象的矢量空间上, 并且对应的标量可以是复数. 但是, 将这些性质逐句照搬到这种情形却会导致矛盾! 具体问题是这样的: 考虑非零矢量 $|a\rangle$, 利用前述对两个因子的线性性质, 我们就得到

$$g(i|a\rangle, i|a\rangle) = i^2 g(|a\rangle, |a\rangle) = -g(|a\rangle, |a\rangle). \quad (2.7)$$

而上式的左右两边必有一边为负. 但是, 前文性质中提到矢量的长度是非负的: 对所有的非零矢量 $|a\rangle$, 都得有 $g(|a\rangle, |a\rangle) > 0$. 自然 $i|a\rangle$ 这个非零矢量也得满足条件, 从而产生矛盾! 产生这一问题的原因在于对两个因子都是线性的. 如果我们让『[\(2.7\) 式](#)』中的两个 i 中的一个变成其复共轭, 这个问题也就消失了. 换言之, 我们要将前文对应的那条性质改写成“对其中一个因子是线性的, 而对另一个因子是共轭线性的”. 至于对哪个因子是共轭线性的, 这取决于具体领域的习惯. 而我们选择让其对第一个因子是共轭线性的[‡]. 因此我们就有

$$g(\alpha|a\rangle + \beta|b\rangle, |c\rangle) = \alpha^* g(|a\rangle, |c\rangle) + \beta^* g(|b\rangle, |c\rangle),$$

[†] 在一些书籍(特别是数学文献)中, 则约定对第二个因子是共轭线性的.

¹⁴译者注: 在物理学中我们经常会遇到矢量这个词, 比如经典力学中我们会说位置矢量、速度矢量等, 在量子力学里面我们会说态矢量, 在相对论里面我们又会说 3 矢量、4 矢量等. 但是它们除了共享矢量空间的结构以外, 还各种带有特定的结构, 物理学家往往不会将这些额外结构纳入到对这些对象的称呼上, 这就导致了物理学语境下的“矢量”一词具有相当广泛的外延: A 领域和 B 领域都用到了“矢量”一词, 但它们的性质却不尽相同(因此泛泛而谈某个对象是不是矢量实际上不是个好说法, 而应该明确具体是哪类矢量——即指明对应的矢量空间及其附带结构). 在学习物理学时应在心中注意同一名词在不同领域中的内涵.

其中 α^* 表示复数 α 的复共轭。考虑到一致性，对称性这个条件也得对应修改¹⁵。事实上，我们必须要求 $g(|a\rangle, |b\rangle) = g(|b\rangle, |a\rangle)^*$ ，由此即可很快推出 $g(|a\rangle, |a\rangle)$ 为实数¹⁶——这也是其非负性的必要条件。

矢量空间上内积的存在性是高等分析学中的一个深奥问题。一般而言，若内积存在，则在矢量空间上引入它的方式有很多种¹⁷。然而，正如我们将在『2.2.4节』中看到的那样，有限维矢量空间上总是可以定义内积，并且这个内积是唯一的*。因此，对于实践中遇到的所有场景，我们可以说有限维矢量空间上的“那个”内积¹⁸。另外，和二维以及三维场景中一样，我们可以略去字母 g ，而使用仅涉及矢量的符号表示内积。这种记号有很多备选可供使用，本书采用的符号约定是 Dirac 的 bra(c)ket 记号：我们将 $g(|a\rangle, |b\rangle)$ 记作 $\langle a|b\rangle$ 。借助这套记号，我们可以将内积定义如下：

定义 2.2.1 (内积). 矢量空间 V 上两个矢量 $|a\rangle$ 和 $|b\rangle$ 的内积 (*inner product*) 是个复数 $\langle a|b\rangle$ ，它需要满足下述三个条件：

1. $\langle a|b\rangle = \langle b|a\rangle^*$,
2. $\langle a|(\beta|b\rangle + \gamma|c\rangle) = \beta\langle a|b\rangle + \gamma\langle a|c\rangle$,
3. $\langle a|a\rangle \geq 0$ ，并且 $\langle a|a\rangle = 0$ 的充要条件是 $|a\rangle = |0\rangle$.

最后一个关系称作内积的正定性 (*positive definite property*)。正定的实内积也称作 *Euclid* 内积 (*Euclidean inner product*)，不满足正定性的实内积则称作伪 *Euclid* 内积 (*pseudo-Euclidean inner product*)。

读者可能发现了，在上面的定义中，没有要求对第一个因子是线性的。因为我们早先已经解释了，这样做会和陈述内积“对称性”的第一条性质相矛盾。由于额外加入了取复共轭这一操作，这使得对第一个因子满足线性时不可能的。也是因为取复共轭的这个操作，复矢量空间上的内积就不是双线性的，而通常称作半双线性的 (sesquilinear)，或者 *Hermite* 的 (Hermitian)。

在处理矢量线性组合的内积时，我们会引入下述有用的简写符号：

¹⁵译者注：如果不做修改，那么对称性和对其中一个因子线性必然可以推出对另一个因子线性，这就与共轭线性的条件矛盾了。进一步，我们还可以推断出此时应该修改为共轭对称性（即后文所述性质）。具体方式是这样的：根据共轭线性条件可知 $g(\alpha|a\rangle, |b\rangle) = \alpha^*g(|a\rangle, |b\rangle)$ 。另一方面，交换两个因子顺序后有 $g(|b\rangle, \alpha|a\rangle) = \alpha g(|b\rangle, |a\rangle)$ 。比较一下即可看到如果我们交换两因子位置后取其结果的复共轭，就能让 α 和 α^* 对应匹配，还不会产生任何矛盾！但是注意，这只是个合理猜想，靠共轭线性没法直接推出共轭对称性，但反过来我们可以靠共轭对称性和其中一个因子的线性/共轭线性条件推出对另一个因子的共轭线性/线性条件。

¹⁶译者注：令 $|a\rangle = |b\rangle$ ，则我们看到 $g(|a\rangle, |a\rangle)$ 与其复共轭相等，因此必然是实数。

¹⁷译者注：也就是说，若存在则通常不唯一。举个例子，若 $g : V \times V \rightarrow \mathbb{C}$ 是个内积，那么对任意非零实数 α ，可以断定 αg 也是个内积。

¹⁸译者注：这里的“那个”对应英文中的定冠词“the”。原文用以强调有限维矢量空间上内积的唯一性（而不必以不定冠词“an”指代）。但后续翻译时我们不再强调这一点。

* 这里的唯一性是指在某种等价意义下的唯一性，但这里我们不予讨论。

约定 2.2.2. 将『[定义 2.2.1](#)』中性质 2 的左边简写作 $\langle a|\beta b + \gamma c\rangle$.

采用这一约定的好处在于可以将线性组合视作是单个矢量. 如此一来, 『[定义 2.2.1](#)』中的性质 2 就可以表述为: 若复标量恰好处在一个右矢 (ket) 内, 则它们可以不受影响地“分裂”出来, 即

$$\langle a|\beta b + \gamma c\rangle = \beta \langle a|b\rangle + \gamma \langle a|c\rangle. \quad (2.8)$$

另一方面, 如果复标量出现在第一个因子中 (也就是左矢内), 则将其“分裂”出来时要额外取复共轭:

$$\langle \beta b + \gamma c|a\rangle = \beta^* \langle b|a\rangle + \gamma^* \langle c|a\rangle. \quad (2.9)$$

矢量空间 \mathcal{V} 上如果定义了内积, 则称其为内积空间 (inner product space). 正如上面提到的那样, 有限维矢量空间总是可以变成内积空间.

例 2.2.3. 在这个例子中, 我们介绍一些常见的内积. 读者应当自行验证这些情况下确实得到的是个内积.

- (\mathbb{C}^n 的自然内积) 设 \mathbb{C}^n 中矢量 $|a\rangle, |b\rangle$ 分别为 $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 和 $|b\rangle = (\beta_1, \beta_2, \dots, \beta_n)$, 定义 \mathbb{C}^n 上的内积为

$$\langle a|b\rangle := \alpha_1^* \beta_1 + \alpha_2^* \beta_2 + \cdots + \alpha_n^* \beta_n = \sum_{i=1}^n \alpha_i^* \beta_i.$$

容易验证, 这个乘积确实满足内积所需的所有性质. 比如说, 若 $|b\rangle = |a\rangle$, 则有

$$\langle a|a\rangle = |\alpha_1|^2 + |\alpha_2|^2 + \cdots + |\alpha_n|^2,$$

很明显这是非负的.

- 类似地, 对于 $|a\rangle, |b\rangle \in \mathbb{R}^n$, 亦可以同 \mathbb{C}^n 的方式定义满足内积所有性质的乘积 (此时不必取复共轭).
- 对 $|a\rangle, |b\rangle \in \mathbb{C}^\infty$, 其自然内积定义为 $\langle a|b\rangle = \sum_{i=1}^\infty \alpha_i^* \beta_i$. 这个无穷和的收敛性问题是『[习题 2.18](#)』的任务.
- 设 $x(t), y(t) \in \mathcal{P}^c[t]$, 即关于 t 的复系数多项式. 定义

$$\langle x|y\rangle := \int_a^b w(t)x^*(t)y(t) dt, \quad (2.10)$$

其中 a, b 是使得积分存在的实数 (或者无穷大), 这里的 $w(t)$ 是区间 (a, b) 上严格为正的一个实值连续函数, 称作权函数 (weight function). 如此一来, 『[\(2.10\) 式](#)』就定义了一个内积. 这个内积依赖于权函数 $w(t)$, 是故在无穷维空间 $\mathcal{P}^c[t]$ 上可以定义很多不同的内积.

- (复函数的自然内积) 设 $f, g \in \mathbb{C}(a, b)$ ¹⁹, 定义它上面的内积为

$$\langle f | g \rangle := \int_a^b w(x) f^*(x) g(x) dx.$$

容易验证 $\langle f | g \rangle$ 满足内积的所有要求. 和多项式的那个例子一样, 权函数 $w(x)$ 在区间 (a, b) 上总是正的. 这就是 $\mathbb{C}(a, b)$ 上的标准内积 (*standard inner product*).

¹⁹译者注: 作者在此先没有定义过这个符号, 但根据上下文它应当表示定义在区间 (a, b) 上的复值函数构成的空间.

2.2.1. 正交性

在解析几何以及微积分中, 矢量通常是通过沿坐标轴的单位矢量(即单位长度且互相垂直的矢量)予以描述的. 这样的矢量在抽象的内积空间中同样重要.

定义 2.2.4 (正交、正交归一基). 若 $|a\rangle, |b\rangle \in \mathcal{V}$ 满足 $\langle a|b\rangle = 0$, 则称它们正交 (*orthogonal*). 所谓规范矢量 (*normal vector*), 或者说归一化矢量 (*normalized vector*), 就是满足 $\langle e|e\rangle = 1$ 的矢量 $|e\rangle$.²⁰ 给定 N 维矢量空间 \mathcal{V} 的一组基 $B = \{|e_i\rangle\}_{i=1}^N$, 如果

$$\langle e_i|e_j\rangle = \delta_{ij} := \begin{cases} 1 & \text{若 } i = j, \\ 0 & \text{若 } i \neq j, \end{cases} \quad (2.11)$$

则称其为正交归一基²¹ (*orthonormal basis*), 此处最后定义的 δ_{ij} 称作 *Kronecker delta*.

例 2.2.5. 设 \mathcal{U} 和 \mathcal{V} 为内积空间, 令 $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$. 则在 \mathcal{W} 上可以通过 \mathcal{U} 和 \mathcal{V} 上的内积定义一个新内积. 事实上, 容易验证若 $|w_i\rangle = (|u_i\rangle, |v_i\rangle)$, $i = 1, 2$, 则

$$\langle w_1|w_2\rangle = \langle u_1|u_2\rangle + \langle v_1|v_2\rangle \quad (2.12)$$

就定义了一个 \mathcal{W} 上的内积. 不仅如此, 借助之前给出的等同方式²²

$$\mathcal{U} = \{(|u\rangle, |0\rangle_{\mathcal{V}}) \mid |u\rangle \in \mathcal{U}\} \quad \text{和} \quad \mathcal{V} = \{(|0\rangle_{\mathcal{U}}, |v\rangle) \mid |v\rangle \in \mathcal{V}\},$$

可以看到 \mathcal{U} 中的任意矢量都和 \mathcal{V} 中的任意矢量正交.

例 2.2.6. 这里是一些正交归一基的例子.

- \mathbb{R}^n (或者 \mathbb{C}^n) 的标准基

$$|e_1\rangle = (1, 0, \dots, 0), \dots, |e_n\rangle = (0, 0, \dots, 1)$$

在这些空间上的自然内积下是正交归一的.

²⁰译者注: normal vector 表示 $\langle e|e\rangle = 1$ 的矢量并不是个常见术语 (事实上在几何中 normal (法向) 和 tangent (切向) 是一对概念, 后加 vector 变成我们熟悉的法矢量和切矢量, 它的含义和这里完全不同, 是故这里我将其翻译为规范矢量, 用以匹配规范正交基的说法), 反倒是后面的归一化矢量是我们常说的 (当然更常见的说法是单位矢量). 因此, 如果后面遇到这个术语, 我们会采用归一化矢量或者单位矢量这种译法.

²¹译者注: 数学文献中也常译作标准正交基, 或者规范正交基.

²²译者注: 这里有点滥用符号, 读者应当清楚此处的等号不是相等的意思, 而是说它们可以视作等同, 也可以将其理解为一个赋值号. 总之 “=” 左右两边的矢量空间不是同一个意思.

- 设 $|e_k\rangle = e^{ikx}/\sqrt{2\pi}$ 为 $\mathbb{C}(0, 2\pi)$ 中的函数，并且 $w(x) = 1$. 则

$$\langle e_k | e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} e^{ikx} dx = 1,$$

并且当 $l \neq k$ 时,

$$\langle e_l | e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ilx} e^{ikx} dx = \frac{1}{2\pi} \int_0^{2\pi} e^{i(k-l)x} dx = 0.$$

因此 $\langle e_l | e_k \rangle = \delta_{lk}$.

2.2.2. Gram-Schmidt 过程

通过取合适的线性组合, 我们总是可以将 \mathcal{V} 的任意一组基变成一组正交归一基. 实现这一目标的过程之一就是 **Gram-Schmidt 正交化** (Gram-Schmidt orthonormalization). 考察一组基 $B = \{|a_i\rangle\}_{i=1}^N$. 我们打算这样取 $|a_i\rangle$ 的线性组合, 以使其结果是正交归一的. 首先, 我们令 $|e_1\rangle = |a_1\rangle / \sqrt{\langle a_1|a_1\rangle}$, 不难发现此时 $\langle e_1|e_1\rangle = 1$. 如果我们从 $|a_2\rangle$ 中减去其沿着 $|e_1\rangle$ 的投影, 那么就可以得到一个和 $|e_1\rangle$ 正交的矢量 (见『图 2.1』).

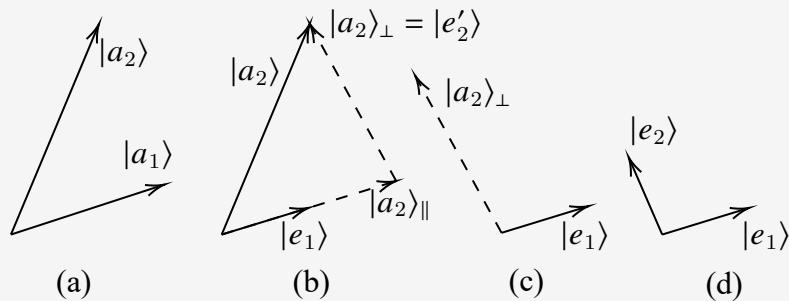


图 2.1: Gram-Schmidt 过程的本质在二维情况下得到了很好的说明. 本图详细描述了构造两个正交矢量的步骤.

上述流程后我们就得到的矢量 $|e'_2\rangle$ 就是 $|e'_2\rangle = |a_2\rangle - \langle e_1|a_2\rangle |e_1\rangle$, 用更为对称的写法就是 $|e'_2\rangle = |a_2\rangle - |e_1\rangle \langle e_1|a_2\rangle$. 很明显, 这个矢量与 $|e_1\rangle$ 正交. 要将 $|e'_2\rangle$ 归一化, 我们只需将其除以 $\sqrt{\langle e'_2|e'_2\rangle}$. 这样得到的矢量 $|e_2\rangle = |e'_2\rangle / \sqrt{\langle e'_2|e'_2\rangle}$ 就是与 $|e_1\rangle$ 正交的归一化矢量. 从 $|a_3\rangle$ 中减去其沿着已有的两个单位矢量方向的投影, 就可以得到矢量

$$|e'_3\rangle = |a_3\rangle - |e_1\rangle \langle e_1|a_3\rangle - |e_2\rangle \langle e_2|a_3\rangle = |a_3\rangle - \sum_{i=1}^2 |e_i\rangle \langle e_i|a_3\rangle,$$

它与 $|e_1\rangle$ 和 $|e_2\rangle$ 都正交 (见『图 2.2』):

$$\langle e_1|e'_3\rangle = \langle e_1|a_3\rangle - \overbrace{\langle e_1|e_1\rangle}^{=1} \langle e_1|a_3\rangle - \overbrace{\langle e_1|e_2\rangle}^{=0} \langle e_2|a_3\rangle = 0.$$

类似地, 就有 $\langle e_2|e'_3\rangle = 0$.

人物志 (Erhard Schmidt)



Erhard Schmidt
(厄尔哈德·施密特)
1876-1959

Erhard Schmidt (1876–1959) 在 David Hilbert 指导下取得了他的博士学位. 他的主要兴趣是积分方程以及 Hilbert 空间. 他就是 **Gram-Schmidt 正交化过程** (即从矢量空间的一组基出发构造出一组正交归一基的过程) 中的那个“Schmidt”. (早在 Gram 和 Schmidt 之前, Laplace 就已经给出了这个过程的一个特殊情形.)

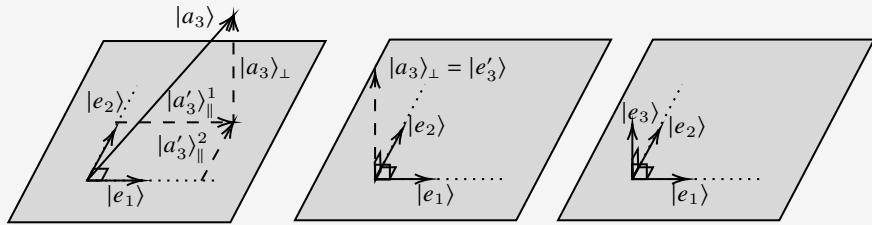


图 2.2: 一旦得到了平面内一对正交归一矢量, 第三个正交归一矢量就容易构造出来了.

1908 年, Schmidt 研究了有无穷多个未知数的无穷多个方程, 并引入了很多几何符号和术语, 现在仍将其用于描述函数空间. Schmidt 的这些想法带来了 Hilbert 空间的几何学. 这受启发于对积分方程(见『第??章』)的研究, 以及对其抽象化的尝试.

早些时候, Hilbert 认为函数是由其 Fourier 系数给出的. 这些系数满足 $\sum_{k=1}^{\infty} a_k^2$ 有限的条件. 进而他引入了使得 $\sum_{n=1}^{\infty} x_n^2$ 有限的实数序列 $\{x_n\}$. Riesz 以及 Fischer 证明了平方可积函数和它们 Fourier 系数的平方可和序列之间一一对应. 在 1907 年, Schmidt 和 Fréchet 则证明了, 若将平方可和序列视作是某个无穷维空间(这是 n 维 Euclid 空间的一个推广)里面点的坐标, 则可以得到一个一致的理论. 于是, 函数就可以视作是某个空间里面的点, 这个空间现在就称作 **Hilbert 空间**(Hilbert space).

一般地, 如果我们已经算得了 m 个正交归一的矢量 $|e_1\rangle, \dots, |e_m\rangle$, 其中 $m < N$. 那么我们就可以通过下述关系式算出下一个:

$$\begin{aligned} |e'_{m+1}\rangle &= |a_{m+1}\rangle - \sum_{i=1}^m |e_i\rangle \langle e_i|a_{m+1}\rangle, \\ |e_{m+1}\rangle &= \frac{|e'_{m+1}\rangle}{\sqrt{\langle e'_{m+1}|e'_{m+1}\rangle}}. \end{aligned} \tag{2.13}$$

尽管我们一直讨论的都是有限维矢量空间, 但『(2.13) 式』给出的这个过程也适用于无穷维空间²³. 读者要注意这样一点, 在 Gram-Schmidt 过程的每一阶段, 我们都是在取最初那些矢量的线性组合.

²³译者注: 我们应当看到, 这个过程无论如何都是可数的, 因此我们没法通过这个过程得到不可数多个正交归一矢量.

2.2.3. Schwarz 不等式

这一小节我们来关注一个同时成立于有限维和无穷维矢量空间的重要不等式。在二维和三维场景中，这个不等式等价于说任意两个矢量之间夹角的余弦值总是小于 1。

定理 2.2.7 (Schwarz 不等式). 对内积空间 \mathcal{V} 里面任意一对矢量 $|a\rangle$ 和 $|b\rangle$, 均有 **Schwarz 不等式 (Schwarz inequality)** 成立, 即有

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

当 $|a\rangle$ 和 $|b\rangle$ 成比例时^a取等号。

^a译者注: 即存在标量 α 使得 $|a\rangle = \alpha|b\rangle$ 或 $|b\rangle = \alpha|a\rangle$.

证. 令 $|c\rangle = |b\rangle - (\langle a|b\rangle / \langle a|a\rangle)|a\rangle$, 则有 $\langle a|c\rangle = 0$. 此时就有 $|b\rangle = (\langle a|b\rangle / \langle a|a\rangle)|a\rangle + |c\rangle$, 取它与自身的内积:

$$\langle b|b\rangle = \left| \frac{\langle a|b\rangle}{\langle a|a\rangle} \right|^2 \langle a|a\rangle + \langle c|c\rangle = \frac{|\langle a|b\rangle|^2}{\langle a|a\rangle} + \langle c|c\rangle.$$

因为 $\langle c|c\rangle \geq 0$, 我们就有

$$\langle b|b\rangle \geq \frac{|\langle a|b\rangle|^2}{\langle a|a\rangle} \Rightarrow \langle a|a\rangle \langle b|b\rangle \geq |\langle a|b\rangle|^2.$$

等号成立的充要条件是 $\langle c|c\rangle = 0$, 它等价于 $|c\rangle = 0$. 而由 $|c\rangle$ 的定义可知, 要让等号成立, $|a\rangle$ 和 $|b\rangle$ 必然成比例²⁴. □

还请读者注意抽象的力量: 我们仅通过内积空间自身最基本的那些假设就导出了 Schwarz 不等式, 而不必考虑内积的具体形式. 因此, 我们不必在每次遇到新的内积空间时再证明对应空间里面的 Schwarz 不等式.



Karl Herman Amandus Schwarz
(卡尔·赫尔曼·阿曼杜斯·施瓦兹)
1843-1921

人物志 (Karl Herman Amandus Schwarz)

Karl Herman Amandus Schwarz (1843–1921) 是位建筑师之子, 出生于现在的波兰索比钦 (Sobiecin). 文理高中²⁵毕业以后, Schwarz 先在柏林大学学了一段时间的化学, 之后转向学习数学, 并在 1864 年取得博士学位. 他深受当时德国当权 (reigning) 数学家的影响, 特别是 Kummer 和 Weierstrass. Schwarz 当年在 Weierstrass 讲述积分学时

²⁴译者注: 这个证明很好地揭示了 Schwarz 不等式的几何意义: $|c\rangle$ 不过是 $|b\rangle$ 去除其沿着 $|a\rangle$ 方向的投影后剩下的那个分量, 而 Schwarz 不等式是说这个分量最多为零.

²⁵译者注: 文理中学 (gymnasium), 是德国教育体系中的说法 (9 年制), 是进入大学的必经之路, 其高中部具有大学预科性质, 通常是那些最优秀的学生才能选择入学文理高中.

所做笔记仍现存于世。他最先在哈雷大学²⁶任职，之后则在苏黎世大学和哥廷根大学任职，再然后于 1892 年在柏林大学获得 Weierstrass 继任者的任命。后面的这几年里，他的生活中充斥着学生与讲座，这并不是他最高产的几年，但他早期的论文确立了他在数学史上的地位。Schwarz 最喜欢的工具是几何学，他很快就用这项工具去研究分析学。他结论性地证明了一些 Riemann 的结果，这些结果早先受到质疑（而且是有理由的质疑）。他的主要结果就是断言平面上的每个单连通区域都可以保角满映²⁷到一个圆形区域上。以此为基础可以得到好几个有名的结论，现在均与 Schwarz 的名字联系在一起，包括反射原理²⁸（principle of reflection）和 Schwarz 引理²⁹。他还研究了最小面积曲面，这是那些涉及到肥皂泡的人所钟爱的几何学分支。

Schwarz 最重要的工作就是他于 Weierstrass 七十诞辰之际，再次展开的对最小面积曲面的研究，特别是一个极小曲面³⁰是否给出最小面积。在此过程中，Schwarz 展示了多重积分的二次变分，利用逐次逼近构造了一个函数，还证明了某些微分方程“最小”本征值的存在性。这项工作还包括了数学中最有名的不等式，该不等式以他的名字命名。

Schwarz 的成功显然源于其天赋与训练与那个时代的数学问题相匹配。然而，他也有个或可看作正面，亦可看作负面的特质——无论问题平凡与否，他都会以相同的程度去关注其细节。这可能至少一定程度解释了为什么他晚年的产出会下降。

Schwarz 对数学以外的东西也很感兴趣——虽说他的婚姻是一场数学婚姻：他娶了数学家 Kummer 的女儿。在数学之外，他还是当地志愿消防队的队长，并且还在当地火车站协助其站长关闭火车门。

²⁶译者注：回忆一下，之前讲述 Cantor 时也提到了这个地方，而 Cantor 同样受到 Weierstrass 的影响。

²⁷译者注：保角映射（conformal mapping）也译作共形映射，前者强调映射前后角度不变（保角），后者强调形状不变（共形）。对此内容详见『??节』。

²⁸译者注：这个原理是说，若 D 和 D' 是复平面上关于实轴对称的两个区域，在 D 上定义有解析函数 $f(z)$ 。若 D 可以连续过渡到实轴上，且 $f(z)$ 在实轴上取实值，则这个函数的定义域也可以解析延拓到 D' 上，延拓后的函数 $\tilde{f}(z)$ 在 D' 上满足 $\tilde{f}(z^*) = [f(z)]^*$ 。

²⁹译者注：这个引理是说，若 $f : D \rightarrow \overline{D}$ 是单位开圆盘 $D = \{z \in \mathbb{C} \mid |z| < 1\}$ 到单位闭圆盘 $\overline{D} = \{z \in \mathbb{C} \mid |z| \leq 1\}$ 的解析函数，且满足 $f(0) = 0$ ，那么 $|f(z)| \leq |z|$ 且 $|f'(0)| \leq 1$ 。

³⁰译者注：极小曲面（minimal surface）指平均曲率为零的曲面。

2.2.4. 矢量的长度

在处理平面和空间中的有向线段这样的对象时, 我们用直观的矢量长度概念去定义点积³¹. 然而, 有时候反过来操作会更方便, 即先引入内积, 然后再定义长度. 这就是这一小节我们要干的事.

定义 2.2.8 (范数). 内积空间里面矢量 $|a\rangle$ 的范数 (norm) $\|a\|$ 定义为 $\|a\| := \sqrt{\langle a|a\rangle}$. 范数可以直观理解为长度 (length). 我们用符号 $\|\alpha a + \beta b\|$ 表示矢量 $\alpha|a\rangle + \beta|b\rangle$ 的范数.

我们可以很轻松地验证范数具有下述性质:

1. 零矢量的范数为零: $\|0\| = 0$.
2. $\|a\| \geq 0$, 并且 $\|a\| = 0$ 的充要条件是 $|a\rangle = |0\rangle$.
3. 对任意复数 α , 均有 $\|\alpha a\| = |\alpha| \|a\|$. *
4. $\|a + b\| \leq \|a\| + \|b\|$. 这个性质也称作三角不等式 (triangle inequality).

一般而言, 矢量空间上任意满足上述四条性质的函数都叫做范数 (norm), 而定义了范数的矢量空间就称作赋范线性空间 (normed linear space), 简称赋范空间 (normed space)³². 我们不必非得有内积才能定义范数.

在赋范空间中, 我们就可以引入两个矢量的“距离”这个想法. 具体而言, 矢量 $|a\rangle$ 和 $|b\rangle$ 之间的距离 $d(a, b)$ 不过就是它们差值的范数, 即 $d(a, b) := \|a - b\|$. 人们很快就能验证这一定义满足距离函数 (或者说度量函数) 所需要的所有性质 (见『[定义 1.3.1](#)』). 然而, 我们也不必非得有范数才能定义距离. 举个例子, 正如『[第一章](#)』中曾经解释的那样, 我们可以定义球面上两点之间的距离, 但是球面上两点相加却是未经定义的, 而加法是矢量空间必须有的结构. 因此, 球面上的点构成度量空间, 但是不构成矢量空间.

内积空间自然是赋范空间, 但是反过来一般却并不成立: 存在一些满足前述范数性质 1-4 的空间, 它们没法提升为内积空间³³. 然而, 如果范数满足平行四边形法则 (parallelogram law):

$$\|a + b\|^2 + \|a - b\|^2 = 2\|a\|^2 + 2\|b\|^2, \quad (2.14)$$

³¹译者注: 回忆高中知识, 我们定义两个有向线段 \vec{a}, \vec{b} 的点积为 $\vec{a} \cdot \vec{b} = |\vec{a}||\vec{b}| \cos(\vec{a}, \vec{b})$. 也就是我们先有矢量长度概念, 然后才有点积.

³²译者注: 由于范数的定义中同时出现了加法和数乘这两种运算, 因此必然定义在矢量空间 (线性空间) 上, 因此可以略去定语“线性”, 而只强调赋范.

³³译者注: 我们将 $\|a\| = \sqrt{\langle a|a\rangle}$ 这个范数称作内积诱导的范数 (induced norm). 此处的断言确切地说, 应当表述为“不存在某个内积使得该内积诱导的范数是给定范数”. 这就是文中“提升”(be promoted to) 的含义.

那么我们就可以定义³⁴

$$\langle a|b \rangle := \frac{1}{4} \{ \|a+b\|^2 - \|a-b\|^2 - i(\|a+ib\|^2 - \|a-ib\|^2)\}. \quad (2.15)$$

可以证明『(2.15) 式』确实定义了一个内积。事实上，我们有下述定理（证明可见 [Friedman, 1982], pp. 203-204）：

定理 2.2.9. 赋范空间为内积空间的充要条件是这个范数满足平行四边形法则。

现考察任意 N 维矢量空间 \mathcal{V} 。选取 \mathcal{V} 的一组基 $\{|a_i\rangle\}_{i=1}^N$ 。对任意矢量 $|a\rangle$ ，设其在这组基下的分量为 $\{\alpha_i\}_{i=1}^N$ ，现定义

$$\|a\|^2 := \sum_{i=1}^N |\alpha_i|^2.$$

读者可以验证这样就定义了一个范数，且该范数满足平行四边形法则。如此一来，根据『定理 2.2.9』即可得到下述结论：

定理 2.2.10. 所有有限维矢量空间都可以变成内积空间。

例 2.2.11 (\mathbb{C}^n 上可以有多种不同的距离函数)。考察空间 \mathbb{C}^n 。它上面的自然内积定义了一个范数。对于矢量 $|a\rangle = (\alpha_1, \dots, \alpha_n)$ ，这个范数就是

$$\|a\| = \sqrt{\langle a|a \rangle} = \left(\sum_{i=1}^n |\alpha_i|^2 \right)^{1/2}.$$

对矢量 $|a\rangle$ 以及 $|b\rangle = (\beta_1, \dots, \beta_n)$ ，上述范数给出如下距离函数：

$$d(a, b) = \|a - b\| = \sqrt{\langle a-b|a-b \rangle} = \left(\sum_{i=1}^n |\alpha_i - \beta_i|^2 \right)^{1/2}.$$

人们也可以定义另外的范数，比如说 $\|a\|_1 := \sum_{i=1}^n |\alpha_i|$ ，它也满足范数的所有性质，并给出下述距离：

$$d_1(a, b) = \|a - b\|_1 = \sum_{i=1}^n |\alpha_i - \beta_i|.$$

在 \mathbb{C}^n 上还可定义如下范数：

$$\|a\|_p := \left(\sum_{i=1}^n |\alpha_i|^p \right)^{1/p},$$

³⁴译者注：『(2.15) 式』也称作极化恒等式 (polarization identity)。

其中 p 是个正整数³⁵. 在高等的分析课程中³⁶, 我们会证明 $\|\bullet\|_p$ 满足范数的所有性质. (证明中非平凡的部分就是验证三角不等式.) 与之相关联的距离为

$$d_p(a, b) = \|a - b\|_p = \left(\sum_{i=1}^n |\alpha_i - \beta_i|^p \right)^{1/p}.$$

前文介绍的那两个范数不过是 $p = 2$ 和 $p = 1$ 的特例.

³⁵译者注: 实际上此处的 p 可以是任意正数, 甚至无穷大. 所得的范数就称作 p 范数, 这是我们最常遇到的范数.

³⁶译者注: 通常是在实变函数论或者泛函分析中

2.3. 线性映射

在丰富矢量空间的结构这一层面上, 我们已经取得了长足的进展, 在它上面增添了范数以及内积. 然而, 这种丰富方式虽然重要, 但是如果仅仅局限在单个矢量空间上, 那么它们的价值也就没有多高了. 我们还想为矢量空间赋予运动自由度的性质, 从而它们可以从一个空间到另一个空间. 承载这一性质的工具就是本节的主题, 即所谓线性映射或者线性变换. 在具体进入主题之前, 我们先考虑与当前讨论内容有关的一些例子, 用以回顾映射这个概念(见『第一章』).

例 2.3.1. 下面是一些比较熟悉的映射的例子.

1. 设 $f : \mathbb{R} \rightarrow \mathbb{R}$ 为 $f(x) = x^2$.
2. 设 $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ 为 $g(x, y) = x^2 + y^2 - 4$.
3. 设 $F : \mathbb{R}^2 \rightarrow \mathbb{C}$ 为 $F(x, y) = U(x, y) + iV(x, y)$, 其中 $U : \mathbb{R}^2 \rightarrow \mathbb{R}$, $V : \mathbb{R}^2 \rightarrow \mathbb{R}$ 为二元实值实变函数.
4. 设 $T : \mathbb{R} \rightarrow \mathbb{R}^2$ 为 $T(t) = (t + 3, 2t - 5)$.
5. 质点在三维空间中的运动可以视作是映射 $M : [a, b] \rightarrow \mathbb{R}^3$, 其中 $[a, b]$ 是实直线³⁷上的一个区间. 现在对每个 $t \in [a, b]$, 我们定义 $M(t) = (x(t), y(t), z(t))$, 其中 $x(t), y(t), z(t)$ 是 t 的实值函数. 如果我们将 t 等同于时间, 它取值于区间 $[a, b]$, 那么 $M(t)$ 就描述了质点的路径(作为时间的函数), 而 a, b 就分别是开始运动的时刻和运动结束的时刻.

现在让我们考察从一个矢量空间 \mathcal{V} 到另一个矢量空间 \mathcal{W} 的任意映射 $F : \mathcal{V} \rightarrow \mathcal{W}$. 我们需要假定这两个矢量空间都定义在相同的标量域上, 比如说它们都是复矢量空间. 现考察 \mathcal{V} 中的矢量 $|a\rangle, |b\rangle$ 以及 \mathcal{W} 中的矢量 $|x\rangle, |y\rangle$. 它们之间满足 $F(|a\rangle) = |x\rangle$ 且 $F(|b\rangle) = |y\rangle$. 一般而言, F 并不能维持矢量空间结构不变. 也就是说, 通常矢量线性组合的像不等于其像的线性组合, 即

$$F(\alpha|a\rangle + \beta|b\rangle) \neq \alpha|x\rangle + \beta|y\rangle.$$

『例 2.3.1』中所有情形在一般情况下都是如此. 不过, 在许多实际应用中, 却需要那些保矢量空间结构(即保线性组合)的映射.

³⁷译者注: 实数集 \mathbb{R} 上有很多结构, 比如线性结构、域结构、群结构、拓扑结构等. 习惯上, 若我们关注的是 \mathbb{R} 的拓扑结构(连续、开集、闭集等), 则称其为实直线(real line). 如果关注域结构, 则称其为实数域(real field). 如果关注其基数性质, 则称其为连续统(continuum).

定义 2.3.2 (线性映射/变换). 从复矢量空间 \mathcal{V} 到复矢量空间 \mathcal{W} 的线性映射 (*linear map*) 就是使得

$$\begin{aligned}\mathbf{T}(\alpha|a\rangle + \beta|b\rangle) &= \alpha\mathbf{T}(|a\rangle) + \beta\mathbf{T}(|b\rangle) \\ \forall|a\rangle, |b\rangle \in \mathcal{V}, \forall\alpha, \beta \in \mathbb{C}\end{aligned}$$

的映射 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$, 也称其为线性变换 (*linear transformation*). ³⁸ 如果线性变换 \mathbf{T} 的定义域和陪域都是 \mathcal{V} , 则称其为 \mathcal{V} 上的自同态 (*endomorphism*), 亦或者 \mathcal{V} 上的线性算符 (*linear operator*). ³⁹ 线性变换在一个矢量上的作用通常会略去括号, 即写作 $\mathbf{T}(|a\rangle) := \mathbf{T}|a\rangle$.

这一定义同样适用于实矢量空间. 注意在定义中要求变换前后的两个矢量空间有相同的标量域, 因为同样的标量会在左边乘以 \mathcal{V} 中的矢量, 而在右边会和 \mathcal{W} 中的矢量相乘.

从 \mathcal{V} 到 \mathcal{W} 所有线性映射的集合就记作 $\mathcal{L}(\mathcal{V}, \mathcal{W})$, 这个集合恰好也是个矢量空间. 这个矢量空间中的零矢量就是零变换 (zero transformation) $\mathbf{0}$, 它将 \mathcal{V} 中的每个矢量都变成 \mathcal{W} 中的零矢量. 两个线性变换 \mathbf{T} 和 \mathbf{U} 的和就是线性变换 $\mathbf{T} + \mathbf{U}$, 它在 $|a\rangle \in \mathcal{V}$ 上的作用定义为 $(\mathbf{T} + \mathbf{U})|a\rangle := \mathbf{T}|a\rangle + \mathbf{U}|a\rangle$. 类似地, 定义 $\alpha\mathbf{T}$ 为 $(\alpha\mathbf{T})|a\rangle := \alpha(\mathbf{T}|a\rangle) = \alpha\mathbf{T}|a\rangle$. 空间 \mathcal{V} 上所有自同态的集合记作 $\mathcal{L}(\mathcal{V})$ 或者 $\text{End}(\mathcal{V})$, 而非 $\mathcal{L}(\mathcal{V}, \mathcal{V})$. 我们将这些观察结果总结如下:

结论 2.3.3. $\mathcal{L}(\mathcal{V}, \mathcal{W})$ 是个矢量空间. 特别地, 单个矢量空间上自同态的集合 $\mathcal{L}(\mathcal{V}) := \text{End}(\mathcal{V}) := \mathcal{L}(\mathcal{V}, \mathcal{V})$ 也是个矢量空间.

* 此处将 $\mathbf{T}|a\rangle$ 简写为 $|\mathbf{T}a\rangle$ 会为我们带来方便. 这种写法进一步允许我们将它的对偶矢量 (见后文) 写成 $\langle \mathbf{T}a|$, 这样就能强调其是和 $|\mathbf{T}a\rangle$ 相关联的那个左矢.

定义 2.3.4 (等距映射). 设 \mathcal{V}, \mathcal{U} 为内积空间. 若线性映射 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{U}$ 满足 *

$$\langle \mathbf{T}a | \mathbf{T}b \rangle = \langle a | b \rangle, \quad \forall |a\rangle, |b\rangle \in \mathcal{V},$$

则称其为等距映射 (*isometric map*). 如果 $\mathcal{U} = \mathcal{V}$, 则称 \mathbf{T} 为线性等距 (*linear isometry*), 或简称为 \mathcal{V} 上的等距. 复/实矢量空间上的等距通常也称作幺正算符/正交算符 (*unitary operator/orthogonal operator*). ⁴⁰

³⁸译者注: 线性变换到底指代什么在不同书中又有不同约定, 有些书将其视作线性映射的代名词 (比如本书), 还有一些书将其视作线性算符的代名词 (比如 *Linear Algebra Done Right*).

³⁹译者注: 所谓同态 (*homomorphism*), 就是保结构的映射 (一些书中也将前面定义的线性映射称作线性同态). 从而自同态就是到自己的保结构映射. 也是因为这个名字, 一些书中会用 $\text{End}(\mathcal{V})$ 表示 \mathcal{V} 上所有线性算符的集合, 用 $\text{Hom}_{\mathbb{C}}(\mathcal{V}, \mathcal{W})$ 表示 \mathcal{V} 到 \mathcal{W} 所有线性映射的集合 (这套符号随着范畴论的发展, 正逐渐成为数学的主流符号). 另外, 线性算符在数学文献中通常译作线性算子, 在物理文献中则称作算符.

⁴⁰译者注: 在数学文献中通常将 *unitary* 翻译为“酉 (的)”, 这个译法是华罗庚先生的提议. 物理文献则偏好幺正这个词, 其中幺的含义是单位, 正的含义是正交.

例 2.3.5. 下面的例子是不同矢量空间中的线性算符. 所有情形下线性的证明都很简单, 因此留给读者充当习题.

1. 设 \mathcal{V} 是个一维空间 (比如说 $\mathcal{V} = \mathbb{C}$). 则 \mathcal{V} 上任意自同态 \mathbf{T} 都具有 $\mathbf{T}|x\rangle = \alpha|x\rangle$ 的形式, 其中 α 是个标量⁴¹. 特别地, 若 \mathbf{T} 是等距, 则 $|\alpha|^2 = 1$. 如果 $\mathcal{V} = \mathbb{R}$ 且 \mathbf{T} 为等距, 则 $\mathbf{T}|x\rangle = \pm|x\rangle$.
2. 设 π 是整数 $\{1, 2, \dots, n\}$ 的一个置换 (或说洗牌, *shuffling*)⁴². 若 $|x\rangle = (\eta_1, \eta_2, \dots, \eta_n)$ 是 \mathbb{C}^n 中的矢量, 我们记

$$\mathbf{A}_\pi|x\rangle = (\eta_{\pi(1)}, \eta_{\pi(2)}, \dots, \eta_{\pi(n)}).$$

这个映射 \mathbf{A}_π 就是个线性算符.

3. 对任意 $|x\rangle \in \mathcal{P}^c[t]$, 其中 $x(t) = \sum_{k=0}^n \alpha_k t^k$. 令 $|y\rangle = \mathbf{D}|x\rangle$, 其中 $|y\rangle$ 的定义为 $y(t) = \sum_{k=1}^n k \alpha_k t^{k-1}$. 这样定义的 \mathbf{D} 就是导数算符 (*derivative operator*), 它也是线性算符.
4. 对任意 $|x\rangle \in \mathcal{P}^c[t]$, 其中 $x(t) = \sum_{k=0}^n \alpha_k t^k$. 令 $|y\rangle = \mathbf{S}|x\rangle$, 其中 $|y\rangle \in \mathcal{P}^c[t]$ 定义为 $y(t) = \sum_{k=0}^n [\alpha_k / (k+1)] t^{k+1}$. 这样定义的 \mathbf{S} 就是积分算符 (*integration operator*), 它是个线性算符.
5. 设 $\mathcal{C}^n[a, b]$ 为区间 $[a, b]$ 上所有前 n 阶导数存在并连续的实值函数之集合. 对任意 $|f\rangle \in \mathcal{C}^n[a, b]$, 定义 $|u\rangle = \mathbf{G}|f\rangle$, 其中 $u(t) = g(t)f(t)$ 且 $g(t)$ 是 $\mathcal{C}^n[a, b]$ 中的固定函数. 那么 \mathbf{G} 就是线性的. 特别地, 乘以 t 这个操作对应的算符 \mathbf{T} (即 t 乘子) 也是线性的.

由『**定义 2.3.2**』可以直接得到下述结论:

结论 2.3.6. 两个线性变换 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$ 和 $\mathbf{U} : \mathcal{V} \rightarrow \mathcal{W}$ 相等的充要条件是对所有 $|a_i\rangle$ 均有 $\mathbf{T}|a_i\rangle = \mathbf{U}|a_i\rangle$, 其中 $\{|a_i\rangle\}$ 是 \mathcal{V} 的某一组基. 因此, 线性变换由其在定义域某个基上的作用唯一确定.

上述结论中的相等就是『**第一章**』中所述集合论意义下映射的相等.

如果矢量空间上还带有内积结构, 那么还可以通过一些更为方便的方法建立起算符的相等关系. 下面的两个定理就包含了这些方法的本质.

定理 2.3.7. 内积空间上的自同态 \mathbf{T} 为零算符 $\mathbf{0}$ 的充要条件是对所有的 $|a\rangle, |b\rangle$ 有 $\langle b|\mathbf{T}|a\rangle := \langle b|\mathbf{T}a\rangle = 0$.

⁴¹译者注: 我们以后将具有这样形式的算符称作乘子 (非标准术语). 当我们说 \mathbf{T} 是个 α 乘子时, 我们就是在说它在任意矢量 $|a\rangle$ 上的作用相当于乘了一个标量 α , 即 $\mathbf{T}|a\rangle = \alpha|a\rangle$.

⁴²译者注: 我们将集合 A 上的任意双射 (特别是有限集上的双射) 称作置换 (*permutation*). 它相当于对 A 中元素的排列顺序做了一个调整.

证. 很明显, 若 $\mathbf{T} = \mathbf{0}$, 则 $\langle b|\mathbf{T}|a\rangle = 0$. 反过来, 若对所有 $|a\rangle, |b\rangle$ 均有 $\langle b|\mathbf{T}|a\rangle = 0$, 则取 $|b\rangle = \mathbf{T}|a\rangle = |\mathbf{T}a\rangle$ 就有

$$\langle \mathbf{T}a|\mathbf{T}a\rangle = 0 \quad \forall|a\rangle.$$

而根据内积的正定性这等价于对所有 $|a\rangle$ 均有 $\mathbf{T}|a\rangle = |0\rangle$, 也就是说 $\mathbf{T} = \mathbf{0}$. \square

定理 2.3.8. 内积空间上的线性算符 \mathbf{T} 为零算符 $\mathbf{0}$ 的充要条件是对所有 $|a\rangle$ 均有 $\langle a|\mathbf{T}|a\rangle = 0$.

证. 很明显, 若 $\mathbf{T} = \mathbf{0}$, 则 $\langle a|\mathbf{T}|a\rangle = 0$. 反过来, 选择矢量 $\alpha|a\rangle + \beta|b\rangle$, 在这个矢量和它的左矢之间插入 \mathbf{T} , 并对所得结果重排以凑出极化恒等式⁴³

$$\begin{aligned} \alpha^*\beta\langle a|\mathbf{T}|b\rangle + \alpha\beta^*\langle b|\mathbf{T}|a\rangle &= \langle \alpha a + \beta b|\mathbf{T}|\alpha a + \beta b\rangle \\ &\quad - |\alpha|^2\langle a|\mathbf{T}|a\rangle - |\beta|^2\langle b|\mathbf{T}|b\rangle. \end{aligned}$$

根据定理的假设条件, 可知上式右边为零. 因此, 如果我们令 $\alpha = \beta = 1$, 就可以得到 $\langle a|\mathbf{T}|b\rangle + \langle b|\mathbf{T}|a\rangle = 0$. 类似地, 取 $\alpha = 1$ 和 $\beta = i$ 就得到 $i\langle a|\mathbf{T}|b\rangle - i\langle b|\mathbf{T}|a\rangle = 0$. 这两个式子联立就可以得到 $\langle a|\mathbf{T}|b\rangle = 0$ 对所有 $|a\rangle, |b\rangle$ 成立. 进而由『定理 2.3.7』可得 $\mathbf{T} = \mathbf{0}$. \square

要证明内积空间上两个算符 \mathbf{U} 和 \mathbf{T} 相等, 我们要么让它们作用在任意矢量上去证明它们给出相同结果, 要么利用上面给出的两个定理去验证 $\mathbf{U} - \mathbf{T}$ 为零算符. 等价地说, 就是证明对所有 $|a\rangle, |b\rangle$ 有 $\langle a|\mathbf{T}|b\rangle = \langle a|\mathbf{U}|b\rangle$ 或者 $\langle a|\mathbf{T}|a\rangle = \langle a|\mathbf{U}|a\rangle$.

⁴³译者注: 此处的极化恒等式和『(2.15) 式』形式有点差异, 但是利用诱导范数的定义可以将二者联系起来. 事实上, 这里所述极化恒等式就是『(2.15) 式』的构造思路. 赋予 α, β 正文证明中所提及的两对特殊值, 并消去交叉项即可由此处的关系式得到『(2.15) 式』.

2.3.1. 线性映射的核

由『[定义 2.3.2](#)』立刻可以得到, \mathcal{V} 中零矢量的像为 \mathcal{W} 中的零矢量⁴⁴. 这对一般的映射来说并不成立, 但它是线性映射的必要条件. 当 \mathcal{V} 中的零矢量映到 \mathcal{W} 的零矢量上时, \mathcal{V} 中的其它矢量也有可能一并变成零矢量. 事实上, 我们有如下定理:

定理 2.3.9 (线性变换的核). 设 $T : \mathcal{V} \rightarrow \mathcal{W}$ 是矢量空间 \mathcal{V} 到 \mathcal{W} 的线性变换. 那么, \mathcal{V} 中满足 $T|v\rangle = |0\rangle$ 的所有矢量 $|v\rangle$ 构成 \mathcal{V} 的一个子空间, 称该子空间为 T 的核 (*kernal*), 也称作其零空间 (*null space*), 记作 $\ker T$.

证. 留作习题. □

核 $\ker T$ 的维数也称作 \mathcal{V} 的零化度 (*nullity*).

下面这个定理的证明同样留作习题.

定理 2.3.10 (线性变换的秩). 线性映射 $T : \mathcal{V} \rightarrow \mathcal{W}$ 的值域 $T(\mathcal{V})$ 是 \mathcal{W} 的子空间, 它的维数称作 T 的秩 (*rank*).

借助上述概念我们可以判断一个线性映射是单射还是满射. 具体判定由下述几个定理构成.

定理 2.3.11 (单射判定). 线性映射 T 为单射的充要条件是 $\ker T = 0$ (我们称其为零核)^a.

^a因为 $\ker T$ 是个集合, 零核就应当写作 $\{|0\rangle\}$, 至少也要写作 $\ker T = |0\rangle$. 然而, 在不至于导致混淆的时候, 我们之后直接记 $\{|0\rangle\} = |0\rangle = 0$.

证. 必要性显然⁴⁵. 至于充分性, 我们假设 $T|a_1\rangle = T|a_2\rangle$, 则因 T 是线性映射, 就得到 $T(|a_1\rangle - |a_2\rangle) = 0$. 由于 $\ker T = 0$, 我们必然有 $|a_1\rangle = |a_2\rangle$. □

定理 2.3.12. 线性等距映射是单射.

证. 设 $T : \mathcal{V} \rightarrow \mathcal{U}$ 为线性等距. 取 $|a\rangle \in \ker T$, 则

$$\langle a|a\rangle = \langle Ta|Ta\rangle = \langle 0|0\rangle = 0.$$

因此 $|a\rangle = 0$. 由『[定理 2.3.11](#)』可知 T 为单射. □

⁴⁴译者注: 考察 T 在 $|0\rangle = |a\rangle - |a\rangle$ 上的作用即可.

⁴⁵译者注: 若为单射, 则不同的像必然对应不同的原像. 因此零矢量的原像只能是零矢量, 从而 $\ker T = 0$.

假设我们从 $\ker \mathbf{T}$ 的一组基出发, 向其中添加足够多的线性无关矢量, 从而得到 \mathcal{V} 的一组基. 不失一般性, 我们假设这组基中前 n 个矢量构成 $\ker \mathbf{T}$ 的一组基. 具体地, 设 $B = \{|a_1\rangle, \dots, |a_N\rangle\}$ 为 \mathcal{V} 的一组基, 而 $B' = \{|a_1\rangle, \dots, |a_n\rangle\}$ 是 $\ker \mathbf{T}$ 的一组基. 此处 $N = \dim \mathcal{V}$ 而 $n = \dim \ker \mathbf{T}$. 可以直接证明 $\{\mathbf{T}|a_{n+1}\rangle, \dots, \mathbf{T}|a_N\rangle\}$ 构成 $\mathbf{T}(\mathcal{V})$ 的一组基⁴⁶. 由此, 我们就得到下面这个结果 (参见本小节末尾).

定理 2.3.13. 设 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$ 为线性映射, 则⁴⁷

$$\dim \mathcal{V} = \dim \ker \mathbf{T} + \dim \mathbf{T}(\mathcal{V}).$$

⁴⁶ 回忆一下, 矢量空间的维数依赖于所选择的标量域. 尽管这里我们处理的是不同的矢量空间, 但是线性映射的定义保证了它们都是同一个标量域上的 (复数或实数), 因此此处维数的概念上不会产生混淆.

上面这个定理称作维数定理 (dimension theorem)⁴⁷. 根据维数公式可知, 单自同态自动是满的, 反过来也成立⁴⁸.

命题 2.3.14. 设 $\mathbf{T} \in \text{End}(\mathcal{V})$, 其中 \mathcal{V} 是有限维矢量空间, 则接下来的三个叙述等价: (a) \mathbf{T} 是双射; (b) \mathbf{T} 是单射; (c) \mathbf{T} 是满射.

很明显, 维数定理仅对有限维矢量空间有用⁴⁹. 特别地, 对于无穷维空间, 无论是单射还是满射, 都无法推出其为双射.

例 2.3.15. 让我们试着求解线性映射 $\mathbf{T} : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ 的核, 该映射的定义为

$$\begin{aligned} \mathbf{T}(x_1, x_2, x_3, x_4) \\ = (2x_1 + x_2 + x_3 - x_4, x_1 + x_2 + 2x_3 + 2x_4, x_1 - x_3 - 3x_4). \end{aligned}$$

⁴⁶ 译者注: 对任意 $|a\rangle \in \mathcal{V}$, 我们总是可以将其在 B 下展开, 但是 $\mathbf{T}|a\rangle$ 在前 n 个基矢下结果为零, 最终只保留了在后 $N - n$ 个基矢上的作用. 由此可知 $\{\mathbf{T}|a_{n+1}\rangle, \dots, \mathbf{T}|a_N\rangle\}$ 张成了 $\mathbf{T}(\mathcal{V})$. 再让其线性组合等于零, 则利用 \mathbf{T} 的线性条件可将像的线性组合转换为原本 $N - n$ 个基矢的线性组合. 结合后面这 $N - n$ 个基矢不在核中, 即可推得这 $N - n$ 个基矢的线性组合为零, 而它们是基矢, 因此对应的系数必然全为零. 由此就证明了线性无关.

⁴⁷ 译者注: 在有些书中称其为秩-零化度定理 (rank-nullity theorem), 因为等式右边的两项恰好分别是线性映射 \mathbf{T} 的零度和秩. 另外维数公式也用于另一场景, 指代 $\dim(\mathcal{U} + \mathcal{V}) = \dim \mathcal{U} + \dim \mathcal{V} - \dim(\mathcal{U} \cap \mathcal{V})$ 这个公式 (当然替换为直和就是这里的结论了), 这也是本章的一个习题. 最后, 此处给出的这个定理也是线性映射同构基本定理 (见『定理 2.3.23』) 的推论 (证明可见『例 2.3.22』).

⁴⁸ 译者注: 若为单自同态, 则立刻有 $\dim \ker \mathbf{T} = 0$, 从而 $\dim \mathbf{T}(\mathcal{V}) = \dim \mathcal{V}$, 而有限维矢量空间 \mathcal{V} 除自己以外的子空间维数严格小于自身维数 (假设相等, 但子空间是真子空间, 那么就存在某个矢量没法用子空间的基线性表示, 从而与子空间基矢线性无关, 于是我们得到了多余线性空间维数个数的线性无关矢量. 这与维数的定义矛盾.), 这说明值域等于 \mathcal{V} . 反过来若 $\dim \mathcal{V} = \dim \mathbf{T}(\mathcal{V})$, 则有 $\dim \ker \mathbf{T} = 0$, 从而 \mathbf{T} 零核, 于是它是单射.

⁴⁹ 译者注: 对无穷维空间而言, 其基是个无穷集 (而且基的定义也有很多种——和矢量空间上附加的结构有关), 而『第一章』提到过无穷集可以和它的子集有相同基数. 因此由 $\dim \mathbf{T}(\mathcal{V}) = \dim \mathcal{V} = \infty$ 没法保证 $\mathbf{T}(\mathcal{V}) = \mathcal{V}$. 这就导致维数定理即便成立, 也无甚大用.

解：需要找到使得 $\mathbf{T}(x_1, x_2, x_3, x_4) = (0, 0, 0)$ 的 (x_1, x_2, x_3, x_4) . 这等同于解方程组

$$2x_1 + x_2 + x_3 - x_4 = 0,$$

$$x_1 + x_2 + 2x_3 + 2x_4 = 0,$$

$$x_1 - x_3 - 3x_4 = 0.$$

它的“解”为 $x_1 = x_3 + 3x_4$, $x_2 = -3x_3 - 5x_4$. 因此, 若 \mathbb{R}^4 中的矢量要落入 $\ker \mathbf{T}$, 则它必然具有如下形式:

$$(x_3 + 3x_4, -3x_3 - 5x_4, x_3, x_4) = x_3(1, -3, 1, 0) + x_4(3, -5, 0, 1),$$

其中 x_3, x_4 是任意实数. 因此, $\ker \mathbf{T}$ 由可以写成 $(1, -3, 1, 0)$ 和 $(3, -5, 0, 1)$ 线性组合的矢量构成, 而这两个矢量线性无关. 因此, $\dim \ker \mathbf{T} = 2$. 接着『定理 2.3.13』就指出 $\dim \mathbf{T}(\mathcal{V}) = 2$, 即 \mathbf{T} 的值域是二维的. 注意到

$$\begin{aligned} \mathbf{T}(x_1, x_2, x_3, x_4) \\ = (2x_1 + x_2 + x_3 - x_4)(1, 0, 1) + (x_1 + x_2 + 2x_3 + 2x_4)(0, 1, -1), \end{aligned}$$

因此 $\mathbf{T}(\mathcal{V})$ 中任意矢量 $\mathbf{T}(x_1, x_2, x_3, x_4)$ 就可以写成两个线性无关矢量 $(1, 0, 1)$ 和 $(0, 1, -1)$ 的线性组合, 这就说明 $\dim \mathbf{T}(\mathcal{V})$ 确实等于 2. ■

2.3.2. 线性同构

在很多时候, 两个矢量空间可能“看上去”不一样, 但是它们(作为矢量空间)实际上是一回事. 举个例子, 复数集 \mathbb{C} 和 \mathbb{R}^2 一样, 都是实数域上的二维矢量空间. 尽管对这两个空间中的矢量我们赋予不同的名字, 但它们具有非常相似的性质. 此处“相似”这个说法更精确的表述就是如下定义:

定义 2.3.16 (线性同构). 若矢量空间 \mathcal{V} 和 \mathcal{W} 之间存在既单且满的线性映射(或者说线性双射) $T : \mathcal{V} \rightarrow \mathcal{W}$, 则称它们同构 (*isomorphic*), 记作 $\mathcal{V} \cong \mathcal{W}$; 并称此处的 T 为同构映射 (*isomorphism*, 也简称为同构)^{*}. 从 \mathcal{V} 到自身的线性双射就称作 \mathcal{V} 的自同构 (*automorphism*). 自同构也称作可逆 (*invertible*) 线性映射. 矢量空间 \mathcal{V} 所有自同构的集合就记作 $GL(\mathcal{V})$.

* 正如我们后面将看到的那样, “同构”这个词会出现在很多代数结构当中. 为了区分它们, 就需要给不同的同构加以修饰. 就当前语境, 我们讨论的就是线性同构 (linear isomorphism). 我们会在必要的时候加入修饰词. 然而, 提到同构时通常都可以通过上下文明确到底是哪一种.

由于等距为单射, 结合『[命题 2.3.14](#)』, 我们可以立刻得到下述结果:

命题 2.3.17. 有限维矢量空间上的等距是该空间的自同构.

实际上, 两个同构的矢量空间是“同一个”矢量空间的不同表现形式. 在前文给出的例子(即 \mathbb{R}^2 和 \mathbb{C})中, 由 $T(x + iy) = (x, y)$ 给出的对应关系 $T : \mathbb{C} \rightarrow \mathbb{R}^2$ 就建立起了这两个矢量空间之间的同构. 需要强调的是, 仅当我们将其视作是矢量空间时, 才能说 \mathbb{C} 和 \mathbb{R}^2 同构. 如果我们超出了矢量空间的结构, 这两个集合就大相径庭了. 比如说, \mathbb{C} 里面的元素有着自然的乘法, 但是 \mathbb{R}^2 没有. 下面的三个定理给出了同构的判据. 其证明比较简单, 故留作习题.

定理 2.3.18. 线性满射 $T : \mathcal{V} \rightarrow \mathcal{W}$ 为同构的充要条件是其零化度等于 0.

定理 2.3.19. 线性单射 $T : \mathcal{V} \rightarrow \mathcal{W}$ 将线性无关的矢量组满映到一个线性无关的矢量组上.

定理 2.3.20. 两个有限维矢量空间同构的充要条件是它们的维数相同.

根据『[定理 2.3.20](#)』, 所有的 N 维实矢量空间都同构于 \mathbb{R}^N , 而所有的 N 维复矢量空间都同构于 \mathbb{C}^N . 因此, 实际上我们只有两个 N 维矢量空间: \mathbb{R}^N 和 \mathbb{C}^N .⁵⁰

⁵⁰译者注: 正是因为这个原因, 很多线性代数教科书会花费百分之八十以上的篇幅去讲述 \mathbb{R}^n 的理论, 而只留很少的篇幅给线性空间的抽象理论——因为有限维线性空间本质上就是 \mathbb{R}^n . 但这并不意味着这种讲法是合理的. 事实上, 这种讲法严重偏离物理学的需求, 还会带来一些不太好的惯性思维.

假设 $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$, \mathbf{T} 是使得 \mathcal{V}_1 不变 (invariant) 的 \mathcal{V} 自同构, 即 $\mathbf{T}(\mathcal{V}_1) = \mathcal{V}_1$. 那么 \mathbf{T} 同样使得 \mathcal{V}_2 不变. 理由如下: 首先注意到若 $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$ 且 $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}'_2$, 那么必有 $\mathcal{V}_2 = \mathcal{V}'_2$. 这是因为若我们将 \mathcal{V}_1 的一组基扩充为 \mathcal{V} 的一组基, 那么新扩充的这些基矢必然同时落入 \mathcal{V}_2 和 \mathcal{V}'_2 , 并张成这两个子空间, 进而二者相同. 现在注意, 由于 $\mathbf{T}(\mathcal{V}) = \mathcal{V}$ 且 $\mathbf{T}(\mathcal{V}_1) = \mathcal{V}_1$, 我们就必须有⁵¹

$$\begin{aligned}\mathcal{V}_1 \oplus \mathcal{V}_2 &= \mathcal{V} = \mathbf{T}(\mathcal{V}) = \mathbf{T}(\mathcal{V}_1 \oplus \mathcal{V}_2) \\ &= \mathbf{T}(\mathcal{V}_1) \oplus \mathbf{T}(\mathcal{V}_2) = \mathcal{V}_1 \oplus \mathbf{T}(\mathcal{V}_2).\end{aligned}$$

进而, 根据前面的论述, 就有 $\mathbf{T}(\mathcal{V}_2) = \mathcal{V}_2$. 我们将这里的分析总结如下⁵²:

命题 2.3.21. 若 $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$, 则使得其中一个被加项不变的自同构也会使得另一个被加项不变.

例 2.3.22 (维数定理的另一证明). 设 $\mathbf{T}, \mathcal{V}, \mathcal{W}$ 如『定理 2.3.13』所述. 现在这样定义一个线性映射 $\mathbf{T}' : \mathcal{V}/\ker \mathbf{T} \rightarrow \mathbf{T}(\mathcal{V})$. 若 $[\![a]\!]$ 是 $|a\rangle$ 的等价类, 则定义 $\mathbf{T}'([\![a]\!]) = \mathbf{T}|a\rangle$. 首先我们需要证明这个映射是良定的, 即若 $[\![a']\!] = [\![a]\!]$, 那么 $\mathbf{T}'([\![a']\!]) = \mathbf{T}'([\![a]\!])$. 而这显然成立, 因为由 $[\![a']\!] = [\![a]\!]$ 可知存在 $|z\rangle \in \ker \mathbf{T}$ 使得 $|a'\rangle = |a\rangle + |z\rangle$. 从而

$$\begin{aligned}\mathbf{T}'([\![a']\!]) &:= \mathbf{T}|a'\rangle = \mathbf{T}(|a\rangle + |z\rangle) \\ &= \mathbf{T}|a\rangle + \underbrace{\mathbf{T}|z\rangle}_{=|0\rangle} = \mathbf{T}|a\rangle = \mathbf{T}'([\![a]\!]).\end{aligned}$$

另外, 很容易看到 \mathbf{T}' 是线性的.

现在我们来证明 \mathbf{T}' 是个同构. 假设 $|x\rangle \in \mathbf{T}(\mathcal{V})$. 那么就存在 $|y\rangle \in \mathcal{V}$ 使得 $|x\rangle = \mathbf{T}|y\rangle = \mathbf{T}'([\![y]\!])$. 这就表明 \mathbf{T}' 是满射. 要证明它还是个单射, 我们令 $\mathbf{T}'([\![y]\!]) = \mathbf{T}'([\![x]\!])$, 从而 $\mathbf{T}|y\rangle = \mathbf{T}|x\rangle$, 也就是说 $\mathbf{T}(|y\rangle - |x\rangle) = 0$. 这就表明 $|y\rangle - |x\rangle \in \ker \mathbf{T}$, 从而 $[\![y]\!] = [\![x]\!]$.

既然 \mathbf{T}' 是同构, 那么自然就有 $\dim(\mathcal{V}/\ker \mathbf{T}) = \dim \mathbf{T}(\mathcal{V})$. 最后利用『(2.2) 式』就能重新得到维数公式.

这个例子实际上证明了如下定理⁵³:

⁵¹译者注: 因为线性单射只会将零矢量变成零矢量, 我们就有 $\mathbf{T}(\mathcal{V} \oplus \mathcal{U}) = \mathbf{T}(\mathcal{V}) \oplus \mathbf{T}(\mathcal{U})$. 具体理由如下: 线性条件保证了 $\mathbf{T}(\mathcal{V} + \mathcal{U}) = \mathbf{T}(\mathcal{V}) + \mathbf{T}(\mathcal{U})$, 是故只需证明后两者交集为零. 现若 $\mathbf{T}(\mathcal{V}) \cap \mathbf{T}(\mathcal{U}) \neq 0$, 则存在非零的 $|v\rangle \in \mathcal{V}$ 和 $|u\rangle \in \mathcal{U}$ 使得 $\mathbf{T}|v\rangle = \mathbf{T}|u\rangle$, 进而 $\mathbf{T}(|v\rangle - |u\rangle) = 0$. 单射条件保证了此时必须有 $|v\rangle = |u\rangle$, 而它们非零, 这就与 $\mathcal{V} \cap \mathcal{U} = 0$ 矛盾了.

⁵²译者注: 这里自同构的条件没法减弱. 因为证明时 $\mathbf{T}(\mathcal{V}) = m\mathcal{V}$ 用到了满射性, $\mathbf{T}(\mathcal{V}_1 \oplus \mathcal{V}_2) = \mathbf{T}(\mathcal{V}_1) \oplus \mathbf{T}(\mathcal{V}_2)$ 则用到了单射性与线性.

⁵³译者注: 这个定理也叫作线性同态基本定理, 或者 (线性) 同构基本定理. 同态基本定理是群论中的重点内容, 以群为底空间的结构大多都继承并强化了这个定理. 此处所述线性同态基本

定理 2.3.23. 设 \mathcal{V} 和 \mathcal{W} 为矢量空间, $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$ 为线性映射. 定义 $\mathbf{T}' : \mathcal{V}/\ker \mathbf{T} \rightarrow \mathbf{T}(\mathcal{V})$ 为 $\mathbf{T}'([\![a]\!]) = \mathbf{T}|a\rangle$, 其中 $[\![a]\!]$ 是 $|a\rangle$ 的等价类. 那么, \mathbf{T}' 就是个良定的同构.

现在设 $\mathcal{U}, \mathcal{V}, \mathcal{W}$ 是复矢量空间. 考察线性映射

$$\mathbf{T} : (\mathcal{U} \oplus \mathcal{V}) \otimes \mathcal{W} \rightarrow (\mathcal{U} \otimes \mathcal{W}) \oplus (\mathcal{V} \otimes \mathcal{W}),$$

其定义为

$$\mathbf{T}((|u\rangle + |v\rangle) \otimes |w\rangle) = |u\rangle \otimes |w\rangle + |v\rangle \otimes |w\rangle.$$

容易验证 \mathbf{T} 是个同构. 因此我们就有

$$(\mathcal{U} \oplus \mathcal{V}) \otimes \mathcal{W} \cong (\mathcal{U} \otimes \mathcal{W}) \oplus (\mathcal{V} \otimes \mathcal{W}). \quad (2.16)$$

译者注: 要验证 \mathbf{T} 为同构可能不是那么容易. 为了方便书写, 我们将 $\mathcal{U} \oplus \mathcal{V}$ 中的元素写作 (u, v) . 于是 \mathbf{T} 的定义就是

$$\mathbf{T}((u, v) \otimes w) = (u \otimes w, v \otimes w).$$

但这个定义实际有问题. 因为 $\mathcal{U} \otimes \mathcal{V}$ 中的元素并不是都能写成 $u \otimes v$ 的形式, 而是这种形式的线性组合. 因此能够覆盖所有 $(\mathcal{U} \oplus \mathcal{V}) \otimes \mathcal{W}$ 元素的定义应当是 (但我们经常习惯于写上面那种, 因为它实际上定义了基上的作用形式, 然后强制将其线性化即可自然得到一个线性映射, 这使得我们后续只需验证非线性的性质即可. 后面很多时候定义线性空间相关的概念时都会采用这种方法, 是故之后不再强调)

$$\mathbf{T}\left(\sum_i (u_i, v_i) \otimes w_i\right) = \left(\sum_i u_i \otimes w_i, \sum_i v_i \otimes w_i\right).$$

设 $x = (u, v) \otimes w$, $y = (u', v') \otimes w'$. 再设 α 是个标量. 接下来我们验证上面定义的 \mathbf{T} 是同构.

- 因为张量积满足 $\alpha(u \otimes v) = (\alpha u) \otimes v = u \otimes (\alpha v)$, 于是 $\alpha x = (u, v) \otimes (\alpha w)$. 进而

$$\begin{aligned} \mathbf{T}(\alpha x) &= \mathbf{T}((u, v) \otimes (\alpha w)) = (u \otimes (\alpha w), v \otimes (\alpha w)) \\ &= (\alpha(u \otimes w), \alpha(v \otimes w)) = \alpha(u \otimes w, v \otimes w) \\ &= \alpha \mathbf{T}(x). \end{aligned}$$

这里用到了直和的定义 $\alpha(u, v) = (\alpha u, \alpha v)$.

定理就是一例. 另外, 原书中将这个命题提升了一下, 将定理中的 $\ker \mathbf{T}$ 更换为任意子空间 \mathcal{U} , 而其他内容不变. 但我认为这个推广有问题, 因为任意子空间的情况下 \mathbf{T}' 不再良定: 若 $[\![a]\!] = [\![b]\!]$, 那么存在 $|u\rangle \in \mathcal{U}$ 使得 $|a\rangle = |b\rangle + |u\rangle$, 从而 $\mathbf{T}|a\rangle$ 和 $\mathbf{T}|b\rangle$ 之间相差一个 $\mathbf{T}|u\rangle$. 这就导致 \mathbf{T}' 依赖于代表元的选法. 换个角度看, 若书中命题成立, 那么我们恒有 $\dim \mathcal{V} - \dim \mathcal{U} = \dim \mathbf{T}(\mathcal{V})$. 然而 $\dim \mathcal{V}$ 和 $\dim \mathbf{T}(\mathcal{V})$ 是个固定的值, 而子空间的维数当然不固定. 这就矛盾了! 因此在翻译时我对该定理做了修改.

- 直和满足 $(x, y) + (x', y') = (x + x', y + y')$, 现在采用我们给出的扩展定义就有

$$\begin{aligned}\mathbf{T}(x+y) &= (u \otimes w + u' \otimes w', v \otimes w + v' \otimes w') \\ &= (u \otimes w, v \otimes w) + (u' \otimes w', v' \otimes w') \\ &= \mathbf{T}(x) + \mathbf{T}(y).\end{aligned}$$

- 接下来我们来证明 \mathbf{T} 是单射. 为此我们需要证明若 $\mathbf{T}(x) = (u \otimes w, v \otimes w) = (0_{U \otimes W}, 0_{V \otimes W})$, 则 x 必为 $(U \oplus V) \otimes W$ 里的零矢量. 若 $w \neq 0_W$, 则我们必然有 $u = 0_U, v = 0_V$. 从而 $x = (0_U, 0_V) \otimes w = 0_{(U \oplus V) \otimes W}$. 若 $w = 0_W$, 则我们就有 $x = (u, v) \otimes 0_W = 0_{(U \oplus V) \otimes W}$.
- 最后证明 \mathbf{T} 是满射. 为此我们考察 $(U \otimes V) \oplus (V \otimes W)$ 中的任意元素 $(\sum_i u_i \otimes w_i, \sum_j v_j \otimes w'_j)$. 我们要给出其原像. 事实上, 注意到

$$\begin{aligned}\mathbf{T}\left(\sum_i (u_i, 0) \otimes w_i + \sum_j (0, v_j) \otimes w'_j\right) \\ &= \left(\sum_i u_i \otimes w_i + \sum_j 0 \otimes w'_j, \sum_i 0 \otimes w_i + \sum_j v_j \otimes w'_j\right) \\ &= \left(\sum_i u_i \otimes w_i, \sum_j v_j \otimes w'_j\right),\end{aligned}$$

这就说明我们开始给出的那个就是其原像.

不难看到, 这个验证的核心就是那个扩展定义, 如果只采用书中的定义方式会在证明加法性质和满射性质时卡住.

根据张量积的维数公式 $\dim(U \otimes V) = \dim U \dim V$, 我们得到

$$U \otimes V \cong V \otimes U. \quad (2.17)$$

不仅如此, 因为 $\dim \mathbb{C} = 1$, 所以 $\dim(\mathbb{C} \otimes V) = \dim V$. 是故

$$\mathbb{C} \otimes V \cong V \otimes \mathbb{C} \cong V. \quad (2.18)$$

类似地, 对于实矢量空间 V 就有

$$\mathbb{R} \otimes V \cong V \otimes \mathbb{R} \cong V. \quad (2.19)$$

2.4. 复结构

到目前为止, 我们处理矢量空间时总是避免改变标量域. 当我们声明一个矢量空间是复矢量空间时, 对应的标量总是复的; 而当我们使用实标量时, 总是要把实数看作是复数的子集.

在这一节, 我们来探究更改标量域的可能性, 并探究在更改过后矢量空间上其它结构对应的改变. 这里我们感兴趣的情形是将标量域从实数域改成复数域.

在讨论变更标量域(以及其它问题的形式化处理)时, 推广内积的概念会有所帮助. 尽管对物理学而言, 内积的正定性至关重要, 但对于其它场景, 这个要求限制性太高. 因此, 我们放松之前对内积的要求, 定义一个新的内积. 我们约定一下:

约定 2.4.1. 除本小节以外, 除非明确声明, 不然所有的复内积空间都如『[定义 2.2.1](#)』那样, 是半双线性的.

定义 2.4.2 (双线性内积). 设 \mathbb{F} 为 \mathbb{C} 或者 \mathbb{R} , \mathcal{V} 是个 \mathbb{F} 线性空间. 我们将满足下述性质的映射 $g : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{F}$ 称作双线性内积 (*bilinear inner product*):

- (a) 对称性 (*symmetry*): $g(|a\rangle, |b\rangle) = g(|b\rangle, |a\rangle);$
- (b) 双线性 (*bilinearity*):

$$\begin{aligned} g(|x\rangle, \alpha|a\rangle + \beta|b\rangle) &= \alpha g(|x\rangle, |a\rangle) + \beta g(|x\rangle, |b\rangle) \\ g(\alpha|a\rangle + \beta|b\rangle, |x\rangle) &= \alpha g(|a\rangle, |x\rangle) + \beta g(|b\rangle, |x\rangle) \end{aligned}$$

- (c) 非退化性 (*nondegeneracy*): 若 $g(|x\rangle, |a\rangle) = 0$ 对所有 $|x\rangle \in \mathcal{V}$ 成立, 则 $|a\rangle = |0\rangle;$

其中 $\alpha, \beta \in \mathbb{F}$, $|a\rangle, |b\rangle, |x\rangle \in \mathcal{V}$.

非退化性亦可这样表述: 对任意非零矢量 $|a\rangle \in \mathcal{V}$, 存在至少一个矢量 $|x\rangle \in \mathcal{V}$ 使得 $g(|x\rangle, |a\rangle) \neq 0$. 它陈述了这样一个事实: 内积空间中与所有矢量都正交的矢量有且仅有零矢量.

和之前一样, 我们采用 Dirac 记号来描述这个内积. 不过, 为了和之前定义的内积区分开来, 我们给它加个下标 \mathbb{F} . 如此一来, 『[定义 2.4.2](#)』中的三条性质就可以表示如下:

- (a) 对称性: $\langle a|b \rangle_{\mathbb{F}} = \langle b|a \rangle_{\mathbb{F}};$
- (b) 双线性: $\langle x|\alpha a + \beta b \rangle_{\mathbb{F}} = \alpha \langle x|a \rangle_{\mathbb{F}} + \beta \langle x|b \rangle_{\mathbb{F}},$
 $\langle \alpha a + \beta b|x \rangle_{\mathbb{F}} = \alpha \langle a|x \rangle_{\mathbb{F}} + \beta \langle b|x \rangle_{\mathbb{F}};$
- (c) 非退化性: $\langle x|a \rangle_{\mathbb{F}} = 0 \quad \forall |x\rangle \in \mathcal{V} \quad \Rightarrow \quad |a\rangle = |0\rangle.$

注意当 $\mathbb{F} = \mathbb{R}$ 时就有 $\langle \bullet | \bullet \rangle_{\mathbb{F}} = \langle \bullet | \bullet \rangle$.⁵⁴

定义 2.4.3 (伴随). 算符 $\mathbf{A} \in \text{End}(\mathcal{V})$ 的伴随算符 (*adjoint operator*) 记作 \mathbf{A}^T , 其由下式定义:

$$\langle \mathbf{A}a | b \rangle_{\mathbb{F}} = \langle a | \mathbf{A}^T b \rangle_{\mathbb{F}},$$

这个条件也可以写作⁵⁵

$$\langle a | \mathbf{A}^T | b \rangle_{\mathbb{F}} = \langle b | \mathbf{A} | a \rangle_{\mathbb{F}}.$$

如若 $\mathbf{A}^T = \mathbf{A}$, 则称它是自伴的 (*self-adjoint*), 若 $\mathbf{A}^T = -\mathbf{A}$, 则称它是斜的 (*skew*).

由定义以及 $\langle \bullet | \bullet \rangle_{\mathbb{F}}$ 的非退化性, 可以得到⁵⁶

$$(\mathbf{A}^T)^T = \mathbf{A}. \quad (2.21)$$

命题 2.4.4. 算符 $\mathbf{A} \in \text{End}(\mathcal{V})$ 为斜算符的充要条件是, 对所有 $|x\rangle \in \mathcal{V}$ 均有 $\langle x | \mathbf{A}x \rangle_{\mathbb{F}} := \langle x | \mathbf{A} | x \rangle_{\mathbb{F}} = 0$.

证. 若 \mathbf{A} 是斜的, 则

$$\langle x | \mathbf{A} | x \rangle_{\mathbb{F}} = \langle x | \mathbf{A}^T | x \rangle_{\mathbb{F}} = -\langle x | \mathbf{A} | x \rangle_{\mathbb{F}}.$$

由此即可推出 $\langle x | \mathbf{A} | x \rangle_{\mathbb{F}} = 0$.

⁵⁴译者注: 注意这里的 $\langle \bullet | \bullet \rangle$ 要去除正定性条件 (也就是我们不强求它是 Euclid 内积). 当然, 这里的等号也可以理解为定义, 即 $\mathbb{F} = \mathbb{R}$ 时我们将推广的内积就定义为原本的内积. 实内积显然是非退化的对称双线性形式 (只需验证非退化性: 若 $\langle x | a \rangle = 0$ 对所有 $|x\rangle \in \mathcal{V}$ 成立, 取 $|x\rangle = |a\rangle$ 就可以得到 $\langle a | a \rangle = 0$, 进而正定性条件保证了此时 $|a\rangle = |0\rangle$). 反例也很经典, 就是狭义相对论中会遇到的 Minkowsky 空间. 我们考察由 $\langle x | y \rangle_{\mathbb{F}} = -x_1 y_1 + x_2 y_2$ 定义的函数 $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$, 其中 $|x\rangle = (x_1, x_2)$, $|y\rangle = (y_1, y_2)$. 很明显这是对称双线性函数, 现在我们验证它非退化: 设 $|a\rangle = (a_1, a_2)$ 且对任意 $|x\rangle = (x_1, x_2)$ 有 $\langle x | a \rangle_{\mathbb{F}} = -x_1 a_1 + x_2 a_2 = 0$. 我们取 $x_1 = -1, x_2 = 0$ 就得到 $a_1 = 0$, 同理可得 $a_2 = 0$. 从而 $|a\rangle = |0\rangle$. 但是它不满足正定性: 对任意 $|x\rangle$ 恒有 $\langle x | x \rangle_{\mathbb{F}} = -x_1^2 + x_2^2$. 在 $|x_1| > |x_2|$ 时为负值, 而且 $|x_1| = |x_2|$ 时它就等于零. 从而正定性的两个条件它都不满足.

⁵⁵译者注: 请注意这里我们用到了对称性, 将 $\mathbf{A}a$ 从左边换到了右边.

⁵⁶译者注: 根据定义就有

$$\langle x | (\mathbf{A}^T)^T | a \rangle_{\mathbb{F}} = \langle a | \mathbf{A}^T | x \rangle_{\mathbb{F}} = \langle x | \mathbf{A} | a \rangle_{\mathbb{F}}$$

对任意 $|a\rangle$ 成立. 或者说有

$$\left\langle x \middle| [(\mathbf{A}^T)^T - \mathbf{A}]a \right\rangle_{\mathbb{F}} = 0 \quad \forall |x\rangle \in \mathcal{V}.$$

如此一来非退化性就指出此处 $[(\mathbf{A}^T)^T - \mathbf{A}]|a\rangle = |0\rangle$, 而 $|a\rangle$ 又是任意的, 于是 $(\mathbf{A}^T)^T - \mathbf{A} = \mathbf{0}$, 这样就得到了『(2.21) 式』.

反过来, 假设对所有 $|x\rangle \in \mathcal{V}$ 均有 $\langle x|\mathbf{A}|x\rangle = 0$, 那么对非零的 $\alpha, \beta \in \mathbb{F}$ 以及非零矢量 $|a\rangle, |b\rangle \in \mathcal{V}$ 就有

$$\begin{aligned} 0 &= \langle \alpha a + \beta b |\mathbf{A}|\alpha a + \beta b \rangle_{\mathbb{F}} \\ &= \underbrace{\alpha^2 \langle a |\mathbf{A}|a \rangle_{\mathbb{F}}}_{=0} + \alpha\beta \langle a |\mathbf{A}|b \rangle_{\mathbb{F}} + \alpha\beta \langle b |\mathbf{A}|a \rangle_{\mathbb{F}} + \underbrace{\beta^2 \langle b |\mathbf{A}|b \rangle_{\mathbb{F}}}_{=0} \\ &= \alpha\beta (\langle b |\mathbf{A}|a \rangle_{\mathbb{F}} + \langle b |\mathbf{A}^T|a \rangle_{\mathbb{F}}). \end{aligned}$$

因为 $\alpha\beta \neq 0$, 我们必然有 $\langle b |(\mathbf{A} + \mathbf{A}^T)|a\rangle_{\mathbb{F}} = 0$ 对所有非零矢量 $|a\rangle, |b\rangle \in \mathcal{V}$ 成立. 接下来利用内积的非退化性即得 $(\mathbf{A} + \mathbf{A}^T)|a\rangle = |0\rangle$. 由于这对所有 $|a\rangle \in \mathcal{V}$ 成立, 从而 $\mathbf{A}^T = -\mathbf{A}$. \square

将这个命题和『定理 2.3.8』比较, 就可以看出正定性对内积施加的限制是何等之强.

定义 2.4.5 (复结构). 实矢量空间 \mathcal{V} 上的复结构 (*complex structure*) 是个线性算符 \mathbf{J} , 它要满足 $\mathbf{J}^2 = -\mathbf{1}$ 且对所有的 $|a\rangle, |b\rangle \in \mathcal{V}$ 有 $\langle \mathbf{J}a | \mathbf{J}b \rangle = \langle a | b \rangle$. 这里 $\mathbf{1}$ 为恒等算符 (*identity operator*; 或称单位算符), 定义为 $\mathbf{1}|x\rangle = |x\rangle, \forall |x\rangle \in \mathcal{V}$.

命题 2.4.6. 复结构 \mathbf{J} 是斜的.

证. 设 $|a\rangle \in \mathcal{V}, |b\rangle \in \mathbf{J}|a\rangle$. 回忆一下有 $\langle \bullet | \bullet \rangle_{\mathbb{R}} = \langle \bullet | \bullet \rangle$. 现在一方面有

$$\langle a | \mathbf{J}a \rangle = \langle a | b \rangle = \langle \mathbf{J}a | \mathbf{J}b \rangle = \langle \mathbf{J}a | \mathbf{J}^2 a \rangle = -\langle \mathbf{J}a | a \rangle.$$

另一方面还有

$$\langle a | \mathbf{J}a \rangle = \langle a | b \rangle = \langle b | a \rangle = \langle \mathbf{J}a | a \rangle.$$

结合这两个结果即可得到 $\langle a | \mathbf{J}a \rangle = 0$ 对所有 $|a\rangle \in \mathcal{V}$ 成立. 因此, 根据『命题 2.4.4』可知 \mathbf{J} 是斜的. \square

现在设 $|a\rangle$ 是 N 维实内积空间中的任意矢量. 将 $|a\rangle$ 归一化以得到单位矢量 $|e_1\rangle$. 由『命题 2.4.4』以及『命题 2.4.6』可知 $\mathbf{J}|e_1\rangle$ 与 $|e_1\rangle$ 正交. 接下来令 $|e_2\rangle = \mathbf{J}|e_1\rangle$. 则 $\langle e_2 | e_2 \rangle = \langle \mathbf{J}e_1 | \mathbf{J}e_1 \rangle = \langle e_1 | e_1 \rangle = 1$, 从而 $|e_2\rangle$ 也是单位矢量. 若 $N > 2$, 设 $|e_3\rangle$ 为和 $|e_1\rangle, |e_2\rangle$ 均正交的任意单位矢量. 接着 $|e_4\rangle := \mathbf{J}|e_3\rangle$ 显然与 $|e_3\rangle$ 正交. 我们进一步断言说它和 $|e_1\rangle, |e_2\rangle$ 也都正交:

$$\begin{aligned} \langle e_1 | e_4 \rangle &= \langle \mathbf{J}e_1 | \mathbf{J}e_4 \rangle = \langle \mathbf{J}e_1 | \mathbf{J}^2 e_3 \rangle \\ &= -\langle \mathbf{J}e_1 | e_3 \rangle = -\langle e_2 | e_3 \rangle = 0, \\ \langle e_2 | e_4 \rangle &= \langle \mathbf{J}e_2 | \mathbf{J}e_4 \rangle = \langle e_1 | e_3 \rangle = 0. \end{aligned}$$

重复这个过程⁵⁷, 我们就可以得到下述结论:

⁵⁷译者注: 注意每当我们得到 $|e\rangle_{2k}$ 后, 就没法通过前面已有的矢量得到新的正交矢量. 比如

定理 2.4.7. 前面构造出的矢量 $\{|e_i\rangle, \mathbf{J}|e_i\rangle\}_{i=1}^m$ 构成 N 维实矢量空间 \mathcal{V} 在内积 $\langle \bullet | \bullet \rangle_{\mathbb{R}} = \langle \bullet | \bullet \rangle$ 下的一组正交归一基, 这里 $N = 2m$. 进而, 只有偶数维的 \mathcal{V} 才会有复结构 \mathbf{J} .

译者注: 我们看一下 \mathbb{R}^2 上的复结构. 我们知道, \mathbb{R}^2 和 \mathbb{C} 之间存在下述对应关系:

$$(a, b) \longleftrightarrow a + ib.$$

而众所周知, i 就是复结构的核心所在. 我们现在让 i 作用在 $a + ib$ 上, 其结果为 $-b + ia$. 现在我们将其转回到 \mathbb{R}^2 . 由此我们看到 i 诱导出映射 $\mathbf{J}: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 如下:

$$\mathbf{J}: (a, b) \rightarrow (-b, a).$$

这个映射就是 \mathbb{R}^2 上的复结构. 因为很明显

$$\mathbf{J}^2(a, b) = \mathbf{J}(-b, a) = (-a, -b) = -(a, b).$$

于是 $\mathbf{J}^2 = -1$. 另一方面,

$$\langle \mathbf{J}(a, b) | \mathbf{J}(x, y) \rangle = \langle (-b, a) | (-y, x) \rangle = by + ax = \langle (a, b) | (x, y) \rangle.$$

定义 2.4.8 (复化). 设 \mathcal{V} 是实矢量空间, 在张量积空间 $\mathbb{C} \otimes \mathcal{V}$ 上装备复数乘法规则⁵⁸:

$$\alpha(\beta \otimes |a\rangle) = (\alpha\beta) \otimes |a\rangle, \quad \alpha, \beta \in \mathbb{C}.$$

所得空间称作 \mathcal{V} 的复化 (*complexification*), 记作 $\mathcal{V}^{\mathbb{C}}$. 特别地, $(\mathbb{R}^n)^{\mathbb{C}} := \mathbb{C} \otimes \mathbb{R}^n \cong \mathbb{C}^n$.

注意, $\dim_{\mathbb{C}} \mathcal{V}^{\mathbb{C}} = \dim_{\mathbb{R}} \mathcal{V}$ 且 $\dim_{\mathbb{R}} \mathcal{V}^{\mathbb{C}} = 2 \dim_{\mathbb{R}} \mathcal{V}$. 事实上, 若 $\{|a_k\rangle\}_{k=1}^N$ 是 \mathcal{V} 的基, 则它亦是 $\mathcal{V}^{\mathbb{C}}$ 作为复矢量空间时的一组基, 而 $\{|a_k\rangle, i|a_k\rangle\}_{k=1}^N$ 则是其作为实矢量空间时的一组基⁵⁹.

在将带内积 $\langle \bullet | \bullet \rangle_{\mathbb{R}} = \langle \bullet | \bullet \rangle$ 的实矢量空间 \mathcal{V} 复化之后, 我们就可以在它上面定义一个半双线性 (或者 Hermite 的) 的内积如下:

$$\langle \alpha \otimes a | \beta \otimes b \rangle := \alpha^* \beta \langle a | b \rangle$$

说, $\mathbf{J}|e_2\rangle = \mathbf{J}^2|e_1\rangle = -|e_1\rangle$ 就和 $|e_1\rangle$ 成比例. 于是就必须再降一个新的正交矢量, 没法再降的时候这个流程就终止了. 由于每次得到的 $|e_k\rangle, \mathbf{J}|e_k\rangle$ 都和前面已有的矢量正交, 因此这个过程只会在已有偶数个正交归一矢量的基础上终止. 这就是下面这个定理中要求 $N = 2m$ 的理由.

⁵⁸译者注: 请注意, 这个空间中元素的一般形式为 $\sum_i \alpha_i \otimes |v_i\rangle$, 而不是 $\alpha \otimes |v\rangle$. 下面的这个乘法规则也要对应扩展到一般形式才可以, 即得有 $\alpha(\sum_i \alpha_i \otimes |a_i\rangle) = \sum_i (\alpha \alpha_i) \otimes |a_i\rangle$.

⁵⁹译者注: 请注意, $\mathcal{V}^{\mathbb{C}}$ 作为复矢量空间的基实际上指的是 $\{1 \otimes |a_k\rangle\}_{k=1}^N$ (见『2.1.4节』的论断), 但是习惯上我们会略去此处的张量积符号. 实矢量空间情形也是类似的.

请读者验证这样定义的内积确实满足『[定义 2.2.1](#)』中所有的性质.

要复化一个实矢量空间 \mathcal{V} , 我们就必须将其与复数集“相乘”: $\mathcal{V}^{\mathbb{C}} = \mathbb{C} \otimes \mathcal{V}$. 这样的结果就是给出一个两倍于原有维数的实矢量空间. 那么这个过程能否反过来, 也就是对某个实矢量空间(当然必须是偶数维的)做“除法”呢? 也就是说, 从一个偶数维的实矢量空间出发, 能否得到一个维数为原来一半的复矢量空间?

设 \mathcal{V} 是 $2m$ 维实矢量空间, \mathbf{J} 是 \mathcal{V} 上的一个复结构. 紧接着我们进一步设 $\{|e_i\rangle, \mathbf{J}|e_i\rangle\}_{i=1}^m$ 是『[定理 2.4.7](#)』构造出的那组正交归一基. 在 \mathcal{V} 的子空间 $\mathcal{V}_1 := \text{Span}\{|e_i\rangle\}_{i=1}^m$ 上, 我们定义其与复数的乘法为

$$(\alpha + i\beta) \otimes |v_1\rangle := (\alpha \mathbf{1} + \beta \mathbf{J})|v_1\rangle, \quad \alpha, \beta \in \mathbb{R}, |v_1\rangle \in \mathcal{V}_1. \quad (2.22)$$

可以很直接地证明这一过程就将 $2m$ 维的实矢量空间 \mathcal{V} 变成一个 m 维的复矢量空间 $\mathcal{V}_1^{\mathbb{C}}$.

2.5. 线性泛函

当线性变换的陪域恰好是对应标量域 (\mathbb{C} 或 \mathbb{R}) 时, 我们就称它是个线性泛函 (linear functional), 这是很重要的一类线性变换. 线性泛函的集合 $\mathcal{L}(\mathcal{V}, \mathbb{C})$ (若 \mathcal{V} 是复矢量空间) 或 $\mathcal{L}(\mathcal{V}, \mathbb{R})$ (若 \mathcal{V} 是实矢量空间) 称作 \mathcal{V} 的对偶空间 (dual space), 记作 \mathcal{V}^* .

例 2.5.1. 这里是一些线性泛函的例子.

(a) 设 $|a\rangle = (\alpha_1, \alpha_2, \dots, \alpha_n)$ 属于 \mathbb{C}^n . 定义 $\phi : \mathbb{C}^n \rightarrow \mathbb{C}$ 如下:

$$\phi(|a\rangle) = \sum_{k=1}^n \alpha_k.$$

容易验证 ϕ 确实是个线性泛函.

(b) 用 μ_{ij} 表示 $m \times n$ 矩阵 M 的矩阵元. 定义 $\omega : M^{m \times n} \rightarrow \mathbb{C}$ 如下:

$$\omega(M) = \sum_{i=1}^m \sum_{j=1}^n \mu_{ij}.$$

同样容易验证 ω 是个线性泛函.

(c) 将 $n \times n$ 矩阵 M 的矩阵元记作 μ_{ij} . 定义 $\theta : M^{n \times n} \rightarrow \mathbb{C}$ 如下:

$$\theta(M) = \sum_{j=1}^n \mu_{jj},$$

即矩阵对角元之和. 同样例行可证 θ 是个线性泛函.

(d) 定义算符 $\text{int} : C^0(a, b) \rightarrow \mathbb{R}$ 如下:

$$\text{int}(f) = \int_a^b f(t) dt.$$

那么 int 就是矢量空间 $C^0(a, b)$ 上的线性泛函.

(e) 设 \mathcal{V} 是个复内积空间. 固定一个 $|a\rangle \in \mathcal{V}$, 定义 $\gamma_a : \mathcal{V} \rightarrow \mathbb{C}$ 如下:

$$\gamma_a(|b\rangle) = \langle a|b \rangle.$$

读者可以证明 γ_a 是个线性泛函.

(f) 设 $\{|a_1\rangle, |a_2\rangle, \dots, |a_m\rangle\}$ 是 \mathcal{V} 中任意有限个矢量组成的集合, 而 $\{\phi_1, \phi_2, \dots, \phi_m\}$ 是 \mathcal{V} 上线性泛函组成的任意集合. 现在定义

$$\mathbf{A} := \sum_{k=1}^m |a_k\rangle \phi_k \in \text{End}(\mathcal{V})$$

如下:

$$\mathbf{A}|x\rangle = \sum_{k=1}^m |a_k\rangle \phi_k(|x\rangle) = \sum_{k=1}^m \phi_k(|x\rangle) |a_k\rangle.$$

则 \mathbf{A} 是 \mathcal{V} 上的线性算符.

在矢量空间与它的对偶空间之间有个线性同构的例子, 这就是接下来我们要讨论的内容. 考察一个 N 维矢量空间, 它的一组基为 $B = \{|a_1\rangle, |a_2\rangle, \dots, |a_N\rangle\}$. 对于任意 N 个给定标量 $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$, 定义线性泛函 ϕ_α 为 $\phi_\alpha|a_i\rangle = \alpha_i$. 当 ϕ_α 作用在 \mathcal{V} 中任意矢量 $|b\rangle = \sum_{i=1}^N \beta_i |a_i\rangle$ 上时, 其结果就是

$$\phi_\alpha|b\rangle = \phi_\alpha\left(\sum_{i=1}^N \beta_i |a_i\rangle\right) = \sum_{i=1}^N \beta_i \phi_\alpha|a_i\rangle = \sum_{i=1}^N \beta_i \alpha_i. \quad (2.23)$$

这个表达式指出 $|b\rangle$ 可以表示为以 $\beta_1, \beta_2, \dots, \beta_N$ 为元素的列矢量, 而将 ϕ_α 表示为以 $\alpha_1, \alpha_2, \dots, \alpha_N$ 为元素的行矢量. 如此一来, $\phi_\alpha|b\rangle$ 不过就是行矢量(在左边)与列矢量(在右边)的矩阵乘法*.

线性泛函 ϕ_α 由集合 $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$ 唯一确定. 换言之, 对每个这样的标量集合(由 N 个标量组成), 都存在唯一一个与之对应的线性泛函. 而这就将我们引向一组特殊的泛函 $\phi_1, \phi_2, \dots, \phi_N$; 它们分别对应 $\{1, 0, 0, \dots, 0\}, \{0, 1, 0, \dots, 0\}, \dots, \{0, 0, 0, \dots, 1\}$.⁶⁰ 具体而言, 就是有

$$\phi_1|a_1\rangle = 1 \quad \text{且若 } j \neq 1 \text{ 则 } \phi_1|a_j\rangle = 0,$$

$$\phi_2|a_2\rangle = 1 \quad \text{且若 } j \neq 2 \text{ 则 } \phi_2|a_j\rangle = 0,$$

⋮

$$\phi_N|a_N\rangle = 1 \quad \text{且若 } j \neq N \text{ 则 } \phi_N|a_j\rangle = 0,$$

亦或者紧凑地写作

$$\phi_i|a_j\rangle = \delta_{ij}, \quad (2.24)$$

其中 δ_{ij} 是 Kronecker delta.

由『(2.24) 式』定义的这些泛函就构成了对偶空间 \mathcal{V}^* 的一组基. 如欲证之, 则可取任意泛函 $\gamma \in \mathcal{V}^*$, 它由其在基 $B = \{|a_i\rangle\}_{i=1}^N$ 上的作用唯一确定. 现设 $\gamma|a_i\rangle = \gamma_i \in \mathbb{C}$. 则我们断言有 $\gamma = \sum_{i=1}^N \gamma_i \phi_i$. 事实上, 考虑 \mathcal{V} 中任意矢量 $|a\rangle$, 设其关于基 B 的分量为 $(\alpha_1, \alpha_2, \dots, \alpha_N)$. 那么, 一方面有

$$\gamma|a\rangle = \gamma\left(\sum_{i=1}^N \alpha_i |a_i\rangle\right) = \sum_{i=1}^N \alpha_i \gamma|a_i\rangle = \sum_{i=1}^N \alpha_i \gamma_i.$$

另一方面又有

$$\begin{aligned} \left(\sum_{i=1}^N \gamma_i \phi_i\right) &= \left(\sum_{i=1}^N \gamma_i \phi_i\right) \left(\sum_{j=1}^N \alpha_j |a_j\rangle\right) \\ &= \sum_{i=1}^N \gamma_i \sum_{j=1}^N \alpha_j \phi_j|a_j\rangle = \sum_{i=1}^N \gamma_i \sum_{j=1}^N \alpha_j \delta_{ij} = \sum_{i=1}^N \gamma_i \alpha_i. \end{aligned}$$

⁶⁰译者注: 读者可能会发现, 集合论语境下这几个集合都是同一个集合 $\{0, 1\}$, 因为重复元素不算数. 不过此处我们应当将其理解为一个有限序列而非普通的集合, 只是用了集合的表示符号罢了.

既然 γ 和 $\sum_{i=1}^N \gamma_i \phi_i$ 在任意 $|a\rangle$ 上的作用结果一致, 它们自然相等, 即有 $\gamma = \sum_{i=1}^N \gamma_i \phi_i$. 换言之, $\{\phi_i\}_{i=1}^N$ 就张成了 \mathcal{V}^* . 至此, 我们就得到了下述结果:

定理 2.5.2. 对 \mathcal{V} 的任意一组基 $B = \{|a_j\rangle\}_{j=1}^N$, 在 \mathcal{V}^* 中都对应了唯一的一组基 $B^* = \{\phi_i\}_{i=1}^N$ 使得 $\phi_i|a_j\rangle = \delta_{ij}$.

根据这一定理, N 维矢量空间的对偶空间同样是 N 维的, 进而二者同构. 由『(2.24) 式』定义的这组基 B^* 称作 B 的对偶基 (dual basis). 由『定理 2.5.2』可以得到这样一个推论: \mathcal{V} 中的每个矢量都唯一对应了 \mathcal{V}^* 中的一个线性泛函. 如欲证之, 则只需注意到每个矢量 $|a\rangle$ 均由其在某一组基 B 下的分量 $(\alpha_1, \alpha_2, \dots, \alpha_N)$ 唯一确定. 然后与 $|a\rangle$ 对应的那个唯一的线性泛函 ϕ_a 就是 $\sum_{i=1}^N \alpha_i \phi_i$, 其中 $\phi_i \in B^*$. 这个对应的线性泛函通常称作 $|a\rangle$ 的对偶 (矢量).

定义 2.5.3 (零化子). 矢量 $|a\rangle \in \mathcal{V}$ 的零化子 (annihilator) 是个线性泛函 $\phi \in \mathcal{V}^*$, 它要满足 $\phi|a\rangle = 0$.⁶¹ 设 \mathcal{W} 为 \mathcal{V} 的子空间, 对偶空间 \mathcal{V}^* 中所有能够零化 \mathcal{W} 中所有矢量的线性泛函构成的集合记作 \mathcal{W}^0 .

可以验证, \mathcal{W}^0 是 \mathcal{V}^* 的子空间. 不仅如此, 如果我们将 \mathcal{W} 的一组基 $\{|a_i\rangle\}_{i=1}^k$ 扩充为 \mathcal{V} 的一组基 $B = \{|a_i\rangle\}_{i=1}^N$. 那么, 从 B 的对偶基 $B^* = \{\phi_j\}_{j=1}^N$ 中选出的泛函 $\{\phi_j\}_{j=k+1}^N$ 就张成了 \mathcal{W}^0 .⁶² 由此就有

$$\dim \mathcal{V} = \dim \mathcal{W} + \dim \mathcal{W}^0. \quad (2.25)$$

之后讨论辛几何的时候我们就会用到零化子.

至此, 我们已经讨论了矢量的对偶、基的对偶、矢量空间的对偶. 还有一个东西的对偶还未曾提及, 那就是线性变换的对偶.

定义 2.5.4 (线性变换的对偶). 设 $T : \mathcal{V} \rightarrow \mathcal{U}$ 为线性映射, 定义 $T^* : \mathcal{U}^* \rightarrow \mathcal{V}^*$ 如下^{*}:

$$[T^*(\gamma)]|a\rangle = \gamma(T|a\rangle) \quad \forall |a\rangle \in \mathcal{V}, \gamma \in \mathcal{U}^*.$$

这样得到的 T^* 就称作 T 的对偶 (dual), 亦称作其拉回 (pullback).

⁶¹译者注: 此时称 ϕ 零化 $|a\rangle$. 这一说法可以推广到 \mathcal{V} 的任意子集上 (即零化子集中的所有矢量), 讨论能零化该子集的泛函 (即该子集的零化子).

⁶²译者注: 注意到若 ϕ 和 ϕ' 都能零化 \mathcal{W} , 则对任意 $|w\rangle \in \mathcal{W}$ 以及标量 α, β 必有 $(\alpha\phi + \beta\phi')|w\rangle = 0$, 这就说明 \mathcal{W}^0 对加法和标量乘法封闭, 从而是 \mathcal{V}^* 的子空间. 进一步, 由对偶基的定义可知 $\phi_i|a_j\rangle = \delta_{ij}$, 因此 $k+1 \leq i \leq N$ 时必有 $\phi_i|a_j\rangle = 0$ 对所有 $1 \leq j \leq k$ 成立, 进而这些 ϕ_i 就零化了 \mathcal{W} . 从而 $\text{Span}\{\phi_i\}_{i=k+1}^N \subseteq \mathcal{W}^0$. 另一方面, 任取 \mathcal{W} 的零化子 γ , 我们将其在 ϕ_i 下展开为 $\gamma = \sum_{i=1}^N \gamma_i \phi_i$. 现将其作用于 \mathcal{W} 的基矢 $\{|a_i\rangle\}_{i=1}^k$ 就有 $\gamma|a_i\rangle = \sum_{j=1}^N \gamma_j \delta_{ij} = \gamma_i$. 而 γ 是零化子, 它作用在 \mathcal{W} 上必须为零, 它的基矢自然也不例外, 于是 $\gamma_i = 0$ 对所有 $1 \leq i \leq k$ 成立. 这就说明 $\gamma \in \text{Span}\{\phi_i\}_{i=k+1}^N$. 结合两个方向的包含关系, 就得到 $\mathcal{W}^0 = \text{Span}\{\phi_i\}_{i=k+1}^N$.

* 不要将这里的“*”和复共轭混淆.

读者很快就能验证 $\mathbf{T}^* \in \mathcal{L}(\mathcal{U}^*, \mathcal{V}^*)$, 即 \mathbf{T}^* 是 \mathcal{U}^* 到 \mathcal{V}^* 的线性变换.⁶³

正如定义暗示的, \mathbf{T}^* 的性质会和 \mathbf{T} 密切相关. 我们先考虑 \mathbf{T}^* 的核. 很明显, $\gamma \in \ker \mathbf{T}^*$ 的充要条件是它能零化所有具有 $\mathbf{T}|a\rangle$ 形式的矢量, 即它得零化 $\mathbf{T}(\mathcal{V})$.⁶⁴ 由此可知 $\gamma \in \mathbf{T}(\mathcal{V})^*$. 特别地, 若 \mathbf{T} 是满射, 那么 $\mathbf{T}(\mathcal{V}) = \mathcal{U}$, 从而 γ 就零化 \mathcal{U} 中的所有矢量, 也就是说它就是零线性泛函. 这就说明此时 $\ker \mathbf{T}^* = 0$, 于是 \mathbf{T}^* 就是单射. 类似地, 读者可以证明若 \mathbf{T} 是单的, 则 \mathbf{T}^* 就是满的.⁶⁵ 我们将此处的讨论总结为下述结果:

命题 2.5.5. 设 \mathbf{T} 为线性变换, \mathbf{T}^* 是其拉回, 则 $\ker \mathbf{T}^* = \mathbf{T}(\mathcal{V})^0$. 此外, 若 \mathbf{T} 是满射(单射), 则 \mathbf{T}^* 就是单射(满射). 特别地, \mathbf{T} 是同构时 \mathbf{T}^* 也是同构.

在内积和泛函之间建立起联系会有助于我们. 为此, 我们考虑一组基 $\{|a_i\rangle\}_{i=1}^N$, 并令 $\alpha_i = \langle a|a_i\rangle$. 而如前文注意到的那样, 标量集合 $\{\alpha_i\}_{i=1}^N$ 唯一地定义了一个满足 $\gamma_a|a_i\rangle = \alpha_i$ 的线性泛函 γ_a (见『例 2.5.1』). 由于 $\langle a|a_i\rangle$ 同样等于 α_i , 就有一种自然的看法: 将 γ_a 等同于符号 $\langle a|$, 并记该等同为 $\gamma_a \mapsto \langle a|$.

⁶³译者注: 只需注意到

$$\begin{aligned} [\mathbf{T}^*(\alpha\gamma + \beta\phi)]|a\rangle &= (\alpha\gamma + \beta\phi)(\mathbf{T}|a\rangle) \\ &= \alpha\gamma(\mathbf{T}|a\rangle) + \beta\phi(\mathbf{T}|a\rangle) \\ &= \alpha[\mathbf{T}^*(\gamma)]|a\rangle + \beta[\mathbf{T}^*(\phi)]|a\rangle \end{aligned}$$

对所有矢量 $|a\rangle$ 以及标量 α, β 成立即可. 另外原书中这里将 \mathbf{T}^* 表述为 \mathcal{U}^* 上的线性算符, 但这与之前对算符的定义不一致(要求陪域和定义域一样——至少得是子集关系), 因此翻译时做了修改.

⁶⁴译者注: 若 $\gamma \in \ker \mathbf{T}^*$, 则有 $\mathbf{T}^*\gamma = 0$. 按定义它等价于说 $[\mathbf{T}^*(\gamma)]|a\rangle = \gamma(\mathbf{T}|a\rangle) = 0$.

⁶⁵译者注: 书中所述“类似可证”需要用到一个和矢量的零化子对偶的概念: 泛函的零化子. 设 ϕ 为线性泛函, 若 $|a\rangle$ 满足 $\phi|a\rangle = 0$, 则称此时 $|a\rangle$ 零化 ϕ . 对应地, 对于 \mathcal{V}^* 的任意子空间 \mathcal{W}^* , 我们规定 $(\mathcal{W}^*)^0$ 为能够零化 \mathcal{W}^* 的矢量构成的集合. 有了这个概念之后, 我们就能证明对 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{U}$ 有 $\ker \mathbf{T} = (\mathbf{T}^*(\mathcal{U}^*))^0$. 这是因为

$$\begin{aligned} (\mathbf{T}^*(\mathcal{U}^*))^0 &= \{|a\rangle \in \mathcal{V} \mid [\mathbf{T}^*(\gamma)]|a\rangle = 0 \quad \forall \gamma \in \mathcal{U}^*\} \\ &= \{|a\rangle \in \mathcal{V} \mid \gamma(\mathbf{T}|a\rangle) = 0 \quad \forall \gamma \in \mathcal{U}^*\} \\ &= \{|a\rangle \in \mathcal{V} \mid \mathbf{T}|a\rangle = 0\} = \ker \mathbf{T}. \end{aligned}$$

这里我们要用到一个结论: 若对任意 $\phi \in \mathcal{V}^*$ 均有 $\phi|a\rangle = 0$, 则 $|a\rangle = |0\rangle$. 理由也很简单, 我们令 $|a\rangle = \sum_{i=1}^N \alpha_i|a_i\rangle$, 其中 $B = \{|a_i\rangle\}_{i=1}^N$ 是 \mathcal{V} 的一组基. 现在取 B 的对偶基 $B^* = \{\phi_j\}_{j=1}^N$. 于是 $\phi_j|a\rangle = \sum_{i=1}^N \alpha_i\phi_j|a_i\rangle = \alpha_j = 0$. 这就说明了 $|a\rangle = |0\rangle$. 接下来就和文中操作类似了: 若 \mathbf{T} 是单射, 则 $\ker \mathbf{T} = 0$, 进而 $(\mathbf{T}^*(\mathcal{U}^*))^* = 0$. 这表明, 不存在非零矢量可以零化所有 $\mathbf{T}^*(\mathcal{U}^*)$ 中的泛函. 现在假设 $\mathbf{T}^*(\mathcal{U}^*) \neq \mathcal{V}^*$, $\{\psi_i\}_{i=1}^k$ 是 $\mathbf{T}^*(\mathcal{U}^*)$ 的一组基, 将其扩充为 \mathcal{V}^* 的一组基 $\{\psi_i\}_{i=1}^N$. 现在我们取 $\{|b_i\rangle\}_{i=1}^N$ 使得 $\psi_i|b_j\rangle = \delta_{ij}$. 我们断言这样的 $|b_j\rangle$ 一定存在. 事实上, 我们可以取 $\{|a_i\rangle\}_{i=1}^N$ 为 \mathcal{V} 的任意一组基, 然后让 $\{\phi_i\}_{i=1}^N$ 为其对偶基. 由于 $\mathbf{T}^*(\mathcal{U}^*)$ 是 \mathcal{V}^* 的子空间, 所以任意的 ϕ_i 要么落入其中, 要么不属于它. 不失一般性, 可以设 $1 \leq i \leq k$ 对应的 ϕ_i 落入 $\mathbf{T}^*(\mathcal{U}^*)$, 并将它们设定为 ψ_i . 然后不落入其中的则规定为扩充的那些基矢. 现在注意, 只要我们取 $k+1 \leq j \leq N$, 则必有 $\psi_i|b_j\rangle = 0$ 对所有 $1 \leq i \leq k$ 成立, 从而这些 $|b_j\rangle$ 就会零化 $\mathbf{T}^*(\mathcal{U}^*)$, 然而它们是基矢, 是非零的, 这就得到了矛盾.

此外, 引入下述记号^{*}也会带来方便:

^{*} 这个记号的重要性会在『??节』得以明晰.

$$(|a\rangle)^\dagger := \langle a|, \quad (2.26)$$

这里符号 \dagger (剑标) 的意思是“某某的对偶”, 读作 dagger. 现在我们发问: dagger 这个运算作用在矢量的线性组合上会怎样? 为此, 设 $|c\rangle = \alpha|a\rangle + \beta|b\rangle$, 并取 $|c\rangle$ 与任意矢量 $|x\rangle$ 的内积, 利用内积对第二个因子线性的性质, 就有 $\langle x|c\rangle = \alpha\langle x|a\rangle + \beta\langle x|b\rangle$. 对其两边同时取复共轭, 并利用内积的(半双)对称性 (sesquisymmetry) 就有

$$\begin{aligned} (\text{LHS})^* &= \langle x|c\rangle^* = \langle c|x\rangle, \\ (\text{RHS})^* &= \alpha^*\langle x|a\rangle^* + \beta^*\langle x|b\rangle^* = \alpha^*\langle a|x\rangle + \beta^*\langle b|x\rangle \\ &= (\alpha^*\langle a| + \beta^*\langle b|)|x\rangle. \end{aligned}$$

由于其对所有 $|x\rangle$ 均成立, 我们必然就有 $(|c\rangle)^\dagger := \langle c| = \alpha^*\langle a| + \beta^*\langle b|$. 因此, 在 dagger 这个“运算”下, 复标量必须取复共轭. 于是, 我们就有

$$(\alpha|a\rangle + \beta|b\rangle)^\dagger = \alpha^*\langle a| + \beta^*\langle b|. \quad (2.27)$$

职是之故, 不像 $|a\rangle \mapsto \phi_a$ 这个对应关系是线性的⁶⁶, 上面给出的对应 $\gamma_a \mapsto \langle a|$ 不是线性的, 而是半线性的 (sesquilinear)⁶⁷:

$$\gamma_{\alpha a+\beta b} \mapsto \alpha^*\langle a| + \beta^*\langle b|.$$

最后, 我们可以采用这样一种方便法门: 将 $|a\rangle \in \mathbb{C}^n$ 表述为列矢量

$$|a\rangle = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

那么, 复内积的定义就指出 $|a\rangle$ 的对偶必须表示成行矢量, 并且其元素要取复共轭:

$$\langle a| = \begin{bmatrix} \alpha_1^* & \alpha_2^* & \cdots & \alpha_n^* \end{bmatrix}. \quad (2.28)$$

而内积现在就可以表述为下述(矩阵)乘法:^{*}

⁶⁶译者注: 这里 ϕ_a 的定义当参见『定理 2.5.2』下方, 其效果等同于提取 $|a\rangle$ 在基 $\{|a_i\rangle\}_{i=1}^N$ 下的分量, 而 $|a\rangle$ 线性组合的分量自然就是分量的线性组合. 原书中此处写的是 $|a\rangle \mapsto \gamma_a$, 但是如果将 γ_a 理解为满足 $\langle a|a_i\rangle = \gamma_a|a_i\rangle$ 的泛函, 那么很明显它不是线性的:

$$\gamma_{\lambda a}|a_i\rangle = \langle \lambda a|a_i\rangle = \lambda^* \langle a|a_i\rangle = \lambda^* \gamma_a|a_i\rangle \Rightarrow \gamma_{\lambda a} = \lambda^* \gamma_a.$$

因此在翻译时我作了修改.

⁶⁷译者注: 最开始我将 sesquilinear 这个词翻译为半双线性, 因为它描述的是具有两个因子的

^{*} 现在将『(2.28) 式』与『(2.23) 式』下的说明作比较. 『(2.28) 式』中的复共轭是对应关系 $|a\rangle \leftrightarrow \langle a|$ 半线性的结果.

$$\langle a|b\rangle = \begin{bmatrix} \alpha_1^* & \alpha_2^* & \dots & \alpha_n^* \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = \sum_{i=1}^n \alpha_i^* \beta_i.$$

内积, 但这里这个映射却只有一个因子, 因此翻译时我去掉了副词“双”. 需要说明的是, 一些书将其称作共轭线性, 前面所述半双对称性 (sesquisymmetry) 则称作共轭对称性, 之前的译者注中我用到的就是这套术语. 它的好处就是含义直观. 但还有一些书将这组概念称作反线性和反对称性——尽管反对称性有着另一种常见的解释 (交换位置反号; 当然此时相关书籍会用“斜 (skew)” 这个术语来指代交换位置反号).

2.6. 多重线性映射

线性泛函有个非常有用的推广, 它在本书后面处理张量时非常重要. 不过, 在讨论行列式时, 我们会用到其应用的一个受限版本, 而我们就从这里开始.

定义 2.6.1 (p 线性映射). 设 \mathcal{V} 和 \mathcal{U} 为矢量空间. 用 \mathcal{V}^p 表示 \mathcal{V} 的 p 次 Descartes 积. 所谓从 \mathcal{V} 到 \mathcal{U} 的 p 线性映射 (p -linear map), 就是一个映射 $\theta : \mathcal{V}^p \rightarrow \mathcal{U}$, 它对每个自变量都是线性的:

$$\begin{aligned}\theta(|a_1\rangle, \dots, \alpha|a_j\rangle + \beta|b_j\rangle, \dots, |a_p\rangle) \\ = \alpha\theta(|a_1\rangle, \dots, |a_j\rangle, \dots, |a_p\rangle) + \beta\theta(|a_1\rangle, \dots, |b_j\rangle, \dots, |a_p\rangle).\end{aligned}$$

从 \mathcal{V} 到 \mathbb{R} 或者 \mathbb{C} 的 p 线性映射就称作 \mathcal{V} 内的 p 线性函数 (p -linear function).

现在来看一个例子. 设 $\{\phi_i\}_{i=1}^p$ 是 \mathcal{V} 上的线性泛函. 现由下式定义 θ :

$$\theta(|a_1\rangle, \dots, |a_p\rangle) = \phi_1(|a_1\rangle) \dots \phi_p(|a_p\rangle), \quad |a_i\rangle \in \mathcal{V}.$$

很明显这里 θ 就是 p 线性的.

设 σ 是 $\{1, 2, \dots, p\}$ 的一个置换, ω 是个 p 线性映射. 现定义 p 线性映射 $\sigma\omega$ 如下⁶⁸:

$$\sigma\omega(|a_1\rangle, \dots, |a_p\rangle) = \omega(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle).$$

定义 2.6.2 (斜对称 p 线性映射). 设 ω 是从 \mathcal{V} 到 \mathcal{U} 的 p 线性映射, 若 $\sigma\omega = \varepsilon_\sigma \cdot \omega$. 即若

$$\omega(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) = \varepsilon_\sigma \omega(|a_1\rangle, \dots, |a_p\rangle),$$

则称 ω 是斜对称的 (skew-symmetric). 这里 ε_σ 是 σ 的符号: 当 σ 为偶置换时等于 $+1$, 为奇置换时等于 -1 .⁶⁹ 所有从 \mathcal{V} 到 \mathcal{U} 的斜对称 p 线性映射构成的集合记作 $\Lambda^p(\mathcal{V}, \mathcal{U})$; 而 \mathcal{V} 内所有斜对称 p 线性函数的集合则记做 $\Lambda^p(\mathcal{V})$, 即 $\Lambda^p(\mathcal{V}) := \Lambda^p(\mathcal{V}, \mathbb{F})$ (其中 \mathbb{F} 是 \mathbb{R} 或者 \mathbb{C}).

⁶⁸译者注: 不难看到 σ 对 ω 是线性的:

$$\begin{aligned}\sigma(\alpha\omega + \beta\theta)(|a_1\rangle, \dots, |a_p\rangle) &= (\alpha\omega + \beta\theta)(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) \\ &= \alpha\omega(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) + \beta\theta(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) \\ &= (\alpha\sigma\omega + \beta\sigma\theta)(|a_1\rangle, \dots, |a_p\rangle).\end{aligned}$$

在『(2.30) 式』下的证明中我们就用到了这个结果.

⁶⁹译者注: 如果我们将 $(a_{\sigma(1)}, \dots, a_{\sigma(n)})$ 做 n 次对换 (所谓对换, 就是交换两个元素的位置) 可以恢复成 (a_1, \dots, a_n) , 则称 n 为偶数时 σ 是偶置换, n 为奇数时 σ 是奇置换. (可以证明, 虽然这里 n 不唯一, 但是其奇偶性是固定的).

有些时候, 置换的符号 ε_σ 也会写作

$$\varepsilon_\sigma = \varepsilon_{\sigma(1)\sigma(2)\dots\sigma(p)} := \varepsilon_{i_1 i_2 \dots i_p}, \quad (2.29)$$

其中 $i_k := \sigma(k)$.

所有的 p 线性映射都可以转化为斜对称的 p 线性映射. 事实上, 若 θ 是 p 线性映射, 则

$$\omega := \sum_{\pi} \varepsilon_{\pi} \cdot \pi \theta \quad (2.30)$$

就是斜对称的, 因为

$$\begin{aligned} \sigma \omega &= \sigma \sum_{\pi} \varepsilon_{\pi} \cdot \pi \theta = \sum_{\pi} \varepsilon_{\pi} \cdot (\sigma \pi) \theta = (\varepsilon_\sigma)^2 \sum_{\pi} \varepsilon_{\pi} \cdot (\sigma \pi) \theta \\ &= \varepsilon_\sigma \sum_{\pi} (\varepsilon_\sigma \varepsilon_{\pi}) \cdot (\sigma \pi) \theta = \varepsilon_\sigma \sum_{\sigma \pi} \varepsilon_{\sigma \pi} \cdot (\sigma \pi) \theta = \varepsilon_\sigma \cdot \omega, \end{aligned}$$

这里我们用到了这样的事实: 两个置换乘积的符号等于其符号的乘积, 并且若 \sum_{π} 对所有置换求和, 则 $\sum_{\sigma \pi}$ 也是如此⁷⁰.

利用置换的性质可以证明如下定理:

定理 2.6.3. 设 $\omega \in \Lambda^p(\mathcal{V}, \mathcal{U})$, 则下面的陈述等价:

1. 只要 $|a_i\rangle = |a_j\rangle$ 对某对 $i \neq j$ 成立, 则有

$$\omega(|a_1\rangle, \dots, |a_p\rangle) = 0.$$

2. 对 $\{1, 2, \dots, p\}$ 的任意置换 σ 以及 \mathcal{V} 中的任意一组矢量 $|a_1\rangle, \dots, |a_p\rangle$ 有

$$\omega(|a_{\sigma(1)}\rangle, \dots, |a_{\sigma(p)}\rangle) = \varepsilon_\sigma \omega(|a_1\rangle, \dots, |a_p\rangle).$$

3. 只要 $\{|a_k\rangle\}_{k=1}^p$ 线性相关, 就有 $\omega(|a_1\rangle, \dots, |a_p\rangle) = 0$.

命题 2.6.4. 设 $N = \dim \mathcal{V}$ 且 $\omega \in \Lambda^N(\mathcal{V}, \mathcal{U})$. 则 ω 由其作用在 \mathcal{V} 基上的值唯一确定. 特别地, 若 ω 在某组基上等于零, 则 $\omega = 0$.

证. 设 $\{|e_k\rangle\}_{k=1}^N$ 是 \mathcal{V} 的一组基. 现在我们取 \mathcal{V} 中任意 N 个矢量 $\{|a_j\rangle\}_{j=1}^N$, 并记 $|a_j\rangle = \sum_{k=1}^N \alpha_{jk} |e_k\rangle$, $j = 1, \dots, N$. 那么, 利用 ω 对每个自变量线性就有

$$\begin{aligned} \omega(|a_1\rangle, \dots, |a_N\rangle) &= \sum_{1 \leq k_1, \dots, k_N \leq N} \alpha_{1k_1} \dots \alpha_{Nk_N} \omega(|e_{k_1}\rangle, \dots, |e_{k_N}\rangle) \\ &:= \sum_{\pi} \alpha_{1\pi(1)} \dots \alpha_{N\pi(N)} \omega(|e_{\pi(1)}\rangle, \dots, |e_{\pi(N)}\rangle) \\ &= \left(\sum_{\pi} \varepsilon_{\pi} \alpha_{1\pi(1)} \dots \alpha_{N\pi(N)} \right) \omega(|e_1\rangle, \dots, |e_N\rangle). \end{aligned}$$

⁷⁰译者注: 此处可以直接承认它们. 置换的性质可见『第??章』.

由于出现在括号中的那一项是常量, 我们就证明了该结果. \square

定义 2.6.5 (行列式函数). 设 $\dim \mathcal{V} = N$, 则 \mathcal{V} 里面的斜对称 N 线性函数 (即 $\Lambda^N(\mathcal{V})$ 中的元素) 就称作 \mathcal{V} 里的行列式函数 (*determinant function*).

⁷¹

现在设 $B = \{|e_k\rangle\}_{k=1}^N$ 是 \mathcal{V} 的一组基, $B^* = \{\boldsymbol{\varepsilon}_j\}_{j=1}^N$ 是其对偶基. 对 N 中任意 N 个矢量 $\{|a_k\rangle\}_{k=1}^N$, 我们由下式定义这么一个 N 线性函数 θ :

$$\theta(|a_1\rangle, \dots, |a_N\rangle) = \boldsymbol{\varepsilon}_1(|a_1\rangle) \dots \boldsymbol{\varepsilon}_N(|a_N\rangle).$$

现在注意,

$$\pi\theta(|e_1\rangle, \dots, |e_N\rangle) := \theta(|e_{\pi(1)}\rangle, \dots, |e_{\pi(N)}\rangle) = \delta_{\iota\pi},$$

这里 ι 是恒等置换 (即所有元素保持不变), 而 $\delta_{\iota\pi}$ 和 Kronecker delta 含义类似: 若 $\pi = \iota$ 则等于 1, 反之则等于 0. 现在定义 Δ 为 $\Delta := \sum_{\pi} \varepsilon_{\pi} \cdot \pi\theta$. 那么, 根据『(2.30) 式』可知 $\Delta \in \Lambda^N(\mathcal{V})$, 也就是说 Δ 是个行列式函数. 不仅如此, 我们还有

$$\begin{aligned} \Delta(|e_1\rangle, \dots, |e_N\rangle) &= \sum_{\pi} \varepsilon_{\pi} \cdot \pi\theta(|e_1\rangle, \dots, |e_N\rangle) \\ &= \sum_{\pi} \varepsilon_{\pi} \delta_{\iota\pi} = \varepsilon_{\iota} = 1. \end{aligned}$$

由此我们就有如下结论:

结论 2.6.6. 每个有限维矢量空间中都存在一个不恒等于零的行列式函数.

命题 2.6.7. 设 $\omega \in \Lambda^N(\mathcal{V}, \mathcal{U})$. 令 Δ 为 \mathcal{V} 中一个固定的行列式函数. 则 ω 唯一确定了一个矢量 $|u_{\Delta}\rangle \in \mathcal{U}$ 使得

$$\omega(|v_1\rangle, \dots, |v_N\rangle) = \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |u_{\Delta}\rangle.$$

证. 因为 Δ 是非零行列式函数, 故可以取 \mathcal{V} 的一组基 $\{|v_k\rangle\}_{k=1}^N$ 使得 $\Delta(|v_1\rangle, \dots, |v_N\rangle) \neq 0$. 既然如此, 我们就可以在两边除以这个非零值 (除给其中一个基矢或者除给 Δ 自己), 进而假定 $\Delta(|v_1\rangle, \dots, |v_N\rangle) = 1$. 此时我们将 $\omega(|v_1\rangle, \dots, |v_N\rangle)$ 定义为 $|u_{\Delta}\rangle$. 现在我们注意到 $\omega - \Delta \cdot |u_{\Delta}\rangle$

⁷¹译者注: 如果直译的话应当将 determinant 翻译为决定式. 这个术语本身来自古典代数学, 人们发现用这个东西可以决定一个方程组是否有解. 因此这样命名, 不过当时的 determinant 大多表现为一个从数表 (矩阵) 到数的映射, 它还可以表示成矩阵各行各列元素乘积的组合, 因此取了“行列式”这个直观的名字. 但后来理论的发展和行列关系越来越远, 反而更加侧重“决定”的含义. 是故有些人认为应当溯本归源, 改回决定式的译法. 不过考虑到行列式已经流传甚广, 这种已经成为固定术语的词更改译名得不偿失, 是故这里还是按照传统译作行列式.

作用在 $\{|v_k\rangle\}_{k=1}^N$ 这组基上会给出零. 从而由『[命题 2.6.4](#)』可知它必然恒等于零. ⁷²

□

推论 2.6.8. 设 Δ 是 \mathcal{V} 中一个固定的非零行列式函数. 那么, 每个行列式函数都是 Δ 乘以一个标量倍数.

证. 在『[命题 2.6.7](#)』中置 \mathcal{U} 为 \mathbb{C} (或者 \mathbb{R}) 即可.

□

命题 2.6.9. 设 Δ 是 N 维矢量空间 \mathcal{V} 中一个行列式函数. 若 $|v\rangle$ 和 $\{|v_k\rangle\}_{k=1}^N$ 是 \mathcal{V} 中的矢量, 那么

$$\begin{aligned} & \sum_{j=1}^N (-1)^{j-1} \Delta \left(|v\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle \right) \cdot |v_j\rangle \\ & = \Delta (|v_1\rangle, \dots, |v_N\rangle) \cdot |v\rangle, \end{aligned}$$

其中矢量上加 *hat* 表示删去这个矢量.

证. 见『[习题 2.37](#)』.

□

⁷²译者注: 注意命题中的 $\{v_i\}_{i=1}^N$ 是任意矢量. 但是『[命题 2.6.4](#)』指出只要在基上的结果确定, 那么在任意矢量上的结果也就确定了. 因此证明中就让其是一组基. 另外, 这里只证明了存在性, 没有证明唯一性. 这是因为唯一性显然: 这里 $\Delta(|v_1\rangle, \dots, |v_N\rangle)$ 是非零标量, 而 $\alpha|u\rangle = \alpha|v\rangle$ 的充要条件就是 $|u\rangle = |v\rangle$.

2.6.1. 线性算符的行列式

设 \mathbf{A} 是 N 维矢量空间 \mathcal{V} 上的线性算符. 选定某一非零行列式函数 Δ . 对于 \mathcal{V} 的一组基 $\{v_i\}_{i=1}^N$, 定义 Δ_A 如下:

$$\Delta_A(|v_1\rangle, \dots, |v_N\rangle) := \Delta(\mathbf{A}|v_1\rangle, \dots, \mathbf{A}|v_N\rangle). \quad (2.31)$$

显然, Δ_A 同样是个行列式函数. 而根据『推论 2.6.8』可知, 它是 Δ 的倍数. 即有 $\Delta_A = \alpha\Delta$. 不仅如此, 这个乘数 α 还与这个非零行列式函数的选法无关. 因为若 Δ' 是另一个非零行列式函数, 则再次利用『推论 2.6.8』就有 $\Delta' = \lambda\Delta$, 从而

$$\Delta'_A = \lambda\Delta_A = \lambda\alpha\Delta = \alpha\Delta'.$$

这就表明 α 仅由 \mathbf{A} 确定, 而与非零行列式函数以及基的选法无关. 由此就引出了如下定义:

定义 2.6.10 (算符的行列式). 设 $\mathbf{A} \in \text{End}(\mathcal{V})$, Δ 是 \mathcal{V} 中一个非零行列式函数, 并如『(2.31) 式』所述定义 Δ_A . 那么, 使得

$$\Delta_A = \det \mathbf{A} \cdot \Delta \quad (2.32)$$

成立的那个系数 $\det \mathbf{A}$ 就称作 \mathbf{A} 的行列式 (*determinant*).

利用『(2.32) 式』, 我们就有下面的一些个定理, 它们的证明均留作习题.

定理 2.6.11. 线性算符 \mathbf{A} 的行列式具有如下性质:

1. 若 $\mathbf{A} = \lambda\mathbf{1}$, 则 $\det \mathbf{A} = \lambda^N$.
2. \mathbf{A} 可逆的充要条件是 $\det \mathbf{A} \neq 0$.^a
3. $\det(\mathbf{A} \circ \mathbf{B}) = \det \mathbf{A} \det \mathbf{B}$.

^a译者注: 这就是之前提到将其命名为决定式 (determinant) 的根本原因: 它决定了算符是否可逆.

2.6.2. 古典伴随

设 \mathcal{V} 是 N 维矢量空间, Δ 是 \mathcal{V} 中一个行列式函数, $\mathbf{A} \in \text{End}(\mathcal{V})$. 对 $|v\rangle, |v_i\rangle \in \mathcal{V}$, 定义 $\Phi : \mathcal{V}^N \rightarrow \text{End}(\mathcal{V})$ 如下:

$$\begin{aligned} \Phi(|v_1\rangle, \dots, |v_N\rangle)|v\rangle \\ = \sum_{j=1}^N (-1)^{j-1} \Delta \left(|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle \right) \cdot |v_j\rangle. \end{aligned}$$

很明显, Φ 是斜对称的. 因此, 根据『[命题 2.6.7](#)』可知存在唯一的线性算符 $\text{ad}(\mathbf{A})$ 使得

$$\Phi(|v_1\rangle, \dots, |v_N\rangle) = \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \text{ad}(\mathbf{A}),$$

也就是说

$$\begin{aligned} & \sum_{j=1}^N (-1)^{j-1} \Delta \left(|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle \right) \cdot |v_j\rangle \\ &= \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \text{ad}(\mathbf{A})|v\rangle. \end{aligned} \quad (2.33)$$

这个式子表明 $\text{ad}(\mathbf{A})$ 与行列式函数的选法无关, 我们称其为 \mathbf{A} 的古典伴随 (classical adjoint). ⁷³

命题 2.6.12. 古典伴随满足如下关系:

$$\text{ad}(\mathbf{A}) \circ \mathbf{A} = \det \mathbf{A} \cdot \mathbf{1} = \mathbf{A} \circ \text{ad}(\mathbf{A}), \quad (2.34)$$

其中 $\mathbf{1}$ 为单位算符.

证. 将『[\(2.33\) 式](#)』中的 $|v\rangle$ 替换为 $\mathbf{A}|v\rangle$ 就有

$$\begin{aligned} & \sum_{j=1}^N (-1)^{j-1} \Delta \left(\mathbf{A}|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle \right) \cdot |v_j\rangle \\ &= \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot \text{ad}(\mathbf{A}) \circ \mathbf{A}|v\rangle. \end{aligned}$$

另一方面, 根据算符行列式的定义, 上式的左边可以这样改写:

$$\begin{aligned} \text{LHS} &= \det \mathbf{A} \cdot \sum_{j=1}^N (-1)^{j-1} \Delta \left(|v\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle \right) \cdot |v_j\rangle \\ &= \det \mathbf{A} \cdot \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |v\rangle, \end{aligned}$$

这里最后一个等号用到了『[命题 2.6.9](#)』. 由于 $|v\rangle$ 是任意的, 我们就证得了本命题的第一个等号.

欲证第二个等号, 我们在『[\(2.33\) 式](#)』两边作用算符 \mathbf{A} . 那么根据『[命题 2.6.9](#)』, 其结果的左边就变成

$$\begin{aligned} \text{LHS} &= \sum_{j=1}^N (-1)^{j-1} \Delta \left(|v\rangle, \mathbf{A}|v_1\rangle, \dots, \widehat{\mathbf{A}|v_j\rangle}, \dots, \mathbf{A}|v_N\rangle \right) \cdot \mathbf{A}|v_j\rangle \\ &= \Delta(\mathbf{A}|v_1\rangle, \dots, \mathbf{A}|v_N\rangle) \cdot |v\rangle = \det \mathbf{A} \cdot \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot |v\rangle, \end{aligned}$$

⁷³译者注: 一些读者可能会在学习线性代数时遇到一个用矩阵的代数余子式定义的矩阵, 一些书中将其命名为伴随矩阵 (adjugate matrix). 它其实抽象化后就是这里定义的古典伴随. 由于一些书中同样把这个伴随矩阵称作 adjoint matrix, 为了避免混淆, 现在通常采用古典伴随这个说法 (事实上, 这是伴随这个概念最先提出时对应的定义). 同样为了避免混淆, 一些作者也称 adjugate matrix 为余子式矩阵 (cofactor matrix).

而右边则变成

$$\text{RHS} = \Delta(|v_1\rangle, \dots, |v_N\rangle) \cdot A \circ \text{ad}(\mathbf{A})|v\rangle.$$

同样由于两边对任意 $|v\rangle$ 成立, 我们就得到了本命题的第二个等号. \square

推论 2.6.13. 若 $\det \mathbf{A} \neq 0$, 则 \mathbf{A} 可逆, 并且

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \cdot \text{ad}(\mathbf{A}).$$

2.7. 本章习题

习题 2.1. 用 \mathbb{R}^+ 表示所有正实数的集合. 定义 \mathbb{R}^+ 中两个元素的“和”为它们的常规乘积; 定义其元素与 \mathbb{R} 中元素的标量乘法为 $r \cdot p = p^r$, 其中 $r \in \mathbb{R}$, $p \in \mathbb{R}^+$. 求证: 在这两个运算下, \mathbb{R}^+ 是 \mathbb{R} 上的矢量空间.

习题 2.2. 求证: 两个子空间的交集还是子空间.

习题 2.3. 判断下述 \mathbb{R}^3 的子集是否亦为其子空间:

- (a) $\{(x, y, z) \in \mathbb{R}^3 \mid x + y - 2z = 0\};$
- (b) $\{(x, y, z) \in \mathbb{R}^3 \mid x + y - 2z = 3\};$
- (c) $\{(x, y, z) \in \mathbb{R}^3 \mid xyz = 0\}.$

习题 2.4. 求证: 矢量在给定基下的分量是唯一的.

习题 2.5. 证明下面的矢量构成 \mathbb{C}^n (或 \mathbb{R}^n) 的一组基.

$$|a_1\rangle = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}, \quad |a_2\rangle = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \dots, \quad |a_n\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}.$$

习题 2.6. 证明『定理 2.1.6』.

习题 2.7. 设 \mathcal{W} 为 \mathbb{R}^5 的子空间, 其定义如下:

$$\mathcal{W} = \{(x_1, \dots, x_5) \in \mathbb{R}^5 \mid x_1 = 3x_2 + x_3, x_2 = x_5, x_4 = 2x_3\}.$$

求出 \mathcal{W} 的一组基.

习题 2.8. 设 \mathcal{U}_1 和 \mathcal{U}_2 是 \mathcal{V} 的子空间. 求证:

- (a) $\dim(\mathcal{U}_1 + \mathcal{U}_2) = \dim \mathcal{U}_1 + \dim \mathcal{U}_2 - \dim(\mathcal{U}_1 \cap \mathcal{U}_2).$

提示: 设 $\{|a_i\rangle\}_{i=1}^m$ 是 $\mathcal{U}_1 \cap \mathcal{U}_2$ 的一组基. 接下来将其扩充为 \mathcal{U}_1 的一组基 $\{|a_i\rangle\}_{i=1}^m \cup \{|b_i\rangle\}_{i=1}^k$, 也将其扩充为 \mathcal{U}_2 的一组基 $\{|a_i\rangle\}_{i=1}^m \cup \{|c_i\rangle\}_{i=1}^l$. 现在证明 $\{|a_i\rangle\}_{i=1}^m \cup \{|b_i\rangle\}_{i=1}^k \cup \{|c_i\rangle\}_{i=1}^l$ 是 $\mathcal{U}_1 + \mathcal{U}_2$ 的一组基.

- (b) 若 $\mathcal{U}_1 + \mathcal{U}_2 = \mathcal{V}$ 且 $\dim \mathcal{U}_1 + \dim \mathcal{U}_2 = \dim \mathcal{V}$, 则 $\mathcal{V} = \mathcal{U}_1 \oplus \mathcal{U}_2$.
- (c) 若 $\dim \mathcal{U}_1 + \dim \mathcal{U}_2 > \dim \mathcal{V}$, 则 $\mathcal{U}_1 \cap \mathcal{U}_2 \neq \{0\}$.

习题 2.9. 求证: 『(2.5) 式』定义的矢量张成了 $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$.

习题 2.10. 求证: 任意矢量与零矢量 $|0\rangle$ 的内积均为零.

习题 2.11. 给出一组标量 $a_0, b_0, b_1, c_0, c_1, c_2$ 使得多项式 $a_0, b_0 + b_1t, c_0 + c_1t + c_2t^2$ 在区间 $[0, 1]$ 内相互正交. 对应的多项式内积如『例 2.2.3』定义, 并取 $w(t) = 1$.

习题 2.12. 给定 $\mathcal{P}^c[t]$ 中一组线性无关的矢量 $x(t) = t^n, n = 0, 1, \dots$ 在下述内积下利用 Gram-Schmidt 方法找出前三个正交多项式 $e_0(t), e_1(t)$ 和 $e_2(t)$.

- (a) 定义内积为 $\langle x|y \rangle = \int_{-1}^1 x^*(t)y(t) dt$.
- (b) 定义内积为带非平凡权函数 $w(t) = e^{-t^2}$ 的内积, 即

$$\langle x|y \rangle = \int_{-\infty}^{\infty} e^{-t^2} x^*(t)y(t) dt.$$

提示: 利用下述结论:

$$\int_{-\infty}^{\infty} e^{-t^2} t^n dt = \begin{cases} \sqrt{\pi} & \text{若 } n = 0, \\ 0 & \text{若 } n \text{ 为奇数,} \\ \sqrt{\pi} \frac{1 \cdot 3 \cdot 5 \cdots (n-1)}{2^{n/2}} & \text{若 } n \text{ 为偶数.} \end{cases}$$

习题 2.13. 请完成下述要求:

- (a) 利用 Gram-Schmidt 方法从 $(1, -1, 1), (-1, 0, 1), (2, -1, 2)$ 这三个矢量得到一个正交矢量组.
- (b) 上面给出的这三个矢量线性无关吗? 如果不是, 请通过 (a) 找到它们的零线性组合 (即给出使得其线性组合为零的那些系数).

习题 2.14. 请完成下述要求:

- (a) 利用 Gram-Schmidt 方法基于下述三个矢量得到一个正交矢量组: $(1, -1, 2), (-2, 1, -1), (-1, -1, 4)$.
- (b) 上面给出的这三个矢量线性无关吗? 如果不是, 请通过 (a) 找到它们的零线性组合.

习题 2.15. 证明下述积分不等式:

$$\begin{aligned} & \int_{-\infty}^{\infty} (t^{10} - t^6 + 5t^4 - 5)e^{-t^4} dt \\ & \leq \sqrt{\int_{-\infty}^{\infty} (t^4 - 1)^2 e^{-t^4} dt} \sqrt{\int_{-\infty}^{\infty} (t^6 + 5)^2 e^{-t^4} dt}. \end{aligned}$$

提示: 定义合适的内积后使用 Schwarz 不等式.

习题 2.16. 证明下述不等式:

$$\begin{aligned} & \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy (x^5 - x^3 + 2x^2 - 2)(y^5 - y^3 + 2y^2 - 2) e^{-(x^4+y^4)} \\ & \leq \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy (x^4 - 2x^2 + 1)(y^6 + 4y^3 + 4) e^{-(x^4+y^4)}. \end{aligned}$$

提示: 定义合适的内积后使用 Schwarz 不等式.

习题 2.17. 求证, 对任意 n 个复数 $\alpha_1, \alpha_2, \dots, \alpha_n$, 我们都有

$$|\alpha_1 + \alpha_2 + \dots + \alpha_n|^2 \leq n(|\alpha_1|^2 + |\alpha_2|^2 + \dots + |\alpha_n|^2).$$

提示: 对 $(1, 1, \dots, 1)$ 和 $(\alpha_1, \alpha_2, \dots, \alpha_n)$ 使用 Schwarz 不等式.

习题 2.18. 利用 Schwarz 不等式证明: 若 $\{\alpha_i\}_{i=1}^{\infty}$ 和 $\{\beta_i\}_{i=1}^{\infty}$ 均属于 \mathbb{C}^{∞} , 则 $\sum_{i=1}^{\infty} \alpha_i^* \beta_i$ 收敛.

习题 2.19. 由 $\mathbf{T}(x, y) = (x^2 + y^2, x + y, 2x - y)$ 定义映射 $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$. 请说明这不是个线性映射.

习题 2.20. 验证『例 2.3.5』中所有的变换都是线性的.

习题 2.21. 设 π 是将 $(1, 2, 3)$ 变成 $(3, 1, 2)$ 的那个置换. 给出

$$\mathbf{A}_{\pi}|e_i\rangle, \quad i = 1, 2, 3,$$

其中 $\{|e_i\rangle\}_{i=1}^3$ 是 \mathbb{R}^3 (或 \mathbb{C}^3) 的标准基, \mathbf{A}_{π} 则如『例 2.3.5』那样定义.

习题 2.22. 求证: 若 $\mathbf{T} \in \mathcal{L}(\mathbb{C}, \mathbb{C})$, 则存在 $\alpha \in \mathbb{C}$ 使得 $\mathbf{T}|a\rangle = \alpha|a\rangle$ 对所有 $|a\rangle \in \mathbb{C}$ 成立.

习题 2.23. 求证: 若 $\{|a_i\rangle\}_{i=1}^n$ 张成 \mathcal{V} 且 $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, 那么 $\{\mathbf{T}|a_i\rangle\}_{i=1}^n$ 就张成了 $\mathbf{T}(\mathcal{V})$. 特别地, 若 \mathbf{T} 是满射, 则 $\{\mathbf{T}|a_i\rangle\}_{i=1}^n$ 就张成了 \mathcal{W} .

习题 2.24. 构造一个使得

$$f(\alpha|a\rangle) = \alpha f(|a\rangle) \quad \forall \alpha \in \mathbb{R}, \forall |a\rangle \in \mathbb{R}^2$$

的非线性函数 $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.

提示: 考虑一个 1 次齐次函数.

习题 2.25. 证明下面的变换都是线性的:

(a) \mathcal{V} 为实数上的 \mathbb{C} , 而 $\mathbf{C}|z\rangle = |z^*\rangle$. 如果用复数代替实数作标量, 那么 \mathbf{C} 是线性的吗?

(b) \mathcal{V} 为 $\mathcal{P}^c[t]$, 而 $\mathbf{T}|x(t)\rangle = |x(t+1)\rangle - |x(t)\rangle$.

习题 2.26. 求证: 线性变换 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$ 的核是 \mathcal{V} 的子空间, 它的像 $\mathbf{T}(\mathcal{V})$ 是 \mathcal{W} 的子空间.

习题 2.27. 设 \mathcal{V} 和 \mathcal{W} 是有限维线性空间. 求证: 若 $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ 是满射, 则 $\dim \mathcal{W} \leq \dim \mathcal{V}$.

习题 2.28. 假设 \mathcal{V} 是有限维的, 并且 $\mathbf{T} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$ 非零. 求证: 存在 \mathcal{V} 的子空间 \mathcal{U} , 使得 $\ker \mathbf{T} \cap \mathcal{U} = \{0\}$ 且 $\mathbf{T}(\mathcal{V}) = \mathbf{T}(\mathcal{U})$.

习题 2.29. 利用『定理 2.3.11』来证明『定理 2.3.18』.

习题 2.30. 利用『定理 2.3.11』来证明『定理 2.3.19』.

习题 2.31. 设 $B_V = \{|a_i\rangle\}_{i=1}^N$ 是 \mathcal{V} 的一组基, $B_W = \{|b_i\rangle\}_{i=1}^N$ 是 \mathcal{W} 的一组基. 现通过 $\mathbf{T}|a_i\rangle = |b_i\rangle, i = 1, 2, \dots, N$ 定义线性变换 $\mathbf{T} : \mathcal{V} \rightarrow \mathcal{W}$. 证明 \mathbf{T} 是同构, 并以此证明『定理 2.3.20』.

习题 2.32. 求证: 由『定义 2.4.3』给出的伴随满足 $(\mathbf{A}^\top)^\top = \mathbf{A}$.

习题 2.33. 证明 \mathcal{W}^0 是 \mathcal{V}^* 的子空间, 并且

$$\dim \mathcal{V} = \dim \mathcal{W} + \dim \mathcal{W}^0.$$

习题 2.34. 证明 N 维矢量空间 \mathcal{V}^* 中的每个矢量都有 $N - 1$ 个线性无关的零化子. 换言之, 任意线性泛函都会将 $N - 1$ 个线性无关的矢量变成零.

习题 2.35. 证明 \mathbf{T} 和 \mathbf{T}^* 有相同的秩. 特别地, 证明若 \mathbf{T} 是单射, 则 \mathbf{T}^* 就是满射.

提示: 利用 \mathbf{T} 和 \mathbf{T}^* 的维数定理以及『(2.25) 式』.

习题 2.36. 证明『定理 2.6.3』.

习题 2.37. 证明『命题 2.6.9』.

提示: 首先证明若 $\{|v_k\rangle\}_{k=1}^N$ 线性无关, 则我们两边得到的都是零. 接下来假定它们线性无关, 并将其取作基, 然后将 $|v\rangle$ 在这组基下展开, 并注意除非 $i = j$, 不然总有

$$\Delta(|v\rangle, |v_1\rangle, \dots, \widehat{|v_j\rangle}, \dots, |v_N\rangle) = 0.$$

习题 2.38. 证明『定理 2.6.11』.

提示: 对于定理的第二部分, 可以利用这样一个事实: 可逆的线性映射 \mathbf{A} 会将线性无关的矢量组映满到线性无关的矢量组.

3

代数

在很多物理学场景中, 矢量空间 \mathcal{V} 会有个自然的“乘积”, 也就是一个二元运算 $\mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$, 我们称其为乘法. 这种矢量空间的主要实例就是矩阵空间. 因此, 考虑一下带有这种乘积的矢量空间是有用武之地的.

3.1. 从矢量空间到代数

在这一节, 我们来定义什么叫代数, 给出我们熟悉的一些例子, 并讨论其一些基本性质.

定义 3.1.1 (代数以及相关概念). 所谓 \mathbb{C} (或者 \mathbb{R}) 上的代数 (*algebra*) \mathcal{A} , 就是一个 \mathbb{C} (或者 \mathbb{R}) 上的矢量空间, 它上面还带有一个称作乘法 (*multiplication*) 的二元运算 $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$. 在乘法这个映射下, $(\mathbf{a}, \mathbf{b}) \in \mathcal{A} \times \mathcal{A}$ 的像就记作 \mathbf{ab} , * 对所有的 $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{A}$ 以及 $\beta, \gamma \in \mathbb{C}$ (或者 \mathbb{R}) 它要满足如下两个关系:

$$\mathbf{a}(\beta\mathbf{b} + \gamma\mathbf{c}) = \beta\mathbf{ab} + \gamma\mathbf{ac},$$

$$(\beta\mathbf{b} + \gamma\mathbf{c})\mathbf{a} = \beta\mathbf{ba} + \gamma\mathbf{ca}.$$

底矢量空间的维数就称作该代数的维数 (*dimension of the algebra*). 如果乘积满足结合律, 即有 $\mathbf{a}(\mathbf{bc}) = (\mathbf{ab})\mathbf{c}$, 则称该代数为结合代数 (*associative algebra*); 如果乘积满足交换律, 即有 $\mathbf{ab} = \mathbf{ba}$, 则称其为交换代数 (*commutative algebra*). 若存在元素 $\mathbf{1}$ 使得对所有 $\mathbf{a} \in \mathcal{A}$ 都有 $\mathbf{a}\mathbf{1} = \mathbf{1}\mathbf{a} = \mathbf{a}$, 则称该代数为带单位元的代数 (*algebra with identity*). 对含幺代数而言, 若 \mathbf{b} 满足 $\mathbf{ba} = \mathbf{1}$, 则称 \mathbf{b} 是 \mathbf{a} 的左逆 (*left inverse*); 若 $\mathbf{ab} = \mathbf{1}$, 则称 \mathbf{b} 是 \mathbf{a} 的右逆 (*right inverse*). 带单位元的代数中, 单位元 $\mathbf{1}$ 也称作幺元 (*unit*), 是故它也得名含幺代数 (*unital algebra*).

我们有时候需要用不同的符号来表示代数的幺元. 特别是当我们同时讨论多个代数之时. 除了 $\mathbf{1}$ 以外, 另一个常用于表示幺元的符号就是 \mathbf{e} .

* 本章我们很大程度上会弃用 Dirac 的左右矢记号体系, 因为在这一主题下它显得过于笨拙. 作为替代, 我们用黑体的拉丁字母来表示矢量.

3.1.1. 一般性质

在『[定义 3.1.1](#)』中取 $\beta = 1 = -\gamma$, 并令 $\mathbf{b} = \mathbf{c}$, 那么立刻就能得到下面的结果:

$$\mathbf{a}\mathbf{0} = \mathbf{0}\mathbf{a} = \mathbf{0} \quad \forall \mathbf{a} \in \mathcal{A}.$$

代数的单位元若存在则必唯一. 这是因为若有两个单位元 $\mathbf{1}$ 和 \mathbf{e} , 那么由于 $\mathbf{1}$ 是单位元就有 $\mathbf{1}\mathbf{e} = \mathbf{e}$; 由于 \mathbf{e} 是单位元就有 $\mathbf{1}\mathbf{e} = \mathbf{1}$. 两相结合则得 $\mathbf{1} = \mathbf{e}$.

若 \mathcal{A} 是结合代数, 且 $\mathbf{a} \in \mathcal{A}$ 同时存在左逆 \mathbf{b} 和右逆 \mathbf{c} , 那么它们必然相等. 因为 \mathbf{bac} 有两种不同的计算方法, 它们分别给出 \mathbf{b} 和 \mathbf{c} :

$$\begin{aligned} \mathbf{bac} &= (\mathbf{ba})\mathbf{c} = \mathbf{1}\mathbf{c} = \mathbf{c}, \\ \mathbf{bac} &= \mathbf{b}(\mathbf{ac}) = \mathbf{b}\mathbf{1} = \mathbf{b}. \end{aligned}$$

因此, 在结合代数中, 若左逆和右逆均存在, 我们就只提逆(元)这个说法, 而不指定到底是左逆还是右逆. 不仅如此, 容易证明(双边)逆存在则唯一. 由此, 我们就有

定理 3.1.2. 设 \mathcal{A} 为含幺结合代数. 若 $\mathbf{a} \in \mathcal{A}$ 同时存在左逆和右逆, 则二者必然相等, 且满足该条件的元素唯一(称作逆(元), *inverse*), 记作 \mathbf{a}^{-1} . 若 \mathbf{a} 和 \mathbf{b} 均可逆, 则 \mathbf{ab} 也可逆, 并且

$$(\mathbf{ab})^{-1} = \mathbf{b}^{-1}\mathbf{a}^{-1}.$$

定理中最后那个陈述的证明很直接, 直接代入逆的定义验证即可.

定义 3.1.3 (子代数). 设 \mathcal{A} 为代数, \mathcal{A}' 是其子空间. 若 \mathcal{A}' 在乘法下封闭, 即若 $\mathbf{a}, \mathbf{b} \in \mathcal{A}'$, 则有 $\mathbf{ab} \in \mathcal{A}'$. 那么, 就称 \mathcal{A}' 是 \mathcal{A} 的子代数(*subalgebra*).

很明显, 结合(交换)代数的子代数同样是结合(交换)的.

设 \mathcal{A} 为结合代数, S 是其子集. 所谓由 S 生成的子代数(*subalgebra generated by S*), 就是

$$\mathbf{s}_1\mathbf{s}_2 \dots \mathbf{s}_k, \quad \mathbf{s}_i \in S$$

所有线性组合的集合. 如果 S 是由 \mathbf{s} 构成的单点集(*singleton*), 即 $S = \{\mathbf{s}\}$, 则 S 生成的子代数就是 \mathbf{s} 的多项式集合.

例 3.1.4 (\mathbb{C} 是任意复含幺代数的子代数). 设 \mathcal{A} 是含幺代数, 则矢量空间

$$\text{Span}\{\mathbf{1}\} = \{\alpha\mathbf{1} \mid \alpha \in \mathbb{C}\}$$

就是 \mathcal{A} 的一个子代数. 由于 $\text{Span}\{\mathbf{1}\}$ 和 \mathbb{C} 无法区分, 我们有时也称 \mathbb{C} 是 \mathcal{A} 的子代数.

定义 3.1.5 (中心). 设 \mathcal{A} 为代数, 与 \mathcal{A} 中所有元素均对易¹的元素构成 \mathcal{A} 的中心 (center), 记作 $\mathcal{Z}(\mathcal{A})$.

容易证明 $\mathcal{Z}(\mathcal{A})$ 是 \mathcal{A} 的子空间, 并且若 \mathcal{A} 结合, 则 $\mathcal{Z}(\mathcal{A})$ 还是其子代数.

定义 3.1.6 (中心代数). 若含幺代数 \mathcal{A} 满足 $\mathcal{Z}(\mathcal{A}) = \text{Span}\{\mathbf{1}\}$, 则称其为中心代数 (central algebra).

例 3.1.7. 考察以 $\{\mathbf{e}_i\}_{i=0}^3$ 为基的代数, 其乘法表由『表 3.1』给出. 这里用 \mathbf{e}_0 表示单位元纯粹是出于审美.

我们现在看看该代数的中心有和特征. 令 $\mathbf{a} \in \mathcal{Z}(\mathcal{S})$, 则对任意 $\mathbf{b} \in \mathcal{S}$ 我们必须有 $\mathbf{ab} = \mathbf{ba}$. 现在设

$$\mathbf{a} = \sum_{i=0}^3 \alpha_i \mathbf{e}_i, \quad \mathbf{b} = \sum_{i=0}^3 \beta_i \mathbf{e}_i.$$

利用乘法表, 我们可以直接算得

$$\begin{aligned} \mathbf{ab} &= (\alpha_0\beta_0 + \alpha_1\beta_1 - \alpha_2\beta_2 + \alpha_3\beta_3)\mathbf{e}_0 \\ &\quad + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)\mathbf{e}_1 \\ &\quad + (\alpha_0\beta_2 + \alpha_1\beta_3 + \alpha_2\beta_0 - \alpha_3\beta_1)\mathbf{e}_2 \\ &\quad + (\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_0)\mathbf{e}_3. \end{aligned}$$

将上式中的 α, β 互换则可得到 \mathbf{ba} 的表达式. 简单比较一下就会看到 $\mathbf{ab} = \mathbf{ba}$ 的充要条件是

$$\alpha_2\beta_3 = \alpha_3\beta_2 \quad \text{且} \quad \alpha_1\beta_3 = \alpha_3\beta_1.$$

仅当 $\alpha_1 = \alpha_2 = \alpha_3 = 0$ 且 α_0 任意时上式对所有 \mathbf{b} 成立.² 因此, $\mathbf{a} \in \mathcal{Z}(\mathcal{S})$ 的充要条件是 \mathbf{a} 是 \mathbf{e}_0 的倍数, 也就等价于说 $\mathbf{a} \in \text{Span}\{\mathbf{e}_0\}$. 职是之故, \mathcal{S} 是中心代数.

设 A 和 B 是代数 \mathcal{A} 的子集. 我们用 AB 表示所有可以写成 A 中元素与 B 中元素乘积之和的元素³. 换言之,

$$AB := \left\{ \mathbf{x} \in \mathcal{A} \mid \mathbf{x} = \sum_k \mathbf{a}_k \mathbf{b}_k, \mathbf{a}_k \in A, \mathbf{b}_k \in B \right\}. \quad (3.1)$$

¹译者注: commute 在数学文献中通常翻译为交换, 在物理文献中通常翻译为对易. 和将 vector 翻译为矢量而非向量的理由一样 (毕竟这是一本给物理人的数学书), 这里我们选择对易的说法. 但这个说法对交换律、交换代数等描述不成立, 我们不会将其翻译为对易律以及对易代数.

²译者注: 取 $\beta_1 = \beta_2 = \beta_3 = 1$ 则推得 $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$. 进一步我们看到 $\alpha\beta_1 = \alpha\beta_2$ 需要对任意 β_1, β_2 成立, 若 $\beta_1 \neq \beta_2$, 则推得 $\alpha = 0$. 另外, 这里得到的结论是充要的: 若这个条件满足, 则必然也有 $\mathbf{ab} = \mathbf{ba}$, 是故下文从必要条件直接快进到了充要条件.

³译者注: 理论上这里后面还有一句“构成的集合”, 但这样就导致句子修饰成分太多, 读着拗口. 之后但凡我们说“A 表示 (所有) 具有性质 $P(x)$ 的元素”, 都是指 $A = \{x \mid P(x)\}$ 这个集合.

	\mathbf{e}_0	\mathbf{e}_1	\mathbf{e}_2	\mathbf{e}_3
\mathbf{e}_0	\mathbf{e}_0	\mathbf{e}_1	\mathbf{e}_2	\mathbf{e}_3
\mathbf{e}_1	\mathbf{e}_1	\mathbf{e}_0	\mathbf{e}_3	\mathbf{e}_2
\mathbf{e}_2	\mathbf{e}_2	$-\mathbf{e}_3$	$-\mathbf{e}_0$	\mathbf{e}_1
\mathbf{e}_3	\mathbf{e}_3	$-\mathbf{e}_2$	$-\mathbf{e}_1$	\mathbf{e}_0

表 3.1: \mathcal{S} 的乘法表

特别地, 称

$$\mathcal{A}^2 := \left\{ \mathbf{x} \in \mathcal{A} \mid \mathbf{x} = \sum_k \mathbf{a}_k \mathbf{b}_k, \mathbf{a}_k, \mathbf{b}_k \in \mathcal{A} \right\} \quad (3.2)$$

为 \mathcal{A} 的导出代数 (derived algebra).⁴

定义 3.1.8 (反代数). 给定任意代数 \mathcal{A} , 其乘法为 $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{ab}$. 由此可以得到底矢量空间相同, 但乘法变成 $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{ba}$ 的代数 \mathcal{A}^{op} . 记

$$(\mathbf{ab})^{\text{op}} = \mathbf{ba},$$

并称 \mathcal{A}^{op} 为 \mathcal{A} 的反代数 (*opposite algebra of \mathcal{A}* , 或 *algebra opposite to \mathcal{A}*).

容易看出, 若 \mathcal{A} 是结合代数, 那么 \mathcal{A}^{op} 亦是如此; 若 \mathcal{A} 是交换代数, 则 $\mathcal{A}^{\text{op}} = \mathcal{A}$.

例 3.1.9. 这里我们给出一些代数的例子.

- 在 \mathbb{R}^2 上定义下述乘积:

$$(x_1, x_2)(y_1, y_2) = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1).$$

请读者证明在上述乘积下 \mathbb{R}^2 为结合代数.

- 类似地, \mathbb{R}^3 上的矢量积 (叉积) 却使 \mathbb{R}^3 变成非结合、非交换的代数.
- 所有代数的范式 (*paradigm*) 都是矩阵代数 (*matrix algebra*), 其二元运算就是 $n \times n$ 矩阵的常规乘法. 该代数是结合的, 但不交换.
- 设 \mathcal{A} 为所有 $n \times n$ 矩阵的集合. 在它上面定义二元运算 $\bullet : \mathcal{A} \rightarrow \mathcal{A}$ 如下:

$$\mathbf{A} \bullet \mathbf{B} := \mathbf{AB} - \mathbf{BA}, \quad (3.3)$$

这里右边的乘法就是常规的矩阵乘法. 读者可以验证, \mathcal{A} 连带着上面的这个乘法就是个非结合、非交换的代数.

- 设 \mathcal{A} 为所有 $n \times n$ 上三角矩阵 (也就是所有对角线下方的元素都是零的矩阵) 的集合. 在常规的矩阵乘法下, 这个集合就变成一个非对易的结合代数. 读者自可验证其成立.
- 还是用 \mathcal{A} 表示所有上三角矩阵的集合. 此时按照『(3.3) 式』定义上面的二元运算. 读者可以验证, 此时 \mathcal{A} 连带着它上面的这个运算就是个非结合非交换的代数. 而 \mathcal{A} 的导出代数 \mathcal{A}^2 就由那些 $n \times n$ 的严格上三角矩阵 (*strictly upper triangular matrix*) 构成, 也就是其元素为那些对角元均为零的上三角矩阵.

⁴译者注: derived algebra 通常翻译为导代数. 但是由于 \mathcal{A}^2 本身是子代数, 因此也可能会有 derived subalgebra 这样的说法, 从而就有导子代数的自然译法. 可是后面我们会遇到 derivation algebra, 而 derivation 译作导子. 从而我们会有两个意思不同的导子代数. 为了避免这种混淆, 我们将 derived 译作“导出的”.

- 我们之前说过, 从 \mathcal{V} 到 \mathcal{W} 所有线性变换的集合是个矢量空间. 现在我们试着在它上面定义一个乘法. 而最佳候选人自然就是线性变换的复合. 若 $T : \mathcal{V} \rightarrow \mathcal{U}$ 且 $S : \mathcal{U} \rightarrow \mathcal{W}$ 均为线性变换, 则它们的复合 $S \circ T : \mathcal{V} \rightarrow \mathcal{W}$ 也是线性变换 (读者可以验证确实如此). 可是, 这个乘积并不是定义在单个矢量空间上, 而是将 $\mathcal{L}(\mathcal{V}, \mathcal{U})$ 中的元素, 和 $\mathcal{L}(\mathcal{U}, \mathcal{W})$ 中另一个元素, 变成 $\mathcal{L}(\mathcal{V}, \mathcal{W})$ 中的元素. 代数要求这里只有一个矢量空间, 而这可以通过让 $\mathcal{V} = \mathcal{U} = \mathcal{W}$ 实现. 从而这三个线性变换的空间就都坍缩为单个空间 $\mathcal{L}(\mathcal{V}, \mathcal{V})$, 也就是 \mathcal{V} 同态的集合, 之前我们说过用 $\mathcal{L}(\mathcal{V})$ 或者 $\text{End}(\mathcal{V})$ 予以缩写. 这样一来, $T, S, ST := S \circ T$ 和 $TS := T \circ S$ 就都属于 $\text{End}(\mathcal{V})$ 了.
- 上面所有的例子都是有限维代数. 而 $C^r(a, b)$ (即实区间 (a, b) 上所有直到 r 阶导数都存在的实值函数) 则是无限维代数的范例. 它的乘法是逐点定义的: 若 $f \in C^r(a, b)$ 且 $g \in C^r(a, b)$, 则

$$(fg)(t) := f(t)g(t) \quad \forall t \in (a, b).$$

这个代数是结合的交换代数, 并带有单位元 $f(t) = 1$.

- 无穷维代数的另一个例子就是多项式代数*. 这个代数是结合且交换的含幺代数.

* 很明显, 多项式代数不可能是有限维的.

定义 3.1.10 (代数的直和). 设 \mathcal{A} 和 \mathcal{B} 均为代数. 在它们的矢量空间直和 $\mathcal{A} \oplus \mathcal{B}$ 基础上定义下述乘积:

$$(\mathbf{a}_1 \oplus \mathbf{b}_1)(\mathbf{a}_2 \oplus \mathbf{b}_2) = (\mathbf{a}_1 \mathbf{a}_2) \oplus (\mathbf{b}_1 \mathbf{b}_2),$$

所得结果称作这两个代数的直和 (*algebra direct sum*).

注意, 若 $\mathbf{a} \in \mathcal{A}$, 则它同样可以视作是 $\mathcal{A} \oplus \mathcal{B}$ 中的元素 $\mathbf{a} \oplus \mathbf{0}$. 类似地, 若 $\mathbf{b} \in \mathcal{B}$, 则它同样可以视作是 $\mathcal{A} \oplus \mathcal{B}$ 中的元素 $\mathbf{0} \oplus \mathbf{b}$. 因此, \mathcal{A} 中任意元素与 \mathcal{B} 中任意元素之积就是零, 亦即 $\mathcal{AB} = \mathcal{BA} = \{\mathbf{0}\}$. 正如我们后面会看到的那样, 这一条件是给定代数是其子代数直和的必要条件.

现在研究 $\mathcal{A} \oplus \mathcal{B}$ 的中心. 若 $\mathbf{a} \oplus \mathbf{b} \in Z(\mathcal{A} \oplus \mathcal{B})$, 则对所有 $\mathbf{x} \in \mathcal{A}$ 以及 $\mathbf{y} \in \mathcal{B}$ 必有

$$(\mathbf{a} \oplus \mathbf{b})(\mathbf{x} \oplus \mathbf{y}) = (\mathbf{x} \oplus \mathbf{y})(\mathbf{a} \oplus \mathbf{b}),$$

换言之, 需要有

$$\mathbf{ax} \oplus \mathbf{by} = \mathbf{xa} \oplus \mathbf{yb} \Leftrightarrow (\mathbf{ax} - \mathbf{xa}) \oplus (\mathbf{by} - \mathbf{yb}) = \mathbf{0}.$$

他要成立, 我们就必须同时有

$$\mathbf{ax} - \mathbf{xa} = \mathbf{0} \quad \text{且} \quad \mathbf{by} - \mathbf{yb} = \mathbf{0},$$

这就是说 $\mathbf{a} \in \mathcal{Z}(\mathcal{A})$ 且 $\mathbf{b} \in \mathcal{Z}(\mathcal{B})$. 因此,

$$\mathcal{Z}(\mathcal{A} \oplus \mathcal{B}) = \mathcal{Z}(\mathcal{A}) \oplus \mathcal{Z}(\mathcal{B}). \quad (3.4)$$

定义 3.1.11 (代数的张量积). 设 \mathcal{A} 和 \mathcal{B} 均为代数, 在矢量空间张量积 $\mathcal{A} \otimes \mathcal{B}$ 基础上, 定义⁵

$$(\mathbf{a}_1 \otimes \mathbf{b}_1)(\mathbf{a}_2 \otimes \mathbf{b}_2) = \mathbf{a}_1 \mathbf{a}_2 \otimes \mathbf{b}_1 \mathbf{b}_2,$$

其所得结果就称作这两个代数的张量积 (*algebra tensor product*). 由于存在同构 $\mathcal{A} \otimes \mathcal{B} \cong \mathcal{B} \otimes \mathcal{A}$, 我们要求对所有 $\mathbf{a} \in \mathcal{A}$ 与 $\mathbf{b} \in \mathcal{B}$ 有 $\mathbf{a} \otimes \mathbf{b} = \mathbf{b} \otimes \mathbf{a}$.

当我们将给定代数写成其两个子代数 \mathcal{B} 和 \mathcal{C} 的张量积时, 上述定义中的最后一个条件就变得很重要了. 在这种情况下, \otimes 就和 \mathcal{A} 里面的乘法重合了, 从而那个条件就变成要求 \mathcal{B} 中所有元素和 \mathcal{C} 中所有元素对易, 即 $\mathcal{BC} = \mathcal{C}\mathcal{B}$.

定义 3.1.12 (结构常数). 给定代数 \mathcal{A} 以及其底矢量空间的一组基 $B = \{\mathbf{e}_i\}_{i=1}^N$, 我们就可以设

$$\mathbf{e}_i \mathbf{e}_j = \sum_{k=1}^N c_{ij}^k \mathbf{e}_k, \quad c_{ij}^k \in \mathbb{C}. \quad (3.5)$$

这里矢量 $\mathbf{e}_i \mathbf{e}_j$ 在基 B 下的分量 c_{ij}^k 就称作 \mathcal{A} 的结构常数 (*structure constants*).

结构常数决定了任意两个矢量之积——只要我们将其用 B 中的基矢表示出来. 反过来, 给定任意 N 维矢量空间 \mathcal{V} , 我们也可以将其变成一个代数: 只要选定一组基, 取 N^3 个标量 $\{c_{ij}^k\}$ 后用『(3.5) 式』定义这些基矢的乘积即可.

例 3.1.13. 设代数 \mathcal{A} 和 \mathcal{B} 在它们各自的基 $\{\mathbf{e}_i\}_{i=1}^M$ 和 $\{\mathbf{f}_n\}_{n=1}^N$ 下对应的结构常数分别为 $\{a_{ij}^k\}_{i,j,k=1}^M$ 和 $\{b_{mn}^l\}_{l,m,n=1}^N$. 从而

$$\begin{aligned} \mathbf{e}_i \mathbf{e}_j &= \sum_{i,j=1}^M a_{ij}^k \mathbf{e}_k, \\ \mathbf{f}_m \mathbf{f}_n &= \sum_{m,n=1}^N b_{mn}^l \mathbf{f}_l. \end{aligned}$$

⁵译者注: 再次注意, 不同于直和空间中的元素必然可以写成两个矢量的直和, 张量积空间中的元素只能说是写成张量积的线性组合. 当我们按照下面这种方式定义运算时, 实际上暗含了一个自然的线性扩张.

现在构造一个 MN 维的代数 \mathcal{C} , 它的基为 $\{\mathbf{v}_{kl}\}_{k,l=1}^{M,N}$, 对应这组基的结构常数为 $c_{im,jn}^{kl} = a_{ij}^k b_{mn}^l$. 从而

$$\mathbf{v}_{im}\mathbf{v}_{jn} = \sum_{i,j=1}^M \sum_{m,n=1}^N c_{im,jn}^{kl} \mathbf{v}_{kl} = \sum_{i,j=1}^M \sum_{m,n=1}^N a_{ij}^k b_{mn}^l \mathbf{v}_{kl}.$$

该代数与 $\mathcal{A} \otimes \mathcal{B}$ 同构. 事实上, 如果我们将右边的 \mathbf{v}_{kl} 等同于 $\mathbf{e}_k \otimes \mathbf{f}_l$, 则

$$\begin{aligned} \mathbf{v}_{im}\mathbf{v}_{jn} &= \sum_{i,j=1}^M \sum_{m,n=1}^N c_{im,jn}^{kl} \mathbf{e}_k \otimes \mathbf{f}_l = \sum_{i,j=1}^M \sum_{m,n=1}^N a_{ij}^k b_{mn}^l \mathbf{e}_k \otimes \mathbf{f}_l \\ &= \left(\sum_{i,j=1}^M a_{ij}^k \mathbf{e}_k \right) \otimes \left(\sum_{m,n=1}^N b_{mn}^l \mathbf{f}_l \right) = (\mathbf{e}_i \mathbf{e}_j) \otimes (\mathbf{f}_m \mathbf{f}_n), \end{aligned}$$

它与 $\mathbf{v}_{im} := \mathbf{e}_i \otimes \mathbf{f}_m$, $\mathbf{v}_{jn} := \mathbf{e}_j \otimes \mathbf{f}_n$ 以及两代数张量积的乘法规则一致.

定义 3.1.14 (可除代数). 所谓可除代数 (*division algebra*), 就是所有非零元素均存在逆的含幺代数.

例 3.1.15. 设 $\{\mathbf{e}_1, \mathbf{e}_2\}$ 是 \mathbb{R}^2 的一组基. 设这组基下的结构常数为

$$\begin{aligned} c_{11}^1 &= -c_{22}^1 = c_{12}^2 = c_{21}^2 = 1, \\ c_{12}^1 &= -c_{21}^1 = c_{11}^2 = c_{22}^2 = 0. \end{aligned}$$

那么, 它对应的乘法规则就是

$$\mathbf{e}_1^2 = -\mathbf{e}_2^2 = \mathbf{e}_1, \quad \mathbf{e}_1 \mathbf{e}_2 = \mathbf{e}_2 \mathbf{e}_1 = \mathbf{e}_2.$$

容易验证, 这样构造的代数不过就是 \mathbb{C} . 我们需要做的不过是将 \mathbf{e}_1 等同于 1, 而将 \mathbf{e}_2 等同于 $\sqrt{-1}$. 显然, \mathbb{C} 是可除代数.

例 3.1.16 (四元数代数). 在 \mathbb{R}^4 的标准基 $\{\mathbf{e}_i\}_{i=0}^3$ 下, 选定结构常数如下:

$$\begin{aligned} \mathbf{e}_0^2 &= -\mathbf{e}_1^2 = -\mathbf{e}_2^2 = -\mathbf{e}_3^2 = \mathbf{e}_0, \\ \mathbf{e}_0 \mathbf{e}_i &= \mathbf{e}_i \mathbf{e}_0 = \mathbf{e}_i \quad i = 1, 2, 3, \\ \mathbf{e}_i \mathbf{e}_j &= \sum_{k=1}^3 \varepsilon_{ijk} \mathbf{e}_k \quad i, j = 1, 2, 3 \text{ 且 } i \neq j. \end{aligned}$$

这里 ε_{ijk} 对它的所有下标是全反对称的 (因此任意两个下标相同时其结果就是零), 并且 $\varepsilon_{123} = 1$. ⁶ 读者可以验证这一关系使得 \mathbb{R}^4 成为非交换的结合代数. 这个代数称作四元数代数 (*algebra of quaternions*), 记作 \mathbb{H} . 在这个语境下, \mathbf{e}_0 通常记作 1, 而 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ 则分别记作 $\mathbf{i}, \mathbf{j}, \mathbf{k}$. 如

⁶译者注: 这个符号称作 Levi-Civita 符号 (Levi-Civita symbol). 在研究张量时还会再次提到它, 彼时它以张量的身份出现.

此一来, \mathbb{H} 中的任意元素就可以写作 $q = x + \mathbf{i}y + \mathbf{j}z + \mathbf{k}w$ 的形式. 很明显, \mathbb{H} 是 \mathbb{C} 的推广. 和 \mathbb{C} 相类比, 就将 x 称作 q 的实部 (*real part*), 将 (y, z, w) 称作 q 的纯部 (*pure part*).⁷ 类似地, q 的共轭 (*conjugate*) 就定义为 $q^* = x - \mathbf{i}y - \mathbf{j}z - \mathbf{k}w$.

我们可以将四元数方便地写作 $q = x_0 + \mathbf{x}$, 其中 \mathbf{x} 是个三维矢量. 那么其共轭就是 $q^* = x_0 - \mathbf{x}$. 不仅如此, 我们可以证明, 若 $q = x_0 + \mathbf{x}$, $p = y_0 + \mathbf{y}$, 那么

$$qp = \underbrace{x_0y_0 - \mathbf{x} \cdot \mathbf{y}}_{qp \text{ 的实部}} + \underbrace{x_0\mathbf{y} + y_0\mathbf{x} + \mathbf{x} \times \mathbf{y}}_{qp \text{ 的纯部}}. \quad (3.6)$$

在上式中将 \mathbf{x} 和 \mathbf{y} 分别替换为 $-\mathbf{x}$ 和 $-\mathbf{y}$, 我们就得到

$$q^* p^* = x_0y_0 - \mathbf{x} \cdot \mathbf{y} - x_0\mathbf{y} - y_0\mathbf{x} + \mathbf{x} \times \mathbf{y},$$

它并不等于 $(qp)^*$. 事实上, 容易证明 $(qp)^* = p^*q^*$.

在『(3.6) 式』中把 p 替换为 q^* , 我们就得到 $qq^* = x_0^2 + |\mathbf{x}|^2$. 与复数绝对值 (模长) 的定义相似, q 的绝对值 (*absolute value*) 就定义为 $|q| = \sqrt{qq^*}$. 如果 $q \neq 0$, 则 $q^*/(x_0^2 + |\mathbf{x}|^2)$ 就是 q 的逆. 因此, 四元数代数是可除代数.

最后, 不难证明

$$q^n = \sum_{k=0}^{[n/2]} (-1)^k \binom{n}{2k} x_0^{n-2k} |\mathbf{x}|^{2k} + \sum_{k=0}^{[n/2]} (-1)^k \binom{n}{2k+1} x_0^{n-2k-1} |\mathbf{x}|^{2k} \mathbf{x}, \quad (3.7)$$

这里 $[x]$ 表示不超过 x 的最小整数.⁸

现在我们来看看 $\mathcal{A} \otimes \mathcal{B}$ 的中心. 假设 $\mathbf{a} \otimes \mathbf{b}$ 属于其中心, 那么对所有 $\mathbf{x} \in \mathcal{A}$ 以及 $\mathbf{y} \in \mathcal{B}$ 我们必然有

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{x} \otimes \mathbf{y}) = (\mathbf{x} \otimes \mathbf{y})(\mathbf{a} \otimes \mathbf{b}),$$

或者说得有

$$\mathbf{ax} \otimes \mathbf{by} = \mathbf{xa} \otimes \mathbf{yb}.$$

要让上式成立, 我们必须让

$$\mathbf{ax} = \mathbf{xa} \quad \text{且} \quad \mathbf{by} = \mathbf{yb},$$

⁷译者注: 一些书中也称这里的纯部为虚部 (*imaginary part*). 也有将这对概念命名为标量部分 (*scalar part*) 和矢量部分 (*vector part*) 的. 事实上, 我们现在所述的矢量以及与之适配的矢量分析就是从四元数代数中分离出来的.

⁸译者注: 原书中把求和上限写作 $[n]/2$, 然后定义 n 为偶数时 $[n] = n$, n 为奇数时 $[n] = n - 1$. 这种写法和绝大多数数学书都不一致, 是故翻译时采用了常规的取整函数来写.

也就是说得有 $\mathbf{a} \in \mathcal{Z}(\mathcal{A})$ 且 $\mathbf{b} \in \mathcal{Z}(\mathcal{B})$. 最终我们就看到

$$\mathcal{Z}(\mathcal{A} \otimes \mathcal{B}) = \mathcal{Z}(\mathcal{A}) \otimes \mathcal{Z}(\mathcal{B}). \quad (3.8)$$

设 \mathcal{A} 是结合代数, S 是 \mathcal{A} 的子集. 如果 \mathcal{A} 中的每个元素都可以写成 S 里面一些元素乘积的线性组合, 则称 S 是 \mathcal{A} 的生成元 (generator). 很明显, 底矢量空间 \mathcal{A} 的任意一组基就是这个代数的一个生成元. 然而, 它并非是最小的生成元, 因为一些基矢可以写成其他基矢的乘积, 从而这个生成元可以进一步缩小. 比如说, 由三维矢量的叉积所给出的代数 (\mathbb{R}^3, \times) 里基 $\{\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y, \hat{\mathbf{e}}_z\}$ 就是一个生成元, 但它可以进一步缩减为 $\{\hat{\mathbf{e}}_x, \hat{\mathbf{e}}_y\}$ (或者其他任意一对基矢), 因为我们有 $\hat{\mathbf{e}}_z = \hat{\mathbf{e}}_x \times \hat{\mathbf{e}}_y$.

3.1.2. 同态

矢量空间之间通过线性变换相连接, 对线性变换稍微加以修饰, 使其适应于对应代数的二元乘法运算, 就可以得到下面的这个概念:

定义 3.1.17 (代数同态). 设 \mathcal{A}, \mathcal{B} 为代数, $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 是它们之间的线性映射.* 如果 $\phi(\mathbf{ab}) = \phi(a)\phi(\mathbf{b})$, 则称 ϕ 就是个代数同态 (*algebra homomorphism*). 类似地, 若代数同态为单射、满射、双射, 则分别对应称其为单同态 (*monomorphism*)、满同态 (*epimorphism*)、同构 (*isomorphism*). 映满到自身的代数同构就称作自同构 (*automorphism*).

例 3.1.18. 设 $\mathcal{A} = \mathbb{R}^3$, \mathcal{B} 为所有具有下述形式的矩阵:

$$\mathbf{A} = \begin{bmatrix} 0 & a_1 & -a_2 \\ -a_1 & 0 & a_3 \\ a_2 & -a_3 & 0 \end{bmatrix}.$$

现在定义映射 $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 如下:

$$\phi(\mathbf{a}) = \phi(a_1, a_2, a_3) = \begin{bmatrix} 0 & a_1 & -a_2 \\ -a_1 & 0 & a_3 \\ a_2 & -a_3 & 0 \end{bmatrix}.$$

可以证明, 这是个线性同构. 接下来将 $\mathcal{A} = \mathbb{R}^3$ 里的二元乘法定义为叉积, 这样 \mathcal{A} 就变成了一个代数. 对于 \mathcal{B} , 我们用『(3.3) 式』定义其二元乘法. 读者可以验证, 在这些运算下, ϕ 可以扩张为一个代数同构.

命题 3.1.19. 设 \mathcal{A} 和 \mathcal{B} 为代数, $\{\mathbf{e}_i\}$ 是 \mathcal{A} 的一组基, $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 为线性变换. 那么, ϕ 是代数同态的充要条件是

$$\phi(\mathbf{e}_i \mathbf{e}_j) = \phi(\mathbf{e}_i)\phi(\mathbf{e}_j).$$

证. 若 $\mathbf{a} = \sum_i \alpha_i \mathbf{e}_i$, $\mathbf{b} = \sum_j \beta_j \mathbf{e}_j$. 那么在命题中条件成立的前提下就有

$$\begin{aligned} \phi(\mathbf{ab}) &= \phi\left(\left(\sum_i \alpha_i \mathbf{e}_i\right)\left(\sum_j \beta_j \mathbf{e}_j\right)\right) = \phi\left(\sum_i \alpha_i \sum_j \beta_j \mathbf{e}_i \mathbf{e}_j\right) \\ &= \sum_i \alpha_i \sum_j \beta_j \phi(\mathbf{e}_i \mathbf{e}_j) = \sum_i \alpha_i \sum_j \beta_j \phi(\mathbf{e}_i)\phi(\mathbf{e}_j) \\ &= \sum_i \alpha_i \phi(\mathbf{e}_i) \sum_j \beta_j \phi(\mathbf{e}_j) = \phi\left(\sum_i \alpha_i \mathbf{e}_i\right) \phi\left(\sum_j \beta_j \mathbf{e}_j\right) \\ &= \phi(\mathbf{a})\phi(\mathbf{b}). \end{aligned}$$

从而 ϕ 就是代数同态, 由此充分性得证. 至于必要性, 直接取 \mathbf{a}, \mathbf{b} 为基矢即可. \square

* 我们通常用 ϕ, ψ 这样的希腊字母表示代数之间的线性映射, 而不像之前用 \mathbf{T}, \mathbf{U} 这样的符号.

例 3.1.20. 设 \mathcal{A} 和 \mathcal{B} 为代数, $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 为同态. 那么『定理 2.3.10』就保证了 $\phi(\mathcal{A})$ 是 \mathcal{B} 的子空间. 现在取 $\mathbf{b}_1, \mathbf{b}_2 \in \phi(\mathcal{A})$. 那么就存在 $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{A}$ 使得 $\mathbf{b}_1 = \phi(\mathbf{a}_1)$ 且 $\mathbf{b}_2 = \phi(\mathbf{a}_2)$. 由此进一步可以看到

$$\mathbf{b}_1 \mathbf{b}_2 = \phi(\mathbf{a}_1) \phi(\mathbf{a}_2) = \phi(\mathbf{a}_1 \mathbf{a}_2),$$

这就表明 $\mathbf{b}_1 \mathbf{b}_2 \in \phi(\mathcal{A})$. 因此, $\phi(\mathcal{A})$ 就是 \mathcal{B} 的子代数.

例 3.1.21 (\mathbb{R} (或 \mathbb{C}) 是任意含幺代数的子代数). 设 \mathcal{A} 为带单位元 $\mathbf{1}$ 的实代数. 定义线性映射 $\phi : \mathbb{R} \rightarrow \mathcal{A}$ 为 $\phi(\alpha) = \alpha \mathbf{1}$. 将 \mathbb{R} 视作它上面的代数, 我们就有

$$\phi(\alpha\beta) = \alpha\beta \mathbf{1} = (\alpha \mathbf{1})(\beta \mathbf{1}) = \phi(\alpha)\phi(\beta).$$

这就表明 ϕ 是个代数同态. 不仅如此, 若 $\phi(\alpha_1) = \phi(\alpha_2)$, 则有 $\alpha_1 \mathbf{1} = \alpha_2 \mathbf{1}$, 进而 $(\alpha_1 - \alpha_2) \mathbf{1} = \mathbf{0}$. 这就表明 $\alpha_1 = \alpha_2$. 因此, ϕ 是单同态. 于是乎, 我们就可以将 \mathbb{R} 等同于 $\phi(\mathbb{R})$, 从而将 \mathbb{R} 视作是 \mathcal{A} 的子代数. 这与我们在『例 3.1.4』中所得结论一致.

定义 3.1.22 (单位同态). 设 \mathcal{A} 和 \mathcal{B} 均为含幺代数. 如果同态 $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 保幺, 即有 $\phi(\mathbf{1}_A) = \mathbf{1}_B$, 则称其为幺同态 (*unital homomorphism*). ⁹

可以证明如下结论¹⁰:

命题 3.1.23. 设 \mathcal{A}, \mathcal{B} 均为含幺代数. 若 $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 是满同态, 则 ϕ 是幺同态.

『例 3.1.9』引入了 \mathcal{V} 上的自同态(算符)代数 $\mathcal{L}(\mathcal{V})$. 这个代数存在单位元 $\mathbf{1}$ (即恒等变换), 它让每个矢量自身保持不变.

定义 3.1.24 (对合). 设 \mathcal{V} 为矢量空间, 若其自同态 ω 的平方等于 $\mathbf{1}$, 则称其为对合 (*involution*).

特别地, $\mathbf{1} \in \text{End}(\mathcal{V})$ 就是个对合. 若 ω_1, ω_2 是满足 $\omega_1 \circ \omega_2 = \omega_2 \circ \omega_1$ 的两个对合, 则 $\omega_1 \circ \omega_2$ 同样是对合.

对于代数而言, 我们要求对合得是同态, 而不仅仅是线性映射. 设 \mathcal{A} 是个代数, 用 $\mathcal{H}(\mathcal{A})$ 表示 \mathcal{A} 同态的集合. 那么代数的对合 $\omega \in \mathcal{H}(\mathcal{A})$ 自然就需要满足 $\omega \circ \omega = \iota \in \mathcal{H}(\mathcal{A})$. ^{*} 现在假设 \mathcal{A} 带有单位元 \mathbf{e} , 那么 $\omega(\mathbf{e})$

⁹译者注: 为了防止和 identity homomorphism 混淆 (即映射 $\iota : \mathbf{x} \mapsto \mathbf{x}$), 这里采用了幺同态的译法, 而不是单位同态.

¹⁰译者注: 如果 ϕ 是满同态, 则任意 $\mathbf{y} \in \mathcal{B}$ 均存在原像 \mathbf{x} . 现在注意到 $\mathbf{y} = \phi(\mathbf{x}) = \phi(\mathbf{x}\mathbf{1}_A) = \phi(\mathbf{x})\phi(\mathbf{1}_A) = \mathbf{y}\phi(\mathbf{1}_A)$ 对任意 $\mathbf{y} \in \mathcal{B}$ 都成立. 同理可以证明 $\mathbf{y} = \phi(\mathbf{1}_A)\mathbf{y}$ 对所有 $\mathbf{y} \in \mathcal{B}$ 成立. 这就表明 $\phi(\mathbf{1}_A)$ 是 \mathcal{B} 的单位元, 而单位元唯一, 故 $\phi(\mathbf{1}_A) = \mathbf{1}_B$.

^{*} 为保持符号的一致性, 我们用 ι 表示代数 \mathcal{A} 上的恒等同态 (或者说单位同态、单位算符等).

就必须等于 \mathbf{e} . 实际上, 若令 $\omega(\mathbf{e}) = \mathbf{a}$, 那么由于 $\omega \circ \omega = \iota$ 我们自然就有 $\omega(\mathbf{a}) = \mathbf{e}$, 进而

$$\omega(\mathbf{ea}) = \omega(\mathbf{e})\omega(\mathbf{a}) = \omega(\mathbf{e})\mathbf{e} = \omega(\mathbf{e}).$$

现在两边同时作用 ω 就有 $\mathbf{ea} = \mathbf{e}$. 这仅在 $\mathbf{a} = \mathbf{e}$ 时才有可能.

定理 3.1.25. 设 \mathcal{U} 和 \mathcal{V} 是彼此同构的矢量空间. 则 $\mathcal{L}(\mathcal{U})$ 和 $\mathcal{L}(\mathcal{V})$ 作为代数也是同构的.

证. 设 $\phi : \mathcal{U} \rightarrow \mathcal{V}$ 是矢量空间的同构 (线性同构). 现在定义 $\Phi : \mathcal{L}(\mathcal{U}) \rightarrow \mathcal{L}(\mathcal{V})$ 如下:

$$\Phi(\mathbf{T}) = \phi \circ \mathbf{T} \circ \phi^{-1}.$$

容易验证 Φ 是个代数同构.¹¹

□

由『定理 3.1.25』以及『定理 2.3.20』可以得到这样一个结论: 任意实矢量空间 \mathcal{V} 上的线性变换代数 $\mathcal{L}(\mathcal{V})$ 都同构于 $\mathcal{L}(\mathbb{R}^N)$, 其中 $N = \dim \mathcal{V}$. 类似地, 若 \mathcal{V} 是 N 维复矢量空间, 则 $\mathcal{L}(\mathcal{V})$ 就同构于 $\mathcal{L}(\mathbb{C}^N)$.

¹¹译者注: 线性变换的复合自然还是线性变换, 因此关键是证明 Φ 保乘法. 而这就要用到一个经典技巧了: 插入单位算符. 以这里的问题为例, 我们就有 $\Phi(\mathbf{TS}) = \phi \circ \mathbf{TS} \circ \phi^{-1}$, 而 $\mathbf{TS} = \mathbf{T}\mathbf{1}_U \mathbf{S}$, 又由于 $\mathbf{1}_U = \phi^{-1} \circ \phi$, 将其代入我们就发现 $\Phi(\mathbf{TS}) = \Phi(\mathbf{T})\Phi(\mathbf{S})$.

3.2. 理想

子代数是那些在与自身元素相乘下不变的子空间,¹² 也就是说子代数中元素的乘积不会离开这个子代数. 而在代数理论中更为重要的则是那些在于整个代数的元素相乘下不变的子空间.

定义 3.2.1 (理想). 设 A 为代数, B 是其子空间. 若对所有的 $\mathbf{a} \in A$ 以及 $\mathbf{b} \in B$, 均有 $\mathbf{ab} \in B$, 则称 B 是 A 的左理想 (*left ideal*). 利用『(3.1) 式』可以将其表述为 $AB \subseteq B$. 类似地, 若有 $BA \subseteq B$, 则称 B 是 A 的右理想 (*right ideal*). 若 B 同时是 A 的左理想和右理想, 则称其为双边理想 (*two-sided ideal*), 亦或者简称为理想 (*ideal*).

由定义很明显可以看出理想必然是子代数, 并且对于含幺代数而言, 仅有一个可以包含单位元 (或者可逆元) 的理想, 那就是这个代数子集.

例 3.2.2. 设 A 为结合代数且 $\mathbf{a} \in A$. 用 $\mathcal{L}(\mathbf{a})$ 表示那些使得 $\mathbf{x}\mathbf{a} = \mathbf{0}$ 的元素 $\mathbf{x} \in A$. 对任意 $\mathbf{x} \in \mathcal{L}(\mathbf{a})$ 以及任意 $\mathbf{y} \in A$, 我们有

$$(\mathbf{yx})\mathbf{a} = \mathbf{y}(\mathbf{x}\mathbf{a}) = \mathbf{0},$$

从而 $\mathbf{yx} \in \mathcal{L}(\mathbf{a})$. 由此可知 $\mathcal{L}(\mathbf{a})$ 是 A 的左理想. 称其为 \mathbf{a} 的左零化子 (*left annihilator of \mathbf{a}*). 类似可构造 \mathbf{a} 的右零化子 $\mathcal{R}(\mathbf{a})$.

例 3.2.3. 设 $C^r(a, b)$ 是区间 (a, b) 上所有 r 次可微实值函数构成的代数 (见『例 3.1.9』). 在给定的固定点 $c \in (a, b)$ 处为零的所有函数就构成了 $C^r(a, b)$ 的一个理想. 由于这个代数是交换的, 是故该理想是双边理想.

更一般的来说, 用 M_n 表示矩阵元 $f_{ij} \in C^r(a, b)$ 的矩阵构成的 (非交换) 代数. 则以在给定点 $c \in (a, b)$ 处等于零的函数为矩阵元的矩阵就构成 M_n 的一个双边理想.

设 A 和 B 为代数, $\phi : A \rightarrow B$ 为同态. 根据『定理 2.3.9』可知 $\ker \phi$ 是 A 的子空间. 现在取 $\mathbf{x} \in \ker \phi$ 和 $\mathbf{a} \in A$. 那么就有

$$\phi(\mathbf{x}\mathbf{a}) = \phi(\mathbf{x})\phi(\mathbf{a}) = \mathbf{0}\phi(\mathbf{a}) = \mathbf{0},$$

这就说明 $\mathbf{x}\mathbf{a} \in \ker \phi$. 由此可知 $\ker \phi$ 是 A 的一个右理想.

¹²译者注: 不变 (invariant) 是个重要的概念: 通常而言, 给定一个映射 $f : A \rightarrow A$ 以及 A 的子集 B , 若有 $f(B) \subseteq B$, 则称 B 在 f 下不变, 或称 B 是 A 的一个 f 不变子集. 更进一步, 我们可以讨论某个映射集 $\mathcal{F} = \{f_\alpha : A \rightarrow A\}_{\alpha \in I}$ 下的不变子集 B : 我们要求对任意 $\alpha \in I$ 均有 $f_\alpha(B) \subseteq B$. 借用这个概念我们重新审视子代数的概念就得到了这里的说法: 考虑任意 $\mathbf{a} \in B \subseteq A$. 若 B 是 A 的子代数, 则它需要对乘法封闭, 于是 B 就是 \mathbf{a} 诱导的映射 $\mathbf{x} \mapsto \mathbf{ax}$ 以及 $\mathbf{x} \mapsto \mathbf{xa}$ 下的不变子集. 原书中采用 stable (稳定) 这个说法, 但这个说法翻译出来总觉得怪怪的, 是故改用不变这一术语.

定理 3.2.4. 设 $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 是代数同态, 则 $\ker \phi$ 是 \mathcal{A} 的(双边)理想.

我们很容易就能构造出结合代数 \mathcal{A} 的左理想: 取任意 $x \in \mathcal{A}$, 然后考察集合

$$\mathcal{A}x := \{ax \mid a \in \mathcal{A}\}.$$

读者可以验证 $\mathcal{A}x$ 确实是左理想.¹³ 类似地, $x\mathcal{A}$ 就是右理想, 而集合

$$\mathcal{A}x\mathcal{A} := \{axx \mid a, b \in \mathcal{A}\}$$

是双边理想. 我们分别称它们为 x 生成的 (generated by x) 左理想、右理想、双边理想.

定义 3.2.5 (极小理想). 设 \mathcal{M} 为代数 \mathcal{A} 的左(右、双边)理想, 若含于 \mathcal{M} 的每个 \mathcal{A} 的左(右、双边)理想都与 \mathcal{M} 重合, 则称 \mathcal{M} 是 \mathcal{A} 的极小理想.

定理 3.2.6. 设 \mathcal{L} 是 \mathcal{A} 的左理想, 则下面的命题等价:

- (a) \mathcal{L} 是极小左理想.
- (b) 对所有 $x \in \mathcal{L}$ 均有 $\mathcal{A}x = \mathcal{L}$.
- (c) 对所有 $x \in \mathcal{L}$ 均有 $\mathcal{L}x = \mathcal{L}$.

对极小右理想也有类似条件.

证. 由理想和极小理想的定义直接可以得到该结果.¹⁴

□

定理 3.2.7. 设 \mathcal{A}, \mathcal{B} 是代数, $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 是满同态, \mathcal{L} 是 \mathcal{A} 的一个(极小)左理想. 那么, $\phi(\mathcal{L})$ 就是 \mathcal{B} 的(极小)左理想. 特别地, 代数的任意自同构都是其极小理想之间的同构.

证. 设 b 是 \mathcal{B} 的任意元素, y 是 $\phi(\mathcal{L})$ 的任意元素. 那么就存在 $a \in \mathcal{A}$ 以及 $x \in \mathcal{L}$ 使得 $b = \phi(a)$, $y = \phi(x)$. 进一步就有

$$by = \phi(a)\phi(x) = \phi(ax) \in \phi(\mathcal{L}),$$

这里最后的属于关系是因为 \mathcal{L} 是左理想, 从而 $ax \in \mathcal{L}$. 由此可知 $\phi(\mathcal{L})$ 就是 \mathcal{B} 的左理想.

¹³译者注: 任取 $a \in \mathcal{A}$ 以及 $bx \in \mathcal{A}x$. 则因 \mathcal{A} 是代数, 故存在 $c \in \mathcal{A}$ 使得 $ab = c$. 进而 $a(bx) = (ab)x = cx \in \mathcal{A}x$. 这就说明 $\mathcal{A}(\mathcal{A}x) \subseteq \mathcal{A}x$, 即 $\mathcal{A}x$ 是左理想.

¹⁴译者注: 若 \mathcal{L} 是左理想, 则立刻得到 $\mathcal{A}x \subseteq \mathcal{L}$; 另外, 理想必为子代数, 进而 $\mathcal{L}x \subseteq \mathcal{L}$. 不仅如此, 我们断言此时 $\mathcal{L}x$ 也是左理想, $\mathcal{A}\mathcal{L} \subseteq \mathcal{L}$, 进而 $\mathcal{A}(\mathcal{L}x) \subseteq \mathcal{L}x$. 现在假设 \mathcal{L} 是极小的, 那么由于 $\mathcal{L}x \subseteq \mathcal{L}$ 是左理想, 极小条件就要求 $\mathcal{L}x = \mathcal{L}$. 这样我们就证明了 (a) \Rightarrow (c). 现在假设 $\mathcal{L}x = \mathcal{L}$, 则因 $\mathcal{L} \subseteq \mathcal{A}$ 就有 $\mathcal{L} = \mathcal{L}x \subseteq \mathcal{A}x \subseteq \mathcal{L}$. 这就说明中间的包含都是相等, 从而 $\mathcal{L} = \mathcal{A}x$. 这就证明了 (c) \Rightarrow (b). 最后, 假设 $\mathcal{A}x = \mathcal{L}$, 而 $\mathcal{I} \subseteq \mathcal{L}$ 是个左理想. 我们现在证明 $\mathcal{I} = \mathcal{L}$, 这只需证明 $\mathcal{L} \subseteq \mathcal{I}$. 任取 $y \in \mathcal{I}$, 则 $y \in \mathcal{L}$ 从而 $\mathcal{L} = \mathcal{A}y \subseteq \mathcal{I}$. 这样我们就证明了 (b) \Rightarrow (a).

现在进一步假设 \mathcal{L} 是极小的. 由于 ϕ 是满同态, 我们就有 $\mathcal{B} = \phi(\mathcal{A})$. 现在取 $\mathbf{u} \in \phi(\mathcal{L})$, 则存在 $\mathbf{t} \in \mathcal{L}$ 使得 $\mathbf{u} = \phi(\mathbf{t})$, 从而

$$\mathcal{B}\mathbf{u} = \phi(\mathcal{A})\phi(\mathbf{t}) = \phi(\mathcal{A}\mathbf{t}) = \phi(\mathcal{L}).$$

这里我们用到了『定理 3.2.6』中的 $(a) \Rightarrow (b)$ 以给出 $\mathcal{A}\mathbf{t} = \mathcal{L}$. 另一方面, 上式也指出 $\mathcal{B}\mathbf{u} = \phi(\mathcal{L})$ 对任意 $\mathbf{u} \in \phi(\mathcal{L})$ 成立, 反过来利用『定理 3.2.6』中的 $(b) \Rightarrow (a)$ 则可以证得 $\phi(\mathcal{L})$ 是极小的.

这个定理的最后一个结论则由 $\ker \phi$ 是 \mathcal{A} 的理想(『定理 3.2.4』)这一事实得出. \square

定义 3.2.8(可约代数). 设 \mathcal{B}, \mathcal{C} 是代数 \mathcal{A} 的子代数. 如果作为矢量空间时我们有 $\mathcal{A} = \mathcal{B} \oplus \mathcal{C}$, 并且 $\mathcal{B}\mathcal{C} = \mathcal{C}\mathcal{B} = \{\mathbf{0}\}$, 则称 \mathcal{A} 是子代数 \mathcal{B} 和 \mathcal{C} 的直和(*direct sum of subalgebras*). 此时将 \mathcal{B} 和 \mathcal{C} 称作 \mathcal{A} 的分量(*components*). 很明显, 代数可以有多个分量. 如果一个代数可以写成其子代数的直和, 则称其可约(*reducible*).

正如我们在『定义 3.1.10』下所述, 如果要将 \mathcal{B} 和 \mathcal{C} 自然地分别等同于 $\mathcal{B} \oplus \{\mathbf{0}\}$ 和 $\{\mathbf{0}\} \oplus \mathcal{C}$, 那么 $\mathcal{B}\mathcal{C} = \mathcal{C}\mathcal{B} = \{\mathbf{0}\}$ 这个条件就是必要的.

命题 3.2.9. 中心代数一定不可约.

证. 假设中心代数(注意它一定是含幺代数) \mathcal{A} 可约. 那么单位元就必然在 \mathcal{A} 的每个构成子代数中都存在分量. 很明显, 这些分量是线性无关的, 并且都属于中心. 这就矛盾了!¹⁵ \square

例 3.2.10. 考虑『例 3.1.7』中所述代数 \mathcal{S} . 现在构造它的一组新基:^{*}

$$\begin{aligned} \mathbf{f}_1 &= \frac{1}{2}(\mathbf{e}_0 + \mathbf{e}_3), & \mathbf{f}_2 &= \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_2), \\ \mathbf{f}_3 &= \frac{1}{2}(\mathbf{e}_0 - \mathbf{e}_3), & \mathbf{f}_4 &= \frac{1}{2}(\mathbf{e}_1 + \mathbf{e}_2). \end{aligned} \quad (3.9)$$

* 建议读者证明这里的 $\{\mathbf{f}_i\}_{i=1}^4$ 是线性无关矢量组.

在这组新基下, \mathcal{S} 对应的乘法表可见『表 3.2』, 读者可以自行验证.

在恒等式 $\mathbf{e}_0 = \mathbf{f}_1 + \mathbf{f}_3$ 两边同时乘以 \mathcal{S} 中的任意元素, 我们就看到, 任意元素都可以写成左理想 $\mathcal{L}_1 := \mathcal{S}\mathbf{f}_1$ 中一个矢量, 与左理想 $\mathcal{L}_3 := \mathcal{S}\mathbf{f}_3$ 中一个矢量之和. 而 \mathcal{L}_1 中的任意矢量都可以写作 \mathcal{S} 中某个矢量与 \mathbf{f}_1 之

¹⁵译者注: 设 $\mathcal{A} = \mathcal{B} \oplus \mathcal{C}$, 且 $\mathbf{1} = \mathbf{b} + \mathbf{c}$, 其中 $\mathbf{b} \in \mathcal{B}, \mathbf{c} \in \mathcal{C}$. 则对任意 $\mathbf{x} \in \mathcal{A}$ 就有 $(\mathbf{b} + \mathbf{c})\mathbf{x} = \mathbf{x}(\mathbf{b} + \mathbf{c})$, 即 $\mathbf{bx} + \mathbf{cx} = \mathbf{xb} + \mathbf{xc}$. 亦或者说 $\mathbf{0} = (\mathbf{xb} - \mathbf{bx}) + (\mathbf{xc} - \mathbf{cx})$. 接下来我们注意这样一个事实: 若 $\mathcal{A} = \mathcal{B} \oplus \mathcal{C}$, 则 \mathcal{B}, \mathcal{C} 一定是 \mathcal{A} 的双边理想. 这是因为对任意 $\mathbf{a} \in \mathcal{A}$ 以及 $\mathbf{b} \in \mathcal{B}$, 我们可以将 \mathbf{a} 写成 $\mathbf{a} = \mathbf{a}_B + \mathbf{a}_C$, 其中 $\mathbf{a}_B \in \mathcal{B}, \mathbf{a}_C \in \mathcal{C}$. 从而 $\mathbf{ab} = \mathbf{a}_B\mathbf{b} + \mathbf{a}_C\mathbf{b} = \mathbf{a}_B\mathbf{b} \in \mathcal{B}$. 这里用到了 $\mathcal{B}\mathcal{C} = \{\mathbf{0}\}$ 的条件. 类似可证 $\mathbf{ba} = \mathbf{b}\mathbf{a}_B \in \mathcal{B}$. 有了这个结果我们就看到前面的 $\mathbf{xb} - \mathbf{bx} \in \mathcal{B}$ 而 $\mathbf{xc} - \mathbf{cx} \in \mathcal{C}$. 这就给出了 $\mathbf{0}$ 的直和分解, 而我们有分解 $\mathbf{0} = \mathbf{0} + \mathbf{0}$, 利用分解的唯一性可知必有 $\mathbf{xb} = \mathbf{bx}$ 以及 $\mathbf{xc} = \mathbf{cx}$. 这就说明 $\mathbf{1}$ 的分量必然属于 \mathcal{A} 的中心. 然而 \mathcal{A} 的中心里面不存在线性无关的元素, 从而这里就得到了矛盾.

	\mathbf{f}_1	\mathbf{f}_2	\mathbf{f}_3	\mathbf{f}_4
\mathbf{f}_1	\mathbf{f}_1	\mathbf{f}_2	$\mathbf{0}$	$\mathbf{0}$
\mathbf{f}_2	$\mathbf{0}$	$\mathbf{0}$	\mathbf{f}_2	\mathbf{f}_1
\mathbf{f}_3	$\mathbf{0}$	$\mathbf{0}$	\mathbf{f}_3	\mathbf{f}_4
\mathbf{f}_4	\mathbf{f}_4	\mathbf{f}_3	$\mathbf{0}$	$\mathbf{0}$

表 3.2: \mathcal{S} 的新乘法表

积. 现在设 $\mathbf{a} = \sum_{i=1}^4 \alpha_i \mathbf{f}_i$ 是 S 的任意元素. 那么 \mathcal{L}_1 中的任意矢量就可写作下述形式:

$$\mathbf{a}\mathbf{f}_1 = (\alpha_1 \mathbf{f}_1 + \alpha_2 \mathbf{f}_2 + \alpha_3 \mathbf{f}_3 + \alpha_4 \mathbf{f}_4)\mathbf{f}_1 = \alpha_1 \mathbf{f}_1 + \alpha_4 \mathbf{f}_4,$$

也就是说 \mathbf{f}_1 和 \mathbf{f}_4 就张成了 \mathcal{L}_1 . 类似地, \mathcal{L}_3 就由 \mathbf{f}_2 和 \mathbf{f}_3 张成. 由此可得 $\mathcal{L}_1 \cap \mathcal{L}_3 = \{\mathbf{0}\}$. 因此, 我们就有

$$S = \mathcal{L}_1 \oplus_V \mathcal{L}_3, \quad \mathcal{L}_1 = \text{Span}\{\mathbf{f}_1, \mathbf{f}_4\}, \quad \mathcal{L}_3 = \text{Span}\{\mathbf{f}_2, \mathbf{f}_3\},$$

这里 \oplus_V 表示的是矢量空间的直和. 请注意, 这个直和分解和 S 为代数并不矛盾. 这是因为上面的这个直和并不是代数的直和, 原因在于 $\mathcal{L}_1 \mathcal{L}_3 \neq \{\mathbf{0}\}$.

设 $\mathbf{x} = \gamma_1 \mathbf{f}_1 + \gamma_4 \mathbf{f}_4$ 是 \mathcal{L}_1 中的任意非零元素. 那么自然有 $S\mathbf{x} \subseteq \mathcal{L}_1$. 为证明 $\mathcal{L}_1 \subseteq S\mathbf{x}$, 我们取 \mathcal{L}_1 中的 $\mathbf{y} = \beta_1 \mathbf{f}_1 + \beta_4 \mathbf{f}_4$. 那么是否存在 $\mathbf{z} \in S$ 使得 $\mathbf{y} = \mathbf{z}\mathbf{x}$ 呢? 为此我们设 $\mathbf{z} = \sum_{i=1}^4 \eta_i \mathbf{f}_i$, 然后注意

$$\begin{aligned} \mathbf{z}\mathbf{x} &= (\eta_1 \mathbf{f}_1 + \eta_2 \mathbf{f}_2 + \eta_3 \mathbf{f}_3 + \eta_4 \mathbf{f}_4)(\gamma_1 \mathbf{f}_1 + \gamma_4 \mathbf{f}_4) \\ &= (\eta_1 \gamma_1 + \eta_2 \gamma_4) \mathbf{f}_1 + (\eta_3 \gamma_4 + \eta_4 \gamma_1) \mathbf{f}_4. \end{aligned}$$

于是我们就需要找到一组 η 使得

$$\eta_1 \gamma_1 + \eta_2 \gamma_4 = \beta_1 \quad \text{且} \quad \eta_3 \gamma_4 + \eta_4 \gamma_1 = \beta_4.$$

如果 $\gamma_1 \neq 0$, 那么 $\eta_1 = \beta_1/\gamma_1$, $\eta_2 = 0 = \eta_3$, $\eta_4 = \beta_4/\gamma_1$ 就给出了 \mathbf{z} 的一组解. 如果 $\gamma_4 \neq 0$, 则 $\eta_2 = \beta_1/\gamma_4$, $\eta_1 = 0 = \eta_4$, $\eta_3 = \beta_3/\gamma_4$ 就给出了 \mathbf{z} 的一组解. 于是就得到了 $\mathcal{L}_1 = S\mathbf{x}$, 进而根据『定理 3.2.6』可知 \mathcal{L}_1 是极小的. 类似地, \mathcal{L}_3 同样是极小的.

如果 $\mathcal{A} = \mathcal{B} \oplus \mathcal{C}$, 那么在其两边同时右乘以 \mathcal{B} 就得到

$$\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{B} \oplus \mathcal{C}\mathcal{B} = \mathcal{B}\mathcal{B} \oplus \{\mathbf{0}\} = \mathcal{B}\mathcal{B} \subset \mathcal{B},$$

这就表明 \mathcal{B} 是 \mathcal{A} 的左理想. 类似地, 在两边左乘以 \mathcal{B} 就表明 \mathcal{B} 同样是 \mathcal{A} 的右理想. 是故 \mathcal{B} 是 \mathcal{A} 的理想. 类似地, \mathcal{C} 也是 \mathcal{A} 的理想. 不仅如此, 因为这些子代数不共享任何非零元素, 从而 \mathcal{A} 的任意理想都必须含于这些子代数. 由此就得到下面的结论:

命题 3.2.11. 若 \mathcal{A} 可以写作代数的直和, 则它的每个分量(或者一些分量的直和)都是 \mathcal{A} 的理想. 不仅如此, \mathcal{A} 的任意其他理想都必须整个包含在其中某个分量内.

那些没有真理想¹⁶的代数在对所有代数进行分类时至关重要.

定义 3.2.12 (单代数). 若代数 \mathcal{A} 的理想仅有 \mathcal{A} 和 $\{\mathbf{0}\}$, 则称其为单(纯)代数 (*simple algebra*).

请注意, 当我们用理想这个词的时候都是指双边理想. 因此, 单代数可以有真左理想和真右理想. 下面的这个例子就说明了这一点.

例 3.2.13. 现在让我们回到『例 3.2.10』中的代数 \mathcal{S} 上, 在那个例子中我们看到 $\mathcal{S} = \mathcal{L}_1 \oplus \mathcal{L}_3$, 其中 $\mathcal{L}_1, \mathcal{L}_3$ 是极小左理想, 而 \oplus_V 表示这是矢量空间直和. 那么 \mathcal{S} 是否存在真双边理想呢? 现在设 \mathcal{J} 是这样的一个理想, 并设 $\mathbf{a} \in \mathcal{J}$ 非零. 利用 \mathcal{S} 的分解就可以将 \mathbf{a} 写作 $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_3$, 其中 $\mathbf{a}_1 \in \mathcal{L}_1, \mathbf{a}_3 \in \mathcal{L}_3$, 并且 $\mathbf{a}_1, \mathbf{a}_3$ 里面至少有一个非零. 假设 $\mathbf{a}_1 \neq \mathbf{0}$. 那么 $\mathcal{S}\mathbf{a}_1$ 就是个非零左理想, 并含于 \mathcal{L}_1 . 由于 \mathcal{L}_1 是极小的, 是故 $\mathcal{S}\mathbf{a}_1 = \mathcal{L}_1$. 由于 $\mathbf{f}_1 \in \mathcal{L}_1$, 进而必然存在 $\mathbf{b} \in \mathcal{S}$ 使得 $\mathbf{b}\mathbf{a}_1 = \mathbf{f}_1$, 从而

$$\mathbf{b}\mathbf{a} = \mathbf{b}\mathbf{a}_1 + \mathbf{b}\mathbf{a}_3 = \mathbf{f}_1 + \mathbf{b}\mathbf{a}_3.$$

在它两边同时右乘以 \mathbf{f}_1 , 并注意到 $\mathbf{f}_1^2 = \mathbf{f}_1$ 以及 $\mathcal{L}_3\mathbf{f}_1 = \{\mathbf{0}\}$ (见『表 3.2』), 我们就得到 $\mathbf{b}\mathbf{a}\mathbf{f}_1 = \mathbf{f}_1$. 由于 \mathcal{J} 是双边理想而 $\mathbf{a} \in \mathcal{J}$, 我们就有 $\mathbf{b}\mathbf{a}\mathbf{f}_1 \in \mathcal{J}$, 从而 $\mathbf{f}_1 \in \mathcal{J}$.

等式 $\mathcal{S}\mathbf{a}_1 = \mathcal{L}_1$ 还可以推出存在 $\mathbf{c} \in \mathcal{S}$ 使得 $\mathbf{c}\mathbf{a}_1 = \mathbf{f}_4$, 从而

$$\mathbf{c}\mathbf{a} = \mathbf{c}\mathbf{a}_1 + \mathbf{c}\mathbf{a}_3 = \mathbf{f}_4 + \mathbf{c}\mathbf{a}_3.$$

在它两边同时右乘 \mathbf{f}_1 , 并注意 $\mathbf{f}_4\mathbf{f}_1 = \mathbf{f}_4$ 以及 $\mathcal{L}_3\mathbf{f}_1 = \{\mathbf{0}\}$, 我们就得到 $\mathbf{c}\mathbf{a}\mathbf{f}_1 = \mathbf{f}_4$. 由于 \mathcal{J} 是双边理想, 我们就必须有 $\mathbf{f}_4 \in \mathcal{J}$. 考虑到 $\mathbf{f}_1\mathbf{f}_2 = \mathbf{f}_2, \mathbf{f}_4\mathbf{f}_2 = \mathbf{f}_3$, 这就说明所有的基矢都在 \mathcal{J} 中. 因此, $\mathcal{J} = \mathcal{S}$. 当 $\mathbf{a}_3 \neq \mathbf{0}$ 时我们可以得到同样的结果. 是故 \mathcal{S} 没有真理想, 也就是说 \mathcal{S} 是单代数.

利用『定义 3.2.12』以及『定理 3.2.4』立刻可以得到下面的结果:

命题 3.2.14. 设 \mathcal{A} 是单代数, \mathcal{B} 是任意其他代数, 则从 \mathcal{A} 到 \mathcal{B} 的非平凡同态¹⁷必然是单射.

证. 对任意同态 $\phi: \mathcal{A} \rightarrow \mathcal{B}$, 我们知道 $\ker \phi$ 是 \mathcal{A} 的理想. 由于 \mathcal{A} 没有真理想, 故 $\ker \phi = \mathcal{A}$ 或者 $\ker \phi = \{\mathbf{0}\}$. 如果 ϕ 非平凡, 则 $\ker \phi \neq \mathcal{A}$, 于是 $\ker \phi = \{\mathbf{0}\}$. 这就说明 ϕ 是单射. \square

¹⁶译者注: 很明显 \mathcal{A} 和 $\{\mathbf{0}\}$ 都是 \mathcal{A} 的理想, 这两个理想称作平凡理想 (trivial ideal). 有些书中将不等于 \mathcal{A} 的理想称作真理想, 这样命名是继承了真子集的说法 (同样还有真子空间的说法). 但还有些书将非平凡的理想称作真理想, 也就是我们还进一步要求真理想非零. 本书没有严格定义过这个概念, 使用时偶尔会混用两种观点.

¹⁷译者注: 很明显 $\phi: \mathbf{a} \mapsto \mathbf{0}$ 是个同态, 这个同态就是所谓平凡同态 (trivial homomorphism).

3.2.1. 商代数

设 \mathcal{A} 是个代数, \mathcal{B} 是 \mathcal{A} 的子代数. 在『2.1.2节』中我们展示了如何去构造商空间 \mathcal{A}/\mathcal{B} . 那么这个商空间是否可以变成代数呢? 我们设 $[\![\mathbf{a}]\!]$ 和 $[\![\mathbf{a}']\!]$ 属于 \mathcal{A}/\mathcal{B} . 那么使得 \mathcal{A}/\mathcal{B} 变成代数最自然的乘法法则就是

$$[\![\mathbf{a}]\!] [\![\mathbf{a}']\!] = [\![\mathbf{aa}']\!]. \quad (3.10)$$

那么怎样的条件可以保证上述乘法有意义呢? 由于对所有的 $\mathbf{b}, \mathbf{b}' \in \mathcal{B}$ 有 $[\![\mathbf{a}]\!] = [\![\mathbf{a} + \mathbf{b}]\!]$ 和 $[\![\mathbf{a}']\!] = [\![\mathbf{a}' + \mathbf{b}']\!]$, 要让『(3.10)式』成立, 我们就得有

$$(\mathbf{a} + \mathbf{b})(\mathbf{a}' + \mathbf{b}') = \mathbf{aa}' + \mathbf{b''},$$

其中 $\mathbf{b''} \in \mathcal{B}$. 取 $\mathbf{a} = \mathbf{0} = \mathbf{a}'$ 就得到 $\mathbf{bb'} = \mathbf{b''}$. 这就说明 \mathcal{B} 必须得是 \mathcal{A} 的一个子代数. 取 $\mathbf{a}' = \mathbf{0}$ 就得到 $\mathbf{ab'} + \mathbf{bb'} = \mathbf{b''}$ 对所有 $\mathbf{a} \in \mathcal{A}, \mathbf{b}, \mathbf{b}' \in \mathcal{B}$ 以及某个 $\mathbf{b''} \in \mathcal{B}$ 成立, 而这就意味着 \mathcal{B} 必须是 \mathcal{A} 的左理想. 类似地, 令 $\mathbf{a} = \mathbf{0}$ 就能推出 \mathcal{B} 必须是 \mathcal{A} 的右理想. 综上所述, 我们就得到下面的结论:

命题 3.2.15. 设 \mathcal{A} 是个代数, \mathcal{B} 是其子空间. 则商空间 \mathcal{A}/\mathcal{B} 在乘法 $[\![\mathbf{a}]\!] [\![\mathbf{a}']\!] = [\![\mathbf{aa}']\!]$ 下可以成为代数的充要条件就是 \mathcal{B} 为 \mathcal{A} 的理想. 这样构造所得代数就称作 \mathcal{A} 关于理想 \mathcal{B} 的商代数 (*quotient algebra/factor algebra*).

例 3.2.16. 设 \mathcal{A} 和 \mathcal{B} 为代数, $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 是个代数同态. 那么『例 3.1.20』以及『定理 3.2.4』就表明 $\phi(\mathcal{A})$ 是 \mathcal{B} 的一个子代数, $\ker \phi$ 是 \mathcal{A} 的理想. 现在考虑『例 2.3.22』中定义的线性映射 $\bar{\phi} : \mathcal{A}/\ker \phi \rightarrow \phi(\mathcal{A})$, 它满足 $\bar{\phi}([\![\mathbf{a}]\!]) = \phi(\mathbf{a})$. 我们很容易就可以证明 $\bar{\phi}$ 是个代数同态. 而『例 2.3.22』又指出 $\bar{\phi}$ 是线性同构, 二者相结合就说明 $\bar{\phi}$ 是代数同构.

3.3. 全矩阵代数

考察 $n \times n$ 矩阵的矢量空间, 并取其标准基 $\{\mathbf{e}_{ij}\}_{i,j=1}^n$, 这里 \mathbf{e}_{ij} 指在 (i, j) 位置为 1, 其余位置均等于零的矩阵. 这也就是说, $(\mathbf{e}_{ij})_{lk} = \delta_{il}\delta_{jk}$, 并且

$$\begin{aligned} (\mathbf{e}_{ij}\mathbf{e}_{kl})_{mn} &= \sum_{r=1}^n (\mathbf{e}_{ij})_{mr}(\mathbf{e}_{kl})_{rn} = \sum_{r=1}^n \delta_{im}\delta_{jr}\delta_{kr}\delta_{ln} \\ &= \delta_{im}\delta_{jk}\delta_{ln} = \delta_{jk}(\mathbf{e}_{il})_{mn}, \end{aligned}$$

这就说明

$$\mathbf{e}_{ij}\mathbf{e}_{kl} = \delta_{jk}\mathbf{e}_{il}.$$

从中可以读出结构常数为 $c_{ij,kl}^{mn} = \delta_{im}\delta_{jk}\delta_{ln}$. 注意我们需要两对指标才能标记这些常数.

以 $\{\mathbf{e}_{ij}\}_{i,j=1}^n$ 为基, 并以上面的结构常数为乘法表的代数 (这里不要求将 \mathbf{e}_{ij} 具体化为矩阵) 就称作全矩阵代数 (total matrix algebra). 用 \mathbb{F} 表示 \mathbb{R} 或者 \mathbb{C} . 那么 \mathbb{F} 上的全矩阵代数就记作 $\mathbb{F} \otimes \mathcal{M}_n$ 或者 $\mathcal{M}_n(\mathbb{F})$. 这是个结合代数, 并与实 (或者复) 矩阵代数同构, 不过其元素不必非得是 $n \times n$ 矩阵. 如果不强调矩阵的维数, 那么我们就简单写作 $\mathbb{F} \otimes \mathcal{M}$ 或者 $\mathcal{M}(\mathbb{F})$.

现在我们来构造这个代数的左理想. 取 \mathbf{e}_{pq} 并给它左乘 $\mathcal{M}_n(\mathbb{F})$ 中的任意元素 $\sum_{i,j=1}^n \alpha_{ij}\mathbf{e}_{ij}$. 这就给出

$$\left(\sum_{i,j=1}^n \alpha_{ij}\mathbf{e}_{ij} \right) \mathbf{e}_{pq} = \sum_{i,j=1}^n \alpha_{ij}\mathbf{e}_{ij}\mathbf{e}_{pq} = \sum_{i,j=1}^n \alpha_{ij}\delta_{jp}\mathbf{e}_{iq} = \sum_{i=1}^n \alpha_{ip}\mathbf{e}_{iq},$$

这对应了一个除了第 q 列, 其余列都是零的矩阵. 将满足这一条件的所有矩阵记作 \mathcal{L} . 现在用一个一般的矩阵 $\sum_{l,m=1}^n \beta_{lm}\mathbf{e}_{lm}$ 去乘 \mathcal{L} 中的元素, 我们就得到*

$$\begin{aligned} \left(\sum_{l,m=1}^n \beta_{lm}\mathbf{e}_{lm} \right) \left(\sum_{i=1}^n \gamma_i \mathbf{e}_{iq} \right) &= \sum_{i,l,m=1}^n \beta_{lm}\gamma_i \mathbf{e}_{lm}\mathbf{e}_{iq} = \sum_{i,l,m=1}^n \beta_{lm}\gamma_i \delta_{mi}\mathbf{e}_{lq} \\ &= \sum_{l,m=1}^n \beta_{lm}\gamma_m \mathbf{e}_{lq} = \sum_{l=1}^n \underbrace{\left(\sum_{m=1}^n \beta_{lm}\gamma_m \right)}_{:=\eta_l} \mathbf{e}_{lq} \\ &= \sum_{l=1}^n \eta_l \mathbf{e}_{lq}. \end{aligned}$$

* 在最后的结果中指标 p 不重要, 因为固定 q 以后不管如何变动 p 都给出同一个矩阵. 因此我们将 \mathcal{L} 中的矩阵写作 $\sum_{i=1}^n \gamma_i \mathbf{e}_{iq}$ 而非前面的 $\sum_{i=1}^n \gamma_i p \mathbf{e}_{iq}$.

这就表明 \mathcal{L} 是个左理想. 不仅如此, \mathcal{L} 的这一构造还推出它满足『定理 3.2.6』的条件 (b). 和前文构造类似, 现在我们给 \mathbf{e}_{pq} 右乘 $\mathcal{M}_n(\mathbb{F})$ 中任意元素, 这样就可以得到一个右理想, 它由那些除了第 p 行以外其余行均

为零的矩阵构成. 这个右理想就满足右极小理想版本下『定理 3.2.6』的条件 (b). 综上所述, 我们得到下述定理:

定理 3.3.1. 全矩阵代数 $\mathbb{R} \otimes \mathcal{M}$ 或 $\mathbb{C} \otimes \mathcal{M}$ 的极小左(右)理想由那些除了一列(行)以外其余列(行)都为零的矩阵构成.

现在给 \mathbf{e}_{pq} 分别左乘和右乘一个任意矩阵. 读者很容易就能证明, 这样做会恢复回整个全矩阵代数. 这就表明这个代数没有真理想. 后文的『例 3.3.3』则指出 $\mathcal{M}_n(\mathbb{F})$ 的中心就是 $\text{Span}\{\mathbf{1}_n\}$, 其中 $\mathbf{1}_n$ 是 $\mathcal{M}_n(\mathbb{F})$ 的单位元. 由此我们就得到下面的结论:

定理 3.3.2. 全矩阵代数 $\mathcal{M}_n(\mathbb{F})$ 为中心单代数.

例 3.3.3. 设 $\mathbf{a} = \sum_{i,j=1}^n \alpha_{ij} \mathbf{e}_{ij}$ 属于 $\mathbb{F} \otimes \mathcal{M}_n$ 的中心. 那么

$$\begin{aligned} \mathbf{a}\mathbf{e}_{kl} &= \sum_{i,j=1}^n \alpha_{ij} \mathbf{e}_{ij} \mathbf{e}_{kl} = \sum_{i,j=1}^n \alpha_{ij} \delta_{jk} \mathbf{e}_{il} = \sum_{i=1}^m \alpha_{ik} \mathbf{e}_{il}, \\ \mathbf{e}_{kl}\mathbf{a} &= \sum_{i,j=1}^n \mathbf{e}_{kl} \alpha_{ij} \mathbf{e}_{ij} = \sum_{i,j=1}^n \alpha_{ij} \delta_{il} \mathbf{e}_{kj} = \sum_{j=1}^n \alpha_{lj} \mathbf{e}_{kj}. \end{aligned}$$

由于这两个式子要相等, 我们就必须有

$$\sum_{i=1}^n (\alpha_{ik} \mathbf{e}_{il} - \alpha_{li} \mathbf{e}_{ki}) = 0.$$

在这个求和中令 $l = k$, 并利用 \mathbf{e}_{ij} 之间的线性无关性, 我们就看到 $i \neq k$ 时必须有 $\alpha_{ik} = 0$. 这就表明 \mathbf{a} 必须是个对角矩阵. 现在记 $\mathbf{a} = \sum_{k=1}^n \lambda_k \mathbf{e}_{kk}$, 并设 $\mathbf{b} = \sum_{i,j=1}^n \beta_{ij} \mathbf{e}_{ij}$ 是 $\mathbb{F} \otimes \mathcal{M}_n$ 中任意元素. 那么,

$$\begin{aligned} \mathbf{ab} &= \sum_{i,j,k=1}^n \lambda_k \beta_{ij} \mathbf{e}_{kk} \mathbf{e}_{ij} = \sum_{i,j,k=1}^n \lambda_k \beta_{ij} \delta_{ki} \mathbf{e}_{kj} = \sum_{i,j=1}^n \lambda_i \beta_{ij} \mathbf{e}_{ij}, \\ \mathbf{ba} &= \sum_{i,j,k=1}^n \lambda_k \beta_{ij} \mathbf{e}_{ij} \mathbf{e}_{kk} = \sum_{i,j,k=1}^n \lambda_k \beta_{ij} \delta_{jk} \mathbf{e}_{ik} = \sum_{i,j=1}^n \lambda_j \beta_{ij} \mathbf{e}_{ij}. \end{aligned}$$

由于这两个表达式必须相等, 并且 \mathbf{e}_{ij} 是线性无关的, 我们看到对所有 i, j 以及 β_{ij} 必须有 $\lambda_j \beta_{ij} = \lambda_i \beta_{ij}$. 使其成立的唯一办法就是让 $\lambda_i = \lambda_j$ 对所有 i, j 成立. 这就推出 $\mathbf{a} = \lambda \mathbf{1}_n$, 其中 $\mathbf{1}_n = \sum_{k=1}^n \mathbf{e}_{kk}$ 是 $\mathcal{M}_n(\mathbb{F})$ 的单位元. 因此, $\mathcal{M}_n(\mathbb{F})$ 就是个中心代数.

3.4. 代数的导子

『例 3.1.9』中最后两个例子有个在所有代数中都尤为重要的特性，那就是微分的乘法法则。

定义 3.4.1 (导子). 若矢量空间的自同态 $\mathbf{D} : \mathcal{A} \rightarrow \mathcal{A}$ 满足

$$\mathbf{D}(\mathbf{ab}) = [\mathbf{D}(\mathbf{a})]\mathbf{b} + \mathbf{a}[\mathbf{D}(\mathbf{b})]$$

这一额外性质，则称其为 \mathcal{A} 上的导子 (*derivation*)。

例 3.4.2. 设 $\mathcal{C}^r(a, b)$ 如『例 3.1.9』所述，并设 \mathbf{D} 就是常规的微分映射 $\mathbf{D} : f \mapsto f'$ ，其中 f' 是 f 的导函数。那么微分的乘法法则就表明 \mathbf{D} 是代数 $\mathcal{C}^r(a, b)$ 上的导子。

例 3.4.3. 考虑所有 $n \times n$ 矩阵在『(3.3) 式』所给乘法下所得代数。设 A 是个固定的矩阵，现在定义线性变换

$$\mathbf{D}_A = A \bullet B.$$

那么我们就有

$$\begin{aligned}\mathbf{D}_A(B \bullet C) &= A \bullet (B \bullet C) = A(B \bullet C) - (B \bullet C)A \\ &= A(BC - CB) - (BC - CB)A \\ &= ABC - ACB - BCA + CBA.\end{aligned}$$

另一方面，我们又有

$$\begin{aligned}(\mathbf{D}_A B) \bullet C + B \bullet (\mathbf{D}_A C) &= (A \bullet B) \bullet C + B \bullet (A \bullet C) \\ &= (AB - BA) \bullet C + B \bullet (AC - CA) \\ &= (AB - BA)C - C(AB - BA) \\ &\quad + B(AC - CA) - (AC - CA)B \\ &= ABC + CBA - BCA - ACB.\end{aligned}$$

二者相等，这就说明 \mathbf{D}_A 是 \mathcal{A} 上的导子。

定理 3.4.4. 设 $\{\mathbf{e}_i\}_{i=1}^N$ 是代数 \mathcal{A} 的一组基。那么矢量空间自同态 $\mathbf{D} : \mathcal{A} \rightarrow \mathcal{A}$ 是它上面一个导子的充要条件是

$$\mathbf{D}(\mathbf{e}_i \mathbf{e}_j) = \mathbf{D}(\mathbf{e}_i) \mathbf{e}_j + \mathbf{e}_i \mathbf{D}(\mathbf{e}_j) \quad \forall i, j = 1, 2, \dots, N.$$

证。其证明很简单，故留给读者用作习题。 \square

若 \mathcal{A} 带单位元 \mathbf{e} , 则有 $\mathbf{D}(\mathbf{e}) = \mathbf{0}$. 原因在于

$$\mathbf{D}(\mathbf{e}) = \mathbf{D}(\mathbf{e}\mathbf{e}) = \mathbf{D}(\mathbf{e})\mathbf{e} + \mathbf{e}\mathbf{D}(\mathbf{e}) = 2\mathbf{D}(\mathbf{e}).$$

这也表明 $\mathbf{e} \in \ker \mathbf{D}$. 容易证明 $\ker \mathbf{D}$ 是 \mathcal{A} 的子代数.¹⁸

命题 3.4.5. 所有导子都满足 *Leibniz* 公式 (*Leibniz formula*):

$$\mathbf{D}^n(\mathbf{ab}) = \sum_{k=0}^n \binom{n}{k} \mathbf{D}^k(\mathbf{A}) \mathbf{D}^{n-k}(\mathbf{b}).$$

证. 通过数学归纳法即可证明, 并且证明过程和『例 1.5.2』中证明二项式定理极为相似. 是故我们将具体细节留作习题. \square

代数 \mathcal{A} 上的导子本身是其底矢量空间的自同态, 从而所有导子就是 $\text{End}(\mathcal{A})$ 的子集. 若 \mathbf{D}_1 和 \mathbf{D}_2 均为导子, 那么很容易证明其线性组合 $\alpha_1 \mathbf{D}_1 + \alpha_2 \mathbf{D}_2$ 也是导子. 因此, 代数 \mathcal{A} 上所有导子的集合 $\mathcal{D}(\mathcal{A})$ 就构成一个矢量空间, 从而是 $\text{End}(\mathcal{A})$ 的子空间. 那么 $\mathcal{D}(\mathcal{A})$ 是否还进一步构成 $\text{End}(\mathcal{A})$ 的子代数呢? 换言之, $\mathbf{D}_1 \mathbf{D}_2$ 是否是导子? 现在让我们具体计算一下!

$$\begin{aligned} \mathbf{D}_1 \mathbf{D}_2(\mathbf{ab}) &= \mathbf{D}_1([\mathbf{D}_2(\mathbf{a})]\mathbf{b} + \mathbf{a}[\mathbf{D}_2(\mathbf{b})]) \\ &= [\mathbf{D}_1 \mathbf{D}_2(\mathbf{a})]\mathbf{b} + \mathbf{D}_2(\mathbf{a})\mathbf{D}_1(\mathbf{b}) + \mathbf{D}_1(\mathbf{a})\mathbf{D}_2(\mathbf{b}) + \mathbf{a}[\mathbf{D}_1 \mathbf{D}_2(\mathbf{b})]. \end{aligned}$$

这就表明两个导子的乘积不再是导子, 因为上式中间两项通常非零. 不过, 由于中间这两项对其下标是对称的, 我们就可以通过作差来消除它们, 也就是去考察 $\mathbf{D}_1 \mathbf{D}_2 - \mathbf{D}_2 \mathbf{D}_1$. 问题是这样做是否确实会给出一个导子. 现在交换上式的下标, 就可以得到

$$\mathbf{D}_2 \mathbf{D}_1(\mathbf{ab}) = [\mathbf{D}_2 \mathbf{D}_1(\mathbf{a})]\mathbf{b} + \mathbf{D}_1(\mathbf{a})\mathbf{D}_2(\mathbf{b}) + \mathbf{D}_2(\mathbf{a})\mathbf{D}_1(\mathbf{b}) + \mathbf{a}[\mathbf{D}_2 \mathbf{D}_1(\mathbf{b})].$$

将其与前面的结果相减, 就得到

$$\begin{aligned} (\mathbf{D}_1 \mathbf{D}_2 - \mathbf{D}_2 \mathbf{D}_1)(\mathbf{ab}) &= [\mathbf{D}_1 \mathbf{D}_2(\mathbf{a})]\mathbf{b} + \mathbf{a}[\mathbf{D}_1 \mathbf{D}_2(\mathbf{b})] - [\mathbf{D}_2 \mathbf{D}_1(\mathbf{a})]\mathbf{b} - \mathbf{a}[\mathbf{D}_2 \mathbf{D}_1(\mathbf{b})] \\ &= [(\mathbf{D}_1 \mathbf{D}_2 - \mathbf{D}_2 \mathbf{D}_1)(\mathbf{a})]\mathbf{b} + \mathbf{a}[(\mathbf{D}_1 \mathbf{D}_2 - \mathbf{D}_2 \mathbf{D}_1)(\mathbf{b})]. \end{aligned}$$

正因如此, 如果我们定义新的乘积为

$$\mathbf{D}_1 \bullet \mathbf{D}_2 := \mathbf{D}_1 \mathbf{D}_2 - \mathbf{D}_2 \mathbf{D}_1, \tag{3.11}$$

那么 $\mathcal{D}(\mathcal{A})$ 就成了一个代数.

¹⁸译者注: 取 $\mathbf{a}, \mathbf{b} \in \ker \mathbf{D}$, 则 $\mathbf{D}(\mathbf{ab}) = \mathbf{0}\mathbf{b} + \mathbf{a}\mathbf{0} = \mathbf{0}$, 是故 $\mathbf{ab} \in \ker \mathbf{D}$. 这就表明 $\ker \mathbf{D}$ 确实是子代数.

定理 3.4.6(导子代数). 代数 \mathcal{A} 上导子的集合 $\mathcal{D}(\mathcal{A})$ 在『(3.11)式』给出的乘法下构成一个代数, 称作 \mathcal{A} 的导子代数 (*derivation algebra*).

定义 3.4.7 (ϕ 导子). 设 \mathcal{A} 和 \mathcal{B} 均为代数, $\phi : \mathcal{A} \rightarrow \mathcal{B}$ 为同态. 若矢量空间自同态 $\mathbf{D} : \mathcal{A} \rightarrow \mathcal{B}$ 满足

$$\mathbf{D}(\mathbf{a}_1\mathbf{a}_2) = \mathbf{D}(\mathbf{a}_1)\phi(\mathbf{a}_2) + \phi(\mathbf{a}_1)\mathbf{D}(\mathbf{a}_2), \quad \mathbf{a}_1, \mathbf{a}_2 \in \mathcal{A},$$

则称 \mathbf{D} 是个 ϕ 导子 (ϕ -derivation).

例 3.4.8. 设 \mathbf{D}_A 是 \mathcal{A} 中一个导子. 那么 $\mathbf{D} := \phi \circ \mathbf{D}_A$ 就是个 ϕ 导子, 原因在于

$$\begin{aligned}\phi \circ \mathbf{D}_A(\mathbf{a}_1\mathbf{a}_2) &= \phi[\mathbf{D}_A(\mathbf{a}_1)\mathbf{a}_2 + \mathbf{a}_1\mathbf{D}_A(\mathbf{a}_2)] \\ &= \phi[\mathbf{D}_A(\mathbf{a}_1)]\phi(\mathbf{a}_2) + \phi(\mathbf{a}_1)\phi[\mathbf{D}_A(\mathbf{a}_2)] \\ &= \phi \circ \mathbf{D}_A(\mathbf{a}_1)\phi(\mathbf{a}_2) + \phi(\mathbf{a}_1)\phi \circ \mathbf{D}_A(\mathbf{a}_2).\end{aligned}$$

类似地, 若 \mathbf{D}_B 是 \mathcal{B} 中的导子, 则 $\mathbf{D}_B \circ \phi$ 也是个 ϕ 导子.

更具体一点, 我们将 \mathcal{A} 取作 r 次可微函数的代数 $\mathcal{C}^r(a, b)$, 将 \mathcal{B} 取作实数代数 \mathbb{R} . 定义 $\phi_c : \mathcal{C}^r(a, b) \rightarrow \mathbb{R}$ 为函数在定点 $c \in (a, b)$ 处的函数值, 即 $\phi_c(f) = f(c)$. 若将 $\mathbf{D}_c : \mathcal{C}^r(a, b) \rightarrow \mathbb{R}$ 定义为 $\mathbf{D}_c(f) = f'(c)$, 那么读者很快就能验证 \mathbf{D}_c 是个 ϕ_c 导子.

定义 3.4.9 (反导子). 设 \mathcal{A} 为含幺代数, ω 是 \mathcal{A} 的对合. 若线性变换 $\mathbf{\Omega} \in \mathcal{L}(\mathcal{A})$ 满足

$$\mathbf{\Omega}(\mathbf{a}_1\mathbf{a}_2) = \mathbf{\Omega}(\mathbf{a}_1)\mathbf{a}_2 + \omega(\mathbf{a}_1)\mathbf{\Omega}(\mathbf{a}_2),$$

则称 $\mathbf{\Omega}$ 是关于 ω 的反导子 (*antiderivation*). 特别地, 任意导子都是关于恒等算符的反导子.

和导子的情况一样, 我们可以证明 (1) $\ker \mathbf{\Omega}$ 是 \mathcal{A} 的子代数; (2) 若 \mathbf{e} 是 \mathcal{A} 的单位元, 则 $\mathbf{\Omega}(\mathbf{e}) = \mathbf{0}$; (3) $\mathbf{\Omega}$ 完全由它在 \mathcal{A} 生成元上的作用效果决定.

定理 3.4.10. 设 $\mathbf{\Omega}_1$ 和 $\mathbf{\Omega}_2$ 分别是关于对合 ω_1 和 ω_2 的反导子. 假设 $\omega_1 \circ \omega_2 = \omega_2 \circ \omega_1$, 并进一步假定

$$\omega_1\mathbf{\Omega}_2 = \pm \mathbf{\Omega}_2\omega_1 \quad \text{且} \quad \omega_2\mathbf{\Omega}_1 = \pm \mathbf{\Omega}_1\omega_2.$$

那么 $\mathbf{\Omega}_1\mathbf{\Omega}_2 \mp \mathbf{\Omega}_2\mathbf{\Omega}_1$ 就是关于对合 $\omega_1 \circ \omega_2$ 的反导子.

证. 借助『**定义 3.4.9**』展开 $\mathbf{\Omega}_1$ 和 $\mathbf{\Omega}_2$ 对 $\mathbf{a}_1\mathbf{a}_2$ 的作用, 然后将其代入 $\mathbf{\Omega}_1\mathbf{\Omega}_2 \mp \mathbf{\Omega}_2\mathbf{\Omega}_1$ 直接计算即可. 我们将具体的计算过程留给读者. \square

我们对上述定理的下述特殊情形感兴趣:

- 设 Ω 是关于 ω 的反导子, D 是使得 $\omega D = D\omega$ 的导子. 那么 $D\Omega - \Omega D$ 就是关于 ω 的反导子.
- 若 Ω_1 和 Ω_2 是关于同一个对合 ω 的反导子, 并满足 $\omega\Omega_i = -\Omega_i\omega$, $i = 1, 2$. 那么, $\Omega_1\Omega_2 + \Omega_2\Omega_1$ 就是一个导子.
- 上述情形还有一个特殊例子: 若 Ω 是关于对合 ω 的反导子, 并且 $\omega\Omega = -\Omega\omega$, 那么 Ω^2 就是个导子.

3.5. 代数的分解

在『2.1.3节』中, 我们将一个矢量空间分解为一些更小的矢量空间. 与之类似, 将代数分解为“更小”的代数也很有用. 在这一节, 我们来探索一下在何种条件下存在这样的分解, 并探究该分解的性质. 本节出现的均为结合代数.

定义 3.5.1 (幂零元、幂等元). 若存在正整数 k 使得非零元 $\mathbf{a} \in A$ 满足 $\mathbf{a}^k = \mathbf{0}$, 则称 \mathbf{a} 为幂零元 (*nilpotent*). 使得这个式子成立的最小正整数就称作 \mathbf{a} 的幂零指数 (*index*), 或简称为指数. 若 A 的子代数 B 内所有元素都是幂零元, 则称这个子代数是零化的 (*nil*).¹⁹ 如若 $B^\nu = \{\mathbf{0}\}$ 且 $B^{\nu-1} \neq \{\mathbf{0}\}$, 则称非零子代数 B 幂零, 且其幂零指数为 ν . * 若非零元素 $\mathbf{P} \in A$ 满足 $\mathbf{P}^2 = \mathbf{P}$, 则称其为幂等元 (*idempotent*).

命题 3.5.2. 单位元是可除代数中仅有的幂等元.

证. 设 $\mathbf{a}^2 = \mathbf{a}$, 因为是可除代数, 我们可以在两边同时乘以 \mathbf{a} 的逆元 \mathbf{a}^{-1} , 由此即得 $\mathbf{a} = \mathbf{aa}^{-1} = \mathbf{1}$. \square

如果 \mathbf{P} 是幂等的, 那么对任意正整数 k 均有 $\mathbf{P}^k = \mathbf{P}$. 因此, 幂零子代数里面不可能包含幂等元.

接下来这个定理很有用, 但是证明却是技术性的, 因此略去其证明. (证明可见 [Blyth, 1990], pp. 191.)

定理 3.5.3. 零化理想是幂零的.

例 3.5.4. 所有 $n \times n$ 上三角矩阵的集合是 $n \times n$ 矩阵代数的子代数, 因为两个上三角矩阵的乘积仍旧是上三角矩阵 (这很容易就能验证).

严格上三角矩阵都是幂零的. 我们以 4×4 矩阵为例. 设

$$\mathbf{A} = \begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ 0 & 0 & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

¹⁹译者注: 习惯上应该将 nil 翻译成零的, 但后面我们会遇到 nil ideal, 翻译成零理想会和 $\{\mathbf{0}\}$ 这个零理想 (zero ideal 或作 null ideal) 混淆, 于是这里改称其为零化的.

那么很容易就能算出

$$A^2 = \begin{bmatrix} 0 & 0 & a_{12}a_{23} & a_{12}a_{24} + a_{13}a_{34} \\ 0 & 0 & 0 & a_{23}a_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A^3 = \begin{bmatrix} 0 & 0 & 0 & a_{12}a_{23}a_{34} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

这就说明 4×4 的严格上三角矩阵是幂零指数为 4 的幂零元.²⁰ 事实上, 我们可以证明 4×4 严格上三角矩阵构成的子代数是幂零的, 且幂零指数为 4.

读者可以确信 $n \times n$ 严格上三角矩阵是幂零指数为 n 的幂零元, 由 $n \times n$ 严格上三角矩阵构成的子代数同样是幂零的, 且幂零指数为 n .

3.5.1. 代数的根

幂零子代数在代数分类问题中起基础性作用. 值得注意的是, 一个代数中所有的幂零左理想、幂零右理想、幂零双边理想都包含于某个特定的幂零理想, 我们现在就来探究这个特定理想.

引理 3.5.5. 设 \mathcal{L} 以及 \mathcal{M} 是代数 A 的两个幂零左(右)理想, 它们的幂零指数分别是 λ 和 μ . 那么, $\mathcal{L} + \mathcal{M}$ 也是幂零的, 并且幂零指数最多为 $\lambda + \mu - 1$.

证. 我们对左理想情形证明该引理. 显然, $\mathcal{L} + \mathcal{M}$ 也是左理想, 它里面任意元素的 k 次方都可以写成形式为 $\mathbf{a}_1\mathbf{a}_2 \dots \mathbf{a}_k$ 的元素之线性组合, 其中 \mathbf{a}_i 要么属于 \mathcal{L} , 要么属于 \mathcal{M} . 现在假设在这个乘积中有 l 项属于 \mathcal{L} , 有 m 项属于 \mathcal{M} . 并设 j 是使得 $\mathbf{a}_j \in \mathcal{L}$ 的最大整数. 现在从 \mathbf{a}_j 开始向左移动, 直到我们遇到另一个属于 \mathcal{L} 的元素(记作 \mathbf{a}_r). 那么, 从 \mathbf{a}_{r+1} 到 \mathbf{a}_{j-1} 的所有项都属于 \mathcal{M} . 考虑到 \mathcal{L} 是左理想, 于是

$$\underbrace{\mathbf{a}_{r+1} \dots \mathbf{a}_{j-1} \mathbf{a}_j}_{\in \mathcal{A}} =: \mathbf{a}'_j \in \mathcal{L}.$$

²⁰译者注: 严格来说, 是指数最多为 4. 因为如果矩阵元够特殊, 它会提前变成零矩阵. 比如说 $a_{23} = 0$, 那么此时 $A^3 = 0$.

这就将乘积 $\mathbf{a}_r \mathbf{a}_{r+1} \dots \mathbf{a}_{j-1} \mathbf{a}_j$ 缩减为 $\mathbf{a}_r \mathbf{a}'_j$, 而且这两个因子都属于 \mathcal{L} . 重复这个过程, 我们就得到

$$\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_k = \mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_l \mathbf{c}, \quad \mathbf{b}_i \in \mathcal{L}, \mathbf{c} \in \mathcal{M}.$$

类似地, 我们从属于 \mathcal{M} 的元素中下标最大的那个开始重复上述操作, 就得到

$$\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_k = \mathbf{c}_1 \mathbf{c}_2 \dots \mathbf{c}_m \mathbf{b}, \quad \mathbf{b} \in \mathcal{L}, \mathbf{c}_i \in \mathcal{M}.$$

由于 $k = l + m$, 若 $k = \mu + \lambda - 1$, 则 $(\mu - m) + (\lambda - l) = 1$. 这就表明若 $m < \mu$, 则 $l \geq \lambda$; 若 $l < \lambda$ 则 $m \geq \mu$. 在这两种情况下, 由上面的两个式子就能看出 $\mathbf{a}_1 \dots \mathbf{a}_k = \mathbf{0}$.²¹ 这就说明, $\mathcal{L} + \mathcal{M}$ 是幂零指数最多为 $\mu + \lambda - 1$ 的左理想. 右理想情形的证明与此雷同. \square

引理 3.5.6. 设 \mathcal{L} 是代数 \mathcal{A} 的幂零左理想. 那么 $\mathcal{J} = \mathcal{L} + \mathcal{LA}$ 就是一个幂零双边理想.

证. 因为 \mathcal{L} 是左理想, 于是 $\mathcal{AL} \subseteq \mathcal{L}$. 进而

$$\mathcal{AJ} = \mathcal{AL} + \mathcal{AL}\mathcal{A} \subseteq \mathcal{L} + \mathcal{LA} = \mathcal{J},$$

这就表明 \mathcal{J} 是左理想. 另一方面,

$$\mathcal{JA} = \mathcal{LA} + \mathcal{LA}\mathcal{A} \subseteq \mathcal{LA} + \mathcal{LA} = \mathcal{LA} \subseteq \mathcal{J},$$

这就表明 \mathcal{J} 还是右理想.

现在考察 \mathcal{LA} 中 k 个元素的乘积:

$$\mathbf{l}_1 \mathbf{a}_1 \mathbf{l}_2 \mathbf{a}_2 \dots \mathbf{l}_k \mathbf{a}_k = \mathbf{l}_1 \mathbf{l}'_2 \mathbf{l}'_3 \dots \mathbf{l}'_k \mathbf{a}_k, \quad \mathbf{l}_j \in \mathcal{L}, \mathbf{a}_j \in \mathcal{A},$$

其中 $\mathbf{l}'_i := \mathbf{a}_{i-1} \mathbf{l}_i \in \mathcal{L}$, $i > 1$. 这就表明, 若 k 等于 \mathcal{L} 的幂零指数, 则这个乘积就等于零, 进而 \mathcal{LA} 是幂零的. 另外注意, 由于有些 \mathbf{a} 可能属于 \mathcal{L} , 所以 \mathcal{LA} 和 \mathcal{A} 的幂零指数最多相等. 最后利用『引理 3.5.5』就证明了这个引理. \square

前述两条引理旨在证明下面的结果:

定理 3.5.7. 设 \mathcal{A} 为代数, 则存在唯一的幂零理想使得 \mathcal{A} 所有的幂零左、右、双边理想都包含于它.

²¹译者注: 因为此时有 l 个 \mathcal{L} 中元素或者 m 个 \mathcal{M} 中元素彼此相邻并相乘, 而 \mathcal{L} 和 \mathcal{M} 的幂零指数分别为 $\lambda \leq l$ 和 $\mu \leq m$, 由定义可知此时这个乘积必须为零.

证. 设 \mathcal{N} 是维数最大的幂零理想. 现取任意幂零理想 \mathcal{M} . 根据『引理 3.5.5』, $\mathcal{N} + \mathcal{M}$ 就同时是幂零的左理想和右理想, 从而是幂零理想. 根据假设, 我们有²² $\mathcal{N} + \mathcal{M} \subseteq \mathcal{N}$, 这就表明 $\mathcal{M} \subseteq \mathcal{N}$, 从而 \mathcal{N} 包含了所有的幂零理想. 现在假设有另一个维数最大的理想 \mathcal{N}' , 则有 $\mathcal{N}' \subseteq \mathcal{N}$ 以及 $\mathcal{N} \subseteq \mathcal{N}'$, 从而 $\mathcal{N}' = \mathcal{N}$, 即 \mathcal{N} 唯一.

若 \mathcal{L} 是幂零的左理想, 那么根据『引理 3.5.6』可知 $\mathcal{J} = \mathcal{L} + \mathcal{LA}$ 是幂零理想, 进而根据上面已有的结果可知 $\mathcal{L} \subseteq \mathcal{J} \subseteq \mathcal{N}$. 这就表明 \mathcal{N} 同时包含了所有的幂零左理想. 类似地, 可以证明 \mathcal{N} 包含所有的幂零右理想. \square

根据上面这个定理, 我们引入如下说法:

定义 3.5.8 (代数的根). 对于代数 \mathcal{A} 而言, 由『定理 3.5.7』保证的那个唯一的最大理想就称作 \mathcal{A} 的根 (*radical*, 亦译作根基), 记作 $\text{Rad}(\mathcal{A})$.

我们已经看到, 幂零代数中不可能包含幂等元. 事实上, 这个推理反过来同样成立. 欲证其成立, 我们需要下述引理帮助:

引理 3.5.9. 若存在 $\mathbf{a} \in \mathcal{A}$ 使得 $\mathcal{A}\mathbf{a}^k = \mathcal{A}\mathbf{a}^{k-1}$ 对某个正整数 k 成立, 则 \mathcal{A} 中存在幂等元.

证. 记 $\mathcal{B} := \mathcal{A}\mathbf{a}^{k-1}$, 那么 \mathcal{B} 就是 \mathcal{A} 的左理想, 并且满足 $\mathcal{B}\mathbf{a} = \mathcal{B}$. 两边同时右乘 \mathbf{a} , 就得到

$$\mathcal{B}\mathbf{a}^2 = \mathcal{B}\mathbf{a} = \mathcal{B}, \quad \mathcal{B}\mathbf{a}^3 = \mathcal{B}\mathbf{a} = \mathcal{B},$$

以此类推可得对任意正整数 k 均有 $\mathcal{B}\mathbf{a}^k = \mathcal{B}$. 考虑到 $\mathcal{B} := \mathcal{A}\mathbf{a}^{k-1}$, 是故 $\mathbf{a}^k \in \mathcal{B}$. 现在令 $\mathbf{b} = \mathbf{a}^k$, 我们就有 $\mathcal{B}\mathbf{b} = \mathcal{B}$. 而这就意味着必然存在 $\mathbf{P} \in \mathcal{B}$ 使得 $\mathbf{P}\mathbf{b} = \mathbf{b}$, 亦或者说 $(\mathbf{P}^2 - \mathbf{P})\mathbf{b} = \mathbf{0}$. 根据『习题 3.32』就有 $\mathbf{P}^2 = \mathbf{P}$. 这就说明 \mathcal{B} 中存在幂等元, 而 $\mathcal{B} \subseteq \mathcal{A}$, 这就证明了这个引理. \square

命题 3.5.10. 一个代数是幂零代数的充要条件是其不含幂等元.

证. 必要性已在『定义 3.5.1』下予以说明, 是故只需证充分性. 也就是说, 我们要证明若 \mathcal{A} 不含幂等元, 则它必然幂零. 注意到一般情况下总有 $\mathcal{A}\mathbf{a} \subseteq \mathcal{A}$, 递推则得对所有 k 有 $\mathcal{A}\mathbf{a}^k \subseteq \mathcal{A}\mathbf{a}^{k-1}$. 这个包含关系必然无法取等, 因为假设可以取等, 则『引理 3.5.9』就断定 \mathcal{A} 内有幂等元, 与假设矛盾. 从而有 $\mathcal{A}\mathbf{a}^k \subsetneq \mathcal{A}\mathbf{a}^{k-1}$. 这对所有 k 都成立, 我们就得到如下递降链:

$$\mathcal{A} \supsetneq \mathcal{A}\mathbf{a} \supsetneq \mathcal{A}\mathbf{a}^2 \supsetneq \cdots \supsetneq \mathcal{A}\mathbf{a}^k \supsetneq \cdots.$$

²²译者注: 假设 $\mathcal{N} + \mathcal{M}$ 比 \mathcal{N} 大, 那么 $\mathcal{N} + \mathcal{M}$ 的维数一定大于 \mathcal{N} 的维数, 这与假设矛盾!

由于每个 $\mathcal{A}\mathbf{a}^k$ 都是 \mathcal{A} 的子空间, 而 \mathcal{A} 是有限维的, 因此必然存在某个正整数 r 使得 $\mathcal{A}\mathbf{a}^r = \{\mathbf{0}\}$ 对所有 $\mathbf{a} \in \mathcal{A}$ 成立. 特别地, 我们就发现对所有 $\mathbf{a} \in \mathcal{A}$ 均有某个正整数 r 使得 $\mathbf{a}^{r+1} = \mathbf{0}$. 这就表明 \mathcal{A} 是零化的, 进而根据『定理 3.5.3』可知其幂零. \square

设 \mathbf{P} 是 \mathcal{A} 的幂等元. 我们考察其左零化子 $\mathcal{L}(\mathbf{P})$ (见『例 3.2.2』), 那么对任意 $\mathbf{a} \in \mathcal{A}$ 均有 $(\mathbf{a} - \mathbf{a}\mathbf{P}) \in \mathcal{L}(\mathbf{P})$.²³ 不仅如此, 若 $\mathbf{a} \in \mathbf{P}\mathcal{L}(\mathbf{P})$, 则存在某个 $\mathbf{x} \in \mathcal{L}(\mathbf{P})$ 使得 $\mathbf{a} = \mathbf{Px}$. 进而这样的 \mathbf{a} 就满足 $\mathbf{Pa} = \mathbf{a}$ 以及 $\mathbf{a}\mathbf{P} = \mathbf{0}$.

²⁴

类似地, 考察 \mathbf{P} 的右零化子 $\mathcal{R}(\mathbf{P})$, 则有 $(\mathbf{a} - \mathbf{Pa}) \in \mathcal{R}(\mathbf{P})$ 对任意 $\mathbf{a} \in \mathcal{A}$ 成立. 并且若 $\mathbf{a} \in \mathcal{R}(\mathbf{P})\mathbf{P}$, 则存在 $\mathbf{x} \in \mathcal{R}(\mathbf{P})$ 使得 $\mathbf{a} = \mathbf{xP}$. 这样的 \mathbf{a} 就满足 $\mathbf{a}\mathbf{P} = \mathbf{a}$ 以及 $\mathbf{Pa} = \mathbf{0}$.

现在令 $\mathcal{J}(\mathbf{P}) = \mathcal{L}(\mathbf{P}) \cap \mathcal{R}(\mathbf{P})$. 那么很明显 $\mathcal{J}(\mathbf{P})$ 是一个双边理想, 它里面的元素 \mathbf{a} 均满足 $\mathbf{a}\mathbf{P} = \mathbf{Pa} = \mathbf{0}$. 除此以外, 我们还需要考虑子代数 $\mathbf{P}\mathcal{A}\mathbf{P}$, 它里面的元素 \mathbf{a} 都满足 $\mathbf{Pa} = \mathbf{a}\mathbf{P} = \mathbf{a}$. 综上所述, 我们就有

$$\begin{aligned} \mathbf{P}\mathcal{A}\mathbf{P} &= \{\mathbf{a} \in \mathcal{A} \mid \mathbf{Pa} = \mathbf{a}\mathbf{P} = \mathbf{a}\}, \\ \mathbf{P}\mathcal{L}(\mathbf{P}) &= \{\mathbf{a} \in \mathcal{A} \mid \mathbf{Pa} = \mathbf{a}, \mathbf{a}\mathbf{P} = \mathbf{0}\}, \\ \mathcal{R}(\mathbf{P})\mathbf{P} &= \{\mathbf{a}\mathbf{P} = \mathbf{a}, \mathbf{Pa} = \mathbf{0}\}, \\ \mathcal{J}(\mathbf{P}) &= \{\mathbf{a} \in \mathcal{A} \mid \mathbf{a}\mathbf{P} = \mathbf{Pa} = \mathbf{0}\}. \end{aligned} \tag{3.12}$$

除此以外, 我们还有如下结论:

定理 3.5.11 (Peirce 分解). 设 \mathcal{A} 是带有幂等元 \mathbf{P} 的代数. 则有 \mathcal{A} 的 Peirce 分解 (Peirce decomposition):

$$\mathcal{A} = \mathbf{P}\mathcal{A}\mathbf{P} \oplus_V \mathbf{P}\mathcal{L}(\mathbf{P}) \oplus_V \mathcal{R}(\mathbf{P})\mathbf{P} \oplus_V \mathcal{J}(\mathbf{P}),$$

其中 \oplus_V 表示这是矢量空间直和, 此外出现在直和中的每个因子都是子代数.

证. 根据『(3.12) 式』, 这里的每个求和项都是代数. 另外不难看到, 任意两个求和项中仅有的共同矢量就是零矢量. 因此, 这里的每个求和都是子空间的直和. 最后只需注意到对任意 $\mathbf{a} \in \mathcal{A}$ 有下述恒等式:

$$\mathbf{a} \equiv \mathbf{PaP} + \underbrace{\mathbf{P}(\mathbf{a} - \mathbf{a}\mathbf{P})}_{\in \mathcal{L}(\mathbf{P})} + \underbrace{(\mathbf{a} - \mathbf{Pa})\mathbf{P}}_{\in \mathcal{R}(\mathbf{P})} + \underbrace{(\mathbf{a} - \mathbf{Pa} - \mathbf{a}\mathbf{P} + \mathbf{PaP})}_{\in \mathcal{J}(\mathbf{P})}.$$

²³译者注: 直接验证就有 $(\mathbf{a} - \mathbf{a}\mathbf{P})\mathbf{P} = \mathbf{a}\mathbf{P} - \mathbf{a}\mathbf{P}^2 = \mathbf{a}\mathbf{P} - \mathbf{a}\mathbf{P} = \mathbf{0}$, 从而 $(\mathbf{a} - \mathbf{a}\mathbf{P}) \in \mathcal{L}(\mathbf{P})$.

²⁴译者注: 因为 $\mathbf{Pa} = \mathbf{P}(\mathbf{Px}) = \mathbf{P}^2\mathbf{x} = \mathbf{Px} = \mathbf{a}$. 另外由于 $\mathbf{x} \in \mathcal{L}(\mathbf{P})$, 故 $\mathbf{xP} = \mathbf{0}$, 进而 $\mathbf{a}\mathbf{P} = \mathbf{PxP} = \mathbf{P}\mathbf{0} = \mathbf{0}$. 需要强调的是, 这两个性质也完全刻画了 $\mathbf{P}\mathcal{L}(\mathbf{P})$, 也就是说若 \mathbf{a} 满足这两个性质则必有 $\mathbf{a} \in \mathbf{P}\mathcal{L}(\mathbf{P})$. 换言之, 存在 $\mathbf{x} \in \mathcal{L}(\mathbf{P})$ 使得 $\mathbf{a} = \mathbf{Px}$. 事实上, 根据开始的论证, 我们就看到 $\mathbf{a} - \mathbf{a}\mathbf{P} = \mathbf{a} \in \mathcal{L}(\mathbf{P})$, 进而 $\mathbf{Pa} = \mathbf{a}$ 这个条件保证了它确实属于 $\mathbf{P}\mathcal{L}(\mathbf{P})$.

我们将具体的证明细节留作习题(见『习题 3.33』). \square

定义 3.5.12 (与幂等元正交、主幂等元). 设 $\mathbf{P} \in \mathcal{A}$ 是幂等元, 若 $\mathbf{a} \in \mathcal{A}$ 满足 $\mathbf{a}\mathbf{P} = \mathbf{P}\mathbf{a} = \mathbf{0}$, 则称 \mathbf{a} 与 \mathcal{P} 正交. 从而 $\mathcal{J}(\mathbf{P})$ 就是与 \mathbf{P} 正交的所有元素. 若 $\mathcal{J}(\mathbf{P})$ 中不含幂等元, 则称幂等元 \mathbf{P} 为主幂等元 (*principle idempotent*).

设 \mathbf{P}_0 为幂等元. 如果它不是主幂等元, 则 $\mathcal{J}(\mathbf{P}_0)$ 中就存在与之正交的幂等元 \mathbf{q} . 现在令 $\mathbf{P}_1 = \mathbf{P}_0 + \mathbf{q}$. 由于 $\mathbf{P}_0\mathbf{q} = \mathbf{q}\mathbf{P}_0\mathbf{0}$, 我们就看到 \mathbf{P}_1 是幂等的, 并且

$$\mathbf{P}_1\mathbf{P}_0 = \mathbf{P}_0\mathbf{P}_1 = \mathbf{P}_0, \quad \mathbf{P}_1\mathbf{q} = \mathbf{q}\mathbf{P}_1 = \mathbf{q}. \quad (3.13)$$

若 $\mathbf{x} \in \mathcal{J}(\mathbf{P}_1)$, 则 $\mathbf{x}\mathbf{P}_1 = \mathbf{P}_1\mathbf{x} = \mathbf{0}$, 接下来『(3.13) 式』的第一个式子中就给出 $\mathbf{x}\mathbf{P}_0 = \mathbf{P}_0\mathbf{x} = \mathbf{0}$, 也就是说 $\mathbf{x} \in \mathcal{J}(\mathbf{P}_0)$. 由此就得到 $\mathcal{J}(\mathbf{P}_1) \subseteq \mathcal{J}(\mathbf{P}_0)$. 由于 $\mathbf{q} \in \mathcal{J}(\mathbf{P}_0)$ 但是 $\mathbf{q} \notin \mathcal{J}(\mathbf{P}_1)$, 所以 $\mathcal{J}(\mathbf{P}_1)$ 是 $\mathcal{J}(\mathbf{P}_0)$ 的真子集. 若 \mathbf{P}_1 不是主幂等元, 那么 $\mathcal{J}(\mathbf{P}_1)$ 就包含某个幂等元 \mathbf{r} . 然后让 $\mathbf{P}_2 = \mathbf{P}_1 + \mathbf{r}$. 同样的过程就表明 \mathbf{P}_2 是幂等的, 并且 $\mathcal{J}(\mathbf{P}_2)$ 是 $\mathcal{J}(\mathbf{P}_1)$ 的真子集. 重复这一过程, 我们就得到

$$\mathcal{J}(\mathbf{P}_0) \supsetneq \mathcal{J}(\mathbf{P}_1) \supsetneq \mathcal{J}(\mathbf{P}_2) \supsetneq \cdots \supsetneq \mathcal{J}(\mathbf{P}_k) \supsetneq \cdots.$$

由于 $\mathcal{J}(\mathbf{P}_0)$ 维数有限, 上述递降链必然有尽头. 这就说明存在某个正整数 n 使得 $\mathcal{J}(\mathbf{P}_n)$ 不含幂等元, 即 \mathbf{P}_n 是主幂等元. 至此, 我们就证明了下面的结果:

命题 3.5.13. 每个非幂零代数都存在主幂等元.

定义 3.5.14 (本原幂等元). 不能写成两个彼此正交的幂等元之和的幂等元称作本原幂等元 (*primitive idempotent*).

命题 3.5.15. \mathbf{P} 是本原幂等元的充要条件是其为 \mathbf{PAP} 中仅有的幂等元.

证. 假设 \mathbf{P} 不是本原的. 那么就存在彼此正交的幂等元 \mathbf{P}_1 和 \mathbf{P}_2 使得 $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2$. 容易证明此时有 $\mathbf{PP}_i = \mathbf{P}_i\mathbf{P} = \mathbf{P}_i$, $i = 1, 2$. 因此, 根据『(3.12) 式』的第一个式子可知 $\mathbf{P}_i \in \mathbf{PAP}$, 从而 \mathbf{P} 不是 \mathbf{PAP} 中仅有的幂等元.

反过来, 假设 \mathbf{P} 不是 \mathbf{PAP} 中仅有的幂等元, 那么 \mathbf{PAP} 中就包含另一个幂等元 \mathbf{P}' . 利用『(3.12) 式』的第一个式子就有 $\mathbf{PP}' = \mathbf{P}'\mathbf{P} = \mathbf{P}'$. 这就表明

$$(\mathbf{P} - \mathbf{P}')\mathbf{P}' = \mathbf{P}'(\mathbf{P} - \mathbf{P}') = \mathbf{0},$$

$$(\mathbf{P} - \mathbf{P}')\mathbf{P} = \mathbf{P}(\mathbf{P} - \mathbf{P}') = \mathbf{P} - \mathbf{P}'.$$

而这就说明 $(\mathbf{P} - \mathbf{P}') \in \mathbf{PAP}$ 与 \mathbf{P}' 正交. 而

$$(\mathbf{P} - \mathbf{P}')^2 = \mathbf{P}^2 - \mathbf{PP}' - \mathbf{P}'\mathbf{P} + \mathbf{P}'^2 = \mathbf{P} - \mathbf{P}' - \mathbf{P}' + \mathbf{P}' = \mathbf{P} - \mathbf{P}',$$

于是 $\mathbf{P} - \mathbf{P}'$ 也是幂等元. 进而, $\mathbf{P} = (\mathbf{P} - \mathbf{P}') + \mathbf{P}'$ 说明 \mathbf{P} 可以写成两个相互正交的幂等元之和, 这说明它不是本原的. \square

现在设 \mathbf{P} 幂等且非本原. 我们记 $\mathbf{P} = \mathbf{P}_1 + \mathbf{Q}$, 其中 \mathbf{P}_1, \mathbf{Q} 是正交的幂等元. 如果 \mathbf{P}_1, \mathbf{Q} 中有一个不是本原的 (比如说 \mathbf{Q}), 那就令 $\mathbf{Q} = \mathbf{P}_2 + \mathbf{P}_3$, 其中 $\mathbf{P}_2, \mathbf{P}_3$ 是正交的幂等元. 根据『习题 3.34』, 集合 $\{\mathbf{P}_i\}_{i=1}^3$ 是彼此正交的幂等元, 且 $\mathbf{P} = \mathbf{P}_1 + \mathbf{P}_2 + \mathbf{P}_3$. 我们可以不断进行这一操作, 直到所有的 \mathbf{P}_i 都是本原的. 职是之故, 我们得到下述定理:

定理 3.5.16. 代数 A 的每个幂等元都可以表示成有限个互相正交的本原幂等元之和.

3.5.2. 半单代数

在代数分类问题中, 没有幂零理想的代数有着重要作用.

定义 3.5.17 (半单代数). 如果一个代数的根为零, 那么就称其半单 (*semi-simple*).

由于 $\text{Rad}(\mathcal{A})$ 包含了 \mathcal{A} 的所有幂零左、右、双边理想, 是故若 \mathcal{A} 半单, 则它自然就没有幂零的左、右、双边理想.

命题 3.5.18. 单代数是半单的.

证. 假设单代数 \mathcal{A} 不是半单的, 那么它就存在幂零理想. 我们现在来证明这种情况不可能发生. 由于 \mathcal{A} 是单代数, 它的非零理想仅有可能是 \mathcal{A} 自己. 因此我们仅需证明 \mathcal{A} 不可能幂零. 注意到 \mathcal{A}^2 是 \mathcal{A} 的理想, 于是 $\mathcal{A}^2 = \mathcal{A}$ 或者 $\mathcal{A}^2 = \{\mathbf{0}\}$. 假设 \mathcal{A} 幂零, 则必然有 $\mathcal{A}^2 = \{\mathbf{0}\}$. 因为若 $\mathcal{A}^2 = \mathcal{A}$, 则对任意 k 都有 $\mathcal{A}^k = \mathcal{A}$, 这就与 \mathcal{A} 幂零的假设矛盾了. 既然 $\mathcal{A}^2 = \{\mathbf{0}\}$, 那么 \mathcal{A} 的任意真子空间都是 \mathcal{A} 的非零真理想²⁵. 但这不可能, 因为 \mathcal{A} 是单代数, 它没有非平凡的理想. \square

引理 3.5.19. 设 \mathcal{A} 是半单代数, \mathbf{P} 是 \mathcal{A} 的任意主幂等元, 则 $\mathcal{A} = \mathbf{P}\mathcal{A}\mathbf{P}$.

证. 因为 \mathcal{A} 不是幂零的, 根据『[命题 3.5.13](#)』就知道它存在主幂等元 \mathbf{P} . 由于 \mathbf{P} 是主幂等元, 所以『[定理 3.5.11](#)』中的 $\mathcal{J}(\mathbf{P})$ 就不含幂等元, 进而根据『[命题 3.5.10](#)』可知其必然幂零. 由于 \mathcal{A} 半单, 它不含幂零理想, 于是 $\mathcal{J}(\mathbf{P}) = \{\mathbf{0}\}$. 现在注意『[定理 3.5.11](#)』中的 $\mathcal{R}(\mathbf{P})\mathcal{L}(\mathbf{P})$ 这个集合, 它由所有从左右两个方向都会被 \mathbf{P} 零化的元素构成. 因此, $\mathcal{R}(\mathbf{P})\mathcal{L}(\mathbf{P})$ 就是 $\mathcal{J}(\mathbf{P})$ 的子集, 从而 $\mathcal{R}(\mathbf{P})\mathcal{L}(\mathbf{P}) = \{\mathbf{0}\}$. 这就表明若 $\mathbf{r} \in \mathcal{R}(\mathbf{P})$ 且 $\mathbf{l} \in \mathcal{L}(\mathbf{P})$, 则 $\mathbf{rl} = \mathbf{0}$. 另一方面, 对任意 $\mathbf{l} \in \mathcal{L}(\mathbf{P})$ 以及 $\mathbf{r} \in \mathcal{R}(\mathbf{P})$, 我们有

$$(\mathbf{lr})^2 = \underbrace{\mathbf{l}(rl)}_{=0} \mathbf{r} = \mathbf{0}.$$

这就说明 $\mathcal{L}(\mathbf{P})\mathcal{R}(\mathbf{P})$ 这个理想(见『[习题 3.10](#)』)里面都是指数为 2 的幂零元, 从而它是零化理想, 而根据『[定理 3.5.3](#)』可知它是幂零的. 进而由 \mathcal{A} 的半单性可知 $\mathcal{L}(\mathbf{P})\mathcal{R}(\mathbf{P}) = \{\mathbf{0}\}$. 我们给 Peirce 分解两边左乘以 $\mathcal{L}(\mathbf{P})$, 然后利用前面给出的这些结果以及 $\mathcal{L}(\mathbf{P})\mathbf{P} = \{\mathbf{0}\}$, 就有

$$\mathcal{L}(\mathbf{P})\mathcal{A} = \mathcal{L}(\mathbf{P})\mathcal{R}(\mathbf{P})\mathbf{P} = \{\mathbf{0}\}.$$

特别地, 就有 $\mathcal{L}(\mathbf{P})\mathcal{L}(\mathbf{P}) = \{\mathbf{0}\}$, 从而 $\mathcal{L}(\mathbf{P})$ 是幂零的, 进而 $\mathcal{L}(\mathbf{P}) = \{\mathbf{0}\}$. 类似地, 可以得到 $\mathcal{R}(\mathbf{P}) = \{\mathbf{0}\}$. 最后 \mathcal{A} 的 Peirce 分解就只剩下 $\mathbf{P}\mathcal{A}\mathbf{P}$ 这一项了. \square

²⁵译者注: 此时对任意真子空间 \mathcal{B} 均有 $\mathcal{B}\mathcal{A} = \mathcal{A}\mathcal{B} = \{\mathbf{0}\} \subseteq \mathcal{B}$, 这就说明 \mathcal{B} 是理想.

定理 3.5.20. 若 \mathcal{A} 是半单代数, 则它必然含幺. 不仅如此, 它的主理想唯一, 且就是单位元.

证. 设 \mathbf{P} 是 \mathcal{A} 的主理想. 若 $\mathbf{b} \in \mathcal{A}$, 则根据『引理 3.5.19』可知 $\mathbf{b} \in \mathbf{P}\mathcal{A}\mathbf{P}$, 从而存在 $\mathbf{a} \in \mathcal{A}$ 使得 $\mathbf{b} = \mathbf{P}\mathbf{a}\mathbf{P}$. 进而

$$\begin{aligned}\mathbf{P}\mathbf{b} &= \mathbf{P}^2\mathbf{a}\mathbf{P} = \mathbf{P}\mathbf{a}\mathbf{P} = \mathbf{b}, \\ \mathbf{b}\mathbf{P} &= \mathbf{P}\mathbf{a}\mathbf{P}^2 = \mathbf{P}\mathbf{a}\mathbf{P} = \mathbf{b}.\end{aligned}$$

由于这对所有的 $\mathbf{b} \in \mathcal{A}$ 都成立, 这就说明 \mathbf{P} 是 \mathcal{A} 的单位元, \mathbf{P} 的任意性和单位元的唯一性则说明主理想唯一. \square

幂等元在某种意义上保代数的半单性, 具体含义见下述命题:

命题 3.5.21. 若 \mathcal{A} 是半单的, 那么对任意幂等元 $\mathbf{P} \in \mathcal{A}$, 子代数 $\mathbf{P}\mathcal{A}\mathbf{P}$ 同样是半单的.

证. 设 $\mathcal{N} = \text{Rad}(\mathbf{P}\mathcal{A}\mathbf{P})$, 并取 $\mathbf{x} \in \mathcal{N} \subseteq \mathbf{P}\mathcal{A}\mathbf{P}$. 则有 \mathcal{A} 的左理想 $\mathcal{A}\mathbf{x}$, 并且根据『(3.12) 式』可知 $\mathbf{x}\mathbf{P} = \mathbf{P}\mathbf{x} = \mathbf{x}$. 这样一来就有如下集合之间的恒等式:

$$\begin{aligned}(\mathcal{A}\mathbf{x})^{v+1} &= \mathcal{A}\mathbf{x}\mathcal{A}\mathbf{x} \dots \mathcal{A}\mathbf{x}\mathcal{A}\mathbf{x} = \mathcal{A}\mathbf{x}(\mathbf{P}\mathcal{A}\mathbf{P})\mathbf{x} \dots (\mathbf{P}\mathcal{A}\mathbf{P})\mathbf{x}(\mathbf{P}\mathcal{A}\mathbf{P})\mathbf{x} \\ &= \mathcal{A}\mathbf{x}(\mathbf{P}\mathcal{A}\mathbf{P}\mathbf{x})^v.\end{aligned}$$

由于 \mathcal{N} 是 $\mathbf{P}\mathcal{A}\mathbf{P}$ 的理想, 我们就有 $\mathbf{P}\mathcal{A}\mathbf{P}\mathbf{x} \subseteq \mathcal{N}$. 并且若 v 是 \mathcal{N} 的幂零指数, 则 $(\mathbf{P}\mathcal{A}\mathbf{P}\mathbf{x})^v = \{\mathbf{0}\}$. 于是 $\mathcal{A}\mathbf{x}$ 就是幂零的. 由于 \mathcal{A} 半单, 故必有 $\mathcal{A}\mathbf{x} = \{\mathbf{0}\}$. 这就表明对任意非零的 $\mathbf{a} \in \mathcal{A}$ 均有 $\mathbf{a}\mathbf{x} = \mathbf{0}$. 特别地, 就有 $\mathbf{P}\mathbf{x} = \mathbf{x} = \mathbf{0}$. 由于 \mathbf{x} 是 $\text{Rad}(\mathbf{P}\mathcal{A}\mathbf{P})$ 里的任意元素, 我们就证明了 $\text{Rad}(\mathbf{P}\mathcal{A}\mathbf{P}) = \{\mathbf{0}\}$. 这就说明 $\mathbf{P}\mathcal{A}\mathbf{P}$ 是半单的. \square

命题 3.5.22. 设 \mathcal{A} 为半单代数且 \mathbf{P} 是 \mathcal{A} 的幂等元. 那么 $\mathbf{P}\mathcal{A}\mathbf{P}$ 是可除代数的充要条件是 \mathbf{P} 为本原幂等元.

证. 假设 $\mathbf{P}\mathcal{A}\mathbf{P}$ 是可除代数. 根据『命题 3.5.2』, 它里面仅有的幂等元就是单位元. 而 \mathbf{P} 就是 $\mathbf{P}\mathcal{A}\mathbf{P}$ 的单位元, 因此 \mathbf{P} 就是 $\mathbf{P}\mathcal{A}\mathbf{P}$ 里面仅有的幂等元, 进而根据『命题 3.5.15』可知其为本原幂等元.

反过来, 设 \mathbf{P} 是本原幂等元. 现在设 $\mathbf{x} \in \mathbf{P}\mathcal{A}\mathbf{P}$ 非零. 则因 $\mathbf{P}\mathcal{A}\mathbf{P}$ 半单(『命题 3.5.21』), 左理想 $\mathcal{L} := (\mathbf{P}\mathcal{A}\mathbf{P})\mathbf{x}$ 不可能是幂零的. 于是根据『命题 3.5.13』可知它里面存在幂等元. 可是 \mathcal{L} 中的幂等元同样是 $\mathbf{P}\mathcal{A}\mathbf{P}$ 中的幂等元. 『命题 3.5.15』指出 \mathbf{P} 是 $\mathbf{P}\mathcal{A}\mathbf{P}$ 内唯一的幂等元, 从而也是 \mathcal{L} 内唯一的幂等元. 作为 \mathcal{L} 的元素, 我们可以将 \mathbf{P} 写作 $\mathbf{P} = \mathbf{a}\mathbf{x}$, 其中 $\mathbf{a} \in \mathbf{P}\mathcal{A}\mathbf{P}$. 由于 \mathbf{P} 是 $\mathbf{P}\mathcal{A}\mathbf{P}$ 里的单位元, 所以 \mathbf{x} 可逆. 由于 \mathbf{x} 是任选的, 故 $\mathbf{P}\mathcal{A}\mathbf{P}$ 中所有非零元都可逆, 进而它是可除代数. \square

直观上看, 单代数显然要比半单代数更为基本. 我们已经看到单代数必然是半单的, 而反过来当然不成立. 如果单代数确实更为基本, 那么半单代数就应当可以通过单代数“构建”出来. 要做到这一点, 我们就需要做一些准备工作.

引理 3.5.23. 设 \mathcal{B} 是代数 \mathcal{A} 的含幺理想, 其单位元为 $\mathbf{1}_B$. 那么, $\mathcal{A} = \mathcal{B} \oplus \mathcal{J}(\mathbf{1}_B)$, 其中 $\mathcal{J}(\mathbf{1}_B)$ 就是 \mathcal{A} 的 Peirce 分解中出现的那个理想.

证. 因为 $\mathbf{1}_B$ 是 \mathcal{A} 中的幂等元,^{*} 我们就可以得到下述 Peirce 分解:

$$\mathcal{A} = \mathbf{1}_B \mathcal{A} \mathbf{1}_B \oplus_V \mathbf{1}_B \mathcal{L}(\mathbf{1}_B) \oplus_V \mathcal{R}(\mathbf{1}_B) \mathbf{1}_B \oplus_V \mathcal{J}(\mathbf{1}_B) =: \mathcal{S}(\mathbf{1}_B) \oplus_V \mathcal{J}(\mathbf{1}_B),$$

其中 $\mathcal{S}(\mathbf{1}_B) = \mathbf{1}_B \mathcal{A} \mathbf{1}_B \oplus_V \mathbf{1}_B \mathcal{L}(\mathbf{1}_B) \oplus_V \mathcal{R}(\mathbf{1}_B) \mathbf{1}_B$. 因为 \mathcal{B} 是理想, 所以 $\mathcal{S}(\mathbf{1}_B)$ 中的每个求和项都是 \mathcal{B} 的子集, 从而 $\mathcal{S}(\mathbf{1}_B) \subseteq \mathcal{B}$. 若 $\mathbf{b} \in \mathcal{B}$, 则 $\mathbf{b} \in \mathcal{A}$, 从而根据上面的分解就有 $\mathbf{b} = \mathbf{b}_1 + \mathbf{b}_2$, 其中 $\mathbf{b}_1 \in \mathcal{S}(\mathbf{1}_B)$, $\mathbf{b}_2 \in \mathcal{J}(\mathbf{1}_B)$. 在它两边同时右乘 $\mathbf{1}_B$ 就得到

$$\mathbf{b} \mathbf{1}_B = \mathbf{b}_1 \mathbf{1}_B + \mathbf{b}_2 \mathbf{1}_B.$$

由于 $\mathbf{1}_B$ 是 \mathcal{B} 里的单位元, 而且 $\mathcal{J}(\mathbf{1}_B)$ 与 $\mathbf{1}_B$ 正交, 故上式指出 $\mathbf{b} = \mathbf{b}_1$. 这就说明 $\mathbf{b} \in \mathcal{S}(\mathbf{1}_B)$, 从而 $\mathbf{B} \subseteq \mathcal{S}(\mathbf{1}_B)$. 两相结合就有 $\mathcal{B} = \mathcal{S}(\mathbf{1}_B)$, 进而 $\mathcal{A} = \mathcal{B} \oplus_V \mathcal{J}(\mathbf{1}_B)$. 因为 $\mathcal{J}(\mathbf{1}_B)\mathcal{B} = \mathcal{B}\mathcal{J}(\mathbf{1}_B) = \{\mathbf{0}\}$, 我们可以进一步将 \oplus_V 替换为 \oplus . \square

引理 3.5.24. 半单代数的非零理想是半单的.

证. 设 \mathcal{A} 是半单代数, \mathcal{B} 是 \mathcal{A} 的非零理想. 由于 $\text{Rad}(\mathcal{B})$ 是 \mathcal{B} 的理想, 是故 $\mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B} \subseteq \text{Rad}(\mathcal{B})$. 不仅如此, 因为 \mathcal{B} 是 \mathcal{A} 的理想, 故 $\mathcal{A}\mathcal{B} \subseteq \mathcal{B}$ 且 $\mathcal{B}\mathcal{A} \subseteq \mathcal{B}$. 由此即可推出 $\mathcal{A}(\mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B})\mathcal{A} = (\mathcal{A}\mathcal{B}) \text{Rad}(\mathcal{B})(\mathcal{B}\mathcal{A}) \subseteq \mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B}$, 也就是说 $\mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B}$ 是 \mathcal{A} 的理想. 不仅如此, 由于它还含有 $\text{Rad}(\mathcal{B})$, 故幂零. 而 \mathcal{A} 的半单性则意味着 $\mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B} = \{\mathbf{0}\}$. 因为 $\text{Rad}(\mathcal{B}) \subseteq \mathcal{B}$, 所以 $\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \subseteq \mathcal{B}$, 进而

$$\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \subseteq \mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B}.$$

现在注意

$$\begin{aligned} (\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A})^2 &= \mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \\ &\subseteq \mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} \\ &\subseteq \mathcal{B} \text{Rad}(\mathcal{B})\mathcal{B} = \{\mathbf{0}\}, \end{aligned}$$

这就说明 $\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A}$ 幂零. 考虑到它是 \mathcal{A} 的理想, 而 \mathcal{A} 半单, 故

$$\mathcal{A} \text{Rad}(\mathcal{B})\mathcal{A} = \{\mathbf{0}\}.$$

根据『定理 3.5.20』可知 \mathcal{A} 存在单位元, 从而就有 $\text{Rad}(\mathcal{B}) = \{\mathbf{0}\}$, 这就说明 \mathcal{B} 半单. \square

* 注意 $\mathbf{1}_B$ 不是 \mathcal{A} 的单位元. 它仅满足 $\mathbf{x} \in \mathcal{B}$ 时有 $\mathbf{x}\mathbf{1}_B = \mathbf{1}_B\mathbf{x} = \mathbf{x}$.

定理 3.5.25. 一个代数是半单代数的充要条件是其可以写作单代数的直和.

证. 若 \mathcal{A} 是单代数的直和, 则根据『[命题 3.2.11](#)』可知 \mathcal{A} 的理想要么就是这些分量的直和, 要么含于这些分量. 无论是何种情况, 这些理想都不可能是零的, 因为单代数是半单的. 是故 \mathcal{A} 半单.

反过来, 假设 \mathcal{A} 半单. 如果它不含真理想, 那么它自然就是单的, 从而半单, 那么不必再做什么了. 因此, 我们假设 \mathcal{A} 存在非零的真理想 \mathcal{B} . 那么根据『[引理 3.5.24](#)』可知 \mathcal{B} 半单, 并且根据『[定理 3.5.20](#)』可知 \mathcal{B} 有单位元 $\mathbf{1}_B$. 接下来利用『[引理 3.5.23](#)』我们就有 $\mathcal{A} = \mathcal{B} \oplus \mathcal{J}(\mathbf{1}_B)$. 如果这两个分量里面有一个不是单的, 那么就重复这个过程. 最后就将 \mathcal{A} 写成了一些单代数的直和. \square

定理 3.5.26. 在不计诸分量求和顺序意义下, 从单代数到半单代数的分解是唯一的.

证. 假设 $\mathcal{A} = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_r$, 其中 \mathcal{A}_i 为单代数. 那么 \mathcal{A} 的单位元就是各个分量中单位元之和: $\mathbf{1} = \mathbf{1}_1 + \dots + \mathbf{1}_r$. 现在设 $\mathcal{A} = \mathcal{A}'_1 \oplus \cdots \oplus \mathcal{A}'_r$ 是另一个分解. 在单位元的分解式两边左乘 \mathcal{A}'_j 就有

$$\mathcal{A}'_j = \mathcal{A}'_j \mathbf{1}_1 + \cdots + \mathcal{A}'_j \mathbf{1}_r =: \mathcal{A}'_{ji} + \cdots + \mathcal{A}'_{jr} = \sum_{i=1}^r \mathcal{A}'_{ji}.$$

因为 $\mathbf{1}_i \in \mathcal{A}_i$, 而 \mathcal{A}_i 是 \mathcal{A} 的理想, 故 $\mathcal{A}'_{ji} \subseteq \mathcal{A}_i$. 又因为 \mathcal{A}_i 之间不相交, 从而 \mathcal{A}'_{ji} 之间也是不相交的. 由于 \mathcal{A}'_j 是个理想, 很容易验证 $\mathcal{A}'_j \mathbf{1}_i$ 是个代数. 不仅如此, 因为 $i \neq j$ 时 $\mathbf{1}_i \mathbf{1}_k = \mathbf{0}$, 上面的那个求和就是代数的直和. 因此, 根据『[命题 3.2.11](#)』可知 \mathcal{A}'_{ji} 是个理想, 而且因为它是 \mathcal{A}_i 的子集, 它就是 \mathcal{A}_i 的子理想. 由于 \mathcal{A}_i 是单的, 故 $\mathcal{A}'_{ji} = \mathcal{A}_i$ 或者 $\mathcal{A}'_{ji} = \{\mathbf{0}\}$. 因为 \mathcal{A}'_j 是单的, 它的分量仅有一个是非零的, 从而这个非零分量自然就是某个 \mathcal{A}_i . \square

3.5.3. 单代数的分类

『定理 3.5.25』以及『定理 3.5.26』通过将所有半单代数表示为单代数的直和实现了对半单代数(也就是零根代数)的分类。那么一个一般的代数是否可以用它的根以及一个半单代数表示呢？可以证明，任意一个有非零根 $\text{Rad}(\mathcal{A})$ 的代数 \mathcal{A} 都可以写成直和 $\mathcal{A} = \text{Rad}(\mathcal{A}) \oplus (\mathcal{A}/\text{Rad}(\mathcal{A}))$ ，也就是写成它的根加其模掉根后的商代数。由于在 $\mathcal{A}/\text{Rad}(\mathcal{A})$ 中，我们已经将根从 \mathcal{A} “提取出来”了，所以这个商代数确实是半单的。这个结果就是著名的 **Wedderburn 主结构定理** (Wedderburn principle structure theorem)，它将对代数的研究简化成了对单代数的研究。单代数则可以进一步分解(其证明可见 [Benn and Tucker, 1987], pp. 330-332)：

定理 3.5.27 (Wedderburn 分解). 代数 \mathcal{A} 是单代数的充要条件是

$$\mathcal{A} \cong \mathcal{D} \otimes \mathcal{M}_n \cong \mathcal{M}_n(\mathcal{D}),$$

其中 \mathcal{D} 是一个可除代数，而 $\mathcal{M}_n(\mathcal{D})$ 是 \mathcal{D} 上的全矩阵代数，其中 n 是某个非负整数。这里 \mathcal{D} 和 $\mathcal{M}_n(\mathcal{D})$ 在相差相似变换的意义下是唯一的。

我们将 \mathcal{M}_n 的中心记作 \mathcal{Z}_n 。因为 \mathcal{M}_n 是中心代数，故根据『定理 3.3.2』可知 $\mathcal{Z}_n = \text{Span } \mathbf{1}_n$ 。另一方面，由『(3.8) 式』有

$$\mathcal{Z}(\mathcal{A}) = \mathcal{Z}(\mathcal{D}) \otimes \mathcal{Z}_n \cong \mathcal{Z}(\mathcal{D}), \quad (3.14)$$

这一关系指出我们可以通过对 \mathcal{A} 中心的了解来确定 \mathcal{D} 。

命题 3.5.28. 复数集 \mathbb{C} 上的可除代数只有 \mathbb{C} 自己。

证. 设 \mathcal{D} 是 \mathbb{C} 上的可除代数， \mathbf{x} 是 \mathcal{D} 里的一个非零元素。因为 \mathcal{D} 是有限维的，所以就存在 \mathbf{x} 的多项式使得*

$$f(\mathbf{x}) = \mathbf{x}^n + \alpha_{n-1}\mathbf{x}^{n-1} + \cdots + \alpha_1\mathbf{x} + \alpha_0\mathbf{1} = \mathbf{0}.$$

设 n 是使其成立的最小整数。根据代数学基本定理(见『??节』)， $f(\mathbf{x})$ 至少存在一个根 λ 。于是我们就有

$$f(\mathbf{x}) = (\mathbf{x} - \lambda\mathbf{1})g(\mathbf{x}) = \mathbf{0}.$$

此时 $g(\mathbf{x})$ 最多是 $n-1$ 次的，并且根据假设它不能是零²⁶。由于 \mathcal{D} 是可除代数，所以 $g(\mathbf{x})$ 可逆。进而 $\mathbf{x} - \lambda\mathbf{1} = \mathbf{0}$ ，这就说明 \mathcal{D} 的每个元素都是 $\mathbf{1}$ 的倍数，从而 \mathcal{D} 同构于 \mathbb{C} 。□

²⁶译者注：如果 $g(\mathbf{x}) = \mathbf{0}$ ，则我们就找到了次数小于 n 且使得 $p(\mathbf{x}) = \mathbf{0}$ 的多项式 $p = g$ ，这就与前面 n 是使其成立的最小整数矛盾了。

结合『[命题 3.5.28](#)』、『[定理 3.5.27](#)』, 并注意到 $\mathcal{M}_n(\mathbb{C})$ 是中心代数的事实(『[定理 3.3.2](#)』), 我们就得到下面的结果:

定理 3.5.29. 复数集 \mathbb{C} 上的任意单代数 \mathcal{A} 都和某个 n 下的 $\mathcal{M}_n(\mathbb{C})$ 同构, 从而 \mathcal{A} 必然是中心单代数.

复单代数一定是中心代数这个结果也可以通过『[\(3.14\) 式](#)』以及『[命题 3.5.28](#)』得出.

抽象代数中有个 **Frobenius 定理** (Frobenius theorem), 它指出 \mathbb{R} 上的可除代数有且仅有 $\mathbb{R}, \mathbb{C}, \mathbb{H}$, 而且因为两个可除代数的张量积也是可除代数, 所有还有 $\mathbb{C} \otimes \mathbb{H}$. * 不仅如此, \mathbb{C} 的中心就是整个 \mathbb{C} , 因为它是交换代数. 而另一方面, \mathbb{H} 是中心代数, 也就是说它的中心由单位元张成 (请读者验证之), 因此 \mathbb{H} 的中心同构于 \mathbb{R} .

现在我们考察 \mathbb{R} 上的单代数 \mathcal{A} . 若 \mathcal{A} 是中心代数, 即若 $\mathcal{Z}(\mathcal{A}) = \mathbb{R}$, 则『[\(3.14\) 式](#)』就给出

$$\mathbb{R} \cong \mathcal{Z}(\mathcal{D}) \Rightarrow \mathcal{D} = \mathbb{R} \text{ 或 } \mathbb{H}.$$

若 $\mathcal{Z}(\mathcal{A}) = \mathbb{C}$, 则

$$\mathbb{C} \cong \mathcal{Z}(\mathcal{D}) \Rightarrow \mathcal{D} = \mathbb{C} \text{ 或 } \mathbb{C} \otimes \mathbb{H}.$$

上述结果, 连带着 Frobenius 定理和 Wedderburn 定理就给出

定理 3.5.30. 实数集 \mathbb{R} 上的任意单代数 \mathcal{A} 都同构于 $\mathcal{D} \otimes \mathcal{M}_n$, 其中 n 是某个正整数. 若 \mathcal{A} 的中心同构于 \mathbb{C} , 则 \mathcal{D} 要么是 \mathbb{C} , 要么是 $\mathbb{C} \otimes \mathbb{H}$; 若 \mathcal{A} 是中心代数 (也就是其中心同构于 \mathbb{R}), 则 \mathcal{D} 就是 \mathbb{R} 或者 \mathbb{H} .

我们接下来进一步来刻画单代数, 并建立起本原幂等元与极小左理想之间的关系, 并以此结束对代数分解的讨论.

定义 3.5.31 (相似幂等元). 设 \mathbf{P}, \mathbf{P}' 是代数 \mathcal{A} 的幂等元, 若存在可逆元 $\mathbf{s} \in \mathcal{A}$ 使得 $\mathbf{P}' = \mathbf{s}\mathbf{P}\mathbf{s}^{-1}$, 则称这两个幂等元相似 (*similar*).

下面这个定理的证明比较难, 故略去. (参见 [[Benn and Tucker, 1987](#)], pp. 332-334).

定理 3.5.32. 若 \mathbf{P} 是代数 \mathcal{A} 的幂等元, 则必然存在相互正交的本原幂等元 $\{\mathbf{P}_i\}_{i=1}^r$ 使得 $\mathbf{P} = \sum_{i=1}^r \mathbf{P}_i$. 这里的整数 r 是唯一的, 称作 \mathbf{P} 的秩 (*rank*). 两个幂等元相似的充要条件是它们的秩相等.

最后这个定理揭示了本原幂等元和极小左 (右) 理想之间的关系.

* 因为 \mathbb{C} 是 \mathbb{H} 的子代数, 这个张量积实际上是冗余的. 不过, 在本书后面对 Clifford 代数进行分类时, \mathbb{C} 有时会被显式地提取出来.

定理 3.5.33. 设 \mathbf{P} 是半单代数 \mathcal{A} 的本原幂等元. 那么 \mathcal{AP} 和 \mathbf{PA} 就分别是 \mathcal{A} 的极小左理想和极小右理想.

证. 因为半单代数是一些单代数的直和, 并且每个分量都和其他分量独立, 所以不失一般性, 可以假设 \mathcal{A} 是单代数. 假设 $\mathcal{L} := \mathcal{AP}$ 不是极小的. 那么 \mathcal{L} 就包含了某个 \mathcal{A} 的非零左理想 \mathcal{L}_1 . 由于 \mathcal{A} 是(半)单的, 所以 \mathcal{L}_1 不是幂零的. 于是根据『[命题 3.5.10](#)』可知其包含一个幂等元 \mathbf{P}_1 . 如果 $\mathbf{P}_1 = \mathbf{P}$, 那么

$$\mathcal{L} = \mathcal{AP} = \mathcal{AP}_1 \subseteq \mathcal{L}_1,$$

从而 $\mathcal{L} = \mathcal{L}_1$, 我们就证完了. 因此假设 $\mathbf{P}_1 \neq \mathbf{P}$. 这样根据『[定理 3.5.16](#)』就有

$$\mathbf{P}_1 = \mathbf{Q}_1 + \cdots + \mathbf{Q}_r,$$

其中 \mathbf{Q}_i 都是本原幂等元, 并且彼此正交. 因为 \mathbf{Q}_1 和 \mathbf{P} 的秩都是 1, 根据『[定理 3.5.32](#)』可知它们相似, 于是存在可逆元 $s \in \mathcal{A}$ 使得 $\mathbf{P} = s\mathbf{Q}_1s^{-1}$. 必要情况下可以将 \mathbf{P}_1 替换为 $s\mathbf{P}_1s^{-1}$, * 我们可以假定 $\mathbf{Q}_1 = \mathbf{P}$, 从而

$$\mathbf{P}_1 = \mathbf{P} + \mathbf{Q}_2 + \cdots + \mathbf{Q}_r,$$

并且 \mathbf{P} 与所有的 \mathbf{Q}_i 正交. 在上式两边左乘 \mathbf{P} , 我们就得到 $\mathbf{PP}_1 = \mathbf{P}$, 从而

$$\mathcal{L} = \mathcal{AP} = \mathcal{APP}_1 \subseteq \mathcal{L}_1,$$

这就推出 $\mathcal{L} = \mathcal{L}_1$. 右理想的情形类似. □

* 这等价于将 \mathcal{L} 替换为 $s\mathcal{L}s^{-1}$. 根据『[定理 3.2.7](#)』以及『[定理 3.5.27](#)』中的非唯一性部分(可以相差相似变换), 这一操作是允许的.

3.6. 多项式代数

设 \mathcal{A} 是带单位元 $\mathbf{1}$ 的结合代数. 对 \mathcal{A} 中任意固定元素 \mathbf{a} , 我们考虑集合 $\mathcal{P}[\mathbf{a}]$, 它由 \mathcal{A} 内那些具有

$$p(\mathbf{a}) := \sum_{k=0}^{\infty} \alpha_k \mathbf{a}^k, \quad \alpha_k \in \mathbb{C}$$

形式的元素构成, 并且我们要求这里的求和中只有有限项是非零的. 它们很明显是 \mathbf{a} 的多项式, 其中的加法和乘法都继承自 \mathcal{A} 的代数结构.

定义 3.6.1 (多项式代数). 设 \mathcal{A} 是带单位元 $\mathbf{1}$ 的结合代数. 对任意固定元素 $\mathbf{a} \in \mathcal{A}$, 集合 $\mathcal{P}[\mathbf{a}]$ 是个带单位元的交换代数, 称作 \mathbf{a} 生成的多项式代数 (*polynomial algebra*). 将多项式 $p(\mathbf{a}) = \sum_{k=0}^{\infty} \alpha_k \mathbf{a}^k$ 的非零项中 \mathbf{a} 幂次最高的那一项称作 p 的首项 (*leading term*), 其系数就是首项系数 (*leading coefficient*); 首项里面 \mathbf{a} 的幂次称作 p 的次数 (*degree*), 记作 $\deg p$; 称 α_0 为标量项 (*scalar term*). 首项系数为 1 的多项式就称作首一多项式 (*monic polynomial*). 形式为 $\alpha_n \mathbf{a}^n$ 的非零多项式称作 n 次单项式 (*monomial*).

很明显 $\{\mathbf{a}^k\}_{k=1}^{\infty}$ 是多项式代数 $\mathcal{P}[\mathbf{a}]$ 的一组基.

若 $p(\mathbf{a}) := \sum_{k=0}^{\infty} \alpha_k \mathbf{a}^k$, $q(\mathbf{a}) := \sum_{j=0}^{\infty} \beta_j \mathbf{a}^j$, 那么

$$(p+q)(\mathbf{a}) = \sum_{k=0}^{\infty} (\alpha_k + \beta_k) \mathbf{a}^k,$$

$$(pq)(\mathbf{a}) = \sum_{i=0}^{\infty} \gamma_i \mathbf{a}^i, \quad \text{其中 } \gamma_i = \sum_{j+k=i} \alpha_j \beta_k.$$

考察任意两个非零多项式 $p(\mathbf{a})$ 以及 $q(\mathbf{a})$. 那么显然就有

$$\deg(p+q) \leq \max(\deg p, \deg q),$$

$$\deg(pq) = \deg p + \deg q.$$

定义 3.6.2 (多项式代数里的微分映射). 称由 $d\mathbf{a}^k = k\mathbf{a}^{k-1}$, $k \geq 1$ 以及 $d\mathbf{a}^0 := d\mathbf{1} = \mathbf{0}$ 定义的线性映射 $d : \mathcal{P}[\mathbf{a}] \rightarrow \mathcal{P}[\mathbf{a}]$ 为 $\mathcal{P}[\mathbf{a}]$ 里的微分映射 (*differentiation map*).

定理 3.6.3. 『[定义 3.6.2](#)』定义的微分映射 d 是 $\mathcal{P}[\mathbf{a}]$ 上的导子. 我们将 dp 记作 p' .

证. 证明很简单, 留作习题 ([『习题 3.35』](#)).

□

设 p, q 为多项式, 那么 $d(pq) = d(p)q + pd(q)$, 特别地就有

$$d(q^2) = 2q d(q),$$

进一步归纳可证

$$\mathbf{d}(q^k) = kq^{k-1}\mathbf{d}(q), \quad k \geq 1, \mathbf{d}(q^0) = \mathbf{0}.$$

由于 q 是 \mathcal{A} 中的元素, 它同样也可以生成一个多项式. 比如说, 我们就可以通过将 \mathbf{a} 替换为 q 构造出多项式 $p(q)$:

$$p(q) = \sum_{k=0}^{\infty} \alpha_k q^k.$$

不难证明 (见『习题 3.36』)

$$\mathbf{d}(p(q)) = p'(q)q'.$$

这就是多项式微分的链式法则 (chain rule).

定义 3.6.4 (多项式的高阶导数). 将多项式 $\mathbf{d}^r(p)$ 称作 p 的 r 阶导数 (r th derivative), 记作 $p^{(r)}$. 并约定 $p^{(0)} = p$.

很明显, 若 $r > \deg p$, 那么 $p^{(r)} = 0$.

现在考察单项式 \mathbf{a}^n , 那么就有

$$\mathbf{d}^r(\mathbf{a}^n) = \frac{n!}{(n-r)!} \mathbf{a}^{n-r},$$

亦或者说

$$\mathbf{a}^{n-r} = \frac{(n-r)!}{n!} \mathbf{d}^r(\mathbf{a}^n).$$

接下来利用二项式定理就有²⁷

$$(\mathbf{a} + \mathbf{b})^n = \sum_{r=0}^n \binom{n}{r} \mathbf{a}^{n-r} \mathbf{b}^r = \sum_{r=0}^n \frac{1}{r!} \mathbf{d}^r(\mathbf{a}^n) \mathbf{b}^r.$$

上式左边是多项式 $p(\mathbf{a} + \mathbf{b})$ 中任意一项. 因此, 取这些项的线性组合, 我们就得到

$$p(\mathbf{a} + \mathbf{b}) = \sum_{r=0}^n \frac{p^r(\mathbf{a})}{r!} \mathbf{b}^r.$$

这就是 p 的 Taylor 公式 (Taylor formula).

若标量 $\lambda \in \mathbb{C}$ 满足 $p(\lambda) = \sum_{k=0}^n \eta_k \lambda^k = 0$, 则称 λ 是 n 次多项式 $p(\mathbf{a}) = \sum_{k=0}^n \eta_k \mathbf{a}^k$ 的根. 代数基本定理^{*}指出 \mathbb{C} 是代数闭域, 也就是说任意复系数多项式都可以分解为一些一次复系数多项式的乘积. 于是

$$p(\mathbf{a}) = \eta_n (\mathbf{a} - \lambda_1 \mathbf{1})^{k_1} \cdots (\mathbf{a} - \lambda_s \mathbf{1})^{k_s}, \quad (3.15)$$

* 该定理的证明可见『??节』.

²⁷译者注: 这里可能有点小问题. 二项式定理成立的前提是 \mathbf{a} 和 \mathbf{b} 对易, 即 $\mathbf{a}\mathbf{b} = \mathbf{b}\mathbf{a}$. 但前面我们只要求 \mathcal{A} 是结合代数, 没有进一步要求其为交换代数. 不过在多项式理论里面我们通常都要求交换性.

其中 $\eta_n \neq 0$, $\{\lambda_i\}_{i=1}^s$ 是这个多项式的互异复根. 上式中的 k_i 称作根 λ_i 的重数 (multiplicity), 这是一个非负整数, 且满足 $\sum_{i=1}^s k_i = n$.

正如最简单的例子 $p(\mathbf{a}) = \mathbf{a}^2 + 1$ 指出的那样, \mathbb{R} 并不是代数闭域. 虽说如此, 实系数多项式总是可以分解为实系数的一次或者二次多项式之积. 要说明这一点, 我们首先注意到若 λ 是某个实系数多项式的复根, 那么它的复共轭 $\bar{\lambda}$ 同样是这个多项式的一个根. 这是因为若 $\sum_{k=0}^n \eta_k \lambda^k = 0$, 则在其两边取复共轭就有 $\sum_{k=0}^n \bar{\eta}_k \bar{\lambda}^k = 0$. 现在注意 η_k 为实数从而 $\bar{\eta}_k = \eta_k$, 这就表明 $\bar{\lambda}$ 同样是根. 不仅如此, λ 和 $\bar{\lambda}$ 必然也有相同的重数, 不然那些无法匹配的因式相乘就会给出一个复系数的多项式, 而那些能够相互匹配的因式相乘则总是实系数的多项式. 最终二者乘积就是复系数的, 这与 $p(\mathbf{a})$ 是实系数多项式矛盾.

在上述论证的基础上, 我们现在对『(3.15) 式』进行重排, 使得每个复根对应的因式都和其共轭复根对应的因式相乘. 注意到若 $\lambda_m = \gamma_m + i\xi_m$, 则有

$$\begin{aligned} (\mathbf{a} - \lambda_m \mathbf{1})^{k_m} (\mathbf{a} - \bar{\lambda}_m \mathbf{1})^{k_m} &= (\mathbf{a} - \gamma_m \mathbf{1} - i\xi_m \mathbf{1})^{k_m} (\mathbf{a} - \gamma_m \mathbf{1} + i\xi_m \mathbf{1})^{k_m} \\ &= (\mathbf{a}^2 - 2\gamma_m \mathbf{a} + \gamma_m^2 \mathbf{1} + \xi_m^2 \mathbf{1})^{k_m} \\ &=: (\mathbf{a} + \alpha_m \mathbf{a} + \beta_m \mathbf{1})^{k_m}, \quad \alpha_m^2 < 4\beta_m. \end{aligned}$$

这里的不等号保证了 $\xi_m \neq 0$, 即这里的根确实是复根. 至此, 我们就证明了下面的结果:

定理 3.6.5. 任意一个 n 次实系数多项式 $p(\mathbf{a}) = \sum_{k=0}^n \eta_k \mathbf{a}^k$ 都具有如下因式分解:

$$p(\mathbf{a}) = \eta_n \prod_{i=1}^r (\mathbf{a} - \lambda_i \mathbf{1})^{k_i} \prod_{j=1}^R (\mathbf{a}^2 + \alpha_j \mathbf{a} + \beta_j \mathbf{1})^{K_j}, \quad \alpha_j^2 < 4\beta_j.$$

这里 $\lambda_i, \alpha_j, \beta_j \in \mathbb{R}$, $k_i, K_j \in \mathbb{N}$, 并且 λ_i 互异, 数对 (α_j, β_j) 互异, 且 $2 \sum_{j=1}^R K_j + \sum_{i=1}^r k_i = n$.

推论 3.6.6. 奇数次的实系数多项式都至少有一个实根.

3.7. 本章习题

习题 3.1. 证明下述结论:

(a) 若在 \mathbb{R}^2 上定义下述乘积:

$$(x_1, x_2)(y_1, y_2) = (x_1y_1 - x_2y_2, x_1y_2 + x_2y_1),$$

那么 \mathbb{R}^2 就成了结合的交换代数.

(b) 在 \mathbb{R}^3 上以叉积为乘法, 其结果是个非结合非交换的代数.

习题 3.2. 证明白代数的中心是其子空间. 并且若代数还是结合的, 则其中心进一步是子代数.

习题 3.3. 求证: A 的导出代数 A^2 确实是个代数.

习题 3.4. 证明所有 $n \times n$ 矩阵的集合 A , 在『(3.3) 式』定义的乘积下, 是个非结合非交换的代数.

习题 3.5. 证明所有 $n \times n$ 上三角矩阵的集合 A , 在常规的矩阵乘法下, 是个非交换的结合代数; 而在『(3.3) 式』定义的乘法下则是非结合非交换的代数, 并证明在这种情况下其导出代数 $A^2 := B$ 就是严格上三角矩阵的集合. 此外, B 的导出代数 B^2 是怎样的?

习题 3.6. 证明『命题 3.1.23』.

习题 3.7. 设 $\omega \in \mathcal{L}(\mathcal{V})$ 满足 $\omega(\mathbf{a}) = -\mathbf{a}$ 对所有 $\mathbf{a} \in \mathcal{V}$ 成立. 请问 ω 是 \mathcal{V} 的对合吗? 现在假设 \mathcal{V} 是个代数. 这样定义的 ω 依旧是代数 \mathcal{V} 的对合吗? 注意代数的对合必须是这个代数的同态.

习题 3.8. 求证: 含幺代数的真左(右)理想不可能包含存在左(右)逆的元素.

习题 3.9. 设 A 是结合代数, $\mathbf{x} \in A$. 证明 $A\mathbf{x}$ 是个左理想, $\mathbf{x}A$ 是个右理想, $A\mathbf{x}A$ 是个双边理想.

习题 3.10. 设 \mathcal{L} 是左理想, \mathcal{R} 是右理想. 证明 $\mathcal{L}\mathcal{R}$ 就是个双边理想.

习题 3.11. 证明『定理 3.1.25』中定义的 Φ 确实是个代数同构.

习题 3.12. 证明『例 3.1.18』中给出的那个线性变换还是代数 A 和 B 之间的同构.

习题 3.13. 设 A 是带单位元 $\mathbf{1}_A$ 的代数, ϕ 是 A 映满到另一个代数 B 上的满同态. 证明 $\phi(\mathbf{1}_A)$ 就是 B 的单位元.

习题 3.14. 证明 \mathcal{A} 的导出代数是 \mathcal{A} 里的理想.

习题 3.15. 证明四元数代数是中心代数.

习题 3.16. 写出四元数代数的所有结构常数. 并证明这个代数是结合的.

习题 3.17. 证明一个四元数是纯四元数的充要条件是其平方为非正实数.

习题 3.18. 设 p, q 是两个四元数. 求证:

- (a) $(pq)^* = q^* p^*$.
- (b) $q \in \mathbb{R}$ 的充要条件是 $q^* = q$; $q \in \mathbb{R}^3$ 的充要条件是 $q^* = -q$.
- (c) $qq^* = q^*q$ 是个非负实数.

习题 3.19. 证明『(3.7) 式』.

习题 3.20. 证明『例 3.2.16』中的 $\bar{\phi}$ 是个代数同态.

习题 3.21. 证明『定理 3.3.2』.

习题 3.22. 设代数 \mathcal{A} 的一组基为 $\{\mathbf{1}, \mathbf{e}\}$, 并满足 $\mathbf{e}^2 = \mathbf{1}$.

- (a) 若 $\mathbf{f}_1 = \frac{1}{2}(\mathbf{1} + \mathbf{e})$, $\mathbf{f}_2 = \frac{1}{2}(\mathbf{1} - \mathbf{e})$, 则 $\{\mathbf{f}_1, \mathbf{f}_2\}$ 同样是一组基.
- (b) 证明 $\mathcal{A} = \mathcal{L}_1 \oplus_V \mathcal{L}_2$, 其中 $\mathcal{L}_i = \mathcal{A}\mathbf{f}_i$, $i = 1, 2$, 并且 \oplus_V 表示矢量空间的直和.
- (c) 证明 $\mathcal{L}_1, \mathcal{L}_2$ 实际上是双边理想, 并且 $\mathcal{L}_1\mathcal{L}_2 = \{\mathbf{0}\}$. 因此 $\mathcal{A} = \mathcal{L}_1 \oplus \mathcal{L}_2$.
- (d) 用 \mathcal{A} 的任意元素与 \mathcal{L}_i , $i = 1, 2$ 里任意元素相乘, 以此证明 $\mathcal{L}_i = \text{Span}\{\mathbf{f}_i\}$, $i = 1, 2$. 因此, $\mathcal{L}_i \cong \mathbb{R}$, $i = 1, 2$. 进而就有 $\mathcal{A} \cong \mathbb{R} \oplus \mathbb{R}$.

习题 3.23. 设 \mathcal{A} 是个代数, \mathbf{D} 是 \mathcal{A} 中的导子. 证明中心 $\mathcal{Z}(\mathcal{A})$ 和导出代数 \mathcal{A}^2 都在 \mathbf{D} 下不变. 也即是说, 若 $\mathbf{a} \in \mathcal{Z}(\mathcal{A})$, 则 $\mathbf{D}(\mathbf{a}) \in \mathcal{Z}(\mathcal{A})$; 若 $\mathbf{a} \in \mathcal{A}^2$, 则 $\mathbf{D}(\mathbf{a}) \in \mathcal{A}^2$.

习题 3.24. 设 $\mathbf{D} : \mathcal{A} \rightarrow \mathcal{A}$ 为导子. 证明 $\ker \mathbf{D}$ 是 \mathcal{A} 的子代数.

习题 3.25. 证明两个导子的线性组合还是导子.

习题 3.26. 固定一个矢量 $\mathbf{a} \in \mathbb{R}^3$, 并定义线性变换 $\mathbf{D}_{\mathbf{a}} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ 为 $\mathbf{D}_{\mathbf{a}}(\mathbf{b}) = \mathbf{a} \times \mathbf{b}$. 证明 $\mathbf{D}_{\mathbf{a}}$ 是 \mathbb{R}^3 在叉积这个乘法下的导子.

习题 3.27. 在 $C^r(a, b)$ 上定义 $\mathbf{D}(f) = f'(c)$, 其中 $a < c < b$. 求证: 若将 ϕ_c 定义为演化映射 $\phi_c(f) = f(c)$, 则 \mathbf{D} 就是个 ϕ_c 导子.

习题 3.28. 设 $\Omega \in \text{End}(\mathcal{A})$ 是 \mathcal{A} 关于 ω 的反导子. 求证: $\ker \Omega$ 是 \mathcal{A} 的子代数, 并且若 \mathcal{A} 是含幺代数, 则 $\Omega(\mathbf{e}) = \mathbf{0}$.

习题 3.29. 导出 Leibniz 公式『(3.4.5) 式』.

习题 3.30. 证明『定理 3.4.10』.

习题 3.31. 求证: $n \times n$ 严格上三角矩阵的代数是指数为 n 的幂零代数.

习题 3.32. 设 \mathbf{b} 是代数 \mathcal{B} 里的固定元素. 考察由 $\mathbf{T}_b(\mathbf{x}) = \mathbf{x}\mathbf{b}$ 定义的线性变换 $\mathbf{T}_b : \mathcal{B} \rightarrow \mathcal{B}$. 利用维数定理, 证明若 $\mathcal{B}\mathbf{b} = \mathcal{B}$, 则有 $\ker \mathbf{T}_b = \mathbf{0}$.

习题 3.33. 设 \mathcal{A} 是带幂等元 \mathbf{P} 的代数. 证明 \mathbf{PAP} 由那些使得 $\mathbf{aP} = \mathbf{Pa} = \mathbf{a}$ 的元素 \mathbf{a} 构成. 对『定理 3.5.11』中的子空间, 令 $\mathcal{A}_1 := \mathbf{PAP}$, $\mathcal{A}_2 := \mathbf{PL}(\mathbf{P})$, $\mathcal{A}_3 := \mathbf{R}(\mathbf{P})\mathbf{P}$, $\mathcal{A}_4 := \mathbf{J}(\mathbf{P})$. 证明 $\{\mathcal{A}_i\}_{i=1}^3$ 是 \mathcal{A} 的子代数, 并且 $\mathcal{A}_i \cap \mathcal{A}_j = \{\mathbf{0}\}$, 但是对所有 $i \neq j, i, j = 1, \dots, 4$ 均有 $\mathcal{A}_i\mathcal{A}_j \neq \{\mathbf{0}\}$. 因此, Peirce 分解就是个矢量空间直和, 但不是代数直和.

习题 3.34. 设 \mathbf{p} 和 \mathbf{q} 是正交的幂等元. 假设 $\mathbf{q} = \mathbf{q}_1 + \mathbf{q}_2$, 其中 \mathbf{q}_1 和 \mathbf{q}_2 是正交幂等元. 证明 $\mathbf{qq}_i = \mathbf{q}_i\mathbf{q} = \mathbf{q}_i, i = 1, 2$. 利用这一结果, 证明 $i = 1, 2$ 时 $\mathbf{pq}_i = \mathbf{q}_i\mathbf{p} = \mathbf{0}$.

习题 3.35. 在 $\mathcal{P}[\mathbf{a}]$ 的基 $\{\mathbf{a}^k\}_{k=0}^\infty$ 上应用『定理 3.4.4』, 以此证明『定义 3.6.2』中定义的微分映射是个导子.

习题 3.36. 导出链式法则『(3.6) 式』.

4

算符代数

在『[第三章](#)』中我们介绍了代数(也就是可以通过两个矢量相乘得到第三个矢量的矢量空间)这个概念。在这一章, 我们来探究一个具体的代数: 线性变换的代数。

4.1. 自同态代数 $\text{End}(\mathcal{V})$

在矢量空间 \mathcal{V} 的自同态代数 $\text{End}(\mathcal{V})$ 中, 乘法定义为映射的复合。除了所有代数中都会存在的零元以外, $\text{End}(\mathcal{V})$ 还有个单位元 **1**, 它满足对所有 $|a\rangle \in \mathcal{V}$ 均有 $\mathbf{1}|a\rangle = |a\rangle$ 。

1, 1

名词中英对照

ϕ 导子 (ϕ -derivation), 109

p 线性函数 (p -linear function), 75

p 线性映射 (p -linear map), 75

r 阶导数 (r th derivative), 126

Cantor 集 (Cantor set), 18

Cauchy 列 (Cauchy sequence), 15

Derscates 矢量 (Cartesian vector), 27

Frobenius 定理 (Frobenius theorem), 123

Leibniz 公式 (Leibniz formula), 108

Peirce 分解 (Peirce decomposition), 115

Taylor 公式 (Taylor formula), 126

Wedderburn 主结构定理 (Wedderburn principle structure theorem), 122

一一对应 (one-to-one correspondence), 10

三角不等式 (triangle inequality), 50

不可数的 (uncountable), 17

中心 (center), 89

中心代数 (central algebra), 89

主幂等元 (principle idempotent), 116

乘法 (multiplication), 87

乘法单位元 (multiplicative identity), 28

二元运算 (binary operation), 12

二项式定理 (binomial theorem), 19

交换代数 (commutative algebra), 87

交集 (intersection), 4

代数 (algebra), 87

代数同态 (algebra homomorphism), 96

代数的张量积 (algebra tensor product), 92

代数的直和 (algebra direct sum), 91

代表元 (representative), 6

伴随算符 (adjoint operator), 65

值域 (range), 10

像 (image), 9

像集 (image), 10

元素 (element), 3

全矩阵代数 (total matrix algebra), 105

全集 (universal set), 5

共轭 (conjugate), 94

关系 (relation), 6

内积 (inner product), 41

内积空间 (inner product space), 42

函数 (function), 9

分量 (components), 31

到达域 (target space), 9

加法单位元 (additive identity), 28

半单 (semi-simple), 118

半双线性的 (sesquilinear), 41

单 (纯) 代数 (simple algebra), 103

单位圆 (unit circle), 12

单同态 (monomorphism), 96

单射 (injection), 10

单点集 (singleton), 4

单的 (injective), 10

- 单项式 (monomial), 125
 原像集 (preimage), 10
 双射 (injection), 10
 双线性 (bilinear), 40
 双线性内积 (bilinear inner product), 64
 双边理想 (two-sided ideal), 99
 反导子 (antiderivation), 109
 古典伴随 (classical adjoint), 80
 可数无穷的 (countably infinite), 17
 可约 (reducible), 101
 可逆 (invertible), 60
 可除代数 (division algebra), 93
 右理想 (right ideal), 99
 右逆 (right inverse), 87
 同构 (isomorphic), 60
 同构 (isomorphism), 96
 同构映射 (isomorphism, 也简称为同构), 60
 含幺代数 (unital algebra), 87
 和 (sum), 27
 商代数 (quotient algebra/factor algebra), 104
 四元数代数 (algebra of quaternions), 93
 图象 (graph), 9
 域 (field), 28
 基 (basis, 也称作基底), 31
 复化 (complexification), 67
 复合映射 (composition mapping), 10
 复合运算 (composition), 10
 复矢量空间 (complex vector space), 28
 复结构 (complex structure), 66
 多项式代数 (polynomial algebra), 125
 子代数 (subalgebra), 88
 子集 (subset), 4
 完备度量空间 (complete metric space), 15
 定义域 (domain), 9
 实矢量空间 (real vector space), 28
 实部 (real part), 94
 对偶 (dual), 71
 对偶基 (dual basis), 71
 对偶空间 (dual space), 69
 对合 (involution), 97
 导出代数 (derived algebra), 90
 导子 (derivation), 107
 导子代数 (derivation algebra), 109
 导数算符 (derivative operator), 55
 左理想 (left ideal), 99
 左逆 (left inverse), 87
 带单位元的代数 (algebra with identity), 87
 幂等元 (idempotent), 111
 幂集 (power set), 4
 幂零元 (nilpotent), 111
 幂零指数 (index), 111
 平凡同态 (trivial homomorphism), 103
 平凡理想 (trivial ideal), 103
 平行四边形法则 (parallelogram law), 50
 并集 (union), 4
 幺元 (unit), 87
 幺同态 (unital homomorphism), 97
 序列 (sequence), 14
 序对 (ordered pair), 5
 底空间 (underlying space), 7
 底集 (underlying set), 7
 度量空间 (metric space), 13
 张成子空间 (span), 31
 张量积 (tensor product), 39
 归一化矢量 (normalized vector), 44
 归纳定义 (inductive definition), 20
 微分映射 (differentiation map), 125
 恒等映射 (identity map), 9
 投影映射 (projection), 11
 收敛到 (converge to), 14
 数学归纳法 (mathematical induction), 19
 斜对称的 (skew-symmetric), 75
 斜的 (skew), 65
 既单且满 (bijective), 10

- 映射 (map), 9
 映满的 (onto), 10
 本原幂等元 (primitive idempotent), 116
 权函数 (weight function), 42
 极化恒等式 (polarization identity), 51
 标准内积 (standard inner product), 43
 标准基 (standard basis), 32
 标量项 (scalar term), 125
 核 (kernal), 57
 根 (radical, 亦译作根基), 114
 次数 (degree), 125
 正交 (orthogonal), 44
 满同态 (epimorphism), 96
 满射 (surjection), 10
 满的 (surjective), 10
- 点 (point), 3
 理想 (ideal), 99
 生成元 (generator), 95
 直和 (direct sum), 35
 直积 (direct product), 38
 真子集 (proper subset), 4
 矢量空间 (vector space), 27
 矩阵代数 (matrix algebra), 90
 秩 (rank), 57, 123
 秩-零化度定理 (rank-nullity theorem), 58
 积分算符 (integraion operator), 55
 空集 (empty set), 4
 等价关系 (equivalence relation), 6
 等价类 (equivalence class), 6
 等同 (identification), 7
 等距映射 (isometric map), 54
 纯部 (pure part), 94
 线性变换 (linear transformation), 54
 线性同构 (linear isomorphism), 60
- 线性无关 (linear independent), 30
 线性映射 (linear map), 54
 线性泛函 (linear functional), 69
 线性相关 (linear dependent), 30
 线性等距 (linear isometry), 54
 线性算符 (linear operator), 54
 线性组合 (linear combination), 30
 结合代数 (associative algebra), 87
 结构常数 (structure constants), 92
 绝对值 (absolute value), 94
 维数 (dimension), 31
 维数定理 (dimension theorem), 58
- 自伴的 (self-adjoint), 65
 自同态 (endomorphism), 54
 自同构 (automorphism), 60, 96
 范数 (norm), 50
 行列式 (determinant), 79
 行列式函数 (determinant function), 77
 规范矢量 (normal vector), 44
 赋范线性空间 (normed linear space), 50
- 逆像集 (inverse image), 10
 逆映射 (inverse), 11
 重数 (multiplicity), 127
 链式法则 (chain rule), 126
 陪域 (codomain), 9
 集合 (set), 3
 零化子 (annihilator), 71
 零化度 (nullity), 57
 零化的 (nil), 111
 零矢量 (zero vector, 或者 null vector), 27
 零空间 (null space), 57
 首一多项式 (monic polynomial), 125
 首项 (leading term), 125
 首项系数 (leading coefficient), 125

参考文献

[Axler, 1996] Axler, S. (1996). *Linear Algebra Done Right*. Springer, Berlin.

[Benn and Tucker, 1987] Benn, I. M. and Tucker, R. W. (1987). *An Introduction to Spinors and Geometry with Applications in Physics*. Adam Hilger, Bristol.

[Blyth, 1990] Blyth, T. (1990). *Module Theory*. Oxford University Press, Oxford.

[Friedman, 1982] Friedman, A. (1982). *Foundations of Modern Analysis*. Dover, New York.