

纠缠理论的数学基础

作者：邓嘉驹

时间：September 20, 2026

目录

第 1 章 预备知识 **1**
1.1 基本概念 . . . 1
1.2 代数结构 . . . 7
1.3 度量空间 . . . 8
1.4 实数基础 . . . 13

第一部分　高等线性代数 **21**

第 2 章 向量空间 **22**
2.1 向量空间及其子空间 . . . 22
2.2 生成集与线性无关 . . . 23
2.3 向量空间的维数 . . . 25
2.4 外直和与内直和 . . . 26

第 3 章 线性变换 **29**
3.1 线性变换及其核与像 . . . 29
3.2 同构与秩-零化度定理 . . . 31
3.3 矩阵变换与基变换 . . . 32
3.4 等价性与相似性 . . . 34
3.5 不变子空间与约化对 . . . 35
3.6 投影算子 . . . 36

第 4 章 同构定理 **39**
4.1 商空间 . . . 39
4.2 商空间的泛性质与第一同构定理 . . . 40
4.3 线性泛函与对偶基 . . . 41

第 5 章 实内积空间与复内积空间 **43**
5.1 内积、范数和距离 . . . 43
5.2 等距性与正交性 . . . 45
5.3 正交集与标准正交集 . . . 47
5.4 最佳逼近与投影定理 . . . 48
5.5 Riesz 表示定理 . . . 50

第 6 章 本征值与本征向量 **52**
6.1 本征值与本征向量 . . . 52
6.2 迹与行列式 . . . 53

6.3 可对角化算子 54

第 7 章 正规算子的结构理论 **55**
7.1 线性算子的伴随 55
7.2 正交投影 57
7.3 正规算子 58
7.4 泛函演算 62
7.5 正算子与极分解 64

第 8 章 Hilbert 空间 **66**
8.1 Hilbert 空间 66
8.2 最佳逼近与 Fourier 展开 67
8.3 Hilbert 基与 Hilbert 维数 70
8.4 Riesz 表示定理 71

第二部分 凸几何 **73**

第 9 章 仿射几何 **74**
9.1 基本概念 74
9.2 仿射组合与仿射包 75
9.3 仿射子空间格 76
9.4 仿射无关 78
9.5 仿射变换 79

第 10 章 凸几何 **80**
10.1 代数性质 80
10.2 组合性质 83
10.3 拓扑性质 85
10.4 支撑与分离 86
10.5 极值表示 89

第三部分 测度论 **92**

第 11 章 测度 **93**
11.1 代数与 σ-代数 93
11.2 测度 95
11.3 外测度 97
11.4 Lebesgue 测度的性质 100
11.5 完备性与正则性 103

第 12 章 函数与积分 **106**
12.1 可测函数 106
12.2 积分 110
12.3 极限定理 116

第 13 章 Haar 测度 **118**

13.1 局部紧空间 . . . 118
13.2 拓扑群 . . . 119
13.3 Haar 测度的存在唯一性 . . . 121
13.4 Haar 测度的性质 . . . 122

第四部分 C^*-代数 **124**

第 14 章 Banach 空间 **125**
14.1 赋范空间 . . . 125
14.2 常见 Banach 空间 . . . 126
14.3 有界线性算子的基本性质 . . . 129
14.4 范数的等价性 . . . 131
14.5 有限维赋范空间 . . . 132
14.6 无穷维赋范空间 . . . 133
14.7 商空间 . . . 134

A 拓扑基础 **135**
A.1 拓扑的定义 . . . 135
A.2 收敛性与连续性 . . . 136
A.3 同胚 . . . 138
A.4 拓扑的基与子基 . . . 139
A.5 乘积拓扑 . . . 141
A.6 诱导拓扑 . . . 142

附录 B 奇异值和 Moore-Penrose 逆 **143**
B.1 奇异值 . . . 143
B.2 Moore-Penrose 逆 . . . 144
B.3 最小二乘逼近 . . . 144

附录 C Schatten 范数 **145**
C.1 从向量范数到 Schatten 范数 . . . 145
C.2 范数的基本性质 . . . 145
C.3 特殊 Schatten 范数 . . . 147

参考文献 **148**

第 1 章 预备知识

目录

1.1 基本概念 1
1.2 代数结构 7
1.3 度量空间 8
1.4 实数基础 13

1.1 基本概念

1.1.1 集合与选择公理

集合及其衍生的（真）子集、空集、补集等概念我们已经熟知，其幂集（*power set*）定义为集合 X 的所有子集构成的集合，通常记为 $\mathcal{P}(X)$ 或 2^X。对于并集和交集，需要进行一些简要说明：为了避免“集合的集合”这种说法带来歧义，我们常换用“集族”这一术语，其中“族”的英文对应“collection”或“family”。设 $\mathcal{S}$ 是一个集族，那么 $\mathcal{S}$ 的并集定义为属于 $\mathcal{S}$ 中至少一个集合的点构成的集合，记作 $\cup_{S\in\mathcal{S}}S$；$\mathcal{S}$ 的交集定义为属于 $\mathcal{S}$ 中每个集合的点构成的集合，记作 $\cap_{S\in\mathcal{S}}S$。若 $\mathcal{S}$ 中任意两个不同的集合交集为空，则称 $\mathcal{S}$ 不相交（*disjoint*）。需要注意的是，这里不关心 $\mathcal{S}$ 是否有限，甚至不关心其是否可数，常见的自然数集和实数集对应的幂集都是集族。此外，集族 $\mathcal{S}$ 还满足如下著名的 De Morgan 恒等式：

$$X\backslash(\cup_{S\in\mathcal{S}}S)=\bigcap_{S\in\mathcal{S}}(X\backslash S),\quad X\backslash(\cap_{S\in\mathcal{S}}S)=\bigcup_{S\in\mathcal{S}}(X\backslash S), \tag{1.1.1}$$

其中 $X\backslash S$ 表示集合 X 关于其子集 S 的补集。在集合 X 在上下文明确时，我们也会将补集简记为 S^c。除去第10章通过 Minkowski 和定义的差集，其余章节中集合 $A-B$ 均表示集合 $\{x\in X|x\in A,x\notin B\}=A\cap B^c$。

有了集合的定义，我们可定义一种特殊的“集合”。

定义 1.1 (多重集)

设 S 为非空集合，那么以 S 为基础集（underlying set）的多重集（multiset）M 是由有序对组成的集合：

$$M=\{(s_i,n_i)|s_i\in S,\ n_i\in\mathbb{Z}^+,\ s_i\neq s_j \text{ for } i\neq j\}, \tag{1.1.2}$$

其中 $\mathbb{Z}^+$ 为正整数集，自然数 n_i 称为 M 中元素的重数（multiplicity）。若多重集的基础集有限，我们称多重集也是有限的，其大小为其所有元素重数之和。

♣

注 我们要在第1.4节才公理化定义实数，进而引入正整数的概念，这里体现的是我们多年来应用正整数的熟悉性。另外，多重集并集定义与集合并集定义的微妙差异在于，前者要求不同集合中相同元素在并集后得到该元素的重数和。例如对于多重集 $A=\{1,1,2,2,2\}$ 和 $B=\{1,1,1,2\}$，其并集为 $A\cup B=\{1,1,1,1,1,2,2,2,2\}$。

定义 1.2 (函数)

设 $S:f\to T$ 为从集合 S 到集合 T 的函数（function）。

1. f 的定义域（domain）为 S；
2. f 的像（image）或值域（range）是集合 $\mathrm{im}(f)=\{f(s)|s\in S\}$；
3. 若 $x\neq y\ \Rightarrow\ f(x)\neq f(y)$，则称 f 是单射（injective）或嵌入（embedding）；
4. 若 $\mathrm{im}(f)=T$，则称 f 是到 T 上的满射（surjective）；
5. 若 f 既是单射又是满射，则称 f 是双射（bijective）；
6. 假定 $0\in T$，则 f 的支撑集（support）为 $\mathrm{supp}(f)=\{s\in S|f(s)\neq 0\}$。

♣

注

1. 对于集合 T 的任意子集 B，有时记号 $f^{-1}(B)=\{s\in S|f(s)\in B\}$ 称为 B 的原像（*inverse image*）而非逆映射，其具有保持交集、并集与取补集运算的良好性质：
$$f^{-1}(\cup_i B_i)=\cup_i f^{-1}(B_i),\quad f^{-1}(\cap_i B_i)=\cap_i f^{-1}(B_i),\quad f^{-1}(B^c)=(f^{-1}(B))^c. \tag{1.1.3}$$
上述等式对于任意指标集中的指标 i 成立，可通过集合相互包含的语言证明；
2. 对于函数 $f:S\to T$ 以及 S 的任意子集 A，f 在 A 上的限制（*restriction*）定义为映射 $f|_A:A\to T$ 其将 A 中的任意元素 a 映射到 $f(a)$。显然，单射的限制仍为单射。

本文关注的理论都默认建立在带选择公理的 Zermelo-Fraenkel 集合论上；事实上，说选择公理是现代数学的主流语言和公理基础都不为过。因此有必要简要介绍选择公理的内容，其等价表述放在第1.1.3小节介绍。

定义 1.3 (选择函数)

设 $\mathcal{S}$ 为非空集合构成的非空集族。$\mathcal{S}$ 上的选择函数（choice function）f 是从 $\mathcal{S}$ 到 $\cup_{S\in\mathcal{S}}S$ 的函数，其对于 $\mathcal{S}$ 中的每个集合 S，$f(S)$ 是 S 的一个元素。

♣

公理 1.1 (选择公理)

设 $\mathcal{S}$ 为非空集合构成的非空集族，则存在 $\mathcal{F}$ 上的选择函数。

♡

1.1.2 等价关系

等价关系是一类重要的关系，同构、同胚等关系事实上都是等价关系。

定义 1.4 (等价关系)

设 S 为非空集合，若 S 上的二元关系 $\sim$ 满足

1. 自反性（reflexivity）：对于任意 $a\in S$，$a\sim a$；
2. 对称性（symmetry）：对于任意 $a,b\in S$，$a\sim b\Rightarrow b\sim a$；
3. 传递性（transitivity）：对于任意 $a,b,c\in S$，$a\sim b,\ b\sim c\Rightarrow a\sim c$；

则称 $\sim$ 为 S 上的等价关系（equivalence relation）。

♣

注 我们熟知的（实数、集合等）相等就是一种等价关系；线性代数中的矩阵相似也是一种等价关系

定义 1.5 (等价类)

设 $\sim$ 为 S 上的等价关系。对于 S 上的元素 a，所有与 a 等价的元素集合记为 $[a]=\{b\in S|b\sim a\}$，称其为 a 的等价类（equivalence class）。

♣

注 考虑一种等价关系：两个实数相减为有理数，那么有理数集自身就构成一个等价类。

命题 1.1

设 $\sim$ 为 S 上的等价关系，则

1. $b\in[a]\Leftrightarrow a\in[b]\Leftrightarrow[a]=[b]$；
2. 对于任意 $a,b\in S$，要么 $[a]=[b]$，要么 $[a]\cap[b]=\varnothing$。

♠

证明

1. 若 $b\in[a]$，则 $b\sim a$，即 $a\in[b]$，反之类似。现在仅需证明 $b\in[a]\Leftrightarrow[a]=[b]$，其中必要性是显然的；对于充分性，由于 $b\in[a]$，那么对于任意 $c\in[b]$，利用等价关系的传递性，$c\sim b\sim a$，因此 $[b]\subseteq[a]$；类似地 $[a]\subseteq[b]$，因此 $[a]=[b]$；
2. $[a]\cap[b]=\varnothing$ 的情形无需讨论；若 $[a]\cap[b]\neq\varnothing$，设存在 $c\in[a]\cap[b]$，则 $c\in[a]$ 且 $c\in[b]$，进而 $a\sim c\sim b$。根据上一条结论，此时 $[a]=[b]$。

定义 1.6 (集合的划分)

若非空子集 S 的非空子集族 $\{A_1, \cdots, A_n\}$ 满足

1. 对于任意 $i \neq j$，$A_i \cap A_j = \varnothing$；
2. $S = A_1 \cup \cdots \cup A_n$；

则称其为 S 的一种划分（partition），其中每个 A_i 称为划分的块（block）。

命题 1.2

1. 设 $\sim$ 为 S 上的等价类，那么关于 $\sim$ 的不同等价类集合构成了 S 的一个划分的不同块；
2. 反之，若 $\mathcal{P}$ 为 S 上的一个划分，则如下定义的二元关系构成 S 上的一个等价关系：

$$a \sim b \text{ if } a \text{ and } b \text{ lie in the same block of } \mathcal{P},$$

其等价类恰好为 $\mathcal{P}$ 的块。

证明

1. 根据命题 1.1，对于任意 $a, b \in S$，若两者不等价，则 $[a] \cap [b] = \varnothing$，满足划分的第一条要求；另外，显然 $\cup_{a \in S}[a] \subseteq S \subseteq \cup_{a \in S}[a]$，即 $S = \cup_{a \in S}[a]$，满足划分的第二条要求。因此不同等价类构成 S 的一个划分；
2. 请读者自行验证该关系是等价关系，从而结论得证。

关于等价关系最重要的一个问题是找到一种有效的方法来确定两个元素何时等价。不幸的是，在大多数情况中，这种定义并不提供一种有效的等价性检验方法，由此引出了如下概念。

定义 1.7 (不变量与不变量系统)

设 $\sim$ 为 S 上的等价关系。对于任一集合 T，若函数 $f : S \to T$ 在每个等价类上均为常数，即

$$a \sim b \implies f(a) = f(b), \tag{1.1.4}$$

则称其为不变量（invariant）；若其在每个等价类上均为常数且常数互不相同，即

$$a \sim b \iff f(a) = f(b), \tag{1.1.5}$$

则称其为完全不变量（complete invariant）。若不变量族 $\{f_1, \cdots, f_n\}$ 满足

$$a \sim b \iff f_i(a) = f_i(b) \text{ for all } i = 1, \cdots, n, \tag{1.1.6}$$

则称其为完整不变量系统（complete system of invariants）。

定义 1.8 (规范形)

设 $\sim$ 为 S 上的等价类，C 为 S 的某个子集。若对于每个 $s \in S$，存在唯一的 $c \in C$ 使得 $c \sim s$，则称 C 为关于 $\sim$ 的规范形集（set of canonical form）。换言之，$\sim$ 下的每个等价类唯一包含 C 中的一个元素。

注 线性代数中的 Jordan 标准型就是一种规范形。

1.1.3 偏序集与 Zorn 引理

定义 1.9 (偏序集)

对于非空集合 P 与二元关系 $\leq$，$(P, \leq)$ 构成一个偏序集（partially ordered set），满足

1. 自反性：对于任意 $a \in P$，$a \leq a$；
2. 反对称性：对于任意 $a, b \in P$，$a \leq b$, $b \leq a \Rightarrow a = b$；
3. 传递性：对于任意 $a, b, c \in P$，$a \leq b$, $b \leq c \Rightarrow a \leq c$。

这里 $\leq$ 称为偏序（partial order），读作“小于等于”。偏序集通常也简称为 poset。

注 当偏序在上下文明确时，习惯上采用“设 P 为偏序集”的说法。

定义 1.10

设 P 为偏序集。

1. 若 P 中不存在比元素 m 大的元素，即对于任意 $p \in P$，$m \leq p \Rightarrow m = p$，则称 m 为极大元（maximal element）；
2. 若 P 中不存在比元素 n 小的元素，即对于任意 $p \in P$，$p \leq n \Rightarrow p = n$，则称 n 为极小元（minimal element）；
3. 设 $a, b \in P$，则如果元素 $u \in P$ 满足 $a \leq u$ 且 $b \leq u$，则称 u 是关于 a 和 b 的上界（upper bound）。若存在关于 a 和 b 的唯一最小上界，则称其为 a 和 b 的最小上界（least upper bound）或上确界（supremum），记作 $\text{lub}\{a, b\}$；
4. 设 $a, b \in P$，则如果元素 $\ell \in P$ 满足 $\ell \leq a$ 且 $\ell \leq b$，则称 ℓ 是关于 a 和 b 的下界（lower bound）。若存在关于 a 和 b 的唯一最大下界，则称其为 a 和 b 的最大下界（greatest lower bound）或下确界（infimum），记作 $\text{glb}\{a, b\}$；

♣

注 需要注意的是，这里术语用的是“极大（小）”而非“最大（小）”，是因为并非偏序集中所有元素都能比大小。我们将在介绍完“全序集”的概念后给出几个说明性的例子。

定义 1.11 (全序集)

若偏序集内每对元素都是可比较的，则称其为全序集（totally ordered set）或线性序集（linearly ordered set）。偏序集 P 的任一全序子集称为 P 的一条链（chain）。

♣

例 1.1

1. 实数集 $\mathbb{R}$ 在通常的二元关系 $\leq$ 下构成一个偏序集，同时也是全序集，并且不含最大元素；
2. 自然数集 $\mathbb{N} = \{0, 1, \cdots\}$ 在整除意义下构成一个偏序集，习惯上用 $n|m$ 表示 n 整除 m；$\mathbb{N}$ 中所有 2 的幂次构成的子集是 $\mathbb{N}$ 的一条链。集合 $P = \{2, 4, 8, 3, 9, 27\}$ 是整除意义下的偏序集，其极大元为 8 和 27；而偏序集 $Q = \{2, 3, 5, 7, 11\}$ 内每个元素均为极大元和极小元；
3. 任意非空集合的幂集在集合包含的意义下构成一个偏序集。

有了偏序的概念后，我们就可以介绍选择公理 1.1的等价表述：Zorn 引理。

定理 1.1 (Zorn 引理)

若偏序集 P 的每条链都有上界，则 P 有极大元。

♡

我们在此不证明 Zorn 引理，将其视为与选择公理地位相同的一条公理，只能被接受或拒绝。

1.1.4 基数

若存在集合 S 与 T 之间的双射，则称 S 与 T 具有相同的基数（*cardinality*），记作 $|S| = |T|$。一个稍微有些反直觉的结果是自然数集、整数集与有理数集具有相同的级数，或称三者等势（*equipotent*），即 $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$。这一结论我们将在第1.4节定义了有理数后进行证明。

若 S 与 T 的某个子集构成一一映射，将其记为 $|S| \leq |T|$；若 S 与 T 的某个真子集但不与整个 T 构成一一映射，则记为 $|S| < |T|$。需要强调的是，第二个条件是必要的，例如 $\mathbb{N}$ 能与 $\mathbb{Z}$ 的真子集（例如 $\mathbb{N}$ 自身）构成一一映射，但同时也与 $\mathbb{Z}$ 构成一一映射，因此 $|\mathbb{N}| = |\mathbb{Z}|$。

定义 1.12

1. 若某个集合能与形如 $\mathbb{Z}_n = \{0, 1, \cdots, n-1\}$ 的集合构成一一映射，则称其有限（finite），其中 n 是某个非负整数。非有限的集合称为无限的（infinite），有限集的基数就是集合的元素数；

2. 自然数集 $\mathbb{N}$ 的基数为 $\aleph_0$，读作“阿列夫零（aleph nought）”，其中 $\aleph$ 是希伯来（Hebrew）字母表的第一个字母，因此 $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = \aleph_0$；
3. 任何具有基数 $\aleph_0$ 的集合称为可数无穷集（countably infinite set），任何有限或可数无穷的集合称为可数集（countable set）。非可数的无穷集称为不可数的（uncountable）。

♣

注 要证明 $|\mathbb{R}| > |\mathbb{N}|$，严格的证明我们放到第1.4节，这里我们先介绍 1891 年由 Cantor 提出的一种更直观的方法：对角论证法。

1. 假设区间 $[0,1]$ 是可数无穷的，并且已知区间内的每个数字都能用小数表达；
2. 将这些数字 $r_1, \cdots, r_n$ 排成序列（不要求按通常的序排列），并如下构造一个区间 $[0,1]$ 上的实数 x：若 r_k 的小数点后第 k 位不是 1，则取 x 的第 k 位为 1；反之取 x 的第 k 位为 2。显然 x 是 $[0,1]$ 内的小数；
3. 由于 x 的特殊定义，x 与序列中的每个小数 x_n 必然在第 n 位不同，即 x 不在序列 $(r_1, \cdots, r_n)$ 内。但这与“$(r_1, r_2, \cdots)$ 包含所有区间 $[0,1]$ 内的实数”的假设矛盾，因此 $[0,1]$ 不是可数无穷的，进而 $\mathbb{R}$ 也不是。

若 S 和 T 是有限集，则偏序的定义要求 $|S| \leq |T|, |T| \leq |S| \Rightarrow |S| = |T|$；另外，$S$ 的幂集 $\mathcal{P}(S)$ 的基数 $2^{|S|}$ 显然大于 $|S|$。对于无限集而言，上述结果依然成立。

定理 1.2

1. Schrödinger-Berstein 定理：对于任意集合 S 与 T，$|S| \leq |T|, |T| \leq |S| \Rightarrow |S| = |T|$；
2. Cantor 定理：对于任意集合 S，$|S| < |\mathcal{P}(S)|$；
3. 若 $\mathcal{P}_0(S)$ 表示无限集 S 的所有有限子集构成的族，则 $|S| = |\mathcal{P}_0(S)|$。

♡

现在我们可以陈列一些基数算术（*cardinal arithmetic*）的基本结论，首先介绍笛卡尔积（*Cartesian product*）的概念：设 S 和 T 为集合，那么其笛卡尔积 $S \times T$ 定义为由所有 S 和 T 中元素的有序对 (s,t) 构成的集合，即 $S \times T = \{(s,t) | s \in S, t \in T\}$。所有从 T 到 S 的函数构成的集合记为 S^T。

定义 1.13 (基数算术)

设 κ 和 λ 为基数，S 和 T 为任意满足 $|S| = \kappa$ 和 $|T| = \lambda$ 的不相交集合。

1. 和 $\kappa + \lambda$ 是 $|S \cup T|$ 的基数；
2. 积 $\kappa\lambda$ 是 $S \times T$ 的基数；
3. 幂 κ^λ 是 S^T 的基数。

♣

注 通过基数算术的定义，记号 S^T 的含义也更加清晰：对于有限集 S 和 T，T 中的每个元素独立地有 $|S| = \kappa$ 种映射的可能，因此 $\left|S^\lambda\right| = \kappa^\lambda$。特别地，当 S 只包含两个元素（不妨记为 0 和 1）时，T 中的元素映射到 0 和 1 可分别理解为该元素是否出现。将出现的元素置于一个新的集合，此即 T 的子集。因此所有可能的映射恰好与 T 的所有子集形成一一对应，因此其幂集可记为 2^T。根据 Cantor 定理 1.2，$|T| < \left|2^T\right| = 2^{|T|}$。

通过上述定义，容易验证基数算术的以下基本性质。

命题 1.3

设 κ, λ, μ 为基数，则其具有下列性质。

1. 结合律（associativity）：$\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$, $\kappa(\lambda\mu) = (\kappa\lambda)\mu$；
2. 交换律（commutativity）：$\kappa + \lambda = \lambda + \kappa$, $\kappa\lambda = \lambda\kappa$；
3. 分配律（distributivity）：$\kappa(\lambda + \mu) = \kappa\lambda + \kappa\mu$；
4. 指数性质：
   (1) $\kappa^{\lambda+\mu} = \kappa^\lambda \kappa^\mu$；
   (2) $(\kappa^\lambda)^\mu = \kappa^{\lambda\mu}$；
   (3) $(\kappa\lambda)^\mu = \kappa^\mu \lambda^\mu$。

♠

尽管上述性质反映出基数算术与通常算术的相似之处，下述命题又反映出两者间的细微差别。

命题 1.4

设 κ 和 λ 为基数，其中至少一个无穷，则 $\kappa+\lambda=\kappa\lambda=\max\{\kappa,\lambda\}$。

事实上，$\aleph_0$ 是最小的无穷势，即如果基数 $\kappa<\aleph_0$，则 κ 为自然数。另外，可以证明[1]实数集 $\mathbb{R}$ 与幂集 $\mathcal{P}(\mathbb{N})$ 间存在一一映射，因此 $|\mathbb{R}|=2^{\aleph_0}$。有时实轴上的点称为连续统（*continuum*），因此 $2^{\aleph_0}$ 称为连续统势（*power of the continuum*），记作 c。

命题 1.5

1. 在基数 $\aleph_0$ 上作用可数次加法或有限次乘法，并不产生任何高于 $\aleph_0$ 的基数。特别地，对于任一非零自然数 n，我们有
$$\aleph_0\cdot\aleph_0=\aleph_0 \quad \text{and} \quad \aleph_0^n=\aleph_0. \tag{1.1.7}$$
2. 在基数 $2^{\aleph_0}$ 上作用可数次加法和乘法，并不产生高于 $2^{\aleph_0}$ 的基数，即
$$\aleph_0\cdot 2^{\aleph_0}=2^{\aleph_0} \quad \text{and} \quad (2_0^{\aleph})^{\aleph_0}=2^{\aleph_0}. \tag{1.1.8}$$

证明

1. 根据命题 1.4，$\aleph_0\cdot\aleph_0=\max\{\aleph_0,\aleph_0\}=\aleph_0$，另一结论类似；
2. 根据命题 1.4，第一条结论显然；再利用命题 1.3，$(2^{\aleph_0})^{\aleph_0}=2^{\aleph_0\cdot\aleph_0}=2^{\aleph_0}$。

利用此命题，我们可以得到一些特殊的关系。例如根据 $2<\aleph_0<2^{\aleph_0}$ 可得 $2^{\aleph_0}\le\aleph_0^{\aleph_0}\le(2^{\aleph_0})^{\aleph_0}=2^{\aleph_0}$，因此 $\aleph_0^{\aleph_0}=2^{\aleph_0}$。不过一般而言，估计 κ^λ 是极其困难的。

命题 1.6

设 $\{A_k|k\in K\}$ 是以 K 为指标集的集族，其中 $|K|=\kappa$。若对于任意 $k\in K$，$|A_k|\le\lambda$，则有 $|\cup_{k\in K}A_k|\le\lambda\kappa$。

最后，我们描述一些著名集合的基数。

命题 1.7

1. 下列集合具有基数 $\aleph_0$：
   (1) 有理数集 $\mathbb{Q}$ 和整数集 $\mathbb{Z}$；
   (2) $\mathbb{N}$ 的所有有限子集构成的集合；
   (3) 可数个可数无穷集合的并集；
   (4) 所有有序整数 n 元组构成的集合 $\mathbb{Z}^n$；
2. 下列集合具有基数 $2^{\aleph_0}$：
   (1) $\mathbb{R}^n$ 中的所有点构成的集合；
   (2) 所有自然数的无穷序列构成的集合 $\{(a_1,a_2,\cdots)|a_i\in\mathbb{N}\}$；
   (3) 所有实数的无穷序列构成的集合；
   (4) $\mathbb{R}$ 的所有有限子集构成的集合；
   (5) 全体无理数构成的集合。

证明

1. (1) 该结果显然；
   (2) 源于定理 1.2；
   (3) 不妨设集族 $\{A_k|k\in K\}$，其中 $|K|\le\aleph_0$ 且 $|A_k|=\aleph_0$。根据命题 1.4和1.6，$|\cup_{k\in K}A_k|\le|K|\cdot\aleph_0=\max\{|K|,\aleph_0\}=\aleph_0$；另一方面，对于任意 $k\in K$，$A_k\subseteq\cup_{k\in K}A_k$，因此 $|\cup_{k\in K}A_k|\ge|A_k|=\aleph_0$。因

[1]一种简单的方法是利用 Cantor 集中的元素 $x=\sum_{n=1}^{\infty}2a_n/3$（其中 $a_n\in\{0,1\}$）将区间 $[0,1]$ 与自然数集的幂集 $2^{\mathbb{N}}$ 关联起来，不过 Cantor 集的详细介绍将放到“测度论”这一部分，届时读者可以回头重新审视这一结论

此 $|\cup_{k\in K}A_k|=\aleph_0$；

(4) 源于命题 1.5；

2. (1) 源于命题 1.5；

(2) 源于命题 1.5后的讨论，有 $\aleph_0^{\aleph_0}=2^{\aleph_0}$；

(3) 这里的“无穷序列”实际上是一种枚举，而枚举对应的指标集恰好为自然数集。枚举的定义将放到第1.4节介绍。根据命题 1.5，$(2^{\aleph_0})^{\aleph_0}=2^{\aleph_0}$；

(4) 源于定理 1.2；

(5) 由于有理数集与无理数集不相交，设后者的基数为 κ，则根据命题 1.4，

$$2^{\aleph_0}=|\mathbb{R}|=|\mathbb{Q}\cup(\mathbb{R}\backslash\mathbb{Q})|=\aleph_0+\kappa=\max\{\aleph_0,\kappa\},\tag{1.1.9}$$

从而 $|\mathbb{R}\backslash\mathbb{Q}|=\kappa=2^{\aleph_0}$。

1.2 代数结构

定义 1.14 (群)

非空集合 G 在二元运算 $*$ 下可构成一个群（group），满足

1. 封闭性：对于任意 $a,b\in G$，$a*b\in G$；
2. 结合律：对于任意 $a,b,c\in G$，$(a*b)*c=a*(b*c)$；
3. 单位元（identity）：存在 $e\in G$，使得对于任意 $a\in G$，$a*e=e*a=a$；
4. 逆元（inverse）：对于任意 $a\in G$，存在 $a^{-1}\in G$，使得 $a*a^{-1}=a^{-1}*a=e$。

♣

定义 1.15 (Abelian 群)

若对于群 G 中任意元素 a,b 均有 $a*b=b*a$，则称 G 是阿贝尔的（Abelian）或交换的（commutative）。对于 Abelian 群，习惯上将运算 $*$ 记为 $+$，并将单位元称为零元（zero element），将逆元称为负元（negative），记为 $-a$。

♣

例 1.2

1. 从 S 到 S 的所有双射在复合运算（*composition*）下构成一个群（通常非 Abelian）；
2. $\mathbb{R}$ 上所有 $m\times n$ 矩阵构成的集合，在矩阵加法下构成一个 Abelian 群，其零元为零矩阵 $0_{m\times n}$。不过 $\mathbb{R}$ 上所有 $n\times n$ 矩阵构成的集合，在矩阵乘法下并不构成群，因为并非所有矩阵都具有乘法逆；当然如果局限在同规模可逆矩阵构成的集合，其在矩阵乘法下仍然构成（非 Abelian）群。

定义 1.16 (环)

非空集合 R 在二元运算加法（记为 $+$）和乘法（用并置表示）下构成一个环（ring），满足

1. R 在加法下构成一个 Abelian 群；
2. 乘法结合律：对于任意 $a,b,c\in R$，$(ab)c=a(bc)$；
3. 左右分配律：对于任意 $a,b,c\in R$，$(a+b)c=ac+bc$ 且 $c(a+b)=ca+cb$。

若对于任意 $a,b\in R$ 有 $ab=ba$，则称环是交换的；若环 R 具有乘法单位元，则称其为含幺环（ring with identity），通常该单位元记为 1。

♣

例 1.3

1. 集合 $\mathbb{Z}_n=\{0,1,\cdots,n-1\}$ 在模 n 的加法和乘法下个构成交换环，即

$$a\oplus b=(a+b)\mod n\quad\text{and}\quad a\odot b=ab\mod n,\tag{1.2.1}$$

其中元素 $1\in\mathbb{Z}_n$ 为单位元；

2. 全体偶数构成的集合 E 在 $\mathbb{Z}$ 的通常运算下构成交换环，但没有单位元；

3. $\mathbb{R}$ 上所有 $n\times n$ 矩阵在矩阵加法和乘法下构成非交换环，其单位元 I_n 为单位矩阵；
4. 所有以 x 为自变量的实系数多项式，在通常的多项式加法和乘法运算下构成交换环。

定义 1.17 (域)

至少含有两个元素的集合 F 在二元运算加法（记为 $+$）和乘法（用并置表示）下构成一个域（field），满足

1. F 在加法下构成一个 Abelian 群；
2. 所有非零元素构成的集合 F^* 在乘法下构成一个 Abelian 群；
3. 左右分配律：对于任意 $a,b,c\in F$，$(a+b)c=ac+bc$ 且 $c(a+b)=ca+cb$。

♣

注 我们要求 F 至少含有两个元素，是为了避免出现“$0=1$”的病态情况。

1. 全体有理数、实数和复数到的集合 $\mathbb{Q},\mathbb{R}$ 和 $\mathbb{C}$ 在通常的加法和乘法下各自构成域；
2. 环 $\mathbb{Z}_n$ 构成域当且仅当 n 为素数，以后将该素数记为 p。

若每个定义在域 F 上的非常数多项式均在 F 上有根，则称 F 代数封闭（*algebraically closed*）。这等价于每个非常值多项式可在 F 上分解为一系列线性因子，例如复数域 $\mathbb{C}$ 代数封闭，但实数域 $\mathbb{R}$ 则不是。我们不加证明地指出，每个域 F 都包含在一个代数闭域（*algebraically closed field*）$\bar{F}$ 内，称 $\bar{F}$ 为 F 的代数闭包（*algebraic closure*）。

1.3 度量空间

这一节我们引入度量空间中的拓扑性质，尽可能避免一般的拓扑空间（参见附录A）的讨论。

1.3.1 度量空间

定义 1.18 (度量空间)

度量空间（metric space）是一组有序对 (X,d)，其中 X 为非空集合，实值函数 $d:X\times X\to\mathbb{R}$ 称为 X 上的度量（metric），满足

1. 正定性（positive definiteness）：对于任意 $x,y\in X$，$d(x,y)\geq 0$，等号成立当且仅当 $x=y$；
2. 对称性（symmetry）：对于任意 $x,y\in X$，$d(x,y)=d(y,x)$；
3. 三角不等式（triangle inequality）：对于任意 $x,y,z\in X$，$d(x,y)\leq d(x,z)+d(z,y)$。

♣

注

1. “正定性”中的“定”性是重要的，例如对于区间 $[0,1]$ 上全体可微函数 $f:[0,1]\to\mathbb{R}$ 构成的集合，积分
$$\int_0^1 |f'(x)-g'(x)|\,\mathrm{d}x \tag{1.3.1}$$
并不构成度量，因为积分为零时 $f(x)$ 和 $g(x)$ 可能分别为不同的实常数；
2. 若 (X,d) 为度量空间，则每个子集 $Y\subseteq X$ 在继承自 (X,d) 的度量下也构成一个度量空间，此时称 Y 上的度量为 X 上的度量诱导的度量。

例 1.4 度量空间的一个最简单例子是 $\mathbb{R}$，其度量恰好为绝对值：$d(x,y)=|x-y|$。$\mathbb{R}^n$ 中的标准度量为 *Euclidean* 度量（*Euclidean metric*），定义为
$$d(\boldsymbol{x},\boldsymbol{y})=\|\boldsymbol{x}-\boldsymbol{y}\|_2=\left(\sum_{i=1}^n |x_i-y_i|^2\right)^{1/2}\quad\text{for all }\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n. \tag{1.3.2}$$

定义 1.19 (开球和闭球)

设 (X,d) 为度量空间，$x\in M$ 且 $r>0$ 为正实数。

1. 以 x 为中心、r 为半径的开球（open ball）为
$$B(x,r)=\{y\in X|d(x,y)<r\}. \tag{1.3.3}$$

2. 以 x 为中心、r 为半径的闭球（closed ball）为

$$\bar{B}(x,r)=\{y\in X|d(x,y)\leq r\}. \tag{1.3.4}$$

利用开球我们可以定义度量空间上的开集，从而定义拓扑。

定义 1.20 (开集与闭集)

设 (X,d) 为度量空间，对于子集 $U\subseteq X$ 内每个点 x，均存在 $r>0$ 使得 $B(x,r)\subseteq U$，则称 U 为开集（open set）。若子集 U 的补集 $X\backslash U$ 为开集，则称 U 为闭集（closed set）。

注 根据上述定义，开球 $B(x,r)$ 显然是开集：对于任意 $y\in B(x,r)$，总存在 $\epsilon=(r-d(x,y))/2>0$ 使得 $B(y,\epsilon)$ 中的每个点 z 均满足 $d(x,z)\leq d(x,y)+d(y,z)<d(x,y)+\epsilon=(r+d(x,y))/2<r$，因此 $B(y,\epsilon)\subseteq B(x,r)$。

根据上述定义，容易验证度量空间中的开集具有如下基本性质。

定理 1.3

度量空间 (X,d) 中的开集具有如下性质：

1. $\varnothing,X$ 为开集；
2. 开集的有限交集为开集；
3. 开集的任意并集为开集。

证明 性质 1 是显然的；对于性质 2，考虑开集 U_1 和 U_2，对于任意 $x\in U_1\cap U_2$，存在开球 $B(x,r_1)$ 和 $B(x,r_2)$，使得 $B(x,r_1)\subseteq U_1$ 且 $B(x,r_2)\subseteq U_2$。取 $r=\min\{r_1,r_2\}$，则 $B(x,r)\subseteq U_1\cap U_2$，因此 $U_1\cap U_2$ 为开集；对于性质 3，考虑开集族 $\{U_k|k\in K\}$，对于任意 $x\in\cup_{k\in K}U_k$，存在指标 k 使得 $x\in U_k$，因此存在开球 $B(x,r)$ 使得 $B(x,r)\subseteq U_k\subseteq\cup_{k\in K}$，因此 $\cup_{k\in K}U_k$ 为开集。

对比拓扑的定义 A.1，容易看出 X 的所有开子集构成了 X 上的一个拓扑。

定义 1.21 (邻域)

设 (X,d) 为度量空间。若对于 $x\in X$ 和集合 U，存在开集 $V\subseteq U$ 使得 $x\in V$，则称 U 是 x 的一个邻域。

有了邻域的概念，可以定义一些与度量空间内的给定子集有关的集合。

定义 1.22 (内部、闭包与边界)

设 U 为度量空间 (X,d) 的一个子集。若 U 是 x 的邻域，则称点 $x\in U$ 为 U 的内点（interior point），并称

1. $\mathrm{Int}(U)=\{x\in U|x\text{ is an interior of }U\}$ 为 U 的内部（interior）；
2. $\bar{U}=\{x\in X|U\cap V\neq\varnothing\text{ for every neighbourhood }V\text{ of }x\}$ 为 U 的闭包（closure）；
3. $\partial U=\bar{U}\backslash\mathrm{Int}(U)$ 为 U 的边界（boundary）。

注 容易证明，U 为开集当且仅当 $U=\mathrm{Int}(U)$，而 U 为闭集当且仅当 $\bar{U}=U$。

有时考察（有限）多个度量空间的乘积是方便的，同时也有可能在这样的积空间上定义度量。

命题 1.8 (积空间与积度量)

设 $(X_i,d_i)\ (i=1,\cdots,n)$ 为度量空间，则 $X=X_1\times\cdots\times X_n$ 是具有度量

$$d(x,y)=\sum_{i=1}^{n}d_i(x_i,y_i) \tag{1.3.5}$$

的度量空间，其中 $x=(x_1,\cdots,x_n)\in X$ 且 $y=(y_1,\cdots,y_n)\in Y$。由此引入的度量称为积度量（product metric），相应的度量空间 (X,d) 称为积空间（product space）。

1.3.2 极限

定义 1.23 (极限)

设 $(x_n)_{n\in\mathbb{N}}$ 是度量空间 (X,d) 的一个序列。若对于 x 的每个邻域 U，存在 $n_0\in\mathbb{N}$ 使得 $x_n\in U$ 对所有 $n\geq n_0$ 成立，则称 x 是 (x_n) 的极限（limit），记为

$$x=\lim_{n\to\infty}x_n \quad \text{or} \quad x_n\to x \text{ and } n\to\infty. \tag{1.3.6}$$

若序列有极限，则称其收敛（convergent），否则称其发散（divergent）。

对于度量空间（而非一般拓扑空间），其收敛性描述存在更直观的等价表述。

命题 1.9

设 (x_n) 为度量空间 (X,d) 的一个序列，且 $x\in X$，则以下表述等价：

1. $\lim_n x_n=x$；
2. 对于每个 $\epsilon>0$，存在 $n_0\in\mathbb{N}$ 使得 $d(x_n,x)<\epsilon$ 对所有 $n\geq n_0$ 成立。

证明 若表述 1 成立，则根据定义，对于任意 $\epsilon>0$，$B(x,\epsilon)$ 是 x 的一个邻域。因此存在 $n_0\in\mathbb{N}$，使得 $x_n\in B(x,\epsilon)$ 对所有 $n\geq n_0$ 成立，即 $d(x_n,x)<\epsilon$；反之若表述 2 成立，则根据定义，对于 x 的每个邻域 U，存在开集 $V\subseteq U$ 使得 $x\in V$，进而存在 $r>0$ 使得 $B(x,r)\subseteq V$。那么存在 $n_0\in\mathbb{N}$，使得 $x_n\in B(x,r)\subseteq U$ 对所有 $n\geq n_0$ 成立。

命题 1.10 (极限的唯一性)

度量空间 (X,d) 内的序列至多有一个极限。

证明 假设 (x_n) 在 (X,d) 中有两个极限值 x 和 y，则根据命题 1.9，对于给定的 $\epsilon>0$，总存在 $n_1\in\mathbb{N}$ 和 $n_2\in\mathbb{N}$，使得 $d(x_n,x)<\epsilon/2$ 对所有 $n\geq n_1$ 成立，且 $d(x_n,y)<\epsilon/2$ 对所有 $n\geq n_2$ 成立。取 $n_0=\max\{n_1,n_2\}$，则 $d(x,y)\leq d(x,x_n)+d(x_n,y)<\epsilon$ 对所有 $n\geq n_0$ 成立。因此 $y=x$。

此外，我们还可以用序列刻画闭包。

定理 1.4

设 U 为度量空间 (X,d) 内的子集，则 $x\in\bar{U}$ 当且仅当存在 U 内的序列 (x_n)，使得 $n\to\infty$ 时 $x_n\to x$。

证明 若 $x\in\bar{U}$，则 $B(x,1/n)$ 是 x 的一个邻域，其中 n 为正整数。根据定义，$U\cap B(x,1/n)\neq\varnothing$，因此总可以选择 $x_n\in U\cap B(x,1/n)$，因此 $d(x_n,x)<1/n$。利用 Archimedean 性质1.19，对于任意 $\epsilon>0$，总存在 $n_0\in\mathbb{N}$ 使得 $1/n_0<\epsilon$，故 $\lim_n x_n=x$；反之对于 x 的每个邻域 V，总存在开球 $B(x,r)\subseteq V$，进而对于该开球，存在 $n_0\in\mathbb{N}$ 使得 $d(x_n,x)<r$ 对所有 $n\geq n_0$ 成立，即存在 $x_n\in U\cap V$，因此 $x\in\bar{U}$。

事实上，还有另一个与收敛强相关的概念：Cauchy 列。相比于研究序列本身的极限值，在 Cauchy 列中我们通常仅关心其逐项特征，从而避免了通常繁琐的极限值讨论。

定义 1.24 (Cauchy 列)

设 (x_n) 为度量空间 (X,d) 内的序列。若对于每个 $\epsilon>0$，存在 $n_0\in\mathbb{N}$ 使得 $d(x_n,x_m)<\epsilon$ 对所有 $m,n\geq n_0$ 成立，则称 (x_n) 为 Cauchy 列（Cauchy sequence）。

命题 1.11

设 (X,d) 为度量空间，则每条收敛列都是 Cauchy 列。

证明 对于收敛到 x 的序列 (x_n) 和任意正数 ϵ，总存在 $n_0\in\mathbb{N}$ 使得 $d(x_n,x)<\epsilon/2$ 对所有 $n\geq n_0$ 成立。因此 $d(x_n,x_m)\leq d(x_n,x)+d(x,x_m)<\epsilon$ 对所有 $n,m\geq n_0$ 成立，从而 (x_n) 是 Cauchy 列。

不过反向结论通常不成立，即在一般的度量空间中，并非所有 Cauchy 列都有极限，从而引出如下定义。

定义 1.25 (完备度量空间)

若度量空间中的每条 Cauchy 列都具有极限，则称其完备（complete）。

♣

完备度量空间具有如下重要性质。

命题 1.12

设 (X,d) 为度量空间。

1. X 的完备子空间是闭集；
2. 若 X 完备，则 X 的子空间完备当且仅当其为闭集。

♠

证明

1. 若 X 的子空间 S 完备，则对于 S 内收敛到 x 的序列 (x_n)，则根据命题 1.11，其为 Cauchy 列。由于 S 完备，(x_n) 收敛到 S 内某个点 x'。由极限的唯一性（命题 1.10），$x=x'\in S$。进而根据定理 1.4，S 为闭集；
2. 这里仅需证明必要性。当 X 完备且其度量子空间 S 是闭集时，考虑 S 内的一组 Cauchy 列 (x_n)。由于 (x_n) 也是 X 内的 Cauchy 列，其收敛到某个 $x\in X$。由于 S 时闭集，故极限值 $x\in S$，从而 S 完备。

最后，我们考察积空间的完备性，其结论是直观的。

命题 1.13 (积空间的完备性)

设 $(X_i,d_i)\ (i=1,\cdots,n)$ 为完备度量空间，则相应的积空间关于积度量是完备的。

♠

证明 对于积空间 $X=X_1\times\times X_k$ 和积度量 $d(x,y)=\sum_{i=1}^k d(x_i,y_i)$，其内每条 Cauchy 列 (x_n) 中每个元素记为 $x_n=(x_{n,1},\cdots,x_{n,k})$。那么对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$，使得 $d_i(x_{n,i},x_{m,i})\le d(x_n,x_m)<\epsilon$。对所有 $n,m\ge n_0$ 成立。可见对于给定的 i，$(x_{n,i})_{n\in\mathbb{N}}$ 构成一条 Cauchy 列，因此其收敛到某个 $y_i\in X_i$。设 $y=(y_1,\cdots,y_k)\in X$，那么对于任意正数 ϵ，总存在 $n_i\in\mathbb{N}$，使得 $d(x_{n,i},y_i)<\epsilon/k$。取 $n_0=\max_i\{n_i\}$，则 $d(x_n,y)<\epsilon$ 对所有 $n\ge n_0$ 成立，从而 (X,d) 也是完备的。

1.3.3 紧致性

定义 1.26 (开覆盖与紧致性)

设 (X,d) 为度量空间，若存在 X 内的开集族 $(U_k)_{k\in K}$ 使得 $X\subseteq\cup_{k\in K}U_k$，则称其为 X 的开覆盖（open cover）。若对于每个开覆盖 $(U_k)_{k\in K}$，存在有限多个 $k_i\in K\ (i=1,\cdots,m)$ 使得 $(U_{k_i})_{i=1,\cdots,m}$ 是 X 的开覆盖，则称空间 X 是紧致的（compact），并称 $\cup_{i=1}^m U_{k_i}$ 为 X 的有限子覆盖（finite subcover）。

♣

例 1.5 Euclidean 空间 $\mathbb{R}^n$ 中的有界闭子集紧致。此性质源于 Heine-Borel 定理 1.10，我们将在第1.4.4小节介绍。

有了紧致性的概念后，一个关心的问题是：紧致集的子集是否紧致。以下命题将给出答案。

命题 1.14

紧致度量空间的闭子集是紧致的。

♠

证明 设 C 为紧致度量空间 X 的一个闭子集。若 $(U_k)_{k\in K}$ 为 C 的一个开覆盖，则将开集 $X\backslash C$ 加入该并集，可得 X 的一个开覆盖。由于 X 紧致，因此该开覆盖存在一个有限子覆盖，即存在有限个指标 k 使得 $X\subseteq(X\backslash C)\cup(\cup_{k=1}^m U_k)$，进而 $C=C\cap X\subseteq C\cap(\cup_{k=1}^m U_k)=\cup_{k=1}^m U_k$。因此 C 也是紧致的。

最后我们证明紧致度量空间的有限积也是紧致的。

命题 1.15

设 $(X_i,d_i)\ (i=1,\cdots,n)$ 为紧致度量空间，则 $X=X_1\times\cdots\times X_n$ 关于积度量紧致。

♠

证明 根据定理 1.3，对于任意 $x \in X$，存在开球 $B(x, r_x) \subseteq X$，进而 $\cup_{x\in X} B(x, r_x) \subseteq X$；另一方面，对于任意 $x \in X$，显然 $x \in B(x, r_x) \subseteq \cup_{x\in X} B(x, r_x)$，从而 $X \subseteq \cup_{x\in X} B(x, r_x)$。因此 $X = \cup_{x\in X} B(x, r_x)$。采用完全归纳法，当 $n = 1$ 时该命题的结论显然成立。假设该结论对于任意 $k < n$ 个紧致度量空间的笛卡尔积成立，那么设 X_1 的有限开覆盖为 $\cup_{i=1}^{k} B_1(x_{1,i}, r_{1,i})$，那么 $\cup_{i=1}^{k} B_1(x_{1,i}, r_{1,i}) \subseteq \cup_{x\in X_1} B(x, r_x) = X_1 \subseteq \cup_{i=1}^{k} B_1(x_{1,i}, r_{1,i})$，故 $X_1 = \cup_{i=1}^{k} B_1(x_{1,i}, r_{1,i})$；类似地，设 $X_2 \times \cdots \times X_n$ 的有限开覆盖为 $\cup_{j=1}^{m} B_2((x_{2,j}, \cdots, x_{n,j}), r_{2,j})$。对于任意 $y = (y_1, \cdots, y_n) \in X$，存在指标 i 使得 $d_1(x_{1,i}, y_1) < r_{1,i}$；存在指标 j 使得 $\sum_{i=2}^{n} d_i(x_{i,j}, y_i) < r_{2,j}$。因此存在 $z_{i,j} = (x_{1,i}, x_{2,j}, \cdots, x_{n,j}) \in X$ 使得 $d(y, z_{i,j}) < r_{1,i} + r_{2,j}$，即 $X \subseteq \cup_{i,j} B(z_{i,j}, r_{1,i} + r_{2,j})$，从而 X 紧致。

1.3.4 连续函数

定义 1.27 (连续函数)

对于两个度量空间 X 和 Y 之间的函数 $f : X \to Y$，若对于每个 $f(x)$ 的邻域 $V \subseteq Y$，存在 x 的邻域 $U \subseteq X$ 使得 $f(U) \subseteq V$，则称函数 f 在点 $x \in X$ 处连续（continuous at x）。若映射 $f : X \to Y$ 在所有 $x \in X$ 上连续，则称其连续（continuous），并记所有从 X 到 Y 的连续映射构成的集合记为 $C(X, Y)$。

关于逐点连续，存在几种等价表述。

命题 1.16 (逐点连续的刻画)

设 X, Y 为度量空间，$f : X \to Y$ 是一个函数，则下列表述等价：

1. f 在 $x \in X$ 处连续；
2. 对于每个 $\epsilon > 0$，存在 $\delta > 0$ 使得 $d_Y(f(x), f(y)) < \epsilon$ 对所有满足 $d_X(x, y) < \delta$ 的 $y \in X$ 成立；
3. 对于每个 X 中满足 $x_n \to x$ 的序列 (x_n)，当 $n \to \infty$ 时有 $f(x_n) \to f(x)$。

证明

1. 当表述 1 成立时，对于每个 $\epsilon > 0$ 以及 $f(x)$ 的邻域 $B(f(x), \epsilon)$，存在 x 的邻域 U 使得 $f(U) \subseteq B(f(x), \epsilon)$。根据定义，存在开集 $V \subseteq X$ 使得 $x \in V$，故存在开球 $B(x, \delta) \subseteq V$，因此 $f(B(x, \delta)) \subseteq B(f(x), \epsilon)$，表述 2 成立；
2. 当表述 2 成立时，对于 X 中每条收敛到 x 的序列 (x_n)，希望存在 $\delta > 0$，使得 $d_Y(f(x), f(x_n)) < \epsilon$ 对所有满足 $d_X(x, x_n) < \delta$ 的 $y \in X$ 成立。由 Archimedean 性质1.19，存在 $n_0 \in \mathbb{N}$ 使得 $1/n_0 < \delta$。因此 $d_Y(f(x), f(x_n)) < \epsilon$ 对于所有 $n \geq n_0$ 成立，即表述 3 成立；
3. 当表述 3 成立时，要证明表述 1 成立，可证明其逆否命题：若 f 在 $x \in X$ 处不连续，即存在 $f(x)$ 的某个邻域 $V \subseteq Y$，对于 x 的任意邻域 $U \subseteq X$，均有 $f(U) \not\subseteq V$。因此取 $U_n = \{y | d(x, y) < 1/n\}$，从中选择 x_n 使得 $f(x_n) \notin V$。此时由 Archimedean 性质（命题 1.19），对于任意正数 ϵ，存在自然数 n_0 使得 $1/n_0 < \epsilon$，即 $d(x, x_n) < \epsilon$ 对所有 $n \geq n_0$ 成立。那么 $(x_n) \to x$，但 $d(f(x), f(x_n)) > \sup\{d(f(x), y) | y \in V\} > 0$ 对所有自然数 n 成立，因此 $(f(x_n)) \nrightarrow f(x)$，表述 3 不成立。

类似地，关于连续函数，也存在如下等价表述。

命题 1.17 (连续函数的刻画)

设 X, Y 为度量空间，则下列表述等价：

1. $f \in C(X, Y)$；
2. 对于每个开集 O，$f^{-1}(O)$ 也是开集；
3. 对于每个闭集 C，$f^{-1}(C)$ 也是闭集；
4. 对于每个 $x \in X$ 和 $f(x)$ 的每个邻域 V，存在 x 的邻域 $U \subseteq X$ 使得 $f(U) \subseteq V$；
5. 对于每个 $x \in X$ 和 $\epsilon > 0$，存在 $\delta > 0$ 使得 $d_Y(f(x), f(y)) < \epsilon$ 对满足 $d_X(x, y) < \delta$ 的 $y \in X$ 成立。

证明 注意到(1.1.3)式提到的原像保持取补运算的性质，表述 2 和 3 显然等价。

1. 若表述 1 成立，则对于每个开集 O，任取 $x \in f^{-1}(O)$，则 O 是 $f(x)$ 的一个邻域。根据定义，$f^{-1}(O)$ 是 x 的一个邻域，进而存在开球 $B(x,r) \subseteq f^{-1}(O)$。结合 x 的任意性，可知 $f^{-1}(O)$ 为开集；
2. 若表述 2 成立，则对于每个 x 和 $f(x)$ 的每个邻域 V，存在开集 $V' \subseteq V$ 使得 $f(x) \in V'$，从而 $f^{-1}(V') \subseteq f^{-1}(V)$ 也是开集。由于 $x \in f^{-1}(V')$，$U = f^{-1}(V')$ 是 x 的一个邻域，且满足 $f(U) \subseteq V$，表述 4 成立；
3. 若表述 4 成立，则对于任意正数 ϵ，取 $V = B(x,\epsilon)$。那么存在 x 的邻域 $U \subseteq X$ 使得 $f(U) \subseteq B(x,\epsilon)$。根据定义，存在开球 $B(x,\delta) \subseteq U$，因此 $f(B(x,\delta)) \subseteq B(x,\epsilon)$，表述 5 成立；
4. 若表述 5 成立，则对于每个 $x \in X$ 和 $f(x)$ 的每个邻域 V，存在开球 $B(f(x),r) \subseteq V$。那么存在开球 $B(x,\delta)$ 使得 $f(B(x,\delta)) \subseteq B(f(x),\epsilon)$，同时 $B(x,\delta)$ 也是 x 的一个邻域，故表述 1 成立。

我们在紧致集上继续讨论连续函数的性质。

命题 1.18

若 $f \in C(X,Y)$ 且 X 紧致，则像 $f(X)$ 在 Y 内紧致。

♠

证明 设 $f(X)$ 存在开覆盖 $\cup_{k\in K}U_k$，其中每个 U_k 为开集。那么根据命题 1.17，每个 $f^{-1}(U_k)$ 也是开集，因此 $f^{-1}(\cup_{k\in K}U_k) = \cup_{k\in K}f^{-1}(U_k)$ 同样为开集。由于 X 紧致，要求存在有限多个开集 U_i，使得 $X \subseteq \cup_{i=1}^n f^{-1}(U_i) = f^{-1}(\cup_{i=1}^n U_i)$，进而 $f(X) \subseteq \cup_{i=1}^n U_i$，可见 Y 紧致。

根据上述命题，我们可导出实值连续函数的一个重要性质。

定理 1.5 (极值定理)

设 X 为紧致度量空间，$f \in C(X,\mathbb{R})$，则 f 可达到极大值和极小值，即存在 $x_1, x_2 \in X$ 使得 $f(x_1) = \inf\{f(x)|x \in X\}$ 且 $f(x_2) = \sup\{f(x)|x \in X\}$。

♡

证明 利用命题 1.18和 Heine-Borel 定理 1.10，$f(X)$ 是有界闭集，因此 $y = \inf\{f(x)|x \in X\} < +\infty$。根据下确界的定义，对于任意 $n \in \mathbb{N}$，总存在 $x_n \in X$ 使得 $y + 1/n > f(x_n) \geq y$，即 $|f(x_n) - y| = f(x_n) - y < \frac{1}{n}$，因此当 $n \to \infty$ 时 $f(x_n) \to y$。根据定理 1.4，$y \in f(X)$，即存在 $x_1 \in X$ 使得 $f(x_1) = y = \inf\{f(x)|x \in X\}$。另一条结论类似可证。

1.4 实数基础

这一节我们对实数进行更严格、更系统的公理化定义，并补充一些在后文会经常用到的实数相关的性质。

1.4.1 域公理、正性公理和完备性公理

我们假设实数集 $\mathbb{R}$ 已给定，并且对于每对实数 a 和 b，存在已定义的实数 $a+b$ 和 ab，分别称为 a 与 b 的和与积，其满足如下域公理、正性公理和完备性公理。

公理 1.2 (域公理)

实数集 $\mathbb{R}$ 在加法和乘法运算下构成一个域。

♡

公理 1.3 (正性公理)

存在一个称为正数（positive number）集的实数集，其满足以下两条性质：

1. *若 a 和 b 为正，则 ab 和 $a+b$ 也为正；*
2. *对于一个实数 a，下列三种情况有且仅有一种成立：a 是正数、$-a$ 是正数和 $a = 0$。*

♡

正性公理自然地引出了实数的一种序关系：对于实数 a 和 b，我们定义 $a > b$ 为 $a - b$ 为正，定义 $a \geq b$ 为 $a > b$ 或 $a = b$。随后可定义 $a < b$ 和 $b \leq a$。

利用域公理和正性公理，可以形式化地建立不等式中一些熟悉的性质。给定满足 $a<b$ 的实数 a 和 b，我们定义 $(a,b)=\{x|a<x<b\}$，并称 (a,b) 中的点存在于 a 和 b 之间。若对于非空实数集 I 中任意两点，所有介于这两点之间的所有点属于 I，则称其为区间（*interval*）。当然，集合 (a,b) 就是一个区间，以下集合也是区间：

$$[a,b]=\{x|a\le x\le b\},\quad [a,b)=\{x|a\le x<b\},\quad (a,b]=\{x|a<x\le b\}. \tag{1.4.1}$$

为了引入三角不等式，我们首先引入实数 x 的绝对值（*absolute value*）$|x|$ 的概念：若 $x\ge 0$，则 $|x|=x$；若 $x<0$，则 $|x|=-x$。事实上，对于任意实数 x 和 y，容易验证绝对值 $|x-y|$ 实际上是实数集 $\mathbb{R}$ 上的度量。根据此定义，容易证明三角不等式（*triangle inequality*）：对于任意实数对 a 和 b，$|a+b|\le|a|+|b|$。

公理 1.4 (完备性公理)

设 E 为上有界[a]的非空实数集，则在集合 E 的所有上界组成的集合中存在一个最小上界。

[a]若非空集合 E 存在一个实数 b，使得对于所有 $x\in E$ 均有 $x\le b$，则称 E 上有界（bounded above），实数 b 称为 E 的上界

♡

注 事实上，公理要求的完备性可导出完备度量空间定义 1.25所提及的完备性：对于任意 Cauchy 列 (x_n) 和任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $|x_n-x_m|<\epsilon$ 对所有 $n,m\ge n_0$ 成立。因此该 Cauchy 列构成的非空实数集 $E=\{x_n|n\in\mathbb{N}\}$ 上有界 $\max\{\max\{x_n|n<n_0\},x_{n_0}+\epsilon\}$。那么 E 具有上确界 $x=\sup\{x_n|n\in\mathbb{N}\}$，从而对于任意正数 ϵ'，存在 $m\in\mathbb{N}$ 使得 $x_m+\epsilon'>x\ge x_m$，即 $|x_m-x|=x-x_m<\epsilon$。这意味着该 Cauchy 列存在聚点，因此收敛。

另外，引入符号 ∞ 和 $-\infty$ 并对所有实数 x 写为 $-\infty<x<\infty$ 将带来便利。我们称 $\mathbb{R}\cup\{-\infty,\infty\}$ 为扩充实数（*extended real number*）。若非空实数集 E 不是上有界的，我们定义其在扩充实数集内的上确界为 ∞，并约定 $-\infty$ 是空集的上确界。因此，每个实数集都具有属于扩充实数的上确界。类似地，我们也可以推广下确界的概念。我们定义 $(-\infty,\infty)=\mathbb{R}$，对于 $a,b\in\mathbb{R}$ 定义

$$(a,\infty)=\{x\in\mathbb{R}|a<x\},\qquad (-\infty,b)=\{x\in\mathbb{R}|x<b\}, \tag{1.4.2}$$

$$[a,\infty)=\{x\in\mathbb{R}|x\le x\},\qquad (-\infty,b]=\{x\in\mathbb{R}|x\le b\}. \tag{1.4.3}$$

1.4.2 自然数和有理数

为了引入自然数的概念，我们首先需要介绍归纳集。

定义 1.28 (归纳集)

若实数集 E 包含 1，并且只要实数 $x\in E$，就有 $x+1\in E$，则称 E 是归纳的。

♣

整个实数集是归纳的。根据不等式 $1>0$，我们推出集合 $\{x\in\mathbb{R}|x\ge 0\}$ 和 $\{x\in\mathbb{R}|x\ge 1\}$ 都是归纳的。自然数（*natural number*）集（记为 $\mathbb{N}$）定义为$\mathbb{R}$ 中所有归纳子集的交集。容易证明 $\mathbb{N}$ 也是归纳的。

定理 1.6 (数学归纳法原理)

对于每个自然数 n，设 $S(n)$ 为某个数学命题。假定 $S(1)$ 为真，并假定当 k 为使得 $S(k)$ 为真的自然数时，$S(k+1)$ 同样为真。那么 $S(n)$ 对所有自然数 n 为真。

♡

证明 定义自然数构成的集合 $A=\{k\in\mathbb{N}|S(k)\text{ is true}\}\subseteq\mathbb{N}$，那么定理中关于 $S(k)$ 的约束意味着 A 是一个归纳集，因此 $A\subseteq\mathbb{N}$。故 $A=\mathbb{N}$，即 $S(n)$ 对所有自然数 n 为真。

定理 1.7

每个由自然数构成的非空集合都具有最小元素。

♡

证明 设 E 为非空自然数集。由于集合 $\{x\in\mathbb{R}|x\ge 1\}$ 是归纳的，自然数下有界 1，从而 E 下有界 1。作为完备性公理的结果，E 有下确界 $c=\inf E$。由于 $c+1$ 不是 E 的下界，存在 $m\in E$ 使得 $m<c+1$。假设 m 不是 E 的最小元素，则存在 $n\in E$ 使得 $n<m$。由于 $n\in E$，满足 $c\le n<m<c+1$，这意味着 $m-n<1$，即自然数

$m \in (n, n+1)$。设数学命题 $S(n)$ 为“$\mathbb{N} \cap (n, n+1) = \varnothing$ 对自然数 n 成立”。对于 $S(1)$，已知 $E' = \{1, 2, \cdots\}$ 是一个归纳集，若存在 $a \in \mathbb{N} \cap (1, 2)$，则 $a \notin E'$，这与 $\mathbb{N} \subseteq E'$ 矛盾，故 $S(1)$ 为真；假设 $S(k)$ 对某个自然数 k 为真，那么若存在 $a \in \mathbb{N} \cap (k+1, k+2)$，则 $a - 1 \in \mathbb{N} \cap (k, k+1)$，与 $S(k)$ 矛盾，故 $S(k+1)$ 为真。综上，区间 $(n, n+1)$ 内不存在其他自然数，从而 m 为 E 的最小元素。

命题 1.19 (Archimedean 性质)

对于每对正实数 a 和 b，存在自然数 n 使得 $na > b$。

♠

证明 定义 $c = b/a > 0$，我们通过反证法证明该结论。若对于所有 $n \in \mathbb{N}$ 均有 $n \leq c$，则 c 是自然数的上界。根据完备性公理，自然数具有上确界 $c_0 = \sup \mathbb{N}$，从而 $c_0 - 1$ 不是自然数的上界。选择自然数 n 使得 $n > c_0 - 1$，由于 $\mathbb{N}$ 是归纳集，$n+1 \in \mathbb{N}$ 且 $n+1 > c_0$，矛盾。故存在自然数 n 使得 $n > c$。

注 通常我们将 $\mathbb{R}$ 的 Archimedean 性质改写为：对于任意正数 ϵ，存在自然数 n 使得 $1/n < \epsilon$。

我们将整数（*integer*）集（记作 $\mathbb{Z}$）定义为由 0、自然数及其逆元组成的集合。有理数（*rational number*）集（记作 $\mathbb{Q}$）定义为由整数的商构成的集合。非有理的实数称为无理的（*irrational*）。

定义 1.29 (实数的稠密性)

若对于任意两个实数，存在其间一个点位于实数集 E 内，则称 E 在 $\mathbb{R}$ 中稠密。

♣

定理 1.8

有理数在 $\mathbb{R}$ 中稠密。

♡

证明 设 a 和 b 满足 $a < b$ 的实数。首先假定 $a \geq 0$，则根据 Archimedean 性质1.19，存在正有理数 q 使得 $1/q < b - a$。再次利用 Archimedean 性质1.19，自然数集 $S = \{n \in \mathbb{N} | n/q \geq b\}$ 非空。根据定理 1.7，S 具有最小元素 p，因此 $p \geq qb > 1 + qa > 1$，从而 $p - 1 \in \mathbb{N}$。利用 p 的最小性，$p - 1 < qb$，进而

$$a = b - (b - a) < \frac{p}{q} - \frac{1}{q} = \frac{p-1}{q} < b. \tag{1.4.4}$$

因此，存在有理数 $r = (p-1)/q$ 介于 a 和 b 之间。若 $a < 0$，则根据 Archimedean 性质1.19，存在自然数 $n > -a$，因此存在有理数 r 使得 $n + a < r < n + b$，即存在有理数 $r - n$ 介于 a 和 b 之间。

1.4.3 可数集与不可数集

我们现在回到第1.1.4小节介绍的可数集与不可数集，并关于实数证明几个简单的结论。

定理 1.9

可数集的子集是可数的。特别地，每个由自然数构成的集合是可数的。

♡

证明 对于可数集 S，其基数 $|S| \leq \aleph_0$，那么其任一子集 T 的基数满足 $|T| \leq |S| \leq \aleph_0$，故 T 可数。

推论 1.1

以下集合是可数无穷的：

1. 对每个自然数 n，笛卡尔积 $\underbrace{\mathbb{N} \times \cdots \times \mathbb{N}}_{n \text{ times}}$；
2. 有理数集 $\mathbb{Q}$。

♡

证明

1. 该结论在 $n = 1$ 时显然成立。假设结论在 $n = k$ 时成立，那么当 $n = k + 1$ 时，定义映射

$$g : \underbrace{\mathbb{N} \times \cdots \times \mathbb{N}}_{(k+1) \text{ times}} \to \underbrace{\mathbb{N} \times \cdots \times \mathbb{N}}_{k \text{ times}}, \quad (n_1, \cdots, n_{k+1}) \mapsto ((n_2 + n_1)^2 + n_1, \cdots, (n_{k+1} + n_1)^2 + n_1). \tag{1.4.5}$$

取 $n_1=0$ 可知 g 为满射；若存在两组自然数有序对 $(n_1,\cdots,n_{k+1})$ 和 $(n_1',\cdots,n_{k+1}')$ 使得 g 的像相同，则

$$|n_1-n_1'|=|(n_i+n_1)+(n_i'+n_1')|\cdot|(n_i+n_1)-(n_i'+n_1')| \text{ for any } i=2,\cdots,k+1. \tag{1.4.6}$$

若 $n_1\neq n_1'$，则 $|n_1-n_1'|$ 整除 $|(n_i+n_1)+(n_i'+n_1)|>|n_1-n_1'|$，这是不可能的，因此 $n_1=n_1'$，进而 $n_i=n_i'$ 对所有 $i=2,\cdots,k+1$ 成立。故 g 为单射，从而 g 为双射。根据归纳假设，$\underbrace{\mathbb{N}\times\cdots\mathbb{N}}_{(k+1)\text{ times}}$ 与 $\underbrace{\mathbb{N}\times\cdots\mathbb{N}}_{k\text{ times}}$ 等势，因此前者可数无穷。从而 $\underbrace{\mathbb{N}\times\cdots\mathbb{N}}_{n\text{ times}}$ 对所有自然数 n 都是可数无穷的；

2. 一个不太严格但十分直观的证明方法是，将自然数有序对 (m,n) 置入一个二维无穷大表格中，其中每行、每列均为从小到大排列的全体自然数。按照 $(1,1)\to(1,2)\to(2,1)\to(3,1)\to(2,2)\to(1,3)\to\cdots$ 的顺序将这些有序对与自然数一一对应。剔除掉重复的商 m/n，记实数集 $E=\{m/n|m,n\in\mathbb{N}\}$，那么 $|E|\leq|\mathbb{N}|$；另一方面，全体自然数都包含在 E 内（取 $n=1$），因此 $|\mathbb{N}|\leq|E|$。综上，$|E|=|\mathbb{N}|=\aleph_0$。

1.4.4 实数的开集与闭集

我们已经在第1.3节讨论了一般度量空间中开集与闭集的性质，现在我们将注意力转向实数集这一特殊情形上的开集与闭集。由于实数具有更丰富的公理化要求，在实数域上的开集与闭集具有一些更优美的性质。

定义 1.30 (开集)

若对于实数集 U 中的每个元素 x，存在 $r>0$ 使得区间 $(x-r,x+r)$ 包含在 U 内，则称 U 为开的（open）。♣

注 对于实数 $a<b$，区间 (a,b) 是开集。这是因为对于区间内的任一实数 x，可取 $r=\min\{b-x,x-a\}$，使得 $(x-r,x+r)\subseteq(a,b)$。故称区间 (a,b) 是有界开区间，并且每个有界开区间都具有此形式。

命题 1.20

实数集 $\mathbb{R}$ 和非空集合 $\varnothing$ 是开集；任意有限多开集族的交集为开集；任意开集族的并集为开集。♠

证明 显然 $\mathbb{R}$ 和 $\varnothing$ 是开集。对于有限多开集 $A_1,\cdots,A_n$ 以及任意实数 $x\in\cap_{i=1}^n A_i$，存在正数 r_i 使得 $(x-r_i,x+r_i)\subseteq A_i$ 对所有指标 i 成立。取 $r=\min\{r_i|i=1,\cdots,n\}$，则 $(x-r,x+r)\subseteq\cap_{i=1}^n A_i$，从而 $\cap_{i=1}^n A_i$ 为开集；对于开集族 $\{A_k|k\in K\}$ 以及任意实数 $x\in\cup_k A_k$，存在指标 k 使得 $x\in A_k$，因此存在正数 r 使得 $(x-r,x+r)\subseteq A_k\subseteq\cup_k A_k$，故 $\cup_k A_k$ 为开集。

注 需要注意的是，并非任意开集族的交集为开集。例如对于每个自然数 n，考虑开区间族 $\{A_n|n\in\mathbb{N}\}$，其中 $A_n=(-1/n,1+1/n)$ 那么根据 Archimedean 性质1.19，$\cap_n A_n=[0,1]$ 是闭区间。

命题 1.21

每个非空开集都是可数个不相交开区间族的并集。♠

证明 设 U 为 $\mathbb{R}$ 的非空开子集且实数 $x\in U$，那么根据定义，存在 $y>x$ 使得 $(x,y)\in U$，且存在 $z<x$ 使得 $(z,x)\in U$。如下定义扩充实数 a_x 和 b_x：

$$a_x=\inf\{z|(z,x)\subseteq U\},\quad b_x=\sup\{y|(x,y)\subseteq U\}, \tag{1.4.7}$$

则 $I_x=(a_x,b_x)$ 是包含 x 的一个开区间。我们断言

$$I_x\subseteq U\quad\text{but}\quad a_x\notin U,\ b_x\notin U. \tag{1.4.8}$$

设 $w\in I_x$，不妨取 $x<w<b_x$。根据 b_x 的定义，存在 $y>w$ 使得 $(x,y)\subseteq U$，否则所有满足 $(x,y)\subseteq U$ 的 y 都应满足 $y\leq w$，从而由上确界的定义，$b_x\leq w$，矛盾。因此 $w\in(x,y)\subseteq U$，当 $a_x<w<x$ 时证明类似，因此 $I_x\subseteq U$。现在假设 $b_x\in U$，则存在正数 r 使得 $(b_x-r,b_x+r)\subseteq U$。根据上述证明和命题 1.20，$(x,b_x+r)=(x,b_x)\cup(b_x-r,b_x+r)$ 也是 U 内的开集。但这与 b_x 的极大性矛盾，故 $b_x\notin U$，类似地 $a_x\notin U$。

现在考虑开集族 $\{I_x\}_{x\in U}$。首先由 $I_x\subseteq U$ 对任意 $x\in U$ 成立可知 $\cup_{x\in U}I_x\subseteq U$；其次由于 U 的每个元素

x 都在 $I_x \subseteq \cup_{x\in U} I_x$ 内，故 $U \subseteq \cup_{x\in U} I_x$。综上，$U = \cup_{x\in U} I_x$。假设存在 $x, y \in U$ 使得 $I_x \cap I_y \neq \varnothing$，则有 $a_y \in I_x \subseteq U$ 或 $a_x \in I_y \subseteq U$，矛盾。故 $\{I_x\}_{x\in U}$ 时不相交开集族。

最后证明该族是可数的。利用有理数的稠密性（定理 1.8），可知每个开区间包含一个有理数，因此存在一个从有理数到该区间族的满射，即区间族的基数不大于 $|\mathbb{Q}| = \aleph_0$，故存在可数区间族。

定义 1.31 (闭包)

对于实数集 E，若每个包含实数 x 的开区间也包含 E 中的一个点，则称 x 为 E 的闭包点（point of closure）。E 的闭包点族称为 E 的闭包（closure），记作 $\bar{E}$。 ♣

注 显然总有 $E \subseteq \bar{E}$。若 E 包含其所有闭包点，即 $E = \bar{E}$，则称集合 E 是闭的（closed）。

命题 1.22

对于实数集 E，其闭包 $\bar{E}$ 是闭的。此外，$\bar{E}$ 是包含 E 的最小闭集，即如果 F 是闭的且 $E \subseteq F$，则 $\bar{E} \subseteq F$。 ♠

证明 设 x 为 $\bar{E}$ 的一个闭包点，因此考虑包含 x 的一个开区间 I_x。根据定义，存在 $x' \in \bar{E} \cap I_x$，那么 x' 是 E 的闭包点且 I_x 是包含 x' 的开区间，从而存在 $x'' \in E \cap I_x$。因此每个包含 x 的开区间 I_x 包含 E 内的一个点 x''，进而 $x \in \bar{E}$，即 $\bar{E}$ 的闭包包含在 $\bar{E}$ 内。那么 $\bar{E}$ 与其闭包相等，故 $\bar{E}$ 是闭的。若 $E \subseteq F$，容易证明 $\bar{E} \subseteq \bar{F} = F$。

我们已经了解了一般度量空间中闭集的定义，因此需要证明两种定义在 $\mathbb{R}$ 上的等价性。

命题 1.23

非空实数集是开的当且仅当其在 $\mathbb{R}$ 中的补集是闭的。 ♠

证明 对于实数集 U，其补集 $U^c = \mathbb{R}\backslash U$ 是闭的当且仅当 $\overline{\mathbb{R}\backslash U} \subseteq \mathbb{R}\backslash U$。若 U 是开的，那么对于任意 $x \in U$，存在正数 r 使得 $(x-r, x+r) \subseteq U$。假设 $x \in \overline{\mathbb{R}\backslash U}$，则 $(\mathbb{R}\backslash U) \cap (x-r, x+r) \neq \varnothing$，但 $(\mathbb{R}\backslash U) \cap (x-r, x+r) \subseteq (\mathbb{R}\backslash U) \cap U = \varnothing$，矛盾。因此 $x \notin \overline{\mathbb{R}\backslash U}$，即 $U \subseteq \mathbb{R}\backslash(\overline{\mathbb{R}\backslash U})$，进而 $\overline{\mathbb{R}\backslash U} \subseteq \mathbb{R}\backslash U$；反之对于必要性，考虑 U 内的元素 x，则 $x \notin \mathbb{R}\backslash U = \overline{\mathbb{R}\backslash U}$，那么存在包含 x 的区间 I_x，使得 $I_x \cap (\mathbb{R}\backslash U) = \varnothing$，取补集得 $(\mathbb{R}\backslash I_x) \cup U = \mathbb{R}$。因此 $I_x = I_x \cap \mathbb{R} = I_x \cap U$，即 $I_x \subseteq U$。对于端点分别为 $a < b$ 的区间 I_x，其必然包含区间 (a, b)，进而存在 $r = \min\{b-x, x-a\}$ 使得 $(x-r, x+r) \subseteq U$，故 U 为开集。

对于闭集，我们也有与命题 1.20类似的结论，根据上述命题取补集即可完成证明。

命题 1.24

实数集 $\mathbb{R}$ 和非空集合 $\varnothing$ 闭集；任意有限多闭集族的并集为闭集；任意闭集族的交集为闭集。 ♠

我们已经在一般拓扑空间中定义了紧致性，但其概念相对抽象。而在实数集中，紧致性有了具体的刻画。

定理 1.10 (Heine-Borel 定理)

实数集紧致当且仅当其为有界闭集。 ♡

证明 若实数集 E 紧致，对于每个 $x \in E$ 和给定的 $y \notin E$，存在正数 r_x 和 r_y 使得 $r_x + r_y < |x-y|$。记开区间 $I_x = (x-r_x, x+r_x)$ 且 $I_y = (y-r_y, y+r_y)$，那么 $\cup_{x\in E} I_x$ 是 E 的一个开覆盖。由于 E 紧致，存在 $x_1, \cdots, x_k \in E$ 使得 $E \subseteq \cup_{i=1}^k I_{x_i}$。记 $a = \inf\{x_i - r_{x_i} | 1 \leq i \leq k\}$ 且 $b = \sup\{x_i + r_{x_i} | 1 \leq i \leq k\}$，则 $E \subseteq (a, b)$，故 E 有界。另一方面，根据命题 1.20，$U = \cap_{i=1}^k I_{y_i}$ 也是开集。由于 $I_{x_i} \cap (\cap_{j=1}^k I_{y_j}) \subseteq I_{x_i} \cap I_{y_i} = \varnothing$ 对所有指标 i 成立，故 $(\cup_{i=1}^k I_{x_i}) \cap (\cap_{j=1}^k I_{y_j}) = \cup_{i=1}^k [I_{x_i} \cap (\cap_{j=1}^k I_{y_j})] = \varnothing$。可见对于每个 $y \in \mathbb{R}\backslash E$，均存在包含 y 的开区间 U 使得 $U \subseteq \mathbb{R}\backslash E$。根据命题 1.23，$\mathbb{R}\backslash E$ 是开集，进而 E 是闭集。充分性得证。

反之，若 E 是有界闭集，设 $\cup_{k\in K} U_k$ 为 E 的开覆盖。首先考虑 $E = [a, b]$ 为闭区间的情形，其中 $a < b$。设 F 由所有使得 $[a, x]$ 可被 $\{U_k | k \in K\}$ 中有限开集覆盖的元素 $x \in E$ 构成。由于 $a \in F$，$F \neq \varnothing$；由于 F 上有界 b，根据完备性公理，存在上确界 $c = \sup F$ 且 $c \in E$，进而存在指标 k 使得 $c \in U_k$。由于 U_k 为开集，存在正数 ϵ 使得 $(c-\epsilon, c+\epsilon) \subseteq U_k$。注意到 $c - \epsilon$ 不是 F 的上界，存在 $x \in F$ 使得 $x > c - \epsilon$。又因为 $x \in F$，存在 $[a, x]$ 的有

限子覆盖 $\cup_{i=1}^{n}U_{k_i}$，进而 $U_k\cup(\cup_{i=1}^{n}U_{k_i})$ 是 $[a,c+\epsilon)$ 的有限开覆盖。若 $c<b$，则存在 $\epsilon'=\min\{\epsilon,b-c\}/2>0$，使得 $[a,c+\epsilon']$ 可被 $U_k\cup(\cup_{i=1}^{n}U_{k_i})$ 覆盖，这与 c 的极大性矛盾。因此 $c=b$，从而 E 存在有限开覆盖。若 E 为任意闭集，设 $\{U_k|k\in K\}$ 的开覆盖。由于 E 有界，其必然包含在某个闭区间 $[a,b]$ 内。根据命题 1.23，$U=\mathbb{R}\backslash E$ 为开集，因此 $[a,b]\subseteq U$，从而 $U\cup(\cup_{k\in K}U_k)$ 是 $[a,b]$ 的一个开覆盖。结合前述证明，该开覆盖存在有限子覆盖。若有限开覆盖不包含 U，则 $\cup_{i=1}^{n}U_{k_i}$ 覆盖 $[a,b]$，自然也覆盖 E；反之若其包含 U，即存在自然数 n 使得 $[a,b]\subseteq U\cup(\cup_{i=1}^{n}U_{k_i})$，那么

$$E=E\cap[a,b]\subseteq E\cap[U\cup(\cup_{i=1}^{n}U_{k_i})]=E\cap(\cup_{i=1}^{n}U_{k_i})\subseteq(\cup_{k\in K}U_k)\cap(\cup_{i=1}^{n}U_{k_i})=\cup_{i=1}^{n}U_{k_i}. \tag{1.4.9}$$

因此无论是何种情况，E 的每个开覆盖都存在有限子覆盖，故 E 是紧致的。

对于可数集 $\{E_n\}_{n=1}^{\infty}$，若 $E_{n+1}\subseteq E_n$ 对每个自然数 n 成立，则称其递减（*descending*）；若 $E_n\subseteq E_{n+1}$ 对每个自然数 n 成立，则称其递增（*ascending*）。

定理 1.11 (闭区间套定理)

若 $\{E_n\}_{n=1}^{\infty}$ 是递减可数非空闭实数集，并且 E_1 有界，则 $\cap_{n=1}^{\infty}E_n\neq\varnothing$。

♡

证明 假设 $\cap_{n=1}^{\infty}E_n=\varnothing$，则 $\mathbb{R}=\cup_{n=1}^{\infty}E_n^c$。根据命题 1.23，$E_n^c$ 是开集，因此 $\cup_{n=1}^{\infty}E_n^c$ 是 $\mathbb{R}$ 也是 E_1 的一个开覆盖。利用 Heine-Borel 定理 1.10，E_1 存在有限子覆盖 $\cup_{i=1}^{k}E_{n_i}^c$。由于 $\{E_n\}_{n=1}^{\infty}$ 递减，$\{E_n^c\}_{n=1}^{\infty}$ 递增，从而 $E_1\subseteq E_{n_k}^c$，但此时 $E_{n_k}\cap E_1\subseteq E_{n_k}\cap E_{n_k}^c=\varnothing$，这与 $E_{n_k}\subseteq E_1$ 矛盾。因此假设不成立，即 $\cap_{n=1}^{\infty}E_n\neq\varnothing$。

1.4.5 实数列

实数的序列（*sequence*）是一个定义在自然数上的实值函数 $f:\mathbb{N}\to\mathbb{R}$。但习惯上我们不采用函数表示，而用下标 a_n 代替 $f(n)$，并将序列记为 $\{a_n\}$。自然数 n 称为序列的指标（*index*），对应于指标 n 的数 a_n 称为该序列的第 n 项。对于序列 $\{a_n\}$，若存在 $c\geq 0$ 使得 $|a_n|\leq c$ 对所有 n 成立，则称其有界（*bounded*）。若 $a_n\leq a_{n+1}$ 对所有 n 成立，则称该序列递增（*increasing*）；若 $\{-a_n\}$ 递增，则称 $\{a_n\}$ 递减（*decreasing*）；若序列要么递增要么递减，则称其单调（*monotone*）。实数列的一般收敛性质已经在第1.3.2小节讨论过，这里我们介绍其特有性质。

定理 1.12 (实数列的单调收敛准则)

单调实数列收敛当且仅当其有界。

♡

证明 设实数列 $\{a_n\}$ 递增。若其收敛到 a，则对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $|a_n-a|<\epsilon$ 对所有 $n\geq n_0$ 成立。因此设 $c_+=\max\{a_{n_0},a+\epsilon\}$ 且 $c_-=\min\{a_1,a-\epsilon\}$，故 $c_-\leq a_n\leq c_+$ 对所有自然数 n 成立，即 $\{a_n\}$ 有界；反之若存在 $c\geq 0$ 使得 $|a_n|\leq c$ 对所有 $n\in\mathbb{N}$ 成立，那么根据完备性公理，集合 $S=\{a_n|n\in\mathbb{N}\}$ 具有上确界 $a=\sup S$。那么对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $a\geq a_{n_0}\geq a-\epsilon$，即 $|a_n-a|=a-a_n\leq a-a_{n_0}<\epsilon$ 对所有 $n\geq n_0$ 成立，故 a 为 $\{a_n\}$ 的极限。

对于实数列 $\{a_n\}$ 和严格单调增的自然数列 $\{n_k\}$，称第 k 项为 a_{n_k} 的序列 $\{a_{n_k}\}$ 为 $\{a_n\}$ 的子列（*subsequence*）。

定理 1.13 (Bolzano-Weierstrass 定理)

每个有界实数列都具有一个收敛子列。

♡

证明 设 $\{a_n\}$ 为有界实数列，并取 $c\geq 0$ 使得 $|a_n|\leq c$ 对所有 $n\in\mathbb{N}$。对于自然数 n，定义闭包 $E_n=\overline{\{a_j|j\geq n\}}$，并且根据定义 $E_n\subseteq[-c,c]$。因此 $\{E_n\}$ 是 $\mathbb{R}$ 的一个有界非空闭子集构成的递减序列。利用闭区间套定理 1.11，存在实数 $a\in\cap_{n=1}^{\infty}E_n$。因此对于每个自然数 k，$a\in E_k$；同时根据定理 1.4，存在 ${}_an_k\in\{a_j|j\geq k\}$ 使得 $|a_{n_k}-a|<1/k$。因此 $\{a_{n_k}\}$ 是 $\{a_n\}$ 中收敛到 a 的子列。

定理 1.14 (实数列的 Cauchy 收敛准则)

实数列收敛当且仅当其为 Cauchy 的。

♡

证明 充分性已由命题 1.11给出；对于必要性，设 $\{a_n\}$ 为 Cauchy 列，那么对于任意正数 ϵ，存在 $n_0 \in \mathbb{N}$ 使得 $|a_n - a_m| < \epsilon/2$ 所所有 $n, m \geq n_0$ 成立。仿照上述定理的证明，对每个自然数 n，定义闭包 $E_n = \overline{\{a_j | j \geq n\}}$。利用闭区间套定理 1.11，存在实数 $a \in \cap_{n=1}^{\infty} E_n$，从而 $a \in E_{n_0}$，即存在 a_m 使得 $|a_m - a| < \epsilon/2$，其中 $m \geq n_0$。因此 $|a_n - a| \leq |a_n - a_m| + |a_m - a| < \epsilon$ 对每个 $n \geq n_0$ 成立，即 Cauchy 列 $\{a_n\}$ 收敛到 a。

定理 1.15 (实数序列收敛的线性性和单调性)

设 $\{a_n\}$ 和 $\{b_n\}$ 为收敛实数列，则对于每对实数 α 和 β，序列 $\{\alpha a_n + \beta b_n\}$ 收敛，且

$$\lim_{n\to\infty} (\alpha a_n + \beta b_n) = \alpha \lim_{n\to\infty} a_n + \beta \lim_{n\to\infty} b_n. \tag{1.4.10}$$

此外，若 $a_n \leq b_n$ 对所有 $n \in \mathbb{N}$ 成立，则 $\lim_n a_n \leq \lim_n b_n$。

证明 设 $a = \lim_n a_n, b = \lim_n b_n$，则对于任意正数 ϵ，存在自然数 n_1 和 n_2，使得 $|a_n - a| < \epsilon/2\alpha$ 对所有 $n \geq n_1$ 成立、$|b_n - b| < \epsilon/2\beta$ 对所有 $n \geq n_2$ 成立。记 $n_0 = \max\{n_1, n_2\}$，则

$$|(\alpha a_n + \beta b_n) - (\alpha a + \beta b)| \leq \alpha |a_n - a| + \beta |b_n - b| < \epsilon \tag{1.4.11}$$

对所有 $n \geq n_0$ 成立，即 $\lim_n(\alpha a_n + \beta b_n) = \alpha a + \beta b$。当 $a_n \leq b_n$ 对所有自然数 n 成立时，记 $c_n = b_n - a_n \geq 0$。根据上述结果，实数列 $\{c_n\}$ 收敛到 $c = b - a$。那么对于任意正数 ϵ，存在 $n_0 \in \mathbb{N}$ 使得 $-\epsilon < c - c_n < \epsilon$ 对所有 $n \geq n_0$ 成立，即 $c > c_n - \epsilon \geq -\epsilon$ 对所有正数 ϵ 成立。故 $c \geq 0$，即 $\lim_n a_n \leq \lim_n b_n$。

实数列 $\{a_n\}$ 收敛到无穷（*converges to infinity*）意味着对于每个实数 c，均存在 $n_0 \in \mathbb{N}$，使得 $a_n \geq c$ 对所有 $n \geq n_0$ 成立，记作 $\lim_n a_n = \infty$；收敛到负无穷的定义类似。有了这一推广的收敛概念，根据单调收敛准则（定理 1.12），任何单调实数列 $\{a_n\}$ 都在扩充实数内有界，因而收敛。对于任意单调实数列，其推广的收敛性概念允许我们引出如下定义。

定义 1.32 (上极限与下极限)

设 $\{a_n\}$ 为实数列，则 $\{a_n\}$ 的上极限（limit superior）和下极限（limit inferior）分别定义为

$$\limsup\{a_n\} = \lim_{n\to\infty} (\sup\{a_k | k \geq n\}), \tag{1.4.12}$$

$$\liminf\{a_n\} = \lim_{n\to\infty} (\inf\{a_k | k \geq n\}). \tag{1.4.13}$$

注 由于 $\sup\{a_k | k \geq n\}$ 和 $\inf\{a_k | k \geq n\}$ 分别为递减和递增列，根据定理 1.12，两者（在扩充实数内）必然收敛。

命题 1.25

设 $\{a_n\}$ 和 $\{b_n\}$ 为实数列。

1. $\limsup\{a_n\} = \infty$ 当且仅当 $\{a_n\}$ 不是上有界的；
2. $\limsup\{a_n\} = -\liminf\{-a_n\}$；
3. 实数列 $\{a_n\}$ 收敛到扩充实数 a 当且仅当
$$\liminf\{a_n\} = \limsup\{a_n\} = a. \tag{1.4.14}$$
4. 若 $a_n \leq b_n$ 对所有 n 成立，则
$$\limsup\{a_n\} \leq \limsup\{b_n\}. \tag{1.4.15}$$

证明

1. 若 $\{a_n\}$ 上有界 M，则根据完备性公理，$\{a_n\}$ 有上确界 $c = \sup\{a_n\}$，显然 $\sup\{a_k | k \geq n\} \leq c$。利用收敛的单调性（定理 1.15），$\limsup\{a_n\} \leq c < \infty$；反之若 $\{a_n\}$ 不是上有界的，即对于任意正数 M，存在 $n_0 \in \mathbb{N}$ 使得 $a_n > M$ 对所有 $n \geq n_0$ 成立。那么 $\sup\{a_k | k \geq n_0\} \geq a_{n_0} > M$，即 $\limsup\{a_n\} = \infty$；
2. 根据定义，$\limsup\{a_n\} = \lim_n(\sup\{a_k | k \geq n\}) = \lim_n(-\inf\{-a_k | k \geq n\}) = -\liminf\{-a_n\}$；
3. 若 $\lim_n a_n = a$，那么对于任意正数 ϵ，存在 $n_0 \in \mathbb{N}$ 使得 $|a_n - a| < \epsilon$，即
$$a - \epsilon < a_n \leq \sup\{a_k | k \geq n\} \leq a + \epsilon \implies |\sup\{a_k | k \geq n\} - a| \leq \epsilon. \tag{1.4.16}$$

因此 $\limsup\{a_n\}=a$，类似地 $\liminf\{a_n\}=a$；反之若(1.4.14)式成立，对于任意正数 ϵ，存在自然数 n_1 和 n_2，使得 $|\sup\{a_k|k\geq n\}-a|<\epsilon$ 对所有 $n\geq n_1$ 成立、$|\inf\{a_k|k\geq n\}-a|<\epsilon$ 对所有 $n\geq n_2$ 成立。因此取 $n_0=\max\{n_1,n_2\}$，可知

$$a-\epsilon<\inf\{a_k|k\geq n\}\leq a_n\leq\sup\{a_k|k\geq n\}<a+\epsilon, \tag{1.4.17}$$

进而 $\lim_n a_n=a$，必要性得证；

4. 由于 $a_k\leq\sup\{b_k|k\geq n\}$ 对所有满足 $k\geq n$ 的自然数 k 和 n 成立，因此 $\sup\{a_k|k\geq n\}\leq\sup\{b_k|k\geq n\}$。利用收敛的单调性（定理 1.15），$\limsup\{a_n\}\leq\limsup\{b_n\}$。

利用上述基本性质，我们可得到上极限和下极限的另一种等价定义。

命题 1.26

设 $\{a_n\}$ 为实数列，则

$$\limsup\{a_n\}=\inf_n(\sup\{a_k|k\geq n\}), \tag{1.4.18}$$

$$\liminf\{a_n\}=\sup_n(\inf\{a_k|k\geq n\}). \tag{1.4.19}$$

♠

证明 若 $\limsup\{a_n\}=a$，则对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $|\sup\{a_k|k\geq n\}-a|<\epsilon$ 对所有 $n\geq n_0$ 成立。由于 $\sup\{a_k|k\geq n\}$ 是递减列，$\sup\{a_k|k\geq n\}>a-\epsilon$ 对所有自然数 n 成立，那么 $\inf_n(\sup\{a_k|k\geq n\})\geq a-\epsilon$ 对任意正数 ϵ 成立，从而 $\inf_n(\sup\{a_k|k\geq n\})\geq a$；另一方面，显然有 $\inf_n(\sup\{a_k|k\geq n\})\leq\sup\{a_k|k\geq n\}<a+\epsilon$ 对任意正数 ϵ 成立，进而 $\inf_n(\sup\{a_k|k\geq n\})\leq a$。综上(1.4.18)式得证。对于后者，利用命题 1.25性质 2 易证。

1.4.6 实变量的连续实值函数

我们已在第1.3.4小节接触了连续函数在度量空间内的一般性质。此处仅补充连续实值函数特有的一条结论。

定理 1.16 (介值定理)

设 f 为有界闭区间 $[a,b]$ 上使得 $f(a)<c<b$ 的连续实值函数，则存在 (a,b) 中的点 x_0 使得 $f(x_0)=c$。

♡

证明 这里采用二分法的思想：记 $a_1=a$ 且 $b_1=b$，取 $[a_1,b_1]$ 的中点 m_1。若 $c<f(m_1)$，则记 $a_2=a_1$ 且 $b_2=m_1$；若 $c\geq f(m_1)$，则记 $a_2=m_1$ 且 $b_2=b_1$。因此 $f(a_2)\leq c\leq f(b_2)$，并且 $b_2-a_2=(b_1-a_1)/2$。归纳地继续此二分过程，进而得到递减闭区间族 $\{[a_n,b_n]\}_{n=1}^{\infty}$，使得

$$f(a_n)\leq c\leq f(b_n)\text{ and }b_n-a_n=\frac{b-a}{2^{n-1}}\text{ for all }n. \tag{1.4.20}$$

利用闭区间套定理 1.11，存在实数 $x_0\in\cap_{n=1}^{\infty}[a_n,b_n]$，并且

$$|a_n-x_0|\leq|b_n-a_n|=\frac{b-a}{2^{n-1}}\text{ for all }n. \tag{1.4.21}$$

因此 $\lim_n a_n=x_0$，类似地 $\lim_n b_n=x_0$。根据命题 1.16和定理 1.15，$f(x_0)\leq c\leq f(x_0)$，即存在 $x_0\in[a,b]$ 使得 $f(x_0)=c$。

第一部分

高等线性代数

第 2 章 向量空间

目录

2.1 向量空间及其子空间 22
2.2 生成集与线性无关 23
2.3 向量空间的维数 25
2.4 外直和与内直和 26

2.1 向量空间及其子空间

本节我们首先介绍向量空间的概念，并引出其子结构。

定义 2.1 (向量空间)

设 F 为域，其元素称为标量（scalar）。F 上的向量空间（vector space）V 是一个非空集合，其元素称为向量（vector）。在该集合上定义了两种运算：第一种运算称为加法（addition），用“$+$”表示，该运算为 V 中每一对向量 (u,v) 赋予 V 中的一个向量 $u+v$；第二种运算称为标量乘法（scalar multiplication），用并置表示，该运算为每一对 $(r,u)\in F\times V$ 赋予 V 中的一个向量 ru。此外，上述运算还需要满足下列性质。

1. V 在加法下构成 Abelian 群；
2. 对于任意 $a,b\in F$ 和任意 $u,v\in V$，标量乘法满足
   (1) 分配律：$a(u+v)=au+av$ 且 $(a+b)u=au+bu$；
   (2) 结合律：$(ab)u=a(bu)$；
   (3) 单位元存在性：存在向量 $1\in V$ 使得 $1u=u$。

♣

注 在域 F 上的向量空间有时也称为F-空间（*F-space*），在实数域上的向量空间称为实向量空间（*real vector space*），在复数域上的向量空间则称为复向量空间（*complex vector space*）。

定义 2.2 (线性组合)

设 S 为向量空间的一个非空子集，则 S 上向量的线性组合（linear combination）具有形式 $a_1v_1+\cdots+a_nv_n$，其中 $v_1,\cdots,v_n\in S$ 且 $a_1,\cdots,a_n\in F$。标量 a_i 称为线性组合的系数（coefficient）。若每个系数 a_i 均为零，则称该线性组合平凡（trivial），否则称其非平凡（nontrivial）。

♣

注 需要注意的是，线性组合在定义上天然要求参与求和的项数有限，因此级数展开并不属于线性组合。

例 2.1

1. 设 F 为域，则所有从 F 到 F 的函数集合 F^F 是 F 上的一个向量空间，其加法和标量乘法分别定义为
$$(f+g)(x)=f(x)+g(x)\quad\text{and}\quad(af)(x)=af(x). \tag{2.1.1}$$
2. 所有元素取自 F 的 $m\times n$ 矩阵集合 $\mathcal{M}_{m,n}(F)$ 构成 F 上的一个向量空间，其加法和标量乘法为通常的矩阵加法和标量乘法；
3. 所有元素取自 F 的 n 元组 F^n 构成 F 上的一个向量空间，其加法和标量乘法按分量定义：
$$(a_1,\cdots,a_n)+(b_1,\cdots,b_n)=(a_1+b_1,\cdots,a_n+b_n)\quad\text{and}\quad c(a_1,\cdots,a_n)=(ca_1,\cdots,ca_n). \tag{2.1.2}$$
为了简便，我们也用列向量的形式写出 F^n 中的元素；
4. 所有元素取自域 F 的无穷序列构成的集合是一个向量空间，其运算也按分量定义
$$(s_n)+(t_n)=(s_n+t_n)\quad\text{and}\quad a(s_n)=(as_n). \tag{2.1.3}$$

类似地，所有收敛到 0 的复数列构成的集合、所有有界复数列构成的集合、所有满足 $\sum_{n=1}^{\infty}|s_n|^p<\infty$（$p$ 为正整数）的复数列 (s_n) 构成的集合，均构成一个向量空间。

定义 2.3 (子空间)

设 S 是 V 的一个子集，若将 V 上的运算限制在 S 上，并且 S 在该运算下构成向量空间，则称 S 是 V 的一个子空间（subspace）。通常用 $S\leq V$ 表示 S 是 V 的子空间，而用 $S<V$ 表示 S 是 V 的真子空间（proper subspace），即 $S\leq V$ 但 $S\neq V$。V 的零子空间（zero subspace）为 $\{0\}$。

由于加法与标量乘法的性质在非空子集 S 中必然成立，因此可通过 S 在 V 的运算下是否封闭确定其是否构成一个子空间。我们有如下结果。

命题 2.1

向量空间 V 的非空子集 S 构成 V 的一个子空间当且仅当 S 在加法和标量乘法下封闭，或者等价地，S 在线性组合下封闭，即

$$a,b\in F \text{ and } u,v\in S \quad\Longrightarrow\quad au+bv\in S. \tag{2.1.4}$$

向量空间 V 的所有子空间构成的集合 $\mathcal{S}(V)$ 在集合包含意义下构成偏序集。零空间 $\{0\}$ 是 $\mathcal{S}(V)$ 中最小的元素，而全空间 V 是最大的元素。若 $S,T\in\mathcal{S}(V)$，则 $S\cap T$ 是 V 中同时包含于 S 和 T 的最大子空间。就集合包含而言，$S\cap T$ 为 S 和 T 的下确界，记作 $S\cap T=\text{glb}\{S,T\}$；类似地，若 $\{S_k|k\in K\}$ 为 V 的任一子空间族，那么其交集 $\cap_{k\in K}S_k=\text{glb}\{S_k|k\in K\}$ 为子空间的下确界。另一方面，若 $S,T\in\mathcal{S}(V)$ 且 F 无限，那么 $S\cup T\in\mathcal{S}(V)$ 当且仅当 $S\subseteq T$ 或 $T\subseteq S$。那么为了确定 V 内包含 S 和 T 的最小子空间，我们引出如下定义。

定义 2.4 (子空间的和)

设 S 和 T 为 V 的子空间，其和定义为 $S+T=\{u+v|u\in S,v\in T\}$。更一般地，任意子空间族 $\{S_k|k\in K\}$ 的和为所有来自并集 $\cup_k S_k$ 的向量的有限和构成的集合，即

$$\sum_{k\in K}S_k=\{s_1+\cdots+s_n|s_j\in\cup_{k\in S}S_k,n\in\mathbb{N}\}, \tag{2.1.5}$$

其中 s_j 是某个子空间内的向量嵌入更大向量空间 $\cup_{k\in K}S_k$ 后的表示。

不难验证 V 内任意子空间族的和为 V 的一个子空间，并且就集合包含而言，该和为上确界，即 $S+T=\text{lub}\{S,T\}$。更一般地，$\sum_{k\in K}S_k=\text{lub}\{S_k|k\in K\}$。若偏序集 P 具有每对元素均具有上确界和下确界的性质，则称 P 为格（*lattice*）。若 P 具有最小元和最大元（存在即唯一），并且每个元素族均有上确界和下确界，则称 P 为完全格（*complete lattice*）。因此我们有如下结论。

命题 2.2 (子空间格)

向量空间 V 的所有子空间构成的集合 $\mathcal{S}(V)$ 在集合包含下构成完全格，其最小元为 $\{0\}$、最大元为 V，即

$$\text{glb}\{S_k|k\in K\}=\bigcap_{k\in K}S_k \quad\text{and}\quad \text{lub}\{S_k|k\in K\}=\sum_{k\in K}S_k. \tag{2.1.6}$$

2.2 生成集与线性无关

定义 2.5 (生成集)

由 V 中一个向量积 S 张成（span）或生成（generate）的子空间为 S 中所有线性组合构成的集合，即

$$\langle S\rangle=\text{spac}(S)=\{r_1v_1+\cdots+r_nv_n|r_i\in F,v_i\in V\}. \tag{2.2.1}$$

当 $S=\{v_1,\cdots,v_n\}$ 是一个有限集时，我们用符号 $\langle v_1,\cdots,v_n\rangle$ 或 $\text{span}(v_1,\cdots,v_n)$ 表示。若 $V=\text{span}(S)$，

则称 V 中的集合 S 张成/生成 V，即每个 $v \in V$ 可写为

$$v = r_1 v_1 + \cdots + r_n v_b, \quad \text{where } r_i \in F, v_i \in S. \tag{2.2.2}$$

注 显然任何生成集的超集（superset）均为生成集，并且所有向量都具有生成集（自身作为生成集）。

定义 2.6 (线性无关与线性相关)

对于 V 中非空向量集 S 中的任意元素 $v_1, \cdots, v_n$，若其满足

$$r_1 v_1 + \cdots + r_n v_n = 0 \quad \Longrightarrow \quad r_1 = \cdots = r_n = 0, \tag{2.2.3}$$

则称 S 线性无关（linearly independent）；反之称其线性相关（linearly dependent）。

注 根据上述定义，线性无关集的任意非空子集也是线性无关的。

命题 2.3

设 S 为 V 内的向量集，则下列表述等价：

1. S 线性无关；
2. $\mathrm{span}(S)$ 中每个向量可唯一表示为 S 内向量的线性组合；
3. S 内没有向量可表示为其余向量的线性组合。

证明

1. 若表述 1 成立，对于任意向量 $v \in \mathrm{span}(S)$，若存在 S 的两组向量和 F 上的两组系数使得

$$v = r_1 u_1 + \cdots + r_n u_n = s_1 v_1 + \cdots + s_m v_m. \tag{2.2.4}$$

当 $v = 0$ 时，则自然所有系数均为零，该平凡表示唯一；反之存在指标 i 和 j 使得 $u_i = t v_j$，其中 $t \in F$ 非零，则将其移到等式右边，合并得到项 $(s_j - t r_i) v_j$。如此操作后得到形如 $\sum_i r_i u_i + \sum_j (-s_j) v_j = 0$ 的等式，其中所有 s_i 和 v_j 均线性无关。因此 $r_i = 0$ 和 $s_j = 0$ 对所有系数成立，从而该表示唯一；

2. 若表述 2 成立，假设存在 $v_i \in S$ 使得 $v_i = \sum_{j \neq i} r_j v_j$ 对某组 $r_j \in F$ 成立，其中 v_j 为 S 内其余向量。那么 $\mathrm{span}(S)$ 内的零向量至少存在两种表示：$\sum_j 0 \cdot v_j$ 和 $-v_i + \sum_{j \neq i} r_j v_j$，矛盾。因此表述 3 成立；

3. 若表述 3 成立，那么对于 S 内的向量 $v_1, \cdots, v_n$，存在 $r_1, \cdots, r_n \in F$ 使得 $r_1 v_1 + \cdots + r_n v_n = 0$。假设存在某个 $r_i \neq 0$，那么 $v_i = \sum_{j \neq i} (-r_j / r_i) v_j$，这与表述 3 矛盾。因此所有 $r_i = 0$，即 S 线性无关。

有了上述结果，我们可将生成集与线性无关的概念相联系。

定理 2.1

设 S 为 V 内的向量积，则下列表述等价：

1. S 线性无关且张成 V；
2. 对于每个向量 $v \in V$，存在 S 内唯一向量集 $\{v_1, \cdots, v_n\}$ 以及 F 中唯一的标量集 $\{r_1, \cdots, r_n\}$，使得 $v = r_1 v_1 + \cdots + r_n v_n$；
3. S 是极小生成集（minimal spanning set），即 S 可张成 V，但任意 S 的真子集均无法张成 V；
4. S 是极大线性无关集（maximal linearly independent set），即 S 线性无关，但 S 的真超集线性相关。

证明 根据命题 2.3，显然表述 1 和 2 等价。

1. 若表述 1 成立，已知 S 可张成 V，不妨假设存在 $v \in S$，使得 $\mathrm{span}(S \backslash \{v\}) = V$ 且 $S \backslash \{v\}$ 线性无关。那么对于 $v \in V$ 而言，存在 $S \backslash \{v\}$ 内唯一向量集 $\{v_1, \cdots, v_n\}$ 和 F 上唯一标量集 $\{r_1, \cdots, r_n\}$，使得 $v = \sum_{i=1}^n r_i v_i$。但这与 S 线性无关矛盾，故任意 S 的真子集无法张成 V，表述 3 成立；反之若存在 S 内的向量 v_i 可表示为其余向量的线性组合，记为 $\sum_{j \neq i} r_j v_j$，那么对于任意 $v \in V$，存在 $s_j \in F$ 使得

$$v = \sum_j s_j v_j = s_i \sum_{j \neq i} r_j v_j + \sum_{j \neq i} s_j v_j = \sum_{j \neq i} (s_i r_j + s_j) v_j \in \mathrm{span}(S \backslash \{v_i\}), \tag{2.2.5}$$

这与条件矛盾，因此 S 线性无关；

2. 若表述 1 成立，则对于任意 $v \in V$，存在 $r_i \in F$ 使得 $v = \sum_i r_i v_i$，其中 $v_i \in S$。可见 $S \cup \{v\}$ 线性相关，即 S 的任意真超集线性相关；反之若存在向量 $v \in V$ 不属于 $\mathrm{span}(S)$，那么 $S \cup \{v\}$ 也是线性无关集，矛盾。因此表述 1 成立。

根据此结果，我们可以定义向量空间的基。

定义 2.7 (基)

向量空间 V 中满足定理 2.1的任一条件的向量集称为 V 的基（basis）。 ♣

例 2.2 F^n 中的第 i 个标准向量（*standard vector*）时除了第 i 个坐标为 1、其余坐标为 0 的向量 e_i，例如 $e_1 = (1, 0, \cdots, 0)$。集合 $\{e_1, \cdots, e_n\}$ 称为 S 的标准基（*standard basis*）。

作为 Zorn 引理的直接结论，每个向量空间都有一组基。

定理 2.2

设 V 为非空向量空间，I 为 V 中的线性无关集，S 为包含 I 的一个生成集，则存在 V 的一组基 $\mathcal{B}$ 使得 $I \subseteq \mathcal{B} \subseteq S$。特别地，

1. 任意向量空间（除去零空间）具有一组基；
2. V 中任意线性无关集包含在一组基内；
3. V 中任意生成集包含一组基。 ♡

证明 记 V 中包含 I 且包含于 S 的所有线性无关子集构成的集族为 $\mathcal{A}$。由于 $I \in \mathcal{A}$，故 $\mathcal{A}$ 非空。若 $\mathcal{C} = \{I_k | k \in K\}$ 为 $\mathcal{A}$ 内的一条链，则其并集 $U = \cup_{k \in K} I_k \in \mathcal{A}$，否则存在指标 k 使得 $I_k \notin \mathcal{A}$，这与条件矛盾。根据 Zorn 引理（定理 1.1），$\mathcal{A}$ 内的每条链都具有属于 $\mathcal{A}$ 的一个上界 U，因此 $\mathcal{A}$ 必然包含一个极大元 $\mathcal{B}$，且 $\mathcal{B}$ 同样线性无关，因此 $I \subseteq \mathcal{B} \subseteq S$。此时仍需证明[1] $\mathcal{B}$ 可张成 V。假设存在向量 $v \in V$ 使得 $v \notin \langle \mathcal{B} \rangle$，根据命题 2.3，$\mathcal{B} \cup \{v\}$ 线性无关，但这与 $\mathcal{B}$ 是极大元的事实矛盾。因此 $\langle \mathcal{B} \rangle = V$，由定理 2.1，$\mathcal{B}$ 是 V 内一组基。任意向量空间内任意向量自身构成一个线性无关集，同时全空间构成一个生成集，从而该空间总是具有一组基；剩余结论显然。

2.3 向量空间的维数

有了基的概念之后，几个很自然的问题是：一个向量空间内是否存在不止一组基？这些基之间具有什么关系？为了解答这些疑问，我们介绍一个引理。

引理 2.1

设 V 为向量空间，向量 $v_1, \cdots, v_n$ 线性无关且 $s_1, \cdots, s_m$ 张成 V，则 $n \leq m$。 ♡

证明 我们将 v_1 移至生成集内，则 v_1 可表示为 $s_1, \cdots, s_m$ 的线性组合，其中至少一个 s_i 的系数为零。不妨设该元素为 s_1（若不为 s_1，则将元素重排），那么可将 s_1 剔除，替换为 v_1，从而构成新的生成集 $\{v_1, s_2, \cdots, s_m\}$。再将 v_2 移至新的生成集内，则 v_2 可表示为 $v_1, s_2, \cdots, s_m$ 的线性组合，其中至少一个 s_j 的系数为零，否则 $v_2 = r v_1$ 对某个 $r \in F$ 成立。设 s_j 为 v_2，将 s_2 剔除并替换为 v_2，得到新的生成集 $\{v_1, v_2, s_3, \cdots, s_m\}$，依此类推。若 $n > m$，则最终所有 s 耗尽，得到生成集 $\{v_1, \cdots, v_m\}$ 和剩余线性无关集 $\{v_{m+1}, \cdots, v_n\}$，但线性无关性要求 $v_n \notin \langle v_1, \cdots, v_m \rangle$（命题 2.3），这与生成集的要求矛盾，故 $n \leq m$。

定理 2.3

若 V 为向量空间，则 V 中任意两组基具有相同的基数。 ♡

证明 对于有限生成的 V 及其两组基 $\mathcal{B}_1$ 和 $\mathcal{B}_2$，引理 2.1要求 $|\mathcal{B}_1| \leq |\mathcal{B}_2|$ 且 $|\mathcal{B}_2| \leq |\mathcal{B}_1|$，进而 $|\mathcal{B}_1| = |\mathcal{B}_2|$。

[1] 由于 $\mathcal{B}$ 是包含于 S 的极大线性无关集，不能直接采用定理 2.1的结论

对于无穷集，设 $\mathcal{B}=\{b_i|i\in I\}$ 为 V 的一组基，而 $\mathcal{C}$ 为另一组基，则其中任一元素 c 可表示为 $\mathcal{B}$ 中向量的线性组合，即存在有限指标集 $U_c\subseteq K$ 和 $r_i\in F$ 使得 $c=\sum_{i\in U_c}r_ib_i$。显然 $\cup_{c\in\mathcal{C}}U_c\subseteq I$，不妨假设存在指标 $i\notin\cup_{c\in\mathcal{C}}U_c$，那么 $\mathcal{C}$ 中的元素均可表示为 $\mathcal{B}\backslash\{b_i\}$ 中元素的线性组合，进而 $V=\langle\mathcal{C}\rangle\subseteq\langle\mathcal{B}\backslash\{b_i\}\rangle$。但 $b_i\notin\langle\mathcal{B}\backslash\{b_i\}\rangle=V$，矛盾。因此 $\cup_{c\in\mathcal{C}}U_c=I$。由于 c 可表示为 $\mathcal{B}$ 内向量的线性组合，故 $|U_c|<\aleph_0$。根据命题 1.4和1.6，

$$|\mathcal{B}|=|I|=|\cup_{c\in\mathcal{C}}U_c|\leq\aleph_0\cdot|\mathcal{C}|=\max\{\aleph_0,|\mathcal{C}|\}=|\mathcal{C}|, \tag{2.3.1}$$

其中已利用 $\aleph_0$ 是最小无穷势的事实。类似地，$|\mathcal{C}|\leq|\mathcal{B}|$，进而 $|\mathcal{B}|=|\mathcal{C}|$。

由此我们可以正式引入向量空间的维数的定义。

定义 2.8 (向量空间的维数)

若向量空间 V 为零空间或具有有限个数的基，则称其有限维（finite-dimensional）；其余所有向量空间都是无穷维的（infinite-dimensional）。零空间的维数是 0，任意非零向量空间 V 的维数定义为 V 中任一组基的基数。换言之，若向量空间 V 有一组基数为 κ 的基，则称 V 是 κ 维的（κ-dimensional），记作 $\dim(V)=\kappa$。♣

即便有了维数的概念，有时在一组基下的表示在重排的意义下也是不唯一的，由此引出如下概念。

定义 2.9 (有序基)

设 V 为 n 维向量空间，则 V 的一组有序基（ordered basis）为有序 n 元向量组 $(v_1,\cdots,v_n)$，其中集合 $\{v_1,\cdots,v_n\}$ 为 V 的一组基。♣

若 $\mathcal{B}=(v_1,\cdots,v_n)$ 为 V 的一组有序基，那么对于每个 $v\in V$，总存在唯一有序 n 元标量组 $(r_1,\cdots,r_n)$，满足 $v=r_1v_1+\cdots r_nv_n$。据此我们可定义坐标映射（*coordinate map*）

$$\phi_{\mathcal{B}}(v):V\to F^n \text{ for which } \phi_{\mathcal{B}}(v)=[v]_{\mathcal{B}}=(r_1,\cdots,r_n)^T, \tag{2.3.2}$$

其中列向量 $[v]_{\mathcal{B}}$ 是熟知的 v 关于有序基 $\mathcal{B}$ 的坐标矩阵（*coordinate matrix*）。显然，知道 $[v]_{\mathcal{B}}$ 等价于知道了 v。

2.4 外直和与内直和

在本节中，我们会介绍一种从旧向量空间构建新向量空间的方式：直和。张量积我们将放在附录第**??**章介绍。

2.4.1 外直和

定义 2.10 (外直和)

设 $V_1,\cdots,V_n$ 为域 F 上的向量空间，则 $V_1,\cdots,V_n$ 的外直和（external direct product）$V_1\boxplus\cdots\boxplus V_n$ 构成一个新的向量空间 V，其元素为有序 n 元组：

$$V=\{(v_1,\cdots,v_n)|v_i\in V_i,\ i=1,\cdots,n\}, \tag{2.4.1}$$

其按分量定义的运算为

$$(u_1,\cdots,u_n)+(v_1,\cdots,v_n)=(u_1+v_1,\cdots,u_n+v_n)\quad\text{and}\quad r(v_1,\cdots,v_n)=(rv_1,\cdots,rv_n). \tag{2.4.2}$$

♣

例 2.3 向量空间 F^n 为 n 个副本的外直和，即 $F^n=\underbrace{F\boxplus\cdots\boxplus F}_{n\text{ terms}}$。

定义 2.11 (直积)

设 $\mathcal{F}=\{V_k|k\in K\}$ 为域 F 上的向量空间族，则 $\mathcal{F}$ 的直积（direct sum）为向量空间

$$\prod_{k\in K}V_k=\{f:K\to\cup_{k\in K}V_k|f(k)\in V_k\},\tag{2.4.3}$$

其被视为所有从 K 到 $\cup_k V_k$ 的函数构成的向量空间的子空间。

♣

有了直积的概念，我们就可以引入支撑集的概念。

定义 2.12 (支撑集)

设 $\mathcal{F}=\{V_k|k\in K\}$ 为域 F 上的向量空间族，则函数 $f:K\to\cup_k V_k$ 的支撑集（support）定义为集合 $\mathrm{supp}(f)=\{k\in K|f(k)\neq 0\}$。若函数 f 除去有限个指标 $k\in K$ 外均取值为 0，则称 f 具有有限支撑（finite support）。族 $\mathcal{F}$ 的外直和为向量空间

$$\bigoplus_{k\in K}V_k=\{f:K\to\cup_{k\in K}V_k|f(k)\in V_k,\ f\text{ has finite support}\},\tag{2.4.4}$$

其可视为所有从 K 到 $\cup_k V_k$ 的函数构成的向量空间的子空间。

♣

注 对于向量空间的有限族（即指标集 K 有限），直积与外直和相同。

2.4.2 内直和

定义 2.13 (内直和)

设 V 为向量空间，对于 V 的子空间族 $\mathcal{F}=\{S_k|k\in K\}$，若 V 内的每个向量均可以唯一（不考虑顺序）写为 $\mathcal{F}$ 中子空间内向量的有限和，则称 V 是 $\mathcal{F}$ 的内直和（internal dircet sum），简称为直和。换言之，对于任意 V 内的向量 v，若存在 $u_k,w_k\in S_k$ 使得 $v=u_1+\cdots+u_n=w_1+\cdots+w_m$，则重排后 $m=n$ 且 $w_i=u_i$ 对所有 $i=1,\cdots,n$ 成立。若 V 是 $\mathcal{F}$ 的直和，那么记 $V=\oplus_{k\in K}S_k$，并称每个 S_k 为 V 的直和项（direct summand）。若 $\mathcal{F}=\{S_1,\cdots,S_n\}$ 为有限族，记 $V=S_1\oplus\cdots\oplus S_n$。若 $V=S\oplus T$，则 T 称为 V 中 S 的补空间（complement）。

♣

注 后面我们将证明内直和与外直和本质上是等价（同构）的。出于这个考虑，我们不加限定地采用术语“直和”。

定理 2.4

任意向量空间的子空间都具有补空间，即如果 S 为 V 的子空间，则存在子空间 T 使得 $V=S\oplus T$。

♡

证明 设 S 的一组基为 $\mathcal{B}_1=\{b_i|i\in I\}$，那么根据定理 2.2，存在 V 中包含 $\mathcal{B}_1$ 的一组基 $\mathcal{B}$，记 $\mathcal{B}_2=\mathcal{B}\backslash\mathcal{B}_1=\{b_j|j\in J\}$ 且 $T=\langle\mathcal{B}_2\rangle$。首先有 $V=\langle\mathcal{B}\rangle=\langle\mathcal{B}_1\cup\mathcal{B}_2\rangle=\langle\mathcal{B}_1\rangle+\langle\mathcal{B}_2\rangle\subseteq V$，因此 $V=\langle\mathcal{B}_1\rangle+\langle\mathcal{B}_2\rangle$。对于向量 $v\in\langle\mathcal{B}_1\rangle\cap\langle\mathcal{B}_2\rangle$，存在 $r_i,r_j\in F$，使得 $v=\sum_{i\in I}r_ib_i=\sum_{j\in J}(-r_j)b_j$，即 $\sum_{i\in I\cup J}r_ib_i=0$。由于 $\mathcal{B}=\{b_i|i\in I\cup J\}$ 是 V 的基，故 $r_i=0$ 对所有 $i\in I\cup J$ 成立，进而 $\langle\mathcal{B}_1\rangle\cap\langle\mathcal{B}_2\rangle=\{0\}$。综上，存在子空间 $T=\langle\mathcal{B}_2\rangle$，使得 $V=S\oplus T$。

定理 2.5

向量空间 V 可表示为子空间族 $\mathcal{F}=\{S_i|i\in I\}$ 的直和当且仅当

1. V 为 S_i 的和：$V=\sum_{i\in I}S_i$；
2. 对于每个 $i\in I$，$S_i\cap(\sum_{j\neq i}S_j)=\{0\}$。

♡

证明 根据直和的定义，充分性是显然的；反之对于必要性，$\mathcal{F}$ 中任意两个不同的子空间 S_i 和 S_j 满足 $S_i\cap S_j\subseteq S_i\cap(\sum_{j\neq i}S_j)=\{0\}$，因此 $S_i\cap S_j=\{0\}$，结论得证。

在本节的最后，我们给出几条与向量空间的基与维数的结论。

定理 2.6

设 V 为向量空间。

1. 若 $\mathcal{B}$ 为 V 的一组基，同时 $\mathcal{B}=\mathcal{B}_1\cup\mathcal{B}_2$ 且 $\mathcal{B}_1\cap\mathcal{B}_2=\varnothing$，那么 $V=\langle\mathcal{B}_1\rangle\oplus\langle\mathcal{B}_2\rangle$；
2. 设 $V=S\oplus T$，若 $\mathcal{B}_1$ 和 $\mathcal{B}_2$ 分别为 S 和 T 的一组基，那么 $\mathcal{B}_1\cap\mathcal{B}_2=\varnothing$，并且 $\mathcal{B}=\mathcal{B}_1\cup\mathcal{B}_2$ 构成 V 的一组基。

♡

证明

1. 设 $\mathcal{B}_1=\{b_i|i\in I\},\mathcal{B}_2=\{b_j|j\in J\}$，其中 $I\cap J=\varnothing$。那么对于任意 $v\in V$，存在 $r_i\in I$，使得

$$v=\sum_{i=1}^{m+n}r_ib_i=\sum_{i=1}^{m}r_ib_i+\sum_{j=m+1}^{m+n}r_jb_j\quad\Longrightarrow\quad V\subseteq\langle\mathcal{B}_1\rangle+\langle\mathcal{B}_2\rangle\subseteq V\quad\Longrightarrow\quad V=\langle\mathcal{B}_1\rangle+\langle\mathcal{B}_2\rangle. \tag{2.4.5}$$

另外，考虑 $u\in\langle\mathcal{B}_1\rangle\cap\langle\mathcal{B}_2\rangle$，存在 $s_i,s_j\in F$ 使得 $u=\sum_{i=1}^{m}s_ib_i=\sum_{j=m+1}^{m+n}(-s_j)b_j$，即 $\sum_{i=1}^{m+n}s_ib_i=0$。因此 $s_i=0$ 对所有指标 i 成立，即 $\langle\mathcal{B}_1\rangle\cap\langle\mathcal{B}_2\rangle=\{0\}$；

2. 同上设 $\mathcal{B}_1$ 和 $\mathcal{B}_2$，假设存在 $v\in\mathcal{B}_1\cap\mathcal{B}_2$，那么 $v\in\langle\mathcal{B}_1\rangle\cap\langle\mathcal{B}_2\rangle=S\cap T=\{0\}$，但零向量与其余任意向量均线性相关，这与 $\mathcal{B}_1$ 和 $\mathcal{B}_2$ 为基矛盾，因此 $\mathcal{B}_1\cap\mathcal{B}_2=\varnothing$。另外，对于任意 $v\in V$，存在 $s\in S$ 和 $t\in T$，使得 $v=s+t\in\langle\mathcal{B}_1\rangle+\langle\mathcal{B}_2\rangle=\langle\mathcal{B}_1\cup\mathcal{B}_2\rangle$，即 $\mathcal{B}=\mathcal{B}_1\cup\mathcal{B}_2$ 生成 V。假设 $\mathcal{B}$ 线性相关，即存在 $b_i\in\mathcal{B}$ 和 $r_j\in F$，使得 $b_i=\sum_{j\neq i}r_jb_j$。若 $b_i\in\mathcal{B}_1$，设该表达式中分别属于 $\mathcal{B}_1$ 和 $\mathcal{B}_2$ 的向量对应的指标集分别为 I' 和 J'，那么 $\sum_{j\in J'}r_jb_j=b_i-\sum_{i\in I'}r_ib_i\in S$，因此 $r_j=0$ 对所有 $j\in J'$ 成立。此时 $b_i=\sum_{i\in I'}r_ib_i$，这与 $\mathcal{B}_1$ 为基的条件矛盾；若 $b_i\in\mathcal{B}_2$，类似分析可知其与 $\mathcal{B}_2$ 为基他的条件矛盾。因此 $\mathcal{B}$ 线性无关，根据定理 2.1，$\mathcal{B}$ 为 V 的一组基。

推论 2.1

V 中的有限向量集 $S=\{v_1,\cdots,v_n\}$ 为基当且仅当 $V=\langle v_1\rangle\oplus\cdots\oplus\langle v_n\rangle$。

♡

证明 充分性是显然的；对于必要性，首先设 $S_1=\langle v_1\rangle,T_1=\langle v_2\rangle\oplus\cdots\oplus\langle v_n\rangle$，那么 $V=S_1\oplus T_1$ 且 v_1 为 S 的基。若存在非零 v_i $(i\geq 2)$ 和 $r_j\in F$，使得 $v_i=\sum_{j\neq i}r_jv_j$，那么至少存在一个 $r_j\neq 0$，进而 $\langle v_i\rangle\subseteq\langle v_j\rangle$，即 $\{0\}\subseteq\langle v_i\rangle=\langle v_i\rangle\cap\langle v_j\rangle$，矛盾。因此 $v_2,\cdots,v_n$ 线性无关，构成 T 的一组基。利用上述定理，$v_1,\cdots,v_n$ 构成 V 的一组基。

定理 2.7

设 S 和 T 为向量空间的子空间，则

$$\dim(S)+\dim(T)=\dim(S+T)+\dim(S\cap T). \tag{2.4.6}$$

特别地，若 T 为 S 在 V 中的补空间，则 $\dim(S)+\dim(T)=\dim(V)=\dim(S\oplus T)$。

♡

证明 设 $\mathcal{B}=\{b_i|i\in I\}$ 为 $S\cap T$ 的一组基。根据定理 2.2，可将 $\mathcal{B}$ 分别扩张为子空间 S 和 T 内的基 $\mathcal{A}\cup\mathcal{B}$ 和 $\mathcal{B}\cup\mathcal{C}$，其中 $\mathcal{A}=\{a_j|j\in J\}$ 和 $\mathcal{C}=\{c_k|k\in K\}$ 与 $\mathcal{B}$ 均不相交。显然有 $\langle\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}\rangle=\langle\mathcal{A}\cup\mathcal{B}\rangle+\langle\mathcal{B}\cup\mathcal{C}\rangle=S+T$。考虑线性组合

$$\sum_{j\in J}r_ja_j+\sum_{i\in I}r_ib_i+\sum_{k\in K}r_kc_k=0\quad\Longrightarrow\quad\sum_{j\in J}r_ja_j=\sum_{i\in I}r_ib_i-\sum_{k\in K}r_kc_k\in\langle\mathcal{B}\cup\mathcal{C}\rangle=T, \tag{2.4.7}$$

同时 $x=\sum_{j\in J}r_ja_j\in\langle\mathcal{A}\rangle\subseteq S$，因此 $x\in S\cap T=\langle\mathcal{B}\rangle$。根据定理 2.6，$\langle\mathcal{A}\rangle\cap\langle\mathcal{B}\rangle=\{0\}$，从而 $x=\sum_{j\in J}r_ja_j=0$。利用 $\mathcal{A}$ 的线性无关性，$r_j=0$ 对所有 $j\in J$ 成立。类似地，$r_k=0$ 对所有 $k\in K$ 成立，进而 $r_i=0$ 对所有 $i\in I$ 成立。可见 $\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}$ 线性无关。根据定理 2.2，$\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}$ 是 $S+T$ 的一组基。结合基数算术的定义，

$$\begin{aligned}\dim(S)+\dim(T)&=|\mathcal{A}\cup\mathcal{B}|+|\mathcal{B}\cup\mathcal{C}|=|\mathcal{A}|+2|\mathcal{B}|+|\mathcal{C}|\\&=|\mathcal{A}\cup\mathcal{B}\cup\mathcal{C}|+|\mathcal{B}|=\dim(S+T)+\dim(S\cap T).\end{aligned} \tag{2.4.8}$$

注 需要强调的是，只有当 $S+T$ 维数有限时，上述结论才可写为 $\dim(S+T)=\dim(S)+\dim(T)-\dim(S\cap T)$，因为无穷势的减法未定义。

第 3 章 线性变换

目录

3.1 线性变换及其核与像 29
3.2 同构与秩-零化度定理 31
3.3 矩阵变换与基变换 32
3.4 等价性与相似性 34
3.5 不变子空间与约化对 35
3.6 投影算子 36

3.1 线性变换及其核与像

定义 3.1 (线性变换)

设 V 和 W 是定义在域 F 上的向量空间，若对于任意 $r,s\in F$ 和向量 $u,v\in V$，函数 $\tau:V\to W$ 满足

$$\tau(ru+sv)=r\tau(u)+s\tau(v), \tag{3.1.1}$$

则称其为线性变换（linear transformation）；线性变换 $\tau:V\to V$ 称为 V 上的线性算子（linear operator）。所有从 V 到 W 的线性变换构成的集合记为[a] $\mathcal{L}(V,W)$；所有 V 上线性算子的集合记为 $\mathcal{L}(V)$。

[a] 这一部分仅讨论代数结构自身的性质，暂且不考虑其拓扑性质。在第14章中，我们会将此处的 $\mathcal{L}(V,W)$ 记为 $\mathrm{Hom}(V,W)$，而将其中连续线性变换构成的集合记为 $\mathcal{L}(V,W)$。

以下术语也经常被使用。

定义 3.2

1. 同态（homomorphism）对应线性变换；
2. 自同态（endomorphism）对应线性算子；
3. 单同态（monomorphism）或嵌入（embedding）对应单射线性变换；
4. 满同态（epimorphism）对应满射线性变换；
5. 同构（isomorphism）对应双射线性变换；
6. 自同构（automorphism）对应双射线性算子。

例 3.1

1. 微分算符 $D:V\to V$ 是 $\mathbb{R}$ 上所有无穷可微函数构成的向量空间 V 上的线性算子；
2. 设矩阵 $A\in\mathcal{M}_{m,n}(F)$，函数 $\tau_A:F^n\to F^m$ 定义为 $\tau_A(v)=Av$，其中所有向量 v 均写成列向量，那么 τ_A 是从 F^n 到 F^m 的线性变换；
3. n 维向量空间的坐标映射 $\phi:V\to F^n$ 为从 V 到 F^n 的线性变换。

命题 3.1

1. 集合 $\mathcal{L}(V,W)$ 在普通的函数加法和函数关于 F 上元素的标量乘法下构成一个向量空间；
2. 若 $\sigma\in\mathcal{L}(U,V),\tau\in\mathcal{L}(V,W)$，则复合（composition）$\tau\sigma\in\mathcal{L}(U,W)$；
3. 若 $\tau\in\mathcal{L}(V,W)$ 为双射，则 $\tau^{-1}\in\mathcal{L}(V,W)$；
4. 向量空间 $\mathcal{L}(V)$ 是一个代数[a]，其乘法为函数复合，单位映射 $\iota\in\mathcal{L}(V)$ 为乘法单位元，零映射 $0\in\mathcal{L}(V)$ 为加法单位元。

[a]代数 $\mathcal{A}$ 上定义了加法、乘法和标量乘法三种运算，要求其在加法和标量乘法下构成向量空间，在加法和乘法下构成环，并且要求对于任意 $r \in F$ 和 $a, b \in \mathcal{A}$，$r(ab) = (ra)b = a(rb)$。

证明

1. 根据定义容易验证该结论；
2. 对于任意 $u, v \in U$ 和 $r, s \in F$，
$$\begin{aligned}(\tau\sigma)(ru+sv) &= \tau(\sigma(ru+sv)) = \tau(r\sigma(u)+s\sigma(v)) \\ &= r\tau(\sigma(u)) + s\tau(\sigma(v)) = r(\tau\sigma)(u) + s(\tau\sigma)(v),\end{aligned} \tag{3.1.2}$$
因此 $\tau\sigma \in \mathcal{L}(U, W)$；
3. 对于任意 $w_1, w_2 \in W$ 和 $r, s \in F$，存在 $v_1, v_2 \in V$ 使得 $w_1 = \tau(v_1)$ 且 $w_2 = \tau(v_2)$，因此
$$\begin{aligned}\tau^{-1}(rw_1+sw_2) &= \tau^{-1}(r\tau(v_1)+s\tau(v_2)) = \tau^{-1}(\tau(rv_1+sv_2)) \\ &= rv_1 + sv_2 = r\tau^{-1}(w_1) + s\tau^{-1}(v_2),\end{aligned} \tag{3.1.3}$$
因此 $\tau^{-1} \in \mathcal{L}(W, V)$；
4. 根据代数的定义容易证明该结论。

定义线性变换最简单的方式之一是给定其在一组基上的值。下述定理说明，我们可以在基上任意赋值，并通过线性延拓得到整个定义域上唯一的线性变换。

定理 3.1

设 V 和 W 为向量空间，$\mathcal{B} = \{v_i | i \in I\}$ 为 V 的一组基，则可以为所有 $v_i \in \mathcal{B}$ 任意指定 $\tau(v_i) \in W$ 的值，然后利用线性性将 τ 的定义域延拓到 V 上，即
$$\tau(a_1v_1 + \cdots + a_nv_n) = a_1\tau(v_1) + \cdots + a_n\tau(v_n). \tag{3.1.4}$$
该过程唯一定义了一个线性变换；换言之，若 $\tau, \sigma \in \mathcal{L}(V, W)$ 对任意 $v_i \in \mathcal{B}$ 满足 $\tau(v_i) = \sigma(v_i)$，则 $\tau = \sigma$。

证明 根据命题 2.3，对于任意 $u, v \in V$，若 $u = v$，则两者具有相同的表示，即存在 $a_i \in F$ 使得 $u = v = \sum_{i=1}^n a_iv_i$，进而 $\tau(u) = \tau(v)$，即该线性延拓是良定义的。现在假设存在 $\sigma \in \mathcal{L}(V, W)$ 使得 $\tau(v_i) = \sigma(v_i)$ 对所有 $i \in I$ 成立。那么对于任意 $v \in V$，存在 $a_i \in F$ 使得 $v = \sum_{i=1}^n a_iv_i$，进而
$$\tau(v) = \tau\left(\sum_{i=1}^n a_iv_i\right) = \sum_{i=1}^n a_i\tau(v_i) = \sum_{i=1}^n a_i\sigma(v_i) = \sigma\left(\sum_{i=1}^n a_iv_i\right) = \sigma(v). \tag{3.1.5}$$
利用 v 的任意性，$\tau = \sigma$。

注 若 $\tau \in \mathcal{L}(V, W)$ 且 S 为 V 的一个子空间，据此可验证 τ 在 S 上的限制 $\tau|_S$ 为从 S 到 W 的线性变换。

定义 3.3 (线性变换的核与像)

设 $\tau \in \mathcal{L}(V, W)$，子空间 $\ker(\tau) = \{v \in V | \tau(v) = 0\}$ 与 $\mathrm{im}(\tau) = \{\tau(v) | v \in V\}$ 分别称为 τ 的核（kernel）与像（image）。$\dim(\ker(\tau))$ 称为 τ 的零化度（nullity），记作 $\mathrm{null}(\tau)$；$\dim(\mathrm{im}(\tau))$ 称为 τ 的秩[a]，记作 $\mathrm{rk}(\tau)$。

[a]我们尚未证明行秩（行空间的维数）等于列秩（列空间的维数），不过这点容易由初等行变换不改变秩（向量集的线性组合不改变其中线性无关集的基数）这一事实证明，具体可参见 [1]。

注 $\ker(\tau)$ 和 $\mathrm{im}(\tau)$ 分别为 V 和 W 的子空间这一结论，虽然出现在定义中，但容易验证。

命题 3.2

设 $\tau \in \mathcal{L}(V, W)$，则

1. τ 为满射当且仅当 $\mathrm{im}(\tau) = W$；
2. τ 为单射当且仅当 $\ker(\tau) = \{0\}$。

证明 结论 1 恰好与满射的定义相一致；对于结论 2，任取 $u,v\in V$，那么 $\tau(u)=\tau(v)\Leftrightarrow\tau(u-v)=0\Leftrightarrow u-v\in\ker(\tau)$。因此若 $\ker(\tau)=\{0\}$，则 $u=v$，即 τ 为单射；若 τ 为单射，则对于任意 $v\in\ker(\tau)$，$\tau(v)=0=\tau(0)$，因此 $v=0$，即 $\ker(\tau)=\{0\}$。

3.2 同构与秩-零化度定理

定义 3.4 (同构)

双射线性变换 $\tau:V\to W$ 称为从 V 到 W 的同构（isomorphism）。当从 V 到 W 的同构存在时，我们称 V 与 W 同构，记作 $V\cong W$。

例 3.2 设 $\dim(V)=n$，则对于 V 中任意有序基 $\mathcal{B}$，坐标映射 $\phi_{\mathcal{B}}:V\to F^n$ 将每个 V 中的向量映射到其坐标矩阵 $[v]_{\mathcal{B}}\in F^n$，从而 $\phi_{\mathcal{B}}$ 为同构。因此，F 上任意 n 维向量空间同构于 F^n。

正如以下命题所示，同构向量空间具有许多共同的性质。若 $\tau\in\mathcal{L}(V,W)$ 且 $S\subseteq V$，则 $\tau(S)=\{\tau(s)|s\in S\}$。

命题 3.3

设 $\tau\in\mathcal{L}(V,W)$ 为同构，$S\subseteq V$，则

1. S 张成 V 当且仅当 $\tau(S)$ 张成 W；
2. S 在 V 中线性无关当且仅当 $\tau(S)$ 在 W 中线性无关；
3. S 是 V 的基当且仅当 $\tau(S)$ 是 W 的基。

上述结果请自行验证，据此我们直接得到如下结论。

推论 3.1

线性变换 $\tau\in\mathcal{L}(V,W)$ 是同构当且仅当存在 V 的一组基 $\mathcal{B}$，使得 $\tau(\mathcal{B})$ 是 W 的一组基。在这种情况下，τ 将 V 中任意基映射到 W 内的一组基。

推论 3.2

设 V 和 W 为 F 上的向量空间，则 $V\cong W$ 当且仅当 $\dim(V)=\dim(W)$。

在例3.2中，我们看到任意 n 维向量空间都同构于 F^n。现在假设 B 是一个基数为 κ 的集合，设 $(F^B)_0$ 为所有从 B 到 F 且具有有限支撑的函数构成的向量空间。容易证明，对于任意 $b\in B$，函数

$$\delta_b(x)=\begin{cases}1, & \text{if } x=b,\\ 0, & \text{if } x\neq b\end{cases}\tag{3.2.1}$$

构成 $(F^B)_0$ 的一组基，称为标准基（*standard basis*）。此时任意函数 $f\in(F^B)_0$ 可表示为 $\sum_{b\in B}f(b)\delta_b$，故 $\dim((F^B)_0)=|B|=\kappa$。当然，对于任意基数 κ，存在维度为 κ 的向量空间与 $(F^B)_0$ 同构。

定理 3.2

若 n 为自然数，则 F 上任意 n 维向量空间同构于 F^n。若 κ 为任意基数，且 B 是基数为 κ 的集合，那么 F 上任意 κ 维向量空间同构于所有从 B 到 F 的具有有限支撑的函数构成的向量空间 $(F^B)_0$。

有了上述结果，我们可以证明关于秩和零化度之间的一个重要关系，称为秩-零化度定理。

定理 3.3 (秩-零化度定理)

设 $\tau\in\mathcal{L}(V,W)$。

1. $\ker(\tau)$ 的任意补空间同构于 $\mathrm{im}(\tau)$；
2. 秩-零化度定理：$\dim(\ker(\tau))+\dim(\mathrm{im}(\tau))=\dim(V)$，或者 $\mathrm{rk}(\tau)+\mathrm{null}(\tau)=\dim(V)$。

证明 设 $\tau \in \mathcal{L}(V,W)$，根据定理 2.4，任何子空间都具有补空间。不妨记 $V = \ker(\tau) \oplus \ker(\tau)^c$，其中 $\ker(\tau)^c$ 为 V 中 $\ker(\tau)$ 的补空间。对于 τ 的 $\ker(\tau)^c$ 上的限制 $\tau^c : \ker(\tau)^c \to W$，

$$\ker(\tau^c) = \{v \in \ker(\tau)^c | \tau(v) = 0\} = \ker(\tau)^c \cap \ker(\tau) = \{0\}. \tag{3.2.2}$$

根据命题 3.2，τ^c 为单射。对于任意向量 $v \in V$，存在 $u \in \ker(\tau)$ 和 $w \in \ker(\tau)^c$，使得

$$\tau(v) = \tau(u+w) = \tau(u) + \tau(w) = \tau^c(w) \quad \Longrightarrow \quad \mathrm{im}(\tau) \subseteq \mathrm{im}(\tau^c). \tag{3.2.3}$$

根据定义，$\mathrm{im}(\tau^c) \subseteq \mathrm{im}(\tau)$ 是显然的，从而 $\mathrm{im}(\tau^c) = \mathrm{im}(\tau)$。由此可见从 $\ker(\tau)^c$ 到 $\mathrm{im}(\tau^c)$ 的映射 τ^c 既是单射又是满射，故 $\ker(\tau)^c \cong \mathrm{im}(\tau^c) = \mathrm{im}(\tau)$。那么根据定理 2.7，$\dim(V) = \dim(\ker(\tau)) + \dim(\ker(\tau)^c) = \dim(\ker(\tau)) + \dim(\mathrm{im}(\tau))$。

上述定理有一条重要的推论。

推论 3.3

设 $\tau \in \mathcal{L}(V,W)$，其中 $\dim(V) = \dim(W) < \infty$，则 τ 为单射当且仅当其为满射。 ♡

证明 根据命题 3.2，τ 为单射当且仅当其为零空间，当且仅当 $\dim(\ker(\tau)) = 0$。再根据定理 3.3，τ 为单射当且仅当 $\dim(\mathrm{im}(\tau)) = \dim(V) = \dim(W)$。利用定理 3.2，$\tau$ 为单射当且仅当 $\mathrm{im}(\tau)$ 与 W 同构，即 τ 为满射。

3.3 矩阵变换与基变换

3.3.1 从 F^n 到 F^m 的线性变换

在例3.1中，对于域 F 上任意 $m \times n$ 矩阵 A，乘法映射 $\tau_A(v) = Av$ 是一个线性变换。事实上，任意线性变换 $\tau \in \mathcal{L}(F_n, F^m)$ 都具有此形式。设矩阵 A 第 i 列为 $\tau(e_i)$，即 $A = (\tau(e_1)|\cdots|\tau(e_n))$，其中 e_i 为第 i 个标准基，则

$$Ae_i = (\tau(e_1)|\cdots|\tau(e_n))^{(i)} = \tau(e_i), \tag{3.3.1}$$

其中 $A^{(i)}$ 表示 A 的第 i 列，因此 $\tau = \tau_A$。

命题 3.4

1. 设 A 为 F 上的一个 $m \times n$ 矩阵，则 $\tau_A \in \mathcal{L}(F^n, F^m)$；
2. 若 $\tau \in \mathcal{L}(F^n, F^m)$，则 $\tau = \tau_A$，其中 $A = (\tau(e_1)|\cdots|\tau(e_n))$ 称为 τ 的矩阵。 ♠

若 $A \in \mathcal{M}_{m,n}(F)$，则 τ_A 的像为 A 的列向量的线性组合，即 A 的列空间，因此秩-零化度定理 3.3给出

$$\dim(\ker(\tau_A)) + \mathrm{rk}(A) = \dim(F^n) = n. \tag{3.3.2}$$

由此对于矩阵 A，可得到类似命题 3.2的结论。

命题 3.5

设 A 为 F 上的 $m \times n$ 矩阵。

1. $\tau_A : F^n \to F^m$ 为单射当且仅当 $\mathrm{rk}(A) = n$；
2. $\tau_A : F^n \to F^m$ 为满射当且仅当 $\mathrm{rk}(A) = m$。 ♠

证明

1. 根据命题 3.2和(3.3.2)式，τ_A 为单射当且仅当 $\ker(\tau_A) = \{0\}$，当且仅当 $\mathrm{rk}(A) = n$；
2. 根据命题 3.2，τ_A 为满射当且仅当 $\mathrm{im}(\tau_A) = F^m$，当且仅当 $\mathrm{rk}(A) = \dim(\mathrm{im}(\tau_A)) = m$（定理 3.2）。

3.3.2 基变换矩阵

设 $\mathcal{B}=(b_1,\cdots,b_n)$ 和 $\mathcal{C}=(c_1,\cdots,c_n)$ 为向量空间 V 的有序基，一个自然的疑问是，坐标矩阵 $[v]_{\mathcal{B}}$ 和 $[v]_{\mathcal{C}}$ 如何关联起来。将 $[v]_{\mathcal{B}}$ 化为 $[v]_{\mathcal{C}}$ 的映射 $\phi_{\mathcal{B},\mathcal{C}}$ 满足

$$[v]_{\mathcal{C}}=\phi_{\mathcal{C}}(v)=\phi_{\mathcal{C}}\cdot\phi_{\mathcal{B}}^{-1}([v]_{\mathcal{B}})=(\phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1})[v]_{\mathcal{B}},\tag{3.3.3}$$

即 $\phi_{\mathcal{B},\mathcal{C}}=\phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}$，称其为基/坐标变换算子（*change of basis/coordinates operator*）。利用上一小节的讨论，$\phi_{\mathcal{C},\mathcal{B}}$ 为 F^n 上的线性算子，因此其矩阵表示为

$$A=(\phi_{\mathcal{B},\mathcal{C}}(e_1)|\cdots|\phi_{\mathcal{B},\mathcal{C}}(e_n))=(\phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}(e_1)|\cdots|\phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}(e_n))=([b_1]_{\mathcal{C}}|\cdots|[b_n]_{\mathcal{C}}).\tag{3.3.4}$$

以后将此处 A 记为 $M_{\mathcal{B},\mathcal{C}}$，并称其为从 $\mathcal{B}$ 到 $\mathcal{C}$ 的基变换矩阵。

定理 3.4

设 $\mathcal{B}=(b_1,\cdots,b_n)$ 和 $\mathcal{C}$ 为向量空间 V 的两组有序基，则基变换算子 $\phi_{\mathcal{B},\mathcal{C}}=\phi_{\mathcal{C}}\phi_{\mathcal{B}}^{-1}$ 为 F^n 上的自同构，其标准矩阵为

$$M_{\mathcal{B},\mathcal{C}}=([b_1]_{\mathcal{C}},\cdots,[b_n]_{\mathcal{C}}),\tag{3.3.5}$$

进而 $[v]_{\mathcal{C}}=M_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$，并且 $M_{\mathcal{C},\mathcal{B}}=M_{\mathcal{B},\mathcal{C}}^{-1}$。

证明 已知坐标映射 $\phi_{\mathcal{B}}$ 和 $\phi_{\mathcal{C}}$ 均为同构，容易验证 $\phi_{\mathcal{B},\mathcal{C}}$ 为自同构。另外，

$$M_{\mathcal{B},\mathcal{C}}^{-1}=(\phi_{\mathcal{B},\mathcal{C}}^{-1}(e_1),\cdots,\phi_{\mathcal{B},\mathcal{C}}^{-1}(e_n))=(\phi_{\mathcal{C},\mathcal{B}}(e_1),\cdots,\phi_{\mathcal{C},\mathcal{B}}(e_n))=M_{\mathcal{C},\mathcal{B}}.\tag{3.3.6}$$

推论 3.4

若给定以下任意两者：

1. 可逆 $n\times n$ 矩阵 A；
2. F^n 的有序基 $\mathcal{B}$；
3. F^n 的有序基 $\mathcal{C}$；

则第三者可由方程 $A=M_{\mathcal{B},\mathcal{C}}$ 唯一确定。

3.3.3 线性变换的矩阵

设 $\tau:V\to W$ 为线性变换，其中 $\dim(V)=n$ 且 $\dim(W)=m$。设 $\mathcal{B}=(b_1,\cdots,b_n)$ 和 $\mathcal{C}$ 分别为 V 和 W 的有序基，则映射 $\theta:[v]_{\mathcal{B}}\to[\tau(v)]_{\mathcal{C}}$ 为线性变换 $\tau:F^n\to F^m$ 的表示（*representation*）。当然，这一表示取决于有序基 $\mathcal{B}$ 和 $\mathcal{C}$ 的选择。由于 θ 从 F^n 和 F^m 的线性变换，它无非是一个 $m\times n$ 的矩阵 A 的乘法作用，即 $[\tau(v)]_{\mathcal{C}}=A[v]_{\mathcal{B}}$。事实上，考虑到 $[b_i]_{\mathcal{B}}=e_i$，类似(3.3.1)式可如下得到 A 的列：

$$A^{(i)}=Ae_i=A[b_i]_{\mathcal{B}}=[\tau(b_i)]_{\mathcal{C}}.\tag{3.3.7}$$

定理 3.5

设 $\tau\in\mathcal{L}(V,W)$，$\mathcal{B}=(b_1,\cdots,b_n)$ 和 $\mathcal{C}$ 分别为 V 和 W 的有序基，则 τ 可关于 $\mathcal{B}$ 和 $\mathcal{C}$ 表示为矩阵乘法，即 $[\tau(v)]_{\mathcal{C}}=[\tau]_{\mathcal{B},\mathcal{C}}[v]_{\mathcal{B}}$，其中

$$[\tau]_{\mathcal{B},\mathcal{C}}=([\tau(b_1)]_{\mathcal{C}},\cdots,[\tau(b_n)]_{\mathcal{C}})\tag{3.3.8}$$

称为 τ 关于基 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵。当 $V=W$ 且 $\mathcal{B}=\mathcal{C}$ 时，我们将 $[\tau]_{\mathcal{B},\mathcal{B}}$ 简记为 $[\tau]_{\mathcal{B}}$，从而 $[\tau(v)]_{\mathcal{B}}=[\tau]_{\mathcal{B}}[v]_{\mathcal{B}}$。

以下结果说明，采用线性变换或与其表示矩阵（给定有序基 $\mathcal{B}$ 和 $\mathcal{C}$）进行操作是同样方便的。这不仅适用于加法和标量乘法，也适用于矩阵乘法。

定理 3.6

设 V 和 W 为域 F 上的向量空间，其分别包含有序基 $\mathcal{B}=(b_1,\cdots,b_n)$ 和 $\mathcal{C}=(c_1,\cdots,c_m)$。

1. 映射 $\mu:\mathcal{L}(V,W)\to\mathcal{M}_{m,n}(F)$ 定义为 $\mu(\tau)=[\tau]_{\mathcal{B},\mathcal{C}}$，则该映射是同构，因此 $\mathcal{L}(V,W)\cong\mathcal{M}_{m,n}(F)$；
2. 若 $\sigma\in\mathcal{L}(U,V)$ 且 $\tau\in\mathcal{L}(V,W)$，$\mathcal{B},\mathcal{C},\mathcal{D}$ 分别为 U,V,W 的有序基，则

$$[\tau\sigma]_{\mathcal{B},\mathcal{D}}=[\tau]_{\mathcal{C},\mathcal{D}}[\sigma]_{\mathcal{B},\mathcal{C}}, \tag{3.3.9}$$

因此乘积（复合）$\tau\sigma$ 的矩阵为 τ 和 σ 各自矩阵的乘积。事实上，这正是矩阵乘法定义的主要动机。♡

证明

1. 首先证明该映射是线性的，即对于任意 $\tau,\sigma\in\mathcal{L}(V,W)$ 和 $r,s\in F$，根据定理 3.5以及坐标映射线性的事实，

$$\begin{aligned}\mu(r\tau+s\sigma)e_i&=[r\tau+s\sigma]_{\mathcal{B},\mathcal{C}}[b_i]_{\mathcal{B}}=[(r\tau+s\sigma)(b_i)]_{\mathcal{C}}=[r\tau(b_i)+s\sigma(b_i)]_{\mathcal{C}}\\&=r[\tau(b_i)]_{\mathcal{C}}+s[\sigma(b_i)]_{\mathcal{C}}=(r[\tau]_{\mathcal{B},\mathcal{C}}+s[\sigma]_{\mathcal{B},\mathcal{C}})e_i.\end{aligned} \tag{3.3.10}$$

利用定理 3.1，μ 是线性变换。根据定理 3.2，可由 $[\tau]_{\mathcal{B},\mathcal{C}}$ 和有序基 $\mathcal{C}$ 确定 $\tau(b_i)$，进而确定 τ。因此 μ 为满射；根据推论 3.3，此时 μ 为双射，进而为同构；

2. 根据上述结果，

$$[\tau\sigma]_{\mathcal{B},\mathcal{D}}[b_i]_{\mathcal{B}}=[(\tau\sigma)(b_i)]_{\mathcal{D}}=[\tau]_{\mathcal{C},\mathcal{D}}[\sigma(b_i)]_{\mathcal{C}}=[\tau]_{\mathcal{C},\mathcal{D}}[\sigma]_{\mathcal{B},\mathcal{C}}[b_i]_{\mathcal{B}} \tag{3.3.11}$$

对任意 $b_i\in\mathcal{B}$ 成立。根据定理 3.1，$[\tau\sigma]_{\mathcal{B},\mathcal{D}}=[\tau]_{\mathcal{B},\mathcal{C}}[\sigma]_{\mathcal{D}}$。

3.3.4 线性变换的基变换

定理 3.7

设 $\tau\in\mathcal{L}(V,W)$，$(\mathcal{B},\mathcal{C})$ 和 $(\mathcal{B}',\mathcal{C}')$ 分别为 V 和 W 的两对有序基，则

$$[\tau]_{\mathcal{B}',\mathcal{C}'}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B}',\mathcal{B}}. \tag{3.3.12}$$

♡

证明 根据定理 3.4和3.5，对于任意 $b_i'\in\mathcal{B}'$，

$$[\tau]_{\mathcal{B}',\mathcal{C}'}[b_i']_{\mathcal{B}'}=[\tau(b_i')]_{\mathcal{C}'}=M_{\mathcal{C},\mathcal{C}'}[\tau(b_i')]_{\mathcal{C}}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}[b_i']_{\mathcal{B}}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B}',\mathcal{B}}[b_i']_{\mathcal{B}'}, \tag{3.3.13}$$

因此有 $[\tau]_{\mathcal{B}',\mathcal{C}'}=M_{\mathcal{C},\mathcal{C}'}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{B}',\mathcal{B}}$，结论得证。

对于 $V=W$ 以及 $\mathcal{B}'=\mathcal{C},\mathcal{C}'=\mathcal{B}$ 的特殊情形，我们可直接得到以下推论。

推论 3.5

设 $\tau\in\mathcal{L}(V)$，$\mathcal{B}$ 和 $\mathcal{C}$ 为 V 的两组有序基，则 $[\tau]_{\mathcal{C}}=M_{\mathcal{B},\mathcal{C}}[\tau]_{\mathcal{B}}M_{\mathcal{B},\mathcal{C}}^{-1}$。♡

3.4 等价性与相似性

3.4.1 矩阵的等价

由于基变换矩阵实际上为可逆矩阵，定理 3.7的内容正是引出“矩阵等价”这一概念的动机。

定义 3.5 (矩阵等价)

若存在可逆矩阵 P 和 Q，使得矩阵 A 和 B 之间满足 $B=PAQ^{-1}$，则称 A 和 B 等价（equivalent）。♣

定理 3.8

设 V 和 W 分别为维度为 n 和 m 的向量空间，则两个 $m\times n$ 矩阵 A 和 B 等价当且仅当其（可能在不同有序基下）表示同一个线性变换 $\tau\in\mathcal{L}(V,W)$。在这种情况下，A 和 B 表示 $\mathcal{L}(V,W)$ 中相同的线性变换集。♡

证明 若 A 和 B 均表示 τ，存在有序基对 $(\mathcal{B},\mathcal{C})$ 和 $(\mathcal{B}',\mathcal{C}')$ 使得 $A=[\tau]_{\mathcal{B},\mathcal{C}}$ 且 $B=[\tau]_{\mathcal{B}',\mathcal{C}'}$。根据定理 3.7，存在 $P=M_{\mathcal{C},\mathcal{C}'}$ 和 $Q=M_{\mathcal{B},\mathcal{B}'}$ 使得 $B=PAQ^{-1}$，从而 A 和 B 等价；反之若 A 和 B 等价，存在可逆矩阵 P 和 Q 使得 $B=PAQ^{-1}$。根据定理 3.5，可通过 A 和 W 的一组任意给定的有序基 $\mathcal{C}$ 确定 V 的一组有序基 $\mathcal{B}$，使得 $A=[\tau]_{\mathcal{B},\mathcal{C}}$，其中 τ 为从 V 到 W 的某个线性变换。类似地，根据定理 3.4，可通过 P 和 Q 以及给定的有序基 $\mathcal{B}$ 和 $\mathcal{C}$ 确定相应的有序基 $\mathcal{B}'$ 和 $\mathcal{C}'$，使得 $P=M_{\mathcal{C},\mathcal{C}'}$ 且 $Q=M_{\mathcal{B},\mathcal{B}'}$。那么利用定理 3.7，$B=PAQ^{-1}=[\tau]_{\mathcal{B}',\mathcal{C}'}$。

3.4.2 矩阵的相似

定义 3.6 (矩阵相似)

若存在可逆矩阵 P，使得矩阵 A 和 B 满足 $B=PAP^{-1}$，则称 A 和 B 相似（similar）。与矩阵相似有关的等价类称为相似类（similar class）。

♣

作为定理 3.8和推论 3.5的推论，我们有如下结果。

推论 3.6

设 V 为 n 维向量空间，那么两个 $n\times n$ 矩阵 A 和 B 相似当且仅当其（可能在不同的有序基下）表示同一个线性算子 $\tau\in\mathcal{L}(V)$。在这种情况下，A 和 B 表示 $\mathcal{L}(V)$ 中相同的线性算子集。

♡

3.4.3 算子的相似

定义 3.7

若存在自同构 $\phi\in\mathcal{L}(V)$ 使得线性算子 $\tau,\sigma\in\mathcal{L}(V)$ 满足 $\sigma=\phi\tau\phi^{-1}$，则称 τ 和 σ 相似。与算子相似相关的等价类称为相似类。

♣

定理 3.9

设 V 为 n 维向量空间，那么 V 上的两个线性算子相似当且仅当存在一个矩阵 $A\in\mathcal{M}_n$ 可同时（可能在不同的有序基下）表示两个算子，在这种情况下，τ 和 σ 可用 $\mathcal{M}_n$ 中相同的矩阵集表示。

♡

证明 若存在 V 的有序基 $\mathcal{B}$ 和 $\mathcal{C}$ 使得 $[\tau]_{\mathcal{B}}=A=[\sigma]_{\mathcal{C}}$，那么可定义同构 $\phi(b_i)=c_i$，从而根据定理 3.4，

$$[\phi]_{\mathcal{B}}=([\phi(b_1)]_{\mathcal{B}},\cdots,[\phi(b_n)]_{\mathcal{B}})=([c_1]_{\mathcal{B}},\cdots,[c_n]_{\mathcal{B}})=M_{\mathcal{C},\mathcal{B}},\tag{3.4.1}$$

因此根据定理 3.6和推论 3.5，

$$[\sigma]_{\mathcal{B}}=M_{\mathcal{C},\mathcal{B}}[\sigma]_{\mathcal{C}}M_{\mathcal{C},\mathcal{B}}^{-1}=[\phi]_{\mathcal{B}}[\tau]_{\mathcal{B}}[\phi^{-1}]_{\mathcal{B}}=[\phi\tau\phi^{-1}]_{\mathcal{B}}.\tag{3.4.2}$$

可见对于任意给定的有序基，$\sigma=\phi\tau\phi^{-1}$，即 σ 和 τ 等价；反之若 τ 和 σ 等价，存在自同构 ϕ 使得 $\sigma=\phi\tau\phi^{-1}$。对于任意给定的有序基 $\mathcal{B}$，可通过 $\phi(b_i)=c_i$ 定义一组有序基 $\mathcal{C}$。根据(3.4.1)式，$[\phi]_{\mathcal{B}}=M_{\mathcal{C},\mathcal{B}}$，进而

$$[\tau]_{\mathcal{B}}=[\phi^{-1}\sigma\phi]_{\mathcal{B}}=[\phi]_{\mathcal{B}}^{-1}[\sigma]_{\mathcal{B}}[\phi]_{\mathcal{B}}=M_{\mathcal{B},\mathcal{C}}[\sigma]_{\mathcal{B}}M_{\mathcal{B},\mathcal{C}}^{-1}=[\sigma]_{\mathcal{C}}.\tag{3.4.3}$$

3.5 不变子空间与约化对

一般而言，线性算子 $\tau\in\mathcal{L}(V)$ 在 V 的子空间 S 上的限制未必是线性算子（封闭性），由此引出如下定义。

定义 3.8 (不变子空间)

设 $\tau\in\mathcal{L}(V)$。对于 V 的一个子空间 S，若 $\tau(S)\subseteq S$，则称 S 在 τ 下不变（invariant under τ）或者 τ-不变（τ-invariant）；换言之，对于任意 $s\in S$，均有 $\tau(s)\in S$，即限制 $\tau|_S$ 是 S 上的线性算子。

♣

例 3.3 在 $\mathbb{R}^2$ 中定义线性算子 $\tau(x,y)=(x+y,y)$，那么子空间 $S=\{(x,0)|x\in\mathbb{R}\}$ 是 τ-不变的，但子空间 $T=\{(0,y)|y\in\mathbb{R}\}$ 不是。这说明一般而言，即便 V 的子空间 S 是 τ-不变的，其补空间 T 也未必是 τ-不变的。

定义 3.9 (约化)

设 $\tau\in\mathcal{L}(V)$。若 $V=S\oplus T$ 且 S 和 T 均为 τ-不变的，称 (S,T) 约化（reduce）τ。 ♣

约化对可用于将线性算子如下分解为直和形式。

定义 3.10 (算子直和)

设 $\tau\in\mathcal{L}(V)$。若 (S,T) 约化 τ，记 $\tau=\tau|_S\oplus\tau|_T$，并称 τ 为 $\tau|_S$ 与 $\tau|_T$ 的直和。因此，表达式 $\rho=\sigma\oplus\tau\in\mathcal{L}(V)$ 意味着存在 V 的子空间 S 和 T，使得 (S,T) 约化 ρ，并且 $\sigma=\rho|_S,\tau=\rho|_T$。 ♣

3.6 投影算子

3.6.1 定义与基本性质

定义 3.11 (投影)

设 $V=S\oplus T$，则线性算子 $\rho_{S,T}:V\to V$ 定义为 $\rho_{S,T}(s+t)=s$，其中 $s\in S,t\in T$。$\rho_{S,T}$ 称为沿 T 到 S 上的投影（projection onto S along T）。 ♣

以下是投影算子的基本性质。

命题 3.6

设 V 为向量空间，$\rho\in\mathcal{L}(V)$。

1. 若 $V=S\oplus T$，则 $\rho_{S,T}+\rho_{T,S}=\iota$；
2. 若 $\rho=\rho_{S,T}$，则 $\mathrm{im}(\rho)=S$ 且 $\ker(\rho)=T$，进而 $V=\mathrm{im}(\rho)\oplus\ker(\rho)$。此外 $v\in\mathrm{im}(\rho)\Leftrightarrow\rho v=v$；
3. 若 $\sigma\in\mathcal{L}(V)$ 满足 $V=\mathrm{im}(\sigma)\oplus\ker(\sigma)$ 且 $\sigma|_{\mathrm{im}(\sigma)}=\iota$，则 σ 为沿 $\ker(\sigma)$ 到 $\mathrm{im}(\sigma)$ 上的投影。 ♠

证明

1. 对于 $V=S\oplus T$ 内任意向量 v，存在 $s\in S$ 和 $t\in T$ 使得 $v=s+t$，那么 $(\rho_{S,T}+\rho_{T,S})(v)=s+t=v$，即 $\rho_{S,T}+\rho_{T,S}=\iota$；
2. 对于任意 $v\in V$，存在 $s\in S$ 和 $t\in T$ 使得 $v=s+t$，因此 $\rho(v)=s$，即 $\mathrm{im}(\rho)\subseteq S$；另一方面，对于任意 $s\in S$，由于 $0\in T$，因此存在 $v\in V$，使得 $s=s+0=v\in V$，那么 $s=\rho(s)$，即 $S\subseteq\mathrm{im}(\rho)$。综上，$\mathrm{im}(\rho)=S$。类似分析可知 $\ker(\rho)=T$，因此 $V=S\oplus T=\mathrm{im}(\rho)\oplus\ker(\rho)$。那么 $v\in\mathrm{im}(\rho)=S$ 当且仅当 $\rho v=v\rho(v+0)=v$；
3. 记 $S=\mathrm{im}(\sigma),T=\ker(\sigma)$，则对于任意 $s\in S$ 和 $t\in T$，首先有 $\sigma(t)=0$，因此
$$\sigma(s+t)=\sigma(s)+\sigma(t)=\sigma|_S(s)+0=s, \tag{3.6.1}$$
因此 σ 为沿 $T=\ker(\sigma)$ 到 $S=\mathrm{im}(\sigma)$ 的投影。

事实上，除去上述定义，投影算子存在一种更简单的等价描述。

定义 3.12 (幂等)

若线性算子 $\tau\in\mathcal{L}(V)$ 满足 $\tau^2=\tau$，则称其幂等（idempotent）。 ♣

定理 3.10

线性算子 $\rho\in\mathcal{L}(V)$ 是投影当且仅当其幂等。 ♡

证明 若 $\rho=\rho_{S,T}$，其中 S 和 T 分别为 V 的子空间，那么对于任意 $v\in V$，存在 $s\in S$ 和 $t\in T$ 使得 $v=s+t$。那么 $\rho^2v=\rho(\rho(s+t))=\rho(s+0)=s=\rho(v)$，即 $\rho^2=\rho$；反之若 $\rho^2=\rho$，对于任意 $v\in\mathrm{im}(\rho)$，总存在 $x\in V$ 使得 $v=\rho x=\rho^2x=\rho v$，因此 $\rho|_{\mathrm{im}(\rho)}=\iota$。考虑 $u\in\mathrm{im}(\rho)\cap\ker(\rho)$，那么 $u=\rho u=0$，即 $\mathrm{im}(\rho)\cap\ker(\rho)=\{0\}$；另外对于任意 $v\in V$，总有 $v=\rho v+(v-\rho v)\in\mathrm{im}(\rho)+\ker(\rho)$，因此 $V=\mathrm{im}(\rho)\oplus\ker(\rho)$。根据命题 3.6，$\rho$ 是沿 $\ker(\rho)$ 到 $\mathrm{im}(\rho)$ 上的投影。

3.6.2 正交投影与单位算子分解

定义 3.13 (正交投影)

对于投影 $\rho,\sigma\in\mathcal{L}(V)$，若 $\rho\sigma=\sigma\rho=0$，则称两者正交（orthogonal），记作 $\rho\perp\sigma$。

命题 3.7

投影 $\rho,\sigma\in\mathcal{L}(V)$ 正交当且仅当 $\mathrm{im}(\rho)\subseteq\ker(\sigma)$ 且 $\mathrm{im}(\sigma)\subseteq\ker(\rho)$。

证明 若 ρ 与 σ 正交，则对于任意 $v\in\mathrm{im}(\rho)$，根据命题 3.6，$\sigma v=\sigma(\rho v)=0$，因此 $\mathrm{im}(\sigma)\subseteq\ker(\sigma)$。类似地 $\mathrm{im}(\sigma)\subseteq\ker(\rho)$，必要性得证；反之若必要性成立，对于任意 $v\in V$，$\rho v\in\mathrm{im}(\rho)$，因此 $\sigma\rho v=0$，类似地 $\rho\sigma v=0$，故 $\rho\sigma=\sigma\rho=0$。

下面的例子将说明，仅有 $\rho\sigma=0$ 的条件是不充分的，甚至可能在此情况下 $\sigma\rho$ 不是投影。

例 3.4 设 $V=F^2$，考虑 x 轴和 y 轴以及对角线

$$X=\{(x,0)|x\in F\},\quad Y=\{(0,y)|y\in F\},\quad D=\{(x,x)|x\in F\},\tag{3.6.2}$$

那么对于任意 $(x,y)\in F^2$，容易验证

$$\rho_{D,X}\rho_{D,Y}(x,y)=\rho_{D,X}\rho_{D,Y}((x,x)+(0,y-x))=\rho_{D,X}(x,x)=(x,x),\tag{3.6.3}$$

$$\rho_{D,Y}\rho_{D,X}(x,y)=\rho_{D,Y}\rho_{D,X}((y,y)+(x-y,0))=\rho_{D,Y}(y,y)=(y,y).\tag{3.6.4}$$

可见若 ρ,σ 为投影，其乘积 $\rho\sigma$ 和 $\sigma\rho$ 可能仍为投影，但两者不相等。此外，

$$\rho_{Y,X}\rho_{X,D}(x,y)=\rho_{Y,X}\rho_{X,D}((x-y,0)+(y,y))=\rho_{Y,X}(x-y,0)=0,\tag{3.6.5}$$

$$\rho_{X,D}\rho_{Y,X}(x,y)=\rho_{X,D}(0,y)=\rho_{X,D}((-y,0)+(y,y))=(-y,0),\tag{3.6.6}$$

可见 $\rho_{Y,X}\rho_{X,D}=0$，但 $(\rho_{X,D}\rho_{Y,X})^2(x,y)=\rho_{X,D}\rho_{Y,X}(-y,0)=(0,0)\neq\rho_{X,D}\rho_{Y,X}(x,y)$，即 $\rho_{X,D}\rho_{Y,X}$ 非投影。

由于投影 ρ 幂等，可将单位算子 ι 写为两个正交投影的和：$\iota=\rho+(\iota-\rho)$，并将此概念推广至更多投影。

定义 3.14 (单位算子的分解)

V 上单位算子的分解（resolution of the identity）为求和形式

$$\rho_1+\cdots+\rho_k=\iota,\tag{3.6.7}$$

其中 ρ_i 为两两（pairwise）正交的投影，即当 $i\neq j$ 时 $\rho_i\perp\rho_j$。

V 上单位算子的分解与 V 的直和分解存在如下关系。

定理 3.11

设 V 为向量空间，V 上单位算子的分解与 V 的直和分解（direct sum decompositions）存在如下对应关系：

1. 若 $\rho_1+\cdots+\rho_k=\iota$ 为单位算子的一种分解，则 $V=\mathrm{im}(\rho_1)\oplus\cdots\oplus\mathrm{im}(\rho_k)$，其中 ρ_i 是沿 $\ker(\rho_i)=\oplus_{j\neq i}\mathrm{im}(\rho_j)$ 到 $\mathrm{im}(\rho_i)$ 上的投影；
2. 若 $V=S_1\oplus\cdots\oplus S_k$ 且 ρ_i 为沿直和 $\oplus_{j\neq i}S_j$ 到 S_i 上的投影，则 $\rho_1+\cdots+\rho_k=\iota$ 是单位算子的分解。

证明

1. 对于任意 $v\in V$，由条件知 $v=(\rho_1+\cdots+\rho_k)v\in\mathrm{im}(\rho_1)+\cdots+\mathrm{im}(\rho_k)$，因此 $V=\mathrm{im}(\rho_1)+\cdots+\mathrm{im}(\rho_k)$。对于任意 $i\neq j$，考虑 $u\in\mathrm{im}(\rho_i)\cap\mathrm{im}(\rho_j)$，那么根据定义，$u=\rho_j u=\rho_i\rho_j u=0$，即 $\mathrm{im}(\rho_i)\cap\mathrm{im}(\rho_j)=\{0\}$，因此 $V=\mathrm{im}(\rho_1)\oplus\cdots\oplus\mathrm{im}(\rho_k)$，并且 $\ker(\rho_i)=\oplus_{j\neq i}\mathrm{im}(\rho_j)$；
2. 根据命题 3.6，对于任意指标 $i\neq j$，$\mathrm{im}(\rho_i)=S_i\subseteq\oplus_{k\neq j}S_k=\ker(\rho_j)$，类似地 $\mathrm{im}(\rho_j)\subseteq\ker(\rho_i)$。根据命题 3.7，$\rho_i$ 两两正交。最后对于任意 $v\in V$，存在 $s_i\in S_i$ 使得 $v=s_1+\cdots+s_k$，进而

$$\rho_i v=\rho_i\left(s_i+\sum_{j\neq i}s_j\right)=s_i\quad\Longrightarrow\quad(\rho_1+\cdots+\rho_k)v=s_1+\cdots+s_k=v,\tag{3.6.8}$$

因此 $\rho_1+\cdots+\rho_k=\iota$。

3.6.3 投影代数

命题 3.8

设 V 为特征[a]不为 2 的域 F 上的向量空间，ρ,σ 为投影。

1. 和 $\rho+\sigma$ 是投影当且仅当 $\rho\perp\sigma$，在这种情况下

$$\mathrm{im}(\rho+\sigma)=\mathrm{im}(\rho)\oplus\mathrm{im}(\sigma),\quad\ker(\rho+\sigma)=\ker(\rho)\cap\ker(\sigma).\tag{3.6.9}$$

2. 差 $\rho-\sigma$ 是投影当且仅当 $\rho\sigma=\sigma\rho=\sigma$，在这种情况下

$$\mathrm{im}(\rho-\sigma)=\mathrm{im}(\rho)\cap\ker(\sigma),\quad\ker(\rho-\sigma)=\ker(\rho)\oplus\mathrm{im}(\sigma).\tag{3.6.10}$$

3. 若 ρ 和 σ 可交换，则 $\rho\sigma$ 是投影，在这种情况下

$$\mathrm{im}(\rho\sigma)=\mathrm{im}(\rho)\cap\mathrm{im}(\sigma),\quad\ker(\rho\sigma)=\ker(\rho)+\ker(\sigma).\tag{3.6.11}$$

[a]简言之，特征是使得乘法单位元 1 加和为加法单位元 0 的加项数，例如比特在按位加法和按位乘法下形成特征为 2 的域（$1+1=0$）

证明

1. $\rho+\sigma$ 是投影当且仅当 $\rho+\sigma=(\rho+\sigma)^2=\rho+\sigma+\rho\sigma+\sigma\rho$，等价于 $\rho\sigma=-\sigma\rho$。若 $\rho\perp\sigma$，该条件显然成立；反之可将其改写为 $\rho\sigma=\rho^2\sigma=\rho(-\sigma\rho)=(\sigma\rho)\rho=\sigma\rho$。结合 $\rho\sigma=-\sigma\rho$ 可知，$\rho\sigma=\sigma\rho=0$。现在对于任意 $v\in\mathrm{im}(\rho+\sigma)$，根据定理 3.10，$v=(\rho+\sigma)v\in\mathrm{im}(\rho)+\mathrm{im}(\sigma)$。并且类似定理 3.11的证明，$\rho\perp\sigma$ 时 $\mathrm{im}(\rho)\cap\mathrm{im}(\sigma)=\{0\}$，因此 $\mathrm{im}(\rho+\sigma)\subseteq\mathrm{im}(\sigma)\oplus\mathrm{im}(\sigma)$。另一方面，对于任意 $u\in\mathrm{im}(\rho)\oplus\mathrm{im}(\sigma)$，总存在 $s\in\mathrm{im}(\rho)$ 和 $t\in\mathrm{im}(\sigma)$ 使得 $u=s+t$。那么 $\rho u=s$ 且 $\rho u=t$，进而 $(\rho+\sigma)u=s+t=u$，可见 $\mathrm{im}(\rho)\oplus\mathrm{im}(\sigma)\subseteq\mathrm{im}(\rho+\sigma)$。综上，$\mathrm{im}(\rho+\sigma)=\mathrm{im}(\rho)\oplus\mathrm{im}(\sigma)$。对于后一结论，$\ker(\rho)\cap\ker(\sigma)\subseteq\ker(\rho+\sigma)$ 是显然的。对于任意 $v\in\ker(\rho+\sigma)$，有 $\rho v=-\sigma v$，因此 $\rho v=\rho^2 v=-\rho\sigma v=0$，类似地 $\sigma v=0$，即 $v\in\ker(\rho)\cap\ker(\sigma)$，充分性得证；
2. $\rho-\sigma$ 是投影当且仅当 $\rho-\sigma=(\rho-\sigma)^2=\rho+\sigma-\rho\sigma-\sigma\rho$，即 $(\iota-\rho)\sigma=-\sigma(\iota-\rho)$。根据上一结论，这等价于 $\sigma\perp(\iota-\rho)$，因此 $\sigma(\iota-\rho)=(\iota-\rho)\sigma=0$，即 $\sigma=\rho\sigma=\sigma\rho$。在此条件下，显然有 $\mathrm{im}(\rho)\cap\ker(\sigma)\subseteq\mathrm{im}(\rho-\sigma)$；对于任意 $v\in\mathrm{im}(\rho-\sigma)$，要求 $v=(\rho-\sigma)v$，即 $\sigma v=\sigma(\rho-\sigma)v=0$，回代得 $v=\rho v$。因此 $v\in\mathrm{im}(\rho)\cap\ker(\sigma)$，即 $\mathrm{im}(\rho-\sigma)\subseteq\mathrm{im}(\rho)\cap\ker(\sigma)$。对于后一结论，首先考虑 $u\in\ker(\rho)\cap\mathrm{im}(\sigma)$，有 $u=\sigma u$ 且 $\rho u=0$，因此 $0=\rho u=\rho\sigma u=\sigma u$，回代得 $u=0$，即 $\ker(\rho)\cap\mathrm{im}(\sigma)=\{0\}$。对于任意 $v\in\ker(\rho)\oplus\mathrm{im}(\sigma)$，存在 $s\in\ker(\rho)$ 和 $t\in\mathrm{im}(\sigma)$ 使得 $v=s+t$。已知 $\rho s=0$ 且 $\sigma t=t$，那么 $0=\sigma\rho s=\sigma s$，$\rho t=\rho\sigma t=\sigma t=t$。因此 $\rho v=\rho(s+t)=s+t=v$ 且 $\sigma v=\sigma(s+t)=0$，进而 $(\rho-\sigma)v=v$，即 $\ker(\rho)\oplus\mathrm{im}(\sigma)\subseteq\ker(\rho-\sigma)$；另一方面，对于任意 $v\in\ker(\rho-\sigma)$，已知 $\rho v=\sigma v$，那么 $v=(v-\rho v)+\rho v=(v-\rho v)+\sigma v\in\ker(\rho)+\mathrm{im}(\sigma)$，因此 $\ker(\rho-\sigma)\subseteq\ker(\rho)\oplus\mathrm{im}(\sigma)$，充分性得证；
3. 根据条件，$(\rho\sigma)^2=\rho(\rho\sigma)\sigma=\rho\sigma$，因此 $\rho\sigma$ 是投影。显然有 $\mathrm{im}(\rho)\cap\mathrm{im}(\sigma)\subseteq\mathrm{im}(\rho\sigma)$；对于任意 $v\in\mathrm{im}(\rho\sigma)$，有 $v=\rho(\sigma v)=\sigma(\rho v)\in\mathrm{im}(\rho)\cap\mathrm{im}(\sigma)$，故 $\mathrm{im}(\rho\sigma)\subseteq\mathrm{im}(\rho)\cap\mathrm{im}(\sigma)$。对于后一结论，显然有 $\ker(\rho)+\ker(\sigma)\subseteq\ker(\rho\sigma)$；反之对于任意 $v\in\ker(\rho\sigma)$，已知 $\rho\sigma v=0$，那么 $v=\sigma v+(v-\sigma v)\in\ker(\rho)+\ker(\sigma)$，即 $\ker(\rho\sigma)\subseteq\ker(\rho)+\ker(\sigma)$，结论得证

第 4 章 同构定理

目录

4.1 商空间 39
4.2 商空间的泛性质与第一同构定理 40
4.3 线性泛函与对偶基 41

4.1 商空间

4.1.1 基本概念

设 S 为向量空间 V 的一个子空间，u 和 v 为 V 内的向量。容易验证，定义为 $u-v\in S$ 的二元关系 $u\equiv v$ 是一个等价关系。当 $u\equiv v$ 时，我们称 u 和 v 模 S 同余（congruent modulo S）。术语“mod”是 modulo 的口语化表述，$u\equiv v$ 通常写为 $u\equiv v \mod S$。当子空间在上下文明确时，我们简单将其写为 $u\equiv v$。

为了看清等价类的结构，我们记

$$\{u\in V|u\equiv v\}=\{u\in V|u-v\in S\}=\{v+s|s\in S\}=v+S, \tag{4.1.1}$$

因此集合 $v+S=\{v+s|s\in S\}$ 称为 V 在 S 中的一个陪集（*coset*），v 称为 $(v+S)$ 的一个陪集代表元（*coset representative*）。V 中 S 的所有陪集构成集合称为V 模 S 的商空间（*quotient space of V modulo S*），读作“V 模 S”，记作 $V/S=\{v+S|v\in V\}$。当然，这里的术语“空间”提示我们将定义 V/S 上向量空间的运算。一个自然的选择是

$$(u+S)+(v+S)=(u+v)+S,\quad r(u+S)=(ru)+S. \tag{4.1.2}$$

不过仍需要检查这些运算是否良定义（*well-defined*），即运算结果是否与代表元的选择无关：若 $u_1+S=u_2+S, v_1+S=v_2+S$，我们希望 $(u_1+v_1)+S=(u_2+v_2)+S$ 且 $(ru_1)+S=(ru_2)+S$。根据条件，$u_1-u_2\in S$ 且 $v_1-v_2\in S$，因此

$$(u_1+v_1)+S=(u_2+v_2)+(u_1-u_2)+(v_1-v_2)+S=(u_2-v_2)+S, \tag{4.1.3}$$

$$(ru_1)+S=(ru_2)+r(u_1-u_2)+S=(ru_2)+S. \tag{4.1.4}$$

可见加法和标量乘法是良定义的，以上结果也可以用等价关系 $\equiv$ 更简洁地表述为

- $u_1\equiv u_2, v_1\equiv v_2\ \Rightarrow\ (u_1+v_1)\equiv(u_2+v_2)$；
- $u_1\equiv u_2\ \Rightarrow\ ru_1\equiv ru_2$。

这一结果在代数中会反复出现。在一个代数结构上（例如群、环、向量空间），若一个等价关系保持代数运算，则称其为同余关系（*congurence relation*）。

定理 4.1 (商空间)

设 S 为 V 的子空间，二元关系 $u\equiv v\ \Leftrightarrow\ u-v\in S$ 为 V 上的等价关系，其等价类为陪集 $v+S=\{v+s|s\in S\}$。V 中 S 的所有陪集构成的集合 V/S 称为 V 模 S 的商空间，其自身构成一个向量空间，包含良定义的运算：

$$(u+S)+(v+S)=(u+v)+S,\quad r(u+S)=ru+S. \tag{4.1.5}$$

V/S 的零向量为陪集 $0+S=S$。

♡

4.1.2 自然投影与对应定理

若 S 为 V 的一个子空间，则可以定义将每个向量嵌入包含其陪集的映射 $\pi_S : V \to V/S,\ v \mapsto v+S$。该映射称为从 V 到 V/S 的典范/自然投影（*canonical/natural projection*），或者简单称为模 S 的投影（*projection modulo* S）。注意这里“投影”并非上一章中“投影算子”的“投影”含义。

定理 4.2 (典范投影)

典范投影 $\pi_S : V \to V/S$ 定义为 $\pi_S(v) = v + S$，该投影为满射线性变换，其核为 $\ker(\pi_S) = S$。♡

证明 容易证明 π_S 是满射。对于任意 $u, v \in V$ 和 $r, s \in F$，

$$\pi_S(ru+sv) = (ru+sv)+S = (ru+S)+(sv+S) = r(u+S)+s(v+S) = r\pi_S(u)+s\pi_S(v), \tag{4.1.6}$$

因此 π_S 为线性变换，最后

$$\ker(\pi_S) = \{v \in V | v+S = S\} = \{v \in V | v \in S\} = S. \tag{4.1.7}$$

4.2 商空间的泛性质与第一同构定理

设 S 为 V 的子空间，那么 $(V/S, \pi_S)$ 具有一个常称为泛性质（*university*）[1]的特殊性质，具体如下。

定理 4.3

设 S 为 V 的一个子空间，$\tau \in \mathcal{L}(V,W)$ 满足 $S \subseteq \ker(\tau)$，则存在唯一线性变换 $\tau' : V/S \to W$，使得 $\tau' \circ \pi_S = \tau$。在这种情况下，$\ker(\tau') = \ker(\tau)/S$ 且 $\mathrm{im}(\tau') = \mathrm{im}(\tau)$。♡

证明 若希望构造映射 $\tau' : V/S \to W$，首先要求同余关系的等价类的像与该等价类的代表元选取无关；换言之，若 $u \equiv v \mod S$，则希望 $\tau'(u+S) = \tau'(v+S)$。对于线性变换，该式等价于

$$0 = \tau'(u+S) - \tau'(v+S) = \tau'(u-v+S) = \tau' \circ \pi_S(u-v) = \tau(u-v) \quad \Leftrightarrow \quad u-v \in \ker(\tau). \tag{4.2.1}$$

因此当 $S \subseteq \ker(\tau)$ 时，上式显然满足，此时 τ' 良定义，并且

$$\mathrm{im}(\tau') = \{\tau'(v+S) | v+S \in V/S\} = \{\tau(v) | v \in V\} = \mathrm{im}(\tau), \tag{4.2.2}$$

$$\ker(\tau') = \{v+S | \tau'(v+S) = 0\} = \{v+S | \tau(v) = 0\} = \ker(\tau)/S. \tag{4.2.3}$$

由于 τ' 在 V 内每个向量上的取值完全由 τ 确定，故 τ' 唯一。

上述定理具有一个非常重要的推论，称为第一同构定理（*first isomorphism theorem*）。

定理 4.4 (第一同构定理)

设 $\tau : V \to W$ 为线性变换，则线性变换

$$\tau' : V/\ker(\tau) \to W \text{ defined by } \tau'(v+\ker(\tau)) = \tau(v) \tag{4.2.4}$$

为单射，并且 $V/\ker(\tau) \cong \mathrm{im}(\tau)$。♡

证明 在定理 4.3 中取 $S = \ker(\tau)$，那么存在唯一线性变换 $\tau' : V/\ker(\tau) \to W$ 使得 $\tau' \circ \pi_S = \tau$，其将 $v+\ker(\tau)$ 映射到 $\tau(v)$。因此 $\ker(\tau') = \ker(\tau)/\ker(\tau) = \ker(\tau)$ 为 $V/\ker(\tau)$ 中的零向量。根据命题 3.2，τ' 为单射，并且 $\mathrm{im}(\tau') = \mathrm{im}(\tau)$ 意味着 τ' 同时也是从 $V/\ker(\tau)$ 到 $\mathrm{im}(\tau)$ 的满射，因此 $V/\ker(\tau) \cong \mathrm{im}(\tau)$。

根据第一同构定理，任何 V 上的线性变换 $\tau : V \to \mathrm{im}(\tau)$ 同构于 V 上的一个商空间；反之，V 的任意商空间 V/S 是 V 上典范投影 π_S 的像。因此，在同构意义下，商空间等价于同态像。

[1]关于泛性质的严格定义，请参见附录第??章介绍

推论 4.1

设 S 为 V 的一个子空间，则 V 中 S 的所有补空间同构于 V/S，进而彼此同构。♡

证明 设 T 为 S 的一个补空间，将第一同构定理 4.4应用于投影算子 $\rho_{S,T}:V\to S$，得到 $T\cong V/S$。

推论 4.2

设 S 为向量空间 V 的一个子空间，则 $\dim(V)=\dim(S)+\dim(V/S)$。♡

证明 对于 V 的子空间 S 的任意补空间 T，投影算子 $\rho_{S,T}$ 的核与像分别为 T 和 S。根据秩-零化度定理 3.3，$\dim(V)=\dim(S)+\dim(T)$。再利用推论 3.2和4.1，$\dim(T)=\dim(V/S)$，从而结论得证。

注 需要注意的是，由于基数减法未定义，不能将上述结论写为 $\dim(V/S)=\dim(V)-\dim(S)$。

根据此推论，我们引入一个可能会偶尔用到的概念。

定义 4.1 (余维数)

若 S 是 V 的一个子空间，则 $\dim(V/S)$ 称为 V 中 S 的余维数（codimension），记作 $\operatorname{codim}(S)$ 或 $\operatorname{codim}_V(S)$。♣

另一方面，需要指出的是，商空间补空间的同构性可能并没有预期的那么好。

例 4.1

1. 考虑 $F[x]$ 上的一组基 $\{1,x,x^2,\cdots\}$，取
$$A=\langle x,x^2,x^3,\cdots\rangle,\quad B=\langle 1\rangle,\quad C=\langle 1,x^2,x^4,\cdots\rangle,\quad D=\langle x,x^3,x^5\rangle. \tag{4.2.5}$$
定义映射 $f:A\to D$，使得 $f(x^{n+1})=x^{2n}$ $(n\in\mathbb{N})$，可证明 f 为同构；但另一方面，$\dim(B)=1\neq\aleph_0=\dim(D)$。这意味着此时 $A\oplus B=F[x]=C\oplus D$ 且 $A\cong C$，但 $B\ncong D$；
2. 考虑 $F[x,y]$ 上的一组基 $\{x^my^n\}_{m,n\in\mathbb{N}}$，取
$$S=\langle x,x^2,\cdots\rangle,\quad B=\langle 1\rangle,\quad D=\langle 1,y,y^2,\cdots\rangle. \tag{4.2.6}$$
定义映射 $g:V\to W$，使得 $g(x^n)=\begin{cases}x^{(n-1)/2}, & \text{if } n \text{ is odd},\\ y^{n/2+1}, & \text{if } n \text{ is even},\end{cases}$ 可证明 g 是同构；但另一方面，$\dim(B)=1\neq\aleph_0=\dim(D)$。这意味着此时 $S\oplus B\cong S\oplus D$，但 $B\ncong D$。

4.3 线性泛函与对偶基

4.3.1 线性泛函

定义 4.2

设 V 为域 F 上的向量空间。若线性变换 $f\in\mathcal{L}(V,F)$ 取值于基域 F，则称其为 V 上的线性泛函（linear functional），或者简称为泛函。V 上所有线性泛函构成的向量空间记为 V^*，称其为 V 的代数对偶空间（algebraic dual space）。♣

注 此处术语“代数”是必要的，因为在第8.4节我们将介绍另一种对偶空间：连续对偶空间；不过目前术语“对偶空间”仅仅指代数对偶空间。为了区别于其他类型的线性变换，常用小写意大利字母 f,g,h 表示线性泛函。

例 4.2

1. 由 $f(p(x))=p(0)$ 定义的映射 $f:F[x]\to F$ 是一个线性泛函，称为 *0 点取值泛函（evaluation at 0）*；
2. 记 $\mathcal{C}[a,b]$ 为在 $[a,b]\subseteq\mathbb{R}$ 上所有连续函数构成的线性空间，映射 $f:\mathcal{C}[a,b]\to\mathbb{R}$ 定义为
$$f(\alpha(x))=\int_a^b\alpha(x)\,\mathrm{d}x, \tag{4.3.1}$$
则 f 是一个线性泛函，记作 $f\in\mathcal{C}[a,b]^*$。

对于任意 $f \in V^*$，秩-零化度定理 3.3给出 $\dim(\ker(f)) + \dim(\mathrm{im}(f)) = \dim(V)$。由于 $\mathrm{im}(f) \subseteq F$，首先有 $0 \in \mathrm{im}(F)$；若存在其他非零元 r，则 $\langle r\rangle = \{rs | s \in F\} = F$。这意味着 $\mathrm{im}(f) = \{0\}$ 或 F，即非零线性泛函是满射。下面我们列举线性泛函其他的一些基本性质。

命题 4.1

1. 对于 V 中任意非零向量 $v \in V$，存在线性泛函 $f \in V^*$ 使得 $f(v) \neq 0$；
2. 向量 $v \in V$ 为零当且仅当 $f(v) = 0$ 对于任意 $f \in V^*$ 成立；
3. 设 $f \in V^*$，若 $f(x) \neq 0$，则 $V = \langle x\rangle \oplus \ker(f)$；
4. 两个非零线性泛函 $f, g \in V^*$ 具有相同的核当且仅当存在非零标量 λ，使得 $f = \lambda g$。

证明 前两条性质是显然的。对于性质 3，首先考虑 $u \in \langle x\rangle \cap \ker(f)$，那么存在 $r \in F$，使得 $u = rx$，从而对于任意线性泛函 $f \in V^*$，$0 = fu = f(rx) = rf(x)$。根据性质 2，这要求 $x = 0$，因此 $\langle x\rangle \cap \ker(f) = \{0\}$。$\langle x\rangle \oplus \ker(f) \subseteq V$ 是显然的；对于反向包含，对任意 $v \in V$ 定义 $r = f(v)/f(x)$，那么 $f(v - rx) = f(v) - rf(x) = 0$，即 $v = rx + (v - rx) \in \langle x\rangle \oplus \ker(f)$，因此 $V \subseteq \langle x\rangle \oplus \ker(f)$。

对于性质 4，必要性是显然的；对于充分性，在 $K = \ker(f) = \ker(g)$ 内，f 与 g 的取值相同。对于任意 $x \notin K$，根据性质 3，$V = \langle x\rangle \oplus K$。取 $\lambda = f(x)/g(x)$，则 $f|_{\langle x\rangle} = \lambda g|_{\langle x\rangle}$。可见 $f(x) = \lambda g(x)$ 对所有 $x \in V$ 成立。

4.3.2 对偶基

设 V 为包含一组基 $\mathcal{B} = \{v_i | i \in I\}$ 的向量空间。对于每个指标 $i \in I$，我们可通过正交性条件

$$v_i^*(v_j) = \delta_{i,j}, \quad \text{where } \delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j \end{cases} \tag{4.3.2}$$

定义一个线性泛函 $v_i^* \in V^*$，其中 $\delta_{i,j}$ 为 *Kronecker* δ-函数（*Kronecker delta function*）。那么对于 $a_{i_k} \in F$，

$$0 = a_{i_1} v_{i_1}^* + \cdots + a_{i_n} v_{i_n}^* \quad \Longrightarrow \quad 0 = \sum_{j=1}^{n} a_{i_j} v_{i_j}^*(v_{i_k}) = a_{i_k} \tag{4.3.3}$$

对所有 $k = 1, \cdots, n$ 成立，因此 $\mathcal{B}^* = \{v_i^* | i \in I\}$ 线性无关。

定理 4.5 (对偶基)

设 V 为具有基 $\mathcal{B} = \{v_i | i \in I\}$ 的线性空间。

1. 集合 $\mathcal{B}^* = \{v_i^* | i \in I\}$ 线性无关，其中 v_i^* 按(4.3.2)式定义；
2. 若 V 是有限维的，则 $\mathcal{B}^*$ 是 V^* 的一组基，称为 $\mathcal{B}$ 的对偶基（dual basis）。

证明 第一条结论已证明。对于任意 $f \in V^*$，设 $\dim(V) = n$，我们有 $\sum_{j=1}^{n} f(v_j) v_j^*(v_i) = f(v_i)$。因此对于任意 $v \in V$，存在唯一一组 $r_i \in F$，使得 $v = \sum_{i=1}^{n} r_i v_i$，进而

$$f(v) = \sum_{i=1}^{n} r_i f(v_i) = \sum_{i=1}^{n} r_i \left(\sum_{j=1}^{n} f(v_j) v_j^*(v_i) \right) = \sum_{j=1}^{n} f(v_j) v_j^* \left(\sum_{i=1}^{n} r_i v_i \right) = \sum_{j=1}^{n} f(v_j) v_j^*(v), \tag{4.3.4}$$

即 $f = \sum_{j=1}^{n} f(v_j) v_j^*$。因此 $V^* = \langle \mathcal{B}^* \rangle$，根据定理 2.1，$\mathcal{B}^*$ 是 V^* 的基。

根据上述定理，若 $\dim(V) < \infty$，则 $\dim(V^*) = \dim(V) < \infty$，这是因为对偶向量同样形成了 V^* 的一组基。但对于无限维情形，情况要复杂得多。我们先给出一个直观的例子，最后不加证明地给出结论。

例 4.3 设 V 为域 $F = \mathbb{Z}_2 = \{0, 1\}$ 上的一个无限维向量空间，其基为 $\mathcal{B}$。由于线性组合是有限和，故 $V \subseteq \mathcal{P}_0(\mathcal{B})$。根据定理 1.2，$\dim(V) \leq |\mathcal{P}_0(\mathcal{B})| = |\mathcal{B}|$。另一方面，由于每个 $f \in V^*$ 都可将每个 $v_i \in \mathcal{B}$ 映射独立映射到 0 和 1，因此 $\dim(V^*) = |\mathcal{P}(\mathcal{B})|$。根据定理 1.2，$\dim(V^*) > \dim(V)$。

定理 4.6

设 V 为向量空间，则 $\dim(V) \leq \dim(V^*)$，等号成立当且仅当 $\dim(V) < \infty$。

第 5 章 实内积空间与复内积空间

在本章中，我们仅聚焦于 F 是实数域 $\mathbb{R}$ 和复数域 $\mathbb{C}$ 的情形，将 $r \in \mathbb{C}$ 的复共轭记为 $\bar{r}$。在这样的向量空间上，我们定义了一个称为“内积”新的函数。

目录

5.1 内积、范数和距离 43
5.2 等距性与正交性 45
5.3 正交集与标准正交集 47
5.4 最佳逼近与投影定理 48
5.5 Riesz 表示定理 50

5.1 内积、范数和距离

5.1.1 内积

定义 5.1 (内积与内积空间)

设 V 为 F 上的向量空间，V 上的内积（inner product）是一个函数 $\langle \cdot, \cdot \rangle : V \times V \to F$，满足

1. 正定性（positive definiteness）：对于所有的 $v \in V$，均有 $\langle v, v \rangle \geq 0$，并且等号成立当且仅当 $v = 0$；
2. 对于 $F = \mathbb{C}$，有共轭对称性（conjugate symmetry）：$\langle u, v \rangle = \overline{\langle v, u \rangle}$；对于 $F = \mathbb{R}$，有对称性（symmetry）：$\langle u, v \rangle = \langle v, u \rangle$；
3. 关于首个坐标线性（linearity in the first coordinate）：对于所有 $u, v, w \in V$ 和 $r, s \in F$，均有

$$\langle ru + sv, w \rangle = r \langle u, w \rangle + s \langle v, w \rangle . \tag{5.1.1}$$

赋予内积的实（复）向量空间 V 称为一个实（复）内积空间（real/complex inner product space）。

注 若 $X, Y \subseteq V$，我们记 $\langle X, Y \rangle = \{\langle x, y \rangle \, | x \in X, y \in Y\}, expvalv, X = \{\langle v, x \rangle \, | x \in X\}$。

若将内积空间 V 上的内积限制在向量子空间 S 上，则 S 在此内积下也构成一个内积空间。若 $F = \mathbb{R}$，那么性质 2 和 3 意味着内积空间在各坐标上都是线性的，也就是说内积空间是双线性的（*bilinear*）；若 $F = \mathbb{C}$，则有

$$\langle w, ru + sv \rangle = \overline{\langle ru + sv, w \rangle} = \bar{r}\overline{\langle u, w \rangle} + \bar{s}\overline{\langle v, w \rangle} = \bar{r} \langle w, u \rangle + \bar{s} \langle w, v \rangle . \tag{5.1.2}$$

这称为在第二个坐标上的共轭线性性（*conjugate linearty*），也称半双线性性（*sesquilinearity*）。

例 5.1

1. 向量空间 $\mathbb{R}^n$ 在标准内积（*standard inner product*）或点积（*dot product*）下构成内积向量空间，其定义为

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n) \rangle = r_1 s_1 + \cdots + r_n s_n. \tag{5.1.3}$$

向量空间 $\mathbb{R}^n$ 常称为n 维 *Euclidean* 空间（*n-dimensional Euclidean space*）；

2. 向量空间 $\mathbb{C}^n$ 在标准内积

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n) \rangle = r_1 \overline{s_1} + \cdots + r_n \overline{s_n} \tag{5.1.4}$$

下构成一个内积空间，常称为n 维幺正空间（*n-dimensional unitary space*）；

3. 在闭区间 $[a, b]$ 上的所有连续复值函数构成一个向量空间 $\mathbb{C}[a, b]$。在内积

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)} \, \mathrm{d}x \tag{5.1.5}$$

下，该空间构成一个复内积空间。

例 5.2 最重要的内积空间之一是满足 $\sum_n |s_n|^2 < \infty$ 的全体实（复）数列 (s_n) 构成的向量空间 ℓ^2，内积定义为

$$\langle (s_n), (t_n) \rangle = \sum_{n=0}^{\infty} s_n \overline{t_n}. \tag{5.1.6}$$

这种序列称为平方可和的（*square summable*）。为了使这样的内积空间有意义，右侧的和必须收敛：

$$0 \le (|s_n| - |t_n|)^2 = |s_n|^2 + |t_n|^2 - 2|s_n||t_n| \quad \Longrightarrow \quad |s_n t_n| \le (|s_n|^2 + |t_n|^2)/2. \tag{5.1.7}$$

引理 5.1

若 V 是内积空间，并且 $\langle u, x \rangle = \langle v, x \rangle$ 对所有 $x \in V$ 成立，则 $u = v$。

♡

下一结论指出了实内积空间与复内积空间之间的主要差异直积，并将在后文起到关键作用。

命题 5.1

设 V 为内积空间，$\tau \in \mathcal{L}(V)$。

1. 若 $\langle \tau v, w \rangle = 0$ 对所有 $v, w \in V$ 成立，则 $\tau = 0$；
2. 若 V 是复内积空间，则 $\langle \tau v, v \rangle = 0$ 对所有 $v \in V$ 成立 $\Rightarrow \tau = 0$。

♠

证明

1. 由于 $\tau v \in V$，可取 $w = \tau v$，因此根据定义，$\langle \tau v, \tau v \rangle = 0$ 当且仅当 $\tau v = 0$ 对所有 $v \in V$ 成立，故 $\tau = 0$；
2. 设 $v = rx + y$，其中 $x, y \in V$ 且 $r \in F$，则

$$0 = \langle \tau v, v \rangle = \langle \tau(rx + y), rx + y \rangle = |r|^2 \langle \tau x, x \rangle + \langle \tau y, y \rangle + r \langle \tau x, y \rangle + \bar{r} \langle \tau y, x \rangle . \tag{5.1.8}$$

分别取 $r = 1$ 和 i，可得

$$\begin{cases} \langle \tau x, y \rangle + \langle \tau y, x \rangle = 0, \\ \langle \tau x, y \rangle - \langle \tau y, x \rangle \end{cases} \quad \Longrightarrow \quad \langle \tau x, y \rangle = \langle \tau y, x \rangle = 0 \text{ for all } x, y \in V \quad \Longrightarrow \quad \tau = 0. \tag{5.1.9}$$

5.1.2 范数

定义 5.2 (范数)

若 V 是内积空间，则 $v \in V$ 的范数（norm）或长度（length）定义为 $\|v\| = \sqrt{\langle v, v \rangle}$。若 $\|v\| = 1$，则向量 v 称为单位向量（unit vector）。

♣

以下是范数的基本性质。

定理 5.1

1. 正定性：对所有 $v \in V$，$\|v\| \ge 0$，并且等号成立当且仅当 $v = 0$；
2. 线性性：对所有 $r \in F$ 和 $v \in V$，均有 $\|rv\| = |r|\|v\|$；
3. 三角不等式：对所有 $u, v \in V$，均有 $\|u + v\| \le \|u\| + \|v\|$，等号成立当且仅当 u 是 v 的非负标量倍。

♡

在一般向量空间中，满足上述三条性质的函数称为其范数。我们将证明留给读者，其中三角不等式的证明在证明了下列性质中的 Cauchy-Schwarz 不等式后将迎刃而解。被赋予范数的向量空间称为赋范线性空间（*normed linear space*）。因此，在内积诱导的范数（定义 5.2）下，所有内积空间都是赋范线性空间。

命题 5.2

1. Cauchy-Schwarz 不等式：对所有 $u, v \in V$，$|\langle u, v \rangle| \le \|u\|\|v\|$，等号成立当且仅当 u 是 v 的标量倍；
2. 对所有 $u, v \in V$，均有 $|\|u\| - \|v\|| \le \|u - v\|$；
3. 平行四边形法则（the parallelogram law）：对所有 $u, v \in V$，均有 $\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$。

♠

证明 这里我们仅证明 Cauchy-Schwarz 不等式。对于任意 $u,v\in V$ 和 $r\in F$，考虑 $u-rv$ 的范数：

$$0\leq\|u-rv\|^2=\langle u-rv,u-rv\rangle=(\langle u,u\rangle-\bar{r}\langle u,v\rangle)-r(\langle v,u\rangle-\bar{r}\langle v,v\rangle). \tag{5.1.10}$$

取 $r=\langle u,v\rangle/\langle v,v\rangle$，则有

$$0\leq\langle u,u\rangle-|\langle u,v\rangle|^2/\langle v,v\rangle\quad\Longrightarrow\quad|\langle u,v\rangle|^2\leq\|u\|\|v\|, \tag{5.1.11}$$

等号成立当且仅当 $u=rv$，即 u 是 v 的标量倍。

一个有趣的事实是 V 上的内积可通过范数恢复，两种概念在实（复）内积空间内是等价的。

定理 5.2 (极化恒等式)

1. 若 V 为实向量空间，则
$$\langle u,v\rangle=\frac{1}{4}\left(\|u+v\|^2-\|u-v\|^2\right). \tag{5.1.12}$$
2. 若 V 为复向量空间，则
$$\langle u,v\rangle=\frac{1}{4}\left(\|u+v\|^2-\|u-v\|^2\right)+\frac{\mathrm{i}}{4}\left(\|u+\mathrm{i}v\|^2-\|u-\mathrm{i}v\|^2\right). \tag{5.1.13}$$

♡

5.1.3 距离

事实上，范数还可以用于定义内积空间中任意两个向量的距离。

定义 5.3 (距离)

设 V 为内积空间，V 中任意两个向量的距离（distance）$d(u,v)$ 定义为 $d(u,v)=\|u-v\|$。

♣

以下是距离的基本性质。

定理 5.3

1. 正定性：对于所有 $u,v\in V$，$d(u,v)\geq 0$，等号成立当且仅当 $u=v$；
2. 对称性：对于所有 $u,v\in V$，$d(u,v)=d(v,u)$；
3. 三角不等式：对于所有 $u,v,w\in V$，$d(u,v)\leq d(u,w)+d(w,v)$。

♡

对比第1.3.1小节的定义，容易看出 d 实际上是 V 上的度量，有序对 (V,d) 构成度量空间。因此在这样的距离定义下，所有内积空间都是度量空间。

5.2 等距性与正交性

5.2.1 等距映射

向量统建的同构保持向量空间的运算。对于被赋予拓扑结构的内积空间，其相应的概念是等距映射。

定义 5.4 (等距映射)

设 V 和 W 为内积空间，$\tau\in\mathcal{L}(V,W)$。

1. 若 τ 保持内积，即 $\langle\tau u,\tau v\rangle=\langle u,v\rangle$ 对所有 $u,v\in V$ 成立，则 τ 是等距映射（isometry）；
2. 双射等距映射称为等距同构（isometric isomorphism），此时称 V 和 W 等距同构。

♣

注 对于任意 $u\in\ker(\tau)$（τ 为等距映射），$\tau u=0$ 意味着 $\langle u,u\rangle=\langle\tau u,\tau u\rangle=0$，因此 $u=0$，即 $\ker(\tau)=\{0\}$。因此等距映射是单射，当其为满射时自然为等距同构。此外，若 $\dim(V)=\dim(W)<\infty$，根据推论 3.3，此时 τ 自然为等距同构。

由于范数确定了内积，我们自然有以下结论。

定理 5.4

线性变换 $\tau \in \mathcal{L}(V,W)$ 是等距映射当且仅当其保持范数，即当且仅当 $\|\tau v\| = \|v\|$ 对所有 $v \in V$ 成立。 ♡

证明 若 τ 为等距映射，取 $u = v$ 自然有 $\|\tau v\| = \|v\|$；反之根据极化恒等式（定理 5.2），在实内积空间中，

$$\langle \tau u, \tau v\rangle = \frac{1}{4}\left(\|\tau(u+v)\|^2 - \|\tau(u-v)\|^2\right) = \frac{1}{4}\left(\|u+v\|^2 - \|u-v\|^2\right) = \langle u, v\rangle , \tag{5.2.1}$$

此时 τ 为等距映射。复内积空间的情形类似。

5.2.2 正交性

定义 5.5

设 V 为内积空间。

1. 若两个向量 $u, v \in V$ 满足 $\langle u, v\rangle = 0$，则称两者正交（orthogonal），记作 $u \perp v$’
2. 若两个子集 $X, Y \subseteq V$ 满足 $\langle X, Y\rangle = \{0\}$，即 $x \perp y$ 对所有 $x \in X$ 和 $y \in Y$ 成立，则称 X 与 Y 正交，记作 $X \perp Y$；同时用 $v \perp X$ 代替 $\{v\} \perp X$ 的写法；
3. 子集 $X \subseteq V$ 的正交补（orthogonal complement）是集合 $X^\perp = \{v \in V | v \perp X\}$。 ♣

容易验证以下性质。

命题 5.3

设 V 为内积空间。

1. 任意子集 $X \subseteq V$ 的正交补 $X^\perp$ 是 V 的一个子空间；
2. 对于 V 的任意子空间 S，$S \cap S^\perp = \{0\}$。 ♠

证明

1. 对于任意 $u, v \in X^\perp$ 和 $r, s \in F$，$\langle ru + sv, X\rangle = r\langle u, X\rangle + s\langle v, X\rangle = 0$，因此 $ru + sv \in X^\perp$。根据命题 2.1，$X^\perp$ 为子空间；
2. 对于任意 $v \in S \cap S^\perp$，$\langle v, v\rangle = 0$，因此 $v = 0$，即 $S \cap S^\perp = \{0\}$。

定义 5.6 (正交直和)

若内积空间 $V = S \oplus T$ 且 $S \perp T$，则称 V 是 S 与 T 的正交直和（orthogonal direct sum），记作 $V = S \odot T$。更一般地，若

$$V = S_1 \oplus \cdots \oplus S_n \quad \text{and} \quad S_i \perp S_j \text{ for } i \neq j, \tag{5.2.2}$$

则称 V 为 $S_1, \cdots, S_n$ 的正交直和，记作 $V = S_1 \odot \cdots \odot S_n$。 ♣

正交直和具有以下等价描述。

定理 5.5

设 V 为内积空间，S 与 T 均为 V 的子空间，则下列表述等价：

1. $V = S \odot T$；
2. $V = S \oplus T$ 且 $T = S^\perp$。 ♡

证明 若 $V = S \odot T$，那么根据定义，$T \subseteq S^\perp$。对于任意 $v \in S^\perp \subseteq V$，存在 $s \in S$ 和 $t \in T$ 使得 $v = s + t$，因此 $0 = \langle s, v\rangle = \langle s, s\rangle + 0$，进而 $s = 0$，故 $v = t \in T$，反向包含证毕。由表述 2 证明表述 1 是显然的。

5.3 正交集与标准正交集

5.3.1 基本概念

定义 5.7 (正交集与标准正交集)

对于内积空间内的一个非空向量集 $\mathcal{O}=\{u_i|i\in I\}$，若 $u_i\perp u_j$ 对任意不相同的指标 $i,j\in I$ 成立，则称其为正交集（orthogonal set）。另外，若每个向量 u_i 都是单位向量，那么 $\mathcal{O}$ 是标准正交集（orthonormal set）。因此，若 $\langle u_i,u_j\rangle=\delta_{i,j}$ 对所有指标 $i,j\in I$ 成立，则称该集合标准正交，其中 $\delta_{i,j}$ 是 Kronecker δ-函数。 ♣

注 给定任意非零向量 $v\in V$，可通过对 v 乘上其倒数得到一个单位向量 $u=v/\|v\|$。该过程称为向量 v 的归一化（*normalization*）。因此，从非零向量的正交集获得标准正交集是一件容易的事。另外，若 $u\perp v$，则 $\|u+v\|^2=\|u\|^2+\|v\|^2$；当 $F=\mathbb{R}$ 时，其逆命题也成立。

命题 5.4

V 内由任意非零向量构成的正交集是线性无关的。 ♠

证明 考虑正交集 $\mathcal{O}=\{u_i|i\in I\}$ 以及线性组合 $\sum_{j=1}^n r_{i_j}u_{i_j}=0$，利用正交性可知 $0=\left\langle u_{i_k},\sum_{j=1}^n r_{i_j}u_{i_j}\right\rangle=r_{i_k}\langle u_{i_k},u_{i_k}\rangle$，即 $r_{i_k}=0$ 对所有 $k=1,\cdots,n$ 成立，因此 $\mathcal{O}$ 线性无关。

5.3.2 Gram-Schmidt 正交化

Gram-Schmidt 过程可用于将向量序列变换为正交序列。我们首先从以下定义入手。

定理 5.6 (Gram-Schmidt 增广)

设 V 为内积空间，$\mathcal{O}=\{u_1,\cdots,u_n\}$ 为 V 内一组正交向量集。若 $v\notin\langle u_1,\cdots,u_n\rangle$，则存在非零向量 $u\in V$ 使得 $\{u_1,\cdots,u_n,u\}$ 正交，且

$$\operatorname{span}\{u_1,\cdots,u_n,u\}=\operatorname{span}\{u_1,\cdots,u_n,v\}. \tag{5.3.1}$$

特别地，

$$u=v-\sum_{i=1}^n r_iu_i,\quad \text{where } r_i=\begin{cases}0, & \text{if } u_i=0,\\ \langle v,u_i\rangle/\langle u_i,u_i\rangle, & \text{if } u_i\neq 0.\end{cases} \tag{5.3.2}$$

♡

上述结论可直接验证(5.3.2)式证明。尽管 Gram-Schmidt 增广（augmentation）常用于一系列线性无关的向量，但事实上也能应用于任意向量序列。

定理 5.7

设 $\mathcal{B}=\{v_1,v_2,\cdots\}$ 是一个向量列，则可通过 Gram-Schmidt 增广定义序列 $\mathcal{O}=(u_1,u_2,\cdots)$：

$$u_k=v_k=\sum_{i=1}^{k-1}r_{k,i}u_i,\quad \text{where } u_1=v_1 \text{ and } r_{k,i}=\begin{cases}0, & \text{if } u_i=0,\\ \langle v_k,u_i\rangle/\langle u_i,u_i\rangle, & \text{if } u_i\neq 0.\end{cases} \tag{5.3.3}$$

这样的 $\mathcal{O}$ 构成 V 内一组正交序列，满足

$$\operatorname{span}\{u_1,\cdots,u_k\}=\operatorname{span}\{v_1,\cdots,v_k\} \tag{5.3.4}$$

对所有 $k>0$ 成立。同样地，$u_k=0$ 当且仅当 $v_k\in\operatorname{span}\{v_1,\cdots,v_{k-1}\}$。 ♡

证明 当 $k=1$ 时，该结论显然成立。假定该结论在 $k-1$ 时成立，若

$$v_k\in\operatorname{span}\{v_1,\cdots,v_{k-1}\}=\operatorname{span}\{u_1,\cdots,u_{k-1}\}, \tag{5.3.5}$$

那么存在 $a_i \in F$ 使得 $v_k = \sum_{i=1}^{k-1} a_i u_i$，进而 $\langle v_k, u_i\rangle = a_i \langle u_i, u_i\rangle$。因此当 $u_i \neq 0$ 时 $a_i = r_{k,i}$，即 $u_k = v_k - \sum_{i=1}^{k-1} r_{k,i} u_i = 0$，即

$$\operatorname{span}\{u_1, \cdots, u_k\} = \operatorname{span}\{u_1, \cdots, u_{k-1}, 0\} = \operatorname{span}\{v_1, \cdots, v_{k-1}\} = \operatorname{span}\{v_1, \cdots, v_k\}. \tag{5.3.6}$$

反之若 $v_k \notin \operatorname{span}\{v_1, \cdots, v_{k-1}\}$，则 Gram-Schmidt 增广过程（定理 5.6）已给出结果。综上，该结论对任意自然数 k 成立，从而可得到一条正交向量列。

5.3.3 Hilbert 基与 Hamel 基

定义 5.8 (Hilbert 基)

在内积空间内的极大标准正交集（maximal orthonormal set）称为 V 的 Hilber 基（Hilbert basis）。 ♣

Zorn 引理（定理 1.1）可用于证明任意非平凡内积空间具有 *Hilbert* 基：对于 V 内任意标准正交集构成的链 $\{E_\alpha\}$，可取其并集 $E = \cup_\alpha E_\alpha$。对于任意 $u, v \in E$，总存在 α, β 使得 $u \in E_\alpha, v \in E_\beta$。不妨设 $E_\alpha \subseteq E_\beta$，那么在 E_β 内，$\langle u, v\rangle = 0$，同时 $\|u\| = \|v\| = 1$。因此 E 为标准正交集，同时 E 为该链的上界。根据 Zorn 引理（定理 1.1），V 内含有极大标准正交集，即具有 Hilbert 基。

需要小心的是，不能将向量空间内的基与内积空间内的 *Hilbert* 基混淆。为了避免这种混淆，我们称前者为 *Hamel* 基（*Hamel basis*），称标准正交的 Hamel 基为标准正交基。

由于每个标准正交集都是线性无关的，因此标准正交基是 Hilbert 基，因为它不能真包含于其他正交集中；否则未包含的元素与标准正交基的并集构成新的线性无关集，这与 Hamel 基的极大性矛盾。对于有限维内积空间，这两种基是相同的。

定理 5.8

设 V 为内积空间，则 V 的有限子集 $\mathcal{O} = \{u_1, \cdots, u_k\}$ 是 V 的标准正交基当且仅当其为 V 的 Hilbert 基。 ♡

证明 对于充分性，假设存在 $v \in V$ 使得 $\mathcal{O} \cup \{v\}$ 构成一组标准正交集。由于 $\mathcal{O}$ 是 Hamel 基，存在 $r_i \in F$ 使得 $v = \sum_{i=1}^{k} r_i u_i$，进而 $0 = \langle v, u_i\rangle = r_i$ 对所有指标 $i = 1, \cdots, k$ 成立，因此 $v = 0$，即 $\mathcal{O}$ 为 Hilbert 基。

反之对于必要性，假设存在 $v \in V$ 使得 $\mathcal{O} \cup \{v\}$ 构成一组线性无关集。利用 Gram-Schmidt 增广过程（定理 5.6）可得到一组更大的标准正交集，但这与 $\mathcal{O}$ 的极大性矛盾，因此 $\mathcal{O}$ 为 Hamel 基。

下面这个例子说明了为什么上述结果在无穷维内积空间内会失效。

例 5.3 设 $V = \ell^2$ 且 M 为具有形如

$$e_i = (\underbrace{0, \cdots, 0}_{(i-1)\text{ terms}}, 1, 0, \cdots) \tag{5.3.7}$$

的所有向量构成的集合。显然 M 是一组正交集，若存在 $v = (x_n) \in \ell^2$ 且 $v \perp M$，则 $0 = \langle v, e_i\rangle = x_i$ 对所有指标 i 成立，这意味着 $v = 0$。因此 M 也是极大标准正交集，从而是内积空间 ℓ^2 的一组 Hilbert 基。

但另一方面，由 M 张成的向量空间是 ℓ^2 内所有具有有限支撑的序列构成的子空间 S，显然 $S \neq \ell^2$，例如 $(s_n) = (1/n)$ 在 ℓ^2 内但不在 S 内。

5.4 最佳逼近与投影定理

5.4.1 最佳逼近

标准正交基相较于任意基具有显著的实用优势。从计算的角度看，若 $\mathcal{B} = \{v_1, \cdots, v_n\}$ 是 V 的一组基，则每个 $v \in V$ 都具有形式 $v = r_1 v_1 + \cdots + r_n v_n$。一般而言，确定坐标 r_i 需要求解 $n \times n$ 规模的线性方程组。但如果 $\mathcal{O} = \{u_1, \cdots, u_n\}$ 是 V 的一组标准正交基，并且 $v = r_1 u_1 + \cdots + r_n u_n$，则系数 r_i 可直接通过 $\langle v, u_i\rangle$ 得到。

即使 $\mathcal{O}=\{u_1,\cdots,u_n\}$ 仅仅是任意标准正交集，我们仍可以考虑其展开：

$$\hat{v}=\langle v,u_1\rangle u_1+\cdots+\langle v,u_n\rangle u_n. \tag{5.4.1}$$

定理 5.9 (Fourier 展开与最佳逼近)

设 $\mathcal{O}=\{u_1,\cdots,u_k\}$ 为内积空间 V 的标准正交子集，子空间 $S=\langle\mathcal{O}\rangle$，则向量 $v\in V$ 关于 $\mathcal{O}$ 的 Fourier 展开（Fourier expansion）为

$$\hat{v}=\langle v,u_1\rangle u_1+\cdots+\langle v,u_k\rangle u_k, \tag{5.4.2}$$

每个系数 $\langle v,u_i\rangle$ 称为 v 关于 $\mathcal{O}$ 的 Fourier 系数（Fourier coefficient）。向量 $\hat{v}$ 可如下刻画：

1. $\hat{v}$ 是唯一满足 $(v-s)\perp S$ 的向量 $s\in S$；
2. $\hat{v}$ 为 S 内对 v 的最佳逼近（best approximation）；换言之，$\hat{v}$ 是在范数意义下唯一最接近 v 的向量 $s\in S$，即 $\|v-\hat{v}\|<\|v-s\|$ 对所有 $s\in S\backslash\{\hat{v}\}$ 成立；
3. Bessel 不等式对所有 $v\in V$ 成立，即 $\|\hat{v}\|\le\|v\|$。

♡

证明

1. 首先验证 $\langle v-\hat{v},u_i\rangle=\langle v,u_i\rangle-\langle v,u_i\rangle=0$，因此对于任意 $s\in S=\langle\mathcal{O}\rangle$，$\langle v-\hat{v},s\rangle=0$，即 $(v-\hat{v})\perp S$。设存在另一个 $s\in S$ 使得 $(v-s)\perp S$，由于 S 和 $S^\perp$ 均为子空间（命题 5.3），故 $s-\hat{v}\in S$，同时 $s-\hat{v}=(v-\hat{v})-(v-s)\in S^\perp$，从而 $s-\hat{v}\in S\cap S^\perp=\{0\}$，即 $s=\hat{v}$，唯一性得证；
2. 对于任意 $s\in S\backslash\{\hat{v}\}$，由于 $(v-\hat{v})\perp S$ 且 $s,\hat{v}\in S$，有
$$\|v-s\|^2=\|(v-\hat{v})+(\hat{v}-s)\|^2=\|v-\hat{v}\|^2+\|\hat{v}-s\|^2\ge\|v-\hat{v}\|^2, \tag{5.4.3}$$
等号成立当且仅当 $s=\hat{v}$，因此 $\|v-\hat{v}\|<\|v-s\|$ 对所有 $s\in S\backslash\{\hat{v}\}$ 成立；
3. 当 $\mathcal{O}$ 为 Hilbert 基时，假设存在 $v\in V$ 使得 $v\notin S$，那么可通过 Gram-Schmidt 增广过程（定理 5.6）得到一组新的标准正交基，但这与 $\mathcal{O}$ 的极大性矛盾。因此存在 $r_i\in F$ 使得 $v=\sum_{i=1}^k r_iu_i$ 对任意 $v\in V$ 成立，其中 $r_i=\langle v,u_i\rangle$。故 $v=\hat{v}$，自然有 $\|v\|=\|\hat{v}\|$。当 $\mathcal{O}$ 不为 Hilbert 基时，存在 $v\in V$ 满足 $v\notin S$，可通过 Gram-Schmidt 增广过程（定理 5.6）得到非零单位向量 $u_{k+1}=v-\sum_{i=1}^k r_iu_i$，其中 $r_i=\langle v,u_i\rangle$。因此 $u_{k+1}=v-\hat{v}\in S^\perp$，进而 $\|v\|^2=\|\hat{v}\|^2+\|v-\hat{v}\|^2>\|\hat{v}\|^2$。综上，$\|v\|\ge\|\hat{v}\|$ 对所有 $v\in V$ 成立。

5.4.2 投影定理

定理 5.10 (投影定理)

若 S 是内积空间 V 的有限维子空间，则 $V=S\odot S^\perp$，由此 $\dim(V)=\dim(S)+\dim(S^\perp)$。特别地，若 $v\in V$，则 $v=\hat{v}+(v-\hat{v})\in S\odot S^\perp$。

♡

证明 取 S 的一组有限基 $\mathcal{B}$，将其通过 Gram-Schmidt 增广过程（定理 5.6）得到一组标准正交基 $\mathcal{O}=\{u_1,\cdots,u_k\}$。再结合定理 5.9构造任意 $v\in V$ 的 Fourier 展开 $\hat{v}$，可知 $(v-\hat{v})\in S^\perp$，因此 $V=S+S^\perp$。根据命题 5.3和定理 5.5，$V=S\odot S^\perp$。维数关系源于推论 4.2。

当 S 不是有限维时，上述结论未必成立。不过即使在无穷维情形，只要 S 是完备子空间，上述结论依然成立。关于这部分的讨论，我们放到第8章。

例 5.4 仍采用例5.3中的例子。设 $S=\mathrm{span}(M)$，那么若 $x=(x_n)\in S^\perp$，则 $0=\langle x,e_i\rangle=x_i$ 对所有指标 i 成立，从而 $x=0$，即 $S^\perp=\{0\}$；然而 $S\odot S^\perp=S\neq\ell^2$，可见投影定理失效。

命题 5.5

设 V 为内积空间，S 为 V 的一个有限维子空间。

1. $S^{\perp\perp}=S$；
2. 若 $X\subseteq V$ 且 $\dim(\langle X\rangle)<\infty$，则 $X^{\perp\perp}=\langle X\rangle$。

♠

证明

1. 显然对于任意 $s \in S$，$s \perp S^\perp$，因此 $S \subseteq S^{\perp\perp}$；若 $v \in S^{\perp\perp} \subseteq V$，根据投影定理 5.10，存在 $s \in S$ 和 $t \in S^\perp$ 使得 $v = s + t$。那么 $0 = \langle v, t\rangle = \langle t, t\rangle$，进而 $t = 0$。可见 $v = s \in S$，即 $S^{\perp\perp} \subseteq S$。综上，$S^{\perp\perp} = S$；
2. 根据上述结果，只要证明 $X^\perp = \langle X\rangle^\perp$ 则结论得证。显然 $\langle X\rangle^\perp \subseteq X^\perp$；对于反向包含，由于 $X^\perp$ 是 V 的子空间（命题 5.3），考虑 $v \in X^\perp$ 以及 $\langle X\rangle$ 的一组有限基 $\mathcal{B}$。由于 $v \perp X$，显然 $v \perp \mathcal{B}$，进而 $v \perp \langle\mathcal{B}\rangle = \langle X\rangle$，可见 $X^\perp \subseteq \langle X\rangle^\perp$。综上，$X^\perp = \langle X\rangle^\perp$，结论得证。

5.4.3 Hilbert 基的刻画

命题 5.6

设 $\mathcal{O} = \{u_1, \cdots, u_k\}$ 为内积空间 V 的标准正交子集，子空间 $S = \langle\mathcal{O}\rangle$，则下列表述等价：

1. $\mathcal{O}$ 为 V 的标准正交基；
2. $S^\perp = \langle\mathcal{O}\rangle^\perp = \{0\}$；
3. 每个向量等于其自身的 Fourier 展开，即对于所有 $v \in V$，均有 $\hat{v} = v$；
4. Bessel 恒等式对所有 $v \in V$ 成立，即 $\|\hat{v}\| = \|v\|$；
5. Parseval 恒等式对所有 $v, w \in V$ 成立，即

$$\langle v, w\rangle = [\hat{v}]_{\mathcal{O}} \cdot [\hat{w}]_{\mathcal{O}} = \langle v, u_1\rangle \overline{\langle w, u_1\rangle} + \cdots + \langle v, u_k\rangle \overline{\langle w, u_k\rangle}. \tag{5.4.4}$$

证明

1. 若表述 1 成立，则定理 5.9中结论 3 的证明已表明，对于任意 $v \in S^\perp \subseteq V$，存在 $r_i \in F$ 使得 $v = \sum_{i=1}^k r_i u_i$，因此 $\langle v, u_i\rangle = 0$ 对所有指标 $i = 1, \cdots, k$ 成立。故 $v = 0$，即 $S^\perp = \{0\}$，表述 2 成立；
2. 若表述 2 成立，根据定理 5.9，$(v - \hat{v}) \in S^\perp = \{0\}$，因此 $\hat{v} = v$，表述 3 成立；
3. 若表述 3 成立，表述 4 显然成立；
4. 若表述 4 成立，那么对于实内积空间内的任意向量 v，

$$\|v\|^2 = \|\hat{v}\|^2 = \left\|\sum_{i=1}^k \langle v, u_i\rangle u_i\right\|^2 = \sum_{i=1}^k \langle v, u_i\rangle \langle u_i, v\rangle . \tag{5.4.5}$$

因此利用极化恒等式（定理 5.2），

$$\begin{aligned}\langle v, w\rangle &= \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2\right)\\ &= \frac{1}{4}\sum_{i=1}^k \left(\langle v + w, u_i\rangle \langle u_i, v + w\rangle - \langle v - w, u_i\rangle \langle u_i, v - w\rangle\right)\\ &= \frac{1}{2}\sum_{i=1}^k \left(\langle w, u_i\rangle \langle u_j, v\rangle + \langle v, u_i\rangle \langle u_j, w\rangle\right)\\ &= \sum_{i=1}^k \langle w, u_i\rangle \langle u_j, v\rangle = \langle v, u_1\rangle \overline{\langle w, u_1\rangle} + \cdots + \langle v, u_k\rangle \overline{\langle w, u_k\rangle}.\end{aligned} \tag{5.4.6}$$

复内积空间的情形类似，故表述 1 成立；

5. 若表述 5 成立，假设存在 $v \in V$ 使得 $\mathcal{O} \cup \{v\}$ 为标准正交集，那么 $0 = \langle v, u_i\rangle$ 对所有指标 $i = 1, \cdots, k$ 成立，进而对于任意 $w \in V$，

$$\langle v, w\rangle = \langle v, u_1\rangle \overline{\langle w, u_1\rangle} + \cdots + \langle v, u_k\rangle \overline{\langle w, u_k\rangle} = 0, \tag{5.4.7}$$

即 $v \perp V$，这要求 $v = 0$，否则 $\langle v, v\rangle \neq 0$ 推出矛盾。因此 $\mathcal{O}$ 为极大标准正交集，即标准正交基。

5.5 Riesz 表示定理

我们在本章初提及复向量空间上的内积关于第二个坐标具有共轭线性，在此我们给出其严格化的一般定义。

定义 5.9 (共轭线性)

对于定义在复向量空间 V 和 W 上的函数 $\sigma : V \to W$，若其满足

$$\sigma(v_1 + v_2) = \sigma(v_1) + \sigma(v_2) \quad \text{and} \quad \sigma(rv) = \bar{r}\sigma(v), \tag{5.5.1}$$

其中 $r \in \mathbb{C}$ 且 $v_1, v_2, v \in V$，则称其共轭线性（conjugate linear）。共轭同构（conjugate isomorphism）是双射共轭线性映射。

♣

对于给定的 $x \in V$，内积函数 $\langle \cdot, x\rangle : V \to F$ 定义为 $\langle \cdot, x\rangle v = \langle v, x\rangle$，该函数是 V 上的线性泛函。因此，若定义线性映射 $\tau : V \to V^*$ 为 $\tau x = \langle \cdot, x\rangle$，则其为共轭线性的。

定理 5.11 (Riesz 表示定理)

设 V 为有限维内积空间。

1. 若定义映射 $\tau : V \to V^*$ 为 $\tau x = \langle \cdot, x\rangle$，则其为共轭同构。特别地，对于每个 $f \in V^*$，存在唯一的向量 $x \in V$ 使得 $f = \langle \cdot, x\rangle$，即 $fv = \langle v, x\rangle$；
2. 若定义映射 $R : V^* \to V$ 为 $Rf = R_f$，则其为 τ 的逆，因而同样为共轭同构。我们称此映射为 Riesz 映射（Riesz map）。若 $V = \mathbb{R}^n$，则 $R_f = (f(e_1), \cdots, f(e_n))$，其中 $e_1, \cdots, e_n$ 为 $\mathbb{R}^n$ 的标准基。

♡

证明

1. 前已证明 τx 共轭线性，根据引理 5.1，$\tau x = \tau y$ 即 $\langle \cdot, x\rangle = \langle \cdot, y\rangle$ 意味着 $x = y$，因此 τ 为单射。对于任意 $f \in V^*$，根据命题 4.1，$K = \ker(f)$ 的余维数为 1。根据投影定理 5.10，对于任意给定的非零向量 $w \in V$，$V = \langle w\rangle \odot K$，这要求 $w \in K^\perp$。取 $x = \alpha w$，其中 $\alpha \in F$，我们希望对于任意 $v \in V$ 证明 $f(v) = \langle v, \alpha w\rangle$。若 $v \in K$，则 $0 = f(v) = \bar{\alpha}\langle v, w\rangle = \langle v, \alpha w\rangle$；若 $v \in K^\perp$，则存在 $r \in F$ 使得 $v = rw$，仅需证明 $f(w) = \langle w, \alpha w\rangle$。不妨取 $\alpha = \overline{f(w)}/\langle w, w\rangle$，则上式成立。综上，对于每个 $f \in V^*$，总可以选择非零向量 $w \in V$ 和 $\alpha = \overline{f(w)}/\|w\|^2 \in F$，使得 $V = \langle w\rangle \odot K$。那么对于任意 $v \in V$，总存在 $r \in F$ 和 $u \in K$，使得 $v = rw + u$，进而 $f(v) = rf(w) + f(u) = \bar{\alpha}r\langle w, w\rangle = \langle rw, \alpha w\rangle = \langle v, x\rangle$，其中 $x = \alpha w$ 唯一。
2. 对于上述 $f \in V^*$ 和非零向量 w，可定义 $Rf = R_f = x = \alpha w$，首先验证其共轭线性：

$$\langle v, R_{rf+sg}\rangle = \left\langle v, \overline{rf(w) + sg(w)}(w/\|w\|^2)\right\rangle = \langle v, \bar{r}R_f + \bar{s}R_g\rangle \tag{5.5.2}$$

对任意 $v \in V$、$r \in F$ 和 $f, g \in V^*$ 成立，因此根据引理 5.1，$R_{rf+sg} = \bar{r}R_f + \bar{s}R_g$，故 R 具有共轭线性。另外，对于任意 $v \in V$，$(R \circ \tau)v = R\langle \cdot, v\rangle = v$；对于任意 $f \in V^*$，存在 $x \in V$ 使得 $f = \langle \cdot, x\rangle$，因此 $(\tau \circ R)f = \tau x = \langle \cdot, x\rangle = f$。可见 R 是 τ 的逆映射，因而也是共轭同构，证毕。

注 事实上，在量子力学的 Dirac 符号体系下，所有左矢 $\langle \cdot|$ 构成的空间正是复内积空间的对偶空间；而每个左矢作为线性泛函，根据此定理可用唯一的形式化的态矢量表示，它将任意右矢映射到某个确定的内积值。简而言之，在常用的幺正空间 $\mathbb{C}^n$ 内，共轭同构 τ 和 R 均为共轭转置操作。

第 6 章 本征值与本征向量

由于线性算子 $\tau \in \mathcal{L}(V)$ 的特征多项式背后的动机涉及较为复杂的模论（参见 [1]），我们将采用与基本线性代数相似的思路：设 $A = [\tau]_{\mathcal{B}}$ 为 τ 在有序基 $\mathcal{B}$ 下的坐标矩阵，将行列式 $\det(xI - A)$ 视为 τ 的特征多项式 $c_\tau(x)$ 或 $c_A(x)$ 的定义。由此出发，我们简单介绍本征值与本征向量的概念，对于涉及模论的部分则专注于结论、不纠结于证明细节。在本章中，所有向量空间均假定为有限维的。

目录

6.1 本征值与本征向量 52
6.2 迹与行列式 53
6.3 可对角化算子 54

6.1 本征值与本征向量

根据特征多项式的定义 $c_A(x) = \det(xI - A)$，标量 $\lambda \in F$ 是 $c_A(x)$ 的根当且仅当 $\det(\lambda I - A) = 0$。特别地，若 $\dim(V) = n$，则上式成立当且仅当 $(\lambda I - A)x = 0$，或者等价地，$\tau_A x = \lambda x$，这里 x 是 n 维列向量。若根据 $[v]_{\mathcal{B}} = x$ 定义向量 $v \in V$，则 $\lambda[v]_{\mathcal{B}} = [\tau]_{\mathcal{B}}[v]_{\mathcal{B}} = [\tau v]_{\mathcal{B}}$。这引出了如下定义。

定义 6.1 (本征值与本征向量)

设 V 为域 F 上的向量空间，$\tau \in \mathcal{L}(V)$。

1. 若存在非零向量 $v \in V$，使得标量 $\lambda \in F$ 满足 $\tau v = \lambda v$，则称 λ 为 τ 的本征值（eigenvalue）或特征值（characteristic value）。在这种情况下，v 称为 τ 的与 λ 相关的本征向量（eigenvector）或特征向量（characteristic vector）；
2. 若存在非零列向量 x，使得标量 $\lambda \in F$ 满足 $Ax = \lambda x$，则称 λ 为矩阵 A 的本征值。在这种情况下，x 称为 A 的与 λ 相关的本征向量或特征向量；
3. 与给定的本征值 λ 相关的所有本征向量集合（包括零向量）构成 V 的一个子空间，称为 λ 的本征空间（eigenspace），记为 $\mathcal{E}_\lambda$。此概念适用于线性算子和矩阵；
4. 算子或矩阵的全体本征值构成的集合称为算子或矩阵的谱（spectrum），记作 $\mathrm{Spec}(\tau)$。

♣

根据上述定义，可以得到以下结果。

命题 6.1

设 $\lambda_1, \cdots, \lambda_k$ 为线性算子 $\tau \in \mathcal{L}(V)$ 的不同本征值。

1. 与不同本征值相关联的本征向量线性无关，即如果非零的 $v_i \in \mathcal{E}_{\lambda_i}$，则集合 $\{v_1, \cdots, v_k\}$ 线性无关；
2. 子空间和 $\mathcal{E}_{\lambda_1} + \cdots + \mathcal{E}_{\lambda_k}$ 是直和，即 $\mathcal{E}_{\lambda_i} \cap \mathcal{E}_{\lambda_j} = \{0\}$ 对所有 $i \neq j$ 成立。

♠

证明

1. 假设 $\{v_1, \cdots, v_k\}$ 线性相关，考虑其所有非平凡线性组合 $r_{i_1}v_{i_1} + \cdots + r_{i_j}v_{i_j} = 0$，其中 $r_{i_1}, \cdots, r_{i_j} \in F$。对于满足上述条件的全体自然数 i_j 构成的集合，根据定理 1.7，其中必然存在一个最小自然数 i_j。将线性算子 τ 作用在方程两端，可得 $r_{i_1}\lambda_{i_1}v_{i_1} + r_{i_j}\lambda_{i_j}v_{i_j} = 0$。两式消去 r_{i_1} 得 $r_{i_2}(\lambda_{i_2} - \lambda_{i_1}) + \cdots + r_{i_j}(\lambda_{i_j} - \lambda_{i_1})v_{i_j} = 0$。但显然该方程是具有更少项的非平凡线性组合，矛盾。因此 $\{v_1, \cdots, v_k\}$ 线性无关；
2. 当 $i \neq j$ 时，设 $v \in \mathcal{E}_{\lambda_i} \cap \mathcal{E}_{\lambda_j}$，那么 $\lambda_i v = \tau v = \lambda_j v$，而 $\lambda_i \neq \lambda_j$，故 $v = 0$，即 $\mathcal{E}_{\lambda_i} \cap \mathcal{E}_{\lambda_j} = \{0\}$。

定理 6.1 (谱映射定理)

设 V 为代数闭包 F 上的一个向量空间，$\tau \in \mathcal{L}(V)$ 且 $p(x) \in F[x]$，那么

$$\operatorname{Spec}(p(\tau)) = p(\operatorname{Spec}(\tau)) = \{p(\lambda) | \lambda \in \operatorname{Spec}(\tau)\}. \tag{6.1.1}$$

♡

证明 对于任意 $\lambda \in \operatorname{Spec}(\tau)$ 和 n 次多项式 $p(x) = \sum_{k=0}^{n} a_k x^k \in F[x]$，利用归纳法容易证明 $\tau^k v = \lambda^k v$ 对所有整数 $k \geq 1$ 成立，因此 $p(\tau)v = p(\lambda)v$，即 $p(\lambda) \in \operatorname{Spec}(p(\tau))$。这意味着 $p(\operatorname{Spec}(\tau)) \subseteq \operatorname{Spec}(p(\tau))$；对于反向包含，设 $\lambda \in \operatorname{Spec}(p(\tau))$，则 $(p(\tau) - \lambda)v = 0$。由于 F 是代数闭包，总可将 $p(x)$ 分解为

$$p(x) - \lambda = (x - r_1)^{e_1} \cdots (x - r_k)^{e_k}, \tag{6.1.2}$$

其中 $r_i \in F$，正整数 e_i 的和为 $\deg(p(x)) = n$。对于 $p(\tau) - \lambda$ 分解得到的一系列算子，我们可从最右侧的算子 $(x - r_k)$ 开始，逐个将其左侧相邻的算子并入。如此至多 $(n-1)$ 次操作后，必然存在某个向量 $u = (\tau - r_i)^{f_i}(\tau - r_{i+1})^{e_{i+1}} \cdots (\tau - r_k)^{e_k} v$ $(f_i \leq e_i)$ 使得 $u \neq 0$ 但 $(\tau - r_i)u = 0$ $(f_i < e_i)$ 或 $(\tau - r_{i-1})u = 0$ $(f_i = e_i)$，否则 $(p(\tau) - \lambda)v \neq 0$，矛盾。将 r_i 或 r_{i-1} 统一用 r 表示，那么 r 是 τ 的本征值，进而 $\lambda = p(r) \in p(\operatorname{Spec}(\tau))$，得证。

6.2 迹与行列式

定义 6.2 (矩阵的迹)

矩阵 $A \in \mathcal{M}_n(F)$ 的迹（trace）为 A 的主对角元上的元素，记作 $\operatorname{tr}(A)$。

♣

以下是迹的基本性质。

命题 6.2

设 $A, B \in \mathcal{M}_n(F)$。

1. 对于任意 $r \in F$，$\operatorname{tr}(rA) = r \operatorname{tr}(A)$；
2. $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$；
3. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$；
4. $\operatorname{tr}(ABC) = \operatorname{tr}(CAB) = \operatorname{tr}(BCA)$；
5. 迹在相似变换下不变；
6. 若 F 是代数闭包，则 $\operatorname{tr}(A)$ 是 A 的本征值之和（包含重数）。若其特征多项式表示为 $c_A(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1(x) + c_0$，则 $\operatorname{tr}(A) = -c_{n-1}$。

♠

证明 这里仅证明最后两条性质，其余性质可由矩阵加法、矩阵乘法和标量乘法法则直接验证。对于矩阵 $A \in \mathcal{M}_n(F)$ 和任意可逆矩阵 $P \in \mathcal{M}_n(F)$，容易验证

$$\begin{aligned} \operatorname{tr}(PAP^{-1}) &= \sum_i (PAP^{-1})_{ii} = \sum_{i,j,k} P_{i,j} A_{j,k} (P^{-1})_{k,i} \\ &= \sum_{j,k} A_{j,k} \sum_i (P^{-1})_{k,i} P_{i,j} = \sum_{j,k} \delta_{j,k} A_{j,k} = \operatorname{tr}(A). \end{aligned} \tag{6.2.1}$$

设 $A = (a_{i,j}) \in \mathcal{M}_n(F)$，那么根据行列式的展开规则，除去主对角线上元素的连乘，其余元素的乘积得到的多项式至多为 $(n-2)$ 次。因此 $c_A(x)$ 中 x^{n-1} 项的贡献完全源于表达式 $(x - a_{1,1}) \cdots (x - a_{n,n})$ 中 x^{n-1} 项，即 $c_{n-1} = -\sum_{i=1}^{n} a_{i,i} = -\operatorname{tr}(A)$；另一方面，由于 F 是代数闭包，可将 $c_A(x)$ 表示为 $(x - \lambda_1) \cdots (x - \lambda_n)$，其中 λ_i 为 τ 的本征值，因此 $c_{n-1} = -\sum_{i=1}^{n} \lambda_i$，结论得证。

由于迹在相似变换下不变，我们可引入如下定义。

定义 6.3 (算子的迹)

线性算子 $\tau \in \mathcal{L}(V)$ 的迹为任意可表示 τ 的矩阵的迹。

♣

6.3 可对角化算子

定义 6.4 (可对角化)

如果对于线性算子 $\tau \in \mathcal{L}(V)$，存在 V 的有序基 $\mathcal{B} = (v_1, \cdots, v_n)$ 使得矩阵 $[\tau]_{\mathcal{B}}$ 是对角的，或者等价地，$\tau v_i = \lambda v_i$ 对所有 $i = 1, \cdots, n$ 成立，则称 τ 是可对角化的（diagonalizable）。 ♣

根据上述定义，容易看出可对角化的性质存在以下等价刻画。

命题 6.3

设 $\tau \in \mathcal{L}(V)$，则下列表述等价：

1. τ 是可对角化的；
2. V 具有一组完全由 τ 的本征向量构成的一组基；
3. V 具有直和分解 $V = \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_k}$，其中 $\lambda_1, \cdots, \lambda_k$ 是 τ 的不同本征值。 ♠

我们已经在第3.6.2小节讨论过，向量空间 V 上单位算子的分解对应于 V 的直和分解。对于 V 上任意可对角化的线性算子 τ，我们有类似结论。

定理 6.2 (谱分解)

线性算子 $\tau \in \mathcal{L}(V)$ 可对角化当且仅当其具有谱分解（spectral resolution）：

$$\tau = \lambda_1 \rho_1 + \cdots + \lambda_k \rho_k, \tag{6.3.1}$$

其中 $\rho_1 + \cdots + \rho_k = \iota$ 是单位算子的分解，$\lambda_i \in F$ 互不相同。在此情形下，$\{\lambda_1, \cdots, \lambda_k\}$ 是 τ 的谱，并且

$$\mathrm{im}(\rho_i) = \mathcal{E}_{\lambda_i} \quad \text{and} \quad \ker(\rho_i) = \oplus_{j \neq i} \mathcal{E}_{\lambda_j}. \tag{6.3.2}$$

♡

证明 对于必要性，定理 3.11表明 $V = \mathrm{im}(\rho_1) \oplus \cdots \oplus \mathrm{im}(\rho_k)$，其中 ρ_i 是沿 $\oplus_{j \neq i} \mathrm{im}(\rho_j)$ 到 $\mathrm{im}(\rho_i)$ 上的投影。任意取定非零向量 $v \in V$，可得到 $\tau(\rho_i v) = \lambda_i \rho_i (\rho_i v) = \lambda_i (\rho_i v)$。可见 $\rho_i v \in \mathcal{E}_{\lambda_i}$，即 $\mathrm{im}(\rho_i) \subseteq \mathcal{E}_{\lambda_i}$，因此根据命题 6.1，

$$V = \mathrm{im}(\rho_1) \oplus \cdots \oplus \mathrm{im}(\rho_k) \subseteq \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_k} \subseteq V. \tag{6.3.3}$$

故 $V = \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_k}$，根据命题 6.3，$\tau$ 可对角化。

反之对于充分性，若 $V = \mathcal{E}_{\lambda_1} \oplus \cdots \oplus \mathcal{E}_{\lambda_k}$ 且 λ_i 互不相同，那么将 ρ_i 取为沿 $\oplus_{j \neq i} \mathcal{E}_{\lambda_j}$ 到 $\mathcal{E}_{\lambda_i}$ 上的投影算子。根据定理 3.11，$\rho_1 + \cdots + \rho_k = \iota$ 是单位算子的分解。对于任意 $v \in V$，由于 $\rho_i v \in \mathrm{im}(\rho_i) = \mathcal{E}_{\lambda_i}$，容易证明

$$\tau v = \tau(\rho_1 + \cdots + \rho_k) v = \lambda_1 (\rho_1 v) + \cdots + \lambda_k (\rho_k v) = (\lambda_1 \rho_1 + \cdots + \lambda_k \rho_k) v, \tag{6.3.4}$$

因此 $\tau = \lambda_1 \rho_1 + \cdots + \lambda_k \rho_k$，其中 λ_i 互不相同。

第 7 章 正规算子的结构理论

在本章中，若不作特别说明，所有向量空间均假定为有限维的，此外域 F 为 $\mathbb{R}$ 或 $\mathbb{C}$。

目录

7.1 线性算子的伴随 55
7.2 正交投影 57
7.3 正规算子 58
7.4 泛函演算 62
7.5 正算子与极分解 64

7.1 线性算子的伴随

定理 7.1 (算子伴随)

设 V 和 W 为 F 上的有限维内积空间，$\tau \in \mathcal{L}(V,W)$，那么对于所有 $v \in V$ 和 $w \in W$，存在唯一的函数 $\tau^* : W \to V$ 满足

$$\langle \tau v, w\rangle = \langle v, \tau^* w\rangle . \tag{7.1.1}$$

该函数在 $\mathcal{L}(W,V)$ 内，称为 τ 的伴随（adjoint）。♡

证明 我们首先寻找满足条件的线性映射 $\tau^* : W \to V$。若 $\tau^* w$ 存在，则其形式上接近将任意 $v \in V$ 映射到 $\langle \tau v, w\rangle$ 的线性泛函。根据 Riesz 表示定理 5.11，任何线性泛函均可用 Riesz 向量表示。具体地，对于每个 $w \in W$，定义线性泛函 $f_w \in V^*$ 为 $f_w v = \langle \tau v, w\rangle$。那么根据 Riesz 表示定理 5.11，存在 Riesz 向量 $R_{f_w} \in V$ 使得 $f_w v = \langle v, R_{f_w}\rangle$。记 Riesz 映射为 R，那么定义线性变换 $\tau^* : W \to V$ 为

$$\tau^* w = R_{f_w} = R(f_w) \quad \Longrightarrow \quad \langle v, \tau^* w\rangle = \langle v, R_{f_w}\rangle = f_w v = \langle \tau v, w\rangle , \tag{7.1.2}$$

其线性性源于 Riesz 映射的线性性。最后验证其唯一性：假设存在满足条件的另一线性变换 $\sigma : W \to V$，则 $\langle v, (\tau^* - \sigma)w\rangle = 0$ 对所有 $v, w \in V$ 成立，即 $(\tau^* - \sigma)w = 0$ 对所有 $w \in V$ 成立，因此 $\sigma = \tau^*$。

注 在量子力学的 Dirac 符号体系下，算子伴随对应算子的共轭转置，即 $(|\psi\rangle , A|\varphi\rangle) = \langle \psi|A|\varphi\rangle = (A^\dagger |\psi\rangle , |\varphi\rangle)$。

以下是算子伴随的基本性质。

命题 7.1

设 V 和 W 为有限维内积空间。对于每个 $\sigma, \tau \in \mathcal{L}(V,W)$ 和 $r \in F$，

1. $(\sigma + \tau)^* = \sigma^* + \tau^*$；
2. $(r\tau)^* = \bar{r}\tau^*$；
3. $\tau^{**} = \tau$ 且 $\langle \tau^* v, w\rangle = \langle v, \tau w\rangle$；
4. 若 $V = W$，则 $(\sigma\tau)^* = \tau^*\sigma^*$；
5. 若 τ 可逆，则 $(\tau^{-1})^* = (\tau^*)^{-1}$；
6. 若 $V = W$ 且 $p(x) = F[x]$，则 $p(\tau)^* = p(\tau^*)$。

此外，若 $\tau \in \mathcal{L}(V)$ 且 S 为 V 的子空间，则

1. S 是 τ-不变的当且仅当 $S^\perp$ 是 τ^* 不变的；
2. $(S, S^\perp)$ 约化 τ 当且仅当 $S^\perp$ 同时是 τ-不变和 τ^*-不变的，此时 $(\tau|_S)^* = (\tau^*)|_S$。♠

证明 我们仅证明最后两条性质。若 S 是 τ-不变的，那么对于任意 $v\in S$ 和 $w\in S^\perp$，$\tau v\in S$，因此

$$0=\langle v,\tau w\rangle=\langle \tau^* w,v\rangle \text{ for all } v\in S \implies \tau w^*\in S^\perp \text{ for all } w^*\in S^\perp \implies \tau^* S^\perp\subseteq S^\perp. \tag{7.1.3}$$

反之若 $S^\perp$ 是 τ^* 不变的，类似可证明该结论。

若 $(S,S^\perp)$ 约化 τ，那么对于任意 $v\in S$ 和 $w\in S^\perp$，$\tau v\in S$ 且 $\tau w\in S^\perp$，进而

$$0=\langle \tau v,w\rangle=\langle v,\tau^* w\rangle \text{ for all } v\in S \implies \tau^* w\in S^\perp \text{ for all } w\in S^\perp \implies \tau^* S^\perp\subseteq S^\perp. \tag{7.1.4}$$

反之若 $S^\perp$ 是 τ-不变且 τ^*-不变的，那么 $0=\langle \tau^* w,v\rangle=\langle w,\tau v\rangle$，即 $\tau v\in S^{\perp\perp}=S$（命题 5.5）对所有 $v\in S$ 成立，因此 S 是 τ-不变的。对于后一论述，考虑 $s,t\in S$，有

$$\langle s,(\tau^*)|_S t\rangle=\langle s,\tau^* t\rangle=\langle \tau s,t\rangle=\langle (\tau|_S)s,t\rangle=\langle s,(\tau|_S)^* t\rangle. \tag{7.1.5}$$

利用引理 5.1，$(\tau^*)|_S t=(\tau|_S)^* t$ 对所有 $t\in S$ 成立，故 $(\tau^*)|_S=(\tau|_S)^*$。

现在我们将线性变换的核与像与其伴随关联起来。

命题 7.2

设 $\tau\in\mathcal{L}(V,W)$，其中 V 和 W 为有限维内积空间。

1. $\ker(\tau^*)=\operatorname{im}(\tau)^\perp$ 且 $\operatorname{im}(\tau^*)=\ker(\tau)^\perp$，进而 τ 单射（满射）$\Leftrightarrow$ τ^* 满射（单射）；
2. $\ker(\tau^*\tau)=\ker(\tau)$ 且 $\ker(\tau\tau^*)=\ker(\tau^*)$；
3. $\operatorname{im}(\tau^*\tau)=\operatorname{im}(\tau^*)$ 且 $\operatorname{im}(\tau\tau^*)=\operatorname{im}(\tau)$；
4. $(\rho_{S,T})^*=\rho_{T^\perp,S^\perp}$。

证明

1. 根据定义，
$$\begin{aligned}\ker(\tau^*)&=\{u\in V|\tau^* u=0\}=\{u\in V|\langle \tau V,u\rangle=\langle V,\tau^* u\rangle=\{0\}\}\\&=\{u\in V|u\in\operatorname{im}(\tau)^\perp\}=\operatorname{im}(\tau)^\perp.\end{aligned} \tag{7.1.6}$$
根据命题 5.5，$\ker(\tau^*)^\perp\subseteq\operatorname{im}(\tau)^{\perp\perp}=\operatorname{im}(\tau)$，并且 $\operatorname{im}(\tau)\subseteq\ker(\tau^*)^\perp$，故 $\operatorname{im}(\tau^*)=\ker(\tau)^\perp$；
2. 显然 $\ker(\tau)\subseteq\ker(\tau^*\tau)$；对于任意 $v\in\ker(\tau^*\tau)$，$0=\langle v,(\tau^*\tau)v\rangle=\langle \tau v,\tau v\rangle$，因此 $\tau v=0$，即 $v\in\ker(\tau)$。可见反向包含成立，$\ker(\tau\tau^*)=\ker(\tau)$。根据命题 7.1，$\ker(\tau^*)=\ker((\tau^*)^*\tau^*)=\ker(\tau\tau^*)$；
3. 结合前两条结论与 $(\tau^*\tau)^*=\tau^*\tau$ 的事实，
$$\operatorname{im}(\tau^*\tau)=\ker(\tau^*\tau)^\perp=\ker(\tau)^\perp=\operatorname{im}(\tau^*),\quad \operatorname{im}(\tau\tau^*)=\operatorname{im}((\tau^*)^*)=\operatorname{im}(\tau). \tag{7.1.7}$$
4. 根据性质 1 和命题 3.6，$\ker((\rho_{S,T})^*)=\operatorname{im}(\rho_{S,T})^\perp=S^\perp$ 且 $\operatorname{im}((\rho_{S,T})^*)=\ker(\rho_{S,T})^\perp=T^\perp$。另外，对于任意 $v\in V$ 和 $w\in W$，根据投影算子的幂等性（定理 3.10），
$$\langle v,((\rho_{S,T})^*)^2 w\rangle=\langle \rho_{S,T}^2 v,w\rangle=\langle \rho_{S,T}v,w\rangle=\langle v,(\rho_{S,T})^* w\rangle. \tag{7.1.8}$$
利用引理 5.1，$(\rho_{S,T}^*)^2=(\rho_{S,T})^*$，因此其为投影算子，即 $(\rho_{S,T})^*=\rho_{T^\perp,S^\perp}$。

最后，在实（复）内积空间中，线性算子伴随对应的矩阵也有一种常用的表示。

命题 7.3

设 $\tau\in\mathcal{L}(V,W)$，其中 V 和 W 为有限维内积空间。若 $\mathcal{B}$ 和 $\mathcal{C}$ 分别为 V 和 W 的有序标准正交基，则

$$[\tau^*]_{\mathcal{C},\mathcal{B}}=([\tau]_{\mathcal{B},\mathcal{C}})^*. \tag{7.1.9}$$

简而言之，伴随 τ^* 的矩阵是 τ 的矩阵的伴随（共轭转置）。

证明 容易验证在 Euclidean 空间或幺正空间上的共轭转置满足伴随的定义要求，那么利用伴随的唯一性，可确定矩阵伴随恰好为共轭转置。不妨设 $\mathcal{B}=(b_1,\cdots,b_n),\mathcal{C}=(c_1,\cdots,c_m)$，那么

$$([\tau^*]_{\mathcal{C},\mathcal{B}})_{i,j}=([\tau^*(c_j)]_{\mathcal{B}})_i=\langle \tau^*(c_j),b_i\rangle=\overline{\langle \tau(b_i),c_j\rangle}=\overline{([\tau(b_i)]_{\mathcal{C}})_j}=\overline{([\tau]_{\mathcal{B},\mathcal{C}})_{j,i}}, \tag{7.1.10}$$

因此 $[\tau^*]_{\mathcal{C},\mathcal{B}}=([\tau]_{\mathcal{B},\mathcal{C}})^*$，结论得证。

7.2 正交投影

7.2.1 基本概念及其性质

在内积空间中，我们可以选择一些特殊的投影算子。

定义 7.1 (正交投影)

形如 $\rho_{S,S^\perp}$ 的投影称为正交投影（orthogonal projection）。等价地，若 $\ker(\rho)\perp\mathrm{im}(\rho)$，则称投影 ρ 正交。♣

注 小心区分彼此正交的投影与正交投影这两个不同的概念。

我们已经知道算子 ρ 是投影当且仅当其幂等（定理 3.10），而正交投影也具有类似的特征。

定理 7.2

设 V 为有限维内积空间，以下关于 $\mathcal{L}(V)$ 内的算子 ρ 的表述等价：

1. ρ 是正交投影；
2. ρ 幂等且自伴，即 $\rho=\rho^*$；
3. ρ 幂等且不增加长度，即 $\|\rho v\|\le\|v\|$ 对所有 $v\in V$ 成立。

♡

证明

1. 若表述 1 成立，则 ρ 幂等，并且存在子空间 S 使得 $\rho=\rho_{S,S^\perp}$。利用命题 5.5和7.2，$\rho^*=\rho_{(S^\perp)^\perp,S^\perp}=\rho_{S,S^\perp}=\rho$，从而 ρ 自伴，表述 2 成立；
2. 若表述 2 成立，那么对于任意 $v\in V$，$\langle v-\rho v,\rho v\rangle=\langle\rho^*(v-\rho v),v\rangle=\langle 0,v\rangle=0$，进而
$$\|v\|^2=\|(v-\rho v)+\rho v\|^2=\|v-\rho v\|^2+\|\rho v\|^2\ge\|\rho v\|^2\quad\Longrightarrow\quad\|\rho v\|\le\|v\|,\tag{7.2.1}$$
等号成立当且仅当 $v=\rho v$，即 $v\in\mathrm{im}(\rho)$，表述 3 成立；
3. 若表述 3 成立，则 ρ 为投影算子，并且投影定理 5.10表明，对于任意 $w\in\mathrm{im}(\rho)\subseteq V$，存在 $x\in\ker(\rho)$ 和 $y\in\ker(\rho)^\perp$ 使得 $w=x+y$，因此
$$x+y=w=\rho w=\rho(x+y)=\rho y\quad\Longrightarrow\quad\|x\|^2+\|y\|^2=\|w\|^2=\|\rho y\|^2\le\|y\|^2,\tag{7.2.2}$$
这要求 $\|x\|=0$，即 $x=0$，因此 $\mathrm{im}(\rho)\subseteq\ker(\rho)^\perp$；根据推论 4.1，$\mathrm{im}(\rho)\cong\ker(\rho)^\perp$。若 $\ker(\rho)^\perp\backslash\mathrm{im}(\rho)\neq\varnothing$，则 $\dim(\mathrm{im}(\rho))<\dim(\ker(\rho)^\perp)<\infty$。根据推论 3.2，这与两者同构的要求矛盾。因此 $\mathrm{im}(\rho)=\ker(\rho)^\perp$，根据命题 3.6，$\rho$ 为正交投影，表述 1 成立。

7.2.2 单位算子的正交分解

我们在第3.6.2小节讨论了 V 上单位算子的分解，其与 V 的直和分解形成一一对应。当投影算子还是正交投影时，对应直和为正交直和。

定义 7.2 (单位算子的正交分解)

单位算子的正交分解（orthogonal resolution of the identity）首先是单位算子的分解 $\rho_1+\cdots+\rho_k=\iota$，其次要求每个投影算子 ρ_i 是正交投影。♣

以下定理指出了 V 的正交直和分解与单位算子的正交分解之间的对应关系。

定理 7.3

设 V 为内积空间，V 上单位算子的正交分解与 V 的正交直和分解按如下方式对应：

1. 若 $\rho_1+\cdots+\rho_k=\iota$ 是单位算子的正交分解，则 $V=\mathrm{im}(\rho_1)\odot\cdots\odot\mathrm{im}(\rho_k)$；
2. 若 $V=S_1\odot\cdots\odot S_k$ 且 ρ_i 为到 S_i 上两两正交的投影，则 $\rho_1+\cdots+\rho_k=\iota$ 是单位算子的正交分解。

♡

证明

1. 根据定理 3.11，仅需证明当 $i \neq j$ 时 $\mathrm{im}(\rho_i) \perp \mathrm{im}(\rho_j)$。对于任意 $v_i \in \mathrm{im}(\rho_i), v_j \in \mathrm{im}(\rho_j)$，根据定理 7.2，
$$\langle v_i, v_j \rangle = \langle \rho_i v_i, \rho_j v_j \rangle = \langle \rho_j \rho_i v_i, v_j \rangle = \langle 0, v_j \rangle = 0, \tag{7.2.3}$$
因此 $\langle \mathrm{im}(\rho_i), \mathrm{im}(\rho_j) \rangle = \{0\}$，结论得证；
2. 根据定理 3.11，仅需证明 ρ_i 为正交投影，即 $S_i^\perp = \odot_{j \neq i} S_j$ 对所有 i 成立。由于 $S_j \perp S_i$ 对所有 $j \neq i$ 成立，显然有 $\odot_{j \neq i} S_j \subseteq S_i^\perp$。根据推论 4.1，$S_i^\perp \cong \odot_{j \neq i} S_j$，因此两者相同。

7.3 正规算子

7.3.1 一般性质

定义 7.3 (正规算子)

1. 若内积空间 V 上的线性算子 τ 与其伴随可交换，即 $\tau\tau^* = \tau^*\tau$，则称其正规（normal）；
2. 对于矩阵 $A \in \mathcal{M}_n(F)$，若 A 与其伴随 A^* 可交换，则称其正规。

♣

注 若 τ 正规且 $\mathcal{O}$ 是 V 的一组有序标准正交基，则
$$[\tau]_{\mathcal{O}}[\tau]_{\mathcal{O}}^* = [\tau]_{\mathcal{O}}[\tau^*]_{\mathcal{O}} = [\tau\tau^*]_{\mathcal{O}} \quad \text{and} \quad [\tau]_{\mathcal{O}}^*[\tau]_{\mathcal{O}} = [\tau^*]_{\mathcal{O}}[\tau]_{\mathcal{O}} = [\tau^*\tau]_{\mathcal{O}}. \tag{7.3.1}$$
因此 τ 是正规算子当且仅当 $[\tau]_{\mathcal{O}}$ 在 V 的某组标准正交基下正规。注意对于非正交归一的基，该结论未必成立。

命题 7.4

设 $\tau \in \mathcal{L}(V)$ 为正规算子。

1. (1) 若 τ 约化 $(S, S^\perp)$，则 $\tau|_S$ 是正规算子；
   (2) τ^* 是正规算子；
   (3) 若 τ 可逆，则 τ^{-1} 是正规算子；
   (4) 对任意多项式 $p(x) \in F[x]$，$p(\tau)$ 是正规算子；
2. 对任意 $v, w \in V$，$\langle \tau v, \tau w \rangle = \langle \tau^* v, \tau^* w \rangle$。特别地，$\|\tau v\| = \|\tau^* v\|$，因此 $\ker(\tau^*) = \ker(\tau)$；
3. 对任意整数 $k \geq 1$，$\ker(\tau^k) = \ker(\tau)$；
4. $\tau v = \lambda v \Leftrightarrow \tau^* v = \bar{\lambda} v$；
5. 若 λ 和 μ 为 τ 的不同本征值，则 $\mathcal{E}_\lambda \perp \mathcal{E}_\mu$，其中 $\mathcal{E}_\lambda$ 为对应本征值 λ 的本征子空间。

♠

证明

1. 根据命题 7.1，对于任意 $s \in S$，
$$(\tau|_S)^*(\tau|_S)s = (\tau^*)|_S \tau s = \tau^*\tau s = \tau\tau^* s = (\tau|_S)(\tau|_S)^* s, \tag{7.3.2}$$
因此 $\tau|_S$ 也是正规算子；根据定义，τ^* 显然是正规算子；根据命题 7.1，$(\tau^{-1})^*(\tau^{-1}) = (\tau^*)^{-1}(\tau^{-1}) = (\tau\tau^*)^{-1} = (\tau^*\tau)^{-1} = (\tau^{-1})(\tau^{-1})^*$，可见 τ^{-1} 是正规算子；根据命题 7.1，通过归纳法容易证明 $(\tau^m)^*\tau^n = (\tau^*)^m\tau^n = \tau^n(\tau^m)^*$，因此 $p(\tau)$ 也是正规算子；
2. 对于任意 $v, w \in V$，$\langle \tau v, \tau w \rangle = \langle \tau^*\tau v, w \rangle = \langle \tau\tau^* v, w \rangle = \langle \tau^* v, \tau^* w \rangle$；
3. 采用归纳法，该结论在 $k = 1$ 时显然成立。假设结论对于 $k-1$ 成立，显然有 $\ker(\tau) \subseteq \ker(\tau^k)$；对于任意 $v \in \ker(\tau^k)$，$0 = \tau^k v = \tau^{k-1}(\tau v)$，因此 $\tau v \in \ker(\tau^{k-1}) = \ker(\tau) = \ker(\tau^*)$（性质 2），即 $\tau^*\tau v = 0$，从而 $0 = \langle \tau^*\tau v, v \rangle$，即 $v \in \ker(\tau)$。因此 $\ker(\tau^k) \subseteq \ker(\tau)$，反向包含成立。那么该结论对所有整数 $k \geq 1$ 成立；
4. 根据性质 2，$v \in \ker(\tau - \lambda\iota) = \ker((\tau - \lambda\iota)^*) = \ker(\tau^* - \bar{\lambda}\iota)$，结论得证；
5. 设 $v \in \mathcal{E}_\lambda, w \in \mathcal{E}_\mu$，其中至少一个本征值非零。不妨设 $\lambda \neq 0$，那么结合性质 5，$\lambda \langle v, w \rangle = \langle \tau v, w \rangle = \langle v, \tau^* w \rangle = \langle v, \mu^* w \rangle = \mu \langle v, w \rangle$。由于 $\lambda \neq \mu$，要求 $\langle v, w \rangle = 0$，而 v, w 的任意性进一步要求 $\mathcal{E}_\lambda \perp \mathcal{E}_\mu$。

7.3.2 特殊正规算子

定义 7.4

设 V 为内积空间。

1. 若 $\tau \in \mathcal{L}(V)$ 满足 $\tau^* = \tau$，则称其为自伴的（self-adjoint）。另外，在复数和实数情形，τ 也分别称为厄米的（Hermitian）和对称的（symmetric）；
2. 若 $\tau \in \mathcal{L}(V)$ 可逆且满足 $\tau^* = \tau^{-1}$，则在复数和实数情形中分别称其为幺正的（unitary）和正交的（orthogonal）。

♣

当然也存在上述定义的矩阵版本，可通过简单将 τ 替换为矩阵 A 得到。此外，算子 τ 自伴当且仅当 τ 在任意一组有序标准正交基 $\mathcal{O}$ 下的坐标矩阵自伴。

7.3.2.1 自伴算子

与线性算子 τ 相关的二次型（*quadratic form*）为函数

$$Q_\tau : V \to F \text{ defined by } Q_\tau(v) = \langle \tau v, v \rangle . \tag{7.3.3}$$

根据命题 5.1，在复内积空间中，$\tau = 0$ 当且仅当 $Q_\tau(v) = 0$，但这对实内积空间一般不成立。不过对于实内积空间上的自伴算子，该结论依然成立。

命题 7.5

设 V 为有限维内积空间，$\sigma, \tau \in \mathcal{L}(V)$。

1. 若 τ 和 σ 自伴，则
   (1) $(\sigma + \tau)$ 自伴；
   (2) 若 τ 可逆，则 τ^{-1} 自伴；
   (3) 对任意实多项式 $p(x) \in \mathbb{R}[x]$，$p(\tau)$ 自伴；
2. 复算子 τ 是 Hermitian 算子当且仅当 $Q_\tau(v)$ 对所有 $v \in V$ 均为实数；
3. 若 τ 是复算子或实对称算子，则 $\tau = 0 \Leftrightarrow Q_\tau(v) = 0$。

♠

证明

1. 读者可结合命题 7.1 自行证明该结论；
2. $Q_\tau(v)$ 对所有 $v \in V$ 为实数当且仅当

$$\langle \tau v, v \rangle = Q_\tau(v) = \overline{Q_\tau(v)} = \langle v, \tau v \rangle = \langle \tau^* v, v \rangle \iff \langle (\tau - \tau^*)v, v \rangle = 0. \tag{7.3.4}$$

   若 $\tau = \tau^*$，则上式显然成立；反之若 $Q_\tau(v)$ 为实数，那么根据命题 5.1，在复内积空间中要求 $\tau^* = \tau$；
3. 复算子情形的证明与前文基本一致。在实数情形中，充分性是显然的；对于必要性，设 $v = x + y$，那么

$$0 = Q_\tau v = \langle \tau(x+y), x+y \rangle = Q_\tau(x) + Q_\tau(y) + \langle \tau x, y \rangle + \langle \tau y, x \rangle = \langle \tau x, y \rangle + \langle \tau^* x, y \rangle = 2 \langle \tau x, y \rangle , \tag{7.3.5}$$

   因此 $\langle \tau x, y \rangle = 0$ 对任意 $x, y \in V$ 成立。根据命题 5.1，$\tau = 0$。

7.3.2.2 幺正算子与等距映射

幺正算子具有一个良好性质：它是将一组标准正交集映射到另一组标准正交基上的正规算子。

命题 7.6

设 V 为有限维内积空间，$\sigma, \tau \in \mathcal{L}(V)$。

1. 若 τ 和 σ 是幺正（正交）算子，则
   (1) 对于满足 $|r| = 1$ 的 $r \in \mathbb{C}$，$r\tau$ 幺正；

(2) $\sigma\tau$ 是幺正算子；
(3) 若 τ 可逆，则 τ^{-1} 是幺正算子；
2. τ 幺正（正交）当且仅当其为等距同构；
3. τ 幺正当（正交）且仅当其将任意标准正交基映射到另一组标准正交基上；
4. 若 τ 幺正（正交），则 τ 的本征值具有绝对值 1。

♠

证明

1. 读者可结合命题 7.1 自行证明该结论；
2. 若 τ 幺正（正交），则对于任意 $u,v\in V$，$\langle\tau u,\tau v\rangle=\langle\tau^*\tau u,v\rangle=\langle u,v\rangle$，因此 τ 为等距同构；反之 $\langle\tau^*\tau u,v\rangle=\langle u,v\rangle$ 对任意 $u,v\in V$ 成立，根据引理 5.1，$\tau^*\tau u=u$ 对任意 $u\in V$ 成立，即 $\tau^*\tau=\iota$。由于 τ 为双射，其存在逆映射，故 $\tau^*=\tau^{-1}$，结论得证；
3. 利用 τ 是等距同构的性质容易证明；
4. 设 $\tau v=\lambda v$ 且 $v\neq 0$，则 $|\lambda|^2\langle v,v\rangle=\langle\lambda v,\lambda v\rangle=\langle\tau v,\tau v\rangle=\langle\tau^*\tau v,v\rangle=\langle v,v\rangle$，故 $|\lambda|=1$。

类似地，关于幺正（正交）矩阵也有如下结果。

命题 7.7

设 A 为 $F=\mathbb{C}$ 或 $\mathbb{R}$ 上的 $n\times n$ 矩阵。

1. 下列表述等价：
(1) A 幺正（正交）；
(2) A 的列组成 F^n 中一组标准正交基；
(3) A 的行组成 F^n 中一组标准正交基；
2. 若 A 幺正，则 $|\det(A)|=1$。若 A 正交，则 $\det(A)=\pm1$。

♠

证明

1. 矩阵 A 幺正当且仅当 $AA^*=I$，这等价于 A 的行标准正交；类似地，$A^*A=I$ 等价于 A 的列标准正交；
2. 利用 $1=\det(A)\det(A^*)=\det(A)\overline{\det(A)}$ 可得到结果。

幺正（正交）矩阵在基变换中起到了重要的作用。若 $\mathcal{B}=(u_1,\cdots,u_n)$ 为有序标准正交基，且

$$\begin{cases}v=a_1u_1+\cdots+a_nu_n,\\ w=b_1u_1+\cdots+b_nu_n\end{cases}\quad\Longrightarrow\quad\langle v,w\rangle=a_1\overline{b_1}+\cdots+a_n\overline{b_n}=[v]_{\mathcal{B}}\cdot[w]\mathcal{B},\tag{7.3.6}$$

其中右侧为 F^n 的标准内积，因此 $v\perp w$ 当且仅当 $[v]_{\mathcal{B}}\perp[w]_{\mathcal{B}}$，从而可给出类似推论 3.4 的结果。

定理 7.4

若给定以下任意两者：

1. 幺正（正交）的 $n\times n$ 矩阵 A；
2. F^n 的有序标准正交基 $\mathcal{B}$；
3. F^m 的有序标准正交基 $\mathcal{C}$；

则第三者可由方程 $A=M_{\mathcal{B},\mathcal{C}}$ 唯一确定。

♡

最后回顾一下算子（矩阵）相似的内容。当有序基正交归一时，会有一些更良好的性质。

定义 7.5 (幺正等价与正交等价)

1. 若对于两个复矩阵 A 和 B，若存在幺正矩阵 U 使得 $B=UAU^{-1}=UAU^*$，则称 A 和 B 幺正相似（unitarily similar），也称为幺正等价（unitarily equivalent）。与幺正相似相关的等价类称为幺正相似类（unitary similarity class）；
2. 若对于两个实矩阵 A 和 B，若存在正交矩阵 O 使得 $B=OAO^{-1}=OAO^T$，则称 A 和 B 正交相似

（orthogonally similar），也称为正交等价（orthogonally equivalent）。与正交相似相关的等价类称为正交相似类（orthogonal similarity class）。

类似推论 3.6的结果如下，其证明留给读者（可结合定理 7.4）。

定理 7.5

设 V 为 n 维内积空间，那么两个 $n\times n$ 矩阵 A 和 B 幺正（正交）相似当且仅当其关于（可能不同的）有序标准正交基表示同一个线性算子 $\tau\in\mathcal{L}(V)$。在这种情况下，A 和 B 恰好表示 $\mathcal{L}(V)$ 中（关于有序标准正交基）相同的线性算子集。

7.3.3 正规算子的结构

鉴于我们未讨论有理规范形和 Jordan 标准型的动机与形式，在此我们直接将结论列举如下。

7.3.3.1 算子版本

定理 7.6 (正规算子的结构定理)

1. 复数情形：设 V 为有限维复内积空间。
   (1) 对于 $\tau\in\mathcal{L}(V)$，以下表述等价：
      I. τ 是正规的；
      II. τ 是幺正可对角化的；
      III. τ 有正交谱分解：
      $$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k. \tag{7.3.7}$$
   (2) 在正规算子中，Hermitian 算子恰好为那些所有复本征值均为实数的算子；
   (3) 在正规算子中，幺正算子恰好为那些所有复本征值范数均为 1 的算子；
2. 实数情形：设 V 为有限维实内积空间。
   (1) $\tau\in\mathcal{L}(V)$ 正规当且仅当
   $$V=\mathcal{E}_{\lambda_1}\odot\cdots\odot\mathcal{E}_{\lambda_k}\odot W_1\odot\cdots\odot W_m, \tag{7.3.8}$$
   其中 $\{\lambda_1,\cdots,\lambda_k\}$ 为 τ 的谱，每个 W_j 是二维不可分解 τ-不变子空间，其内包含一组有序基 $\mathcal{B}_j$，满足
   $$[\tau]_{\mathcal{B}_j}=\begin{bmatrix}a_j & -b_j\\ b_j & a_j\end{bmatrix} \tag{7.3.9}$$
   (2) 在实正规算子中，对称算子为那些上述分解中不出现子空间 W_i 的算子。因此，对于 $\tau\in\mathcal{L}(V)$，以下表述等价：
      I. τ 是对称的；
      II. τ 是正交可对角化的；
      III. τ 有正交谱分解：
      $$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k. \tag{7.3.10}$$
   (3) 在是正规算子中，正交算子恰好为那些特征值为 ±1 且分解出的 $[\tau]_{\mathcal{B}_i}$ 的行和列范数均为 1，即存在 $\theta\in\mathbb{R}$ 使得
   $$[\tau]_{\mathcal{B}_i}=\begin{bmatrix}\sin\theta & -\cos\theta\\ \cos\theta & \sin\theta\end{bmatrix} \tag{7.3.11}$$

7.3.3.2 矩阵版本

定理 7.7 (正规矩阵的结构定理)

1. 复数情形：
   (1) 复矩阵 A 正规当且仅当其幺正可对角化，即当且仅当存在幺正矩阵 U 使得
   $$UAU^* = \mathrm{diag}(\lambda_1, \cdots, \lambda_k). \tag{7.3.12}$$
   (2) 复矩阵 A Hermitian 当且仅当表述 (1) 成立，其中各本征值 λ_i 为实数；
   (3) 复矩阵 A 幺正当且仅当表述 (1) 成立，其中所有本征值 λ_i 的范数为 1；
2. 实数情形：
   (1) 实矩阵 A 正规当且仅当存在正交矩阵 O 使得
   $$OAO^T = \mathrm{diag}\left(\lambda_1, \cdots, \lambda_k, \begin{bmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{bmatrix}, \cdots \begin{bmatrix} a_m & -b_m \\ b_m & a_m \end{bmatrix}\right) \tag{7.3.13}$$
   (2) 实矩阵 A 对称当且仅当其正交可对角化，即当且仅当存在正交矩阵 O 使得
   $$OAO^T = \mathrm{diag}(\lambda_1, \cdots, \lambda_k). \tag{7.3.14}$$
   (3) 实矩阵 A 正交当且仅当存在正交矩阵 O 和 $\theta_1, \cdots, \theta_m \in \mathbb{R}$ 使得
   $$OAO^T = \mathrm{diag}\left(\lambda_1, \cdots, \lambda_k, \begin{bmatrix} \sin\theta_1 & -\cos\theta_1 \\ \cos\theta_1 & \sin\theta_1 \end{bmatrix}, \cdots, \begin{bmatrix} \sin\theta_m & -\cos\theta_m \\ \cos\theta_m & \sin\theta_m \end{bmatrix}\right) \tag{7.3.15}$$

♡

7.4 泛函演算

7.4.1 概念与动机

设 τ 为有限维内积空间 V 上的正规算子，τ 具有谱分解（定理 6.2）

$$\tau = \lambda_1\rho_1 + \cdots + \lambda_k\rho_k. \tag{7.4.1}$$

由于每个 ρ_i 都是幂等的，$\rho_i^m = \rho_i$ 对任意整数 $m \geq 1$ 成立。利用归纳法，可证明投影算子的两两正交性意味着 $\tau^n = \lambda_1^n\rho_1 + \cdots + \lambda_k^n\rho_k$ 对所有正整数 n 成立。因此对于 F 上任意多项式 $p(x)$，均有

$$p(\tau) = p(\lambda_1)\rho_1 + \cdots + p(\lambda_k)\rho_k. \tag{7.4.2}$$

对于任意 $(k-1)$ 次多项式 $p(x)$，若其在 $p(\lambda_1), \cdots, p(\lambda_k)$ 这 k 个点的取值确定，则该多项式唯一确定。

例 7.1

1. 对于给定的指标 $1 \leq j \leq k$，设多项式 $p_j(x)$ 在 λ_i 处取值为 $\delta_{i,j}$，那么
   $$p_j(\tau) = \sum_{i=1}^{k} p(\lambda_i)\rho_i = \sum_{i=1}^{k} \delta_{i,j}\rho_i = \rho_j. \tag{7.4.3}$$
   因此每个投影算子 ρ_i 都可以唯一表示为关于 τ 的多项式；
2. 若 τ 可逆，设多项式 $p(x)$ 在 λ_i 处取值为 λ_i^{-1}，那么
   $$p(\tau) = \lambda_1^{-1}\rho_1 + \cdots + \lambda_k^{-1}\rho_k \quad \Longrightarrow \quad \tau p(\tau) = \rho_1 + \cdots + \rho_k = \iota = p(\tau)\tau. \tag{7.4.4}$$
   因此逆映射 $\tau^{-1} = p(\tau)$ 也可以唯一表示为关于 τ 的多项式；
3. 设多项式 $p(x)$ 在 λ_i 处取值为 $\overline{\lambda_i}$，那么
   $$p(\tau) = \overline{\lambda_1}\rho_1 + \cdots + \overline{\lambda_k}\rho_k = \tau^*. \tag{7.4.5}$$
   因此 τ 的伴随 τ^* 同样可以唯一表示为关于 τ 的多项式。

我们可以将该想法进一步推广。对于任意函数 $f:\{\lambda_1,\cdots,\lambda_k\}\to F$，可如下定义线性算子：

$$f(\tau)=f(\lambda_1)\rho_1+\cdots+f(\lambda_k)\rho_k, \tag{7.4.6}$$

其中 $f(\tau)$ 只是一种形式记号，并不意味着 f 的定义域为 $\mathcal{L}(V)$。由此可定义例如 $\sqrt{\tau},\tau^{-1},\mathrm{e}^{\tau}$ 等函数。不过需要注意的是，在此定义中我们仅关心 f 在有限个点 $\lambda_1,\cdots,\lambda_k$ 上的取值而不关心其具体形式。因此，采用多项式之外的函数讨论 $f(\tau)$ 的表示并不带来任何额外便利：由于 τ 的谱分解是有限和，借助插值的思想，通过 $\{(\lambda_i,f(\lambda_i))\}_{i=1}^{k}$ 可唯一确定一个 $(k-1)$ 次多项式 $p(x)$，从而 $f(\tau)=p(\tau)$；换言之，对于任意形式上非多项式的函数 f，其在 τ 上的作用与多项式函数 p 在 τ 上的作用结果无异。研究算子 τ 的函数称为 τ 的*泛函演算*（*functional calculus*）。

7.4.2 交换性

泛函演算可应用于算子交换性的研究，以下是两个简单例子。

定理 7.8

设 V 为有限维复内积空间。对于 $\tau,\sigma\in\mathcal{L}(V)$，设其谱分解分别为

$$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k, \tag{7.4.7}$$

$$\sigma=\mu_1\nu_1+\cdots+\mu_m\nu_m, \tag{7.4.8}$$

并以 $\tau\leftrightarrow\sigma$ 的记号表示 τ 与 σ 可交换。

1. 对于任意 $\mu\in\mathcal{L}(V)$，$\mu\leftrightarrow\tau\ \Leftrightarrow\ \mu\leftrightarrow\rho_i$ 对所有 i 成立；
2. $\tau\leftrightarrow\sigma\ \Leftrightarrow\ \rho_i\leftrightarrow\nu_j$ 对所有 i,j 成立；
3. 若 $f:\{\lambda_1,\cdots,\lambda_k\}\to F$ 和 $g:\{\mu_1,\cdots,\mu_m\}\to F$ 为单射，则 $f(\tau)\leftrightarrow g(\sigma)\ \Leftrightarrow\ \tau\leftrightarrow\sigma$。

♡

证明

1. 必要性是显然的；对于充分性，根据例7.1，存在多项式 $p_i(x)$ 使得 $\rho_i=p_i(\tau)$，因此 $\mu\leftrightarrow\rho_i$；
2. 必要性是显然的；对于充分性，根据例7.1，存在多项式 $p_i(x)$ 和 $q_j(x)$ 使得 $\rho_i=p_i(\tau)$ 且 $\nu_j=q_j(\sigma)$，因此 $\tau\leftrightarrow\sigma$；
3. 对于必要性，结论 2 要求 $\tau\leftrightarrow\sigma$ 时 $\rho_i\leftrightarrow\nu_j$ 对所有 i,j 成立，因此 $f(\tau)\leftrightarrow g(\sigma)$；对于充分性，记 $\Lambda=\{\lambda_1,\cdots,\lambda_k\}$。由于 $f:\Lambda\to F$ 是单射，故 $f:\Lambda\to\mathrm{im}(f)$ 是双射，即存在逆映射 $f^{-1}:\mathrm{im}(f)\to\Lambda$ 使得 $f^{-1}(f\tau)=\tau$。当然 $f^{-1}\in F^{\Lambda}$，从而也是泛函演算，即 $\tau=f^{-1}(f\tau)$ 是关于 $f(\tau)$ 的多项式。类似地，σ 是关于 $g(\sigma)$ 的多项式，从而 $\tau\leftrightarrow\sigma$。

定理 7.9

设 τ 和 σ 为有限维复内积空间 V 上的正规算子，则 τ 和 σ 可交换当且仅当其具有如下形式：

$$\tau=p(r(\tau,\sigma)),\quad \sigma=q(r(\tau,\sigma)), \tag{7.4.9}$$

其中 $p(x),q(x)$ 和 $r(x,y)$ 均为多项式。

♡

证明 必要性是显然的；对于充分性，不妨设 τ 和 σ 的谱分解分别为 $\lambda_1\rho_1+\cdots+\lambda_k\rho_k$ 和 $\mu_1\nu_1+\cdots+\mu_m\nu_m$。根据7.8，$\tau\leftrightarrow\sigma$ 意味着 $\rho_i\leftrightarrow\nu_j$ 对所有 i,j 成立，因此

$$\tau^r\sigma^s=(\lambda_1^r\rho_1+\cdots+\lambda_k^r\rho_k)(\mu_1^s\nu_1+\cdots+\mu_m^s\nu_m)=\sum_{i,j}\lambda_i^r\mu_j^s\rho_i\nu_j\quad\Longrightarrow\quad r(\tau,\sigma)=\sum_{i,j}r(\lambda_i,\mu_j)\rho_i\nu_j \tag{7.4.10}$$

对任意双变量多项式 $r(x,y)$ 成立。选择合适的多项式 $r(x,y)$，使得 $\alpha_{i,j}=r(\lambda_i,\mu_j)$ 关于不同的有序对 (i,j) 不同。那么存在唯一的 $(km-1)$ 次多项式 $p(x)$ 和 $q(y)$，使得 $p(\alpha_{i,j})=\lambda_i$ 且 $q(\alpha_{i,j})=\mu_j$。由于 $\rho_i\leftrightarrow\nu_j$，可知

$$p(r(\tau,\sigma))=\sum_{i,j}p(\alpha_{i,j})(\rho_i\nu_j)=\sum_i\lambda_i\rho_i\sum_j\nu_j=\sum_i\lambda_i(\rho_i\iota)=\tau, \tag{7.4.11}$$

类似地 $q(r(\tau,\sigma))=\sigma$，结论得证。

7.5 正算子与极分解

7.5.1 正算子与正定算子

定义 7.6 (正算子与正定算子)

对于自伴算子 $\tau \in \mathcal{L}(V)$。

1. 若二次型 $Q_\tau(v) = \langle \tau v, v\rangle \geq 0$ 对所有 $v \in V$ 成立，则称 τ 是正的（positive）；
2. 若二次型 $Q_\tau(v) = \langle \tau v, v\rangle > 0$ 对所有非零的 $v \in V$ 成立，则称 τ 是正定的（positive definite）。

♣

根据定义，我们立刻得到正算子与正定算子的等价描述。

命题 7.8

在有限维内积空间上的自伴算子 τ 是

1. 正的当且仅当其所有本征值非负；
2. 正定的当且仅当其所有本征值为正。

♠

证明

1. 对于充分性，取非零向量 v 为 τ 的本征值为 λ 的本征向量，则有 $\lambda\langle v, v\rangle = \langle \tau v, v\rangle \geq 0$，即 $\lambda \geq 0$；对于必要性，根据定理 6.2和7.6，设 τ 的谱分解为 $\tau = \lambda_1\rho_1 + \cdots + \lambda_k\rho_k$，那么

$$\begin{aligned} Q_\tau(v) &= \langle (\lambda_1\rho_1 + \cdots + \lambda_k\rho_k)v, (\rho_1 + \cdots + \rho_k)v\rangle = \sum_{i,j} \lambda_i \langle \rho_i v, \rho_j v\rangle \\ &= \sum_{i,j} \lambda_i \langle \rho_j\rho_i v, v\rangle = \sum_i \lambda_i \langle \rho_i^2 v, v\rangle = \sum_i \lambda_i \langle \rho_i v, \rho_i v\rangle \geq 0, \end{aligned} \tag{7.5.1}$$

   因此 τ 为正算子，结论得证；
2. 对于充分性，仿上可知 $\lambda\langle v, v\rangle > 0$ 对 τ 的所有非零本征向量 v 成立，因此 $\lambda > 0$；对于必要性，仿上可知 $Q_\tau(v) = \sum_i \lambda_i \langle \rho_i v, \rho_i v\rangle$ 中至少存在一个 $\langle \rho_i v, \rho_i v\rangle > 0$，对应的 $\lambda_i > 0$，因此 $Q_\tau(v) > 0$。

若 τ 为具有谱分解 $\tau = \lambda_1\rho_1 + \cdots + \lambda_k\rho_k$ 的正算子，那么我们可取 τ 的正平方根（*positive square root*）：

$$\sqrt{\tau} = \sqrt{\lambda_1}\rho_1 + \cdots + \sqrt{\lambda_k}\rho_k, \tag{7.5.2}$$

其中 $\sqrt{\lambda_i}$ 是 λ_i 的非负平方根，容易验证 $(\sqrt{\tau})^2 = \tau$；并且不难看出 $\sqrt{\tau}$ 是唯一平方为 τ 的正算子：若存在正算子 σ 使得 $\tau = \sigma^2$，那么对于 τ 的本征值 λ_i 及其对应的本征向量 v_i，$\tau(\sigma v_i) = \sigma^3 v_i = \sigma(\tau v_i) = \lambda_i(\sigma v_i)$，可见 $\mathcal{E}_{\lambda_i}$ 也是 σ-不变子空间，这意味着

$$\mathcal{E}_{\lambda_i} = \tau\mathcal{E}_{\lambda_i} = \sigma(\sigma\mathcal{E}_{\lambda_i}) \subseteq \sigma\mathcal{E}_{\lambda_i} \subseteq \mathcal{E}_{\lambda_i}. \tag{7.5.3}$$

因此 $\mathcal{E}_{\lambda_i} = \sigma\mathcal{E}_{\lambda_i}$。根据定理 6.2，$\sigma$ 与 τ 的正交谱分解具有相同的投影算子，可设 $\sigma = \mu_1\rho_1 + \cdots + \mu_k\rho_k$；根据命题 7.8，$\mu_i \geq 0$。利用 $\sigma^2 = \tau$ 可确定 $\mu_i = \sqrt{\lambda_i}$。反之，若 τ 有正平方根，即存在正算子 σ 使得 $\tau = \sigma^2$，则 τ 显然是正的。我们将上述讨论整理为如下结果。

定理 7.10

设 $\tau \in \mathcal{L}(V)$，则以下表述等价：

1. τ 为正算子；
2. τ 有唯一的正平方根；
3. 存在算子 σ 使得 $\tau = \sigma^*\sigma$。

♡

证明

1. 若表述 1 成立，则根据上述讨论，表述 2 也成立；
2. 若表述 2 成立，容易证明 τ 的正平方根 $\sqrt{\tau}$ 满足 $(\sqrt{\tau})^* = \sqrt{\tau}$，进而 $\tau = (\sqrt{\tau})^2 = (\sqrt{\tau})^*(\sqrt{\tau})$，表述 3 成立；

3. 若表述 3 成立，τ 的二次型 $Q_\tau(v)=\langle\sigma^*\sigma v,v\rangle=\langle\sigma v,\sigma v\rangle\geq 0$ 对所有 $v\in V$ 成立，因此 τ 为正算子。

下述定理恰好是正平方根的应用。

命题 7.9

若 τ 和 σ 为正算子且 $\tau\sigma=\sigma\tau$，则 $\tau\sigma$ 是正的。

♠

证明 设 τ 和 σ 的谱分解分别为 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$ 和 $\sigma=\mu_1\nu_1+\cdots+\mu_m\nu_m$。根据定理 7.8，此时 ρ_i 与 ν_j 对任意 i,j 可交换，那么其正平方根 $\sqrt{\tau}$ 与 $\sqrt{\sigma}$ 同样可交换，进而

$$(\sqrt{\tau}\sqrt{\sigma})^2=(\sqrt{\tau}\sqrt{\sigma})(\sqrt{\tau}\sqrt{\sigma})=\sqrt{\tau}(\sqrt{\tau}\sqrt{\sigma})\sqrt{\sigma}=\tau\sigma. \tag{7.5.4}$$

根据定理 7.10，$\tau\sigma$ 也是正算子。

7.5.2 算子的极分解

类似复数的极形式（*polar form*）$z=re^{i\theta}$，有限维复内积空间上的任意非零线性算子 τ 也有类似的分解。

引理 7.1

设 τ 为有限维复内积空间 V 上的非零线性算子。

1. 存在正算子 ρ 和幺正算子 ν 使得 $\tau=\nu\rho$，其中 ρ 唯一。此外，若 τ 可逆，则 ν 也唯一；
2. 存在正算子 σ 和幺正算子 μ 使得 $\tau=\sigma\mu$，其中 σ 唯一。此外，若 τ 可逆，则 μ 也唯一。

♡

证明 这里我们仅证明第一条结论，第二条结论类似可证。假设 τ 可分解为 $\tau=\nu\rho$ 的形式，那么

$$\tau^*\tau=(\rho^*\nu^*)(\nu\rho)=\rho(\nu^{-1}\nu)\rho=\rho^2 \tag{7.5.5}$$

那么根据定理 7.10，正算子 $\tau^*\tau$ 具有唯一正平方根 ρ，并且对于任意 $v\in V$，

$$\|\rho v\|^2=\langle\rho v,\rho v\rangle=\langle\rho^2 v,v\rangle=\langle\tau^*\tau v,v\rangle=\|\tau v\|^2. \tag{7.5.6}$$

根据命题 7.6，幺正算子是等距同构，因此可定义等距映射 $\nu:\mathrm{im}(\rho)\to\mathrm{im}(\tau)$ 为 $\nu(\rho v)=\tau v$。首先需要检验 ν 是良定义的：当 $\rho x=\rho y$ 对某些 $x,y\in V$ 成立时，$\|\tau(x-y)\|=\|\rho(x-y)\|=0$，因此 $\tau x=\tau y$。反之可验证 τ 是单射，因此若 $\mathcal{B}=\{b_1,\cdots,b_k\}$ 是 $\mathrm{im}(\rho)$ 的一组标准正交基，那么 $\nu\mathcal{B}=\{\nu b_1,\cdots,\nu b_k\}$ 也是一组标准正交集，并且 $\langle\nu\mathcal{B}\rangle=\mathrm{im}(\tau)$，即 $\nu\mathcal{B}$ 是 $\mathrm{im}(\tau)$ 的一组标准正交基。根据定理 5.7，对于有限维内积空间，总可以将 $\mathcal{B}$ 扩张为 V 的一组标准正交基，从而 $\nu\mathcal{B}$ 也是 V 的一组标准正交基，此时 $\nu:V\to\mathrm{im}(\nu)$ 是等距同构。

当 τ 可逆时，$\nu^*\tau=\nu^{-1}(\nu\rho)=\rho$ 表明 ρ 也是可逆的。利用 ρ 的唯一性，此时 $\nu=\tau\rho^{-1}$ 唯一。

我们不加证明地指出任何幺正算子都具有 $\mu=e^{i\sigma}$ 的形式，其中 σ 自伴，从而有如下推论。

定理 7.11 (极分解)

设 τ 为有限维复内积空间上的非零线性算子，则存在正算子 ρ 和 σ 使得 τ 具有极分解（polar decomposition）$\tau=\rho e^{i\sigma}$，其中 ρ 唯一确定。此外，若 τ 可逆，则 σ 也唯一。

♡

该定理还有一个直接的推论。

推论 7.1

设 $\tau=\rho e^{i\sigma}$ 为非零线性算子的极分解，则 τ 正规当且仅当 $\rho\sigma=\sigma\rho$。

♡

证明 已知 $\tau\tau^*=(\rho e^{i\sigma})(e^{-i\sigma}\rho)=\rho^2$ 且 $\tau^*\tau=e^{-i\sigma}\rho^2e^{i\sigma}$，因此 τ 正规当且仅当 $\tau\tau^*=\tau^*\tau$，即 $e^{i\sigma}\rho^2=\rho^2e^{i\sigma}$。设 τ 和 σ 的谱分解分解为 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$ 和 $\sigma=\mu_1\nu_1+\cdots+\mu_m\nu_m$，从而 $e^{i\sigma}=e^{i\mu_1}\nu_1+\cdots+e^{i\mu_m}\nu_m$。根据定理 7.8，$\rho$ 与 $e^{i\sigma}$ 可交换当且仅当 ρ_i 与 ν_j 对所有 i,j 可交换，当且仅当 ρ 与 σ 可交换。

第 8 章 Hilbert 空间

我们曾在第5章详细讨论了实内积空间与复内积空间的基本结构，不过彼时讨论仅限于有限维情形。在赋予内积空间特定的拓扑结构后，许多性质对于无穷维情形依然适用，因此可以放宽对维数的限制。在本章中，F 同样表示 $\mathbb{R}$ 或 $\mathbb{C}$。

目录

8.1 Hilbert 空间 66
8.2 最佳逼近与 Fourier 展开 67
8.3 Hilbert 基与 Hilbert 维数 70
8.4 Riesz 表示定理 71

8.1 Hilbert 空间

在引入 Hilbert 空间的概念以前，我们需要以下两个与连续性有关的基本事实。

命题 8.1

设 V 为内积空间，则对于内积诱导的范数。

1. $(x_n) \to x, (y_n) \to y \Rightarrow \langle x_n, y_n\rangle \to \langle x, y\rangle$；
2. $(x_n) \to x \Rightarrow \|x_n\| \to x$。

♠

证明 我们仅需证明第一条结论，第二条结论是第一条的直接推论。由于 $(x_n) \to x$，存在正数 M 使得 $\|x_n\| \leq M$ 对所有自然数 n 成立。因此对于任意正数 ϵ，存在 $n_1, n_2 \in \mathbb{N}$，使得 $\|x_n - x\| < \epsilon/2\|y\|$ 对所有 $n \geq n_1$ 成立、$\|y_n - y\| < \epsilon/2M$ 对所有 $n \geq n_2$ 成立。取 $n_0 = \max\{n_1, n_2\}$，利用三角不等式和 Cauchy-Schwarz 不等式，

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &= |\langle x_n, y_n\rangle - \langle x_n, y\rangle + \langle x_n, y\rangle - \langle x, y\rangle| \leq |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| < M \cdot \frac{\epsilon}{2M} + \frac{\epsilon}{2\|y\|} \cdot \|y\| = \epsilon \end{aligned} \tag{8.1.1}$$

对所有 $n \geq n_0$ 成立，因此 $\langle x_n, y_n\rangle \to \langle x, y\rangle$。

由于内积空间是度量空间，可根据完备度量空间的概念1.25定义 Hilbert 空间。

定义 8.1 (Hilbert 空间)

在内积诱导的度量下完备的内积空间称为 Hilbert 空间（Hilbert space）。

♣

例 8.1 Hilbert 空间的一个重要例子是 ℓ^2，其内积定义为 $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{n=1}^{\infty} x_n\overline{y_n}$，因此内积诱导的度量为

$$d(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|_2 = \left(\sum_{n=1}^{\infty} |x_n - y_n|^2\right)^{1/2}. \tag{8.1.2}$$

根据定理 1.14，实（复）内积空间内的 Cauchy 列收敛，因此是完备的，即 ℓ^2 是 Hilbert 空间。

内积空间的子空间也有与命题 1.12类似的结论。

命题 8.2

1. 内积空间的任意完备子空间是闭集；
2. Hilbert 空间的子空间是 Hilbert 空间当且仅当其为闭集；
3. 任意内积空间的有限维子空间是完备闭集。

♠

证明 前两条结论是命题 1.12的直接结果。对于第三条结论，设 (x_n) 是有限维子空间 S 中收敛到 x 的序列且 $x\notin S$。设 $\mathcal{B}=\{b_1,\cdots,b_m\}$ 是 S 的一组标准正交 Hamel 基，那么 x 的 Fourier 展开 $\hat{x}=\sum_{i=1}^m\langle x,b_i\rangle b_i\in S$。根据定理 5.9，$(x-\hat{x})\perp S$ 但 $x\neq\hat{x}$。因此，如果记 $y_n=x_n-\hat{x}\in S, y=x-\hat{x}\perp y_n$，我们有 $\|y_n-y\|^2=\|y_n\|^2+\|y\|^2\geq\|y\|^2>0$，这与 $(x_n)\to x$ 矛盾。因此 $x\in S$，即 S 是闭集。根据定理 1.14，S 内任意 Cauchy 列 (x_n) 收敛到某个内积空间内的向量 x。由于 S 是闭集，$x\in S$，故 S 完备。

8.2 最佳逼近与 Fourier 展开

8.2.1 最佳逼近

设 V 是任意维数的内积空间，S 是 V 的一个子集。能否对任意 $x\in V$，找到 S 中在内积诱导的度量下最接近 x 的向量是一个我们相当关心的问题。这就是 V 的逼近问题（*approximation problem*）。

定理 8.1

设 V 为内积空间，S 为 V 的完备凸子集[a]，则对于任意 $x\in V$，存在唯一的 $\hat{x}\in S$ 使得

$$\|x-\hat{x}\|=\inf_{s\in S}\|x-s\|,\tag{8.2.1}$$

向量 $\hat{x}$ 称为 S 内 x 的最佳逼近（best approximation）。

[a]凸集的概念可直接参考定义 10.1

证明 当 $x\in S$ 时，显然 $\|x-\hat{x}\|=0$ 当且仅当 $\hat{x}=x$ 唯一，该情形是平凡的；当 $x\notin S$ 时，记 $\delta=\inf_{s\in S}\|x-s\|\in\mathbb{R}$，那么对于任意自然数 n，存在存在 $s_n\in S$ 使得 $\delta\leq\delta_n=\|x-s_n\|<\delta+1/n$，可见 $\delta_n\to\delta$。因此存在正数 M 使得 $|\delta_n|<M$ 对所有自然数 n 成立。对于任意正数 ϵ，存在自然数 n_0 使得 $|\delta_n-\delta|<\epsilon/4(\delta+M)$ 对所有 $n\geq n_0$ 成立。记 $y_n=x-s_n$，那么根据平行四边形法则（命题 5.2），

$$\|y_n-y_m\|^2=2(\|y_n\|^2+\|y_m\|^2)-\|y_n+y_m\|^2.\tag{8.2.2}$$

对任意自然数 n,m 成立。由于 S 是凸集，$(s_n+s_m)/2\in S$，从而 $\|y_n+y_m\|^2=4\|x-(s_n+s_m)/2\|^2\geq4\delta^2$，即

$$\|s_m-s_n\|^2=\|y_n-y_m\|^2\leq2(\delta_n^2-\delta^2)+2(\delta_m^2-s^2)<\frac{\epsilon}{4(\delta+M)}\cdot4\left(\delta+\frac{\delta_n+\delta_m}{2}\right)\leq\epsilon\tag{8.2.3}$$

对所有 $n,m\geq n_0$ 成立，可见 (s_n) 是 S 内的一条 Cauchy 列，其收敛到某个 $\hat{x}\in S$。利用范数的连续性（命题 8.1），

$$\|x-\hat{x}\|=\left\|x-\lim_{n\to\infty}s_n\right\|=\lim_{n\to\infty}\|x-s_n\|=\lim_{n\to\infty}\delta_n=\delta=\inf_{s\in S}\|x-s\|.\tag{8.2.4}$$

最后证明 $\hat{x}$ 的唯一性。设存在 $\hat{x}'\in S$ 使得 $\|x-\hat{x}'\|=\delta=\|x-\hat{x}\|$，则根据平行四边形法则（命题 5.2），

$$\|\hat{x}-\hat{x}'\|^2=2(\|x-\hat{x}\|^2+\|x-\hat{x}'\|^2)-4\left\|x-\frac{\hat{x}+\hat{x}'}{2}\right\|^2\leq2(\delta^2+\delta^2)-4\delta^2=0,\tag{8.2.5}$$

上式成立当且仅当 $\|\hat{x}-\hat{x}'\|=0$，即 $\hat{x}'=\hat{x}$。

由于内积空间 V 的任意子空间是代数封闭的，因此是凸的，我们可将上述定理应用于完备子空间。

定理 8.2

设 V 为内积空间，S 为 V 的完备子空间，则对于任意 $x\in V$，$\hat{x}$ 是 S 内 x 的最佳逼近当且仅当 $\hat{x}$ 是唯一使得 $(x-s)\perp S$ 的向量 $s\in S$。

证明 若存在 $\hat{x}\in S$ 使得 $(x-\hat{x})\perp S$，那么 $\langle x-\hat{x},\hat{x}-s\rangle=0$ 对任意 $s\in S$ 成立，进而

$$\|x-s\|^2=\|(x-\hat{x})+(\hat{x}-s)\|^2=\|x-\hat{x}\|^2+\|\hat{x}-s\|^2\geq\|x-\hat{x}\|^2,\tag{8.2.6}$$

即 $\|x-s\|\geq\|x-\hat{x}\|$ 对任意 $s\in S$ 成立，故 $\hat{x}$ 是 x 的在最佳逼近。

根据定理 8.1，存在唯一的 $\hat{x}\in S$ 使得 $\|x-\hat{x}\|=\inf_{s\in S}\|x-s\|=\delta$。对于任意 $s\in S$ 和 $r\in F$，首先有

$$\|x-rs\|^2=\|x\|^2+\left(|r|^2\|s\|^2-r\langle s,x\rangle-\bar{r}\langle x,s\rangle\right)=\|x\|^2+\|s\|^2\left|r-\frac{\langle x,s\rangle}{\|s\|^2}\right|^2-\frac{|\langle x,s\rangle|^2}{\|s\|^2}. \tag{8.2.7}$$

因此当 r 取 $r_0=\langle x,s\rangle/\|s\|^2$ 时上式第二项为零，此时

$$\|x-\hat{x}-r_0s\|^2=\|x-\hat{x}\|^2-\frac{|\langle x-\hat{x},s\rangle|^2}{\|s\|^2}\leq\|x-\hat{x}-r_0s\|^2-\frac{|\langle s-\hat{x},s\rangle|^2}{\|s\|^2}. \tag{8.2.8}$$

这要求 $|\langle x-\hat{x},s\rangle|^2\leq 0$，即 $\langle x-\hat{x},s\rangle=0$ 对所有 $s\in S$ 成立，故 $(x-\hat{x})\perp S$。根据前述讨论，若存在 $\hat{x}'\in S$ 使得 $(x-\hat{x}')\perp S$，那么 $\hat{x}'$ 也是 x 的最佳逼近。利用最佳逼近的唯一性（定理 8.1），$\hat{x}'=\hat{x}$。

根据上述两条定理，可直接得到以下结果。

定理 8.3 (投影定理)

若 S 是内积空间 V 的完备子空间，则 $V=S\odot S^\perp$。 ♡

证明 对于任意 $v\in V$，根据定理 8.1和8.2，存在 $\hat{v}\in S$ 使得 $(v-\hat{v})\perp S$，因此 $v=\hat{v}+(v-\hat{v})\in S\odot S^\perp$，即 $V\subseteq S\odot S^\perp\subseteq V$，最终得到 $V=S\odot S^\perp$。

命题 8.3

设 S,T 和 T' 均为内积空间 V 的子空间。

1. 若 $V=S\odot T$，则 $T=S^\perp$；
2. 若 $S\odot T=S\odot T'$，则 $T=T'$。 ♠

证明

1. $T\perp S$ 意味着 $T\subseteq S^\perp$。对于任意 $v\in S^\perp\subseteq V$，存在 $s\in S$ 和 $t\in T$ 使得 $v=s+t$，进而
$$0=\langle v,s\rangle=\langle s,s\rangle\quad\Longrightarrow\quad s=0\quad\Longrightarrow\quad v=t\in T, \tag{8.2.9}$$
从而 $S^\perp\subseteq T$，反向包含得证；
2. 根据上述结果，$T=S^\perp=T'$、

命题 8.4

设 H 为 Hilbert 空间，若 A 为 H 的子集，则 A 张成空间的闭包 $\operatorname{cspan}(A)=A^{\perp\perp}$。 ♠

证明 容易证明 $\operatorname{cspan}(A)^\perp\subseteq A^\perp$；对于任意 $v\in A^\perp$ 和任意 $u\in\operatorname{cspan}(A)$，存在 $a_1,\cdots,a_n\in A$ 和 $r_1,\cdots,r_n\in F$ 使得 $u=\sum_{i=1}^n r_ia_i$，进而 $\langle u,v\rangle=\sum_{i=1}^n r_i\langle a_i,v\rangle=0$，即 $A^\perp\subseteq\operatorname{cspan}(A)^\perp$。因此 $\operatorname{cspan}(A)^\perp=A^\perp$。根据投影定理 8.3，$\operatorname{cspan}(A)$ 是 V 的子空间意味着 $V=\operatorname{cspan}(A)\odot\operatorname{cspan}(A)^\perp=\operatorname{cspan}(A)\odot A^\perp$；而 $A^\perp$ 是 V 的子空间意味着 $V=A^\perp\odot A^{\perp\perp}$。根据命题 8.3，$A^{\perp\perp}=\operatorname{cspan}(A)$。

8.2.2 Fourier 展开

现在我们希望进一步考察最佳逼近。我们的目标是对 Hilbert 空间内的任意向量 x，从闭子空间 S 找到关于 x 最佳逼近的显式表达。这里我们仅讨论最常见的 S 维度有限和可数无穷的情形，不可数无穷维的情形参见 [1]。

8.2.2.1 有限维情形

关于有限维情形的结论，我们已经在第5.4.1小节进行了充分讨论，这里仅给出结论。

定理 8.4

设 $\mathcal{O}=\{u_1,\cdots,u_n\}$ 为 Hilbert 空间 H 内的有限标准正交集。对于任意 $x\in H$，x 的 Fourier 展开 $\hat{x}=\sum_{k=1}^{n}\langle x,u_k\rangle u_k$ 为 x 在 $\operatorname{span}(\mathcal{O})$ 内的最佳逼近，满足 Bessel 不等式 $\|\hat{x}\|\leq\|x\|$，等号成立当且仅当 $x\in\operatorname{span}(\mathcal{O})$。

8.2.2.2 可数无穷维情形

由于在可数无穷维情形中可能涉及无穷求和，因此会引出收敛性的问题。我们从以下结论入手。

定理 8.5

设 $\mathcal{O}=\{u_1,u_2,\cdots\}$ 为 Hilbert 空间 H 内的一组可数无穷标准正交集，则级数 $\sum_{k=1}^{\infty}r_ku_k$ 在 H 内收敛当且仅当 $\sum_{k=1}^{\infty}|r_k|^2$ 在 $\mathbb{R}$ 内收敛。若该级数收敛，则其无条件收敛，即任何重排这些项顺序形成的新级数也收敛。最后，若级数 $\sum_{k=1}^{\infty}r_ku_k$ 收敛，则 $\left\|\sum_{k=1}^{\infty}r_ku_k\right\|^2=\sum_{k=1}^{\infty}|r_k|^2$。

证明 设 $s_n=\sum_{k=1}^{n}r_ku_k,p_n=\sum_{k=1}^{n}|r_k|^2$，那么对于自然数 $m\leq n$，$\|s_n-s_m\|^2=\left\|\sum_{k=m+1}^{n}r_ku_k\right\|^2=\sum_{k=m+1}^{n}|r_k|^2=|p_n-p_m|$。因此 (s_n) 是 H 内的 Cauchy 列当且仅当 (p_n) 是 $\mathbb{R}$ 内的 Cauchy 列。由于 H 和 $\mathbb{R}$ 均完备（定理 1.14），(s_n) 收敛当且仅当 (p_n) 收敛，此时级数重排不影响 $\sum_{k=1}^{\infty}|r_k|^2$ 的结果。

定理 8.6

设 $\mathcal{O}=\{u_1,u_2,\cdots\}$ 为 Hilbert 空间 H 内一组可数无穷的标准正交集。对于任意 $x\in H$，x 的 Fourier 展开 $\hat{x}=\sum_{k=1}^{\infty}\langle x,u_k\rangle u_k$ 无条件收敛，并且是 x 在 $\operatorname{cspan}(\mathcal{O})$ 内的最佳逼近，满足 Bessel 不等式 $\|\hat{x}\|\leq\|x\|$，等号成立当且仅当 $x\in\operatorname{cspan}(\mathcal{O})$。

证明 对于任意 $y\in\operatorname{cspan}(\mathcal{O})$，存在 $r_{i_k}\in F$ 使得 $y=\sum_{k=1}^{n}r_{i_k}u_{i_k}$，因此

$$\langle x-\hat{x},y\rangle=\sum_{k=1}^{n}\overline{r_{i_k}}\langle x-\hat{x},u_{i_k}\rangle=\sum_{k=1}^{n}\overline{r_{i_k}}(\langle x,u_{i,k}\rangle-\langle x,u_{i_k}\rangle)=0,\tag{8.2.10}$$

即 $(x-\hat{x})\perp\operatorname{cspan}(\mathcal{O})$。根据定理 8.2，$\hat{x}$ 是 x 在 $\operatorname{cspan}(\mathcal{O})$ 内的最佳逼近，并且由于 $\hat{x}\in\operatorname{cspan}(\mathcal{O})$，

$$\|x\|^2=\|(x-\hat{x})+\hat{x}\|^2=\|x-\hat{x}\|^2+\|\hat{x}\|^2\geq\|\hat{x}\|^2,\tag{8.2.11}$$

等号成立当且仅当 $\|x-\hat{x}\|=0$，即 $x=\hat{x}\in\operatorname{cspan}(\mathcal{O})$。另外，$\|x\|^2\geq\|\hat{x}\|^2\geq\left\|\sum_{k=1}^{n}\langle x,u_k\rangle u_k\right\|^2=\sum_{k=1}^{n}|\langle x,u_k\rangle|^2$ 对所有自然数 n 成立。利用 $\mathbb{R}$ 上极限的单调性（定理 1.15），$\sum_{k=1}^{\infty}|\langle x,u_k\rangle|^2=\lim_n\sum_{k=1}^{n}|\langle x,u_k\rangle|^2\leq\|x\|^2$，故级数 $\sum_{k=1}^{\infty}|\langle x,u_k\rangle|^2$ 收敛。根据定理 8.5，级数 $\sum_{k=1}^{\infty}\langle x,u_k\rangle u_k$ 无条件收敛。

8.2.2.3 任意维数情形

在本小节中，我们略去所有证明细节，只给出最终结论。

定义 8.2 (收敛)

设 $\mathcal{K}=\{x_k|k\in K\}$ 为向量空间 V 内的任意向量族，若对于任意正数 ϵ，存在有限集 $S\subseteq K$，使得任意包含 S 的有限集 T 满足 $\left\|\sum_{k\in T}x_k-x\right\|\leq\epsilon$，则称求和 $\sum_{k\in K}x_k$ 收敛到向量 $x\in V$，记作 $x=\sum_{k\in K}x_k$。

上述定义保持收敛列的线性性与连续性。

命题 8.5

设 $\mathcal{K}=\{x_k|k\in K\}$ 为内积空间 V 中的任意向量族。若 $\sum_{k\in K}x_k=x$ 且 $\sum_{y\in K}y_k=y$，则其满足

1. 线性性：对于任意 $r,s\in F$，$\sum_{k\in K}(rx_k+sy_k)=rx+sy$；
2. 连续性：$\sum_{k\in K}\langle x_k,y\rangle=\langle x,y\rangle$ 且 $\sum_{k\in K}\langle y,x_k\rangle=\langle y,x\rangle$。

在此收敛性的定义下，同样有类似定理 8.5的结论。

命题 8.6

设 $\mathcal{O}=\{u_k|k\in K\}$ 为 Hilbert 空间内的任意标准正交向量族，则级数 $\sum_{k\in K}r_ku_k$ 和 $\sum_{k\in K}|r_k|^2$ 同敛散。若两者均收敛，则 $\left\|\sum_{k\in K}r_ku_k\right\|^2=\sum_{k\in K}|r_k|^2$。

关于最佳逼近，其结论与有限维和可数无穷维情形的结论基本一致。

定理 8.7

设 $\mathcal{O}=\{u_k|k\in K\}$ 为 Hilbert 空间 H 内的标准正交向量族。对于任意 $x\in H$，x 的 Fourier 展开 $\hat{x}=\sum_{k\in K}\langle x,u_k\rangle u_k$ 在 H 内收敛，并且是 $\operatorname{cspan}(\mathcal{O})$ 内 x 的唯一最佳逼近，满足 Bessel 不等式 $\|\hat{x}\|\leq\|x\|$，等号成立当且仅当 $x\in\operatorname{cspan}(\mathcal{O})$。

8.3 Hilbert 基与 Hilbert 维数

8.3.1 Hilbert 基

在5.3.3小节中我们定义了有限维内积空间中的 Hilbert 基，现在给出无穷维情形下 Hilbert 基的等价表述。

定理 8.8

设 $\mathcal{O}=\{u_k|k\in K\}$ 为 Hilbert 空间 H 内的标准正交向量族，则下列表述等价：

1. $\mathcal{O}$ 是 Hilbert 基；
2. $\mathcal{O}^\perp=\{0\}$；
3. $\mathcal{O}$ 是 H 的完全子集（total subset），即 $\operatorname{cspan}(\mathcal{O})=H$；
4. $x=\hat{x}$ 对所有 $x\in H$ 成立；
5. Bessel 不等式的等号对所有 $x\in H$ 成立，即 $\|x\|=\|\hat{x}\|$ 对所有 $x\in H$ 成立；
6. Parseval 恒等式：$\langle x,y\rangle=\langle\hat{x},\hat{y}\rangle$ 对所有 $x,y\in H$ 成立，即 $\langle x,y\rangle=\sum_{k\in K}\langle x,u_k\rangle\overline{\langle y,u_k\rangle}$。

证明

1. 若表述 1 成立，假设 $\mathcal{O}^\perp$ 存在其他非零向量 v，则 $\mathcal{O}\cup\{v\}$ 构成新的标准正交集，但这与 $\mathcal{O}$ 的极大性矛盾。故 $\mathcal{O}^\perp=\{0\}$，表述 2 成立；
2. 若表述 2 成立，则根据投影定理 8.3和命题 8.4的证明，$H=\operatorname{cspan}(\mathcal{O})\odot\operatorname{cspan}(\mathcal{O})^\perp=\operatorname{cspan}(\mathcal{O})\odot\mathcal{O}^\perp=\operatorname{cspan}(\mathcal{O})\odot\{0\}=\operatorname{cspan}(\mathcal{O})$，表述 3 成立；
3. 若表述 3 成立，根据定理 8.2和8.7，$(x-\hat{x})\perp H$。假设 $v=x-\hat{x}\in H$ 非零，则 $\langle v,v\rangle\neq0$，这与 $v\perp H$ 矛盾。因此 $x=\hat{x}$，表述 4 成立；
4. 若表述 4 成立，则表述 5 显然成立；
5. 若表述 5 成立，则在实内积空间中，利用极化恒等式（定理 5.2）和命题 8.6，

$$\begin{aligned}\langle x,y\rangle&=\frac{1}{4}\left(\|x+y\|^2-\|x-y\|^2\right)=\frac{1}{4}\left(\left\|\sum_{k\in K}\langle x+y,u_k\rangle u_k\right\|^2-\left\|\sum_{k\in K}\langle x-y,u_k\rangle u_k\right\|^2\right)\\&=\frac{1}{4}\left(\sum_{k\in K}|\langle x+y,u_k\rangle|^2-\sum_{k\in K}|\langle x-y,u_k\rangle|^2\right)=\frac{1}{4}\sum_{k\in K}\left(|\langle x+y,u_k\rangle|^2-|\langle x-y,u_k\rangle|^2\right)\\&=\frac{1}{2}\sum_{k\in K}\left(\langle x,u_k\rangle\langle y,u_k\rangle+\langle y,u_k\rangle\langle x,u_k\rangle\right)=\sum_{k\in K}\langle x,u_k\rangle\overline{\langle y,u_k\rangle}=\langle\hat{x},\hat{y}\rangle,\end{aligned}\tag{8.3.1}$$

其中已利用实向量空间中内积的对称性。复向量空间的情形类似可证明，因此表述 6 成立；

6. 若表示 6 成立，假设存在标准向量 $v\in H$ 使得 $v\in\mathcal{O}^\perp$，那么根据定理 8.2和8.7，$v-\hat{v}\in\mathcal{O}^\perp$。利用子空

间的代数封闭性，$\hat{v} \in \mathcal{O}^{\perp} = \operatorname{cspan}(\mathcal{O})^{\perp}$；但同时 $\hat{v} \in \operatorname{cspan}(\mathcal{O})$，因此 $\hat{v} = 0$。这意味着 $\langle v, v\rangle = \langle \hat{v}, \hat{v}\rangle = 0$，即 $v = 0$，矛盾。因此 $\mathcal{O}$ 为极大标准正交向量族，表述 1 成立。

8.3.2 Hilbert 维数

类似 Hamel 基，我们希望 Hilbert 空间内的所有 Hilbert 基具有相同的基数，可据此定义其 Hilbert 维数。

定理 8.9 (Hilbert 维数)

Hilbert 空间 H 的所有 Hilbert 基具有相同的基数，该基数称为 H 的 Hilbert 维数（Hilbert dimension），记作 $\operatorname{hdim}(H)$。

♡

证明 对于有限维情形，根据定理 5.8，所有的 Hilbert 基都具有相同的基数 $\dim(H)$；对于无穷维情形，设 $\mathcal{B} = \{b_k | k \in K\}$ 和 $\mathcal{C} = \{c_j | j \in J\}$ 为 H 的 Hilbert 基。那么根据定理 8.8，$H = \operatorname{cspan}(\mathcal{C})$，这意味着存在可数指标集 J_k 和 $r_j \in F$ 使得 $b_k = \sum_{j\in J_k} r_j c_j$。对于任意 $j \in J_k$，$\langle b_k, c_j\rangle = \sum_{j'\in J_k} r_{j'} \langle c_{j'}, c_j\rangle = r_j$，因此 $b_k = \sum_{j\in J_k} \langle b_k, c_j\rangle c_j$。若存在某个 $c_j \perp \mathcal{B}$，则 $\mathcal{B} \cup \{c_j\}$ 也是标准正交向量族，这与 $\mathcal{B}$ 的极大性矛盾，因此必然存在 $k \in K$ 使得 $j \in J_k$。这意味着 $J \subseteq \cup_{k\in K} J_k \subseteq J$，即 $J = \cup_{k\in K} J_k$。根据命题 1.4和1.6以及 $\aleph_0$ 是最小无穷势的事实，$|J| \leq \aleph_0 \cdot |K| = \max\{\aleph_0, |K|\} = |K|$。对称地 $|K| \leq |J|$，那么 Schrödinger-Berstein 定理（定理 1.2）给出 $|J| = |K|$。

定理 8.10

Hilbert 空间等距同构当且仅当其含有相同的 Hilbert 维数。

♡

证明 设两个 Hilbert 空间分别为 H_1 和 H_2，$\mathcal{B} = \{b_k | k \in K\}$ 为 H_1 的一组 Hilbert 基，那么 $\mathcal{B}^{\perp} = \{0\}$。对于充分性，设等距同构为 τ，考虑 $v \in (\tau\mathcal{B})^{\perp} = \operatorname{cspan}(\tau\mathcal{B})^{\perp}$，$\langle v, b_k\rangle = 0$ 对所有 $k \in K$ 成立，因此存在有限指标集 K_v 和 $r_k \in F$，使得 $\tau^{-1}v = \sum_{k\in K_v} r_k b_k$，即 $v = \sum_{k\in K_v} r_k(\tau b_k) \in \operatorname{cspan}(\tau\mathcal{B})$。这意味着 $v = 0$，即 $(\tau\mathcal{B})^{\perp} = \{0\}$，因此 $\tau\mathcal{B}$ 是 H_2 的 Hilbert 基，双射要求 $|\mathcal{B}| = |\tau\mathcal{B}|$。

对于必要性，设 $\mathcal{C} = \{c_j | j \in J\}$ 为 H_2 的一组 Hilbert 基，则 $|J| = |K|$，即存在双射使得对于任意 $j \in J$，存在唯一的 $k \in K$ 与之对应。定义双射 $\tau : \mathcal{C} \to \mathcal{B}$ 为 $\tau(c_j) = b_k$，并通过线性性将其定义域延拓到 H_2 上。那么根据定理 8.8，对于任意 $v \in H_2$，

$$\|\tau v\|^2 = \left\|\sum_{j\in J} \langle v, c_j\rangle \tau c_j\right\|^2 = \left\|\sum_{j\in J} \langle v, c_j\rangle b_k\right\|^2 = \sum_{j\in J} |\langle v, c_j\rangle|^2 = \left\|\sum_{j\in J} \langle v, c_j\rangle c_j\right\|^2 = \|v\|^2. \tag{8.3.2}$$

根据定理 5.4，τ 为等距映射，进而为等距同构。

8.4 Riesz 表示定理

这一节我们讨论任意维数 Hilbert 空间内的 Riesz 表示定理。事实上，并非所有 Hilbert 空间上线性的泛函都具有有限维情形的 Riesz 表示定理 5.11中的内积形式。对于任意给定的非零向量 $y \in H$，定义函数 $f_y : H \to F$ 为 $f_y(x) = \langle x, y\rangle$。显然 f_y 是 H 上的线性泛函，并且利用 Cauchy-Schwarz 不等式，$|f_y(x)|/\|x\| \leq \|y\|$ 对任意 $x \in H$ 成立，因此 $\sup_{x\neq 0}(|f_y(x)|/\|x\|) = \|y\|$，等号在 $y = x$ 时成立。这引出了如下定义。

定义 8.3 (算子的有界性)

设 $\tau : H_1 \to H_2$ 为从 H_1 到 H_2 的线性变换，若 $\sup_{x\neq 0}(\|\tau x\|/\|x\|) < \infty$，则称 τ 有界（bounded）。若 τ 有界，则将该上确界记为 $\|\tau\|$，并称其为 τ 的范数。特别地，H 上的有界线性泛函 $f : H \to F$ 构成的集合称为 H 的连续对偶空间（continuous dual space）或共轭空间（conjugate space）。

♣

以下命题给出了有界线性变换的等价范数表示。

命题 8.7

设 $\tau: H_1 \to H_2$ 为有界线性变换，则

$$\|\tau\| = \sup_{\|x\|=1} \|\tau x\| = \sup_{\|x\|\leq 1} \|\tau x\| = \inf\{c \in \mathbb{R} | \|\tau x\| \leq c\|x\| \text{ for all } x \in H\}. \tag{8.4.1}$$

♠

证明 首先利用范数的线性性，$\|\tau\| = \sup_{x\neq 0} \|\tau(x/\|x\|)\| \xlongequal{x/\|x\|\to x} \sup_{\|x\|=1} \|\tau x\|$，因此 $\|\tau\| \leq \sup_{\|x\|\leq 1} \|\tau x\|$。容易验证该上确界不在 $x = 0$ 处取到（对于零映射，任取 $x \neq 0$），$\|\tau x\| = \|\tau(x/\|x\|)\| \cdot \|x\| \leq \|\tau\| \cdot 1$，因此 $\sup_{\|x\|\leq 1} \|\tau x\| \leq \|\tau\|$。综上，$\|\tau\| = \sup_{\|x\|\leq 1} \|\tau x\|$。最后记 $\delta = \inf\{c \in \mathbb{R} | \|\tau x\| \leq c\|x\| \text{ for all } x \in H\}$，显然 $\|\tau x\|/\|x\| = \|\tau(x/\|x\|)\| \leq \|\tau\|$ 对任意 $x \in H\backslash\{0\}$ 成立，故 $\delta \leq \|\tau\|$；另一方面对于任意正数 ϵ，存在 $c \in \mathbb{R}$ 使得 $\delta + \epsilon > c \geq \|\tau x\|/\|x\|$ 对所有 $x \in H$ 成立，故 $\delta + \epsilon \geq \sup_{\|x\|=1} \|\tau x\| = \|\tau\| \geq \delta$。利用 ϵ 的任意性，$\|\tau\| = \delta$。

有界线性变换的重要性之一体现在其保持连续性上。

命题 8.8

设 $\tau: H_1 \to H_2$ 为线性变换，则 τ 有界当且仅当 τ 连续。

♠

证明 若 τ 有界，那么对于任意 $x, y \in H_1$，$\|\tau x - \tau y\| = \|\tau(x-y)\| \leq \|\tau\|\|x-y\|$。因此对于任意正数 ϵ，当 $\|x-y\| < \epsilon/\|\tau\|$ 时，$\|\tau x - \tau y\| < \epsilon$，因此根据命题 1.17，$\tau$ 连续；反之若 τ 连续，则根据命题 1.17，存在 $\delta > 0$ 使得 $\|y\| \leq \delta$ 时 $\|\tau y\| \leq 1$。那么对于任意满足 $\|x\| = 1$ 的 $x \in H_1$，$\|\tau x\| = \|\tau(\delta x)\|/\delta\|x\| \leq 1/\delta\|x\| = 1/\delta$，因此 $\|\tau\| = \sup_{\|x\|=1} \|\tau x\| \leq 1/\delta$，即 τ 有界。

现在我们证明任意维度 Hilbert 空间的 Riesz 表示定理。

定理 8.11 (Riesz 表示定理)

设 H 为 Hilbert 空间。对于 H 上任意有界线性泛函 f，存在唯一的 $z_0 \in H$ 使得 $f(x) = \langle x, z_0\rangle$ 对所有 $x \in H$ 成立。此外，$\|f\| = \|z_0\|$。

♡

证明 若 $f = 0$，取 $z_0 = 0$ 即可。对于非零的 f，记 $K = \ker(f) \neq H$。根据命题 8.8，f 连续，因此对于 K 中任意收敛到 $x \in H$ 的序列 (x_n)，均有 $f(x) = \lim_n f(x_n) = 0$，因此 $x \in K$，即 K 为闭集。根据投影定理 8.3，$J = K \odot K^\perp$。根据第一同构定理 4.4以及命题 4.1之前的讨论，$V/K \cong \mathrm{im}(f) = F$；再根据推论 3.2和4.1，$K^\perp \cong V/K \cong F$，因此 $\dim(K^\perp) = \dim(F) = 1$。

显然 $\langle x, z\rangle = 0 = f(x)$ 对任意 $x \in K$ 成立当且仅当 $z \in K^\perp$，因此我们希望限制在在 $K^\perp$ 上找到一个合适的 z_0，同时使得 $f(x) = \langle x, z_0\rangle$ 对任意 $x \in K^\perp$ 也成立。由于 $\dim(K^\perp) = 1$，任意非零的 $z \in K^\perp$ 均为作为基。首先希望 $f(z_0) = \langle z_0, z_0\rangle$，由此定义了 z_0 的范数。对于任意 $x \in K^\perp$，存在唯一的 $r \in F$ 使得 $x = rz_0$，因此 $f(x) = rf(z_0) = r\langle z_0, z_0\rangle = \langle rz_0, z_0\rangle = \langle x, z_0\rangle$。综上，对于任意 $x \in H$，存在 $r = \langle x, z_0\rangle / \langle z_0, z_0\rangle \in F$ 使得 $\langle x - rz_0, z_0\rangle = 0$，因此 $x = (x - rz_0) + rz_0 \in K \odot K^\perp$，进而

$$f(x) = 0 + rf(z_0) = r\langle z_0, z_0\rangle = \langle rz_0, z_0\rangle = \langle x - (x - rz_0), z_0\rangle = \langle x, z_0\rangle. \tag{8.4.2}$$

若存在 $z_0' \in H$ 也使得 $f(x) = \langle x, z_0'\rangle$ 对所有 $x \in H$ 成立，那么 $0 = f(x) - f(x) = \langle x, z_0 - z_0'\rangle$ 对所有 $x \in H$ 成立，即 $z_0' = z_0$，故 z_0 唯一。

最后，利用 Cauchy-Schwarz 不等式，$|f(x)|/\|x\| = |\langle x, z_0\rangle|/\|x\| \leq \|z_0\|$，等号成立当且仅当 $x \in K^\perp$，那么 $\|f\| \leq \|z_0\|$；另一方面，$\|z_0\| = |f(z_0)|/\|z_0\| \leq \|f\|$，因此 $\|f\| = \|z_0\|$。

第二部分

凸几何

第 9 章 仿射几何

在本章中，假设所有的向量空间均为有限维的。

目录

9.1 基本概念 74
9.2 仿射组合与仿射包 75
9.3 仿射子空间格 76
9.4 仿射无关 78
9.5 仿射变换 79

9.1 基本概念

定义 9.1 (仿射几何)

设 S 为向量 V 的子空间，陪集 $v+S=\{v+s|s\in S\}$ 称为 V 中的一个平坦[a]（flat），其中 S 是该平坦的基底子空间（base），v 是平坦代表元（flat representation）；有时也称 $v+S$ 为 S 的平移（translate）。V 中所有平坦构成的集合 $\mathcal{A}(V)$ 称为 V 的仿射几何（affine geometry）。$\mathcal{A}(V)$ 的维数 $\dim(\mathcal{A}(V))$ 定义为 $\dim(V)$。

[a]有时也将其意译为“仿射子空间”，不过这一术语在英文中有更明确的“affine subspace”作为对应

♣

注 尽管一个平坦可能有许多代表元，该平坦也只有一个基底。不妨考虑某个平坦同时具有 $x+S$ 和 $y+T$ 两种表示，其中 $x\neq y$，那么由于 $0\in S$，有 $x\in y+T$，因此 $S=(y-x)+T=T$。

定义 9.2

平坦 $x+S$ 的维数定义为 $\dim(S)$。维数为 k 的平坦称为 k-平坦（k-flat）。0-平坦是一个点（point），1-平坦是一条直线（line），2-平坦是一个平面（plane）。维数为 $\dim(\mathcal{A}(V))-1$ 的平坦称为超平面（hyperplane）。

♣

注 当 $v=0$ 时，1-平坦和 2-平坦分别退化为过零点的直线和平面，也就是该向量空间的子空间。

定义 9.3 (平行)

若两个平坦 $X=x+S$ 和 $Y=y+T$ 满足 $S\subseteq T$ 或 $T\subseteq S$，则称其平行（parrallel），记作 $X\|T$。

♣

习惯上我们用 S,T 表示 V 的子空间，而用 X,Y 表示 V 的平坦。以下是平坦的基本交集性质。

命题 9.1

设 S 和 T 为 V 的子空间，$X=x+S$ 和 $Y=y+T$ 是 V 中的平坦。

1. 以下表述等价：
   (1) 存在 X 的平移在 Y 中，即存在 $w\in V$ 使得 $w+X\subseteq Y$；
   (2) 存在 S 的平移在 T 中，即存在 $v\in V$ 使得 $v+S\subseteq T$；
   (3) $S\subseteq T$；
2. 以下表述等价：
   (1) X 和 Y 互为平移，即存在 $w\in V$ 使得 $w+X=Y$；
   (2) S 和 T 互为平移，即存在 $v\in V$ 使得 $v+S=T$；
   (3) $S=T$；

3. $X\cap Y\neq\varnothing, S\subseteq T \Leftrightarrow X\subseteq Y$；
4. $X\cap Y\neq\varnothing, S=T \Leftrightarrow X=Y$；
5. 若 $X\|Y$，则 $X\subseteq Y, Y\subseteq X$ 或 $X\cap Y=\varnothing$；
6. $X\|Y$ 当且仅当存在这些平坦之一的平移，其包含在另一平坦内。

♠

证明

1. 若表述 1 成立，即存在 $w\in V$ 使得 $w+X\subseteq Y$，则存在 $v=w+x-y$ 使得 $v+S\subseteq T$，表述 2 成立；若表述 2 成立，则由于 $0\in S$，我们有 $v\in T$，进而 $S\subseteq T-v=T$，表述 3 成立；若表述 3 成立，则存在 $w=x-y$ 使得 $w+X\subseteq Y$，表述 1 成立；
2. $S=T$ 当且仅当 $S\subseteq T$ 且 $T\subseteq S$，根据性质 1 得证；
3. 对于充分性，性质 1 给出存在 $w\in V$ 使得 $w+X\subseteq Y$。对于任意 $v\in X\cap Y$，存在 $y\in Y$ 使得 $w+v=y$，进而 $w=y-v\in Y$，从而 $X\subseteq Y-w=Y$；对于必要性，性质 1 直接给出 $S\subseteq T$，并且 $X\cap Y=X\neq\varnothing$；
4. $S=T$ 当且仅当 $S\subseteq T$ 且 $T\subseteq S$，根据性质 3 得证；
5. $X\|Y$ 意味着 $S\subseteq T$ 或 $T\subseteq S$，根据性质 1，此时存在 $w,v\in V$ 使得 $w+X\subseteq Y$ 或 $v+Y\subseteq Y$。若 $w\in Y$ 或 $v\in X$，那么 $X\subseteq Y-w=Y$ 或 $Y\subsetneq X-v=X$；反之若 $w\notin Y$ 且 $v\notin X$，首先考虑 $S\subseteq T$ 的情况。假设存在 $u\in X\cap Y$，那么存在 $y\in Y$ 使得 $w+u=y$，即 $w=y-u\in Y$，矛盾。因此 $X\cap Y=\varnothing$，这对于 $T\subseteq S$ 的情形也成立；
6. 根据性质 1 和性质 5，该等价性是显然的。

9.2 仿射组合与仿射包

在向量空间中，可通过向量集的线性组合生成一个子空间；而在仿射子空间中，同样存在相应的概念。

9.2.1 仿射组合

类似线性组合，我们可定义仿射组合。

定义 9.4 (仿射组合)

设 V 为向量空间，$x_i\in V$，那么线性组合

$$r_1x_1+\cdots+r_nx_n,\quad \text{where } r_i\in F \text{ and } \sum_{i=1}^{n}r_i=1 \tag{9.2.1}$$

称为向量集 $\{x_1,\cdots,x_n\}$ 的仿射组合（affine combination）。

♣

若 V 的非空子集 X 在 X 中向量的任意仿射组合下方便，则称其仿射封闭（*affine closed*）；若 X 在 X 中任意两个向量的仿射组合下封闭，则称其 2-仿射封闭（*two-affine closed*）。

命题 9.2

设 V 是特征不为 2 的域 F 上的向量空间，则 V 的子集 X 仿射封闭当且仅当其 2-仿射封闭。

♠

证明 充分性是显然的；对于必要性，采用归纳法，其中 $n=2$ 的情形自然成立。当 $n>2$ 时，假设结论对于任意不多于 $(n-1)$ 个向量的仿射组合成立，那么考虑仿射组合 $z=r_1x_1+\cdots+r_nx_n$。若存在某个 $r_i=1$，则其余 $r_j=0\,(j\neq i)$，此时 $z=x_i\in V$，该情况平凡；因此考虑所有 r_i 均不为 1 的情况，此时可将 z 改写为

$$z=r_1x_1+(1-r_1)\left(\frac{r_2}{1-r_1}x_2+\cdots+\frac{r_n}{1-r_1}x_n\right). \tag{9.2.2}$$

由于 $\sum_{i=2}^{n}r_i/(1-r_1)=(1-r_1)/(1-r_1)=1$，根据归纳假设，$y=\sum_{i=2}^{n}r_ix_i/(1-r_1)\in X$，进而 2-仿射封闭性要求 $z=r_1x+(1-r_1)y\in X$。因此 X 内任意仿射组合均在 X 内，即 X 仿射封闭。

注 上述命题中，特征不为 2 的要求是必须的。对于 $F=\mathbb{Z}_2$，存在 F^2 的子集 $X=\{(0,0),(1,0),(0,1)\}$ 是 2-仿射封闭的，但 $(0,0)+(1,0)+(1,0)=(1,1)\notin X$，即 X 不是仿射封闭的。

定理 9.1

向量空间的非空子集 X 是平坦当且仅当 X 是仿射封闭的。此外，若 F 特征不为 2，则 X 是平坦当且仅当 X 是 2-仿射封闭的。

♡

证明 若 X 是平坦，设其基底子空间为 S，x 为某个平坦代表元。对于任意 $x_i\in X$，存在 $s_i\in S$ 使得 $x_i=x+s_i$。因此利用 S 的封闭性，这些向量的仿射组合 $\sum_{i=1}^n r_ix_i=\sum_{i=1}^n r_i(x+s_i)=x+\sum_{i=1}^n r_is_i\in x+S=X$，其中已利用 $\sum_{i=1}^n r_i=1$，因此 X 仿射封闭；反之若 X 仿射封闭，任取 $x\in X$，记 $S=X-x$。对于 X 内的任意向量集 $\{x_1,\cdots,x_n\}$，记 $s_i=x_i-x$。那么对于任意 $r_i\in F$，利用 X 的仿射封闭性，

$$x+\sum_{i=1}^n r_is_i=\left(1-\sum_{i=1}^n r_i\right)x+\sum_{i=1}^n r_ix_i\in X, \tag{9.2.3}$$

因此 $\sum_{i=1}^n r_is_i\in X-x=S$，即 S 代数封闭，故 S 为子空间，从而 X 为平坦。

9.2.2 仿射包

类似向量集张成的子空间，我们可定义仿射包。

定义 9.5

设 X 为 V 中的非空向量集。

1. X 的仿射包（affine hull）是包含 X 的最小平坦，记作 $\mathrm{affhull}(X)$；
2. X 的仿射张成（affine span）是 X 中向量的所有仿射组合构成的集合，记作 $\mathrm{affspan}(X)$。

♣

类似向量集张成的子空间是包含该向量集的最小子空间这一事实，仿射包与仿射张成实际上是等价的。

定理 9.2

设 X 为 V 的非空子集，则

$$\mathrm{affhull}(X)=\mathrm{affspan}(X)=x+\mathrm{span}(X-x), \tag{9.2.4}$$

此时 $\dim(\mathrm{affspan}(X))=\dim(\mathrm{span}(X-x))$。以后对仿射包与仿射张成不加区分，统一记为 $\mathrm{aff}(X)$。

♡

证明 根据定理 9.1,X 的仿射包是仿射封闭的，即 X 的任意仿射组合均在 $\mathrm{affhull}(X)$ 内，因此 $\mathrm{affspan}(X)\subseteq\mathrm{affhull}(X)$；对于反向包含，我们希望证明 $\mathrm{affspan}(X)$ 是包含 X 的一个平坦，那么根据仿射包的定义，$\mathrm{affhull}(X)\subseteq\mathrm{affspan}(X)$，从而结论得证。任取 $x\in X=\{x_j|j\in J\}$，记 $S=\mathrm{affspan}(X)-x$。对于任意 $y,y_i\in\mathrm{affspan}(X)$，存在 $s_{i,j}\in F$ 使得 $y_j=\sum_j s_{i,j}x_j$，其中 $\sum_j s_{i,j}=1$ 对所有 $i=1,\cdots,n$ 成立。因此对于任意和为 1 的 $r_i\in F$，

$$\sum_{i=1}^n r_iy_i=\sum_{i,j}r_is_{i,j}x_j,\quad \text{where } \sum_{i,j}r_is_{i,j}=\sum_i r_i\sum_j s_{i,j}=\sum_i r_i=1, \tag{9.2.5}$$

因此 $\sum_{i=1}^n r_iy_i\in\mathrm{affspan}(X)$，即 $\mathrm{affspan}(X)$ 仿射封闭。根据定理 9.1，$\mathrm{affspan}(X)$ 是包含 X 的一个平坦，故 $\mathrm{affhull}(X)\subseteq\mathrm{affspan}(X)$，反向包含成立。

9.3 仿射子空间格

子空间的交集仍然是子空间，尽管它可能是平凡的零空间。对于平坦，我们将证明：若其平坦的交集非空，则其也是平坦。不过平坦的交集可能为空集，我们需要对其稍作修正使其重新成为格（参见命题 2.2前的定义）。

定理 9.3

设 V 为向量空间，V 中所有平坦与空集共同构成的集族

$$\mathcal{A}_0(V)=\mathcal{A}(V)\cup\{\varnothing\} \tag{9.3.1}$$

在交集下构成完全格。特别地，

1. $\mathcal{A}_0(V)$ 在任意交集下是封闭的。若集族 $\mathcal{F}=\{X_k|k\in K\}=\{x_k+S_k|k\in K\}$ 具有非空交集，则
$$\bigcap\mathcal{F}=\bigcap_{k\in K}X_k=x+\bigcap_{k\in K}S_k, \tag{9.3.2}$$
其中 $x\in\cap_{k\in K}X_k$；换言之，交集的基底子空间是基底子空间的交集；
2. 集族 $\mathcal{F}=\{X_k|k\in K\}=\{x_k+S_k|k\in K\}$ 并集的上确界是包含 $\mathcal{F}$ 内所有元素的最小平坦，即
$$\bigvee\mathcal{F}=\operatorname{aff}\left(\bigcup_{k\in K}X_k\right). \tag{9.3.3}$$
特别地，若 $X=x+S$ 和 $y+T$ 为 V 中的平坦，则
$$X\vee Y=x+(\langle x-y\rangle+S+T)=\operatorname{aff}\{x,y\}+(S+T). \tag{9.3.4}$$
若 $X\cap Y\neq\varnothing$，则进一步有
$$X\vee Y=x+(S+T). \tag{9.3.5}$$

证明

1. $\cap_{k\in K}X_k=\varnothing\in\mathcal{A}_0(V)$ 是平凡的；若 $\cap_{k\in K}X_k\neq\varnothing$，其包含的任意向量集必然包含在所有 X_k 内。由于 X_k 是平坦，根据定理 9.1，该向量集在所有 X_k 内仿射封闭，因此其仿射组合在 $\cap_{k\in K}X_k$ 内。由此可知 $\cap_{k\in K}X_k$ 也是仿射封闭的，根据定理 9.1，$\cap_{k\in K}X_k\in\mathcal{A}(V)\subseteq\mathcal{A}_0(V)$。综上，$\mathcal{A}_0(V)$ 在任意交集下封闭。对于任意 $x\in\cap_{k\in K}$，$x-x_k\in S_k$ 对所有 $k\in K$ 成立，因此 $x+S_k=x_k+(x-x_k)+S_k=x_k+S_k$ 对所有 $k\in K$ 成立，从而
$$\bigcap_{k\in K}X_k=\bigcap_{k\in K}(x_k+S_k)=\bigcap_{k\in K}(x+S_k)=x+\bigcap_{k\in K}S_k. \tag{9.3.6}$$
2. 对于 $\mathcal{A}_0(V)$ 内的任意集族 $\mathcal{F}=\{X_k|k\in X\}$，首先 $\cup_{k\in K}X_k\subseteq\operatorname{aff}(\cup_{k\in K}X_k)$，因此 $\operatorname{aff}(\cup_{k\in K}X_k)\in\mathcal{A}_0(V)$ 是该集族的上界。另外根据仿射包的定义，任意包含 $\cup_{k\in K}X_k$ 的平坦也包含 $\operatorname{aff}(\cup_{k\in K}X_k)$，因此 $\operatorname{aff}(\cup_{k\in K}X_k)$ 为集族 $\mathcal{F}$ 的最小上界，即上确界 $\vee\mathcal{F}=\operatorname{aff}(\cup_{k\in K}X_k)$。
特别地，若 $X=x+S$ 和 $Y=y+T$ 为 V 中的平坦，显然 $x,y\in X\cup Y\subseteq\operatorname{aff}(X\cup Y)$。根据仿射几何定义 9.1后的讨论，$\operatorname{aff}(X\cup Y)$ 具有唯一基底子空间 U，即 $\operatorname{aff}(X\cup Y)=x+U=y+U$，因此 $x-y\in U$。根据命题 9.1，$X\subseteq X\cup Y\subseteq\operatorname{aff}(X\cup Y)$ 等价于 $S\subseteq U$，类似地 $T\subseteq U$。利用子空间的封闭性，$\langle x-y\rangle+S+T\subseteq U$，从而 $x+(\langle x-y\rangle+S+T)\subseteq x+U=\operatorname{aff}(X\cup Y)$；另一方面，容易验证 $x+(\langle x-y\rangle+S+T)$ 是包含 X 和 Y 的平坦，因此 $\operatorname{aff}(X\cup Y)\subseteq x+(\langle x-y\rangle+S+T)$。综上，
$$X\vee Y=\operatorname{aff}(X\cup Y)=x+(\langle x-y\rangle+S+T)=\operatorname{aff}\{x,y\}+S+T. \tag{9.3.7}$$
当 $X\cap Y\neq\varnothing$ 时，那么对于交集内的元素，存在 $s\in S$ 和 $t\in T$ 使得 $x+s=y+t$，即 $\langle x-y\rangle\subseteq S+T$，此时可将表达式简化为 $X\vee Y=x+(S+T)$。

根据此结果，我们可以很方便地讨论两个平坦的并集的维数。

推论 9.1

设 $X=x+S$ 和 $Y=y+T$ 为 V 的平坦。

1. 若 $X\cap Y\neq\varnothing$，则 $\dim(X\vee Y)=\dim(S+T)=\dim(X)+\dim(Y)-\dim(X\cap Y)$；
2. 若 $X\cap Y=\varnothing$，则 $\dim(X\vee Y)=\dim(S+T)+1$。

证明

1. 根据定理 2.7、推论 3.2和定理 9.3，此时 $\dim(X\vee Y)=\dim(S+T)=\dim(S)+\dim(T)-\dim(S\cap T)$；

2. 考虑 $v \in \langle x-y\rangle \cap (S+T)$，那么存在 $r \in F$ 以及 $s \in S$ 和 $t \in T$，使得 $v = r(x-y) = s+t$。假设 $r \neq 0$，那么 $x - s/r \in x+S = X$ 且 $y + t/r \in y+T = Y$，这意味着 $X \cap Y \neq \varnothing$，矛盾。因此 $r = 0$，即 $\langle x-y\rangle \cap (S+T) = \{0\}$，进而根据定理 2.7、推论 3.2和定理 9.3，此时

$$\dim(X \vee Y) = \dim(\langle x-y\rangle) + \dim(S+T) - \dim(\langle x-y\rangle) = \dim(S+T) + 1. \tag{9.3.8}$$

9.4 仿射无关

对应于向量的线性无关，也存在仿射无关的概念。

定理 9.4 (仿射无关)

设 X 为 V 的一个非空子集，则以下表述等价：

1. 对于所有 $x \in X$，集合 $(X-x)\backslash\{0\}$ 线性无关；
2. 对于所有 $x \in X$，$x \notin \text{aff}(X\backslash\{x\})$；
3. 对于任意向量 $x_i \in X$，
$$\sum_{i=1}^{n} r_i x_i = 0 \text{ and } \sum_{i=1}^{n} r_i = 0 \quad \Longrightarrow \quad r_i = 0 \text{ for all } i = 1, \cdots, n. \tag{9.4.1}$$
4. 对于所有 X 中向量的仿射组合，
$$\sum_{i=1}^{n} r_i x_i = \sum_{i=1}^{n} s_i x_i \quad \Longrightarrow \quad r_i = s_i \text{ for all } i = 1, \cdots, n. \tag{9.4.2}$$
5. 当 $X = \{x_1, \cdots, x_n\}$ 有限时，$\dim(\text{aff}(X)) = n-1$。

称满足上述任一条件的集合 X 为仿射无关的（affinely independent）。

♡

证明

1. 若表述 1 成立，假设存在 $x \in X$ 使得 $x \in \text{aff}(X\backslash\{x\})$，即存在不等于 x 的 $x_i \in X$ 和 $r_i \in F$，使得 $x = \sum_{i=1}^{n} r_i x_i$，其中 $\sum_{i=1}^{n} r_i = 1$，因此至少一个 r_i 非零。移项得 $\sum_{i=1}^{n} r_i(x_i - x) = 0$，可见向量集 $\{x_1 - x, \cdots, x_n - x\}$ 线性相关，矛盾。故 $x \notin \text{aff}(X\backslash\{x\})$ 对所有 $x \in X$ 成立，即表述 2 成立；
2. 若表述 2 成立，那么对于任意 $x_i \in X$ 以及和为 0 的系数 $r_i \in F$，当 $\sum_{i=1}^{n} r_i x_i = 0$ 时，假设存在 $r_i \neq 0$，那么 $x_i = \sum_{j\neq i}(-r_j/r_i)x_j$。注意到 $\sum_{j\neq i}(-r_j/r_i) = -(-r_i)/r_i = 1$ 且 $x_j \neq x_i$ $(j \neq i)$，因此 $x_i \in \text{aff}(X\backslash\{x_i\})$，矛盾。这意味着所有 r_i 均为零，表述 3 成立；
3. 若表述 3 成立，将表述 4 改写为 $\sum_{i=1}^{n}(r_i - s_i)x_i = 0$，则其显然成立；
4. 若表述 4 成立，根据定理 9.2，$\text{aff}(X) = x_1 + \text{span}(X - x_1) = x_1 + \text{span}((X-x_1)\backslash\{0\})$。假设向量集 $(X-x_1)\backslash\{0\}$ 线性相关，即存在不全为零 $r_i \in F$ 使得 $\sum_{i=2}^{n} r_i(x_i - x_1) = 0$，不妨设 $r_i \neq 0$，那么
$$x_i = x_1 - \sum_{j\neq i} \frac{r_j}{r_i}(x_j - x_1) = \left(1 + \sum_{j\neq i} \frac{r_j}{r_i}\right) x_1 - \sum_{j\neq i} \frac{r_j}{r_i} x_j \in \text{aff}(X), \tag{9.4.3}$$
但等式两侧 x_i 的系数不同，矛盾。因此 $(X-x_1)\backslash\{0\}$ 线性无关，进而 $\dim(\text{aff}(X)) = n-1$，表述 5 成立；
5. 若表述 5 成立，假设存在 $x \in X$ 使得 $(X-x)\backslash\{0\}$ 线性相关，即存在不全为零的 $r_i \in F$ 和 $x_i \in X$ 使得 $\sum_{i=1}^{n} r_i(x_i - x) = 0$。那么在有限子集 $X' = \{x, x_1, \cdots, x_n\}$ 中，根据定理 9.2，$n = \dim(\text{aff}(X')) = \dim(\text{span}(X'-x)) < |(X'-x)\backslash\{0\}| = n$，矛盾。因此 $(X-x)\backslash\{0\}$ 对所有 $x \in X$ 成立，即表述 1 成立。

类比向量空间中基的概念，对于平坦 $A = x + S$，若其子集 $\mathcal{B}$ 仿射无关且 $\text{aff}(\mathcal{B}) = A$，则称 $\mathcal{B}$ 为 A 的仿射基（*affine basis*）。从几何上看，由于仿射子空间是过原点的“平面”经平移得到的“平面”，因此仿射基对应平移前该子空间内一组基。下面我们对这一直觉进行严格的描述与证明。

定理 9.5

设 $A = x + S$ 为 n 维平坦，$\mathcal{B} = (x_1, \cdots, x_n)$ 为 S 的一组有序基，$(\mathcal{B} + x) \cup \{x\}$ 为 A 的有序仿射基，则每个 $v \in A$ 具有唯一仿射组合表示

$$v = r_1(x_1 + x) + \cdots + r_n(x_n + x) + r_{n+1}x, \tag{9.4.4}$$

系数 r_i 称为 v 关于有序仿射基 $\mathcal{B}$ 的重心坐标（barycentric coordinate）。

证明 对于任意 $v \in A$，存在唯一的 $r_i \in F$ 使得 $v - x = \sum_{i=1}^n r_i x_i \in S$，或者等价地写为 $v = \sum_{i=1}^n r_i(x_i + x) + r_{n+1}x \in \mathrm{aff}((\mathcal{B} + x) \cup \{x\})$，其中 $r_{n+1} = 1 - \sum_{i=1}^n r_i$。因此 $A = \mathrm{aff}((\mathcal{B} + x) \cup \{x\})$，同时 $(\mathcal{B} + x) \cup \{x\}$ 的仿射无关性由定理 9.4保证，因此 $(\mathcal{B} + x) \cup \{x\}$ 是 A 的仿射基。

9.5 仿射变换

定义 9.6 (仿射变换)

保持仿射组合的函数 $f : V \to V$ 称为仿射变换（affine transformation），即

$$\sum_{i=1}^n r_i = 1 \quad \Longrightarrow \quad f\left(\sum_{i=1}^n r_i x_i\right) \sum_{i=1}^n r_i f(x_i). \tag{9.5.1}$$

根据定义可以立即得到命题 9.2的对应结果。

命题 9.3

若 F 的特征不为 2，则函数 $f : V \to V$ 是仿射变换当且仅当其保持每对点的仿射组合，即

$$f(rx + (1 - r)y) = rf(x) + (1 - r)f(y). \tag{9.5.2}$$

证明 充分性是显然的；必要性可通过归纳法采用与命题 9.2类似的讨论直接证明。

定义 9.7 (平移)

设 $v \in V$，则对于所有 $v \in V$，仿射映射 $T_v : V \to V,\ x \mapsto x + v$ 称为沿 v 的平移（translation by v）。

对于任意 $\tau \in \mathcal{L}(V)$，容易验证任意复合 $T_v \circ \tau$ 是仿射的。有趣的是，任意仿射变换也具有此形式。

定理 9.6

函数 $f : V \to V$ 是仿射变换当且仅当存在 $v \in V$ 和 $\tau \in \mathcal{L}(V)$ 使得 $f = T_v \circ \tau$。

证明 必要性是显然的；对于充分性，设 $f(0) = -z$，取 $\tau = T_z \circ f$，那么对于任意 $u, v \in V$ 和 $r, s \in F$

$$\begin{aligned} \tau(ru + sv) &= (T_z \circ f)(ru + sv) = f(ru + sv) + z = f(ru + sv + (1 - r - s) \cdot 0) + z \\ &= rf(u) + sf(v) + (1 - r - s)(-z) + z = r(f(u) + z) + s(f(v) + z) \\ &= r(T_z \circ f)(u) + s(T_z \circ f)(v) = r\tau(u) + s\tau(v), \end{aligned} \tag{9.5.3}$$

因此 $\tau : V \to V$ 是线性变换，进而 $(T_{-z} \circ \tau)(v) = \tau(v) - z = f(v)$ 对所有 $v \in V$ 成立，故 $f = T_{-z} \circ \tau$。

推论 9.2

1. 两个仿射变换的复合是仿射变换；
2. 仿射变换 $T_v \circ \tau$ 是双射当且仅当 τ 是双射；
3. V 上所有双射仿射变换构成的集合在映射复合下构成一个群，称为 V 的仿射群（affine group）。

第 10 章 凸几何

在凸几何中，我们常常在 n 维 Euclidean 空间内处理问题，因此本章讨论的向量空间仅限于 $\mathbb{R}^n$。尽管凸性本身是纯代数性质，不过凸集本身不具有代数封闭的良好性质；在凸集上讨论更多的是一些特有的代数性质与拓扑性质，例如极值表示、支撑与分离，其中前者与高维 Bell 不等式的违背密切相关。

目录

10.1 代数性质 80
10.2 组合性质 83
10.3 拓扑性质 85
10.4 支撑与分离 86
10.5 极值表示 89

10.1 代数性质

定义 10.1 (凸集)

对于集合 $A \subseteq \mathbb{R}^n$，若对于 $x, y \in A$ 和 $\alpha \in [0,1]$ 均有 $(1-\alpha)x + \alpha y \in A$，则称其为凸的（convex）。 ♣

例 10.1

1. 对于不同的 $x, y \in \mathbb{R}^n$，最简单的非平凡凸集是开线段（*open segment*）$(x,y) = \{(1-\alpha)x + \alpha y | \alpha \in (0,1)\}$ 和闭线段（*close segment*）$[x,y] = \{(1-\alpha) + \alpha y | \alpha \in [0,1]\}$。类似可定义半开线段 $(x,y]$ 和 $[x,y)$；
2. $\mathbb{R}^n$ 中的平坦也是凸集。更具体地，给定线性形式 $f \neq 0$ 和某个 $\alpha \in \mathbb{R}$，则 $\{f = \alpha\}$ 是凸的。若 $\{f = \alpha\}$ 表示超平面，则开半空间（*open halfspace*）$\{f < \alpha\}, \{f > \alpha\}$ 和闭半空间（*close halfspace*）也是凸集；
3. 定义 1.19中的开球与闭球也是凸的，并且在 $B(x,r)$ 的边界上添加任意子集得到的集合仍是凸的。

定义 10.2 (凸组合)

设 $k \in \mathbb{N}$，$x_1, \cdots, x_k \in \mathbb{R}^n$，并且 $\alpha_1, \cdots, \alpha_k \in [0,1]$，满足 $\alpha_1 + \cdots + \alpha_k = 1$，则 $\alpha_1 x_1 + \cdots + \alpha_k x_k$ 称为点 $x_1, \cdots, x_k$ 的凸组合（convex combination）。 ♣

定理 10.1

集合 $A \subseteq \mathbb{R}^n$ 是凸集当且仅当 A 中所有凸组合均在 A 内。 ♡

证明 充分性是显然的；必要性可通过归纳法采用与命题 9.2类似的讨论直接证明。

根据先前的经验，代数结构的“生成”是一个极其重要的概念，例如闭包、生成向量空间（可以视为一种“线性包”）、仿射包。采用类似的思路，在这里我们希望定义“凸包”。对于 $\mathbb{R}^n$ 中的任意凸集族 $\{A_k | k \in K\}$，对于其交集 $\cap_{k\in K} A_k$ 内的任意两点 x 和 y，由于 $(1-\alpha)x + \alpha y \in A_k$ 对所有 $\alpha \in [0,1]$ 和所有 $k \in K$ 成立，因此 $\cap_{k\in K} A_k$ 也是凸集。特别地，对于给定的 $A \subseteq \mathbb{R}^n$，所有包含 A 的凸集的交集是凸的。由此可以引出如下定义。

定义 10.3 (凸包)

对于 $A \subseteq \mathbb{R}^n$，集合

$$\operatorname{conv}(A) = \bigcap \{C \subseteq \mathbb{R}^n | A \subseteq C,\ C \text{ is convex}\} \tag{10.1.1}$$

称为 A 的凸包（convex hull）。 ♣

注 若 $A,B\subseteq\mathbb{R}^n$ 且 $A\subseteq B$，由于 $\mathrm{conv}(A)$ 是包含 A 的最小凸包，因此 $A\subseteq B\subseteq\mathrm{conv}(B)$ 意味着 $\mathrm{conv}(A)\subseteq\mathrm{conv}(B)$；以后我们经常会用到此结论。此外还有 $\mathrm{conv}(\varnothing)=\varnothing$ 和 $\mathrm{conv}(\mathbb{R}^n)=\mathbb{R}^n$。

类似定理 9.2，凸包不出意外地可由凸组合刻画。

定理 10.2

对于 $A\subseteq\mathbb{R}^n$，$\mathrm{conv}(A)$ 是 A 中向量的所有凸组合构成的集合。

证明 记 A 中向量的所有凸组合构成的集合为 B。根据定理 10.1，A 的凸包含有 A 中向量的所有凸组合，因此 $B\subseteq\mathrm{conv}(A)$。现在我们希望证明 B 是包含 A 的一个凸集，那么作为包含 A 的最小凸集，$\mathrm{conv}(A)\subseteq B$。设 $A=\{x_i|i\in I\}$，则对于任意 $y_i\in B$，存在 $\beta_{i,j}\in[0,1]$ 满足 $\sum_j\beta_{i,j}=1$，同时使得 $y_i=\sum_j s_{i,j}x_j$，因此凸组合

$$\sum_{i=1}^{n}\alpha_i y_i=\sum_{i,j}\alpha_i\beta_{i,j}x_j\in B,\quad \text{where } \alpha_i\beta_{i,j}\in[0,1] \text{ and } \sum_{i,j}\alpha_i\beta_{i,j}=\sum_i\alpha_i\sum_j\beta_{i,j}=\sum_i\alpha_i=1, \tag{10.1.2}$$

因此 B 也是凸集，从而 $\mathrm{conv}(A)\subseteq B$，结论得证。

注 该定理的一个平凡结论是，A 为凸集当且仅当 $\mathrm{conv}(A)=A$，这实际上就是定理 10.1的内容。

定义 10.4 (Minkowski 组合)

对于集合 $A,B\subseteq\mathbb{R}^n$ 和 $\alpha,\beta\in\mathbb{R}$，定义

$$\alpha A+\beta B=\{\alpha x+\beta y|x\in A,y\in B\}. \tag{10.1.3}$$

集合 $\alpha A+\beta B$ 称为集合 A,B 的 Minkowski 组合（Minkowski combination），运算 $+$ 称为 Minkowski 加法（Minkowski addition）。A 与 B 的差集（difference）$A-B$ 定义为 $A+(-B)$。

命题 10.1

若 $A,B\subseteq\mathbb{R}^n$ 是凸集且 $\alpha,\beta\in\mathbb{R}$，则 $\alpha A+\beta B$ 是凸集。

证明 任取 $\alpha A+\beta B$ 中的元素 $\alpha x_1+\beta y_1$ 和 $\alpha x_2+\beta y_2$，其中 $x_1,x_2\in A$ 且 $y_1,y_2\in B$。那么对于任意 $\lambda\in[0,1]$，$\lambda(\alpha x_1+\beta y_1)+(1-\lambda)(\alpha x_2+\beta y_2)=\alpha[\lambda x_1+(1-\lambda)x_2]+\beta[\lambda y_1+(1-\lambda)y_2]\in\alpha A+\beta B$。可见 $\alpha A+\beta B$ 为凸集。

命题 10.2

设集合 $A\subseteq\mathbb{R}^n$。

1. $(\alpha+\beta)A\subseteq\alpha A+\beta A$ 对所有 $\alpha,\beta\in\mathbb{R}$ 成立；
2. 若 A 为凸集，则 $\alpha A+\beta A=(\alpha+\beta)A$ 对所有 $\alpha,\beta\geq 0$ 成立；
3. 若 $(1-\lambda)A+\lambda A=A$ 对所有 $\lambda\in[0,1]$ 成立，则 A 为凸集。

证明

1. 对于任意 $a\in A$，$(\alpha+\beta)a=\alpha a+\beta a\in\alpha A+\beta A$，即 $(\alpha+\beta)A\subseteq\alpha A+\beta A$；
2. 这里仅需证明反向包含成立。当 A 为凸集时，对于任意 $\alpha x+\beta y\in\alpha A+\beta A$，容易看出 $(\alpha x+\beta y)/(\alpha+\beta)\in A$，故 $\alpha A+\beta A\subseteq(\alpha+\beta)A$；
3. 对于任意 $x,y\in A$，$(1-\lambda)x+\lambda y\in(1-\lambda)A+\lambda A=A$，因此 A 为凸集。

虽然上一章中我们未对仿射变换作过多讨论，但还是有必要介绍，因为其具有保凸的良好性质。

定理 10.3

设 $A\subseteq\mathbb{R}^n$ 和 $B\subseteq\mathbb{R}^n$ 均为凸集，$f:\mathbb{R}^n\to\mathbb{R}^m$ 为仿射变换，则

$$f(A)=\{f(x)|x\in A\}\quad\text{and}\quad f^{-1}(B)=\{x\in\mathbb{R}^n|f(x)\in B\} \tag{10.1.4}$$

均为凸集。

证明 对于 $f(A)$ 中的任意两个元素，存在 $x,y\in A$ 使其分别等于 $f(x)$ 和 $f(y)$。由于 A 为凸集，$(1-\lambda)x+\lambda y\in A$

对所有 $\lambda\in[0,1]$ 成立，从而 $(1-\lambda)f(x)+\lambda f(y)=f((1-\lambda)x+\lambda y)\in f(A)$，故 $f(A)$ 为凸集；$f^{-1}(B)$ 类似。

定义 10.5

1. 有限多个闭半空间的交集称为多面体集（polyhedral set）；
2. 有限多个点 $x_1,\cdots,x_k\in\mathbb{R}^n$ 的凸包称为（凸）多胞形（polytype）；
3. 仿射无关点集的凸包称为单纯形（simplex）。r-单纯形为 $(r+1)$ 个仿射无关点的凸包。

♣

注 空集和全空间均为多面体集。根据 Heine-Borel 定理 1.10，多胞形总是紧致集；多面体集是闭集，但可能无界。

定义 10.6 (顶点)

若多胞形 P 内的点 x 使得 $P\backslash\{x\}$ 仍为凸集，则称其为顶点（vertex）。P 的顶点构成的集合记为 $\mathrm{vert}(P)$。

♣

定理 10.4

设 P 为 $\mathbb{R}^n$ 中的多胞形，$x_1,\cdots,x_k\in\mathbb{R}^n$ 为不同的点。

1. 若 $P=\mathrm{conv}\{x_1,\cdots,x_k\}$，则 x_1 是 P 的顶点当且仅当 $x_1\notin\mathrm{conv}\{x_2,\cdots,x_k\}$；
2. P 为其顶点的凸包。

♡

证明

1. 若 x_1 为 P 的顶点，则 $P\backslash\{x_1\}$ 是凸集且 $x_1\notin P\backslash\{x_1\}$。根据凸包的定义，$\{x_2,\cdots,x_k\}\subseteq P\backslash\{x_1\}$ 意味着 $\mathrm{conv}\{x_2,\cdots,x_k\}\subseteq P\backslash\{x_1\}$，因此 $x_1=P\backslash(P\backslash\{x_1\})\in P\backslash\mathrm{conv}\{x_2,\cdots,x_k\}$，即 $x_1\notin\mathrm{conv}\{x_2,\cdots,x_k\}$；反之，若 x_1 不是 P 的顶点，那么存在 $a,b\in P\backslash\{x_1\}$ 和 $\lambda(0,1)$ 使得 $x_1=(1-\lambda)a+\lambda b$（否则 $P\backslash\{x_1\}$ 为凸集，矛盾）。设凸组合 $a=\sum_{i=1}^k\alpha_i,b=\sum_{i=1}^k\beta_ix_i$，那么 $x_1=(1-\lambda)a+\lambda b=\sum_{i=1}^k[(1-\lambda)\alpha_i+\lambda\beta_i]x_i$，其中 $0\leq(1-\lambda)\alpha_i+\lambda\beta_i\leq(1-\lambda)+\lambda=1$ 对所有 i 成立。当 $\alpha_1=\beta_1=1$ 时上式恒等，但此时 $a=b=x_1$，矛盾；因此 $(1-\lambda)\alpha_1+\lambda\beta_1\in[0,1)$，将其改写为 $x_1=[1-(1-\lambda)\alpha_1-\lambda\beta_1]^{-1}\sum_{i=2}^k[(1-\lambda)\alpha_i+\lambda\beta_i]x_i\in\mathrm{conv}\{x_2,\cdots,x_k\}$，结论得证；
2. 根据定义，设存在有点点集 $A=\{x_1,\cdots,x_n\}\in P$ 使得 $P=\mathrm{conv}(A)$。若存在 $x_i\in P$ 使其可表示为其余向量的凸组合，即 $x_i=\sum_{j\neq i}\beta_jx_j$。那么对于任意 $x\in P$，设其凸组合为 $x=\sum_{i=1}^k\alpha_ix_i$，我们有
$$x=\alpha_i\sum_{j\neq i}\beta_jx_j+\sum_{j\neq i}\alpha_jx_j=\sum_{j\neq i}(\alpha_j+\alpha_i\beta_j)x_j\in\mathrm{conv}(A\backslash\{x_i\}),\tag{10.1.5}$$
从而 $P\subseteq\mathrm{conv}(A\backslash\{x_i\})\subseteq\mathrm{conv}(A)=P$，即 P 为 $A\backslash\{x_i\}$ 的凸包。如此操作至多有限次，最终剩余点均为 P 的顶点，并且 $P=\mathrm{conv}(\mathrm{vert}(P))$。

推论 10.1

设 $M\subseteq\mathbb{R}^n$ 是一个集合，$x\in\mathbb{R}^n\backslash M$。记 $A=\mathrm{conv}(\{x\}\cup M)$，则 $A\backslash\{x\}$ 是凸的当且仅当 $x\notin\mathrm{conv}(M)$。

♡

多胞形的性质在集合的基本运算下保持，即有如下结果。

命题 10.3

若 P,Q 为 $\mathbb{R}^n$ 中的多胞形，则以下集合也是多胞形：

1. $\mathrm{conv}(P\cup Q)$ 和 $P\cap Q$；
2. $\alpha P+\beta Q$ 对任意 $\alpha,\beta\in\mathbb{R}$；
3. $f(P)$ 对任意仿射变换 $f:\mathbb{R}^n\to\mathbb{R}^m$。

♠

证明

1. 由于 $\mathrm{vert}(P)\cup\mathrm{vert}(Q)\subseteq P\cup Q$，$\mathrm{conv}(\mathrm{vert}(P)\cup\mathrm{vert}(Q))\subseteq\mathrm{conv}(P\cup Q)$；另一方面，对于任意 $x\in P\cup Q$，根据定理 10.4，x 可表示为 $\mathrm{vert}(P)$ 或 $\mathrm{vert}(Q)$ 的凸组合，进而可表示为 $\mathrm{vert}(P)\cup\mathrm{vert}(Q)$ 的凸组合（多余的向量补足系数 0），因此 $P\cup Q\subseteq\mathrm{conv}(\mathrm{vert}(P)\cup\mathrm{vert}(Q))$，进而 $\mathrm{conv}(P\cup Q)\subseteq\mathrm{conv}(\mathrm{vert}(P)\cup\mathrm{vert}(Q))$。综

上，$\mathrm{conv}(P\cup Q)=\mathrm{conv}(\mathrm{vert}(P)\cup\mathrm{vert}(Q))$。多胞形的定义保证了 $|\mathrm{vert}(P)\cup\mathrm{vert}(Q)|<\infty$，故 $\mathrm{conv}(P\cup Q)$ 也是多胞形。至于 $P\cap Q$，在第10.4节证明了多胞形与多面体集的等价性后该结论自然成立；

2. 由于 $\alpha\mathrm{vert}(P)+\beta\mathrm{vert}(Q)\subseteq\alpha P+\beta Q$，因此根据命题 10.1，$\mathrm{conv}(\alpha\mathrm{vert}(P)+\beta\mathrm{vert}(Q))\subseteq\alpha P+\beta Q$；另一方面，对于任意 $\alpha x+\beta y\in\alpha P+\beta Q$，设 $\mathrm{vert}(P)=\{x_1,\cdots,x_k\},\mathrm{vert}(Q)=\{y_1,\cdots,y_l\}$，则存在凸组合 $x=\sum_{i=1}^k\alpha_ix_i$ 和 $y=\sum_{j=1}^l\beta_jy_j$，使得
$$\begin{aligned}\alpha x+\beta y&=\alpha\left(\sum_{j=1}^l\beta_j\right)\left(\sum_{i=1}^k\alpha_ix_i\right)+\beta\left(\sum_{i=1}^k\alpha_i\right)\left(\sum_{j=1}^l\beta_jy_j\right)\\&=\sum_{i,j}\alpha_i\beta_j(\alpha x_i+\beta y_j)\in\mathrm{conv}(\alpha\mathrm{vert}(P)+\beta\mathrm{vert}(Q)),\end{aligned}\tag{10.1.6}$$
即 $\alpha P+\beta Q\subseteq\mathrm{conv}(\alpha\mathrm{vert}(P)+\beta\mathrm{vert}(Q))$，反向包含成立。此时 $|\alpha\mathrm{vert}(P)+\beta\mathrm{vert}(Q)|\leq|\mathrm{vert}(P)||\mathrm{vert}(Q)|<\infty$，即 $\alpha\mathrm{vert}(P)+\beta\mathrm{vert}(Q)$ 为有限集，故 $\alpha P+\beta Q$ 为多胞形；
3. 由于 $f(\mathrm{vert}(P))\subseteq f(P)$，根据定理 10.3，$\mathrm{conv}(f(\mathrm{vert}(P)))\subseteq f(P)$；另一方面，设 $\mathrm{vert}(P)=\{x_1,\cdots,x_k\}$，则对于任意 $x\in P$，存在凸组合 $x=\sum_{i=1}^k\alpha_ix_i$，使得 $f(x)=\sum_{i=1}^k\alpha_if(x_i)\in\mathrm{conv}(f(\mathrm{vert}(P)))$，即 $f(P)\subseteq\mathrm{conv}(f(\mathrm{vert}(P)))$，反向包含成立。此时 $f(\mathrm{vert}(P))$ 为有限集，故 $f(P)$ 为多胞形。

命题 10.4

若 P 是仿射无关点 $x_0,\cdots,x_r$ 的凸包，则 P 是 r-单纯形，并且每个点 x_i 都是 P 的顶点。r-单纯形具有维数 $\dim(P)=r$，这里维数定义为将该集合嵌入其仿射包后该仿射包的维数。

♠

证明 根据定义，P 是 r-单纯形，因此必然为多胞形。记 $X=\{x_0,\cdots,x_r\}$，根据定理 9.4，X 仿射无关意味着 $x_i\notin\mathrm{aff}(X\backslash\{x_i\})$，显然有 $x_i\notin\mathrm{conv}(X\backslash\{x_i\})$；因此利用定理 10.4，所有 x_i 均为 P 的顶点，且 $\dim(P)=r$。

单纯形可通过凸组合的唯一性刻画。

定理 10.5

凸集 $A\subseteq\mathbb{R}^n$ 是单纯形当且仅当存在 $x_0,\cdots,x_k\in A$，使得每个 $x\in A$ 具有作为 $x_0,\cdots,x_k$ 的凸组合的唯一表示。

♡

证明 若 A 为单纯形，那么存在仿射无关集 $X=\{x_0,\cdots,x_k\}\subseteq A$ 使得 $A=\mathrm{conv}(X)$。根据定理 9.4，A 内任意点的关于 X 的仿射组合唯一，进而关于 X 的凸组合唯一；反之当该表示存在时，首先利用 $X\subseteq A$ 可知 $A\subseteq\mathrm{conv}(X)\subseteq A$，即 $A=\mathrm{conv}(X)$；现在利用该表示的唯一性，假设存在 $x_i\in X$ 使得 $x_i\in\mathrm{aff}(X\backslash\{x_i\})$，即存在仿射组合 $x_i=\sum_{j\neq i}r_jx_j$。分别记指标集 $J_+=\{j\neq i|r_j>0\}$ 和 $J_-=\{j\neq i|r_j<0\}$，并记 $\lambda^{-1}=\sum_{j\in J^+}r_j>0$，那么 $1-\sum_{j\in J^-}r_j=1-(1-\sum_{j\in J^+}r_j)=\lambda^{-1}$，同时唯一性要求以下的凸组合
$$\lambda x_i+\lambda\sum_{j\in J_-}(-r_j)x_j=\lambda\sum_{j\in J^+}r_jx_j\tag{10.1.7}$$
中 x_i 同系数，但 $\lambda>0$ 且 $i\notin J_\pm$，矛盾。因此 $x_i\notin\mathrm{aff}(X\backslash\{x_i\})$。根据定理 9.4，$X$ 仿射无关，即 A 是单纯形。

10.2 组合性质

凸几何中这一部分称为组合几何（*combinatorial geometry*），我们仅介绍其中最重要的三个定理。

定理 10.6 (Radon 定理)

设 $x_1,\cdots,x_m\in\mathbb{R}^n$ 为仿射相关的点，则存在划分 $\{1,\cdots,m\}=I\cup J$，其中 $I\cap J=\varnothing$，使得
$$\mathrm{conv}\{x_i|i\in I\}\cap\mathrm{conv}\{x_j|j\in J\}\neq\varnothing.\tag{10.2.1}$$

♡

证明 设 $x_1,\cdots,x_m\in\mathbb{R}^n$ 仿射相关，那么根据定理 9.4，存在不全为零的 $\alpha_1,\cdots,\alpha_m\in\mathbb{R}$ 使得 $\sum_{i=1}^m\alpha_ix_i=0$ 且 $\sum_{i=1}^m\alpha_i=0$。定义指标集 $I=\{i|\alpha_i\geq0\}$ 和 $J=\{1,\cdots,m\}\backslash I$，显然 $I\cap J=\varnothing$。记 $\lambda^{-1}=\sum_{j\in J}(-\alpha_j)>0$，

那么 $\sum_{i\in I}\alpha_i=-\sum_{j\in J}\alpha_j=\lambda^{-1}$，进而存在凸组合

$$\lambda\sum_{i\in I}\alpha_i=\lambda\sum_{j\in J}(-r_j)\alpha_j \quad\Longrightarrow\quad \operatorname{conv}\{x_i|i\in I\}\cap\operatorname{conv}\{x_j|j\in J\}\neq\varnothing. \tag{10.2.2}$$

Helly 定理研究的是这样一类问题：若集族 $\{A_k|k\in K\}$ 中任意有限多集合都共享某种性质，那么是否所有集合都共享这种性质？事实上，我们在第1章证明的闭区间套定理 1.11就是一例。下面我们仅针对特定性质讨论。

定理 10.7 (Helly 定理)

1. 设 $A_1,\cdots,A_m$ 为 $\mathbb{R}^n$ 中的凸集，其中 $m\geq n+1$。若集合 $A_1,\cdots,A_m$ 中任意 $(n+1)$ 个均具有非空交集，则 $\cap_{i=1}^m A_i\neq\varnothing$；
2. 设 $\mathcal{A}=\{A_k|k\in K\}$ 为 $\mathbb{R}^n$ 中至少 $(n+1)$ 个紧凸集构成的集族。若 $\mathcal{A}$ 中任意 $(n+1)$ 个集合均具有非空交集，则 $\cap_{k\in K}A_k\neq\varnothing$。

♡

证明

1. 我们对 $m-n$ 进行归纳。当 $m-n=1$ 时，结论显然成立。当 $m-n\geq 2$ 时，假设结论对于 $A_1,\cdots,A_m$ 中任意 $m-1$ 个集合成立，那么对于任意 $i=1,\cdots,m$，存在 $x_i\in\cap_{j\neq i}A_j$，并记 $X=\{x_1,\cdots,x_m\}$。若 X 仿射无关，那么根据定理 9.4，$\dim(\operatorname{aff}(X))=m-1\geq n+1>\dim(\mathbb{R}^n)$，矛盾。因此 X 仿射无关。根据 Radon 定理 10.6，存在指标集 I,J 满足 $I\cup J=\{1,\cdots,m\}$ 且 $I\cap J=\varnothing$，并使得交集 $\operatorname{conv}\{x_i|i\in I\}\cap\operatorname{conv}\{x_j|j\in J\}$ 非空。取该交集内的元素 x。注意到 $I\cap J=\varnothing$，则对于任意 $i\in I$ 和 $j\in J$，$x_i\in\cap_{k\neq i}A_k=(\cap_{k\neq i,j}A_k)\cap A_j\subseteq A_j$，进而 $x\in\operatorname{conv}\{x_i|i\in I\}\subseteq A_j$ 对所有 $j\in J$ 成立，因此 $x\in\cap_{j\in J}A_j$。类似地，$x\in\cap_{i\in I}A_i$，最后得到 $x\in(\cap_{i\in I}A_i)\cap(\cap_{j\in J}A_j)=\cap_{i=1}^m A_i$，即 $\cap_{i=1}^m A_i\neq\varnothing$；
2. 根据上一条结论，集族 $\mathcal{A}$ 内任意有限多集合的交集非空。根据 Heine-Borel 定理 1.10，A_k 是有界闭集。假设 $\cap_{k\in K}A_k=\varnothing$，那么 $\cup_{k\in K}A_k^c=\mathbb{R}^n$ 构成 $\mathbb{R}^n$ 的一个开覆盖。由于任意集合 $A\in\mathcal{A}$ 紧致，存在有限指标集 K_0 使得 $A\subseteq\cup_{k\in K_0}A_k^c$ 对所有 $k\in K$ 成立，即 $\cap_{k\in K_0}A_k\subseteq A^c$。但此时有限交集 $(\cap_{k\in K_0}A_k)\cap A=\varnothing$，矛盾。因此 $\cap_{k\in K}A_k\neq\varnothing$，结论得证。

注 上述第二条结论中紧致性是必要的。考虑 $\mathbb{R}$ 上的开线段 $A_n=(-1/n,1-1/n)$。那么对于任意两条不同的开线段 A_n 和 A_m，不妨设 $n<m$，那么 $A_n\cap A_m=(-1/m,1-1/n)\neq\varnothing$；但是 $\cap_{k=1}^n A_k=(-1/n,0)$。对于任意正数 ϵ，由 Archimedean 性质1.19，存在自然数 n_0 使得 $-\epsilon<-1/n_0$，即 $-\epsilon\notin\cap_{n=n_0}^\infty A_n$，进而 $\cap_{n=1}^\infty A_n=\varnothing$。

定理 10.8 (Carathéodory 定理)

对于集合 $A\subseteq\mathbb{R}^n$ 和 $x\in\mathbb{R}^n$，以下两种表述等价：

1. $x\in\operatorname{conv}(A)$；
2. 对于某个 $r\in\{0,\cdots,n\}$，存在一个顶点在 A 上的 r-单纯形 P，使得 $x\in P$。

♡

证明 对于必要性，$\operatorname{vert}(P)\subseteq A$ 意味着 $x\in P=\operatorname{conv}(\operatorname{vert}(P))\subseteq\operatorname{conv}(A)$；对于充分性，存在 $x_1,\cdots,x_k\in A$ 使得 x 可表示为凸组合 $\sum_{i=1}^k\alpha_i x_i$。对于满足上述条件的全体自然数构成的集合，根据定理 1.7，其中必然存在一个最小自然数 k，此时任意 $(k-1)$ 个点的凸包不包含 x，希望证明这 k 个点仿射无关。

假设其仿射相关，即存在不全为零的 $\beta_1,\cdots,\beta_k\in\mathbb{R}$ 使得 $\sum_{i=1}^k\beta_k x_i=0$ 且 $\sum_{i=1}^k\beta_i=0$。那么存在非空指标集 $J=\{i|\beta_i>0\}$，从中选择指标 j_0 使得 $\alpha_{j_0}/\beta_{j_0}=\min_{j\in J}(\alpha_j/\beta_j)$，进而可构造

$$x=x-\frac{\alpha_{j_0}}{\beta_{j_0}}\sum_{i=1}^k\beta_i x_i=\sum_{i\neq j_0}\left(\alpha_i-\frac{\alpha_{j_0}}{\beta_{j_0}}\beta_i\right)x_i, \tag{10.2.3}$$

其中 $\alpha_i-\alpha_{j_0}\beta_i/\beta_{j_0}\geq 0$ 对任意 i 成立，并且 $\sum_{i=1}^k(\alpha_i-\alpha_{j_0}\beta_i/\beta_{j_0})=1-0=1$，因此 x 可由 $(k-1)$ 个非零向表示，这与 k 的极小性矛盾。因此 $x_1,\cdots,x_k$ 仿射无关，此时 $P=\operatorname{conv}\{x_1,\cdots,x_k\}$ 是包含 x 的一个单纯形。

注 上述定理表明，若 x 是仿射相关点 $x_1,\cdots,x_k$ 的凸组合，则至少有一个点是冗余的。重复此操作有限次，我们可得到 x 作为 $\{x_1,\cdots,x_k\}$ 的仿射无关子集的凸组合表示。特别地，$A\subseteq\mathbb{R}^n$ 的凸包内任一点是 A 中至多 $(n+1)$ 个点的凸组合。显然，这些点的选择依赖于 x。

10.3 拓扑性质

尽管凸性是纯代数性质，但该性质仍然带来了一些特别的拓扑结果。

命题 10.5

若对于整数 $k\in\{1,\cdots,n\}$，$P=\mathrm{conv}\{x_0,\cdots,x_k\}$ 是 $\mathbb{R}^n$ 中的一个 k-单纯形，则其相对内部[a]

$$\mathrm{relint}(P)=\{\alpha_0x_0+\cdots+\alpha_kx_k\in\mathbb{R}^n|\alpha_i\in(0,1),\ \alpha_0+\cdots+\alpha_k=1\}. \tag{10.3.1}$$

特别地，$(x_0+\cdots+x_k)/(k+1)\in\mathrm{relint}(P)$。

[a]集合 A 的相对内部 $\mathrm{relint}(A)$ 定义为将 $\mathrm{aff}(A)$ 的基底子空间视为整个向量空间时 A 的内部



证明 记 P 的基底子空间为 $S=\mathrm{aff}(P)-x_0$，$s_i=x_i-x_0$ 对所有 $1\le i\le k$ 成立。根据定理 9.4，$\{x_0,\cdots,x_k\}$ 的仿射无关性要求 $\mathcal{B}=\{s_1,\cdots,s_k\}$ 线性无关，那么 $\mathcal{B}$ 构成 S 的一组基，此时等价于证明

$$\mathrm{relint}(P)=\{\alpha_1s_1+\cdots+\alpha_ks_k\in S|\alpha_i\in(0,1),\ \alpha_1+\cdots+\alpha_k\in(0,1)\}. \tag{10.3.2}$$

考虑(2.3.2)式定义的坐标映射 $\phi_{\mathcal{B}}$，其为向量空间之间的同构。对于任意 $x=\sum_{i=1}^k\alpha_is_i,y=\sum_{i=1}^k\beta_is_i\in S$ 以及正数 ϵ，当 $\sum_{i=1}^k|\alpha_i-\beta_i|^2<\epsilon^2/(\sum_{i=1}^k\|s_i\|^2)$ 时，由 Cauchy-Schwarz 不等式，

$$\|x-y\|=\left\|\sum_{i=1}^k(\alpha_i-\beta_i)s_i\right\|\le\sum_{i=1}^k|\alpha_i-\beta_i|\|s_i\|\le\left(\sum_{i=1}^k|\alpha_i-\beta_i|^2\right)^{1/2}\left(\sum_{i=1}^k\|s_i\|^2\right)^{1/2}<\epsilon. \tag{10.3.3}$$

根据命题 1.16，$\phi_{\mathcal{B}}^{-1}:\mathbb{R}^k\to S$ 是连续的，从而 $\phi_{\mathcal{B}}$ 是同胚（参见定义 A.9）。对于任意 $x\in\phi_{\mathcal{B}}(\mathrm{relint}(P))$，存在 S 内的开球 $B(\phi_{\mathcal{B}}^{-1}(x),r)\subseteq P-x_0$，根据命题 1.17，存在 x 的邻域 $U\subseteq\phi_{\mathcal{B}}(B(\phi^{-1}(x),r))\subseteq\phi_{\mathcal{B}}(P-x_0)$，即 $x\in\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))$，从而 $\phi_{\mathcal{B}}(\mathrm{relint}(P))\subseteq\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))$；反之对于任意 $x\in\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))$，存在开球 $B'(x,r)\subseteq\phi_{\mathcal{B}}(P-x_0)$，根据命题 1.17，存在包含 $\phi_{\mathcal{B}}^{-1}(x)$ 的邻域 $V\subseteq\phi_{\mathcal{B}}^{-1}(B'(\phi_{\mathcal{B}}^{-1},r))\subseteq P-x_0$，即 $\phi_{\mathcal{B}}^{-1}(x)\in\mathrm{relint}(P)$，从而 $\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))\subseteq\phi_{\mathcal{B}}(\mathrm{relint}(P))$。因此

$$\phi_{\mathcal{B}}(\mathrm{relint}(P))=\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))=\mathrm{relint}\big\{(\alpha_1,\cdots,\alpha_k)\in\mathbb{R}^k|\alpha_i\in[0,1],\ \alpha_1+\cdots+\alpha_k\in[0,1]\big\}. \tag{10.3.4}$$

记集合 $U=\big\{(\alpha_1,\cdots,\alpha_k)\in\mathbb{R}^k|\alpha_i\in(0,1),\ \alpha_1+\cdots+\alpha_k\in(0,1)\big\}$。对于任意 $x=(\alpha_1,\cdots,\alpha_k)\in U$，取正数 ϵ 为 $\min_i\alpha_i$ 和 $(1-\sum_i\alpha_i)/k$ 中的较小值。那么对于任意 $y=(\beta_1,\cdots,\beta_k)\in B(x,\epsilon)$，首先有 $|\beta_i-\alpha_i|\le(\sum_i|\beta_i-\alpha_i|^2)^{1/2}=\|x-y\|<\epsilon$ 对所有 i 成立，因此 $\beta_i>\alpha_i-\epsilon\ge\alpha_i-\min_i\alpha_i\ge0$，同时 $\sum_i\beta_i<\sum_i(\alpha_i+\epsilon)\le\sum_i\alpha_i+(1-\sum_i\alpha_i)=1$，因此 $B(x,\epsilon)\subseteq\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))$；对于反向包含，设 $x=(\alpha_1,\cdots,\alpha_k)\in\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))$，即存在开球 $B(x,r)\subseteq\phi_{\mathcal{B}}(P-x_0)$。我们希望证明 $\alpha_i>0$ 对所有 i 成立，并且 $\sum_i\alpha_i<1$。假设存在 $\alpha_i=0$，那么 $y=(\alpha_1,\cdots,\alpha_{i-1},-r/2,\alpha_{i+1},\cdots,\alpha_k)\in B(x,r)$，但显然 $y\notin\phi_{\mathcal{B}}(P-x_0)$，矛盾，因此 $\alpha_i>0$ 对所有 i 成立；假设 $\sum_i\alpha_i=1$，那么 $y=(\alpha_1+r/2\sqrt{k},\cdots,\alpha_k+r/2\sqrt{k})\in B(x,r)$，但显然 $\sum_i(\alpha_i+r/2\sqrt{k})=1+r\sqrt{k}/2>1$，矛盾，因此 $\sum_i\alpha_i<1$。因此 $\alpha_i\in(0,1)$ 对所有 i 成立，并且 $\sum_i\alpha_i\in(0,1)$，即 $\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))\subseteq U$。结合(10.3.2)式，$\phi_{\mathcal{B}}(\mathrm{relint}(P))=\mathrm{relint}(\phi_{\mathcal{B}}(P-x_0))=U$，即

$$\mathrm{relint}(P)=\phi^{-1}(U)=\left\{\sum_i\alpha_is_i\in S|\alpha_i\in(0,1),\ \sum_i\alpha_i\in(0,1)\right\}. \tag{10.3.5}$$

定理 10.9

若非空子集 $A\subseteq\mathbb{R}^n$ 是凸的，则 $\mathrm{relint}(A)\neq\varnothing$。



证明 根据 Carathéodory 定理 10.8，对于任意 $x\in A$，存在顶点在 A 上的单纯形 P 使得 $x\in P$，其中 $P=\mathrm{conv}(\mathrm{vert}(P))\subseteq\mathrm{conv}(A)=A$。再根据命题 10.5，$\mathrm{relint}(P)$ 非空，考虑 $y\in\mathrm{relint}(P)$，那么存在包含 y 的邻域 $U\subseteq P\subseteq A$，故 $y\in\mathrm{relint}(A)$，即 $\mathrm{relint}(A)\neq\varnothing$。

上述定理说明，在研究给定凸集 A 时，将其仿射包作为基底空间是有用的：尽管 A 在全空间内未必具有内点（例如 $\mathbb{R}^3$ 中 xOy 平面上的圆），但其在仿射包中必然具有内点。因此，后文我们会不失一般性地假定 A（在 $\mathbb{R}^n$ 的某个子空间 $\mathbb{R}^k$ 中）内部非空。

命题 10.6

若 $A \subseteq \mathbb{R}^n$ 是凸的，并且 $x \in \mathrm{cl}(A), y \in \mathrm{relint}(A)$，则 $[y, x) \in \mathrm{relint}(A)$。

♠

证明 设 $z = (1-\lambda)x + \lambda y$，其中 $\lambda \in (0,1]$。由于 $x \in \mathrm{cl}(A)$，对于每个自然数 k，均存在 $x_k \in B(x, 1/k) \cap A$，那么 $(x_k) \to x$。由于 $y \in \mathrm{relint}(A)$，存在 $\mathrm{aff}(A)$ 上的开球 $B(y,r) \subseteq A$。定义 $\mathbb{R}^n$ 内的序列 (y_k) 中每个元素为 $y_k = [z-(1-\lambda)x_k]/\lambda$，则 $(y_k) \to y$。因此存在 $k_0 \in \mathbb{N}$ 使得 $y_k \in B(y,r) \subseteq A$ 对所有 $k \geq k_0$ 成立。根据命题 10.2，$z = (1-\lambda)x_k + \lambda y_k \in (1-\lambda)x_k + \lambda B(x, 1/k) \subseteq (1-\lambda)A + \lambda A = A$，因此 $z \in \mathrm{relint}(A)$，即 $[y, x) \subseteq \mathrm{relint}(A)$。

定理 10.10

若 $A \subseteq \mathbb{R}^n$ 是（相对）开的，则 $\mathrm{conv}(A)$ 是（相对）开的；若 $A \subseteq \mathbb{R}^n$ 是紧致的，则 $\mathrm{conv}(A)$ 是紧致的。

♡

证明

1. 对于任意 $x \in \mathrm{conv}(A)$，根据 Carathéodory 定理 10.8，存在仿射无关的点集 $\{x_0, \cdots, x_r\} \subseteq A$ 使得 x 可表示为凸组合 $\sum_{i=0}^{r} \alpha_i x_i$，其中 $r \leq n$。由于 A 在 $\mathrm{aff}(A)$ 上是开集，故对于每个点 x_i，存在开球 $B(x_i, r_i) \subseteq A$。取 $r = \min_i r_i$，那么 $x_i + B(0,r) \subseteq A$ 对所有 i 成立。对于任意 $y \in B(0,r)$，利用 $\mathrm{conv}(A)$ 的凸性可知
$$x + y = \sum_{i=0}^{r} \alpha_i x_i + y = \sum_{i=0}^{r} \alpha_i (x_i + y) \in \mathrm{conv}(A), \tag{10.3.6}$$
因此 $x \in x + B(0,r) \subseteq \mathrm{conv}(A)$，因此 $\mathrm{conv}(A)$ 在 $\mathrm{aff}(A)$ 上是开集；
2. 根据 Heine-Borel 定理 1.10，A 是有界闭集，即存在开球 $B(x,r)$ 使得 $A \subseteq B(x,r)$，根据定义有 $\mathrm{conv}(A) \subseteq B(x,r)$，因此 $\mathrm{conv}(A)$ 有界；设 $\dim(\mathrm{aff}(A)) = m$，考虑 $\mathrm{conv}(A)$ 内一组收敛到 x 的序列 (x_k)，根据 Carathéodory 定理 10.8，对于每个 $k \in \mathbb{N}$，存在 A 中一组点 $x_{k,0}, \cdots, x_{k,m}$ 使得
$$x_k = \alpha_{k,0} x_{k,0} + \cdots + \alpha_{k,n} x_{k,m}, \quad \text{where } \alpha_{k,i} \in [0,1] \text{ and } \sum_{i=0}^{n} \alpha_{k,i} = 1. \tag{10.3.7}$$
由于 $A \subseteq \mathbb{R}^n$ 和 $[0,1]$ 均紧致，根据 Bolzano-Weierstrass 定理 1.13，存在 $\{\alpha_{k,0}\}_{k\in\mathbb{N}}$ 的一个子列收敛；在该子列中，继续寻找 $\{x_{k,0}\}$ 的收敛子列；重复这样的操作 $(2k+2)$ 次，最后得到 $\mathbb{N}$ 中的子列 $(k_r)_{r\in\mathbb{N}}$ 使得所有 $\{\alpha_{k_r,j}\}_{r\in\mathbb{N}}$ 和 $\{x_{k_r,j}\}_{r\in\mathbb{N}}$ 分别收敛到 $\beta_j \in [0,1]$ 和 $y_j \in A$，那么子列 $\{\sum_{i=0}^{m} \alpha_{k_r,j} x_{k_r,j}\}_{r\in\mathbb{N}}$ 收敛到 $\sum_{j=0}^{m} \beta_j y_j$。利用 $\mathbb{R}$ 上极限的线性性和单调性（定理 1.15），$\beta_j = \lim_r \alpha_{k_r,j} \in [0,1]$ 对所有 j 成立，$\sum_{j=0}^{m} \beta_j = \lim_r \sum_{j=0}^{m} \alpha_{k_r,j} = 1$；再利用极限的唯一性（命题 1.10），$x = \sum_{j=0}^{m} \beta_j y_j \in \mathrm{conv}(A)$，因此 $\mathrm{conv}(A)$ 也是有界闭集。根据 Heine-Borel 定理 1.10。

注 特别地，上述定理表明，凸多胞形 P 是紧致的。

10.4 支撑与分离

凸集是包含其内所有元素及其凸组合的集合。在本节中，我们考虑凸集另一种具有对偶性质的描述：凸集可从半空间的交集获得。为了得到这样的结果，所研究的凸集必须是闭集。

定理 10.11

设 $A \subseteq \mathbb{R}^n$ 为非空闭凸集，那么对于每个 $x \in \mathbb{R}^n$，存在唯一的点 $p(A,x) \in A$ 满足
$$\|p(A,x) - x\| = \inf_{y\in A} \|y - x\|. \tag{10.4.1}$$

♡

证明 该定理的证明思路与定理 8.1类似，区别在于定理 8.1中研究的集合同时还是子空间，因此天然是闭集。显然对于 $x \in A$，显然有 $p(A,x) = x$；对于 $x \notin A$，由于 $\|y - x\| \geq 0$ 对任意 $y \in A$ 成立，完备性公理 1.4要求其下确界 $d = \inf_{y\in A} \|y - x\|$ 存在。另外，对于任意 $y \in A$，存在 $r > \|x - y\| > 0$ 使得 $A \cap \bar{B}(x,r) \neq \varnothing$，因此 $d \leq \inf_{y\in A\cap\bar{B}(x,r)} \|y - x\|$。另一方面，对于任意 $\epsilon > 0$，存在 $y_\epsilon \in A \cap \bar{B}(x,\epsilon)$ 使得 $\|y_\epsilon - x\| < d + \epsilon$，因此 $\inf_{y\in A\cap\bar{B}(x,r)} \|y - x\| \leq \|y_\epsilon - x\| < d + \epsilon$。综合上述结果，$d = \inf_{y\in A\cap\bar{B}(x,r)} \|y - x\|$。

根据命题 1.24，$A\cap\bar{B}(x,r)$ 是有界闭集，因此根据 Heine-Borel 定理 1.10，$A\cap\bar{B}(x,r)$ 是紧致的；再利用命题 1.18和8.1，$\mathbb{R}^n$ 上范数的像 $\|A\cap\bar{B}(x,r)\|$ 也是紧致的。根据极值定理 1.5，存在 $y_0\in A\cap\bar{B}(x,r)$ 使得

$$\|y_0-x\|=\inf_{y\in A\cap\bar{B}(x,r)}\|y-x\|=\inf_{y\in A}\|y-x\|. \tag{10.4.2}$$

最后，假设存在 $y_1\neq y_2$ 达到该下确界 d，根据凸性有 $y_0=(y_1+y_2)/2\in A$，那么

$$\langle y_2-y_0,y_0-x\rangle=\frac{1}{4}\langle(y_2-x)-(y_1-x),(y_2-x)+(y_1-x)\rangle=\frac{1}{4}(\|y_2-x\|^2-\|y_1-x\|^2)=0. \tag{10.4.3}$$

由于 $y_1\neq y_2$，我们有 $y_2\neq y_0$，这意味着

$$\|y_2-x\|^2=\|(y_2-y_0)+(y_0-x)\|^2=\|y_2-y_0\|^2+\|y_0-x\|^2\geq\|y_0-x\| \tag{10.4.4}$$

与 y_2 达到下确界矛盾，因此达到下确界的 $y_0\triangleq p(A,x)\in A$ 唯一。

根据该定理，我们可引出度量投影的定义。

定义 10.7 (度量投影)

映射 $p(A,\cdot):\mathbb{R}^n\to A$ 称为到非空闭凸集 $A\subseteq\mathbb{R}^n$ 上的度量投影（metric projection）。

♣

注 事实上，最近邻点 $p(A,x)$ 的存在性关于所有闭集 A 成立，而其唯一性则由 A 的凸性保证。

定义 10.8

设 $A\subseteq\mathbb{R}^n$ 是闭凸集，$E=\{f=\alpha\}$ 为超平面，那么如果 $A\cap E\neq\varnothing$ 且 A 包含在两个闭半空间 $\{f\leq\alpha\}$ 或 $\{f\geq\alpha\}$ 之一，则称 E 为 A 的支撑超平面（supporting hyperplane）。包含 A 且以 A 的支撑超平面为边界的半空间称为 A 的支撑半空间（supporting halfspace）；集合 $A\cap E$ 称为支撑集（supporting set），并且任意 $x\in A\cap E$ 称为支撑点（supporting point）。若 E 是 A 的支撑超平面，我们简称 E 支撑 A。

♣

例 10.2 集合 $A=\{(a,b)\in\mathbb{R}^2|b\geq a^{-1},a>0\}$ 是闭凸集；直线 $g=\{a+b=2\}$ 是支撑直线，但直线 $\{a=0\}$ 和 $\{b=0\}$ 不是。

定理 10.12

设 $A\subseteq\mathbb{R}^n$ 为非空闭凸集，$x\in\mathbb{R}^n\backslash A$，那么通过 $p(A,x)$ 且正交于 $x-p(A,x)$ 的超平面 E 支撑 A。此外，以 E 为边界且不包含 x 的半空间 H 是一个支撑半空间。

♡

证明 若 $x\in E$，那么 $p(A,x)\in E$ 意味着其线性形式 $x-p(A,x)\in E$；另一方面，已知 $(x-p(A,x))\perp E$，那么 $\langle x-p(A,x),x-p(A,x)\rangle=0$ 意味着 $x=p(A,x)\in A\cap E\subseteq A$，矛盾。对于 $x\notin E$，通过内积定义线性泛函 $f:\mathbb{R}^n\to\mathbb{R}$ 为 $f(y)=\langle y,x-p(A,x)\rangle$。记 $\alpha=\langle p(A,x),x-p(A,x)\rangle$，那么通过 $p(A,x)$ 且正交于 $x-p(A,x)$ 的向量集为 $E=\{y\in\mathbb{R}^n|\langle y-p(A,x),x-p(A,x)\rangle=0\}=\{f=\alpha\}$。根据定理 9.1，容易验证 E 是仿射封闭的，因此 E 是平坦。通过 Gram-Schmidt 增广过程（定理 5.6），可在 $\mathbb{R}^n$ 中得到一组与 x 正交的标准正交集；根据定理 5.8，该集合基数为 $n-1$，即 $\dim(E)=n-1$，因此 E 是超平面。

现在仅需证明 $A\subsetneq H$。为了体现该结论的直观性，后续证明我们多通过几何图像进行说明，证明细节请读者自行补全。假设存在点 $y\in A$ 但 $y\notin H$，由于 $\langle x-p(A,x),x-p(A,x)\rangle>0$ 且 $x\notin H$，可证明 $\langle y-p(A,x),x-p(A,x)\rangle>0$；从几何上看，$x$ 与 H 被 E 分开，因此 $y\notin H$ 意味着 x 与 y 位于 E 同侧，进而上述关于 x 的投影点 $p(A,x)$ 的向量内积为正。设 x 在通过 $p(A,x)$ 和 y 上的闭线段投影为 $\bar{y}$，当 $\bar{y}$ 在 $p(A,x)$ 和 y 之间时，$\bar{y}\in A$ 且 $x,p(A,x)$ 与 $\bar{y}$ 构成直角三角形的顶点，其中 $\bar{y}$ 为垂足，因此 $\|x-p(A,x)\|>\|x-\bar{y}\|$，但这与 $p(A,x)$ 的极小性矛盾；当 y 在 $p(A,x)$ 与 $\bar{y}$ 之间时，类似有 $\|x-p(A,x)\|>\|x-y\|$，而这同样与 $p(A,x)$ 的极小性矛盾。综上，只要 $y\in A$ 就有 $y\in H$，即 $A\subseteq H$。

推论 10.2

每个非空闭凸真子集 $A\subsetneq\mathbb{R}^n$ 都是所有包含 A 的闭半空间的交集。更具体地，A 是其所有支撑半空间的交集。

♡

证明 由于 A 包含在其任一支撑半空间内，因此其在支撑半空间的交集 B 内。对于 $x \notin A$，根据定理 10.12，存在 x 的支撑半空间 H 使得 $x \notin H$，因此 $\mathbb{R}^n \backslash A \subseteq \mathbb{R}^n \backslash B$，即 $B \subseteq A$。

为了方便后续讨论，我们首先证明度量投影的连续性。

定理 10.13

设 $A \subseteq \mathbb{R}^n$ 是非空凸集，那么对于任意 $x, y \in \mathbb{R}^n$，$\|p(A,x) - p(A,y)\| \leq \|x - y\|$。 ♡

证明 若 $x \in A$ 或 $y \in A$，那么 $p(A,x) = x$ 或 $p(A,y) = y$，定理 10.12使得该不等式立即成立；若 $x \notin A$ 且 $y \notin A$，对于任意 $z \in \mathbb{R}^n$ 定义线性泛函 $f_x(z) = \langle z - p(A,x), x - p(A,x) \rangle$，因此 $f_x(x) > 0$；由于 $x \notin H_x$，那么对于任意 $z \in H_x$，$f_x(z) \leq 0$。类似定义 f_y，从而 $p(A,x), p(A,y) \in A$ 意味着 $f_x(p(A,y)) \leq 0, f_y(p(A,x)) \leq 0$，因此

$$\|x - y\|^2 = \|(x - p(A,x)) + (p(A,y) - y)\|^2 + \|p(A,x) - p(A,y)\|^2 - 2f_x(p(A,y)) - 2f_y(p(A,x)), \tag{10.4.5}$$

即 $\|x - y\| \geq \|p(A,x) - p(A,y)\|$ 对任意 $x, y \notin A$ 成立。综上，结论得证。

定理 10.14 (支撑定理)

设 $A \subseteq \mathbb{R}^n$ 为闭凸集，那么通过 A 的每个边界点均存在一个支撑超平面。 ♡

证明 对于任意给定的 $x \in \partial A$ 和任意自然数 k，存在 $x_k \in B(x, 1/k)$ 满足 $x_k \notin A$，那么根据定理 10.13，

$$\|x - p(A,x_k)\| = \|p(A,x) - p(A,x_k)\| \leq \|x - x_k\| < 1/k, \tag{10.4.6}$$

即 $x_k, p(A,x_k) \in B(x, 1/k) \subseteq B(x,1)$，因此直观上看，存在唯一边界点 $y_k \in B(x,1)$（证明留给读者）使得 x_k 介于 $p(A,x_k)$ 和 y_k 之间。对于过 $p(A,x_k)$ 的支撑超平面 E_k，由于 y_k 是 $p(A,x_k)$ 与 x_k 的仿射组合，根据定理 10.12，$(x - p(A,x_k)) \perp E_k$ 等价于 $(y_k - p(A_k)) \perp E_k$。对于超平面而言，与其正交的标准向量唯一（否则不构成极大标准正交集），因此 $p(A,y_k) = p(A,x_k)$。根据 Bolzano-Weierstrass 定理 1.13，实数列 (y_k) 存在收敛到某个 y 的子列 $(y_{k_r})_{r \in \mathbb{N}}$；利用度量投影的连续性（定理 10.13），$p(A,y_{k_r}) \to p(A,y)$，并且 $p(A,y_{k_r}) = p(A,x_{k_r}) \to p(A,x) = x$；由极限的唯一性（命题 1.10），$x = p(A,y)$。结合范数的连续性（命题 8.1），$\|y - x\| = \lim_r \|y_{k_r} - x\| = 1$，因此 $x \neq y$。换言之，存在 $y \notin A$ 使得 $x = p(A,y)$，即过 x 存在于 $y - p(A,y)$ 正交的支撑超平面，结论得证。

定义 10.9 (分离与真分离)

对于两个集合 $A, B \subseteq \mathbb{R}^n$ 和超平面 $E = \{f = \alpha\}$，若 $A \subseteq \{f \leq \alpha\}, B \subseteq \{f \geq \alpha\}$ 或 $A \subseteq \{f \leq \alpha\}, B \subseteq \{f \leq \alpha\}$ 两者之一成立，则称 E 分离（seperate）A 和 B。若存在分离超平面不同时包含 A 和 B，则称凸集 A 和 B 可被真分离（properly seperate）。 ♣

定理定理 10.12说明闭凸集 A 和点 $x \notin A$ 可被超平面分离（甚至存在一个与 A 和 x 都要正距离的分离超平面）。这一结果可推广到紧凸集 B 而不仅仅是单点集 $\{x\}$ 上。此外，支撑定理 10.14表明对于 A 的每个边界点，都存在一个超平面使其与 A 分离。分离定理则提供了两个非空凸集真分离的充要条件。

定理 10.15 (分离定理)

设 $A, B \subseteq \mathbb{R}^n$ 为 $\mathbb{R}^n$ 中的非空凸集，则 $\operatorname{relint}(A) \cap \operatorname{relint}(B) = \varnothing$ 当且仅当 A 和 B 可被真分离。 ♡

结合该定理之前的一小段讨论，我们已经能接受该结论在直观上的正确性。有鉴于此，我们在这里仅给出结果，不再对其进行严格证明。如果对详细的推导有兴趣，请参见 [2]。以下推论同样仅给出结果

推论 10.3 (强分离)

设 $A, B \subseteq \mathbb{R}^n$ 是凸集，此外 A 是闭集、B 是紧致集，并且 $A \cap B = \varnothing$，则存在超平面 $\{f = \gamma\}$ 和 $\epsilon > 0$ 使得 $A \subseteq \{f \geq \gamma + \epsilon\}$ 且 $B \subseteq \{f \leq \gamma - \epsilon\}$。此时称 A 和 B 被超平面 $\{f = \gamma\}$ 强分离（strongly seperate）。 ♡

注 注意这里紧致性的要求是必须的，这是为了防止 A 和 B 在渐近情形不可强分离。例如 $A = \{(x,y) \in \mathbb{R}^2 : y \geqslant \mathrm{e}^x\}$ 和 $B = \{(x,y) : y \leqslant 0\}$ 均为非空闭凸集，但只存在分离直线 $\{(x,y) : y = 0\}$，不存在可强分离的直线。

在本节的最后，我们讨论凸多胞形与多面体集的等价性。我们先从支撑集的维数特征入手。

定义 10.10 (边和面)

若闭凸集 $A\subseteq\mathbb{R}^n$ 的支撑集 F 的维数 $\dim(F)=k\in\{0,\cdots,n-1\}$，则称其为 k-支撑集（k-support）。A 的 1-支撑集称为边（edge），$(n-1)$-支撑集称为（facet）；有时也称 A 的 $(\dim(A)-1)$-支撑集为 A 的面。♣

定理 10.16

多胞形 $P\subseteq\mathbb{R}^n$ 的 0-支撑集恰好是单点集 $\{x\}$，其中 $x\in\mathrm{vert}(P)$。♡

证明 对于充分性，存在支撑超平面 $\{f=\beta\}$ 使得 $P\subseteq\{f\leq\beta\}$ 且 $P\cap\{f=\beta\}=\{x\}$。由于开半平面是凸集，故 $P\backslash\{x\}=P\cap\{f<\beta\}$ 也是凸集，因此 $x\in|(P)$。

反之设 $\mathrm{vert}(P)\backslash\{x\}=\{x_1,\cdots,x_k\}$ $(k\geq1)$。那么根据定理 10.4，$x\notin\mathrm{conv}\{x_1,\cdots,x_k\}\overset{\Delta}{=\!=}Q$。假设 $x_1\in\mathrm{conv}\{x_2,\cdots,x_k\}$，那么 x_1 不是 P 的顶点，矛盾。因此 $\{x_1,\cdots,x_k\}\subseteq\mathrm{vert}(Q)$。若存在 $x\in\mathrm{vert}(Q)\backslash\{x_1,\cdots,x_k\}$，则 $x\notin\mathrm{conv}\{x_1,\cdots,x_k\}=Q$，矛盾。综上，$\mathrm{vert}(Q)=\mathrm{conv}\{x_1,\cdots,x_k\}$，进而 Q 也是闭凸集。根据定理 1.14，存在通过 $p(Q,x)$ 且正交于 $x-p(Q,x)$ 的支撑超平面 $\{f=\alpha\}$ 且 $Q\subseteq\{f\leq\alpha\}$，同时记 $\beta=f(x)>\alpha$。设 $y\in P$，那么存在 $\alpha_i\in[0,1]$ 满足 $\sum_{i=1}^{k+1}\alpha_i=1$，使得

$$y=\sum_{i=1}^{k}\alpha_ix_i+\alpha_{k+1}x\quad\Longrightarrow\quad f(y)=\sum_{i=1}^{k}\alpha_if(x_i)+\alpha_{k+1}f(x)\leq\sum_{i=1}^{k}\alpha_i\alpha+\alpha_{k+1}\beta\leq\beta,\tag{10.4.7}$$

等号成立当且仅当 $\alpha_{k+1}=1$，即 $y=x$。因此 $\{f\leq\beta\}$ 是支撑半空间，且 $P\cap\{f=\beta\}=\{x\}$，可见 $\{x\}$ 是 P 的一个 0-支撑集。

注 出于这种考虑，后面我们不再严格区分 0-支撑集和顶点，尽管前者是单点集、后者单纯是点。

定理 10.17

设 $P\subseteq\mathbb{R}^n$ 是具有顶点 $\mathrm{vert}(P)=\{x_1,\cdots,x_k\}$ 的多胞形，F 是 P 的支撑集，则 $F=\mathrm{conv}\{x_i|x_i\in F\}$。♡

证明 设 $F=P\cap\{f=\alpha\}$ 且 $P\subseteq\{f\leq\alpha\}$。假设不存在 $x_i\in F$，即 $f(x_i)<\alpha$ 对所有 i 成立；那么对于任意 $x\in P$，其凸组合 $\sum_{i=1}^{k}\alpha_ix_i$ 满足 $f(x)=\sum_{i=1}^{k}\alpha_if(x_i)<\alpha$，即 $F=P\cap\{f=\alpha\}=\varnothing$，这与支撑定理 10.14矛盾。因此恰当重排指标后，存在某个 $m\in\{1,\cdots,k\}$ 使得 $x_1,\cdots,x_m\in F$，并且 $x_{m+1},\cdots,x_k\notin F$。不妨设 $f(x_i)=\alpha-\delta_i$ $(i=m+1,\cdots,k)$，其中 $\delta_i>0$。对于任意 $x\in P$，其凸组合 $\sum_{i=1}^{k}\alpha_ix_i$ 满足 $f(x)=\sum_{i=1}^{m}\alpha_if(x_i)=\alpha-\sum_{i=m+1}^{k}\alpha_i\delta_i\leq\alpha$，等号成立当且仅当 $\alpha_i=0$ 对所有 $i=m+1,\cdots,k$ 成立。因此 $F=\mathrm{conv}\{x_1,\cdots,x_m\}=\mathrm{conv}\{x_i|x_i\in F\}$。

注 特别地，上述定理表明，多胞形 P 的支撑集也是由 $\mathrm{vert}(P)$ 组成的多胞形，因此 P 只存在有限多支撑集。

综合以上分析，我们最终可以得到多胞形是多面体集这一重要结论。考虑到其严格推导较为冗杂，这里我们仅了解定理内容，对证明感兴趣的读者可参见 [2]。

定理 10.18

每个多胞形 $P\subseteq\mathbb{R}^n$ 都是多面体集。♡

10.5 极值表示

在先前的内容中，我们已经看到闭凸集 $A\subseteq\mathbb{R}^n$ 存在一种平凡表示：A 是所有包含 A 的闭凸集的交集，即 $A=\mathrm{conv}(A)$；该表示可改进为一种非平凡表示：A 是所有 A 的支撑半空间的交集。因此，讨论如何寻找一个尽可能小的真子集 $B\subseteq A$ 使得 $A=\mathrm{conv}(B)$ 是很自然的，而这是一个非平凡问题。尽管存在一些关于一般闭凸集的普遍结果，我们还是仅将注意力放在紧致情形，此时我们可以得到该问题的一个简单解。

定义 10.11 (极值点)

设 $A \subseteq \mathbb{R}^n$ 是闭凸集，若点 $x \in A$ 不能表示为 A 中点的非平凡凸组合，即对于 $y, z \in A$ 和 $\alpha(0,1)$，$x = \alpha y + (1-\alpha)z$ 必然给出 $x = y = z$，则称其为 A 的极值点（extreme point）。A 中所有极值点的集合记为 $\mathrm{ext}(A)$。

♣

注 为了对极值点有直观理解，容易证明 $\mathrm{ext}(A) \neq \varnothing$ 当且仅当集合 $A \subseteq \mathbb{R}^n$ 不包含任意直线（参见笔记《Convex Geometry》）。特别地，若 A 是闭半空间且 $n \geq 2$，则 $\mathrm{ext}(A) = \varnothing$。

命题 10.7

若 $\{x\}$ 是 A 的一个支撑集，则 $x \in \mathrm{ext}(A)$。

♠

证明 设 $\{x\} = A \cap \{f = \beta\}$ 且 $A \subseteq \{f \leq \beta\}$，并且存在 $y, z \in A$ 使得 $x = \alpha y + (1-\alpha)z$ 对某个 $\alpha \in (0,1)$ 成立。那么 $f(x) = \alpha f(y) + (1-\alpha) f(z) \leq \beta$，等号成立当且仅当 $f(y) = f(z) = \eta$，即 $y, z \in A \cap \{f = \beta\} = \{x\}$。因此 $y = z = x$，这意味着 $x \in \mathrm{ext}(A)$。

注 反之该结论并不成立。考虑类似 400 米跑道的图形围成的闭凸集，那么从“直道”到“弯道”的曲率变化点是极值点，但它不是支撑集。这引出了如下定义。

定义 10.12 (暴露点)

设 $A \subseteq \mathbb{R}^n$ 为闭凸集，若点 $x \in A$ 构成的单点集 $\{x\}$ 是 A 的一个 0-支撑集，则称其为暴露点（exposed point）。A 的暴露点集合记为 $\exp(A)$。

♣

注 根据命题 10.7，$\exp(A) \subseteq \mathrm{ext}(A)$。

命题 10.8

对于凸集 A 内的点 x，$x \in \mathrm{ext}(A)$ 当且仅当 $A \backslash \{x\}$ 是凸集。

♠

证明 对于充分性，考虑任意 $y, z \in A \backslash \{x\}$，假设存在 $\alpha \in [0,1]$ 使得 $x = \alpha y + (1-\alpha)z$。根据定义，$\alpha \in (0,1)$ 时 $y = z = x$；$\alpha = 0$ 和 1 时分别有 $z = x$ 和 $y = x$。以上情况均与 $y, z \in A \backslash \{x\}$ 矛盾，因此 y, z 的任意凸组合在 $A \backslash \{x\}$ 内，即 $A \backslash \{x\}$ 为凸集；反之若存在 $y, z \in A$ 和 $\alpha \in (0,1)$ 使得 $x = \alpha y + (1-\alpha)z$，假设 $y, z \in A \backslash \{x\}$，那么 $A \backslash \{x\}$ 的凸性要求 $x \in A \backslash \{x\}$，矛盾。因此 y 和 z 中至少有一个等于 x，进而 $y = z = x$，故 $x \in \mathrm{ext}(A)$。

根据此结论，对于多胞形 P，其顶点恰好为极值点，即 $\mathrm{vert}(P) = \mathrm{ext}(P)$。因此根据定理 10.4，$P = \mathrm{conv}(\mathrm{vert}(P)) = \mathrm{conv}(\mathrm{ext}(P))$。尽管对于一般的凸紧集 K，前者通常不成立，但 $K = \mathrm{conv}(K)$ 仍成立。

定理 10.19 (Minkowski 定理)

设 $K \subseteq \mathbb{R}^n$ 是凸紧集，$A \subseteq K$，则 $K = \mathrm{conv}(A)$ 当且仅当 $\mathrm{ext}(K) \subseteq A$。特别地，$K = \mathrm{conv}(\mathrm{ext}(K))$。

♡

证明 首先 $A \subseteq K$ 自然给出 $\mathrm{conv}(A) \subseteq K$，因此仅需证明反向包含。若 $K \subseteq \mathrm{conv}(A)$，那么对于任意 $x \in \mathrm{ext}(K) \subseteq K$，存在凸组合 $x = \alpha y + (1-\alpha)z$，其中 $y, z \in A$。假设存在 $x \notin A$，那么 $\alpha \in (0,1)$，极值点的定义要求 $x = y = z \in A$，矛盾。因此 $\mathrm{ext}(K) \subseteq A$；反之若 $\mathrm{ext}(K) \subseteq A$，对 n 采用归纳假设证明该结论。

当 $n = 1$ 时，由于 K 紧致，根据极值定理 1.5，存在 $a, b \in K$ 使得 $a = \inf(K)$ 且 $b = \sup(K)$，因此 $a \leq x \leq b$ 对任意 $x \in K$ 成立。容易证明 $a, b \in \mathrm{ext}(K)$，因此 $K \subseteq [a,b] \subseteq \mathrm{ext}(K) \subseteq A \subseteq \mathrm{conv}(A)$。假设结论在 $\mathbb{R}^{n-1}$ 上成立，考虑 $x \in K$ 和任意过 x 的直线 g。由于 K 紧致，根据 Heine-Borel 定理 1.10，K 为有界闭集，因此 $g \cap K$ 有界；对于 $g \cap K$ 内任意收敛到 x 的序列 (x_n)，首先 $x \in K$；利用收敛的线性性定理 1.15，$x \in g$，进而 $g \cap K$ 也是有界闭集。根据极值定理 1.5，存在 $y, z \in g \cap K$ 使其在 g 上关于给定的 $x_1, x_2 \in g$ 的仿射组合系数分别取最小值和最大值，因此 $g \cap K \subseteq [y,z]$；$g \cap K$ 的凸性要求 $[y,z] \subseteq g \cap K$，因此 $g \cap K = [y,z]$，并且容易证明 $y, z \in \partial K$。因此根据支撑定理 10.14，y, z 为支撑点，相应的支撑超平面分别记为 E_y 和 E_z，则 $y \in K \cap E_y \triangleq K_1$ 且 $z \in K \cap E_z \triangleq K_2$，其中 $\dim(E_y) = \dim(E_z) = n-1$。

根据归纳假设，凸紧集 $K_1 = \mathrm{conv}(\mathrm{ext}(K_1))$ 且 $K_2 = \mathrm{conv}(\mathrm{ext}(K_2))$。对于任意 $x' \in \mathrm{ext}(K_1)$，若存在 $y', z' \in K_1$ 和 $\alpha \in (0,1)$ 使得 $x' = \alpha y' + (1-\alpha)z'$，假设 y' 和 z' 至少有一个不在 $K \backslash K_1$ 内。不妨设 $y' \in K \backslash K_1 = K \cap E_y^c$，利用超平面的线性容易证明 $x' \in K \cap E_y^c$，矛盾。因此 $y', z' \in K_1$，进而极值点的定义要求 $x' = y' = z'$，这意味着 $\mathrm{ext}(K_1) \subseteq \mathrm{ext}(K)$；类似地 $\mathrm{ext}(K_2) \subseteq \mathrm{ext}(K)$。注意到 K_1 可表示为 $\mathrm{ext}(K_1)$ 的凸组合，因此可表示为 $\mathrm{ext}(K_1) \cup \mathrm{ext}(K_2) \subseteq \mathrm{ext}(K)$ 的凸组合，即 $K_1 \subseteq \mathrm{conv}(\mathrm{ext}(K))$；类似地 $K_2 \subseteq \mathrm{conv}(\mathrm{ext}(K))$，进而

$$x \in [y,z] = \mathrm{conv}\{y,z\} \subseteq \mathrm{conv}(K_1 \cup K_2) \subseteq \mathrm{conv}(\mathrm{ext}(K)) \subseteq \mathrm{conv}(A) \subseteq \mathrm{conv}(K) = K. \tag{10.5.1}$$

利用 $x \in K$ 的任意性，此时 $K = \mathrm{conv}(A) = \mathrm{conv}(\mathrm{ext}(K))$。因此对于任意自然数 n，该结论成立。

推论 10.4

设 $P \subseteq \mathbb{R}^n$ 是凸紧集，则 P 是多胞形当且仅当 $\mathrm{ext}(P)$ 有限。

♡

证明 若 P 为多胞形，根据命题 10.8后的讨论，$\mathrm{ext}(P)$ 有限；反之若 $\mathrm{ext}(P)$ 有限，根据定理 10.19，$K = \mathrm{conv}(\mathrm{ext}(P))$ 是有限多个点的凸包，因此 K 为多胞形。

有了上述结果，现在我们可以补充定理 10.18的反向结论，从而证明多胞形与有界多面体集的等价性。

定理 10.20

设 $P \subseteq \mathbb{R}^n$ 为有界多面体集，则 P 是多胞形。

♡

证明 显然 P 是凸紧集，根据推论 10.4，此时仅需证明 $\mathrm{ext}(P)$ 有限。设 $x \in \mathrm{ext}(P)$，并且 $P = \cap_{i=1}^k H_i$，其中 H_i 是以超平面 E_i 为边界的闭半空间。考虑凸集

$$D = \bigcap_{i=1}^{k} A_i, \quad \text{where } A_i = \begin{cases} E_i, & \text{if } x \in E_i, \\ \mathrm{relint}(H_i), & \text{if } x \notin E_i. \end{cases} \tag{10.5.2}$$

若 $x \in \cap_{i=1}^k E_i$，那么根据定理 9.3，D 也是仿射子空间。假设存在不同于 x 的点 $y \in D$，则其仿射组合也在 D 内。结合极值点定义 10.11后的讨论，此时 $\mathrm{ext}(P) = \varnothing$，矛盾。因此 $D = \{x\}$；若存在 $A_i = \mathrm{relint}(H_i)$，那么 $E_i \cap \mathrm{relint}(H_j)$ 与 $\mathrm{relint}(H_i) \cap \mathrm{relint}(H_j)$ 是相对开的，进而 D 是相对开的。假设 $\dim(D) > 0$，则在 $\mathrm{aff}(D)$ 上存在 $B(x,\epsilon) \subseteq D$，设 $\bar{B}(x, \epsilon/2)$ 与通过 x 的直线相交于 y, z，则 $x = (y+z)/2$，这与 $x \in \mathrm{ext}(P)$ 矛盾。因此无论在上述何种情况下，均有 $D = \{x\}$。由于 D 的可能数目有限（不超过 2^k），$\mathrm{ext}(P)$ 也有限，结论得证。

推论 10.5

有限多个多胞形的交集仍为多胞形。

♡

最后，在 Minkowski 定理 10.19中，若将 $\mathrm{ext}(K)$ 替换为 $\exp(K)$，则结论通常不成立。即使在平面中，凸紧集的暴露点集的凸包也未必紧致；不过该凸包的闭包可还原为 K。

定理 10.21

设 $K \subseteq \mathbb{R}^n$ 是凸紧集，则 $K = \mathrm{cl}(\mathrm{conv}(\exp(K)))$。

♡

证明 假设 K 不是单点集。由于 K 紧致，根据极值定理 1.5和命题 8.1，对于每个 $x \in \mathbb{R}^n$，存在 $y_x \in K$ 使得 $\|y_x - x\| = \sup_{y \in K} \|y - x\|$。若 $y_x \in \mathrm{relint}(K)$，则存在 $\mathrm{aff}(K)$ 上的 $B(y,\epsilon) \subseteq K$。那么存在不同于 y 的 $y_\epsilon \in B(x,\epsilon)$，使得 y_x 是 y_ε 和 x 的凸组合，从而 $\|y_\epsilon - x\| > \|y_x - x\|$，这与 y_x 的极大性矛盾。因此 $y_x \in \partial K$。对于通过 y_x 且与 $y_x - x$ 正交的超平面 E，若存在其他的 $y \in E \cap K$，那么

$$\|x-y\|^2 = \|(x-y_x) + (y_x - y)\|^2 = \|x-y_x\|^2 + \|y_x - y\|^2 > \|x - y_x\|^2, \tag{10.5.3}$$

这与 y_x 的极大性矛盾，因此 $E \cap K = \{y_x\}$，进而 E 是 K 的支撑超平面，故 $y_x \in \exp(K)$。根据前述分析，$\{y_x | x \in \mathbb{R}^n\} \subseteq \exp(K) \subseteq K$，因此 $\hat{K} \triangleq \mathrm{cl}(\{y_x | x \in \mathbb{R}^n\}) \subseteq K$。由于 K 紧致且 $\hat{K}$ 有界，根据 Heine-Borel 定理 1.10，$\hat{K}$ 紧致。后续证明稍微有些琐碎，易懂但书写较为麻烦，故再次略去，详见 [2]。

第三部分

测度论

第 11 章 测度

目录

11.1 代数与 σ-代数 93
11.2 测度 95
11.3 外测度 97
11.4 Lebesgue 测度的性质 100
11.5 完备性与正则性 103

11.1 代数与 σ-代数

定义 11.1 (代数)

设 X 为任意集合，若 X 的子集族（collection）$\mathscr{A}$ 满足

1. $X \in \mathscr{A}$；
2. 对每个集合 $A \in \mathscr{A}$，集合 $A^c \in \mathscr{A}$；
3. 对每个有限集合序列 $A_1, \cdots, A_n \in \mathscr{A}$，集合 $\cup_{i=1}^n A_i \in \mathscr{A}$；
4. 对每个有限集合序列 $A_1, \cdots, A_n \in \mathscr{A}$，集合 $\cap_{i=1}^n A_i \in \mathscr{A}$；

则称 $\mathscr{A}$ 是 X 上的代数（algebra）。

注 事实上，代数 $\mathscr{A}$ 关于取补集操作的封闭性表明，定义中 $X \in \mathscr{A}$ 的要求与 $\varnothing \in \mathscr{A}$ 是等价的；并且 $\mathscr{A}$ 关于有限交集和有限并集操作的封闭性是等价的，因此两者取其一作为定义即可。

定义 11.2 (σ-代数)

设 X 为任意集合，若 X 的子集族（collection）$\mathscr{A}$ 满足

1. $X \in \mathscr{A}$；
2. 对每个集合 $A \in \mathscr{A}$，集合 $A^c \in \mathscr{A}$；
3. 对 $\mathscr{A}$ 中每个无穷集合序列 $\{A_i\}$，集合 $\cup_{i=1}^\infty A_i \in \mathscr{A}$；
4. 对 $\mathscr{A}$ 中每个无穷集合序列 $\{A_i\}$，集合 $\cap_{i=1}^\infty A_i \in \mathscr{A}$；

则称 $\mathscr{A}$ 是 X 上的 σ-代数（σ-algebra）。

注 类似地，关于 $\mathscr{A}$ 取补集操作的封闭性表明，定义中 $X \in \mathscr{A}$ 的要求与 $\varnothing \in \mathscr{A}$ 是等价的；并且 $\mathscr{A}$ 关于可数交集和可数操作的封闭性是等价的，因此两者取其一作为定义即可。

定义 11.3 (可测性)

若 $\mathscr{A}$ 是集合 X 上的一个 σ-代数，若 X 的某个子集属于 $\mathscr{A}$，则称该子集 $\mathscr{A}$-可测（$\mathscr{A}$-measurable）。

例 11.1 这里介绍一些是或不是 σ-代数的例子。

1. 设 X 是一个集合，$\mathscr{A}$ 为 X 的幂集，则 $\mathscr{A}$ 是 X 上的 σ-代数；
2. 设 X 是一个集合，$\mathscr{A} = \{\varnothing, X\}$，则 $\mathscr{A}$ 是 X 上的 σ-代数；
3. 设 X 是一个无穷集，$\mathscr{A} = \{A \subseteq X | A \text{ finite or } A^c \text{ finite}\}$，则 $\mathscr{A}$ 是 X 上的代数，但不是 $X\sigma$-代数。

命题 11.1

设 X 是一个集合，则 X 上任意非空 σ-代数族的交集是 X 上的 σ-代数。

证明 考虑任意非空 σ-代数族 $\{\mathscr{A}_k|k\in K\}$，首先 $X\in\mathscr{A}_k$ 对所有 $k\in K$ 成立，因此 $X\in\cap_k A_k$；对于任意子集 $A\in\cap_k\mathscr{A}_k$，其补集 $A^c\in\mathscr{A}^k$ 对所有 $k\in K$ 成立，故 $A^c\in\cap_k\mathscr{A}_k$；对于任意子集族 $\{A_i|i\in I\}\subseteq\cap_k\mathscr{A}_k$，由于其并集 $\cup_i A_i\in\mathscr{A}_k$ 对所有 $k\in K$ 成立，故 $\cup_i A_i\in\cap_k\mathscr{A}_k$。综上，$\cap_k\mathscr{A}_k$ 为 σ-代数。

注 不过，一般而言 σ-代数族的并集不是 σ-代数。

类似“生成”的概念，上述命题可导出如下结论，该结论是构造 σ-代数的基本工具。

推论 11.1

设 X 是一个代数，$\mathscr{F}$ 是 X 的一个子集族，则存在 X 上包含 $\mathscr{F}$ 的唯一最小代数。

证明 首先任意集合 X 的幂集都包含 X 的子集族，并且自身构成一个 σ-代数，因此必然存在 X 上包含 $\mathscr{F}$ 的 σ-代数，记该 σ-代数族为 $\{\mathscr{A}_k|k\in K\}$。根据命题 11.1，$\mathscr{A}=\cap_k\in K\mathscr{A}_k$ 是包含 $\mathscr{F}$ 的 σ-代数。若存在包含 $\mathscr{F}$ 的最小 σ-代数 $\mathscr{B}$，那么根据定义，$\mathscr{A}\subseteq\mathscr{B}$ 且 $\mathscr{B}\subseteq\mathscr{A}$，即 $\mathscr{B}=\mathscr{A}$。因此这样的 σ-代数唯一。

利用此结论，我们可以从 X 的任意子集族出发生成一个 σ-代数。

定义 11.4 (生成 σ-代数)

对于 X 的任意子集族 $\mathscr{F}$，包含 $\mathscr{F}$ 的最小 σ-代数称为由 $\mathscr{F}$ 生成的 σ-代数（σ-algebra generated by $\mathscr{F}$），通常记为 $\sigma(\mathscr{F})$。

利用此定义，我们将定义一个贯穿整个测度论部分的重要 σ-代数。

定义 11.5 (Borel σ-代数)

在 $\mathbb{R}^d$ 上由 $\mathbb{R}^d$ 的开集族生成的 σ-代数称为 Borel σ-代数（Borel σ-algebra），记为 $\mathscr{B}(\mathbb{R}^d)$。$\mathbb{R}^d$ 的 Borel 子集（Borel set）是属于 $\mathscr{B}(\mathbb{R}^d)$ 的集合。对于 $d=1$ 的情形，一般将 $\mathscr{B}(\mathbb{R}^1)$ 简记为 $\mathscr{B}(\mathbb{R})$。

在进一步讨论 Borel σ-代数 $\mathscr{B}(\mathbb{R})$ 的性质之前，有必要说明其重要性主要源于以下两点：

- $\mathscr{B}(\mathbb{R})$ 几乎包含了分析中所有感兴趣的 $\mathbb{R}$ 的子集；
- $\mathscr{B}(\mathbb{R})$ 足够小，以至于可以用构造性（显式）的方式来处理。

命题 11.2

$\mathbb{R}$ 的所有 Borel 子集构成的 σ-代数 $\mathscr{B}(\mathbb{R})$ 可由以下任一集族生成：

1. $\mathbb{R}$ 的所有闭子集构成的族；
2. $\mathbb{R}$ 的所有形如 $(-\infty,b]$ 的子区间构成的族；
3. $\mathbb{R}$ 的所有形如 $(a,b]$ 的子区间构成的族。

证明 设 $\mathscr{B}_1,\mathscr{B}_2$ 和 $\mathscr{B}_3$ 分别对应命题中三个集族。我们将证明 $\mathscr{B}_3\subseteq\mathscr{B}_2\subseteq\mathscr{B}_1\subseteq\mathscr{B}(\mathbb{R})$，随后证明 $\mathscr{B}(\mathbb{R})\subseteq\mathscr{B}_3$ 以形成闭环。利用 $\mathscr{B}_2$ 在取补集和可数并集操作下封闭的性质，任意 $(a,b]=(-\infty,b)\cap(-\infty,a]^c\in\mathscr{B}_2$，因此 $\mathscr{B}_3\subseteq\mathscr{B}_2$；形如 (b,∞) 的集合是开的，因此 $(-\infty,b]$ 是闭的，从而属于 $\mathscr{B}_1$，即 $\mathscr{B}_2\subseteq\mathscr{B}_1$；利用 $\mathscr{B}(\mathbb{R})$ 在取补集操作下封闭的性质，$\mathbb{R}$ 的任意闭子集的补集属于 $\mathscr{B}(\mathbb{R})$，因此 $\mathscr{B}_1\subseteq\mathscr{B}(\mathbb{R})$。最后，根据命题 1.21，每个开集都可以表示为可数个形如 (a,b) 的开区间的并集，同时 $(a,b)=\cup_n(a,b-1/n]$。根据命题 1.7，可数个可数半开区间的并集是可数的，因此每个开区间都在 $\mathscr{B}_3$ 内，因此 $\mathscr{B}(\mathbb{R})\subseteq\mathscr{B}_3$。

对于 d 维 Euclidean 空间，我们有类似的结果。

命题 11.3

$\mathbb{R}^d$ 的所有 Borel 子集构成的 σ-代数可由以下任一集族生成：

1. $\mathbb{R}^d$ 的所有闭子集族；
2. $\mathbb{R}^d$ 的所有形如 $\{(x_1,\cdots,x_d)|x_i\leq b_i\}$ 的闭半空间，其中 i 为某个指标，$b\in\mathbb{R}$；
3. $\mathbb{R}^d$ 的所有形如 $\{(x_1,\cdots,x_d)|a_i<x_i\leq b_i \text{ for } i=1,\cdots,d\}$ 的矩形族。

我们现在考察 $\mathscr{B}(\mathbb{R}^d)$ 中集合的更多细节。设 $\mathscr{G}$ 是 $\mathbb{R}^d$ 中所有开子集构成的族，同时 $\mathscr{F}$ 是 $\mathbb{R}^d$ 中所有闭子集构成的族。另外设 $\mathscr{G}_\delta$ 是 $\mathscr{G}$ 中所有集合列的交集构成的族，$\mathscr{F}_\sigma$ 是 $\mathscr{F}$ 中所有集合列的并集构成的族。$\mathscr{G}_\delta$ 中的集合常称为 G_δ 集，$\mathscr{F}_\sigma$ 中的集合常称为 F_σ 集。字母 G 和 F 推测代表德语单词 *Gebiet*（区域）和法语单词 *fermé*（闭的），而字母 σ 和 δ 分别代表德语单词 *Summe*（和）与 *Durchschnitt*（交）。

命题 11.4

在 $\mathbb{R}^d$ 中，每个闭子集都是 G_δ 集，每个开子集都是 F_σ 集。

♠

证明 设 F 是 $\mathbb{R}^d$ 中的一个闭子集，我们希望构造 $\mathbb{R}^d$ 中的一个开子集列 $\{U_n\}_{n\in\mathbb{N}}$ 使得 $F=\cap_n U_n$。定义 $U_n=\left\{x\in\mathbb{R}^d \middle| \|x-y\|<1/n \text{ for some } y\in F\right\}$，显然每个 U_n 都是开集，并且 $F\subseteq U_n$ 对每个自然数 n 都成立，因此 $F\subseteq\cap_n U_n$；另一方面，对于任意 $x\in\cap_n U_n$，存在 $x_n\in F$ 使得 $\|x-x_n\|<1/n$ 对任意自然数 n 成立。因此 F 内的序列 (x_n) 收敛到 x，而 F 的闭性要求 $x\in F$，故 $\cap_n U_n\subseteq F$。综上，$F=\cap_n U_n$。

若 U 为开集，则 U^c 为闭集，因而存在开集列 $\{U_n\}$ 使得 $U^c=\cap_n U_n$，从而 $U=\cup_n U_n^c$ 为 F_σ 集。

11.2 测度

定义 11.6 (测度)

设 X 是一个集合，$\mathscr{A}$ 是 X 上的一个 σ-代数。对于一个定义域为 σ-代数 $\mathscr{A}$、值属于扩充半直线 $[0,+\infty]$ 的函数 μ，若其对于 $\mathscr{A}$ 中每个无穷不相交集合列 $\{A_i\}$ 均满足

$$\mu\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}\mu(A_i). \tag{11.2.1}$$

则称其可数可加（countably additive）。$\mathscr{A}$ 上的测度（measure）是满足 $\mu(\varnothing)=0$ 的可数可加函数 $\mu:\mathscr{A}\to[0,+\infty]$。

♣

注 由于 $\mu(A_i)$ 对每个 i 非负，因此实数列 $\{\sum_{i=1}^n\mu(A_i)\}_{n\in\mathbb{N}}$ 关于 n 是递增的。根据实数列的单调收敛准则（定理 1.12），该数列在扩充实数上收敛，因此 $\sum_{i=1}^{\infty}\mu(A_i)\le\infty$ 总是存在的。

与之相应的还有一种性质稍弱的测度。

定义 11.7 (有限可加测度)

设 $\mathscr{A}$ 是 X 上的一个代数（未必是 σ-代数），则对于一个定义域为 $\mathscr{A}$、值属于扩充半直线 $[0,+\infty]$ 的函数，若其对于 $\mathscr{A}$ 中每个有限不相交集合列 $A_1,\cdots,A_n$ 满足

$$\mu\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}\mu(A_i). \tag{11.2.2}$$

则称其限可加（finitely additive）。代数 $\mathscr{A}$ 上的有限可加测度（finitely additive measure）是满足 $\mu(\varnothing)=0$ 的有限可加函数 $\mu:\mathscr{A}\to[0,+\infty]$。

♣

注 有限可加性粗看似乎是一种比可数可加性更自然的性质。不过，一方面可数可加测度对于几乎所有应用都足够；另一方面，其支撑的积分理论比有限可加测度要强大的多。因此我们在后续的内容中，将几乎所有的注意力放在可数可加测度上。若不作特别说明，这一部分的所有“测度”默认指“可数可加测度”。

定义 11.8 (测度空间与可测空间)

若 X 是一个集合，$\mathscr{A}$ 是 X 上的一个 σ-代数，μ 是 $\mathscr{A}$ 上的测度，则三元组 $(X,\mathscr{A},\mu)$ 常称为测度空间（measure space）。类似地，若 X 是一个集合，$\mathscr{A}$ 是 X 上的一个 σ-代数，则有序对 $(X,\mathscr{A})$ 常称为可测空间（measurable space）。若 $(X,\mathscr{A},\mu)$ 是一个测度空间，则常称 μ 是 $(X,\mathscr{A})$ 上的一个测度；当 σ-代数在上下文明确时，也称其为 X 上的一个测度。

♣

例 11.2

1. 设 X 为任意集合，$\mathscr{A}$ 为 X 上的 σ-代数。如下定义函数 $\mu:\mathscr{A}\to[0,+\infty]$：若 A 是具有 n 个元素的有限集，则 $\mu(A)=n$；若 A 是无穷集，则 $\mu(A)=+\infty$。那么称 μ 为 $(X,\mathscr{A})$ 上的计数测度（*counting measure*）；
2. 设 X 是非空集合，$\mathscr{A}$ 为 X 上的 σ-代数。设 x 为 X 的一个元素，如下定义函数 $\delta_x:\mathscr{A}\to[0,+\infty]$：若 $x\in A$，则 $\delta_x(A)=1$；若 $x\notin A$，则 $\delta_x(A)=0$。那么称 δ_x 为集中在 x 处的质点（*point mass*）；
3. 设 X 为正整数集，$\mathscr{A}$ 为 X 中所有使得 A 或 A^c 有限的子集族。根据例11.1中的讨论，它是一个代数，但不是 σ-代数。如下定义函数 $\mu:\mathscr{A}\to[0,+\infty]$：若 A 无穷，则 $\mu(A)=1$；若 A 有限，则 $\mu(A)=0$。容易验证 μ 是有限可加测度，但不可能将 μ 扩张为由 $\mathscr{A}$ 生成的 σ-代数上的可数可加测度。例如考虑 $A_k=\{k\}$，$\mu(\cup_{k=1}^{\infty}A_k)=\mu(X)=1$，但 $\sum_{k=1}^{\infty}\mu(A_k)=0$；
4. 设 X 是任一集合，$\mathscr{A}$ 是 X 上的任意 σ-代数。如下定义函数 $\mu:\mathscr{A}\to[0,+\infty]$：若 $A\neq\varnothing$，则 $\mu(A)=+\infty$；若 $A=\varnothing$，则 $\mu(A)=0$。那么 μ 是一个测度。

命题 11.5 (测度的单调性)

设 $(X,\mathscr{A},\mu)$ 为测度空间，若 X 的子集 $A,B\in\mathscr{A}$ 满足 $A\subseteq B$，则 $\mu(A)\leq\mu(B)$。若 A 额外满足 $\mu(A)<\infty$，则 $\mu(B-A)=\mu(B)-\mu(A)$。

证明 由于 $A\cap(B-A)=\varnothing$，$\mu(B)=\mu(A\cup(B-A))=\mu(A)+\mu(B-A)\geq\mu(A)$。

定义 11.9 (σ-有限)

若 X 是集合列 $\{A_n\}_{n\in\mathbb{N}}$ 的并集，其中对于每个 i，A_i 均属于 $\mathscr{A}$ 且满足 $\mu(A_i)<\infty$，则称 μ 是 σ-有限的（σ-finite）。更一般地，若 $\mathscr{A}$ 中的一个集合是属于 $\mathscr{A}$ 且在 μ 下具有有限测度的集合列的并集，则其在 μ 下是 σ-有限的；同时也称相应的测度空间 $(X,\mathscr{A},\mu)$ 是 σ-有限的。

注 若测度空间 $(X,\mathscr{A},\mu)$ 是 σ-有限的，则 X 是属于 $\mathscr{A}$ 且在 μ 下具有有限测度的不相交集合列的并集：先按 μ 的 σ-有限性选择集合列 $\{A_n\}$，再设 $B_1=A_1,B_n=A_n-(\cup_{k=1}^{n}A_k)$ $(n\geq2)$ 以得到不相交集合列 $\{B_n\}$。

下面几个命题给出了测度的一些基本但有用的性质。

命题 11.6 (测度的可数次可加性)

设 $(X,\mathscr{A},\mu)$ 是一个测度空间。若 $\{A_n\}$ 是属于 $\mathscr{A}$ 的任意一个集合列，则 $\mu(\cup_k A_k)\leq\sum_n\mu(A_k)$。

证明 设 $B_1=A_1,B_k=A_k-(\cup_{i=1}^{k}A_i)$ $(k\geq2)$，首先 $B_1\in\mathscr{A}$，其次利用 $\mathscr{A}$ 关于交集和取补集操作的封闭性可知 $B_k=A_k\cap(\cup_{i=1}^{k}A_i)^c\in\mathscr{A}$。因此每个 $B_k\in\mathscr{A}$ 且 $B_k\subseteq A_k$。利用 μ 的可数可加性与单调性（命题 1.5），

$$\mu\left(\bigcup_k A_k\right)=\mu\left(\bigcup_k B_k\right)=\sum_k\mu(B_k)\leq\sum_k\mu(A_k). \tag{11.2.3}$$

注 上述定理表明，μ（关于不相交集合列的）可数可加性意味着 μ 的可数次可加性（*countable subadditivity*）。

命题 11.7

设 $(X,\mathscr{A},\mu)$ 为测度空间。

1. 若 $\{A_k\}$ 是属于 $\mathscr{A}$ 的递增集合列，则 $\mu(\cup_k A_k)=\lim_k\mu(A_k)$；
2. 若 $\{A_k\}$ 是属于 $\mathscr{A}$ 的递减集合列，且 $\mu(A_n)<\infty$ 对某个 n 成立，则 $\mu(\cap_k A_k)=\lim_k\mu(A_k)$。

证明

1. 对于递增集合列 $\{A_k\}$，设 $B_1=A_1,B_k=A_k-(\cup_{i=1}^{k}A_i)$ $(k\geq2)$，那么 $\{B_k\}$ 不相交且属于 $\mathscr{A}$，因此
$$\mu\left(\bigcup_k A_k\right)=\mu\left(\bigcup_k B_k\right)=\sum_k\mu(B_k)=\lim_k\sum_{i=1}^{k}\mu(B_k)=\lim_k\mu\left(\bigcup_{i=1}^{k}B_i\right)=\lim_k\mu(A_k). \tag{11.2.4}$$
2. 对于递减集合列 $\{A_k\}$，利用测度的单调性（命题 1.5），$\mu(A_1)\leq\mu(A_n)<\infty$。因此 $C_k=A_1-A_k=$

$A_1 \cap A_k^c \in \mathscr{A}$ 递增，并且 $\cup_k C_k = A_1 \cap (\cup_k A_k^c) = A_1 \cap (\cap_k A_k)^c = A_1 - \cap_k A_k$。结合命题 11.5和上述结论，

$$\mu(A_1) - \mu\left(\bigcap_k A_k\right) = \mu\left(A_1 - \bigcap_k A_k\right) = \mu\left(\bigcup_k C_k\right) = \lim_k \mu(C_k) = \mu(A_1) - \lim_k \mu(A_k), \tag{11.2.5}$$

整理得 $\mu(\cap_k A_k) = \lim_k \mu(A_k)$。

上述命题具有如下部分逆命题，这有时对验证有限可加测度的可数可加性有所帮助。

命题 11.8

设 $(X, \mathscr{A})$ 为可测空间，μ 是 $(X, \mathscr{A})$ 上的一个有限可加测度，则以下任一条件满足时，μ 为测度：

1. $\lim_k \mu(A_k) = \mu(\cup_k A_k)$ 对每个属于 $\mathscr{A}$ 的递增集合列 $\{\mathscr{A}_k\}$ 成立；
2. $\lim_k \mu(A_k) = 0$ 对每个属于 $\mathscr{A}$ 且满足 $\cap_k A_k = \varnothing$ 的递减集合列 $\{A_k\}$ 成立。

♠

证明

1. 设 $\{B_k\}$ 是属于 $\mathscr{A}$ 的不相交集合列，那么所有 $A_k = \cup_{i=1}^k B_i$ 构成的集合列 $\{A_k\}$ 递增，进而

$$\mu\left(\bigcup_k B_k\right) = \mu\left(\bigcup_k A_k\right) = \lim_k \mu(A_k) = \lim_k \mu\left(\bigcup_{i=1}^k B_i\right) = \lim_k \sum_{i=1}^k \mu(B_i) = \sum_k \mu(B_k), \tag{11.2.6}$$

可见 μ 也是可数可加的，进而 μ 为测度；

2. 设 $\{B_k\}$ 是属于 $\mathscr{A}$ 的不相交集合列，那么所有 $A_k = \cap_{i=k}^\infty B_i$ 构成的集合列 $\{A_k\}$ 递减。假设存在 $x \in \cap_k A_k$，那么 $x \in A_k$ 对每个 k 都成立。对于给定的 k，根据定理 1.7，自然数构成的集合 $\{i \in \mathbb{N} | x \in B_i, i \geq k\}$ 存在最小元素 $j \geq k$。但是 $\{B_k\}$ 不相交意味着 $x \notin B_{j+1}$，进而 $x \notin \cup_{i=j+1}^\infty B_i = A_{j+1}$，这与 x 在所有 A_k 中矛盾，因此 $\cap_k A_k = \varnothing$。那么根据已有条件，

$$\mu\left(\bigcup_k B_k\right) = \mu\left(\left(\bigcup_{i=1}^k B_i\right) \cup A_{k+1}\right) = \sum_{i=1}^k \mu(B_i) + \mu(A_{k+1}) \text{ for each } k \in \mathbb{N}. \tag{11.2.7}$$

等式两端令 $k \to \infty$，可得 $\mu(\cup_k B_k) = \sum_k \mu(B_k)$，因此 μ 是测度。

作为本节的结尾，我们引入一些术语。在 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 上的测度常称为 $\mathbb{R}^d$ 上的 *Borel* 测度（*Borel measure*）。一般地，若 X 是 $\mathbb{R}^d$ 的 Borel 子集，$\mathscr{A}$ 是由 $\mathbb{R}^d$ 中包含在 X 内的 Borel 子集构成的 σ-代数，则 $(X, \mathscr{A})$ 上的测度称为 X 上的 *Borel* 测度。现在假设 $(X, \mathscr{A})$ 是一个可测空间，它使得每个 X 中的元素 x 对应的单点集 $\{x\}$ 属于 $\mathscr{A}$。那么对于 $(X, \mathscr{A})$ 上有限或 σ-有限的测度 μ，若 $\mu(\{x\}) = 0$ 对每个 $x \in X$ 均成立，则称其连续（*continuous*）；若存在 X 的可数子集 D 使得 $\mu(D^c) = 0$，则称其离散（*discrete*）。

11.3 外测度

本节我们介绍一种构造测度的标准技巧，随后将其用于构造 $\mathbb{R}^d$ 上著名的 Lebesgue 测度。

定义 11.10 (外测度)

设 X 是一个集合，$\mathscr{P}(X)$ 是 X 的幂集。若 X 上的函数 $\mu^* : \mathscr{P}(X) \to [0, +\infty]$ 满足如下条件：

1. $\mu^*(\varnothing) = 0$；
2. 若 $A \subseteq B \subseteq X$，则 $\mu^*(A) \leq \mu^*(B)$；
3. 若 $\{A_n\}$ 是 X 的一个无穷子集列，则 $\mu^*(\cup_n A_n) \leq \sum_n \mu^*(A_n)$；

则称其为外测度（outer measure）。

♣

注 X 上的外测度是从 $\mathscr{P}(X)$ 到 $[0, +\infty]$ 且在 $\varnothing$ 处取值为 0 的函数，满足单调性（*monotone*）和可数次可加性。

需要注意的是，外测度一般不满足可数可加性，因此不为测度；另一方面，测度一般也不是外测度。当测度的定义域为 $\mathscr{P}(X)$ 时，显然测度 μ 为外测度；当测度的定义域为 $\mathscr{P}(X)$ 的真子集族 $\mathscr{A}$ 时，存在 X 的子集 A 不在 $\mathscr{A}$ 内，因此 μ 不是外测度。

例 11.3

1. 设 X 为集合，如下定义 $\mathscr{P}(X)$ 上的 μ^*：若 $A=\varnothing$，则 $\mu^*(A)=0$；否则 $\mu^*(A)=1$。那么 μ^* 为外测度；
2. 设 X 为集合，如下定义 $\mathscr{P}(X)$ 上的 μ^*：若 A 可数，则 $\mu^*(A)=0$；否则 $\mu^*(A)=1$。那么 μ^* 为外测度；
3. $\mathbb{R}$ 上的 *Lebesgue* 外测度（*Lebesgue outer measure*）通常记为 λ^*，定义如下：对于 $\mathbb{R}$ 的每个子集 A，设 $\mathscr{C}_A$ 包含所有使得 $A\subseteq\cup_i(a_i,b_i)$ 的有界开区间列 $\{(a_i,b_i)\}$，那么 $\lambda^*:\mathscr{P}(\mathbb{R})\to[0,+\infty]$ 定义为

$$\lambda^*(A)=\inf\left\{\sum\nolimits_i(b_i-a_i)\middle|\{(a_i,b_i)\}\in\mathscr{C}_A\right\}. \tag{11.3.1}$$

由于 $\mathbb{R}=\cup_n(n-\epsilon,n+1+\epsilon)$，故 $\mathscr{C}_A\neq\varnothing$。我们将通过下述命题证明 λ^* 确实是一个外测度。

命题 11.9

$\mathbb{R}$ 上的 Lebesgue 测度 λ^* 是一个外测度，并且其赋予 $\mathbb{R}$ 的每个子区间以其长度。

证明 对于任意正数 ϵ，存在开区间列 $\{(a_i,b_i)\}\in\mathscr{C}_A$，使得 $b_i-a_i=\epsilon/2^{i+1}$ 对所有 i 成立，因此 $\cup_i(b_i-a_i)=\epsilon/2<\epsilon$，进而 $0\leq\lambda^*(A)<\epsilon$。由 ϵ 的任意性，$\lambda^*(\varnothing)=0$；对于任意 $A\subseteq B\subseteq X$ 以及任意正数 ϵ，存在有界开区间列 $\{(a_i,b_i)\}\in\mathscr{C}_B\subseteq\mathscr{C}_A$ 使得 $\lambda^*(B)+\epsilon>\sum_i(b_i-a_i)\geq\lambda^*(A)$。由 ϵ 的任意性，$\lambda^*(B)\geq\lambda^*(A)$；设 $\{A_n\}$ 是 X 的一个无穷子集列，若 $\sum_n\lambda^*(A_n)=\infty$，则 $\lambda^*(\cup_nA_n)\leq\sum_n\lambda^*(A_n)$ 显然成立；因此假定 $\sum_n\lambda^*(A_n)<\infty$，那么对于任意正数 ϵ，存在开区间列 $\{(a_{n,i},b_{n,i})\}_{i=1}^\infty\in\mathscr{C}_{A_n}$，使得 $\lambda^*(A_n)+\epsilon/2^n>\sum_i(b_{n,i}-a_{n,i})$。根据命题 1.7，存在从 $\mathbb{N}\times\mathbb{N}$ 到 $\mathbb{N}$ 的双射，因此可将开区间列 $\{(a_{n,i},b_{n,i})\}_{n,i\in\mathbb{N}}$ 的指标用单个指标 $k\in\mathbb{N}$ 表示为 $\{(a_k,b_k)\}_{k\in\mathbb{N}}$，进而 $\cup_nA_n\subseteq\cup_{n,i}(a_{n,i},b_{n,i})=\cup_k(a_k,b_k)$，并且 $\sum_k(b_k-a_k)=\sum_{n,i}(b_{n,i}-a_{n,i})<\sum_n(\lambda^*(A_n)+\epsilon/2^n)=\sum_n\lambda^*(A_n)+\epsilon$。由 ϵ 的任意性，$\sum_n\lambda^*(A)_n\geq\sum_k(b_k-a_k)\geq\lambda^*(\cup_nA_n)$。

最后，对于 $\mathbb{R}$ 的有界闭区间 $[a,b]$ 以及任意正数 $\epsilon<b-a$，可取 $(a_1,b_1)=(a-\epsilon/4,b+\epsilon/4)$，而其余 $(a_i,b_i)\subseteq(a+\epsilon/2,b-\epsilon/2)$，同时满足 $\sum_i(b_i-a_i)<\epsilon/2^i$，因此 $[a,b]\subseteq\cup_i(a_i,b_i)$ 且 $\sum_i(b_i-a_i)=(b-a)+\epsilon$，因此 $\lambda^*([a,b])\leq b-a$；另一方面，对于任意正数 ϵ，存在开区间列 $\{(a_i,b_i)\}\in\mathscr{C}_{[a,b]}$ 使得 $\lambda^*([a,b])+\epsilon>\sum_i(b_i-a_i)$。根据 Heine-Borel 定理 1.10，区间 $[a,b]$ 紧致，因此存在有限指标集 I 使得 $[a,b]\subseteq\cup_{i\in I}(a_i,b_i)$。将 $\{(a_i,b_i)\}_{i\in I}$ 重排后可将上述关系写为 $[a,b]\subseteq\cup_{i=1}^n(a_i,b_i)$，其中 n 为某个自然数，利用归纳法可证明 $b-a\leq\sum_{i=1}^n(b_i-a_i)$。因此 $b-a\leq\sum_{i=1}^\infty(b_i-a_i)<\lambda^*([a,b])+\epsilon$。由 ϵ 的任意性，$\lambda^*([a,b])\geq b-a$，因此反向不等式也成立。

对于开区间 (a,b) 以及任意正数 $\epsilon<b-a$，由于 $[a+\epsilon/2,b-\epsilon/2]\subseteq(a,b)\subseteq[a-\epsilon/2,b+\epsilon/2]$，外测度的单调性要求 $(b-a)-\epsilon\leq\lambda^*((a,b))\leq(b-a)+\epsilon$。由 ϵ 的任意性，$\lambda^*(a,b)=b-a$。其余半开区间类似。

例 11.4 $\mathbb{R}^d$ 上的 *Lebesgue* 外测度通常记为 λ_d^* 或者 λ^*，定义如下：d 维区间（*d-dimensional interval*）是 $\mathbb{R}^d$ 的一个形如 $I_1\times\cdots\times I_d$ 的一个子集，其中 $I_1,\cdots,I_d$ 是 $\mathbb{R}$ 的子区间，并且

$$I_1\times\cdots\times I_d=\{(x_1,\cdots,x_d):x_i\in I_i\text{ for }i=1,\cdots,d\}. \tag{11.3.2}$$

注意到区间 $I_1,\cdots,I_d$ 可能为开的、闭的或者非开非闭的，从而 $I_1\times\cdots\times I_d$ 也如此。d 维区间的体积（*volume*）是区间 $I_1,\cdots,I_d$ 的长度之积，记为 $\mathrm{vol}(I_1\times\cdots\times I_d)$。对于 $\mathbb{R}^d$ 的每个子集 A，设 $\mathscr{C}_A$ 包含所有使得 $A\subseteq\cup_{i=1}^\infty R_i$ 的 d 维有界开区间列 $\{R_i\}$。那么 A 的外测度 $\lambda^*(A)$ 是集合 $\{\sum_{i=1}^\infty\mathrm{vol}(R_i)|\{R_i\}\in\mathscr{C}_A\}$ 的下确界。类似命题 11.9，我们可以证明上述 Lebesgue 外测度也是外测度。

命题 11.10

$\mathbb{R}^d$ 上的 Lebesgue 外测度是一个外测度，并且其赋予 $\mathbb{R}^d$ 的每个 d 为区间以其体积。

上述结论的证明与命题 11.9一致，故在此不进行赘述。现在我们引入 Lebesgue 测度的概念。

定义 11.11 (μ^*-可测)

设 X 是一个集合，μ^* 是 X 上的外测度。若 X 的某个子集 B 使得

$$\mu^*(A)=\mu^*(A\cap B)+\mu^*(A\cap B^c) \tag{11.3.3}$$

对 X 的每个子集 A 均成立，则称 B 是 μ^*-可测的（μ^*-measurable）或关于 μ^* 可测（measurable with respect to μ^*）。$\mathbb{R}$ 或 $\mathbb{R}^d$ 的 Lebesgue 可测（Lebesgue measurable）子集是关于 Lebesgue 外测度可测的子集。

注 对于 X 的任意子集 A 和 B，其可数次可加性天然要求 $\mu^*(A)\leq\mu^*(A\cap B)+\mu^*(A\cap B^c)$。因此如果希望验证某个子集 B 是 μ^*-可测的，只需对于测度有限的子集 A 验证反向不等式。

命题 11.11

设 X 为集合，μ^* 是 X 上的外测度，则 X 的每个满足 $\mu^*(B)=0$ 或 $\mu^*(B^c)=0$ 的子集 B 是 μ^*-可测的。♠

证明 若 $\mu^*(B)=0$，那么利用外测度的单调性与非负性，$\mu^*(A\cap B)\leq\mu^*(B)=0$，进而 $\mu^*(A\cap B)=0$；那么 $\mu^*(A\cap B)+\mu^*(A\cap B^c)\leq 0+\mu^*(A)$，根据上述讨论，$B$ 是 μ^*-可测的。$\mu(B^c)=0$ 的情形类似。

下述定理是关于外测度的重要基本事实，并且将是许多测度构造的关键。

定理 11.1

设 X 是一个集合，μ^* 是 X 上的外测度，$\mathscr{M}_{\mu^*}$ 是 X 的所有 μ^*-可测的子集构成的族，则

1. $\mathscr{M}_{\mu^*}$ 是一个 σ-代数；
2. μ^* 在 $\mathscr{M}_{\mu^*}$ 上的限制是 $\mathscr{M}_{\mu^*}$ 上的测度。

♡

证明

1. 根据命题 11.11，$\mu^*(\varnothing)=0$，因此 $X\in\mathscr{M}_{\mu^*}$；对于任意 $B\in\mathscr{M}_{\mu^*}$，μ^*-可测定义要求 $B^c\in\mathscr{M}_{\mu^*}$；对于任意不相交子集列 $\{B_n\}$ 和 X 的任意子集 A，

$$\mu^*(A)=\sum_{i=1}^{n}\mu^*(A\cap B_i)+\mu^*(A\cap(\cap_{i=1}^{n}B_i^c)). \tag{11.3.4}$$

显然上式在 $n=1$ 时成立；假设上式对某个自然数 n 成立，由于 $B_i\cap B_j=\varnothing$ 对任意 $i\neq j$ 成立，那么 $B_i\cap B_j^c=B_i\backslash(B_i\cap B_j)=B_i$，因此

$$\begin{aligned}\mu^*(A)&=\sum_{i=1}^{n}\mu^*(A\cap B_i)+\mu^*(A\cap(\cap_{i=1}^{n}B_i^c)\cap B_{n+1})+\mu^*(A\cap(\cap_{i=1}^{n+1}B_i^c))\\&=\sum_{i=1}^{n}\mu^*(A\cap B_i)+\mu^*(A\cap B_{n+1})+\mu^*(A\cap(\cap_{i=1}^{n+1}B_i^c))\\&=\sum_{i=1}^{n+1}\mu^*(A\cap B_i)+\mu^*(A\cap(\cap_{i=1}^{n+1}B_i^c)).\end{aligned} \tag{11.3.5}$$

因此对于任意自然数 n，(11.3.4)式成立。利用外测度的单调性，$\cap_{i=1}^{\infty}B_i^c\subseteq\cap_{i=1}^{n}B_i^c$ 意味着

$$\mu^*(A)\geq\sum_{i=1}^{n}\mu^*(A\cap B_i)+\mu^*(A\cap(\cap_{i=1}^{\infty}B_i^c)). \tag{11.3.6}$$

利用实数列收敛的单调性（定理 1.15）以及外测度的可数次可加性，令上述不等式两端 $n\to\infty$ 可得

$$\mu^*(A)\geq\sum_{n}\mu^*(A\cap B_n)+\mu^*(A\cap(\cap_n B_n^c))\geq\mu^*(A\cap(\cup_n B_n))+\mu^*(A\cap(\cup_n B_n)^c)\geq\mu^*(A), \tag{11.3.7}$$

因此上述不等号均取等，即 $\mathscr{M}_{\mu^*}$ 中任意不相交子集列的并集也属于 $\mathscr{M}_{\mu^*}$。

现在考虑 $\mathscr{M}_{\mu^*}$ 内任意子集列 $\{C_n\}$，首先考虑其中两个子集的并集：

$$\begin{aligned}\mu^*(A\cap(C_1\cup C_2))&=\mu^*(A\cap(C_1\cup C_2)\cap C_1)+\mu^*(A\cap(C_1\cup C_2)\cap C_1^*)\\&=\mu^*(A\cap C_1)+\mu^*(A\cap C_1^c\cap C_2).\end{aligned} \tag{11.3.8}$$

因此可证明 $C_1\cup C_2$ 也是 μ^*-可测的：

$$\begin{aligned}\mu^*(A)&=\mu^*(A\cap C_1)+\mu^*(A\cap C_1^c)=\mu^*(A\cap C_1)+\mu^*(A\cap C_1^c\cap C_2)+\mu^*(A\cap C_1^c\cap C_2^c)\\&=\mu^*(A\cap(C_1\cup C_2))+\mu^*(A\cap(C_1\cup C_2)^c).\end{aligned} \tag{11.3.9}$$

假设 $\cup_{i=1}^{n}C_i$ 对某个自然数 n 是 μ^*-可测的，注意到 $\cup_{i=1}^{n+1}C_i^{n+1}=(\cup_{i=1}^{n}C_i)\cup C_{n+1}$，采用完全相同的思路可证明 $\cup_{i=1}^{n+1}C_i$ 也是 μ^*-可测的。因此任意有限多 μ^*-可测的子集的并集也是 μ^*-可测的。结合开头证明的 $\mathscr{M}_{\mu^*}$ 在取补集操作下封闭的性质，通过 $B_1=C_1, B_n=C_n-(\cap_{i=1}^{n-1}C_i)\ (n\geq 2)$ 定义的不相交子集列 $\{B_n\}$

也是 μ^*-可测的。因此 $\cup_n C_n = \cup_n B_n \in \mathscr{M}_{\mu^*}$，即 $\mathscr{M}_{\mu^*}$ 在可数并集操作下是封闭的。综上，$\mathscr{M}_{\mu^*}$ 是 σ-代数；

2. (11.3.7)式表明，对于 $\mathscr{M}_{\mu^*}$ 内的不相交子集列 $\{B_n\}$ 以及 X 的任意子集 A，

$$\mu^*(A) = \sum_n \mu^*(A \cap B_n) + \mu^*(A \cap (\cup_n B_n)^c) \quad \Longrightarrow \quad \mu^*(\cup_n B_n) = \sum_n \mu^*(B_n), \tag{11.3.10}$$

因此 μ^* 在 $\mathscr{M}_{\mu^*}$ 上是可数可加的，并且之前已证明 $\mu^*(\varnothing) = 0$，因此 μ^* 是 $\mathscr{M}_{\mu^*}$ 上的测度。

现在我们应用定理 11.1以构造 Lebesgue 测度，首先记 $\mathbb{R}$ 的 Lebesgue 可测子集族为 $\mathscr{M}_{\lambda^*}$。

命题 11.12

$\mathbb{R}$ 的每个 Borel 子集都是 Lebesgue 可测的。

♠

证明 首先验证每个形如 $B = (-\infty, b]$ 的子集是 Borel 可测的。根据 μ^*-可测定义 11.11后的讨论，仅需讨论满足 $\lambda^*(A) < \infty$ 的子集 A。设 ϵ 为任一正数，那么存在开区间列 $\{(a_n, b_n)\}$ 使得 $\lambda^*(A) + \epsilon > \sum_n (b_n - a_n)$，其中 $A \subseteq \cup_n (a_n, b_n)$。由于集合 $(a_n, b_n) \cap B$ 与 $(a_n, b_n) \cap B^c$ 为有界不相交区间（可能为空），根据命题 11.9，

$$\lambda^*((a_n, b_n)) = b_n - a_n = (b_n - b) + (b - a_n) = \lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c). \tag{11.3.11}$$

再利用外测度的单调性与可数次可加性，

$$\lambda^*(A) \leq \lambda^*(\cup_n (a_n, b_n)) \leq \sum_n \lambda^*((a_n, b_n)) = \sum_n (b_n - a_n) < \lambda^*(A) + \epsilon. \tag{11.3.12}$$

由 ϵ 的任意性，上述不等式的等号均成立，因此

$$\lambda^*(A) \leq \lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \sum_n \lambda^*((a_n, b_n) \cap B) + \sum_n \lambda^*((a_n, b_n) \cap B^c) = \lambda^*(A), \tag{11.3.13}$$

此处不等式的等号也都成立，即 B 是 Lebesgue 可测的。根据定理 2.1，Lebesgue 可测集族 $\mathscr{M}_{\lambda^*}$ 是 $\mathbb{R}$ 上的 σ-代数且包含所有形如 $(-\infty, b]$ 的区间；又根据命题 11.2，$\mathscr{B}(\mathbb{R})$ 是 $\mathbb{R}$ 上包含所有形如 $(-\infty, b]$ 的区间的最小 σ-代数，因此 $\mathscr{B}(\mathbb{R}) \subseteq \mathscr{M}_{\lambda^*}$，从而每个 Borel 子集都是 Lebesgue 可测的。

仿照上述证明，在 d 维 Euclidean 空间也有类似的结论。

命题 11.13

$\mathbb{R}^d$ 的每个 Borel 子集都是 Lebesgue 可测的。

♠

由此我们可以正式定义 Lebesgue 测度。

定义 11.12 (Lebesgue 测度)

将 $\mathbb{R}^d$ 上的 Lebesgue 外测度限制在 $\mathbb{R}^d$ 的 Lebesgue 可测子集族 $\mathscr{M}_{\lambda^*}$ 上得到的测度称为 Lebesgue 测度（Lebesgue measure），记作 λ 或 λ_d；同时，将 Lebesgue 外测度限制在 $\mathscr{B}(\mathbb{R}^d)$ 上也得到相同的 Lebesgue 测度。一般通过指明可测空间是 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 还是 $(\mathbb{R}^d, \mathscr{M}_{\lambda^*})$ 来区分两种 Lebesgue 测度。

♣

注 根据上述定义，自然产生了两个问题：是否 $\mathbb{R}$ 的每个子集都是 Lebesgue 可测的？以及是否每个 Lebesgue 可测集是 Borel 集？事实上两者的回答都是否定的，我们将在本章稍后对前一个问题进行讨论。

11.4 Lebesgue 测度的性质

命题 11.14

设 A 为 $\mathbb{R}^d$ 的 Lebesgue 可测子集，那么

1. $\lambda(A) = \inf\{\lambda(U) | U \text{ is open and } A \subseteq U\}$；
2. $\lambda(A) = \sup\{\lambda(K) | K \text{ is compact and } K \subseteq A\}$。

♠

证明 对于 Lebesgue 可测子集 A，对于所有满足 $K\subseteq A\subseteq U$ 的开集 U 和紧致集 K，λ 的单调性自然要求

$$\lambda(K)\leq\lambda(A)\leq\lambda(U)\quad\Longrightarrow\quad\sup\{\lambda(K)|\text{compact } K\subseteq A\}\leq\lambda(A)\leq\inf\{\lambda(U)|\text{open } U\supseteq A\},\tag{11.4.1}$$

因此仅需证明其反向不等式。

1. 当 $\lambda(A)=\infty$ 时，显然有 $\sup\{\lambda(K)|\text{compact } K\subseteq A\}\leq\lambda(A)$；因此假设 $\lambda(A)<\infty$。设 ϵ 为任意正数，那么对于任意 Lebesgue 可测子集 A，存在 d 维开区间列 $\{R_n\}$ 使得 $\lambda(A)+\epsilon>\sum_n\text{vol}(R_n)$，其中 $A\subseteq\cup_n R_n$。取开集 $U=\cup_n R_n$（命题 1.20），那么利用测度的可数次可加性（命题 11.6）和命题 11.10可知
$$\inf\{\lambda(U)|\text{open } U\supseteq A\}\leq\lambda(U)\leq\sum_n\lambda(R_n)=\sum_n\text{vol}(R_n)<\lambda(A)+\epsilon,\tag{11.4.2}$$
由 ϵ 的任意性，$\inf\{\lambda(U)|\text{open } U\supseteq A\}\leq\lambda(A)$ 成立；
2. 首先处理 A 有界的情形。设 C 为包含 A 的有界闭集，ϵ 为任意正数，那么根据上一条结论，存在包含有界集合 $C-A$ 的开集 U 使得 $\lambda(C-A)+\epsilon>\lambda(U)$。根据命题 11.5，$\lambda(C)<\infty$ 意味着 $\lambda(C-A)=\lambda(C)-\lambda(A)$。那么对于闭集 $K=C\cap U^c=C-U$，$\lambda(K)=\lambda(C)-\lambda(U)>\lambda(A)+\epsilon$。由 ϵ 的任意性，$\lambda(A)\leq\lambda(K)\leq\sup\{\lambda(K)|\text{compact } K\subseteq A\}$。现在考虑 A 无界的情形，设 b 是小于 $\lambda(A)$ 的任意实数。设 $\{A_k\}$ 是 A 的一个递增有界可测子集列，使得 $A=\cup_k A_k$（例如取 $A_k=(-k,k)\cap A$）。根据命题 11.7，$\lambda(A)=\lambda(\cup_k A_k)=\lim_k\lambda(A_k)$，因此存在 $k_0\in\mathbb{N}$ 使得 $\lambda(A_k)>b$ 对所有 $k\geq k_0$ 成立。根据前述证明，对于有界的 A_k 和正数 b，存在 $\mathbb{R}$ 的紧子集 K 使得 $\lambda(K)>\lambda(A_k)+b$。由 b 的任意性和极限的单调性（定理 1.15），$\sup\{\lambda(K)|\text{compact } K\subseteq A\}\geq\lambda(K)\geq\lim_k\lambda(A_k)+0=\mu(A)$。

注 这两条性质实际上称为测度的正则性，我们将在第11.5节详细讨论。

引理 11.1

$\mathbb{R}^d$ 的每个开子集都是可数个形如

$$\{(x_1,\cdots,x_d)|j_i2^{-k}\leq x_i<(j_i+1)2^{-k}\text{ for }i=1,\cdots,d\},\quad\text{where}j_1,\cdots,j_d\in\mathbb{Z}\text{ and }k\in\mathbb{N}\tag{11.4.3}$$

的不相交半开立方体的并集。♡

证明 对于每个正整数 k，设 $\mathscr{C}_k$ 为所有形如(11.4.3)式的半开立方体构成的族，其中 k 给定。对于给定的 k，显然对于不同的 j_i，(11.4.3)式对应的半开立方体不相交；并且对于任意 $x_i\in\mathbb{R}$，Archimedean 性质（命题 1.19）保证存在自然数 n 使得 $x_i<n\cdot 2^{-k}$，并且定理 1.7保证存在最小的自然数 j_i 使得 $x_i<(j_i+1)2^{-k}$，即 $j_i2^{-k}\leq x_i$。因此 $\mathbb{R}^d=\cup\mathscr{C}_k$。结合上述两点，对于给定的 k，$\mathscr{C}_k$ 构成 $\mathbb{R}^d$ 的一个划分（回顾定义 1.6）。并且若 $k_1<k_2$，则 $\mathscr{C}_{k_2}$ 中每个对应指标为 j_i 的立方体都包含在对应指标为 $j_i2^{k_1-k_2}$ 的立方体内。

根据上述讨论，假定 U 是 $\mathbb{R}^d$ 的一个开子集，我们可以归纳地（类似定义 1.28）构造半开立方体族 $\mathscr{D}$，使得 $\cup\mathscr{D}\subseteq U$。对于空集，该结论显然成立。若 $\mathscr{D}$ 中包含 $\mathscr{C}_k$ 内的可数个不相交半开立方体的并集，那么选择 $\mathscr{C}_{k+1}$ 中与 $\mathscr{D}$ 中已有立方体不相交且包含在 U 内的立方体，将其并入 $\mathscr{D}$。由此得到的 $\mathscr{D}$ 满足 $\cup\mathscr{D}\subseteq U$。现在证明反向包含：对于任意 $x\in U$，由于 x 是开集，存在 $B(x,r)\subseteq U$。此证明第一段分析了给定 k 时 j_i 的存在性，因此对于充分大的 k，存在 $\mathscr{D}$ 内的半开立方体包含在 $B(x,r)$ 内，进而包含在 U 内，从而 $x\in\cup\mathscr{D}$，即 $U\subseteq\cup\mathscr{D}$。

命题 11.15

Lebesgue 测度是 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上唯一赋予每个 d 维区间或形如(11.4.3)式的半开立方体以体积的测度。♠

证明 根据命题 11.10，我们已知 Lebesgue 测度赋予每个 d 维区间以体积，并且对于每个形如(11.4.3)式的半开立方体，可将每个分量对应的区间嵌入两个开区间 $(j_i2^{-k}-\epsilon_i,(j_i+1)2^{-k})$ 和 $(j_i2^{-k}+\epsilon_i,(j_i+1)2^{-k})$，因此半开立方体也具有此体积。我们假定 μ 是 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上赋予每个形如(11.4.3)式的半开立方体以体积的测度，希望证明 $\mu=\lambda$。首先假定 U 是 $\mathbb{R}^d$ 的开子集，那么根据引理 11.1，存在可数个不相交半开立方体族 $\{C_j\}$ 使其并集为 U，因此 $\mu(U)=\sum_n\mu(C_n)=\sum_n\text{vol}(C_n)=\sum_n\lambda(C_n)=\lambda(U)$。因此 μ 和 λ 在 $\mathbb{R}^d$ 的所有开子集上相同。

现在假定 A 是 $\mathbb{R}^d$ 的任意 Borel 子集，若 U 是 $\mathbb{R}^d$ 内任意包含 A 的一个开子集，则测度的单调性（命题 11.5）要求 $\mu(A)\leq\mu(U)=\lambda(U)$，进而 $\mu(A)\leq\inf\{\lambda(U)|U\text{ is open and }A\subseteq U\}=\lambda(A)$。对于反向不等式，首先考虑

A 有界的情形，设 V 是一个包含 A 的有界开集，则 $\mu(V)=\mu(A)+\mu(V-A)\leq\lambda(A)+\lambda(V-A)=\lambda(V)$。而我们最开始已经证明 $\mu(V)=\lambda(V)$ 对所有开集 V 成立，这要求 $\mu(A)=\lambda(A)$。对于无界 Borel 子集 A，总可以找到一组的有界的不相交可测 Borel 子集列 $\{A_k\}$ 使得 $A=\cup_k A_k$，例如 $A_k=A\cap((-(k+1),-k]\cup[k,k+1))$。利用测度的可数可加性，$\mu(A)=\sum_k\mu(A_k)=\sum_k\lambda(A_k)=\lambda(\cup_k A_k)=\lambda(A)$。因此在 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上 $\mu=\lambda$。

对于 $\mathbb{R}^d$ 的每个元素 x 和子集 A，我们定义 $\mathbb{R}^d$ 的子集 $A+x=\{a+x\in\mathbb{R}^d|a\in A\}$。集合 $A+x$ 称为 A 沿 x 的平移（*translate*）。我们现在证明 Lebesgue 测度在这样的平移下具有不变性。

命题 11.16

$\mathbb{R}^d$ 上的 Lebesgue 外测度是平移不变的，即如果 $x\in\mathbb{R}^d$ 且 $A\subseteq\mathbb{R}^d$，则 $\lambda^*(A)=\lambda^*(A+x)$。此外，$\mathbb{R}^d$ 的子集 B 是 Lebesgue 可测的当且仅当 $B+x$ 是 Lebesgue 可测的。

♠

证明 根据 λ^* 的定义以及体积的平移不变性，自然有 $\lambda^*(A)=\lambda^*(A+x)$。若 B 是 Lebesgue 可测的，注意到 $(A-x)\cap B=\{b\in B|b+x\in A\}=\{a\in A|a\in B+x\}=A\cap(B+x)$，类似地 $(A-x)\cap B^c=A\cap(B^c+x)=A\cap(B+x)^c$，因此

$$\lambda^*(A)=\lambda^*(A-x)=\lambda^*((A-x)\cap B)+\lambda^*((A-x)\cap B^c)=\lambda^*(A\cap(B+x))+\lambda^*(A\cap(B+x)^c),\quad(11.4.4)$$

因此 $B+x$ 也是 Lebesgue 可测的；反之将 x 取为 $-x$，类似可证明必要性。

在常背书意义下，$(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上的Lebesgue测度可由以下结果刻画。对于更一般情况下的结果，参见第13章。

命题 11.17

设 μ 为 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上的非零测度，其在 $\mathbb{R}^d$ 的有界 Borel 子集上有限且平移不变，即 $\mu(A)=\mu(A+x)$ 对 $\mathscr{B}(\mathbb{R}^d)$ 中的每个 A 和 $\mathbb{R}^d$ 中的每个 x 成立。那么存在正数 c 使得 $\mu(A)=c\lambda(A)$ 对 $\mathscr{B}(\mathbb{R}^d)$ 中每个 A 成立。

♠

证明 设 $C=\{(x_1,\cdots,x_d)|0\leq x_i<1\text{ for each }i\}$，记 $c=\mu(C)$。假设 $c=0$，那么 $\mathbb{R}^d$ 作为 C 与其可数（命题 1.7）平移的并集，测度的可数可加性与平移不变性（命题 11.16）要求其测度 $\mu(\mathbb{R}^d)=0$，这与 μ 非零矛盾。因此 $c>0$，因此可对每个 Borel 子集 A 定义测度 $\nu(A)=c^{-1}\mu(A)$。对于任意形如(11.4.3)式的半开立方体 D，利用测度的可数可加性与平移不变性，$(2^d)^d\nu(D)=\nu(C)=\lambda(C)=2^{dk}\lambda(D)$，因此 $\nu(D)=\lambda(D)$，即 ν 赋予 d 维半开立方体以体积。根据命题 11.15，$\nu=\lambda$，即 $\mu=c\lambda$。

定理 11.2

存在 $\mathbb{R}$ 的一个子集不是 Lebesgue 可测的。

♡

证明 定义 $\mathbb{R}$ 上的一个等价关系：$x\sim y$ 当且仅当 $x-y\in\mathbb{Q}$，因此每个等价关系下的等价类都具有形式 $Q+x$。根据定理 1.8，对于任意实数 a 和 b，存在有理数 $r\in(a-x,b-x)$，因此每个 $Q+x$ 在 $\mathbb{R}$ 中均稠密，从而均与区间 $(0,1)$ 相交。又因为每个等价类不相交（命题 1.1），可通过选择公理 1.1取 $(0,1)$ 的一个子集 E，使其恰好包含每个等价类中的一个元素。我们将证明规范形 E 不是 Lebesgue 可测的。设 $\{r_n|n\in\mathbb{N}\}$ 为有理数在区间 $(-1,1)$ 内的一个枚举，并且对于每个 $n\in\mathbb{N}$ 取 $E_n=E+r_n$，可验证

- 集合 E_n 彼此不相交；
- $\cup_n E_n$ 包含在区间 $(-1,2)$ 内；
- 区间 $(0,1)$ 包含在 $\cup_n E_n$ 内。

若上述三条基本事实为真，假设 E 是 Lebesgue 可测的，那么 Lebesgue 测度的平移不变性（命题 11.16）要求每个 $E_n=E+r_n$ 也是 Lebesgue 可测的，从而 $\lambda(\cup_n E_n)=\sum_n\lambda(E_n)=\sum_n\lambda(E)$。若 $\lambda(E)=0$，则 $\lambda(\cup_n E_n)=0$，这与事实 3 矛盾；若 $\lambda(E)>0$，则 $\lambda(\cup_n E_n)=\infty$，这与事实 2 矛盾。因此 $E\subseteq(0,1)$ 不是 Lebesgue 可测的。

现在我们来证明上述三条事实。对于事实 1，若存在不同的 n 和 m 使得 $E_n\cap E_m\neq\varnothing$，则 $r_n-r_m\in E$。E 的定义要求 $r_n=r_m$，即 $E_n=E_m$，矛盾。因此 $\{E_n\}$ 彼此不相交；对于事实 2，由于 $E\subseteq(0,1)$ 且 $r_n\in(-1,1)$，$E_n=r_n+E\subseteq(-1,2)$ 对所有 $n\in\mathbb{N}$ 成立，进而 $\cup_n E_n\subseteq(-1,2)$；对于事实 3，对于任意 $x\in(0,1)$，设 e 是 E 中满足 $x\sim e$ 的元素，那么有理数 $x-e\in(-1,1)$，从而存在 r_n 使得 $x=e+r_n\in E_n\subseteq\cup_n E_n$，即 $(0,1)\subseteq\cup_n E_n$。

11.5 完备性与正则性

定义 11.13 (测度的完备性)

设 $(X,\mathscr{A},\mu)$ 为测度空间，若 $A\in\mathscr{A},\mu(A)=0$ 以及 $B\subseteq A$ 的关系共同给出 $B\in\mathscr{A}$ 的结论，则称测度 μ 或测度空间 $(X,\mathscr{A},\mu)$ 完备（complete）。有时也称这样的子集 $B\subseteq X$ 是 μ-可忽略的（μ-negligible）。♣

注 根据定理 11.1，若 μ^* 是集合 X 上的外测度，那么其单调性与非负性要求 $\mu^*(A)=0$ 且 $B\subseteq A$ 时 $\mu^*(B)=0$。此时对于任意子集 $C\subseteq X$，$\mu^*(C\cap B)\leq\mu^*(B)=0$，从而 $\mu^*(C\cap B)+\mu^*(C\cap B^c)\leq 0+\mu^*(C)$。根据定义 11.11后的讨论，$B$ 也是 μ^*-可测的。因此 μ^* 在 $\mathscr{M}_{\mu^*}$ 上的限制是完备的。特别地，测度空间 $(\mathbb{R}^d,\mathscr{M}_{\lambda^*},\mu^*)$ 是完备的，但测度空间 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d),\mu^*)$ 不是。我们将在第12章证明后者。

为了引出完备化的定义，我们首先需要一个引理。

引理 11.2

设 $(X,\mathscr{A},\mu)$ 为测度空间，对于 X 的任意子集 A，若存在 $E,F\in\mathscr{A}$ 使得 $E\subseteq A\subseteq F$ 且 $\mu(F-E)=0$，那么 $\mu(E)=\mu(F)=\sup\{\mu(B)|B\subseteq A,B\in\mathscr{A}\}$，即该值仅取决于测度 μ 和子集 $A\subseteq X$。♡

证明 利用可数可加性容易证明 $\mu(F)=\mu(E)+\mu(F-E)=\mu(E)$。为了说明该值与 E 和 F 的选择无关，记 $c=\sup\{\mu(B)|B\subseteq A,B\in\mathscr{A}\}$。对于 A 的任意子集 $B\in\mathscr{A}$，单调性给出 $\mu(B)\leq\mu(A)\leq\mu(F)$，因此 $c\leq\mu(F)$；另一方面，由于 B 可取为 A，故 $\mu(F)=\mu(E)\leq\mu(A)\leq c$。综上，$\mu(E)=\mu(F)=c$，并且 c 仅取决于测度 μ 和集合 $A\in\mathscr{A}$。

定义 11.14 (完备化)

设 $(X,\mathscr{A})$ 为可测空间，μ 是 $\mathscr{A}$ 上的一个测度，那么 $\mathscr{A}$ 在 μ 下的完备化（completion）是满足如下条件的子集 $A\subseteq X$ 构成的族 $\mathscr{A}_\mu$：存在 $E,F\in\mathscr{A}$ 使得 $E\subseteq A\subseteq F$ 且 $\mu(F-E)=0$。属于 $\mathscr{A}_\mu$ 的集合有时也称为 μ-可测的；μ 的完备化指的是函数 $\bar{\mu}:\mathscr{A}_\mu\to[0,+\infty]$，其将上述 $A\in\mathscr{A}$ 映射到 $\mu(E)$ 或 $\mu(F)$。♣

测度及其完备化具有如下关系。

命题 11.18

设 $(X,\mathscr{A})$ 为可测空间，μ 为 $\mathscr{A}$ 上的测度，那么 $\mathscr{A}_\mu$ 是 X 上包含 $\mathscr{A}$ 的 σ-代数，并且 $\bar{\mu}$ 是 $\mathscr{A}_\mu$ 上的完备测度，其在 $\mathscr{A}$ 上的限制是 μ。♠

证明 由于可取 $E=F=A$，因此任意 $A=E\in\mathscr{A}$ 在 $\mathscr{A}_\mu$ 内，即 $X\in\mathscr{A}\subseteq\mathscr{A}_\mu$；对于子集 $A\in\mathscr{A}_\mu$，存在 $E,F\in\mathscr{A}$ 使得 $E\subseteq A\subseteq F$ 且 $\mu(F-E)=0$。那么 $F^c,E^c\in\mathscr{A}$，同时满足 $F^c\subseteq A^c\subseteq E^c$ 和 $\mu(E^c-F^c)=\mu(E^c\cap F)=\mu(F-E)=0$，因此 $A^c\in\mathscr{A}_\mu$；对于子集族 $\{A_n\}\subseteq\mathscr{A}_\mu$ 中的每个 A_n，存在 $E_n,F_n\in\mathscr{A}$ 使得 $E_n\subseteq A_n\subseteq F_n$ 且 $\mu(F_n-E_n)=0$。记 $E=\cup_nE_n\in\mathscr{A},F=\cap_nF_n\in\mathscr{A}$，那么 $E\subseteq\cup_nA_n\subseteq F$，并且 $\mu(F-E)=\mu(\cap_n(F_n\cap E_n^c))\leq\mu(F_n-E_n)=0$ 对所有 n 成立，故 $\mu(F-E)=0$。因此 $\mathscr{A}_\mu$ 是包含 $\mathscr{A}$ 的 σ-代数。

现在证明 $\bar{\mu}$ 是 $\mathscr{A}_\mu$ 上的测度。根据引理 11.2，显然 $\bar{\mu}(A)\geq 0$ 且 $\bar{\mu}(\varnothing)=0$（取 $E=F=\varnothing$）；对于不相交子集列 $\{A_n\}\subseteq\mathscr{A}_\mu$，存在 $E_n,F_n\in\mathscr{A}$ 使得 $E_n\subseteq A_n\subseteq F_n$ 且 $\mu(F_n-E_n)=0$ 对所有 n 成立。而 A_n 不相交性意味着 E_n 不相交，因此根据 $\bar{\mu}$ 的定义，

$$\bar{\mu}(\cup_nA_n)=\mu(\cup_nE_n)=\sum_n\mu(E_n)\leq\sum_n\mu(A_n)\leq\sum_n\mu(F_n)=\sum_n\mu(E_n),\tag{11.5.1}$$

这要求上述不等式中等号均成立，因此 $\bar{\mu}(\cup_nA_n)=\sum_n\mu(E_n)=\sum_n\bar{\mu}(A_n)$。综上，$\bar{\mu}$ 为测度，并且对于任意 $A\in\mathscr{A}$，取 $E=F=A$ 可得 $\bar{\mu}(A)=\mu(A)$。最后，对于满足 $\bar{\mu}(A)=0$ 的子集 $A\in\mathscr{A}_\mu$，存在 $E,F\in\mathscr{A}$ 使得 $E\subseteq A\subseteq F$ 且 $\mu(F-E)=0$。因此对于任意 $B\subseteq A$，$\varnothing,F\in\mathscr{A}$ 且 $\varnothing\subseteq B\subseteq A\subseteq F$，同时满足 $\mu(F-\varnothing)=\mu(F)=\bar{\mu}(A)=0$。可见 $B\in\mathscr{A}$，即 $\bar{\mu}$ 是完备的。

为了讨论 Lebesgue 测度的完备性，我们首先需要一个引理。

引理 11.3

设 A 为 $\mathbb{R}^d$ 的一个 Lebesgue 可测子集，则存在 $\mathbb{R}^d$ 的 Borel 子集 E 和 F，使得 $E \subseteq A \subseteq F$ 且 $\lambda(F-E)=0$。♡

证明 设 A 是 $\mathbb{R}^d$ 的一个有界 Lebesgue 可测子集，根据命题 11.14，对于任意自然数 n，存在开集 $U_n \supseteq A$ 使得 $\lambda(A)+1/n > \lambda(U_n)$，也存在紧子集 $K_n \subseteq A$ 使得 $\lambda(A)-1/n < \lambda(K_n)$。取 $\mathbb{R}^d$ 的 Borel 子集 $U=\cap_n U_n, K=\cup_n K_n$，那么首先有 $K \subseteq A \subseteq U$，并且 $\lambda(U_n) < \lambda(A)+1/n < \infty$，根据命题 11.5，$\lambda(U_n-K_n)=\lambda(U_n)-\lambda(K_n)<2/n$。因此 $\lambda(U-K)=\lambda(\cap_n(U_n-K_n)) \le \lambda(U_n-K_n)<2/n$ 对所有 $n\in\mathbb{N}$ 成立。根据 Archimedean 性质1.19与测度的非负性，$\lambda(U-K)=0$。

对于任意 Lebesgue 可测子集 A，存在有界 Lebesgue 可测集列 $\{A_n\}$，使得 $A=\cup_n A_n$，例如 $A_n=A\cap(-n,n)$。对于每个 A_n，存在 $\mathbb{R}^d$ 的 Borel 子集 E_n 和 F_n 使得 $E_n \subseteq A_n \subseteq F_n$ 且 $\lambda(F_n-E_n)=0$。记 Borel 子集 $E=\cup_n E_n, F_n=\cap_n F_n$，那么 $\lambda(F-E)=\lambda(\cap_n(F_n-E_n)) \le \lambda(F_n-E_n)=0$ 对所有 n 成立，即 $\lambda(F-E)=0$。

命题 11.19

$(\mathbb{R}^d, \mathscr{M}_{\lambda^*})$ 上的 Lebesgue 测度是 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 上的 Lebesgue 测度的完备化。♠

证明 设 λ 和 λ_m 分别为 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 和 $(\mathbb{R}^d, \mathscr{M}_{\lambda^*})$ 上的 Lebesgue 测度，$\bar{\lambda}$ 为 λ 的完备化。引理 11.3表明 $\mathscr{M}_{\lambda^*}$ 包含在 $\mathscr{B}(\mathbb{R}^d)$ 的完备化内，因此 $\bar{\lambda}$ 在 $\mathscr{M}_{\lambda^*}$ 的限制为 μ_m。反之，对于每个 $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ 的完备化内的子集 A，存在 Borel 子集 E 和 F，使得 $E \subseteq A \subseteq F$ 且 $\lambda(F-E)=0$。根据命题 11.13和11.18，$F-E$ 是 Lebesgue 可测的，因此 $\lambda_m(F-E)=\bar{\lambda}(F-E)=0$，而 $\bar{\lambda}$ 的完备性要求 $\mu_m(A-E)=0$，因此 $A-E\in\mathscr{M}_{\lambda^*}$，进而 $A=(A-E)\cup E \in \mathscr{M}_{\lambda^*}$。结论得证。

不过，上述结论的反向结果并不成立。我们将在第12.1节证明：

- 存在 $\mathbb{R}^d$ 的 Lebesgue 可测子集不是 Borel 子集；
- Lebesgue 测度在 $\mathscr{B}(\mathbb{R}^d)$ 上的限制不是完备的。

不过，尽管在测度空间的完备化内讨论可避免一些困难，但同时也引入了其他困难：一方面，测度空间的完备化可能比原测度空间复杂得多；另一方面，定义在同一可测空间上的不同测度的完备化可能并不相同。因此，如无必要，我们在后文将尽可能避免引入测度空间的完备化。

定义 11.15 (外测度与内测度)

设 $(X, \mathscr{A})$ 为可测空间，μ 为 $\mathscr{A}$ 上的测度，A 为 X 的任意子集。

1. A 的外测度（outer measure）$\mu^*(A)$ 定义为 $\mu^*(A)=\inf\{\mu(B)|A\subseteq B, B\in\mathscr{A}\}$；
2. A 的内测度（inner measure）$\mu_*(A)$ 定义为 $\mu_*(A)=\sup\{\mu(B)|B\subseteq A, B\in\mathscr{A}\}$。♣

注 容易验证 $\mu_*(A)\le\mu^*(A)$ 对 X 的任意子集 A 成立；当 $\mu^*(A)<+\infty$ 时，$\mu_*(A)=\mu^*(A)$ 当且仅当 $A\in\mathscr{A}_\mu$。

命题 11.20

设 $(X, \mathscr{A}, \mu)$ 为测度空间，则定义 11.15中的外测度是定义 11.10中的一种外测度。♠

证明 显然 $\mu^*(\varnothing)=0$；对于子集 $A\subseteq B$，由于所有包含 B 的集合构成的族包含在所有包含 A 的集合构成的族内，因此 $\mu^*(A)\le\mu^*(B)$；对于无穷子集列 $\{A_n\}$ 和任意正数 ϵ，显然 $\sum_n\mu^*(A_n)=+\infty$ 时 $\mu(\cup_n A_n)\le\sum_n\mu(A_n)$ 成立。若 $\sum_n\mu^*(A)<+\infty$，存在 $B_n\in\mathscr{A}$ 使得 $\mu^*(A_n)+\epsilon/2^n>\mu(B_n)$ 对所有 $n\in\mathbb{N}$ 成立，其中 $A_n\subseteq B_n$。利用测度的可数次可加性，$\mu^*(\cup_n A_n)\le\mu(\cup_n B_n)\le\sum_n\mu(B_n)<\sum_n\mu^*(A_n)+\epsilon$。利用 ϵ 的任意性，$\mu^*(\cup_n A_n)\le\sum_n\mu^*(A_n)$。

定义 11.16 (正则性)

设 $\mathscr{A}$ 为 $\mathbb{R}^d$ 上包含 $\mathscr{B}(\mathbb{R}^d)$ 的 σ-代数，若 $(\mathbb{R}^d, \mathscr{A})$ 上的测度满足如下条件，则称其正则（regular）：

1. $\mathbb{R}^d$ 的每个紧子集 K 满足 $\mu(K)<+\infty$；

2. $\mathscr{A}$ 中每个集合满足 $\mu(A)=\inf\{\mu(U)|U \text{ is open and } A\subseteq U\}$；
3. $\mathbb{R}^d$ 中每个开子集 U 满足 $\mu(U)=\sup\{\mu(K)|K \text{ is compact and } K\subseteq U\}$。

♣

注 命题 11.14表明 Lebesgue 测度在 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 和 $(\mathbb{R}^d,\mathscr{M}_{\lambda^*})$ 上均为正则的。

引理 11.4

设 μ 为 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上的有限测度，则 $\mathbb{R}^d$ 的每个 Borel 子集 A 满足

$$\mu(A)=\inf\{\mu(U)|A\subseteq U \text{ and } U \text{ is open}\}, \tag{11.5.2}$$

$$\mu(A)=\sup\{\mu(C)|C\subseteq A \text{ and } C \text{ is closed}\}. \tag{11.5.3}$$

♡

证明 记首先证明该结论对于任意开集 $A\subseteq\mathbb{R}^n$ 成立。根据 μ 的单调性（命题 11.5），$\sup\{\mu(C)|\text{closed } C\subseteq A\}\leq\mu(A)\leq\inf\{\mu(U)|\text{open } A\subseteq U\}$；另一方面，存在 $U=A$ 使得 $\mu(U)=\mu(A)$，因此 $\inf\{\mu(U)|\text{open } U\subseteq A\}\leq\mu(A)$。从而(11.5.2)式成立。对于(11.5.3)式，根据命题 11.4，存在闭子集列 $\{C_n\}$ 使得 $A=\cup_n C_n$。将 $\cup_{i=1}^n C_i$ 取为新的 C_n，那么 $\{C_n\}$ 为递增列。根据命题 11.7，$\mu(A)=\mu(\cup_n C_n)=\lim_\mu(C_n)$。因此对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $\mu(A)<\mu(C_n)+\epsilon\leq\sup\{\mu(C)|\text{closed } C\subseteq A\}+\epsilon$ 对所有 $n\geq n_0$ 成立。因此(11.5.3)式对开集 A 也成立。

由于结论涉及 Borel 子集，典型的证明思路是构造一个包含所有开集的 σ-代数，从而其包含 Borel σ-代数。现在证明满足该引理条件的所有子集 $A\subseteq\mathbb{R}^d$ 构成的集族 $\mathscr{R}$ 是一个 σ-代数。首先 $\varnothing,\mathbb{R}^d$ 是开集，因此属于 $\mathscr{R}$；对于任意 $A\in\mathscr{R}$，和任意正数 ϵ，存在开集 U 和闭集 C，使得 $C\subseteq A\subseteq U$ 且 $\mu(U)-\epsilon/2<\mu(A)<\mu(C)+\epsilon/2<\infty$，因此 $\mu(U-C)=\mu(U)-\mu(C)<\epsilon$；那么 U^c 为闭集、C^c 为开集，同时 $U^c\subseteq A^c\subseteq C^c$ 且 $\mu(C^c-U^c)=\mu(C^c\cap U)=\mu(U-C)<\epsilon$，故 $A^c\in\mathscr{R}$；对于子集族 $\{A_n\}\subseteq\mathscr{R}$，其中对于每个 A_n 均存在开集 U_n 和闭集 C_n，使得 $C_n\subseteq A_n\subseteq U_n$ 且 $\mu(C_n-U_n)<\epsilon$。记开集 $U=\cup_n U_n$，闭集 $C=\cap_n C_n$，那么 $C\subseteq\cup_n A_n\subseteq U$，同时测度的单调性要求 $\mu(C-U)=\mu(\cap_n(C_n-U_n))\leq\mu(U_n-C_n)<\epsilon$，因此 $\cup_n A_n\in\mathscr{R}$。综上，$\mathscr{R}$ 是一个 σ-代数。

命题 11.21

设 μ 为 $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上的有限测度，则 μ 是正则的。此外，$\mathbb{R}^d$ 的每个 Borel 子集 A 满足

$$\mu(A)=\sup\{\mu(K)|K\subseteq A \text{ and is compact}\}. \tag{11.5.4}$$

♠

证明 μ 的有限性使得正则性第一条要求自然满足；根据引理 11.4，第二条要求满足；根据引理 11.4，对于任意开集 U 和任意正数 ϵ，存在闭子集 $C\subseteq A$ 使得 $\mu(A)-\epsilon<\mu(C)$。选择有界递增闭集列 $\{C_n\}$ 使得 $C=\cup_n C_n$，例如 $C_n=C\cap\{x\in\mathbb{R}^d|\|x\|_2\leq n\}$。根据命题 11.7，$\mu(C)=\lim_n\mu(C_n)$。因此当 n 充分大时，根据 Heine-Borel 定理 1.10，存在紧子集 C_n 使得 $\mu(A)<\mu(C_n)+\epsilon\leq\sup\{\mu(K)|\text{compact } K\subseteq A\}+\epsilon$。因此(11.5.4)式成立，这意味着正则性的第三条要求满足。综上，$(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$ 上的有限测度 μ 是正则的。

第 12 章 函数与积分

目录

12.1 可测函数 106
12.2 积分 110
12.3 极限定理 116

12.1 可测函数

在引入可测函数之前，我们先介绍一个基本结果。

命题 12.1

设 $(X,\mathscr{A})$ 是一个可测空间，$A\in\mathscr{A}$ 是 X 的一个子集。对于函数 $f:A\to[-\infty,+\infty]$，以下条件等价：

1. 对于每个实数 t，集合 $\{x\in A|f(x)\le t\}\in\mathscr{A}$；
2. 对于每个实数 t，集合 $\{x\in A|f(x)<t\}\in\mathscr{A}$；
3. 对于每个实数 t，集合 $\{x\in A|f(x)\ge t\}\in\mathscr{A}$；
4. 对于每个实数 t，集合 $\{x\in A|f(x)>t\}\in\mathscr{A}$。

证明

1. 若表述 1 成立，则 $\{x\in A|f(x)<t\}=\cup_n\{x\in A|f(x)\le t-1/n\}\in\mathscr{A}$，表述 2 成立；
2. 若表述 2 成立，则 $\{x\in A|f(x)\ge t\}=\{x\in A|f(x)<t\}^c\in\mathscr{A}$，表述 3 成立；
3. 若表述 3 成立，则 $\{x\in A|f(x)>t\}\in\mathscr{A}=\cap_n\{x\in A|f(x)\ge t+1/n\}\in\mathscr{A}$，表述 4 成立；
4. 若表述 4 成立，则 $\{x\in A|f(x)\le t\}=\{x\in A|f(x)>t\}^c\in\mathscr{A}$，表述 1 成立。

定义 12.1 (可测函数)

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的一个子集。若函数 $f:A\to[-\infty,+\infty]$ 满足命题 12.1的任一条件，则称其关于 $\mathscr{A}$ 可测（measurable with respect to $\mathscr{A}$）。关于 $\mathscr{A}$ 可测的函数有时也称为 $\mathscr{A}$-可测（$\mathscr{A}$-measurable）；当 σ-代数在上下文明确时，简称为可测。关于 $\mathscr{B}(\mathbb{R}^d)$ 可测的函数称为 Borel 可测（Borel measurable）或 Borel 函数（Borel function），关于 $\mathscr{M}_{\lambda^*}$ 可测的函数称为 Lebesgue 可测（Lebesgue measurable）。

注 根据命题 11.13，显然每个 $\mathbb{R}^d$ 上的 Borel 可测函数是 Lebesgue 可测的。

我们先讨论一些可测函数的例子，随后讨论可测函数的基本性质。

例 12.1

1. 设 $f:\mathbb{R}^d\to\mathbb{R}$ 连续，则对于每个实数 t，$f^{-1}((-\infty,t))=\left\{x\in\mathbb{R}^d|f(x)<t\right\}$ 是开集，故 f 是 Borel 可测的；
2. 设 I 是 $\mathbb{R}$ 的一个子区间，$f:I\to\mathbb{R}$ 是非减的。那么对于每个实数 t，集合 $\{x\in I|f(x)<t\}\subseteq I$（可能为空）是一个 Borel 集，因此 f 是 Borel 可测的；
3. 设 $(X,\mathscr{A})$ 是一个可测空间，B 是 X 的一个子集。那么 B 的示性函数 χ_B 是 $\mathscr{A}$-可测的当且仅当[1] $B\in\mathscr{A}$；
4. 若一个函数只有有限多个非零值，则称其为简单函数（*simple function*）。设 $(X,\mathscr{A})$ 是一个可测空间，$f:X\to[-\infty,+\infty]$ 是简单函数，$\alpha_1,\cdots,\alpha_n$ 是 f 的值，则 f 是 $\mathscr{A}$-可测的当且仅当[2] $\{x\in X|f(x)=\alpha_i\}\in\mathscr{A}$ 对所有 $i=1,\cdots,n$ 成立。

[1] 若 $B\in\mathscr{A}$，对于任意实数 t，$\{x\in X|\chi_B(x)<t\}$ 可能为 $\varnothing$、B^c 或 X，因此属于 $\mathscr{A}$，即 χ_B 是 $\mathscr{A}$-可测的；反之若 χ_B 是 $\mathscr{A}$-可测的，$B^c=\{x\in X|\chi_B(x)<1\}\in\mathscr{A}$，因此 $B=(B^c)^c\in\mathscr{A}$。

[2] 对于充分性，$\{x\in X|f(x)=\alpha_i\}=\{x\in X|f(x)\le\alpha_i\}\cap\{x\in X|f(x)\ge\alpha_i\}\in\mathscr{A}$；反之对于必要性，考虑任意实数 t，记有限指标集 $I_t=\{i|\alpha_i<t\}$，那么 $\{x\in X|f(x)<t\}=\cup_{i\in I_t}\{x\in X|f(x)=\alpha_i\}\in\mathscr{A}$，因此 f 可测。

命题 12.2

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的子集，f 和 g 为 A 上取值于 $[-\infty,+\infty]$ 的可测函数。那么集合 $\{x\in A|f(x)<g(x)\},\{x\in A|f(x)\leq g(x)\}$ 和 $\{x\in A|f(x)=g(x)\}$ 均属于 $\mathscr{A}$。

证明 利用有理数的稠密性（定理 1.8）及其可数性（命题 1.7），

$$\{x\in A|f(x)<g(x)\}=\bigcup_{r\in\mathbb{Q}}(\{x\in A|f(x)<r\}\cup\{x\in A|r<g(x)\})\in\mathscr{A},\tag{12.1.1}$$

因此我们有

$$\{x\in A|f(x)\leq g(x)\}=\{x\in A|g(x)<f(x)\}^c\in\mathscr{A},\tag{12.1.2}$$

$$\{x\in A|f(x)=g(x)\}=\{x\in A|f(x)\leq g(x)\}\cap\{x\in A|g(x)\leq f(x)\}\in\mathscr{A}.\tag{12.1.3}$$

定义 12.2

设 f 和 g 为 A 上取值于 $[-\infty,+\infty]$ 的函数，(f_n) 为 A 上取值于 $[-\infty,+\infty]$ 的函数列。

1. f 和 g 最大值和最小值分别定义为

$$(f\vee g)(x)=\max\{f(x),g(x)\}\quad\text{and}\quad(f\wedge g)(x)=\min\{f(x),g(x)\}.\tag{12.1.4}$$

2. (f_n) 的上确界和下确界分别定义为

$$(\sup_n f_n)(x)=\sup\{f_n(x)|n\in\mathbb{N}\}\quad\text{and}\quad(\inf_n f_n)(x)=\inf\{f_n(x)|n\in\mathbb{N}\}.\tag{12.1.5}$$

3. (f_n) 的上极限和下极限分别定义为

$$(\limsup_n f_n)(x)=\limsup\{f_n(x)|n\in\mathbb{N}\}\quad\text{and}\quad(\liminf_n f_n)(x)=\liminf\{f_n(x)|n\in\mathbb{N}\}.\tag{12.1.6}$$

4. (f_n) 的极限定义为 $(\lim_n f_n)(x)=\lim\{f_n(x)|n\in\mathbb{N}\}$。

注 根据命题 1.25，$(\lim_n f_n)(x)$ 定义在集合 $\{x\in A|\limsup_n f_n(x)=\liminf_n f_n(x)\}$ 上，未必定义在整个 A 上。

命题 12.3

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的一个子集，f 和 g 为 A 上取值于 $[-\infty,+\infty]$ 的函数，则 $f\vee g$ 和 $f\wedge g$ 均为 $\mathscr{A}$-可测的。

证明 $f\vee g$ 和 $f\wedge g$ 的可测性源于如下恒等式：

$$\{x\in A|(f\vee g)(x)<t\}=\{x\in A|f(x)<t\}\cap\{x\in A|g(x)<t\}\in\mathscr{A},\tag{12.1.7}$$

$$\{x\in A|(f\wedge g)(x)<t\}=\{x\in A|f(x)<t\}\cup\{x\in A|g(x)<t\}\in\mathscr{A}.\tag{12.1.8}$$

命题 12.4

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的一个子集，(f_n) 为 A 上取值于 $[-\infty,+\infty]$ 的可测函数列，那么

1. 函数 $\sup_n f_n$ 和 $\inf_n f_n$ 是 $\mathscr{A}$-可测的；
2. 函数 $\limsup_n f_n$ 和 $\liminf_n f_n$ 是 $\mathscr{A}$-可测的；
3. 定义在 $A_0=\{x\in A|\limsup_n f_n(x)=\liminf_n f_n(x)\}$ 上的函数 $\lim_n f_n$ 是 $\mathscr{A}$-可测的。

证明

1. $\sup_n f_n$ 和 $\inf_n f_n$ 的可测性源于如下恒等式：

$$\{x\in A|\sup_n f_n(x)<t\}=\bigcap_n\{x\in A|f_n(x)<t\}\in\mathscr{A},\tag{12.1.9}$$

$$\{x\in A|\inf_n f_n(x)<t\}=\bigcup_n\{x\in A|f_n(x)<t\}\in\mathscr{A}.\tag{12.1.10}$$

2. 对于每个正整数 k，可定义 $f_k=\sup\{f_n|n\geq k\}$ 和 $g_k=\inf\{f_n|n\geq k\}$。根据命题 1.26，

$$\{x\in A|\limsup_n f_n(x)<t\}=\{x\in A|\inf_k f_k(x)<t\}\in\mathscr{A},\tag{12.1.11}$$

$$\{x\in A|\liminf_n f_n(x)<t\}=\{x\in A|\sup_k g_k(x)<t\}\in\mathscr{A}.\tag{12.1.12}$$

3. 根据上述结论与命题 12.2，$\limsup_n f_n$ 和 $\limsup_n g_n$ 是 $\mathscr{A}$-可测的，因此 $A_0\in\mathscr{A}$，进而

$$\{x\in A_0|\lim_n f_n(x)<t\}=A_0\cap\{x\in A|\limsup_n f_n(x)<t\}\in\mathscr{A}.\tag{12.1.13}$$

在以下两个命题中，我们讨论取值于 $[0,+\infty]$ 的可测函数在 $\mathbb{R}$ 上的算术运算；对于取值于 $[-\infty,+\infty]$ 的函数，其算术运算更棘手，也较少使用。

命题 12.5

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的一个子集，f 和 g 为 A 上取值于 $[0,+\infty]$ 的可测函数，α 为非负实数，那么 αf 和 $f+g$ 是可测的。

证明 当 $\alpha=0$ 时，$\{x\in A|\alpha f(x)\leq t\}=A\in\mathscr{A}$ 对所有 $t\geq 0$ 成立；当 $\alpha>0$ 时，$\{x\in A|\alpha f(x)\leq t\}=\{x\in A|f(x)\leq t/\alpha\}\in\mathscr{A}$，因此 αf 可测。利用有理数的稠密性（定理 1.8）及其可数性（命题 1.7），

$$\{x\in A|(f+g)(x)<t\}=\bigcup_{r\in\mathbb{Q}}(\{x\in A|f(x)<r\}\cap\{x\in A|r<t-g(x)\})\in\mathscr{A},\tag{12.1.14}$$

因此 $f+g$ 也是可测的。

命题 12.6

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的一个子集，f 和 g 为 A 上的实值可测函数，α 为实数，那么 $\alpha f,f+g,f-g,fg$ 和 f/g 可测，其中 f/g 的定义域为 $\{x\in A|g(x)\neq 0\}$。

证明 仿照命题 12.5的证明，分 α 为正、为负、为零的三种情况讨论，可证明 $\alpha f,f+g$ 可测。另外，$-g=(-1)g$ 也可测，因此 $f-g=f+(-g)$ 可测。下面证明 h 可测时 h^2 也可测。

首先对于实数 $t\leq 0$，$\{x\in A|h^2(x)<t\}=\varnothing\in\mathscr{A}$；对于 $t>0$，$\{x\in A|h^2(x)<t\}=\{x\in A|h(x)<\sqrt{t}\}\cap\{x\in A|h(x)>-\sqrt{t}\}\in\mathscr{A}$。因此 h^2 可测，进而 $fg=[(f+g)^2-(f^2+g^2)]/2$ 可测。现在设 $A_0=\{x\in A|g(x)\neq 0\}$ 为 f/g 的定义域，那么 $\{x\in A_0|(f/g)(x)<t\}$ 恰好为

$$(\{x\in A|g(x)>0\}\cap\{x\in A|f(x)<tg(x)\})\cup(\{x\in A|g(x)<0\}\cap\{x\in A|f(x)>tg(x)\}),\tag{12.1.15}$$

而括号内的两个集合各自属于 $\mathscr{A}$，故 f/g 可测。

定义 12.3 (可测函数的正部与负部)

设 A 是一个集合，f 是 A 上的扩充实值函数，那么其正部（positive part）与负部（negative part）分别是定义为 $f^+=f\vee 0$ 和 $f^-=-f\vee 0$ 的扩充实值函数，即对于任意 $x\in A$，

$$f^+(x)=\max\{f(x),0\}\quad\text{and}\quad f^-(x)=\min\{f(x),0\}.\tag{12.1.16}$$

注 若 $(X,\mathscr{A})$ 是可测空间，那么命题 12.3和12.6表明，那么定义在 X 的子集 A 上的函数 f 可测当且仅当 f^+ 和 f^- 均可测，此时 $|f|=f^++f^-$ 也可测。

命题 12.7

设 $(X,\mathscr{A})$ 为可测空间，$A\in\mathscr{A}$ 是 X 的一个子集，f 为 A 上取值于 $[-\infty,+\infty]$ 的可测函数。

1. 若 f 是 $\mathscr{A}$-可测的，$B\in\mathscr{A}$ 是 A 的一个子集，则 f 在 B 上的限制是 $\mathscr{A}$-可测的；
2. 若 $\{B_n\}$ 是属于 $\mathscr{A}$ 的一个集合列且 $A=\cup_n B_n$，同时对于每个 n，f 在 B_n 上的限制是 $\mathscr{A}$-可测的，则 f 是 $\mathscr{A}$-可测的。

证明 该结论源于以下两个恒等式（请读者自行证明）：

$$\{x \in B | f(x) < t\} = B \cap \{x \in A | f(x) < t\} \in \mathscr{A}, \tag{12.1.17}$$

$$\{x \in A | f(x) < t\} = \bigcup_n \{x \in B_n | f(x) < t\} \in \mathscr{A}. \tag{12.1.18}$$

在本章后续讨论中，我们将反复用到以下基本结果。

命题 12.8

设 $(X, \mathscr{A})$ 为可测空间，$A \in \mathscr{A}$ 是 X 的一个子集，f 为 A 上取值于 $[0, +\infty]$ 的可测函数，那么存在 A 上取值于 $[0, +\infty)$ 的可测简单函数列，其满足

$$f_1(x) \leq f_2(x) \leq \cdots \quad \text{and} \quad f(x) = \lim_n f_n(x) \text{ for each } x \text{ in } A. \tag{12.1.19}$$

证明 对于每个正整数 n 和正整数 $k \in \{1, \cdots, n2^n\}$，$f$ 可测要求 $A_{n,k} = \{x \in A | (k-1)/2^n \leq f(x) < k/2^n\} \in \mathscr{A}$。定义从 A 到 $\mathbb{R}$ 的函数列 (f_n)，使得 f_n 在 $A_{n,k}$ 上每个点取值为其左端点值 $(k-1)/2^n$，而在 $A - \cup_k A_{n,k}$ 上每个点取值为 n。可见每个 $f_n(x)$ 只有 $n2^n + 1$ 个取值，是简单函数；根据例12.1中的讨论，f_n 可测当且仅当 $\{x \in A | f_n(x) = (k-1)/2^n\} = A_{n,k} \in \mathscr{A}$ 且 $\{x \in A | f_n(x) = n\} = A - \cup_k A_{n,k} \in \mathscr{A}$。这意味着对于任意实数 t，$\{x \in A | f_n(x) < t\}$ 包含至多可数个 $A_{n,k}$ 关于 k 的并集，因此 f_n 可测。容易验证函数列 (f_n) 关于 n 在每个 $x \in A$ 上单调递增，下面证明该函数列收敛到 f。

若 $f(x) < +\infty$，根据 Archimedean 性质（命题 1.19）以及定理 1.7，存在最小的正整数 n 使得 $n-1 \leq f(x) < n$，即 $x \in \cup_k A_{n,k}$，那么存在正整数 k 使得 $x \in A_{n,k}$，使得 $|f_n(x) - f(x)| = f(x) - f_n(x) < n - (n - 1/2^n) = 1/2^n$，因此 $\lim_n f_n(x) = x$；若 $f(x) = +\infty$，那么根据 $A_{n,k}$ 的定义，$x \in A - \cup_k A_{n,k}$ 对所有正整数 n 成立，因此 $f_n(x) = n$，从而 $\lim_n f_n(x) = +\infty = f(x)$。

推论 12.1

设 $(X, \mathscr{A})$ 为可测空间，$A \in \mathscr{A}$ 是 X 的一个子集，f 为 A 上取值于 $[0, +\infty]$ 的可测函数，那么存在 A 上取值于 $\mathbb{R}$ 的可测简单函数列，使得 $f(x) = \lim_n f_n(x)$ 对每个 $x \in A$ 成立。

证明 根据命题 12.3，f 的正部 f^+ 和负部 f^- 均可测，因此根据命题 12.8，分别存在取值于 $[0, +\infty)$ 的简单函数列 (f_n^+) 和 (f_n^-)，使得 $f^+(x) = \lim_n f_n^+(x)$ 和 $f^-(x) = \lim_n f_n^-(x)$ 对每个 $x \in A$ 成立。根据命题 12.6，对每个 n 定义可测简单函数 $f_n = f_n^+ - f_n-$，从而 $\lim_n f_n(x) = f(x)$ 对每个 $x \in A$ 成立。

以下命题给出了一些考虑可测函数的其他方法，其中最后一种可用于处理更一般的情形。

命题 12.9

设 $(X, \mathscr{A})$ 为可测空间，$A \in \mathscr{A}$ 是 X 的一个子集。对于函数 $f : A \to \mathbb{R}$，以下条件等价：

1. f 关于 $\mathscr{A}$ 可测；
2. 对于 $\mathbb{R}$ 的每个开集 U，集合 $f^{-1}(U) \in \mathscr{A}$；
3. 对于 $\mathbb{R}$ 的每个闭集 C，集合 $f^{-1}(C) \in \mathscr{A}$；
4. 对于 $\mathbb{R}$ 的每个 Borel 子集 B，集合 $f^{-1}(B) \in \mathscr{A}$。

证明 由于涉及 Borel 子集，我们需要构造一个包含所有开集的 σ 代数。设子集族 $\mathscr{F} = \{B \subseteq \mathbb{R} | f^{-1}(B) \in \mathscr{A}\}$，那么首先有 $f^{-1}(\mathbb{R}) = A \in \mathscr{A}$；对于任意 $B \in \mathscr{F}$，$f^{-1}(B^c) = (f^{-1}(B))^c \in \mathscr{A}$，因此 $B^c \in \mathscr{F}$；对于任意子集族 $\{B_n\} \in \mathscr{F}$，$f^{-1}(\cup_n B_n) = \cup_n f^{-1}(B_n) \in \mathscr{A}$，因此 $\cup_n B_n \in \mathscr{A}$。综上，$\mathscr{F}$ 构成一个 σ-代数。

根据命题 12.1，f 关于 $\mathscr{A}$ 可测当且仅当 $\{x \in A | f(x) \leq b\} = f^{-1}((-\infty, b]) \in \mathscr{A}$，因此 $(-\infty, b] \in \mathscr{F}$；根据命题 11.2，$\mathscr{F}$ 包含所有形如 $(-\infty, b]$ 的区间当且仅当 $\mathscr{F}$ 包含 Borel σ-代数，因此表述 1 和表述 4 等价。再次利用命题 11.2表述 2、表述 3 和表述 4 彼此等价。

在本节先前的所有讨论中，我们仅考虑 f 是取值于扩充实数的函数，即 f 是从可测空间 $(X, \mathscr{A})$ 到可测空间 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 的映射。若函数 f 的像为一般可测空间，其可测性也有类似的推广定义。

定义 12.4 (可测函数)

设 $(X,\mathscr{A})$ 和 $(Y,\mathscr{B})$ 为可测空间。对于函数 $f:X\to Y$ 以及 $\mathscr{B}$ 中的每个 B，若集合 $f^{-1}(B)\in\mathscr{A}$，则称 f 关于 $\mathscr{A}$ 和 $\mathscr{B}$ 可测（measurable with respect to $\mathscr{A}$ and $\mathscr{B}$）；有时也称 f 是从 $(X,\mathscr{A})$ 到 $(Y,\mathscr{B})$ 的可测函数（measurable function），或者简单称 $f:(X,\mathscr{A})\to(Y,\mathscr{B})$ 是可测的。类似地，对于子集 $A\subseteq X$ 以及函数 $f:A\to Y$，若 $A\in\mathscr{A}$ 且 $f^{-1}(B)\in\mathscr{A}$ 对所有 $B\in\mathscr{B}$ 成立，则称 f 可测。

根据定义，我们很容易证明复合函数的可测性。

命题 12.10

设 $(X,\mathscr{A}),(Y,\mathscr{B})$ 和 $(Z,\mathscr{C})$ 为可测空间，$f:(Y,\mathscr{B})\to(Z,\mathscr{C})$ 和 $g:(X,\mathscr{A})\to(Y,\mathscr{B})$ 均可测，则 $f\circ g:(X,\mathscr{A})\to(Z,\mathscr{C})$ 也是可测的。

下述结论对于验证函数的可测性往往十分有用。

命题 12.11

设 $(X,\mathscr{A})$ 和 $(Y,\mathscr{B})$ 为可测空间，$\mathscr{B}_0$ 为 Y 的子集族，满足 $\sigma(\mathscr{B}_0)=\mathscr{B}$，那么函数 $f:X\to Y$ 关于 $\mathscr{A}$ 和 $\mathscr{B}$ 可测当且仅当 $f^{-1}(B)\in\mathscr{A}$ 对每个 $\mathscr{B}_0$ 中的 B 成立。

证明 充分性是显然的；对于必要性，设 $f^{-1}(B)\in\mathscr{A}$ 对 $\mathscr{B}_0$ 中每个 B 成立，$\mathscr{F}=\{B\subseteq Y|f^{-1}(B)\in\mathscr{A}\}$。类似命题 12.9的证明，可知 $\mathscr{F}$ 是一个 σ-代数。由于 $\mathscr{F}$ 包含 $\mathscr{B}_0$，故 $\mathscr{B}=\sigma(\mathscr{B}_0)\subseteq\mathscr{F}$。

12.2 积分

在正式开始本章内容之前，我们需要先引入“几乎处处成立”的概念；这一概念有时放宽了 f 在整个定义域上的要求。

定义 12.5 (μ-几乎处处成立)

设 $(X,\mathscr{A},\mu)$ 为测度空间，若 X 中使得某个性质不成立的点集是 μ-可忽略的（μ-negligible），则称该性质在 X 上 μ-几乎处处成立（hold μ-almost everywhere）；换言之，若存在集合 $N\in\mathscr{A}$ 满足 $\mu(N)=0$ 且包含性质不成立的每个点，则称该性质 μ-几乎处处成立。更一般地，若 E 是 X 的一个子集，那么使得 E 中该性质不成立的点集是 μ-可忽略的性质称为 μ-几乎处处成立。μ-几乎处处成立常简写为 μ-a.e. 或 a.e.$[\mu]$。当测度 μ 在上下文明确时，也常几乎处处成立简写为 a.e.。

注 对于几乎处处成立的性质，设 F 为 X 中使得该性质不成立的点构成的集合，需要注意 F 不必属于 $\mathscr{A}$，只要求存在包含 F 的 $\mathscr{A}$-可测子集 N。当然，若 μ 完备，则有 $F\in\mathscr{A}$。

12.2.1 简单函数的积分

定义 12.6 (积分)

设 $(X,\mathscr{A},\mu)$ 为测度空间，$\mathscr{S}$ 为 X 上所有简单实值 $\mathscr{A}$-可测函数构成的族，$\mathscr{S}_+$ 为 $\mathscr{S}$ 中所有非负函数构成的族。若 $f\in\mathscr{S}_+$ 且由 $f=\sum_{i=1}^m a_x\chi_{A_i}$ 给出，其中 $a_1,\cdots,a_m$ 为非负实值，$A_1,\cdots,A_m\in\mathscr{A}$ 是 X 的不相交子集，那么 f 关于 μ 的积分（integral）$\int f\,\mathrm{d}\mu$ 定义为 $\sum_{i=1}^m a_i\mu(A_i)$。

注 这里需要说明的是，积分 $\int f\,\mathrm{d}\mu$ 仅依赖于简单函数 f 本身而不依赖于其展开表示 $\sum_{i=1}^m a_i\mu(A_i)$。设 f 还可以表示为 $\sum_{j=1}^n b_j\chi_{B_j}$，其中 $b_1,\cdots,b_n$ 是非负实值，$B_1,\cdots,B_n\in\mathscr{A}$ 是 X 的不相交子集。若 $\cup_{i=1}^m A_i\subseteq\cup_{j=1}^n B_j$，则记 $A_{m+1}=(\cup_{i=1}^m A_i)\backslash(\cup_{j=1}^n B_j)$ 并取 $a_{m+1}=0$；反之类似。因此不失一般性，总可以设 $\cup_{i=1}^m A_i=\cup_{j=1}^n B_j$，进

而 $A_i = A_i \cap (\cup_{i=1}^m A_i) = A_i \cap (\cup_{j=1}^m B_j) = \cup_{j=1}^m (A_i \cap B_j)$，其中 $A_i \cap B_j$ 关于不同的 i 和 j 不相交。另外，在非空的 $A_i \cap B_j$ 上，$a_i = a_i \chi_{A_i}(A_i \cap B_j) = b_j \chi_{B_j}(A_i \cap B_j) = b_j$，因此

$$\sum_{i=1}^m a_i \mu(A_i) = \sum_{i=1}^m a_i \sum_{j=1}^n \mu(A_i \cap B_j) = \sum_{j=1}^n b_j \sum_{i=1}^m \mu(A_i \cap B_j) = \sum_{j=1}^n b_j \mu(B_j). \tag{12.2.1}$$

可见 $\int f \,\mathrm{d}\mu$ 不依赖于 f 的具体形式表达。

首先证明一些关于简单函数积分的基本性质。

命题 12.12

设 $(X, \mathscr{A}, \mu)$ 为可测空间，$f, g \in \mathscr{S}_+$，α 为非负实数，则

1. $\int \alpha f \,\mathrm{d}\mu = \alpha \int f \,\mathrm{d}\mu$；
2. $\int (f+g) \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu$；
3. 若 $f(x) \le g(x)$ 对 X 中每个点成立，则 $\int f \,\mathrm{d}\mu \le \int g \,\mathrm{d}\mu$。

♠

证明 设 $f = \sum_{i=1}^m a_i \chi_{A_i}$，其中 a_i 是非负实值，$A_i \in \mathscr{A}$ 是 X 的不相交子集。类似可设 $g = \sum_{j=1}^n b_j \chi_{B_j}$。性质 1 是显然的，对于性质 2 和 3，利用测度的可数可加性与非负性，

$$\begin{aligned}\int (f+g) \,\mathrm{d}\mu &= \sum_{i,j} (a_i + b_j) \mu(A_i \cap B_j) = \sum_{i=1}^m a_i \sum_{j=1}^n \mu(A_i \cap B_j) + \sum_{j=1}^n b_j \sum_{i=1}^m \mu(A_i \cap B_j) \\ &= \sum_{i=1}^m a_i \mu(A_i) + \sum_{j=1}^n b_j \mu(B_j) = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu, \end{aligned} \tag{12.2.2}$$

$$\int g \,\mathrm{d}\mu = \int [f + (g-f)] \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int (g-f) \,\mathrm{d}\mu \ge \int f \,\mathrm{d}\mu. \tag{12.2.3}$$

命题 12.13

设 $(X, \mathscr{A}, \mu)$ 为测度空间，$f \in \mathscr{S}_+$，(f_n) 为 $\mathscr{S}_+$ 中的非减序列，且使得 $f(x) = \lim_n f_n(x)$ 对每个 X 中的 x 成立，那么 $\int f \,\mathrm{d}\mu = \lim_n \int f_n \,\mathrm{d}\mu$。

♠

证明 根据命题 12.12，首先有 $\int f_1 \,\mathrm{d}\mu \le \int f_2 \,\mathrm{d}\mu \le \cdots$。根据单调收敛准则（定理 1.12），实数列 $\left(\int f_n \,\mathrm{d}\mu\right)$ 极限存在。假设存在 $n_0 \in \mathbb{N}$ 使得 $f_{n_0}(x) > f(x)$，那么存在正数 $\epsilon_0 < f_0(x) - f(x)$，使得 $|f_n(x) - f(x)| \ge \epsilon$ 对所有 $n \ge n_0$ 成立，这与 $\lim_n f_n(x) = f(x)$ 矛盾。因此 $f_n(x) \le f(x)$，根据命题 12.12，$\int f_n \,\mathrm{d}\mu \le \int f \,\mathrm{d}\mu$ 对所有 $n \in \mathbb{N}$ 成立。利用极限的单调性（定理 1.15），$\lim_n \int f_n \,\mathrm{d}\mu \le \int f \,\mathrm{d}\mu$；对于反向不等式，设 $\epsilon \in (0,1)$，我们希望构造 $\mathscr{S}_+$ 中的非减函数列 (g_n)，使得 $g_n \le f_n$ 对每个 n 成立，同时满足 $\lim_n g_n \,\mathrm{d}\mu = (1-\epsilon) \int f \,\mathrm{d}\mu$。

设 $a_1, \cdots, a_m$ 为 f 的非零值，A_i 为 f 取值为 a_i 时对应的 X 的子集。仿照命题 12.8的构造，设 $A_{n,i} = \{x \in A_i | f_n(x) \ge (1-\epsilon) a_i\} \subseteq A_i$。根据例12.1的讨论，$f$ 可测当且仅当 $A_i = \{x \in X | f(x) = a_i\} \in \mathscr{A}$ 对所有 i 成立，那么 $A_{n,i}$ 是有限个 $A_i \in \mathscr{A}$ 的并集，因而 $A_{n,i} \in \mathscr{A}$。由于 (f_n) 非减，对于给定的 i，集合 $\{A_{n,i}\}_{n=1}^\infty$ 也是非减的。对于任意 $x \in A_i$，存在 $n_0 \in \mathbb{N}$ 使得 $|f(x) - f_n(x)| \le \epsilon a_i$ 对所有 $n \ge n_0$ 成立，即 $x \in \cup_{n \ge n_0} A_{n,i} \subseteq \cup_n A_{n,i}$；反之 $A_{n,i} \subseteq A_i$ 显然给出 $\cup_n A_{n,i} \subseteq A_i$。因此 $A_i = \cup_n A_{n,i}$，定义 $g_n = (1-\epsilon) \sum_{i=1}^m a_i \chi_{A_{n,i}} \in \mathscr{S}_+$，那么对于任意 $x \in A_{n,i}$，$g_n(x) \le (1-\epsilon) a_i \le f_n(x)$；对于任意 $x \in X \backslash (\cup_{i=1}^m A_{n,i})$，$g_n(x) = 0 \le f_n(x)$。综上，$g_n \le f_n$ 对每个 n 成立；并且根据命题 11.7，$\{A_{n,i}\}$ 关于 n 递增意味着 $\lim_n \mu(A_{n,i}) = \mu(\cup_n A_{n,i}) = \mu(A_i)$，因此

$$\lim_n \int g_n \,\mathrm{d}\mu = (1-\epsilon) \lim_n \sum_{i=1}^m a_i \mu(A_{n,i}) = (1-\epsilon) \sum_{i=1}^m a_i \mu(A_i) = (1-\epsilon) \int f \,\mathrm{d}\mu. \tag{12.2.4}$$

利用命题 12.12和极限的单调性，$(1-\epsilon) \int f \,\mathrm{d}\mu \le \lim_n \int f_n \,\mathrm{d}\mu \le \int f \,\mathrm{d}\mu$ 对任意 $\epsilon \in (0,1)$ 成立，因此等号成立。

注 该结论是下一节中单调收敛定理 12.1的弱化版，不过现在我们尚且需要利用此结论定义积分。

12.2.2 取值于 $[0,+\infty]$ 函数的积分

定义 12.7 (积分)

设 $(X,\mathscr{A},\mu)$ 为测度空间，对于 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数 f，其 Lebesgue 积分定义为

$$\int f\,\mathrm{d}\mu=\sup\left\{\int g\,\mathrm{d}\mu\,\middle|\,g\in\mathscr{S}_+\text{ and }g\leq f\right\}. \tag{12.2.5}$$

♣

类似地，我们也证明一些关于取值于 $[0,+\infty]$ 的可测函数积分的基本性质。

命题 12.14

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 为 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数，(f_n) 为 $\mathscr{S}_+$ 中使得 $f(x)=\lim_n f_n(x)$ 对 X 中每个 x 成立的非减简单函数列，则 $\int f\,\mathrm{d}\mu=\lim_n\int f_n\,\mathrm{d}\mu$。

♠

证明 作为命题 12.13的推广，仿照其证明首先有 $\lim_n\int f_n\,\mathrm{d}\mu\leq\int f\,\mathrm{d}\mu$；对于反向不等式，设 g 为 $\mathscr{S}_+$ 中任意满足 $g\leq f$ 的 $\mathscr{A}$-可测函数，那么 $(g\wedge f_n)$ 是 $\mathscr{S}_+$ 中的非减函数列。另外对于 X 中的每个 x 以及任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $|f_n(x)-f(x)|<\epsilon$ 对所有 $n\geq n_0$ 成立。若 $g(x)=f(x)$，则 $|(g\wedge f_n)(x)-g(x)|=f(x)-\min\{f(x),f_n(x)\}=\max\{0,f(x)-f_n(x)\}<\epsilon$ 对所有 $n\geq n_0$ 成立；若 $g(x)<f(x)$，则取 $\epsilon<f(x)-g(x)$，使得 $|(g\wedge f_n)(x)-g(x)|=|\max\{0,g(x)-f_n(x)\}|=0<\epsilon$ 对所有 $n\geq n_0$ 成立。因此 $\lim_n(g\wedge f_n)=g$。根据命题 12.13以及极限的单调性（定理 1.15），$\int g\,\mathrm{d}\mu=\lim_n\int(g\wedge f_n)\,\mathrm{d}\mu\leq\lim_n\int f_n\,\mathrm{d}\mu$。

命题 12.15

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 和 g 为 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数，α 为非负实数，则

1. $\int\alpha f\,\mathrm{d}\mu=\alpha\int f\,\mathrm{d}\mu$；
2. $\int(f+g)\,\mathrm{d}\mu=\int f\,\mathrm{d}\mu+\int g\,\mathrm{d}\mu$；
3. 若 $f(x)\leq g(x)$ 对 X 中每个点成立，则 $\int f\,\mathrm{d}\mu\leq\int g\,\mathrm{d}\mu$。

♠

证明 根据命题 12.8，存在 A 上取值于 $[0,+\infty)$ 的非减可测简单函数列 (f_n) 和 (g_n)，使得 $f\lim_n f_n$ 且 $g=\lim_n g_n$。根据命题 12.12，$\int\alpha f_n\,\mathrm{d}\mu=\alpha\int f_n\,\mathrm{d}\mu$ 和 $\int(f_n+g_n)\,\mathrm{d}\mu=\int f_n\,\mathrm{d}\mu+\int g_n\,\mathrm{d}\mu$ 对每个 $n\in\mathbb{N}$ 成立。根据命题 12.14，

$$\int\alpha f\,\mathrm{d}\mu=\lim_n\int\alpha f_n\,\mathrm{d}\mu=\alpha\lim_n\int f_n\,\mathrm{d}\mu=\alpha\int f\,\mathrm{d}\mu\,, \tag{12.2.6}$$

$$\int(f+g)\,\mathrm{d}\mu=\lim_n\int(f_n+g_n)\,\mathrm{d}\mu=\lim_n\left(\int f_n\,\mathrm{d}\mu+\int g_n\,\mathrm{d}\mu\right)=\int f\,\mathrm{d}\mu+\int g\,\mathrm{d}\mu\,. \tag{12.2.7}$$

对于最后一条，利用定义以及 $\left\{\int h\,\mathrm{d}\mu\,|h\in\mathscr{S}_+\text{ and }h\leq f\right\}\subseteq\left\{\int h\,\mathrm{d}\mu\,|h\in\mathscr{S}_+\text{ and }h\leq g\right\}$ 可证。

12.2.3 扩充实值函数的积分

定义 12.8 (积分)

设 $(X,\mathscr{A},\mu)$ 为测度空间，对于 X 上取值于 $[-\infty,+\infty]$ 的 $\mathscr{A}$-可测函数，若 $\int f^+\,\mathrm{d}\mu$ 和 $\int f^-\,\mathrm{d}\mu$ 均有限，则称 f 为 μ-可积的（μ-integrable）或可和的（summable），其 Lebesgue 积分定义为

$$\int f\,\mathrm{d}\mu=\int f^+\,\mathrm{d}\mu-\int f^-\,\mathrm{d}\mu\,. \tag{12.2.8}$$

若 $\int f^+\,\mathrm{d}\mu$ 和 $\int f^-\,\mathrm{d}\mu$ 至少一个有限，则称 f 的极限存在（exist）；在这种情况下，$\int f\,\mathrm{d}\mu$ 的定义不变。对于以上两种情况，有时用 $\int f(x)\mu(\mathrm{d}x)$ 或 $\int f(x)\,\mathrm{d}\mu(x)$ 的写法代替 $\int f\,\mathrm{d}\mu$。

特别地，对于 $X=\mathbb{R}^d$ 且 $\mu=\lambda$ 的情形，通常称上述积分为 Lebesgue 积分（Lebesgue integral），将其记为 $\int f(x)\,\mathrm{d}x$；区间 $[a,b]$ 上的积分常用 $\int_a^b f(x)\,\mathrm{d}x$ 表示。若要强调 Lebesgue 积分，则用 $(L)\int_a^b f(x)\,\mathrm{d}x$ 表示。

♣

注 设 $f:X\to[-\infty,+\infty]$ 是 $\mathscr{A}$-可测的，$A\in\mathscr{A}$。若函数 $f\chi_A$ 是可积的，则称 f 在 A 上可积（*integrable over A*），

将积分 $\int_A f\,\mathrm{d}\mu$ 定义为 $\int f\chi_A\,\mathrm{d}\mu$。一般地，在 $\mu(A^c)=0$ 的情形中，常用 $\int f\,\mathrm{d}\mu$ 代替 $\int_A f\,\mathrm{d}\mu$，并称 f 是可积的。

为了证明形式上类似命题 12.14和12.15的结论，我们首先需要一个引理。

引理 12.1

设 $(X,\mathscr{A},\mu)$ 为测度空间，f_1,f_2 和 g_1,g_2 为 X 上满足 $f_1-f_2=g_1-g_2$ 的非负实值可积函数，则 $\int f_1\,\mathrm{d}\mu-\int f_2\,\mathrm{d}\mu=\int g_1\,\mathrm{d}\mu-\int g_2\,\mathrm{d}\mu$。♡

证明 根据条件，$f_1+g_2=f_2+(f_1-f_2)+g_2=f_2+(g_1-g_2)+g_2=f_2+g_1$；那么根据命题 12.15，

$$\int f_1\,\mathrm{d}\mu+\int g_2\,\mathrm{d}\mu=\int f_2\,\mathrm{d}\mu+\int g_1\,\mathrm{d}\mu\quad\Longrightarrow\quad\int f_1\,\mathrm{d}\mu-\int f_2\,\mathrm{d}\mu=\int g_1\,\mathrm{d}\mu-\int g_2\,\mathrm{d}\mu\,.\tag{12.2.9}$$

命题 12.16

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 和 g 为 X 上的实值可积函数，α 为实数，则

1. αf 和 $f+g$ 可积；
2. $\int\alpha f\,\mathrm{d}\mu=\alpha\int f\,\mathrm{d}\mu$；
3. $\int(f+g)\,\mathrm{d}\mu=\int f\,\mathrm{d}\mu+\int g\,\mathrm{d}\mu$；
4. 若 $f(x)\leq g(x)$ 对 X 中每个点成立，则 $\int f\,\mathrm{d}\mu\leq\int g\,\mathrm{d}\mu$。♠

证明 设 f 的正部与负部分别为 f^+ 和 f^-，因此当 $\alpha=0$ 时，显然 $\int\alpha f\,\mathrm{d}\mu=\alpha\int f\,\mathrm{d}\mu$ 成立；若 $\alpha>0$，则根据命题 12.15，$\int(\alpha f)^+\,\mathrm{d}\mu=\int\alpha f^+\,\mathrm{d}\mu=\alpha\int f^+\,\mathrm{d}\mu<+\infty$，类似地 $\int(\alpha f)^-\,\mathrm{d}\mu=\alpha\int f^-\,\mathrm{d}\mu<+\infty$，因此 αf 可积，且 $\int\alpha f\,\mathrm{d}\mu=\alpha(\int f^+\,\mathrm{d}\mu-\int f^-\,\mathrm{d}\mu)=\alpha\int f\,\mathrm{d}\mu$；若 $\alpha<0$，利用 $\alpha=-(-\alpha)$ 类似可证。

由于 $(f+g)^+(x)=\max\{f(x)+g(x),0\}=[(f(x)+g(x))+|f(x)+g(x)|]/2\leq[(f(x)+g(x))+(|f(x)|+|g(x)|)]/2=\max\{f(x),0\}+\max\{g(x),0\}=f^+(x)+g^+(x)$ 对所有 $x\in X$ 成立，因此 $(f+g)^+\leq f^++g^+$，类似地 $(f+g)^-\leq f^-+g^-$。根据命题 12.15，$\int(f+g)^+\,\mathrm{d}\mu\leq\int(f^++g^+)\,\mathrm{d}\mu=\int f^+\,\mathrm{d}\mu+\int g^+\,\mathrm{d}\mu<+\infty$，类似地 $\int(f+g)^-\,\mathrm{d}\mu<+\infty$，因此 $f+g$ 可积。注意到 $(f+g)^+-(f+g)^-=f+g=(f^+-f^-)+(g^+-g^-)=(f^++g^+)-(f^-+g^-)$，根据引理 12.1，

$$\begin{aligned}\int(f+g)\,\mathrm{d}\mu&=\int(f^++g^+)\,\mathrm{d}\mu-\int(f^-+g^-)\,\mathrm{d}\mu\\&=\int f^+\,\mathrm{d}\mu+\int g^+\,\mathrm{d}\mu-\int f^-\,\mathrm{d}\mu-\int g^-\,\mathrm{d}\mu\\&=\int f\,\mathrm{d}\mu-\int g\,\mathrm{d}\mu\,.\end{aligned}\tag{12.2.10}$$

若 $f(x)\leq g(x)$ 对 X 中的每个 x 成立，那么根据性质 1，$g-f=g+(-1)f$ 是非负可积函数，因此 $\int(g-f)\,\mathrm{d}\mu=\int(g-f)^+\,\mathrm{d}\mu\geq0$；再利用性质 3，$\int g\,\mathrm{d}\mu=\int[f+(g-f)]\,\mathrm{d}\mu=\int f\,\mathrm{d}\mu+\int(g-f)\,\mathrm{d}\mu\geq\int f\,\mathrm{d}\mu$。

例 12.2

1. 若 μ 是有限测度，则 $(X,\mathscr{A},\mu)$ 上的每个有界可测函数都是可积的。这是因为此时存在 $M>0$ 使得 $|f(x)|<M$，即 $f^+(x),f^-(x)<M$ 对所有 $x\in X$ 成立。那么根据命题 12.15，$\int f^+\,\mathrm{d}\mu\leq\int M\,\mathrm{d}\mu=M\mu(X)<+\infty$，类似地 $\int f^-\,\mathrm{d}\mu<+\infty$，因此 f 可积；
2. 特别地，每个 $[a,b]$ 上的有界 Borel 函数是 Lebesgue 可积的，进而每个连续函数都是 Lebesgue 可积的。这是因为此时存在 $M>0$ 使得有界 Borel 函数 f 满足 $|f(x)|<M$，即 $f^+(x),f^-(x)<M$ 对所有 $x\in X$ 成立；并且对于任意满足 $g\leq f^+$ 的简单函数 g，$\int g\,\mathrm{d}\lambda<M\lambda([a,b])=M(b-a)<+\infty$。因此 $\int f^+\,\mathrm{d}\lambda=\sup\left\{\int g\,\mathrm{d}\lambda\,|g\in\mathscr{S}^+\text{ and }g\leq f^+\right\}\leq M(b-a)<+\infty$，类似地 $\int f^-\,\mathrm{d}\lambda<+\infty$，从而 f 可积；
3. 设 $\mathscr{A}$ 是 $\mathbb{N}$ 上所有 $\mathbb{N}$ 的子集构成的 σ-代数，μ 是 $\mathscr{A}$ 上的计数测度（参见例11.2），则 $\mathbb{N}$ 上的非负函数 f 是 μ-可积的当且仅当无穷级数 $\sum_n f(n)$ 收敛。这是因为 $\mathbb{N}$ 上的非负函数 f 可表示为 $f=\sum_n f(n)\chi_{\{n\}}$，而定义为 $f_k(n)=\sum_{k=1}^n f(k)\chi_{\{k\}}$ 的简单函数列 (f_n) 非减，并且 $f=\lim_n f_n$。根据命题 12.14，$\int f\,\mathrm{d}\mu=\lim_n\int f_n\,\mathrm{d}\mu=\lim_n\sum_{k=1}^n f(k)=\sum_n f(n)$。可见对于非负函数 f，由于 $f=f^+$，f 可积当且仅当 $\int f\,\mathrm{d}\mu<+\infty$，即当且仅当 $\sum_n f(n)<+\infty$。一般地，定义在 $\mathbb{N}$ 上的函数 f 可积当且仅当 $\sum_n f(n)$ 绝对收敛；

4. 几乎处处为零的简单可测函数是可积的，其积分为 0。对于几乎处处为零的任意可测函数 f，设其正部和负部分别为 f^+ 和 f^-。根据命题 12.2，对于测度空间 $(X,\mathscr{A},\mu)$，$A=\{x\in X|f(x)>0\}\in\mathscr{A}$。现在定义函数 g：若 $x\notin A$，那么 $g(x)=0$，否则 $g(x)=+\infty$。定义函数列 (g_n) 为 $g_n(x)=n\chi_A$，那么 $g=\lim_n g_n$ 且 $f^+\leq g$。根据命题 12.14和12.15，$\int f^+\,\mathrm{d}\mu\leq\int g\,\mathrm{d}\mu=\lim_n\int g_n\,\mathrm{d}\mu=\lim_n n\mu(A)=0$；而命题 12.15同时表明 $\int f^+\,\mathrm{d}\mu\geq 0$，因此 $\int f^+\,\mathrm{d}\mu=0$；类似地 $\int f^-\,\mathrm{d}\mu$，从而 $\int f\,\mathrm{d}\mu=0$。对于反向结论，可参见推论 12.3。

12.2.4 积分的基本性质

命题 12.17

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 是 X 上取值于 $[-\infty,+\infty]$ 的 $\mathscr{A}$-可测函数，那么 f 可积当且仅当 $|f|$ 可积。若 f 或 $|f|$ 可积，则 $\left|\int f\,\mathrm{d}\mu\right|\leq\int|f|\,\mathrm{d}\mu$。

证明 f 可积当且仅当 f^+ 和 f^- 可积，根据命题 12.15，f 可积时 $\int|f|\,\mathrm{d}\mu=\int f^+\,\mathrm{d}\mu+\int f^-\,\mathrm{d}\mu<+\infty$，即 $|f|$ 可积；反之若 $|f|$ 可积，根据命题 12.15，$\int f^+\,\mathrm{d}\mu\leq\int|f|\,\mathrm{d}\mu<+\infty$，类似地 $\int f^-\,\mathrm{d}\mu<+\infty$，从而 f 可积。此时

$$\left|\int f\,\mathrm{d}\mu\right|=\left|\int f^+\,\mathrm{d}\mu-\int f^-\,\mathrm{d}\mu\right|\leq\int f^+\,\mathrm{d}\mu+\int f^-\,\mathrm{d}\mu=\int(f^++f^-)\,\mathrm{d}\mu=\int|f|\,\mathrm{d}\mu\,.\tag{12.2.11}$$

注 需要注意的是，该结论对于不可测函数不成立；换言之，存在不可积但绝对可积的不可测函数。考虑从 $\mathbb{R}$ 到 $\mathbb{R}$ 的函数 $f=\chi_E-\chi_{[0,1]\setminus E}$，其中 E 为定理 11.2中构造的非 Lebesgue 可测集。假设 f 是 Lebesgue 可测的，根据命题 12.2，$E=\{x\in[0,1]|f(x)=1\}\in\mathscr{M}_{\lambda^*}$，矛盾。因此 f 不是 Lebesgue 可测的；但是 $|f|=\chi_E+\chi_{[0,1]\setminus E}=\chi_{[0,1]}$ 是 Lebesgue 可积的。

命题 12.18

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 和 g 为 X 上取值于 $[-\infty,+\infty]$ 且几乎处处相同的 $\mathscr{A}$-可测函数。若 $\int f\,\mathrm{d}\mu$ 或 $\int g\,\mathrm{d}\mu$ 存在，则两者均存在，且 $\int f\,\mathrm{d}\mu=\int g\,\mathrm{d}\mu$。

证明 首先考虑 f 和 g 均非负的情形。根据命题 12.2，$A=\{x\in X|f(x)\neq g(x)\}\in\mathscr{A}$，那么 $\mu(A)=0$。定义函数

$$h(x)=\begin{cases}+\infty, & \text{if } x\in A,\\ 0, & \text{if } x\notin A\end{cases}\tag{12.2.12}$$

以及非减函数列 (h_n)，其中 $h_n=n\chi_A$ 对每个 $n\in\mathbb{N}$ 成立，则 $h=\lim_n h_n$。根据命题 12.15，$\int h\,\mathrm{d}\mu=\lim_n\int h_n\,\mathrm{d}\mu=\lim_n n\mu(A)=0$。因此结合命题 12.14，$\int f\,\mathrm{d}\mu\leq\int(g+h)\,\mathrm{d}\mu=\int g\,\mathrm{d}\mu+\int h\,\mathrm{d}\mu=\int g\,\mathrm{d}\mu$；对称地 $\int g\,\mathrm{d}\mu\leq\int f\,\mathrm{d}\mu$，从而 $\int f\,\mathrm{d}\mu=\int g\,\mathrm{d}\mu$。

对于任意 f 和 g，f^+ 和 g^+、f^- 和 g^- 在 A^c 上相同，因此积分相等，进而 $\int f\,\mathrm{d}\mu=\int g\,\mathrm{d}\mu$。

命题 12.19

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 为 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数。若 t 是正实数，$A_t=\{x\in X|f(x)\geq t\}$，则 $\mu(A_t)\leq t^{-1}\int_{A_t}f\,\mathrm{d}\mu\leq t^{-1}\int f\,\mathrm{d}\mu$。

证明 利用 $0\leq t\chi_{A_t}\leq f\chi_{A_t}\leq f$ 的关系，根据命题 12.15，$t\mu(A_t)\leq\int_{A_t}f\,\mathrm{d}\mu\leq\int f\,\mathrm{d}\mu$。

推论 12.2

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 是 X 上取值于 $[-\infty,+\infty]$ 的可积函数，则 $\{x\in X|f(x)\neq 0\}$ 在 μ 下 σ-有限。

证明 设 $A_n=\{x\in X||f(x)|\geq 1/n\}$，则根据命题 12.17和12.19，$\mu(A_n)\leq n\int|f|\,\mathrm{d}\mu<+\infty$。另外，容易验证 $\cup_n A_n=\{x\in A||f(x)|>0\}=\{x\in A||f(x)|\neq 0\}$，因此结论得证。

推论 12.3

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 是 X 上取值于 $[-\infty,+\infty]$ 且满足 $\int|f|\,\mathrm{d}\mu=0$ 的 $\mathscr{A}$-可测函数，那么 f 是 μ-几乎处处为零的。 ♡

证明 类似推论 12.2的证明，此时 $\mu(A_n)\leq n\int|f|\,\mathrm{d}\mu=0$ 即 $\mu(A_n)=0$ 对每个正整数 n 成立。根据命题 11.7，$\mu(\{x\in X|f(x)\neq 0\})=\mu(\{x\in X||f(x)|\})=\mu(\cup_n A_n)=\lim_n\mu(A_n)=0$。

推论 12.4

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 为 X 上取值于 $[-\infty,+\infty]$ 且使得 $\int_A f\,\mathrm{d}\mu\geq 0$ 对所有 $A\in\mathscr{A}$ 成立的可积函数，则 $f\geq 0$ 是 μ-几乎处处成立的。 ♡

证明 设 $A=\{x\in X|f(x)<0\}\in\mathscr{A}$，则 $\int f\chi_A\,\mathrm{d}\mu=\int_A f\,\mathrm{d}\mu\leq 0$。根据条件，$\int f\chi_A\,\mathrm{d}\mu=0$；而推论 12.3表明 $f\chi_A$ 几乎处处为零，即 $B=\{x\in X|(f\chi_A)(x)\neq 0\}\in\mathscr{A}$ 的测度为 0。又因为 $A\subseteq B$，故 $\mu(A)=0$，即 $f\geq 0$ 是 μ-几乎处处成立的。

推论 12.5

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 是 X 上取值于 $[-\infty,+\infty]$ 的可积函数，则 $|f(x)|<+\infty$ 对于 X 中几乎每个 x 成立。 ♡

证明 设 $A_n=\{x\in A||f(x)|=n\}$，根据命题 12.17和12.19，$\mu(A_n)\leq n^{-1}\int|f|\,\mathrm{d}\mu<+\infty$。利用极限的单调性（定理 1.15），$0\leq\lim_n\mu(A_n)\leq 0$，因此 $\lim_n\mu(A_n)=0$。根据命题 11.7，$\mu(\{x\in X||f(x)|=+\infty\})=\mu(\cup_n A_n)=\lim_n\mu(A_n)=0$。

12.2.5 复值函数

定义 12.9 (复值函数可积性)

设 $(X,\mathscr{A},\mu)$ 为测度空间，若 X 上复值函数 f 的实部 $\Re(f)$ 和 $\Im(f)$ 分别可积，则称 f 可积（integrable）；若 f 可积，则其积分定义为 $\int f\,\mathrm{d}\mu=\int\Re(f)\,\mathrm{d}\mu+\mathrm{i}\int\Im(f)\,\mathrm{d}\mu$。 ♣

类似命题 12.16，我们有如下结论。

命题 12.20

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 和 g 为 X 上的复制可积函数，α 为复数，则

1. αf 和 $f+g$ 可积；
2. $\alpha f\,\mathrm{d}\mu=\alpha\int f\,\mathrm{d}\mu$；
3. $\int(f+g)\,\mathrm{d}\mu=\int f\,\mathrm{d}\mu+\int g\,\mathrm{d}\mu$。 ♠

证明 设 f 的实部和虚部分别为 f_1 和 f_2，α 的实部和虚部分别为 α_1 和 α_2。根据命题 12.16，$\int\Re(\alpha f)\,\mathrm{d}\mu=\int(\alpha_1 f_1-\alpha_2 f_2)\,\mathrm{d}\mu=\alpha_1\int f_1\,\mathrm{d}\mu-\alpha_2\int f_2\,\mathrm{d}\mu<+\infty$；类似地 $\int\Im(\alpha f)\,\mathrm{d}\mu<+\infty$，因此 αf 可积，此时

$$\begin{aligned}\int\alpha f\,\mathrm{d}\mu&=\left(\alpha_1\int f_1\,\mathrm{d}\mu-\alpha_2\int g_2\,\mathrm{d}\mu\right)+\left(\alpha_1\int f_2\,\mathrm{d}\mu+\alpha_2\int f_1\,\mathrm{d}\mu\right)\\&=(\alpha_1+\mathrm{i}\alpha_2)\left(\int f_1\,\mathrm{d}\mu+\mathrm{i}\int f_2\,\mathrm{d}\mu\right)=\alpha\int f\,\mathrm{d}\mu\,.\end{aligned}\tag{12.2.13}$$

设 g 的实部和虚部分别为 g_1 和 g_2，根据命题 12.16，$\int\Re(f+g)\,\mathrm{d}\mu=\int(f_1+g_1)\,\mathrm{d}\mu=\int f_1\,\mathrm{d}\mu+\int g_1\,\mathrm{d}\mu<+\infty$；类似地 $\int\Im(f+g)\,\mathrm{d}\mu<+\infty$，因此 $f+g$ 可积，此时容易证明 $\int(f+g)\,\mathrm{d}\mu=\int f\,\mathrm{d}\mu+\int g\,\mathrm{d}\mu$。

对于复值可积函数，也有类似命题 12.17控制收敛的结论。

命题 12.21

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 为 X 上关于 $\mathscr{A}$ 和 $\mathscr{B}(\mathbb{C})$ 可测[a]的复值函数，那么 f 可积当且仅当 $|f|$ 可积。若 f 或 $|f|$ 可积，则 $\left|\int f\,\mathrm{d}\mu\right| \leq \int |f|\,\mathrm{d}\mu$。

[a]对于一般拓扑空间上的 Borel σ-代数，我们将放在第13章介绍；这里我们先将 $\mathscr{B}(\mathbb{C})$ 理解为所有复平面上的开集构成的集族。

证明 由于复数的模运算与 $\mathbb{R}^2$ 中内积诱导的范数具有相同的形式，故函数 $g:\mathbb{R}^2\to\mathbb{C},\ (x,y)\mapsto x+\mathrm{i}y$ 是从 $\mathbb{R}^2$ 到 $\mathbb{C}$ 的同胚。那么 f 关于 $\mathscr{A}$ 和 $\mathbb{C}$ 的可测性等价于 f 关于 $\mathscr{A}$ 和 $\mathbb{R}^2$ 的可测性。为了书写简便，后续证明将 f 视为 $\mathbb{R}^2$ 上的函数[3]。这意味着对于任意形如 $B_n=(-\infty,b]\times(-\infty,n]$ 的闭半空间，$f^{-1}(B_n)\in\mathscr{A}$，其中 $b\in\mathbb{R}$ 且 $n\in\mathbb{N}$。因此 $\{x\in X|\Re(f)(x)\leq b\}=\cup_n f^{-1}(B_n)\in\mathscr{A}$，即 $\Re(f)$ 是 $\mathscr{A}$ 可测的；类似地 $\Im(f)$ 也是 $\mathscr{A}$-可测的。

现在利用有理数的稠密性（定理 1.8），对于任意正实数 $t\in\mathbb{R}$，$\{x\in X||f(x)|<t\}$ 可表示为

$$\bigcup_{r\in\mathbb{Q}}\left(\{x\in X|-r<\Re(f)(x)<r\}\cap\{x\in X|-\sqrt{t^2-r^2}<\Im(f)(x)<\sqrt{t^2-r^2}\}\right)\in\mathscr{A}, \tag{12.2.14}$$

其中允许任意集合为空集，并集属于 $\mathscr{A}$ 利用了 $\mathbb{Q}$ 的可数性（命题 1.7）；对于任意 $t\leq 0$，$\{x\in X||f(x)|<t\}=\varnothing\in\mathscr{A}$。因此 $|f|$ 也是 $\mathscr{A}$-可测的。若 f 可积，则根据命题 12.17，$|\Re(f)|$ 和 $|\Im(f)|$ 均可积；再利用命题 12.16，$\int|f|\,\mathrm{d}\mu\leq\int(|\Re(f)|+|\Im(f)|)\,\mathrm{d}\mu=\int|\Re(f)|\,\mathrm{d}\mu+\int|\Im(f)|\,\mathrm{d}\mu<+\infty$，从而 $|f|$ 可积。反之若 $|f|$ 可积，则根据命题 12.16，$\int|\Re(f)|\,\mathrm{d}\mu,\int|\Im(f)|\,\mathrm{d}\mu\leq\int|f|\,\mathrm{d}\mu<+\infty$；根据命题 12.17，$\Re(f)$ 和 $\Im(f)$ 均可积，进而 f 可积。当 f 和 $|f|$ 可积时，记 $w=\int f\,\mathrm{d}\mu\,/\left|\int f\,\mathrm{d}\mu\right|\in\mathbb{C}$，那么根据命题 12.16和12.20，

$$\left|\int f\,\mathrm{d}\mu\right|=w^{-1}\int f\,\mathrm{d}\mu=\int(w^{-1}f)\,\mathrm{d}\mu\leq\int\left|w^{-1}f\right|\mathrm{d}\mu=\int|f|\,\mathrm{d}\mu\,. \tag{12.2.15}$$

12.3 极限定理

这一节内容主要用于对积分的一些特征进行更细致的刻画，与纠缠理论不甚相关；并且这些结论与《数学分析》课程中学习的关于 Riemann 积分的结论大体一致。有鉴于此，读者可根据自身能力与需求阅读本节。

定理 12.1 (单调收敛定理)

设 $(X,\mathscr{A},\mu)$ 为测度空间，f 和 $f_1,f_2,\cdots$ 是 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数，若

$$f_1(x)\leq f_2(x)\leq\cdots\quad\text{and}\quad f(x)=\lim_n f_n(x) \tag{12.3.1}$$

在 X 中几乎每个 x 点成立，则 $\int f\,\mathrm{d}\mu=\lim_n\int f_n\,\mathrm{d}\mu$。

证明 首先假定(12.3.1)式对每个 $x\in X$ 成立，那么仿照命题 12.13的证明，可知 $\lim_n\int f_n\,\mathrm{d}\mu\leq\int f\,\mathrm{d}\mu$；对于反向不等式，根据命题 12.8，对每个 $n\in\mathbb{N}$ 可选择取值于 $[0,+\infty)$ 的可测非减简单函数列 $(g_{n,k})$，使得 $f_n=\lim_k g_{n,k}$。定义函数 $h_n=\max\{g_{1,n},\cdots,g_{n,n}\}$，那么 (h_n) 是取值于 $[0,+\infty)$ 的非零简单函数列，并且 $h_n\leq f_n$，否则存在 $g_{k,n}(x)>f_n(x)\geq f_k(x)$，使得 $\lim_n g_{k,n}(x)=f_k(x)$ 不成立。此外，

$$|h_n(x)-f(x)|=f(x)-\max\{g_{1,n},\cdots,g_{n,n}\}\leq f(x)-g_{k,n}(x)\ \text{for each}\ k=1,\cdots,n. \tag{12.3.2}$$

对于任意正数 ϵ，存在自然数 n_1 和 n_2，分别使得 $|f_n(x)-f(x)|<\epsilon/2$ 对所有 $n\geq n_1$ 成立、$|f_k(x)-g_{k,n}(x)|<\epsilon/2$ 对所有 $n\geq n_2$ 成立。因此在(12.3.2)式中取 $k=n$，则存在自然数 $n_0=\max\{n_1,n_2\}$，使得

$$|h_n(x)-f(x)|\leq|f(x)-g_{n,n}(x)|\leq|f(x)-f_n(x)|+|f_n(x)-g_{n,n}(x)|<\epsilon/2+\epsilon/2=\epsilon \tag{12.3.3}$$

对所有 $n\geq n_0$ 成立，即 $\lim_n h_n=f$。根据命题 12.14和12.15以及收敛的单调性（定理 1.15），$\int f\,\mathrm{d}\mu=\lim_n\int h_n\,\mathrm{d}\mu\leq\lim_n\int f_n\,\mathrm{d}\mu$。

现在设 $N\in\mathscr{A}$ 包含使得(12.3.1)式中至少一个条件不成立的所有点，同时满足 $\mu(N)=0$。那么 $f\chi_{N^c}$ 和函数列 $(f_n\chi_{N^c})$ 满足前述条件，故 $\int_{N^c}f\,\mathrm{d}\mu=\int(f\chi_{N^c})\,\mathrm{d}\mu=\lim_n\int(f_n\chi_{N^c})\,\mathrm{d}\mu=\lim_n\int_{N^c}f_n\,\mathrm{d}\mu$。由于 f 与 $f\chi_{N^c}$、

[3]尽管 $\mathbb{C}$ 与 $\mathbb{R}^2$ 上的函数存在解析性质上的差异，但就函数可测性而言，可以将两者视为等同（参见 [3]）。

f_n 与 $f_n\chi_{N^c}$ 几乎处处相等，根据命题 12.18，$\int f\,\mathrm{d}\mu = \lim_n \int f_n\,\mathrm{d}\mu$。

注 该结论是命题 12.13的推广，将其中 f 和 (f_n) 均简单可测的条件放宽到取值于 $[0,+\infty]$ 且可测；值得注意的是，该定理中函数 f 以及函数列 (f_n) 只假定非负且可测，并不要求其可积。

推论 12.6 (Beppo Levi 定理)

设 $(X,\mathscr{A},\mu)$ 为测度空间，$\sum_{k=1}^{\infty}$ 是无穷级数，其各项为 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数，则

$$\int \sum_{k=1}^{\infty} f_k\,\mathrm{d}\mu = \sum_{k=1}^{\infty} \int f_k\,\mathrm{d}\mu\,. \tag{12.3.4}$$

♡

证明 设函数 $f=\sum_{k=1}^{\infty} f_k$，函数列 (f_n) 中每一项为 $f_n=\sum_{k=1}^{\infty} f_k$，因此 (f_n) 非减且 $f=\lim_n f_n$。利用单调收敛定理 12.1，$\int \sum_{k=1}^{\infty} f_k\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu = \lim_n \int f_n\,\mathrm{d}\mu = \lim_n \sum_{k=1}^{n} \int f_k\,\mathrm{d}\mu = \sum_{k=1}^{\infty} \int f_k\,\mathrm{d}\mu$。

例 12.3 上述推论可用于构造如下一类测度：设 $(X,\mathscr{A},\mu)$ 为测度空间，$f: X\to[0,+\infty]$，定义函数 $\nu:\mathscr{A}\to[0,+\infty]$ 为 $\nu(A)=\int_A f\,\mathrm{d}\mu$，则 ν 是可测空间 $(X,\mathscr{A})$ 上的测度。这是因为对于不相交子集族 $\{A_n\}\subseteq\mathscr{A}$，$\nu(\cup_n A_n) = \int (f\chi_{\cup_n A_n})\,\mathrm{d}\mu = \int f(\sum_n \chi_{A_n})\,\mathrm{d}\mu = \sum_n \nu(A_n)$。该等式还表明，$\nu$ 是有限测度当且仅当 f 是 μ-可积的。

下一结论常用于证明函数是可积的，或者提供积分值的上界。

定理 12.2 (Fatous 引理)

设 $(X,\mathscr{A},\mu)$ 为测度空间，(f_n) 为 X 上取值于 $[0,+\infty]$ 的 $\mathscr{A}$-可测函数列，则

$$\int \liminf_n f_n\,\mathrm{d}\mu \le \liminf_n \int f_n\,\mathrm{d}\mu\,. \tag{12.3.5}$$

♡

证明 对于每个正整数 n，定义 $g_n = \inf\{f_k | k\ge n\}$，那么根据命题 12.4，$g_n$ 是非减且 $\mathscr{A}$-可测的。设 $g = \liminf_n f_n = \lim_n g_n$，则根据单调收敛定理 12.1，$\int g\,\mathrm{d}\mu = \lim_n \int g_n\,\mathrm{d}\mu$。另外，根据命题 12.15，$\int g_n\,\mathrm{d}\mu \le \int f_k\,\mathrm{d}\mu$ 对所有 $k\ge n$ 成立，因此关于 k 取下确界得 $\int g_n\,\mathrm{d}\mu \le \inf\{\int f_k\,\mathrm{d}\mu\,|k\ge n\}$。最终得到 $\int \liminf_n f_n\,\mathrm{d}\mu \le \lim_n \inf\{\int f_k\,\mathrm{d}\mu\,|k\ge n\} = \liminf_n \int f_n\,\mathrm{d}\mu$。

定理 12.3 (Lebesgue 控制收敛定理)

设 $(X,\mathscr{A},\mu)$ 为测度空间，g 为 X 上取值于 $[0,+\infty]$ 的可积函数，f 和 $f_1,f_2,\cdots$ 为 X 上取值于 $[-\infty,+\infty]$ 的 $\mathscr{A}$-可测函数，使得

$$f(x)=\lim_n f_n(x) \quad \text{and} \quad |f_n(x)|\le g(x) \text{ for } n=1,2,\cdots \tag{12.3.6}$$

对 X 中几乎每个 x 成立，则 f 和 $f_1,f_2,\cdots$ 是可积的，并且 $\int f\,\mathrm{d}\mu = \lim_n \int f_n\,\mathrm{d}\mu$。

♡

证明 首先证明函数 f 和函数列 (f_n) 的可积性。根据命题 12.15，$\int |f_n|\,\mathrm{d}\mu \le \int g\,\mathrm{d}\mu < +\infty$；因此根据命题 12.17，$f_n$ 也是可积的。又因为 $||f_n|-|f|| \le |f_n-f|$ 对每个 $n\in\mathbb{N}$ 成立，结合极限的单调性（定理 1.15），$|f(x)| = \lim_n |f_n(x)| \le g(x)$。再次利用命题 12.17，可证明 f 可积。

现在假定 $g(x)<+\infty$ 对 X 中每个 x 成立，那么根据命题 12.3，$(g+f_n)$ 是非负 $\mathscr{A}$-可测函数列，极限的线性性（定理 1.15）要求 $\lim_n(g+f_n) = g+\lim_n f_n = f$。因此根据 Fatou 引理12.2和命题 1.25，

$$\int (g+f)\,\mathrm{d}\mu = \int \lim_n (g+f_n)\,\mathrm{d}\mu = \int \liminf_n (g+f_n)\,\mathrm{d}\mu \le \liminf_n \int (g+f_n)\,\mathrm{d}\mu\,, \tag{12.3.7}$$

因此 $\int f\,\mathrm{d}\mu \le \liminf_n \int f_n\,\mathrm{d}\mu$；类似地，$\int (g-f)\,\mathrm{d}\mu \le \liminf_n \int (g-f_n)\,\mathrm{d}\mu$，从而 $\int f\,\mathrm{d}\mu \ge \limsup_n \int f_n\,\mathrm{d}\mu$。根据命题 1.25，$\liminf_n \int f_n\,\mathrm{d}\mu = \limsup_n \int f_n\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu$ 意味着 $(\int f_n\,\mathrm{d}\mu)$ 收敛到 $\int f\,\mathrm{d}\mu$。

最后设 $N\in\mathscr{A}$ 包含使得(12.3.6)式中至少一个条件不成立的所有点，同时满足 $\mu(N)=0$。类似单调收敛定理 12.1最后一段论述，可证明此时结论也成立。

第 13 章 Haar 测度

这一章会涉及较多一般拓扑空间的性质，如文中有提及，请及时查阅附录第A章。

目录

13.1 局部紧空间 118
13.2 拓扑群 119
13.3 Haar 测度的存在唯一性 121
13.4 Haar 测度的性质 122

13.1 局部紧空间

定义 13.1 (局部紧)

若拓扑空间的每个点都具有闭包紧致的开邻域，则称该拓扑空间是局部紧的（locally compact）。 ♣

例 13.1 *Hausdorff* 空间（*Hausdorff space*）指的是其中的点“由邻域分离”的拓扑空间 $(X,\mathcal{T})$；具体而言，对于任意 $x,y\in X$，分别存在 x 和 y 的邻域 U 和 V，使得 $U\cap V=\varnothing$。局部紧 Hausdorff 空间的例子包括：

- Euclidean 空间 $\mathbb{R}^d$，对于任意实数 $a<b$，有理数的稠密性（定理 1.8）要求存在 $r\in\mathbb{Q}$ 使得 $a<r<b$，因此该空间是 Hausdorff 的；另外根据 Heine-Borel 定理 1.10保证 x 的每个开邻域 $B(x,r)$ 的闭包 $\bar{B}(x,r)$ 紧致；
- 具有离散拓扑（参见例A.1）的空间，在离散度量下 $B(x,1/2)=\{x\}$ 对任意 $x\in X$ 成立，因此其为 Hausdorff 的；另外 $\{x\}=B(x,1/2)=\bar{B}(x,1/2)$，因此对于 $\bar{B}(x,1/2)$ 的任意开覆盖 $\cup_{i\in I}U_i$，存在 i 使得 $x\in U_i$，故 U_i 为 $\bar{B}(x,1/2)$ 的有限子覆盖，从而 $\bar{B}(x,1/2)$ 紧致。综合以上分析，该拓扑空间也是局部紧的；
- 紧 Hausdorff 空间，它显然是局部紧的。

不过例5.2中的 ℓ^2 空间在赋以范数 $(x_n)\mapsto(\sum_n|x_n|^2)^{1/2}$ 诱导的拓扑时不是局部紧的。考虑包含原点的开邻域 $B(0,1)$，其闭包为闭球 $\bar{B}(0,1)$。在该闭球上恰好存在 ℓ^2 的一组 Hilbert 基 $\{e_n\}$，其中 e_n 在第 n 个坐标为 1、在其余坐标为 0。显然每个 $e_n\in\ell^2$，但是 $\|e_n-e_m\|=\sqrt{2}$ 对所有 $n\neq m$ 成立。因此序列 (e_n) 不是 Cauchy 列，进而闭球 $\bar{B}(0,1)$ 不是序列紧的[1]，那么它自然也不可能是紧致的。对于包含原点的任意开球 $B(0,r)$，我们可以采用完全相同的思路，证明其每个开邻域的闭包都不是序列紧的，因而 ℓ^2 不是局部紧的。

我们在第11章定义了 $\mathbb{R}^d$ 上的 Borel σ-代数。对于一般的集合 X，可将此概念推广。

定义 13.2 (Borel σ-代数)

设 X 为 Hausdorff 拓扑空间，则 X 上的 Borel σ-代数（Borel σ-algebra）是 X 的所有开（闭）子集生成的 σ 代数；X 的 Borel 子集（Borel subset）是属于 $\mathscr{B}(X)$ 的集合。 ♣

有了推广的 Borel σ-代数的概念，我们可以进一步定义正则 Borel 测度；它是后续讨论 Haar 测度的基础。

定义 13.3 (正则 Borel 测度)

设 X 为 Hausdorff 拓扑空间，则 X 上的 Borel 测度（Borel measure）是以 $\mathscr{B}(X)$ 为定义域的测度。设 $\mathscr{A}$ 为 X 上使得 $\mathscr{B}(X)\subseteq\mathscr{A}$ 的 σ-代数，则如果 $\mathscr{A}$ 上的正测度 μ 满足

1. 每个 X 的紧子集 K 满足 $\mu(K)<+\infty$；
2. 每个 $\mathscr{A}$ 中的集合 A 满足 $\mu(A)=\inf\{\mu(U)|A\subseteq U \text{ and } U \text{ is open}\}$；

[1]我们在这里并没有详细展开聚点的概念，因此也没有给出度量空间内紧致性与序列紧致性的等价性，详见 [4]。

3. 每个 X 中的开子集 U 满足 $\mu(U)=\sup\{\mu(K)|K\subseteq U \text{ and } K \text{ is compact}\}$。

则称 μ 正则（regular）。X 上的正则 Borel 测度（regular Borel measure）是以 $\mathscr{B}(X)$ 为定义域的正则测度。♣

为了便于后续概念的引入，我们还要先证明一些关于局部紧空间的基本结论。

引理 13.1

设 X 和 Y 为 Hausdorff 拓扑空间，$f:X\to Y$ 为连续函数，那么 f 是 Borel 可测的，即关于 $\mathscr{B}(X)$ 和 $\mathscr{B}(Y)$ 可测。♡

证明 根据命题 1.17，f 连续意味着对于 Y 的任意开集 U，$f^{-1}(U)$ 是开集，因而也在 $\mathscr{B}(X)$ 内。根据命题 12.11，由于 $\mathscr{B}(Y)$ 是由所有开集 U 生成的 σ-代数，故 f 是 Borel 可测的。

定理 13.1 (Riesz 表示定理)

设 X 为局部紧 Hausdorff 空间，I 为 $\mathscr{K}(X)$ 上的正线性泛函，那么存在 X 上唯一的正则 Borel 测度 μ，使得 $I(f)$ 对 $\mathscr{K}(X)$ 中的每个 f 成立。♡

鉴于我们已经了解过内积空间中的 Riesz 表示定理 5.11和8.11，在这里我们直观上相信它是正确的；并且缺少该定理并不影响 Haar 测度的构造，只影响其部分性质，因此略去证明。有兴趣的读者可参见 [3] 第 7 章。

13.2 拓扑群

定义 13.4 (拓扑群)

拓扑群（topological group）是具有群结构和拓扑结构的集合 G，其运算 $(x,y)\mapsto xy$ 和 $x\mapsto x^{-1}$ 均连续。♣

注 首先，此处有序对 (x,y) 所在的集合 $G\times G$ 被视为使用乘积拓扑（参见命题 A.2）得到的拓扑空间；其次，$(x,y)\mapsto xy$ 是从乘积空间 $G\times G$ 到 G 的连续函数，我们要求 $(x,y)\in G\times G$ 的像 $xy\in G$ 关于 x 和 y “联合连续”，而不仅仅是固定 x 时关于 y 连续，或者固定 y 时关于 x 连续。

定义 13.5 (局部紧群和紧群)

局部紧拓扑群（locally compact topological group）或局部紧群（locally compact group）是拓扑局部紧的 Hausdorff 拓扑群。紧群（compact group）是拓扑紧的 Hausdorff 拓扑群。♣

例 13.2

1. 集合 $\mathbb{R}$ 连同其通常的拓扑和作为群运算的加法，构成一个局部紧群。类似地，$\mathbb{R}^d$、$\mathbb{Z}^d$ 也是局部紧群；
2. 非零实数 $\mathbb{R}^*$ 连同其从 $\mathbb{C}$ 继承的子空间拓扑（参见命题 A.1），在作为群运算的乘法下构成一个局部紧群；
3. 设 $\mathbb{T}=\{z\in\mathbb{C}||z|=1\}$，则 $\mathbb{T}$ 连同其从 $\mathbb{C}$ 继承的子空间拓扑，在作为群运算的乘法下构成一个紧群；
4. 有理数集 $\mathbb{Q}$ 连同其从 $\mathbb{R}$ 继承的子空间拓扑，在作为群运算的加法下不是局部紧的；
5. 任意群 G 连同其使得 G 的每个子集为开集的拓扑构成一个局部紧群；G 紧致当且仅当 G 有限。

定义 13.6 (邻域系的基)

设 X 为拓扑空间，x 为 X 的一个元素。若 X 的子集族 $\mathscr{U}$ 满足：

1. $\mathscr{U}$ 的每个元素是 x 的一个开邻域；
2. 对于 x 的每个开邻域 V，存在属于 $\mathscr{U}$ 且包含于 V 内的集合。

则称 $\mathscr{U}$ 为 x 的邻域系[a]的基（base for the family of neighborhoods of x）

[a]邻域以及邻域系的概念参见定义 A.2

♣

设 G 是一个群，若 $a \in G$ 且 $B \subseteq G$，那么乘积 aB 和 Ba 分别定义为 $aB = \{ab|b \in B\}$ 和 $Ba = \{ba|b \in B\}$。类似地，若 B 和 C 是 G 的子集，那么 BC 和 B^{-1} 分别定义为 $BC = \{bc|b \in B \text{ and } c \in C\}$ 和 $B^{-1} = \{b^{-1}|b \in B\}$。若 $B = B^{-1}$，则称集合 B 是对称的（*symmetric*）。因此 B 对称当且仅当条件 $x \in B$ 等价于条件 $x^{-1} \in B$。

命题 13.1

设 G 为拓扑群，e 为 G 的单位元，a 为 G 的任意元素。

1. 函数 $x \mapsto ax, x \mapsto xa$ 和 $x \mapsto x^{-1}$ 是从 G 到 G 上的同胚；
2. 若 $\mathscr{U}$ 是 e 的一组邻域系的基，则 $\{aU|U \in \mathscr{U}\}$ 和 $\{Ua|U \in \mathscr{U}\}$ 均为 a 的邻域系的基；
3. 若 K 和 L 是 G 的紧子集，则 aK, Ka, KL 和 K^{-1} 均为 G 的紧子集。

♠

证明

1. 定义映射 $f_\alpha : G \to G \times G$ 为 $f_a(x) = (a, x)$，则 $f_a = (c_a, \mathrm{id})$，其中常映射 c_a 和单位映射 id 在 G 上均连续。利用乘积拓扑的泛性质（定理 A.2），f_a 连续。结合复合映射的连续性（命题 A.3）与拓扑群的定义，映射 $x \mapsto ax$ 是连续的。类似地，$x \mapsto xa$ 也连续，而 $x \mapsto x^{-1}$ 的连续性源于拓扑群的定义。对于逆映射，其相当于将 $a \in G$ 替换为其逆元 a^{-1}，完全相同的思路可证明逆映射的连续性。因此这些映射为同胚；
2. 对于任意 $U \in \mathscr{U}$，首先 U 是 e 的开邻域，其次存在 $V \in \mathscr{U}$ 使得 $V \subseteq U$。根据上一条结论，$x \mapsto ax$ 是开映射，因此 aU 也是开集，同时 $a = ae \in aU$；也存在 aV 使得 $aV \subseteq aU$。因此 $\{aU|U \in \mathscr{U}\}$ 构成 a 的邻域系的基；类似地，$\{Ua|U \in \mathscr{U}\}$ 也是 a 的邻域系的基；
3. 设 $\cup_{i=1}^n U_i$ 和 $\cup_{j=1}^m V_j$ 分别为 K 和 L 的有限开覆盖，那么对于任意 $x \in aK$，存在指标 i 使得 $a^{-1}x \in U_i$，因此 $x \in aU_i \subseteq \cup_{i=1}^n (aU_i)$。根据第一条结论，$x \mapsto ax$ 是开映射，故 aU_i 也是开集，从而 $\cup_{i=1}^n (aU_i)$ 是 aK 的有限开覆盖，即 aK 紧致。类似地，Ka 和 K^{-1} 也是 G 的紧子集。对于任意 $a \in L$，$\{U_i a\}_{i=1}^n$ 构成 Ka 的有限子覆盖；而 $a \in L \subseteq \cup_{j=1}^m V_j$ 意味着存在指标 j 使得 $a \in V_j$。因此对于任意 $x \in KL$，存在 $a \in L$ 使得 $xa^{-1} \in \cup_{i=1}^m (U_i)$，即 $x \in \cup_{i=1}^m (U_i a) \subseteq \cup_{i=1}^m (U_i V_j) \subseteq \cup_{i,j} U_i V_j$。可见 KL 紧致。

命题 13.2

设 G 为拓扑群，e 为 G 的单位元，U 为 e 的开邻域，那么

1. 存在 e 的开邻域 V 使得 $VV \subseteq U$；
2. 存在 e 的对称开邻域包含在 U 内。

♠

证明

1. 由于映射 $(x, y) \mapsto xy$ 连续，该映射关于 U 的原像 $W = \{(x, y)|xy \in W\}$ 是开集，并且 $(e, e) \in W$。根据命题 A.2，存在 e 的开邻域 $V_1, V_2 \subseteq G$ 使得 $V_1 \times V_2 \subseteq W$. 那么集合 $V = V_1 \cap V_2$ 也是 e 的开邻域，满足 $V \times V \subseteq W$，进而在映射 $(x, y) \mapsto xy$ 下，$VV \subseteq U$；
2. 由于映射 $x \mapsto x^{-1}$ 连续，如果 U 是 e 的开邻域，那么 U^{-1} 也是 e 的开邻域，从而 $V = U \cap U^{-1} \subseteq U$ 是 e 的对称开邻域。

定义 13.7 (一致连续)

设 G 为拓扑群，f 为 G 上的实(复)值函数，那么对于每个正数 ϵ，若存在 e 的开邻域 U 使得 $|f(x) - f(y)| < \epsilon$ 对每对满足 $y \in xU$ 的 $x, y \in G$ 成立，则称 f 左一致连续（left uniformly continuous）；类似地，若存在 e 的开邻域 U 使得 $|f(x) - f(y)| < \epsilon$ 对每对满足 $y \in Ux$ 的 $x, y \in G$ 成立，则称 f 右一致连续。

♣

注 根据命题 13.2，对于 e 的开邻域 U，我们总可以选择其对称开邻域 $V = U \cap U^{-1}$；在这样的开邻域 V 内，$y \in xV$ 等价于 $y^{-1}x = (x^{-1}y)^{-1} \in V$，即 $x \in yV$。因此在这样的定义中，x 和 y 实际上是对称的。

命题 13.3

设 G 为局部紧群，定义 $\mathscr{K}(G)$ 为所有使得 $\mathrm{supp}(f)$ 紧致的连续函数 $f : G \to \mathbb{R}$ 构成的集合，则 $\mathscr{K}(G)$ 中的每个函数都是左一致连续且右一致连续的。

♠

证明 设 $f\in\mathscr{K}(G)$ 且 $K=\operatorname{supp}(f)$ 紧致，ϵ 为任意正数。由于 f 连续，对于 K 中的每个 x，可以如下选择 e 的开邻域 U_x：根据命题 13.1，xU_x 是 x 的开邻域，因此存在 U_x 使得 $|f(x)-f(y)|<\epsilon/2$ 对任意 $y\in xU_x$ 成立。利用命题 13.2，对于这样的开邻域 U_x，可选择开邻域 V_x，使得 $V_xV_x\subseteq U_x$；而命题 13.1表明 xV_x 为 x 的开邻域，因此 $\cup_{x\in K}(xV_x)$ 构成 K 的一个开覆盖。K 的紧致性表明，存在 K 中有限个点 $x_1,\cdots,x_n$ 以及对应的开邻域 $V_1,\cdots,V_n$，使得 $K\subseteq\cup_{i=1}^n(x_iV_i)$。根据命题 13.2，存在包含在 e 的开邻域 $\cap_{i=1}^nV_i$ 内的对称开邻域 V，我们希望证明：若 $x,y\in G$ 且 $y\in xV$，则 $|f(x)-f(y)|<\epsilon$。

对于任意 $x,y\in G$，当 $x\notin K$ 且 $y\notin K$ 时，显然 $f(x)=f(y)=0$，进而 $|f(x)-f(y)|=0<\epsilon$ 自然成立；当 $x\in K$ 时，对于任意 $y\in xV$，由于 $x\in K\subseteq\cup_{i=1}^n(x_iV_i)$，此时存在指标 i 使得 $x\in x_iV_i\subseteq x_iU_{x_i}$，同时 $y\in xV\subseteq(x_iV_i)V_i\subseteq x_iU_{x_i}$。这意味着存在指标 i 使得 $x,y\in x_iU_{x_i}$；回顾 U_x 的构造，可知此时 $|f(x_i)-f(x)|<\epsilon/2$ 且 $|f(x_i)-f(y)|<\epsilon/2$，进而 $|f(x)-f(y)|<\epsilon$；当 $y\in K$ 时，对于任意 $y\in xV$，由于 V 是 e 的对称开邻域，此时该条件等价于 $x\in yV$ 且 $y\in K$。采用与前文相同的证明（仅交换符号 x 和 y），同样得到 $f(y)-f(x)<\epsilon$。因此 f 是左一致连续的，其右一致连续性类似可证。

推论 13.1

设 G 为局部紧群，μ 为 G 上的正则 Borel 测度，$f\in\mathscr{K}(G)$，则函数 $x\mapsto\int f(xy)\mu(\mathrm{d}y)$ 和 $x\mapsto\int f(yx)\mu(\mathrm{d}y)$ 都是连续的。♡

证明 设 $K=\operatorname{supp}(f)$；对于任意 $x_0\in G$，根据局部紧的定义，设 W 为 x_0 的闭包紧致的开邻域。根据命题 13.1以及复合映射的连续性（命题 A.3），对于 W 中的每个 x，映射 $y\mapsto f(yx)$ 是连续的；此外，当 $y\notin K\bar{W}^{-1}$ 时，假设存在 $x\in W$ 使得 $f(yx)\neq0$，那么 $yx\in K$，进而 $y\in Kx^{-1}\subseteq KW^{-1}\subseteq K\bar{W}^{-1}$，矛盾。因此 f 在 $K\bar{W}^{-1}$ 外的函数值为 0。根据命题 13.1，$\bar{W}^{-1}$ 紧致，进而 $K\bar{W}^{-1}$ 紧致，因此正则 Borel 测度要求 $\mu(K\bar{W}^{-1})<+\infty$。设 ϵ 为任意正数，那么存在正数 $\epsilon'<\epsilon/\mu(K\bar{W}^{-1})$。根据命题 13.3，$f$ 左一致连续，因此存在 e 的开邻域 V 和任意满足 $s\in tV$ 的 $s,t\in G$，使得 $|f(s)-f(t)|<\epsilon'$。那么对于每个 $x\in W\cap(x_0V)$ 和 $y\in G$，首先有 $yx\in y(x_0V)=(yx_0)V$；根据命题 12.16和12.17，取 $s=yx,t=yx_0$ 可得

$$\begin{aligned}\left|\int f(yx)\mu(\mathrm{d}y)-\int f(yx_0)\mu(\mathrm{d}y)\right|&\leq\int|f(yx)-f(yx_0)|\chi_{K\bar{W}^{-1}}\mu(\mathrm{d}y)\\&\leq\epsilon'\cdot\mu(K)=\epsilon'\cdot\mu(K\bar{W}^{-1})=\epsilon.\end{aligned}\tag{13.2.1}$$

综上，对于任意 $x_0\in G$，在其闭包紧致的开邻域内存在开邻域 U，使得 $\left|\int f(yx)\mu(\mathrm{d}y)-\int f(yx_0)\mu(\mathrm{d}y)\right|<\epsilon$ 对任意正数 ϵ 和任意 $x\in U$ 成立；可见映射 $x\mapsto\int f(xy)\mu(\mathrm{d}y)$ 连续。类似地，映射 $x\mapsto\int f(yx)\mu(\mathrm{d}y)$ 也连续。

13.3 Haar 测度的存在唯一性

根据引理 13.1和命题 13.1，若 $x\in G$ 且 $A\in\mathscr{B}(G)$，则映射 $y\mapsto xy$ 和 $y\mapsto yx$ 均为同胚，进而这两个映射都是 Borel 可测的，即 $xA\in\mathscr{B}(G)$。由此我们可以给出如下定义。

定义 13.8 (Haar 测度)

设 G 为局部紧群，μ 为 G 上的非零正则 Borel 测度。若 μ 在左平移下不变（invariant under left translation），即 $\mu(xA)=\mu(A)$ 对每个 $x\in G$ 和每个 $A\in\mathscr{B}(G)$ 成立，则称其为左 Haar 测度（left Haar measure），或简称为 Haar 测度。类似地，若 $\mu(Ax)=\mu(A)$ 对每个 $x\in G$ 和每个 $A\in\mathscr{B}(G)$ 成立，则称其为右 Haar 测度。♣

例 13.3

1. 结合命题 11.16和例13.2，$\mathbb{R}$ 或 $\mathbb{R}^d$ 上的 Lebesgue 测度同时是左 Haar 测度和右 Haar 测度；
2. 若 G 是被赋予离散拓扑的群，那么 G 上的计数测度同时是左 Haar 测度和右 Haar 测度；直观上看，G 的任意有限子集在左乘任一群元后不可能增加新元素，而左乘的可逆性要求左乘后得到的新元素不可能彼此相同。特别地，整数群 $\mathbb{Z}$ 上的计数测度是 Haar 测度。

定理 13.2 (左 Haar 测度的存在性)

设 G 为局部紧群，则存在 G 上的左 Haar 测度。 ♡

该结论的构造性证明过于复杂，有兴趣的读者请参见 [3]，这里我们仅介绍一些关于 Haar 测度的基本性质。

引理 13.2

设 G 为局部紧群，μ 为 G 上的左 Haar 测度，那么存在 G 的非空开子集满足 $\mu(U)>0$，并且每个属于 $\mathscr{K}(G)$ 且不恒为零的非负函数 f 满足 $\int f\,\mathrm{d}\mu>0$。 ♡

证明 由于 μ 正则且非零，假设 $\mu(K)=0$ 对所有紧致集成立，那么 μ 的正则性要求开集 G 的测度 $\mu(G)=0$，与 μ 非零矛盾。因此存在紧致集 K 使得 $\mu(K)>0$，设 U 为 G 的非空开子集，则 $\cup_{x\in G}(xU)$ 是 K 的一个开覆盖。那么存在 $x_1,\cdots,x_n\in G$ 使得 $K\subseteq\cup_{i=1}^n(x_iU)$。根据命题 11.5和11.6以及 Haar 测度的左平移不变性，$\mu(K)\leq\mu(\cup_{i=1}^n x_iU)\leq\sum_{i=1}^n\mu(x_iU)=\sum_{i=1}^n\mu(U)=n\mu(U)$，因此 $\mu(U)>0$。现在设 $f\in\mathscr{K}(G)$ 为不恒为零的非负函数，那么存在正数 ϵ 和非空开集 U 使得 $f\geq\epsilon\chi_U$。因此根据命题 12.16，$\int f\,\mathrm{d}\mu\geq\int\epsilon\chi_U\,\mathrm{d}\mu=\epsilon\mu(U)>0$。

定理 13.3 (Haar 测度的唯一性)

设 G 为局部紧群，μ 和 ν 为 G 上的左 Haar 测度，则存在正实数 c 使得 $\nu=c\mu$。 ♡

证明 设 $g\in\mathscr{K}(G)$ 为不恒为零的非负函数，$f\in\mathscr{K}(G)$ 为任意函数。根据引理 13.2，$\int g\,\mathrm{d}\mu>0$，同时可证明（参见 [3]）$\int f\,\mathrm{d}\mu\,/\int g\,\mathrm{d}\mu$ 仅依赖于函数 f 和 g、不依赖于特定的 Haar 测度 μ。因此对于不同的 Haar 测度 μ 和 ν，记正实数 $c=\int g\,\mathrm{d}\mu\,/\int g\,\mathrm{d}\mu$，上述定理表明 $\int f\,\mathrm{d}\nu=c\int f\,\mathrm{d}\mu$。根据 Riesz 表示定理 13.1，上式对每个 $\mathscr{K}(G)$ 中的 f 成立，因此测度唯一，即 $\nu=c\mu$。

13.4 Haar 测度的性质

设 G 为局部紧群，μ 为 G 上的正则 Borel 测度。根据命题 13.1，映射 $x\mapsto x^{-1}$ 是从 G 到 G 的同胚；因此根据引理 13.1，G 的 Borel 子集 A 恰好是那些使得 $A^{-1}\in\mathscr{B}(G)$ 的子集。那么在 $\mathscr{B}(G)$ 上定义函数 $\check{\mu}(A)=\mu(A^{-1})$，容易验证该函数也是正则 Borel 测度。定义函数 $\check{f}:G\to G$ 为 $t\mapsto t^{-1}$，那么我们有如下结论。

命题 13.4

设 G 为局部紧群，μ 为 G 上的左 Haar 测度。若 f 是非负 Borel 函数或 $\check{\mu}$-可积的函数，则 $\int f\,\mathrm{d}\check{\mu}=\int\check{f}\,\mathrm{d}\mu$。 ♠

证明 首先考虑 f 为简单可测函数的情形，不妨设 $f=\sum_{i=1}^m a_i\chi_{A_i}$，那么根据命题 12.12，

$$\int f\,\mathrm{d}\check{\mu}=\sum_{i=1}^m a_i\check{\mu}(A_i)=\sum_{i=1}^m a_i\mu(A_i^{-1})=\sum_{i=1}^m a_i\int\chi_{A_i^{-1}}\,\mathrm{d}\mu=\int\left(\sum_{i=1}^m a_i\chi_{A_i^{-1}}\right)\mathrm{d}\mu\,. \tag{13.4.1}$$

另一方面，根据定义容易逐点验证 $\check{f}=\sum_{i=1}^m a_i\chi_{A_i^{-1}}$，因此 $\int f\,\mathrm{d}\check{\mu}=\int\check{f}\,\mathrm{d}\mu$。

若 f 为 G 上的非负 Borel 函数，根据定义 12.7，对于任意满足 $g\leq f$ 的简单函数 g，考虑任意 $x\in G$，有 $\check{g}(x)=g(x^{-1})\leq f(x^{-1})=\check{f}$，因此 $\check{g}\leq\check{f}$ 也成立。由于 $\int\check{g}\,\mathrm{d}\mu=\int g\,\mathrm{d}\check{\mu}\leq\int f\,\mathrm{d}\check{\mu}$ 对所有满足 $\check{g}\leq\check{f}$ 的简单函数 $\check{g}$ 成立，故 $\int\check{f}\,\mathrm{d}\mu\leq\int f\,\mathrm{d}\check{\mu}$；类似地 $\int f\,\mathrm{d}\check{\mu}\leq\int\check{f}\,\mathrm{d}\mu$，从而两者相等。对于 $\check{\mu}$-可积的函数，其正部 f^+ 和负部 f^- 分别满足 $\int f^+\,\mathrm{d}\check{\mu}=\int\check{f}^+\,\mathrm{d}\mu$ 和 $\int f^-\,\mathrm{d}\check{\mu}=\int\check{f}^-\,\mathrm{d}\mu$，因此 $\int f\,\mathrm{d}\check{\mu}=\int\check{f}\,\mathrm{d}\mu$。

命题 13.5

设 G 为局部紧群，μ 为 G 上的正则 Borel 测度，那么 μ 是左 Haar 测度当且仅当 $\check{\mu}$ 是右 Haar 测度；μ 是右 Haar 测度当且仅当 $\check{\mu}$ 是左 Haar 测度。 ♠

证明 利用恒等式 $(Ax)^{-1}=\{(ax)^{-1}|a\in A\}=\{x^{-1}a^{-1}|a\in A\}=x^{-1}A^{-1}$，$\check{\mu}(Ax)=\mu(x^{-1}A^{-1})=\mu(A^{-1})=\check{\mu}(A)$ 对每个 $x\in G$ 和每个 $A\in\mathscr{B}(G)$ 成立当且仅当 μ 是左 Haar 测度。利用 $\check{\check{\mu}}=\mu$ 可证明另一结论。

推论 13.2

设 G 为局部紧群，则 G 上存在一个右 Haar 测度，并且在常倍数下唯一。 ♡

证明 结合定理 13.2和13.3和命题 13.5可直接证明该结论。

命题 13.6

设 G 为局部紧群，μ 为 G 上的左 Haar 测度，则 μ 有限当且仅当 G 紧致。 ♠

证明 若 G 紧致，则 μ 的正则性要求 $\mu(G) < +\infty$，必要性得证；若 μ 有限，根据引理 13.2的证明，存在 G 的紧子集 K 使得 $\mu(K) > 0$。$\mu(G)$ 的有限性意味着对于那些使得集族 $\{x_iK\}_{i=1}^n$ 内元素不相交的序列 $\{x_i\}_{i=1}^n$，$\mu(\cup_{i=1}^n x_iK) = n\mu(K) \leq n\mu(G) < +\infty$。因此 n 有限，从中选择最大者。那么对于任意 $x \in G$，$xK \cap (\cup_{i=1}^n x_iK) \neq \varnothing$，即 $x \in (\cup i = 1^n x_iK)K^{-1}$。因此 $G \subseteq (\cup_{i=1}^n x_iK)K^{-1} \subseteq G$，即 $G = (\cup_{i=1}^n x_iK)K^{-1}$ 紧致（命题 13.1）。

注 对于每个紧致群 G 上的任意 Haar 测度 μ，常常通过 $\nu = \mu(G)^{-1}\mu$ 定义一个所谓“归一化”的 Haar 测度。

命题 13.7 (模函数)

设 G 为局部紧群，μ 为 G 上的左 Haar 测度。对于每个 $x \in G$ 和 $A \in \mathscr{B}(G)$，定义映射 $\mu_x : G \to \mathbb{R}$ 为 $\mu_x(A) = \mu(Ax)$，则 μ_x 也是 G 上的左 Haar 测度；此时模函数（modular）$\Delta : G \to \mathbb{R}$ 定义为 $\Delta(x) = \mu_x/\mu$；该函数仅由群 G 确定，不依赖于特定的左 Haar 测度 μ。 ♠

证明 根据命题 13.1，映射 $u \mapsto ux$ 是从 G 到 G 的同胚，因此是 G 上 Borel σ-代数之间的双射；因此容易验证 μ_x 为正则 Borel 测度。而 μ 的左平移不变性意味着 $\mu_x(yA) = \mu(yAx) = \mu(Ax) = \mu_x(A)$ 对任意 $x, y \in G$ 和 $A \in \mathscr{B}(G)$ 成立，那么 μ_x 也是 G 上的左 Haar 测度。根据定理 13.3，对每个 x 存在一个正数 $\Delta(x)$，使得 $\mu_x = \Delta(x)\mu$。设 ν 是 G 上的另一个左 Haar 测度，根据定理 13.2，存在正实数 c 使得 $\nu = c\mu$，那么 $\nu_x = c\mu_x = c\Delta(x)\mu = \Delta(x)\nu$ 对每个 $x \in G$ 成立，从而 Δ 不依赖于特定的 Haar 测度。

命题 13.8

设 G 为局部紧群，Δ 为 G 的模函数，则

1. Δ 连续；
2. $\Delta(xy) = \Delta(x)\Delta(y)$ 对每个 $x, y \in G$ 成立。 ♠

证明

1. 设 μ 为 G 上的左 Haar 测度，$f \in \mathscr{K}(G)$ 是不恒为零的非负函数。根据引理 13.2，$\int f \,\mathrm{d}\mu > 0$。类似命题 13.4的思路，可证明 $\int f_x \,\mathrm{d}\mu = \Delta(x) \int f \,\mathrm{d}\mu$。最后利用推论 13.1，可证明映射 $x \mapsto \int f(yx) \,\mathrm{d}\mu = \Delta(x) \int f \,\mathrm{d}\mu$ 连续，即取定 f 后 Δ 连续；
2. 直接验证：$\Delta(xy)\mu(A) = \mu_{xy}(A) = \mu(Axy) = \Delta(y)\mu(Ax) = \Delta(x)\Delta(y)\mu(A)$ 对任意 $A \in \mathscr{B}(G)$ 成立，因此 $\Delta(xy) = \Delta(x)\Delta(y)$。

若局部紧群 G 的模函数对于每个 $x \in G$ 都满足 $\Delta(x) = 1$，则称 G 为幺模的（*unimodular*）。因此局部紧群 G 幺模当且仅当 G 上每个左 Haar 测度都是右 Haar 测度；当然每个局部紧的交换群都是幺模的。

命题 13.9

每个紧群都是幺模的。 ♠

证明 设 G 为紧群，Δ 为其模函数。根据命题 13.8，容易由归纳法证明 $\Delta(x^n) = \Delta(x)^n$ 对每个 $x \in G$ 和每个 $n \in \mathbb{N}$ 成立。根据命题 13.6，G 的紧致性要求 μ 有限。假设 $\Delta(x) > 1$，那么根据 Archimedean 性质（命题 1.19），对于每个 $A \in \mathscr{B}(G)$ 和任意正数 M，存在 $n \in \mathbb{N}$ 使得 $\mu_{x^n}(A) = (\Delta(x))^n\mu(A) > M$，矛盾。因此 $\Delta(x) \leq 1$；另一方面，若存在 $x \in G$ 使得 $0 < \Delta(x) < 1$，则根据命题 13.8，$\Delta(x^{-1}) = \Delta(x)^{-1} > 1$，同样与有限性矛盾。因此只能有 $\Delta(x) = 1$ 对所有 $x \in G$ 成立，即紧群 G 是幺模的。

第四部分

C^*-代数

第 14 章 Banach 空间

本章的目标是引入一类以 $\mathbb{R}^n$ 为模型的向量空间。除去基本的代数性质，我们在 $\mathbb{R}^n$ 上定义了“范数”，这使得我们可度量点之间的距离。我们将范数的概念推广到一般向量空间，并证明这些“赋范空间”的一些性质。我们将看到，所有有限维赋范空间本质上是 $\mathbb{R}^n$ 或 $\mathbb{C}^n$。若空间是无穷维的，则情况会更复杂。

目录

14.1 赋范空间 125
14.2 常见 Banach 空间 126
14.3 有界线性算子的基本性质 129
14.4 范数的等价性 131
14.5 有限维赋范空间 132
14.6 无穷维赋范空间 133
14.7 商空间 134

14.1 赋范空间

我们已经在第5.1.2小节接触了内积空间上的范数，现在将这一概念推广到一般向量空间。

定义 14.1 (范数)

设 X 为向量空间，若映射 $\|\cdot\| : X \to \mathbb{R}$ 满足

1. 正定性：$\|x\| \geq 0$ 对所有 $x \in X$ 成立，并且 $\|x\| = 0$ 当且仅当 $x = 0$；
2. 线性性：$\|\alpha x\| = |\alpha|\|x\|$ 对所有 $x \in X$ 和 $\alpha \in \mathbb{F}$ 成立；
3. 三角不等式：$\|x + y\| \leq \|x\| + \|y\|$ 对所有 $x, y \in X$ 成立；

则称映射 $\|\cdot\|$ 是 X 上的范数，$(X, \|\cdot\|)$ 或 X 是赋范空间（normed space）。

由上述定义可以导出一个简单而十分重要的结果。

命题 14.1 (反向三角不等式)

设 $(X, \|\cdot\|)$ 为赋范空间，那么对于所有 $x, y \in X$，$\|x - y\| \geq |\|x\| - \|y\||$。

证明 三角不等式给出 $\|x\| = \|(x - y) + y\| \leq \|x - y\| + \|y\|$，类似地 $\|y\| \leq \|x - y\| + \|x\|$，综合两式可得结论。

引理 14.1

设 $(X, \|\cdot\|)$ 为赋范空间，并且对于所有 $x, y \in X$ 定义 $d(x, y) = \|x - y\|$，那么 (X, d) 是一个度量空间。

注 该引理可通过验证度量的定义直接证明。从该引理可以看出，相比代数性质，范数更侧重于刻画向量空间的拓扑性质，因此习惯上用字母 X, Y 而非 V, W 表示。

事实上，我们总是为赋范空间赋予一个由范数诱导的度量所生成的拓扑；在此基础上讨论范数的连续性是有意义的。以下定理表明，该拓扑与向量空间结构是相容的。

命题 14.2

给定赋范空间 $(X, \|\cdot\|)$，下列映射关于乘积拓扑（参见命题 A.2）是连续的。

1. 范数的连续性：对于所有 $x \in X$，$x \mapsto \|x\|$；

2. 加法的连续性：对于所有 $x,y\in X$，$(x,y)\mapsto x+y$；
3. 标量乘法的连续性：对于所有 $\alpha\in\mathbb{F}$ 和 $x\in X$，$(\alpha,x)\mapsto \alpha x$。

♠

证明

1. 对于任意正数 ϵ 和 $x\in X$，根据反向三角不等式（命题 14.1），当 $d(x,y)=\|x-y\|<\epsilon$ 时，$|\|x\|-\|y\||\le\|x-y\|<\epsilon$。根据命题 1.16，$\|\cdot\|$ 连续；
2. 对于任意正数 ϵ 和 $x,y\in X$，根据三角不等式，当 $d(x,a)=\|x-a\|<\epsilon/2$ 且 $d(y,b)=\|y-b\|<\epsilon/2$ 时，$d(x+y,a+b)\le\|x-a\|+\|y-b\|<\epsilon$。根据命题 1.16，该映射连续；
3. 对于任意正数 ϵ 和 $x\in X$，根据线性性与三角不等式，当 $|\alpha-\beta|<\epsilon/2\|x\|$ 时，$|\beta|\le|\alpha|+|\alpha-\beta|<|\alpha|+\epsilon/2\|x\|$。那么当 $\|x-y\|<(\|x\|+2|\alpha|/\epsilon)^{-1}$ 时，

$$\|\alpha x-\beta y\|\le\|\alpha x-\beta x\|+\|\beta x+\beta y\|=|\alpha-\beta|\|x\|+|\beta|\|x-y\|<\epsilon. \tag{14.1.1}$$

根据命题 1.16，该映射连续。

在研究度量空间时，一类相当重要的度量空间是完备空间；这在赋范空间中也类似。

定义 14.2

关于范数诱导的度量完备的赋范空间称为 Banach 空间（Banach space）。

♣

Banach 空间最简单的例子就是 $\mathbb{F}^n$。我们就级数性质给出 Banach 空间的一种刻画。

定理 14.1

赋范空间 X 完备当且仅当 X 中每个绝对收敛的级数收敛。

♡

证明 对于充分性，设 $\sum_n u_n$ 为 X 中绝对收敛的级数，记部分和为 $s_n=\sum_{k=1}^n u_k$。由于 $\sum_n\|u_n\|$ 收敛，其部分和 $\sum_{k=1}^n\|u_k\|$ 为 Cauchy 列，即对于任意正数 ϵ，存在自然数 n_0 使得 $\sum_{k=m+1}^n\|u_k\|<\epsilon$ 对所有 $n>m\ge n_0$ 成立。利用三角不等式，对于这样的自然数 n 和 m，$\|s_n-s_m\|\le\sum_{k=m+1}^n\|u_k\|<\epsilon$，因此 (s_n) 也是 Cauchy 列。X 的完备性要求 (s_n) 收敛，即 $\sum_n u_n$ 收敛。

反之对于必要性，设 (x_n) 为 X 中的 Cauchy 列，那么存在自然数 n_k 使得 $|x_n-x_m|<2^{-k}$ 对所有 $n,m\ge n_k$ 成立。不妨取 $n=n_{k+1},m=n_k$，那么 $\left\|x_{n_{k+1}}-x_{n_k}\right\|<2^{-k}$，这意味着 $\sum_{k=1}^m\left\|x_{n_{k+1}}-x_{n_k}\right\|<1-2^{-m}$ 对所有自然数 m 成立。实数列极限的单调性（定理 1.15）表明 $\sum_k\left\|x_{n_{k+1}}-x_{n_k}\right\|\le1<\infty$，因此级数 $\sum_k(x_{n_{k+1}}-x_{n_k})$ 绝对收敛，进而收敛。该级数部分和为 $\sum_{k=1}^m(x_{n_{k+1}}-x_{n_k})=x_{n_{m+1}}-x_{n_1}$，因此存在 $x=\lim_m x_{n_{m+1}}-x_{n_1}\in X$。极限的唯一性（命题 1.10）要求 $\lim_n x_n=x+x_{n_1}\in X$，因此赋范空间 $(X,\|\cdot\|)$ 完备。

14.2 常见 Banach 空间

在本节中，我们从基本不等式和 $\mathbb{F}^n$ 中的范数族入手，给出 Banach 空间的一些例子。

14.2.1 有限维内积空间

对于 $1\le p\le\infty$，我们定义 $x=(x_1,\cdots,x_n)\in\mathbb{F}^n$ 上的 p-范数为

$$\|x\|_p=\begin{cases}(\sum_{i=1}^n|x_i|^p)^{1/p}, & \text{if } 1\le p<\infty,\\ \max\{|x_i||1\le i\le n\}, & \text{if } p=\infty.\end{cases} \tag{14.2.1}$$

为了证明 p-范数确为范数，并且导出这些范数之间的关系，首先我们需要 Young 不等式。

引理 14.2 (Young 不等式)

设 $p,q\in(1,\infty)$ 使得 $1/p+1/q=1$，那么对于所有的 $a,b\ge0$，$ab\le a^p/p+b^q/q$。

♡

证明 $a=0$ 或 $b=0$ 时，该不等式平凡成立，因此假定 $a,b>0$，此时 Jensen 不等式给出

$$\ln(ab)=\frac{1}{p}\ln a^p+\frac{1}{q}\ln b^q\leq\ln\left(\frac{a^p}{p}+\frac{b^q}{q}\right)\quad\Longrightarrow\quad ab\leq\frac{a^p}{p}+\frac{b^q}{q}.\tag{14.2.2}$$

上述引理相当重要；我们称引理中的 q 是与 p 对偶的指数（*exponent dual to* p）或p 的共轭指数（*conjugate exponent*）。若 $p=1$，我们取 $q=\infty$，反之类似。通过 Young 不等式，我们可以得到 Hölder 不等式。

定理 14.2 (Hölder 不等式)

设 $1\leq p\leq\infty$，q 为与之对偶的指数，那么对于所有 $x,y\in\mathbb{F}^n$，$\|xy\|_1\leq\|x\|_p\|y\|_q$，或者具体写为

$$\sum_{i=1}^n|x_iy_i|\leq\left(\sum_{i=1}^n|x_i|^p\right)^{1/p}\left(\sum_{i=1}^n|y_i|^q\right)^{1/q}.\tag{14.2.3}$$

♡

证明 根据引理 14.2，对于每个 $i=1,\cdots,n$，

$$\frac{|x_i|}{(\sum_{i=1}^n|x_i|^p)^{1/p}}\cdot\frac{|y_i|}{(\sum_{i=1}^n|y_i|^q)^{1/q}}\leq\frac{|x_i|^p}{p(\sum_{i=1}^n|x_i|^p)}+\frac{|x_i|^q}{q(\sum_{i=1}^n|y_i|^q)}.\tag{14.2.4}$$

对上式求和得 $\sum_{i=1}^n|x_iy_i|/(\sum_{i=1}^n|x_i|^p)^{1/p}(\sum_{i=1}^n|y_i|^q)^{1/q}\leq 1/p+1/q=1$，结论得证。

现在我们可以得到如下不等式，进而证明 p-范数确为范数。

定理 14.3 (Minkowski 不等式)

设 $p>1$ 且 $x,y\in\mathbb{F}^n$，那么 $\|x+y\|_p\leq\|x\|_p+\|y\|_p$，或者具体写为

$$\left(\sum_{i=1}^n|x_i+y_i|^p\right)^{1/p}\leq\left(\sum_{i=1}^n|x_i|^p\right)^{1/p}+\left(\sum_{i=1}^n|y_i|^p\right)^{1/p}.\tag{14.2.5}$$

♡

证明 p 的共轭指数 $q=p/(p-1)$，因此根据 Hölder 不等式（定理 14.2）和三角不等式，

$$\begin{aligned}\|x+y\|_p^p&=\sum_{i=1}^n|x_i+y_i|^p\leq\sum_{i=1}^n|x_i+y_i|^{p-1}(|x_i|+|y_i|)=\sum_{i=1}^n|x_i+y_i|^{p-1}|x_i|+\sum_{i=1}^n|x_i+y_i|^{p-1}|y_i|\\&\leq\left(\sum_{i=1}^n|x_i+y_i|^{(p-1)q}\right)^{1/q}\left(\sum_{i=1}^n|x_i|^p\right)^{1/p}+\left(\sum_{i=1}^n|x_i+y_i|^{(p-1)q}\right)^{1/q}\left(\sum_{i=1}^n|y_i|^p\right)^{1/p}\\&=\left(\sum_{i=1}^n|x_i+y_i|^p\right)^{1-1/p}(\|x\|_p+\|y\|_p)=\|x+y\|_p^{p-1}(\|x\|_p+\|y\|_p).\end{aligned}\tag{14.2.6}$$

对于非零的 $x,y\in\mathbb{F}^n$，上述不等式可化简为 $\|x+y\|_p\leq\|x\|_p+\|y\|_p$。

综合以上结论，我们有以下定理。

定理 14.4

设 $1\leq p\leq\infty$，则 $\|\cdot\|_p$ 是 $\mathbb{F}^n$ 上的范数。此外，若 $1\leq p\leq q\leq\infty$，那么 $\|x\|_q\leq\|x\|_p\leq n^{(q-p)/qp}\|x\|_q$ 对所有 $x\in\mathbb{F}^n$ 成立，并且该不等式是最优的（等号可成立）。

♡

证明 结合 Minkowski 不等式（定理 14.3），容易验证 $\|\cdot\|$ 为范数。若 $q=\infty$，不妨考虑 $1\leq p<\infty$ 的情形，此时 $|x_i|\leq(\sum_{i=1}^n|x_i|^p)^{1/p}$ 对所有 i 成立，因此 $\|x\|_\infty=\max_i|x_i|\leq\|x\|_p$，等号成立当且仅当存在唯一非零的 x_i；另一方面 $|x_i|\leq\|x\|_\infty$ 对所有 i 成立，因此 $\|x\|_p=(\sum_{i=1}^n|x_i|^p)^{1/p}\leq(n\|x\|_\infty)^{1/p}=n^{1/p}\|x\|_\infty$，等号成立当且仅当所有 $|x_i|$ 相同。该不等式得证。

若 $q<\infty$，已知 $|x_i|\leq\|x\|_p$ 对任意 $1\leq p\leq q$ 成立，因此

$$\left(\frac{\|x\|_q}{\|x\|_p}\right)^q=\sum_{i=1}^n\left(\frac{|x_i|}{\|x\|_p}\right)^q\leq\sum_{i=1}^n\left(\frac{|x_i|}{\|x\|_p}\right)^p=1\quad\Longrightarrow\quad\|x\|_q\leq\|x\|_p,\tag{14.2.7}$$

存在唯一非零 x_i 时等号成立。另一方面，根据 Hölder 不等式（定理 14.2），取 $r=q/p$ 的共轭指数 $s=r/(r-1)=$

$q/(q-p)$，可得

$$\|x\|_p^p=\sum_{i=1}^n|x_i|^p\le\left(\sum_{i=1}^n|x_i|^{pr}\right)^{1/r}\left(\sum_{i=1}^n 1^s\right)^{1/s}=\|x\|_q^p\cdot n^{(q-p)/q}\quad\Longrightarrow\quad\|x\|_p\le n^{(q-p)/qp}\|x\|_q,\tag{14.2.8}$$

当所有 $|x_i|$ 相同时等号成立。

14.2.2 序列空间

类似例5.2中的向量空间 ℓ^2，我们可定义 X 上任意序列 (x_n) 的 p-范数为

$$\|(x_n)\|_p=\begin{cases}(\sum_{n=1}^\infty|x_n|^p)^{1/p}, & \text{if } 1\le p<\infty,\\ \sup\{|x_n||n\in\mathbb{N}\}, & \text{if } p=\infty.\end{cases}\tag{14.2.9}$$

不过这些 p-范数未必对所有序列有限，因此我们以如下方式定义向量空间

$$\ell^p=\ell^p(\mathbb{F})\xlongequal{\Delta}\left\{(x_n)\in\mathbb{F}^{\mathbb{N}}\middle|\|x\|_p<\infty\right\}\quad(1\le p\le\infty),\tag{14.2.10}$$

$$c_0=c_0(\mathbb{F})\xlongequal{\Delta}\left\{(x_n)\in\mathbb{F}^{\mathbb{N}}\middle|\lim_{n\to\infty}|x_n|=0\right\}.\tag{14.2.11}$$

序列的 p-范数具有与 $\mathbb{F}^n$ 中 p-范数类似的性质；事实上，大部分性质可视为上一小节结论到无穷维的推广。

定理 14.5

设 $1\le p\le\infty$，则 $(\ell^p,\|\cdot\|_p)$ 和 $(c_0,\|\cdot\|_\infty)$ 为 Banach 空间。此外，若 $1<p<q<\infty$，则 $\ell_1\subsetneq\ell^p\subsetneq\ell^q\subsetneq c_0\subsetneq\ell^\infty$。最后，若 $1\le p<q\le\infty$，则 $\|(x_n)\|_q\le\|(x_n)\|_p$ 对所有 $x\in\ell^p$ 成立。

证明 当 $1\le p<\infty$ 时，利用 Minkowski 不等式（定理 14.3），对于任意 $(x_n),(y_n)\in\ell^p$，$(\sum_{i=1}^n|x_i+y_i|^p)^{1/p}\le(\sum_{i=1}^n|x_i|^p)^{1/p}+(\sum_{i=1}^n|y_i|^p)^{1/p}\le\|(x_n)\|_p+\|(y_n)\|_p$ 对所有自然数 n 成立，因此根据实数列的单调收敛准则（定理 1.12），$\|(x_n+y_n)\|_p\le\|(x_n)\|_p+\|(y_n)\|_p$；当 $p=\infty$ 时，$|x_n+y_n|\le|x_n|+|y_n|\le\|(x_n)\|_\infty+\|(y_n)\|_\infty$ 对所有自然数 n 成立，因此 $\|(x_n+y_n)\|_\infty\le\|(x_n)\|_\infty+\|(y_n)\|_\infty$。可见 $\|\cdot\|_p$ $(1\le p\le\infty)$ 确为范数，同时 ℓ^p 关于加法封闭，因此 $(\ell^p,\|\cdot\|_p)$ 为赋范空间。而对于 $(x_n),(y_n)\in c_0$，实数列极限的单调性与线性性（定理 1.15）给出 $\lim_n|x_n+y_n|\le\lim_n(|x_n|+|y_n|)=0+0=0$，因此 c_0 也关于加法封闭，从而 $(c_0,\|\cdot\|_\infty)$ 也是赋范空间。

现在证明 $(\ell^p,\|\cdot\|)$ 和 $(c_0,\|\cdot\|_\infty)$ 的完备性。考虑 ℓ^p 内的任意 Cauchy 列 $x_1,x_2,\cdots$，这里 $x_n=(x_{n,1},x_{n,2},\cdots)$。对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $|x_{n,i}-x_{m,i}|\le\|x_n-x_m\|_p<\epsilon$ 对所有 $n,m\ge n_0$ 和每个 i 成立，因此对于每个给定的 i，$(x_{n,i})_{n\in\mathbb{N}}$ 也构成一组 Cauchy 列。利用实数列的完备性（定理 1.14），每条实数列 $(x_{n,i})_{n\in\mathbb{N}}$ 收敛到某个 $y_i\in\mathbb{F}$。记 $y=(y_1,y_2,\cdots)$，希望证明 $y\in\ell^p$。已知 $\sum_{i=1}^r|x_{n,i}-x_{m,i}|^p\le\|x_n-x_m\|_p^p<\epsilon^p$ 对所有 $n,m\ge n_0$ 成立。令 $m\to\infty$，则范数的连续性（命题 14.2）要求 $\sum_{i=1}^r|x_{n,i}-y_i|^p\le\epsilon^p$ 对所有自然数 r 成立。由实数列极限的单调性（定理 1.15），$\|x_n-y\|_p^p\le\epsilon^p$，可见 $y-x_n\in\ell^p$，进而 $y=(y-x_n)+x_n\in\ell^p$；这意味着 Cauchy 列 $x_1,x_2,\cdots$ 收敛到 $y\in\ell^p$，故 $(\ell^p,\|\cdot\|_p)$ 完备。对于 $(c_0,\|\cdot\|_\infty)$ 内的任意 Cauchy 列 $x_1,x_2,\cdots$，类似可证 $|x_{n,i}-x_{m,i}|\le\|x_n-x_m\|_\infty<\epsilon/2$ 对所有 $n,m\ge n_1$ 和每个 i 成立。令 $m\to\infty$，可得 $|x_{n,i}-y_i|\le\epsilon/2$；同时 $\lim_i|x_{n,i}|=0$ 要求存在 $n_2\in\mathbb{N}$，使得 $|x_{n,i}|<\epsilon/2$。因此存在 $n_0=\max\{n_1,n_2\}$，使得 $|y_i|\le|x_{n,i}|+|x_{n,i}-y_i|<\epsilon$ 对每个 i 成立，因此 $y\in c_0$，进而可证明 Cauchy 列 $x_1,x_2,\cdots$ 收敛到 $y\in c_0$，即 $(c_0,\|\cdot\|_\infty)$ 也完备。

最后证明各向量空间的包含关系。根据定理 14.4，$\sup_i|x_i|=\max|x_i|\le(\sum_{i=1}^n|x_i|^q)^{1/q}\le(\sum_{i=1}^n|x_i|^p)^{1/p}$ 对所有自然数 n 成立，因此利用实数列极限的单调性（定理 1.15），$\|(x_n)\|_\infty\le\|(x_n)\|_q\le\|(x_n)\|_p$。这意味着只要较大的范数有界，那么较小的范数必然有界，因此 $\ell^1\subsetneq\ell^p\subsetneq\ell^q\subsetneq\ell^\infty$。对于任意 $(x_n)\in c_0$ 和任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $|x_n|<\epsilon$ 对所有 $n\ge n_0$ 成立，因此 $|x_n|\le\max\{\max_{n\le n_0}|x_n|,\epsilon\}<\infty$ 对所有 $n\in\mathbb{N}$ 成立，进而 $\|(x_n)\|_\infty<\infty$；同时容易构造 $(x_n)=(1-1/n)\in\ell^\infty$ 不在 c_0 内，因此 $c_0\subsetneq\ell^\infty$。而对于任意 $(x_n)\in\ell^q$，通过反证法容易验证 $\lim_n|x_n|=0$，同时可构造 $(x_n)=(1/n^{1/q})\in c_0$ 不在 ℓ^q 内，因此 $\ell^q\subsetneq c_0$。

14.2.3 Lebesgue 空间

这里我们不加证明地给出 Lebesgue 空间（简称为 L_p-空间）的主要性质，可将其视为 ℓ^p-空间的某种连续版本。设 $X \subseteq \mathbb{R}^n$ 是开集（或者仅仅是 Borel 可测集），那么对于可测函数 $u : X \to \mathbb{F}$，定义

$$\|u\|_p = \begin{cases} (\int_X |u(x)|^p \,\mathrm{d}x)^{1/p}, & \text{if } 1 \le p < \infty, \\ \operatorname{ess\,sup}\{|u(x)| | x \in X\}, & \text{if } p = \infty, \end{cases} \tag{14.2.12}$$

其中函数 f 的本质上确界（*essential supremum*）定义为 $\operatorname{ess\,sup} f = \inf\{a \in \mathbb{R} | \mu(\{x | f(x) > a\}) = 0\}$。我们将 X 上所有使得 $\|u\|_p$ 有限的可测函数构成的空间记为 $L_p(X, \mathbb{F}) = L_p(X)$。若两个函数几乎处处相等，则视其等同。

命题 14.3 (Hölder 不等式)

设 $1 \le p \le \infty$，若 $u \in L_p(X)$ 且 $v \in L_q(X)$，则 $\int_X |u||v| \,\mathrm{d}x \le \|u\|_p \|v\|_q$。

此外还有如下事实。

定理 14.6

设 $1 \le p \le \infty$，那么 $(L_p(X), \|\cdot\|_p)$ 是一个 Banach 空间。若 X 具有有限测度且 $1 < p < q < \infty$，则有 $L_\infty(X) \subsetneq L_q(X) \subsetneq L_p(X) \subsetneq L_1(X)$；若 X 具有无穷测度，则 $p \ne q$ 时，在 $L_p(X)$ 与 $L_q(X)$ 之间不存在集合包含关系。

14.3 有界线性算子的基本性质

在介绍有界算子之前，需要先引入有界集的概念

定义 14.3 (有界集)

若存在 $M > 0$，使得赋范空间 X 的子集 $U \subseteq B(0, M)$，则称 U 有界（bounded）。

定义 14.4

设 X 和 Y 为两个赋范空间。

1. 所有从 X 到 Y 的线性算子构成的集合记作 $\operatorname{Hom}(X, Y)$。这里“Hom”源于线性算子时向量空间之间的同态（homomorphism），同时记 $\operatorname{Hom}(X) = \operatorname{Hom}(X, X)$；
2. 连续线性算子集 $\mathcal{L}(X, Y) = \{T \in \operatorname{Hom}(X, Y) | T \text{ continuous}\}$，同时记 $\mathcal{L}(X) = \mathcal{L}(X, X)$；
3. 若 $T \in \operatorname{Hom}(X, Y)$ 将 X 的每个有界子集映射到 Y 的一个有界子集上，则称其有界（bounded）。

类似命题 1.17，我们对连续有界线性算子进行刻画。

命题 14.4

对于 $T \in \operatorname{Hom}(X, Y)$，下列表述等价：

1. $T \in \mathcal{L}(X, Y)$；
2. T 在 $x = 0$ 处连续；
3. T 有界；
4. 对于所有 $x \in X$，存在 $\alpha > 0$ 使得 $\|Tx\|_Y \le \alpha \|x\|_X$。

证明

1. 若表述 1 成立，显然 T 在 $x = 0$ 处连续，表述 2 成立；
2. 若表述 2 成立，由于 T 在 $x = 0$ 处连续，故存在 $\delta > 0$ 使得 $\|x\|_X \le \delta$ 时 $\|Tx\|_Y \le 1$。对于 X 的任意有界子集 U，$M = \sup_{x \in U} \|x\|_X < \infty$，那么对于每个 $x \in U$，$\|(\delta/M)x\|_X = (\delta/M)\|x\|_X \le \delta$。因此

$1 \geq \|T(\delta x/M)\|_X = (\delta/M)\|Tx\|_Y$，即 $TU \subseteq B(0, M/\delta)$，因此 T 有界，表述 3 成立；

3. 若表述 3 成立，那么存在 $\alpha > 0$ 使得 $\|x\|_X \leq 1$ 时 $\|Tx\|_Y \leq \alpha$。对于任意 $x \in X$，$\alpha \geq \|T(x/\|x\|_X)\|_Y = \|Tx\|_Y/\|x\|_X$，即 $\|Tx\|_Y \leq \alpha\|x\|_X$，表述 4 成立；
4. 若表述 4 成立，对于任意 $x, y \in X$ 和任意正数 ϵ，当 $\|x-y\|_X \leq \epsilon/\alpha$ 时，$\|Tx - Ty\|_Y = \|T(x-y)\|_Y \leq \alpha\|x-y\|_X < \epsilon$，因此 $T \in \mathcal{L}(X,Y)$，表述 1 成立。

上述结论中表述 4 对我们有所启发：可以寻找一个充分小的 α 以刻画算子 T 的“长度”。

定义 14.5 (算子范数)

对于 $T \in \mathcal{L}(X,Y)$，定义 $\|T\|_{\mathcal{L}(X,Y)} = \inf\{\alpha > 0 | \|Tx\|_Y \leq \alpha\|x\|_X \text{ for all } x \in X\}$ 为 X 的算子范数。 ♣

注 原则上，我们可以对所有 $T \in \mathrm{Hom}(X,Y)$ 定义算子范数。不过命题 14.4表明，算子范数有界当且仅当 T 连续。

在证明算子范数确为范数前，我们先给出其等价刻画。

命题 14.5

设 $T \in \mathcal{L}(X,Y)$，则 $\|Tx\|_Y \leq \|T\|_{\mathcal{L}(X,Y)}\|x\|_X$ 对所有 $x \in X$ 成立。此外，

$$\|T\|_{\mathcal{L}(X,Y)} = \sup_{x \in X\setminus\{0\}} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X=1} \|Tx\|_Y = \sup_{\|x\|_X<1} \|Tx\|_Y = \sup_{\|x\|_X\leq 1} \|Tx\|_Y. \tag{14.3.1}$$

♠

证明 根据定义，对于任意正数 ϵ，存在 $\alpha > 0$ 使得 $\|T\|_{\mathcal{L}(X,Y)} + \epsilon > \alpha$，因此 $\|Tx\|_Y \leq \alpha\|x\|_X < (\|T\|_{\mathcal{L}(X,Y)} + \epsilon)\|x\|_X$ 对所有 $x \in X$ 成立。利用 ϵ 的任意性，可知 $\|Tx\|_Y \leq \|T\|_{\mathcal{L}(X,Y)}\|x\|_X$ 对所有 $x \in X$ 成立。

对于第一个等号，$\lambda = \sup_{X\setminus\{0\}} \|Tx\|_Y/\|x\|_X \geq \|Tx\|_Y/\|x\|_X$ 对所有 $x \in X$ 成立，因此 $\|T\|_{\mathcal{L}(X,Y)} \leq \lambda$；反之已证明 $\|Tx\|_Y/\|x\|_X \leq \|T\|_{\mathcal{L}(X,Y)}$ 对所有 $x \in X\setminus\{0\}$ 成立，因此 $\lambda = \sup_{x\in X\setminus\{0\}} \|Tx\|_Y/\|x\|_X \leq \|T\|_{\mathcal{L}(X,Y)}$。因此 $\lambda = \|T\|_{\mathcal{L}(X,Y)}$。对于任意 $x \in X\setminus\{0\}$，将 $x/\|x\|_X$ 重新记为 x 可直接证明第二个等号。对于第三个等号，由于 $\|Tx\|_Y \leq \|Tx\|_Y/\|x\|$ 对所有满足 $\|x\|_X < 1$ 的 $x \in X$ 成立，故 $\beta \triangleq \sup_{\|x\|_X<1} \|Tx\|_Y \leq \sup_{\|x\|_X<1} \|T(x/\|x\|_X)\|_Y = \lambda$；另一方面，对于任意正数 ϵ 和 $x \in X$，$\beta \geq \|Tx/(\|x\|_X + \epsilon)\|_Y$，即 $\|Tx\|_Y \leq \beta(\|x\|_X + \epsilon)$，因此 $\|Tx\|_Y \leq \beta\|x\|_X$ 对所有 $x \in X$ 成立，这意味着 $\beta \geq \lambda$，从而 $\beta = \lambda$。最后，首先有 $\lambda = \beta \leq \sup_{\|x\|_X\leq 1} \|Tx\|_Y$；对于反向不等式，类似有 $\sup_{\|x\|_X\leq 1} \|Tx\|_Y \leq \sup_{x\in X\setminus\{0\}} \|Tx\|_Y/\|x\|_X = \lambda$，得证。

有了上述结果，现在我们可证明 $\|\cdot\|_{\mathcal{L}(X,Y)}$ 是范数。

命题 14.6

空间 $(\mathcal{L}(X,Y), \|\cdot\|_{\mathcal{L}(X,Y)})$ 是赋范空间。 ♠

证明 对于任意 $T \in \mathcal{L}(X,Y)$ 和任意正数 ϵ，根据定义，存在 $\alpha > 0$ 使得 $\|T\|_{\mathcal{L}(X,Y)} + \epsilon > \alpha > 0$，因此 ϵ 的任意性要求 $\|T\|_{\mathcal{L}(X,Y)} \geq 0$。当 $\|T\|_{\mathcal{L}(X,Y)} = 0$ 时，根据命题 14.5，$\|Tx\|_Y = 0$ 对所有 $x \in X$ 成立，因此 $T \equiv 0$，正定性得证；对于任意 $\lambda \in \mathbb{F}$，将 λT 代入命题 14.5中任一表达式可证明该结论，线性性得证；对于任意 $S \in \mathcal{L}(X,Y)$，$\|(S+T)x\|_Y = \|Sx + Tx\|_Y \leq \|Sx\|_Y + \|Tx\|_Y$ 对任意满足 $\|x\|_X = 1$ 的 $x \in X$ 成立，因此 $\|S+T\|_{\mathcal{L}(X,Y)} \leq \sup_{\|x\|_X=1}(\|Sx\|_Y + \|Tx\|_Y) \leq \|S\|_{\mathcal{L}(X,Y)} + \|T\|_{\mathcal{L}(X,Y)}$，三角不等式得证。

定理 14.7

若 Y 是一个 Banach 空间，则 $\mathcal{L}(X,Y)$ 是关于算子范数的 Banach 空间。 ♡

证明 设 (T_n) 是 $\mathcal{L}(X,Y)$ 内的 Cauchy 列，那么对于任意正数 ϵ 和 $x \in X$，存在 $n_0 \in \mathbb{N}$，使得 $\|T_n - T_m\|_{\mathcal{L}(X,Y)} < \epsilon$ 对所有 $n, m \geq n_0$ 成立。根据命题 14.5，对于任意给定的 $x \in X$，$\|T_n x - T_m x\|_Y \leq \|T_n - T_m\|_{\mathcal{L}(X,Y)}\|x\|_X < \epsilon\|x\|_X$，因此 $(T_n x)$ 是 Banach 空间 Y 内的 Cauchy 列，其收敛到 Y 内某个向量。根据定理 3.1，可以通过指定 $T \in \mathrm{Hom}(X,Y)$ 在 X 的每个基向量上的值唯一确定 T，并通过线性延拓使得 $Tx = \lim_n T_n x$ 对所有 $x \in X$ 成立。下面证明 $T \in \mathcal{L}(X,Y)$。利用范数的连续性（命题 14.2）与极限的单调性（定理 1.15），$\|T_n x - Tx\| =$

$\lim_m \|T_n x - T_m x\| \le \epsilon \|x\|_X$ 对所有 $x \in X$ 和 $n \ge n_0$ 成立。根据命题 14.4，$T - T_n \in \mathcal{L}(X,Y)$。命题 3.1表明，$\mathrm{Hom}(X,Y)$ 是一个向量空间，而复合映射的连续性（简单验证或参见命题 A.3）要求 $\mathcal{L}(X,Y)$ 为 $\mathrm{Hom}(X,Y)$ 的一个子空间。因此 $T = (T - T_n) + T_n \in \mathcal{L}(X,Y)$。

14.4 范数的等价性

根据所考察的特定问题，采用不同的范数处理可能会带来方便。这里先介绍一些定义。

定义 14.6 (等价范数)

设 X 为向量空间，$\|\cdot\|_1$ 和 $\|\cdot\|_2$ 是 X 上的两个范数。

1. 若存在常数 $C > 0$ 使得 $\|x\|_2 \le C\|x\|_1$ 对所有 $x \in X$ 成立，则称 $\|\cdot\|_1$ 强于 $\|\cdot\|_2$，或者 $\|\cdot\|_2$ 弱于 $\|\cdot\|_1$；
2. 若存在常数 $C_1, C_2 > 0$ 使得 $C_1\|x\|_1 \le \|x\|_2 \le C_2\|x\|_1$ 对所有 $x \in X$ 成立，则称两个范数等价。

♣

注 我们已经知道任意范数 $\|\cdot\|$ 总可以定义一个度量 $d(x,y) = \|x - y\|$，并且定理 1.3表明任意度量总可以诱导一个拓扑。因此范数的强弱性是与拓扑的强弱性（参见定义 A.4）直接相关的一个概念。此外，容易验证等价范数也是一种等价关系，这种等价性体现在两者生成相同的拓扑。我们稍后将证明这一点。

例 14.1

1. 根据定理 14.4，$\mathbb{F}^n$ 上所有的 p-范数等价；
2. 根据定理 14.5，当 $p \le q$ 时，ℓ^p 上的 $\|\cdot\|_p$ 范数强于 ℓ^q 上的 $\|\cdot\|_q$ 范数；此时不存在等价范数。

命题 14.7

1. 范数等价是一种等价关系；
2. 若 $\|\cdot\|_1$ 强于 $\|\cdot\|_2$，则 X 上由 $\|\cdot\|_1$ 诱导的拓扑 $\mathcal{T}_1$ 强于 $\|\cdot\|_2$ 诱导的拓扑 $\mathcal{T}_2$，即开集族 $\mathcal{T}_2 \subseteq \mathcal{T}_1$。特别地，等价范数生成 X 上的相同拓扑；
3. 对于赋范空间 $X_1 = (X, \|\cdot\|_1)$ 和 $X_2 = (X, \|\cdot\|_2)$，$\|\cdot\|_1$ 强于 $\|\cdot\|_2$ 当且仅当单位映射 $\mathrm{id} : X_1 \to X_2, x \mapsto x$ 是有界线性算子，即 $\mathrm{id} \in \mathcal{L}(X_1, X_2)$。特别地，该范数等价当且仅当 $\mathrm{id}^{-1} = \mathrm{id} \in \mathcal{L}(X_2, X_1)$。

♠

证明

1. 自反性是显然的；若存在 $C_1, C_2 > 0$ 使得 $C_1\|x\|_1 \le \|x\|_2 \le C_2\|x\|_1$ 对所有 $x \in X$ 成立，那么 $C_2^{-1}\|x\|_2 \le \|x\|_1 \le C_1^{-1}\|x\|_2$，对称性成立；若存在 $C_1, C_2, C_1', C_2' > 0$ 使得 $C_1\|x\|_1 \le \|x\|_2 \le C_2\|x\|_1$ 且 $C_1'\|x\|_2 \le \|x\|_3 \le C_2'\|x\|_2$，那么 $C_1C_1'\|x\|_1 \le \|x\|_3 \le C_2C_2'\|x\|_1$，传递性成立；
2. 对于任意开集 $U \in \mathcal{T}_2$，根据定义，对于任意 $x \in U$，存在 $B_2(x,r) = \{y \in U \mid \|x - y\|_2 < r\} \subseteq U$。那么对于任意 $y \in B_1(x, r/C) = \{y \in U \mid \|x - y\|_1 < r/C\}$，$\|x - y\|_2 \le C\|x - y\|_1 < r$，即 $B_1(x, r/C) \subseteq B_2(x,r) \subseteq U$。可见 $U \in \mathcal{T}_1$，即 $\mathcal{T}_2 \subseteq \mathcal{T}_1$；
3. 根据命题 14.4，id 有界当且仅当存在 $C > 0$ 使得 $\alpha\|x\|_2 \ge \|\mathrm{id}(x)\|_1 = \|x\|_1$ 对所有 $x \in X$ 成立，因此当且仅当 $\|\cdot\|_1$ 强于 $\|\cdot\|_2$。

引理 14.3

若 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 为等价范数，则 $X_1 = (X, \|\cdot\|_1)$ 完备当且仅当 $X_2 = (X, \|\cdot\|_2)$ 完备。

♡

证明 若 X_1 完备，那么对于 X_2 上的任意 Cauchy 列 (x_n) 和任意正数 ϵ，存在 $C > 0$ 和 $n_0 \in \mathbb{N}$ 使得 $\|x_n - x_m\|_1 \le C\|x_n - x_m\|_2 < C \cdot (\epsilon/C) = \epsilon$ 对所有 $n, m \ge n_0$ 成立。因此 (x_n) 也是 X_1 上的 Cauchy 列，进而收敛到某个 $x \in X$，即 X_2 也完备；反之类似。

定义 14.7 (图范数)

设 X, Y 为赋范空间，$T \in \mathrm{Hom}(X,Y)$，则称 $\|u\|_T = \|u\|_X + \|Tu\|_Y$ 是与 T 有关的图范数（graph norm）。

♣

注 我们对术语“图范数”稍作解释。利用 T 的线性性，容易验证 $\|\cdot\|_T$ 是 X 上的范数；它被称为“图范数”是因为它确实是 T 的图（*graph*）$\mathrm{graph}(T)=\{(u,Tu)|u\in X\}\subseteq X\times Y$ 上的范数。这使其成为一个赋范空间。

命题 14.8

设 $T\in\mathrm{Hom}(X,Y)$，将 X 上的图范数记为 $\|\cdot\|_T$，则 $T\in\mathcal{L}(X,Y)$ 当且仅当 $\|\cdot\|_T$ 等价于 $\|\cdot\|_X$。♠

证明 由于 $\|u\|_X\leq\|u\|_T$ 显然成立，后续证明只考虑 $\|\cdot\|_X$ 是否强于 $\|\cdot\|_T$。对于充分性，命题 14.4表明存在 $\alpha>0$ 使得 $\|Tu\|_Y\leq\alpha\|u\|_X$ 对所有 $u\in X$ 成立，因此 $\|u\|_T\leq(1+\alpha)\|u\|_X$，故 $\|\cdot\|_T$ 等价于 $\|\cdot\|_X$；反之对于必要性，存在 $C>0$ 使得 $\|u\|_T=\leq C\|u\|_X$ 对所有 $u\in X$ 成立，即 $\|Tu\|_Y\leq(C-1)\|u\|_X$。这要求 $C>1$（$C=1$ 时 T 为零映射），从而 T 有界。根据命题 14.4，$T\in\mathcal{L}(X,Y)$。

14.5 有限维赋范空间

定理 14.8

设 X 是有限维向量空间，那么 X 上所有范数都是等价的。此外，X 关于每个范数均完备。♡

证明 设 $\dim(X)=n$，那么存在同构（推论 3.2）$T:\mathbb{F}^n\to X$，使得 $Te_i=b_i$，其中 $\mathcal{B}=(b_1,\cdots,b_n)$ 是 X 的一组有序基，而 e_i 是 $\mathbb{F}^n$ 的标准基；再通过线性延拓将 T 的定义域延拓到 X。设 $\|\cdot\|_X$ 是 X 上的范数，那么对于任意 $x=(x_1,\cdots,x_n)\in\mathbb{F}^n$，Cauchy-Schwarz 不等式（命题 5.2）给出

$$\|Tx\|_X=\left\|T\sum_{i=1}^n x_ie_i\right\|_X=\left\|\sum_{i=1}^n x_ib_i\right\|_X\leq\sum_{i=1}^n|x_i|\|b_i\|_X\leq\left(\sum_{i=1}^n\|b_i\|_X^2\right)^{1/2}\left(\sum_{i=1}^n|x_i|^2\right).\tag{14.5.1}$$

记 $C=(\sum_{i=1}^n\|b_i\|_X^2)^{1/2}$，则上式可写为 $\|Tx\|_X\leq C\|x\|_2$。根据命题 14.4，$T\in\mathcal{L}(\mathbb{F}^n,X)$。利用范数的连续性（命题 14.2）和复合映射的连续性（命题 A.3），映射 $x\mapsto\|Tx\|_X$ 是从 $\mathbb{F}^n$ 到 $\mathbb{R}$ 的连续映射。特别地，该映射在单位球面 $S=\{x\in\mathbb{F}^n|\|x\|_2=1\}$ 上连续。由于 S 是 $\mathbb{F}^n$ 的紧子集，根据极值定理 1.5，存在 $y\in\mathbb{F}^n$ 使得 $\|Ty\|_X\leq\|Tx\|_X$ 对所有 $x\in S$ 成立。假设 $\|Ty\|_X=0$，那么范数的正定性要求 $Ty=0$，从而 Ty 在 $\mathcal{B}$ 下的展开系数 y_i 均为零，即 $y=0$，这与 $\|y\|_2=1$ 矛盾。故 $c=\|Ty\|_X>0$，那么对于任意非零的 $x\in\mathbb{F}^n$，

$$\left\|T\frac{x}{\|x\|_2}\right\|_X\geq\|Ty\|_X=c\quad\Longrightarrow\quad\|Tx\|_X\geq c\|x\|_2\quad\Longrightarrow\quad\left\|T^{-1}(Tx)\right\|_2\leq c^{-1}\|Tx\|_X.\tag{14.5.2}$$

对于任意 $u\in X$，存在 $x=T^{-1}u\in\mathbb{F}^n$，因此根据命题 14.4，上式表明 $T^{-1}\in\mathcal{L}(X,\mathbb{F}^n)$。根据命题 14.8，$X$ 上的任意范数 $\|\cdot\|_X$ 等价于 $\|\cdot\|_{T^{-1}}$，而 T 仅取决于向量空间 $\mathbb{F}^n$ 和 X 的代数性质而不依赖于具体的范数。因此对于给定的 X，所有范数 $\|\cdot\|_X$ 彼此等价。

现在证明其完备性。设 (x_n) 为 X 内的 Cauchy 列，那么对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $\|x_n-x_m\|_X\leq c\epsilon$ 对所有 $n,m\geq n_0$ 成立。根据(14.5.2)式，$\left\|T^{-1}x_n-T^{-1}x_m\right\|_2\leq c^{-1}\|x_n-x_m\|_X<\epsilon$，因此 $(T^{-1}x_n)$ 也构成 $\mathbb{F}^n$ 的 Cauchy 列。利用 $\mathbb{F}^n$ 的完备性（请读者结合定理 1.14，仿照定理 14.5的方法自行证明），类似定理 14.7的分析可知，存在 $x\in X$ 使得 $\lim_n(T^{-1}x_n)=T^{-1}x$。$T\in\mathcal{L}(\mathbb{F}^n,X)$ 表明 $x_n\to x\in X$，因此 $(X,\|\cdot\|_X)$ 完备。

上述定理存在一些有用的结果。首先是与任意赋范空间的有限维子空间的结果。

推论 14.1

赋范空间 X 的每个有限维子空间均是 X 内的完备闭集。♡

证明 若 S 是 X 的有限维子空间，那么 S 的封闭性保证其为范数继承自 X 的赋范空间。根据定理 14.8，S 是完备的。对于 S 内任意收敛到 $x\in X$ 的序列 (x_n)，根据命题 1.11，(x_n) 也是 S 内的 Cauchy 列，因而收敛到某个 $x'\in X$。由极限的唯一性（命题 1.10），$x=x'\in S$，因此 S 也是闭集。

下一个结果表明，任意有限维赋范空间上的线性算子是连续的。

推论 14.2

设 X,Y 为赋范空间且 $\dim(X)<\infty$，若 $T:X\to Y$ 是线性的，则 T 是有界的。

♡

证明 对于任意 $u\in X$，根据图范数定义 14.7后的讨论，只要 $T\in\mathrm{Hom}(X,Y)$ 就能定义图范数 $\|u\|_T=\|u\|_X+\|Tu\|_Y$。根据定理 14.8，$\|\cdot\|_T$ 与 $\|\cdot\|_X$ 等价；再根据命题 14.8，$T\in\mathcal{L}(X,Y)$。

最后证明 Heine-Borel 定理 1.10的在一般的有限维赋范空间内的推广；而在下一节中，我们将证明反向结论也为真。这给出了有限维赋范空间的一个拓扑刻画。

推论 14.3 (一般 Heine-Borel 定理)

设 X 为有限维赋范空间，则 $A\subseteq X$ 紧致当且仅当 A 为有界闭集。

♡

证明 延续定理 14.8的证明，已知对于任意 n 维赋范空间 X，同构 $T\in\mathcal{L}(\mathbb{F}^n,X)$ 且 $T^{-1}\in\mathcal{L}(X,\mathbb{F}^n)$。若 $A\subseteq X$ 紧致，则根据命题 1.18，$T^{-1}(A)$ 紧致，因此 Heine-Borel 定理 1.10表明 $T^{-1}A$ 是有界闭集。根据命题 1.17，$A=(T^{-1})^{-1}(T^{-1}A)$ 是闭集；根据命题 14.4，$A=T(T^{-1}A)$ 是有界的。因此 A 也是有界闭集；反之类似。

14.6 无穷维赋范空间

定理 14.9

设 X 为赋范空间，S 是 X 的真闭子空间，则对于每个 $\epsilon\in(0,1)$，存在满足 $\|x_\epsilon\|=1$ 的 $x_\epsilon\in X$，使得

$$\mathrm{dist}(x_\epsilon,S)=\inf_{x\in S}\|x-x_\epsilon\|\geq 1-\epsilon. \tag{14.6.1}$$

♡

证明 由于 S 是 X 的真闭子空间，那么存在 $x\in X\backslash S$。假设 $\mathrm{dist}(x,S)=0$，那么对于任意正数 ϵ，存在 $y\in S$ 使得 $\epsilon>\|x-y\|$，即 $B(x,\epsilon)\cap S\neq\varnothing$；这意味着 $x\in\bar{S}=S$，但这与 $x\in X\backslash S$ 矛盾。因此设 $\alpha=\mathrm{dist}(x,S)>0$。任取 $\epsilon\in(0,1)$，那么存在 $s_\epsilon\in S$ 使得 $\alpha/(1-\epsilon)>\|x-s_\epsilon\|$。定义 $x_\epsilon=(x-s_\epsilon)/\|x-s_\epsilon\|$，那么对于每个 $s\in S$，

$$\|x_\epsilon-s\|=\left\|\frac{x-s_\epsilon}{\|x-s_\epsilon\|}-s\right\|=\frac{1}{\|x-s_\epsilon\|}\|x-s_\epsilon-\|x-s_\epsilon\|s\|>\frac{1-\epsilon}{\alpha}\|x-(s_\epsilon+\|x-s_\epsilon\|s)\|. \tag{14.6.2}$$

利用 S 的封闭性，$s_\epsilon+\|x-s_\epsilon\|s\in S$，因此 $\|x_\epsilon-s\|\geq(1-\epsilon)\mathrm{dist}(x,S)/\alpha=1-\epsilon$ 对所有 $s\in S$ 成立，进而 $\mathrm{dist}(x_\epsilon,S)\geq 1-\epsilon$。

推论 14.4

设 X 具有闭子空间 S_i $(i\in\mathbb{N})$，如果 $S_1\subsetneq S_2\subsetneq\cdots\subsetneq S_n\subsetneq S_{n+1}$ 对所有 $n\in\mathbb{N}$ 成立，那么存在范数为 1 的 $s_n\in S_n$，使得 $\mathrm{dist}(s_n,S_{n-1})\geq 1/2$ 对所有 $n\in\mathbb{N}$ 成立；类似地，如果 $S_1\supsetneq S_2\supsetneq\cdots\supsetneq S_n\supsetneq S_{n+1}$ 对所有 $n\in\mathbb{N}$ 成立，那么存在范数为 1 的 $s_n\in S_n$，使得 $\mathrm{dist}(s_n,S_{n+1})\geq 1/2$ 对所有 $n\in\mathbb{N}$ 成立。

♡

有了这一推论，我可可以给出有限维和无穷维空间的刻画。

定理 14.10

赋范空间 X 是有限维的当且仅当单位球面 $S=\{x\in X|\|x\|=1\}$ 是紧致的。

♡

证明 若 $\dim(X)=n<\infty$，显然 S 是有界集。对于 S 上任意收敛到 $x\in X$ 的序列 (x_n)，范数的连续性（命题 14.2）给出 $\|x\|=\|\lim_n x_n\|=\lim_n\|x_n\|=1$，因此 $x\in S$，从而 S 是闭集。根据一般 Heine-Borel 定理（推论 14.3），S 紧致；反之若 $\dim(X)=\infty$，由于 $\aleph_0$ 是最小无穷势，因此 X 内至少存在一组可数线性无关集 $\{e_n|n\in\mathbb{N}\}$。设子空间 $S_n=\mathrm{span}\{e_k|k=1,\cdots,n\}$，那么显然 $\dim(S_n)=n$。根据推论 14.1，每个 S_n 都是完备闭集；又根据推论 14.4，$S_1\subsetneq S_2\subsetneq\cdots\subsetneq S_n\subsetneq S_{n+1}$ 对每个 $n\in\mathbb{N}$ 成立，因此存在 $s_n\in S_n$，使得 $\mathrm{dist}(s_n,S_{n-1})\geq 1/2$ 对所有 $n\in\mathbb{N}$ 成立。后续分析与例13.1类似，因为 (s_n) 不是序列紧的，因此 S 不是紧致的。序列紧的定义及其与紧致性的等价性参见 [4]。

14.7 商空间

我们已经在第4.1节了解了向量空间 X 的某个子空间 S 对应的商空间 X/S 定义。先我们希望引入定义在商空间上的范数，使其称为一个赋范空间。结合定理 14.9，这在一定条件下总是可以做到的。

命题 14.9

设 X 为赋范空间，S 是 X 的一个子空间，那么 X/S 是具有范数 $\|\cdot\|_{X/S} : X/S \to \mathbb{R},\ x+S \mapsto \inf_{z\in S}\|x-z\|_X$ 的赋范空间当且仅当 S 是闭集。

♠

证明 首先验证函数 $\|\cdot\|_{X/S}$ 的线性性与三角不等式。对于前者，

$$\|\alpha(x+S)\|_{X/S} = \|\alpha x+S\|_{X/S} = \inf_{z\in S}\|\alpha x - z\|_X = \inf_{z/\alpha\in S}\|\alpha(x-z/\alpha)\|_X = |\alpha|\|x+S\|_{X/S}. \tag{14.7.1}$$

而对于后者，考虑任意正数 ϵ，存在 $z_x, z_y \in S$，使得 $\inf_{z\in S}\|x-z_x\|_X + \epsilon/2 > \|x-z_x\|_X$ 且 $\inf_{z\in S}\|y-z\|_X + \epsilon/2 > \|y-z_y\|_X$，因此

$$\begin{aligned}\|(x+S)+(y+S)\|_{X/S} = \|(x+y)+S\|_{X/S} = \inf_{z\in S}\|(x+y)-z\|_X \le \|(x+y)-(z_x+z_y)\|_X \\ \le \|x-z_x\|_X + \|y-z_y\|_X \le \|x+S\|_{X/S} + \|y+S\|_{X/S} + \epsilon\end{aligned} \tag{14.7.2}$$

对所有正数 ϵ 成立，因此 $\|(x+S)+(y+S)\|_{X/S} \le \|x+S\|_{X/S} + \|y+S\|_{X/S}$，即三角不等式成立。

对于正定性，首先 $\|x-z\|_X \ge 0$ 对所有 $z\in S$ 成立，因此 $\|x+S\|_{X/S} \ge 0$。在 $S=X$ 的平凡情形中，$\|\cdot\|_{X/S}$ 为零映射；不妨设 $S \subsetneq X$。若 S 是闭集，根据定理 14.9的证明，$\inf_{z\in S}\|x-z\| > 0$ 对每个 $x\in X\backslash S$ 成立。因此 $\|x+S\|_{X/S}=0$ 当且仅当 $x\in S$，即 $x+S=S$ 是商空间 X/S 内的零向量；反之若 $\|\cdot\|_{X/S}$ 正定，考虑 S 内任意收敛到某个 $x\in X$ 的序列 (z_n)。对于任意正数 ϵ，存在 $n_0\in\mathbb{N}$ 使得 $\inf_{z\in S}\|x-z\| \le \|x-z_n\| < \epsilon$ 对所有 $n\ge n_0$ 成立。由 ϵ 的任意性，$\inf_{z\in S}\|x-z\|=0$，因此正定性要求 $x\in S$，故 S 是闭集。

定义 14.8 (商范数)

若 S 是赋范空间 X 的闭子空间，则范数 $\|\cdot\|_{X/S}$ 称为 X/S 上的商范数（quotient norm）。

♣

定理 14.11

设 X 为 Banach 空间，S 是 X 的闭子空间，则 X/S 关于商范数构成 Banach 空间。

♡

证明 这里我们采用定理 14.1中对 Banach 空间的刻画：对于商空间 V/S 内任意绝对收敛的序列 (x_n+S)，存在 $z_n\in S$ 使得 $\|x_n+S\|_{X/S} + 1/2^n > \|x_n-z_n\|_X$ 对每个自然数 n 成立，因此

$$\sum_{k=1}^{n}\|x_k-z_k\|_X < \sum_{k=1}^{n}\|x_n+S\|_{X/S} + \left(1-\frac{1}{2^n}\right) < \sum_{n=1}^{\infty}\|x_n+S\|_{X/S} + 1 < \infty \tag{14.7.3}$$

也对所有自然数 n 成立。利用实数列的单调收敛准则（定理 1.12），级数 $\sum_{n=1}^{\infty}(x_n-z_n)$ 绝对收敛。根据定理 14.1，X 完备意味着 $\sum_{n=1}^{\infty}(x_n-z_n)$ 收敛到某个 $x\in X$。那么对于任意正数 ϵ，存在自然数 n_0 使得

$$\begin{aligned}\left\|\sum_{k=1}^{n}(x_k+S)-(x+S)\right\|_{X/S} = \left\|\left(\sum_{k=1}^{n}x_k-x\right)+S\right\|_{X/S} = \inf_{z\in S}\left\|\left(\sum_{k=1}^{n}x_k-x\right)-z\right\|_X \\ \le \left\|\sum_{k=1}^{n}x_k - x - \sum_{k=1}^{n}z_k\right\|_X = \left\|\sum_{k=1}^{n}(x_k-z_k)-x\right\|_X < \epsilon\end{aligned} \tag{14.7.4}$$

对所有 $n\ge n_0$ 成立，因此 $\sum_{n=1}^{\infty}(x_n+S)$ 收敛到 $(x+S)$。根据定理 14.1，X/S 完备。

附录 A 拓扑基础

目录

A.1 拓扑的定义 135
A.2 收敛性与连续性 136
A.3 同胚 138
A.4 拓扑的基与子基 139
A.5 乘积拓扑 141
A.6 诱导拓扑 142

A.1 拓扑的定义

定义 A.1 (拓扑)

设 X 为非空集合，若 X 的子集族 $\mathcal{T}$ 满足

1. $\varnothing \in \mathcal{T}$ 且 $X \in \mathcal{T}$；
2. 若 $U_1, U_2 \in \mathcal{T}$，则 $U_1 \cap U_2 \in \mathcal{T}$；
3. 若 $\{U_i | i \in I\}$ 是 $\mathcal{T}$ 的子集族，则 $\cup_{i\in I} U_i \in \mathcal{T}$；

则称 $\mathcal{T}$ 是 X 上的一个拓扑（topology），其中每个元素称为开集（open set）。给定 X 上的一个拓扑结构 $\mathcal{T}$，我们称 $(X, \mathcal{T})$ 为一个拓扑空间（topology space）。

定义 A.2 (邻域)

设 $(X, \mathcal{T})$ 为拓扑空间，$x \in X$，N 是 X 的一个子集。若存在开集 $U \in \mathcal{T}$ 使得 $x \in U \subseteq N$，则称集合 N 为 x 的一个邻域（neighborhood），相应的邻域系（family of neighborhoods）为

$$\mathcal{N}(x) = \{N \subseteq X | \exists U \in \mathcal{T} \text{ for which } x \in U \text{ and } U \subseteq N\}. \tag{A.1.1}$$

定义 A.3 (闭集)

设 U 为拓扑空间 $(X, \mathcal{T})$ 的一个子集，若 U 的补集 $U^c = X \backslash U$ 为开集，则称 U 为闭集（close set）。

例 A.1

1. 度量拓扑：设 (X, d) 是任意度量空间，$\mathcal{T}_{\text{metric}} = \{U \subseteq X | \forall x \in U, \exists r > 0 \text{ for which } B(x, r) \subseteq U\}$，则 $\mathcal{T}_{\text{metric}}$ 是 X 上的一个拓扑，称为度量拓扑（*metric topology*）；
2. 离散拓扑：设 X 是任意集合，$\mathcal{T}_{\text{discrete}} = \mathcal{P}(X) = \{Y | Y \subseteq X\}$，则称其为 X 上的一个拓扑；并且不难发现其为关于 X 上离散度量[1]的度量拓扑；
3. 平凡拓扑：设 X 是任意集合，$\mathcal{T}_{\text{trivial}} = \{\varnothing, X\}$，则其为 X 上的拓扑；不过只要 $|X| > 1$，那么该拓扑不是度量拓扑。

定义 A.4 (拓扑的强弱性)

设 $\mathcal{T}_1$ 和 $\mathcal{T}_2$ 是 X 上的两个拓扑。若 $\mathcal{T}_1 \subseteq \mathcal{T}_2$，则称 $\mathcal{T}_1$ 弱于 $\mathcal{T}_2$，或称 $\mathcal{T}_2$ 强于 $\mathcal{T}_1$。

注 在任意集合 X 上，$\mathcal{T}_{\text{trivial}}$ 是最弱拓扑，而 $\mathcal{T}_{\text{discrete}}$ 是最强拓扑。另外，由于集合包含通常不是全序关系，也存在不可比较的拓扑。

[1] 对于任意 $x, y \in X$，离散度量将有序对 (x, y) 映射到 $\delta_{x,y}$。

命题 A.1 (子空间拓扑)

设 $(X,\mathcal{T})$ 为拓扑空间，$Y\subseteq X$ 是一个子集，则集族 $\mathcal{T}_Y=\{U\cap Y|U\in\mathcal{T}\}$ 是 Y 上的一个拓扑，称为子空间拓扑（subspace topology）。

证明 首先 $\varnothing,X\in\mathcal{T}$，因此 $\varnothing=\varnothing\cap Y\in\mathcal{T}_Y$ 且 $Y=X\cap Y\in\mathcal{T}_Y$；对于任意 $V_1,V_2\in\mathcal{T}_Y$，存在 $U_1,U_2\in\mathcal{T}$ 使得 $V_1=U_1\cap Y,V_2=U_2\cap Y$，因此 $V_1\cap V_2=(U_1\cap U_2)\cap Y\in\mathcal{T}_Y$；对于任意 $\{V_i|i\in I\}\subseteq\mathcal{T}_Y$，其中每个元素对应一个 $U_i\in\mathcal{T}$ 使得 $V_i=U_i\cap Y$，因此 $\cup_{i\in I}V_i=(\cup_{i\in I}U_i)\cap Y\in\mathcal{T}_Y$。综上，$\mathcal{T}_Y$ 构成 Y 上的一个拓扑。

命题 A.2 (乘积拓扑)

设 $(X,\mathcal{T}_X)$ 和 $(Y,\mathcal{T}_Y)$ 为拓扑空间，则

$$\mathcal{T}_{X\times Y}=\{W\subseteq X\times Y|\forall(x,y)\in W,\ \exists U\in\mathcal{T}_X \text{ and } V\in\mathcal{T}_Y \text{ for which } (x,y)\in U\times V\subseteq W\}\tag{A.1.2}$$

是 $X\times Y$ 上的一个拓扑，称为乘积拓扑（product topology）。

证明 显然 $\varnothing,X\times Y\in\mathcal{T}_{X\times Y}$；对于任意 $W_1,W_2\in\mathcal{T}_{X\times Y}$ 以及任意 $(x,y)\in W_1\cap W_2$，存在 $U_1,U_2\in\mathcal{T}_X$ 和 $V_1,V_2\in\mathcal{T}_Y$,，使得 $(x,y)\in U_1\times V_1\subseteq W_1$ 且 $(x,y)\in U_2\times V_2\in W_2$。因此 $(x,y)\in(U_1\cap U_2)\times(V_1\cap V_2)\subseteq W_1\cap W_2$，即 $W_1\cap W_2\in\mathcal{T}_{X\times Y}$；对于任意 $\{W_i|i\in I\}\subseteq\mathcal{T}_{X\times Y}$ 以及任意 $(x,y)\in\cup_{i\in I}W_i$，存在 $i\in I$ 使得 $(x,y)\in W_i$，即存在 $U_i\in\mathcal{T}_X$ 和 $V_i\in\mathcal{T}_Y$ 使得 $(x,y)\in U_i\times V_i\subseteq W_i\subseteq\cup_{i\in I}W_i$，因此 $\cup_{i\in I}W_i\in\mathcal{T}_{X\times Y}$。故 $\mathcal{T}_{X\times Y}$ 是一个拓扑。

A.2 收敛性与连续性

定义 A.5 (收敛)

设 (x_n) 是拓扑空间 $(X,\mathcal{T})$ 中的序列。若存在 $x_0\in X$，使得对于 x_0 的任意邻域 N，存在正整数 k 使得 $x_n\in N$ 对所有 $n\geq k$ 成立，则称 (x_n) 收敛到 x，记作 $(x_n)\to x_0$。

例 A.2

1. 度量拓扑下的收敛：按度量收敛和按度量拓扑收敛在度量空间内是一致的；
2. 离散拓扑下的收敛：由于离散拓扑空间中的开球 $B(x,1)=\{x\}$，容易看出 $(x_n)\to x_0$ 当且仅当存在正整数 k，使得 $x_n=x_0$ 对所有 $n\geq k$ 成立。直观地说，在离散拓扑空间内，只有“最终常值”的序列收敛；
3. 平凡拓扑下的收敛：由于平凡拓扑空间内唯一的非空开集为集合 X，因此 X 中任意序列 $\{x_n\}$ 都是收敛的，并且任意元素 x_0 均为极限（不唯一性）。

定义 A.6 (序列连续映射)

若对于拓扑空间之间的映射 $f:(X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$ 与 X 中任意收敛序列 $(x_n)\to x_0$，在 Y 中均有 $(f(x_n))\to f(x_0)$，则称 f 在 x_0 处序列连续（sequentially continuous at x_0）。若 f 在 X 中每个 x_0 处都序列连续，则称 f 为序列连续映射（sequentially continuous map）。

定义 A.7 (连续映射)

对于拓扑空间之间的映射 $f:(X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$，若 Y 中点 $f(x_0)$ 的任意邻域 B 的原像 $f^{-1}(B)$ 都是 X 中点 x_0 的邻域，则称 f 在点 x_0 处连续。若 f 在 X 中每个 x_0 处均连续，则称 f 为连续映射。

命题 A.3 (复合映射的连续性)

设 X,Y,Z 为拓扑空间。

1. 若 $f:X\to Y$ 和 $g:Y\to Z$ 分别在 x_0 和 $f(x_0)$ 处连续，则 $g\circ f:X\to Z$ 在 x_0 处连续；
2. 若 $f:X\to Y$ 和 $g:Y\to Z$ 分别在 x_0 和 $f(x_0)$ 处序列连续，则 $g\circ f:X\to Z$ 在 x_0 处序列连续。

证明 根据1.1.3式，取原像的操作保持并集开集，该性质容易验证。

命题 A.4

若 $f:(X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$ 在 x_0 处连续，则其在 x_0 处序列连续。特别地，任意连续映射都是序列连续的。♠

证明 对于任意收敛到 x_0 的序列 (x_n)，在 Y 中取 $f(x_0)$ 的任意邻域 B，则 $f^{-1}(B)$ 是 X 中点 x_0 的邻域。利用收敛性的定义，存在正整数 k 使得 $x_n\in f^{-1}(B)$，因此 $f(x_n)\in B$。

命题 A.5 (连续性的刻画)

设 $f:(X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$ 是一个映射，则以下表述等价：

1. f 是连续的；
2. 对于任意开集 $U\in\mathcal{T}_Y$，原像 $f^{-1}(U)$ 在 X 中是开集；
3. 对于任意闭集 $C\in\mathcal{T}_Y$，原像 $f^{-1}(C)$ 在 X 中是闭集。♠

证明 容易证明表述 2 和表述 3 等价，下面证明表述 1 和表述 2 等价。

1. 若表述 1 为真，对于开集 $U\subseteq\mathcal{T}_Y$，考虑任意 $x\in f^{-1}(U)$，那么 U 是 $f(x)$ 的开邻域。连续性的定义要求 $f^{-1}(U)$ 是 $f(x)$ 的邻域，因此存在开集 $V_x\subseteq f^{-1}(U)$ 使得 $x\in V_x$。这表明 $f^{-1}(U)\subseteq\cup_{x\in X}V_x\subseteq f^{-1}(U)$，即 $f^{-1}(U)=\cup_{x\in X}V_x$ 为开集，表述 2 为真；
2. 若表述 2 为真，对于任意 $x\in X$ 对应的像 $f(x)$ 的邻域 U，存在开集 $V\subseteq U$ 使得 $f(x)\in V$。根据条件，$f^{-1}(V)$ 为 x 的一个开邻域，因此 $f^{-1}(U)\supseteq f^{-1}(V)$ 也是 x 的邻域，进而 f 连续，表述 1 为真。

定义 A.8 (开映射与闭映射)

对于拓扑空间之间的映射 $f:(X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$，

1. 若 X 中任意开集 U 的像 $f(U)$ 在 Y 中是开集，则称映射 f 为开映射（open map）；
2. 若 X 中任意闭集 C 的像 $f(C)$ 在 Y 中是闭集，则称映射 f 为闭映射（closed map）；♣

命题 A.6 (子空间的嵌入映射)

设 $(X,\mathcal{T}_X)$ 为拓扑空间，赋予 $A\subseteq X$ 以子空间拓扑，则包含映射 $\iota:A\hookrightarrow X$ 是连续的，且子空间拓扑是 A 上使得 ι 连续的最弱拓扑。♠

证明 对于 X 中任意开集 U，显然 $\iota^{-1}(U)=\varnothing$ 在子空间拓扑内；当 $\iota^{-1}(U)\neq\varnothing$ 时，对于任意 $x\in\iota^{-1}(U)\subseteq A$，显然 $x=\iota(x)\in U$，因此 $\iota^{-1}(U)\subseteq A\cap U$；另一方面，对于任意 $x\in A\cap U$，$\iota(x)=x\in U$，即 $x\in\iota^{-1}(U)$，因此 $A\cap U\subseteq\iota^{-1}(U)$。综上，$\iota^{-1}(U)=A\cap U$ 在子空间拓扑内，因此是一个开集。根据命题 A.5，ι 连续。

对于 A 上任意使得 $\iota:(A,\mathcal{T})\to(X,\mathcal{T}_X)$ 连续的拓扑 $\mathcal{T}$ 以及任意开集 $U\in\mathcal{T}_X$，$\iota^{-1}(U)=A\cap U\in\mathcal{T}$。那么子空间拓扑包含在 $\mathcal{T}$ 内，即子空间拓扑是所有使得 ι 连续的拓扑中最弱的。

推论 A.1

设 $(X,\mathcal{T}_X)$ 和 $(Y,\mathcal{T}_Y)$ 为拓扑空间，赋予 $A\subseteq X$ 为子空间拓扑。

1. 若映射 $f:X\to Y$ 连续，则 $f|_A:A\to Y$ 连续；
2. 映射 $g:Y\to A$ 是连续当且仅当 $\iota\circ g:Y\to X$ 连续。♡

证明

1. 根据命题 A.3和A.6，$f|_A=f\circ\iota$ 连续；
2. 若 g 连续，则根据命题 A.3和A.6，$\iota\circ g$ 连续；若 $\iota\circ g$ 连续，则对于 A 的任意开子集 U，$A\cap U=U$，进而

$$(\iota\circ g)^{-1}(U)=\{y\in Y|(\iota\circ g)(y)\in U\}=\{y\in Y|g(y)\in\iota^{-1}(U)=A\cap U=U\}=g^{-1}(U)\tag{A.2.1}$$

是开集，故根据命题 A.5，g 连续。

命题 A.7 (乘积拓扑的投影映射)

设 $(X,\mathcal{T}_X)$ 和 $(Y,\mathcal{T}_Y)$ 为拓扑空间，$(X\times Y,\mathcal{T}_{X\times Y})$ 为其乘积拓扑空间，则投影映射

$$\pi_X : X\times Y\to X,\quad (x,y)\mapsto x, \tag{A.2.2}$$

$$\pi_Y : X\times Y\to Y,\quad (x,y)\mapsto y \tag{A.2.3}$$

都是连续映射，也是开映射。

♠

证明 根据命题 A.2，对于 X 的任意开子集 U，$\pi_X^{-1}(U)=U\times Y\in\mathcal{T}_{X\times Y}$，因此 π_X 连续；类似地 π_Y 也连续。对于 $X\times Y$ 内的任意开集 $W\in\mathcal{T}_{X\times Y}$ 以及任意 $x\in\pi_X(W)$，存在 $y\in Y$ 使得 $(x,y)\in W$。此时存在 $U\in\mathcal{T}_X$ 和 $V\in\mathcal{T}_Y$ 使得 $(x,y)\in U\times V\subseteq W$，因此 $x=\pi_X((x,y))\in\pi_X(U\times V)=U\subseteq\pi_X(W)$。利用 x 的任意性，将 U 记为 U_x，则 $\pi_X(W)\subseteq\cup_{x\in\pi_X(W)}U_x\subseteq\pi_X(W)$，因此 $\pi_X(W)=\cup_{x\in\pi_X(W)}U_x$ 为开集，从而 π_X 为开映射。

A.3 同胚

定义 A.9 (同胚)

设 X 和 Y 为拓扑空间，若存在可逆映射 $f:X\to Y$ 使得 f 和 f^{-1} 都是连续映射，则称拓扑空间 X 和 Y 同胚（homeomorphic），记作 $X\simeq Y$；映射 f 称为 X 与 Y 之间的同胚。在同胚下保持的性质称为拓扑性质（topological property）。

♣

注 根据命题 A.5，集合的开性与闭性在同胚下保持，因此是最基本的拓扑性质。

命题 A.8

同胚是拓扑空间之间的等价关系。

♠

证明 显然 ι 是拓扑空间 $(X,\mathcal{T}_X)$ 与其自身的同胚，自反性得证；对于拓扑空间 $X\simeq Y$，存在同胚 $f:X\to Y$，因此同胚的定义要求 $f^{-1}:Y\to X$ 也是同胚，即 $Y\simeq X$，对称性得证；对于拓扑空间 $X\simeq Y$ 和 $Y\simeq Z$，存在同胚 $f:X\to Y$ 和 $g:Y\to Z$，根据命题 A.3，$f\circ g:X\to Z$ 也是同胚，传递性得证。

命题 A.9

设 f 是一个连续双射，若 f 是开映射或闭映射，则 f 是同胚。

♠

证明 设 f 是从拓扑空间 X 到 Y 的连续双射。当 f 是开映射时，对于任意开集 $U\subseteq X$，$(f^{-1})^{-1}(U)$ 为开集。根据命题 A.5，f^{-1} 是连续映射，因此 f 为同胚；当 f 是闭映射时，对于任意开集 $U\subseteq X$，$(f^{-1}(U))^c=f^{-1}(U^c)$ 为闭集，即 $f^{-1}(U)$ 为开集，回到前述讨论情形。

定义 A.10 (拓扑向量空间)

设 X 为实（复）向量空间，并倍赋予了一个拓扑。若加法映射

$$\mathcal{A}:X\times X\to X,\quad (x,y)\mapsto x+y, \tag{A.3.1}$$

和标量乘法映射

$$\mathcal{M}:\mathbb{F}\times X\to X,\quad (\lambda,x)\mapsto \lambda x \tag{A.3.2}$$

在 $X\times X$ 和 $\mathbb{F}\times X$ 上的乘积拓扑中是连续映射，则称 X 为拓扑向量空间（topological vector space）。

♣

A.4 拓扑的基与子基

定义 A.11 (拓扑基)

若集族 $\mathcal{B}\in\mathcal{P}(X)$ 满足：

1. 对于任意 $x\in X$，存在 $B\in\mathcal{B}$ 使得 $x\in B$；
2. 对于任意 $B_1,B_2\in\mathcal{B}$ 和任意 $x\in B_1\cap B_2$，存在 $B\in\mathcal{B}$ 使得 $x\in B\subseteq B_1\cap B_2$；

则称 $\mathcal{B}$ 为 X 的一组拓扑基（basis for a topological），并称集族

$$\mathcal{T}_{\mathcal{B}}=\{U\subseteq X|\forall x\in U,\ \exists B\in\mathcal{B}\text{ for which }x\in B\subseteq U\}\tag{A.4.1}$$

是由基 $\mathcal{B}$ 生成的拓扑（topology generated by the basis $\mathcal{B}$）。

♣

注 这里我们对 $\mathcal{T}_{\mathcal{B}}$ 构成拓扑稍加说明：显然 $\varnothing,X\in\mathcal{T}_{\mathcal{B}}$，同时 $\mathcal{T}_{\mathcal{B}}$ 关于有限交集的封闭性由上述定义给出；设 $\{U_i|i\in I\}\subseteq\mathcal{T}_{\mathcal{B}}$ 为开集族，那么对于任意 $x\in\cup_iU_i$，存在指标 i 使得 $x\in U_i$。根据定义，存在 $B_i\in\mathcal{B}$ 使得 $x\in B_i\subseteq U_i\subseteq\cup_iU_i$，因此 $\cup_iU_i\in\mathcal{T}_{\mathcal{B}}$。可见 $\mathcal{T}_{\mathcal{B}}$ 的确构成一个拓扑。

注 不同的拓扑基可以生成相同的拓扑，例如 $\mathcal{B}_1=\{B(x,r)|x\in\mathbb{R}^2,r>0\}$、$\mathcal{B}_2=\{B(x,r)|x\in\mathbb{Q}^2,r\in\mathbb{Q}_{>0}\}$ 和 $\mathcal{B}_3=\{(a,b)\times(c,d)|a,b,c,d\in\mathbb{R}\}$ 是 $\mathbb{R}^2$ 的三种不同的拓扑基，但它们生成相同的拓扑；这一点容易通过相互包含验证。另外，这三种拓扑基的基数也不尽相同，其中只有 $\mathcal{B}_2$ 是可数的。

例 A.3 对于任意给定的拓扑空间 $(X_i,\mathcal{T}_i)$ $(i\in I)$，类似命题 A.2定义其乘积拓扑 $\prod_iX_i=\{(x_i)|x_i\in X_i\}$。考虑集族 $\mathcal{B}=\{\prod_iU_i|U_i\in\mathcal{T}_i\}$。对于任意 $x=(x_i)\in\prod_iX_i$，由于 $X_i\in\mathcal{T}_i$，必有 $x_i\in X_i$，进而 $x\in\prod_{\mathrm{i}}X_{\mathrm{i}}\in\mathcal{B}$；另一方面，对于任意 $B_1=\prod_{\mathrm{i}}U_{\mathrm{i}}$ 和 $B_2=\prod_{\mathrm{i}}V_{\mathrm{i}}$，以及任意 $x=(x_i)\in B_1\cap B_2$，根据定义，$x_{\mathrm{i}}\in U_{\mathrm{i}}\cap V_{\mathrm{i}}\in\mathcal{T}_{\mathrm{i}}$，故存在 $B=\prod_{\mathrm{i}}(U_{\mathrm{i}}\cap V_{\mathrm{i}})\in\mathcal{B}$，使得 $x\in B=B_1\cap B_2$。因此 $\mathcal{B}$ 是 $\prod_{\mathrm{i}}X_{\mathrm{i}}$ 的一个拓扑基，其生成的拓扑

$$\mathcal{T}_{\text{box}}=\left\{U\subseteq\prod_{\mathrm{i}}X_{\mathrm{i}}\middle|\forall(x_{\mathrm{i}})\in U,\ \exists U_{\mathrm{i}}\in\mathcal{T}_{\mathrm{i}}\text{ for which }(x_{\mathrm{i}})\in\prod_{\mathrm{i}}U_{\mathrm{i}}\subseteq U\right\}\tag{A.4.2}$$

称为 $X=\prod_{\mathrm{i}}X_{\mathrm{i}}$ 上的箱拓扑（*box topology*）。

命题 A.10

若 $\mathcal{B}$ 是 X 的一个拓扑基，那么由其生成的拓扑为 $\mathcal{T}_{\mathcal{B}}=\{\cup_{B\in\mathcal{B}'}B|\mathcal{B}'\subseteq\mathcal{B}\}$。

♠

证明 根据定义，显然 $\mathcal{B}\subseteq\mathcal{T}_{\mathcal{B}}$。那么对于任意 $B\in\mathcal{B}'\subseteq\mathcal{B}\subseteq\mathcal{T}_{\mathcal{B}}$，均有 $\cup_{B\in\mathcal{B}'}B\in\mathcal{T}_{\mathcal{B}}$，因此 $\{\cup_{B\in\mathcal{B}'}B|\mathcal{B}'\subseteq\mathcal{B}\}$；另一方面，对于任意 $U\in\mathcal{T}_{\mathcal{B}}$ 以及 U 内的元素 x，存在 $B_x\in\mathcal{B}$ 使得 $x\in B_x\subseteq U$，那么 $U\subseteq\cup_{x\in U}B_x\subseteq U$，即 $U=\cup_{x\in U}B_x$，因此 $\mathcal{T}_{\mathcal{B}}\subseteq\{\cup_{B\in\mathcal{B}'}B|\mathcal{B}'\subseteq\mathcal{B}\}$。

推论 A.2

设 $\mathcal{B}$ 为 X 的一组拓扑基，拓扑 $\mathcal{T}'$ 满足 $\mathcal{B}\subseteq\mathcal{T}'$，则 $\mathcal{T}_{\mathcal{B}}\subseteq\mathcal{T}'$。

♡

证明 根据命题 A.10，对于 $\mathcal{T}_{\mathcal{B}}$ 内的任意元素，存在 $\mathcal{B}'\subseteq\mathcal{B}$ 使其可表示为 $\cup_{B\in\mathcal{B}'}B$。因此若 $\mathcal{B}\subseteq\mathcal{T}'$，则 $B\in\mathcal{T}'$，进而 $\cup_{B\in\mathcal{B}'}B\in\mathcal{T}'$，最终得到 $\mathcal{T}_{\mathcal{B}}\subseteq\mathcal{T}'$。

类似向量空间生成集的概念，我们也可以对一般子集族定义生成拓扑的概念。

定义 A.12 (生成拓扑)

对于任意 X 的子集构成的族 $\mathcal{S}\subseteq\mathcal{P}(X)$，称 X 上包含 $\mathcal{S}$ 的所有拓扑的交集是由 $\mathcal{S}$ 生成的拓扑（topology generated by $\mathcal{S}$），通常记作 $\mathcal{T}_{\mathcal{S}}$。

♣

注 容易验证拓扑的交集也是拓扑；根据定义，在使得 $\mathcal{S}$ 中所有集合均为开集的拓扑中，$\mathcal{T}_{\mathcal{S}}$ 是最弱的那个。

命题 A.11

设子集族 $\mathcal{S}\subseteq\mathcal{P}(X)$，并记 $\mathcal{B}=\{B\subseteq X|\exists S_1,\cdots,S_m\in\mathcal{S}\text{ for which }B=S_1\cap\cdots\cap S_m\}$，那么

1. 若 $\cup_{S\in\mathcal{S}}S=X$，则 $\mathcal{B}$ 是 X 的一组拓扑基，并且 $\mathcal{T}_{\mathcal{S}}=\mathcal{T}_{\mathcal{B}}$；
2. 一般地，若 $X'=\cup_{S\in\mathcal{S}}S\subseteq X$，则 $\mathcal{B}$ 是 X' 上的一组拓扑基，并且 $\mathcal{T}_{\mathcal{S}}=\{X\}\cup\mathcal{T}_{\mathcal{B}}$。

♠

证明 我们仅对第二条性质进行证明。对于任意 $x\in X'$，存在子集 $S\in\mathcal{S}$ 使得 $x\in S$，而 $S\in\mathcal{B}$；对于任意 $B_1,B_2\in\mathcal{B}$ 和任意 $x\in B_1\cap B_2$，存在 $S_1,\cdots,S_m\in\mathcal{S}$ 和 $T_1,\cdots,T_n\in\mathcal{S}$，使得 $B_1=\cap_{i=1}^m S_i$ 且 $B_2=\cap_{j=1}^n T_j$。因此 $x\in(\cap_{i=1}^m S_i)\cap(\cap_{j=1}^n S_i)\subseteq B_1\cap B_2$。综上，$\mathcal{B}$ 是 X' 的一组拓扑基。容易验证 $\{X\}\cup\mathcal{T}_{\mathcal{B}}$ 是 X 上的一个拓扑，并且 $\mathcal{S}\subseteq\mathcal{B}\subseteq\{X\}\cup\mathcal{T}_{\mathcal{B}}$。根据推论 A.2，$\mathcal{T}_{\mathcal{S}}\subseteq\{X\}\cup\mathcal{T}_{\mathcal{B}}$；另一方面，显然 $\mathcal{B}\subseteq\mathcal{T}_{\mathcal{S}}$，根据推论 A.2，$\mathcal{T}_{\mathcal{B}}\subseteq\mathcal{T}_{\mathcal{S}}$。由于 $\mathcal{T}_{\mathcal{S}}$ 是 X 上的拓扑，自然有 $X\in\mathcal{T}_{\mathcal{S}}$，故 $\{X\}\cup\mathcal{T}_{\mathcal{B}}\subseteq\mathcal{T}_{\mathcal{S}}$，反向包含成立。

根据上述结论，我们可以引出拓扑子基的概念。

定义 A.13 (拓扑子基)

若集族 $\mathcal{S}\subseteq\mathcal{P}(X)$ 满足 $\cup_{S\in\mathcal{S}}S=X$，则称 $\mathcal{S}$ 是 X 的一组拓扑子基（subbasis for a topology），称 $\mathcal{T}_{\mathcal{S}}$ 为由子基 $\mathcal{S}$ 生成的拓扑（topology generated by the subbasis $\mathcal{S}$）。

♣

例 A.4

1. 对于 $\mathbb{R}$ 上的标准 Euclidean 拓扑，$\mathcal{B}=\{(a,b)|a<b\}$ 是一组拓扑基，而 $\mathcal{S}=\{(-\infty,a),(a,+\infty)|a\in\mathbb{R}\}$ 是一组拓扑子基；
2. 对于 $X\times Y$ 上的乘积拓扑，$\mathcal{B}=\{U\times V|U\in\mathcal{T}_X,V\in\mathcal{T}_Y\}$ 是一组拓扑基，而 $\mathcal{S}=\{U\times Y|U\in\mathcal{T}_X\}\cup\{X\times V|V\in\mathcal{T}_Y\}$ 是一组拓扑子基。

命题 A.12 (拓扑基的判据)

设 $(X,\mathcal{T})$ 为拓扑空间，那么集族 $\mathcal{B}\subseteq\mathcal{P}(X)$ 是生成 $\mathcal{T}$ 的一组拓扑基当且仅当 $\mathcal{B}\subseteq\mathcal{T}$，并且对于任意 $U\in\mathcal{T}$ 和任意 $x\in U$，存在 $B\in\mathcal{B}$ 使得 $x\in B\subseteq U$。

♠

证明 充分性是显然的；对于必要性，拓扑基的第一条要求可通过取 $U=X$ 证明。对于任意 $B_1,B_2\in\mathcal{B}\subseteq\mathcal{T}$ 和任意 $x\in B_1\cap B_2$，首先有 $B_1\cap B_2\in\mathcal{T}$。根据条件，存在 $B\in\mathcal{B}$ 使得 $x\in B\subseteq B_1\cap B_2$。综上，$\mathcal{B}$ 是 X 的一组拓扑基，根据推论 A.2，$\mathcal{T}_{\mathcal{B}}\subseteq\mathcal{T}$。另外，类似命题 A.10的证明，对于任意 $U\in\mathcal{T}$，$U=\cup_{x\in U}B_x\in\mathcal{T}_{\mathcal{B}}$，其中 $B_x\in\mathcal{B}$。因此 $\mathcal{T}\subseteq\mathcal{T}_{\mathcal{B}}$，反向包含成立。

命题 A.13 (拓扑子基的判据)

设 $(X,\mathcal{T})$ 为拓扑空间，那么集族 $\mathcal{S}\subseteq\mathcal{P}(X)$ 是生成 $\mathcal{T}$ 的一组拓扑子基当且仅当 $\mathcal{S}\subseteq\mathcal{T}$，并且对于任意 $U\in\mathcal{T}$ 和任意 $x\in U$，存在 $S_1,\cdots,S_m\in\mathcal{S}$ 使得 $x\in\cap_{i=1}^n S_i\subseteq U$。

♠

证明 对于充分性，根据命题 A.11，可通过拓扑子基 $\mathcal{S}$ 构造 $\cup_{S\in\mathcal{S}}S=X$ 的一组拓扑基 $\mathcal{B}$，使得 $\mathcal{T}=\mathcal{T}_{\mathcal{S}}=\mathcal{T}_{\mathcal{B}}$。那么显然有 $\mathcal{S}\subseteq\mathcal{T}_{\mathcal{S}}=\mathcal{T}$；对于任意 $U\in\mathcal{T}=\mathcal{T}_{\mathcal{B}}$ 和任意 $x\in U$，存在 $B\in\mathcal{B}$ 使得 $x\in B\subseteq U$。根据命题 A.11，这里的 B 为 $\mathcal{S}$ 中有限多个集合的交集，因此充分性得证。

对于必要性，根据命题 A.11，可通过拓扑子基 $\mathcal{S}$ 构造 $X'=\cup_{S\in\mathcal{S}}S$ 的一组拓扑基 $\mathcal{B}$，使得 $\mathcal{T}_{\mathcal{S}}=\{X\}\cup\mathcal{T}_{\mathcal{B}}$。根据推论 A.2，$\mathcal{S}\subseteq\mathcal{T}$ 意味着 $\mathcal{T}_{\mathcal{S}}\subseteq\mathcal{T}$；对于反向包含，考虑任意 $U\in\mathcal{T}$ 和任意 $x\in U$，条件表明存在 $B_x\in\mathcal{B}$ 使得 $x\in B_x\subseteq U$，因此 $U\subseteq\cup_{x\in U}B_x\subseteq U$，即 $U=\cup_{x\in U}B_x$。根据命题 A.10，这意味着 $U\in\mathcal{T}_{\mathcal{B}}$ 对任意 $U\in\mathcal{T}$ 成立，那么 $X\in\mathcal{T}_{\mathcal{B}}$。根据推论 A.2，$O\subseteq\mathcal{T}_{\mathcal{B}}=\{X\}\cup\mathcal{T}_{\mathcal{B}}=\mathcal{T}_{\mathcal{S}}$。

定理 A.1 (连续映射的刻画：拓扑基与子基)

设 $\mathcal{B}$ 是 $\mathcal{T}_Y$ 的一个拓扑基，$\mathcal{S}$ 是 $\mathcal{T}_Y$ 的一个拓扑子基，那么以下几种表述等价：

1. 映射 $f:(X,\mathcal{T}_X)\to(Y,\mathcal{T}_Y)$ 是连续映射；

2. 对于任意 $B \in \mathcal{B}$，其原像 $f^{-1}(B)$ 在 X 中是开集；
3. 对于任意 $S \in \mathcal{S}$，其原像 $f^{-1}(S)$ 在 X 中是开集。

A.5 乘积拓扑

有了拓扑基的概念，现在我们可以重新审视命题 A.2中乘积拓扑的定义，将其定义推广到任意指标。

定义 A.14 (乘积拓扑)

设 $(X_\alpha, \mathcal{T}_\alpha)$ 为一族拓扑空间，称 $\prod_\alpha X_\alpha$ 上由子基 $\mathcal{S}_{\text{product}} = \cup_\beta \{\pi_\beta^{-1}(V_\beta) | V_\beta \in \mathcal{T}_\beta\}$ 生成的拓扑 $\mathcal{T}_{\text{product}}$ 为 $\prod_\alpha X_\alpha$ 的乘积拓扑（product topology），其中投影映射 $\pi_\beta : \prod_\alpha X_\alpha \to X_\beta$ 将 (x_α) 映射到 x_β。

注 根据命题 A.13，容易验证对于两个拓扑空间，该定义可导出命题 A.2中的结果。

注 回顾例A.3中关于箱拓扑的定义。对于任意指标 β 和 $V_\beta \in \mathcal{T}_\beta$，取 $\pi_\beta^{-1}(V_\beta) \in \mathcal{S}$。那么对于任意 $(x_\alpha) \in \pi_\beta^{-1}(V_\beta)$，分量 $x_\beta \in V_\beta \subseteq X_\beta$，其余分量 $x_\alpha \in X_\alpha$ $(\alpha \neq \beta)$，因此 $(x_\alpha) \in \pi_\beta^{-1}(V_\beta)$，即 $\mathcal{S} \subseteq \mathcal{T}_{\text{box}}$。根据推论 A.2，$\mathcal{T}_{\text{product}} \subseteq \mathcal{T}_{\text{box}}$，即乘积拓扑弱于箱拓扑。尽管乘积拓扑不如箱拓扑那样自然，但事实证明前者更重要：乘积具有很多良好的拓扑性质，而箱拓扑在拓扑学中被广泛用作反例。当然，对于有限乘积空间，这两种拓扑相同。

在此定义下，我们可以得到类似命题 A.7的结论。

命题 A.14

设 $(X_\alpha, \mathcal{T}_\alpha)$ 为一族拓扑空间，那么无论赋予 $\prod_\alpha X_\alpha$ 是乘积拓扑还是箱拓扑，对于任意指标 β，投影映射 $\pi_\beta : \prod_\alpha X_\alpha \to X_\beta$ 都是连续的开映射。

证明 由于 $\mathcal{T}_{\text{product}} \subseteq \mathcal{T}_{\text{box}}$，对于任意 $V_\beta \in \mathcal{T}_\beta$，只要 $\pi_\beta^{-1}(V_\beta) \in \mathcal{T}_{\text{product}}$ 就有 $\pi_\beta^{-1}(V_\beta) \in \mathcal{T}_{\text{box}}$，因此仅需在 $\mathcal{T}_{\text{product}}$ 中证明映射的连续性；类似地，仅需在 $\mathcal{T}_{\text{box}}$ 中证明 π_β 为开映射。前者是显然的；对于后者，回顾箱拓扑的定义式(A.4.2)，对于任意开集 $U \in \mathcal{T}_{\text{box}}$ 以及任意 $(x_\alpha) \in U$，存在 $U_\alpha \in \mathcal{T}_\alpha$ 使得 $(x_\alpha) \in \prod_\alpha U_\alpha \subseteq U$，因此 $x_\beta = \pi_\beta((x_\alpha)) \in U_\beta \subseteq \pi_\beta(U)$。利用 x 的任意性，可知 $\pi_\beta(U)$ 为开集，即 π_β 为开映射。

命题 A.15 (乘积拓扑的刻画)

乘积拓扑 $\mathcal{T}_{\text{product}}$ 是在 $\prod_\alpha X_\alpha$ 上使得所有投影映射 π_β 都连续的拓扑中最弱的拓扑。

证明 π_η 关于 $\mathcal{T}_{\text{product}}$ 的连续性源于命题 A.14。设 $\mathcal{T}$ 为在 $\prod_\alpha X_\alpha$ 使得所有典范投影 π_β 均连续的拓扑。那么根据命题 A.5，对于任意 $V_\beta \in \mathcal{T}_\beta$，$\pi^{-1}(V_\beta) \in \mathcal{T}$，这意味着 $\mathcal{S}_{\text{product}} \subseteq \mathcal{T}$。根据推论 A.2，$\mathcal{T}_{\text{product}} \subseteq \mathcal{T}$。

定理 A.2 (乘积拓扑的泛性质)

设 X 和每个 X_α 均为拓扑空间，赋予空间 $\prod_\alpha X_\alpha$ 以乘积拓扑。定义映射 $f : X \to \prod_\alpha X_\alpha$ 为 $f(x) = (f_\alpha(x))$，则 f 是连续映射当且仅当所有 $f_\alpha = \pi_\alpha \circ f$ 都是连续映射。此外，乘积拓扑是 $\prod_\alpha X_\alpha$ 上唯一满足此性质的拓扑。

证明 若 f 连续，那么利用 π_β 的连续性（命题 A.14）和复合映射的连续性（命题 A.3），$f_\alpha = \pi_\alpha \circ f$ 也是连续的；若 f_α 连续，那么对于任意 $V_\beta \in \mathcal{T}_\beta$ 和相应子基 $\pi_\beta^{-1}(V_\beta)$，

$$f^{-1}(\pi_\beta^{-1}(V_\beta)) = \left\{x \in X \middle| f(x) \in \pi_\beta^{-1}(V_\beta)\right\} = \{x \in X | (\pi_\beta \circ f)(x) \in V_\beta\} = f_\beta^{-1}(V_\beta) \in \mathcal{T}_X, \tag{A.5.1}$$

根据命题 A.5，此时 f 也是连续映射。泛性质的证明从略。

注 事实上，箱拓扑就是上述不满足泛性质的一例。例如设 $X = \mathbb{R}^{\mathbb{N}}$，定义映射 $f : \mathbb{R} \to \mathbb{R}^{\mathbb{N}}$ 为 $f(t) = (t, t, \cdots)$，那么 f 的每个分量都是 $\mathbb{R}$ 上元素 t 到其自身的恒等映射，从而 f 的每个分量都连续；但是如果赋予 X 以箱拓扑，那么 $(-1, 1) \times (-1/2, 1/2) \times \cdots$ 在箱拓扑中是开集，但其原像 $\{0\}$ 不是 X 中的开集。

A.6 诱导拓扑

事实上，我们可以从命题 A.1中的子空间拓扑和命题 A.15中乘积拓扑的刻画中得到启发，利用“使得给定映射都连续的最弱拓扑”的特征，在原像空间上构造一个新的拓扑。这就是诱导拓扑。

定义 A.15 (诱导拓扑)

设 $(Y_\alpha,\mathcal{T}_\alpha)$ 为一族拓扑空间，$\mathcal{F}=\{f_\alpha|X\to(Y_\alpha,\mathcal{T}_\alpha)\}$ 是一族映射，那么在 X 上使得所有 f_α 都是连续映射的拓扑中，最弱的拓扑 $\mathcal{T}_\mathcal{F}$ 称为 X 的 $\mathcal{F}$-诱导拓扑（$\mathcal{F}$-induced topology）或弱拓扑（weak topology）。♣

注

1. 如果赋予 X 最强的拓扑，即离散拓扑，则任何映射 $f_\alpha:(X,\mathcal{T}_{\text{discrete}})\to(Y_\alpha,\mathcal{T}_\alpha)$ 都是连续的，该情形平凡。考虑在 X 上的一族拓扑 $\{\mathcal{T}_\beta\}$，若该拓扑族使得每个 f_α 关于每个 $\mathcal{T}_\beta$ 连续，那么 $\mathcal{T}=\cap_\beta\mathcal{T}_\beta$ 也是 X 上的一个拓扑，并且 f_α 关于 $\mathcal{T}$ 连续。因此，在 X 上存在唯一一个最弱拓扑，使得每个 f_α 连续；
2. 事实上，$\mathcal{T}_\mathcal{F}$ 是由子基 $\mathcal{S}_\mathcal{F}=\cup_\alpha\{f_\alpha^{-1}(V_\alpha)|V_\alpha\in\mathcal{T}_\alpha\}$ 生成的拓扑：首先对于任意 $f_\alpha\in\mathcal{F}$ 和任意 $V_\alpha\in\mathcal{T}_\alpha$，定义要求 $f_\alpha^{-1}(V_\alpha)\in\mathcal{T}_\mathcal{F}$，因此 $\mathcal{S}_\mathcal{F}\subseteq\mathcal{T}_\mathcal{F}$，进而推论 A.2要求 $\mathcal{T}_\mathcal{S}\subseteq\mathcal{T}_\mathcal{F}$；其次 $f_\alpha^{-1}(V_\alpha)\in\mathcal{S}_\mathcal{F}$ 对所有 α 成立，这意味着包含 $\mathcal{S}_\mathcal{F}$ 的拓扑是使得 $\mathcal{F}$ 连续的拓扑，从而 $\mathcal{T}_\mathcal{F}\subseteq\mathcal{T}_\mathcal{S}$。综上，$\mathcal{T}_\mathcal{F}=\mathcal{T}_\mathcal{S}$。

命题 A.16 (诱导拓扑的泛性质)

设 $(Y_\alpha,\mathcal{T}_\alpha)$ 为一族拓扑空间，$\mathcal{F}=\{f_\alpha|X\to(Y_\alpha,\mathcal{T}_\alpha)\}$ 为一族映射。赋予 X 以 $\mathcal{F}$-诱导拓扑，那么对于任意拓扑空间 Z，映射 $f:Z\to X$ 连续当且仅当所有 $f_\alpha\circ f:Z\to Y_\alpha$ 均连续。此外，$\mathcal{F}$-诱导拓扑是 X 上唯一满足该性质的拓扑。♠

证明 若映射 $f:Z\to X$，则由复合函数的连续性（命题 A.3），$f_\alpha\circ f$ 也是连续的；反之若 $f_\alpha\circ f$ 连续，则对于每个拓扑子基 $f_\beta^{-1}(V_\beta)$，$f^{-1}(f_\beta^{-1}(V_\beta))=(f_\beta\circ f)^{-1}(V_\beta)\in\mathcal{T}_\mathcal{F}$。根据命题 A.5，$f$ 连续。泛性质的证明从略。

例 A.5

1. 作为诱导拓扑的度量拓扑：对于任意度量空间 (X,d)，度量拓扑是由所有度量球 $\{B(x,r)|x\in X,r>0\}$ 生成的；换言之，度量拓扑是由映射族 $\{d_x|x\in X\}$ 生成的诱导拓扑，其中 $d_x:X\to\mathbb{R}$ 定义为 $d_x(y)=d(x,y)$；
2. 弱拓扑和弱-$*$ 拓扑：设 X 为拓扑向量空间，X^* 为其对偶空间。此时可以在 X 上定义一个新的拓扑，同时在 X^* 上定义一个自然的拓扑：
   (1) X 上的弱拓扑（*weak topology*）是 X^* 生成的诱导拓扑，即 X 上所有使得 $f\in X^*$ 连续的拓扑中最弱的拓扑，因此 X 上的弱拓扑比 X 上的原始拓扑更弱；
   (2) X^*-上的弱-$*$ 拓扑（*weak-$*$topology*）是由 $\{ev_x|x\in X\}$，即使得 X^* 上所有赋值映射 $ev_x:X^*\to\mathbb{R},\ l\mapsto l(x)$ 都连续的拓扑中最弱的拓扑。

附录 B 奇异值和 Moore-Penrose 逆

目录

B.1 奇异值 143
B.2 Moore-Penrose 逆 144
B.3 最小二乘逼近 144

B.1 奇异值

设 U 和 V 是域 $\mathbb{C}$ 或 $\mathbb{R}$ 上的 n 维内积空间，$\tau \in \mathcal{L}(U,V)$。根据定理 7.10，$\tau^*\tau$ 是正算子，此时 U 具有由 $\tau^*\tau$ 的本征向量构成的有序标准正交基 $\mathcal{B} = (u_1, \cdots, u_r, u_{r+1}, \cdots, u_n)$，其中 $n = \dim(U)$，对应的本征值满足 $\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n$。根据命题 7.2，向量集 $(u_1, \cdots, u_r)$ 和 $(u_{r+1}, \cdots, u_n)$ 分别是 $\ker(\tau^*\tau)^\perp = \mathrm{im}((\tau^*\tau)^*) = \mathrm{im}(\tau^*\tau) = \mathrm{im}(\tau^*) = \ker(\tau)^\perp$ 和 $\ker(\tau^*\tau) = \ker(\tau)$ 的有序标准正交基。对于 $i = 1, \cdots, r$，正数 $s_i = \sqrt{\lambda_i}$ 称为 τ 的奇异值（*singular value*）；对于 $i = r+1, \cdots, n$，取 $s_i = 0$。那么对于每个 $i \leq r$ 和 $u_i \in U$，可定义相应的 $v_i = s_i^{-1}\tau u_i \in V$。由此我们可以得到 τ 与 τ^* 之间的某种“对称性”：

$$\tau u_i = \begin{cases} s_i v_i, & i \leq r, \\ 0, & i > r \end{cases} \quad \text{and} \quad \tau^* v_i = s_i^{-1}\tau^*\tau u_i = s_i u_i = \begin{cases} s_i u_i, & i \leq r, \\ 0, & i > r. \end{cases} \tag{B.1.1}$$

此外，根据上述定义，各 v_i 也是标准正交的：当 $i, j \leq r$ 且 $i \neq j$ 时，

$$\langle v_i, v_j \rangle = \frac{1}{s_i \overline{s_j}} \langle \tau u_i, \tau u_j \rangle = \frac{1}{s_i s_j} \langle u_i, \tau^*\tau u_j \rangle = \frac{s_j}{s_i} \langle u_i, u_j \rangle = \delta_{i,j}, \tag{B.1.2}$$

同时可利用 Gram-Schmidt 增广（定理 5.6）将有序标准正交集 $(v_1, \cdots, v_r)$ 延拓为 V 上一组标准正交基。最后，注意到 $(\tau\tau^*)v_i = \tau(s_i u_i) = s_i(s_i v_i) = \lambda_i v_i$，因此 $(v_1, \cdots, v_r)$ 也是 $\tau\tau^*$ 的本征向量，对应的本征值同样为 $\lambda_1, \cdots, \lambda_r$。将上述讨论整理后，可得到如下定理。

定理 B.1

设 U 和 V 为 $\mathbb{C}$ 或 $\mathbb{R}$ 上的 n 维内积空间，$\tau \in \mathcal{L}(U,V)$ 具有秩 r，则分别存在 U 和 V 的有序标准正交基（ordered orthonormal basis, abbreviated as ONB）$\mathcal{B}$ 和 $\mathcal{C}$，使得

$$\mathcal{B} = (\underbrace{u_1, \cdots, u_r}_{\text{ONB for } \mathrm{im}(\tau^*)}, \underbrace{u_{r+1}, \cdots, u_n}_{\text{ONB for } \ker(\tau)}), \tag{B.1.3}$$

$$\mathcal{C} = (\underbrace{v_1, \cdots, v_r}_{\text{ONB for } \mathrm{im}(\tau)}, \underbrace{v_{r+1}, \cdots, v_n}_{\text{ONB for } \ker(\tau^*)}). \tag{B.1.4}$$

此外，对于 $1 \leq k \leq r$，$\tau u_i = s_i v_i$ 且 $\tau^* v_i = s_i u_i$，其中 $s_i > 0$ 称为 τ 的奇异值，定义为 $(\tau^*\tau)u_i = s_i^2 u_i$。向量 $u_1, \cdots, u_r$ 称为 τ 的右奇异向量（right singular vector），而向量 $v_1, \cdots, v_r$ 称为 τ 的左奇异向量（left singular vector）。

该结论的一个重要结果是奇异值分解（*singular value decomposition*）。设 $A \in \mathcal{M}_{m,n}(F)$，$\mathcal{B} = (u_1, \cdots, u_n)$ 和 $\mathcal{C} = (v_1, \cdots, v_m)$ 分别为来自 U 和 V 的标准正交基，那么

$$[\tau]_{\mathcal{B},\mathcal{C}} = ([\tau(u_1)]_{\mathcal{C}}, \cdots, [\tau(u_n)]_{\mathcal{C}}) = \mathrm{diag}(s_1, \cdots, s_r, \underbrace{0, \cdots, 0}_{(n-r)\text{ terms}}) \triangleq \Sigma. \tag{B.1.5}$$

根据定理 3.7，设 $A = [\tau]_{\mathcal{E}_n,\mathcal{E}_m}$，那么 $A = M_{\mathcal{C},\mathcal{E}_m}[\tau]_{\mathcal{B},\mathcal{C}}M_{\mathcal{E}_n,\mathcal{B}} = P\Sigma Q^*$，其中 $P = M_{\mathcal{C},\mathcal{E}_m}$ 和 $Q = M_{\mathcal{B},\mathcal{E}_n}$ 是幺正或正交的。这就是奇异值分解。至于分解的唯一性，设 $A = P\Sigma Q^*$，则 $A^*A = Q(\Sigma^*\Sigma)Q^*$，其中 $\Sigma^*\Sigma = \mathrm{diag}(s_1, \cdots, s_n^2)$。那么 $(A^*A)Q^{(j)} = s_j^2 Q^{(j)}$，即 s_j^2 是 A^*A 的本征值，因此 Σ 由 A 唯一确定（但 P 和 Q 未必唯一）。

注 我们不加证明地给出如下唯一性事实：若 $n \leq m$ 且本征值 λ_i 不同，则 P 唯一确定，最多右乘一个形如 $D = \mathrm{diag}(z_1, \cdots, z_m)$ 的对角阵，其中 $|z_i| = 1$；若 $n < m$，则 Q 不唯一确定；若 $m = n = r$，则对于任意给定的 P，存在唯一的 Q。由此我们可以看出，通常奇异值分解是不唯一的。

B.2 Moore-Penrose 逆

根据上一节的讨论，我们定义线性变换 $\tau^+ : V \to U$ 为 $\tau^+ v_i = \begin{cases} s_i^{-1} u_i, & \text{if } i \leq r, \\ 0, & \text{if } i > r. \end{cases}$ 若 $n = m = r$，则 $\tau^+ = \tau^{-1}$。线性变换 τ^+ 称为 τ 的 *Moore-Penrose 广义逆*（*Moore-Penrose generalized inverse*）或 *Moore-Penrose 伪逆*（*Moore-Penrose pseudoinverse*）；后面简称 MP 逆。

定理 B.2

设 $\tau \in \mathcal{L}(U, V)$，则 τ 的 MP 逆完全由如下性质刻画：

1. $\tau\tau^+\tau = \tau$；
2. $\tau^+\tau\tau^+ = \tau^+$；
3. $\tau\tau^+$ 厄米；
4. $\tau^+\tau$ 厄米。

♡

证明 该结论的充要性容易证明，其中必要性可通过 $(\tau\tau^+)v_i = v_i$ 以及设 $\tau v_i = \sum_j r_j u_j$ 证明。现在我们证明其唯一性：对于任意满足上述条件的 MP 逆 ρ 和 σ，

$$\rho = \rho\tau\rho = (\rho\tau)^*\rho = \tau^*\rho^*\rho = (\tau\sigma\tau)^*\rho^*\rho = (\sigma\tau)^*\tau^*\rho^*\rho = (\sigma\tau)(\rho\tau)^*\rho = \sigma\tau(\rho\tau\rho) = \sigma\tau\rho, \tag{B.2.1}$$

$$\sigma = \sigma\tau\sigma = \sigma(\tau\sigma)^* = \sigma\sigma^*\tau = \sigma\sigma^*(\tau\rho\tau)^* = \sigma\sigma^*\tau^*(\tau\rho)^* = \sigma(\tau\sigma)^*(\tau\rho) = (\sigma\tau\sigma)\tau\rho = \sigma\tau\rho. \tag{B.2.2}$$

可见 $\rho = \sigma$，即满足上述性质的 MP 逆唯一。

B.3 最小二乘逼近

MP 逆最重要的应用是最小二乘问题。考虑线性方程组 $Ax = v$，其中 $A \in \mathcal{M}_{m,n}(F)$，通常 $F = \mathbb{C}$ 或 $\mathbb{R}$。该方程组有解当且仅当 $v \in \mathrm{im}(\tau_A)$；若方程组无解，则求解方程 $Ax = \hat{v}$ 具有重要的实际意义，其中 $\hat{v}$ 是 $\mathrm{im}(\tau_A)$ 中按幺正距离或 Euclidean 距离唯一最接近 v 的向量（定理 5.9）。此问题称为*线性最小二乘*（*linear least squares*）问题，该方程组 $Ax = \hat{v}$ 的任意解称为方程组 $Ax = v$ 的*最小二乘解*（*least square solution*）。换言之，$Ax = v$ 的最小二乘解是使得 $\|Ax - v\|$ 最小化的向量 x。

设 w 和 z 是 $Ax = v$ 的最小二乘解，则 $Aw = \hat{v} = Az$，因此 $w - z \in \ker(\tau_A)$。这意味着如果 w 是某个特定的最小二乘解，那么所有最小二乘解构成的集合为 $w + \ker(A)$。在所有解中，我们最感兴趣的是具有极小范数的解。如果存在处于 $\ker(A)^\perp$ 内的最小二乘解 w，那么对于任意 $z \in \ker(A)$，我们有 $\|w + z\|^2 = \|w\|^2 + \|z\|^2 \geq \|w\|^2$。因此 w 将是唯一具有极小范数的最小二乘解。

定理 B.3

设 $A \in \mathcal{M}_{m,n}(F)$，那么在方程组 $Ax = \hat{v}$ 的所有最小二乘解中，存在唯一极小范数解 $A^+ v$，其中 A^+ 是 A 的 MP 逆。

♡

证明 向量 w 是最小二乘解当且仅当 $Aw = \hat{v} \in \mathrm{im}(\tau_A)$，那么根据定理 5.9，该条件等价于 $(Aw - v) \perp \mathrm{im}(\tau_A)$。根据命题 7.2，$\mathrm{im}(\tau_A)^\perp = \ker(\tau_A^*)$，因此上述条件等价于 $Aw - v \in \ker(\tau_A^*)$，即 $0 = \tau_A^*(Aw - v)$，整理得 $(A^*A)w = A^*v$。该方程称为 $Ax = v$ 的正规方程（normal equation），其解恰好为方程组 $Ax = v$ 的最小二乘解。

对于 V 中 AA^* 的本征向量构成的一组基 $\mathcal{C} = (v_1, \cdots, v_m)$，代入 $w = A^+ v_i$ 知 $(A^*A)(A^+ v_i) = s_i^{-1} AA^* u_i = s_i u_i = A^* v_i$。利用该方程的线性，可知 $(A^*A)(A^+ v) = A^* v$。根据定理 5.9，该解存在即唯一。

附录 C Schatten 范数

我们在第14章了解了内积空间 $\mathbb{F}^n$ 上的一些具体范数，并引出了算子范数的一种定义；事实上，量子信息论中涉及的算子也存在类似 p-范数的其他算子范数。本章我们主要关注这些算子范数在应用层面的结论，对其动机则不多作说明。在本章中，域 $\mathbb{F}$ 表示 $\mathbb{R}$ 或 $\mathbb{C}$。

目录

C.1 从向量范数到 Schatten 范数 145
C.2 范数的基本性质 145
C.3 特殊 Schatten 范数 147

C.1 从向量范数到 Schatten 范数

与 Euclidean 空间内的 p-范数相对应，这里我们最关心的范数族为 *Schatten p-范数*（*Schatten p-norm*），因为它包含了量子信息中最常见的谱范数（*spectral norm*）、*Frobenius* 范数（*Frobenius norm*）和迹范数（*trace norm*）。对于任意算子 $A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ 和实数 $p \in [1, \infty]$，A 的 Schatten p-范数定义为

$$\|A\|_p = \begin{cases} [\mathrm{Tr}(|A|^p)]^{1/p}, & \text{if } 1 \le p < \infty \\ \sup\{\|Au\|_2 | u \in \mathbb{F}^n, \|u\|_2 \le 1\}, & \text{if } p = \infty, \end{cases} \tag{C.1.1}$$

其中 $|A| = \sqrt{A^*A}$，而 $\|A\|_\infty$ 正是我们在第14.3节定义的算子范数。事实上，该范数之所以称为 p-范数，是因为其与算子 A 的奇异值构成的 r-元组的 p-范数紧密相关。

命题 C.1

设 $A \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$，$s(A)$ 为 A 的奇异值构成的 r-元组，那么 $\|A\|_p = \|s(A)\|_p$ 对所有 $1 \le p \le \infty$ 成立。

证明 当 $1 \le p < \infty$ 时，根据定理 B.1，任意算子 A 都存在奇异值分解 $A = P\Sigma Q^*$，其中 $\Sigma = \mathrm{diag}(s_1, \cdots, s_r)$ 且 $Q = (v_1, \cdots, v_m)$。结合第B.1节的讨论，v_i 是 $A^*A = Q\Sigma^2 Q^*$ 的本征向量，因此其存在谱分解 $A^*A = \sum_{i=1}^m s_i^2 v_i v_i^*$。根据第7.4节对于算子函数定义的讨论与命题 6.3，

$$|A|^p = (A^*A)^{p/2} = \sum_{i=1}^m s_i^p v_i v_i^* \quad \Longrightarrow \quad |A|^p v_i = s_i^p v_i \quad \Longrightarrow \quad \mathrm{Tr}(|A|^p) = \sum_{i=1}^r s_i^p, \tag{C.1.2}$$

可见 $\|A\|_p = (\sum_{i=1}^r s_i^p)^{1/p} = \|s(A)\|_p$，这里 $s(A)$ 表示 A 的奇异值构成的 r-元组。

当 $p = \infty$ 时，设 $P = (u_1, \cdots, u_n)$，那么将任意向量 $u \in \mathbb{F}^n$ 表示为 $\sum_i r_i u_i$，其中 $\sum_i |r_i|^2 = \|u\|_2^2 \le 1$。因此 $\|Au\|_2^2 = \|\sum_i r_i A u_i\|_2^2 = \|\sum_i r_i s_i v_i\|_2^2 = \sum_i |r_i| s_i^2 \le \max_i s_i^2$ 对所有 u 成立，因此 $\|A\|_\infty \le s_1$（此处 s_i 按非增顺序排列）；另一方面，$\|Au_1\|_2 = \|s_1 v_1\| = s_1$，因此 $\|A\|_\infty \ge s_1$。综上，$\|A\|_\infty = \max_i s_i = \|s(A)\|_\infty$。

C.2 范数的基本性质

定义 *Hilbert-Schmidt* 内积（*Hilbert-Schmidt inner product*）为 $\langle A, B\rangle = \mathrm{Tr}(B^*A)$（请读者自行验证该映射为内积）。在证明 Schatten 范数的基本性质之前，我们需要先证明一个引理 [5]。

引理 C.1 (von Neumann 迹不等式)

对于算子 $A, B \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$，$|\langle A, B\rangle| \le \sum_{i=1}^r s_i t_i$，其中 $s_1 \ge \cdots \ge s_a > 0 = s_{a+1}$ 和 $t_1 \ge \cdots \ge t_b > 0 = t_{b+1}$ 分别为 A 和 B 的奇异值，$r = \min\{a, b\}$。

证明 设 A 和 B 的奇异值分解分别为 $P_1\Sigma_1Q_1^*$ 和 $P_2\Sigma_2Q_2^*$。记幺正算子 $P=P_2^*P_1, Q=Q_2^*Q_1$，那么

$$\begin{aligned}\langle A,B\rangle &= \mathrm{Tr}(Q_2\Sigma_2^*P_2^*P_1\Sigma_1Q_1^*)=\mathrm{Tr}(\Sigma_2^*P\Sigma_1Q^*)=\sum_{i,j,k,l}(\Sigma_2^*)_{i,j}P_{j,k}(\Sigma_1)_{k,l}(Q^*)_{l,i}\\ &=\sum_{i,j,k,l}t_i\delta_{i,j}\cdot P_{j,k}\cdot s_k\delta_{k,l}\cdot\overline{Q_{i,l}}=\sum_{i,j}s_it_jP_{j,i}\overline{Q_{j,i}},\end{aligned}\tag{C.2.1}$$

因此 $|\langle A,B\rangle|\le\sum_{i,j}|P_{j,i}||Q_{j,i}|s_it_j\le\sum_{i,j}(|P_{j,i}|^2+|Q_{j,i}|^2)s_it_j/2$。我们希望证明 $\sum_{i,j}|P_{j,i}|^2s_it_j\le\sum_{i,j}\delta_{i,j}s_it_j$。

由于 P 幺正，$\sum_{i\le k}|P_{j,i}|^2\le\sum_i|P_{j,i}|^2=1$ 和 $\sum_{j\le l}|P_{j,i}|^2\le\sum_j|P_{j,i}|^2=1$ 对所有 i,j,k,l 成立，因此 $\sum_{i\le k}\sum_{j\le l}|P_{j,i}|^2\le\min\{k,l\}$。对于任意 $i\in\{1,\cdots,a\}$ 和 $j\in\{1,\cdots,b\}$，定义 $\xi_i=s_i-s_{i+1}\ge0,\eta_j=t_j-t_{j+1}\ge0$，即 $s_i=\sum_{k\ge i}\xi_k, t_j=\sum_{l\ge j}\eta_l$，从而

$$\begin{aligned}\sum_{i,j}(\delta_{i,j}-|P_{j,i}|^2)s_it_j&=\sum_{i,j}(\delta_{i,j}-|P_{j,i}|^2)\cdot\sum_{k\ge i}\xi_k\cdot\sum_{l\ge j}\eta_l=\sum_{k,l}\xi_k\eta_l\sum_{i\le k}\sum_{j\le l}(\delta_{i,j}-|P_{j,i}|^2)\\ &\ge\sum_{k,l}\xi_k\eta_l(\min\{k,l\}-\min\{k,l\})=0.\end{aligned}\tag{C.2.2}$$

因此 $\sum_{i,j}|P_{j,i}|^2s_it_j\le\sum_i s_it_i$，类似地 $\sum_{i,j}|Q_{j,i}|^2s_it_j\le\sum_i s_it_i$，故结论得证。

命题 C.2

设 $A\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$，实数 p,q 满足 $1\le p\le q\le\infty$，并假设其余算子规模恰当，那么

1. 范数等价性：$\|A\|_q\le\|A\|_p\le\mathrm{rank}(A)^{(q-p)/qp}\|A\|_q$；
2. 幺正不变性：对于任意幺正算符 U,V，$\|A\|_p=\|UAV^*\|_p$；
3. 范数对偶性：设 $p^*=p/(p-1)$ 为 p 的共轭指数，那么 Schatten p-范数与 p^*-范数是对偶的，即
$$\|A\|_p=\sup\left\{|\langle A,B\rangle|\middle|B\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m),\|B\|_{p^*}\le1\right\}.\tag{C.2.3}$$
特别地，对于一般的 $B\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$，Hölder 不等式 $|\langle A,B\rangle|\le\|A\|_p\|B\|_{p^*}$ 成立；
4. 次可乘性（submultiplicativity）：$\|AB\|_p\le\|A\|_\infty\|B\|_p$ 且 $\|AB\|_p\le\|A\|_p\|B\|_\infty$。
5. $\|A\|_p=\|A^*\|_p=\left\|A^T\right\|_p=\left\|\bar A\right\|_p$。

证明

1. 根据命题 C.1和定理 14.4，$\|A\|_q=\|s(A)\|_q\le\|s(A)\|_p=\|A\|_p\le\mathrm{rank}(A)^{(q-p)/qp}\|A\|_q$；
2. 根据命题 C.1和奇异值分解，A 左乘和右乘幺正算符并不影响其奇异值，从而范数不变；
3. 我们先证明 Hölder 不等式。根据 von Neumann 迹不等式（引理 C.1）、Hölder 不等式（定理 14.2）和命题 C.1，$|\langle A,B\rangle|\le\|s(A)\|_p\|s(B)\|_{p^*}=\|A\|_p\|B\|_{p^*}$。当 $\|B\|_{p^*}\le1$ 时，$|\langle A,B\rangle|\le\|A\|_p$ 恒成立，因此
$$\sup\left\{|\langle A,B\rangle|\middle|B\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m),\|B\|_{p^*}\le1\right\}\le\|A\|_p.\tag{C.2.4}$$
反之我们希望找到一个 $B_0\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ 使得 $|\mathrm{Tr}(B_0^*A)|=\|A\|_p=[\mathrm{Tr}(|A|^p)]^{1/p}$ 在 $\|B_0\|_{p^*}\le1$ 时成立。当 $1<p<\infty$ 时，对比 $\mathrm{Tr}(\cdot)$ 中的项，不妨设 $B_0=cA|A|^{p-2}$，其中算子指数的定义参考第7.4节，常数 c 可由下式确定：
$$1=\|B_0\|_{p^*}=\left[\mathrm{Tr}\left(|B_0|^{p^*}\right)\right]^{1/p^*}=|c|\left\{\mathrm{Tr}\left[\left(|A|^{p-1}\right)^{p/(p-1)}\right]\right\}^{1-1/p}=|c|\|A\|_p^{p-1}.\tag{C.2.5}$$
取 $c=\|A\|_p^{1-p}$，则 $|\mathrm{Tr}(B_0^*A)|=\|A\|_p^{1-p}\cdot[\mathrm{Tr}\,(|A|^p)]=\|A\|_p$，因此反向不等式成立。当 $p=1$ 时，设 A 的奇异值分解为 $P\Sigma Q^*$，那么取 $B_0=PQ^*$，使得 $\mathrm{Tr}(B_0^*A)=\mathrm{Tr}(QP^*P\Sigma Q^*)=\mathrm{Tr}(\Sigma)=\mathrm{Tr}\sqrt{(Q\Sigma Q^*)^2}=\|A\|_1$，因此反向不等式成立。当 $p=\infty$ 时，设 A 的最大奇异值 s_1 对应的左右奇异向量分别为 v_1 和 u_1，那么根据定理 B.1和命题 C.1，取 $B_0=u_1v_1^*$，使得 $\mathrm{Tr}(B_0^*A)=\mathrm{Tr}(u_1v_1^*A)=v_1^*Au_1=s_1v_1^*v_1=s_1=\|s(A)\|_\infty=\|A\|_\infty$，因此反向不等式成立。综上，该结论成立；
4. 我们先证明一个引理：$s_i(AB)\le\|B\|_\infty s_i(A)$。设 B 的奇异值 s_i 对应的左奇异向量为 v_i，则 $BB^*=\sum_i s_iv_iv_i^*$，进而 $\|B\|_\infty^2\mathrm{id}-BB^*=\sum_i(s_1^2-s_i^2)v_iv_i^*\ge0$；这意味着 $(AB)(AB^*)\le\|B\|_\infty^2AA^*$ 对任意 $A\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ 成立。由于 A 的奇异值恰好为 AA^* 或 A^*A 本征值的正平方根，上式给出 $s_i(AB)\le$

$\|B\|_\infty s_i(A)$。因此，该引理和命题 C.1表明，当 $1\le p<\infty$ 时，$\|AB\|_p=\|s(AB)\|_p=(\sum_i s_i(AB)^p)^{1/p}\le\|B\|_\infty(\sum_i s_i(A)^p)^{1/p}=\|B\|_\infty\|A\|_p$；当 $p=\infty$ 时，$s_i(AB)\le\|B\|_\infty s_i(A)\le\|B\|_\infty\|A\|_\infty$ 对所有 i 成立，因此 $\|AB\|_\infty=\|s(AB)\|_\infty=\sup_i s_i(AB)\le\|B\|_\infty\|A\|_\infty$。综上，$\|AB\|_p\le\|B\|_\infty\|A\|_p$ 对所有 $1\le p\le\infty$ 成立；对称地，$\|AB\|_p\le\|A\|_\infty\|B\|_p$。根据此结论和性质 1，$\|AB\|_p\le\|A\|_p\|B\|_\infty\le\|A\|_p\|B\|_p$；

5. 根据性质 3，$\|A^*\|_p=\sup\{|\langle A^*,B\rangle|\,|\,\|B\|_{p^*}\le1\}=\sup\{|\langle B^*,A\rangle|\,|\,\|B^*\|_{p^*}\le1\}=\|A\|_p$。设 A 的奇异值 s 对应的右奇异向量为 u，那么 $(A^*A)u=s^2u$，逐分量表示为 $s^2u_i=(A^*Au)_i=\sum_j(A^*A)_{i,j}u_j$。由于 $\overline{(A^*A)_{i,j}}=\sum_k\overline{(A^*)_{i,k}}\cdot\overline{A_{k,j}}=\sum_k A_{k,i}\overline{A_{k,j}}=(\bar A^*\bar A)_{i,j}$，那么对其取复共轭可得 $s^2\overline{u_i}=\sum_{i,j}\overline{(A^*A)_{i,j}}\,\overline{u_j}=(\bar A^*\bar A\bar u)_i$，即 s 也是 $\bar A^*\bar A$ 的奇异值。根据命题 C.1，$\|\bar A\|_p=\|s(\bar A)\|_p=\|s(A)\|_p=\|A\|_p$，进而 $\|A^T\|_p=\|(\bar A)^*\|_p=\|A\|_p$。

有了上述性质，现在我们可以证明 Schatten p-范数确为范数，以此结束本节。

定理 C.1

设 $A,B\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$，那么 $\|A+B\|_p\le\|A\|_p+\|B\|_p$ 对所有 $1\le p\le\infty$ 成立。 ♡

证明 根据命题 C.2性质 3，对于任意正数 ϵ，存在 $C\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$，使得 $\|C\|_{p^*}\le1$ 且 $\|A+B\|_p-\epsilon<|\langle A+B,C\rangle|=|\langle A,C\rangle+\langle B,C\rangle|\le|\langle A,C\rangle|+|\langle B,C\rangle|\le\|A\|_p+\|B\|_p$。利用 ϵ 的任意性，$\|A+B\|_p\le\|A\|_p+\|B\|_p$ 成立。

C.3 特殊 Schatten 范数

在量子信息论中，常见的三种 Schatten 范数对应 $p=1,2,\infty$，分别称为迹范数（*trace form*）、*Frobenius* 范数（*Frobenius norm*）和谱范数（*spectral norm*）或算子范数（*operator norm*）。除去命题 C.2给出的普遍性质，这三种范数还有一些额外性质。

C.3.1 迹范数

$A\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ 的迹范数定义为 $\|A\|_1=\operatorname{Tr}|A|=\operatorname{Tr}\sqrt{A^*A}$，它常用于定义密度算符 ρ 和 σ 的迹距离（*trace distance*）：$\|\rho-\sigma\|_1$。对于任意单位向量 u,v 和非负实数 α,β，设 $\alpha uu^*-\beta vv^*$ 的本征值向量为 $u+xv$，那么

$$(\alpha uu^*-\beta vv^*)(u+xv)=\alpha(1+cx)u-\beta(x+\bar c)v=\lambda(u+xv)\quad\Longrightarrow\quad x=\frac{-(\alpha+\beta)\pm\sqrt{\Delta}}{2\alpha c},\tag{C.3.1}$$

其中 $\Delta=(\alpha+\beta)^2-4\alpha\beta|\langle u,v\rangle|^2$，故本征值 $\lambda_\pm=[(\alpha-\beta)\pm\sqrt{\Delta}]/2$。由于 $\alpha uu^*-\beta vv^*$ 是自伴算子，根据第7.4节的讨论，$|\alpha uu^*-\beta vv^*|=\sqrt{(\alpha uu^*-\beta vv^*)^2}$ 的本征值为 $|\lambda_\pm|$，进而

$$\begin{aligned}\|\alpha uu^*-\beta vv^*\|_1&=\operatorname{Tr}|\alpha uu^*-\beta vv^*|=|\lambda_+|+|\lambda_-|=\sqrt{(\lambda_++\lambda_-)^2+2(|\lambda_+\lambda_-|-\lambda_+\lambda_-)}\\&=\sqrt{(\alpha-\beta)^2+2\cdot2\alpha\beta(1-|\langle u,v\rangle|^2)}=\sqrt{(\alpha+\beta)^2-4\alpha\beta|\langle u,v\rangle|^2}.\end{aligned}\tag{C.3.2}$$

C.3.2 Frobenius 范数

$A\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ 的 Frobenius 范数定义为 $\|A\|_2=\sqrt{\operatorname{Tr}|A|^2}=\sqrt{\operatorname{Tr}(A^*A)}=\langle A,A\rangle$。该范数的优点在于具有关于 A 的逐分量表示，不必关心其本征值或奇异值：

$$\|A\|_2^2=\operatorname{Tr}(A^*A)=\sum_i(A^*A)_i=\sum_i(A^*)_{i,j}A_{j,i}=\sum_{i,j}|A_{i,j}|^2\quad\Longrightarrow\quad\|A\|_2=\|(A_{i,j})\|_2.\tag{C.3.3}$$

C.3.3 谱范数

$A\in\mathcal{L}(\mathbb{F}^n,\mathbb{F}^m)$ 的谱范数定义为 $\|A\|_\infty=\sup\{\|Au\|_2|u\in\mathbb{F}^n,\|u\|_2\le1\}$，其主要性质我们已经在第14.3节进行了详细讨论。需要补充的是，根据命题 C.2，$\|A^*A\|_\infty\le\|A^*\|_\infty\|A\|_\infty=\|A\|_\infty^2$；反之，对于任意正数 ϵ，存在 $u\in\mathbb{F}^n$，使得 $\|u\|_2\le1$ 且 $\|Au\|_2-\epsilon<\|Au\|_2$。根据 Hölder 不等式（命题 C.2），$\|Au\|_2^2=u^*A^*Au=\operatorname{Tr}(A^*Auu^*)\le\|A^*A\|_\infty\|uu^*\|_1=\|A^*A\|_\infty\|u\|_2=\|A^*A\|_\infty$。由 ϵ 的任意性，$\|A\|_\infty^2\le\|A^*A\|_\infty$。类似地 $\|AA^*\|_\infty^2=\|A\|_\infty^2$。

参考文献

[1] Steven Roman. *Advanced linear algebra*. Springer, 2005.

[2] Daniel Hug, Wolfgang Weil, et al. *Lectures on convex geometry*. Vol. 286. Springer, 2020.

[3] Donald L Cohn. *Measure theory*. Vol. 2. Springer, 2013.

[4] Daniel Daners. “Introduction to functional analysis”. In: *The University of Sydney* (2008).

[5] Leon Mirsky. “A trace inequality of John von Neumann”. In: *Monatshefte für mathematik* 79.4 (1975), pp. 303–306.