测度论与概率论基础

学习笔记

何憾[1]

版本 1.0.0 Beta

[1]访问我的个人主页, 查看其他笔记

目录

第一章　测度与测度空间

简而言之, 测度论可以理解为在抽象空间建立类似于实变函数中测度, 积分和导数那样的分析系统.

测度, 对我们来说并不陌生, 像线段的长度, 平面上某些曲线围成的面积和容器的容积等都是测度. 但是仅仅讨论这些由直接经验建立的测度是远远不够的, 例如, 概率从抽象角度看是对形形色色的事件发生的可能性进行测量, 因而只有在抽象空间的集合上建立了测度, 才有可能真正解决概率论的问题, 在抽象空间建立测度是没有什么直接经验可循的, 只能采用公理化方法, 当然, 归根结底, 公理化方法中的那些公理也是从实际中提炼出来的.

1.1　集合系

以空间 X 中的一些集合为元素组成的集合称为 X 上的集合系. 为了在集合 X 上建立测度, 必须确定出一些可测集, 而这些可测集的全体就组成了一个集合系. 下面, 我们先引入在抽象空间确定可测集时所必须引进的一些特殊集合系. 若未特别声明, 我们总是假定讨论的是全空间 X 上的集合系.

1.1.1　一些常用的集合系

定义 1. 称集合系 $\mathcal{P}$ 为 π **系** (π-system), 若其对有限交运算封闭. 换言之, 对任意 $A, B \in \mathcal{P}$, 必有

$$A \cap B \in \mathcal{P}\,.$$

定义 2. 称集合系 $\mathcal{Q}$ 为 **半环** (semiring), 若 $\mathcal{Q}$ 为 π 系, 且对任何 $A, B \in \mathcal{Q}$, 存在有限个互不相交的集合 $C_1, C_2, \cdots, C_n \in \mathcal{Q}$, 使得

$$A \backslash B = \bigcup_{j=1}^{n} C_j\,.$$

注. 观察到 $A\backslash B = A\backslash(A\cap B)$. 故若想验证 π 系 $\mathcal{Q}$ 为半环, 只需验证对任意 $A,B\in\mathcal{Q}$, 且 $B\subset A$, 存在有限个互不相交的集合 $C_1,C_2,\cdots,C_n\in\mathcal{Q}$, 使得 $A\backslash B=\cup_{j=1}^n C_j$.

例 1. $\mathbb{R}$ 上集合系

$$\mathcal{P}_{\mathbb{R}} \coloneqq \{(-\infty,a] : a\in\mathbb{R}\}$$

是一个 π 系. 而

$$\mathcal{Q}_{\mathbb{R}} \coloneqq \{(a,b] : a,b\in R\}$$

是一个半环.

定义 3. 称集合系 $\mathcal{R}$ 为 **环** (ring), 若 $\mathcal{R}$ 对有限并运算和差运算封闭. 换言之, 对任何 $A,B\in\mathcal{R}$, 总有

$$A\cup B\in\mathcal{R};\ A\backslash B\in\mathcal{R}.$$

注. 观察到 $A\cup B=A\cup(B\backslash A)$. 故若验证 $\mathcal{R}$ 为环, 只需验证 $\mathcal{R}$ 有限不交并运算和差运算封闭.

命题 1. *环必是半环. 换言之, 环对交, 并, 差运算皆封闭.*

证明. 只需证环 $\mathcal{R}$ 必是 π 系. 对任何 $A,B\in\mathcal{R}$, 注意

$$A\cap B=(A\cup B)\backslash(A\backslash B\cup B\backslash A).$$

于是知 $A\cap B\in\mathcal{R}$. □

定义 4. 称集合系 $\mathcal{A}$ 为 **代数** (algebra) 或者 **域** (field), 若 $\mathcal{A}$ 是包含全空间 X 的环.

注. 由上述定义知, 代数对交, 并, 差, 补运算皆封闭. 此外, 注意到 $A\cup B=(A^c\cap B^c)^c$, $A\backslash B=A\cap B^c$. 故 $\mathcal{A}$ 是代数, 当且仅当 $X\in\mathcal{A}$, 且 $\mathcal{A}$ 对补运算和有限交运算封闭.

定义 5. 称集合系 $\mathcal{M}$ 为 **单调类** (monotone class), 若 $\mathcal{M}$ 对单调集列的极限运算封闭.

定义 6. 称集合系 $\mathcal{D}$ 是 λ **系** (λ system) [1], 若 $\mathcal{D}$ 满足

[1]有些书上也称为 **d 系**或 Dynkin system .

1. $X \in \mathcal{D}$,
2. $A, B \in \mathcal{D}$ 且 $B \subset A$, 则 $A \backslash B \in \mathcal{D}$,
3. $\{A_n\}$ 是 $\mathcal{D}$ 中的递增集列, 则 $\bigcup_{n=1}^{\infty} A_n \in D$.

注. 除了上述常用定义外, λ 系还有一等价定义: $\mathcal{D}$ 是 λ 系当且仅当 $X \in \mathcal{D}$, 且 $\mathcal{D}$ 对补运算和可列不交并运算封闭.

命题 2. *λ 系必是单调类.*

证明. 由定义知 λ 系 $\mathcal{D}$ 对递增集列的极限运算封闭. 若 $\{A_n\}$ 是递减集列, 则 $\{A_n^c\}$ 是递增的, 于是

$$\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{D}.$$

故知 λ 系 $\mathcal{D}$ 是单调类. □

定义 7. 称集合系 $\mathcal{R}$ 是 σ **环**, 若 $\mathcal{R}$ 对差运算和可列并运算封闭.

注. 实际上要验证 $\mathcal{R}$ 为 σ 环, 只需验证 $\mathcal{R}$ 对差运算和可列不交并运算和封闭.

对 σ 环 $\mathcal{R}$ 中的集列 $\{A_n\}$, 注意到

$$A_1 \backslash \bigcap_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} (A_1 \backslash A_n) \in \mathcal{R}.$$

于是知

$$\bigcap_{n=1}^{\infty} A_n = A_1 \backslash \left(A_1 \backslash \bigcap_{n=1}^{\infty} A_n\right) \in \mathcal{R}.$$

即 σ 环 $\mathcal{R}$ 对可列交运算也封闭.

定义 8. 称 σ 环 $\mathcal{F}$ 是 σ **代数**或 σ **域**, 若 $\mathcal{F}$ 包含全空间 X.

注. $\mathcal{F}$ 是 σ 代数当且仅当 $X \in \mathcal{F}$, 且 $\mathcal{F}$ 对补运算和可列交运算封闭.

由上述定义可知, σ 代数对交, 并, 差, 补, 可列交, 可列并运算皆封闭. 可见 σ 代数是代数, 且是 λ 系. 相反的, 有下面的定理.

定理 3. *$\mathcal{F}$ 是一个集合系, 则下面三个命题等价:*

1. $\mathcal{F}$ 为 σ 代数.

2. $\mathcal{F}$ *既是单调系又是域.*

3. $\mathcal{F}$ *既是* λ *系又是* π *系.*

证明. 1 蕴涵 2,3 已证, 先证 2 蕴涵 1. 任取 $\mathcal{F}$ 中集列 $\{A_n\}$, 令 $B_1 = A_1$,

$$B_n = A_n \setminus \bigcup_{j=1}^{n-1} A_j \,, n = 2, 3, \cdots$$

由 $\mathcal{F}$ 为域知 $B_n \in \mathcal{F}$, $\forall n \in \mathbb{N}^+$. 且 $\{B_n\}$ 为递增集列. 于是

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{F}\,.$$

故 $\mathcal{F}$ 是 σ 代数.

再证 3 蕴涵 1. 由 $\mathcal{F}$ 是 π 系知, 其对有限交运算封闭. 由于 $\mathcal{F}$ 是 λ 系知其为单调系, 且含全空间 X, 对补运算封闭. 故 $\mathcal{F}$ 是代数又是单调系. 故 3 蕴涵 2, 自然蕴含 1. □

总结上面的讨论, 我们得到下面的关系:

$$\pi \text{ 系} \to \text{ 半环} \to \text{环} \to \text{ 代数} \to \sigma \text{ 代数}$$
$$\text{单调系} \to \lambda \text{ 系} \to \sigma \text{ 代数}$$

1.1.2 生成集合系

上面我们没有给出环, 域, 单调系, λ 系, σ 代数的具体了例子, 因为这些是较为复杂的集合系. 我们自然可以考虑由简单的集合系生成复杂的集合系. 首先来严格定义一下怎样算是 "生成".

定义 9. 称 $\mathcal{R}$ 是由集合系 $\mathcal{E}$ **生成**的环 (单调系,λ 系,σ 代数), 若 $\mathcal{R}$ 是包含 $\mathcal{E}$ 的**最小的环** (或单调系, λ 系, σ 代数). 换言之,

1. $\mathcal{R}$ 是包含 $\mathcal{E}$ 的环 (或单调系, λ 系, σ 代数).

2. 对任何包含 $\mathcal{E}$ 的环 (或单调系, λ 系, σ 代数) $\mathcal{R}'$, 皆有 $\mathcal{R} \subset \mathcal{R}'$.

注. 容易证明由集合系 $\mathcal{E}$ 生成的环 (单调系, λ 系,σ 代数) 皆是存在的. 这是因为任意多个环 (单调系, λ 系, σ 代数) 的交仍然是环 (单调系, λ 系, σ 代数). 于是**集合系 $\mathcal{E}$ 生成的环 (域, 单调系, λ 系, σ 代数) 恰好是所有包含 $\mathcal{E}$ 的环 (单调系, λ 系, σ 代数) 的交.**

为了方便, 我们把集合系 $\mathcal{E}$ 生成的环, 单调系, λ 系, σ 代数分别记为 $r(\mathcal{E}), m(\mathcal{E}), l(\mathcal{E}), \sigma(\mathcal{E})$. 本节的主要结论是下面三个定理.

定理 4. *若 $\mathcal{Q}$ 是一个半环, 则其生成的环为*

$$r(\mathcal{Q}) = \left\{ \bigcup_{j=1}^{n} C_j : C_1, \cdots, C_n \in \mathcal{Q} \text{ 且两两不交}; n \in \mathbb{N}^+ \right\}. \tag{1.1}$$

证明. 只需证明 (1.1) 右边, 记为 $\mathcal{R}$, 是一个环, 显然对 $\mathcal{R}$ 有限交运算封闭, 故 $\mathcal{R}$ 是一个 π 系. 同样容易看出 $\mathcal{R}$ 对不交并运算封闭. 现只需证明其对差运算封闭.

设 $\{A_i\}_{i=1}^n, \{B_j\}_{j=1}^m$ 皆是 $\mathcal{Q}$ 中互不相交的集列, 注意

$$\bigcup_{i=1}^{n} A_i \backslash \left(\bigcup_{j=1}^{m} B_j \right) = \bigcup_{i=1}^{n} \bigcap_{j=1}^{m} A_i \backslash B_j \, .$$

对任意 $1 \leqslant i \leqslant n$, 和 $1 \leqslant j \leqslant m$, 成立

$$A_i \backslash B_j \in \mathcal{R}$$

再由 $\mathcal{R}$ 对有限交运算和不交并运算封闭得到 $A \backslash B \in \mathcal{R}$. 故 $\mathcal{R}$ 其对差运算封闭. □

定理 5 (单调类定理)**.** *若 $\mathcal{A}$ 是域, 则 $\mathcal{A}$ 生成的单调类*

$$m(\mathcal{A}) = \sigma(\mathcal{A}) \, .$$

证明. 由定理 3 知, 只需证明 $m(\mathcal{A})$ 是域. 又因 $X \in \mathcal{A} \subset m(\mathcal{A})$, 只需证任何 $A, B \in m(\mathcal{A})$, 有

$$A \backslash B \in m(\mathcal{A}) \, , \; A \cup B \in m(\mathcal{A}) \, .$$

现在任取 $A \in \mathcal{A}$, 令

$$\mathcal{G}_A = \{ B \in m(\mathcal{A}) : A \backslash B \in m(\mathcal{A}), A \cup B \in m(\mathcal{A}) \} \, .$$

现在显然有 $\mathcal{A} \subset \mathcal{G}_A$. 注意到 $\mathcal{G}_A$ 是一单调类 , 故 $m(\mathcal{A}) \subset \mathcal{G}_A$. 因此我们得到

$$m(\mathcal{A}) = \mathcal{G}_A \, .$$

即任何 $A \in \mathcal{A}, B \in m(\mathcal{A})$ 有

$$A \backslash B \in m(\mathcal{A}) \, , \; A \cup B \in m(\mathcal{A}) \, .$$

再任取 $B \in m(\mathcal{A})$, 令

$$\mathcal{G}_B = \{A \in m(\mathcal{A}) : A \backslash B \in m(\mathcal{A}), A \cup B \in m(\mathcal{A})\}.$$

我们已经得到 $\mathcal{A} \subset \mathcal{G}_B$. 注意到 $\mathcal{G}_B$ 是一单调类 , 故 $m(\mathcal{A}) \subset \mathcal{G}_B$. 因此我们得到

$$m(\mathcal{A}) = \mathcal{G}_B.$$

即任何 $A \in m(\mathcal{A}), B \in m(\mathcal{A})$ 有

$$A \backslash B \in m(\mathcal{A}),\ A \cup B \in m(\mathcal{A}). \qquad \square$$

定理 6 (Dynkin's $\pi - \lambda$ theorem)**.** *若 $\mathcal{P}$ 是 π 系, 则 $\mathcal{P}$ 生成的 λ 系为*

$$l(\mathcal{P}) = \sigma(\mathcal{P}).$$

证明. 由定理 3 知, 只需证明 $l(\mathcal{P})$ 是 π 类. 现在任取 $A \in \mathcal{P}$, 令

$$\mathcal{G}_A = \{B \in l(\mathcal{P}) : A \cap B \in l(\mathcal{P})\}.$$

显然 $\mathcal{P} \subset \mathcal{G}_A$. 注意 $\mathcal{G}_A$ 也是 λ 系 . 故 $l(\mathcal{P}) \subset \mathcal{G}_A$, 因此

$$l(\mathcal{P}) = \mathcal{G}_A.$$

即任何 $A \in \mathcal{P}, B \in l(\mathcal{P})$ 有

$$A \cap B \in l(\mathcal{P}).$$

再任取 $B \in l(\mathcal{P})$, 令

$$\mathcal{G}_B = \{A \in l(\mathcal{P}) : A \cap B \in l(\mathcal{P})\}.$$

已证 $\mathcal{P} \subset \mathcal{G}_B$. 注意 $\mathcal{G}_B$ 也是 λ 系 . 故 $l(\mathcal{P}) \subset \mathcal{G}_B$, 因此

$$l(\mathcal{P}) = \mathcal{G}_B.$$

即任何 $A \in l(\mathcal{P}), B \in l(\mathcal{P})$ 有

$$A \cap B \in l(\mathcal{P}). \qquad \square$$

例 2. 由简单集合系生成 σ 代数的一个重要例子是

$$\mathcal{B}_{\mathbb{R}} := \sigma(\mathcal{P}_{\mathbb{R}}) = \sigma(\mathcal{Q}_{\mathbb{R}}).$$

称为 $\mathbb{R}$ 上的 **Borel 代数**，其中的集合称为 $\mathbb{R}$ 中的 **Borel 集**，以 $\mathcal{O}_{\mathbb{R}}$ 记由 $\mathbb{R}$ 中开集组成的集合系, 容易证明 $\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{O}_{\mathbb{R}})$.

1.2 测度的定义与性质

本节我们先来介绍抽象空间上测度的公理化定义, 进而通过公理化定义研究研究测度的性质.

给定空间 X 上的集合系 $\mathcal{E}$, 且 $\varnothing \in \mathcal{E}$. 称定义在集合系 $\mathcal{E}$ 上, 取值于 $[0,\infty]$ 的函数为 $\mathcal{E}$ 上的 **非负集函数**.

定义 10. $\mathcal{E}$ 上非负集函数 μ 称为 $\mathcal{E}$ 上的**测度** (measure), 若 $\mu(\varnothing)=0$ 且 μ 满足**可列可加性**: 对 $\mathcal{E}$ 中任何两两不交的集列 $\{A_n\}$, 且 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{E}$, 有

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)\,.$$

称 μ 为**有限**的, 若 $\mu(A)<\infty$ 对任意 $A\in\mathcal{E}$ 成立; 称 μ 为 σ **有限**的, 若任意 $A\in\mathcal{E}$, 存在 $\mathcal{E}$ 中集列 $\{B_n\}$, 满足 $A\subset \cup_{n=1}^{\infty} B_n$ 且 $\mu(B_n)<\infty$ 对任意正整数 n 皆成立.

注. 在许多书籍文献上, 谈及测度, 一般指的其定义域为某个 σ 域; 当定义域不是 σ 域, 而是某个半环/环/域, 一般称为称为预测度 (pre-measure). 但是下一节我们会看到, 我们总可以将其定义延拓到该半环生成的 σ 域上.

由测度的可列可加性和在空集处取零值很容易得到下面的命题.

命题 7. *μ 是 $\mathcal{E}$ 上的测度, 则 μ 具有***有限可加性***: $\mathcal{E}$ 中任意有限个互不相交的集合 $\{A_j\}_{j=1}^{n}$, 若 $\bigcup_{j=1}^{n} A_j \in \mathcal{E}$ 则*

$$\mu(\bigcup_{j=1}^{n} A_j) = \sum_{j=1}^{n} \mu(E_j)\,.$$

由有限可加性质容易得到下面的简单推论.

推论 8. *μ 是 $\mathcal{E}$ 上的测度, 则 μ 具有***可减性*** : $\mathcal{E}$ 中任意两集合 A,B , $B\subset A$ 且 $A\backslash B\in\mathcal{E}$, 若 $\mu(B)<\infty$, 则*

$$\mu(A\backslash B) = \mu(A)-\mu(B)\,.$$

现在, 让我们看一些测度的具体例子.

例 3. X 是一非空集合, $\mathcal{F}$ 是由 X 的一切子集组成的集合系. 以 $|A|$ 记集合 A 中元素的个数, 并令 $\mu(A)=|A|$, $\forall A\in\mathcal{F}$. 则 μ 是 $\mathcal{F}$ 上的测度, 称为 **计数测度**. 若 X 是有限集，则 μ 是有限测度; 若 X 是可列集, 则 μ 是 σ 有限测度.

例 4. X 是一集合, $\mathcal{F}$ 是 X 上的 σ 域. $x \in X$ 是一给定元素. 对每个 $A \in \mathcal{F}$, 令

$$\delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & x \in A \end{cases}$$

则 δ_x 是 $\mathcal{F}$ 上的测度. 更进一步, 若 $x_1, x_2, \cdots, x_n \in X$, 则

$$\mu = \sum_{i=1}^{n} \delta_{x_i}$$

仍是 $\mathcal{F}$ 上的测度. 这二者都属于**点测度**.

为了方便进一步研究测度的性质, 我们需要使集合系 $\mathcal{E}$ 对集合交, 并, 差运算有适当的封闭性. 于是下面我们来讨论环 $\mathcal{R}$ 或 σ 域 $\mathcal{F}$ 上的测度.

命题 9. *μ 是环 $\mathcal{R}$ 上的测度. 则 μ 具有***单调性***: $\mathcal{E}$ 中任意两集合 A, B , 且 $B \subset A$, 则*

$$\mu(B) \leqslant \mu(A)\,.$$

证明. 因 $A \backslash B \in \mathcal{R}$, 由测度的有限可加性可得

$$\mu(A) = \mu(B) + \mu(A \backslash B)\,.$$

因此 $\mu(B) \leqslant \mu(A)$. □

命题 10. *μ 是环 $\mathcal{R}$ 上的测度. 则 μ 具有***次可列可加性***: 任意 $\mathcal{R}$ 中集列 $\{A_n\}$, 若 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$, 则*

$$\mu(\bigcup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} \mu(A_n).$$

证明. 考虑集列 $\{B_n\}$, 其中 $B_1 = A_1$,

$$B_n = A_n \backslash \bigcup_{j=1}^{n-1} A_j\,,\ n \geqslant 2\,.$$

可见 $\{B_n\}$ 是 $\mathcal{R}$ 中两两无交的集列. 且对任何正整数 n 有 $B_n \subset A_n$, 因而 $\mu(B_n) \leqslant \mu(A_n)$. 同时 $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. 于是有

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) \leqslant \sum_{n=1}^{\infty} \mu(A_n)\,.$$ □

命题 11. μ 是环 $\mathcal{R}$ 上的测度. 则 μ 具有**下连续性**: $\mathcal{R}$ 中递增集列 $\{A_n\}$, 若 $\lim A_n \in \mathcal{R}$, 则

$$\mu(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$$

和**上连续性**: $\mathcal{R}$ 中递减集列 $\{A_n\}$, 若 $\lim A_n \in \mathcal{R}$ 且 $\mu(A_1) < \infty$, 则

$$\mu(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mu(A_n)\,.$$

证明. 先证下连续性. 由 $\{A_n\}$ 是递增集列, 知极限 $\lim \mu(A_n)$ 存在, 且有

$$\lim_{n\to\infty} \mu(A_n) \leqslant \mu(\lim_{n\to\infty} A_n)\,.$$

故不妨设 $\lim \mu(A_n) < \infty$. 考虑集列 $\{B_n\}$, 其中 $B_1 = A_1$,

$$B_n = A_n \backslash A_{n-1}\,,\ n \geqslant 2\,.$$

则 $\{B_n\}$ 是 $\mathcal{R}$ 中集列, 且互不相交, 且有 $\bigcup_{n=1}^{\infty} B_n = \lim A_n$. 当 $n \geqslant 2$ 时, 由 $\lim \mu(A_n) < \infty$ 得到 $\mu(B_n) = \mu(A_n) - \mu(A_{n-1})$, 于是

$$\begin{aligned}
\mu(\lim_{n\to\infty} A_n) &= \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n) \\
&= \sum_{n=2}^{\infty} \mu(A_n) - \mu(A_{n-1}) + \mu(A_1) \\
&= \lim_{n\to\infty} \mu(A_n)\,.
\end{aligned}$$

再证上连续性. 易见极限 $\lim \mu(A_n)$ 存在. 考虑集列 $\{B_n\}$, 其中

$$B_n = A_n \backslash A_{n+1}\,,\ n \geqslant 1\,.$$

则集列 $\{B_n\}$ 是互不相交的, 且 $\bigcup_{n=1}^{\infty} B_n = A_1 \backslash \lim A_n$. 于是有

$$\mu(A_1) - \mu(\lim_{n\to\infty} A_n) = \mu(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu(B_n)$$
$$= \sum_{n=1}^{\infty} \mu(A_n) - \mu(A_{n+1}) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n)\,.$$

于是得到 $\mu(\lim A_n) = \lim \mu(A_n)$. □

推论 12. μ 是 σ 域 $\mathcal{F}$ 上的测度. 则 μ 具有**下半连续性**: 对 $\mathcal{F}$ 中任意集列 $\{A_n\}$,

$$\mu(\varliminf_{n\to\infty} A_n) \leqslant \varliminf_{n\to\infty} \mu(A_n)$$

和**上半连续性**: 对 $\mathcal{F}$ 中任意集列 $\{A_n\}$, 且 $\mu(\bigcup_{n=1}^{\infty} A_n) < \infty$, 则

$$\mu(\varlimsup_{n\to\infty} A_n) \geqslant \varlimsup_{n\to\infty} \mu(A_n)$$

证明. μ 在 σ 域 $\mathcal{F}$ 上定义, 保证了下面出现的集合皆是可测的. 我们先证明下半连续性, 这是测度的下连续性的推论. 对任意 $\mathcal{F}$ 中集列 $\{A_n\}$,

$$\begin{aligned}
\mu(\varliminf_{n\to\infty} A_n) &= \mu(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k) \\
&= \lim_{n\to\infty} \mu(\bigcap_{k=n}^{\infty} A_k) \\
&\leqslant \varliminf_{n\to\infty} \mu(A_n)\,.
\end{aligned}$$

同样, 上半连续性也可用上连续性证明. 对 $\mathcal{F}$ 中任意集列 $\{A_n\}$, 且 $\mu(\bigcup_{n=1}^{\infty} A_n) < \infty$,

$$\begin{aligned}
\mu(\varlimsup_{n\to\infty} A_n) &= \mu(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k) \\
&= \lim_{n\to\infty} \mu(\bigcup_{k=n}^{\infty} A_k) \\
&\geqslant \varlimsup_{n\to\infty} \mu(A_n)\,.
\end{aligned}$$

□

推论 13 (Borel-Cantelli 引理)**.** *μ 是 σ 域 $\mathcal{F}$ 上的测度. 对 $\mathcal{F}$ 中任意集列 $\{A_n\}$, 若*

$$\sum_{n=1}^{\infty} \mu(A_n) < \infty\,,$$

则有

$$\mu(\varlimsup_{n\to\infty} A_n) = 0\,.$$

证明. 注意 $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, 于是

$$\begin{aligned}
\mu(\varlimsup_{n\to\infty} A_n) = \mu(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k) = \lim_{n\to\infty}\mu(\bigcup_{k=n}^{\infty} A_k) \\
\leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} \mu(A_k) = 0\,.
\end{aligned}$$

□

何时环上有限可加非负集函数为测度?

本节在最后, 我们来探讨一个问题: 若环 $\mathcal{R}$ 上已定义了一个非负集函数 μ, 其在空集处为零, 且满足有限可加性. 再加上什么条件, 可以使 μ 成为环 $\mathcal{R}$ 上的测度? 换言之, 再加上什么条件, 使 μ 满足可列可加性?

定理 14. *μ 是环 $\mathcal{R}$ 上非负集函数, 在空集处为零, 且满足有限可加性. μ 是环 $\mathcal{R}$ 上测度与下面命题等价:*

1. *μ 下连续.*
2. *μ 次可列可加.*
3. *μ 在 $\varnothing$ 处连续, 且任意环 $\mathcal{R}$ 上递增集列 $\{A_n\}$, 若 $\lim A_n \in \mathcal{R}$, 且 $\lim \mu(A_n) < \infty$, 则有 $\mu(\lim A_n) < \infty$.*

注. 若 μ 在 $\mathcal{R}$ 上有限, 则上述 3 可简洁的表示为 μ 在 $\varnothing$ 处连续. 在第五章 (Tulea 定理,Kolmogorov 相容定理) 中, 我们要在无穷维空间上建立测度中我们将用到上面的定理先定义好环上的测度, 在将其延拓.

证明. 只证若 3 成立, 则 μ 是 $\mathcal{R}$ 上测度, 其他是显然的. 设 $\{A_n\}$ 是 $\mathcal{R}$ 中互不相交的集列, 且 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}$, 我们欲证

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)\,.$$

由非负集函数 μ 有限可加, 在空集处为零, 知 μ 有单调性, 从而有

$$\mu(\bigcup_{n=1}^{\infty} A_n) \geqslant \sum_{n=1}^{\infty} \mu(A_n)\,.$$

于是不妨设 $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. 由 3 知此时 $\mu(\bigcup_{n=1}^{\infty} A_n) < \infty$. 于是考虑集列 $\{B_n\}$, 其中

$$B_n = \bigcup_{k=n}^{\infty} A_k\,,\ \forall n \in \mathbb{N}\,.$$

则 $\{B_n\}$ 单调递减收敛于 $\varnothing$, 且 $\mu(B_1) < \infty$. 由 μ 在 $\varnothing$ 处上连续, 故有

$$\lim_{n\to\infty} \mu(B_n) = 0\,.$$

注意到对任意 $N \in \mathbb{N}^{+}$, 有

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{N} \mu(A_n) + \mu(B_N)\,.$$

令 $N \to \infty$, 即得

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n) .$$

□

有时我们只在半环 $\mathcal{Q}$ 上定义了测度, 它是否具有上面提到的环上测度的性质呢? 一个自然的想法是将半环上的测度延拓到它生成的环上, 那半环上的测度自然也具有单调性, 次可列可加性和上下连续性.

练习 1. μ 是半环 $\mathcal{Q}$ 上的测度, 则 μ 可以唯一的延拓到环 $r(\mathcal{Q})$ 上.

1.3 测度延拓

本节探讨如何在 σ 域上建立测度. 但由于 σ 域的复杂结构, 我们很难直接在上面定义测度. 还好, 由之前的知识, 我们知道可以用较简单的集合系, 例如半环, 来生成 σ 域. 自然我们希望能先在较简单的集合系上定义测度, 之后将其延拓到简单集合系生成的 σ 域上.

我们先来探讨一个基本的:

1.3.1 外测度与测度延拓

定义 11. X 的子集全体 2^X 上的非负集函数 τ 称为 **外测度**, 若

1. $\tau(\varnothing) = 0$.
2. τ 有**单调性**: 对任意 $B \subset A \subset X$, 有 $\tau(B) \leqslant \tau(A)$.
3. τ 有**次可列可加性**: 对任意 X 中集列 $\{A_n\}$, 有

$$\tau(\bigcup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} \tau\left(A_n\right) .$$

我们希望能将外测度限制在比 2^X 小的某个 σ 域上, 使之成为该 σ 域上的测度. 这也是我们后面能进行测度延拓的关键.

定义 12. τ 为 X 上外测度. 称 X 的子集 A 是 τ 可测集, 若 A **分割测量外测度** τ, 其定义为: 对任意 $D \subset X$,

$$\tau(D) = \tau(D \cap A) + \tau\left(D \cap A^c\right) .$$

我们将全体 τ 可测集记为 $\mathcal{F}_\tau$.

定理 15 (Caratheodory 定理)**.** *τ 为 X 上外测度. 则 τ 可测集全体 $\mathcal{F}_\tau$ 是 σ 域, 且 τ 为 $\mathcal{F}_\tau$ 上测度.*

证明. 第一步. 我们先证明 $\mathcal{F}_\tau$ 是 σ 域. 不难证明 $\mathcal{F}_\tau$ 是一个域, 我们下面证明 τ 对可数的不交并封闭. 设 $\{A_n\}$ 是 $\mathcal{F}_\tau$ 中互不相交的集列, 任取 $D \subset X$, 由归纳法知对任意 k,

$$\tau(D) = \sum_{n=1}^{k} \tau(D \cap A_n) + \tau(D \setminus \cup_{n=1}^{k} A_n)\,.$$

由外测度的单调性与次可列可加性,

$$\begin{aligned}\tau(D) &\geq \sum_{n=1}^{\infty} \tau(D \cap A_n) + \tau(D \setminus \cup_n A_n) \\ &\geq \tau(D \cap \cup_n A_n) + \tau(D \setminus \cup_n A_n) \geq \tau(D)\,.\end{aligned} \tag{1.2}$$

于是可见 $\cup_n A_n \in \mathcal{F}_\tau$. 从而 $\mathcal{F}_\tau$ 是 σ 域.

第二步. 我们证明 τ 限制在 $\mathcal{F}_\tau$ 上是测度. 只需证明可列可加性, $\{A_n\}$ 是 $\mathcal{F}_\tau$ 中互不相交的集列, 则在 (1.2) 中取 $D = \cup_n A_n$ 得到

$$\tau\big(\bigcup_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} \tau(A_n)\,.$$

于是 τ 是 $\mathcal{F}_\tau$ 上的测度. □

我们可以利用半环 $\mathcal{Q}$ 上的测度生成外测度.

引理 16. *$\mathcal{Q}$ 是 X 上的半环. μ 是 $\mathcal{Q}$ 上测度. 对任意 $A \subset X$, 令*

$$\mu^*(A) = \inf\left\{\sum_{n=1}^{\infty} \mu(B_n) : A \subset \bigcup_{n=1}^{\infty} B_n, B_n \in \mathcal{Q}, n = 1, 2, \cdots\right\}$$

则 μ^* 是 X 上的外测度, 称为 μ **生成的外测度**.

证明. 显然 μ^* 是 2^X 上非负, 单调的集函数. 我们证明次可列可加性. 对任意集列 $\{A_n\}$, 且不妨设

$$\sum_n \mu^*(A_n) < \infty\,.$$

任意取定 $\epsilon > 0$, 对任意给定 n, 存在 $\mathcal{Q}$ 中集列 $\{B_k^n\}$, 与 ϵ 有关, 使得 $A_n \subset \cup_k B_k^n$ 且

$$\sum_k \mu(B_k^n) \leq \mu^*(A_n) + \frac{\epsilon}{2^n}\,.$$

注意到 $\cup_n A_n \subset \cup_{n,k} B_k^n$, 得到

$$\mu^*(\cup_n A_n) \leq \sum_{n,k} \mu(B_k^n) \leq \sum_n \mu^*(A_n) + \epsilon\,.$$

由于 ϵ 的任意性, 知次可列可加性成立. $\square$

利用引理 16 和 Caratheodory 定理, 可得如下的测度扩张定理:

定理 17. *$\mathcal{Q}$ 是 X 上的半环. μ 是 $\mathcal{Q}$ 上测度, 则存在 μ 在 $\sigma(\mathcal{Q})$ 上延拓. 换言之, 存在 $\sigma(\mathcal{Q})$ 上的测度 τ, 满足*

$$\tau(A) = \mu(A),\ \forall\, A \in \mathcal{Q}\,.$$

证明. 首先, 容易证明 μ^* 在 $\mathcal{Q}$ 上的限制等于 μ. 其此, 再证明 $\mathcal{Q} \subset \mathcal{F}_{\mu^*}$. 进而得到 $\sigma(\mathcal{Q}) \subset \mathcal{F}_{\mu^*}$. 令 τ 为 μ^* 在 $\sigma(\mathcal{Q})$ 上的限制即可. $\square$

考虑上面测度延拓的具体过程, 我们会问如下的问题:

- μ 在 $\sigma(\mathcal{Q})$ 上的延拓是唯一的吗?
- 既然可以将 μ 延拓到一个可能更大的 σ 域 $\mathcal{F}_{\mu^*}$ 上, $\mathcal{F}_{\mu^*}$ 与 $\sigma(\mathcal{Q})$ 中的集合多了些什么样的集合?

1.3.2 延拓的唯一性

先来回答第一个问题, μ 在 $\sigma(\mathcal{Q})$ 上的延拓是否唯一. 一般情形下并不是, 看下面的例子:

例 5. 考虑有理数全体 $\mathbb{Q}$ 和 $\mathbb{Q}$ 上的半环

$$\mathcal{Q} = \{\mathbb{Q} \cap (a, b] : -\infty < a \leqslant b < \infty\}\,.$$

定义 $\mathcal{Q}$ 上测度 μ 为:

$$\mu(A) = \begin{cases} 0, & A = \varnothing \\ \infty, & A \neq \varnothing \end{cases} \quad \text{for all } A \in \mathcal{Q}\,.$$

则对任何 $a > 0$, $\tau(A) = a|A|$, $\forall A \in \sigma(\mathcal{Q})$. 都是 μ 在 $\sigma(\mathcal{Q})$ 上的延拓.

可以注意到上述例子中 μ 在 $\mathcal{Q}$ 的定义很无趣, 我们自然想如果对测度做一点限制, 排除掉这种情况, 能不能得到唯一的延拓呢?

定理 18. $\mathcal{P}$ 是 X 上的 π 系. 若 μ,ν 皆是 $\sigma(\mathcal{P})$ 上的测度, 且满足

1. 对任何 $A\in\mathcal{P}$, 有 $\mu(A)=\nu(A)$.
2. 存在 X 的分划 (或穷竭) $\{A_n\}_{n=1}^{\infty}$, [2] 对任意正整数 n, $A_n\in\mathcal{P}$, 且 $\mu(A_n)<\infty$.

则对任何 $A\in\sigma(\mathcal{P})$, 有 $\mu(A)=\nu(A)$.

证明. 只需证任意 $n\in\mathbb{N}^+$ 和 $B\in\sigma(\mathcal{P})$ 有 $\mu(B\cap A_n)=\nu(B\cap A_n)$. 之后利用可列可加性 (或连续性) 可得, 对任意 $B\in\sigma(\mathcal{P})$

$$\mu(B)=\sum_{n=1}^{\infty}\mu(B\cap A_n)=\sum_{n=1}^{\infty}\nu(B\cap A_n)=\nu(B)\,.$$
$$(\text{ or } \mu(B)=\lim_{n\to\infty}\mu(B\cap A_n)=\lim_{n\to\infty}\nu(B\cap A_n)=\nu(B)\,.)$$

用 "好集原理" 来论证. 任取正整数 n , 定义

$$\mathcal{G}=\{B\in\sigma(\mathcal{P}):\mu(B\cap A_n)=\nu(B\cap A_n)\}\,.$$

我们知道 $\mathcal{P}\subset\mathcal{G}$. 再证明 $\mathcal{G}$ 是 λ 系即得 $\mathcal{G}=\sigma(\mathcal{P})$. □

注. 若想不用条件 2 直接证明, 另

$$\mathcal{G}=\{B\in\sigma(\mathcal{P}):\mu(B)=\nu(B)\}\,.$$

我们发现无法证明 $\mathcal{G}$ 是 λ 系, 除非 $\mu(X)<\infty$.

当我们讨论概率测度的时候, 定理会变得简单的多, 这是因为概率测度是有限的, 且全空间的测度总是为 1.

定理 19. $\mathcal{P}$ 是 Ω 上的 π 系. 若 $\mathbb{P}_1,\mathbb{P}_1$ 皆是 $\sigma(\mathcal{P})$ 上的概率测度, 且任何 $A\in\mathcal{P}$, 有 $\mathbb{P}_1(A)=\mathbb{P}_2(A)$. 则任何 $\mathbb{P}_1=\mathbb{P}_2$.

1.3.3 $\sigma(\mathcal{Q})$ 与 $\mathcal{F}_{\mu^*}$

我们来考虑 $\sigma(\mathcal{Q})$ 与 $\mathcal{F}_{\mu^*}$ 这两个 σ 代数的区别. 仍记半环 $\mathcal{Q}$ 上的测度 μ 在 $\mathcal{F}_{\mu^*}$ 上的延拓为 μ. 先来看下面的引理:

引理 20. 任意 $A\subset X$, 有 $B\in\sigma(\mathcal{Q})$ 使得 $A\subset B$ 且 $\mu(A)=\mu(B)$.

---

[2] X 分划的指 $\{A_n\}$ 两两不交且 $\cup_{n=1}^{\infty}A_n=X$; X 穷竭的指 $A_n\uparrow X$.

证明. 布不失一般性, 设 $\mu(A) < \infty$. 对任意正整数 n, 存在 $\mathcal{Q}$ 中集列 $\{B_k^{(n)}\}_{k=1}^{\infty}$, 使得

$$\sum_{k=1}^{\infty}\mu(B_k^{(n)}) < \mu^*(A) + \frac{1}{n}\,,$$

再令

$$B = \bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\infty} B_k^{(n)}\,,$$

则 B 满足要求. □

定理 21. 若 μ 是 $\mathcal{F}_{\mu^*}$ 上的 σ 有限测度, 则存在 $B_1, B_2 \in \sigma(\mathcal{Q})$, 满足 $B_2 \subset A \subset B_1$ 且 $\mu(B_1\backslash A) = \mu(A\backslash B_2) = 0$.

证明. 由 μ 是 $\mathcal{F}_{\mu^*}$ 上的 σ 有限测度, 则有 X 的划分 $\{A_n\}$, 对任意正整数 n, $A_n \in \mathbb{F}_\mu$ 且 $\mu(A_n) < \infty$. 由上面引理, 存在 $E_n \in \mathcal{Q}$, $A \cap A_n \subset E_n$ 且 $\mu(A\cap A_n) = \mu(E_n)$, 令 $B_1 = \cup_{n=1}^{\infty} E_n$, 则 B_1 满足要求.

由上述证明, 存在 $B_2 \in \sigma(\mathcal{Q})$, 使得 $A^c \subset B_2^c$ 且 $\mu(B_2^c\backslash A^c) = 0$, 可见 B_2 满足要求. □

上面的定理在 μ 是 σ 有限测度的情形下回答了我们的问题: $\mathcal{F}_{\mu^*}$ 中的集合总可以表示为和 $\sigma(\mathcal{Q})$ 上的集合差去或并上某个 μ 零测集. 注意到这一种现象: 测度空间 $(X, \sigma(\mathcal{Q}), \mu)$ 上的零测集, 其子集不一定可测. 但测度空间 $(X, \mathcal{F}_{\mu^*}, \mu)$ 上的零测集, 其子集一定是可测的. 为了描述这种区别, 我们引入了完全的测度空间的概念.

定义 13. 测度空间 $(X, \mathcal{F}, \mu)$ 称为**完全的**, 若其中任何零测集的子集也可测. 换言之, 对任何 $N \in \mathcal{F}$, $\mu(N) = 0$, 成立 $A \in \mathcal{F}$, $\forall\, A \subset N$.

于是我们说 $(X, \sigma(\mathcal{Q}), \mu)$ 是完全的, 但 $(X, \sigma(\mathcal{Q}), \mu)$ **不一定**是完全的. 我们由 $(X, \sigma(\mathcal{Q}), \mu)$ 再得到 $(X, \mathcal{F}_{\mu^*}, \mu)$ 的过绍, 是一种**完全化**的过程. 但我们总可以将测度空间完全化.

练习 2. 对任何测度空间 $(X, \mathcal{F}, \mu)$, 定义集合系

$$\tilde{\mathcal{F}} = \{A \cup N : A \in \mathcal{F}, N \text{ is a null set }\}\,.$$

及其上非负集函数

$$\tilde{\mu}(A \cap N) = \mu(A)\,,$$

则 $(X, \tilde{\mathcal{F}}, \tilde{\mu})$ 是一完全的测度空间, 且对每个 $A \in \mathcal{F}$, $\tilde{\mu}(A) = \mu(A)$.

下面的练习在高等概率论中证明 Hewitt-Sage 零一律时用到.

练习 3. $(X, \sigma(\mathcal{A}), \mu)$ 是 σ 有限的测度空间, 其中 $\mathcal{A}$ 是一代数. 则任取 $B \in \sigma(\mathcal{A})$, 给定 $\epsilon > 0$, 存在 $A \in \mathcal{A}$, 使得

$$\mu(A \Delta B) < \epsilon.$$

第二章　可测函数

本章讨论可测空间上定义的一类特殊的函数, 及可测函数. 为了极限运算的方便, 我们允许函数值为广义实数. 在可测空间引入测度后, 我们还会考虑可测函数的几种收敛性. 最后一节讨论概率空间中的实值可测函数, 也就是所谓的随机变量.

2.1　定义与基本性质

本章我们来研究一个可测空间到另一个可测空间的映射, 这个映射当然要和可测集产生关系, 不然这与研究两个普通的集合 (上面没有可测结构) 之间的映射没什么区别. 我们最关心, 也是最基本最常用的情况当然是到实数空间 (或者广义实数空间) 的映射, 即是所谓的可测函数.

2.1.1　可测映射

定义 14. 可测空间 $(X.\mathcal{F})$ 到另一可测空间 $(Y,\mathcal{S})$ 的映射 $f: X \to Y$ 被称为 $(X,\mathcal{F})$ 到 $(Y,\mathcal{S})$ 的**可测映射**或**随机元**, 若

$$f^{-1}\mathcal{S} \subset \mathcal{F},$$

容易验证 $f^{-1}\mathcal{S}$ 是 X 上的 σ 域. 我们称之 f **生成**的 σ 域, 记为 $\sigma(f)$, 它也是使 f 可测的最小的 σ 域.

要按照定义验证映射 f 是否可测有些麻烦. 还好在绝大部分情况下, 我们有一个简单的判别方法:

引理 22. *f 是 X 到可测空间 $(Y,\mathcal{S})$ 的映射. 若 $\mathcal{S}$ 是由集合系 $\mathcal{E}$ 生成的, 即 $\mathcal{S}=\sigma(\mathcal{E})$, 则*

$$f^{-1}\sigma(\mathcal{E}) = \sigma\left(f^{-1}\mathcal{E}\right) .$$

证明. 我们已经指出了 $f^{-1}\sigma(\mathcal{E})$ 是 σ 域. 于是有

$$\sigma(f^{-1}\mathcal{E}) \subset f^{-1}\sigma(\mathcal{E}).$$

只需证 $f^{-1}\sigma(\mathcal{E}) \subset \sigma(f^{-1}\mathcal{E})$. 令

$$\mathcal{G} = \{A \in \sigma(\mathcal{E}) : f^{-1}A \in \sigma(f^{-1}\mathcal{E})\}.$$

容易验证 $\mathcal{G}$ 是 σ 域. 注意到 $\mathcal{E} \subset \mathcal{G} \subset \sigma(\mathcal{E})$, 故有 $\mathcal{G} = \sigma(\mathcal{E})$. 于是得到 $f^{-1}\sigma(\mathcal{E}) \subset \sigma(f^{-1}\mathcal{E})$. □

利用上述引理, 容易得到下面的命题.

命题 23. *f 是 $(X, \mathcal{F})$ 到 $(Y, \sigma(\mathcal{E}))$ 的可测映射当且仅当 $f^{-1}\mathcal{E} \subset \mathcal{F}$.*

回忆数学分析中我们学过的, 连续函数的复合仍是连续函数, 自然地要问, 可测映射的复合是否还是可测映射? 答案是肯定的, 证明不难, 故略去了.

命题 24. *设 f 是可测空间 $(X, \mathcal{F})$ 到 $(Y, \mathcal{S})$ 的可测映射, g 是 $(Y, \mathcal{S})$ 到可测空间 $(Z, \mathcal{Z})$ 的可测映射, 则 $g \circ f$ 是 $(X, \mathcal{F})$ 到 $(Z, \mathcal{Z})$ 的可测映射.*

2.1.2　可测函数

下面我们来考虑最基本也最重要的情形, 也就是可测函数的情形. 实变函数中已经介绍过广义实数集 $\overline{\mathbb{R}}$ 和其中的运算. $\overline{\mathbb{R}}$ 上的 Borel 集定义为

$$\mathcal{B}(\overline{\mathbb{R}}) \coloneqq \sigma(\mathcal{B}(\mathbb{R}), \{-\infty\}, \{\infty\})$$

很容易证明有 $\mathcal{B}(\overline{\mathbb{R}}) = \sigma([-\infty, a] : a \in \mathbb{R}) = \sigma([-\infty, a) : a \in \mathbb{R})$.

定义 15. 可测空间 $(X, \mathcal{F})$ 到 $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ 的可测映射称为 $(X, \mathcal{F})$ 上的 **可测函数**. 特别地, $(X, \mathcal{F})$ 到 $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ 的可测映射称为上的 **有限值可测函数**或 **随机变量**.

现在我们立刻得到: 一个取广义实数值的 $(X, \mathcal{F})$ 上的函数 f 可测, 当且仅当对任意实数 a, $\{f \leqslant a\}$ 是可测集, 即 $\{f \leqslant a\} \in \mathcal{F}$.

可测函数的运算　可测函数与可测映射最大的不同在于, 可测函数可以进行运算, 这是因为广义实数集上定义好了加减乘除和极限运算. 一个自然的问题是, 定义在可测空间 $(X, \mathcal{F})$ 上的可测函数, 在经过 $\overline{\mathbb{R}}$ 中的运算以后其可测性是否仍然可以保持? 我们下面来回答这个问题.

命题 25. *$f_n, n \in \mathbb{N}^+$ 是 $(X,\mathcal{F})$ 上的一列可测函数，则 $\inf_n f_n$, $\sup_n f_n$, $\underline{\lim} f_n$, $\overline{\lim} f_n$ 皆是可测函数.*

证明. 只需证明 $\inf_n f_n$ 可测, 因为 $\sup_n f_n = -\inf_n(-f_n)$,

$$\underline{\lim}_{n\to\infty} f_n = \sup_{k\geqslant 1}\inf_{n\geqslant k} f_n \ , \ \overline{\lim}_{n\to\infty} f_n = \inf_{k\geqslant 1}\sup_{n\geqslant k} f_n \, .$$

任取实数 a, 有

$$\left\{\inf_n f_n > a\right\} = \bigcap_n \{f_n > a\} \in \mathcal{F} \, .$$

故 $\inf_n f_n$ 是可测函数. □

命题 26. *若 f, g 是 $(X,\mathcal{F})$ 上的可测函数，则*

1. *对任意实数 a, af 是可测函数.*
2. *若 $f+g$ 有意义, 即任何 $x \in X, f(x)+g(x)$ 有意义, 则其是可测函数.*
3. *fg 是可测函数.*
4. *若 g 不取 0 值, 则 f/g 是可测函数.*

证明. $1:$ 显然.

$2:$ 不妨先设 f, g 皆为实值可测函数, 则对任意实数 a,

$$\{f+g<a\} = \bigcup_{r\in\mathbb{Q}} \{f<r\} \cap \{g<a-r\} \in \mathcal{F} \, .$$

对于 f, g 取广义实数值的情形, 令 $\tilde{f} = f1_{|f|<\infty}$, $\tilde{g} = g1_{|g|<\infty}$, 则由 $f+g$ 有意义,

$$f+g = (\tilde{f}+\tilde{g}) + \infty 1_{\{f=\infty, g>-\infty\}} + (-\infty)1_{\{f=-\infty, g<\infty\}}$$

易见 $f+g$ 为可测函数.

$3:$ 任取实数 a, 注意

$$\begin{aligned}\{fg<a\} = (\{fg<a\}\cap\{g=0\}) &\cup (\{fg<a\}\cap\{g>0\}) \\ &\cup (\{fg<a\}\cap\{g<0\}) \, .\end{aligned}$$

显然 $\{fg<a\}\cap\{g=0\} \in \mathcal{F}$, 注意

$$\{fg<a\}\cap\{g>0\} = \bigcup_{r\in\mathbb{Q}} \{f<r\} \cap \left\{g<\frac{a}{r}\right\} \in \mathcal{F} \, ,$$

同理 $\{fg<a\}\cap\{g<0\} \in \mathcal{F}$, 故见 fg 可测.

$4:$ 由 $g\neq 0$ 易证 $\frac{1}{g}$ 可测, 再由 3 即得. □

可测函数的结构 我们首先来考虑最简单的可测函数. 称 $(X,\mathcal{F})$ 上的可测函数 f 为**简单函数**, 若 f 的值域由有限个实数组成. 显然, 对一个简单函数 f, 存在 X 的有限可测分割 $\{A_j : j=1,2,\cdots,n\}$ 和实数 $\{a_1,a_2,\cdots,a_n\}$, 使得

$$f=\sum_{j=1}^{n} a_i I_{A_j}\,.$$

换言之, 简单函数就是可测集的示性函数的线性组合. 容易看出上述表示不是唯一的.

命题 27. *f 是 $(X,\mathcal{F})$ 上非负可测函数, 则存在非负简单可测函数列 $\{\phi_n\}_{n\in\mathbb{N}^+}$, 使得 $\phi_n\uparrow f$.*

证明. 对任意正整数 n, 定义

$$\phi_n \coloneqq \sum_{k=0}^{n2^n-1}\frac{k}{2^n} I_{\left\{\frac{k}{2^n}\leqslant f<\frac{k+1}{2^n}\right\}}+nI_{\{f\geqslant n\}}$$

易见 ϕ_n 满足条件 . □

可以发现, 若 f 是非负有界可测函数, 则 ϕ_n 一致收敛于 f. 同时, 注意任何可测函数可以分解为其正部和负部之差:

$$f=f^+-f^-\,.$$

故利用上述命题, 容易得到: *任何 $(X,\mathcal{F})$ 上可测函数 f, 存在简单可测函数列 $\{\phi_n\}_{n\in\mathbb{N}^+}$, 使得 $\phi_n\to f$.*

$(X,\sigma(g))$ 上可测函数的表示

定理 28. *设 g 是 X 到 $(Y,\mathcal{S})$ 的映射. 则对任一 $(X,\sigma(g))$ 上可测函数 f, 存在 $(Y,\mathcal{S})$ 上的可测函数 h, 使得 $f=h\circ g$.*

证明. 首先, 假设 f 的 $(X,\sigma(g))$ 上某可测集的示性函数, 注意 $\sigma(g)=g^{-1}\mathcal{S}$, 则存在 $A\in\mathcal{S}$,

$$f=I_{g^{-1}(A)}=I_A\circ g\,.$$

令 $h=I_A$, 此时定理成立.

进而假设 f 是 $(X,\sigma(g))$ 上简单函数, 即存在 $A_1,\cdots,A_n\in\mathcal{S}$ 和实数 $a_1,\cdots,a_n$ 使得

$$f=\sum_{j=1}^{n} a_i I_{g^{-1}(A_j)}=\sum_{j=1}^{n} a_i(I_{A_j}\circ g)=\left(\sum_{j=1}^{n} a_i I_{A_j}\right)\circ g\,.$$

令 $h=\sum_{j=1}^n a_i I_{A_j}$, 此时定理成立.

最后考虑 f 是 $(X,\sigma(g))$ 上一般的可测函数, 这时取简单函数列 $\{\phi_n\}$, 使得 $\phi_n\to f$, 由上一步的结论, 存在存在 $(Y,\mathcal{S})$ 上的可测函数 h_n, 使得

$$\phi_n = h_n\circ g\,,\ \text{for all } n\,.$$

于是

$$f=\lim_{n\to\infty}\phi_n=\lim_{n\to\infty}(h_n\circ g)$$

定义 Y 上的函数

$$h=\begin{cases}\lim\limits_{n\to\infty} h_n\,, & \text{if } \lim\limits_{n\to\infty} h_n \text{ exists}\,,\\ 0\,, & \text{otherwise}\,.\end{cases}$$

易见 h 是 $(Y,\mathcal{S})$ 上可测函数, 且 $f=h\circ g$, 故定理成立. □

注. 注意, 从定理中证明中不难发现,

1. h 不一定是唯一的. 但是若 g 是 X 到 Y 的满射, h 必然是唯一的.
2. 若 f 为实值可测函数, 则 h 亦可取为实值可测函数; 若 f 为有界可测函数, 则 h 亦可取为有界可测函数.

请注意定理证明过程中具有典型意义的方法! 在测度论和概率论中, 为了证明一个关于可测函数的命题, 常常分解为如下几个比较容易的步骤来进行:

1. 证明该命题对最简单的函数, 即指示函数成立.
2. 证明该命题对非负简单函数, 即指示函数的线性组合成立.
3. 证明该命题对非负可测函, 即递增的非负简单函数列的极限成立.
4. 证明命题对一般可测函数, 即两个非负可测函数之差成立.

按上述步骤证明命题的方法叫做测度论中的 **典型方法**. 典型方法符合人们的认识过程, 是一种具有普遍意义的, 行之有效的方法.

把典型方法分别和单调类定理和 Dynkin's $\pi-\lambda$ 定理结合起来, 可以得到函数形式的 “单调类定理” 和 “ Dynkin's $\pi-\lambda$ 定理”.

练习 4 (函数形式单调类定理). $\mathcal{M}$ 是一个由 X 上的非负广义实值函数组成的单调类, 即它是 X 上具有下列性质的由非负广义实值函数组成的集合:

1. 对任意 $f, g \in \mathcal{M}$ 和实数 $a, b \geqslant 0$, $af + bg \in \mathcal{M}$.

2. 对任何 $\mathcal{M}$ 中递增的函数列 $\{f_n\}$, $\lim_n f_n \in \mathcal{M}$.

设 $\mathcal{A}$ 是一个域, 若任意 $A \in \mathcal{A}$, $I_A \in \mathcal{M}$, 则 $(X, \sigma(\mathcal{A}))$ 上的非负可测函数全体皆属于 $\mathcal{M}$.

练习 5 (函数形式 π-λ 定理). $\mathcal{L}$ 是 X 上的非负广义实值函数组成的 λ 系, 即是说

1. $I_X \in \mathcal{L}$.

2. 任何 $f, g \in \mathcal{L}$ 和 $a, b \in \mathbb{R}$, 若 $af + bg \geqslant 0$, 则 $af + bg \in \mathcal{L}$.

3. $\{f_n\} \subset \mathcal{L}$, 若 $f_n \uparrow f$, 则 $f \in \mathcal{L}$.

设 $\mathcal{P}$ 是 X 上的 π 系, 若任何 $A \in \mathcal{P}$, $I_A \in \mathcal{L}$, 则 $(X, \sigma(\mathcal{P}))$ 上的非负可测函数全体皆属于 $\mathcal{L}$.

2.2 可测函数的收敛

本节我们来介绍可测函数 (序列) 几种收敛的定义, 及其相互的关系. 若无特别声明, 我们总是假定可测函数定义在某个测度空间 $(X, \mathcal{F}, \mu)$ 上.

2.2.1 几乎处处收敛

为了介绍几乎处处收敛, 我们先严格的说明什么 “几乎处处”.

$(X, \mathcal{F}, \mu)$ 是一测度空间, $\mathrm{P}(x)$ 是与 X 中的点有关的某个命题. 若存在一零测集 N, 当 $x \in N^c$ 时, 命题 P 皆成立, 我们称命题 P (关于测度 μ) **几乎处处 (almost everywhere)** 成立. (关于测度 μ) 几乎处处成立常常简写为 (μ-)a.e. , 在不至于混淆时科我们不写出关于哪一个测度几乎处处.

注意 $\{x : P(x)\text{不成立}\}$ 不一定是零测集, 因为其不一定可测. 但若 $(X, \mathcal{R}, \mu)$ 为完全的测度空间, $\{x \in X : \mathrm{P}(x)\text{不成立}\}$ 是零测集.

下面我们介绍可测函数序列几乎处处收敛的概念:

定义 16. $f, f_n, n = 1, 2, \cdots$ 皆是可测函数. 若

$$\mu\left\{\lim_{n\to\infty} f_n \neq f \text{ 或 } \lim_{n\to\infty} f_n\text{不存在}\right\} = 0\,.$$

则称 f_n **几乎处处以 f 为极限**. 若 f 几乎处处有限, 则称 $\{f_n\}$ **几乎处处收敛于 f.** 记为 $f_n \xrightarrow{\text{a.e.}} f$.

注意. 请验证 $\left\{\lim_{n\to\infty} f_n \neq f \text{ 或 } \lim_{n\to\infty} f_n\text{不存在}\right\}$ 是可测集.

我们主要用到的是几乎处处有限的可测函数, 而这样的函数可以当作有限可测函数来处理. 对于几乎处处有限的可测函数列, 其几乎处处收敛有如下的充要判定条件.

定理 29. f 是几乎处处有限的可测函数. 可测函数列 $f_n \xrightarrow{\text{a.e.}} f$ 当且仅当对任意 $\epsilon > 0$, 成立

$$\mu\left(\limsup_{n\to\infty}\{|f_n - f| \geqslant \epsilon\}\right) = 0\,. \tag{2.1}$$

注意. 因为我们假定了 f 几乎处处有限, 故无论认为 $\{|f_n - f| \geqslant \epsilon\}$ 是否包含使得 $f_n(x) - f(x)$ 无意义的点, 对式 (2.1) 没有影响 . 故今后总认为 $\{|f_n - f| \geqslant \epsilon\}$ 不包含使得 $f_n(x) - f(x)$ 无意义的点.

Proof. 不失一般性, 可设 f 为实值可测函数. 注意到

$$\lim_{n\to\infty} f_n(x) = f(x) \Leftrightarrow \limsup_{n\to\infty} |f_n(x) - f(x)| = 0\,.$$

于是有

$$\left\{\lim_{n\to\infty} f_n \neq f \text{ 或 } \lim_{n\to\infty} f_n\text{不存在}\right\} = \{\limsup_{n\to\infty} |f_n - f| > 0\}\,.$$

注意

$$\begin{aligned}\{\limsup_{n\to\infty} |f_n - f| > 0\} &= \bigcup_{k=1}^{\infty}\left\{\limsup_{n\to\infty} |f_n - f| > \frac{1}{k}\right\} \\ &= \bigcup_{k=1}^{\infty}\limsup_{n\to\infty}\left\{|f_n - f| > \frac{1}{k}\right\}\ .\end{aligned}$$

因此有

$$\begin{aligned}&\mu\left(\left\{\lim_{n\to\infty} f_n \neq f \text{ 或 } \lim_{n\to\infty} f_n\text{不存在}\right\}\right) = 0 \\ \Longleftrightarrow\ &\mu\left(\{\limsup_{n\to\infty} |f_n - f| > 0\}\right) = 0 \\ \Longleftrightarrow\ &\mu\left(\bigcup_{k=1}^{\infty}\limsup_{n\to\infty}\left\{|f_n - f| > \frac{1}{k}\right\}\right) = 0 \\ \Longleftrightarrow\ &\mu\left(\limsup_{n\to\infty}\left\{|f_n - f| > \frac{1}{k}\right\}\right) = 0,\ \forall\, k \in \mathbb{N}^+ \\ \Longleftrightarrow\ &\mu\left(\limsup_{n\to\infty}\{|f_n - f| \geqslant \epsilon\}\right) = 0,\ \forall\, \epsilon > 0\,.\end{aligned}$$

□

2.2.2 几乎一致收敛

我们引入另一种可测函数的收敛, 回忆实变函数课程中著名的 Egoroff 定理, 这种收敛概念便是从中抽象出来的.

定义 17. $f, f_n(n = 1, 2, \cdots)$ 是可测函数, 若对任意 $\delta > 0$, 存在可测集 $A = A_\delta$, 满足 $\mu(A) < \delta$ 且 f_n 在 A^c 上一致收敛于 f, 即

$$\sup_{x \in A^c} |f_n(x) - f(x)| \to 0\,, n \to \infty\,. \tag{2.2}$$

则称可测函数列 $\{f_n\}$ **几乎一致 (almostly uinform) 收敛于** f, 记为 $f_n \xrightarrow{\text{a.u.}} f$.

注意. f_n 在 A^c 上一致收敛于 f, 蕴涵了 f 在 A^c 上取实值. 因此若 f_n 几乎处处收敛于 f, 则 f 必是几乎处处有限的可测函数.

关于可测函数列的几乎一致收敛, 我们有如下的判定条件:

定理 30. *f 是几乎处处有限的可测函数. 可测函数列 $f_n \xrightarrow{\text{a.u.}} f$ 的当且仅当对任意 $\epsilon > 0$, 成立*

$$\lim_{n \to \infty} \mu \left\{ \sup_{k \geqslant n} |f_k - f| \geqslant \epsilon \right\} = 0\,. \tag{2.3}$$

Proof. *必要性.* 由 $f_n \xrightarrow{\text{a.u.}} f$, 任取 $\delta > 0$, 存在可测集 A, 使得 $\mu(A) < \delta$, 使得在 A^c 上有 $\sup_{k \geqslant n} |f_k - f| \to 0$, 当 $n \to \infty$. 由此可知当 n 充分大时,

$$A^c \cap \left\{ \sup_{k \geqslant n} |f_k - f| \geqslant \epsilon \right\} = \varnothing\,.$$

从而有

$$\lim_{n \to \infty} \mu \left\{ \sup_{k \geqslant n} |f_k - f| \geqslant \epsilon \right\} \leqslant \delta\,.$$

由 δ 的任意性知, (2.3) 式成立.

充分性. 若 (2.3) 式成立, 则任取 $\delta > 0$ 和正整数 j, 必存在正整数 $N_j = N_j(\delta)$, 使得

$$\mu \left\{ \sup_{k \geqslant N_j} |f_k - f| \geqslant \frac{1}{j} \right\} \leqslant \frac{\delta}{2^j}\,.$$

令

$$A = \bigcup_{j=1}^{\infty} \left\{ \sup_{k \geqslant N_j} |f_k - f| \geqslant \frac{1}{j} \right\}\,.$$

由次可加性知 $\mu(A) < \delta$. 注意

$$A^c = \bigcap_{j=1}^{\infty} \left\{ \sup_{k \geqslant N_j} |f_k - f| < \frac{1}{j} \right\}.$$

可见在 A^c 上 f_n 一致收敛于 f. 从而有 $f_n \xrightarrow{\text{a.u.}} f$. □

注. 结合后面对可测函数序列依测度收敛的概念, 本定理即是

$$f_n \xrightarrow{\text{a.u.}} f \Leftrightarrow \sup_{k \geqslant n} |f_k - f| \xrightarrow{\mu} 0 \quad \text{as } n \to \infty .$$

2.2.3 依测度收敛

定义 18. $f_n (n = 1, 2, \cdots)$ 是一列可测函数, f 是几乎处处有限的可测函数. 若对任意 $\epsilon > 0$, 有

$$\lim_{n \to \infty} \mu \{ |f_n - f| \geqslant \epsilon \} = 0 .$$

则称可测函数列 $\{f_n\}$ **依测度收敛于** f , 记为 $f_n \xrightarrow{\mu} f$.

换言之, $\{f_n\}$ 依测度收敛于 f 当且仅当, 对任意 $\epsilon, \delta > 0$, 存在 $N = N(\epsilon, \delta) \in \mathbb{N}^+$, 使得对任意的 $n \geqslant N$

$$\mu \{ |f_n - f| \geqslant \epsilon \} < \delta .$$

对于依测度收敛, 我们有如下的判定条件:

定理 31. f 是几乎处处有限的可测函数. 可测函数列 $f_n \xrightarrow{\mu} f$ 当且仅当 $\{f_n\}$ 的任意子列 $\{f_{k_n}\}$, 存在其子列 $\{f_{k_{l_n}}\}$ 使得 $f_{k_{l_n}} \xrightarrow{\text{a.u.}} f$.

Proof. 必要性. 只需证明若 $f_n \xrightarrow{\mu} f$, 必存在 $\{f_n\}$ 的子列 $\{f_{k_n}\}$ 几乎一致收敛于 f. 由 $f_n \xrightarrow{\mu} f$ 知, 有 $\{f_n\}$ 的子列 $\{f_{k_n}\}$, 满足

$$\mu \left\{ |f_{k_n} - f| \geqslant \frac{1}{n} \right\} \leqslant \frac{1}{2^n} .$$

任取 $\epsilon > 0$, 当 $\frac{1}{m} < \epsilon$ 时,

$$\begin{aligned} \mu \left\{ \sup_{n \geqslant m} |f_{k_n} - f| \geqslant \epsilon \right\} &\leqslant \mu \left\{ \sup_{n \geqslant m} |f_{k_n} - f| \geqslant \frac{1}{n} \right\} \\ &\leqslant \frac{1}{2^n} = \frac{1}{2^{m-1}} . \end{aligned}$$

故知 $\{f_{k_n}\}$ 近一致收敛于 f.

充分性. 用反证法, 假设 $f_n \xrightarrow{\mu} f$ 不成立. 则存在 $\epsilon_0, \delta_0 > 0$ 和 $\{f_n\}$ 的子列 $\{f_{k_n}\}$, 使得

$$\mu\{|f_{k_n} - f| \geqslant \epsilon_0\} \geqslant \delta_0 .$$

于是对 $\{f_{k_n}\}$ 的任何子列 $\{f_{k_{l_n}}\}$ 也成立

$$\mu\left\{\left|f_{k_{l_n}} - f\right| \geqslant \epsilon_0\right\} \geqslant \delta_0 .$$

故 $f_{k_{l_n}}$ 不依测度收敛于 f, 进而不几乎一致收敛于 f. 这是一个矛盾! □

练习 6. $f_n (n = 1, 2, \cdots)$ 是一列实值可测函数. 若对任意 $\epsilon > 0$, 有

$$\lim_{n,m\to\infty} \mu\{|f_n - f_m| \geqslant \epsilon\} = 0 .$$

则称 $\{f_n\}$ 为**依测度 Cauchy 列**. 证明, $\{f_n\}$ 为依测度 Cauchy 列当且仅当存在实值可测函数 f 使得 $f_n \xrightarrow{\mu} f$.

学完积分的定义后, 有下面的练习.

练习 7. 设 $(X, \mathcal{F}, \mu)$ 为有限测度空间, 记其上实值可测函数全体为 $\mathcal{L}$. 定义 $\mathcal{L}$ 上的二元函数 d, 为

$$d(f,g) = \int_X \frac{|f-g|}{1+|f-g|}\,\mathrm{d}\mu\,, \text{ for all } f, g \in \mathcal{L} .$$

则 d 是一个距离, 满足对 $f_n, f \in \mathcal{L}$

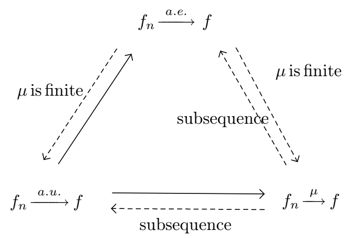
$$d(f_n, f) \to 0 \Leftrightarrow f_n \xrightarrow{\mu} f .$$

从而可见 $(\mathcal{L}, d)$ 是完备的距离空间. (实际上 $\mathcal{L}$ 是线性空间, 距离平移不变, 加法与数乘皆连续, 0 有局部凸邻域基, 故 $\mathcal{L}$ 为 Frechet 空间.)

三种收敛的蕴涵关系 由定理 29, 定理 30, 定理 31 立得下面的定理:

定理 32.

- $f_n \xrightarrow{\text{a.u.}} f$ 蕴涵了 $f_n \xrightarrow{\text{a.e.}} f$ 和 $f_n \xrightarrow{\mu} f$.
- μ 为有限测度时, $f_n \xrightarrow{\text{a.e.}} f$ 等价于 $f_n \xrightarrow{\text{a.u.}} f$.
- $f_n \xrightarrow{\mu} f$ 当且仅当 $\{f_n\}$ 的任意子列 $\{f_{k_n}\}$, 存在其子列 $\{f_{k_{l_n}}\}$ 使得 $f_{k_{l_n}} \xrightarrow{\text{a.u.}} f$.



此外, 我们举出几个反例, 说明上面定理所加的条件是必须的.

例 6. $\{f_n\}$ 几乎处处收敛于 f, 但 f_n 不依测度收敛于 f, 自然也不几乎一致收敛于 f. 考虑测度空间 $(\mathbb{R}, \mathcal{B}, m)$, 其中 $\mathcal{B}$ 是 Borel 代数, m 是 Lebesgue 测度. 任意 $n \in \mathbb{N}^+$, 定义函数 f_n 为:

$$f_n(x) = \begin{cases} 0, & |x| \leqslant n. \\ 1, & |x| > n. \end{cases}$$

容易发现 f_n 处处收敛于 0. 但显然 f_n 不是依测度收敛于 0 的.

例 7. 有限测度空间上 $\{f_n\}$ 依测度收敛于 f, 但 f_n 不几乎处处收敛于 f, 自然也不几乎一致收敛于 f. 考虑有限测度空间 $((0,1], \mathcal{B} \cap (0,1], m)$. 对任何正整数 k, 我们有区间的划分

$$(0,1] = (\frac{0}{2^k}, \frac{1}{2^k}] \cup (\frac{1}{2^k}, \frac{2}{2^k}] \cup \cdots \cup (\frac{2^k-1}{2^k}, \frac{2^k}{2^k}]$$

对任何 $n \in \{1, 2, \cdots, 2^k\}$, 定义

$$g_k^{(n)}(x) = I_{(\frac{n-1}{2^k}, \frac{n}{2^k}]}(x), \ x \in (0,1].$$

现在按照字典排序, 将 $g_k^{(n)}$ 重新排列为

$$g_1^{(1)}, g_1^{(2)}, \cdots, g_k^{(1)}, g_k^{(2)}, \cdots, g_k^{(2^k)}, \cdots$$

并且记为函数列 $\{f_n\}$. 显然有 $\{f_n\}$ 依测度收敛于 0, 但 $\{f_n\}$ 在每个点 $x \in (0,1]$ 都不收敛.

2.3 随机变量的依分布收敛

首先回忆一下, 按照之前的定义, 对于概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的 r.v. $\{X_n\}$ 和 X, 有

$$X_n \xrightarrow{\text{a.e.}} X \Leftrightarrow \mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1\,;$$
$$X_n \xrightarrow{\mathbb{P}} X \Leftrightarrow \lim_{n\to\infty} \mathbb{P}\left(|X_n - X| < \varepsilon\right) = 1, \forall\, \varepsilon > 0\,.$$

在概率论中 $X_n \xrightarrow{\text{a.e.}} X$ 我们更乐意称之为**几乎必然 (almost surely) 收敛**于 X, 记为 $X_n \xrightarrow{\text{a.s.}} X$. $X_n \xrightarrow{\mathbb{P}} X$ 也不泛泛地用依测度收敛这个词, 而称之为 $\{X_n\}$ **依概率收敛**于 X. 由定理 32 直接得到依概率收敛和几乎必然收敛的关系.

定理 33. *随机变量 $\{X_n\}$ 依概率收敛于 X 的充要条件是对 $\{X_n\}$ 的任意子列 $\{X_{k_n}\}$, 存在其子列 $\{X_{k_{l_n}}\}$ 使得 $X_{k_{l_n}} \xrightarrow{\text{a.s.}} X$.*

除了上面对一般测度空间定义过的那些收敛性以外，概率论中还有一种很重要的收敛性，叫做**依分布收敛**. 本节我们主要来讨论概率空间中随机变量的的依分布收敛, 以及它和其他收敛的关系.

2.3.1 分布函数

我们将满足

$$F(-\infty) \coloneqq \lim_{x\to-\infty} F(x) = 0\,;\ F(+\infty) \coloneqq \lim_{x\to\infty} F(x) = 1$$

的准分布函数 (即单调递增右连续) 称为**分布函数**. 记为 d.f. (distribution function). 不难验证: 对 r.v. X, 定义

$$F(x) = \mathbb{P}(X \leqslant x)\,,\ \forall\, x \in \mathbb{R}\,,$$

则 F 是一分布函数, 称为 r.v. X 的分布函数, 也说 X **服从** F, 记为 $X \sim F$.

若 X 是从概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 到另一可测空间 $(Y, \mathcal{S})$ 的随机元. 它在 $(Y, \mathcal{S})$ 自然诱导出的概率测度

$$(\mathbb{P}_X)(B) \coloneqq \mathbb{P}(X^{-1}B)\,,\ \forall\, B \in \mathcal{S}$$

称为随机元 X 的**概率分布**或**分布**. 容易看出, r.v. X 的 d.f. F 只不过是它的概率分布在 π 系 $\mathcal{P}_{\mathbb{R}}$ 上的值.

设 F 是一准分布函数. 对任何 $t \in (F(-\infty)\ F(\infty))$, 令

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \geqslant t\} = \sup\{x \in \mathbb{R} : F(x) < t\}\,,$$

我们称 $F^{\leftarrow}$ 为 F 的**左连续逆**. 显然分布函数的左连续逆在 $(0,1)$ 上皆有定义. 很容易证明左连续逆有下列性质.

命题 34. *对任何准分布函数 F , 有*

1. *$F^{\leftarrow}$ 是 $(F(-\infty), F(\infty))$ 上单调递增左连续的实值函数, 且*

$$F^{\leftarrow}(t^{+}) = \inf\{x : F(x) > t\} = \sup\{x : F(x) \leqslant t\}\,.$$

2. *对任意 $t \in (F(-\infty), F(\infty))$ 和 $x \in \mathbb{R}$, 有 $F(F^{\leftarrow}(t)) \geqslant t$, 更详细地,*

$$F(x) \geqslant t \Leftrightarrow x \geqslant F^{\leftarrow}(t)\,;\ x < F^{\leftarrow}(t) \Leftrightarrow F(x) < t\,.$$

有一个自然的问题, 任何一个 d.f. F, 是否一定有某个概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 以及它上面的 r.v. X 使得 $X \sim F$ 呢? 利用 d.f. F 的左连续逆我们能给出一个构造性的肯定回答.

考虑概率空间 $((0,1), \mathcal{B} \cap (0,1), m)$, 以及 r.v. U

$$U(t) = t,\ \forall\, t \in (0,1)\,.$$

易见, U 是初等概率论中讲过的在区间 $(0,1)$ 上均匀分布的随机变量. 因此, 对任意给定的 d.f. F, 考虑 r.v. $F^{\leftarrow} \circ U$ 的分布

$$\begin{aligned} P\left(F^{\leftarrow} \circ U \leqslant x\right) &= P\left(\{t \in (0,1), F^{\leftarrow}(U(t)) \leqslant x\}\right) \\ &= P\left(\{t \in (0,1) : F^{\leftarrow}(t) \leqslant x\}\right) \\ &= P(t \in (0,1), t \leqslant F(x)) = F(x) \end{aligned}$$

这样我们就证明了一个基本而有用的命题:

定理 35. 任何一个 d.f. F, 存在一个概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 以及它上面的 r.v. X 使得 $X \sim F$.

注. 实际上就是说, $((0,1), \mathcal{B} \cap (0,1), m)$ 到 $(\mathbb{R}, \mathcal{B})$ 的可测映射 $F^{\leftarrow}$ 在 $(\mathbb{R}, \mathcal{B})$ 上的像测度 (分布) 就是 F.

2.3.2 依分布收敛

设 $F_n, n=1,2,3,\cdots$ 和 F 都是单调递增的实函数. 若 $F_n(x)\to F(x)$ 对 F 的每一个连续点都成立, 则称 $\{F_n\}$ **弱收敛**到 F. 记为 $F_n \xrightarrow{w} F$. 分布函数作为特殊的单调实值函数, 当然也可以谈论弱收敛的问题.

定义 19. $\{X_n \sim F_n\}$ 是概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 上的随机变量序列, F 是一分布函数. 若 $F_n \xrightarrow{w} F$, 则称随机变量序列 $\{X_n\}$ 依分布收敛于 F, 记为 $X_n \xrightarrow{d} F$; 若还有随机变量 $X\sim F$, 则称随机变量序列 $\{X_n\}$ 依分布收敛于 X, 记为 $X_n \xrightarrow{d} f$.

我们来依分布收敛与其他收敛性之间的关系. 首先我们指出, 依概率收敛蕴涵了依分布收敛.

定理 36. $X, X_n, n=1,2,\cdots$ 是概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 上的随机变量, 则

$$X_n \xrightarrow{\mathbb{P}} X \Rightarrow X_n \xrightarrow{d} X\,.$$

证明. 记 X 分布函数为 F. 任取 $x\in\mathbb{R}$ 和 $\epsilon>0$,

$$\begin{aligned}\mathbb{P}(X_n\leqslant x) &= \mathbb{P}(X_n\leqslant x, |X_n-X|\leqslant\epsilon)+\mathbb{P}(X_n\leqslant x, |X_n-X|>\epsilon)\\ &\leqslant \mathbb{P}(f\leqslant x+\epsilon)+\mathbb{P}(|X_n-X|>\epsilon)\,.\end{aligned}$$

先令 $n\to\infty$, 再令 $\epsilon\to 0$, 便得到

$$\limsup_{n\to\infty}\mathbb{P}\left(X_n\leqslant x\right)\leqslant F(x)\,.$$

同样的方法, 有

$$\begin{aligned}\mathbb{P}(X_n> x) &= \mathbb{P}(X_n> x, |X_n-X|\leqslant\epsilon)+\mathbb{P}(X_n> x, |X_n-X|>\epsilon)\\ &\leqslant \mathbb{P}(f> x-\epsilon)+\mathbb{P}(|X_n-X|>\epsilon)\,.\end{aligned}$$

先令 $n\to\infty$, 再令 $\epsilon\to 0$, 得到

$$\liminf_{n\to\infty}\mathbb{P}\left(X_n\leqslant x\right)\geqslant F(x^-)\,.$$

由此即知, 当 x 是 F 的连续点时成立

$$\lim_{n\to\infty}\mathbb{P}\left(X_n\leqslant x\right)=F(x)\,.$$ □

注. 用同样的方法也可以证明此定理的一个推广: $X_n\xrightarrow{\mathbb{P}}0$, $Y_n\xrightarrow{d}Y$, 则有 $X_n+Y_n\xrightarrow{d}Y$. 这个命题在中心极限定理的相关内容中会用到.

依分布收敛和几乎必然收敛的关系是引人注意的: **依分布收敛可以在另一个概率空间用几乎必然收敛表示出来.** 其证明基于下列引理.

引理 37. *$F, F_n, n \in \mathbb{N}^+$ 皆为分布函数, 且 $F_n \xrightarrow{w} F$. 则有 $F_n^{\leftarrow} \xrightarrow{w} F^{\leftarrow}$.*

证明. 任取 t 为 $F^{\leftarrow}$ 的连续点,(反证) 若 $\lim_n F_n^{\leftarrow}(t) = F^{\leftarrow}(t)$ 不成立, 则成立以下二者之一

$$\limsup_{n\to\infty} F_n^{\leftarrow}(t) > F^{\leftarrow}(t)\,;\ \liminf_{n\to\infty} F_n^{\leftarrow}(t) < F^{\leftarrow}(t)\,.$$

- 若前者成立, 取 y 为 F 的连续点使得

  $$F^{\leftarrow}(t) < y < \limsup_{n\to\infty} F_n^{\leftarrow}(t)$$

  于是有子列 $\{n_k\}$ 使得 $F_{n_k}(y) < t$, 且 $t \leqslant F(y)$. 令 $k \to \infty$ 就有 $F(y) = t$. 但这与 t 点 $F^{\leftarrow}$ 右连续矛盾: 我们可取 $\epsilon > 0$ 充分小使 $F^{\leftarrow}(t+\epsilon) < y$, 这说明 $F(y) \geqslant t+\epsilon$, 矛盾.

- 若后者成立, 取 y 为 F 的连续点使得

  $$\liminf_{n\to\infty} F_n^{\leftarrow}(t) < y < F^{\leftarrow}(t)$$

  于是有子列 $\{n_k\}$ 使得 $F_{n_k}(y) \geqslant t$, 且 $t < F(y)$. 令 $k \to \infty$ 就有 $t < F(y) = t$, 矛盾.

于是我们用反证法证明了命题. □

定理 38 (Skorokhod)**.** *$X, X_n, n = 1, 2, \cdots$ 是概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量, 且 $X_n \xrightarrow{d} f$. 则存在概率空间 $(\widetilde{X}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, 和在它上面定义的随机变量 $\widetilde{X}, \widetilde{X}_n, n = 1, 2, \cdots$ 满足*

$$\widetilde{X} \stackrel{d}{=} X\,;\ \widetilde{X}_n \stackrel{d}{=} X_n, \forall\, n \in \mathbb{N}^+ \tag{2.4}$$

且

$$\widetilde{X}_n \xrightarrow{\text{a.s.}} \widetilde{X}\,. \tag{2.5}$$

证明. 设 X_n 的分布函数为 F_n, X 的分布函数为 F. 考察概率空间 $((0,1), \mathcal{B} \cap (0,1), m)$. 令

$$\widetilde{X} = F^{\leftarrow},\ \widetilde{X}_n = F_n^{\leftarrow}\,,$$

由定理 35 知 (2.4) 成立. 由引理 37, 并注意 $\widetilde{X} = F^{\leftarrow}$ 的不连续点至多可数, 知 (2.5) 成立. □

在学习完积分后, 我们可以给出依分布收敛的定价定义, 这个等价的定义十分重要, 是向更一般的空间 (Polish 空间) 中推广分布的弱收敛的出发点.

定理 39. $F, F_n, n = 1, 2, \cdots$ 皆是 $\mathbb{R}$ 上的分布函数, 则 F_n 弱收敛于 F 当且仅当对任意 $\mathbb{R}$ 上有界连续函数 g, 有

$$\int_{\mathbb{R}} g \, \mathrm{d}F_n \to \int_{\mathbb{R}} g \, \mathrm{d}F \ (n \to \infty) .$$

证明. 必要性. 任取 $\mathbb{R}$ 上有界连续函数 g. 实数 $a < b$ 是 F 的两个连续点. 我们先证明 $\int_{(a,b]} g \, \mathrm{d}F_n \to \int_{(a,b]} g \, \mathrm{d}F$.

由 g 在 $[a, b]$ 上的一致连续性, 任取 $\Delta > 0$, 存在 $\delta = \delta_\Delta$, 对任何 $\xi, \eta \in [a, b]$, 且 $|\xi - \eta| < \delta$, 必成立 $|g(\xi) - g(\eta)| < \Delta$. 取 $(a, b]$ 的分划,

$$a = x_0 < x_1 < \cdots < x_k = b ,$$

使得 $\{x_j\}$ 皆为 F 的连续点, 且 $\sup_j |x_j - x_{j-1}| < \delta$. 我们构造阶梯函数来逼近 g. 定义

$$\phi \coloneqq \sum_{j=1}^{k} g(x_j) I_{(x_{j-1}, x_j]} ,$$

可见在 $(a, b]$ 上总有 $|g - \phi| < \Delta$. 注意到 $\int_{(a,b]} \phi \, \mathrm{d}F_n \to \int_{(a,b]} \phi \, \mathrm{d}F$, 于是

$$\limsup_{n\to\infty} \left| \int_{(a,b]} g \, \mathrm{d}F_n - \int_{(a,b]} g \, \mathrm{d}F \right| \leqslant \limsup_{n\to\infty} \left| \int_{(a,b]} g \, \mathrm{d}F_n - \int_{(a,b]} \phi \, \mathrm{d}F_n \right|$$
$$+ \limsup_{n\to\infty} \left| \int_{(a,b]} \phi \, \mathrm{d}F_n - \int_{(a,b]} \phi \, \mathrm{d}F \right| + \limsup_{n\to\infty} \left| \int_{(a,b]} \phi \, \mathrm{d}F - \int_{(a,b]} g \, \mathrm{d}F \right| \leqslant 2\Delta$$

由 Δ 任意性知 $\limsup \left| \int_{(a,b]} g \, \mathrm{d}F_n - \int_{(a,b]} g \, \mathrm{d}F \right| = 0$.

现在很容易证明原来的命题. 任取 $\epsilon > 0$, 可取实数 a, b (与 $\epsilon > 0$ 有关), 皆为 F 的连续点, 且

$$\int_{\mathbb{R}\setminus(a,b]} g \, \mathrm{d}F < \epsilon.$$

由 $F_n \xrightarrow{w} F$, 可知存在正整数 $N = N_\epsilon$, 当 $n \geqslant N_\epsilon$ 时, 有

$$\int_{\mathbb{R}\setminus(a,b]} g \, \mathrm{d}F_n < \epsilon .$$

于是

$$\limsup_{n\to\infty} \left| \int_{\mathbb{R}} g \, \mathrm{d}F_n - \int_{\mathbb{R}} g \, \mathrm{d}F \right| \leqslant \limsup_{n\to\infty} \left| \int_{(a,b]} g \, \mathrm{d}F_n - \int_{(a,b]} g \, \mathrm{d}F \right|$$
$$+ \limsup_{n\to\infty} \left| \int_{\mathbb{R}\setminus(a,b]} g \, \mathrm{d}F_n - \int_{\mathbb{R}\setminus(a,b]} g \, \mathrm{d}F \right| \leqslant 2\epsilon .$$

由 ϵ 任意性知 $\limsup \left| \int_{\mathbb{R}} g \, \mathrm{d}F_n - \int_{\mathbb{R}} g \, \mathrm{d}F \right| = 0$. $\square$

注. 实际上利用 Skorokhod 定理, 后面要学的积分的变量替换公式, Lebesgue 控制收敛定理可以立刻得到本结果.

第三章 积分

积分是测度论最重要的概念之一. 和实变函数论中的 Lebesgue 积分一样, 一般测度空间上可测函数的积分也是通过典型方法定义的. 因此, Lebesgue 积分的许多重要性质可以拷贝过来. 概率空间中的积分就是数学期望. 学习了这一章以后, 初等概率论中没有, 也不可能给出证明的一些求期望的公式将得到严格的证明.

3.1 积分的建立

测度空间 $(X,\mathcal{F},\mu)$ 上可测函数积分的定义是用典型方法经过三个步骤来完成的.

3.1.1 非负简单可测函数的积分

f 是一个非负简单可测函数, $f=\sum_{j=1}^{n} a_j I_{A_j}$ 是其一个表达式, 其中 $\{A_j, j=1,\cdots,n\}$ 是 X 的可测分割, $a_j, j=1,\cdots,n$ 是非负实数. 称

$$\int_X f \ \mathrm{d}\mu \coloneqq \sum_{j=1}^{n} a_j \mu\left(A_j\right)$$

为 f 的**积分**. 容易证明上面的定义是良定 (well-defined).

下面, 我们来介绍非负简单函数积分的性质, 证明是简单的.

命题 40. *f,g 皆是非负简单函数, 则有*

1. $\int_X (af)\,\mathrm{d}\mu = a\int_X f\,\mathrm{d}\mu, \forall a\in[0,\infty)$.
2. $\int_X (f+g)\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu + \int_X g\,\mathrm{d}\mu$.
3. *若* $f\leqslant g$, *则* $\int_X f d\mu \leqslant \int_X g d\mu$.
4. *若可测集列* $\{A_n\}$ *是* X *的穷竭, 即* $A_n\uparrow X$, *则* $\int_X f 1_{A_n} d\mu \uparrow \int_X f d\mu$.

3.1.2 非负可测函数的积分

对于测度空间 $(X,\mathcal{F},\mu)$ 上的非负可测函数 f , 称

$$\int_X f\,\mathrm{d}\mu := \sup\left\{\int_X g\ \mathrm{d}\mu : g \text{ 是非负简单函数, 且 } g \leqslant f.\right\}$$

为 f 的积分. 容易证明用此定义给出的非负简单函数的积分与原定义是一致的.

定理 41 (Levi 引理)**.** *$f, f_n, n \in \mathbb{N}^+$ 皆是非负可测函数, 且 $f_n \uparrow f$. 则 $\int_X f_n\,\mathrm{d}\mu \uparrow \int_X f\,\mathrm{d}\mu$.*

证明. 由定义可以看出, 积分具有单调性. 故立得 $\{\int_X f_n\,\mathrm{d}\mu\}$ 是单调递增数列, 且

$$\lim_{n\to\infty}\int_X f_n\,\mathrm{d}\mu \leqslant \int_X f\,\mathrm{d}\mu\,.$$

现在不妨设 $\int_X f\,\mathrm{d}\mu > 0$ 且 $\lim_n \int_X f_n\,\mathrm{d}\mu < \infty$. 任取一非负简单可测函数 $g \leqslant f$, 和实数 $\lambda \in (0,1)$, 由于 $f_n \uparrow f$, 令

$$A_n = \{f_n \geqslant \lambda g\}\,.$$

则 $A_n \uparrow X$,

$$\int_X f_n\,\mathrm{d}\mu \geqslant \int_X f_n 1_{A_n}\,\mathrm{d}\mu \geqslant \int_X \lambda g 1_{A_n}\ \mathrm{d}\mu\,.$$

令 $n\to\infty$, 由命题 40 中 4, 则有

$$\lim_{n\to\infty}\int_X f_n\,\mathrm{d}\mu \geqslant \lambda \int_X g\,\mathrm{d}\mu\,.$$

取上确界, 并且令 $\lambda \uparrow 1$, 得到

$$\lim_{n\to\infty}\int_X f_n\,\mathrm{d}\mu \geqslant \int_X f\,\mathrm{d}\mu\,.$$

于是原定理得证. □

上面的 Levi 引理是十分有用的定理. 利用它, 我们把一些关于非负可测函数积分的命题转化为关于简单函数积分的问题, 例如, 证明下面的命题.

命题 42. *f, g 皆是非负可测函数, 则有*

1. $\int_X (af)\,\mathrm{d}\mu = a\int_X f\,\mathrm{d}\mu, \forall\, a \in [0,\infty)$.
2. $\int_X (f+g)\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu + \int_X g\,\mathrm{d}\mu$.
3. *若* $f \leqslant g$, *则* $\int_X f d\mu \leqslant \int_X g d\mu$.
4. *若* $\int_X f\,\mathrm{d}\mu = 0$, *则* $f = 0$ *a.e.* .

3.1.3 一般可测函数的积分

对一般的可测函数 f, 若 $\min\left\{\int_X f^+ \,\mathrm{d}\mu, \int_X f^- \,\mathrm{d}\mu\right\} < \infty$, 则称其积分存在或积分有意义, 且定义 f 的**积分**为

$$\int_X f \,\mathrm{d}\mu \coloneqq \int_X f^+ \,\mathrm{d}\mu - \int_X f^- \,\mathrm{d}\mu\,.$$

若 f 积分为有限值, 称 f 是**可积**的. 不至于混淆时, 我们将 $\int_X f \,\mathrm{d}\mu$ 简写为 $\int f \,\mathrm{d}\mu$.

注. 显然, f 可积当且仅当 $|f|$ 可积, 且 f 必然是几乎处处有限的. 同时, 我们有 $\int f \,\mathrm{d}\mu \leqslant \int |f| \,\mathrm{d}\mu$.

我们还可以定义 f 在可测集上的积分. 设 $A \in \mathcal{F}$. 只要可测函数 fI_A 的积分存在或可积, 我们就分别说 f 在集合 A 上**积分存在**或**可积**, 并且把

$$\int_A f \,\mathrm{d}\mu \coloneqq \int_X fI_A \,\mathrm{d}\mu$$

叫做 f **在集合 A 上的积分**. 显然只要可测函数 f 的积分存在, 则它在任何 $A \in \mathcal{F}$ 上的积分也存在. 此外, f 在零测集上的积分必然是零.

下面来考察积分的性质. 最自然也最重要的, 自然是线性性质.

命题 43. *f, g 是 $(X, \mathcal{F}, \mu)$ 上积分存在的可测函数. 则*

1. *对任意 $a \in \mathbb{R}$, af 的积分存在且 $\int_X (af) \,\mathrm{d}\mu = a \int_X f \,\mathrm{d}\mu$.*
2. *若 $\int f d\mu + \int g d\mu$ 有意义, 则 $f + g$ 几乎处处有定义, 且其积分存, 并成立*
$$\int (f+g) \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu\,.$$

证明. 1. 是显然的.

2. 不失一般性, 设 f^- 与 g^- 皆可积. 因此 f^-, g^- 皆是几乎处处有限的, 故 $f + g$ 几乎处处有定义. 且由

$$(f+g)^- \leqslant f^- + g^-\,.$$

知其可积, 故 $f + g$ 积分存在. 欲证

$$\int (f+g) \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu + \int g \,\mathrm{d}\mu$$

只需证

$$\int (f+g)^+ \,\mathrm{d}\mu - \int (f+g)^- \,\mathrm{d}\mu = \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu + \int g^+ \,\mathrm{d}\mu - \int g^- \,\mathrm{d}\mu\,.$$

移项就是,

$$\int (f+g)^+ \,\mathrm{d}\mu + \int f^- \,\mathrm{d}\mu + \int g^- \,\mathrm{d}\mu = \int (f+g)^- \,\mathrm{d}\mu + \int f^+ \,\mathrm{d}\mu + \int g^+ \,\mathrm{d}\mu\,.$$

由 $(f+g)^+ + f^+ + g^+ = (f+g)^- + f^- + g^-$ 和命题 42 的 2, 得证. □

下面我们考虑积分的单调性, 其证明是容易的.

命题 44. *f, g 是 $(X, \mathcal{F}, \mu)$ 上积分存在的可测函数. 则*

1. 若 $f \leqslant g$ a.e. , 则 $\int f \,\mathrm{d}\mu \leqslant \int g \,\mathrm{d}\mu$.

2. 若 $f = g$ a.e. , 则 $\int f \,\mathrm{d}\mu \leqslant \int g \,\mathrm{d}\mu$.

此外, 上面命题的 "逆命题" 也成立: 如果两个可积函数在任意可测集上积分有相同的大小关系, 则这两个函数 (几乎处处) 保持此大小关系.

命题 45. *f, g 是 $(X, \mathcal{F}, \mu)$ 上可积函数. 则有*

1. 若任何 $A \in \mathcal{F}$, 有 $\int_A g \,\mathrm{d}\mu \leqslant \int_A f \,\mathrm{d}\mu$, 则 $g \leqslant f$ a.e. .

2. 若任何 $A \in \mathcal{F}$, 有 $\int_A g \,\mathrm{d}\mu = \int_A f \,\mathrm{d}\mu$, 则 $g = f$ a.e. .

证明. 我们指出 1 已经蕴含了 2. 欲证 1, 只需证明

$$\mu\{g > f\} = 0\,.$$

由 $\int_{\{g>f\}} g \,\mathrm{d}\mu \leqslant \int_{\{g>f\}} f \,\mathrm{d}\mu$ 和 f, g 可积, 即有

$$\int (g-f) 1_{\{g>f\}} \,\mathrm{d}\mu = 0\,.$$

再由命题 42 的 4, 得证. □

上面我们假设了 f, g 可积. 若对 f, g 积分存在的情形, 我们必须对测度空间有一定要求才能使得上述命题成立

推论 46. *若 $(X, \mathcal{F}, \mu)$ 是 σ 有限的测度空间, f, g 是两积分存在的可测函数. 若任意 $A \in \mathcal{F}$,*

$$\int_A f \,\mathrm{d}\mu = \int_A g \,\mathrm{d}\mu\,.$$

则 $f = g$ a.e.

证明. 首先, 由 $(X, \mathcal{F}, \mu)$ 是 σ 有限的, 存在对 $\{|f| < \infty, |g| < \infty\}$ 的可测分划 $\{E_n : n \in \mathbb{N}_+\}$, 使得 f, g 在 E_n 上可积. 由命题 45 知, 对任意正整数 n, 在集 E_n 上 f 与 g 几乎处处相等. 故在 $\{|f| < \infty, |g| < \infty\}$ 上 f 与 g 几乎处处相等.

其次, 注意到

$$\begin{aligned}\{f \neq g\} \cap \{|f| < \infty, |g| < \infty\}^c = \{f = \infty, g < \infty\} \cup \{f < \infty, g = \infty\} \\ \cup \{f = -\infty, g > -\infty\} \cup \{f > -\infty, g = -\infty\}\end{aligned}$$

等式右边的四个集合皆为零测集, 以 $\{f = \infty, g < \infty\}$ 为例, 若其测度非零, 则存在正整数 N, 使得

$$\mu\{f = \infty, g < N\} > 0 \, .$$

由于 μ 为 σ 有限测度, 存在 $\{f = \infty, g < N\}$ 的可测子集 E, 使得 $0 < \mu(E) < \infty$. 但是

$$\int_E f \, \mathrm{d}\mu = \infty = \int_E g \, \mathrm{d}\mu \leqslant N\mu(E) \, .$$

这是一个矛盾! □

此外, 可积函数具有某种连续性: 在测度很小的可测集上的积分值也会很小.

定理 47 (绝对连续性). *设 f 可积. 则任何 $\epsilon > 0$, 存在 $\delta = \delta_\epsilon > 0$, 使得任何 $A \in \mathcal{F}$ 且 $\mu(A) < \delta$, 成立*

$$\int_A |f| \, \mathrm{d}\mu < \epsilon \, .$$

证明. 由 $|f|$ 可积, 对上述 $\epsilon > 0$, 存在非负简单可测函数 $g \leqslant |f|$ 使得

$$\int (|f| - g) \, \mathrm{d}\mu < \frac{\epsilon}{2} \, .$$

于是对任意可测集 A,

$$\int_A |f| \, \mathrm{d}\mu < \int_A g \, \mathrm{d}\mu + \frac{\epsilon}{2} \, .$$

设 g 的某个上界为 $M > 0$, 令 $\delta = \frac{\epsilon}{2M+1}$ 即可. □

3.2 积分号下取极限

本节我们来给出积分号和函数极限可以交换的几个充分条件. 他们分别是 Levi 引理, Fatou 引理和 Lebesgue 控制收敛定理.

在上一小节定理 41, 我们已经证明了非负可测函数的 Levi 引理. 它可以很轻易的推广为如下稍一般的形式.

定理 48 (Levi's lemma)**.** *$\{f_n\}$ 是递增的可测函数列, $f_n \uparrow f$. 若存在可积函数 g, 对任意正整数 n, $f_n \geqslant g$ a.e. , 则*

$$\int f_n \, \mathrm{d}\mu \uparrow \int f \, \mathrm{d}\mu.$$

注. 很容易给出对单调递减的可测函数列的情形: $\{f_n\}$ 是递减的可测函数列, $f_n \downarrow f$. 若存在可积函数 g, 对任意正整数 n, $f_n \leqslant g$. 则

$$\int f_n \, \mathrm{d}\mu \downarrow \int f \, \mathrm{d}\mu\,.$$

下面的介绍的 Fatou 引理更好应用, 但实际上它与 Levi 引理等价.

定理 49 (Fatou's lemma)**.** *$\{f_n\}$ 是一列可测函数. 若存在可积函数 g, 对任意正整数 n, $f_n \geqslant g$ a.e. . 则*

$$\int \liminf_{n\to\infty} f_n \ \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \ \mathrm{d}\mu\,.$$

证明. 对任意 $n \in \mathbb{N}^+$, 记 $h_n = \inf_{k\geqslant n} f_k$. 注意 $h_n \uparrow \liminf f_n$, 且对任意正整数 n, $h_n \geqslant g$ a.e. 于是由 Levi 引理得,

$$\lim_{n\to\infty} \int h_n \, \mathrm{d}\mu = \int \liminf_{n\to\infty} f_n \ \mathrm{d}\mu.$$

注意对任意正整数 n, $h_n \leqslant f_n$, 于是

$$\int h_n \, \mathrm{d}\mu \leqslant \int f_n \, \mathrm{d}\mu\,.$$

令 $n \to \infty$, 得到

$$\int \liminf_{n\to\infty} f_n \ \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \ \mathrm{d}\mu\,. \qquad \square$$

注. 自然可以考虑 $\{f_n\}$ 有 "上控" 函数的情形: $\{f_n\}$ 是一列可测函数列. 若存在可积函数 g, 对任意正整数 n, $f_n \leqslant g$. 则

$$\int \limsup_{n\to\infty} f_n \ \mathrm{d}\mu \geqslant \limsup_{n\to\infty} \int f_n \ \mathrm{d}\mu\,.$$

Fatou 引理还有一种有趣的情形, 就是依测度收敛形式的 Fatou 引理.

定理 50. $\{f_n\}$ 是一列可测函数, 且 f_n 依测度收敛于可测函数 f. 若存在可积函数 g, 对任意正整数 n, $f_n \geqslant g$. 则

$$\int f \,\mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu\,.$$

证明. 取 $\{f_n\}$ 的子列 $\{f_{k(n)}\}$ 使得

$$\lim_{n\to\infty} \int f_{k(n)} \,\mathrm{d}\mu = \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu\,.$$

由 $\{f_n\}$ 依测度收敛于 f, 存在 $\{f_{k(n)}\}$ 的子列, 记为 $\{h_n\}$, 几乎处处收敛于 f. 由对任意正整数 n, $h_n \geqslant g$, 应用 Fatou 引理得到

$$\int f \,\mathrm{d}\mu = \int \liminf_{n\to\infty} h_n \,\mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int h_n \,\mathrm{d}\mu = \liminf_{n\to\infty} \int f_n \,\mathrm{d}\mu\,. \qquad \square$$

注. 同样地, 若存在可积函数 g, 对任意正整数 n, $f_n \leqslant g$. 则有

$$\limsup_{n\to\infty} \int f \,\mathrm{d}\mu \leqslant \int f \,\mathrm{d}\mu\,.$$

最后我们来介绍 Lebesgue 控制收敛定理, 和它的一个推广. 因为证明方法类似, 我们仅给出推广形式的证明.

定理 51 (Lebesgue 控制收敛定理)**.** $\{f_n\}, f$ 皆是可测函数, $f_n \xrightarrow{\text{a.e.}} f$ 或 $f_n \xrightarrow{\mu} f$. 若存在非负可积函数 g, 对任意正整数 n 有 $|f_n| \leqslant g$. 则

$$\lim_{n\to\infty} \int |f_n - f| \,\mathrm{d}\mu = 0\,.$$

实际上, 我们有一个更强形式的 Lebesgue 控制收敛定理. 在之后的一节 L^p 空间中, 我们将看到它的威力.

定理 52. $\{f_n\}, \{g_n\}$ 皆是可积函数列, 且对任意正整数 n, $|f_n| \leqslant g_n$. $f_n \xrightarrow{\text{a.e.}} f$ 或 $f_n \xrightarrow{\mu} f$, $g_n \xrightarrow{\text{a.e.}} g$ 或 $g_n \xrightarrow{\mu} g$. 若 $\int g_n \,\mathrm{d}\mu \to \int g \,\mathrm{d}\mu$, 则

$$\lim_{n\to\infty} \int |f_n - f| \,\mathrm{d}\mu = 0\,.$$

证明. 先证 $f_n \xrightarrow{\text{a.e.}} f, g_n \xrightarrow{\text{a.e.}} g$ (或 $f_n \xrightarrow{\mu} f, g_n \xrightarrow{\mu} g$) 的情形. 对任意正整数 n, 定义

$$h_n := g_n + g - |f_n - f|.$$

可见 $h_n \geqslant 0$, 且 $h_n \xrightarrow{\text{a.e.}} 2g$ (或 $h_n \xrightarrow{\mu} 2g$), 由 Fatou 引理 (或依测度 Fatou 引理), 可见

$$\liminf_{n\to\infty} \int h_n \, \mathrm{d}\mu \geqslant 2 \int g \, \mathrm{d}\mu \,.$$

注意 $\int g_n \, \mathrm{d}\mu \to \int g \, \mathrm{d}\mu$, 就得到

$$\limsup_{n\to\infty} \int |f_n - f| \, \mathrm{d}\mu = 0 \,.$$

对于 $f_n \xrightarrow{\text{a.e.}} f, g_n \xrightarrow{\mu} g$ 的情形 ($f_n \xrightarrow{\mu} f, g_n \xrightarrow{\text{a.e.}} g$ 是类似的), 对 $\{g_n\}$ 的任意子列 $\{g_{k(n)}\}$, 存在该子列的子列 $\{g_{l(n)}\}$ 使得 $g_{l(n)} \xrightarrow{\text{a.e.}} g$, 同时 $f_{l(n)} \xrightarrow{\text{a.e.}} f$, 由第一步的证明,

$$\int |f_{l(n)} - f| \, \mathrm{d}\mu \to 0$$

由子列 $\{k(n)\}$ 的任意性,

$$\int |f_n - f| \, \mathrm{d}\mu \to 0 \,.$$

□

积分极定理的应用举例 利用本节所讲的积分的极限定理和之前常用的典型方法, 我们来证明下面在概率论中经常用到的 *积分变量替换定理*.

定理 53. *g 是测度空间 $(X, \mathcal{F}, \mu)$ 到可测空间 $(Y, \mathcal{S})$ 的可测映射.*

1. 定义 $(Y, \mathcal{S})$ 上的集函数 ν 如下

$$\nu(A) := \mu\left(g^{-1}A\right)$$

则 ν 是 $(Y, \mathcal{S})$ 上的测度, 我们将其记为 $\mu \circ g^{-1}$.

2. 对 $(Y, \mathcal{S})$ 上的任何可测函数 f, 积分 $\int_Y f \, \mathrm{d}(\mu \circ g^{-1})$ 存在当且仅当 $\int_X f \circ g \, \mathrm{d}\mu$ 存在, 且此时成立

$$\int_Y f \, \mathrm{d}(\mu \circ g^{-1}) = \int_X f \circ g \, \mathrm{d}\mu \,.$$

证明. 1 是简单的.

2 用典型方法来证. 对非负简单可测函数, 命题显然成立. 对非负可测函数 f, 取一列非负简单可测函数 $\{\phi_n\}$ 使得 $\phi_n \uparrow f$, 则由 Levi 引理

$$\int_Y f \, \mathrm{d}(\mu \circ g^{-1}) = \lim_{n\to\infty} \int_Y \phi_n \, \mathrm{d}(\mu \circ g^{-1}) = \lim_{n\to\infty} \int_X \phi_n \circ g \, \mathrm{d}\mu = \int_X f \circ g \, \mathrm{d}\mu \,.$$

命题成立. 对一般的可测函数 f, 由

$$\int_Y f^+ \, \mathrm{d}(\mu \circ g^{-1}) = \int_X f^+ \circ g \, \mathrm{d}\mu \,; \quad \int_Y f^- \, \mathrm{d}(\mu \circ g^{-1}) = \int_X f^- \circ g \, \mathrm{d}\mu$$

命题得证. □

3.3 空间 L^p

本开开始, 我们先引入一个方便的记号. 若 f 是 $(\Omega,\mathcal{F},\mu)$ 上积分存在的可测函数, 我们记

$$\mu(f)=\int f\,\mathrm{d}\mu\,.$$

设 $p\in(0,\infty)$. 我们将测度空间 $(X,\mathcal{F},\mu)$ 上的所有 p 幂可积函数全体记为 $L^p(X,\mathcal{F},\mu)$, 或简记为 L^p. 换言之,

$$L^p(X,\mathcal{F},\mu)\coloneqq\big\{f\text{ 为}(X,\mathcal{F})\text{ 上可测函数}:\mu(|f|^p)<\infty\big\}\,.$$

下面我们总是将两种情形: 对 $p\geqslant 1$, 定义 L^p 上的函数 $\|\cdot\|_p$ 为

$$\|f\|_p\coloneqq[\,\mu\,(|f|^p)\,]^{\frac{1}{p}}\ \text{ for all } f\in L^p\,.$$

对于 $p\in(0,1)$, 定义 L^p 上的函数 $((\cdot))_p$ 为

$$((f))_p\coloneqq\mu\,(|f|^p)\ \text{ for all } f\in L^p\,.$$

我们将要证明对 $p\geqslant 1$, $(L^p,\|\cdot\|_p)$ 是 Banach 空间; 对于 $p\in(0,1)$, 由 $((\cdot))_p$ 可诱导 L^p 上的距离 d_p 使得 L^p 称为完备的距离空间. 并且我们还要研究当 $p=\infty$ 时的情形: 空间 L^∞.

下面, 我们需要如下两个不等式, 由简单的微积分既可证明.

引理 54. *a,b 为任意实数. r, 为正实数, $1<p,q<\infty$ 且 $\frac{1}{p}+\frac{1}{q}=1$. 则*

$$|a+b|^r\leqslant C_r(|a|^r+|b|^r)\,,$$

$$|ab|\leqslant\frac{|a|^p}{p}+\frac{|b|^q}{q}\,.\qquad\text{(Young 不等式)}$$

由上面的引理, 对任意 $p\in(0,\infty)$, $f,g\in L^p$,

$$\mu\,(|f+g|^p)\leqslant C_p\,\mu\,(|f|^p+|g|^p)\,.$$

于是可见 $p\in(0,\infty)$, L^p 是线性空间.

3.3.1 空间 L^p, $p\in[1,\infty)$

下面我们来证明当 $p\geqslant 1$ 时, $\|\cdot\|_p$ 确为 L^p 上的范数. 显然正定性[1], 绝对齐次性皆成立. 我们要证明三角不等式. 当 $p=1$ 时是显然的. 当 $p>1$ 时, 利用上面的 Young 不等式, 我们容易得到 Hölder 不等式, 进而容易证明.

[1]需将 a.e. 相等的函数视为同一, 下同

引理 55. 设 $p, q \in (1, \infty)$, 且 $\frac{1}{p} + \frac{1}{q} = 1$. 则对可测函数 f, g, 有

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q . \quad \text{(Hölder 不等式)}$$

$$\|f+g\|_p \leqslant \|f\|_p + \|g\|_p . \quad \text{(Minkowski 不等式)}$$

证明. Hölder 不等式易证.为证明 Minkowski 不等式, 设不等式右边为正数. 注意

$$\begin{aligned}\int |f+g|^p d\mu &= \int |f+g||f+g|^{p-1} d\mu \\ &\leq \int |f||f+g|^{p-1} d\mu + \int |g||f+g|^{p-1} d\mu .\end{aligned}$$

而由 Hölder 不等式, 注意 $(p-1)q = p$

$$\int |f||f+g|^{p-1} d\mu \leqslant \|f\|^p \mu(|f+g|^p)^{\frac{1}{q}} ,$$

$$\int |g||f+g|^{p-1} d\mu \leqslant \|g\|^p \mu(|f+g|^p)^{\frac{1}{q}} ,$$

得到

$$\int |f+g|^p d\mu \leq \|f\|^p \mu(|f+g|^p)^{\frac{1}{q}} + \|g\|^p \mu(|f+g|^p)^{\frac{1}{q}} .$$

从而得证. □

由上面的定理, 可见 $(L^p, \|\|_p)$ 是赋范线性空间. 若 $\{f_n\}$ 是其中点列, 我们称其**依 L^p 收敛**, 若存在 $f \in L^p$, 使得

$$\|f_n - f\|_p \to 0 .$$

显然, L^p 收敛的极限是唯一确定的 (在 μ-a.e. 相等意义下), 此外, L^p 收敛蕴含依测度收敛. 事实上, 对任意 $\epsilon > 0$,

$$\mu\left(|f_n - f| \geqslant \epsilon\right) = \mu\left(|f_n - f|^p \geqslant \epsilon^p\right) \leqslant \frac{1}{\epsilon^p} \mu\left(|f_n - f|^p\right)$$

定理 56. 对任意 $p \geqslant 1$, $(L_p, \|\cdot\|_p)$ 是 *Banach* 空间.

证明. 现在设 $\{f_n\}$ 是 L^p 中的 Cauchy 列, 显然 $\{f_n\}$ 为依测度 Cauchy 列, 于是存在实值可测函数 f 使得 $f_n \xrightarrow{\mu} f$. 由 Fatou 引理知, 对任意正整数 n

$$\mu\left(|f_n - f|^p\right) \leqslant \varliminf_{m \to \infty} \mu\left(|f_n - f_m|^p\right)$$

可见 $\|f_n - f\|_p \to 0$, 同时得到 $f \in L^p$. □

L^p 空间中的收敛

我们来研究 L^p 空间依距离的收敛. 设 $f, f_n, n \in \mathbb{N}^+ \in L^p$. 若 $\{f_n\}$ 依 $(L^p, \|\cdot\|_p)$ 空间上的距离收敛于 f, 我们称 $\{f_n\}$ **依** L^p **收敛于** f 或 p **阶平均收敛于** f, 记为 $f_n \xrightarrow{L^p} f$.

定理 57. $\{f_n\} \subset L^p$, $f \in L^p$, 则

$$f_n \xrightarrow{L^p} f \Leftrightarrow f_n \xrightarrow{\mu} f \text{ 且} \|f_n\|_p \to \|f\|_p .$$

证明. 若 $f_n \xrightarrow{L^p} f$, 显然 $f_n \xrightarrow{\mu} f$ 且 $\|f_n\|_p \to \|f\|_p$.

若 $f_n \xrightarrow{\mu} f$ 且 $\|f_n\|_p \to \|f\|_p$, 注意到

$$|f_n - f|^p \leqslant 2^{p-1}\left(|f_n|^p + |f|^p\right)$$

从而可积函数列 $\{|f_n - f|^p\}$ 被可积函数列 $\{2^{p-1}(|f_n|^p + |f|^p)\}$ 控制, 且 $|f_n - f|^p \xrightarrow{\mu} 0$, $2^{p-1}(|f_n|^p + |f|^p) \xrightarrow{\mu} 2^p|f|^p$,

$$\int 2^{p-1}(|f_n|^p + |f|^p)\,\mathrm{d}\mu \to \int 2^p|f|^p\,\mathrm{d}\mu$$

由定理 52, 得到 $\|f_n - f\|_p \to 0$. □

定理 58. $\{f_n\} \subset L^p$, $f \in L^p$, 且 $f_n \xrightarrow{\text{a.e.}} f$. 则

$$f_n \xrightarrow{L^p} f \Leftrightarrow \|f_n\|_p \to \|f\|_p .$$

证明. 左边推出右边是显然的, 右边推出左边, 若 $f_n \xrightarrow{\mu} f$ 且 $\|f_n\|_p \to \|f\|_p$, 注意到

$$|f_n - f|^p \leqslant 2^{p-1}\left(|f_n|^p + |f|^p\right)$$

从而可积函数列 $\{|f_n - f|^p\}$ 被可积函数列 $\{2^{p-1}(|f_n|^p + |f|^p)\}$ 控制, 且 $|f_n - f|^p \xrightarrow{\text{a.e.}} 0$, $2^{p-1}(|f_n|^p + |f|^p) \xrightarrow{\text{a.e.}} 2^p|f|^p$,

$$\int 2^{p-1}(|f_n|^p + |f|^p)\,\mathrm{d}\mu \to \int 2^p|f|^p\,\mathrm{d}\mu$$

由定理 52, 得到 $\|f_n - f\|_p \to 0$. □

当 L^p 空间为 Banach 空间时, 我们知道除了依距离的收敛外, 还有弱收敛的概念, 首先我们介绍下面两个结果, 证明见 A.1 节.

- 当 $p > 1$ 时, $L^p(X, \mathcal{F}, \mu)$ 的对偶空间为 $L^q(X, \mathcal{F}, \mu)$, 其中 $\frac{1}{p} + \frac{1}{q} = 1$.

- 若 $(X,\mathcal{F},\mu)$ 为 σ 有限的测度空间, 则 $L^1(X,\mathcal{F},\mu)$ 的对偶空间为 $L^\infty(X,\mathcal{F},\mu)$[2]

于是, 我们引入函数序列弱收敛的定义如下.

定义 20. $\{f_n\}\subset L^p$, $f\in L^p$. 称 $\{f_n\}$ 在 L^p 中 **弱收敛**于 f, 若任意 $g\in L^q$, $\mu(f_n g)\to\mu(fg)$, 记为 $f_n\xrightarrow{w}f$.

注. 由 Hölder 不等式可见依 L^p 收敛蕴涵了弱收敛.

我们给出弱收敛的一个充分条件来结束本节.

定理 59. 设 $p>1$. 若 $\{f_n\}$ 是 L^p 中的有界序列且 $f_n\xrightarrow{\mu}f$ 或 $f_n\xrightarrow{\text{a.e.}}f$ 对某可测函数 f 成立, 则 $f\in L^p$, 且 $f_n\xrightarrow{w}f$.

证明. 第一步. 记 $M=\sup_n\|f_n\|_p$, 由 $f_n\xrightarrow{\text{a.e.}}f$ 和 Fatou 引理 (或依测度 Fatou 引理) 可见

$$\int|f|^p\mathrm{d}\mu\leqslant\liminf_{n\to\infty}\int|f_n|^p\mathrm{d}\mu\leqslant M^p<\infty$$

从而 $f\in L^p$.

第二步. 任意取定 $g\in L^q$. 我们断言: 对任意 $\epsilon>0$, 存在有限可测集 E, 与 ϵ 有关使得

$$\int_{E^c}|g|^q\,\mathrm{d}\mu<\epsilon\,.$$

为此, 考虑集列 $E_k=\{\frac{1}{k}\leqslant|g|\leqslant k\}$, $k\geqslant1$. 由 $|g|^q$ 可积, 可见 E_k 测度皆有限; 注意到 $E_k\uparrow\{|g|>0\}$, 有单调收敛定理

$$\int_{E_k}|g|^q\,\mathrm{d}\mu\uparrow\int|g|^q\,\mathrm{d}\mu$$

从而对给定的 $\epsilon>0$, 存在正整数 K, 与 ϵ 有关, 使得

$$\int|g|^q\,\mathrm{d}\mu-\int_{E_K}|g|^q\,\mathrm{d}\mu=\int_{E_K^c}|g|^q\,\mathrm{d}\mu<\epsilon\,.$$

断言得证.

第三步. 考察

$$\int|f_n-f||g|\,\mathrm{d}\mu=\int_E|f_n-f||g|\,\mathrm{d}\mu+\int_{E^c}|f_n-f||g|\,\mathrm{d}\mu\,.$$

第二项由 Hölder 不等式可以直接控制:

$$\int_{E^c}|f_n-f||g|\,\mathrm{d}\mu\leqslant\|f_n-f\|_p\,\|gI_{E^c}\|_q\leqslant 2M\epsilon^{1/q}\,.$$

[2] L^∞ 定义见 3.3.2 小节.

再来控制第一项, 注意在有限测度空间 $(E, E\cap\mathcal{F}, \mu)$ 上 $f_n \xrightarrow{\mu} f$ 与 $f_n \xrightarrow{\text{a.e.}} f$ 等价. 由 $|g|^q$ 的绝对连续性, 对上述 $\epsilon > 0$ 存在 δ_ϵ, 对任意可测集 A, 若 $\mu(A) < \delta_\epsilon$, 则

$$\int_A |g|^q \,\mathrm{d}\mu < \epsilon\,.$$

由 $f_n \xrightarrow{\mu} f$, 存在 $N = N_\epsilon$, 对任意 $n \geqslant N$,

$$\mu(E\cap\{|f_n - f| > \epsilon\}) < \delta_\epsilon\,.$$

于是

$$\begin{aligned}\int_E |f_n - f||g|\,\mathrm{d}\mu &= \int_{E\cap\{|f_n-f|\leqslant\epsilon\}} |f_n - f||g|\,\mathrm{d}\mu + \int_{E\cap\{|f_n-f|>\epsilon\}} |f_n - f||g|\,\mathrm{d}\mu \\ &\leqslant \int_E \epsilon|g|\,\mathrm{d}\mu + \|f_n - f\|_p \left\|gI_{E\cap\{|f_n-f|>\epsilon\}}\right\|_q \\ &\leqslant \epsilon\|g\|_q + 2M\epsilon^{1/q}\,.\end{aligned}$$

于是当 $n \geqslant N$ 时

$$\int |f_n - f||g|\,\mathrm{d}\mu \leqslant \epsilon\|g\|_q + 4M\epsilon^{1/q}\,.$$

从而有 $\mu(f_n g) \to \mu(fg)$. □

注. 本命题对 $p = 1$ 的情形不成立! 例如考虑 $([0,1], \mathcal{B}[0,1], \lambda)$, λ 为 Lebesgue 测度. 令 $f_n = nI_{[0,\frac{1}{n}]}$, 则 $\|f_n\|_1 \equiv 1$, $f_n \xrightarrow{\text{a.e.}} 0$, 但可见 $\{f_n\}$ 不弱收敛到 0.

练习 8. 考虑测度空间 $(\mathbb{N}, 2^{\mathbb{N}}, \mu)$, 其中 μ 为计数测度. 证明: L^1 中依范数收敛 (强收敛) 与弱收敛等价.

关于 Banach 空间 L^p 的可分性, 我们有

练习 9. $(X, \mathcal{F}, \mu)$ 为 σ 有限测度空间. 若 $\mathcal{F}$ **可分** [3], 则 L^p 为可分 Banach 空间.

3.3.2 空间 L^∞

我们将测度空间 $(X, \mathcal{F}, \mu)$ 上的几乎处处有界 (也成为本性有界) 的可测函数全体记为 $L^\infty(X, \mathcal{F}, \mu)$, 简记为 L^∞. 显然 L^∞ 是线性空间. 对任意 $f \in L^\infty$, 令

$$\|f\|_\infty \coloneqq \inf\left\{M \in \mathbb{R}^+ : |f| \leqslant M \text{ a.e.}\right\}\,.$$

[3] 即存在可数集类 $\mathcal{E}$ 使得 $\mathcal{F} = \sigma(\mathcal{E})$.

显然有 $f \leqslant \|f\|_\infty$ a.e. 从而容易看出, $\|\cdot\|_\infty$ 为一个范数. 同时我们有 "Hölder" 不等式: 若 $f \in L^\infty$, $g \in L_1$, 则

$$\|fg\|_1 \leqslant \|f\|_\infty \|g\|_1 .$$

下一命题的证明是不足道的.

命题 60. $(L^\infty, \|\cdot\|_\infty)$ *为一 Banach 空间.*

证明. 设 $\{f_n\}$ 为 $(L^\infty, \|\cdot\|_\infty)$ 中的 Cauchy 列. 令

$$N = \bigcup_{n,m\in\mathbb{N}} \{|f_n - f_m| > \|f_n - f_m\|_\infty\}$$

则 $\mu(N) = 0$. 且在 N^c 上, $\{f_n\}$ 为 Cauchy 列. 于是存在实值可测函数 f, $\{f_n\}$ 在 N^c 上一致收敛于 f, 显然 $f \in L_\infty$. □

易见, $(L^\infty, \|\cdot\|_\infty)$ 上依范数的收敛收敛是很简单的, 所以这里没什么值得研究的. 而 $(L^\infty, \|\cdot\|_\infty)$ 对偶空间较复杂, 我们这里不研究弱收敛. 关于 L^∞ 的对偶空间的表示, 可以参看严家安先生, 测度论讲义一书.

3.3.3 L^p 空间, $p \in (0, 1)$

对任意 $f \in L^p$, 定义

$$((f))_p \coloneqq \int |f|^p \mathrm{d}\mu . \tag{3.1}$$

易见此时成立正定性和三角不等式, 但不成立绝对齐次性:

$$\|af\|_p = |a|^p \|f\|_p, a \in \mathbb{R} .$$

实际上 $((\cdot))_p$ 是一个准范数 (paranorm). 虽然我们仅仅定义出 L^p 上的一个准范数, 但它仍可以诱导一个距离 d_p, 定义如下: 任何 $f, g \in L^p$, 令

$$d_p(f, g) \coloneqq ((f - g))_p .$$

容易看出, 准范数的三角不等式保证了 d 的三角不等式成立. 而正定性和对称性是显然的. 同样的, 若 $\{f_n\}$ 依照此距离收敛于 f, 蕴含了 $f_n \xrightarrow{\mu} f$. 从而与 $p \geqslant 1$ 情形一样, 距离都是完备的.

命题 61. 设 $p \in (0, 1)$. (L_p, d_p) *是完备的距离空间.*

证明. 现在设 $\{f_n\}$ 是 L^p 中的 Cauchy 列, 则 $\{f_n\}$ 为依测度 Cauchy 列, 于是存在实值可测函数 f 使得 $f_n \xrightarrow{\mu} f$. 由 Fatou 引理知, 对任意正整数 n

$$\mu\left(|f_n - f|^p\right) \leqslant \varliminf_{m\to\infty} \mu\left(|f_n - f_m|^p\right)$$

可见 $\|f_n - f\|_p \to 0$, 同时得到 $f \in L^p$. □

注. 实际上 $p \in (0,1)$ 时, L^p 为 F-空间[4], 但一般来说不是 Frechet 空间: L^p 是线性空间, 距离 d_p 满足平移不变性, 且加法与数乘皆连续, 故 L^p 为 F-空间. $L^p(0,1)$ 中非空凸开集只有它本身, 故不是局部凸的.(证明见 J.Conway. *A First Cousre in Funtional Analysis*, Chapter IV, 3.16 Example).

我们来研究 L^p 空间依距离的收敛. 设 $f, f_n, n \in \mathbb{N}^+ \in L^p$. 若 $\{f_n\}$ 依 L^p 上的距离收敛于 f, 我们也称 $\{f_n\}$ p **阶平均收敛于** f, 记为 $f_n \xrightarrow{L^p} f$. 下面的与 $p \geqslant 1$ 时是一样的, 证明不需要改变.

命题 62. *$\{f_n\} \subset L^p$, $f \in L^p$, 则*

$$f_n \xrightarrow{L^p} f \Leftrightarrow f_n \xrightarrow{\mu} f \text{ 且} \|f_n\|_p \to \|f\|_p \,.$$

命题 63. *$\{f_n\} \subset L^p$, $f \in L^p$, 且 $f_n \xrightarrow{\text{a.e.}} f$. 则*

$$f_n \xrightarrow{L^p} f \Leftrightarrow \|f_n\|_p \to \|f\|_p \,.$$

当 $p \in (0,1)$ 时, 我们上面说明了 $L^p(0,1)$ 中的非空凸开集只有它本身, 可以证明 L^p 上的连续线性泛函只有 0. 研究弱收敛也就没有意义了.

3.4 概率空间上的积分

概率论的中的数学期望, 用测度论的语言来说, 就是在特殊的测度空间—概率空间上的积分. 若概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量 (或可测函数) X 积分存在, 也它的数学期望存在, 称

$$\mathbb{E}X := \int_\Omega X \mathrm{d}\mathbb{P}$$

为 X 的**期望**. 若 X 是可积的, 也说它的期望是有限的.

关于期望的计算, 由积分的变量替换, 定理 53, 立刻得到

[4] 一个拓扑线性空间, 其拓扑由一个平移不变的完备距离诱导, 称为 F-空间.

命题 64. X 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量, 其分布记为 $\mathbb{P}_X$. 则对 $\mathbb{R}$ 上的任何 *Borel* 可测函数 f, $f(X)$ 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的可测函数, 且只要

$$\mathbb{E}[f(X)] = \int_{\mathbb{R}} f(x)\, \mathbb{P}_X(\mathrm{d}x)\,.$$

的一端有意义, 另一端也有意义且等式成立.

在概率空间的情形, 我们有一个特殊的不等式.

定理 65 (Jensen 不等式)**.** X 为 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间上可积的随机变量, φ 为 $\mathbb{R}$ 上下凸函数. 则 $\varphi(X)$ 期望存在, 且满足

$$\varphi(\mathbb{E}X) \leqslant \mathbb{E}[\varphi(X)]\,.$$

证明. 用 φ'_+ 表示 φ 的右导数, 则由 φ 下凸, 我们有

$$\varphi(x) \geqslant \varphi(\mathbb{E}X) + \varphi'_+(\mathbb{E}X)(x - \mathbb{E}X)\,,\ \text{for } x \in \mathbb{R}\,.$$

于是

$$\varphi(X) \geqslant \varphi(\mathbb{E}X) + \varphi'_+(\mathbb{E}X)(X - \mathbb{E}X)\,.$$

可见 $\varphi(X)$ 期望存在, 且两边取期望得 Jensen 不等式. □

显然, 以前对一般测度空间得到的那些结论对概率空间都是成立的. 但由于概率空间是有限的测度空间, 所以它还有一些特殊的性质. 例如, 对于概率空间而言, 有下列关系

$$L_t \subset L_s,\ \forall\, 0 < s \leqslant t < \infty\,.$$

且利用 Hölder 不等式可以证明 $\|X\|_s \leqslant \|X\|_t$ 对任意 $1 \leqslant s < t$ 成立.

3.4.1 一致可积

本节的最后进行概率空间中随机变量序列各种收敛性之间关系的讨论. 在这些讨论中, 一致可积的概念起着重要作用. 我们将先介绍这个概念, 给出它的判别方法.

定义 21. 称随机变量族 $\{X_t, t \in T\}$ **一致可积**. 若

$$\sup_{t\in T} \mathbb{E}|X_t| 1_{\{|X_t| > \lambda\}} \to 0,\ \text{as } \lambda \to \infty\,.$$

注. 显然, 有限族的可积随机变量是一致可积的. $\{X_t, t \in T\}$ 有 (一致) 的可积控制函数时必然一致可积.

与一致可积有关的是随机变量族 $\{X_t, t \in T\}$ 一致绝对连续的概念.

定义 22. 称随机变量族 $\{X_t, t \in T\}$ **一致绝对连续**, 若任意 $\epsilon > 0$, 存在 $\delta = \delta_\epsilon > 0$, 使对任何可测集 A, 只要 $\mathbb{P}(A) < \delta$, 就有

$$\sup_{t \in T} \mathbb{E}\,|X_t|\, 1_A < \epsilon$$

由积分的绝对连续性, 我们知道有限个可积的随机变量组成的集合一定是一致绝对连续的.

定理 66. *一族随机变量 $\{X_t, t \in T\}$ 一致可积的充要条件是, 它既一致绝对连续又 L^1 有界.*

证明. 若 $\{X_t, t \in T\}$ 一致可积, 对给定的 $\epsilon > 0$, 可取 λ_ϵ 充分大, 使得对任意 $t \in T$,

$$\mathbb{E}\,|X_t|\, 1_{\{|X_t| > \lambda\}} =< \frac{\epsilon}{2}\,.$$

于是对任意 $A \in \mathcal{F}$,

$$\mathbb{E}\,|X_t|\, 1_A = \mathbb{E}\,|X_t|\, 1_{A \cap \{|X_t| > \lambda\}} + \mathbb{E}\,|X_t|\, 1_{A \cap \{|X_t| \leqslant \lambda\}} \leqslant \lambda_\epsilon\, \mathbb{P}(A) + \frac{\epsilon}{2}\,.$$

令 $A = \Omega$, 知 $\{X_t\}$ 为 L^1 有界. 同时可见 $\{X_t\}$ 一致绝对连续.

若 $\{X_t, t \in T\}$ 一致绝对连续又 L^1 有界, 记 $M = \sup_{t \in T} \|X_t\|_1$. 对任意 $\epsilon > 0$, 由一致绝对连续性, 存在 $\delta = \delta_\epsilon > 0$, 使对任何可测集 A, 只要 $\mathbb{P}(A) < \delta$, 就有

$$\mathbb{E}\,|X_t|\, 1_A < \epsilon\,, \text{ for all } t \in T\,.$$

注意到

$$\mathbb{P}(|X_t| > \lambda) \leqslant \frac{\mathbb{E}|X_t|}{\lambda} \leqslant \frac{M}{\lambda}\,.$$

可见当 λ 充分大时, $\mathbb{P}(|X_t| > \lambda) \leqslant \delta$, 从而得到

$$\mathbb{E}\,|X_t|\, 1_{\{|X_t| > \lambda\}} < \epsilon\,, \text{ for all } t \in T\,.$$ □

一致可积的概念可以刻画概率空间中依 L^p 收敛.

定理 67. *$p \in [1, \infty)$, $X, X_n (n \in \mathbb{N}) \in L^p$. 则*

$$X_n \xrightarrow{L^p} X \iff X_n \xrightarrow{\mathbb{P}} X \text{ 且} \{|X_n|^p\} \text{ 一致可积.}$$

证明. 若 $X_n \xrightarrow{L^p} X$, 则显然 $X_n \xrightarrow{\mathbb{P}} X$, 且 $\{|X_n|^p\}$ L^1 有界. 对任意 $\epsilon > 0$ 和 $A \in \mathcal{F}$,

$$\mathbb{E}|X_n|^p 1_A \leqslant 2^{p-1}\left[\mathbb{E}\big(|X_n - X|^p 1_A\big) + \mathbb{E}\,|X|^p 1_A\right]$$

可见 $\{|X_n|^p\}$ 一致绝对连续. 由定理 66 知 $\{|X_n|^p\}$ 一致可积.

若 $\{|X_n|^p\}$ 一致可积, 且 $X_n \xrightarrow{\mathbb{P}} X$. 则任取 $\epsilon > 0$, 注意到

$$\mathbb{E}|X_n - X|^p = \mathbb{E}\big(|X_n - X|^p 1_{\{|X_n - X| > \epsilon\}}\big) + \mathbb{E}\big(|X_n - X|^p 1_{\{|X_n - X| \leqslant \epsilon\}}\big)$$

易知 $\{|X_n - X|^p\}$ 一致绝对连续, 于是存在 $\delta = \delta_\epsilon$, 对任意 $A \in \mathcal{F}$ 满足 $\mathbb{P}(A) \leqslant \delta$, 有

$$\mathbb{E}\big(|X_n - X|^p 1_A\big) \leqslant \epsilon\,, \text{ for all } n\,.$$

注意当 n 充分大时, 由 $X_n \xrightarrow{\mathbb{P}} X$,

$$\mathbb{P}(|X_n - X| > \epsilon) \leqslant \delta\,.$$

可见 $\mathbb{E}|X_n - X|^p \to 0$. □

推论 68. *$X, X_n (n \in \mathbb{N}) \in L^1$, 且 $X_n \xrightarrow{\mathbb{P}} X$ 或 $X_n \xrightarrow{\text{a.s.}} X$. 则*

$$\|X_n - X\|_1 \to 0 \Leftrightarrow \{X_n\} \text{ 一致可积} \Leftrightarrow \|X_n\| \to \|X\|\,.$$

下一定理给出了一致可积性的又一准则, 其证明是十分容易的.

定理 69. *$\{X_t, t \in T\}$ 为一随机变量族. φ 是定义在 $\mathbb{R}_+$ 上的非负 Borel 函数, 满足*

$$\frac{\varphi(t)}{t} \to \infty\,,\ t \to \infty\,.$$

若

$$\sup_{t \in T} \mathbb{E}\,\varphi(|X_t|) < \infty$$

则 $\{X_t, t \in T\}$ 一致可积.

注. 值得注意的是, 本定理的逆命题也是成立的: 若 $\{X_t, t \in T\}$ 一致可积, 则存在上述要求的 φ 使得 $\{\varphi(|X_t|)\}$ 是 L^1 有界的. φ 的构造可参见严家安先生测度论讲义 7.4.5 定理.

练习 10. f, g 是 $\mathbb{R}_+$ 上 Borel 可测函数, 其中 g 非负, 且

$$\frac{|f(t)|}{g(t)} \to 0\,,\ t \to \infty\,.$$

若

$$\sup_{t\in T} \mathbb{E}\, g(|X_t|) < \infty$$

证明: $\{f(X_t), t \in T\}$ 一致可积.

第四章　符号测度

在微积分中, 如果函数 f 在区间 $[a,b]$ 上连续, 那么

$$F(x)=\int_a^x f(y)\,\mathrm{d}y$$

称为不定积分, 而 f 是 F 的导数. 现在设 f 是测度空间 $(X,\mathcal{F},\mu)$ 上积分存在的可测函数, 我们可以在 $\mathcal{F}$ 上定义集函数 φ 为

$$\varphi(A)\coloneqq\int_A f\,\mathrm{d}\mu\,,\ \forall\ A\in\mathcal{F}$$

称之为 f 关于 μ 的不定积分, 记为 $\varphi=f.\mu$, 而 f 也叫做 φ 对测度 μ 的导数. 我们注意, φ 满足测度的如下两条性质:

1. 空集上为 $0:\varphi(\varnothing)=0$.

2. 可数可加性: 若 $\{A_n\}$ 是 $\mathcal{F}$ 中两两不交的集列, 则

$$\varphi(A)=\sum_{n=1}^{\infty}\varphi(A_n)\,.$$

但是 φ 不一定满足非负性. 若定义 φ^+, φ^- 分别为

$$\varphi^+(A)\coloneqq\int_A f^+\,\mathrm{d}\mu\,,\ \varphi(A)^-\coloneqq\int_A f^-\,\mathrm{d}\mu\,,\forall\ A\in\mathcal{F}$$

则 φ^+, φ^- 皆为 $\mathcal{F}$ 上的测度, 且其中有一个为有限测度, 同时

$$\varphi=\varphi^+-\varphi^-\,.$$

一般地, 若有 $\mathcal{F}$ 上的集函数, 也像 ϕ 一样, 除了非负性以外满足测度的所有其他性质, 是否一定就有导数呢? 是否也有上面的分解式呢? 本章将围绕这一中心问题展开讨论. 学完这一章以后, 概率论中几乎所有的重要概念都奠定了严格的数学基础.

4.1 符号测度, Hahn 分解与 Jordan 分解

定义 23. $(X,\mathcal{F})$ 是一个可测空间, 从 $\mathcal{F}$ 到 $[-\infty,\infty]$ 的集函数 φ 称为**符号测度**, 若它满足

1. 空集上为 $0:\varphi(\varnothing)=0$.
2. 可数可加性: 若 $\{A_n\}$ 是 $\mathcal{F}$ 中两两不交的集列, 则

$$\varphi(A)=\sum_{n=1}^{\infty}\varphi(A_n)\,.$$

我们注意, 虽然定义中 φ 的值域为 $[-\infty,\infty]$, 但实际上只可能有下面两种情况:

$$-\infty<\varphi\leqslant\infty\,,$$
$$-\infty\leqslant\varphi<\infty\,.$$

若不然, 存在 $A,B\in\mathcal{F}$ 使得 $\varphi(A)=\infty$, $\varphi(B)=-\infty$, 注意到

$$\varphi(A\cup B)=\varphi(B)+\varphi(A\backslash B)\,,$$
$$\varphi(A\cup B)=\varphi(A)+\varphi(B\backslash A)\,.$$

为了使得第一行有意义, 需要 $\varphi(A\cup B)=\infty$, 为了使得第二行有意义, 需要 $\varphi(A\cup B)=-\infty$, 这是矛盾的!

例 8. 首先, 本章开始提到的不定积分都是符号测度. 若设 μ,ν 皆为 $(X,\mathcal{F})$ 上的测度, 且其中一个是有限测度, 则 $\mu-\nu$ 也是符号测度.

注意上面例子中的符号测度, 都可以表示为两个测度之差, 且其中之一为有限测度. 这一结论对任意的符号测度成立吗? 我们下面来研究这个问题.

4.1.1 Hahn 分解与 Jordan 分解

考虑 f 对 μ 的不定积分 φ, 他是怎样分解为两个测度的差的呢?

$$\varphi^+(A)=\int_A f^+\,\mathrm{d}\mu=\varphi(A\cap\{f\geqslant 0\})\,,$$
$$\varphi^+(A)=\int_A f^-\,\mathrm{d}\mu=-\varphi(A\cap\{f<0\})\,.$$

若我们记 $X^+=\{f\geqslant 0\}$, $X^-=\{f<0\}$, 则显然他是 X 的可测分割, 满足

$$A\in\mathcal{F},A\subset X^+\Rightarrow\varphi(A)\geqslant 0\,,$$
$$A\in\mathcal{F},A\subset X^-\Rightarrow\varphi(A)\leqslant 0\,.$$

我们来效仿这个思路, 若能找出类似的分解, 问题就十分容易了.

定理 70 (Hahn 分解)**.** *对 $(X,\mathcal{F})$ 上的符号测度 φ, 存在 $\mathcal{F}$ 可测集 X^+, X^- 为 X 的一个划分, 且*

$$\varphi(A\cap X^+)\geqslant 0,\ \varphi(A\cap X^-)\geqslant 0\ ,\forall\ A\in\mathcal{F}\,.$$

证明. 我们要分几步来证明这个定理. 不失一般性, 假设 $\varphi > -\infty$.

第一步. 考虑 X^+ 具有的性质, 我们引入一个概念: 称 A 为*正集*, 若 $A\in\mathcal{F}$, 且对 A 的任何可测子集 B, $\varphi(B)\geqslant 0$. 类似的, 有*负集*的概念. 记 $\mathcal{F}_+$, $\mathcal{F}_-$ 为所有正集, 负集组成的集类. 显然 $\mathcal{F}_+$, $\mathcal{F}_-$ 是环. 进一步, 由 φ 的可数可加性知, $\mathcal{F}_+$, $\mathcal{F}_-$ 是 σ 环.

第二步. 我们找出 "最大的" 负集. 令

$$\alpha=\inf\{\varphi(A): A\in\mathcal{F}^-\}\,.$$

故存在负集列 $\{A_n\}$ 使得 $\varphi(A_n)\to\alpha$. 令 $X^-=\cup_{n=1}^\infty A_n$, 则 X^- 为负集且

$$\alpha\leqslant\varphi(X^-)=\varphi(A_n)+\varphi(X^-\backslash A_n)\leqslant\varphi(A_n),\ \forall n\in\mathbb{N}^+$$

令 $n\to\infty$ 知, $\varphi(X^-)=\alpha$, 从而可见 $\alpha>-\infty$.

第三步. 令 $X^+=X\backslash X^-$, 我们只需要证明 X^+ 是一正集. 直接证明是困难的, 但我们不难发现 X^+ 具有如下的性质: X^+ 的可测子集中没有 "真正的" 负集: 不存在 $A\subset X^+$ 为负集, 且 $\varphi(A)<0$.

否则 $X^-\cup A$ 是负集, 且 $\varphi(X^-\cup A)=\varphi(X^-)+\varphi(A)<\alpha$, 但是这与 α 的定义矛盾!

第四步. 我们证明利用在第三步中指出的 X^+ 的性质, 证明 X^+ 是正集.

反证: 假设 X^+ 不是正集, 则存在 X^+ 可测子集 A_0 使得, $\varphi(A_0)<0$. 由 A_0 不是负集, 故 $\{B\subset A_0:\varphi(B)>0\}$ 非空. 令

$$\alpha_0=\sup\{\varphi(B): B\subset A_0,\varphi(B)>0\}>0\,,$$

取 A_0 的可测子集 B_0, 使得 $\varphi(B_0)>\min\{1,\frac{\alpha_0}{2}\}$.

注意到 $\varphi(A_0\backslash B_0)<0$, 故 A_0-B_0 亦不为负集, 故 $\{B\subset A_0-B_0:\varphi(B)>0\}$ 非空. 令

$$\alpha_1=\sup\{\varphi(B): B\subset A_0-B_0,\varphi(B)>0\}\,,$$

取 $A_0 - B_0$ 的可测子集 B_1, 使得 $\varphi(B_1) > \min\{1, \frac{\alpha_1}{2}\}$.

现在设 B_n 已由上面定义: $B_n \subset A_0 - \cup_{k=0}^{n-1} B_k$, $\varphi(B_n) > 0$. 因此 $A_0 - \cup_{k=0}^{n} B_k$ 不为负集, 从而 $\{B \subset A_0 - \cup_{k=0}^{n} B_k : \varphi(B) > 0\}$ 非空. 令

$$\alpha_{n+1} = \sup\{B \subset \cup_{k=0}^{n} B_k : \varphi(B) > 0\}$$

则存在 $A_0 - \cup_{k=0}^{n} B_k$ 的可测子集 B_{n+1} , $\varphi(B_n) > \min\{1, \frac{\alpha_n}{2}\} > 0$.

注意到 $\{B_n\}$ 是互不相交的 A_0 的可测子集, 且 $\varphi(A_0) < 0$, 于是

$$\sum_{n=0}^{\infty} \min\{1, \frac{\alpha_n}{2}\} \leqslant \sum_{n=0}^{\infty} \varphi(B_n) < \infty\,.$$

这蕴涵了 $\alpha_n \to 0$. 我们考虑

$$A_1 \coloneqq A_0 - \bigcup_{n=0}^{\infty} B_n$$

显然 $\varphi(A_1) < 0$, 由 A_1 不为负集, 存在其可测子集 B, $\varphi(B) > 0$. 但是注意到对任意 $n \in \mathbb{N}_0$

$$B \subset A_1 = A_0 - \bigcup_{k=0}^{\infty} B_k \subset A_0 - \bigcup_{k=0}^{n} B_k$$

由 α_n 的定义, 有

$$0 < \varphi(B) \leqslant \alpha_n$$

令 $n \to \infty$ 可得 $0 < \varphi(B) \leqslant 0$, 这是一个矛盾. □

注. Hahn 分解也具有如下意义下的唯一性: 若 $\tilde{X}^+$, $\tilde{X}^-$ 也是 φ 的 Hahn 分解, 则 $X^+ \Delta \tilde{X}^+ = X^- \Delta \tilde{X}^-$ 的可测子集皆为零测集. 其证明是容易的.

对于一个任意给定的符号测度 φ, Hahn 分解把空间 X 分解为两部分 X^+ 和 X^- : 在 X^+ 上 φ 只取正值; 在 X^- 上 φ 只取负值. 于是, 只要令

$$\varphi^+(A) = \varphi(A \cap X^+)\,,\ \varphi^-(A) = \varphi(A \cap X^-)\,,\ \forall\ A \in \mathcal{F}$$

就得到了分解式, 从而回答了本小节开始提出的问题. 注意, 任意 $A \in \mathcal{F}$,

$$\varphi(A \cap X^+) = \sup\{\varphi(B) : B \subset A, B \in \mathcal{F}\}\,.$$

我们把这个结论写成下面的形式.

定理 71 (Jordan 分解)**.** 对可测空间 $(X,\mathcal{F})$ 上的符号测度 φ, 存在测度 φ^+ 和 φ^-, 其中有一个为有限测度, 使得对任意 $A\in\mathcal{F}$,

$$\begin{aligned}\varphi^+(A)&=\sup\{\varphi(B):B\subset A,B\in\mathcal{F}\}\,,\\ \varphi^-(A)&=\sup\{-\varphi(B):B\subset A,B\in\mathcal{F}\}\,,\end{aligned}\tag{4.1}$$

且有

$$\varphi=\varphi^+-\varphi^-\,.$$

我们称上面的分解为符号测度 φ 的 **Jordan 分解**. 显然, 由 (4.1) 可知 Jordan 分解是唯一的. 但是注意, 若我们想找两个测度 μ,ν 使得 $\varphi=\mu-\nu$, 这样的 μ,ν 显然不是唯一的. 容易证明, Jordan 分解满足

$$\varphi^+\leqslant\mu\;,\;\varphi^-\leqslant\nu\,.$$

我们将测度 φ^+, φ^- 分别称为 φ 的**正部**与**负部**. 并且定义测度

$$|\varphi|\coloneqq\varphi^++\varphi^-$$

为 φ 的**全变差测度**. 称 $|\varphi|(X)$ 为 φ 的全变差, 记为 $\|\varphi\|_{var}$.

定义 24. 称符号测度 φ 为**有限**的, 若 $|\varphi|$ 为有限测度; 称 φ 为 **σ-有限**的, 若 $|\varphi|$ 为 σ-有限测度.

显然, φ 是有限符号测度当且仅当 $|\varphi(X)|<\infty$; φ 是 σ-有限符号测度当且仅当存在 X 的可测分割 $\{A_n\}$, 使得对任意 n 有 $|\varphi(A_n)|<\infty$.

4.2 Lebesgue 分解与 Radon-Nikodym 定理

$(X,\mathcal{F},\mu)$ 是一个测度测度, 考虑该测度空间上的符号测度 φ. 本节我们将要讨论的问题就是: 什么时候 φ 是某个积分存在的可测函数 f 关于 μ 的不定积分? 显然 $\varphi=f.\mu\ll\mu$, 显然他们满足如下的关系:

$$\mu(A)=0\Rightarrow|\varphi|\,(A)=0\,,\;\;\forall\,A\in\mathcal{F}\,.$$

下面我们将抽象出这个性质, 将引入符号测度之间绝对连续性. 另一方面, 很容易举出不是 μ 的不定积分的符号测度: 假设 $x\in X$ 满足 $\mu\{x\}=0$, 对不定积分 $f.\mu$, 符号测度

$$\delta_x+f.\mu$$

就不再是 μ 的不定积分了. 我们将 δ_x 与 μ 的关系抽象出来, 引入了符号测度之间的相互奇异的概念.

定义 25. 设 φ_1, φ_2 是可测空间 $(X, \mathcal{F})$ 上的符号测度, 我们称 φ_2 关于 φ_1 **绝对连续**, 记为 $\varphi_2 \ll \varphi_1$, 若

$$|\varphi_2|(A) = 0 \Rightarrow |\varphi_1|(A) = 0, \ \forall A \in \mathcal{F}.$$

若 $\varphi_2 \ll \varphi_1$ 且 $\varphi_2 \ll \varphi_1$, 则称 φ_1 与 φ_1 **等价**, 记为 $\varphi_1 \sim \varphi_2$. 若存在 $N \in \mathcal{F}$, 使得

$$|\varphi_1|(N^c) = |\varphi_2|(N) = 0$$

则称 φ_1 与 φ_1 **相互奇异**, 记为 $\varphi_1 \perp \varphi_2$.

设 φ 为 $(X, \mathcal{F})$ 上的符号测度. 若 $N \in \mathcal{F}$, 使得 $|\varphi|(N^c) = 0$, 则称 N 为 φ 的**支撑**, 一般说来, 支撑并非唯一确定.

注.

1. $\varphi_2 \ll \varphi_1$ 当且仅当 φ_1 的支撑必是 φ_2 的支撑; $\varphi_1 \perp \varphi_2$ 当且仅当 φ_1 与 φ_2 有不相交支撑.
2. 若 $\varphi_2 \ll \varphi_1$ 且 $\varphi_1 \perp \varphi_2$, 则 $\varphi_2 = 0$.
3. μ, ν 皆为测度, 其中一个是有限测度. 则 $\mu \perp \nu$ 当且仅当 $(\mu - \nu)^+ = \mu$, $(\mu - \nu)^- = \nu$.

设 h 是 $(X, \mathcal{F}, \mu)$ 上的非负可测函数, 则不定积分 $h.\mu$ 是一个测度. 可测函数 f 关于这个测度的积分, 与关于 μ 的积分存在什么联系吗? 利用典型方法, 容易得到类似于微积分基本定理的如下引理.

引理 72. *设 $(X, \mathcal{F}, \mu)$ 为测度空间, h 是一非负可测函数, 令 $h.\mu$ 表示 h 关于 μ 的不定积分. 对任何可测函数 f, f 关于测度 $h.\mu$ 积分存在当且仅当 gh 关于 μ 的积分存在. 这时有*

$$\int f \, \mathrm{d}(h.\mu) = \int fh \, \mathrm{d}\mu.$$

这个引理在下面的定理中的应用基于如下的事实: 我们总可以将一个 σ 有限测度看作一个有限测度的不定积分. 因此, 对这个 σ 有限测度的积分可以转化为对有限测度的积分.

4.2.1 Lebesgue 分解

定理 73 (Lebesgue 分解)**.** *φ 是 σ 有限测度空间 $(X, \mathcal{F}, \mu)$ 上 σ-有限符号测度. 则 φ 有如下的唯一分解*

$$\varphi = \varphi_c + \varphi_s,$$

其中 φ_c, φ_s 为 σ 有限符号测度, 满足 $\varphi_c \ll \mu$, $\varphi_s \perp \mu$.

证明. 考虑 μ 非零的情形. 由 μ 的 σ 有限性知, 存在 X 的可测划分 $\{A_n\}$ 使得对任意 n, $\mu(A_n) \in (0, \infty)$. 令

$$h = \sum_{n=1}^{\infty} \frac{1}{2^n \mu\left(A_n\right)} I_{A_n}$$

则 h 处处严格正, 且 $\mu(h) = 1$. 令 $\tilde{\mu} = h.\mu$, 则 $\tilde{\mu}$ 为有限测度. 由引理 72 知, 可以用 $\tilde{\mu}$ 代替 μ 来证明定理的结论. 因此不妨设 μ 为有限测度.

情况一. 先考虑 φ 也是有限测度的情形.

第一步. 为了将 φ 关于 μ 绝对连续的部分分离出来, 我们考虑

$$\mathcal{H} = \left\{h \in L^1 : h \geqslant 0, \int_A h \,\mathrm{d}\mu \leqslant \varphi(A)\,, \forall A \in \mathcal{F}\right\}.$$

若 $h_1, h_2 \in \mathcal{H}$, 则 $h_1 \vee h_2 \in \mathcal{H}$:

$$\begin{aligned}\int_A h_1 \vee h_2 \,\mathrm{d}\mu &= \int_{A \cap \{h_1 \geqslant h_2\}} h_1 \,\mathrm{d}\mu + \int_{A \cap \{h_1 < h_2\}} h_2 \,\mathrm{d}\mu \\ &\leqslant \varphi\left(A \cap \{h_1 \geqslant h_2\}\right) + \varphi\left(A \cap \{h_1 < h_2\}\right) = \varphi(A)\end{aligned},$$

即 $\mathcal{H}$ 对有限上端运算封闭. 此外若 $h_n \in \mathcal{H}$, $h_n \uparrow h$, 则 $h \in \mathcal{H}$:

$$\int_A h \,\mathrm{d}\mu = \lim_{n \to \infty} \int_A h_n \,\mathrm{d}\mu \leqslant \varphi(A)\,.$$

于是存在 $f_n \in \mathcal{H}$ 使得, 满足 $f_n \uparrow f$, 且

$$\int f \,\mathrm{d}\mu = \sup\left\{\int h \,\mathrm{d}\mu : h \in \mathcal{H}\right\}.$$

我们定义 φ_c 为 f 关于 μ 的不定积分 $g.\mu$, 再定义 $\varphi_s = \varphi - \varphi_c$, 则由 $f \in \mathcal{H}$ 知 φ_s 是有限测度.

第二步. 现在只需证明 $\varphi_s \perp \mu$. 任取一正整数 n, 考虑符号测度

$$\varphi_s - \frac{1}{n}\mu$$

令 $\{D_n, D_n^c\}$ 为其 Hahn 分解, 则任意 $A \in \mathcal{F}$,

$$\frac{1}{n}\mu(A \cap D_n) = \int_A \frac{1}{n} I_{D_n} \,\mathrm{d}\mu \leqslant \varphi_s(A \cap D_n) \leqslant \varphi_s(A)\,,$$

从而得到

$$\int_A g + \frac{1}{n} I_{D_n} \,\mathrm{d}\mu \leqslant \varphi(A)\,.$$

由 g 的选取, 必然有 $\mu(D_n)=0$. 同时, 注意到任意 $A\in\mathcal{F}$, 且 $A\in D_n^c$

$$\varphi_s(A)\leqslant\frac{1}{n}\mu(A)\,.$$

现在令 $N=\cup_{n=1}^{\infty}D_n$, 则 $\mu(N)=0$, 且 $N^c=\cap_{n=1}^{\infty}D_n^c\subset D_n^c$,

$$\varphi_s(N^c)\leqslant\frac{1}{n}\mu(N^c)$$

令 n 趋向于无穷, 得到 $\varphi_s(N^c)=0$.

第三步. 我们证明分解是唯一的. 若 φ_s',φ_c' 是另一 Lebesgue 分解, 则由于 $\varphi_s-\varphi_s'=\varphi_c-\varphi_c'$ 知 $\varphi_s-\varphi_s'$ 即关于 μ 绝对连续, 又与 μ 相互奇异, 故为 0. 唯一性得证.

情况二. 考虑 φ 为 σ 有限测的情形.

第四步. 取 X 的可测划分 $\{A_n\}$ 使得对任意 $n,\varphi(A_n)<\infty$. 定义测度 φ^n 为

$$\varphi^n(A)=\varphi\left(A\cap A_n\right)\,,\ \text{for}\ A\in\mathcal{F}\,.$$

则 φ^n 为有限测度. 在情况一中我们已经证下面分解的存在唯一性:

$$\varphi^n=\varphi_c^n+\varphi_s^n\,,\ n\in\mathbb{N}\,.$$

其中 $\varphi_s^n\perp\mu,\varphi_c^n\ll\mu$, 且存在非负, 关于 μ 可积的可测函数 f_n 使得

$$\varphi_c^n=f_n.\mu\quad.$$

显然 f_n 在 A_n^c 上可取为 0, 令

$$\varphi_c=\sum_{n=1}^{\infty}\varphi_c^n,\ \varphi_s=\sum_{n=1}^{\infty}\varphi_s^n,\ f=\sum_{n=1}^{\infty}f_n\,.$$

则 φ_s 与 φ_c 皆为 σ 有限测度, g 是关于 μ a.e. 有限的非负可测函数, 且有 $\varphi_s\perp\mu$, $\varphi_c\ll\mu$, $\varphi=\varphi_c+\varphi_c$. 特别的, $\varphi_c=f.\mu$.

情况三. 考虑 φ 是 σ 有限的符号测度的情形.

第五步. 分别考虑 φ^+ 以及 φ^-, 不妨设 φ^- 是有限测度. 易知 φ^+ 是 σ 有限测度; φ^+, φ^- 皆关于 μ 绝对连续. 由情况一与情况二, 存在分解

$$\varphi^{\pm}=\varphi_c^{\pm}+\varphi_s^{\pm}\,,$$

其中 φ_c^+,φ_s^+ 是 σ 有限测度, φ_c^-,φ_s^- 是有限测度, 满足

$$\varphi_c^{\pm}\ll\mu,\varphi_s^{\pm}\perp\mu\,,$$

且存在非负 μ-a.e. 有限的可测函数 f^+ 使得 $\varphi_c^+ = f^+.\mu$, 存在非负 μ-可积函数 f^- 使得 $\varphi_c^- = f^-.\mu$, 令

$$\varphi_c = \varphi_c^+ - \varphi_c^- \,,\ \varphi_s = \varphi_s^+ - \varphi_s^- \,,\ f = f^+ - f^- \,.$$

则 φ_s 与 φ_c 皆为 σ 有限测度, f 是关于 μ a.e. 有限的积分存在的可测函数, 满足

$$\varphi = \varphi_c + \varphi_c \,,\ \varphi_c \ll \mu \,,\ \varphi_s \perp \mu \,.$$

特别的, $\varphi_c = f.\mu$. □

经过简单的推到不难证明, Lebesgue 分解可以写成下面更一般的形式.

推论 74. *对可测空间 $(X, \mathcal{F})$ 上的 σ 有限符号测度 φ 与 ψ, 存在 φ 是分解*

$$\varphi = \varphi_s + \varphi_c$$

其中 φ_c, φ_s 为 σ 有限测度, 满足 $\varphi_c \ll \psi$, $\varphi_s \perp \psi$.

例 9. Lebesgue 分解可以用来解释初等概率论中随机变量的分类问题.

设 X 是概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量, 其概率分布记为 $\mathbb{P}_X$. 根据 Lebesgue 分解, $(\mathbb{R}, \mathcal{B}(\mathbb{R})$ 上概率测度 $\mathbb{P}_X$ 对 σ 有限的 Lebesgue 测度 λ 有如下分解:

$$\mathbb{P}_X = \mu_c + \mu_s \ ;\ \mu_c \ll \lambda \,,\ \mu_s \perp \lambda \,.$$

4.2.2 Radon-Nikodym 定理

定义 26. φ 是测度空间 $(X, \mathcal{F}, \mu)$ 上的符号测度, 若存在 μ-a.e. 意义下唯一的, 关于 μ 积分存在可测函数 f, 使得

$$\varphi = f.\mu \ ,$$

则称 f 是 φ 关于 μ 的 **R-N 导数**, 记为 $\dfrac{\mathrm{d}\varphi}{\mathrm{d}\mu} := f$.

正如微积分中并非所有的函数都可以求导一样, 也不是每一个符号测度都有 R-N 导数. 什么样的符号测度才有 R-N 导数呢? 上一小节, 我们已经指出了只有当 φ 对 μ 绝对连续时才有可能. 并且证明了若 μ , φ 皆为 σ 有限测度时, R-N 导数的存在性. 下面的定理表明, φ 是 σ 有限的要求可以去掉. 只要 μ 为 σ 有限测度时, R-N 导数一定存在.

定理 75 (Radon-Nikodym 定理)**.** $(X, \mathcal{F}, \mu)$ 是 σ 有限测度空间. φ 是 $(X, \mathcal{F})$ 上的符号测度 *(不必 σ-有限)*. 若 φ 关于 μ 绝对连续, 则存在 $(X, \mathcal{F}, \mu)$ 上 *a.e.* 唯一的, 积分存在的可测函数 f, 满足对任意 $A \in \mathcal{F}$,

$$\varphi(A) = \int_A f \, \mathrm{d}\mu \, .$$

特别的, 若 φ 也是 σ 有限符号测度, 则 f 是 μ-*a.e.* 有限的.

证明. 考虑 μ 非零的情形. 由 μ 的 σ 有限性知, 存在 X 的可测划分 $\{A_n\}$ 使得对任意 n, $\mu(A_n) \in (0, \infty)$. 令

$$h = \sum_{n=1}^{\infty} \frac{1}{2^n \mu(A_n)} I_{A_n}$$

则 h 处处严格正, 且 $\mu(h) = 1$. 令 $\tilde{\mu} = h.\mu$, 则 $\tilde{\mu}$ 为有限测度. 由引理 72 知, 可以用 $\tilde{\mu}$ 代替 μ 来证明定理的结论. 因此不妨设 μ 为有限测度.

第一步. 当 φ 是 σ 有限的测度时, 由定理 73 的证明知, 满足要求的 f 的存在, 此时 f 关于 μ 积分存在, 且 μ-a.e. 有限. 命题 45 保证了唯一性.

第二步. 当 φ 为一般的测度时, 考虑

$$\mathcal{G} = \{A \in \mathcal{F} : \varphi(A) < \infty\} \, ,$$

可见 $\mathcal{G}$ 对有限并运算封闭. 于是存在 $C_n \in \mathcal{G}$, $C_n \uparrow C$, 使得

$$\mu(C) = \sup\{\mu(G) : G \in \mathcal{G}\} \, .$$

令

$$\varphi'(A) = \varphi(A \cap C), \quad \varphi''(A) = \varphi(A \cap C^c) \, , \ \forall A \in \mathcal{F} \, .$$

则显然 φ' 为 σ 有限测度, 且 $\varphi' \ll \mu$. 于是存在 μ-a.e. 有限的非负可测函数 f', 使得 $\varphi' = f'.\mu$; 另一方面, 由 $\mathcal{G}$ 的定义以及 C 的选取, 以及绝对连续性, 任意 $A \in \mathcal{F}$,

$$\mu(A \cap C^c) > 0 \Rightarrow \varphi(A \cap C^c) = \varphi''(A) = \infty \, ,$$
$$\mu(A \cap C^c) = 0 \Rightarrow \varphi(A \cap C^c) = \varphi''(A) = 0 \, .$$

因此令 $f'' = (\infty) I_{C^c}$, 则 $\varphi'' = f''.\mu$; 从而令 $f = f' + f''$, 则可见

$$\varphi = \varphi' + \varphi'' = f.\mu \quad .$$

第三步. 当 φ 是符号测度时, 不妨设 φ^- 为有限测度. 利用第一, 二步结果立得. □

例 10. μ 是 σ 有限的这个条件是不能去掉的. 我们给出一个反例.

考虑测度空间 $(\mathbb{R}, \mathcal{F}, \mu)$, 其中 $\mathcal{F} = \{A \subset \mathbb{R} : A \text{ 或} A^c \text{ 至多可数}\}$, μ 是计数测度. 我们定义 φ 为

$$\varphi(A) = \begin{cases} 0\,, & A \text{ 至多可数} \\ 1\,, & A^c \text{ 至多可数} \end{cases}\,, \quad \text{for } A \in F\,.$$

可见 φ 是一个测度, 且 $\varphi \ll \mu$. 但不存在 f 使得 $\varphi = f.\mu$. 因为如果存在, 对任意 $x \in \mathbb{R}$, 有

$$0 = \varphi(\{x\}) = \int_{\{x\}} f \mathrm{d}\mu = f(x)\mu(\{x\}) = f(x)\,,$$

从而 $f = 0$. 于是

$$\varphi(\mathbb{R}) = \int_{\mathbb{R}} f \mathrm{d}\mu = \int_{\mathbb{R}} 0 \mathrm{d}\mu = 0$$

与 $\varphi(\mathbb{R}) = 1$ 矛盾!

本节的最后, 我们重新叙述下引理 72.

定理 76. *设 $(X, \mathcal{F}, \mu)$ 是 σ 有限的测度空间. ν 是 $\mathcal{F}$ 上测度, 且 $\nu \ll \mu$. 对任意 $(X, \mathcal{F})$ 上可测函数 f, 只要下式的一端有意义, 另一端也就有意义并且等号成立:*

$$\int_X f \mathrm{d}\nu = \int_X f \frac{d\nu}{d\mu} \mathrm{d}\mu\,.$$

练习

练习 11. $(X, \mathcal{F}, \mu)$ 是 σ 有限的测度空间. 证明导数的运算法则:

(1) 若 φ 与 $\mathcal{F}$ 上符号测度, $\varphi \ll \mu$, 则对任何实数 a

$$\frac{\mathrm{d}(a\varphi)}{\mathrm{d}\mu} = a \frac{\mathrm{d}\varphi}{\mathrm{d}\mu}\,, \quad \mu\text{-a.e.}.$$

(2) 若 φ, ψ 与 $\mathcal{F}$ 上符号测度, $\varphi, \psi \ll \mu$, 则

$$\frac{\mathrm{d}(\varphi + \psi)}{\mathrm{d}\mu} = \frac{\mathrm{d}\varphi}{\mathrm{d}\mu} + \frac{d\psi}{\mathrm{d}\mu}\,, \quad \mu\text{-a.e.}.$$

练习 12. φ 是 $(X, \mathcal{F})$ 上符号测度, μ, ν 是 σ 有限测度且 $\varphi \ll \nu \ll \mu$. 证明

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\mu} = \frac{\mathrm{d}\varphi}{\mathrm{d}\nu} \cdot \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \quad \mu\text{-a.e.}\,.$$

练习 13. $(X,\mathcal{F})$ 上符号测度, μ, ν 是 σ 有限测度且 $\nu \ll \mu$. 证明: $\nu \sim \mu$ 当且仅当

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu} > 0\,, \quad \mu\text{-a.e.},$$

且此时

$$\frac{\mathrm{d}\mu}{\mathrm{d}\nu} = 1\Big/\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\,, \quad \text{a.e.}.$$

4.3 条件期望与条件概率

条件概率是概率论最重要的概念之一. 本节的目的是用 R-N 定理来给出条件期望和条件概率的严格定义并讨论它们的性质. 设 $(\Omega,\mathcal{F},\mathbb{P})$ 是一个概率空间, X 是它上面积分存在的随机变量. 又设 $\mathcal{G}$ 是 $\mathcal{F}$ 的子 σ 域. 将 $\mathbb{P}$ 在 $\mathcal{G}$ 上的限制记为 $\mathbb{P}|_{\mathcal{G}}$. 我们将不定积分也限制在 $\mathcal{G}$ 上,

$$(X.\mathbb{P})|_{\mathcal{G}}$$

是 $\mathcal{G}$ 上的 σ 有限的符号测度, 且关于 $\mathbb{P}|_{\mathcal{G}}$ 绝对连续. 因此由 R-N 定理知, 存在 $(\Omega,\mathcal{G},\mathbb{P})$ 上 a.s. 意义下唯一的可测函数, 现记为 $E(X|\mathcal{G})$, 使得

$$\int_A X\,\mathrm{d}\mathbb{P} = \int_A E(X|\mathcal{G})\,\mathrm{d}\mathbb{P}\,, \text{ for all } A \in \mathcal{G}\,.$$

换言之

$$E(X|\mathcal{G}) \coloneqq \frac{\mathrm{d}(X.\mathbb{P})|_{\mathcal{G}}}{\mathrm{d}\mathbb{P}|_{\mathcal{G}}}\,.$$

遗憾的是, 我们之前虽然证明了 R-N 导数的存在性, 但没有给出具体算法, 也就是说, 想要计算条件期望, 还要另寻他路, 我们将 $E(X|\mathcal{G})$ 的性质抽象出来, 把条件期望和条件概率公理化地定义如下.

定义 27. X 是 $(\Omega,\mathcal{F},\mathbb{P})$ 上积分存在的随机变量, 称 Y 为 X 关于 $\mathcal{F}$ 的子 σ 域 $\mathcal{G}$ 的 **条件期望** (conditional expectation), 若

1. Y 是 $(\Omega,\mathcal{G},\mathbb{P})$ 上积分存在的随机变量.
2. 对任意 $A \in \mathcal{G}$, 有
$$\int_A X\,\mathrm{d}\mathbb{P} = \int_A Y\,\mathrm{d}\mathbb{P}\,.$$

满足上述要求的随机变量 Y 称为 $E(X|\mathcal{G})$ 的一个版本 (version). 此外, 对 $A \in \mathcal{F}$, 我们定义 A 关于 $\mathcal{G}$ 的 **条件概率**为

$$P(A|\mathcal{G}) \coloneqq E(1_A|\mathcal{G})\,.$$

注. 正如本节开头所指出的, R-N 定理保证了上述定义对象的存在性. 由于条件期望和条件概率都是在 a.s. 意义下惟一确定的, 所以关于它们的式子后面通常都要加上 a.s. 的字样. 但是要特别注意的是, 这个 a.s. 不是指相差 $\mathcal{F}$ 中的零测集, 而是指相差 $\mathcal{G}$ 中的零测集.

又注. 另一点有可能用到的是, 若 X 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上取广义实值的随机变量, 其关于 $\mathcal{F}$ 的子 σ 域 $\mathcal{G}$ 的条件期望也可以同上定义, 但是要将 1 中的随机变量改为广义实值随机变量, 其存在唯一性也可由 R-N 保证.

在应用中, 我们常常遇到下面的情况: σ 域 $\mathcal{G}$ 是由某个随机元生成的. 设 η 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 到某个抽象可测空间 $(S, \mathcal{S})$ 的随机元, 我们分别称

$$E(X|\eta) := E(X|\sigma(\eta))\ ;\ P(A|Y) := P(A|\sigma(\eta))$$

为随机变量 X 关于 η 的条件期望和事件 A 关于 η 的条件概率. 注意, 由 $E(X|\eta)$ 关于 $\sigma(\eta)$ 可测和定理 28, 存在 $(S, \mathcal{S})$ 上实值可测函数, 记为 h, 使得

$$E(X|\eta) = h(\eta)\,,$$

而且

$$\int_{\eta\in B} X\,\mathrm{d}\mathbb{P} = \int_{\eta\in\mathcal{B}} h(\eta)\,\mathrm{d}\mathbb{P} = \int_B h(x)\mathbb{P}_\eta(\mathrm{d}x)\,.$$

可见这样的 h 是关于测度 $\mathbb{P}_\eta$ 几乎处处唯一的. 我们给出以下定义.

定义 28. 可测空间 $(S, \mathcal{S})$ 上的可测函数 h 称为 X **关于** η **给定值的条件期望**, 如果对任何 $B \in \mathcal{S}$,

$$\int_{\eta\in B} X\,\mathrm{d}\mathbb{P} = \int_B h(x)\mathbb{P}_\eta(\mathrm{d}x)\,.$$

满足上述要求实可测函数 $h(\cdot)$ 称为 $E(X|\eta = \cdot)$ 的一个版本. 此外, 事件 $A \in \mathcal{F}$ 关于 η 给定值的条件概率自然地定义为

$$P(A \mid \eta = \cdot) := E(1_A|\eta = \cdot)\,.$$

注. 显然, 若 h 是 X 关于 η 给定值的条件期望, 则 $h(\eta)$ 就是 X 关于 η 的条件期望, 在实际的计算中我们利用这样的得到条件期望.

最后, 我们提出一个问题: 条件概率是不是概率 (测度) ? 如果是, 条件期望能不能表示成条件概率的积分? 下一节我们将会考虑这个问题. 现在先让来看一些关于条件期望的例子.

4.3.1 例子

通过符号测度的 R-N 导数来定义条件期望和条件概率显得比较抽象. 但是, 问题的本质不在于抽象本身, 而在于抽象得合理不合理. 为此, 让我们通过一些简单的例子来说明这里定义的条件概率确实是在初等概率论中学过的条件概率的一种合理的抽象.

直观地，我们认为 $\mathcal{G}$ 描述了我们所拥有的信息: 对于 $\mathcal{G}$ 中的每个 A, 我们知道 A 是否已经发生. $E(X|\mathcal{G})$ 是根据我们所拥有的信息对 X 进行的"最佳猜测".

例 11. 若 X 是 $\mathcal{G}$ 可测的, 那么

$$E(X|\mathcal{G}) = X \quad \text{a.s. .}$$

这即是说, 如果我们知道 X 的全部信息, 那么我们对 X 的最佳预测自然应该是 X 本身.

例 12. 另一极端情况是无相关信息时进行预测. 设 X 与 $\mathcal{G}$ 相互独立[1], 则

$$E(X|\mathcal{G}) = \mathbb{E}X \quad \text{a.s. .}$$

这是说, 如果我们知道的信息皆与 X 无关, 那么对 X 最佳的猜测就是他的期望 $\mathbb{E}X$. 更一般的, 若 $\mathcal{G}'$ 与 $\sigma(X,\mathcal{G})$ 独立[2], 则

$$E[\,X|\sigma(\mathcal{G},\mathcal{G}')\,] = E(X|\mathcal{G}) \quad \text{a.s..}$$

例 13. 设 $\{X_n\}$ 为 i.i.d. 随机变量, X_1 可积. 令 $S_n = X_1 + X_2 + \cdots + X_n$, 且令

$$\mathcal{G}_n = \sigma\left(S_n, S_{n+1}, \ldots\right) = \sigma\left(S_n, X_{n+1}, \ldots\right)$$

则

$$\mathbb{E}\left[X_1|\mathcal{G}_n\right] = \mathbb{E}\left[X_1|S_n\right] = \frac{S_n}{n} \quad \text{a.s..}$$

例 14. 在此例中, 我们将条件期望的抽象定义和在初等概率论中学习到的定义联系到了一起. 假设 $\{\Omega_1, \Omega_2, \cdots\}$ 是 Ω 的至多可数的可测划分, 且皆有正概率. 令 $\mathcal{G} = \sigma(\Omega_1, \Omega_2, \cdots)$, 则

$$E(X|\mathcal{G}) = \frac{\mathbb{E}\left(X 1_{\Omega_i}\right)}{\mathbb{P}\left(\Omega_i\right)} = \mathbb{E}(X \mid \Omega_i) \quad \text{on } \Omega_i\,.$$

[1] 独立的定义见高等概率论笔记

[2] 注意这与 $\mathcal{G}'$ 既与 X 独立, 又与 $\mathcal{G}$ 独立并不等价: 考虑两两独立但不相互独立的集合.

且

$$P(A|\mathcal{G}) = \frac{\mathbb{P}\left(A \cap \Omega_i\right)}{\mathbb{P}\left(\Omega_i\right)} = \mathbb{P}(A \mid \Omega_i) \quad \text{ on } \Omega_i\,.$$

接下来的两个例子, 需要用到积分交换顺序, 其正确性由第五章中的 Fubini 定理保证.

例 15. 初等概率论中, 关于某个连续性的随机变量的条件期望是如此定义的: 假设 (X, Y) 有联合密度 f, 即

$$\mathbb{P}((X,Y) \in B) = \int_B f(x,y)\,\mathrm{d}x\,\mathrm{d}y \quad \text{ for any } B \in \mathcal{B}(\mathbb{R}^2)\,.$$

为简单起见, 这里我们假设对任意 $y \in \mathbb{R}$, $\int f(x,y)\,\mathrm{d}x > 0$. 我们当时定义了条件所谓的条件密度: 对任意 y, 定义

$$f_{X|Y}(x|y) \coloneqq \frac{f(x,y)}{\int f(x,y)\,\mathrm{d}x}\,, \text{ for all } x \in \mathbb{R}\,.$$

直观上说,

$$\frac{f(x,y)}{\int f(x,y)\,\mathrm{d}x} = \frac{\mathbb{P}(X=x, Y=y)}{\mathbb{P}(X=x \mid Y=y)} = \mathbb{P}(X = x \mid Y = y)\,.$$

对任意 (Borel 可测) 函数 g, 若 $\mathbb{E}|g(X)| < \infty$, 对任意 $y \in \mathbb{R}$, 定义了 "$g(X)$ 关于 $Y = y$ 的条件期望"

$$E[g(X)|Y=y] \equiv h(y) \coloneqq \int g(x) f_{X|Y}(x|y)\,\mathrm{d}x = \frac{\int g(x) f(x,y)\,\mathrm{d}y}{\int f(x,y)\,\mathrm{d}x}\,.$$

则 $g(X)$ 关于 Y 的条件期望定义为

$$E[g(X)|Y] \coloneqq h(Y)\,.$$

现在我们来说明上面初等概率论中的定义确实本节讲的一般定义: 首先 h 是一个实值可测函数 (证明蕴含在第五章 Fubini 定理的证明中), 其次对任意 $B \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\int_B h(y)\mathbb{P}_Y(\mathrm{d}y) &= \int_B h(y) \int_{\mathbb{R}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_B \int_{\mathbb{R}} g(x) f(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathbb{R}^2} g(x) 1_{\{y\in B\}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{Y\in B} g(X)\,\mathrm{d}\mathbb{P}\,.\end{aligned}$$

第三个等号用到了第五章中的 Funibi 定理.

注. 若 $\int f(x,y)\,\mathrm{d}x > 0$ 并不对所有 $y \in \mathbb{R}$ 成立, 我们修改 $f_{X|Y}(x|y)$ 的定义为

$$f_{X|Y}(x|y) \int f(x,y)\,\mathrm{d}x = f(x,y) \ \text{ a.e.}$$

换言之, 对 $\int f(x,y)\,\mathrm{d}x = 0$ 的 y, 我们可以令 $f_{X|Y}(x|y)$ 为任意实函数. 容易看出, 这样修改后, 上面的证明仍然成立.

例 16. 这是一个很符合直觉的例子. 设 X, Y 相互独立, 令 ϕ 为 $\mathbb{R}^2$ 上 Borel 可测函数, 且 $\phi(X,Y)$ 可积. 那么,

$$E[\phi(X,Y) \mid X = x] \equiv h(x) = \begin{cases} \mathbb{E}\,\phi(x,Y), \ \text{if } \mathbb{E}\,|\varphi(x,Y)| < \infty\,. \\ 0\ , \ \text{else}\ \ . \end{cases}$$

且 $E[\phi(X,Y) \mid X] = h(X)$.

下面来验证, 首先由 $\phi(X,Y)$ 可积,

$$\int_{\mathbb{R}^2} |\phi(x,y)|\,\mathbb{P}_X(\mathrm{d}x)\mathbb{P}_Y(\mathrm{d}y) < \infty\,.$$

于是

$$\mathbb{E}|\phi(x,Y)| = \int_{\mathbb{R}} |\phi(x,y)|\mathbb{P}_Y(\mathrm{d}y) < \infty\,, \quad \mathbb{P}_X\text{-a.e.}$$

可见 h 为 Borel 可测函数. 故 $h(X)$ 为 $\sigma(X)$ 可测的. 再注意到对 $B \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\int_{X\in B} h(X)\,\mathrm{d}\mathbb{P} &= \int_B h(x)\mathbb{P}_X(\mathrm{d}x) = \int_B \int_{\mathbb{R}} \phi(x,y)\mathbb{P}_Y(\mathrm{d}y)\mathbb{P}_X(\mathrm{d}x) \\ &= \int_{\mathbb{R}^2} \phi(x,y)1_{\{x\in B\}}\mathbb{P}_X(\mathrm{d}x)\mathbb{P}_Y(\mathrm{d}y) = \int_{X\in B} \phi(X,Y)\,\mathrm{d}\mathbb{P}\,.\end{aligned}$$

可见 $h(X)$ 确为条件期望.

4.3.2 条件期望的性质

条件期望有一些与普通期望相同的性质, 比如线性, 单调性. 证明只需机械的验证定义, 故略去了.

命题 77. *设 X, Y 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上积分存在的随机变量. $\mathcal{G}$ 是 $\mathbb{F}$ 子 σ 域.*

1. *线性: 对任意实数 a, b, 若 $a\mathbb{E}X + b\mathbb{E}Y$ 有意义, 则*

$$E(aX + Y|\mathcal{G}) = aE(X|\mathcal{G}) + E(Y|\mathcal{G}) \ \ a.s.$$

2. *单调性: 若 $X \leqslant Y$ a.s. , 则*

$$E(X|\mathcal{G}) \leq E(Y|\mathcal{G}) \ \ a.s.$$

对于取期望, 我们有三个常用的极限定理, 现在依然成立.

命题 78. $\{X_n\}$ 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上积分存在的随机变量. $\mathcal{G}$ 是 $\mathcal{F}$ 子 σ 域.

1. *(单调收敛定理)* 若 $X_n \uparrow X$, 且存在可积随机变量 Y, 对任意正整数 n, $X_n \geqslant Y$ *a.s.* , 则

$$\lim_{n\to\infty} E\left(X_n|\mathcal{G}\right) = E(X|\mathcal{G}) \quad a.s.\ ,$$

2. *(Fatou)* 若存在可积随机变量 Y, 对任意正整数 n, $X_n \geqslant Y$ *a.s.* 则

$$E\left(\liminf_{n\to\infty} X_n \mid \mathcal{G}\right) \leqslant \liminf_{n\to\infty} E(X_n|\mathcal{G}) \quad a.s.\ .$$

3. *(Lebesgue)* 若存在可积随机变量 Y, 对任意正整数 n, $|X_n| \leqslant Y$ *a.s.* , 且 $X_n \to X$ *a.s.* 则

$$\lim_{n\to\infty} E\left(|X_n - X| \mid \mathcal{G}\right) = 0 \quad a.s.\ .$$

注. 要注意依测度形式的 Fatou 引理不一定成立了.

证明. 1. 显然 $\{E\left(X_n|\mathcal{G}\right)\}$ 几乎必然单调递增, 且 $\lim\limits_{n\to\infty} E\left(X_n|\mathcal{G}\right)$ 关于 $\mathcal{G}$ 可测. 由 Levi 单调收敛定理, 对任意 $A \in \mathcal{G}$,

$$\begin{aligned}\int_A \lim_{n\to\infty} E\left(X_n|\mathcal{G}\right) \mathrm{d}\mathbb{P} &= \lim_{n\to\infty} \int_A E\left(X_n|\mathcal{G}\right) \mathrm{d}\mathbb{P} = \lim_{n\to\infty} \int_A X_n \,\mathrm{d}\mathbb{P} \\ &= \int_A \lim_{n\to\infty} X \,\mathrm{d}\mathbb{P}.\end{aligned}$$

2. 对任意正整数 n, 令 $Y_n = \inf_{k\geqslant n} X_k$, 则 $Y_n \geqslant Y$ 且 $Y_n \uparrow \liminf_n X_n$, 由单调收敛定理,

$$E\left(\liminf_{n\to\infty} X_n \mid \mathcal{G}\right) = \lim_{n\to\infty} E\left(Y_n|\mathcal{G}\right) \leqslant \liminf_{n\to\infty} E\left(X_n|\mathcal{G}\right) .$$

3. 考虑非负随机变量序列 $\{2Y - |X_n - X|\}$, 显然 $2Y - |X_n - X| \to 2Y$ a.s. , 由 2 得到

$$E(2Y|\mathcal{G}) \leqslant E(2Y|\mathcal{G}) - \limsup_{n\to\infty} E\big(|X_n - X|\big|\mathcal{G}\big) \quad \text{a.s.}\ .$$

于是得证. □

例 17. 设 $\{X_n\}$ 是一致可积的随机变量序列, 且几乎必然收敛于 X. 我们之前已经证明了在 L^1 范数下 $X_n \to X$, 特别的

$$\mathbb{E}X_n \to \mathbb{E}X\,.$$

但是现在不成立

$$E\left(X_n|\mathcal{G}\right) \to E\left(X|\mathcal{G}\right) \quad \text{a.s.}\,.$$

不过很容易证明在 L^1 范数下, $E\left(X_n|\mathcal{G}\right) \to E(X|\mathcal{G})$.

令 $Y_1, Y_2, \ldots$ 与 $Z_1, Z_2, \ldots$ 为相互独立的随机变量, 且

$$\mathbb{P}\left(Y_n = 1\right) = 1/n\,, \quad \mathbb{P}\left(Y_n = 0\right) = 1 - 1/n\,,$$
$$\mathbb{P}\left(Z_n = n\right) = 1/n\,, \quad \mathbb{P}\left(Z_n = 0\right) = 1 - 1/n\,.$$

令 $X_n = Y_n Z_n$. 可见 $\{X_n\}$ 相互独立, 且 $\mathbb{P}\left(X_n > 0\right) = 1/n^2$, 由 B-C 引理, $X_n \to 0$ a.s. 同时 $\{X_n\}$ 一致可积, 因为

$$\mathbb{E}\left(X_n 1_{|X_n|\geq 1}\right) = n/n^2,$$

令 $\mathcal{G} = \sigma\left(Y_1, Y_2, \ldots\right)$, 则

$$E\left(X_n|\mathcal{G}\right) = Y_n E\left(Z_n|\mathcal{G}\right) = Y_n \mathbb{E}Z_n = Y_n\,.$$

但是 $Y_n \to 0$ in L^1 而不 a.s., 我们给出了反例.

我们知道, 求期望下有 Jensen 不等式: 若 φ 为 $\mathbb{R}$ 上的下凸函数, 且随机变量 X, $\varphi(X)$ 皆可积, 则有

$$\varphi(\mathbb{E}X) \leqslant \mathbb{E}\,\varphi(X)\,.$$

对于条件期望, 仍然成立此不等式.

命题 79. *X 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上随机变量. $\mathcal{G}$ 是 $\mathcal{F}$ 子 σ 域. φ 为 $\mathbb{R}$ 上的下凸函数. 若 X, $\varphi(X)$ 皆可积, 则有*

$$\varphi(E[X|\mathcal{G}]) \leqslant E[\varphi(X)|\mathcal{G}] \quad a.s.\,,$$

证明. 若 φ 线性, 结论是平凡的. 我们假设 φ 非线性. 令

$$S = \{(a,b) : a, b \in \mathbf{Q}, ax + b \leqslant \varphi(x) \text{ for all } x \in \mathbb{R}\},$$

由下凸函数的性质, 则有 $\varphi(x) = \sup\{ax + b : (a,b) \in S\}$. 若 $(a,b) \in S$, 则

$$aE(X|\mathcal{G}) + b \leqslant E(\varphi(X)|\mathcal{G}) \quad \text{a.s.}\,,$$

遍历 $(a,b) \in S$ 取上确界得到

$$\varphi(E(X|\mathcal{G})) \leqslant E(\varphi(X)|\mathcal{G}) \quad \text{a.s.}\,. \qquad \square$$

注. 这里我们通过不等式写出 a.s. 来强调对于每个 (a,b) 都有一个例外集, 所以我们要使得 (a,b) 属于某个可数集.

推论 80. *X 是 $(\Omega,\mathcal{F},\mathbb{P})$ 上 p 幂可积的随机变量. $\mathcal{G}$ 是 $\mathcal{F}$ 子 σ 域. 则*

$$|E(X|\mathcal{G})|^p \leqslant E\left(|X|^p|\mathcal{G}\right) \quad a.s. \quad , \ \mathbb{E}[|E(X|\mathcal{G})|^p] \leqslant \mathbb{E}|X|^p \, .$$

练习 14. X, Y 是 $(\Omega,\mathcal{F},\mathbb{P})$ 上随机变量. $\mathcal{G}$ 是 $\mathcal{F}$ 子 σ 域, 证明下面两个不等式.

i) (Hölder 不等式) 若 $p,q>1$ 满足 $\frac{1}{p}+\frac{1}{q}=1$, 则

$$E(|XY||\mathcal{G}) \leqslant [E\left(|X|^p|\mathcal{G}\right)]^{1/p}[E\left(|Y|^q|\mathcal{G}\right)]^{1/q} \quad \text{a.s.} \, .$$

ii) (Minkovski 不等式) 若 $p \geqslant 1$, 则

$$[E\left(|X+Y|^p|\mathcal{G}\right)]^{1/p} \leqslant [E\left(|X|^p|\mathcal{G}\right)]^{1/p} + [E\left(|Y|^p|\mathcal{G}\right)]^{1/p} \quad \text{a.s.} \, .$$

条件期望还以有一些特别的性质.

命题 81 (塔性质)**.** *X 是 $(\Omega,\mathcal{F},\mathbb{P})$ 上积分存在的随机变量. $\mathcal{G}_1,\mathcal{G}_1$ 是 $\mathcal{F}$ 子 σ 域, 且 $\mathcal{G}_1 \subset \mathcal{G}_2$, 则*

$$E[E(X|\mathcal{G}_1)|\mathcal{G}_2] = E[E(X|\mathcal{G}_2)|\mathcal{G}_1] = E(X|\mathcal{G}_1) \quad a.s.$$

证明. 由于 $E(X|\mathcal{G}_1)$ 关于 $\mathcal{G}_1$ 可测, 必关于 $\mathcal{G}_2$ 可测, 从而有

$$E[E(X|\mathcal{G}_1)|\mathcal{G}_2] = E(X|\mathcal{G}_1) \quad \text{a.s.} \, .$$

为证明另一个等式, 注意到对任意 $A \in \mathcal{G}_1$,

$$\int_A E[E(X|\mathcal{G}_2)|\mathcal{G}_1] \, \mathrm{d}\mathbb{P} = \int_A E(X|\mathcal{G}_2) \, \mathrm{d}\mathbb{P} = \int_A X \, \mathrm{d}\mathbb{P} \, .$$

由定义即得. □

下面的结果说明, 当 $\mathcal{G}$-可测的随机变量求关于 $\mathcal{G}$ 的条件期望时, 就像常数一样: 他们可以直接提到积分号外面.

定理 82. *X,Y 是 $(\Omega,\mathcal{F},\mathbb{P})$ 上随机变量, 且 X, XY 的积分存在. $\mathcal{G}$ 是 $\mathcal{F}$ 子 σ 域, 那么*

$$E(XY|\mathcal{G}) = XE(Y|\mathcal{G}) \quad a.s. \, .$$

证明. 首先 $XE(Y|\mathcal{G})$ 关于 $\mathcal{G}$ 可测, 只需证明对任意 $A \in \mathcal{G}$,

$$\int_A XE(Y|\mathcal{G})\,\mathrm{d}\mathbb{P} = \int_A XY\,\mathrm{d}\mathbb{P}. \tag{4.2}$$

用典型方法来证即可. 若 X 为示性函数 1_B, 其中 $B \in \mathcal{G}$, 显然 (4.2) 成立. 由积分的线性知, X 为非负简单 $\mathcal{G}$-可测函数时 (4.2) 成立.

当 X 为非负 $\mathcal{G}$-可测函数时, 先假设 Y 亦非负, 由单调收敛定理知 (4.2) 成立; 对一般的 Y, 将其分解为正部与负部, (4.2) 成立. 当 X 为一般的 $\mathcal{G}$-可测函数时, 将 X 分解成其正负部即可. □

我们用一个有意义的推论来结束本小节, 它给出了一个关于条件期望的几何解释. 同时阐明了条件期望确实是 (某种意义下)"最佳" 的预测.

推论 83. *$X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{G}$ 是 $\mathcal{F}$ 的子 σ 域. 则 $E(X|\mathcal{G})$ 是 X 在闭子空间 $L^2(\Omega, \mathcal{F}, \mathbb{P})$ 上的投影. 换言之,*

$$\|X - E[X|\mathcal{G}]\|_2 = \mathrm{dist}(X, L^2(\Omega, \mathcal{G}, \mathbb{P})) = \inf\left\{\|Y - X\|_2 : Y \in L^2(\Omega, \mathcal{G}, P)\right\}$$

证明. 注意定理 82 蕴含了对任意 $Y \in L^2(\Omega, \mathcal{G}, \mathbb{P})$,

$$\langle X - E[X|\mathcal{G}], Y\rangle = \int (X - E[X|\mathcal{G}])Y\,\mathrm{d}\mathbb{P} = 0\,.$$

由泛函分析的知识即得. □

4.4 正则条件分布

4.4.1 动机

条件概率是不是概率测度? 如果是, 条件期望能不能表示成条件概率的积分? 我们利用上一节讲的条件期望的性质来讨论这个问题. 设 $(\Omega, \mathcal{F}, \mathbb{P})$ 为一概率空间, $\mathcal{G}$ 为 $\mathcal{F}$ 的子 σ 域. 由条件期望的性质有

- $P(\varnothing|\mathcal{G}) = 0$ a.s. , $P(\Omega|\mathcal{G}) = 1$ a.s. .
- 对任意 $A \in \mathcal{F}$, $P(A|\mathcal{G}) \geqslant 0$ a.s. .
- 对任意互不相交的 $\mathcal{F}$- 可测集列 $\{A_n\}$,

$$P\big(\cup_{n=1}^{\infty} A_n \,\big|\, \mathcal{G}\big) = \sum_{n=1}^{\infty} P\left(A_n|\mathcal{G}\right) \text{ a.s. .}$$

这些性质与概率测度的性质已经非常相似了, 可惜的地方在于有个 a.s. 的尾巴, 也就是出现了例外集. 我们能否通过选择条件概率的合适的版本, 去掉这些例外集? 这个问题就是说, 我们能否做到: 对任意 $A \in \mathcal{F}$, 选定 $P(A|\mathcal{G})$ 的一个版本 (它是一随机变量), 记为 $P(\cdot, A)$, 进而得到一个定义在 $\Omega \times \mathcal{F}$ 上的二元函数 $P(\cdot,\cdot)$. 使得对每个 $\omega \in \Omega$, $P(\omega,\cdot)$ 是 $(\Omega,\mathcal{F})$ 上的概率测度?

一般来说, 回答这个问题十分困难. 对于能做出肯定回答的, 我们抽象出了如下的定义.

定义 29. $(\Omega,\mathcal{F},\mathbb{P})$ 为概率空间, $\mathcal{G}$ 是 $\mathcal{F}$ 的子 σ 域. 二元函数 $P:\Omega\times\mathcal{F}\to[0,1]$ 称为 $\mathbb{P}$ **关于** $\mathcal{G}$ **的正则条件概率**, 如果

1. 对任意 $\omega \in \Omega$, $P(\omega,\cdot)$ 是 $(\Omega,\mathcal{F})$ 上的概率测度.
2. 对任意 $A \in \mathcal{F}$, $P(\cdot, A)$ 是条件概率 $P(A|\mathcal{G})$ 的一个版本.

正则条件概率的第一个应用是: 条件期望算子成了关于正则条件概率的积分. 证明是非常典型的典型方法.

定理 84. *X 是概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 上的积分存在的随机变量, $\mathcal{G}$ 是 $\mathcal{F}$ 的子 σ 域, P 是 $\mathbb{P}$ 关于 $\mathcal{G}$ 的正则条件概率. 则对几乎处处 ω, X 关于测度 $P(\omega,\cdot)$ 的可积, 且*

$$E[X|\mathcal{G}](\omega)=\int_\Omega X\left(\omega'\right)P\left(\omega,\mathrm{d}\omega'\right) \qquad a.s.\ \omega$$

证明. 用典型方法, 当 X 为 $\mathcal{F}$ 可测集的示性函数时, 由正则条件概率的定义知定理成立. 由积分的线性知, X 为非负简单可测函数时, 定理成立. 由单调收敛定理知, X 为非负随机变量时, 定理成立.

当 X 为积分存在的随机变量时, 不妨设 X^- 可积. 由上一步的结论,

$$E[X^+|\mathcal{G}](\omega)=\int_\Omega X^+\left(\omega'\right)P\left(\omega,\mathrm{d}\omega'\right) \qquad \text{a.s. } \omega$$

$$E[X^-|\mathcal{G}](\omega)=\int_\Omega X^-\left(\omega'\right)P\left(\omega,\mathrm{d}\omega'\right) \qquad \text{a.s. } \omega$$

可见对几乎处处的 ω, X^+, X^- 关于 $P\left(\omega,\cdot\right)$ 可积, 即 X 关于 $P\left(\omega,\cdot\right)$ 可积. 再将上面两式做差, 定理得证. □

直接讨论正则条件概率的存在与否虽然是一个很困难的问题, 我们可以考虑一个简单一些的情况, 不要求 $P(\omega,\cdot)$ 是 $\mathcal{F}$ 上的概率测度了, 改成要求其为一个小一点的 σ 域上的概率测度, 比如由某个 $(\Omega,\mathcal{F})$ 上取值于 $(E,\mathcal{E})$ 的随机元 ξ 生成的 σ 域. 假设 $P:\Omega\times\sigma(\xi)\to[0,1]$, 满足

- 对任意 $\omega \in \Omega$, $P(\omega, \cdot)$ 是 $(\Omega, \sigma(\xi))$ 上的概率测度.
- 对任意 $A \in \sigma(\xi)$, $P(\cdot, A)$ 是条件概率 $P(A|\mathcal{G})$ 的一个版本.

通过 P, 我么可以诱导出 $\Omega \times \mathcal{E}$ 上的一个二元函数:

$$\mu(\omega, B) = P\left(\omega, \xi^{-1}(B)\right) \text{ for } \omega \in \Omega,\, B \in \mathcal{E}\,.$$

这时, 显然对任何 ω, $\mu(\omega, \cdot)$ 就是 $\mathcal{E}$ 上的概率测度了. 我们进而考虑证明对形如 $f(\xi)$ 的积分存在的随机变量 (f 为 $(E, \mathcal{E})$ 上的 Borel 实可测函数), 是否有

$$E[f(\xi)|\mathcal{G}](\omega) = \int_E f(x)\mu\left(\omega, \mathrm{d}x\right) \qquad \text{a.s. } \omega \tag{4.3}$$

在许多情况下, 正则条件概率并不存在, 但是满足 (4.3) 的二元函数 μ 却是存在的. 于是我们引入了正则条件分布.

4.4.2 ξ 关于 $\mathcal{G}$ 的正则条件分布

定义 30. $(\Omega, \mathcal{F}, \mathbb{P})$ 是一概率空间, G 是 $\mathcal{F}$ 的子 σ 域, ξ 是 $(\Omega, \mathcal{F})$ 上取值于另一可测空间 $(E, \mathcal{E})$ 的随机元. 称 $\mu : \Omega \times \mathcal{E} \to [0, 1]$ 为 **ξ 关于 $\mathcal{G}$ 的正则条件分布**, 若

1. 对任意 $\omega \in \Omega$, $\mu(\omega, \cdot)$ 是 $(E, \mathcal{E})$ 上的概率测度.
2. 对任意 $B \in \mathcal{E}$, $\mu(\cdot, B)$ 是 $P(\xi \in B|\mathcal{G})$ 的一个版本.

注. 若 $(E, \mathcal{E}) = (\Omega, \mathcal{F})$, 且 ξ 为 Ω 上的恒等映射, 则 ξ 关于 $\mathcal{G}$ 的正则条件分布正是 $\mathbb{P}$ 关于 $\mathcal{G}$ 的正则条件概率. 因此, 正则条件概率存在性的研究可以归结为正则条件分布存在性的研究, 当然, 存在性的难度与 $(E, \mathcal{E})$ 的复杂程度有关系, 我们可以挑一些简单的情况来研究.

例 18. (接例 15). 我们定义了所谓条件密度 $f_{X|Y}(x|y)$, 可以发现其确实为一概率密度. 我们令 $\nu(y, \cdot)$ 来表示其对应的测度, 即

$$\nu(y, B) = \int_B f_{X|Y}(x|y)\,\mathrm{d}x\,, \quad \text{for } B \in \mathcal{B}(\mathbb{R})\,.$$

则 ν 是 $\mathbb{R} \times \mathcal{B}(\mathbb{R}) \to [0, 1]$ 的映射. 我们定义 $\mu : \Omega \times \mathcal{B}(\mathbb{R}) \to [0, 1]$ 为

$$\mu(\omega, B) = \nu(Y(\omega), B)\,, \text{ for } \omega \in \Omega,\, B \in \mathcal{B}(\mathbb{R})\,.$$

则 μ 是 X 关于 $\sigma(Y)$ 的正则条件分布.

下面来验证这一点. 显然对任意 ω, $\mu(\omega,\cdot)$ 是 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 上概率测度. 对给定的 $B\in\mathcal{B}(\mathbb{R})$, 易见 $\nu(\cdot,B)$ Borel 可测, 于是 $\mu(\cdot,B)$ 关于 $\sigma(Y)$ 可测. 又对任意 $A\in\mathcal{B}(\mathbb{R})$,

$$\begin{aligned}\int_{Y\in A}\mu(\cdot,B)\,\mathrm{d}\mathbb{P}=\int_{Y\in A}\nu(Y,B)\,\mathrm{d}\mathbb{P}=\int_A\nu(y,B)\int_{\mathbb{R}}f(x,y)\,\mathrm{d}x\,\mathrm{d}y\\=\int_A\int_B f(x,y)\,\mathrm{d}x\,\mathrm{d}y=\mathbb{P}(X\in B,Y\in A)\,.\end{aligned}$$

可见 $\mu(\cdot,B)$ 是 $P[X\in B|\sigma(Y)]$ 的一个版本, 验证完成.

之前在引入正则条件分布时, 我们说过这么做的目的是, 把条件期望表示为正则条件分布的积分. 下面的定理也用典型方法即可证明, 上面已经展示过一次了, 本次从略.

定理 85. *设 ξ 是 $(\Omega,\mathcal{F},\mathbb{P})$ 上取值于 $(E,\mathcal{E})$ 的随机元. $\mathcal{G}$ 是 $\mathcal{F}$ 的子 σ 域. μ 是 ξ 关于 $\mathcal{G}$ 的正则条件分布. 对任意 $(E,\mathcal{E})$ 上的 Borel 实值可测函数 f, 若 $f(\xi)$ 的积分存在, 则对几乎处处 ω, f 关于测度 $\mu(\omega,\cdot)$ 的可积, 且*

$$E[f(\xi)|\mathcal{G}](\omega)=\int_E f(x)\mu\left(\omega,\mathrm{d}x\right)\qquad a.s.\ \omega\,.$$

例 19. (接例 15,18) 在例 18 中我们已经求出了 X 关于 $\sigma(Y)$ 的正则条件分布, 而例 15 中我们是在计算 $g(X)$ 关于 Y 的条件期望. 现在我们应用定理 85 来再次验证这一点, 这也体现了初等概率论中计算方法的来源.

由定理 85,

$$\begin{aligned}E[g(X)|Y](\omega)=\int_{\mathbb{R}}g(x)\nu\left(\omega,\mathrm{d}x\right)=\int_{\mathbb{R}}g(x)\mu\left(Y(\omega),\mathrm{d}x\right)\\=\int_{\mathbb{R}}g(x)f_{X|Y}(x|Y(\omega))\,\mathrm{d}x\,.\end{aligned}$$

这再次验证了例 15 中的结果.

下面我们来讨论正则条件分布的存在性. 可见我们本质上是在考虑 $\{P(\xi\in B|\mathcal{G}):B\in\mathcal{E}\}$, 由条件期望的性质

- $P(\varnothing|\mathcal{G})=0$ a.s. , $P(\Omega|\mathcal{G})=1$ a.s. .
- 对任意 $B\in\mathcal{E}$, $P(\xi\in B|\mathcal{G})\geqslant 0$ a.s. .
- 对任意互不相交的 $\mathcal{E}$- 可测集列 $\{B_n\}$,

$$P\big(\xi\in\cup_{n=1}^{\infty}B_n\,\big|\,\mathcal{G}\big)=\sum_{n=1}^{\infty}P\left(B_n|\mathcal{G}\right)\ \text{a.s.}\ .$$

现在问题变成了, 我们能否做到: 对任意 $B \in \mathcal{E}$, 选定 $P(\xi \in B|\mathcal{G})$ 的一个版本 (它是一随机变量), 记为 $\mu(\cdot, A)$, 进而得到一个定义在 $\Omega \times \mathcal{E}$ 上的二元函数 $\mu(\cdot,\cdot)$. 使得对每个 $\omega \in \Omega$, $\mu(\omega,\cdot)$ 是 $(E,\mathcal{E})$ 上的测度?

上面列出的性质已经很接近了, 但可惜每个集合 B, 都有一个 a.s. 的尾巴, 也就是有例外的 ω. 可以看到, 这个问题的困难与否在于 $\mathcal{E}$ 中集合的多少. 一个极端的例子是, 假设 E 有一个真子集 B, $\mathcal{E} = \{\varnothing, B, B^c, E\}$. 这样我们选 $\mu(\cdot,\varnothing) \equiv 0$, $\mu(\cdot, E) \equiv 1$, $\mu(\cdot, B)$ 和 $\mu(\cdot, B^c)$ 分别为 $0 \leqslant P(\xi \in B|\mathcal{G}) \leqslant 1$ 的某个版本和 $1 - P(\xi \in B|\mathcal{G})$ 即可.

我们给出一个关于正则条件分布存在性的基本但十分重要的结果.

定理 86. $(\Omega, \mathcal{F}, \mathbb{P})$ 是一概率空间, G 是 $\mathcal{F}$ 的子 σ 域, X 是 $(\Omega,\mathcal{F})$ 上随机变量. 则 X 关于 $\mathcal{G}$ 的正则条件分布存在.

思路. 要找正则条件分布, 就是挑选 $P(X \in B|\mathcal{G})$, $B \in \mathcal{B}(\mathbb{R})$ 的合适版本. 注意 $\mathcal{B}(\mathbb{R})$ 中的集合太多了, 我们可以先在 π 系 $\mathcal{P}_{\mathbb{R}}$ 中来挑, 使得其有概率分布函数的性质, 即先找出 "正则条件分布函数", 再模仿 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上的 L-S 测度构造的过程, 用测度延拓的方式找出正则条件分布.

证明. 对任意有理数 r, 任选 $P(X \leqslant r|\mathcal{G})$ 的一个版本, 记为 $G(\cdot, r)$. 令

$$
\begin{aligned}
N_1 &= \bigcup_{\substack{r_1, r_2 \in \mathbb{Q} \\ r_1 \leqslant r_2}} \{\omega : G(\omega, r_1) > G(\omega, r_2)\} , \\
N_2 &= \bigcup_{r \in \mathbb{Q}} \left\{\omega : \lim_{n\to\infty} G(x, r + \frac{1}{n}) \neq G(x, r)\right\} , \\
N_3 &= \left\{\omega : \lim_{n\to\infty} G(x, n) \neq 1\right\} \cup \left\{\omega : \lim_{n\to\infty} G(x, -n) \neq 0\right\} .
\end{aligned}
$$

由条件概率的性质, $N = N_1 \cup N_2 \cup N_3 \in \mathcal{G}$ 且为零测集. 选定一个分布函数 H, 定义

$$
F(\omega, x) = \begin{cases} \inf\{G(\omega, r) : r \in \mathbb{Q}, r > x\}, & \omega \notin N. \\ H(x)\,, \ \omega \in N. \end{cases}
$$

则对任意 ω, $F(\omega,\cdot)$ 是概率分布函数, 设 $\mu(\omega,\cdot)$ 其对应的 L-S 测度. 按照下面的步骤既可完成证明.

1. 证明, 对任意 $x \in \mathbb{R}$, $F(\cdot, x)$ 是 $P(X \leqslant x|\mathcal{G})$ 的一个版本.

2. 用 $\pi - \lambda$ 定理证明, 对任意 $B \in \mathcal{B}(\mathbb{R})$, $\mu(\omega, B)$ 是 $P(X \in B|\mathcal{G})$ 的一个版本. □

注. 本定理有很简单的一个推广, 若 $(E,\mathcal{E})$ 与 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 可测同构 (如 Polish 空间), 则上面定理对取值于 $(E,\mathcal{E})$ 中的随机元 ξ 也成立. 对于取值于更一般的可测空间 $(E,\mathcal{E})$, 正则条件分布存在性的结果, 可参考严加安先生测度论讲义一书.

4.4.3 ξ 关于 η 的给定值的正则条件分布

对上文中 ξ 关于 σ 域 $\mathcal{G}$ 的正则条件分布, 我们也可以考虑子 $\mathcal{G}$ 是由 $(\Omega,\mathcal{F},\mathbb{P})$ 上某取值与 $(S,\mathcal{S})$ 的随机元 η 生成的. 我们也可考虑 ξ 关于 η 的给定值的正则条件分布. 实际上在我们经常会遇到如下的情况: 对随机变量 X,Y 考虑在条件 $\{Y=y\}$ 下 X 的分布, 这就是指 X 关于 Y 给定值的正则条件分布.

定义 31. $(\Omega,\mathcal{F},\mathbb{P})$ 是一概率空间, ξ,η 是 $(\Omega,\mathcal{F})$ 上取值于 $(E,\mathcal{E})$ 和 $(S,\mathcal{S})$ 的随机元. 称 $\nu: S\times\mathcal{E}\to[0,1]$ 为 ξ **关于** η **给定值的正则条件分布**, 若

1. 对任意 $y\in S$, $\nu(y,\cdot)$ 是 $(E,\mathcal{E})$ 上的概率测度.
2. 对任意 $B\in\mathcal{E}$, $\nu(\cdot,B)$ 是 $P(\xi\in B\mid\eta=\cdot)$ 的一个版本.

注. 注意对给定的 $B\in\mathcal{E}$, $\nu(\cdot,B)$ 是 $P(\xi\in B\mid\eta=\cdot)$ 的一个版本等价于 $\nu(\eta,B)$ 是 $P(\xi\in B\mid\eta)$ 的一个版本. 我们定义 $\Omega\times\mathcal{E}$ 上的二元函数 μ 为

$$\mu(\omega,B):=\nu(\eta(\omega),B)\ \text{ for } \omega\in\Omega,\ B\in\mathcal{E}\,.$$

可见, ν 是 ξ 关于 η 给定值的正则条件分布 当且仅当 μ 是 ξ 关于 $\sigma(\eta)$ 的正则条件分布.

例 20. (接例 15,18,19). 对任意 $y\in\mathbb{R}$, 我们已经定义了条件密度 $f_{X|Y}(x|y)$, 令 $\nu(y,\cdot)$ 来表示其对应的测度, 则

$$\nu(y,B)=\int_B f_{X|Y}(x|y)\,\mathrm{d}x\,,\quad \text{for } y\in\mathbb{R},\ B\in\mathcal{B}(\mathbb{R})\,.$$

例 19 已经蕴含了这一点.

定理 87. *设 ξ,η 分别是 $(\Omega,\mathcal{F},\mathbb{P})$ 上取值于 $(E,\mathcal{E})$ 和 $(S,\mathcal{S})$ 的随机元, ν 为 ξ 关于 η 给定值的正则条件分布. 对任意 $(E,\mathcal{E})$ 上的实值 Borel 可测函数 f, 若 $f(\xi)$ 的积分存在, 则对 (关于 $\mathbb{P}_\eta$) 几乎处处的 $y\in S$, f 关于概率测度 $\nu(y,\cdot)$ 可积, 且成立:*

$$E[f(\xi)\mid\eta=y]=\int_E f(x)\,\nu\,(y,\mathrm{d}x)\quad \mathbb{P}_\eta\text{-}a.e.\ y\,.$$

证明. 由定理 85, 和前文中的 "注", 对几乎处处 ω, f 关于测度 $\nu(\eta(\omega),\cdot)$ 的可积, 且

$$E[f(\xi)\mid\eta](\omega)=\int_E f(x)\,\nu\,(\eta(\omega),\mathrm{d}x)\quad \text{a.s. } \omega\,.$$

于是对 (关于 $\mathbb{P}_\eta$) 几乎处处的 $y\in S$, f 关于概率测度 $\nu(y,\cdot)$ 可积, 且

$$E[f(\xi)\mid\eta=y]=\int_E f(x)\,\nu\,(y,\mathrm{d}x)\quad \mathbb{P}_\eta\text{-a.e. } y\,.$$ □

最后, 利用定理 86 和和前文中的 "注", 我们立刻得到下面关于存在性的结果.

定理 88. *X 是 $(\Omega,\mathcal{F},\mathbb{P})$ 是上随机变量, η 是取值于 $(S,\mathcal{S})$ 中的随机元. 则 X 关于 η 给定值的正则条件分布存在.*

注. 同样的, 若 $(E,\mathcal{E})$ 与 $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ 可测同构, 则上面定理对取值于 $(E,\mathcal{E})$ 中的随机元 ξ 也成立.

第五章　乘积空间

一族可测空间相乘即构成所谓的乘积可测空间. 本章将讨论在有限维乘积空间上如何通过转移函数产生测度, 在可列维乘积空间上如何通过概率转移函数产生概率测度和在任意无穷维乘积空间上如何通过相容的有限维分布族产生概率测度. 概率论中的随机向量和随机过程等重要概念都与乘积空间有关. 在开始本章内容之前, 先来回忆下笛卡尔积的定义. 令 $(X_t)_{t\in T}$ 是一指标集族. 我们称

$$\prod_{t\in T} X_t = \left\{ x : T \to \bigcup_{t\in T} X_t : x(t) \in X_t, \forall\, t \in T \right\}$$

是 $(X_t)_{t\in T}$ 的**乘积**. 我们也将乘积空间中的元记为 $(x_t)_{t\in T}$, 以表示映射 x 作用在 $t \in T$ 上的像是 $x_t \in X_t$. 特别的, 当指标集 T 分别是 $\{1, 2, \cdots, n\}, \mathbb{N}$ 时, 我们也将乘积空间分别写成

$$\prod_{k=1}^{n} X_k \quad , \quad \prod_{k=1}^{\infty} X_k \ ,$$

其中的元素分别写为 $(x_1, \cdots, x_n)$, $(x_1, x_2, \cdots)$.

在乘积空间上我们有投影的概念: 设 S 是 T 的子集, 我们称 π_S 是 $\Pi_{t\in T} X_t$ **到** $\Pi_{t\in S} X_t$ **的投影**, 如果

$$\pi_S : (x_t)_{t\in T} \mapsto (x_t)_{t\in S} \, .$$

特别的, 将 $\prod_{t\in T} X_t$ 到 X_t 的投影 $\pi_{\{t\}}$ 简记为 π_t. 此外, 若 $S_2 \subset S_1 \subset T$, 我们将 $\Pi_{t\in S_1} X_t$ 到 $Pi_{t\in S_2} X_t$ 的投影记为 $\pi_{S_2}^{S_1}$, 即

$$\pi_{S_2}^{S_1} : (x_t)_{t\in S_1} \mapsto (x_t)_{t\in S_2} \, .$$

5.1 乘积可测空间

5.1.1 有限维乘积空间

我们先来介绍有限个可测空间的乘积. 设 $(X_k, \mathcal{F}_k)$, $k = 1, \cdots, n$ 皆为可测空间, 我们要在其乘积空间上规定一个 σ 域. 一个自然的要求是, 使得如下集合系可测:

$$\mathcal{Q} \coloneqq \left\{ \prod_{k=1}^{n} A_k : A_k \in \mathcal{F}_k \right\}.$$

我们称 $\mathcal{Q}$ 中的元素为**可测矩形**, 容易看出 $\mathcal{Q}$ 是乘积空间上的半代数. 我们将 $\mathcal{Q}$ 生成的 σ 域, 也就是 $\sigma(\mathcal{Q})$, 称为 $\mathcal{F}_1, \cdots, \mathcal{F}_k$ 的**乘积 σ 域**, 并记之为

$$\prod_{i=1}^{n} \mathcal{F}_i \coloneqq \sigma(\mathcal{Q}).$$

注. 在点集拓扑的课程中, 我们在乘积空间上定义的拓扑是使得全体投影映射连续的最小的拓扑. 实际上容易看出, 乘积空间上的乘积 σ 域是使得全体投影可测的最小的 σ 域, 即

$$\prod_{k=1}^{n} \mathcal{F}_k = \sigma(\pi_1, \cdots, \pi_n).$$

截口的可测性

回忆在数学分的课程中, 我们定义了有界区域 D 上的二重黎曼积分, 且我们证明了重积分等于累次积分, 累次积分的顺序是可以交换的:

$$\iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int \left(\int f(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int \left(\int f(x,y)\,\mathrm{d}y \right) \mathrm{d}x.$$

我们后文中会证明, 这个定理在抽象的测度积分下仍然成立. 但是现在我们先解决一个基本的问题, 就是当我们固定 x 时, 需要讨论函数 $f(x,\cdot)$ 的积分, 那我们首先要证明它必须是可测函数才行. 首先我们引入如下的定义.

定义 32. 设 f 是 $\Pi_{k=1}^{n} X_k$ 到集合 Y 的映射, S 是 $\{1, \cdots, n\}$ 的非空子集, $(x_k)_{k\in S} \in \Pi_{k\in S} X_k$. 我们称 $\Pi_{k\in S^c} X_k$ 到 Y 的映射

$$f|_{(x_k)_{k\in S}} : (x_k)_{k\in S^c} \mapsto f(x_1, \cdots, x_n)$$

为 f 在 $(x_k)_{k\in S}$ 处的**截口**. 特别的, 对 $A \in \Pi_{k=1}^{n} X_k$, 我们称 $\Pi_{k\in S^c} X_k$ 的子集

$$A|_{(x_k)_{k\in S}} \coloneqq \{(x_k)_{k\in S^c} : (x_1, \cdots, x_n) \in A\}.$$

为 A 在 $(x_k)_{k\in S}$ 处的**截口**.

如同所期望的那样, 我们很容易证明截口的可测性.

命题 89. 设 $A \in \prod_{k=1}^n \mathcal{F}_k$, 则 A 在 $(x_k)_{k\in S}$ 处的截口 $A|_{(x_k)_{k\in S}} \in \prod_{k\in S^c} \mathcal{F}_k$; 设 f 是 $\prod_{k=1}^n X_k$ 到 $(Y, \mathcal{S})$ 的可测映射, 则 f 在 $(x_k)_{k\in S}$ 处的截口 $f|_{(x_k)_{k\in S}}$ 是 $\prod_{k\in S^c} X_k$ 到 $(Y, \mathcal{S})$ 的可测映射.

证明. 利用好集原理来证. 对给定的 $(x_k)_{k\in S} \in \prod_{k\in S} X_k$, 我们记

$$\mathcal{G} = \{A \in \prod_{k=1}^n \mathcal{F}_k : A|_{(x_k)_{k\in S}} \in \prod_{k\in S^c} \mathcal{F}_k\}$$

显然可测矩形全体 $\mathcal{Q} \subset \mathcal{G}$. 易见 $\mathcal{G}$ 是 σ 域. 因此有 $\sigma(\mathcal{Q}) = \prod_{k=1}^n \mathcal{F}_k \subset \mathcal{G}$. 从而 $\sigma(\mathcal{Q}) = \mathcal{G}$.

若 f 是 $\prod_{k=1}^n X_k$ 到 $(Y, \mathcal{S})$ 的可测映射, 对任意 $B \in \mathcal{S}$, 注意

$$\{f|_{(x_k)_{k\in S}} \in B\} = \{f \in B\}|_{(x_k)_{k\in S}} \in \prod_{k\in S^c} \mathcal{F}_k$$

从而完成了证明. □

练习 15. 设 $\mathbb{R}$ 为实数集, 证明 $\mathcal{B}(\mathbb{R}) \times \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

5.1.2 任意维的乘积空间

下面我们仿照有限维的情形, 建立任意无穷维的乘积可测空间.

$(X_t, \mathcal{F}_t)_{t\in T}$ 是一族可测空间. 我们规定乘积空间 $\prod_{t\in T} X_t$ 上的 σ 域为使得投影 $\{\pi_t\}_{T\in t}$ 可测的最小 σ 域, 称值为 $\{\mathcal{F}_t\}_{t\in T}$ 的乘积 σ 域, 记为 $\prod_{t\in T} \mathcal{F}_t$. 换言之,

$$\prod_{t\in T} \mathcal{F}_t \coloneqq \sigma(\pi_t\,, t \in T) = \sigma\big(\bigcup_{t\in T} \pi_t^{-1}\mathcal{F}_t\big)\,.$$

注. 我们可用 “更少” 的集合生成乘积 σ 域. 设对任何 t, $\mathcal{F}_t$ 可由其子集合系 $\mathcal{E}_t$ 生成, 即 $\mathcal{F}_t = \sigma(\mathcal{E}_t)$. 则

$$\prod_{t\in T} \mathcal{F}_t = \sigma\big(\bigcup_{t\in T} \pi_t^{-1}(\mathcal{E}_t)\big)\,.$$

一个自然的问题是, 投影 π_S 是乘积空间上的可测映射吗? 我们下面来考察这一点.

命题 90. 设 S 是 T 的非空子集, 则投影 π_S 是 $(\prod_{t\in T} X_t, \prod_{t\in T}\mathcal{F}_t)$ 到 $(\prod_{t\in S} X_t, \prod_{t\in S}\mathcal{F}_t)$ 的可测映射.

证明. 由定义知

$$\prod_{t\in S}\mathcal{F}_t \coloneqq \sigma(\pi_t^S, t\in S) = \sigma\big(\bigcup_{t\in S}(\pi_t^S)^{-1}\mathcal{F}_t\big)\,.$$

只需证, 对任意 $t\in S$,

$$\pi_S^{-1}\big((\pi_t^S)^{-1}\mathcal{F}_t\big) \subset \prod_{t\in T}\mathcal{F}_t\,.$$

注意到 $\pi_S^{-1}\circ(\pi_t^S)^{-1} = (\pi_t^S\circ\pi_S)^{-1} = \pi_t^{-1}$ 即得. □

更一般的, 我们有

定理 91. 设 f 是 $(\Omega,\mathcal{F})$ 到乘积可测空间 $(\prod_{t\in T} X_t, \prod_{t\in T}\mathcal{F}_t)$ 的映射. 则 f 为可测映射当且仅当其分量 $f_t \coloneqq \pi_t\circ f$ 是到 $(X_t,\mathcal{F}_t)$ 的可测映射.

证明. 由 π_t 皆可测知, 若 f 可测, 其任意分量 f_t 可测. 若任意 $t\in T$, f_t 可测, 注意

$$f^{-1}(\pi_t^{-1}\mathcal{F}_t) = (f^{-1}\circ\pi_t^{-1})\mathcal{F}_t = f_t^{-1}\mathcal{F}_t \subset \mathcal{F}\,,$$

于是

$$f^{-1}\big(\bigcup_{t\in T}\pi_t^{-1}\mathcal{F}_t\big) \subset \mathcal{F}\,.$$

于是 f 可测. □

注. 我们将从 $(\Omega,\mathcal{F})$ 到 $(\mathbb{R}^n,\mathcal{B}_{\mathbb{R}}^n)$, $(\mathbb{R}^\infty,\mathcal{B}_{\mathbb{R}}^\infty)$, $(\mathbb{R}^T,\mathcal{B}_{\mathbb{R}}^T)$ 的可测映射 X 分别称为 **n 维随机向量**, **随机变量序列**, T **上的随机过程**. 由上述定理我们知道, 这等价于要求 X 的任何分量皆是随机变量.

此外, 我们还经常用到下面几种集合系来生成乘积 σ 域 $\prod_{t\in T}\mathcal{F}_t$. 记 $\mathcal{Q}$ 为**有限维可测矩形柱集**全体:

$$\mathcal{Q} \coloneqq \left\{\pi_S^{-1}\left(\prod_{t\in S}A_t\right) : A_t\in\mathcal{F}_t,\, t\in S,\, S\subset T \text{ is finite}\right\}.$$

记 $\mathcal{A}$ 为**有限维可测柱集**全体:

$$\mathcal{A} \coloneqq \bigcup_{S\subset T \text{ is finite}} \pi_S^{-1}\left(\prod_{t\in S}\mathcal{F}_t\right).$$

命题 92. $\mathcal{Q}, \mathcal{A}$ 分别是 $\prod_{t\in T} X_t$ 上的半代数和代数, 且

$$\sigma(\mathcal{Q}) = \sigma(\mathcal{A}) = \prod_{t\in T} \mathcal{F}_t \,.$$

证明. $\mathcal{Q}, \mathcal{A}$ 分别是 $\prod_{t\in T} X_t$ 上的半代数和代数是容易验证的, 同时有

$$\bigcup_{t\in T} \pi_t^{-1}\mathcal{F}_t \subset \mathcal{Q} \subset \mathcal{A} \subset \sigma(\mathcal{A}) = \prod_{t\in T} \mathcal{F}_t \,,$$

于是得证. □

下面的定理是有用的, 他说明了乘积 σ 代数的构造特点.

定理 93. T 是无穷集, 则乘积 σ 域恰是可列维可测柱集全体.

$$\prod_{t\in T} \mathcal{F}_t = \bigcup_{I\subset T \text{ is countable}} \pi_I^{-1}\left(\prod_{t\in I} \mathcal{F}_t\right) \tag{5.1}$$

证明. 注意到

$$\bigcup_{t\in T} \pi_t^{-1}\mathcal{F}_t \subset \bigcup_{I\subset T \text{ is countable}} \pi_I^{-1}\left(\prod_{t\in I} \mathcal{F}_t\right) \subset \prod_{t\in T} \mathcal{F}_t \,,$$

只需证式 (5.1) 右端 (下面记为 RHS, right hand side) 为 σ 域. 显然其包含 $\prod_{t\in T} X_t$, 且对取余运算封闭. 下证其对可列交运算封闭.

对任意正整数 n, 取 $\pi_{I_n}^{-1}(A_n) \in \text{RHS}$, 其中 $A_n \in \prod_{t\in I_n} \mathcal{F}_t$. 要证

$$\bigcap_{n=1}^{\infty} \pi_{I_n}^{-1}(A_n) \in \text{ RHS } .$$

令 $I = \cup_{n=1}^{\infty} I_n$, 则 I 可数, 且

$$\bigcap_{n=1}^{\infty} \pi_{I_n}^{-1}(A_n) = \pi_I^{-1}\left(\bigcap_{n=1}^{\infty} \left(\pi_{I_n}^{I}\right)^{-1}(A_n)\right) \in \text{ RHS } .$$

于是定理得证. □

我们给出定理 93 的一个应用, 在学习布朗运动时将会遇到这个问题.

练习 16. 考虑乘积可测空间 $(\mathbb{R}^{[0,\infty)}, \mathcal{B}(\mathbb{R})^{[0,\infty)})$. 我们记 $C[0,\infty)$ 是 $[0,\infty)$ 到 $\mathbb{R}$ 的连续函数全体. 证明

$$C[0,\infty) \notin \mathcal{B}(\mathbb{R})^{[0,\infty)} \,.$$

5.2 二维乘积测度

我们希望讨论在有限维乘积空间上如何定义测度. 为了清楚简单, 我们本节只讨论二维乘积空间, 只要二维的情况说清楚了, 就不难把所得到的结果推广到一般的有限维乘积可测空间上去, 推广会在下一节给出.

设 $(X_1, \mathcal{F}_1, \mu_1)$ 与 $(X_2, \mathcal{F}_2, \mu_2)$ 是两个测度空间. 我们考虑如何借助 μ_1 与 μ_2 在乘积可测空间 $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2)$ 上自然地定义一个测度. 注意 $\mathcal{F}_1 \times \mathcal{F}_2$ 可由半代数, 即可测矩形全体

$$\mathcal{Q} = \{A_1 \times A_2 : A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\},$$

生成. 自然的, 我们想到用 1.3 节中的*测度延拓*的方法, 只要先在半代数 $\mathcal{Q}$ 上定义好了测度, 再由定理 17 就能将其延拓到 $\mathcal{F}_1 \times \mathcal{F}_2$ 上. 故我们先在半代数 $\mathcal{Q}$ 上定义测度.

联想到平面的中矩形的面积等于长与宽的乘积, 我们自然的将可测矩形 $A_1 \times A_2$ 的 "面积", 也就是测度, 定义成两个边 A_1, A_2 的长度之积, 也就是 $\mu_1(A_1)\mu_2(A_2)$. 我们把这个集函数记作 $\mu_1 \times \mu_2$. 换言之,

$$(\mu_1 \times \mu_2)(A_1 \times A_2) \coloneqq \mu_1(A_1)\mu_2(A_2), \ \forall\, A_1 \times A_2 \in \mathcal{Q}.$$

下面来证明 $\mu_1 \times \mu_2$ 确实是 $\mathcal{Q}$ 上的测度. 只需要验证可列可加性. 设 $\{A_n \times B_n\}$ 是 $\mathcal{Q}$ 中互不相交的集列, 且 $A \times B = \cup_{n=1}^{\infty}(A_n \times B_n)$ 我们只需证明

$$\mu_1(A)\mu_2(B) = \sum_{n=1}^{\infty} \mu_1(A_n)\mu_2(B_n).$$

我们注意到对任意 $(x_1, x_2) \in X_1 \times X_2$,

$$I_A(x_1)I_B(x_2) = \sum_{n=1}^{\infty} I_{A_n}(x_1)I_{B_n}(x_2).$$

任取 $x_1 \in X_1$ 固定, 对 $x_2 \in X_2$ 积分就有

$$I_A(x_1)\mu_2(B) = \sum_{n=1}^{\infty} I_{A_n}(x_1)\mu_2(B_n).$$

再对 $x_1 \in X_1$ 积分, 就证明了可列可加性.

定理 94. *$(X_1, \mathcal{F}_1, \mu_1)$ 与 $(X_2, \mathcal{F}_2, \mu_2)$ 是两个测度空间. 在乘积可测空间 $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2)$ 存在测度 $\mu_1 \times \mu_2$, 满足对任何 $A_1 \times A_2 \in \mathcal{Q}$, 有*

$$(\mu_1 \times \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2). \tag{5.2}$$

若 μ_1, μ_2 皆是 σ 有限的, 在乘积可测空间上满足 (5.2) 的测度是唯一的, 且是 σ 有限的.

证明. 注意当 μ_1, μ_2 皆是 σ 有限测度时, 存在 $X_1 \times X_2$ 的划分 $\{A_n \times B_n\}$, 满足对任意 n, $A_n \times B_n \in \mathcal{Q}$ 且 $(\mu_1 \times \mu_2)(A_n \times B_n) < \infty$, 于是由定理 18 得到唯一性. □

为了保证定义是唯一的, 我们假设 μ_1, μ_2 都是 σ 有限测度. 这样好处在于测度 $\mu_1 \times \mu_2$ 是唯一指定的. 虽然如此, 我们对 $A \in \sigma(\mathcal{Q})$ 这样的集合的测度 $(\mu_1 \times \mu_2)(A)$ 却没给出一个显示的计算方法. 下面我们就解决这个问题. 注意测度总是示性函数的积分,

$$(\mu_1 \times \mu_2)(A) = \int_{X_1 \times X_2} 1_A \, \mathrm{d}(\mu_1 \times \mu_2) \, ,$$

这样问题变成了上面积分如何计算. 回忆微积分中, 在进行二重黎曼积分的计算时, 我们总是用累次积分来算的, 是不是现在仍然有累次积分的公式呢? 这就是下面的 Fubini 定理, 它表明乘积测度空间上的积分可以表示为累次积分的形式.

定理 95 (Fubini 定理)**.** *μ_1 与 μ_2 是 σ 有限测度. f 是乘积测度空间 $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ 上积分存在的可测函数, 则*

1. *$x_1 \mapsto \int_{X_2} f(x_1, x_2)\, \mu_2(\mathrm{d}x_2)$ 是 μ_1-a.e. 定义的可测函数; $x_2 \mapsto \int_{X_1} f(x_1, x_2)\, \mu_1(\mathrm{d}x_1)$ 是 μ_2-a.e. 定义的可测函数*

2. *累次积分可交换顺序, 且等于重积分:*

$$\begin{aligned}\int_{X_1 \times X_2} f \, \mathrm{d}(\mu_1 \times \mu_2) &= \int_{X_1} \mu_1(\mathrm{d}x_1) \int_{X_2} f(x_1, x_2)\, \mu_2(\mathrm{d}x_2) \\ &= \int_{X_2} \mu_2(\mathrm{d}x_2) \int_{X_1} f(x_1, x_2)\, \mu_1(\mathrm{d}x_1) \, .\end{aligned}$$

证明. 本定理是证明是标准的: 好集原理与典型方法.

1. 对可测矩形 $\mathcal{Q}$ 中的集对应的示性函数证明.

2. 对任何可测集对应的示性函数成立, 这里用了 μ_1, μ_2 σ 有限.

3. 对任何非负简单可测函数成立; 对任何非负可测函数成立.

4. 对积分存在的可测函数成立. □

我们将指出, 若 μ_1 与 μ_2 不全是 σ 有限时, 会产生额外的麻烦. 于是若想应用 Fubini 定理, 我们必须要求两个测度 σ 有限.

例 21. 设 $X_1 = X_2 = (0,1)$, $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}(0,1)$, μ_1 是 Lebesgue 测度, μ_2 是计数测度. 可见 μ_1 是 σ 有限的, μ_2 不是. 此时考虑对角线

$$D = \{(x_1, x_2)\,; x_1 = x_2; x_1, x_2 \in (0,1)\}$$

就有

$$\int_{(0,1)} \mu_2\left(D|_{x_1}\right)\mu_1(\mathrm{d}x_1) = 1 \neq 0 = \int_{(0.1)} \mu_1\left(D|_{x_2}\right)\mu_2(\mathrm{d}x_2)\ .$$

练习 17. 设 $(X_1, \mathcal{F}_1)$ 与 $(X_2, \mathcal{F}_2)$ 皆是可测空间, $\mu_1 \ll \nu_1$ 皆是 $(X_1, \mathcal{F}_1)$ 上 σ 有限测度, $\mu_2 \ll \nu_2$ 皆是 $(X_2, \mathcal{F}_2)$ 上 σ 有限测度. 证明 $\nu_1 \times \nu_2 \ll \mu_1 \times \mu_2$, 且对 R-N 导数有

$$\frac{\mathrm{d}\left(\nu_1 \times \nu_2\right)}{\mathrm{d}\left(\mu_1 \times \mu_2\right)}(x, y) = \frac{\mathrm{d}\nu_1}{\mathrm{d}\mu_1}(x)\,\frac{\mathrm{d}\nu_2}{\mathrm{d}\mu_2}(y), \quad \mu_1 \times \mu_2\text{-a.e..}$$

Fubini 定理允许我们交换累次积分的顺序, 应用十分广泛. 我们给出用 Fubini 定理推到 L-S 积分的分部积分公式的例子.

练习 18 (分部积分公式). 设 F, G 是准分布函数 (单调递增, 右连续), 证明对任何实数 $a \leqslant b$,

$$\int_{(a,b]} F(x)G(\mathrm{d}x) = FG|_a^b - \int_{(a,b]} G(x-)F(\mathrm{d}x)$$

其中 $FG|_a^b = F(b)G(b) - F(a)G(a)$.

5.3 由 σ 有限核产生的测度

本节通过引入所谓 “核”, 亦称为 “测度转移函数” 的概念来推广上一节在二维空间中定义测度, 给出 Fubini 定理. 最后将所有结论推广到有限维乘积空间上.

定义 33. $(X, \mathcal{F})$ 与 $(Y, \mathcal{S})$ 是两个可测空间. 称 $K : X \times \mathcal{S} \to [0, \infty]$ 是 $(X, \mathcal{F})$ 到 $(Y, \mathcal{S})$ 的**核** (kernel) 或**测度转移函数**, 如果

1. 任何 $x \in X$, $K(x, \cdot)$ 是 $(Y, \mathcal{S})$ 上的测度.

2. 任何 $A \in \mathcal{S}$, $K(\cdot, A)$ 是 $(X, \mathcal{F})$ 上的非负可测函数.

称 K 是**有限核**, 若 $\{K(x, \cdot)\}$ 是有限测度族; 称 K 是 σ **有限核**, 如果 $\{K(x, \cdot)\}$ "一致地" σ 有限: 存在 Y 的划分 $\{A_n\}$, 满足 $A_n \in \mathcal{S}$, 且 $p(x, A_n) < \infty$ 对任何 $x \in X$ 和 $n \in \mathbb{N}$ 成立. 此外, 若 $\{K(x, \cdot)\}$ 是概率测度族, 就称 K 是**概率转移函数**或**概率核**, 并且一般写为 p 或 P.

例 22. 可测空间 $(X, \mathcal{F})$ 到测度空间 $(Y, \mathcal{S}, \nu)$ 可以定义如下的核: 对任何 $x \in X$ 和 $A \in \mathcal{S}$, 定义 $K(x, A) = \nu(A)$.

例 23. 离散时间, 可数状态空间 I 上的 Markov 链的 n 步转移矩阵 P_n 皆是 S 到自身的概率转移函数.

例 24. 回忆第四章中所讲的正则条件分布: ξ 是 $(\Omega, \mathcal{F})$ 到 $(E, \mathcal{E})$ 的随机元, $\mathcal{G}$ 是 $\mathcal{F}$ 的子 σ 代数, 则 ξ 关于 $\mathcal{G}$ 的正则条件分布就是 $(\Omega, \mathcal{G})$ 到 $(E, \mathcal{E})$ 的概率转移函数.

下面的定理说明, $(X, \mathcal{F})$ 到 $(Y, \mathcal{S})$ 的核, 可以将 $(X, \mathcal{F})$ 的测度 "转移" 到 $(Y, \mathcal{S})$ 上, 这也是为什么叫它测度转移函数的原因.

定理 96. *设 K 是 $(X, \mathcal{F})$ 到 $(Y, \mathcal{S})$ 的核. μ 是 $(X, \mathcal{F})$ 的测度, 则 $\mathcal{S}$ 上定义的如下集函数 ν 是一测度:*

$$\nu(B) := \int K(x, B)\mu(\mathrm{d}x), \quad \textit{for all } B \in \mathcal{S}.$$

特别的, 若 μ 是概率测度, K 是概率核, 则 ν 亦为概率测度.

证明. 首先给定 B 后, 有核的定义知 $K(\cdot, B)$ 是 $(X, \mathcal{F})$ 上非负可测函数, 于是积分有定义. 由 Levi 单调收敛定理立刻得到 ν 的可数可加性. □

定理 97. *设 K 是 $(X, \mathcal{F})$ 到 $(Y, \mathcal{S})$ 的核, μ 是 $(X, \mathcal{F})$ 的测度, ν 的定义同定理 96, f 是 $(Y, \mathcal{S}, \nu)$ 上积分存在的可测函数. 那么*

1. *$x \mapsto \int_Y f(y)K(x, \mathrm{d}y)$ 是 $(X, \mathcal{F}, \mu)$ 上 a.e. 有定义的可测函数.*

2. *我们有*
$$\int_Y f(y)\nu(\mathrm{d}y) = \int_X \mu(\mathrm{d}x) \int_Y f(y)K(x, \mathrm{d}y).$$

证明. 当 $f = I_B$, 其中 $B \in \mathcal{S}$ 时, 由 K 的定义与定理 96 知成立; 于是对 f 为简单非负可测函数时成立; 当 f 是非负可测函数时, 取一列递增的非负简单函数逼近 f, 由单调收敛定理知定理成立. 同时可见 1 中的函数对任意 x 皆有定义. 我们下面会指出, 对一般 f 只能做到 μ-a.e. 有定义.

当 f 为积分存在的可测函数时, 不妨设 f^- 可积. 于是有

$$\int_Y f^-(y)\nu(\mathrm{d}y) = \int_X \mu(\mathrm{d}x) \int_Y f^-(y)K(x, \mathrm{d}y) < \infty\,.$$

从而

$$\int_Y f^-(y)K(x, \mathrm{d}y) < \infty,\ \mu\text{ -a.e.}.$$

于是可见

$$\int_Y f(y)K(x, \mathrm{d}y) = \int_Y f^+(y)K(x, \mathrm{d}y) - \int_Y f^-(y)K(x, \mathrm{d}y)$$

仅仅 μ-a.e. 有意义, 同时无意义的点是可测集, 于是 $x \mapsto \int_Y f(y)K(x, \mathrm{d}y)$ 可测. 再由积分的线性性质,

$$\begin{aligned}\int_Y f(y)\nu(\mathrm{d}y) &= \int_Y f^+(y)\nu(\mathrm{d}y) - \int_Y f^-(y)\nu(\mathrm{d}y)\\ &= \int_X \mu(\mathrm{d}x)\int_Y f^+(y)K(x, \mathrm{d}y) - \int_X \mu(\mathrm{d}x)\int_Y f^-(y)K(x, \mathrm{d}y)\\ &= \int_X \mu(\mathrm{d}x)\int_Y f(y)K(x, \mathrm{d}y)\,.\end{aligned}$$

于是完成了证明. □

下面我们利用核来构造乘积空间上的测度. 在接下来的证明中可以看到, 我们必须用 σ 有限核才能在乘积空间中构造测度, 并推广 Funibi 定理.

我们的想法是这样的: 考虑可测空间 $(X_1, \mathcal{F}_1)$ 和 $(X_2, \mathcal{F}_2)$, μ_1 是前者上的测度, K 是前者到后者的 σ 核. 对 $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2)$ 中的可测矩形 $A_1 \times A_2$, 对 $x_1 \in A$, 这一点矩形的 “高” 为 $K(x_1, A_2)$, 因为高的度量随着 x_1 的不同取值是在变动的, 我们不能像之前一样直接相乘; 而是应该用 “积分” 的想法. 于是可将矩形 $A_1 \times A_2$ 的测度定义为

$$\int_{A_1} K(x_1, A_2)\mu_1(\mathrm{d}x_1)$$

类似的, 对任何 $A \in \sigma(\mathcal{Q})$, 我们可以类似的定义其测度为

$$\int_{X_1} K(x_1, A|_{x_1})\mu(\mathrm{d}x_1)\,.$$

现在首当其冲的问题是, $x_1 \mapsto K(x_1, A|_{x_1})$ 是否为 $(X_1, \mathcal{F}_1)$ 上可测函数? 细心的我们发现, 固定 $x_1 \in X_1$，$A \mapsto K(x_1, A|_{x_1})$ 是 $\mathcal{F}_1 \times \mathcal{F}_2$ 上的测度. 于是我们的问题等价于如下定义的函数 N 是否为 $(X_1, \mathcal{F}_1)$ 到 $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2)$ 的 (σ 有限) 核?

$$N(x_1, A) \coloneqq K(x_1, A|_{x_1}), \text{ for } x_1 \in X_1, A \in \mathcal{F}_1 \times \mathcal{F}_2 .$$

下面我们就来回答这个问题.

引理 98. K 是 $(X_1, \mathcal{F}_1)$ 到 $(X_2, \mathcal{F}_2)$ 的 σ 有限核. 对 $A \in \mathcal{F}_1 \times \mathcal{F}_2$, 定义

$$N(x_1, A) \coloneqq K(x_1, A|_{x_1})$$

则 $N(\cdot, A)$ 是 $(X_1, \mathcal{F}_1)$ 上的非负可测函数.

证明. 证明方法自然是典型的好集原理. 为了书写方便, 我们在证明中把 $N(\cdot, A)$ 写为 N_A.

当 $A = A_1 \times A_2 \in \mathcal{Q}$ 时, 可见

$$N_A(x_1) = I_{A_1}(x_1) K(x_1, A_2), \forall x_1 \in X_1 .$$

由测度转移函数定义知 N_A 是 $(X_1, \mathcal{F}_1)$ 上的非负可测函数. 现在我们希望能用 $\pi - \lambda$ 定理将这样的性质扩到 $\mathcal{F}_1 \times \mathcal{F}_2$ 上. 令

$$\mathcal{G} = \{A \in \mathcal{F}_1 \times \mathcal{F}_2 : N_A \text{ 在 } (X_1, \mathcal{F}_1) \text{ 非负可测. }\}$$

显然 $\mathcal{Q} \subset \mathcal{G}$. 下面只需证 $\mathcal{G}$ 是一个 λ 系. 显然 $X_1 \times X_2 \in \mathcal{G}$, 且 $\mathcal{G}$ 对递增集列极限封闭是容易的: 若 $\{A_n\} \subset \mathcal{G}$, $A_n \uparrow A$, 此时 $N_{A_n} \uparrow N_A$, 故有 $A \in \mathcal{G}$. 下面证明 $\mathcal{G}$ 对真差运算封闭: 若 $A, B \in \mathcal{G}$, 且 $B \subset A$, 要证 $A \backslash B \in \mathcal{G}$.

注意这时 $K\left(x_1, (A\backslash B)|_{x_1}\right) = K\left(x_1, A|_{x_1}\right) - K\left(x_1, B|_{x_1}\right)$ 对 $x_1 \in X_1$ 不一定成立，除非 $K(x_1, \cdot)$ 是有限测度. 但通过假设核 K 的 σ 有限性，可将 K 写可数个有限核之和，从而可以完成证明.

先假设 K 是有限核, 则对任意 $x_1 \in X_1$,

$$K\left(x_1, (A\backslash B)|_{x_1}\right) = K\left(x_1, A|_{x_1}\right) - K\left(x_1, B|_{x_1}\right)$$

于是 $N_{A\backslash B} = N_A - N_B$ 可测, 故 $A\backslash B \in \mathcal{G}$. 于是对有限核 K, 命题成立.

当 K 为 σ 有限核时, 设 D_n 是 X_2 的划分, 使得 $D_n \in \mathcal{F}_2$, $K(x_1, D_n)$ 有限对任何 $x_1 \in X_1$ 和任何正整数 n 成立. 对任意 n, 令

$$K_n(x_1, D) = K(x_1, D \cap D_n), \text{ for } x_1 \in X_1, D \in \mathcal{F}_2 .$$

则 $\{K_n\}$ 是一列有限核. 由上文, 对 $A \in \mathcal{F}_1 \times \mathcal{F}_2$, $K_n\left(x_1, A|_{x_1}\right)$ 是关于 x_1 的非负可测函数. 同时

$$K\left(x_1, A|_{x_1}\right) = \sum_{k=1}^{\infty} K\left(x_1, D_k \cap A|_{x_1}\right) = \sum_{k=1}^{\infty} K_n\left(x_1, A|_{x_1}\right)$$

故 N_A 是 $(X_1, \mathcal{F}_1)$ 上非负可测函数. □

利用定理96,, 很容易得到:

定理 99. *$(X_1, \mathcal{F}_1, \mu_1)$ 是一测度空间. K 是 $(X_1, \mathcal{F}_1)$ 到 $(X_2, \mathcal{F}_2)$ 的 σ 有限核. 定义 $\mathcal{F}_1 \times \mathcal{F}_2$ 上的集函数 μ 为*

$$\mu(A) \coloneqq \int_{X_1} K(x_1, A|_{x_1})\mu_1(\mathrm{d}x_1)\,, \ \textit{for}\ A \in \mathcal{F}_1 \times \mathcal{F}_2\,.$$

则 μ 是一个测度. 特别的, 对任何可测矩形 $A_1 \times A_2$,

$$\mu\left(A_1 \times A_2\right) = \int_{A_1} K\left(x_1, A_2\right)\mu_1\left(dx_1\right)\,. \tag{5.3}$$

若 μ_1 也 σ 有限的, 则在乘积可测空间上满足使得 (5.3) 成立的测度 μ 是唯一的, 且是 σ 有限的.

证明. 直接验证可知 μ 是一测度. 下证唯一性. 设 μ_1 是 σ 有限测度. 若两个测度 μ, ν 皆是满足 (5.3) 的测度, 于是 μ, ν 在 $\mathcal{Q}$ 上相等. 易见存在可测矩形列 $\{A_1^{(n)} \times A_2^{(n)}\}$, 它是 X 的划分, 且 $\mu(A_1^{(n)} \times A_2^{(n)}) < \infty$. 用第二章定理18, 即得 $\mu = \nu$. 同时可见 μ 是 σ 有限的. □

在定理 97 中, 令 $X = X_1, \mathcal{F} = \mathcal{F}_1$, $Y = (X_1, X_2), \mathcal{S} = \mathcal{F}_1 \times \mathcal{F}_2$. 我们就得到下面的累次积分 的公式.

$$\int_{X_1 \times X_2} f\,\mathrm{d}\mu = \int_{X_1} \mu_1\left(\mathrm{d}x_1\right) \int_Y f\left(y\right) N\left(x_1, \mathrm{d}y\right)\,.$$

注意到对任意使得下式左端有意义的 x_1, 必然有

$$\int_Y f\left(y\right) N\left(x_1, \mathrm{d}y\right) = \int_{X_2} f\left(x_1, x_2\right) K\left(x_1, \mathrm{d}x_2\right)\,,$$

我们得到如下的形式的 Fubini 定理.

定理 100 (Fubini)**.** *f 是乘积测度空间 $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu)$ 上积分存在的可测函数, 则 $x_1 \mapsto \int_{X_2} f\left(x_1, x_2\right) K\left(x_1, \mathrm{d}x_2\right)$ 是 $(X_1, \mathcal{F}_1)$ 上 μ_1-a.e. 定义的可测函数, 且有*

$$\int_{X_1 \times X_2} f\,\mathrm{d}\mu = \int_{X_1} \mu_1\left(\mathrm{d}x_1\right) \int_{X_2} f\left(x_1, x_2\right) K\left(x_1, \mathrm{d}x_2\right)\,.$$

有限维空间

设 $(X_j,\mathcal{F}_j)$, $j=1,2,\cdots,n$ 是可测空间. μ_1 是 $(X_1,\mathcal{F}_1)$ 上的测度. 对 $2\le j\le n$, 设 $K(x_1,\cdots,x_{j-1},\mathrm{d}x_j)$ 是从 $\left(\Pi_{i=1}^{j-1}X_i,\Pi_{i=1}^{j-1}\mathcal{F}_i\right)$ 到 $(X_j,\mathcal{F}_j)$ 的 σ 有限核. 定义乘积空间 $\left(\Pi_{j=1}^{n}X_j,\Pi_{j=1}^{n}\mathcal{F}_j\right)$ 上的集函数 μ 为: 对任何可测集 $A\in\Pi_{j=1}^{n}\mathcal{F}_j$, 令

$$\begin{aligned}\mu(A)\coloneqq&\int_{X_1}\mu_1\left(\mathrm{d}x_1\right)\int_{X_2}K\left(x_1,\mathrm{d}x_2\right)\\&\cdots\int_{X_n}1_A\left(x_1,\cdots,x_n\right)K\left(x_1,\cdots,x_{n-1},\mathrm{d}x_n\right),\end{aligned}$$

那么我们有如下的结论:

定理 101. μ 是一个测度. 特别的, 对任意可测矩形 $\Pi_{j=1}^{n}A_j$,

$$\mu\left(\prod_{j=1}^{n}A_j\right)=\int_{A_1}\mu_1\left(\mathrm{d}x_1\right)\int_{A_2}K(x_1,\mathrm{d}x_2)\cdots\int_{A_n}K\left(x_1,\cdots,x_{n-1},\mathrm{d}x_n\right)\tag{5.4}$$

此外, 若 μ_1 是 σ 有限测度, 满足 (5.4) 的测度 μ 是唯一的, 且是 σ 有限的.

注. 下一节会用到, 若 μ_1 是概率测度, K 皆是概率核, 则 μ 也是概率测度.

定理 102. 对 $\left(\Pi_{j=1}^{n}X_j,\Pi_{j=1}^{n}\mathcal{F}_j,\mu\right)$ 上积分存在可测函数 f, 有累次积分公式

$$\begin{aligned}\int_{\Pi_{j=1}^{n}X_j}f\,\mathrm{d}\mu\coloneqq&\int_{X_1}\mu_1\left(\mathrm{d}x_1\right)\int_{X_2}K\left(x_1,\mathrm{d}x_2\right)\\&\cdots\int_{X_n}f\left(x_1,\cdots,x_n\right)K\left(x_1,\cdots,x_{n-1},\mathrm{d}x_n\right).\end{aligned}$$

5.4 可列维乘积空间上的概率测度

在概率论中, 我们经常要讨论任意有限多个试验 (不一定相互独立). 为了能在同一概率空间中考虑它们, 我们需要在无穷乘积可测空间上构造概率测度. 本节我们讨论一个在可列维乘积空间上建立概率测度的方法, 它为离散时间的随机过程存在性奠定了理论基础.

设 $(\Omega_j,\mathcal{F}_j)$, $j=1,2,\cdots$ 是一列可测空间. 为了下面书写方便, 我们将 $(\Pi_{j=1}^{k}\Omega_i,\Pi_{j=1}^{k}\mathcal{F}_i)$ 记为 $(\Omega_{(k)},\mathcal{F}_{(k)})$, 将 $(\Pi_j\Omega_j,\Pi_j\mathcal{F}_j)$ 记为 $(\Omega,\mathcal{F})$. Ω 到 $\Omega_{(n)}$ 的投影记为 $\pi_{(n)}$, 即

$$\pi_{(n)}\left(\omega_1,\omega_2,\cdots\right)=\left(\omega_1,\cdots,\omega_n\right).$$

令 $P(\omega_1,\cdots,\omega_{j-1},\mathrm{d}\omega_j)$ 是从 $(\Omega_{(j-1)},\mathcal{F}_{(j-1)})$ 到 $(\Omega_j,\mathcal{F}_j)$ 上的概率核.

设 $\mathbb{P}_1$ 是 $(\Omega_1,\mathcal{F}_1)$ 上的概率测度, 则由定理 101, 对任何 n, 我们定义了 n 维空间 $(\Omega_{(n)},\mathcal{F}_{(n)})$ 上的概率测度 $\mathbb{P}_n$ 为, 对 $A_{(n)}\in\mathcal{F}_{(n)}$,

$$\begin{aligned}\mathbb{P}_n(A_{(n)}) = &\int_{\Omega_1}\mathbb{P}_1(\mathrm{d}\omega_1)\int_{\Omega_2}P(\omega_1,\mathrm{d}\omega_2)\\ &\cdots\int_{\Omega_n}I_{A_{(n)}}(\omega_1,\cdots,\omega_n)P(\omega_1,\cdots,\omega_{n-1},\mathrm{d}\omega_n)\,.\end{aligned}$$

定理 103 (Tulcea 定理)**.** *存在 $(\Omega,\mathcal{F})$ 上的唯一概率测度 $\mathbb{P}$, 使得对所有正整数 n, 成立*

$$\mathbb{P}_{\pi_{(n)}}\equiv\mathbb{P}\circ\pi_{(n)}^{-1}=\mathbb{P}_n\,. \tag{5.5}$$

证明. 注意有限维可测柱集

$$\mathcal{A}=\bigcup_{n=1}^{\infty}\pi_{(n)}^{-1}\mathcal{F}_{(n)}$$

是生成 $\mathcal{F}$ 的域. 为了使 (5.5) 成立, 自然地在 $\mathcal{A}$ 上定义集函数 $\mathbb{P}$ 如下: 对 $\pi_{(n)}^{-1}(A_{(n)})\in\mathcal{A}$, 令

$$\mathbb{P}\left(\pi_{(n)}^{-1}\left(A_{(n)}\right)\right):=\mathbb{P}_n(A_{(n)})\,.$$

易知 $\mathbb{P}$ 是良定的, 且显然 $\mathbb{P}$ 非负, 在空集处为 0, 有限可加. 于是只需要证 $\mathbb{P}$ 在空集处上连续, 就得到 $\mathbb{P}$ 是 $\mathcal{A}$ 上测度. 利用测度延拓定理和唯一性, 就证明了 Tulcea 定理.

反证, 假设存在 $B_n\in\mathcal{A}$, $n\geq 1$, $B_n\downarrow\varnothing$, 使得 $\lim\mathbb{P}(B_n)>0$. 必要时在序列 $\{B_n\}$ 首项前添加若干项 Ω, 且在两个集合 B_n 与 B_{n+1} 之间适当重复若干项 B_n, 我们进一步假定 $B_n=\pi_{(n)}^{-1}(A_{(n)})$, 其中 $A_{(n)}\in\mathcal{F}_{(n)}$. 由 $B_{n+1}\subset B_n$, 我们有

$$A_{(n+1)}\subset A_{(n)}\times\Omega_{n+1}$$

此外, 对任意 $n>1$,

$$\mathbb{P}(B_n)=\int_{\Omega_1}g_n^{(1)}(\omega_1)\,\mathbb{P}_1(d\omega_1)$$

其中

$$g_n^{(1)}(\omega_1)=\int_{\Omega_2}P(\omega_1,d\omega_2)\cdots\int_{\Omega_n}I_{A_{(n)}}(\omega_1,\cdots\omega_n)\,P(\omega_1,\cdots,d\omega_n)$$

由于

$$I_{A_{(n+1)}}(\omega_1,\cdots,\omega_{n+1})\leq I_{A_{(n)}}(\omega_1,\cdots\omega_n)$$

得到 $g_n^{(1)}$ (逐点的) 单调递减, 设极限为 h_1. 由控制收敛定理

$$\int_{\Omega_1} h_1(\omega_1)\,\mathbb{P}_1(d\omega_1) = \lim_{n\to\infty} \mathbb{P}(B_n) > 0$$

于是存在 $\omega_1' \in \Omega_1$, 使得 $h_1(\omega_1') > 0$. 实际上, 必有是 $\omega_1' \in A_{(1)}$. 斗则对任何 $n > 1$ 有 $I_{A_{(n)}}(\omega_1', \omega_1, \cdots, \omega_n) = 0$ 从而 $g_n^{(1)}(\omega_1') = 0$, 得到 $h_1(\omega_1') = 0$.

现在设 $n > 2$, 则

$$g_n^{(1)}(\omega_1') = \int g_n^{(2)}(\omega_2)\,P(\omega_1', d\omega_2)$$

其中

$$\begin{aligned} g_n^{(2)}(\omega_2) = &\int_{\Omega_3} P(\omega_1', \omega_2, d\omega_3) \\ &\cdots \int_{\Omega_n} I_{A_{(n)}}(\omega_1', \omega_2, \cdots, \omega_n)\,P(\omega_1', \omega_2, \cdots, \omega_{n-1}, d\omega_n) \end{aligned}$$

如上所证, 可知 $g_n^{(2)}(\omega_2) \downarrow h_2(\omega_2)$. 由于 $g_n^{(1)}(\omega_1') \to h_1(\omega_1') > 0$, 故存在 $\omega_2' \in \Omega_2$ 使得 $h_2(\omega_2') > 0$, 同上可知, $(\omega_1', \omega_2') \in A_{(2)}$.

最后, 由归纳法得到点列 $\{\omega_1', \omega_2', \cdots\}$, 使得 $\omega_j' \in \Omega_j$ 且 $(\omega_1', \cdots, \omega_n') \in A_{(n)}$, 因此

$$(\omega_1', \omega_2', \cdots) \in \bigcap_{n=1}^{\infty} A_{(n)} = \varnothing\,.$$

矛盾! □

推论 104 (Kolmogorov 定理)**.** $(\Omega_j, \mathcal{F}_j, \mathbb{P}_j), k = 1, 2, \cdots$ 是一列概率空间. 则在 $(\Omega, \mathcal{F})$ 上有唯一的概率测度 $\mathbb{P}$, 使得对任意 $n \in \mathbb{N}$,

$$\mathbb{P}\left(\pi_{(n)}^{-1}\left(\prod_{k=1}^{n} A_k\right)\right) = \prod_{k=1}^{n} \mathbb{P}_k(A_k)$$

其中 $A_j \in \mathcal{F}_j, 1 \le j \le n$.

练习 19. $\{(\Omega_t, \mathcal{F}_t, \mathbb{P}_t), t \in T\}$ 是一族概率空间. 证明, 存在 $(\Pi_t\Omega_t, \Pi_t\mathcal{F}_t)$ 上的唯一概率测度 $\mathbb{P}$, 使得对 T 的非空有限子集 S, 成立

$$\mathbb{P}_{\pi_S} \equiv \mathbb{P} \circ {\pi_S}^{-1} = \prod_{t\in S} \mathbb{P}_t\,.$$

5.5 任意维乘积空间上的概率测度

本节将给出 Kolmogorov 相容性定理的一般形式, 为此我们先需要紧概率测度的概念, 它是 $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上概率测度的某种推广. 但是在此之前, 我们先来看一下经典的 Kolmogorov 相容性定理.

经典 Kolmogorov 相容性定理

设 T 是任意无穷集. 经典的 kolmogorov 定理是说, 给定 $\left(\mathbb{R}^T, \mathcal{B}_{\mathbb{R}}^T\right)$ 上的一族有限维概率测度族后, 如何在它上面产生概率测度?

定义 34. 称 T 上一族有限维概率测度族

$$\{\mathbb{P}_{t_1,\cdots,t_n} : t_1,\cdots,t_n \in T; n = 1,2,\cdots\} \tag{5.6}$$

是**相容**的, 若

1. 对任意正整数 n 和 $t_1,\cdots,t_n \in T$, $\mathbb{P}_{t_1,\cdots,t_n}$ 是 $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n)$ 上的概率测度,
2. 对任意正整数 n, $t_1,\cdots,t_n \in T$, $A_{t_1},\cdots A_{t_n} \in \mathcal{B}$, 和 $t_1,\cdots,t_n$ 的重排 $t(1),\cdots,t(n)$,
$$\mathbb{P}_{t_1,\cdots,t_n}\left(\prod_{i=1}^n A_{t_i}\right) = \mathbb{P}_{t_{(1)},\cdots,t_{(n)}}\left(\prod_{i=1}^n A_{t(i)}\right).$$
3. 对任意正整数 n, $t_1,\cdots,t_n,t_{n+1} \in T$, $A_{t_1},\cdots A_{t_n} \in \mathcal{B}$,
$$\mathbb{P}_{t_1,\cdots,t_n,t_{n+1}}\left(\prod_{i=1}^n A_{t_i} \times \mathbb{R}\right) = \mathbb{P}_{t_1,\cdots,t_n}\left(\prod_{i=1}^n A_{t_i}\right).$$

注. 直接定义相容性的概念为 $\{\mathbb{P}_S : S \subset T \text{ is finite }\}$, 其中 $\mathbb{P}_S$ 是 $\left(\mathbb{R}^S, \mathcal{B}_{\mathbb{R}}^S\right)$ 上概率测度, 且满足 $\mathbb{P}_{S_2} = \mathbb{P}_{S_1} \circ (\pi_{S_2}^{S_1})^{-1}, S_2 \subset S_1$. 这样的定义显然是更简洁的. 程士宏先生书中用上述定义, 我个人认为是为了将 $\left(\mathbb{R}^S, \mathcal{B}_{\mathbb{R}}^S\right)$ 换为我们更熟悉的 $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n)$. (当 $S = \{1,\cdots,n\}$ 时二者相同, 否则二者不同.) 为了二者之间的双射, 也就是对 S 中元素编号, 有 $n!$ 种, 于是对不同的编号定义了 $(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n)$ 上不同的测度, 但他们皆是由 $\mathbb{P}_S$ 诱导的像测度, 为了保证这一点, 需要额外要求条件 2.

定理 105. *对任意无穷集 T 上的相容的有限维概率测度族 (5.6), 存在 $(\mathbb{R}^T, \mathcal{B}(\mathbb{R})^T)$ 上唯一概率测度 $\mathbb{P}$, 对任意 T 的有限子集 S, 有*

$$\mathbb{P}_{\pi_S} \equiv \mathbb{P} \circ {\pi_S}^{-1} = \mathbb{P}_S.$$

5.5.1 紧类, 紧概率测度

定义 35. Ω 上的集类 $\mathcal{C}$ 称为**紧类**, 若对任何 $\mathcal{C}$ 中集列 $\{C_n\}$ 且 $\cap_n C_n = \varnothing$, 必然存在 $\mathbb{N}$ 的有限子集 S, 使得

$$\bigcap_{n\in S} C_n = \varnothing\,.$$

例 25. 可数集 Ω 的有限子集全体是其上的紧类.

例 26. $\mathbb{R}^n$ 中的紧集全体是紧类; 一般的, Hausdorff 空间的中的紧集全体也是紧类.

引理 106. *若 $\mathcal{C}$ 为紧类, 则 $\mathcal{C}_{\Sigma f}$ 为紧类, 其中 $\mathcal{C}_{\Sigma f}$ 是用有限不交并运算封闭 $\mathcal{C}$ 所得的集类.*

证明. 设 $D_n = \cup_{m=1}^{M_n} C_{n,m}$ 是 $\mathcal{C}_{\Sigma f}$ 中的集列, 满足任意 $p \geq 1$, $\cap_{n\leq p} D_n$ 非空. 令 $J = \{(m_n) : 1 \leq m_n \leq M_n \text{ for all } n\}$. 则

$$\bigcap_{n\leq p} D_n = \bigcup_{(m_n)\in J} \left(\bigcap_{n\leq p} C_{n,m_n} \right)\,.$$

令 $J_p = \{(m_n) \in J : \cap_{n\leq p} C_{n,m_n} \neq \varnothing\}$, 显然 $\{J_p\}$ 是非空单调递减集列. 我们证明 $\cap_p J_p$ 非空.

对任意 $p \geq 1$, 取 $(m_n^{(p)}) \in J_p$. 注意对固定的 n, $1 \leq m_n^{(p)} \leq M_n$ 对所有 p 成立, 因此有无穷多个 p, 使得 $m_n^{(p)}$ 取相同的值. 我们用归纳法可以构造出序列 $(l_n) \in \cap_p J_p$, 从而对任意 p, 使得

$$\bigcap_{n\leq p} C_{n,l_n} \neq \varnothing\,,$$

再由 $\mathcal{C}$ 为紧类, 知 $\cap_n C_{n,l_n}$ 非空, 从而 $\cap_n D_n$ 非空, 于是 $\mathcal{C}_{\Sigma f}$ 为紧类. □

引理 107. *设 μ 是域 $\mathcal{A}$ 上任何非负, 空集处为 0, 有限可加的集函数. 若存在紧类 $\mathcal{C} \subset \mathcal{A}$, 满足对任意 $A \in \mathcal{A}$ 和 $\epsilon > 0$, 存在 $C \in \mathcal{C}$ 满足 $C \subset A$ 且 $\mu(A\backslash C) < \epsilon$. 则 μ 是 $\mathcal{A}$ 上的测度.*

证明. 我们证明 μ 在空集处 "强上连续": 任何 $A_n \in \mathcal{A}$, 且 $A_n \downarrow \varnothing$, 必然有 $\mu(A_n) \to 0$. 现在任取 $\epsilon > 0$, 令 $\epsilon_n = \epsilon/3^n, n \in \mathbb{N}$.

取 $C_n \subset A_n$ 且 $\mu(A_n \backslash C_n) \leqslant \epsilon_n$, 则 $\bigcap_{n=1}^{\infty} C_n = \varnothing$. 于是存在正整数 m, 使得 $\bigcap_{n=1}^{m} C_n = \varnothing$.

$$\begin{aligned}\mu(A_m) &= \mu\left(A_m \backslash \bigcap_{n=1}^{m} C_n\right) = \mu\left(\bigcup_{n=1}^{m}(A_m \backslash C_n)\right) \\ &\leqslant \sum_{n=1}^{m} \mu\left(A_m \backslash C_n\right) \leqslant \sum_{n=1}^{m} \mu\left(A_n \backslash C_n\right) \leqslant \sum_{n=1}^{m} \epsilon_n < \epsilon.\end{aligned}$$

由 $\mu(A_n)$ 的单调性, ϵ 的任意性, 知 $\mu(A_n) \to 0$.

由此空集处的 “强上连续” 性, 容易证明 μ 是测度. □

注. 若加上 $\mu(\Omega) < \infty$ 的条件, 则上述引理对 μ 的要求可改为 “内正则条件” 的形式: 对任意 $A \in \mathcal{A}$,

$$\mu(A) = \sup\{\mu(C) : C \subset A\,, C \in \mathcal{C}\}\,.$$

定义 36. 对概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$, 称 $\mathbb{P}$ **紧概率测度**, 若存在紧类 $\mathcal{C} \subset \mathcal{F}$, 使得对任意 $A \in \mathcal{F}$, 成立

$$\mathbb{P}(A) = \sup\{\,\mathbb{P}(C) : C \subset A\,, C \in \mathcal{C}\}\,.$$

例 27. $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上的任何概率测度都是紧概率测度.

例 28. Ω 是一可数集, $\mathcal{F} = 2^{\Omega}$. 则 $(\Omega, \mathcal{F})$ 上的概率测度是紧概率测度.

5.5.2 相容性定理

T 是非空指标集. $\{(\Omega_t, \mathcal{F}_t), t \in T\}$ 是一族可测空间. 我们记乘积空间

$$(\Omega, \mathcal{F}) = (\prod_{t \in T} \Omega_t\,, \prod_{t \in T} \mathcal{F}_t)\ ,$$

对任何 $S \subset T$, 记

$$(\Omega_S, \mathcal{F}_S) = (\prod_{t \in S} \Omega_t\,, \prod_{t \in S} \mathcal{F}_t)\ ,$$

定理 108 (Kolmogorov 相容性定理)**.** *若对 T 的非空有限子集 S, 定义好了 $(\Omega_S, \mathcal{F}_S)$ 上的概率测度 $\mathbb{P}_S$, 满足*

1. 任何 $t \in T$, $\mathbb{P}_t$ 是 $(\Omega_t, \mathcal{F}_t)$ 上的紧测度.

2. 相容性条件: 对任意 T 的非空有限子集 $S_2 \subset S_1$, 有

$$\mathbb{P}_{S_2} = \mathbb{P}_{S_1} \circ (\pi_{S_2}^{S_1})^{-1} .$$

则存在 $(\Omega, \mathcal{F})$ 上的唯一概率测度 $\mathbb{P}$, 满足对任何 T 的非空有限子集 S, 有

$$\mathbb{P}_{\pi_S} \equiv \mathbb{P} \circ \pi_S^{-1} = \mathbb{P}_S \ . \tag{5.7}$$

证明. 回忆第一节中, 我们定义了有限维可测矩形柱集全体 $\mathcal{Q}$, 为

$$\mathcal{Q} = \left\{ \pi_S^{-1}(\prod_{t\in S} A_t) : A_t \in \mathcal{F}_t,\, t \in S\,; S \subset T \text{ is finite} \right\}.$$

且证明了 $\mathcal{Q}$ 是半代数, $\mathcal{F} = \sigma(\mathcal{Q})$. 为了找到 $(\Omega, \mathcal{F})$ 上满足 (5.7) 的概率测度, 先定义 $\mathcal{Q}$ 上的集函数 $\mathbb{P}$ 为

$$\mathbb{P}\big(\pi_S^{-1}(\prod_{t\in S} A_t)\big) = \mathbb{P}_S\big(\prod_{t\in S} A_t\big) .$$

由相容性条件知 $\mathbb{P}$ 是良定的. 若能证明 $\mathbb{P}$ 是 $\mathcal{Q}$ 上的 (概率) 测度, 则由测度延拓定理与概率测度的唯一性定理, 就证明了相容性定理.

显然 $\mathbb{P}$ 是 $\mathcal{Q}$ 上非负, 空集处取 0, 有限可加的集函数. 如果我们能找到 $\mathcal{Q}$ 的子集合系 $\mathcal{C}$ 是紧的, 且 $\mathcal{Q}$ 中任意集合的的概率可由 $\mathcal{C}$ 中集合逼近, 由引理 107 就证明了定理. 自然的想法是令

$$\mathcal{C} = \left\{ \pi_S^{-1}(\prod_{t\in S} C_t) : C_t \in \mathcal{C}_t,\, t \in S;\ S \subset T \text{ is finite} \right\}.$$

其中 $\mathcal{C}_t$ 是 $\mathcal{F}_t$ 中的紧类, 使得 $\mathbb{P}_t$ 是紧概率测度.

首先证明 $\mathcal{C}$ 是紧类. 注意 $\mathcal{C}$ 对有限交运算封闭, 因此只需要证明对任何 $A_n = \pi_{t_n}^{-1}(C_{t_n}), n \in \mathbb{N}$, 若 $\cap_n A_n = \varnothing$, 必存在正整数的有限子集 J 使得

$$\bigcap_{n\in J} A_n = \varnothing .$$

注意 t_n 之间可能有重复的, 我们令 $I = \{t_n : n \in \mathbb{N}\}$. 对任意 $t \in I$, 令

$$B_t = \bigcap_{n:\, t_n = t} C_{t_n} ,$$

于是

$$\bigcap_{n\in\mathbb{N}} A_n = \pi_I^{-1}\big(\prod_{t\in I} B_t\big) = \varnothing .$$

因此, 存在 $t' \in I$ 使得

$$B_{t'} = \bigcap_{n:\, t_n = t'} C_{t_n} = \varnothing\,.$$

由 $\mathcal{C}_{t'}$ 为紧, 故存在 $\{n : t_n = t'\}$ 的有限子集 J 使得 $\cap_{n\in J} C_{t_n} = \varnothing$. 于是

$$\bigcap_{n\in J} A_n = \bigcap_{n\in J} \pi_{t'}^{-1}(C_{t_n}) = \pi_{t'}^{-1}(\bigcap_{n\in J} C_{t_n}) = \varnothing\,.$$

其次, 由引理 106 和引理 107, 只需证明 $\mathcal{C}$ 中集合可任意逼近 $\mathcal{Q}$ 中集合, 因为将 $\mathbb{P}$ (唯一) 延拓到 $r(\mathcal{Q})$ 上后, $\mathcal{C}_{\Sigma f}$ 是 $r(\mathcal{Q})$ 中紧类, 且相应的 “内正则条件” 成立.

于是任取 $\pi_S^{-1}(\Pi_{t\in S} A_t) \in \mathcal{Q}$. 和 $\epsilon > 0$, 由 $\mathbb{P}_t$ 是 $(\Omega_t, \mathcal{F}_t)$ 上的紧测度, 存在 $C_t \in \mathcal{C}_t$, 使得 $C_t \subset A_t$ 且

$$\mathbb{P}_t(A_t) - \mathbb{P}_t(C_t) \leqslant \frac{\epsilon}{|S|}\,.$$

其中 $|S|$ 表示 S 中元素个数. 则可见 $\Pi_{t\in S} C_t \subset \Pi_{t\in S} A_t$, 同时注意到 $\Pi_{t\in S} A_t \backslash \Pi_{t\in S} C_t \subset \cup_{t\in S} \pi_t^{-1}(A_t \backslash C_t)$, 因此

$$\mathbb{P}(\Pi_{t\in S} A_t) - \mathbb{P}(\Pi_{t\in S} C_t) \leqslant \sum_{t\in S} \mathbb{P}_t(A_t) - \mathbb{P}_t(C_t) \leqslant \epsilon\,.$$

于是定理得证. □

附录 A

A.1 L^p 对偶空间的表示 *

本节我们给出 L^p 对偶空间的表示的详细证明. 首先我们来看 $p>1$ 的情形.

定理 109. *设 $p,q\in(1,\infty)$ 满足 $\frac{1}{p}+\frac{1}{q}=1$. $(X,\mathcal{F},\mu)$ 是测度空间. 则 L^{p*} 与 L^q 等距同构: 若 $g\in L^q$, 令*

$$T_g(f)=\mu(fg),\quad \mathit{for}\ f\in L^p\,,$$

则 $T_g\in L^{p*}$. 且映射 $g\mapsto T_g$ 为 L^q 到 L^{p*} 的双射, 满足 $\|g\|_q=\|T_g\|$.

证明. 第一步. 设 $g\in L^q$, 由 Hölder 不等式知, 上面定义的 T_g 为 L^p 上连续线性泛函, 且 $\|T_g\|\leqslant\|g\|_q$. 往证 $\|T_g\|\geqslant\|g\|_q$. 不失一般性设 $\|g\|_q>0$, 令

$$f=|g|^{q-1}\mathrm{sgn}(g)\,,$$

其中 sgn 为符号函数, 可见 $fg=|g|^q$ 且 $|f|^p=|g|^{p(q-1)}=|g|^q$, 于是 $\|f\|_p=\|g\|_q^{q-1}$.

$$T_g(f)=\mu\left(|g|^q\right)=\|g\|_q^q=\|g\|_q\|f\|_p\,.$$

从而 $\|T_g\|\geqslant\|g\|_q$. 于是得到 $g\mapsto T_g$ 为等距嵌入. 下面只需证明这是一满射.

第二步. 我们假设 $(X,\mathcal{F},\mu)$ 是有限测度空间, 注意此时对任意 $A\in\mathcal{F}$, $I_A\in L^p$. 任意取定 $T\in L^{p*}$. 定义 $(X,\mathcal{F})$ 上的集函数 ν 为

$$\nu(A)=T(I_A)\,,\text{for } A\in\mathcal{F}\,.$$

容易验证 ν 是 $(X,\mathcal{F},\mu)$ 上的有限符号测度. 由 R-N 定理, 存在 $(X,\mathcal{F},\mu)$ 上可积函数 g, 使得

$$\nu(A)=T(I_A)=\int gI_A\,\mathrm{d}\mu\,,\ \text{for } A\in\mathcal{F}\,.$$

由典型方法, 易证对任意 $f\in L^p$, fg 可积且

$$T(f)=\int fg\,\mathrm{d}\mu\,.$$

下面只需要证明 $g\in L^q$. 我们先做截断: 给定正整数 n, 令 $B_n=\{|g|\leqslant n\}$, 则 $g_n\coloneqq gI_{B_n}\in L^q$, 于是对任意 $f\in L^p$,

$$T_{g_n}(f)=\mu(g_nf)=\mu(fgI_{B_n})=T(fI_{B_n})\,.$$

得到

$$\|T_{g_n}\|=\|g_n\|_q\leqslant\|T\|\,.$$

注意 $|g_n|\uparrow|g|$, 由单调收敛定理得 $\|g\|_q\leqslant\|T\|$.

第三步. 对一般的情况, 令 $\mathcal{E}=\{E\in\mathcal{F}:\mu(E)<\infty\}$. 对每个 $E\in\mathcal{E}$, 令 $\mu|_E(\cdot)\coloneqq\mu(\cdot\cap E)$, 则 $(X,\mathcal{F},\mu|_E)$ 为有限测度空间. 记 $L^p(\mu|_E)=L^p(X,\mathcal{F},\mu|_E)$. 定义

$$T_E(f)=T(fI_E)\ \text{ for }f\in L^p(\mu|_E)\,.$$

则 $T_E\in L^p(\mu|_E)^*$ 且 $\|T_E\|\leqslant\|T\|$. 于是存在 $g_E\in L^q(\mu|_E)$, 使得 $\|g_E\|_q=\|T_E\|$, 且对 $f\in L^p(\mu|_E)$

$$T_E(f)=T(fI_E)=\int fg_E\,\mathrm{d}\mu|_E=\int fI_Eg_E\,\mathrm{d}\mu\,.$$

注意若 $E,D\in\mathcal{E}$, 且 $E\cap D=\varnothing$, 则有

$$g_E=g_D=g_{D\cap E}\ \text{ a.e. on }\ D\cap E\,.$$

我们定义函数 g, 使得在 $E\in\mathcal{E}$ 上, $g=g_E$; 在 $\cup_{E\in\mathcal{E}}E$ 的补集上, $g=0$. 现在出现了一个新的问题, g 是可测函数吗?

令 $\sigma\coloneqq\sup_{E\in\mathcal{E}}\|g_E\|_q\leqslant\|T\|$, 同时 $\|g_D\|\leqslant\|g_E\|_q$ 若 $D\subset E$, 我们可以选取 $\mathcal{E}$ 中的递增集列 $\{E_n\}$, 使得

$$\|g_{E_n}\|_q\uparrow\sigma=\sup_{E\in\mathcal{E}}\|g_E\|_q\,.$$

令 $G=\cup_nE_n$. 若 $E\in\mathcal{E}$ 且 $E\cap G=\varnothing$, 则

$$\|g_{E\cup E_n}\|_q^q=\|g_E\|_q^q+\|g_{E_n}\|_q^q\uparrow\|g_E\|_q^q+\sigma^q$$

从而必有 $\|g_E\|_q=0$, 故不妨令 $g_E=0$. 于是在 G^c 上 g 恒为 0. 因此实际上 $g=\lim_n g_{E_n}I_{E_n}$. 特别的, $g\in L^q$ 且 $\|g\|_q=\sigma$.

若 $f \in L^p$, 则集合 $\{|f| > 0\}$ σ 有限, 于是存在递增集列 $\{D_n\} \subset \mathcal{E}$, 使得 $\{|f| > 0\} = \cup_n D_n$. 于是 $f I_{D_n} \xrightarrow{L^p} f$,

$$T(f) = \lim_{n\to\infty} T(f I_{D_n}) = \lim_{n\to\infty} \int f I_{D_n} g_{D_n} \,\mathrm{d}\mu = \lim_{n\to\infty} \int f I_{D_n} g \,\mathrm{d}\mu = \int f g \,\mathrm{d}\mu$$

因此 $T = T_g$. □

仿照上面的定理, 容易证明下面 $p = 1$ 的情形:

定理 110. $(X, \mathcal{F}, \mu)$ *是 σ 有限测度空间. 则 L^{1*} 与 L^∞ 等距同构: 若 $g \in L^\infty$, 令*

$$T_g(f) = \mu(fg), \quad \textit{for } f \in L^1,$$

则 $T_g \in L^{1*}$. 且映射 $g \mapsto T_g$ 为 L^∞ 到 L^{1*} 的双射, 满足 $\|g\|_\infty = \|T_g\|$.

注. 上面定理中的 σ 有限的假设不能去掉. 考虑测度空间 $([0,1], \mathcal{B}([0,1]), \mu)$, 其中

$$\mu(A) = \begin{cases} \lambda(A), & 0 \notin A. \\ \infty, & 0 \in A. \end{cases}$$

上面 λ 表示 Lebesgue 测度. 测度 μ 在 0 处又一个无穷的原子, 因此不是 σ 有限的. 令 $g = I_{\{0\}}$. 则 $\phi \in L^\infty(\mu)$ 且 $\|\phi\|_\infty = 1$. 我们断言 $T_\phi = 0$, 所以 $\|M_\phi\| = 0 < 1 = \|\phi\|_\infty$.

为证明上面断言, 注意任何 $f \in L^p(\mu)$ 必然有 $f(0) = 0$, 于是 $M_\phi f = f(0) I_{\{0\}} = 0$.