Contents

Part I

测度论

1

可测空间

1.1 可测集

定义 1.1. 集合 E 上的 σ-域 $\mathcal{A}$ 指的是 E 的一个子集族，其满足下面的性质：

1. $E \in \mathcal{A}$；
2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$；
3. 如果一列子集 $A_n \in \mathcal{A}$，那么 $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

$\mathcal{A}$ 的元素被称为**可测集**, $(E,\mathcal{A})$ 被称为**可测空间**. 根据定义, 我们很容易得出下面的结果:

- $\emptyset = E^c \in \mathcal{A}$.
- 如果一列子集 $A_n \in \mathcal{A}$, 那么

$$\bigcap_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}} A_n\right)^c \in \mathcal{A}.$$

- $\mathcal{A}$ 对有限并和有限交也是封闭的, 只需要从某一项 A_n 开始全部取空集即可.

例 1.2. 根据可测集的定义, 很容易构造出一些最简单的例子:

1. $\mathcal{A} = \mathcal{P}(E)$, 当 E 是有限集或者可数集的时候我们通常会使用这样的 σ-域, 其他情况则很少使用.
2. $\mathcal{A} = \{\emptyset, E\}$, 平凡 σ-域.
3. E 的所有至多可数的子集以及所有补集至多可数的子集构成 E 上的一个 σ-域.

为了产生更多的例子, 我们注意到 E 上任意 σ-域的交集仍然是 σ-域, 这导出了下面的定义.

定义 1.3. 令 $\mathcal{C}$ 是 $\mathcal{P}(E)$ 的子集，E 上包含 $\mathcal{C}$ 的最小的 σ-域被记为 $\sigma(\mathcal{C})$，不难看出其是所有包含 $\mathcal{C}$ 的 σ-域的交集. 我们称 $\sigma(\mathcal{C})$ 是由 $\mathcal{C}$ 生成的 σ-域.

定义 1.4. 设 $(E,\mathcal{O})$ 是拓扑空间，所有开集 $\mathcal{O}$ 生成的 σ-域 $\sigma(\mathcal{O})$ 被称为 E 上的 Borel σ-域，记为 $\mathcal{B}(E)$.

E 上的 Borel σ-域是包含所有开集的最小的 σ-域. $\mathcal{B}(E)$ 的元素被称为 E 的 **Borel 子集**. 显然, E 中的闭集也都是 Borel 子集.

例 1.5 (ℝ 上的 Borel σ-域)**.** 记 $\mathcal{C}_1$ 为 ℝ 中开区间的集合:

$$\mathcal{C}_1 = \{(a,b) \mid a,b \in \mathbb{R}, a < b\},$$

显然有 $\mathcal{C}_1 \subseteq \mathcal{B}(\mathbb{R})$, 于是 $\sigma(\mathcal{C}_1) \subseteq \mathcal{B}(\mathbb{R})$. 下面我们说明 $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C}_1)$. 我们不加证明地使用一个结论 (Lindelöf 定理): ℝ 的任意开子集 U 都是开区间的可数并. 那么根据 σ-域的定义, 任意开区间都在 $\sigma(\mathcal{C}_1)$ 中, 故 $\mathcal{B}(\mathbb{R}) \subseteq \sigma(\mathcal{C}_1)$. 这表明 $\mathcal{B}(\mathbb{R})$ 可以由所有开区间生成.

此外, 如果注意到

$$(a,b) = (-\infty, b) \cap (-\infty, a)^c,$$

还可以证明 $\mathcal{B}(\mathbb{R})$ 由 $\mathcal{C}_2$ 生成, 其中

$$\mathcal{C}_2 = \{(-\infty, a) \mid a \in \mathbb{R}\}.$$

在后文中, 每当我们考虑拓扑空间 (例如 ℝ 或者 $\mathbb{R}^d$) 时, 除非有特别说明, 否则我们总是假设它们配备 Borel σ-域.

下一个非常重要的 σ-域是乘积 σ-域.

定义 1.6. 令 $(E_1, \mathcal{A}_1)$ 和 $(E_2, \mathcal{A}_2)$ 是可测空间，定义 $E_1 \times E_2$ 上的 σ-域 $\mathcal{A}_1 \otimes \mathcal{A}_2$ 为

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma\big(\{A_1 \times A_2 \mid A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}\big).$$

引理 1.7. 设 E 和 F 是可分 (有可数的稠密子集) 的拓扑空间，$E \times F$ 配备积拓扑，那么 $\mathcal{B}(E \times F) = \mathcal{B}(E) \otimes \mathcal{B}(F)$.

1.2 正测度

令 $(E, \mathcal{A})$ 是可测空间.

定义 1.8. $(E, \mathcal{A})$ 上的正测度指的是一个映射 $\mu : \mathcal{A} \to [0, \infty]$，其满足下面的性质：

1. $\mu(\emptyset) = 0$；
2. (σ-可加性) 对于任意可数个不相交的可测集序列 $(A_n)_{n\in\mathbb{N}}$，有

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

此时，三元组 $(E, \mathcal{A}, \mu)$ 被称为**测度空间**. 值 $\mu(E)$ 被称为测度 μ 的总质量.

需要注意的是, 我们允许 μ 的值为 $+\infty$, 此时级数 $\sum_{n\in\mathbb{N}} \mu(A_n)$ 作为正向级数在 $[0, \infty]$ 中总是有意义的. 根据 σ-可加性, 如果我们令 $n > n_0$ 开始 $A_n = \emptyset$, 便可以得到有限可加性.

命题 1.9 (测度的性质). 根据定义，测度 μ 满足下面的性质：

1. 如果 $A \subseteq B$，那么 $\mu(A) \le \mu(B)$. 此外，如果还满足 $\mu(A) < \infty$，那么

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

2. 如果 $A, B \in \mathcal{A}$, 那么
$$\mu(A)+\mu(B)=\mu(A\cup B)+\mu(A\cap B).$$
3. 如果 $A_n \in \mathcal{A}$ 且 $A_n \subseteq A_{n+1}$, 那么
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)=\lim_{n\to\infty}\mu(A_n).$$
4. 如果 $B_n \in \mathcal{A}$ 且 $B_{n+1} \subseteq B_n$, $\mu(B_1)<\infty$, 那么
$$\mu\left(\bigcap_{n\in\mathbb{N}} B_n\right)=\lim_{n\to\infty}\mu(B_n).$$
5. 如果 $A_n \in \mathcal{A}$, 那么
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)\le\sum_{n\in\mathbb{N}}\mu(A_n).$$

Proof. (1) 若 $A \subseteq B$, 那么 $B = A\bigcup(B\setminus A)$ 是无交并, 所以
$$\mu(B)=\mu(A)+\mu(B\setminus A)\ge\mu(A).$$

(2) 若 $\mu(A), \mu(B)$ 中有至少一个为无穷, 那么根据 (1), $\mu(A\cup B)$ 为无穷, 所以结论成立. 下面假设 $\mu(A), \mu(B)$ 均有限, 记 $C = A\cap B$, 那么 $A\cup B = (A\setminus C)\cup C\cup(B\setminus C)$ 是无交并, 所以
$$\mu(A\cup B)=\mu(A\setminus C)+\mu(C)+\mu(B\setminus C)=\mu(A)+\mu(B)-\mu(C),$$
结论 (2) 成立.

(3) 令 $C_1 = A_1$, 对于 $n \ge 2$ 的时候, 令
$$C_n = A_n \setminus A_{n-1},$$
那么 $A_n = \bigcup_{k\le n} C_k$ 是无交并, 所以
$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right)=\mu\left(\bigcup_{n\in\mathbb{N}} C_n\right)=\sum_{n\in\mathbb{N}}\mu(C_n)=\lim_{n\to\infty}\sum_{k=1}^{n}\mu(C_k)=\lim_{n\to\infty}\mu(A_n).$$

(4) 令 $A_n = B_1\setminus B_n$, 那么 $A_n \subseteq A_{n+1}$, 此时
$$\mu\left(\bigcap_{n\in\mathbb{N}} B_n\right)=\mu(B_1)-\mu\left(B_1\setminus\bigcap_{n\in\mathbb{N}} B_n\right)=\mu(B_1)-\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right),$$
再根据 (3), 就有
$$\mu\left(\bigcap_{n\in\mathbb{N}} B_n\right)=\mu(B_1)-\lim_{n\to\infty}\mu(A_n)=\lim_{n\to\infty}\mu(B_1\setminus A_n)=\lim_{n\to\infty}\mu(B_n).$$

(5) 令 $C_1 = A_1$, 对于 $n \geq 2$ 的时候, 令

$$C_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k,$$

那么 C_n 之间互不相交, 所以

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \mu\Big(\bigcup_{n\in\mathbb{N}} C_n\Big) = \sum_{n\in\mathbb{N}} \mu(C_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n).$$ □

例 1.10 (常见的测度)**.**

1. 令 $E = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\mathbb{N})$, 定义**计数测度**为

$$\mu(A) = \mathrm{card}(A).$$

2. 如果 A 是 E 的子集, 定义 A 的指示函数 $\mathbf{1}_A : E \to \{0, 1\}$ 为

$$\mathbf{1}_A(x) = \begin{cases} 1 & x \in A, \\ 0 & x \notin A. \end{cases}$$

令 $(E, \mathcal{A})$ 是可测空间, 固定 $x \in E$. 对于每个 $A \in \mathcal{A}$, 令 $\delta_x(A) = \mathbf{1}_A(x)$, 这给出了 $(E, \mathcal{A})$ 上的一个测度, 被称为 ***x* 处的 Dirac 测度**.

3. 可以证明, 在 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上存在唯一的正测度 λ 使得: 对于每个开区间 $[a, b]$, 有 $\lambda\big([a, b]\big) = b - a$. 这个测度 λ 被称为 **Lebesgue 测度**.

如果 μ 是 $(E, \mathcal{A})$ 上的正测度, $C \in \mathcal{A}$, 那么可以定义 μ 在 C 上的**限制** ν 为:

$$\nu(A) = \mu(A \cap C), \quad \forall A \in \mathcal{A}.$$

不难验证 ν 还是 $(E, \mathcal{A})$ 上的正测度.

定义 1.11.

- 如果 $\mu(E) < \infty$，那么我们说测度 μ 是**有限的**.
- 如果 $\mu(E) = 1$，那么我们说测度 μ 是**概率测度**，$(E, \mathcal{A}, \mu)$ 是**概率空间**.
- 如果存在一列可测集 $(E_n)_{n\in\mathbb{N}}$ 使得 $E = \bigcup_n E_n$ 以及每个 $\mu(E_n) < \infty$，那么我们说测度 μ 是 **σ-有限的**.
- 如果 $x \in E$ 使得单点集 $\{x\} \in \mathcal{A}$ 并且 $\mu(\{x\}) > 0$，那么我们说 x 是测度 μ 的一个**原子**.
- 如果测度 μ 没有原子，那么我们说 μ 是**扩散测度**.

如果 $(A_n)_{n\in\mathbb{N}}$ 是一列可测集, 类比数列的上下极限, 我们可以定义集合列的上下极限分别为:

$$\limsup A_n = \bigcap_{n=1}^{\infty}\Big(\bigcup_{k=n}^{\infty} A_k\Big), \quad \liminf A_n = \bigcup_{n=1}^{\infty}\Big(\bigcap_{k=n}^{\infty} A_k\Big).$$

注意到对于任意 m, 都有

$$\bigcup_{n=1}^{m}\left(\bigcap_{k=n}^{\infty} A_k\right)=\bigcap_{k=m}^{\infty} A_k,\quad \bigcap_{n=1}^{m}\left(\bigcup_{k=n}^{\infty} A_k\right)=\bigcup_{k=m}^{\infty} A_k,$$

所以显然有 $\liminf A_n \subseteq \limsup A_n$.

引理 1.12. 令 μ 是 $(E,\mathcal{A})$ 上的测度，那么

$$\mu(\liminf A_n)\le \liminf \mu(A_n).$$

如果 μ 是有限测度，或者更一般地，$\mu\left(\bigcup_{n=1}^{\infty} A_n\right)<\infty$，那么

$$\mu(\limsup A_n)\ge \limsup \mu(A_n).$$

Proof. 对于任意的 n, 有

$$\mu\left(\bigcap_{k=n}^{\infty} A_k\right)\le \inf_{k\ge n}\mu(A_k),$$

所以

$$\mu(\liminf A_n)=\lim_{n\to\infty}\mu\left(\bigcap_{k=n}^{\infty} A_k\right)\le \lim_{n\to\infty}\inf_{k\ge n}\mu(A_k)=\liminf \mu(A_n).$$

第二个结论同理. □

1.3 可测函数

定义 1.13. 令 $(E,\mathcal{A})$ 和 $(F,\mathcal{B})$ 是两个可测空间，如果映射 $f:E\to F$ 满足：

$$\forall B\in\mathcal{B},\ f^{-1}(B)\in\mathcal{A},$$

那么我们说 f 是**可测映射**. 当 E,F 是两个配备了 Borel σ-域的拓扑空间时，我们说 f 是 **Borel 可测的**.

显然, 可测映射的复合是可测映射.

命题 1.14. 令 $(E,\mathcal{A})$ 和 $(F,\mathcal{B})$ 是两个可测空间，映射 $f:E\to F$. f 可测当且仅当对于某个生成 $\mathcal{B}$ 的子集族 $\mathcal{C}$ (即 $\mathcal{B}=\sigma(\mathcal{C})$)，有 $f^{-1}(B)\in\mathcal{A}$ $(\forall B\in\mathcal{C})$.

Proof. 只需证明充分性. 记

$$\mathcal{G}=\{B\in\mathcal{B}\mid f^{-1}(B)\in\mathcal{A}\},$$

直接验证可知 $\mathcal{G}$ 是一个 σ-域, 又因为 $\mathcal{C}\subseteq\mathcal{G}$, 所以 $\mathcal{B}=\sigma(\mathcal{C})\subseteq\mathcal{G}\subseteq\mathcal{B}$, 所以 $\mathcal{G}=\mathcal{B}$, 这就表明 f 是可测的. □

例 1.15. 若 $(F,\mathcal{B})=(\mathbb{R},\mathcal{B}(\mathbb{R}))$, 要证明 f 是可测的, 只需说明集合 $f^{-1}((a,b))$ 是可测的, 或者 $f^{-1}((-\infty,a))$ 是可测的.

推论 1.16. 设 E,F 是两个配备 Borel σ-域的拓扑空间，那么连续映射 $f:E\to F$ 都是可测的.

Part II

概率论

2

概率论基础

2.1 一般定义

2.1.1 概率空间

令 $(\Omega, \mathcal{A})$ 是可测空间, $\mathbb{P}$ 是 $(\Omega, \mathcal{A})$ 上的概率测度, 我们说 $(\Omega, \mathcal{A}, \mathbb{P})$ 是**概率空间**.

因此, 概率空间是测度空间的一个特例. 然而, 概率论的观点与测度论有很大不同. 在概率论中, 我们的目标是一个"随机实验"的数学模型:

- Ω 表示实验的所有可能的结果的集合.
- $\mathcal{A}$ 是所有"事件"的集合. 这里的事件指的是 Ω 的一个子集, 其概率可以被计算 (也就是可测集). 我们应当把事件 A 视为满足某一属性的所有 $\omega \in \Omega$ 构成的子集.
- 对于每个 $A \in \mathcal{A}$, $\mathbb{P}(A)$ 表示事件 A 发生的概率.

当然, 一个自然的疑问是, 为什么需要考虑事件域 $\mathcal{A}$? 换句话说, 为什么不能对 Ω 的任意子集都计算一个概率? 原因在于, 一般不可能在 Ω 的幂集 $\mathcal{P}(\Omega)$ 上定义我们感兴趣的概率测度 (除开 Ω 是可数集这一简单情况). 例如, 取 $\Omega = [0, 1]$, 配备 Borel σ-域和 Lebesgue 测度, 但是, 可以证明不可能将 Lebesgue 测度扩展到 $[0, 1]$ 的任意子集上使得其仍然满足测度的定义.

例 2.1. 一些常见的概率模型.

1. 考虑扔两次骰子这一实验, 那么
$$\Omega = \{1, 2, \ldots, 6\}^2, \quad \mathcal{A} = \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \frac{\mathrm{card}(A)}{36}.$$
这里概率 $\mathbb{P}$ 的选取意味着让所有结果都有相同的概率. 更一般地, 如果 Ω 是有限集, $\mathcal{A} = \mathcal{P}(\Omega)$, 概率测度 $\mathbb{P}(\{\omega\}) = 1/\mathrm{card}(\Omega)$ 被称为 Ω 上的**均匀概率测度**.
2. 现在我们考虑实验: 扔骰子, 直到出现 6 为止. 由于得到 6 所需的投掷次数是无界的 (即使你扔了 1000 次骰子, 仍有可能没有得到 6), 所以 Ω 的正确选择是想象我们扔了无限次骰子:
$$\Omega = \{1, 2, \ldots, 6\}^{\mathbb{N}}.$$
Ω 上的 σ-域 $\mathcal{A}$ 被定义为包含形如
$$\{(\omega_1, \omega_2, \ldots, \omega_n) \in \Omega \mid \}$$

2.2 随机变量

在本章的剩余部分，我们都考虑一个概率空间 $(\Omega, \mathcal{A}, \mathbb{P})$，并且所有随机变量都将在这个概率空间上定义.

定义 2.2. 令 $(E, \mathcal{E})$ 是可测空间，值在 E 中的**随机变量**指的是一个可测映射 $X : \Omega \to E$.