



随机过程

作者：Huang

时间：May 19, 2025

目录

第 1 章 预备知识

1.1 σ 代数

集合 \mathbb{R} 表示实数集, \mathbb{Q} 表示有理数集, \mathbb{Z} 表示整数集, \mathbb{N} 表示自然数集, 下标 $+$ 表示非负元素全体, 如 \mathbb{R}_+ 表示非负实数集, 其他类似。

用 Ω 表示抽象空间, 其上的任一元素用 ω 表示。

集类: Ω 中子集构成的集合, 以后用花体字母 $\mathcal{A}, \mathcal{B}, \mathcal{C}$ 来表示。例 $\Omega := \{1, 2\}$ 。则 Ω 有不同的集合类。比如:

$\mathcal{A} := \{\{1\}, \{2\}\}$ 为一个集类。 $\mathcal{B} := \{\{1\}, \{2\}, \{1, 2\}\}$, $\mathcal{C} := \{\{1\}, \{2\}, \{1, 2\}, \emptyset\}$ 均为一个集类。

2^Ω 表示 Ω 中的所有子集构成的集类, 称为幂集, Ω 的一个子集类是指 2^Ω 的一个子集。

如: $\Omega := \{1, 2\}$, 则 $2^\Omega := \{\{1\}, \{2\}, \{1, 2\}, \emptyset\}$ 。

我们说一个子集类对集合的某种运算封闭, 是指此子集类中的集合经过此种运算后得到的集合还在此子集类内。

集合常见运算:

- 余运算: A^c
 - 交: $A \cap B$
 - 并: $A \cup B$
 - 可列并: $\bigcup_{j=1}^{\infty} A_j$
- 例 $\Omega := \{1, 2\}$ 。 $\mathcal{A} := \{\{1\}, \{2\}\}$ (余运算封闭, 交与并不封闭)
 $\mathcal{C} := \{\{1\}, \{2\}, \{1, 2\}, \emptyset\}$ (余运算、交与并均封闭)

定义 1.1.1

Ω 中的非空子集类 \mathcal{A} 称为代数, 若:

1. 若 $A \in \mathcal{A}$, 则 $A^c \in \mathcal{A}$;
2. 若 $A \in \mathcal{A}, B \in \mathcal{A}$, 则 $A \cup B \in \mathcal{A}$ 。

定理 1.1.1

设 \mathcal{A} 是代数, 则

1. $\Omega \in \mathcal{A}, \emptyset \in \mathcal{A}$;
2. 若 $A, B \in \mathcal{A}$, 则 $A \cap B \in \mathcal{A}$;
3. 若 $A_j \in \mathcal{A}, 1 \leq j \leq n$, 则 $\bigcup_{j=1}^n A_j \in \mathcal{A}, \bigcap_{j=1}^n A_j \in \mathcal{A}$ 。

证明 用代数的定义。由 \mathcal{A} 非空。任取 $A \in \mathcal{A}$, 则 $A^c \in \mathcal{A}$ 。 $\Omega = A \cup A^c \in \mathcal{A}$, 故 $\emptyset = \Omega^c \in \mathcal{A}$ 。

另一方面, 任取 $A, B \in \mathcal{A}$, 则 $A^c, B^c \in \mathcal{A}$ 。则 $A \cap B = (A^c \cup B^c)^c$ 。

有归纳法, 若 $A_j \in \mathcal{A}, 1 \leq j \leq n$, 则 $\bigcup_{j=1}^n A_j \in \mathcal{A}, \bigcap_{j=1}^n A_j \in \mathcal{A}$ 。

定义 1.1.2

设 $\{A_n, n \geq 1\}$ 为一集合序列。令

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k; \quad \liminf_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$

分别称其为 $\{A_n\}$ 的上极限和下极限 (上极限有时也记为 $\{A_n, i.o.\}$) 显然有

$$\limsup_{n \rightarrow \infty} A_n = \{w | w \text{ 属于无穷多个 } A_n\} = \{w | \forall n \in \mathbb{N}, \exists k \geq n, \text{ 使 } w \in A_k\}$$

$$\liminf_{n \rightarrow \infty} A_n = \{w | w \text{ 至多不属于有限多个 } A_n\} = \{w | \exists n \in \mathbb{N}, \forall k \geq n, \text{ 有 } w \in A_k\}$$

从而恒有 $\liminf_{n \rightarrow \infty} A_n \subseteq \limsup_{n \rightarrow \infty} A_n$ 。



若 $\liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n$, 则称 $\{A_n\}$ 的极限存在, 并用 $\lim_{n \rightarrow \infty} A_n$ 表示, 即令 $\lim_{n \rightarrow \infty} A_n = \liminf_{n \rightarrow \infty} A_n = \limsup_{n \rightarrow \infty} A_n$ 。

特别地, 若对每个 n , 有 $A_n \subset A_{n+1}$ (相应地, $A_n \supset A_{n+1}$), 则称 $\{A_n\}$ 为单调增 (相应地, 单调降)。对单调增或单调降序列 $\{A_n\}$, 我们分别令 $A = \bigcup_n A_n$ 或 $A = \bigcap_n A_n$, 称 A 为 $\{A_n\}$ 的极限, 通常记为 $A_n \uparrow A$ 或 $A_n \downarrow A$ 。

定义 1.1.3 (σ -代数)

Ω 中的非空子集类 \mathcal{F} 称为是 σ -代数, 若:

1. 若 $A \in \mathcal{F}$, 则 $A^c \in \mathcal{F}$;
2. 若 $A_n \in \mathcal{F}$, $n \geq 1$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ 。此时称 (Ω, \mathcal{F}) 为可测空间。



1. 2^Ω 是 σ -代数; (想想为什么?)
2. $\{\emptyset, \Omega, A, A^c\}$ 也是一个 σ -代数。

命题 1.1.1

设 I 为指标集 (I 有限或无限), 对任意的 i , \mathcal{F}_i 为 σ -代数, 则 $\bigcap_{i \in I} \mathcal{F}_i$ 为 σ -代数。(自行验证)



证明 $\emptyset \in \bigcap_{i \in I} \mathcal{F}_i$, 故 $\bigcap_{i \in I} \mathcal{F}_i$ 非空。

若 $A \in \bigcap_{i \in I} \mathcal{F}_i$, 则对任意的 $i \in I$, $A \in \mathcal{F}_i$, 故 $A^c \in \mathcal{F}_i$ 。因此, $A^c \in \bigcap_{i \in I} \mathcal{F}_i$ 。

若对任意的 n , $A_n \in \bigcap_{i \in I} \mathcal{F}_i$, 则 $\bigcup_n A_n \in \mathcal{F}_i$ 。因此, $\bigcup_n A_n \in \bigcap_{i \in I} \mathcal{F}_i$ 。

定理 1.1.2

设 $\mathcal{C} \subseteq 2^\Omega$ 为一非空集类, 则存在 σ -代数 \mathcal{F}_0 , 使得

1. $\mathcal{C} \subseteq \mathcal{F}_0$ 。
2. 若 \mathcal{F} 为 σ -代数且 $\mathcal{C} \subseteq \mathcal{F}$ 。则 $\mathcal{F}_0 \subseteq \mathcal{F}$ 。即 \mathcal{F}_0 为包含 \mathcal{C} 的最小 σ -代数, 称为由 \mathcal{C} 生成的 σ -代数, 记为 $\sigma(\mathcal{C})$ 。

(思考如何证明存在这样的最小 σ -代数)。



证明 设 $C(\mathcal{C})$ 是包含 \mathcal{C} 的 σ -代数全体构成的集类, 则 $2^\Omega \in C(\mathcal{C})$, 所以 $C(\mathcal{C})$ 是非空的。又因为任意个包含 \mathcal{C} 的 σ -代数的交仍是一个包含 \mathcal{C} 的 σ -代数 (按 σ -代数的定义逐条验证), 所以若取

$$\mathcal{F}_0 = \bigcap_{\mathcal{B} \in C(\mathcal{C})} \mathcal{B},$$

则 \mathcal{F}_0 就是所要求的。

定义 1.1.4

1. 数直线 \mathbb{R} 上由开集全体产生的 σ -代数称为直线上的 Borel 代数, 记为 \mathcal{B} 。 \mathcal{B} 中的集称为一维 Borel 集。也等于 $\mathcal{H} = \{(-\infty, a) | a \in \mathbb{R}\}$ 生成的 σ -代数称为直线 \mathbb{R} 上的伯雷尔代数, 记为 $\mathcal{B}(\mathbb{R})$ 。
2. 若存在子集列 $\{A_n\}$, 使得 $\sigma(\{A_n\}) = \mathcal{A}$, 则称 \mathcal{A} 是可列生成的。



定理 1.1.3

设 $\mathbb{R} = (-\infty, +\infty)$ 表示数直线, 则下列集类生成相同的 σ -代数。

1. $\{(a, b) : a, b \in \mathbb{R}\}$;

2. $\{(a, b) : a, b \in \mathbb{R}\}$;
3. $\{[a, b] : a, b \in \mathbb{R}\}$;
4. $\{(-\infty, b], b \in \mathbb{R}\}$;
5. $\{(r_1, r_2) : r_1, r_2 \text{ 为有理数}\}$;
6. $\{G : G \text{ 为 } \mathbb{R} \text{ 中的开集}\}$;
7. $\{F : F \text{ 为 } \mathbb{R} \text{ 中闭集}\}$ 。



为什么?

因为代表元素可以互相表示。

$$(a, b) = \bigcup_{n=1}^{\infty} (a, b - \frac{1}{n}]$$

定义 1.1.5

设 Ω 为一空间, $\mathcal{G} \subset 2^\Omega$ 为某个集合类, \mathcal{G} 上广义实值 (可取 $\pm\infty$) 函数 μ 称为集函数。



若对每个 $A \in \mathcal{G}$, $|\mu(A)| < \infty$, 称 μ 为有限的;

若对任意 $A, B \in \mathcal{G}$, $A \cap B = \emptyset$, 且 $A \cup B \in \mathcal{G}$, 都有 $\mu(A \cup B) = \mu(A) + \mu(B)$, 则称 μ 在 \mathcal{G} 上为有限可加的。

备注记号: 假设 $\{A_\alpha\}$ 为集合序列, 且 $A_i \cap A_j = \emptyset (i \neq j)$, 则 $\sum_\alpha A_\alpha := \bigcup_\alpha A_\alpha$

若对任意 $\{A_n, n \geq 1\} \subset \mathcal{G}$, $A_i \cap A_j = \emptyset, i \neq j$, 且 $\sum_{i=1}^{\infty} A_i \in \mathcal{G}$, 则 $\mu(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$, 则称 μ 在 \mathcal{G} 上为 σ 可加的或可列可加的。

定义 1.1.6

设 (Ω, \mathcal{F}) 为一可测空间, \mathcal{F} 上集函数 μ 称为测度或正测度, 若它满足:

1. $\mu(\emptyset) = 0$;
2. μ 为非负的, 即 $\mu(A) \geq 0, \forall A \in \mathcal{F}$;
3. μ 为可列可加的。

若存在 $\{A_n, n \geq 1\} \in \mathcal{F}$, 使 $\Omega = \bigcup_n A_n$, 且对每个 n , $\mu(A_n) < \infty$, 称 μ 为在 \mathcal{F} 上为 σ 有限的, 简称为 σ 有限的。



定义 1.1.7

设 (Ω, \mathcal{F}) 是可测空间, $P(\cdot)$ 是定义在 \mathcal{F} 上的实值函数。如果

1. $P(\Omega) = 1$;
2. $\forall A \in \mathcal{F}, 0 \leq P(A) \leq 1$;
3. 对两两互不相容事件 A_1, A_2, \dots , (即当 $i \neq j$ 时, $A_i \cap A_j = \emptyset$) 有

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

则称 P 是 (Ω, \mathcal{F}) 上的概率, (Ω, \mathcal{F}, P) 称为概率空间, \mathcal{F} 中的元素称为事件, $P(A)$ 称为事件 A 的概率。



定理 1.1.4

若 μ 为 (Ω, \mathcal{F}) 上的测度, 则

1. μ 是单调的, 即当 $A \subset B$, 必有 $\mu(A) \leq \mu(B)$;
2. μ 是有限次可加的: 若 $A \subset \bigcup_{m=1}^n A_m$, 则 $\mu(A) \leq \sum_{m=1}^n \mu(A_m)$;
3. μ 是 (从下连续): 对每个递增序列 $\{A_n\}$, 便有 $\lim_{n \rightarrow \infty} \mu(A_n) = \mu(\bigcup_n A_n)$ 。
4. 若 μ 是 (从上连续) 的: 对每个递减序列 $\{A_n\}$, 且存在 n_0 , 使 $\mu(A_{n_0}) < \infty$, 便有 $\lim_{n \rightarrow \infty} \downarrow \mu(A_n) = \mu(\bigcap_n A_n)$ 。

5. 若 μ 是有限测度, 则有半可列可加性: 若 $A \subset \bigcup_{m=1}^{\infty} A_m$, 则 $\mu(A) \leq \sum_{m=1}^{\infty} \mu(A_m)$;

定义 1.1.8

设 $(\Omega, \mathcal{F}, \mu)$ 为测度空间。

设 μ 为 σ 代数 \mathcal{F} 上的测度,

$$\mathcal{L} = \{A : A \in \mathcal{F}, \mu(A) = 0\},$$

$$\mathcal{N} = \{N \in 2^{\Omega} : \exists A \in \mathcal{L}, \text{使 } N \subset A\}$$

称 \mathcal{N} 中元素为 μ 可略集。若 $\mathcal{N} \subseteq \mathcal{F}$, 则称 μ 在 \mathcal{F} 上为完备的。(什么意义呢?)

当 $(\Omega, \mathcal{F}, \mu)$ 为概率空间, \mathcal{N} 中的元素简称为“可略集”。由此定义可见, 完备性的要求与 \mathcal{F} 及测度 μ 都是有关的。

定理 1.1.5 (完备化扩张)

若 $(\Omega, \mathcal{F}, \mu)$ 为测度空间, \mathcal{N} 为 μ 可略集全体, 则

1. $\overline{\mathcal{F}} = \{A \cup N : A \in \mathcal{F}, N \in \mathcal{N}\}$ 为 σ 代数, $\overline{\mathcal{F}} \supset \mathcal{F}$;
2. 在 $\overline{\mathcal{F}}$ 上, 令 $\overline{\mu}(A \cup N) = \mu(A)$, 则 $\overline{\mu}$ 是 $\overline{\mathcal{F}}$ 上的测度, $\overline{\mu}|_{\mathcal{F}} = \mu$, 当 μ 为概率测度时 $\overline{\mu}$ 亦然。
3. $(\Omega, \overline{\mathcal{F}}, \overline{\mu})$ 是完备测度空间, 即 $\overline{\mu}$ 在 $\overline{\mathcal{F}}$ 上是完备的。

定义 1.1.9

设 f 为 Ω_1 到 Ω_2 的映照 (即对每个 $\omega_1 \in \Omega_1$, $f(\omega_1)$ 在 Ω_2 中是唯一的), 对 $A_2 \subseteq \Omega_2$,

$$f^{-1}(A_2) = \{\omega_1 \in \Omega_1 : f(\omega_1) \in A_2\}$$

称为 A_2 的原象。对 2^{Ω_2} 的子类 \mathcal{A}_2 ,

$$f^{-1}(\mathcal{A}_2) = \{f^{-1}(A_2) : A_2 \in \mathcal{A}_2\}$$

称为 \mathcal{A}_2 的原象。

定理 1.1.6

设 f 为 Ω_1 到 Ω_2 的任一映照, 则有

1. $f^{-1}(\emptyset) = \emptyset$, $f^{-1}(\Omega_2) = \Omega_1$, $(f^{-1}(A))^c = f^{-1}(A^c)$;
2. $f^{-1}(\bigcup_{\alpha} A_{\alpha}) = \bigcup_{\alpha} f^{-1}(A_{\alpha})$, $f^{-1}(\bigcap_{\alpha} A_{\alpha}) = \bigcap_{\alpha} f^{-1}(A_{\alpha})$;
3. $\sum_{\alpha} f^{-1}(A_{\alpha}) = f^{-1}(\sum_{\alpha} A_{\alpha})$ 。

2^{Ω_2} 的子类 \mathcal{A}_2 ,

$$f^{-1}(\mathcal{A}_2) = \{f^{-1}(A_2) : A_2 \in \mathcal{A}_2\}$$

称为 \mathcal{A}_2 的原象。

定理 1.1.7

由定理 ?? 可知, f^{-1} 与集合的并、交、余集、差、对称差等运算都是可交换的, 而且也不限于可列运算。由此推出, Ω_2 上任一 σ -代数 \mathcal{G} 的原象 $f^{-1}(\mathcal{G})$ 是 Ω_1 的 σ 代数。

证明 以证明可列并封闭为例:

设 $\{B_n\} \subset f^{-1}(\mathcal{G})$ 。则对任意的 n , 存在 $A_n \in \mathcal{G}$, 使得 $B_n = f^{-1}(A_n)$ 。则 $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} f^{-1}(A_n) = f^{-1}(\bigcup_{n=1}^{\infty} A_n) \in f^{-1}(\mathcal{G})$ 。

定义 1.1.10

设 $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ 为可测空间, f 为 Ω_1 到 Ω_2 的映照。若 $f^{-1}(\mathcal{F}_2) \subset \mathcal{F}_1$ (即 $f^{-1}(A) \in \mathcal{F}_1, \forall A \in \mathcal{F}_2$), 则称 f 为 $(\Omega_1, \mathcal{F}_1)$ 到 $(\Omega_2, \mathcal{F}_2)$ 的可测映照, 记为 $f \in \mathcal{F}_1/\mathcal{F}_2$ 或 $f \in \mathcal{F}_1$, 或称 f 为 \mathcal{F}_1 可测的。并记 $\sigma(f) = f^{-1}(\mathcal{F}_2)$, 称为由 f 生成的 σ 代数。



由于我们以后常常会遇到取值包括 $\pm\infty$ 的函数, 因此将上述定义推广一下会带来许多方便。即: 将 $\{-\infty\}, \{+\infty\}$ 也称为 **Borel** 集。因此我们以后将这两个集合和 $\mathcal{B}(\mathbb{R})$ 共同产生的二代数也称为 (扩张) **Borel** 代数并以 $\mathcal{B}(\overline{\mathbb{R}})$ 表示之, 即: $\mathcal{B}(\overline{\mathbb{R}}) := \sigma(\mathcal{B}(\mathbb{R}), \{-\infty\}, \{+\infty\})$ 。

定义 1.1.11

由 (Ω, \mathcal{F}) 到 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (或 $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$) 的可测映照称为可测函数。特别, 当 (Ω, \mathcal{F}) 为概率可测空间时, (Ω, \mathcal{F}) 到 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (或 $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$) 可测映照 X 称为 (实值) 随机变量 (或广义实值随机变量), 也记为 $X \in \mathcal{F}$ 。

**例题 1.1**

1. 若 $(\Omega_1, \mathcal{F}_1) = (\Omega_1, 2^{\Omega_1})$, 则 $(\Omega_1, \mathcal{F}_1)$ 到 $(\Omega_2, \mathcal{F}_2)$ 到的任一映照都是可测的。任取 $B \in \mathcal{F}_2$, 可知 $f^{-1}(B) \in 2^{\Omega_1}$, 因此 $f^{-1}(\mathcal{F}_2) \subseteq 2^{\Omega_1} = \mathcal{F}_1$ 。故可测。
2. 若 $(\Omega_1, \mathcal{F}_1) = (\Omega_1, \{\emptyset, \Omega_1\})$, 则 $(\Omega_1, \mathcal{F}_1)$ 到 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 的可测映照 f 在 Ω_1 上只取同一个值。

设 f 不为常值函数, 则存在 $\omega_1, \omega_2 \in \Omega_1$, 使得 $f(\omega_1) = a_1 < a_2 = f(\omega_2)$ 。则 $\omega_1 \in f^{-1}((-\infty, \frac{a_1+a_2}{2}))$ 。由于 f 可测, 则 $f^{-1}((-\infty, \frac{a_1+a_2}{2})) \in \mathcal{F}_1$ 。但 $\omega_2 \notin f^{-1}((-\infty, \frac{a_1+a_2}{2})) = \Omega_1$ 。

定理 1.1.8

若 $E = \{r_n\}$ 为 \mathbb{R} 中稠密集, 则 X 为随机变量的充要条件是对每个 $r_n \in E$, $\{\omega : X(\omega) \leq r_n\} \in \mathcal{F}$ 。



例题 1.2 设 (Ω, \mathcal{F}) 为可测空间, $A \in \mathcal{F}$, 则

$$f(x) := I_A(x) = \begin{cases} 1 & x \in A; \\ 0 & x \notin A. \end{cases}$$

为可测函数。因为

$$\{\omega | f \leq a\} = \begin{cases} \emptyset & a < 0; \\ A^c, & 0 \leq a < 1; \\ \Omega, & a \geq 1. \end{cases}$$

定理 1.1.9

设 (Ω, \mathcal{F}) 为可测空间, $f, g, \{f_n, n \geq 1\}$ 为可测函数, c 是常数。假定下面出现的所有运算均有意义。则 $f + g, cf, f^{-1}, fg, \sup_{n \geq 1} f_n, \inf_{n \geq 1} f_n, \lim_{n \rightarrow \infty} f_n, \overline{\lim_{n \rightarrow \infty} f_n}$, 均是可测函数。

**定理 1.1.10**

设 $(\Omega_i, \mathcal{F}_i), i = 1, 2, 3$, 为可测空间。若 $g \in \mathcal{F}_1/\mathcal{F}_2, f \in \mathcal{F}_2/\mathcal{F}_3$, 则 $f \circ g \in \mathcal{F}_1/\mathcal{F}_3$, 其中 $f \circ g(\omega_1) \triangleq f(g(\omega_1))$ 。



证明 由复合映照的定义, 对任意 $A \in \mathcal{F}_3$, 有 $f^{-1}(A) \in \mathcal{F}_2$,

$$(f \circ g)^{-1}(A) = \{\omega_1 : f(g(\omega_1)) \in A\} = \{\omega_1 : g(\omega_1) \in f^{-1}(A)\} = g^{-1}(f^{-1}(A)) \in \mathcal{F}_1$$

故 $f \circ g \in \mathcal{F}_1/\mathcal{F}_3$ 。

定理 1.1.11

对 Ω 的子集类 \mathcal{C} , 若以 $\mathcal{C} \cap A \triangleq \{B \cap A : B \in \mathcal{C}\}$, 则

$$\sigma_{\Omega}(\mathcal{C}) \cap A = \sigma_A(\mathcal{C} \cap A),$$

这里, $\sigma_A(\mathcal{C} \cap A)$ 表示以 A 为全集, 且由 $\mathcal{C} \cap A$ 生成的 σ -代数。当然, $\sigma_{\Omega}(\mathcal{C}) = \sigma(\mathcal{C})$ 。

$$\sigma_{\Omega}(\mathcal{C}) \cap A := \{B \cap A \mid B \in \sigma_{\Omega}(\mathcal{C})\}.$$



设 $\{f_{\lambda} : \lambda \in \Lambda\}$ 是 $\Omega \rightarrow \Omega'$ 的映射族, \mathcal{F}' 是 Ω' 上的 σ -代数, 那么 Ω 上存在唯一一个使得映射 $\{f_{\lambda} : \lambda \in \Lambda\}$ 都可测的最小 σ -代数 \mathcal{F} , 即

1. 每个 f_{λ} 是可测映射;
2. 如果 Ω 上另外一个 σ -代数 \mathcal{F}_1 使得每个 f_{λ} 都可测, 那么 $\mathcal{F} \subset \mathcal{F}_1$ 。事实上, 不难验证 $\mathcal{F} = \sigma(\bigcup_{\lambda \in \Lambda} f_{\lambda}^{-1}(\mathcal{F}'))$, 记为 $\sigma(\{f_{\lambda} : \lambda \in \Lambda\})$ 。

1.2 积分

在本节中, 设测度空间 $(\Omega, \mathcal{F}, \mu)$ 固定, μ 是有限测度。

定义 1.2.1

若 $f(\omega) = \sum_{i=1}^n a_i I_{A_i}(\omega)$, 则称 f 为阶梯函数, 其中 $a_i \in \mathbb{R}, a_i \neq a_j$ 。称 $\sum_i a_i \mu(A_i)$ 为 f 关于 μ 的积分, 记为: $\mu(f)$



注 当 f 为广义实值随机变量且不会同时取 $+\infty$ 及 $-\infty$ 时, 若约定 $0 \cdot (\pm\infty) = 0$, 则仍可如上规定 $\mu(f)$ 。

命题 1.2.1

记 \mathcal{G}_+ 表示非负阶梯型函数全体,

$$\mathcal{G}_+ = \{f = \lim_n \uparrow f_n : f_n \in \mathcal{G}_+, n \geq 1\}.$$

对 $f \in \mathcal{G}_+$, 若 $f = \lim_n \uparrow f_n, f_n \in \mathcal{G}_+, n \geq 1$, 令

$$\mu(f) \triangleq \lim_n \mu(f_n), \quad (1.2.1)$$

则

1. \mathcal{G}_+ 为 (Ω, \mathcal{F}) 上非负可测函数全体;
2. 由 ?? 式规定的 $\mu[f]$ 是完全确定的;

**定义 1.2.2**

广义实值可测函数 f , 若 $\mu[f^+] < \infty, \mu[f^-] < \infty$, 则称 f 为可积的, 且以 $\mu(f) = \mu(f^+) - \mu(f^-)$ 表示 f 对 μ 的积分, 记为 $\int f d\mu$ 。

较为一般地, 若 $\mu(f^+), \mu(f^-)$ 中至少有一个取有限值, 则称 f 积分存在的, 用 $\mu(f) = \mu[f^+] - \mu[f^-]$ 表示 f 关于 μ 的积分或期望。

**定义 1.2.3**

实值可测函数 f , 关于定义在 $(a, b]$ 区间上的单调函数 g 的 Riemann-Stieltjes 积分定义为

$$\int_a^b f dg = \int_{(a,b]} f(t) dg(t) := \lim_{\delta \rightarrow 0} \sum_{i=1}^n f(\xi_i^n)(g(t_i^n) - g(t_{i-1}^n)).$$

其中 $\{a = t_0^n < t_1^n \cdots < t_n^n = b\}$, $|\delta| := \sup_{1 \leq i \leq n} |t_i^n - t_{i-1}^n|$ 。



定义 1.2.4

若 g 为有限变差函数, 则由实变函数可知 $g(t) = a(t) - b(t)$ 其中 $a(t) = V_g(t)$, $b(t) = V_g(t) - g(t)$ 均为单调增函数。若 $\int_0^t |f(s)| da(s) = \int_0^t |f(s)| dV_g(s) := \int_0^t |f(s)| |dg(s)| < \infty$ 。则 f 关于 g 的 Stieltjes 积分定义为

$$\int_{(0,t]} f(s) dg(s) := \int_{(0,t]} f(s) da(s) - \int_{(0,t]} f(s) db(s).$$



注 记号说明: $\int_a^b f(s) dg(s) = \int_{(a,b]} f(s) dg(s)$ 。

$$\int_{(0,t)} dg(s) = g(t-) - g(0).$$

定理 1.2.1 (Protter)

若给定区间 $[a, b]$, 定义划分 $\Delta := \{a = t_0^n < t_1^n \dots < t_n^n = b\}$ 。 $\delta_n := \sup_{1 \leq i \leq n} |t_i^n - t_{i-1}^n|$, $\delta_n = \max_i (t_i^n - t_{i-1}^n)$ 。若 $\lim_{\delta_n \rightarrow 0} \sum_{i=1}^n f(t_{i-1}^n) [g(t_i^n) - g(t_{i-1}^n)]$ 对任意的连续函数 f 存在, 则 g 一定为 $[a, b]$ 上的有限变差函数



故若 g 为无穷变差函数, 则积分不存在。

定理 1.2.2

设 f 为可测函数, $A \in \mathcal{F}$, 若 f 积分存在, 则记

$$\int_A f d\mu = \mu[fI_A].$$

$\varphi(A) = \int_A f d\mu$ ($A \in \mathcal{F}$), 看作 $A \in \mathcal{F}$ 的函数时, 称为 f 关于测度 μ 的不定积分。

**定理 1.2.3 (单调收敛定理)**

1. 若 $f_n \uparrow f$ a.e., 且对某个 n_0 使 $\mu(f_{n_0}) > -\infty$, 则 $\lim_n \mu(f_n) = \mu(f)$;
2. 若 $f_n \downarrow f$ a.e., 且对某个 n_0 使 $\mu(f_{n_0}) < +\infty$, 则 $\lim_n \mu(f_n) = \mu(f)$ 。

**定理 1.2.4 (Lebesgue 控制收敛定理)**

若 $\{f_n, n \geq 1\}$ 为可测函数序列, $|f_n| \leq Y$, Y 可积, 且 $\lim_n f_n = f$ 存在, 则

$$\lim_n \mu(f_n) = \mu(f).$$

**定理 1.2.5 (Fatou 引理)**

设 $\{f_n, n \geq 1\}$ 为可测函数序列, Y, Z 均为可积可测函数, 则

1. 若 $f_n \geq Z$ ($n \geq n_0$), 则

$$\mu \left[\lim_n f_n \right] \leq \lim_n \mu(f_n). \quad (1.2.2)$$

2. 若 $f_n \leq Y, n \geq n_0$, 则

$$\mu \left[\lim_n f_n \right] \geq \lim_n \mu(f_n). \quad (1.2.3)$$

设 λ 是 $[0, 1]$ 上 Lebesgue 测度。定义 $f_{2n-1} := 1_{[0, \frac{1}{2})}$, $f_{2n} := 1_{(\frac{1}{2}, 1]}$

**定理 1.2.6**

设 f, g 积分存在

1. $\forall \alpha \in \mathbb{R}$, αf 的积分存在, 且 $\mu(\alpha f) = \alpha \mu(f)$;
2. 若 $f + g$ 处处有定义, 且 $\mu(f) + \mu(g)$ 有意义, 则 $f + g$ 的积分存在, 且有 $\mu(f + g) = \mu(f) + \mu(g)$;

3. $|\mu(f)| \leq \mu(|f|)$;
4. 若 N 为一零测集, 则 $\mu(fI_N) = 0$;
5. 若 $f \leq g$, a.e., 则 $\mu(f) \leq \mu(g)$;
6. 若 f 为非负实值可测函数, 则 $f = 0$, a.e. 等价于 $\mu(f) = 0$;
7. 若 f 可积, $\forall \varepsilon > 0$, 存在 $\delta > 0$, s.t. $\forall A \in \mathcal{F}$, 只需 $\mu(A) \leq \delta$ 就有 $|\int_A f d\mu| \leq \int_A |f| d\mu \leq \varepsilon$.



定理 1.2.7 (变量代换)

若 $(\Omega, \mathcal{F}, \mu)$ 为测度空间, Y 为 (Ω, \mathcal{F}) 到可测空间 (E, \mathcal{E}) 的可测映照, μY^{-1} 为 Y 在 (E, \mathcal{E}) 上的导出测度, 又 f 是 (E, \mathcal{E}) 上的可测函数, 则下列两端任一端存在 (有限) 必可推出另一端也存在 (有限), 且有:

$$\int_E f(x) \mu Y^{-1}(dx) = \int_{\Omega} f(Y(\omega)) \mu(d\omega). \quad (1.2.4)$$

其中任取 $A \in \mathcal{E}$, $\mu Y^{-1}(A) := \mu(Y^{-1}(A))$.

且

$$\int_A f(x) \mu Y^{-1}(dx) = \int_{Y^{-1}(A)} f(Y(\omega)) \mu(d\omega), \forall A \in \mathcal{E} \quad (1.2.5)$$



1.3 随机变量的收敛性

令 $\mathcal{L}^p(\Omega), p \geq 1$ 表示所有使得 $E[|X|^p] < \infty$ 的随机变量 (等价类) 全体, 简记为 \mathcal{L}^p .

定义 1.3.1

1. 设 $\{X_n, n \geq 1\}$ 是随机变量序列, 若存在随机变量 X 使得

$$P\{\omega \in \Omega : X(\omega) = \lim_{n \rightarrow \infty} X_n(\omega)\} = 1$$

则称随机变量序列 $\{X_n, n \geq 1\}$ 几乎必然收敛 (或以概率 1 收敛) 于 X , 记为 $X_n \rightarrow X, a.s.$ 或 $X_n \xrightarrow{X} X$.

2. 设 $\{X_n, n \geq 1\}$ 是随机变量序列, 若存在随机变量 X 使得 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} P\{|X_n - X| \geq \varepsilon\} = 0$$

则称随机变量序列 $\{X_n, n \geq 1\}$ 依概率收敛于 X , 记为 $X_n \xrightarrow{P} X$.

3. 设随机变量序列 $\{X_n\} \subset \mathcal{L}^p, p \geq 1, X \in \mathcal{L}^p$, 若有

$$\lim_{n \rightarrow \infty} E[|X_n - X|^p] = 0$$

则称随机变量序列 $\{X_n, n \geq 1\}$ p 次平均收敛于 X , 或称 $\{X_n\}$ 在 \mathcal{L}^p 中强收敛于 X 。当 $p = 2$ 时, 称为均方收敛。

4. 设 $\{F_n(x)\}$ 是分布函数列, 如果存在一个单调不减函数 $F(x)$, 使得在 $F(x)$ 的所有连续点 x 上均有

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

则称 $\{F_n(x)\}$ 弱收敛于 $F(x)$, 记为 $F_n(x) \xrightarrow{W} F(x)$ 。

设随机变量 X_n, X 的分布函数分别为 $F_n(x)$ 及 $F(x)$, 若

$$F_n(x) \xrightarrow{W} F(x)$$

则称 $\{X_n\}$ 依分布收敛于 X , 记为 $X_n \xrightarrow{L} X$ 。



定理 1.3.1

1. 随机变量序列 $X_n \xrightarrow{X}$ 的充分必要条件是 $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} P \left\{ \sup_{m \geq n} |X_m - X| \geq \varepsilon \right\} = 0$$

2. 随机变量序列 $X_n \xrightarrow{P} X$ 的充分必要条件是 $\{X_n\}$ 的任意子序列都包含几乎必然收敛于 X 的子序列。♡

随机变量序列的这 4 种收敛性之间的关系可以总结为下面的关系图：

几乎必然收敛 \implies 依概率收敛 \implies 依分布收敛；

p 次平均收敛 \implies 依概率收敛 \implies 依分布收敛。

注 几乎必然收敛与 p 阶矩收敛之间没有蕴含关系。

1.4 随机变量和分布函数

定义 1.4.1

设 (Ω, \mathcal{F}, P) 是 (完备的) 概率空间, X 是定义在 Ω 上取值于实数集 \mathbb{R} 的函数, 如果对任意实数 $x \in \mathbb{R}$, $\{\omega : X(\omega) \leq x\} \in \mathcal{F}$, 则称 $X(\omega)$ 是 \mathcal{F} 上的随机变量, 简称为随机变量。

$$F(x) = P(\omega : X(\omega) \leq x), \quad -\infty < x < \infty$$

称为随机变量 X 的分布函数。

如果存在函数 $f(x)$, 满足

$$F(x) = \int_{-\infty}^x f(t) dt$$

则称 $f(x)$ 为随机变量 X 或其分布函数 $F(x)$ 的分布密度。如果 X 具有分布密度, 则称 X 为连续型随机变量；

如果 X 最多以正概率取可数多个值, 则称 X 为离散型随机变量。♣

定义 1.4.2

两个随机变量 X 与 Y , 如果满足 $P(\omega \in \Omega : X(\omega) \neq Y(\omega)) = 0$, 则称它们是等价的。♣

注

对于两个等价的随机变量, 我们视为同一。

定理 1.4.1

下列命题等价：

1. X 是随机变量；
2. $\{\omega : X(\omega) \geq a\} \in \mathcal{F}, \quad \forall a \in \mathbb{R}$;
3. $\{\omega : X(\omega) > a\} \in \mathcal{F}, \quad \forall a \in \mathbb{R}$;
4. $\{\omega : X(\omega) < a\} \in \mathcal{F}, \quad \forall a \in \mathbb{R}$ 。

♡

注

习惯上将 $\{\omega : X(\omega) \geq a\}$ 记为 $\{X \geq a\}$

定理 1.4.2

1. 若 X, Y 是随机变量, 则 $\{X < Y\}, \{X \leq Y\}, \{X = Y\}$ 及 $\{X \neq Y\}$ 都属于 \mathcal{F} ;
 2. 若 X, Y 是随机变量, 则 $X \pm Y$ 与 XY 亦然;
 3. 若 $\{X_n\}$ 是随机变量序列, 则 $\sup_n X_n, \inf_n X_n, \limsup_{n \rightarrow \infty} X_n$ 和 $\liminf_{n \rightarrow \infty} X_n$ 都是随机变量。
- 映射 $X: \Omega \rightarrow \mathbb{R}^d$, 表示为 $\mathbf{X} = (X_1, \dots, X_d)$, 若对所有的 $k, 1 \leq k \leq d$, X_k 都是随机变量, 则称 \mathbf{X} 为随机向量。



常用的两种类型随机变量:

1. 离散型随机变量 X 的概率分布用分布列描述:

$$p_k = P(X = x_k), \quad k = 1, 2, \dots$$

其分布函数 $F(x) = \sum_{x_k \leq x} p_k$.

2. 连续型随机变量 X 的概率分布用概率密度 $f(x)$ 描述, 其分布函数

$$F(x) = \int_{-\infty}^x f(t) dt.$$

3. 对于随机向量 $X = (X_1, \dots, X_d)$, 它的 (d 维) 分布函数 (或联合分布函数) 定义为

$$F(x_1, \dots, x_d) = P(X_1 \leq x_1, \dots, X_d \leq x_d)$$

这里 d 为正整数, $x_k \in \mathbb{R}, k = 1, 2, \dots, d$.

定理 1.4.3

若 $F(x_1, \dots, x_d)$ 是联合分布函数, 则

1. $F(x_1, \dots, x_d)$ 对每个变量都是单调的;
2. $F(x_1, \dots, x_d)$ 对每个变量都是右连续的;
3. 对 $i = 1, 2, \dots, d$

$$\lim_{x_i \rightarrow -\infty} F(x_1, \dots, x_i, \dots, x_d) = 0,$$

$$\lim_{x_1, x_2, \dots, x_d \rightarrow \infty} F(x_1, x_2, \dots, x_d) = 1.$$

**注**

如果 $f(x_1, \dots, x_d) = \frac{\partial^d F}{\partial x_1 \cdots \partial x_d}$ 对所有的 $(x_1, \dots, x_d) \in \mathbb{R}^d$ 存在, 则称函数 $f(x_1, \dots, x_d)$ 为 $F(x_1, \dots, x_d)$ 或 $X = (X_1, \dots, X_d)$ 的联合密度函数, 并且

$$F(x_1, \dots, x_d) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_d} f(t_1, \dots, t_d) dt_d \cdots dt_1$$

设 $F(x_1, \dots, x_d)$ 为 X_1, \dots, X_d 的联合分布函数, $1 \leq k_1 < \dots < k_n \leq d$, 则 X_1, \dots, X_d 的边缘分布 $F_{k_1, \dots, k_n}(x_{k_1}, \dots, x_{k_n})$ 定义为

$$F_{k_1, \dots, k_n}(x_{k_1}, \dots, x_{k_n}) = F(\infty, \dots, \infty, x_{k_1}, \infty, \dots, \infty, x_{k_2}, \infty, \dots, \infty, x_{k_n}, \infty, \dots, \infty)$$

1.5 分布函数及其生成的测度**定理 1.5.1**

若 $F(x)$ 为有限实值随机变量 X 的分布函数, 则

1. $F(x)$ 是不减的;
2. $F(x)$ 是右连续的;
3. $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$.



定理 1.5.2

(a) 若 $F(x)$ 为 \mathbb{R} 上的右连续不减有界函数, 则在 $(\mathbb{R}, \mathcal{B})$ 必存在唯一的有限测度 μ , 使得:

$$\mu((a, b]) = F(b) - F(a), \quad -\infty \leq a < b < +\infty.$$

(b) 若 $F(x)$ 为 \mathbb{R} 上的右连续不减的实值函数, 则必存在在 $(\mathbb{R}, \mathcal{B})$ 上唯一的 σ 有限的测度 μ 使得

$$\mu((a, b]) = F(b) - F(a), \quad -\infty \leq a < b < +\infty.$$



定理 1.5.3

若 $F(x)$ 为 \mathbb{R} 上满足定理 ??(1)-(3) 的函数, 则必存在概率空间 (Ω, \mathcal{F}, P) 及其上的随机变量 X , 使得

$$P(X \leq x) = F(x).$$



定义 1.5.1

若 F 为 \mathbb{R} 上有限右连续不减函数, 则由 F 在 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上按定理 ??(b) (即 $\mu((a, b]) = F(b) - F(a), -\infty \leq a < b < +\infty$) 生成的 σ 有限完备测度 μ 称为由 F 生成的 Lebesgue-Stieltjes 测度, 简称为 L-S 测度. 特别, 当 $F(t) = t$ (或同样的 $f(t) = t + c$), 由此产生的完备化测度称为 Lebesgue 测度. 由 \mathcal{B} 按 Lebesgue 测度扩张的完备 σ 代数 $\mathcal{B}(\mathbb{R})$ 中的集合都称为 Lebesgue 可测集.



1.6 数字特征、矩母函数与特征函数

1.6.1 数字特征

定义 1.6.1

1. 取值为 $\{s_k\}$ 的离散型随机变量 X 的数学期望 (简称为期望) $E[X]$ 定义为

$$E[X] = \sum_k s_k p_k = \sum_k s_k P(X = s_k)$$

如果 $\sum |s_k| p_k < \infty$.

2. 连续型随机变量 X 的数学期望 $E[X]$ 定义为

$$E[X] = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x) dx$$

如果 $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, 这里 $F(x)$ 是 X 的分布函数, $f(x)$ 是其密度函数.

利用 Riemann-Stieltjes 积分, 我们可以对离散型随机变量和连续型随机变量的期望给出一个统一的表达式:

$$E[X] = \int_{-\infty}^{+\infty} x dF(x)$$

3. 设 X 为任一随机变量, 对正整数 k , 称 $m_k = E[X^k]$ 为 X 的 k 阶原点矩. 数学期望是一阶原点矩.

4. 设 X 为任一随机变量, 对正整数 k , 称 $c_k = E[X - E[X]]^k$ 为 X 的 k 阶中心矩. 方差是二阶中心矩.



1.6.2 矩母函数

定义 1.6.2

若随机变量 X 的分布函数为 $F_X(x)$, 则称

$$\phi_X(t) = E[e^{tX}] = \int_{\Omega} e^{tX(\omega)} P(d\omega) = \int_{-\infty}^{\infty} e^{tx} dF_X(x)$$

为 X 的矩母函数。



注

X 的各阶矩与矩母函数的关系 (假设对 $\phi(t)$ 求导时, 求导运算与求期望运算可以交换次序)

$$\phi'(t) = E(Xe^{tX})$$

$$\phi''(t) = E(X^2e^{tX})$$

$$\vdots$$

$$\phi^{(n)}(t) = E(X^n e^{tX})$$

令 $t = 0$, 得到 $\phi^{(n)}(0) = E[X^n]$, $n \geq 1$. 当矩母函数存在时, 它唯一地决定分布, 因此我们能够用矩母函数刻画随机变量的概率分布. 但有时随机变量的矩母函数不一定存在, 在这种情况下, 更方便的是特征函数。

1.6.3 特征函数

定义 1.6.3

若随机变量 X 的分布函数为 $F_X(x)$, 则称

$$\psi_X(t) = E[e^{itX}] = \int_{\Omega} e^{itX(\omega)} P(d\omega) = \int_{-\infty}^{\infty} e^{itx} dF_X(x)$$

为 X 的特征函数。如果 F_X 有密度 $f(x)$, 则 $\psi_X(t)$ 就是 $f(x)$ 的 Fourier 变换

$$\psi_X(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx$$



1.7 独立性与条件期望

1.7.1 独立性

定义 1.7.1

1. 设 A, B 为两个事件, 若 $P(A \cap B) = P(A)P(B)$, 则称 A 与 B 独立. 更一般地, 设 A_1, A_2, \dots, A_n 为 n 个事件, 如果对任何 $m \leq n$ 及 $1 \leq k_1 < k_2 < \dots < k_m \leq n$, 有

$$P\left(\bigcap_{j=1}^m A_{k_j}\right) = \prod_{j=1}^m P(A_{k_j})$$

则称 A_1, A_2, \dots, A_n 相互独立. A_1, A_2, \dots, A_n 两两独立不一定相互独立.

2. 设 $\{A_i, i \in I\}$ 是一族事件, 若对 I 的任意有限子集 $\{i_1, \dots, i_k\} \neq \emptyset$ 有

$$P\left(\bigcap_{j=1}^k A_{i_j}\right) = \prod_{j=1}^k P(A_{i_j}) \quad (1.7.1)$$

则称 $\{A_i, i \in I\}$ 是相互独立的.

3. 设 $\{A_i, i \in I\}$ 是一族事件类, 如果对 I 的任意有限子集 $\{i_1, \dots, i_k\} \neq \emptyset$, 任意 $A_i \in A_{i_j}$, 有 (1.7.1) 式成立, 则称 $\{A_i, i \in I\}$ 是独立事件类.
4. 设 $\{X_i, i \in I\}$ 是 Ω 上一族随机变量, 如果 σ 代数 $\{X_i, i \in I\}$ 是独立事件类, 则称 $\{X_i, i \in I\}$ 相互独立.



容易证明随机变量 X_1, \dots, X_n 独立的充分必要条件是它们的联合分布函数可以分解为

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$$

定理 1.7.1

设 $\{X_t, t \in T\}$ 为独立随机变量族, $\{T_\alpha, \alpha \in J\}$ 为 T 的互不相交子集, $\{f_\alpha(X_t, t \in T_\alpha), \alpha \in J\}$ 为 Borel 函数族, 则 $\{Y_\alpha = f_\alpha(X_t, t \in T_\alpha), \alpha \in J\}$ 为独立随机变量。

**1.7.2 条件期望**

设 B 是一个事件, 且 $P(B) > 0$. 则事件 B 发生的条件下

事件 A 发生的条件概率为

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

(全概率公式) 设 $\{B_n\}$ 是 Ω 的一个分割, 且使得 $P(B_n) > 0, \forall n$. 如果 $A \in \mathcal{F}$, 则

$$P(A) = \sum_n P(B_n)P(A|B_n)$$

(Bayes 公式) 设 $\{B_n\}$ 是 Ω 的一个分割, 且使得 $P(B_n) > 0, \forall n$. 如果 $P(A) > 0$, 则

$$P(B_k|A) = \frac{P(B_k)P(A|B_k)}{\sum_n P(B_n)P(A|B_n)}, \quad n \geq 1$$

如果 X 与 Y 是离散型随机变量, 对一切使得 $P\{Y = y\} > 0$ 的 y , 给定 $Y = y$ 时, X 的条件概率定义为:

$$P\{X = x|Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}}$$

X 的条件分布定义为:

$$F(x|y) = P\{X \leq x|Y = y\}$$

X 的条件期望定义为:

$$E[X|Y = y] = \int x dF(x|y) = \sum_x x P\{X = x|Y = y\}$$

如果 X 与 Y 有联合概率密度函数 $f(x, y)$, 则对一切使得 $f_Y(y) > 0$ 的 y , 给定 $Y = y$ 时, X 的条件概率密度函数定义为:

$$f(x|y) = \frac{f(x, y)}{f_Y(y)}$$

X 的条件分布定义为:

$$F(x|y) = P\{X \leq x|Y = y\} = \int_{-\infty}^x f(z|y) dz$$

X 的条件期望定义为:

$$E[X|Y = y] = \int x dF(x|y) = \int x f(x|y) dx$$

我们以 $E[X|Y]$ 表示随机变量 Y 的函数, 它在 $Y = y$ 时, 取值为 $E[X|Y = y]$. 条件期望的一个重要性质是对一切随机变量 X 和 Y , 当期望存在时, 有

$$E[X] = E[E[X|Y]] = \int E[X|Y = y] dF_Y(y) \quad (1.7.2)$$

当 Y 为一个离散随机变量时, ?? 式为

$$E[X] = \sum_y E[X|Y = y] P\{Y = y\}$$

当 Y 为一个连续随机变量时, ?? 式为

$$E[X] = \int_{-\infty}^{+\infty} E[X|Y = y] f(y) dy$$

例题 1.3 随机个随机变量之和 设 X_1, X_2, \dots 是一列与 X 独立同分布的随机变量; 设 N 为一非负整值随机变量, 且与序列 X_1, X_2, \dots 独立. 求 $Y = \sum_{i=1}^N X_i$ 的均值和方差。

证明 首先在对 N 取条件的情况下来计算 $Y = \sum_{i=1}^N X_i$ 的矩母函数, 即

$$E[\exp\{t \sum_{i=1}^N X_i\} | N = n] = E[\exp\{t \sum_{i=1}^n X_i\}] = (\phi_X(t))^n$$

其中 $\phi_X(t)$ 是随机变量 X 的矩母函数, 因此

$$E[\exp\{t \sum_{i=1}^N X_i\} | N] = (\phi_X(t))^N$$

从而

$$\phi_Y(t) = E[\exp\{t \sum_{i=1}^N X_i\}] = E[(\phi_X(t))^N]$$

现在对 $\phi_Y(t)$ 求导得

$$\phi_Y'(t) = E[N(\phi_X(t))^{N-1} \phi_X'(t)]$$

再求一次导数得

$$\phi_Y''(t) = E[N(N-1)(\phi_X(t))^{N-2} (\phi_X'(t))^2 + N(\phi_X(t))^{N-1} \phi_X''(t)]$$

计算在 $t = 0$ 点的值, 得

$$E[Y] = E[NE[X]] = E[N]E[X]$$

及

$$E[Y^2] = E[N(N-1)(E[X])^2 + NE[X^2]] = E[N]\text{Var}[X] + E[N^2](E[X])^2$$

因此有

$$\text{Var}[Y] = E[Y^2] - (E[Y])^2 = E[N]\text{Var}[X] + (E[X])^2\text{Var}[N]$$

定义 1.7.2

线性空间 \mathcal{H} 上的一个共轭双线性函数 $(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow K$ 称为一个内积, 若

1. $(x, y) = \overline{(y, x)} \forall x, y \in \mathcal{H}$;
2. $(x, x) \geq 0, \forall x \in \mathcal{H}$ 且 $(x, x) = 0 \Leftrightarrow x = 0$.

完备的内积空间称为 Hilbert 空间。



定义 1.7.3

内积空间 \mathcal{H} 上的两个元素 x, y 称为正交的, 若 $(x, y) = 0$, 记为 $x \perp y$ 。

设 M 是 \mathcal{H} 上的一个非空子集, $x \in \mathcal{H}$ 。若 $\forall y \in M$ 都有 $x \perp y$ 则称 x 与 M 正交, 记为 $x \perp M$ 。

记 $\{x \in \mathcal{H} | x \perp M\}$ 为 M 的正交补空间, 记为 M^\perp 。



定理 1.7.2

设 $f \in L^2(\Omega, \mathcal{F}, P)$, $g \in L^2(\Omega, \mathcal{G}, P)$, 则下列条件等价:

1. $g = E[f|\mathcal{G}]$;
2. $\forall h \in L^2(\Omega, \mathcal{G}, P), \int fhdP = \int gh dP$;
3. $\forall h \in L^\infty(\Omega, \mathcal{G}, P), \int fhdP = \int gh dP$;
4. $\forall A \in \mathcal{G}, \int_A f dP = \int_A g dP$;



定理 1.7.3

$f \rightarrow E[f|\mathcal{G}]$ 可唯一扩充为 $L^1(\Omega, \mathcal{F}, P)$ 到 $L^1(\Omega, \mathcal{G}, P)$ 的连续线性映射。



定理 1.7.4

设 $f \in L^1(\Omega, \mathcal{F}, P)$, $g \in L^1(\Omega, \mathcal{G}, P)$, 则下列条件等价:

1. $g = E[f|\mathcal{G}]$;
2. $\forall A \in \mathcal{G}, \int_A f dP = \int_A g dP$;
3. $\forall h \in L^\infty(\Omega, \mathcal{G}, P), \int f h dP = \int g h dP$;

**定义 1.7.4**

设 X 是随机变量且 $E[|X|] < \infty$. 若对每个子 σ -代数 $\mathcal{G} \subset \mathcal{F}$, 存在唯一的 (L^1 意义下唯一的) 随机变量 X^* , 有 $E[|X^*|] < \infty$, 使得 X^* 是 \mathcal{G} 可测随机变量 (即对任何 $a \in \mathbb{R}$, 有 $\{X^* \leq a\} \in \mathcal{G}$), 且

$$E[X^* I_B] = E[X I_B], \quad \forall B \in \mathcal{G}$$

则称随机变量 X^* 为 X 在给定 \mathcal{G} 下的条件期望, 记为 $X^* = E[X|\mathcal{G}]$, 即

$$\int_B E[X|\mathcal{G}] dP = \int_B X dP, \quad \forall B \in \mathcal{G}$$

特别地, 当 $\mathcal{G} = \sigma(Y)$ 时, 则记 $E(X|Y)$ 为 $E[X|Y]$ 并称为 X 关于 Y 的条件期望。

**定理 1.7.5**

条件期望有如下基本性质:

1. 设 σ -代数 $\mathcal{G} \subset \mathcal{F}$, 则 $E[E[X|\mathcal{G}]] = E[X]$.
2. 若 X 是 \mathcal{G} 可测, 则 $E[X|\mathcal{G}] = X$, a.s.
3. 设 $\mathcal{G} = \{\emptyset, \Omega\}$, 则 $E[X|\mathcal{G}] = E[X]$, a.s.
4. $E[X|\mathcal{G}] = E[X^+|\mathcal{G}] - E[X^-|\mathcal{G}]$, a.s.
5. 若 $X \leq Y$, a.s., 则 $E[X|\mathcal{G}] \leq E[Y|\mathcal{G}]$, a.s.
6. 若 a, b 为常数, $X, Y, aX + bY$ 的期望存在, 则 $E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}]$, a.s.

如果右端和式有意义。

7. $|E[X|\mathcal{G}]| \leq E[|X||\mathcal{G}]$, a.s.
8. 设 $0 \leq X_n \uparrow X$, a.s., $E[X_n|\mathcal{G}] \uparrow E[X|\mathcal{G}]$, a.s.
9. 设 X 及 XY 的期望存在, 且 Y 为 \mathcal{G} 可测, 则 $E[XY|\mathcal{G}] = YE[X|\mathcal{G}]$, a.s.
10. 若 X 与 \mathcal{G} 相互独立 (即 $\sigma(X)$ 与 \mathcal{G} 相互独立), 则有 $E[X|\mathcal{G}] = E[X]$, a.s.
11. 若 $\mathcal{G}_1, \mathcal{G}_2$ 是两个子 σ -代数, 使得 $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$, 则 $E[E[X|\mathcal{G}_2]|\mathcal{G}_1] = E[X|\mathcal{G}_1]$, a.s.
12. 若 X, Y 是两个独立的随机变量, 函数 $g(x, y)$ 使得 $E[|g(X, Y)|] < +\infty$, 则有

$$E[g(X, Y)|Y] = E[g(X, y)|y] \cdots, \quad \text{a.s.}$$

这里 $E[g(X, y)|y]$ 的意义是, 先将 y 视为常数, 求得数学期望 $E[g(X, y)]$ 后再将随机变量 Y 代入到的位置。

**定义 1.7.5**

设 $\mathcal{G} \subset \mathcal{F}$, $A \in \mathcal{F}$, 则称 $E[I_A|\mathcal{G}]$ 为 A 关于 \mathcal{G} 的条件概率, 记为 $P(A|\mathcal{G})$ 。



1.8 随机过程的基本概念

定义 1.8.1

随机过程是概率空间 (Ω, \mathcal{F}, P) 上的一族随机变量 $\{X(t), t \in T\}$, 其中 t 是参数, 它属于某个指标集 T , 可称为参数集。



注

- 当 $T = \{0, 1, 2, \dots\}$ 时称之为随机序列或时间序列。
- 随机过程 $\{X(t), \omega\}, t \in T, \omega \in \Omega$ 决定定义在 $T \times \Omega$ 上的二元函数。

随机过程的分类:

1. $X(t)$ 表示系统在时刻 t 所处的状态。
2. $X(t)$ 的所有可能状态构成的集合为状态空间, 记为 S 。
3. 依据状态空间可分为连续状态和离散状态;
4. 依据参数集, 可分为离散参数过程和连续参数过程。

例如: $X(t)$ 表示时刻一个队列中排队的人数, 则 $S = \{0, 1, 2, \dots\}$; $X(t)$ 表示时刻某只股票的价格, 则 $S = [0, \infty)$ 。

注 一般如果不作说明都认为状态空间是实数集 \mathbb{R} 的子集。

例题 1.4 随机游动 一个醉汉在路上行走, 以概率 p 前进一步, 以概率 $1 - p$ 后退一步 (假定其步长相等), 以 $X(t)$ 记他在路上的位置, 则 $X(t)$ 就是线上的随机游动。

例题 1.5 Brown 运动 英国植物学家 Brown 注意到悬浮在液面上的微小粒子不断进行无规则的运动, 这种运动后来称为 Brown 运动。它是分子大量随机碰撞的结果。若以 $X(t), Y(t)$ 为粒子在平面坐标上的位置, 则它又平面上的 Brown 运动。

例题 1.6 排队模型 顾客来到服务站要求服务。当服务站中的服务员都忙碌, 即服务员都在为别的顾客服务时, 来到的顾客就要排队等候。顾客的到来、每个顾客所需的服务时间都是随机的, 所以如果用 $X(t)$ 表示时刻的队长, 用 $Y(t)$ 表示时刻到来的顾客所需的等待时间, 则 $\{X(t), t \in T\}, \{Y(t), t \in T\}$ 都是随机过程

1.9 有限维分布与 Kolmogorov 定理

设 (Ω, \mathcal{F}, P) 是概率空间, $X = \{X_t, t \geq 0\}$ 是其上的以 (E, \mathcal{E}) 为状态空间的随机过程, 用 \mathcal{F}_T 表示 T 的有序的有限子集全体, 即

$$\mathcal{F}_T := \{(t_1, \dots, t_n) : n \geq 1, t_1, \dots, t_n \in T\},$$

$\mathcal{N}_T = (t_1, \dots, t_n) \in \mathcal{F}_T$, 记乘积空间 E^n 为 E^T , 定义 X_T 为下列映射

$$\omega \mapsto (X_{t_1}(\omega), \dots, X_{t_n}(\omega)),$$

则 X_T 是 Ω 到 E^T 的可测映射。用 μ_T 表示空间 (E^T, \mathcal{E}^T) 上由 X_T 诱导的测度 $P \circ X_T^{-1}$, 即 P 在 X_T 下的像测度。

即对任何 $A_1, \dots, A_n \in \mathcal{E}$,

$$\mu_{t_1, \dots, t_n}(A_1 \times \dots \times A_n) = P(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n).$$

用记 $D_X = \{\mu_T : t \in \mathcal{F}_T\}$, 它称为随机过程 X 的有限维分布族。

给定随机过程 X , 其有限维分布族由过程唯一决定且满足下面的性质:

分布族的性质:

1. 对称性: 对 $(1, 2, \dots, n)$ 的任一排列 $\sigma = (\sigma_1, \dots, \sigma_n) = (j_1, j_2, \dots, j_n)$, 有

$$\mu_{t_1, \dots, t_n}(F_{j_1}, \dots, F_{j_n}) = \mu_{t_{\sigma_1}, \dots, t_{\sigma_n}}(F_1, \dots, F_n).$$

2. 相容性: 对于 $m < n$, 有

$$\mu_{t_1, \dots, t_m}(F_1, \dots, F_m, \mathbb{R}^n, \dots, \mathbb{R}^n) = \mu_{t_1, \dots, t_n}(F_1, \dots, F_m).$$

定理 1.9.1

设分布族 $\{\mu_{t_1, \dots, t_n}, t_1, \dots, t_n \in T, n \geq 1\}$ 满足上述的对称性和相容性, 则必存在一个概率空间 (Ω, \mathcal{F}, P) 及上的随机过程 $\{X(t), t \in T\}$, 使

$$\mu_{t_1, \dots, t_n}(F_1, \dots, F_n) = P(X(t_1) \in F_1, \dots, X(t_n) \in F_n).$$



注 随机过程的有限维分布函数族是随机过程概率特征的完整描述，它是证明随机过程存在性的有力工具。但是在实际问题中，要知道随机过程的全部有限维分布是不可能的，因此，人们想到了用随机过程的某些数字特征来刻画随机过程。

定义 1.9.1

设 $\{X(t), t \in T\}$ 是一随机过程。

1. 称 $X(t)$ 的期望 $\mu_X(t) = E[X(t)]$ 为过程的均值函数（如果存在的话）。
2. 如果 $\forall t \in T, E[X^2(t)]$ 存在，则称随机过程 $\{X(t), t \in T\}$ 为二阶矩过程。

函数 $\gamma(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))]$, $t_1, t_2 \in T$ 为过程的协方差函数；称 $\text{Var}[X(t)] = \gamma(t, t)$ 为过程的方差函数。

1.10 随机过程的基本类型

1.10.1 平稳过程

定义 1.10.1

称 \mathcal{F} 的子 σ -域序列 $\{F_t, t \in \mathbb{R}_+\}$ 为一个流，若 $\forall 0 \leq s < t$ 有 $F_s \subseteq F_t$ 。记 $F_\infty = \bigvee_{t \geq 0} F_t$,

$$F_t = \bigcap_{s > t} F_s, \quad t \geq 0, \quad F_\infty = \bigvee_{s < t} F_s = \sigma \left(\bigcup_{s < t} F_s \right), \quad t \geq 0.$$

一个流 $\{F_t, t \in \mathbb{R}_+\}$ 称为右连续的，若 $\forall t \geq 0$ 有 $F_t = F_{t+}$ 。

定义 1.10.2

如果随机过程 $\{X(t), t \in T\}$ 对任意的 $t_1, \dots, t_n \in T$ 和任意的 h (使得 $t_n + h \in T$) 有，

$$(X(t_1 + h), \dots, X(t_n + h)) \stackrel{d}{=} (X(t_1), \dots, X(t_n))$$

具有相同的联合分布，记为

$$(X(t_1 + h), \dots, X(t_n + h)) \stackrel{d}{=} (X(t_1), \dots, X(t_n))$$

则称 $\{X(t), t \in T\}$ 为严平稳的。

当参数 t 仅取整数值 $0, \pm 1, \pm 2, \dots$ 或 $0, 1, 2, \dots$ 时，称平稳过程为平稳序列。

1.10.2 独立增量过程

定义 1.10.3

设 $X = \{X(t, \cdot), t \in \mathbb{R}_+\}$ 为随机过程，对每个固定的 t , $X(t, \cdot)$ 为一个随机变量；而对每个固定的 $\omega \in \Omega$, $X(\cdot, \omega)$ 是关于 $t \in T$ 的映射，称 $X(\cdot, \omega)$ 为样本轨道（或样本函数，一个实现）。若 X 的全部轨道是连续的（右连续的，左连续的），则称 X 为连续（右连续，左连续）过程。如果 X 的全部轨道是右连续并有左极限的（简称右连左极）。

定义 1.10.4

随机过程 $X = \{X_t, t \geq 0\}$ 为 $(\mathcal{F}_t, t \in \mathbb{R}_+)$ 适应的，若 $\forall t$, X_t 为 \mathcal{F}_t 可测的。

定义 1.10.5

在带流的概率空间 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ 上的适应过程 $X = \{X_t, t \geq 0\}$ 称为具有独立增量性, 如果对任意的 $t > s \geq 0$, $X_t - X_s$ 与 \mathcal{F}_s 独立。



第 2 章 Poisson 过程

2.1 Poisson 过程的定义

假设 r.v. 定义在 (Ω, \mathcal{F}, P) 。

定义 2.1.1

设 $T \subset (-\infty, +\infty)$ 。若对每个 $t \in T$, $X(t, \cdot)$ 是随机变量, 则称随机变量集 $X_T := \{X(t, \cdot) | t \in T\}$ 为随机过程, 称 T 为该随机过程的指标集。

用映射表示为:

$$X_T : (t, \omega) \mapsto X(t, \omega).$$

对每个固定的 t , $X(t, \cdot)$ 为一个随机变量; 而对每个固定的 $\omega \in \Omega$, $X(\cdot, \omega)$ 是关于 $t \in T$ 的映射, 称 $X(\cdot, \omega)$ 为样本轨道 (或样本函数, 一个实现)。



定义 2.1.2

若 $\{N(t), t \geq 0\}$ 表示 $[0, t]$ 内某事件发生次数, 则称 $\{N(t); t \geq 0\}$ 是计数过程, 或称为 $\{N(t)\}$, 满足:

1. 对 $\forall t \geq 0$, $N(t)$ 取非负整数值 $N(t) \in \mathbb{Z} := \{0, 1, \dots\}$.
2. 对 $t > s \geq 0$, $N(t) \geq N(s)$,
3. $\{N(t)\}$ 轨道单调不减右连续阶梯函数。



定义 2.1.3

计数过程 $\{N(t), t \geq 0\}$ 称为参数为 $\lambda (\lambda > 0)$ 的 Poisson 过程, 如果

1. $N(0) = 0$;
2. 过程有独立增量;
3. 对任意的 $s, t \geq 0$,

$$P(N(t+s) - N(s) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots$$



注

1. Poisson 过程是独立平稳增量的计数过程;
2. 由于 $E[N(t)] = \lambda t$, 于是可认为 λ 是单位时间内发生的事件的平均次数, 故一般称 λ 是 Poisson 过程的强度或速率。

例题 2.1 Poisson 过程在排队论中的应用 在随机服务系统中排队现象的研究中, 经常用到 Poisson 过程模型, 例如, 到达电话总机的呼叫数目, 到达某服务设施的顾客数, 都可以用 Poisson 过程来描述, 以某火车站售票处为例, 设从早上 8:00 开始, 此售票处连续售票, 乘客依 10 人/小时的平均速率到达, 则从 9:00 到 10:00 这 1 小时内最多有 5 名乘客来此购票的概率是多少? 从 10:00-11:00 没有人来买票的概率是多少?

证明 我们用一个 Poisson 过程来描述。设 8:00 为 0 时刻, 则 9:00 为 1 时刻, 参数 $\lambda = 10$ 。由 Poisson 过程的平稳性知

$$P(N(2) - N(1) \leq 5) = \sum_{n=0}^5 e^{-10} \cdot \frac{(10 \cdot 1)^n}{n!},$$
$$P(N(3) - N(2) = 0) = e^{-10} \cdot \frac{(10)^0}{0!} = e^{-10}.$$

为什么实际中有这么多的现象可以用 Poisson 过程来反映呢? 其根据是小概率事件原理。我们在概率论的学习中已经知道, Bernoulli 试验中, 每次试验成功的概率很小而试验的次数很多时, 二项分布会逼近 Poisson 分

布。这一想法很自然地推广到随机过程情况。比如上面提到的事故发生的例子，在很短的时间内发生事故的概率是很小的，但假如考虑很多个这样很短的时间的连接，事故的发生将会有一个大致稳定的速率，这很类似于 Bernoulli 试验以及二项分布逼近 Poisson 分布时的假定。

Poisson 过程的另一等价定义：

定义 2.1.4

设 $\{N(t), t \geq 0\}$ 是一个计数过程，若满足

1. $N(0) = 0$;
2. 过程有平稳独立增量;
3. 存在 $\lambda > 0$, 当 $h \downarrow 0$ 时

$$P(N(t+h) - N(t) = 1) = \lambda h + o(h);$$

4. 当 $h \downarrow 0$ 时,

$$P(N(t+h) - N(t) \geq 2) = o(h).$$

则称 $\{N(t), t \geq 0\}$ 为 Poisson 过程。



事实上，把 $[0, t]$ 划分为 n 个相等的时间区间，则由条件 (4) 可知，当 $n \rightarrow \infty$ 时，在每个小区间内事件发生两次或两次以上的概率趋于 0，因此，事件发生一次的概率

$$p \approx \lambda \frac{t}{n} \text{ (显然 } p \text{ 会很小) 事件发生的概率为 } 1 - p \approx 1 - \lambda \frac{t}{n}$$

这恰好是一次 Bernoulli 试验。其中事件发生一次即为试验成功，不发生即为失败，再由条件 (2) 给出的平稳独立增量性， $N(t)$ 就相当于 n 次独立 Bernoulli 试验中试验成功的总次数，由 Poisson 分布的二项分布逼近可知 $N(t)$ 将服从参数为 λt 的 Poisson 分布。

Poisson 过程两定义等价的严格的数学证明：

定理 2.1.1

满足上述条件 (1)–(4) 的 $\{N(t), t \geq 0\}$ 是 Poisson 过程，反过来 Poisson 过程一定满足这 4 个条件。



第3章 离散时间马尔可夫链

3.1 马氏链及其转移概率

有一类随机过程，它具备所谓的“无后效性”(Markov 性)，即要确定过程将来的状态，知道它此刻的情况就足够了，并不需要对它以往状况的认识，这类过程称为 Markov 过程. 我们将介绍离散时间的 Markov 链（简称马氏链）。

本章假定： $T = \{0, 1, \dots\}$ ， $S = \{0, 1, 2, \dots, N\}$ （或者 $S := \mathbb{N}$ ），所有 r.v. 均定义在同一个概率空间上。用 i, j 表示 S 中元素。

定义 3.1.1 (离散时间马尔可夫链)

随机过程 $\{X_n, n = 0, 1, 2, \dots\}$ 称为 Markov 链，若它只取有限或可列个值（若无特别说明，通常取非负整数集 $\{0, 1, 2, \dots\}$ ），并且对任意 $n \geq 0$ 及任意状态 $i_0, i_1, \dots, i_{n-1}, i, j$ ，有

$$P\{X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i\} = P\{X_{n+1} = j | X_n = i\} \quad (3.1)$$

其中 $X_n = i$ 表示过程在时刻 n 处于状态 i ，称为 S 。式 (1.1) 刻画了 Markov 链的特性，称为 Markov 性，或马氏性，或无记忆性。

定义 3.1.2 (转移概率)

设 $\{X_n, n = 0, 1, \dots\}$ 为马氏链，称

$$P\{X_{n+1} = j | X_n = i\} =: p_{ij}(n)$$

为 n 时刻的一步转移概率。若它与 n 无关，则记作 p_{ij} ，并称相应的马氏链为齐次的或时齐的。令 $P = (p_{ij})$ ，称 P 为齐次马氏链的转移概率矩阵，简称为转移矩阵， p_{ij} 为一步转移概率。我们只考虑齐次马氏链。

$$\begin{aligned} P\{X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i\} \\ = P\{X_{n+1} = j | X_n = i\} \text{ 马尔可夫性} \\ = P\{X_1 = j | X_0 = i\} \text{ 齐次} \end{aligned}$$

设 $\{X_n, n = 0, 1, \dots\}$ 是齐次马氏链，具有转移矩阵 $P = (p_{ij})$ ，则有

$$\begin{aligned} p_{ij} \geq 0 \quad \forall i, j \in S \text{ 且} \\ \sum_{j \in S} p_{ij} = \sum_{j \in S} P(X_1 = j | X_0 = i) = P(X_1 \in S | X_0 = i) = 1 \quad \forall i \in S. \end{aligned}$$

定义 3.1.3 (随机矩阵)

称矩阵 $A = (a_{ij})_{S \times S}$ 为随机矩阵，若 $a_{ij} \geq 0 (\forall i, j \in S)$ ，且 $\sum_{j \in S} a_{ij} = 1 (\forall i \in S)$ 。

由该定义知转移矩阵是随机矩阵。

例题 3.1 赌徒破产问题 系统的状态是 $0 \sim n$ ，反映赌博者在赌博期间拥有的钱数，当他输光或拥有钱数为 n 时，赌博停止，否则他将持续赌博。每次以概率 p 赢得 1，以概率 $q = 1 - p$ 输掉 1。则每个时刻，该赌徒拥有的钱数服从马尔可夫性吗？能否写出对应的转移概率矩阵？

证明 这个系统的转移矩阵为

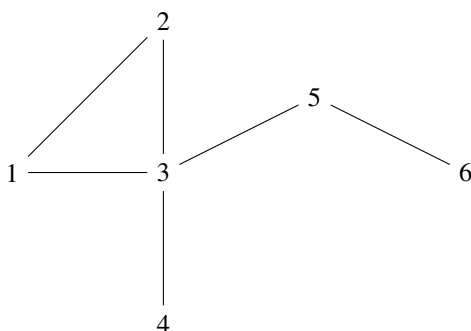
$$P = \begin{array}{c|cccccccc} & 0 & 1 & 2 & 3 & \cdots & n-2 & n-1 & n \\ \hline 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 2 & 0 & q & 0 & p & \cdots & 0 & 0 & 0 \\ 3 & 0 & 0 & q & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ n-2 & 0 & 0 & 0 & 0 & \cdots & 0 & p & 0 \\ n-1 & 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ n & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{array}$$

例题 3.2 简单随机游动 质点在直线的整数点上作简单随机游动：质点到达某个状态后，下次向右移动一步的概率是 p ，向左移动一步的概率是 q ，在原地不动的概率为 r ，且 $p + q + r = 1$ 。 X_0 表示初始状态， X_n 表示质点在时间 n 的状态。假设初始状态与每次移动相互独立。则 $\{X_n\}$ 是马氏链，

证明

$$\begin{cases} p_{i,i-1} = P(X_{n+1} = i-1 | X_n = i) = q \\ p_{i,i+1} = P(X_{n+1} = i+1 | X_n = i) = p \\ p_{i,i} = P(X_{n+1} = i | X_n = i) = r \end{cases}$$

例题 3.3 设有一蚂蚁在下图爬行，当两个结点相临时，蚂蚁将爬向它临近的一点，并且爬向任何一个邻居的概率是相同的。



证明 此 Markov 链的转移矩阵为

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

定理 3.1.1

设 A, B, C 为三个随机事件，则 $P(BC|A) = P(B|A)P(C|AB)$.



证明

$$P(BC|A) = \frac{P(ABC)}{P(A)} = \frac{P(AB)P(C|AB)}{P(A)P(AB)} = P(B|A)P(C|AB).$$

注 令 $P(\cdot|A) := P_A$ ，则应用乘法公式 $P(BC|A) = P_A(BC) = P_A(C|B) \cdot P_A(B) = P(C|BA)P(B|A)$.

定理 3.1.2

对于事件 A, B, C , 当 $P(AB) > 0$, 条件

$$P(C|BA) = P(C|B),$$

和条件

$$P(AC|B) = P(A|B)P(C|B)$$

等价。



证明

$$\frac{P(ACB)}{P(B)} = \frac{P(AB)}{P(B)} \frac{P(BC)}{P(AB)} \text{ 可知 } \frac{P(ACB)}{P(AB)} = \frac{P(BC)}{P(B)}. \text{ 即 } P(C|BA) = P(C|B).$$

定理 3.1.3

对于事件 A, B, C , 当 $P(AB) > 0$, 条件

$$P(C|BA) = P(C|B),$$

和条件

$$P(AC|B) = P(A|B)P(C|B)$$

等价。



马氏性的解释:

过去: $A = (X_0 = i_0, \dots, X_{n-1} = i_{n-1})$,

现在: $B = (X_n = i_n)$,

将来: $C = (X_{n+1} = i_{n+1})$ 。

马氏性代表在已知现在的情况下, 将来与过去无关。

定理 3.1.4

设 S 是马氏链 $\{X_n\}$ 的状态空间, 则有

1. 对任意的 $n, m \geq 1$ 有

$$\begin{aligned} P(X_{n+1} = i_{n+1}, \dots, X_{n+m} = i_{n+m} | X_0 = i_0, \dots, X_n = i) \\ = P(X_{n+1} = i_{n+1}, X_{n+2} = i_{n+2}, \dots, X_{n+m} = i_{n+m} | X_n = i) \end{aligned}$$

2. 对任意的 $n, m \geq 1$, 以及 $C \subset S^m, A \subset S^n$ 有

$$\begin{aligned} P((X_{n+1}, X_{n+2}, \dots, X_{n+m}) \in C | (X_0, \dots, X_{n-1}) \in A, X_n = i) \\ = P((X_{n+1}, X_{n+2}, \dots, X_{n+m}) \in C | X_n = i) \end{aligned}$$

3. 对任意的 $k, m \geq 1$, 以及 $t_0 < t_1 < \dots < t_k < t_{k+1} < \dots < t_{k+m}, i \in S, C \subset S^m, A \subset S^k$ 有

$$\begin{aligned} P((X_{t_{k+1}}, X_{t_{k+2}}, \dots, X_{t_{k+m}}) \in C | (X_{t_0}, \dots, X_{t_k}) \in A, X_{t_k} = i) \\ = P((X_{t_{k+1}}, X_{t_{k+2}}, \dots, X_{t_{k+m}}) \in C | X_{t_k} = i) \end{aligned}$$



证明 (1) 对 m 进行归纳证明。当 $m = 1$ 时有马氏性即得。现在设对 $m = k$ 成立, 即已知

$$\begin{aligned} P(X_{n+1} = i_{n+1}, \dots, X_{n+k} = i_{n+k} | X_0 = i_0, \dots, X_n = i) \\ = P(X_{n+1} = i_{n+1}, X_{n+2} = i_{n+2}, \dots, X_{n+k} = i_{n+k} | X_n = i) \end{aligned}$$

$$\begin{aligned}
& P(X_{n+1} = i_{n+1}, \dots, X_{n+k+1} = i_{n+k+1} | X_0 = i_0, \dots, X_n = i) \\
&= P(X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_0 = i_0, \dots, X_{n+1} = i_{n+1}) \\
&\quad \cdot P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i) \\
&= P(X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_{n+1} = i_{n+1}) \\
&\quad \cdot P(X_{n+1} = i_{n+1} | X_n = i) \\
&= P(X_{n+1} = i_{n+1}, X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_n = i) \\
&= P(X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_n = i, X_{n+1} = i_{n+1}) \\
&= \sum_{i_{n-1} \in S, \dots, i_0 \in S} P(X_{n+k+1} = i_{n+k+1}, \dots, X_{n+2} = i_{n+2}, X_{n+1} = i_{n+1}, \\
&\quad X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) / P(X_{n+1} = i_{n+1}, X_n = i) \\
&= \sum_{i_{n-1} \in S, \dots, i_0 \in S} P(X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_{n+1} = i_{n+1}, \\
&\quad X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) P(X_{n+1} = i_{n+1}, \\
&\quad X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) / P(X_{n+1} = i_{n+1}, X_n = i) \\
&= \sum_{i_{n-1} \in S, \dots, i_0 \in S} P(X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_{n+1} = i_{n+1}) \\
&\quad P(X_{n+1} = i_{n+1}, X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&\quad / P(X_{n+1} = i_{n+1}, X_n = i) \\
&= P(X_{n+2} = i_{n+2}, \dots, X_{n+k+1} = i_{n+k+1} | X_{n+1} = i_{n+1})
\end{aligned}$$

(2)

$$\begin{aligned}
& P((X_{n+1}, \dots, X_{n+m}) \in C | (X_0, \dots, X_{n-1}) \in A, X_n = i) \\
&= \frac{P((X_{n+1}, \dots, X_{n+m}) \in C, X_n = i, (X_{n-1}, \dots, X_0) \in A)}{P((X_0, \dots, X_{n-1}) \in A, X_n = i)} \\
&= \sum_{(i_{n+1}, i_{n+2}, \dots, i_{n+m}) \in C} \sum_{(i_{n-1}, \dots, i_0) \in A} P(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1}, \\
&\quad X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) / P(X_n = i, (X_{n-1}, \dots, X_0) \in A) \\
&= \sum_{(i_{n+1}, i_{n+2}, \dots, i_{n+m}) \in C} \sum_{(i_{n-1}, \dots, i_0) \in A} P(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1} | X_n = i, \\
&\quad X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \cdot P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&\quad / P(X_n = i, (X_{n-1}, \dots, X_0) \in A) \\
&= \sum_{(i_{n+1}, i_{n+2}, \dots, i_{n+m}) \in C} P(X_{n+m} = i_{n+m}, \dots, X_{n+1} = i_{n+1} | X_n = i) \\
&\quad \cdot \sum_{(i_{n-1}, \dots, i_0) \in A} P(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\
&\quad / P(X_n = i, (X_{n-1}, \dots, X_0) \in A) \\
&= P((X_{n+1}, X_{n+2}, \dots, X_{n+m}) \in C | X_n = i)
\end{aligned}$$

(3) 用增补变量的方法: 只举特殊情形证明, 其余类似. 证明 $P((X_7, X_5) \in C | X_3 = i, X_1 \in A) = P((X_7, X_5) \in C | X_3 = i)$

令 $\tilde{C} := \{(i_7, i_6, i_5, i_4) | (i_7, i_5) \in C, i_6 \in S, i_4 \in S\}$, $\tilde{A} := \{(i_2, i_1, i_0) | i_2 \in S, i_1 \in A, i_0 \in S\}$

$$\begin{aligned}
P((X_7, X_5) \in C | X_3 = i, X_1 \in A) \\
&= P((X_7, X_6, X_5, X_4) \in \tilde{C} | X_3 = i, (X_2, X_1, X_0) \in \tilde{A}) \\
&= P((X_7, X_6, X_5, X_4) \in \tilde{C} | X_3 = i) \\
&= P((X_7, X_5) \in C | X_3 = i)
\end{aligned}$$

注 设 S 是马氏链 $\{X_n\}$ 的状态空间, 对任意的 $k, m \geq 1$, 以及 $t_0 < t_1 < \dots < t_k < t_{k+1} < \dots < t_{k+m}$, $B \subset S$, $A \subset S^k$, $C \subset S^m$, 有

$$P((X_{t_{k+1}}, X_{t_{k+2}}, \dots, X_{t_{k+m}}) \in C | X_{t_k} \in B, (X_{t_{k-1}}, \dots, X_{t_0}) \in A) \neq P((X_{t_{k+1}}, X_{t_{k+2}}, \dots, X_{t_{k+m}}) \in C | X_{t_k} \in B)$$

注 看不懂也没关系, 不影响后面的学习

例题 3.4 质点在直线的整数点上作简单随机游动: 质点到达某个状态后, 下次向右移动一步的概率是 p , 向左移动一步的概率是 $q = 1 - p$. X_0 表示初始状态, X_n 表示质点在时间 n 的状态. 假设初始状态与每次移动相互独立. 则 $\{X_n\}$ 是马氏链,

$$\begin{cases} p_{i,i-1} = P(X_{n+1} = i-1 | X_n = i) = q \\ p_{i,i+1} = P(X_{n+1} = i+1 | X_n = i) = p \end{cases}$$

设初始分布 $P(X_0 = 0) = P(X_0 = 2) = \frac{1}{2}$, $D := \{1, 3\}$. 证明: 当 $p \neq q$ 时, $P(X_2 = 2 | X_0 = 0, X_1 \in D) \neq P(X_2 = 2 | X_1 \in D)$.

证明

$$\begin{aligned}
P(X_1 = 1) &= P(X_0 = 0)P(X_1 = 1 | X_0 = 0) + P(X_0 = 2)P(X_1 = 1 | X_0 = 2) = \frac{1}{2}p + \frac{1}{2}q = \frac{1}{2}, \\
P(X_1 = 3) &= P(X_1 = 3, X_0 = 2) + P(X_1 = 3, X_0 = 0) = P(X_0 = 2)P(X_1 = 3 | X_0 = 2) = \frac{1}{2}p.
\end{aligned}$$

故

$$\begin{aligned}
&P(X_2 = 2 | X_1 \in D) \\
&= \frac{P(X_2 = 2, X_1 \in D)}{P(X_1 \in D)} \\
&= \frac{P(X_2 = 2, X_1 = 1) + P(X_2 = 2, X_1 = 3)}{P(X_1 \in D)} \\
&= \frac{P(X_1 = 1)P(X_2 = 2 | X_1 = 1) + P(X_1 = 3)P(X_2 = 2 | X_1 = 3)}{P(X_1 = 1) + P(X_1 = 3)} \\
&= \frac{\frac{1}{2}p + \frac{1}{2}pq}{\frac{1}{2} + \frac{1}{2}p} = p \frac{1+q}{1+p}
\end{aligned}$$

$$\begin{aligned}
&P(X_2 = 2 | X_1 \in D, X_0 = 0) \\
&= \frac{P(X_2 = 2, X_1 \in D, X_0 = 0)}{P(X_1 \in D, X_0 = 0)} \\
&= \frac{P(X_2 = 2, X_1 = 1, X_0 = 0)}{P(X_1 \in D, X_0 = 0)} \\
&= \frac{P(X_2 = 2 | X_1 = 1, X_0 = 0)}{P(X_1 \in D, X_0 = 0)} = p
\end{aligned}$$

$P(X_2 = 2 | X_0 = 0, X_1 \in D) = P(X_2 = 2 | X_0 = 0, X_1 = 1) = p$. 当 $p \neq q$ 时, $P(X_2 = 2 | X_0 = 0, X_1 \in D) \neq P(X_2 = 2 | X_1 \in D)$.

定理 3.1.5

设随机过程 $\{X_n, n \geq 0\}$ 满足:

1. $X_n = f(X_{n-1}, \xi_n) (n \geq 1)$, 其中 $f: S \times S \rightarrow S$, 且 ξ_n 取值在 S 上,

2. $\{\xi_n, n \geq 1\}$ 为独立同分布随机变量, 且 X_0 与 $\{\xi_n, n \geq 1\}$ 也相互独立, 则 $\{X_n, n \geq 0\}$ 是马尔可夫链, 而且其一步转移概率为

$$p_{ij} = P(f(i, \xi_1) = j).$$



注 这个定理讲的是如何生成一个马尔可夫链。

例题 3.5 质点在直线的整数点上作简单随机游动: 质点到达某个状态后, 下次向右移动一步的概率是 p , 向左移动一步的概率是 $q = 1 - p$. X_0 表示初始状态, X_n 表示质点在时间 n 的状态。假设初始状态与每次移动相互独立。证明 $\{X_n\}$ 是马氏链, 且

$$\begin{cases} p_{i,i-1} = P(X_{n+1} = i-1 | X_n = i) = q \\ p_{i,i+1} = P(X_{n+1} = i+1 | X_n = i) = p \end{cases}$$

证明 令 X_n 为质点在时刻 $n \geq 0$ 的位置,

$$\xi_n := \begin{cases} 1 & \text{第 } n \text{ 次向右移动} \\ -1 & \text{第 } n \text{ 次向左移动} \end{cases}$$

$$X_n = X_{n-1} + \xi_n.$$

转移概率:

$$p_{ij} = P(i + \xi_1 = j) = P(\xi_1 = j - i) = \begin{cases} q & j = i - 1 \\ p & j = i + 1 \end{cases}$$

注 这道例题就是讲述如何生成一个马尔可夫链。

设 $\{X_n, n = 0, 1, \dots\}$ 是齐次马氏链, 具有转移矩阵 $P = (p_{ij})$, 则有

$$p_{ij} \geq 0 \quad \forall i, j \in S \text{ 且 } \sum_{j \in S} p_{ij} = 1 \quad \forall i \in S.$$

3.2 转移概率矩阵

定义 3.2.1 (随机矩阵)

称矩阵 $A = (a_{ij})_{S \times S}$ 为随机矩阵, 若 $a_{ij} \geq 0 (\forall i, j \in S)$, 且 $\sum_{j \in S} a_{ij} = 1 (\forall i \in S)$.



注 随机矩阵就是转移矩阵。

特别地, 记 $P^0 = I$ (S 上的单位矩阵),

$$p_{ij}^{(0)} = \delta_{ij} = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$$

且

$$p_{ij}^{(n)} := P(X_n = j | X_0 = i) = P(X_{n+m} = j | X_m = i) \quad \text{与 } m \text{ 无关!}$$

表示从 i 出发经 n 步到达 j 的概率。称 $P^{(n)} := (p_{ij}^{(n)})_{i,j \in S}$ 为 $\{X_n\}$ 的 n 步转移概率矩阵。显然 $P^{(n)}$ 为随机矩阵。

定理 3.2.1 (Chapman-Kolmogorov 方程)

设 $\{X_n\}$ 是齐次马氏链, 具有转移矩阵 P , 则对任意的 $m, n \geq 0$, 有

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)} \quad \forall i, j \in S, m, n \geq 0.$$

$$P^{(n+m)} = P^{(m)} P^{(n)} = P^{n+m}, \text{ 其中 } P^{n+m} \text{ 表示矩阵 } P \text{ 的 } n+m \text{ 次乘积}$$

证明

$$\begin{aligned} p_{ij}^{(n+m)} &= P(X_{n+m} = j | X_0 = i) = P(X_{n+m} = j, X_n \in S | X_0 = i) \\ &= \sum_{k \in S} P(X_n = k, X_{n+m} = j | X_0 = i) \\ &= \sum_{k \in S} P(X_n = k | X_0 = i) P(X_{n+m} = j | X_0 = i, X_n = k) \\ &= \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}. \end{aligned}$$

从状态 i 出发经 $n+m$ 步到达 j 的概率可以由转移矩阵 P 及归纳法证明。

定理 3.2.2

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)} \quad \forall i, j \in S, m, n \geq 0.$$

$$P^{(n+m)} = P^{(m)} P^{(n)} = P^{n+m}, \text{ 其中 } P^{n+m} \text{ 表示矩阵 } P \text{ 的 } n+m \text{ 次乘积}$$

推论 3.2.1

对任意的正整数 $n, m, k, n_1, n_2, \dots, n_k$ 和状态 i, j, l , 有

1. $p_{ij}^{(n+m)} \geq p_{il}^{(n)} p_{lj}^{(m)}$;
2. $p_{ii}^{(n+m+k)} \geq p_{ij}^{(n)} p_{jl}^{(k)} p_{li}^{(m)}$;
3. $p_{ii}^{(n_1+n_2+\dots+n_k)} \geq p_{ii}^{(n_1)} p_{ii}^{(n_2)} \dots p_{ii}^{(n_k)}$;
4. $p_{ii}^{(nk)} \geq (p_{ii}^{(n)})^k$.

证明

$$P^{(n)} = P^{(n-1)} P = P^{(n-2)} P \cdot P = \dots = P^n.$$

例题 3.6 已知马氏链的一步转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$$

求 $P^{(2)}, P^{(4)}$ 。

证明

$$\mathbf{P}^{(2)} = \mathbf{P} \cdot \mathbf{P} = \begin{pmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{pmatrix}$$

$$\mathbf{P}^{(4)} = \mathbf{P}^{(2)} \cdot \mathbf{P}^{(2)} = \begin{pmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{pmatrix}$$

例题 3.7 系统的状态是 $0 \sim n$, 反映赌博者在赌博期间拥有的钱数, 当他输光或拥有钱数为 n 时, 赌博停止, 否

则他将持续赌博。每次以概率 p 赢得 1, 以概率 $q = 1 - p$ 输掉 1。这个系统的转移矩阵为

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}_{(n+1) \times (n+1)}$$

$n = 3, p = q = \frac{1}{2}$ 。赌博者从 2 元赌金开始赌博, 求解他经过 4 次赌博之后输光的概率。

证明 这个概率为 $p_{20}^{(4)} = P\{X_4 = 0 | X_0 = 2\}$, 一步转移矩阵为

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

利用矩阵乘法得

$$\mathbf{P}^{(4)} = \mathbf{P}^4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{5}{16} & \frac{1}{16} & 0 & \frac{5}{16} \\ \frac{5}{16} & 0 & \frac{1}{16} & \frac{5}{8} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

故 $p_{20}^{(4)} = \frac{5}{16}$ ($\mathbf{P}^{(4)}$ 中第 3 行第 1 列)。

例题 3.8 甲乙两人进行某种比赛, 设每局甲胜的概率是 p , 乙胜的概率是 q , 和局的概率是 r , $p + q + r = 1$ 。设每局比赛后, 胜者记“+1”分, 负者记“-1”分, 和局不记分, 且当两人中有一人获得 2 分时结束比赛。以 X_n 表示比赛至第 n 局时甲获得的分数, 则 $\{X_n, n = 0, 1, 2, \cdots\}$ 为时齐 Markov 链, 求在甲获得 1 分的情况下, 不超过两局可结束比赛的概率。

证明 $\{X_n, n = 0, 1, 2, \cdots\}$ 的一步转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & q & r & p & 0 & 0 & 0 \\ 0 & 0 & q & r & p & 0 & 0 \\ 1 & 0 & 0 & 0 & q & r & p \\ 2 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

两步转移概率矩阵为

$$\mathbf{P}^{(2)} = \mathbf{P} \cdot \mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ q + rq & r^2 + pq & 2pr & p^2 & 0 & 0 & 0 \\ q^2 & 2rq & r^2 + 2pq & 2pr & p^2 & 0 & 0 \\ 0 & q^2 & 2qr & r^2 + pq & p + pr & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

故在甲获得 1 分的情况下, 不超过两局可结束比赛的概率为

$$p_{1,2}^{(2)} + p_{1,-2}^{(2)} = p + pr$$

例题 3.9 质点在直线的整数点上作简单随机游动: 质点到达某个状态后, 下次向右移动一步的概率是 p , 向左移动一步的概率是 $q = 1 - p$ 。 X_0 表示初始状态, X_n 表示质点在时间 n 的状态。假设初始状态与每次移动相互独立。则 $\{X_n\}$ 是马氏链, 求 $P^{(n)}$ 。

$$\begin{cases} p_{i,i-1} = P(X_{n+1} = i-1 | X_n = i) = q \\ p_{i,i+1} = P(X_{n+1} = i+1 | X_n = i) = p \end{cases}$$

证明 从 i 经过 n 步到 j , 其中向左走了 x 步, 向右走了 y 步。则有

$$\begin{cases} x + y = n \\ i - x + y = j \end{cases}$$

故 $x = \frac{n-(j-i)}{2}$ 且 $y = \frac{n+j-i}{2}$, 其中 $n+j-i$ 必须是偶数 $(n+j-i) + [n-(j-i)] = 2n$ 。

$$p_{ij}^{(n)} = \begin{cases} C_n^{(n+j-i)/2} p^{(n+j-i)/2} q^{(n-j+i)/2} & n+j-i \text{ 为偶数} \\ 0 & n+j-i \text{ 为奇数} \end{cases}$$

回顾 $p_{ij}^{(n)} = P(X_n = j | X_0 = i)$ 与 $P(X_n = j)$

$$\pi_i(n) = P(X_n = i), i \in \mathcal{S},$$

$$\pi(n) = (\pi_i(n), i \in \mathcal{S}).$$

即 $\pi(n)$ 表示 n 时刻 X_n 的概率分布, 称 $\pi(0) := (\pi_i(0), i \in \mathcal{S})$ 为马氏链 $\{X_n, n = 0, 1, \dots\}$ 的初始分布。

对任意的 $n \geq 0$,

$$\sum_{i \in \mathcal{S}} \pi_i(n) = P(X_n \in \mathcal{S}) = 1.$$

定理 3.2.3

$$\pi(n+1) = \pi(n)\mathbf{P}, \quad \pi(n) = \pi(0)\mathbf{P}^n,$$

其中 \mathbf{P}^n 是 \mathbf{P} 的 n 次幂, 特别有 $\pi(n) = \pi(k)\mathbf{P}^{n-k}, 0 \leq k \leq n$.

证明

$$\begin{aligned} \pi_j(n+1) &= P(X_{n+1} = j) = P(X_n \in \mathcal{S}, X_{n+1} = j) \\ &= \sum_{i \in \mathcal{S}} P(X_n = i, X_{n+1} = j) \\ &= \sum_{i \in \mathcal{S}} P(X_n = i) P(X_{n+1} = j | X_n = i) = \sum_{i \in \mathcal{S}} \pi_i(n) p_{ij}. \end{aligned}$$

令 $P_{:,j} := (p_{ij}, i \in \mathcal{S})'$ 为一步转移概率矩阵的第 j 列。

则 $\pi_j(n+1) = \pi(n)P_{:,j}$ 。写成向量形式, 即为 $\pi(n+1) = \pi(n)\mathbf{P}$ 。

由马氏性可得: 对任意 $i_0, i_1, \dots, i_m \in \mathcal{S}$, 和 $n_0 < n_1 < \dots < n_m$ 有

$$\begin{aligned} &P(X_{n_0} = i_0, X_{n_1} = i_1, \dots, X_{n_{m-1}} = i_{m-1}, X_{n_m} = i_m) \\ &= P(X_{n_0} = i_0, X_{n_1} = i_1, \dots, X_{n_{m-1}} = i_{m-1}) \cdot P(X_{n_m} = i_m | X_{n_0} = i_0, X_{n_1} = i_1, \dots, X_{n_{m-1}} = i_{m-1}) \\ &= P(X_{n_0} = i_0, X_{n_1} = i_1, \dots, X_{n_{m-1}} = i_{m-1}) P(X_{n_m} = i_m | X_{n_{m-1}} = i_{m-1}) \\ &= P(X_{n_0} = i_0, \dots, X_{n_{m-2}} = i_{m-2}) \cdot P(X_{n_{m-1}} = i_{m-1} | X_{n_{m-2}} = i_{m-2}, \dots, X_{n_1} = i_1, X_{n_0} = i_0) p_{i_{m-1} i_m}^{(n_m - n_{m-1})} \\ &= \dots \\ &= \pi_{i_0}(n_0) p_{i_0 i_1}^{(n_1 - n_0)} p_{i_1 i_2}^{(n_2 - n_1)} \dots p_{i_{m-1} i_m}^{(n_m - n_{m-1})}. \end{aligned}$$

从而知, 马氏链 $\{X_n, n = 0, 1, \dots\}$ 的任何有限维联合分布由转移矩阵 \mathbf{P} 及其初始分布完全确定。

例题 3.10

$$\begin{aligned}
& P(X_2 = 1, X_5 = 3, X_7 = 2) \\
&= P(X_5 = 3, X_2 = 1)P(X_7 = 2 | X_5 = 3, X_2 = 1) \\
&= P(X_2 = 1)P(X_5 = 3 | X_2 = 1)P(X_7 = 2 | X_5 = 3) \\
&= \pi_1(2)p_{13}^{(3)}p_{32}^{(2)}
\end{aligned}$$

上面讨论可知, 给定齐次马氏链, 可得一随机矩阵 \mathbf{P} , 而且该马氏链的若干转移概率均可由 \mathbf{P} 确定。反之, 若给定一随机矩阵 \mathbf{P} , 我们可否确定一齐次马氏链, 使其转移概率为给定的 \mathbf{P} 相同? 答案是肯定的!

定理 3.2.4 (马氏链的存在性)

对任给的随机矩阵 $\mathbf{P} = (p_{ij})_{S \times S}$, S 上的概率分布为 $\mu = \{\mu_i, i \in S\}$, 存在唯一的概率空间 $(\Omega, \mathcal{F}, \tilde{P})$ 及其上的随机过程 $\{X_n, n = 0, 1, \dots\}$ 使得:

1. μ 为该过程的初始分布, 即 $\tilde{P}(X_0 = i) = \mu_i, i \in S$;
2. $\{X_n, n = 0, 1, \dots\}$ 是以 \mathbf{P} 为转移矩阵的齐次马氏链, 即

$$\tilde{P}(X_{n+1} = j | X_n = i) = p_{ij} \quad \forall i, j \in S, n \geq 0.$$



3.3 状态的分类及性质

定义 3.3.1

设 S 是 Markov Chain $\{X_n\}$ 的状态空间

1. 如果 $p_{ii} = 1$, 则称 i 是吸收状态 (即可知 $p_{ii}^{(n)} \equiv 1, n \geq 1$);
2. 称状态 i 可达状态 j , 如果存在状态 $i_1, \dots, i_n \in S$, 使得 $p_{i_1 i_2} p_{i_2 i_3} \cdots p_{i_n j} > 0$ 。或等价地, 存在 $n \geq 0$, 使得 $p_{ij}^{(n)} > 0$, 则称 i 通 j , 记为 $i \rightarrow j$ 。
3. 如果 $i \rightarrow j$ 且 $j \rightarrow i$, 则称 i 与 j 是互通的, 记为 $i \leftrightarrow j$ 。若对一切 $i, j \in S$, 均有 $i \leftrightarrow j$ 成立, 则称转移矩阵 (或马氏链) 是不可约。



注 \rightarrow 具有传递性。即若 $i \rightarrow j$ 且 $j \rightarrow k$, 则 $i \rightarrow k$ 。理由: 存在 $n, m, s.t. p_{ij}^{(n)} > 0$ 且 $p_{jk}^{(m)} > 0$, 则由 $C-K$ 方程 $p_{ik}^{(n+m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0$ 。

命题 3.3.1

互通有下列关系

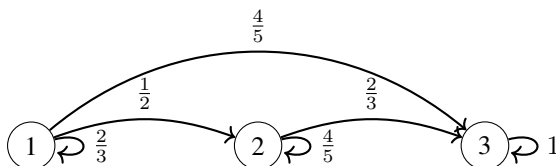
1. 对称性 $i \leftrightarrow j$, 则 $j \leftrightarrow i$;
2. 传递性 $i \leftrightarrow j, j \leftrightarrow k$, 则 $i \leftrightarrow k$;



例题 3.11 记某设备的状态为 1, 2, 3, 其中 1 表示设备运行良好, 2 表示运行正常, 3 表示设备失效。以 X_n 表示设备在时刻 n 的状态, 且假设 $\{X_n, n = 0, 1, \dots\}$ 是齐次马氏链。在有维修及更换条件下, 其转移矩阵为

$$\mathbf{P} = \begin{pmatrix} \frac{2}{3} & \frac{3}{12} & \frac{1}{12} \\ 0 & \frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 1 \end{pmatrix}$$

试说明: $1 \rightarrow 2, 2 \rightarrow 3, 3$ 是吸收的。



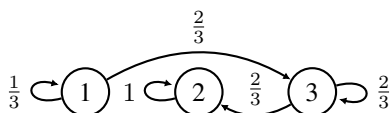
证明 如图

例题 3.12 记马氏链 $\{X_n, n \geq 0\}$ 的状态空间 $S = \{1, 2, 3\}$ 。对应的一步转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

则状态 1 是否可达状态 2?

证明



虽然 $p_{12} = 0$, 但 $p_{12}^{(2)} \geq p_{13}p_{32} = \frac{2}{9} > 0$ 。故仍然有 $1 \rightarrow 2$ 。

定义 3.3.2 (周期)

集合 $\{n, p_{ii}^{(n)} > 0\}$ 非空, $d(i)$ 为 $\{n, p_{ii}^{(n)} > 0\}$ 的最大公约数, 则称 $d(i)$ 为 i 的周期。特别地, 当 $d(i) = 1$ 时, 称 i 为非周期的。若集合 $\{n, p_{ii}^{(n)} > 0\} = \emptyset$, 则称 i 的周期为 ∞ 。



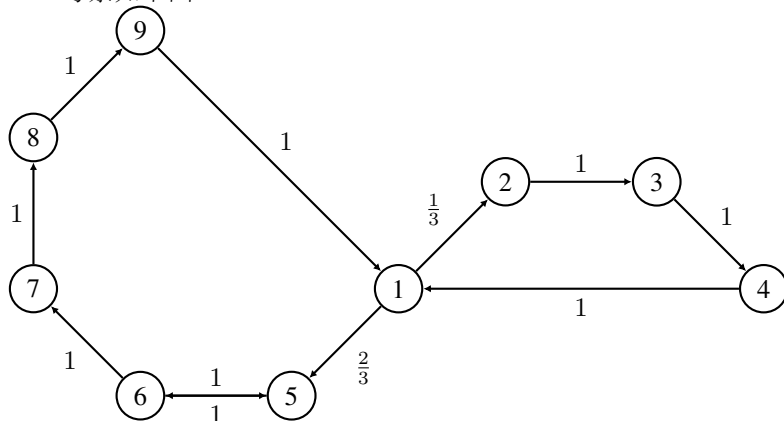
假设 i 状态为 d , 能不能推出对所有的 n , $p_{ii}^{(nd)} > 0$?

注

由定义?? 知道, 虽然有周期 d , 但并不是对所有的 n , $p_{ii}^{(nd)}$ 都大于 0。

但是可以证明, 当 n 充分大之后一定有 $p_{ii}^{(dn)} > 0$ 。

例题 3.13 考察如下图



由状态 1 出发再回到状态 1 的可能步长为 $T = \{4, 6, 8, 10, \dots\}$, 它的最大公约数是 2, 虽然从状态 1 出发 2 步并不能回到状态 1, 我们仍然称 2 是状态 1 的周期。

例题 3.14 如果质点每次向前、向后移动一步的概率是 $\frac{1}{3}$, 向后移动两步概率为 $\frac{1}{3}$, 则每个状态均为非周期的。

证明 对任意的 i , $p_{i,i+1} = p_{i,i-1} = p_{i,i-2} = \frac{1}{3}$ 。

$$p_{ii}^{(2)} \geq p_{i,i+1}p_{i+1,i} = \frac{1}{3} \cdot \frac{1}{3} > 0.$$

$$p_{ii}^{(3)} \geq p_{i,i-2}p_{i-2,i-1}p_{i-1,i} = \left(\frac{1}{3}\right)^3 > 0.$$

由于 2, 3 互素, 可知 i 非周期。

例题 3.15 在直线上, 如果质点每次向前移动一步的概率为 p , 向后移动 5 步概率是 $q = 1 - p$, $pq > 0$, 则每个状态周期均为 6。

证明 任取 n 满足 $p_{ii}^{(n)} > 0$, 则记向前移动次数为 x , 向后移动的步数记为 y 。

$$\begin{cases} x + y = n \\ x = 5y \end{cases}$$

故可知 $n = 6y$ 。同时 $p_{ii}^{(6)} \geq p_{i,i+1}p_{i+1,i+2}p_{i+2,i+3}p_{i+3,i+4}p_{i+4,i+5}p_{i+5,i} > 0$ 。故周期为 6。

定理 3.3.1

设 $i \leftrightarrow j$, 则 $d(i) = d(j)$ 。

证明 $d(i)|d(j)$ 且 $d(j)|d(i)$ 。证明 $d(i)$ 为 $\{k|p_{jj}^{(k)} > 0\}$ 的公约数。



证明 令 $p_{ii}^{(r)} > 0, p_{ij}^{(s)} > 0 (r, s \geq 1)$ 。任取 $n \in \{k|p_{jj}^{(k)} > 0\}$ 。则有

$$p_{ii}^{(r+n+s)} \geq p_{ij}^{(s)} p_{jj}^{(n)} p_{ji}^{(r)} > 0, \quad p_{ii}^{(r+s)} \geq p_{ij}^{(s)} p_{ji}^{(r)} > 0.$$

从而由周期的定义知 $d(i)|(r+n+s), d(i)|(r+s)$, 故有 $d(i)|n$ 。于是 $d(i)$ 为 $\{k|p_{jj}^{(k)} > 0\}$ 的公约数。因此, $d(i)|d(j)$ 。同理可得 $d(j)|d(i)$ 。

引理 3.3.1

设 $m \geq 2$, 正整数 s_1, s_2, \dots, s_m 的最大公因数为 d , 则存在正整数 N , 使得 $n > N$ 时, 必有非负整数 c_1, c_2, \dots, c_m 使得 $nd = \sum_{i=1}^m c_i s_i$ 。



定理 3.3.2

设 $i \in S, d(i) \geq 1$, 则存在 $N \geq 1$ 使得 $p_{ii}^{(nd(i))} > 0$ 对一切 $n \geq N$ 成立。



证明 将集合 $\{n \geq 1 : p_{ii}^{(n)} > 0\}$ 按其中元素递增的顺序重新排列成 $\{n_1, \dots, n_k, \dots\}$ 。令 \hat{d}_k 为 $\{n_1, n_2, \dots, n_k\}$ 的最大公约数, 则有

$$\hat{d}_1 \geq \hat{d}_2 \geq \dots \geq \hat{d}_k \geq \hat{d}_{k+1} \geq \dots \geq d(i)$$

因 \hat{d}_i 和 $d(i)$ 均有限, 故存在 n_0 使得: $\hat{d}_k = d(i)$ 对一切 $k > n_0$ 。则由初等数论知: $\exists N \geq 1$, 当 $n \geq N$ 时, 有非负整数 $s_1 \geq 0, \dots, s_{n_0} \geq 0$ 使得 $nd(i) = s_1 n_1 + \dots + s_{n_0} n_{n_0}$ 。故

$$p_{ii}^{(nd(i))} \geq (p_{ii}^{(n_1)})^{s_1} \dots (p_{ii}^{(n_{n_0})})^{s_{n_0}} > 0$$

3.4 常返性

定义 3.4.1 (首达时间)

设对 $j \in S$, 令

$$T_{ij} := \begin{cases} \min\{n : n \geq 1, X_n = j, X_0 = i\} & \text{若 } \{n \geq 1 | X_n = j, X_0 = i\} \neq \emptyset \\ +\infty & \text{若 } \{n \geq 1 | X_n = j, X_0 = i\} = \emptyset. \end{cases}$$

若 $X_0 = i$, 则称 T_{ij} 为从 i 出发首次到达 j 的时间, 而 T_{ii} 则表示从 i 出发首次回到 i 的时间。



定义 3.4.2 (首达概率)

对 $i, j \in S$, 令

$$f_{ij}^{(n)} = P(T_{ij} = n | X_0 = i), \quad n \geq 1.$$

称 $f_{ij}^{(n)}$ 为从 i 出发经过 n 步首次到达 j 的概率。 $f_{ii}^{(n)}$ 为从 i 出发经过 n 步首次回到 i 的概率。



由定义易知:

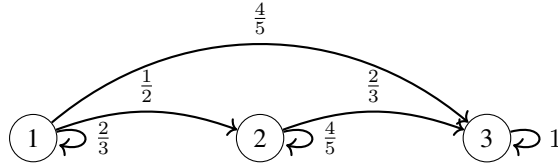
$$\begin{aligned}
 f_{ij}^{(n)} &= P(X_k \neq j, 1 \leq k \leq n-1, X_n = j | X_0 = i) \\
 &= \sum_{i_k \neq j, 1 \leq k \leq n-1} P(X_k = i_k, 1 \leq k \leq n-1, X_n = j | X_0 = i) \\
 &= \sum_{i_k \neq j, 1 \leq k \leq n-1} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} \\
 &= \sum_{i_{n-1} \neq j} \sum_{i_{n-2} \neq j} \cdots \sum_{i_1 \neq j} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}
 \end{aligned}$$

例题 3.16 记某设备的状态为 1, 2, 3, 其中 1 表示设备运行良好, 2 表示运行正常, 3 表示设备失效。以 X_n 表示设备在时刻 n 的状态, 且假设 $\{X_n, n = 0, 1, \dots\}$ 是齐次马氏链。在有维修及更换条件下, 其转移矩阵为

$$\mathbf{P} = \begin{pmatrix} \frac{2}{3} & \frac{3}{12} & \frac{1}{12} \\ 0 & \frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 1 \end{pmatrix}$$

求 $f_{13}^{(1)}$, $f_{13}^{(2)}$, $f_{13}^{(3)}$ 。

证明



1. $f_{13}^{(1)} = P(T_{13} = 1 | X_0 = 1) = p_{13} = \frac{1}{12}$;
2. $f_{13}^{(2)} = p_{11}p_{13} + p_{12}p_{23} = \frac{19}{180}$;
3. $P(T_{13} \geq 3 | X_0 = 1) = 1 - f_{13}^{(1)} - f_{13}^{(2)} = \frac{73}{90}$. 表示设备在 $[0, 3]$ 内运行的可靠性。

$$\begin{aligned}
 f_{13}^{(3)} &= P(T_{13} = 3 | X_0 = 1) \\
 &= P(X_3 = 3, X_2 \neq 3, X_1 \neq 3 | X_0 = 1) \\
 &= P(X_3 = 3, X_2 = 1, X_1 = 1 | X_0 = 1) \\
 &\quad + P(X_3 = 3, X_2 = 1, X_1 = 2 | X_0 = 1) \\
 &\quad + P(X_3 = 3, X_2 = 2, X_1 = 1 | X_0 = 1) \\
 &\quad + P(X_3 = 3, X_2 = 2, X_1 = 2 | X_0 = 1) \\
 &= \frac{1}{27} + 0 + \frac{1}{30} + \frac{1}{25} = \frac{149}{1350}
 \end{aligned}$$

$$\begin{aligned}
 f_{ij} &:= \sum_{n=1}^{\infty} f_{ij}^{(n)} \\
 &= \sum_{n=1}^{\infty} P(X_k \neq j, 1 \leq k \leq n-1, X_n = j | X_0 = i) \\
 &= \sum_{n=1}^{\infty} P(T_{ij} = n | X_0 = i) = P(T_{ij} < \infty | X_0 = i)
 \end{aligned}$$

$f_{ij} := \sum_{n=1}^{\infty} f_{ij}^{(n)}$ 表示从 i 出发经有限步首次到达 j 的概率。同理可解释 f_{ii}

定义 3.4.3 (常返状态)

若 $f_{ii} = 1$, 则称 i 为常返。否则, 称 i 为非常返 (或暂留, 瞬时的)。



注

吸收状态 i 满足 $f_{ii} = f_{ii}^{(1)} = 1$

下面分析各状态的性质，以及如何根据转移矩阵 P 来判断其常返性，不可约性等。

定理 3.4.1

对任意 $i, j \in S, n \geq 1$ ，有：

(a) $p_{ij}^{(n)} = \sum_{m=1}^n f_{ij}^{(m)} p_{ij}^{(n-m)}$ (注意, $p_{ij}^{(0)} \equiv 1$);

(b) $f_{ij}^{(n)} = \sum_{k \neq j} p_{ik} f_{kj}^{(n-1)} I_{\{n>1\}} + p_{ij} I_{\{n=1\}}$;

(c) $i \rightarrow j \iff f_{ij} > 0; i \leftrightarrow j \iff f_{ij} f_{ji} > 0$ 。



注 由 (a) 可知 $p_{ij}^{(n)} \geq f_{ij}^{(n)}$

证明

(a) 证明用的是首次进入方法。即依照首次进入状态 j 的时刻进行分解 (即下面的第二个等式)：

$$\begin{aligned} p_{ij}^{(n)} &= P(X_n = j | X_0 = i) = P(X_n = j, T_{ij} \leq n | X_0 = i) \\ &= \sum_{m=1}^n P(T_{ij} = m, X_n = j | X_0 = i) \\ &= \sum_{m=1}^n P(T_{ij} = m | X_0 = i) P(X_n = j | X_0 = i, T_{ij} = m) \\ &= \sum_{m=1}^n f_{ij}^{(m)} P(X_n = j | X_0 = i, X_1 \neq j, \dots, X_{m-1} \neq j, X_m = j) \\ &= \sum_{m=1}^n f_{ij}^{(m)} p_{ij}^{(n-m)}. \end{aligned}$$

(b) 当 $n = 1$ 时，由 $f_{ij}^{(1)} = p_{ij}$ ，结论显然成立。对 $n \geq 2$ ，因为

$$\begin{aligned} \{T_{ij} = n\} &= \{X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j\} \\ &= \bigcup_{k \neq j} \{X_1 = k, X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j\}. \end{aligned}$$

$$\begin{aligned} f_{ij}^{(n)} &= P(T_{ij} = n | X_0 = i) \\ &= \sum_{k \neq j} P(X_1 = k, X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j | X_0 = i) \\ &= \sum_{k \neq j} p_{ik} \cdot P(X_2 \neq j, \dots, X_{n-1} \neq j, X_n = j | X_0 = i, X_1 = k) \\ &= \sum_{k \neq j} p_{ik} f_{kj}^{(n-1)}. \end{aligned}$$

(马氏性与齐次性)

(c) 由 (a) 知 $p_{ij}^{(n)} \geq f_{ij}^{(n)} (\forall i \in S, n \geq 1)$ 。当 $f_{ij} > 0$ 时，有 $i \rightarrow j$ 。另一方面，若 $i \rightarrow j$ ，则有 $n \geq 1$ 使得 $p_{ij}^{(n)} > 0$ 。当 $n = 1$ 时，知 $f_{ij}^{(1)} = p_{ij}^{(1)} > 0$ ，从而 $f_{ij} > 0$ 。当 $n \geq 2$ 时，由 (a) 知，

$$p_{ij}^{(n)} = \sum_{m=1}^n f_{ij}^{(m)} p_{ij}^{(n-m)} > 0.$$

故有 $1 \leq m \leq n$ 使得 $f_{ij}^{(m)} > 0$ ，从而 $f_{ij} > 0$ 。

下面有一个有意思的结论，务必要记下来

命题 3.4.1

$$\begin{aligned}
f_{ij}^{(k)} &= P(X_k \neq j, 1 \leq k \leq n-1, X_n = j | X_0 = i) \\
&= \sum_{i_k \neq j, 1 \leq k \leq n-1} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} \\
&= \sum_{i_{n-1} \neq j} \sum_{i_{n-2} \neq j} \cdots \sum_{i_1 \neq j} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j} \\
&= \sum_{i_{n-1} \neq j} \left(\sum_{i_{n-2} \neq j} \cdots \left(\sum_{i_2 \neq j} \left(\sum_{i_1 \neq j} p_{ii_1} p_{i_1 i_2} \right) p_{i_2 i_3} \right) \cdots p_{i_{n-2} i_{n-1}} \right) p_{i_{n-1} j} \\
&= \sum_{i_1 \neq j} p_{ii_1} \left(\sum_{i_2 \neq j} \cdots \sum_{i_{n-3} \neq j} \left(\sum_{i_{n-2} \neq j} \left(\sum_{i_{n-1} \neq j} p_{i_{n-2} i_{n-1}} p_{i_{n-1} j} \right) \right) \right) \right).
\end{aligned}$$

用矩阵表示：对任意的 $k \geq 2$,

$$f_{ij}^{(k)} = \begin{pmatrix} p_{i1} & p_{i2} & \cdots & p_{i,j-1} & p_{i,j+1} & \cdots & p_{in} \end{pmatrix} \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1,j-1} & p_{1,j+1} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2,j-1} & p_{2,j+1} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{j-1,1} & p_{j-1,2} & \cdots & p_{j-1,j-1} & p_{j-1,j+1} & \cdots & p_{j-1,n} \\ p_{j+1,1} & p_{j+1,2} & \cdots & p_{j+1,j-1} & p_{j+1,j+1} & \cdots & p_{j+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{n,j-1} & p_{n,j+1} & \cdots & p_{nn} \end{pmatrix}^{k-2} \begin{pmatrix} p_{1j} \\ p_{2j} \\ \vdots \\ p_{j-1,j} \\ p_{j+1,j} \\ \vdots \\ p_{nj} \end{pmatrix}$$

注 左边行向量是转移矩阵的第 i 行去掉第 j 个元素，右边列向量是转移矩阵的第 j 列去掉第 i 个元素，中间的矩阵是转移矩阵去掉第 j 行和第 j 列。

例题 3.17

$$\mathbf{P} = \begin{pmatrix} \frac{2}{3} & \frac{3}{12} & \frac{1}{12} \\ 0 & \frac{4}{5} & \frac{1}{5} \\ 0 & 0 & 1 \end{pmatrix}$$

求 $f_{13}^{(3)}$ 。

证明

$$\begin{aligned}
f_{13}^{(3)} &= \begin{pmatrix} \frac{2}{3} & \frac{3}{12} \end{pmatrix} \begin{pmatrix} \frac{2}{3} & \frac{3}{12} \\ 0 & \frac{4}{5} \end{pmatrix} \begin{pmatrix} \frac{1}{12} \\ \frac{1}{5} \end{pmatrix} \\
&= \frac{149}{1350}
\end{aligned}$$

定理 3.4.2

- (1) $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}}$
- (2) i 常返 $\iff \sum_{n=0}^{\infty} p_{ii}^{(n)} = +\infty$
- (3) i 非常返 $\iff \sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1-f_{ii}} < \infty$

证明 (1,2,3) 引进母函数：

$$P_{ij}(s) \triangleq \sum_{n=0}^{\infty} s^n p_{ij}^{(n)}, \quad F_{ij}(s) \triangleq \sum_{n=1}^{\infty} s^n f_{ij}^{(n)}, \quad s \in (0, 1).$$

于是有:

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} s^n = \sum_{n=1}^{\infty} \left[\sum_{m=1}^n f_{ij}^{(m)} p_{ij}^{(n-m)} \right] s^n = \sum_{m=1}^{\infty} \left[f_{ij}^{(m)} s^m \sum_{n=m}^{\infty} p_{ij}^{(n-m)} s^{n-m} \right],$$

$$P_{ij}(s) - \delta_{ij} = P_{ij}(s) F_{ij}(s).$$

这样有: $F_{ii}(s) = 1 - \frac{1}{P_{ii}(s)}$ 。任取 $N > 1$, 则有

$$\sum_{n=0}^N p_{ii}^{(n)} s^n < P_{ii}(s) < \sum_{n=0}^{\infty} p_{ii}^{(n)},$$

令 $s \uparrow 1$, 在令 $N \rightarrow \infty$ 由单调收敛定理可知 $P_{ii}(s) \rightarrow \sum_{n=0}^{\infty} p_{ii}^{(n)}$, 令 $f_{ii} = \lim_{s \uparrow 1} F_{ii}(s)$, 及 $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \lim_{s \uparrow 1} P_{ii}(s)$ 可得: 若 $f_{ii} = 1$, 则有 $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$ 。故 (1)(2)(3) 得证。

推论 3.4.1

若 j 为非常返状态, 则对任意的 $i \in S$,

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty, \quad \lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0.$$

证明 对给定的 N ,

$$\begin{aligned} \sum_{n=1}^N p_{ij}^{(n)} &= \sum_{n=1}^N \sum_{l=1}^n f_{ij}^{(l)} p_{ij}^{(n-l)} \\ &= \sum_{l=1}^N \sum_{n=l}^N f_{ij}^{(l)} p_{ij}^{(n-l)} \\ &= \sum_{l=1}^N f_{ij}^{(l)} \sum_{m=0}^{N-l} p_{ij}^{(m)} \\ &\leq \sum_{l=1}^N f_{ij}^{(l)} \sum_{n=0}^N p_{ij}^{(n)}. \end{aligned}$$

令 $N \rightarrow \infty$,

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} \leq \sum_{l=1}^{\infty} f_{ij}^{(l)} \left(1 + \sum_{n=1}^{\infty} p_{ij}^{(n)} \right) \leq 1 + \sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty.$$

由于 $p_{ij}^{(n)} \geq 0$, 故 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 。

推论 3.4.2

若 j 为常返状态, 则当 $i \rightarrow j$ 时,

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty.$$

证明 由于 $i \rightarrow j$, 故存在 $m > 0$ 使得 $p_{ij}^{(m)} > 0$, 故

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)} \geq p_{ij}^{(m)} p_{jj}^{(n)}$$

故 $\sum_{n=1}^{\infty} p_{ij}^{(m+n)} \geq p_{ij}^{(m)} \sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$ 。

$$I_n(i) := \begin{cases} 1 & X_n = i \\ 0 & X_n \neq i \end{cases}$$

$$S(i) := \sum_{n=0}^{\infty} I_n(i)$$

则可证明

$$\begin{aligned} E(S(i)|X_0 = i) &= E\left[\sum_{n=0}^{\infty} I_n(i)|X_0 = i\right] \\ &= \sum_{n=0}^{\infty} E[I_n(i)|X_0 = i] \\ &= \sum_{n=0}^{\infty} p_{ii}^{(n)}. \end{aligned}$$

即由 i 出发回到 i 的平均次数。

$$g_{ij} := P(\text{有无穷多个 } n \geq 1 \text{ 使得 } X_n = j | X_0 = i)$$

定理 3.4.3

对一切的 i, j

- (i) $g_{ii} = \lim_{n \rightarrow \infty} (f_{ii})^n$, $g_{ij} = f_{ij}g_{jj}$.
- (ii) 设 i 为常返态, $i \rightarrow j$, 则 $g_{ij} = f_{ij} = 1$.



证明 记 $g_{ij}(m) = P(\text{至少有 } m \text{ 个 } n \geq 1 \text{ 使得 } X_n = j | X_0 = i)$ 。显然 $g_{ij}(1) = f_{ij}$, 且

$$\{\{\text{有无穷多个 } n \geq 1 \text{ 使得 } X_n = j\}\} = \bigcap_{m=1}^{\infty} \{\text{至少有 } m \text{ 个 } n \geq 1 \text{ 使得 } X_n = j\}.$$

故 $g_{ij}(m) \downarrow g_{ij}$

$$\begin{aligned} g_{ij}(m+1) &= P(\text{至少有 } m+1 \text{ 个 } n \geq 1 \text{ 使得 } X_n = j | X_0 = i) \\ &= \sum_{k=1}^{\infty} P(T_{ij} = k, \text{至少有 } m+1 \text{ 个 } n \geq 1 \text{ 使得 } X_n = j | X_0 = i) \\ &= \sum_{k=1}^{\infty} P(X_k = j, X_v \neq j, 0 < v < k, \text{至少有 } m \text{ 个 } n \geq k+1 \text{ 使得 } X_n = j | X_0 = i) \\ &= \sum_{k=1}^{\infty} P(X_k = j, X_v \neq j, 0 < v < k | X_0 = i) \cdot P(\text{至少有 } m \text{ 个 } n \geq k+1 \text{ 使得 } X_n = j | X_k = j, X_v \neq j, 0 < v < k, X_k = j) \\ &= \sum_{k=1}^{\infty} f_{ij}^{(k)} P(\text{至少有 } m \text{ 个 } n \geq 1 \text{ 使得 } X_n = j | X_0 = i) \\ &= \sum_{k=1}^{\infty} f_{ij}^{(k)} g_{ij}(m) = f_{ij} g_{ij}(m) \end{aligned}$$

令 $m \rightarrow \infty$, 则有 $g_{ij} = f_{ij}g_{ij}$

由 $g_{ij}(m+1) = f_{ij}g_{ij}(m)$ 可知有

$$g_{ij}(m+1) = f_{ij}g_{ij}(m) = \cdots (f_{ij})^{m+1}.$$

故 (i) $g_{ii} = \lim_{n \rightarrow \infty} g_{ii}(n) = \lim_{n \rightarrow \infty} (f_{ii})^n$, $g_{ij} = f_{ij}g_{ij}$.

定理 3.4.4

设 i 为常返态, $i \rightarrow j$, 则 $g_{ij} = f_{ij}g_{ii} = 1$.



证明 对任意的 $m \geq 1$ 以及 $l \in S$

$$\begin{aligned}
 g_{il} &= P(\text{有无穷多个 } n \geq 1 \text{ 使得 } X_n = l | X_0 = i) \\
 &= P(X_m \in S, \text{有无穷多个 } n \geq 1 \text{ 使得 } X_n = l | X_0 = i) \\
 &= \sum_k P(X_m = k; \text{有无穷多个 } n \geq 1 \text{ 使得 } X_n = l | X_0 = i) \\
 &= \sum_k P(X_m = k; \text{有无穷多个 } n \geq m+1 \text{ 使得 } X_n = l | X_0 = i) \\
 &= \sum_k P(X_m = k) P(\text{有无穷多个 } n \geq m+1 \text{ 使得 } X_n = l | X_m = k, X_0 = i) \\
 &= \sum_k p_{ik}^{(m)} P(\text{有无穷多个 } n \geq m+1 \text{ 使得 } X_n = l | X_m = k, X_0 = i) \\
 &= \sum_k p_{ik}^{(m)} g_{kl}
 \end{aligned}$$

由于 i 常返,

$$0 = 1 - g_{ii} = \sum_k p_{ik}^{(m)} (1 - g_{kl}).$$

从而对一切 $m \geq 1$ 以及 $k \in S$, $p_{ik}^{(m)} (1 - g_{kl}) = 0$. 若 $i \rightarrow j$, 则存在 $m \geq 1$ 使得 $p_{ij}^{(m)} > 0$, 此时有 $g_{ij} = 1$, 但 $f_{ij} \geq g_{ij}$.

定理 3.4.5

设 i 常返, 若 $i \rightarrow j$, 则 $g_{ij} = f_{ij} = f_{ii} = 1$, $j \leftrightarrow i$, 并且 j 也是常返的.

证明 设 n, m 使得 $p_{ij}^{(m)} p_{ji}^{(n)} > 0$. 对任意的 $s \geq 1$,

$$p_{ij}^{(m+s+n)} \geq p_{ij}^{(n)} p_{ii}^{(s)} p_{ij}^{(m)}.$$

两边对 s 求和得

$$\sum_{s=1}^{\infty} p_{ij}^{(m+s+n)} \geq p_{ij}^{(n)} p_{ij}^{(m)} \sum_{s=1}^{\infty} p_{ii}^{(s)} = \infty.$$

故由 (2) 可知状态 j 常返。

例题 3.18 考虑直线上无限制的随机游动, 状态空间为 $S = \{0, \pm 1, \pm 2, \dots\}$, 转移概率为 $p_{i,i+1} = 1-p, p_{i,i-1} = p, i \in S (0 < p < 1)$. 对于状态 0, 可知 $p_{00}^{(2n+1)} = 0, n = 1, 2, \dots$, 即从 0 出发奇数次不可能返回到 0. 而

$$p_{00}^{(2n)} = \binom{2n}{n} p^n (1-p)^n = \frac{(2n)!}{n!n!} [p(1-p)]^n$$

即经过偶数次回到 0 当且仅当它向左、右移动距离相同。

由 Stirling 公式知, 当 n 充分大时, $n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}$, 则 $p_{00}^{(2n)} \sim \frac{[4p(1-p)]^n}{\sqrt{\pi n}}$. 而 $p(1-p) \leq \frac{1}{4}$ 且 $p(1-p) = \frac{1}{4} \iff p = \frac{1}{2}$. 于是 $p = \frac{1}{2}$ 时, $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$, 否则 $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$, 即当 $p \neq \frac{1}{2}$ 时状态 0 是非常返状态, $p = \frac{1}{2}$ 时是常返状态. 显然, 过程的各个状态都是相通的, 故以此可得其他状态的常返性. (请读者自己考虑它们的周期是什么?)

定义 3.4.4 (正常返状态)

设 i 常返, 若

$$\mu_i := \sum_{n=1}^{\infty} n f_{ii}^{(n)} = \sum_{n=1}^{\infty} n P(T_{ii} = n | X_0 = i) < \infty,$$

则称 i 为正常返的; 否则称之为零常返的

明显: $E[T_{ii} | X_0 = i] = \mu_i$ 为从 i 出发回到 i 的平均时间。

定义 3.4.5 (遍历性)

若状态 i 是正常返且非周期的, 则称它是遍历的。



例题 3.19 记 $S = \{1, 2, 3, 4\}$.

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$

判断状态 1, 2, 3, 4 是常返性还是非常性。若是常返, 进一步判断是正常返还是零常返?

证明 $f_{444}^{(n)} = 0, f_{444} = 0 \quad \forall n \geq 1$;

$$f_{33}^{(1)} = \frac{2}{3}, f_{33}^{(n)} = 0, f_{33} = f_{33}^{(1)} = \frac{2}{3} \quad \forall n \geq 2;$$

因此: 3 和 4 非常返。

$$f_{11}^{(1)} = \frac{1}{2}, f_{11}^{(2)} = p_{12}p_{21} = \frac{1}{2}, f_{11}^{(n)} = 0, \forall n \geq 3; f_{11} = \sum_{n=1}^{\infty} f_{11}^{(n)} = 1.$$

$$f_{22}^{(1)} = 0, f_{22}^{(2)} = p_{21}p_{12} = \frac{1}{2}, f_{22}^{(3)} = p_{21}p_{11}p_{12} = \frac{1}{2} \cdot \frac{1}{2}, \dots, f_{22}^{(n+1)} = \frac{1}{2^n}, \forall n \geq 1; f_{22} = \sum_{n=1}^{\infty} f_{22}^{(n)} = 1.$$

$$\text{因此: } \mu_1 = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}; \mu_2 = \sum_{n=1}^{\infty} n f_{22}^{(n)} = \sum_{n=2}^{\infty} n \cdot \frac{1}{2^{n-1}} = 3.$$

于是: 1 和 2 均正常返, 3 和 4 非常返。

定理 3.4.6

设 P 不可约, 周期 $d > 1$, 则有

(a) 对任给定的 $i, j \in S$, 若 $p_{ij}^{(m)} > 0, p_{ij}^{(n)} > 0$, 则有 $d | (n - m)$ 。同时存在唯一的 r , 使得只要 $p_{ij}^{(n)} > 0$ 就有 $n = kd + r$ 。

(b) 状态空间 S 可分成 d 个不相交的集合的并:

$$S = G_1 \cup G_2 \cup \dots \cup G_d,$$

其中, 从任一 G_m 中的状态出发, 下一步到达 $G_{m+1 \pmod d}$ 中某状态的概率大于 0;

(c) 链 P^d 是非周期的, 且 $\sum_{k \in G_m} p_{ik}^{(d)} = 1 \quad \forall i \in G_m$ 。且将 P^d 限制在 G_m 上时, 构成一个不可约非周期的子链。



证明 (a) 由 $j \rightarrow i$, 存在 $k \geq 1$ 使得 $p_{ij}^{(k)} > 0$, 于是有

$$p_{ii}^{(n+k)} \geq p_{ij}^{(n)} p_{ji}^{(k)} > 0, \quad p_{ii}^{(m+k)} \geq p_{ij}^{(m)} p_{ji}^{(k)} > 0$$

故 $d | (n + k)$ 且 $d | (m + k)$, 于是有 $d | (n - m)$ 。因此存在唯一的 r , 使得 $n = k_1 d + r, m = k_2 d + r$ 。

(b) 先固定 $i \in S$, 令

$$G_1 \triangleq \{k : \exists n, p_{ik}^{(nd+1)} > 0\},$$

$$G_2 \triangleq \{k_2 : \exists k_1 \in G_1 \text{ 使得 } p_{k_1 k_2} > 0\},$$

\vdots

$$G_m \triangleq \{k_m : \exists k_{m-1} \in G_{m-1} \text{ 使得 } p_{k_{m-1} k_m} > 0\} \quad (m = 2, \dots, d).$$

对任意的 $k_1 \in G_1$, 由 $\sum_{j \in S} p_{k_1 j} = 1$ 知存在 $k_2 \in S$ 使得 $p_{k_1 k_2} > 0$, 故 $G_2 \neq \emptyset$ 。依次类推, $G_m \neq \emptyset (m = 2, \dots, d)$ 。

由 (a) 知 G_1, \dots, G_d 互不相交。若存在 $j \in G_{d_1} \cap G_{d_2}$, 则有

$$p_{ij}^{(nd+d_1)} > 0, \quad p_{ij}^{(md+d_2)} > 0$$

因此 $d | (d_2 - d_1)$, 故 $G_{d_1} = G_{d_2}$ 。

由 P 不可约知 $G_1 \cup \dots \cup G_d = S$ (即 $i \rightarrow i_1 \rightarrow \dots \rightarrow i_n \rightarrow j$)。事实上, 任取 $j \in S$, 存在 m 使得 $p_{ij}^{(m)} > 0$ 。设 $m = kd + r$, 则 $j \in G_r$ 。

(c) 由定理 ?? 知 P^d 是非周期的。令 $\tilde{P} := P^d$ 为一步转移矩阵。对任意 $j \in S$, 存在 N 使得当 $n \geq N$ 时 $p_{ij}^{(nd)} > 0$ 。取

$$\tilde{p}_{ij}^{(N)} := p_{ij}^{(Nd)} > 0, \quad \tilde{p}_{ij}^{(N+1)} := p_{ij}^{((N+1)d)} > 0$$

其中 $\tilde{p}_{ik} := p_{ik}^{(d)}$ 。由于 $(N, N+1) = 1$, 故 \tilde{P} 非周期。

由 (b) 知 $\sum_{k \in G_m} p_{jk}^{(d)} = 1$ ($\forall j \in G_m, G_m \in \{G_1, \dots, G_d\}$)。因此

$$P^d|_{G_m} := \left(p_{ij}^{(d)} \right)_{i,j \in G_m}$$

为随机矩阵。对任意 $j, k \in G_m$, 由不可约性存在 $l \in \mathbb{N}$ 使得 $p_{jk}^{(ld)} > 0$, 即在 $P^d|_{G_m}$ 中 l 步可达, 故 $P^d|_{G_m}$ 不可约。

注 (c) 将周期矩阵 P^d 限制在 G_m 上时, 构成一个不可约非周期的子链。

例题 3.20 记 $S = \{1, 2, 3, 4, 5, 6\}$.

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} & 0 \end{pmatrix}$$

证明 取初始状态 1, 则有

$$G_1 := \{3, 5\}.$$

$$G_2 := \{2\}.$$

$$G_3 := \{1, 4, 6\}.$$

例题 3.21

$$\mathbf{P}^{(3)} = \begin{pmatrix} \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{7}{12} & 0 & \frac{5}{12} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{7}{12} & 0 & \frac{5}{12} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{pmatrix}$$

证明

$$P^{(3)}|_{G_1} = \begin{pmatrix} \frac{7}{12} & \frac{5}{12} \\ \frac{7}{12} & \frac{5}{12} \end{pmatrix}$$

$$P^{(3)}|_{G_2} = (1)$$

$$P^{(3)}|_{G_3} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

3.5 不变分布

若知极限 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j$ 存在且不依赖于 i , 则由

$$P^{n+1} = P^n P \quad (p_{ij}^{(n+1)} = \sum_{l \in S} p_{il}^{(n)} p_{lj})$$

及 Fatou 引理知:

$$\pi_j \geq \sum_i \pi_i p_{ij} \quad \forall j \in S \text{ 且 } \sum_j \pi_j \leq 1.$$

从而设存在一状态 $k \in S$ 使得 $\pi_k > \sum_i \pi_i p_{ik}$ 。

$$\sum_j \pi_j > \sum_j \sum_i \pi_i p_{ij} = \sum_i \pi_i \sum_j p_{ij} = \sum_i \pi_i,$$

矛盾。

故

$$\pi_j = \sum_i \pi_i p_{ij} \quad \forall j \in S.$$

记 $\pi := (\pi_i, i \in S)$, 则有 $\pi = \pi P$.

定义 3.5.1

称 $v = \{v_i, i \in S\}$ 为 P 的不变测度, 如果 $0 \leq v_i < \infty (i \in S)$ 且 $v = vP, v \neq 0$.

$$v_j = \sum_i v_i p_{ij} \quad \forall j \in S.$$



定义 3.5.2 (不变分布)

若 $\{\pi_i, i \in S\}$ 为 S 上的概率分布, 且它满足对任意的 $j \in S$ 有 $\pi_j = \sum_i \pi_i p_{ij}$, 则称 $\{\pi_i, i \in S\}$ 为 P (或该马氏链) 的不变分布或平稳分布。



回顾 $p_{ij}^{(n)} = P(X_n = j | X_0 = i)$ 与 $P(X_n = j)$

$$\pi_i(n) = P(X_n = i), i \in S,$$

$$\pi(n) = (\pi_i(n), i \in S).$$

即 $\pi(n)$ 表示 n 时刻 X_n 的概率分布, 称 $\pi(0) := (\pi_i(0), i \in S)$ 为马氏链 $\{X_n, n = 0, 1, \dots\}$ 的初始分布。

$$\pi(n+1) = \pi(n)P,$$

$$\pi(n) = \pi(0)P^n,$$

设 $\pi := (\pi_i, i \in S)$ 为不变分布, 即 $\pi = \pi P$ 。

若令马氏链的初始分布 $\pi(0) := \pi$ 。会怎样? 则有递归关系:

$$\pi = \pi P = \pi P \cdot P = \pi \cdot P^2 = \dots = \pi P^n.$$

即 $\pi(0) = \pi(0)P^n = \pi(n)$ 。即马氏链任意时刻的分布都为初始分布。

现在考虑

$$e_{ji}^{(n)} = P(X_n = i, X_m \neq j, 0 < m < n | X_0 = j) \quad (5.2)$$

$$e_{ji} := \sum_{n=1}^{\infty} e_{ji}^{(n)}.$$

定理 3.5.1

若 j 常返, 则有: $V := (v_i := e_{ji}, i \in \mathcal{S})$ 为不变测度, 且

$$e_{jj} = \sum_{n=1}^{\infty} f_{jj}^{(n)} = 1.$$



注

$$e_{jj}^{(n)} = f_{jj}^{(n)}.$$

证明 由定义知

$$e_{ji}^{(n)} = P(X_n = i, X_m \neq j, 0 < m < n | X_0 = j) \quad e_{ji}^{(1)} = p_{ji}. \quad (5.3)$$

$$\begin{aligned} e_{ji}^{(n+1)} &= P(X_{n+1} = i, X_v \neq j, 0 < v < n+1 | X_0 = j) \\ &= \sum_{k \neq j} P(X_{n+1} = i, X_n = k, X_v \neq j, 0 < v < n | X_0 = j) \\ &= \sum_{k \neq j} P(X_n = k, X_v \neq j, 0 < v < n | X_0 = j) p_{ki} \\ &= \sum_{k \neq j} e_{jk}^{(n)} p_{ki}. \end{aligned}$$

$$\begin{aligned} e_{ji} &= \sum_{n=1}^{\infty} e_{ji}^{(n)} \\ &= \sum_{n=2}^{\infty} \sum_{k \neq j} e_{jk}^{(n-1)} p_{ki} + p_{ji} \\ &= \sum_{k \neq j} \sum_{n=1}^{\infty} e_{jk}^{(n)} p_{ki} + p_{ji} \\ &= \sum_{k \neq j} e_{jk} p_{ki} + p_{ji} \end{aligned}$$

由于 $e_{jj}^{(n)} = f_{jj}^{(n)}$, 故 $e_{jj} = \sum_{n=1}^{\infty} f_{jj}^{(n)} = f_{jj} = 1 > 0$ 。

令 $v := (v_i := e_{ji}, i \in \mathcal{S})$ 。下证 v 为不变测度。由上式可知 $e_{ji} := \sum_k e_{jk} p_{ki}$ 即 $v_i := \sum_k v_k p_{ki}$ 。同时迭代可知

$$e_{ji} = \sum_{k \in \mathcal{S}} e_{jk} p_{ki} = \sum_{k \in \mathcal{S}} \left(\sum_{s \in \mathcal{S}} e_{js} p_{sk} \right) p_{ki} = \sum_{s \in \mathcal{S}} e_{js} \sum_{k \in \mathcal{S}} p_{sk} p_{ki} = \sum_{s \in \mathcal{S}} e_{js} p_{si}^{(2)}.$$

且由 $e_{ii} = 1 > 0$ 可知 $\{e_{ii}, i \in \mathcal{S}\}$ 非负且不全为 0。

$$\begin{aligned} e_{ji} &= \sum_{k \in \mathcal{S}} e_{jk} p_{ki} = \sum_{k \in \mathcal{S}} \left(\sum_{s \in \mathcal{S}} e_{js} p_{sk} \right) p_{ki} = \sum_{s \in \mathcal{S}} e_{js} \sum_{k \in \mathcal{S}} p_{sk} p_{ki} = \sum_{s \in \mathcal{S}} e_{js} p_{si}^{(2)} \\ &= \sum_{s \in \mathcal{S}} e_{js} p_{si}^{(n)} \end{aligned}$$

且由 $e_{ii} > 0$ 可知 $\{e_{ii}, i \in \mathcal{S}\}$ 非负且不全为 0。下证 $v_i = e_{ji} < \infty$ 。

若 j 不可达状态 i , 则有 $v_i := e_{ji} = 0$ 。

若 $j \rightarrow i$, 则由常返可知, $i \leftrightarrow j$ 。故存在 $n \geq 1$ 使得 $p_{ij}^{(n)} > 0$ 。

故 $1 = v_i = e_{ji} = \sum_{k \in \mathcal{S}} e_{jk} p_{ki}^{(n)} \geq e_{ji} p_{ij}^{(n)}$ 。

因此, $e_{ji} \leq \frac{1}{p_{ij}^{(n)}} < \infty$ 。

引理 3.5.1

设马氏链有不变测度 $V = (v_i, i \in S)$ 。若 $v_i > 0, i \rightarrow j$, 则 $v_j > 0$ 。特别地, 若链不可约, 则 $v_j > 0, j \in S$ 。



证明 由于 $v_i > 0, i \rightarrow j$, 则存在 $n \geq 1$ 使得 $p_{ij}^{(n)} > 0$ 。

由 $v_j = \sum_k v_k p_{kj}^{(n)} \geq v_i p_{ij}^{(n)} > 0$ 。

由于不变测度不恒为 0, 至少存在 $v_i > 0$ 。由于该链不可约, 则对任意的 $j \in S$ 有 $i \rightarrow j$ 。故 $v_j > 0$ 。

引理 3.5.2

设链常返不可约, 则不计一个常数因子, 不变测度唯一。



证明

证明可见何声武《随机过程引论》定理 5.3。

即证明任取 $\mu := (\mu_i, i \in S)$ 为另一不变测度。则有对任意的 $i \in S$ 有 $\mu_i = \mu_j e_{ji}$ 。

即 $(\mu_i, i \in S)$ 与 $(e_{ji}, i \in S)$ 只差一个常数。

引理 3.5.3

设 $V = (v_i, i \in S)$ 为马氏链 $\{X_n\}$ 的平稳分布, $v_j > 0$ 则 j 为常返态。



证明 反证法: 设 j 为非常返态。由推论 4.6 可知对任意的 $i \in S$ 有 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 。

由于 $v_j = \sum_{i \in S} v_i p_{ij}^{(n)}$, 令 $n \rightarrow \infty$, 由控制收敛定理可知 $v_j = 0$ 。矛盾。

引理 3.5.4

$$\sum_{i \in S} e_{ji} = \begin{cases} \mu_j := E[T_{ij} | X_0 = j] & \text{若 } j \text{ 为常返态} \\ \infty & \text{若 } j \text{ 为非常返态} \end{cases}$$



证明 对任意的 $n \geq 1$ 以及 $i, j \in S$ 。

$$\begin{aligned} \sum_{i \in S} e_{ji}^{(n)} &= \sum_{i \in S} P(X_n = i, X_v \neq j, 0 < v < n | X_0 = j) \\ &= P(X_v \neq j, 0 < v < n | X_0 = j) \\ &= P(T_{ij} \geq n | X_0 = j) \\ &= P(T_{ij} = \infty | X_0 = j) + \sum_{v=n}^{\infty} P(T_{ij} = v | X_0 = j) \\ &= (1 - f_{jj}) + \sum_{v=n}^{\infty} f_{jj}^{(v)}. \end{aligned}$$

引理 3.5.5

$$\sum_{i \in S} e_{ji} = \begin{cases} \mu_j := E[T_{ij} | X_0 = j] & \text{若 } j \text{ 为常返态} \\ \infty & \text{若 } j \text{ 为非常返态} \end{cases}$$



证明

$$\sum_{i \in S} e_{ji} = \sum_{i \in S} \sum_{n=1}^{\infty} e_{ji}^{(n)} = \sum_{n=1}^{\infty} \left[(1 - f_{jj}) + \sum_{v=n}^{\infty} f_{jj}^{(v)} \right].$$

故若 j 为非常返态, $f_{jj} < 1$, 则 $\sum_{i \in S} e_{ji} = \infty$ 。

若 j 为常返态, $f_{jj} = 1$,

则 $\sum_{i \in \mathcal{S}} e_{ji} = \sum_{n=1}^{\infty} \sum_{v=n}^{\infty} f_{jj}^{(v)} = \sum_{v=1}^{\infty} v f_{jj}^{(v)} = \mu_j$.

从上述结论可知若 j 是正常返的, 则若令 $\nu_i := \frac{e_{ji}}{\mu_j}$ 。则有

$$\sum_{i \in \mathcal{S}} \nu_i = \sum_{i \in \mathcal{S}} \frac{e_{ji}}{\mu_j} = 1,$$

且由 $e_{ji} = \sum_{k \in \mathcal{S}} e_{jk} p_{ki}$, 可知 $\nu_i = \sum_{k \in \mathcal{S}} \nu_k p_{ki}$ 。即 $\nu := (\nu_i, i \in \mathcal{S})$ 为平稳分布。

设 $P = (P_1, P_2, \dots, P_m)$ 是马氏链的一步转移概率矩阵, P_j 是 P 的第 j 列, 设 $\pi := (\pi_1, \pi_2, \dots, \pi_m)$ 为平稳分布。则方程组

$$\pi = \pi P, \quad \sum_{j=1}^m \pi_j = 1 \quad (5.1)$$

和

$$(\pi_1, \pi_2, \dots, \pi_{m-1}) = \pi(P_1, P_2, \dots, P_{m-1}), \quad \sum_{j=1}^m \pi_j = 1 \quad (5.2)$$

等价。实际上,

$$\pi = \pi P \text{ 即为 } (\pi_1, \pi_2, \dots, \pi_{m-1}, \pi_m) = \pi(P_1, P_2, \dots, P_{m-1}, P_m).$$

例题 3.22 记 $\mathcal{S} = \{1, 2, 3\}$.

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{3}{4} \\ 1 & 0 & 0 \end{pmatrix}$$

求平稳分布。

证明

$$\begin{cases} (\pi_1, \pi_2, \pi_3) \mathbf{P} = (\pi_1, \pi_2, \pi_3) \\ \pi_1 + \pi_2 + \pi_3 = 1 \\ \pi_i \geq 0. \quad (i = 1, 2, 3) \end{cases}$$

$$\begin{cases} \pi_3 = \pi_1, \\ \frac{1}{4}\pi_2 + \frac{1}{2}\pi_1 = \pi_2 \\ \pi_1 + \pi_2 + \pi_3 = 1 \\ \pi_i \geq 0. \end{cases}$$

解得 $\pi_1 = \frac{3}{8}, \pi_2 = \frac{1}{4}, \pi_3 = \frac{3}{8}$.

例题 3.23 记 $\mathcal{S} = \{0, 1, 2, 3, 4\}$.

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

求平稳分布。

证明

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

$$\begin{cases} \pi_0 = \frac{1}{2}\pi_0 + \frac{1}{2}\pi_1 \\ \pi_2 = \frac{1}{2}\pi_4 \\ \pi_3 = \pi_2 \\ \pi_4 = \pi_3 + \frac{1}{2}\pi_4 \\ \pi_0 + \pi_1 + \pi_2 + \pi_3 + \pi_4 = 1 \\ \pi_i \geq 0, i = 0, 1, 2, 3, 4 \end{cases}$$

解得 $\pi = (\pi_0, \pi_0, \frac{1}{4} - \frac{1}{2}\pi_0, \frac{1}{4} - \frac{1}{2}\pi_0, \frac{1}{2} - \pi_0)$

$$\begin{cases} 0 \leq \pi_0 \leq 1 \\ 0 \leq \frac{1}{4} - \frac{1}{2}\pi_0 \leq 1, \end{cases}$$

其中 $0 \leq \pi_0 \leq \frac{1}{2}$.

3.6 状态空间的分解

3.6.1 质点在常返等价类中的转移

定理 3.6.1

设 i 为常返状态, 则

$$\lim_{n \rightarrow \infty} p_{ii}^{(nd(i))} = \frac{d(i)}{\mu_i}$$

定理 3.6.2

设 i 是常返状态, 则

1. i 是零常返的充分必要条件是 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$;
2. 当 i 是零常返, $i \rightarrow j$ 时, j 也是零常返的;
3. 若 i 是正常返, $i \rightarrow j$ 时, j 也是正常返的。

证明 (1) 设 i 为零常返态, 则 $\lim_{n \rightarrow \infty} p_{ii}^{(nd(i))} = 0$ 。由周期的定义可知, 当 n 不能被 $d(i)$ 整除时, $p_{ii}^{(n)} = 0$ 。故有 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$ 。

反之, 若 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$ 。下用反证法证明。假设 i 是正常返态, 由此可知 $\lim_{n \rightarrow \infty} p_{ii}^{(nd(i))} > 0$, 故 i 是零常返态。

(2) 设 $i \leftrightarrow j$, 取 m_0, n_0 使得 $p_{ij}^{(n_0)} > 0, p_{ji}^{(m_0)} > 0$ 。则有:

$$p_{ii}^{(n+m_0+n_0)} \geq p_{ij}^{(n_0)} p_{ji}^{(m_0)} p_{ii}^{(n)}, \quad n \geq 1.$$

由 $p_{ii}^{(n)} \rightarrow 0$, 可知 $p_{ii}^{(n)} \leq \frac{1}{p_{ij}^{(n_0)} p_{ji}^{(m_0)}} p_{ii}^{(n+m_0+n_0)} \rightarrow 0$

由此及 (1): 若 i 零常返, 则有 j 亦零常返。由 (2) 和定理??知 (3) 成立。

定理 3.6.3 (Lebesgue 控制收敛定理)

若 $\{f_n, n \geq 1\}$ 为可测函数序列, $|f_n| \leq Y$, Y 可积, 且对任意的 x $\lim_n f_n(x) = f(x)$ 存在, 则

$$\lim_n \mu(f_n) = \mu(f).$$

例题 3.24 如果 j 不是正常返, 则对任意状态 i , $p_{ij}^{(n)} \rightarrow 0, n \rightarrow \infty$ 。

证明 若 j 非常返, 由定理 ??(2) 可知 $p_{ij}^{(n)} \rightarrow 0$ 。对零常返 j , 由定理 ??(2) 可知 $p_{ij}^{(n)} \rightarrow 0$ 。对任意的 i ,

$$\begin{aligned} p_{ij}^{(n)} &= \sum_{k=1}^n f_{ij}^{(k)} p_{ij}^{(n-k)} \\ &= \sum_{k=1}^{\infty} f_{ij}^{(k)} p_{ij}^{(n-k)} I_{\{k \leq n\}}. \end{aligned}$$

由 $\sum_{k=1}^{\infty} f_{ij}^{(k)} \leq 1$ 。由控制收敛定理可知, 令 $n \rightarrow \infty$ 可知结论成立。(令 $g_n(k) := p_{ij}^{(n-k)} I_{\{k \leq n\}}$ 。则对任意的 $k \in \mathcal{S}$ 有, $\lim_{n \rightarrow \infty} g_n(k) = 0$ 。且 $|g_n(k)| \leq 1$)

定义 3.6.1 (闭集)

设 \mathcal{S} 是马氏链 $\{X_n\}$ 的状态空间, $i \in \mathcal{S}$ 。和 j 互通的状态记为

(a)

$$C(i) := \{i\} \cup \{j \in \mathcal{S}, j \leftrightarrow i\}.$$

若 i 不与别的状态互通, 则 $C(i) = \{i\}$ 。称 $C(i)$ 是一个等价类;

(b) 集合 $C \subseteq \mathcal{S}$ 称为闭的, 如果

$$p_{ij} = 0 \quad \forall i \in C, j \notin C.$$

即 $\sum_{k \in C} p_{ik} = 1 \quad \forall i \in C$ 。闭集 C 称为不可约的, 若 C 中的任意两个状态均是互通的。闭集 A 称为极小的, 若 A 的任意真子集不是闭集。

定理 3.6.4

集合 $C \subseteq \mathcal{S}$ 为闭的充要条件是对任意的 $i \in C$ 及 $j \notin C$, $n \geq 1$, 都有 $p_{ij}^{(n)} = 0$ 。

证明 下证必要性。若 C 为闭集, 当 $n = 1$ 时成立。

设当 $n = l$ 时成立。即对任意 $i \in C, j \notin C$, 有 $p_{ij}^{(l)} = 0$ 。

则对任意 $i \in C, j \notin C$

$$\begin{aligned} p_{ij}^{(l+1)} &= \sum_{k \in C} p_{ik}^{(l)} p_{kj} + \sum_{k \notin C} p_{ik}^{(l)} p_{kj} \\ &= \sum_{k \in C} p_{ik}^{(l)} \cdot 0 + \sum_{k \notin C} 0 \cdot p_{kj} = 0. \end{aligned}$$

定理 3.6.5

集合 $C \subseteq \mathcal{S}$ 为闭的充要条件是对任意的 $i \in C$ 及 $j \notin C, n \geq 1$, 都有 $p_{ij}^{(n)} = 0$ 。

注 集合 $C \subseteq \mathcal{S}$ 为闭充要条件是对任意的 $i \in C$ 有, $\sum_{j \in C} p_{ij}^{(n)} = 1$ 。整个马氏链也是一个闭集。

引理 3.6.1

设 C 为闭集, 只考虑 C 上的 m 步转移子矩阵 $P_C^{(m)} := (p_{ij}^{(m)}), i, j \in C$ 。则它为随机矩阵。

例题 3.25

$$P = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{pmatrix}$$

求极小闭集。

证明 $\{1, 3, 5\}$ 是极小闭集。

定理 3.6.6

设 C 是一个等价类，则

1. 不同等价类互不相交。
2. C 中的状态有相同的类型：或都是正常返的，或都是零常返的，或都是非常返的。在任何情况下， C 中的状态有相同的周期。
3. 所有常返状态构成一个闭集。同时常返等价类是闭集：质点不能走出常返等价类。
4. 零常返等价类含有无穷个状态。
5. 非常返等价类如果是闭集，则含有无穷个状态。
6. 设 T 为非常返状态全体。若 T 为闭集，则必含有无穷个状态。



证明 设 C 和 C_1 都是等价类，如果有 $i \in C \cap C_1$ ，由互通的传递性可知 i 和 $C \cup C_1$ 中的所有状态互通，于是 C 和 C_1 中的任意状态互通。故 $C = C_1$ 。

(2) 显然。

(3) 假设 C 为常返等价类或者所有常返态构成的集合。如果存在 $i \in C, j \notin C$ 且 $i \rightarrow j$ 则可知 $j \longleftrightarrow i$ 与 $j \notin C$ 矛盾。

(4,5) 同时证明：假设 C 为非常返的或零常返等价类。反证法，由 C 为闭集可知 $\sum_{j \in C} p_{ij}^{(n)} = 1$ 。则对 $\forall i, j \in C$ 有 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 。不妨假设 C 中有有限个点 (N) 。故令 $n \rightarrow \infty$,

$$1 = \lim_{n \rightarrow \infty} \sum_{j=1}^N p_{ij}^{(n)} = \sum_{j=1}^N \lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0,$$

矛盾。

定理 3.6.7

设 $C \neq \emptyset$ ，则它可以分成若干个互不相交的闭集 $\{C_n\}$ ，使得 $C = C_1 \cup C_2 \cup \dots$ ，且有

1. C_n 中任意两个状态互通；
2. $C_h \cap C_l = \emptyset (h \neq l)$ 。



证明 因 $C \neq \emptyset$ ，任取 $i_1 \in C$ ，令 $C_1 = \{i : i \longleftrightarrow i_1 \in C\}$ 。

若 $C - C_1 \neq \emptyset$ ，再任取 $i_2 \in C - C_1$ ，令 $C_2 = \{i : i \longleftrightarrow i_2 \in C - C_1\}$ ，...

若 $C - \bigcup_{l=1}^n C_l \neq \emptyset$ ，取 $i_{n+1} \in C - \bigcup_{l=1}^n C_l$ ，令 $C_{n+1} = \{i : i \longleftrightarrow i_{n+1} \in C - \bigcup_{l=1}^n C_l\}$ ，...

利用等价关系可将 S 分解：

$$S = \bigcup_{j=1}^m C_j \cup T, \quad m \leq \infty,$$

其中 C_j 为常返状态构成的不可约闭集， T 为非常返状态集。

注 对任意的 $k \in T$ ，

1. k 可能转移到 C_l 中，然后永远在 C_l 中。
2. T 若为有限集， T 一定非闭集，并且 T 中的任意质点一定走出 T 进入某闭集 C_l 。

例题 3.26 记 $S = \{1, 2, 3, 4, 5\}$ 。

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

证明 $\{1, 2\}$ 和 $\{3, 4, 5\}$ 均为等价类。

故 $\{1, 2\}$ 和 $\{3, 4, 5\}$ 均为正常返等价类。如果不是则由这两个闭, 可推出有无穷多个状态

例题 3.27 记 $S = \{1, 2, 3, 4, 5\}$ 。

$$\mathbf{P} = \begin{pmatrix} 0.6 & 0.1 & 0 & 0.3 & 0 \\ 0.2 & 0.5 & 0.1 & 0.2 & 0 \\ 0.2 & 0.2 & 0.4 & 0.1 & 0.1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

证明 故 $\{1, 2, 3\}$ 、 $\{4\}$ 、 $\{5\}$ 均为等价类。

故 $\{1, 2, 3\}$ 为非常返等价类、 $\{4\}$ 、 $\{5\}$ 为正常返等价类。

注 因为常返等价类 C 是一个不可约闭集, 故令 $P_C := (p_{ij})_{i,j \in C}$ 时, C 在 P_C 下为一个子马氏链。

定理 3.6.8

设常返等价类 C 有周期 $d > 1$ 。取定 C 中的状态 i , 对于 $j \in C$,

1. 有唯一的 $r \in \{1, 2, \dots, d\}$, 使得只要 $p_{ij}^{(n)} > 0$, 则有 $n = kd + r$;
2. 对于 (1) 中的 r , 存在 N_j 使得 $n > N_j$ 时, $p_{ij}^{(nd+r)} > 0$;
3. $f_{ij} = \sum_{n=0}^{\infty} f_{ij}^{(nd+r)} = 1$ 。



证明 (1) 由 $i \longleftrightarrow j$, 可知必存在 n 使得 $p_{ij}^{(n)} > 0$ 。显然存在 k 及 r 使得 $n = kd + r$ 。下证对于满足 $p_{ij}^{(n)} > 0$ 的 r 是唯一的。现假设 n, m 使得 $p_{ij}^{(n)} p_{ij}^{(m)} > 0$, 及 l 使得 $p_{jl}^{(l)} > 0$ 。则

$$\begin{aligned} p_{il}^{(n+l)} &\geq p_{ij}^{(n)} p_{jl}^{(l)} > 0, \\ p_{il}^{(m+l)} &\geq p_{ij}^{(m)} p_{jl}^{(l)} > 0, \end{aligned}$$

由于 $d \mid n+l$, $d \mid m+l$, 故 $d \mid (m+l) - (n+l)$, 即 $d \mid (m-n)$ 。故 n, m 跟 d 相除有相同的余数。

(2) 任取 N 使得 $p_{ij}^{(Nd+r)} > 0$ 。由定理 ?? 可知, 存在 M_j 使得 $\forall m > M_j$ 时, $p_{ij}^{(md)} > 0$ 。当 $n > N + M_j$ 时, $n - N > M_j$, 故

$$p_{ij}^{(nd+r)} \geq p_{ij}^{(Nd+r)} p_{ij}^{(nd-Nd)} = p_{ij}^{(Nd+r)} p_{ij}^{(md)} > 0.$$

即 $N_j := N + M_j$ 。

(3) 由于 $f_{ij}^{(n)}$ 是质点 i 出发第 n 步首次到达 j 的概率, 故 $f_{ij}^{(n)} \leq p_{ij}^{(n)}$ 。于是只要 n 使得 $f_{ij}^{(n)} > 0$ 就有 $p_{ij}^{(n)} > 0$, 从而 $n = kd + r$ 。由定理 ??(4) 可知

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} = \sum_{n=0}^{\infty} f_{ij}^{(nd+r)} = 1.$$

定理 3.6.9

设常返等价类 C 有周期 $d > 1$ 。

(a) C 可分成 d 个不相交的集合的并:

$$C = G_1 \cup G_2 \cup \cdots \cup G_d,$$

其中, 从任一 G_m 中的状态出发, 下一步到达 $G_{m+1(\bmod d)}$ 中某状态的概率大于 0;

(b) P^d 限制在 C 是非周期的, 且 $\sum_{k \in G_m} p_{ik}^{(d)} = 1 \forall i \in G_m$ 。且将 P^d 限制在 G_m 上时, 构成一个不可约非周期的子链。



定理 3.6.10

常返等价类是极小闭集。



3.6.2 不可约马氏链的极限定理

例题 3.28 设 Markov 链 $\{X_n, n = 0, 1, \dots\}$ 状态空间为 $S := \{1, 2\}$, 其转移矩阵为

$$\mathbf{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad 0 < p, q < 1$$

现在考虑 $\mathbf{P}^{(n)}$ 当 $n \rightarrow \infty$ 时的情况。

证明 由 $\mathbf{P}^{(n)} = \mathbf{P}^n$ 知, 只需计算 \mathbf{P} 的 n 重乘积的极限。求 \mathbf{P} 所对应的特征根: $|\lambda \mathbf{I} - \mathbf{P}| = 0$ 可得 $\lambda_1 = 1$ 以及 $\lambda_2 = 1 - p - q$ 。解得 $\lambda_1 = 1$ 是特征向量为 $(1, 1)^T$, $\lambda_2 = 1 - p - q$ 时特征向量为 $(-p, q)^T$ 。故令

$$\mathbf{Q} = \begin{pmatrix} 1 & -p \\ 1 & q \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 1 - p - q \end{pmatrix}$$

则

$$\mathbf{Q}^{-1} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ -\frac{1}{p+q} & \frac{1}{p+q} \end{pmatrix}, \quad \mathbf{P} = \mathbf{Q} \mathbf{D} \mathbf{Q}^{-1}$$

从而

$$\begin{aligned} \mathbf{P}^n &= (\mathbf{Q} \mathbf{D} \mathbf{Q}^{-1})^n = \mathbf{Q} \begin{pmatrix} 1 & 0 \\ 0 & 1 - p - q \end{pmatrix}^n \mathbf{Q}^{-1} \\ &= \begin{pmatrix} \frac{q+p(1-p-q)^n}{p+q} & \frac{p-p(1-p-q)^n}{p+q} \\ \frac{q-q(1-p-q)^n}{p+q} & \frac{p+q(1-p-q)^n}{p+q} \end{pmatrix} \end{aligned}$$

由于 $|1 - p - q| < 1$, (6.1) 式的极限为

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}$$

可见此 Markov 链的 n 步转移概率有一个稳定的极限。

定理 3.6.11 (不可约马氏链的极限定理)

设 P 不可约, 则下述断言等价:

1. P 是正常返 (不可约假设下, 一个状态正常返则所有状态正常返!);
2. P 有平稳分布, 进而平稳分布唯一: 对一切 j , $\pi_j = \frac{1}{\mu_j}$ (此处, $\mu_j = E[T_{jj}|X_0 = j]$);
3. 若 P 还是非周期的, 则 (1) 和 (2) 又都分别等价于 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{1}{\mu_j} > 0$ (不依于 i)



我们回到之前那一题

例题 3.29 设 Markov 链 $\{X_n, n = 0, 1, \dots\}$ 状态空间为 $S := \{1, 2\}$, 其转移矩阵为

$$\mathbf{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad 0 < p, q < 1$$

现在考虑 $\mathbf{P}^{(n)}$ 当 $n \rightarrow \infty$ 时的情况。

证明

$$\begin{cases} (\pi_1, \pi_2)\mathbf{P} = (\pi_1, \pi_2) \\ \pi_1 + \pi_2 = 1 \\ \pi_i \geq 0, (i = 1, 2) \end{cases} \quad (3.2)$$

解得 $\pi_1 = \frac{q}{q+p}, \pi_2 = \frac{p}{q+p}$ 。

故该马氏链平稳分布存在, 且马氏链不可约非周期。由不可约马氏链极限定理??可知

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}$$

例题 3.30 设马氏链的状态是 $S = \{1, 2\}$, 转移矩阵是

$$\mathbf{P} = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{5}{8} & \frac{3}{8} \end{pmatrix}$$

(1) 计算不变分布 π 和极限 $\lim_{n \rightarrow \infty} \mathbf{P}^n$; (2) 计算状态 1, 2 的期望返回时间 μ_1, μ_2 。

证明

$$\begin{cases} (\pi_1, \pi_2)\mathbf{P} = (\pi_1, \pi_2) \\ \pi_1 + \pi_2 = 1 \\ \pi_i \geq 0, (i = 1, 2) \end{cases}$$

$\pi_1 = \frac{5}{7}$ 以及 $\pi_2 = \frac{2}{7}$ 。

(2) 由于该马氏链为不可约非周期马氏链, 且平稳分布存在唯一。故该马氏链为正常返链。由不可约马氏链极限定理??可知

$$\begin{aligned} \lim_{n \rightarrow \infty} p_{11}^{(n)} &= \lim_{n \rightarrow \infty} p_{21}^{(n)} = \frac{5}{7}, \\ \lim_{n \rightarrow \infty} p_{12}^{(n)} &= \lim_{n \rightarrow \infty} p_{22}^{(n)} = \frac{2}{7}. \end{aligned}$$

(3) 由 (2) 分析以及不可约马氏链极限定理??可知 $\mu_1 = \frac{1}{\pi_1} = \frac{7}{5}, \mu_2 = \frac{1}{\pi_2} = \frac{7}{2}$ 。

3.6.3 一般情况下马氏链的极限定理

定义 3.6.2 (本质状态)

称状态 i 为本质的, 若 $i \rightarrow j$ 时必有 $j \rightarrow i$ 。否则称为非本质。



定理 3.6.12

常返状态是本质状态。



例题 3.31

$$S = \{0, 1, 2, 3, 4\}$$

$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

证明 0, 1, 2 是常返态, 4 是常返态, 3 非常返态。0, 1, 2 是本质态, 4 是本质态; 3 非本质态。

引理 3.6.2

设 C 为某个常返等价类。

记 $\tilde{\mathbf{P}} = (\tilde{p}_{ij})_{i,j \in C} := \mathbf{P}_C$, 其中 $i, j \in C$ 时 $\tilde{p}_{ij} = p_{ij}$ 。 $i, j \in C$ 时, 记 $\tilde{f}_{ij}^{(n)} = \tilde{\mathbf{P}}(T_{ij} = n | X_0 = i)$, 其中 $T_{ij} = n$ 为在 $\tilde{\mathbf{P}}$ 下状态 i 到达状态 j 的首达时间, 即说明 $\tilde{f}_{ij}^{(n)}$ 为在子马氏链 C 中状态 i 经过 n 步首次到达状态 j 的概率。则 $\tilde{f}_{ij}^{(n)} = f_{ij}^{(n)}$, 平均返回时间 $\tilde{\mu}_i = \mu_i$ 。

证明 用数学归纳法: 当 $n = 1$ 时,

$$\tilde{f}_{ij}^{(1)} = \tilde{p}_{ij} = p_{ij} = f_{ij}^{(1)}.$$

假设当 $n = k$ 时, 有 $\tilde{f}_{ij}^{(k)} = f_{ij}^{(k)}$ 。则当 $n = k + 1$ 时,

$$\begin{aligned} \tilde{f}_{ij}^{(k+1)} &= \sum_{x \in C, x \neq j} \tilde{p}_{ix} \tilde{f}_{xj}^{(k)} \\ &= \sum_{x \in C, x \neq j} p_{ix} f_{xj}^{(k)} + \sum_{x \notin C, x \neq j} p_{ix} f_{xj}^{(k)} \\ &= f_{ij}^{(k+1)}. \end{aligned}$$

引理 3.6.3

设 $\{a_n, n \geq 0\}$ 为一个不全为零的非负数列, 且满足条件

$$\lim_{n \rightarrow \infty} \frac{a_n}{\sum_{m=0}^n a_m} = 0,$$

$\{b_n, n \geq 0\}$ 为一个收敛数列, 则

$$\lim_{n \rightarrow \infty} \frac{\sum_{m=0}^n a_m b_{n-m}}{\sum_{m=0}^n a_m} = \lim_{n \rightarrow \infty} b_n.$$

$$\text{约定 } \mu_j = \begin{cases} \mu_j & \text{常返} \\ \infty & \text{非常返} \end{cases}$$

定理 3.6.13

1. 如果 j 非常返, 则对一切 $i \in S$, 有 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 。
2. 如果 j 常返, 其周期为 $d(j)$, 则对一切 $i \in S$ 和 $1 \leq r \leq d(j)$, 有

$$\lim_{n \rightarrow \infty} p_{ij}^{(nd(j)+r)} = \frac{d(j)}{\mu_j} \sum_{m=0}^{\infty} f_{ij}^{(md(j)+r)}.$$

特别地,

- a j 为零常返时 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 。
- b j 为本质状态 (或 i 为常返态), 若 $i \notin C(j)$, 则对任意的 n 有 $p_{ij}^{(n)} = 0$, 则 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 。

c 若 $i \in C(j)$ 且 j 属于 $C(i)$ 周期性分解中第 r 个子集 (即若 $p_{ij}^{(m)} > 0$ 有 $m = nd(j) + r$), 则

$$\lim_{n \rightarrow \infty} p_{ij}^{(nd(j)+r)} = \frac{d(j)}{\mu_j}.$$

否则 $p_{ij}^{(n)} = 0$, 当 $n \neq md + r$ 。



推论 3.6.1

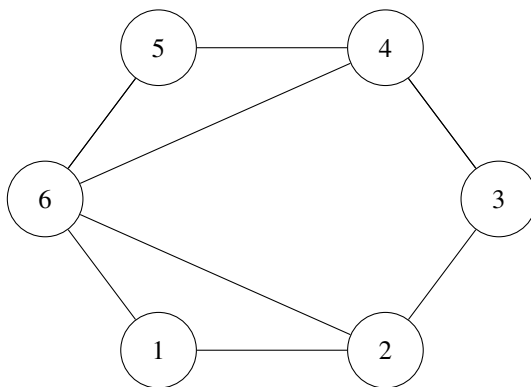
设 j 为非周期状态。对任意的 $i \in S$,

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{f_{ij}}{\mu_j},$$

其中当 j 为非常返时, $\mu_j := \infty$ 。



例题 3.32 设有 6 个车站, 车站中间的公路连接情况如图所示。



汽车每天可以从一个站驶向与之直接相邻的车站, 并在夜晚到达车站留宿, 次日凌晨重复相同的活动。设每天凌晨汽车开往临近的任何一个车站都是等可能的, 试说明很长时间后, 各站每晚留宿的汽车比例趋于稳定。求出这个比例以便正确地设置各站的服务规模。

证明 以 $\{X_n, n = 0, 1, \dots\}$ 记第 n 天某辆汽车留宿的车站号, 这是一个 Markov 链, 转移概率矩阵为

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \end{pmatrix}$$

解方程

$$\begin{cases} \pi P = \pi \\ \sum_{i=1}^6 \pi_i = 1 \end{cases}$$

其中 $\pi = (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6)$, 可得 $\pi = (\frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{3}{16}, \frac{1}{8}, \frac{1}{4})$ 。故无论开始汽车从哪一个车站出发在很长时间后他在任一个车站留宿的概率都是固定的, 所有的汽车也将以一个稳定的比例在各车站留宿。

例题 3.33 设甲袋中有 k 个白球和 1 个黑球, 乙袋中有 $k+1$ 个白球, 每次从两袋中各任取一球, 交换后放入对方的袋中。证明经过 n 次交换后, 黑球仍在甲袋中的概率 p_n 满足 $\lim_{n \rightarrow \infty} p_n = \frac{1}{2}$ 。

证明 以 X_n 表示第 n 次取球后甲袋中的黑球数, 则 $\{X_n, n = 0, 1, 2, \dots\}$ 是状态空间为 $S = \{0, 1\}$ 的时齐 Markov 链, 一步转移概率矩阵为

$$P = \begin{pmatrix} \frac{k}{k+1} & \frac{1}{k+1} \\ \frac{1}{k+1} & \frac{k}{k+1} \end{pmatrix}$$

则它的平稳分布满足

$$\begin{cases} \pi_0 = \frac{k}{k+1}\pi_0 + \frac{1}{k+1}\pi_1 \\ \pi_1 = \frac{1}{k+1}\pi_0 + \frac{k}{k+1}\pi_1 \end{cases}$$

且有 $\pi_0 + \pi_1 = 1$,

求解得 $\pi = (\pi_0, \pi_1) = (\frac{1}{2}, \frac{1}{2})$ 。故

$$\lim_{n \rightarrow \infty} P^n = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

由于初始分布 $\pi(0) = (0, 1)$,

$$\lim_{n \rightarrow \infty} \pi(0)P^n = \pi(0) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \left(\frac{1}{2}, \frac{1}{2}\right)$$

故经过 n 次交换后, 黑球仍在甲袋中的概率 p_n 满足

$$\lim_{n \rightarrow \infty} p_n = \lim_{n \rightarrow \infty} P\{X_n = 1\} = \pi_1 = \frac{1}{2}$$

如果把关于 n 的逐点极限改为平均极限, 则结果十分简洁, 并有下列结果:

定理 3.6.14

对一切 $i, j \in S$, 有

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ij}^{(m)} = \frac{f_j}{\mu_j}.$$

其中当 j 非常返时 $\mu_j := \infty$ 。



证明 当 j 为非常返时, 由 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$,

故 $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ij}^{(m)} = \lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$ 且 $\mu_j = \infty$ (约定), 则上述结论明显成立。

下面考虑 j 为常返的情形。为此, 需一个数学分析中的结论。

引理 3.6.4

设有正整数 d 和数列 $\{a_n : n \geq 1\}$ 满足

$$\lim_{n \rightarrow \infty} a_{nd+r} = b_r \quad 1 \leq r \leq d.$$

则有:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n a_m = \frac{1}{d} \sum_{m=1}^d b_m.$$



当 j 为常返时, 令 $d = d(j)$, 由定理 ?? (b) 得:

$$\lim_{n \rightarrow \infty} p_{ij}^{(nd+r)} = \frac{d}{\mu_j} \sum_{m=0}^{\infty} f_{ij}^{(md+r)} \quad 1 \leq r \leq d.$$

现再由引理得:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n p_{ij}^{(m)} &= \frac{1}{d} \sum_{r=1}^d \left[\lim_{n \rightarrow \infty} p_{ij}^{(nd+r)} \right] \\ &= \frac{1}{d} \sum_{r=1}^d \left[\frac{d}{\mu_j} \sum_{m=0}^{\infty} f_{ij}^{(md+r)} \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\mu_j} \sum_{r=1}^d \sum_{m=0}^{\infty} f_{ij}^{(md+r)} \\
&= \frac{1}{\mu_j} f_{ij}.
\end{aligned}$$

定理 3.6.15

对任意的 $j \in S$, $\{f_{ij}, i \in S\}$ 是线性方程组

$$z_i = \sum_{k \neq j} p_{ik} z_k + p_{ij}, \quad i \in S,$$

即

$$f_{ij} = \sum_{k \neq j} p_{ik} f_{kj} + p_{ij}, \quad i \in S,$$

的解. (见《随机过程引论》何声武 49 页)

**定理 3.6.16**

设 $j, k \in S$ 属于同一常返类, 则对任意的 $i \in T$ (T 表示非常返状态全体) 有 $f_{ij} = f_{ik}$. (见《随机过程引论》何声武 50 页)



例题 3.34 记 $S = \{0, 1, 2, 3, 4, 5\}$.

$$\mathbf{P} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 0 & 1/4 & 1/4 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1/3 & 2/3 \end{pmatrix}$$

求 f_{ij} 以及 $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=1}^n p_{ij}^{(\nu)}$.

证明

$$\mathbf{P}|_{G_1} = \begin{pmatrix} 1/4 & 1/4 & 1/2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{P}|_{G_2} = \begin{pmatrix} 0 & 1 \\ 1/3 & 2/3 \end{pmatrix}$$

$G_1 := \{1, 2, 3\}$, $G_2 := \{4, 5\}$ 为正常返等价类. $\{0\}$ 为非常返等价类. 记 (π'_1, π'_2, π'_3) 以及 (π'_4, π'_5) 为在 $\mathbf{P}|_{G_1}$ 和 $\mathbf{P}|_{G_2}$ 上得到平稳分布. 则

$$\begin{cases} (\pi'_1, \pi'_2, \pi'_3) \mathbf{P}|_{G_1} = (\pi'_1, \pi'_2, \pi'_3) \\ \pi'_1 + \pi'_2 + \pi'_3 = 1 \\ \pi'_i \geq 0. \quad (i = 1, 2, 3) \end{cases}$$

解得 $\pi'_1 = \frac{4}{9}, \pi'_2 = \frac{1}{3}, \pi'_3 = \frac{2}{9}$.

记 (π'_1, π'_2, π'_3) 以及 (π'_4, π'_5) 为在 $\mathbf{P}|_{G_1}$ 和 $\mathbf{P}|_{G_2}$ 上得到平稳分布. 则

$$\begin{cases} (\pi'_4, \pi'_5) \mathbf{P}|_{G_2} = (\pi'_4, \pi'_5) \\ \pi'_4 + \pi'_5 = 1 \\ \pi'_i \geq 0. \quad (i = 4, 5) \end{cases}$$

解得 $\pi'_4 = \frac{1}{4}, \pi'_5 = \frac{3}{4}$. 猜猜 f_{ij} 取多少???

$$(\mathbf{f}_{ij})_{i,j \in S} = \begin{pmatrix} * & * & * & * & * & * \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$f_{00} = \frac{1}{6}.$$

$$f_{01} = p_{00}f_{01} + p_{02}f_{21} + p_{03}f_{31} + p_{04}f_{41} + p_{05}f_{51} + p_{01} = \frac{1}{6}f_{01} + \frac{1}{2}. \text{ 解得 } f_{01} = \frac{3}{5}.$$

$$f_{04} = p_{00}f_{04} + p_{01}f_{14} + p_{02}f_{24} + p_{03}f_{34} + p_{05}f_{54} + p_{04} = \frac{1}{6}f_{04} + \frac{1}{3}. \text{ 解得 } f_{04} = \frac{2}{5}.$$

则

$$(\mathbf{f}_{ij})_{i,j \in S} = \begin{pmatrix} 1/6 & 3/5 & 3/5 & 3/5 & 2/5 & 2/5 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$f_{01} = \frac{3}{5}, \frac{1}{\mu_1} = \pi'_1 = \frac{4}{9}. \text{ 由 } \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=1}^n p_{ij}^{(\nu)} = \frac{f_{ij}}{\mu_j} \text{ 则}$$

$$\left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\nu=1}^n \mathbf{p}_{ij}^{(\nu)} \right)_{i,j \in S} = \begin{pmatrix} 0 & 4/15 & 1/5 & 2/15 & 1/10 & 3/10 \\ 0 & 4/9 & 1/3 & 2/9 & 0 & 0 \\ 0 & 4/9 & 1/3 & 2/9 & 0 & 0 \\ 0 & 4/9 & 1/3 & 2/9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 0 & 0 & 0 & 1/4 & 3/4 \end{pmatrix}$$

注

- 有限不可约马氏链的状态都是常返状态。
- 有限状态马氏链没有零常返态。
- 不可约的有限马氏链的状态都是正常返状态。
- 有限马氏链必存在正常返态。

定理 3.6.17

存在平稳分布的充要条件是存在正常返类。若正常返类记为 D_1, D_2, \dots , 则平稳分布具有下列形式:

$$\pi_i = \begin{cases} \lambda_n \frac{1}{\mu_i} & i \in D_n \\ 0 & i \notin \bigcup_n D_n \end{cases}$$

其中 $\sum_n \lambda_n = 1$, 且对任意的 n , $\lambda_n \geq 0$



定理 3.6.18

设 C_+ 是马氏链 $\{X_n\}$ 的所有正常返状态。

1. $\{X_n\}$ 平稳 (不变) 分布存在的充分必要条件是 C_+ 非空;
2. $\{X_n\}$ 有唯一的平稳 (不变) 分布的充分必要条件是 C_+ 是等价类;
3. 状态有限的马氏链必有平稳 (不变) 分布。



第4章 鞅论

4.1 离散时间鞅

例题 4.1 在赌局中, 某人根据前面各次赌博结果决定下次赌博的金额。设本金为 X_0 , 记第 n 次赌博后本金记为 X_n 。设 $\{Y_n, n \geq 1\}$ 为独立同分布的随机变量序列,

$$P(Y_n = 1) = P(Y_n = -1) = \frac{1}{2},$$

表示每次赌博的结果。

证明 令 $\{Y_n, n = 1, 2, \dots\}$, 是一列独立同分布的随机变量, 表示每次赌博的结果

$$P\{Y_n = 1\} = P\{Y_n = -1\} = \frac{1}{2}$$

这里 $\{Y_n = 1\}$ ($\{Y_n = -1\}$) 表示赌博者在第 n 次赌博时的赢(输)。如果赌博者采用的赌博策略(即所下赌注)依赖于前面的赌博结果, 那么他的赌博可以用下面的随机变量序列

$$b_n = b_n(Y_1, \dots, Y_{n-1}), \quad n = 2, 3, \dots$$

描述, 其中 $b_n < \infty$ 是第 n 次的赌注, 若赌赢则获利 b_n , 否则输掉 b_n 。

设 X_0 是该赌博者的初始赌资, 则

$$X_n = X_0 + \sum_{i=1}^n b_i Y_i$$

是他在第 n 次赌博后的赌资。可以断言

$$E[X_{n+1}|Y_1, \dots, Y_n] = X_n.$$

事实上, 由上式我们可以得到

$$X_{n+1} = X_n + b_{n+1}Y_{n+1},$$

因此

$$\begin{aligned} E[X_{n+1}|Y_1, \dots, Y_n] &= E[X_n|Y_1, \dots, Y_n] + E[b_{n+1}Y_{n+1}|Y_1, \dots, Y_n] \\ &= X_n + b_{n+1}E[Y_{n+1}|Y_1, \dots, Y_n] \quad (\text{因为 } X_n \text{ 与 } b_{n+1} \text{ 由 } Y_1, \dots, Y_n \text{ 确定}) \\ &= X_n + b_{n+1}E[Y_{n+1}] \quad (\text{因为 } \{Y_n\} \text{ 是独立随机变量序列}) \\ &= X_n \quad (\text{因为 } E[Y_{n+1}] = 0, \forall n \geq 0). \end{aligned}$$

这证明了, 如果每次赌博的输赢机会是均等的, 并且赌博策略是依赖于前面的赌博结果, 则赌博是“公平的”。因此任何赌博者都不可能将公平的赌博通过改变赌博策略使得赌博变成有利于自己的赌博。

定义 4.1.1 (鞅)

称 $(X_n)_{n \geq 0}$ 关于 $(Y_n)_{n \geq 0}$ 为鞅, 若 $\forall n \in \mathbb{N}$,

- X_n 可积;
- $E(X_{n+1}|Y_0, Y_1, \dots, Y_n) = X_n$ a.s.



注 由 $E(X_{n+1}|Y_0, Y_1, \dots, Y_n) = X_n$ 可知, X_n 为 Y_0, Y_1, \dots, Y_n 的函数。

定义 4.1.2 (上鞅)

称 $(X_n)_{n \geq 0}$ 关于 $(Y_n)_{n \geq 0}$ 上鞅, 若 $\forall n \in \mathbb{N}$,

- $E(X_n^-) < \infty$; 其中 $x^- := -\min\{x, 0\}$
- $E(X_{n+1}|Y_0, Y_1, \dots, Y_n) \leq X_n$ a.s.

- X_n 为 Y_0, Y_1, \dots, Y_n 的函数



定义 4.1.3 (下鞅)

称 $(X_n)_{n \geq 0}$ 关于 $(Y_n)_{n \geq 0}$ 下鞅, 若 $\forall n \in \mathbb{N}$,

- $E(X_n^+) < \infty$; 其中 $x^+ := \max\{x, 0\}$
- $E(X_{n+1}|Y_0, Y_1, \dots, Y_n) \geq X_n$ a.s.
- X_n 为 Y_0, Y_1, \dots, Y_n 的函数



4.1.1 条件期望的性质

定理 4.1.1

1. 若 X, Y 均可积, α, β 为常数, 则

$$E[\alpha X + \beta Y | \mathcal{F}_0] = \alpha E[X | \mathcal{F}_0] + \beta E[Y | \mathcal{F}_0]; \text{ a.e.}$$

2. $E[1 | \mathcal{F}_0] = 1$; a.e.
3. 当 $X \geq 0$ 时, 则 $E[X | \mathcal{F}_0] \geq 0$ a.e.
($X \geq Y \implies E[X | \mathcal{F}_0] \geq E[Y | \mathcal{F}_0]$ a.e.);
4. $|E[X | \mathcal{F}_0]| \leq E[|X| | \mathcal{F}_0]$ a.e.
5. $E[E[X | \mathcal{F}_0]] = E[X]$ a.e.



定理 4.1.2

1. 若 Y 为 \mathcal{F}_0 可测, 且 X, XY 可积, 则

$$E[XY | \mathcal{F}_0] = Y E[X | \mathcal{F}_0] \text{ a.e.};$$

2. 若 $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}$, 则有

$$E[E[X | \mathcal{F}_2] | \mathcal{F}_1] = E[X | \mathcal{F}_1] = E[E[X | \mathcal{F}_1] | \mathcal{F}_2] \text{ a.e.}$$

3. 若 X 可积, $\sigma(X)$ 与 \mathcal{F}_0 独立, 则 $E[X | \mathcal{F}_0] = E[X]$ a.e.;
特别, 当 X, Y 相互独立时, $E[X | Y] = E[X]$ a.e..



本节假定所有 r.v. 均定义在同一个概率空间 (Ω, \mathcal{F}, P) 上. $\mathbb{N} :=$ 非负整数全体

定义 4.1.4 (σ -代数流)

称 \mathcal{F} 的子 σ -域序列 $(\mathcal{F}_n, n \in \mathbb{N})$ 为一个流, 若 $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$.

记 $\mathcal{F}_\infty = \bigvee_{n \geq 0} \mathcal{F}_n = \sigma(\bigcup_n \mathcal{F}_n)$



定义 4.1.5

随机过程 $X = (X_n)_{n \geq 0}$ 为 $(\mathcal{F}_n, n \in \mathbb{N})$ 适应的, 若 $\forall n, X_n$ 为 \mathcal{F}_n 可测的。(回忆可测的定义? \forall 实数 a , $\{\omega | X_n(\omega) \leq a\} \in \mathcal{F}_n$)



定义 4.1.6

一个 $(\mathcal{F}_n, n \in \mathbb{N})$ -适应的随机序列 $X = (X_n)_{n \geq 0}$ 称为 $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -鞅, 若 $\forall n \in \mathbb{N}$,

- $E[|X_n|] < \infty$;
- $E(X_{n+1} | \mathcal{F}_n) = X_n$ a.s.



定义 4.1.7

一个 $(\mathcal{F}_n, n \in \mathbb{N})$ -适应的随机序列 $X = (X_n)_{n \geq 0}$ 称为 $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -上鞅 (下鞅), 若 $\forall n \in \mathbb{N}$,

- $E[X_n] < \infty; (E[X_n^+] < \infty)$
- $E(X_{n+1}|\mathcal{F}_n) \leq X_n, (\geq X_n) \text{ a.s.}$



注 当 $\{X_n\}$ 分别为 $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -鞅、上鞅、下鞅时, $\forall m > n$ 有

- $E(X_m|\mathcal{F}_n) = X_n (\leq X_m, \geq X_m) \text{ a.s.}$
- $E(X_m) = E(X_n) (\leq E(X_n), \geq E(X_n))$

注 若 $X = \{X_n, n \geq 0\}$ 为 $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -上鞅 (下鞅), 则 $-X := \{-X_n, n \geq 0\}$ 为 $(\mathcal{F}_n)_{n \in \mathbb{N}}$ -下鞅 (上鞅)

例题 4.2 设 X_1, X_2, \dots 是一族零均值独立随机变量序列, 且 $E[|X_i|] < \infty$, 令 $S_0 = 0, S_n = \sum_{k=1}^n X_k$, 则 $\{S_n\}$ 是关于 σ -代数流 $\{\mathcal{F}_n\}$ 的鞅, 其中 $\mathcal{F}_n := \sigma(X_1, X_2, \dots, X_n)$. 另外, 若 $X_n (n = 1, 2, \dots)$ 均值为 $\mu \neq 0$, 则 $\{M_n = S_n - n\mu\}$ 是 (关于 $\{\mathcal{F}_n\}$) 的鞅。

证明 当 $E[X_n] = 0, (n = 1, 2, \dots)$ 时,

- S_n 是 X_1, \dots, X_n 的函数, 故 S_n 关于 \mathcal{F}_n 可测的, 故 $\{S_n\}$ 关于 $\{\mathcal{F}_n\}$ 适应
- $E[|S_n|] \leq \sum_{i=1}^n E[|X_i|] < \infty$,

$$\begin{aligned} E[S_{n+1}|\mathcal{F}_n] &= E[X_1 + X_2 + \dots + X_n + X_{n+1}|\mathcal{F}_n] \\ &= E[X_1 + \dots + X_n|\mathcal{F}_n] + E[X_{n+1}|\mathcal{F}_n] \\ &= S_n + E[X_{n+1}] = S_n \end{aligned}$$

从而 $\{S_n\}$ 是一个关于 $\{\mathcal{F}_n\}$ 的鞅。

例题 4.3 在例 1.3 中设 $E[X_k] = \mu \neq 0, E[|X_k|] < \infty, (k = 1, 2, \dots)$, 则有 $E[|S_n|] < \infty$,

$$E[S_{n+1}|\mathcal{F}_n] = E\left[\sum_{i=1}^n X_i + X_{n+1}|\mathcal{F}_n\right] = S_n + \mu$$

显然, 若 $\mu > 0 (\mu < 0)$, 则 $\{S_n\}$ 是一关于 $\{\mathcal{F}_n\}$ 的下鞅 (上鞅)。

例题 4.4 Polya 坛子抽样模型 考虑一个装有红、黄两色球的坛子. 假设最初坛子中装有红、黄两色球各一个, 每次都按如下规则有放回地随机抽取: 如果拿出的是红色的球, 则放回的同时再加入一个同色的球; 如果拿出的是黄色的球也采取同样的方法. 以 X_n 表示第 n 次抽取后坛子中的红球数, 则 $X_0 = 1$, 且 $\{X_n\}$ 是一个非时齐的 Markov 链, 转移概率为

$$\begin{aligned} P\{X_{n+1} = k+1 | X_n = k\} &= \frac{k}{n+2} \\ P\{X_{n+1} = k | X_n = k\} &= \frac{n+2-k}{n+2} \end{aligned}$$

证明 显然, 若 $\mu > 0 (\mu < 0)$, 则 $\{S_n\}$ 是一关于 $\{\mathcal{F}_n\}$ 的下鞅 (上鞅)。

令 M_n 表示第 n 次抽取后红球所占的比例, 则 $M_n = \frac{X_n}{n+2}$, 并且 $\{M_n\}$ 是一个关于 $\{\mathcal{F}_n\}$ 的鞅, 其中 $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. 这是因为

$$\begin{aligned} E[X_{n+1}|X_n = k] &= \frac{(k+1)k}{n+2} + \frac{k(n+2-k)}{n+2} = k + \frac{k}{n+2} \\ E[X_{n+1}|X_n] &= X_n + \frac{X_n}{n+2} \text{ a.e.} \end{aligned}$$

- 对任意的 n , X_n 关于 \mathcal{F}_n 可测, 故 M_n 关于 \mathcal{F}_n 可测. 因此 $\{M_n\}$ 关于 $\{\mathcal{F}_n\}$ 适应.
- $E[|M_n|] \leq 1$.
- 由于 $\{X_n\}$ 是一个 Markov 链, 所以

$$E[M_{n+1}|\mathcal{F}_n] = E[M_{n+1}|X_n] = E\left[\frac{X_{n+1}}{n+1+2}|X_n\right] = \frac{1}{n+3}E[X_{n+1}|X_n] = \frac{1}{n+3}\left(X_n + \frac{X_n}{n+2}\right) = \frac{X_n}{n+2} = M_n \text{ a.e.}$$

本例研究的模型是 Polya 首次引入的, 它适用于描述群体增值和传染病的传播等现象.

例 4.5 设 ξ 为一个可积随机变量, 令 $X_n := E[\xi|\mathcal{F}_n]$, 则 $\{X_n\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -鞅.

证明

- 显然由条件期望的性质可知, $X_n = E[\xi|\mathcal{F}_n]$ 关于 \mathcal{F}_n 可测. 故 $\{X_n\}$ 关于 $\{\mathcal{F}_n\}$ 适应.
- $E[|X_n|] = E[|E[\xi|\mathcal{F}_n]|] \leq E[E[|\xi||\mathcal{F}_n]] = E[|\xi|] < \infty$.
- $E[X_{n+1}|\mathcal{F}_n] = E[E[\xi|\mathcal{F}_{n+1}]|\mathcal{F}_n] = E[\xi|\mathcal{F}_n] = X_n$. a.e.

例 4.6 设 (ξ_n) 为 \mathcal{F}_n -适应可积随机变量序列, 若对任意的 $n \geq 0$, ξ_{n+1} 与 \mathcal{F}_n 独立. 若对任意的 $n \geq 1$, $E(\xi_n) = 0$ ($< 0, \geq 0$), 令 $X_n := \sum_{i=0}^n \xi_i$, 则 $\{X_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}$ 鞅 (上, 下鞅).

证明 证鞅的情形

- 对给定的 n , 由于每个 ξ_i 关于 \mathcal{F}_i 可测, 且当 $i \leq n$ 时, $\mathcal{F}_i \subset \mathcal{F}_n$. 故 $X_n := \sum_{i=0}^n \xi_i$ 关于 \mathcal{F}_n 可测.
- $E[|X_n|] = E[|\sum_{i=0}^n \xi_i|] \leq \sum_{i=0}^n E[|\xi_i|] < \infty$.
- $E[X_{n+1}|\mathcal{F}_n] = E[X_n + \xi_{n+1}|\mathcal{F}_n] = X_n + E[\xi_{n+1}|\mathcal{F}_n] = X_n$. a.e.

设 $\{Y_n, n \geq 0\}$ 是马尔可夫链, 具有转移概率矩阵 $P = (p_{ij})$, f 满足 $f(i) \geq 0$, 且 $f(i) = \sum_{j \in S} p_{ij}f(j)$, $|f(i)| < M, i \in S$. 令 $X_n = f(Y_n)$, 则 $\{X_n, n \geq 0\}$ 关于 $\{Y_n, n \geq 0\}$ 是鞅.

$\mathcal{F}_n := \sigma(Y_i, 0 \leq i \leq n)$. X_n 是 Y_n 的函数, 故 X_n 关于 \mathcal{F}_n 可测的, 故 $\{X_n\}$ 关于 $\{\mathcal{F}_n\}$ 适应

$$E[|X_n|] < \infty,$$

$$\begin{aligned} E[X_{n+1}|Y_0, Y_1, \dots, Y_n] &= E[f(Y_{n+1})|Y_0, Y_1, \dots, Y_n] \\ &= E(f(Y_{n+1})|Y_n) \\ &= \sum_{j \in S} f(j)P(Y_{n+1} = j|Y_n) \\ &= \sum_{j \in S} f(j)P_{Y_n j} = f(Y_n) = X_n. \end{aligned}$$

4.1.2 鞅的性质

定义 4.1.8

对随机过程 $(X_n)_{n \geq 0}$, 令

$$\mathbb{F}^0(X) = (\mathcal{F}_n^0(X), n \geq 0), \mathcal{F}_n^0(X) = \sigma(X_0, X_1, \dots, X_n), n \geq 0.$$

$\mathbb{F}^0(X)$ 为 X 的自然流, 即使得 X 适应的最小流。



定义 4.1.9

简称随机过程 $(X_n)_{n \geq 0}$ 为鞅 (上鞅, 下鞅), 若 $(X_n)_{n \geq 0}$ 为 $(\mathcal{F}_n^0)_{n \in \mathbb{N}}$ -鞅 (上鞅, 下鞅)。



定理 4.1.3

设 $X = \{X_n, n \geq 0\}, Y = \{Y_n, n \geq 0\}$ 为两个 $\{\mathcal{F}_n\}$ 鞅 (上鞅), 则 $X + Y = \{X_n + Y_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}$ 鞅 (上鞅), $X \wedge Y = \{X_n \wedge Y_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}$ 上鞅。



证明 假设 $X = \{X_n, n \geq 0\}, Y = \{Y_n, n \geq 0\}$ 为两个 $\{\mathcal{F}_n\}$ 上鞅, 只证 $X \wedge Y = \{X_n \wedge Y_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}$ 上鞅. 对任意的 n ,

- $(X_n \wedge Y_n)^- := \max\{-(X_n \wedge Y_n), 0\} = \max\{(-X_n) \vee (-Y_n), 0\}$
- $\leq \max\{(-X_n), 0\} + \max\{(-Y_n), 0\} = X_n^- + Y_n^-$

故 $E[(X_n \wedge Y_n)^-] \leq E[X_n^-] + E[Y_n^-] < \infty$,

对任意的 n ,

- $X_n \wedge Y_n = \frac{X_n + Y_n - |X_n - Y_n|}{2}$, 故 $X_n \wedge Y_n$ 关于 \mathcal{F}_n 可测, 即 $X \wedge Y$ 关于 $\{\mathcal{F}_n\}$ 适应.
- $E[X_{n+1} \wedge Y_{n+1} | \mathcal{F}_n] \leq E[X_{n+1} | \mathcal{F}_n] \wedge E[Y_{n+1} | \mathcal{F}_n] \leq X_n \wedge Y_n$ a.e.

定理 4.1.4

设 $X = \{X_n, n \geq 0\}, Y = \{Y_n, n \geq 0\}$ 为两个 $\{\mathcal{F}_n\}$ 上鞅, a, b 为两个正数, 则 $\{aX_n + bY_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}$ 上鞅。

**定理 4.1.5**

设 f 为 $(a, b) \subset (-\infty, +\infty)$ 上的连续有限凸函数, X 为取值 (a, b) 的可积随机变量, 则有

$$f(E[X]) \leq E[f(X)].$$

**定理 4.1.6**

若 X 及 $f(X)$ 为可积随机变量, $f(x)$ 为 \mathbb{R} 上的有限连续凸函数, 则

$$\begin{aligned} f(E[X | \mathcal{F}_0]) &\leq E[f(X) | \mathcal{F}_0] \\ (E[X])^2 &\leq E[X^2]. (E[X | \mathcal{F}_0])^2 \leq E[X^2 | \mathcal{F}_0]. \end{aligned}$$

**定义 4.1.10**

设 f 为定义在区间 I 上的函数, 若对 I 上的任意两点 x_1, x_2 和任意的实数 $\lambda \in (0, 1)$, 总有 $f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$ 则称为 I 上的凸函数。



判定方法可利用定义法以及函数的二阶导数

对于实数集上的凸函数, 一般的判别方法是求它的二阶导数, 如果其二阶导数在区间上非负, 就为凸函数。

定理 4.1.7

若 $X = \{X_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -鞅 (下鞅), f 为 \mathbb{R} 上的连续凸 (凸增) 函数。若 $f(X_n)$ 可积 ($\forall n \geq 0$), 则 $\{f(X_n), n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -下鞅。



证明 只证下鞅的情形:

- 由 $\{X_n\}$ 为下鞅, $f(X_n)$ 关于 \mathcal{F}_n 可测, 故 $\{f(X_n)\}$ 关于 $\{\mathcal{F}_n\}$ 适应.
- $E[|f(X_n)|] < \infty$.
- 由于 f 凸增函数, $f(X_n) \leq f(E[X_{n+1} | \mathcal{F}_n]) \leq E[f(X_{n+1}) | \mathcal{F}_n]$ a.e.

推论 4.1.1

设 $X = \{X_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -下鞅, $X \vee a := \{X_n \vee a, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -下鞅. 特别地 $a = 0$ 时, $X^+ := \{X_n^+, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -下鞅.



证明 任取 $0 \leq \lambda \leq 1, x, y$, 则可知

$$\begin{aligned} \lambda(x \vee a) + (1 - \lambda)(y \vee a) &= (\lambda x \vee a) + [(1 - \lambda)y \vee (1 - \lambda)a] \\ &\geq [\lambda x + (1 - \lambda)y] \vee a \end{aligned}$$

综合可知 $x \vee a$ 为一个凸函数。

定义 4.1.11

若 $X = \{X_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -鞅 (上, 下鞅), 若 $X_n \in L^p(\Omega, \mathcal{F}, P) (p \geq 1)$, 则称 $X = \{X_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ - L^p 鞅 (上, 下鞅)。若 $X = \{X_n, n \geq 0\}$ 满足 $\sup_n E[|X_n|^p] < \infty$, 则称为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ - L^p 有界鞅 (上, 下鞅)。



推论 4.1.2

设 $X = \{X_n, n \geq 0\}$ 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}-L^p$ 鞅, $|X|^p := \{|X_n|^p, n \geq 0\}$ 为下鞅。



证明 当 $p = 1$ 时, 任取 x, y 以及 $0 \leq \lambda \leq 1$ 可知,

$$|\lambda x + (1 - \lambda)y| \leq \lambda|x| + (1 - \lambda)|y|, \text{ 故此时为凸函数.}$$

当 $p > 1$ 时, 求二阶导数, 可知任意点二阶导数非负, 故仍为凸函数.

定义 4.1.12

随机序列 $X = \{X_n, n \geq 0\}$ 称为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ -可料的, 如果 X_0 为 \mathcal{F}_0 -可测, 且对任意的 $n \geq 1$, X_n 为 \mathcal{F}_{n-1} 可测的. $A = \{A_n, n \geq 0\}$ 称为增序列, 如果对任意的 $n \geq 0$, $0 \leq A_n \leq A_{n+1}$ a.s.. 定义 $A_\infty := \lim_{n \rightarrow \infty} A_n$, 增序列 $\{A_n\}$ 称为可积的, 若 $E[A_\infty] < \infty$.

**定理 4.1.8**

设 $X = \{X_n, n \geq 0\}$ 为一 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 下鞅, 则 X 可唯一的分解为 $X_n = M_n + A_n$, 其中 $\{M_n, n \geq 0\}$ 为一 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 鞅, $\{A_n, n \geq 0\}$ 为一 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 可料增过程, 且 $A_0 = 0$, 称为 X 的 Doob-分解.



4.2 停时和鞅的停时定理

$\mathbb{N} :=$ 非负整数全体, $\bar{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$, 本节假定所有 r.v. 均定义在同一个带流的概率空间 $(\Omega, \mathcal{F}, (\mathcal{F}_n), P)$ 上.

定义 4.2.1

在 \mathbb{N} 中取值的随机变量 T 叫做 \mathcal{F}_n -停时 (或可选时), 若对任意的 $n \geq 0$ 有 $\{\omega | T(\omega) = n\} \in \mathcal{F}_n$ 或等价的 $\{\omega | T(\omega) \leq n\} \in \mathcal{F}_n$.

**引理 4.2.1**

设 T 是取值于 $\{0, 1, 2, \dots, \infty\}$ 的随机变量, 则下述三者等价

1. 对任意的 n , $\{T = n\} \in \mathcal{F}_n$;
2. 对任意的 n , $\{T \leq n\} \in \mathcal{F}_n$;
3. 对任意的 n , $\{T > n\} \in \mathcal{F}_n$.



证明 只要注意到如下等式, 即可证明 (1),(2),(3) 的等价性。

$$\begin{aligned} \{T \leq n\} &= \bigcup_{k=0}^n \{T = k\} \\ \{T > n\} &= \Omega - \{T \leq n\} \\ \{T = n\} &= \{T \leq n\} - \{T \leq n-1\} \end{aligned}$$

由定义我们知道事件 $\{T = n\}$ 或 $\{T \neq n\}$ 都应该由 n 时刻及其之前的信息完全确定, 而不需要也无法借助将来的情况。仍然回到公平博弈的例子, 赌博者决定何时停止赌博只能以他已经赌过的结果为依据, 而不能说, 如果我下次要输我现在就停止赌博, 这是对停止时刻 T 的第一个要求: 它必须是一个停时。

显然 $T = n$ 是停时。

对任意的 k

$$\{T = k\} = \begin{cases} \Omega & \text{若 } n = k \\ \emptyset & \text{若 } n \neq k \end{cases}$$

则 T 为停时。

例题 4.7 首次时 $\{X_n, n \geq 0\}$ 是一个随机变量序列, A 是一个事件集, 令

$$T(A) = \inf\{n, X_n \in A\}, \text{ 并约定 } T(\emptyset) = \inf\{n, X_n \in \emptyset\} = \infty,$$

可见 $T(A)$ 是 $\{X_n, n \geq 0\}$ 首次进入 A (即发生了 A 中所含事件) 的时刻, 称 $T(A)$ 是 $\{X_n, n \geq 0\}$ 到集合 A 的首达时, 可以证明 $T(A)$ 是关于 $\{X_n, n \geq 0\}$ 的停时。事实上

$$\{T(A) = n\} = \{X_0 \notin A, X_1 \notin A, \dots, X_{n-1} \notin A, X_n \in A\}$$

显然 $\{T(A) = n\}$ 完全由 X_0, X_1, \dots, X_n 决定, 从而 $T(A)$ 是关于 $\{\mathcal{F}_n\}$ 的停时, 其中 $\mathcal{F}_n := \sigma(X_0, X_1, \dots, X_n)$ 。

例题 4.8 设 $\{X_n, n \geq 0\}$ 为关于 $\{\mathcal{F}_n\}$ 适应的随机变量序列, $B \in \mathbb{B}(\mathbb{R})$, S 为停时, 令 $A := \{m : m \geq S(\omega), X_m(\omega) \in B\}$

$$T(\omega) = \begin{cases} \inf\{m : m \geq S(\omega), X_m(\omega) \in B\} & \text{若 } A \neq \emptyset \\ \infty & \text{若 } A = \emptyset \end{cases}$$

则 T 为 $\{\mathcal{F}_n\}$ 停时。

证明 当 $S = k < n = T$ 时, 即 $\bigcap_{k \leq m < n} \{X_m \in B^c\} \cap \{X_n \in B\} \in \mathcal{F}_n$ 。故把 S 的所有可能性都考虑,

$$\{T = n\} = \bigcup_{k=0}^n \left[\{S = k\} \bigcap_{k \leq m < n} \{X_m \in B^c\} \cap \{X_n \in B\} \right] \subseteq \mathcal{F}_n$$

设 T 为 $\{\mathcal{F}_n\}$ -停时, 令 $\mathcal{F}_T := \{A \in \mathcal{F}_\infty | \forall n \geq 0, A \cap \{T \leq n\} \in \mathcal{F}_n\}$ 称为 T 前事件 σ -域。

显然 $\mathcal{F}_T := \{A \in \mathcal{F}_\infty | \forall n \geq 0, A \cap \{T = n\} \in \mathcal{F}_n\}$ 。

\mathcal{F}_T 为 σ -代数。且当 $T = t$ 时, $\mathcal{F}_T = \mathcal{F}_t$ 。

定理 4.2.1

设 S, T 为 $\{\mathcal{F}_n\}$ 停时, (S_k) 为 $\{\mathcal{F}_n\}$ 停时列。

1. $\bigwedge_k S_k, \bigvee_k S_k$ 为 $\{\mathcal{F}_n\}$ 停时。
2. $S \leq T$, 则有 $\mathcal{F}_S \subset \mathcal{F}_T$ 。

其中 $\bigwedge_k S_k := \min_k \{S_k\}, \bigvee_k S_k := \max_k \{S_k\}$ 。



定理 4.2.2

设 S, T 为 $\{\mathcal{F}_n\}$ 停时, (S_k) 为 $\{\mathcal{F}_n\}$ 停时列。

1. $\bigwedge_k S_k, \bigvee_k S_k$ 为 $\{\mathcal{F}_n\}$ 停时。
2. $S \leq T$, 则有 $\mathcal{F}_S \subset \mathcal{F}_T$ 。



证明

1. 对任意的 n , $\{\bigwedge_k S_k > n\} = \bigcap_k \{S_k > n\} \in \mathcal{F}_n$, $\{\bigvee_k S_k \leq n\} = \bigcap_k \{S_k \leq n\} \in \mathcal{F}_n$ 。
2. 对任意的 $A \in \mathcal{F}_S$, 则 $A \in \mathcal{F}_\infty$,

$$A \cap \{T \leq n\} = A \cap \{S \leq n\} \cap \{T \leq n\} \in \mathcal{F}_n$$

故 $A \in \mathcal{F}_T$ 。

定理 4.2.3

设 $\{X_n, n \geq 0\}$ 为一 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 鞅 (上鞅, 下鞅), S, T 为两个 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 有界停时, 且 $S \leq T$, 则 $E[X_T | \mathcal{F}_S] = X_S (\leq X_S, \geq X_S)$ a.e.



定理 4.2.4

设 $\{X_n, n \geq 0\}$ 为一 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 鞅, T 为 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 停时, 且满足

1. $P(T < \infty) = 1$;
2. $E[|X_T|] < \infty$;
3. $\lim_{n \rightarrow \infty} E[|X_n I_{\{T > n\}}|] = 0$.

则 $E[X_T] = EX_0$.



例题 4.9 质点在 $\{0, 1, 2, \dots, N\}$ 上作带吸收壁的随机游动: 当质点移动到状态 $0, N$ 时就永远停留在该位置, 即 $p_{00} = 1, p_{NN} = 1$. 质点到达某个状态后, 下次向右移动一步的概率是 $\frac{1}{2}$, 向左移动一步的概率是 $\frac{1}{2}$. $X_0 = a$ 表示初始状态, X_n 表示质点在时间 n 的状态. 假设初始状态与每次移动相互独立. 则 $\{X_n\}$ 是马氏链, 且是一个 $\{\mathcal{F}_n\}$ 鞅. 令 $T = \min\{j : X_j = 0 \text{ 或 } X_j = N\}$, 则 T 是一个 $\{\mathcal{F}_n\}$ 停时, 其中 $\mathcal{F}_n := \sigma(X_0, X_1, \dots, X_n)$. 求 $P(X_T = N | X_0 = a)$.

证明

- 对任意的 n , 显然 X_n 关于 \mathcal{F}_n 可测.
- $E[|X_n|] \leq N < \infty$.

$$E[X_{n+1} | \mathcal{F}_n] = E[X_{n+1} | X_n] = \begin{cases} \frac{1}{2}(X_n + 1) + \frac{1}{2}(X_n - 1) = X_n & \text{if } X_n \neq 0, N \\ X_n & \text{if } X_n = 0, N \end{cases}$$

由于 $A := \{1, 2, \dots, N-1\}$ 为非常返等价类, $C := \{0, N\}$ 为常返状态全体. 由马氏链结论可知 A 中的状态一定走出 A 进入 C 中, 故 $P(T < \infty) = 1$. $E[|X_T|] \leq N$. 由于 $T < \infty$ a.e. 故 $I_{\{T > n\}} \rightarrow 0$ (as $n \rightarrow \infty$). 且 $|X_n I_{\{T > n\}}| \leq N$, 由控制收敛定理可知 $\lim_{n \rightarrow \infty} E[|X_n I_{\{T > n\}}|] = 0$.

由停止定理可知 $E(X_T) = E(X_0) = a$. 此时 X_T 只取 $N, 0$ 两个值, 有 $E(X_T) = N \cdot P(X_T = N | X_0 = a) + 0 \cdot P(X_T = 0 | X_0 = a)$, 故 $P(X_T = N | X_0 = a) = \frac{E(X_T)}{N} = \frac{a}{N}$.

定理 4.2.5

设 $\{X_n, n \geq 0\}$ 为一 $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ 鞅 (下鞅), $m \geq 0$. 则对任意的 $\lambda > 0$ 有

1. $\lambda P(\sup_{0 \leq n \leq m} X_n \geq \lambda) \leq E[X_m I_{\{\max_{0 \leq n \leq m} X_n \geq \lambda\}}] \leq E[X_m^+]$
2. $\lambda P(\inf_{0 \leq n \leq m} X_n \leq -\lambda) \leq E[|X_m|] + E[|X_0|]$
3. 若 $p \geq 1$ 且 X 为 L^p 鞅, 则有

$$P(\max_{0 \leq n \leq m} |X_n| \geq \lambda) \leq \lambda^{-p} E[|X_m|^p].$$

4. 若 $p > 1$, q 为其共轭指数 ($\frac{1}{p} + \frac{1}{q} = 1$), 则有

$$E(\max_{0 \leq n \leq m} |X_n|^p) \leq q^p E[|X_m|^p]$$



4.3 鞅收敛定理

$\mathbb{T} = \mathbb{R}_+ := [0, \infty)$. 本节假定所有 r.v. 均定义在同一个带流的概率空间 (Ω, \mathcal{F}, P) 上.

定义 4.3.1

称 \mathcal{F} 的子 σ -域序列 $\{\mathcal{F}_t, t \in \mathbb{R}_+\}$ 为一个流, 若 $\forall 0 \leq s < t$ 有 $\mathcal{F}_s \subseteq \mathcal{F}_t$. 记 $\mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t$,

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s, t \geq 0, \mathcal{F}_{t-} = \bigvee_{s < t} \mathcal{F}_s = \sigma(\bigcup_{s < t} \mathcal{F}_s), t \geq 0.$$

一个流 $\{\mathcal{F}_t, t \in \mathbb{R}_+\}$ 称为右连续的, 若 $\forall t \geq 0$ 有 $\mathcal{F}_t = \mathcal{F}_{t+}$.



定义 4.3.2

设 $X = (X(t, \cdot), t \in \mathbb{R}_+)$ 为随机过程, 对每个固定的 t , $X(t, \cdot)$ 为一个随机变量; 而对每个固定的 $\omega \in \Omega$, $X(\cdot, \omega)$ 是关于 $t \in T$ 的映射, 称 $X(\cdot, \omega)$ 为样本轨道 (或样本函数, 一个实现)。若 X 的全部轨道是连续的 (右连续的, 左连续的), 则称 X 为连续 (右连续, 左连续) 过程。如果 X 的全部轨道是右连续并有左极限的。 (简称右连左极)。

**定理 4.3.1**

设 $(X_t, t \geq 0)$ 为一 $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ 右连续下鞅。则对任意的 $\lambda > 0$, $s_0 \leq t_0 \in T$ 有

1. $\lambda P(\sup_{s_0 \leq t \leq t_0} X_t \geq \lambda) \leq E[X_{t_0} I_{\{\sup_{s_0 \leq t \leq t_0} X_t \geq \lambda\}}] \leq E[X_{t_0}^+]$
2. $\lambda P(\sup_{s_0 \leq t \leq t_0} X_t > \lambda) \leq E[X_{t_0} I_{\{\sup_{s_0 \leq t \leq t_0} X_t > \lambda\}}] \leq E[X_{t_0}^+]$
3. $\lambda P(\inf_{s_0 \leq t \leq t_0} X_t \leq -\lambda) \leq E[|X_{t_0}|] + E[|X_{s_0}|]$
4. $\lambda P(\inf_{s_0 \leq t \leq t_0} X_t < -\lambda) \leq E[|X_{t_0}|] + E[|X_{s_0}|]$
5. 若 $p \geq 1$ 且 X 为 L^p 鞅, 则有

$$P(\sup_{0 \leq t \leq u} |X_t| \geq \lambda) \leq \lambda^{-p} E[|X_u|^p].$$

6. 若 $p > 1$, q 为其共轭指数, 则有

$$E(\sup_{0 \leq t \leq u} |X_t|^p) \leq q^p E[|X_u|^p].$$

**定义 4.3.3**

两个定义在同一个概率空间上具有同一个状态空间的随机过程 X, X' 称为互为修正, 如果对任意的 $t \in T$, $X_t = X'_t$ a.s. 即对任意的 $t \in T$, $P(X_t = X'_t) = 1$ 。

**定义 4.3.4**

设 $X := (X(t, \omega), t \geq 0), Y := (Y(t, \omega), t \geq 0)$ 为两个随机过程, 称 X, Y 无区别, 若对几乎所有的 ω , 轨道 $X(\cdot, \omega), Y(\cdot, \omega)$ 一致。即 $P(X(t, \omega) = Y(t, \omega), \forall t \in T) = 1$ 。

**定理 4.3.2**

设 $(X_t, t \geq 0)$ 为一 $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ 下鞅, 则 X 存在右连左极修正的充分必要条件是: $E[X_t]$ 作为 t 的函数在 \mathbb{R}_+ 上右连续。特别的, 一切鞅存在右连左极的修正。



第5章 布朗运动

5.1 布朗运动定义

我们从讨论简单的随机游动开始。设有一个粒子在直线上随机游动，在每个单位时间内等可能地向左或向右移动一个单位的长度。现在加速这个过程，在越来越小的时间间隔中走越来越小的步子。若能以正确的方式趋于极限，我们就得到 Brown 运动。详细地说就是令此过程每隔 Δt 时间等概率地向左或向右移动 Δx 的距离，如果以 $X(t)$ 记时刻 t 粒子的位置，则

$$X(t) = \Delta x(X_1 + \cdots + X_{[t/\Delta t]}) \quad (5.1)$$

其中 $[t/\Delta t]$ 表示 $t/\Delta t$ 的整数部分，其中

$$X_i = \begin{cases} +1, & \text{如果第 } i \text{ 步向右} \\ -1, & \text{如果第 } i \text{ 步向左} \end{cases}$$

且假设诸 X_i 相互独立

$$P\{X_i = 1\} = P\{X_i = -1\} = \frac{1}{2}$$

由于 $E[X_i] = 0$, $\text{Var}[X_i] = E[X_i^2] = 1$ 及??, 我们有 $E[X(t)] = 0$, $\text{Var}[X(t)] = (\Delta x)^2[t/\Delta t]$ 。现在要令 Δx 和 Δt 趋于零，并使得极限有意义。如果取 $\Delta x = \Delta t$ ，令 $\Delta t \rightarrow 0$ ，则 $\text{Var}[X(t)] \rightarrow 0$ ，从而 $X(t) = 0$, a.s. 如果取 $\Delta t = (\Delta x)^3$ ，则 $\text{Var}[X(t)] \rightarrow \infty$ ，这是不合理的。因为粒子的运动是连续的，不可能在很短时间内远离出发点。因此，我们作下面的假设： $\Delta x = \sigma\sqrt{\Delta t}$ ， σ 为某个正常数，从上面的讨论可见，当 $\Delta t \rightarrow 0$ 时， $E[X(t)] = 0$ ， $\text{Var}[X(t)] \rightarrow \sigma^2 t$ 。

下面来看这一极限过程的一些直观性质。由式 (1.1) 及中心极限定理可得：

1. $X(t)$ 服从均值为 0，方差为 $\sigma^2 t$ 的正态分布。此外，由于随机游动的值在不相重叠的时间区间中的变化是独立的，所以有
2. $\{X(t), t \geq 0\}$ 有独立增量。又因为随机游动在任一时间区间中的位置变化的分布只依赖于区间的长度，可见
3. $\{X(t), t \geq 0\}$ 有平稳增量。

本节假设 $\mathcal{T} := [0, \infty)$,

定义 5.1.1

如果对任何 $t_1, t_2, \dots, t_n \in \mathcal{T}, t_1 < t_2 < \dots < t_n$ 随机变量 $X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$ 是相互独立的，则称 $X(t)$ 为独立增量过程。如果对任何 t_1, t_2 ，有 $X(t_1 + h) - X(t_1) \stackrel{d}{=} X(t_2 + h) - X(t_2)$ ，则称 $\{X(t), t \in \mathcal{T}\}$ 为是平稳增量过程。兼有独立增量和平稳增量的过程称为平稳独立增量的过程。

本节假定所有 r.v. 均定义在同一个概率空间 (Ω, \mathcal{F}, P) 上。

定义 5.1.2

若一个随机过程 $\{X(t), t \geq 0\}$ 满足

1. $X(t)$ 是独立增量过程；
2. $\forall s, t > 0, X(t+s) - X(s) \sim N(0, \sigma^2 t)$ ；其中 σ 为给定的常数。
3. $X(t)$ 关于 t 是连续函数。则称其为布朗运动。

注 若 $\sigma = 1$ ， $\{X(t), t \geq 0\}$ 为标准布朗运动。即 $X(0) = 0$, $X(t+s) - X(s) \sim N(0, t)$ 。

本章讨论标准布朗运动，即 $\{B(t), t \geq 0\}$ 。则在 t 时刻 $B(t) = B(t) - B(0)$ 的概率密度函数为

$$f(x, t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}.$$

例题 5.1 设 $\{B(t), t \geq 0\}$ 为标准布朗运动, $B(0) = 0$ 。求

1. $P(B(2) \leq 0)$
2. $P(B(t) \leq 0; t = 0, 1, 2)$

证明 (1) $P(B(2) \leq 0) = \frac{1}{2}$.
(2).

$$\begin{aligned}
 P(B(t) \leq 0; t = 0, 1, 2) &= P\{B(1) \leq 0, B(2) \leq 0\} \\
 &= P\{B(1) \leq 0, B(1) + (B(2) - B(1)) \leq 0\} \\
 &= P\{B(1) \leq 0, B(2) - B(1) \leq -B(1)\} \\
 &= E[\mathbf{1}_{(-\infty, 0]}(B(1))\mathbf{1}_{(-\infty, -B(1))}(B(2) - B(1))] \\
 &= E[E[\mathbf{1}_{(-\infty, 0]}(B(1))\mathbf{1}_{(-\infty, -B(1))}(B(2) - B(1)) | B(1)]] \\
 &= E[\mathbf{1}_{(-\infty, 0]}(B(1))E[\mathbf{1}_{(-\infty, -B(1))}(B(2) - B(1)) | B(1)]] \\
 &= E\left[\mathbf{1}_{(-\infty, 0]}(B(1))P(B(2) - B(1) \leq -x) \Big|_{x=B(1)}\right] \\
 &= \int_{-\infty}^{+\infty} \mathbf{1}_{(-\infty, 0)}(x)P\{B(2) - B(1) \leq -x\}f(x)dx
 \end{aligned}$$

其中 $f(x)$ 为标准正态的密度函数, $\Phi(x)$ 为标准正态的分布函数。

$$\begin{aligned}
 P\{B(1) \leq 0, B(2) \leq 0\} &= P\{B(1) \leq 0, B(1) + (B(2) - B(1)) \leq 0\} \\
 &= P\{B(1) \leq 0, B(2) - B(1) \leq -B(1)\} \\
 &= \int_{-\infty}^0 P\{B(2) - B(1) \leq -x\}f(x)dx \\
 &= \int_{-\infty}^0 \Phi(-x)f(x)dx \\
 &= \int_0^{\infty} \Phi(x)f(-x)dx = \int_0^{\infty} \Phi(x)f(x)dx \\
 &= \int_0^{\infty} \Phi(x)d\Phi(x) = \int_{\frac{1}{2}}^1 ydy = \frac{3}{8}.
 \end{aligned}$$

其中 $f(x)$ 为标准正态的密度函数。

定理 5.1.1

设 (Ω, \mathcal{F}, P) 为概率空间, \mathcal{G} 为 \mathcal{F} 的一个子 σ 代数; (S, \mathcal{S}) 和 (E, \mathcal{E}) 为可测空间, X 为一 \mathcal{G} 可测 S 值随机元, Y 为一 E 值随机元。假定 Y 和 \mathcal{G} 独立。令 $g(x, y)$ 为 $S \times E$ 上的 $S \times \mathcal{E}$ 可测函数, 使得 $E[|g(X, Y)|] < \infty$, 则 $E[g(X, Y) | \mathcal{G}] = E[g(x, Y) |_{x=X}]$ 。



设 $X = (X_1, X_2, \dots, X_n)$ 有联合密度 $f_X(x_1, x_2, \dots, x_n)$, n 元可测函数 $y = (y_1, \dots, y_n)$, 其中 $y_i(x_1, \dots, x_n), i = 1, \dots, n$ 满足:

- 除 \mathbb{R}^n 中的 L -零测集 N 外, 存在至多可数个两两不交的可求积区域 $\{D_{x^k}\}_{k \geq 1}$, 使 $\mathbb{R}^n - N = \sum_{k=1}^{\infty} D_{x^k}$,
- 对一切 k , 函数 y 把 D_{x^k} 中的点一对一地映射到可求积区域 D_{y^k} 上,
- 在每一个 D_{x^k} 上, y 的逆映射为 $x^{(k)} = (x_1^{(k)}, \dots, x_n^{(k)})$, 其中 $x_i^{(k)} = x_i^{(k)}(y_1, \dots, y_n), i = 1, \dots, n$ 是 D_{y^k} 上的连续函数, 且有连续的偏导数, 函数行列式 $J_k(y_1, \dots, y_n) = \frac{D(x_1, \dots, x_n)}{D(y_1, \dots, y_n)} \neq 0$ 。

若 $\eta = (\eta_1, \dots, \eta_n)$, 其中 $\eta_i = y_i(X_1, \dots, X_n), i = 1, \dots, n$, 则 η 是连续型随机变量, 其密度函数为

$$f_{\eta}(y_1, \dots, y_n) = \sum_{k=1}^{\infty} |D_{y^k}(y_1, \dots, y_n) f_X(x_1^{(k)}(y_1, \dots, y_n), \dots, x_n^{(k)}(y_1, \dots, y_n))| |J_k|$$

定理 5.1.2

设 $B(0) = 0$, $\{B(t), t \geq 0\}$ 为标准布朗运动。令 $x_0 = 0, t_0 = 0$, 则当 $B(0) = 0$ 时, $\forall 0 < t_1 < t_2 < \dots < t_n$, $(B(t_1), B(t_2), \dots, B(t_n))$ 的联合密度函数为

$$\begin{aligned} f_{B(t_1), \dots, B(t_n)}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp \left\{ -\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})} \right\} \\ &= \prod_{i=1}^n p(x_i - x_{i-1}, t_i - t_{i-1}) \end{aligned}$$

即对任意的 $F_k \in \mathcal{B}(\mathbb{R}), k = 1, 2, \dots, n$,

$$P^X(B(t_1) \in F_1, B(t_2) \in F_2, \dots, B(t_n) \in F_n) = \int_{F_1} p(t_1, x_1) \int_{F_2} p(t_2 - t_1, x_2 - x_1) \cdots \int_{F_n} p(t_n - t_{n-1}, x_n - x_{n-1}) dx_1 dx_2 \dots dx_n$$

证明 令 $Y_1 = B(t_1), Y_i = B(t_i) - B(t_{i-1})$, 则有 $B(t_i) = \sum_{k=1}^i Y_k$ 。令 $y_1 = x_1, y_i = x_i - x_{i-1}$, 则 $x_i = \sum_{k=1}^i y_k$ 。故 Y_i 独立同分布, 服从 $N(0, t_i - t_{i-1})$ 。

$$f_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp \left(-\frac{y_i^2}{2(t_i - t_{i-1})} \right)$$

故对应的雅可比行列式为

$$J = \left| \frac{\partial y}{\partial x} \right| = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{vmatrix}$$

$$f_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp \left(-\frac{y_i^2}{2(t_i - t_{i-1})} \right)$$

故对应的雅可比行列式为

$$J = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

$$f(x_1, x_2, \dots, x_n) = \frac{1}{\sqrt{2\pi t_1}} \exp \left(-\frac{x_1^2}{2t_1} \right) \cdot \prod_{i=2}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp \left(-\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})} \right)$$

设 $B(0) = 0$, $\{B(t), t \geq 0\}$ 为标准布朗运动。则当 $B(0) = 0$ 时, $\forall 0 < t_1 < t_2$, $(B(t_1), B(t_2))$ 的联合密度函数为

$$f_{B(t_1), B(t_2)}(x_1, x_2) = \frac{1}{\sqrt{2\pi t_1}} \exp \left(-\frac{x_1^2}{2t_1} \right) \cdot \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp \left(-\frac{(x_2 - x_1)^2}{2(t_2 - t_1)} \right)$$

注

在 $B(t_1) = x_1$ 条件下, $B(t_2)$ 条件概率密度为

$$f(x_2 | x_1, t_2 - t_1) = \frac{1}{\sqrt{2\pi(t_2 - t_1)}} \exp \left\{ -\frac{(x_2 - x_1)^2}{2(t_2 - t_1)} \right\}$$

故在 $B(t_1) = x_1$ 条件下 $B(t_1 + t)$ 条件概率密度为

$$f(x | x_1, t) = \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{(x - x_1)^2}{2t} \right\}$$

$$P(B(t + t_0) > x_0 | B(t_0) = x_0) = P(B(t + t_0) \leq x_0 | B(t_0) = x_0) = \frac{1}{2}.$$

即在给定初始条件 $B(t_0) = x_0$ 下, $\forall t > 0$, $t_0 + t$ 时刻高或低于初始时刻概率相等均为 $1/2$ 。

1. 若 $X \sim N(\mu, \sigma^2)$, 则 $Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$, 即 $X = \mu + \sigma Z$ 。
2. 若 $X = (X_1, X_2, \dots, X_n)' \sim N(\bar{\mu}, \Sigma)$, 其中 $\bar{\mu} = (\mu_1, \dots, \mu_n)'$, Σ 为 $n \times n$ 正定阵。则存在 $n \times n$ 矩阵 A 使得 $\Sigma = AA'$ 且 $|A| \neq 0$, $X = \bar{\mu} + AZ$, 其中 $Z = (Z_1, Z_2, \dots, Z_n)'$ 为 n 维独立标准正态分布。记 $\bar{x} = (x_1, x_2, \dots, x_n)'$, 则 X 的概率密度函数为

$$f_X(\bar{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (\bar{x} - \bar{\mu})' \Sigma^{-1} (\bar{x} - \bar{\mu}) \right\}$$

例题 5.2 取 $\mu = (0, 0)'$,

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

$X = (X_1, X_2) \sim N(\mu, \Sigma)$, 给出其密度函数。

证明 令 $\mathbf{X} = A\mathbf{Z}$, 其中 $\mathbf{Z} = (Z_1, Z_2)$, Z_1, Z_2 相互独立服从标准正态分布

$$A = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix}$$

满足 $\Sigma = AA^T$ 。

$$\begin{cases} x_1 = z_1 \\ x_2 = \rho z_1 + \sqrt{1 - \rho^2} z_2 \end{cases}$$

则可得逆变换

$$\begin{cases} z_1 = x_1 \\ z_2 = \frac{x_2 - \rho x_1}{\sqrt{1 - \rho^2}} \end{cases}$$

$$\left| \frac{\partial z_1}{\partial x} \frac{\partial z_2}{\partial y} \right| = \left| \begin{vmatrix} 1 & 0 \\ 0 & \frac{1}{\sqrt{1 - \rho^2}} \end{vmatrix} \right|$$

\mathbf{Z} 的概率密度函数为 $f_{\mathbf{Z}}(z_1, z_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(z_1^2 + z_2^2)}$,

则 \mathbf{X} 的密度函数为 $f_{\mathbf{X}}(z_1, z_2) = \frac{1}{2\pi\sqrt{1 - \rho^2}} e^{-\frac{1}{2(1 - \rho^2)}(x_1^2 - 2\rho x_1 x_2 + x_2^2)}$ 。

若随机向量 (X_1, \dots, X_n) 的分布函数为 $F(x_1, \dots, x_n)$, 与随机变量相仿, 类似地定义它的特性函数

$$\psi(t_1, \dots, t_n) = E \left[e^{i \sum_{k=1}^n t_k X_k} \right] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{i(t_1 x_1 + \dots + t_n x_n)} dF(x_1, \dots, x_n)$$

若 \mathbf{X} 服从多元正态分布 $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 则特征函数为

$$\psi_{\mathbf{X}}(\mathbf{t}) = \exp \left\{ i \mathbf{t}' \boldsymbol{\mu} - \frac{1}{2} \mathbf{t}' \boldsymbol{\Sigma} \mathbf{t} \right\}$$

定理 5.1.3

若 $\{X(t_i), i = 1, \dots, n\}$ 为多元正态分布当且仅当对任意的实数 $\alpha_k (k = 1, 2, \dots, n)$, 随机变量 $Y := \sum_{k=1}^n \alpha_k X(t_k)$ 为一维的正态分布。



证明 充分性: 设 $\{X(t_i), i = 1, \dots, n\}$ 的任意线性组合为一维正态分布。则 $\{X(t_i), i = 1, \dots, n\}$ 的联合特征函

数为 $\psi(\alpha_1, \dots, \alpha_n)$

$$\begin{aligned}\psi(\alpha_1, \dots, \alpha_n) &= E \left[e^{i \sum_{k=1}^n \alpha_k X(t_k)} \right] = E \left[e^{iY} \right] \\ &= e^{iE \left[\sum_{k=1}^n \alpha_k X(t_k) \right] - \frac{1}{2} E \left[\sum_{k=1}^n \alpha_k (X(t_k) - E(X(t_k))) (X(t_j) - E(X(t_j))) \right]} \\ &= e^{i \sum_{k=1}^n \alpha_k E[X(t_k)] - \frac{1}{2} \sum_{k,j} \alpha_k \alpha_j E[(X(t_k) - E(X(t_k)))(X(t_j) - E(X(t_j)))]}\end{aligned}$$

故 $\psi(\alpha_1, \dots, \alpha_n)$ 为多元正态分布的特征函数, 故 $\{X(t_i), i = 1, \dots, n\}$ 为多元正态分布。

必要性: 设 $\{X(t_i), i = 1, \dots, n\}$ 为多元正态分布则其联合特征函数为

$$\begin{aligned}\psi(\lambda_1, \dots, \lambda_n) &= E \left[e^{i \sum_{k=1}^n \lambda_k X(t_k)} \right] \\ &= e^{i \sum_{k=1}^n \lambda_k \mu_k - \frac{1}{2} \sum_{k,j} \lambda_k \lambda_j \sigma_{kj}}\end{aligned}$$

其中 $\mu_k = E[X(t_k)]$, $\sigma_{kj} = E[(X(t_k) - E(X(t_k)))(X(t_j) - E(X(t_j)))]$ 。取 $\lambda_k = \alpha_k$, 则 $Y := \sum_k \alpha_k X(t_k)$ 有特征函数

$$\begin{aligned}\psi(\lambda) &= E \left[e^{i\lambda Y} \right] \\ &= E \left[e^{i \sum_{k=1}^n \lambda \alpha_k X(t_k)} \right] \\ &= e^{i \sum_{k=1}^n \lambda \alpha_k \mu_k - \frac{1}{2} \lambda^2 \sum_{k,j} \alpha_k \alpha_j \sigma_{kj}} \\ &= e^{i\lambda EY - \frac{1}{2} \lambda^2 E[Y - EY]^2}\end{aligned}$$

故其特征函数为一元正态分布的分布函数, 故 Y 为一元正态分布。

定义 5.1.3

若 $\{X(t), t \in \mathcal{T}\}$, $\forall t \in \mathcal{T} (i = 1, 2, \dots, n) (X(t_1), X(t_2), \dots, X(t_n))$ 联合分布为 n 维正态分布, 则称 $\{X(t), t \in \mathcal{T}\}$ 为正态过程。



定理 5.1.4

若 $\{X(t), t \in \mathcal{T}\}$ 为正态过程当且仅当任意有限个随机变量的线性组合都是一维的正态分布, 即对任意的 $t_1, t_2, \dots, t_n \in \mathcal{T} (n \geq 1)$, 实数 α_k , 随机变量 $\sum_{k=1}^n \alpha_k X(t_k)$ 为一维的正态分布。



定理 5.1.5

若 $\{X(t), t \in \mathcal{T}\}$ 为正态过程, 则

1. $\{X(t), t \in \mathcal{T}\}$ 中的随机变量相互独立 (即任意有限个相互独立), 当且仅当 $\forall s \neq t (s, t \in \mathcal{T})$,

$$\sigma(s, t) := E[(X(t) - E[X(t)])(X(s) - E[X(s)])] = 0$$

2. $\{X(t), t \in \mathcal{T}\}$ 中的任意一个 $X(t_0)$ 与 $\{X(t), t \neq t_0\}$ 相互独立, 当且仅当 $\forall t \neq t_0 (t \in \mathcal{T})$,

$$\sigma(t_0, t) = 0$$



定理 5.1.6

设 $B(0) = 0$, $\{B(t), t \geq 0\}$ 为标准布朗运动。 $(B(t_1), B(t_2), \dots, B(t_n))$ 其中 $t_1 < t_2 < \dots < t_n$ 是均值向量为 0, 协方差矩阵满足 $Cov(B(s), B(t)) = \min\{t, s\}$ 的正态分布。



证明 任取实数 α_k ,

$$\begin{aligned}\sum_k \alpha_k B(t_k) &= \alpha_n B(t_n) + \alpha_{n-1} B(t_{n-1}) + \dots + \alpha_1 B(t_1) \\ &= \alpha_n (B(t_n) - B(t_{n-1})) + (\alpha_n + \alpha_{n-1}) (B(t_{n-1}) - B(t_{n-2})) \\ &\quad + \dots + (\alpha_n + \alpha_{n-1} + \dots + \alpha_2) (B(t_2) - B(t_1)) \\ &\quad + (\alpha_n + \alpha_{n-1} + \dots + \alpha_1) B(t_1),\end{aligned}$$

由于 $B(t_i) - B(t_{i-1})$ 与 $B(t_j) - B(t_{j-1}) (i \neq j)$ 相互独立, 可知 $\sum_k \alpha_k B(t_k)$ 为一维正态分布。

当 $s < t$ 时

$$\begin{aligned} \text{Cov}(B(s), B(t)) &= E[B(s)B(t)] - E[B(s)]E[B(t)] \\ &= E[B(s)(B(t) - B(s) + B(s))] \\ &= E[B^2(s)] + E[(B(t) - B(s))B(s)] \\ &= s \end{aligned}$$

例题 5.3 设 $B(0) = 0$, $\{B(t), t \geq 0\}$ 为标准布朗运动。求 $B(1) + B(2) + B(3) + B(4)$ 的分布。

证明

$$X = (B(1), B(2), B(3), B(4))'$$

X 是多元正态分布, 均值向量为 $(0, 0, 0, 0)$ 和协方差矩阵,

$$E(XX') = \sum = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix}$$

$A = (1, 1, 1, 1)$, 则 $AX = B(1) + B(2) + B(3) + B(4)$, 均值向量为 0 , 方差为 $E(AXX'A') = AE(XX')A' = 30$ 。

故 $B(1) + B(2) + B(3) + B(4)$ 均值为 0 , 方差为 30 。故 $B(1) + B(2) + B(3) + B(4) \sim N(0, 30)$ 。

定理 5.1.7

设 $\{B(t), t \geq 0\}$ 为正态过程, 轨道连续, $B(0) = 0$, $\forall s, t > 0$ 有 $E[B(t)] = 0$, $E[B(t)B(s)] = t \wedge s$ 则当且仅当 $\{B(t), t \geq 0\}$ 为标准布朗运动。

证明 充分性: 由定理 ??, ?? 知, $\{B(t), t \geq 0\}$ 为正态过程, 故 $E[B(t)] = 0$,

$$\begin{aligned} E[B(s)B(t)] &= E[B(s)(B(t) - B(s) + B(s))] \\ &= E[B^2(s)] + E[(B(t) - B(s))B(s)] \\ &= s \end{aligned}$$

必要性: 设 $\{B(t), t \geq 0\}$ 为正态过程, 设 $s < t$, $E[B(t) - B(s)] = 0$ 。

$$\begin{aligned} E[B(t) - B(s)]^2 &= E[B^2(t)] + E[B^2(s)] - 2E[B(t)B(s)] \\ &= t + s - 2(t \wedge s) = t - s \end{aligned}$$

由于 $B(t) - B(s)$ 为正态分布, 故 $B(t) - B(s) \sim N(0, t - s)$ 。对任意的 $s_1 < t_1 \leq s_2 < t_2$,

$$\begin{aligned} \text{Cov}(B(t_1) - B(s_1), B(t_2) - B(s_2)) &= E[(B(t_1) - B(s_1))(B(t_2) - B(s_2))] \\ &= E[B(t_1)B(t_2)] - E[B(t_1)B(s_2)] - E[B(s_1)B(t_2)] + E[B(s_1)B(s_2)] \\ &= t_1 - t_1 - s_1 + s_1 = 0 \end{aligned}$$

故 $B(t_1) - B(s_1)$ 与 $B(t_2) - B(s_2)$ 相互独立。

定理 5.1.8

设 $\{B(t), t \geq 0\}$ 为标准布朗运动。

1. 对任意的 $\tau > 0$, $\{B(t + \tau) - B(\tau), t \geq 0\}$
2. 对任意的 $\lambda > 0$, $\left\{\frac{1}{\sqrt{\lambda}}B(\lambda t), t \geq 0\right\}$
3. $\{tB\left(\frac{1}{t}\right), t \geq 0\}$ 其中 $tB\left(\frac{1}{t}\right)|_{t=0} := 0$
4. 对任意的 $t_0 > 0$, $\{B(t_0 - s) - B(t_0), 0 \leq s \leq t_0\}$;

仍为标准布朗运动。

定理 5.1.9

对任意的 $a \geq 0$, 过程 $X = \{X(t) = B(t+a) - B(a), t \geq 0\}$ 是标准布朗运动与 σ -代数 $\mathcal{F}_a^B = \sigma(B(t), 0 \leq t \leq a)$ 独立。

**定理 5.1.10**

$\mathcal{F}_t := \sigma(B(u), u \leq t)$. 则

1. $\{B(t), t \geq 0\}$ 为 $\{\mathcal{F}_t\}$ 鞅;
2. $\{B^2(t) - t, t \geq 0\}$ 为 $\{\mathcal{F}_t\}$ 鞅;
3. $\{e^{\lambda B(t) - \frac{1}{2}\lambda^2 t}, t \geq 0\}$ 为 $\{\mathcal{F}_t\}$ 鞅;
4. $\{e^{i\lambda B(t) + \frac{1}{2}\lambda^2 t}, t \geq 0\}$ 为 $\{\mathcal{F}_t\}$ 鞅;

**定义 5.1.4**

设有概率空间 (Ω, \mathcal{F}, P) 上的以 (E, \mathcal{E}) 为状态空间的随机过程 $\xi = \{\xi_t, t \in T\}$ 及其 σ -代数流 $\{\mathcal{F}_t, t \in T\}$ 满足 $\mathcal{F}_s \subset \mathcal{F}_t$. 设 ξ 对 $\{\mathcal{F}_t, t \in T\}$ 是适应的。此时称 $(\Omega, \mathcal{F}, P, \xi, \mathcal{F}_t)$ 是一个 $\{\mathcal{F}_t, t \in T\}$ -Markov 过程, 若对任意的 $s < t \in T, B \in \mathcal{E}$ 有 $P(\xi_t \in B | \mathcal{F}_s) = P(\xi_t \in B | \xi_s)$ 。特别的当 $\mathcal{F}_t = \sigma(\xi_s, s \leq t)$, 则称 ξ 为 (Ω, \mathcal{F}, P) 上的马氏过程。

**定义 5.1.5**

设有概率空间 (Ω, \mathcal{F}, P) 上的以 (E, \mathcal{E}) 为状态空间的随机过程 $\xi = \{\xi_t, t \in T\}$ 及其 σ -代数流 $\{\mathcal{F}_t, t \in T\}$ 满足 $\mathcal{F}_s \subset \mathcal{F}_t$. 设 ξ 对 $\{\mathcal{F}_t, t \in T\}$ 是适应的。此时称 $(\Omega, \mathcal{F}, P, \xi, \mathcal{F}_t)$ 是一个 $\{\mathcal{F}_t, t \in T\}$ -Markov 过程, 若对任意的 $s < t \in T, B \in \mathcal{E}$ 有 $P(\xi_t \in B | \mathcal{F}_s) = P(\xi_t \in B | \xi_s)$ 。特别的当 $\mathcal{F}_t = \sigma(\xi_s, s \leq t)$, 则称 ξ 为 (Ω, \mathcal{F}, P) 上的马氏过程。

**定理 5.1.11**

$\{B(t), t \geq 0\}$ 为 (Ω, \mathcal{F}, P) 上的马氏过程。

**定理 5.1.12**

设 (Ω, \mathcal{F}, P) 为概率空间, \mathcal{G} 为 \mathcal{F} 的一个子 σ -代数; (S, \mathcal{S}) 和 (E, \mathcal{E}) 为可测空间, X 为一 \mathcal{G} 可测 S 值随机元, Y 为一 E 值随机元。假定 Y 和 \mathcal{G} 独立。令 $g(x, y)$ 为 $S \times E$ 上的 $\mathcal{S} \times \mathcal{E}$ 可测函数, 使得 $E[|g(X, Y)|] < \infty$, 则 $E[g(X, Y) | \mathcal{G}] = E[g(x, Y)]|_{x=X}$ 。

**定理 5.1.13**

$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ 上 R^d 适应过程若具有独立增量性, 则 $X = (X_t, t \geq 0)$ 为马氏过程。



证明 对任意的 $B \in \mathcal{B}(R^d)$, $s < t$, 令 $A := \{(x, y) | x, y \in R^d, x + y \in B\}$.

$$\begin{aligned} P(X_t \in B | \mathcal{F}_s) &= P((X_t - X_s, X_s) \in A | \mathcal{F}_s) \\ &= P((X_t - X_s, y) \in A)_{y=X_s} \end{aligned}$$

$$\begin{aligned} P(X_t \in B | X_s) &= P((X_t - X_s, X_s) \in A | X_s) \\ &= P((X_t - X_s, y) \in A)_{y=X_s} \end{aligned}$$

故可知

$$P(X_t \in B | \mathcal{F}_s) = P(X_t \in B | X_s).$$

定义 5.1.6

马氏过程 $\{X_t, t \geq 0\}$ 称为具有强马氏性, 若对任意的 $t \in T$ 以及 $A \in \mathcal{B}(\mathbb{R})$ 以及 $\{\mathcal{F}_t\}$ 停时 τ 有

$$P(\{X_{t+\tau} \in A\} \cap \{\tau < \infty\} | \mathcal{F}_\tau) = P(\{X_{t+\tau} \in A\} \cap \{\tau < \infty\} | X_\tau).$$



令 $\mathcal{F}_{t+} := \bigcap_{u>t} \mathcal{F}_u$.

$\mathcal{F}_{\tau+} := \{B \in \sigma(\bigcup_n \mathcal{F}_n) | \forall t > 0, B \cap \{\tau \leq t\} \in \mathcal{F}_{t+}\}$.

定理 5.1.14

令 $\mathcal{F}_{t+} := \bigcap_{u>t} \mathcal{F}_u$. 标准布朗运动 $\{B(t), t \geq 0\}$ 为 $\{\mathcal{F}_{t+}\}$ 的强马氏过程, 即对任意的 $\{\mathcal{F}_{t+}\}$ 停时 τ 以及有界可测函数 f 以及非负 t

$$E_x[f(B(\tau+t)) | \mathcal{F}_{\tau+}] = E_{B(\tau)}[f(B(t)) | \{\tau < \infty\}].$$



注 去任意的 $A \in \mathcal{B}(\mathbb{R})$, 把上述定理中的 $f(x) := I_A(x)$, 则有

$$P(\{X_{t+\tau} \in A\} \cap \{\tau < \infty\} | \mathcal{F}_\tau) = P(\{X_{t+\tau} \in A\} \cap \{\tau < \infty\} | X_\tau).$$

定理 5.1.15

标准布朗运动 $\{B(t)\}$ 为 $\{\mathcal{F}_t\}$ 的强马氏过程. 即对任意的 $\{\mathcal{F}_t\}$ 停时 τ 以及有界可测函数 f 以及非负 t

$$E_x[f(B(\tau+t))I_{\{\tau < \infty\}} | \mathcal{F}_\tau] = E_{B(\tau)}[f(B(t))I_{\{\tau < \infty\}}]$$

**定理 5.1.16**

设 $B = \{B(t), t \geq 0\}$ 是一个标准布朗运动, 若 $\tau < \infty$ a.s. 为 $\{\mathcal{F}_{t+}\}$ 停时, $\{B'(t) := B(t+\tau) - B(\tau) | t \geq 0\}$ 仍为一个布朗运动且与 \mathcal{F}_τ 独立.

**定理 5.1.17**

标准布朗运动 $\{B(t), t \geq 0\}$ 及 $\{\mathcal{F}_t\}$ 停时 τ , 其中 $\mathcal{F}_t := \sigma(B(u), u \leq t)$. 则 $\tilde{B}(t) := B(t \wedge \tau) - (B(t) - B(t \wedge \tau)), t \geq 0$ 与 $B(t)$ 有相同的分布.



证明 假设 $\tau < \infty$, 令 $B^\tau(t) := B(t \wedge \tau)$ 以及 $B'(t) = B(\tau+t) - B(\tau)$. 则由定理 1.30 可知, $\{B'(t), t \geq 0\}$ 为标准布朗运动, 且与 \mathcal{F}_τ 相互独立. 由于 $B^\tau := \{B^\tau(t), t \geq 0\}$ 关于 \mathcal{F}_τ 可测, 故 $\{B'(t), t \geq 0\}$ 与 (τ, B^τ) 相互独立. 由于 $-B'$ 与 B' 同分布, 故 (τ, B^τ, B') 与 $(\tau, B^\tau, -B')$ 同分布. 由于

$$B(t) = B(t \wedge \tau) + (B(t) - B(t \wedge \tau))$$

$$\tilde{B}(t) = B(t \wedge \tau) - (B(t) - B(t \wedge \tau))$$

故 $B(t)$ 与 $\tilde{B}(t)$ 同分布.

5.2 首中时和最大时

令 $M_t := \max_{0 \leq u \leq t} B(u)$, $T_a := \inf\{t > 0, B(t) = a\}$. $\tilde{B}(t) := B(t \wedge T_a) - (B(t) - B(t \wedge T_a))$

定理 5.2.1

设 $B = \{B(t), t \geq 0\}$ 是一个标准布朗运动, 对任意的 $a > 0$ 以及 $x \leq a$

1. $P(M_t \geq a, B(t) < x) = P(B(t) \geq 2a - x)$;

2. M_t 与 $|B(t)|$ 同分布, 且

$$P(M_t \geq a) = 2P(B(t) \geq a) = 2 \int_{\frac{a}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 2(1 - \Phi(\frac{a}{\sqrt{t}}))$$

3. $P(T_a \leq t) = P(M_t \geq a) = 2(1 - \Phi(\frac{a}{\sqrt{t}}))$



证明 (1)

$$\begin{aligned} P(M_t \geq a, B(t) \leq x) &= P(T_a \leq t, B(t) \leq x) \\ &= P(T_a \leq t, \tilde{B}(t) \leq x) \\ &= P(T_a \leq t, 2a - B(t) \leq x) \\ &= P(T_a \leq t, B(t) \geq 2a - x) \\ &= P(B(t) \geq 2a - x) \end{aligned}$$

(2)

$$\begin{aligned} P(M_t \geq a) &= P(M_t \geq a, B(t) \geq a) + P(M_t \geq a, B(t) < a) \\ &= P(B(t) \geq a) + P(B(t) \geq 2a - a) \\ &= 2P(B(t) \geq a) \\ &= P(|B(t)| \geq a) \end{aligned}$$

标准化 $\frac{B(t)}{\sqrt{t}} \sim N(0, 1)$, 故

$$2P(B(t) \geq a) = 2P\left(\frac{B(t)}{\sqrt{t}} \geq \frac{a}{\sqrt{t}}\right) = 2 \int_{\frac{a}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

(3)

$$\begin{aligned} P(T_a \leq t) &= P(M_t \geq a) \\ &= 2P(B(t) \geq a) \end{aligned}$$

定理 5.2.2

设 $B = \{B(t), t \geq 0\}$ 是一个标准布朗运动, $\forall a < 0$, 则有

$$P\left(\min_{0 \leq s \leq t} B(s) \leq a\right) = 2P(B(t) \geq -a) = 2P(B(t) \leq a).$$



证明 令 $m(t) = \min_{0 \leq s \leq t} B(s)$, $-\min_{0 \leq s \leq t} B(s) = \max_{0 \leq s \leq t} (-B(s))$.

$$\begin{aligned} P(m(t) \leq a) &= P(-\max_{0 \leq s \leq t} (-B(s)) \leq a) \\ &= P(\max_{0 \leq s \leq t} (-B(s)) \geq -a) \\ &= P(\max_{0 \leq s \leq t} B(s) \geq -a) \\ &= 2P(B(t) \geq -a) \end{aligned}$$

定理 5.2.3

$B = \{B(t), t \geq 0\}$ 是一个标准布朗运动, $\forall a$,

$$P(T_a \leq t) = 2 \int_{\frac{|a|}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 2(1 - \Phi(\frac{|a|}{\sqrt{t}}))$$



证明 当 $a < 0$ 时,

$$\begin{aligned} P(T_a \leq t) &= P\left(\min_{0 \leq s \leq t} B(s) \leq a\right) = 2P(B(t) \geq -a) \\ &= 2P\left(\frac{B(t)}{\sqrt{t}} \geq -\frac{a}{\sqrt{t}}\right) \\ &= 2(1 - \Phi(\frac{|a|}{\sqrt{t}})) \end{aligned}$$

当 $a \geq 0$ 时, $P(T_a \leq t) = 2(1 - \Phi(\frac{a}{\sqrt{t}}))$ 。

注

1. $P(T_a < \infty) = 1$;
2. $E[T_a] = \infty$.

例题 5.4 若 $\varphi(x)$ 为非负可测函数, X 为非负随机变量, 又 $\Phi(x) = \int_0^x \varphi(t)dt$, 则

$$E[\Phi(X)] = \int_0^\infty \varphi(t)P(X > t)dt.$$

证明 若用 $F(x)$ 记 X 的分布函数, 则有

$$\begin{aligned} E[\Phi(X)] &= \int_0^\infty \Phi(x)dF(x) \\ &= \int_0^\infty \int_0^x \varphi(t)dt dF(x) \\ &= \int_0^\infty \varphi(t) \left(\int_t^\infty dF(x) \right) dt \\ &= \int_0^\infty \varphi(t)P(X > t)dt. \end{aligned}$$

5.2.1 布朗运动的轨道性质

定义 5.2.1

若 g 为实函数, 则 g 在 $[a, b]$ 上的变差定义为

$$V_g([a, b]) = \sup_{\Delta} S^{\Delta}([a, b]),$$

其中

$$S^{\Delta}([a, b]) := \sum_{i=1}^n |g(t_i^n) - g(t_{i-1}^n)|,$$

其中 $\Delta := \{a = t_0^n < t_1^n < \cdots < t_n^n = b\}$. $|\Delta| := \sup_{1 \leq i \leq n} |t_i^n - t_{i-1}^n|$.



若 Δ' 为 Δ 的加细, 即若 $t \in \Delta$ 有 $t_i \in \Delta'$, 则由三角不等式 $S^{\Delta}([a, b]) \leq S^{\Delta'}([a, b])$. 故 $V_g([a, b]) = \lim_{|\Delta| \rightarrow 0} S^{\Delta}([a, b])$.

定义 5.2.2

若 g 为实函数, 则 g 在 $[a, b]$ 上的变差定义为

$$V_g([a, b]) = \sup_{\Delta} \sum_{i=1}^n |g(t_i^n) - g(t_{i-1}^n)|,$$

其中 $a = t_0^n < t_1^n < \cdots < t_n^n = b$ 表示 $[a, b]$ 上的一种划分.

若 g 为定义在 $[0, \infty)$ 上的函数, 则定义函数 g 的变差函数为关于 t 的实函数, $V_g(t) := V_g([0, t])$.



定义 5.2.3

若 g 为定义在 $[0, \infty)$ 上的函数, 称 g 是有限变差的, 若对任意的 t 有 $V_g(t) < \infty$. 称 g 是有界变差的, 若 $\sup_t V_g(t) < \infty$.



例题 5.5

1. 若 $g(t)$ 为单调增函数, 则 $V_g(t) = g(t) - g(0)$
2. 若 $g(t)$ 为单调减函数, 则 $V_g(t) = g(0) - g(t)$
3. 若 $g(t)$ 可微且有连续导数 $g'(t)$, 且 $g(t) = \int_0^t g'(s) ds$ 且 $\int_0^t |g'(s)| ds < \infty$, 则 $V_g(t) = \int_0^t |g'(s)| ds$.

证明 (1) 记 $0 = t_0^n < t_1^n < \dots < t_n^n = t$,

$$S^n[0, t] = \sum_{i=1}^n |g(t_i^n) - g(t_{i-1}^n)| = \sum_{i=1}^n (g(t_i^n) - g(t_{i-1}^n)) = g(t) - g(0). \text{ 故 } V_g(t) = g(t) - g(0)$$

(3)

记 $0 = t_0^n < t_1^n < \dots < t_n^n = t$, 有中值定理可知 $\int_{t_{i-1}}^{t_i} g'(s) ds = g'(\xi_i)(t_i - t_{i-1})$, $\xi_i \in (t_{i-1}, t_i)$.

故 $\left| \int_{t_{i-1}}^{t_i} g'(s) ds \right| = |g'(\xi_i)|(t_i - t_{i-1})$. 且

$$\begin{aligned} V_g(t) &= \lim_{|\Delta| \rightarrow 0} \sum_{i=1}^n |(g(t_i^n) - g(t_{i-1}^n))| \\ &= \lim_{|\Delta| \rightarrow 0} \sum_{i=1}^n \left| \int_{t_{i-1}}^{t_i} g'(s) ds \right| \\ &= \lim_{|\Delta| \rightarrow 0} \sum_{i=1}^n |g'(\xi_i)|(t_i - t_{i-1}) \\ &= \int_0^t |g'(s)| ds \end{aligned}$$

定义 5.2.4

若 g 为定义在 $[0, \infty)$ 上的函数, 则 g 在 $[0, t]$ 上的平方变差定义为

$$|g|(t) = \lim_{|\Delta| \rightarrow 0} \sum_{i=1}^n (g(t_i^n) - g(t_{i-1}^n))^2,$$

其中 $\Delta := \{0 = t_0^n < t_1^n < \dots < t_n^n = t\}$ 表示 $[0, t]$ 上的任意一种划分序列, $|\Delta| := \max_{1 \leq i \leq n} (t_i^n - t_{i-1}^n)$.



定义 5.2.5

实值过程 X 称为有限平方变差的, 若存在一个随机过程 $\langle X, X \rangle$ 使得对任意的 t 和 $[0, t]$ 上的任意的分割 $\Delta := \{0 = t_0^n < t_1^n < \dots < t_n^n = t\}$, 当

$$\lim_{|\Delta| \rightarrow 0} \sum_{i=1}^n (X(t_i^n) - X(t_{i-1}^n))^2 = \langle X, X \rangle_t.$$

(在依概率收敛意义下) 此时, 称 $\langle X, X \rangle$ 为 X 的平方变差过程。



定理 5.2.4

设 $[a, b]$ 上的任意的分割 $\Delta := \{a = t_0^n < t_1^n < \dots < t_n^n = b\}$, 则

$$\lim_{|\Delta| \rightarrow 0} \sum_{i=1}^n (B(t_i^n) - B(t_{i-1}^n))^2 = b - a$$

在 L^2 收敛意义下。



证明

$$\begin{aligned}
& E \left[\left(\sum_{k=1}^n (B(t_k^n) - B(t_{k-1}^n))^2 - (b-a)^2 \right) \right] \\
&= \text{Var} \left[\sum_{i=1}^n (B(t_i^n) - B(t_{i-1}^n))^2 \right] \\
&= \sum_{i=1}^n \text{Var} [(B(t_i^n) - B(t_{i-1}^n))^2] \\
&= \sum_{i=1}^n E [B(t_i^n) - B(t_{i-1}^n)]^4 - \left(E [B(t_i^n) - B(t_{i-1}^n)]^2 \right)^2 \\
&= \sum_{i=1}^n [3(t_i^n - t_{i-1}^n)^2 - (t_i^n - t_{i-1}^n)^2] \\
&= 2 \sum_{i=1}^n (t_i^n - t_{i-1}^n)^2 \\
&\leq 2 \max_j |t_j^n - t_{j-1}^n| \sum_{i=1}^n |t_i^n - t_{i-1}^n| \\
&= 2 \max_j |t_j^n - t_{j-1}^n| (b-a) \rightarrow 0
\end{aligned}$$

注

由于 L^2 收敛可推出依概率收敛, 故 $\langle B, B \rangle_t = t$.

推论 5.2.1

在任意有限区间 $[a, b]$ 上, 布朗运动轨道几乎处处无穷变差, 即 $V_B([a, b]) = \infty a.s.$



证明 由于

$$\lim_{|\Delta| \rightarrow 0} E \left[\left(\sum_k (B(t_k^n) - B(t_{k-1}^n))^2 - (b-a)^2 \right)^2 \right] = 0$$

由于 L^2 收敛可以推出依概率收敛, 故可以得到下列几乎处处收敛。即存在 $\Omega_0 \subseteq \Omega, P(\Omega_0) = 1$, 存在分割 Δ^n 使得 $|\Delta^n| \rightarrow 0$ 。且对任意的 $\omega \in \Omega_0$ 有

$$\lim_{n \rightarrow \infty} \sum_{t_i^n \in \Delta^n} (B(t_i^n) - B(t_{i-1}^n))^2 = b - a \text{ a.e..}$$

若 $\{B(t)\}$ 在 $[a, b]$ 上的全变差 $V_B[a, b] < \infty$, 则有

$$\sum_i (B(t_i^n) - B(t_{i-1}^n))^2 \leq \sup_i |B(t_i^n) - B(t_{i-1}^n)| V_B[a, b]$$

由于轨道连续, 则在 $[a, b]$ 上一致连续, 则 $\sup_i |B(t_i^n) - B(t_{i-1}^n)| \rightarrow 0$ ($\exists m \rightarrow \infty$ 时, $\max(t_i^n - t_{i-1}^n) \rightarrow 0$)。故

$$\sum_i (B(t_i^n) - B(t_{i-1}^n))^2 \rightarrow 0.$$

矛盾即 $V_B[a, b] = \infty$.

作为 t 的函数, Brown 运动的几乎所有路径 $B(t)$ 都具有下面性质:

1. 是 t 的连续函数;
2. 在任何区间 (无论区间多小) 上都不是单调的;
3. 在任何点都不是可微的;
4. 在任何区间 (无论区间多小) 上都是无限变差的;
5. 对任意 t , 在 $[0, t]$ 上的二次变差等于 t 。

5.3 Ito 积分

定理 5.3.1

(a) 若 $F(x)$ 为 \mathbb{R} 上的右连续不减有界函数, 则在 $(\mathbb{R}, \mathcal{B})$ 必存在唯一的有限测度 μ , 使得:

$$\mu((a, b]) = F(b) - F(a), \quad -\infty \leq a < b < +\infty.$$

(b) 若 $F(x)$ 为 \mathbb{R} 上的右连续不减的实值函数, 则必存在 $(\mathbb{R}, \mathcal{B})$ 上唯一的 σ 有限的测度 μ 使得

$$\mu((a, b]) = F(b) - F(a), \quad -\infty \leq a < b < +\infty.$$



问题 $\int_0^t f(s)dB(s)$?

定义 5.3.1

设 $B(t)$ 是一维布朗运动, 定义 $\mathcal{F}_t := \sigma(B(u), u \leq t)$. 即 \mathcal{F}_t 是包含所有形如

$$\{\omega | B(t_1, \omega) \in F_1, \dots, B(t_k, \omega) \in F_k\}$$

的集合的最小 σ -代数. $t_j \leq t, F_j \in \mathcal{B}(\mathbb{R})$.



定义 5.3.2

设 $\mathcal{V} = \mathcal{V}(\mathcal{S}, T) := \{f(t, \omega) : [0, \infty) \times \Omega \rightarrow \mathbb{R}\}$ 且其中 f 满足:

- (a) $(t, \omega) \rightarrow f(t, \omega)$ 是 $\mathcal{B}(\mathbb{R}) \times \mathcal{F}$, 其中 $\mathcal{B}(\mathbb{R})$ 是 $[0, \infty)$ 上的 Borel σ -代数;
- (b) $f(t, \omega)$ 是 \mathcal{F}_t 适应的;
- (c) $E \left[\int_{\mathcal{S}} f(t, \omega)^2 dt \right] < \infty$.



定义 5.3.3

称 $\phi \in \mathcal{V}$ 为一个简单过程, 若

$$\phi(t, \omega) = \sum_J \theta_j(\omega) I_{[t_j, t_{j+1})}(t).$$

其中 $\theta_j(\cdot)$ 为 \mathcal{F}_{t_j} 可测的, $\mathcal{S} = t_0 < t_1 < \dots < t_n = T$. 定义 ϕ 的随机积分如下:

$$\int_{\mathcal{S}}^T \phi(t, \omega) dB(t, \omega) = \sum_J \theta_j(\omega) [B(t_{j+1}, \omega) - B(t_j, \omega)].$$



引理 5.3.1 (Ito 等距)

如果 $\phi \in \mathcal{V}$ 为一个简单过程, 则

$$E \left[\left(\int_s^T \phi(t, \omega) dB(t, \omega) \right)^2 \right] = E \left[\int_s^T \phi^2(t, \omega) dt \right]$$



证明 记 $\Delta B_j = B(t_{j+1}) - B(t_j)$, $\phi(t, \omega) = \sum_j \theta_j(\omega) I_{[t_j, t_{j+1})}(t)$.

$$\begin{aligned} E \left[\left(\int_s^T \phi(t, \omega) dB(t, \omega) \right)^2 \right] &= E \left[\left(\sum_j \theta_j(\omega) (B(t_{j+1}) - B(t_j)) \right)^2 \right] \\ &= E \left[\sum_{i,j} \theta_i \theta_j (B(t_{i+1}) - B(t_i)) (B(t_{j+1}) - B(t_j)) \right] \\ &= \sum_{i,j} E [\theta_i \theta_j (B(t_{i+1}) - B(t_i)) (B(t_{j+1}) - B(t_j))] \end{aligned}$$

当 $j > i$

$$\begin{aligned}
 & E [\theta_i \theta_j (B(t_{i+1}) - B(t_i))(B(t_{j+1}) - B(t_j))] \\
 &= E [E [\theta_i \theta_j (B(t_{i+1}) - B(t_i))(B(t_{j+1}) - B(t_j)) \mid \mathcal{F}_{t_i}]] \\
 &= E [\theta_i \theta_j E [(B(t_{j+1}) - B(t_j)) \mid \mathcal{F}_{t_i}] E [B(t_{i+1}) - B(t_i) \mid \mathcal{F}_{t_i}]] \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 E \left[\sum_j \theta_j^2 (B(t_{j+1}) - B(t_j))^2 \right] &= \sum_j E [E [\theta_j^2 (B(t_{j+1}) - B(t_j))^2 \mid \mathcal{F}_{t_j}]] \\
 &= \sum_j E [\theta_j^2 E [(B(t_{j+1}) - B(t_j))^2 \mid \mathcal{F}_{t_j}]] \\
 &= \sum_j E [\theta_j^2 E [(B(t_{j+1}) - B(t_j))^2]] \\
 &= \sum_j E [\theta_j^2 (t_{j+1} - t_j)] \\
 &= E \left[\int_s^T \phi^2 dt \right]
 \end{aligned}$$

定理 5.3.2

设 $X = \{X(t), t \geq 0\} \in \mathcal{V}$, 则存在简单过程 $\{X^m\} \subset \mathcal{V}$ 使得

$$\lim_{m \rightarrow \infty} E \left[\int_s^T |X^m(t) - X(t)|^2 dt \right] = 0.$$



定义 5.3.4

设 $f \in \mathcal{V}(\mathcal{S}, T)$, 则 f 的 Itô 积分定义为

$$\int_s^T f(t, \omega) dB(t, \omega) = \lim_{n \rightarrow \infty} \int_s^T \phi_n(t, \omega) dB(t, \omega)$$

在 L^2 中的极限, 其中 ϕ_n 为简单过程, 满足:

$$E \left[\int_s^T (f(t, \omega) - \phi_n(t, \omega))^2 dt \right] \rightarrow 0, \text{ 当 } n \rightarrow \infty.$$



注

$$\begin{aligned}
 & E \left[\left(\int_s^T \phi_n(t, \omega) dB(t, \omega) - \int_s^T \phi_m(t, \omega) dB(t, \omega) \right)^2 \right] \\
 &= E \left[\left(\int_s^T (\phi_n - \phi_m) dB(t, \omega) \right)^2 \right] \\
 &= E \left[\int_s^T (\phi_n - \phi_m)^2 dt \right] \\
 &\leq E \left[\int_s^T (\phi_n - f)^2 dt \right] + E \left[\int_s^T (\phi_m - f)^2 dt \right] \\
 &\rightarrow 0.
 \end{aligned}$$

故 $\left\{ \int_s^T \phi_n(t, \omega) dB(t, \omega) \right\}$ 存在极限。

5.3.1 Ito 积分性质

定理 5.3.3

设 $f, g \in \mathcal{V}(0, T)$, $0 \leq S < U < T$.

- (a) $\int_S^T f(t, \omega) dB(t, \omega) = \int_S^U f(t, \omega) dB(t, \omega) + \int_U^T f(t, \omega) dB(t, \omega)$ a.s.
- (b) 对任意的常数 c , $\int_S^T (cf(t, \omega) + g(t, \omega)) dB(t, \omega) = c \int_S^T f(t, \omega) dB(t, \omega) + \int_S^T g(t, \omega) dB(t, \omega)$ a.s.
- (c) $E \left[\int_S^T f(t, \omega) dB(t, \omega) \right] = 0$
- (d) $\int_S^T f(t, \omega) dB(t, \omega)$ 为 \mathcal{F}_T 可测的。



注 Ito 积分不具有单调性。

反例 $\int_0^1 1 dB(t) = B(1)$, $\int_0^1 2 dB(t) = 2B(1)$.

例题 5.6 求 $E \left(\int_0^1 e^{B(t)} dB(t) \right)$, $E \left(\left[\int_0^1 e^{B(t)} dB(t) \right]^2 \right)$

证明 由于 e^x 连续, 故有定义。因此, $e^{B(t)}$ 关于 $\mathcal{B}(\mathbb{R}) \times \mathcal{F}$ 可测。故 $E \left(\int_0^1 e^{B(t)} dB(t) \right) = 0$ 。

$$\begin{aligned} E[e^{2B(t)}] &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} e^{2x} dx \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-2t)^2}{2t}} e^{2t} dx \\ &= e^{2t}. \end{aligned}$$

$$\begin{aligned} E \left(\left[\int_0^1 e^{B(t)} dB(t) \right]^2 \right) &= E \left(\int_0^1 e^{2B(t)} dt \right) \\ &= \int_0^1 E[e^{2B(t)}] dt \\ &= \int_0^1 e^{2t} dt = \frac{1}{2}(e^2 - 1) \end{aligned}$$

例题 5.7 求 $E \left(\int_0^1 B(t) dB(t) \right)$, $E \left(\left[\int_0^1 B(t) dB(t) \right]^2 \right)$

证明 由于 $B(t)$ 关于 $\mathcal{B}(\mathbb{R}) \times \mathcal{F}$ 可测。故 $E \left(\int_0^1 B(t) dB(t) \right) = 0$ 。

$$\begin{aligned} E \left[\int_0^1 B(t) dB(t) \right]^2 &= E \left[\int_0^1 B^2(t) dt \right] \\ &= \int_0^1 E[B^2(t)] dt \\ &= \int_0^1 t dt = \frac{1}{2}. \end{aligned}$$

定理 5.3.4

设 $f \in \mathcal{V}(0, T)$, 则存在 $\int_0^t f(s, \omega) dB(s, \omega)$, $0 \leq t \leq T$ 的一个连续修正, 即存在一个关于 (Ω, \mathcal{F}, P) 上的 t 连续的随机过程 J_t 使得

$$P[J_t = \int_0^t f dB] = 1, \forall 0 \leq t \leq T.$$



例题 5.8 设 $B(0) = 0$, $B = \{B(t), t \geq 0\}$ 是一个标准布朗运动, 则

$$\int_0^t B(s) dB(s) = \frac{1}{2} B^2(t) - \frac{1}{2} t.$$

证明 (1) 构造简单过程: $\phi_n(s) := \sum_j B(t_j) I_{[t_j^n, t_{j+1}^n)}(s) \rightarrow B(s)$,

$$\begin{aligned}
 E \left[\int_0^t (\phi_n(s) - B(s))^2 ds \right] &= E \left[\sum_j \int_{t_j^n}^{t_{j+1}^n} (\phi_n(s) - B(s))^2 ds \right] \\
 &= E \left[\sum_j \int_{t_j^n}^{t_{j+1}^n} (B(t_j^n) - B(s))^2 ds \right] \\
 &= \sum_j \int_{t_j^n}^{t_{j+1}^n} E[(B(t_j^n) - B(s))^2] ds \\
 &= \sum_j \int_{t_j^n}^{t_{j+1}^n} (s - t_j^n) ds \\
 &= \sum_j \frac{1}{2} (t_{j+1}^n - t_j^n)^2 \rightarrow 0, \text{ 当 } |\Delta^n| := \max_j (t_{j+1}^n - t_j^n) \rightarrow 0
 \end{aligned}$$

(2) 求简单过程 Ito 积分:

$$\int_0^t \phi_n(t) dB(t) = \sum_j B(t_j^n) [B(t_{j+1}^n) - B(t_j^n)].$$

(3) 取极限:

$$\int_0^t \phi_n(t) dB(t) = \sum_j B(t_j^n) [B(t_{j+1}^n) - B(t_j^n)] \rightarrow \frac{1}{2} B^2(t) - \frac{1}{2} t.$$

$$\begin{aligned}
 B^2(t_{j+1}^n) - B^2(t_j^n) &= [B(t_{j+1}^n) - B(t_j^n)]^2 + 2B(t_j^n)[B(t_{j+1}^n) - B(t_j^n)] \\
 B^2(t) &= \sum_{j=1}^{n-1} [B^2(t_{j+1}^n) - B^2(t_j^n)] \\
 &= \sum_{j=1}^{n-1} [B(t_{j+1}^n) - B(t_j^n)]^2 + 2B(t_j^n)[B(t_{j+1}^n) - B(t_j^n)]
 \end{aligned}$$

即

$$\sum_{j=1}^{n-1} B(t_j^n) [B(t_{j+1}^n) - B(t_j^n)] = \frac{1}{2} B^2(t) - \frac{1}{2} \sum_{j=1}^{n-1} [B(t_{j+1}^n) - B(t_j^n)]^2$$

定义 5.3.5

设 $\{B(t)\}$ 是 (Ω, \mathcal{F}, P) 上的一维布朗运动, 一个 1 维 Itô 过程是 (Ω, \mathcal{F}, P) 上具有如下形式的随机过程 $X(t)$:

$$X(t) := X(0) + \int_0^t u(s, \omega) ds + \int_0^t v(s, \omega) dB(s, \omega).$$

其中 $v \in \mathcal{W}$:

- $(t, \omega) \rightarrow v(t, \omega)$ 是 $\mathcal{B}(\mathbb{R}) \times \mathcal{F}$, 其中 $\mathcal{B}(\mathbb{R})$ 是 $[0, \infty)$ 上的 Borel σ -代数;
- $v(t, \omega)$ 是 \mathcal{F}_t 适应的;
- $P \left[\int_0^t v^2(s, \omega) ds < \infty, \forall t \geq 0 \right] = 1$.

$X(t)$ 的微分形式:

$$dX(s) = u(s, \omega) ds + v(s, \omega) dB(s, \omega).$$

u 满足:

- $(t, \omega) \rightarrow u(t, \omega)$ 是 $\mathcal{B}(\mathbb{R}) \times \mathcal{F}$, 其中 $\mathcal{B}(\mathbb{R})$ 是 $[0, \infty)$ 上的 Borel σ -代数;

- $u(t, \omega)$ 是 \mathcal{F}_t 适应的;
- $P \left[\int_0^t |u(s, \omega)| ds < \infty, \forall t \geq 0 \right] = 1$.



定理 5.3.5

设 $X(t)$ 为一个一维的 Ito 过程:

$$X(t) := X(0) + \int_0^t u(s) ds + \int_0^t v(s) dB(s).$$

其中 $g(t, x) \in C^2([0, \infty) \times \mathbb{R})$, 即 g 为 $[0, \infty) \times \mathbb{R}$ 上的二阶连续可微函数, 则 $Y(t) := g(t, X(t))$ 也为 Ito 过程, 且有

$$dY(t) = \frac{\partial g}{\partial t}(t, X(t))dt + \frac{\partial g}{\partial x}(t, X(t))dX(t) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X(t))(dX(t))^2$$

其中 $dt \cdot dt = 0$, $dt \cdot dB(t) = 0$, $dB(t) \cdot dt = 0$, $dB(t) \cdot dB(t) = dt$.

故

$$dY(t) = \left[\frac{\partial g}{\partial t}(t, X(t)) + u(t) \frac{\partial g}{\partial x}(t, X(t)) + \frac{1}{2} v^2(t) \frac{\partial^2 g}{\partial x^2}(t, X(t)) \right] dt + v(t) \frac{\partial g}{\partial x}(t, X(t)) dB(t).$$



注

$$X(t) := X(0) + \int_0^t u(s) ds + \int_0^t v(s) dB(s),$$

$$\text{则 } dX(t) := u(t)dt + v(t)dB(t).$$

$g(t, x) \in C^2([0, \infty) \times \mathbb{R})$, 令 $Y(t) := g(t, X(t))$, 则

$$\begin{aligned} dY(t) &= \frac{\partial g}{\partial t}(t, X(t))dt + \frac{\partial g}{\partial x}(t, X(t))dX(t) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X(t))(dX(t))^2 \\ &= \frac{\partial g}{\partial t}(t, X(t))dt + \frac{\partial g}{\partial x}(t, X(t))(u(t)dt + v(t)dB(t)) \\ &\quad + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X(t))(u(t)dt + v(t)dB(t))(u(t)dt + v(t)dB(t)) \\ &= \frac{\partial g}{\partial t}(t, X(t))dt + \frac{\partial g}{\partial x}(t, X(t))(u(t)dt + v(t)dB(t)) \\ &\quad + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X(t))(v^2(t)dt) \\ &= \frac{\partial g}{\partial t}(t, X(t)) + u(t) \frac{\partial g}{\partial x}(t, X(t)) + \frac{1}{2} v^2(t) \frac{\partial^2 g}{\partial x^2}(t, X(t))dt \\ &\quad + v(t) \frac{\partial g}{\partial x}(t, X(t))dB(t). \end{aligned}$$

例题 5.9 求 $I := \int_0^t B(s)dB(s)$.

证明 取 $f(x) = x^2$, 则 $f'(x) = 2x$, $f''(x) = 2$.

由 Itô 公式,

$$\begin{aligned} dB^2(t) &= f'(B(t))dB(t) + \frac{1}{2}f''(B(t))(dB(t))^2 \\ &= 2B(t)dB(t) + 1dt \end{aligned}$$

两边从 0 积分到 t , 得

$$\begin{aligned} B^2(t) - B^2(0) &= 2 \int_0^t B(s)dB(s) + \int_0^t 1ds \\ B^2(t) &= 2 \int_0^t B(s)dB(s) + t \end{aligned}$$

因此,

$$\int_0^t B(s) dB(s) = \frac{1}{2} B^2(t) - \frac{1}{2} t$$

例题 5.10 设 $f(x) = \sin x$, $Y(t) = f(B(t))$, 求 $dY(t)$ 。

证明

$$f(x) = \sin x, \quad f'(x) = \cos x, \quad f''(x) = -\sin x.$$

故

$$\begin{aligned} d(\sin(B(t))) &= \cos(B(t)) dB(t) - \frac{1}{2} \sin(B(t)) (dB(t))^2 \\ &= \cos(B(t)) dB(t) - \frac{1}{2} \sin(B(t)) dt \end{aligned}$$

但不会有这种情况:

例题 5.11 设 $B(t)$ 是一维布朗运动, 对下列 $Y(t)$ 给出其微分和积分形式。

$$Y(t) = e^{\frac{1}{2}t} \cos(B(t))$$

证明

$$\text{取 } f(t, x) = e^{\frac{1}{2}t} \cos x, \quad \frac{\partial f}{\partial t} = \frac{1}{2} e^{\frac{1}{2}t} \cos x, \quad \frac{\partial f}{\partial x} = -e^{\frac{1}{2}t} \sin x, \quad \frac{\partial^2 f}{\partial x^2} = -e^{\frac{1}{2}t} \cos x. \text{ 取 } X(t) = B(t).$$

$$\begin{aligned} dY(t) &= \frac{1}{2} e^{\frac{1}{2}t} \cos B(t) dt - (e^{\frac{1}{2}t} \sin B(t) dB(t) - \frac{1}{2} e^{\frac{1}{2}t} \cos B(t) (dB(t))^2) \\ &= \frac{1}{2} e^{\frac{1}{2}t} \cos B(t) dt - (e^{\frac{1}{2}t} \sin B(t) dB(t) - \frac{1}{2} e^{\frac{1}{2}t} \cos B(t) dt) \\ &= -e^{\frac{1}{2}t} \sin B(t) dB(t) \end{aligned}$$