

目 录

0.1 概率与测度	1
0.2 乘积空间	1
0.3 离散型随机变量	3
0.4 一般型随机变量	4
0.5 条件概率、条件分布与条件期望	6
0.6 随机过程	8
0.7 建议阅读书目	8
0.8 记号	8
第一章 马氏链	9
1.1 定义与例子	9
1.2 不变分布与可逆分布	21
1.3 状态的分类	33
1.4 首达时与强马氏性	35
1.5 常返性	42
1.6 *可配称马氏链的常返性	60
1.7 遍历定理与正常返	68
1.8 强遍历定理	82
1.9 收敛速度	91
1.10 一维简单随机游动	94
1.11 分支过程	106
1.12 综合练习题	111
第二章 跳过程	113
2.1 泊松过程	113
2.2 跳过程的构造及其转移概率	123
2.3 跳过程的性质	130
2.4 排队系统	139
2.5 接触过程	142
2.6 *最小过程	145
第三章 布朗运动	151
3.1 高斯分布与高斯过程	151

3.2 布朗运动的定义与Levi构造	152
3.3 不变原理概述	158
3.4 布朗轨道的性质	160
3.5 位势理论	171
3.6 布朗桥与OU 过程	180
3.7 随机积分与随机微分方程简介	184
第四章 附录	189
参考文献	193

第零章 预备知识

§0.1 概率与测度

考虑一个随机试验, 其中的每一个试验结果被称为一个样本, 记为 ω . 称所有样本组成的集合为**样本空间**, 记为 Ω , 它是一个非空集合. 称 Ω 的子集为事件, 通常记为 A, B, \dots . 称一个以事件为元素的集合为**集合系**. 最大的集合系由 Ω 的所有子集组成, $\mathcal{T}_\Omega := \{A : A \subseteq \Omega\}$. 假设 $\mathcal{F} \subseteq \mathcal{T}_\Omega$ 满足下面三条性质:

- (i) $\Omega \in \mathcal{F}$;
- (ii) 若 $A \in \mathcal{F}$, 则 $A^c \in \mathcal{F}$;
- (iii) 若 $A_n \in \mathcal{F}, n = 1, 2, \dots$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$,

那么, 称 \mathcal{F} 为 Ω 上的一个 σ -代数, 称 (Ω, \mathcal{F}) 为一个可测空间. \mathcal{F} 中的元素被称为可测事件, 由于本书中所考虑的都是可测事件, 因此在下文中我们将省略“可测”二字. 进一步, 假设 $P : \mathcal{F} \rightarrow \mathbb{R}$ 满足下面三条性质:

- (i) 非负性: 对任意 $A \in \mathcal{F}$ 均有 $P(A) \geq 0$;
- (ii) 规范性: $P(\Omega) = 1$;
- (iii) 可列可加性: 若 $A_n \in \mathcal{F}, n = 1, 2, \dots$ 两两不交, 则

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

那么, 称 P 为 \mathcal{F} 上的一个概率, 也称 P 为 Ω 上的概率. 称 $P(A)$ 为事件 A (发生) 的概率. 称 (Ω, \mathcal{F}, P) 这个三元组为一个概率空间.

一般地, 假设 (Ω, \mathcal{F}) 为一个可测空间, $\mu : \mathcal{F} \rightarrow \mathbb{R}_+ \cup \{\infty\}$. 如果 $\mu(\emptyset) = 0$ 并且 μ 满足可列可加性, 那么称 μ 为 \mathcal{F} 上的测度, 也称 μ 为 Ω 上的测度. 如果 $\mu(\Omega) < \infty$, 那么我们总可以将它进行归一化, 从而得到概率

$$P(A) := \frac{\mu(A)}{\mu(\Omega)}, \quad A \in \mathcal{F}.$$

§0.2 乘积空间

假设 I 是指标集¹. 假设 $(\Omega_\alpha, \mathcal{F}_\alpha), \alpha \in I$ 是一族可测空间. 给定一族样本 $\omega_\alpha \in \Omega_\alpha, \alpha \in I$, 记 $\omega = (\omega_\alpha : \alpha \in I)$. 令

$$\Omega = \prod_{\alpha \in I} \Omega_\alpha := \{\omega : \omega_\alpha \in \Omega_\alpha\}.$$

¹指标集非空, 可以是可数集, 也可以是不可数集.

例0.2.1. 若 $I = \{1, \dots, n\}$, 则 $\Omega = \Omega_1 \times \dots \times \Omega_n$, $\omega = (\omega_1, \dots, \omega_n)$ 为一个 n 维向量. ω 的第 r 个分量(即, 第 r 维坐标)满足: $\omega_r \in \Omega_r$.

若 $I = \{1, 2, \dots\}$, 则 $\omega = (\omega_1, \omega_2, \dots)$ 为一个无穷维向量, 或一个序列. 此时, 也将 Ω 记为 $\Omega_1 \times \Omega_2 \times \dots$.

假设 \mathcal{E} 是 Ω 上的一个非空的集合系. 则

$$\sigma(\mathcal{E}) := \bigcap_{\mathcal{F} \text{ 是 } \sigma \text{ 代数}, \mathcal{F} \supseteq \mathcal{E}} \mathcal{F}$$

是包含 \mathcal{E} 的最小的 σ 代数, 称其为 \mathcal{E} 生成的 σ 代数, 记为 $\sigma(\mathcal{E})$. 取

$$\mathcal{E} = \{\{\omega : \omega_\alpha \in D\}, \alpha \in I, D \in \mathcal{F}_\alpha\}, \quad \mathcal{F} = \prod_{\alpha \in I} \mathcal{F}_\alpha := \sigma(\mathcal{E}).$$

称 (Ω, \mathcal{F}) 为 $(\Omega_\alpha, \mathcal{F}_\alpha)$, $\alpha \in I$ 的乘积空间.

进一步, 假设对任意 $\alpha \in I$, 我们还有 Ω_α 上的概率 P_α . 由附录中的定理 4.0.13, 可以推出在 \mathcal{F} 上存在唯一的概率 P 满足如下条件: 对任意 $n \geq 1$, $\alpha_1, \dots, \alpha_n \in I$ 两两不同, $D_1 \in \mathcal{F}_{\alpha_1}, \dots, D_n \in \mathcal{F}_{\alpha_n}$,

$$P(\{\omega : \omega_{\alpha_1} \in D_1, \dots, \omega_{\alpha_n} \in D_n\}) = \prod_{r=1}^n P_{\alpha_r}(A_r).$$

将此概率记为 $\prod_{\alpha \in I} P_\alpha$. 称 (Ω, \mathcal{F}, P) 为 $(\Omega_\alpha, \mathcal{F}_\alpha, P_\alpha)$, $\alpha \in I$ 的乘积空间.

当 $I = \{1, \dots, n\}$ 或 $\{1, 2, \dots\}$, 也将 \mathcal{F} 记为 $\mathcal{F}_1 \times \dots \times \mathcal{F}_n$ 或 $\mathcal{F}_1 \times \mathcal{F}_2 \times \dots$; 将 P 记为 $P_1 \times \dots \times P_n$ 或 $P_1 \times P_2 \times \dots$. 又若 $\Omega_r = \Omega_1$, 也将 Ω 记为 Ω_1^n 或 Ω_1^∞ , 类似地有 \mathcal{F}_1^n , \mathcal{F}_1^∞ ; P_1^n , P_1^∞ .

假设 $A_1, \dots, A_n \in \mathcal{F}$, 若对任意 $1 \leq k \leq n$, $1 \leq i_1 < \dots < i_k \leq n$,

$$P(A_{i_1} \cdots A_{i_k}) = P(A_{i_1}) \cdots P(A_{i_k}),$$

则称 A_1, \dots, A_n 相互独立. 不难验证, 在乘积空间中, $\{\omega : \omega_{\alpha_1} \in A_1\}, \dots, \{\omega : \omega_{\alpha_n} \in A_n\}$ 这 n 个事件是相互独立的. 事实上, 如果 $(\Omega_\alpha, \mathcal{F}_\alpha, P_\alpha)$, $\alpha \in I$ 刻画了一族小型随机试验, 那么, (Ω, \mathcal{F}, P) 则刻画了相互独立的小型随机试验组成的大型随机试验.

例0.2.2 (伯努利试验). 伯努利(Bernoulli)试验指的是独立地抛一枚硬币的随机试验. 用 H 代表抛到正面(即, 国徽朝上), 用 T 代表抛到反面(即, 国徽朝上). 并假设抛到正面的概率为 p , 其中, $0 \leq p \leq 1$. 假设只抛一次, 那么该小型随机试验对应的概率空间 $(\Omega_1, \mathcal{F}_1, P_1)$ 如下: $\Omega_1 = \{H, T\}$, $\mathcal{F}_1 = \mathcal{T}_{\Omega_1}$,

$$P_1(\emptyset) = 0, \quad P_1(\{H\}) = p, \quad P_1(\{T\}) = 1 - p, \quad P_1(\{H, T\}) = 1.$$

那么 $(\Omega_1^n, \mathcal{F}_1^n, P_1^n)$ 就刻画了独立地抛 n 次硬币的随机试验, 而 $(\Omega_1^\infty, \mathcal{F}_1^\infty, P_1^\infty)$ 则刻画了(不停地)独立地抛这枚硬币的随机试验.

习题

1. 证明: $\mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ 是所有形如 $A_1 \times \cdots \times A_n$ 的事件生成的 σ 代数, $\mathcal{F}_1 \times \mathcal{F}_2 \times \cdots$ 是所有形如 $A_1 \times \cdots \times A_n \times \Omega_{n+1} \times \Omega_{n+2} \times \cdots$ 的事件生成的 σ 代数, 其中, $n \geq 1$, $A_r \in \mathcal{F}_r$, $r = 1, \dots, n$.

2* 假设 I 不可数. 取 $\alpha_1, \alpha_2, \dots \subseteq I$, 两两不交. 令

$$\mathcal{F}_{\alpha_1, \alpha_2, \dots} := \left\{ \{\omega : (\omega_{\alpha_1}, \omega_{\alpha_2}, \dots) \in D\}, \quad D \in \prod_{n=1}^{\infty} \mathcal{F}_{\alpha_n} \right\}.$$

证明: $\mathcal{F}_{\alpha_1, \alpha_2, \dots}$ 为 σ 代数, 并且

$$\mathcal{F} = \bigcup_{\alpha_1, \alpha_2, \dots \subseteq I, \text{ 两两不交}} \mathcal{F}_{\alpha_1, \alpha_2, \dots}$$

3* 利用定理 4.0.13, 证明: $\prod_{\alpha \in I} P_{\alpha}$ 存在唯一.

§0.3 离散型随机变量

在面对实际问题时, 人们往往是将所关注的量当做随机变量来处理. 假如我们并不关心这个量的期望和方差等数字特征, 那么我们可以将处理的对象拓宽为取值不一定是实数的“量”. 只要其可能取值是可数个, 它本质上就和离散型随机变量一样. 为了不引入过多的词语, 我们仍然称之为一个离散型随机变量.

以下假设 S 是一个非空的可数集, 它表示离散型随机变量的取值范围, 被称为状态空间. S 中的每个点都被称为一个状态, 通常用 i, j, k, \dots 表示. 离散型随机变量指的是从某个样本空间 Ω 到 S 的映射, 通常用 X, Y, \dots 等表示, 它须满足如下条件:

$$\{X = i\} := \{\omega : X(\omega) = i\} \in \mathcal{F}, \quad \forall i \in S.$$

离散型随机变量常用于表示一个系统所处的状态. 我们用粒子来代表这个系统, 当 $\{X = i\}$ 发生时, 我们称粒子处于状态 i .

为区分 \mathcal{F} 上的概率, 我们称 \mathcal{T}_S 上的概率为 S 上的一个分布. 若

$$\mu_i \geq 0, \quad \forall i \in S \quad \text{且} \quad \sum_{i \in S} \mu_i = 1,$$

则称 $\{\mu_i : i \in S\}$ 为 S 上的一个分布列. 不难发现, 分布 μ 与分布列 $\{\mu_i : i \in S\}$ 在下式的意义下是等价的, 因此我们也称 $\{\mu_i : i \in S\}$ 为 S 上的一个分布:

$$\begin{cases} \mu_i = \mu(\{i\}), & i \in S, \\ \mu(A) = \sum_{i \in A} \mu_i, & A \in \mathcal{T}_S. \end{cases}$$

类似地, 如果 $\mu_i \in \mathbb{R}_+ \cup \{\infty\}$, $i \in S$, 那么也称 $\{\mu_i : i \in S\}$ 为 S 上的一个测度, 因为它对应着 \mathcal{T}_S 上的一个测度 $\mu(A) := \sum_{i \in A} \mu_i$, $A \in \mathcal{T}_S$.

给定 \mathcal{F} 上的一个概率 P , 随机变量 X 便诱导了 \mathcal{T}_S 上的一个分布:

$$\mu_i = P(X = i), \quad \forall i \in S,$$

称其为 X 的分布. 此时, 称 X 服从分布 μ . 于是, X 的所有统计学性质都由其分布 μ 完全决定.

例0.3.1. 将十字路口的交通信号灯所处的状态视为一个随机变量, 如果用 R , G , Y 分别表示信号灯处于红灯、绿灯、黄灯亮起的状态, 则状态空间是 $S = \{R, G, Y\}$. 比如说, 在某一时刻, 信号灯处于红灯、绿灯、黄灯的概率分布为 $0.45, 0.45, 0.1$. 那么,

$$\mu_R = 0.45, \quad \mu_G = 0.45, \quad \mu_Y = 0.1.$$

用 X 表示信号灯(视为粒子)所处的状态. 那么, X 是离散型随机变量, $X \sim \mu$, 因此, 事件 $\{X = R\}$ 的概率为 $P(X = R) = 0.45$.

在很多模型中, 我们需要将离散型随机变量 X 理解为一个粒子所处的位置. 因此, 状态 i 就是粒子的一个可能的位置, S 就是所有可能的位置. 此时, $\{X = i\}$ 即是该粒子位于位置 i 的这一事件.

例0.3.2. 将一个粒子随机地放置在 $S = \mathbb{Z}_+ := \{0, 1, 2, \dots\}$ 中, 将它放在 i 处的概率为

$$\mu_i = \frac{\lambda^i}{i!} e^{-\lambda}, \quad i = 0, 1, 2, \dots$$

其中, $\lambda > 0$. 将该粒子的位置记为 X , 则 X 是一个离散型随机变量, 其分布 μ 就是参数为 λ 的泊松分布.

§0.4 一般型随机变量

假设 Ω, \mathcal{F} 和 (S, \mathcal{S}) 为可测空间, 其中, S 可以是不可数的. 假设 $X : \Omega \rightarrow S$. 若

$$\{X \in D\} := \{\omega : X(\omega) \in D\} \in \mathcal{F}, \quad \forall D \in \mathcal{S},$$

则称 X 为取值于 S 的随机变量, 称 S 为 X 的状态空间或值空间. 又若 P 为 Ω 上的概率, 则称

$$\mu : \mathcal{S} \rightarrow \mathbb{R}, \quad D \mapsto P(X \in D)$$

为 X 的分布, 称 X 服从 μ , 记为 $X \sim \mu$. 若两个随机变量 X 与 Y 服从相同的分布, 则称它们同分布, 记为 $X \stackrel{d}{=} Y$.

给定整数 $n \geq 2$. 假设 X_1, \dots, X_n 是 Ω 上的随机变量, 分别取值于 S_1, \dots, S_n , 其上的 σ 代数分别为 $\mathcal{S}_1, \dots, \mathcal{S}_n$. 那么, 称 $\mathbf{X} = (X_1, \dots, X_n)$ 为一个 n 维随机向量. 事实上, 取 $\mathcal{S} := \prod_{r=1}^n \mathcal{S}_r$, 那么 \mathbf{X} 是取值于 $S := \prod_{r=1}^n S_r$ 的随机变量, 称 \mathbf{X} 的分布为 X_1, \dots, X_n 的联合分布. 对任意 $1 \leq k < n$, $1 \leq i_1 < \dots < i_k \leq n$, 称 $(X_{i_1}, \dots, X_{i_k})$ 为 \mathbf{X} 的一个 k 维边缘, 称其联合分布为 \mathbf{X} 的一个 k 维边缘分布. 若联合分布等于所有 1 维边缘分布的乘积, 即, 对任意 $A_1 \in \mathcal{S}_1, \dots, A_n \in \mathcal{S}_n$,

$$P(X_1 \in A_1, \dots, X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n),$$

则称 X_1, \dots, X_n 相互独立. 假设 X_1, X_2, \dots 是一列随机变量.

例0.4.1 (随机向量的相互独立性). 注意到随机向量本质上是取值于乘积空间的一个随机变量, 上面的 X_i 还可以取为随机向量. 具体地, 假设 $n \geq 2$, 对任意 $i = 1, \dots, n$, $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,m_i})$ 是一个 m_i 维随机向量, 取值于 $S_i = \prod_{r=1}^{m_i} S_{i,r}$, 其上的 σ 代数为 $\mathcal{S}_i = \prod_{r=1}^{m_i} \mathcal{S}_{i,r}$. 那么, $\mathbf{X}_1, \dots, \mathbf{X}_n$ 相互独立当且仅当若对任意 $A_1 \in \mathcal{S}_1, \dots, A_n \in \mathcal{S}_n$,

$$P(\mathbf{X}_1 \in A_1, \dots, \mathbf{X}_n \in A_n) = P(\mathbf{X}_1 \in A_1) \cdots P(\mathbf{X}_n \in A_n).$$

假设 I 是一个指标集. 假设对任意 $\alpha \in I$, $(S_\alpha, \mathcal{S}_\alpha)$ 是可测空间, X_α 是取值于 S_α 的随机变量. 于是, 我们得到一族随机变量 $\mathbf{X} := \{X_\alpha : \alpha \in I\}$. 取 $\mathcal{S} := \prod_{\alpha \in I} \mathcal{S}_\alpha$, 那么 \mathbf{X} 事实上是取值于 $S = \prod_{\alpha \in I} S_\alpha$ 的随机变量. 对任意 $n \geq 1$, $\alpha_1, \dots, \alpha_n \in I$ 两两不等, $(X_{\alpha_1}, \dots, X_{\alpha_n})$ 都是 \mathbf{X} 的一个有限维边缘, 称其分布为 \mathbf{X} 的有限维联合分布. 假设 $\mathbf{Y} := \{Y_\alpha : \alpha \in I\}$ 也是取值于 S 的随机变量, 若它的任意有限维联合分布都与 \mathbf{X} 的相同, 即

$$(X_{\alpha_1}, \dots, X_{\alpha_n}) = (Y_{\alpha_1}, \dots, Y_{\alpha_n}), \quad \forall n \geq 1, \alpha_1, \dots, \alpha_n \in I,$$

则 \mathbf{Y} 与 \mathbf{X} 同分布, 即, $\mathbf{Y} \stackrel{d}{=} \mathbf{X}$.

假设 $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}$ 是 n 族随机变量, 如果对 $\mathbf{X}^{(i)}$ 的任意有限维边缘 \mathbf{X}_i , $i = 1, \dots, n$, 均有 $\mathbf{X}_1, \dots, \mathbf{X}_n$ 相互独立, 则称 $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}$ 相互独立. 假设 $\mathbf{X}_1, \mathbf{X}_2, \dots$ 是一列随机变量族. 若对任意 $n \geq 2$, $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}$ 相互独立, 则称 $\mathbf{X}_1, \mathbf{X}_2, \dots$ 相互独立. 若有限或可列个随机变量相互独立, 并且分布都相同, 则称它们独立同分布(independent and identically distributed).

例0.4.2. 假设 $\mathbf{X} := \{X_\alpha : \alpha \in I\}$ 与 $\mathbf{Y} := \{Y_\beta : \beta \in J\}$ 是两族随机变量, 状态空间及其上的 σ 代数分别为 $(S_\alpha, \mathcal{S}_\alpha)$, $\alpha \in I$ 与 $(S_\beta, \mathcal{S}_\beta)$, $\beta \in J$. 那么, 它们相互独立当且仅当对任意 $n, m \geq 1$, $\alpha_1, \dots, \alpha_n \in I$, $\beta_1, \dots, \beta_m \in J$, $A_1 \in \mathcal{S}_{\alpha_1}, \dots, A_n \in \mathcal{S}_{\alpha_n}$, $B_1 \in \mathcal{S}_{\beta_1}, \dots, B_m \in \mathcal{S}_{\beta_m}$, 记

$$A = \{X_{\alpha_1} \in A_1, \dots, X_{\alpha_n} \in A_n\}, \quad B = \{Y_{\beta_1} \in B_1, \dots, Y_{\beta_m} \in B_m\}.$$

则 $P(AB) = P(A)P(B)$.

例0.4.3 (离散型). 给定整数 $n \geq 2$. 假设 X_1, \dots, X_n 是 Ω 上的 n 个离散型随机变量, 分别取值于 S_1, \dots, S_n . 则称 $\mathbf{X} = (X_1, \dots, X_n)$ 为一个 n 维离散型随机向量. 此时, X_1, \dots, X_n 相互独立当且仅当对任意 $i_1 \in S_1, \dots, i_n \in S_n$,

$$P(X_1 = i_1, \dots, X_n = i_n) = P(X_1 = i_1) \cdots P(X_n = i_n).$$

命题0.4.4. 假设 X, Y 是两个离散型随机变量, 分别取遍 S_1, S_2 . 如果

$$P(X = i, Y = j) = \mu_i P(Y = j), \quad \forall i \in S_1, j \in S_2,$$

那么, X 与 Y 相互独立, 且 $X \sim \mu$.

证: 上式两边对 j 求和, 我们推出 $X \sim \mu$, 再将其再代入上式便知 X 与 Y 相互独立. \square

例0.4.5 (连续型). 称 \mathbb{R} 上由所有左开右闭的区间 $(a, b]$, $-\infty < a < b < \infty$ 生成 σ 代数为波莱尔(Borel) σ 代数, 记为 \mathcal{B} . 于是 $(\mathbb{R}, \mathcal{B})$ 为可测空间. 假设 (Ω, \mathcal{F}) 是可测空间, $X : \Omega \rightarrow \mathbb{R}$. 若存在函数 $\rho : \mathbb{R} \rightarrow \mathbb{R}$ 使得:

$$P(a < X \leq b) = \int_a^b \rho(x) dx, \quad a, b \in \mathbb{R}, \quad a < b,$$

则称 X 是连续型随机变量, 称 ρ 为 X 的概率密度函数, 简称密度, 记为 p_X . 上面的 ρ 满足:

$$\rho(x) \geq 0, \quad x \in \mathbb{R} \quad \text{且} \quad \int_{-\infty}^{\infty} \rho(x) dx = 1.$$

称满足上式的函数 ρ 为 \mathbb{R} 上的一个概率密度函数, 简称密度. 于是, 若 X 是连续型随机变量, 则 p_X 是密度.

一般地, 假设 $\mathbf{X} = (X_1, \dots, X_n)$ 是一个维随机向量, 其中的每个 X_i 都取值于 \mathbb{R} . 如果存在函数 $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ 使得: 对任意 $a_i < b_i, i = 1, \dots, d$, 记 $D = (a_1, b_1] \times \cdots \times (a_d, b_d]$, 则

$$P(\mathbf{X} \in D) = \int_D p(x_1, \dots, x_n) dx_1 \cdots dx_n,$$

那么称 \mathbf{X} 为 d 维连续型随机向量, 称 ρ 为 X_1, \dots, X_n 的联合(概率)密度, 记为 $p_{\mathbf{X}}$. 此时, X_1, \dots, X_d 都是连续型, 它们相互独立当且仅当

$$p_{\mathbf{X}}(x_1, \dots, x_d) = p_{X_1}(x_1) \cdots p_{X_d}(x_d), \quad \forall x_1, \dots, x_d \in \mathbb{R}.$$

§0.5 条件概率、条件分布与条件期望

假设 (Ω, \mathcal{F}, P) 为一个概率空间. 假设 $P(A) > 0$. 称

$$\frac{P(AB)}{P(A)}$$

为在 A 发生的条件下, 事件 B 的条件概率, 记为 $P(B|A)$ 或 $P_A(B)$.

注0.5.1. 假设 $P(A) > 0$. $P_A(\cdot) : \mathcal{F} \rightarrow \mathbb{R}$, $B \mapsto P(B|A)$ 是一个概率.

假设 X 是离散型随机变量, 取值于 S . 若

$$P(X = i|A) = \mu_i, \quad \forall i \in S,$$

则称在 A 发生的条件下, X 服从分布 μ . 进一步, 假设 Y 也是离散型随机变量, 取值于 \hat{S} . 若

$$P(X = i, Y = j|A) = P(X = i|A)P(Y = j|A), \quad \forall i \in S, j \in \hat{S}.$$

则称在事件 A 发生的条件下, X 与 Y 相互独立.

假设 $S \subseteq \mathbb{R}$. 如果

$$\sum_{i \in S} i P(X = i)$$

绝对收敛, 则称 X 的期望存在, 称上式为 X 的期望, 记为 EX .

假设 X 的期望存在. 进一步, 假设 Y 是离散型随机变量, 状态空间为 \hat{S} . 对任意 $j \in \hat{S}$, 令

$$E(X|Y = j) := \sum_{i \in S} i P(X = i|Y = j).$$

记 $\varphi : \hat{S} \rightarrow \mathbb{R}$, $j \mapsto E(X|Y = j)$, 称 $\varphi(Y)$ 为 X 关于 Y 的条件期望, 记为 $E(X|Y)$. 需要特别注意的是, 条件期望 $E(X|Y)$ 是一个随机变量, 它是 Y 的函数.

假设 X 是连续型的随机变量. 如果

$$\int_{-\infty}^{\infty} x p_X(x) dx$$

绝对可积, 则称 X 的期望存在, 称上式为 X 的期望, 记为 EX . 假设 X 的期望存在. 进一步, 假设 (X, Y) 是 2 维连续型的随机向量, 联合密度为 $p_{X,Y}$. 固定 $y \in \mathbb{R}$. 假设 $p_Y(y) > 0$. 称

$$p_{X|Y}(x|y) := \frac{p_{X,Y}(x,y)}{p_Y(y)}, \quad x \in \mathbb{R}$$

为在 $Y = y$ 的条件下, X 的条件密度. 令

$$\varphi : \mathbb{R} \rightarrow \mathbb{R}, \quad y \mapsto E(X|Y = y) := \int_{-\infty}^{\infty} x p_{X|Y}(x|y) dx.$$

称 $\varphi(Y)$ 为 X 关于 Y 的条件期望, 记为 $E(X|Y)$. 对于上述离散型和连续型的两种情形, 下面的重期望公式均成立.

命题0.5.2 (重期望公式). 若 X 的期望存在, 则 $E(X|Y)$ 的期望存在, 且

$$EX = EE(X|Y).$$

§0.6 随机过程

假设 (Ω, \mathcal{F}, P) 为概率空间, (S, \mathcal{S}) 为可测空间, I 为指标集. 将 I 视为时间参数, 称一族取值于 S 的随机变量 $\mathbf{X} = \{X_\alpha : \alpha \in I\}$ 为一个随机过程(stochastic process). 当 $I = \mathbb{Z}_+ = \{0, 1, 2, \dots\}$ 或 $\{0, 1, \dots, n\}$ 时, 称 \mathbf{X} 为离散时间参数的随机过程; 当 $I = \mathbb{R}_+ = [0, \infty)$ 或区间 $[a, b]$ 时, 称 \mathbf{X} 为连续时间参数的随机过程. 考虑乘积空间

$$(S^I, \mathcal{S}^I) := \left(\prod_{\alpha \in I} S_\alpha, \prod_{\alpha \in I} \mathcal{S}_\alpha \right),$$

称 S^I 中的元素为一条“轨道”, 那么, 随机过程 \mathbf{X} 本质上是一条随机轨道, 因为它是取值于 S^I 的随机变量. \mathbf{X} 的分布称为轨道分布. 在本书中, 我们将介绍三类随机过程: 马氏链(一个离散时间参数的离散型随机变量族), 跳过程(一个连续时间参数的离散型随机变量族), 布朗运动(一个连续时间参数的连续型随机变量族). 在这三类过程中, 随机过程的轨道分布都被 \mathbf{X} 的所有的有限维联合分布完全确定, 这一结论的证明超出了本书的要求范围, 有兴趣的同学可以参阅 [1] 中的第五章, 或附录中的定理 4.0.14.

§0.7 建议阅读书目

也许放前言好一点?

[11] 是为美国大学生写的课外读物. [25] 与 [16] 是当代欧洲的大学生课本, 很有借鉴意义. [21] 的内容更深入一些, 适合做为大学生科研的出发点.

随机游动是马氏链的一个特例. 与一般马氏链相比, 人们对随机游动的认识更加深入、丰富、完整. 计划深入学习的同学可以参阅 [13] 的第三章. [27] 是这方面的权威参考书之一. 如果对群上的随机游动感兴趣, 则可参阅 [10] 和 [29].

§0.8 记号

也许放最后好一点?

\mathbb{R} : 实数集, $\mathbb{R}_+ = [0, \infty)$ 非负实数集.

\mathbb{Z} : 整数集, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ 非负整数集.

$A \cup B$: 集合 A 与 B 的并集; $A \cap B$ 或 AB : 集合 A 与 B 的交集.

$x \wedge y = \min\{x, y\}$, $x \vee y := \max\{x, y\}$, $x, y \in \mathbb{R}$.

$\lfloor \cdot \rfloor$ 表示实数的整数部分, 即, 对任意 $x \in \mathbb{R}$, $\lfloor x \rfloor$ 表示不超过 x 的最大的整数.

$\|\vec{x}\|$: 假设 $d \geq 1$, 对任意 d 维向量 $\vec{x} \in \mathbb{R}^d$, $\|\vec{x}\|$ 表示 \vec{x} 在 \mathbb{R}^d 中的欧氏模, 即, 若 $\vec{x} = (x_1, \dots, x_d)$, 则 $\|\vec{x}\| = \sqrt{x_1^2 + \dots + x_d^2}$.

第一章 马氏链

§1.1 定义与例子

一、定义.

马尔可夫链(Markov Chain), 简称马氏链, 是最基本的随机过程之一. 它十分简单, 且适用面广. 本书只介绍时齐的马氏链, 它刻画了一个在非空可数集 S 上运动的粒子, 称 S 为状态空间, 称 S 中的元素为状态, 它也是粒子所处的位置, 记为 i, j, \dots . 为了直观地理解马氏链, 我们先引入骰子的概念. 在一个抽象的骰子中, 每一个面对应着 S 中的一个状态, 因此它的所有投掷结果就对应着集合 S . 对任意 $i \in S$, 在位置 i 上放一个骰子, 它投到 j 的概率为 p_{ij} , $j \in S$. 然后, 我们可以将时齐马氏链理解为下面的数学模型. 让一个粒子在 S 中按照下面的运动规则一步一步地跳跃: 如果粒子位于位置 i , 那么我们(独立地)投掷 i 处的骰子, 根据投掷的结果让粒子跳跃, 具体地, 若投到位置 j , 则让粒子下一步跳至 j . 对任意 $n \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$, 将粒子 n 时刻(即, 第 n 步)所处的位置记为 X_n . 则粒子运动的轨道就是 $\{X_n : n \geq 0\}$, 简记为 $\{X_n\}$.

投一次骰子的随机试验会产生 S 上的一个分布, 反过来, S 上的一个分布也对应着投一次骰子的随机试验. 对任意 S 上的分布 μ , 我们将 μ 视为一个行向量, 它的第 j 个分量就是 μ_j . 例如, 若 $S = \{1, \dots, n\}$, 则 $\mu = (\mu_1, \dots, \mu_n)$. 位置 i 上的骰子对应的分布为 $\{p_{ij} : j \in S\}$, 它刻画的是位置 i 上的粒子下一步所在随机位置的分布. 将其视为行向量并放在矩阵的第 i 行, 我们便得到一个矩阵 $\mathbf{P} = (p_{ij})_{S \times S}$, 它的第 i 行、第 j 列的元素就是 p_{ij} . 它符合如下定义.

定义1.1.1. 假设

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

将 p_{ij} 作为第 i 行第 j 列的元素, 得到矩阵 $\mathbf{P} = (p_{ij})_{S \times S}$. 称 \mathbf{P} 为 S 上的一个转移概率矩阵, 简称转移矩阵. 称 p_{ij} 为从 i 到 j 的转移概率(*transition probability*).

注1.1.2. 我们在交代转移矩阵时, 只需提及转移概率中的非零项, 未提及的转移概率默认为 0. 有时, 为了记号更清晰, 我们也将 p_{ij} 记为 $p_{i,j}$.

在上面描述的运动规则下, 只要粒子在某时刻 n 位于位置 i , 那么, 无论它前 n 步是如何跳到位置 i 的, 也就是无论它前 $n - 1$ 步依次经历过哪些位置, 我们都知道, 它下一步跳到位置 j 的概率都是 p_{ij} , 因为我们是独立地投掷位置 i 处的投骰子. 这一性质成为马氏链的定义中的关键性要求, 用条件概率的语言写出来就是下面的 (1.1.1). 假设 S 是一个可数集, $\{X_n : n \geq 0\}$ 是取值于 S 的离散时间参数的随机过程.

定义1.1.3. 若转移矩阵 \mathbf{P} 使得: $\forall n \geq 0, \forall i_0, \dots, i_{n-1}, i, j \in S$,

$$P(X_{n+1} = j | X_n = i, X_0 = i_0, \dots, X_{n-1} = i_{n-1}) = p_{ij}, \quad (1.1.1)$$

则称 $\{X_n\}$ 是 S 上的一个(时齐的)马尔可夫链(Markov chain), 简称马氏链. 也称 $\{X_n\}$ 是马氏的. 此时, 也称 \mathbf{P} 为 $\{X_n\}$ 的转移矩阵.

注1.1.4. 鉴于 $P(B|A)$ 在 $P(A) = 0$ 时没有定义, 因此, (1.1.1) 成立的意思是: 当等式两边有意义时, 等式成立. 换句话说, 只有在

$$P(X_n = i, X_0 = i_0, \dots, X_{n-1} = i_{n-1}) > 0,$$

时才需验证 (1.1.1) 是否成立. 下文类似, 只要出现验证条件概率的等式, 都是指仅在该条件概率有意义的时候才验证等号成立.

注1.1.5. 由全概公式可以推出: 对任意 $n \geq 0, i_0, \dots, i_{n-1}, i, j \in S$,

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

更一般地, 上式被称为马氏性, 称满足上式的随机过程 $\{X_n\}$ 为马氏链. 当上式右边依赖于 n 时, 称该马氏链是非时齐的.

二、例子.

例1.1.6 (随机游动及其步长). 考虑一个在 \mathbb{Z} 上运动的粒子. 它每一次等可能地向左或向右移动一步. 可以认为每个位置 i 上放置了一个骰子, 它投到 $i+1$ 的概率和投到 $i-1$ 的概率都是 $\frac{1}{2}$, 因此, $S = \mathbb{Z}$,

$$p_{i,i+1} = p_{i,i-1} = \frac{1}{2}, \quad \forall i \in \mathbb{Z}.$$

我们也可以用独立地抛一枚公平的硬币的随机试验来刻画这个模型, 就是每次抛到正面则让粒子往右走(位置加1), 抛到反面则让粒子往左走(位置减1). 我们还可以用随机变量序列来进行刻画, 假设 ξ_1, ξ_2, \dots 独立同分布,

$$P(\xi_1 = 1) = P(\xi_1 = -1) = \frac{1}{2}.$$

假设 S_0 为取整数值的随机变量, 且它独立于 ξ_1, ξ_2, \dots . 令

$$S_n := S_0 + \xi_1 + \dots + \xi_n = S_{n-1} + \xi_n.$$

那么, $\{S_n : n \geq 0\}$ 为从 S_0 出发的(一维)简单随机游动.

一般地, 令 $S = \mathbb{Z}^d$. 设 ξ_1, ξ_2, \dots 取值于 \mathbb{Z}^d , 独立同分布, 则称如上定义的 $\{S_n : n \geq 1\}$ 为 d 维随机游动, 称 ξ_1, ξ_2, \dots 为步长, 称其共同的分布为步长分布. 将第 i 个坐标取 1, 其他坐标取 0 的 d 维向量记为 e_i . 若

$$P(\xi_1 = e_i) = P(\xi_1 = -e_i) = \frac{1}{2d}, \quad i = 1, \dots, d,$$

则称 $\{S_n\}$ 为 d 维简单随机游动.

例1.1.7 (两状态马氏链). 状态空间为 $S = \{0, 1\}$, 转移概率为

$$p_{01} = 1 - p_{00} = p, \quad p_{10} = 1 - p_{11} = q,$$

其中 $0 < p, q < 1$. 可以说两个状态的马氏链是最简单的马氏链. 当年马尔可夫提出马氏链这一概念时, 他关心的是新模型的数学性质, 并不在意该模型是否有用. 不过他还是给出一个具体的例子: 普希金诗歌Eugenie Onegin 里某一段(长度约为 12 页)中元音字母和辅音字母交替出现的规律. 把元音字母记为 0, 辅音字母记为 1, 他还估计了 p_{01} 和 p_{10} .

例1.1.8 (Ehrenfest模型). 设有 N 个不同编号的球和 A, B 两个纸箱. 首先, 随机地把这 N 个球分装在两个纸箱中, 其中纸箱 A 中有 X_0 个球. 然后, 每次独立地以 $1/N$ 的概率选定某一个球, 把它从其所在的纸箱中拿出并放到另一个纸箱中. 将 n 次操作后纸箱 A 中球的个数记为 X_n . 则 $\{X_n\}$ 是一个马氏链, 状态空间为 $S = \{0, 1, 2, \dots, N\}$, 转移概率为:

$$p_{i,i+1} = 1 - \frac{i}{N}, \quad 0 \leq i \leq N-1, \quad p_{i,i-1} = \frac{i}{N}, \quad 1 \leq i \leq N.$$

除了用表达式给出转移矩阵, 我们还可以用概率转移图来给出转移矩阵. 在概率转移图中, 每个顶点代表一个状态. 如果 $p_{ij} > 0$, 那么我们画一条从 i 到 j 的有向边, 将 p_{ij} 标记在这条边上. 例如, Ehrenfest 模型的概率转移图如下.

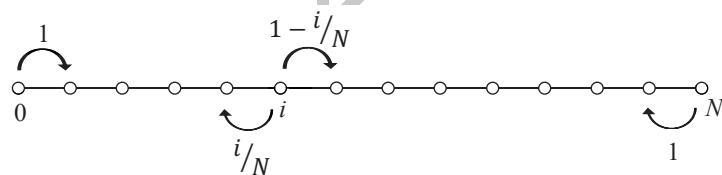


图 1.1 Ehrenfest 模型的转移概率图.

例1.1.9 (更新过程). 假设 $L = L_1, L_2, \dots$ 是独立同分布的取正整数的随机变量列, L_r 表示第 r 个灯泡的寿命. 假设时刻 0 放上第一个灯泡, 灯泡坏的时候立刻换上一个新灯泡. 则 $\{0, L_1, L_1 + L_2, \dots\}$ 表示换灯泡的时间, 它们被称为更新时刻. 如下图所示.

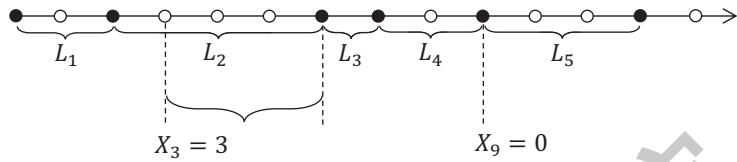


图 1.2 更新时刻用实心圆标记, 其他时刻用空心圆圈标记.

记 X_n 为在时刻 n 正在使用的那个灯泡的余寿, 即它的总寿命减去已经使用的时间. 在图上, X_n 就是时刻 n 与它右边最近的更新时刻之间的距离. 例如, 图中的 $X_3 = 3$. 当 n 为更新时刻, $X_n = 0$, 它表示被换下的旧灯泡的余寿. $\{X_n\}$ 被称为更新过程, 它是一个马氏链, $S = \mathbb{Z}_+$, 转移概率为:

$$p_{i,i-1} = 1, \forall i \geq 1; \quad p_{0,i} = P(L = i+1), \forall i \geq 0,$$

概率转移图如下图所示.

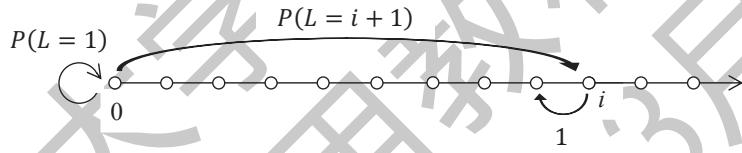


图 1.3 更新过程的转移概率图.

例1.1.10 (图上的随机游动). 假设 V 是一个非空的可数集. 称 V 中元素为顶点或结点. E 是以 V 的某些点对为元素的集合, 若 $i \neq j$ 且 $\{i, j\} \in E$, 则在 i 与 j 之间连一条(无向的)边, 并且称 i 与 j 相邻, 或称它们是邻居. 称顶点 i 的邻居的数目为 i 的度, 记为 d_i . 称 $G = (V, E)$ 为一个简单图. 假设 $i \neq j$, 若 $n \geq 1, i_1, \dots, i_{n-1}$ 使得 i_{r+1} 与 i_r 相邻, $r = 0, \dots, n-1$, 其中 $i = i_0$, $j = i_n$, 则 (i_0, i_1, \dots, i_n) 为一条连接 i, j 的路径, 称 n 为该路径的长度. 若对任意 $i \neq j$ 都存在连接 i, j 的路径, 则称 G 是连通的. 例如, 图 1.4 中的蜂窝图和三角形图是连通的简单图.

假设 G 是连通的简单图, G 上的随机游动是一个马氏链, 状态空间为 $S = V$, 转移概率为:

$$p_{ij} = \begin{cases} 1/d_i, & \text{若 } j \text{ 与 } i \text{ 相邻,} \\ 0, & \text{其它,} \end{cases}$$

粒子每一步在它所处顶点的所有邻居中随机等可能地选择一个, 然后跳过去. 特别地, 取 $V = \mathbb{Z}^d$, 当且仅当两个顶点之间的欧氏距离为 1 时在它们之间连一条边. 那么我们得到一个图, 称之为 d 维格点. 在不引起歧义的情况下, 我们仍然用 \mathbb{Z}^d 表示这个图, 该图上的随机游动就是例 1.1.6 中介绍的 d 维简单随机游动.

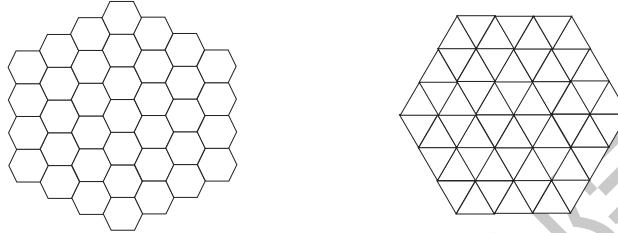
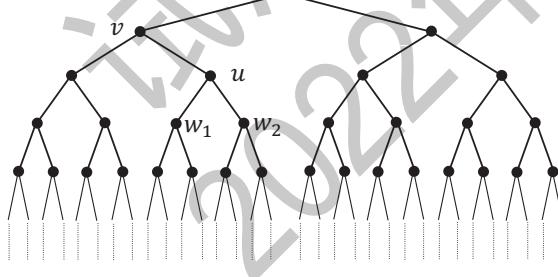


图 1.4 左图为蜂窝图, 右图为三角形图

若图中没有圈, 即任意两点之间仅有一条路径相连, 则称该图为树, 记为 \mathbb{T} . 选定 \mathbb{T} 中的一点作为根点, 将其记为 o , 我们便得到一个有根点的树. 若根点的度是 d , 其余每个顶点的度均为 $d + 1$, 则称其为规则树, 记为 \mathbb{T}^d . 例如, \mathbb{T}^2 如图 1.5 所示. 对任意顶点 $u \neq o$, 将连接 u 与 o 的路径的长度记为 $|u|$. 若 $|u| = n$, 则称 u 位于树的第 n 层. 若 u 与根点 o 相邻, 则称 u 为 o 的儿子, 此时, $|u| = 1$. 当 $u \neq o$ 时, 它有且仅有一个人邻居 v 满足 $|v| = |u| - 1$, 称 v 为 u 的父结点, 称 u 的其他邻居为 u 的子结点, u 的任意子结点 w 都满足 $|w| = |u| + 1$. 例如, 图中的 w_1, w_2 是 u 的子结点. 假设 $\{X_n\}$ 是 \mathbb{T}^d 上的随机游动. 不难看出,

$$P(X_{n+1} = X_n \text{ 的父结点} | X_n \neq o) = \frac{1}{d+1}, \quad P(X_{n+1} = X_n \text{ 的子结点} | X_n \neq o) = \frac{d}{d+1}.$$

图 1.5 规则树 \mathbb{T}^2

三、有限维联合分布

如前所述, 马氏链等价于一个在 S 中运动的粒子, 其运动规则通过转移矩阵 \mathbf{P} 给定. 称 X_0 的分布为初分布. 给了初分布 μ , 我们就可以根据 (1.1.1) 计算出轨道的任意有限维联合分布. 具体地, 对任意 $n \geq 0$ 以及 $i_0, \dots, i_n \in S$,

$$P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mu_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}. \quad (1.1.2)$$

反过来, 如果上式对任意 $n \geq 0$ 以及 $i_0, \dots, i_n \in S$ 均成立, 那么根据条件概率的定义不难验证 $\{X_n\}$ 是以 \mathbf{P} 为转移矩阵的马氏链.

命题1.1.11 (马氏性). 取定 $n \geq 1, i \in S$. 令 $Y_m = X_{n+m}, \forall m \geq 0$. 那么, 在 $X_n = i$ 的条件下, $\{Y_m\}$ 是从 i 出发的以 \mathbf{P} 为转移矩阵的马氏链, 并且, 它与 $\vec{Z} = (X_0, \dots, X_{n-1})$ 相互独立.

证: 将条件概率 $P(\cdot | X_n = i)$ 简记为 $\hat{P}(\cdot)$. 对任意 $m \geq 0, j_0, \dots, j_m \in S$, 以及任意 $i_0, \dots, i_{n-1} \in S$, 记

$$\vec{Y} = (Y_0, \dots, Y_m), \quad \vec{j} = (j_0, \dots, j_m), \quad \vec{i} = (i_0, \dots, i_{n-1}).$$

当 $j_0 \neq i$ 时, $\{\vec{Y} = \vec{j}\}$ 与 $\{X_n = i\}$ 不交; 当 $j_0 = i$ 时, 根据 (1.1.2),

$$\begin{aligned} & P(\vec{Z} = \vec{i}, X_n = i, \vec{Y} = \vec{j}) \\ &= P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i, X_{n+1} = j_1, \dots, X_{n+m} = j_m) \\ &= \mu_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}} p_{i_{n-1} i} \times p_{j_0 j_1} \cdots p_{j_{m-1} j_m}. \end{aligned}$$

再根据 (1.1.2),

$$\mu_{i_0} p_{i_0 i_1} \cdots p_{i_{n-2} i_{n-1}} p_{i_{n-1} i} = P(\vec{Z} = \vec{i}, X_n = i) = P(X_n = i) \hat{P}(\vec{Z} = \vec{i}).$$

因此, 无论 j_0 是否等于 i , 我们总有

$$P(\vec{Z} = \vec{i}, X_n = i, \vec{Y} = \vec{j}) = P(X_n = i) \hat{P}(\vec{Z} = \vec{i}) \times (\mathbf{1}_{\{j_0=i\}} p_{j_0 j_1} \cdots p_{j_{m-1} j_m}).$$

上式两边同时除以 $P(X_n = i)$ 即得

$$\hat{P}(\vec{Z} = \vec{i}, \vec{Y} = \vec{j}) = \hat{P}(\vec{Z} = \vec{i}) \times \mathbf{1}_{\{j_0=i\}} \times p_{j_0 j_1} \cdots p_{j_{m-1} j_m}.$$

根据命题 0.4.4, 由 m, \vec{j}, \vec{i} 的任意性知命题成立. \square

假设 $r \geq 1, 0 \leq n_1 < \dots < n_r < n < n + m, i_1, \dots, i_r, i, j \in S$. 根据马氏性, 在已知 $\{X_n = i\}$ 的条件下, $(X_{n_1}, \dots, X_{n_r})$ 与 X_{n+m} 相互独立, 因为前者可以视为命题 1.1.11 中的过去轨道 \vec{Z} 的函数, 而后者可以视为命题 1.1.11 中的将来轨道 $\{Y_n\}$ 的函数. 于是,

$$P(X_{n+m} = j | X_n = i, X_{n_1} = i_1, \dots, X_{n_r} = i_r) = P(X_{n+m} = j | X_n = i). \quad (1.1.3)$$

根据时齐性, $P(X_{m+n} = j | X_n = i)$ 是一个不依赖于 n 的常值, 将其记为 $p_{ij}^{(m)}$. 有时为了更清晰, 也记为 $p_{i,j}^{(m)}$. 不难看出, $p_{ij}^{(0)} = \mathbf{1}_{\{i=j\}}$, $p_{ij}^{(1)} = p_{ij}$. 一般地, 我们可以按照 X_n 的取值划分样本空间, 并由全概公式与 (1.1.3) 得到如下命题.

命题1.1.12 (Chapman-Kolmogorov 等式). 对任何 $i, j \in S, m, n \geq 0$,

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

由数学归纳法, Chapman-Kolmogorov 等式蕴含着

$$p_{ij}^{(n)} = (\mathbf{P}^n)_{ij}. \quad (1.1.4)$$

定义1.1.13. 称 $p_{ij}^{(n)}$ 为从 i 到 j 的 n 步转移概率. 称矩阵 $\mathbf{P}^n = (p_{ij}^{(n)})_{S \times S}$ 为马氏链的 n 步转移(概率)矩阵.

根据 (1.1.3), 我们可以把(1.1.1) 进行如下推广: 对任意 $0 = n_0 < n_1 < \dots < n_r$, $i_0, i_1, \dots, i_r \in S$,

$$P(X_0 = i_0, X_{n_1} = i_1, X_{n_2} = i_2, \dots, X_{n_r} = i_r) = \mu_{i_0} p_{i_0 i_1}^{(m_1)} p_{i_1 i_2}^{(m_2)} \cdots p_{i_{r-1} i_r}^{(m_r)}, \quad (1.1.5)$$

其中 $m_s = n_s - n_{s-1}, \forall s$.

四、马氏链的构造.

根据马氏链的定义, 若已知现在的状态是 i , 即知道 $X_n = i$ 则下一步的随机变量 X_{n+1} 与过去的随机轨道 (X_0, \dots, X_{n-1}) 是相互独立的. 换句话说, X_{n+1} 的分布被 X_n 完全确定. 由于再下一步 X_{n+2} 的分布又可以由 X_{n+1} 的取值完全确定, 因此我们可以利用递归的办法来构造马氏链.

将 S 中的状态编号, 不妨设 $S = \{1, 2, \dots, N\}$ 或 $\{1, 2, \dots\}$. 假设 \mathbf{P} 为一个转移矩阵, 即 $p_{ij} \geq 0, i, j \in S$ 且 $\sum_{j \in S} p_{ij} = 1$; 假设 μ 为 S 上的一个分布, 即 $\mu_i \geq 0, i \in S$ 且 $\sum_{i \in S} \mu_i = 1$.

取独立同分布的随机变量序列 U_0, U_1, U_2, \dots , 使得 U_1 服从区间 $(0, 1)$ 上的均匀分布 $U(0, 1)$. 取 $g : (0, 1] \rightarrow S$ 使得: 对任意 $i \geq 1$,

$$g(u) := i, \quad \forall u \in \left(\sum_{r=1}^{i-1} \mu_j, \sum_{r=1}^i \mu_j \right].$$

那么,

$$P(g(U_0) = i) = P \left(\sum_{r=1}^{i-1} \mu_j < U \leq \sum_{r=1}^i \mu_j \right) = \mu_i, \quad \forall i \geq 1.$$

即, $g(U_0) \sim \mu$.

对任意 $i \in S$, 取 $f(i, \cdot) : (0, 1] \rightarrow S$ 使得: 对任意 $i \geq 1$,

$$f(i, u) := j, \quad \forall u \in \left(\sum_{r=1}^{j-1} p_{ir}, \sum_{r=1}^j p_{ir} \right].$$

那么, 对任意 $n \geq 1$,

$$P(f(i, U_n) = i) = P\left(\sum_{r=1}^{j-1} p_{ir} < U \leq \sum_{r=1}^j p_{ir}\right) = p_{ij}, \quad \forall j \geq 1.$$

令 $X_0 := g(U_0)$, 并递归地定义

$$X_{n+1} := f(X_n, U_{n+1}), \quad n = 0, 1, 2, \dots$$

命题1.1.14. $\{X_n\}$ 是以 μ 为初分布, 以 \mathbf{P} 为转移矩阵的马氏链.

补充知识.

马氏链是这样一个模型: 我们在空间 S 的每一个位置上放一个骰子, 按某个初分布 μ 将粒子放置在 S 中的某个随机位置, 然后让粒子在 S 中一步一步地跳跃. 这个模型本质上在做一个很大的随机试验, 试验结果 ω , 也就是样本, 其实就是粒子运动的轨道 (x_0, x_1, \dots) , 其中 x_n 是粒子在 n 时刻的位置. 因此这个随机试验的样本空间 Ω 其实就是轨道空间, 即, 由所有轨道组成的集合 $S^{\mathbb{Z}_+}$:

$$\Omega = S_+^{\mathbb{Z}} = \{\omega = (x_0, x_1, \dots) : x_n \in S, \forall n \in \mathbb{Z}_+\}.$$

鉴于 $\Omega = S^{\mathbb{Z}_+}$, 它又被称为样本轨道空间, 其中的 ω 又被称为样本轨道.

粒子在时刻 n 的位置 X_n 就是轨道的第 n 个坐标, 即

$$X_n(\omega) = x_n, \quad \forall \omega = (x_0, x_1, \dots).$$

这样定义的 $\{X_n\}$ 被称为样本轨道空间 Ω 上的坐标过程. 进一步, 因为 ω 其实就是 $\vec{X}(\omega)$, 所以, Ω 上的任意随机变量都可以视为 \vec{X} 的函数.

我们可以将马氏链 $\{X_n\}$ 理解为一条随机轨道, 即, 取值于轨道空间 $S^{\mathbb{Z}_+}$ 的一个抽象的随机变量. 也可以将 $\{X_n\}$ 理解为一个无穷随机向量 $\vec{X} = (X_0, X_1, X_2, \dots)$. 就像对于一个随机变量 X 而言, 最重要的是 X 的分布一样, 对于马氏链来说, 如果将其视为一条随机轨道, 最重要的就是这条随机轨道的分布, 即轨道分布. 值得注意的是, 尽管 S 是可数的, 但是, 所有轨道组成的集合 $S^{\mathbb{Z}_+}$ 却是不可数的, 因此, $\{X_n\}$ 往往不再是离散型, 谈论它恰好取到某条特定轨道的概率是没有意义的, 因为这往往是0. 从 \vec{X} 的角度看, 轨道分布指的是这个无穷随机向量的联合分布. 附录中的定理 4.0.14 指出, 轨道分布 \vec{X} 的任意有限步联合分布完全决定. 换句话说, 只要对所有的 $n \geq 0, i_0, \dots, i_n \in S$ 给出 $\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\}$ 的概率即可. 这个事件, 用样本 ω 来表达就是

$$\{\omega = (x_0, x_1, \dots) : x_0 = i_0, \dots, x_n = i_n\}.$$

这样的事件被称为柱集, 记为 $C_{i_0 i_1 \dots i_n}$. 于是, 马氏链其实在说, Ω 上的概率 P 满足:

$$P(C_{i_0 i_1 \dots i_n}) = \mu_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}, \quad \forall n \geq 0, i_0, \dots, i_n \in S. \quad (1.1.6)$$

而附录中的定理 4.0.14 则表明满足上述条件的概率 P 是存在唯一的.

基于上述讨论, 我们以后在有需要时可以不妨假设考虑的就是上面的随机试验, 即, Ω 是样本轨道空间, P 是 Ω 上满足 (1.1.6) 的那个唯一的概率, $\{X_n\}$ 是 Ω 上的坐标过程.

其次, 马氏性的本意是在已知现在状态的条件下, 过去与将来独立, 这种独立性体现在上述证明中的 U_1, U_2, \dots 之间相互独立上. 但是, 马氏性其实并不要求粒子每一步跳跃的规律必须一样, 即 p_{ij} 可以依赖于时间的, 也就是说, 在上面的证明中, 每次用于迭代的函数 f 可以是不一样的. 如果 p_{ij} 不依赖于时间, 这就是我们本章讨论的情形, 我们称之为时齐的马氏链; 如果 p_{ij} 依赖于时间, 我们称之为非时齐的马氏链. 在非时齐的马氏链中, 在任意时刻都有一个转移矩阵, 换句话说, 我们有一列转移矩阵 $\mathbf{P}_n = (p_{n;i,j})_{S \times S}, \forall n \geq 1$, 以 μ 为初分布的马氏链有限步轨道的联合分布为

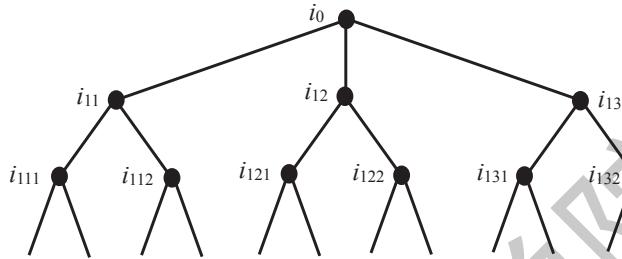
$$P(X_0 = i_0, \dots, X_n = i_n) = \mu_{i_0} p_{1;i_0 i_1} p_{2;i_1 i_2} \cdots p_{n;i_{n-1} i_n}.$$

这个式子是 (1.1.2) 的推广. 事实上, 在第 §1.2 节中, 若原过程 $\{X_n\}$ 的初分布不是不变分布, 其逆过程 $\{Y_m : 0 \leq m \leq N\}$, 其中 $Y_m := X_{N-m}$, 就是一个非时齐的马氏链.

最后, 马氏链是可数个随机变量, 在刚才的叙述中, 我们也看到, 逆过程 $\{Y_m : 0 \leq m \leq N\}$ 指有限个随机变量. 一般地, 一个离散时间参数、离散状态空间上的随机过程指的是有限或可数个随机变量放在一起. 它有一个取值空间 S , 就是本章所说的状态空间; 还有一个时间参数空间 T , 它在本章通常被取为 $\{0, 1, 2, \dots\}$, 在讨论逆过程时被取为 $\{0, 1, \dots, N\}$. S 中的点被称为状态, 通常记为 i ; T 中的点被称为时间, 通常记为 n . 所谓离散状态空间是指 S 是可数集, 所谓离散时间参数是指 T 是可数集. 一般地, 我们不要求时间参数空间 I 中有全序关系. 例如: 在 $\{(n, m) : n, m \in \mathbb{Z}\}$ 的每个顶点上放一个独立的离散型随机变量, 就得到一个随机过程; 在有根点的规则树 \mathbb{T}^d 的每个顶点放一个取值于 S 的随机变量, 规则是: 在某顶点 $x \in \mathbb{T}^d$ 处取值为 $i \in S$ 的条件下, 其子代顶点上独立放一个同分布的随机变量, 取值为 j 的概率为 p_{ij} . 这样可以递归定义出一族随机变量, 它也是一个随机过程, 被称为树上的马氏场. 例如, 在 \mathbb{T}^2 的马氏场中, 前两层看到如图 §1.1 所示状态的概率为

$$\mu_{i_0} \cdot (p_{i_0 i_{11}} \cdot p_{i_0 i_{12}} \cdot p_{i_0 i_{13}}) \cdot ((p_{i_{11} i_{111}} p_{i_{11} i_{112}})) \cdot ((p_{i_{12} i_{121}} p_{i_{12} i_{122}})) \cdot ((p_{i_{13} i_{131}} p_{i_{13} i_{132}})).$$

定义 1.1.15. 设 (Ω, \mathcal{F}) 为一个概率空间, T, S 为一个可数集. 若对任意 $n \in T$, $X_n : \Omega \rightarrow S$ 满足 $\{X_n = i\} \in \mathcal{F}$, 则称 $\{X_n : n \in T\}$ 为一个随机过程. 称 (Ω, \mathcal{F}) 为其样本空间, S 为其状态空间, T 为时间参数空间.

图 1.6 \mathbb{T}^2 的马氏场

我们仍然将随机过程 $\{X_n\}$ 视为取值于 $S^T = \{(s_n)_{n \in I} : s_n \in S, \forall n \in I\}$ 的随机变量, 于是, 随机过程的本质就是它在 S^T 上的分布. 若 $|I| < \infty$, 那么随机过程就是一个有限维的随机向量, 其分布就是此随机向量的联合分布; $|I| = \infty$, 问题会变得复杂很多. 在马氏链中, 我们实际上只给出了过程在任意有限个时刻上的联合分布 $P(X_{n_1} = i_1, \dots, X_{n_r} = i_r)$, $\forall n_1, \dots, n_r \in T, \forall i_1, \dots, i_r \in S$, 我们可以用独立同分布的均匀分布随机变量构造马氏链. 但是, 在一般的过程中, 问题是: 是否在 S^T 上唯一地存在一个分布, 使得它在集合 $\{(s_n)_{n \in I} : s_{n_1} = i_1, \dots, s_{n_r} = i_r\}$ 上的值就是我们给出的 $P(X_{n_1} = i_1, \dots, X_{n_r} = i_r)$? 这个问题的答案是肯定的, 见 Kolmogorov 相容性定理(附录中的定理 4.0.14), 它已经超出了本书的要求范围, 有兴趣的读者可以参阅 [1] 第四章第3节.

习题

- 假设 X_0, X_1, X_2, \dots 独立同分布, X_0 是离散型随机变量, 分布为 μ . 证明: $\{X_n\}$ 是马氏链, 求其转移概率.
- 假设 $\{S_n\}$ 是一维简单随机游动. 令 $X_n = \max_{0 \leq k \leq n} S_k, n \geq 0$. 试问: $\{X_n\}$ 是马氏链吗? 请证明你的结论.
- 在例 1.1.9 中, 记 Y_n 为时刻 n 与它左边最近的更新时刻之间的距离, 即, 时刻 n 正在用的那个灯泡已经使用的时间, 特别地, 当 n 为更新时刻, $Y_n = 0$. 证明 $\{Y_n\}$ 是一个马氏链, 并求其转移概率. (注: 称 $\{Y_n\}$ 为一个老化过程.)
- 某篮球运动员投球成功的概率取决于他前两次的投球成绩. 如果两次都成功, 则下次投球成功的概率为 $3/4$; 如果两次都失败, 则下次投球成功的概率为 $1/2$; 如果两次投球中有一次成功一次失败, 则下次投球成功的概率为 $2/3$. (1) 试用一个马氏链来刻画该球员的连续投球, (2) 将他第 n 次投球成功的概率记为 p_n , 求 $\lim_{n \rightarrow \infty} p_n$.

5. 某车间最多可以安放 s 台机床, 如果第 n 个星期开始上班时共有 $X_n = i$ 台机床可以使用, 车间主任就订购 $s - i$ 台新机床, 可于周末到货, 下星期开始上班时投入使用. 而在这个星期又有 Y_n 台车床报废, 已知 $P(Y_n = j|X_n = i) = 1/(1+i)$, $0 \leq j \leq i$. 因此 $X_{n+1} = s - Y_n$. 于是 $\{X_n\}$ 为一个马氏链. (1) 写出其转移概率, (2) 令 $a_n = EX_n$. 试给出序列 a_0, a_1, a_2, \dots 的递推式, 并求 $\lim_{n \rightarrow \infty} EX_n$.
6. 假设某加油站给一辆车加油需要一个单位时间(比如, 五分钟). ξ_n 是第 n 个单位时间来加油的汽车数. 假设 ξ_1, ξ_2, \dots 独立同分布, 取值非负整数, $P(\xi_1 = k) = p_k$, $k \geq 0$. 在任意时刻 n , 如果加油站有车, 那么加油站为其中一辆加油(耗时一个单位时间, 然后该汽车在时刻 $n+1$ 离开加油站); 否则, 加油站什么都不做. 将 n 时刻加油站中的汽车数记为 X_n . 于是 $\{X_n\}$ 为一个马氏链. 写出其状态空间与转移概率.
7. 一个粒子在三角形的三个顶点之间跳跃, 它按顺时针方向移动的概率为 $p \in (0, 1)$, 按逆时针方向移动的概率为 $1 - p$. 试求 n 步之后该粒子恰好处在出发点的概率 p_n , 并计算 $\lim_{n \rightarrow \infty} p_n$.
8. 设 $S = \{1, 2, 3, 4\}$, 转移矩阵 \mathbf{P} 如下, (1) 试求 \mathbf{P}^2 , (2) 求 $p_{ii}^{(2n)}$, $\forall i$.

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.2 & 0.8 \\ 0.25 & 0.75 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix}.$$

9. 某数据通信系统由 n 个中继站组成, 从上一站向下一站传送信号0或1时, 接收的正确率为 p . 如用 X_0 表示初始站发出的数字. 用 X_k 表示第 k 个中继站接收到的数字. 试证: $\{X_k : 0 \leq k \leq n\}$ 是一个马氏链,

$$P(X_0 = 1|X_n = 1) = \frac{\alpha + \alpha(p - q)^n}{1 + (2\alpha - 1)(p - q)^n},$$

其中 $\alpha = P(X_0 = 1)$, $q = 1 - p$. 并请说明上述条件概率的实际意义.

10. 假设 $\{X_n\}$ 是规则树 \mathbb{T}^d 上的随机游动, 取 $Y_n = |X_n|$. 证明 $\{Y_n\}$ 是马氏链, 并求其状态空间与转移概率.

- 11* 假设 \mathbf{P} 是 S 上的转移矩阵, \hat{S} 是可数集, $f : S \rightarrow \hat{S}$ 是满射. 证明: 如果对所有 $i, i' \in S$, 若 $f(i) = f(i')$, 则

$$\sum_{j:f(j)=k} p_{ij} = \sum_{j:f(j)=k} p_{i'j}, \quad \forall k \in \hat{S},$$

那么, 对任意 S 上以 \mathbf{P} 为转移矩阵的马氏链 $\{X_n\}$ 均有 $\{f(X_n)\}$ 是 \hat{S} 上的马氏链.

12. 对于马氏链, $\forall n < m$, $\forall B_0, \dots, B_{n-1}, A \subseteq S$, $i \in S$, 证明: $P(X_{n+1} \in A | X_n = i, X_0 \in B_0, \dots, X_{n-1} \in B_{n-1}) = P(X_{n+1} \in A | X_n = i)$.
13. 证明命题 1.1.14.

14* 假设对任意 $n \geq 0$, $i \in S$, $f(n, i, \cdot) : [0, 1] \rightarrow S$ 使得对任意 $U \sim U(0, 1)$,

$$P(f(n; i, U) = j) = p_{n;i,j}, \quad j \in S.$$

(1) 证明: 对任意 $n \geq 0$, $\mathbf{P}_n = (p_{n;i,j})_{S \times S}$ 为转移矩阵. (2) 假设 X_0, U_1, U_2, \dots 相互独立, $U_n \sim U(0, 1)$, $n \geq 1$, X_0 取值于 S . 递归定义 $X_{n+1} = f(n+1; X_n, U_{n+1})$, $n = 0, 1, 2, \dots$. 那么, $\{X_n\}$ 是马氏链. (注: $\{X_n\}$ 可以为非时齐的.)

§1.2 不变分布与可逆分布

一、不变分布.

假设 $\{X_n\}$ 是以 \mathbf{P} 为转移概率的马氏链, 初始位置为 i , 即 $X_0 = i$. 此时, 我们称 $\{X_n\}$ 是从 i 出发的马氏链, 将其对应的概率记为 P_i , 即

$$P_i(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mathbf{1}_{\{i_0=i\}} \cdot p_{i_0 i_1} \cdots p_{i_{n-1} i_n},$$

如果 $\{X_n\}$ 初分布为 μ , 那么我们将其对应的概率记为 P_μ . 根据全概公式,

$$P_\mu(A) = \sum_{i \in S} \mu_i P_i(A).$$

分别将 P_i 与 P_μ 对应的数学期望记为 E_i 和 E_μ . 根据等式(1.1.2), $\{X_n\}$ 的有限长的轨道分布被其初分布 μ 和转移矩阵 \mathbf{P} 完全决定. 例如, 我们可以算出

$$P_\mu(X_1 = j) = \sum_{i \in S} P_\mu(X_0 = i, X_1 = j) = \sum_{i \in S} \mu_i p_{ij}. \quad (1.2.1)$$

现在, 我们将 S 上的分布 μ 视为一个行向量, 第 i 个分量(或, 第 i 列的分量)就是 μ_i . 类似于矩阵理论中的行向量右乘矩阵得到一个新的行向量, 我们知道 $\mu\mathbf{P}$ 还是 S 上的一个分布, 其第 j 列的分量就是 $\sum_{i \in S} \mu_i p_{ij}$. 那么, 根据上面的公式 (1.2.1), 若 $X_0 \sim \mu$, 则 $X_1 \sim \mu\mathbf{P}$.

定义1.2.1. 假设 $\pi = \{\pi_j : j \in S\}$ 为 S 上的测度. 若 π 满足如下不变方程:

$$\pi_j = \sum_{k \in S} \pi_k p_{kj}, \quad \forall j \in S,$$

则称 π 为 \mathbf{P} 的一个不变测度. 进一步, 若 π 还是分布, 则称 π 为一个不变分布 (invariant distribution). 也称 π 为以 \mathbf{P} 为转移矩阵的马氏链的不变测度或不变分布.

注1.2.2. 显然, $\pi_i \equiv 0$ 是不变测度, 它是平凡的不变测度. 若 π 是非平凡的不变测度, 并且 $\sum_{i \in S} \pi_i < \infty$, 那么, 我们总可以将其归一化, 得到一个不变分布 $\pi_i / \sum_{j \in S} \pi_j$, $i \in S$.

注1.2.3. 假设 π 为不变分布. 如果 $X_0 \sim \pi$, 那么,

$$(X_0, X_1, \dots, X_n) \stackrel{d}{=} (X_m, X_{m+1}, \dots, X_{m+n}), \quad n, m \geq 1,$$

其中, “ $\stackrel{d}{=}$ ”表示同分布. 满足上式的过程被称为平稳过程. 对于一个平稳过程, 任意时刻 m 都可以被视为时间起点. 对于马氏链而言, 因为同样的初分布带来同样的有限步轨道分布. 因此, 不变分布也被称为平稳分布(stationary distribution). 另外, 不变分布在物理中被称为稳态(steady state), 含义为: 此时系统有某种稳定性.

如果 S 有限, 不妨设 $|S| = N$. 那么, 从线性代数的观点来看, 非平凡的不变测度对应的行向量是 $N \times N$ 矩阵 \mathbf{P} 的特征值为 1 的左特征向量, 视为行向量时, 不变方程就可以改写为

$$\pi = \pi \mathbf{P}.$$

于是, 我们可以通过解线性方程组来求不变分布. 我们也可以用数学分析的知识证明不变分布存在(见本节习题8), 在后续的课程中, 我们还将从概率论的角度证明不变分布存在.

注1.2.4. 当 S 有限时, 根据线性代数中的Perron-Frobenius定理(附录中的定理 4.0.15), 非平凡的不变测度总存在, 将它归一化以后便得到不变分布.

例1.2.5 (两状态马氏链, 例 1.1.7续). $S = \{0, 1\}$, $p_{01} = p$, $p_{10} = q$, 其中 $0 < p, q < 1$. 如果写成矩阵, 第一行/列对应状态 0, 第二行/列对应状态 1, 那么, 转移矩阵 \mathbf{P} 形如:

$$\begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}.$$

此时, 不变分布 π 所满足的方程组为:

$$\begin{cases} \pi_0 p_{00} + \pi_1 p_{10} = \pi_0, \\ \pi_0 p_{01} + \pi_1 p_{11} = \pi_1. \end{cases}$$

这两个方程本质是同一个, 因为分别对它们的左右两边求和是一个恒等式 $\pi_0 + \pi_1 = \pi_0 + \pi_1$. 一般地, 如果 $S = |N|$, 那么, 不变分布的方程组 $\pi \mathbf{P} = \pi$ 本质只有 $N - 1$ 个方程. 这 $N - 1$ 个方程, 结合归一化条件 $\sum_{i \in S} \pi_i = 1$ 可解出 π . 在这个例题中, 我们解得 $\pi_0 = \frac{q}{p+q}$, $\pi_1 = \frac{p}{p+q}$.

当 S 可列或者 N 很大时, 通过解方程组 $\pi = \pi \mathbf{P}$ 来寻找不变分布的办法就不太可行了. 不过, 对于一些比较特殊的转移矩阵, 我们可以从转移概率图的结构着手来寻找不变分布. 事实上, 对 S 的任意子集 A ,

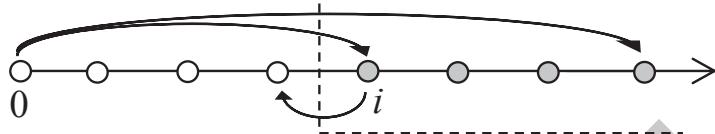
$$\sum_{i \notin A, j \in A} \pi_i p_{ij} = \sum_{i \in A, j \notin A} \pi_i p_{ij}. \quad (1.2.2)$$

上式的证明留为习题(本节习题2). 反过来, 如果上式对 S 的任意子集 A 都成立, 那么我们可以将 A 取为任意单点集 $\{i\}$, 便知 π 是不变分布. 换句话说, π 是不变分布当且仅当上式对任意 $A \subseteq S$ 成立. 这个等价条件有时候也可以帮助我们寻找不变分布.

例1.2.6 (更新过程, 例 1.1.9续). $S = \mathbb{Z}_+$, 转移概率为:

$$p_{i,i-1} = 1, \quad \forall i \geq 1; \quad p_{0,i} = P(L = i + 1), \quad \forall i \geq 0.$$

其中 L 代表(第一个)灯泡的寿命.

图 1.7 $A = \{i, i + 1, i + 2, \dots\}$.

假设 π 为不变分布. 固定 $i \geq 1$. 取

$$A = \{i, i + 1, i + 2, \dots\},$$

它即是图 1.7 中虚线画出的范围. 由于粒子从 A 中跳到 A^c 中时只能是从 i 跳到 $i - 1$, 因此 (1.2.2) 的右边等于 $\pi_i p_{i,i-1} = \pi_i$; 而粒子从 A^c 中跳到 A 中, 只能是从 0 跳至 A 中的某个状态 j , 因此 (1.2.2) 的左边等于 $\pi_0 p_{0,i} + \pi_0 p_{0,i+1} + \dots = \pi_0 P(L \geq i + 1)$. 根据 i 的任意性, π 必须满足:

$$\pi_i = \pi_0 P(L \geq i + 1), \quad \forall i \geq 1.$$

进一步, π 还须满足归一化条件 $\sum_{i=0}^{\infty} \pi_i = 1$, 即

$$1 = \pi_0 (1 + P(L \geq 2) + P(L \geq 3) + \dots) = \pi_0 EL.$$

当 $EL = \infty$, 归一化条件不能被满足, 因此不变分布不存在. 当 $EL < \infty$ 时, 如果 π 是不变分布, 它必须形如

$$\pi_0 = \frac{1}{EL}, \quad \pi_i = \frac{P(L \geq i + 1)}{EL}, \quad i = 1, 2, \dots$$

反过来, 不难验证, 上式定义的 π 的确满足不变方程, 因此它是(唯一的)不变分布.

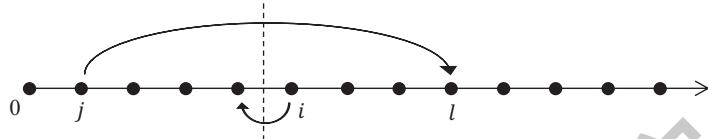
例1.2.7. 假设 $S = \mathbb{Z}_+$, 转移概率为 $p_{0,1} = 1$; 对任意 $i \geq 1$,

$$p_{i,i-1} = \frac{\lambda}{\lambda + 1}, \quad p_{i,i+k} = \frac{p_k}{\lambda + 1}, \quad \forall k \geq 1,$$

其中, $\sum_{k=1}^{\infty} p_k = 1$. 求该马氏链的不变分布.

解: 假设 π 是不变分布. 给定 $i \geq 2$. 取 $A = \{i, i + 1, i + 2, \dots\}$. 与上一个例题类似, 由于粒子从 A 中跳到 A^c 中时只能是从 i 跳到 $i - 1$, 因此 (1.2.2) 的右边等于 $\pi_i p_{i,i-1} = \pi_i$; 粒子从 A^c 中跳至 A 中, 可以从状态 $i - 1$ 及其左边的某个状态 j 跳至状态 i 及其右边的某个状态 l . 记 $f_r := \sum_{k=r}^{\infty} p_k$, 则 (1.2.2) 左边等于

$$\sum_{j=1}^{i-1} \sum_{l=i}^{\infty} \pi_j p_{jl} = \sum_{j=1}^{i-1} \pi_j \sum_{k=i-j}^{\infty} \frac{p_k}{\lambda + 1} = \frac{1}{\lambda + 1} \sum_{j=1}^{i-1} \pi_j f_{i-j}.$$

图 1.8 $A = \{i, i+1, i+2, \dots\}$, $l = j+k$.

粒子从 A 中跳至 A^c 中, 只能从状态 i 跳至状态 $i-1$, 因此的概率流很简单, 就是从 i 流向 $i-1$, 总流量为 $\pi_i \frac{\lambda}{\lambda+1}$. 由 i 的任意性, 不变分布必须满足如下方程组:

$$\pi_i = \frac{1}{\lambda} \sum_{j=1}^{i-1} \pi_j f_{i-j}, \quad \forall i \geq 2. \quad (1.2.3)$$

将上式对 i 求和可得

$$1 - \pi_0 - \pi_1 = \frac{1}{\lambda} \sum_{i=2}^{\infty} \sum_{j=1}^{i-1} \pi_j f_{i-j} = \frac{1}{\lambda} \sum_{j=1}^{\infty} \sum_{i=j+1}^{\infty} \pi_j f_{i-j} = \frac{1}{\lambda} \left(\sum_{j=1}^{\infty} \pi_j \right) \left(\sum_{r=1}^{\infty} f_r \right).$$

由于

$$\sum_{r=1}^{\infty} f_r = \sum_{r=1}^{\infty} \sum_{k=r}^{\infty} p_k = \sum_{k=1}^{\infty} \sum_{r=1}^k p_k = \sum_{k=1}^{\infty} kp_k = m,$$

因此当

$$m := \sum_{k=1}^{\infty} kp_k = \infty$$

时, 不变分布不存在. 当 $m < \infty$ 时, 我们推出不变分布必须满足:

$$1 - \pi_0 - \pi_1 = \frac{m}{\lambda} (1 - \pi_0).$$

进一步, 根据状态 0 处的不变方程,

$$\pi_0 = \sum_{i \in S} \pi_i p_{i0} = \pi_1 \frac{\lambda}{\lambda+1}.$$

因此,

$$\pi_0 = \frac{\lambda - m}{\lambda - m + \lambda + 1}, \quad \pi_1 = \frac{\lambda - m}{\lambda - m + \lambda + 1} \cdot \frac{\lambda + 1}{\lambda}.$$

最后, 根据递推式 (1.2.3) 可得到 π_i , $i \geq 2$. 当 $\lambda \leq m$ 时, 这显然不是不变分布, 因此不变分布不存在. 当 $\lambda > m$ 时, 不难验证这就是不变分布.

对某些特殊的模型, 我们还可以利用空间的平移不变性寻找不变分布.

例1.2.8 (离散圆周上的随机游动). 假设 $N \geq 3$. 考虑有 N 个顶点的离散圆周 $\mathbb{S}_N = \{0, 1, \dots, N-1\}$ 上的马氏链 $\{X_n\}$, 其转移概率为:

$$p_{0,1} = p_{1,2} = \dots = p_{N-2,N-1} = p_{N-1,0} = p, \quad p_{0,N-1} = p_{1,0} = \dots = p_{N-1,N-2} = 1-p,$$

其中 $0 < p < 1$. 当 $p = \frac{1}{2}$ 时, $\{X_n\}$ 就是 \mathbb{S}_n 上的随机游动; 当 $p \neq \frac{1}{2}$ 时, $\{X_n\}$ 被称为 \mathbb{S}_n 上的有偏随机游动. 下面, 我们利用空间的平移不变性寻找不变分布. 将 N 等同于 0, 并且将 -1 等同于 $N-1$, 我们有不变方程等价于,

$$\pi_i = p\pi_{i-1} + (1-p)\pi_{i+1}$$

对所有状态 i 都成立. 因此, 不难验证 $\pi_i \equiv 1$ 满足不变方程. 将其归一化得到 $\pi_i = \frac{1}{N}, i \in S$, 代入上述方程便知 π 就是不变分布.

二、可逆分布.

细致平衡条件:

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad \forall i, j \in S, \tag{1.2.4}$$

定义1.2.9. 假设 \mathbf{P} 不可约, π 为 S 上的一个测度. 若细致平衡条件 (1.2.4) 成立, 则称 π 为 \mathbf{P} 的配称测度 (symmetric measure). 此时, 称 \mathbf{P} 为可配称的. 进一步, 若 π 是一个分布, 则称 π 为 \mathbf{P} 的可逆分布 (reversible distribution). 此时, 称 \mathbf{P} 为可逆的 (reversal).

注1.2.10. 配称测度是不变测度. 可逆分布是满足细致平衡条件的不变分布, 也是满足归一化条件的配称测度.

注1.2.11. 假设 π 是不变分布, $\{X_n\}$ 是以 π 为初分布、 \mathbf{P} 为转移矩阵的马氏链. 固定 N , 令 $Y_n := X_{N-n}$, 则 $\{Y_n : 0 \leq n \leq N\}$ 被称为 $\{X_n : 0 \leq n \leq N\}$ 的时间倒逆过程, 简称逆过程. 那么, $\{Y_n : 0 \leq n \leq N\}$ 也是一个马氏链, 其转移概率为

$$\tilde{p}_{ij} = \frac{\pi_j p_{ji}}{\pi_i}, \quad \forall i, j \in S. \tag{1.2.5}$$

因此, π 是可逆分布指的是 $\tilde{\mathbf{P}} = \mathbf{P}$, 即, 逆过程 $\{Y_n : 0 \leq n \leq N\}$ 与原过程 $\{X_n : 0 \leq n \leq N\}$ 具有相同的初分布和转移概率. 此时, 我们也称以 π 为初分布、 \mathbf{P} 为转移矩阵的马氏链是可逆的.

可逆分布的优越性之一是容易计算. 我们先看一个简单的例子.

例1.2.12 (生灭链). 生灭链刻画某群体中的个体数目, 假设现在有 i 个个体, 经过一个单位时间后, 群体中会增加一个个体(即, 一个新个体出生), 概率为 b_i , 或减少一个个体(即, 一个个

体灭亡), 概率为 d_i . 那么, 群体中的个体数是一个马氏链, 其状态空间为 $S = \mathbb{Z}_+$, 转移概率为

$$p_{i,i+1} = b_i, \quad \forall i \geq 0, \quad p_{i,i-1} = d_i, \quad \forall i \geq 1,$$

其中 $b_0 = 1$, $b_i + d_i = 1$, $\forall i \geq 1$.

在生灭链中, 我们进一步假设 $b_i, d_i > 0$, $i \geq 1$. 取 $j = i - 1$, 那么 (1.2.4) 便转化为

$$\pi_i d_i = \pi_{i-1} b_{i-1}, \quad \forall i \geq 1.$$

这表明 $\pi_i = \pi_{i-1} \frac{b_{i-1}}{d_i}$, $i \geq 1$. 迭代后, 我们有可逆分布必须满足如下方程组:

$$\pi_i = \pi_0 \frac{b_0 \cdots b_{i-1}}{d_1 \cdots d_i}, \quad \forall i \geq 1.$$

事实上, 上述方程组已经包含了所有使得 $p_{ij} > 0$ 的方程组, 因此它的解就是配称测度.

当 $C := 1 + \sum_{i=1}^{\infty} \frac{b_0 \cdots b_{i-1}}{d_1 \cdots d_i} < \infty$ 时,

$$\pi_0 = \frac{1}{C}, \quad \pi_i = \frac{1}{C} \times \frac{b_0 \cdots b_{i-1}}{d_1 \cdots d_i}, \quad i \geq 1$$

是可逆分布. 当 $C = \infty$ 时, 可逆分布不存在, 事实上, 此时不变分布也不存在.

一般地, 假设对任意 $i, j \in S$, 都有 $p_{ij} > 0 \Leftrightarrow p_{ji} > 0$ (否则可逆分布不存在). 那么, 我们总可以通过下面的流程(i) ~ (iii) 来求可逆分布.

- (i) 取定 o , 记 $S_0 = \{o\}$.
- (ii) 通过细致平衡条件, 归纳定义从 o 出发经过 n 步才能到达的状态 i 的 π_i . 具体地, 对 $n \geq 1$, 记

$$S_n = \{i \in S : p_{oi}^{(n)} > 0\} \setminus (S_0 \cup \dots \cup S_{n-1}),$$

假设 $\cup_{n=0}^{\infty} S_n = S$ (否则情况比较复杂, 我们以后再讨论). 对任意 $i \in S_n$, 随便取一个 $k \in S_{n-1}$ 使得 $p_{ki} > 0$, 并令

$$\pi_i = \frac{\pi_k p_{ki}}{p_{ik}}. \quad (1.2.6)$$

- (iii) 验证细致平衡条件 (1.2.4), 并进一步将 π 归一化, 结论为以下三种情形之一.
- 情形一、存在两个状态 i, j 使得 $\pi_i p_{ij} \neq \pi_j p_{ji}$. 那么, 结论是: \mathbf{P} 没有配称测度, 从而也没有可逆分布.

情形二、细致平衡条件 (1.2.4) 成立, 但 π 不可以归一化. 那么, 结论是: \mathbf{P} 有配称测度, 但没有可逆分布.

情形三、细致平衡条件 (1.2.4) 成立, 且 π 可以归一化. 那么, 结论是: 可取到恰当的 π_o 使得 π 就是 \mathbf{P} 的可逆分布.

例1.2.13 (有限图上的随机游动, 例 1.1.10 续). 假设 $\{X_n\}$ 是连通图 $G = (V, E)$ 上的随机游动, 其中 $|V| < \infty$. 假设顶点 i 的度为 d_i . 对所有顶点 i , 令 $\mu_i = d_i$. 那么,

$$\mu_i p_{ij} = \mu_j p_{ji} = \begin{cases} 1, & \text{若 } i, j \text{ 是邻居;} \\ 0, & \text{否则,} \end{cases}$$

因此, 细致平衡条件成立, 即, μ 是配称测度. 因为 $|V| < \infty$, 所以我们可以将它归一化, 即令 $\pi_i = \frac{d_i}{\sum_{j \in V} d_j}$, $i \in V$, 那么, π 是可逆分布.

可逆分布的优越性之二是配称测度具有继承性: 假设 $D \subseteq S$. 现在, 我们禁止所有离开区域 D 的跳跃, 将其改为粒子的原地跳跃. 即, 令

$$\begin{cases} \tilde{p}_{ij} = p_{ij}, & \forall i, j \in D, j \neq i, \\ \tilde{p}_{ii} = p_{ii} + \sum_{j \notin D} p_{ij}, & \forall i \in D. \end{cases}$$

那么, $\tilde{\mathbf{P}} = (\tilde{p}_{ij})_{D \times D}$ 是 D 上的转移矩阵. 若 π 是 \mathbf{P} 的可逆分布, 则 $\pi|_D$ 仍然满足细致平衡条件(1.2.4), 于是将其归一化即得到 $\tilde{\mathbf{P}}$ 的可逆分布; 但是, 若 π 仅仅是 \mathbf{P} 的不变分布而没有配称性, 则 $\mu|_D$ 的归一化未必是 $\tilde{\mathbf{P}}$ 的不变分布.

例1.2.14 (例 1.1.6 续). 考虑一维简单随机游动, $S = \mathbb{Z}$, $p_{i,i+1} = p$, $p_{i,i-1} = 1 - p =: q$, $\forall i \in \mathbb{Z}$. 其中, $0 < p < 1$.

首先, 可逆分布不存在. 否则, 根据细致平衡条件及其归纳法, 我们有, $\forall i \geq 1$,

$$\pi_1 = \pi_0 \frac{p}{q}, \dots, \pi_i = \pi_0 \frac{p^i}{q^i}, \dots; \quad \pi_{-1} = \pi_0 \frac{q}{p} = \pi_0 \frac{p^{-1}}{q^{-1}}, \dots, \pi_{-i} = \pi_0 \frac{p^{-i}}{q^{-i}}, \dots$$

不难验证, 上面定义的 π 满足细致平衡条件, 即, 它是配称测度. 然而, $1 + \sum_{i=1}^{\infty} \frac{p^i}{q^i} + \sum_{i=1}^{\infty} \frac{q^i}{p^i} = \infty$, $\forall 0 < p < 1$. 因此, 上面定义的 π 不能进行归一化. 从而可逆分布不存在. 进一步, 如果我们只考虑正半轴, 那么该配称测度基本上仍然适用, 具体地, 见本节习题3.

其次, 事实上不变分布也不存在. 否则, 取 $A = \{i, i+1, \dots\}$, 那么 (1.2.2) 就转化为细致平衡条件. 这表明不变分布必须是可逆分布, 从而它不存在.

最后, 我们已经发现, 一方面, 上面定义的 $\pi_i = \pi_0 \frac{p^i}{q^i}$, $\forall i \in \mathbb{Z}$ 满足配称性, 因此它是不变测度. 另一方面, 由空间的齐次性, 不难得出 $\pi_i = \pi_0$, $\forall i$ 也是不变测度. 当 $p = q$ 时, 它们本质是同一类不变测度, 即它们只相差常数倍. 当 $p \neq q$ 时, 它们本质是不同的不变测度, 因此, 我们其实找到了两种不同类型的不变测度.

三、访问频率

假设 L_1, L_2, \dots 独立同分布, 取非负整数, $P(L_1 = 0) < 1$. 令

$$S_0 = 0, \quad S_r = L_1 + \dots + L_r, \quad r \geq 1.$$

令

$$R_n = \max\{r \geq 0 : S_r \leq n\}.$$

定理1.2.15 (更新定理).

$$P\left(\lim_{n \rightarrow \infty} \frac{R_n}{n} = \frac{1}{EL_1}\right) = 1.$$

证: 我们可以将 S_1, S_2, \dots 标记在时间轴上, 如下图所示的实心点. $R_n = r$ 当且仅当 $S_r \leq n < S_{r+1}$, 例如, 图示为 $r = 4$.

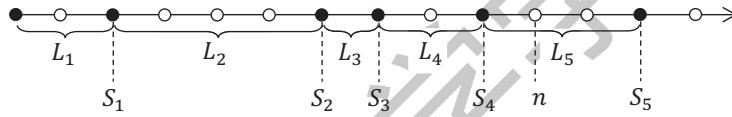


图 1.9 $R_n = 4$.

一方面, 不难看出, R_n 单调上升至无穷大, 虽然它并不是严格单调上升的, 或者说, 上升得很慢. 严格地说, $P(\lim_{n \rightarrow \infty} R_n = \infty) = 1$. 另一方面, 根据强大数定律, $P(\lim_{r \rightarrow \infty} \frac{S_r}{r} = EL) = 1$, 其中 $L = L_1$. 换句话说, 记

$$\Omega_1 = \left\{ \omega : \lim_{n \rightarrow \infty} R_n(\omega) = \infty \right\}, \quad \Omega_2 = \left\{ \omega : \lim_{r \rightarrow \infty} \frac{S_r(\omega)}{r} = EL \right\},$$

那么, $P(\Omega_1) = P(\Omega_2) = 1$.

给定 $\omega \in \Omega_1 \Omega_2$. 因为 $\omega \in \Omega_2$, 所以, 当 $r \rightarrow \infty$ 时, 实数序列 $a_r = \frac{S_r(\omega)}{r} \rightarrow EL$. 于是, 由数学分析的知识, 我们知道, 如果 $\lim_{n \rightarrow \infty} r_n = \infty$, 那么也有 $\lim_{n \rightarrow \infty} a_{r_n} = EL$. 特别地, 我们取 $r_n = R_n(\omega)$. 因为由 $\omega \in \Omega_2$ 知 $\lim_{n \rightarrow \infty} r_n = \infty$, 所以 $\lim_{n \rightarrow \infty} \frac{S_{r_n}(\omega)}{r_n} \rightarrow EL$. 同理, $\frac{S_{r_n+1}(\omega)}{r_n} = \frac{S_{r_n+1}(\omega)}{r_n+1} \times \frac{r_n+1}{r_n} \rightarrow EL$. 这结合 $S_{r_n} \leq n < S_{r_n+1}$ 表明 $\lim_{n \rightarrow \infty} \frac{n}{r_n} = EL$, 即 $\lim_{n \rightarrow \infty} \frac{R_n(\omega)}{n} = \frac{1}{EL}$.

综上, 我们有 $P(\Omega_1 \Omega_2) = 1$ 且 $\Omega_1 \Omega_2 \subseteq \left\{ \omega : \lim_{n \rightarrow \infty} \frac{R_n(\omega)}{n} = \frac{1}{EL} \right\}$, 于是结论成立. \square

下面, 我们要在更新过程和两状态马氏链这两个例子中, 利用更新定理看出不变分布与状态的访问频率之间的关系.

例1.2.16 (更新过程, 例 1.2.6 续). 在更新过程中, L_r 就是第 r 个灯泡的寿命, $S_r = L_1 + \dots + L_r$ 就是第 r 个更新时刻(换灯泡的时刻). 因为 X_n 表示灯泡余寿, 所以 $X_n = 0$ 当且仅当 n 是一个更新时刻. 换句话说, $\{X_n\}$ 每访问一次状态 0, 就意味着使用一个新灯泡, 也就是图 1.9 的时间轴上出现一个实心原点. 因此, 在时刻 n 之前总共出现了 $1 + R_n$ 次状态 0. 状态 0 出现的频率就是 $\frac{1+R_n}{n}$. 根据更新定理, 状态 0 出现的频率几乎必然收敛, 它的极限 $\frac{1}{EL}$ 就是在不变分布下, 状态 0 出现的概率 π_0 .

例1.2.17 (两状态马氏链, 例 1.2.5 续). 在两状态马氏链中, 考察 $X_0 = 0$ 的轨道. 直观上, 如果我们将粒子下一步跳到状态 1 视为成功, 将粒子在状态 0 原地跳视为失败, 假设粒子在状

态0 原地跳跃 $\xi - 1$ 次，并在第 ξ 次跳至状态1，那么， ξ 服从几何分布，参数是 $p_{01} = p$. 同理，如果粒子位于状态1，它在状态1 原地跳跃 $\eta - 1$ 次，并在第 η 次跳至状态0，那么， η 服从几何分布，参数为 $p_{10} = q$. 因此，我们可以用相互独立的服从几何分布的随机变量序列来构造两状态马氏链. 具体的构造如下.

假设 $\xi_1, \xi_2, \dots, \eta_1, \eta_2, \dots$ 是相互独立的随机变量，分布列如下：

$$P(\xi_r = m) = (1-p)^{m-1}p, \quad m = 1, 2, \dots; \quad P(\eta_r = m) = (1-q)^{m-1}q, \quad m = 1, 2, \dots.$$

记 $S_0 = T_0 = 0$, $S_r = \xi_1 + \dots + \xi_r$, $T_r = \eta_1 + \dots + \eta_r$. 那么，一个从状态0 出发的马氏链看上去就是：粒子原地跳，直到时刻 S_1 跳至状态1，然后原地跳，直到时刻 $S_1 + T_1$ 跳至状态0，然后原地跳，直到时刻 $S_1 + T_1 + S_2$ 跳至状态1，，然后原地跳，直到时刻 $S_1 + T_1 + S_2 + T_2$ 跳至状态0, 具体地，令 $L_0 = 0$, $L_r = S_r + T_r$, 令

$$X_{L_r} = \dots = X_{L_r + S_{r+1}-1} = 0, \quad X_{L_r + S_{r+1}} = \dots = X_{L_{r+1}-1} = 1, \quad \forall r \geq 0.$$

那么， $\{X_n\}$ 就是一个从0 出发的马氏链. 同理，令

$$Y_{L_r} = \dots = Y_{L_r + T_{r+1}-1} = 0, \quad Y_{L_r + T_{r+1}} = \dots = Y_{L_{r+1}-1} = 1, \quad \forall r \geq 0.$$

那么， $\{Y_n\}$ 就是一个从0 出发的马氏链. 读者可以根据 (1.1.1) 直接验证这个结论.

如果我们将 L_1, L_2, \dots 视为换灯泡的时间. 根据更新定理，在时刻 n 之前换了 R_n 个灯泡，因此，状态0 出现的次数 $V_0(n)$ 介于 S_{R_n} 与 S_{R_n+1} 之间. 根据强大数定律与更新定理，状态0 出现的频率 $\frac{1}{n}V_0(n)$ 几乎必然收敛于

$$E\xi_1 \cdot \frac{1}{EL} = \frac{E\xi_1}{E\xi_1 + E\eta_1} = \frac{1/p}{1/p + 1/q} = \frac{q}{p+q} = \pi_0.$$

同理，状态1 出现的频率几乎必然收敛于 π_1 .

四、访问概率的收敛性.

例1.2.18 (两状态马氏链，例 1.2.5 续). 将两状态马氏链的转移矩阵 \mathbf{P} 通过相似变换进行对角化，表达如下：

$$\mathbf{P} = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \mathbf{A},$$

其中， $\lambda = 1 - p - q$,

$$\mathbf{A} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{1}{p+q} & -\frac{1}{p+q} \end{pmatrix}, \quad \mathbf{A}^{-1} = \begin{pmatrix} 1 & p \\ 1 & -q \end{pmatrix}.$$

于是,

$$\mathbf{P}^n = \mathbf{A}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & \lambda^n \end{pmatrix} \mathbf{A}.$$

由 $0 < p + q < 2$ 知 $|\lambda| < 1$. 从而, 当 $n \rightarrow \infty$ 时, $\lambda^n \rightarrow 0$. 于是,

$$\hat{\mathbf{P}} := \lim_{n \rightarrow \infty} \mathbf{P}^n = \mathbf{A}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{A} = \begin{pmatrix} \frac{q}{p+q} & \frac{p}{p+q} \\ \frac{q}{p+q} & \frac{p}{p+q} \end{pmatrix}.$$

由于 $\hat{\mathbf{P}}$ 的每一行的行向量都是不变分布 $\pi = (\pi_0, \pi_1)$, 因此上面的结论表明: 对任意 i 都有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j, \quad \forall j.$$

进而, 对任意初分布 μ , 都有当 $n \rightarrow \infty$ 时, $P_\mu(X_n = j) = \mu_0 p_{0j}^{(n)} + \mu_1 p_{1j}^{(n)} \rightarrow \pi_j, \forall j$. 最后,

$$\mathbf{P}^n - \hat{\mathbf{P}} = \mathbf{A}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \lambda^n \end{pmatrix} \mathbf{A} = \lambda^n \begin{pmatrix} \frac{p}{p+q} & -\frac{p}{p+q} \\ -\frac{q}{p+q} & \frac{q}{p+q} \end{pmatrix}.$$

这表明, $p_{ij}^{(n)} - \pi_j = O(\lambda^n)$, 即 $p_{ij}^{(n)} - \pi_j$ 以指数速度趋于 0.

马氏链可以没有不变分布(例如, 一维简单随机游动, 见例 1.2.14), 也可以有多个不变分布(例如, $p_{ii} = 1, \forall i$). 马氏链研究中的一个基本问题是找到所有的不变分布. 因此, 在后续的章节中, 我们将要研究不变分布的存在唯一性. 对于不变分布 π 存在唯一的马氏链, 我们还将进一步研究下列问题, 它们可以说是在从不同的角度刻画马氏链 $\{X_n\}$ 的某种极限行为.

1. 假设在前 n 步中, 状态 i 出现了 $V_i(n)$ 次. 当 $n \rightarrow \infty$ 时, 频率 $\frac{V_i(n)}{n}$ 是否趋于 π_i ? 就像在例 1.2.17 与例 1.2.16 中看到的那样, 状态 0 出现的频率几乎必然收敛到 π_0 .
2. 在什么条件下, 对任意初分布 μ , 都有 $\mu \mathbf{P}^n$ 收敛到不变分布? 就像在例 1.2.18 中看到的那样.
3. 如果 $\mu \mathbf{P}^n$ 收敛到不变分布, 那么收敛速度有多快? 什么情况下会如同例 1.2.18 那样表现出指数速度?

补充知识.

事实上, \mathbf{P} 可配称当且仅当 **Kolmogorov** 准则成立: 若 $i_0, i_1, \dots, i_n, i_{n+1}$, 其中 $n \geq 1$, 形成概率转移图中的一个圈, 即, $i_{n+1} = i_0, p_{i_r i_{r+1}} > 0$, 则

$$p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_n i_{n+1}} = p_{i_{n+1} i_n} p_{i_n i_{n-1}} \cdots p_{i_1 i_0}. \quad (1.2.7)$$

本质上, Kolmogorov 准则描述的是: 若从状态 o 出发可以有两条不同的路径到达状态 i , 假设它们的中间状态分别是 i_1, \dots, i_n 和 j_1, \dots, j_m , 其中 $n, m \geq 0$ 且 $n + m > 0$, 则沿着这两条路径通过 (1.2.6) 算出的 μ_i 是一样的.

习题

1. 通过解方程直接求如下转移矩阵的不变分布.

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

2. 证明: 若 π 为不变分布, 则对任意 $A \subseteq S$, (1.2.2) 成立.

3. 设 $\{X_n\}$ 是取值于 \mathbb{Z}_+ 的马氏链, 其转移概率为:

$$p_{01} = 1 - p_{00} = q, \quad p_{i,i+1} = 1 - p_{i,i-1} = p, \quad \forall i = 1, 2, \dots.$$

其中 $0 < q \leq 1$, $0 < p < 1$. 当 $p < 1/2$ 时, 求不变分布 π 并计算 $E_\pi X_{100}$. (注: 当 $q = 1$ 时, $\{X_n\}$ 被称为 \mathbb{Z}_+ 上的带反射壁的随机游动; 当 $q = 0$ 时, $\{X_n\}$ 被称为 \mathbb{Z}_+ 上的带吸收壁的随机游动, 当 $0 < q < 1$ 时, $\{X_n\}$ 被称为 \mathbb{Z}_+ 上的带粘滞边界的随机游动.)

4. 假设 $S = \{1, 2, 3\}$, 转移矩阵如下. (1) 求不变分布, (2) 求 $\lim_{n \rightarrow \infty} P(X_n = 1)$.

$$\mathbf{P} = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}.$$

5. 证明: 转移矩阵 \mathbf{P} 的不变分布全体构成一个凸集, 即, 若 μ, π 都是 \mathbf{P} 的不变分布, $0 < p < 1$, 则 $p\mu + (1-p)\pi$ 也是 \mathbf{P} 的不变分布.

6. 假设 π 是马氏链的不变分布. 假设存在常数 $c > 0$ 使得 $p_{ki} = cp_{kj}, \forall k$. 证明 $\pi_i = c\pi_j$.

7. 若转移矩阵 $\mathbf{P} = (p_{ij})_{S \times S}$ 满足 $\sum_{i \in S} p_{ij} = 1$, 则称 \mathbf{P} 为双重随机的 (double stochastic).

证明: (1) 若 \mathbf{P} 是双重随机的, 则对任何 n , \mathbf{P}^n 也是双重随机的; (2) 若 \mathbf{P} 是双重随机的, 则 $\mu \equiv 1$ 是其不变测度.

8. 假设 S 有限, \mathbf{P} 是 S 上的转移矩阵. 固定 $i \in S$. 证明: (1) 存在正整数的子列 n_1, n_2, \dots 使得 $\lim_{r \rightarrow \infty} \frac{1}{n_r} \sum_{m=0}^{n_r-1} p_{ij}^{(m)}$ 存在(将此极限记为 π_j), $\forall j \in S$. (2) 证明: $\pi = (\pi_j, j \in S)$ 是 \mathbf{P} 的不变分布.

9. 证明: 可逆分布是不变分布.

10. 验证注 1.2.11 中的 $\{Y_n, 0 \leq n \leq N\}$ 是马氏链, 并且, 其初分布仍为 π , 转移概率由 (1.2.5) 给出.
11. 假设 S 有限; $p_{ij} > 0, \forall i \neq j$; 并且 $p_{ij}p_{jk}p_{ki} = p_{ik}p_{kj}p_{ji}, \forall i, j, k$. 证明: \mathbf{P} 是可逆的.
12. 假设 π 为 \mathbf{P} 的不变分布. 假设 $\{X_n\}$ 与 $\{Y_n\}$ 都是 S 上的马氏链, 满足: (i) 转移概率分别为 \mathbf{P} 与 $\tilde{\mathbf{P}}$ (由 (1.2.5) 定义), (ii) $Y_0 = X_0 \sim \pi$, (iii) 在已知 $X_0 = Y_0 = i$ 的条件下, $\{X_n : n \geq 1\}$ 与 $\{Y_n : n \geq 1\}$ 相互独立. 令

$$Z_n = \begin{cases} X_n, & \text{若 } n \geq 0; \\ Y_{-n}, & \text{若 } n < 0 \end{cases}$$

证明: 给定 $N \in \mathbb{Z}$, 令 $W_n = Z_{N+n}, \forall n \geq 0$. 那么, $\{W_n\}$ 是以 π 为初分布, \mathbf{P} 为转移矩阵的马氏链.

§1.3 状态的分类

前面我们已经提到过, 马氏链等价于一个在 S 中运动的粒子. 在接下来的这几节中, 我们将研究该粒子的轨道性质, 例如, 沿这条随机的轨道, 我们是否能观察到某个特定的状态, 即, 粒子是否会到达(或访问)某个状态.

定义1.3.1. 若 $P_i(\exists n \geq 0 \text{ 使得 } X_n = j) > 0$, 则称 i 可达 j , 记为 $i \rightarrow j$.

在研究可达这一性质时, 概率转移简图是一个基本工具, 它就是在转移概率图中不标记 p_{ij} 的值, 只对 $p_{ij} > 0$ 的状态对 (i, j) 画出从 i 到 j 的箭头(即, 有向边). 例如:

$$\mathbf{P}_1 = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

对应的转移概率图如下.

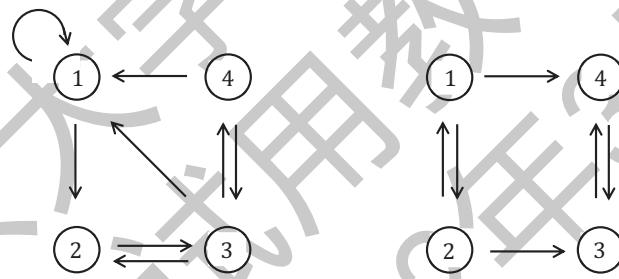


图 1.10 概率转移简图

在概率转移简图中, i 可达 j 可直接读解为: 从 i 出发, 可以顺着箭头方向到达 j , 这即是下面命题中的(1).

命题1.3.2. 假设 $i \neq j$. 那么 $i \rightarrow j$ 与下列两条等价:

- (1) 存在 $n \geq 1$, 及 $n+1$ 个互不相同的状态 $i_0, \dots, i_n \in S$, 使得 $i_0 = i$, $i_n = j$ 且 $p_{i_0 i_1} > 0$, $p_{i_1 i_2} > 0, \dots, p_{i_{n-1} i_n} > 0$.
- (2) 存在 $n \geq 1$ 使得 $p_{ij}^{(n)} > 0$.

在图 1.10 中, 左图满足这样的性质: 从任意顶点 i 出发都可以顺着箭头方向都可以到达任意的另一个顶点 j , 即任意顶点对 i, j 都有 i 可达 j . 而右图则不满足此性质.

定义1.3.3. 若 $i \rightarrow j$ 且 $j \rightarrow i$, 则称 i 与 j 互通(communicate), 记为 $i \leftrightarrow j$. 若 S 中任意两个状态都互通, 则称该马氏链不可约(irreducible); 否则称该马氏链可约(reducible).

注1.3.4. 可达与互通本质上是由转移矩阵 \mathbf{P} 决定的性质, 因此, 我们也说 \mathbf{P} 可约或不可约. 在明确转移矩阵的时候, 我们也说 S 可约或不可约.

可达这一性质具有传递性: 若 $i \rightarrow j$ 且 $j \rightarrow k$, 则 $i \rightarrow k$. 于是我们推出:

$$i \leftrightarrow i; \quad i \leftrightarrow j \text{ 则 } j \leftrightarrow i; \quad i \leftrightarrow j \text{ 且 } j \leftrightarrow k, \text{ 则 } i \leftrightarrow k.$$

这表明互通关系是 S 上的一个等价关系. 将 S 按照互通关系划分为互不相交的等价类, 每个等价类被称为一个互通类, 简称类(class). 于是, 马氏链不可约就是指只有一个互通类, 例如上面的转移矩阵 \mathbf{P}_1 . 在上面的 \mathbf{P}_2 对应的马氏链中, 有两个互通类: $C_1 = \{1, 2\}$, $C_2 = \{3, 4\}$. 从 C_2 出发的粒子永远在 C_2 中, 这表明 C_2 有某种封闭性. 但是, C_1 没有这种封闭性, 因为从 C_1 出发的粒子可以离开 C_1 . 根据互通类的定义, 粒子一旦离开 C_1 就不能再回到 C_1 中, 于是, 接下来 C_1 中的状态将再也不会出现. 从这个角度看, 像 C_2 这样有封闭性的互通类就更值得研究一些.

定义1.3.5. 若 $A \subseteq S$ 满足: $\sum_{j \in A} p_{ij} = 1, \forall i \in A$, 则称 A 为一个闭集.

命题1.3.6. 假设 A 为闭集, 则 $P_i(X_n \in A, \forall n \geq 0) = 1, \forall i \in A$.

如果 A 是闭集, 那么我们可以将状态空间缩小为 $\hat{S} = A$, 转移概率就是 $\hat{\mathbf{P}} = (p_{ij})_{i,j \in A}$, 这个矩阵当然是转移矩阵, 因为 A 是闭集. 特别地, 如果 A 是一个闭的互通类, 那么 $\hat{\mathbf{P}}$ 还是一个不可约马氏链. 不过, 并不是所有的马氏链都有闭的互通类, 例如: $S = \mathbb{Z}, p_{i,i+1} = 1, \forall i \in S$.

例1.3.7 (生灭链, 例 1.1.8 与 1.2.12 续). 在 Ehrenfest 模型中, 补充定义 $b_i = 0, d_i = 1, \forall i \geq N+1$, 便得到一个生灭链. 在该生灭链中, 状态0 所在的互通类 A 是一个闭集. 而 Ehrenfest 模型则可视为该生灭链在 A 上的限制.

习题

1. 将小白鼠放在如下图所示的迷宫中作随机的移动. 即, 当它处于某一格子且此格子有 k 个门通入其它格子, 则小白鼠以 $1/k$ 的概率选择任一条路径. 假如小白鼠每次移动一个格子, 试写出它的转移概率, 并按互通的关系分解状态空间.
2. 证明命题1.3.2 和命题1.3.6.
3. 假设 A 是闭集, C 是一个互通类. 证明: 或者 $C \subseteq A$, 或者 $C \cap A = \emptyset$.
4. 证明: 状态空间 S 可约当且仅当 S 有非空、闭的真子集.
5. 假设状态空间 S 有限. 证明: 存在闭的互通类.

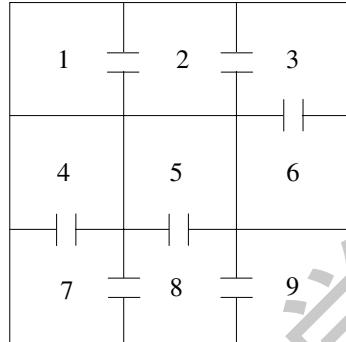


图 1.11 习题1 之图

§1.4 首达时与强马氏性

假设 $\{X_n\}$ 是一个马氏链, $i \in S$, 令

$$\tau_i := \inf\{n \geq 0 : X_n = i\}, \quad (1.4.1)$$

其中, 我们约定 $\inf \emptyset = \infty$. 那么, τ_i 表示马氏链 $\{X_n\}$ 首次访问状态*i*的时间, 它被称为 $\{X_n\}$ 的首达时 (hitting time). 由于 X_n 都是随机变量, 因此首达时是一个可以取值非负整数和 ∞ 的(广义)随机变量. 有时, 为了强调它是随机变量, 也就是说, 为了突出样本 ω , 我们也将 ω 写出来; 有时, 为了强调它是马氏链 $\{X_n\}$ 的首达时, 我们也在上标中写出“(X)”. 具体地,

$$\tau_i(\omega) = \tau_i^{(X)} = \inf\{n \geq 0 : X_n(\omega) = i\}.$$

我们还可以定义

$$\sigma_i := \inf\{n \geq 1 : X_n = i\}, \quad (1.4.2)$$

它被称为 $\{X_n\}$ 的首入时, 表示马氏链 $\{X_n\}$ 首次从某处进入位置*i*的时间, 即在 $n = \sigma_i$ 的时刻, 粒子(首次从某状态 X_{n-1})跳至状态 $X_n = i$. 需要注意的是, 总有 $\sigma_i \geq 1$. 假设初始状态 $X_0 = i_0$. 一般而言, 首达时 τ_i 多用于 $i_0 \neq i$ 的情况, 它表示从 i_0 出发的粒子首次到达状态*i*的时间, 此时 $\sigma_i = \tau_i$; 而首入时则多用于初始状态 $i_0 = i$ 的情况, 它表示一个从*i*出发的粒子首次回访状态*i*的时间, 此时 $\tau_i = 0 < \sigma_i$. 类似地, 对任意 $D \subseteq S$, 我们可以定义

$$\tau_D := \inf\{n \geq 0 : X_n \in D\}, \quad \sigma_D := \inf\{n \geq 1 : X_n \in D\}.$$

类似地, 我们可以更具体地写 $\tau_D(\omega)$, $\tau_D^{(X)}$, $\sigma_D(\omega)$, $\sigma_D^{(X)}$. 从定义看, τ_i 与 σ_i 都是在轨道中搜寻第一次访问状态*i*的时刻, 它们的区别是: 搜寻的时间起点是0还是1.

例1.4.1. 令 $Y_n = X_{1+n}$, $n = 0, 1, 2, \dots$. 那么, $\sigma_i^{(X)} = 1 + \tau_i^{(Y)}$, $\forall i \in S$. 注意到 $\{Y_n\}$ 是在 $\{X_n\}$ 中扔掉初始状态 X_0 而得到的过程, 或者说, $\{Y_n\}$ 是在 $\{X_n\}$ 的轨道中将时刻 1 视为新的初始时刻而看到的后续轨道. 例如, 图 1.12 是 $S = \mathbb{Z}$ 上一条从 0 出发的简单随机游动的轨道, 横坐标是时间轴, 纵坐标是空间 S , $\{Y_n\}$ 的轨道就是其中的实线部分, 它是一条从 -1 出发的轨道. 从图上可以看出, $\sigma_0^{(X)} = 1 + \tau_0^{(Y)}$.

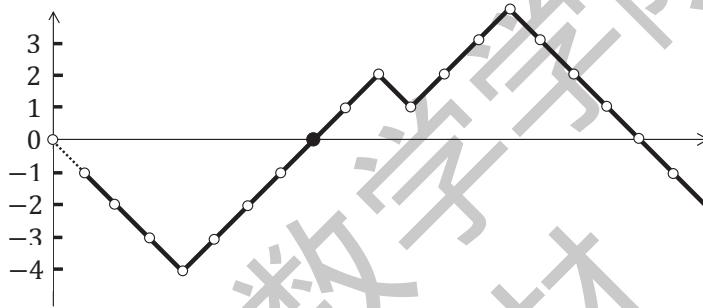


图 1.12 $\sigma_0^{(X)} = 8$, $\tau_0^{(Y)} = 7$.

下面, 我们取定 i , 将 τ_i 简记为 τ . 假设 $X_0 \sim \mu$, 将 P_μ 简记为 P . 假设 $P(\tau < \infty) > 0$. 在事件 $\{\tau < \infty\}$ 上, 我们可以将 $\{X_n\}$ 的轨道分成两段: 时刻 τ 之前的轨道, 记为 \vec{Z} ; 时刻 τ 之后的轨道, 记为 $\vec{Y} = \{Y_m\}$. 例如, 图 1.13 中的实线部分是 \vec{Y} , 虚线部分是 \vec{Z} . 具体地, 令

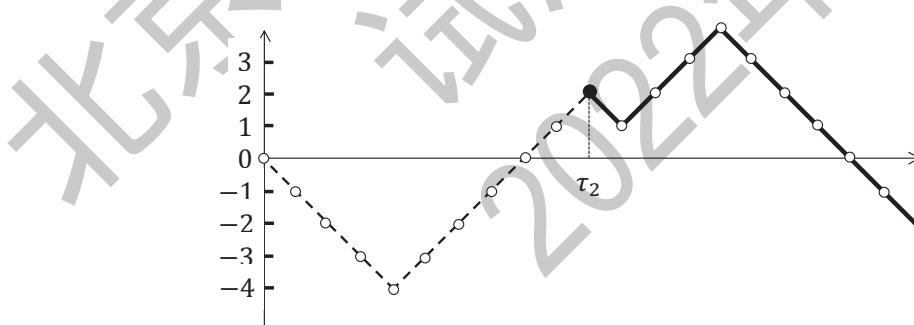


图 1.13 $i = 2$. $\vec{Y} = (2, 1, 2, 3, 4, 3, 2, 1, 0, -1, \dots)$; $\vec{Z} = (0, -1, -2, -3, -4, -3, -2, -1, 0, 1, 2)$.

$$Y_m := X_{\tau+m}, \quad \forall m \geq 0; \quad \vec{Z} := (X_0, \dots, X_\tau). \quad (1.4.3)$$

下面我们要考虑的是在事件 $\tau < \infty$ 发生的条件下的轨道性质, 也就是在条件概率 $P(\cdot | \tau < \infty)$ 下考虑. 鉴于此, 可以认为 $\{Y_m\}$ 和 \vec{Z} 的定义已经完整, 换句话说, 我们不需要在事件 $\{\tau = \infty\}$ 上对它们进行补充定义. 在 $\tau < \infty$ 发生的条件下, 我们可以将 \vec{Z} 视为一条随机轨道, 它的长

度也是随机变量. 由于长度有限, 并且状态空间 S 是离散的, 因此, \vec{Z} 可以被视为一个离散型随机变量, 它的所有可能取值组成的集合记为 I , 定义如下: 令

$$I_n := \left\{ (i_0, i_1, \dots, i_n) \in S^{n+1} : i_0, \dots, i_{n-1} \neq i \text{ 且 } i_n = i \right\}, \quad \forall n \geq 0,$$

即 I_n 是所有满足 “在时刻 n 首次到达状态 i ” 这一要求的有限步轨道, 它是 S^{n+1} 的子集; 那么, \vec{Z} 取值于 $I := \cup_{n=0}^{\infty} I_n$. 进一步,

$$\left\{ \vec{Z} = \vec{i} \right\} = \left\{ (X_0, \dots, X_n) = \vec{i} \right\}, \quad \forall \vec{i} = (i_0, \dots, i_n) \in I_n. \quad (1.4.4)$$

这是因为: 当 $\vec{i} \in I_n$ 时, 根据 \vec{Z} 的定义, 等号左边的事件蕴含着 $\tau = n$, 进而蕴含着 $\vec{Z} = (X_0, \dots, X_n)$, 从而蕴含着等号右边的事件; 反过来, 根据 I_n 的定义及其 $\vec{i} \in I_n$, 等号右边的事件蕴含着 $\tau_i = n$, 即 $\tau = n$, 进而蕴含着 $\vec{Z} = (X_0, \dots, X_n)$, 从而蕴含着等号左边的事件. 事实上, 当 \vec{i} 取遍 I_n 时, 上述事件(等号两边的事件)就是 $\tau = n$ 的划分. 于是,

$$P(\tau < \infty) = \sum_{n=0}^{\infty} P(\tau = n) = \sum_{n=0}^{\infty} \sum_{\vec{i} \in I_n} P(\vec{Z} = \vec{i}) = \sum_{n=0}^{\infty} \sum_{\vec{i} \in I_n} P((X_0, \dots, X_n) = \vec{i}).$$

进一步, 根据 (1.1.2),

$$P((X_0, \dots, X_n) = \vec{i}) = \mu_{i_0} p_{i_1 i_1} \cdots p_{i_{n-1} i_n}.$$

因此, 我们还可以得到 $P(\tau < \infty)$, 以及在 $\tau < \infty$ 的条件下, Z 的条件分布列:

$$P(\vec{Z} = \vec{i} \mid \tau < \infty) = \frac{1}{P(\tau < \infty)} \mu_{i_0} p_{i_1 i_1} \cdots p_{i_{n-1} i_n}, \quad \forall n \geq 0, \vec{i} \in I_{n+1}.$$

命题1.4.2. 在 $\{\tau_i < \infty\}$ 发生的条件下, $\{Y_m\}$ 是一个从 i 出发的, 以 \mathbf{P} 为转移矩阵的马氏链, 并且它与 \vec{Z} 相互独立.

证: 我们先证明在 $\{\tau = n\}$ 发生的条件下结论成立, 然后再证明在 $\tau < \infty$ 发生的条件下结论成立.

给定 $n \geq 0$. 假设 $P(\tau = n) > 0$, 将条件概率 $P(\cdot \mid \tau = n)$ 简记为 $P_n(\cdot)$. (注: 在此证明中, 暂时使用 $P_n(\cdot)$ 这个符号, 相信读者不会将其与从状态 n 出发的马氏链对应的概率混淆.) 注意到若 $\tau = n$, 则 $\vec{Z} \in I_n$. 根据命题 0.4.4, 为证明在 $\tau = n$ 发生的条件下结论成立, 我们只需验证

$$P_n(\vec{Y} = \vec{j}, \vec{Z} = \vec{i}) = g(\vec{j}) P_n(\vec{Z} = \vec{i}) \quad (1.4.5)$$

对任意 $m \geq 0$, $\vec{j} \in S^{m+1}$ 以及任意 $\vec{i} = (i_0, \dots, i_n) \in I_n$ 都成立, 其中 $\vec{Y} = (Y_0, \dots, Y_m)$, $\vec{j} = (j_0, \dots, j_m)$, $g(\vec{j}) = \mathbf{1}_{\{j_0=i\}} p_{j_0 j_1} p_{j_1 j_2} \cdots p_{j_{m-1} j_m}$. 下面, 我们验证之. 考察事件 $\{\vec{Y} =$

$\vec{j}, \vec{Z} = \vec{i}, \tau = n\}$. 当 $\tau = n$ 发生时, 一方面, $Y_0 = X_\tau = X_n, Y_1 = X_{\tau+1} = X_{n+1}, \dots$, 因此, $\{\vec{Y} = \vec{j}\}$ 可改写为

$$A := \{X_n = j_0, \dots, X_{n+m} = j_m\};$$

另一方面, 因为 $\vec{i} \in I_n$, 所以根据(1.4.4), $\{\vec{Z} = \vec{i}\}$ 可改写为

$$B := \{X_0 = i_0, \dots, X_n = i_n\}.$$

因此,

$$\{\vec{Y} = \vec{j}, \vec{Z} = \vec{i}, \tau = n\} = A \cap B \cap \{\tau = n\}.$$

进一步, 根据 I_n 的定义, 由 $\vec{i} \in I_n$ 知 $i_n = i$, 并且事件 B 蕴含着 $\tau = n$, 也就是说

$$\{X_0 = i_0, \dots, X_n = i_n\} \subseteq \{\tau = n\}, \quad \forall \vec{i} \in I_n. \quad (1.4.6)$$

换句话说, $B = B \cap \{\tau = n\} = \{\vec{Z} = \vec{i}, \tau = n\}$. 于是,

$$\begin{aligned} P(\vec{Y} = \vec{j}, \vec{Z} = \vec{i}, \tau = n) &= P(A \cap B \cap \{\tau = n\}) = P(AB) = P(B)g(\vec{j}) \\ &= g(\vec{j})P(B \cap \{\tau = n\}) = g(\vec{j})P(\vec{Z} = \vec{i}, \tau = n). \end{aligned} \quad (1.4.7)$$

左右两边同时除以 $P(\tau = n)$ 即得(1.4.5).

最后, 考虑 $\{\tau < \infty\}$, 它可以划分为 $\{\tau = n\}, n = 0, 1, \dots$. 对任意 $\vec{i} \in I = \bigcup_{n=0}^{\infty} I_n$, 存在唯一的非负整数 n 使得 $\vec{i} \in I_n$. 仍然 $\{X_0 = i_0, \dots, X_n = i_n\}$. 此时, 根据(1.4.7), 我们有

$$\begin{aligned} P(\vec{Y} = \vec{j}, \vec{Z} = \vec{i}, \tau < \infty) &= P(\vec{Y} = \vec{j}, \vec{Z} = \vec{i}, \tau = n) \\ &= g(\vec{j})P(B \cap \{\tau = n\}) = g(\vec{j})P(B \cap \{\tau < \infty\}) \\ &= g(\vec{j})P(\vec{Z} = \vec{i}, \tau < \infty). \end{aligned}$$

两边同时除以 $P(\tau < \infty)$, 再结合命题0.4.4便知结论成立. \square

注1.4.3. 命题1.4.2是强马氏性的特殊情形, 其一般情形请见本节补充知识中的命题1.4.11.

推论1.4.4. 若 $P(\tau_i < \infty) = 1$, 那么 $\{Y_m\}$ 是从 i 出发的马氏链, 并且与 \vec{Z} 相互独立.

推论1.4.5. 将 τ_i 改为 σ_i , 则命题1.4.2与推论1.4.4仍然成立.

证: 令 $\tilde{X}_n := X_{n+1}, n \geq 0$. 对 $\{\tilde{X}_n\}$ 应用命题1.4.2与推论1.4.4即可. \square

注1.4.6. 马氏性指把某个固定的时刻 n 当作现在, 那么在已知现在处于状态 i 的条件下, 将来是从 i 出发的马氏链, 并且与过去的轨道相互独立. 强马氏性则是说, 将随机时刻 τ_i (或 σ_i)当作现在, 那么我们已经知道现在的状态是 i , 在此条件下, 将来是从 i 出发的马氏链, 并且与过去的轨道 \vec{Z} 相互独立.

补充知识.

强马氏性不仅对 τ_i 和 σ_i 成立, 事实上它对更多的随机时刻成立. 在命题 1.4.2 的证明过程中, 我们可以看到, 至关重要的是公式 (1.4.6). 将前 n 步轨道记为 $\vec{i} = (i_0, \dots, i_n)$. 公式 (1.4.6) 则是说: 由 $\vec{i} \in I_n$ 可判断事件 $\tau_i = n$ 成立. 事实上, 我们还可以看出, 由 $\vec{i} \in S^{n+1} \setminus I_n$ 可判断事件 $\tau_i = n$ 不成立. 满足这一性质的随机时间被称为停时, 其数学定义如下.

定义1.4.7. 设 $\{X_n\}$ 是 S 上的一个随机过程, τ 是一个取值于 $\{0, 1, 2, \dots\} \cup \{\infty\}$ 的广义随机变量. 若对任意 $n \geq 0$, $i_0, i_1, \dots, i_n \in S$, 下列两种情况中有且仅有一种成立, 则称 τ 为 $\{X_n\}$ 的一个停时(stopping time), 简称 τ 是停时.

$$\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subseteq \{\tau \leq n\}; \quad (1.4.8)$$

$$\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subseteq \{\tau > n\}. \quad (1.4.9)$$

命题1.4.8. τ 为停时当且仅当对任意 $n \geq 0$, $i_0, i_1, \dots, i_n \in S$, 以下两条中有且仅有一条成立.

$$\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subseteq \{\tau = n\}, \quad (1.4.10)$$

$$\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subseteq \{\tau \neq n\}. \quad (1.4.11)$$

证: 假设 τ 是停时. 如果 (1.4.9) 成立, 那么 (1.4.11) 成立. 否则, (1.4.8) 成立, 此时, 根据停时的定义, $\{X_0 = i_0, \dots, X_{n-1} = i_{n-1}\}$ 包含于 $\{\tau \leq n-1\}$ 或 $\{\tau > n-1\}$, 这结合 (1.4.8) 表明 $\{X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i_n\}$ 包含于 $\{\tau \leq n-1, \tau \leq n\} = \{\tau \leq n-1\} \subseteq \{\tau \neq n\}$ 或 $\{\tau > n-1, \tau \leq n\} = \{\tau = n\}$, 即 (1.4.11) 或 (1.4.10) 成立.

反过来, 若 (1.4.10) 成立, 则 (1.4.8) 成立. 否则, (1.4.11) 成立, 此时, 若 $\{X_0 = i_0\} \subseteq \{\tau = 0\}$, $\{X_0 = i_0, X_1 = i_1\} \subseteq \{\tau = 1\}$, \dots , $\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subseteq \{\tau = n\}$ 之一成立, 则 (1.4.8) 成立. 否则, $\{X_0 = i_0, X_1 = i_1, \dots, X_n = i_n\} \subseteq \{\tau \neq 0, \tau \neq 1, \dots, \tau \neq n\} = \{\tau > n\}$, 即 (1.4.9) 成立. \square

例1.4.9. 对任意 $D \subseteq S$, τ_D 与 σ_D 都是停时.

例1.4.10. 假设 $\{S_n\}$ 是 \mathbb{Z}^1 上的随机游动, $p_{i,i+1} = 1 - p_{i,i-1} = p \in (\frac{1}{2}, 1)$. 对任意 $i \geq 0$, $\varepsilon_i := \sup\{n \geq 0 : X_n = i\}$ 不是停时.

命题1.4.11 (强马氏性). 假设 $\{X_n\}$ 是一个马氏链, 转移矩阵为 \mathbf{P} . 假设 τ 是 $\{X_n\}$ 的一个停时. 在事件 $\{\tau < \infty\}$ 上, 定义

$$Y_m := X_{\tau+m}, \quad \forall m \geq 0; \quad \vec{Z} = (X_0, \dots, X_\tau).$$

那么, 在 $\{\tau < \infty, X_\tau = i\}$ 发生的条件下, 或者在 $\{\tau = n, X_\tau = i\}$ 发生的条件下, 我们都有如下结论: $\{Y_m\}$ 是一个从 i 出发的, 以 \mathbf{P} 为转移矩阵的马氏链, 且它与 \vec{Z} 相互独立.

证: 记 $E_n = \{\tau = n, X_\tau = i\}$, $E = \{\tau < \infty, X_\tau = i\}$. 我们仍然先证明在事件 E_n 发生的条件下, 结论成立, 然后再证明在事件 E 发生的条件下, 结论成立.

假设 $P(E_n) > 0$, 并将条件概率 $P(\cdot|E_n)$ 简记为 $P_n(\cdot)$. 若 E_n 发生, 则 $\tau = n$. 这意味着则 $\vec{Z} = (X_0, \dots, X_n)$. 根据命题 1.4.8, 此时, \vec{Z} 只能取值于

$$I_n = \{\vec{i} = (i_0, \dots, i_n) \in S^{n+1} : (1.4.10) \text{ 成立}\}.$$

并且, 对任意 $\vec{i} \in S^{n+1} \setminus I_n$ 都有 (1.4.11) 成立. 为证明在事件 E_n 发生的条件下, 结论成立, 我们只需验证 (1.4.5), 等价地, 对任意 $m \geq 0$, 任意 $\vec{j} = (j_0, \dots, j_m) \in S^{m+1}$, 任意 $\vec{i} = (i_0, \dots, i_n) \in I_n$,

$$P(\vec{Y} = \vec{j}, \vec{Z} = \vec{i}, E_n) = g(\vec{j})P(\vec{Z} = \vec{i}, E_n). \quad (1.4.12)$$

其中 $\vec{Y} = (Y_0, \dots, Y_m)$, $g(\vec{j}) = \mathbf{1}_{\{j_0=i\}} p_{j_0 j_1} p_{j_1 j_2} \cdots p_{j_{m-1} j_m}$.

当事件 E_n 发生时, 一方面, $\tau = n$, 从而 $\vec{Z} = (X_0, \dots, X_n)$ 且 $\vec{Y} = (X_n, \dots, X_{n+m})$, 另一方面, $X_n = X_\tau = i$. 于是, (1.4.12) 实际上就是

$$\begin{aligned} & P(X_n = j_0, \dots, X_{n+m} = j_m; X_0 = i_0, \dots, X_n = i_n; \tau = n, X_n = i) \\ &= g(\vec{j})P(X_0 = i_0, \dots, X_n = i_n; \tau = n, X_n = i). \end{aligned}$$

由于 $\vec{i} \in I_n$, 即 (1.4.10) 成立, 因此, 上式两边的 $\tau = n$ 都可以去掉. 当 $i_n \neq i$ 时, 上式成立, 因为等号两边都是0; 当 $i_n = i$ 时, 两边的 $X_n = i_n$ 都可以去掉, 于是上式转化为

$$\begin{aligned} & P(X_n = j_0, \dots, X_{n+m} = j_m; X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= g(\vec{j})P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i). \end{aligned}$$

由 $\{X_n\}$ 是马氏链知该式成立, 从而 (1.4.12) 成立.

最后, 仿照命题 1.4.2 的证明的最后一段, 可以证明在 E 发生的条件下, 结论成立. \square

在 $\tau < \infty$ 发生的条件下, 我们可以将 \vec{Z} 视为一条随机的有限步轨道, 其步长是随机变量, 其所有可能取值为 $\cup_{n=0}^{\infty} I_n$ (而不是 $\cup_{n=0}^{\infty} S^{n+1}$), 具体地, 对任意 $\vec{i} = (i_0, \dots, i_n) \in I_n$,

$$P(\vec{Z} = \vec{i}) = P(X_0 = i_0, \dots, X_n = i_n),$$

因为等号两边的事件都蕴含着 $\tau = n$, 从而蕴含着 $\vec{Z} = (X_0, \dots, X_n)$. 当 $\vec{i} \in J_n$ 时, 根据 \vec{Z} 的定义, 左边的事件 $\vec{Z} = \vec{i}$ 蕴含着 $\tau = n$, 从而蕴含着 $\vec{Z} = (X_0, \dots, X_n)$, 因此蕴含着右边的事件 $X_0 = i_0, \dots, X_n = i_n$, 这意味着 $\tau \neq n$, 因为 $\vec{i} \in J_n$. 于是, $P(\vec{Z} = \vec{i}) = 0, \forall n, \forall \vec{i} \in J_n$.

需要注意的是, 在运用强马氏性时, 一定要检查所涉及的随机时刻满足停时的定义. 譬如说, 例 1.4.10 中的 ε_i 不是停时, 而 ε_i 的后续轨道虽然是从 i 出发, 但是它不是随机游动, 因为按照 ε_i 的定义, 其后续轨道的第一步必须往正方向跳.

习题

1. 假设 i, j 是两个不同的状态, $P_i(\tau_j < \infty) = P_j(\tau_i < \infty) = 1$ 且 $P_i(\tau_j < \sigma_i) = P_j(\tau_i < \sigma_j)$.
假设从 i 出发的马氏链在回到 i 之前, 访问了状态 j 总共 ξ 次. 求 ξ 的分布列与期望.

2* 取 $n \in \mathbb{Z}_+$, 令 $\tau \equiv n$. 证明: τ 是停时.

3* 证明 $\tau = \inf\{n \geq n_0 : X_n \in A\}$ 是停时.

4* 假设 τ 为停时, $A \subseteq S$. 令 $\sigma_1 := \inf\{n \geq \tau : X_n \in A\}$,

$$\sigma_2 := \begin{cases} \inf\{n \geq \tau + 1 : X_n \in A\}, & \text{若 } \tau < \infty; \\ \infty, & \text{若 } \tau = \infty. \end{cases}$$

证明: σ_1, σ_2 都是停时.

5* 假设 τ, σ 是停时. 试证明下列随机时间都是停时: $\min\{\tau, \sigma\}, \max\{\tau, \sigma\}, \tau + \sigma$.

§1.5 常返性

在第 §1.3 节中, 我们讨论了从状态 i_0 出发的马氏链是否有可能(即, 以正概率)访问状态 i , 这是可达的概念. 在本节, 我们将研究的问题是: 马氏链是否一定(即, 概率为1)访问状态 i ? 在整条轨道上, 总共访问状态 i 多少次? 访问 i 的所有时间有什么性质?

一、回访时间.

给定状态 i . 沿用 (1.4.1) 和 (1.4.2) 定义的符号. 记

$$\tau_i = \inf\{n \geq 0 : X_n = i\}, \quad \sigma_i = \inf\{n \geq 1 : X_n = i\}.$$

我们考虑从 i 出发的粒子, 即假设 $X_0 = i$. 它首次返回状态 i 的时间(即, 回访时间)就是 σ_i , 将它记为 $T_{i,1}$, 然后我们归纳定义粒子第 r 次回访时间. 具体地, 令 $T_{i,1} := \sigma_i$,

$$T_{i,r} := \begin{cases} \inf\{n \geq T_{i,r-1} + 1 : X_n = i\}, & \text{若 } T_{i,r-1} < \infty; \\ \infty, & \text{若 } T_{i,r-1} = \infty, \end{cases} \quad \forall r \geq 2.$$

根据定义, $T_{i,r}$ 表示马氏链第 r 次回访状态 i 的时间. 假设 $\sigma_i < \infty$, 我们知道粒子在时刻 $T_{i,1}$ 返回状态 i . 根据强马氏性, 粒子的后续运动仍然是一个从 i 出发的马氏链, 并且它跟 σ_i 之前的轨道(特别是, 跟这段轨道的长度 σ_i)是相互独立的. 因此, 它再次返回状态 i 所需的时间长 $T_{i,2} - T_{i,1}$ 与 σ_i 具有相同的分布, 并且是相互独立的. 严格地说, 令 $T_{i,0} := 0$,

$$\sigma_{i,r} := \begin{cases} T_{i,r} - T_{i,r-1} & \text{若 } T_{i,r-1} < \infty; \\ \infty, & \text{若 } T_{i,r-1} = \infty, \end{cases} \quad \forall r \geq 1,$$

其中 $\sigma_{i,1}$ 就是 σ_i .

命题 1.5.1. 对任意 $r \geq 1$ 以及任意 $m_1, \dots, m_r \geq 1$,

$$P_i(\sigma_{i,1} = m_1, \sigma_{i,2} = m_2, \dots, \sigma_{i,r} = m_r) = \prod_{s=1}^r P_i(\sigma_i = m_s). \quad (1.5.1)$$

进一步, 若 $P_i(\sigma_i < \infty) = 1$, 则 $\sigma_{i,1}, \sigma_{i,2}, \sigma_{i,3}, \dots$ 独立同分布.

证: 记

$$A = \{\sigma_{i,2} = m_2, \dots, \sigma_{i,r} = m_r\}.$$

那么, (1.5.1) 的左边就是 $P_i(\sigma_i = m_1, A)$, 根据乘法公式, 它等于 $P_i(\sigma_i = m_1)P(A|\sigma_i = m_1)$. 在事件 $\{\sigma_i < \infty\}$ 上, 令 $Y_n = X_{\sigma_i+n}$, $n \geq 0$. 根据定义, 我们有 $\sigma_{i,1}^{(Y)} = \sigma_{i,2}, \dots, \sigma_{i,r-1}^{(Y)} = \sigma_{i,r}$. 因此, 事件 A 就可以改写为

$$\tilde{A} = \{\sigma_{i,1}^{(Y)} = m_2, \dots, \sigma_{i,r-1}^{(Y)} = m_r\}.$$

请注意, A 就是 \tilde{A} , 我们换符号只是为了强调该事件原本是用 $\{X_n\}$ 表达, 而现在则改用 $\{Y_n\}$ 来表达. 根据强马氏性(推论 1.4.5), 在 $\sigma_i = m_1$ 上, $\{Y_n\}$ 是从 i 出发的马氏链, 因此事件 \tilde{A} 的条件概率就等于事件

$$\hat{A} = \{\sigma_{i,1} = m_2, \dots, \sigma_{i,r-1} = m_r\}$$

的概率. 换句话说,

$$P(\tilde{A} | \sigma_i = m_1) = P_i(\hat{A}).$$

因为 $A = \tilde{A}$, 所以 $P(A | \sigma_i = m_1) = P_i(\hat{A})$, 从而

$$P_i(\sigma_i = m_1, A) = P_i(\sigma_i = m_1)P(A | \sigma_i = m_1) = P_i(\sigma_i = m_1)P_i(\hat{A}).$$

最后, 利用归纳法, 我们可以证明结论成立.

进一步, 若 $P_i(\sigma_i < \infty) = 1$, 则 $\sigma_{i,1}, \sigma_{i,2}, \dots$ 都是取值正整数的随机变量. 于是, (1.5.1) 表明它们独立同分布. \square

注1.5.2. 我们其实可以直接证明 (1.5.1). 我们以 $r = 2$ 为例进行证明. 假设 $\sigma_{i,1} = m, \sigma_{i,2} = n$, 即, 粒子在 m 时刻首次返回状态 i , 之后又经历了时间长 n 第二次返回 i . 这等价于粒子在 $0, m, m+n$ 这三个时刻访问 i , 并且在时间段 $[0, m+n]$ 中的其它时刻访问其它状态. 于是, 我们可以将这个事件用前 $m+n$ 步轨道写出. 将 m 视为现在, 则这个事件可以写为一个过去及现在的事件 C 和一个将来的事件 B 的交集, 其中

$$\begin{cases} B = \{X_{m+r} \neq i, 1 \leq r \leq n-1; X_{m+n} = i\}, \\ C = \{X_0 = i; X_r \neq i, \forall 1 \leq r \leq m-1; X_m = i\}. \end{cases}$$

于是

$$P_i(\sigma_{i,1} = m, \sigma_{i,2} = n) = P_i(C \cap B) = P_i(C)P_i(B|C).$$

一方面, $C = \{\sigma_i = m\}$. 另一方面, 由马氏性(命题 1.1.11),

$$\begin{aligned} P_i(B|C) &= P(X_{m+r} \neq i, 1 \leq r \leq n-1; X_{n+m} = i | X_m = i) \\ &= P(X_r \neq i, 1 \leq r \leq n-1; X_n = i | X_0 = i) = P_i(\sigma_i = n). \end{aligned}$$

于是,

$$P_i(\sigma_{i,1} = m, \sigma_{i,2} = n) = P_i(\sigma_i = m)P_i(\sigma_i = n).$$

这即是 (1.5.1) 的 $r = 2$ 的情形.

考虑状态 i 的回访概率

$$\rho_i := P_i(\sigma_i < \infty). \tag{1.5.2}$$

当 $\rho_i < 1$ 时, $\sigma_{i,1}, \sigma_{i,2}, \dots$ 可以取值 ∞ . 此时, (1.5.1) 不能表明它们独立同分布, 事实上, 根据定义, $\sigma_i = \infty$ 蕴含着 $\sigma_{i,2} = \infty$, 因此它们不是相互独立的. 在(1.5.1) 中, 让 m_1, \dots, m_r 取遍所有正整数, 然后求和, 便得到下面的推论. 它主要针 $\rho_i < 1$ 的情形, 因为当 $\rho_i = 1$ 时, $\sigma_{i,r} < \infty$, $\forall r$, 因此 $T_{i,r} < \infty, \forall r$, 于是等式是平凡的, 因为其两边都等于1.

推论1.5.3. $P_i(T_{i,r} < \infty) = \rho_i^r$.

二、回访总次数.

令

$$V_i = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=i\}} = |\{n \geq 0 : X_n = i\}|.$$

它表示马氏链沿其轨道访问 i 的总次数. 有时, 为了强调所涉及的马氏链是 $\{X_n\}$, 我们也将 V_i 记为 $V_i^{(X)}$. 若马氏链从 i 出发, 则 $V_i + 1$ 就是该马氏链回访 i 的总次数. 直观上, 若回访概率 $\rho_i = 1$, 则粒子出游后必然返回 i , 于是它将一而再、再而三的返回 i , 因此 $V_i = \infty$; 若 $\rho_i < 1$, 则粒子每次出游都以(正)概率 $1 - \rho_i$ 不返回 i , 于是终有一次出游时它不会返回, 因此 $V_i < \infty$. 为此, 我们做一个抛硬币的随机试验, 正面表示回访成功(即, 回访 i), 反面表示回访失败(即, 不回访 i). 与独立抛硬币试验不同的是, 在现在的随机试验中, 只有抛到正面才可以进行下一次抛硬币(因为只有回访成功, 谈及下一次回访的时间才有意义); 一旦抛到反面, 试验就结束. 在该随机试验中, 我们关心的回访次数就是抛到反面之前正面出现的次数, 这与独立抛硬币试验的结果是一样的. 换句话说, V_i 服从参数为 ρ_i 的几何分布. 这个结论本质上就是推论 1.5.3, 因为至少回访 r 次(即, 至少访问 $r+1$ 次)等价于第 r 次回访时间 $T_{i,r}$ 有限, 从而

$$P_i(V_i \geq r+1) = P_i(T_{i,r} < \infty) = \rho_i^r, \quad \forall r \geq 0,$$

而上式表明 V_i 服从参数为 ρ_i 的几何分布.

根据上面的分析, 我们得到如下总结: 当回访概率 $\rho_i = 1$ 时, $P_i(V_i = \infty) = 1$, 此时 $E_i V_i = \infty$; 当回访概率 $\rho_i < 1$ 时, $P_i(V_i < \infty) = 1$, 此时 $E_i V_i = \frac{1}{1-\rho_i} < \infty$. 于是, 如下的二择一法则成立:

$$\begin{aligned} P_i(\sigma_i < \infty) = 1 &\iff P_i(V_i = \infty) = 1 \iff E_i V_i = \infty, \\ P_i(\sigma_i = \infty) = 0 &\iff P_i(V_i < \infty) = 1 \iff E_i V_i < \infty. \end{aligned} \tag{1.5.3}$$

此二择一法则表明, “回访 i 无穷多次”这一事件有0-1律, 即 $P_i(V_i = \infty) = 0$ 或 1. 若此概率为 1, 我们认为从 i 出发的马氏链会经常返回 i , 因此称这样的状态 i 为一个常返态.

定义1.5.4. 若 $P_i(V_i = \infty) = 1$, 则称状态 i 是一个常返态, 或者说 i 是常返的(recurrent). 否则称 i 是一个暂态, 或称 i 是非常返的(transient).

假设 i 是常返态, 等价地, $P_i(\sigma_i < \infty) = 1$, 并且 $P(T_{i,r} < \infty) = 1, \forall r \geq 1$. 根据命题 1.5.1, $\sigma_{i,1}, \sigma_{i,2}, \dots$ 独立同分布. 事实上, 我们还可以证明如下更强的结论.

命题1.5.5. 假设 i 常返, $X_0 = i$. 令

$$\vec{Z}^{(r)} := (X_{T_{i,r-1}}, \dots, X_{T_{i,r}}), \quad \forall r \geq 1.$$

则 $\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots$ 独立同分布.

证: 由 i 常返以及 $X_0 = i$ 知 $\sigma_i < \infty$. 记 $Y_n = X_{\sigma_i+n}$, $n \geq 0$. 那么 $\{Y_n\}$ 第 r 次返回 i 的时间 $T_{i,r}^{(Y)}$ 等于 $T_{i,r+1} - \sigma_i$. 在 $\vec{Z}^{(r)}$ 的定义中, 将 $\{X_n\}$ 改为 $\{Y_n\}$, 于是 $T_{i,r}$ 相应地改为 $T_{i,r}^{(Y)}$, 我们得到 $\vec{Z}^{(\vec{Y},r)}$. 那么, 对任意 $r \geq 1$,

$$\vec{Z}^{(\vec{Y},r)} = (Y_{T_{i,r-1}^{(Y)}}, \dots, Y_{T_{i,r}^{(Y)}}) = (X_{\sigma_i+T_{i,r-1}^{(Y)}}, \dots, X_{\sigma_i+T_{i,r}^{(Y)}}) = (X_{T_{i,r}}, \dots, X_{T_{i,r+1}}) = \vec{Z}^{(r+1)}.$$

根据强马氏性(推论 1.4.5), $\{Y_n\}$ 也是从 i 出发的以 \mathbf{P} 为转移矩阵的马氏链, 且 $\{Y_n\}$ 与 $\vec{Z}^{(1)}$ 相互独立. 根据 $(\vec{Z}^{(2)}, \vec{Z}^{(3)}, \dots) = (\vec{Z}^{(\vec{Y},1)}, \vec{Z}^{(\vec{Y},2)}, \dots)$, 我们得到下面的两条结论:

结论 (a): $(\vec{Z}^{(2)}, \vec{Z}^{(3)}, \dots)$ 与 $(\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots)$ 同分布, 因为 $\{Y_n\}$ 与 $\{X_n\}$ 同分布, 它们都是从 i 出发的马氏链;

结论 (b): $(\vec{Z}^{(2)}, \vec{Z}^{(3)}, \dots)$ 与 $\vec{Z}^{(1)}$ 相互独立, 因为 $\{Y_n\}$ 与 $\vec{Z}^{(1)}$ 相互独立.

一方面, 根据结论(a), $(\vec{Z}^{(r+1)}, \vec{Z}^{(r+2)}, \dots)$ 与 $(\vec{Z}^{(r)}, \vec{Z}^{(r+1)}, \dots)$ 同分布, 因为它们分别是结论(a)中提到的两个序列的从第 r 项开始的后续序列. 由归纳法, 我们得到

结论 (c): $(\vec{Z}^{(r)}, \vec{Z}^{(r+1)}, \dots)$ 与 $(\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots)$ 同分布. 从而, $\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots$ 同分布.

另一方面, 结论(b)描述的性质是序列的第一项与所有后续项相互独立. 由结论(c), 对任意 $r \geq 1$, $(\vec{Z}^{(r)}, \vec{Z}^{(r+1)}, \dots)$ 都满足该性质, 从而 $\vec{Z}^{(r)}$ 与 $(\vec{Z}^{(r+1)}, \vec{Z}^{(r+2)}, \dots)$ 相互独立, $\forall r \geq 1$. 这结合归纳法表明 $\vec{Z}^{(1)}, \dots, \vec{Z}^{(r)}$ 相互独立, $\forall r \geq 1$. 这即是说, $\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots$ 相互独立.

综上, $\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots$ 独立同分布. □

下面我们研究 $\vec{Z}^{(1)}$ 的分布列. 首先, 它的可能取值为

$$E_i = \left\{ \vec{i} = (i_0, i_1, \dots, i_n) : n \geq 1; i_0 = i_n = i; i_1, \dots, i_{n-1} \neq i \right\}.$$

其次, 对任意 $\vec{i} = (i_0, i_1, \dots, i_n) \in E_i$, 都有 $(X_0, \dots, X_n) = \vec{i}$ 或者 $\vec{Z}^{(1)} = \vec{i}$ 都蕴含着 $\sigma_i = n$, 从而蕴含着 $\vec{Z}^{(1)} = (X_0, \dots, X_n)$. 因此,

$$P(\vec{Z}^{(1)} = \vec{i}) = P((X_0, \dots, X_n) = \vec{i}) = p_{\vec{i}},$$

其中

$$p_{\vec{i}} = p_{i_0 i_1} \cdots p_{i_{n-1} i_n}, \quad \forall \vec{i} = (i_0, i_1, \dots, i_n) \in E_i.$$

我们称服从此分布列的有限长的随机轨道为一个状态 i 的游弋. 简单地说, 游弋指的是一个从 i 出发的马氏链, 首次回到 i 之前经历的(有限长的)轨道. 例如, 命题 1.5.5 中的 $\vec{Z}^{(1)}, \vec{Z}^{(2)}, \dots$ 就是独立同分布的游弋.

不难验证,

$$P_i(\sigma_i = n) = \sum_{\vec{i} \in S^{n+1} \cap E_i} p_{\vec{i}}, \quad P_i(\sigma_i < \infty) = \sum_{\vec{i} \in E_i} p_{\vec{i}}.$$

这表明当 i 非常返时, $\pi_{\vec{i}}, \forall i \in E_i$ 给出的不是分布列, 因为它不满足求和等于1 的归一化条件.

三、常返与互通类.

命题 1.5.6. 假设 i 常返且 $i \rightarrow j$. 则 j 常返且 $j \rightarrow i$. 进一步, $P_i(V_j = \infty) = P_j(V_i = \infty) = 1$. 从而, $P_i(\tau_j < \infty) = P_j(\tau_i < \infty) = 1$.

证: 当 $j = i$ 时, 根据常返的定义, 结论成立. 下面假设 $j \neq i$. 我们首先证明结论1: $P_i(V_j = \infty) = 1$, 这表明 $P_i(\tau_j < \infty) = 1$, 然后证明结论2: $P_j(V_i = \infty) = 1$, 这表明 $P_j(\tau_i < \infty) = 1$, 从而 $j \rightarrow i$, 最后证明结论3: j 常返.

考虑从 i 出发的轨道, 即, 假设 $X_0 = i$. 考虑事件 A_r = “状态 j 出现在状态 i 的第 r 个游弋中”, 即

$$A_r := \{\exists n \in [T_{i,r-1}, T_{i,r}] \text{ 使得 } X_n = j\}.$$

由命题 1.5.5, $\mathbf{1}_{A_1}, \mathbf{1}_{A_2}, \dots$ 独立同分布, 从而, $E\mathbf{1}_{A_n} = P(A_n) = p, \forall n$, 其中, $p = P(A_1)$.

往证 $p > 0$. 这是因为

$$\{\tau_j < \infty\} = \{\exists n \geq 0 \text{ 使得 } X_n = j\} = \bigcup_{r=1}^{\infty} A_r,$$

而 $i \rightarrow j$ 表明上述事件的概率为正. 根据概率的次可列可加性,

$$\sum_{r=1}^{\infty} P_i(A_r) \geq P_i(\tau_j < \infty) > 0,$$

从而 $p > 0$.

由强大数定律,

$$P_i \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{r=1}^n \mathbf{1}_{A_r} = p \right) = 1.$$

由 $V_j \geq \sum_{r=1}^{\infty} \mathbf{1}_{A_r}$ 可推出 $P_i(V_j = \infty) = 1$. 也就是说, 状态 j 将会出现在无穷多个游弋中, 从而出现无穷多次. 这即是结论1.

由于 $V_j = \infty$ 蕴含着 $\tau_j < \infty$, 因此 $P_i(\tau_j < \infty) \geq P_i(V_j = \infty) = 1$. 令 $Y_n = X_{\tau_j + n}, n \geq 0$. 那么,

$$V_i^{(Y)} := \sum_{n=0}^{\infty} \mathbf{1}_{\{Y_n=i\}} = \sum_{n=\tau_j}^{\infty} \mathbf{1}_{\{X_n=i\}}.$$

因为 $V_i^{(Y)} \leq V_i \leq \tau_j + V_i^{(Y)}$ 且 $\tau_j < \infty$, 所以 $V_i = \infty$ 当且仅当 $V_i^{(Y)} = \infty$. 根据强马氏性, $P_i(V_i^{(Y)} = \infty | \tau_j < \infty) = P_j(V_i = \infty)$. 再由 i 常返且 $\tau_j < \infty$ 知,

$$1 = P_i(V_i = \infty, \tau_j < \infty) = P_i(\tau_j < \infty)P_i(V_i^{(Y)} = \infty | \tau_j < \infty) = P_i(\tau_j < \infty)P_j(V_i = \infty),$$

其中 $P_i(\tau_j < \infty) = 1$. 从而, $P_j(V_i = \infty) = 1$. 这即是结论2.

类似地, 考虑 $V_j^{(Y)}$, 我们也有 $V_j = \infty$ 当且仅当 $V_j^{(Y)} = \infty$. 由已证的结论1 以及 $\tau_j < \infty$ 知

$$1 = P_j(V_j = \infty, \tau_j < \infty) = P_i(\tau_j < \infty)P_i(V_j^{(Y)} = \infty | \tau_j < \infty) = P_i(\tau_j < \infty)P_j(V_j = \infty).$$

这表明 $P_j(V_j = \infty) = 1$, 即 j 常返. 这即是结论3. \square

上面的命题 1.5.6 表明常返是互通类的性质, 若某互通类中的所有状态都常返, 则我们称之为常返类, 否则称之为非常返类. 进一步, 常返类一定是闭集. 具体地, 若一个互通类 C 不是闭集, 那么存在 $i \in C, j \notin C$ 使得 $p_{ij} > 0$, 因此 $i \rightarrow j$ 但 $P_j(\tau_i < \infty) = P_j(V_i = \infty) = 0$, 由命题 1.5.6 知 i 非常返. 基于此, 我们只需要考虑闭的互通类的常返性. 此时, 我们可以将马氏链限制在这个闭的互通类上, 得到一个状态空间小一点的不可约马氏链. 因此, 在讨论常返性时, 我们不妨假设马氏链不可约. 此时, 所有状态或者都常返, 或者都非常返, 我们也可称该马氏链是常返的, 或非常返的. 由 (1.5.3), 我们可以通过求回访概率 $P_i(\sigma_i < \infty)$ 或求平均访问时间 $E_i V_i$ 来判别状态 i 的常返性, 所用的工具分别为吸收概率和格林函数.

四、吸收概率与首步分析法.

记

$$\rho_{ij} = P_i(\tau_j < \infty).$$

ρ_{ij} 就是定义 1.3.1 中涉及的概率, 即, $i \rightarrow j$ 指的是 $\rho_{ij} > 0$.

定义 1.5.7. 如果 $p_{ii} = 1$, 则称 i 为一个吸收态.

状态 i 是吸收态的意思就是: 粒子一旦到达状态 i , 就被 i 吸收了. 因为一旦有 $X_n = i$, 我们就知道 $X_{n+m} = i, \forall m \geq 0$. 在研究 ρ_{ij} 时, 我们可以将状态 j 想象为一个吸收态, 换句话说, 考虑以

$$\hat{p}_{ik} = p_{ik}, \quad \forall i \neq j, \forall k, \quad \hat{p}_{jj} = 1$$

为转移矩阵的马氏链 $\{\hat{X}_n\}$ 及其首达时 $\tau_j^{(\hat{X})}$. 那么,

$$P_i(\tau_j < \infty) = P_i(\tau_j^{(\hat{X})} < \infty), \quad \forall i. \tag{1.5.4}$$

上式的验证并不困难, 留为习题. 因此, ρ_{ij} 表示从 i 出发的粒子最终被状态 j 吸收的概率. 它被称为吸收概率.

取定一个状态 o . 首先, 我们考虑从 i 出发的马氏链能到达状态 o 的概率 $P_i(\tau_o < \infty)$. 如前所述, 若 $i = o$, 则 $\tau_o = 0$, 从而 $P_i(\tau_o < \infty) = 1$. 下面假设 $i \neq o$, 此时 $\tau_o = \sigma_o$. “首步分析法” 指让马氏链先走一步, 根据第一步 X_1 的具体状态进行分情况讨论. 运用全概公式, 得到

$$P_i(\tau_o < \infty) = P_i(\sigma_o < \infty) = \sum_{j \in S} P_i(X_1 = j) P_i(\sigma_o < \infty | X_1 = j).$$

考虑 $Y_n = X_{1+n}$, $\forall n \geq 0$. 根据例 1.4.1, $\sigma_o = \sigma_o^{(X)} = 1 + \tau_o^{(Y)}$, 从而 $\sigma_o < \infty$ 当且仅当 $\tau_o^{(Y)} < \infty$. 根据马氏性(命题 1.1.11), 在 $X_1 = j$ 的条件下, $\{Y_n\}$ 是从 j 出发的马氏链. 于是,

$$P_i(\sigma_o < \infty | X_1 = j) = P_i(\tau_o^{(Y)} < \infty | X_1 = j) = P_j(\tau_o < \infty).$$

将 $x_i = P_i(\tau_o < \infty)$, $\forall i \in S$ 视为一组未知数. 根据上面的分析, 它们满足下面的方程组:

$$x_i = \sum_{j \in S} p_{ij} x_j, \quad \forall i \neq o; \quad x_o = 1, \quad (1.5.5)$$

其中, $x_o = 1$ 可被视为边界条件. 注意 $x_o = P_o(\tau_o < \infty)$, 而不是 $P_o(\sigma_o < \infty)$.

事实上, 对任意 S 的子集 D , 我们都可以将 D 想象为一个由吸收态组成的区域. 那么, $x_i := P_i(\tau_D < \infty)$ 就表示从 i 出发的粒子最终被区域 D 吸收的概率, 它也被称为吸收概率, 其中 τ_D 是粒子首达 D 的时刻. 类似地, 对 S 的任意子集 D , 这一组吸收概率满足如下方程组:

$$x_i = \sum_{j \in S} p_{ij} x_j, \quad \forall i \notin D; \quad x_i = 1, \quad \forall i \in D. \quad (1.5.6)$$

其中, $x_i = 1, \forall i \in D$ 是边界条件. 当 $D = \{o\}$ 时, 上式就是 (1.5.5).

命题 1.5.8. 吸收概率是方程组 (1.5.6) 最小的非负解, 即, 若 $\{\tilde{x}_i, i \in S\}$ 也是方程组 (1.5.6) 的非负解, 那么, $\tilde{x}_i \geq x_i, \forall i$.

证: 假设 $\{\tilde{x}_i, i \in S\}$ 也是方程组 (1.5.6) 的非负解. 那么, 根据边界条件, 对任意 $i \in D$, $\tilde{x}_i = x_i = 1$ 从而 $\tilde{x}_i \geq x_i$ 成立. 下面假设 $i \notin D$, 往证 $\tilde{x}_i \geq x_i$. 根据方程组,

$$\tilde{x}_i = \sum_{j \in S} p_{ij} \tilde{x}_j = \sum_{j \in D} p_{ij} \tilde{x}_j + \sum_{j \notin D} p_{ij} \tilde{x}_j.$$

在右边第一项中代入边界条件 $\tilde{x}_j = 1, \forall j \in D$, 并对第二项利用方程组进行迭代处理, 我们有

$$\tilde{x}_i = \sum_{j \in D} p_{ij} + \sum_{j \notin D} p_{ij} \left(\sum_{k \in S} p_{jk} \tilde{x}_k \right).$$

继续将 $k \in S$ 分成 $k \in D$ 和 $k \notin D$, 并对前者代入边界条件, 对后者进行迭代, …… 我们推出, 对任意 $n \geq 1$,

$$\begin{aligned} \tilde{x}_i &= \sum_{i_1 \in D} p_{ii_1} + \sum_{i_1 \notin D, i_2 \in D} p_{ii_1} p_{i_1 i_2} + \cdots + \sum_{i_1, \dots, i_{n-1} \notin D, i_n \in D} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \\ &\quad + \sum_{i_1, \dots, i_n \notin D} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} \tilde{x}_{i_n}. \end{aligned}$$

根据 $\tilde{x}_{i_n} \geq 0$ 的假设, 上式最后一项非负. 由于

$$\begin{aligned}\sum_{i_1 \in D} p_{ii_1} &= P_i(X_1 \in D) = P_i(\tau_D = 1), \\ \sum_{i_1 \notin D, i_2 \in D} p_{ii_1} p_{i_1 i_2} &= P_i(X_1 \notin D, X_2 \in D) = P_i(\tau_D = 2), \\ &\dots \\ \sum_{i_1, \dots, i_{n-1} \notin D, i_n \in D} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} &= P_i(X_1, \dots, X_{n-1} \notin D, X_n \in D) = P_i(\tau_D = n),\end{aligned}$$

因此

$$\tilde{x}_i \geq P_i(\tau_D = 1) + P_i(\tau_D = 2) + \cdots + P_i(\tau_D = n) = P_i(\tau_D \leq n), \quad \forall n.$$

最后, 令 $n \rightarrow \infty$ 知 $\tilde{x}_i \geq P_i(\tau_D < \infty) = x_i, \forall i \notin D$. \square

类似地, 我们可以考虑首达区域 D 的平均时间 $E_i \tau_D$. 当 $D = \{o\}$ 时, 就是首达状态 o 的平均时间. 若 $i \in D$, 则 $E_i \tau_D = 0$; 若 $i \notin D$, 则采用首步分析法, 考虑 $Y_n = X_{1+n}, \forall n$. 因为 $\sigma_D = 1 + \tau_D^{(Y)}$, 所以

$$E_i \tau_D = E_i \sigma_D = \sum_{j \in S} p_{ij} E_i(\sigma_D | X_1 = j),$$

其中 $E_i(\sigma_D | X_1 = j) = E_i(1 + \tau_D^{(Y)} | X_1 = j) = 1 + E_i(\tau_D^{(Y)} | X_1 = j) = 1 + E_i \tau_D$. 于是, 将 $y_i := E_i \tau_D, \forall i \in S$ 视为一组未知数, 则它满足下面的方程组:

$$y_i = 1 + \sum_{j \in S} p_{ij} y_j, \quad \forall i \notin D; \quad y_i = 0, \quad \forall i \in D. \quad (1.5.7)$$

事实上, $y_i, i \in S$ 是(1.5.7) 最小的非负解, 证明与命题 1.5.8 的类似, 留为习题.

下面我们讨论如何用吸收概率判别常返性. 如前所述, 我们不妨假设马氏链不可约, 否则该马氏链非常返. 取定状态 o . 一方面, 如果 o 常返, 由命题 1.5.6, 那么 $P_o(\tau_o < \infty) = 1, \forall i$, 即, 所有的吸收概率都是1. 另一方面, 如果所有的吸收概率都为1, 那么, 根据首步分析法我们有

$$P_o(\sigma_o < \infty) = \sum_{i \in S} p_{oi} P_o(\tau_o^{(Y)} < \infty | X_1 = i) = \sum_{i \in S} p_{oi} P_i(\tau_o < \infty) = 1.$$

因此, o 常返当且仅当所有的吸收概率都是1.

由于吸收概率 $P_i(\tau_o < \infty), \forall i$ 是方程组 (1.5.5) 最小的非负解, 因此它们恒为1 等价于该方程没有其它 $([0,1]$ 中的)解, 也就是说, o 常返当且仅当方程 (1.5.5) (在 $[0,1]$ 中)仅有恒为1 的解.

例1.5.9 (例 1.2.12 续). 假设 $\{X_n\}$ 为 \mathbb{Z}_+ 上的不可约生灭链, 出生概率为 b_i , 死亡概率为 d_i . 记 $x_i = P_i(\tau_0 < \infty)$, 则根据 (1.5.5), $x_i = b_i x_{i+1} + d_i x_{i-1}, \forall i \geq 1$, 从而

$$x_i - x_{i+1} = (x_{i-1} - x_i) \frac{d_i}{b_i} = \cdots = (1 - x_1) \frac{d_1 \cdots d_i}{b_1 \cdots b_i}. \quad (1.5.8)$$

等式两边对 i 求和知

$$1 - x_{i+1} = R_i(1 - x_1),$$

其中

$$R_0 := 1, \quad R_i := 1 + \sum_{k=1}^i \frac{d_1 \cdots d_k}{b_1 \cdots b_k}.$$

当 $i \rightarrow \infty$ 时, R_i 单调上升到

$$R := 1 + \sum_{k=1}^{\infty} \frac{d_1 \cdots d_k}{b_1 \cdots b_k}.$$

若 $R < \infty$, 取 $x_1 = \frac{R-1}{R}$, 则 $0 < x_1 < 1$, 并且 $0 < R_i(1 - x_1) = \frac{R_i}{R} < 1$, 于是

$$x_i := 1 - \frac{R_{i-1}}{R}, \quad i = 1, 2, \dots$$

就是 (1.5.5) 的一个在 $[0, 1]$ 中的不恒为 1 的解, 从而该马氏链非常返.

若 $R = \infty$, 则

$$1 - x_1 = \frac{1}{R_i}(1 - x_{i+1}) \leq \frac{1}{R_i} \xrightarrow{i \rightarrow \infty} 0.$$

于是 $x_1 = 1$, 从而 x_i 恒为 1, 马氏链常返. 因此, 该马氏链常返当且仅当 $R = \infty$.

例1.5.10 (例 1.1.10 续, 规则树 \mathbb{T}^d 上的 λ -biased 随机游动). 规则树 \mathbb{T}^d 上的 λ -biased 随机游动指一个马氏链, 其转移概率为: 对根点 o 的任意子结点 j , $p_{oj} := 1/d$; 对任意 $i \neq o$,

$$p_{ij} = \frac{\lambda}{\lambda + d}, \quad j = i_0; \quad p_{ij} = \frac{1}{\lambda + d}, \quad j \in \{i_1, \dots, i_d\},$$

其中 i_0 是 i 的父结点, i_1, \dots, i_d 是 i 的 d 个子结点.

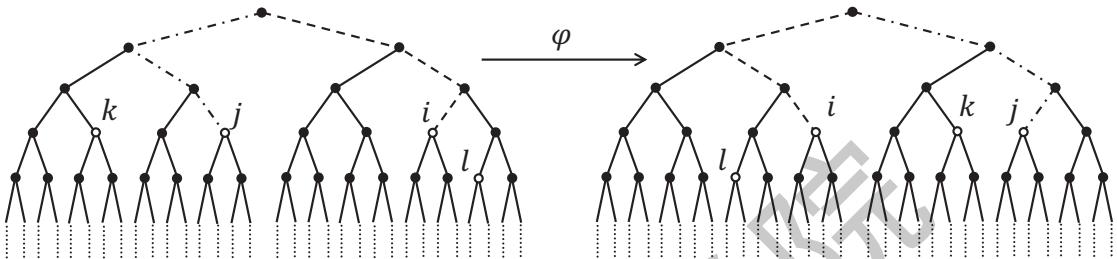
如图 1.14 所示, \mathbb{T}^d 具有如下的球对称性: 假设结点 i, j 满足 $|i| = |j|$, 那么, i, j 的位置可以交换. 具体地, 我们可以建立 \mathbb{T}^d 到自身的一一映射 φ , 交换 i 与 j 的位置. 譬如, 图 1.14 显示的是 $d = 2$ 的情形, 左图是原图, 右图为它在 φ 映射下的像. 于是, 根点 o 到 i 的路径(图中的虚线)与 o 到 j 的路径(图中的点划线)也交换了. 进一步, 如果 $\{X_n\}$ 是从 i 出发的马氏链, $Y_n = \varphi(X_n)$, $\forall n \geq 0$, 那么 $\{Y_n\}$ 就是从 j 出发的马氏链.

根据首达时的定义, $\tau_o^{(X)} = \tau_o^{(Y)}$. 这表明对应的吸收概率相等, 即 $P_i(\tau_o < \infty) = P_j(\tau_o < \infty)$. 因此, 吸收概率可以简化为 $|i|$ 的函数, 即存在 a_0, a_1, a_2, \dots 使得 $P_i(\tau_o < \infty) = a_{|i|}$. 而方程 (1.5.5) 则变为

$$a_n = \frac{\lambda}{\lambda + d} a_{n-1} + \frac{d}{\lambda + d} a_{n+1}, \quad \forall n \geq 1; \quad a_0 = 1.$$

我们将其中的递归式变形为 $db_n = \lambda b_{n-1}$, 其中 $b_n = a_n - a_{n+1}$. 令 $\kappa = \frac{\lambda}{d}$. 解得 $b_n = \kappa b_{n-1} = \dots = \kappa^n b_0$. 因此,

$$a_0 - a_n = b_0 + b_1 + \dots + b_{n-1} = (1 + \kappa + \dots + \kappa^{n-1}) b_0.$$

图 1.14 \mathbb{T}^2 , $|i| = |j| = 3$.

当 $\lambda < d$ 时, $\kappa < 1$. 此时, 取 $b_0 = 1 - \kappa$, 得到 $a_n = \kappa^n$, $n \geq 1$, 这便是方程 (1.5.5) 的非负最小解, 因此该马氏链非常返. 进一步, 由命题 1.5.8, $P_i(\tau_o < \infty) = \kappa^{|i|}$, $\forall i \in \mathbb{T}^d$.

当 $\lambda \geq d$ 时, 为保证 $0 \leq a_n \leq 0$, $\forall n$, 只能取 $b_0 = 0$. 换句话说, 方程 (1.5.5) 的非负最小解就是 $a_n \equiv 1$, $\forall n$. 因此, 该马氏链常返.

最后, 我们介绍一个结论. 虽然它与常返性没有太直接的关系, 但是我们可以在证明中看到首步分析法的巧妙应用以及 \mathbb{Z}^2 上简单随机游动的对称性.

命题1.5.11. 考虑 \mathbb{Z}^2 上的简单随机游动, 记 $o = (0, 0)$, $e_1 = (1, 0)$. 那么,

$$P_o(\tau_{e_1} < \sigma_o) = \frac{1}{2}.$$

证: 考虑从 o 出发的简单随机游动 $\{S_n\}$. 由于 S_1 在 $B = \{e_1, e_2, -e_1, -e_2\}$ 中等可能选取, 其中 $e_2 = (0, 1)$, 因此根据首步分析法,

$$P_o(\sigma_o < \tau_{e_1}) = \frac{1}{4} \sum_{i \in B} P_i(\tau_0 < \tau_{e_1}). \quad (1.5.9)$$

记 $B_n = \{y \in \mathbb{Z}^2 \setminus \{0\} : |y_1|, |y_2| \in [n, 2n]\}$, 并将其上的均匀分布记为 μ_n . 由 \mathbb{Z}^2 常返知, $P_{\mu_n}(\tau_B < \infty) = 1$. 下面, 我们考虑 S_{τ_B} . 若 S_0 服从 μ_n , 则该简单随机游动在东、南、西、北四个方向上是对称的. 从而

$$P_{\mu_n}(S_{\tau_B} = i) = \frac{1}{4}, \quad \forall i \in B.$$

因为 $S_0 \in B_n$ 且 $e_1 \in B$, 所以 $\tau_0 \geq \tau_B$ 且 $\tau_{e_1} \geq \tau_B$. 由强马氏性(命题 1.4.11), 在 $S_{\tau_B} = i$ 的条件下, τ_B 之后的过程 $Y_n = S_{\tau_B+n}$, $n \geq 0$ 是一个从 i 出发的简单随机游动. 由全概公式,

$$\begin{aligned} P_{\mu_n}(\tau_0 < \tau_{e_1}) &= \sum_{i \in B} P_{\mu_n}(S_{\tau_B} = i) P_{\mu_n}(\tau_0 < \tau_{e_1} | S_{\tau_B} = i) \\ &= \frac{1}{4} \sum_{i \in B} P_i(\tau_0 < \tau_{e_1}) = P_o(\sigma_o < \tau_{e_1}), \end{aligned}$$

其中, 最后一个等号就是(1.5.9). 故 $P_{\mu_n}(\tau_0 > \tau_{e_1}) = P_o(\sigma_o > \tau_{e_1})$. 往证

$$\lim_{n \rightarrow \infty} (P_{\mu_n}(\tau_0 < \tau_{e_1}) - P_{\mu_n}(\tau_{e_1} < \tau_0)) = 0. \quad (1.5.10)$$

如果上式成立, 那么 $P_o(\sigma_o < \tau_{e_1}) = P_o(\tau_{e_1} < \sigma_o)$, 从而 $P_o(\sigma_o < \tau_{e_1}) = \frac{1}{2}$.

下面, 我们证明 (1.5.10). 首先, 根据 μ_n 在东、西方向上的对称性, $P_{\mu_n}(\tau_{e_1} < \tau_0) = P_{\mu_n}(\tau_{-e_1} < \tau_0)$. 进一步, 根据空间的平移不变性, $P_j(\tau_{-e_1} < \tau_0) = P_{j+e_1}(\tau_0 < \tau_{e_1})$. 当 j 取遍 B_n 时, $j + e_1$ 取遍 $C_n = \{j + e_1 : j \in B_n\}$. 从而,

$$P_{\mu_n}(\tau_{e_1} < \tau_0) = \frac{1}{|B_n|} \sum_{j \in C_n} P_j(\tau_0 < \tau_{e_1}),$$

其中, 我们用到了 $|C_n| = |B_n|$. 然后,

$$\begin{aligned} |P_{\mu_n}(\tau_0 < \tau_{e_1}) - P_{\mu_n}(\tau_{e_1} < \tau_0)| &= \frac{1}{|B_n|} \left| \sum_{j \in B_n} P_j(\tau_0 < \tau_{e_1}) - \sum_{j \in C_n} P_j(\tau_0 < \tau_{e_1}) \right| \\ &= \frac{1}{|B_n|} \left| \sum_{j \in B_n \setminus C_n} P_j(\tau_{-e_1} < \tau_0) - \sum_{j \in C_n \setminus B_n} P_j(\tau_{-e_1} < \tau_0) \right| \leq \frac{|B_n \setminus C_n| + |C_n \setminus B_n|}{|B_n|}. \end{aligned}$$

最后, 令 $n \rightarrow \infty$, 上式右边趋于0, 从而 (1.5.10) 成立. \square

五、格林函数.

对任意 $i, j \in S$, 令

$$G_{ij} = E_i V_j = E_i \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=j\}} = \sum_{n=0}^{\infty} P_i(X_n = j) = \sum_{n=0}^{\infty} p_{ij}^{(n)},$$

并称之为马氏链的格林函数. 根据二择一法则 (1.5.3), 状态 i 常返当且仅当 $G_{ii} = \infty$.

例1.5.12. (例1.1.6续) 假设 $\{S_n\}$ 是 d 维简单随机游动. 那么, 当 $d = 1$ 或 2 时, $\{S_n\}$ 常返; 当 $d \geq 3$ 时, $\{S_n\}$ 非常返.

解: 将 \mathbb{Z}^d 的原点记为0. 我们将证明, 当 $d = 1$ 或 2 时, $G_{00} = \infty$; 当 $d \geq 3$ 时, $G_{00} < \infty$. 不妨假设 $S_0 = 0$.

不难看出, d 维简单随机游动经过奇数步回到原点的概率为0, 即 $S_{2n+1} = 0, \forall n \in \mathbb{Z}_+$. 如果它经过 $2n$ 步回到原点, 那么对任意 $r \in \{1, \dots, d\}$, 它在第 r 个方向上往前、往后必须走相等的步数, 记为 n_r 步. 于是,

$$p_{00}^{(2n)} = P_0(S_{2n} = 0) = \sum_{n_1 + \dots + n_d = n} \frac{(2n)!}{(n_1!)^2 \cdots (n_d!)^2} \cdot \frac{1}{(2d)^{2n}}. \quad (1.5.11)$$

在上面的求和以及下文中, n_1, n_2, \dots, n_d 都只取非负整数.

当 $d = 1$ 时, 我们推出

$$p_{00}^{(n)} = C_{2n}^n \frac{1}{2^{2n}} = \frac{(2n)!}{(n!)^2} \cdot \frac{1}{2^{2n}}.$$

对任意 $\varepsilon > 0$, 存在 $\delta > 0$ 使得

$$1 - \varepsilon < \frac{1 - \delta}{(1 + \delta)^2} \quad \text{且} \quad \frac{1 + \delta}{(1 - \delta)^2} < 1 + \varepsilon.$$

根据Stirling公式:

$$\lim_{n \rightarrow \infty} \frac{n!}{(n/e)^n \sqrt{2\pi n}} = 1. \quad (1.5.12)$$

存在 n_0 使得

$$1 - \delta \leq \frac{n!}{(n/e)^n \sqrt{2\pi n}} \leq 1 + \delta, \quad \forall n \geq n_0.$$

于是,

$$(1 - \varepsilon) \frac{1}{\sqrt{\pi n}} \leq C_{2n}^n \frac{1}{2^{2n}} \leq (1 + \varepsilon) \frac{1}{\sqrt{\pi n}}, \quad \forall n \geq n_0. \quad (1.5.13)$$

从而 $G_{00} = \infty$, 即 $\{S_n\}$ 常返.

当 $d = 2$ 时,

$$\begin{aligned} p_{00}^{(2n)} &= \sum_{n_1+n_2=n} \frac{(2n)!}{(n_1!)^2 (n_2!)^2} \cdot \frac{1}{4^{2n}} = \frac{(2n)!}{n! n!} \cdot \frac{1}{4^{2n}} \sum_{n_1+n_2=n} \frac{n!}{n_1! n_2!} \cdot \frac{n!}{n_2! n_1!} \\ &= C_{2n}^n \cdot \frac{1}{4^{2n}} \sum_{n_1=0}^n C_n^{n_1} C_n^{n-n_1} = C_{2n}^n \cdot \frac{1}{4^{2n}} \cdot C_{2n}^n = \left(C_{2n}^n \frac{1}{2^{2n}} \right)^2. \end{aligned}$$

这结合 (1.5.13) 表明 $G_{00} = \infty$, 即 $\{S_n\}$ 常返.

当 $d \geq 3$ 时, 我们可以证明当 n 充分大时,

$$P_0(S_{2n} = 0) \leq C_d \cdot n^{-d/2}, \quad (1.5.14)$$

其中 C_d 是依赖于 d 的常数. 我们以 $d = 3$ 为例进行证明. 将 \mathbb{Z}^d 中的三个方向分别称为 x, y, z . 考虑一个有三个面的公平色子, 投掷结果可能为 x, y, z , 以及三个公平硬币, 分别称为 x -硬币, y -硬币, z -硬币. 下面, 我们独立地投掷色子, 然后抛这三个硬币之一, 并按照按如下操作运行 \mathbb{Z}^3 上的简单随机游动. 每一次, 先看投掷色子的结果, 如果投掷结果为 w (其中 w 可为 x, y, z 之一), 那么再抛 w -硬币, 抛到正面则粒子在 w 方向上前进一步, 抛到反面则在 w 方向上后退一步. 将分别由硬币 x, y, z 产生的一维简单随机游动记为 $\{X_m\}, \{Y_m\}, \{Z_m\}$, 它们都从 0 出发. 假设 $2n$ 次操作后, 粒子在 x, y, z 方向上分别运动了 K, L, M 步. 那么, K, L, M 都服从二项分布 $B(2n, \frac{1}{3})$. 请注意它们不相互独立, 因为它们满足一个约束条件 $K + L + M = 2n$. 值得注意的是, 投色子, 抛 x -硬币, 抛 y -硬币, 抛 z -硬币是相互独立地进行的, 因此, (K, L, M) , $\{X_m\}, \{Y_m\}, \{Z_m\}$ 是相互独立的. 由于 $S_{2n} = (X_K, Y_L, Z_M)$, 我们有 $S_n = 0$ 当且仅当 $X_K = Y_L = Z_M = 0$.

一方面,

$$\begin{aligned} p_{00}^{(2n)} &= P(X_K = Y_L = Z_M = 0) \\ &= \sum_{n_1+n_2+n_3=n} P(K = 2n_1, L = 2n_2, M = 2n_3)P(X_{2n_1} = Y_{2n_2} = Z_{2n_3}), \end{aligned} \quad (1.5.15)$$

其中, 根据 (K, L, M) 服从多项分布,

$$P(K = 2n_1, L = 2n_2, M = 2n_3) = \frac{(2n)!}{(2n_1)!(2n_2)!(2n_3)!} \cdot \frac{1}{3^{2n}},$$

又根据 $\{X_m\}, \{Y_m\}, \{Z_m\}$ 是相互独立的一维简单随机游动,

$$P(X_{2n_1} = Y_{2n_2} = Z_{2n_3} = 0) = C_{2n_1}^{n_1} \cdot \frac{1}{2^{2n_1}} \times C_{2n_2}^{n_2} \cdot \frac{1}{2^{2n_2}} \times C_{2n_3}^{n_3} \cdot \frac{1}{2^{2n_3}}. \quad (1.5.16)$$

将上面两式代入 (1.5.15), 我们得到

$$\begin{aligned} p_{00}^{(2n)} &= \sum_{n_1+n_2+n_3=n} \frac{(2n)!}{(2n_1)!(2n_2)!(2n_3)!} \cdot \frac{1}{3^{2n}} \times C_{2n_1}^{n_1} \cdot \frac{1}{2^{2n_1}} \times C_{2n_2}^{n_2} \cdot \frac{1}{2^{2n_2}} \times C_{2n_3}^{n_3} \cdot \frac{1}{2^{2n_3}} \\ &= \sum_{n_1+n_2+n_3=n} \frac{(2n)!}{(n_1!)^2(n_2!)^2(n_3!)^2} \cdot \frac{1}{(2 \cdot 3)^{2n}}, \end{aligned}$$

这即是 (1.5.11) 在 $d = 3$ 时的表达式.

另一方面, 由于 $K \sim B(2n, \frac{1}{3})$, 其期望为 $EK = \frac{2n}{3}$, 因此它不低于 $\frac{n}{3}$ 的概率是很大的, 同理, L 与 M 也一样. 令

$$A = \left\{ K, L, M \geq \frac{n}{3} \right\},$$

并进行如下分解:

$$P_0(S_{2n} = 0) = P(A^c, S_{2n} = 0) + P(A, S_{2n} = 0). \quad (1.5.17)$$

对上式中的第一项, 我们将证明其概率不超过 $n^{-3/2}$; 对第二项, 我们将用一维简单随机游动的结果进行估计. 具体地, 关于 (1.5.17) 右边的第一项, 我们用 4 阶切比雪夫不等式, 得到

$$P(A^c) \leq 3P\left(K < \frac{n}{3}\right) \leq 3P\left(|K - EK| > \frac{n}{3}\right) \leq \frac{3E(K - EK)^4}{(n/3)^4} \leq \frac{3000}{n^2}.$$

其中, 在第一个等式中我们用到 K, L, M 同分布, 而最后一个不等式是因为 $K \sim B(2n, \frac{1}{3})$, 计算的细节留为习题. 需要说明的是: 我们的目标是证明 (1.5.14), 而 4 阶切比雪夫不等式只能证明 $P(A^c) = O(n^{-2})$. 因此, 当 $d \geq 5$ 时, 需要用别的办法给出更精细的估计, 这部分证明留为习题.

关于 (1.5.17) 右边的第二项, 假设 K, L, M 分别取值 $2n_1, 2n_2, 2n_3$. 当 $2n_1, 2n_2, 2n_3 \geq 2n/3$ 时,

$$\begin{aligned} &P(K = 2n_1, L = 2n_2, M = 2n_3, S_{2n} = 0) \\ &= P(K = 2n_1, L = 2n_2, M = 2n_3)P(X_{2n_1} = Y_{2n_2} = Z_{2n_3} = 0), \end{aligned} \quad (1.5.18)$$

取 $\varepsilon = \frac{1}{2}$. 根据 (1.5.13) 与 (1.5.16), 当 $n \geq 3n_0$ 时,

$$P(X_{2n_1} = Y_{2n_2} = Z_{2n_3} = 0) \leq 8 \cdot \frac{1}{\sqrt{\pi n_1}} \cdot \frac{1}{\sqrt{\pi n_2}} \cdot \frac{1}{\sqrt{\pi n_3}} \leq \frac{8}{\sqrt{n^3}}, \quad (1.5.19)$$

其中, 最后一个不等式用到了 $n_i \geq \frac{n}{3}$ 从而 $\pi n_i \geq n$. 将 (1.5.19) 代入 (1.5.18), 并对 n_1, n_2, n_3 求和. 我们有

$$P(A, S_n = 0) \leq \frac{8}{\sqrt{n^3}} P(A) \leq \frac{8}{\sqrt{n^3}}.$$

综上, 当 $n \geq \max\{2n_0, 3000^2\}$ 时,

$$p_{00}^{(2n)} \leq \frac{3000}{n^2} + \frac{3}{\sqrt{n^3}} \leq \frac{4}{\sqrt{n^3}}.$$

这即是 (1.5.14). 它表明 $G_{00} < \infty$, 从而 $\{S_n\}$ 非常返.

注1.5.13. 上面的例题表明一维和二维的简单随机游动常返, 而三维以上的简单随机游动非常返. 日本学者Kakutani 对此给出了一个有趣的注解: “A drunk man can find way home, a drunk bird can not”.

除了判别马氏链的常返性以外, 我们还可以用格林函数证明另外的两个结论. 一个是, 常返性是互通类的性质, 我们把它写成个推论的形式, 因为它是命题 1.5.6 的部分结论. 另一个是, 不变分布(若存在, 则)在非常返的状态上没有权重.

推论1.5.14. 假设 i 常返, $i \rightarrow j$. 那么, j 常返.

证: 若 $j = i$ 则命题成立. 下面假设 $j \neq i$.

首先, 我们用反证法证明 $j \rightarrow i$. 假设 $j \rightarrow i$ 不成立, 即 $P_j(\tau_i = \infty) = 1$. 由命题 1.3.2, 存在 $n \geq 1$ 与 $i_1, \dots, i_{n-1} \in S \setminus \{i\}$ 使得 $p_{i_0 i_1} > 0, \dots, p_{i_{n-1} i_n} > 0$, 其中, $i_0 = i, i_n = j$. 于是,

$$\begin{aligned} P_i(X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = j, X_{n+m} \neq i, \forall m \geq 1) \\ = p_{i_0 i_1} \cdots p_{i_{n-1} i_n} P_j(\tau_i = \infty) > 0. \end{aligned}$$

然而, 根据 i_1, \dots, i_{n-1} 都不是 i 并且 $j \neq i$, 上式等号左边中的事件可推出 $X_n \neq i, \forall n \geq 1$, 即 $\sigma_i = \infty$. 这表明回访概率 $P_i(\sigma_i < \infty) < 1$. 由 (1.5.3), i 非常返, 矛盾! 因此假设不成立, 我们有 $j \rightarrow i$.

然后, 我们得到 $i \leftrightarrow j$. 根据命题 1.3.2, 存在 $n, m \geq 1$ 使得 $p_{ij}^{(n)}, p_{ji}^{(m)} > 0$. 根据 Chapman-Kolmogorov 等式, 即命题 1.1.12, 我们有

$$p_{jj}^{(m+r+n)} = \sum_{k,l \in S} p_{jk}^{(m)} p_{kl}^{(r)} p_{lj}^{(n)} \geq p_{ji}^{(m)} p_{ii}^{(r)} p_{ij}^{(n)}, \quad \forall r \geq 0.$$

两边同时对 r 求和, 得 $G_{jj} \geq p_{ji}^{(m)} G_{ii} p_{ij}^{(n)} = \infty$, 其中, 最后一个等号是由于 i 常返, 即 $G_{ii} = \infty$. 因此, j 常返. \square

命题1.5.15. $G_{ii} = P_i(\tau_j < \infty)G_{jj}$.

证: 假设 $i \neq j$, 我们研究从 i 出发的马氏链 $\{X_n\}$ 访问 j 的总次数 $V_j = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=j\}}$ 的期望. 在 $\tau_j = \infty$ 上, $V_j = 0$; 在 $\tau_j < \infty$ 上, 令 $Y_n = X_{\tau_j+n}$, $\forall n \geq 0$, 根据首达时的定义,

$$V_j = V_j^{(Y)} = \sum_{n=0}^{\infty} \mathbf{1}_{\{Y_n=j\}}.$$

于是, 当 $m \geq 1$ 时, $V_j = m$ 蕴含着 $\tau_j < \infty$, 这结合在事件 $\tau_j < \infty$ 上, $V_j = V_j^{(Y)}$ 表明

$$\begin{aligned} P_i(V_j = m) &= P_i(V_j^{(Y)} = m, \tau_j < \infty) \\ &= P_i(\tau_j < \infty)P_i(V_j^{(Y)} = m | \tau_j < \infty) = P_i(\tau_j < \infty)P_j(V_j = m), \end{aligned}$$

其中, 在最后一个等号中, 我们用到了强马氏性. 两边乘以 m , 再对 $m \geq 1$ 求和, 我们得到

$$E_i V_j = P_i(\tau_j < \infty) \cdot E_j V_j.$$

即 $G_{ij} = P_i(\tau_j < \infty)G_{jj}$. □

推论1.5.16. 若 j 非常返, 则 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$, $\forall i$.

证: 一方面, 由命题 1.5.15, $G_{ij} \leq G_{jj}$. 另一方面, j 非常返表明 $G_{jj} < \infty$. 因此,

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} < \infty.$$

这蕴含着结论成立. □

推论1.5.17. 若 π 是一个不变分布, j 非常返, 则 $\pi_j = 0$.

证: 由命题 1.5.15, $G_{ij} \leq G_{jj}$. 两边同时乘以 π_i , 并对 i 求和, 我们有: 若 π 是不变分布, 而 j 非常返, 那么

$$\sum_{i \in S} \pi_i \sum_{n=0}^{\infty} p_{ij}^{(n)} \leq G_{jj} < \infty.$$

然而, 交换求和次序, 我们有

$$\sum_{i \in S} \pi_i \sum_{n=0}^{\infty} p_{ij}^{(n)} = \sum_{n=0}^{\infty} \sum_{i \in S} \pi_i p_{ij}^{(n)} = \sum_{n=0}^{\infty} \pi_j,$$

其中, 最后一个等号是因为 π 是不变分布. 综上, 我们有 $\infty \cdot \pi_j \leq G_{jj} < \infty$. 因此, $\pi_j = 0$. □

在下一节中, 我们将证明, 对于一个不可约马氏链, 就像我们在例 1.2.16 与例 1.2.17 中看到的那样, 如果存在不变分布 π , 那么该马氏链访问 j 的频率就会几乎必然收敛到 π_j , 这一

结论请见下一节的公式 (1.7.5). 从这个角度看, 非常返的状态被访问的总次数有限, 从而访问频率几乎必然趋于0, 于是 $\pi_j = 0$ 就是个很自然的结论.

补充知识.

关于随机运动的一些历史注记如下. 在1905年, 英国学者Karl Pearson在Nature杂志上发表文章“*The problem of random walker*”, 首次用了随机游动这一名称. 他的问题如下: A man starts from a point 0 and walks l yards in a straight line: he then turns through any angle whatever and walks another l yards in a second straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and $r + \delta r$ from his starting-point 0. The problem is one of considerable interest, but I have only succeeded in obtaining an integrated solution for two stretches. I think, however, that a solution ought to be found, if only in the form of a series in power of $1/n$, where n is large. 其实, 随机游动的某些特性早已被人们研究过. Lord Rayleigh 在1880年和1889年发表的文章就解答了Pearson的问题. Pearson所考虑的是二维随机游动.

历史上, 随机游动也曾被译为随机徘徊或随机游走. 随机游动不仅是研究其它随机过程的基础工具, 还是当今十分活跃的研究课题. 它大致分为两个方面, 一是对简单随机游动的十分精准的刻画, 二是研究群上的随机游动, 例如考察某类矩阵组成群上的随机游动.

据说苏黎世的高等理工学校周围街区相当规则, 像是二维格点 \mathbb{Z}^2 , 上个世纪二十年代Polya在那里随意漫步时一而再、再而三地遇到一对情侣. 为了说明自己并不在盯梢人家, Polya把随意漫步看成是二维格点上的简单随机游动, 而两个独立的简单随机游动 $\{S_n\}$ 与 $\{T_n\}$ 相遇等价于一个随机游动 $\{S_n - T_n\}$ 回到原点, 这引发了人们对常返性的研究.

习题

1. 证明公式 (1.5.4).
2. 假设 S 不可约, 常返, $A, B \subseteq S$ 且 $A \cap B = \emptyset$. 记 $x_i = P_i(\tau_A < \tau_B)$, 写出 $x_i, i \in S$ 满足的方程组.
3. 制造某种产品需要经过前后两道工序. 在完成第一道工序之后, 10% 的加工件成了废品, 20% 的加工件需要返工, 剩余的70% 则进入第二道工序. 在完成第二道工序之后, 5% 的加工件成了废品, 5% 的加工件需要返回到第一道工序, 10% 的加工件需要返回到第二道工序; 剩余的80% 可以出厂. (1) 试建立一个马氏链模拟此系统, (2) 利用吸收概率求整个生产过程的废品率.
4. 一只青蛙在正立方体的顶点上做随机游动, 每次以 $1/4$ 的概率停留不动, 以 $1/4$ 的概率选

- 取一条边并跳至相邻的顶点. 试求: (1) 从正方体的一个顶点 v 出发首次回到 v 的平均时间; (2) 从 v 出发首次到达对径点 w 的平均时间.
5. 某赌徒参加公平博弈, 每次输、赢的概率均为 $\frac{1}{2}$. 他的押注方式如下: 如果口袋里钱不超过 5 元, 他将全部押为赌注; 如果口袋里钱数 i 超过 5 元但不足 10 元, 则他押注 $10 - i$ 元; 如果口袋里没有钱或有 10 元钱, 他就结束赌博回家. 假设他最初有 2 元钱, 求: 他回家时口袋里有 10 元钱的概率.
 6. 假设 $\{S_n\}$ 是一维随机游动, 步长分布为 $P(\xi = k) = \frac{1}{6}$, $k = 1, \dots, 6$. 令 $T = \min\{n \geq 1 : S_n - 1 \text{ 可以被 } 8 \text{ 整除}\}$. 试求: $E_0 T$.
 7. 设 $\{X_n\}$ 是一维随机游动, $X_0 = 0$, 步长分布为 $P(\xi = 1) = 1 - P(\xi = -1) = p \in (\frac{1}{2}, 1)$.
 - (1) 对于正整数 k , 试求 $E_0 \tau_k$.
 - (2) 求 $Y = \min\{X_0, X_1, X_2, \dots\}$ 的概率分布.
 8. 某商家为促销, 在产品包装里面夹塞进一张小画片. 假设全套共有 N 种小画片, 每种小画片的出现概率都相同. 假设某人每天购买一包该产品, 第 n 天见过 X_n 种不同的小画片. (1) 写出 $\{X_n\}$ 的转移概率, (2) 假设此人总共花了 τ 天收集齐整套小画片, 利用吸收概率求 $E\tau$.
 9. 设 $\{X_n\}$ 为随机游动, 步长分布为 $P(\xi = 2) = P(\xi = -1) = 1/2$. 令 $\phi(s) = E_1 s^{\tau_0}$, $\forall s \in (0, 1)$. 证明: $s\phi(s)^3 - 2\phi(s) + s = 0$.
 10. 证明 $y_i = E_i \sigma_D$, $i \in S$ 是 (1.5.7) 最小的非负解.
 11. 研究更新过程(例 1.1.9)的常返性.
 12. 证明规则树 \mathbb{T}^d 上的简单随机游动非常返, 其中, $d \geq 2$.
 13. 考虑 $\{0, 1, 2, \dots\}$ 上的马氏链 $\{X_n : n \geq 0\}$, 其转移概率为 $p_{01} = 1$;

$$p_{i,i+1} = \frac{i^2 + 2i + 1}{2i^2 + 2i + 1}, \quad p_{i,i-1} = \frac{i^2}{2i^2 + 2i + 1} \quad i \geq 1;$$
 若 $|i - j| \geq 2$, 则 $p_{ij} = 0$. 证明该马氏链是非常返的, 并计算 $\rho_i = P_i(\sigma_0 < \infty)$. (提示: $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$.)
 14. 设 ξ_1, ξ_2, \dots 是独立同分布的离散型随机变量序列, $E\xi_1 \neq 0$. 令 $S_0 = 0$, $S_n = \xi_1 + \xi_2 + \dots + \xi_n$. 证明: 随机游动 $\{S_n\}$ 是非常返的.
 15. 证明: (1) 若 A 是有限闭集, 则存在常返类 C 使得 $C \subseteq A$; (2) 有限状态空间上的马氏链有常返态.

16. 设马氏链的状态空间为 $\{0, 1, \dots, 6\}$, 转移矩阵如下. 试确定哪些状态是常返的, 哪些是非常返的, 并求在每个常返类上的不变分布.

$$\begin{pmatrix} 1/2 & 0 & 1/8 & 1/4 & 1/8 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}.$$

17. 假设 ξ_1, \dots, ξ_n 独立同分布, $P(\xi_1 = 1) = 1 - P(\xi_1 = 0) = p \in (0, 1)$. 记 $K = \xi_1 + \dots + \xi_n$.
证明: $E(K - EK)^4 = \sum_{r=1}^n E(\xi_1 - p)^4 + C_n^2 C_4^2 (\text{Var}(\xi_1))^2$.
18. 记号同上题. 假设 $0 < q < p$, $a > 0$. 证明: (1) $P(K < qn) \leq e^{-aqn}(pe^a + 1 - p)^n$. (2)
求 $\varphi(a) = e^{-aq}(pe^a + 1 - p)$ 的最小值点 a^* , 并证明 $\varphi(a^*) < 1$.
19. 假设 $d \geq 3$. 证明: 存在常数 C_d 使得 $P_0(S_{2n} = 0) \leq C_d \cdot n^{-d/2}$. (提示: 仿照 (1.5.18)
与(1.5.19), 利用上题结论.)
20. 证明: 对任意 $i \neq j$, 下面三条等价: $\rho_{ij} > 0$; $i \rightarrow j$; $G_{ij} > 0$.
21. 证明: $\rho_{ii} = 1 - 1/G_{ii}$.
22. 假设马氏链从 i 出发, 记 F_{ij} 为 τ_j 的母函数, 即 $F_{ij}(s) = \sum_{n=0}^{\infty} P_i(\tau_j = n)s^n$. 证明: $G_{ij}(s) = F_{ij}(s)G_{jj}(s)$.
- 23* 对任意 $(i, j) \in \mathbb{Z}^2$, 独立抛一枚公平硬币, 若抛到正面, 则在 (i, j) 与 $(i+1, j)$ 之间连一条边, 否则, 在 (i, j) 与 $(i, j+1)$ 之间连一条边. 于是, 我们得到二维格点的一个随机子图 G , 以 \mathbb{Z}^2 为顶点. 证明: $P(G \text{ 连通}) = 1$.

§1.6 *可配称马氏链的常返性

本节的主要目的是从物理中的电网的视角给出可配称马氏链常返的一个判别法. 假设 \mathbf{P} 不可约、可配称, π 为一个配称测度, 满足: $\pi_i > 0, \forall i \in S$, 且如下细致平衡条件成立.

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad \forall i, j \in S.$$

一、可配称马氏链与电网.

将 S 中的任意状态视为电网中的一个结点. 对任意 $i \neq j$, 若 $p_{ij} > 0$, 则我们在 i, j 之间连一条导线, 其电阻为

$$r_{ij} = r_{i,j} := \frac{1}{\pi_i p_{ij}}. \quad (1.6.1)$$

根据 \mathbf{P} 可配称并且 π 是 \mathbf{P} 的配称测度知, 此时, $p_{ji} > 0$, 并且 $r_{ji} = r_{ij}$. 因此, 以上操作是合理的. 当 $p_{ij} = 0$ 时, 补充定义 $r_{ij} = \infty$. 由 \mathbf{P} 不可约, 以上操作完毕后, 我们便得到一个连通的电网. 如果无视电阻的大小, 只看是否有电线连接, 那么电网就退化为一个图, 记为 $G = (V, E)$.

例1.6.1 (例 1.1.10 续). 如果 \mathbf{P} 就是某个连通图 $G = (V, E)$ 上的随机游动, 那么, 可以取 $\pi_i = d_i$, 其中 d_i 为结点 i 的度. 此时, $r_{ij} = r_{ji} = 1$. 也就是说, \mathbf{P} 对应的电网图就是 $G = (V, E)$, 其中 E 中的每条边的电阻都是 1. 若图 G 上的随机游动常返, 则我们称 G 常返, 否则称 G 非常返.

现在, 取定某顶点, 记为 o . 在物理上, 如果在 o 点接电势为 1 的电源, 其它地方没有电源(无穷远接地), 用 V_i 表示电网中结点 i 的电势, 则 $0 \leq V_i \leq 1, \forall i$, 并且 $V_i, i \in S$ 满足的方程为:

$$\sum_{j \in S} \frac{1}{r_{ij}} (V_j - V_i) = 0, \quad \forall i \neq o; \quad V_o = 1.$$

将(1.6.1) 代入并整理, 该方程便可改写为

$$V_i = \sum_{j \in S} p_{ij} V_j, \quad \forall i \neq o; \quad V_o = 1.$$

这就是吸收概率 $P_i(\tau_o < \infty)$ 满足的方程组(1.5.5). 因此, 猜测

$$V_i = P_i(\tau_j < \infty), \quad \forall i \in S$$

是具有合理性. 事实上, 数学家们证明了下面的命题 1.6.2, 它表明上式是正确的. 根据第 §1.5 中常返性的吸收概率判别法, 该马氏链是否常返等价于 V_i 是否恒为 1. 用电网的语言来描述就是: 该电网中结点 o 与无穷远点之间的“总有效电阻 R_o ”是否为无穷大. 根据功率等于 $I^2 R$, 其中 I 是电流, R 是电阻, 并且电流的功率是最小的, 我们可以通过功率来定义电阻 R . 为此,

我们先引入“单位流”的概念. 一个从 o 出发的单位流指的是 S^2 上满足如下要求的函数 f , 其中 $f_{ij} = f(i, j)$:

$$f_{ij} = 0, \text{ 若 } i \not\sim j; \quad f_{ij} = -f_{ji}, \forall i \neq j; \quad \sum_{j \sim i} f_{ij} = \mathbf{1}_{\{i=o\}}.$$

假设 $i \sim j$. 若 $f_{ij} = 0$, 则在连接 i, j 的边上没有流通过; 若 $f_{ij} > 0$, (此时, $f_{ji} < 0$), 则有流量为 f_{ij} 的流从 i 流向 j . 关于最后一条, 当 $i = o$ 时, $\sum_{j \sim o} f_{oj} = 1$, 这表明 o 是源, 流量为1; 当 $i \neq o$ 时, $\sum_{j \sim i} f_{ij} = 0$, 等价地,

$$\sum_{j \sim i: f_{ij} > 0} f_{ij} = \sum_{j \sim i: f_{ij} < 0} f_{ji}.$$

这表明结点 i 即不是源也不是汇. 在 i 处的流入总量(等号右边)等于流出总量(等号左边). 将 S 上所有从 o 出发的单位流组成的集合记为 \mathbb{F}_o . 对任意 $f \in \mathbb{F}_o$, 将 f 带来的功率定义为

$$I(f) := \frac{1}{2} \sum_{i,j} f_{ij}^2 r_{ij},$$

其中, 系数 $\frac{1}{2}$ 是因为每条边因着端点 i, j 的顺序不同而被计算了两次. 当 f 为从 o 出发的单位电流时, 它就是 R_o . 根据电流做功最小原理, 结点 o 与无穷远点之间的有效电阻被定义为

$$R_o := \inf \{I(f) : f \in \mathbb{F}_o\}. \quad (1.6.2)$$

命题1.6.2. 假设 \mathbf{P} 不可约、可配称. 那么, 下面三条等价: (1) \mathbf{P} 常返, (2) 存在 $o \in S$ 使得 $R_o = \infty$, (3) $R_o = \infty, \forall o \in S$.

例1.6.3. 对于一个有限的电网络, $\mathcal{F}_o = \emptyset$, 于是 $R_o = \inf \emptyset$, 根据约定, $\inf \emptyset = \infty$. 因此, 有限马氏链总是常返的.

推论1.6.4. 假设 G_1 是 G_2 的子图. 若 G_2 常返, 则 G_1 也常返.

证: 取定 G_1 中的一个点 o . 对 $i = 1, 2$, 将 G_i 中 o 与无穷原点之间的有效总电阻记为 $R_o(G_i)$, 将在 G_i 中从 o 出发的流组成的集合记为 $\mathbb{F}_o(G_i)$. 因为 G_1 是 G_2 的子图, 所以对任意 $f \in \mathbb{F}_o(G_1)$, 我们都可以将 f 放在 G_2 中看, 即 $\mathbb{F}_o(G_1) \subseteq \mathbb{F}_o(G_2)$. 这结合例 1.6.1 与 (1.6.2) 表明 $R_o(G_1) \geq R_o(G_2)$. 因此, 结论成立. \square

例1.6.5 (生灭链的常返性. 例 1.2.12, 1.3.7 和 1.5.9 续). 假设 π 满足细致平衡条件, 那么, 根据 (1.6.1), 我们有

$$\frac{r_{i,i+1}}{r_{i-1,i}} = \frac{r_{i,i+1}}{r_{i,i-1}} = \frac{\pi_i p_{i,i-1}}{\pi_i p_{i,i+1}} = \frac{p_{i,i-1}}{p_{i,i+1}}. \quad (1.6.3)$$

因此,

$$r_{i,i+1} = \frac{d_i}{b_i} r_{i-1,i} = \cdots = \frac{d_1 \cdots d_i}{b_1 \cdots b_i} r_{01}.$$

取 $r_{0,1} = 1$. 此电网为所有导线串联, 因此总电阻为

$$R_0 = \sum_{i=0}^{\infty} r_{i,i+1} = 1 + \frac{d_1}{b_1} + \frac{d_1 d_2}{b_1 b_2} + \dots$$

由命题 1.6.2 知生灭链常返当且仅当 $\sum_{i=1}^{\infty} \frac{d_1 \cdots d_i}{b_1 \cdots b_i}$ 发散.

例1.6.6. $S = \{0, 1, 2, \dots\}$, $p_{0,1} = 1$;

$$p_{i,i+1} = \frac{1}{2}, \quad p_{i,i-1} = \frac{\exp\{-ci^{-\alpha}\}}{2}, \quad p_{ii} = \frac{1 - \exp\{-ci^{-\alpha}\}}{2}, \quad \forall i \geq 1.$$

与例 1.6.5 类似, (1.6.3) 仍然成立. 此时,

$$r_{i,i+1} = \exp\{-ci^{-\alpha}\} r_{i-1,i} = \dots = \exp\left\{-c \sum_{j=1}^i j^{-\alpha}\right\} r_{0,1}.$$

取 $r_{0,1} = 1$, 我们有

$$R_0 = 1 + \sum_{i=1}^{\infty} \exp\left\{-c \sum_{j=1}^i j^{-\alpha}\right\}.$$

从而, 该马氏链常返当且仅当上面的级数发散. 若 $c = 0$, 则 $R_0 = \infty$, 此马氏链常返, 它就是带反射壁的简单随机游动. 下设 $c > 0$.

当 $\alpha > 1$ 时, $\sum_{j=1}^i j^{-\alpha} < \sum_{j=1}^{\infty} j^{-\alpha} < \infty$, 从而 $\exp\left\{-c \sum_{j=1}^i j^{-\alpha}\right\} \geq \varepsilon > 0$, $\forall i \geq 1$. 于是, $R_0 = \infty$, 即, 该马氏链常返.

当 $0 \leq \alpha < 1$ 时, $\sum_{j=1}^i j^{-\alpha} \geq \int_1^{i+1} x^{-\alpha} dx = \frac{1}{1-\alpha}((i+1)^{1-\alpha} - 1)$. 于是存在 i_0 使得 $\exp\left\{c \sum_{j=1}^i j^{-\alpha}\right\} \geq i^2$, $\forall i \geq i_0$. 这表明 $R_0 < \infty$, 即, 该马氏链非常返.

当 $\alpha = 1$ 时, 一方面, $\sum_{j=1}^i j^{-1} \leq 1 + \int_2^i x^{-1} dx = 1 + \ln i$, $\forall i \geq 2$. 另一方面, $\sum_{j=1}^i j^{-1} \geq \int_1^{i+1} x^{-1} dx = \ln(i+1) \geq \ln i$, $\forall i \geq 1$. 因此, $\exp\left\{-c \sum_{j=1}^i j^{-1}\right\} \in [(ek)^{-c}, k^{-c}]$. 因此, 当 $c \leq 1$ 时, $R_0 = \infty$, 即, 该马氏链常返; 当 $c > 1$ 时, $R_0 < \infty$, 即, 该马氏链非常返.

总结起来, 该马氏链常返的充要条件是: $\alpha > 1$; 或者 $c = 0$; 或者 $\alpha = 1$ 且 $c \leq 1$. 常返性对参数 α, c 的依赖关系如图 1.15 所示, 其中, 实线表示常返, 虚线表示非常返.

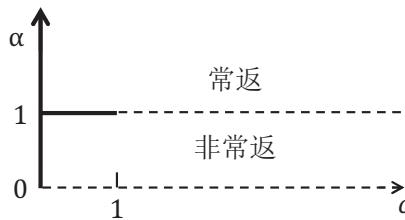


图 1.15 常返: $\alpha > 1$; $\alpha = 1$ 且 $c \leq 1$; $\alpha < 1$ 且 $c > 1$.

在上面的两个例子中, 马氏链对应的电网结构非常简单, 就是一些导线串联起来. 当马氏链对应的电网结构很复杂时, 一般而言我们是无法直接计算其总电阻的. 这使得我们需要一些其它办法来判断图的常返性. 下面, 我们将介绍两种方法: Thomson 原理和Dirichlet 原理. 其中, Thomsom 原理(命题 1.6.7)用于证明某马氏链非常返, 而Dirichlet 原理(命题 1.6.10)用于证明某马氏链常返. 值得一提的是, 虽然命题 1.6.2 和命题 1.6.7 是用物理的语言描述的, 但是它们都有严格的数学证明. 有兴趣的读者可以参阅 [11].

二、Thomson 原理.

根据命题 1.6.2, 我们得到如下结论, 它通常用于证明某马氏链非常返.

命题1.6.7 (Thomsom 原理). 若存在从 o 出发的单位流 f 使得 $I(f) < \infty$, 则该马氏链非常返.

例1.6.8. 在例 1.5.12 中, 我们看到, 当 $d \geq 3$ 时, \mathbb{Z}^d 非常返. 现在, 我们用Thomson 原理来证明这个结论.

当 $d \geq 3$ 时, \mathbb{Z}^3 是 \mathbb{Z}^d 的子图. 由推论 1.6.4, 我们只需证明 \mathbb{Z}^3 非常返即可.

考虑3 维格点图, 其中, 任意边的电阻都为1. 即, 取

$$r_{(i,j,k)(i+1,j,k)} = r_{(i,j,k)(i,j+1,k)} = r_{(i,j,k)(i,j,k+1)} = 1, \quad \forall (i,j,k) \in \mathbb{Z}^3.$$

下面, 我们考虑一个单位流 f , 它从 $o = (1, 1, 1)$ 出发, 在第一象限流向无穷远. 具体地, 对任意 $n \geq 0$, 记

$$S_n := \{(i, j, k) \in \mathbb{Z}^3 : i, j, k \geq 1, i + j + k = 3 + n\}, \quad S = \bigcup_{n=0}^{\infty} S_n. \quad (1.6.4)$$

不难看出, $S_0 = \{(1, 1, 1)\}$, 且

$$|S_n| = \frac{1}{2}(n+1)(n+2).$$

下面, 我们让 f 从 S_n 流向 S_{n+1} . 具体地, 对任意 $n \geq 0$, 任意 $(i, j, k) \in S_n$, 流可以从 (i, j, k) 流向 $(i+1, j, k)$, $(i, j+1, k)$, $(i, j, k+1)$, 流量分别为

$$\begin{aligned} f_{(i,j,k),x} &:= \frac{1}{|S_n|} \cdot \frac{i}{i+j+k}, \\ f_{(i,j,k),y} &:= \frac{1}{|S_n|} \cdot \frac{j}{i+j+k}, \\ f_{(i,j,k),z} &:= \frac{1}{|S_n|} \cdot \frac{k}{i+j+k}, \end{aligned}$$

其中, 角标的 x 、 y 、 z 分别表示流的方向是沿着 x 轴、 y 轴、 z 轴. 不难验证, f 确实是一个从 $(1, 1, 1)$ 出发的单位流. 进一步,

$$f_{(i,j,k),x}^2 + f_{(i,j,k),y}^2 + f_{(i,j,k),z}^2 = \frac{1}{|S_n|^2} \cdot \frac{i^2 + j^2 + k^2}{(i+j+k)^2} \leq \frac{1}{|S_n|^2},$$

其中, 不等号是因为 $i^2 + j^2 + k^2 \leq (i + j + k)^2$. 因此,

$$I_n(f) = \sum_{(i,j,k) \in S_n} \left(f_{(i,j,k),x}^2 + f_{(i,j,k),y}^2 + f_{(i,j,k),z}^2 \right) \leq |S_n| \cdot \frac{1}{|S_n|^2} = \frac{2}{(n+1)(n+2)}.$$

于是, $I(f) = \sum_{n=0}^{\infty} I_n(f) < \infty$. 这表明 \mathbb{Z}^3 非常返.

上述例题中的流 f 有其具体的概率含义, 它其实就是下面介绍的 Polya 坛子模型中的概率流.

例1.6.9 (Polya 坛子). 假设坛中最初有一个红球、一个黑球和一个白球. 每一步, 从坛中随机拿出一个球, 再将此球连同一个与之同色的球一起放回坛中. 假设 n 步后坛中有 R_n 个红球, B_n 个黑球, W_n 个白球, 那么 $R_n + B_n + W_n = 3 + n, \forall n \geq 0$. 令 $X_n = (R_n, B_n, W_n)$, 则 $\{X_n\}$ 是一个马氏链, 状态空间 S 由 (1.6.4) 给出. 转移概率如下:

$$\begin{aligned} p_{(i,j,k),(i+1,j,k)} &:= \frac{i}{i+j+k}, \\ p_{(i,j,k),(i,j+1,k)} &:= \frac{j}{i+j+k}, \\ p_{(i,j,k),(i,j,k+1)} &:= \frac{k}{i+j+k}. \end{aligned}$$

由初始状态为 $X_0 = (1, 1, 1)$ 不难验证 X_n 服从 S_n 上的均匀分布, 其中 S_n 也由 (1.6.4) 给出. 于是, 当 $(i, j, k) \in S_n$ 时, 从 (i, j, k) 到 $(i+1, j, k)$ 的概率流就是

$$P(X_n = (i, j, k)) p_{(i,j,k),(i+1,j,k)} = \frac{1}{|S_n|} \cdot \frac{i}{i+j+k} = f_{(i,j,k),x}.$$

同理, 从 (i, j, k) 到 $(i, j+1, k)$ 的概率流为 $f_{(i,j,k),y}$, 到 $(i, j, k+1)$ 的概率流为 $f_{(i,j,k),z}$.

三、Dirichlet 原理.

现在我们介绍用于证明马氏链常返的一个工具——Dirichlet 原理. 如果 S 有限, 那么它常返. 下面假设 S 可列. 取定 $o \in S$. 任取包含 o 的有限连通子集 A . “连通”指的是: 对任意 $i, j \in A$, 存在 $n \geq 1$ 及其 A 中的状态 i_1, \dots, i_{n-1} 使得 $p_{i_{r-1}i_r} > 0, r = 1, \dots, n$, 其中 $i_0 = i, i_n = j$. 那么, 我们可以考虑一个新的马氏链 $\{\tilde{X}_n\}$, 它的状态空间为 $\tilde{S} = A \cup \{\partial\}$, 转移概率为:

$$\tilde{p}_{ij} = p_{ij}, \quad \forall i, j \in A; \quad \tilde{p}_{i\partial} = \sum_{j \notin A} p_{ij}; \quad \tilde{p}_{\partial\partial} = 1.$$

它其实就是在 S 中把 A^c 中的所有状态捏为一个点 ∂ , 并让 ∂ 为一个吸收态. 令

$$\varepsilon_A := \inf\{n \geq 0 : X_n \notin A\} = \inf\{n \geq 0 : \tilde{X}_n = \partial\} = \tilde{\tau}_{\partial}.$$

它被称为马氏链的逃逸时间, 即, 逃离区域 A 的时间. 于是, $P_i(\varepsilon_A < \infty) = \tilde{P}_i(\tilde{\tau}_\partial < \infty) = 1$, $\forall i \in A$. 从而,

$$P_o(\sigma_o = \infty) \leq P_o(\varepsilon_A < \sigma_o). \quad (1.6.5)$$

下面的命题可以认为是给出了 $P_o(\varepsilon_A \leq \sigma_o)$ 的上界.

命题1.6.10. (Dirichlet原理) 假设 π 是 \mathbf{P} 的配称测度, $\pi_o = 1$. 假设 A 有限, $o \in A \subseteq S$, $P_o(\varepsilon_A < \infty) = 1$, 那么

$$2\pi_o P_o(\varepsilon_A < \sigma_o) = \inf_{f \in \mathbb{F}} \Phi(f) = \Phi(g),$$

其中, \mathbb{F} 是 S 上满足如下性质的函数 $f : S \rightarrow \mathbb{R}$, $i \mapsto f_i$.

$$0 \leq f_i \leq 1, \forall i \in S; \quad f_o = 0; \quad f_i = 1, \forall i \notin A, \quad (1.6.6)$$

$$\Phi(f) = \sum_{i,j \in S} \pi_i p_{ij} (f_j - f_i)^2; \quad g_i = P_i(\varepsilon_A < \sigma_o), \quad i \in S.$$

证: 我们分三步进行证明.

第一步, 对任意 $f \in \mathbb{F}$, 由于 $f|_{A^c} \equiv 1$, 即, $1 - f$ 在 A^c 上恒为 0, 因此我们可以将 $1 - f$ 视为 A 上的函数. 往证 $\Phi(f) < \infty$. 事实上,

$$\Phi(f) = 2 \sum_{i \in A} \pi_i (1 - f_i)^2 - 2 \sum_{i,j \in A} \pi_i p_{ij} (1 - f_i)(1 - f_j). \quad (1.6.7)$$

将 $f_j - f_i$ 改写为 $(1 - f_i) - (1 - f_j)$, 便有

$$\begin{aligned} \Phi(f) &= \sum_{i,j \in S} \pi_i p_{ij} ((1 - f_i) - (1 - f_j))^2 \\ &= \sum_{i,j \in S} \pi_i p_{ij} (1 - f_i)^2 + \sum_{i,j \in S} \pi_i p_{ij} (1 - f_j)^2 - 2 \sum_{i,j \in S} \pi_i p_{ij} (1 - f_i)(1 - f_j). \end{aligned}$$

一方面, 因为 $f|_{A^c} \equiv 1$, 所以上式右边第三个级数实际上是有限求和. 具体地,

$$\sum_{i,j \in S} \pi_i p_{ij} (1 - f_i)(1 - f_j) = \sum_{i,j \in A} \pi_i p_{ij} (1 - f_i)(1 - f_j).$$

另一方面, 根据细致平衡条件, 第一个级数和第二个是相等的. 具体地,

$$\sum_{i,j \in S} \pi_i p_{ij} (1 - f_j)^2 = \sum_{i,j \in S} \pi_j p_{ji} (1 - f_j)^2 = \sum_{i,j \in S} \pi_i p_{ij} (1 - f_i)^2, \quad (1.6.8)$$

其中, 第一个等号用到了细致平衡条件, 而第二个等号则是交换了求和哑元 i, j 的符号. 进一步,

$$\sum_{i,j \in S} \pi_i p_{ij} (1 - f_i)^2 = \sum_{i \in S} \pi_i (1 - f_i)^2 = \sum_{i \in A} \pi_i (1 - f_i)^2.$$

总结起来, 我们有 (1.6.7) 成立. 进一步, 因为 A 有限, 所以 $\Phi(f) < \infty$.

第二步, 往证 $\Phi(g) \leq \Phi(f)$, $\forall f \in \mathbb{F}$.

$$g_i = \sum_{j \in S} p_{ij} g_j, \quad \forall i \in A \setminus \{o\}. \quad (1.6.9)$$

记 $h = g - f$. 因为 $f, g \in \mathbb{F}$, 所以

$$h_o = 0; \quad h_i = 0, \quad \forall i \notin A.$$

仿照第一步的推导, 可以证明 $\Phi(h) < \infty$. 从而,

$$\Phi(f) = \Phi(g) + \Psi(h) + 2a,$$

其中,

$$\Psi(h) = 2 \sum_{i \in A} \pi_i h_i^2 - 2 \sum_{i,j \in A} \pi_i p_{ij} h_i h_j = \Phi(h) \geq 0,$$

而 a 为交叉项的求和, 即

$$a := 2 \sum_{i \in A} \pi_i (1 - g_i) h_i - \sum_{i,j \in A} \pi_i p_{ij} (1 - g_i) h_j - \sum_{i,j \in A} \pi_i p_{ij} h_i (1 - g_j).$$

仿照 (1.6.8) 可以证明, 上式的最后两个求和相等. 由 $g|_{A^c} \equiv 1$ 知

$$\sum_{j \in A} p_{ij} (1 - g_j) = \sum_{j \in S} p_{ij} (1 - g_j) = 1 - \sum_{j \in S} p_{ij} g_j.$$

进一步, 若还有 $i \in A \setminus \{o\}$, 则根据 (1.6.9), 上式右边还等于 $1 - g_i$. 因此,

$$\begin{aligned} \sum_{i,j \in A} \pi_i p_{ij} h_i (1 - g_j) &= \sum_{i \in A} \pi_i h_i \sum_{j \in A} p_{ij} (1 - g_j) \\ &= \pi_o h_o \left(1 - \sum_{j \in S} p_{oj} g_j \right) + \sum_{i \in A \setminus \{o\}} \pi_i h_i (1 - g_i). \end{aligned} \quad (1.6.10)$$

将上述结论代入 a 的表达式. 由 $h_o = 0$ 便可推出 $a = 0$. 综上, $\Phi(f) = \Phi(g) + \Psi(h) \geq \Phi(g)$. 又因为 $g \in \mathbb{F}$, 所以 $\inf_{f \in \mathbb{F}} \Phi(f) = \Phi(g)$.

第三步, 往证 $\Phi(g) = 2\pi_o P_o(\epsilon_A < \sigma_o)$. 将 (1.6.7) 中的 f 取为 g . 我们先计算 (1.6.7) 右边的第二项. 仿照 (1.6.10), 具体地, 将 (1.6.10) 中的 h_i 改为 $1 - g_i$. 由 $g_o = 0$ 便可推出

$$\sum_{i,j \in A} \pi_i p_{ij} (1 - g_i)(1 - g_j) = \pi_o \left(1 - \sum_{j \in S} p_{oj} g_j \right) + \sum_{i \in A \setminus \{o\}} \pi_i (1 - g_i)^2.$$

将其代入 (1.6.7), 我们有

$$\Phi(g) = 2\pi_o \sum_{j \in S} p_{oj} g_j = 2\pi_o \sum_{j \in S} p_{oj} P_j(\epsilon_A < \sigma_o) = 2\pi_o P_o(\epsilon_A < \sigma_o),$$

其中, 我们在第一个等号中用到了 $g_o = 0$, 在第二个等号中代入了 g 的定义, 在第三个等号中用到了首步分析法. \square

命题 1.6.10 结合 (1.6.5) 可用于证明马氏链常返. 具体的方法如下: (i) 找一列包含 o 有限连通子集 A_1, A_2, \dots ; (ii) 对每个 n , 取 $A = A_n$ 并构造一个满足 (1.6.6) 的函数 f , 记为 $f^{(n)}$, 使得 $\Phi(f^{(n)}) \rightarrow 0$; (iii) 于是, 根据命题 1.6.10 与 (1.6.5), $P_o(\sigma_o = \infty) = 0$, 即 \mathbf{P} 常返.

例1.6.11. 在例 1.5.12 中, 我们看到, \mathbb{Z}^1 与 \mathbb{Z}^2 都常返. 现在, 我们用 *Dirichlet* 原理来证明这个结论. 根据推论 1.6.4, 只需证明 \mathbb{Z}^2 常返即可.

将 \mathbb{Z}^2 中的顶点 i 的横坐标记为 i_1 , 纵坐标记为 i_2 , 即 $i = (i_1, i_2)$. 取配称分布 $\pi_i = 1, \forall i \in \mathbb{Z}^2$. 考虑最大模 $|i| := \max\{|i_1|, |i_2|\}$. 取 o 为原点. 给定 $n \geq 1$, 取 $A = \{j : |j| \leq n\}$, $f_j = a_{|j|}$, 其中 $a_0 = 0, a_n = 1$, 而 a_1, \dots, a_{n-1} 是待定常数, 它们都将在 $[0, 1]$ 中选取. 由 *Dirichlet* 原理,

$$2P_o(\epsilon_A < \sigma_o) \leq \Phi(f) = \sum_{r=1}^n (2r-1)(a_r - a_{r-1})^2.$$

特别地, 取

$$a_r = \frac{\sum_{s=1}^r 1/(2s-1)}{\sum_{t=1}^n 1/(2t-1)}, \quad r = 1, \dots, n-1.$$

那么,

$$2P_o(\epsilon_A < \sigma_o) \leq \sum_{r=1}^n (2r-1) \left(\frac{1/(2r-1)}{\sum_{t=1}^n 1/(2t-1)} \right)^2 = \frac{1}{\sum_{t=1}^n 1/(2t-1)}.$$

上式对任意 $n \geq 1$ 都成立. 令 $n \rightarrow \infty$, 我们便得到 $P_o(\epsilon_A < \sigma_o) = 0$, 即, \mathbb{Z}^2 常返.

§1.7 遍历定理与正常返

在接下来的几节中, 我们讨论第 §1.2 节最后的几个关于不变分布的问题. 在本节, 我们研究不变分布的存在性和表达式. 由上节的推论 1.5.17, 若 π 为不变分布, 则 π 在非常返的状态上全是 0. 所以, 我们只研究常返类. 由命题 1.5.6, 常返类是闭集. 于是, 我们可以将马氏链限制在每一个常返类上. 也就是说, 因为本节在讨论不变分布, 所以我们总假设马氏链不可约, 甚至假设马氏链不可约、常返.

一、频率的极限.

令

$$V_i(n) := \sum_{m=0}^{n-1} \mathbf{1}_{\{X_m=i\}} = |\{0 \leq m \leq n-1 : X_m = i\}|,$$

它表示马氏链在前 n 步(即, 第 0 步, 第 1 步, …, 第 $n-1$ 步)中访问 i 的总次数.

定理 1.7.1. 假设 \mathbf{P} 不可约. 那么, 对任意 $i \in S$, 访问 i 的频率 $\frac{V_i(n)}{n}$ 几乎必然收敛于 $\frac{1}{E_i \sigma_i}$. 具体地, 对任意初分布 μ , 都有

$$P_\mu \left(\lim_{n \rightarrow \infty} \frac{V_i(n)}{n} = \frac{1}{E_i \sigma_i} \right) = 1, \quad \forall i \in S.$$

证: 假设 \mathbf{P} 非常返. 那么,

$$V_i = \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=i\}} < \infty.$$

于是结论成立. 具体地, 一方面, $V_i = \infty$ 表明 $\tau_i < \infty$, 此时, 令 $Y_n = X_{\tau_i+n}$, $\forall n \geq 0$. 于是 $V_i = V_i^{(Y)}$. 根据强马氏性以及非常返性, 对任意状态 $j \in S$,

$$P_j(V_i = \infty | \tau_i < \infty) = P_j(V_i^{(Y)} = \infty | \tau_i < \infty) = P_i(V_i = \infty) = 0.$$

于是

$$P_j(V_i = \infty) = P_j(\tau_i < \infty) P_i(V_i = \infty) = 0.$$

因此, $P_\mu(V_i < \infty) = \sum_j \mu_j P_j(V_i < \infty) = 1$. 由于 $V_i(n) \leq V_i$, $\forall n$, 因此 $\{V_i < \infty\} \subseteq \left\{ \lim_{n \rightarrow \infty} \frac{V_i(n)}{n} = 0 \right\}$. 从而,

$$P_\mu \left(\lim_{n \rightarrow \infty} \frac{V_i(n)}{n} = 0 \right) = 1.$$

另一方面, 因为状态 i 非常返, 所以 $P_i(\sigma_i = \infty) > 0$, 这表明 $E_i \sigma_i = \infty$, 即 $\frac{1}{E_i \sigma_i} = 0$. 从而命题成立.

下面假设 \mathbf{P} 常返. 取定 i . 记

$$A = \left\{ \lim_{n \rightarrow \infty} \frac{V_i(n)}{n} = \frac{1}{E_i \sigma_i} \right\}.$$

往证对任意初分布 μ , 都有 $P_\mu(A) = 1$ 成立.

首先, 假设 $X_0 = i$, 即 $\mu_i = 1$. 根据命题 1.5.5, 粒子回访状态 i 的时间间隔 $\sigma_{i,1}, \sigma_{i,2}, \dots$ 是独立同分布的随机变量序列. 在更新定理(定理 1.2.15) 中, 将 L_i 取为 $\sigma_{i,r}$, 那么 $V_i(n)$ 就是更新定理中的 $1 + R_{n-1}$, 于是

$$\lim_{n \rightarrow \infty} \frac{V_i(n)}{n} = \lim_{n \rightarrow \infty} \frac{1 + R_{n-1}}{n} = \frac{1}{EL_1},$$

其中 $EL_1 = E_i\sigma_i$. 即, $P_i(A) = 1$.

其次, 假设 $X_0 = j$, 即 $\mu_j = 1$, 其中 $j \neq i$. 由于马氏链不可约、常返, 根据命题 1.5.6, 我们有 $P_j(\tau_i < \infty) = 1$. 于是我们可以先让粒子到达状态 i , 再在后续的运动中考虑访问 i 的次数或频率. 具体地, 令 $Y_n = X_{\tau_i+n}$, $\forall n \geq 0$. 因为在 $[0, \tau_i - 1]$ 这个时间段内, 粒子不会访问状态 i , 所以

$$V_i(n) = V_i^{(Y)}(n - \tau_i),$$

其中 $V_i^{(Y)}(m) = \sum_{\ell=0}^{m-1} \mathbf{1}_{\{Y_\ell=i\}}$, $\forall m \geq 0$. 由强马氏性(命题 1.4.2), 对任意给定的 $m \geq 1$, 在 $\tau_i = m$ 的条件下, $\{Y_n\}$ 是从 i 出发的以 \mathbf{P} 为转移矩阵的马氏链. 这表明

$$P_j \left(\lim_{n \rightarrow \infty} \frac{V_i^{(Y)}(n)}{n} = \frac{1}{E_i\sigma_i} \mid \tau_i = m \right) = P_i(A) = 1.$$

于是, 在 $\tau_i = m$ 的条件下,

$$\frac{V_i(n)}{n} = \frac{V_i^{(Y)}(n - \tau_i)}{n} = \frac{V_i^{(Y)}(n - m)}{n - m} \cdot \frac{n - m}{n} \xrightarrow{n \rightarrow \infty} \frac{1}{E_i\sigma_i}.$$

即, $P_j(A | \tau_i = m) = 1$. 根据全概公式,

$$P_j(A) = \sum_{m=1}^{\infty} P_j(\tau_i = m) \times P_j(A | \tau_i = m) = \sum_{m=1}^{\infty} P_j(\tau_i = m) \times 1 = 1.$$

最后, 对于任意初分布 μ , 由全概公式, $P_\mu(A) = \sum_{j \in S} \mu_j P_j(A) = 1$. 从而命题成立. \square

上述命题表明: 在不可约、常返的马氏链中, 对任意 $i \in S$, 粒子访问状态 i 的频率 $\frac{V_i(n)}{n}$ 几乎必然有极限, 其极限就是 $\frac{1}{E_i\sigma_i}$. 进一步, 我们将在下面的命题中证明, 如果不变分布 π 存在, 那么访问 i 的频率 $\frac{V_i(n)}{n}$ 的极限就是 π_i . 因此, 不变分布具有唯一性.

命题1.7.2. 假设 \mathbf{P} 不可约, π 为 \mathbf{P} 的一个不变分布. 则对所有状态 i 均有 $\pi_i > 0$ 且

$$\pi_i = \frac{1}{E_i\sigma_i}. \tag{1.7.1}$$

证: 我们先证明 $\pi_i > 0$, $\forall i \in S$. 首先, 存在 $o \in S$ 使得 $\pi_o > 0$, 否则 $\sum_i \pi_i = 0$, 这与分布列的归一化条件 $\sum_i \pi_i = 1$ 矛盾. 对任意 $i \in S$, 因为 \mathbf{P} 不可约, 所以 $o \rightarrow i$, 于是根据命题 1.3.2, 存在 $n \geq 0$ 使得 $p_{oi}^{(n)} > 0$. 根据 π 是不变分布, 我们有 $\pi_i = \sum_j \pi_j p_{ji}^{(n)} \geq \pi_o p_{oi}^{(n)} > 0$.

下面, 考虑以 π 为初分布的马氏链. 一方面, 因为 π 是不变分布, 所以

$$E_\pi \frac{V_i(n)}{n} = E_\pi \frac{1}{n} \sum_{m=0}^{n-1} \mathbf{1}_{\{X_m=i\}} = \frac{1}{n} \sum_{m=0}^{n-1} P_\pi(X_m=i) = \pi_i.$$

另一方面, 定理 1.7.1 表明 $P_\pi \left(\lim_{n \rightarrow \infty} \frac{V_i(n)}{n} = \frac{1}{E_i \sigma_i} \right) = 1$. 因为 $0 \leq \frac{V_i(n)}{n} \leq 1, \forall n$, 所以根据有界收敛定理(附录中的定理 4.0.16)我们可以推出

$$\lim_{n \rightarrow \infty} E_\pi \frac{V_i(n)}{n} = E_\pi \left(\lim_{n \rightarrow \infty} \frac{V_i(n)}{n} \right) = \frac{1}{E_i \sigma_i}.$$

因此, $\pi_i = \frac{1}{E_i \sigma_i}$. □

上面的命题表明, 对于一个不可约马氏链, 不变分布如果存在则唯一. 将这个唯一的不变分布记为 π , 那么对任意状态*i*, 都有 $E_i \sigma_i = \frac{1}{\pi_i} < \infty$. 我们将在下面的定理 1.7.5 中证明其逆命题, 也就是反过来, 若 $E_i \sigma_i < \infty$, 令 $\pi_i = \frac{1}{E_i \sigma_i}, i \in S$, 则 π 就是不变分布.

二、正常返与不变分布.

定义1.7.3. 若 $E_i \sigma_i < \infty$, 则称状态*i* 是正常返态, 或说*i* 是正常返的(positive recurrent); 若*i* 常返但 $E_i \sigma_i = \infty$, 则称状态*i* 为零常返态, 或说*i* 是零常返的(null recurrent).

取定状态 o , 令

$$\mu_i := E_o \sum_{n=0}^{\sigma_o-1} \mathbf{1}_{\{X_n=i\}}, \quad \forall i \in S. \quad (1.7.2)$$

它表示在一个状态 o 的游弋中, 粒子访问状态*i* 的平均次数.

命题1.7.4. 假设 o 常返. 那么, μ 满足下面的不变方程:

$$\sum_{j \in S} \mu_j p_{ji} = \mu_i, \quad \forall i \in S.$$

证: 根据 (1.7.2), 对任意 $j \in S$,

$$\mu_j = E_o \sum_{n=0}^{\sigma_o-1} \mathbf{1}_{\{X_n=j\}} = E_o \sum_{n=0}^{\infty} \mathbf{1}_{\{X_n=j, \sigma_o > n\}} = \sum_{n=0}^{\infty} P_o(X_n=j, \sigma_o > n). \quad (1.7.3)$$

然后, 将时刻 n 视为现在. 因为 $\{\sigma_o > n\} = \{X_1 \neq o, \dots, X_n \neq o\}$, 这是一个关于过去及现在的状态的事件, 所以根据马氏性可以推出

$$p_{ji} = P(X_{n+1}=i | X_n=j) = P(X_{n+1}=i | X_n=j, \sigma_o > n).$$

于是,

$$\begin{aligned} \sum_{j \in S} \mu_j p_{jk} &= \sum_{j \in S} \sum_{n=0}^{\infty} P_o(X_n=j, \sigma_o > n) P(X_{n+1}=k | X_n=j, \sigma_o > n) \\ &= \sum_{n=0}^{\infty} \sum_{j \in S} P_o(X_n=j, \sigma_o > n, X_{n+1}=k) = \sum_{n=0}^{\infty} P_o(\sigma_o > n, X_{n+1}=k). \end{aligned}$$

现在, 仿照 (1.7.3), 将 $\sum_{n=0}^{\infty} P_o(\sigma_o > n, X_{n+1} = i)$ 写回粒子访问状态 i 的平均次数的形式:

$$\sum_{n=0}^{\infty} P_o(\sigma_o > n, X_{n+1} = i) = E_o \sum_{n=0}^{\infty} \mathbf{1}_{\{\sigma_o > n, X_{n+1} = i\}} = E_o \sum_{n=0}^{\sigma_o-1} \mathbf{1}_{\{X_{n+1} = i\}} = E_o \sum_{n=1}^{\sigma_o} \mathbf{1}_{\{X_n = i\}},$$

其中, 最后一个等号是将其左边的 $n + 1$ 视为新的求和哑元 n . 综上,

$$\sum_{j \in S} \mu_j p_{jk} = E_o \sum_{n=1}^{\sigma_o} \mathbf{1}_{\{X_n = i\}}. \quad (1.7.4)$$

由于 o 常返, 因此 $P_o(\sigma_o < \infty) = 1$. 如果 $i = o$, 那么在 $[0, \sigma_o]$ 这个时间段中, 粒子只有在时刻 0 和时刻 σ_o 的状态为 i , 在其他时刻的状态都不是 i , 于是, $\sum_{n=1}^{\sigma_o} \mathbf{1}_{\{X_n = i\}} = 1 = \sum_{n=0}^{\sigma_o-1} \mathbf{1}_{\{X_n = i\}}$; 如果 $i \neq o$, 那么 X_0 与 X_{σ_o} 都不是 i , 于是 $\sum_{n=1}^{\sigma_o} \mathbf{1}_{\{X_n = i\}} = \sum_{n=1}^{\sigma_o-1} \mathbf{1}_{\{X_n = i\}} = \sum_{n=0}^{\sigma_o-1} \mathbf{1}_{\{X_n = i\}}$. 总结起来, 对任意 i , 我们总有

$$\sum_{n=1}^{\sigma_o} \mathbf{1}_{\{X_n = k\}} = \sum_{n=0}^{\sigma_i - o} \mathbf{1}_{\{X_n = k\}}.$$

这结合 (1.7.2) 表明 (1.7.4) 中等号的右边就是 μ_i . 从而, 不变方程成立. \square

定理1.7.5. 假设 \mathbf{P} 不可约, 则下列几条等价: (1) 所有状态都是正常返的; (2) 存在正常返态; (3) 不变分布存在.

证: 我们将证明 $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$, 于是它们等价. 首先, $(1) \Rightarrow (2)$ 是显然的.

$(2) \Rightarrow (3)$: 假设(2)成立. 任取一个常返态, 记为 o . 根据命题 (1.7.4), (1.7.2) 定义的 μ 是不变测度. 由于 $\sum_{i \in S} \mu_i = E_o \sigma_o < \infty$, 因此 $\pi_i = \frac{\mu_i}{E_o \sigma_o}, i \in S$ 是不变分布. 因此(3)成立.

$(3) \Rightarrow (1)$: 假设(3)成立. 由命题 1.7.2, 不变分布唯一, 其表达式就是 $\pi_i = \frac{1}{E_i \sigma_i}, \forall i$. 并且, 对任意 $i \in S$ 都有 $\pi_i > 0$, 即 $E_i \sigma_i < \infty$. 根据定义, 这即是说 i 是正常返态. 因此, (1) 成立.

\square

注1.7.6. 在定义正常返和零常返时, 并不需要假设马氏链不可约. 由于它们都蕴含着常返, 因此定理 1.7.5 表明对于一般的马氏链, 正常返也是互通类的性质.

推论1.7.7. 假设状态空间 S 有限、不可约. 那么, 不变分布存在.

证: 由定理 1.7.1, $\frac{V_i(n)}{n} \rightarrow 1/E_i \sigma_i, \forall i$. 由 S 有限知 $\sum_{i \in S} \frac{V_i(n)}{n} \rightarrow \sum_{i \in S} \frac{1}{E_i \sigma_i}$, 因为有限的求和号可以和极限符号交换. 由 $\sum_{i \in S} \frac{V_i(n)}{n} = 1$ 知 $\sum_{i \in S} \frac{1}{E_i \sigma_i} = 1$. 从而, 存在 $i \in S$ 使得 $E_i \sigma_i < \infty$, 即 i 正常返. 这表明存在正常返态, 由定理 1.7.5, 不变分布存在. \square

三、遍历定理.

结合定理 1.7.1 1.7.5 和命题 1.7.2, 我们有: 如果马氏链不可约、正常返, 那么

$$P_\mu \left(\lim_{n \rightarrow \infty} \frac{1}{n} V_i(n) = \pi_i \right) = 1, \quad \forall i \in S. \quad (1.7.5)$$

这即是说, 访问状态 i 的频率几乎必然收敛于 i 在(唯一的)不变分布 π 下的概率 π_i . 事实上, 我们还可以得到频率(作为分布列)是 ℓ^1 收敛到不变分布的这个更强结论(命题 1.7.9), 并进一步得到遍历定理(定理 1.7.10). 为证明这些结论, 我们先给出一个引理.

引理1.7.8. 假设 $\pi, \pi^{(1)}, \pi^{(2)}, \dots$ 是分布列, 满足 $\lim_{n \rightarrow \infty} \pi_i^{(n)} = \pi_i, \forall i \in S$. 那么,

$$\lim_{n \rightarrow \infty} \sum_{i \in S} |\pi_i^{(n)} - \pi_i| = 0.$$

证: 若 S 有限,

$$\lim_{n \rightarrow \infty} \sum_{i \in S} |\pi_i^{(n)} - \pi_i| = \sum_{i \in S} \lim_{n \rightarrow \infty} |\pi_i^{(n)} - \pi_i| = 0.$$

下设 S 可列. $\forall \varepsilon > 0$, 存在 S 的有限子集 D 使得 $\sum_{i \in D} \pi_i > 1 - \varepsilon$, 即

$$\sum_{i \notin D} \pi_i < \varepsilon. \quad (1.7.6)$$

因为 D 有限, 所以

$$\lim_{n \rightarrow \infty} \sum_{i \in D} |\pi_i^{(n)} - \pi_i| = \sum_{i \in D} \lim_{n \rightarrow \infty} |\pi_i^{(n)} - \pi_i| = 0. \quad (1.7.7)$$

进一步, 上式表明存在 $N \geq 1$ 使得当 $n \geq N$ 时, 我们有

$$\left| \sum_{i \in D} \pi_i^{(n)} - \sum_{i \in D} \pi_i \right| \leq \sum_{i \in D} |\pi_i^{(n)} - \pi_i| < \varepsilon.$$

因为 $\pi^{(n)}$ 与 π 都是分布, 所以 $\sum_i \pi_i^{(n)} = \sum_i \pi_i = 1$, 从而,

$$\sum_{i \in D^c} \pi_i^{(n)} - \sum_{i \in D^c} \pi_i = \sum_{i \in D} \pi_i - \sum_{i \in D} \pi_i^{(n)}.$$

因此,

$$\left| \sum_{i \in D^c} \pi_i^{(n)} - \sum_{i \in D^c} \pi_i \right| = \left| \sum_{i \in D} \pi_i^{(n)} - \sum_{i \in D} \pi_i \right| < \varepsilon.$$

这结合 (1.7.6) 表明 $\sum_{i \in D^c} \pi_i^{(n)} < 2\varepsilon$. 进一步, 这又结合 (1.7.6) 与 (1.7.7) 表明

$$\sum_{i \in S} |\pi_i^{(n)} - \pi_i| = \sum_{i \in D} |\pi_i^{(n)} - \pi_i| + \sum_{i \notin D} |\pi_i^{(n)} - \pi_i| \leq \sum_{i \in D} |\pi_i^{(n)} - \pi_i| + \sum_{i \notin D} \pi_i^{(n)} + \sum_{i \notin D} \pi_i < 4\varepsilon.$$

因此, 结论成立. \square

命题1.7.9. 假设 \mathbf{P} 不可约、正常返, π 是不变分布. 对任意 $i \in S$, 任意初分布 μ , 都有

$$P_\mu \left(\lim_{n \rightarrow \infty} \sum_{i \in S} \left| \frac{V_i(n)}{n} - \pi_i \right| = 0 \right) = 1, \quad \lim_{n \rightarrow \infty} \sum_{i \in S} \left| \frac{1}{n} \sum_{m=0}^{n-1} P_\mu(X_m = i) - \pi_i \right| = 0.$$

特别地, $\lim_{n \rightarrow \infty} \sum_{i \in S} \left| \frac{1}{n} \sum_{m=0}^{n-1} p_{ji}^{(m)} - \pi_i \right| = 0$.

证: 因为 $0 \leq \frac{V_i(n)}{n} \leq 1$, 所以根据有界收敛定理可以推出 $E_\mu \frac{V_i(n)}{n} \rightarrow \pi_i$. 由于

$$E_\mu \frac{V_i(n)}{n} = \frac{1}{n} \sum_{m=0}^{n-1} P_\mu(X_m = i) =: a_i(n),$$

因此 $a_i(n) \rightarrow \pi_i$, $\forall i$, 并且 $a_i(n)$, $i \in S$ 是分布列. 由引理 1.7.8 知 $\sum_i |a_i(n) - \pi_i| \rightarrow 0$. 这即是第二个结论 $\sum_{j \in S} \left| \frac{1}{n} \sum_{m=0}^{n-1} P_\mu(X_m = j) - \pi_j \right| \rightarrow 0$. 特别地, 取 $\mu_j = 1$, $\mu_k = 0$, $\forall k \notin j$, 则 $P_\mu(X_m = i) = p_{ji}^{(m)}$. 因此, 第三个结论 $\sum_{i \in S} \left| \frac{1}{n} \sum_{m=0}^{n-1} p_{ji}^{(m)} - \pi_i \right|$ 成立.

下面, 我们证明第一个结论. 为了符号更清晰, 我们将样本 ω 写出来, 记

$$V_i(n, \omega) = V_i(n) = \sum_{m=0}^{n-1} \mathbf{1}_{\{X_m(\omega) = i\}}.$$

一方面, 当 ω 取定时, $\frac{V_i(n, \omega)}{n}$, $i \in S$ 是分布列. 另一方面, 记

$$\Omega_0 = \left\{ \lim_{n \rightarrow \infty} \frac{1}{n} V_i(n) = \pi_i, \forall i \in S \right\}.$$

根据(1.7.5) 与概率的次可列可加性,

$$P_\mu(\Omega_0^c) \leq \sum_i P_\mu \left(\lim_{n \rightarrow \infty} \frac{1}{n} V_i(n) = \pi_i \text{ 不成立} \right) = 0, \text{ 即 } P_\mu(\Omega_0) = 1.$$

于是, 对任意 $\omega \in \Omega_0$, 都有 $\frac{1}{n} V_i(n, \omega) \rightarrow \pi_i$, $\forall i \in S$.

根据引理 1.7.8, 结合上面的两个方面, 我们有 $\sum_{j \in S} \left| \frac{V_j(n, \omega)}{n} - \pi_j \right| \rightarrow 0$, $\forall \omega \in \Omega_0$. 这表明,

$$\Omega_0 \subseteq \left\{ \lim_{n \rightarrow \infty} \sum_{i \in S} \left| \frac{V_i(n)}{n} - \pi_i \right| = 0 \right\} =: \Omega_1. \quad (1.7.8)$$

因此 $P_\mu(\Omega_1) \geq P_\mu(\Omega_0) = 1$, 从而 $P_\mu(\Omega_1) = 1$. \square

定理 1.7.10 (遍历定理). 假设 \mathbf{P} 不可约、正常返, π 是不变分布. 假设 f 是 S 上的有界函数.

那么, 对任意初分布 μ , 都有

$$P_\mu \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} f(X_m) = \sum_{i \in S} \pi_i f(i) \right) = 1.$$

证: 沿用 (1.7.8) 中定义的 Ω_1 . 由命题 1.7.9, $P_\mu(\Omega_1) = 1$. 因为

$$\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) = \frac{1}{n} \sum_{i \in S} f(i) V_i(n) = \sum_{i \in S} \frac{V_i(n)}{n} f(i),$$

所以在 Ω_1 上,

$$\left| \frac{1}{n} \sum_{m=0}^{n-1} f(X_m) - \sum_{i \in S} \pi_i f(i) \right| \leq M \sum_{i \in S} \left| \frac{V_i(n)}{n} - \pi_i \right| \rightarrow 0,$$

其中, $M = \sup_{i \in S} |f(i)| < \infty$. 因此, 结论成立. \square

事实上, 定理 1.7.10 中的函数 f 可以更一般, 只需要满足 $\sum_i \pi_i f(i)$ 收敛即可(见定理 1.7.11). 由于该定理的证明有一定难度, 所以我们将它放在本节的补充知识中.

定理1.7.11 (遍历定理). 假设 \mathbf{P} 不可约、正常返, π 是不变分布. 假设 f 是 S 上的函数, 满足 $\sum_i \pi_i |f(i)| < \infty$. 那么, 对任意初分布 μ , 都有

$$P_\mu \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} f(X_m) = \sum_{i \in S} \pi_i f(i) \right) = 1.$$

四、应用.

1. 公式 (1.7.1) 的应用.

假设 \mathbf{P} 不可约、正常返. 结合命题 1.7.2 和定理 1.7.5, 我们推出公式 (1.7.1) 成立, 即

$$\pi_i = \frac{1}{E_i \sigma_i}.$$

此关系式可以有如下的两类应用. 一类是通过求解不变分布来得到 $E_i \sigma_i$, 并判断正常返性, 比如, 例 1.7.12; 另一类是通过求解平均回访时间 $E_i \sigma_i$ 来判别不变分布是否存在, 或者计算不变分布, 比如, 例 1.7.13 与例 1.5.10.

例1.7.12 (例 1.1.8 续, Ehrenfest 模型的不变分布). 因为 Ehrenfest 模型是生灭链, 所以我们可以仿照例 1.2.12 计算其不变分布, 并得到如下表达式:

$$\pi_i = \frac{N(N-1)\cdots(N-(i-1))}{i!} \pi_0 = C_N^i \pi_0, \quad \forall i \in S = \{0, 1, \dots, N\},$$

由归一化条件 $\sum_{i=0}^N \pi_i = 1$ 知 $\pi_0 = 2^{-N}$. 事实上, π 就是参数为 $N, \frac{1}{2}$ 的二项分布 $B(N, \frac{1}{2})$. 进一步, 我们有

$$E_i \sigma_i = \frac{1}{\pi_i} = \frac{2^N}{C_N^i}.$$

特别地, 设 $N = 2M$, $M = 10^4$, 单位时间为 1 秒, 则

$$E_0 \sigma_0 = \frac{1}{\pi_0} = 2^{20000} \text{ (秒)}, \quad E_M \sigma_M = \frac{1}{\pi_M} = \frac{2^{2M}}{C_{2M}^M} \approx 100\sqrt{\pi} \text{ (秒)}.$$

其中, 最后一个约等号用到了 Stirling 公式 (1.5.12), 其含义是左右两边之商约为 1.,

例1.7.13 (老化过程). 设 $S = \mathbb{Z}_+$. 对任意 $i \geq 0$, 转移概率为 $p_{i,i+1} = 1 - p_{i,0} = p_i$, $\forall i \geq 0$. 其中, $0 < p_i < 1$, $\forall i \geq 0$.

由于粒子只能往前跳 1 步或者跳回 0, 粒子也可以在状态 0 原地跳跃. 考虑一个从 0 出发的粒子. 它在跳跃 n 步时首次返回 0 当且仅当它先往前跳 $n-1$ 步 (于是到达状态 $n-1$), 然后在第 n 步跳回 0. 换句话说,

$$P_0(\sigma_0 = 1) = 1 - p_0, \quad P_0(\sigma_0 = n) = p_0 \cdots p_{n-2}(1 - p_{n-1}), \quad n \geq 2.$$

于是

$$\begin{aligned} E_0\sigma_0 &= \sum_{n=1}^{\infty} nP_0(\sigma_0 = n) \\ &= 1 - p_0 + \sum_{n=2}^{\infty} np_0 \cdots p_{n-2}(1 - p_{n-1}) = 1 + \sum_{n=1}^{\infty} p_0 \cdots p_{n-1}. \end{aligned}$$

将上式右边的级数记为 R . 则该马氏链正常返(等价于不变分布存在)当且仅当 $R < \infty$, 此时 $\pi_0 = \frac{1}{E_0\sigma_0}$.

例1.7.14. 在中国象棋的棋盘上有一只“马”, 起点是象棋开局时它所在位置. 每经过一个单位时间, 它移动一步(在象棋规则允许的范围内等可能选一个目的地, 然后跳跃). 经过 T 步后它(首次)回到起点. 试求 ET .

因为马的运动可以视为图 $G = (V, E)$ 上的随机游动, 其中 V 就是棋盘中的所有 9×10 个交叉点, 所以在居中的 5×6 个交叉点上, 马有 8 个(规则允许的)目的地, 也就是说, 这些交叉点的度是 8; 在 4 个角上, 马只有 2 个目的地, 也就是说, 这 4 个交叉点的度为 2. 依据象棋规则, 我们还可以数出, 度分别为 3、4、6 的交叉点的数目分别为 8, 26, 22. 我们将每一类的交叉点的数目乘以度再求和, 得到

$$\sum_{j \in V} d_j = 4 \times 2 + 8 \times 3 + 26 \times 4 + 22 \times 6 + 30 \times 8 = 508.$$

将马的起点记为 o , 则 $d_o = 2$. 因此, 由例 1.2.13, $\pi_o = \frac{d_o}{508} = \frac{1}{254}$, 从而 $ET = \frac{1}{\pi_o} = 254$.

例1.7.15 (例 1.1.10, 1.5.10 续, 规则树 \mathbb{T}^d 上的 λ -biased 随机游动). 在例 1.5.10 中, 我们看到, 若 $|i| = |j|$, 则存在 \mathbb{T}^d 到自身的一一映射 φ 使得: 如果 $\{X_n\}$ 是从 i 出发的马氏链, $Y_n = \varphi(X_n)$, $\forall n \geq 0$, 那么 $\{Y_n\}$ 就是从 j 出发的马氏链. 根据首达时的定义, $\sigma_i^{(X)} = \sigma_j^{(Y)}$. 这表明 $E_i\sigma_i = E_j\sigma_j$. 因此, 规则树 \mathbb{T}^d 上的 λ -biased 随机游动的不变分布 π 如果存在, 则它也具有球对称性, 即, 如果 $|i| = |j|$, 那么 $\pi_i = \pi_j$. 换句话说, 存在 c_n , $n \geq 0$ 使得 $\pi_i = c_{|i|}$, $\forall i \in \mathbb{T}^d$. 取 $A = \{i : |i| \geq n\}$, 公式 (1.2.2) 就变为

$$(d^{n-1}c_{n-1}) \cdot \frac{d}{\lambda + d} = (d^n c_n) \cdot \frac{\lambda}{\lambda + d},$$

即 $c_n = c_{n-1} \cdot \frac{1}{\lambda}$, $\forall n \geq 1$. 迭代后, 我们有 $c_n = c_0 \frac{1}{\lambda^n}$, $\forall n \geq 1$. 最后, 归一化条件 $\sum_{n=0}^{\infty} d^n c_n = 1$ 表明 $c_0 = \frac{1}{R}$, 其中 $R = \sum_{n=0}^{\infty} (\frac{d}{\lambda})^n$. 综上, 当 $\lambda \leq d$ 时, $R = \infty$, 该马氏链不存在不变分布; 当 $\lambda > d$ 时, $R < \infty$, 不难验证 $\pi_i = \frac{1}{R} \cdot \frac{1}{\lambda^{|i|}}$, $\forall i \in \mathbb{T}^d$ 就是不变分布, 此时, 该马氏链正常返.

2. 遍历定理的应用

例1.7.16. 假设 $\{X_n\}$ 是一个马氏链, 状态空间为 S , 转移矩阵 \mathbf{P} 不可约、正常返. 对任意 $i, j \in S$, 考虑从 i 到 j 的跳跃出现的频率

$$f_{ij}(n) = \frac{1}{n} \sum_{m=0}^n \mathbf{1}_{\{X_m=i, X_{m+1}=j\}}.$$

证明: $f_{ij}(n)$ 几乎必然收敛于 $\pi_i p_{ij}$.

证: 记 $Y_n = (X_n, X_{n+1})$, $\forall n$. 不难验证, $\{Y_n\}$ 是一个不可约、正常返马氏链, 状态空间为 $\tilde{S} = \{(i, j) : p_{ij} > 0\}$, 转移概率为

$$\tilde{p}_{(i,j),(j,k)} = p_{jk}, \quad \tilde{p}_{(i,j),(\ell,k)} = 0, \text{ 若 } \ell \neq k. \quad (1.7.9)$$

不变分布为 $\tilde{\pi}_{(i,j)} = \pi_i p_{ij}$, $\forall (i, j) \in \tilde{S}$. 对 $\{Y_n\}$ 用遍历定理便知

$$f_{ij}(n) = \frac{1}{n} \sum_{m=0}^n \mathbf{1}_{\{Y_m=(i,j)\}} \xrightarrow{a.s.} \tilde{\pi}_{(i,j)} = \pi_i p_{ij}.$$

假设 \mathbf{P} 不可约、正常返, π 为其不变分布. 我们知道, 当 $n \rightarrow \infty$ 时, 任意状态 i 出现的频率会几乎必然收敛到 π_i , 即, 存在 Ω_0 使得 $P(\Omega_0) = 1$, 且 $\forall \omega \in \Omega_0$, $\lim_{n \rightarrow \infty} \frac{1}{n} V_i(n, \omega) = \pi_i$. 特别地, 给定 ω , 对正整数的任意子列 n_1, n_2, \dots 也有

$$\lim_{r \rightarrow \infty} \frac{1}{n_r} V_i(n_r, \omega) = \pi_i. \quad (1.7.10)$$

从下面的例题可以看到, 适当地选取 n_1, n_2, \dots 可以帮助我们巧妙地解决问题.

例1.7.17. 假设马氏链不可约, 正常返. 对给定正整数 m , 令 $\sigma := \inf\{n \geq m : X_n = i\}$. 证明:

$$E_i \sum_{n=0}^{\sigma-1} \mathbf{1}_{\{X_n=j\}} = \pi_j E_i \sigma. \quad (1.7.11)$$

证: 首先, 我们验证等号两边所涉及的期望都是有限的. 因为 $\sum_{n=0}^{\sigma-1} \mathbf{1}_{\{X_n=j\}} \leq \sigma$, 所以我们需要验证 $E_i \sigma < \infty$ 即可. 考虑一个从 i 出发的粒子, 将其第 r 次回访 i 的时间记为 T_r . 那么, $T_m \geq m$, 这表明 $\sigma \leq T_{m+1}$. 因此, $E_i \sigma \leq E_i T_{m+1} < (m+1) E_i \sigma_i < \infty$, 其中, 最后一步用到了该马氏链正常返.

其次, 因为 σ 是一个停时, $X_\sigma = i$, 并且该马氏链常返, 所以我们可以将轨道切割成独立同分布的有限步轨道, 具体操作如下: 记 $S_0 = 0$, $S_r = \inf\{n \geq m + S_{r-1} : X_n = i\}$, $r \geq 1$. 令 $\xi_r = S_r - S_{r-1}$, $r \geq 1$. 根据强马氏性(命题 1.4.11), ξ_1, ξ_2, \dots 独立同分布, 事实上, $(X_{S_{r-1}}, \dots, X_{S_r-1})$, $r \geq 1$ 是独立同分布的.

然后, 记 $\eta_r = \sum_{m=S_{r-1}}^{S_r-1} \mathbf{1}_{\{X_m=i\}}$, 我们有 η_1, η_2, \dots 独立同分布. 将(1.7.10)中的 n_r 取为 S_r . 我们有 $V_i(S_r, \omega) = \eta_1 + \dots + \eta_r$, 因此,

$$\lim_{r \rightarrow \infty} \frac{\eta_1 + \dots + \eta_r}{\xi_1 + \dots + \xi_r} = \lim_{r \rightarrow \infty} \frac{V_i(S_r, \omega)}{S_r} = \pi_i.$$

最后, 根据强大数定律,

$$\lim_{r \rightarrow \infty} \frac{\eta_1 + \cdots + \eta_r}{\xi_1 + \cdots + \xi_r} = \lim_{r \rightarrow \infty} \frac{\frac{1}{n}(\eta_1 + \cdots + \eta_r)}{\frac{1}{n}(\xi_1 + \cdots + \xi_r)} = \frac{E\eta_1}{E\xi_1}.$$

因此, $E\eta_1 = \pi_i E\xi_1$. 由 $E\xi_1 = E_i \sigma$ 而 $E\eta_1 = E_i \sum_{n=0}^{\sigma-1} \mathbf{1}_{\{X_n=j\}}$, 我们便可推出结论成立.

我们还可以结合遍历定律与大数定律, 进行推广应用, 譬如, 下面的例题.

例1.7.18. 假设某男同学每天中午去甲食堂或乙食堂之一吃饭. 如果他某天去甲食堂, 则第二天去乙食堂的可能性为0.7; 如果他某天去乙食堂, 则第二天去甲食堂的可能性为0.5. 假设他每次去甲食堂, 餐费(单位: 元)独立地等可能为14, 15, 16; 每次去乙食堂, 餐费独立地等可能为9, 10, 11. 请问: 他一顿午餐平均花费多少元?

解: 我们可以建立一个马氏链描述该同学前往的食堂, 将甲食堂记为0, 乙食堂记为1. 将他第 n 天去的食堂记为 X_n , 则 $\{X_n\}$ 是一个马氏链, 状态空间为 $S = \{0, 1\}$, 转移概率为 $p_{01} = 0.7$, $p_{00} = 0.3$; $p_{10} = p_{11} = 0.5$. 我们可以解出其对应的不变分布为 $\pi_0 = \frac{5}{12}$, $\pi_1 = \frac{7}{12}$.

将他每次去农园的花费记为 ξ_1, ξ_2, \dots , 它们独立同分布. 假设他在 n 天中去了农园 $V_0(n)$ 次, 那么他去农园的总消费就为 $F(n) := \xi_1 + \cdots + \xi_{V_0(n)}$. 于是,

$$\frac{1}{n} F(n) = \frac{V_0(n)}{n} \times \frac{\xi_1 + \cdots + \xi_{V_0(n)}}{V_0(n)} \xrightarrow{a.s.} \pi_0 \times E\xi_1,$$

其中, 第一项用了遍历定理, 而第二项用了强大数定律.

同理, 我们可以分析他去燕南的平均花费. 因此, 所求为

$$\frac{5}{12} \times \frac{14 + 15 + 16}{3} + \frac{5}{12} \times \frac{9 + 10 + 11}{3} = \frac{145}{12}.$$

补充知识.

假设 \mathbf{P} 不可约、常返. 给定状态 o , 考虑公式 (1.7.2) 定义的 μ , 即,

$$\mu_i := E_o \sum_{n=0}^{\sigma_o-1} \mathbf{1}_{\{X_n=i\}}, \quad \forall i \in S,$$

由命题 1.7.4, μ 满足不变方程. 下面的命题 1.7.20 表明, 在忽略常数倍的意义下, μ 实际上是 \mathbf{P} 唯一的不变测度. 为此, 我们先给出一个引理.

引理1.7.19. 假设 \mathbf{P} 不可约. 假设 $0 \leq \lambda_i \leq \infty$, $\forall i$, 且 $\lambda = \{\lambda_i, i \in S\}$ 满足不变方程. 那么, 以下三种情况之一属实: (1) $0 < \lambda_i < \infty$, $\forall i$; (2) $\lambda_i = 0$, $\forall i$; (3) $\lambda_i = \infty$, $\forall i$.

证: 首先, 存在 o 使得 $0 < \lambda_o < \infty$, 那么(1) 属实. 这是因为, 根据不变方程, $\lambda = \lambda \mathbf{P}$, 进一步 $\lambda = \lambda \mathbf{P}^n, \forall n \geq 1$. 换句话说,

$$\lambda_i = \sum_{j \in S} \lambda_j p_{ji} = \sum_{j \in S} \lambda_j p_{ji}^{(n)}, \quad \forall i, j.$$

一方面, 取 $j = o$, 由不可约知对任意 i , 存在 n 使得 $p_{oi}^{(n)} > 0$, 因此, $\lambda_i \geq \lambda_o p_{oi}^{(n)} > 0$. 另一方面, 取 $i = o$. 由不可约知对任意 j , 存在 n 使得 $p_{jo}^{(n)} > 0$, 因此, $\lambda_o \geq \lambda_j p_{jo}^{(n)}$. 这表明 $\lambda_j < \infty$. 综合这两方面, 我们有 $0 < \lambda_i < \infty, \forall i$.

然后, 如果存在 o 使得 $\lambda_o = 0$, 那么(2) 属实. 这是因为对任意 j , 都有 $0 = \lambda_o \geq \lambda_j p_{jo}^{(n)}, \forall n$. 于是, 对任意 j , 取 n 使得 $p_{jo}^{(n)} > 0$, 则根据不变方程知 $0 = \nu_o \geq \nu_j p_{jo}^{(n)}$. 这表明 $\nu_j = 0, \forall j$.

最后, 结合上面的结论, 反正法可以证明, 如果存在 o 使得 $\lambda_o = \infty$, 那么(3) 属实. \square

命题1.7.20. 假设 \mathbf{P} 不可约、常返. 假设 $0 \leq \lambda_i \leq \infty, \forall i$; $\lambda = \{\lambda_i, i \in S\}$ 满足不变方程, 并且 $\lambda_o = 1$. 那么, $\lambda_i = \mu_i, \forall i \in S$.

证: 首先, 根据命题 1.7.4 与引理 1.7.19, 对所有 i 都有 $0 < \lambda_i < \infty$.

然后, 我们证明 $\lambda_i \geq \mu_i, \forall i \in S$. 根据不变方程,

$$\lambda_i = \sum_{j \in S} \lambda_j p_{ji} = \lambda_o p_{oi} + \sum_{j \neq o} \lambda_j p_{ji} = p_{oi} + \sum_{j \neq o} \lambda_j p_{ji},$$

其中, 最后一个等号是因为命题假设 $\lambda_o = 1$. 下面, 我们用迭代, 当 $j \neq o$ 时, 对 λ_j 进行上面的分解操作. 于是, 我们有

$$\lambda_i = p_{oi} + \sum_{j \neq o} \left(p_{oj} + \sum_{k \in S, k \neq o} \lambda_k p_{kj} \right) p_{ji} = p_{oi} + \sum_{j \neq o} p_{oj} p_{ji} + \sum_{j \neq o} \sum_{k \neq o} \lambda_k p_{kj} p_{ji}.$$

将上式中最后一个求和号中的 λ_k 再迭代一次, 我们有

$$\lambda_i = p_{oi} + \sum_{j \neq o} p_{oj} p_{ji} + \sum_{j, k \neq o} \lambda_o p_{ok} p_{kj} p_{ji} + \sum_{j, k, \ell \neq o} \lambda_\ell p_{\ell k} p_{kj} p_{ji}. \quad (1.7.12)$$

将上式右边第一个求和号中的 j 改写成 i_1 , 第二个求和号中的 k, j 分别改写成 i_1, i_2 . 根据 $\lambda_\ell \geq 0, \forall \ell$, 我们有

$$\lambda_i \geq p_{oi} + \sum_{i_1 \neq o} p_{oi_1} p_{i_1 i} + \sum_{i_1, i_2 \neq o} \lambda_o p_{oi_1} p_{i_1 i_2} p_{i_2 i}.$$

在右边的三项中, 分别将 i 分别视为 X_1, X_2, X_3 的值, 那么, 上面的不等式就可以改写为

$$\begin{aligned} \lambda_i &\geq P_o(X_1 = i) + P(X_1 \neq o, X_2 = i) + P_o(X_1, X_2 \neq o, X_3 = i) \\ &= P_o(\sigma_o > 1, X_1 = i) + P(\sigma_o > 2, X_2 = i) + P_o(\sigma_o > 3, X_3 = i). \end{aligned}$$

我们可以将 (1.7.12) 中的 λ_ℓ 再改写, 继续迭代下去……便知, 对任意 $n \geq 1$ 都有

$$\lambda_i \geq P_o(\sigma_o > 1, X_1 = i) + P_o(\sigma_o > 2, X_2 = i) + \cdots + P_o(\sigma_o > n, X_n = i).$$

令 $n \rightarrow \infty$. 我们有,

$$\lambda_i \geq \sum_{n=1}^{\infty} P_o(\sigma_o > n, X_n = i) = E_o \sum_{n=1}^{\infty} \mathbf{1}_{\{\sigma_o > n, X_n = i\}} = E_o \sum_{n=1}^{\sigma_o - 1} \mathbf{1}_{\{X_n = i\}} = \mu_i, \forall i.$$

最后, 我们证明 $\lambda = \mu$. 这是因为, 如果我们令 $\nu_i = \lambda_i - \mu_i, \forall i$, 那么 $\nu = \{\nu_i, i \in S\}$ 也满足不变方程, 且 $\nu_o = 0$. 根据引理 1.7.19, $\nu_j = 0, \forall j$. \square

根据上述命题, 对于不可约、常返的马氏链, 如果不变方程有一个非负解 λ , 那么存在常数 c 使得 $\lambda_i = c\mu_i, \forall i$. 事实上, $c = \lambda_o$. 于是, 进一步, 如果 λ 可以归一化, 那么我们将其归一化便得到不变分布; 如果 λ 不可以归一化, 那么不变分布不存在. 这即是如下推论.

推论 1.7.21. 假设 \mathbf{P} 不可约. 如果 $\lambda_i, i \in S$ 是不变测度, 但 $\sum_{i \in S} \lambda_i = \infty$. 那么, 不变分布不存在.

证: 假设 \mathbf{P} 非常返, 那么, 由推论 1.5.17 知不变分布不存在.

假设 \mathbf{P} 常返, 那么, 我们用反证法进行证明. 假设不变分布 π 存在. 由引理 1.7.19 知 $\pi_i, \lambda_i > 0$ 对所有 i 都成立. 特别地, $\pi_o, \lambda_o > 0$, 因此 $\{\frac{1}{\pi_o} \pi_i, i \in S\}$ 与 $\{\frac{1}{\lambda_o} \lambda_i, i \in S\}$ 都是不变测度. 于是, 由命题 1.7.20,

$$\frac{\pi_i}{\pi_o} = \mu_i = \frac{\lambda_i}{\lambda_o}, \quad \forall i \in S.$$

这表明

$$\pi_i = \frac{\pi_o}{\lambda_o} \lambda_i, \quad \forall i \in S.$$

然而, 左边的求和为 1, 但右边的求和发散. 矛盾! 因此假设不成立, 即不变分布不存在. \square

值得一提的是, 在命题 1.7.20 中, \mathbf{P} 常返的这一假设不能去掉. 反例如下: 考虑 \mathbb{Z} 上的非对称紧邻随机游动, 比如说,

$$p_{i,i+1} = 1 - p_{i,i-1} = p \in (0, 0.5).$$

那么, $\pi_i = 1$ 和 $\mu_i = p^i / (1-p)^i, \forall i \in \mathbb{Z}$ 都是 \mathbf{P} 不变的, 但它们之间不是相差常数倍的关系.

定理 1.7.11 的证明: 假设 $X_0 = i$. 将轨道分解成独立同分布的 i 游戏. 具体地, 将马氏链第 r 次返回 i 的时刻记为 T_r , 即

$$T_0 = 0, T_1 = \sigma_i = \inf\{n \geq 1 : X_n = i\}, T_{r+1} = \inf\{n \geq T_r : X_n = i\}.$$

记

$$\xi_s = \sum_{m=T_{s-1}}^{T_s-1} f(X_m), \quad \eta_s = \sum_{m=T_{s-1}}^{T_s-1} |f(X_m)|.$$

那么, ξ_1, ξ_2, \dots 独立同分布, η_1, η_2, \dots 独立同分布. 因为

$$E\eta_1 = \sum_{j \in S} |f(j)|\mu_j,$$

其中 μ_j 由(1.7.2) 定义, 即 $\mu_j = E_i \sum_{m=0}^{\sigma_i-1} \mathbf{1}_{\{X_m=j\}}$, 所以根据定理 1.7.5 可推出 $\mu_j = \pi_j E_i \sigma_i$, 从而 $E\eta_1 = E_i \sigma_i \times \sum_j \pi_j |f(j)| < \infty$. 这表明 ξ_1 的期望存在, 同理, $E\xi_1 = E_i \sigma_i \times \sum_j \pi_j f(j)$.

将 $V_i(n) = \sum_{m=0}^{n-1} \mathbf{1}_{\{X_m=i\}}$ 简记为 r . 那么, $T_{r-1} \leq n < T_r$, 并且, 根据强大数律, 当 $n \rightarrow \infty$ 时, n/r 几乎必然收敛于 $E_i \sigma_i$. 由

$$\sum_{m=0}^{n-1} f(X_m) = \sum_{s=1}^r \xi_s + \sum_{m=T_{i,s}}^{n-1} f(X_m)$$

我们可以推出

$$\frac{1}{n} \sum_{m=0}^{n-1} f(X_m) = \frac{r}{n} \times \frac{1}{r} \sum_{s=1}^r \xi_s + \frac{1}{n} \sum_{m=T_s}^{n-1} f(X_m).$$

令 $n \rightarrow \infty$, 强大数律表明

$$\frac{r}{n} \times \frac{1}{r} \sum_{s=1}^r \xi_s \xrightarrow{\text{a.s.}} \frac{1}{E_i \sigma_i} E\xi_1 = \sum_j \pi_j f(j).$$

而引理 4.0.25 表明

$$\frac{1}{n} \left| \sum_{m=T_s}^{n-1} f(X_m) \right| \leq \frac{\eta_n}{n} \xrightarrow{\text{a.s.}} 0.$$

从而, 结论成立. □

习题

1. 马氏链不可约是不变分布唯一的必要条件吗? 如果是, 请证明之; 如果不是, 请改进之.
2. 假设状态空间 S 有限. (1) 证明存在正常返态, (2) 证明存在不变分布.
3. 仿照命题 1.7.4, 证明例 1.7.17 中的 $E_i \sum_{n=0}^{\sigma_i-1} \mathbf{1}_{\{X_n=j\}}$, $i \in S$ 是不变测度. (注: 将其归一化可得 (1.7.11) 的另外一个证明.)
4. 某考试有 100 道判断题. 假设某题的正确答案为“是”(或者, 为“否”), 则下一题的正确答案为“是”的概率为 0.6 (或者, 为 0.5). 某学生所有题都独立地以概率 p 回答“是”(以概率 $1-p$ 回答“否”). (1) 试建立一个马氏链模型刻画该生每道题回答正确与否, (2) 估计该生能得多少分, (3) 求 p 的最优选择.

5. 假设 $\{S_n\}$ 是一维随机游动, 步长分布为 $P(\xi = k) = \frac{1}{6}$, $k = 1, \dots, 6$. 令 A_n = “ S_n 能被13整除”. 试求: $\lim_n P_0(A_n)$.
6. 假设 $\{X_n\}$ 是离散圆周 \mathbb{S}_N 上的随机游动(参见例 1.2.8). 用两种不同的方法求 $E_0\sigma_0$.
7. 假设 $\{X_n\}$ 是 \mathbb{Z} 上的马氏链, $p_{i,i+1} = 1 - p_{i,i-1} = p_i$, $\forall i \in \mathbb{Z}$; $p_{i+d} = p_i$, $\forall i \in \mathbb{Z}$; 且 $p_0, \dots, p_{d-1} \in (0, 1)$. 证明: $\frac{1}{n}X_n$ 几乎必然收敛. (提示: 取 $Y_n \in S = \{0, 1, \dots, d-1\}$ 满足 $Y_n \equiv X_n \pmod{d}$, 则 $\{Y_n\}$ 是 S 上的马氏链.)
8. 假设 $\{X_n\}$ 是 S 上以 \mathbf{P} 为转移矩阵的马氏链. 令 $Y_n = (X_n, X_{n+1})$, $\forall n \geq 0$. 证明: (1) $\{Y_n\}$ 是 $\tilde{S} := \{(i, j) : p_{ij} > 0\}$ 上的马氏链, 转移概率由(1.7.9)给出, (2)如果 $\{X_n\}$ 不可约(常返, 或正常返), 那么, $\{Y_n\}$ 也相应地不可约(常返, 或正常返).
9. 假设 $\{X_n\}$ 是不可约、正常返马氏链, π 为其不变分布. 用两种方法证明: 对任意 $\ell \geq 1$, $i_0, \dots, i_\ell \in S$, 都有
- $$\frac{1}{n} \left| \{0 \leq m \leq n-1 : X_m = i_0, X_{m+1} = i_1, \dots, X_{m+\ell} = i_\ell\} \right| \xrightarrow{\text{a.s.}} \pi_{i_0} p_{i_0 i_1} \cdots p_{i_{\ell-1} i_\ell}.$$
10. 假设马氏链不可约, 其转移阵 \mathbf{P} 是幂等的, 即 $\mathbf{P} = \mathbf{P}^2$. 证明: 对于任意 i, j , $p_{ij} = p_{jj}$.
11. 假设 $\{X_n\}$ 是 N 个顶点的完全图上的随机游动. (1) 求 $P_i(\sigma_i = n)$, $\forall n \geq 1$, 并利用之计算 $E_i\sigma_i$. (2) 根据不变分布的定义列方程并解出 π , 然后验证(1.7.1). (注: 在完全图中, 任意两个不同的顶点之间有且仅有一条边相连.)
12. 在例 1.7.16 中, 考虑从 i 出发的马氏链第 r 次回访 i 的时间 T_r . 令 $\xi_r = \mathbf{1}_{\{X_{T_{r-1}+1}=j\}}$, $r \geq 1$. 试仿照例 1.7.18 证明 $f_{ij}(n)$ 几乎必然收敛于 $\pi_i p_{ij}$.
- 13* 假设 $\{X_n\}$ 是 N 个顶点的完全图上的随机游动. 将 $\{X_n\}$ 走遍所有顶点的时间记为 T , 即 $T = \max_{i \in S} \sigma_i$. 求 $E_i T$.
- 14* 假设某马氏链不可约、正常返. (1) 假设观察该马氏链 n 步, 依次得到状态 i_0, \dots, i_n , 求该马氏链的转移概率矩阵的最大似然估计 $\hat{\mathbf{P}} = (\hat{p}_{ij})$, (2) 证明最大似然估计 $\hat{\mathbf{P}}$ 具有强相合性.

§1.8 强遍历定理

本节研究当 $n \rightarrow \infty$ 时, $p_{ij}^{(n)}$ 的极限问题. 假设 \mathbf{P} 正常返, 不变分布为 π . 那么根据遍历定理, 若 $p_{ij}^{(n)}$ 的极限存在, 则它应该为 π_i , 于是极限存在的必要条件是当 n 充分大时, $p_{ij}^{(n)} > 0$.

定义1.8.1. 若存在 N 使得对任意 $n \geq N$ 都有 $p_{ii}^{(n)} > 0$, 则称 i 为非周期(aperiodic)的.

命题1.8.2. 假设 \mathbf{P} 不可约, 存在非周期态. 那么, 所有状态都非周期, 并且对任意 $i, j \in S$, 存在 N 使得 $p_{ij}^{(n)} > 0, \forall n \geq N$.

证: 假设 i 非周期, 当 $m \geq N$ 时 $p_{ii}^{(m)} > 0$. 对任意 $j \in S$, 取 r, s 使得 $p_{ij}^{(r)}, p_{ji}^{(s)} > 0$. 于是, 当 $n \geq N + r + s$ 时, $m := n - r - s \geq N$, $p_{jj}^{(n)} \geq p_{ji}^{(s)} p_{ii}^{(m)} p_{ij}^{(r)} > 0$, 从而 j 非周期, 因此所有状态都非周期. 进一步, i, j, N, r 同上, 那么当 $n \geq N + r$ 时, $m := n - r \geq N$, $p_{ij}^{(n)} \geq p_{ii}^{(m)} p_{ij}^{(r)} > 0$, 从而命题成立. \square

定理1.8.3 (强遍历定理). 假设 \mathbf{P} 不可约、正常返、非周期. 那么, 对任意初分布 μ 都有

$$\lim_{n \rightarrow \infty} \sum_{j \in S} |P_\mu(X_n = j) - \pi_j| = 0,$$

其中, π 为 \mathbf{P} 的不变分布. 特别地,

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j, \quad \forall i, j \in S.$$

证: 考虑状态空间 $\tilde{S} = S \times S$ 以及 \tilde{S} 上的转移矩阵 \mathbf{R} :

$$r_{(i,j)(k,\ell)} := p_{ik} p_{j\ell}, \quad \forall (i, j), (k, \ell) \in \tilde{S}.$$

假设 $\{Z_n\}$ 是 \tilde{S} 上的以 \mathbf{R} 为转移矩阵的马氏链. 记 $Z_n = (W_n, Y_n)$. 为方便起见, 将 $\{Z_n\}$ 的初分布记为 ν . 我们在之后的证明中会选取特定的初分布, 不过目前 ν 可以是 \tilde{S} 上的任意分布.

下面, 我们分几步证明定理的结论.

第一步, 不难验证,

$$r_{(i,j)(k,\ell)}^{(n)} := p_{ik}^{(n)} p_{j\ell}^{(n)}, \quad \forall (i, j), (k, \ell) \in \tilde{S}. \quad (1.8.1)$$

对任意 $j \in S$,

$$\begin{aligned} P(W_n = k) &= \sum_{\ell \in S} P(W_n = k, Y_n = \ell) = \sum_{(i,j) \in \tilde{S}, \ell \in S} \nu_{(i,j)} j p_{ik}^{(n)} p_{j\ell}^{(n)} \\ &= \sum_{(i,j) \in \tilde{S}} \nu_{(i,j)} p_{ik}^{(n)} \left(\sum_{\ell \in S} p_{j\ell}^{(n)} \right) = \sum_{i \in S} \left(\sum_{j \in S} \nu_{(i,j)} \right) p_{ik}^{(n)}. \end{aligned}$$

同理,

$$P(Y_n = \ell) = \sum_{j \in S} \left(\sum_{i \in S} \nu_{(i,j)} \right) p_{j\ell}^{(n)}.$$

第二步, 令 $\tilde{\pi}_{(i,j)} := \pi_i \pi_j$, $\forall (i,j) \in \tilde{S}$. 不难验证, μ 是 \mathbf{R} 的不变分布.

第三步, \mathbf{R} 不可约. 对任意 $i, j, k, \ell \in S$, 由于 \mathbf{P} 不可约、非周期, 由定理 1.8.6 中的(2), 存在 $N_1, N_2 \geq 0$ 使得 $p_{ik}^{(n_1)}, p_{j\ell}^{(n_2)} > 0$ 对任意 $n_1 \geq N_1, n_2 \geq N_2$ 都成立. 于是 $r_{(i,j)(k,\ell)}^{(n)} > 0$ 对任意 $n \geq \max\{N_1, N_2\}$ 成立. 这表明 (i,j) 可达 (k,ℓ) , 从而 \mathbf{R} 不可约.

第四步, 假设 $\{Z_n\}$ 是 \tilde{S} 上以 \mathbf{R} 为转移矩阵的马氏链. 由 \mathbf{R} 有不变分布 $\tilde{\pi}$ 知 \mathbf{R} 正常返. 从而, \mathbf{R} 常返. 记 $Z_n = (W_n, Y_n)$. 令

$$\tau := \inf\{n \geq 0 : W_n = Y_n\}.$$

它表示 $\{W_n\}$ 与 $\{Y_n\}$ 首次相遇的时刻. 这即是二维马氏链 $\{Z_n\}$ 首达对角线

$$\tilde{D} = \{(i,i) : i \in S\}$$

的时间, 它不超过 $\{Z_n\}$ 首达某个具体的状态 (i,i) 的时间. 具体地, 取定 $i \in S$, 则 $\tau \leq \tau_{(i,i)} := \{n \geq 0 : Z_n = (i,i)\}$. 因为 $\{Z_n\}$ 常返, 由命题 1.5.6, $P(\tau_{(i,i)} < \infty) = 1$, 从而 $P(\tau < \infty) = 1$. 换句话说, $\{W_n\}$ 与 $\{Y_n\}$ 必然相遇.

第五步, 往证在 $\{W_n\}$ 与 $\{Y_n\}$ 相遇之后, 它们有同样的分布, 即

$$P(W_n = j, \tau \leq n) = P(Y_n = j, \tau \leq n).$$

假设 $\{W_n\}$ 与 $\{Y_n\}$ 在时刻 m 相遇于状态 i , 那么直观上, 从时刻 m 开始, 它们后续的运动都是从 i 出发、以 \mathbf{P} 为转移矩阵的马氏链, 从而它们在任何时刻有相同的分布. 严格的证明如下:

$$\begin{aligned} P(W_n = j, \tau \leq n) &= \sum_{m=0}^n \sum_{i \in S} P(W_n = j, \tau = m, Z_m = (i,i)) \\ &= \sum_{m=0}^n \sum_{i \in S} P(\tau = m, Z_m = (i,i)) P(Z_n \in A_j | \tau = m, Z_m = (i,i)), \end{aligned}$$

其中 $A_j = \{(j,k) : j \in S\}$. 如果我们将时刻 m 视为现在, 那么

$$\{\tau = m, Z_m = (i,i)\} = \{Z_0, \dots, Z_{m-1} \notin \tilde{D}, Z_m = (i,i)\}$$

是过去及现在的事件, 而 $Z_n \in A_j$ 则可视为一个将来的事件. 于是, 由 $\{Z_n\}$ 是马氏链知,

$$\begin{aligned} P(Z_n \in A_j | \tau = m, Z_m = (i,i)) &= P(Z_n \in A_j | Z_m = (i,i)) \\ &= \sum_{k \in S} r_{(i,i)(j,k)}^{(n-m)} = \sum_{k \in S} p_{ij}^{(n-m)} \cdot p_{ik}^{(n-m)} = p_{ij}^{(n-m)}. \end{aligned}$$

综上,

$$P(W_n = j, \tau \leq n) = \sum_{m=0}^n \sum_{i \in S} P(\tau = m, Z_m = (i, i)) \cdot p_{ij}^{(n-m)}.$$

同理, 令 $B_j = \{(k, j) : k \in S\}$, 我们有

$$\begin{aligned} P(Z_n \in B_j | \tau = m, Z_m = (i, i)) &= P(Z_n \in B_j | Z_m = (i, i)) \\ &= \sum_{k \in S} r_{(i,i)(k,j)}^{(n-m)} = \sum_{k \in S} p_{ik}^{(n-m)} \cdot p_{ij}^{(n-m)} = p_{ij}^{(n-m)}. \end{aligned}$$

从而, 也有

$$P(Y_n = j, \tau \leq n) = \sum_{m=0}^n \sum_{i \in S} P(\tau = m, Z_m = (i, i)) \cdot p_{ij}^{(n-m)}.$$

因此 $P(W_n = j, \tau \leq n) = P(Y_n = j, \tau \leq n)$.

第六步, 由于 $P(W_n = j) = P(W_n = j, \tau \leq n) + P(W_n = j, \tau > n)$, 并且对 Y_n 类似, 所以根据第五步的结论我们可以推出

$$P(W_n = j) - P(Y_n = j) = P(W_n = j, \tau > n) - P(Y_n = j, \tau > n).$$

从而,

$$\begin{aligned} &\sum_{j \in S} |P(W_n = j) - P(Y_n = j)| \\ &\leq \sum_{j \in S} (P(W_n = j, \tau > n) + P(Y_n = j, \tau > n)) = 2P(\tau > n). \end{aligned}$$

当 $n \rightarrow \infty$ 时, 第四步的结论 $P(\tau < \infty) = 1$ 表明上式右边趋于 0.

最后, 将 $\{Z_n\}$ 的初分布设置为

$$\nu_{(i,j)} = \mu_i \pi_j. \quad (1.8.2)$$

那么, 根据第一步结论, $P(W_n = j) = P_\mu(X_n = j)$, 可得 $P(Y_n = j) = P_\pi(X_n = j) = \pi_j, \forall j$, 其中, 我们用到 π 是 \mathbf{P} 的不变分布. 将其代入第六步的结论中就可推出定理成立. \square

例1.8.4 (Wright-Fisher模型). 该模型来源于遗传学, 不过, 此处我们借用社交网相互学习这个模型中的语言来描述它. 假设现在有 N 个人, 他们可以对某件事进行表态, “支持”或“反对”. 每一天, 每个人(称之为甲)独立地在 N 个人中随机挑选一人(称之为乙, 乙可以是甲自己). 如果乙前一天表态支持, 那么甲今天以概率 u 改为表态反对, 以概率 $1-u$ 跟着表态支持; 如果乙前一天表态反对, 那么甲今天以概率 v 改为表态支持, 以概率 $1-v$ 跟着表态反对, 其中 $0 < u, v < 1$. 记 X_n 为第 n 天表态支持的人数. 那么, $\{X_n\}$ 为马氏链, 取值于 $\{0, 1, 2, \dots, N\}$, 转移概率如下:

$$p_{ij} = C_N^j \rho_i^j (1 - \rho_i)^{N-j}, \quad \text{其中 } \rho_i = (1 - u) \frac{i}{N} + v \frac{N-i}{N}.$$

下面我们考虑选择支持的平均人数, 即 $EX_n = \sum_{j=0}^N jP(X_n = j)$. 因为转移概率全部为正, 所以该马氏链不可约、非周期, 因此 $\lim_{n \rightarrow \infty} P(X_n = j) = \pi_j$, 其中 π 为该马氏链唯一的不变分布, 其存在性由推论 1.7.7 保证. 故 $\lim_{n \rightarrow \infty} EX_n = \sum_{j=0}^N j\pi_j =: a$. 因为

$$\begin{aligned} a &= \sum_{j=0}^N j \sum_{i=0}^N \pi_i p_{ij} = \sum_{i=0}^N \pi_i \sum_{j=0}^N j p_{ij} = \sum_{i=0}^N \pi_i \sum_{j=0}^N j C_N^j \rho_i^j (1 - \rho_i)^{N-j} \\ &= \sum_{i=0}^N \pi_i N \rho_i = \sum_{i=0}^N \pi_i [(1-u)i + v(N-i)] = (1-u-v)a + vN. \end{aligned}$$

解得 $a = \frac{vN}{u+v}$. 特别地, 当 $u = v = \frac{1}{2}$ 时, 对任意 i , $\{p_{ij} : j \in S\}$ 都是二项分布 $B(N, \frac{1}{2})$. 于是, X_1, X_2, \dots 独立同分布, 它们都服从 $B(N, \frac{1}{2})$. 故, π 就是 $B(N, \frac{1}{2})$.

例1.8.5 (例 1.1.8 和例 1.7.12 续). 考虑 Ehrenfest 模型, 其中 $S = \{0, 1, \dots, N\}$, 转移概率为 $p_{0,1} = p_{N,N-1} = 1$, $p_{i,i-1} = 1 - p_{i,i+1} = \frac{i}{N}$, $1 \leq i \leq N-1$. 则 \mathbf{P} 不可约, 且在例 1.7.12 中得到不变分布 $\pi_i = C_N^i 2^{-N}$, $i = 0, 1, \dots, N$. 假设马氏链从状态 0 出发. 若经过奇数步, 则它向前与向后的跳跃次数必然不等, 于是它不可能回到状态 0. 这说明 $p_{00}^{(2n+1)} = 0$, $n = 0, 1, 2, \dots$. 因此, $\lim_{n \rightarrow \infty} p_{00}^{(n)} = \pi_0$ 不成立. 不难发现若 i 与 j 的奇偶性不同, 则 $p_{ij}^{(2n+1)}$, $n \geq N$ 且 $p_{ij}^{(2n)} = 0$, $n \geq 0$; 若 i 与 j 的奇偶性相同, 则 $p_{ij}^{(2n)}$, $n \geq N$ 且 $p_{ij}^{(2n+1)} = 0$, $n \geq 0$.

定理1.8.6. 假设 \mathbf{P} 不可约, 那么, 存在唯一的正整数 d 以及 S 的一个分割 D_0, D_1, \dots, D_{d-1} 使得下面的结论成立. 补充定义 $D_{nd+r} := D_r$, $n \geq 1$, 则

- (1) 对任意 $r \geq 0$, $i \in D_r$ 以及 $s \geq 0$, 均有 $\sum_{j \in D_{r+s}} p_{ij}^{(s)} = 1$;
- (2) 对任意 $r \geq 0$ 和 $i, j \in D_r$, 存在 $N \geq 0$ 使得对任意 $n \geq N$ 均有 $p_{ij}^{(nd)} > 0$.

上述定理的证明过程比较复杂, 也没有太多概率思想, 因此我们把它放在本节的补充知识中.

定理 1.8.6 中的(1) 表明从某一个区域 D_r 中的任意状态出发, 走一步必然到下一个区域 D_r 中, 取 $s = 1$ 即可. 如图 1.16 中的左图所示. 定理 1.8.6 中的(2) 则是为了保证 d 确实是这个马氏链的周期. 例如: 在图 1.16 中, 带圆圈的数字 i 表示状态 i . 在右图中, $S = \{0, 1, \dots, 5\}$, 转移概率为: $p_{i,i+1} = 1$, $i \neq 6$; $p_{6,0} = 1$. 那么, 该马氏链的周期应该为 6, $D_i = \{i\}$, $\forall i \in S$. 然而, 如果我们按照图示将 S 分成 $\tilde{d} = 3$ 个区域: $\tilde{D}_0 = \{0, 3\}$, $\tilde{D}_1 = \{1, 4\}$, $\tilde{D}_2 = \{2, 5\}$, 那么, 定理 1.8.6 中的(1) 仍然对 \tilde{d} 与 $\tilde{D}_1, \tilde{D}_2, \tilde{D}_3$ 也成立. 此时, $p_{00}^{((2m+1)\tilde{d})} = p_{00}^{(6m+3)} = 0$, 于是 $p_{00}^{(nd)} \rightarrow \pi_0$ 还是不成立. 本质原因是, 如果考虑 $Y_0 = X_0 = 0$, $Y_n = X_{nd}$, 那么, $\{Y_n\}$ 仍然具有周期性, 其周期为 2. 总结起来, 定理 1.8.6 中的(2) 在说, 如果我们考虑 $Y_n = X_{nd}$, $n = 0, 1, 2, \dots$, 那么 $\{Y_n\}$ 不再具有周期性, 这才使得 $p_{ii}^{(nd)} \rightarrow \pi_i$ 成为可能的事.

定义1.8.7. 定理 1.8.6 中的 d 被称为 \mathbf{P} 的周期 (period), 也称为 S 或马氏链的周期.

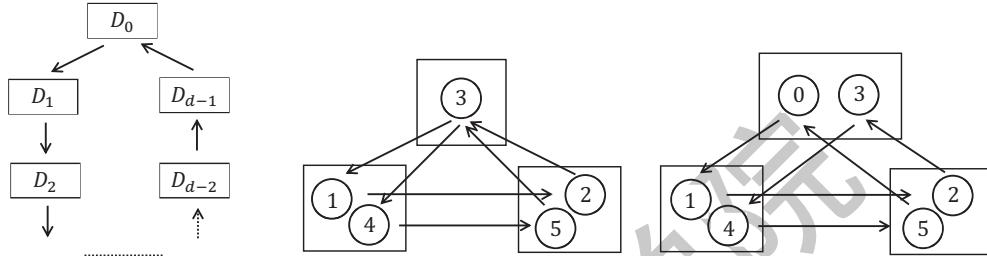


图 1.16 周期

例1.8.8. 假设 $S = \{1, 2, 3, 4\}$,

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.2 & 0.8 \\ 0.25 & 0.75 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{pmatrix}.$$

不难验证

$$\mathbf{P}^2 = \begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{B} \end{pmatrix}, \quad \text{其中 } \mathbf{A} = \begin{pmatrix} 0.35 & 0.65 \\ 0.45 & 0.55 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix}.$$

因为 \mathbf{A} 是 $\{1, 2\}$ 上不可约、非周期的转移矩阵; \mathbf{B} 是 $\{3, 4\}$ 上不可约、非周期的转移矩阵, 所以 \mathbf{P} 的周期为 2. 又因为 \mathbf{A} 与 \mathbf{B} 的不变分布分别为 $(\frac{9}{22}, \frac{13}{22})$, $(\frac{4}{11}, \frac{7}{11})$. 由定理 1.8.3,

$$\lim_{n \rightarrow \infty} \mathbf{P}^{2n} = \lim_{n \rightarrow \infty} \begin{pmatrix} \mathbf{A}^n & 0 \\ 0 & \mathbf{B}^n \end{pmatrix} = \begin{pmatrix} \frac{9}{22} & \frac{13}{22} & 0 & 0 \\ \frac{9}{22} & \frac{13}{22} & 0 & 0 \\ 0 & 0 & \frac{4}{11} & \frac{7}{11} \\ 0 & 0 & \frac{4}{11} & \frac{7}{11} \end{pmatrix}.$$

不难验证 $(\frac{9}{44}, \frac{13}{44}, \frac{4}{22}, \frac{7}{22})$ 为 \mathbf{P} 的不变分布.

补充知识.

1. 定理 1.8.6 的证明.

取定状态 o . 令

$$d := \min\{m - n : m, n \in \mathbb{N} \text{ 且 } m > n\}, \quad \text{其中 } \mathbb{N} := \{n \geq 0 : p_{oo}^n > 0\}.$$

假设最小值在 n_0 处达到, 即, 假设

$$p_{oo}^{(n_0)}, p_{oo}^{(n_0+d)} > 0. \quad (1.8.3)$$

对 $r \in \{0, 1, \dots, d-1\}$, 令

$$D_r := \{i \in S : \exists n \geq 0 \text{ 使得 } p_{oi}^{(nd+r)} > 0\}.$$

那么, 由 S 不可约知

$$S = D_0 \cup \dots \cup D_{d-1}.$$

首先, 我们证明: 事实上, D_0, \dots, D_{d-1} 互不相交, 从而它们是 S 的一个分割. 也就是说, 如果 $r, s \in \{0, 1, \dots, d-1\}$ 且 $r \neq s$, 那么 $D_r \cap D_s = \emptyset$. 反证法: 若不然, 存在 $i \in D_r \cap D_s$. 根据 D_r, D_s 的定义, 存在 $n, \hat{n} \geq 0$ 使得

$$p_{oi}^{(nd+r)}, p_{oi}^{(\hat{n}d+s)} > 0. \quad (1.8.4)$$

不妨假设 $n \geq \hat{n}$, 否则交换 r 与 s . 因为 S 互通, 所以存在 m 使得

$$p_{io}^{(m)} > 0. \quad (1.8.5)$$

于是, 如果令

$$\begin{aligned} n_1 &= nd + r + m + n_0(n - \hat{n}), \\ n_2 &= \hat{n}d + s + m + (n_0 + d)(n - \hat{n}), \end{aligned}$$

那么, 结合 (1.8.3), (1.8.4) 与 (1.8.5), 我们推出

$$p_{oo}^{(n_1)} \geq p_{oi}^{(nd+r)} p_{io}^{(m)} \left(p_{oo}^{(n_0)} \right)^{n-\hat{n}} > 0.$$

即 $n_1 \in \mathbf{N}$. 同理, $n_2 \in \mathbf{N}$. 然而,

$$n_1 - n_2 = nd + r + m + n_0(n - \hat{n}) - \hat{n}d - s - d(n - \hat{n}) = r - s.$$

这表明 $\tilde{m} := \max\{n_1, n_2\}$ 与 $\tilde{n} := \min\{n_1, n_2\}$ 都属于 \mathbf{N} , 并且 $\tilde{m} > \tilde{n}$, 但 $\tilde{m} - \tilde{n} = |r - s| < d$, 这与 d 的定义矛盾. 因此, 假设不成立. 即, D_0, \dots, D_{d-1} 互不相交.

下面, 我们证明(i). 等价地, 我们需要验证: 如果 $i \in D_r$, 且 $p_{ij}^{(s)} > 0$, 那么 $j \in D_{r+s}$. (此即 $\sum_{j \in S} p_{ij}^{(s)} = \sum_{j \in D_{r+s}} p_{ij}^{(s)}$, 即(i).) 根据 D_r 的定义以及 $i \in D_r$, 可以取 $n \geq 0$ 使得 $p_{oi}^{(nd+r)} > 0$. 于是,

$$p_{oj}^{(nd+r+s)} \geq p_{oi}^{(nd+r)} p_{ij}^{(s)} > 0.$$

这表明 $j \in D_{r+s}$.

最后, 我们证明(ii). 一方面, 结合(i) 与 (1.8.3), 我们有 $d|n_0$ (即, n_0 是 d 的整数倍). 当 $nd \geq n_0^2$ 时, 存在 $q \geq n_0$, $r \in \{0, 1, \dots, n_0 - 1\}$ 使得 $nd = qn_0 + r$. 这表明, $r = nd - qn_0$ 是 d 的整数倍, 将其写成 md , 其中 $m \geq 0$. 于是,

$$nd = qn_0 + r = qn_0 + md = (q - m)n_0 + m(n_0 + d).$$

由 $m \leq r < n_0 \leq q$ 知

$$p_{oo}^{(nd)} \geq \left(p_{oo}^{(n_0)}\right)^{q-m} \left(p_{oo}^{(n_0+d)}\right)^m > 0.$$

另一方面, 对任意 $r \geq 0, i, j \in D_r$, 由 S 不可约, 可以取 a, b 使得 $p_{io}^{(a)}, p_{oj}^{(b)} > 0$.

$$p_{ij}^{(a+b)} \geq p_{io}^{(a)} p_{oj}^{(b)} > 0,$$

这结合(i) 表明存在非负整数 k 使得 $a + b = kd$. 综合以上两个方面, 取 $N = k + \frac{n_0^2}{d}$, 那么, 当 $n \geq N$ 时, $nd - (a + b) = (n - k)d \geq n_0^2$, 于是,

$$p_{ij}^{(nd)} \geq p_{io}^{(a)} p_{oo}^{(n-k)d} p_{oj}^{(b)} > 0.$$

因此, (ii) 成立. \square

2. 零常返情形.

命题 1.8.9. 若 \mathbf{P} 不可约、零常返, 则对任何 $i, j \in S$, $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

证: 假定该马氏链是非周期的, 否则可以考虑 \mathbf{P}^d . 考虑定理 1.8.3 证明中的马氏链 $\{(W_n, Y_n)\}$, 其状态空间 \tilde{S} 与转移概率 \mathbf{R} 如下:

$$\tilde{S} = S \times S, \quad r_{(i,j)(k,\ell)} = p_{ij} p_{k\ell}.$$

由定理 1.8.3 证明的第一步, \mathbf{R} 不可约. 下面我们分 \mathbf{R} 非常返、常返两种情形讨论.

情形 I. 假设 \mathbf{R} 非常返. 由推论 1.5.16 与 (1.8.1) 知, 当 $n \rightarrow \infty$ 时,

$$\left(p_{ij}^{(n)}\right)^2 = r_{(i,i)(j,j)}^{(n)} \rightarrow 0.$$

从而 $p_{ij}^{(n)} \rightarrow 0$.

情形 II. 假设 \mathbf{R} 常返. 则它必然碰到对角线. 与定理 1.8.3 证明的第六步类似, 我们此时仍然有

$$\lim_{n \rightarrow \infty} |P(W_n = j) - P(Y_n = j)| = 0. \quad (1.8.6)$$

特别地, 我们取如下的初始分布:

$$P((W_0, Y_0) = (k, j)) = \mathbf{1}_{\{k=i\}} \cdot p_{ij}^{(m)}, \quad \forall (k, j) \in \tilde{S}.$$

即,

$$W_0 = i, \quad P(Y_0 = j) = p_{ij}^{(m)}, \quad \forall j \in S.$$

由命题 1.1.12 与定理 1.8.3 证明的第一步, 我们有

$$P(W_n = j) = p_{ij}^{(n)}, \quad P(Y_n = j) = p_{ij}^{(n+m)}, \quad \forall j \in S.$$

这结合 (1.8.6) 表明

$$\lim_{n \rightarrow \infty} |p_{ij}^{(n)} - p_{ij}^{(n+m)}| = 0, \quad \forall m \geq 1. \quad (1.8.7)$$

为了证明 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$, 即, 对任意 $\varepsilon > 0$, 存在 $N \geq 0$ 使得 $p_{ij}^{(n)} < \varepsilon, \forall n \geq N$. 由 (1.8.7), 我们只需证明:

$$\exists M \geq 1 \text{ 使得 } p_{ij}^{(n+m)}, 0 \leq m \leq M-1 \text{ 中至少有一个不超过 } \frac{\varepsilon}{2}, \forall n \geq 0. \quad (1.8.8)$$

这是因为: 如果上式成立, 那么, 由 (1.8.7) 知存在 N 使得 $|p_{ij}^{(n)} - p_{ij}^{(n+m)}| < \frac{\varepsilon}{2}, 0 \leq m \leq M-1, \forall n \geq N$. 由 (1.8.8) 知存在 m 使得 $p_{ij}^{(n+m)} \leq \frac{\varepsilon}{2}$, 于是

$$p_{ij}^{(n)} \leq p_{ij}^{(n+m)} + |p_{ij}^{(n)} - p_{ij}^{(n+m)}| < \varepsilon, \quad \forall n \geq N.$$

最后, 我们证明 (1.8.8). 考虑 S 上以 \mathbf{P} 为转移矩阵的马氏链 $\{X_n\}$. 考察事件 $A = “\{X_n\}$ 在时间区间 $[n, n+M-1]$ 上访问过状态 $j”$. 即,

$$A = \bigcup_{m=0}^{M-1} \{X_{n+m} = j\}.$$

当 A 发生时, 将 $\{X_n\}$ 最后一次访问状态 j 的时刻记为 κ , 那么,

$$\{\kappa = n+m\} = \{X_{n+m} = j, X_{n+r} \neq j, \forall m+1 \leq r \leq M-1\}.$$

将时刻 $n+m$ 视为现在, 则后续的轨道从 j 出发, 它经过了 $M-1-n$ 的时间还没有回到 j . 由马氏性

$$\begin{aligned} P_i(\kappa = n+m) &= P_i(X_{n+m} = j) P_i(X_{n+m+r} \neq j, \forall 1 \leq r \leq M-1-m | X_{n+m} = j) \\ &= p_{ij}^{(n+m)} P_j(\sigma_j > M-1-m). \end{aligned}$$

由于上面的事件互不相交, 事实上, 它们是事件 A 的划分, 因此

$$\sum_{m=0}^{M-1} p_{ij}^{(n+m)} P_j(\sigma_j > M-1-m) = \sum_{m=0}^{M-1} P_i(\tau = n+m) = P_i(A) \leq 1.$$

下面我们用反证法进行证明. 假设 (1.8.8) 不成立. 那么, 对任意 M , 存在 n 使得

$$p_{ij}^{(n+m)} > \frac{\varepsilon}{2}, \quad 0 \leq m \leq M-1.$$

于是,

$$\frac{\varepsilon}{2} \cdot \sum_{\ell=0}^{M-1} P_j(\sigma_j > \ell) = \frac{\varepsilon}{2} \cdot \sum_{m=0}^{M-1} P_j(\sigma_j > M-1-m) \leq \sum_{m=0}^{M-1} p_{ij}^{(n+m)} P_j(\sigma_j > M-1-m) \leq 1.$$

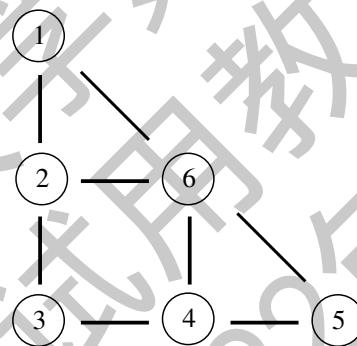
然而, 令 $M \rightarrow \infty$, 我们有

$$\sum_{\ell=0}^{M-1} P_j(\sigma_j > \ell) \rightarrow \sum_{\ell=0}^{\infty} P_j(\sigma_j > \ell) = E_j \sigma_j = \infty,$$

其中, 最后一个等号是因为 j 是零常返的. 上面的两个式子是相互矛盾的, 从而假设不成立, 即 (1.8.8) 成立. \square

习题

- 设有 6 个车站, 它们之间有公路连接, 连接情况如图 1 所示. 如果汽车每天可以从一个站驶到与之直接有公路相连接的相邻车站, 在夜间到达车站留宿并接受加油、清洗、检修等服务, 次日清晨各站按相同比例将各汽车派往其相邻车站. (1) 试说明: 在运行了很多日子以后, 各站每晚留宿的汽车比例趋于稳定, (2) 求出这些稳定值, 以便正确地设置各站的服务规模.



- (1) 求蜂窝图和三角形图(见图 1.4)上的简单随机游动的周期; (2) 对任意连通图, 试讨论其上的简单随机游动的周期.
 - 在 Ehrenfest 模型(例 1.1.8 与 1.7.12)中, 设 $N = 8, X_0 = 0$, 描述 n 很大时 X_n 的分布, (按照 n 的奇偶分别讨论).
 - 假设 \mathbf{P} 不可约、非周期. 证明: 定理 1.8.3 的证明中定义的转移矩阵 \mathbf{R} 也是非周期的.
 - 假设 \mathbf{P} 不可约、正常返, 周期 $d \geq 2$. 若 $i \in D_r, j \in D_{r+s}$, 则 $\lim_{n \rightarrow \infty} p_{ij}^{(nd+s)} = d\pi_j$, 其中 π 为 \mathbf{P} 的不变分布.
 - 证明公式(1.8.1)成立.
- 7* 证明: 定理 1.8.3 的证明中定义的 $\{W_n\}$ 与 $\{Y_n\}$ 都是 S 上以 \mathbf{P} 为转移矩阵的马氏链.

§1.9 收敛速度

将 S 上的所有分布组成的集合记为 \mathcal{M} . 对任意 $\mu, \nu \in \mathcal{M}$, 令

$$d_{TV}(\mu, \nu) := \frac{1}{2} \sum_{i \in S} |\mu_i - \nu_i|,$$

并称之为 μ 与 ν 之间的全变差距离.

注1.9.1. 引理 1.7.8 给出一列分布 $\pi^{(1)}, \pi^{(2)}, \dots$ 在全变差距离下收敛于极限分布 π 的充分必要条件.

假设 \mathbf{P} 不可约、非周期、正常返. 由强遍历定理, $\lim_{n \rightarrow \infty} d_{TV}(\mu \mathbf{P}^n, \pi) = 0$, 其中 π 为 \mathbf{P} 的不变分布. 本节讨论 $d_{TV}(\mu \mathbf{P}^n, \pi)$ 趋于 0 的速度. 不难验证

$$d_{TV}(\mu \mathbf{P}, \nu \mathbf{P}) \leq d_{TV}(\mu, \nu). \quad (1.9.1)$$

上式中的等号有可能成立. 例如 \mathbf{P} 的周期 $d \geq 2$, μ 和 ν 分别集中在 D_r 与 D_s 上, 其中 $0 \leq r < s < d$, 而 D_r 由定理 1.8.6 中定义, 此时, (1.9.1) 的左右两边总为 1.

命题1.9.2. (Dobrushin准则) 假设 S 有限, 转移矩阵 $\mathbf{P} = (p_{ij})$ 满足 $p_{ij} > 0, \forall i, j \in S$. 则存在常数 $0 \leq \alpha < 1$, 使得

$$d_{TV}(\mu \mathbf{P}, \nu \mathbf{P}) \leq \alpha d_{TV}(\mu, \nu), \quad \mu, \nu \in \mathcal{M}.$$

特别地, 若 π 是 \mathbf{P} 的不变分布, 则 $d_{TV}(\mu \mathbf{P}^n, \pi) \leq \alpha^n, \forall n \geq 0$.

证: 记 $\delta := \min_{i,j \in S} p_{ij}$. 那么 $\delta > 0$, 并且 $N\delta \leq \sum_{j \in S} p_{ij} = 1$, 其中 $N = |S|$. 因为

$$\sum_{i \in S} \mu_i p_{ij} = \delta + \sum_{i \in S} \mu_i (p_{ij} - \delta),$$

所以

$$\begin{aligned} d_{TV}(\mu \mathbf{P}, \nu \mathbf{P}) &= \frac{1}{2} \sum_{j \in S} \left| \sum_{i \in S} \mu_i p_{ij} - \sum_{i \in S} \nu_i p_{ij} \right| \\ &= \frac{1}{2} \sum_{j \in S} \left| \sum_{i \in S} \mu_i (p_{ij} - \delta) - \sum_{i \in S} \nu_i (p_{ij} - \delta) \right| \\ &\leq \frac{1}{2} \sum_{j \in S} \sum_{i \in S} |\mu_i - \nu_i| (p_{ij} - \delta) = \frac{1}{2} \sum_{i \in S} |\mu_i - \nu_i| \sum_{j \in S} (p_{ij} - \delta). \end{aligned}$$

取

$$\alpha := \sum_{j \in S} (p_{ij} - \delta) = 1 - N\delta < 1,$$

即得第一个不等式. 进一步, 假设 π 是不变分布, 那么 $\pi = \pi\mathbf{P}^n$. 取 $\nu = \pi$, 将第一个不等式迭代 n 次便知

$$d_{TV}(\mu\mathbf{P}^n, \pi) = d_{TV}(\mu\mathbf{P}^n, \pi\mathbf{P}^n) \leq \alpha d_{TV}(\mu\mathbf{P}^{n-1}, \pi\mathbf{P}^{n-1}) \leq \cdots \leq \alpha^n d_{TV}(\mu, \pi).$$

这结合 $d_{TV}(\mu, \pi) \leq 1$ 便推出第二个不等式. \square

在上面的命题中, 转移矩阵里的元素都为正. 这样的条件往往难以满足. 一种放宽条件的做法是先用 \mathbf{P}^m 代替 \mathbf{P} .

命题1.9.3. 假设 S 有限, \mathbf{P} 不可约、非周期, π 为 \mathbf{P} 的不变分布, 则存在常数 $C > 0$, $\beta > 0$, 使得对任意 $\mu \in \mathcal{M}$,

$$d_{TV}(\mu\mathbf{P}^n, \pi) \leq C e^{-\beta n}, \quad \forall n \geq 0.$$

证: 因为 \mathbf{P} 不可约、非周期, 所以对任意 $i, j \in S$, 存在 N_{ij} , 当 $n \geq N_{ij}$ 时, $p_{ij}^{(n)} > 0$. 取 $m = \max_{i,j \in S} N_{ij}$. 那么, 对所有的 $i, j \in S$, 都有 $p_{ij}^{(m)} > 0$. 将任意正整数写为 $rm + s$ 的形式, 其中 $0 \leq s \leq m - 1$. 由 Dobrusin 准则和 (1.9.1), 存在 $0 \leq \alpha < 1$ 使得

$$d_{TV}(\mu\mathbf{P}^{rm+s}, \pi) = d_{TV}((\mu\mathbf{P}^{rm})\mathbf{P}^s, \pi\mathbf{P}^s) \leq d_{TV}(\mu(\mathbf{P}^m)^r, \pi) \leq \alpha^r.$$

取 $C = \frac{1}{\alpha}$, $\beta = \frac{1}{m} \ln \frac{1}{\alpha}$ 便知命题成立. \square

命题 1.9.3 表明: 从任意初分布出发, X_n 的分布都将收敛到不变分布, 而且收敛速度是指数阶的. 这几乎是人们能指望的最好结论. 在很长一段时间, 人们热衷于讨论指数阶的大小, 即 β 的大小, 尤其是 β 随着状态数目 $|S|$ 的增大而下降的规律, 特别关心 β 是否下降到 0. 然而在上个世纪八十年代, Persi Diaconis 发现 $d_{TV}(\pi, \nu\mathbf{P}^k)$ 的典型曲线不是一条指数曲线, 而会出现门限现象(threshold). 比如说, 通常的洗牌模型, 一副牌有 54 张, 摆成一叠可设想为 1 至 54 的一个排列, 所以状态空间为 1 至 54 的所有排列, $|S| = 54!$. 洗牌就是从一个排列变为另一个排列, 不变分布就是所有 1 至 54 的排列上的均匀分布. 在此情形, 人们发现只要洗 7 次效果就很好了. 再多洗效果改进不大. 这在实践中很有价值, 尤其在随机算法的收敛方面.

习题

1. 证明: d_{TV} 是 \mathcal{M} 上的距离, 即满足: 非负性 $d_{TV}(\mu, \nu) \geq 0$, 等号成立当且仅当 $\mu = \nu$; 对称性 $d_{TV}(\mu, \nu) = d_{TV}(\nu, \mu)$; 三角不等式 $d_{TV}(\mu, \nu) \leq d_{TV}(\mu, \pi) + d_{TV}(\pi, \nu)$.
2. 证明下列对偶表达式成立, 并给出达到极值的子集 A 和函数 f .

$$d_{TV}(\mu, \nu) = \sup_{A \subseteq S} |\mu(A) - \nu(A)| = \sup_{0 \leq f_i \leq 1, i \in S} \left| \sum_{i \in S} \mu_i f_i - \sum_{i \in S} \nu_i f_i \right|.$$

3. 证明不等式(1.9.1) 成立.
4. 假设存在 $\mu \in \mathcal{M}$ 使得 $\lim_{n \rightarrow \infty} d_{TV}(\mu \mathbf{P}^n, \pi) = 0$. 证明: π 是不变分布.

北京大学数学教材
试用教材
2022年3月

§1.10 一维简单随机游动

本节考察一维简单随机游动. 在这个特定的模型中, 我们将得到一系列有趣的特殊结论. 假设 ξ_1, ξ_2, \dots 独立同分布,

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

$$S_0 = 0, \quad S_n = \xi_1 + \dots + \xi_n, \quad \forall n \geq 1.$$

一、首达时.

我们首先考虑粒子首次到达 i 的时刻, 即

$$\tau_i := \inf\{n \geq 0 : S_n = i\}.$$

根据对称性, 我们只用研究 $i \geq 1$ 的情形.

命题1.10.1 (反射原理). 固定 $n \geq 1$. 对任意 $1 \leq i \leq n$, 我们有

$$P_0(\tau_i < n, S_n = i + j) = P_0(\tau_i < n, S_n = i - j), \quad \forall j \geq 1. \quad (1.10.1)$$

证: 考虑前 n 步轨道, 并研究事件 $\{\tau_i < n\}$. 如果该事件发生, 那么轨道必然在 n 之前的某个时刻 m (即, $m \leq n - 1$) 已经(首次)到达过 i . 接下来, 我们可以将时刻 m 后的轨道关于水平高度 i 做反射. 如图 1.17 所示, 全为实线的轨道在反射后变为时刻 m 之前为实线, 而时刻 m 到时刻 n 之间为虚线的轨道. 考虑两组 n 步轨道, 一组满足 $\{\tau_i \leq n, S_n = i + j\}$, 而另一

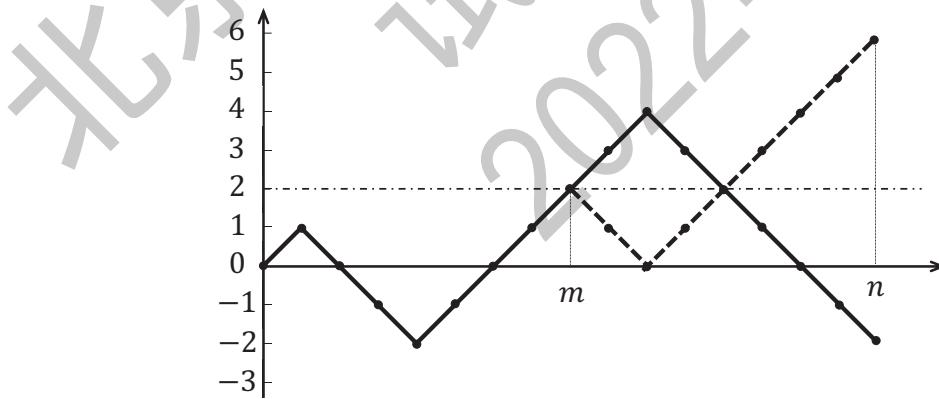


图 1.17 反射原理

组满足 $\{\tau_i \leq n, S_n = i - j\}$. 那么, 上述反射带来了这两组轨道之间的一一对应关系. 由于每条轨道出现的概率都是 $\frac{1}{2^n}$, 因此两个事件中含有的轨道数相等就意味着这两个事件的概率相等. 于是, 我们得到 (1.10.1). \square

命题1.10.2. 假设 $i, n \geq 1$. 那么,

$$P_0(\tau_i = n) = \frac{i}{n} P_0(S_n = i).$$

证: 当 $n + i$ 为奇数时, 上式左右两边都为0, 从而成立. 下面假设 $n + i$ 为偶数. 若 $\tau_i = n$, 则 $S_n = i$, 即

$$\{\tau_i = n\} \subseteq \{S_n = i\}. \quad (1.10.2)$$

我们考虑所有从0 到 i 的步长为 n 的轨道, 即所有满足 $S_n = i$ 的 n 步轨道. 根据反射原理(命题 1.10.1),

$$P(\tau_i < n, S_{n-1} = i+1) = P(\tau_i < n, S_{n-1} = i-1).$$

将 $n-1$ 视为现在, 那么 $\{\tau_i < n\} = \cup_{m=0}^{n-1} \{S_m = i\}$ 是过去与现在的事件. 根据马氏性, 左右两边乘以 $\frac{1}{2}$, 即分别乘以 $P(S_n = i | S_{n-1} = i+1)$ 与 $P(S_n = i | S_{n-1} = i-1)$, 我们有

$$P(\tau_i < n, S_{n-1} = i+1, S_n = i) = P(\tau_i < n, S_{n-1} = i-1, S_n = i).$$

由于上式左、右两边的事件形成 $\{\tau_i < n, S_n = i\}$ 的划分, 因此

$$P(S_n = i, \tau_i < n) = 2P(\tau_i < n, S_{n-1} = i+1, S_n = i).$$

进一步, 因为粒子每次的跳跃幅度都是1, 所以, $S_{n-1} = i+1$ 表明粒子在时间区间 $[0, n-1]$ 上必然穿过 i , 即, $\tau_i < n$. 也就是说,

$$\{S_{n-1} = i+1\} \subseteq \{\tau_i < n\}.$$

从而,

$$P(S_n = i, \tau_i < n) = 2P(S_{n-1} = i+1, S_n = i) = P(S_{n-1} = i+1).$$

这结合 (1.10.2) 表明

$$P_0(\tau_i = n) = P_0(S_n = i) - P_0(S_n = i, \tau_i < n) = P_0(S_n = i) - P_0(S_{n-1} = i+1).$$

由 $n+i$ 为偶数, 不妨将其记为 $2k$, 我们推出

$$P_0(S_{n-1} = i+1) = \frac{(n-1)!}{k!(n-1-k)!} \cdot \frac{1}{2^{n-1}} = \frac{2(n-k)}{n} \times \frac{n!}{k!(n-k)!} \cdot \frac{1}{2^n} = \frac{n-i}{n} P_0(S_n = i),$$

其中, 在最后一个等式中, 我们用到了 $2k = n+i$, 即 $2(n-k) = n-i$. 结合上面两个式子, 我们便可以推出结论. \square

推论1.10.3. $P_0(\tau_1 > 2n-1) = P_0(S_{2n} = 0)$.

证: 由反射原理,

$$P_0(\tau_1 < 2n, S_{2n} = 1 + j) = P_0(\tau_1 < 2n, S_{2n} = 1 - j), \quad \forall j \geq 1.$$

左、右两边对 j 求和知

$$P_0(\tau_1 < 2n, S_{2n} \geq 2) = P_0(\tau_1 < 2n, S_{2n} \leq 0).$$

由于左、右两边的事件是 $\{\tau_1 < 2n\}$ 的划分, 且 $S_{2n} \geq 2$ 蕴含着 $\tau_1 < 2n$, 因此

$$P_0(\tau_1 < 2n) = 2P_0(\tau_1 < 2n, S_{2n} \geq 2) = 2P_0(S_{2n} \geq 2).$$

一方面, 根据对称性,

$$P_0(S_{2n} \geq 2) = P_0(S_{2n} \leq -2).$$

另一方面, S_{2n} 是偶数. 因此,

$$2P_0(S_{2n} \geq 2) = P_0(S_{2n} \geq 2) + P_0(S_{2n} \leq -2) = P_0(S_{2n} \neq 0).$$

又 $\{\tau_1 > 2n - 1\}^c = \{\tau_1 \leq 2n - 1\} = \{\tau_1 < 2n\}$. 综上,

$$P_0(\tau_1 > 2n - 1) = 1 - P_0(\tau_1 < 2n) = 1 - P_0(S_{2n} \neq 0) = P_0(S_{2n} = 0).$$

因此, 结论成立. □

推论1.10.4. $P_0(\tau_1 < \infty) = 1, E_0 \tau_1 = \infty$.

证: 根据推论1.10.3, $P_0(\tau_1 \geq n)$ 与 $1/\sqrt{n}$ 为等价无穷小. 从而

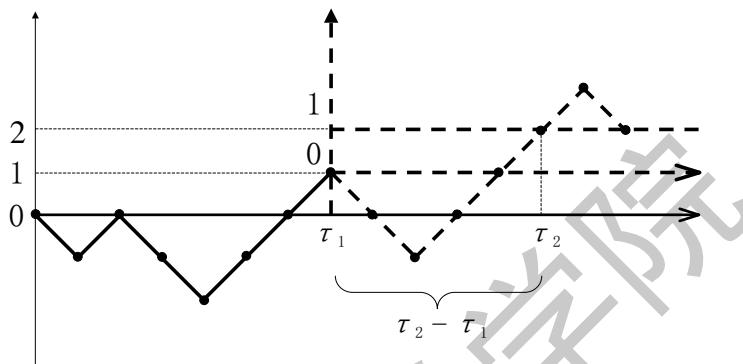
$$P_0(\tau_1 = \infty) = \lim_{n \rightarrow \infty} P_0(\tau_1 \geq n) = 0,$$

并且, $E_0 \tau_1 = \sum_{n=0}^{\infty} P_0(\tau_1 > n) = \infty$. □

注1.10.5. 上述推论的概率意义是很清楚的. 从原点出发的一维简单随机游动迟早要到达1, 但平均所用的时间为无穷大. 进一步, 先让随机游动走一步, 我们有

$$P_0(\sigma_0 > 2n) = P_1(\sigma_0 > 2n - 1) = P_0(\sigma_1 > 2n - 1).$$

即, 对于一个从0 出发的简单随机游动, 从分布的角度而言, $\sigma_0 \stackrel{d}{=} 1 + \tau_1$. 事实上, $\sigma_0 = 1 + \tilde{\tau}_1$, 其中 $\tilde{\tau}_1$ 是 $X_n = S_{1+n}, n \geq 0$ 这个时刻1 的后续轨道首达 $-S_1$ 的时间. 因此, 上述推论还蕴含着简单随机游动是零常返的.

图 1.18 $\tau_2 - \tau_1$ 与 τ_1 独立同分布

下面我们研究 τ_i 的性质. 为此我们先观察 τ_2 与 τ_1 的关系. 由下图, 当粒子首次到达 1 后, 后续轨道为从 1 出发的简单随机游动. 根据空间的平移不变性, 它还需要经过一个同分布于 τ_1 的随机时间到达 2, 并且, 由马氏性, 这个随机时间与 τ_1 相互独立.

于是, 类似于命题 1.5.1, 我们可以证明如下的命题.

命题 1.10.6. 假设 $S_0 = 0$, 那么 $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$ 独立同分布.

该命题提供了证明 $E_0\tau_1 = \infty$ 的另一个方法: 当 $S_1 = 1$ 时, $\tau_1 = 1$; 当 $S_1 = -1$ 时, $\tau_1 = 1 + \tilde{\tau}_2$, 其中 $\tilde{\tau}_2$ 是一个从 -1 出发的随机游动首达 1 的时间, 它与从 0 出发首达 2 的时间 τ_2 同分布. 从而

$$E_0\tau_1 = \frac{1}{2} \times 1 + \frac{1}{2} \times (1 + E_0\tau_2) = 1 + \frac{1}{2}E_0\tau_2 = 1 + E_0\tau_1.$$

这表明 $E_0\tau_1 = \infty$.

二. 反正弦律

首达时考虑的是 $\{S_n\}$ 首次达到某状态的时刻, 一种特殊情形等价于第一次返回原点的时刻 σ_0 (见注 1.10.5). 下面的反正弦律考虑的则是(在 n 之前)最后一次返回原点的时刻. 具体地, 将时刻 n 之前最后一次碰到 0 的时刻记为 ϵ_n , 即

$$\epsilon_n := \max\{m \leq n : S_m = 0, S_{m+1}, \dots, S_n \neq 0\}.$$

命题 1.10.7 (反正弦律). 假设 $0 < \delta < 1$. 那么,

$$\lim_{n \rightarrow \infty} P_0(\epsilon_n \leq \delta n) = \frac{2}{\pi} \arcsin \sqrt{\delta}.$$

证: 记

$$p_n := P_0(\epsilon_n < \delta n) = \sum_{m=0}^{[\delta n/2]} P_0(\epsilon_n = 2m). \quad (1.10.3)$$

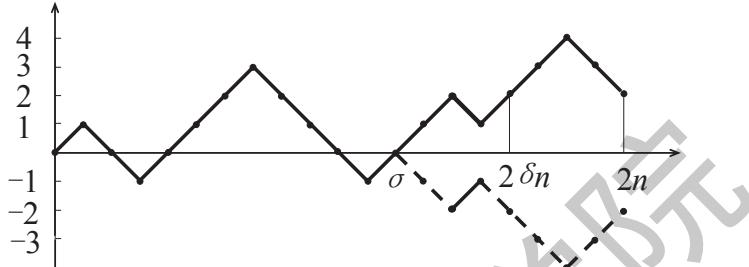


图 1.19 反正弦律

给定 $m \geq 0$. 由于 $Y_r := S_{2m+r} - S_{2m}$, $r \geq 0$ 是简单随机游动, 且与 S_{2m} 相互独立, 因此根据 ϵ_n 的定义可推出

$$P_0(\epsilon_n = 2m) = P_0(S_{2m} = 0; Y_r \neq 0, r = 1, \dots, n-2m) = P_0(S_{2m} = 0)P_0(\sigma_0 > n-2m). \quad (1.10.4)$$

下面, 我们计算 $P_0(\sigma_0 > n-2m)$. 根据首步分析法, 我们有

$$P_0(\sigma_0 > n-2m) = \frac{1}{2}P_1(\sigma_0 > n-2m-1) + \frac{1}{2}P_{-1}(\sigma_0 > n-2m-1).$$

根据空间的对称性与平移不变性,

$$P_1(\sigma_0 > n-2m-1) = P_{-1}(\sigma_0 > n-2m-1) = P_0(\tau_1 > n-2m-1).$$

综上, 由命题 1.10.3 可以推出

$$P_0(\sigma_0 > n-2m) = P_0(\tau_1 > n-2m-1) = P_0(S_{2n'-2m} = 0), \quad (1.10.5)$$

其中,

$$n' = \begin{cases} \frac{1}{2}n, & \text{若 } n \text{ 为偶数,} \\ \frac{1}{2}(n-1), & \text{若 } n \text{ 为奇数.} \end{cases}$$

将其代入 (1.10.4), 再结合 (1.10.3), 我们推出

$$p_n = \sum_{m=0}^{[\delta n/2]} P_0(S_{2m}=0)P_0(S_{2(n'-m)}=0).$$

由 Stirling 公式 (1.5.13), 对任意 $\varepsilon > 0$, 存在 n_0 使得

$$(1-\varepsilon)\frac{1}{\sqrt{\pi\ell}} \leq P_0(S_{2\ell}=0) \leq (1+\varepsilon)\frac{1}{\sqrt{\pi\ell}}, \quad \forall \ell \geq n_0.$$

于是, 一方面, 当 $n' \geq \max\{2n_0, \frac{1}{1-\delta}n_0\}$ 时,

$$\begin{aligned} p_n &\leq \sum_{m=0}^{n_0-1} P_0(S_{2(n'-m)} = 0) + \sum_{m=n_0}^{[\delta n/2]} P_0(S_{2m=0})P_0(S_{2(n'-m)} = 0) \\ &\leq (1+\varepsilon) \sum_{m=0}^{n_0-1} \frac{1}{\sqrt{\pi(n'-m)}} + (1+\varepsilon)^2 \sum_{m=n_0}^{[\delta n/2]} \frac{1}{\pi\sqrt{m(n'-m)}}. \end{aligned}$$

令 $n \rightarrow \infty$, 关于上式右边第二项, 我们得到

$$\begin{aligned} \sum_{m=n_0}^{[\delta n/2]} \frac{1}{\pi\sqrt{m(n'-m)}} &\leq \sum_{m=1}^{[\delta n/2]} \frac{1}{\pi\sqrt{m(n'-m)}} = \frac{1}{\pi} \sum_{m=0}^{[\delta n/2]} \frac{1}{\sqrt{\frac{m}{n'}(1-\frac{m}{n'})}} \cdot \frac{1}{n'} \\ &\rightarrow \frac{1}{\pi} \int_0^\delta \frac{1}{\sqrt{x(1-x)}} dx = \frac{2}{\pi} \int_0^{\sqrt{\delta}} \frac{1}{\sqrt{1-y^2}} dy = \frac{2}{\pi} \arcsin \sqrt{\delta}. \end{aligned}$$

最后, 令 $\varepsilon \rightarrow 0$ 便知

$$\limsup_{n \rightarrow \infty} p_n \leq \frac{2}{\pi} \arcsin \sqrt{\delta}.$$

另一方面,

$$\begin{aligned} p_n &\geq \sum_{m=n_0}^{[\delta n/2]} P_0(S_{2m=0})P_0(S_{2(n'-m)} = 0) \\ &\geq (1-\varepsilon)^2 \sum_{m=n_0}^{[\delta n/2]} \frac{1}{\pi\sqrt{m(n'-m)}} \geq (1-\varepsilon)^2 \sum_{m=1}^{[\delta n/2]} \frac{1}{\pi\sqrt{m(n'-m)}} - (1-\varepsilon)^2 \sum_{m=1}^{n_0-1} \frac{1}{\sqrt{n'-m}}. \end{aligned}$$

令 $n \rightarrow \infty$, 上式右边第二项趋于 0, 第一项趋于 $(1-\varepsilon)^2 \frac{2}{\pi} \arcsin \sqrt{\delta}$. 再令 $\varepsilon \rightarrow 0$, 便知

$$\liminf_{n \rightarrow \infty} p_n \geq \frac{2}{\pi} \arcsin \sqrt{\delta}.$$

结合以上两个方面, 可以推出结论成立. \square

注1.10.8. 反正弦律显示, 简单随机游动的游弋非常地长, 这是导致它零常返的本质原因. 在第三章, 我们可以看到, 作为简单随机游动的连续版本, 布朗运动也有一样的反正弦律, 见定理 3.4.16.

三、赌徒模型.

假设某赌徒最初有赌资 i 元, 其中 $i \geq 1$, 他/她参加某公平赌博(公平赌博指每一局独立地等可能赢或输), 每一局下注 1 元, 对手的赌资为无穷大. 那么, 该赌徒的赌资的变化就是一个简单随机游动. 假设不允许出现负资产, 也就是说, 当赌徒破产(即, 其赌资为 0) 时, 赌博结束. 那么, 根据推论 1.10.4 与命题 1.10.6, 无论该赌徒的原始赌资有多少, 他/她都以概率 1 破产, 只不过破产所需的平均时间为无穷.

现在, 我们修改一下模型的假设. 假设对手最初有 j 元赌资, 其中 $j \geq 1$. 该赌徒破产或完胜(即, 对手破产)时, 赌博结束. 此时, 赌博仍然概率1会结束. 我们关心下面的三个问题: (1) 该赌徒破产和完胜的概率分别是多少? (2) 在赌博结束时, 该赌徒平均能赢多少钱? (3) 这场赌博平均要进行多长时间?

(1) 该赌徒完胜和破产的概率分别是多少?

这就是所谓的赌徒破产问题. 赌徒的赌资是一个从 i 出发的简单随机游动, 将其记为 $\{S_n\}$. 那么, 赌博结束的时刻就是 $\{S_n\}$ 首次到达0或 $i+j$ 的时刻. 记

$$\tau := \min\{\tau_0, \tau_{i+j}\} \triangleq \tau_0 \wedge \tau_{i+j},$$

其中 τ_k 是 $\{S_n\}$ 首达 k 的时刻. 根据推论 1.10.4, 命题 1.10.6, 空间的平移不变性, 我们有

$$P_i(\tau < \infty) \geq P_i(\tau_{i+j} < \infty) = P_0(\tau_j < \infty) = 1.$$

即, 赌博以概率1会结束. 赌徒破产和完胜的事件分别为

$$\{\tau_0 < \tau_{i+j}\} = \{S_\tau = 0\}, \quad \{\tau_0 > \tau_{i+j}\} = \{S_\tau = i+j\}.$$

为计算它们的概率, 我们采取首步分析法. 现在我们考虑从 k 出发的简单随机游动首次到达边界0或 $i+j$ 时的位置 S_τ , 其中 $0 \leq k \leq i+j$. 若 $k \neq 0$ 或 $i+j$, 我们按照粒子一步后的位置 S_1 的值进行划分, 并运用全概公式, 得到

$$P_k(S_\tau = 0) = \frac{1}{2}P_{k-1}(S_\tau = 0) + \frac{1}{2}P_{k+1}(S_\tau = 0), \quad 1 \leq k \leq i+j-1.$$

这结合边界条件 $P_0(S_\tau = 0) = 1$ 和 $P_{i+j}(S_\tau = 0) = 0$ 表明

$$P_k(S_\tau = 0) = 1 - \frac{k}{i+j}.$$

特别地, 赌徒破产概率和完胜概率分别为

$$P_i(S_\tau = 0) = \frac{j}{i+j}, \quad P_i(S_\tau = i+j) = \frac{i}{i+j}. \quad (1.10.6)$$

(2) 赌博结束时, 该赌徒平均能赢多少钱?

赌博结束时, 该赌徒的平均赌资为

$$E_i S_\tau = 0 \times P_i(S_\tau = 0) + (i+j)P_i(S_\tau = i+j) = i. \quad (1.10.7)$$

由于该赌徒的初始赌资为 i 元, 因此上面的式子表明在赌博结束时, 该赌徒赢到的钱 $S_\tau - i$ 的平均值为0. 从这个角度看, 赌博是公平的.

(3) 这场赌博平均要进行多长时间?

我们首先证明这场赌博的平均时间是有限的.

引理1.10.9. $E_k \tau < \infty, \forall 0 \leq k \leq i+j$.

证: 记 $N = i + j$. 若从 k 出发, 向右连续走 N 步(该事件的概率为 2^{-N}), 则途中一定能遇到 N . 因此, $P_k(\tau \leq N) \geq 2^{-N}$, 即

$$P_k(\tau > N) \leq 1 - \frac{1}{2^N}, \quad \forall 0 \leq k \leq N. \quad (1.10.8)$$

下面我们用归纳法证明: 对任意 $n \geq 0$,

$$P_k(\tau > nN) \leq \left(1 - \frac{1}{2^N}\right)^n. \quad (1.10.9)$$

首先, 当 $n = 0, 1$ 时, (1.10.9) 成立. 其次, 假设 (1.10.9) 对 n 成立, 往证它对 $n+1$ 也成立. 如果 $\tau > (n+1)N$, 那么粒子在 nN 时刻还没有到达边界 $\{0, N\}$, 即 $1 \leq S_{nN} \leq N-1$. 于是, 我们按照 S_{nN} 的值进行划分便知

$$P_k(\tau > (n+1)N) = \sum_{r=0}^N P_k(\tau > (n+1)N, S_{nN} = r). \quad (1.10.10)$$

将 nN 时刻视为现在, 将 $\tilde{S}_m = S_{nN+m}$, $m \geq 0$ 视为后续轨道. 那么,

$$\{\tau > (n+1)N\} = \{\tau > nN\} \cap \{\tilde{\tau} > N\},$$

其中, $\{\tau > nN\}$ 是关于过去与现在的事件, 如图 1.20 所示, 实线部分一直不碰到边界 $\{0, N\}$; 而 $\tilde{\tau}$ 是 $\{\tilde{S}_m\}$ 首达边界 $\{0, N\}$ 的时间, 如图 1.20 所示, 虚线部分表示 $\{\tilde{S}_m\}$ 的轨道, $\{\tilde{\tau} > N\}$ 就是虚线的前 N 步一直不碰到边界. 综上,



图 1.20 $N = 5, n = k = r = 2$

$$\{\tau > (n+1)N, S_{nN} = r\} = \{\tau > nN, S_{nN} = r\} \cap \{\tilde{\tau} > N\},$$

其中, $\{S_{nN} = r, \tau > (n+1)N\}$ 是关于过去与现在的事件, 而 $\{\tilde{\tau} > N\}$ 是关于将来的事件. 由马氏性,

$$P_k(S_{nN} = r, \tau > (n+1)N) = P_k(\tau > nN, S_{nN} = r)P_r(\tau > N),$$

其中, 由 (1.10.8) 知

$$P_r(\tau > N) \geq 1 - \frac{1}{2^N}.$$

将上面的两个代入 (1.10.10), 并利用归纳假设 (1.10.9), 我们有

$$\begin{aligned} P_k(\tau > (n+1)N) &= \sum_{r=0}^N P_k(\tau > nN, S_{nN} = r) P_r(\tau > N) \\ &\leq \sum_{r=0}^N P_k(\tau > nN, S_{nN} = r) \left(1 - \frac{1}{2^N}\right) \\ &= P_k(\tau > nN) \left(1 - \frac{1}{2^N}\right) \leq \left(1 - \frac{1}{2^N}\right)^{n+1}. \end{aligned}$$

这即为 (1.10.9). 综上, (1.10.9) 对任意 $n \geq 0$ 都成立.

进一步,

$$\begin{aligned} E_k \tau &= \sum_{m=0}^{\infty} P_k(\tau > m) = \sum_{n=0}^{\infty} \sum_{s=0}^{N-1} P_k(\tau > nN + s) \\ &\leq \sum_{n=0}^{\infty} N P_k(\tau > nN) \leq N \sum_{n=0}^{\infty} \left(1 - \frac{1}{2^N}\right)^n < \infty, \end{aligned}$$

从而命题成立. \square

由首步分析法, 当 $1 \leq k \leq i+j-1$ 时,

$$E_k \tau = 1 + \frac{1}{2} E_{k+1} \tau + \frac{1}{2} E_{k-1} \tau.$$

结合引理 1.10.9 和边界条件 $E_0 \tau = E_{i+j} \tau = 0$, 我们可以推出

$$E_k \tau = k(i+j-k). \quad (1.10.11)$$

特别地, 取 $k = i$ 便得到赌博平均进行的时间.

定理1.10.10. 若 $0 \leq i \leq N$, 则

$$E_i \tau_0 \wedge \tau_N = i \times j.$$

根据 (1.10.6), S_τ 服从两点分布. 进一步, 我们可以计算其方差

$$\text{Var}_i(S_\tau) = E_i S_\tau^2 - (E_i S_\tau)^2 = 0^2 \times \frac{j}{i+j} + (i+j)^2 \times \frac{i}{i+j} - i^2 = i \times j.$$

由上面的结论, 我们发现 $\text{Var}_i(S_\tau) = E_i \tau$. 进一步, 根据随机游动 $\{S_n\}$ 的步长 ξ_n 的方差为 1, 我们推出

$$\text{Var}_i(S_\tau) = E_i \tau \times \text{Var}(\xi_1). \quad (1.10.12)$$

事实上, 公式 (1.10.7) 与 (1.10.12) 分别是Wald 等式与Wald 第二等式的特殊情形. 在补充知识中, 我们将介绍这两个等式.

补充知识.

假设 $\xi = \xi_1, \xi_2, \dots$ 是一列独立同分布的、非退化的随机变量. 令

$$S_0 = 0, \quad S_n = \xi_1 + \dots + \xi_n, \quad n \geq 1.$$

假设 τ 是 $\{S_n\}$ 的停时(参见定义 1.4.7).

定理1.10.11 (Wald 等式). 若 $E\tau < \infty, E|\xi| < \infty$, 则 $ES_\tau = E\tau \cdot E\xi$.

证: 我们需要先验证 S_τ 的期望存在, 即 $E|S_\tau| < \infty$. 由于

$$S_\tau = \sum_{n=1}^{\tau} \xi_n = \sum_{n=1}^{\infty} \xi_n \mathbf{1}_{\{\tau \geq n\}},$$

因此

$$|S_\tau| \leq |\xi_1| + \dots + |\xi_\tau| = \sum_{n=1}^{\infty} |\xi_n| \mathbf{1}_{\{\tau \geq n\}}.$$

因为 τ 是停时, 所以 $\{\tau \geq n\} = \{\tau \leq n-1\}^c$ 是否发生只依赖于 S_0, S_1, \dots, S_{n-1} 的值, 即, 只依赖于 ξ_1, \dots, ξ_{n-1} 的值. 从而 $\mathbf{1}_{\{\tau \geq n\}}$ 与 ξ_n 相互独立. 于是我们推出

$$E|S_\tau| \leq \sum_{n=1}^{\infty} E|\xi_n| \mathbf{1}_{\{\tau \geq n\}} = \sum_{n=1}^{\infty} E|\xi_n| \cdot P(\tau \geq n) = E\tau \cdot E|\xi| < \infty.$$

于是, 由Fubini定理(见附录中的定理 4.0.20), 我们可以交换求期望与求和的顺序, 便得到

$$ES_\tau = E \sum_{n=1}^{\infty} \xi_n \mathbf{1}_{\{\tau \geq n\}} = \sum_{n=1}^{\infty} E(\xi_n \mathbf{1}_{\{\tau \geq n\}}) = \sum_{n=1}^{\infty} E\xi_n \cdot P(\tau \geq n) = E\xi \cdot E\tau.$$

从而结论成立. \square

例1.10.12. 对于第 §1.10 节研究的赌徒模型. 假设赌徒与其对手的初始赌资分别为 i, j . 则赌徒的赌资扣除 i 就是从 0 出发的随机游动, 记为 $\{S_n\}$. 赌博结束的时间为 $\{S_n\}$ 首达 $-i$ 或 j 的时间, 记为 τ , 它是停时. 根据引理 1.10.9, $E\tau < \infty$. 由定理 1.10.11,

$$ES_\tau = E\xi \times E\tau = 0 \times E\tau = 0.$$

于是, 我们得到方程组

$$\begin{cases} P_0(S_\tau = -i) + P_0(S_\tau = j) = 1 \\ -iP_0(S_\tau = -i) + jP_0(S_\tau = j) = E_0 S_\tau = 0. \end{cases}$$

解得

$$P_0(S_\tau = -i) = \frac{j}{i+j}, \quad P_0(S_\tau = j) = \frac{i}{i+j}.$$

这给出求赌徒破产概率与完胜概率的另一种方法.

注1.10.13. 在定理 1.10.11 中, $E\tau < \infty$ 的假设不能去掉. 例如: 考虑从 0 出发的一维简单随机游动, 我们有 $E\xi_1 = 0$, 但此时不能认为所有停时都满足 $E_0 S_\tau = 0$. 比如说, 对于首达 1 的时刻 τ_1 , 总有 $S_{\tau_1} = 1$, 从而 $E_0 S_{\tau_1} = 1 \neq 0$.

定理1.10.14 (Wald 第二等式). 假设 $E\tau < \infty$, $E\xi = 0$, $E\xi^2 < \infty$, 且存在 $M \geq 0$ 使得 $|S_n| \leq M$, $\forall n \leq \tau$. 则 $ES_\tau^2 = E\tau \cdot E\xi^2$.

证: 不妨设 $E\xi^2 = 1$, 否则用 $\eta_n := \frac{\xi}{\sqrt{E\xi^2}}$ 与 $T_n := \frac{S_n}{\sqrt{E\xi^2}}$ 分别代替 ξ_n 与 S_n 即可. 对任意 $n \geq 1$, 考虑 $\tau \wedge n = \min\{\tau, n\}$. 它是停时, 且

$$S_{\tau \wedge n} = \sum_{i=1}^{\tau \wedge n} \xi_i = \sum_{i=1}^n \xi_i \mathbf{1}_{\{\tau \geq i\}}.$$

于是

$$ES_{\tau \wedge n}^2 = E \left(\sum_{i=1}^n \xi_i \mathbf{1}_{\{\tau \geq i\}} \right)^2 = E \sum_{i=1}^n \xi_i^2 \mathbf{1}_{\{\tau \geq i\}} + 2E \sum_{1 \leq i < j \leq n} \xi_i \xi_j \mathbf{1}_{\{\tau \geq j\}}.$$

由定理 1.10.11, 右边第一项就是 $E(\tau \wedge n) \cdot E\xi^2$. 关于第二项, 由于 $\{\tau \geq j\}$ 是否发生仅依赖于 ξ_1, \dots, ξ_{j-1} 的值, 因此 $\xi_i \mathbf{1}_{\{\tau \geq j\}}$ 与 ξ_j 相互独立, 从而 $E(\xi_i \xi_j \mathbf{1}_{\{\tau \geq j\}}) = E(\xi_i \mathbf{1}_{\{\tau \geq j\}}) \cdot E\xi_j = 0$. 于是,

$$ES_{\tau \wedge n}^2 = E(\tau \wedge n) \cdot E\xi^2.$$

下面, 令 $n \rightarrow \infty$. 一方面, 由有界收敛定理(见附录中的定理 4.0.19), 因为 $0 \leq S_{\tau \wedge n}^2 \leq M$ 且 $S_{\tau \wedge n} \rightarrow S_\tau$, 所以

$$\lim_{n \rightarrow \infty} ES_{\tau \wedge n}^2 = ES_\tau^2.$$

另一方面, 由单调收敛定理 $\lim_{n \rightarrow \infty} E(\tau \wedge n) = E\tau$. 从而结论成立. \square

注1.10.15. Wald 等式刻画了 ES_τ 的数学期望, Wald 第二等式则刻画了 S_τ 的方差. 它需要更多的假设条件, 除了要求 ξ 的方差存在外, 为了保证证明中 $n \rightarrow \infty$ 时, 极限号可以和求期望的符号交换, 我们还要求随机游动在直到停时 τ 为止, 不能超出一个给定的有界区域 $[-M, M]$.

例1.10.16 (例 1.10.12 续). 对于赌徒赌博模型, τ 为赌博停止的时间, 则 M 可取为 $\max\{i, j\}$. Wald 等式给出了赌徒破产和完胜的概率(见例 1.10.12), 再利用 Wald 第二等式, 我们有

$$\begin{aligned} E\tau = ES_\tau^2 &= (-i)^2 P(S_\tau = -i) + j^2 P(S_\tau = j) \\ &= i^2 \frac{j}{i+j} + j^2 \frac{i}{i+j} = i \times j. \end{aligned}$$

这也可以给出了 (1.10.11) 的另一个证明.

习题

1. 假设 $\{Y_n\}$ 为离散圆周 \mathbb{S}_N 上的简单随机游动. 试求 $\{Y_n\}$ 在首次回到初始点之前走遍所有顶点的概率.
2. 假设 $\{S_n\}$ 为 \mathbb{Z} 上的紧邻随机游动, (即步长

$$P(\xi_n = 1) = 1 - P(\xi_n = -1) = p,$$

定义见第 §1.1 节.) 假设 $0 < p < 1$, 且 $p \neq \frac{1}{2}$. N 和 M 为正整数. 令 $\tau = \tau_{-M} \wedge \tau_N$. (1) 试求 $P_0(\tau_N = n)$, 与 $E_0 \tau_N$, (2) 证明: $P_0(\tau < \infty) = 1$, (3) 求 $P_0(S_\tau = N)$, (4) 求 $E_0 \tau$.

3. 假设 $\{S_n\}$ 为一维简单随机游动, $M, N \in \mathbb{Z}$, $N - M \geq 2$. 记

$$\tau = \tau_M \wedge \tau_N = \min\{n : S_n = M \text{ 或 } N\}.$$

假设 $i, j \in [M + 1, N - 1] \cap \mathbb{Z}$, 求: $E_i \sum_{n=0}^{\tau-1} \mathbf{1}_{\{X_n=j\}}$.

- 4* 某研究员每隔一段独立同分布的随机时间观察一次实验进度, 间隔时间 ξ 等概率地为 1 分钟, 2 分钟, … 30 分钟. 假设他/她在某整点进行了一次观察. 请问: 平均多长时间后他/她再一次恰好在整点进行观察?

- 5* 将定理 1.10.11 中的“独立同分布”这一假设改为“相互独立, 且具有相同的期望”, 证明: 该定理的结论仍然成立.

- 6* 假设 $\{S_n\}$ 为一维简单随机游动. 令 $M_n^{(1)} = S_n$, $M_n^{(2)} = S_n^2 - n$, $M_n^{(3)} = S_n^3 - 3nS_n$, $M_n^{(4)} = S_n^4 - 6nS_n^2 + 3n^2 + 2n$. 证明: $EM_n^{(k)} = 0$, $k = 1, 2, 3, 4$.

- 7* 在上题中, 假设 τ 是 $\{S_n\}$ 停时. 试给出一个充分条件, 使得 $EM_\tau^{(k)} = 0$. (注: $k = 1, 2$ 时分别对应 Wald 等式与 Wald 第二等式.)

§1.11 分支过程

在1874年, 英国人Francis Galton 和Reverend H.W. Watson 发现一些英国的古老姓氏消失了. 他们在研究这个问题时建立了这个模型. 因此, 分支过程又被称为**Galton-Watson 过程**, 或简称为**GW 过程**. 当然, 此模型也可用于描述某动物种群的个体数量变化, 或核反应中裂变的原子个数. 因为这一原因, 第二次世界大战以后, 美、苏等国都对此非常重视.

假设所有孩子都跟父亲姓. 在某个家族中, 我们用 $\xi_{n,i}$ 表示第 n 代第 i 个男丁的儿子数, 用 X_n 表示第 n 代的男丁总数. 则我们有如下的递归式:

$$X_{n+1} := \sum_{i=1}^{X_n} \xi_{n,i}. \quad (1.11.1)$$

我们假设 $\{\xi_{ni} : n \geq 0, i \geq 1\}$ 独立同分布, 这蕴含着 $\{X_n\}$ 为一个马氏链. 此时, 称 $\{X_n\}$ 为一个**分支过程**. 分支过程的另外一个常用描述如下: 将每个男丁视为一个粒子, 假设每个粒子过一个单位时间死亡, 它死亡时立刻分裂为若干个新粒子(新粒子的数目也可以为0). 如果没有特别声明, 我们总假设 $X_0 = 1$, 并将 P_1 简记为 P .

在分支过程中, 我们称一个男丁的儿子为其子代, 其所有子孙为其后代. 将 $\xi := \xi_{0,1}$ 的分布记为 p_k , $k \geq 0$, 即

$$p_k = P(\xi = k), \quad \forall k \geq 0.$$

该分布(列)被称为分支过程的子代分布(列). 于是, 分支过程的转移概率由其子代分布完全确定. 换句话说, 欲刻画一个分支过程, 只需交代其子代分布即可. 为排除平凡情形, 下面我们总假设 $p_1 \neq 1$, (否则 $X_n \equiv 1, \forall n$).

如果在某个时刻 n , 发现 $X_n = 0$, 那么, 对任何 $m \geq 1$, $X_{n+m} = 0$. 即, 一旦在某一代(第 n 代)已经没有男丁, 那么之后就再也不会有男丁, 过程就停止不动了, 处于状态0. 换句话说, 状态0 为一个吸收态. 称

$$\tau_0 := \inf\{n \geq 0 : X_n = 0\}$$

为该分支过程的灭绝时间, 因为整个家族在第 τ_0 代灭绝. 若 $\tau_0 < \infty$, 我们称该过程灭绝, 否则称该过程存活. 注意, 灭绝和存活都是个事件. 将灭绝这一事件发生的概率

$$\rho := P(\tau_0 < \infty)$$

称为该过程的灭绝概率. 由于除0 外的其它状态都是暂态(这一结论留为习题), 因此若过程存活, 则 $X_n \rightarrow \infty$. 换句话说, $\{\tau = \infty\} \subseteq \{X_n \rightarrow \infty\}$.

在分支过程的研究中, 母函数是一个最重要而基本的工具. 将随机变量 X 的母函数记为 f_X , 即

$$f_X(s) = Es^X,$$

将 ξ 的母函数简记为 f , 即

$$f(s) = \sum_{k=0}^{\infty} p_k s^k = p_0 + p_1 s + p_2 s^2 + \dots$$

命题1.11.1. $f_{X_n}(s) = f^{(n)}(s)$, $n \geq 1$.

证: 我们用归纳法进行证明. 当 $n = 1$ 时, 由 $X_1 = \xi_{0,1} = \xi$ 知 $f_{X_1} = f$, 即命题成立. 假设 $f_{X_n}(s) = f^{(n)}(s)$, 往证 $f_{X_{n+1}}(s) = f^{(n+1)}(s)$. 根据归纳定义式 1.11.1 与全概公式,

$$f_{X_{n+1}}(s) = E s^{X_{n+1}} = E s^{\sum_{i=1}^{X_n} \xi_{n,i}} = \sum_{k=0}^{\infty} P(X_n = k) E\left(s^{\sum_{i=1}^k \xi_{n,i}} | X_n = k\right).$$

由于 X_n 仅依赖于 $\{\xi_{m,i} : 0 \leq m \leq n-1, i \geq 1\}$, 它与 $\{\xi_{n,i} : i \geq 1\}$ 相互独立, 因此

$$E\left(s^{\sum_{i=1}^k \xi_{n,i}} | X_n = k\right) = E s^{\sum_{i=1}^k \xi_{n,i}} = f(s)^k,$$

其中, 我们用到 $\xi_{n,i}, i \geq 1$ 独立同分布, 且都与 ξ 同分布. 将其代入上面的式子便得到

$$f_{X_{n+1}}(s) = \sum_{k=0}^{\infty} P(X_n = k) f(s)^k = f_{X_n}(f(s)).$$

最后, 由归纳假设, 上式右边就是 $f^{(n+1)}(s)$. 综上, 命题成立. \square

在判别分支过程是否以概率1 灭绝概率时, 一个至关重要的指标就是平均子代数

$$m = E\xi = \sum_{k=0}^{\infty} kp_k.$$

命题1.11.2. $EX_n = m^n, \forall n \geq 0$. 进一步, 若 $m < 1$, 则 $\rho = 1$.

证: 当 $n = 0$ 和 1 时, $EX_0 = 1 = m^0$, $EX_1 = E\xi = m$.

$$\begin{aligned} EX_{n+1} &= E \sum_{j=1}^{X_n} \xi_{nj} = \sum_{k=0}^{\infty} P(X_n = k) E\left(\sum_{j=1}^k \xi_{nj} | X_n = k\right) \\ &= \sum_{k=0}^{\infty} P(X_n = k) k E\xi = m EX_n. \end{aligned}$$

由归纳法知 $EX_n = m^n, \forall n \geq 0$.

进一步, 如果 $m < 1$, 那么, 由马尔可夫不等式知, 当 $n \rightarrow \infty$ 时

$$P(\tau_0 = \infty) \leq P(\tau_0 \geq n) = P(X_n \geq 1) \leq EX_n = m^n \rightarrow 0.$$

于是 $P(\tau_0 = \infty) = 0$, 即 $\rho = 1$. \square

我们还可以用“分支树”来更精细地刻画分支过程. 具体地, 如图 1.21 所示, 每个男丁就是一个顶点, 父子关系就是边, 第0 代的男丁是根点, 由它往下连出来 $\xi_{0,1}$ 条边(从左到右

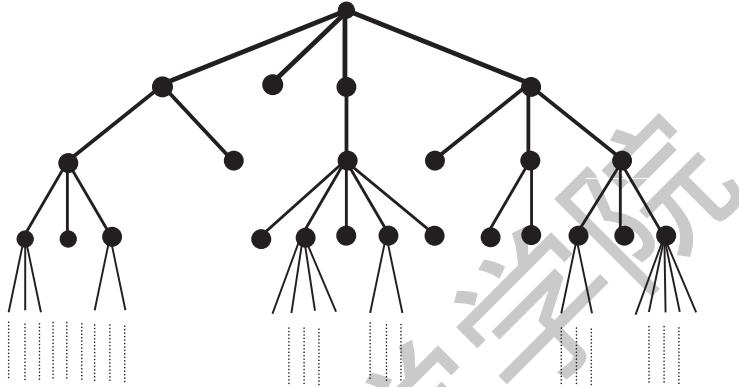


图 1.21 $\xi_{0,1} = 4; \xi_{1,1} = 2, \xi_{1,2} = 0, \xi_{1,3} = 1, \xi_{1,4} = 3; \xi_{2,1} = 3, \xi_{2,2} = 0, \xi_{2,3} = 5, \xi_{2,4} = 0, \xi_{2,5} = 2, \xi_{2,6} = 3, \dots$.

画), 其中, 最左边的边连接着根点的大儿子, 他有 $\xi_{1,1}$ 个儿子, 因此, 从它往下又连出 $\xi_{1,1}$ 条边, 左起第二条边连接着根点的二儿子, 他有 $\xi_{1,2}$ 个儿子, 因此, 从它往下又连出 $\xi_{1,2}$ 条边, (图示 $\xi_{1,2} = 0$, 因此实际上图中并没有往下画边), \dots . 因为 $\xi_{n,i}, n \geq 0, i \geq 1$ 都是随机变量, 所以按上述方法画出来的是一棵随机的树, 记为 \mathbb{T} . 将根点的第 i 个儿子记为 v_i . 考虑 v_i 的子树 \mathbb{T}_i , 即, 断开 v_i 与根点之间的边以后, v_i 与其子孙所在的连通分支, 其中, v_i 被视为新的根点. 那么, 由 $\xi_{n,i}, n \geq 0, i \geq 1$ 独立同分布, 不难看出, 如果 $\xi_{0,i} = n \geq 1$, 那么 $\mathbb{T}_1, \dots, \mathbb{T}_n$ 独立同分布, 此时, 每个 \mathbb{T}_i 的结构仍然是一个子代分布为 $p_k, k \geq 0$ 的分支树, 即, \mathbb{T}_i 与 \mathbb{T} 的原始结构是相似的. 这是分支树的自相似性. 我们可以由此自相似得到很多递归式, 例如, 下面的命题中的结论 $\rho = f(\rho)$.

命题1.11.3. ρ 是方程 $s = f(s)$ 的最小非负解. 于是, $\rho < 1$ 当且仅当 $m > 1$.

证: 由全概公式,

$$\rho = \sum_{k=0}^{\infty} P(\tau_0 < \infty, X_1 = k) = p_0 + \sum_{k=1}^{\infty} p_k P_k(\tau_0 < \infty).$$

下面我们研究 $P_k(\tau_0 < \infty), k \geq 1$. 如果最初有 k 个粒子, 记 A_r = “第 r 个粒子对应的分支灭绝”, 即, “ \mathbb{T}_r 有限”. 一方面 $P_k(A_r) = P(\tau_0 < \infty) = \rho, r = 1, \dots, k$; 另一方面, 由于粒子的分裂具有独立性, 我们知 A_1, \dots, A_k 相互独立. 故

$$P_k(\tau_0 < \infty) = P_k(A_1 \cap \dots \cap A_k) = P_k(A_1) \cdots P_k(A_k) = \rho^k.$$

于是我们推出

$$\rho = p_0 + \sum_{k=1}^{\infty} p_k \rho^k = f(\rho).$$

因为 $p_1 \neq 1$, 所以 $f(\rho) = \rho$ 的解就是图 1.22 中的曲线与斜线的交点.



图 1.22 灭绝概率

因为 $m = f'(1)$, 所以当 $m \leq 1$ 时(如图1.22 的左图), 方程 $f(s) = s$ 在 $[0, 1]$ 上只有一个不动点, 该不动点就是1, 所以 $\rho = 1$; 当 $m > 1$ 时(如图1.22 的右图), 方程 $s = f(s)$ 有两个不动点: ρ_0 和1, 其中 $0 \leq \rho_0 < 1$. 因为

$$\begin{aligned} \rho &= P(\tau_0 < \infty) = \lim_{n \rightarrow \infty} P(\tau_0 \leq n) = \lim_{n \rightarrow \infty} P_1(X_n = 0) \\ &= \lim_{n \rightarrow \infty} f_{X_n}(0) = \lim_{n \rightarrow \infty} f^{(n)}(0) \leq \lim_{n \rightarrow \infty} f^{(n)}(\rho_0) = \rho_0. \end{aligned}$$

因此 $\rho = \rho_0$.

□

注1.11.4. 图 1.22 的右图中的折线反应了 $f^{(n)}(0)$ 的迭代过程.

定义1.11.5. 根据 $m < 1$, $m = 1$ 和 $m > 1$, 我们分别称对应的分支过程为次临界的(或下临界的), 临界的, 和超临界的(或上临界的).

根据命题 1.11.3, 次临界和临界的分支过程都以概率1 灭绝, 而超临界的分支过程以正概率存活, 其存活概率为 $1 - \rho$, 灭绝概率为 ρ .

例1.11.6. 假设子代分布列为 $p_k = p(1-p)^k$, $k \geq 0$, 其中, $0 < p < 1$. 试求灭绝概率 ρ .

解: 首先, 计算平均后代表数.

$$m = \sum_{k=0}^{\infty} kp_k = \sum_{k=0}^{\infty} kp(1-p)^k = \frac{p(1-p)}{(1-(1-p))^2} = \frac{1-p}{p}.$$

然后判别 $m > 1$ 是否成立并求 ρ . 当 $p \geq 1/2$ 时, $m \leq 1$, 故 $\rho = 1$. 而当 $p < 1/2$ 时, $m > 1$,

$$\rho = f(\rho) = \sum_k \rho^k p_k = \sum_k \rho^k p(1-p)^k = \frac{p}{1-\rho(1-p)}.$$

解得 $\rho = \frac{p}{1-p}$ 或 1, 舍弃 1 这个根, 得到 $\rho = \frac{p}{1-p}$.

习题

1. 证明: 分支过程中, 除状态0 以外的其它状态都是暂态.
2. 假设子代分布为 $B(2, p)$, (1) 求灭绝概率 $\rho = P_1(\tau_0 < \infty)$; (2) 求 $P_1(\tau_0 = 3)$; (3) 假设 X_0 服从参数为 λ 的泊松分布, 试求 $P(\tau_0 < \infty)$.
3. 假设平均子代数 $m = E\xi < 1$. 证明: $E_1\tau_0 < \infty$.
4. 设分支过程 $\{X_n\}$ 的子代分布的均值为 μ , 方差为 σ^2 . 试证:

$$E(X_{n+1}^2) = \mu^n \sigma^2 + \mu^2 E X_n^2.$$

5. 假设子代分布为 $p_1 = p_2 = \frac{q}{2}$, $p_0 = 1 - q$. (1) 试问: q 取何值时灭绝概率 $\rho = 1$? (2) 假设 $q = 0.9$, 初值 $X_0 = 4$. 试求存活概率 $P_4(\tau_0 = \infty)$.
6. 假设某物种的子代数目 ξ 服从参数为 λ 的泊松分布, 每个个体独立地以概率 p 能活到成年并生子. 将第 n 代中成年的个体数记为 Y_n , $n \geq 0$. 假设 $Y_n = 1$, 那么 $\{Y_n\}$ 也是分支过程. (1) 求 $\{Y_n\}$ 的子代分布, (2) 求 $\{Y_n\}$ 为超临界的充分必要条件(用 λ, p 的多项式来表示).

7* 考虑带移民的分支过程 $\{X_n\}$, 即 $X_{n+1} = \sum_{i=1}^{X_n} \xi_i^n + Y_{n+1}$. 其中每个个体以概率 p_k 产生 k 个后代, 每一代以概率 h_k 进来 k 个新移民. 设 $P(\xi_k = i) = p_i$, $P(Y_n = j) = h_j$. 进一步设 $\sum k p_k < 1$, $\sum j h_j < \infty$. (1) 证明 $\{X_n\}$ 具有不变分布 π . (2) 记 $\pi(s) = \sum_i \pi_i s^i$, $\phi(s) = \sum_i p_i s^i$, $h(s) = \sum_i h_i s^i$. 证明 $\pi(\phi(s))h(s) = \pi(s)$.

§1.12 综合练习题

1. 试举出三个生灭过程的例子, 分别是非常返的, 正常返的和零常返的.
2. 假设 S 有限、 \mathbf{P} 不可约. (1) 证明: 若 \mathbf{P} 可逆的, 则 \mathbf{P} 的所有特征根为实数. (注: \mathbf{P} 与对称矩阵 \mathbf{Q} 相似, 其中 $q_{ij} = \sqrt{\pi_i} p_{ij} / \sqrt{\pi_j}$.) (2) 举例说明: 所有特征根为实数的转移矩阵未必是可逆的.
3. 假设某地区的天气可分为晴天、阴天、雨天. 若某天是晴天, 则第二天等可能为阴天或雨天, 若某天为阴天(或雨天), 则第二天有 $\frac{1}{2}$ 的可能继续为阴天(或雨天), 有 $\frac{1}{4}$ 的可能为晴天, 有 $\frac{1}{4}$ 的可能为雨天(或阴天). (1) 试建立一个马氏链描述该地区的天气变化. (2) 这个马氏链有可逆分布吗? (3) 试求出平均而言, 该地区出现晴天、阴天、雨天的比例. (4) 两年后的元旦节下雨的概率大概是多少?
4. 假设 $\{X_n\}$ 是 \mathbb{Z}_+ 上的马氏链, 转移概率为 $p_{0,1} = 1$,

$$p_{i,i+1} + p_{i,i-1} = 1, \quad p_{i,i+1} = \left(\frac{i+1}{i}\right)^\alpha p_{i,i-1}, \quad i \geq 1.$$

试根据 α 的取值判断该马氏链是否常返; 若是正常返, 试求不变分布; 若是非常返, 试求 $P_0(\sigma_0 < \infty)$.

5. 设 $\{X_n\}$ 是状态空间 $\{1, 2, 3, 4, 5\}$ 上的马氏链, 其转移矩阵

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

(1) 该马氏链是否不可约? 是否为非周期? (2) 求其不变分布; (3) 设 $X_0 = 1$, 求再次回到状态 1 所需要的平均步数; (3) 求从状态 1 出发到达状态 4 所需要的平均步数. (4) 设 $X_0 = 1$, 求该马氏链到达状态 3 之前到达状态 5 的概率.

6. 设 J_1, J_2, \dots 为一列独立同分布随机变量, $P(J_1 = 1) = p$, $P(J_1 = 0) = 1 - p$, 其中 $0 < p < 1$. (1) 设 $X_0 = 0$, 对任意 $n \geq 1$, 若

$$J_n = \dots = J_{n-k+1} = 1, \text{ 且 } J_{n-k} = 0,$$

则令 $X_n = k$. (1) 试解释 $\{X_n\}$ 是马氏链, 给出状态空间和转移概率. (2) 证明该马氏链不可约、正常返, 并求其不变分布. (3) 对于正整数 k , 令

$$T_k = \min\{n \geq k : J_n = J_{n-1} = \dots = J_{n-k+1} = 1\},$$

试求 ET_k .

7* 假设马氏链 $\{X_n\}$ 不可约, 有可逆分布. 试证明下列两条命题等价. (a) $P_i(X_n = i) = P_j(X_n = j), \forall i, j \in S, \forall n \geq 0$; (b) $P_i(\tau_j = n) = P_j(\tau_i = n), \forall i, j \in S, \forall n \geq 0$.

第二章 跳过程

这一章我们所讨论的跳过程是一个连续时间参数的随机过程, 它的许多内容与前一章是平行的, 只是时间参数不同而已. 因此, 我们将重点介绍跳过程与马氏链的不同之处. 在本章中, t, s, \dots 表示 $\mathbb{R}_+ := [0, \infty)$ 中的点, 它们为连续时间的参数; n, m, \dots 表示 $\mathbb{Z}_+ := \{0, 1, 2, \dots\}$ 中的点, 它们为时间参数. 我们仍然将 $\{X_t : t \in \mathbb{R}_+\}$ 简记为 $\{X_t\}$, 将 $\{Y_n : n \in \mathbb{Z}_+\}$ 简记为 $\{Y_n\}$.

§2.1 泊松过程

泊松过程(Poisson Process)的结构很简单, 但其性质相当丰富. 它是最简单的跳过程, 也是讨论其它跳过程的基础. 泊松过程还常常用于构造更复杂的连续时间参数的过程. 在本节中, 我们将从两个角度介绍泊松过程. 第一个角度基于指数分布, 侧重于强调指数分布的无记忆性蕴含着泊松过程的马氏性; 第二个角度则基于泊松分布与均匀分布, 侧重于介绍泊松流的叠加和细分.

一、从指数分布到泊松过程.

假设有一个闹钟, 它等待 ξ_1 时间后响一下, 然后再等待 ξ_2 时间后再响一下, 然后再等待 ξ_3 时间后再响一下, …… 我们称 ξ_1, ξ_2, \dots 为 $\{X_t\}$ 的等待时间. 于是, 对任意 $n \geq 1$, 闹钟第 n 次响起的时刻为

$$S_n = \xi_1 + \xi_2 + \dots + \xi_n.$$

补充定义 $S_0 := 0$. 对任意 $t \geq 0$, 将时刻 t 之前闹钟响起的总次数记为 X_t . 那么,

$$X_t := |\{n \geq 1 : S_n \leq t\}| = \sup\{n \geq 0 : S_n \leq t\}, \quad \forall t \geq 0. \quad (2.1.1)$$

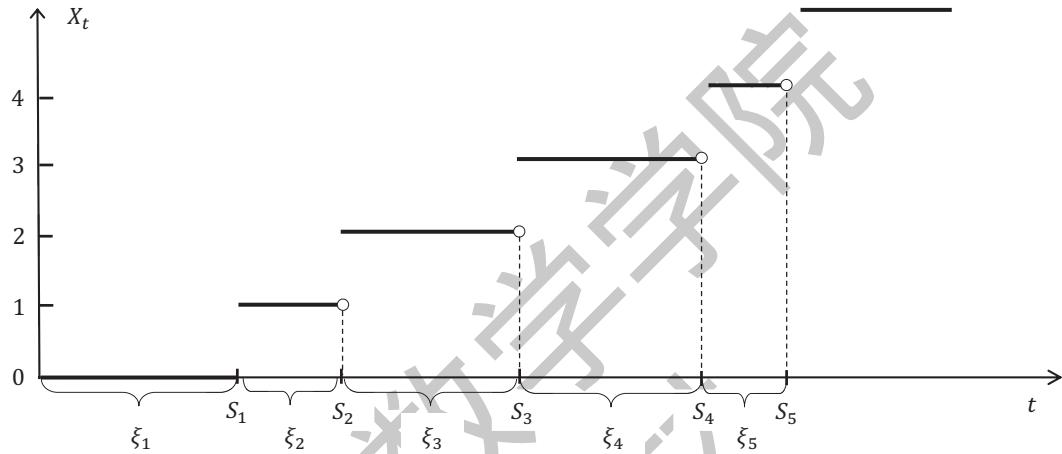
定义2.1.1. 如果 ξ_1, ξ_2, \dots 独立同分布, 并且 $\xi_1 \sim \text{Exp}(\lambda)$, 那么, 称如上定义的 $\{X_t : t \geq 0\}$ 为一个泊松过程, 称 λ 为其参数, 强度, 或者速率.

若 $\{X_t\}$ 满足:

- (i) 非负性: $X_0 = 0$, X_t 取非负整数值;
 - (ii) 单调性: X_t 关于 t 单调上升;
 - (iii) 右连左极性(continue à droite, limites à gauche, 简称 càdlàg): X_t 右连续, 即, 对任意 $t \geq 0$, 都有 $\lim_{s \searrow t^+} X_s = X_t$; 并且 X_t 具有左极限, 即, 对任意 $t > 0$, 都有 $\lim_{s \nearrow t} X_s$ 存在,
- 则称 $\{X_t\}$ 为一个计数过程. 泊松过程是计数过程的特例. 如果一个计数过程 $\{X_t\}$ 如图 §2.1 所示, 我们可以把发生时刻 S_1, S_2, \dots 与等待时间 ξ_1, ξ_2, \dots 求出来. 具体地,

$$S_n = \inf\{t \geq 0 : X_t \geq n\}, \quad \forall n \geq 0, \quad \xi_n = S_n - S_{n-1}, \quad \forall n \geq 1.$$

当 ξ_1, ξ_2, \dots 独立同分布且 $\xi_1 \sim \text{Exp}(\lambda)$ 时, 我们称上面的闹钟为一个指数闹钟. 换句话说, 一个指数闹钟和一个泊松过程是等价的. 我们也将 λ 称为该指数闹钟的参数.



以下总假设 $\{X_t\}$ 是参数为 λ 的泊松过程, 即, 假设 ξ_1, ξ_2, \dots 独立同分布, 且 $\xi_1 \sim \text{Exp}(\lambda)$, 且 X_t 由 (2.1.1) 定义. 下面的命题表明, 对任意 $t > 0$, X_t 均服从泊松分布, 这也是 $\{X_t\}$ 被称为泊松过程的缘由. 证明主要用到了 $\{X_t\}$ 与 $\{S_n\}$ 的如下关系式:

$$\{X_t = k\} = \{S_k \leq t < S_{k+1}\}, \quad \forall t \in \mathbb{R}_+, \forall k \in \mathbb{Z}_+. \quad (2.1.2)$$

该关系式可以将一个关于 $\{X_t\}$ 的事件转化为关于 $\{S_n\}$ 的事件, 继而转化为关于 ξ_1, ξ_2, \dots 的事件. 于是, 我们便可以求出该事件的概率.

定理2.1.2. 对任意 $t \geq 0$, X_t 服从参数为 λt 的泊松分布.

证: 固定 $t \geq 0$. 我们计算 X_t 的分布列.

当 $k = 0$ 时, $P(X_t = 0) = P(\xi_1 > t) = e^{-\lambda t}$.

下面假设 $k \geq 1$. 由 (2.1.2), 我们需要计算事件 $\{S_k \leq t < S_{k+1}\}$ 的概率. 注意到该事件只涉及 ξ_1, \dots, ξ_{k+1} . 为此, 我们用 $\vec{x} = (x_1, \dots, x_k)$ 表示 (ξ_1, \dots, ξ_k) 的具体取值, 同时, 用 x_{k+1} 表示 ξ_{k+1} 的具体取值. 将 $(x_1, \dots, x_k, x_{k+1})$ 简记为 (\vec{x}, x_{k+1}) . 对任意 $n \geq 1$, 将 $x_1 + \dots + x_n$ 记为 s_n . 令

$$\begin{aligned} C(k; t) &= \{\vec{x} \in \mathbb{R}^k : x_1, \dots, x_k \geq 0, \text{ 且 } s_k \leq t\}, \\ D(k; t) &= \{(\vec{x}, x_{k+1}) : \vec{x} \in C(k; t), \text{ 且 } x_{k+1} > t - s_k\}. \end{aligned} \quad (2.1.3)$$

那么,

$$\{S_k \leq t < S_{k+1}\} = \{(\xi_1, \dots, \xi_{k+1}) \in D(k; t)\}.$$

上述事件的概率就是 $(\xi_1, \dots, \xi_{k+1})$ 的联合密度在 $D(k; t)$ 上的积分. 因此, 对任意 $k \geq 1$,

$$\begin{aligned} P(X_t = k) &= P(S_k \leq t < S_{k+1}) \\ &= \int_{(\vec{x}, x_{k+1}) \in D(k; t)} \lambda^{k+1} e^{-\lambda s_{k+1}} dx_1 \cdots dx_{k+1} \end{aligned} \quad (2.1.4)$$

$$\begin{aligned} &= \int_{\vec{x} \in C(k; t)} \lambda^k e^{-\lambda s_k} \times \left(\int_{t-s_k}^{\infty} \lambda e^{-\lambda x_{k+1}} dx_{k+1} \right) dx_1 \cdots dx_k \\ &= \int_{\vec{x} \in C(k; t)} \lambda^k e^{-\lambda s_k} \times e^{-\lambda(t-s_k)} dx_1 \cdots dx_n \\ &= \int_{\vec{x} \in C(k; t)} \lambda^k e^{-\lambda t} dx_1 \cdots dx_k = \lambda^k e^{-\lambda t} \times \frac{t^k}{k!}, \end{aligned} \quad (2.1.5)$$

其中, 最后的 $t^k/k!$ 就是 k 维椎体 $C(k, t)$ 的体积.

综上, X_t 服从参数为 λt 的泊松分布. \square

进一步, 给定 $T > 0$. 我们还可以考虑对任意 $t > 0$, 在 $(T, T+t]$ 时间段内指数闹钟响起的总次数. 具体地, 令

$$Y_t := X_{T+s} - X_T, \quad t \geq 0.$$

下面, 我们考察 $\{Y_t\}$ 的等待时间序列. 如果从时刻 T 开始, 需要等待 η_1 的时间闹钟才会响起, 接着再等待 η_2 的时间闹钟再次响起……那么, 如图 2.1 所示, 当 $\{X_T = k\}$ 发生时, $\eta_1 = \xi_{k+1} - (t - S_k)$, $\eta_2 = \xi_{k+2} - \dots$. 换句话说,

$$\eta_1 = \xi_{X_T+1} - (t - S_{X_T}), \quad \eta_2 = \xi_{X_T+2}, \quad \eta_3 = \xi_{X_T+3}, \dots \quad (2.1.6)$$

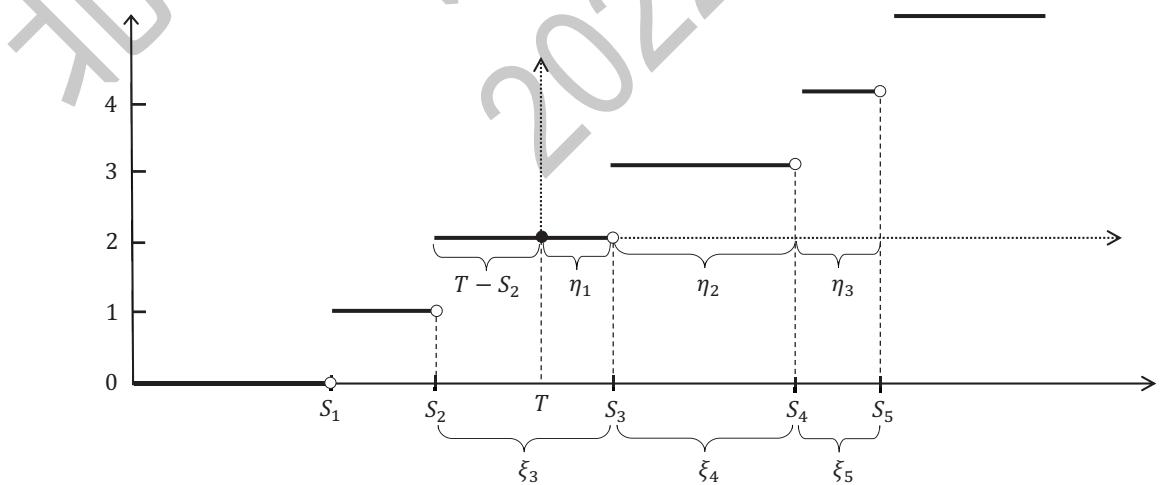


图 2.1 $X_t = 2, \eta_1, \eta_2, \dots$ 独立同分布.

定理2.1.3. 给定 $T > 0$. $X_T, \eta_1, \eta_2, \dots$ 相互独立. 并且, 对任意 $n \geq 1$ 都有 $\eta_n \sim \text{Exp}(\lambda)$.

证: 对任意 $k \geq 0, n \geq 1, t_1, \dots, t_n > 0$, 根据 η_1, η_2, \dots 的定义,

$$\{X_t = k, \eta_1 > t_1, \dots, \eta_n > t_n\} = \{S_k \leq t, S_{k+1} > t + t_1, \xi_{k+2} > t_2, \dots, \xi_{k+n} > t_n\}.$$

于是,

$$P(X_t = k, \eta_1 > t_1, \dots, \eta_n > t_n) = P(S_k \leq t, S_{k+1} > t + t_1) \times e^{-\lambda(t_2 + \dots + t_n)}.$$

我们可仿照 (2.1.6) 计算, 并得到

$$P(S_k \leq t, S_{k+1} > t + t_1) = \int_{\vec{x} \in C(k; t)} \lambda^k e^{-\lambda s_k} \times e^{-\lambda(t+t_1-s_k)} dx_1 \dots dx_{k+1} = e^{-\lambda t_1} P(X_t = k).$$

因此,

$$P(X_t = k, \eta_1 > t_1, \dots, \eta_n > t_n) = P(X_t = k) e^{-\lambda(t_1 + \dots + t_n)}.$$

从而结论成立. \square

推论2.1.4. 给定 $T > 0$. 对任意 $t \geq 0$, 令 $Y_t := X_{T+t} - X_T$. 那么, $\{Y_t\}$ 也是参数为 λ 的泊松过程. 并且, 对任意 $r \geq 1, 0 < t_1 < \dots < t_r$ 都有 $(Y_{t_1}, \dots, Y_{t_r})$ 与 X_T 相互独立.

证: $\{Y_t\}$ 的等待时间是 η_1, η_2, \dots . 根据定理 2.1.3, $\{Y_t\}$ 是参数为 λ 的泊松过程. 并且, 对任意非负整数 $i_1 \leq \dots \leq i_r$, 我们可以仿照 (2.1.2), 将 $\{(Y_{t_1}, \dots, Y_{t_r}) = (i_1, \dots, i_r)\}$ 改写为关于 $\eta_1, \dots, \eta_{i_r+1}$ 的事件, 由定理 2.1.3, 该事件与关于 X_T 的任意事件 $\{X_T = i\}$ 都独立. 因此, $(Y_{t_1}, \dots, Y_{t_r})$ 与 X_T 相互独立. \square

推论2.1.5. 假设 $r \geq 1, 0 = t_0 < t_1 < \dots < t_r$. 那么, 对任意非负整数 k_1, \dots, k_r ,

$$P(X_{t_i} - X_{t_{i-1}} = k_i, i = 1, \dots, r) = \prod_{i=1}^r P(X_{t_i} - X_{t_{i-1}} = k_i).$$

证: 当 $r = 1$ 时, 结论自然成立. 假设结论关于 r 成立, 往证结论关于 $r + 1$ 也成立.

对任意 $t \geq 0$, 令 $Y_t = X_{t_1+t} - X_{t_1}$. 记 $s_i = t_{i+1} - t_1, i = 0, 1, \dots, r$. 那么,

$$\{X_{t_i} - X_{t_{i-1}} = k_i, i = 2, \dots, r+1\} = \{Y_{s_i} - Y_{s_{i-1}} = k_{i+1}, i = 1, \dots, r\}.$$

一方面, 由推论 2.1.4,

$$\begin{aligned} P(X_{t_i} - X_{t_{i-1}} = k_i, i = 1, \dots, r+1) &= P(X_{t_1} = k_1) P(Y_{s_i} - Y_{s_{i-1}} = k_{i+1}, i = 1, \dots, r) \\ &= P(X_{t_1} = k_1) P(X_{s_i} - X_{s_{i-1}} = k_{i+1}, i = 1, \dots, r). \end{aligned}$$

根据归纳假设，并注意到 $s_i - s_{i-1} = t_{i+1} - t_i$, $i = 1, \dots, r$, 我们推出

$$P(X_{s_i} - X_{s_{i-1}} = k_{i+1}, i = 1, \dots, r) = \prod_{i=1}^r P(X_{s_i - s_{i-1}} = k_{i+1}) = \prod_{i=2}^{r+1} P(X_{t_i - t_{i-1}} = k_i).$$

综上，

$$P(X_{t_i} - X_{t_{i-1}} = k_i, i = 1, \dots, r+1) = P(X_{t_1} = k_1) \prod_{i=2}^{r+1} P(X_{t_i - t_{i-1}} = k_i) = \prod_{i=1}^{r+1} P(X_{t_i - t_{i-1}} = k_i).$$

由归纳法，命题成立。 \square

注2.1.6. 反过来，假设 $\{X_t\}$ 是一个计数过程，并且对任意 $r \geq 1$, $0 = t_0 < t_1 < \dots < t_r$, 以及非负整数 k_1, \dots, k_r , 都有

$$P(X_{t_i} - X_{t_{i-1}} = k_i, i = 1, \dots, r) = \prod_{i=1}^r \frac{(\lambda(t_i - t_{i-1}))^{k_i}}{k_i!} e^{-\lambda(t_i - t_{i-1})}.$$

那么, $\{X_t\}$ 是参数为 λ 的泊松过程。

推论2.1.7 (马氏性). 对任意 $r \geq 1$, 任意 $0 = t_0 < t_1 < \dots < t_r < t < t+s$, 以及任意单调上升的非负整数 $0 \leq k_1 \leq \dots \leq k_r \leq k \leq \ell$, 都有

$$\begin{aligned} & P(X_{t+s} = \ell | X_t = k, X_{t_1} = k_1, \dots, X_{t_r} = k_r) \\ &= P(X_{t+s} = \ell | X_t = k) = P(X_s = \ell - k). \end{aligned}$$

二、从均匀分布与泊松分布到泊松过程.

给定 $t > 0$. 假设在 $[0, t]$ 时间段上闹钟响过恰好 k 次, 其中 $k \geq 1$. 现在我们考察这 k 次闹钟响起的时刻所服从的联合分布.

命题2.1.8. 对任意 $0 < s < t$, $P(S_1 \leq s | X_t = 1) = s/t$.

证: 根据条件概率的定义,

$$\begin{aligned} P(S_1 \leq s | X_t = 1) &= \frac{P(S_1 \leq s, X_t = 1)}{P(X_t = 1)} \\ &= \frac{P(\xi_1 \leq s, \xi_2 > t - \xi_1)}{P(X_t = 1)} = \frac{\int_0^s \int_{t-u}^{\infty} \lambda e^{-\lambda u} \times \lambda e^{-\lambda v} dv du}{e^{-\lambda t} \lambda t} \\ &= \frac{\int_0^s \lambda e^{-\lambda u} e^{-\lambda(t-u)} du}{e^{-\lambda t} \lambda t} = \frac{\int_0^s \lambda e^{-\lambda t} du}{e^{-\lambda t} \lambda t} = \frac{s}{t}. \end{aligned}$$

\square

注意到在 $\{X_t = 1\}$ 发生的条件下, S_1 只能取值于 $(0, t)$. 上面的命题表明, 如果在 $[0, t]$ 上闹钟响过恰好一次, 那么, 我们可以用一个服从 $U(0, t)$ 的随机变量 U_1 来标记闹钟响起的时

刻. 一般地, 如果在 $(0, t]$ 上闹钟恰好响了 k 次, 其中 $k \geq 1$, 那么, 我们需要 k 个服从 $U(0, 1)$ 的随机变量 U_1, \dots, U_k 来标记闹钟响起的时刻, 事实上, 它们还是相互独立的. 该结论的严格表述为下面的命题 2.1.9, 因为如果 U_1, \dots, U_k 独立同分布, 都服从 $U(0, t)$, 那么 U_1, \dots, U_k 的顺序统计量 $U_{(1)}, \dots, U_{(k)}$ 的联合密度就是 $\frac{k!}{t^k} \times \mathbf{1}_{\{0 < s_1 < \dots < s_k < t\}}$, 其中 s_i 代表 $U_{(i)}$ 的取值. 命题 2.1.9 的证明与 $k = 1$ 的情形类似, 只是证明过程相对繁琐, 我们将之留为习题.

命题 2.1.9. 在 $\{X_t = k\}$ 发生的条件下, (S_1, \dots, S_k) 的条件密度为

$$\frac{k!}{t^k} \times \mathbf{1}_{\{0 < s_1 < \dots < s_k < t\}}.$$

接下来, 我们将 $[0, \infty)$ 切割为一系列长度为 1 的区间, $[0, 1), [1, 2), \dots$. 将闹钟在第 n 个时间段 $[n-1, n)$ 上响起的次数记为 Z_n . 那么, 由推论 2.1.5 知 Z_n 服从参数为 λ 的泊松分布. 再由命题 2.1.4 与命题 2.1.9, 在 $[n-1, n)$ 这个时间段上, 闹钟响起的时刻应该是相互独立且都服从 $U(n-1, n)$ 的随机变量序列的前 Z_n 项的顺序统计量. 我们用 $U_{n,1}, U_{n,2}, \dots$ 来代表这个随机变量序列. 于是, 所有闹钟响起的时刻就是

$$\Xi := \bigcup_{n=1}^{\infty} \{U_{n,1}, \dots, U_{n,Z_n}\}.$$

命题 2.1.10. 假设 $Z_n, U_{n,i}, n, i = 1, 2, \dots$ 全都相互独立. 假设对任意 $n \geq 1$, Z_n 服从参数为 λ 的泊松分布, 且 $U_{n,i} \sim U(n-1, n)$, $i = 1, 2, \dots$. 那么, $\{|[0, t] \cap \Xi|\}$ 是参数为 λ 的泊松分布.

上述命题的证明比较繁琐, 我们将其留为习题. 根据上述命题, 我们也可以通过考察 $\{|[0, t] \cap \Xi|\}$ 来研究泊松过程.

三、泊松过程的合并与细分.

本节介绍两个定理及其直观含义. 读者可以根据泊松过程的定义或命题 2.1.10 证明这两个定理. 验证的过程比较繁琐, 我们将其留为习题.

定理 2.1.11. (泊松过程的合并) 假设 $\{X_t\}$ 与 $\{Y_t\}$ 是相互独立的泊松过程, 参数分别为 λ_1 和 λ_2 . 令 $Z_t = X_t + Y_t$. 那么, $\{Z_t\}$ 是一个参数为 $\lambda_1 + \lambda_2$ 的泊松流.

为方便起见, 我们将 $\{X_t\}$, $\{Y_t\}$ 与 $\{Z_t\}$ 对应的闹钟分别称为 X 闹钟, Y 闹钟与 Z 闹钟. 在图 2.2 中的第一条时间轴上, 我们用“叉”标记了 X 闹钟响起时刻, 在第二条轴上, 我们用“点”标记了 Y 闹钟响起时刻. 第三条轴是前两条的合并, 它则标记了 Z 闹钟的响起时刻.

反过来, 如果 $\{Z_t\}$ 是一个参数为 $\text{Exp}(\lambda)$ 的泊松过程, 我们将 Z 闹钟每次响起的时刻独立地随机标记为“叉”或“点”, 概率分别为 p 和 $q = 1 - p$. 具体地, 假设 ϕ_1, ϕ_2, \dots 独立同分布, 且与 $\{Z_t\}$ 独立, $P(\phi_1 = 1) = p$, $P(\phi_1 = 0) = q$. 令

$$T_0^{(1)} := 0, \quad T_0^{(2)} := 0,$$

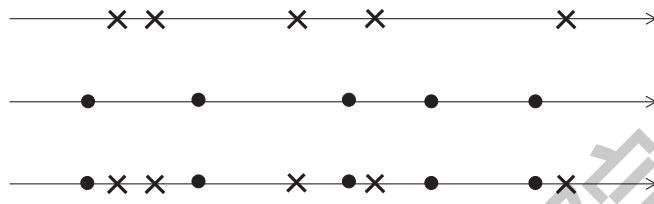


图 2.2 泊松过程的叠加与细分

$$T_n^{(1)} := \inf\{n \geq T_{n-1}^{(1)} + 1 : \phi_n = 1\}, \quad T_n^{(2)} := \inf\{n \geq T_{n-1}^{(2)} + 1 : \phi_n = 0\},$$

$$S_n := \inf\{t \geq 0 : Z_t \geq n\}, \quad \forall n \geq 1,$$

$$S_n^{(1)} := S_{T_n^{(1)}}, \quad S_n^{(2)} := S_{T_n^{(2)}}, \quad \forall n \geq 1.$$

那么, $\{S_n^{(1)}\}$ 就对应着所有带“叉”的时刻, 它如图 2.2 的第一条时间轴所示; 而 $\{S_n^{(2)}\}$ 则对应着所有带“点”的时刻, 它如第二条时间轴所示. 将它们对应的闹钟分别称为 X 闹钟与 Y 闹钟. 即, 对任意 $t \geq 0$, 令

$$X_t := |\{n \geq 1 : S_n^{(1)} \leq t\}|, \quad Y_t := |\{n \geq 1 : S_n^{(2)} \leq t\}|.$$

定理2.1.12. (泊松流的细分) 假设 $\{Z_t\}$ 是一个参数为 $\text{Exp}(\lambda)$ 的泊松过程. 那么, 如上定义的 $\{X_t\}$ 与 $\{Y_t\}$ 是相互独立的泊松过程, 参数分别为 λp 与 λq .

泊松过程的合并与细分表明: 两个独立的小的指数闹钟(即, 上述 X 闹钟与 Y 闹钟)放在一起就等价于一个大的指数闹钟(即, Z 闹钟)与一个硬币, 这里的大与小指的是参数的大与小. 我们还可以将此结论进行推广, 可数个小的指数闹钟放在一起(只要它们的参数之和有限), 就等价于一个大的指数闹钟与一个多面骰子.

例2.1.13. 某人钓鱼, 从下钩到钓到鱼的时间服从指数分布, 平均时间为半小时. 钓到草鱼的概率为 0.4, 钓到青鱼的概率为 0.6. 假设他连续钓鱼 2.5 小时. 试求他恰好钓到一条草鱼和两条青鱼的概率.

解: 将 1 小时视为单位时间. 假设直到 t 为止钓鱼者钓到了 Z_t 条鱼, 其中, 草鱼有 X_t 条, 青鱼有 Y_t 条. 由题意, $\{Z_t\}$ 是参数为 $\lambda = 2$ 的泊松过程. 根据命题 2.1.12, $\{X_t\}$ 与 $\{Y_t\}$ 是相互独立的泊松过程, 参数分别为 $\lambda \times 0.4 = 0.8$ 和 $\lambda \times 0.6 = 1.2$. 接下来, 根据定理 2.1.2, 2.5 小时后, 他钓到的草鱼数 $X_{2.5}$ 和青鱼数 $Y_{2.5}$ 相互独立, 都服从泊松分布, 参数分别为 $0.8 \times 2.5 = 2$ 和 $1.2 \times 2.5 = 3$. 因此, 所求概率为 $e^{-2} 2^1 / e^{-3} 3^2 / 2! = 9e^{-5}$.

例2.1.14 (复合泊松过程). 假设 $\{X_t\}$ 是一个参数为 λ 的泊松过程, ϕ_1, ϕ_2, \dots 是独立同分布的随机变量序列, 且与 $\{X_t\}$ 独立. 令 $Y_t = \sum_{n=1}^{X_t} \phi_n$. 则 $\{Y_t\}$ 被为复合泊松过程.

譬如, 假设某商场每过一个独立的 $\text{Exp}(\lambda)$ 时间会有一个旅行团前来购物, 第 n 个旅行团中有 ϕ_n 个人. 那么, 该商场在时刻 t 之前到达的顾客总数就是 Y_t . 我们有两种方法计算 Y_t . 一种是按照时间发展进行累计, 即, 定义中的 $\sum_{n=1}^{X_t} \phi_n$; 另一种则是先按照旅行团中的顾客数进行分类, 然后再累计, 即, $Y_t = \sum_{k=1}^{\infty} k X_t^{(k)}$, 其中 $X_t^{(k)} = \sum_{n=1}^{X_t} \mathbf{1}_{\{\phi_n=k\}}$ 表示时刻 t 之前到达的恰有 k 个人的旅行团的数目. 第二种方法的好处在于, 根据命题 2.1.12, $\{X_t^{(1)}\}, \{X_t^{(2)}\}, \dots$ 是相互独立的泊松过程. 因此, 当我们只关心某些特殊团体(例如, 至少有 50 人的旅行团)时, 其总人数仍形成一个复合泊松过程.

例2.1.15 (保险理赔). 假设某保险公司有 K 种保险计划, 第 k 种计划的理赔金额为 a_k , 理赔数目形成一个独立的泊松过程 $\{X_t^{(k)}\}$, $k = 1, \dots, K$. 那么, 该保险公司到时刻 t 为止所赔付的金额总数为 $Y_t = \sum_{k=1}^K a_k X_t^{(k)}$. 如果我们只想计算大额赔付(例如, 单笔超过 M 元)的金额, 那么我们只需找出 $I = \{k : 1 \leq k \leq K, a_k \geq M\}$, 并计算 $Z_t := \sum_{k \in I} a_k X_t^{(k)}$. 按照定义, $\{Y_t\}$ 与 $\{Z_t\}$ 都是复合泊松过程.

补充知识.

定义2.1.16. 假设 $\{X_t, t \geq 0\}$ 是一个取值于 \mathbb{R} 的过程.

若对任意 $r \geq 2$, 以及任意的 $0 < t_1 < t_2 < \dots < t_r$, 都有 $X_0, X_{t_1} - X_0, X_{t_2} - X_{t_1}, \dots, X_{t_r} - X_{t_{r-1}}$ 相互独立, 则称 $\{X_t\}$ 是一个独立增量过程.

若对任何 $t, s > 0$, $X_s - X_0$ 与 $X_{t+s} - X_t$ 同分布, 则称 $\{X_t\}$ 是一个平稳增量过程.

根据推论 2.1.5, 泊松过程具有独立平稳增量性, 即, 泊松过程既是独立增量过程, 又是平稳增量的过程. 并且, 对任意 $t > 0, s \geq 0$, 增量 $X_{t+s} - X_s$ 都服从参数为 λt 泊松分布. 反过来, 任何一个满足如上条件的计数过程都是一个泊松过程. 事实上, 如下更强的命题成立.

命题2.1.17. 假设 $\{X_t\}$ 是一个具有独立增量性的计数过程. 如果对任意 $t \geq 0$,

$$\lim_{s \searrow 0} \frac{P(X_{t+s} - X_t \neq 0)}{s} = \lim_{s \searrow 0} \frac{P(X_{t+s} - X_t = 1)}{s} = \lambda,$$

那么, $\{X_t\}$ 是一个参数为 λ 的泊松过程.

下面, 我们考虑指数闹钟响起的时刻 S_1, S_2, \dots , 它们是半直线 $\mathbb{R}_+ = [0, \infty)$ 上的一系列随机的点. 将它们组成的集合记为 Ξ . 即,

$$\Xi := \{S_1, S_2, \dots\}.$$

它也被称为泊松点过程. 我们用 Ξ 重述推论 2.1.5, 便得到下面的推论.

推论2.1.18. 对任意 $0 \leq s < t$, $|\Xi \cap (s, t)|$ 服从参数为 $\lambda(t-s)$ 的泊松分布, 并且对任意 $0 \leq s < t < u < v$, $|\Xi \cap (s, t)|$ 与 $|\Xi \cap (u, v)|$ 相互独立.

进一步, 取与 Ξ 相互独立的泊松点过程 $\hat{\Xi} = \{T_1, T_2, \dots\}$. 令 $\bar{\Xi} = \{-T_1, -T_2, \dots\}$. 那么, 由命题 2.1.10 可推出 $\Xi \cup \hat{\Xi}$ 也满足上面的推论 2.1.18, 其中 $s \geq 0$ 的限制取消. 事实上, 推论 2.1.18 可以推广到更一般的空间, 作为泊松点过程的定义.

假设 S 是非空集合, \mathcal{G} 是其上的 σ 代数, μ 是测度. 假设 Ξ 是 S 的随机子集, 满足:

- (1) 若 $\mu(A) < \infty$, 则 $|\Xi \cap A|$ 服从参数为 $\mu(A)$ 的泊松分布;
- (2) 若 $A \cap B = \emptyset$, 则 $|\Xi \cap A|$ 与 $|\Xi \cap B|$ 相互独立.

那么, Ξ 被称为 S 上的一个泊松点过程, μ 被称为其强度. 例如, 夜空里某一等级的星星可以看成是二维球面上的泊松点过程, 北大勺海里的荷花也是一个泊松点过程.

练习题

1. 假设 $\xi \sim \text{Exp}(\lambda)$. 证明: 对任意 $t, s > 0$,

$$P(\xi - t > s | \xi > t) = P(\xi > s).$$

2. 假设 ξ, η 相互独立, 并且 $\xi \sim \text{Exp}(\lambda_1)$, $\eta \sim \text{Exp}(\lambda_2)$. 证明: (1) $\min\{\xi, \eta\} \sim \text{Exp}(\lambda_1 + \lambda_2)$,
(2) $P(\xi < \eta) = \frac{\lambda_1}{\lambda_1 + \lambda_2}$. (注: 试从此题结论读解泊松过程的叠加.)
3. 假设 $V, \zeta_1, \zeta_2, \dots$ 相互独立, $P(V = k) = (1 - p)^{k-1}p$, $k = 1, 2, \dots$, 并且 $\zeta_n \sim \text{Exp}(\lambda)$,
 $n = 1, 2, \dots$. 令 $\xi = \zeta_1 + \dots + \zeta_V$. 证明: $\xi \sim \text{Exp}(\lambda p)$. (注: 试从此题结论读解泊松过程的细分.)
4. 假设某公交车站有甲、乙两路公共车. 将 1 小时视为单位时间. 将时刻 t 之前甲路车到达的数量记为 X_t , 乙路车到达的数量记为 Y_t . 假设 $\{X_t\}$ 与 $\{Y_t\}$ 为相互独立的泊松过程, 参数分别为 λ_1 与 λ_2 . 求: (1) 下午一点至两点的这一小时里恰好来了三辆公交车的概率,
(2) 某人在车站等甲路车, 在他等甲路车的时间段内, 恰好经过三辆乙路车的概率.
- 5* 在例 2.1.14 中, 假设 ϕ_1 是非负的离散型随机变量. 证明: 推论 2.1.5 和推论 2.1.7 对于复合泊松过程 $\{Y_t\}$ 也成立.
6. 将 1 小时视为单位时间. 将某房产公司在 t 时刻之前发布的售楼信息的数量记为 X_t . 假设 $\{X_t\}$ 是参数为 λ 的泊松过程. 假设每条信息的房价服从 $U(300, 2000)$ (单位: 万元). 假设某先生只读不超过房价不超过 800 万的信息, 他读每条信息所花时间是一个独立的随机变量, 服从 $U(1, 2)$. 将某 30 天中发布的与他有关的信息的数目记为 X , 他读完这些信息所花的总时间记为 Y . 试求: (1) X 的分布, (2) $E \exp(aY)$, 其中 a 为常数. (注: 他读完所有信息的时间有可能超过 30 天, 即, 可能出现 $Y \geq 30$ 的情形.)
7. 证明推论 2.1.7

8. 证明命题 2.1.9.
9. 证明定理 2.1.11 和定理 2.1.12.
- 10* 证明命题 2.1.17 和命题 2.1.18.

北京大学数学学院
试用教材
2022年3月

§2.2 跳过程的构造及其转移概率

一、跳过程的定义

假设 S 为一个非空可数集, 且 $|S| \geq 2$. 我们考虑一个独立指数闹钟模型: 对任意 $i, j \in S$, $j \neq i$, 我们在状态 i 放置一个指向 j 的指数闹钟, 其速率记为 q_{ij} 或 $q_{i,j}$. 其中, q_{ij} 可以为 0, 该情形意味着在 i 处其实并没有指向 j 的闹钟. 假设所有的指数闹钟都是相互独立的. 现在, 假设有一个粒子在 S 中按如下规则跳跃: 如果它位于状态 i , 那么它等待 i 上的闹钟响起, 当其中某个闹钟响起时, 不妨假设是指向状态 j 的指数闹钟响起了, 那么, 该粒子就在闹钟响起的时刻跳跃至状态 j ; 然后, 它在状态 j 等待 j 上的闹钟响. 记

$$q_i := \sum_{j \neq i} q_{ij}, \quad \forall i \in S.$$

为了上述模型能运行, 我们还需要假设 $q_i < \infty$, $i \in S$. 补充定义

$$q_{ii} := -q_i.$$

我们便得到一个实矩阵 $\mathbf{Q} = (q_{ij})_{i,j \in S}$.

定义2.2.1. 假设 $\mathbf{Q} = (q_{ij})_{i,j \in S}$ 为实矩阵, 满足:

$$(1) \quad q_{ij} \geq 0, \forall j \neq i; \quad (2) \quad \sum_{j \in S} q_{ij} = 0, \forall i \in S.$$

那么, 称 \mathbf{Q} 为一个(转移)速率矩阵.

下面总假设 \mathbf{Q} 是一个转移速率矩阵. 根据泊松过程的细分性, 如果 $q_i > 0$, 那么可以认为在每个状态 i 放置了一个参数为 q_i 的大的指数闹钟, 以及一个多面骰子, 它投到第 j 个面的概率为

$$\hat{p}_{ij} := \begin{cases} \frac{q_{ij}}{q_i}, & \forall j \neq i, \\ 0, & j = i. \end{cases}$$

如果 $q_i = 0$, 那么, 事实上状态 i 上并没有放置任何闹钟. 因此粒子在到达状态 i 后就永远停留在状态 i . 此时, 我们称 i 为一个吸收态, 此时定义 $\hat{p}_{ii} = 1$; $\hat{p}_{ij} = 0$, $j \in S \setminus \{i\}$. 于是, 我们得到如下转移矩阵.

$$\hat{\mathbf{P}} := (\hat{p}_{ij})_{S \times S}.$$

从现在开始, 我们总假设 $\{\hat{X}_n : n = 0, 1, 2, \dots\}$ 是一个以 $\hat{\mathbf{P}}$ 为转移矩阵的马氏链, $\xi_0, \xi_1, \xi_2, \dots$ 是与 $\{\hat{X}_n\}$ 相互独立的独立同分布随机变量序列, 其中 $\xi_0 \sim \text{Exp}(1)$.

下面, 我们通过 $\{\hat{X}_n\}$ 和 $\{\xi_n\}$ 定义 X_t , 用它来描述独立指数闹钟模型中粒子在 t 时刻的位置. 令

$$\eta_n = \xi_n / q_{\hat{X}_n}, \quad n = 0, 1, 2, \dots$$

$$S_0 := 0; \quad S_n = \eta_0 + \eta_1 + \cdots + \eta_{n-1}, \quad n = 1, 2, \dots$$

假设

$$P\left(\lim_{n \rightarrow \infty} S_n = \infty \mid \hat{X}_0 = i\right) = 1, \quad \forall i \in S.$$

对 $n = 0, 1, 2, \dots$ 分段定义

$$X_t := \hat{X}_n, \quad \forall t \in [S_n, S_{n+1}). \quad (2.2.1)$$

我们称如上定义的 $\{X_t\}$ 为以 \mathbf{Q} 为速率矩阵的跳过程, 称 $\{\hat{X}_n\}$ 为该跳过程的嵌入链¹, 称 S_n 为第 n 次跳跃时刻, 称 η_1, η_2, \dots 为等待时间. 将从 i 出发的跳过程对应的概率记为 P_i . 如果 $X_0 \sim \mu$, 则对应的概率记为 P_μ . 将它们对应的期望分别记为 E_i 和 E_μ . 根据全概公式, $P_\mu = \sum_{i \in S} \mu_i P_i$. 下面的命题及其推论给出了 $P_i\left(\lim_{n \rightarrow \infty} S_n = \infty\right) = 1, i \in S$ 的若干充分条件, 在下一节中我们还将给出更多的充分条件.

命题2.2.2. 假设 ζ_1, ζ_2, \dots 相互独立, 且 $\zeta_n \sim \text{Exp}(\lambda_n), n = 1, 2, \dots$

$$\begin{aligned} &\text{若 } \sum_{n=0}^{\infty} \frac{1}{\lambda_n} < \infty, \text{ 则 } P\left(\sum_{n=1}^{\infty} \zeta_n < \infty\right) = 1; \\ &\text{若 } \sum_{n=0}^{\infty} \frac{1}{\lambda_n} = \infty, \text{ 则 } P\left(\sum_{n=1}^{\infty} \zeta_n = \infty\right) = 1. \end{aligned}$$

证: 令 $T_n = \zeta_1 + \cdots + \zeta_n$. 注意到 T_n 非负, 并且单调上升到 $T_\infty := \sum_{n=0}^{\infty} \zeta_n$. 由单调收敛定理, $ET_\infty = \lim_{n \rightarrow \infty} ET_n$. 若 $\sum_{n=0}^{\infty} 1/\lambda_n < \infty$, 则 $ET_\infty < \infty$, 故 $P(T_\infty < \infty) = 1$. 如果 $\sum_{n=0}^{\infty} 1/\lambda_n = \infty$, 那么 $\prod_{n=1}^{\infty} (1 + 1/\lambda_n) \geq \sum_{n=1}^{\infty} 1/\lambda_n = \infty$. 注意到若 $\zeta \sim \text{Exp}(\lambda)$, 则

$$Ee^{-\zeta} = \int_0^\infty \lambda e^{-\lambda x} e^{-x} dx = \frac{\lambda}{1+\lambda} = \frac{1}{1+1/\lambda}.$$

于是, 由有界收敛定理,

$$Ee^{-T_\infty} = \prod_{n=1}^{\infty} Ee^{-\zeta_n} = \prod_{n=1}^{\infty} \frac{1}{1+1/\lambda_n} = \frac{1}{\prod_{n=1}^{\infty} (1+1/\lambda_n)} = 0.$$

注意到 $e^{-T_\infty} \geq 0$. 上式表明 $P(e^{-T_\infty} = 0) = 1$, 即 $P(T_\infty = \infty) = 1$. \square

推论2.2.3. 给定 $o \in S$. 假设存在 $\lambda > 0$ 使得

$$q_i \leq \lambda|i|, \quad \forall i \in S \setminus \{o\}, \quad \text{其中 } |i| := \inf \left\{ n \geq 0 : p_{oi}^{(n)} > 0 \right\},$$

那么, $P_o\left(\lim_{n \rightarrow \infty} S_n = \infty\right) = 1$.

¹注: 跳过程 $\{X_t : t \geq 0\}$ 的时间参数是连续的; 嵌入链 $\{\hat{X}_n : n = 0, 1, 2, \dots\}$ 是马氏链, 时间参数是离散的. 在本章中, $\hat{\cdot}$ 符号是嵌入链的专用符号, 例如, $\hat{\tau}_i$ 表示嵌入链 $\{\hat{X}_n\}$ 首次访问状态 i 的时刻.

证: 不妨假设 $q_o > 0$, 否则 $X_t \equiv o$, 结论自然成立. 根据 $|i|$ 的定义, 对任意 $n \geq 1$, 都有 $|Y_n| \leq n$. 于是 $q_{Y_n} \leq \lambda n$, 从而

$$\eta_n = \frac{\xi_n}{q_{Y_n}} \geq \frac{1}{\lambda} \times \frac{\xi_n}{n}, \quad n = 1, 2, \dots$$

将命题 2.2.2 中的 ζ_n 取为 $\frac{\xi_n}{n}$ 便知 $P\left(\sum_{n=1}^{\infty} \frac{\xi_n}{n} = \infty\right) = 1$. 从而, 结论成立. \square

二、例子

例2.2.4 (纯生过程). 假设 $S = \{1, 2, \dots\}$, $\lambda_1, \lambda_2, \dots$ 是一列正数. 令

$$q_i = q_{i,i+1} = \lambda_i, \quad i = 1, 2, \dots$$

将命题 2.2.2 中的 ζ_n 取为 η_{n-1} 便知: $P_i\left(\lim_{n \rightarrow \infty} S_n = \infty\right) = 1, i \in S$ 当且仅当 $\sum_{n=0}^{\infty} \frac{1}{\lambda_n} = \infty$. 此时, 对应的跳过程被称为纯生过程. 注意到若 $\hat{X}_0 = i$, 那么 $\hat{X}_n = i + n, n = 1, 2, \dots$.

例2.2.5 (Yule 过程). 取

$$\lambda_i = \lambda i, \quad i = 1, 2, \dots$$

那么, 对应的纯生过程被称为 **Yule 过程**. Yule 过程有一个非常直观的含义: 假设某系统中有若干细胞, 每个细胞都自带一个独立的指数闹钟, 参数均为 λ . 当闹钟响起时, 该细胞分裂为两个(分裂出的新细胞自带一个独立的指数闹钟). 则系统中的细胞数是一个 Yule 过程. 这是因为当系统中有 i 个细胞时, 根据泊松过程的叠加性, 整个系统中相当于有一个参数为 λi 的大闹钟, 于是, 细胞数从状态 i 变为 $i+1$ 的等待时间服从 $\text{Exp}(\lambda i)$. 在这个观点中, 我们把闹钟放在系统中的每个细胞上, 而不是系统的状态(即, 细胞总数) i 上. 基于此观点, 从 i 出发的 Yule 过程等价于 i 个相互独立的从 1 出发的 Yule 过程之和. 换句话说, 假设 $\{X_t^{(1)}\}, \dots, \{X_t^{(i)}\}$ 是相互独立的从 1 出发的 Yule 过程, 令 $X_t = X_1 + \dots + X_t^{(i)}$, 那么, $\{X_t\}$ 是从 i 出发的 Yule 过程.

例2.2.6 (生灭过程). 假设 $S = \mathbb{Z}_+$,

$$q_{i,i+1} = \beta_i, \quad i = 0, 1, 2, \dots; \quad q_{i,i-1} = \delta_i, \quad i = 1, 2, \dots$$

若 $|j - i| \geq 2$, 则 $q_{ij} = 0$. 则 **Q** 对应的跳过程 $\{X_t\}$ 被称为生灭过程², 因为 X_t 可用于表示某生物群体在 t 时刻的个体数目, 其中, β_i 被称为出生速率, δ_i 被称为死亡速率. 由推论 2.2.3, $P_i\left(\lim_{n \rightarrow \infty} S_n = \infty\right) = 1$ 的一个充分条件是存在常数 $c > 0$, 使得 $\beta_i, \delta_i \leq ci, i = 1, 2, \dots$.

² 在本书中, “生灭链”用于第一章的马氏链, “生灭过程”用于本章的跳过程. 于是, 生灭过程的嵌入链就是生灭链.

例2.2.7 (连续时间的分支过程). 在生灭过程中, 取

$$\beta_i = \beta i, \quad \delta_i = \delta i.$$

那么, 我们可以认为每个个体的寿命是一个服从 $\text{Exp}(\beta + \delta)$ 的独立的随机变量, 个体死亡时产生 0 个或 2 个子代, 概率分别为

$$p_0 = \frac{\delta}{\beta + \delta}, \quad p_2 = \frac{\beta}{\beta + \delta}.$$

因此, 该生灭过程也被称为一个连续时间参数的分支过程.

三、转移概率与转移速率

对任意 $i, j \in S, t \geq 0$, 记

$$p_{ij}(t) := P_i(X_t = j).$$

它即是跳过程的转移概率. 假设粒子从 i 出发, 直到时刻 t 共跳跃了 N_t 次, 即

$$N_t := \sup\{n \geq 0 : S_n \leq t\}.$$

当 $n = 0$ 时,

$$P_i(X_t = j, N_t = 0) = P_i(S_1 > t) \times \mathbf{1}_{\{j=i\}} = e^{-q_i t} \mathbf{1}_{\{j=i\}}.$$

当 $n \geq 1$ 时, 记 $\vec{x} = (x_0, \dots, x_{n-1})$, 它们代表跳过程的等待时间的值; 记

$$\Delta_n := \{\vec{x} = (x_0, \dots, x_{n-1}) : x_r > 0, r = 0, 1, \dots, n-1; x_0 + \dots + x_{n-1} < t\}.$$

取 $i_0 = i$. 对任意 $i_1, \dots, i_n \in S, D \subseteq \Delta_n$,

$$\begin{aligned} & P_i(N_t = n, \hat{X}_0 = i_0, \dots, \hat{X}_n = i_n, (S_0, \dots, S_{n-1}) \in D) \\ &= \prod_{r=0}^{n-1} p_{i_r, i_{r+1}} \times \int_D \left(\prod_{r=0}^{n-1} q_{i_r} e^{-q_{i_r} x_r} \right) e^{-q_{i_n} (t - \sum_{r=0}^{n-1} x_r)} dx_0 \cdots dx_{n-1}. \end{aligned}$$

因此, 我们得到 $p_{ij}(t)$ 的表达式:

$$p_{ij}(t) = e^{-q_i t} \mathbf{1}_{\{j=i\}} + \sum_{n=1}^{\infty} \sum_{i_1, \dots, i_{n-1} \in S} \left(\prod_{r=0}^{n-1} q_{i_r, i_{r+1}} \right) \int_{\Delta_n} \left(\prod_{r=0}^n e^{-q_{i_r} x_r} \right) e^{-q_{i_n} (t - \sum_{r=0}^{n-1} x_r)} dx_0 \cdots dx_n,$$

其中, 在给定 n 时, 取 $i_0 = i, i_n = j$.

命题2.2.8 (马氏性). 对任意 $n \geq 2, 0 \leq t_1 < t_2 < \dots < t_n < t, s \geq 0, i_1, \dots, i_n, i, j \in S$,

$$P(X_{t+s} = j | X_t = i, X_{t_1} = i_1, \dots, X_{t_n} = i_n) = p_{ij}(s).$$

推论2.2.9 (Chapman-Kolmogorov 等式). 对任意 $i, j \in S$, $t, s \geq 0$,

$$p_{ij}(t+s) = \sum_{k \in S} p_{ik}(t)p_{kj}(s).$$

我们已经看到通过速率矩阵 \mathbf{Q} 如何表达转移概率 $p_{ij}(t)$, 下面我们要反过来通过转移概率表达转移速率.

定理2.2.10. 假设 $i, j \in S$ 且 $j \neq i$. 当 t 趋于 0 时,

$$p_{ij}(t) \rightarrow 0, \quad \frac{p_{ij}(t)}{t} \rightarrow q_{ij}; \quad p_{ii}(t) \rightarrow 1, \quad \frac{1 - p_{ii}(t)}{t} \rightarrow q_i.$$

证. 首先, 我们证明转移概率在 $t = 0$ 连续.

$$p_{ij}(t) \leq 1 - p_{ii}(t) = P_i(X_t \neq i) = P_i(S_1 < t) = 1 - e^{-q_i t}.$$

因此, 当 $t \rightarrow 0$ 时, $p_{ii}(t) \rightarrow 1$ 且 $p_{ij}(t) \rightarrow 0$.

然后, 我们证明当 t 趋于 0 时,

$$\frac{1}{t} P_i(N_t = 1, X_t = j) \rightarrow q_{ij}, \quad \frac{1}{t} P_i(N_t = 1) \rightarrow q_i, \quad \frac{1}{t} P_i(N_t \geq 2) \rightarrow 0. \quad (2.2.2)$$

由 (??),

$$\begin{aligned} P_i(N_t = 1, X_t = j) &= q_{ij} q_j \int_{x_1, x_2 \geq 0, x_1 \leq t < x_1 + x_2} e^{-q_i x_1} e^{-q_j x_2} dx_1 dx_2 \\ &= q_{ij} \int_0^t e^{-q_i x_1 - q_j(t-x_1)} dx_1 = q_{ij} e^{-q_j t} \int_0^t e^{(q_j - q_i)x} dx. \end{aligned}$$

从而, 当 $t \rightarrow 0+$ 时, $\frac{1}{t} P_i(N_t = 1, X_t = j) \rightarrow q_{ij}$. 一方面,

$$\begin{aligned} &\liminf_{t \rightarrow 0+} \frac{1}{t} P_i(N_t \geq 1) \geq \liminf_{t \rightarrow 0+} \frac{1}{t} P_i(N_t = 1) \\ &= \liminf_{t \rightarrow 0+} \sum_{j \neq i} \frac{1}{t} P_i(N_t = 1, X_t = j) \geq \sum_{j \neq i} \liminf_{t \rightarrow 0+} \frac{1}{t} P_i(N_t = 1, X_t = j) = \sum_{j \neq i} q_{ij} = q_i; \end{aligned}$$

其中, 第二个 “ \geq ” 是依据 Fatou 引理(附录中的引理4.0.17). 另一方面,

$$\begin{aligned} &\limsup_{t \rightarrow 0+} \frac{1}{t} P_i(N_t = 1) \leq \limsup_{t \rightarrow 0+} \frac{1}{t} P_i(N_t \geq 1) \\ &= \limsup_{t \rightarrow 0+} \frac{1}{t} P_i(S_1 \leq t) = \limsup_{t \rightarrow 0+} \frac{1 - e^{-q_i t}}{t} = q_i. \end{aligned}$$

从而,

$$\lim_{t \rightarrow 0+} \frac{1}{t} P_i(N_t = 1) = \lim_{t \rightarrow 0+} \frac{1}{t} P_i(N_t = 1) = q_i.$$

进一步,

$$\frac{1}{t} P_i(N_t \geq 2) = \frac{1}{t} P_i(N_t \geq 1) - \frac{1}{t} P_i(N_t = 1) \rightarrow 0.$$

综上, (2.2.2) 成立.

最后, 我们求转移概率在 $t = 0$ 的导数. 注意到

$$\begin{aligned} p_{ij}(t) &= P_i(X_t = j) = P_i(N_t = 1, X_t = j) + P_i(N_t \geq 2, X_t = j), \\ 1 - p_{ii}(t) &= P_i(X_t \neq i) = P_i(N_t = 1) + P_i(N_t \geq 2, X_t \neq i). \end{aligned}$$

由 (2.2.2) 知, 当 t 趋于 0 时, $\frac{p_{ij}(t)}{t} \rightarrow q_{ij}$, 且 $\frac{1-p_{ii}(t)}{t} \rightarrow q_i$. \square

命题2.2.11. 对任意 $i, j \in S$, $p_{ij}(t)$ 关于 t 可微, 并且

$$\begin{aligned} p'_{ij}(t) &= \sum_{k \in S} p_{ik}(t)q_{kj}. \quad (\text{Kolmogorov 前进方程}) \\ p'_{ij}(t) &= \sum_{k \in S} q_{ik}p_{kj}(t). \quad (\text{Kolmogorov 后退方程}) \end{aligned}$$

在实际问题中, 我们往往是通过解前进方程或后退方程求出 $p_{ij}(t)$.

例2.2.12 (两个状态的马氏链). 假设 $S = \{0, 1\}$, $q_{01} = \lambda$, $q_{10} = \mu$. 根据 Kolmogorov 后退方程,

$$\begin{cases} p'_{00}(t) = -\lambda p_{00}(t) + \lambda p_{10}(t), \\ p'_{11}(t) = -\mu p_{11}(t) + \mu p_{01}(t). \end{cases}$$

再利用 $p_{10}(t) = 1 - p_{11}(t)$ 与 $p_{01}(t) = 1 - p_{00}(t)$, 解得

$$\begin{aligned} p_{00}(t) &= \frac{\mu + \lambda e^{-(\mu+\lambda)t}}{\mu + \lambda}, & p_{01}(t) &= \frac{\lambda - \lambda e^{-(\mu+\lambda)t}}{\mu + \lambda}. \\ p_{11}(t) &= \frac{\lambda + \mu e^{-(\mu+\lambda)t}}{\mu + \lambda}, & p_{10}(t) &= \frac{\mu - \mu e^{-(\mu+\lambda)t}}{\mu + \lambda}. \end{aligned}$$

注意到

$$\lim_{t \rightarrow \infty} p_{00}(t) = \lim_{t \rightarrow \infty} p_{10}(t) = \frac{\mu}{\mu + \lambda}, \quad \lim_{t \rightarrow \infty} p_{11}(t) = \lim_{t \rightarrow \infty} p_{01}(t) = \frac{\lambda}{\mu + \lambda}.$$

在下一节, 我们将介绍, $\pi_0 = \frac{\mu}{\mu+\lambda}$, $\pi_1 = \frac{\lambda}{\mu+\lambda}$ 其实是该过程的不变分布, 而上式其实是跳过程的强遍历定理, 这些性质与上一章的内容基本上是平行的.

补充知识

习题

- 三人在同一办公室相互独立工作, 各有一部电话, 第 i 个人每次打电话的时间长度服从指数分布, 参数为 μ_i ; 放下电话到下一次打电话之间的间隔时间长度也服从指数分布, 参数为 λ_i . 试用一个跳过程来刻画该办公室内的工作状况, 并写出状态空间和转移速率矩阵.

2. 点灯模型. 在 \mathbb{Z}^1 的每个顶点放置一盏灯, 每盏灯有“亮”和“灭”两个状态, 初始时刻所有灯都是灭的. 某“点灯人”在 \mathbb{Z}^1 上进行连续时间的简单随机游动, 往左和往右跳跃的速率都是 $1/2$. 假设点灯人初始时刻位于原点, 每跳跃一次就会改变目的地处灯的亮或灭状态. 试建立一个跳过程描述此系统, 并写出状态空间和转移速率矩阵.
3. 假设 $\{X_t\}$ 是Yule 过程, 且 $X_0 = 1$. 对任意 $t > 0$, 求 X_t 的分布列.
4. 假设 $\{X_t\}$ 是从1出发的Yule 过程. (1) 证明: 在 $(0, t]$ 上共跳跃 n 次的条件下, 跳跃时刻为 n 个独立同分布随机变量 U_1, \dots, U_n 的顺序统计量, (2) 求 U_1 的分布.
5. 写出关于泊松过程的转移速率矩阵, Kolmogorov 前进方程和后退方程, 并用之求出 $p_{ij}(t)$.
6. 考虑 $\{1, 2, \dots, N\}$ 上的跳过程. 证明: 对任何 $t \geq 0$, $\det(\mathbf{P}(t)) > 0$.
7. (传染模型) 假设一个动物群体中共有 N 个个体, 每两个个体之间以速率 λ 接触, 接触导致某种疾病的传染. 用 X_t 表示 t 时刻被感染的个体数, 设 $X_0 = 1$. 则 $\{X_t\}$ 是一个跳过程. (1) 写出状态空间和转移速率, (2) 将所有个体全被感染的时刻记为 τ . 求 $E\tau_N$.
8. 如果 S 有限, 验证 $\mathbf{P}(t) = e^{t\mathbf{Q}}$ 是方程 $\mathbf{P}'(t) = \mathbf{QP}(t)$ 的解, 满足初值条件 $\mathbf{P}(0) = \mathbf{I}$. 其中

$$e^{t\mathbf{Q}} = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{Q}^n.$$

9. 考虑生灭过程 $\{X_t\}$, 出生速率为 $\beta_i = \lambda i$, $i \geq 0$; 死亡速率为 $\delta_n = \delta i$, $i \geq 1$.

- (1) 记 $h(t) = P_1(X_t = 0)$. 证明: $P_2(X_t = 0) = (h(t))^2$.
- (2) 假设 $\beta \neq \delta$. 证明: $h(t) = \int_0^t e^{-(\lambda+\mu)s} [\mu + \lambda h^2(t-s)] ds$.
- (3) 利用前面结论证明:

$$h(t) = \frac{\mu e^{\mu t} - \mu e^{\lambda t}}{\mu e^{\mu t} - \lambda e^{\lambda t}}.$$

- 10* 利用数学归纳法证明: $p_{ij}^{(n)}(t) := P_i(X_t = j, N_t < n)$ 满足如下递归方程:

$$p_{ij}^{(n+1)}(t) := e^{-q_i t} \delta_{ij} + \int_0^t e^{-q_i(t-s)} \sum_{k \neq i} q_{ik} p_{kj}^{(n)}(s) ds,$$

并利用之证明 $p_{ij}(t)$ 满足如下的积分方程: $\forall i, j \in S$,

$$p_{ij}(t) = e^{-q_i t} \delta_{ij} + \int_0^t e^{-q_i(t-s)} \sum_{k \neq i} q_{ik} p_{kj}(s) ds.$$

11. 对任意 $i, j \in S$, $p_{ij}(t)$ 关于 t 一致连续.

- 12* 假设 f 是 S 上的一个有界函数, 证明: $\frac{d}{dt} \left(\sum_j p_{ij}(t) f(j) \right) = \sum_{j,k} q_{ik} p_{kj}(t) f(j)$, $\forall i$.

§2.3 跳过程的性质

我们沿用上一节的记号. 假设 $\{X_t\}$ 是上一节定义的跳过程, 转移速率矩阵为 $\mathbf{Q} = (q_{ij})_{S \times S}$. 嵌入链 $\{\hat{X}_n\}$ 为一个马氏链, 转移矩阵为 $\hat{\mathbf{P}}$. 假设 $\xi_0, \xi_1, \xi_2, \dots$ 独立同分布, 都服从 $\text{Exp}(1)$, 且它们与 $\{\hat{X}_n\}$ 相互独立. 则跳过程的等待时间为 $\eta_n = q_{\hat{X}_n} \xi_n$, $n = 0, 1, 2, \dots$ 跳跃时刻为 S_1, S_2, \dots . 补充定义 $S_0 = 0$.

不难发现跳过程依次经历的状态其实就是其嵌入链依次经历的状态. 因此, 对任意 $i, j \in S$, 如果在对应的嵌入链中, i 可达 j , 那么我们称在跳过程中 i 可达 j , 仍记为 $i \rightarrow j$. 类似地, 有 i 与 j 互通. 如果对应的嵌入链不可约, 则称跳过程不可约, 也称 \mathbf{Q} 不可约. 以下假设 \mathbf{Q} 不可约, $|S| \geq 2$. 于是, $q_i > 0, i \in S$.

一、常返

现在我们研究跳过程停留在某个给定状态 i 的时间. 将嵌入链 $\{\hat{X}_n\}$ 的访问状态 i 的时刻依次记为 $\hat{T}_{i,0}, \hat{T}_{i,1}, \hat{T}_{i,2}, \dots$, 访问 i 的总次数记为 \hat{V}_i . 记

$$T_{i,r} = S_{\hat{T}_{i,r}}, r = 0, 1, 2, \dots, \zeta_{i,r} := \eta_{\hat{T}_{i,r-1}} = q_i \xi_{\hat{T}_{i,r-1}}, r = 1, 2, \dots$$

如图 2.3 所示, 跳过程依次在如下的时间区间中停留在状态 i :

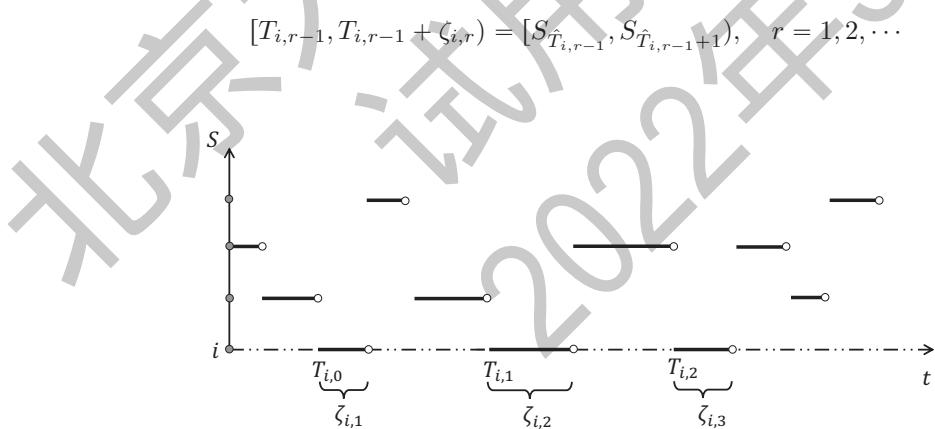


图 2.3 跳过程访问状态 i 的时间区间.

引理2.3.1. 以下两条成立.

- (1) 在 $\{\hat{V}_i = \infty\}$ 发生的条件下, $\zeta_{i,1}, \zeta_{i,2}, \dots$ 独立同分布, 且 $\zeta_{i,1} \sim \text{Exp}(q_i)$.
- (2) 在 $\{\hat{V}_i = n\}$ 发生的条件下, $\zeta_{i,1}, \dots, \zeta_{i,n}$ 独立同分布, 且 $\zeta_{i,1} \sim \text{Exp}(q_i)$.

推论2.3.2. 假设 \mathbf{Q} 不可约, 嵌入链常返, 那么 $P_i \left(\lim_{n \rightarrow \infty} S_n = \infty \right) = 1$.

证: 因为嵌入链常返, 所以 $P_i(\hat{V}_i = \infty) = 1$. 当 $\{\hat{V}_i = \infty\}$ 发生时, $\lim_{n \rightarrow \infty} S_n = \lim_{r \rightarrow \infty} S_{\hat{T}_{i,r}} = \sum_{r=1}^{\infty} \zeta_{i,r} = \infty$. 因此结论成立. \square

若 $\hat{V}_i = \infty$, 则可以找到一列单调上升到无穷的时刻 $s_r := S_{\hat{T}_{i,r}}$, $r = 0, 1, 2, \dots$ 使得 $X_{s_r} = i$; 若 $\hat{V}_i < \infty$, 则嵌入链在 $n := \hat{T}_{i,\hat{V}_i-1} + 1$ 之后不再访问 i , 于是跳过程在 S_n 之后不再访问 i . 因此, $\{\forall t > 0, \exists s > t, \text{s.t. } X_s = i\} = \{\hat{V}_i = \infty\}$.

定义2.3.3. 若

$$P_i(\forall t > 0, \exists s > t, \text{s.t. } X_s = i) = 1,$$

则称 i 为 $\{X_t\}$ 的一个常返态, 或者说 i 是常返的; 否则称 i 为暂态, 或者说 i 是非常返的.

于是, i 是跳过程的常返态当且仅当它是嵌入链的常返态. 令

$$G_{ij} := \int_0^\infty p_{ij}(t) dt.$$

称其为跳过程的格林函数. 将粒子在状态 i 停留的总时间记为 V_i , 即,

$$V_i := \int_0^\infty \mathbf{1}_{\{X_t=i\}} dt = \sum_{r=1}^{\hat{V}_i} \zeta_{i,r}.$$

根据引理 2.3.1, $E_i(V_j | \hat{V}_j = n) = \frac{1}{q_j} \times n$. 进一步, 根据重期望公式,

$$G_{ij} = E_i V_j = \frac{1}{q_j} E_i \hat{V}_j = \frac{1}{q_j} \hat{G}_{ij},$$

其中, \hat{G}_{ij} 为嵌入链 $\{\hat{X}_n\}$ 的格林函数. 我们将以上结论总结为如下命题.

命题2.3.4. i 常返与以下两条等价: (1) i 是 $\hat{\mathbf{P}}$ 的常返态. (2) $G_{ii} = \infty$.

根据命题 2.3.4, 若 i 与 j 互通, 则 i 与 j 或者都常返, 或者都非常返. 进一步, 假设 \mathbf{Q} 不可约, 那么, 或者所有状态都常返(此时, 我们称该跳过程常返, 或者 \mathbf{Q} 常返); 或者所有状态都非常返(此时, 我们称该跳过程非常返, 或者 \mathbf{Q} 非常返).

二、正常返与不变分布

取定状态 i . 记

$$\tau_i := \inf\{t \geq 0 : X_t = j\}, \quad \sigma_i := \inf\{t \geq S_1 : X_t = i\}.$$

假设 $X_0 = i$. 若 $q_i = 0$, 则 $X_t \equiv i$, 这是平凡的情形. 下面假设 $q_i > 0$. 此时, $T_{i,0} = \tau_i = 0$, $T_{i,1} = \sigma_i$. 将 $T_{i,r}$ 理解为第 r 次回访 i 的时刻. 令 $\sigma_{i,r} := T_{i,r} - T_{i,r-1}$, 它表示从第 $r-1$ 回访到第 r 次回访 i 所经历的时间. 在这两次回访之间的时间段中, 粒子在状态 i 停留的时间长度为 $\zeta_{i,r}$. 在引理 2.3.1 中, 我们已经看到, $\zeta_{i,1}, \zeta_{i,2}, \dots$ 独立同分布, 事实上, 根据跳过程的构造还可以推出如下更强的结论.

引理2.3.5. $(\sigma_{i,1}, \zeta_{i,1}), (\sigma_{i,2}, \zeta_{i,2}), \dots$ 为独立同分布的随机向量.

命题2.3.6 (访问频率). 假设 \mathbf{Q} 不可约、常返. 那么对任意初分布 μ ,

$$P_\mu \left(\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbf{1}_{\{X_s=i\}} ds = \frac{1}{q_i E_i \sigma_i} \right) = 1.$$

证: 一方面, 根据引理 2.3.5, $\sigma_{i,1}, \sigma_{i,2}, \dots$ 独立同分布. 于是, 由强大数定律,

$$P_i \left(\lim_{r \rightarrow \infty} \frac{T_{i,r}}{r} = E_i \sigma_i \right) = 1.$$

对任意 $t > 0$, 令

$$M_t := \sup\{r \geq 0 : T_{i,r} \leq t\}.$$

那么, $T_{i,M_t} \leq t < T_{i,M_t+1}$, 从而当 t 趋于无穷时, $\frac{t}{M_t}$ 趋于 $E_i \sigma_i$, 即

$$\frac{M_t}{t} \rightarrow \frac{1}{E_i \sigma_i}.$$

另一方面, $\sum_{r=1}^{M_t} \zeta_{i,r} \leq \int_0^t \mathbf{1}_{\{X_s=i\}} ds \leq \sum_{r=1}^{M_t+1} \zeta_{i,r}$, 由引理 2.3.5, 当 t 趋于无穷时,

$$\frac{1}{M_t} \int_0^t \mathbf{1}_{\{X_s=i\}} ds \rightarrow E_i \zeta_{i,1} = \frac{1}{q_i}.$$

综上, 当 t 趋于无穷时,

$$\frac{1}{t} \int_0^t \mathbf{1}_{\{X_s=i\}} ds = \frac{M_t}{t} \times \frac{1}{M_t} \int_0^t \mathbf{1}_{\{X_s=i\}} ds \rightarrow \frac{1}{E_i \sigma_i} \times \frac{1}{q_i} = \frac{1}{q_i E_i \sigma_i}.$$

若 $X_0 \neq i$, 令 $Y_s := X_{\tau_i+s}$, 则 $\{Y_s\}$ 是从 i 出发的跳过程. 当 $t \geq \tau_i$ 时, $\int_0^t \mathbf{1}_{\{X_s=i\}} ds = \int_{\tau_i}^t \mathbf{1}_{\{X_s=i\}} ds = \int_0^{t-\tau_i} \mathbf{1}_{\{Y_s=i\}} ds$, 于是 $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \mathbf{1}_{\{X_s=i\}} ds = \lim_{t \rightarrow \infty} \frac{t-\tau_i}{t} \times \frac{1}{t-\tau_i} \int_0^{t-\tau_i} \mathbf{1}_{\{Y_s=i\}} ds = 1 \times \frac{1}{q_i E_i \sigma_i}$. 因此, 命题成立. \square

定义2.3.7. 若 i 常返且 $E_i \sigma_i < \infty$, 则称 i 是正常返的; 若 i 常返且 $E_i \sigma_i = \infty$, 则称 i 是零常返的.

给定状态 o . 令

$$\mu_i := E_o \int_0^{\sigma_o} \mathbf{1}_{\{X_t=i\}} dt, \quad \forall i \in S. \quad (2.3.1)$$

注2.3.8. 将嵌入链 $\{\hat{X}_n\}$ 首次回到 o 的时间记为 $\hat{\sigma}_o$. 令

$$\hat{\mu}_i := E_o \sum_{n=0}^{\hat{\sigma}_o-1} \mathbf{1}_{\{\hat{X}_n=j\}},$$

则 $\mu_i = \frac{1}{q_i} \hat{\mu}_i$.

命题2.3.9. \mathbf{Q} 不可约、常返. 那么对任意 $s > 0$,

$$\sum_{j \in S} \mu_j p_{ji}(s) = \mu_i, \quad \forall i \in S.$$

证: 因为 o 常返, 所以 $P_o(\sigma_o < \infty) = 1$. 根据跳过程的构造, 不难发现 $\{X_{\sigma_o+t} : t \geq 0\}$ 也是从 i 出发的以 \mathbf{Q} 为转移速率矩阵的跳过程. 于是, 对任意 $s > 0$,

$$E_o \int_0^s \mathbf{1}_{\{X_t=i\}} dt = E_o \int_0^s \mathbf{1}_{\{X_{\sigma_o+t}=i\}} dt = E_o \int_{\sigma_o}^{\sigma_o+s} \mathbf{1}_{\{X_t=i\}} dt.$$

用 $E_o \int_0^{\sigma_o+s} \mathbf{1}_{\{X_t=i\}} dt$ 分别减去上式两边, 便知

$$E_o \int_s^{\sigma_o+s} \mathbf{1}_{\{X_t=i\}} dt = E_o \int_0^{\sigma_o} \mathbf{1}_{\{X_t=i\}} dt = \mu_i.$$

往证上式的左边即为 $\sum_{j \in S} p_{ji}(s) \mu_j$.

$$\begin{aligned} E_o \int_s^{\sigma_o+s} \mathbf{1}_{\{X_t=i\}} dt &= E_o \int_0^{\sigma_o} \mathbf{1}_{\{X_{t+s}=i\}} dt = E_o \int_0^{\infty} \mathbf{1}_{\{X_{t+s}=i, t < \sigma_o\}} dt \\ &= \sum_{j \in S} \int_0^{\infty} P_o(X_{t+s}=i, X_t=j, t < \sigma_o) dt. \end{aligned}$$

可以证明,

$$P_o(X_{t+s}=i, X_t=j, t < \sigma_o) = p_{ji}(s) P_o(X_t=j, t < \sigma_o). \quad (2.3.2)$$

于是,

$$\begin{aligned} E_o \int_s^{\sigma_o+s} \mathbf{1}_{\{X_t=i\}} dt &= \sum_{j \in S} \int_0^{\infty} p_{ji}(s) P_o(X_t=j, t < \sigma_o) dt \\ &= \sum_{j \in S} p_{ji}(s) E_o \int_0^{\infty} \mathbf{1}_{\{X_t=j, t < \sigma_o\}} dt \\ &= \sum_{j \in S} p_{ji}(s) E_o \int_0^{\sigma_o} \mathbf{1}_{\{X_t=j\}} dt = \sum_{j \in S} p_{ji}(s) \mu_j. \end{aligned}$$

因此, 结论成立. \square

定义2.3.10. 假设 S 上的测度 $\pi = \{\pi_i : i \in S\}$ 满足: 对任意 $t \geq 0$ 都有

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}(t), \quad \forall j \in S,$$

则称 π 为 \mathbf{Q} 的不变测度, 也称为跳过程的不变测度. 进一步, 若 π 还是分布, 则称 π 为 \mathbf{Q} 的不变分布, 也称为跳过程的不变分布.

命题2.3.11. 设 \mathbf{Q} 不可约, 则以下几条等价.

- (1) 所有状态都是正常返的; (2) 存在正常返态; (3) 不变分布存在.

特别地, 若 π 是不变分布, 则

$$\pi_i = \frac{1}{q_i E_i \sigma_i} = \frac{1}{E_o \sigma_o} \mu_i, \quad \forall i \in S.$$

证: (1) \Rightarrow (2): 显然.

(2) \Rightarrow (3): 假设 o 正常返. 那么, $\sum_{i \in S} \mu_i = E_o \sigma_o < \infty$. 由命题 2.3.9, $\mu = \{\mu_i, i \in S\}$ 是不变测度. 将 μ 归一化后即得到不变分布 π , 其表达式为 $\pi_i = \frac{1}{E_o \sigma_o} \mu_i, \forall i \in S$.

(3) \Rightarrow (1). 假设 π 是一个不变分布. 对任意 i , $P_\pi(X_s = i) \equiv \pi_i$. 于是

$$\pi_i = \frac{1}{t} \int_0^t P_\pi(X_s = i) ds, \quad t > 0.$$

令 t 趋于无穷. 由命题 2.3.6 和有界收敛定理,

$$\frac{1}{t} \int_0^t P_\pi(X_s = i) ds = E_\pi \frac{1}{t} \int_0^t \mathbf{1}_{\{X_s=i\}} ds \rightarrow \frac{1}{q_i E_i \sigma_i}.$$

因此, $\pi_i = \frac{1}{q_i E_i \sigma_i}$. 由 \mathbf{Q} 不可约, 不难看出对任意 k 以及任意 $s > 0$, 总有 $p_{ki}(s) > 0$. 因此 $\pi_i > 0$. 从而, $E_i \sigma_i = \frac{1}{\pi_i q_i} < \infty$, 即 i 正常返. \square

利用上述命题, 我们可以通过求不变分布来判断跳过程是否正常返. 虽然在不变分布的定义中涉及到的是转移矩阵 $(p_{ij}(t))_{S \times S}$, 但是对于跳过程而言, 最根本的却是其速率矩阵 \mathbf{Q} . 因此, 我们需要通过速率矩阵 \mathbf{Q} 直接求出不变分布.

推论2.3.12. 假设 \mathbf{Q} 不可约、常返. 那么下列两条成立.

- (1) 若 π 是 S 上的测度, 则 $\pi \mathbf{Q} = 0$ 当且仅当存在常数 c 使得 $\pi_i = c \mu_i, i \in S$.
(2) 若 π 是 S 上的分布, 则 $\pi \mathbf{Q} = 0$ 当且仅当 π 是不变分布.

证: (1) 考虑嵌入链 $\{\hat{X}_n\}$ 的转移概率: $\hat{p}_{ij} = q_{ij}/q_i, j \neq i$. 首先, 当 $j \neq i$ 时, 将 q_{ij} 改写为 $q_i \hat{p}_{ij}$. 那么, $\pi \mathbf{Q} = 0$ 则改写为 $\sum_{j \neq i} \pi_j q_j \hat{p}_{ji} = \pi_i q_i = 0$. 换句话说, $\pi \mathbf{Q} = 0$ 当且仅当 $\pi_i q_i, i \in S$ 是嵌入链的不变测度. 其次, 根据命题 1.7.20, $\pi_i q_i, i \in S$ 是嵌入链的不变测度当且仅当存在常数 c 使得 $\pi_i q_i = c \hat{\mu}_i, i \in S$. 最后, 因为 $\mu_i = \frac{1}{q_i} \hat{\mu}_i, i \in S$. 因此, $\pi_i q_i = c \hat{\mu}_i, i \in S$ 当且仅当 $\pi_i = c \mu_i, i \in S$. 综上, $\pi \mathbf{Q} = 0$ 当且仅当 $\pi_i = c \mu_i, i \in S$.

(2) 如果 $\pi \mathbf{Q} = 0$, 那么 $\pi_i = c \mu_i, i \in S$. 由命题 2.3.9 知 π 是不变测度. 反过来, 如果 π 是不变分布, 由命题 2.3.11, $\pi_i = c \mu_i, i \in S$, 其中 $c = \frac{1}{E_o \sigma_o}$. 于是 $\pi \mathbf{Q} = 0$. 综上, $\pi \mathbf{Q} = 0$ 当且仅当 π 是不变分布. \square

例2.3.13. 取 $S = \mathbb{Z}_+$. 假设 $\{X_t\}$ 是生灭过程, 将其出生速率与死亡速率分别记为 β_i 与 δ_i . 其嵌入链 $\{\hat{X}_n\}$ 是一个生灭链, 将其出生概率与死亡概率分别记为 b_i 与 d_i .

(1) 取

$$b_0 = 1, \quad d_1 = b_1 = \frac{1}{2}; \quad b_i = \frac{i-1}{2i-1}, \quad d_i = \frac{i}{2i-1}, \quad i = 2, 3, \dots$$

则 $\hat{\{X_n\}}$ 零常返. 取 $q_0 = 1, q_i = i, i = 1, 2, \dots$. 则 $\beta_i = ib_i, i = 0, 1, 2, \dots, \delta_i = id_i, i = 1, 2, \dots$. 解方程 $\pi\mathbf{Q} = 0$ 得到 $\pi_0 = \frac{1}{2}c, \pi_1 = c, \pi_2 = \frac{3}{4}c, \pi_i = \frac{2i-1}{i(i-1)}c, i = 3, 4, \dots$, 其中 $c = (\frac{9}{4} + \sum_{i=3}^{\infty} \frac{2i-1}{i(i-1)})^{-1}$. 因为 π 是 \mathbf{Q} 的不变分布, 所以 $\{X_t\}$ 正常返.

(2) 取

$$b_0 = 1; \quad b_i = \frac{1}{3}, \quad d_i = \frac{2}{3}, \quad i = 1, 2, \dots$$

则 $\hat{\{X_n\}}$ 正常返, 其不变分布 μ 的表达式为: $\mu_0 = \frac{1}{4}, \mu_1 = \frac{3}{8}, \mu_i = \frac{3}{2^{i+2}}, i = 2, 3, \dots$. 取 $q_i = \mu_i, i = 0, 1, 2, \dots$. 则 $\pi\mathbf{Q} = 0$ 的解形如 $\pi_i \equiv c$, 其中 c 是常数. 于是 $\{X_t\}$ 没有不变分布, 从而它是零常返的.

虽然跳过程与嵌入链的互通性与常返性是相同的, 但是上面的例题表明这两者的正常返性是不同的.

三、可逆分布

假设 \mathbf{Q} 不可约、正常返, π 为不变分布. 令

$$\tilde{q}_{ij} = \frac{\pi_j q_{ji}}{\pi_i}, \quad \forall i, j \in S; \quad \tilde{\mathbf{Q}} = (\tilde{q}_{ij})_{S \times S}.$$

由 $\pi\mathbf{Q} = 0$ 知, 对任意 $i \in S$, $\sum_{j \neq i} \pi_j q_{ji} = \pi_i q_i$. 于是,

$$\tilde{q}_i := \sum_{j \neq i} \tilde{q}_{ij} = \sum_{j \neq i} \frac{\pi_j q_{ji}}{\pi_i} = q_i = -\tilde{q}_{ii}$$

因此 $\tilde{\mathbf{Q}}$ 仍然是转移速率矩阵. 对任意 $i, j \in S$,

$$\hat{q}_{ij} := \frac{\tilde{q}_{ij}}{q_i} = \frac{\pi_j q_j \hat{p}_{ji}}{\pi_i q_i}.$$

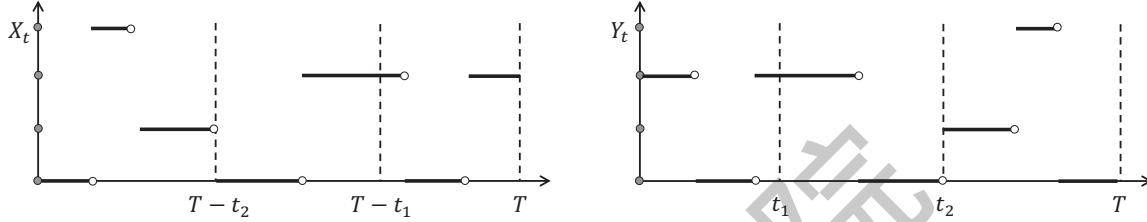
令 $\mu_i = \pi_i q_i, i \in S$. 那么, $\mu_i \hat{q}_{ij} = \mu_j \hat{p}_{ji}, i, j \in S$. 这蕴含着 $\mu_i \hat{q}_{ij}^{(n)} = \mu_j \hat{p}_{ji}^{(n)}$, 特别地, $\hat{q}_{ii}^{(n)} = \hat{p}_{ii}^{(n)}$, $n \geq 1$. 因此 $(\hat{q}_{ij})_{S \times S}$ 常返. 根据推论 2.3.2, 我们可以定义以 $\tilde{\mathbf{Q}}$ 为速率矩阵的跳过程, 记为 $\{Z_t\}$, 并且由命题 2.3.4 知 $\tilde{\mathbf{Q}}$ 常返. 进一步, 不难验证 $\pi \tilde{\mathbf{Q}} = 0$, 因此, π 也是 $\tilde{\mathbf{Q}}$ 的不变分布.

假设 $X_0 \sim \pi$. 对任意 $s > 0$, 将 $\{X_t\}$ 在时刻 s 的左极限记为 X_{s-} , 即, $X_{s-} := \lim_{u \nearrow s} X_u$. 给定 $T > 0$, 令

$$Y_0 = X_T, \quad Y_t = X_{(T-t)-}. \quad (2.3.3)$$

称 $\{Y_t : 0 \leq t \leq T\}$ 为 $\{X_t : 0 \leq t \leq T\}$ 的时间倒逆过程, 简称逆过程³. $\{X_t : 0 \leq t \leq T\}$ 与 $\{Y_t : 0 \leq t \leq T\}$ 的轨道对应关系如图 2.4 所示.

³如上定义是为了保持轨道的右连续性. 注意: 若在时刻 $T - t$, $\{X_u\}$ 从发生了跳跃, 不妨假设从状态 i 跳跃至状态 j . 那么, 根据轨道的右连续型, $X_{T-t} = j$, 但 \tilde{X}_t 却定义为 i .

图 2.4 例如, $Y_{t_1} = X_{T-t_1}$, $Y_{t_2} \neq X_{T-t_2}$.

命题2.3.14. 假设 $Z_0 \sim \pi$, 则 $\{Z_t : 0 \leq t \leq T\}$ 与 $\{Y_t : 0 \leq t \leq T\}$ 同分布.

证: 将 $\{Y_t : 0 \leq t \leq T\}$ 跳跃的总次数记为 N_Y , 依次经历的状态记为 $\hat{Y}_0, \hat{Y}_1, \dots$, 它跳跃的时刻依次记为 T_1, T_2, \dots . 那么, $P(Y_0 = 0, N_Y = 0) = P(\hat{X}_0 = i, N_X = 0) = \pi_i e^{-q_i T} = P(Z_0 = i, N_Z = 0)$, 其中, N_Z 是 $\{Z_t\}$ 在 T 之前跳跃的总次数. 当 $n \geq 1$ 时, 若 $\{N = n\}$ 发生, 则 $T_r = T - S_{n+1-r}$. 对任意 $0 < u_1 < \dots < u_n < u_{n+1} := T$, $0 < \delta_r < u_{r+1} - u_r$, $r = 1, \dots, n$, 令

$$\Delta_Y := \{(y_0, \dots, y_{n-1}) : u_r < t_r < u_{r+1}, r = 1, \dots, n\},$$

$$\Delta_X := \{(x_0, \dots, x_{n-1}) : T - (u_{n+1-r} + \delta_{n+1-r}) < s_r < T - u_{n+1-r}, r = 1, \dots, n\},$$

其中 $s_r = x_0 + \dots + x_{r-1}$, $t_r = y_0 + \dots + y_{r-1}$. 那么,

$$\{N_Y = n, (T_1, \dots, T_n) \in \Delta_Y\} = \{N_X = n, (S_1, \dots, S_n) \in \Delta_X\}.$$

对任意 $i_0, \dots, i_n \in S$,

$$\begin{aligned} & P(\hat{Y}_0 = i_0, \dots, \hat{Y}_n = i_n, N_Y = n, (T_1, \dots, T_n) \in D_Y) \\ &= P(\hat{X}_0 = i_n, \dots, \hat{X}_n = i_0, N_X = n, (S_1, \dots, S_n) \in D_X) \\ &= \pi_{i_n} q_{i_n i_{n-1}} \cdots q_{i_1 i_0} \int_{\Delta_X} \left(\prod_{r=0}^{n-1} e^{-q_{i_{n+1-r}} x_r} \right) e^{-q_{i_0}(T-s_n)} dx_0 \cdots dx_{n-1}. \end{aligned}$$

令 $y_0 = T - s_n$, $y_1 = x_{n-1}, \dots, y_{n-1} = x_1$. 那么, $x_0 = T - (y_0 + \dots + y_{n-1})$, 并且

$$\left(\prod_{r=0}^{n-1} e^{-q_{i_{n+1-r}} x_r} \right) e^{-q_{i_0}(T-s_n)} = \left(\prod_{r=0}^{n-1} e^{-q_{i_r} y_r} \right) e^{-q_{i_n}(T-t_n)}.$$

又因为 $(x_0, \dots, x_{n-1}) \in D_X$ 等价于 $(y_0, \dots, y_{n-1}) \in D_Y$; $q_i = \tilde{q}_i$, $i \in S$; 并且 $\pi_{i_n} q_{i_n i_{n-1}} \cdots q_{i_1 i_0} = \pi_{i_0} \tilde{q}_{i_0 i_1} \cdots \tilde{q}_{i_{n-1} i_n}$, 所以

$$\begin{aligned} & P(\hat{Y}_0 = i_0, \dots, \hat{Y}_n = i_n, N_Y = n, (T_1, \dots, T_n) \in D_Y) \\ &= \pi_{i_0} \tilde{q}_{i_0 i_1} \cdots \tilde{q}_{i_{n-1} i_n} \int_{\Delta_Y} \left(\prod_{r=0}^{n-1} e^{-\tilde{q}_{i_r} y_r} \right) e^{-\tilde{q}_{i_n}(T-t_n)} dy_0 \cdots dy_{n-1}, \end{aligned}$$

这即是 $\{Z_t\}$ 的分布. 因此, 结论成立. \square

定义2.3.15. 若 π 是不变分布, 并且满足如下细致平衡条件:

$$\pi_i q_{ij} = \pi_j q_{ji}, \quad \forall i, j \in S,$$

则称 π 为 \mathbf{Q} 的可逆分布.

若 π 是可逆分布, $X_0 \sim \pi$, 则对任意 $T > 0$, $\{Y_t : 0 \leq t \leq T\}$ 与 $\{Z_t : 0 \leq t \leq T\}$ 同分布. 此时, 称 $\{X_t\}$ 称为可逆过程. 假设 \mathbf{Q} 常返, π 是 S 上的任一分布. 若细致平衡条件成立, 则 $\pi \mathbf{Q} = 0$, 于是 π 是不变分布, 进而是可逆分布.

注2.3.16. 在 \mathbf{Q} 非常返的情形中, 即便 π 是满足细致平衡条件的分布, π 也不是可逆分布, 因为它不是不变分布.

习题

1. 假设 $i \neq j$. 证明: $i \rightarrow j$ 与下列几条均等价. (1) 存在 $n \geq 1$, i_0, \dots, i_n (其中 $i_0 = i, i_n = j$) 使得 $q_{i_0 i_1}, \dots, q_{i_{n-1} i_n} > 0$; (2) $p_{ij}(t) > 0, t > 0$; (3) 存在 $t > 0$ 使得 $p_{ij}(t) > 0$. (注: 对于不可约的跳过程而言, 不存在马氏链中的周期或是非周期的讨论.)
2. 给定 $D \subseteq S$. 令 $\tau_D := \inf\{t \geq 0 : X_t \in D\}$. 记 $x_i = P_i(\tau_D < \infty), i \in S; y_i = E_i \tau_D, i \in S$. 仿照马氏链的首步分析法, 分别写出 $x_i, i \in S$ 与 $y_i, i \in S$ 满足的方程. (注: 事实上, 它们是对应方程的最小非负解.)
3. 在例 2.2.7 中, 假设 $\beta < \delta$. (1) 求: $P_1(\tau_0 < \infty)$. (2) 假设 $\beta < \delta$. 求: $E_1 \tau_0$.
4. 假设状态 i 非常返, $X_0 = i$. 那么, $\{X_t\}$ 在状态*i* 的总停留时间服从指数分布. (提示: 注意到 \hat{V}_i 服从几何分布, 利用引理 2.3.1 和第 §2.1 节习题3.)
5. 证明: i 常返与下列两条等价. (1) 对任意 $\delta > 0$ 均有 $\sum_{n=0}^{\infty} p_{ii}(n\delta) = \infty$. (2) 存在 $\delta > 0$ 使得 $\sum_{n=0}^{\infty} p_{ii}(n\delta) = \infty$.
6. 证明引理 2.3.5.
7. 假设 \mathbf{Q} 不可约. 若存在 S 上的测度 λ 满足 $\lambda \mathbf{Q} = 0$ 且 $\sum_{i \in S} \lambda_i = \infty$, 则不变分布不存在.
8. 假设 \mathbf{Q} 不可约、正常返, π 是不变分布. 证明: 对任意 $i \in S$ 和初分布 μ 都有

$$\mathbf{P}_{\mu} \left(\lim_{t \rightarrow \infty} \sum_{i \in S} \left| \frac{1}{n} \int_0^t \mathbf{1}_{\{X_s=i\}} ds - \pi_i \right| = 0 \right) = 1, \quad \lim_{t \rightarrow \infty} \sum_{i \in S} \left| \frac{1}{t} \int_0^t P_{\mu}(X_s=i) ds - \pi_i \right| = 0.$$

9. 假设 \mathbf{Q} 不可约、正常返, 其不变分布记为 π . 假设 f 是 S 上的函数. 分别在以下两种假设下, 证明: 对任意初分布 μ ,

$$P_\mu \left(\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X_s) ds = \sum_{i \in S} \pi_i f(i) \right) = 1.$$

(1) 假设 f 有界. (2) 假设 $\sum_{i \in S} \pi_i |f(i)| < \infty$.

10. 假设 \mathbf{Q} 不可约、正常返, π 为其不变分布. 证明: $\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j$, $i, j \in S$. (提示: 先证明 $\lim_{n \rightarrow \infty} p_{ij}(n\delta) = \pi_j$, 再利用命题 ??.)

§2.4 排队系统

在日常生活中, 我们经常要排队, 例如排队买火车票或者某场火爆演出的门票. 人们到银行办理业务, 进门后需要取一个号码, 其实就是进入一个排队系统. 在生产调度中, 排队现象更是屡见不鲜, 譬如繁忙的机场常有飞机等候起飞或降落. 这类系统对应着一个数学模型, 称为排队系统, 其要素包括: 顾客的到达时刻, 服务时间, 服务员的个数.

一、M/M/1 模型⁴

将 t 时刻的队列长度记为 X_t , 那么, $\{X_t\}$ 是一个生灭过程, 出生率就是新顾客到达的速度, 记为 λ ; 死亡率就是服务员的服务速率, 记为 α . 即,

$$S = \mathbb{Z}_+ = \{0, 1, 2, \dots\}, \quad \beta_i = \lambda, \quad i = 0, 1, 2, \dots \quad \delta_i = \alpha, \quad i = 1, 2, \dots$$

因为 $q_i, i \in \mathbb{Z}_+$ 有界, 所以该生灭过程非爆炸. 进一步, 该跳过程正常返当且仅当 $\lambda < \alpha$. 下面, 假设 $\lambda < \alpha$. 此时, 不变分布为

$$\pi_i = \left(\frac{\lambda}{\alpha}\right)^i \frac{\alpha - \lambda}{\alpha}, \quad \forall i = 0, 1, 2, \dots \quad (2.4.1)$$

注意到细致平衡条件成立, 因此 π 也是可逆分布.

假设 $X_0 \sim \pi$. 那么, 对任意时刻 t , 都有 $X_t \sim \pi$. 因此, 平均的队列长度总为

$$E_\pi X_t = \sum_{i=0}^{\infty} i \pi_i = \frac{\lambda}{\alpha - \lambda}.$$

由于服务员在队列长度到达 0 时开始休息, 然后, 经过指数时间 $\xi \sim \text{Exp}(\lambda)$ 后, 服务员开始工作, 直到下一次队列长度到达 0 时可以再次开始休息. 因此, 服务员连续工作的时间为 $\sigma_0 - \xi$, 其期望为

$$E_0 \sigma_0 - E_0 S_1 = \frac{1}{q_0 \pi_0} - \frac{1}{q_0} = \frac{1}{\delta - \beta}.$$

最后, 对任意 t , 将进入系统的顾客数和离开系统的顾客数分别记为 I_t 和 L_t , 具体地, 令

$$I_t := |\{0 \leq s \leq t : X_s = X_{s-} + 1\}|, \quad L_t := |\{0 \leq s \leq t : X_s = X_{s-} - 1\}|. \quad (2.4.2)$$

注意到 $\{X_t\}$ 是可逆过程. 因此, 对任意 $T > 0$, $\{L_t : 0 \leq t \leq T\}$ 与 $\{I_t : 0 \leq t \leq T\}$ 同分布. 这表明 $\{L_t : t \geq 0\}$ 是参数为 λ 的泊松过程.

二、M/M/s 模型与 M/M/ ∞ 模型.

⁴ “M/M/1”中的第一个字母“M”表示顾客到达时刻具有无记忆性(Memoryless), 即顾客到达时刻为一个泊松流; 第二个字母“M”表示服务员的服务时间也具有无记忆性, 即他(她)服务一位顾客的时间服从指数分布; 最后的数字“1”表示系统中有一位服务员.

将 t 时刻的队列长度记为 X_t . 那么, $\{X_t\}$ 仍然为一个生灭过程.

在 $M/M/s$ 模型中, 共有 s 个服务员, 因此, 出生率和死亡率如下:

$$\beta_i = \lambda, i = 0, 1, 2, \dots \quad \delta_i = \alpha \times \min\{i, s\}, i = 1, 2, \dots$$

在 $M/M/\infty$ 模型中, 共有无穷多个服务员, 因此, 出生率和死亡率如下:

$$\beta_i = \lambda, i = 0, 1, 2, \dots \quad \delta_i = \alpha \times i, i = 1, 2, \dots$$

在现实生活中, 不可能有无穷多个服务员, 所以我们从另外的角度来理解 $M/M/\infty$ 模型. 例如, 假设在一个参数为 λ 的指数闹钟响起的时刻就有一位游客进入某公园, 每位游客的游玩时间相互独立, 都服从 $Exp(\alpha)$, 游玩结束后自行离去. 那么, 我们就可以将公园中的游客数读解为 $M/M/\infty$ 模型中的队列长度.

根据推论 2.2.3, 这两个模型都非爆炸.

三、有两个服务项目的排队系统

在这个系统中有两个服务项目, 因此两个队列. 假设每个项目中有一名服务员. 新顾客到达后排在第一个队列等待第一项服务, 接受完第一项服务后排在第二个队列等待第二项服务. 假设新顾客到达的速率为 λ , 第一个服务员的服务速率为 α , 第二个服务员的服务速率为 β . 将 t 时刻第一个队列的长度记为 X_t , 第二个队列的长度记为 Y_t . 记 $Z_t = (X_t, Y_t)$. 那么, $\{Z_t\}$ 是一个跳过程, 状态空间为 $\mathbb{Z}_+ \times \mathbb{Z}_+$, 转移速率如下:

$$\begin{cases} q_{(i,j)(i+1,j)} = \lambda, & i, j \geq 0; \\ q_{(i,j)(i-1,j+1)} = \alpha, & i \geq 1, j \geq 0; \\ q_{(i,j)(i,j-1)} = \beta, & i \geq 0, j \geq 1. \end{cases}$$

在此模型中, $\{X_t\}$ 是一个 $M/M/1$ 模型. 假设 $\lambda < \alpha$, 那么, (2.4.1) 定义的 π 是 $\{X_t\}$ 有不变分布. 并且, 若 $X_0 \sim \pi$, 则 $\{X_t\}$ 是可逆过程此时, $\{L_t\}$ 是参数为 λ 的泊松过程, 其中, L_t 表示直到 t 时刻为止, 离开第一个项目的人数, 如 (2.4.2) 中定义. 于是, 第二个队列也变成一个 $M/M/1$ 模型. 当 $\lambda < \beta$ 时, 它有不变分布

$$\mu_i = \left(\frac{\lambda}{\beta}\right)^i \frac{\beta - \lambda}{\beta}, \quad \forall i = 0, 1, 2, \dots$$

事实上, 当 $\beta < \delta, \alpha$ 时, $\pi \times \mu$ 是 $\{Z_t\}$ 的不变分布.

四、其他排队系统

假设队列长度不能超过 s , 将 t 时刻的队列长度记为 X_t . 例如, 假设某投诉处共有 s 台电话, 当 s 台座机全都在通话时, 新电话无法拨入, 那么 X_t 则表示 t 时刻正在通话的电话数目.

$\{X_t\}$ 是一个 $S = \{0, 1, \dots, s\}$ 上的生灭过程, 出生率和死亡率如下

$$\lambda_i = \beta, \quad 0 \leq i \leq s-1; \quad \delta_i = d \times i, \quad 1 \leq i \leq s.$$

练习题

1. 设在 $M/M/1$ 排队系统中顾客到达速率为 λ , 服务速率为 α . 当队伍长度为 n 时新来的顾客以概率 $\frac{n+1}{n+2}$ 加入到等候的队伍, 以概率 $\frac{1}{n+2}$ 离开. 试问: (1) 在什么条件下该系统有平稳分布? (2) 当该排队系统达到平稳时平均队伍长度是多少? 顾客加入等候队列的概率是多少?
2. 试求 $M/M/s$ 和 $M/M/\infty$ 排队系统的平稳分布.
3. 假设某理发店只有一名理发师, 为每个顾客的服务时间服从指数分布, 平均时间 $\mu < 1$. 顾客到达时刻为泊松流, 参数为 1. (a) 如果顾客看见理发师正忙, 调头就走. 这样丢失的潜在顾客占全部到达该店顾客的比例是多大? (b) 如果该店增设 K 把椅子, 顾客来了以后, 若没有其他顾客, 则马上理发; 若理发师正忙但有空椅子, 则留下来等待, 否则调头就走. 增设 K 把椅子之后该店丢失的潜在顾客占全部顾客的比例是多大?

§2.5 接触过程

本节介绍 \mathbb{Z} 上的有限接触过程, 简称接触过程. 设想每个顶点 $i \in \mathbb{Z}$ 上有一个粒子, 它或是健康, 或是被(某种疾病的)感染. 假设 i 上的个体被感染, 那么, 它以速率1 自动恢复健康; 在恢复健康之前, 它以速率 λ 感染 $i + 1$ 位置上的邻粒子, 也以速率 λ 感染 $i - 1$ 位置上的粒子. 这个模型中有一个参数 λ , 它表示传染病的传播强度. 假定受感染粒子的康复以及疾病在邻居之间的传播都是相互独立的. 接触过程研究的是被感染粒子的群体随时间演化的规律. 将时刻 t 受感染的粒子所处顶点组成的集合记为 X_t . 换句话说, 假设在 t 时刻, X_t 中的粒子被感染, $\mathbb{Z} \setminus A_t$ 中的粒子是健康的. 那么, 接触过程是一个跳过程, 其状态空间为

$$S = \{A : A \subset \mathbb{Z}, |A| < \infty\},$$

转移速率如下:

$$\begin{aligned} q_{A, A \setminus \{i\}} &= 1, && \text{若 } i \in A; \\ q_{A, B} &= \lambda |A \cap \{i-1, i+1\}|, && \text{若 } i \notin A; \\ q_{A, B} &= 0, \text{ 其它 } B \neq A. \end{aligned}$$

由推论 2.2.3, $\{X_t\}$ 非爆炸.

在读解接触过程时, 我们当然可以在任意 \mathbb{Z} 的有限子集 A 上放置若干独立的指数闹钟, 但是, 这样做会将接触过程本身的物理含义完全丢失, 从而丧失所有直观. 下面我们介绍一种非常直观而有效的构造方法——Harris的图表示法. 如图 §2.5 所示, 对任意 $i \in \mathbb{Z}$, 考虑二维平面上从 $(i, 0)$ 出发向上画一条射线 $\{(i, t) : t \geq 0\}$, 表示时间的方向. 然后, 考虑三个独立的指数闹钟, 其中, 第一个闹钟的参数为 λ , 在它响起的任意时刻 t , 画一个左箭头“ \leftarrow ”从 (i, t) 指向 $(i-1, t)$; 第二个闹钟的参数也为 λ , 在它响起的任意时刻 t , 画一个右箭头“ \rightarrow ”从 (i, t) 指向 $(i+1, t)$; 第三个闹钟的参数为1, 在它响起的任意时刻 t , 在 (i, t) 画一个“ \times ”. 如果从 $(i, 0)$ 出发, 顺着时间增长方向及其箭头方向, 有一条不经过“ \times ”的通路到达 (j, t) , 那么, 记 $(i, 0) \rightarrow (j, t)$. 假设最初有且仅有 i 位置上的粒子被感染, 即, $X_0 = \{i\}$, 那么, t 时刻被感染的群体就是

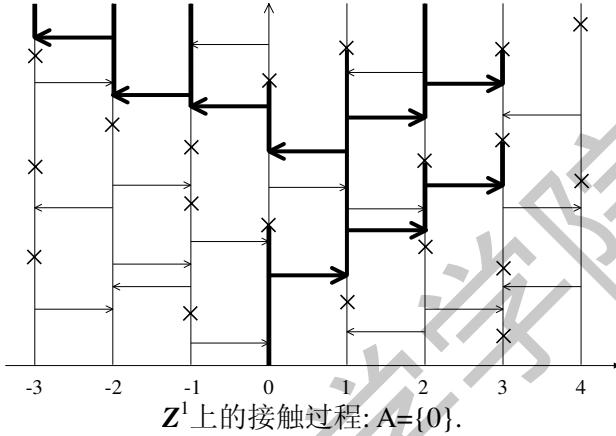
$$X_{i,t} = \{y \in \mathbb{Z} : (i, 0) \rightarrow (y, t)\}.$$

对 \mathbb{Z} 的任意有限子集 A , 假设 $X_0 = A$, 那么, t 时刻被感染的群体就是

$$X_{A,t} = \cup_{i \in A} X_{i,t}.$$

从图表示中, 我们可以看出过程的可叠加性:

$$X_{A \cup B, t} = X_{A,t} \cup X_{B,t}.$$



命题2.5.1 (单调性). 假设 $G_1 \subset G_2$; A, B 分别为 G_1, G_2 中的两个有限的顶点子集, 且 $A \subset B$; $\lambda_1 \leq \lambda_2$. 则存在耦合 $\{(A_t, B_t) : t \geq 0\}$ 使得 $\{A_t : t \geq 0\}$ 是 G_1 上从 A 出发的参数为 λ_1 的接触过程, $\{B_t : t \geq 0\}$ 是 G_2 上从 B 出发的参数为 λ_2 的接触过程, 且 $A_t \subset B_t$.

下面, 我们讨论接触过程的性质. 注意空集 \emptyset 是唯一的吸收态, $S \setminus \{\emptyset\}$ 是互通类. 于是我们要问, 是否从任意 A 出发, 都一定到达 \emptyset . 与分支过程类似, 称 $P_A(\tau_\emptyset < \infty)$ 为(疾病的)灭绝概率, $P_A(\tau_\emptyset = \infty)$ 为(疾病的)存活概率. 注意到, 当参数 λ 变大时, 我们只需要在图上增加箭头即可, 于是, X_t 会变大. 因此, 不难看出, 灭绝概率 $P_A(\tau_\emptyset < \infty)$ 关于参数 λ 单调下降. 由此可以定义接触过程的临界值. 将参数为 λ 的接触过程对应的概率记为 $P^{(\lambda)}$. 令

$$\lambda_c = \inf\{\lambda > 0 : P_{\{\emptyset\}}^{(\lambda)}(\tau_\emptyset = \infty) > 0\}.$$

命题2.5.2. $\lambda_c \geq 1$.

证: 令 $\hat{R}_0 = 0$. 将 $\{X_t\}$ 的嵌入链记为 $\{\hat{X}_n\}$. 假设 $\hat{X}_n = A \neq \emptyset$, 记 $i = \max\{j : j \in A\}$. 令

$$\hat{R}_{n+1} = \begin{cases} \hat{R}_n + 1, & \text{若 } \hat{X}_{n+1} = A \cup i + 1; \\ \hat{R}_n - 1, & \text{若 } \hat{X}_{n+1} = A \setminus i + 1; \\ \hat{R}_n, & \text{其他.} \end{cases}$$

如果 $\{\tau_\emptyset < \infty\}$, 此时, 存在 N 使得 $S_N = \tau$; 如果 $\{\tau_\emptyset = \infty\}$, 那么, 令 $N = \infty$. 对任意 $0 \leq n < N$, 对所有的 $t \in [S_n, S_{n+1})$ 都令 $R_t = R_n$. 那么, R_t 是一个随机游动, $q_{i,i+1} = \lambda$, $q_{i,i-1} = 1$. 并且, 对任意 $t < \tau_\emptyset$,

$$R_t \geq \sup\{i \in \mathbb{Z} : i \in X_t\}.$$

如果 $\lambda < 1$, 那么, 当 t 趋于无穷时, R_t 趋于负无穷, 因此,

$$\sup\{i \in \mathbb{Z} : i \in X_t\} \rightarrow -\infty.$$

根据对称性,

$$\inf\{i \in \mathbb{Z} : i \in X_t\} \rightarrow +\infty.$$

因此, 必然有 $\tau_{\emptyset} < \infty$, 即 $\lambda < \lambda_c$. 由 λ 的任意性, $\lambda_c \geq 1$. □

二、图上的接触过程

考虑区间 $\{1, 2, \dots, N\}$ 上的接触过程. 取初状态 $A_0 = \{1, 2, \dots, N\}$. 那么,

$$\begin{cases} \frac{1}{\log N} E_A \tau_{\emptyset} \rightarrow C, & \lambda < \lambda_c; \\ \frac{1}{N} \log E_A \tau_{\emptyset} \rightarrow C, & \lambda > \lambda_c; \\ C_1 N^2 \leq E_A \tau_{\emptyset} \leq C_2 N^4, & \lambda = \lambda_c. \end{cases}$$

此结论是Durrett等人于1988-1989年得到的[14]. □

在图表示的构造方法中, 我们令 $A_t^A = \cup_{x \in A} A_t^x$. 这其实并不要求 A 有限. 换句话说, 对 V 的任意子集 A , 都可以定义从 A 出发的接触过程. 只不过, 状态空间变得不可数, 因而它不是我们熟悉的跳过程. 虽然 $|A_0| = \infty$ 将意味着 $\tau_{\emptyset} = \infty$, 不过人们仍然可以给过程存活还是灭绝一个合理的定义, 并讨论临界值、相变等问题. 接触过程只是相互作用无穷粒子系统中的一种, 其它的还有伊辛模型、排它过程、选举模型等. 有兴趣的同学可参阅[22].

练习题

1. 证明: 有限图上的接触过程灭绝.
2. 试比较完全图上的接触过程和生灭过程, 并研究该接触过程的性质.
3. 假设 A 为 \mathbb{Z} 的非空、有限子集. 证明:

$$\lambda_c = \inf\{\lambda : P_A^{(\lambda)}(\tau_{\emptyset} = \infty) > 0\} = \sup\{\lambda : P_A^{(\lambda)}(\tau_{\emptyset} < \infty) = 1\}.$$

§2.6 *最小过程

沿用第 §2.2 节的符号: 假设 $\mathbf{Q} = (q_{ij})_{S \times S}$ 是转移矩阵, 且 $q_i > 0, i \in S$. 记 $\hat{p}_{ij} = \frac{q_{ij}}{q_i}, i, j \in S$. $\{\hat{X}_n\}$ 是以 $\hat{\mathbf{P}} = (\hat{p}_{ij})_{S \times S}$ 为转移矩阵的马氏链, $\xi_0, \xi_1, \xi_2, \dots$ 是与 $\{\hat{X}_n\}$ 相互独立的独立同分布随机变量序列, 其中 $\xi_0 \sim \text{Exp}(1)$, $\eta_n = \xi_n/q_{\hat{X}_n}$, $S_n = \eta_0 + \eta_1 + \dots + \eta_{n-1}$. 给定 $n \geq 0$, 对任意 $t \in [S_n, S_{n+1})$ 定义 $X_t := \hat{X}_n$.

一、最小过程

令

$$\tau_\infty := \lim_{n \rightarrow \infty} S_n = \eta_0 + \eta_1 + \dots$$

当事件 $\{\tau_\infty < \infty\}$ 发生时, 为了定义的完备性, 我们添加的一个新状态, 记为 ∂ , 并补充定义 $X_t = \partial, t \in [\tau_\infty, \infty)$. 称这样定义的 $\{X_t\}$ 为 **最小过程**, 称 τ_∞ 为 **爆炸时**. 若 $P_i(\tau_\infty = \infty) = 1, i \in S$, 则我们称该跳过程 **非爆炸**, 也称 \mathbf{Q} **非爆炸**, 此时, 最小过程就是之前定义的跳过程; 否则, 称该跳过程 **爆炸**, 也称 \mathbf{Q} **爆炸**.

例2.6.1. 取 $S = \mathbb{Z}_+$,

$$q_{i,i+1} = 2 \times 3^i, \quad i = 0, 1, 2, \dots, \quad q_{i,i-1} = 3^i, \quad i = 1, 2, \dots$$

该跳过程的嵌入链 $\{\hat{X}_n\}$ 的转移概率为: $p_{0,1} = 1, p_{i,i+1} = \frac{2}{3}, p_{i,i-1} = \frac{1}{3}$, 因此它是非常返的.

令

$$\pi_i = \frac{1}{3} \times \left(\frac{2}{3}\right)^i, \quad i = 0, 1, 2, \dots$$

不难验证 $\pi \mathbf{Q} = 0$. 然而, 它不是不变分布. 事实上, $\{X_t\}$ 是爆炸的.

不难看出 $\tau_\infty = \sum_{i \in S} V_i$, 因此, $E_0 \tau_\infty = \sum_{i=0}^\infty G_{0i}$. 由于从 0 出发的跳过程一定会经历所有状态, 因此, $G_{0i} = G_{ii} = \frac{\hat{G}_{ii}}{q_i}$. 可以验证, 存在 M , 使得 $\hat{G}_{ii} \leq M, i \in S$, 证明留为习题. 于是, $E_0 \tau_\infty \leq M \sum_{i=0}^\infty \frac{1}{q_i} < \infty$. 因此, $P_0(\tau_\infty < \infty) = 1$.

命题2.6.2. 假设 \mathbf{Q} 不可约, π 为 S 上的分布. 则 π 为不变分布当且仅当 $\pi \mathbf{Q} = 0$ 且 \mathbf{Q} 非爆炸.

证: 必要性已证, 因此下面我们只证明充分性. 令 $\tilde{\mathbf{Q}} = (\tilde{q}_{ij})_{S \times S}$, 其中 $\tilde{q}_{ij} = \frac{\pi_j q_{ji}}{\pi_i}, i, j \in S$. 往证 π 为 \mathbf{Q} 的不变分布. 为此, 我们先证明若 $\pi \mathbf{Q} = 0$ 且 \mathbf{Q} 非爆炸, 则 $\tilde{\mathbf{Q}}$ 非爆炸且 π 为 $\tilde{\mathbf{Q}}$ 的不变分布. 推导如下: 记 $\{\tilde{X}_t\}$ 为 $\tilde{\mathbf{Q}}$ 对应的最小过程. 对任意 t , 记 $N_T := \sup\{n \geq 0 : S_n \leq t\}$ 为 $\{X_t\}$ 在 T 之前跳跃的次数, 若 $\tau_\infty \leq T$, 则 $N_T = \infty$. 类似地有 \tilde{N}_T . 与命题 2.3.14 类似地可以证明: 对任意 $n = 0, 1, 2, \dots, i, j \in S, \pi_i P_i(N_T = n, X_T = j) = \pi_j P_j(\tilde{N}_T = n, \tilde{X}_T = i)$. 两边对 n 求和得到 $\pi_i P_i(N_T < \infty, X_T = j) = \pi_j P_j(\tilde{N}_T < \infty, \tilde{X}_T = i)$. 将左右两边对 $i, j \in S$ 求和,

则由**Q** 非爆炸知左边之和为1. 于是, $\sum_{i,j \in S} \pi_j P_j(\tilde{N}_T < \infty, \tilde{X}_T = i) = 1$. 这表明对任意 $j \in S$, $P_j(\tilde{N}_T < \infty) = 1$. 由 T 的任意性, $\tilde{\mathbf{Q}}$ 也非爆炸. 并且对任意 $T > 0$,

$$\sum_{j \in S} \pi_j P_j(\tilde{X}_T = i) = \sum_{j \in S} \pi_i P_i(X_T = j) = \pi_i, \quad \forall i \in S.$$

因此, π 是 $\tilde{\mathbf{Q}}$ 的不变分布.

最后, 因为 $q_{ij} = \frac{\pi_j \tilde{q}_{ji}}{\pi_i}$, $i, j \in S$, 所以可以交换 \mathbf{Q} 与 $\tilde{\mathbf{Q}}$ 的位置. 我们已经证明 $\pi \tilde{\mathbf{Q}} = 0$ 且 $\tilde{\mathbf{Q}}$ 非爆炸, 因此, 根据上面的推理, π 是 \mathbf{Q} 的不变分布. \square

二、连续时间参数的马氏链

假设 \mathbf{Q} 爆炸. 我们还可以在爆炸的时候, 让粒子在 S 中重生. 比如说, 假设 μ 是 S 上的分布. 取一列独立同分布的以 \mathbf{Q} 为转移速率, μ 为初分布的跳过程 $\{X_t^{(1)}\}, \{X_t^{(2)}\}, \dots$. 将 $\{X_t^{(r)}\}$ 的爆炸时记为 $\tau_\infty^{(r)}$, 则 $\tau_\infty^{(1)}, \tau_\infty^{(2)}, \dots$ 独立同分布, 且 $E\tau_\infty^{(1)} > 0$. 令 $T_0 = 0$,

$$T_r = \tau_\infty^{(1)} + \dots + \tau_\infty^{(r)}, \quad r = 1, 2, \dots$$

根据强大数定律, $P\left(\lim_{r \rightarrow \infty} T_r = \infty\right) = 1$. 对任意 $t \geq 0$, 存在唯一的 $r \geq 0$ 使得 $T_r \leq t < T_{r+1}$, 令

$$X_t = X_{t-T_r}^{(r)}.$$

称 $\{X_t\}$ 是可重生的过程. 对任意 $i, j \in S$, 记

$$p_{ij}(t) := P_i(X_t = j).$$

可以验证, $\{X_t\}$ 符合下面的定义.

定义2.6.3. 若对任意 $0 \leq t_1 < \dots < t_n < t < t+s$, 及其 $i_1, \dots, i_n, i, j \in S$,

$$P(X_{t+s} = j | X_t = i, X_{t_1} = i_1, \dots, X_{t_n} = i_n) = P(X_{t+s} = j | X_t = i),$$

则称 $\{X_t\}$ 是一个(连续时间参数的)马氏链. 若 $P(X_{t+s} = j | X_t = i)$ 不依赖于 s , 则称 $\{X_t\}$ 是时齐的, 此时称 $p_{ij}(t) := P(X_{t+s} = j | X_s = i)$ 为过程的转移概率, 称 $\mathbf{P}(t) := (p_{ij}(t))_{i,j \in S}$ 为转移(概率)矩阵.

定义2.6.4. 设 $\{\mathbf{P}(t) : t \geq 0\}$ 是一族转移概率矩阵, 满足

- (1) $\mathbf{P}(0) = \mathbf{I}$;
- (2) Chapman-Kolmogorov方程: $\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s)$;
- (3) 连续性: $\lim_{t \rightarrow 0+} p_{ij}(t) = \mathbf{1}_{\{i=j\}}$,

则称 $\{\mathbf{P}(t) : t \geq 0\}$ 为马氏半群⁵.

⁵ 在泛函分析里, 人们习惯把(无穷维)矩阵看作一个算子. 所以在许多场合, 马氏半群又被称作马氏算子半群.

定义2.6.5. 设 $\mathbf{Q} = (q_{ij})_{i,j \in S}$ 满足:

$$(1) 0 \leq q_{ij} < \infty; (2) 0 \leq q_i = -q_{ii} \leq \infty; (3) \sum_{j \neq i} q_{ij} \leq q_i,$$

则称为 S 上的一个(转移)速率矩阵.

注2.6.6. 定义 2.2.1 中的 \mathbf{Q} 被称为保守的速率矩阵.

命题2.6.7. 假设 $\{\mathbf{P}(t) : t \geq 0\}$ 为马氏半群. 那么, 对任意 $i, j \in S$,

$$q_{ij} := \lim_{t \searrow 0} \frac{p_{ij}(t) - p_{ij}(0)}{t}$$

存在(可以等于正无穷或负无穷). 并且, $\mathbf{Q} = (q_{ij})_{S \times S}$ 是速率矩阵.

证: 我们先证明当 $j \neq i$ 时, $\lim_{t \rightarrow 0^+} p_{ij}(t)/t$ 存在.

取定 $i, j \in S, j \neq i$. 因为 $\lim_{t \rightarrow 0} \mathbf{P}(t) = \mathbf{I}$, 所以对任意 $\epsilon > 0$, 存在 $T > 0$ 使得当 $t < T$ 时, $p_{ii}(t), p_{jj}(t) > 1 - \epsilon$. 这表明 $p_{ij}(t) < \epsilon$. 设 $0 < s < t < T$ 满足 $t > 2s$. 往证

$$\frac{p_{ij}(s)}{s} \leq \frac{p_{ij}(t)}{(t - 2s)(1 - 3\epsilon)}. \quad (2.6.1)$$

在上式中, 先令 $s \rightarrow 0^+$, 得

$$\limsup_{s \rightarrow 0^+} \frac{p_{ij}(s)}{s} \leq \frac{p_{ij}(t)}{t} \frac{1}{1 - 3\epsilon} < \infty,$$

再令 $t \rightarrow 0^+$, 得

$$\limsup_{s \rightarrow 0^+} \frac{p_{ij}(s)}{s} \leq \left(\liminf_{t \rightarrow 0^+} \frac{p_{ij}(t)}{t} \right) \frac{1}{1 - 3\epsilon},$$

最后令 $\epsilon \rightarrow 0$, 我们推出 $q_{ij} := \lim_{t \rightarrow 0^+} p_{ij}(t)/t$ 存在, 非负且有限.

下面我们证明(2.6.1). 取整数 n 使得 $ns \leq t < (n + 1)s$. 首先

$$\begin{aligned} \varepsilon &\geq p_{ij}(t) \geq P_i(X_t = j, \text{ 在 } X_s, \dots, X_{ns} \text{ 中出现过 } j) \\ &= \sum_{k=1}^n P_i(X_s, \dots, X_{(k-1)s} \neq j, X_{ks} = j, X_t = j) \\ &= \sum_{k=1}^n P_i(X_s, \dots, X_{(k-1)s} \neq j, X_{ks} = j) p_{jj}(t - ks) \\ &\geq (1 - \varepsilon) \sum_{k=1}^n P_i(X_s, \dots, X_{(k-1)s} \neq j, X_{ks} = j) \\ &= (1 - \varepsilon) P_i(\text{在 } X_s, \dots, X_{ns} \text{ 中出现过 } j). \end{aligned}$$

这表明

$$P_i(\text{在 } X_s, \dots, X_{ns} \text{ 中出现过 } j) \leq \frac{\varepsilon}{1 - \varepsilon}. \quad (2.6.2)$$

其次,

$$\begin{aligned} p_{ij}(t) &\geq \sum_{k=1}^{n-1} P_i(X_s, \dots, X_{(k-1)s} \neq j, X_{ks} = i, X_{(k+1)s} = j, X_t = j) \\ &= \sum_{k=1}^{n-1} P_i(X_s, \dots, X_{(k-1)s} \neq j; X_{ks} = i) p_{ij}(s) p_{jj}(t - (k+1)s). \end{aligned} \quad (2.6.3)$$

注意到

$$\begin{aligned} &P_i(X_s, \dots, X_{(k-1)s} \neq j; X_{ks} = i) \\ &\geq P_i(X_{ks} = i) - P_i(\text{在 } X_s, \dots, X_{(k-1)s} \text{ 中出现过 } j) \\ &\geq 1 - \varepsilon - P_i(\text{在 } X_s, \dots, X_{ns} \text{ 中出现过 } j) \\ &\geq 1 - \varepsilon - \frac{\varepsilon}{1 - \varepsilon} \geq \frac{1 - 3\varepsilon}{1 - \varepsilon}, \end{aligned}$$

其中第三个不等号就是(2.6.2). 将上式代入(2.6.3)中, 我们推出

$$p_{ij}(t) \geq \sum_{k=1}^{n-1} \frac{1 - 3\varepsilon}{1 - \varepsilon} p_{ij}(s) (1 - \varepsilon) \geq (n-1)(1 - 3\varepsilon) p_{ij}(s).$$

于是,

$$\frac{p_{ij}(s)}{s} \leq \frac{p_{ij}(t)}{(n-1)s(1-3\varepsilon)} \leq \frac{p_{ij}(t)}{(t-2s)(1-3\varepsilon)}.$$

即(2.6.1)成立.

最后, 我们证明 $p'_{ii}(0)$ 存在. 令 $q_i = \liminf_{t \rightarrow 0^+} (1 - p'_{ii}(t))/t$. 若 $q_i = \infty$, 则 $p'_{ii}(0) = -\infty$. 下设 $q_i < \infty$. 往证 $\forall t > 0$, $(1 - p_{ii}(t))/t \leq q_i$, 于是 $\limsup_{t \rightarrow \infty} (1 - p_{ii}(t))/t \leq q_i$. 这表明 $p'_{ii}(0) = -q_i$. 给定 $t > 0$. 对任意 $\varepsilon > 0$, 取 h 充分小使得

$$\frac{1}{h}(1 - p_{ii}(h)) \leq q_i + \frac{\varepsilon}{2}; \quad h(q_i + \frac{\varepsilon}{2}) < 1; \quad 1 - p_{ii}(s) \leq \frac{\varepsilon t}{2}, \quad \forall 0 \leq s \leq h.$$

设 $t = nh + s$, $0 \leq s < h$, 于是

$$\begin{aligned} p_{ii}(t) &\geq p_{ii}(h)^n p_{ii}(s) \geq [1 - h(q_i + \varepsilon/2)]^n (1 - \varepsilon t/2) \\ &\geq 1 - nh(q_i + \varepsilon/2) - \varepsilon t/2 \geq 1 - t(q_i + \varepsilon/2) - \varepsilon t/2 = 1 - t(q_i + \varepsilon), \end{aligned}$$

从而 $(1 - p_{ii}(t))/t \leq q_i + \varepsilon$. 令 $\varepsilon \rightarrow 0$ 知 $(1 - p_{ii}(t))/\varepsilon \leq q_i$.

对 S 的任意有限子集 T , $1 - p_{ii}(t) \geq \sum_{j \in T, j \neq i} p_{ij}(t)$, 所以 $q_i \geq \sum_{j \in T, j \neq i} q_{ij}$. 令 $T \rightarrow S$ 知 $q_i \geq \sum_{j \in S, j \neq i} q_{ij}$. \square

定义2.6.8. 假设 $\{\mathbf{P}(t) : t \geq 0\}$ 是马氏半群. 那么, \mathbf{Q} 称为 $\{\mathbf{P}(t) : t \geq 0\}$ 的转移速率矩阵.

命题2.6.9. 假设 $\{\mathbf{P}(t) : t \geq 0\}$ 是马氏半群, $\mathbf{Q} = (q_{ij})_{S \times S}$ 是其转移速率矩阵. 那么,

$$p'_{ij}(t) = \sum_{k \in S} q_{ik} p_{kj}(t), \quad \forall i, j \in S, \forall t \geq 0 \quad (\text{Kolmogorov后退方程})$$

成立的充要条件是 \mathbf{Q} 保守; 对于给定的 $j \in S$, 若 $q_j < \infty$ 且当 t 趋于0 时, $p_{kj}(t)/t$ 在 $k \in S \setminus \{j\}$ 上一致趋于 q_{kj} , 则,

$$p'_{ij}(t) = \sum_{k \in S} p_{ik}(t) q_{kj}, \quad \forall i \in S \quad (\text{Kolmogorov前进方程}).$$

例2.6.10 (最小过程). 假设 \mathbf{Q} 是保守的转移速率矩阵, 且 \mathbf{Q} 爆炸. 跳过程的转移概率是 $p_{ij}^{(1)}(t) = P_i(X_t^{(1)} = j)$; 可重生的跳过程的转移概率是 $p_{ij}(t) = P_i(X_t = j)$. 记 $\mathbf{P}(t) = (p_{ij}(t))_{S \times S}$. 那么, $\{\mathbf{P}(t)\}$ 的转移速率矩阵就是 \mathbf{Q} . 记 $\mathbf{P}^{(1)}(t) = (p_{ij}^{(1)}(t))_{S \times S}$. 那么, $\{\mathbf{P}(t) : t \geq 0\}$ 与 $\{\mathbf{P}^{(1)}(t) : t \geq 0\}$ 都满足Kolmogorov 后退方程, 并且 $p_{ij}(t) \geq p_{ij}^{(1)}(t)$. 事实上, $\{p_{ij}^{(1)}(t) : i, j \in S, t \geq 0\}$ 是Kolmogorov 前进方程的非负最小解, 也是Kolmogorov 后退方程的非负最小解

练习题

1* 证明: 在例 2.6.1 中, $\hat{G}_{ii} \leq M$, $i \in S$. (提示: 假设 $\{\tilde{X}_n\}$ 是 \mathbb{Z} 上的马氏链, 转移概率为 $p_{i,i+1} = \frac{2}{3}$, $p_{i,i-1} = \frac{1}{3}$, $i \in \mathbb{Z}$, 格林函数为 \hat{G}_{ij} , 那么, 可取 $M = \hat{G}_{00}$.)

北京大学数学学院
试用教材
2022年3月

第三章 布朗运动

§3.1 高斯分布与高斯过程

设 X_1, \dots, X_n 是取值于 \mathbb{R} 的随机变量, 记 $\mathbf{X} = (X_1, \dots, X_n)^T$. 若 \mathbf{X} 是连续型随机向量, 并且联合密度形如

$$\frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{m})^T \Sigma^{-1} (\vec{x} - \vec{m}) \right\}$$

其中, $\vec{x} = (x_1, \dots, x_n)^T \in \mathbb{R}^n$, $\vec{m} = (m_1, \dots, m_n) \in \mathbb{R}^n$, $\Sigma = (\sigma_{ij})_{n \times n}$ 是正定矩阵, 则 \mathbf{X} 服从 n 维正态分布, 记为 $\mathbf{X} \sim N(\vec{m}, \Sigma)$. 记 $\vec{t} = (t_1, \dots, t_n)^T \in \mathbb{R}^n$, 则 $N(\vec{m}, \Sigma)$ 对应的特征函数具有如下形式:

$$f_{\mathbf{X}}(\vec{t}) := E e^{\sqrt{-1}(t_1 X_1 + \dots + t_n X_n)} = \exp \left\{ \sqrt{-1} \vec{m}^T \vec{t} + \frac{1}{2} \vec{t}^T \Sigma \vec{t} \right\}.$$

一般地, 假设 Σ 是 $n \times n$ 半正定矩阵. 若 \mathbf{X} 的特征函数具有上面表达式, 则称 \mathbf{X} 为一个 n 维高斯向量, 或 \mathbf{X} 服从 n 维高斯分布, 记为 $\mathbf{X} \sim N(\vec{m}, \Sigma)$.

假设 $\mathbf{X} \sim N(\vec{m}, \Sigma)$. 不难验证,

$$E X_i = m_i, \quad \sigma_{ij} = \text{Cov}(X_i, X_j) = E(X_i - m_i)(X_j - m_j), \quad i, j = 1, \dots, n.$$

因此, 对于高斯向量, 给出其数字特征期望和协方差矩阵, 其联合分布就完全确定了.

定义3.1.1. 假设 $\{X_\alpha : \alpha \in I\}$ 是一族取值于 \mathbb{R} 的随机变量. 若对任意 $n \geq 1$, $\alpha_1, \dots, \alpha_n \in I$, $(X_{\alpha_1}, \dots, X_{\alpha_n})$ 都服从 n 维高斯分布, 则称 $\{X_\alpha : \alpha \in I\}$ 为一个高斯系或高斯过程.

假设 I_1, \dots, I_n 是互不相交的指标集, 对 $r = 1, \dots, n$, $\mathbf{X}_r := \{X_\alpha : \alpha \in I_r\}$ 是一族取值于 \mathbb{R} 的随机变量. 记

$$\mathbf{X} = \bigcup_{r=1}^n \mathbf{X}_r = \left\{ X_\alpha : \alpha \in \bigcup_{r=1}^n I_r \right\}.$$

命题3.1.2. 假设 \mathbf{X} 是高斯系. 若

$$\text{Cov}(X_\alpha, X_\beta) = 0, \quad \forall r \neq s, \alpha \in I_r, \beta \in I_s,$$

则 $\mathbf{X}_1, \dots, \mathbf{X}_n$ 相互独立.

练习题.

- 假设 $\mathbf{X} = (X_1, \dots, X_n)^T$ 服从 n 维高斯分布. 证明: (1) 存在服从 n 维正态分布的随机向量 $\mathbf{Z} = (Z_1, \dots, Z_n)^T$ 和 $n \times n$ 矩阵 \mathbf{M} 使得 $\mathbf{X} = \mathbf{M}\mathbf{Z}$. (2) 对任意 $m \times n$ 矩阵 \mathbf{M} , \mathbf{MX} 都服从 m 维高斯分布.

2. 假设 $\{X_\alpha : \alpha \in I\}$ 是一个高斯系, J 是一个指标集. 假设对任意 $\beta \in J$, 存在 $n \geq 1$, $\alpha_1, \dots, \alpha_n \in I$ 以及 $a_1, \dots, a_n \in \mathbb{R}$ 使得 $Y_\beta = a_1 X_{\alpha_1} + \dots + a_n X_{\alpha_n}$. 证明: $\{Y_\beta, \beta \in J\}$ 也是高斯系.
3. 证明命题 3.1.2.
4. 假设 $\mathbf{X}_1, \mathbf{X}_2, \dots$ 是一列 d 维高斯向量, 且对任意 $\mathbf{t} \in \mathbb{R}^d$, 它们的特征函数点点收敛, 即 $f(\vec{t}) = \lim_{n \rightarrow \infty} f_{\mathbf{X}_n}(\vec{t})$ 存在且有限. 证明: $f(\vec{t})$ 仍然是 d 为高斯向量的特征函数.

§3.2 布朗运动的定义与Levi构造

假设对任意 $t \geq 0$, X_t 是取值于 \mathbb{R} 的随机变量. 我们也称 $\{X_t : t \geq 0\}$ 为一个取值于 \mathbb{R} 的随机过程, 简记为 $\{X_t\}$.

定义3.2.1. 设 $\{X_t : t \geq 0\}$ 是一个随机过程, $X_0 = 0$. 若对任意 $t \geq 0, s > 0$ 都有 $X_{t+s} - X_t \stackrel{d}{=} X_s$, 则称 $\{X_t\}$ 为一个平稳增量过程; 若对任意 $0 < t_1 < \dots < t_n$ 都有 $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ 相互独立, 则称 $\{X_t\}$ 是一个独立增量过程.

定义3.2.2. 设 $\{B_t : t \geq 0\}$ 是一个随机过程, $B_0 = 0$. 如果它满足以下三条,

- (1) $B_{t+s} - B_t \sim N(0, s), \forall t \geq 0, s > 0$.
- (2) 对任意 $0 < t_1 < \dots < t_n$, $B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ 相互独立;
- (3) $P(B_t \text{ 关于 } t \text{ 连续}) = 1$,

则称 $\{B_t : t \geq 0\}$ 为(一维)标准布朗运动.

注3.2.3. 我将标准布朗运动对应的概率记为 P_0 . 对任意 $x \in \mathbb{R}$, 称 $\{x + B_t\}$ 为从 x 出发的标准布朗运动, 对应的概率记为 P_x . 一般地, 也称 $\{x + \sigma B_t\}$ 为从 x 出发的布朗运动. 不过, 在下文中, 如果不特别声明, “布朗运动”都特指从 0 出发的标准布朗运动.

注3.2.4. 定义 3.2.2 中的(1)和(2) 表明布朗运动是一个独立、平稳增量的过程, 且它在长度为 s 的区间上的增量服从 $N(0, s)$. 这给出了布朗运动的任意有限维联合分布, 即, 它们刻画的是随机向量的分布. 而(3) 刻画的则是轨道的分析性质, 称其为轨道连续性.

命题3.2.5. 对任意 $0 < t_1 < t_2 < \dots < t_n$, $(B_{t_1}, B_{t_2}, \dots, B_{t_n})$ 服从 n 维正态分布, 其联合分布密度为

$$p_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = \prod_{k=1}^n p_{t_k - t_{k-1}}(x_{k-1}, x_k), \quad (3.2.1)$$

其中

$$p_t(x, y) = \frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{(y-x)^2}{2\sigma^2 t}}. \quad (3.2.2)$$

证: 将 $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}})$ 简记为 $\vec{Y} = (Y_1, \dots, Y_n)$. 由定义 3.2.2 中的(1)与(2), 的联合密度为 $p_{\vec{Y}}(y_1, \dots, y_n) = \prod_{k=1}^n p_{t_k - t_{k-1}}(y_k)$. 注意到 $(B_{t_1}, \dots, B_{t_n})$ 是 \vec{Y} 的线性变换, 且行列式为1, 便不难看出 (3.2.1) 成立. \square

上面的证明只是重复了《概率论》课程中一个重要结论: 服从高维正态分布的随机向量经过非退化的线性变换后仍然服从高维正态分布. 为了将 $B_0 = 0$ 的这一特殊的退化情形也包含进来, 也为了避免每次都验证“非退化”的这个条件, 我们将“正态分布”放宽为“高斯分布”.

定义3.2.6. 如果对任意整数 $n \geq 1$ 与任意 $t_1, \dots, t_n \geq 0$, $(X_{t_1}, \dots, X_{t_n})$ 服从高斯分布. 那么, 称 $\{X_t : t \geq 0\}$ 为一个高斯过程.

注意到高维正态分布的联合密度的具体表达式完全由其数字特征(期望和协方差矩阵)决定, 因此, 定义 3.2.2 中的(1) 和(2) 可转化为验证 $\{B_t\}$ 是高斯过程, 并刻画 $(B_{t_1}, \dots, B_{t_n})$ 的期望和协方差矩阵: 对所有的 t 都有 $EB_t = 0$; 当 $s \geq t \geq 0$ 时,

$$EB_t B_s = EB_t (B_t + (B_s - B_t)) = EB_t^2 + EB_t (B_s - B_t) = t.$$

命题3.2.7. 假设 $\{X_t : t \geq 0\}$ 是高斯过程, 并且

$$EX_t = 0, \forall t \geq 0; \quad \text{Cov}(X_t, X_s) = t, \forall s \geq t \geq 0.$$

那么, $\{X_t\}$ 满足定义 3.2.2中的(1) 与(2). 进一步, 若 $\{X_t\}$ 还具有轨道连续性, 那么 $\{X_t\}$ 是标准布朗运动.

例3.2.8 (尺度变换性质(scaling property)). 假设 $\{B_t : t \geq 0\}$ 是标准布朗运动, $c > 0$. 令

$$X_t = \frac{1}{c} B_{c^2 t}, \quad t \geq 0.$$

那么, $\{X_t : t \geq 0\}$ 也是标准布朗运动, 理由如下: 显然 $X_0 = 0$, $\{X_t\}$ 的有限维分布都是正态的, 并且对任意 $t \geq 0$, $EX_t = 0$. 下面我们计算协方差: 对任意 $s \geq t \geq 0$,

$$EX_t X_s = E\left(\frac{1}{c} B_{c^2 t} \times \frac{1}{c} B_{c^2 s}\right) = \frac{1}{c^2} \times EB_{c^2 t} B_{c^2 s} = \frac{1}{c^2} \times c^2 t = t.$$

进一步, $\{X_t\}$ 也具有轨道连续性, 因为作为时间 t 的函数, X_t 是 B_t 与连续函数的复合. 从而, $\{X_t\}$ 是标准布朗运动.

假设 $d \geq 2$, $\{B_t^{(1)} : t \geq 0\}, \{B_t^{(2)} : t \geq 0\}, \dots, \{B_t^{(d)} : t \geq 0\}$ 是 d 个相互独立的标准布朗运动. 记

$$\vec{B}_t = (B_t^{(1)}, B_t^{(2)}, \dots, B_t^{(d)})^T,$$

其中上标 T 表示矩阵的转置.

定义3.2.9. 称 $\{\vec{B}_t : t \geq 0\}$ 为一个 d 维标准布朗运动.

命题3.2.10. 假设 $\{\vec{B}_t : t \geq 0\}$ 是一个 d 维标准布朗运动, $\mathbf{O} = (o_{ij})_{d \times d}$ 是一个 n 维正交矩阵, 则 $\{\mathbf{O}\vec{B}_t : t \geq 0\}$ 仍然是一个 d 维标准布朗运动.

证: 记 $\vec{X}_t = (X_t^{(1)}, X_t^{(2)}, \dots, X_t^{(d)})^T = \mathbf{O}\vec{B}_t$. 注意到 $\{X_t^{(i)} : i = 1, \dots, d; t \geq 0\}$ 是轨道连续的, 其任何有限维分布都是正态分布, 且对任意 $t \geq 0$, $EX_t^{(i)} = 0$. 因而, 我们只需计算所有的协方差. 对任意 $1 \leq i, j \leq d$, $t > s \geq 0$,

$$\begin{aligned} EX_t^{(i)} X_s^{(j)} &= E \left(\sum_{k=1}^d o_{ik} B_t^{(k)} \right) \left(\sum_{l=1}^d o_{jl} B_s^{(l)} \right) = \sum_{k=1}^d \sum_{l=1}^d o_{ik} o_{jl} EB_t^{(k)} B_s^{(l)} \\ &= \sum_{k=1}^d \sum_{l=1}^d o_{ik} o_{jl} \times s \times \mathbf{1}_{\{k=l\}} = s \sum_{k=1}^d o_{ik} o_{jk} = s \times \mathbf{1}_{\{i=j\}}. \end{aligned}$$

对任意 i , 在上式中取 $j = i$ 便知 $\{X_t^{(i)} : t \geq 0\}$ 是一个标准布朗运动. 注意到高维高斯分布的协方差为0 等价于相互独立, 我们有 $\{X_t^{(1)}\}, \{X_t^{(2)}\}, \dots, \{X_t^{(n)}\}$ 相互独立. 因此, $\{\vec{X}_t\}$ 仍然是一个 n 维标准布朗运动. \square

布朗运动的Levi构造. 研究对象的存在性是数学理论中要处理的一个基本问题. 否则, 后续的所有研究都有可能毫无意义. 在第一章, 我们用独立同分布的均匀分布序列构造马氏链, 在第二章, 我们用马氏链和指数分布序列构造跳过程. 现在, 我们给出布朗运动的构造.

引理3.2.11. 假设 $X \sim N(0, \sigma^2)$. 若 \tilde{X} 与 X 独立同分布, 那么,

$$Y := \frac{1}{2}(X + \tilde{X}), \quad Z = \frac{1}{2}(X - \tilde{X})$$

都服从 $N(0, \sigma^2/2)$, 并且 Y 与 Z 相互独立.

注意到 $X = Y + Z$, 因此我们总可以利用上述引理, 将一个零均值的正态变量分解为两个独立同分布的正态变量之和.

下面, 我们构造 $\{B_t : 0 \leq t \leq 1\}$, 方法是利用引理 3.2.11 进行插值, 递归地给出所有二分点上的值. 具体地, 令 $B_0 = 0$, 任取 $\xi = \xi_{0,1} \sim N(0, 1)$, 令 $B_1 := \xi_{0,1}$. 第一步, 取 $\tilde{\xi} = \tilde{\xi}_{0,1}$ 与 $\xi_{0,1}$ 独立同分布, 记

$$\xi_{1,1} = \frac{1}{2}(\xi + \tilde{\xi}), \quad \xi_{1,2} = \frac{1}{2}(\xi - \tilde{\xi}).$$

令 $B_{1/2} := \xi_{1,1}$, 于是, 自然有 $B_1 - B_{1/2} = \xi_{1,2}$. 假设经过 n 步后, 我们得到 $\xi_{n,i}$, $B_{i/2^n}$, $i = 1, \dots, 2^n$, 使得 $\xi_{n,i} = B_{i/2^n} - B_{(i-1)/2^n}$, $1 \leq i \leq 2^n$ 独立同分布, 都服从 $N(0, 2^{-n})$. 那么, 在第 $n+1$ 步, 我们取独立于 $\xi_{n,i}$, $1 \leq i \leq 2^n$ 的独立同分布随机变量族 $\tilde{\xi}_{n,i}$, $1 \leq i \leq 2^n$, 令

$$\xi_{n+1,2i-1} = \frac{1}{2}(\xi_{n,i} + \tilde{\xi}_{n,i}), \quad \xi_{n+1,2i} = \frac{1}{2}(\xi_{n,i} - \tilde{\xi}_{n,i}),$$

并令 $B_{i/2^{n+1}} = \xi_{n+1,1} + \cdots + \xi_{n+1,i}$, $i = 1, \dots, 2^{n+1}$. 最终, 我们得到 $[0, 1]$ 上的所有二分点上的随机变量 $B(\frac{i}{2^n})$, $\forall n \geq 1$, $i = 0, 1, \dots, 2^n$. 最后, 我们需说明, 这些二分点上的随机变量几乎必然确定了一条关于 t 连续的随机轨道, 它就是我们要找的布朗运动. 具体地, 对任意 n , 取 $B_t^{(n)}$ 使得它在 $t = \frac{i}{2^n}$ 时取 $B(\frac{i}{2^n})$, $0 \leq i \leq 2^n$, 在 $[\frac{i-1}{2^n}, \frac{i}{2^n}]$ 上是 t 的线性函数, $1 \leq i \leq 2^n$. 即, 对 $i = 1, \dots, 2^n$, $t \in [(i-1)/2^n, i/2^n]$, 令

$$B^{(n)}(t) := B_{(i-1)/2^n} + 2^n \left(t - \frac{i-1}{2^n} \right) \xi_{n,i}.$$

在区间 $[\frac{i-1}{2^n}, \frac{i}{2^n}]$ 中, 比较 $B_t^{(n)}$ 与 $B_t^{(n+1)}$ 的值, 我们发现误差最大的地方出现在区间中点, 即

$$\max_{(i-1)/2^n \leq t \leq i/2^n} |B_t^{(n+1)} - B_t^{(n)}| = |B_{(2i-1)/2^n}^{(n+1)} - B_{(2i-1)/2^n}^{(n)}| = \frac{1}{2} |\tilde{\xi}_{n,i}|.$$

这表明

$$\max_{0 \leq t \leq 1} |B_t^{(n+1)} - B_t^{(n)}| = \frac{1}{2} \times \max_{1 \leq i \leq 2^n} |\tilde{\xi}_{n,i}|.$$

令

$$A_n := \left\{ \max_{1 \leq i \leq 2^n} |\tilde{\xi}_{n,i}| \geq 2^{-n/4} \sqrt{n} \right\}.$$

给定 $n \geq 1$, 注意到 $\tilde{\xi}_{n,i}$, $1 \leq i \leq 2^n$ 独立同分布, 都服从 $N(0, 2^{-n})$. 我们推出

$$P(A_n) \leq 2^n P(|\tilde{\xi}_{n,1}| \geq 2^{-n/4} \sqrt{n}) = 2^n P(|Z| \geq 2^{n/4} \sqrt{n}) \leq 2^n \frac{E|Z|^4}{2^n n^2} = \frac{EZ^4}{n^2}.$$

其中 $Z = 2^{n/2} \tilde{\xi}_{n,1} \sim N(0, 1)$. 因此, $\sum_{n=1}^{\infty} P(A_n) < \infty$. 由Borel-Cantelli 引理, A_n 发生无穷多次的概率为0, 即, $P(\cup_{N=1}^{\infty} \cap_{n=N}^{\infty} A_n^c) = 1$. 在事件 $\cap_{n=N}^{\infty} A_n^c$ 上, 当 $n \geq N$ 时,

$$\max_{0 \leq t \leq 1} |B_t^{(n+1)} - B_t^{(n)}| \leq \frac{1}{2} \times 2^{-n/4} \sqrt{n}.$$

这表明 $\sum_{n=1}^{\infty} \max_{0 \leq t \leq 1} |B_t^{(n+1)} - B_t^{(n)}| < \infty$, 这蕴含着 $B_t^{(n)}$, $t \in [0, 1]$ 一致收敛到某连续函数 $B(t)$, $t \in [0, 1]$. 令 $N \rightarrow \infty$ 便知

$$P\left(B_t^{(n)} \xrightarrow{t \in [0,1]} B_t\right) = 1.$$

最后, 利用上述办法, 构造一列独立同分布的 $[0, 1]$ 上的随机连续函数, 将它们拼接起来便得到 $B(t)$, $t \geq 0$. 不难验证, 若 t, s, t_1, \dots, t_n 都是二分点, 那么定义 3.2.2 中的(1) 和(2) 成立. 进一步, 利用 B_t 关于 t 连续, 可以证明定义 3.2.2 中的(1) 和(2) 成立. 于是, $\{B_t\}$ 是一个标准布朗运动.

补充知识

1. 布朗运动的起源

1827 年植物学家布朗(Robert Brown)向英国皇家学会报告了他在显微镜下观察到的奇怪现象: 漂浮在液体表面的花粉(中的)大颗粒在毫无规律的随机移动. 后来此现象被命名为布朗运动(Brownian Motion). 十九世纪末分子理论正在形成, 人们认识到花粉大颗粒运动是因为它受到液体分子的不断碰撞. 当时的实验条件并不能直接观察分子, 人们可以通过观察布朗运动来检验有关分子的假设是否正确. 1905年, 瑞士专利局的年轻职员爱因斯坦发表了5篇文章, 其中一篇是关于狭义相对论的, 一篇则给出了布朗运动的物理解释. 此前, 法国的Louis Bachelier于1900年在著名数学家庞加莱指导下完成了他的博士论文, 他把股价变化解释为布朗运动, 分析了布朗运动的一些特性. 或许是题目的数学味不够浓厚, 他的文章当时并没有在数学界产生应有的影响, 而是很多年后才被人注意到. 今天, 世界性的金融数学学会就以他的名字命名. 1923年, 在MIT任讲师的维纳(Norbert Wiener)给出了布朗运动的严格数学基础. 他提出了一个概率空间, 用来定义布朗运动, 后人称之为维纳空间. 这项研究工作恐怕是美国本土产生的第一个具有世界水平的数学研究成果. 二次世界大战期间, 日本学者角谷静夫(Shizuo Kakutani)在饥饿困顿中发现了布朗运动与牛顿位势之间的联系, 他的同事伊藤清(Kiyosi Itô)则提出有关布朗运动的一套计算理论, 被称为Itô积分. 之后, 布朗运动的研究者纷至沓来, 有关文献汗牛充栋, 尤以法国Paul Lévy的工作最为丰富而深刻.

爱因斯坦的解释: 花粉大颗粒受到了液体分子的不断碰撞, 而液体分子本身在做不规则的运动. 将花粉大颗粒经过时间 t 之后的位移记为 X_t , 则 $X_0 = 0$; X_t 各向同性; $X_{t+s} - X_t$ 与 X_t 相互独立, 且与 X_s 具有相同的分布; 并且, X_t 关于 t 连续. 假设 X_t 的方差均存在, 那么,

$$EX_{t+s}^2 = \text{Var}((X_{t+s} - X_t) + X_t) = \text{Var}(X_{t+s} - X_t) + \text{Var}(X_t) = EX_s^2 + EX_t^2.$$

在连续性假设下,

$$EX_t^2 = \sigma^2 t,$$

其中 σ^2 为常数. 这个等式称为爱因斯坦方程. 事实上, $\{X_t/\sigma\}$ 为标准布朗运动.

2. 维纳(Wiener)的贡献.

取值于 \mathbb{R} 的随机过程的轨道空间为:

$$\Omega = \mathbb{R}^{[0, \infty)} = \{\omega : \omega \text{ 为 } [0, \infty) \text{ 到 } \mathbb{R} \text{ 的映射}\}.$$

其上的乘积 σ 代数记为 \mathcal{F} . 注意到定义 3.2.2 中的(1) 和(2) 给出了过程的任意有限维联合分布, 根据测度论中的Kolmogorov 相容性定理(附录中的定理 4.0.14), 在 (Ω, \mathcal{F}) 上存在概率 P , 使得坐标过程 $X_t(\omega) := \omega(t)$, $t \geq 0$ 满足(1) 和(2). 我们称这样的 P 为布朗分布. 下面考虑连续函数的全体:

$$\Omega_0 := C[0, \infty) := \{\omega : \omega \text{ 为 } [0, \infty) \text{ 到 } \mathbb{R} \text{ 的连续映射}\}.$$

麻烦的事情是, $\Omega_0 \notin \mathcal{F}$, 因此我们不能谈论 Ω_0 的概率. 下面考虑 $\mathcal{F}_0 := \{A \cap \Omega_0 : A \in \mathcal{F}\}$. 不难验证, 它是 Ω_0 上的 σ 代数. 维纳则成功地构造出一个 $(\Omega_0, \mathcal{F}_0)$ 上的概率 P_0 , 使得坐标过

程满足(1) 和(2). 因此, 这个概率也称为维纳测度, 布朗运动也被称为维纳过程, 有时也记为 $\{W_t : t \geq 0\}$. 于是, 我们可取 $\Omega_0 = C[0, \infty)$ 作为布朗运动的经典轨道空间.

(此处感觉有点突兀!) [9]写的深入浅出, 非常适合本科生阅读; 更深入的内容可参阅[13]第7章或[24]. 有关布朗运动对物理学发展的意义, 可参阅??, 有关Wiener的贡献可参阅[18].

练习题.

假设 $\{B_t : t \geq 0\}$ 是标准布朗运动.

1. 设 $\{W_t : t \geq 0\}$ 是标准布朗运动, 且与 $\{B_t\}$ 相互独立. (1) 设 $\xi_t = aB_t + bW_t$, 若 $\{\xi_t\}$ 也是标准布朗运动, 那么 a 和 b 应满足什么条件? (2) 取 $\eta_t = B(2t) - B(t)$, 试问 $\{\eta_t\}$ 是标准布朗运动吗? 论证你的结论.
2. 设 $0 < t_1 < t_2 < t_3 < t_4$, 计算 $E[B(t_1)B(t_2)B(t_3)B(t_4)]$;
3. 设 $s < t$, 试证

$$P(B_s > 0, B_t > 0) = \frac{1}{4} + \frac{1}{2\pi} \arcsin \sqrt{\frac{s}{t}}.$$

4. 证明: (1) d 维标准布朗运动的转移密度

$$p_t(x, y) = \prod_{i=1}^d p_t(x_i, y_i) = \frac{1}{(\sqrt{2\pi t})^d} e^{-\sum_{i=1}^d (y_i - x_i)^2 / (2t)}.$$

(2) 对 $d \geq 3$, 格林函数

$$G(x, y) := \int_0^\infty p_t(x, y) dt = \frac{\Gamma(d/2 - 1)}{2\pi^{d/2}} \frac{1}{|x - y|^{d-2}}.$$

(注: 忽略前面的系数, 格林函数正是物理中的牛顿位势.)

5. 验证布朗运动的转移密度满足如下偏微分方程.

$$\frac{\partial p_t(x, y)}{\partial t} = \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 p_t(x, y)}{\partial x_i^2} = \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 p_t(x, y)}{\partial y_i^2}.$$

6. 证明下列随机过程都是标准布朗运动.

- (1) $\{-B_t : t \geq 0\}$,
- (2) $\{B_{t+u} - B_u : t \geq 0\}$,
- (3) $\{B_{T-t} - B_T : 0 \leq t \leq T\}$ (仅有限时间段),

§3.3 不变原理概述

假设 $\{S_n : n \geq 0\}$ 是一维简单随机游动. 具体地, 假设 ξ_1, ξ_2, \dots 独立同分布, $P(\xi_1 = 1) = P(\xi_1 = -1) = \frac{1}{2}$. 记 $S_0 = 0$, $S_n = \xi_1 + \dots + \xi_n$, $n \geq 1$. 对任意整数 $m \geq 1$, 我们利用线性插值补充定义

$$S_{m+r} := (1-r)S_m + rS_{m+1}, \quad 0 < r < 1.$$

固定 n , 令

$$S_t^{(n)} = \frac{S_{nt}}{\sqrt{n}}, \quad \forall t \geq 0.$$

则 $\{S_t^{(n)} : t \geq 0\}$ 是一个轨道连续的过程, 它的轨道是由一维简单随机游动(补充线性插值后)的轨道做时空尺度变换得到, 其中, 微观的单位时间等于 $\frac{1}{n}$, 微观的单位空间等于 $\frac{1}{\sqrt{n}}$. 不变原理说的是, 当 $n \rightarrow \infty$ 时, 随机轨道 $\{S_t^{(n)} : t \geq 0\}$ 的分布越来越像随机布朗轨道 $\{B_t : t \geq 0\}$ 的分布. 为了把这个结论说得更严格, 我们需要引入如下的一些定义.

任意给定 $T > 0$, 将 $[0, T]$ 上所有的连续函数(即时间长度为 T 的所有连续轨道)组成的集合记为 $C[0, T]$. 为符号的统一性, 对任意 $\varphi \in C[0, T]$, 我们将 t 在 φ 映射下的函数值记为 φ_t . 令

$$d_T(\varphi, \psi) := \max_{0 \leq t \leq T} |\varphi_t - \psi_t|, \quad \forall \varphi, \psi \in C[0, T].$$

则 d_T 是 $C[0, T]$ 上的一个距离, 且 $(C[0, T], d_T)$ 是一个完备可分的距离空间. 设 $f : C[0, T] \rightarrow \mathbb{R}$, 鉴于 f 是以轨道(即时间的函数)为自变量的函数, 我们称之为一个泛函. 若存在 $M > 0$ 使得

$$|f(\varphi)| \leq M, \quad \forall \varphi \in C[0, T],$$

则称 f 为 $C[0, T]$ 上的一个有界泛函. 如果 f 连续, 即对任意 $\varphi \in C[0, T]$,

$$\forall \varepsilon > 0, \exists \delta > 0, \text{ 使得: 若 } \psi \in C[0, T] \text{ 满足 } d(\varphi, \psi) < \delta, \text{ 则 } |f(\psi) - f(\varphi)| < \varepsilon,$$

则称 f 为 $C[0, T]$ 上的一个连续泛函.

定理3.3.1. (不变原理) 对任意 $T > 0$, 任意 $C[0, T]$ 上的任意有界连续泛函 f ,

$$\lim_{n \rightarrow \infty} Ef\left(\{S_t^{(n)} : 0 \leq t \leq T\}\right) = Ef(\{B_t : 0 \leq t \leq T\}).$$

例3.3.2. 对任意 $\lambda \in \mathbb{R}$, 取 $f_\lambda(\varphi) = e^{i\lambda\varphi(1)}$. 则 f_λ 是 $C[0, 1]$ 上的有界连续函数, 而

$$Ef_\lambda(\{C_t : 0 \leq t \leq 1\}) = Ee^{i\lambda S_n/\sqrt{n}},$$

此即 S_n/\sqrt{n} 的特征函数, 它收敛到 $Ef_\lambda(\{B_t : 0 \leq t \leq 1\}) = Ee^{i\lambda B_1}$, 即 B_1 的特征函数. 因此 S_n/\sqrt{n} 依分布收敛到 B_1 , 即中心极限定理成立.

例3.3.3. 当 $n \rightarrow \infty$ 时,

$$\frac{1}{\sqrt{n}} \sup_{0 \leq m \leq n} S_m \xrightarrow{d} \max_{0 \leq t \leq 1} B_t.$$

理由如下: 取 $T = 1$. 对任意 $\varphi \in C[0, 1]$, 令 $\text{Max}(\varphi) = \max_{0 \leq t \leq 1} \varphi(t)$. 则 Max 是 $C[0, 1]$ 上的连续函数. 这是因为, 若 $d_1(\varphi, \psi) < \varepsilon$, 则 $M(\varphi) = \varphi(t_0) \leq \psi(t_0) + \varepsilon \leq M(\psi) + \varepsilon$, 其中 t_0 为 φ 的(某个)最大值点. 同理 $M(\psi) \leq \text{Max}(\varphi) + \varepsilon$, 于是 $|\text{Max}(\varphi) - \text{Max}(\psi)| < \varepsilon$. 对 \mathbb{R} 上的任意有界连续函数 F , 我们有 $F \circ \text{Max}$ 是 $C[0, 1]$ 上的有界连续函数. 由不变原理 $EF \circ \text{Max}(\{S_t^{(n)} : 0 \leq t \leq 1\}) \rightarrow EF \circ \text{Max}(\{B_t : 0 \leq t \leq 1\})$, 即 $EF(M_n) \rightarrow EF(M)$, 其中, $M_n = \text{Max}(\{S_t^{(n)} : 0 \leq t \leq 1\}) = \frac{1}{\sqrt{n}} \sup_{0 \leq m \leq n} S_m$, $M_n = \text{Max}(\{B_t : 0 \leq t \leq 1\}) = \max_{0 \leq t \leq 1} B_t$. 由 F 的任意性, 这表明 $M_n \xrightarrow{d} M$.

补充知识

不变原理叙述的事实就是: 简单随机游动产生的随机连续轨道, 在适当尺度变换下, 依分布收敛于布朗运动产生的随机连续轨道. 如同中心极限定理一样, 不变原理的适用范围并不限于简单随机游动, 但证明涉及测度的知识, 有兴趣的同学可参阅[7]第137页.

不变原理传递了这么一个观点: 标准布朗运动是简单随机游动的极限. 这暗示着标准布朗运动与简单随机游动具有许多相同的性质. 例如, 反正弦律(命题 1.10.7 与命题 3.4.16)与下面的重对数律, 这些结论在简单随机游动和标准布朗运动中相映成趣.

定理3.3.4 (重对数律, Khintchine, 1924).

$$P\left(\limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \ln \ln n}} = 1\right) = P\left(\limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = 1\right) = 1.$$

练习题

- 设 \mathbf{f} 是 \mathbb{R}^n 到 \mathbb{R} 的连续函数, $0 \leq t_1, t_2, \dots, t_n \leq T$. 令

$$f : C[0, T] \rightarrow \mathbb{R}, \quad \varphi \mapsto \mathbf{f}(\varphi_{t_1}, \varphi_{t_2}, \dots, \varphi_{t_n}).$$

证明: f 是一个连续泛函.

§3.4 布朗轨道的性质

一、首达时

与马氏链和跳过程类似, 我们定义布朗运动的首达时

$$\tau_A := \inf\{t \geq 0 : B_t \in A\},$$

其中, A 可以是开集, 闭集, 或更一般的可测集. 将 $\tau_{\{a\}}$ 简记为 τ_a . 对任意 $r \geq 0$, 我们还可以考虑时刻 r 之后首达 A 的时间

$$\tau_A^{(r)} := \inf\{t \geq r : B_t \in A\}.$$

假设 τ 为一个首达时(即, 形如 $\tau_A^{(r)}$), 令

$$\hat{B}_t = B_{\tau+t} - B_\tau.$$

定理3.4.1 (强马氏性). 假设 $\{B_t : t \geq 0\}$ 是布朗运动, τ 为一个首达时. 则在已知 $\{\tau < \infty\}$ 的条件下, $\{\hat{B}_t : t \geq 0\}$ 是布朗运动, 且与 B_τ 相互独立.

推论3.4.2 (反射原理). 对任意 $a > 0$,

$$P_0(\tau_a < t, B_t > a) = P_0(\tau_a < t, B_t < a).$$

命题3.4.3. 对任意 $a > 0$, $P_0(\tau_a \leq t) = 2P_0(B_t > a)$.

证: 因为 $P_0(B_t = a) = 0$, 所以

$$P_0(\tau_a \leq t) = P_0(\tau_a < t, B_t > a) + P_0(\tau_a < t, B_t < a).$$

由反射原理, $P_0(\tau_a \leq t) = 2P_0(\tau_a < t, B_t > a)$. 进一步, 因为布朗运动具有轨道连续性, 并且 $a > 0$, 所以 $\{B_t > a\} \subseteq \{\tau_a < t\}$, 从而 $P_0(\tau_a < t, B_t > a) = 2P_0(B_t > a)$. 因此, 命题成立. \square

推论3.4.4. 对任意 $a > 0$, $P_0(\tau_a < \infty) = 1$, 并且, τ_a 是连续型随机变量, 其密度函数为

$$p(t) = \begin{cases} \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}, & t > 0, \\ 0, & \text{其他.} \end{cases}$$

证: 因为 B_t/\sqrt{t} 与 B_1 同分布, 所以

$$P_0(B_t > a) = P_0\left(B_1 > \frac{a}{\sqrt{t}}\right) \xrightarrow{t \rightarrow \infty} \frac{1}{2}.$$

再由命题3.4.3, 可以推出 $P_0(\tau_a < \infty) = 1$. 进一步, 将 B_1 的密度函数记为 $p_{B_1}(\cdot)$, 那么, τ_a 的密度为: 对任意 $t > 0$,

$$\begin{aligned} p(t) &= \frac{d}{dt} P_0(\tau_a \leq t) = 2 \frac{d}{dt} P_0\left(B_1 > \frac{a}{\sqrt{t}}\right) \\ &= 2 \times p_{B_1}\left(\frac{a}{\sqrt{t}}\right) \times \frac{d}{dt}\left(\frac{a}{\sqrt{t}}\right) = 2 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2t}} \frac{a}{2\sqrt{t^3}} = \frac{a}{\sqrt{2\pi t^3}} e^{-\frac{a^2}{2t}}. \end{aligned}$$

从而命题成立. \square

例3.4.5. 对任意 $x, a \in \mathbb{R}$, 从 x 出发的布朗运动首达 a 的时间等价于从 0 出发的布朗运动首达 $a-x$ 的时间, 因此, $P_x(\tau_a < \infty) = 1$.

例3.4.6. 对任意 $c, a > 0$. 由例 3.2.8, $\hat{B}_t := \frac{1}{c} B_{c^2 t}$, 则 $\{\hat{B}_t\}$ 也是标准布朗运动. 将 $\{\hat{B}_t\}$ 的首达时记为 $\hat{\tau}_a$, 那么, $\hat{\tau}_a \stackrel{d}{=} \tau_{ca}$. 注意到

$$\hat{\tau}_a = \inf\{t : \hat{B}_t = a\} = \inf\{t : B_{a^2 t} = ca\} = \frac{1}{a^2} \inf\{a^2 t : B_{a^2 t} = ca\} = \frac{1}{a^2} \inf\{s : B_s = ca\} = \frac{1}{a^2} \tau_{ca}.$$

因此,

$$\tau_a \stackrel{d}{=} \frac{1}{c^2} \tau_{ca}.$$

将 a 视为“时间”, 我们得到首达时过程 $\{\tau_a, a \geq 0\}$, 它的时间 a 对应于布朗运动的取值; 而它的取值 τ_a 则对应于布朗运动的时间. 特别地, 令 $\tilde{B}_t := B_{t+\tau_a} - B_t$, 那么, 对任意 $b > 0$,

$$\tau_{a+b} - \tau_a = \hat{\tau}_b \stackrel{d}{=} \tau_b,$$

并且, $\tau_{a+b} - \tau_a$ 与 τ_b 相互独立. 更一般地, 不难看出: $\{\tau_a, a \geq 0\}$ 也是独立、平稳增量过程.

例3.4.7. 对任意 $\lambda, a > 0$, 试计算 $Ee^{-\lambda\tau_a}$. 当然, 我们也可以直接用 τ_a 的密度函数进行计算, 并得到 $Ee^{-\lambda\tau_a} = e^{-\sqrt{2\lambda}a}$. 不过, 我们可以先完全回避 τ_a 的密度函数, 利用上例, 推出如下结论: 存在 $\beta < 0$ 使得

$$f(\lambda, a) := Ee^{-\lambda\tau_a} = e^{\beta \times \sqrt{\lambda} \times a}. \quad (3.4.1)$$

推理如下: 首先, 固定 λ . 由上例,

$$f(\lambda, a+b) = Ee^{-\lambda\tau_{a+b}} = Ee^{-\lambda\tau_a} Ee^{-\lambda(\tau_{a+b}-\tau_a)} = Ee^{-\lambda\tau_a} Ee^{-\lambda\tau_b} = f(\lambda, a)f(\lambda, b).$$

这表明 $\ln f(\lambda, a+b) = \ln f(\lambda, a) + \ln f(\lambda, b)$. 又因为 $f(\lambda, a)$ 关于 a 单调下降, 因此 $\ln f(\lambda, a)$ 是 a 的线性函数, 即存在 $\beta(\lambda) < 0$ 使得 $\ln f(\lambda, a) = \beta(\lambda) \times a$, $a \geq 0$. 事实上, $\beta(\lambda) = \ln f(\lambda, 1)$. 进一步, 由上例, $\tau_a \stackrel{d}{=} a^2 \tau_1$, 因此, $f(a^2 \lambda; 1) = Ee^{-\lambda a^2 \tau_1} = Ee^{-\lambda \tau_a} = f(\lambda, a)$, 即, $\beta(a^2 \lambda) = \beta(\lambda) \times a$. 在这个式子中, 将 λ 设为 1, 将 a^2 视为 λ , 则可推出 $\beta(\lambda) = \beta(\sqrt{\lambda}^2 \times 1) = \beta(1) \times \sqrt{\lambda}$. 记 $\beta = \beta(1)$.

综上, 即便不利用 τ_a 的密度函数, 我们也可以从布朗运动的强马氏性和尺度变换性质推出 (3.4.1) 成立.

当然, 为求出常数 β , 我们还是需要利用 τ_a 的密度函数. 具体地,

$$Ee^{-2\tau_1} = \int_0^\infty e^{-2t} \frac{1}{\sqrt{2\pi t^3}} e^{-\frac{1}{2t}} dt = \frac{1}{\sqrt{\pi}} \times I,$$

其中, 我们做变量替换 $s = 2t$, 并令

$$I := \int_0^\infty e^{-(s+\frac{1}{s})} \frac{1}{\sqrt{s^3}} ds.$$

将 $\frac{1}{s}$ 视为新的变量 s , 则

$$I = \int_0^\infty e^{-(\frac{1}{s}+s)} \sqrt{s}^2 d\frac{1}{\sqrt{s}} = \int_0^\infty e^{-(s+\frac{1}{s})} \frac{1}{\sqrt{s}} ds.$$

令 $u = \sqrt{s} - \frac{1}{\sqrt{s}}$, 那么, $u^2 = s + \frac{1}{s} - 2$, 且 $du = \frac{1}{2}(\frac{1}{\sqrt{s}} + \frac{1}{s^2})$. 因此,

$$I = \int_0^\infty e^{-(s+\frac{1}{s})} \times \frac{1}{2}(\frac{1}{\sqrt{s^3}} + \frac{1}{\sqrt{s}}) ds = \int_{-\infty}^\infty e^{-(u^2+2)} du = \sqrt{\pi} e^{-2}.$$

于是, $e^{\beta \times \sqrt{2} \times 1} = Ee^{-2\tau_1} = e^{-2}$, 即, $\beta = -\sqrt{2}$.

二、最大值

因为 B_s 在 $s \in [0, t]$ 上连续, 因此它可以达到最大值. 令

$$M_t = \max_{0 \leq s \leq t} B_s.$$

命题3.4.8. M_t 是连续型随机变量, 且它与 $|B_t|$ 同分布.

证: 显然 $M_t \geq B_0 = 0$. 由命题 3.4.3, 对任意 $a > 0$,

$$P_0(M_t \geq a) = P_0(\tau_a \leq t) = 2P_0(B_t > a) = P_0(|B_t| > a).$$

从而命题成立. \square

注意到 $P_0(M_t = 0) = 0$, 我们可以推出: 最大值 M_t 不能在区间 $[0, t]$ 的边界取到. 这是因为: 若最大值在左端点取到, 即 $M_t = B_0$, 那么, $M_t = 0$, 该事件的概率为 0; 若最大值在右端点取到, 即 $M_t = B_t$, 那么我们可以考虑 $\tilde{B}_s = B_{t-s} - B_t$, $0 \leq s \leq t$, 这是一个(有限时间区间上的)标准布朗运动, $M_t = B_t$ 表明 $\tilde{M}_t := \max_{0 \leq s \leq t} \tilde{B}_s = 0$, 该事件的概率也为 0. 换句话说, 最大值点只能出现在区间内部 $(0, t)$. 下面的推论表明, 最大值点是唯一的.

推论3.4.9. 对任意 $t > 0$,

$$P_0(\text{存在 } u, v \in (0, t), u \neq v \text{ 使得 } B_u = B_v = M_t) = 0.$$

证：若存在两个不相等的时刻 u 与 v ，不妨设 $u < v$ ，我们就能在 (u, v) 中找到一个有理数 r 。于是，上述事件蕴含着：存在有理数 r 使得 A_r 发生，其中

$$A_r := \{\text{存在 } u \in (0, r), v \in (r, t) \text{ 使得 } B_u = B_v = M_t\}$$

往证 $r \in \mathbb{Q}_+ \cap (0, t)$ 都有 $P(A_r) = 0$ 。固定 r 。令

$$\hat{B}_s = B_{r-s} - B_r, \quad 0 \leq s \leq r; \quad \tilde{B}_s = B_{s+r} - B_r, \quad s \geq 0.$$

则 $\{\hat{B}_s : 0 \leq s \leq r\}$ 与 $\{\tilde{B}_s : s \geq 0\}$ 是相互独立的布朗运动。令

$$\hat{M}_r := \max_{0 \leq s \leq r} \hat{B}_s, \quad \tilde{M}_{t-r} := \max_{0 \leq s \leq t-r} \tilde{B}_s.$$

那么， \hat{M}_r 与 \tilde{M}_{t-r} 相互独立。注意到它们都是连续型随机变量，因此 $(\hat{M}_r, \tilde{M}_{t-r})$ 是二维连续型随机向量。于是，

$$P_0(A_r) = P_0(\hat{M}_r = \tilde{M}_{t-r}) = 0.$$

最后，

$$\begin{aligned} & P_0(\exists u, v \in (0, t), u \neq v \text{ 使得 } B_u = B_v = M_t) \\ & \leq P_0(\cup_{r \in \mathbb{Q}_+ \cap (0, t)} A_r) \leq \sum_{r \in \mathbb{Q}_+ \cap (0, t)} P_0(A_r) = 0. \end{aligned}$$

因此，该推论成立。 □

推论 3.4.10. $P_0(\lim_{t \rightarrow \infty} \frac{1}{t} B_t = 0) = 1$ 。

证：注意到

$$B_n = (B_1 - B_0) + (B_2 - B_1) + \cdots + (B_n - B_{n-1}),$$

其中， $B_m - B_{m-1}$, $m = 1, 2, \dots$ 是独立同分布的随机变量序列。由强大数定律，

$$\frac{1}{n} B_n \xrightarrow{\text{a.s.}} E B_1 = 0.$$

对任意 $n \geq 0$ ，令

$$X_n = \sup_{n \leq t \leq n+1} (B_t - B_n), \quad Y_n = \sup_{n \leq t \leq n+1} (B_n - B_t),$$

则 X_0, X_1, X_2, \dots 是独立同分布的随机变量序列，且 X_0 与 M_1 同分布，即它与 $|B_1|$ 同分布。由引理 4.0.25，

$$\frac{1}{n} X_n \xrightarrow{\text{a.s.}} 0, \quad \frac{1}{n} Y_n \xrightarrow{\text{a.s.}} 0.$$

注意到当 $n \leq t \leq n+1$ 时， $|B_t| \leq |B_n| + X_n + Y_n$ 。我们便可以推出 $\frac{1}{t} B_t \xrightarrow{\text{a.s.}} 0$ 。 □

令

$$W_t = \begin{cases} tB_{1/t}, & t > 0; \\ 0, & t = 0. \end{cases}$$

推论3.4.11. $\{W_t\}$ 是一维标准布朗运动.

证: 根据上面的推论, W_t 轨道连续. 不难看出, $\{W_t\}$ 是高斯过程; 对任意 $t \geq 0$, $EW_t = 0$; 并且对任意 $s \geq t = 0$, $EW_t W_s = EW_0 W_s = 0$, 对任意 $s \geq t > 0$,

$$EW_t W_s = EtB_{1/t} \times sB_{1/s} = ts \times EB_{1/t} B_{1/s} = ts \times \frac{1}{s} = t.$$

因此, $\{W_t\}$ 是一维标准布朗运动. \square

三、轨道性质

命题3.4.12. $P_0(\forall a < b, B_t \text{ 在区间 } [a, b] \text{ 中不是单调函数}) = 0$.

证: 给定 $r < s$, 记

$$A_{r,s} := \{B_t \text{ 在区间 } [r, s] \text{ 中单调}\}.$$

对任意正整数 n , 记 $r_{n,m} = r + \frac{m}{n}(s-r)$, $m = 0, 1, \dots, n$. 那么, $B_{r_{n,m}} - B_{r_{n,m-1}}$, $m = 1, \dots, n$ 独立同分布, 都服从 $N(0, \frac{s-r}{n})$. 若 $A_{r,s}$ 发生, 则 $B_{r_{n,m}} - B_{r_{n,m-1}}$, $m = 1, \dots, n$ 全部同号. 因此,

$$P(A_{r,s}) \leq P(B_{r_{n,m}} - B_{r_{n,m-1}} \geq 0, m = 1, \dots, n) + P(B_{r_{n,m}} - B_{r_{n,m-1}} \leq 0, m = 1, \dots, n) \leq 2 \cdot 2^{-n}.$$

由 n 的任意性, $P(A_{r,s}) = 0$.

对任意 $a < b$, 存在有理数 r, s 满足 $a < r < s < b$. 若 B_t 在区间 $[a, b]$ 中单调, 则 $A_{r,s}$ 发生. 于是,

$$\begin{aligned} P_0(\exists a < b \text{ 使得 } B_t \text{ 在区间 } (a, b) \text{ 中单调}) \\ \leq P_0(\cup_{r,s \in \mathbb{Q}, 0 < r < s} A_{r,s}) \leq \sum_{r,s \in \mathbb{Q}, 0 < r < s} P_0(A_r) = 0. \end{aligned}$$

因此, 命题成立. \square

命题3.4.13. $P_0(\forall t_0 \geq 0, B_t \text{ 在 } t_0 \text{ 不可微}) = 1$.

证: 先考虑 $s \in [0, 1]$. 记

$$A_N = \left\{ \exists t_0 \in [N, N+1] \text{ 使得 } \lim_{t \in [n, n+1], t \rightarrow t_0} \frac{B_t - B_{t_0}}{s - t_0} \text{ 存在} \right\}.$$

先考虑 $N = 0$ 的情形. 假设 A_0 发生, 且 $t_0 \in [0, 1]$ 使得 $\lim_{t \in [n, n+1], t \rightarrow t_0} \frac{B_t - B_{t_0}}{s - t_0}$ 存在. 那么, 存在 $M > 0$ 和充分小的 $\delta > 0$ 使得

$$|B_t - B_{t_0}| \leq M|t - t_0|, \quad \forall t \in [0, 1] \cap [t_0 - \delta, t_0 + \delta].$$

当 $4/n < \delta$ 时, 存在 $0 \leq k \leq n - 3$ 使得 $\frac{k+i}{n} \in [0, 1] \cap [t_0 - \delta, t_0 + \delta]$, $i = 0, 1, 2, 3$. 于是,

$$|B_{(k+i)/n} - B_{t_0}| \leq \frac{4M}{n}, \quad i = 0, 1, 2, 3.$$

这蕴含着下面定义的事件 $A_{n,k}$ 发生,

$$A_{n,k} = \left\{ |B_{(k+i)/n} - B_{(k+i-1)/n}| \leq \frac{8M}{n}, \quad i = 1, 2, 3. \right\}$$

综上,

$$A_0 \subseteq \bigcup_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \bigcup_{k=0}^{n-3} A_{n,k}.$$

往证上式右端的事件的概率为 0. 对任意 $M, N \geq 1$, $n \geq N$, 固定 $k = 0, \dots, n-3$, 注意到 $B_{(k+i)/n} - B_{(k+i-1)/n}$, $i = 1, 2, 3$ 是相互独立的随机变量, 它们都与 $\frac{1}{\sqrt{n}}B_1$ 同分布, 并且

$$P_0 \left(\frac{1}{\sqrt{n}} |B_1| \leq \frac{8M}{n} \right) = \int_{-8M/\sqrt{n}}^{8M/\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \leq \frac{8M}{\sqrt{n}}.$$

于是 $P(A_{n,k}) \leq (8M/\sqrt{n})^3$. 进一步,

$$\sum_{k=0}^{n-3} P(A_{n,k}) \leq n(8M/\sqrt{n})^3 \xrightarrow{n \rightarrow \infty} 0.$$

这表明 $P_0(\bigcap_{n=N}^{\infty} \bigcup_{k=0}^{n-3} A_{n,k}) = 0$. 证毕. □

对任意 $\delta > 0$, 令

$$\text{osc}(\delta) := \max_{t, s \in [0, 1], |t-s| \leq \delta} |B_t - B_s|.$$

它表示布朗运动在单位时间区间内的 δ 振幅(oscillation).

命题3.4.14. $P_0 \left(\limsup_{\delta \rightarrow 0} \frac{\text{osc}(\delta)}{(-\delta \log \delta)^{1/2}} \leq 6 \right) = 1$.

证: 对任意整数 $n, m \geq 0$, 考虑二分区间 $I_{n,m} = [\frac{m-1}{2^n}, \frac{m}{2^n}]$, $m = 1, \dots, 2^n$. 令

$$\Delta_{n,m} := \max_{s \in I_{n,m}} |B_s - B_{m/2^n}|.$$

由布朗运动的平稳、独立增量性, $\Delta_{n,1}, \dots, \Delta_{n,2^n}$ 独立同分布. 根据布朗运动的尺度变换性质(具体地, 在例 3.2.8 中取 $c = 2^{-n}$)知

$$\Delta_{n,1} \stackrel{d}{=} 2^{-n/2} \max_{0 \leq u \leq 1} |B_u|.$$

对任意 $x \geq \frac{4}{\sqrt{2\pi}}$,

$$\begin{aligned} P_0(\Delta_{n,m} > 2^{-n/2}x) &= P_0\left(\max_{0 \leq u \leq 1} |B_u| > x\right) = P_0\left(\max_{0 \leq u \leq 1} B_u > x \text{ 或 } \max_{0 \leq u \leq 1} (-B_u) > x\right) \\ &= 2P_0(M_1 > x) = 2P_0(|B_1| > x) = 4P_0(B_1 > x) \\ &= \int_x^\infty \frac{4}{\sqrt{2\pi}} e^{-z^2/2} dz \leq \int_x^\infty z e^{-z^2/2} dz = e^{-x^2/2}. \end{aligned}$$

特别地, 取定 $\varepsilon > 0$, 并令

$$x_n = \sqrt{2(1+\varepsilon)n \ln 2},$$

令

$$A_n := \left\{ \text{存在 } 1 \leq m \leq 2^n \text{ 使得 } \Delta_{n,m} > 2^{-n/2}x_n \right\}.$$

当 $n \geq 2$ 时,

$$P_0(A_n) = 2^n \times 2^{-(1+\varepsilon)n} = 2^{-\varepsilon n}.$$

由于 $\sum_{n=1}^\infty 2^{-\varepsilon n} < \infty$, 根据 Borel-Cantelli 引理, $P(\cap_{N=1}^\infty \cup_{n=N}^\infty A_n) = 0$. 即,

$$P(\cup_{N=1}^\infty \cap_{n=N}^\infty A_n^c) = 1.$$

下面假设事件 $\cap_{n=N}^\infty A_n^c$ 发生. 记 $n = \lfloor \frac{-\ln \delta}{\ln 2} \rfloor$, 即, $2^{-(n+1)} < \delta \leq 2^{-n}$. 当 $\delta \leq 2^{-N}$ 时, $n \geq N$. 假设 $t, s \in [0, 1]$, $|t - s| < \delta$. 假设 $s \in I_{n,m}$, $t \in I_{n,m'}$. 不妨假设 $t > s$, 即, $m' \geq m$. 根据 $t - s < \delta \leq 2^{-n}$, m' 只可能为 m 或 $m + 1$. 根据三角不等式,

$$\left\{ \begin{array}{l} |B_t - B_s| \leq |B_t - B_{m/2^n}| + |B_{m/2^n} - B_s| \\ \quad \leq 2\Delta_{n,m}, \quad \text{若 } m' = m; \\ |B_t - B_s| \leq |B_t - B_{(m+1)/2^n}| + |B_{(m+1)/2^n} - B_{m/2^n}| + |B_{m/2^n} - B_s| \\ \quad \leq \Delta_{n,m+1} + 2\Delta_{n,m}, \quad \text{若 } m' = m + 1. \end{array} \right.$$

于是, 我们总有:

$$|B_t - B_s| \leq 3 \times 2^{-n/2}x_n \leq 3 \times \sqrt{2\delta} \sqrt{2(1+\varepsilon) \ln(\delta^{-1})} = 6\sqrt{1+\varepsilon} \cdot \sqrt{-\delta \ln \delta}.$$

综上, 如果事件 $\cap_{n=N}^\infty A_n^c$ 发生, 那么对任意 $\delta \leq 2^{-N}$ 都有 $\text{osc}(\delta)/\sqrt{\delta \log_2(\delta^{-1})} \leq 6\sqrt{1+\varepsilon}$.

令 $\delta \rightarrow 0$ 知,

$$\cap_{n=N}^\infty A_n^c \subseteq \left\{ \limsup_{\delta \rightarrow 0} \frac{\text{osc}(\delta)}{(-\delta \ln \delta)^{1/2}} \leq 6\sqrt{1+\varepsilon} \right\}.$$

最后, 令 $N \rightarrow \infty$ 知上式右边的事件的概率等于 1, 再由 ε 的任意性知结论成立. \square

注3.4.15. 对任意 $\alpha \in (0, \frac{1}{2})$, 存在随机变量 η 使得

$$P_0(|B_t - B_s| \leq \eta \cdot |t - s|^\alpha, \forall t, s \in [0, 1]) = 1.$$

但此结论在 $\alpha = 1/2$ 时不成立.

四、零点

将布朗轨道在 $[0, t]$ 上的最后一个零点记为 L_t . 根据布朗运动的轨道连续性, $B_{L_t} = 0$. 因此,

$$L_t := \sup\{s \leq t : B_s = 0\} = \max\{s \leq t : B_s = 0\}.$$

命题3.4.16 (反正弦律). 对任意 $t \geq s \geq 0$,

$$P_0(L_t \leq s) = \frac{2}{\pi} \arcsin \sqrt{\frac{s}{t}}.$$

L_t 为连续型随机变量, 其密度为

$$p(s) = \begin{cases} \frac{1}{\pi \sqrt{t-s}}, & s \in (0, t), \\ 0, & \text{其他.} \end{cases}$$

证: 注意到 L_t 只取值于区间 $[0, t]$, 因此我们只用考虑 $s \in (0, t)$ 的情形. 固定 s . 对任意 $u \geq 0$, 令 $B_u = B_{u+s} - B_s$. 则 $\{B_u\}$ 是布朗运动, 并且它与 B_s 相互独立. 将 $\{\tilde{B}_u\}$ 首达 a 的时间记为 $\tilde{\tau}_a$. 那么,

$$\{L_t \leq s\} = \{\tilde{\tau}_{-B_s} > t-s\}.$$

于是,

$$P_0(L_t \leq s) = \int_{-\infty}^{\infty} p_s(0, x) P_0(\tilde{\tau}_{-x} > t-s) dx.$$

根据反射原理,

$$P_0(\tilde{\tau}_{-x} > t-s) = P_0(|\tilde{B}_{t-s}| \leq |x|).$$

于是

$$P_0(L_t \leq s) = \int_{-\infty}^{\infty} p_s(0, x) P_0(|\tilde{B}_{t-s}| \leq |x|) dx = P_0(|\tilde{B}_{t-s}| \leq |B_s|).$$

记

$$W = B_s / \sqrt{s}, \quad Z = \tilde{B}_{t-s} / \sqrt{t-s}.$$

那么, W 与 Z 相互独立, 都服从 $N(0, 1)$. 将 (Z, W) 用极坐标表达为 $W = R \cos U$, $Z = R \sin U$, 其中 $R > 0$, $U \in (-\pi, \pi)$. 那么, $U \sim U(-\pi, \pi)$. 于是,

$$\begin{aligned} P_0(L_t \leq s) &= P(\sqrt{t-s}|Z| \leq \sqrt{s}|W|) \\ &= P(|U| \leq \arcsin \sqrt{s/t}) + P(|U| \geq \pi - \arcsin \sqrt{s/t}) \\ &= 4 \times \frac{1}{2\pi} \arcsin \sqrt{s/t} = \frac{2}{\pi} \arcsin \sqrt{s/t}. \end{aligned}$$

将上式右边对 s 求导, 我们得到 L_t 的密度函数在 s 的取值, 所以当 $s \in (0, t)$ 时,

$$p(s) = \frac{2}{\pi} \frac{1}{\sqrt{1 - (\sqrt{s/t})^2}} \frac{d\sqrt{s/t}}{ds} = \frac{2}{\pi} \sqrt{\frac{t}{t-s}} \frac{1}{2\sqrt{st}} = \frac{1}{\pi \sqrt{s(t-s)}}.$$

□

记

$$\sigma_0 = \inf\{t > 0 : B_t = 0\}.$$

推论3.4.17. $P_0(\sigma_0 = 0) = 1$.

证: 对任意整数 $n \geq 1$, 根据反正弦律, $P_0(0 < L_{1/n} < 1/n) = 1$. 因此, 事件 $A := \cap_{n=1}^{\infty} \{0 < L_{1/n} < 1/n\}$ 的概率为 1. 当事件 A 发生时, $L_{1/n} > 0$ 且 $\lim_{n \rightarrow \infty} L_{1/n} = 0$, 于是, $\sigma_0 = 0$. 从而, $P_0(\sigma_0 = 0) \geq P(A) = 1$. □

注3.4.18. 上面的推论表明布朗运动事实上没有“首次返回的时间”. 进一步, 对任意 $a \neq 0$, 令 $\tilde{B}_t = B_{\tau_a+t} - B_t$, $t \geq 0$; $\tilde{\sigma}_0 = \inf\{t > 0 : \tilde{B}_t = 0\}$. 根据强马氏性, $\{\tilde{B}_t\}$ 是从 0 出发的布朗运动, 因此 $P_0(\tilde{\sigma}_0 = 0) = 1$. 从而, $\{B_t\}$ 没有“第二次到达 a ”的时间.

令

$$\mathcal{Z} := \{t \geq 0 : B_t = 0\}.$$

它是布朗轨道的零点集, 是 $[0, \infty)$ 的一个随机子集. 因为布朗运动的轨道是连续的, 所以 \mathcal{Z} 几乎必然为一个闭集.

定义3.4.19. 假设 D 为 \mathbb{R} 中的一个闭集. 如果任意 $t \in D$ 都存在一列 $t_1, t_2, \dots \in D$ 使得对所有的 $n \geq 1$ 都有 $t_n \neq t$, 并且 $\lim_{n \rightarrow \infty} t_n = t$, 那么称 D 为一个完全闭集.

推论3.4.20. $P_0(\mathcal{Z} \text{ 是完全闭集}) = 1$.

证: 记

$$\mathcal{Z}' := \{t \in \mathcal{Z} : \exists t_1, t_2, \dots \in \mathcal{Z} \cap [0, t) \text{ 使得 } \lim_{n \rightarrow \infty} t_n = t\}.$$

往证 $P_0(\mathcal{Z} \subseteq \mathcal{Z}') = 1$. 令

$$\begin{aligned} C_- &:= \{t > 0 : \exists t_1, t_2, \dots \in \mathcal{Z} \cap [0, t) \text{ 使得 } \lim_{n \rightarrow \infty} t_n = t\}, \\ C_+ &:= \{t > 0 : \exists \delta > 0 \text{ s.t. } B_s \neq 0, \forall s \in (t - \delta, t)\}. \end{aligned}$$

那么,

$$\mathcal{Z} = \{0\} \cup C_- \cup C_+,$$

由 $P_0(\sigma_0 = 0) = 1$ 知 $P_0(0 \in \mathcal{Z}') = 1$; 根据 C_- 的定义, $P_0(C_- \subseteq \mathcal{Z}') = 1$. 因此, 我们只需证明 $P_0(C_+ \subseteq \mathcal{Z}') = 1$. 对任意 $t \in C_+$, 存在 $r \in \mathbb{Q} \cap (0, t)$ 使得当 $s \in [r, t)$ 时 $B_s \neq 0$. 于是 t 是时刻 r 之后的第一个零点, 即

$$t = \tau_0^{(r)} := \inf\{s \geq r : B_s = 0\}.$$

注意到 $B_{\tau_0^{(r)}} = 0$. 令

$$\hat{B}_t = B_{\tau_0^{(r)} + t}, \quad \forall t \geq 0.$$

根据强马氏性, $\{\hat{B}_t\}$ 是布朗运动. 记 $\hat{\sigma}_0 = \inf\{t > 0 : \hat{B}_t = 0\}$. 当事件 $\{\hat{\sigma}_0 = 0\}$ 发生时, 存在 $s_1, s_2, \dots > 0$ 使得 $\lim_{n \rightarrow \infty} s_n = 0$, 并且 $\hat{B}_{s_n} = 0$. 记 $t_n = \tau_0^{(r)} + s_n$, 我们有 $t_1, t_2, \dots \in \mathcal{Z}$, 且 $\lim_{n \rightarrow \infty} t_n = \tau_0^{(r)}$. 因此, $\tau_0^{(r)} \in \mathcal{Z}'$. 综上,

$$P_0(\tau_0^{(r)} \in \mathcal{Z}') \geq P_0(\hat{\sigma}_0 = 0) = 1.$$

从而, 命题成立. \square

练习题

1. 设 $\{W_t\}$ 是一维标准布朗运动, 且与 $\{B_t\}$ 相互独立, $b > a > 0$. 令 $Y_a = W_{\tau_a}$. (1) 求 Y_a 的分布, 并证明 Y_a 与 aY_1 同分布; (2) 证明: Y_a 与 $Y_b - Y_a$ 相互独立. (注: $\{Y_a : a \geq 0\}$ 被称为柯西(Cauchy)过程, 它是独立平稳增量的, 但轨道不连续.)
2. 试证: 对任何 $t > 0$, M_t 与 $|B_t|$ 同分布.
3. 设 $x > y$, 试求 $P(M_t \geq x, B_t \leq y)$. (用积分表达)
4. 证明 (M_t, B_t) 是二维连续型随机变量, 并求其联合密度函数.
- 5* 将 $[0, t]$ 上唯一的最大值点记为 Λ_t . 证明 (M_t, B_t, Λ_t) 是三维连续型随机变量, 并求其联合密度函数.
6. 证明: $P_0(M_t > a | B_t = M_t) = e^{-a^2/2t}$.
7. 假设 $C > 0$. 证明:

$$P\left(\sup_{0 \leq u \leq t} |B_u| > C\right) \leq \frac{2t}{C^2}.$$

8* 证明上题中, 不等号右边的常数2 可以改为1.

9. 证明: $P_0(\lim_{y \rightarrow x} \tau_y = \tau_x) = 1$.

10. 利用上题证明 $P_0(\sigma_0 = 0) = 1$.

11. 用两种方法证明一维标准布朗运动是点常返的, 即,

$$P_0(\forall t > 0, \exists s > t, \text{使得 } B_s = 0) = 1.$$

12. 设 $0 < t_0 \leq t_1 \leq t_2$. 试求 $P_0(B_s \neq 0, \forall s \in (t_0, t_2) | B_s \neq 0, \forall s \in (t_0, t_1))$.

13. 若 $t = \tau_1$, 取 $X_t = 0$; 若 $t \neq \tau_1$, 令 $X_t = B_t$. 验证: $\{X_t\}$ 满足定义 3.2.2 中的(1) 和(2), 但不满足其中的(3). (注: 满足(1) 和(2) 的过程并不唯一).
14. 证明: $R := \inf\{t > 1 : B_t = 0\}$ 的概率密度为

$$\rho(t) = \begin{cases} \frac{1}{\pi t \sqrt{t-1}}, & t > 1; \\ 0, & \text{其他.} \end{cases}$$

15. 证明布朗轨道的局部极值点是稠集, 即

$$\{t \geq 0 : \exists \delta > 0 \text{ 使得若 } |s - t| < \delta, s \neq t, \text{ 则 } B_s < B_t\}$$

以概率1 是 \mathbb{R} 中的稠集.

- 16* 证明: 对任意 $\alpha \in (0, 1/2)$, 存在依赖于 α 的随机变量 η_α 使得

$$P_0(|B_t - B_s| \leq \eta_\alpha |t - s|^\alpha, \forall t, s \in [0, 1]) = 1.$$

- 17* 令

$$X_n = \left| \left\{ i : 1 \leq i \leq n, \mathcal{Z} \cap \left[\frac{i-1}{n}, \frac{i}{n} \right] \neq \emptyset \right\} \right|.$$

证明: $\lim_{n \rightarrow \infty} E X_n / \sqrt{n} = 1$. (注: \mathcal{Z} 的Hausdorff 维数几乎必然为 $\frac{1}{2}$.)

§3.5 位势理论

一、一维情形

在一维情形, 我们总假设

$$a < x < b, \quad \{B_t : t \geq 0\} \text{ 是从 } x \text{ 出发的布朗运动.}$$

对任意 $c \in \mathbb{R}$, 记

$$\tau_c := \inf\{t \geq 0 : B_t = c\}, \quad \tau = \min\{\tau_a, \tau_b\}.$$

命题3.5.1. 对任意 $a \leq x \leq b$,

$$P_x(\tau_b < \tau_a) = \frac{x-a}{b-a}, \quad P_x(\tau_a < \tau_b) = \frac{b-x}{b-a}.$$

证: 记 $\varphi(x) = P_x(\tau_b < \tau_a)$. 假设 $a < x < b$, $\delta > 0$ 使得 $a \leq x - \delta < x < x + \delta \leq b$.

令 $\sigma := \tau_{x-\delta} \wedge \tau_{x+\delta}$. 一方面, 由对称性,

$$P_x(B_\sigma = x + \delta) = P_x(\tau_{x+\delta} < \tau_{x-\delta}) = \frac{1}{2}.$$

另一方面, 根据轨道连续性, $\sigma \leq \tau_a \wedge \tau_b$. 记 $\hat{B}_t = B_{\sigma+t}$, 根据强马氏性, 在 $\{B_\sigma = x + \delta\}$ 的条件下, $\{\hat{B}_t\}$ 是从 $x + \delta$ 出发的布朗运动, 而在 $\{B_\sigma = x - \delta\}$ 的条件下, $\{\hat{B}_t\}$ 是从 $x - \delta$ 出发的布朗运动, 并且, 总有

$$\tau_a = \tau + \hat{\tau}_a, \quad \tau_b = \tau + \hat{\tau}_b.$$

因此,

$$\begin{aligned} \varphi(x) &= P(B_\sigma = x + \delta)P(\tau_a < \tau_b | B_\sigma = x + \delta) + P(B_\sigma = x - \delta)P(\tau_a < \tau_b | B_\sigma = x - \delta) \\ &= \frac{1}{2}P(\hat{\tau}_b < \hat{\tau}_a | B_\sigma = x + \delta) + \frac{1}{2}P(\hat{\tau}_b < \hat{\tau}_a | B_\sigma = x - \delta) \\ &= \frac{1}{2}P_{x+\delta}(\tau_b < \tau_a) + \frac{1}{2}P_{x-\delta}(\tau_b < \tau_a) = \frac{1}{2}(\varphi(x + \delta) + \varphi(x - \delta)). \end{aligned}$$

注意到边界值 $\varphi(a) = 0$, $\varphi(b) = 1$. 于是, 根据上式, $\varphi(\frac{a+b}{2}) = \frac{1}{2}(\varphi(a) + \varphi(b)) = \frac{1}{2}$. 然后, 利用归纳法不难证明, 对区间 $[a, b]$ 中的二分点 x (即存在整数 $n \geq 1$, $1 \leq m \leq 2^n$, 使得 $x = a + \frac{m}{2^n}(b - a)$ 的点), 总有 $\varphi(x) = (x - a)/(b - a)$. 进一步, 由强马氏性, 当 $a < x < y < b$ 时,

$$\varphi(x) = P_x(\tau_y < \tau_a)\varphi(y) \leq \varphi(y).$$

因此, $\varphi(x)$ 是单调上升的. 因此, 对任意 $x \in (a, b)$, 我们可以取一列二分点 x_1, x_2, \dots 单调上升到 x ; 再取一列二分点 y_1, y_2, \dots 单调下降到 x , 于是推出 $\varphi(x) = (x - a)/(b - a)$ 成立. 进一步, 由推论 3.4.4,

$$P_x(\tau_a < \infty) = P_x(\tau_b < \infty) = 1.$$

因此, 必有 $\tau_a \neq \tau_b$, 于是 $P_x(\tau_a < \tau_b) = P_x(\tau_b < \tau_a) = 1$. 从而结论成立. \square

例3.5.2. 假设 $\{B_t\}$ 是从0出发一维布朗运动. 令 $\sigma_0 = 0$. 对于 $n \geq 1$, 递归地定义

$$\sigma_n := \inf\{t \geq \sigma_{n-1}, |B_t - B_{\sigma_{n-1}}| = 1\}.$$

则 $\sigma_1, \sigma_2 - \sigma_1, \sigma_3 - \sigma_2, \dots$ 独立同分布. 令 $S_0 := 0, S_n = B_{\sigma_n}, n \geq 1$. 那么, 由强马氏性(定理3.4.1)和上述命题, $\{S_n\}$ 是一维简单随机游动.

推论3.5.3 (Wald引理). $E_x B_\tau = x$.

证: $E_x B_\tau = bP_x(B_\tau = b) + aP_x(B_\tau = a) = x$. □

引理3.5.4. $\sup_{a \leq x \leq b} E_x \tau^\alpha < \infty$ 对任意 $\alpha > 0$ 都成立.

证: 记 $\delta = \sup_{a \leq x \leq b} P_x(B_1 \in [a, b])$. 注意到 $f(x) := P_x(B_1 \in [a, b])$ 关于 x 连续, 所以它在区间 $[a, b]$ 上可以取到最大值, 即存在 $x_0 \in [a, b]$ 使得 $\delta = f(x_0)$. 因此, $\delta < 1$. 进一步, 对任意整数 $n \geq 1$, 注意到事件 $\{\tau > n\}$ 是否成立, 仅依赖于 $\{B_t : t \leq n\}$, 因此它与 $\hat{B}_1 = B_{n+1} - B_n$ 相互独立. 在已知 $\{B_t : t \leq n\}$ 的条件下, 如果 $\{\tau > n\}$ 为真, 那么 $B_n \in [a, b]$, 于是 $\{B_{n+1} \in [a, b]\}$ 的(条件)概率不超过 δ . 综上,

$$P_x(\tau > n+1) \leq P_x(\tau > n, B_{n+1} \in [a, b]) \leq \delta P_x(\tau > n).$$

由归纳法, 对任意 $n \geq 1$ 都有 $P_x(\tau > n) \leq \delta^n$. 显然, 此不等式对 $n = 0$ 也成立. 因此, 对任意 $\alpha > 0$,

$$\begin{aligned} E_x \tau^\alpha &= \int_0^\infty P_x(\tau^\alpha > t) dt = \sum_{n=0}^{\infty} \int_{n^\alpha}^{(n+1)^\alpha} P_x(\tau^\alpha > t) dt \\ &\leq \sum_{n=0}^{\infty} (n+1)^\alpha P_x(\tau > n) \leq \sum_{n=0}^{\infty} (n+1)^\alpha \delta^n < \infty. \end{aligned}$$

证毕. □

引理3.5.5 (Wald第二引理). $E_x(B_\tau - x)^2 = E_x \tau$.

由于该引理的证明已经超过本书的要求范围, 故此略去. 读者可比照简单随机游动的结论进行理解, 有兴趣的读者可以参阅??.

推论3.5.6. $E_x \tau = (x - a)(b - x)$.

证: 由命题3.5.1和上述命题,

$$E_x \tau = E_x(B_\tau - x)^2 = \frac{x-a}{b-a}(b-x)^2 + \frac{b-x}{b-a}(a-x)^2 = (x-a)(b-x).$$

□

假设 $f : \{a, b\} \rightarrow \mathbb{R}$ 是区间边界上的函数, 令

$$\varphi(x) := E_x f(B_\tau),$$

假设 $g : (a, b) \rightarrow \mathbb{R}$ 是区间内部的函数, 令

$$\psi(x) := E_x \int_0^\tau g(B_t) dt.$$

命题3.5.7. φ 是如下狄利克莱 (Dirichlet) 问题的解:

$$\begin{cases} \varphi''(x) = 0, & \forall x \in (a, b), \\ \lim_{y \in (a, b), y \rightarrow x} \varphi(y) = f(x), & x = a \text{ 或 } b. \end{cases}$$

假设 g 有界连续. 那么, ψ 是如下泊松问题的解:

$$\begin{cases} \psi''(x) = -2g(x), & \forall x \in (a, b), \\ \lim_{y \in (a, b), y \rightarrow x} \psi(y) = 0, & \forall x \in \{a, b\}. \end{cases}$$

证: 对于第一个方程. 由推论3.4.4及命题3.5.1,

$$\varphi(x) = f(a)P_x(B_\tau = a) + f(b)P_x(B_\tau = b) = f(a)\frac{b-x}{b-a} + f(b)\frac{x-a}{b-a},$$

故结论成立.

对于第二个方程. 首先, 我们证明 ψ 在 (a, b) 上连续. 假设 $M > 0$ 使得对所有的 $x \in (a, b)$ 都有 $|g(x)| \leq M$. 由引理 3.5.4, $\psi(x) \leq \|g\|E_x \tau < \infty$, 其中 $\|g\| = \sup_{x \in (a, b)} |g(x)|$. 首先, 我们证明 ψ 在 (a, b) 中连续. 对任意 $x \in (a, b)$, 取 $\delta > 0$ 使得 $a < x - \delta$ 且 $x + \delta < b$. 令 $\sigma = \min\{\tau_{x-\delta}, \tau_{x+\delta}\}$. 于是, 根据强马氏性, 当 $y \in (x - \delta, x + \delta)$ 时,

$$\psi(y) = E_y \int_0^\sigma g(B_t) dt + P_y(B_\sigma = x - \delta)\psi(x - \delta) + P_y(B_\sigma = x + \delta)\psi(x + \delta),$$

由推论 3.5.6,

$$\left| E_y \int_0^\sigma g(B_t) dt \right| \leq M |E_y \sigma| \leq M \delta^2.$$

令 $y \rightarrow x$. 注意到 $P_y(B_\sigma = x - \delta)$ 在 $(x - \delta, x + \delta)$ 上连续. 因此,

$$\limsup_{y \rightarrow x} |\psi(y) - \psi(x)| \leq 2M\delta^2.$$

再令 $\delta \rightarrow 0$, 便知 ψ 在 x 连续.

其次, 我们证明对任意 $x \in (a, b)$ 都有 $\psi''(x) = g(x)$. 取 δ 与 σ 同上. 那么,

$$E_x \int_0^\tau g(B_t) dt = E_x \int_0^\sigma g(B_t) dt + E_x \int_\sigma^\tau g(B_t) dt.$$

由强马氏性,

$$E_x \int_{\sigma}^{\tau} g(B_t) dt = \frac{1}{2}(\psi(x + \delta) + \psi(x - \delta)).$$

于是

$$\psi(x + \delta) - 2\psi(x) + \psi(x - \delta) = -2E_x \int_0^{\sigma} g(B_t) dt.$$

记 $M_{\delta} := \sup_{x-\delta \leq y \leq x+\delta} |g(x) - g(y)|$. 注意到 $E_x \int_0^{\sigma} g(x) dt = g(x)E_x \sigma = \delta^2 g(x)$. 我们推出:

$$\begin{aligned} & |\psi(x + \delta) - 2\psi(x) + \psi(x - \delta) - (-2\delta^2 g(x))| \\ &= 2 \left| E_x \int_0^{\sigma} (g(B_t) - g(x)) dt \right| \leq 2E_x \int_0^{\sigma} |g(B_t) - g(x)| dt \\ &\leq 2M_{\delta} \times E_x \sigma = 2M_{\delta} \delta^2. \end{aligned}$$

令 $\delta \rightarrow 0$. 由 g 的连续性, $M_{\delta} \rightarrow 0$, 因此,

$$\lim_{\delta \rightarrow 0} \frac{\psi(x + \delta) - 2\psi(x) + \psi(x - \delta)}{\delta^2} = -2g(x).$$

又因为 φ 是连续函数, 由 Schwarz 定理(附录中的定理 4.0.22), 结论成立.

最后, 我们证明当 y 趋于边界 a 或 b 时, $\psi(y) \rightarrow 0$. M 同上, 我们有

$$|\psi(y)| \leq M E_y \tau = M(y - a)(b - y).$$

右边在 $y \rightarrow a$ 或 $y \rightarrow b$ 时都收敛到 0, 因此结论成立. \square

注 3.5.8. 事实上, g 的“有界连续”的要求可以放宽, 例如, 假设 g 有界, 且仅有有限个不连续点 x_1, \dots, x_n . 那么, 结论则改为: ψ 在 (a, b) 中连续可导, ψ' 在 (a, b) 中仅在 x_1, \dots, x_n 不可导, 且 $\psi''(x) = -2g(x)$ 对任意 $x \in (a, b) \setminus \{x_1, \dots, x_n\}$ 成立.

二、高维情形及其应用

称 \mathbb{R}^d 中的非空、有界、连通的开集为一个区域, 记为 D . 称 $\partial D = \bar{D} \setminus D$ 为 D 的边界, 其中 \bar{D} 为 D 的闭包. 在高维时, 我们需要对区域边界 ∂D 加一些所谓的“正则性”假设, 以保证当 x 趋于某个 $a \in \partial D$ 时, $B_{\tau_{\partial D}}$ 收敛到 a , 于是下面定义的 φ 和 ψ 可以连续到边界(读者可以对照一维情形的证明来理解). 正则性假设是至关重要的, 光滑的边界都是正则的, 此处我们只罗列结论, 有兴趣的同学可阅读 [9].

假设 D 是 \mathbb{R}^d 中的一个边界光滑的区域. 假设 f 是 ∂D 上的连续函数. g 是 D 上的有界连续函数, 令

$$\varphi(x) := E_x f(B_{\tau}), \quad \psi(x) := E_x \int_0^{\tau} g(B_t) dt.$$

记 $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$.

命题3.5.9. φ 是下列狄利克莱问题的解:

$$\begin{cases} \Delta\varphi(x) = 0, & \forall x \in D, \\ \lim_{y \in D, y \rightarrow x} \varphi(y) = f(x), & \forall x \in \partial D. \end{cases}$$

ψ 是下列泊松问题的解:

$$\begin{cases} \Delta\psi(x) = -2g(x), & \forall x \in D, \\ \lim_{y \in D, y \rightarrow x} \psi(y) = 0, & \forall x \in \partial D. \end{cases}$$

注3.5.10. 更一般地, f 和 g 还可以是分段连续函数. 那么, 关于 φ 的结论修改为: 若 f 在 x 连续, 则 $\lim_{y \in D, y \rightarrow x} \varphi(y) = \varphi(x)$; 而关于 ψ 的结论则修改为: $\psi(x)$ 在 D 内一阶连续可导, 在 g 的连续点 x 存在二阶导数, 并且 $\Delta\psi(x) = -2g(x)$.

注3.5.11. 满足 $\Delta\varphi = 0$ 的函数 φ 被称为调和函数. 因此, φ 是边界函数 f 在区间内部的调和延拓.

例3.5.12 (d 维标准布朗运动的常返性). 假设 $\{\vec{B}_t\}$ 是 d 维标准布朗运动. 对任意 $r > 0$, 令

$$\tau_r := \inf\{t \geq 0 : \|\vec{B}_t\| = r\},$$

其中 $\|\cdot\|$ 表示 d 为欧氏模. 也就是说, τ_r 表示 $\{\vec{B}_t\}$ 首次击中 \mathbb{R}^d 中半径为 r 的球面的时间. 假设 $\vec{x} \in \mathbb{R}^d$, $0 < \varepsilon < \|\vec{x}\| < R$, 我们考虑

$$\varphi(\vec{x}) = P_{\vec{x}}(\tau_\varepsilon < \tau_R).$$

令 $D = \{\vec{y} : \varepsilon < \|\vec{y}\| < R\}$, 则 $\varphi(\vec{x})$ 是狄利克莱问题的解. 由于高维布朗运动在旋转(即, 正交变换)下仍是布朗运动, 所以 φ 是径向对称的函数, 即, 其函数值只依赖于 $r := \|\vec{x}\|$. 令 $z = r^2$, 记 $\varphi(\vec{x}) = F(z)$, 并令 $F' = G$. 于是

$$\Delta\varphi(\vec{x}) = 4F''(z) + 2d \times F'(z) = 4G'(z) + 2d \times G(z).$$

由 $\Delta\psi$ 在区域 D 内恒等于 0, 我们推出

$$2G'(z) + d \times G(z) = 0, \quad \forall z \in (\varepsilon^2, R^2).$$

解得 $G' = Cz^{-d/2}$, 其中 C 是待定常数. 进一步, 由 $F' = G$ 推出

$$F(r^2) = \begin{cases} C_1r + C_2, & d = 1, \\ C_1 \ln r + C_2, & d = 2, \\ C_1r^{2-d}, & d \geq 3. \end{cases}$$

根据边界条件, 当 $\|\vec{x}\| = R$ 时, $F(R^2) = \varphi(\vec{x}) = 1$; 当 $\|\vec{x}\| = \varepsilon$ 时, $F(\varepsilon^2) = \varphi(\vec{x}) = 0$. 于是, 我们推出

$$\varphi(\vec{x}) = \begin{cases} (R - \|\vec{x}\|)/(R - \varepsilon), & d = 1, \\ (\ln R - \ln \|\vec{x}\|)/(\ln R - \ln \varepsilon), & d = 2, \\ (R^{2-d} - \|\vec{x}\|^{2-d})/(R^{2-d} - \varepsilon^{2-d}), & d \geq 3. \end{cases}$$

当 $d = 1$ 或 2 时, 固定 ε , 令 $R \rightarrow \infty$, 则

$$P_{\vec{x}}(\tau_{\varepsilon} < \infty) \geq P_{\vec{x}}(\tau_{\varepsilon} < \tau_R) \xrightarrow{R \rightarrow \infty} 1.$$

这表明布朗运动以概率1 击中半径为 ε 的小球. 事实上, $d = 1$ 时, 由之前的结论 $P_0(\tau_a < \infty) = 1$ 及其轨道连续性, 我们已经知道布朗运动以概率1 取遍 \mathbb{R} .

当 $d \geq 3$ 时, 首先, $\{\vec{B}_t\}$ 的第一维坐标 $\{B_t^{(1)}\}$ 是一维布朗运动, 它以概率1 能到达 $R + 1$, 此时, $\{\vec{B}_t\}$ 离开 \mathbb{R}^d 中半径为 R 的球, 因此, $P_{\vec{x}}(\tau_R < \infty) = 1$. 其次, 令 $R \rightarrow \infty$, 则 τ_R 单调上升, 将其极限记为 τ . 那么, 对任意 $T > 0$, 若 $\{\tau \leq T\}$ 发生, 则在 $[0, T]$ 这个有限的闭区间上, B_t 没有上界, 这与轨道连续性是矛盾的, 因此, 该事件的概率为0. 令 $T \rightarrow \infty$, 便可推出 $P_{\vec{x}}(\tau = \infty) = 1$. 综上, 我们得到如下结论:

$$P_{\vec{x}}(\tau_R < \infty) = 1, \forall R \geq \|\vec{x}\|, \text{ 且 } P_{\vec{x}}\left(\lim_{R \rightarrow \infty} \tau_R = \infty\right) = 1.$$

最后, 固定 ε , 对整数 $n \geq \|\vec{x}\|$, 记 $A_n = \{\tau_{\varepsilon} < \tau_n\}$. 那么,

$$P_{\vec{x}}(\tau_{\varepsilon} < \infty) = P_{\vec{x}}(\text{存在 } n \geq \|\vec{x}\| \text{ 使得 } \tau_{\varepsilon} < \tau_n) = \lim_{n \rightarrow \infty} P_{\vec{x}}(\tau_n < \tau_{\varepsilon}) = \frac{\varepsilon^{d-2}}{\|\vec{x}\|^{d-2}}.$$

因此, 布朗运动以正概率不能击中半径为 ε 的小球. 最后, 令 $\tau_{\vec{y}, \varepsilon} := \inf\{t \geq 0 : \|\vec{B}_t - \vec{y}\| \leq \varepsilon\}$, 我们可以得到如下总结:

- 当 $d = 1$ 时, 对任意 $x, y \in \mathbb{R}$, $P_x(\tau_y < \infty) = 1$ 都成立;
- 当 $d = 2$ 时, 对任意 $\vec{x}, \vec{y} \in \mathbb{R}^2$ 及任意 $\varepsilon > 0$, $P_{\vec{x}}(\tau_{\vec{y}, \varepsilon} < \infty) = 1$ 都成立;
- 当 $d = 3$ 时, 对任意 $\vec{x}, \vec{y} \in \mathbb{R}^3$ 及任意 $\varepsilon > 0$, 如下结论成立: 若 $\|\vec{x} - \vec{y}\| \leq \varepsilon$, 则 $P_{\vec{x}}(\tau_{\vec{y}, \varepsilon} < \infty) = 1$; 若 $\|\vec{x} - \vec{y}\| > \varepsilon$, 则

$$P_{\vec{x}}(\tau_{\vec{y}, \varepsilon} < \infty) = \frac{\varepsilon^{d-2}}{\|\vec{x} - \vec{y}\|^{d-2}}.$$

注3.5.13. 我们称一维布朗运动为“点”常返的, 二维布朗运动为“集合”常返的, 三维以上布朗运动为非常返的。

例3.5.14. 考虑 $D = [0, 1]$, 记 $\tau = \tau_{\partial D} = \min\{\tau_0, \tau_1\}$. 给定 $0 \leq y < z \leq 1$. 对任意 $x \in [0, 1]$, 令

$$\psi(x) := E_x \int_0^\tau \mathbf{1}_{\{y \leq B_t \leq z\}} dt.$$

它是泊松问题的解, 其中 $g(x) = \mathbf{1}_{\{y \leq x \leq z\}}$ 是分段连续函数. 因此, ψ 在 $[y, z]$ 上为二次函数, 在 $[0, y]$ 与 $[z, 1]$ 上都为线性函数. 根据边值条件 $\psi(0) = \psi(1) = 0$, 不难得到

$$\psi(x) = \begin{cases} -x^2 + ax + b, & y \leq x \leq z; \\ \frac{x}{y}h(y), & 0 \leq x \leq y; \\ \frac{1-x}{1-z}h(z), & z \leq x \leq 1, \end{cases}$$

其中 a, b 为待定常数. 进一步, 根据 ψ 的连续可导, 我们可通过 ψ 在 y 和 z 的左右导数相等解出 a, b . 具体地, a, b 满足如下方程:

$$\begin{cases} \psi'(y) = h'(y) = \frac{h(y)}{y}; \\ \psi'(z) = h'(z) = -\frac{h(z)}{1-z}. \end{cases}$$

解得

$$a = 2z - (z^2 - y^2), \quad b = -y^2.$$

进一步, 对 $0 \leq u \leq w \leq 1$, 令

$$p(u, w) = p(w, u) := 2u(1-w).$$

那么,

$$\psi(x) = \int_y^z p(x, w) dw.$$

例* 3.5.15 (区域上的格林函数). 假设 $\{\vec{B}_t\}$ 为 d 维布朗运动, D_0 为 \mathbb{R}^d 中的一个边界光滑的区域, $\vec{x} \in D$. 记 $\tau = \tau_{\partial D_0}$. 对任意 D_0 的子区域 D , 我们考虑从 \vec{x} 出发的布朗运动在到达边界之前花在 D 中的平均总时间

$$\mu(D) := E_{\vec{x}} \int_0^\tau \mathbf{1}_D(\vec{B}_t) dt = E_{\vec{x}} \int_0^\infty \mathbf{1}_{\{\vec{B}_t \in D, \tau > t\}} dt = \int_0^\infty P_{\vec{x}}(\vec{B}_t \in D, \tau > t) dt.$$

可以证明存在 $\rho_t(x, w)$ 使得对所有子区域 D 都有

$$P_{\vec{x}}(\vec{B}_t \in D, \tau > t) = \int_{w \in D} \rho_t(\vec{x}, \vec{w}) d\vec{w},$$

其中, 我们将 $dw_1 \cdots dw_d$ 简写为 $d\vec{w}$. 这个证明需要用到测度论中的知识点(Radon-Nikodym 导数), 已经超出本书的要求范围, 故此略去, 有兴趣的读者可以参阅 [1] 第四章第3节. 令

$$G(\vec{x}, \vec{w}) = \int_0^\infty \rho_t(\vec{x}, \vec{w}) dt.$$

它被称为区域 D_0 上的格林函数. 那么,

$$\mu(D) = \int_0^\infty \int_D \rho_t(\vec{x}, \vec{w}) d\vec{w} dt = \int_{\vec{w} \in D} G(\vec{x}, \vec{w}) d\vec{w}.$$

特别地, 若 $d = 1$, $D_0 = [y, z]$, 由上例知 $G(x, w) = 2 \min\{x, w\}(1 - \max\{x, w\})$.

进一步, $\mu(\cdot)$ 给出了 D_0 上的一个概率. 假设 $\vec{W} \sim \mu$. 那么, 对形如 $\mathbf{1}_D$ 的示性函数 g , 我们有

$$E_{\vec{x}} \int_0^{\tau} g(\vec{B}_t) dt = \int_{\vec{w} \in D} G(\vec{x}, \vec{w}) g(\vec{w}) d\vec{w}.$$

事实上, 利用测度论的知识点(典型方法), 上式对更广一类的函数 g 都成立, 这个证明已经超出本书的要求范围, 故此略去, 有兴趣的读者可以参阅 [1] 第一章第5节. 于是, $\psi(\vec{x}) = \int_{\vec{w} \in D} G(\vec{x}, \vec{w}) g(\vec{w}) d\vec{w}$ 给出了泊松问题 $\Delta\psi(x) = -2g(x)$, $x \in D$ 的解. 因此, $G(\vec{x}, \vec{w})$ 在微分方程中被称为泊松问题的基本解.

补充知识

假设 $\{B_t\}$ 是从 0 出发一维布朗运动. 给定 $N \geq 1$. 令 $\sigma_0 = 0$. 对于 $n \geq 1$, 递归地定义

$$\sigma_n := \inf\{t \geq \sigma_{n-1}, |B_t - B_{\sigma_{n-1}}| = \frac{1}{\sqrt{N}}\}.$$

令 $S_0 := 0$, $S_n = \sqrt{N}B_{\sigma_n}$, $n \geq 1$. 仿照例 3.5.2, 我们推出 $\{S_n^{(N)}\}$ 是一维简单随机游动. 仿照第 3.3 节, 将 $\{S_n^{(N)}\}$ 线性插值得到 $\{S_t^{(N)}\}$, 并令

$$B_t^{(N)} := \frac{S_{\lfloor Nt \rfloor}^{(N)}}{\sqrt{N}}, \quad t \geq 0.$$

那么, 对任意 $T > 0$,

$$\max_{0 \leq t \leq T} \|B_t^{(N)} - B_t\| \xrightarrow{P} 0.$$

由于该证明超过了本书的要求范围, 故此略去. 作为上式的推论, $\{B_t^{(N)} : 0 \leq t \leq T\}$ 在 $C[0, T]$ 依分布收敛于 $\{B_t : 0 \leq t \leq T\}$, 因此不变原理成立.

练习题

1. 记 $\tau := \inf\{t > 0 : |B_t| = 1\}$. 试求 τ 的方差.
2. 试求 $P_0(\tau_1 < \tau_{-1} < \tau_2)$.
3. 假设 $\eta_t = B_t + ct$. 称 $\{\eta_t : t \geq 0\}$ 为带漂移的布朗运动, 其中 $c > 0$ 是常数. 设 $\tau := \inf\{t > 0 : \eta_t \notin (a, b)\}$, $a < 0 < b$. 试求 $P_0(\eta_\tau = a)$ 和 $E_0\tau$.
4. 假设 $\{\vec{B}_t : t \geq 0\}$ 为 d 维标准布朗运动, $\vec{x} \in \mathbb{R}^d$ 且 $\|\vec{x}\| < 1$. 令 $\tau := \inf\{t \geq 0 : |\vec{B}_t| = 1\}$. 试求: $E_{\vec{x}}\tau$.

5. 试用布朗运动及其首中时来表示下列微分方程的解:

$$\begin{cases} f''(x) = g(x), & x \in (a, b), \\ f(a) = c_1, \quad f(b) = c_2, \end{cases}$$

其中 $g(x)$ 是 (a, b) 上的有界连续函数, c_1, c_2 为常数.

6. 假设 $\{\vec{W}_t : t \geq 0\}$ 为二维标准布朗运动. $L \subseteq \mathbb{R}^2$, $\sigma = \inf\{t \geq 0 : \vec{W}_t \in L\}$. 在以下情形中, 证明: $P_0(\sigma < \infty) = 1$. (1) L 是一条直线; (2) L 是一条线段; (3) L 是一条曲线.

7. 假设 $d \geq 3$, $\{\vec{B}_t\}$ 是 d 为布朗运动. 证明:

$$P_0\left(\lim_{t \rightarrow \infty} \|\vec{B}_t\| = \infty\right) = 1.$$

(注: 这也表明三维以上布朗运动是非常返的.)

§3.6 布朗桥与OU 过程

一、布朗桥

假设 $\{B_t : t \geq 0\}$ 是一维标准布朗运动, ($B_0 = 0, B_t^2 = t$). 令

$$X_t = B_t - tB_1, \quad 0 \leq t \leq 1.$$

对任意 $0 \leq t \leq s \leq 1$,

$$\begin{aligned} EX_t X_s &= E(B_t - tB_1)(B_s - sB_1) \\ &= EB_t B_s - tEB_s B_1 - sEB_t B_1 + stEB_1^2 \\ &= t - 2st + st = t(1-s). \end{aligned}$$

定义3.6.1. 假设 $\{W_t : 0 \leq t \leq 1\}$ 是一个轨道连续的高斯过程, 满足:

$$EW_t = 0, \quad \forall t \in [0, 1]; \quad EW_t W_s = t(1-s), \quad \forall 0 \leq t \leq s \leq 1.$$

那么, 称 $\{W_t : 0 \leq t \leq 1\}$ 为布朗桥.

按照定义, $\{X_t : 0 \leq t \leq 1\}$ 是布朗桥. 注意到 $\{X_t : 0 \leq t \leq 1\} \cup \{B_1\}$ 也是高斯系, 且

$$EX_t B_1 = EB_t B_1 - tEB_1^2 = t - t = 0, \quad \forall t \in [0, 1].$$

我们推出: $\{X_t : 0 \leq t \leq 1\}$ 与 $\{B_1\}$ 相互独立, 因为高斯系中, 不相关等价于相互独立(命题 3.1.2). 在 $\{B_1 = 0\}$ 的条件下, $B_t = X_t$ 对所有 $t \in [0, 1]$ 都成立. 因此, 对任意 $n \geq 1$, $0 < t_1 < \dots < t_n < 1$, $(X_{t_1}, \dots, X_{t_n})$ 的联合密度为:

$$\hat{p}_{t_1, \dots, t_n}(x_1, \dots, x_n) = \frac{p_{t_1, \dots, t_n, 1}(x_1, \dots, x_n, 0)}{p_1(0)},$$

其中, $p_{s_1, \dots, s_m}(y_1, \dots, y_m)$ 为布朗运动的有限维联合分布, 表达式见 (3.2.1).

固定 $t \in (0, 1)$, 令

$$Y_u := X_u - \frac{u}{t}X_t, \quad u \in [0, t], \quad Z_s := X_s - \frac{1-s}{1-t}X_t, \quad s \in [t, 1].$$

那么, $\{X_v, v \in [0, 1]\} \cup \{Y_u, u \in [0, t]\} \cup \{Z_s, s \in [t, 1]\}$ 是高斯系, 其中所有随机变量的期望都是0, 协方差如下: 对任意 $u \leq t \leq s$,

$$EY_u X_s = EX_u X_s - \frac{u}{t} EX_t X_s = u(1-s) - \frac{u}{t} \times t(1-s) = 0,$$

因此, $\{Y_u, u \in [0, t]\}$ 与 $\{X_t\} \cup \{Z_s, s \in [t, 1]\}$ 相互独立; 又

$$EX_u Z_s = EX_u X_s - \frac{1-s}{1-t} EX_u X_t = u(1-s) - \frac{1-s}{1-t} \times u(1-t) = 0,$$

因此, $\{Z_s, s \in [t, 1]\}$ 与 $\{X_t\}$ 这三个高斯系相互独立. 从而, $X_t, \{Y_u, u \in [0, t]\}, \{Z_s, s \in [t, 1]\}$ 是相互独立的.

下面, 我们考察 $\{Y_u, u \in [0, t]\}$. 注意到

$$Y_u = X_u - \frac{u}{t}X_t = B_u - uB_1 - \frac{u}{t}(B_t - tB_1) = B_u - \frac{u}{t}B_t.$$

它满足: $EY_u = 0, u \in [0, t]$; 且

$$EY_s Y_u = \frac{s(t-u)}{t}, \quad \forall 0 \leq s \leq u \leq t.$$

定义3.6.2. 假设 $\{W_t : 0 \leq t \leq T\}$ 是一个轨道连续的高斯过程, 满足:

$$EW_t = 0, \forall t \in [0, T]; \quad EW_t W_s = \frac{t(T-s)}{T}, \quad \forall 0 \leq t \leq s \leq T.$$

那么, 称 $\{W_t : 0 \leq t \leq T\}$ 为(长度为 T 的)布朗桥.

例3.6.3. 之前定义的 $\{Y_u, u \in [0, t]\}$ 和 $\{Z_s, s \in [t, 1]\}$ 都是布朗桥, 长度分别为 t 和 $1-t$.

二、Ornstein-Uhlenbeck 过程

假设 $\{B_t : t \geq 0\}$ 是布朗运动. 令

$$X_t = e^{-\alpha t} B_{e^{2\alpha t}}, \quad \forall t \in \mathbb{R}.$$

对任意 $s \in \mathbb{R}, t > 0$,

$$\begin{aligned} X_{t+s} &= e^{-\alpha(t+s)} B_{e^{2\alpha(t+s)}} \\ &= e^{-\alpha(t+s)} B_{e^{2\alpha s}} + e^{-\alpha(t+s)} (B_{e^{2\alpha(t+s)}} - B_{e^{2\alpha s}}). \end{aligned}$$

记

$$Y := e^{-\alpha(t+s)} (B_{e^{2\alpha s}} - B_{e^{2\alpha s}})$$

则, Y 与 $\{X_u : 0 \leq u \leq s\}$ 相互独立, $Y \sim N(0, 1 - e^{-2\alpha t})$, 且

$$X_{t+s} = e^{-\alpha t} X_s + Y.$$

. 将 $N(0, 1)$ 的密度记为 ϕ , 令

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

令

$$q_t(x, y) = \frac{1}{\sqrt{2\pi(1 - e^{-2\alpha t})}} \exp \left\{ -\frac{1}{2(1 - e^{-2\alpha t})} (y - e^{-\alpha t} x)^2 \right\}.$$

那么, 对任意 $n \geq 1, 0 = t_0 < t_1 < \cdots < t_n < 1$, $(X_0, X_{t_1}, \dots, X_{t_n})$ 的联合密度为:

$$q_{0,t_1,\dots,t_n}(x_0, x_1, \dots, x_n) = \phi(x_0) \prod_{i=1}^n q_{t_i-t_{i-1}}(x_{i-1}, x_i).$$

定义3.6.4. 称以如上定义的 $q_t(x, y)$, $t \geq 0$, $x, y \in \mathbb{R}$ 为转移概率密度的时齐马氏过程为 *Ornstein-Uhlenbeck* 过程, 简称OU 过程.

注意到 $X_t \sim N(0, 1)$ 对任意 $t \geq 0$ 都成立. 这表明, $N(0, 1)$ 是OU 过程的不变分布. 那么, 不难验证,

$$\phi(x)q_t(x, y) = \phi(y)q_t(x, y), \quad \forall t \geq 0, \forall x, y \in \mathbb{R}.$$

换句话说, $N(0, 1)$ 还是可逆分布.

例3.6.5. 令 $W_0 := 0$, $W_t := tB_{1/t}$, $t > 0$. 根据推论 3.4.11, $\{W_t\}$ 是布朗运动. 令 $Y_t = X_{-t}$. 那么, 对任意 $t \in \mathbb{R}$,

$$Y_t = e^{\alpha t}B_{e^{-2\alpha t}} = e^{\alpha t} \times \frac{1}{e^{2\alpha t}}W_{e^{2\alpha t}} = e^{-\alpha t}W_{e^{2\alpha t}}.$$

因此, $\{Y_t\}$ 也是OU 过程, 且它与 $\{X_t\}$ 同分布.

练习题

以下假设 $\{B_t, t \geq 0\}$ 是标准布朗运动, $B_0 = 0$.

1. 令 $Y_t = e^{B_t}$, 称 $\{Y_t, t \geq 0\}$ 为几何布朗运动. 假设 $y > 0$. 试求:

$$\lim_{h \rightarrow 0+} \frac{E(Y(t+h) - Y(t)|Y(t) = y)}{h} \quad \text{和} \quad \lim_{h \rightarrow 0+} \frac{E((Y(t+h) - Y(t))^2|Y(t) = y)}{h}.$$

2. 令 $X_t = B_t - [B_t]$, 称 $\{X_t : t \geq 0\}$ 为圆周上的布朗运动. 证明: $\{X_t\}$ 是时齐马氏过程, 并求其转移密度函数.

3. 假设 $\{\vec{B}_t : t \geq 0\}$ 为 d 维标准布朗运动. 证明: $\{\|\vec{B}_t\|, t \geq 0\}$ 是马氏过程. (注: 它被称为贝塞尔(Bessel)过程.)

4. 对任意 $R \geq 0$, $T > 0$, 令

$$W_t := B_{R+t} - \left(\frac{R-t}{R}B_R + \frac{t}{R}B_{R+T} \right), \quad 0 \leq t \leq T.$$

证明: $\{W_t, 0 \leq t \leq T\}$ 是长度为 T 的布朗桥, 并且它与 (B_R, B_{R+T}) 相互独立.

5. 假设 $\{W_t : 0 \leq t \leq T\}$ 是长度为 T 的布朗运动或布朗桥, $n \geq 1$, $t_0 := 0 < t_1 < \dots < t_n < t_{n+1} := T$. 对 $r = 1, 2, \dots, n+1$, 令 $T_r = T_{r+1} - T_r$,

$$X_t^{(r)} := X_{t_{r-1}+t} - \left(\frac{T_r-t}{T_r}X_{t_{r-1}} + \frac{t}{T_r}X_{t_r} \right), \quad 0 \leq t \leq T_r.$$

证明: (1) $\{X_t^{(r)}, 0 \leq t \leq T_r\}$ 是长度为 T_r 的布朗桥, $r = 1, 2, \dots, n+1$; (2) $\{X_t^{(1)}, 0 \leq t \leq T_1\}, \{X_t^{(2)}, 0 \leq t \leq T_2\}, \dots, \{X^{(n+1)}, 0 \leq t \leq T_{n+1}\}, (W_{t_1}, \dots, W_{t_n}, W_{t_{n+1}})$ 这 $n+2$ 个高斯系相互独立.

6. 假设 $\{X_t, 0 \leq t \leq 1\}$ 是长度为1的布朗桥. 令

$$X_t^{(T)} := \sqrt{T} X_{t/T}, 0 \leq t \leq T.$$

那么, $\{X_t^{(T)}, 0 \leq t \leq T\}$ 是长度为T的布朗桥.

7. 令 $U_t := (1-t)B_{t/(1-t)}, 0 \leq t < 1; U_1 := 0$. 证明: $\{U_t : 0 \leq t \leq 1\}$ 是布朗桥.

8. 假设 $\{U_t : 0 \leq t \leq 1\}$ 是布朗桥. 令 $W_t = (t+1)U_{t/(t+1)}, 0 \leq t \leq 1$. 证明: $\{W_t : t \geq 0\}$ 是标准布郎运动.

9. 设 $\{X_t : 0 \leq t \leq 1\}$ 是布朗桥. 证明:

$$P\left(\max_{0 \leq t \leq 1} X_t > b\right) = e^{-2b^2}, \quad \forall b > 0.$$

10. 假设 $q_t(x, y)$ 是OU过程的转移概率密度, $\phi(x)$ 是标准正态分布的密度函数. 证明:

$$\lim_{t \rightarrow \infty} q_t(x, y) = \phi(y), \quad \text{且} \quad \int q_t(x, y) \phi(x) dx = \phi(y).$$

§3.7 随机积分与随机微分方程简介

假设 $\{f_t : t \geq 0\}$ 是一个随机过程. 在本节, 我们将介绍随机积分 $\int_0^T f_t dB_t$ 的定义和性质. 当然, 这需要 $\{f_t\}$ 满足一定的条件. 给定 t . 仿照黎曼积分的定义, 直观上我们应该将 $[0, T]$ 进行划分, 令

$$\Delta = \Delta_T : 0 = t_0 < t_1 < \dots < t_n = T.$$

然后, 任取 $s_i \in [t_i, t_{i+1}]$, $i = 0, 1, \dots, n$, 令

$$X_\Delta := \sum_{i=0}^{n-1} f_{s_i}(B_{t_{i+1}} - B_{t_i});$$

最后, 令

$$|\Delta| := \max_{0 \leq i \leq n-1} (t_{i+1} - t_i),$$

并将 $\int_0^T f_t dB_t$ 定义为 $\lim_{\Delta \rightarrow 0} X_\Delta$.

然而, 这样的定义会遇到两个本质性的困难. 一方面, 如前所述, 布朗轨道几乎必然是处处不可微的. 这导致上述第三步拟定义的极限在几乎必然收敛的意义下不存在. 另一方面, 我们将看到: 在 $[t_i, t_{i+1}]$ 中选择不同的 s_i 将带来不同的极限. 本节介绍的伊藤积分是选择 $s_i = t_i$. 此时, X_Δ 在 L^2 意义下收敛.

定义3.7.1. 假设 X, X_1, X_2, \dots 是一列随机变量, 当 $n \rightarrow \infty$ 时 $E(X_n - X)^2 \rightarrow 0$, 那么我们称 X_n 依 L^2 收敛于 X , 记为 $X_n \xrightarrow{L^2} X$.

假设:

- (1) 对任意 $t \geq 0$, f_t 的值仅依赖于 $\{B_s : s \leq t\}$;
- (2) f_t 关于 t 连续;
- (3) 对任意 $T > 0$, $\lim_{|\Delta_T| \rightarrow 0} \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} E(f_u - f_{t_i})^2 du = 0$.

那么, 可以证明存在一个轨道连续的过程 $\{X_t\}$ 使得:

$$\sum_{i=0}^{n-1} f_{t_i}(B_{t_{i+1}} - B_{t_i}) \xrightarrow{L^2} X_t, \quad \forall t \geq 0.$$

此时, X_t 也被记为 $\int_0^t f_u dB_u$.

定义3.7.2. 称 X_t 为 $\{f_t\}$ 的伊藤积分.

可以证明随机积分满足如下性质:

$$\int_0^t (f_u + g_u) dB_u = \int_0^t f_u dB_u + \int_0^t g_u dB_u;$$

$$E \int_0^t f_u dB_u = 0;$$

$$E \int_0^s f_u dB_u \int_0^t g_u dB_u = \int_0^s E(f_u g_u) du, \quad \forall 0 \leq s \leq t.$$

例3.7.3 (布朗轨道的时间变换). 假设 f_t 是确定性函数, 则 $X_t = \int_0^t f_u dB_u$ 是高斯过程. 记 $\varphi(t) = \int_0^t f_u^2 du$. 根据定义, φ 是单调上升函数. 假设 φ 严格单调上升, 其逆函数记为 ψ . 令 $Y_t = X_{\psi^{-1}(t)}$, 则 $\{Y_t\}$ 是轨道连续的高斯过程; 对任意 $t \geq 0$, $EY_t = 0$; 对任意 $s \geq t \geq 0$, 则

$$EY_t Y_s = EX_{\psi(t)} X_{\psi^{-1}(s)} = \varphi(\psi(t)) = t.$$

这表明 $\{Y_t\}$ 是标准布朗运动.

例3.7.4. 求 $\int_0^t B_u dB_u$.

解: 考虑 $[0, t]$ 的划分 $\Delta: 0 = t_0 < t_1 < \dots < t_n = t$. 直接计算得到

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} E(B_u - B_{t_i})^2 du = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} (u - t_i) du = \sum_{i=0}^{n-1} \frac{1}{2} (t_{i+1} - t_i)^2 \leq \frac{1}{2} |\Delta| t \xrightarrow{|\Delta| \rightarrow 0} 0.$$

下面, 我们计算 $\sum_{i=0}^{n-1} B_{t_i} (B_{t_{i+1}} - B_{t_i})$. 注意到 $(x+y)^2 = x^2 + y^2 + 2xy$. 取 $x = B_{t_i}$, $y = B_{t_{i+1}} - B_{t_i}$ 知 $2B_{t_i} (B_{t_{i+1}} - B_{t_i}) = B_{t_{i+1}}^2 - B_{t_i}^2 - (B_{t_{i+1}} - B_{t_i})^2$. 从而

$$\sum_{i=0}^{n-1} 2B_{t_i} (B_{t_{i+1}} - B_{t_i}) = B_t^2 - B_0^2 - \sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2.$$

往证 $\sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2$ 依 L^2 收敛于 t . 注意到 $\sum_{i=0}^{n-1} E(B_{t_{i+1}} - B_{t_i})^2 = \sum_{i=0}^{n-1} (t_{i+1} - t_i) = t$. 于是

$$\begin{aligned} E \left(\sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2 - t \right)^2 &= \text{Var} \left(\sum_{i=0}^{n-1} (B_{t_{i+1}} - B_{t_i})^2 \right) \\ &= \sum_{i=0}^{n-1} \text{Var} ((B_{t_{i+1}} - B_{t_i})^2) = \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 \text{Var}(Z^2) \leq \text{Var}(Z^2) \times |\Delta| \times t \rightarrow 0, \end{aligned}$$

其中 $Z \sim N(0, 1)$, 并且在第二个等号中, 我们用到了布朗运动的独立增量性. 因此,

$$\int_0^t B_u dB_u = \frac{1}{2} (B_t^2 - t).$$

注3.7.5. 上例表明 $\sum_{i=1}^n B_{t_{i+1}} (B_{t_{i+1}} - B_{t_i}) = \sum_{i=0}^{n-1} B_{t_i} (B_{t_{i+1}} - B_{t_i}) + \sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2$ 依 L^2 收敛于 $\frac{1}{2} (B_t^2 - t) + t = \frac{1}{2} (B_t^2 + t)$. 这表明选取不同的 s_i , 会使得 $\sum_{i=0}^{n-1} f_{s_i} (B_{t_{i+1}} - B_{t_i})$ 有不同的极限.

我们已经说明 s_i 不能在 $[t_i, t_{i+1}]$ 中随意选取, 否则极限不存在. 一般地, 统一取左端点 $s_i = t_i$ (或右端点 $s_i = t_{i+1}$) 时, 累计量在 L^2 中有极限, 此极限被称为

给定初值 X_0 (与布朗运动 $\{B_t\}$ 相互独立),

$$X_t := X_0 + \int_0^t \sigma_u dB_u + \int_0^t b_u du, \quad (3.7.1)$$

定义了一个轨道连续的随机过程, 它对应的随机微分形式如下:

$$dX_t = \sigma_t dB_t + b_t dt. \quad (3.7.2)$$

换句话说, X_t 满足(3.7.2) 的意思是: 存在某个 X_0 使得(3.7.1) 成立.

根据例 3.7.4, 我们知道 $X_t := B_t^2$ 是满足如下随机微分方程:

$$dX_t = 2B_t dB_t + dt.$$

如果我们令 $\varphi(x) = x^2$, 则上式可以写成

$$d\varphi(B_t) = \varphi'(B_t)dB_t + \frac{1}{2}\varphi''(B_t)dt, \quad (3.7.3)$$

这与我们在微积分中常见的公式 $d\varphi(x) = \varphi'(x)dx$ 不一致(视 $x = B_t$). 其原因在于, 如果我们对 φ 在进行泰勒展开, φ 的增量可以写为

$$\varphi(x + \Delta x) - \varphi(x) = \varphi'(x)(\Delta x) + \frac{1}{2}\varphi''(x)(\Delta x)^2 + o(\Delta x^2),$$

其中 Δx 相当于 $B_{t+1} - B_t$. 由例 3.7.4, $(\Delta x)^2$ 的累积效果不是 0 而是 t . 这是导致 (3.7.3) 成立的原因. 事实上, 如果 X_t 满足(3.7.2), $\varphi(t, x)$ 关于 t 连续可导, 关于 x 二阶连续可导, 则 $Y_t = \varphi(t, X_t)$ 满足如下随机微分方程:

$$dY_t = \frac{\partial \varphi}{\partial x}(t, X_t)dX_t + \frac{1}{2}\frac{\partial^2 \varphi}{\partial x^2}(t, X_t)(dX_t)^2 + \frac{\partial \varphi}{\partial t}(t, X_t)dt \quad (3.7.4)$$

其中 $(dX_t)^2 := \sigma_t^2 dt$. 上式被称为是 Itô 公式, 它是求解随机微分方程的重要工具. 所谓求解随机微分方程, 就是要对于一个 X_t 的随机微分方程, 给出一个形如(3.7.1) 的解, 满足: $\{\sigma_t\}$, $\{b_t\}$ 都是不依赖于 $\{X_t\}$ 的已知的随机过程. 下面, 我们给出一些例子.

例3.7.6. (带漂移项的布朗运动) 方程 $dX_t = \sigma dB_t + \mu dt$ 的解为 $X_t = X_0 + \sigma B_t + \mu t$.

例3.7.7 (OU 过程). 求随机微分方程

$$dX_t = \sigma dB_t - bX_t dt$$

的解. 请注意, 不能认为该方程的解是 $X_0 + \sigma B_t - \int_0^t bX_u du$, 因为 $b_u = bX_u$ 是一个未知的随机过程.

解: 将随机微分方程改写为 $dX_t + bX_t dt = \sigma dB_t$. 令 $\varphi(t, x) = e^{bt}x$, $Y_t = \varphi(t, X_t)$. 则 Y_t 满足方程

$$dY_t = e^{bt}(dX_t + bX_t dt) = e^{bt}\sigma dB_t.$$

因此 $Y_t = Y_0 + \sigma \int_0^t e^{bs} dB_s$. 于是

$$X_t = e^{-bt}(Y_0 + \sigma \int_0^t e^{bs} dB_s) = e^{-bt}Y_0 + \sigma \int_0^t e^{-b(t-s)} dB_s,$$

其中我们用到 $Y_0 = \varphi(0, X_0) = X_0$.

特别地, 取 $\sigma = \sqrt{2\alpha}$, $b = \alpha$, 其中 α 为严格正的常数. 则 $dX_t = \sqrt{2\alpha} dB_t - \alpha X_t dt$ 的解为

$$X_t = e^{-\alpha t} X_0 + \sqrt{2\alpha} e^{-\alpha t} \int_0^t e^{\alpha u} dB_u.$$

假设 $X_0 \sim N(0, 1)$, 并且与 $\{B_t\}$ 相互独立. 那么, $\{X_t\}$ 是轨道连续的高斯过程; $EX_t = 0$ 对任意 $t \geq 0$ 成立; $EX_s X_t = e^{-\alpha(t-s)}$ 对任意 $t \geq s \geq 0$ 成立. 这即是说: $\{X_t\}$ 是 OU 过程.

例3.7.8 (Black-Scholes 模型). 在金融中, 用 X_t 表示 t 时刻某产品的价格, 那么, $X_t \geq 0$ 对所有 $t \geq 0$ 都成立, 并且 $\{X_t\}$ 满足如下随机微分方程:

$$dX_t = X_t(\sigma dB_t + \mu dt),$$

其中, σ 是波动率, μ 是平均收益率. 令 $Y_t = \ln X_t$, 那么,

$$\begin{aligned} dY_t &= \frac{1}{X_t} dX_t - \frac{1}{2} \cdot \frac{1}{X_t^2} (dX_t)^2 = (\sigma dB_t + \mu dt) - \frac{1}{2} \cdot \frac{1}{X_t^2} (\sigma X_t)^2 dt \\ &= \sigma dB_t + (\mu - \frac{1}{2}\sigma^2)dt. \end{aligned}$$

因此, $Y_t = Y_0 + \sigma B_t + (\mu - \frac{1}{2}\sigma^2)t$. 从而, $X_t = X_0 e^{\sigma B_t + (\mu - \frac{1}{2}\sigma^2)t}$.

例3.7.9 (分部积分). 记 $X_t = \int_0^t f_u dB_u$, 求: $\int_0^t X_u du$.

解: 取 $\varphi(t, x) = tx$, 并令 $Y_t = \varphi(t, X_t)$. 则 $Y_0 = 0$ 且 $dY_t = X_t dt + t dX_t = X_t dt + t f_t dB_t$. 于是 $\int_0^t X_u du = Y_t - \int_0^t u f_u dB_u$, 即

$$\int_0^t \int_0^u f_s dB_s du = t X_t - \int_0^t u f_u dB_u.$$

直观上, 这就是分部积分公式 $\int_0^t X_u du = u X_u|_0^t - \int_0^t u dX_u$.

仿照(3.7.4), 我们可以考虑多个过程的函数. 假设

$$dX_{i,t} = \sigma_{i,t} dB_t + b_{i,t} dt, \quad i = 1, \dots, d.$$

记 $\vec{X}_t = (X_{1,t}, \dots, X_{d,t})$. 则 $Y_t = \varphi(t, \vec{X}_t)$ 满足如下随机微分方程:

$$dY_t = \sum_{i=1}^d \frac{\partial \varphi}{\partial x_i}(t, \vec{X}_t) dX_{i,t} + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(t, \vec{X}_t) (dX_{i,t})(dX_{j,t}) + \frac{\partial \varphi}{\partial t}(t, \vec{X}_t) dt,$$

其中 $(dX_{i,t})(dX_{j,t}) = \sigma_{i,t} \sigma_{j,t} dt$.

例3.7.10. 求 $\int_0^t f_u dB_u \int_0^t g_u dB_u$ 满足的随机微分方程.

解: 令 $X_t = \int_0^t f_u dB_u$, $Y_t = \int_0^t g_u dB_u$, $Z_t = X_t Y_t$. 则

$$dZ_t = Y_t dX_t + X_t dY_t + dX_t dY_t = (f_t Y_t + X_t g_t) dB_t + f_t g_t dt.$$

进一步, 写成积分形式我们有

$$Z_t = \int_0^t (f_u Y_u + X_u g_u) dB_u + \int_0^t f_u g_u du.$$

因此, $EZ_t = \int_0^t f_u g_u du$.

例3.7.11. (例 3.7.7 续). 假设 $dX_t = \sigma dB_t - bX_t dt$, $dY_t = X_t dt$, $X_0 = x_0$, $Y_0 = y_0$. 求解 Y_t .

解: 根据例 3.7.7, $X_t = e^{-bt} X_0 + \sigma e^{-bt} \int_0^t e^{bs} dB_s$. 令 $Z_t = \int_0^t e^{bs} dB_s$, 那么

$$Y_t = y_0 + x_0 \frac{1 - e^{-bt}}{b} + \sigma \int_0^t e^{-bu} Z_u du. \quad (3.7.5)$$

其中 $dZ_t = e^{bt} dB_t$. 下面我们计算 $\int_0^t e^{-bu} Z_u du$. 令 $\varphi(t, z) = e^{-bt} z$, 则 $W_t = \varphi(t, Z_t)$ 满足 $W_0 = 0$ 以及方程 $dW_t = e^{-bt} dZ_t - be^{-bt} Z_t dt = dB_t - be^{-bt} Z_t dt$. 于是

$$\int_0^t e^{-bu} Z_u du = \frac{1}{b} (B_t - W_t) b = \frac{1}{b} (B_t - e^{-bt} Z_t) = \frac{1}{b} \int_0^t (1 - e^{-b(t-s)}) dB_u.$$

结合 Z_t 的定义与(3.7.5), 我们有

$$Y_t = y_0 + \frac{1}{b} (1 - e^{-bt}) x_0 - \frac{\sigma}{b} \int_0^t (1 - e^{-b(t-s)}) dB_u.$$

这即是 Y_t 的解.

练习题

1. 设 Δ 为 $[0, t]$ 的划分. 证明: 当 $|\Delta| \rightarrow 0$ 时,

$$\sum_{i=0}^{n-1} B\left(\frac{t_i + t_{i+1}}{2}\right) (B(t_{i+1}) - B(t_i))$$

在 L^2 中的极限存在, 并求此极限. 此极限被称为Stratonovich积分.

2. 在给定 $B_t = x$ 的条件下求 $Y_t = \int_0^t B_s ds$ 的分布.
3. 在干嘛??? 假设 $\{S_n : n \geq 0\}$ 是 \mathbb{Z} 上的简单随机游动, $S_0 = 0$. 记 $W_n = \sum_{m=1}^n S_m^2 / n^2$. (1) 证明 W_n 依分布收敛, (2) 将其极限记为 W , 试求 EW , (3) 你能推广此命题吗? 请详述之.

第四章 附录

引理4.0.12. 假设 \mathcal{E} 对交运算封闭, $\Omega \in \mathcal{E}$, 且 $\mathcal{F} = \sigma(\mathcal{E})$, P 与 \tilde{P} 都是 \mathcal{F} 上的概率. 若 $P(E) = \tilde{P}(E)$ 对所有 $E \in \mathcal{E}$ 均成立, 则 $P(A) = \tilde{P}(A)$ 对所有 $A \in \mathcal{F}$ 均成立.

此引理是[1]中的命题2.3.1的一个推论.

定理4.0.13. 假设 $(\Omega_n, \mathcal{F}_n, P_n)$, $n = 1, 2, \dots$ 是一列概率空间. 则在 $(\prod_{n=1}^{\infty} \Omega_n, \prod_{n=1}^{\infty} \mathcal{F}_n, \prod_{n=1}^{\infty} P_n)$ 上存在唯一的概率 P 使得对任意 $n = 1, 2, \dots$, $A_1 \in \mathcal{F}_1, \dots, A_n \in \mathcal{F}_n$, 均有

$$P\left(\prod_{k=1}^n A_k \times \prod_{n=k+1}^{\infty} \Omega_k\right) = \prod_{k=1}^n P_k(A_k).$$

定理4.0.14. (Kolmogorov 相容性定理) 假设 S 为可数集, I 为指标集(可能不可数). 令

$$\Omega = S^I = \{\omega = (\omega_i)_{i \in I} : \omega_i \in S, \forall i \in I\},$$

\mathcal{F} 为有限柱集生成的 σ 代数, 即使得 $\{\omega_i = s\}$, $i \in I$, $s \in S$ 都可测的最小 σ 代数. 若

$$\{p_{i_1, \dots, i_n}(s_1, \dots, s_n) : n \geq 1; i_1, \dots, i_n \in I \text{ 互不相等}; s_1, \dots, s_n \in S\}$$

为一族相容的概率值, 即:

- (1) $p_{i_1, \dots, i_n}(s_1, \dots, s_n) \geq 0$;
- (2) $\sum_{s_1, \dots, s_n \in S} p_{i_1, \dots, i_n}(s_1, \dots, s_n) = 1$;
- (3) $\sum_{s \in S} p_{i_1, \dots, i_n, j}(s_1, \dots, s_n, s) = p_{i_1, \dots, i_n}(s_1, \dots, s_n)$, $\forall j \in S \setminus \{i_1, \dots, i_n\}$,

则在 (Ω, \mathcal{F}) 上存在唯一的概率 P 使得

$$P(\omega_{i_1} = s_1, \dots, \omega_{i_n} = s_n) = p_{i_1, \dots, i_n}(s_1, \dots, s_n).$$

定理4.0.15. (Perron-Frobenius定理) 设 $A = (a_{ij})$ 为 $n \times n$ 不可约实矩阵, 所有元素均非负, $a_{ij} \geq 0$. 则下列结论成立.

- (1) 存在一个实特征值 r , 其他特征值 λ 的模均不超过 r , 即 $|\lambda| \leq r$.
- (2) 与 r 对应的左(右)特征向量的所有元素非负.
- (3) $\min_i \sum_j a_{ij} \leq r \leq \max_i \sum_j a_{ij}$.

定理4.0.16. (有界收敛定理) 设 X_1, X_2, \dots 为一列随机变量, 依概率收敛于 X . 若存在 $M > 0$ 使得 $P(|X_n| \leq M) = 1$, $\forall n \geq 1$, 则 $\lim_{n \rightarrow \infty} EX_n = EX$.

证: 容易看出, $P(|X| \leq M) = 1$. 对任意 $\varepsilon > 0$, $E|X_n - X| = E|X_n - X|1_{|X_n - X| > \varepsilon} + E|X_n - X|1_{|X_n - X| \leq \varepsilon} \leq 2MP(|X_n - X| > \varepsilon) + \varepsilon$, 令 $n \rightarrow \infty$ 知 $\limsup_{n \rightarrow \infty} E|X_n - X| \leq \varepsilon$, $\forall \varepsilon > 0$. 从而, 结论成立. \square

定理4.0.17. (Fatou引理) 设 $i \in I$ 可数. $\lim_{n \rightarrow \infty} a_{n,i} \rightarrow a_i, \forall i$. 则 $\liminf_{n \rightarrow \infty} \sum_i a_{n,i} \geq \sum_i a_i$, 其中, n 可以为整数或实数.

证: 任取有限子集 $A \in I$. $\liminf_{n \rightarrow \infty} \sum_i a_{n,i} \geq \liminf_{n \rightarrow \infty} \sum_{i \in A} a_{n,i} = \sum_{i \in A} a_i$. 故命题成立. \square

推论4.0.18. 假设 $a_j(\varepsilon) \geq 0, \forall j$ 满足: $\liminf_{\varepsilon \rightarrow 0} a_j(\varepsilon) = a_j, \forall j$ 且 $\lim_{\varepsilon \rightarrow 0} \sum_j a_j(\varepsilon) = \sum_j a_j < \infty$. 那么, $\lim_{\varepsilon \rightarrow 0} a_j(\varepsilon) = a_j, \forall j$.

证: 任意给定 i , $\limsup_{\varepsilon \rightarrow 0} a_i(\varepsilon) = \limsup_{\varepsilon \rightarrow 0} \left(\sum_j a_j(\varepsilon) - \sum_{j \neq i} a_j(\varepsilon) \right) = \lim_{\varepsilon \rightarrow 0} \sum_j a_j(\varepsilon) - \liminf_{\varepsilon \rightarrow 0} \sum_{j \neq i} a_j(\varepsilon)$. 由Fatou 引理, $\liminf_{\varepsilon \rightarrow 0} \sum_{j \neq i} a_j(\varepsilon) \geq \sum_{j \neq i} \liminf_{\varepsilon \rightarrow 0} a_j(\varepsilon) = \sum_{j \neq i} a_j$, 从而, $\limsup_{\varepsilon \rightarrow 0} a_i(\varepsilon) \leq \sum_j a_j - \sum_{j \neq i} a_j = a_i$. 这加上 $\liminf_{\varepsilon \rightarrow 0} a_j(\varepsilon) = a_j$ 知 $\lim_{\varepsilon \rightarrow 0} a_i(\varepsilon) = a_i$.

定理4.0.19. (单调收敛定理) 设 X_1, X_2, \dots 为一列非负随机变量, 单调上升到 X , 则 $\lim_{n \rightarrow \infty} EX_n = EX$.

定理4.0.20. (Fubini定理) 若 $\iint |f(x,y)| \mu(dx)\nu(dy) < \infty$, 则

$$\iint f(x,y) \mu(dx)\nu(dy) = \int \int f(x,y) \nu(dy) \mu(dx).$$

定理4.0.21. (Borel-Cantelli 引理)

若 $\sum_n P(A_n) < \infty$, 则 $P(A_n \text{ i.o.}) = 0$;

若 $\sum_n P(A_n) = \infty$ 且 A_1, A_2, \dots 相互独立, 则 $P(A_n \text{ i.o.}) = 1$.

定理4.0.22. (Schwarz定理) 设 $f, \varphi : (a, b) \rightarrow \mathbb{R}$ 连续, 且

$$\lim_{h \rightarrow 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} = \varphi(x), \quad \forall x \in (a, b).$$

则 f 在 (a, b) 上二阶连续可微, 且 $f'' = \varphi$.

定理4.0.23. (Kolmogorov强大数定律) 假设 X_1, X_2, \dots 独立同分布, EX_1 存在. 令 $S_n = X_1 + \dots + X_n$, 那么 $S_n/n \xrightarrow{a.s.} 0$.

强大数定律的主要证明工具是如下的引理, 但证明过程已超过本教材的范围.

引理4.0.24 (Borel-Cantelli引理). 假设 A_1, A_2, \dots 是随机事件, 记 $\{A_n \text{ i.o.}\} := \cap_{N=1}^{\infty} \cup_{n=N}^{\infty} A_n$. 那么, (1) 若 $\sum_{n=1}^{\infty} P(A_n) < \infty$, 则 $P(A_n \text{ i.o.}) = 0$; (2) 若 $\sum_{n=1}^{\infty} P(A_n) = \infty$ 且 A_1, A_2, \dots 相互独立, 则 $P(A_n \text{ i.o.}) = 1$.

证：注意到事件 $\cup_{n=N}^{\infty} A_n$ 关于 N 单调下降，我们有 $P(A_n \text{ i.o.}) = \lim_{N \rightarrow \infty} P(\cup_{n=N}^{\infty} A_n)$.

(1) 对任意 $N \geq 1$, $P(A_n \text{ i.o.}) \leq P(\cup_{n=N}^{\infty} A_n) \leq \sum_{n=N}^{\infty} P(A_n)$. 令 $N \rightarrow \infty$ 知结论成立.

(2) 固定 N , 事件 $\cup_{n=N}^M A_n$ 关于 M 单调上升, 因此 $P(\cup_{n=N}^{\infty} A_n) = \lim_{M \rightarrow \infty} P(\cup_{n=N}^M A_n)$.
由 A_1, A_2, \dots 相互独立, $1 - P(\cup_{n=N}^M A_n) = P(\cap_{n=N}^M A_n^c) = \prod_{n=N}^M P(A_n^c) = \prod_{n=N}^M (1 - P(A_n))$.
注意到 $1 - p \leq e^{-p}$, $\forall p \geq 0$. 因此 $1 - P(\cup_{n=N}^M A_n) \leq e^{-\prod_{n=N}^M P(A_n)}$. 令 $M \rightarrow \infty$, 有 $1 - P(\cup_{n=N}^{\infty} A_n) = \lim_{M \rightarrow \infty} (1 - P(\cup_{n=N}^M A_n)) \leq e^{-\lim_{M \rightarrow \infty} \prod_{n=N}^M P(A_n)} = 0$, 其中, 最后一个等号用到了 $\sum_{n=1}^{\infty} P(A_n) = \infty$. 因此, $P(\cup_{n=N}^{\infty} A_n) = 1$, $\forall N$. 这表明 $P(A_n \text{ i.o.}) = 1$. \square

引理4.0.25. 设 ξ_1, ξ_2, \dots 是一列独立同分布的随机变量, $E|\xi_1| < \infty$. 则

$$P\left(\lim_{n \rightarrow \infty} \frac{\xi_n}{n} = 0\right) = 1.$$

证法一：我们可以利用强大数定律进行证明. 令 $S_n = \xi_1 + \dots + \xi_n$. 由强大数定律,
 $S_n/n \xrightarrow{\text{a.s.}} E\xi_1$. 当然, $S_{n-1}/n = \frac{S_{n-1}}{n-1} \cdot \frac{n-1}{n} \xrightarrow{\text{a.s.}} E\xi_1$. 于是 $\xi_n/n = S_n/n - S_{n-1}/n \xrightarrow{\text{a.s.}} E\xi_1 - E\xi_1 = 0$.

证法二：我们可以直接通过Borel-Cantelli引理证明该结论. 因为 $E|\xi_1| < \infty$ 所以 $\sum_n P(|\xi_n| > n\varepsilon) < \infty$, $\forall \varepsilon > 0$. 由Borel-Cantelli引理(见附录4.0.21), $P(|\xi_n| > n\varepsilon \text{ i.o.}) = 0$. 从而 $\xi_n/n \rightarrow 0$ a.s.. \square

北京大学数学学院
试用教材
2022年3月

参考文献

- [1] 程士宏,《测度论与概率论基础》,北京大学出版社, 2004.
- [2] 龚光鲁,《随机微分方程及其应用概要》,清华大学出版社, 2008.
- [3] 李贤平,《概率论基础》(第二版),高等教育出版社, 1997.
- [4] 钱敏平, 龚光鲁,《应用随机过程》, 北京大学出版社, 1998.
- [5] 吴立德, 汪嘉冈, 李贤平, 卞国瑞,《概率论》第三册《随机过程》, 高等教育出版社, 1980.
- [6] Alon, N. & Spencer, J. H., *The Probabilistic Method*, 2nd ed., Wiley, New York, 2000.
- [7] Billingsley, P., *Weak Convergence of Probability Measures*, Wiley & Sons, New York, 1968.
- [8] Chen, D. Feng, J. & Qian, M.,(1996), The metastability of exponentially perturbed Markov chains, *Science in China, Series A*, Vol.39, 7-28.
- [9] Chung, K. L., *Green, Brown and Probability & Brownian Motion on the Line*, World Scientific, Singapore, 2002.
- [10] Dembo, A., Simple random covering, disconnection, late and favorite points, in *Proceedings of the International Congress of Mathematicians*, Madrid, Spain, 2006
- [11] Doyle, P. G. & Snell, J. L., *Random Walks and Electrical Networks*, (Carus Mathematical Monographs, No 22) Mathematical Association of America, 1984. <http://www-ee.technion.ac.il/~adam/FUN/RWEN.pdf>
- [12] Durrett, R., *Essentials of Stochastic Processes*, Springer, New York 1999.
- [13] Durrett, R., *Probability: Theory and Examples*, 3rd ed., 世界图书出版社, 2007.
- [14] Durrett, R. and Liu, X. F. (1988). The contact process on a finite set. *Ann. Probab.* **16** 1158–1173; Durrett, R. and Schonmann, R. H. (1988). The contact process on a finite set II. *Ann. Probab.* **16** 1570–1583; Durrett, R., Schonmann, R. H. and Tanaka, N. I. (1989). The contact process on a finite set III: The critical case. *Ann. Probab.* **17** 1303–1321.
- [15] Harris, T. E., *The Theory of Branching Processes*, Springer, Berlin, 1963 & Dover Publications, Inc., New York, 1989.
- [16] Häggström, O., *Finite Markov Chains and Algorithmic Applications*, Cambridge University Press, Cambridge, 2002.

- [17] Hoel, P. G., Port, S. C. & Stone, C. J., *Introduction to Stochastic Processes*, Houghton Mifflin Company, Boston, 1972.
- [18] Jerison, D. & Stroock, D., Norbert Wiener, *AMS Notices* (1995), **42**(4), 430–438.
- [19] Kemeny, J. G. & Snell, J. L., *Finite Markov Chains*, Springer, New York, 1983.
- [20] Karlin, S. & Taylor, H., *A First Course in Stochastic Processes*, 2nd ed., Academic Press, 1975.
- [21] Levin, D., Peres Y., & Wilmer, E.L., *Markov Chains and Mixing Times*, American Mathematical Society, 2008. <http://www.uoregon.edu/~dlevin/MARKOV>
- [22] Liggett, T.M., *Interacting Particle Systems*, Springer, Berlin, 1985.
- [23] Montroll, E. L., & Shlesinger, M., On the wonderful world of random walk, in “*From stochastic to hydrodynamics*”, Lebowitz, J. L. & Montroll, E. W., editors, North-Holland Physics Publishing, 1984.
- [24] Morters, P. & Peres, Y., *Brownian Motion*, 书稿, 可下载.
<http://www.stat.berkeley.edu/users/peres/bmbook.pdf>
- [25] Norris, J. R., *Markov Chains*, Cambridge University Press, 1997.
- [26] Ross, S. M., *Stochastic Processes*, John Wiley & Sons, 1983, (有中译本, 随机过程, 何声武, 谢盛荣, 程依明译, 中国统计出版社, 1997).
- [27] Spitzer, F., *Principles of Random Walk*, 2nd Edition, GTM34, Springer, New York, 1976. 世界图书出版社, 200X.
- [28] Stroock, D.W., *An Introduction to Markov Processes*, Springer, New York, 2005.
- [29] Woess, W., *Random Walks on Infinite Graphs and Groups*, Cambridge University Press, Cambridge, 2000.