

南开大学研究生课程讲义

随机过程

January 9, 2010



This document is typeset using lua \TeX .

\TeX and $\mathcal{A}\mathcal{M}\mathcal{S}\text{-}\text{\TeX}$ are trademarks of the American Mathematical Society; METAFONT is a trademark of Addison–Wesley Publishing Company; PostScript, Portable Document Format and Acrobat are trademarks of Adobe Systems Incorporated; all other product names are trademarks of their producers.

©2008--2009 Longmin Wang

Contents

1	随机过程简介	
1.1	随机过程的分类	1
1.2	简单随机游动	2
2	离散时间马尔可夫链	
2.1	定义与转移矩阵	5
2.2	马氏链状态的分类	6
2.2.1	互通性 6	
2.2.2	周期性 7	
2.3	常返性	7
2.4	极限定理与平稳分布	11
2.4.1	离散更新方程 11	
2.4.2	定理 2.7 的证明 15	
2.5	一些例子	17
2.6	吸收概率	24
2.7	常返性准则	27
2.8	离散排队模型	29
2.9	马氏链的模拟	33
2.10	蒙特卡罗	34
3	Poisson 过程	
3.1	定义	38
3.2	另一个等价定义	40
3.3	Poisson过程的其它性质	43
3.3.1	顺序统计量 43	
3.3.2	过程的稀疏 47	
3.4	复合Poisson过程及应用	48
3.4.1	复合Poisson过程 48	
3.4.2	古典风险模型 49	
3.5	Poisson 过程的其它扩展	51
3.5.1	非齐次 Poisson 过程 51	
3.5.2	条件 Poisson 过程 52	
3.5.3	Poisson 随机测度 52	
4	连续时间马氏链	
4.1	定义	54
4.1.1	马氏性与等价条件 54	
4.1.2	转移概率 56	
4.2	标准转移矩阵	58
4.3	向前与向后方程组	62
4.4	Q 矩阵的概率意义	67
4.5	嵌入链	72
4.6	马氏链的构造	75
4.7	生灭过程	83
4.8	常返性与不变测度	86

5	更新理论	
5.1	Laplace 变换	94
5.2	基本概念	96
5.3	更新方程	100
5.4	平稳更新过程	108
5.5	更新定理	112
5.5.1	直接黎曼可积 113	
5.5.2	更新定理 116	
5.5.3	离散更新理论 120	
5.5.4	Improper Renewal Equation 122	
5.6	再生过程	124
5.7	其它更新过程	127
5.7.1	有偿更新过程 127	
5.7.2	可终止更新过程 132	
6	平稳过程简介	
6.1	谱分解	134
6.1.1	相关函数的谱分解 134	
6.1.2	均方积分 137	
6.1.3	过程的谱分解 139	
6.2	均方遍历性	140
6.3	平稳序列的线性预测	143
6.4	强平稳过程与遍历理论	144

1 随机过程简介

设 $T \subset \mathbb{R}$. 对 $t \in T$, X_t 是定义在概率空间 (Ω, \mathcal{F}, P) 取值于 E 的随机变量. 我们称这族随机变量 $\{X_t, t \in T\}$ 为一个随机过程. 我们可以简记为 X_t , 有时也写为 $X(t)$.

在本书中, 如无特别说明, 我们都假设 T 为 \mathbb{N} 或 $[0, \infty)$, 状态空间 E 为 \mathbb{Z} 或 \mathbb{R} .

1.1 随机过程的分类

若状态空间 E 离散, 如 $E = \{0, 1, 2, \dots\}$, 我们称 X_t 为离散状态随机过程.

依照时间指标集 T 离散或连续, 我们称过程 X_t 为离散或连续时间随机过程.

依照随机变量 X_t 之间不同的独立性关系, 我们可以分为:

- 平稳独立增量过程: 对任意的 $t_1 < t_2 < \dots < t_n$,

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}}$$

独立, 且对任意的 $h > 0, t \in T$, $X_{t+h} - X_t$ 的分布不依赖于 t . 如布朗运动与 Poisson 过程.

当 $T = \mathbb{N}$ 时,

$$X_n = Z_0 + Z_1 + \dots + Z_n, \quad n = 0, 1, \dots,$$

其中 $\{Z_n, n \geq 0\}$ 是独立同分布随机变量列. 此时我们称 X_n 为随机游动.

- 马尔可夫过程: 给定 $X_t, X_s, s > t$ 的取值不依赖于 $X_u, u < t$; 即已知当前, 将来与过去无关. 当状态空间离散时, 我们也称为马尔可夫链, 简称马氏链. 马氏链又可分为离散时间和连续时间马氏链.
- 更新过程: 设 $\{X_n, n \geq 1\}$ 独立同分布, 称 $S_n = X_1 + \dots + X_n$ 为更新序列, 通常我们理解为事件发生的时刻. 在 t 时刻之前发生的事件次数 $N(t)$ 被称为更新 (计数) 过程, 即

$$N(t) = \#\{n, S_n \leq t\} = \sum_{n=0}^{\infty} 1_{[0, t]}(S_n).$$

- 平稳过程: 若对于任意的 $h > 0, t_1, \dots, t_n \in T$,

$$(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}) \text{ 与 } (X_{t_1}, X_{t_2}, \dots, X_{t_n})$$

具有相同的分布, 我们称 X_t 为严平稳过程. 若对于任意的 $h > 0, t \in T$, X_t 与 X_{t+h} 具有二阶矩, 且 $\text{Cov}(X_t, X_{t+h}) = E(X_t X_{t+h}) - E(X_t)E(X_{t+h})$ 仅依赖于 h , 则称 X_t 为宽平稳过程. 在考虑平稳过程时, 我们经常取 $T = \mathbb{Z}$ 或 $T = \mathbb{R}$.

- 点过程 设 (S, \mathcal{A}) 是一个可测空间, 点过程 N 是以所有的 $A \in \mathcal{A}$ 为指标, 以 $\mathbb{N} \cup \{\infty\}$ 为状态空间的随机过程. 我们可以理解为在 S 中随机散落着一些点, 对 $A \in \mathcal{A}$, $N(A)$ 表示分布在 A 中点的个数. 通常我们取 S 为 \mathbb{R}^d 的子集. 设 ν 是 (S, \mathcal{A}) 上的 Radon 测度, 称 N 为 *Poisson* 点过程, 若它满足:

- (i) 对任意的 $A \in \mathcal{A}$, $\nu(A) < \infty$, $N(A)$ 服从以 $\nu(A)$ 为参数的 Poisson 分布;
- (ii) 对任意的 $A_1, A_2, \dots, A_n \in \mathcal{A}$, 若它们两两不交, 则 $N(A_1), N(A_2), \dots, N(A_n)$ 相互独立.

Poisson 过程是在 $S = [0, \infty)$ 时 Poisson 点过程的特例.

- 缺 在随机分析课程中讲述.

1.2 简单随机游动

设 $\{X_n, n \geq 1\}$ 是取值于 $\{1, -1\}$ 的独立同分布随机变量列, 且 $P(X_1 = 1) = p$, $P(X_1 = -1) = q$, 其中 $0 \leq p, q \leq 1, p + q = 1$. 令

$$S_0 = 0, S_n = X_1 + X_2 + \dots + X_n, n \geq 1.$$

我们称 S_n 为简单随机游动. 它常用于基本的赌博模型: 掷一个硬币, 如果正面朝上, 你将赢得 1 元; 如果反面朝上, 你将输掉 1 元. S_n 表示在 n 步之后你的收益.

令

$$\tau_1 = \inf\{n \geq 1, S_n = 1\}$$

为首次达到 1 的时间, 令 $\phi_n = P(\tau_1 = n)$. 明显的, $\phi_0 = 0, \phi_1 = p$. 对 $n \geq 2$, 假设随机游动在第 n 步首次到达 1, 则随机游动在第一步只能到 -1 (概率为 q), 从 -1 经 j 步首次回到 0 的概率为 ϕ_j , 在剩下的 $n - j - 1$ 步里, 随机游动以概率 ϕ_{n-j-1} 首次到达 1. 这样,

$$\phi_n = \sum_{j=1}^{n-2} q\phi_j\phi_{n-j-1}, n \geq 2, \phi_0 = 0, \phi_1 = p.$$

上式的严格证明留作练习. 为了求解此递归方程, 我们引入 $\Phi(s) = \sum_{n=0}^{\infty} \phi_n s^n$, $0 \leq s \leq 1$. 于是

$$\sum_{n=2}^{\infty} \phi_n s^n = \sum_{n=2}^{\infty} \sum_{j=0}^{n-2} q\phi_j\phi_{n-j-1} s^n = \sum_{j=0}^{\infty} \sum_{n=j+2}^{\infty} (\phi_{n-j-1} s^{n-j-1}) \phi_j s^j q s = q s \Phi^2(s),$$

即 $\Phi(s) - p s = q s \Phi^2(s)$. 从而

$$\Phi(s) = \frac{1 - \sqrt{1 - 4pqs^2}}{2qs}, \quad 0 \leq s \leq 1. \quad (1.1)$$

展开可得

$$\Phi(s) = \left(1 - \sum_{j=0}^{\infty} \binom{1/2}{j} (-1)^j (4pqs^2)^j \right) / (2qs).$$

这样,

$$\phi_{2j-1} = \binom{1/2}{j} (-1)^{j+1} (4pq)^j / (2q), \quad \phi_{2j} = 0, \quad j \geq 1.$$

由 (1.1) 可知

$$P(\tau_1 < \infty) = \Phi(1) = (1 - |p - q|) / (2q) = \begin{cases} 1, & \text{若 } p \geq q, \\ p/q, & \text{若 } p < q. \end{cases}$$

这说明当 $p \geq q$ 时, 随机游动一定可以到达 1. 而当 $p < q$ 时, 随机游动停留在负半轴的概率为

$$P\left(\bigcap_{n \geq 0} \{S_n \leq 0\}\right) = P(\tau_1 = \infty) = 1 - p/q > 0.$$

此时 $E(\tau_1) = \infty$. 若 $p \geq q$, 则

$$E(\tau_1) = \lim_{s \uparrow 1} \Phi'(s) = \begin{cases} \infty, & \text{若 } p = q = 1/2, \\ (p - q)^{-1}, & \text{若 } p > q. \end{cases}$$

沿用上面的方法, 我们可以研究 $\tau_0 = \inf\{n \geq 1, S_n = 0\}$. 令

$$f_0 = 0, \quad f_n = P(\tau_0 = n), \quad n \geq 1,$$

则当 n 为奇数时, $f_n = 0$. 令

$$F(s) = \sum_{n=0}^{\infty} f_{2n} s^{2n}, \quad 0 \leq s \leq 1.$$

可以看出

$$\tau_0 = \begin{cases} 1 + \inf\{n \geq 1, \sum_{i=1}^n X_{i+1} = 1\}, & \text{若 } X_1 = -1, \\ 1 + \inf\{n \geq 1, \sum_{i=1}^n X_{i+1} = -1\}, & \text{若 } X_1 = 1. \end{cases}$$

令

$$\tau^+ = \inf\{n \geq 1, \sum_{i=1}^n X_{i+1} = 1\}, \quad \tau^- = \inf\{n \geq 1, \sum_{i=1}^n X_{i+1} = -1\}.$$

则 τ^+ 与 τ_1 同分布, 且与 X_1 独立; 由对称性, τ^- 的分布可由 τ_1 的分布中 p 和 q 互换位置得到, 且也与 X_1 独立. 这样,

$$F(s) = E(s^{\tau_0}) = E(s^{1+\tau^+} 1_{\{X_1=-1\}}) + E(s^{1+\tau^-} 1_{\{X_1=1\}}) = 1 - \sqrt{1 - 4pqs^2}.$$

于是,

$$P(\tau_0 < \infty) = F(1) = 1 - |p - q| \begin{cases} = 1, & \text{若 } p = q = 1/2, \\ < 1, & \text{若 } p \neq q. \end{cases}$$

当 $p = q = 1/2$ 时, S_n 的分布是关于原点对称的, 我们称它为对称随机游动. 此时,

$$E(\tau_0) = \lim_{s \uparrow 1} F'(s) = \infty.$$

也就是说, 对称简单随机游动总能够在有限时间内回到原点, 但期望时间为 ∞ .

从直观上来看, 对称简单随机游动在以概率 1 回到原点之后, 继续依照与原过程相同的规律运动, 因此能够以概率 1 再次回到原点, 继续下去, 它将以概率 1 回到原点无穷多次. 而在非对称情形下, 随机运动只能回到原点有限次, 这可由强大数定律导出. 例如, 当 $p > q$ 时, $E(X_1) = p - q > 0$, 故

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = E(X_1) > 0 \text{ a.s.},$$

从而

$$\lim_{n \rightarrow \infty} S_n = \infty.$$

2 离散时间马尔可夫链

2.1 定义与转移矩阵

在本章中, 我们总假定时间指标 $T = \mathbb{N}$, 状态空间 E 可数或者有限, 通常我们可以取为 \mathbb{N} 或者 $\{0, 1, \dots, N\}$. 此时, 马氏性具有下面的形式:

对任意的 $n \geq 0$ 及状态 i_0, \dots, i_{n+1} , 只要 $P(X_0 = i_0, \dots, X_n = i_n) > 0$, 就有

$$P(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = P(X_{n+1} = i_{n+1} | X_n = i_n). \quad (2.1)$$

以后在不引起混淆的情况下, 我们总假定作为条件的事件具有正的概率. 如果已知在时刻 n 是马氏链处于状态 i (记为 $X_n = i$), 则在时刻 $n+1$ 处于状态 j 的概率 (称为单步转移概率) 是 $P(X_{n+1} = j | X_n = i)$, 一般它与 i, j 与 n 都有关. 如果所有的单步转移概率与时间变量 (即 n) 无关, 则称此马氏链为齐次的. 在本书中我们研究的马氏链都是齐次的.

记 $p_{ij} = P(X_{n+1} = j | X_n = i)$, 矩阵 $P = (p_{ij})_{i,j \in E}$ 称为转移矩阵. 显然的, 转移矩阵 P 满足

$$(i) \quad p_{ij} \geq 0, i, j \in E,$$

$$(ii) \quad \sum_{j \in E} p_{ij} = 1, i \in E.$$

马氏链 X_n 完全由转移矩阵 P 及初始分布 $\pi_i = P(X_0 = i)$ 决定. 我们只需要计算出概率族 $P(X_0 = i_0, \dots, X_n = i_n)$, 因为利用全概率公式, 可以计算出任意 X_{j_1}, \dots, X_{j_k} , $j_1 < \dots < j_k$ 的分布.

由条件概率的定义及马氏性, 有

$$\begin{aligned} & P(X_0 = i_0, \dots, X_n = i_n) \\ &= P(X_n = i_n | X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \cdot P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \\ &= P(X_n = i_n | X_{n-1} = i_{n-1}) \cdot P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \\ &= p_{i_{n-1}i_n} P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}). \end{aligned}$$

递推下去, 可以得到

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

对于 $m \geq 1$, 我们引入 $p_{ij}^{(m)} = P(X_{n+m} = j | X_n = i)$, 它表示马氏链从状态 i 出发, 经过 m 步到达 j 的概率, 我们称它为 m 步转移概率. 记 m 步转移矩阵为 $P^{(m)} = (p_{ij}^{(m)})$. 下

面定理在马氏链的研究中起着重要的作用.(2.2) 被称为 C--K (Chapman--Kolmogolov) 方程, 它说明 m 步转移矩阵恰好是转移矩阵 P 的 m 次自乘积. 从直观上来看, 从状态 i 经 $m+k$ 步到达 j 这一过程可以分为两个步骤, 首先从 i 出发, 经过 m 步到达某个状态 r , 然后从 r 再经剩余的 k 步之后到达 j . 或者说, 多步转移概率由始发与终止状态及经过的步数来决定, 而与中间经过的轨道无关, 这即是马氏性的含义.

定理 2.1 $p_{jk}^{(m)}$ 具有性质:

$$(i) 0 \leq p_{ij}^{(m)} \leq 1,$$

$$(ii) \sum_j p_{ij}^{(m)} = 1,$$

(iii) 对任意的 m, k 及状态 i, j , 有

$$p_{ij}^{(m+k)} = \sum_{r \in E} p_{ir}^{(m)} p_{rj}^{(k)}. \quad (2.2)$$

证明 我们只需要证明 (2.2).

$$\begin{aligned} p_{ij}^{(m+k)} &= P(X_{m+k+n} = j | X_n = i) = \sum_{r \in E} P(X_{m+k+n} = j, X_{m+n} = r | X_n = i) \\ &= \sum_{r \in E} P(X_{m+k+n} = j | X_{m+n} = r, X_n = i) P(X_{m+n} = r | X_n = i), \end{aligned}$$

应用下面的引理, 即可得到 (2.2). □

引理 2.1 记 $\mathcal{F}_n = \sigma(X_m, m \leq n)$, $\mathcal{F}^n = \sigma(X_m, m \geq n)$. 则马氏性等价于下面看起来更一般的形式:

对任意的 n , 状态 i 及 $E \in \mathcal{F}_{n-1}$, $G \in \mathcal{F}^{n+1}$, 都有

$$P(G | X_n = i, E) = P(G | X_n = i).$$

上面我们给出了如何决定马氏链的分布, 另一个经常需要考虑的问题是当 $n \rightarrow \infty$ 时 $p_{ij}^{(n)}$ 的渐近行为, 它反映了一个系统经过长时间演化, 能否达到某种平稳态. 在研究这个问题之前, 我们先对状态空间做进一步的研究.

2.2 马氏链状态的分类

2.2.1 互通性

称状态 i 可达 j , 若存在 n , 使得 $p_{ij}^{(n)} > 0$, 记为 $i \rightarrow j$, 从概率上来说, 它等价于马氏链从状态 i 出发, 可以在经过有限个状态之后到达 j . 如果状态 i 和 j 相互可达, 我们称它们是互通的, 记为 $i \leftrightarrow j$. 互通关系是一个等价关系:

- (i) $i \leftrightarrow i$ (反身性),
- (ii) 若 $i \leftrightarrow j$, 则 $j \leftrightarrow i$ (对称性),
- (iii) 若 $i \leftrightarrow j, j \leftrightarrow k$, 则 $i \leftrightarrow k$ (传递性).

传递性的证明如下: 由 $i \leftrightarrow j, j \leftrightarrow k$ 可知, 存在 m, n , 使得 $p_{ij}^{(n)} > 0, p_{jk}^{(m)} > 0$, 这样由 C-K 方程,

$$p_{ik}^{(n+m)} = \sum_{r \in E} p_{ir}^{(n)} p_{rk}^{(m)} \geq p_{ij}^{(n)} p_{jk}^{(m)} > 0.$$

于是, $i \rightarrow k$. 同样可以证明 $k \rightarrow i$.

这样, 我们可以把状态空间 E 依照互通关系分解成若干个等价类, 同一类中两个状态互通. 如果只有一个等价类, 即状态空间中任意两个状态互通, 我们称此马氏链是不可约的.

2.2.2 周期性

考虑集合 $\{n \geq 1, p_{ii}^{(n)} > 0\}$, 其最大公约数 (g.c.d.) 记为 $d(i)$ (如果集合为空, 定义 $d(i) = 0$), 称为状态 i 的周期.

定理 2.2 若 $i \leftrightarrow j$, 则 i 和 j 具有相同的周期.

证明 设 i 的周期为 d, j 周期为 t . 由 $i \leftrightarrow j$, 则存在 s, r , 使得 $p_{ij}^{(s)} > 0, p_{ji}^{(r)} > 0$. 若 $p_{jj}^{(n)} > 0$, 则

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

这样, $n + r + s$ 能被 d 整除. 又 $p_{ii}^{(r+s)} \geq p_{ij}^{(s)} p_{ji}^{(r)} > 0$, 故 $r + s$ 能被 d 整除. 从而 n 能被 d 整除, 由 n 的任意性, d 整除 t . 反过来, 由对称性, t 整除 d . 故 $d = t$. \square

这个定理说明周期在每个互通关系的等价类上是常值函数. 下面结果的证明需要简单的初等数论, 我们省去它的证明

定理 2.3 存在依赖于 i 的正整数 N , 使得当 $n \geq N$ 时, $p_{ii}^{(nd(i))} > 0$.

对于某个状态 i , 如果 $d(i) = 1$, 我们称它是非周期的. 如果一个马氏链中所有状态周期都是 1, 我们称这个马氏链是非周期的. 通常我们先研究非周期的马氏链, 然后把结果推广到周期情形.

2.3 常返性

首先我们引入记号 P_i , 它表示马氏链从状态 i 出发的概率, 即 $P_i(\cdot) = P(\cdot | X_0 = i)$.

对每个 $n \geq 1$, 定义

$$f_{ij}^{(n)} = P_i(X_n = j, X_v \neq j, v = 1, \dots, n-1).$$

或者说, $f_{ij}^{(n)}$ 表示从状态 i 出发, 在第 n 步首次到达 j 的概率. 约定 $f_{ij}^{(0)} = 0$.

显然的, $f_{ij}^{(1)} = p_{ij}^{(1)}$, 对一般的 n , 我们可以通过下式

$$p_{ij}^{(n)} = \sum_{k=0}^n f_{ij}^{(k)} p_{jj}^{(n-k)}, \quad n \geq 1 \quad (2.3)$$

递推得到. 事实上, 假设马氏链从 i 出发在第 n 步到达 j , 这个过程可以分解为两个步骤, 首先在第 k ($1 \leq k \leq n$) 首次到达 j , 然后从 j 出发, 用剩余的 $n-k$ 步再次到达 j . 严格的证明如下:

$$\begin{aligned} p_{ij}^{(n)} &= P_i(X_n = j) = \sum_{k=1}^n P_i(X_v \neq j, v \leq k-1, X_k = j, X_n = j) \\ &= \sum_{k=1}^n P_i(X_n = j | X_k = j, X_v \neq j, v \leq k-1) P_i(X_k = j, X_v \neq j, k \leq k-1) \\ &= \sum_{k=1}^n f_{ij}^{(k)} p_{jj}^{(n-k)} = \sum_{k=0}^n f_{ij}^{(k)} p_{jj}^{(n-k)}. \end{aligned}$$

(2.3) 被称为关于初次进入的分解公式, 它在马氏链的研究中具有广泛的应用, 是本章最重要的公式之一. 我们先给出它的一个直接应用.

命题 2.1 状态 i 可达 j 当且仅当 $f_{ij} > 0$.

证明 设 $i \rightarrow j$, 则存在 $n \geq 1$ 使得 $p_{ij}^{(n)} > 0$. 若 $f_{ij} = 0$, 则对任意的 $k \geq 1$, $f_{ij}^{(k)} = 0$, 由 (2.3) 可得 $p_{ij}^{(n)} = 0$, 矛盾, 故 $f_{ij} > 0$.

反过来, 若 $f_{ij} > 0$, 则存在 $n \geq 1$, 使得 $f_{ij}^{(n)} > 0$. 由 (2.3),

$$p_{ij}^{(n)} = \sum_{k=0}^n f_{ij}^{(k)} p_{jj}^{(n-k)} \geq f_{ij}^{(n)} p_{jj}^{(0)} = f_{ij}^{(n)} > 0,$$

故 $i \rightarrow j$. □

我们引入矩母函数

$$P_{ij}(s) = \sum_{n=0}^{\infty} p_{ij}^{(n)} s^n, \quad |s| < 1$$

及

$$F_{ij}(s) = \sum_{n=0}^{\infty} f_{ij}^{(n)} s^n, \quad |s| \leq 1.$$

由 (2.3),

$$\begin{aligned} P_{ij}(s) &= \delta_{ij} + \sum_{n=1}^{\infty} \sum_{k=0}^n f_{ij}^{(k)} p_{jj}^{(n-k)} s^n = \delta_{ij} + \sum_{k=0}^{\infty} f_{ij}^{(k)} s^k \sum_{n=k}^{\infty} p_{jj}^{(n-k)} s^{n-k} \\ &= \delta_{ij} + F_{ij}(s) P_{jj}(s). \end{aligned}$$

当 $i = j$ 时, 可以得到

$$P_{ii}(s) = \frac{1}{1 - F_{ii}(s)}, \quad |s| < 1. \quad (2.4)$$

而当 $i \neq j$ 时,

$$P_{ij}(s) = \frac{F_{ij}(s)}{1 - F_{jj}(s)}, \quad |s| < 1. \quad (2.5)$$

称状态 i 常返, 若 $f_{ii} = \sum_{n=1}^{\infty} f_{ii}^{(n)} = 1$, 反之, 则称为非常返或者暂留的. 在 (2.4) 中令 $s \rightarrow 1-$, 可以得到

定理 2.4 状态 i 常返当且仅当 $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$.

定义随机变量

$$N_j = \sum_{n=1}^{\infty} 1_{\{X_n=j\}} = \#\{n \geq 1, X_n = j\},$$

它表示马氏链 X_n 到达状态 j 的次数. 在概率 P_i 下 N_j 的期望为 $\sum_{n=1}^{\infty} p_{ij}^{(n)}$. 上面定理表明, 状态 i 常返当且仅当从状态 i 出发, 回到 i 的平均次数为 ∞ , 这与“常返”字面上的意思一致.

推论 2.1 设 $i \leftrightarrow j$, i 常返, 则 j 也是常返的.

证明 由 $i \leftrightarrow j$, 存在 $m, n \geq 1$, 使得 $p_{ij}^{(n)} > 0$, $p_{ji}^{(m)} > 0$. 对任意的 $v \geq 1$, 有 $p_{jj}^{(m+n+v)} \geq p_{ji}^{(m)} p_{ii}^{(v)} p_{ij}^{(n)}$. 对 v 求和, 有

$$\sum_{v=0}^{\infty} p_{jj}^{(v)} \geq \sum_{v=0}^{\infty} p_{jj}^{(m+n+v)} \geq \sum_{v=0}^{\infty} p_{ji}^{(m)} p_{ii}^{(v)} p_{ij}^{(n)} = \infty,$$

故 j 也常返. □

这个推论说明, 类似于周期性, 同一个等价类中状态或者同时常返, 或者同时非常返.

命题 2.2 设 j 非常返, 则对任意状态 i , $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

证明 由 j 非常返可知 $f_{jj} < 1$, 故

$$\sum_{n=0}^{\infty} p_{ij}^{(n)} = P_{ij}(1) = \delta_{ij} + \frac{f_{ij}}{1 - f_{jj}} < \infty,$$

从而 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$. □

推论 2.2 有限状态马氏链必有常返状态.

证明 设状态空间 $E = \{0, \dots, N\}$, 若 E 中所有状态均非常返, 则对任意的 $0 \leq i, j \leq N$, 有 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$. 又对任意的 $n \geq 1$, $\sum_{j=0}^N p_{ij}^{(n)} = 1$, 令 $n \rightarrow \infty$ 即可得到矛盾. □

下面我们进一步明确常返性的概率解释. 我们需要引入 $q_{ij} = P_i(N_j = \infty)$, 它表示从 i 出发, 到达 j 无穷多次的概率.

定理 2.5

$$q_{ij} = \begin{cases} f_{ij}, & \text{若 } j \text{ 常返,} \\ 0, & \text{若 } j \text{ 非常返.} \end{cases}$$

证明 令 $q_{ij}(m) = P_i(N_j \geq m)$ 为从 i 出发到达 j 至少 m 次的概率. 则

$$\begin{aligned} q_{ij}(m+1) &= P_i(X_n = j \text{ } m+1) \\ &= \sum_{k=1}^{\infty} P_i(X_v \neq j, v \leq k-1, X_k = j, X_{k+n} = j \text{ } m) \\ &= \sum_{k=1}^{\infty} P_i(X_v \neq j, v \leq k-1, X_k = j) P_i(X_{k+n} = j \text{ } m | X_k = j) = f_{ij} q_{jj}(m). \end{aligned}$$

这样,

$$q_{ij}(m) = f_{ij} q_{jj}(m-1) = f_{ij} f_{jj} q_{jj}(m-2) = f_{ij} f_{jj}^{m-1}.$$

若 j 非常返, 则 $q_{ij} = \lim_{m \rightarrow \infty} q_{ij}(m) = 0$. 若 j 常返, 则 $q_{ij} = f_{ij}$. □

从这个定理我们可以看出, 状态 i 常返当且仅当从 i 出发, 以概率 1 能够返回 i 无穷多次, 即 $P_i(N_i = \infty) = 1$. 而当 i 非常返时, 只能回到 i 有限多次.

在推论 2.1 中我们证明互通的状态具有相同的常返性. 下面结果说明, 常返状态只能到达常返状态.

定理 2.6 设状态 i 常返, 且 $i \rightarrow j$. 则 $j \rightarrow i$, 从而 j 也是常返的. 进一步的, $f_{ij} = f_{ji} = q_{ij} = q_{ji} = 1$.

证明 我们只需要证明 $j \rightarrow i$. 由 $i \rightarrow j$, 应用命题 2.1, 存在 $m \geq 1$, 使得 $f_{ij}^{(m)} > 0$. 又 i 常返, $q_{ii} = P_i(N_i = \infty) = 1$,

$$\begin{aligned} 0 &= P_i(X_{m+n} \neq i, n \geq 1) \geq P_i(X_v \neq j, v \leq m-1, X_m = j, X_{m+n} \neq i, n \geq 1) \\ &= f_{ij}^{(m)} P_j(X_n \neq i, n \geq 1) = f_{ij}^{(m)} (1 - f_{ji}). \end{aligned}$$

从而 $f_{ji} = 1$. 再次应用命题 2.1 可知 $j \rightarrow i$. \square

2.4 极限定理与平稳分布

2.4.1 离散更新方程

下面定理是分析马氏链极限性质的一个关键工具. 其证明留在第 2.4.2 节.

定理 2.7 设序列 $\{a_k\}, \{b_k\}, k = 0, 1, 2, \dots$ 满足 $a_k \geq 0, \sum a_k = 1, \sum ka_k < \infty, b_k \geq 0, \sum b_k < \infty$, 并且 $\{k, a_k > 0\}$ 的最大公约数为 1. 设有界序列 $\{u_k\}$ 满足更新方程

$$u_n - \sum_{k=0}^n a_{n-k} u_k = b_n, \quad n = 0, 1, 2, \dots,$$

则极限 $\lim_{n \rightarrow \infty} u_n$ 存在, 且

$$\lim_{n \rightarrow \infty} u_n = \frac{\sum_{k=0}^{\infty} b_k}{\sum_{k=0}^{\infty} ka_k}. \quad (2.6)$$

在 $\sum_{k=0}^{\infty} ka_k = \infty$ 情形下, 如果将上式的极限理解为 0, 定理的结论仍旧成立.

注 2.1 我们举一个例子来说明上面的方程为什么被称为更新方程. 假设某事件发生的时间间隔 ξ_n 具有分布

$$P(\xi_n = k) = a_k, \quad k = 0, 1, \dots, \sum_{k=0}^{\infty} a_k = 1,$$

并且各次发生是独立的, 即 $\{\xi_n\}$ 是独立随机变量列. 如果我们从第 n 次事件发生的时刻 $\xi_1 + \xi_2 + \dots + \xi_n$ 重新开始计时, 得到的过程与原来过程具有完全相同的规律. 因此, 我们说在 $\xi_1 + \dots + \xi_n$ 时进行了一次更新. 设 u_n 为时刻 n 之前事件发生的平均次数. 按照第一次事件发生的时刻进行分解, 我们可以得到

$$u_n = \sum_{k=0}^n (1 + u_{n-k}) a_k + 0 \sum_{k=n+1}^{\infty} a_k = \sum_{k=0}^n a_{n-k} u_k + b_n,$$

其中 $b_n = \sum_{k=0}^n a_k$. 这就得到了更新方程.

注意到 $\{n \geq 1, p_{ii}^{(n)} > 0\}$ 与 $\{n \geq 1, f_{ii}^{(n)} > 0\}$ 具有相同的最大公约数. 在定理 2.7 中, 取 $u_n = p_{ii}^{(n)}$, $a_n = f_{ii}^{(n)}$, $b_0 = 1$, $b_n = 0$, $n \geq 1$, 可以得到

定理 2.8 设马氏链 X_n 不可约, 非周期且常返, 则

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = \frac{1}{m_i}.$$

其中 $m_i = \sum_{n=0}^{\infty} n f_{ii}^{(n)}$ 为状态 i 的平均返回时间. 进一步的,

$$\lim_{n \rightarrow \infty} p_{ji}^{(n)} = \lim_{n \rightarrow \infty} p_{ii}^{(n)}.$$

证明 $p_{ii}^{(n)}$ 的极限由定理 2.7 直接推出. 对于 $i \neq j$, 利用关系式

$$p_{ji}^{(n)} = \sum_{v=0}^n f_{ji}^{(v)} p_{ii}^{(n-v)}$$

即可. 更一般的, 设

$$y_n = \sum_{k=0}^n a_{n-k} x_k,$$

其中

$$a_k \geq 0, \sum_{k=0}^{\infty} a_k = 1, \lim_{k \rightarrow \infty} x_k = c,$$

则 $\lim_{n \rightarrow \infty} y_n = c$. 事实上,

$$y_n - c = \sum_{k=0}^n a_{n-k} x_k - c \sum_{k=0}^{\infty} a_k = \sum_{k=0}^n a_{n-k} (x_k - c) - c \sum_{k=n+1}^{\infty} a_k.$$

对任意的 $\varepsilon > 0$, 存在 $K(\varepsilon)$ 使得当 $k \geq K(\varepsilon)$ 时, $|x_k - c| < \varepsilon/3$. 这样,

$$y_n - c = \sum_{k=0}^{K(\varepsilon)} a_{n-k} (x_k - c) + \sum_{k=K(\varepsilon)+1}^n a_{n-k} (x_k - c) - c \sum_{k=n+1}^{\infty} a_k,$$

于是

$$|y_n - c| \leq M \sum_{k=0}^{K(\varepsilon)} a_{n-k} + \frac{\varepsilon}{3} \sum_{k=K(\varepsilon)+1}^n a_{n-k} + |c| \sum_{k=n+1}^{\infty} a_k,$$

其中 $M = \max_{k \geq 0} |x_k - c|$. 选取 $N(\varepsilon)$ 满足

$$|c| \sum_{k=n+1}^{\infty} a_k < \varepsilon/3, \quad \sum_{k=0}^{K(\varepsilon)} a_{n-k} = \sum_{k=n-K(\varepsilon)}^n a_k < \frac{\varepsilon}{3M}, \quad n \geq N(\varepsilon).$$

则

$$|y_n - c| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \quad n \geq N(\varepsilon).$$

这就证明了 $\lim_{n \rightarrow \infty} y_n = c$. □

注 2.2 设 C 是一个常返类. 则对任意的 $n, i \in C, j \notin C$, 有 $p_{ij}^{(n)} = 0$, 即从 C 中状态出发, 不能到达 C 之外. 这样, 把状态空间限制在 C 上, 仍旧可以得到一个马氏链, 它是不可约和常返的. 因此, 上面定理可以应用于非周期的常返类.

注 2.3 当 j 不在 i 的常返类里时, 同上面定理的证明一样, 我们可以得到

$$\lim_{n \rightarrow \infty} p_{ji}^{(n)} = \sum_{k=0}^{\infty} f_{ji}^{(k)} \lim_{n \rightarrow \infty} p_{ii}^{(n)} = \frac{f_{ji}}{m_i}.$$

因此, 当 $m_i = \infty$ 时, 对任意状态 j 都有 $\lim_{n \rightarrow \infty} p_{ji}^{(n)} = 0$.

注 2.4 应用基本的分析可以得到, 对常返非周期状态 i ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n p_{ii}^{(k)} = \lim_{n \rightarrow \infty} p_{ii}^{(n)} = \frac{1}{\sum_{n=0}^{\infty} n f_{ii}^{(n)}} = \frac{1}{m_i}. \quad (2.7)$$

注 2.5 对于周期为 d 的常返状态 i , 可以证明, 若 m 不是 d 的倍数时, $p_{ii}^{(m)} = 0$, 以及

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

容易看出, 此时 (2.7) 仍旧成立.

注 2.6 对于周期为 d 的常返类 C , 我们可以把它分解成 d 个互不相交的集合 G_r , $0 \leq r \leq d-1$, 使得从 G_r 中状态出发, 一步转移后进入 G_{r+1} (这里 $G_d = G_0$). 这样, 在 d 步之后, 又回到集合 G_r . 进一步的, 可以证明, 每个 G_r 都是马氏链 $\{X_{nd}, n \geq 0\}$ 的非周期常返类. 新马氏链的转移矩阵为 $(p_{ij}^{(d)})$. 于是, 当 i 与 j 在同一个 G_r 中时, $\lim_{n \rightarrow \infty} p_{ji}^{(nd)} = d/m_i$, 而在不同的集合 G_r 中时, $\lim_{n \rightarrow \infty} p_{ji}^{(nd)} = 0$.

对常返状态 i , 若 $m_i < \infty$, 称 i 为正常返, 若否, 则称为零常返. 非周期的正常返状态称为遍历的. 应用推论 2.1 中的方法, 可以证明, 若常返类中某个状态是正常返 (零常返或遍历) 的, 则此类中所有状态都是正常返 (零常返或遍历) 的.

推论 2.3 若马氏链有一个零常返状态, 则必定有无穷多个零常返状态.

证明 设状态 i 零常返, C 为 i 所在的零常返类. 假设 C 有限, 则在 $1 = \sum_{j \in C} p_{ij}^{(n)}$ 中令 $n \rightarrow \infty$, 有

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

矛盾. 故 C 是无穷集合. □

推论 2.4 有限状态马氏链的常返状态必为正常返.

定理 2.9 设 $C = \{0, 1, 2, \dots\}$ 是一个遍历类, 则 $\pi_j = \lim_{n \rightarrow \infty} p_{jj}^{(n)}$ 满足

$$\pi_j = \sum_{i \in C} \pi_i p_{ij}, \quad \sum_{i=1} \pi_i = 1.$$

进一步的, $\pi = \{\pi_j\}$ 由方程

$$\pi_i \geq 0, \quad \sum_{i=0}^{\infty} \pi_i = 1, \quad \pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, \quad j \in C \quad (2.8)$$

唯一确定. 我们称满足 (2.8) 的一组 $\{\pi_j\}$ 为马氏链的一个平稳分布.

证明 对任意的 $n, M, 1 = \sum_{j \in C} p_{ij}^{(n)} \geq \sum_{j=0}^M p_{ij}^{(n)}$. 令 $n \rightarrow \infty$, 得 $1 \geq \sum_{j=0}^M \pi_j$. 故 $\sum_{j=0}^{\infty} \pi_j \leq 1$. 又 $p_{ij}^{(n+1)} \geq \sum_{k=0}^M p_{ik}^{(n)} p_{kj}$, 令 $n \rightarrow \infty$, 可知

$$\pi_j \geq \sum_{k=0}^{\infty} \pi_k p_{kj}. \quad (2.9)$$

两边乘以 p_{ji} 并对 j 求和, 应用 (2.9), 得到 $\pi_j \geq \sum_{k=0}^{\infty} \pi_k p_{kj}^{(2)}$. 继续下去, 对任意的 n , $\pi_j \geq \sum_{k=0}^{\infty} \pi_k p_{kj}^{(n)}$. 对 j 求和,

$$\sum_{j=0}^{\infty} \pi_j \geq \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \pi_k p_{kj}^{(n)} = \sum_{k=0}^{\infty} \pi_k \sum_{j=0}^{\infty} p_{kj}^{(n)} = \sum_{k=0}^{\infty} \pi_k.$$

故上面这组不等式均取等号, 即对任意的 j , $\pi_j = \sum_{k=0}^{\infty} \pi_k p_{kj}^{(n)}$. 令 $n \rightarrow \infty$, 由 $\sum \pi_k$ 收敛, 应用控制收敛定理可以得到

$$\pi_j = \sum_{k=0}^{\infty} \pi_k \lim_{n \rightarrow \infty} p_{kj}^{(n)} = \pi_j \sum_{k=0}^{\infty} \pi_k, \quad j = 0, 1, \dots$$

由 $\pi_j > 0$ 知 $\sum_{k=0}^{\infty} \pi_k = 1$.

若 $x = \{x_k\}$ 满足 (2.8), 则

$$x_k = \sum_{j=0}^{\infty} x_j p_{jk} = \sum_{j=0}^{\infty} x_j p_{jk}^{(n)},$$

令 $n \rightarrow \infty$ 可知

$$x_k = \sum_{j=0}^{\infty} x_j \lim_{n \rightarrow \infty} p_{jk}^{(n)} = \pi_k \sum_{j=0}^{\infty} x_j = \pi_k.$$

唯一性得证. \square

注 2.7 若 C 是周期为 d 的正常返类, 应用注 2.6 和定理 2.9 可以证明 $\sum_j \frac{1}{m_j} = 1$. 进一步的, $\{\frac{1}{m_j}\}$ 是马氏链的平稳分布.

定理 2.10 设 $\{\pi_j\}$ 是马氏链的平稳分布. 若对某个状态 j , $\pi_j > 0$, 则 j 是正常返状态.

证明 若 j 非常返或者零常返, 则 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$. 由平稳分布的定义, 对任意的 n , $\pi_j = \sum_{i \in E} \pi_i p_{ij}^{(n)}$. 令 $n \rightarrow \infty$, 由控制收敛定理,

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

矛盾. 故 j 正常返. \square

注 2.8 综合上面结果, 马氏链有平稳分布当且仅当它具有正常返状态. 设状态空间可以分解为 $E = Q \cup H_1 \cup H_2 \cup \dots$, 其中 Q 是所有非常返和零常返状态全体, H_α , $\alpha \geq 1$ 是所有的正常返类. 则 $\{\omega_j\}$ 是马氏链的平稳分布当且仅当

(i) $\omega_j = 0$, $j \in Q$,

(ii) 存在 $\lambda_\alpha \geq 0$, $\sum_\alpha \lambda_\alpha = 1$, 使得当 $j \in H_\alpha$ 时, $\omega_j = \frac{\lambda_\alpha}{m_j}$.

2.4.2 定理 2.7 的证明

我们假设 $a_1 > 0$. 把定理 2.7 中的更新方程改写为

$$u_n - \sum_{k=0}^n a_k u_{n-k} = b_n, \quad n = 0, 1, 2, \dots \quad (2.10)$$

容易看出对所有的 $k \geq 0$, $u_k \geq 0$. 由于我们假设 $\{u_k\}$ 有界, 故上极限 $\lambda = \limsup_{n \rightarrow \infty} u_n < \infty$. 取子列 $\{n_j\}$ 使得 $\lim_{j \rightarrow \infty} u_{n_j} = \lambda$. 我们证明 $\lim_{j \rightarrow \infty} u_{n_j-1} = \lambda$. 用反证法, 如果这个结论不对, 则存在 $\lambda' < \lambda$ 及无穷多个 j 使得 $u_{n_j-1} < \lambda'$. 记 $\varepsilon = [a_1(\lambda - \lambda')]/4$, $M = \sup_{n \geq 0} u_n$. 取 N 充分大使得当 $n \geq N$ 时,

$$\sum_{k=0}^n a_k > 1 - \frac{\varepsilon}{M}. \quad (2.11)$$

选取充分大的 j 使得

$$n_j \geq N, u_{n_j} > \lambda - \varepsilon, u_{n_j-1} < \lambda' < \lambda, 0 \leq b_{n_j} < \varepsilon, u_n < \lambda - \varepsilon, n \geq n_j - N. \quad (2.12)$$

由 λ 和 λ' 的定义, 这些条件都是可以满足的.

由 (2.10), (2.11) 和 (2.12), 可以得到

$$\begin{aligned} u_{n_j} &\leq \sum_{k=0}^{n_j} a_k u_{n_j-k} + \varepsilon < \sum_{k=0}^N a_k u_{n_j-k} + M \sum_{k=N+1}^{n_j} a_k + \varepsilon < \sum_{k=0}^N a_k u_{n_j-k} + 2\varepsilon \\ &< (a_0 + a_2 + a_3 + \cdots + a_{N-1} + a_N)(\lambda + \varepsilon) + a_1 \lambda' + 2\varepsilon \\ &\leq (1 - a_1)(\lambda + \varepsilon) + a_1 \lambda' + 2\varepsilon < \lambda + 3\varepsilon - a_1(\lambda - \lambda') = \lambda - \varepsilon. \end{aligned}$$

但这与 j 的选取矛盾. 这样, 我们就证明了 $\lim_{j \rightarrow \infty} u_{n_j-1} = \lambda$.

重复刚才的步骤, 可以知道对任意的 $d \geq 0$,

$$\lim_{j \rightarrow \infty} u_{n_j-d} = \lambda. \quad (2.13)$$

接下来, 令 $r_n = a_{n+1} + a_{n+2} + \cdots$, 或者等价的, $\sum_{k=0}^{\infty} k a_k = \sum_{n=0}^{\infty} r_n$. 注意我们不需要假设级数的收敛性. 代入 (2.10),

$$r_0 u_n + r_1 u_{n-1} + \cdots + r_n u_0 = r_0 u_{n-1} + r_1 u_{n-2} + \cdots + r_{n-1} u_0 + b_n,$$

记 $A_n = r_0 u_n + \cdots + r_n u_0$, 上式可写为

$$A_0 = r_0 u_0 = (1 - a_0) u_0 = b_0, \quad A_n = A_{n-1} + b_n, \quad n = 1, 2, \dots,$$

这样, $A_n = \sum_{i=0}^n b_i$. 又 $r_n \geq 0, u_n \geq 0$, 对固定的 $N > 0$ 和 $j > 0$, 有

$$r_0 u_{n_j} + r_1 u_{n_j-1} + \cdots + r_N u_{n_j-N} \leq A_{n_j} - \sum_{n=0}^{n_j} b_n.$$

令 $j \rightarrow \infty$, $(r_0 + \cdots + r_N) \lambda \leq \sum_{n=0}^{\infty} b_n$, 等价的, $\lambda \leq \sum_{n=0}^{\infty} b_n / \sum_{n=0}^{\infty} r_n$.

由 N 的任意性,

$$\lambda \leq \frac{\sum_{n=0}^{\infty} b_n}{\sum_{n=0}^{\infty} r_n}. \quad (2.14)$$

由 $u_k \geq 0$, 这就在 $\sum_{n=0}^{\infty} r_n = \infty$ 情形下证明了结果.

若 $\sum_{n=0}^{\infty} r_n < \infty$, 令 $\mu = \liminf_{n \rightarrow \infty} u_n$. 同上确界情形一样, 我们可以证明若 $\lim_{j \rightarrow \infty} u_{n_j} = \mu$, 则对任意 $d \geq 0$, $\lim_{j \rightarrow \infty} u_{n_j-d} = \mu$. 令 $g(N) = \sum_{n=N+1}^{\infty} r_n$, 则 $\lim_{N \rightarrow \infty} g(N) = 0$, 且

$$\sum_{n=0}^{n_j} b_n \leq r_0 u_{n_j} + r_1 u_{n_j-1} + \cdots + r_N u_{n_j-N} + g(N)M.$$

令 $j \rightarrow \infty$, 可得 $\sum_{n=0}^{\infty} b_n \leq (r_0 + \cdots + r_N)\mu + g(N)M$.

这样, 令 $N \rightarrow \infty$, 我们得到

$$\sum_{n=0}^{\infty} b_n \leq \mu \sum_{n=0}^{\infty} r_n \quad \mu \geq \frac{\sum_{n=0}^{\infty} b_n}{\sum_{n=0}^{\infty} r_n}. \quad (2.15)$$

联合 (2.14) 和 (2.15), 得到 $\mu \geq \lambda$. 从而 $\mu = \lambda$, 即 $\lim_{n \rightarrow \infty} u_n$ 极限存在, 并且

$$\lim_{n \rightarrow \infty} u_n = \frac{\sum_{n=0}^{\infty} b_n}{\sum_{n=0}^{\infty} r_n}.$$

对于 $a_1 = 0$ 的情况, 由于 $\{k, a_k > 0\}$ 的最大公约数为 1, 不难得到当 N 充分大时, 对任意的 $n \geq N$, $a_n > 0$. 应用上面的结果, 我们仍旧可以证明定理的结论.

2.5 一些例子

例 2.1 (简单随机游动) 设 ξ_n 是独立同分布随机变量列,

$$P(\xi_n = 1) = p, \quad P(\xi_n = -1) = q, \quad 0 < p, q < 1, \quad p + q = 1.$$

令

$$X_0 = 0, \quad X_n = \xi_1 + \xi_2 + \cdots + \xi_n, \quad n \geq 1.$$

则 X_n 是马氏链, 状态空间为 \mathbb{Z} , 转移概率

$$p_{ij} = P(X_{n+1} = j | X_n = i) = P(\xi_{n+1} = j - i) = \begin{cases} p, & j = i + 1, \\ q, & j = i - 1, \\ 0, & \text{其它情形.} \end{cases}$$

转移概率图如下:

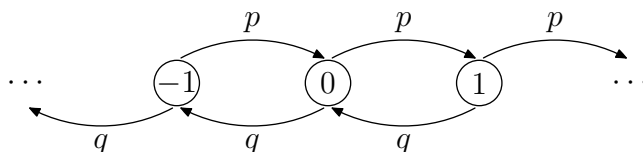


图 2.1 简单随机游动

容易看到, 此马氏链任两个状态是互通的. 在上一章中我们已经证明, $f_{00} = 1 - |p - q|$. 故当 $p \neq q$ 时, $f_{00} < 1$, 0 是非常返的. 而当 $p = q = 1/2$ 时, $f_{00} = 1$, 为常返. 此时, $m_0 = \sum_{n=0}^{\infty} n f_{00}^{(n)} = \infty$, 故为零常返.

我们也可以通过转移概率 $p_{ij}^{(n)}$ 来判断常返性. 事实上,

$$p_{00}^{(2n+1)} = 0, \quad p_{00}^{(2n)} = \binom{2n}{n} p^n q^n.$$

由 Stirling 公式

$$n! \sim \sqrt{2\pi n} e^{-n} n^{n+1/2}, \quad n \rightarrow \infty$$

可得

$$\binom{2n}{n} \sim (\pi n)^{-1/2} 4^n, \quad n \rightarrow \infty,$$

故

$$p_{00}^{(2n)} \sim (\pi n)^{-1/2} (4pq)^n, \quad n \rightarrow \infty.$$

这样, 当 $p \neq q$ 时, $\sum p_{00}^{(n)} < \infty$, 而当 $p = q$ 时, $\sum p_{00}^{(n)} = \infty$.

例 2.2 设 $0 < p < 1$. 考虑带一个吸收壁的随机游动. 状态空间为 $E = \{0, 1, \dots\}$, 转移概率为

$$p_{ij} = \begin{cases} 1, & \text{若 } i = j = 0, \\ p, & \text{若 } j = i + 1, i \neq 0, \\ 1 - p, & \text{若 } j = i - 1, i \neq 0, \\ 0, & \text{其它.} \end{cases}$$

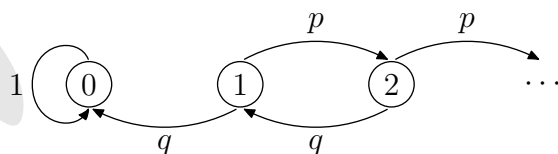


图 2.2 带一个吸收壁的随机游动

显然 0 是遍历状态. 对任意状态 $i \geq 1$, $i \rightarrow 0$, 但 $0 \not\rightarrow i$, 故 i 非常返.

例 2.3 考虑带一个吸收壁和一个反射壁的随机游动, 其状态空间为 $\{0, 1, \dots, n\}$, 转移概率为

$$p_{00} = 1, \quad p_{n,n-1} = 1, \quad p_{i,i+1} = p, \quad p_{i,i-1} = q, \quad 1 \leq i \leq n-1.$$

其中 $0 \leq p \leq 1$, $p + q = 1$. 对 $p < 1$, 转移概率图为:

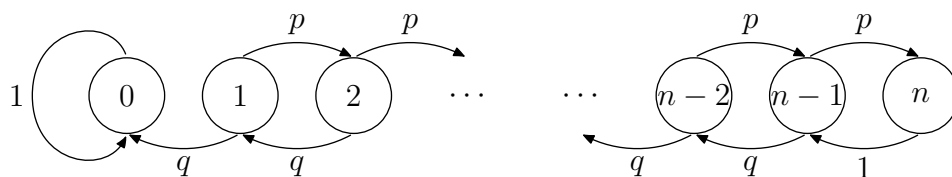


图 2.3 带一个吸收壁和一个反射壁的随机游动

而对于 $p = 1$, 转移概率图为

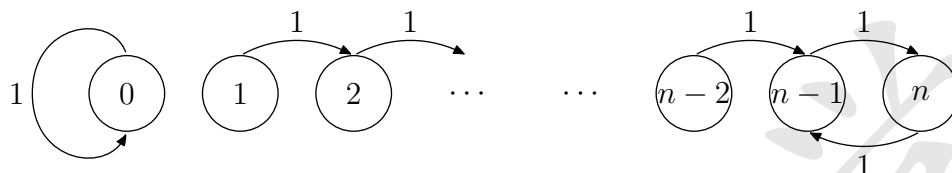


图 2.4 带一个吸收壁和一个反射壁的随机游动

例 2.4 (二维随机游动) 考虑对称的二维随机游动, 其状态空间为平面上格点, 即 $\mathbb{Z} \times \mathbb{Z}$, 从 (i, j) 以 $1/4$ 概率转移到四个点 $(i \pm 1, j), (i, j \pm 1)$ 中之一. 显然此马氏链不可约, 故我们只考虑点 $0 = (0, 0)$. 可以看出

$$p_{00}^{(2n+1)} = 0, \quad p_{00}^{(2n)} = \sum_{k+l=n} \frac{(2n)!}{(k!)^2(l!)^2} \frac{1}{4^{(2n)}}, \quad n \geq 0.$$

分子分母上都乘以 $(n!)^2$, 有

$$p_{00}^{(2n)} = \frac{1}{4^{(2n)}} \binom{2n}{n} \sum_{k=0}^n \binom{n}{k} \binom{n}{n-k} = \left(\frac{1}{4}\right)^{(2n)} \binom{2n}{n}^2.$$

再次应用 Stirling 公式可知 $p_{00}^{(2n)} \sim \frac{1}{\pi n}$. 从而此马氏链是常返的.

例 2.5 考虑马氏链 X_n , 状态空间为 \mathbb{N} , 转移矩阵为

$$\begin{pmatrix} q_0 & p_0 & 0 & 0 & \cdots \\ q_1 & 0 & p_1 & 0 & \cdots \\ q_2 & 0 & 0 & p_2 & \cdots \\ q_3 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

其中 $0 < p_i, q_i < 1, p_i + q_i = 1$. 转移概率图为:

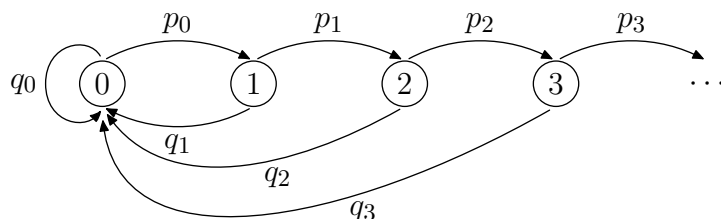


图 2.5 Success runs of binomial trials

这个马氏链可用来描述下面的模型. 为简单起见, 我们假设所有的 p_i 都相同, 记成 p . 独立的重复某一个实验, 每次成功的概率为 p , 失败的概率为 $q = 1 - p$. 记 X_n 为第 n 次实验时连续成功的次数, 即第 n 次实验若失败, 则 $X_n = 0$, 若第 $n - i$ 次为失败, 从 $n - i + 1$ 次到 n 次结果均为成功, 则 $X_n = i$.

显然此马氏链只有一个等价类. 考虑状态 0.

$$f_{00}^{(1)} = q_0 = 1 - p_0,$$

$$f_{00}^{(n)} = \prod_{k=0}^{n-2} p_k(1 - p_{n-1}) = u_{n-2} - u_{n-1},$$

其中 $u_n = p_0 p_1 \cdots p_n$, $u_{-1} = 1$. 这样

$$\sum_{n=1}^{m+1} f_{00}^{(n)} = 1 - u_m.$$

引理 2.2 $\lim_{m \rightarrow \infty} u_m = 0$ 当且仅当 $\sum_{k=0}^{\infty} (1 - p_k) = \infty$.

证明 注意到 $\lim_{m \rightarrow \infty} u_m > 0$ 当且仅当 $\sum_{k=0}^{\infty} -\log p_k < \infty$. 而 $-\log p_k > 1 - p_k$, 从而若 $\lim_{m \rightarrow \infty} u_m > 0$, 则 $\sum_k (1 - p_k) < \infty$.

设 $\sum_k (1 - p_k) = \sum_k q_k < \infty$. 由归纳法可知

$$\prod_{k=n}^m p_k = \prod_{k=n}^m (1 - q_k) > 1 - q_n - \cdots - q_m.$$

取 n 使得 $\sum_{k=n}^{\infty} q_k < 1$, 则 $\prod_{k=n}^{\infty} p_k > 1 - \sum_{k=n}^{\infty} q_k > 0$. □

这样, 马氏链常返当且仅当 $\sum_k (1 - p_k) < \infty$.

下面设 $p_k = p$, $k \geq 0$. 此时马氏链是常返的. 若它存在平稳分布 π_k , 则

$$\sum_{k=0}^{\infty} \pi_k = 1, \quad \pi_j = p\pi_{j-1}, \quad j \geq 1, \quad \pi_0 = \sum_{k=0}^{\infty} (1 - p)\pi_k = 1 - p.$$

从而 $\pi_0 = 1 - p$, $\pi_j = (1 - p)p^j$, $j \geq 1$. 经验证它确实是平稳分布, 从而此马氏链正常返.

例 2.6 (分枝过程) 设某生物种群中个体在死亡时产生 ξ 个后代, ξ 的分布为

$$P(\xi = k) = a_k, \quad a_k \geq 0, \quad \sum_{k=0}^{\infty} a_k = 1.$$

设个体产生的后代数目的是独立的. 记 X_n 为第 n 代时个体数目, 即

$$X_n = \sum_{k=1}^{X_{n-1}} \xi_{n-1,k},$$

其中 $\xi_{n,k}$ 为第 n 代第 k 个个体产生的后代数目. 则 X_n 是马氏链, 转移概率为

$$p_{ij} = P(X_{n+1} = j | X_n = i) = P(\xi_{n,1} + \xi_{n,2} + \cdots + \xi_{n,i} = j).$$

设 ξ 的矩母函数是 $g(s)$, 则 $\xi_{n,1} + \xi_{n,2} + \cdots + \xi_{n,i}$ 的矩母函数为 $g(s)^i$, 而 p_{ij} 为 $g(s)^i$ 展开式中的第 j 项.

假设 $a_1 < 1$. 显然 $p_{00} = 1$, 故 0 是常返状态. 若 $a_0 = 0$, 则 X_n 单增, 且对 $i \geq 1$,

$$f_{ii} = P(X_{n+1} = i | X_n = i) = P(\xi_{n,1} = \xi_{n,2} = \cdots = \xi_{n,i} = 1) = a_1^i < 1.$$

故 i 非常返. 若 $a_0 > 0$, 则

$$p_{i0} = P(X_{n+1} = 0 | X_n = i) = P(\xi_{n,1} = \cdots = \xi_{n,i} = 0) = a_0^i > 0,$$

故 $i \rightarrow 0$, 但从 0 出发不能到达 i , 故 i 非常返.

例 2.7 取状态空间为 $\{0, 1, \dots, 2N\}$, 转移概率为

$$p_{ij} = \binom{2N}{j} p_i^j q_i^{2N-j},$$

其中 $p_i = i/(2N)$, $q_i = 1 - p_i = 1 - i/(2N)$. 此马氏链可用于描述遗传模型. 例如, 假设某种生物具有 $2N$ 个 a 或 A 基因. 如果父代由 j 个 a 基因和 $2N - j$ 个 A 基因构成, 则后代分别以概率

$$p_j = \frac{j}{2N}, \quad q_j = 1 - \frac{j}{2N}$$

选择 a 或是 A 基因. 记 X_n 为第 n 代中 a 基因的数目, 则它是一个马氏链, 并且转移概率如我们上面所述.

容易看到, 状态 0 和 $2N$ 都是吸收态 (即从此状态出发不能到达其它状态), 它们都是常返的. 而对 $1 \leq i \leq 2N - 1$, $p_{i0} > 0$, 故 $i \rightarrow 0$, 从而 i 非常返.

例 2.8 设坛子里装有 $2N$ 个球, 它们是红色或白色的, 每次随机从坛子里拿出一个球, 把它改变颜色后放回去. 设 X_n 为第 n 次后坛中的红球数, 则 X_n 状态空间 $\{1, 2, \dots, 2N\}$, 转移概率为

$$p_{ii} = 0, \quad p_{i,i+1} = \frac{2N-i}{2N}, \quad p_{i,i-1} = \frac{i}{2N}, \quad i = 0, \dots, 2N.$$

显然 X_n 不可约且正常返. 下面我们求它的平稳分布. 令

$$\begin{aligned}\pi_0 &= \frac{\pi_1}{2N}, \\ \pi_k &= \pi_{k-1} \frac{2N-k+1}{2N} + \pi_{k+1} \frac{k+1}{2N}, \quad 1 \leq k \leq 2N-1, \\ \pi_{2N} &= \frac{\pi_{2N-1}}{2N}.\end{aligned}$$

解得

$$\pi_k = \binom{2N}{k} \pi_0.$$

由 $\sum_k \pi_k = 1$ 可知 $\pi_0 = 1/(2^{2N})$. 于是此链的平稳分布为

$$\pi_k = \binom{2N}{k} 2^{-2N}, \quad k = 0, \dots, 2N.$$

例 2.9 (生灭链) 设马氏链 X_n 状态空间 $\{0, 1, \dots\}$, 转移概率为

$$p_{ii} = r_i, \quad p_{i,i+1} = b_i, \quad p_{i,i-1} = a_i,$$

其中 $a_0 = 0$, $a_i + b_i + r_i = 1$. 我们假设 $X_0 = 1$, $a_i > 0$ ($i \geq 1$), $b_i > 0$ ($i \geq 0$). 此时 X_n 不可约.

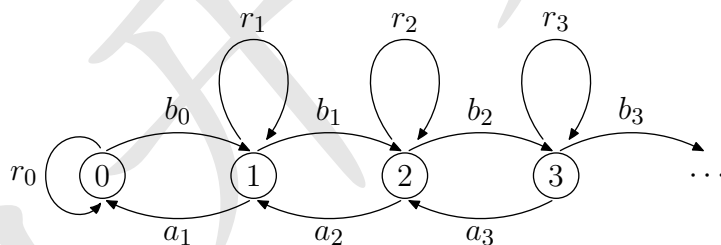


图 2.6 生灭链

(i) X_n 常返当且仅当

$$\sum_{k=1}^{\infty} \frac{a_1 a_2 \cdots a_k}{b_1 b_2 \cdots b_k} = \infty.$$

记

$$\tau_i = \inf\{n, X_n = i\}.$$

对固定的状态 k , 记

$$u(i) = P_i(\tau_0 < \tau_k), \quad 0 < i < k,$$

则

$$u(i) = b_i u(i+1) + a_i u(i-1) + r_i u(i), \quad 0 < i < k.$$

应用 $r_i = 1 - a_i - b_i$,

$$u(i+1) - u(i) = \frac{a_i}{b_i} [u(i) - u(i-1)] = \cdots = \frac{a_1 \cdots a_i}{b_1 \cdots b_i} [u(1) - u(0)].$$

令 $\beta_0 = 1$, $\beta_i = \frac{a_1 \cdots a_i}{b_1 \cdots b_i}$, $u(0) = 1$, 则

$$u(i) - u(i+1) = \beta_i (1 - u(1)), \quad 0 \leq i < k.$$

于是

$$1 = (1 - u(1)) \sum_{i=0}^{k-1} \beta_i,$$

$$u(i) = \sum_{j=i}^{k-1} [u(j) - u(j+1)] = \sum_{j=i}^{k-1} \beta_j / \sum_{j=0}^{k-1} \beta_j.$$

由 $\{\tau_0 < \tau_k\} \uparrow \{\tau_0 < \infty\}$, 当 $\sum_{k=1}^{\infty} \beta_k < \infty$ 时,

$$P_1(\tau_0 < \infty) = \lim_{k \rightarrow \infty} P_1(\tau_0 < \tau_k) = \lim_{k \rightarrow \infty} \left[1 - \left(\sum_{j=0}^{k-1} \beta_j \right)^{-1} \right] = 1.$$

又 $P_1(\tau_0 < \infty) = f_{10}$, 故

$$f_{00} = p_{00} + p_{01} f_{10} = r_0 + b_0 = 1.$$

从而 0 常返.

反过来, 若 0 常返, 由 $f_{10} = 1$ 知 $\sum_{k=1}^{\infty} \beta_k = \infty$.

(ii) 令

$$\gamma_0 = 0, \quad \gamma_k = \frac{b_0 b_1 \cdots b_{k-1}}{a_1 a_2 \cdots a_k}, \quad k \geq 1.$$

则 X_n 正常返当且仅当 $\sum_{k=0}^{\infty} \gamma_k < \infty$.

由 X_n 不可约, 故正常返等价于平稳分布存在. 设

$$\pi_0 = \pi_0 r_0 + \pi_1 a_1,$$

$$\pi_k = \pi_{k-1} b_{k-1} + \pi_k r_k + \pi_{k+1} a_{k+1}, \quad k \geq 1.$$

因 $a_k + r_k + b_k = 1$, 故

$$\begin{aligned} a_1\pi_1 - b_0\pi_0 &= 0, \\ a_{k+1}\pi_{k+1} - b_k\pi_k &= a_k\pi_k - b_{k-1}\pi_{k-1}, \quad k \geq 1. \end{aligned}$$

于是

$$\pi_k = \frac{b_{k-1}\pi_{k-1}}{a_k} = \cdots = \frac{b_0 \cdots b_{k-1}}{a_1 \cdots a_k} \pi_0 = \gamma_k \pi_0.$$

这样,

$$\sum_k \pi_k = \pi_0 \sum_k \gamma_k.$$

这就证明了 π_k 是平稳分布当且仅当 $\sum_k \gamma_k < \infty$.

2.6 吸收概率

记 T 是非常返状态全体. 考虑

$$x_i^{(1)} = \sum_{j \in T} p_{ij}, \quad x_i^{(n)} = \sum_{j \in T} p_{ij} x_j^{(n-1)}, \quad i \in T.$$

容易看到 $x_i^{(n)}$ 是从 i 出发, 到第 n 步仍旧停留在 T 内的概率. 可以证明 $0 \leq x_i^{(n)} \leq x_i^{(n-1)} \leq 1$. 事实上,

$$x_i^{(2)} = \sum_{j \in T} p_{ij} x_j^{(1)} \leq \sum_{j \in T} p_{ij} = x_i^{(1)},$$

若对任意的 $j \in T$, $x_j^{(n-1)} \leq x_j^{(n)}$, 则

$$x_i^{(n+1)} = \sum_{j \in T} p_{ij} x_j^{(n)} \leq \sum_{j \in T} p_{ij} x_i^{(n-1)} \leq x_i^{(n)}.$$

这样, 存在 x_i 使得 $x_i = \lim_{n \rightarrow \infty} x_i^{(n)}$, 并且

$$x_i = \sum_{j \in T} p_{ij} x_j, \quad i \in T. \quad (2.16)$$

如果上面方程有唯一的有界解, 即零向量 $(0, 0, \dots)$, 则从任何非常返状态出发, 最终都会以概率 1 进入常返类. 事实上, x_i 是从状态 i 出发, 永不进入某一个常返类的概率.

设所有的常返类为 C, C_1, C_2, \dots , $\pi_i(C)$ 是从状态 i 出发, 最终进入 C 的概率, $\pi_i^n(C)$ 是从状态 i 出发, 在第 n 步首次进入 C 的概率. 则

$$\begin{aligned}\pi_i(C) &= \sum_{n=0}^{\infty} \pi_i^n(C) \leq 1, \\ \pi_i^1(C) &= \sum_{j \in C} p_{ij}, \\ \pi_i^n(C) &= \sum_{j \in T} p_{ij} \pi_j^{n-1}(C), \quad n \geq 2.\end{aligned}$$

于是

$$\begin{aligned}\pi_i(C) &= \pi_i^1(C) + \sum_{n=2}^{\infty} \pi_i^n(C) = \pi_i^1(C) + \sum_{n=2}^{\infty} \sum_{j \in T} p_{ij} \pi_j^{n-1}(C) \\ &= \pi_i^1(C) + \sum_{j \in T} p_{ij} \pi_j(C).\end{aligned}\tag{2.17}$$

假设方程

$$w_i = \sum_{j \in T} p_{ij} w_j, \quad i \in T$$

有唯一的有界解 (即零解), 则 $\pi_i(C)$ 由方程 2.17 的唯一有界解所确定. 进一步的, 或者对某个 $i \in T$ 有 $\pi_i^1(C) > 0$, 或者对所有的 $i \in T$, $\pi_i(C) = 0$, 从而对所有的 $n \geq 1$, $\pi_i^n(C) = 0$.

定理 2.11 设 C 是非周期正常返类, $j \in C$. 则对 $i \in T$, 有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_i(C) \lim_{n \rightarrow \infty} p_{jj}^{(n)} = \pi_i(C) \pi_j.$$

注 2.9 对于任意的状态 $j \in C$, 有 $\pi_i(C) = f_{ij}$. 直观上来看, 从状态 i 出发, 首次进入 C 时处于某状态 k , 但从状态 k 以概率 1 能够到达状态 j .

例 2.10 设状态空间 $E = \{0, 1, \dots, n\}$, 转移矩阵为

$$\begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

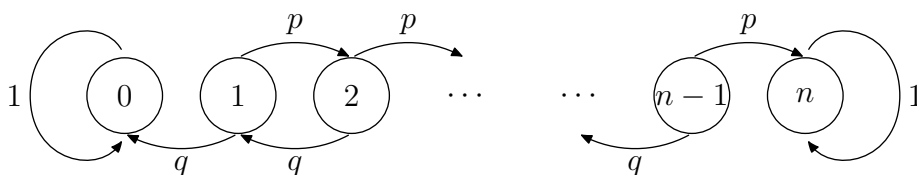


图 2.7 赌徒输光模型

$C_0 = \{0\}$ 和 $C_n = \{n\}$ 是两个常返类. 我们计算 $u_i = \pi_i(C_0)$ 与 $v_i = \pi_i(C_n)$. 方程组 2.17 为

$$\begin{aligned} u_1 &= q + pu_2, \\ u_i &= qu_{i-1} + pu_{i+1}, \quad 2 \leq i \leq n-2, \\ u_{n-1} &= qu_{n-2}. \end{aligned}$$

由于 $u_i - u_{i+1} = (q/p)(u_{i-1} - u_i)$ 可知存在 A, B 使得 $u_r = A + B(q/p)^r$ (当 $p = q$ 时, $(q/p)^r$ 需要换为 r). 当 $q \neq p$ 时, 代入第一个方程, 有

$$A + B\frac{q}{p} = q + p(A + B\frac{q^2}{p^2}),$$

即

$$A = 1 - B.$$

代入最后一个方程,

$$A + B\left(\frac{q}{p}\right)^{n-1} = q\left(A + B\left(\frac{q}{p}\right)^{n-2}\right)$$

即

$$p^n A + q^n B = 0.$$

这样,

$$A = \frac{q^n}{q^n - p^n}, \quad B = \frac{-p^n}{q^n - p^n}.$$

于是

$$u_r = \frac{(q/p)^n - (q/p)^r}{(q/p)^n - 1}.$$

当 $q = p$ 时, $A = 1, B = -1/n$, 故

$$u_r = \frac{n-r}{n}.$$

类似的, 可知 $v_i = 1 - u_i$. 这是明显的, 因为从任意状态 i 出发, 最终都会进入某一个常返类.

当状态空间 $E = \{0, 1, \dots\}$ 时, 类似的有

$$\begin{aligned} u_1 &= q + pu_2, \\ u_i &= qu_{i-1} + pu_{i+1}, i \geq 2 \end{aligned}$$

同样的,

$$u_i = \begin{cases} A + B(q/p)^i, & q \neq p, \\ A + Bi, & p = q = 1/2 \end{cases}$$

若 $q \geq p$, u_i 有界可知 $B = 0$, 从而 $u_i = 1, i \geq 1$. 若 $q < p$, 可以知道 $u_i = (q/p)^i$. 事实上, 在有限状态情形, 令 $n \rightarrow \infty$ 就能得到这结果.

2.7 常返性准则

定理 2.12 设 X_t 是不可约马氏链, $E = \{0, 1, \dots\}$. 则 X_t 非常返的充分必要条件是方程组

$$\sum_{j=0}^{\infty} p_{ij} y_j = y_i, \quad i \neq 0 \quad (2.18)$$

有非常数的有界解.

证明 定义新的转移矩阵

$$\tilde{P} = \tilde{p}_{ij} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{22} & p_{22} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

即将马氏链 X_t 的状态 0 换成一个吸收态. 设 \tilde{X}_t 是以 \tilde{P} 为转移矩阵的马氏链.

设 X_t 是非常返的, 则可以证明存在状态 $j > 0$, 使得 $f_{j0} < 1$. 对马氏链 \tilde{X}_t , $\tilde{\pi}_0(C_0) = 1$, $\tilde{\pi}_j(C_0) = f_{j0} < 1$, 且对任意的 i , $\tilde{\pi}_i(C_0) = \sum_{k=0}^{\infty} \tilde{p}_{ik} \tilde{\pi}_k(C_0)$. 于是对 $i \neq 0$, $\tilde{\pi}_i(C_0) = \sum_{k=0}^{\infty} p_{ik} \tilde{\pi}_k(C_0)$, 从而 $y_k = \tilde{\pi}_k(C_0)$ 即是所需要的有界非常数解.

反过来, 若 $\{y_i\}$ 是方程组 2.18 的有界解. 则

$$\sum_{j=0}^{\infty} \tilde{p}_{ij} y_j = y_i, \quad i \geq 0.$$

迭代下去可得对任意的 $i \geq 0, n \geq 1$,

$$\sum_{j=0}^{\infty} \tilde{p}_{ij}^{(n)} y_j = y_i.$$

若马氏链 X_t 常返, 则 $\lim_{n \rightarrow \infty} \tilde{p}_{i0}^{(n)} = 1$, 且当 $n \rightarrow \infty$ 时,

$$\sum_{j \neq 0} \tilde{p}_{ij}^{(n)} y_j \leq M(1 - \tilde{p}_{i0}^{(n)}) \rightarrow 0,$$

其中 M 是 $\{y_j\}$ 的上界. 于是

$$y_i = \sum_{j \neq 0} \tilde{p}_{ij}^{(n)} y_j + \tilde{p}_{i0}^{(n)} y_0 \rightarrow y_0,$$

从而 $y_i = y_0$ 为常数. □

定理 2.13 设 X_t 不可约, 且存在序列 $\{y_i\}$ 满足

$$\sum_{j=0}^{\infty} p_{ij} y_j \leq y_i \quad i \neq 0, \quad (2.19)$$

且 $\lim_{i \rightarrow \infty} y_i = \infty$, 则 X_t 常返.

证明 应用上面定理的记号. 则对任意的 i ,

$$\sum_{j=0}^{\infty} \tilde{p}_{ij} y_j \leq y_i, \quad \forall i.$$

我们不妨设对任意 $i \geq 0$, $y_i > 0$. 若否, 可以取充分大的 b , 令 $z_i = y_i + b$. 于是对任意的 $m \geq 1$, 有

$$\sum_{j=0}^{\infty} \tilde{p}_{ij}^{(m)} y_j \leq y_i.$$

对任意给定的 $\varepsilon > 0$, 取 $M = M(\varepsilon)$ 满足当 $i \geq M(\varepsilon)$ 时, $1/y_i \leq \varepsilon$. 这样,

$$\sum_{j=0}^{M-1} \tilde{p}_{ij}^{(m)} y_j + \min_{r \geq M} y_r \sum_{j=M}^{\infty} \tilde{p}_{ij}^{(m)} \leq y_i.$$

从而

$$\sum_{j=0}^{M-1} \tilde{p}_{ij}^{(m)} y_j + \min_{r \geq M} y_r (1 - \sum_{j=0}^{M-1} \tilde{p}_{ij}^{(m)}) \leq y_i.$$

注意到当 $j \neq 0$ 时, 由上面定理的证明可知 $\lim_{m \rightarrow \infty} \tilde{p}_{ij}^{(m)} = 0$, 于是

$$\tilde{\pi}_i(C_0)y_0 + \min_{r \geq M} y_r(1 - \tilde{\pi}_i(C_0)) \leq y_i,$$

即

$$1 - \tilde{\pi}_i(C_0) \leq \frac{1}{\min_{r \geq M} y_r} (y_i - \tilde{\pi}_i(C_0)y_0) \leq \varepsilon K,$$

其中

$$K = y_i - \tilde{\pi}_i(C_0)y_0.$$

由 ε 的任意性可知对任意的 i , $\tilde{\pi}_i(C_0) = 1$, 从而马氏链 X_t 是常返的.

2.8 离散排队模型

考虑一个服务队列, 在每个时间段只能对一名顾客进行服务, 而在第 n 个时间段到来的顾客数目 ξ_n 服从分布

$$P(\xi_n = k) = a_k, \quad k = 0, 1, \dots, \quad a_k \geq 0, \quad \sum_{k=0}^{\infty} a_k = 1.$$

我们假设 $\{\xi_n\}$ 是独立的. 记 X_n 为在第 n 个时间段等待的顾客数目. 若 $X_n = i$, 则在第 $n+1$ 个时间段等待的顾客数为

$$X_{n+1} = \begin{cases} i - 1 + \xi_n, & \text{若 } i \geq 1, \\ \xi_n, & \text{若 } i = 0. \end{cases}$$

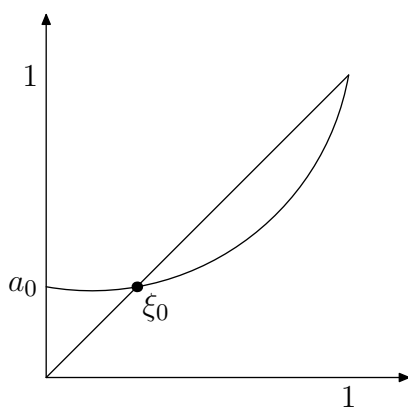
我们也可以写成 $X_{n+1} = (X_n - 1)^+ + \xi_n$. 这样, X_n 是马氏链, 转移矩阵为

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & \cdots \\ a_0 & a_1 & a_2 & a_3 & a_4 & \cdots \\ 0 & a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & 0 & a_0 & a_1 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

下面我们假设 $0 < a_0 < 1$, $a_0 + a_1 < 1$. 容易看到, 此时 X_n 是不可约的. 若 $\sum_{k=0}^{\infty} ka_k > 1$, 我们考虑方程组 2.18 的形如 ξ^i 的非常数有界解. 令 $f(\xi) = \sum_{k=0}^{\infty} a_k \xi^k$. 则方程 2.18 变为

$$\xi^i = \sum_{j=0}^{\infty} p_{ij} \xi^j = \sum_{j=i-1}^{\infty} a_{j-i+1} \xi^j = \xi^{i-1} f(\xi), \quad i \neq 0.$$

于是我们只需要考虑 $f(\xi) = \xi$ 在 $0 < \xi < 1$ 的解. 由于 $f(0) = a_0 > 0$, $f(1) = \sum_{k=0}^{\infty} a_k = 1$, $f'(1) = \sum_{k=0}^{\infty} ka_k > 1$, 从而存在 $0 < \xi_0 < 1$, 使得 $f(\xi_0) = \xi_0$. 这从下图容易看出:

图 2.8 $\xi = f(\xi)$

这样, 我们取 $y_i = \xi_0^i$, $i = 0, 1, \dots$, 则 $\{y_i\}$ 满足方程

$$\sum_{j=0}^{\infty} p_{ij} y_j = y_i, \quad i \neq 0.$$

从而 X_n 非常返.

当 $\sum k a_k \leq 1$ 时, 我们取 $y_j = j$, 则对 $i \neq 0$,

$$\sum_{j=0}^{\infty} p_{ij} j = \sum_{j=i-1}^{\infty} a_{j-i+1} j = \sum_{k=0}^{\infty} k a_k - 1 + i \leq i.$$

应用定理 2.13 可知 X_n 常返.

为了考察 X_n 的正常返性, 我们先考虑下面的问题, 它也具有独立的意义. 设 Y_1, Y_2, \dots 为独立同分布随机变量列, 且

$$P(Y_i = k) = b_k, \quad k = -1, 0, 1, 2, \dots, \quad b_{-1} > 0.$$

令 $S_n = Y_1 + Y_2 + \dots + Y_n$, $Z = \inf\{n, S_n < 0\}$, 记 $\gamma_k = P(Z = k)$, 约定 $\gamma_0 = 0$, 记

$$U(s) = \sum_{k=0}^{\infty} \gamma_k s^k.$$

对非负整数 r , 令 $T_n^{(r)} = r + S_n$, $Z^{(r)} = \inf\{n, T_n^{(r)} < 0\}$. 由于 $X_i \geq -1$, 容易看出 $Z^{(r)} = Z_1 + Z_2 + \dots + Z_{r+1}$, 其中 Z_i 独立且与 Z 分布相同. 这样, $Z^{(r)}$ 的矩母函数为 $(U(s))^{r+1}$, 它的 s^m 项的系数记为 $\gamma_m^{(r+1)}$.

最后, 令

$$G(s) = \frac{b_{-1}}{s} + b_0 + b_1 s + b_2 s^2 + \dots.$$

我们希望使用 $G(s)$ 来表示 $U(s)$. 为此, 我们写出更新关系式

$$\gamma_1 = b_{-1}, \quad \gamma_k = \sum_{j=0}^{\infty} b_j \gamma_{k-1}^{(j+1)}, \quad k \geq 2.$$

第一个等式是明显的, 对于第二个等式, 我们可以把事件 $\{Z = k\}$ 分拆为

$$\{X_1 = j, X_2 + \cdots + X_n + j \geq 0, n = 2, \dots, k-1, X_2 + \cdots + X_k + j = -1\}$$

的并. 这样,

$$\begin{aligned} U(s) &= b_{-1}s + \sum_{n=2}^{\infty} \sum_{j=0}^{\infty} b_j \gamma_{n-1}^{(j+1)} s^n = b_{-1}s + s \sum_{j=0}^{\infty} b_j (U(s))^{j+1} \\ &= b_{-1}s + sU(s) \left[G(U(s)) - \frac{b_{-1}}{U(s)} \right] = sU(s)G(U(s)). \end{aligned}$$

由于 $U(0) = 0$, $U(s)$ 在 $[0, 1]$ 上连续且严格单增, 于是对 $s \in (0, 1]$, $G(U(s)) = 1/s$. 但

$$G''(s) = \frac{2b_{-1}}{s^3} + 2b_2 + 6b_3s + 12b_4s^2 + \cdots > 0, \quad s > 0,$$

即 $G(s)$ 是凸函数, 且 $G(1) = 1$, $\lim_{s \downarrow 0} G(s) = +\infty$. 如下图所示, 对任意的 $s \in [0, 1]$, $G(x) = 1/s$ 至多有两个正根.

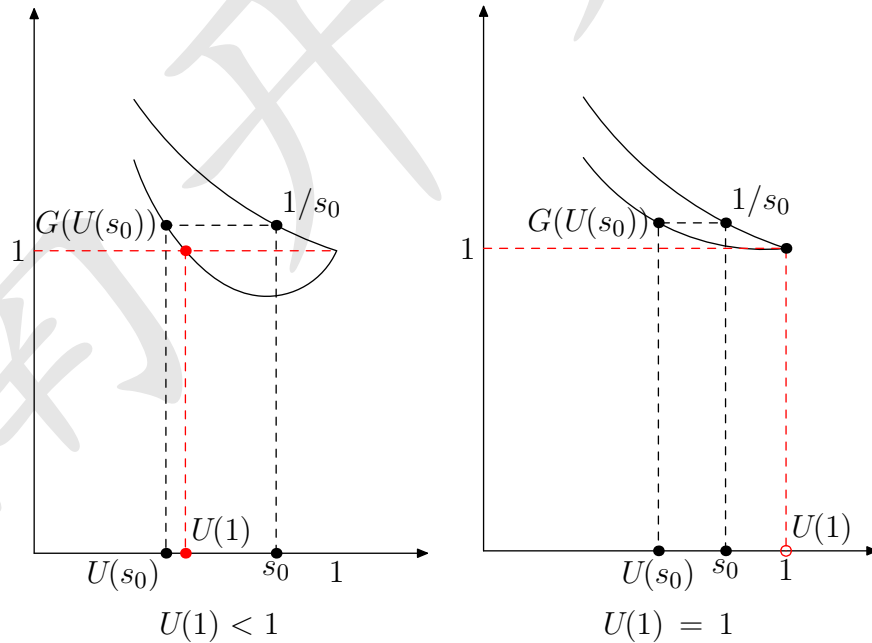


图 2.9 $G(x) = 1/s$

由于 $U(s)$ 的连续性, 当 $G(x) = 1/s$ 有两个解时, $U(s)$ 只能取较小的一个. 下面我们分两种情况来讨论:

- (i) $G'(1) > 0$, 这等价于 $b_{-1} < \sum_{n=0}^{\infty} nb_n$. 容易看出此时 $U(1) = \sum_{k=0}^{\infty} \gamma_k = \xi_0 < 1$. 因此 $\{S_n \geq 0, n \geq 0\}$ 具有正的概率.
- (ii) $G'(1) \leq 0$, 即 $b_{-1} \geq \sum_{n=0}^{\infty} nb_n$. 此时 $U(1) = \sum_{k=0}^{\infty} \gamma_k = 1$, 且对 $0 < s \leq 1$, $G'(U(s))U'(s) = -1/s^2$. 这样, 当 $s \rightarrow 1$ 时, $U(s) \rightarrow 1$. 因此, 若 $G'(1) < 0$, 则

$$E(Z) = \sum_{n=0}^{\infty} n\gamma_n = U'(1) = \frac{-1}{G'(1)} < \infty,$$

而当 $G'(1) = 0$ 时,

$$E(Z) = \sum_{n=0}^{\infty} n\gamma_n = U'(1) = \infty.$$

下面我们回到排队模型. 取 $b_k = a_{k+1}$, 对 $j < i$, 令 Z_{ij} 为从 i 出发, 首次 $< j$ 的时间. 则 $Z_{i,i-1}$ 与 Z 同分布. 由于 $\sum_{i=0}^{\infty} a_i = 1$, 可以得到

$$b_{-1} > \sum_{n=0}^{\infty} nb_n \iff a_0 > \sum_{n=0}^{\infty} na_{n+1} \iff \sum_{n=0}^{\infty} na_n < 1,$$

类似的 $b_{-1} = \sum_{n=0}^{\infty} nb_n$ 等价于 $\sum_{n=0}^{\infty} na_n = 1$. 于是在前一种情况下, $E(Z_{i,i-1}) = \mu < \infty$, 而后种情况下, $E(Z_{i,i-1}) = \mu = \infty$. 又

$$Z_{ij} = Z_{i,i-1} + Z_{i-1,i-2} + \cdots + Z_{j+1,j},$$

从而 $E(Z_{ij}) = (i-j)\mu$, 特别的 $E(Z_{i,0}) = i\mu$.

下面我们考虑状态 0 的平均返回时间. 由马尔可夫性,

$$\begin{aligned} \sum_{n=0}^{\infty} nf_{00}^{(n)} &= a_0 + \sum_{i=1}^{\infty} a_i [E(Z_{i,0}) + 1] = 1 + \sum_{i=1}^{\infty} a_i E(Z_{i,0}) \\ &= 1 + \sum_{i=1}^{\infty} i\mu a_i = 1 + \mu \sum_{i=0}^{\infty} ia_i. \end{aligned}$$

于是, 当 $\mu < \infty$ 时, $\sum_{n=0}^{\infty} nf_{00}^{(n)} < \infty$, X_n 正常返. 而当 $\mu = \infty$ 时, X_n 零常返. 综上,

- 正常返, 若 $\sum_{n=0}^{\infty} na_n < 1$,
- 零常返, 若 $\sum_{n=0}^{\infty} na_n = 1$,
- 非常返, 若 $\sum_{n=0}^{\infty} na_n > 1$.

从直观上来看, $\sum_n na_n$ 为一个时间段到来的顾客的平均数目, 如果 >1 , 则等待队列中人数会越来越多, 如果 <1 , 则等待队列总有为空时.

2.9 马氏链的模拟

在这一节, 给定一个马氏链的初始分布和转移矩阵, 我们使用计算机模拟出马氏链来. 首先我们来看如何模拟一个具有分布函数 $F(x)$ 的随机变量 ξ . 通常, 我们在计算机上容易得到均匀分布的 (伪) 随机数. 假设我们已经得到具有 $[0, 1]$ 上均匀分布的随机变量 U . 令 $\eta = F^{-1}(U)$, 其中

$$F^{-1}(t) = \inf\{s, F(s) > t\}$$

为 F 的右连续逆. 这样

$$P(\eta \leq x) = P(F^{-1}(U) \leq x) = P(U \leq F(x)) = F(x),$$

即 η 的分布函数为 $F(x)$. 例如 ξ 为离散随机变量, 且 $P(\xi = a_i) = p_i$, $a_0 < a_1 < \dots$, 则

$$\eta = \begin{cases} a_0, & U < p_0, \\ a_1, & p_0 \leq U < p_0 + p_1, \\ a_2, & p_0 + p_1 \leq U < p_0 + p_1 + p_2, \\ \dots, & \dots \end{cases}$$

下面设马氏链状态空间为 $E = \{1, 2, \dots, N\}$, 初始分布为 $P(X_0 = i) = p_i$, 转移概率 $P(X_{n+1} = j | X_n = i) = p_{ij}$. 构造函数 $\psi: [0, 1] \rightarrow E$,

$$\psi(x) = \begin{cases} 1, & \text{若 } 0 \leq x < p_1, \\ 2, & \text{若 } p_1 \leq x < p_1 + p_2, \\ \dots, & \dots, \\ N, & \text{若 } p_1 + p_2 + \dots + p_{N-1} \leq x \leq p_1 + \dots + p_N = 1, \end{cases}$$

与 $\phi: E \times [0, 1] \rightarrow E$,

$$\phi(i, x) = \begin{cases} 1, & \text{若 } 0 \leq x < p_{i1}, \\ 2, & \text{若 } p_{i1} \leq x < p_{i1} + p_{i2}, \\ \dots, & \dots, \\ N, & \text{若 } p_{i1} + \dots + p_{i,N-1} \leq x < 1. \end{cases}$$

下面我们生成一系列均匀分布的随机数 U_1, U_2, \dots , 递归的定义

$$X_0 = \psi(U_0), \quad X_{n+1} = \phi(X_n, U_{n+1}), \quad n \geq 0.$$

可以证明, X_n 即是我们所需要的马氏链.

下面是简单随机游动一条轨道的一部分.

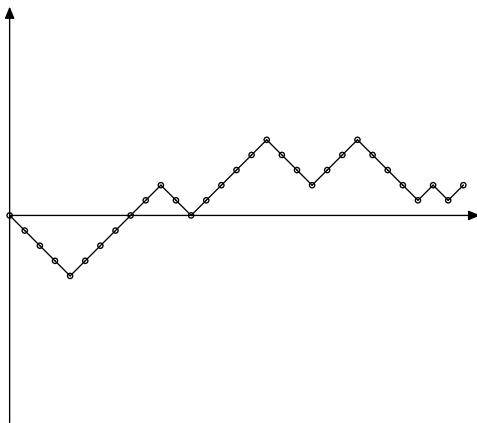


图 2.10 简单随机游动

2.10 蒙特卡罗

给定集合 $S = \{s_1, s_2, \dots, s_k\}$ 上一个分布 π , 我们如何来计算与 π 相关的量呢?例如, 对于 S 上函数 f , 我们需要知道 $\pi(f) = \sum_{i=1}^k f(s_i)\pi_i$. 当 S 非常大时, 这些量往往不容易计算.

例 2.11 (Hard-core 模型) 考虑图 $G = (V, E)$, 其中 $V = \{v_1, v_2, \dots, v_k\}$ 为顶点集, $E = \{e_1, e_2, \dots, e_l\}$ 为边的集合. 我们在图 G 的每个顶点放置 0 或者 1, 使得相邻的顶点 (同一条边上的两个顶点) 不同时为 1. 这样的放置我们称为可行的. 设共有 Z_G 种不同的可行放置方法. 定义 $\{0, 1\}^V$ 上概率分布

$$\mu_G(\xi) = \begin{cases} \frac{1}{Z_G}, & \text{若 } \xi \text{ 可行,} \\ 0, & \text{其它,} \end{cases}$$

它是所有可行放置上的均匀分布. 下图是 8×8 方格上的一个可行放置.

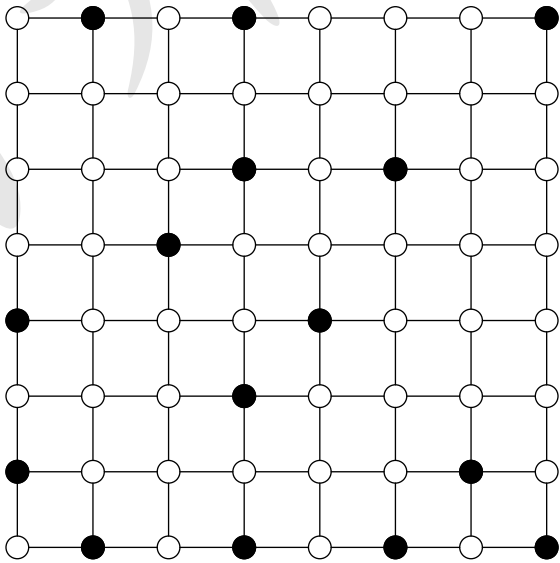


图 2.11 Hard-core 模型. 实心圆表示 1, 空心圆表示 0.

此模型来源于统计物理 (此时 G 为三维格点), 它可以描述气体中半径不可忽略的粒子的分布.

我们希望确定可行放置平均有多少个 1, 即

$$E[n(X)] = \sum_{\xi \in \{0, 1\}^V} n(\xi) \mu_G(\xi) = \frac{1}{\mu_G} \sum_{\xi \in \{0, 1\}^V} n(\xi) 1_{\{\xi\}},$$

其中 X 是分布为 μ_G 的随机变量, $n(\xi)$ 为 ξ 中 1 的数目. 对图进行遍历, 得到所有可行放置的算法并不可取, 因为当图的顶点数增大时, 算法复杂度将会指数增长. 即使在上图的例子中, 所有的放置共有 $2^{64} \approx 1.8447e + 19$ 种, 可行放置数目与此相差的数量级不会多. 而在实际应用中, 图的顶点数还要大的多.

一个可以操作的算法是构造一个不可约, 非周期马氏链 X_n , 使它的状态空间为所有的可行放置, 并且以 μ_G 为平稳分布. 这样, 当 n 非常大时, X_n 的分布渐近于 μ_G , 因此, 我们可以把 X_n 作为 μ_G 的一个样本. 这样的算法我们称为马氏链蒙特卡罗 (Markov chain Monte Carlo) 方法.

构造马氏链 X_n 的一种最简单的算法如下. 我们可以从所有顶点都为 0 的状态开始, 若 X_0, X_1, \dots, X_n 都已知, 我们按照下面的步骤来确定 X_{n+1} :

- (i) 随机的选取一个顶点 v ,
- (ii) 掷一个硬币. 如果正面朝上, 并且 v 的所有相邻顶点均为 0, 则令 $X_{n+1}(v) = 1$, 否则令 $X_{n+1}(v) = 0$,
- (iii) 对除 v 之外的顶点 w 保持不变, 即 $X_{n+1}(w) = X_n(w)$.

容易看到此马氏链是不可约和非周期的. 为证明 μ_G 是它的平稳分布, 我们只需要注意到对所有的可行放置 ξ, ξ' ,

$$\mu_G(\xi) p_{\xi, \xi'} = \mu_G(\xi') p_{\xi', \xi},$$

其中 $p_{\xi, \xi'}$ 是马氏链的转移概率. 更一般的, 我们有

引理 2.3 设马氏链状态空间为 S , 转移概率为 (p_{ij}) , 若 S 上分布 π 满足

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

则 π 是马氏链的平稳分布. 此时我们称马氏链关于分布 π 是可逆的.

证明

$$\sum_{i \in S} \pi_i p_{ij} = \sum_{i \in S} \pi_j p_{ji} = \pi_j,$$

故 π 为平稳分布. □

上面这种算法我们称为 Gibbs 采样. 一般的, 设 π 是 S^V 上的分布. 当 X_0, \dots, X_n 已经确定时, 我们

- (i) 随机选择一个顶点 v ,
- (ii) 依照在分布 π 下, 给定在其它顶点取值与 X_n 相同时的条件概率来确定 $X_{n+1}(v)$ 的取值,
- (iii) 对于其它顶点 w , 取 $X_{n+1}(w) = X_n(w)$.

在我们的例子里, $S = \{0, 1\}$, 取定顶点 v 时, 如果 X_n 在 v 邻接的顶点均为 0, 则在分布 π 下, 顶点 v 以等概率可以取值 0 或者 1, 而如果 X_n 在某个与 v 相邻的顶点取值为 1, 则在概率 π 下, v 只能取 0 值.

对于一般的状态空间 E , 及其上概率分布 π , 我们希望构造不可约, 非周期的马氏链 X_n , 转移概率为 (p_{ij}) , 并且

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

我们可以寻找下面形式的解

$$p_{ij} = q_{ij} \alpha_{ij}, \quad i \neq j, \quad i, j \in E,$$

其中 $Q = (q_{ij})$ 为 E 上一个不可约的转移矩阵, $0 \leq \alpha_{ij} \leq 1$. 也就是说, 如果当前状态为 i , 我们以概率 q_{ij} 来选择新的状态 j , 对 $i \neq j$, 这一选择以概率 α_{ij} 被接受. 下面我们需要选择合适的 α_{ij} . 较为一般的取法的是 Hastings 算法

$$\alpha_{ij} = \frac{s_{ij}}{1 + t_{ij}},$$

其中 $\Sigma = (s_{ij})$ 是对称矩阵,

$$t_{ij} = \frac{\pi_i q_{ij}}{\pi_j q_{ji}}.$$

容易看到这样定义的转移矩阵是关于 π 可逆的. 事实上,

$$\pi_i p_{ij} = \frac{\pi_i q_{ij} s_{ij}}{1 + \pi_i q_{ij} / (\pi_j q_{ji})} = \frac{\pi_i q_{ij} s_{ij} \pi_j q_{ji}}{\pi_i q_{ij} + \pi_j q_{ji}} = \pi_j p_{ji}.$$

当然, 我们需要选择 Σ 使得 $0 \leq \alpha_{ij} \leq 1$, 即

$$s_{ij} \leq 1 + \min(t_{ij}, t_{ji}).$$

例 2.12 (Metropolis 算法) 选取

$$s_{ij} = 1 + \min(t_{ij}, t_{ji}),$$

此时,

$$\alpha_{ij} = \min \left(1, \frac{\pi_j q_{ji}}{\pi_i q_{ij}} \right).$$

其中的一个特例是 q_{ij} 为常数, 即完全随机的选择新的状态 j , 于是

$$\alpha_{ij} = \min \left(1, \frac{\pi_j}{\pi_i} \right).$$

例 2.13 (Barker 算法) 取 $s_{ij} = 1$, 即

$$\alpha_{ij} = \frac{\pi_j q_{ji}}{\pi_j q_{ji} + \pi_i q_{ij}}.$$

当 q_{ij} 为常数时,

$$\alpha_{ij} = \frac{\pi_j}{\pi_i + \pi_j}.$$

3 Poisson 过程

称随机变量 X 服从参数为 λ 的 Poisson 分布, 若 $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, \dots$

称随机变量 X 服从参数为 λ 的指数分布, 若 $P(X > t) = e^{-\lambda t}$. 此时, X 的密度函数为 $\lambda e^{-\lambda t}$, $t > 0$, 分布函数为 $1 - e^{-\lambda t}$, $t > 0$. 指数分布满足无记忆性, 即

$$P(X > t + s) = P(X > t)P(X > s).$$

引理 3.1 设随机变量 X, Y 独立, $f: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ 有界可测. 令 $g(x) = E[f(x, Y)]$. 则 $g(X)$ 可积, 且

$$E[f(X, Y)] = E[g(X)].$$

称 $\{N(t), t \geq 0\}$ 为计数过程, 若 $N(t)$ 表示在时刻 t 之前发生事件的次数. 因此, 计数过程 $N(t)$ 满足:

- (i) $N(t) \geq 0$;
- (ii) $N(t)$ 为整数值;
- (iii) 对 $0 \leq s \leq t$, $N(s) \leq N(t)$;
- (iv) 对 $0 \leq s < t$, $N(t) - N(s)$ 表在区间 $(s, t]$ 发生事件的次数.

3.1 定义

定义 3.1 称 $\{N(t), t \geq 0\}$ 为参数为 λ 的 (齐次) Poisson 过程, 若

- (i) $N(t)$ 是计数过程, $N(0) = 0$;
- (ii) $N(t)$ 具有平稳独立增量, 即对任意的 $0 \leq t_0 < t_1 < \dots < t_n$, $t \geq 0$, $h > 0$, 有 $N(t_1) - N(t_0), \dots, N(t_n) - N(t_{n-1})$ 独立, 且 $N(t+h) - N(t)$ 与 $N(h)$ 同分布;
- (iii) 当 $h \downarrow 0$ 时,

$$P(N(h) = 1) = \lambda h + o(h), \quad P(N(h) \geq 2) = o(h). \quad (3.1)$$

定理 3.1 设 $N(t)$ 是参数为 λ 的 Poisson 过程, 则对任意的 $h > 0$,

$$P(N(t+h) - N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, \dots \quad (3.2)$$

证明 记

$$p_n(t) = P(N(t) = n) = P(N(t+s) - N(s) = n).$$

i) 先考虑 $n = 0$ 的情形. 对 $h > 0$, 有

$$\begin{aligned} p_0(t+h) &= P(N(t+h) = 0) = P(N(t) = 0, N(t+h) - N(t) = 0) \\ &= P(N(t) = 0)P(N(t+h) - N(t) = 0) = p_0(t)p_0(h). \end{aligned}$$

应用

$$\begin{aligned} p_0(h) &= P(N(h) = 0) = 1 - P(N(h) = 1) - P(N(h) \geq 2) \\ &= 1 - \lambda h + o(h), \end{aligned}$$

得

$$p_0(t+h) - p_0(t) = (1 - p_0(h))p_0(t) = \lambda h p_0(t) + o(h).$$

从而 $p_0(t)$ 在 t 右可导, 且右导数为 $-\lambda p_0(t)$. 而

$$\begin{aligned} \frac{p_0(t-h) - p_0(t)}{h} &= \frac{p_0(t-h) - p_0(t-h)p_0(h)}{h} \\ &= \frac{1 - p_0(h)}{h} p_0(t), \end{aligned}$$

令 $h \rightarrow 0$ 可得 $p_0(t)$ 在 t 的左导数也存在, 且为 $-\lambda p_0(t)$. 这样

$$p_0'(t) = -\lambda p_0(t), \quad p_0(0) = 1,$$

于是 $p_0(t) = e^{-\lambda t}$.

ii) 当 $n > 0$ 时

$$\begin{aligned} p_n(t+h) &= P(N(t+h) = n) \\ &= P(N(t) = n, N(t+h) - N(t) = 0) + P(N(t) = n-1, N(t+h) - N(t) = 1) \\ &\quad + P(N(t+h) = n, N(t+h) - N(t) \geq 2) \\ &= p_n(t)p_0(h) + p_{n-1}(t)p_1(h) + o(h) \\ &= (1 - \lambda h)p_n(t) + \lambda h p_{n-1}(t) + o(h). \end{aligned}$$

对 $h > 0$, 有

$$\frac{p_n(t+h) - p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t) + \frac{o(h)}{h},$$

从而 $p_n(t)$ 在 t 的右导数为 $-\lambda p_n(t) + \lambda p_{n-1}(t)$. 类似的可知 $p_n(t)$ 的左导数也存在. 这样

$$p'_n(t) = -\lambda p_n(t) + \lambda p_{n-1}(t), \quad p_n(0) = 0, \quad n \geq 1.$$

上面方程等价于

$$(e^{\lambda t} p_n(t))' = e^{\lambda t} p_{n-1}(t).$$

容易得到

$$p_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

□

这样, Poisson 过程有如下的等价定义.

定义 3.2 称 $\{N(t), t \geq 0\}$ 为参数为 λ 的 Poisson 过程, 若

(i) $N(t)$ 是计数过程, 且 $N(0) = 0$;

(ii) $N(t)$ 是独立增量过程;

(iii) 对任意的 $t \geq 0, h > 0$, 有

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

3.2 另一个等价定义

设 $N(t)$ 是参数为 λ 的 Poisson 过程. 令 $S_0 = 0, S_n = \inf\{t > 0, N(t) \geq n\}, T_n = S_n - S_{n-1}, n = 1, 2, \dots$

定理 3.2 $T_n, n = 1, 2, \dots$ 独立同分布且服从参数 λ 的指数分布.

证明 由

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t},$$

T_1 服从参数为 λ 的指数分布. 对 $0 < t_1 < t_2$ 和充分小的 $h_1, h_2 > 0$,

$$\begin{aligned}
& P(t_1 - h_1 < S_1 \leq t_1 + h_1, t_2 - h_2 < S_2 \leq t_2 + h_2) \\
&= P(N(t_1 - h_1) = 0, N(t_1 + h_1) - N(t_1 - h_1) = 1, \\
&\quad N(t_2 - h_2) - N(t_1 + h_1) = 0, N(t_2 + h_2) - N(t_2 - h_2) = 1) \\
&= e^{-\lambda(t_1 - h_1)} \cdot \lambda 2h_1 e^{-2\lambda h_1} \cdot e^{-\lambda(t_2 - h_2 - t_1 - h_1)} \cdot \lambda 2h_2 e^{-2\lambda h_2} \\
&= 4\lambda^2 h_1 h_2 e^{-\lambda(t_2 + h_2)}.
\end{aligned}$$

所以, (S_1, S_2) 的联合密度函数为

$$g(s_1, s_2) = \begin{cases} \lambda^2 e^{-\lambda s_2}, & \text{若 } 0 < s_1 < s_2; \\ 0, & \text{其它.} \end{cases}$$

由 $T_1 = S_1, T_2 = S_2 - S_1$, (T_1, T_2) 的联合密度函数为

$$f(t_1, t_2) = \begin{cases} \lambda^2 e^{-\lambda(t_1 + t_2)}, & \text{若 } t_i \geq 0; \\ 0, & \text{其它.} \end{cases}$$

这样, T_1, T_2 独立同分布. 一般的情形类似可证. □

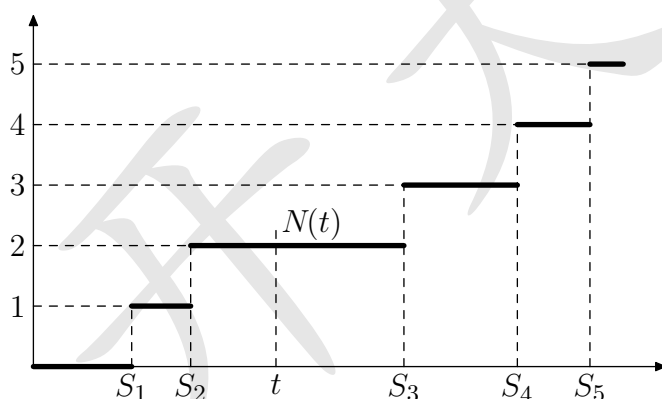


图 3.1 Poisson 过程的构造

定理 3.3 设 T_1, T_2, \dots 独立同分布且同服从参数为 λ 的指数分布. 令 $S_0 = 0, S_n = T_1 + \dots + T_n, n = 1, 2, \dots$ 则 $N(t) = \sup\{n : S_n \leq t\}$ 是参数为 λ 的 Poisson 过程.

证明 当 $h \rightarrow 0$ 时, 有

$$\begin{aligned}
P(N(h) \geq 2) &= P(S_2 \leq h) = \int_0^h \lambda^2 e^{-s_2} ds_2 \int_0^{s_2} ds_1 \\
&= \int_0^h \lambda^2 s e^{-\lambda s} ds \leq \lambda^2 h \int_0^h e^{-\lambda s} ds = o(h).
\end{aligned}$$

及

$$\begin{aligned}
P(N(h) = 1) &= P(S_1 \leq h < S_2) = P(S_1 \leq h) + o(h) \\
&= 1 - e^{-\lambda h} + o(h) = \lambda h + o(h),
\end{aligned}$$

为使得定理成立, 只需要再证明 $N(t)$ 具有平稳独立增量. 我们只证明对任意的 n, k ,

$$P(N(t+s) - N(t) = k, N(t) = n) = P(N(s) = k)P(N(t) = n).$$

一般情形类似可证. 注意到

$$\{S(n) \leq t\} = \{N(t) \geq n\}. \quad (3.3)$$

我们分下面几种情况来讨论.

(i) 设 $k = 0, n = 0$. 由指数分布的无记忆性,

$$\begin{aligned}
P(N(t+s) - N(t) = 0, N(t) = 0) &= P(N(t+s) = 0) \\
&= P(S_1 > t+s) = P(S_1 > s)P(S_1 > t) \\
&= P(N(s) = 0)P(N(t) = 0).
\end{aligned}$$

(ii) 设 $k = 0, n \geq 1$.

$$\begin{aligned}
P(N(t+s) - N(t) = 0, N(t) = n) &= P(S_n \leq t < t+s < S_{n+1}) \\
&= P(S_n \leq t < t+s < S_n + T_{n+1}) = \int_0^t P(t+s < u + T_{n+1}) dP(S_n \leq u) \\
&= \int_0^t P(S_1 > s)P(T_{n+1} > t-u) dP(S_n \leq u) = P(N(s) = 0)P(N(t) = n).
\end{aligned}$$

(iii) 设 $k \geq 1, n = 0$.

$$\begin{aligned}
P(N(t+s) - N(t) = k, N(t) = 0) &= P(t < S_1 \leq S_k \leq t+s < S_{k+1}) \\
&= P(t < S_1 \leq S_1 + \sum_{i=2}^k T_i \leq t+s < S_1 + \sum_{i=2}^{k+1} T_i) \\
&= \int_0^s P(\sum_{i=2}^k T_i \leq s-u < \sum_{i=2}^{k+1} T_i) dP(S_1 \leq t+u).
\end{aligned}$$

由

$$\begin{aligned}
dP(S_1 \leq t+u) &= -dP(S_1 > t+u) = -P(S_1 > t)dP(S_1 > u) \\
&= P(S_1 > t)dP(S_1 \leq u),
\end{aligned}$$

可得

$$\begin{aligned}
 & P(N(t+s) - N(t) = k, N(t) = 0) \\
 &= P(S_1 > t) \int_0^s P\left(\sum_{i=2}^k T_i \leq s - u < \sum_{i=2}^{k+1} T_i\right) dP(S_1 \leq u) \\
 &= P(N(s) = k)P(N(t) = 0).
 \end{aligned}$$

(iv) 设 $k \geq 1, n \geq 1$.

$$\begin{aligned}
 & P(N(t+s) - N(t) = k, N(t) = n) = P(S_n \leq t < S_{n+1} \leq S_{n+k} \leq t+s < S_{n+k+1}) \\
 &= \int_0^t P(t-u < S_1 \leq S_k \leq t+s-u < S_{k+1}) dP(S_n \leq u) \\
 &= P(N(s) = k) \int_0^t P(N(t-u) = 0) dP(S_n \leq u) = P(N(s) = k)P(N(t) = n).
 \end{aligned}$$

□

3.3 Poisson过程的其它性质

3.3.1 顺序统计量

假定 Poisson 过程在时刻 t 之前恰好有一次事件发生, 即 $N(t) = 1$. 可以证明, 事件发生的时刻应服从 $(0, t]$ 上的均匀分布. 事实上,

$$\begin{aligned}
 P(S_1 < s | N(t) = 1) &= \frac{P(S_1 < s, N(t) = 1)}{P(N(t) = 1)} \\
 &= \frac{P(N(s) = 1, N(t) - N(s) = 0)}{P(N(t) = 1)} = \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} \\
 &= \frac{s}{t}.
 \end{aligned}$$

为推广这一结果, 我们引入顺序统计量的概念.

设 Y_1, Y_2, \dots, Y_n 是 n 个随机变量,

$$\{Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}\} = \{Y_1, Y_2, \dots, Y_n\}, \quad Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n)},$$

则称 $Y_{(1)}, \dots, Y_{(n)}$ 为对应于 Y_1, Y_2, \dots, Y_n 的顺序统计量. 若 Y_1, \dots, Y_n 独立同分布, 且具有密度函数 $f(x)$, 则 $Y_{(1)}, \dots, Y_{(n)}$ 的密度函数为

$$f(y_1, \dots, y_n) = n! \prod_{i=1}^n f(y_i), \quad y_1 < \dots < y_n.$$

特别的, 当服从 $(0, t)$ 上均匀分布时, 密度函数为

$$f(y_1, \dots, y_n) = \frac{n!}{t^n}, \quad 0 < y_1 < \dots < y_n < t.$$

定理 3.4 假设在时间 $t > 0$ 前已经发生了 n 次事件, 即已知 $N(t) = n$, 则随机向量 (S_1, S_2, \dots, S_n) 的分布与区间 $[0, t]$ 上 n 个独立均匀分布的顺序统计量具有相同的分布, 即它的联合密度函数为

$$f(t_1, \dots, t_n) = \frac{n!}{t^n}, \quad 0 < t_1 < \dots < t_n < t.$$

证明 我们只对 $n = 2$ 证明. 由于 S_1, \dots, S_n 的联合密度为

$$g(s_1, \dots, s_n) = \lambda^n e^{-\lambda s_n}, \quad s_1 < \dots < s_n,$$

所以, 对 $0 \leq s_1, s_2 \leq t$,

$$\begin{aligned} P(S_1 \leq s_1, S_2 \leq s_2, N(t) = 2) &= P(T_1 \leq s_1, T_1 + T_2 \leq s_2, T_1 + T_2 + T_3 > t) \\ &= \int_0^{s_1} dt_1 \int_0^{s_2-t_1} dt_2 \int_{t-t_1-t_2}^{\infty} \lambda^3 e^{-\lambda(t_1+t_2+t_3)} dt_3 \\ &= \lambda^2 e^{-\lambda t} \int_0^{s_1} dt_1 \int_0^{s_2-t_1} dt_2, \end{aligned}$$

作变量替换 $u_1 = t_1, u_2 = t_1 + t_2$,

$$P(S_1 \leq s_1, S_2 \leq s_2, N(t) = 2) = \lambda^2 e^{-\lambda t} \int_0^{s_1} du_1 \int_{u_1}^{s_2} du_2.$$

因此

$$P(S_1 \leq s_1, S_2 \leq s_2 | N(t) = 2) = \frac{2}{t^2} \int_0^{s_1} du_1 \int_{u_1}^{s_2} du_2.$$

□

例 3.1 设 $N(t)$ 是参数为 λ 的 Poisson 过程, S_n 是第 n 次跳跃的时间, 即 $S_n = \inf\{t, N(t) \geq n\}$. 设对任意的 $n \geq 1$, 在时刻 S_n 发生一次强度为 X_n 的脉冲, 然后它的强度指数递减, 即我们在时刻 t 能够检测到的强度为

$$X_n \exp(-\alpha(t - S_n))_+ = \begin{cases} 0, & \text{若 } t < S_n, \\ X_n \exp(-\alpha(t - S_n)), & \text{若 } t \geq S_n, \end{cases}$$

其中 $\alpha > 0$ 为常数. 假设脉冲的强度是可以叠加的, 则我们在时刻 t 检测到的强度为

$$\eta(t) = \sum_{k=1}^{N(t)} X_k \exp(-\alpha(t - S_k))_+.$$

见图 3.2. 我们需要确定 $\eta(t)$ 的分布函数.

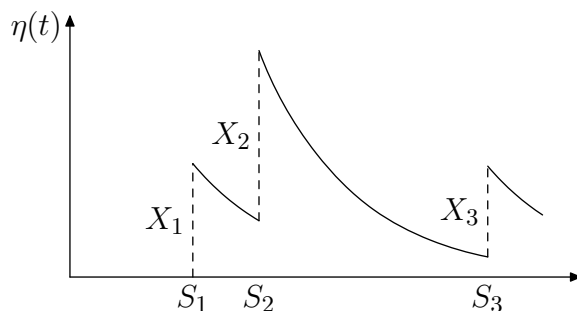


图 3.2

假设 X_k 是独立同分布的, 密度函数为 $h(x)$, 特征函数为

$$\psi(s) = \int_0^\infty e^{isx} h(x) dx.$$

令

$$R(v; t) = P(\eta(t) \leq v) = \sum_{n=0}^{\infty} P(\eta(t) \leq v | N(t) = n) P(N(t) = n).$$

在条件 $N(t) = n$ 下, 前 n 次脉冲到来的时刻 S_k 是 $(0, t)$ 上均匀分布的顺序统计量. 设 $\tau_k, k = 1, 2, \dots, n$ 为独立的 $(0, t)$ 上均匀分布随机变量, 并且以 $S_k, k = 1, \dots, n$ 为顺序统计量.

设 Z_1, \dots, Z_n 为独立随机变量, 与 X_k 同分布, 并且它们是与 $\{\tau_k\}$ 独立的. 定义随机变量

$$Z'_1 = \sum_{j=1}^n Z_j 1_{\{\tau_j = S_1\}}, \quad Z'_2 = \sum_{j=1}^n Z_j 1_{\{\tau_j = S_2\}}, \quad \dots, \quad Z'_n = \sum_{j=1}^n Z_j 1_{\{\tau_j = S_n\}}.$$

即在 $\{\tau_j = S_1\}$ 时, 令 $Z'_1 = Z_j$, 依次类推. 由于 $\{Z_k\}$ 与 $\{\tau_k\}$ 独立, 不能看出 $\{Z'_k\}$ 也是相互独立的随机变量列, 它们都与 Z_k 同分布, 并且与 $\{\tau_k\}$ 也是独立的. 这样, 在条件 $N(t) = n$ 下,

$$\eta(t) = \sum_{k=1}^n Z'_k \exp(-\alpha(t - S_k))_+ = \sum_{k=1}^n Z_k \exp(-\alpha(t - \tau_k))_+.$$

令

$$Y_k(t) = Z_k \exp(-\alpha(t - \tau_k))_+,$$

则对固定的 t , $\{Y_k(t)\}$ 是独立同分布的. 设 $Y_k(t)$ 的密度函数为 $g_t(y; k)$, 则

$$\begin{aligned} \int_0^y g_t(u; k) du &= P(Y_t(k) \leq y) = P(Z_k \exp(-\alpha(t - \tau_k))_+ \leq y) \\ &= \int_0^t P(Z_k \exp(-\alpha(t - \tau_k))_+ \leq y | \tau_k = u) \frac{du}{t} = \int_0^t P(Z_k \leq ye^{\alpha(t-u)}) \frac{du}{t} \\ &= \frac{1}{t} \int_0^t H(ye^{\alpha(t-u)}) du, \end{aligned}$$

其中 $H(y) = \int_0^y h(u) du = P(X_k \leq y)$. 两边对 y 微分,

$$g_t(y; k) = \frac{1}{t} \int_0^t h(ye^{\alpha(t-u)}) e^{\alpha(t-u)} du.$$

这样,

$$\begin{aligned} \theta_t(s) &= \int_0^\infty e^{isy} g_t(y; k) dy = \frac{1}{t} \int_0^t e^{\alpha(t-u)} \left(\int_0^\infty e^{isy} h(ye^{\alpha(t-u)}) dy \right) du \\ &= \frac{1}{t} \int_0^t du \int_0^\infty \exp[is(e^{-\alpha(t-u)}z)] h(z) dz = \frac{1}{t} \int_0^t \psi(se^{-\alpha t} e^{\alpha u}) du = \frac{1}{t} \int_0^t \psi(se^{-\alpha v}) dv. \end{aligned}$$

令 $r(x; t)$ 为 $R(x, t)$ 的密度函数, $\phi_t(w)$ 为特征函数, 则

$$\begin{aligned} \phi_t(w) &= \int_0^\infty e^{iwx} r(x; t) dx \\ &= \sum_{n=0}^\infty \left(\int_0^\infty e^{iwx} \frac{d}{dx} P(\eta(t) \leq x | N(t) = n) dx \right) e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\ &= \sum_{n=0}^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} [\theta_t(w)]^n = \sum_{n=0}^\infty \frac{e^{-\lambda t}}{n!} \left(\lambda \int_0^t \psi(we^{-\alpha v}) dv \right)^n \\ &= \exp \left(- \left(\lambda \int_0^t [1 - \psi(we^{-\alpha v})] dv \right) \right). \end{aligned}$$

特别的,

$$E[\eta(t)] = -i \frac{d}{dw} \phi_t(w) |_{w=0} = \lambda E[X_k] \frac{1 - e^{-\alpha t}}{\alpha}.$$

3.3.2 过程的稀疏

定理 3.5 设 *Poisson* 过程 $N(t)$ 表示到 t 时刻发生的事件个数. 如果每个事件被记录的概率为 p , 且是否被记录是独立的, 则被记录的事件个数 $N_1(t)$ 是强度为 λp 的 *Poisson* 过程.

证明 由全概率公式,

$$\begin{aligned} P(N_1(t) = n) &= \sum_{m=0}^{\infty} P(N_1(t) = n | N(t) = m+n) P(N(t) = m+n) \\ &= \sum_{m=0}^{\infty} \binom{m+n}{n} p^n (1-p)^m e^{-\lambda t} \frac{(\lambda t)^{m+n}}{(m+n)!} \\ &= e^{-\lambda t} \frac{(\lambda p t)^n}{n!} \sum_{m=0}^{\infty} \frac{(\lambda t(1-p))^m}{m!} = e^{-\lambda t} \frac{(\lambda p t)^n}{n!} e^{\lambda t(1-p)} \\ &= e^{-\lambda p t} \frac{(\lambda p t)^n}{n!}. \end{aligned}$$

□

下面假设事件被记录的概率与发生的时间有关. 具体的说, 若某次事件发生在时刻 t , 则以概率 $p(t)$ 被记录, 并且与其它事件是否被记录独立. 记 $N_1(t)$ 为到时刻 t 被记录的事件个数, $N_2(t) = N(t) - N_1(t)$ 为未记录事件的个数.

命题 3.1 $N_1(t)$ 与 $N_2(t)$ 是相互独立的 *Poisson* 随机变量, 且分别以 $\lambda t p$ 和 $\lambda t(1-p)$ 为参数, 其中

$$p = \frac{1}{t} \int_0^t p(s) ds.$$

证明 考虑在 $[0, t]$ 中发生的任一事件, 它发生的时间服从 $[0, t]$ 上的均匀分布. 因此它被记录的概率为

$$p = \frac{1}{t} \int_0^t p(s) ds.$$

这样,

$$P(N_1(t) = n | N(t) = n+k) = \binom{n+k}{n} p^n (1-p)^k.$$

于是

$$\begin{aligned}
P(N_1(t) = n, N_2(t) = k) &= P(N_1(t) = n, N_2(t) = k | N(t) = n+k) P(N(t) = n+k) \\
&= \frac{(n+k)!}{n!k!} p^n (1-p)^k \frac{(\lambda t)^{n+k}}{(n+k)!} e^{-\lambda t} = e^{-\lambda t p} \frac{(\lambda t p)^n}{n!} \cdot e^{-\lambda t (1-p)} \frac{(\lambda t (1-p))^k}{k!}.
\end{aligned}$$

□

例 3.2 (无穷多个服务员的 Poisson 排队系统) 设顾客按照强度为 λ 的 Poisson 过程到达一服务站. 顾客到达后, 不需等待, 即可由某一服务员提供服务, 且服务时间是独立的, 有共同的分布 G .

设在时刻 t 服务完毕的顾客数为 $N_1(t)$, 尚在服务的顾客数为 $N_2(t)$. 若顾客在时刻 s 到达, 如果服务时间小于 $t-s$, 他将会在 t 服务完毕, 这个概率是

$$p(s) = G(t-s), \quad s \leq t.$$

从而由上面命题知 $N_1(t)$ 与 $N_2(t)$ 独立, 且 $N_1(t)$ 和 $N_2(t)$ 分别服从参数为 $\lambda \int_0^t G(y)dy$ 和 $\lambda \int_0^t (1-G(y))dy$ 的 Poisson 分布.

3.4 复合Poisson过程及应用

3.4.1 复合Poisson过程

定义 3.3 设 $Y_n, n \geq 1$ 是一列独立同分布的随机变量, 且与 $N(t)$ 独立. 称过程 $S(t) = \sum_{i=1}^{N(t)} Y_i$ 为复合 Poisson 过程.

定理 3.6 复合 Poisson 过程 $S(t)$ 具有如下性质:

- (i) 它是平稳独立增量过程;
- (ii) 若 $\mu = E[Y_i] < \infty$, 则 $E[S(t)] = \lambda t \mu$; 若 $\text{Var}[Y_i] < \infty$, 则 $\text{Var}(S(t)) = \lambda t E(Y_i^2)$.
- (iii) 对任意 $\xi \in \mathbb{R}$, $E[e^{i\xi S(t)}] = e^{-t\psi(\xi)}$, 其中

$$\psi(\xi) = \lambda \int_{-\infty}^{\infty} (1 - e^{i\xi x}) dF(x).$$

证明 先证平稳独立增量性. 记 F 为 Y_i 的分布函数, F^{*k} 为 F 的 k 重卷积. 对任意的 $0 \leq t_0 < \dots < t_n$ 和 x_i , 有

$$\begin{aligned}
& P(S(t_0) \leq x_0, S(t_1) - S(t_0) \leq x_1, \dots, S(t_n) - S(t_{n-1}) \leq x_n) \\
&= P\left(\sum_{i=1}^{N(t_0)} Y_i \leq x_0, \dots, \sum_{i=N(t_{n-1})+1}^{N(t_n)} Y_i \leq x_n\right) \\
&= \sum_{k_0, \dots, k_n} \prod_{j=0}^n F^{*k_j}(x_j) P(N(t_0) = k_0, N(t_1) - N(t_0) = k_1, \dots, N(t_n) - N(t_{n-1}) = k_n) \\
&= \sum_{k_0, \dots, k_n} \prod_{j=0}^n F^{*k_j}(x_j) P(N(t_0) = k_0) P(N(t_1) - N(t_0) = k_1) \cdots P(N(t_n) - N(t_{n-1}) = k_n) \\
&= P(S(t_0) \leq x_1) P(S(t_1 - t_0) \leq x_2) \cdots P(S(t_n - t_{n-1}) \leq x_n).
\end{aligned}$$

再证 (ii).

$$\begin{aligned}
E[S(t)] &= \sum_{n=1}^{\infty} E\left[\sum_{i=1}^n Y_i | N(t) = n\right] P(N(t) = n) \\
&= \sum_{n=1}^{\infty} n\mu P(N(t) = n) = \mu\lambda t
\end{aligned}$$

相同方法可以得到方差表达式.

(iii) 记 \hat{F} 为分布 F 的特征函数, 即

$$\hat{F}(\xi) = \int_{-\infty}^{\infty} e^{i\xi x} dF(x).$$

$$\begin{aligned}
E[e^{i\xi S(t)}] &= \sum_{n=0}^{\infty} E\left[\exp(i\xi \sum_{j=1}^n Y_j) | N(t) = n\right] P(N(t) = n) \\
&= \sum_{n=0}^{\infty} \hat{F}(\xi)^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \exp\left(-\lambda t \left(\int_{-\infty}^{\infty} (1 - e^{i\xi x}) dF(x)\right)\right).
\end{aligned}$$

□

3.4.2 古典风险模型

我们考虑古典风险模型:

$$R(t) = u + ct - \sum_{i=1}^{N(t)} Y_i,$$

其中 $u > 0, c > 0, Y_n$ 是独立同分布的非负随机变量, $N(t)$ 是强度为 λ 的 Poisson 过程. 记 F 为 Y_i 的分布函数, μ 为 Y_i 的期望, $M(r) = E[e^{rY_i}]$. 假设

- 存在 $0 < \gamma \leq \infty$, 使得当 $r < \gamma$ 时, $M(r) < \infty$, 并且 $\lim_{r \rightarrow \gamma^-} M(r) = \infty$.
- $c > \lambda\mu$.

此时, 关于 r 的方程

$$\lambda M(r) = \lambda + cr$$

有唯一的正实根 R (证明留作练习). 称 R 为 Lundberg 指数 (调节系数).

风险过程 $R(t)$ 的破产时 T 定义为

$$T = \inf\{t > 0, R(t) < 0\}.$$

显然的,

$$\{T < \infty\} = \bigcup_n \{T = S_n\}.$$

下面我们给出 $P(T < \infty)$ 的一个估计. 记

$$A_n = \{T \leq S_n\}, \quad \psi_n(u) = P(A_n),$$

则 $P(T < \infty) = \lim_{n \rightarrow \infty} \psi(u)$,

$$A_n = \{T \leq n\} = \{\exists 1 \leq k \leq n, u + cT_1 + \cdots + \sum_{i=1}^k (cT_i - Y_i) < 0\},$$

即在 S_n 时刻之前破产. 于是

$$\begin{aligned} \psi_{n+1}(u) &= P(A_{n+1}) = P(T = S_1) + P(S_1 < T \leq S_{n+1}) \\ &= P(u + cT_1 - Y_1 < 0) + \\ &\quad P(u + cT_1 - Y_1 \geq 0, \exists 2 \leq k \leq n+1, u + cT_1 - Y_1 + \sum_{i=2}^k (cT_i - Y_i) < 0) \\ &= \int_0^\infty \int_{u+cs}^\infty \lambda e^{-\lambda s} dF(y) ds \end{aligned}$$

$$\begin{aligned}
& + \int_0^\infty \int_0^{u+cs} P(\exists 2 \leq k \leq n+1, u+cs-y + \sum_{i=2}^k (cT_i - Y_i) < 0) \lambda e^{-\lambda s} dF(y) ds \\
& = \int_0^\infty \int_{u+cs}^\infty \lambda e^{-\lambda s} dF(y) ds + \int_0^\infty \int_0^{u+cs} \psi_n(u+cs-y) \lambda e^{-\lambda s} dF(y) ds.
\end{aligned}$$

因为当 $y \geq u+cs$ 时, $e^{-R(u+cs-y)} \geq 1$, 因此

$$\begin{aligned}
\psi_1(u) & = \int_0^\infty \int_{u+cs}^\infty \lambda e^{-\lambda s} dF(y) ds \leq \int_0^\infty \int_{u+cs}^\infty \lambda e^{-\lambda s} e^{-R(u+cs-y)} dF(y) ds \\
& \leq \int_0^\infty \int_0^\infty \lambda e^{-\lambda s} e^{-R(u+cs-y)} dF(y) ds = e^{-Ru} \frac{\lambda M(R)}{\lambda + cR} = e^{-Ru}.
\end{aligned}$$

假设 $\psi_n(u) \leq e^{-Ru}$, 则

$$\psi_{n+1}(u) \leq \int_0^\infty \int_0^\infty \lambda e^{-\lambda s} e^{-R(u+cs-y)} dF(y) ds = e^{-Ru}$$

这就证明了对任意的 u , 都有 $\psi(u) \leq e^{-Ru}$. 从而

命题 3.2

$$P(T < \infty) \leq e^{-Ru}.$$

3.5 Poisson 过程的其它扩展

3.5.1 非齐次 Poisson 过程

定义 3.4 称 $N(t)$ 为具有强度 $\lambda(t)$ 的非齐次 Poisson 过程, 若

(i) $N(t)$ 是计数过程, $N(0) = 0$;

(ii) $N(t)$ 具有独立增量;

(iii) 当 $h \downarrow 0$ 时,

$$P(N(t+h) - N(t) \geq 2) = o(h), \quad P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h).$$

令

$$m(t) = \int_0^t \lambda(s) ds, \quad (3.4)$$

则 $N(t+s) - N(t)$ 服从参数为 $m(t+s) - m(t)$ 的 Poisson 分布. 当强度 $\lambda(t)$ 有界, 即存在 λ 使得 $\lambda(t) \leq \lambda$ 时, 可看作是齐次 Poisson 过程的随机取样. 考虑强度为 λ 的 Poisson 过程, 若某一事件发生在时刻 t , 则以概率 $\lambda(t)/\lambda$ 被记录, 得到的过程即为强度为 $\lambda(t)$ 的非齐次 Poisson 过程 (见命题 3.1).

3.5.2 条件 Poisson 过程

定义 3.5 设正值随机变量 Λ 分布函数为 G , $N(t)$ 为计数过程. 称 $N(t)$ 为条件 Poisson 过程, 若在 $\Lambda = \lambda$ 的条件下, $N(t)$ 是以 λ 为强度的 Poisson 过程.

容易证明, $N(t)$ 具有平稳增量, 但不是独立增量过程, 且

$$P(\Lambda \leq x | N(t) = n) = \frac{\int_0^x e^{-\lambda t} (\lambda t)^n dG(\lambda)}{\int_0^\infty e^{-\lambda t} (\lambda t)^n dG(\lambda)}.$$

3.5.3 Poisson 随机测度

定义 3.6 设 ν 是空间 (S, \mathcal{B}) 上 σ -有限的测度. 称随机测度 N 为以 ν 为强度的 Poisson 随机测度, 若

- (i) 对任意的 $B \in \mathcal{B}$, $\nu(B) < \infty$, 则 $N(B)$ 是参数为 $\nu(B)$ 的 Poisson 随机变量;
- (ii) 对互不相交的集合 $B_1, \dots, B_n \in \mathcal{B}$, $N(B_1), \dots, N(B_n)$ 相互独立.

对 (S, \mathcal{B}) 上非负可测函数 f , 定义

$$\psi_N(f) = E[\exp(-N(f))].$$

可以证明, 若 $\int_S (1 - e^{-f(x)}) \nu(dx) < \infty$, 则

$$\Psi_N(f) = \exp \left\{ - \int_S (1 - e^{-f(x)}) \nu(dx) \right\}.$$

当 $\lambda = \nu(S) < \infty$ 时, 设 ξ_1, ξ_2, \dots 是独立同分布随机变量, 其分布为 $\lambda^{-1}\nu$, 设 X 是参数为 λ 的与 (ξ_k) 独立的 Poisson 随机变量. 令

$$N = \sum_{k=1}^X \delta_{\xi_k},$$

其中 δ_x 是在 $x \in S$ 的单点分布. 则 N 是以 ν 为强度的 Poisson 随机测度. 对于 ν σ -有限情形, 我们只要把 S 分解为可数多个测度有限的自己的并, 然后对相应的 Poisson 随机测度求和即可.

我们经常会考虑乘积空间 $S \times [0, \infty)$ 上的以 $\mu = \nu \otimes dx$ 为强度的 Poisson 随机测度 N . 容易看到, 对于任意的 $B \in \mathcal{B}$, 若 $\nu(B) < \infty$, 则 $N(B \times [0, t])$ 是参数为 $\nu(B)$ 的 Poisson 过程.

这样, $N(S\{t\}) = 0$ 或 $1, \forall t$. 我们按下面的方式来定义 $e(t)$: 若 $N(S\{t\}) = 0$, 令 $e(t) = \partial$, 其中 ∂ 是 E 之外的一个孤立点; 若 $N(S\{t\}) = 1$, 由于 N 可看作是一些单点分布的和, 故限制在 $E \times \{t\}$ 上, N 是一个单点分布, 假设它的负荷集中在 (ϵ, t) 上, 我们令 $e(t) = \epsilon$. 这样,

$$N = \sum_{t \geq 0} \delta_{(e(t), t)}.$$

我们称 $e(t)$ 是一个以 ν 为特征测度的 Poisson 点过程,

$$N(B \times [0, t]) = \#\{s \leq t, e(s) \in B\}, \quad t \geq 0$$

被称为点过程 $e(t)$ 对应的计数过程.

对于复合 Poisson 过程 $S(t) = \sum_{k=1}^{N(t)} Y_k$, 我们令

$$e(t) = S(t) - S(t-) = \begin{cases} Y_k, & \text{若 } t = S_k, \\ 0, & \text{若 } t \notin \{S_1, S_2, \dots\}. \end{cases}$$

则 $e(t)$ 是一个 Poisson 点过程. 而

$$S(t) = \sum_{s \leq t} e(s).$$

4 连续时间马氏链

4.1 定义

4.1.1 马氏性与等价条件

给定概率空间 (Ω, \mathcal{F}, P) , 考虑其上的随机过程 $\{X_t, t \in [0, \infty)\}$, 其状态空间 E 离散. 本章我们总假定时间参数 t 取值于 $[0, \infty)$.

定义 4.1 称 X_t 为连续参数马氏链, 若对任意的 $t_1 < t_2 < \cdots < t_n, i_1, i_2, \dots, i_n \in E$, 只要

$$P(X_{t_1} = i_1, X_{t_2} = i_2, \dots, X_{t_{n-1}} = i_{n-1}) > 0,$$

就有

$$P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) = P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}). \quad (4.1)$$

与前面一样, 我们总假定作为条件的事件具有正的概率.

令 $\mathcal{F}^t = \sigma\{X_s, s \geq t\}$, $\mathcal{F}_t = \sigma\{X_s, s \leq t\}$.

定理 4.1 下列条件等价:

(i) (4.1) 成立.

(ii) 对任意的 $A \in \mathcal{F}^t, B \in \mathcal{F}_t$, 有

$$P(A|B, X_t = i) = P(A|X_t = i).$$

(iii) 对任意的 $A \in \mathcal{F}^t, B \in \mathcal{F}_t$, 有

$$P(AB|X_t = i) = P(A|X_t = i)P(B|X_t = i).$$

(iv) 对任意的 $t_1 < t_2 < \cdots < t_n$, 有

$$P(X_{t_n} = i | X_{t_1}, X_{t_2}, \dots, X_{t_{n-1}}) = P(X_{t_n} = i | X_{t_{n-1}}).$$

(v) 对任意的 $\xi \in \mathcal{F}^t$, 只要 $E(|\xi|) < \infty$, 就有

$$E[\xi | \mathcal{F}_t] = E[\xi | X_t].$$

证明 (1)(ii) 与 (iii) 的等价性由下式

$$\begin{aligned}
P(AB|X_t = i) &= \frac{P(B, X_t = i)}{P(X_t = i)} \frac{P(A, B, X_t = i)}{P(B, X_t = i)} \\
&= P(A|B, X_t = i)P(B|X_t = i)
\end{aligned}$$

即得.

(2) 下面证明 (i) 和 (iv) 的等价性. 由条件期望的定义, 容易知道

$$\begin{aligned}
&P(X_{t_n} = i_n | X_{t_1}, \dots, X_{t_{n-1}}) \\
&= \sum_{i_1, \dots, i_{n-1}} P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) 1_{\{X_{t_1}=i_1, \dots, X_{t_{n-1}}=i_{n-1}\}}.
\end{aligned}$$

事实上, 记

$$Y = \sum_{i_1, \dots, i_{n-1}} P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) 1_{\{X_{t_1}=i_1, \dots, X_{t_{n-1}}=i_{n-1}\}},$$

则对任意的 i_1, \dots, i_{n-1} , 令

$$A = \{X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}\},$$

有

$$\begin{aligned}
&\int_A P(X_{t_n} = i_n | X_{t_1}, \dots, X_{t_{n-1}}) dP = \int_A 1_{\{X_{t_n}=i_n\}} dP \\
&= P(X_{t_1} = i_1, \dots, X_{t_n} = i_n) = \int_A P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) dP = \int_A Y dP.
\end{aligned}$$

进一步的, 上式对任意的 $A \in \sigma(X_{t_1}, \dots, X_{t_{n-1}})$ 成立, 即 $P(X_{t_n} = i_n | X_{t_1}, \dots, X_{t_{n-1}}) = Y$. 这样, 若 (i) 成立, 则

$$\begin{aligned}
&P(X_{t_n} = i_n | X_{t_1}, \dots, X_{t_{n-1}}) \\
&= \sum_{i_1, \dots, i_{n-1}} P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) 1_{\{X_{t_1}=i_1, \dots, X_{t_{n-1}}=i_{n-1}\}} \\
&= \sum_{i_{n-1}} P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}) 1_{\{X_{t_{n-1}}=i_{n-1}\}} = P(X_{t_n} = i_n | X_{t_{n-1}}).
\end{aligned}$$

反过来, 若 (iv) 成立, 则由

$$\int_A P(X_{t_n} = i_n | X_{t_1}, \dots, X_{t_{n-1}}) dP = \int_A P(X_{t_n} = i_n | X_{t_{n-1}}) dP$$

即得 (i).

(3) 接下来证 (iv) \Rightarrow (v). 先证对任意的 $u \geq t$, $B \subset E$, 有

$$P(X_u \in B | \mathcal{F}_t) = P(X_u \in B | X_t),$$

即对任意的 $A \in \mathcal{F}_t$,

$$P(X_u \in B, A) = \int_A P(X_u = i | X_t) dP.$$

由 (iv), 上式对任意的 $A = \{X_{t_1} = i_1, \dots, X_{t_n} = i_n\}$, $t_1 < \dots < t_n \leq t$ 成立, 由 λ - π 方法易知对任意 $A \in \mathcal{F}_t$ 也成立.

进一步的, 应用归纳法可以证明对任意的 $t \leq u_1 < u_2 < \dots < u_m$ 及 $B_1, \dots, B_m \subset E$, 有

$$P(X_{u_1} \in B_1, \dots, X_{u_m} \in B_m | \mathcal{F}_t) = P(X_{u_1} \in B_1, \dots, X_{u_m} \in B_m | X_t).$$

最后应用 \mathcal{L} -系方法可以证明 (v).

(4)(v) \Rightarrow (ii). 设 $A \in \mathcal{F}^t$, $B \in \mathcal{F}_t$. 由 (v),

$$P(A | \mathcal{F}_t) = P(A | X_t).$$

因此

$$\begin{aligned} P(A, B, X_t = i) &= \int_{B \cap \{X_t = i\}} P(A | \mathcal{F}_t) dP \\ &= \int_{B \cap \{X_t = i\}} P(A | X_t) dP = P(A | X_t = i) P(B, X_t = i). \end{aligned}$$

(5)(ii) \Rightarrow (i) 显然. 证毕. □

4.1.2 转移概率

若 $P(X_s = i) > 0$, 记

$$p_{ij}(s, s+t) = P(X_{s+t} = j | X_s = i).$$

若 $p_{ij}(s, s+t)$ 只依赖于 t , 则称相应的马氏链为齐次马氏链. 此时, 记 $p_{ij}(t) = p_{ij}(s, s+t)$, 即

$$p_{ij}(t) = P(X_{s+t} = j | X_s = i).$$

本章我们只考虑齐次马氏链. 记 $P(t) = (p_{ij}(t))$ 为马氏链的转移矩阵.

命题 4.1 转移函数满足

$$(i) 0 \leq p_{ij}(t) \leq 1;$$

$$(ii) \sum_{j \in E} p_{ij}(t) = 1;$$

$$(iii) p_{ij}(s+t) = \sum_{k \in E} p_{ik}(s)p_{kj}(t);$$

$$(iv) p_{ij}(0) = \delta_{ij}.$$

上面命题的证明与离散参数情形下相同. 性质 (iii) 也被称为 C-K 方程.

定理 4.2 设 X_t 是齐次马氏链, 转移概率为 $P(t)$, $q_i = P(X_0 = i)$, 则

$$P(X_{t_1} = i_1, \dots, X_{t_n} = i_n) = \sum_i q_i p_{ii_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1} i_n}(t_n - t_{n-1}).$$

定理 4.3 设 $(P(t))$ 满足上述命题中的四个条件, $q_i \geq 0$, $\sum_i q_i = 1$. 则存在概率空间 (Ω, \mathcal{F}, P) 及其上的齐次马氏链 X_t 满足

$$P(X_0 = i) = q_i, \quad P(X_{s+t} = j | X_s = i) = p_{ij}(t).$$

证明 构造相容的分布函数族

$$F_{t_1, t_2, \dots, t_n}(i_1, i_2, \dots, i_n) = \sum_i q_i p_{ii_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1} i_n}(t_n - t_{n-1}).$$

不妨设 $t_1 < t_2 < \cdots < t_n$. 当 $m < n$ 时,

$$F_{t_1, t_2, \dots, t_m}(i_1, i_2, \dots, i_m) = \sum_{i_{m+1}, \dots, i_n} F_{t_1, t_2, \dots, t_n}(i_1, i_2, \dots, i_n),$$

所以存在概率空间 (Ω, \mathcal{F}, P) 以及 $X_t(\omega)$, 满足

$$P(X_{t_1} = i_1, \dots, X_{t_n} = i_n) = F_{t_1, t_2, \dots, t_n}(i_1, i_2, \dots, i_n).$$

此时 $P(X_0 = i) = \sum_i q_i p_{ii_1}(0) = q_i$, 且

$$\begin{aligned}
& P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) \\
&= \frac{\sum_i q_i p_{ii_1}(t_1) \cdots p_{i_{n-2}i_{n-1}}(t_{n-1} - t_{n-2}) p_{i_{n-1}i_n}(t_n - t_{n-1})}{\sum_i q_i p_{ii_1}(t_1) \cdots p_{i_{n-2}i_{n-1}}(t_{n-1} - t_{n-2})} \\
&= p_{i_{n-1}i_n}(t_n - t_{n-1}) = \frac{\sum_i q_i p_{ii_{n-1}}(t_{n-1}) p_{i_{n-1}i_n}(t_n - t_{n-1})}{\sum_i q_i p_{ii_{n-1}}(t_{n-1})} \\
&= P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}).
\end{aligned}$$

□

4.2 标准转移矩阵

定义 4.2 称 $P(t) = (p_{ij}(t))$ 为标准转移矩阵, 若对任意的 $i, j \in E$,

$$\lim_{t \rightarrow 0+} p_{ij}(t) = \delta_{ij},$$

即 $p_{ij}(t)$ 在 $t = 0$ 是连续的. 这也等价于对任意的 $i \in E$,

$$\lim_{t \rightarrow 0+} p_{ii}(t) = 1.$$

定理 4.4 设 $(p_{ij}(t))$ 标准. 则

(i) 对任意的 $i \in E$ 和 $t \geq 0$, $p_{ii}(t) > 0$.

(ii) 对任意的 $i, j \in E$ 和 $t > 0$, 若 $|h| < t$, 则

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(|h|). \quad (4.2)$$

这样, $p_{ij}(t)$ 在 $t \geq 0$ 上一致连续且关于 $j \in E$ 也是一致的.

(iii) 对任意的 $i, j \in E$, 或者对一切 $t > 0$, $p_{ij}(t) = 0$; 或者对一切 $t > 0$, $p_{ij}(t) > 0$.

证明 (i) 由 $(p_{ij}(t))$ 标准, 对任意的 $i \in E$ 和 $t > 0$, 有正整数 n 使得 $p_{ii}(t/n) > 0$. 这样, 由 C-K 方程, $p_{ii}(t) \geq (p_{ii}(t/n))^n > 0$.

(ii) 对任意的 $h > 0$, 由

$$\begin{aligned}
p_{ij}(t+h) - p_{ij}(t) &= \sum_k p_{ik}(h) p_{kj}(t) - p_{ij}(t) \\
&= \sum_{k \neq i} p_{ik}(h) p_{kj}(t) - p_{ij}(t)(1 - p_{ii}(h)).
\end{aligned}$$

这样,

$$p_{ij}(t+h) - p_{ij}(t) \leq \sum_{k \neq i} p_{ik}(h)p_{kj}(t) \leq \sum_{k \neq i} p_{ik}(h) = 1 - p_{ii}(h),$$

$$p_{ij}(t+h) - p_{ij}(t) \geq -p_{ij}(t)(1 - p_{ii}(h)) \geq -(1 - p_{ii}(h)),$$

从而

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - p_{ii}(h).$$

类似的, 当 $h < 0$ 且 $|h| < t$ 时, 有

$$|p_{ij}(t) - p_{ij}(t+h)| \leq 1 - p_{ii}(|h|).$$

(iii) 证明略. □

定理 4.5 设 $(p_{ij}(t))$ 标准, 则

$$\lim_{t \rightarrow 0+} \frac{1 - p_{ii}(t)}{t} = q_i$$

存在, 但可能等于 $+\infty$.

证明 令 $f(t) = -\log p_{ii}(t)$, 则 $f(t)$ 非负有限. 由 $p_{ii}(s+t) \geq p_{ii}(s)p_{ii}(t)$ 知

$$f(s+t) \leq f(s) + f(t).$$

于是对 $t > h > 0$, 取 n 使得 $t = nh + \epsilon$, $0 \leq \epsilon < h$, 则有

$$\frac{f(t)}{t} \leq \frac{nf(h)}{t} + \frac{f(\epsilon)}{t} = \frac{nh}{t} \frac{f(h)}{h} + \frac{f(\epsilon)}{t}.$$

当 $h \rightarrow 0$ 时, $\frac{nh}{t} \rightarrow 1$, $f(\epsilon) = -\log p_{ii}(\epsilon) \rightarrow 0$, 于是

$$\frac{f(t)}{t} \leq \liminf_{h \rightarrow 0+} \frac{f(h)}{h}.$$

这样,

$$\limsup_{h \rightarrow 0+} \frac{f(h)}{h} \leq \sup_t \frac{f(t)}{t} \leq \liminf_{h \rightarrow 0+} \frac{f(h)}{h}.$$

从而极限

$$\lim_{h \rightarrow 0+} \frac{f(h)}{h} = q_i$$

存在, 其中 $q_i = \sup_{t>0} \frac{f(t)}{t}$. 因此, 当 $h \rightarrow 0+$ 时,

$$\frac{1 - p_{ii}(t)}{t} = \frac{1 - e^{-f(t)}}{t} = (1 + o(1)) \frac{f(t)}{t} \rightarrow q_i.$$

□

定理 4.6 设 $(p_{ij}(t))$ 标准, 则对 $i \neq j$,

$$\lim_{t \rightarrow 0+} \frac{p_{ij}(t)}{t} = q_{ij}$$

存在且有限.

证明 取 $0 < \epsilon < 1/3$, 由标准性知存在 $\delta > 0$ 使当 $t \leq \delta$ 时, $p_{ii}(t) > 1 - \epsilon$, $p_{jj}(t) > 1 - \epsilon$. 先证明对任意的 $t, h > 0$, 只要 $h \leq t \leq \delta$, 就有

$$p_{ij}(h) \leq \frac{p_{ij}(t)}{n} \cdot \frac{1}{1 - 3\epsilon}, \quad (4.3)$$

其中 $n = \lfloor \frac{t}{h} \rfloor$. 为此, 记

$$\begin{aligned} {}_j p_{ik}(h) &= p_{ik}(h), \\ {}_j p_{ik}((l+1)h) &= \sum_{r \neq j} {}_j p_{ir}(lh) p_{rk}(h). \end{aligned}$$

即 ${}_j p_{ik}(mh)$ 表示从 i 出发, 在时刻 mh 处在状态 k , 但在时刻 $h, 2h, \dots, (m-1)h$ 不处于 j 的概率. 这样,

$$p_{ik}(mh) = \sum_{l=1}^{m-1} {}_j p_{ij}(lh) p_{jk}((m-l)h) + {}_j p_{ik}(mh). \quad (4.4)$$

对 $\delta \geq t \geq h$, $n = \lfloor \frac{t}{h} \rfloor$, 有

$$\begin{aligned} \epsilon > 1 - p_{ii}(t) &= \sum_{k \neq i} p_{ik}(t) \geq p_{ij}(t) \\ &\geq \sum_{m=1}^n {}_j p_{ij}(mh) p_{jj}(t - mh) \geq (1 - \epsilon) \sum_{m=1}^n {}_j p_{ij}(mh). \end{aligned}$$

因此

$$\sum_{m=1}^n {}_j p_{ij}(mh) \leq \frac{\epsilon}{1 - \epsilon}. \quad (4.5)$$

其次, 由

$$p_{ii}(mh) = {}_j p_{ii}(mh) + \sum_{l=1}^{m-1} {}_j p_{ij}(lh) p_{ji}((m-1)h)$$

及 (4.5), 得

$${}_j p_{ii}(mh) \geq p_{ii}(mh) - \sum_{l=1}^{m-1} {}_j p_{ij}(lh) \geq 1 - \varepsilon - \frac{\varepsilon}{1 - \varepsilon}. \quad (4.6)$$

从而由 (4.5), (4.6),

$$\begin{aligned} p_{ij}(t) &\geq \sum_{m=1}^n {}_j p_{ii}((m-1)h) p_{ij}(h) p_{jj}(t - mh) \\ &\geq n(1 - \varepsilon - \frac{\varepsilon}{1 - \varepsilon}) p_{ij}(h) (1 - \varepsilon) \geq n(1 - 3\varepsilon) p_{ij}(h). \end{aligned}$$

这就证明了 (4.3).

这样,

$$\frac{p_{ij}(h)}{h} \leq \frac{1}{1 - 3\varepsilon} \frac{p_{ij}(t)}{nh}.$$

当 $h \rightarrow 0+$ 时, $nh \rightarrow t$, 而 $p_{ij}(t)$ 关于 t 连续, 故

$$\limsup_{h \rightarrow 0+} \frac{p_{ij}(h)}{h} \leq \frac{1}{1 - 3\varepsilon} \frac{p_{ij}(t)}{t}.$$

令 $t \rightarrow 0+$, 有

$$\limsup_{h \rightarrow 0+} \frac{p_{ij}(h)}{h} \leq \frac{1}{1 - 3\varepsilon} \liminf_{t \rightarrow 0+} \frac{p_{ij}(t)}{t}.$$

由 ε 的任意性可知极限

$$q_{ij} = \lim_{t \rightarrow 0+} \frac{p_{ij}(t)}{t}$$

存在且有限. □

推论 4.1 设 $(p_{ij}(t))$ 标准, 则

$$\sum_{j \neq i} q_{ij} \leq q_i.$$

证明 因

$$\sum_{j \neq i} \frac{p_{ij}(t)}{t} = \frac{1 - p_{ii}(t)}{t},$$

令 $t \rightarrow 0+$, 由 Fatou 引理,

$$q_i = \liminf_{t \rightarrow 0+} \sum_{j \neq i} \frac{p_{ij}(t)}{t} \geq \sum_{j \neq i} q_{ij}.$$

□

4.3 向前与向后方程组

记 $q_{ii} = -q_i$, $Q = (q_{ij})$. 我们称 Q 为马氏链的密度矩阵.

我们已经证明了转移函数 $p_{ij}(t)$ 在 $t = 0$ 是可导的. 由 C--K 方程

$$p_{ij}(t+h) = \sum_{k \in E} p_{ik}(h)p_{kj}(t) = \sum_{k \in E} p_{ik}(t)p_{kj}(h),$$

两边对 h 求导, 并令 $h = 0$, 如果求导运算与求和号可以交换顺序, 则 $p_{ij}(t)$ 在任意时刻 t 都可导, 并且满足微分方程组

$$p'_{ij}(t) = \sum_{k \in E} q_{ik}p_{kj}(t),$$

及

$$p'_{ij}(t) = \sum_{k \in E} p_{ik}(t)q_{kj}.$$

写成矩阵形式为 $P'(t) = QP(t)$ 及 $P'(t) = P(t)Q$, 分别称为 *Kolmogorov* 向后和向前方程组. 这两个方程组对研究马氏链的性质是重要的. 在实际应用中, 我们往往不能完整的给出转移矩阵 $P(t)$, 但 Q 是较容易确定的, 此时, 我们可以尝试通过上面的两个方程组来解出 $P(t)$. 在非常广泛的条件下, 给定 Q , 总存在转移矩阵 $P(t)$, 使得 $P'(0) = Q$, 即 $p'_{ij}(0) = q_{ij}$, 但这样的 $P(t)$ 可能不是唯一的.

下面我们来考虑向后和向前方程组成立的条件.

定义 4.3 若 $\sum_{j \neq i} q_{ij} = q_i < \infty$, 则称相应的 Q 矩阵是保守的.

由于 $\frac{p_{ij}(t)}{1-p_{ii}(t)} = \frac{q_{ij}t+o(t)}{q_i t+o(t)} \rightarrow \frac{q_{ij}}{q_i}$, 直观上来看, $\frac{q_{ij}}{q_i}$ 表示在 t 时刻离开 i 的条件下, 转移到状态 j 的条件概率, 因此假设 $\frac{\sum_{j \neq i} q_{ij}}{q_i} = 1$ 是一个合理的条件, 也是实际问题中常遇到的情况.

命题 4.2 有限马氏链是保守的.

证明 在

$$\sum_{j \neq i} \frac{p_{ij}(t)}{t} = \frac{1 - p_{ii}(t)}{t}$$

两边直接令 $t \rightarrow 0+$ 即可. □

引理 4.1 若 $f(x)$ 为 (a, b) 上连续函数, 且 $f'_+(x)$ 连续, 则 $f(x)$ 在 (a, b) 上可导.

定理 4.7 设 $(p_{ij}(t))$ 标准, 且一切 $q_i < \infty$. 则 Q 保守当且仅当向后方程组 $P'(t) = QP(t)$ 成立.

证明 设 Q 保守. 由 C-K 方程,

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = -\frac{1 - p_{ii}(h)}{h} p_{ij}(t) + \sum_{k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t).$$

由 Fatou 引理,

$$\liminf_{h \rightarrow 0+} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} \geq \sum_k q_{ik} p_{kj}(t).$$

对任意的 $N > i$, 有

$$\begin{aligned} \limsup_{h \rightarrow 0+} \sum_{k \neq i} \frac{p_{ik}(h)}{h} p_{kj}(t) &\leq \limsup_{h \rightarrow 0+} \sum_{k \neq i, k < N} \frac{p_{ik}(h)}{h} p_{kj}(t) + \limsup_{h \rightarrow 0+} \sum_{k \geq N} \frac{p_{ik}(h)}{h} \\ &\leq \sum_{k \neq i, k < N} q_{ik} p_{kj}(t) + \limsup_{h \rightarrow 0+} \frac{1 - p_{ii}(h)}{h} - \liminf_{h \rightarrow 0+} \sum_{k \neq i, k < N} \frac{p_{ik}(h)}{h} \\ &\leq \sum_{k \neq i, k < N} q_{ik} p_{kj}(t) + q_i - \sum_{k \neq i, k < N} q_{ik}. \end{aligned}$$

令 $N \rightarrow \infty$, 由 Q 的保守性可知

$$\limsup_{h \rightarrow 0+} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} \leq \sum_k q_{ik} p_{kj}(t).$$

这样,

$$\lim_{h \rightarrow 0+} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_k q_{ik} p_{kj}(t).$$

由保守性, 上式右方关于 t 一致收敛, 故为 t 的连续函数. 由上面引理可知向后方程组成立.

下面设向后方程组成立. 由定理 4.5 的证明,

$$q_i = \sup_{t>0} \frac{f(t)}{t},$$

其中 $f(t) = -\log p_{ii}(t)$. 这样,

$$p_{ii}(t) \geq e^{-q_i t} \geq 1 - q_i t. \quad (4.7)$$

对 $t > s \geq 0$, 有

$$p_{ij}(t) \geq p_{ii}(t-s)p_{ij}(s) \geq e^{-q_i(t-s)}p_{ij}(s).$$

故 $e^{q_i t}p_{ij}(t)$ 是关于 t 的增函数. 应用 Fubini 定理 (定理 4.8), 对

$$\sum_j e^{q_i t} p_{ij}(t) = e^{q_i t}$$

逐项求导可得,

$$\sum_j q_i e^{q_i t} p_{ij}(t) + e^{q_i t} \sum_j p'_{ij}(t) = q_i e^{q_i t} \quad \text{a.s.},$$

即

$$\sum_j p'_{ij}(t) = 0, \quad \text{a.s.}$$

在向后方程组

$$p'_{ij}(t) = \sum_k q_{ik} p_{kj}(t)$$

两边对 j 求和, 有

$$0 = \sum_j p'_{ij}(t) = \sum_k q_{ik} \quad \text{a.s.}$$

这样就得到了 Q 的保守性, 并且上式对所有的 t 都成立. □

定理 4.8 若 $f_n(t)$ 单增, 且 $f(t) = \sum_n f_n(t)$ 有限, 则

$$f'(t) = \sum_n f'_n(t) \quad \text{a.s.}$$

定理 4.9 设 Q 保守, $q = \sup_i q_i < \infty$, 则向前方程组 $P'(t) = P(t)Q$ 成立.

证明

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = p_{ij}(t) \frac{p_{jj}(h) - 1}{h} + \sum_{k \neq j} p_{ik}(t) \frac{p_{kj}(h)}{h} \quad (4.8)$$

由于

$$0 \leq \frac{p_{kj}(h)}{h} \leq \sum_{l \neq k} \frac{p_{kl}(h)}{h} = \frac{1 - p_{kk}(h)}{h} \leq q_k \leq q < \infty,$$

应用控制收敛定理, 在 (4.8) 中令 $h \rightarrow 0+$ 即得. □

实际上, 条件 $q = \sup_i q_i < \infty$ 能够推出 Q 的保守性.

推论 4.2 有限状态马氏链向前向后方程组均成立.

这样, 若马氏链状态有限, 则

$$P'(t) = QP(t), \quad P(0) = E,$$

其中 E 是单位矩阵. 容易知道,

$$P(t) = e^{Qt} = \sum_{m=0}^{\infty} \frac{t^m}{m!} Q^m.$$

例 4.1 设马氏链的状态空间 $E = \{0, 1, 2, \dots\}$, 转移矩阵 $(p_{ij}(t))$ 标准, 密度矩阵 Q 满足

$$\begin{cases} q_i = b, & i \geq 0, \\ q_{i,i+1} = b, & i \geq 0, \\ q_{ij} = 0, & j \neq i, j \neq i+1, i \geq 0, j \geq 0, \end{cases}$$

其中 $b > 0$ 为常数.

此马氏链的 Kolmogorov 向前方程组为

$$\begin{cases} p'_{i0}(t) = -bp_{i0}(t), & i \geq 0 \\ p'_{ij}(t) = bp_{i,j-1}(t) - bp_{ij}(t), & i \geq 0, j \geq 1 \\ p_{ij}(0) = \delta_{ij}, & i, j \geq 0 \end{cases}$$

于是 $p_{00}(t) = e^{-bt}$, $p_{i0}(t) = 0, i \geq 1$. 对 $j \geq 1$, 令

$$r_{ij}(t) = e^{bt} p_{ij}(t).$$

则

$$r'_{ij}(t) = be^{bt}p_{ij}(t) + e^{bt}p'_{ij}(t) = br_{i,j-1}(t).$$

于是对 $j < i$,

$$r_{ij}^{(j)}(t) = b^j r_{i0}(t) = 0, \quad r_{ij}^{(k)}(0) = 0, \quad 0 \leq k \leq j,$$

故 $r_{ij}(t) = 0$, 即 $p_{ij}(t) = 0$. 这样 $r'_{ii}(t) = 0$, $r_{ii}(0) = 1$, 从而 $r_{ii}(t) = 1$, $p_{ii}(t) = e^{-bt}$. 对 $j > i$, 有

$$r_{ij}^{(j-i)}(t) = b^{j-i} r_{ii}(t) = b^{j-i},$$

$$r_{ij}^{(k)}(0) = b^k r_{i,j-k}(0) = 0, \quad 0 \leq k \leq j-i+1, \quad r_{ij}^{(j-i)}(0) = b^{j-i},$$

从而 $r_{ij}(t) = \frac{(bt)^{j-i}}{(j-i)!}$, 即 $p_{ij}(t) = e^{-bt} \frac{(bt)^{j-i}}{(j-i)!}$. 这样,

$$p_{ij}(t) = \begin{cases} 0, & j < i \\ e^{-bt} \frac{(bt)^{j-i}}{(j-i)!}, & j \geq i. \end{cases}$$

容易看出, 此马氏链是参数为 b 的 Poisson 过程.

例 4.2 设马氏链状态空间 $E = \{0, 1\}$, 密度矩阵

$$Q = \begin{pmatrix} -p & p \\ q & -q \end{pmatrix},$$

其中 $p > 0, q > 0$.

此马氏链的向后方程组为

$$P'(t) = QP(t), \quad P(0) = E.$$

由 $\det(\lambda I - Q) = \lambda(\lambda + p + q)$ 知 Q 的特征根为 0 和 $-p - q$, 它们的特征向量分别是 $(1, 1)^T$ 和 $(p, -q)^T$. 令

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

则 $T^{-1}QT = \text{diag}(0, -p - q)$. 于是

$$\begin{aligned}
P(t) &= e^{tQ} = T e^{t \operatorname{diag}(0, -p-q)} T^{-1} = T \operatorname{diag}(1, e^{-(p+q)t}) T^{-1} \\
&= \frac{1}{p+q} \begin{pmatrix} q + p e^{-(p+q)t} & p - p e^{-(p+q)t} \\ q - q e^{-(p+q)t} & p + q e^{-(p+q)t} \end{pmatrix}.
\end{aligned}$$

4.4 Q 矩阵的概率意义

在概率论中我们经常需要研究下面这种类型的集合

$$B = \{\omega, \bigcap_{t \in A} (X_t(\omega) \in \Gamma)\}.$$

当 A 不可数时, 通常 B 是不可测的. 这就需要我们引入可分性的概念. 记 $T = [0, \infty)$

定义 4.4 称取值于 $[-\infty, \infty]$ 的函数 $y(t)$, $t \in T$ 关于 T 的可数稠密子集 S 可分, 若对任意 $t \in T$, 有 S 中子列 $\{r_n\}$, 使得 $r_n \rightarrow t$ 且 $y(r_n) \rightarrow y(t)$. 称 S 为此函数的可分集.

定义 4.5 称随机过程 $\{X_t\}_{t \in T}$ 可分, 若存在零概率集 N 和 T 的可数稠密子集 S , 使对一切 $\omega \notin N$, 样本函数 $X_t(\omega)$ 关于 S 可分. 称 S 为可分集, N 为例外集.

定义 4.6 称 $\{X_t(\omega)\}_{t \in T}$ 是完全可分的, 若它关于 T 的任一可数稠密子集可分.

定理 4.10 任意随机过程必存在可分修正.

定义 4.7 称过程 X_t 随机连续, 若对任意的 $t_0 \geq 0$, 当 $t \rightarrow t_0$ 时, X_t 依概率收敛到 X_{t_0} .

定理 4.11 若可分过程 $\{X_t\}$ 随机连续, 则它是完全可分的.

证明 首先, 对任意的 $\{t_i\} \subset T$, $t_i \rightarrow t_0$, 有 $X_{t_i} \xrightarrow{P} X_{t_0}$. 所以存在子列 t'_i 使得 $X_{t'_i} \xrightarrow{a.s.} X_{t_0}$.

由可分性, 存在 T 的可数稠集 S 及零概率集 N , 使得当 $\omega \notin N$ 时, 对任意的 $t \in T$, 有 $s_n \in S$, $s_n \rightarrow t$, 且 $X_{s_n}(\omega) \rightarrow X_t(\omega)$.

设 R 是 T 的任意可数稠集. 此时存在一个零概率集 N_1 使得, 当 $\omega \notin N_1$ 时, 对任意的 $s \in S$, 存在 $r_n \in R$, $r_n \rightarrow s$ 且 $X_{r_n}(\omega) \rightarrow X_s(\omega)$. 这样, 当 $\omega \notin N \cup N_1$ 时, 对上面的 s_n , 有 $r_{mn} \in R$, 使得当 $m \rightarrow \infty$ 时, $r_{mn} \rightarrow s_n$, 且 $X_{r_{mn}}(\omega) \rightarrow X_{s_n}(\omega)$. 容易看出, 存在 $\{r'_{mn}\}$ 的子列 r'_n , 满足 $r'_n \rightarrow t$, 且 $X_{r'_n}(\omega) \rightarrow X_t(\omega)$. 这就证明了 X_t 关于 R 可分. \square

引理 4.2 设 $(p_{ij}(t))$ 标准的, 则 X_t 随机连续.

证明 当 $h < 0$ 时,

$$\begin{aligned}
 P(X_{t+h} = X_t) &= \sum_j P(X_t = j | X_{t+h} = j) P(X_{t+h} = j) \\
 &= \sum_j P(X_{t+h} = j) p_{jj}(-h).
 \end{aligned}$$

由 Fatou 引理,

$$\liminf_{h \rightarrow 0-} P(X_{t+h} = X_t) \geq \sum_j P(X_t = j) = 1.$$

故

$$\lim_{h \rightarrow 0-} P(X_{t+h} = X_t) = 1.$$

而当 $h > 0$ 时, 类似的, 应用控制收敛定理可得

$$\lim_{h \rightarrow 0+} P(X_{t+h} = X_t) = 1.$$

这样,

$$\lim_{h \rightarrow 0} P(X_{t+h} \neq X_t) = 0.$$

从而对任意的 $\varepsilon > 0$,

$$\lim_{h \rightarrow 0} P(|X_{t+h} - X_t| > \varepsilon) \leq \lim_{h \rightarrow 0} P(X_{t+h} \neq X_t) = 0.$$

□

定理 4.12 对任意的 $i \in E$, 有

$$P_i(X_s = i, 0 \leq s \leq t) = e^{-q_i t}.$$

证明 取可分集

$$\left\{ \frac{kt}{2^n}, k = 0, 1, \dots, 2^n; n = 0, 1, \dots \right\},$$

由完全可分性,

$$\begin{aligned}
P_i(X_u = i, 0 \leq u \leq t) &= P_i\left(\bigcap_{n=1}^{\infty} \{X_{\frac{kt}{2^n}} = i, 0 \leq k \leq 2^n\}\right) \\
&= \lim_{n \rightarrow \infty} P_i(X_{t/2^n} = i, 0 \leq k \leq 2^n) = \lim_{n \rightarrow \infty} \left[p_{ii}\left(\frac{t}{2^n}\right)\right]^{2^n} \\
&= \lim_{n \rightarrow \infty} \exp\left\{\frac{\ln p_{ii}\left(\frac{t}{2^n}\right)}{\frac{t}{2^n}} \frac{t}{2^n} 2^n\right\} = e^{-q_i t}.
\end{aligned}$$

□

由上面定理, 若 $q_i = 0$, 则从状态 i 出发, 以概率 1 停留在 i ; 若 $q_i = \infty$, 则从 i 出发立刻离开 i ; 若 $0 < q_i < \infty$, 则从 i 出发, 在 i 停留一段时间之后离开 i .

令

$$\tau = \inf\{t : X_t \neq X_0\}.$$

则 τ 表示马氏链停留在初始状态的时间. 由马氏链可分, 故 τ 是 \mathcal{F}_∞ 可测的, 这里 $\mathcal{F}_\infty = \sigma(X_s, s \geq 0)$.

对任意的 $t \geq 0$, 有

$$\{X_u = i, 0 \leq u \leq t\} \supset \{\tau > t\} \supset \{X_u = i, 0 \leq u \leq t + \frac{1}{n}\}.$$

所以

$$e^{-q_i t} \geq P_i(\tau > t) \geq e^{-q_i(t+1/n)}.$$

令 $n \rightarrow \infty$, 对任意 t ,

$$P_i(\tau > t) = e^{-q_i t}.$$

特别的,

$$E^i[\tau] = q_i^{-1}.$$

因此, q_i^{-1} 是在状态 i 的期望停留时间.

我们不加证明的给出下面定理.

定理 4.13 若马氏链 X_t 的一切 $q_i < \infty$, 则它必存在一个修正, 该修正的全部样本函数在 $E \cup \{\infty\}$ 中右连续.

下面我们总假定 X_t 的样本函数是右连续的. 这样, 对任意 $j \in E$,

$$\{X_\tau = j\} \cap \{\tau < \infty\} \in \mathcal{F}_\infty.$$

定理 4.14 设 $0 < q_i < \infty, j \neq i$. 则对任意 $t \geq 0$, 有

$$P_i(\tau \leq t, X_\tau = j) = (1 - e^{-q_i t}) \frac{q_{ij}}{q_i}. \quad (4.9)$$

证明 对 $n \geq 1$, 令

$$\tau_n = \begin{cases} \frac{kt}{2^n}, & \text{若 } \frac{(k-1)t}{2^n} < \tau \leq \frac{kt}{2^n}, k = 0, 1, \dots, \\ \infty, & \text{若 } \tau = \infty. \end{cases}$$

易见 $\tau_n \in \mathcal{F}_\infty$, 且当 $n \rightarrow \infty$ 时, $\tau_n \downarrow \tau$. 由 $P_i(\tau < \infty) = 1$ 及 X_t 右连续, 故 $\lim_{n \rightarrow \infty} X_{\tau_n} = X_\tau, P^j\text{-a.s.}$ 因此,

$$\begin{aligned} \{\tau < t, X_\tau = j\} &\subset \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} \{\tau_n \leq t, X_{\tau_n} = j\} \\ &= \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{\tau_n \leq t, X_{\tau_n} = j\} \subset \{\tau \leq t, X_\tau = j\}. \end{aligned}$$

注意到 $P_i(\tau = t) = 0$, 故

$$\lim_{n \rightarrow \infty} P_i(\tau_n \leq t, X_{\tau_n} = j) = P_i(\tau \leq t, X_\tau = j).$$

这样,

$$\begin{aligned} P_i(\tau_n \leq t, X_{\tau_n} = j) &= \sum_{l=1}^{2^n} P_i\left(\frac{(l-1)t}{2^n} < \tau \leq \frac{lt}{2^n}, X_{\frac{lt}{2^n}} = j\right) \\ &= \sum_{l=1}^{2^n} P_i\left(\tau > \frac{(l-1)t}{2^n}, X_{\frac{lt}{2^n}} = j\right) = \sum_{l=1}^{2^n} P_i\left(X_s = i, 0 \leq s \leq \frac{(l-1)t}{2^n}, X_{\frac{lt}{2^n}} = j\right) \\ &= \sum_{l=1}^{2^n} e^{-q_i \frac{(l-1)t}{2^n}} p_{ij}\left(\frac{t}{2^n}\right) = \frac{1 - e^{-q_i t}}{1 - e^{-q_i \frac{t}{2^n}}} p_{ij}\left(\frac{t}{2^n}\right) \\ &\rightarrow (1 - e^{-q_i t}) \frac{q_{ij}}{q_i} \quad (n \rightarrow \infty). \end{aligned}$$

□

这样, 令 $t \rightarrow \infty$, 可以得到

$$P_i(X_\tau = j) = \frac{q_{ij}}{q_i}.$$

上式的直观意思是说, 马氏链 X_t 在首次离开初始状态 i 之后, 以概率 $\frac{q_{ij}}{q_i}$ 跳到状态 j . 若

$$q_i = \sum_{j \neq i} q_{ij} < \infty,$$

则

$$P_i(X_\tau \in (E - \{i\})) = \sum_{j \neq i} \frac{q_{ij}}{q_i} = 1,$$

即以 P_i 概率 1, τ 是跳跃点.

前面我们已经给出向后方程组成立的充分必要条件, 下面我们研究它的积分形式.

定理 4.15 向后方程组成立的充分必要条件是对任意的 $i, j \in E, t \geq 0$,

$$p_{ij}(t) = \delta_{ij}e^{-q_i t} + \sum_{k \neq i} \int_0^t e^{-q_i s} q_{ik} p_{kj}(t-s) ds. \quad (4.10)$$

证明 若积分方程组 (4.10) 成立, 只需要在两边对 t 求导, 即可得到向后方程组. 下面我们证明必要性. 设向后方程组成立, 则 Q 矩阵是保守的. 先设 $q_i > 0$. 若 $j \neq i$, 由 $P_i(X_\tau \in E - \{i\}) = 1$ 可知

$$p_{ij}(t) = P_i(\tau \leq t, X_t = j) = \sum_{k \neq i} P_i(\tau \leq t, X_\tau = k, X_t = j). \quad (4.11)$$

应用定理 4.14 的证明方法,

$$P_i(\tau \leq t, X_\tau = k, X_t = j) = \lim_{n \rightarrow \infty} P_i(\tau_n \leq t, X_\tau = k, X_t = j).$$

而

$$\begin{aligned} P_i(\tau_n \leq t, X_\tau = k, X_t = j) &= \sum_{l=1}^{2^n} P_i(\tau > \frac{(l-1)t}{2^n}, X_{lt/2^n} = k, X_t = j) \\ &= \sum_{l=1}^{2^n} e^{-q_i \frac{(l-1)t}{2^n}} p_{ik}(t/2^n) p_{kj}(t - (lt)/2^n) \\ &= \left[\sum_{l=1}^{2^n} e^{-q_i \frac{(l-1)t}{2^n}} p_{kj}(t - lt/2^n) \cdot \frac{t}{2^n} \right] \frac{p_{ik}(t/2^n)}{t/2^n}. \end{aligned}$$

令 $n \rightarrow \infty$, 上式极限为

$$\int_0^t e^{-q_i s} q_{ik} p_{kj}(t-s) ds = \int_0^t e^{-q_i(t-s)} q_{ik} p_{kj}(s) ds,$$

从而

$$p_{ij}(t) = \int_0^t e^{-q_i(t-s)} q_{ik} p_{kj}(s) ds.$$

若 $j = i$, 则

$$\begin{aligned} p_{ii}(t) &= P_i(\tau > t, X_t = i) + P_i(\tau \leq t, X_t = i) \\ &= P_i(\tau > t) + P_i(\tau \leq t, X_t = i) \\ &= e^{-q_i t} + \sum_{k \neq i} \int_0^t e^{-q_i(t-s)} q_{ik} p_{ki}(s) ds. \end{aligned}$$

这样, 积分方程组 (4.10) 成立.

若 $q_i = 0$, 则对 $k \neq i$, $q_{ik} = 0$, 因此对 $j \neq i$, $p_{ij}(t) = 0$, $p_{ii}(t) = 1$. 积分方程组 (4.10) 仍旧成立. \square

(4.10) 可以直观上理解为: 马氏链从状态 i 出发, 在时刻 t 到达 j , 有两种可能的方式, 一种是一直停留在 i , 它的概率是 $\delta_{ij} e^{-q_i t}$, 另一种是在时刻 t 之前先离开 i , 然后再到达 j , 它首次离开 i 的时间发生在 $(s, s + \Delta)$ 内, 并跳跃到另一状态 k , 这个概率是 $(e^{-q_i s} - e^{-q_i(s+\Delta)}) \frac{q_{ik}}{q_i} \approx e^{-q_i s} q_{ik} \Delta$, 最后以概率 $p_{kj}(t-s)$ 转移到状态 j .

4.5 嵌入链

在这一节中, 我们假设 $0 < q_i < \infty$, Q 矩阵保守, 并且 X_t 的所有轨道在 $E \cup \{\infty\}$ 上是右连续的.

同前面一样, 令 $\mathcal{F}_t = \sigma(X_s, 0 \leq s \leq t)$, 则 $\{\mathcal{F}_t\}$ 是一个过滤, 且 X_t 关于此过滤适应, 即对任意 t , $X_t \in \mathcal{F}_t$. 通常我们把所有的零概率集都添加到 \mathcal{F}_t 中, 仍旧记为 \mathcal{F}_t .

定义 4.8 设 τ 是取值于 $[0, \infty]$ 的随机变量. 称 τ 为 $\{\mathcal{F}_t\}$ -停时, 若对任意的 t , $\{\tau \leq t\} \in \mathcal{F}_t$.

设 A 是闭集, 定义

$$D_A = \inf\{t \geq 0, X_t \in A\}.$$

应用 X_t 的右连续性可以证明 D_A 是 $\{\mathcal{F}_t\}$ -停时. 事实上,

$$\{D_A \leq t\} = \left\{ \inf_{\substack{r \leq t \\ r \in \mathbb{Q}}} d(X_r, A) = 0 \right\} \in \mathcal{F}_t,$$

其中 $d(x, y)$ 是距离函数.

对于停时 τ , 我们可以定义 σ -代数

$$\mathcal{F}_\tau = \{A \in \mathcal{F}_\infty, A \cap (\tau \leq t) \in \mathcal{F}_t\}.$$

可以证明若停时 $\tau_1 \leq \tau_2$, 则 $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

在 $\{\tau < \infty\}$ 上, 我们可以定义

$$X_\tau(\omega) = X_{\tau(\omega)}(\omega).$$

可以证明, X_τ 是 \mathcal{F}_τ 可测的.

下面我们证明右连续马氏链的强马氏性.

定理 4.16 设 τ 是停时, 则在 $\{\tau < \infty\}$ 上, 有

$$P_i(X_{\tau+t} = j | \mathcal{F}_\tau) = p_{X_\tau, j}(t) \quad P_i\text{-a.s.}$$

在 $\{\tau < \infty\}$ 上, 我们定义新的过程

$$Y_t = X_{\tau+t}, \quad t \geq 0.$$

则 Y_t 也是一个马氏链, 它与 X_t 具有相同的转移函数, 并且在条件 $X_\tau = j$ 下, 它是与 \mathcal{F}_τ 独立的.

设 τ_1 是马氏链 X_t 首次离开初始状态的时刻, 即

$$\tau_1 = \inf\{t \geq 0, X_t \neq X_0\}.$$

由 $q_i > 0$ 及 Q 矩阵保守, 故 τ_1 是 X_t 的第一个跳跃点. 一般来说, τ_1 不一定是 $\{\mathcal{F}_t\}$ -停时. 但是, 对任意的 t , 我们有

$$\{\tau_1 < t\} = \bigcup_{\substack{r < t \\ r \in R}} \{X_r \neq X_0\} \in \mathcal{F}_t,$$

其中 R 是一个可分集. 这样, 记

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s,$$

则

$$\{\tau_1 \leq t\} = \bigcap_{\varepsilon > 0} \{\tau_1 < t + \varepsilon\} \in \mathcal{F}_{t+},$$

即 τ_1 是 $\{\mathcal{F}_{t+}\}$ -停时, 我们称为宽停时. 由于

$$P(\tau_1 < \infty) = \sum_{i \in E} P(X_0 = i) P_i(\tau_1 < \infty) = 1,$$

我们可以定义

$$Y_t^{(1)} = X_{\tau_1+t}, \quad t \geq 0.$$

应用强马氏性可以得到

(i) $Y_t^{(1)}$ 是马氏链, 它的转移矩阵与 X_t 相同, 且几乎所有的样本函数是右连续的.

(ii) 对任意的 $A \in \mathcal{F}_{\tau_1}$, $C \in \sigma(Y_s^{(1)}, s \geq 0)$, 有

$$P(AC|Y_0^{(1)} = i) = P(A|Y_0^{(1)} = i)P(C|Y_0^{(1)} = i).$$

令 σ_1 是 $Y_t^{(1)}$ 的第一个跳跃点, 则

$$P(\sigma_1 > t | Y_0^{(1)} = i) = e^{-q_i t}.$$

显然的 $\sigma_1 < \infty$. 令 $\tau_2 = \tau_1 + \sigma_1 < \infty$, 则 τ_2 是 X_t 的第二个跳跃点, 所以

$$\{\tau_2 < t\} = \bigcap_{\substack{0 < s_1 < s_2 < t \\ s_1, s_2 \in R}} \{X_{s_1} \neq X_0, X_{s_2} \neq X_{s_1}\} \in \mathcal{F}_t,$$

即 τ_2 是宽停时. 继续下去, 我们可以得到 X_t 的一系列跳跃点

$$0 = \tau_0 < \tau_1 < \tau_2 < \cdots.$$

记 $\tau_\infty = \lim_{n \rightarrow \infty} \tau_n$, 称为马氏链的第一个飞跃点. 如果

$$P(\tau_\infty = \infty) = 1, \tag{4.12}$$

则 X_t 的几乎所有的样本函数是阶梯的.

定理 4.17 若 $q = \sup_{i \in E} q_i < \infty$, 则 (4.12) 成立.

证明 对任意的 $t > 0$,

$$\begin{aligned} P(\tau_{n+1} - \tau_n > t) &= \sum_{k \in E} P(\sigma_n > t | X_{\tau_n} = k) P(X_{\tau_n} = k) \\ &= \sum_{k \in E} P(X_{\tau_n} = k) e^{-q_i t} \geq e^{-qt}. \end{aligned}$$

这样,

$$P(\limsup_{n \rightarrow \infty} (\tau_{n+1} - \tau_n) > t) \geq e^{-qt},$$

即有无穷多 n 使得 $\tau_{n+1} - \tau_n > t$ 的概率不小于 e^{-qt} . 因此

$$P(\tau_\infty = \infty) \geq e^{-qt}.$$

由 t 的任意性可知 (4.12) 成立.

下面我们考虑离散时间随机过程 $Z_n = X_{\tau_n}$. 令

$$u_{ij} = \begin{cases} \frac{q_{ij}}{q_i}, & \text{若 } i \neq j, \\ 0, & \text{若 } i = j. \end{cases}$$

定理 4.18 Z_n 是离散参数马氏链, 状态空间是 E , 转移矩阵是 (u_{ij}) . 我们称 Z_n 为 X_t 的嵌入链.

证明 由强马氏性,

$$\begin{aligned} & P(Z_{n+1} = j | Z_0 = i_0, \dots, Z_{n-1} = i_{n-1}, Z_n = i) \\ &= P(Z_{n+1} = j | Z_n = i) = P(Y_{\sigma_1}^{(n)} = j | Y_n^{(1)} = i) = u_{ij} \end{aligned}$$

□

我们可以通过嵌入链的常返性来描述 (4.12). 事实上, 如果状态 i 对嵌入链 Z_n 是常返的, 则从 i 出发可以返回 i 无穷多次, 设返回时间是 N_1, N_2, \dots , 即

$$N_1 = \inf\{n \geq 1, Z_n = i\}, \quad N_2 = \inf\{n \geq N_1 + 1, Z_n = i\}, \dots$$

这样, 在概率 P_i 下,

$$\tau_\infty \geq \sum_{m=1}^{\infty} (\tau_{N_m+1} - \tau_{N_m}) = \sum_{m=1}^{\infty} \sigma_{N_m} = \infty.$$

这就证明了

定理 4.19 若 i 是嵌入链 Z_n 的常返状态, 则 $P_i(\tau_\infty = \infty) = 1$. 特别的, 若 Z_n 所有状态常返, 则 (4.12) 成立.

可以看到, 当 $\tau_\infty = \infty$ 时, 马氏链具有特别简单的结构, 它可以由嵌入链 Z_n 及宽停时列 τ_n 决定. 反过来, 如果给定符合一定条件的马氏链 Z_n 及 τ_n , 在下一节里, 我们构造出马氏链 X_t , 使得它以 Z_n 为嵌入链, 以 τ_n 为跳跃时间.

4.6 马氏链的构造

给定转移矩阵 (u_{ij}) 与 q_i . 我们假设

$$u_{ii} = 0, \quad 0 < q_i < \infty, \quad i \in E.$$

令

$$q_{ii} = -q_i, \quad q_{ij} = q_i u_{ij}, \quad i \neq j, \quad Q = (q_{ij}).$$

我们希望构造一个以 Q 为密度矩阵的马氏链 X_t . 这样的 X_t 称为 Q 过程.

在某个给定的概率空间 (Ω, \mathcal{F}, P) 上, 我们可以构造

- (i) 离散参数马氏链 Z_n , 其转移概率为 u_{ij} ;
- (ii) 独立同分布随机变量列 Λ_n , 它们服从参数为 1 的指数分布, 并且与 $\{Z_n\}$ 也是独立的.

我们把 q_i 理解为 E 上的函数, 即

$$q: E \rightarrow [0, \infty), \quad q(i) = q_i.$$

令

$$T_{n+1} = \frac{\Lambda_n}{q(Z_n)} = \sum_{i \in E} \frac{\Lambda_n}{q_i} 1_{\{Z_n=i\}},$$

则

$$P(T_n > x | Z_n = i) = e^{-q_i x},$$

即在概率 $P(\cdot | Z_n = i)$ 下, T_n 服务参数为 q_i 的指数分布. 进一步的, 给定条件 Z_0, Z_1, \dots, Z_{n-1} 下, T_1, T_2, \dots, T_n 是独立的指数随机变量, 且参数分别为 $q(Z_0), q(Z_1), \dots, q(Z_{n-1})$.

令 $\tau_0 = 0, \tau_n = \sum_{k \leq n-1} T_k, n \geq 1$. 我们定义随机过程 X_t , 使得它在区间 $[\tau_n, \tau_{n+1})$ 取常值, 而在每个 τ_n 处进行转移, 即

$$X_t = \sum_{n=0}^{\infty} Z_n 1_{[\tau_n, \tau_{n+1})}(t).$$

注 4.1 直观上来看, 过程 X_t 从某个状态 i 出发, 即 $Z_0 = i$, 经历时间 T_1 (服从参数为 q_i 的指数分布) 之后, 以概率 u_{ij} 转移到状态 $j \neq i$, 然后在状态 j 停留 T_2 (服从参数为 q_j 的指数分布) 的时间, 再以概率 u_{jk} 转移到状态 $k \neq j$. 如此继续下去, 就可以得到 X_t .

令 $\tau_\infty = \lim_{n \rightarrow \infty} \tau_n = \sum_{n \geq 1} T_n$. 实际上我们只是在 τ_∞ 之前确定了过程 X_t . 为了使得我们的过程定义在区间 $[0, \infty)$ 上, 我们需要考虑在什么样的条件下 $\tau_\infty = \infty$. 我们需要下面的引理.

引理 4.3 设 Y_n 是独立随机变量列, 分别服从参数为 λ_n 的指数分布, 则 $\sum_n Y_n < \infty$ a.s. 当且仅当 $\sum_n \lambda_n^{-1} < \infty$.

证明 若 $\sum_n \lambda_n^{-1} < \infty$, 则

$$E[\sum_n Y_n] = \sum_n E[Y_n] = \sum_n \lambda_n^{-1} < \infty,$$

故 $\sum_n Y_n < \infty$ a.s..

若 $\sum_n \lambda_n^{-1} = \infty$, 则 $\prod_n (1 + 1/\lambda_n) = \infty$. 由单调收敛定理,

$$\begin{aligned} E[\exp(-\sum_n Y_n)] &= \prod_n E(\exp(-Y_n)) \\ &= \prod_n \left(1 + \frac{1}{\lambda_n}\right)^{-1} = 0. \end{aligned}$$

这样,

$$P(\sum_n Y_n = \infty) = 1.$$

□

这样, 在已知马氏链 $\{Z_n\}$ 的条件下,

$$\tau_\infty = \sum_n T_n = \sum_n \frac{\Lambda_n}{q(Z_n)}$$

是独立指数随机变量的和, 它们的参数分别是 $q(Z_n)$. 由上面的引理,

$$P(\tau_\infty < \infty | \{Z_n\}) = \begin{cases} 1, & \text{若 } \sum_n \frac{1}{q(Z_n)} < \infty, \\ 0, & \text{若 } \sum_n \frac{1}{q(Z_n)} = \infty. \end{cases}$$

于是

$$P(\tau_\infty < \infty | \{Z_n\}) = 1_{\{\sum_n 1/q(Z_n) < \infty\}} \quad \text{a.s..}$$

两边关于概率 P_i 取期望, 则

$$P_i(\tau_\infty < \infty) = P_i\left(\sum_n \frac{1}{q(Z_n)} < \infty\right).$$

定理 4.20 设对任意的 $i \in E$, 在概率 P_i 下以概率 1 成立

$$\sum_{n=0}^{\infty} \frac{1}{q(Z_n)} = \infty, \quad (4.13)$$

则 $P(\tau_\infty = \infty) = 1$. 特别的, 若 $q = \sup_i q_i < \infty$, 则 (4.13) 成立.

下面我们简述连续时间马氏链的另一种构造. 设 $\{\Lambda_n^j, n \geq 1, j \in E\}$ 是一列独立同分布的指数随机变量, 参数为 1. 我们递归的来给出 T_n, Z_n : 若 $Z_n = i$, 则令

$$\begin{aligned}
T_{n+1}^j &= \frac{\Lambda_{n+1}^j}{q_{ij}}, \quad j \neq i, \\
T_{n+1} &= \inf_{j \neq i} T_{n+1}^j, \\
Z_{n+1} &= \begin{cases} j, & \text{若 } T_{n+1}^j = T_{n+1} < \infty, \\ i, & \text{若 } T_{n+1} = \infty. \end{cases}
\end{aligned}$$

在条件 $Z_n = i$ 下, T_{n+1}^j 是独立的指数随机变量, 参数分别为 q_{ij} . 由定理 4.21 可知, T_{n+1} 是参数为 $q_i = \sum_{j \neq i} q_{ij}$ 的指数随机变量, Z_{n+1} 具有分布 $\{u_{ij}, j \in E\}$, 并且 T_{n+1} 与 Z_0, \dots, Z_n, Z_{n+1} 及 T_1, \dots, T_n 独立. 然后应用我们上面的构造即可得到所需的马氏链.

定理 4.21 设 Y_n 是独立的指数随机变量, 参数分别为 λ_n , $0 < \lambda = \sum_n \lambda_n < \infty$. 令 $Y = \inf_n Y_n$, 且极小值在第 N 个随机变量处取得, 即在 $\{Y = Y_n\}$ 上, $N = n$. 则 N 与 Y 独立, Y 服从参数为 λ 的指数分布, $P(N = n) = \lambda_n / \lambda$.

证明 注意到在 $\{N = n\}$ 上, 对任意的 $m \neq n$, $Y_n < Y_m$. 这样

$$\begin{aligned}
P(N = n, Y \geq t) &= P(Y_n \geq t, Y_m > Y_n, m \neq n) \\
&= \int_t^\infty \lambda_n e^{-\lambda_n s} P(Y_m > s, m \neq n) = \int_t^\infty \lambda_n e^{-\lambda_n s} \prod_{m \neq n} e^{-\lambda_m s} \\
&= \int_t^\infty \lambda_n e^{-\lambda s} ds = \frac{\lambda_n}{\lambda} e^{-\lambda t}.
\end{aligned}$$

□

如果不作 (4.13) 的限制, 即允许 $P(\tau_\infty < \infty) > 0$, 此时我们称 Q 是爆炸 (*explosive*) 的. 可以看到, 当 Q 爆炸时, Q 过程不是唯一的. Doob 的著名例子可以说明这个问题.

例 4.3 假设 $P(\tau_\infty < \infty) = 1$, 考虑一系列独立的 Q 过程 $X_t^{(n)}$, 它们都是可分, 右连续的, 具有相同的初始分布 π . 记 $\tau_\infty^{(n)}$ 分别是 $X_t^{(n)}$ 的第一个跳跃点, 令 $\tau_0 = 0$, $\tau_n = \sum_{v=1}^n \tau_\infty^{(v)}$. 定义新的过程

$$X_t = X_{t-\tau_{n-1}}^{(n)}, \quad \tau_{n-1} \leq t < \tau_n.$$

我们称 X_t 为 *Doob* 过程. 由于 X_t 的转移函数完全由 Q 及 π 决定, 我们也称为 (Q, π) 过程.

当 Q 矩阵爆炸时, 我们可以从这无穷多个 Q 过程中选取一个特殊的过程. 设 X_t 是 Q 过程, 称它是极小 (*minimal*) 的, 若当 $t \geq \tau_\infty$ 时, $X_t = \infty$.

我们还需要证明如上构造的 X_t 是以 Q 为密度矩阵的马氏链. 设 X_t 是极小的 Q 过程, 令

$$p_{ij}(t) = P_i(X_t = j), \quad P(t) = (p_{ij}(t)).$$

在条件 $X_0 = i$ 下, τ_1 服从参数为 q_i 的指数分布, X_{τ_1} 的分布为 $(u_{ik}, k \in E)$. 在条件 $\tau_1 = s, X_{\tau_1} = k$ 下, $(X_{s+t})_{t \geq 0}$ 也是一个 Q 过程. 这样,

$$\begin{aligned} P_i(X_t = j, t < \tau_1) &= e^{-q_i t} \delta_{ij}, \\ P_i(\tau_1 \leq t, X_{\tau_1} = k, X_t = j) &= \int_0^t q_i e^{-q_i s} u_{ik} p_{kj}(t-s) ds. \end{aligned}$$

于是,

$$\begin{aligned} p_{ij}(t) &= P_i(X_t = j, t < \tau_1) + \sum_{k \neq i} P_i(\tau_1 \leq t, X_{\tau_1} = k, X_t = j) \\ &= e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} u_{ik} p_{kj}(t-s) ds. \end{aligned} \quad (4.14)$$

这样,

$$e^{q_i t} p_{ij}(t) = \delta_{ij} + \int_0^t \sum_{k \neq i} e^{q_i s} q_{ik} p_{kj}(s) ds. \quad (4.15)$$

这说明 $p_{ij}(t)$ 连续, 故上式右边是一个连续函数的积分, 故 $p_{ij}(t)$ 连续可微. 从而

$$e^{q_i t} (q_i p_{ij}(t) + p'_{ij}(t)) = \sum_{k \neq i} e^{q_i t} q_{ik} p_{kj}(t).$$

这就证明了 $p'_{ij}(t) = \sum_{k \in E} q_{ik} p_{kj}(t)$, 特别的, $p'_{ii}(0) = -q_i, p'_{ij}(t) = q_{ij}$.

为证明 X_t 的马氏性, 只需要证明对任意的 $i_1, \dots, i_n \in E, t_1 < \dots < t_n$,

$$\begin{aligned} &P_i(X_{t_1} = i_1, \dots, X_{t_n} = i_n) \\ &= p_{ii_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1} i_n}(t_n - t_{n-1}). \end{aligned}$$

其证明我们留作练习.

这样, 我们就证明了如上构造的极小过程 X_t 是连续时间马氏链, 并且它的转移函数满足向后方程组. 实际上, 对于给定的 Q 矩阵, 我们给出了向后方程组的一组解, 或者说, 证明了向后方程组的存在性. 注意到当 Q 爆炸时, 向后方程组没有解的唯一性. 设 $\tilde{P}(t) = (\tilde{p}_{ij}(t))$ 是向后方程组的另一组非负解, 则

$$\tilde{p}_{ij}(t) = e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} u_{ik} \tilde{p}_{kj}(t-s) ds. \quad (4.16)$$

由 $\tilde{P}(t) \geq 0$, 对所有的 i, j ,

$$P_i(X_t = j, t < \tau_1) = e^{-q_i t} \delta_{ij} \leq \tilde{p}_{ij}(t).$$

假设当 n 时有

$$P_i(X_t = j, t < \tau_n) \leq \tilde{p}_{ij}(t), \quad \forall i, j, t,$$

类似于 (4.14) 的证明,

$$\begin{aligned} & P(X_t = j, t < \tau_{n+1}) \\ &= e^{-q_i t} \delta_{ij} + \sum_{k \neq i} \int_0^t q_i e^{-q_i s} u_{ik} P_k(X_{t-s} = j, t-s < \tau_n) ds. \end{aligned} \quad (4.17)$$

这样, 比较 (4.16) 与 (4.17) 可知,

$$P(X_t = j, t < \tau_{n+1}) \leq \tilde{p}_{ij}(t).$$

于是, 对任意的 n ,

$$P(X_t = j, t < \tau_n) \leq \tilde{p}_{ij}(t).$$

特别的,

$$p_{ij}(t) = \lim_{n \rightarrow \infty} P(X_t = j, t < \tau_n) \leq \tilde{p}_{ij}(t),$$

即我们构造的 $P(t)$ 是向后方程组的最小解.

定理 4.22 给定 Q 矩阵, 向后方程组

$$P'(t) = QP(t), \quad P(0) = I$$

有最小解 $P(t)$, 并且 $P(t)$ 是极小 Q 过程的转移矩阵, 它满足 $C-K$ 方程

$$P(t+s) = P(t)P(s).$$

接下来我们考虑向前方程组. 我们首先给出一个关于时间逆转的引理.

引理 4.4 对任意的 n, t 及状态 i_0, i_1, \dots, i_n ,

$$\begin{aligned}
& q_{i_n} P(\tau_n \leq t < \tau_{n+1} | Y_0 = i_0, Y_1 = i_1, \dots, Y_n = i_n) \\
& = q_{i_0} P(\tau_n \leq t < \tau_{n+1} | Y_0 = i_n, \dots, Y_{n-1} = i_1, Y_n = i_0).
\end{aligned}$$

证明 在条件 $Y_0 = i_0, \dots, Y_n = i_n$ 下, T_1, \dots, T_{n+1} 是独立的指数随机变量列, 参数分别是 q_{i_0}, \dots, q_{i_n} . 这样, 上式左边为

$$\int_{\Delta(t)} q_{i_n} \exp(-q_{i_n}(t - s_1 - \dots - s_n)) \prod_{k=1}^n q_{i_k} \exp(-q_{i_{k-1}} s_k) ds_k,$$

其中

$$\Delta(t) = \{(s_1, \dots, s_n), s_1 + \dots + s_n \leq t, s_1, \dots, s_n \geq 0\}.$$

做变量替换 $u_1 = t - s_1 - \dots - s_n, u_k = s_{n-k+2}, 2 \leq k \leq n$, 可以得到

$$\begin{aligned}
& q_{i_n} P(\tau_n \leq t < \tau_{n+1} | Y_0 = i_0, \dots, Y_n = i_n) \\
& = \int_{\Delta(t)} q_{i_0} \exp(-q_{i_0}(t - u_1 - \dots - u_n)) \prod_{k=1}^n q_{i_{n-k+1}} \exp(-q_{i_{n-k+1}} u_k) du_k \\
& = q_{i_0} P(\tau_n \leq t < \tau_{n+1} | Y_0 = i_n, \dots, Y_n = i_0).
\end{aligned}$$

□

定理 4.23 向后方程组的最小非负解 $P(t)$ 也是向前方程组

$$P'(t) = P(t)Q, \quad P(0) = I$$

的最小非负解.

证明 设 X_t 是以 Q 为密度矩阵的极小马氏链, 则由定理 4.22,

$$\begin{aligned}
p_{ij}(t) &= P_i(X_t = j) \\
&= \sum_{n=0}^{\infty} \sum_{k \neq j} P_i(\tau_n \leq t < \tau_{n+1}, Z_{n-1} = k, Z_n = j).
\end{aligned}$$

应用引理 4.4, 对 $n \geq 1$,

$$\begin{aligned}
& P_i(\tau_n \leq t < \tau_{n+1} | Z_{n-1} = k, Z_n = j) \\
&= \frac{q_i}{q_j} P_j(\tau_n \leq t < \tau_{n+1} | Z_1 = k, Z_n = i) \\
&= \frac{q_i}{q_j} \int_0^t q_j e^{-q_j s} P_k(\tau_{n-1} \leq t-s < \tau_n | Z_{n-1} = i) ds \\
&= q_i \int_0^t e^{-q_j s} \frac{q_k}{q_i} P_i(\tau_{n-1} \leq t-s < \tau_n | Z_{n-1} = k) ds.
\end{aligned}$$

这样,

$$\begin{aligned}
p_{ij}(t) &= \delta_{ij} e^{-q_i t} + \sum_{n=1}^{\infty} \sum_{k \neq j} \int_0^t P_i(\tau_{n-1} \leq t-s < \tau_n | Z_{n-1} = k) \\
&\quad \cdot P_i(Z_{n-1} = k, Z_n = j) q_k e^{-q_j s} ds \\
&= \delta_{ij} e^{-q_i t} + \sum_{n=1}^{\infty} \sum_{k \neq j} \int_0^t P_i(\tau_{n-1} \leq t-s < \tau_n, Z_{n-1} = k) q_k u_{kj} e^{-q_j s} ds \\
&= \delta_{ij} e^{-q_i t} + \int_0^t \sum_{k \neq j} p_{ik}(t-s) q_{kj} e^{-q_j s} ds,
\end{aligned} \tag{4.18}$$

这即是向前方程组的积分形式. 我们改变下形式

$$p_{ij}(t) e^{q_j t} = \delta_{ij} + \int_0^t \sum_{k \neq j} p_{ik}(u) q_{kj} e^{q_j u} du. \tag{4.19}$$

由 (4.15) 可知 $e^{q_i t} p_{ij}(t)$ 单增. 而对任意 t , 由于 (4.19) 中积分有限, 故

$$\sum_{k \neq j} p_{ik}(t) q_{kj} < \infty,$$

从而 $\sum_{k \neq j} p_{ik}(t) q_{kj}$ 关于 $u \in [0, t]$ 一致收敛. 这样, 在 (4.19) 两边对 t 求导,

$$p'_{ij}(t) + p_{ij}(t) q_j = \sum_{k \neq j} p_{ik}(t) q_{kj},$$

即 $P(t)$ 满足向前方程组.

为证明 $P(t)$ 的极小性, 设 \tilde{p}_{ij} 是向前方程组的一组非负解. 同样的, 有

$$\tilde{p}_{ij}(t) = \delta_{ij}(t) + \sum_{k \neq j} \int_0^t \tilde{p}_{ik}(t-s) q_{kj} e^{-q_j s} ds.$$

类似于 (4.18) 的证明, 对任意的 $n \geq 0$,

$$\begin{aligned}
 & P_i(X_j = j, t < \tau_{n+1}) \\
 &= \delta_{ij}e^{-q_it} + \sum_{k \neq j} \int_0^t P_i(X_t = j, t < \tau_n) q_{kj} e^{-q_j s} ds.
 \end{aligned} \tag{4.20}$$

由 $\tilde{P}(t) \geq 0$,

$$P(X_t = j, t < \tau_0) = 0 \leq \tilde{p}_{ij}(t).$$

应用归纳法可以证明对任意的 $n \geq 0$,

$$P_i(X_t = j, t < \tau_n) \leq \tilde{p}_{ij}(t).$$

于是

$$p_{ij}(t) = \lim_{n \rightarrow \infty} P_i(X_t = j, t < \tau_n) \leq \tilde{p}_{ij}(t).$$

□

下面例子说明当 Q 矩阵不保守时, 马氏链是非常复杂的. 这个例子是 P. Lévy 给出的.

例 4.4 设

$$D_n = \{k2^{-n}, k \in \mathbb{Z}_+\}, \quad n \geq 0,$$

$E = \bigcup_n D_n$. 对每个 $i \in D_n - D_{n-1}$, 设 T_i 是独立的指数随机变量, 参数为 $(2^n)^2$. 由于 $(D_n - D_{n-1}) \cap [0, 1)$ 共有 2^{n-1} 个元素, 对 $i \in E$, 则

$$E\left[\sum_{j \leq i} T_j\right] \leq (i+1) \sum_{n=0}^{\infty} 2^{n-1} (2^{-2n}) < \infty,$$

$$P\left(\lim_{i \rightarrow \infty} \sum_{j \leq i} T_j = \infty\right) = 1.$$

定义

$$X_t = \begin{cases} i, & \text{若对某个 } i, \sum_{j < i} T_j \leq t < \sum_{j \leq i} T_j, \\ \infty, & \text{其它情况.} \end{cases}$$

关于这个例子的详细说明, 请参考 D. Freedman 的 *Markov Chains*.

4.7 生灭过程

设对任意的 i , $\lambda_i + \mu_i > 0$, 定义 Q 矩阵

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & (-\lambda_2 + \mu_2) & \lambda_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}. \quad (4.21)$$

或者说, 当 $h \rightarrow 0$ 时,

$$\begin{aligned} p_{i,i+1}(h) &= \lambda_i h + o(h), \\ p_{i,i-1}(h) &= \mu_i h + o(h), \\ p_{i,i}(h) &= 1 - (\lambda_i + \mu_i)h + o(h), \\ p_{ij}(h) &= o(h), \quad |i - j| \leq 2, \\ p_{ij}(0) &= \delta_{ij}. \end{aligned}$$

我们称以上述 Q 为密度矩阵的连续时间马氏链 X_t 为生灭过程. 其中的 λ_i 被称为出生率, μ_i 为死亡率. 若对任意的 i , $\mu_i = 0$, 则称为纯生过程. 相应的, 若所有的 $\lambda_i = 0$, 称为纯灭过程. 如果 $\lambda_i = \lambda i$, $\mu_i = \mu i$, 则称为线性生灭过程.

生灭过程的命名是基于下面的原因: 在位于状态 i 时, 设 B_i, D_i 是独立的指数随机变量, 参数分别是 λ_i 和 μ_i , 我们将 B_i 理解为人口数为 i 时等待一个生命出生的时间, 而 D_i 为等待一个死亡的时间. 当出生早于死亡, 即 $B_i \leq D_i$ 时, 人口数将会从 i 增加到 $i + 1$, 反之则从 i 减少到 $i - 1$. 这样, 从 i 转移到 $i + 1$ 的概率为

$$P(B_i \leq D_i) = \frac{\lambda_i}{\lambda_i + \mu_i}.$$

容易看到, X_t 的嵌入链 Z_n 是一个生灭链, 见例 2.9, 其中

$$b_i = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad a_i = 1 - b_i, \quad r_i = 0.$$

从前面的讨论可知, 如果 Z_n 常返, 即

$$\sum_{m=0}^{\infty} \frac{a_1 a_2 \cdots a_m}{b_1 b_2 \cdots b_m} = \infty,$$

则 $\tau_{\infty} = \infty$, 以 Q 为密度矩阵的生灭过程是唯一的. 实际上, 我们并不需要这样强的条件. Feller 在 1940 年给出了 Q 过程唯一的充分必要条件, 我们简述如下.

假设 $\lambda_i > 0$, $\mu_i > 0$. 记

$$m_i = \frac{1}{\lambda_i} + \sum_{j=0}^{i-1} \frac{\lambda_i \lambda_{i-1} \cdots \lambda_{i-j}}{\mu_i \mu_{i-1} \cdots \mu_{i-j}}, \quad (4.22)$$

$$R = \sum_{i=0}^{\infty} m_i. \quad (4.23)$$

定理 4.24 Q 过程唯一的充分必要条件是 $R = \infty$.

证明 设 X_t 是一个可分右连续的 Q 过程. 令

$$\xi_i = \inf\{t \geq 0, X_t = i\}, \quad m_i = E_i[\xi_{i+1}].$$

由强马氏性,

$$m_i = \frac{\lambda_i}{\lambda_i + \mu_i} \frac{1}{\lambda_i + \mu_i} + \frac{\mu_i}{\lambda_i + \mu_i} \left(\frac{1}{\lambda_i + \mu_i} + m_{i-1} + m_i \right), \quad i > 0.$$

又 $m_0 = 1/\lambda_0$, 求解上面的递归方程即可得到我们前面定义的 m_i . 即我们证明了 $m_i = E_i[\xi_{i+1}]$. 由单调收敛定理,

$$E_0[\tau_\infty] = \lim_{i \rightarrow \infty} E_0[\xi_i].$$

而

$$\begin{aligned} E_0[\xi_i] &= \sum_{j=0}^{i-1} E_0[\xi_{j+1} - \xi_j] \\ &= \sum_{j=0}^{i-1} E_j[\xi_{j+1}] = \sum_{j=0}^{i-1} m_j. \end{aligned}$$

因此

$$E_0[\tau_\infty] = \sum_{j=0}^{\infty} m_j = R.$$

这样, 如果 $P(\tau_\infty = \infty) = 1$, 则 $P_0(\tau_\infty = \infty) = 1$, 从而 $R = E_0[\tau_\infty] = \infty$.

反之, 若 $P(\tau_\infty < \infty) > 0$, 则 $P_0(\tau_\infty < \infty) > 0$. 从而存在 $T > 0$ 及 $\alpha > 0$, 使得 $P_0(\tau_\infty \leq T) \geq \alpha$. 这样, 对任意的 k , $P_k(\tau > T) \leq P_0(\tau > T) \leq 1 - \alpha$. 于是,

$$\begin{aligned} E_0[\tau_\infty] &= \int_0^\infty P_0(\tau_\infty > s) ds = \sum_{m=0}^{\infty} \int_{mT}^{(m+1)T} P_0(\tau_\infty > s) ds \\ &\leq T \sum_{m=0}^{\infty} P_0(\tau_\infty > mT). \end{aligned}$$

由马氏性可知

$$\begin{aligned} P_0(\tau_\infty > mT) &= E_0[1_{\{\tau_\infty > (m-1)T\}} P_{X_{(m-1)T}}(\tau_\infty > T)] \\ &\leq (1 - \alpha) P_0(\tau_\infty > (m-1)T), \end{aligned}$$

故 $P_0(\tau_\infty > mT) \leq (1 - \alpha)^m$. 这就证明了 $R = E_0[\tau_\infty] < \infty$. \square

4.8 常返性与不变测度

设 X_t 是以 Q 为密度矩阵的马氏链, τ_∞ 是 X_t 的第一个跳跃点. 在这一节里, 我们假设 $0 \leq q_i < \infty$, Q 矩阵保守, 且当 $t \geq \tau_\infty$ 时, $X_t = \infty$. 设 Z_n 是 X_t 的嵌入链, (u_{ij}) 是它的转移矩阵, 则对 $i \neq j$,

$$u_{ij} = \begin{cases} \frac{q_{ij}}{q_i}, & \text{若 } q_i \neq 0, \\ 0, & \text{若 } q_i = 0, \end{cases}$$

对 $i = j$,

$$u_{ii} = \begin{cases} 0, & \text{若 } q_i \neq 0, \\ 1, & \text{若 } q_i = 0. \end{cases}$$

与离散参数情形一样, 我们称状态 i 可达 j , 若

$$P_i(\exists t \geq 0, X_t = j) > 0,$$

记为 $i \rightarrow j$. 如果 $i \rightarrow j$ 且 $j \rightarrow i$, 则称 i 与 j 互通, 记为 $i \leftrightarrow j$. 同离散参数情形一样, 我们可以定互通等价类, 吸收态, 不可约性等.

定理 4.25 对状态 $i \neq j$, 下列条件等价:

(i) $i \rightarrow j$;

(ii) 对嵌入链 Z_n , $i \rightarrow j$;

(iii) 存在状态 $i = i_0, i_1, \dots, i_n = j$ 使得 $q_{i_0 i_1} \cdots q_{i_{n-1} i_n} > 0$;

(iv) 对任意的 $t > 0$, 有 $p_{ij}(t) > 0$;

(v) 存在 $t > 0$, $p_{ij}(t) > 0$.

证明 (iv) \Rightarrow (v) \Rightarrow (i) \Rightarrow (ii) 是明显的. 设 (ii) 成立, 则存在状态 $i = i_0, i_1, \dots, i_n = j$, 使得 $u_{i_0 i_1} \cdots u_{i_{n-1} i_n} > 0$, 于是 (iii) 成立.

若 $q_{ij} > 0$, 则对任意 $t > 0$,

$$\begin{aligned} p_{ij}(t) &\geq P_i(\tau_1 \leq t, Z_1 = j, \tau_2 - \tau_1 > t) \\ &= (1 - e^{-q_i t}) u_{ij} e^{-q_j t} > 0. \end{aligned}$$

这样, 若 (iii) 成立, 则

$$p_{ij}(t) \geq p_{i_0 i_1}(t/n) \cdots p_{i_{n-1} i_n}(t/n) > 0,$$

即 (iv) 成立. □

注 4.2 这里我们证明了定理 4.4 中的 (iii).

定义 4.9

(i) 称状态 i 是常返的, 若

$$P_i(\{t \geq 0, X_t = i\} \text{ 无界}) = 1;$$

(ii) 称状态 i 是暂留(非常返)的, 若

$$P_i(\{t \geq 0, X_t = i\} \text{ 无界}) = 0.$$

容易看到, 如果 $P_i(\tau_\infty < \infty) > 0$, 则 i 不是常返的. 下面定理说明连续时间马氏链的常返性可以由它的嵌入链来刻画.

定理 4.26 同前面一样, 设 Z_n 是马氏链 X_t 的嵌入链. 则

- (i) 若在 Z_n 中, 状态 i 常返, 则对于 X_t , i 也常返;
- (ii) 若在 Z_n 中, 状态 i 非常返, 则对于 X_t , i 也非常返;
- (iii) 任意状态或者常返, 或者非常返;
- (iv) 若状态 i 与 j 互通, 则它们具有相同的常返性.

证明

- (i) 设对于 Z_n , i 是常返的. 由定理 4.19 可知 $P_i(\tau_\infty = \infty) = 1$. 这样, 对无穷多个 n , $X_{\tau_n} = Z_n = i$, 从而 $\{t \geq 0, X_t = i\}$ 以概率 1 无界.
- (ii) 设对于 Z_n , i 是非常返的. 则在概率 P_i 下,

$$N = \sup\{n \geq 0, Z_n = i\} < \infty \text{ a.s..}$$

这样, $\{t \geq 0, X_t = i\}$ 具有上界 τ_{N+1} , 它以概率 1 有限. □

令

$$T_i = \inf\{t \geq \tau_1, X_t = i\}.$$

下面结果给出了与离散情形类似的常返性的判据.

定理 4.27

(i) 若 $q_i = 0$ 或 $P_i(T_i < \infty) = 1$, 则 i 常返且 $\int_0^\infty p_{ii}(t)dt = \infty$;

(ii) 若 $q_i > 0$, 且 $P_i(T_i < \infty) < 1$, 则 i 非常返, $\int_0^\infty p_{ii}(t)dt < \infty$.

证明 若 $q_i = 0$, 则在概率 P_i 下, X_t 一直停留在 i , 于是 i 常返, 且对任意的 $t \geq 0$, $p_{ii}(t) = 1$, 故 $\int_0^\infty p_{ii}(t)dt = \infty$.

设 $q_i > 0$. 记 N_i 是嵌入链 Z_n 首次到达状态 i 的时间, 即

$$N_i = \inf\{n \geq 1, Z_n = i\}.$$

则

$$P_i(N_i < \infty) = P_i(T_i < \infty) = 1.$$

故 i 在 Z_n 中常返. 这就证明了 i 是常返态.

记 $u_{ij}^{(n)}$ 为 Z_n 的 n 步转移概率. 则由 Fubini 定理可知

$$\begin{aligned} \int_0^\infty p_{ii}(t)dt &= \int_0^\infty E_i(1_{\{X_t=i\}})dt = E_i\left[\int_0^\infty 1_{\{X_t=i\}}dt\right] \\ &= E_i\left(\sum_{n=0}^\infty (\tau_{n+1} - \tau_n)1_{\{Z_n=i\}}\right) = \sum_{n=0}^\infty E_i(\tau_{n+1} - \tau_n | Z_n = i)P_i(Z_n = i) \\ &= \frac{1}{q_i} \sum_{n=0}^\infty u_{ii}^{(n)}. \end{aligned}$$

这样, i 常返当且仅当 $\int_0^\infty p_{ii}(t)dt = \infty$. □

由 $p_{ii}(t)$ 的连续性,

$$a_i(s) = \inf_{0 \leq v \leq s} p_{ii}(v) > 0.$$

应用 C-K 方程, 对 $0 \leq v \leq s$,

$$p_{ij}(t+v) \geq p_{ii}(v)p_{ij}(t) \geq a_i(s)p_{ij}(t).$$

取 $ns \leq t \leq (n+1)s$, 则

$$\frac{1}{a_i(s)}p_{ij}((n+1)s) \geq p_{ij}(t) \geq a_i(s)p_{ij}(ns),$$

故

$$\frac{s}{a_i(s)} \sum_{n=1}^{\infty} p_{ij}(ns) \geq \int_0^{\infty} p_{ij}(t) dt \geq a_i(s) s \sum_{n=0}^{\infty} p_{ij}(ns),$$

从而

$$\int_0^{\infty} p_{ij}(t) dt = \infty$$

当且仅当

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

注意到 $(p_{ij}(ns))$ 是离散参数马氏链 $(X_{ns})_{n \geq 0}$ 的 n 步转移矩阵. 我们称 (X_{ns}) 为马氏链 X_t 的 s -骨架.

命题 4.3 状态 i 常返当且仅当它相对于某个 s -骨架过程是常返的.

类似于离散参数马氏链, 我们称状态 i 正常返, 若 $q_i = 0$ 或者 $m_i = E_i[T_i] < \infty$. 对于常返状态, 若不是正常返的, 则称为零常返.

与常返性不同, 马氏链的正常返性不能通过嵌入链来刻画. 我们看下面的两个例子.

例 4.5 设 $p_i > 0, i \geq 1, \sum_i p_i = 1$. 定义嵌入链 Z_n 的转移矩阵为

$$u_{0i} = p_i, u_{i0} = 1, i \geq 1, u_{ij} = 0, i, j \neq 0.$$

则嵌入链 Z_n 是正常返的. 令 $q_i = p_i/2, q_{ij} = q_i u_{ij}, i \neq j$, 则

$$\begin{aligned} E_0[T_0] &= E_0[\tau_n 1_{\{N_0=n\}}] = E_0[\tau_2] = \sum_{i=1}^{\infty} E_0[\tau_2 | Z_1 = i] P_0(Z_1 = i) \\ &\geq \sum_{i=1}^{\infty} \frac{1}{q_i} p_i = \infty. \end{aligned}$$

故 Q 过程不是正常返的.

例 4.6 考虑保守的生灭过程

$$\lambda_i = \mu_i, i \geq 1, \mu_i > 0, i \geq 0.$$

选取 $\mu_i \rightarrow \infty$, 使得 $\sum_i 1/\mu_i < \infty$. 则 Q 过程是正常返的, 但嵌入链不是.

设 π 是 E 上一个 σ -有限的测度, 我们把 π 看作是一个行向量, 这样 π 与 $P(t)$ 可以做乘法, 即

$$(\pi P(t))_j = \sum_{i \in E} \pi_i p_{ij}(t).$$

它是一个新的测度. 当 π 是 X_0 的分布时, $\pi P(t)$ 是 X_t 的分布.

定义 4.10 设 π 是 E 上一个非平凡测度. 称 π 为马氏链 $X(t)$ 的不变测度, 若对任意的 $t > 0$,

$$\pi = \pi P(t).$$

若 π 是 X_t 的不变测度, 由向前方程组的积分形式

$$p_{ij}(t) = \delta_{ij}e^{-q_j t} + \sum_{k \neq j} \int_0^t p_{ik}(t-s)q_{kj}e^{-q_j s} ds$$

可知,

$$\begin{aligned} \pi_j &= \sum_i \pi_i p_{ij}(t) \\ &= \pi_j e^{-q_j t} + \int_0^t \sum_{k \neq j} \sum_i \pi_i p_{ik}(t-s)q_{kj}e^{-q_j s} ds \\ &= \pi_j e^{-q_j t} + \int_0^t \sum_{k \neq j} \pi_k q_{kj}e^{-q_j s} ds \\ &= \pi_j e^{-q_j t} + \frac{1}{q_j} \sum_{k \neq j} \pi_k q_{kj}(1 - e^{-q_j t}), \end{aligned}$$

即

$$\sum_k \pi_k q_{kj}(1 - e^{-q_j t}) = 0,$$

由 t 的任意性, $\pi Q = 0$. 注意到 $q_i(u_{ij} - \delta_{ij}) = q_{ij}$, 令 $\nu_i = q_i \pi_i$, 则

$$(\nu(U - I))_j = \sum_i \nu_i(u_{ij} - \delta_{ij}) = \sum_i \pi q_{ij} = (\pi Q)_j.$$

这样,

命题 4.4 设 π 是 X_t 的不变测度, 则

(i) $\pi Q = 0$;

(ii) 设 $\nu_i = q_i \pi_i$, 则 ν 是嵌入链 Z_n 的不变测度, 即 $\nu U = \nu$.

容易看到, 上面定理中 (i) 和 (ii) 是等价的.

下面我们考虑马氏链不变测度的存在性与唯一性, 这里的唯一性是指差一个常数倍的意义下.

对于离散参数马氏链, 我们不加证明的给出下面定理.

定理 4.28 设 X_n 是离散参数马氏链.

(i) 设状态 i 常返, 记

$$N_i = \inf\{n \geq 1, X_n = i\},$$

$$e_{ij} = E_i\left[\sum_{n=1}^{N_i} 1_{\{X_n=j\}}\right],$$

则 $\{e_{ij}, j \in E\}$ 是 X_n 的不变测度.

(ii) 设 X_n 不可约常返, 则它有唯一的不变测度.

下面我们回到连续时间马氏链. 设状态 i 常返, $0 < q_i < \infty$, 则 $P_i(T_i < \infty) = 1$. 令

$$\mu_j = E_i \left[\int_0^{T_i} 1_{\{X_t=j\}} dt \right].$$

在停时 T_i 应用强马氏性可知

$$E_i \left[\int_0^s 1_{\{X_t=j\}} dt \right] = E_i \left[\int_{T_i}^{T_i+s} 1_{\{X_t=j\}} dt \right],$$

这样,

$$\begin{aligned} \mu_j &= E_i \left[\int_0^{T_i+s} 1_{\{X_t=j\}} dt \right] = \int_0^\infty P_i(X_{s+t} = j, t < T_i) dt \\ &= \int_0^\infty \sum_{k \in E} P_i(X_t = k, t < T_i) p_{kj}(s) dt \\ &= \sum_{k \in E} E_i \left[\int_0^{T_i} 1_{\{X_t=k\}} dt \right] p_{kj}(s) \\ &= \sum_{k \in E} \mu_k p_{kj}(s), \end{aligned}$$

即 μ 是 X_t 的不变测度.

反过来, 设 π 是不可约马氏链 X_t 的不变测度, 即对任意的 t , $\pi P(t) = \pi$. 固定 $t > 0$, 考虑离散参数马氏链 $Y_n = X_{nt}$, 则 π 也是 Y_n 的不变测度. 由于 X_t 不可约, 故对任意状态 i, j , $p_{ij}(t) > 0$, 从而 Y_n 不可约. 又当 X_t 常返时, Y_n 也是常返的. 此时 Y_n 的不变测度唯一. 因此 X_t 的不变测度也是唯一的.

定理 4.29 设 X_t 不可约, 常返, 则它唯一的不变测度由

$$\mu_j = E_i \left[\int_0^{T_i} 1_{\{X_t=j\}} dt \right]$$

确定, 其中 i 是任意给定的一个状态. 此时, X_t 的嵌入链 Z_n 也不可约, 常返. 令 $\nu_j = q_j \mu_j$, 则 ν 是 Z_n 的唯一的不变测度.

实际上, 应用强马氏性可以得到

$$\begin{aligned} \mu_j &= E_i \left[\int_0^{T_i} 1_{\{X_t=j\}} dt \right] = \sum_{n=1}^{\infty} E_i [(\tau_{n+1} - \tau_n) 1_{\{Z_n=j, n < N_i\}}] \\ &= \sum_{n=1}^{\infty} E_i [\tau_{n+1} - \tau_n | Z_n = j] P_i(Z_n = j, n < N_i) \\ &= \frac{1}{q_j} E_i \left[\sum_{n=1}^{N_i} 1_{\{Z_n=j\}} \right] = \frac{e_{ij}}{q_j}. \end{aligned}$$

对于非常返不可约马氏链, 我们知道它不一定存在不变测度, 即使存在, 也不一定是唯一的. 我们可以把不变测度的限制放宽一些.

定义 4.11 称测度 π 为 X_t 的过分测度, 若对任意的状态 j 与 $t > 0$, 有

$$\pi_j \geq \sum_i \pi_i p_{ij}(t).$$

定理 4.30 设马氏链 X_t 不可约非常返, $\alpha = (\alpha_i)$ 是 E 上有限支撑 (即仅在有限个状态上有负荷) 的测度, 令

$$\mu_j = \sum_i \int_0^t \alpha_i p_{ij}(s) ds.$$

则 μ 是 X_t 的过分测度.

证明 由 X_t 非常返, 则对任意的 i, j ,

$$\int_0^{\infty} p_{ij}(s) ds < \infty.$$

应用控制收敛定理,

$$\begin{aligned}
\sum_j \mu_j p_{jk}(t) &= \sum_j \sum_i \int_0^\infty \alpha_i p_{ij}(s) p_{jk}(t) ds \\
&= \sum_i \int_0^\infty \alpha_i \sum_j p_{ij}(s) p_{jk}(t) ds = \sum_i \int_0^\infty \alpha_i p_{ik}(t+s) ds \\
&\leq \mu_k,
\end{aligned}$$

故 μ 是过分测度. □

一个自然的问题是过分测度什么时候是不变测度. 设 μ 是过分测度, 考虑对偶链, 即以 $\bar{p}_{ij}(t) = \mu_j p_{ji}/\mu_i$ 为转移函数的马氏链. 显然的, μ 是不变测度等价于 $\bar{P}(t)1 = 1$, 即对于对偶链, $P(\bar{\tau}_\infty = \infty) = 1$.

5 更新理论

这一章主要研究与更新有关的内容.

考虑下面几个例子:

例 5.1 设 N_t 是参数为 λ 的齐次 Poisson 过程, 跟以前一样, 令

$$S_n = \inf\{t, N_t \geq n\}, \quad T_n = S_n - S_{n-1}.$$

则 T_1, T_2, \dots 为独立同分布随机变量序列, 且服从参数为 λ 的指数分布,

$$S_n = T_1 + T_2 + \dots + T_n.$$

例 5.2 设 X_n 为离散参数马尔可夫链, 状态 i 常返, 令

$$S_0 = \inf\{k, X_k = i\}, \quad S_{n+1} = \inf\{k > S_n, X_k = i\}.$$

则 $S_1 - S_0, S_2 - S_1, \dots$ 独立同分布. 如果马氏链从状态 i 出发, 则 $S_0 = 0$.

在上面例子中, S_n 都可以表示成独立随机变量的和, 并且除去第一项外, 其它因子都是同分布的. 这种序列我们称为更新序列. 进一步的, 通过这些随机时间, 我们可以把过程分成一些随机时间段, 它们是独立同分布的. 这种现象在随机过程里是相当普遍的.

5.1 Laplace 变换

设 $U(x)$ 是 $[0, \infty)$ 上单增函数, $g(x)$ 是 $[0, \infty)$ 上局部有界的函数, 我们可以定义 $g(x)$ 关于 $U(x)$ 的 Lebesgue--Stieltjes 积分, 记为

$$\int_0^\infty g(x) dU(x) = \int_{[0, \infty)} g(x) U(dx).$$

如果不加说明, 我们这一章中考虑的函数和分布, 都是定义在 $\mathbb{R}_+ = [0, \infty)$ 上的. 对局部有界函数 g 和分布 F , 我们可以定义它们的卷积为

$$(F * g)(t) = \int_0^t g(t-s) F(ds), \quad t \geq 0.$$

我们不加证明的给出下面的性质:

(i) $F * g$ 局部有界:

$$\sup_{0 \leq s \leq t} |F * g(s)| \leq \left(\sup_{0 \leq s \leq t} |g(s)| \right) F(t).$$

(ii) 若 g 有界连续, 则 $F * g$ 连续.

(iii) 记 $F^{2*} = F * F$, 则 $F * (F * g) = F^{2*} * g$. 高维情形类似.

(iv) 设 X_1, X_2 独立, 且分布函数分别是 F_1 和 F_2 , 则 $X_1 + X_2$ 的分布函数为 $F_1 * F_2$.

(v) $F_1 * F_2 = F_2 * F_1$.

(vi) 设 F, G 为分布函数. 若 F 绝对连续, 密度为 f . 则 $F * G$ 也绝对连续, 且密度为 $G * f$.

设 X 是非负随机变量, 具有分布函数 F . X (或 F) 的 Laplace 变换定义为

$$\hat{F}(\lambda) = E[e^{-\lambda X}] = \int_0^{\infty} e^{-\lambda x} F(dx), \quad \lambda \geq 0.$$

注意到 $e^{-\lambda x} \leq 1, x \geq 0$, 因此上面积分是有限的. Laplace 变换具有下面的性质:

(i) 若 F 和 G 具有相同的 Laplace 变换, 则 $F = G$.

(ii) 设 X_1 和 X_2 独立, 且分别具有分布 F_1 和 F_2 , 则

$$(\widehat{F_1 * F_2})(\lambda) = \hat{F}_1(\lambda) \hat{F}_2(\lambda).$$

(iii) $\hat{F}(\lambda)$ 在 $(0, \infty)$ 无穷次可微, 且

$$(-1)^n \frac{d^n}{d\lambda^n} \hat{F}(\lambda) = \int_0^{\infty} e^{-\lambda x} x^n F(dx), \quad \lambda > 0.$$

这样,

$$\lim_{\lambda \downarrow 0} (-1)^n \frac{d^n}{d\lambda^n} \hat{F}(\lambda) = \int_0^{\infty} x^n F(dx).$$

特别的, $EX = -\hat{F}'(0)$, $E(X^2) = \hat{F}''(0)$.

(iv)

$$\int_0^{\infty} e^{-\lambda x} F(x) dx = \lambda^{-1} \hat{F}(\lambda), \quad (5.1)$$

$$\int_0^{\infty} e^{-\lambda x} (1 - F(x)) dx = (1 - \hat{F}(\lambda)) / \lambda. \quad (5.2)$$

例 5.3 下表列出了几种常见分布的 Laplace 变换.

分布	Laplace 变换
(0, 1) 上均匀分布	$(1 - e^{-\lambda})/\lambda$
参数为 α 的指数分布	$\alpha/(\alpha + \lambda)$
参数为 α, n 的 Γ 分布	$(\alpha/(\alpha + \lambda))^n$

对于 $[0, \infty)$ 上单增函数 $U(x)$, 如果存在 $a \geq 0$, 使得对任意 $\lambda > a$,

$$\int_0^\infty e^{-\lambda x} U(dx) < \infty,$$

我们可以定义

$$\hat{U}(\lambda) = \int_0^\infty e^{-\lambda x} U(dx), \quad \lambda > a,$$

称为 U 的 Laplace 变换. 容易看出, 分布函数的 Laplace 变换的性质仍旧成立.

例 5.4 设 $U(dx) = e^{ax} dx$, 则

$$\int_0^\infty e^{-\lambda x} U(dx) = \begin{cases} (\lambda - a)^{-1}, & \text{若 } \lambda > a; \\ \infty, & \text{若 } \lambda \leq a. \end{cases}$$

例 5.5 设 $U(dx) = (1 - F(x))dx$, 则

$$\hat{U}(\lambda) = (1 - \hat{F}(\lambda))/\lambda.$$

5.2 基本概念

设 $\{Y_n, n \geq 1\}$ 是独立的非负随机变量序列, $\{Y_n, n \geq 2\}$ 同分布, Y_1 的分布函数为 $G(x)$, $Y_n, n \geq 2$ 的分布函数为 $F(x)$, 设 $F(0) < 1$. 令

$$S_n = Y_1 + Y_2 + \cdots + Y_n, \quad n \geq 1.$$

称 $\{S_n, n \geq 1\}$ 为更新序列. 若分布函数 G 与 F 相同, 则称 S_n 为普通更新序列 (*ordinary renewal sequence*); 若 $G \neq F$, 则称 S_n 为延迟更新序列 (*delay renewal sequence*).

记

$$N(t) = \sum_{n=1}^\infty 1_{[0, t]}(S_n) = \sup\{n : S_n \leq t\}. \tag{5.3}$$

我们称 $N(t)$ 为更新过程. $E[N(t)]$ 在更新理论里是一个非常基本的量, 我们称它为更新函数.

若 $G = F$, 即在普通更新情形, 更新函数

$$\begin{aligned} U(t) &= E\left[\sum_{n=1}^{\infty} 1_{[0, t]}(S_n)\right] = \sum_{n=1}^{\infty} P(S_n \leq t) \\ &= \sum_{n=1}^{\infty} F^{*n}(t). \end{aligned} \quad (5.4)$$

而对延迟更新情形, 更新函数

$$\begin{aligned} V(t) &= \sum_{n=1}^{\infty} P(S_n \leq t) = \sum_{n=1}^{\infty} (G * F^{(n-1)*})(t) \\ &= (G * (U + 1_{[0, \infty)}))(t). \end{aligned} \quad (5.5)$$

注意到 S_n 与 $N(t)$ 是相互决定的, 特别的,

$$\begin{aligned} S_{N(t)} &\leq t < S_{N(t)+1}, \\ \{N(t) < n\} &= \{S_n > t\}, \\ \{N(t) = n\} &= \{S_n \leq t < S_{n+1}\}, \quad n \geq 1. \end{aligned}$$

由 $F(0) < 1$ 知 $\mu = E(Y_2) > 0$. 应用强大数定律,

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = E(Y_2),$$

故 $\lim_{n \rightarrow \infty} S_n = \infty$. 这样

$$\begin{aligned} P(N(t) < \infty) &= \lim_{n \rightarrow \infty} P(N(t) < n) \\ &= \lim_{n \rightarrow \infty} P(S_n > t) = 1. \end{aligned}$$

定理 5.1

(i) 设 $\gamma < 1/F(0)$, 则

$$\sum_{n=0}^{\infty} \gamma^n F^{n*}(t) < \infty, \quad t \geq 0. \quad (5.6)$$

(ii) $N(t)$ 各阶矩有限. 特别的, $U(t) < \infty$.

证明

(i) 固定 $\gamma < 1/F(0)$, 由

$$\begin{aligned}\lim_{\lambda \rightarrow \infty} \hat{F}(\lambda) &= \lim_{\lambda \rightarrow \infty} \left(F(0) + \int_{(0, \infty)} e^{-\lambda x} F(dx) \right) \\ &= F(0),\end{aligned}$$

存在 $\lambda > 0$, 使 $\hat{F}(\lambda)\gamma < 1$. 这样, 应用 Markov 不等式,

$$\begin{aligned}\sum_{n=0}^{\infty} \gamma^n F^{n*}(t) &= \sum_{n=0}^{\infty} \gamma^n P(S_n \leq t) = \sum_{n=0}^{\infty} \gamma^n P(e^{-\lambda S_n} \geq e^{-\lambda t}) \\ &\leq \sum_{n=0}^{\infty} \gamma^n E[e^{-\lambda S_n} e^{\lambda t}] = e^{\lambda t} \sum_{n=0}^{\infty} (\gamma E[e^{-\lambda Y_1}])^n < \infty.\end{aligned}$$

(ii) 取 $1 < \gamma < 1/F(0)$, 由 $\lim_{n \rightarrow \infty} \gamma^n F^{n*}(t) = 0$, 存在 n_0 , 使得当 $n > n_0$ 时,

$$F^{n*}(t) \leq \gamma^{-n} = e^{-(\log \gamma)n}.$$

这样, 对 $n > n_0$,

$$P(N(t) > n) = P(S_n \leq t) = F^{n*}(t) \leq e^{-(\log \gamma)n}.$$

于是存在常数 $K > 0, c > 0$ 使对任意 n ,

$$P(N(t) > n) \leq Ke^{-cn}.$$

注意到对 $\theta < c$,

$$\begin{aligned}E(e^{\theta N(t)}) &= 1 + E \left[\int_0^{N(t)} \theta e^{\theta u} du \right] \\ &= 1 + E \left[\int_0^{\infty} \theta e^{\theta u} 1_{\{u \leq N(t)\}} du \right] = 1 + \int_0^{\infty} \theta e^{\theta u} P(N(t) > u) du \\ &\leq 1 + \theta K \int_0^{\infty} e^{\theta u} e^{-cu} du < \infty.\end{aligned}$$

因此 $N(t)$ 各阶矩有限.

□

例 5.6 (Poisson 过程) 设

$$F(dx) = \alpha e^{-\alpha x} dx, \quad \alpha > 0, \quad x \geq 0.$$

则

$$F^{n*}(dx) = \alpha(\alpha x)^{n-1} \frac{e^{-\alpha x}}{(n-1)!} dx, \quad x \geq 0,$$

于是,

$$\begin{aligned} U(x) \sum_{n=1}^{\infty} F^{n*}(x) &= \sum_{n=1}^{\infty} \int_0^x \alpha(\alpha s)^{n-1} \frac{e^{-\alpha s}}{(n-1)!} ds \\ &= \int_0^x \sum_{n=1}^{\infty} \alpha(\alpha s)^{n-1} \frac{e^{-\alpha s}}{(n-1)!} ds = \alpha x. \end{aligned}$$

所以 $U(x) = \alpha x$.

定理 5.2 设 $\mu = E(Y_2) < \infty$.

(i) 若 $P(Y_1 < \infty) = 1$, 则

$$\lim_{t \rightarrow \infty} \frac{N(t)}{t} = \frac{1}{\mu}, \quad a.s.. \quad (5.7)$$

(ii) 若 $\sigma^2 = \text{Var}(Y_2) < \infty$, 则

$$\lim_{t \rightarrow \infty} P\left(\frac{N(t) - t/\mu}{\sqrt{t\sigma^2\mu^{-3}}} \leq x\right) = N(0,1;x), \quad (5.8)$$

其中 $N(0,1;x)$ 是标准正态分布的分布函数.

证明

(i) 由强大数定律, 当 $n \rightarrow \infty$ 时,

$$\frac{S_n}{n} = \frac{Y_1}{n} + \frac{1}{n} \sum_{i=2}^n Y_i \rightarrow \mu.$$

注意到 $\lim_{t \rightarrow \infty} N(t) = \infty$ 及

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{N(t)+1}}{N(t)},$$

而

$$\lim_{t \rightarrow \infty} \frac{S_{N(t)}}{N(t)} = \lim_{t \rightarrow \infty} \frac{S_{N(t)+1}}{N(t)} = \mu,$$

故当 $t \rightarrow \infty$ 时, $N(t)/t \rightarrow \mu^{-1}$.

(ii) 令

$$h(t) = x\sqrt{\sigma^2 t \mu^{-3}} + t/\mu, \quad z(t) = \frac{t - \mu h(t)}{\sigma h^{1/2}(t)}.$$

注意到当 $t \rightarrow \infty$ 时, $h(t) \rightarrow \infty$, $z(t) \rightarrow -x$. 应用中心极限定理, 当 $t \rightarrow \infty$ 时,

$$\begin{aligned} P\left(\frac{N(t) - t/\mu}{\sqrt{\sigma^2 t \mu^{-3}}} \leq x\right) &= P(N(t) \leq x\sqrt{\sigma^2 t \mu^{-3}} + t/\mu) \\ &= P(N(t) < [h(t)] + 1) = P(S_{[h(t)]+1} > t) \\ &= P\left(\frac{S_{[h(t)]+1} - \mu([h(t)] + 1)}{\sigma\sqrt{[h(t)] + 1}} > \frac{t - \mu([h(t)] + 1)}{\sigma\sqrt{[h(t)] + 1}}\right) \\ &\rightarrow 1 - N(0, 1; -x) = N(0, 1; x). \end{aligned}$$

□

事实上, 在 $\mu = \infty$ 时, 定理 5.2 中 (i) 的结论仍然成立.

定理 5.3 等待时间分布函数 $F(x)$ 和更新函数 $U(t)$ 相互唯一确定.

证明 只需证明 $U(t)$ 唯一确定 $F(x)$. 由于 $\widehat{F^{n*}} = (\hat{F})^n$, $U(t)$ 的 Laplace 变换为

$$\hat{U}(\lambda) = \sum_{n=1}^{\infty} \widehat{F^{n*}}(\lambda) = \frac{\hat{F}(\lambda)}{1 - \hat{F}(\lambda)}.$$

从而

$$\hat{F}(\lambda) = \frac{\hat{U}(\lambda)}{1 + \hat{U}(\lambda)}$$

由 U 唯一确定, 从而 F 也由 U 确定. □

推论 5.1 若 $U(t) = \lambda t$, 则 $N(t)$ 是参数为 λ 的 Poisson 过程.

5.3 更新方程

在本节中, 如果不加声明, 我们设 S_n 为普通更新序列.

定理 5.4 更新函数 $U(t)$ 满足方程 $U = F + U * F$, 即

$$U(t) = F(t) + \int_0^t U(t-y)dF(y).$$

证明 [证明一]

$$\begin{aligned} U(t) &= \sum_{n=1}^{\infty} F^{*n}(t) = F(t) + \sum_{n=2}^{\infty} (F^{*(n-1)} * F)(t) \\ &= F(t) + (U * F)(t). \end{aligned}$$

□

证明 [证明二]注意到

$$N(t) = \sum_{n=1}^{\infty} 1_{[0, t]}(S_n) = \sum_{n=1}^{\infty} 1_{[0, t]}(Y_1 + S'_{n-1}),$$

其中 S'_n 与 S_n 同分布且与 Y_1 独立. 这样,

$$\begin{aligned} E[N(t)|Y_1 = s] &= \sum_{n=1}^{\infty} P(s + S'_{n-1} \leq t) \\ &= \begin{cases} 0, & \text{若 } s > t; \\ 1 + U(t-s), & \text{若 } s \leq t. \end{cases} \end{aligned}$$

于是

$$\begin{aligned} U(t) &= E[N(t)] = \int_0^{\infty} E[N(t)|Y_1 = s]dF(s) \\ &= \int_0^t [1 + U(t-s)]dF(s) = F(t) + \int_0^t U(t-s)dF(s). \end{aligned}$$

□

一般的, 我们有

定义 5.1 设 $z(t)$ 是定义在 $[0, \infty)$ 上的函数, 称关于未知函数 $Z(t)$, $t \in [0, \infty)$ 的方程

$$Z(t) = z(t) + \int_0^t Z(t-s)dF(s) \quad (5.9)$$

为更新方程, 简记为

$$Z = z + F * Z,$$

其中 F 为 $[0, \infty)$ 上广义分布函数, $F(\infty) < \infty$. 若 $F(\infty) = 1$, 则方程称为正规的.

定理 5.5 设 $F(0) < 1$, U 是 F 对应的更新函数. 若 $z(t)$ 局部有界, 则

(i) $z(t) + (U * z)(t) = z(t) + \int_0^t z(t-s) dU(s)$ 为更新方程 (5.9) 的一个局部有界的解.

(ii) 方程 (5.9) 满足 $Z(t) = 0, t < 0$ 的局部有界解唯一.

证明

(i) 对任意 $T > 0$,

$$\sup_{0 \leq t \leq T} |U * z(t)| \leq \int_0^t \sup_{0 \leq t \leq T} |z(t-s)| dU(s) \leq \sup_{0 \leq t \leq T} |z(t)| U(T),$$

故 $z(t) + (U * z)(t)$ 局部有界. 又

$$\begin{aligned} z + F * (z + U * z) &= z + F * (z + \sum_{n=1}^{\infty} F^{n*} * z) \\ &= z + F * z + \sum_{n=2}^{\infty} F^{n*} * z \\ &= z + \sum_{n=1}^{\infty} F^{n*} * z = z + U * z, \end{aligned}$$

于是 $z + U * z$ 是更新方程 (5.9) 的解.

(ii) 下面证明解的唯一性. 设 Z_1, Z_2 为 (5.9) 的两个局部有界解, 令

$$H = Z_1 - Z_2 = F * (Z_1 - Z_2) = F * H = F^{n*} * H.$$

则对任意的 $T > 0$,

$$\begin{aligned} \sup_{0 \leq t \leq T} |H(t)| &= \sup_{0 \leq t \leq T} \left| \int_0^t H(t-s) dF^{n*}(s) \right| \\ &\leq \int_0^T \sup_{0 \leq t \leq T} |H(t)| dF^{n*}(s) = \sup_{0 \leq t \leq T} |H(t)| F^{n*}(T). \end{aligned}$$

令 $m = F(\infty)$, 由

$$m(m^{-1}F)(0) = F(0) < 1,$$

$$U(T) = \sum_{n=1}^{\infty} F^{n*}(T) = \sum_{n=1}^{\infty} m^n (m^{-1}F)^{n*}(T) < \infty.$$

于是

$$\lim_{n \rightarrow \infty} F^{n*}(T) = 0,$$

从而对任意的 $t \geq 0$, $H(t) = 0$.

□

由于 $U = F + U * F$, 故 U 是更新方程 $Z = F + Z * F$ 的唯一局部有界解.

定理 5.6 设 $F(t)$ 绝对连续, 密度函数 $f(t)$ 局部有界, 则

$$U(t) = \int_0^t u(s)ds, \quad t \geq 0,$$

其中

$$u(t) = f(t) + \int_0^t f(t-s)dU(s), \quad t \geq 0.$$

证明 由定理 5.5, $u(t)$ 是更新方程

$$u(t) = f(t) + \int_0^t u(t-s)dF(s)$$

的局部有界解, 所以

$$\begin{aligned} \int_0^t u(s)ds &= \int_0^t f(s)ds + \int_0^t ds \int_0^s u(s-r)dF(r) \\ &= F(t) + \int_0^t \left(\int_r^t u(s-r)ds \right) dF(r) \\ &= F(t) + \int_0^t \left(\int_0^{t-r} u(s)ds \right) dF(r). \end{aligned}$$

因此 $\int_0^t u(s)ds$ 是更新方程 $B = F + B * F$ 的局部有界解. 由唯一性知 $U(t) = \int_0^t u(s)ds$.

□

定义 5.2 设 S_n 为更新序列. 我们分别称

$$\begin{aligned} A(t) &= t - S_{N(t)}, \\ R(t) &= S_{N(t)+1} - t, \\ \beta(t) &= A(t) + R(t) = S_{N(t)+1} - S_{N(t)} \end{aligned}$$

为时刻 t 的年龄, 剩余寿命和总寿命.

容易看出, $A(t)$ 表示上一次更新到当前的时间, $R(t)$ 表示到下一次更新所需的时间. 下面我们研究这三个量所满足的更新方程.

命题 5.1 对 $x > 0$,

$$P(R(t) > x) = 1 - F(t + x) + \int_0^t P(R(t-s) > x) F(ds). \quad (5.10)$$

证明 利用全概率公式

$$P(R(t) > x) = P(R(t) > x, Y_1 > t) + P(R(t) > x, Y_1 \leq t).$$

分别考虑上式右端的两项. 容易看出

$$P(R(t) > x, Y_1 > t) = P(Y_1 > t + x) = 1 - F(t + x).$$

对于第二项, 由于 Y_1, Y_2, \dots 独立同分布, 我们有

$$\begin{aligned} P(R(t) > x, Y_1 \leq t) &= P(S_{N(t)+1} - t > x, Y_1 \leq t) \\ &= P(S_{N(t)+1} - t > x, N(t) \geq 1) \\ &= \sum_{n=1}^{\infty} P(S_{n+1} - t > x, S_n \leq t < S_{n+1}) \\ &= \sum_{n=1}^{\infty} \int_0^t P(s + \sum_{i=2}^{n+1} Y_i - t > x, s + \sum_{i=2}^n Y_i \leq t < s + \sum_{i=2}^{n+1} Y_i) dF(s) \\ &= \sum_{n=1}^{\infty} \int_0^t P(\sum_{i=2}^{n+1} Y_i - (t-s) > x, \sum_{i=2}^n Y_i \leq t-s < \sum_{i=2}^{n+1} Y_i) dF(s) \\ &= \sum_{n=1}^{\infty} \int_0^t P(S_n - (t-s) > x, S_{n-1} \leq t-s < S_n) dF(s) \\ &= \sum_{n=1}^{\infty} \int_0^t P(S_{N(t-s)+1} - (t-s) > x, N(t-s) = n-1) dF(s) \\ &= \int_0^t P(S_{N(t-s)+1} - (t-s) > x) dF(s) \\ &= \int_0^t P(R(t-s) > x) dF(s). \end{aligned}$$

于是命题得证. □

命题 5.2 对 $x > 0$,

$$P(A(t) \leq x) = (1 - F(t))1_{[0, x]}(t) + \int_0^t P(A(t-s) \leq x) dF(s). \quad (5.11)$$

证明 注意到在 $\{Y_1 > t\}$ 上, $A(t) = t$. 类似于上面命题的证明, 我们有

$$\begin{aligned} P(A(t) \leq x) &= P(A(t) \leq x, Y_1 > t) + P(A(t) \leq x, Y_1 \leq t) \\ &= P(t \leq x, Y_1 > t) + P(A(t) \leq x, Y_1 \leq t) \\ &= (1 - F(t))1_{[0, x]}(t) + \int_0^t P(A(t-s) \leq x) dF(s) \end{aligned}$$

□

例 5.7 设 $N(t)$ 是参数为 λ 的 Poisson 过程, $A(t)$ 和 $R(t)$ 分别为 $N(t)$ 在时刻 t 的年龄和剩余寿命. 根据命题 5.1,

$$\begin{aligned} P(R(t) > x) &= 1 - F(t+x) + \lambda \int_0^t (1 - F(x+t-s)) ds \\ &= e^{-\lambda(x+t)} + \lambda \int_0^t e^{-\lambda(x+t-s)} ds \\ &= e^{-\lambda x}. \end{aligned}$$

由命题 5.2,

$$\begin{aligned} P(A(t) \leq x) &= (1 - F(t))1_{[0, x]}(t) + \lambda \int_0^t (1 - F(t-s))1_{[0, x]}(t-s) ds. \end{aligned}$$

这样, 当 $t \leq x$ 时,

$$\begin{aligned} P(A(t) \leq x) &= 1 - F(t) + \lambda \int_0^t (1 - F(t-s)) ds \\ &= e^{-\lambda t} + \lambda \int_0^t e^{-\lambda s} ds = 1. \end{aligned}$$

而当 $t > x$, 则

$$\begin{aligned} P(A(t) \leq x) &= \lambda \int_{t-x}^t (1 - F(t-s)) ds \\ &= \lambda \int_0^x (1 - F(s)) ds = \lambda \int_0^x e^{-\lambda s} ds = 1 - e^{-\lambda x}. \end{aligned}$$

下面考虑另一种方法. 以时刻 t 前最近的跳做条件. 则有

$$\begin{aligned}
 P(R(t) > x) &= P(S_{N(t)+1} - t > x) \\
 &= \sum_{n=0}^{\infty} P(S_{n+1} - t > x; N(t) = n) \\
 &= \sum_{n=0}^{\infty} P(S_{n+1} - t > x; S_n \leq t < S_{n+1}) \\
 &= \sum_{n=0}^{\infty} \int_0^t P(s - t + Y_{n+1} > x; t < s + Y_{n+1}) F^{n*}(ds) \\
 &= \int_0^t P(Y_{n+1} > t + x - s) \sum_{n=0}^{\infty} F^{n*}(ds) \\
 &= 1 - F(t + x) + \int_0^t (1 - F(t + x - s)) U(ds)
 \end{aligned}$$

推论 5.2 更新过程是 *Poisson* 过程的充分必要条件是剩余寿命与 F 同分布.

证明 只需证明充分性. 由于

$$P(R(t) > x) = 1 - F(x),$$

所以由命题 5.1,

$$P(R(t) > x) = \bar{F}(t + x) + \int_0^t P(R(t - s) > x) dF(s).$$

于是

$$\bar{F}(x) = \bar{F}(t + x) + \int_0^t \bar{F}(x) dF(s) = \bar{F}(t + x) + \bar{F}(x) F(t),$$

由此

$$\bar{F}(t + x) = \bar{F}(t) \bar{F}(x),$$

因此 F 是指数分布. □

推论 5.3 更新过程是 *Poisson* 过程的充分必要条件是年龄分布是

$$P(A(t) \leq x) = \begin{cases} F(x), & x < t, \\ 1, & x \geq t. \end{cases}$$

证明 只需证明充分性. 设 $x < t$, 由

$$P(A(t) \leq x) = (1 - F(t))1_{[0, x]}(t) + \int_0^t P(A(t-s) \leq x) dF(s)$$

得

$$\begin{aligned} F(x) &= \int_0^t P(A(t-s) \leq x) dF(s) \\ &= \int_{[0, t-x)} F(x) dF(s) + \int_{[t-x, t]} dF(s) \\ &= F(x)F(t-x-) + F(t) - F(t-x-), \end{aligned}$$

即

$$\bar{F}(t) = \bar{F}(x)(1 - F(t-x-)).$$

在上式中取右极限得

$$\bar{F}(t) = \bar{F}(x)\bar{F}(t-x).$$

由此可知 $F(x)$ 为指数分布. □

命题 5.3 对 $t, x \geq 0$,

$$P(\beta(t) > x) = \bar{F}(\max(t, x)) + \int_0^t \bar{F}(\max(t-s, x)) dU(s),$$

其中 $\bar{F}(t) = 1 - F(t)$.

证明 对第一个更新时刻 Y_1 取条件概率,

$$P(\beta(t) > x | Y_1 = s) = \begin{cases} 1_{(x, \infty)}(s), & \text{若 } s > t; \\ P(\beta(t-s) > x), & \text{若 } s \leq t. \end{cases}$$

这样,

$$\begin{aligned} P(\beta(t) > x) &= \int_0^\infty P(\beta(t) > x | Y_1 = s) dF(s) \\ &= \int_{s>t} 1_{(x, \infty)}(s) dF(s) + \int_0^t P(\beta(t-s) > x) dF(s) \\ &= \bar{F}(\max(t, x)) + \int_0^t P(\beta(t-s) > x) dF(s). \end{aligned}$$

□

5.4 平稳更新过程

在这一节中, 我们首先研究延迟更新过程相关的几个方程.

定理 5.7 设 S_n 是延迟更新序列, 则

$$V(t) = G(t) + \int_0^t V(t-s)dF(s), \quad (5.12)$$

$$V(t) = G(t) + \int_0^t U(t-s)dG(s). \quad (5.13)$$

证明 类似于普通更新情形,

$$V(t) = \sum_{n=1}^{\infty} P(S_n \leq t) = G(t) + \sum_{n=1}^{\infty} G * F^{n*}(t) = G(t) + G * U(t).$$

这样, 由 $G * U = U * G$ 可知 (5.13) 成立. 又由定理 5.5, V 是更新方程 $V = G + V * F$ 的解, 即 (5.12) 成立. □

在延迟更新情形下, $R(t)$, $A(t)$, $\beta(t)$ 满足的方程与普通更新类似. 我们只给出 $\beta(t)$ 的证明.

定理 5.8 设 $R(t)$ 是延迟更新过程的剩余寿命, $R_o(t)$ 是普通更新过程的剩余寿命, 则

$$P(R(t) > x) = \bar{G}(t+x) + \int_0^t P(R_o(t-y) > x)dG(y).$$

定理 5.9 设 $A(t)$ 是延迟更新过程的年龄, $A_o(t)$ 是普通更新的年龄, 则

$$P(A(t) \leq x) = \bar{G}(t)1_{[0, x]}(t) + \int_0^t P(A_o(t-s) \leq x)dF(s).$$

定理 5.10 设 $\beta(t)$ 是延迟更新过程的总寿命函数, $\beta_o(t)$ 是普通更新过程的总寿命. 则对 $t, x \geq 0$,

$$P(\beta(t) > x) = \bar{G}(\max(t, x)) + \int_0^t P(\beta_o(t-s) > x)dG(s),$$

$$P(\beta(t) > x) = \bar{G}(\max(t, x)) + \int_0^t \bar{F}(\max(t-s, x))dV(t).$$

证明 类似于命题 5.3,

$$\begin{aligned}
P(\beta(t) > x) &= \int_0^\infty P(\beta(t) > x | Y_1 = s) dG(s) \\
&= \int_{s>t} 1_{(x, \infty)}(s) dG(s) + \int_0^t P(\beta_o(t-s) > x) dG(s) \\
&= \bar{G}(\max(t, x)) + \int_0^t P(\beta_o(t-s) > x) dG(s).
\end{aligned}$$

又由命题 5.3,

$$P(\beta_o(t) > x) = \bar{F}(\max(t, x)) + \int_0^t \bar{F}(\max(t-s, x)) dU(s).$$

代入上式得

$$\begin{aligned}
P(\beta(t) > x) &= \bar{G}(\max(t, x)) + \bar{F}(\max(t, x)) * (G + G * U)(t) \\
&= \bar{G}(\max(t, x)) + \bar{F}(\max(., x)) * V(t).
\end{aligned}$$

□

对于参数为 λ 的 Poisson 过程, 更新函数 $U(t) = \lambda t$. 一般的, 我们有

定理 5.11 设 $\mu_F = E(Y_2) < \infty$, 则延迟更新过程 $N(t)$ 的更新函数 $V(t) = t/\mu_F$, $t > 0$ 的充分必要条件是

$$G(t) = \frac{1}{\mu_F} \int_0^t (1 - F(s)) ds.$$

证明 记

$$F_e(t) = \frac{1}{\mu_F} \int_0^t (1 - F(s)) ds,$$

则

$$\begin{aligned}
\hat{F}_e(\lambda) &= \int_0^\infty e^{-\lambda x} dF_e(x) = \frac{1}{\mu_F} \int_0^\infty e^{-\lambda x} (1 - F(x)) dx \\
&= \frac{1}{\lambda \mu_F} - \frac{1}{\mu_F} \int_0^\infty e^{-\lambda x} F(x) dx = \frac{1}{\lambda \mu_F} - \frac{1}{\mu_F} \int_0^\infty e^{-\lambda x} dx \int_0^x dF(u) \\
&= \frac{1}{\lambda \mu_F} - \frac{1}{\mu_F} \int_0^\infty \int_u^\infty e^{-\lambda x} dx dF(u) \\
&= \frac{1}{\lambda \mu_F} - \frac{1}{\mu_F} \int_0^\infty \frac{e^{-\lambda u}}{\lambda} dF(u) = \frac{1 - \hat{F}(\lambda)}{\lambda \mu_F}
\end{aligned}$$

若 $G(t) = F_e(t)$, 则

$$\hat{V}(\lambda) = \frac{\hat{G}(\lambda)}{1 - \hat{F}(\lambda)} = \frac{\hat{F}_e(\lambda)}{1 - \hat{F}(\lambda)} = \frac{1}{\lambda\mu_F},$$

从而 $V(t) = t/\mu_F$. 反过来, 若 $V(t) = t/\mu_F$, 则

$$\hat{V}(\lambda) = \frac{1}{\lambda\mu_F}.$$

而

$$\hat{G}(\lambda) = \hat{V}(\lambda)(1 - \hat{F}(\lambda)) = \frac{1 - \hat{F}(\lambda)}{\lambda\mu_F} = \hat{F}_e(\lambda),$$

故 $G = F_e$. □

定义 5.3 设 $\mu_F = \int_0^\infty t dF(t) < \infty$. 记

$$F_e(t) = \frac{1}{\mu_F} \int_0^t [1 - F(s)] ds, \quad t \geq 0$$

若 Y_1 的分布为 $F_e(x)$, 我们称更新过程 $N(t)$ 为平稳 (稳定) 更新过程.

定理 5.12 设 $N(t)$ 平稳, 则

$$P(R(t) > x) = P(R(0) > x) = P(Y_1 > x) = 1 - F_e(x).$$

证明 由定理 5.8,

$$\begin{aligned} P(R(t) > x) &= 1 - G(t+x) + \int_0^t P(R_o(t-y) > x) dG(y) \\ &= 1 - G(t+x) + G * (1 - F(\cdot + x) + U * (1 - F(\cdot + x)))(t) \\ &= 1 - G(t+x) + V * (1 - F(\cdot + x))(t) \\ &= 1 - F_e(t+x) + \frac{1}{\mu_F} \int_0^t (1 - F(t-y+x)) dy \\ &= \frac{1}{\mu_F} \int_{t+x}^\infty (1 - F(y)) dy + \frac{1}{\mu_F} \int_x^{t+x} (1 - F(u)) du \\ &= 1 - F_e(x) = P(R(0) > x) = P(Y_1 > x). \end{aligned}$$

□

定理 5.13 设 $N(t)$ 平稳, 则对任意的 $h > 0$,

(i) $\{S_{N(h)+n} - h, n \geq 1\}$ 和 $\{S_n, n \geq 1\}$ 同分布.

(ii) $\{Y_n, n \geq 1\}$ 与

$$\{S_{N(h)+1} - h, Y_{N(h)+1+k}, k \geq 1\} = \{R(h), Y_{N(h)+1+k}, k \geq 1\}.$$

同分布.

证明 只需要证明 (ii).

$$\begin{aligned} & P(S_{N(h)+1} - h \leq x_0, Y_{N(h)+1+i} \leq x_i, 1 \leq i \leq k) \\ &= \sum_{n=0}^{\infty} P(S_{n+1} - h \leq x_0, Y_{n+1+i} \leq x_i, 1 \leq i \leq k, N(h) = n) \\ &= \sum_{n=0}^{\infty} P(S_{n+1} - h \leq x_0, Y_{n+1+i} \leq x_i, 1 \leq i \leq k, S_n \leq h < S_{n+1}) \\ &= \sum_{n=0}^{\infty} P(S_{n+1} - h \leq x_0, S_n \leq h < S_{n+1}) \prod_{i=1}^k P(Y_{n+1+i} \leq x_i) \\ &= \sum_{n=0}^{\infty} P(S_{n+1} - h \leq x_0, N(h) = n) \prod_{i=1}^k F(x_i) \\ &= P(S_{N(h)+1} - h \leq x_0) \prod_{i=1}^k F(x_i) = P(R(h) \leq x_0) \prod_{i=1}^k F(x_i) \\ &= P(Y_1 \leq x_0, Y_{i+1} \leq x_i, 1 \leq i \leq k). \end{aligned}$$

□

定理 5.14 $N(t)$ 是平稳更新过程的充分必要条件是对任意 $h > 0, k \geq 1$,

$$\{N(t_i, t_i + s_i], i = 1, 2, \dots, k\}$$

与

$$\{N(t_i + h, t_i + s_i + h], i = 1, 2, \dots, k\}$$

同分布.

证明 设 $N(t)$ 是平稳更新过程. 由于

$$\begin{aligned}
& P(N(t_i, t_i + s_i] = m_i, 1 \leq i \leq k) \\
&= P\left(\bigcup_{n=0}^{\infty} (S_n, S_{n+1}, \dots, S_{n+m_i-1} \in (t_i, t_i + s_i], \right. \\
&\quad \left. S_{n-1} \leq t_i, S_{n+m_i} > t_i + s_i), 1 \leq i \leq k\right)
\end{aligned}$$

及

$$\begin{aligned}
& P(N(t_i + h, t_i + s_i + h] = m_i, 1 \leq i \leq k) \\
&= P\left(\bigcup_{n=0}^{\infty} \{S_{N(h)+n}, S_{N(h)+n+1}, \dots, S_{N(h)+n+m_i-1} \in (t_i + h, t_i + s_i + h], \right. \\
&\quad \left. S_{N(h)+n-1} \leq t_i + h, S_{N(h)+n+m_i} > t_i + s_i + h), 1 \leq i \leq k\right) \\
&= P\left(\bigcup_{n=0}^{\infty} (S_{N(h)+n} - h, S_{N(h)+n+1} - h, \dots, S_{N(h)+n+m_i-1} - h \in (t_i, t_i + s_i], \right. \\
&\quad \left. S_{N(h)+n-1} - h \leq t_i, S_{N(h)+n+m_i} - h > t_i + s_i), 1 \leq i \leq k\right),
\end{aligned}$$

应用上面定理即得.

反过来, 则对任意的 $t \geq 0, h > 0, N(t, t+h]$ 与 $N(0, h]$ 同分布, 于是

$$V(t+h) - V(t) = E(N(t, t+h]) = E(N(0, h]) = V(h).$$

这样, 存在 μ 使得 $V(t) = t/\mu$. 由定理 5.11 知 $N(t)$ 是平稳更新过程. \square

5.5 更新定理

在这一节, 我们将研究与更新过程相关的一些量在 $t \rightarrow \infty$ 时的极限性质.

定理 5.15 (Elementary renewal theorem) 设 $P(Y_1 < \infty) = 1, \mu = E(Y_2) \leq \infty$, 则

$$\lim_{t \rightarrow \infty} \frac{U(t)}{t} = \frac{1}{\mu}, \quad \lim_{t \rightarrow \infty} \frac{V(t)}{t} = \frac{1}{\mu}.$$

证明 应用定理 5.2 及 Fatou 引理,

$$\mu^{-1} = E(\liminf_{t \rightarrow \infty} \frac{N(t)}{t}) \leq \liminf_{t \rightarrow \infty} E(\frac{N(t)}{t}) = \liminf_{t \rightarrow \infty} \frac{V(t)}{t}.$$

对 $b > 0$, 令 $Y_i^* = Y_i \wedge b, S_n^* = Y_1^* + \dots + Y_n^*, N^*(t) = \sum_{n=1}^{\infty} 1_{[0, t]}(S_n^*)$, 则

$$V^*(t) = E[N^*(t)] \geq E[N(t)] = V(t).$$

这样,

$$\begin{aligned} \limsup_{t \rightarrow \infty} \frac{V(t)}{t} &\leq \limsup_{t \rightarrow \infty} \frac{V^*(t)}{t} \leq \limsup_{t \rightarrow \infty} \frac{E\{N^*(t) + 1\}}{t} \\ &= \limsup_{t \rightarrow \infty} \frac{E[S_{N^*(t)+1}^*]}{tE\{Y_1^*\}} \leq \limsup_{t \rightarrow \infty} \frac{t+b}{tE(Y_1^*)} = \frac{1}{E(Y_1^*)} \rightarrow \frac{1}{\mu}, \quad (b \rightarrow \infty). \end{aligned}$$

□

5.5.1 直接黎曼可积

设 f 是 $[0, \infty)$ 上的非负函数. 称 f 在 $[0, \infty)$ 上黎曼可积, 若对任意的 $a > 0$, f 在 $[0, a]$ 上黎曼可积, 且极限

$$\lim_{a \rightarrow \infty} \int_0^a f(s) ds$$

存在. 记

$$\int_0^\infty f(s) ds = \lim_{a \rightarrow \infty} \int_0^a f(s) ds.$$

在考虑关键更新定理时, 我们希望对某些函数 z , 极限 $\lim_{t \rightarrow \infty} (U * z)(t)$ 存在. 下面的例子说明, 仅要求 z 黎曼可积是不够的.

例 5.8 设 $0 < \alpha < 1$ 为无理数, F 是 $\{1 - \alpha, \alpha\}$ 上的两点分布, 则 U 负荷集中在集合

$$\{k - n\alpha, n \in \mathbb{Z}, k - n\alpha \geq 0\}$$

上. 设 $z(k - n\alpha) = 1$, z 在点 $k - n\alpha$ 附近的图像是一个底边充分小的三角形, 使得这些三角形两两不交, 且 z 在 $[0, \infty)$ 上黎曼可积. 这样, 对于正整数 r ,

$$\begin{aligned} \int_0^r z(r-y)U(dy) &= \sum_{k, n: k-n\alpha \leq r} z(r-(k-n\alpha))U(\{k-n\alpha\}) \\ &= \sum_{k, n: k-n\alpha \leq r} U(\{k-n\alpha\}) = U(r) \rightarrow \infty. \end{aligned}$$

因此关键更新定理不成立. 其原因是 z 在 ∞ 处振幅过大. 为了解决这个障碍, 我们引入直接黎曼可积的概念.

对 $k \geq 1, h > 0$, 令

$$\underline{m}_k(h) = \inf_{(k-1)h \leq t \leq kh} f(t), \quad \overline{m}_k(h) = \sup_{(k-1)h \leq t \leq kh} f(t),$$

$$\underline{\sigma}(h) = \sum_{k=1}^{\infty} h \underline{m}_k(h), \quad \overline{\sigma}(h) = \sum_{k=1}^{\infty} h \overline{m}_k(h).$$

定义 5.4 称 f 为直接黎曼可积的, 若

(i) 对所有 $h > 0$, $\overline{\sigma}(h) < \infty$;

(ii) $\lim_{h \rightarrow 0} \{\overline{\sigma}(h) - \underline{\sigma}(h)\} = 0$.

注 5.1 若对某个 $h > 0$, $\overline{\sigma}(h) = \infty$, 则对任意的 $h' > 0$, $\overline{\sigma}(h') = \infty$.

注 5.2 若函数 f 具有紧支撑, 即存在 $b > 0$ 使得当 $t > b$ 时, $f(t) = 0$, 则 f 直接黎曼可积当且仅当它是黎曼可积的.

引理 5.1 设 $f \geq 0$ 在 $[0, \infty)$ 上直接黎曼可积, 则它在 $[0, \infty)$ 上黎曼可积, 且

$$\lim_{h \rightarrow 0} \overline{\sigma}(h) = \lim_{h \rightarrow 0} \underline{\sigma}(h) = \int_0^{\infty} f(x) dx.$$

证明 对任意的 $T > 0$, 设 $a_k = \frac{T}{k}$, 则

$$\lim_{k \rightarrow \infty} a_k \sum_{n=1}^k (\overline{m}_n(a_k) - \underline{m}_n(a_k)) \leq \lim_{k \rightarrow \infty} (\overline{\sigma}(a_k) - \underline{\sigma}(a_k)) = 0.$$

所以 f 在 $[0, T]$ 上黎曼可积, 而且

$$\lim_{k \rightarrow \infty} a_k \sum_{n=1}^k \overline{m}_n(a_k) = \lim_{k \rightarrow \infty} a_k \sum_{n=1}^k \underline{m}_n(a_k) = \int_0^T f(x) dx.$$

任取正整数 $M > 0$, 取 $a_k = \frac{1}{2^k}$, 则

$$\begin{aligned} & \left| \overline{\sigma}(a_k) - \int_0^M f(x) dx \right| \\ & \leq \left| \overline{\sigma}(a_k) - a_k \sum_{n=1}^{2^k M} \overline{m}_n(a_k) \right| + \left| a_k \sum_{n=1}^{2^k M} \overline{m}_n(a_k) - \int_0^M f(x) dx \right| \\ & = a_k \sum_{n=2^k M+1}^{\infty} \overline{m}_n(a_k) + \left| a_k \sum_{n=1}^{2^k M} \overline{m}_n(a_k) - \int_0^M f(x) dx \right| \\ & \leq \sum_{n=M+1}^{\infty} \overline{m}_n(1) + \left| a_k \sum_{n=1}^{2^k M} \overline{m}_n(a_k) - \int_0^M f(x) dx \right|. \end{aligned}$$

先令 $k \rightarrow \infty$, 再令 $M \rightarrow \infty$ 即得所要结论. \square

引理 5.2 若 $f \geq 0$ 在 $[0, \infty)$ 上单调下降, 且 $\int_0^\infty f(x)dx < \infty$, 则 f 在 $[0, \infty)$ 上直接黎曼可积.

证明 对任意的 $a > 0$, 有 $\underline{m}_n(a) = f(na)$, $\overline{m}_n(a) = f((n-1)a)$. 所以

$$\begin{aligned} a \sum_{n=1}^{\infty} \underline{m}_n(a) &= a \sum_{n=1}^{\infty} f(na) \geq \int_a^\infty f(x)dx, \\ a \sum_{n=1}^{\infty} \overline{m}_n(a) &= a \sum_{n=1}^{\infty} f((n-1)a) \leq af(0) + \int_0^\infty f(x)dx. \end{aligned}$$

因此可得

$$\lim_{a \rightarrow 0} a \sum_{n=1}^{\infty} \underline{m}_n(a) = \lim_{a \rightarrow 0} a \sum_{n=1}^{\infty} \overline{m}_n(a) = \int_0^\infty f(x)dx.$$

\square

引理 5.3 设

(i) f 在任意有限区间黎曼可积;

(ii) $\overline{\sigma}(1) < \infty$,

则 f 是直接黎曼可积的.

证明 若 $h < 1$, 则 $\underline{\sigma}(h) \leq \overline{\sigma}(h) \leq \overline{\sigma}(1) < \infty$. 其中 $\overline{\sigma}(h) = \sum_{n=1}^{\infty} h \overline{m}_n(h)$. 由假设 (ii), 对任意的 $\varepsilon > 0$, 存在 N_0 , 使得 $\sum_{n > N_0} \overline{m}_n(1) < \varepsilon$. 所以

$$\begin{aligned} & \overline{\sigma}(h) - \underline{\sigma}(h) \\ &= h \sum_{n: nh \leq N_0} (\overline{m}_n(h) - \underline{m}_n(h)) + h \sum_{n: nh > N_0} (\overline{m}_n(h) - \underline{m}_n(h)) \\ &\leq h \sum_{n: nh \leq N_0} (\overline{m}_n(h) - \underline{m}_n(h)) + 2h \sum_{n: nh > N_0} \overline{m}_n(h) \\ &\leq h \sum_{n: nh \leq N_0} (\overline{m}_n(h) - \underline{m}_n(h)) + 2\varepsilon. \end{aligned}$$

令 $h \rightarrow 0$, 则 $h \sum_{n: nh \leq N_0} \overline{m}_n(h) \rightarrow 0$, 因此 f 是直接黎曼可积的. \square

推论 5.4 设 g 直接黎曼可积, 函数 f 在 $[0, \infty)$ 上黎曼可积, 且 $f \leq g$, 则 f 也直接黎曼可积.

证明 注意到

$$\sum_{n=1}^{\infty} \bar{m}_n(1) \leq \sum_{n=1}^{\infty} \sup_{n-1 \leq s < n} g(s) < \infty,$$

应用引理 5.3 即得. □

5.5.2 更新定理

定理 5.16 设 $F(0) < 1$, $F(\infty) = 1$,

$$\begin{aligned} \mu = \mu_F &= \int_0^{\infty} x dF(x) \leq \infty, \\ F_e(x) &= \frac{1}{\mu} \int_0^x (1 - F(y)) dy, \end{aligned}$$

则下列条件等价:

(i) (**Blackwell renewal theorem**) 若 $G(\infty) = 1$, 则对任意的 $b > 0$,

$$\lim_{t \rightarrow \infty} V(t, t+b] = \lim_{t \rightarrow \infty} V(t+b) - V(t) = \frac{b}{\mu}.$$

(ii) (**Smith key renewal theorem**) 若 $z(t)$ 直接黎曼可积, 则

$$\lim_{t \rightarrow \infty} U * z(t) = \frac{1}{\mu} \int_0^{\infty} z(s) ds.$$

(iii) 若 $G(\infty) = 1$, 则对任意的 $x > 0$,

$$\lim_{t \rightarrow \infty} P(R(t) \leq x) = F_e(x).$$

(iv) 若 $G(\infty) = 1$, 则对任意的 $x > 0$,

$$\lim_{t \rightarrow \infty} P(A(t) \leq x) = F_e(x).$$

证明 (iii) 与 (iv) 的等价性由

$$P(R(t) \leq x) = P(N(t, t+x] \geq 1) = P(A(t+x) < x)$$

即可看出. 下面我们证明 (ii) \Rightarrow (iv), (iii) \Rightarrow (i) \Rightarrow (ii).

(1) (ii) \Rightarrow (iv). 对普通更新情形, 记

$$Z(t) = P(A(t) \leq x) = z(t) + U * z(t),$$

其中 $z(t) = (1 - F(t))1_{[0, x]}(t)$ 是直接黎曼可积的, 所以

$$\lim_{t \rightarrow \infty} Z(t) = \frac{1}{\mu} \int_0^\infty z(s) ds = \frac{1}{\mu} \int_0^x (1 - F(s)) ds.$$

对延迟更新情形,

$$P(A(t) \leq x) = P(A(t) \leq x, Y_1 > t) + \int_0^t Z(t-s) dG(s).$$

由 $G(\infty) = 1$, 容易看出

$$\begin{aligned} \lim_{t \rightarrow \infty} P(A(t) \leq x, Y_1 > t) &= 0, \\ \lim_{t \rightarrow \infty} \int_0^t Z(t-s) dG(s) &= \lim_{t \rightarrow \infty} Z(t) = \frac{1}{\mu} \int_0^x (1 - F(s)) ds. \end{aligned}$$

(2) (iii) \Rightarrow (i). 记

$$G_t(x) = P(R(t) \leq x).$$

由

$$V(t, t+b] = G_t * (1_{[0, \infty)} + U)(b)$$

可知

$$\begin{aligned} G_t * (1_{[0, \infty)} + U)(b) &= G_t(b) + \int_0^b G_t(b-s) dU(s) \\ &\rightarrow F_e(b) + \int_0^b F_e(b-s) dU(s) = F_e * (1_{[0, \infty)} + U)(b) = \frac{b}{\mu}. \end{aligned}$$

(3) (i) \Rightarrow (ii) 取 $h > 0$ 使得 $F(h) < 1$. 令

$$\begin{aligned} \bar{z}(t) &= \sum_{n=1}^{\infty} \bar{m}_n(h) 1_{[(n-1)h, nh)}(t), \\ \underline{z}(t) &= \sum_{n=1}^{\infty} \underline{m}_n(h) 1_{[(n-1)h, nh)}(t). \end{aligned}$$

应用引理 5.4, 当 $t \rightarrow \infty$ 时,

$$\begin{aligned}
U * \bar{z}(t) &= \int_0^t \bar{z}(t-s) dU(s) = \sum_1^\infty \bar{m}_n(h) U(t-nh, t-(n-1)h] \\
&\rightarrow \sum_1^\infty \bar{m}_n(h) \frac{h}{\mu} = \frac{\bar{\sigma}(h)}{\mu}.
\end{aligned}$$

由于 $U(t-nh, t-(n-1)h] \leq \frac{1}{1-F(h)}$, 同理 $\lim_{t \rightarrow \infty} U * \underline{z}(t) = \frac{\underline{\sigma}(h)}{\mu}$.

□

引理 5.4 若 $F(b) < 1$, 则对 $t > b$,

$$U(t-b, t] \leq (1-F(b))^{-1},$$

于是

$$\sup_{t \geq 0} U(t, t+b] \leq (1-F(b))^{-1}.$$

证明 因为 $U = F + F * U$, $U * (1-F) = F$, 所以

$$\begin{aligned}
F(t) &= \int_0^t (1-F(t-s)) dU(s) \geq \int_{t-b}^t (1-F(t-s)) dU(s) \\
&\geq (1-F(b)) U(t-b, t].
\end{aligned}$$

这样,

$$U(t-b, t] \leq \frac{F(t)}{1-F(b)} \leq \frac{1}{1-F(b)}.$$

□

对 \mathbb{R} 上分布函数 F , 定义 F 的支集为

$$\text{supp}(F) = \{x \in \mathbb{R} : F(x+\varepsilon) - F(x-\varepsilon) > 0, \forall \varepsilon > 0\}.$$

定义 5.5 称分布函数 $F(x)$ 为算术分布, 若存在 $d > 0$, 使得

$$\text{supp}(F) \subset \{nd : n = 0, \pm 1, \pm 2, \dots\}.$$

如果不存在上述的 $d > 0$, 则称 $F(x)$ 为非算术分布.

我们已经证明了几个更新定理的等价性. 事实上, 在非算术分布情形, 上面的更新定理都是成立的. 我们不加证明的给出下面结果.

定理 5.17 设 F 非算术分布, 则对 $x > 0$,

$$\lim_{t \rightarrow \infty} P(R(t) \leq x) = F_e(x).$$

推论 5.5 若 F 非算术分布, $\mu_F < \infty$, 则

$$\lim_{t \rightarrow \infty} P(\beta(t) > x) = \frac{1}{\mu_F} \int_x^\infty t dF(t). \quad (5.14)$$

证明 由于 $\bar{F}(\max(x, t))$ 单调下降, 且

$$\begin{aligned} \int_0^\infty \bar{F}(\max(x, t)) dt &= \int_0^x \bar{F}(x) dt + \int_x^\infty \bar{F}(t) dt \\ &= x\bar{F}(x) + \int_x^\infty \bar{F}(t) dt = \int_x^\infty t dF(t), \end{aligned}$$

应用引理 5.2, 可得 $\bar{F}(\max(x, t))$ 是直接黎曼可积的. 这样, 由

$$P(\beta_o(t) > x) = \bar{F}(\max(t, x)) + \int_0^t \bar{F}(\max(t-s, x)) dU(s)$$

及更新定理可知在普通更新情形下,

$$\lim_{t \rightarrow \infty} P(\beta_o(t) > x) = \frac{1}{\mu_F} \int_x^\infty t dF(t).$$

对于延迟更新情形, 利用公式

$$P(\beta(t) > x) = \bar{G}(\max(x, t)) + \int_0^t P(\beta_o(t-s) > x) dG(s)$$

即得. □

命题 5.4 设 $F(\infty) = 1$, $\mu = \mu_F < \infty$, 且

$$\sigma^2 = \int_0^\infty (u - \mu)^2 dF(u) < \infty,$$

则

$$\lim_{t \rightarrow \infty} (U(t) - \frac{t}{\mu}) = \frac{1}{2\mu^2} \int_0^\infty x^2 dF(x) = \frac{\sigma^2 - \mu^2}{2\mu^2}.$$

证明

$$U(t) - \frac{t}{\mu} = U(t) - F_e(t) - F_e * U(t) = U * (1 - F_e)(t) - F_e(t).$$

如果能够证明 $1 - F_e(t)$ 直接黎曼可积, 由更新定理,

$$\begin{aligned}
\lim_{t \rightarrow \infty} U(t) - \frac{t}{\mu} &= \frac{1}{\mu} \int_0^\infty (1 - F_e(t)) dt - 1 \\
&= \frac{1}{\mu^2} \int_0^\infty \int_t^\infty (1 - F(s)) ds dt - 1 = \frac{1}{\mu^2} \int_0^\infty \int_0^s (1 - F(s)) t ds - 1 \\
&= \frac{1}{\mu^2} \int_0^\infty s(1 - F(s)) ds - 1 = \frac{1}{\mu^2} \int_0^\infty s \int_s^\infty dF(t) ds - 1 \\
&= \frac{1}{\mu^2} \int_0^\infty \int_0^t s ds dF(t) - 1 = \frac{1}{2\mu^2} \int_0^\infty t^2 dF(t) - 1 = \frac{\sigma^2 - \mu^2}{2\mu^2}.
\end{aligned}$$

因为 $1 - F_e(t)$ 递减, 由上面计算可知 $1 - F_e(t)$ 可积, 应用引理 5.2, $1 - F_e(t)$ 直接黎曼可积, 于是命题成立. \square

5.5.3 离散更新理论

在这一节中, 我们考虑 F 是步长为 d 的算术分布情形下的更新定理. 首先我们假设 $d = 1$, 即 $\{Y_n, n \geq 1\}$ 是独立同分布, 取非负整数值的随机变量列, 且 $\text{supp}(F)$ 的最大公约数为 1. 令

$$f_k = P(Y_1 = k), \quad k \geq 0.$$

假设

$$P(Y_1 < \infty) = \sum_{k=0}^{\infty} f_k = 1, \quad f_0 = 0.$$

记

$$u_n = U(\{n\}) = E\left[\sum_{k=0}^{\infty} 1_{\{n\}}(S_k)\right] = P\left(\bigcup_{k=0}^{\infty} \{S_k = n\}\right).$$

注意到 $U(n) = U([0, n]) = \sum_{i=0}^n u_i$, $u_0 = 1$.

令

$$n_0 = \sup\{n : P(Y_1 = n) > 0\}.$$

定义

$$p_n = \begin{cases} P(Y_1 > n | Y_1 > n-1), & 1 \leq n < n_0; \\ 0 & n \geq n_0, \end{cases} \quad q_n = 1 - p_n.$$

考虑马尔可夫链 A_n , 其转移矩阵为

$$P = \begin{pmatrix} 0 & q_1 & p_1 & 0 & 0 & \cdots \\ 1 & q_2 & 0 & p_2 & 0 & \cdots \\ 2 & q_3 & 0 & 0 & p_3 & \cdots \\ \vdots & \vdots & & \ddots & & \end{pmatrix}.$$

这样, 对 $n \leq n_0$,

$$\begin{aligned} f_{00}^{(n)} &= P^0(A_1 = 1, A_2 = 2, \dots, A_{n-1} = n-1, A_n = 0) \\ &= p_1 p_2 \cdots p_{n-1} q_n = \frac{P(Y_1 = n)}{P(Y_1 > 0)} = P(Y_1 = n) = f_n, \end{aligned}$$

对 $n > n_0$, $f_{00}^{(n)} = f_n = 0$. 如果我们假设 $\mu = E(Y_1) < \infty$, 则 0 是 A_n 的正常返状态. 注意到 0 是非周期的, 故为遍历状态.

事实上, A_n 恰好是普通更新序列 S_n 的年龄过程. 注意到

$$p_{00}^{(n)} = P_0(A_n = 0) = u_n,$$

应用初次进入的分解公式 $p_{00}^{(n)} = \sum_{j=0}^n f_{00}^{(j)} p_{00}^{(n-j)}$, 我们有

$$u_n = \sum_{j=0}^n f_j u_{n-j}.$$

令 $U(s) = \sum_{n=0}^{\infty} u_n s^n$, $F(s) = F_{00}(s) = \sum_{n=0}^{\infty} f_n s^n$, 则

$$U(s) = \frac{1}{1 - F(s)}, \quad 0 \leq s < 1.$$

假设 $\mu = E(Y_1) < \infty$, 则

$$u_n = p_{00}^{(n)} \rightarrow 1/\mu \quad (n \rightarrow \infty),$$

这即是 Blackwell 更新定理. 更准确的, 对正整数 h ,

$$U(n, n+h] = u_n + \cdots + u_{n+h} \rightarrow h/\mu \quad (n \rightarrow \infty).$$

设 $z(k)$ 是定义在 $\{0, 1, \dots\}$ 上的非负函数, 满足 $\sum_{k=0}^{\infty} z(k) < \infty$. 特别的, 当 $k \rightarrow \infty$ 时, $z(k) \rightarrow 0$. 这样,

$$\lim_{n \rightarrow \infty} U * z(n) = \mu^{-1} \sum_{k=0}^{\infty} z(k).$$

这就证明了关键更新定理.

下面设 S_n 是延迟更新序列,

$$P(Y_1 = n) = g_n, \quad P(Y_2 = n) = f_n, \quad n \geq 1.$$

令

$$V(\{n\}) = v_n = P(\exists k, S_k = n),$$

则对 $n \geq 1$,

$$v_n = \sum_{k=1}^{\infty} g_k u_{n-k} = U * g(n).$$

由关键更新定理,

$$\lim_{n \rightarrow \infty} v_n = \sum_{k=1}^{\infty} g_k / E(Y_2) = \mu^{-1} \sum_{k=1}^{\infty} g_k.$$

最后我们考虑 $d \neq 1$ 的情形. 令 $Y_n^\sharp = Y_n/d$, 则 Y_n^\sharp 是步长为 1 的算术分布. 这样, 当 $n \rightarrow \infty$ 时,

$$\begin{aligned} U(\{nd\}) &= u_{nd} = P\left(\bigcup_{k=0}^{\infty} \{S_k = nd\}\right) = P\left(\bigcup_{k=0}^{\infty} \{S_k^\sharp = n\}\right) \\ &= u_n^\sharp \rightarrow \frac{d}{E(Y_2)}. \end{aligned}$$

5.5.4 Improper Renewal Equation

这一节我们考虑在 $F(\infty) \neq 1$ 下更新方程解的极限性质.

命题 5.5 设存在 $\beta \in \mathbb{R}$ 使得 $\int_0^\infty e^{\beta x} F(dx) = 1$. 定义

$$Z^*(t) = e^{\beta t} Z(t), \quad z^*(t) = e^{\beta t} z(t), \quad F^*(dt) = e^{\beta t} F(dt).$$

则 F^* 是正常分布函数, 且 $Z^* = z^* + F^* * Z^*$. 若 z^* 直接黎曼可积, 则

$$\lim_{t \rightarrow \infty} Z^*(t) = \lim_{t \rightarrow \infty} e^{\beta t} Z(t) = \frac{\int_0^\infty e^{\beta t} z(t) dt}{\int_0^\infty t e^{\beta t} dF(t)}. \quad (5.15)$$

证明 由于

$$\begin{aligned}
Z^*(t) &= e^{\beta t} Z(t) = e^{\beta t} z(t) + \int_0^t e^{\beta t} Z(t-s) dF(s) \\
&= e^{\beta t} z(t) + \int_0^t e^{\beta(t-s)} Z(t-s) e^{\beta s} dF(s) \\
&= z^*(t) + \int_0^t Z^*(t-s) dF^*(s) \\
&= z^*(t) + Z^* * F^*(t),
\end{aligned}$$

更新方程 $Z^* = z^* + Z^* * F^*$ 成立. 剩余的结论直接应用关键更新定理即得. \square

下面我们分两种情况来讨论.

- (i) 若 $F(\infty) > 1$, 则一定存在 $\beta < 0$, 使 $\int_0^\infty e^{\beta t} F(dt) = 1$. 因此若 $z(t)$ 直接黎曼可积, 则 z^* 也直接黎曼可积, 从而

$$Z(t) \sim e^{-\beta t} \int_0^\infty e^{\beta s} z(s) ds / \mu^*,$$

其中 $\mu^* = \int_0^\infty x e^{\beta x} dF(x)$. 注意到 $x e^{\beta x}$ 为有界函数, 故 $\mu^* < \infty$.

- (ii) 若 $F(\infty) < 1$, 则不一定存在 $\beta \in R$, 使 $\int_0^\infty e^{\beta t} F(dt) = 1$, 例如 $F(x) = 1/2 - x^{-1}$, $x \geq 2$. 因此我们需要假设:

(1) 存在 $\beta > 0$, 使 $\int_0^\infty e^{\beta x} dF(x) = 1$;

(2) $z(\infty)$ 存在.

此时,

$$\begin{aligned}
1 + U(\infty) &= (1 - F(\infty))^{-1} < \infty, \\
Z(\infty) &= z(\infty) + U * z(\infty) = z(\infty) + U(\infty)z(\infty).
\end{aligned}$$

令

$$Z_1(t) = Z(\infty) - Z(t), z_1(t) = z(\infty) - z(t) + Z(\infty)(F(\infty) - F(t)),$$

则 $Z_1 = z_1 + F * Z_1$. 事实上,

$$\begin{aligned}
& z_1 + F * Z_1 \\
&= z(\infty) - z(t) + Z(\infty)(F(\infty) - F(t)) + \int_0^t Z(\infty) - Z(t-s) dF(s) \\
&= -Z(t) + z(\infty) + Z(\infty)(F(\infty) - F(t)) + \int_0^t Z(\infty) dF(s) \\
&= -Z(t) + z(\infty) + Z(\infty)F(\infty) = -Z(t) + Z(\infty) = Z_1(t).
\end{aligned}$$

因此若 $e^{\beta t} z_1(t)$ 直接黎曼可积, 则

$$\begin{aligned}
\lim_{t \rightarrow \infty} e^{\beta t} (Z(\infty) - Z(t)) &= \frac{\int_0^\infty e^{\beta s} z_1(s) ds}{\int_0^\infty s e^{\beta s} dF(s)} \\
&= \frac{\int_0^\infty e^{\beta s} (z(\infty) - z(s)) ds + z(\infty) \beta^{-1}}{\int_0^\infty s e^{\beta s} dF(s)}.
\end{aligned}$$

5.6 再生过程

再生性描述的是随机过程 $X(t)$ 在某些随机时间 S_n 处重新开始的现象. 例如, 对于从状态 0 出发的离散参数马氏链 X_n , S_n 为第 n 次回到 0 的时间, 对任意的 m , $\{X(S_m + n), n \geq 0\}$ 仍旧是一个从 0 出发的离散时间马氏链, 与原马氏链 X_n 具有相同的分布, 且与 S_0, \dots, S_m 独立. 下面我们给出再生过程的定义.

定义 5.6 称过程 $X(t)$ 为再生过程, 若存在一系列随机时间 S_n 满足

(i) $\{S_n\}$ 是更新序列,

(ii) 对每个 n , 过程 $X(S_n + t)$ 与原过程 $X(t)$ 具有相同的分布, 即对任意的 k , $0 < t_1 < \dots < t_k$,

$$\{X(S_n + t_i), 1 \leq i \leq k\} \stackrel{d}{=} \{X(t_i), 1 \leq i \leq k\},$$

(iii) 对任意 n , 过程 $\{X(S_n + t), t \geq 0\}$ 与 S_1, \dots, S_n 独立.

随机时间 S_n 称为过程的再生点, $\{X(t), S_n \leq t < S_{n+1}\}$ 称为过程的第 n 次 cycle.

例 5.9 设 X_n 是从 0 出发的常返离散参数马氏链, S_n 是第 n 次回到 0 的时间. 则 X_n 是以 S_n 为再生点的再生过程.

例 5.10 设序列 $(\{Y_j(t), t \geq 0\}, Y_j)$ 独立同分布, 设 $S_0 = 0, S_n = \sum_{j=1}^n Y_j$. 则过程

$$X(t) = \sum_{j=1}^{\infty} Y_j(t - S_{j-1}) 1_{[S_{j-1}, S_j)}(t)$$

是以 S_n 为再生点的再生过程.

例 5.11 (存储过程) 设 $A(t) = \sum_{i=1}^{N(t)} X_i$ 为复合 Poisson 过程, X_i 是独立同分布的非负随机变量序列, 设

$$X(t) = A(t) - ct.$$

对于 Borel 集 A , 令

$$Z(t) = P(X(t) \in A).$$

假设 $S_0 = 0$. 记

$$K(t, A) = P(X(t) \in A, t < S_1),$$

它表示在首次 cycle 结束之前在 t 时刻进入 A 的概率. 这样, 我们有

$$Z(t) = P(X(t) \in A, t < S_1) + P(X(t) \in A, t \geq S_1).$$

注意到

$$P(X(t) \in A, t \geq S_1) = \int_0^t P(X(t-s) \in A) F(ds),$$

于是我们有

$$Z(t) = K(t, A) + \int_0^t Z(t-s) F(ds). \quad (5.16)$$

解这个更新方程可以得到

$$Z(t) = (K(\cdot, A) * U)(t). \quad (5.17)$$

定理 5.18 (Smith) 设 $\mu = E(S_1)$, $S_0 = 0$. 设 $K(\cdot, A)$ 黎曼可积.

(i) 若 $\mu < \infty$, 则

$$\begin{aligned} \lim_{t \rightarrow \infty} P(X(t) \in A) &= \mu^{-1} \int_0^\infty K(s, A) ds \\ &= \mu^{-1} E \left[\int_0^{S_1} 1_{\{X(s) \in A\}} ds \right]. \end{aligned}$$

(ii) 若 $\mu = \infty$, 则

$$\lim_{t \rightarrow \infty} P(X(t) \in A) = 0.$$

证明 我们首先证明 $K(\cdot, A)$ 直接黎曼可积. 注意到

$$K(t, A) = P(X(t) \in A, t < S_1) \leq 1 - F(t),$$

而 $1 - F(t)$ 单调可积, 故由引理 5.2 知 $1 - F$ 直接黎曼可积. 又 $K(\cdot, A)$ 黎曼可积, 由推论 5.4 可以得到直接黎曼可积.

这样, 应用关键更新定理可得

$$\begin{aligned} \lim_{t \rightarrow \infty} P(X(t) \in A) &= \lim_{t \rightarrow \infty} (K(\cdot, A) * U)(t) \\ &= \mu^{-1} \int_0^\infty K(s, A) ds = \mu^{-1} E \left[\int_0^{S_1} 1_A(X(s)) ds \right]. \end{aligned}$$

□

注 5.3 我们需要对 F 做一些假设使得关键更新定理成立.

注 5.4 一般的, $K(\cdot, A)$ 不一定黎曼可积. Miller 在 1972 年给出了一个例子: 设 F 在点 n^{-1} 处具有负荷 2^{-n} , 再生过程 $X(t) = 1_{\mathbb{Q}}(t)$. 则对 $0 < t \leq 1$,

$$K(t, \{1\}) = P(X(t) = 1, t < S_1) = 1_{\mathbb{Q}}(t)F(t, 1],$$

它不是黎曼可积的. 此时, 当 $t \rightarrow \infty$ 时, $P(X(t) = 1) = 1_{\mathbb{Q}}(t)$ 不收敛.

下面我们考虑在什么样的条件下, $K(t, A)$ 黎曼可积.

命题 5.6 设更新过程 $X(t)$ 轨道右连左极, A 是开集. 则 $K(\cdot, A)$ 黎曼可积.

证明 记

$$\Gamma = \{t \geq 0, F \text{ 在 } t \text{ 连续, 且 } \lim_{s \rightarrow t} X(s) = X(t) \text{ a.s.}\}.$$

则 $[0, \infty) - \Gamma$ 至多可数, 故为零测集. 对 $t \in \Gamma$, 记

$$\begin{aligned} \Lambda_s &= \{S_1 > t, X(t) \in A\} \cap \{S_1 > s, X(s) \in A\}^c \\ &= \{S_1 > t, X(t) \in A, S_1 \leq s\} \cup \{S_1 > t, X(t) \in A, X(s) \notin A\}. \end{aligned}$$

于是

$$\begin{aligned} P(\Lambda_s) &\leq P(S_1 > t, X(t) \in A, S_1 \leq s) + P(S_1 > t, X(t) \in A, X(s) \notin A). \end{aligned}$$

由于 F 在 t 连续, 上式第一项

$$\leq P(t \wedge s < S_1 \leq t \vee s) \rightarrow 0 \quad (s \rightarrow t).$$

而第二项

$$\leq P(X(t) \in A, X(s) \notin A),$$

由 A 是开集, 当 $X(t) \in A$, 对充分接近 t 的时刻 s , 同样有 $X(s) \in A$. 这样, $\lim_{s \rightarrow t} P(A_s) = 0$. 交换 t 与 s 的位置, 可以得到当 $s \rightarrow t$ 时,

$$P(\{S_1 > t, X(t) \in A\} \triangle \{S_1 > s, X(s) \in A\}) \rightarrow 0.$$

这可以推出 $K(t, A) - K(s, A) \rightarrow 0$. 故 $K(\cdot, A)$ 黎曼可积. \square

推论 5.6 设更新过程 $X(t)$ 轨道右连左极, A 是开集, 则

$$\lim_{t \rightarrow \infty} P(X(t) \in A) = \mu^{-1} \int_0^{\infty} K(s, A) ds.$$

特别的, 若 $X(\infty)$ 是由

$$P(X(\infty) \in A) = \mu^{-1} \int_0^{\infty} K(s, A) ds$$

确定的随机变量, 则当 $t \rightarrow \infty$ 时, $X(t)$ 分布收敛到 $X(\infty)$. $X(\infty)$ 的 Laplace 变换为

$$E[e^{-\zeta X(\infty)}] = \mu^{-1} E\left[\int_0^{S_1} e^{-\zeta X(s)} ds\right]$$

5.7 其它更新过程

5.7.1 有偿更新过程

设随机变量列 $\{X_n, n \geq 0\}$ 满足 $\{X_n, n \geq 1\}$ 独立同分布, (X_n, Y_n) 与 $\{Y_j, j \neq n\}$ 独立.

定义 5.7 称 $R(t) = \sum_{n=0}^{N(t)-1} X_n$ 为有偿更新过程.

引理 5.5 (Wald 等式) 假设 $Y_n, n \geq 1$ 是一列独立同分布的随机变量, $E(Y_1) = \mu < \infty$, T 是关于 $\mathcal{F}_n = \sigma\{Y_i, i \leq n\}$ 的取非负整数值的停时, 即 $\{T \leq n\} \in \mathcal{F}_n$, $E(T) < \infty$, 则

$$E\left(\sum_{i=1}^T Y_i\right) = E(T)E(Y).$$

证明 令

$$I_n = \begin{cases} 1, & T \geq n, \\ 0, & T < n. \end{cases}$$

则 I_n 是 \mathcal{F}_{n-1} 可测的. 于是

$$\begin{aligned} E\left\{\sum_{i=1}^T Y_i\right\} &= E\left\{\sum_{n=1}^{\infty} Y_n I_n\right\} = \sum_{n=1}^{\infty} E(Y_n I_n) = \sum_{n=1}^{\infty} E(Y_n) E(I_n) \\ &= E(Y) \sum_{n=1}^{\infty} P(T \geq n) = E(Y) E(T). \end{aligned}$$

□

推论 5.7 对于普通更新过程,

$$E\left\{\sum_{i=1}^{N(t)+1} g(Y_i)\right\} = E[N(t) + 1] E[g(Y_i)].$$

定理 5.19 设 $E(|X_n|) < \infty$, $\mu = E(Y_1) \in (0, \infty)$. 则

$$(i) \lim_{t \rightarrow \infty} \frac{R(t)}{t} = \frac{E(X_1)}{\mu} \text{ a.s..}$$

$$(ii) \lim_{t \rightarrow \infty} \frac{E(R(t))}{t} = \frac{E(X_1)}{\mu}.$$

证明

(i) 当 $t \rightarrow \infty$ 时,

$$\frac{R(t)}{t} = \frac{\sum_{i=0}^{N(t)-1} X_i}{N(t) - 1} \cdot \frac{N(t) - 1}{t} \rightarrow \frac{EX_1}{\mu}.$$

(ii)

$$\begin{aligned} E(R(t)) &= E(R(t)1_{\{N(t)=0\}}) + E(R(t)1_{\{N(t)>0\}}) = E\left(\sum_{n=0}^{N(t)-1} X_n 1_{\{N(t)>0\}}\right) \\ &= E\left(\sum_{n=0}^{N(t)} X_n 1_{\{N(t)>0\}}\right) - E(X_{N(t)} 1_{\{N(t)>0\}}) = A + B. \end{aligned}$$

根据 Wald 等式, 当 $t \rightarrow \infty$ 时,

$$\begin{aligned}
A &= E\left(\sum_{n=0}^{N(t)} X_n 1_{\{N(t)>0\}}\right) = E(X_0 1_{\{N(t)>0\}}) + E\left(\sum_{n=1}^{\infty} X_n 1_{\{n \leq N(t)\}}\right) \\
&= E(X_0 1_{\{N(t)>0\}}) + E\left(\sum_{n=1}^{\infty} X_n 1_{\{n-1 < N(t)\}}\right) \\
&= E(X_0 1_{\{N(t)>0\}}) + E\left(\sum_{n=1}^{\infty} X_n 1_{\{S_{n-1} \leq t\}}\right) \\
&= E(X_0 1_{\{N(t)>0\}}) + E(X_1) E\left(\sum_{n=1}^{\infty} 1_{\{S_{n-1} \leq t\}}\right) \\
&= E(X_0 1_{\{N(t)>0\}}) + E(X_1) V(t) \\
&\sim t E(X_1) / \mu.
\end{aligned}$$

而对 B , 我们有

$$\begin{aligned}
B &= E(X_{N(t)} 1_{\{N(t)>0\}}) = \sum_{n=1}^{\infty} E(X_n 1_{\{N(t)=n\}}) = \sum_{n=1}^{\infty} E(X_n 1_{\{S_{n-1} \leq t < S_n\}}) \\
&= \sum_{n=1}^{\infty} \int_0^t E(X_n 1_{\{x \leq t < x+Y_n\}}) P(S_{n-1} \in dx) = \int_0^t E(X_1 1_{\{t < x+Y_1\}}) V(dx).
\end{aligned}$$

令

$$z(u) = E(|X_1| 1_{\{u < Y_1\}}),$$

则当 $u \rightarrow \infty$ 时, $z(u) \rightarrow 0$, 且 $z(u)$ 递减, $z(u) \leq E(|X_1|)$. 这样,

$$|B| \leq V * z(t).$$

对任意的 $\varepsilon > 0$, 有 $M > 0$ 使当 $u > M$ 时, $z(u) < \varepsilon$. 于是, 对 $t \geq M$,

$$\begin{aligned}
V * z(t) &= \int_0^t z(t-x) V(dx) \\
&= \int_0^{t-M} z(t-x) V(dx) + \int_{t-M}^t z(t-x) V(dx) \\
&\leq \varepsilon V(t-M) + z(0)(V(t) - V(t-M)).
\end{aligned}$$

由更新定理,

$$\limsup_{t \rightarrow \infty} \frac{V * z(t)}{t} \leq \frac{\varepsilon}{\mu}.$$

由 ε 的任意性, 可知当 $t \rightarrow \infty$ 时, $B/t \rightarrow 0$. 这样,

$$\lim_{t \rightarrow \infty} \frac{E(R(t))}{t} = \frac{E(X_1)}{\mu}.$$

□

定理 5.20 设 $F(x)$ 非算术分布,

$$E(|X_1|) < \infty, E(Y_1) < \infty, E(|X_1 Y_1|) < \infty,$$

则对任意 $a > 0$, 有

$$\lim_{t \rightarrow \infty} \{E(R(t+a)) - E(R(t))\} = \frac{aE(X_1)}{E(Y_1)}.$$

证明 根据 Wald 等式, 我们有

$$\begin{aligned} & E(R(t+a)) - E(R(t)) \\ &= E\left(\sum_{i=0}^{N(t+a)} X_i\right) - E(X_{N(t+a)}) - E\left(\sum_{i=0}^{N(t)} X_i\right) + E(X_{N(t)+1}) \\ &= E(X_1)(E[N(t+a)] - E[N(t)]) - E(X_{N(t+a)}) + E(X_{N(t)}). \end{aligned}$$

由更新定理, 我们只要证明 $\lim_{t \rightarrow \infty} E(X_{N(t)})$ 存在且有限. 不妨设 $X_i \geq 0$, 由

$$E(X_{N(t)}) = \int_0^t E(X_1 1_{\{t-x < Y_1\}}) dU(x),$$

$E(X_1 1_{\{t < Y_1\}})$ 是单调下降函数, 且

$$\int_0^\infty E(X_1 1_{\{t < Y_1\}}) dt = E(X_1 \int_0^\infty 1_{\{t < Y_1\}} dt) = E(X_1 Y_1) < \infty.$$

所以 $E(X_1 1_{\{t < Y_1\}})$ 直接黎曼可积. 于是

$$\lim_{t \rightarrow \infty} E(X_{N(t)}) = \int_0^\infty E(X_1 1_{\{t < Y_1\}}) dt = E(X_1 Y_1) < \infty.$$

□

例 5.12 考虑古典风险过程

$$R_n = u + c \sum_{i=1}^n T_i - \sum_{i=1}^n Z_i.$$

设 $\{T_i\}$ 独立同分布, 且均服从参数为 α 的指数分布, 设 $\{Z_i\}$ 独立同分布, 分布函数为 F , 且与 $\{T_i\}$ 独立. 记 $\Phi(u)$ 为不破产的概率. 则

$$\begin{aligned}
\Phi(u) &= P(R_n \geq 0, n \geq 0) \\
&= P(u + cT_1 - Z_1 \geq 0, u + cT_1 + Z_1 + \sum_{i=2}^n (cT_i - Z_i) \geq 0, n \geq 1) \\
&= \int_0^\infty ds \int_0^{u+cs} \alpha e^{-\alpha s} P(u + cs - z + \sum_{i=2}^n (cT_i - Z_i) \geq 0, n \geq 1) dF(z) \\
&= \int_0^\infty ds \int_0^{u+cs} \alpha e^{-\alpha s} \Phi(u + cs - z) dF(z) \\
&= \int_u^\infty dx \int_0^x \frac{\alpha}{c} e^{-\alpha \frac{x-u}{c}} \Phi(x - z) dF(z) \\
&= \frac{\alpha}{c} e^{\frac{\alpha u}{c}} \int_u^\infty e^{-\frac{\alpha x}{c}} dx \int_0^x \Phi(x - z) dF(z).
\end{aligned}$$

这样,

$$\begin{aligned}
\Phi'(u) &= \frac{\alpha}{c} \Phi(u) - \frac{\alpha}{c} \int_0^u \Phi(u - z) dF(z), \\
\Phi(t) - \Phi(0) &= \frac{\alpha}{c} \int_0^t \Phi(u) du - \frac{\alpha}{c} \int_0^t du \int_0^u \Phi(u - z) dF(z) \\
&= \frac{\alpha}{c} \int_0^t \Phi(u) du - \frac{\alpha}{c} \int_0^t dF(z) \int_z^t \Phi(u - z) du \\
&= \frac{\alpha}{c} \int_0^t \Phi(u) du + \frac{\alpha}{c} \int_0^t d(1 - F(z)) \int_0^{t-z} \Phi(u) du \\
&= \frac{\alpha}{c} \int_0^t \Phi(t - z)(1 - F(z)) dz.
\end{aligned}$$

因此

$$\Phi(u) = \Phi(0) + \frac{\alpha}{c} \int_0^u \Phi(u - z)(1 - F(z)) dz,$$

即

$$\Phi(u) = \Phi(0) + \frac{\alpha\mu}{c} \int_0^u \Phi(u - z) dF_e(z) \quad (5.18)$$

假设 $\frac{\alpha\mu}{c} < 1$, 即安全负荷条件成立, 且存在 $\beta > 0$, 使得

$$\frac{\alpha}{c} \int_0^\infty e^{\beta z}(1 - F(z)) dz = 1.$$

于是,

$$\lim_{t \rightarrow \infty} e^{\beta t} (\Phi(\infty) - \Phi(t)) = \frac{\Phi(0)\beta^{-1}}{\int_0^\infty s e^{\beta s} \frac{\alpha}{c} (1 - F(s)) ds}.$$

下面只需知道 $\Phi(\infty)$ 和 $\Phi(0)$. 因为

$$\frac{ct - \sum_{i=1}^{N(t)} X_i}{t} = c - \alpha\mu > 0,$$

所以

$$P(\inf_{t \geq 0} (ct - \sum_{i=1}^{N(t)} X_i) > -\infty) = 1.$$

从而

$$\lim_{u \rightarrow \infty} P(\inf_{t \geq 0} (ct - \sum_{i=1}^{N(t)} X_i) > -u) = \lim_{u \rightarrow \infty} \Phi(u) = 1,$$

即 $\Phi(\infty) = 1$. 从而由 (5.18) 知 $\Phi(0) = 1 - \frac{\alpha\mu}{c}$.

5.7.2 可终止更新过程

定义 5.8 设 $L = F(\infty) = P(Y_1 < \infty) < 1$, 则称相应的更新过程为可终止的更新过程, $N(\infty)$ 为更新总数,

$$\zeta = S_{N(\infty)} = \sup\{S_n : S_n < \infty\}$$

为最后更新时间.

定理 5.21

$$(i) P(N(\infty) = k) = (1 - L)L^k, \quad k \geq 0.$$

$$(ii) EN(\infty) = L/(1 - L).$$

$$(iii) P(\zeta \leq t) = (1 - L)(1 + EN(t)).$$

$$(iv) E\zeta = \frac{1}{1-L} \int_0^\infty (L - F(s)) ds.$$

证明 只要注意到

$$P(N(\infty) = k) = P(Y_1 < \infty, \dots, Y_k < \infty, Y_{k+1} = \infty).$$

□

6 平稳过程简介

在这一章中, 我们简单介绍弱平稳过程的谱分解与强平稳过程的遍历理论. 我们将略去绝大部分结果的证明, 如果想进行系统的学习, 请查阅相关的文献.

设 $x_t = \xi_t + i\eta_t$, $t \in T$ 是概率空间 (Ω, \mathcal{F}, P) 上复值随机过程, 其中 ξ_t, η_t 都是实值的.

定义 6.1 称 x_t 为强平稳过程, 若对任意的正整数 $n, t_1, t_2, \dots, t_n \in T$ 和 $s \in T$, 有

$$(x_{t_1}, x_{t_2}, \dots, x_{t_n}) \stackrel{d}{=} (x_{t_1+s}, x_{t_2+s}, \dots, x_{t_n+s}).$$

设 x_t 是强平稳过程, 如果对某个 t_0 , $E x_{t_0}$ (或 $E(|x_{t_0}|^2)$) 存在, 则对任意的 $t \in T$, $E(x_t)$ (或 $E(|x_t|^2)$) 都存在. 并且, x_t 的均值函数

$$m(t) = E(x_t) = E(x_{t+s}) = m(t+s).$$

故 $m(t)$ 为常数, 记 $m(t) = m$. 又 x_t 的协方差函数

$$E((x_s - m)\overline{(x_{s+t} - m)}) = E(x_s \overline{x_{s+t}}) - m^2 = E(x_0 \overline{x_t}) - m^2,$$

故 $E((x_s - m)\overline{(x_{s+t} - m)})$ (或等价的 $E(x_s \overline{x_t})$) 与 s 无关.

定义 6.2 称 x_t 为弱平稳过程, 若它满足

- (i) 对一切 $t \in T$, $E(|x_t|^2) < \infty$;
- (ii) 对一切 $t \in T$, $E(x_t) = m$ 为常数;
- (iii) 对任意的 $t, t + \tau \in T$, 相关函数

$$B(\tau) = E(x_t \overline{x_{t+\tau}})$$

与 t 无关.

在本章中, 如果不加说明, 我们将弱平稳过程简称为平稳过程.

例 6.1 (白噪声过程) 设 $\{x_n, n = 0, \pm 1, \pm 2, \dots\}$ 为互不相关的实值随机变量序列, $E(x_n) = 0, 0 < E(|x_n|^2) = \sigma^2 < \infty$, 则 x_n 是平稳过程, 我们称它为白噪声过程.

例 6.2 (滑动平均过程) 设

$$\{x_n, n = 0, \pm 1, \pm 2, \dots\}$$

是互不相关的实值随机变量序列, $E(x_n) = 0, 0 < D(x_n) = \sigma^2 < \infty$. 固定正整数 M , 定义

$$y_n = \sum_{k=0}^M a_k x_{n-k}, \quad n = 0, \pm 1, \dots,$$

其中 a_k 为实数列. 则 y_n 是平稳过程. 事实上,

$$E(y_n) = 0,$$

$$E(y_{n+\tau} y_n) = \begin{cases} \sigma^2(a_0 a_\tau + a_1 a_{\tau+1} + \dots + a_{M-\tau} a_M), & \tau \leq M, \\ 0, & \tau > M. \end{cases}$$

当 $\sum_{k=0}^M a_k = 1$ 时, 称 y_n 为以 $\{a_0, a_1, \dots, a_M\}$ 为权, $M+1$ 为步长的滑动平均.

例 6.3 (三角多项式过程) 设 $\{A_n, 1 \leq n \leq N\}$ 和 $\{B_n, 1 \leq n \leq N\}$ 为实值随机变量序列, 满足

$$E(A_n) = E(B_m) = E(A_n B_m) = 0,$$

$$E(A_n A_m) = E(B_n B_m) = \sigma_n^2 \delta_{nm}.$$

设 $\omega_1, \omega_2, \dots, \omega_N$ 为正实数, 定义

$$x_t = \sum_{k=1}^N (A_k \cos(\omega_k t) + B_k \sin(\omega_k t)), \quad -\infty < t < \infty,$$

则 x_t 是平稳过程.

6.1 谱分解

6.1.1 相关函数的谱分解

称复函数 $z(t)$ 正定, 若对任意的 n , 复数 $\xi_1, \xi_2, \dots, \xi_n$ 及 $t_1, t_2, \dots, t_n \in T$, 都有

$$\sum_{j, k=1}^n z(t_j - t_k) \xi_j \bar{\xi}_k \geq 0.$$

命题 6.1 设 x_t 是平稳过程, 则 $B(\tau)$ 正定, 特别的 $B(0) \geq 0$, 且对任意的 $\tau \in T$, $|B(\tau)| \leq B(0)$, $B(\tau) = \overline{B(-\tau)}$.

证明 对任意的 $\tau_1, \tau_2, \dots, \tau_n \in T$ 及复数 $\xi_1, \xi_2, \dots, \xi_n$,

$$\begin{aligned} \sum_{j, k=1}^n B(\tau_j - \tau_k) \xi_j \bar{\xi}_k &= \sum_{j, k=1}^n \xi_j \bar{\xi}_k E[x_t \overline{x_{t+\tau_j - \tau_k}}] \\ &= \sum_{j, k=1}^n \xi_j \bar{\xi}_k E[x_{t+\tau_k} \overline{x_{t+\tau_j}}] = E \left[\left| \sum_{j=1}^n \bar{\xi}_j x_{t+\tau_j} \right|^2 \right]. \end{aligned}$$

故 $B(\tau)$ 正定. □

下面我们设平稳过程 x_t 均方连续, 即对任意的 $\tau \in T$, 当 $t \rightarrow \tau$ 时, 在 L^2 意义下, $x_t \rightarrow x_\tau$. 容易证明, x_t 均方连续当且仅当 $B(t)$ 是连续的. 又 $B(t)$ 正定, 如果我们假设 $B(0) = 1$, 应用 Bochner 定理, $B(t)$ 为一特征函数, 即存在左连续分布函数 F , 使得

$$B(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda).$$

事实上反过来的结果也是对的. 因此我们有

定理 6.1 为使复值函数 $B(t)$, $t \in T$ 成为均值为 0, 方差为 1 的均方连续平稳过程 $\{x_t, t \in T\}$ 的相关函数, 其充分必要条件是 $B(t)$ 可以表示成

$$B(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda), \quad (6.1)$$

其中 F 是左连续分布函数. 并且, 不计常数之差, F 由 $B(t)$ 唯一决定. 我们称 $F(\lambda)$ 为 x_t 的谱函数.

注 6.1 如果 $T = \{0, \pm 1, \dots\}$, 则 (6.1) 可以改写为

$$B(n) = \int_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda), \quad (6.2)$$

其中 $F(\lambda)$ 是 $[-\pi, \pi]$ 上左连续分布函数.

注 6.2 特别的, 若 $F(\lambda)$ 绝对连续, 即有 $f(t)$, 使得

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

则 (6.1) 可写为

$$B(t) = \int_{-\infty}^{\infty} e^{i\lambda t} f(\lambda) d\lambda. \quad (6.3)$$

我们称 $f(\lambda)$ 为 x_t 的谱密度. 由 Fourier 逆变换公式,

$$f(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda t} B(t) dt.$$

事实上, 若 $B(t)$ 在 $(-\infty, \infty)$ 上可积, 则谱密度 $f(\lambda)$ 是存在的.

注 6.3 定理中条件 $E(x_0) = 0, E(|x_t|^2) = 1$ 不是本质的. 因为我们可以对 x_t 标准化, 即令

$$y_t = \frac{x_t - m}{\sigma},$$

其中 $m = E(x_t), \sigma^2 = E(|x_t - m|^2)$.

当 x_t 为实值过程时, 我们有

定理 6.2 实值函数 $B(t)$ 具有均值为 0, 方差为 1 的均方连续平稳过程 x_t 的相关函数, 其充分必要条件是 $B(t)$ 可以表示成

$$B(t) = \int_{-\infty}^{\infty} \cos(\lambda t) dF(\lambda). \quad (6.4)$$

当 $T = \{0, \pm 1, \dots\}$ 时, (6.4) 可写成

$$B(n) = \int_{-\pi}^{\pi} \cos(\lambda n) dF(\lambda). \quad (6.5)$$

特别的, 若 $F(\lambda)$ 在 $\lambda = 0$ 连续, 则 (6.4), (6.5) 可改写成

$$\begin{aligned} B(t) &= \int_0^{\infty} \cos(\lambda t) dG(\lambda), \\ B(n) &= \int_0^{\infty} \cos(\lambda n) dG(\lambda), \end{aligned}$$

其中 $G(\lambda) = 2F(\lambda), \lambda \geq 0$ 为有界递减左连续函数.

例 6.4 设 $\sigma = 1$, 考虑例 6.1 中白噪声 x_n . 由于 $B(0) = 1, B(n) = 0, n \neq 0$, 故 $\sum_{n=-\infty}^{\infty} |B(n)| < \infty$, 从而谱密度 $f(\lambda)$ 存在, 且

$$f(\lambda) = \frac{1}{2\pi}, \quad -\pi \leq \lambda \leq \pi.$$

谱函数 $F(\lambda)$ 可取为

$$F(\lambda) = \frac{1}{2\pi}(\lambda + \pi), \quad -\pi \leq \lambda \leq \pi.$$

例 6.5 设 $\{y_n, n = 0, \pm 1, \dots\}$ 为互不相关的实值随机变量列, $E(y_n) = 0, E(|y_n|^2) = 1$. 考虑滑动平均过程

$$x_n = \sum_{k=0}^{\infty} a^k y_{n-k}, \quad n = 0, \pm 1, \dots,$$

其中 a 为实数, $|a| < 1$. 则 x_t 为平稳过程, 且

$$B(n) = \sum_{k \geq n} a^k a^{k-n} = \frac{a^n}{1-a^2}, \quad n \geq 0.$$

这样, 令 $A = 1/(1-a^2) > 0$, x_n 的相关函数

$$B(n) = Aa^{|n|}, \quad n = 0, \pm 1, \dots$$

注意到 $\sum_k |B(k)| < \infty$, 于是

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} B(k) e^{-i\lambda k} \\ &= \frac{1}{2\pi(1-2a \cos \lambda + a^2)}, \quad -\pi \leq \lambda \leq \pi. \end{aligned}$$

例 6.6 设 $\lambda > 0$, A 和 B 不相关, 且均值为 0, 方差为 1. 考虑平稳过程

$$x_t = A \cos(\lambda t) + B \sin(\lambda t), \quad -\infty < t < \infty.$$

它的相关函数为

$$B(t) = \cos(\lambda t).$$

于是谱函数可取为

$$F(u) = \begin{cases} 0, & u \leq -\lambda, \\ \frac{1}{2}, & -\lambda < u \leq \lambda, \\ 1, & u \geq \lambda. \end{cases}$$

6.1.2 均方积分

设 $x(t)$, $t \in T$ 是二阶矩过程, 即对任意的 t , $x(t)$ 平方可积.

定义 6.3 称二阶矩过程 $x(t)$, $t \in T$ 为正交增量过程, 若对任意的 $t_1 < t_2 \leq t_3 < t_4$, 有

$$E[(x_{t_2} - x_{t_1})(x_{t_4} - x_{t_3})] = 0.$$

下面我们假设 $T = (-\infty, \infty)$, $x(t)$ 是正交增量过程, 且均方左连续, 即对任意的 t ,

$$\lim_{s \uparrow t} E[|x_s - x_t|^2] = 0.$$

注意到存在一个单调不减的左连续实值函数 $F(t)$, 使得对任意的 $s \leq t$, 有

$$E[|x_t - x_s|^2] = F(t) - F(s).$$

且如果不计常数之差, $F(t)$ 是由 $x(t)$ 性唯一决定的. 事实上, 我们可以取

$$F(t) = \begin{cases} E[|x_t - x_{t_0}|^2], & t \geq t_0; \\ -E[|x_t - x_{t_0}|^2], & t < t_0, \end{cases}$$

其中 t_0 是固定的常数.

记关于 $F(t)$ 平方可积的 Borel 可测函数全体为 $L^2(dF)$, 即

$$L^2(dF) = \{f(t) : \int_{-\infty}^{\infty} |f(t)|^2 dF(t) < \infty\}.$$

下面我们对 $f \in L^2(dF)$ 定义随机积分 $\int_{-\infty}^{\infty} f(t) dx(t)$.

先设 $f(t)$ 是阶梯函数, 即存在常数 $c > 0$, 及 $[-c, c]$ 的分割 $-c \leq t_0 < t_1 < \cdots < t_n = c$, 复数 a_i , 使

$$f(t) = \begin{cases} 0, & t < t_0; \\ a_i, & t_i \leq t < t_{i+1}, \quad 0 \leq i \leq n-1; \\ 0, & t \geq t_n. \end{cases}$$

这样, 我们可以定义随机积分

$$\int_{-\infty}^{\infty} f(t) dx(t) = \sum_{j=0}^{n-1} a_j (x_{t_{j+1}} - x_{t_j}).$$

这样定义出来的积分与 $f(t)$ 的表示形式无关, 且满足

命题 6.2 对任意的阶梯函数 $f(t), g(t)$, 常数 α, β , 有

(i)

$$\int_{-\infty}^{\infty} [\alpha f(t) + \beta g(t)] dx(t) = \alpha \int_{-\infty}^{\infty} f(t) dx(t) + \beta \int_{-\infty}^{\infty} g(t) dx(t).$$

(ii)

$$E \left[\int_{-\infty}^{\infty} f(t) dx(t) \cdot \overline{\int_{-\infty}^{\infty} g(t) dx(t)} \right] = \int_{-\infty}^{\infty} f(t) \overline{g(t)} dF(t).$$

下面设 $f(t) \in L^2(dF)$, 则有一列阶梯函数 $f_n(t)$ 使得

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} |f_n(t) - f(t)|^2 dF(t) = 0.$$

令

$$J_n = \int_{-\infty}^{\infty} f_n(t) dx(t),$$

由命题 6.2,

$$E[|J_n - J_m|^2] = \int_{-\infty}^{\infty} |f_n(t) - f_m(t)|^2 dF(t) \rightarrow 0 \quad (n, m \rightarrow \infty).$$

这样, 存在唯一的平方可积随机变量 J , 使得

$$E[|J_n - J|^2] \rightarrow 0 \quad (n \rightarrow \infty).$$

并且, J 的值与阶梯函数列 $f_n(t)$ 的选取无关, 因此, 我们把 J 定义为 $f(t)$ 关于 $x(t)$ 的随机积分, 记为

$$J = \int_{-\infty}^{\infty} f(t) dx(t).$$

容易看出, 命题 6.2 中结论对 $f(t), g(t) \in L^2(dF)$ 仍旧成立.

对于 $-\infty < a < b < \infty$, 通常我们定义

$$\int_a^b f(t) dx(t) = \int_{-\infty}^{\infty} f(t) 1_{[a, b]}(t) dx(t).$$

6.1.3 过程的谱分解

前面我们已经提到了平稳过程相关函数的谱分解, 这一节我们介绍平稳过程的谱分解.

定理 6.3 设 $x_t, t \in T$ 是平稳过程, 则存在均方左连续正交增量过程 $y(\lambda)$, 使得

$$x_t = \int_{-\infty}^{\infty} e^{i\lambda t} dy(\lambda). \quad (6.6)$$

并且, 不计常数之差, $y(\lambda)$ 由 x_t 唯一确定. 进一步的, $y(\lambda)$ 满足

(i) 对一切 $\lambda, E[y(\lambda)] = 0$;

(ii) 对任意的 $\lambda_1 < \lambda_2$,

$$E[|y(\lambda_2) - y(\lambda_1)|^2] = F(\lambda_2) - F(\lambda_1),$$

其中 $F(\lambda)$ 是 x_t 的相关函数.

如果 $T = \{0, \pm 1, \dots\}$, 则 (6.6) 可化为

$$x_n = \int_{-\pi}^{\pi} e^{in\lambda} dy(\lambda). \quad (6.7)$$

我们称 $y(\lambda)$ 为 x_t 的随机谱函数.

定理 6.4 设 x_t 是实值均方连续平稳过程, $E(x_t) = 0$, $E(x_t^2) = 1$. 则 x_t 可表示成

$$x_t = \hat{\eta}_0 + \int_{0+}^{\infty} \cos(\lambda t) dy(\lambda) + \int_0^{\infty} \sin(\lambda t) dz(\lambda). \quad (6.8)$$

对 $T = \{0, \pm 1, \dots\}$,

$$x_n = \hat{\eta}_0 + \int_{0+}^{\pi} \cos(\lambda n) dy(\lambda) + \int_0^{\pi} \sin(\lambda n) dz(\lambda). \quad (6.9)$$

进一步的,

(i) 设 $F(\lambda)$ 为 x_t 的谱函数, 则 $\hat{\eta}_0$ 是与 t 无关的实值随机变量, 且 $E[|\hat{\eta}_0|^2] = F(0+) - F(0)$;

(ii) $y(\lambda), z(\lambda)$ 为均值为 0 的实值均方左连续正交增量过程, 且不计常数之差, 由 x_t 唯一确定;

(iii) 对任意的 $0 \leq \lambda_1 < \lambda_2 \leq \lambda_3 < \lambda_4$,

$$\begin{aligned} E[(y(\lambda_2) - y(\lambda_1))(y(\lambda_4) - y(\lambda_3))] &= E[(z(\lambda_2) - z(\lambda_1))(z(\lambda_4) - z(\lambda_3))] \\ &= 2(F(\lambda_2) - F(\lambda_3)); \end{aligned}$$

(iv) 对任意的 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$,

$$E[(y(\lambda_2) - y(\lambda_1))(z(\lambda_4) - z(\lambda_3))] = 0.$$

注 6.4 当 x_t 的谱函数 $F(\lambda)$ 在 0 点连续时, $\hat{\eta}_0 = 0$. x_t 可以表为

$$x_t = \int_0^{\infty} \cos(\lambda t) dy(\lambda) + \int_0^{\infty} \sin(\lambda t) dz(\lambda).$$

如果 x_t 有谱密度, 则上述条件成立.

6.2 均方遍历性

若 $x_n, n \geq 1$ 独立同分布, 且 $E(|x_n|) < \infty$, 则强大数定律

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n x_k = m, \text{ a.s.},$$

成立, 其中 $m = E(x_n)$. 对于平稳过程, 我们可以利用谱分解, 证明在均方意义下的收敛性.

定理 6.5 设 $x_t, -\infty < t < \infty$ 是均方连续平稳过程, 记

$$m = E(x_t), K(\tau) = E[(x_{t+\tau} - m)\overline{(x_t - m)}].$$

则存在随机变量 η , 使在 L^2 意义下,

$$\frac{1}{2t} \int_{-t}^t (x_s - m) ds \rightarrow \eta \quad (t \rightarrow \infty), \quad (6.10)$$

且

$$E[|\eta|^2] = \lim_{t \rightarrow \infty} \frac{1}{2t} \int_{-t}^t K(\tau) d\tau.$$

证明 不妨设 $m = 0, K(\tau) = B(\tau)$. 设 $F(\lambda)$ 是 x_t 的谱函数, x_t 的谱分解为

$$x_s = \int_{-\infty}^{\infty} e^{i\lambda s} dy(\lambda),$$

$$E[|y(\lambda_2) - y(\lambda_1)|^2] = F(\lambda_2) - F(\lambda_1), \quad \lambda_2 > \lambda_1.$$

这样,

$$\frac{1}{2t} \int_{-t}^t x_s ds = \int_{-\infty}^{\infty} \left[\frac{1}{2t} \int_{-t}^t e^{i\lambda t} dt \right] dy(\lambda) = \int_{-\infty}^{\infty} \phi_t(\lambda) dy(\lambda),$$

其中

$$\phi_t(\lambda) = \begin{cases} \frac{\sin(\lambda t)}{\lambda t}, & \lambda \neq 0, \\ 1, & \lambda = 0. \end{cases}$$

令 $\eta = y(0+) - y(0) = \int_{-\infty}^{\infty} 1_{\{0\}}(\lambda) dy(\lambda)$. 则

$$\begin{aligned} & E \left[\left| \frac{1}{2t} \int_{-t}^t x_s ds - \int_{-\infty}^{\infty} 1_{\{0\}}(\lambda) dy(\lambda) \right|^2 \right] \\ &= E \left[\left| \int_{-\infty}^{\infty} [\phi_t(\lambda) - 1_{\{0\}}(\lambda)] dy(\lambda) \right|^2 \right] \\ &= \int_{-\infty}^{\infty} [\phi_t(\lambda) - 1_{\{0\}}(\lambda)]^2 dF(\lambda) \rightarrow 0 \quad (t \rightarrow \infty), \end{aligned}$$

故在 L^2 意义下,

$$\frac{1}{2t} \int_{-t}^t x_s ds \rightarrow \eta \quad (t \rightarrow \infty).$$

类似的, 应用 $B(\tau)$ 的谱分解,

$$\lim_{t \rightarrow \infty} \frac{1}{2t} \int_{-t}^t B(\tau) d\tau = F(0+) - F(0).$$

而

$$\begin{aligned} E[|\eta|^2] &= \lim_{t \rightarrow \infty} E \left[\left| \frac{1}{2t} \int_{-t}^t x_s ds \right|^2 \right] \\ &= \lim_{t \rightarrow \infty} \int_{-\infty}^{\infty} \phi_t^2(\lambda) dF(\lambda) = F(0+) - F(0). \end{aligned}$$

定理证毕. □

推论 6.1 延用上定理的记号,

$$\lim_{t \rightarrow \infty} E \left[\left| \frac{1}{2t} \int_{-t}^t x_s ds - m \right|^2 \right] = 0$$

的充分必要条件是 $F(\lambda)$ 在 0 点连续, 或者等价的,

$$\lim_{t \rightarrow \infty} \frac{1}{2t} \int_{-t}^t K(\tau) d\tau = 0.$$

作为应用, 我们考虑如何估计平稳序列 $x_n, n = 0, \pm 1, \dots$ 的相关函数 $B(\tau)$. 不失一般性, 设 $E(x_n) = 0$. 另外, 我们假设 $E(|x_n|^4) < \infty$. 固定 τ , 考虑新的平稳序列

$$\omega_k = x_{k+\tau} \bar{x}_k, \quad k = 0, \pm 1, \dots$$

记相关函数为 $R(m) = E[\omega_{k+m} \bar{\omega}_k]$.

定理 6.6

$$\lim_{n \rightarrow \infty} E \left[\left| \frac{1}{n+1} \sum_{k=0}^n x_{k+\tau} \bar{x}_k - B(\tau) \right|^2 \right] = 0$$

的充分必要条件是

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{m=0}^n R(m) = |B(\tau)|^2. \quad (6.11)$$

注 6.5 上面定理中条件 (6.11) 需要对 x_n 的四阶矩作限制, 这在应用中非常不方便. 但对于正态实值平稳过程情形, 我们只需要对二阶矩作要求, 即将条件 (6.11) 换成

$$\lim_{n \rightarrow \infty} \frac{1}{n+1} \sum_{k=0}^n B(k)^2 = 0.$$

6.3 平稳序列的线性预测

设 $x_n, n = 0, \pm 1, \dots$ 是平稳序列, $E(x_n) = 0$. 在实际应用中, 我们经常需要从已有的数据来预测未来的取值, 即已知 x_{-N}, \dots, x_0 , 求一个与 $\{x_k, -N \leq k \leq 0\}$ 相关的函数 \hat{x}_n ($n > 0$), 使得 \hat{x}_n 与 x_n 在一定意义下误差最小. 在平稳序列情形, 我们经常用均方误差 $E[|\hat{x}_n - x_n|^2]$ 作为衡量预报最优程度的指标.

定义 6.4 设 $\mathcal{F}(N)$ 为关于 $\sigma(x_k, -N \leq k \leq 0)$ 可测的平方可积随机变量全体. 固定 $n > 0$, 称 \hat{x}_n 为 $\{x_k, -N \leq k \leq 0\}$ 对 x_n 的最优预报值, 若

$$E[|\hat{x}_n - x_n|^2] = \inf_{g \in \mathcal{F}(N)} E[|x_n - g|^2].$$

容易看出, 最优预报值是唯一的. 事实上, 若 x_n^* 也是最优预报值, 则 $E[|\hat{x}_n - x_n^*|^2] = 0$.

定理 6.7 $\{x_k, -N \leq k \leq 0\}$ 对 x_n 的最优预报值

$$\hat{x}_n = E[x_n | x_{-N}, \dots, x_0].$$

事实上, 上面定理对一般的二阶矩有限序列都是正确的. 这也可以作为条件期望的一个直观解释.

由于 $E[x_n | x_{-N}, \dots, x_0]$ 的具体表达式很难求出, 在实际应用中, 上述定理意义并不大. 下面我们只考虑 x_{-N}, \dots, x_0 的线性函数.

令 $\mathcal{H}(N)$ 为 x_{-N}, \dots, x_0 的实系数线性组合全体构成的线性空间. 若 $\hat{x}_n \in \mathcal{H}(N)$ 满足

$$E[|x_n - \hat{x}_n|^2] = \inf_{Y \in \mathcal{H}(N)} E[|x_n - Y|^2],$$

则称 \hat{x}_n 为 $\{x_k, -N \leq k \leq 0\}$ 对 x_n 的最优线性预报值.

利用 Hilbert 空间的性质, 容易看出

定理 6.8 设 x 平方可积.

(i) 存在唯一的 $\hat{x} \in \mathcal{H}(N)$, 使得

$$E[|x - \hat{x}|^2] = \inf_{Y \in \mathcal{H}(N)} E[|x - Y|^2]; \quad (6.12)$$

(ii) $\hat{x} \in \mathcal{H}(N)$ 满足 (6.12) 的充分必要条件是任意的 $y \in \mathcal{H}(N)$,

$$E[(x - \hat{x})y] = 0.$$

一般来说, 最优预报值不一定在 $\mathcal{H}(N)$ 中. 但应用下面定理, 对于正态平稳序列, 最优预报值 $E[x_n | x_{-N}, \dots, x_0] \in \mathcal{H}(N)$.

定理 6.9 设 (ξ_1, \dots, ξ_n) 是 n 维正态向量, $E(\xi_k) = 0$, 则存在实数 c_1, \dots, c_{n-1} , 使得

$$E[\xi_n | \xi_1, \dots, \xi_{n-1}] = \sum_{k=1}^{n-1} c_k \xi_k.$$

证明 应用 Schmidt 正交化方法即得. □

6.4 强平稳过程与遍历理论

定义 6.5 设 (S, \mathcal{S}) 是可测空间, 给定其上测度 μ 及可测变换 T . 称 T 是保测变换, 若 $\mu \circ T^{-1} = \mu$. 这里 $\mu \circ T^{-1}$ 是 (S, \mathcal{S}) 上测度, 定义为 $(\mu \circ T^{-1})(B) = \mu(T^{-1}(B))$, $B \in \mathcal{S}$.

这样, 若 ξ 是取值于 S 的随机元, μ 是 ξ 的分布, 则 T 是保测变换当且仅当 $T\xi$ 与 ξ 同分布 (记为 $T\xi \stackrel{d}{=} \xi$).

设 θ 是 S^∞ 上推移算子, 即对任意的 $x_k \in S$, $k \geq 0$, 有 $\theta(x_0, x_1, x_2, \dots) = (x_1, x_2, \dots)$. 随机元 $\xi = (\xi_0, \xi_1, \dots)$ 为强平稳序列的充分必要条件是 $\theta\xi \stackrel{d}{=} \xi$. 下面结果描述了强平稳序列与保测变换的联系.

引理 6.1 对 S 上随机元 ξ 及保测变换 T , $T\xi \stackrel{d}{=} \xi$ 当且仅当 $(T^n\xi)_{n \geq 0}$ 是强平稳序列. 此时, 对任意的可测函数 f , $(f(T^n\xi))$ 也是强平稳的. 反过来, 任意强平稳序列都具有这种表示.

证明 设 $T\xi \stackrel{d}{=} \xi$, 则

$$\theta(f \circ T^n\xi) = (f \circ T^{n+1}\xi) = (f \circ T^n T\xi) \stackrel{d}{=} (f \circ T^n\xi),$$

即 $(f \circ T^n\xi)$ 是强平稳的. 反过来, 如果 $\eta = (\eta_0, \eta_1, \dots)$ 强平稳, 则 $\theta\eta \stackrel{d}{=} \eta$. 令 $\pi_0(x_0, x_1, \dots) = x_0$, 则 $\eta_n = \pi_0(\theta^n\eta)$. □

给定测度空间 (S, \mathcal{S}, μ) 上可测变换 T . 称集合 $I \in \mathcal{S}$ 为不变的, 若 $T^{-1}(I) = I$. 所有的不变集组成 \mathcal{S} 的一个子 σ -代数, 称为不变 σ -代数, 记为 \mathcal{I} . 类似的, 称 S 上可测函数 f 为不变函数, 若 $f \circ T = f$. 可以证明, f 不变当且仅当 f 关于 \mathcal{I} 可测.

显然的, 对于任意的 n , $\mathcal{I} \subset \sigma(\xi_k, k \geq 1)$. 一般的, \mathcal{I} 不是完备的. 记 \mathcal{I} 关于 μ 的完备化为 \mathcal{I}^μ . 则

$$\mathcal{I}^\mu = \{I \in \mathcal{I}, \mu((T^{-1}I) \Delta I) = 0\}.$$

定义 6.6 称测度空间 (S, \mathcal{S}, μ) 上保测变换 T 为关于 μ 遍历的, 若不变 σ -代数 \mathcal{I} 关于 μ 是平凡的, 即对任意的 $I \in \mathcal{I}$, $\mu(I) = 0$ 或 1 . 此时, 我们也称 μ 关于 T 是遍历的. 若随机元 ξ 具有分布 μ , 则称 ξ 是遍历的.

ξ 是遍历的, 当且仅当对任意的 $I \in \mathcal{I}$, $P(\xi \in I) = 0$ 或 1 , 即 σ -代数 $\mathcal{I}_\xi = \xi^{-1}\mathcal{I}$ 关于 P 是平凡的. 特别的, 强平稳序列 $\xi = (\xi_0, \xi_1, \dots)$ 遍历当且仅当推移算子的不变 σ -代数关于 ξ 的分布是平凡的. 类似于前面的结果, 可以证明 ξ 是 T -遍历的, 当且仅当强平稳序列 $(T^n\xi)$ 是 θ -遍历的.

下面是我们这一节的主要定理. 我们采用 Yosida 和 Kakutani 的证明.

定理 6.10 (ergodic theorem, Birkhoff) 设 ξ 是 S 上随机元, 其分布为 μ , T 为 S 上 μ -保测变换. 则对 S 上任意非负可测函数 f , 当 $n \rightarrow \infty$ 时,

$$\frac{1}{n} \sum_{k < n} f(T^k \xi) \rightarrow E[f(\xi) | \mathcal{I}_\xi] \text{ a.s..} \quad (6.13)$$

若对 $p \geq 1$, $f \in L^p(\mu)$, 则上述收敛在 L^p 意义下也成立.

定理的证明基于下面的不等式.

引理 6.2 (maximal ergodic lemma) 设 $\xi = (\xi_1, \xi_2, \dots)$ 是由可积随机变量组成的强平稳序列. 记 $S_n = \xi_1 + \xi_2 + \dots + \xi_n$, 则

$$E[\xi_1 : \sup_n S_n > 0] \geq 0.$$

证明 记 $M_n = S_1 \wedge \dots \wedge S_n$. 则

$$S_k = \xi_1 + S_{k-1} \circ \theta \leq \xi_1 + (M_n \circ \theta)_+, \quad k = 1, \dots, n.$$

于是对任意的 n , $M_n \leq \xi_1 + (M_n \circ \theta)_+$. 由于平稳性,

$$\begin{aligned} E[\xi_1 : M_n > 0] &\geq E[M_n - (M_n \circ \theta)_+ : M_n > 0] \\ &\geq E[(M_n)_+ - (M_n \circ \theta)_+] = 0. \end{aligned}$$

由于 $M_n \uparrow \sup_n S_n$, 由控制收敛定理即得. \square

证明 (定理 6.10 的证明) 先设 $f \in L^1$. 记 $\eta_k = f(T^{k-1}x)$. 由于 $E[\eta_1 | \mathcal{I}_\xi]$ 是 ξ 的不变函数, 序列 $\zeta_k = \eta_k - E[\eta_1 | \mathcal{I}_\xi]$ 也是强平稳的. 令 $S_n = \zeta_1 + \dots + \zeta_n$. 对任意 $\varepsilon > 0$, 定义

$$A_\varepsilon = \{\limsup_{n \rightarrow \infty} \frac{S_n}{n} > \varepsilon\}, \quad \zeta_n^\varepsilon = (\zeta_n - \varepsilon)1_{A_\varepsilon}.$$

注意到 $S_n^\varepsilon = \zeta_1^\varepsilon + \dots + \zeta_n^\varepsilon$ 满足

$$\begin{aligned} \{\sup_n S_n^\varepsilon > 0\} &= \{\sup_n \frac{S_n}{n} > 0\} \\ &= \{\sup_n \frac{S_n}{n} > \varepsilon\} \cap A_\varepsilon = A_\varepsilon. \end{aligned}$$

由于 $A_\varepsilon \in \mathcal{I}_\xi$, (ζ_n^ε) 是强平稳序列,

$$\begin{aligned} 0 &\leq E[\zeta_1^\varepsilon : A_\varepsilon] = E[\zeta_1 - \varepsilon : A_\varepsilon] \\ &= E[E[\zeta_1 | \mathcal{I}_\xi] : A_\varepsilon] - \varepsilon P(A_\varepsilon) = -\varepsilon P(A_\varepsilon). \end{aligned}$$

于是 $P(A_\varepsilon) = 0$. 这样, 由 ε 的任意性, $\limsup_n S_n/n \leq 0$ a.s.. 对 $-S_n$ 应用相同的方法, 有 $\liminf_n S_n/n \geq 0$ a.s., 从而 $S_n/n \rightarrow 0$ a.s..

下面设 $f \in L^p$. 由 Jessen 不等式, 对任意的 $A \in \mathcal{F}$ 和 $r > 0$, 有

$$\begin{aligned} E[1_A |\sum_{k < n} f(T^k \xi)/n|^p] &\leq n^{-1} \sum_{k < n} E[|f(T^k \xi)|^p : A] \\ &\leq r^p P(A) + E[|f(\xi)|^p : |f(\xi)| > r]. \end{aligned}$$

容易证明 $|n^{-1} \sum_{k < n} f(T^k \xi)|^p$ 是一致可积的, 这样即可得到 L^p 意义下的收敛性.

对于一般的非负可测函数 f , 令 $\bar{\eta} = E[f(\xi) | \mathcal{J}_\xi]$. 对任意 $r > 0$, 关于 $\{\bar{\eta} \leq r\}$ 取条件, 可以证明 (6.13) 在 $\{\bar{\eta} < \infty\}$ 上成立. 而在 $\{\bar{\eta} = \infty\}$ 上,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{k \leq n} f(T^k \xi) &\leq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k \leq n} (f(T^k \xi) \vee r) \\ &= E[f(\xi) \vee r | \mathcal{J}_\xi] \rightarrow \infty \quad (r \rightarrow \infty). \end{aligned}$$

□

上面定理说明, 强平稳序列的极限性质依赖于不变 σ -代数的结构. 下面我们看几个例子.

例 6.7 设 $\xi = (\xi_1, \xi_2, \dots)$ 是独立同分布随机变量列. 由 0-1 律, 尾 σ -代数 $\bigcap_n \sigma(\xi_k, k \geq n)$ 是平凡的, 于是 \mathcal{J} 也是平凡的. 故 ξ 是遍历的. 如果 $E[|\xi_1|] < \infty$, 则

$$\lim_{n \rightarrow \infty} \frac{\xi_1 + \dots + \xi_n}{n} = E[\xi_1] \text{ a.s.},$$

这即是强大数定律.

例 6.8 设强平稳序列 $\xi = (\xi_0, \xi_1, \dots)$ 满足 $E[\xi_0] = 0$, $E[|\xi_0|^2] < \infty$. 对非负整数 τ , 令 $\eta_n = \xi_{n+\tau} \xi_n$, 则 $\eta = (\eta_0, \eta_1, \dots)$ 也是强平稳序列. 若 ξ 遍历, 则 η 也是遍历的. 此时

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k < n} \xi_{k+\tau} \xi_k = B(\tau) \text{ a.s.},$$

其中 $B(\tau) = E[\xi_{n+\tau} \xi_n]$.

对于连续时间情形, 我们考虑 S 上一族变换 T_t , 它满足半群性质: $T_{t+s} = T_t T_s$. 我们假设 (T_t) 是可测的, 即对任意的 $x \in S$, 映射 $(x, t) \mapsto T_t x$ 是可测的, 此时我们称 (T_t) 为一个流. 离散时间情形的概念都可以类似的定义. 例如, 不变 σ -代数

$$\mathcal{J} = \{I \in \mathcal{S}, T_t^{-1} I = I, \forall t\},$$

S 上随机元 ξ 称为 (T_t) -强平稳的, 若对任意 t , $T_t \xi \stackrel{d}{=} \xi$. 下面的定理是连续时间下的遍历定理.

定理 6.11 设 ξ 是 S 上随机元, 具有分布 μ , (T_t) 是保测变换组成的流. 记不变 σ -代数 \mathcal{I} . 则对任意的非负可测函数 f ,

$$\lim_{t \rightarrow \infty} t^{-1} \int_0^t f(T_s \xi) ds = E[f(\xi) | \mathcal{I}_\xi] \text{ a.s..}$$

若 $f \in L^p(\mu)$, 则上述收敛在 L^p 意义下也成立.

海外文談