



# Stochastic Process

## 随机过程

作者: Fir1247

组织: USTC

时间: June 22, 2024

联系方式: fa1247@mail.ustc.edn.cn 或者 QQ: 3105292483

有很多故事戛然而止，有很多故事让人辗转反侧，  
另一些..... 最后只能变成一声长长的叹息。  
在没有尽头的荒原上，人们相逢，然后走散，  
但在错身而过的那个瞬间，  
没有人知道，我们是否应该郑重地道出再见。



# 前言

本文档涉及以下中科大 2024 春季课程：随机过程（张土生）<sup>1</sup>、应用随机过程（翟建梁）。参考书：《Probability Theory and Examples (5th edition)》（作者：Durrett）、《Brownian Motion, Martingales, and Stochastic Calculus》（GTM275，作者：Le Gall）等。阅读本文档可能需要的前置知识（按照重要程度排序）：数学分析、线性代数、实分析（主要是测度论）、初等概率论、泛函分析（不是很重要）。

这份文档以随机过程（以下简称“研随”）课程内容为主体，小节标题标注“\*”代表此部分不属于研随课程范围。按照顺序主要可以分成两大部分：离散过程和连续过程。也可以分成前置知识（一些基本的测度论定理，以及条件数学期望）、鞅（离散鞅和连续鞅）与马氏过程（离散马氏链，Poisson 过程、布朗运动则是特殊的连续马氏过程）。鞅是我认为的本课程最有趣的部分，先分析了各种情况下的收敛性，然后介绍了最重要的择停定理，在经过作业和考试的洗礼之后，“给鞅套一个停时再取极限”已经刻在了 DNA 里。马氏过程部分没有讲太应用的例子，大概就是把基本概念介绍了一遍，但也足够精彩，逻辑环环相扣叙述完备，总比会算东西但不明不白强很多。

本文档的另一部分则为应用随机过程（以下简称“应随”），应随部分期中前重点研究离散时间马氏链的状态分类和极限分布，期中后重点研究跳跃过程（通过跳链分析性质，还是蛮有意思的），并对 Poisson 过程、布朗运动、随机积分做了简单介绍。此课程全程拿初等语言讲述，如果强行全部整理进来会造成笔记整体的不和谐，部分深刻内容一带而过（例如随机积分），令人不明觉厉。但应随内容并非完全没有价值，研随虽然足够有深度，却缺乏直观性，也没能提及很多实际研究经常利用的技巧和例子。因此推荐读者当做习题课内容来阅读。在笔记的正文部分，我尝试对应随的内容基于研随已经给出的理论进行重新排版，略去了很多证明（因为懒），与研随部分结合着一起看体验会比较好，也能给自己带来更多直观理解。

现在（2024.06.22）的这个版本是笔记的最初版本，可能会有很多 Typo，也会有叙述不完整的部分，还望读者不吝赐教、提出修改意见。最后，特别感谢 yhb 同学<sup>2</sup>为本笔记的完善做出的贡献。

最后更新：June 22, 2024

---

<sup>1</sup>期末卷子也许不能在明面上发布，悄悄在这里说一下，回忆版本：<http://home.ustc.edu.cn/~fa1247/2024SPFinal.pdf>

<sup>2</sup>完成了应随连续马氏链部分翟老师讲义的“中译中”工作，讲义手稿原文乱七八糟，我是真看不下去。

# 目录

<b>第 1 章 离散时间鞅</b>	<b>1</b>
1.1 条件数学期望	1
1.2 介绍离散时间鞅	4
1.2.1 基本定义与性质	4
1.2.2 例子	5
1.3 停时	7
1.4 鞅的分解定理	9
1.5 鞅的收敛性	9
1.5.1 鞅收敛定理	10
1.5.2 推论与应用	12
1.6 鞅的 $L^p$ 收敛性	13
1.7 鞅的择停定理	15
1.7.1 回顾：一致可积性	15
1.7.2 一致可积鞅	17
1.7.3 择停定理的推论与应用	19
1.8 两个例子	20
1.9 习题	22
1.9.1 第一次作业	22
1.9.2 第二次作业	25
1.9.3 第三次作业	28
1.10 向后鞅*	30
<b>第 2 章 离散时间马氏过程</b>	<b>32</b>
2.1 马氏链的定义	32
2.1.1 基本定义	32
2.1.2 基本性质和应用*	33
2.1.3 转移概率函数	36
2.2 马氏链的构造与性质	37
2.2.1 从转移概率构造马氏链	37
2.2.2 马氏性	39



2.2.3 强马氏性	40
2.3 马氏链状态的分类	41
2.3.1 常返态与瞬时态	41
2.3.2 进一步分类与判断 *	44
2.3.3 实例: 简单对称随机游走	47
2.4 平稳测度与平稳分布	50
2.4.1 基本定义	50
2.4.2 平稳测度的存在唯一性	50
2.4.3 平稳分布的性质	52
2.4.4 利用平稳分布判断状态 *	54
2.4.5 极限定理 *	55
2.5 终止分布问题 (Exit Distributions)*	59
2.6 习题	66
2.6.1 研随第四次作业	66
2.6.2 研随第五次作业	70
2.6.3 应随部分习题	72
2.7 关于本章记号的说明	75
<b>第 3 章 Poisson 过程</b>	<b>76</b>
3.1 Poisson 过程的定义	76
3.2 构造 Poisson 过程	77
3.3 Poisson 的另一种定义 *	80
3.4 Poisson 过程的更多性质 *	81
3.5 生过程 *	82
<b>第 4 章 布朗运动</b>	<b>86</b>
4.1 高斯空间	86
4.2 准布朗运动	88
4.3 连续轨道与布朗运动	90
4.4 布朗运动路径的性质	92
4.5 布朗运动的强马氏性	94
4.6 习题	97
4.7 随机积分介绍 *	101
4.7.1 定义	101
4.7.2 $L^2$ 可积实函数对 B.M. 积分的性质	102
4.7.3 随机过程对 B.M. 积分的定义	102
4.7.4 Itô 公式	104
<b>第 5 章 连续时间鞅</b>	<b>107</b>
5.1 滤流	107
5.2 停时	109
5.3 鞅的定义与基本性质	113
5.4 鞅的收敛性	116
5.4.1 轨道修正	116
5.4.2 右连续鞅的收敛性	118
5.5 鞅的择停定理	120
5.6 例子与习题	124



第 6 章 连续时间马氏过程	131
6.1 马氏过程的定义	131
6.1.1 转移核与转移半群	131
6.1.2 马氏过程	132
6.2 马氏过程的构造	134
6.3 马氏过程的其他要素	134
6.3.1 预解式	134
6.3.2 Feller 半群	137
6.3.3 生成元	138
6.4 马氏性与强马氏性	142
6.5 两个例子：跳跃过程与 Levy 过程	144
6.5.1 跳跃过程	144
6.5.2 Levy 过程	146
6.6 习题	147
6.7 关于跳跃过程的更多内容 *	151
6.7.1 基本概念	151
6.7.2 状态分类、状态之间的关系	152
6.7.3 Kolmogorov 方程	154
6.7.4 平稳分布	156
6.7.5 终止分布问题 (Exit Distributions)	158

# 第 1 章 离散时间鞅

## 1.1 条件数学期望

### 定义 1.1.1

概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上,  $\mathcal{G} \subset \mathcal{F}$  是一个  $\sigma$ -域, 随机变量  $X$  在给定  $\mathcal{G}$  时的条件 (数学) 期望 (Conditional Expectation), 是指满足以下两个条件的随机变量  $Y$ :

1°  $Y$  是  $\mathcal{G}$ -可测的, 即  $Y^{-1}(B) \in \mathcal{G}, \forall B \in \mathcal{R}$ .

2°  $\forall A \in \mathcal{G}$ , 成立关系式:

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$$

我们往往把条件期望  $Y$  记作  $\mathbb{E}[X|\mathcal{G}]$ .

顺带一提, 条件概率  $\mathbb{P}(A|\mathcal{G}) \stackrel{\text{def}}{=} \mathbb{E}[I_A|\mathcal{G}]$ .



### 命题 1.1.1

在几乎处处相等意义下, 条件期望存在且唯一。

**证明** 唯一性: 如果随机变量  $Y'$  也满足定义 1.1.1 的两个条件, 对于  $\forall \varepsilon > 0$ , 取集合  $A = \{\omega : Y(\omega) - Y'(\omega) \geq \varepsilon\}$ , 由于  $Y, Y'$  都是  $\mathcal{G}$ -可测的, 所以  $A \in \mathcal{G}$ , 因此

$$\varepsilon \cdot \mathbb{P}(A) \leq \int_A (Y - Y') d\mathbb{P} = \int_A Y d\mathbb{P} - \int_A Y' d\mathbb{P} = \int_A X d\mathbb{P} - \int_A X d\mathbb{P} = 0$$

所以  $\mathbb{P}(A) = 0$ , 同理  $\mathbb{P}(\{\omega : Y'(\omega) - Y(\omega) \geq \varepsilon\}) = 0$ , 因此  $Y = Y'$  a.s.

存在性: 如果两个有限测度  $\mu, \nu$  满足关系:  $\mu(A) = 0 \Rightarrow \nu(A) = 0$ , 则记作  $\nu \ll \mu$ . Radon-Nikodym 定理表明, 测度空间  $(\Omega, \mathcal{F})$  上, 如果  $\nu \ll \mu$ , 则存在一个  $\mathcal{F}$ -可测的可积函数  $f$  满足

$$\forall A \in \mathcal{F}, \nu(A) = \int_A f(\omega) d\mu$$

一般记作  $f = \frac{d\nu}{d\mu}$ ,  $f$  称为 Radon-Nikodym 导数。言归正传, 概率空间  $(\Omega, \mathcal{G}, \mathbb{P})$  上, 对于非负的随机变量  $X \geq 0$ , 令  $\mu = \mathbb{P}$ ,

$$\nu : \mathcal{G} \rightarrow \mathbb{R}, A \mapsto \int_A X d\mathbb{P}$$

则  $\mu, \nu$  都是  $(\Omega, \mathcal{G})$  上的有限测度并且  $\nu \ll \mu$ , 于是存在  $\mathcal{G}$ -可测的随机变量  $Y$  满足:

$$\int_A X d\mathbb{P} = \nu(A) = \int_A Y d\mathbb{P}$$

这正是我们定义的条件期望。 □

### 例 1.1.1.

给定事件  $A, B \in \mathcal{F}$ , 取  $\sigma$ -域  $\mathcal{G} = \{\emptyset, \Omega, A, A^c\}$  和随机变量  $X = I_B$ , 则

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{P}(B|A) \cdot I_A + \mathbb{P}(B|A^c) \cdot I_{A^c}$$

**证明** 令  $Y = \mathbb{P}(B|A) \cdot I_A + \mathbb{P}(B|A^c) \cdot I_{A^c}$ , 验证其满足定义 1.1.1 的两个条件即可。

1° 显然  $I_A$  和  $I_{A^c}$  是  $\mathcal{G}$ -可测的, 所以  $Y$  也是  $\mathcal{G}$ -可测的。

2° 对  $\mathcal{G}$  中的元素逐个验证即可, 例如

$$\begin{aligned} \int_{\Omega} Y d\mathbb{P} &= \mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|A^c)\mathbb{P}(A^c) = \mathbb{P}(B) = \int_{\Omega} I_B d\mathbb{P} = \int_{\Omega} X d\mathbb{P} \\ \int_A Y d\mathbb{P} &= \int_{\Omega} \mathbb{P}(B|A)I_A d\mathbb{P} = \mathbb{P}(B|A)\mathbb{P}(A) = \mathbb{P}(A \cap B) = \int_A I_B d\mathbb{P} = \int_A X d\mathbb{P} \end{aligned}$$

其余不再赘述。 □

下面的几个定理介绍了条件期望的性质, 注意这些性质都是从定义出发直接得到的, 所以在证明的过程中我们反复使用了定义 1.1.1.

### 定理 1.1.1

a.s. 意义下:

- (1).  $\mathbb{E}[aX_1 + bX_2|\mathcal{G}] = a\mathbb{E}[X_1|\mathcal{G}] + b\mathbb{E}[X_2|\mathcal{G}]$ , 其中  $a, b$  是常数。
- (2). 若  $X \geq 0$ , 则  $\mathbb{E}[X|\mathcal{G}] \geq 0$ . 特别地,  $X_1 \geq X_2 \Rightarrow \mathbb{E}[X_1|\mathcal{G}] \geq \mathbb{E}[X_2|\mathcal{G}]$ .
- (3). 如果随机变量列  $X_n \geq 0$  且  $X_n \nearrow X$  a.s.,  $X$  可积, 那么  $\mathbb{E}[X_n|\mathcal{G}] \xrightarrow{\text{a.s.}} \mathbb{E}[X|\mathcal{G}]$ ; 如果随机变量列  $X_n \rightarrow X$  a.s. 且存在可积的  $Y$  使得  $|X_n| \leq Y$ , 那么  $\mathbb{E}[X_n|\mathcal{G}] \xrightarrow{\text{a.s.}} \mathbb{E}[X|\mathcal{G}]$ . 这是条件期望版本下的单调/控制收敛定理。

### 证明

(1). 利用积分的线性即可得证。

(2).  $\forall \varepsilon > 0$ , 设集合  $A = \{\omega : \mathbb{E}[X|\mathcal{G}] \leq -\varepsilon\}$ , 因为随机变量  $\mathbb{E}[X|\mathcal{G}]$  是  $\mathcal{G}$ -可测的, 所以集合  $A \in \mathcal{G}$ , 因此

$$0 \leq \int_A X d\mathbb{P} = \int_A \mathbb{E}[X|\mathcal{G}] d\mathbb{P} \leq -\varepsilon \mathbb{P}(A)$$

因此  $\mathbb{P}(A) = 0$ , 即  $\mathbb{E}[X|\mathcal{G}] \geq 0$  a.s.

(3). 取随机变量  $Z_n = \mathbb{E}[X_n|\mathcal{G}]$ , 由 (2) 知  $Z_n$  单调递增, 设  $Z_n \nearrow Z$ , 希望证明  $Z = \mathbb{E}[X|\mathcal{G}]$ , 也就是验证定义 1.1.1 中的两个条件即可:

1°  $Z_n = \mathbb{E}[X_n|\mathcal{G}]$  都是  $\mathcal{G}$ -可测的, 所以它们的极限也是  $\mathcal{G}$ -可测的。

2°  $\forall A \in \mathcal{G}$ ,

$$\int_A Z d\mathbb{P} = \lim_{n \rightarrow \infty} \int_A Z_n d\mathbb{P} = \lim_{n \rightarrow \infty} \int_A X_n d\mathbb{P} \stackrel{\text{DCT/MCT of r.v.}}{=} \int_A X d\mathbb{P}$$

□

### 定理 1.1.2

- (1).  $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$ .
- (2). 随机变量  $X$  和  $\mathcal{G}$  独立, 即  $X$  与  $\forall I_A, A \in \mathcal{G}$  是独立的, 那么  $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ , 是一个常数。
- (3). 如果随机变量  $Y$  是  $\mathcal{G}$ -可测的, 那么  $\mathbb{E}[YX|\mathcal{G}] = Y \cdot \mathbb{E}[X|\mathcal{G}]$ .

(4). 如果  $\sigma$ -域  $\mathcal{G}_1 \subset \mathcal{G}$ , 则  $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$ .

**证明**

(1). 全集  $\Omega \in \mathcal{G}$ , 所以

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \int_{\Omega} \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_{\Omega} Xd\mathbb{P} = \mathbb{E}[X]$$

(2). 令  $Y = \mathbb{E}[X]$ , 来验证定义 1.1.1 中的两个条件:

1° 显然,  $Y$  作为一个常数是全集  $\Omega$  的示性函数的倍数, 肯定是  $\mathcal{G}$ -可测的。

2°  $\forall A \in \mathcal{G}$ , 因为  $X$  和  $I_A$  独立, 所以  $\mathbb{E}[X] \cdot \mathbb{E}[I_A] = \mathbb{E}[X \cdot I_A]$ , 那么

$$\int_A Yd\mathbb{P} = \mathbb{P}(A)\mathbb{E}[X] = \mathbb{E}[X] \cdot \mathbb{E}[I_A] = \mathbb{E}[X \cdot I_A] = \int_{\Omega} X \cdot I_Ad\mathbb{P} = \int_A Xd\mathbb{P}$$

(3). 令  $Z = Y \cdot \mathbb{E}[X|\mathcal{G}]$ , 来验证定义 1.1.1 中的两个条件:

1°  $Y$  和  $\mathbb{E}[X|\mathcal{G}]$  是  $\mathcal{G}$ -可测的, 自然其乘积  $Z = Y \cdot \mathbb{E}[X|\mathcal{G}]$  也是  $\mathcal{G}$ -可测的。

2° 还是经典的四步走: 示性函数、简单函数、非负函数、一般函数。

(i). 考虑  $Y = I_B, \forall B \in \mathcal{G}$ , 则对于  $\forall A \in \mathcal{G}$ ,

$$\int_A Zd\mathbb{P} = \int_A I_B \cdot \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_{A \cap B} \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_{A \cap B} Xd\mathbb{P} = \int_A X \cdot I_Bd\mathbb{P} = \int_A XYd\mathbb{P}$$

(ii). 考虑  $Y = \sum_{k=1}^n a_k I_{B_k}$ , 其中  $B_k \in \mathcal{G}$ , 利用定理 1.1.1 的 (1) 即可得证。

(iii). 考虑  $Y \geq 0$ , 取一系列简单函数  $Y_n \nearrow Y$ , 利用定理 1.1.1 的 (3) 可知

$$\mathbb{E}[YX|\mathcal{G}] = \lim_{n \rightarrow \infty} \mathbb{E}[Y_n X|\mathcal{G}] = \lim_{n \rightarrow \infty} Y_n \mathbb{E}[X|\mathcal{G}] = Y \mathbb{E}[X|\mathcal{G}]$$

(iv). 对于一般的随机变量  $Y$ , 拆分成正负部即可。

(4). 令  $Z = \mathbb{E}[X|\mathcal{G}]$  和  $Y = \mathbb{E}[X|\mathcal{G}_1]$ , 希望证明  $Y = \mathbb{E}[Z|\mathcal{G}_1]$ ,

1° 根据定义,  $Y = \mathbb{E}[X|\mathcal{G}_1]$  是  $\mathcal{G}_1$ -可测的。

2°  $\forall A \in \mathcal{G}_1 \subset \mathcal{G}$ ,

$$\begin{aligned} \int_A Zd\mathbb{P} &= \int_A \mathbb{E}[X|\mathcal{G}]d\mathbb{P} = \int_A Xd\mathbb{P} \\ \int_A Yd\mathbb{P} &= \int_A \mathbb{E}[X|\mathcal{G}_1]d\mathbb{P} = \int_A Xd\mathbb{P} \end{aligned}$$

所以二者相等。

□

### 定理 1.1.3 (Jensen's Inequality 的条件期望版本)

$\varphi$  是凸函数,  $X$  和  $\varphi(X)$  可积, 则

$$\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}]$$

**证明** 凸函数的性质是切线在曲线下方, 所以  $\forall x, y \in \mathbb{R}$ ,

$$\varphi'(y)(x - y) + \varphi(y) \leq \varphi(x)$$

把  $x$  替换成随机变量  $X$ ,  $y$  替换成  $\mathbb{E}[X|\mathcal{G}]$ , 即有

$$\varphi'(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G}]) + \varphi(\mathbb{E}[X|\mathcal{G}]) \leq \varphi(X)$$

两边关于  $\mathcal{G}$  同时取条件期望, 不等号仍然成立 (定理 1.1.1(2))。右边变成了  $\mathbb{E}[\varphi(X)|\mathcal{G}]$ , 这是我们所需要的形式, 那左边如何?

$$\mathbb{E}[\varphi'(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G}]) + \varphi(\mathbb{E}[X|\mathcal{G}])|\mathcal{G}] = \mathbb{E}[\varphi'(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G}])|\mathcal{G}] + \mathbb{E}[\varphi(\mathbb{E}[X|\mathcal{G}])|\mathcal{G}]$$

注意  $\mathbb{E}[X|\mathcal{G}]$  是  $\mathcal{G}$ -可测的, 套一个函数  $\varphi'$  仍然是  $\mathcal{G}$ -可测的, 所以根据定理 1.1.2(3) 可以把它提出去, 因此

$$\text{LHS} = \varphi'(\mathbb{E}[X|\mathcal{G}])\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])|\mathcal{G}] + \mathbb{E}[\varphi(\mathbb{E}[X|\mathcal{G}])|\mathcal{G}]$$

一个很显然的推论是

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]\mathbb{E}[1|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]\mathbb{E}[1] = \mathbb{E}[X|\mathcal{G}]$$

因此  $\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}]|\mathcal{G}] = 0$ ,  $\text{LHS} = \mathbb{E}[\varphi(\mathbb{E}[X|\mathcal{G}]|\mathcal{G})]$ , 注意  $\varphi(\mathbb{E}[X|\mathcal{G}])$  也是  $\mathcal{G}$ -可测的,  $\text{LHS} = \varphi(\mathbb{E}[X|\mathcal{G}])$ , 至此定理得证。□

### 推论 1.1.1

取凸函数  $\varphi(x) = |x|^p$ , 则有  $|\mathbb{E}[X|\mathcal{G}]|^p \leq \mathbb{E}[|X|^p|\mathcal{G}]$ .

### 命题 1.1.2

随机变量  $X$  满足  $\mathbb{E}[X^2] < \infty$ , 考虑所有的  $\mathcal{G}$ -可测的随机变量  $Y$ , 使得  $\mathbb{E}[(X - Y)^2]$  达到最小值的就是条件期望  $\mathbb{E}[X|\mathcal{G}]$ , 即

$$\mathbb{E}[X|\mathcal{G}] = \underset{\text{r.v. } Y \in \mathcal{G}}{\operatorname{argmin}} \mathbb{E}[(X - Y)^2]$$

其中  $\text{r.v. } Y \in \mathcal{G}$  代表随机变量  $Y$  是  $\mathcal{G}$ -可测的。

这个结论表明, 条件期望极小化均方误差 (minimize the mean square error), 是对随机变量的“最好估计”。

**证明**  $\forall \text{r.v. } Y \in \mathcal{G}$ ,

$$\begin{aligned} \mathbb{E}[(X - Y)^2] &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}] + \mathbb{E}[X|\mathcal{G}] - Y)^2] \\ &= \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2] + \mathbb{E}[(\mathbb{E}[X|\mathcal{G}] - Y)^2] + 2\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Y)] \end{aligned}$$

关注最后一项, 利用 [定理 1.1.2\(1\)](#), 可知

$$\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Y)] = \mathbb{E}[\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Y)|\mathcal{G}]]$$

注意到  $(\mathbb{E}[X|\mathcal{G}] - Y)$  是  $\mathcal{G}$ -可测的, 利用 [定理 1.1.2\(3\)](#) 可以把它提出来,

$$\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])(\mathbb{E}[X|\mathcal{G}] - Y)] = \mathbb{E}[(\mathbb{E}[X|\mathcal{G}] - Y) \cdot \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}]|\mathcal{G})]]$$

而  $\mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}]|\mathcal{G})] = \mathbb{E}[X|\mathcal{G}] - \mathbb{E}[X|\mathcal{G}] = 0$ , 所以

$$\mathbb{E}[(X - Y)^2] = \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2] + \mathbb{E}[(\mathbb{E}[X|\mathcal{G}] - Y)^2] \geq \mathbb{E}[(X - \mathbb{E}[X|\mathcal{G}])^2]$$

等号成立当且仅当  $Y = \mathbb{E}[X|\mathcal{G}]$  a.s. □

## 1.2 介绍离散时间鞅

### 1.2.1 基本定义与性质

鞅是一种特殊的随机过程, 我们先来定义什么是随机过程。

#### 定义 1.2.1

概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上的一族随机变量  $\{X_t, t \in I\}$  称为随机过程, 其中指标集  $I$  一般是  $[0, +\infty)$  或者  $\{0, 1, \dots\}$ .

随机过程  $\{X_t, t \in I\}$  的有限维分布是指给定有限个时间点  $0 \leq t_1 < t_2 < \dots < t_n$  时随机向量  $(X_{t_1}, \dots, X_{t_n})$  的联合分布。

称两个随机过程同分布, 是指其有相同的有限维分布。

从直观上理解, 随机过程里的指标集就是现实中的“时间轴”, 每一个时刻都有一个相应的随机变量与之对应。

## 定义 1.2.2

给定一系列单调递增的  $\sigma$ -域  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_n \subset \dots$ , 称之为滤流 (filtration), 如果有一列随机变量  $Z_0, Z_1, \dots, Z_n, \dots$  满足: 对于每个  $n$ ,

1°  $Z_n$  是  $\mathcal{F}_n$ -可测的。

2°  $\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = Z_n$ .

那么称随机过程  $\{Z_n, n \in \mathbb{N}\}$  为关于  $\{\mathcal{F}_n, n \in \mathbb{N}\}$  的离散时间鞅, 或者简称为鞅 (martingale)。

将条件 2° 改成  $\mathbb{E}[Z_{n+1}|\mathcal{F}_n] \geq Z_n$  或者  $\mathbb{E}[Z_{n+1}|\mathcal{F}_n] \leq Z_n$ , 则称为下鞅 (submartingale) 或者上鞅 (supermartingale)。



从直观上理解, 一系列单调递增的  $\sigma$ -域  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_n \subset \dots$  就是一个“随时间演化的信息族”。条件 2° 表明, 如果掌握了所有  $t \leq n$  的信息,  $Z_{n+1}$  的期望就是  $Z_n$ , 也就是说时间点  $t = n + 1$  发生的事情, 只取决于  $t \leq n$  已经观测到的所有信息。

例 1.2.1. 随机变量生成的  $\sigma$ -域

我们熟知, 随机变量就是 Borel 可测函数, 因此所有事件

$$X^{-1}(B) = \{X \in B\}, \forall B \in \mathcal{R}$$

就可以描述观测随机变量  $X$  可能得到的所有信息, 或者说与随机变量  $X$  有关的所有事件, 这其实就是我们之前提到的随机变量  $X$  生成的  $\sigma$ -域, 见??。类似地可以定义多个随机变量生成的  $\sigma$ -域:

$$\sigma(X_1, \dots, X_n) \stackrel{\text{def}}{=} \sigma(\{a_k \leq X_k \leq b_k\}, a_k, b_k \in \mathbb{R}, 1 \leq k \leq n)$$

那么, 对于随机过程  $\{Z_n, n \in \mathbb{N}\}$ , 令  $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ , 就得到了一个滤流  $\{\mathcal{F}_n\}$ 。

一个简单的推论是:  $f$  是  $\sigma(X_1, \dots, X_n)$ -可测的当且仅当  $f$  是  $X_1, \dots, X_n$  的 Borel 可测函数。根据条件期望的定义,  $\mathbb{E}[X|\sigma(X_1, \dots, X_n)]$  是  $\sigma(X_1, \dots, X_n)$ -可测的, 这就说明  $\mathbb{E}[X|\sigma(X_1, \dots, X_n)]$  可以写成  $X_1, \dots, X_n$  的 Borel 可测函数! 给定  $\sigma(X_1, \dots, X_n)$  就相当于获取了和  $X_1, \dots, X_n$  相关的所有信息, 所以最后的结果也只会与  $X_1, \dots, X_n$  有关, 这非常符合条件期望的直观印象。我们在初等概率论中定义的条件期望  $\mathbb{E}[X|Y]$ , 其实是  $\mathbb{E}[X|\sigma(Y)]$ 。

## 命题 1.2.1

如果  $\{Z_n, n \in \mathbb{N}\}$  是关于  $\{\mathcal{F}_n, n \in \mathbb{N}\}$  的鞅, 则

$$\mathbb{E}[Z_{n+2}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[Z_{n+2}|\mathcal{F}_{n+1}]|\mathcal{F}_n] = \mathbb{E}[Z_{n+1}|\mathcal{F}_n] = Z_n$$

进而  $\forall m \geq n$ , 都有  $\mathbb{E}[Z_m|\mathcal{F}_n] = Z_n$ , 因此  $\mathbb{E}[Z_m] = \mathbb{E}[\mathbb{E}[Z_m|\mathcal{F}_n]] = \mathbb{E}[Z_n]$ 。

**注** 对于下鞅, 结论就改为:  $\mathbb{E}[Z_m] \geq \mathbb{E}[Z_n], \forall m \geq n$ . 从这里可以看出上/下鞅的命名方式稍微有点反直觉, 上鞅的期望是下降的、下鞅的期望是上升的。

有关凸函数的推论:

## 推论 1.2.1

- (1).  $\{X_n\}$  是一个鞅,  $\varphi$  是凸函数, 且  $\forall \varphi(X_n)$  可积,  $\{\varphi(X_n)\}$  是一个下鞅。
- (2).  $\{X_n\}$  是一个下鞅,  $\varphi$  是单调递增的凸函数, 且  $\forall \varphi(X_n)$  可积,  $\{\varphi(X_n)\}$  仍是一个下鞅。



## 1.2.2 例子

下面我们给出一些鞅的例子。为了叙述方便, 如果没有特殊说明, 默认鞅都是关于滤流  $\{\mathcal{F}_n\}$  的。

**例 1.2.2.**

$X_0, \dots, X_n, \dots$  是独立可积的 r.v., 且  $\forall n, \mathbb{E}[X_n] = 0$ , 令  $S_n = X_0 + \dots + X_n$ ,  $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ , 则  $\{S_n, n \in \mathbb{N}\}$  是鞅。

**证明** 显然  $S_n$  是  $\mathcal{F}_n$ -可测的,

$$\mathbb{E}[S_{n+1} | \mathcal{F}_n] = \mathbb{E}[S_n + X_{n+1} | \mathcal{F}_n] = S_n + \mathbb{E}[X_{n+1} | \mathcal{F}_n]$$

因为  $X_{n+1}$  和  $\mathcal{F}_n$  是独立的, 所以  $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = \mathbb{E}[X_{n+1}] = 0$ , 所以  $\mathbb{E}[S_{n+1} | \mathcal{F}_n] = S_n$ , 因此  $\{S_n, n \in \mathbb{N}\}$  是鞅。  
□

**例 1.2.3.**

给定一个滤流  $\{\mathcal{F}_n, n \in \mathbb{N}\}$  和可积的 r.v.  $Y$ , 令  $Z_n = \mathbb{E}[Y | \mathcal{F}_n]$ , 则  $\{Z_n, n \in \mathbb{N}\}$  是鞅。

**证明** 由条件期望的定义,  $Z_n$  是  $\mathcal{F}_n$ -可测的,

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[Y | \mathcal{F}_{n+1}] | \mathcal{F}_n] = \mathbb{E}[Y | \mathcal{F}_n] = Z_n$$

□

**注** 这个鞅还有个重要性质: 一致可积。见推论 1.7.2。

**例 1.2.4.**

$Y_1, \dots, Y_n, \dots$  是独立可积的 r.v., 记  $a_i = \mathbb{E}[Y_i]$ , 且  $\forall a_i \neq 0$ , 令  $Z_n = \frac{Y_1 \cdots Y_n}{a_1 \cdots a_n}$ ,  $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$ , 则  $\{Z_n, n \in \mathbb{N}_+\}$  是鞅。

**证明** 显然  $Z_n$  是  $\mathcal{F}_n$ -可测的,

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n] = \mathbb{E}\left[\frac{Y_1 \cdots Y_{n+1}}{a_1 \cdots a_{n+1}} \mid \mathcal{F}_n\right] = \frac{Y_1 \cdots Y_n}{a_1 \cdots a_{n+1}} \mathbb{E}[Y_{n+1} | \mathcal{F}_n] = \frac{Y_1 \cdots Y_n}{a_1 \cdots a_{n+1}} \mathbb{E}[Y_{n+1}] = \frac{Y_1 \cdots Y_n}{a_1 \cdots a_n} = Z_n$$

□

**例 1.2.5.**

r.v. 列  $X_1, \dots, X_n, \dots$  满足  $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$ ,  $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ ,  $S_n = X_1 + \dots + X_n$ ,  $Y_n = S_n^2 - n$ , 则  $\{Y_n, n \in \mathbb{N}_+\}$  是鞅。

**证明** 显然  $Y_n$  是  $\mathcal{F}_n$ -可测的,

$$\begin{aligned} \mathbb{E}[Y_{n+1} | \mathcal{F}_n] &= \mathbb{E}[S_{n+1}^2 - (n+1) | \mathcal{F}_n] \\ &= \mathbb{E}[S_n^2 + X_{n+1}^2 + 2S_n X_{n+1} | \mathcal{F}_n] \\ &= S_n^2 + \mathbb{E}[X_{n+1}^2] + 2S_n \mathbb{E}[X_{n+1} | \mathcal{F}_n] - (n+1) \\ &= S_n^2 + 1 - (n+1) = S_n^2 - n = Y_n \end{aligned}$$

□

**例 1.2.6.**

独立同分布 r.v. 列  $X_1, \dots, X_n, \dots$  服从二项分布  $b(p)$ , 令  $S_n = X_1 + \dots + X_n$ ,  $Z_n = \left(\frac{1-p}{p}\right)^{S_n}$ , 则  $\{Z_n, n \in \mathbb{N}_+\}$  是鞅。

**证明** 显然  $Z_n$  是  $\mathcal{F}_n$ -可测的,

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n] = \mathbb{E}\left[Z_n \cdot \left(\frac{1-p}{p}\right)^{X_{n+1}} \mid \mathcal{F}_n\right] = Z_n \mathbb{E}\left[\left(\frac{1-p}{p}\right)^{X_{n+1}}\right] = Z_n \left(\frac{1-p}{p} \cdot p + \frac{p}{1-p} \cdot (1-p)\right) = Z_n$$

□

### 1.3 停时

#### 定义 1.3.1

$\{\mathcal{F}_n, n \in \mathbb{N}\}$  是滤流, r.v.  $T$  只取非负整数值和  $+\infty$ , 如果  $\forall n \geq 0, \{T \leq n\} \in \mathcal{F}_n$ , 则称  $T$  为 (关于  $\{\mathcal{F}_n, n \in \mathbb{N}\}$ ) 的停时 (stopping times).



#### 例 1.3.1.

独立同分布 r.v. 列  $X_1, \dots, X_n, \dots$  服从二项分布  $b(p)$ , 令  $S_n = X_1 + \dots + X_n$ ,  $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$ , 取

$$T = \min\{k \geq 1 : S_k = 100\}$$

即首次  $S_n$  首次达到 100 时的时间, 那么  $T$  是一个停时, 因为

$$\{T \leq n\} = \bigcup_{k=1}^n \{S_k = 100\}$$

右边的  $\{S_k = 100\}$  是  $\mathcal{F}_k$ -可测的,  $k = 1, \dots, n$ , 因此都是  $\mathcal{F}_n$ -可测的, 从而  $\{T \leq n\}$  是  $\mathcal{F}_n$ -可测的, 于是  $T$  是一个停时。

#### 推论 1.3.1

如果  $T$  是停时, 则

- (1).  $\{T > n\} = \{T \leq n\}^c$  是  $\mathcal{F}_n$ -可测的。
- (2).  $\{T = n\} = \{T \leq n\} \cap \{T > n-1\}$  是  $\mathcal{F}_n$ -可测的。



#### 定义 1.3.2

$T$  是一个停时,  $\{Z_n, n \in \mathbb{N}\}$  是一个随机过程, 记  $n \wedge T = \min\{T, n\}$ , 定义  $Y_n \stackrel{\text{def}}{=} Z_{n \wedge T}$ , 则  $\{Y_n, n \in \mathbb{N}\}$  成为一个新的随机过程, 称为随机过程  $\{Z_n\}$  关于停时  $T$  的停止过程 (stopped process).



这个停止过程其实就是:  $\{Y_1, Y_2, \dots, Y_{T-1}, Y_T, Y_T, Y_T, \dots\}$ , 只不过  $T$  不是固定的整数, 而是一个停时 r.v.

#### 定理 1.3.1

$\{Z_n, n \in \mathbb{N}\}$  是关于  $\{\mathcal{F}_n, n \in \mathbb{N}\}$  的 (上/下) 鞅,  $T$  是停时, 则停止过程  $\{Y_n = Z_{n \wedge T}, n \in \mathbb{N}\}$  也是关于  $\{\mathcal{F}_n, n \in \mathbb{N}\}$  的 (上/下) 鞅。

**证明** 验证鞅的定义:

1°  $Y_n$  是  $\mathcal{F}_n$ -可测的, 因为:

$$Y_n = Z_{n \wedge T} = Z_n I_{\{T > n\}} + Z_T I_{\{T \leq n\}}$$

很明显  $Z_n$  和  $I_{\{T > n\}}$  是  $\mathcal{F}_n$ -可测的, 但后面的  $Z_T$  并不明显, 因为  $T$  不是固定的整数, 但我们可以稍作变换:

$$Z_T I_{\{T \leq n\}} = \sum_{k=0}^n Z_T I_{\{T=k\}} = \sum_{k=0}^n Z_k I_{\{T=k\}}$$

这样就能看出  $Z_T I_{\{T \leq n\}}$  是  $\mathcal{F}_n$ -可测的了, 所以  $Y_n$  是  $\mathcal{F}_n$ -可测的。

2° 希望证明:  $\mathbb{E}[Z_{T \wedge (n+1)} | \mathcal{F}_n] = Z_{T \wedge n}$ . 依然是拆分  $Z_{T \wedge (n+1)}$ , 考虑

$$\begin{aligned} Z_{T \wedge (n+1)} &= Z_{n+1} I_{\{T \geq n+1\}} + Z_T I_{\{T < n+1\}} \\ &= Z_{n+1} I_{\{T \geq n+1\}} + Z_T I_{\{T \leq n\}} \end{aligned}$$

$$= Z_{n+1}I_{\{T \geq n+1\}} + \sum_{k=0}^n Z_k I_{\{T=k\}}$$

注意到  $\{T \geq n+1\} = \{T \leq n\}^c$  是  $\mathcal{F}_n$ -可测的, 后面的每一个  $Z_k I_{\{T=k\}}$  是  $\mathcal{F}_k \subset \mathcal{F}_n$ -可测的, 因此

$$\begin{aligned} \mathbb{E}[Z_{T \wedge (n+1)} | \mathcal{F}_n] &= \mathbb{E}[Z_{n+1} I_{\{T \geq n+1\}} | \mathcal{F}_n] + \sum_{k=0}^n \mathbb{E}[Z_k I_{\{T=k\}} | \mathcal{F}_n] \\ &= \mathbb{E}[Z_{n+1} | \mathcal{F}_n] I_{\{T \geq n+1\}} + \sum_{k=0}^n Z_k I_{\{T=k\}} \\ &= Z_n I_{\{T \geq n+1\}} + \sum_{k=0}^n Z_k I_{\{T=k\}} \\ &= Z_{T \wedge n} \end{aligned}$$

上、下鞅同理。

2° 的另一种证明方式: 注意到

$$Z_{T \wedge (n+1)} = Z_{T \wedge n} + I_{\{T \geq n+1\}}(Z_{n+1} - Z_n) \quad (*)$$

因为:

$$RHS = \begin{cases} Z_T & , T \leq n \\ Z_n + (Z_{n+1} - Z_n) = Z_n & , T \geq n+1 \end{cases} = Z_{T \wedge (n+1)} = LHS$$

那么直接就得到

$$\begin{aligned} \mathbb{E}[Z_{T \wedge (n+1)} | \mathcal{F}_n] &= \mathbb{E}[Z_{T \wedge n} + I_{\{T \geq n+1\}}(Z_{n+1} - Z_n) | \mathcal{F}_n] \\ &= Z_{T \wedge n} + I_{\{T \geq n+1\}} \mathbb{E}[Z_{n+1} - Z_n | \mathcal{F}_n] \\ &= Z_{T \wedge n} \end{aligned}$$

这是老师上课讲的方法, 两种方法都是选择拆分  $Z_{(n+1) \wedge T}$ . □

### 定义 1.3.3 (停时前 $\sigma$ -域)

$M$  是关于滤流  $\{\mathcal{F}_n\}$  的停时, 定义:

$$\mathcal{F}_M \stackrel{\text{def}}{=} \{A \in \mathcal{F} : A \cap \{M \leq n\} \in \mathcal{F}_n, \forall n\}$$

称为  $M$  前  $\sigma$ -域。 ♣

### 推论 1.3.2

$X_M$  是  $\mathcal{F}_M$ -可测的。 ♥

**证明**  $\forall A = \{X_M \leq a\}$ ,

$$A \cap \{M \leq n\} = \bigcup_{k=0}^n A \cap \{M = k\} = \bigcup_{k=0}^n \{X_k \leq a\} \cap \{M = k\} \in \mathcal{F}_n$$

□

### 定义 1.3.4

对于  $A \in \mathcal{F}_M$ , 令

$$M^A \stackrel{\text{def}}{=} \begin{cases} M & , \text{on } A \\ +\infty & , \text{on } A^c \end{cases}$$

容易验证  $M^A$  也是一个停时, 称为  $M$  在  $A$  上的限制。 ♣

我们后续会用到这个概念, 相关: [题目 1.7](#)

## 1.4 鞅的分解定理

### 定义 1.4.1

$\{\mathcal{F}_n\}$  是滤流, 随机过程  $\{H_n\}$  称为关于  $\{\mathcal{F}_n\}$  可预测的 (predictable w.r.t.  $\{\mathcal{F}_n\}$ ), 如果  $\forall n, H_n$  是  $\mathcal{F}_{n-1}$ -可测的。

### 定理 1.4.1 (Doob 分解定理)

任何一个下鞅  $\{X_n, n \in \mathbb{N}\}$  都可以被唯一地分解成  $X_n = M_n + A_n$ , 其中  $\{M_n, n \in \mathbb{N}\}$  是鞅,  $\{A_n, n \in \mathbb{N}\}$  是可预测的、单调递增的, 且  $A_0 = 0$ .

**证明** 从结论往前倒推:

$$\begin{aligned} \mathbb{E}[X_n | \mathcal{F}_{n-1}] &= \mathbb{E}[M_n | \mathcal{F}_{n-1}] + \mathbb{E}[A_n | \mathcal{F}_{n-1}] \\ &= M_{n-1} + A_n \\ &= X_{n-1} - A_{n-1} + A_n \\ A_n - A_{n-1} &= \mathbb{E}[X_n | \mathcal{F}_{n-1}] - X_{n-1} \\ A_n &= \sum_{k=1}^n A_k - A_{k-1} = \sum_{k=1}^n (\mathbb{E}[X_k | \mathcal{F}_{k-1}] - X_{k-1}), n \geq 1, A_0 = 0 \\ M_n &= X_n - A_n = X_n - \sum_{k=1}^n (\mathbb{E}[X_k | \mathcal{F}_{k-1}] - X_{k-1}) \end{aligned}$$

此即为  $A_n, M_n$  的构造方法, 根据构造过程可以看出是唯一的, 然后验证  $\{M_n\}$  是鞅以及  $A_n$  满足题目条件即可, 比较简单, 这里省略。□

课程后续似乎没有再用过这个分解定理, 所以记一下具体的构造方式就好了。

## 1.5 鞅的收敛性

回顾各种收敛性:

1. 几乎必然收敛:  $X_n \xrightarrow{\text{a.s.}} X$ , 即

$$\mathbb{P}\left(\left\{\omega : \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\}\right) = 1$$

也叫“依概率 1 收敛”。

2. 依  $L^p$  收敛:  $X_n \xrightarrow{L^p} X$ , 即

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

也简记为  $X_n \xrightarrow{p} X$ .

3. 依概率测度收敛:  $X_n \xrightarrow{\mathbb{P}} X$ , 即

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(\{|X_n - X| > \varepsilon\}) = 0$$

4. 依分布收敛:  $X_n \xrightarrow{F} X$ , 即

$$\forall x, \lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x)$$

还有的地方记作  $X_n \xrightarrow{D} X$ .

## 定义 1.5.1

对于随机过程  $\{X_n\}$ , 如果存在一个滤流  $\{\mathcal{F}_n\}$  满足  $\forall n, X_n$  是  $\mathcal{F}_n$ -可测的, 则称  $\{X_n\}$  为  $\{\mathcal{F}_n\}$ -适应过程 ( $\{\mathcal{F}_n\}$ -adapted process).

我们先约定一些记号: 对于实值的  $\{\mathcal{F}_n\}$ -适应过程  $\{X_n\}$  和区间  $[a, b]$ , 记

$$\begin{aligned} \tau_1 &= \min\{n, X_n \leq a\}, \tau_2 = \min\{n \geq \tau_1 : X_n \geq b\} \\ \tau_3 &= \min\{n \geq \tau_2, X_n \leq a\}, \tau_4 = \min\{n \geq \tau_3 : X_n \geq b\} \\ &\vdots \\ \tau_{2k+1} &= \min\{n \geq \tau_{2k}, X_n \leq a\}, \tau_{2k+2} = \min\{n \geq \tau_{2k+1} : X_n \geq b\} \\ &\vdots \end{aligned}$$

这些  $\tau$  记录了  $X_n$  每次上下交替跳出区间  $[a, b]$  的时间。然后我们令

$$U_N^X(a, b) \stackrel{\text{def}}{=} \max\{k : \tau_{2k} \leq N\}$$

为  $N$  次之前  $\{X_n\}$  在区间  $[a, b]$  的上穿次数。

## 1.5.1 鞅收敛定理

## 定理 1.5.1

$\{X_n, n \in \mathbb{N}_+\}$  是一个上鞅, 则有以下估计: 给定  $N, j \in \mathbb{N}_+, N > j$ ,

(1).

$$\mathbb{P}(U_N^X(a, b) > j) \leq \frac{1}{b-a} \int_{\{U_N^X(a, b) = j\}} (X_n - a)^- d\mathbb{P}$$

(2).

$$\mathbb{E}[U_N^X(a, b)] \leq \frac{1}{b-a} \int \mathbb{E}[(X_n - a)^-]$$

其中  $(X_n - a)^-$  代表  $X_n - a$  的负部。

**证明** 不妨  $a = 0$ , 注意我们本题的研究范围是  $N$  次之前, 所以不妨规定随机过程  $X = \{X_n\}$  停止在  $N$ , 即  $X_N = X_{N+1} = \dots$ , 令

$$S = \tau_{2j+1} \wedge (N+1), T = \tau_{2j+2} \wedge (N+1)$$

$S$  是第  $j+1$  次下穿的时间, 但是因为可能在  $N+1$  次之前没能下穿  $j+1$  次, 所以又与  $N+1$  取最小值;  $T$  则是第  $j+1$  次上穿的事件与  $N+1$  取最小值。注意事件  $\{U_N^X(a, b) > j\}$  代表着至少上穿了  $j+1$  次, 那么

$$\{S \leq N\} = \{\tau_{2j+1} \leq N\} \quad (\star 1)$$

所以在  $\{S \leq N\}$  上  $X_S = X_{\tau_{2j+1}}$ , 同时

$$\{T \leq N\} = \{\tau_{2j+2} \leq N\} = \{U_N^X(a, b) > j\} = \{S < N, X_T \geq b\} \quad (\star 2)$$

前两个等号比较好理解, 但是最后一个等号并不明显:<sup>1</sup>

$$X_T \geq b \Leftrightarrow \begin{cases} X_{N+1} \geq b \\ \text{or } \tau_{2j+2} \leq N \end{cases} \Leftrightarrow \begin{cases} X_N \geq b \\ \text{or } \tau_{2j+2} \leq N \end{cases}$$

<sup>1</sup>还是用人话解释一下吧:  $S < N$  就是  $\tau_{2j+1} < N$ , 即在  $N$  次之前就下穿了  $j+1$  次,  $U_N^X$  的计数是从下穿开始记录的, 所以下穿  $j+1$  次的  $j$  次间隔中已经上穿了  $j$  次, 还差一次, 这就要求  $\tau_{2j+2} \leq N$ , 而  $X_T \geq b$  等价于  $X_{N+1} \geq b$  或者  $\tau_{2j+2} \leq N$ , 而我们前面规定了  $X_N = X_{N+1}$ ,  $X_N \geq b$  意味着  $\tau_{2j+2} \leq N$ , 所以  $\tau_{2j+2} \leq N$  和  $X_T \geq b$  等价, 这就得到了  $(\star 2)$  式。

$$U_N^X(a, b) > j \Leftrightarrow \begin{cases} \tau_{2j+1} < N \\ \tau_{2j+2} \leq N \end{cases} \Leftrightarrow \begin{cases} S < N \\ \tau_{2j+2} \leq N \end{cases} \Rightarrow \begin{cases} S < N \\ X_T \geq b \end{cases}$$

$$\begin{cases} S < N \\ X_N \geq b \end{cases} \Rightarrow \begin{cases} S < N \\ \tau_{2j+2} \leq N \end{cases} \Leftrightarrow U_N^X(a, b) > j$$

从而

$$U_N^X(a, b) > j \Rightarrow \begin{cases} S < N \\ X_T \geq b \end{cases} \Leftrightarrow \begin{cases} S < N \\ X_N \geq b \text{ or } \tau_{2j+2} \leq N \end{cases} \Rightarrow \begin{cases} S < N \\ \tau_{2j+2} \leq N \end{cases} \Leftrightarrow U_N^X(a, b) > j$$

此外还有：

$$\{S < N, X_T < b\} = \{S < N, T = N + 1\} \subset \{U_N^X(a, b) = j\} \quad (\star 3)$$

因此：

$$\begin{aligned} b \cdot \mathbb{P}(U_N^X(a, b) > j) &= \int_{\{U_N^X(a, b) > j\}} b d\mathbb{P} = \int_{\{\{S < N, X_T \geq b\}\}} b d\mathbb{P} \\ &\leq \int_{\{\{S < N, X_T \geq b\}\}} X_T d\mathbb{P} \\ &= \int_{\{\{S < N\}\}} X_T d\mathbb{P} - \int_{\{\{S < N, X_T < b\}\}} X_T d\mathbb{P} \\ &= \int_{\{\{S < N\}\}} X_T d\mathbb{P} - \int_{\{\{S < N, T = N + 1\}\}} X_T d\mathbb{P} \\ &= \int_{\{\{S < N\}\}} X_T d\mathbb{P} - \int_{\{\{S < N, T = N + 1\}\}} X_{N+1} d\mathbb{P} \\ &= \int_{\{\{S < N\}\}} X_T d\mathbb{P} - \int_{\{\{S < N, T = N + 1\}\}} X_N d\mathbb{P} \end{aligned} \quad (\star 4)$$

(\star 4) 式是一个阶段性成果！我们先放着。下面我们证明：

$$\int_{\{S < N\}} X_T d\mathbb{P} \leq \int_{\{S < N\}} X_S d\mathbb{P} \quad (\star 5)$$

因为  $\{S < N\} = \bigcup_{n=0}^{N-1} \{S = n\}$ ，只需证明：

$$\mathbb{E}[X_T I_{\{S=n\}}] = \mathbb{E}[X_S I_{\{S=n\}}], \forall n \leq N - 1 \quad (\star 6)$$

令  $Y_n = X_{T \wedge n} - X_{S \wedge n}$ ，那么容易验证：

$$Y_n - Y_{n-1} = I_{\{T \geq n > S\}}(X_n - X_{n-1})$$

从而

$$\begin{aligned} \mathbb{E}[Y_n | \mathcal{F}_{n-1}] &= \mathbb{E}[Y_{n-1} + I_{\{T \geq n > S\}}(X_n - X_{n-1}) | \mathcal{F}_{n-1}] \\ &= Y_{n-1} + I_{\{T \geq n > S\}} \mathbb{E}[X_n - X_{n-1} | \mathcal{F}_{n-1}] \leq Y_{n-1} \end{aligned}$$

这里用到了  $X$  是上鞅。于是  $\forall n \leq N - 1$ ， $\mathbb{E}[Y_{N+1} I_{\{S=n\}}] \leq \mathbb{E}[Y_n I_{\{S=n\}}]$ ，此时  $Y_n$  只能为 0，从而

$$\mathbb{E}[Y_{N+1} I_{\{S=n\}}] = \int_{\{S=n\}} (X_T - X_S) d\mathbb{P} \leq 0$$

因此 (\star 5) 式得证。由 (\star 4)(\star 5) 可得

$$\begin{aligned} b \cdot \mathbb{P}(U_N^X(a, b) > j) &\leq \int_{\{\{S < N\}\}} X_T d\mathbb{P} - \int_{\{\{S < N, T = N + 1\}\}} X_N d\mathbb{P} \\ &\leq 0 - \int_{\{\{S < N, X_T < b\}\}} X_{N+1} d\mathbb{P} \\ &= - \int_{\{\{S < N, X_T < b\}\}} X_N d\mathbb{P} \end{aligned}$$

$$\begin{aligned} &\leq \int_{\{S < N, T = N+1\}} X_N^- d\mathbb{P} \\ &\leq \int_{\{U_N^X(0,b)=j\}} X_N^- d\mathbb{P} \end{aligned}$$

至此定理的 (1) 得证, 关于 (2):

$$\mathbb{E}[U_N^X(a,b)] = \sum_{j=0}^{\infty} \mathbb{P}(U_N^X(a,b) > j) \leq \frac{1}{b} \sum_{j=0}^{\infty} \int_{\{U_N^X(a,b)=j\}} X_N^- d\mathbb{P} = \frac{1}{b} \int X_N^- d\mathbb{P} = \frac{1}{b} \mathbb{E}[X_N^-]$$

□

有了上述估计, 就可以得到下面鞅的收敛性定理。

**定理 1.5.2 (鞅收敛定理, Martingale Convergence Theorem)**

$\{X_n\}$  是一个上鞅, 且  $\sup_n \mathbb{E}[X_n^-] < +\infty$ , 则极限  $X(\omega) = \lim_{n \rightarrow \infty} X_n(\omega)$  a.s. 存在。

**证明** 令  $U^X(a,b) \stackrel{\text{def}}{=} \lim_{N \rightarrow \infty} U_N^X(a,b)$ , 如果发散则定义为  $+\infty$ , 由于  $U_N^X$  关于  $n$  是个单调递增序列, 所以可以使用 MCT:

$$\mathbb{E}[U^X(a,b)] = \lim_{N \rightarrow \infty} \mathbb{E}[U_N^X(a,b)] \leq \sup_N \frac{1}{b-a} \mathbb{E}[(X_N - a)^-] < \infty$$

因此  $W_{a,b} = \{U^X(a,b) = +\infty\}$  是个零测集,

$$V_{a,b} = \{\liminf_{n \rightarrow \infty} X_n(\omega) < a, \limsup_{n \rightarrow \infty} X_n(\omega) > b\} \subset W_{a,b}$$

也是个零测集, 而

$$\{X_n \text{ 不收敛}\} = \{\liminf_{n \rightarrow \infty} X_n(\omega) < \limsup_{n \rightarrow \infty} X_n(\omega)\} = \bigcup_{a < b, a, b \in \mathbb{Q}} V_{a,b}$$

所以  $X_n$  a.s. 收敛。

□

**1.5.2 推论与应用**

下鞅的版本:

**推论 1.5.1**

如果  $\{X_n\}$  是一个下鞅, 且  $\sup_n \mathbb{E}[X_n^+] < +\infty$ , 则  $X_n$  a.s. 收敛。

♡

**证明** 考虑  $Y_n = -X_n$ ,  $Y_n$  的负部即为  $X_n$  的正部, 所以  $Y_n$  a.s. 收敛, 进而  $X_n$  a.s. 收敛。

□

$L^1$  度量一致有界的鞅的极限是  $L^1$  可积的:

**推论 1.5.2**

$\{X_n\}$  是鞅, 且  $\sup_n \mathbb{E}[|X_n|] < \infty$ , 则  $X_n$  a.s. 收敛到  $X$ , 且  $X$  可积。

♡

下面这道题展示了如何利用停时构造收敛鞅。

**例 1.5.1.**

$\{X_n\}$  是鞅, 存在  $M > 0$  使得  $|X_{n+1} - X_n| \leq M < \infty$ , 令

$$C = \{\omega : \lim_{n \rightarrow \infty} X_n(\omega) \text{ 存在}\}$$

$$D = \{\omega : \limsup_{n \rightarrow \infty} X_n(\omega) = +\infty, \liminf_{n \rightarrow \infty} X_n(\omega) = -\infty\}$$

证明:  $\mathbb{P}(C \cup D) = 1$ .

**证明** 不妨假设  $X_0 = 0$ , 否则考虑  $\{X_n - X_0\}$ . 对于  $\forall k \in \mathbb{N}_+$ , 令  $N_k = \inf\{n : X_n \leq -k\}$ , 为首次落在  $(-\infty, -k]$  的时刻, 则  $\{X_{n \wedge N_k}\}$  成为一个新的鞅:

$$(X_1, X_2, \dots, X_{N_k-1}, X_{N_k}, X_{N_k}, \dots)$$

题目条件表明相邻的两项之差不超过  $M$ , 所以  $X_{N_k} \geq -k - M$ , 进而  $X_{n \wedge N_k} \geq -k - M$ , 有下界就表明  $\sup_n \mathbb{E}[X_{n \wedge N_k}^-] < +\infty$ , 所以由鞅收敛定理 1.5.2 可知  $X_{n \wedge N_k}$  a.s. 收敛, 注意在事件  $\{N_k = +\infty\}$  上  $X_{n \wedge N_k} = X_n$ , 故  $X_n$  在事件  $\{N_k = +\infty\}$  上 a.s. 收敛, 而

$$\{\liminf_{n \rightarrow \infty} X_n(\omega) > -\infty\} = \bigcup_k \{N_k = +\infty\}$$

所以  $X_n$  在事件  $\{\liminf_{n \rightarrow \infty} X_n(\omega) > -\infty\}$  上 a.s. 收敛, 同理考虑  $-X_n$ , 可得  $X_n$  在事件  $\{\limsup_{n \rightarrow \infty} X_n(\omega) < +\infty\}$  上 a.s. 收敛, 于是  $X_n$  在事件  $D$  的补集上 a.s. 收敛, 这说明  $\mathbb{P}(C \cup D) = 1$ .  $\square$

相关: 第一次作业中有相关应用。

## 1.6 鞅的 $L^p$ 收敛性

### 引理 1.6.1

$\{X_n\}$  是一个下鞅,  $N$  是停时, 且存在  $k \in \mathbb{N}_+$  使得  $N \leq k$  a.s., 则  $\mathbb{E}[X_0] \leq \mathbb{E}[X_N] \leq \mathbb{E}[X_k]$ .

**证明** 乍一看这个结论似乎很显然, 但是要注意,  $N$  是随机变量, 并不是一个固定的数字。

由定理 1.3.1,  $\{X_{m \wedge N}, m \in \mathbb{N}\}$  也是一个下鞅, 所以

$$\mathbb{E}[X_{0 \wedge N}] \leq \mathbb{E}[X_{k \wedge N}]$$

也就是

$$\mathbb{E}[X_0] \leq \mathbb{E}[X_N]$$

对于另一边, 拆分  $\Omega = \bigsqcup_{m=0}^k \{\omega : N(\omega) = m\}$  a.s., 所以

$$\begin{aligned} \mathbb{E}[X_N] &= \mathbb{E}\left[\sum_{m=0}^k X_N I_{\{N=m\}}\right] \\ &= \sum_{m=0}^k \mathbb{E}[X_m I_{\{N=m\}}] \\ &\leq \sum_{m=0}^k \mathbb{E}[\mathbb{E}[X_k | \mathcal{F}_m] I_{\{N=m\}}] \\ &= \sum_{m=0}^k \mathbb{E}[\mathbb{E}[X_k I_{\{N=m\}} | \mathcal{F}_m]] \\ &\leq \sum_{m=0}^k \mathbb{E}[X_k I_{\{N=m\}}] \\ &= \mathbb{E}[X_k] \end{aligned} \quad (*)$$

(\*) 是比较关键的一步, 因为  $\mathbb{E}[X_n] \leq \mathbb{E}[X_k]$  并不能直接推出  $\mathbb{E}[X_n I_A] \leq \mathbb{E}[X_k I_A]$ .  $\square$

### 定理 1.6.1 (Doob's Inequality)

$X = \{X_m, m \in \mathbb{N}\}$  是下鞅,  $A = \{\max_{0 \leq m \leq n} X_m^+ \geq \lambda\}, \lambda > 0$ , 则

$$\lambda \mathbb{P}(A) \leq \mathbb{E}[X_n \cdot I_A] \leq \mathbb{E}[X_n^+ I_A] \leq \mathbb{E}[X_n^+]$$

**证明** 后两个不等号是显然的。取停时  $N = \inf\{m : X_m \geq \lambda\} \wedge n$ , 注意到事件  $A$  代表着在  $n$  时刻之前  $X_m$  超过

了  $\lambda$ , 所以  $X_N \geq \lambda$ , 因此

$$\lambda \mathbb{P}(A) = \mathbb{E}[\lambda I_A] \leq \mathbb{E}[X_N I_A]$$

由引理 1.6.1 可知  $\mathbb{E}[X_N] \leq \mathbb{E}[X_n]$ , 注意到  $A^c = \{X_m^+ < \lambda, 0 \leq m \leq n\}$ , 在事件  $A^c$  上  $N = n$ , 所以  $\mathbb{E}[X_N I_{A^c}] = \mathbb{E}[X_n I_{A^c}]$ , 因此

$$\mathbb{E}[X_N I_A] = \mathbb{E}[X_N] - \mathbb{E}[X_N I_{A^c}] = \mathbb{E}[X_N] - \mathbb{E}[X_n I_{A^c}] \leq \mathbb{E}[X_n] - \mathbb{E}[X_n I_{A^c}] = \mathbb{E}[X_n I_A]$$

□

于是我们给出了  $\mathbb{P}(A)$  的上界估计。

### 定理 1.6.2 ( $L^p$ Maximum Inequality)

$X = \{X_m\}$  是下鞅, 记  $\bar{X}_n = \max_{0 \leq m \leq n} X_m^+$ , 则

$$\mathbb{E}[\bar{X}_n^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[(X_n^+)^p]$$

**证明** 注意到  $\forall M > 0$ ,  $\{\bar{X}_n \wedge M \geq \lambda\}$  要么是  $\{\bar{X}_n \geq \lambda\}$ , 要么是  $\emptyset$ .

$$\begin{aligned} \mathbb{E}[(\bar{X}_n \wedge M)^p] &= \mathbb{E}\left[\int_0^{\bar{X}_n \wedge M} p\lambda^{p-1} d\lambda\right] \\ &= \mathbb{E}\left[\int_0^{+\infty} I_{\{\bar{X}_n \wedge M \geq \lambda\}} p\lambda^{p-1} d\lambda\right] \\ &= \int_0^{+\infty} p\lambda^{p-1} \mathbb{E}[I_{\{\bar{X}_n \wedge M \geq \lambda\}}] d\lambda && \text{By Fubini Thm} \\ &= \int_0^{+\infty} p\lambda^{p-1} \mathbb{P}(\bar{X}_n \wedge M \geq \lambda) d\lambda \\ &\leq \int_0^{+\infty} p\lambda^{p-1} \frac{1}{\lambda} \int X_n^+ I_{\bar{X}_n \wedge M \geq \lambda} d\mathbb{P} d\lambda && \text{By 定理 1.6.1} \\ &= \int X_n^+ \int_0^{\bar{X}_n \wedge M} p\lambda^{p-2} d\lambda d\mathbb{P} && \text{By Fubini Thm} \\ &= \frac{p}{p-1} \int X_n^+ (\bar{X}_n \wedge M)^{p-1} d\mathbb{P} \\ &= \frac{p}{p-1} \mathbb{E}[X_n^+ (\bar{X}_n \wedge M)^{p-1}] \\ &\leq \frac{p}{p-1} \mathbb{E}[|X_n^+|^p]^{\frac{1}{p}} \mathbb{E}[|\bar{X}_n \wedge M|^p]^{\frac{p-1}{p}} && \text{By Holder Inequality} \end{aligned}$$

令  $M \rightarrow \infty$ , 则得到

$$\mathbb{E}[\bar{X}_n^p] \leq \frac{p}{p-1} \mathbb{E}[|X_n^+|^p]^{\frac{1}{p}} \mathbb{E}[|\bar{X}_n|^p]^{\frac{p-1}{p}}$$

整理即可得到结论。 □

### 推论 1.6.1

$\{X_n, n \in \mathbb{N}\}$  是鞅,  $\sup_n \mathbb{E}[|X_n|^p] < +\infty, p > 1$ , 则

$$\mathbb{E}\left[\max_{0 \leq m \leq n} |X_m|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_n|^p]$$

♥

**证明** 对  $X_n, -X_n$  都使用定理 1.6.2,

$$\begin{aligned} \mathbb{E}\left[\max_{0 \leq m \leq n} (X_m^+)^p\right] &\leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[(X_n^+)^p] \\ \mathbb{E}\left[\max_{0 \leq m \leq n} (X_m^-)^p\right] &\leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[(X_n^-)^p] \end{aligned}$$

注意到  $|X_n| = X_n^+ I_{\{X_n > 0\}} + X_n^- I_{\{X_n < 0\}}$ , 两式相加即可。 □

**定理 1.6.3 ( $L^p$  收敛定理)**

$\{X_n, n \in \mathbb{N}\}$  是鞅,  $\sup_n \mathbb{E}[|X_n|^p] < +\infty, p > 1$ , 则  $X_n \xrightarrow{\text{a.s.}} X$  且  $X_n \xrightarrow{L^p} X$ .

**证明** 根据 Jensen 不等式,  $\mathbb{E}[|X_n|] \leq \mathbb{E}[|X_n|^p]^{\frac{1}{p}} < +\infty$ , 由鞅收敛定理即可得 a.s. 收敛。

由推论 1.6.1,

$$\mathbb{E}[\max_{0 \leq m \leq n} |X_m|^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_n|^p]$$

由  $n$  的任意性, 可得

$$\mathbb{E}[\sup_n |X_n|^p] \leq \left(\frac{p}{p-1}\right)^p \sup_n \mathbb{E}[|X_n|^p] < +\infty$$

这说明  $\sup_n |X_n|$  是  $L^p$  可积的, 由控制收敛定理可得  $L^p$  收敛。 □

## 1.7 鞅的择停定理

### 1.7.1 回顾: 一致可积性

**定义 1.7.1**

称一族 r.v.  $\{X_i, i \in I\}$  是一致可积的 (Uniformly Integrable, U.I.), 如果

$$\lim_{M \rightarrow \infty} \sup_{i \in I} \mathbb{E}[|X_i|; |X_i| > M] = 0$$

这里

$$\mathbb{E}[X; A] \stackrel{\text{def}}{=} \mathbb{E}[X I_A] = \int_A X d\mathbb{P}$$

**引理 1.7.1 (积分的绝对连续性)**

如果 r.v.  $X$  可积, 则

$$\lim_{\mathbb{P}(A) \rightarrow 0} \int_A |X| d\mathbb{P} = 0$$

**证明**  $X$  是可积的, 则  $X I_{|X| > M} \xrightarrow{\mathbb{P}} 0$ , 因为  $|X I_{|X| > M}| \leq |X|$ , 由 DCT 可知

$$\int |X I_{|X| > M}| d\mathbb{P} \rightarrow 0$$

即  $\forall \varepsilon, M$  充分大时

$$\int |X I_{|X| > n}| d\mathbb{P} < \frac{1}{2} \varepsilon$$

取  $\delta = \frac{\varepsilon}{2M}$ ,  $\mathbb{P}(A) < \delta$  时,

$$\begin{aligned} \int_A |X| d\mathbb{P} &= \int_A |X| I_{\{|X| \geq M\}} + \int_A |X| I_{\{|X| < M\}} d\mathbb{P} \\ &\leq \frac{\varepsilon}{2} + \int_A M I_{\{|X| < M\}} d\mathbb{P} \\ &\leq \frac{\varepsilon}{2} + \mathbb{P}(A) \cdot M = \varepsilon \end{aligned}$$

□

**推论 1.7.1**

一致有界的  $\{X_i, i \in I\}$  是一致可积的。 ♥

**证明** 取充分大的  $M$  使得  $\{|X_i| \geq M\} = \emptyset, \forall i$  即可。  $\square$

### 推论 1.7.2

概率空间  $(\Omega, \mathbb{P}, \mathcal{F})$  上的随机变量  $X$  可积, 考虑 r.v. 族  $\{\mathbb{E}[X|\mathcal{G}] : \mathcal{G} \subset \mathcal{F}\}$ , 则是一致可积的。  $\heartsuit$

**证明** 设  $M > 0$ , 令事件  $A = \{|\mathbb{E}[X|\mathcal{G}]| > M\} = \left\{\frac{|\mathbb{E}[X|\mathcal{G}]|}{M} > 1\right\}$ ,

$$\mathbb{P}(A) = \int_A 1 d\mathbb{P} \leq \int_A \frac{|\mathbb{E}[X|\mathcal{G}]|}{M} d\mathbb{P} \leq \int \frac{|\mathbb{E}[X|\mathcal{G}]|}{M} d\mathbb{P} = \frac{\mathbb{E}[|X|]}{M}$$

由引理 1.7.1,  $\forall \varepsilon > 0, \exists \delta > 0$ , 只要取足够大的  $M$  使得  $\mathbb{P}(A) \leq \mathbb{E}[|X|]/M \leq \delta$ , 就有

$$\int_A |X| d\mathbb{P} < \varepsilon$$

又根据条件期望的定义,  $A \in \mathcal{G}$ , 进而

$$\int_A \mathbb{E}[|X| | \mathcal{G}] d\mathbb{P} = \int_A |X| d\mathbb{P}$$

于是

$$\varepsilon \geq \int_A |X| d\mathbb{P} = \int_A \mathbb{E}[|X| | \mathcal{G}] d\mathbb{P} \geq \int_A |\mathbb{E}[X|\mathcal{G}]| d\mathbb{P}$$

这就证明了一致可积性。  $\square$

### 定理 1.7.1

$\forall \varphi \geq 0$ , 且  $\frac{\varphi(x)}{x} \rightarrow \infty$  as  $x \rightarrow \infty$ , 如果  $\sup_i \mathbb{E}[\varphi(X_i)] < +\infty$ , 则  $\{X_i, i \in I\}$  是一致可积的。

**证明** 设  $M > 0$ , 令

$$\varepsilon_M = \sup \left\{ \frac{x}{\varphi(x)} : x \geq M \right\}$$

则在  $A = \{|X_i| > M\}$  上,

$$\frac{X_i}{\varphi(X_i)} \leq \varepsilon_M$$

因为  $x \rightarrow \infty$  时  $\frac{x}{\varphi(x)} \rightarrow 0$ , 所以  $M \rightarrow \infty$  时  $\varepsilon_M \rightarrow 0$ . 对于  $\forall i \in I$ ,

$$\mathbb{E}[|X_i| I_A] = \int_A \frac{X_i}{\varphi(X_i)} \cdot \varphi(X_i) d\mathbb{P} \leq \varepsilon_M \int_A \varphi(X_i) d\mathbb{P} \leq \varepsilon_M \mathbb{E}[\varphi(X_i)]$$

则

$$\sup_i \mathbb{E}[|X_i| I_A] \leq \varepsilon_M \sup_i \mathbb{E}[\varphi(X_i)] \rightarrow 0$$

所以一致可积性得证。  $\square$

**注** 取  $\varphi(x) = |x|^p, p > 1$  是一个很常见的应用。

### 定理 1.7.2 (Durrett Theorem 5.5.2)

$\mathbb{E}[|X_n|] < +\infty, \forall n$ , 若  $X_n \xrightarrow{\mathbb{P}} X$ , 以下命题等价:

- (1)  $\{X_n, n \geq 0\}$  一致可积;
- (2)  $X_n \xrightarrow{L^1} X$ ;
- (3)  $\mathbb{E}[|X_n|] \rightarrow \mathbb{E}[|X|] < +\infty$ .

**证明** 这段证明用到的前文结论太多, 暂时无力整理, 等前三章补充完整了再说吧。  $\square$

## 1.7.2 一致可积鞅

## 定理 1.7.3

对于一个下鞅  $\{X_n\}$ , 以下命题等价:

- (1).  $\{X_n, n \geq 0\}$  一致可积;
- (2).  $X_n \xrightarrow{L^1} X$  且  $X_n \xrightarrow{\text{a.s.}} X$ ;
- (3).  $X_n \xrightarrow{L^1} X$ .

**证明** 我们利用定理 1.7.2 的结论证明该定理。

(1)  $\Rightarrow$  (2): 一致可积意味着  $\sup_n \mathbb{E}[|X_n|] < +\infty$ , 利用鞅收敛定理可知  $X_n \xrightarrow{\text{a.s.}} X$ , 进而  $X_n \xrightarrow{\mathbb{P}} X$ , 于是由定

理 1.7.2(1)  $\Rightarrow$  (2) 知  $X_n \xrightarrow{L^1} X$ .

(2)  $\Rightarrow$  (3): 显然。

(3)  $\Rightarrow$  (1):  $X_n \xrightarrow{L^1} X$  意味着  $X_n \xrightarrow{\mathbb{P}} X$ , 于是由定理 1.7.2(2)  $\Rightarrow$  (1) 则得证。  $\square$

## 推论 1.7.3

$\{X_n\}$  是 U.I. 鞅,  $X_n \rightarrow X$ , 则  $X_n = \mathbb{E}[X|\mathcal{F}_n], \forall n$ .

**证明** 先回顾一个测度论里的结论:

## 引理 1.7.2

$X_n \in L^1, X_n \xrightarrow{L^1} X$ , 则  $\forall A \subset \Omega$ ,

$$\mathbb{E}[X_n I_A] \rightarrow \mathbb{E}[X I_A]$$

**证明**

$$\mathbb{E}[X_n I_A - X I_A] \leq \mathbb{E}[|X_n - X| I_A] \leq \mathbb{E}[|X_n - X|] = 0$$

$\square$  

设  $X_n \xrightarrow{\text{a.s. and } L^1} X$ , 考虑  $m \geq n$ ,  $X_n = \mathbb{E}[X_m|\mathcal{F}_n]$ , 因此若  $A \in \mathcal{F}_n$ , 就有

$$\mathbb{E}[X_m I_A] = \int_A \mathbb{E}[X_m|\mathcal{F}_n] d\mathbb{P} = \mathbb{E}[X_n I_A]$$

由引理可知,  $\mathbb{E}[X_n I_A] \rightarrow \mathbb{E}[X I_A]$ , 所以  $\forall n$  有  $\mathbb{E}[X I_A] = \mathbb{E}[X_n I_A]$ , 这意味着  $X_n = \mathbb{E}[X|\mathcal{F}_n]$ .  $\square$

## 推论 1.7.4

$\{X_n\}$  是 U.I. 下鞅,  $X_n \rightarrow X$ , 则  $X_n \leq \mathbb{E}[X|\mathcal{F}_n], \forall n$ .

**证明** 没办法直接验证条件期望了, 但测度论中有一个结论:

## 引理 1.7.3

$(\Omega, \mathcal{F}, \mathbb{P})$  上,  $\mathcal{G} \subset \mathcal{F}$ , 两个  $\mathcal{G}$ -可测的随机变量  $X, Y$  满足  $\forall A \in \mathcal{G}, \mathbb{E}[X I_A] \leq \mathbb{E}[Y I_A]$ , 则  $X \leq Y$  a.s. 

于是只需验证  $\forall n, A \in \mathcal{F}_n$  有  $\mathbb{E}[X_n I_A] \leq \mathbb{E}[\mathbb{E}[X|\mathcal{F}_n] I_A] = \mathbb{E}[X I_A]$  即可, 之后的证明过程和上一个推论同理。  $\square$

## 定理 1.7.4 (条件期望随信息演化而收敛)

设  $\mathcal{F}_n \nearrow \mathcal{F}_\infty$ , 则当  $n \rightarrow \infty$  时,

$$\mathbb{E}[X|\mathcal{F}_n] \xrightarrow{\text{a.s. and } L^1} \mathbb{E}[X|\mathcal{F}_\infty]$$

$\mathcal{F}_n \nearrow \mathcal{F}_\infty$  的意思是  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$ , 并且存在  $\mathcal{F}_\infty \stackrel{\text{def}}{=} \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n)$ .

**证明** 注意到  $Y_n = \mathbb{E}[X|\mathcal{F}_n]$  是一致可积鞅, 设  $Y_n \rightarrow Y_\infty$  a.s. 和  $L^1$ , 并且

$$\mathbb{E}[X|\mathcal{F}_n] = Y_n = \mathbb{E}[Y_\infty|\mathcal{F}_n]$$

因此  $\forall A \in \mathcal{F}_n$ ,

$$\mathbb{E}[Y_n I_A] = \mathbb{E}[X I_A] = \mathbb{E}[Y_\infty I_A]$$

考虑到  $\bigcup_n \mathcal{F}_n$  是一个  $\pi$ -系, 根据  $\pi$ - $\lambda$  定理可知对于任意的  $A \in \mathcal{F}_\infty$  上式都成立, 又因为  $Y_\infty$  是  $\mathcal{F}_\infty$ -可测的<sup>2</sup>, 所以  $Y_\infty = \mathbb{E}[X|\mathcal{F}_\infty]$ .  $\square$

### 推论 1.7.5 (Levy 0-1 律)

设  $\mathcal{F}_n \nearrow \mathcal{F}_\infty$ , 事件  $A \in \mathcal{F}$ , 则  $\mathbb{E}[I_A|\mathcal{F}_n] \rightarrow I_A$  a.s.



### 定理 1.7.5 (条件期望的控制收敛定理 (对角线 ver.))

设  $\mathcal{F}_n \nearrow \mathcal{F}_\infty$ ,  $Y_n \xrightarrow{\text{a.s.}} Y$ ,  $|Y_n| \leq Z, \forall n$ ,  $Z$  可积, 则  $\mathbb{E}[Y_n|\mathcal{F}_n] \rightarrow \mathbb{E}[Y|\mathcal{F}_\infty]$ .

**证明** 令  $W_n = \sup\{|Y_n - Y_m| : \forall n, m \geq N\}$ , 则  $W_N \leq 2Z$ , 进而  $\mathbb{E}[W_N] < +\infty$ , 由定理 1.7.4,

$$\limsup_{n \rightarrow \infty} \mathbb{E}[|Y_n - Y| | \mathcal{F}_n] \leq \limsup_{n \rightarrow \infty} \mathbb{E}[W_N | \mathcal{F}_n] = \mathbb{E}[W_N | \mathcal{F}_\infty]$$

因为  $N \rightarrow \infty$  时  $\mathbb{E}[W_N | \mathcal{F}_\infty] \searrow 0$ , 以及  $\mathbb{E}[Y|\mathcal{F}_n] \rightarrow \mathbb{E}[Y|\mathcal{F}_\infty]$ , 所以  $\mathbb{E}[Y_n|\mathcal{F}_n] \rightarrow \mathbb{E}[Y|\mathcal{F}_\infty]$ .  $\square$

### 定理 1.7.6 (Doob's Optional Stopping Theorem, 择停定理)

$\{X_n, n \in \mathbb{N}\}$  是一致可积下鞅, 对于任意停时  $N$ ,  $\{X_{n \wedge N}\}$  是一致可积的。

**证明** 下鞅  $\Rightarrow \mathbb{E}[X_{N \wedge n}^+] \leq \mathbb{E}[X_n^+]$ , 因此

$$\sup_n \mathbb{E}[X_{N \wedge n}^+] \leq \sup_n \mathbb{E}[X_n^+] \leq \sup_n \mathbb{E}[|X_n|] < +\infty$$

另一方面,

$$\mathbb{E}[X_{N \wedge n}^-] = \mathbb{E}[X_{N \wedge n}^+] - \mathbb{E}[X_{N \wedge n}] \leq \mathbb{E}[X_{N \wedge n}^+] - \mathbb{E}[X_0]$$

所以

$$\sup_n \mathbb{E}[X_{N \wedge n}^-] \leq \sup_n \mathbb{E}[X_{N \wedge n}^+] - \mathbb{E}[X_0] < +\infty$$

从而  $\sup_n \mathbb{E}[|X_{N \wedge n}|] < +\infty$ , 由鞅收敛定理可知  $X_{N \wedge n} \xrightarrow{\text{a.s.}} X_N$  且  $\mathbb{E}[|X_N|] < +\infty$ , 于是

$$\begin{aligned} \mathbb{E}[|X_{N \wedge n}| I_{\{|X_{N \wedge n}| > k\}}] &= \mathbb{E}[|X_{N \wedge n}| I_{\{|X_{N \wedge n}| > k, N \leq n\}}] + \mathbb{E}[|X_{N \wedge n}| I_{\{|X_{N \wedge n}| > k, N > n\}}] \\ &= \mathbb{E}[|X_N| I_{\{|X_N| > k, N \leq n\}}] + \mathbb{E}[|X_n| I_{\{|X_n| > k, N > n\}}] \\ &\leq \mathbb{E}[|X_N| I_{\{|X_N| > k\}}] + \mathbb{E}[|X_n| I_{\{|X_n| > k\}}] \end{aligned}$$

$X_N$  可积,  $\{X_n\}$  一致可积, 所以这两项都趋于 0, 故

$$\lim_{k \rightarrow \infty} \mathbb{E}[|X_{N \wedge n}| I_{\{|X_{N \wedge n}| > k\}}] = 0$$

$\square$

<sup>2</sup>这也是测度论里的一个结论: 显然  $Y_n$  都是  $\mathcal{F}_\infty$ -可测的, 而可测函数类对于极限运算封闭, 见实分析笔记定理 1.4.2, 后续记得补上。

## 1.7.3 择停定理的推论与应用

## 定理 1.7.7

$\{X_n\}$  是一个下鞅,  $N$  是停时, 如果  $X_N$  可积且  $\{X_n I_{\{N>n\}}, n \in \mathbb{N}\}$  一致可积, 则  $\{X_{N \wedge n}\}$  一致可积, 从而  $\mathbb{E}[X_0] \leq \mathbb{E}[X_N]$ .

**证明** 设  $Y_n = X_{N \wedge n}$ , 则

$$\begin{aligned} \mathbb{E}[Y_n; |Y_n| > M] &= \int_{\{|Y_n| > M\}} |Y_n| d\mathbb{P} \\ &= \int_{\{|X_n| > M, N > n\}} |X_n| d\mathbb{P} + \int_{\{|X_N| > M, N \leq n\}} |X_N| d\mathbb{P} \\ &\leq \int_{\{|X_n| > M\}} |X_n I_{\{N > n\}}| d\mathbb{P} + \int_{\{|X_N| > M\}} |X_N| d\mathbb{P} \end{aligned}$$

$M \rightarrow \infty$  时这两项都趋于 0, 前面是因为  $\{X_n I_{\{N > n\}}, n \in \mathbb{N}\}$  一致可积的定义, 后面是因为积分的绝对连续性.

因为  $\{X_n\}$  是一个下鞅,  $\{Y_n = X_{N \wedge n}\}$  就是一致可积下鞅, 因此  $Y_n \xrightarrow{\text{a.s. and } L^1} Y_\infty, \forall n$  有  $\mathbb{E}[Y_0] \leq \mathbb{E}[Y_n] \rightarrow \mathbb{E}[Y_\infty]$ , 所以  $\mathbb{E}[X_0] \leq \mathbb{E}[X_N]$ .  $\square$

## 定理 1.7.8

$\{X_n\}$  是一致可积下鞅,  $X_n \rightarrow X_\infty$ , 则对于任意停时  $N$  有

$$\mathbb{E}[X_0] \leq \mathbb{E}[X_N] \leq \mathbb{E}[X_\infty]$$

**证明**  $\{Y_n = X_{n \wedge N}\}$  是一致可积下鞅, 设  $Y_n \rightarrow Y_\infty, \mathbb{E}[Y_0] \leq \mathbb{E}[Y_\infty]$  可得  $\mathbb{E}[X_0] \leq \mathbb{E}[X_N]$ .

另一方面, 考虑

$$\begin{aligned} \mathbb{E}[X_N] &= \sum_{n=0}^{\infty} \mathbb{E}[X_n I_{\{N=n\}}] \\ &\leq \sum_{n=0}^{\infty} \mathbb{E}[\mathbb{E}[X_\infty | \mathcal{F}_n] I_{\{N=n\}}] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[\mathbb{E}[X_\infty I_{\{N=n\}} | \mathcal{F}_n]] \\ &= \sum_{n=0}^{\infty} \mathbb{E}[X_\infty I_{\{N=n\}} | \mathcal{F}_n] \\ &= \mathbb{E}[X_\infty] \end{aligned}$$

这和引理 1.6.1 的证明非常类似.  $\square$

## 推论 1.7.6

$\{X_n\}$  是一致可积下鞅,  $M \leq N$  是两个停时, 则  $\mathbb{E}[X_0] \leq \mathbb{E}[X_M] \leq \mathbb{E}[X_N]$ . ♥

**证明** 考虑  $Y_n = X_{n \wedge N}$  是一致可积下鞅, 则

$$\mathbb{E}[Y_0] \leq \mathbb{E}[Y_M] \leq \mathbb{E}[Y_\infty]$$

相对应的就是

$$\mathbb{E}[X_0] \leq \mathbb{E}[X_M] \leq \mathbb{E}[X_N]$$

$\square$

## 推论 1.7.7

$\{X_n\}$  是一致可积鞅,  $M \leq N$  是两个停时, 则  $\mathbb{E}[X_0] = \mathbb{E}[X_M] = \mathbb{E}[X_N]$ . ♥

证明 同上。 □

**推论 1.7.8**

$\{X_n\}$  是一致可积鞅,  $M \leq N$  是两个停时, 则  $X_M = \mathbb{E}[X_N | \mathcal{F}_M]$ . ♥

证明 回顾定义 1.3.3.  $X_M$  是  $\mathcal{F}_M$ -可测的, 只需验证  $\forall A \in \mathcal{F}_M$ ,

$$\mathbb{E}[X_M I_A] = \mathbb{E}[X_N I_A]$$

$A \in \mathcal{F}_M \subset \mathcal{F}_N$ , 考虑  $M^A \leq N^A$  这两个停时, 应用推论 1.7.7, 就得到

$$\mathbb{E}[X_{M^A}] = \mathbb{E}[X_{N^A}]$$

即

$$\mathbb{E}[X_M I_A] + \mathbb{E}[X_\infty I_{A^c}] = \mathbb{E}[X_N I_A] + \mathbb{E}[X_\infty I_{A^c}]$$

这说明  $\mathbb{E}[X_M I_A] = \mathbb{E}[X_N I_A]$ , 于是得证。 □

**定理 1.7.9**

$\{X_n\}$  是下鞅, 存在常数  $B > 0$  使得  $\mathbb{E}[|X_{n+1} - X_n| | \mathcal{F}_n] \leq B$  a.s., 如果  $N$  是停时,  $\mathbb{E}[N] < +\infty$ , 则  $\{Y_n = X_{N \wedge n}\}$  一致可积, 进而  $\mathbb{E}[X_0] \leq \mathbb{E}[X_N]$ .

证明 记

$$Y = |X_0| + \sum_{m=0}^{\infty} |X_{m+1} - X_m| I_{\{N > m\}}$$

则  $|X_{N \wedge n}| \leq Y$ , 下面只需证明  $Y$  可积即可。

$$\begin{aligned} \mathbb{E}[Y] &= \mathbb{E}[|X_0|] + \sum_{m=0}^{\infty} \mathbb{E}[|X_{m+1} - X_m| I_{\{N > m\}}] \\ &= \mathbb{E}[|X_0|] + \sum_{m=0}^{\infty} \mathbb{E}[\mathbb{E}[|X_{m+1} - X_m| I_{\{N > m\}} | \mathcal{F}_m]] \\ &= \mathbb{E}[|X_0|] + \sum_{m=0}^{\infty} \mathbb{E}[\mathbb{E}[|X_{m+1} - X_m| | \mathcal{F}_m] I_{\{N > m\}}] \\ &\leq \mathbb{E}[|X_0|] + B \cdot \sum_{m=0}^{\infty} \mathbb{E}[I_{\{N > m\}}] \\ &= \mathbb{E}[|X_0|] + B \mathbb{E}[N] < +\infty \end{aligned}$$

□

## 1.8 两个例子

**例 1.8.1.**

赌徒破产: A 和 B 玩抛硬币, 规则是这样的: 硬币均匀, 抛出正反面的概率均为  $\frac{1}{2}$ , 每轮游戏抛一次硬币, 抛出正面时 B 给 A 一块钱, 抛出反面时 A 给 B 一块钱。游戏开始前 A 有  $a$  元, B 有  $b$  元, 游戏会一直持续到一方没有钱为止, 此时另一方获得胜利。那么, A 获胜的概率是多少? 游戏的持续轮数  $T$  的期望是多少?

解答: 设  $\{X_n, n \in \mathbb{N}_+\}$  满足  $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$ , 令  $S_n = X_1 + \dots + X_n$ , 则  $S_n$  表示经过  $n$  轮游戏后 A 获得/失去的钱数,  $S_0$  定义为 0, 那么, 根据例 1.2.2 可知  $\{S_n, n \in \mathbb{N}\}$  是一个鞅。

游戏的持续轮数  $T$  的定义应当是  $S_n$  首次等于  $b$  或者  $-a$  的时刻:

$$T \stackrel{\text{def}}{=} \min\{n : S_n = -a \text{ or } S_n = b\}$$

则  $T$  是一个停时。为了证明  $\mathbb{E}[T] < +\infty$ , 首先要证明下面这个式子:

$$\mathbb{P}(T > m(a+b)) \leq \left(1 - \left(\frac{1}{2}\right)^{a+b}\right)^m, \quad m \geq 1 \quad (\star)$$

先考虑  $m=1$  的情况,

$$\mathbb{P}(T > a+b) = 1 - \mathbb{P}(T \leq a+b) \leq 1 - \mathbb{P}(X_1 = X_2 = \cdots = X_{a+b} = 1) = 1 - \left(\frac{1}{2}\right)^{a+b}$$

归纳:  $m$  时命题成立, 则考虑  $m+1$  时

$$\begin{aligned} \mathbb{P}(T > (m+1)(a+b)) &= \mathbb{P}(T > (m+1)(a+b) | T > m(a+b)) \mathbb{P}(T > m(a+b)) \\ &\leq (1 - \mathbb{P}(S_{(m+1)(a+b)} - S_{m(a+b)} = a+b)) \left(1 - \left(\frac{1}{2}\right)^{a+b}\right)^m \\ &= \left(1 - \left(\frac{1}{2}\right)^{a+b}\right)^{m+1} \end{aligned}$$

也成立, 于是  $(\star)$  式得证, 那么

$$\begin{aligned} \mathbb{E}[T] &= \mathbb{E}[TI_{\{T \leq a+b\}}] + \mathbb{E}[TI_{\{T > a+b\}}] \\ &= \mathbb{E}[TI_{\{T \leq a+b\}}] + \sum_{m=1}^{\infty} \mathbb{E}[TI_{\{m(a+b) < T \leq (m+1)(a+b)\}}] \\ &\leq a+b + \sum_{m=1}^{\infty} (m+1)(a+b) \mathbb{P}(m(a+b) < T \leq (m+1)(a+b)) \\ &\leq a+b + \sum_{m=1}^{\infty} (m+1)(a+b) \mathbb{P}(T > m(a+b)) \\ &\leq a+b + \sum_{m=1}^{\infty} (m+1)(a+b) \left(1 - \left(\frac{1}{2}\right)^{a+b}\right)^m < +\infty \end{aligned}$$

注意到  $S_{T \wedge n} \in [-a, b]$ , 有界  $\Rightarrow$  一致可积, 应用推论 1.7.7 可得

$$\mathbb{E}[S_T] = \mathbb{E}[S_{T \wedge n}] = \mathbb{E}[S_0]$$

其中  $\mathbb{E}[S_0] = 0$ ,  $\mathbb{E}[S_T] = -a\mathbb{P}(S_T = -a) + b\mathbb{P}(S_T = b)$ , 同时又因为  $\mathbb{P}(S_T = -a) + \mathbb{P}(S_T = b) = 1$ , 可解得

$$\mathbb{P}(S_T = -a) = \frac{b}{a+b}, \quad \mathbb{P}(S_T = b) = \frac{a}{a+b}$$

$A$  获胜的概率即为  $\mathbb{P}(S_T = b) = \frac{a}{a+b}$ .

为了计算  $\mathbb{E}[T]$ , 考虑另一个鞅  $\{Y_n = S_n^2 - n, n \in \mathbb{N}\}$  (见例 1.2.5), 那么  $\{Y_{T \wedge n}, n \in \mathbb{N}\}$  也是鞅, 应用推论 1.7.7 可得

$$\mathbb{E}[S_{T \wedge n}^2 - T \wedge n] = \mathbb{E}[Y_0] = 0 \Rightarrow \mathbb{E}[S_{T \wedge n}^2] = \mathbb{E}[T \wedge n]$$

令  $n \rightarrow \infty$  并应用 DCT 得  $\mathbb{E}[S_T^2] = \mathbb{E}[T] = (-a)^2\mathbb{P}(S_T = -a) + b^2\mathbb{P}(S_T = b) = ab$ .

### 例 1.8.2.

随机游走: 数轴上的原点是醉汉的家, 醉汉所处的位置为  $k$ , 位置  $N$  处有一条河,  $0 < k < N$ . 醉汉走路的方向都是随机的, 每一步向右走一个单位的概率为  $p$ , 每一步向左走一个单位的概率为  $q = 1 - p$ , 试求醉汉在掉到河里之前成功走回家的概率。

解答: 设  $\{X_n, n \in \mathbb{N}_+\}$ , 满足  $\mathbb{P}(X_i = 1) = p, \mathbb{P}(X_i = -1) = 1 - p$ , 令  $S_n = X_1 + \cdots + X_n$ , 表示醉汉走了  $n$  步

之后向右移动的距离,

$$T \stackrel{\text{def}}{=} \min\{n \geq 0 : S_T = 0 \text{ or } S_T = N\}$$

代表醉汉首次掉进河或者回家的时刻,  $\mathbb{P}(S_T = 0)$  即为题目所求。

设  $Z_n = \left(\frac{q}{p}\right)^{S_n}$ , 则  $\{Z_n, n \in \mathbb{N}\}$  是一个鞅 (见例 1.2.6), 注意到

$$\begin{aligned} |Z_{T \wedge n}| &= \left(\frac{q}{p}\right)^{S_{T \wedge n}} \leq \max_{0 \leq l \leq N} \left(\frac{q}{p}\right)^l = M \\ \mathbb{E}[Z_{T \wedge n}] &= \mathbb{E}\left[\left(\frac{q}{p}\right)^{S_{T \wedge n}}\right] = \mathbb{E}[Z_0] = \left(\frac{q}{p}\right)^k \end{aligned}$$

令  $n \rightarrow \infty$  可得

$$\mathbb{E}[Z_T] = \left(\frac{q}{p}\right)^k$$

同时

$$\mathbb{E}[Z_T] = \mathbb{P}(S_T = 0) + \left(\frac{q}{p}\right)^N \mathbb{P}(S_T = N)$$

以及  $\mathbb{P}(S_T = 0) + \mathbb{P}(S_T = N) = 1$ , 可得

$$\mathbb{P}(S_T = 0) = \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}$$

## 1.9 习题

### 1.9.1 第一次作业

#### 题目 1.1. Durrett(Exercise 4.2.3)

证明: 如果  $X_n, Y_n$  是关于  $\mathcal{F}_n$  的下鞅, 则  $X_n \vee Y_n$  也是下鞅。

解答: 令  $Z_n = X_n \vee Y_n = X_n I_{X_n \geq Y_n} + Y_n I_{X_n < Y_n}$ , 因为  $X_n, Y_n$  是  $\mathcal{F}_n$ -可测的,

$$\{X_n \geq Y_n\} = \bigcup_{q \in \mathbb{Q}} \{X_n \geq q\} \cup \{q \geq Y_n\} \in \mathcal{F}_n$$

于是  $Z_n$  就是  $\mathcal{F}_n$ -可测的。而

$$\begin{aligned} \mathbb{E}[X_{n+1} \vee Y_{n+1} | \mathcal{F}_n] &\geq \mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n \\ \mathbb{E}[X_{n+1} \vee Y_{n+1} | \mathcal{F}_n] &\geq \mathbb{E}[Y_{n+1} | \mathcal{F}_n] = Y_n \\ \Rightarrow \mathbb{E}[X_{n+1} \vee Y_{n+1} | \mathcal{F}_n] &\geq X_n \vee Y_n \end{aligned}$$

所以  $Z_n$  也是下鞅。

#### 题目 1.2. Durrett(Exercise 4.2.4)

$\{X_n, n \geq 0\}$  是下鞅且  $\sup_n X_n < \infty$ , 令  $\xi_n = X_n - X_{n-1}$ , 若  $\mathbb{E}[\sup_n \xi_n^+] < \infty$ , 证明:  $X_n$  a.s. 收敛。

解答: 定义  $T_m = \inf\{k \geq 0 : X_k > m\}$ , 即  $X_k$  首次超过  $m$  的时刻, 那么  $T_m$  是一个停时。考虑  $Y_n = X_{n \wedge T_m}$  是一个新的下鞅, 注意  $T_m$  是“首次”超过  $m$ , 也就是只有  $X_{T_m}$  超过了  $m$ , 所以

$$\{Y_n(\omega)^+\} = \{X_1(\omega)^+, X_2(\omega)^+, \dots, X_{T_m(\omega)-1}(\omega)^+, X_{T_m(\omega)}(\omega)^+, X_{T_m(\omega)}(\omega)^+, \dots\}$$

里最大的就是  $X_{T_m(\omega)}(\omega)^+$ , 因此

$$\sup_n Y_n^+ = \sup_n X_{n \wedge T_m}^+ \leq X_{T_m}^+ = (X_{T_m-1} + \xi_{T_m})^+ \leq X_{T_m-1}^+ + \xi_{T_m}^+ \leq m + \sup_n \xi_n^+$$

于是

$$\mathbb{E}[\sup_n Y_n^+] \leq m + \mathbb{E}[\sup_n \xi_n^+] < +\infty$$

于是  $Y_n$  a.s. 收敛。注意在  $\{T_m = +\infty\}$  上  $X_n = Y_n$ , 因此  $X_n$  在  $\{T_m = +\infty\}$  上 a.s. 收敛, 而

$$\{T_m = +\infty\} = \{\forall k \geq 0, X_k \leq m\}$$

考虑到  $\sup_n X_n < \infty$ , 存在某个  $m$  使得  $\forall X_n \leq m$ , 这表明  $\{T_m = +\infty\} = \Omega$ , 于是  $X_n$  a.s. 收敛。

### 题目 1.3. Durrett(Exercise 4.2.6)

一系列非负独立同分布 r.v.  $Y_1, Y_2, \dots$  满足  $\mathbb{E}[Y_m] = 1, \mathbb{P}(Y_m = 1) < 1$ , 根据例 1.2.4 可知  $X_n = \prod_{m \leq n} Y_m$  是一个鞅, 证明:  $X_n \rightarrow 0$  a.s.

解答: 注意  $Y_n$  是独立同分布的, 因为  $\mathbb{P}(Y_m = 1) < 1$ , 考虑取  $u > 0$  使得  $\mathbb{P}(|Y_1 - 1| > u) > 0$ , 则对于  $\forall n$  都有  $\mathbb{P}(|Y_n - 1| > u) > 0$ , 于是  $\forall \varepsilon > 0$ ,

$$\begin{aligned} \mathbb{P}(|X_{n+1} - X_n| > \varepsilon u) &= \mathbb{P}(X_n |Y_{n+1} - 1| > \varepsilon u) \\ &= \mathbb{P}(X_n |Y_{n+1} - 1| > \varepsilon u | X_n \geq \varepsilon) \mathbb{P}(X_n \geq \varepsilon) \\ &\geq \mathbb{P}(|Y_{n+1} - 1| > u | X_n \geq \varepsilon) \mathbb{P}(X_n \geq \varepsilon) \\ &= \mathbb{P}(|Y_{n+1} - 1| > u) \mathbb{P}(X_n \geq \varepsilon) \end{aligned}$$

最后一个等号是因为  $X_n$  和  $Y_{n+1}$  独立。于是两边令  $n \rightarrow \infty$ , 可知  $X_n$  a.s. 收敛, 所以左边  $\rightarrow 0$ , 右边的  $\mathbb{P}(|Y_{n+1} - 1| > u) > 0$ , 于是只能  $\mathbb{P}(X_n \geq \varepsilon) \rightarrow 0$ , 所以  $X_n \rightarrow 0$  a.s.

### 题目 1.4. Durrett(Exercise 4.2.8)

正、可积的  $X_n, Y_n$  是关于  $\mathcal{F}_n$  的适应过程, 若

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] \leq (1 + Y_n) X_n$$

且  $\sum Y_n < +\infty$  a.s. 证明  $X_n$  a.s. 收敛。

解答: 观察题目给的条件, 我们尝试把它变成  $\mathbb{E}[Z_{n+1} | \mathcal{F}_n] \leq Z_n$  的形式。

$$\mathbb{E} \left[ \frac{X_{n+1}}{(1+Y_0)(1+Y_1)\cdots(1+Y_n)} \middle| \mathcal{F}_n \right] \leq \frac{X_n}{(1+Y_0)(1+Y_1)\cdots(1+Y_{n-1})}$$

于是可令

$$Z_n = \frac{X_n}{\prod_{i=0}^{n-1} (1+Y_i)}, \quad n \geq 1$$

则  $Z_n$  是一个下鞅, 同时  $Z_n > 0$ , 因此  $Z_n$  a.s. 收敛。考虑

$$\ln \prod_{i=0}^{n-1} (1+Y_i) = \sum_{i=0}^{n-1} \ln(1+Y_i) \leq \sum_{i=0}^{n-1} Y_i < +\infty$$

因此  $Z_n$  的分母  $\prod_{i=0}^{n-1} (1+Y_i)$  收敛, 从而分子  $X_n$  收敛。

### 题目 1.5. Durrett(Exercise 4.2.9)

$X_n^1, X_n^2$  是关于  $\mathcal{F}_n$  的上鞅,  $N$  是停时, 且满足  $X_N^1 \geq X_N^2$ , 令

$$Y_n = X_n^1 I_{\{N > n\}} + X_n^2 I_{\{N \leq n\}}$$

$$Z_n = X_n^1 I_{\{N \geq n\}} + X_n^2 I_{\{N < n\}}$$

则  $Y_n, Z_n$  都是上鞅。

解答:  $Y_n$  的各部分是  $\mathcal{F}_n$ -可测的, 所以  $Y_n$  是可测的。考虑拆分  $Y_{n+1}$ , 然后将其中的示性函数转化为  $\mathcal{F}_n$ -可测的形式,

$$\begin{aligned} Y_{n+1} &= X_{n+1}^1 I_{\{N > n+1\}} + X_{n+1}^2 I_{\{N \leq n+1\}} \\ &= X_{n+1}^1 I_{\{N > n\}} - X_{n+1}^1 I_{\{N = n+1\}} + X_{n+1}^2 I_{\{N = n+1\}} + X_{n+1}^2 I_{\{N \leq n\}} \\ &= X_{n+1}^1 I_{\{N > n\}} + X_{n+1}^2 I_{\{N \leq n\}} + (X_{n+1}^2 - X_{n+1}^1) I_{\{N = n+1\}} \end{aligned}$$

于是取条件期望:

$$\begin{aligned} \mathbb{E}[Y_{n+1} | \mathcal{F}_n] &= \mathbb{E}[X_{n+1}^1 I_{\{N > n\}} | \mathcal{F}_n] + \mathbb{E}[X_{n+1}^2 I_{\{N \leq n\}} | \mathcal{F}_n] + \mathbb{E}[(X_{n+1}^2 - X_{n+1}^1) I_{\{N = n+1\}} | \mathcal{F}_n] \\ &\leq \mathbb{E}[X_{n+1}^1 | \mathcal{F}_n] I_{\{N > n\}} + \mathbb{E}[X_{n+1}^2 | \mathcal{F}_n] I_{\{N \leq n\}} \\ &\leq X_n^1 I_{\{N > n\}} + X_n^2 I_{\{N \leq n\}} \\ &= Y_n \end{aligned}$$

关于  $Z_n$  的证明略有不同,

$$\begin{aligned} Z_{n+1} &= X_{n+1}^1 I_{\{N > n+1\}} + X_{n+1}^2 I_{\{N \leq n+1\}} \\ \mathbb{E}[Z_{n+1} | \mathcal{F}_n] &\leq X_{n+1}^1 I_{\{N > n+1\}} + X_{n+1}^2 I_{\{N \leq n+1\}} \\ &\leq X_n^1 I_{\{N > n+1\}} + X_n^2 I_{\{N \leq n+1\}} + (X_{n+1}^1 - X_n^1) I_{\{N = n+1\}} \\ &\leq X_n^1 I_{\{N \geq n+1\}} + X_n^2 I_{\{N < n+1\}} \\ &= Z_n \end{aligned}$$

### 题目 1.6. Durrett(Exercise 4.3.3)

正、可积的  $X_n, Y_n$  是关于  $\mathcal{F}_n$  的适应过程, 若

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] \leq X_n + Y_n$$

且  $\sum Y_n < \infty$  a.s. 证明  $X_n$  a.s. 收敛。提示: 构造停时

$$N = \inf \left\{ k : \sum_{m=1}^k Y_m > M \right\}$$

解答: 由题意可知

$$\mathbb{E} \left[ X_{n+1} - \sum_{k=0}^n Y_k \right] \leq X_n - \sum_{k=0}^{n-1} Y_k$$

于是令

$$Z_n = X_n - \sum_{k=0}^{n-1} Y_k, \quad n \geq 1$$

是一个上鞅, 取停时

$$N = \inf \left\{ k : \sum_{m=1}^k Y_m > M \right\}$$

于是  $\{Z_{n \wedge N}\}$  是一个新的上鞅, 考虑

$$Z_{n \wedge N} + M = X_{n \wedge N} - \sum_{k=0}^{n \wedge N - 1} Y_k + M$$

注意  $N$  是  $\sum Y_n$  首次超过  $M$ , 而  $n \wedge N - 1 < N$ , 所以上式中的求和不会超过  $M$ , 而  $X_n$  是正的, 所以  $Z_{n \wedge N} + M > 0$ , 这说明  $Z_{n \wedge N}$  a.s. 收敛。

在  $\{N = +\infty\}$  上,  $Z_{n \wedge N} = Z_n$  a.s. 收敛, 考虑到  $\sum Y_n < \infty$  a.s., 因此必然存在某个  $M$  使得  $\{N = +\infty\} = \Omega$  a.s., 于是  $Z_n$  在  $\Omega$  上 a.s. 收敛。

## 1.9.2 第二次作业

首先要补充一个引理 1.6.1 的升级版结论, 这是 Durrett 书上的 Exercise 4.4.2.

### 引理 1.9.1

$X_n$  是下鞅,  $M \leq N$  是停时, 且  $N \leq k$  a.s., 证明:  $\mathbb{E}[X_M] \leq \mathbb{E}[X_N]$ .

**证明** 取  $Y_n = X_{n \wedge N}$  为下鞅, 由引理 1.6.1 可得  $\mathbb{E}[Y_M] \leq \mathbb{E}[Y_k]$ , 即  $\mathbb{E}[X_M] \leq \mathbb{E}[X_N]$ . □

然后补充一个书上的定理 4.4.7, 是一个关于鞅的运算技巧,

### 定理 1.9.1 (Durrett(Theorem 4.4.7))

$X_n$  是鞅,  $\mathbb{E}X_n^2 < +\infty$ , 若  $m \leq n, Y \in \mathcal{F}_m, \mathbb{E}[Y^2] < +\infty$ , 则

$$\mathbb{E}[(X_n - X_m)Y] = 0$$

**证明** 由 Cauchy-Schwarz 不等式可知,

$$\mathbb{E}[(X_n - X_m)Y] \leq \mathbb{E}[(X_n - X_m)^2] \cdot \mathbb{E}[Y^2] < +\infty$$

这确保了  $(X_n - X_m)Y$  的可积性。然后就是很简单的变换:

$$\mathbb{E}[(X_n - X_m)Y] = \mathbb{E}[\mathbb{E}[(X_n - X_m)Y | \mathcal{F}_m]] = \mathbb{E}[\mathbb{E}[(X_n - X_m) | \mathcal{F}_m] Y] = 0$$

□

### 题目 1.7. Durrett(Exercise 4.4.3)

假设  $M \leq N$  是停时, 如果  $A \in \mathcal{F}_M$ , 则

$$L = MI_A + NI_{A^c}$$

也是停时。

**解答:** 验证定义:

$$\{L \leq n\} = (\{M \leq n\} \cap A) \cup (\{N \leq n\} \cap A^c)$$

注意  $A \in \mathcal{F}_M \Rightarrow \{M \leq n\} \cap A \in \mathcal{F}_n$ , 而右侧

$$\{N \leq n\} \cap A^c = \{N \leq n\} \cap \{M \leq n\} \cap A^c$$

其中  $\{N \leq n\} \in \mathcal{F}_n, A^c \in \mathcal{F}_M \Rightarrow \{M \leq n\} \cap A^c \in \mathcal{F}_n$ , 于是  $\{L \leq n\} \in \mathcal{F}_n$ . 由  $n$  的任意性可知  $L$  是停时。

### 题目 1.8. Durrett(Exercise 4.4.4)

利用上一题中的构造, 证明以下结论:  $X_n$  是下鞅,  $M \leq N$  是停时, 且  $\mathbb{P}(N \leq k) = 1$ , 则  $X_M \leq \mathbb{E}[X_N | \mathcal{F}_M]$ .

解答: 我们倒着分析这道题. 从结论出发, 设  $Y = \mathbb{E}[X_N | \mathcal{F}_M]$ ,  $X_M$  和  $Y$  都是  $\mathcal{F}_M$  可测的, 所以只需证明:  $\forall A \in \mathcal{F}_M$ ,

$$\mathbb{E}[X_M I_A] \leq \mathbb{E}[Y I_A] = \mathbb{E}[X_N I_A]$$

定义  $L = N I_A + M I_{A^c}$ , 则上式两边加上  $\mathbb{E}[X_M I_{A^c}]$  得到

$$\mathbb{E}[X_M] \leq \mathbb{E}[X_L]$$

而  $M \leq N \Rightarrow M \leq L$ , 由引理 1.9.1 即可得证。

### 题目 1.9. Durrett(Exercise 4.4.6)

设独立的随机变量列  $X_1, \dots, X_n, \dots$  满足均值为 0、方差  $\text{var}(X_n) = \sigma_n^2$ , 现在设  $S_n = X_1 + \dots + X_n$ ,  $s_n^2 = \sigma_1^2 + \dots + \sigma_n^2$ . 容易证明  $S_n^2 - s_n^2$  是一个鞅 (类似于例 1.2.5), 现在加上条件:  $|X_m| \leq K$ , 证明:

$$\mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| \leq x\right) \leq \frac{(x+K)^2}{s_n^2}$$

提示: 类似于定理 1.6.1 的证明过程。

解答: 设事件  $A = \{\max_{1 \leq m \leq n} |S_m| \leq x\}$ , 因为  $Y_n = S_n^2 - s_n^2$  是鞅, 取停时  $N = \inf\{m \geq 1 : |S_m| > x\}$ , 则  $Z_n = Y_{n \wedge N}$  也是鞅, 从而  $\mathbb{E}[Z_n] = \mathbb{E}[Z_1] = 0$ , 注意到  $A = \{n \leq N\}$ , 所以

$$\begin{aligned} 0 &= \mathbb{E}[Z_n] = \mathbb{E}[Y_n I_A] + \mathbb{E}[Y_N I_{A^c}] \\ &\leq \mathbb{E}[(S_n^2 - s_n^2) I_A] + \mathbb{E}[S_N^2 I_{A^c}] \\ &\leq (x^2 - s_n^2) \mathbb{P}(A) + \mathbb{E}[(S_{N-1} + X_N)^2 I_{A^c}] \\ &\leq (x^2 - s_n^2) \mathbb{P}(A) + \mathbb{E}[(x+K)^2 I_{A^c}] \\ &= (x^2 - s_n^2) \mathbb{P}(A) + (x+K)^2 (1 - \mathbb{P}(A)) \end{aligned}$$

于是

$$\mathbb{P}(A) \leq \frac{(x+K)^2}{(x+K)^2 - x^2 + s_n^2} \leq \frac{(x+K)^2}{s_n^2}$$

### 题目 1.10. Durrett(Exercise 4.4.7)

$X_n$  是鞅, 且  $X_0 = 0, \mathbb{E}[X_n]^2 < \infty$ , 证明:

$$\mathbb{P}\left(\max_{1 \leq m \leq n} X_m \geq \lambda\right) \leq \frac{\mathbb{E}[X_n^2]}{\mathbb{E}[X_n^2] + \lambda^2}$$

解答: 由题设知  $\mathbb{E}[X_n] = 0$ . 不妨设  $\lambda > 0$ , 取常数  $c \geq 0$ , 注意到  $(X_n + c)^2$  是一个非负下鞅, 由定理 1.6.1,

$$\mathbb{P}\left(\max_{1 \leq m \leq n} X_m \geq \lambda\right) = \mathbb{P}\left(\max_{1 \leq m \leq n} (X_m + c)^2 \geq (\lambda + c)^2\right) \leq \frac{1}{(\lambda + c)^2} \mathbb{E}[(X_n + c)^2] = \frac{\mathbb{E}[X_n^2] + c^2}{(\lambda + c)^2}$$

令  $c = \frac{\mathbb{E}[X_n^2]}{\lambda}$ , 就得到结论  $\frac{\mathbb{E}[X_n^2]}{\mathbb{E}[X_n^2] + \lambda^2}$ .

**题目 1.11. Durrett(Exercise 4.4.10)**

$\{X_n, n \geq 0\}$  是鞅, 对于  $n \geq 1$  设  $\xi_n = X_n - X_{n-1}$ , 如果  $\mathbb{E}[X_0^2] < \infty$  且

$$\sum_{m=1}^{\infty} \mathbb{E}[\xi_m^2] < +\infty$$

证明:  $X_n \rightarrow X$  a.s., 且  $X_n \xrightarrow{2} X$ .

解答: 由定理 1.6.3, 只需证明

$$\sup_n \mathbb{E}[X_n^2] < +\infty$$

注意到

$$\begin{aligned} X_n^2 &= \left( \sum_{m=1}^n \xi_m + X_0 \right)^2 \\ &= \sum_{m=1}^n \xi_m^2 + X_0^2 + 2 \sum_{m=1}^n X_0 \xi_m + 2 \sum_{m=1}^n \sum_{k=m+1}^n \xi_m \xi_k \end{aligned}$$

根据定理 1.9.1,

$$X_0 \in \mathcal{F}_{m-1} \Rightarrow \mathbb{E}[X_0 \xi_m] = \mathbb{E}[X_0 (X_m - X_{m-1})] = 0$$

$$X_m - X_{m-1} \in \mathcal{F}_{k-1} \Rightarrow \mathbb{E}[\xi_m \xi_k] = \mathbb{E}[(X_m - X_{m-1})(X_k - X_{k-1})] = 0$$

从而

$$\begin{aligned} \mathbb{E}[X_n^2] &= \sum_{m=1}^n \mathbb{E}[\xi_m^2] + \mathbb{E}[X_0^2] < +\infty \\ \sup_n \mathbb{E}[X_n^2] &\leq \sum_{m=1}^{\infty} \mathbb{E}[\xi_m^2] + \mathbb{E}[X_0^2] < +\infty \end{aligned}$$

**题目 1.12. Durrett(Exercise 4.6.6)**

$X_n \in [0, 1]$  是关于  $\mathcal{F}_n$  的适应过程, 设  $\alpha, \beta > 0$ ,  $\alpha + \beta = 1$ ,

$$\mathbb{P}(X_{n+1} = \alpha + \beta X_n | \mathcal{F}_n) = X_n$$

$$\mathbb{P}(X_{n+1} = \beta X_n | \mathcal{F}_n) = 1 - X_n$$

证明:

$$\mathbb{P}(\lim_{n \rightarrow \infty} X_n = 0 \text{ or } 1) = 1$$

且如果  $X_0 = \theta$ , 则  $\mathbb{P}(\lim_{n \rightarrow \infty} X_n = 1) = \theta$ .

解答: 容易验证  $X_n$  是个鞅, 而且一致有界, 所以一致可积, 进而  $X_n \xrightarrow{L^1 \text{ and a.s.}} X$ , 现在考虑集合:

$$B_n = \{\omega : X_{n+1}(\omega) = \alpha + \beta X_n(\omega)\}, B = \limsup_{n \rightarrow \infty} B_n = \{\omega : X_{n+1}(\omega) = \alpha + \beta X_n(\omega) \text{ 对于无数个 } n \text{ 成立}\}$$

由于  $X_n$  a.s. 收敛, 考虑  $\forall \varepsilon > 0$ , 有充分大的  $n$  使得  $|X_{n+1} - X_n| < \varepsilon$ , 于是在  $B$  上:

$$|X_{n+1} - X_n| = \alpha |1 - X_n| < \varepsilon \text{ 对于无数个 } n \text{ 成立}$$

这说明  $\{X_n(\omega)\}$  有收敛到 1 的子列, 那本身就收敛到 1, 即在  $B$  上  $X = 1$  a.s.

另一方面, 考虑

$$B^c = \liminf_{n \rightarrow \infty} B_n^c = \{\omega : X_{n+1}(\omega) = \beta X_n(\omega) \text{ 只对于有限个 } n \text{ 不成立}\}$$

于是在  $B^c$  上,

$$|X_{n+1} - X_n| = \alpha |X_n| < \varepsilon \text{ 只对于有限个 } n \text{ 不成立}$$

这说明  $X_n \rightarrow 0$ , 即在  $B^c$  上  $X = 0$  a.s.

因为  $\mathbb{E}[X] = \mathbb{E}[X_0] = \theta$ , 所以  $\mathbb{P}(X = 1) = \theta$ .

### 题目 1.13. Durrett(Exercise 4.6.7)

证明: 如果  $\mathcal{F}_n \nearrow \mathcal{F}_\infty$ ,  $Y_n \xrightarrow{L^1} Y$ , 则

$$\mathbb{E}[Y_n | \mathcal{F}_n] \xrightarrow{L^1} \mathbb{E}[Y | \mathcal{F}_\infty]$$

解答: 注意到:

$$\begin{aligned} \int |\mathbb{E}[Y_n | \mathcal{F}_n] - \mathbb{E}[Y | \mathcal{F}_\infty]| d\mathbb{P} &= \int |\mathbb{E}[Y_n | \mathcal{F}_n] - \mathbb{E}[Y | \mathcal{F}_n] + \mathbb{E}[Y | \mathcal{F}_n] - \mathbb{E}[Y | \mathcal{F}_\infty]| d\mathbb{P} \\ &\leq \int |\mathbb{E}[Y_n - Y | \mathcal{F}_n]| d\mathbb{P} + \int |\mathbb{E}[Y | \mathcal{F}_n] - \mathbb{E}[Y | \mathcal{F}_\infty]| d\mathbb{P} \\ &\leq \int \mathbb{E}[|Y_n - Y| | \mathcal{F}_n] d\mathbb{P} + \int |\mathbb{E}[Y | \mathcal{F}_n] - \mathbb{E}[Y | \mathcal{F}_\infty]| d\mathbb{P} \\ &= \mathbb{E}[|Y_n - Y|] + \int |\mathbb{E}[Y | \mathcal{F}_n] - \mathbb{E}[Y | \mathcal{F}_\infty]| d\mathbb{P} \end{aligned}$$

$Y_n \xrightarrow{L^1} Y$  所以前面这项  $\rightarrow 0$ , 后面这项由定理 1.7.4 可知也  $\rightarrow 0$ .

## 1.9.3 第三次作业

### 题目 1.14. Durrett(Exercise 4.8.1)

$L \leq M$  是停时,  $Y_{M \wedge n}$  是一致可积下鞅, 则  $\mathbb{E}[Y_L] \leq \mathbb{E}[Y_M]$ , 且  $Y_L \leq \mathbb{E}[Y_M | \mathcal{F}_L]$ .

解答: 先令  $Z_n = Y_{M \wedge n}$ , 那么  $Z_n \rightarrow Z_\infty$ , 则根据推论 1.7.6 可知  $\mathbb{E}[Z_L] \leq \mathbb{E}[Z_\infty]$ , 即  $\mathbb{E}[Y_L] \leq \mathbb{E}[Y_M]$ .

要证明  $Y_L \leq \mathbb{E}[Y_M | \mathcal{F}_L]$ , 因为  $Y_L$  是  $\mathcal{F}_L$ -可测的, 所以只需证明  $\forall A \in \mathcal{F}_L$  都有  $\mathbb{E}[Y_L I_A] \leq \mathbb{E}[\mathbb{E}[Y_M | \mathcal{F}_L] I_A] = \mathbb{E}[Y_M I_A]$ , 也就是要证明

$$\mathbb{E}[Z_L I_A] \leq \mathbb{E}[Z_\infty I_A]$$

两边加上  $\mathbb{E}[Z_L I_{A^c}]$ ,

$$\mathbb{E}[Z_L] \leq \mathbb{E}[Z_{L \wedge c}]$$

注意  $L \leq L^c$ , 而且  $Z_n$  是一致可积下鞅, 再使用推论 1.7.6 即可得证。

### 题目 1.15. Durrett(Exercise 4.8.3)

$X_1, \dots, X_n, \dots$  独立, 均值为 0, 方差相同, 为  $\text{var}(X_i) = \sigma^2$ ,  $S_n = X_1 + \dots + X_n$ , 则  $S_n^2 - n\sigma^2$  是鞅, 令  $T = \min\{n : |S_n| > a\}$ , 利用定理 1.7.7 证明<sup>a</sup>:  $\mathbb{E}[T] \geq a^2/\sigma^2$ .

<sup>a</sup>说真的, 题目让用定理 1.7.7 证明, 我死活没搞明白怎么证明  $S_T^2 - T\sigma^2$  可积, 最后发现答案根本没用那个定理, 甚至是用的下一题的结论? 实在是...

解答: 不妨假设  $\mathbb{E}[T] < +\infty$ , 然后用下一题的结论即可直接得证:

$$\mathbb{E}[T] = \frac{1}{\sigma^2} \mathbb{E}[S_T^2] \geq \frac{a^2}{\sigma^2}$$

### 题目 1.16. Durrett(Exercise 4.8.4)

条件同上一题, 设停时  $T$  满足  $\mathbb{E}[T] < +\infty$ , 则  $\mathbb{E}[S_T^2] = \sigma^2 \mathbb{E}[T]$ .

解答: 记  $Y_n = S_n^2 - n\sigma^2$ ,  $Z_n = Y_{n \wedge T}$ , 从而  $Z_n$  也是鞅并且  $\mathbb{E}[Z_n] = \mathbb{E}[Z_0] = 0$ , 因此我们有

$$\mathbb{E}[S_{n \wedge T}^2] = \sigma^2 \mathbb{E}[n \wedge T]$$

从而

$$\sup_n \mathbb{E}[S_{n \wedge T}^2] = \sigma^2 \cdot \sup_n \mathbb{E}[n \wedge T] \leq \sigma^2 \mathbb{E}[T] < +\infty$$

注意到  $S_n$  也是一个鞅, 所以由  $L^p$  收敛定理 1.6.3 可知  $S_{n \wedge T} \xrightarrow{L^2 \text{ and a.s.}} S_T$ , 从而

$$\mathbb{E}[S_T^2] = \lim_{n \rightarrow \infty} \mathbb{E}[S_{n \wedge T}^2] = \sigma^2 \lim_{n \rightarrow \infty} \mathbb{E}[n \wedge T] = \sigma^2 \mathbb{E}[T]$$

### 题目 1.17. Durrett(Exercise 4.8.5)

$X_1, \dots, X_n, \dots$  独立, 两点分布, 且  $p < \frac{1}{2}$ , 规定  $S_0 = 0$ , 设  $S_n = X_1 + \dots + X_n + x$ ,  $x$  是给定的整数。令  $V_0 = \min\{n \geq 0 : S_n = 0\}$ , 则  $\mathbb{E}[V_0] = x/(q-p)$ , 我们尝试计算  $V_0$  的方差, 我们令  $Y_i = X_i - (p-q)$ , 注意到  $\mathbb{E}[Y_i] = 0$ ,

$$\text{var}(Y_i) = \text{var}(X_i) = 1 - (p-q)^2$$

因此

$$Z_n = (S_n - (p-q)n)^2 - n(1 - (p-q)^2)$$

是一个鞅。证明:

$$\mathbb{E}[V_0^2] = x \cdot \frac{1 - (p-q)^2}{(p-q)^3}$$

解答: 根据上一题结论,

$$\mathbb{E}[V_0] = \frac{\mathbb{E}[(S_{V_0} - (p-q)V_0)^2]}{1 - (p-q)^2} = \frac{(p-q)^2 \mathbb{E}[V_0^2]}{1 - (p-q)^2} = \frac{x}{q-p}$$

于是

$$\mathbb{E}[V_0^2] = x \frac{1 - (p-q)^2}{(q-p)^3}$$

### 题目 1.18. Durrett(Exercise 4.8.8)

随机变量列  $X_1, \dots, X_n, \dots$  独立,  $\pm 1$  两点分布, 设  $S_n = X_1 + \dots + X_n$ , 设对于某些常数  $\theta_0 < 0$ ,

$$\mathbb{E}[\exp\{\theta_0 \cdot X_1\}] = 1$$

且  $X_1$  不是常数, 于是  $Y_n = \exp\{\theta_0 S_n\}$  是一个鞅, 令  $\tau = \inf\{n : S_n \notin (a, b)\}$ ,  $Z_n = Y_{n \wedge \tau}$ , 证明:  $\mathbb{E}[Y_\tau] = 1$ , 且

$$\mathbb{P}(S_\tau \leq a) \leq \exp\{-\theta_0 a\}$$

解答: 注意到  $S_{\tau \wedge n} \geq a$ , 所以  $Z_n \leq \exp\{\theta_0 a\}$ , 从而是一致可积鞅, 于是  $Z_n \rightarrow Z_\infty = Y_\tau$ ,  $\mathbb{E}[Y_\tau] = \mathbb{E}[Z_0] = 1$ ,

$$1 = \mathbb{E}[Y_\tau] \geq \mathbb{E}[Y_\tau I_{S_\tau \leq a}] \geq e^{a\theta_0} \mathbb{P}(S_\tau \leq a)$$

得证。

**题目 1.19. Durrett(Exercise 4.8.9)**

条件同上一题, 但  $X_i$  的取值范围改为整数值, 且  $\mathbb{P}(X_i < -1) = 0, \mathbb{P}(X_i = -1) > 0, \mathbb{E}[X_i] > 0$ , 令  $T = \inf\{n : S_n = a\}$ , 其中  $a < 0$ , 利用鞅  $Y_n$  来证明:  $\mathbb{P}(T < \infty) = \exp\{-\theta_0 a\}$ .

**解答:** 易证  $Y_{T \wedge n}$  是一致可积鞅, 从而  $\mathbb{E}[Y_T] = \mathbb{E}[Y_0] = 1$ ,

$$1 = \mathbb{E}[Y_T] = \exp\{a\theta_0\}\mathbb{P}(T < +\infty) + \mathbb{E}[Y_\infty; T = +\infty]$$

其中  $Y_\infty = 0$ , 因为根据大数定律,  $S_\infty = +\infty$ , 于是  $\mathbb{P}(T < +\infty) = \exp -a\theta_0$ .

**1.10 向后鞅 \***

其实课上没讲过向后鞅, 但是布置了作业题... 补充一点书上向后鞅的部分内容吧。

**定义 1.10.1**

向后鞅是指指标集为  $\{0, -1, -2, \dots\}$  的关于  $\{\mathcal{F}_n\}$  的适应过程  $\{X_n, n \leq 0\}$ , 且满足:

$$\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n, \forall n \leq -1$$

向后鞅的  $\sigma$ -域随着  $n \rightarrow -\infty$  是递减的。

**定理 1.10.1**

$X_n \rightarrow X_{-\infty}$  a.s. 且  $L^1$ .

**证明** 设

$$U_n = \sum_{-n \leq k \leq 0} I_{\{X_k \notin [a, b]\}}$$

为  $X_{-n}, \dots, X_0$  跳出区间  $[a, b]$  的次数。于是由定理 4.2.10 可得

$$(b - a)\mathbb{E}[U_n] \leq \mathbb{E}[(X_0 - a)^+]$$

令  $n \rightarrow \infty$ , 由 MCT, 可得  $\mathbb{E}[U_\infty] < +\infty$ , 所以  $X_n$  a.s. 收敛。

注意  $X_n = \mathbb{E}[X_0 | \mathcal{F}_n]$ , 直接可得  $X_n$  是一致可积的, 因此  $L^1$  收敛。 □

**定理 1.10.2**

$X_{-\infty} = \lim_{n \rightarrow -\infty} X_n, \mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_n$ , 则  $X_{-\infty} = \mathbb{E}[X_0 | \mathcal{F}_{-\infty}]$ .

**证明** 显然  $X_{-\infty} \in \mathcal{F}_{-\infty}, X_n = \mathbb{E}[X_0 | \mathcal{F}_n]$ , 于是如果  $A \in \mathcal{F}_{-\infty} \subset \mathcal{F}_n$ , 则

$$\int_A X_n d\mathbb{P} = \int_A X_0 d\mathbb{P}$$

上一个定理以及引理 4.6.5 表明  $\mathbb{E}[X_n; A] \rightarrow \mathbb{E}[X_{-\infty}; A]$ , 于是

$$\int_A X_{-\infty} d\mathbb{P} = \int_A X_0 d\mathbb{P}$$

对于  $\forall A \in \mathcal{F}_{-\infty}$  成立, 于是结论得证。 □

**定理 1.10.3**

如果  $\mathcal{F}_n \searrow \mathcal{F}_{-\infty}$ , 则

$$\mathbb{E}[Y | \mathcal{F}_n] \xrightarrow{\text{a.s. and } L^1} \mathbb{E}[Y | \mathcal{F}_{-\infty}]$$

**证明** 考虑  $X_n = \mathbb{E}[Y | \mathcal{F}_n]$  是一个向后鞅, 由前两个定理可得  $n \rightarrow -\infty$  时  $X_n \rightarrow X_{-\infty}$  a.s. 且  $L^1$ , 即

$$X_{-\infty} = \mathbb{E}[X_0 | \mathcal{F}_{-\infty}] = \mathbb{E}[\mathbb{E}[Y | \mathcal{F}_0] | \mathcal{F}_{-\infty}] = \mathbb{E}[Y | \mathcal{F}_{-\infty}]$$

□

**题目 1.20. Durrett(Exercise 4.7.1)**

证明：对于向后鞅  $X_n$ ，如果  $X_0 \in L^p$ ，则所有  $X_n \in L^p$ 。

解答：由  $L^p$  最大值不等式 (定理 1.6.2)，对于  $\forall n \leq 0$ ，

$$\mathbb{E} \left[ \sup_{n \leq m \leq 0} |X_m|^p \right] \leq \left( \frac{p}{p-1} \right)^p \mathbb{E}[|X_0|^p]$$

$n \rightarrow -\infty$  可得  $\sup_{m \leq 0} |X_m| \in L^p$ ，注意到  $|X_n - X_{-\infty}|^p \leq (2 \sup_{m \leq 0} |X_m|)^p$ ，后者  $L^p$  可积，所以由 DCT 可知  $\mathbb{E}[|X_n - X_{-\infty}|^p] \rightarrow 0$ 。

**题目 1.21. Durrett(Exercise 4.7.2)**

设  $Y_n \rightarrow Y_{-\infty}$  a.s.， $|Y_n| \leq Z \in L^1$ ，如果  $\mathcal{F}_n \searrow \mathcal{F}_{-\infty}$ ，则  $\mathbb{E}[Y_n | \mathcal{F}_n] \rightarrow \mathbb{E}[Y_{-\infty} | \mathcal{F}_{-\infty}]$  a.s.

解答：注意到

$$|\mathbb{E}[Y_n | \mathcal{F}_n] - \mathbb{E}[Y_{-\infty} | \mathcal{F}_{-\infty}]| \leq \mathbb{E}[|Y_n - Y_{-\infty}| | \mathcal{F}_n] + |\mathbb{E}[Y_{-\infty} | \mathcal{F}_n] - \mathbb{E}[Y_{-\infty} | \mathcal{F}_{-\infty}]|$$

显然后者趋于 0，对于前者，考虑  $W_N = \sup_{m, n \leq N} |Y_m - Y_n|$ ，则  $W_N \leq 2Z \in L^1$  且  $W_N \rightarrow 0$ ，于是

$$\limsup_{n \rightarrow -\infty} \mathbb{E}[|Y_n - Y_{-\infty}| | \mathcal{F}_n] \leq \limsup_{n \rightarrow -\infty} \mathbb{E}[W_N | \mathcal{F}_n] = \mathbb{E}[W_N | \mathcal{F}_{-\infty}] \rightarrow 0$$

## 第 2 章 离散时间马氏过程

主要内容为随机过程（研课）关于马氏链的部分，标星号的为应用随机过程中提到的新内容。研随课程在前期用测度论的语言给出了严谨的马氏链定义，详细介绍了马氏性与强马氏性以及应用；而应随则全程拿初等语言死磕，着重介绍了状态分类的判断准则和平均返回时间、平均经过次数、首次到达时间期望等各种奇奇怪怪的量的计算方法。应随有些讲得特别不清不白的内容我干脆就删了（强马氏性），以及遍历定理相关的那部分我没听（结果期中考了，淦），也没多大兴趣，不整理了。

### 2.1 马氏链的定义

#### 2.1.1 基本定义

##### 定义 2.1.1 (马氏链)

概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上的离散随机过程  $X = \{X_n, n \geq 0\}$  中的随机变量所有可能的取值组成的集合称为状态空间，记作  $S$ 。

设  $S$  是至多可列集，称  $\{X_n\}$  具有马氏性 (Markov property) 是指：

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

其中  $i, j, i_k \in S, k = 0, 1, \dots, n-1$  是任意的。此时  $X$  就被称为离散时间离散状态马氏链，简称离散马氏链。

马氏性的直观理解：给定  $n$  时刻及以前的信息，则未来 ( $n+1$  时刻) 与过去 ( $< n$  时刻) 无关，只与现在 ( $n$  时刻) 的状态有关。

##### 定义 2.1.2

对于马氏链  $X = \{X_n, n \geq 0\}$ ，如果它还满足

$$p_{ij} \stackrel{\text{def}}{=} \mathbb{P}(X_{n+1} = j | X_n = i)$$

与  $n$  无关，则称其为齐次的马氏链， $p_{ij}$  称为一步转移概率，

我们下文中默认探讨的都是齐次的离散马氏链。齐次有个好处是：我们只需要关心两个时刻之间的时间差，而不需要关心具体的时刻。后文中的 [题目 2.1](#) 给出了一个非齐次马氏链的例子。

## 例 2.1.1. Ehrenfest Chain

有两个瓶子记作 A、B，两个瓶子一共有  $r$  个球，初始时 A 瓶中有  $k$  个球，B 瓶中有  $r-k$  个球。每次操作从  $r$  个球中随机地选一个，并将其放到另外一个瓶子里。记  $X_n$  为  $n$  次操作之后 A 瓶中的球数，则  $\{X_n\}$  是一个马氏链，状态空间  $S = \{0, 1, \dots, r\}$ ，一步转移概率：

$$p_{kj} = \mathbb{P}(X_{n+1} = j | X_n = k) \begin{cases} 0 & , |k-j| \neq 1 \\ \frac{k}{r} & , j = k+1 \\ \frac{r-k}{r} & , j = k-1 \end{cases}$$

## 2.1.2 基本性质和应用 \*

## 命题 2.1.1 (马氏链的等价定义)

对于离散随机过程  $\{X_n\}$ ，以下命题等价：

(1). 具有马氏性： $\forall n \geq 0, \forall i, j, i_{n-1}, \dots, i_0 \in S$ ,

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

(2).  $\forall m, n \geq 0, \forall s, x_m, \dots, x_0 \in S$ ,

$$\mathbb{P}(X_{m+n} = s | X_m = x_m, X_{m-1} = x_{m-1}, \dots, X_0 = x_0) = \mathbb{P}(X_{m+n} = s | X_m = x_m)$$

(3).  $\forall 0 \leq n_1 < n_2 < \dots < n_k \leq n, \forall s, x_{n_1}, \dots, x_{n_k} \in S$ ,

$$\mathbb{P}(X_{n+1} = s | X_{n_k} = x_{n_k}, \dots, X_{n_1} = x_{n_1}) = \mathbb{P}(X_{n+1} = s | X_{n_k} = x_{n_k})$$

## 例 2.1.2.

$\{X_n\}$  是马氏链，请问以下等式成立吗？

(1).  $\mathbb{P}(X_3 = j | X_2 = i, X_1 \in A, X_0 \in B) = \mathbb{P}(X_3 = j | X_2 = i)$ .

(2).  $\mathbb{P}(X_3 = j | X_2 \in C, X_1 \in A, X_0 \in B) = \mathbb{P}(X_3 = j | X_2 \in C)$ .

解答：

(1). 成立，拆开计算验证即可：

$$\begin{aligned} LHS &= \frac{\mathbb{P}(X_3 = j, X_2 = i, X_1 \in A, X_0 \in B)}{\mathbb{P}(X_2 = i, X_1 \in A, X_0 \in B)} \\ &= \sum_{a \in A, b \in B} \frac{\mathbb{P}(X_3 = j, X_2 = i, X_1 = a, X_0 = b)}{\mathbb{P}(X_2 = i, X_1 \in A, X_0 \in B)} \\ &= \sum_{a \in A, b \in B} \frac{\mathbb{P}(X_3 = j | X_2 = i, X_1 = a, X_0 = b) \mathbb{P}(X_2 = i, X_1 = a, X_0 = b)}{\mathbb{P}(X_2 = i, X_1 \in A, X_0 \in B)} \\ &= \sum_{a \in A, b \in B} \frac{\mathbb{P}(X_3 = j | X_2 = i) \mathbb{P}(X_2 = i, X_1 = a, X_0 = b)}{\mathbb{P}(X_2 = i, X_1 \in A, X_0 \in B)} \\ &= \frac{\mathbb{P}(X_3 = j | X_2 = i) \mathbb{P}(X_2 = i, X_1 \in A, X_0 \in B)}{\mathbb{P}(X_2 = i, X_1 \in A, X_0 \in B)} \\ &= \mathbb{P}(X_3 = j | X_2 = i) = RHS \end{aligned}$$

(2). 不一定成立，考虑取  $C = S$ ，则等式化为  $\mathbb{P}(X_3 = j | X_1 \in A, X_0 \in B) = \mathbb{P}(X_3 = j)$ ，这要求  $X_1, X_0$  和  $X_3$  独立。

如果马氏链的状态空间  $S$  是有限集，那么一步转移概率  $p_{ij}$  可以组成转移概率矩阵  $\mathbf{P} = (p_{ij})_{|S| \times |S|}$ .

## 例 2.1.3.

某产品有三种品牌，分别记作 1、2、3，记  $X_n$  为某顾客第  $n$  购买的产品品牌，并设  $\{X_n\}$  是一个马氏链，其转移矩阵为

$$\mathbf{P} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

问：

- (1). 顾客目前购买了品牌 2，请问下一次购买 3、再下一次购买 1 的概率？
- (2). 顾客目前购买了品牌 2，请问下次购买 1 的概率？

解答：

(1). 即  $\mathbb{P}(X_2 = 1, X_1 = 3 | X_0 = 2) = \mathbb{P}(X_2 = 1 | X_1 = 3, X_0 = 2) \mathbb{P}(X_1 = 3 | X_0 = 2) = p_{31} p_{23} = 0.06$ .

(2). 即  $\mathbb{P}(X_2 = 1 | X_0 = 2) = \sum_i \mathbb{P}(X_2 = 1, X_1 = i | X_0 = 2) = \sum_i p_{i1} p_{2i} = 0.16 + 0.12 + 0.06 = 0.34$ .

这个例子表明多步转移概率是可以通过一步转移概率来计算的，那么它们之间具体的关系是什么？

## 定理 2.1.1 (Chapman-Kolmogorov Equation)

设状态空间  $S$  是离散的（即至多可数），则

$$\mathbb{P}_x(X_{m+n} = z) = \sum_{y \in S} \mathbb{P}_x(X_m = y) \mathbb{P}_y(X_n = z), \forall x, z \in S$$

这里  $\mathbb{P}_x(A)$  是指  $\mathbb{P}_{X_0}(A) |_{X_0=x}$ ，即给定初始状态  $X_0 = x$ 。

证明

$$\begin{aligned} \mathbb{P}_x(X_{m+n} = z) &= \sum_{y \in S} \mathbb{P}_x(X_{m+n} = z, X_m = y) \\ &= \sum_{y \in S} \mathbb{P}(X_{m+n} = z | X_m = y) \mathbb{P}_x(X_m = y) \\ &= \sum_{y \in S} \mathbb{P}_y(X_n = z) \mathbb{P}_x(X_m = y) \end{aligned}$$

□

如果状态空间有限，我们可以写出 Chapman-Kolmogorov 方程的矩阵形式：

$$\mathbf{P}(m, m+n+r) = \mathbf{P}(m, m+n) \mathbf{P}(m+n, m+n+r)$$

所以  $n$  步转移概率矩阵  $\mathbf{P}(m, m+n) = \mathbf{P}^n$ ，这是 Chapman-Kolmogorov 方程的一个简单推论。

记  $\mu_i^{(n)} = \mathbb{P}(X_n = i)$ ，为  $X_n$  的质量函数，行向量  $\mu^{(n)} = (\mu_i^{(n)}, i \in S)$  记录了  $X_n$  的分布，那么

## 引理 2.1.1

$$\mu^{(m+n)} = \mu^{(m)} \mathbf{P}^n.$$

♥

证明

$$\begin{aligned} \mu_j^{(m+n)} &= \mathbb{P}(X_{m+n} = j) \\ &= \sum_{i \in S} \mathbb{P}(X_{m+n} = j | X_m = i) \mathbb{P}(X_m = i) \\ &= \sum_{i \in S} \mu_i^{(m)} p_{ij}(m, m+n) \\ &= (\mu^{(m)} \mathbf{P}^n)_j \end{aligned}$$

□

因此, 转移概率矩阵  $\mathbf{P}$  和初始时刻  $X_0$  的分布  $\mu^{(0)}$  决定了马氏链的分布。

### 例 2.1.4.

满足 C-K 方程的随机过程一定是马氏链吗?

解答: 答案是否定的, 下面给出反例: 考虑随机过程  $\{Y_n, n \geq 1\}$ , 其中奇数项  $Y_1, Y_3, Y_5, \dots$  独立同分布,

$$\mathbb{P}(Y_{2k+1} = \pm 1) = \frac{1}{2}, \forall k \in \mathbb{N}$$

而偶数项  $Y_{2k} = Y_{2k-1}Y_{2k+1}$ , 可以证明:

1° 偶数项独立同分布, 分布和奇数项一样。

2°  $\{Y_n\}$  两两独立。

于是

$$p_{ij}(m, m+n) = \mathbb{P}(Y_{m+n} = j | Y_m = i) = \mathbb{P}(Y_{m+n} = j) = \frac{1}{2}$$

回头看 C-K 方程:

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

两边都是  $\frac{1}{2}$ . 但是  $\{Y_n\}$  并不是一个马氏链:

$$\mathbb{P}(Y_{2k+1} | Y_{2k} = -1, Y_{2k-1} = 1) = 0 \neq \mathbb{P}(Y_{2k+1} | Y_{2k} = -1) = \frac{1}{2}$$

### 题目 2.1. 非齐次马氏链的例子

接例 2.1.4, 如果令  $Z_n = (Y_n, Y_{n+1})$ , 证明  $\{Z_n, n \geq 1\}$  是一个 (非齐次) 马氏链, 并写出其一步转移概率。

解答: 状态空间为  $S = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$ , 我们先说明  $\{Z_n\}$  是马氏链: 记状态  $i = (i^{(1)}, i^{(2)})$ , 要证明其为马氏链, 也就是证明

$$\mathbb{P}(Z_n = i | Z_{n-1} = i_{n-1}, \dots, Z_1 = i_1) = \mathbb{P}(Z_n = i | Z_{n-1} = i_{n-1}) \quad (*)$$

考虑把 (\*) 左式展开:

$$\begin{aligned} & \mathbb{P}(Z_n = i | Z_{n-1} = i_{n-1}, \dots, Z_1 = i_1) \\ &= \frac{\mathbb{P}(Y_{n+1} = i^{(2)}, Y_n = i^{(1)} = i_{n-1}^{(2)}, Y_{n-1} = i_{n-1}^{(1)} = i_{n-2}^{(2)}, \dots, Y_2 = i_2^{(1)} = i_1^{(2)}, Y_1 = i_1^{(1)})}{\mathbb{P}(Y_n = i_{n-1}^{(2)}, Y_{n-1} = i_{n-1}^{(1)} = i_{n-2}^{(2)}, \dots, Y_2 = i_2^{(1)} = i_1^{(2)}, Y_1 = i_1^{(1)})} \end{aligned}$$

条件概率的定义规定了条件事件概率不能为零, 所以上式中所有状态相等的那些等式, 都必须是成立的, 唯一一个可能不成立的是  $i^{(1)} = i_{n-1}^{(2)}$ , 不成立时 (\*) 式成立 (两边都是零), 所以接下来也不妨假设是成立的。将原本得到分式简化一下:

$$\begin{aligned} & \frac{\mathbb{P}(Y_{n+1} = i^{(2)}, Y_n = i_{n-1}^{(2)}, Y_{n-1} = i_{n-1}^{(1)}, \dots, Y_2 = i_1^{(2)}, Y_1 = i_1^{(1)})}{\mathbb{P}(Y_n = i_{n-1}^{(2)}, Y_{n-1} = i_{n-1}^{(1)}, \dots, Y_2 = i_2^{(1)}, Y_1 = i_1^{(1)})} \\ &= \mathbb{P}(Y_{n+1} = i^{(2)} | Y_n = i_{n-1}^{(2)}, Y_{n-1} = i_{n-1}^{(1)}, \dots, Y_2 = i_1^{(2)}, Y_1 = i_1^{(1)}) \\ &= \begin{cases} \frac{1}{2}, & n \text{ 是奇数} \\ 1, & n \text{ 是偶数, 且 } i^{(2)} \cdot i_{n-1}^{(1)} = i_{n-1}^{(2)} \\ 0, & n \text{ 是偶数, 且 } i^{(2)} \cdot i_{n-1}^{(1)} \neq i_{n-1}^{(2)} \end{cases} \end{aligned}$$

到这里已经能看出来, (\*) 左式只和  $n-1$  之后发生的事情有关。如果想严谨一些, 就展开 (\*) 右式再讨论即可, 这里就不过多叙述了。

一步转移概率如下:  $n$  为偶数时,

	(1, 1)	(1, -1)	(-1, 1)	(-1, -1)
(1, 1)	$\frac{1}{2}$	$\frac{1}{2}$	0	0
(1, -1)	0	0	$\frac{1}{2}$	$\frac{1}{2}$
(-1, 1)	$\frac{1}{2}$	$\frac{1}{2}$	0	0
(-1, -1)	0	0	$\frac{1}{2}$	$\frac{1}{2}$

$n$  为奇数时,

	(1, 1)	(1, -1)	(-1, 1)	(-1, -1)
(1, 1)	1	0	0	0
(1, -1)	0	0	0	1
(-1, 1)	0	1	0	0
(-1, -1)	0	0	1	0

### 2.1.3 转移概率函数

#### 定义 2.1.3

设  $(S, \beta(S))$  是一个可测空间, 如果函数  $p: S \times \beta(S) \rightarrow \mathbb{R}$  满足:

1° 对于任意固定的  $x \in S$ ,  $p_x \stackrel{\text{def}}{=} p(x, \cdot): \beta(S) \rightarrow \mathbb{R}$  是一个概率测度;

2° 对于任意固定的  $A \in \beta(S)$ ,  $p^A \stackrel{\text{def}}{=} p(\cdot, A): S \rightarrow \mathbb{R}$  是  $(S, \beta(S))$  上的可测函数;

则称  $p$  是一个转移概率函数。

#### 定义 2.1.4

概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上的随机过程  $\{X_n\}$  是关于滤流  $\{\mathcal{F}_n\}$  的马氏链, 状态空间为  $S$ , 再给定可测空间  $(S, \beta(S))$ , 如果  $(S, \beta(S))$  上的转移概率函数  $p$  满足:

$$\mathbb{P}(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B), \quad \forall n, \forall B \in \beta(S) \quad (2.1.1)$$

则称  $\{X_n\}$  是 (关于滤流  $\{\mathcal{F}_n\}$  的) 以  $p$  为转移概率的马氏链。

我们可以通过概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$ 、状态空间  $S$  和转移概率矩阵  $\mathbf{P}$  来描述一条马氏链, 但如果涉及到  $S$  非离散的情况,  $\mathbb{P}(X_n = i) = 0$ , 一步转移概率  $\mathbb{P}(X_{n+1} = j | X_n = i)$  作为条件概率就无法定义。因此我们采用转移概率函数而非转移概率矩阵来描述马氏链。

下面是笔者的一些想法, 尝试寻找转移概率函数和一步转移概率的关系:

$$\{X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0\} \in \mathcal{F}_n$$

根据条件期望的定义, 可得

$$\begin{aligned} & \int_{\{X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}} \mathbb{P}(X_{n+1} \in B | \mathcal{F}_n) d\mathbb{P} \\ &= \int_{\{X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}} \mathbb{E}[I_{\{X_{n+1} \in B\}} | \mathcal{F}_n] d\mathbb{P} \\ &= \int_{\{X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}} I_{\{X_{n+1} \in B\}} d\mathbb{P} \\ &= \mathbb{P}(X_{n+1} \in B, X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \\ &= \mathbb{P}(X_{n+1} \in B | X_n = i) \mathbb{P}(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \end{aligned}$$

然后考虑另一侧,

$$\begin{aligned} \int_{\{X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}} p(X_n, B) d\mathbb{P} &= \int p(X_n, B) I_{\{X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}} d\mathbb{P} \\ &= \int p(i, B) I_{\{X_n=i, X_{n-1}=i_{n-1}, \dots, X_0=i_0\}} d\mathbb{P} \\ &= p(i, B) \mathbb{P}(X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) \end{aligned}$$

这就得到  $p(i, B) = \mathbb{P}(X_{n+1} \in B | X_n = i)$ , 进一步 (如果单点集  $\{j\}$  是可测的话) 得到

$$p(i, j) = \mathbb{P}(X_{n+1} = j | X_n = i)$$

可以发现这正是我们之前定义的一步转移概率。所以我认为, 转移概率函数是一种比一步转移概率更“通用”的表示方法, 正如我们用  $\mathbb{E}[X|\sigma(Y)]$  的语言代替了  $\mathbb{E}[X|Y]$ , 就是为了解决连续型分布落在单个点上的概率为零导致无法定义条件概率的问题。

但是因为我们正在探讨的马氏链很简单: 离散时间、离散状态、齐次, 所以大部分情况下我们可以采用初等的语言来描述 (尽管形式上非常繁琐), 转移概率函数在叙述方式上的“优越性”还没有被充分展现。

## 2.2 马氏链的构造与性质

### 2.2.1 从转移概率构造马氏链

现在研究的问题是: 如果先给出  $(S, \beta(S))$  和转移概率函数  $p$ , 如何给出概率空间  $(\Omega, \mathcal{F})$  及其上的以  $p$  为转移概率的马氏链  $\{X_n\}$ ?

#### 定义 2.2.1

我们现在只给定  $(S, \beta(S))$  和转移概率函数  $p$ , 令  $\Omega_0 = S^\infty$ , 即  $S$  上的序列全体, 对于  $\omega = (\omega_0, \dots, \omega_n, \dots) \in \Omega$ , 定义函数:

$$X_n : \Omega_0 \rightarrow S, \omega \mapsto \omega_n$$

令  $\mathcal{F} = \beta(S)^\infty$ , 那么  $X_n$  就是  $(\Omega_0, \mathcal{F})$  到  $(S, \beta(S))$  的可测函数。

如果在  $(S, \beta(S))$  上给定一个概率测度  $\mu$ , 根据 Kolmogorov 扩张定理,  $(\Omega_0, \mathcal{F})$  上存在唯一概率测度  $\mathbb{P}_\mu$  满足:  $\forall n, \forall B_k \in \beta(S), k = 0, \dots, n$ ,

$$\mathbb{P}_\mu(X_0 \in B_0, \dots, X_n \in B_n) = \int_{B_0} \mu(dX_0) \int_{B_1} p(X_0, dX_1) \int_{B_2} p(X_1, dX_2) \cdots \int_{B_n} p(X_{n-1}, dX_n) \quad (2.2.1)$$

左边是  $\{\omega = (\omega_0, \dots, \omega_n, \dots) : \omega_0 \in B_0, \dots, \omega_n \in B_n\}$  的概率测度, 右边第一项积分是指  $\int_{B_0} X_0 d\mu$ , 即  $X_0$  在  $B_0$  上对测度  $\mu$  积分, 后面的几项有点复杂, 我们回顾定义 2.1.3: 固定  $X_0$  时,  $p(X_0, \cdot)$  是一个概率测度, 所以  $\int_{B_1} p(X_0, dX_1)$  其实就是  $X_1$  在  $B_1$  上对这个概率测度积分, 也就是:

$$\int_{B_1} X_1 dp(X_0, \cdot)$$

这样积分出来是一个与  $X_0$  的值有关的量, 其余项同理。

整理一下思绪: 我们给定了测度空间  $(S, \beta(S), \mu)$  和转移概率函数  $p$ , 定义出了概率空间  $(\Omega_0, \mathcal{F}, \mathbb{P}_\mu)$  和其上的一个随机过程  $\{X_n\}$ ,

$$(\Omega_0, \mathcal{F}, \mathbb{P}_\mu) \xrightarrow{X_n} (S, \beta(S), \mu)$$

$\{X_n\}$  的状态空间是  $S$ , 那么  $\{X_n\}$  是这个概率空间上的以  $p$  为转移概率马氏链  $\{X_n\}$  吗? 答案是肯定的。

## 定理 2.2.1

令  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ , 则定义 2.2.1 中给出的  $\{X_n\}$ , 就是关于滤流  $\{\mathcal{F}_n\}$  的以  $p$  为转移概率的马氏链, 即满足:

$$\mathbb{P}_\mu(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B), \forall n, \forall B \in \beta(S) \quad (2.2.2)$$

**证明** 式 (2.2.2) 也就是:

$$\mathbb{E}[I_{\{X_{n+1} \in B\}} | \mathcal{F}_n] = p(X_n, B), \forall n, \forall B \in \beta(S)$$

固定  $B$  之后,  $p(X_n, B)$  是关于  $X_n$  的可测函数, 当然是  $\mathcal{F}_n$ -可测的, 于是我们只需要验证条件期望的第二条定义:

$$\int_A \mathbb{E}[I_{X_{n+1} \in B} | \mathcal{F}_n] d\mathbb{P}_\mu = \int_A p(X_n, B) d\mathbb{P}_\mu, \forall A \in \mathcal{F}_n$$

实际上左式可以变换为:

$$\int_A \mathbb{E}[I_{X_{n+1} \in B} | \mathcal{F}_n] d\mathbb{P}_\mu = \mathbb{E}[\mathbb{E}[I_{X_{n+1} \in B} | \mathcal{F}_n] I_A] = \mathbb{E}[\mathbb{E}[I_{X_{n+1} \in B} I_A | \mathcal{F}_n]] = \mathbb{E}[I_{X_{n+1} \in B} I_A]$$

所以即证:

$$\mathbb{E}[I_{X_{n+1} \in B} I_A] = \int_A p(X_n, B) d\mathbb{P}_\mu, \forall A \in \mathcal{F}_n \quad (2.2.3)$$

不失一般性, 设  $A = \{X_0 \in B_0, \dots, X_n \in B_n\}$ , 则

$$\begin{aligned} (2.2.3) LHS &= \mathbb{P}_\mu(X_0 \in B_0, \dots, X_n \in B_n, X_{n+1} \in B) \\ &= \int_{B_0} \mu(dX_0) \int_{B_1} p(X_0, dX_1) \cdots \int_{B_n} p(X_{n-1}, dX_n) \int_B p(X_n, dX_{n+1}) \\ &= \int_{B_0} \mu(dX_0) \int_{B_1} p(X_0, dX_1) \cdots \int_{B_n} p(X_n, B) \cdot p(X_{n-1}, dX_n) \end{aligned}$$

实际上我们可以证明: 对于任意的  $\beta(S)$ -可测函数  $f$ , 都有:

$$\int_{B_0} \mu(dX_0) \int_{B_1} p(X_0, dX_1) \cdots \int_{B_n} f(X_n) \cdot p(X_{n-1}, dX_n) = \int_A f(X_n) d\mathbb{P}_\mu \quad (2.2.4)$$

然后我们令  $f(X_n) = p(X_n, B)$ , 式 (2.2.3) 就得证。

至于如何证明式 (2.2.4), 就是经典的四步走: 示性函数  $\xrightarrow{\text{线性组合}}$  简单函数  $\xrightarrow{\text{逼近}}$  非负函数  $\xrightarrow{\text{正负部分离}}$  一般可测函数。我们只验证一下示性函数的情况: 设  $f = \chi_C, C \in \beta(S)$ ,

$$\begin{aligned} (2.2.4) LHS &= \int_{B_0} \mu(dX_0) \int_{B_1} p(X_0, dX_1) \cdots \int_{B_n} \chi_C \cdot p(X_{n-1}, dX_n) \\ &= \int_{B_0} \mu(dX_0) \int_{B_1} p(X_0, dX_1) \cdots \int_{B_{n-1}} p(X_{n-1}, B_n \cap C) p(X_{n-2}, dX_{n-1}) \\ (2.2.4) RHS &= \int_A \chi_C(X_n) d\mathbb{P}_\mu \\ &= \mathbb{P}_\mu(X_0 \in B_0, \dots, X_{n-1} \in B_{n-1}, X_n \in B_n \cap C) \end{aligned}$$

根据式 (2.2.1) 展开就得证。 □

下面是一些应用和推论。先回顾单调类定理:

## 定理 2.2.2

$\mathcal{A}$  是一个包含全集  $\Omega$  的  $\pi$ -系,  $\mathcal{H}$  是某些实值函数的集合, 若

1°  $\forall A \in \mathcal{A} \Rightarrow I_A \in \mathcal{H}$ .

2°  $f, g \in \mathcal{H} \Rightarrow af + bg \in \mathcal{H}$ .

3°  $0 \leq f_n \in \mathcal{H}$  且  $f_n \nearrow f$ , 则  $f \in \mathcal{H}$ .

则  $\mathcal{H}$  包含所有  $\sigma(\mathcal{A})$ -可测的有界实值函数。

根据单调类定理我们可以得到:

$$\mathbb{E}[f(X_{n+1}|\mathcal{F}_n)] = \int_S f(y)p(X_n, dy) \quad (2.2.5)$$

进而得到以下推论:

### 推论 2.2.1

概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上,  $\{X_n\}$  是以  $p$  为转移概率的马氏链, 给定若干有界可测函数  $f_0, \dots, f_m$ , 有:

$$\mathbb{E}[f_0(X_0)f_1(X_1)\cdots f_m(X_m)] = \int_S f_0(X_0)\mu(dX_0) \int_S f_1(X_1)p(X_0, dX_1) \cdots \int_S f_m(X_m)p(X_{m-1}, dX_m) \quad (2.2.6)$$

**证明** 归纳法:  $m=1$  时即为式 (2.2.5), 假设命题对于  $m$  成立, 考虑  $m+1$ :

$$\begin{aligned} \mathbb{E}[f_0(X_0)f_1(X_1)\cdots f_m(X_m)f_{m+1}(X_{m+1})] &= \mathbb{E}[\mathbb{E}[f_0(X_0)f_1(X_1)\cdots f_{m+1}(X_{m+1})|\mathcal{F}_m]] \\ &= \mathbb{E}[f_0(X_0)\cdots f_m(X_m)\cdot\mathbb{E}[f_{m+1}(X_{m+1})|\mathcal{F}_m]] \\ &= \mathbb{E}[f_0(X_0)\cdots f_m(X_m)\underbrace{\int_S f_{m+1}(y)p(X_m, dy)}_{\stackrel{\text{def}}{=} \hat{f}(X_m)}] \\ &= \int_S f_0(X_0)\mu(dX_0) \int_S f_1(X_1)\mu(dX_1) \cdots \int_S p(X_{m-1}, dX_m)\hat{f}(X_m) \end{aligned}$$

令  $y = X_{m+1}$  即得. □

## 2.2.2 马氏性

引入推移算子,

### 定义 2.2.2

给定  $(S, \beta(S))$ , 定义  $\Omega_0 = S^\infty$  上的推移算子:

$$\theta_n : (\omega_0, \dots, \omega_n, \dots) \mapsto (\omega_n, \omega_{n+1}, \dots)$$

对于转移概率函数  $p$ ,  $p(X, \cdot)$  是一个概率测度, 下文中, 我们记

$$\mathbb{P}_X(A) \stackrel{\text{def}}{=} p(X, A), \mathbb{E}_X[Y] \stackrel{\text{def}}{=} \int Y(\omega)p(X, d\omega)$$

### 定理 2.2.3

给定马氏链  $\{X_n\}$  和转移概率  $p$ , 设  $Y$  是  $\sigma(X_0, \dots, X_n, \dots)$ -有界可测的,  $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ . 则马氏性可叙述如下:

$$\mathbb{E}_\mu[Y \circ \theta_m | \mathcal{F}_m] = \mathbb{E}_{X_m}[Y]$$

**证明** 不难看出右式是  $X_m$  的函数, 故是  $\mathcal{F}_m$ -可测的, 现只需验证:  $\forall A \in \mathcal{F}_m$ ,

$$\mathbb{E}_\mu[Y \circ \theta_m \cdot I_A] = \mathbb{E}_\mu[\mathbb{E}_{X_m}[Y] \cdot I_A] \quad (2.2.7)$$

不失一般性, 可以设  $A$  具有以下形式<sup>1</sup>:  $A = \{X_0 \in A_0, X_1 \in A_1, \dots, X_m \in A_m\}$ ,  $Y$  具有以下形式<sup>2</sup>:  $Y = g_0(X_0)g_1(X_1)\cdots g_n(X_n), \forall n \geq 0$ , 其中  $g_i$  都是有界可测的.

$$\begin{aligned} (2.2.7) \text{LHS} &= \mathbb{E}_\mu[g_0(X_m)\cdots g_n(X_{m+n})I_A] \\ &= \mathbb{E}_\mu[g_0(X_m)\cdots g_n(X_{m+n})I_{A_0}(X_0)\cdots I_{A_m}(X_m)] \\ &= \int_{A_0} \mu(dX_1) \int_{A_1} p(X_0, dX_1) \cdots \int_{A_m} g_0(X_m)p(X_{m-1}, dX_m) \end{aligned}$$

<sup>1</sup>所有这样的集合是一个  $\pi$  类, 生成整个  $\mathcal{F}_m$ .

<sup>2</sup>根据单调类定理可得.

$$\begin{aligned}
& \times \int g_1(X_{m+1})p(X_m, dX_{m+1}) \cdots \int g_n(X_{m+n})p(X_{m+n-1}, dX_{m+n}) \\
& = \int_{A_0} \mu(dX_1) \int_{A_1} p(X_0, dX_1) \cdots \int_{A_m} \mathbb{E}_{X_m}[Y]p(X_{m-1}, dX_m) \\
& = \mathbb{E}_\mu[\mathbb{E}_{X_m}[Y] \cdot I_A] = (2.2.7)RHS
\end{aligned}$$

□

### 2.2.3 强马氏性

#### 定义 2.2.3

$N$  是关于  $\{\mathcal{F}_n\}$  的停时, 定义:

$$\begin{aligned}
\mathcal{F}_N &= \{A : A \cap \{N \leq n\} \in \mathcal{F}_n, \forall n\} \\
\theta_N(\omega) &= \begin{cases} \theta_n(\omega) & , \text{ on } \{N = n\} \\ \text{whatever} & , \text{ on } \{N = +\infty\} \end{cases}
\end{aligned}$$



#### 定理 2.2.4 (强马氏性)

r.v.  $Y : \Omega \rightarrow \mathbb{R}$  有界, 则在  $\{N < +\infty\}$  上有

$$\mathbb{E}_\mu[Y \circ \theta_N | \mathcal{F}_N] = \mathbb{E}_{X_N}[Y] \quad (2.2.8)$$

**证明** 右侧是  $X_N$  的函数, 所以是  $\mathcal{F}_N$ -可测的, 接下来验证  $\forall A \in \mathcal{F}_N$ , 有

$$\mathbb{E}_\mu[Y \circ \theta_N \cdot I_A] = \mathbb{E}_\mu[\mathbb{E}_{X_N}[Y] I_A]$$

因为仅考虑  $N < +\infty$  的情况, 将其拆分:

$$\begin{aligned}
\mathbb{E}_\mu[Y \circ \theta_N \cdot I_A] &= \sum_{n=0}^{\infty} \mathbb{E}_\mu[Y \circ \theta_n \cdot \frac{I_{A \cap \{N=n\}}}{\mathbb{1}_{\{N=n\}}}] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_\mu[\mathbb{E}_\mu[Y \circ \theta_n \cdot I_{A \cap \{N=n\}} | \mathcal{F}_n]] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_\mu[\mathbb{E}_\mu[Y \circ \theta_n | \mathcal{F}_n] I_{A \cap \{N=n\}}] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_\mu[\mathbb{E}_{X_n}[Y] I_{A \cap \{N=n\}}] \\
&= \mathbb{E}_\mu[\mathbb{E}_{X_N}[Y]; I_A]
\end{aligned}$$

□

## 2.3 马氏链状态的分类

### 2.3.1 常返态与瞬时态

先约定一些记号:

1.  $k$  次转移概率记为  $p^k(x, y) = \mathbb{P}(X_k = y | X_0 = x) = \mathbb{P}_x(x_k = y)$ .
2. 随机变量  $N(y) = \sum_{n=1}^{\infty} I_{X_n=y}$ , 为回到  $y$  的总次数.
3. 规定  $T_y^0 = 0$ ,

$$T_y^k = \inf\{n > T_y^{k-1} : X_n = y\}, k \geq 1$$

这些都是停时, 为第  $k$  次到达  $y$  的时刻. 简记  $T_y = T_y^1$ , 为首次到达时刻. 对于  $x \in S$ , 记  $\rho_{xy} = \mathbb{P}_x(T_y < +\infty)$ , 为从  $x$  出发、有限时间内到达  $y$  的概率.

#### 定义 2.3.1

如果  $\rho_{yy} = 1$ , 称状态  $y$  是常返的 (recurrent), 否则称为瞬时的 (transient).

#### 定义 2.3.2

考虑状态空间  $S$  的子集  $C$ , 如果

- (1).  $x \in C, \rho_{xy} > 0 \Rightarrow y \in C$ , 称  $C$  封闭 (closed).
- (2).  $x, y \in C \Rightarrow \rho_{xy} > 0$ , 称  $C$  是不可约的 (irreducible).

#### 定理 2.3.1

$$\mathbb{P}_x(T_y^k < +\infty) = \rho_{xy}\rho_{yy}^{k-1} \quad (2.3.1)$$

**证明**  $k = 1$ , 式 (2.3.1) 显然成立, 下面归纳证明  $k > 1$  的情况. 若  $k - 1$  成立, 则令  $N = T_y^{k-1}$ , 注意到  $T_y^k = N + T_y \circ \theta_N$ , 因为:

$$\begin{aligned} T_y^k(\omega) &= \inf\{n > N : X_n(\omega) = \omega_n = y\} \\ &= \inf\{n - N > 0 : X_n(\omega) = \omega_n = y\} \\ &= N + \inf\{m > 0 : X_m(\omega) \circ \theta_N = \omega_n = y\} \\ &= N + T_y \circ \theta_N \end{aligned}$$

然后我们就可以得到结论了:

$$\begin{aligned} \mathbb{P}_x(T_y^k < +\infty) &= \mathbb{P}_x(T_y^k < +\infty, N < +\infty) \\ &= \mathbb{P}_x(T_y \circ \theta_N < +\infty, N < +\infty) \\ &= \mathbb{E}_x[I_{\{N < +\infty\}} \cdot I_{\{T_y \circ \theta_N < +\infty\}}] \\ &= \mathbb{E}_x[I_{T_y < +\infty} \circ \theta_N \cdot I_{\{N < +\infty\}}] \\ &= \mathbb{E}_x[\mathbb{E}_x[I_{T_y < +\infty} \circ \theta_N \cdot I_{\{N < +\infty\}} | \mathcal{F}_N]] \\ &= \mathbb{E}_x[I_{\{N < +\infty\}} \cdot \mathbb{E}_x[I_{\{T_y < +\infty\}} \circ \theta_N | \mathcal{F}_N]] \\ &= \mathbb{E}_x[I_{\{N < +\infty\}} \mathbb{E}_{X_N}[I_{\{T_y < +\infty\}}]] \\ &= \mathbb{E}_x[I_{\{N < +\infty\}} \mathbb{E}_y[I_{\{T_y < +\infty\}}]] \\ &= \mathbb{E}_x[I_{\{N < +\infty\}} \mathbb{P}_y(T_y < +\infty)] \\ &= \rho_{yy} \cdot \mathbb{P}_y(N < +\infty) \end{aligned}$$

By 强马氏性

$$= \rho_{yy}\rho_{xy}\rho_{yy}^{k-2} = \rho_{xy}\rho_{yy}^{k-1}$$

□

**推论 2.3.1**

如果  $\rho_{yy} < 1$ , 则  $N(y)$  的期望有限, 为

$$\mathbb{E}_x[N(y)] = \frac{\rho_{xy}}{1 - \rho_{yy}}$$

♥

**证明** 求和即可:

$$\mathbb{E}_x[N(y)] = \sum_{k=1}^{\infty} \mathbb{P}_x(N(y) \geq k) = \sum_{k=1}^{\infty} \mathbb{P}_x(T_y^k < +\infty) = \sum_{k=1}^{\infty} \rho_{xy}\rho_{yy}^{k-1} = \frac{\rho_{xy}}{1 - \rho_{yy}}$$

□

**推论 2.3.2**

$\forall x, y \in S$ , 记

$$P_{xy} = \sum_{n=1}^{\infty} p^n(x, y)$$

则

- (1).  $\rho_{xy} > 0 \Leftrightarrow P_{xy} > 0$ .<sup>a</sup>
- (2).  $y$  常返  $\Leftrightarrow P_{xy} = +\infty$ .

<sup>a</sup> $P_{xy} > 0$  正是  $x$  可达  $y$  的定义, 随机过程课上没有介绍后文却用到了相关结论, 出于严谨笔者将那些结论补充为这个推论。

♥

**证明** 注意到

$$\frac{\rho_{xy}}{1 - \rho_{yy}} = \mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} \mathbb{E}_x[I_{\{X_n=y\}}] = \sum_{n=1}^{\infty} p^n(x, y) = P_{xy}$$

□

**定理 2.3.2**

$y$  常返  $\Rightarrow N(y) = +\infty$  a.s.

**证明** 如果  $y$  常返, 则  $\forall k \geq 1$ ,

$$\mathbb{P}_y(N(y) \geq k) = \mathbb{P}_y(T_y^k < +\infty) = \rho_{yy}^k = 1$$

这意味着  $N(y) = +\infty$  a.s.

□

**定理 2.3.3**

如果  $x$  常返, 且  $\rho_{yx} > 0$ , 则  $y$  常返且  $\rho_{yx} = 1$ .

**证明** (反证) 假设  $\rho_{yx} < 1$ , 令  $K = \inf\{k : p^k(x, y) > 0\}$ , 注意到

$$\rho_{xy} = \mathbb{P}_x\left(\bigcup_{n=1}^{\infty} \{X_n = y\}\right)$$

所以  $p^K(x, y) > 0 \Rightarrow$  存在一系列  $y_1, \dots, y_{K-1}$  使得

$$p(x, y_1)p(y_1, y_2) \cdots p(y_{K-1}, y) > 0$$

且  $y_1, \dots, y_{K-1} \neq y$ , 因为  $K$  是首次到达  $y$  的时刻。考虑

$$p(x, y_1)p(y_1, y_2) \cdots p(y_{K-1}, y)(1 - \rho_{yx}) > 0$$

这说明路径  $x \rightarrow y_1 \rightarrow y_2 \rightarrow \cdots \rightarrow y \rightarrow x$  的概率不为零, 这与  $x$  常返矛盾。

下面证明  $y$  常返, 由于  $\rho_{yx} > 0$ , 存在  $L$  使得  $p^L(y, x) > 0$ , 于是

$$p^{L+n+K}(y, y) \geq p^L(y, x)p^n(x, x)p^K(x, y)$$

$$\sum_n p^{L+n+K}(y, y) \geq p^L(y, x) \left( \sum_n p^n(x, x) \right) p^K(x, y) = +\infty$$

所以  $y$  常返。  $\square$

**注** 这个定理说明了两个事实: 如果  $x$  常返, 则  $\rho_{xy}$  要么是 1, 要么是 0; 对于不可约的状态子集  $C$ , 里面的状态常返/瞬时都一致, 即有一个常返态则全都是常返态、有一个瞬时态则全都是瞬时态。

#### 定理 2.3.4

$C$  是封闭的有限集, 则  $C$  包含至少一个常返态。(进一步地, 如果  $C$  不可约, 则  $C$  中所有状态都是常返的。)

**证明** (反证) 假设不成立, 即  $\forall y \in C, \rho_{yy} < 1$ , 于是

$$\mathbb{E}_x[N(y)] = \frac{\rho_{xy}}{1 - \rho_{yy}} < +\infty$$

同时又有

$$\mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} \mathbb{E}_x[I_{\{X_n=y\}}] = \sum_{n=1}^{\infty} p^n(x, y)$$

于是

$$+\infty > \sum_{y \in C} \mathbb{E}_x[N(y)] = \sum_{y \in C} \sum_{n=1}^{\infty} p^n(x, y) = \sum_{n=1}^{\infty} 1 = +\infty$$

矛盾。  $\square$

#### 推论 2.3.3

若状态空间  $S$  有限, 对于  $x \in S$ ,

- (1). 如果  $\exists y \in S$  s.t.  $\rho_{xy} > 0$  且  $\rho_{yx} = 0$ , 则  $x$  是瞬时的。
- (2). 如果不存在 (1) 中这样的  $y$ , 即  $\forall y \in S, \rho_{xy} > 0 \Rightarrow \rho_{yx} > 0$ , 则  $x$  是常返的。

如果令  $C_x = \{y : \rho_{xy} > 0\}$ , 则  $C_x$  是不可约、封闭且有限的。  $\heartsuit$

#### 定理 2.3.5

所有常返态  $R = \{x : \rho_{xx} = 1\}$ , 可以将其划分为  $R = \sqcup R_i$ , 每个  $R_i$  闭且不可约。

**证明** 对于每个  $x \in R$ , 我们令

$$C_x = \{z \in R : \rho_{xz} > 0, \rho_{zx} > 0\}$$

那么很显然, 每个  $C_x$  都是闭且不可约的, 只需要证明: 对于  $x, y \in R$ , 要么  $C_x \cap C_y = \emptyset$ 、要么  $C_x = C_y$ , 那么定理就得证。假设  $C_x \cap C_y \neq \emptyset$ , 取  $z \in C_x \cap C_y$ , 则  $\rho_{xy} \geq \rho_{xz}\rho_{zy} > 0$ , 任取  $w \in C_y$ ,

$$\rho_{xw} \geq \rho_{xy}\rho_{yw} > 0 \Rightarrow w \in C_x$$

反过来也成立, 这说明  $C_x = C_y$ 。  $\square$

**注** 这个定理也表明所有马氏链都能划分为瞬时态和一系列闭、不可约的子链, 每一个子链都可以视为一条新的马氏链。

## 2.3.2 进一步分类与判断\*

## 2.3.2.1 常返与瞬时

回顾定义 2.3.1, 常返态的定义为

$$\rho_{ii} = \mathbb{P}(T_i < +\infty | X_0 = i) = \mathbb{P}(\exists n \geq 1, X_n = i | X_0 = i) = 1$$

接下来, 记

$$f_{ij}(n) = \mathbb{P}(T_j = n | X_0 = i), \quad f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$$

## 命题 2.3.1

$f_{ij} = \rho_{ij}$ , 也就是

$$\sum_{n=1}^{\infty} \mathbb{P}(T_j = n | X_0 = i) = \mathbb{P}(\exists n \geq 1, X_n = j | X_0 = i)$$

**证明** 其实非常直观, 只需要注意到事件  $A_n = \{T_j = n, X_0 = i\}$  互不相交就可以了, 这是因为

$$A_n = \{X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j, X_0 = i\}$$

而且

$$\bigsqcup_{n=1}^{\infty} A_n = \{T_j < +\infty, X_0 = i\}$$

然后就可以得到

$$LHS = \sum_{n=1}^{\infty} \frac{\mathbb{P}(A_n)}{\mathbb{P}(X_0 = i)} = \frac{\mathbb{P}(T_j < +\infty, X_0 = i)}{\mathbb{P}(X_0 = i)} = \mathbb{P}(T_j < +\infty | X_0 = i) = RHS$$

□

这样我们就把  $\rho_{ij}$  拆成了级数的形式, 那么就可以通过研究级数的收敛性来判断状态是否常返了。

设生成函数:

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

并规定  $p_{ij}(0) = \delta_{ij}$ ,  $f_{ij}(0) = 0$ , 可以发现:

- (1).  $f_{ij} = F_{ij}(1)$ .
- (2).  $|s| < 1$  时,  $P_{ij}(s), F_{ij}(s) < \infty$ .

于是由 Abel 定理<sup>3</sup>, 可知  $s \nearrow 1$  时  $P_{ij}(s) \rightarrow P_{ij}(1)$ ,  $F_{ij}(s) \rightarrow F_{ij}(1)$ .

## 定理 2.3.6

- (1).  $P_{ii}(s) = 1 + F_{ii}(s)P_{ii}(s)$ .
- (2).  $P_{ij}(s) = F_{ij}(s)P_{jj}(s), i \neq j$ .

**证明** 先来研究一下  $p_{ij}(m)$ , 对于事件  $\{X_m = j, X_0 = i\}$ , 可以利用首达时对其拆分:

$$\{X_m = j, X_0 = i\} = \bigsqcup_{r=1}^m \{X_m = j, T_j = r, X_0 = i\}$$

<sup>3</sup>复分析里提到过, 这里还是说一下: 如果级数

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

的收敛半径为 1, 且级数在  $z = 1$  处收敛于  $S$ , 那么  $f$  在  $z = 1$  有非切向极限  $S$ . 对于本课程, 我们只需要考虑其沿实轴趋于 1 的极限, 这当然是非切向极限。

于是

$$\begin{aligned}
 p_{ij}(m) &= \frac{\mathbb{P}(X_m = j, X_0 = i)}{\mathbb{P}(X_0 = i)} \\
 &= \sum_{r=1}^m \mathbb{P}(X_m = j, T_j = r | X_0 = i) \\
 &= \sum_{r=1}^m \mathbb{P}(X_m = j, X_r = j, X_{r-1} \neq j, \dots, X_1 \neq j | X_0 = i) \\
 &= \sum_{r=1}^m \mathbb{P}(X_m = j | X_r = j, X_{r-1} \neq j, \dots, X_1 \neq j, X_0 = i) \mathbb{P}(X_r = j, X_{r-1} \neq j, \dots, X_1 \neq j | X_0 = i) \\
 &= \sum_{r=1}^m \mathbb{P}(X_m = j | X_r = j, X_{r-1} \neq j, \dots, X_1 \neq j, X_0 = i) f_{ij}(r)
 \end{aligned}$$

回顾命题 2.1.1 的 (2) 和例 2.1.2 的 (1), 可以得到

$$\mathbb{P}(X_m = j | X_r = j, X_{r-1} \neq j, \dots, X_1 \neq j, X_0 = i) = \mathbb{P}(X_m = j | X_r = j) = p_{jj}(m-r)$$

这就得到

$$p_{ij}(m) = \sum_{r=1}^m p_{jj}(m-r) f_{ij}(r), \quad m \geq 1$$

于是回到  $P_{ij}$ ,

$$\begin{aligned}
 P_{ij}(s) &= \sum_{n=1}^{\infty} s^n p_{ij}(n) + p_{ij}(0) \\
 &= \sum_{n=1}^{\infty} \sum_{r=1}^n s^n p_{jj}(n-r) f_{ij}(r) + \delta_{ij} \\
 &= \sum_{r=1}^{\infty} \sum_{n=r}^{\infty} s^n p_{jj}(n-r) f_{ij}(r) + \delta_{ij} && \text{换求和顺序} \\
 &= \sum_{r=1}^{\infty} \sum_{t=0}^{\infty} s^{t+r} p_{jj}(t) f_{ij}(r) + \delta_{ij} && \text{令 } t = n - r \\
 &= \sum_{r=0}^{\infty} \sum_{t=0}^{\infty} s^{t+r} p_{jj}(t) f_{ij}(r) + \delta_{ij} && \text{补一项 } f_{ij}(0) = 0 \\
 &= P_{jj}(s) \cdot F_{ij}(s) + \delta_{ij}
 \end{aligned}$$

于是得证。 □

### 推论 2.3.4

(1).  $j$  常返  $\Leftrightarrow$

$$\sum_{n=0}^{\infty} p_{jj}(n) = +\infty$$

(2).  $j$  常返且  $f_{ij} > 0 \Rightarrow$

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

(3).  $j$  瞬时, 则  $\forall i,$

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

进而  $\lim_{n \rightarrow \infty} p_{ij}(n) = 0.$

**证明** 利用

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

以及 Abel 定理和一些处理级数的基本技巧折腾一下就能倒出来了。□

接下来, 我们引入状态的返回次数这一概念, 进一步研究常返与瞬时。

### 定义 2.3.3

对于随机过程  $\{X_n, n \in \mathbb{N}\}$ , 我们给定条件  $\mathbb{P}(X_0 = i) = 1$ , 可以定义随机变量  $N_i(j)$ , 为  $\{X_1, X_2, \dots\}$  中经过  $j$  的次数, 即

$$N_i(j)(\omega) = \sum_{n=1}^{\infty} I_{\{X_n(\omega)=j\}}$$

回顾定理 2.3.1, 我们用比较初等的语言再证明一遍。

### 命题 2.3.2

$$\mathbb{P}(N_i(j) = n) = \begin{cases} 1 - f_{ij} & , n = 0 \\ f_{ij}(f_{jj})^{n-1}(1 - f_{jj}) & , n \geq 1 \end{cases}$$

**证明**  $n = 0$  显然, 对于  $n \geq 1$ ,

$$\begin{aligned} \mathbb{P}(N_i(j) = n) &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i, X_{k_1} = j, \dots, X_{k_n} = j, X_k \neq j \text{ for other } k \geq 1) \\ &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i) \cdot \mathbb{P}(X_{k_1} = j, \dots, X_{k_n} = j, X_k \neq j \text{ for other } k \geq 1 | X_0 = i) \\ &= (\star 1) \end{aligned}$$

后面那个很复杂的项, 我们沿时间线上的  $k_1$  切开,  $\leq k_1$  为  $A$ ,  $> k_1$  为  $B$ , 后面的条件为  $C$ , 那么根据

$$\mathbb{P}(AB|C) = \mathbb{P}(A|BC) \cdot \mathbb{P}(B|C)$$

我们可以进一步拆分:

$$\begin{aligned} (\star 1) &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i) \cdot \mathbb{P}(X_{k_1} = j, X_k \geq j \text{ for } 1 \leq k \leq k_1 - 1 | X_0 = i) \cdot \\ &\quad \mathbb{P}(X_{k_2} = j, \dots, X_{k_n} = j, X_k \neq j \text{ for other } k \geq k_1 + 1 | X_0 = i, X_{k_1} = j, X_k \geq j \text{ for } 1 \leq k \leq k_1 - 1) \\ &= (\star 2) \end{aligned}$$

注意到第二项就是“ $i$  出发  $k_1$  时首次到达  $j$ ”, 也就是  $f_{ij}(k_1)$ ; 第三项则由马氏性, 可以转化为

$$\begin{aligned} (\star 2) &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i) \cdot f_{ij}(k_1) \cdot \mathbb{P}(X_{k_2} = j, \dots, X_{k_n} = j, X_k \neq j \text{ for other } k \geq k_1 + 1 | X_{k_1} = j) \\ &= (\star 3) \end{aligned}$$

注意这时第三项回到了类似于  $(\star 1)$  式的形式, 我们以此类推继续沿时间线上的  $k_2$  切开, 最后得到:

$$\begin{aligned} (\star 3) &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i) \cdot f_{ij}(k_1) \cdot f_{jj}(k_2 - k_1) \mathbb{P}(X_{k_3} = j, \dots, X_{k_n} = j, X_k \neq j \text{ for other } k \geq k_2 + 1 | X_{k_2} = j) \\ &= \dots \\ &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i) \cdot f_{ij}(k_1) \cdot f_{jj}(k_2 - k_1) f_{jj}(k_3 - k_2) \cdots f_{jj}(k_n - k_{n-1}) \mathbb{P}(X_k \neq j \text{ for } k > k_n | X_{k_n} = j) \\ &= \sum_{1 \leq k_1 < k_2 < \dots < k_n} \mathbb{P}(X_0 = i) \cdot f_{ij}(k_1) \cdot f_{jj}(k_2 - k_1) f_{jj}(k_3 - k_2) \cdots f_{jj}(k_n - k_{n-1}) (1 - f_{jj}) \end{aligned}$$

$$= \mathbb{P}(X_0 = i)(1 - f_{jj}) \sum_{1 \leq k_1 < k_2 < \dots < k_n} f_{ij}(k_1) \cdot f_{jj}(k_2 - k_1) f_{jj}(k_3 - k_2) \cdots f_{jj}(k_n - k_{n-1})$$

后面这一堆求和其实就是  $f_{ij} f_{jj}^{n-1}$ ，可以换元： $t_2 = k_2 - k_1, t_3 = k_3 - k_2$  等等，就能看出来了。<sup>4</sup> □

从直观上理解：注意  $f_{ij}$  代表从  $i$  出发最终到达  $j$  的概率， $n = 0$  即再也不到达  $j$ ，所以概率为  $1 - f_{ij}$ ；对于  $n \geq 1$  的情况， $i$  出发经过了  $j$  共计  $n$  次，首次到达为  $f_{ij}$ ，之后  $j$  出发返回  $j$  共计  $n - 1$  次为  $f_{jj}^{n-1}$ ，最终再也不回来为  $1 - f_{jj}$ 。

对前  $n$  项求和得到： $\mathbb{P}(N_i(i) \leq n) = 1 - f_{ii}^{n+1}$ ，这就说明

$$\mathbb{P}(N_i(i) < +\infty) = \begin{cases} 1 & , f_{ii} < 1 \\ 0 & , f_{ii} = 1 \end{cases}$$

则  $i$  常返  $\Leftrightarrow \mathbb{P}(N_i(i) = +\infty)$ ，这意味着常返态有 1 的概率回到自身无穷多次。

### 2.3.2.2 零常返与正常返

接下来我们进一步对常返态进行分类。

#### 定义 2.3.4

对于状态  $i$ ，平均常返时间：

$$\mu_i \stackrel{\text{def}}{=} \mathbb{E}[T_i | X_0 = i] = \begin{cases} \sum_{n=1}^{\infty} n f_{ii}(n) & , i \text{ 常返} \\ +\infty & , i \text{ 瞬时} \end{cases}$$

如果  $i$  常返且  $\mu_i = +\infty$ ，称其零常返；如果  $i$  常返且  $\mu_i < +\infty$ ，称其非零常返或正常返。

#### 定理 2.3.7

若  $i$  常返，则  $i$  零常返  $\Leftrightarrow \lim_{n \rightarrow \infty} p_{ii}(n) = 0$ 。

这个定理的证明不要求掌握。

#### 定义 2.3.5

对于状态  $i$ ，周期：

$$d(i) = \gcd\{n \geq 1 : p_{ii}(n) > 0\}$$

如果  $d(i) > 1$ ，称  $i$  为周期的，反之为非周期的。

如果状态  $i$  的周期为  $i$ ，说明至少存在两个互质的  $u, v$  使得  $p_{ii}(u), p_{ii}(v) > 0$ ，因此，根据裴蜀定理，存在一个充分大的  $N$  使得  $n > N$  时， $n$  总能写成一些  $u$  和  $v$  的和，即  $n = k_1 u + k_2 v$ ，进而  $p_{ii}(n) \geq p_{ii}(u)^{k_1} p_{ii}(v)^{k_2} > 0$ 。这个结果意味着：对于非周期的状态  $i$ ，在充分大的时刻之后，每一个时间点都有可能返回。有点欧拉筛的感觉。

#### 定义 2.3.6

对于状态  $i$ ，如果它非零常返且非周期，称其为遍历的 (aperiodic)。

### 2.3.3 实例：简单对称随机游走

我们一般利用推论 2.3.2 来判断常返/瞬时，利用定理 2.3.7（或者直接找平稳分布，后文会介绍）来判断零常返/正常返。

下面以简单对称随机游走为例介绍这些定理的应用，注意随机游走的各个状态之间都是相互可达的，所以不可约，我们只需要分析原点的状态即可。

<sup>4</sup>至于  $\mathbb{P}(X_0 = i)$ ，按定义来说应该得是 1 的，结果也如此，但既然如此一开始何必带着这玩意儿算呢？这里属实没搞懂。

## 引理 2.3.1

Stirling 公式:  $n \rightarrow \infty$  时,

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n$$

为同阶无穷大。



## 例 2.3.1.

一维情形: 随机过程  $X = (X_n)_{n \in \mathbb{N}}$ , 状态空间为  $\mathbb{Z}$ ,  $X_0 = 0$  a.s., 给定  $X_n$  状态后,  $X_{n+1}$  等概率地向随机一个方向 (左、右) 移动一步。则  $X$  是一个马氏链, 分析每个状态的常返性 (瞬时/常返/正常返/零常返)。

**证明** 考虑  $2n$  步之后回到原点的概率:

$$p(2n) = 2^{-2n} \frac{(2n)!}{n!n!}$$

由 Stirling 公式,

$$p(2n) \sim n^{-\frac{1}{2}}$$

所以  $\sum p(2n)$  发散, 但  $\lim_{n \rightarrow \infty} p(2n) = 0$ , 所以原点是零常返的, 因此所有状态都是零常返的。□

## 例 2.3.2.

二维情形: 随机过程  $X = (X_n)_{n \in \mathbb{N}}$ , 状态空间为  $\mathbb{Z} \times \mathbb{Z}$ ,  $X_0 = (0, 0)$  a.s., 给定  $X_n$  状态后,  $X_{n+1}$  等概率地向随机一个方向 (左、右、上、下) 移动一步。则  $X$  是一个马氏链, 分析每个状态的常返性 (瞬时/常返/正常返/零常返)。

**证明** 考虑  $2n$  步之后回到原点的概率:

$$p(2n) := \mathbb{P}(X_{2n} = (0, 0)) = 4^{-2n} \sum_{m=0}^n \frac{(2n)!}{m!m!(n-m)!(n-m)!} = 4^{-2n} \frac{(2n)!}{n!n!} \sum_{m=0}^n \frac{n!}{m!(n-m)!} \cdot \frac{n!}{m!(n-m)!}$$

这里要用到一个组合的技巧, 设  $(1+x)^{2n}$  中  $x^n$  的系数为  $C_n$ , 直接展开得到

$$C_n = \binom{2n}{n} = \frac{(2n)!}{n!n!}$$

另一方面, 拆分成  $(1+x)^n \cdot (1+x)^n$ , 设  $(1+x)^n$  中  $x^m$  的系数是  $B_m$ ,

$$C_n = \sum_{m=0}^n B_m \cdot B_{n-m} = \sum_{m=0}^n \binom{n}{m} \cdot \binom{n}{n-m} = \sum_{m=0}^n \frac{n!}{m!(n-m)!} \cdot \frac{n!}{m!(n-m)!}$$

这就说明

$$p(2n) = 4^{-2n} \left(\frac{(2n)!}{n!n!}\right)^2$$

借助 Stirling 公式, 可得

$$p(2n) \sim \frac{1}{n} 4^{2n}$$

所以

$$\sum_{n=0}^{\infty} p(2n) = +\infty, \quad \lim_{n \rightarrow \infty} p(2n) = 0$$

所以原点是零常返的, 从而所有状态都是零常返的。□

从另一个角度看, 一、二维的随机游走是不可能存在平稳分布的, 所以如果常返则肯定是零常返的, 但三维情形下是瞬时的。

## 例 2.3.3.

三维情形：随机过程  $X = (X_n)_{n \in \mathbb{N}}$ ，状态空间为  $\mathbb{Z}^3$ ， $X_0 = (0, 0, 0)$  a.s.，给定  $X_n$  状态后， $X_{n+1}$  等概率地向随机一个方向（前、后、左、右、上、下）移动一步。则  $X$  是一个马氏链，分析每个状态的常返性（瞬时/常返/正常返/零常返）。

**证明** 考虑  $2n$  步之后回到原点的概率：

$$p(2n) = 6^{-2n} \sum_{i+j+k=2n} \frac{(2n)!}{i!j!k!} = 2^{-2n} \frac{(2n)!}{n!n!} \sum_{i+j+k=n} \left( 3^{-n} \frac{n!}{i!j!k!} \right)^2$$

考虑

$$3^n = (1 + 1 + 1)^n = \sum_{i+j+k=n} \frac{n!}{i!j!k!}$$

所以

$$p(2n) \leq 2^{-2n} \frac{(2n)!}{n!n!} \max_{i+j+k=n} 3^{-n} \frac{n!}{i!j!k!}$$

根据 Stirling 公式，

$$2^{-2n} \frac{(2n)!}{n!n!} \sim \frac{1}{\sqrt{n}}$$

下证：

$$\alpha_{i,j,k} = 3^{-n} \frac{n!}{i!j!k!}, \max_{i+j+k=n} \alpha_{i,j,k} \sim n^{-1}$$

事实上，

(1). 如果  $\max_{i+j+k=n} \alpha_{i,j,k} = \alpha_{i',j',k'}$ ，则  $i', j', k' \in \{\lfloor \frac{n}{3} \rfloor, \lceil \frac{n}{3} \rceil + 1\}$ ；

(2). 由 Stirling 公式，

$$\alpha_{i,j,k} \sim 3^{-n} \frac{n^n}{i^{i'} j^{j'} k^{k'}} \sqrt{\frac{n}{i' j' k'}} \cdot \frac{1}{2\pi}$$

(3).

$$i^{i'} j^{j'} k^{k'} \sim \left( \left( \frac{n}{3} \right)^{\frac{n}{3}} \right)^3 = \left( \frac{n}{3} \right)^n$$

$$i' j' k' \sim n^3$$

从而

$$\alpha_{i',j',k'} \sim 3^{-n} \cdot \frac{n^n}{\left( \frac{n}{3} \right)^n} \cdot \sqrt{\frac{n}{n^3}} \cdot \frac{1}{2\pi} \sim n^{-1}$$

因此， $n$  充分大时

$$p(2n) \leq C n^{-\frac{3}{2}}$$

所以  $\sum p(2n)$  收敛，这说明原点是瞬时态，从而所有状态都是瞬时态。  $\square$

直观上来看这是一个很有趣的事实，醉汉在地面上随机游走总能回家，但喝醉的小鸟在空中“随机游走”就回不了家。

## 2.4 平稳测度与平稳分布

### 2.4.1 基本定义

#### 定义 2.4.1

测度  $\mu$  如果满足:

$$\sum_x \mu(x)p(x, y) = \mu(y)$$

则称之为平稳测度。

进一步地, 如果  $\mu$  是概率测度, 则称之为平稳分布; 如果  $\mu$  满足:

$$\mu(x)p(x, y) = \mu(y)p(y, x)$$

则称之为可逆测度。

#### 定理 2.4.1

$\mu$  是平稳分布, 且  $X_0$  的分布也是  $\mu$ , 令  $Y_m = X_{n-m}$ ,  $-\infty < m \leq n$ , 则  $\{Y_m, -\infty < m \leq n\}$  是马氏链, 转移概率  $q(x, y) = \mu(y)p(y, x)/\mu(x)$ .

进一步地, 如果  $\mu$  是可逆测度, 则  $p = q$ .

**证明**  $Y$  是“倒过来”的  $X$ , 利用贝叶斯公式翻转一下:

$$\begin{aligned} q(x, y) &= \mathbb{P}(Y_{m+1} = y | Y_m = x) \\ &= \mathbb{P}(X_{n-m-1} = y | X_{n-m} = x) \\ &= \frac{\mathbb{P}(X_{n-m} = x | X_{n-m-1} = y) \mathbb{P}(X_{n-m-1} = y)}{\mathbb{P}(X_{n-m} = x)} = \frac{\mu(y)p(y, x)}{\mu(x)} \end{aligned}$$

□

### 2.4.2 平稳测度的存在唯一性

#### 定理 2.4.2 (存在性)

$x$  常返,  $T_x = \inf\{m \geq 1 : X_m = x\}$ , 则定义

$$\mu_x(y) \stackrel{\text{def}}{=} \mathbb{E}_x \left[ \sum_{n=0}^{T_x-1} I_{\{X_n=y\}} \right] = \mathbb{E}_x \left[ \sum_{n=0}^{\infty} I_{\{X_n=y, T_x > n\}} \right] = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, T_x > n)$$

**证明:**  $(\mu_x(y), y \in S)$  是一个平稳测度。

**证明** 首先, 如果考虑  $X_0 = x$ , 则返回  $x$  前只会经过  $x$  一次, 即初始状态的一次:

$$\mu_x(x) = \mathbb{E}_x \left[ \sum_{n=0}^{T_x-1} I_{\{X_n=x\}} \right] = \mathbb{E}_x [I_{\{X_0=x\}}] = 1$$

我们的目标是要证明:

$$\mu_x(z) = \sum_{y \in S} \mu_x(y)p(y, z), \quad \forall z \in S \quad (\star 1)$$

注意到

$$\begin{aligned} \mathbb{P}_x(X_n = y, X_{n+1} = z, T_x > n) &= \mathbb{E}_x [I_{\{X_n=y, T_x > n\}} I_{\{X_{n+1}=z\}}] \\ &= \mathbb{E}_x [\mathbb{E}_x [I_{\{X_n=y, T_x > n\}} I_{\{X_{n+1}=z\}} | \mathcal{F}_n]] \\ &= \mathbb{E}_x [I_{\{X_n=y, T_x > n\}} \mathbb{E}_x [I_{\{X_{n+1}=z\}} | \mathcal{F}_n]] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_x[I_{\{X_n=y, T_x>n\}} \mathbb{E}_{X_n}[I_{\{X_1=z\}}]] \\
&= \mathbb{E}_x[I_{\{X_n=y, T_x>n\}} p(X_n, z)] \\
&= \mathbb{P}_x(X_n = y, T_x > n) p(y, z)
\end{aligned}$$

所以

$$\begin{aligned}
(\star)RHS &= \sum_{y \in S} \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, T_x > n) p(y, z) \\
&= \sum_{n=0}^{\infty} \sum_{y \in S} \mathbb{P}_x(X_n = y, X_{n+1} = z, T_x > n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1} = z, T_x > n)
\end{aligned}$$

如果  $z \neq x$ , 则

$$\sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1} = z, T_x > n) = \sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1} = z, T_x > n+1) = \mu_x(z)$$

如果  $z = x$ , 则

$$\sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1} = z, T_x > n) = \mathbb{P}_x(T < +\infty) = 1 = \mu_x(x)$$

得证。 □

上述定理表明, 只要马氏链中存在常返态, 则能构造出平稳测度。

#### 定理 2.4.3 (唯一性)

如果马氏链  $X$  不可约、常返, 则其平稳测度在相差常数倍意义下唯一。

**证明** 取  $a \in S$ , 由定理 2.4.2,  $\{\mu_a(y)\}$  是一个平稳测度, 假设  $\{v(y)\}$  是另一个平稳测度, 我们希望证明: 存在常数  $c > 0$  使得  $v(y) = c\mu_a(y)$ .

一方面, 根据平稳测度的性质, 我们得到

$$\begin{aligned}
v(z) &= \sum_{y \in S} v(y) p(y, z) \\
&= v(a) p(a, z) + \sum_{y \neq a} v(y) p(y, z)
\end{aligned} \tag{\star}$$

把上式求和中的  $v(y)$  拆开, 得到

$$\begin{aligned}
v(z) &= v(a) p(a, z) + \sum_{y \neq a} \left[ v(a) p(a, y) + \sum_{x \neq a} v(x) p(x, y) \right] p(y, z) \\
&= v(a) p(a, z) + \sum_{y \neq a} v(a) p(a, y) p(y, z) + \sum_{y \neq a} \sum_{x \neq a} v(x) p(x, y) p(y, z) \\
&= v(a) \mathbb{P}_a(X_1 = z) + v(a) \sum_{y \neq a} \mathbb{P}_a(X_1 = y, X_2 = z) + \sum_{y \neq a} \sum_{x \neq a} v(x) p(x, y) p(y, z) \\
&= v(a) \mathbb{P}_a(X_1 = z) + v(a) \mathbb{P}_a(X_1 \neq a, X_2 = z) + \sum_{y \neq a} \sum_{x \neq a} v(x) p(x, y) p(y, z)
\end{aligned}$$

再按照  $(\star)$  式把最后一项里的  $v(x)$  拆开,

$$\begin{aligned}
v(z) &= v(a) \mathbb{P}_a(X_1 = z) + v(a) \mathbb{P}_a(X_1 \neq a, X_2 = z) + \sum_{y \neq a} \sum_{x \neq a} \left[ v(a) p(a, x) + \sum_{w \neq a} v(w) p(w, x) \right] p(x, y) p(y, z) \\
&= v(a) \mathbb{P}_a(X_1 = z) + v(a) \mathbb{P}_a(X_1 \neq a, X_2 = z) + v(a) \mathbb{P}_a(X_1 \neq a, X_2 \neq a, X_3 = z) +
\end{aligned}$$

$$\sum_{y \neq a} \sum_{x \neq a} \sum_{w \neq a} v(w) p(w, x) p(x, y) p(y, z)$$

以此类推，把最后一项放缩到 0，我们最终得到

$$\begin{aligned} v(z) &\geq v(a) \mathbb{P}_a(X_1 = z) + v(a) \mathbb{P}_a(X_1 \neq a, X_2 = z) + \cdots + v(a) \mathbb{P}_a(X_k \neq a, 1 \leq k < n, X_n = z) + \cdots \\ &= v(a) \sum_{n=1}^{\infty} \mathbb{P}_a(X_k \neq a, 1 \leq k < n, X_n = z) \\ &= v(a) \sum_{n=1}^{\infty} \mathbb{P}_a(T_a > n, X_n = z) \\ &= v(a) \mu_a(z) \end{aligned}$$

另一方面，考虑

$$\begin{aligned} v(a) &= \sum_{x \in S} v(x) p^n(x, a) \\ &\geq \sum_{x \in S} v(a) \mu_a(x) p^n(x, a) \\ &= v(a) \sum_{x \in S} \mu_a(x) p^n(x, a) \\ &= v(a) \mu_a(a) = v(a) \end{aligned}$$

所以我们得到

$$v(z) = v(a) \mu_a(z)$$

□

### 2.4.3 平稳分布的性质

#### 定理 2.4.4

如果马氏链有平稳分布  $\pi$ ，则  $\pi(y) > 0 \Rightarrow y$  常返。

**证明** 平稳分布满足

$$\pi \mathbf{P}^n = \pi$$

所以

$$\sum_{x \in S} \pi(x) p^n(x, y) = \pi(y) > 0$$

从而

$$\sum_{n=1}^{\infty} \sum_{x \in S} \pi(x) p^n(x, y) = \sum_{n=1}^{\infty} \pi(y) = +\infty$$

假设  $y$  瞬时， $\rho_{yy} < 1$ ，则由推论 2.3.2 可知，

$$\sum_{n=1}^{\infty} p^n(x, y) = \frac{\rho_{xy}}{1 - \rho_{yy}}$$

所以

$$\sum_{n=1}^{\infty} \sum_{x \in S} \pi(x) p^n(x, y) = \sum_{x \in S} \pi(x) \frac{\rho_{xy}}{1 - \rho_{yy}} < +\infty$$

这就矛盾。

□

## 定理 2.4.5

如果不可约马氏链有平稳分布  $\pi$ , 则  $\forall \pi(x) > 0$ , 进一步地,

$$\pi(x) = \frac{1}{\mathbb{E}_x[T_x]}$$

其中  $T_x = \min\{n \geq 1 : X_n = x\}$ .

**证明** 假设存在某个  $\pi(y) = 0$ ,

$$\pi(y) = \sum_{x \in S} \pi(x) p^n(x, y) = 0$$

不可约所以  $\rho_{xy} > 0$ , 从而

$$\sum_{n=1}^{\infty} p^n(x, y) > 0$$

于是存在某个  $N$  使得  $p^N(x, y) > 0$ , 这就和

$$\pi(y) = \sum_{x \in S} \pi(x) p^N(x, y) = 0$$

矛盾。

任取一个  $x$ , 考虑平稳测度  $\mu_x$ , 存在常数  $c$  使得  $\mu_x(y) = c\pi(y)$ , 于是

$$\sum_{y \in S} \mu_x(y) = c$$

同时,

$$\sum_{y \in S} \mu_x(y) = \sum_{n=0}^{\infty} \sum_{y \in S} \mathbb{P}_x(X_n = y, T_x > n) = \sum_{n=0}^{\infty} \mathbb{P}_x(T_x > n) = \mathbb{E}_x[T_x]$$

所以

$$c = \mathbb{E}_x[T_x] = \frac{\mu_x(y)}{\pi(y)}$$

取  $x = y$  得证。 □

## 定义 2.4.2

$\mathbb{E}_x[T_x]$  称为平均返回时间, 如果有限, 则称  $x$  正常返, 否则称为零常返。 ♣

## 定理 2.4.6

对于不可约马氏链, 以下说法等价:

- (1). 存在正常返态;
- (2). 存在平稳分布;
- (3). 所有的状态都正常返。

**证明** 只说明一下 (1)  $\Rightarrow$  (2): 考虑平稳测度  $\mu_x$ ,

$$\sum_{y \in S} \mu_x(y) = \sum_{n=0}^{\infty} \sum_{y \in S} \mathbb{P}_x(X_n = y, T_x > n) = \mathbb{E}_x[T_x] < +\infty$$

和有限, 则取

$$v(y) = \frac{\mu_x(y)}{\mathbb{E}_x[T_x]}$$

就是平稳分布。 □

这个定理告诉我们, 不可约、常返马氏链中的状态要么都是正常返的, 要么都是零常返的, 而且如果存在平稳分布, 就是正常返的。

## 2.4.4 利用平稳分布判断状态 \*

本节来自应随课程，大部分都是重复内容，没有补充太多新东西，可直接跳过。

可以按照以下步骤来构造一个平稳测度/分布：

1. 取一个不可约、常返的马氏链  $X = \{X_n, n \in \mathbb{N}\}$ ，状态空间为  $S$ ，转移矩阵为  $\mathbf{P}$ 。
2. 定义首达时  $T_k$ ，注意  $k$  常返所以  $\mathbb{P}(T_k < +\infty | X_0 = k) = 1$ 。
3. 给定条件  $X_0 = k$  的情况下，定义

$$N_{k,i}(\omega) \stackrel{\text{def}}{=} \sum_{n=1}^{T_k} I_{\{X_n=i\}}$$

为从  $k$  出发回到  $k$  前经过  $i$  的次数。

4.  $\mathbb{P}(N_{k,k} = 1) = 1$ ， $T_k(\omega) = \sum_{i \in S} N_{k,i}(\omega)$ 。

5. 注意到

$$N_{k,i}(\omega) = \sum_{n=1}^{\infty} I_{\{X_n=i, T_k \geq n\}}$$

6. 定义  $\rho_i(k) = \mathbb{E}[N_{k,i} | X_0 = k]$ ，则

$$\rho_i(k) = \sum_{n=1}^{\infty} \mathbb{P}(X_n = i, T_k \geq n | X_0 = k)$$

7. 由 4，得到

$$\mu_k = \mathbb{E}[T_k | X_0 = k] = \sum_{i \in S} \mathbb{E}[N_{k,i} | X_0 = k] = \sum_{i \in S} \rho_i(k)$$

8. 行向量  $\rho(k) = (\rho_i(k), i \in S)$  就是一个平稳测度，如果  $\mu_k < +\infty$ ， $\pi = (\rho_i(k)/\mu_k, i \in S)$  就是一个平稳分布。

我们可以利用平稳分布的存在性判断不可约马氏链是否非零常返。

## 定理 2.4.7

对于不可约马氏链，其存在平稳分布的充要条件为所有状态为非零常返的。此时  $\pi_i = \mu_i^{-1}$ ，也是唯一的。

该定理提供了一个判断不可约马氏链是否（非零）常返的方法，以及通过平均返回时间求平稳分布的方法。那么，无法验证平稳分布存在性的情况下，如何判断不可约马氏链是否常返呢？

## 定理 2.4.8

对于不可约马氏链  $X$ ， $X$  瞬时  $\Leftrightarrow \exists s \in S, \exists \{y_j, j \neq s\}$  满足

$$\forall j, |y_j| \leq 1; \exists j', |y_{j'}| > 0$$

并使得

$$y_i = \sum_{j \neq s} p_{ij} y_j, i \neq s$$

定理 2.4.8 提供了不可约马氏链常返的充要条件：方程

$$y_i = \sum_{j \neq s} p_{ij} y_j, i \neq s$$

的有界解都是 0。

## 定理 2.4.9

不可约马氏链  $X$ ,  $S = \mathbb{N}$ , 若存在  $s \notin S, \{y_j, j \neq s\}$  使得

$$y_i \geq \sum_{j \neq s} p_{ij} y_j, i \neq s$$

且  $i \rightarrow +\infty$  时  $y_i \rightarrow +\infty$ , 则  $X$  常返。

最后是一个实际应用的例子。

## 例 2.4.1. 带边界的随机游走

状态空间  $S = \mathbb{N}$ , 转移矩阵形如:

$$\begin{array}{cccccc} & 0 & 1 & 2 & 3 & \dots \\ 0 & q & p & 0 & 0 & \\ 1 & q & 0 & p & 0 & \\ 2 & 0 & q & 0 & p & \\ 3 & 0 & 0 & q & 0 & \\ \vdots & & & & & \end{array}$$

即向左走概率为  $q$ , 向右走概率为  $p$ ,  $p + q = 1$ , 但是在 0 向左走会被挡住, 只能停留。证明这个马氏链是不可约的, 分析其常返/瞬时的条件。

解答: 显然每个状态之间都是相互可达的, 所以  $X$  是不可约的。解方程  $\pi = \pi \mathbf{P}$  可得  $\pi_j = r^j(1-r)$ , 其中  $r = p/q$ , 因此  $r < 1$  时,  $X$  存在平稳分布  $\pi$ , 进而是非零常返的;  $r > 1$  时, 取  $s = 0, y_j = 1 - r^{-j}$ , 可以判断  $X$  是瞬时的;  $r = 1$  时, 取  $s = 0, y_j = j$ , 由定理 2.4.9 知  $X$  是零常返的。

相关: ??

## 2.4.5 极限定理 \*

本节来自应随课程, 讨论了平稳分布与  $\lim_{n \rightarrow \infty} p_{ij}(n)$  之间的关系。

## 定理 2.4.10

马氏链  $X$  是不可约、非周期的, 则

$$\lim_{n \rightarrow \infty} p_{ij}(n) = \frac{1}{\mu_j}, \forall i, j$$

证明 如果  $X$  是瞬时的, 则由推论 2.3.4(3) 立即得证。下面只考虑  $X$  是常返的。

我们采用“耦合方法”: 取马氏链  $Y = \{Y_n, n \in \mathbb{N}\}$ , 其满足  $Y$  与  $X$  独立, 并且状态空间、转移矩阵和  $X$  都相同, 取  $Z = \{Z_n = (X_n, Y_n)\}$ , 为一个新的马氏链, 称为  $X$  的耦合链。很容易得到其转移概率为:

$$p_{(i,j),(k,l)} = \mathbb{P}(X_1 = k, Y_1 = l | X_0 = i, Y_0 = j) = \mathbb{P}(X_1 = k | X_0 = i) \mathbb{P}(Y_1 = l, Y_0 = j) = p_{ik} p_{jl}$$

因为  $X$  不可约、非周期, 故  $\forall i, j, k, l$ , 在  $n$  充分大时,  $\forall n$  有  $p_{ik}(n), p_{jl}(n) > 0$ , 这说明  $Z$  也是不可约、非周期的。

先考虑  $X$  非零常返, 此时  $X$  存在平稳分布  $\pi$ , 那么

$$v = (v_{ij} = \pi_i \pi_j, (i, j) \in S \times S)$$

就是  $Z$  的平稳分布, 从而  $Z$  也是非零常返的。

$$\lim_{n \rightarrow \infty} |p_{ik}(n) - p_{ik}(n)| = 0, \forall i, j, k$$

Claim 1

现在, 考虑  $\pi \mathbf{P}^n = \mathbf{P}^n$ , 可得

$$\pi_k = \sum_{i \in S} \pi_i p_{ik}(n)$$

于是

$$\pi_k - p_{jk}(n) = \sum_{i \in S} \pi_i p_{ik}(n) - \sum_{i \in S} \pi_i p_{jk}(n) = \sum_{i \in S} \pi_i (p_{ik}(n) - p_{jk}(n)) = (\star 1)$$

任取  $S$  的有限子集  $F$ , 因为  $\forall \pi_i, p_{ij}(n) \in [0, 1]$ , 所以

$$\begin{aligned} (\star 1) &= \sum_{i \in F} \pi_i (p_{ik}(n) - p_{jk}(n)) + \sum_{i \notin F} \pi_i (p_{ik}(n) - p_{jk}(n)) \\ &\leq \sum_{i \in F} (p_{ik}(n) - p_{jk}(n)) + \sum_{i \notin F} \pi_i \end{aligned}$$

根据 Claim 1, 有限和是趋于 0 的, 再令  $F \nearrow S$ , 即可得  $\pi_k - p_{jk}(n) \rightarrow 0$ .

再考虑  $X$  零常返, 此时只需证明  $p_{ij}(n) \rightarrow 0$  即可. 仍然考虑耦合链  $Z$ ,

1. 如果  $Z$  瞬时, 有推论 2.3.4 即可得证.
2. 如果  $Z$  非零常返, 考虑首达时:

$$T_{ii}^Z = \min\{n \geq 1 : Z_n = (i, i)\}$$

$$T_i = \min\{n \geq 1 : X_n = i\}$$

$Z$  非零常返而  $X$  零常返, 故

$$\mathbb{E}[T_{ii}^Z | Z_0 = (i, i)] < +\infty, \mathbb{E}[T_i | Z_0 = (i, i)] = \mathbb{E}[T_i | X_0 = i] = +\infty$$

但是  $T_i \leq T_{ii}^Z$  a.s., 导致矛盾.

3. 如果  $Z$  零常返, Claim 1 同样成立 (下文中 Claim 1 的证明中只用到了  $Z$  常返的条件, 与是否“零”常返无关). 假设 “ $\lim_{n \rightarrow \infty} p_{ij}(n) = 0, \forall i, j$ ” 不成立, 则

$$\forall i, j, \exists \{\alpha_j\}_{j=1}^{\infty}, \alpha_j \text{ s.t. } \lim_{r \rightarrow \infty} p_{ij}(n_r) = \alpha_j \quad \text{Claim 3}$$

其中  $\alpha_j$  不全为零、与  $i$  的选取无关,  $n_r$  与  $i, j$  的选取无关. 于是令行向量  $\alpha = (\alpha_j, j \in S)$ , 那么任取有限集  $F \subset S$ ,

$$\sum_{j \in F} \alpha_j = \lim_{r \rightarrow \infty} \sum_{j \in F} p_{ij}(n_r) \leq 1$$

再令  $F \nearrow S$  即可得  $0 < \sum \alpha_j \leq 1$ .

$$\sum_{k \in S} \alpha_k p_{kj} = \alpha_j, \forall j \in S \quad \text{Claim 4}$$

这说明我们找到了  $X$  的平稳分布, 这与  $X$  零常返矛盾.

Claim 1 的证明: 假设  $X_0 = i, Y_0 = j, Z_0 = (i, j)$ , 选定  $s \in S$ , 设

$$T = \min\{n \geq 1 : Z_n = (s, s)\}$$

由于  $Z$  不可约、常返, 根据题目 2.20,  $\mathbb{P}(T < +\infty) = f_{(i,j),(s,s)} = 1$ , 于是  $\mathbb{P}(T > n) \rightarrow 0$ .

$$\mathbb{P}(X_n \leq x | T < n) = \mathbb{P}(Y_n \leq x | T < n), \forall x \quad \text{Claim 2}$$

现在,

$$\begin{aligned} p_{ik}(n) &= \mathbb{P}(X_n = k) \\ &= \mathbb{P}(X_n = k, T < n) + \mathbb{P}(X_n = k, T \geq n) \\ &= \mathbb{P}(Y_n = k, T < n) + \mathbb{P}(X_n = k, T \geq n) \\ &\leq \mathbb{P}(Y_n = k) + \mathbb{P}(T \geq n) = p_{jk}(n) + \mathbb{P}(T \geq n) \end{aligned}$$

交换  $i, j$  位置可得  $|p_{ik}(n) - p_{jk}(n)| \leq \mathbb{P}(T \geq n) \rightarrow 0$ .

Claim 2 的证明：实际上就是“给定  $T \leq n$ ”条件下  $X_{n+1}$  和  $Y_{n+1}$  分布相同。老师是用初等方法证明的，太复杂所以我不想抄了。不过用测度论的语言就很好理解。回顾定义 2.1.4，我们取滤流

$$\mathcal{F}_n = \sigma(X_0, \dots, X_n, Y_0, \dots, Y_n)$$

就有  $X_n, Y_n \in \mathcal{F}_n$ ，所以

$$\mathbb{P}(X_{n+1} \in B | \mathcal{F}_n) = p(X_n, B)$$

当然对  $Y$  也成立。同时，我们也能得到多步转移概率：

$$\mathbb{P}(X_{n+1} \in B | \mathcal{F}_m) = p_{n+1-m}(X_m, B), \forall m \leq n$$

这说明给定  $m$  时刻之前的信息， $X_{n+1}$  的分布只取决于  $X_m$ 。  $T$  是  $Z$  的停时，同时

$$\{T \leq n\} = \bigcup_{m=1}^n \{T = m\} = \bigcup_{m=1}^n \underbrace{\{(X_1, Y_1) \neq (i, i), \dots, (X_{m-1}, Y_{m-1}) \neq (i, i), X_m = Y_m = i\}}_{\subset \mathcal{F}_m}$$

所以

$$\begin{aligned} \mathbb{P}(X_{n+1} \in B, T \leq n) &= \sum_{m=1}^n \mathbb{P}(X_{n+1} \in B, (X_1, Y_1) \neq (i, i), \dots, (X_{m-1}, Y_{m-1}) \neq (i, i), X_m = Y_m = i) \\ &= \sum_{m=1}^n \mathbb{P}(X_{n+1} \in B | (X_1, Y_1) \neq (i, i), \dots, (X_{m-1}, Y_{m-1}) \neq (i, i), X_m = Y_m = i) \cdot \mathbb{P}(T = m) \\ &= \sum_{m=1}^n p_{n+1-m}(i, B) \cdot \mathbb{P}(T = m) \end{aligned}$$

最后式子只与  $B$  有关，考虑  $Y$  则会得到相同的结果，所以二者同分布。

Claim 3 的证明：

(2). 不妨设  $S \times S = \{s_k, k \geq 1\}$ . 先证引理：对  $\{x_i(n), i \geq 1\}$  这样一组有界实数， $\exists$  序列  $\{n_r\}_{r \geq 1}, \{y_i\}_{i \geq 1}$ ，使得  $\lim_{r \rightarrow \infty} n_r = \infty, \lim_{r \rightarrow \infty} x_i(n_r) = y_i$ .

首先，对  $\{x_1(n)\}_{n \geq 1}$ ，存在子列  $\{n_k^{(1)}\}_{k \geq 1}$  s.t.  $\lim_{k \rightarrow \infty} x_1(n_k^{(1)}) = y_1$ ；对  $\{x_2(n_k^{(1)})\}_{k \geq 1}$ ，存在  $\{n_k^{(2)}\}_{k \geq 1}$ ，s.t.  $\lim_{k \rightarrow \infty} x_2(n_k^{(2)}) = y_2$ ；重复上述步骤，可得序列族  $\{n_k^{(m)}\}_{k \geq 1}, m \geq 1$ ，满足：①  $\{n_k^{(m)}\}$  为  $\{n_k^{(m+1)}\}$  的子列；②  $\lim_{k \rightarrow \infty} n_k^{(m)} = \infty, \forall m \geq 1$ .

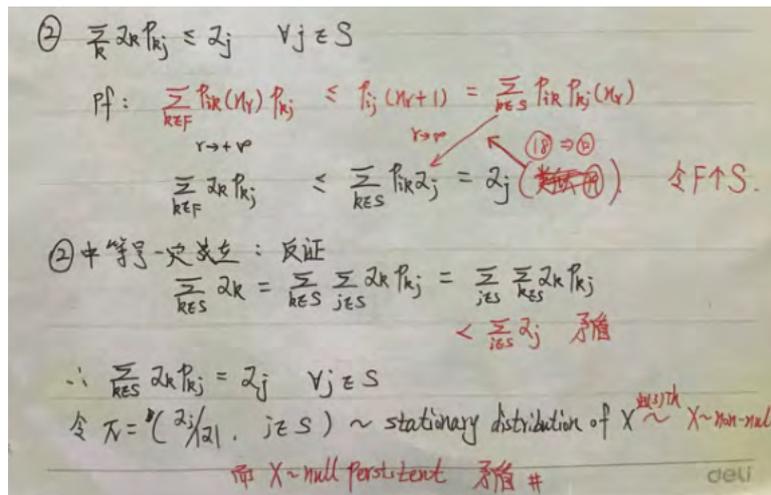
下令  $n_r = n_r^{(m)}$ . 由于  $r \geq m$  时， $\{n_r\}_{r \geq m}$  为  $\{n_k^{(m)}\}_{k \geq 1}$  的子列，则  $\lim_{r \rightarrow \infty} x_m(n_r) = y_m, \forall m \geq 1$ . 引理证毕.

由引理，对  $\{p_{ij}(n), (i,j) \in S \times S, n \geq 1\}$ ， $\exists \{n_r\}_{r \geq 1}, \{\alpha_{ij}\}_{(i,j) \in S \times S}$  使得  $\lim_{r \rightarrow \infty} p_{ij}(n_r) = \alpha_{ij}$ .

下面说明： $\alpha_{ij} = a_{ij}, \forall i, j \in S$ ，且  $\alpha_{ij}$  不全为 0；（因此，令  $\alpha_j = \alpha_{ij}$ ，某  $i \in S$  即可）

- 由于  $Z$  常返，可从证明  $|p_{ij}(n) - p_{ij}(n_r)| \rightarrow 0 (n \rightarrow \infty, \forall i, j \in S)$ ，从而  $\alpha_{ij} = a_{ij}$ ；
- $\alpha_{ij}$  不全为 0 依赖于子列选取：因为  $p_{ij}(n) \rightarrow 0, n \rightarrow \infty, \forall i, j \in S$  不成立，故  $\exists \hat{i}, \hat{j} \in S$  及  $\varepsilon > 0$ ，及子列  $\{\hat{n}_k\}_{k \geq 1}$ ，使得  $p_{\hat{i}\hat{j}}(\hat{n}_k) \geq \varepsilon$ . 此时选取  $s_i = (\hat{i}, \hat{j})$ （相当于首先选取  $p_{\hat{i}\hat{j}}(\hat{n}_k)$  的收敛子列），就可以保证构造出的  $n_r$  满足  $\lim_{r \rightarrow \infty} p_{\hat{i}\hat{j}}(n_r) = \alpha_{\hat{i}\hat{j}} > 0$ . 从而  $\alpha_{ij}$  不全为 0；

Claim 4 的证明：



□

这个定理表明，如果  $X$  是不可约、非零常返、非周期的，那  $p_{ij}(n) \rightarrow \pi_j$ ，这说明  $\lim_{n \rightarrow \infty} p_{ij}(n)$  与初始状态  $X_0$  的分布无关，即

$$\mathbb{P}(X_n = j) = \sum_{i \in S} \mathbb{P}(X_0 = i) \cdot p_{ij}(n) \rightarrow \frac{1}{\mu_j}$$

其现实意义是：即使不知道初始状态，经过足够长的时间之后，我们能够近似地预测此时落在某个的状态的概率。或者说，有多个独立样本参与了随机过程，经过足够长的时间之后，会发现每个状态处的样本数量趋近于稳定，这便是平稳分布。

即使不给定不可约的条件，也有相关的结论。

**定理 2.4.11**

对于马氏链  $X$ ，其任意非周期状态  $j$  都有

$$\lim_{n \rightarrow \infty} p_{jj}(n) = \frac{1}{\mu_j}$$

$$\lim_{n \rightarrow \infty} \frac{f_{ij}}{\mu_j}, i \neq j$$

**推论 2.4.1**

令

$$\tau_{ij}(n) = \frac{1}{n} \sum_{m=1}^n p_{ij}(m)$$

若  $j$  非周期，则

$$\lim_{n \rightarrow \infty} \tau_{ij}(n) = \frac{f_{ij}}{\mu_j}$$



了解即可，不要求掌握证明。

## 2.5 终止分布问题 (Exit Distributions)\*

本小节内, 对于状态空间  $S$  的子集  $A$ , 我们记

$$V_A = \inf\{n \geq 0 : X_n \in A\}, T_A = \inf\{n \geq 1 : X_n \in A\}$$

本节讨论的是系统离开/进入某些状态时的分布。

### 定理 2.5.1

马氏链  $X = \{X_n, n \geq 0\}$  状态空间为  $S$ ,  $A, B \subset S$  满足  $C = S - (A \cup B)$  有限, 如果:

(1).  $h(A) = 1, h(B) = 0$ .

(2). 对于  $\forall x \in C$  满足

$$h(x) = \sum_{y \in S} p_{xy} h(y)$$

(3).  $\mathbb{P}(V_A \wedge V_B < +\infty | X_0 = x) > 0, \forall x \in C$

则  $h(x) = \mathbb{P}(V_A \wedge V_B < +\infty | X_0 = x)$ .

这个定理用来计算“从  $x$  出发, 进入  $B$  之前, 能够进入  $A$  的概率”。

### 例 2.5.1.

甲有  $x$  元, 乙有  $N - x$  元, 两人玩公平的游戏, 每轮游戏的败者需给胜者 1 元, 直到某人输光为止。求甲最终能够获胜 (即乙输光) 的概率。

解答: 马氏链  $X = \{X_n, n \geq 0\}$ ,  $X_n$  代表了玩了  $n$  轮游戏之后, 甲拥有的钱数。状态空间  $S = \{0, 1, \dots, N\}$ ,  $A = \{N\}$ ,  $B = \{0\}$ , 此题目即求

$$h(x) = \mathbb{P}(V_A < V_B | X_0 = x)$$

那么可知,  $h(x)$  需满足:

(1).  $h(0) = 0$ .

(2).  $h(N) = 1$ .

(3).  $h(x) = \frac{1}{2}h(x-1) + \frac{1}{2}h(x+1), x \neq 0, N$ .

可解得

$$h(x) = \frac{x}{N}$$

### 例 2.5.2.

马氏链  $X = \{X_n, n \geq 0\}$ , 状态空间  $S = \{1, 2, \dots, 7\}$ , 转移矩阵为

	1	2	3	4	5	6	7
1	0.7				0.3		
2	0.1	0.2	0.3	0.4			
3			0.5	0.3	0.2		
4				0.5		0.5	
5	0.6				0.4		
6						0.2	.8
7				1			

分析每个状态 (瞬时/正常返/零常返、周期、平均返回时间), 并对每个  $i, j \in S$  求出

$$\lim_{n \rightarrow \infty} p_{ij}(n)$$

解答: 首先对马氏链进行划分, 分析状态之间的可达性, 可知  $T = \{2, 3\}$  为瞬时态,  $A = \{1, 5\}$  和  $B = \{4, 6, 7\}$  为两条常返、不可约、闭的子链; 两条子链的状态空间  $A, B$  都有限, 所以正常返; 状态 2, 3, 4, 5, 6 可以停留在自身, 所以周期为 1, 再根据 1, 5 和 6, 7 相互可达可知所有的状态周期都为 1.

然后我们来求平均返回时间, 注意到 4, 5 瞬时所以  $\mu_4 = \mu_5 = +\infty$ , 子链  $A$  的平稳分布  $(\pi_1, \pi_5)$  需满足:

$$(\pi_1, \pi_5) \begin{pmatrix} 0.7 & 0.3 \\ 0.6 & 0.4 \end{pmatrix} = (\pi_1, \pi_5), \pi_1 + \pi_5 = 1$$

可解得  $\pi_1 = \frac{1}{3}, \pi_5 = \frac{2}{3}$ , 同理可解得子链  $B$  的平稳分布:  $\pi_4 = \frac{8}{17}, \pi_6 = \frac{5}{17}, \pi_7 = \frac{4}{17}$ . 取倒数就得到平均返回时间.

最后我们来分析  $p_{ij}(n)$  的极限. 注意到首先  $i$  不可达  $j$  时极限为 0,  $i, j$  属于同一条子链时极限为  $\pi_j$ ,  $i, j$  都瞬时时极限为 0, 所以目前我们可以得到的是:

$$P^\infty = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{matrix} \pi_1 & 0 & 0 & 0 & \pi_5 & 0 & 0 \\ ? & 0 & 0 & ? & ? & ? & ? \\ ? & 0 & 0 & ? & ? & ? & ? \\ 0 & 0 & 0 & \pi_4 & 0 & \pi_6 & \pi_7 \\ \pi_1 & 0 & 0 & 0 & \pi_5 & 0 & 0 \\ 0 & 0 & 0 & \pi_4 & 0 & \pi_6 & \pi_7 \\ 0 & 0 & 0 & \pi_4 & 0 & \pi_6 & \pi_7 \end{matrix} \end{matrix}$$

我们先来分析  $\lim_{n \rightarrow \infty} p_{21}(n)$ , 它代表的含义是: “从状态 2 出发, 在进入  $B$  之前先进入了子链  $A$ ”, “然后停留在  $A$  中的状态 1”. 前者为  $\mathbb{P}(V_A < V_B | X_0 = 2)$ , 后者为  $\pi_1$ , 两者相乘即得到答案. 其余 ? 处的值的计算同理:

$$\lim_{n \rightarrow \infty} p_{ij}(n) = \mathbb{P}(V_A < V_B | X_0 = i) \cdot \pi_j, i \in T, j \in A$$

相似题目: 题目 2.21.

**定理 2.5.2**

设  $C = S - A$  有限,

- (1).  $g(A) = 0$ .
- (2).  $\forall x \in C$ , 有

$$g(x) = 1 + \sum_{y \in S} p_{xy}g(y)$$

- (3).  $\mathbb{P}(V_A < +\infty | X_0 = x) > 0, \forall x \in C$ .

则  $g(x) = \mathbb{E}[V_A | X_0 = x]$ .

证明 见题目 2.10. □

这个定理用来计算: 从  $x$  出发, 直到首次进入  $A$  的时间的期望.

**题目 2.2. 应随第七次作业 1.**

平均需要抛多少次硬币, 才能连续抛出两次正面?

解答: 记  $Y_n$  为第  $n$  抛硬币的正反, 正记作 1, 反记作 0, 令  $X_n = (Y_n, Y_{n+1}), n \geq 1$ , 规定  $X_0 = 0$ , 则  $X = \{X_n, n \geq 0\}$  是一个马氏链, 状态空间为:  $S = \{0, (1, 1), (1, 0), (0, 1), (0, 0)\}$ , 分别记作  $S = \{0, a, b, c, d\}$ .

于是, 考虑集合  $A = \{a\}$ ,  $g(x) = \mathbb{E}[V_A | X_0 = x]$  需满足

- (1).  $g(a) = 0$ .  
(2).

$$g(0) = 1 + \frac{1}{4}(g(a) + g(b) + g(c) + g(d))$$

$$g(b) = 1 + \frac{1}{2}g(c) + \frac{1}{2}g(d)$$

$$g(c) = 1 + \frac{1}{2}g(a) + \frac{1}{2}g(b)$$

$$g(d) = 1 + \frac{1}{2}g(c) + \frac{1}{2}g(d)$$

可解得  $g(0) = 5$ , 注意  $X_5 = (Y_5, Y_6)$ , 所以题目的最终答案应当为 6.

### 题目 2.3. 应随第七次作业 2.

平均需要抛多少次骰子, 才能每个面都出现至少一次?

解答: 思路:  $\xi_n, n \geq 1$  代表每次抛出的点数, 设  $X = \{X_n, n \geq 0\}$ , 其中  $X_n = |\{\xi_1, \dots, \xi_n\}|$ , 代表第  $n$  次抛出后一共出现几种不同的点数, 并规定  $X_0 = 0$ , 根据题目 2.5, 这是一个马氏链。题目要求的就是  $\mathbb{E}_0[V_6]$ , 令

$$g(x) = \mathbb{E}_x[V_6]$$

然后列方程解出  $g(0)$  即可, 最终答案为  $147/10$ .

### 例 2.5.3. 应随 2020 期中 3.

车站的每两趟车之间的时间间隔等概率地为 10、20、30 分钟, 若有一趟车正好在整点时刻到达, 那么平均需要等多久, 会出现下一趟整点时刻到达的车?

解答: 思路: 整点、10、20、30、40、50 分别记作 0, 1, 2, 3, 4, 5, 记  $X_n$  为第  $n$  趟车的到达时刻, 则其状态空间就是  $S = \{0, 1, 2, 3, 4, 5\}$ , 转移矩阵为:

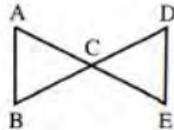
$$\begin{array}{cccccc} & 0 & 1 & 2 & 3 & 4 & 5 \\ \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{array}{c} 0 \\ 0 \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{array} & \begin{array}{c} \frac{1}{3} \\ 0 \\ 0 \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \end{array} & \begin{array}{c} \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ 0 \\ 0 \\ 0 \end{array} & \begin{array}{c} 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ 0 \\ 0 \end{array} & \begin{array}{c} 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \\ 0 \end{array} & \begin{array}{c} 0 \\ 0 \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{array} \end{array}$$

题目要求的就是  $\mathbb{E}_0[T_0]$ , 即 0 的平均返回时间。不难看出平稳分布是  $\frac{1}{6}(1, 1, 1, 1, 1, 1)$ , 所以答案就是 6.

接下来这道习题的后两问更是重量级。本节虽然只介绍了两个定理和其应用, 但你不能只会这两个定理, 比如这道题就需要一些用到马氏性的变换, 生搬硬套就做不了。希望读者能借此搞清楚这类题目的通用计算方法。

## 题目 2.4. 第 7 次作业 3

3.  
**11. Bow ties.** A particle performs a random walk on a bow tie ABCDE drawn beneath on the left, where C is the knot. From any vertex its next step is equally likely to be to any neighbouring vertex. Initially it is at A. Find the expected value of:
- the time of first return to A,
  - the number of visits to D before returning to A,
  - the number of visits to C before returning to A,
  - the time of first return to A, given no prior visit by the particle to E,
  - the number of visits to D before returning to A, given no prior visit by the particle to E.



解答:

(a). 即求 A 的平均返回时间  $\mu_A$ , 不难解得平稳分布是:

$$\pi = (\pi_A, \pi_B, \pi_C, \pi_D, \pi_E) = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{3}, \frac{1}{6}, \frac{1}{6}\right)$$

所以  $\mu_A = \pi_A^{-1} = 6$ .

(b). 即求 A 返回之前经过其他状态的平均次数, 这个是在我们构造平稳测度的时候用到的玩意儿:

$$\rho_A(x) = \mathbb{E}_A \left[ \sum_{n=0}^{T_A-1} I_{\{X_n=x\}} \right]$$

注意  $\rho_A(A) = 1$ , 所以以 A 构造的平稳测度为

$$\rho_A = (1, 1, 2, 1, 1)$$

所求即为  $\rho_A(D) = 1$ .

(c). 即  $\rho_A(C) = 2$ .

(d). 如果从未经过 E, 返回 A 所需的平均时间。即

$$\mathbb{E}_A[T_A | T_A < T_E]$$

Target of (d)

作为准备工作, 首先计算  $g(x) = \mathbb{P}_x[V_A | V_A < V_E]$ , 得到

$$(g(x), x \in S) = \left(1, \frac{3}{4}, \frac{1}{2}, \frac{1}{4}, 0\right)$$

进一步计算:

$$\begin{aligned} \mathbb{P}_A(T_A < T_E) &= p(A, B)\mathbb{P}_B(T_A < T_E) + p(A, C)\mathbb{P}_C(T_A < T_E) \\ &= p(A, B)g(B) + p(A, C)g(C) = \frac{5}{8} \end{aligned}$$

接下来分析  $f(x) = \mathbb{E}_x[V_A | V_A < V_E]$ , 对于  $x \neq A, E$ ,

$$\begin{aligned} f(x) &= \mathbb{E}_x[V_A | V_A < V_E] = \frac{\mathbb{E}_x[V_A \cdot I_{\{V_A < V_E\}}]}{\mathbb{P}_x(V_A < V_E)} \\ &= \frac{\mathbb{E}_x[(V_A \circ \theta_1 + 1) \cdot I_{\{V_A < V_E\}}]}{\mathbb{P}_x(V_A < V_E)} \\ &= \frac{\mathbb{E}_x[(V_A \circ \theta_1) \cdot I_{\{V_A < V_E\}}]}{\mathbb{P}_x(V_A < V_E)} + 1 \\ &= \frac{\mathbb{E}_x[(V_A \cdot I_{\{V_A < V_E\}}) \circ \theta_1]}{\mathbb{P}_x(V_A < V_E)} + 1 \end{aligned}$$

$$\begin{aligned}
&= \frac{\mathbb{E}_x[\mathbb{E}_{X_1}[V_A \cdot I_{\{V_A < V_E\}}]]}{\mathbb{P}_x(V_A < V_E)} + 1 \\
&= \frac{1}{\mathbb{P}_x(V_A < V_E)} \sum_{y \in S} p(x, y) \mathbb{E}_y[V_A \cdot I_{\{V_A < V_E\}}]
\end{aligned}$$

注意给定初值  $y = E$  时  $I_{\{V_A < V_E\}} = 0$ , 初值  $y = A$  时  $V_A = 0$ , 所以

$$\begin{aligned}
f(x) &= \frac{1}{\mathbb{P}_x(V_A < V_E)} \sum_{y \in \{B, C, D\}} p(x, y) \mathbb{E}_y[V_A \cdot I_{\{V_A < V_E\}}] + 1 \\
&= \frac{1}{\mathbb{P}_x(V_A < V_E)} \sum_{y \in \{B, C, D\}} p(x, y) \cdot \mathbb{P}_y(V_A < V_E) \cdot \frac{\mathbb{E}_y[V_A \cdot I_{\{V_A < V_E\}}]}{\mathbb{P}_y(V_A < V_E)} + 1 \\
&= \frac{1}{\mathbb{P}_x(V_A < V_E)} \sum_{y \in \{B, C, D\}} p(x, y) \cdot \mathbb{P}_y(V_A < V_E) \cdot \mathbb{E}_y[V_A | V_A < V_E] + 1 \\
&= \frac{1}{g(x)} \sum_{y \in \{B, C, D\}} p(x, y) \cdot g(y) \cdot f(y) + 1
\end{aligned}$$

这个方程过于美观以至于我想单独列一个式子出来:

$$f(x) = 1 + \frac{1}{g(x)} \sum_{y \in \{B, C, D\}} p(x, y) \cdot g(y) \cdot f(y) \quad (\star)$$

代入  $x = B, C, D$ , 得到三个方程:

$$\begin{aligned}
f(B) &= 1 + \frac{1}{3}f(C) \\
f(C) &= 1 + \frac{3}{8}f(B) + \frac{1}{8}f(D) \\
f(D) &= 1 + f(C)
\end{aligned}$$

解得

$$f(B) = \frac{5}{3}, f(C) = 2, f(D) = 3$$

回头看我们的目标, 和  $(\star)$  的推导过程类似, 我们将其转化为:

$$\begin{aligned}
\mathbb{E}_A[T_A | T_A < T_E] &= \frac{\mathbb{E}_A[T_A \cdot I_{\{T_A < T_E\}}]}{\mathbb{P}_A(T_A < T_E)} \\
&= \frac{\mathbb{E}_A[(T_A \circ \theta_1 + 1) \cdot I_{\{T_A < T_E\}}]}{\mathbb{P}_A(T_A < T_E)} \\
&= 1 + \frac{\mathbb{E}_A[(T_A \circ \theta_1) \cdot I_{\{T_A < T_E\}}]}{\mathbb{P}_A(T_A < T_E)} \\
&= 1 + \frac{\mathbb{E}_A[(T_A \cdot I_{\{T_A < T_E\}}) \circ \theta_1]}{\mathbb{P}_A(T_A < T_E)} \\
&= 1 + \frac{\mathbb{E}_A[\mathbb{E}_{X_1}[T_A \cdot I_{\{T_A < T_E\}}]]}{\mathbb{P}_A(T_A < T_E)} \\
&= 1 + \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in S} p(A, y) \mathbb{E}_y[T_A \cdot I_{\{T_A < T_E\}}] \\
&= 1 + \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in \{B, C\}} p(A, y) \mathbb{E}_y[T_A \cdot I_{\{T_A < T_E\}}] \\
&= 1 + \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in \{B, C\}} p(A, y) \mathbb{E}_y[V_A \cdot I_{\{V_A < V_E\}}] \\
&= 1 + \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in \{B, C\}} p(A, y) f(y) g(y) = \frac{14}{5}
\end{aligned}$$

(e). 如果从未经过  $E$ , 返回  $A$  之前经过状态  $D$  的平均次数。

$$\mathbb{E}_A \left[ \sum_{n=0}^{T_A-1} I_{\{X_n=D\}} \mid T_A < T_E \right] \quad \text{Target of (e)}$$

我们先研究

$$h(x) = \mathbb{E}_x \left[ \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} \mid V_A < V_E \right], x \neq E$$

并注意到  $h(A) = 0$ , 为了简化表达, 令

$$Y = \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} I_{\{V_A < V_E\}}$$

于是

$$h(x) = \frac{\mathbb{E}_x[Y]}{g(x)}$$

当  $x \in \{B, C\}$  时,

$$\begin{aligned} \mathbb{E}_x[Y] &= \mathbb{E}_x \left[ \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} I_{\{V_A < V_E\}} \right] \\ &= \mathbb{E}_x \left[ \sum_{n=1}^{V_A} I_{\{X_n=D\}} I_{\{V_A < V_E\}} \right] \\ &= \mathbb{E}_x \left[ \left( \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} \right) \circ \theta_1 \cdot I_{\{V_A < V_E\}} \right] \\ &= \mathbb{E}_x \left[ \left( \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} \cdot I_{\{V_A < V_E\}} \right) \circ \theta_1 \right] \\ &= \mathbb{E}_x \left[ \mathbb{E}_{X_1} \left[ \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} \cdot I_{\{V_A < V_E\}} \right] \right] \\ &= \mathbb{E}_x [\mathbb{E}_{X_1}[Y]] \\ &= \sum_{y \in S} p(x, y) \mathbb{E}_y[Y] = \sum_{y \in \{B, C, D\}} p(x, y) g(y) h(y) \end{aligned} \quad \text{注意 } \mathbb{E}_E[Y] = 0$$

从而可得

$$h(x) = \frac{1}{g(x)} \sum_{y \in \{B, C, D\}} p(x, y) g(y) h(y), x \in \{B, C\} \quad (\star 1)$$

而当  $x = D$  时, 情况稍有不同:

$$\begin{aligned} \mathbb{E}_D[Y] &= \mathbb{E}_D \left[ \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} I_{\{V_A < V_E\}} \right] \\ &= \mathbb{E}_D \left[ \left( \sum_{n=1}^{V_A} I_{\{X_n=D\}} \right) I_{\{V_A < V_E\}} \right] \\ &= \mathbb{E}_D[Y \circ \theta_1] + g(D) \end{aligned}$$

从而

$$h(D) = 1 + \frac{1}{g(D)} \sum_{y \in \{B, C, D\}} p(D, y) g(y) h(y) \quad (\star 2)$$

联立  $(\star 1)(\star 2)$  可解得

$$h(B) = \frac{1}{18}, h(C) = \frac{1}{6}, h(D) = \frac{7}{6}$$

那么回到我们的目标, 进行类似的变换:

$$\begin{aligned}
 \mathbb{E}_A \left[ \sum_{n=0}^{T_A-1} I_{\{X_n=D\}} \middle| T_A < T_E \right] &= \frac{1}{\mathbb{P}_A(T_A < T_E)} \mathbb{E}_A \left[ \sum_{n=0}^{T_A-1} I_{\{X_n=D\}} I_{\{T_A < T_E\}} \right] \\
 &= \frac{1}{\mathbb{P}_A(T_A < T_E)} \mathbb{E}_A \left[ \mathbb{E}_{X_1} \left[ \sum_{n=0}^{T_A-1} I_{\{X_n=D\}} I_{\{T_A < T_E\}} \right] \right] \\
 &= \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in \{B,C\}} p(A, y) \mathbb{E}_y \left[ \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} I_{\{V_A < T_E\}} \right] \\
 &= \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in \{B,C\}} p(A, y) \mathbb{E}_y \left[ \sum_{n=0}^{V_A-1} I_{\{X_n=D\}} I_{\{V_A < V_E\}} \right] \\
 &= \frac{1}{\mathbb{P}_A(T_A < T_E)} \sum_{y \in \{B,C\}} p(A, y) g(y) h(y) \\
 &= \frac{8}{5} \left( \frac{3}{8} h(B) + \frac{1}{4} h(C) \right) = \frac{1}{10}
 \end{aligned}$$

## 2.6 习题

### 2.6.1 研随第四次作业

#### 题目 2.5. Durrett(Exercise 5.1.1)

$\xi_1, \xi_2, \dots$  独立同分布, 且在  $\{1, 2, \dots, N\}$  上均匀取值, 证明  $X_n = |\{\xi_1, \dots, \xi_n\}|$  是马氏链, 计算其一步转移概率。

#### 题目 2.6. Durrett(Exercise 5.1.3)

$\xi_1, \xi_2, \dots$  独立同分布, 且在  $\{H, T\}$  上取值, 每个取值概率均为  $1/2$ , 证明  $X_n = (\xi_n, \xi_{n+1})$  为马氏链, 计算其一步和两步转移概率  $p, p^2$ 。

这两题没啥意思, 就跳过了。

补充一个结论: 对于马氏链  $X_n$ , 如果

$$\mathbb{P}\left(\bigcup_{m=n+1}^{\infty} \{X_m \in B_m\} | X_n\right) \geq \delta > 0 \text{ on } \{X_n \in A_n\}$$

则

$$\mathbb{P}(\{X_n \in A_n \text{ i.o.}\} - \{X_n \in B_n \text{ i.o.}\}) = 0$$

然后利用这个结论解决下面这道题。

#### 题目 2.7. Durrett(Exercise 5.2.3)

状态  $a$  若满足  $\mathbb{P}_a(X_1 = a) = 1$ , 称之为吸收态, 令

$$D = \{\omega : \exists n \geq 1, X_n(\omega) = a\}$$

令  $h(x) = \mathbb{P}_x(D)$ , 证明: 在  $D^c$  上  $h(X_n) \rightarrow 0$  a.s.

解答: 需要用到 Levy 0-1 律。

#### 题目 2.8. Durrett(Exercise 5.2.5)

证明:

$$\sum_{m=0}^n \mathbb{P}_x(X_m = x) \geq \sum_{m=k}^{n+k} \mathbb{P}_x(X_m = x)$$

解答: 令  $T^k = \inf\{i \geq k : X_i = x\}$ , 那么

$$\begin{aligned} \mathbb{P}_x(X_m = x) &= \sum_{j=k}^m \mathbb{P}(X_m = x, T^k = j | X_0 = x) \\ &= \sum_{j=k}^m \mathbb{P}(X_m = x | T^k = j, X_0 = x) \mathbb{P}(T^k = j | X_0 = x) \\ &= \sum_{j=k}^m \mathbb{P}(X_m = x | X_j = x) \mathbb{P}_x(T^k = j) \\ &= \sum_{j=k}^m \mathbb{P}_x(X_{m-j} = x) \mathbb{P}_x(T^k = j) \end{aligned}$$

两边对  $m$  求和,

$$\begin{aligned}
 \sum_{m=k}^{n+k} \mathbb{P}_x(X_m = x) &= \sum_{m=k}^{n+k} \sum_{j=k}^m \mathbb{P}_x(X_{m-j} = x) \mathbb{P}_x(T^k = j) \\
 &= \sum_{j=k}^{n+k} \sum_{m=j}^{n+k} \mathbb{P}_x(X_{m-j} = x) \mathbb{P}_x(T^k = j) \\
 &= \sum_{j=k}^{n+k} \mathbb{P}_x(T^k = j) \sum_{m=0}^{n-j+k} \mathbb{P}_x(X_m = x) \\
 &\leq \sum_{j=k}^{n+k} \mathbb{P}_x(T^k = j) \sum_{m=0}^n \mathbb{P}_x(X_m = x) \\
 &= \mathbb{P}_x(T^k \leq n+k) \sum_{m=0}^n \mathbb{P}_x(X_m = x) \\
 &\leq \sum_{m=0}^n \mathbb{P}_x(X_m = x)
 \end{aligned}$$

### 题目 2.9. Durrett(Exercise 5.2.8)

$X_n$  为马氏链,  $S = \{0, 1, \dots, N\}$ , 假设  $X_n$  是一个鞅, 令

$$V_x = \min\{n \geq 0 : X_n = x\}$$

设  $\mathbb{P}_x(V_0 \wedge V_N < +\infty) > 0$  对于  $\forall x$  成立, 证明:

$$\mathbb{P}_x(V_N < V_0) = \frac{x}{N}$$

解答: 考虑鞅  $X_{n \wedge V_0 \wedge V_N}$ , 根据其收敛性, 可得

$$\mathbb{E}_x[X_{0 \wedge V_0 \wedge V_N}] = \mathbb{E}_x[X_{+\infty \wedge V_0 \wedge V_N}]$$

即

$$x = \mathbb{E}_x[X_{V_0 \wedge V_N}] = N \cdot \mathbb{P}(V_N < V_0)$$

于是得证。

### 题目 2.10. Durrett(Exercise 5.2.11)

令  $V_A = \inf\{n \geq 0 : X_n \in A\}$ ,

(1). 设  $g(x) = \mathbb{E}_x[V_A]$ , 假设  $S - A$  是有限集, 且  $\forall x \in S - A, \mathbb{P}_x(V_A < +\infty) > 0$ , 证明:

$$g(x) = 1 + \sum_{y \in S} p(x, y)g(y), \quad \forall x \notin A \quad (*)$$

(2). 证明: 如果  $g$  满足 (\*), 则  $g(X(n \wedge V_A)) + n \wedge V_A$  是一个鞅。

(3). 证明: 如果  $g$  满足 (\*), 且对于  $\forall x \in A$  都有  $g(x) = 0$ , 则  $g(x) = \mathbb{E}_x[T_A]$ 。

解答: 这个题的结论十分有用, 可以用来计算首次达时的期望, 在应随部分有进一步介绍。

(1). 设  $x_0 \notin A$ , 给定  $X_0 = x$  的条件下,

$$\begin{aligned}
 V_A(\omega) &= \inf\{n \geq 0 : X_n(\omega) = \omega_n \in A\} \\
 &= \inf\{n \geq 1 : X_n(\omega) = \omega_n \in A\}
 \end{aligned}$$

$$\begin{aligned}
&= 1 + \inf\{n \geq 0 : X_{n+1}(\omega) = \omega_{n+1} = X_n \circ \theta_1(\omega) \in A\} \\
&= 1 + V_A \circ \theta_1(\omega)
\end{aligned}$$

于是

$$\begin{aligned}
g(x) &= \mathbb{E}_x[V_A] = \mathbb{E}_x[1 + V_A \circ \theta_1] \\
&= 1 + \mathbb{E}_x[V_A \circ \theta_1] \\
&= 1 + \mathbb{E}_x[\mathbb{E}_x[V_A \circ \theta_1 | \mathcal{F}_1]] \\
&= 1 + \mathbb{E}_x[\mathbb{E}_{X_1}[V_A]] \\
&= 1 + \mathbb{E}_x[g(X_1)] \\
&= 1 + \sum_{y \in S} \mathbb{E}_x[g(y) I_{\{X_1=y\}}] \\
&= 1 + \sum_{y \in S} p(x, y) g(y)
\end{aligned}$$

(2). 由 (1). 的证明过程, 可知 (\*) 式就是  $g(x) = 1 + \mathbb{E}_x[g(X_1)]$ ,

$$\begin{aligned}
\mathbb{E}[g(X_{(n+1) \wedge V_A}) | \mathcal{F}_n] &= \mathbb{E}[g(X_{n+1}) I_{\{V_A \geq n+1\}} | \mathcal{F}_n] + \mathbb{E}[g(X_{V_A}) I_{\{V_A \leq n\}} | \mathcal{F}_n] \\
&= \mathbb{E}[g(X_{n+1}) | \mathcal{F}_n] I_{\{V_A \geq n+1\}} + \sum_{k=0}^n \mathbb{E}[g(X_k) I_{\{V_A=k\}} | \mathcal{F}_n] \\
&= \mathbb{E}[g(X_1) \circ \theta_n | \mathcal{F}_n] I_{\{V_A \geq n+1\}} + \sum_{k=0}^n g(X_k) I_{\{V_A=k\}} \\
&= \mathbb{E}_{X_n}[g(X_1)] I_{\{V_A \geq n+1\}} + g(X_{V_A}) I_{\{V_A \leq n\}} \\
&= (g(X_n) - 1) I_{\{V_A \geq n+1\}} + g(X_{V_A}) I_{\{V_A \leq n\}} \\
&= g(X_{n \wedge V_A}) - I_{\{V_A \geq n+1\}}
\end{aligned}$$

从而

$$\begin{aligned}
\mathbb{E}[g(X_{(n+1) \wedge V_A}) - (n+1) \wedge V_A | \mathcal{F}_n] &= g(X_{n \wedge V_A}) - I_{\{V_A \geq n+1\}} - (n+1) I_{\{V_A \geq n+1\}} - V_A I_{\{V_A \leq n\}} \\
&= g(X_{n \wedge V_A}) - n I_{\{V_A \geq n+1\}} - V_A I_{\{V_A \leq n\}} \\
&= g(X_{n \wedge V_A}) - n \wedge V_A
\end{aligned}$$

(3). 注意  $X_{V_A} \in A \Rightarrow g(X_{V_A}) = 0$ , 又根据 (2) 中构造的鞅的收敛性, 可得

$$g(x) = \mathbb{E}_x[g(X_0)] = \mathbb{E}_x[g(X_{V_A}) + V_A] = \mathbb{E}_x[V_A]$$

### 题目 2.11. Durrett(Exercise 5.3.1)

假设  $y$  是常返的, 对于  $k \geq 0$ , 令  $R_k = T_y^k$ , 为第  $k$  次返回  $y$ , 对于  $k \geq 1$ , 令  $r_k = R_k - R_{k-1}$ , 为第  $k$  次返回时间。利用强马氏性证明: 在概率测度  $\mathbb{P}_y$  下, 随机向量  $v_k = (r_k, X_{R_{k-1}}, \dots, X_{R_{k-1}}), k \geq 1$  是独立同分布的。

解答: 太复杂, 偷个懒吧, 贴一个学长写的答案:

5.3.1 Proof: Denote  $V = \{(r, s_1, s_2, \dots, s_r) \mid r \geq 1, s_i \in S, 1 \leq i \leq r\}$  to be all the possible values of the vectors. Remark: 设  $S$  为 state space. 则  $V = \prod_{k=1}^{\infty} S^k$  (即  $V$  为卡氏积  $S^k$  的无交并)  
 $V$  is countable, then the  $\sigma$ -field on  $V$  is  $2^V$ .  $V_k: \Omega_0 \rightarrow V$  are random elements. 这样  $V$  上就有自然的  $\sigma$ -代数.  
 $y$  is recurrent  $\Rightarrow P_y(V_k < \infty) = 1, \forall k \geq 1$ . 题设中的  $V_k$  就是良定义的 random element.  
 $V_k = V_1 \circ \theta_{R_k}, \forall v, \tilde{v} \in V, k > m$ .  

$$P_y(V_k = v, V_m = \tilde{v}) = E_y(I_{(V_k=v)}; \{V_m = \tilde{v}\}) = E_y(E_y(I_{(V_k=v)} | \mathcal{F}_{R_k}^y); \{V_m = \tilde{v}\})$$

$$\stackrel{\text{Markov}}{=} E_y(E_y(I_{(V_1=v)} \circ \theta_{R_k} | \mathcal{F}_{R_k}^y); \{V_m = \tilde{v}\}) \quad (\{V_m = \tilde{v}\} \in \mathcal{F}_{R_k}^y)$$

$$= E_y(P_{X_{R_k}}(V_1=v); \{V_m = \tilde{v}\}) = P_y(V_1=v) P_y(V_m = \tilde{v}) \quad (\text{Note that } X_{R_k} = y)$$

So it suffices to prove that  $P_y(V_1=v) = P_y(V_k=v)$  to show that  $V_k, V_m$  are independent.  
 It's trivial to see that  

$$P_y(V_k=v) = E_y E_y(I_{(V_1=v)} \circ \theta_{R_k} | \mathcal{F}_{R_k}^y) = E_y E_{X_{R_k}}(I_{(V_1=v)}) = P_y(V_1=v)$$
 This also verifies that  $V_k$ 's are identically distributed.  
 In general,  $\{V_k\}_{k \geq 1}$  is i.i.d.

**题目 2.12. Durrett(Exercise 5.3.3)**

证明 Ehrenfest chain (例 2.1.1) 中的所有状态都是常返的。

解答: 状态空间  $S = \{0, 1, \dots, r\}$  是有限、封闭的, 易证所有状态相互可达, 所以不可约, 所以所有状态都常返。

**题目 2.13. Durrett(Exercise 5.3.7)**

如果  $f$  满足

$$f(x) \geq \sum_y p(x,y) f(y)$$

则称其为“上调和的”(superharmonic), 等价于  $f(X_n)$  为上鞅。

假设  $p$  不可约, 证明  $p$  常返当且仅当任意非负上调和函数  $f$  都是常数。

解答:

5.3.7 Proof: Provided that  $p$  is recurrent.  
 $\forall x, y \in S, P_{xy} = P_x(T_y < \infty) > 0$  due to the irreducibility of  $p$ .  
 Applying Thm 5.3.2, we conclude  $P_{xy} = 1$  because  $y$  is recurrent.  
 $f(X_{n \wedge T_y})$  is a nonnegative supermartingale and converges  $P_x$ -a.s. to  $f(X_{T_y})$   
 $\Rightarrow f(x) = E_x f(X_{0 \wedge T_y}) \geq E_x f(X_{T_y}) = f(y)$   
 Similarly, we can prove that  $f(x) \leq f(y)$ . Henceforth,  $f$  is constant.  
 If  $p$  is transient, then all states are transient considering  $y$  is irreducible.  
 $\exists x, y, z \in S, \text{ s.t. } P_x(T_y < \infty) \neq P_z(T_y < \infty)$ .  
 Keeping  $y$  fixed, assign  $f(x) = P_x(T_y < \infty)$ .  
 Then  $f$  is a nonconstant nonnegative superharmonic function.

2.6.2 研随第五次作业

题目 2.14. Durrett(Exercise 5.5.2)

设  $w_{xy} = \mathbb{P}_x(T_y < T_x)$ , 证明  $\mu_x(y) = w_{xy}/w_{yx}$ .

解答:

Another proof of Ex 5.5.2  
 Denote  $N(y) = \sum_{n=0}^{T_x-1} \mathbb{1}_{\{X_n=y\}}$ , then  $\mu_x(y) = E_x N(y) = \sum_{n=1}^{\infty} P_x(N(y) \geq n)$   
 Obviously,  $P_x(N(y) \geq 1) = P_x(T_y < T_x) = w_{xy}$   
 And  $P_x(N(y) \geq n) = E_x(\mathbb{1}_{\{T_y < T_x\}}; T_y^{n-1} < T_x) \stackrel{\text{Markov}}{=} E_x(E_{X_{T_y^{n-1}}} \mathbb{1}_{\{T_y < T_x\}}; T_y^{n-1} < T_x) = (1-w_{yx}) P(N(y) \geq n-1)$   
 $\Rightarrow P_x(N(y) \geq n) = w_{xy} (1-w_{yx})^{n-1}$   
 Sum over  $n$  to prove that  $\mu_x(y) = w_{xy}/w_{yx}$ .

题目 2.15. Durrett(Exercise 5.5.3)

证明: 如果  $p$  不可约、常返, 则

$$\mu_x(y)\mu_y(z) = \mu_x(z)$$

解答: 平稳测度相差常数倍, 所以  $\mu_x(z) = \frac{\mu_x(y)}{\mu_y(y)}\mu_y(z) = \mu_x(y)\mu_y(z)$ .

题目 2.16. Durrett(Exercise 5.5.4)

证明: 如果  $p$  不可约、正常返, 则  $E_x[T_y] < +\infty, \forall x, y$

解答:

Another proof of Ex 5.5.4  
 $p(y,x) P_x(T_y \geq n) \stackrel{\text{Markov}}{\leq} P_y(T_y \geq n+1, X_1=x) \leq P_y(T_y \geq n+1)$   
 $\Rightarrow p(y,x) E_x T_y = \sum_{n=1}^{\infty} p(y,x) P_x(T_y \geq n) \leq \sum_{n=1}^{\infty} P_y(T_y \geq n+1) \leq E_y T_y < \infty$  by the positive recurrence.  
 $p(y,x) > 0$  due to the irreducibility  
 $\Rightarrow E_x T_y < \infty$ .

题目 2.17. Durrett(Exercise 5.5.5)

证明: 如果  $p$  不可约, 且有不是平稳分布的平稳测度  $\mu$ , 即  $\sum_x \mu(x) = +\infty$ , 则  $p$  不是正常返的。

解答: 假设  $p$  正常返, 则存在平稳分布, 矛盾。

### 题目 2.18. Durrett(Exercise 5.5.6)

对于简单对称随机游走，证明：

- (1). 两次途径 0 之间，途径状态  $k$  的次数的期望为 1.
- (2). 从  $k$  出发，返回 0 前，途径状态  $k$  的次数的期望为  $2k$ .

解答：简单对称随机游走是不可约、常返的，所以存在唯一平稳测度，不难发现  $\mu(x) = 1, \forall x$  就是一个平稳测度。我们设

$$T_x = \inf\{n \geq 0 : X_n = x\}$$

$$\mu_x(y) = \mathbb{E}_x \left[ \sum_{n=0}^{T_x-1} I_{\{X_n=y\}} \right]$$

则  $\mu_x$  是一个平稳测度，注意  $\mu_x(x) = 1$ ，所以所有的  $\mu_x \equiv 1$ .

(1). 即  $\mu_0(k) = 1$ .

(2). 我们的目标是要证明：

$$\mathbb{E}_k \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] = 2k \quad (\star 1)$$

对于  $k \geq 1$ ，有<sup>5</sup>

$$\begin{aligned} 1 &= \mathbb{E}_0 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] = \mathbb{E}_0 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} I_{\{X_1=1\}} \right] && k \geq 1, \text{ 所以第一步必须向右走} \\ &= \mathbb{E}_0 \left[ \sum_{n=1}^{T_0} I_{\{X_n=k\}} I_{\{X_1=1\}} \right] \\ &= \mathbb{E}_0 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \circ \theta_1 \cdot I_{\{X_1=1\}} \right] \\ &= \mathbb{E}_0 \left[ \mathbb{E}_0 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \circ \theta_1 \middle| \mathcal{F}_1 \right] \cdot I_{\{X_1=1\}} \right] \\ &= \mathbb{E}_0 \left[ \mathbb{E}_{X_1} \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] I_{\{X_1=1\}} \right] \\ &= \mathbb{E}_0 \left[ \mathbb{E}_1 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] I_{\{X_1=1\}} \right] \\ &= p(0,1) \mathbb{E}_1 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] \\ &= \frac{1}{2} \mathbb{E}_1 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] \end{aligned}$$

所以

$$\mathbb{E}_1 \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] = 2 \quad (\star 2)$$

于是  $k=1$  时  $(\star 1)$  成立，下面假设  $k$  时  $(\star 1)$  成立，注意到

$$\mathbb{E}_k \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k\}} \right] = \mathbb{E}_{k+1} \left[ \sum_{n=0}^{T_1-1} I_{\{X_n=k+1\}} \right] \quad (\star 3)$$

<sup>5</sup>第二个等号的原因是： $X_0 = X_{T_0} = 0 \neq k$ ，所以时刻  $0-T_0-1$  中  $k$  的次数和时刻  $1-T_0$  中  $k$  的次数是相等的。

即向右平移了一格, 变为从  $k+1$  出发, 返回 1 前, 途径状态  $k+1$  的次数的期望。而

$$\begin{aligned} \mathbb{E}_{k+1} \left[ \sum_{n=T_1}^{T_0-1} I_{\{X_n=k+1\}} \right] &= \mathbb{E}_{k+1} \left[ \left( \sum_{n=0}^{T_0-1} I_{\{X_n=k+1\}} \right) \circ \theta_{T_1} \right] \\ &= \mathbb{E}_{k+1} \left[ \mathbb{E} \left[ \left( \sum_{n=0}^{T_0-1} I_{\{X_n=k+1\}} \right) \circ \theta_{T_1} \mid \mathcal{F}_{T_1} \right] \right] \\ &= \mathbb{E}_{k+1} \left[ \mathbb{E}_{X_{T_1}} \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k+1\}} \right] \right] \\ &= \mathbb{E}_{k+1} \left[ \mathbb{E}_{X_1} \left[ \sum_{n=0}^{T_0-1} I_{\{X_n=k+1\}} \right] \right] = 2 \end{aligned} \quad (\star 4)$$

( $\star 3$ ) + ( $\star 4$ ) 即可得  $k+1$  时 ( $\star 1$ ) 成立, 归纳得证。

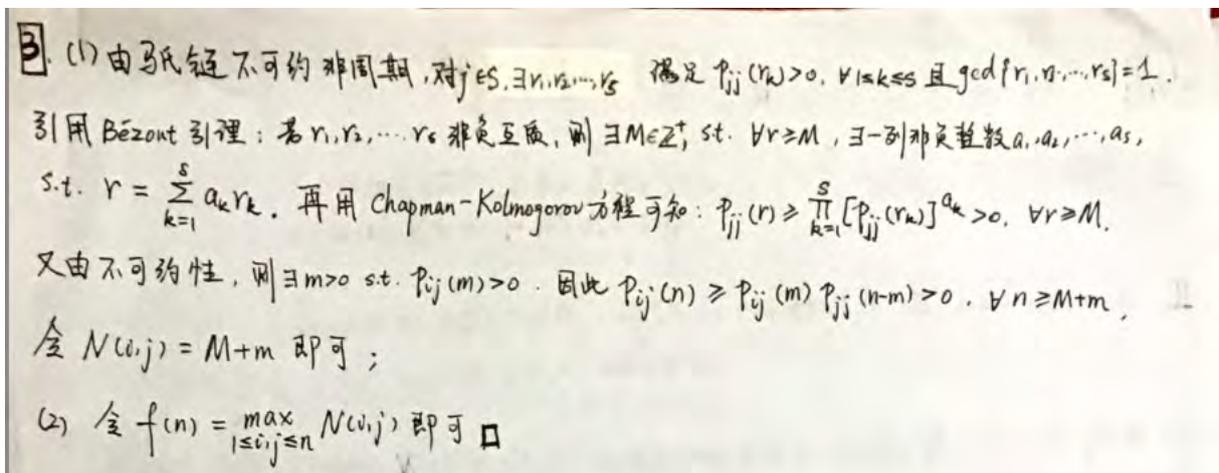
**注** 直观上看, “从  $k$  出发, 返回 0 前, 途径  $k$  的次数  $N_k$ ” 就等于 “从  $k$  出发, 返回 1 前, 途径  $k$  的次数” + “从 1 出发, 返回 0 前, 途径  $k$  的次数”, 后者我们计算出固定为 2, 前者又等于 “从  $k-1$  出发, 返回 0 前, 途径  $k-1$  的次数  $N_{k-1}$ ”, 这说明  $N_k = N_{k-1} + 2$ , 结论就很直观了。

### 2.6.3 应随部分习题

#### 题目 2.19. 第 4 次作业 4

- (1). 对于不可约、非周期马氏链的每一个状态  $i, j$ , 存在  $N$  使得  $p_{ij}(r) > 0, \forall r \geq N$ .
- (2). 假设  $\mathbf{P}$  是某个具有  $n$  个状态的不可约、非周期马氏链的转移矩阵, 证明存在一个函数  $f$ , 使得对于所有的  $i, j \in S, r > f(n)$  都有  $p_{ij}(r) > 0$ .

解答:



#### 题目 2.20. 第 4 次作业 5

证明:  $i \rightarrow j$ , 则 “从  $i$  出发, 在重新回到  $i$  前能够到达  $j$ ” 的概率大于 0. 进一步地, 在不可约、常返马氏链上有  $i \rightarrow j$ , 证明:  $f_{ij} = 1$ .

解答:

首先证明:  $\tilde{L}_{ij} := P(\exists n, \text{ s.t. } X_n=j, \text{ 且 } X_k \neq i, \forall 1 \leq k \leq n-1 \mid X_0=i) > 0$ . 令  $L_{ij}(n) = P(X_n=j, X_k \neq i, \forall 1 \leq k \leq n-1 \mid X_0=i)$   
 则  $\tilde{L}_{ij} \geq L_{ij}(n), \forall n \geq 1$ . 假设  $\tilde{L}_{ij} = 0$ , 则  $L_{ij}(n) = 0, \forall n \geq 1$ . 因  $i \neq j$ , 则  $P_{ij}(m) = \sum_{k=0}^{m-1} P_{ii}(k) L_{ij}(m-k), \forall m \geq 1$   
 所以  $P_{ij}(m) = 0, \forall m \geq 1$ , 与可达  $j$  矛盾;  
 (在第三周作业 Ex6.2.1 中证过).  
 其次要证, 若马氏链不可约且常返, 则  $f_{ij} = 1, \forall i, j \in S$ . 不妨取  $i \neq j$ , 则  $\exists m > 0$  s.t.  $L_{ji}^{(m)} > 0$ .  
 注意到  $P(X_k \neq j, \forall 1 \leq k \leq n \mid X_0=j) \geq P(X_m=i, X_k \neq j, \forall 1 \leq k \leq n \mid X_0=j) \stackrel{\text{马尔可夫性}}{=} L_{ji}^{(m)} \cdot P(X_k \neq j, \forall 1 \leq k \leq n-m \mid X_0=j)$   
 同时  $P(X_k \neq j, \forall 1 \leq k \leq n \mid X_0=i) = 1 - \sum_{k=1}^n f_{ij}(k)$ ; 因为  $P(X_k \neq j, \forall 1 \leq k \leq n-1 \mid X_0=i) = P(X_k \neq j, \forall 1 \leq k \leq n \mid X_0=i) = f_{ij}(n)$   
 这代得:  $P(X_k \neq j, \forall 1 \leq k \leq n \mid X_0=i) = -\sum_{k=2}^n f_{ij}(k) + P(X_1 \neq j \mid X_0=i) \stackrel{\forall n \geq 2}{=} 1 - \sum_{k=1}^n f_{ij}(k)$ ;  
 同理  $P(X_k \neq j, \forall 1 \leq k \leq n \mid X_0=j) = 1 - \sum_{k=1}^n f_{jj}(k)$ ; 因此令  $n \rightarrow \infty$ , 可得  $1 - f_{jj} \geq L_{ji}^{(m)} (1 - f_{ij})$ ,  
 由  $j$  常返, 则  $f_{jj} = 1$ , 因此  $f_{ij} = 1$ . 由  $i, j$  任意, 命题得证  $\square$

题目 2.21. 第 5 次作业 5

5. 设  $X$  为离散时间马氏链, 状态空间  $S = \{1, 2, 3, 4, 5, 6, 7\}$ , 转移矩阵如下图所示:

$$P = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix} \end{matrix}$$

- 问题: (1) 画出转移概率图, 用分解定理对该马氏链进行分解;
- (2) 分析每个状态的性质, 包括瞬时、零常返、正常返、周期;
- (3) 计算每个状态的平均返回时间;
- (4) 对  $i \in \{1, 2, 3, 6, 7\}, j \in S$ , 计算  $\lim_{n \rightarrow \infty} p_{ij}(n)$ .
- (5) [思考题, 不用写到作业本上] 对  $i \in \{4, 5\}, j \in S$ , 考虑  $\lim_{n \rightarrow \infty} p_{ij}(n)$ .

解答:

(1)  $S = \{1, 2, 3\} \cup \{6, 7\} \cup \{4, 5\}$ ,  $\{1, 2, 3\}$  和  $\{6, 7\}$  均为不可约且闭的常返态集,  $\{4, 5\}$  为瞬态态集;  
 转移概率图:

(2)  $\{1, 2, 3\}$  和  $\{6, 7\}$  均正常返, 周期为 1 (因为  $p_{11} > 0, p_{66} > 0$ );  
 $\{4, 5\}$  均瞬态, 5 的周期为 1 ( $p_{55} > 0$ ), 由 4, 5 互相可达, 4 的周期为 1;

(3)  $H_4 = H_5 = \infty$ ; 设  $\{1, 2, 3\}$  对应子链的平稳分布为  $\pi^{(1)} = (\pi_1^{(1)}, \pi_2^{(1)}, \pi_3^{(1)})$ , 则

$$\begin{cases} \frac{1}{6}\pi_2^{(1)} + \frac{1}{2}\pi_3^{(1)} = \pi_1^{(1)} \\ \pi_1^{(1)} + \frac{1}{2}\pi_2^{(1)} + \frac{1}{2}\pi_3^{(1)} = \pi_2^{(1)} \\ \frac{1}{6}\pi_2^{(1)} = \pi_3^{(1)} \\ \pi_1^{(1)} + \pi_2^{(1)} + \pi_3^{(1)} = 1 \end{cases} \Rightarrow \begin{cases} \pi_1 = \frac{1}{5} \\ \pi_2 = \frac{3}{5} \\ \pi_3 = \frac{1}{5} \end{cases}, \quad \text{则 } H_1 = \pi_1^{-1} = 5, H_2 = \pi_2^{-1} = \frac{5}{3}, H_3 = \pi_3^{-1} = 5;$$

设  $\{6, 7\}$  对应子链的平稳分布为  $\pi^{(2)} = (\pi_6^{(2)}, \pi_7^{(2)})$ , 则

$$\begin{cases} \frac{1}{3}\pi_6^{(2)} + \frac{2}{3}\pi_7^{(2)} = \pi_6^{(2)} \\ \frac{2}{3}\pi_6^{(2)} + \frac{1}{3}\pi_7^{(2)} = \pi_7^{(2)} \\ \pi_6^{(2)} + \pi_7^{(2)} = 1 \end{cases} \Rightarrow \begin{cases} \pi_6^{(2)} = \frac{2}{17} \\ \pi_7^{(2)} = \frac{9}{17} \end{cases}$$

则  $H_6 = \frac{17}{2}, H_7 = \frac{17}{9}$ ;

(4)  $i \in \{1, 2, 3\}$ :  $\lim_{n \rightarrow \infty} p_{ij}(n) = \begin{cases} \frac{1}{5}, & j=1 \\ \frac{3}{5}, & j=2 \\ \frac{1}{5}, & j=3 \\ 0, & j \notin \{1, 2, 3\} \end{cases}$  (因为不可达)

$i \in \{6, 7\}$ :  $\lim_{n \rightarrow \infty} p_{ij}(n) = \begin{cases} \frac{2}{17}, & j=6 \\ \frac{9}{17}, & j=7 \\ 0, & j \notin \{6, 7\} \end{cases}$  (因为不可达)

## 2.7 关于本章记号的说明

汇总一下正文部分用过的记号，目的是让读者清楚地看出哪些记号是等价的，以及有些场合下默认使用的记号是什么含义。

1. 给定初始状态的条件期望、条件概率：

$$\mathbb{E}_x(Y) \stackrel{\text{def}}{=} \mathbb{E}[Y|X_0 = x] = \int Y(\omega)p(X_0, d\omega)$$

$$\mathbb{P}_x(A) \stackrel{\text{def}}{=} \mathbb{P}(A|X_0 = x) = p(x, A) = \int I_A(\omega)p(X_0, d\omega)$$

注意给定的条件可以是随机变量：

$$\mathbb{P}_X(A) : \omega \rightarrow \mathbb{P}(A|X_0 = X(\omega)) = p(X(\omega), A) = \int I_A(\theta)p(X_0(\omega), d\theta)$$

$$\mathbb{E}_X[Y] : \omega \rightarrow \mathbb{E}_{X(\omega)}[Y] = \int Y(\theta)p(X(\omega), d\theta)$$

2. 一步、多步转移概率：

$$p_{ij} \stackrel{\text{def}}{=} p(i, j) = \mathbb{P}(X_1 = j|X_0 = i)$$

$$p_{ij}(n) = p^n(i, j) \stackrel{\text{def}}{=} \mathbb{P}(X_n = j|X_0 = i)$$

3.  $\mathbb{E}$  默认为  $(\Omega, \mathbb{P})$  上对  $\mathbb{P}$  积分，也是  $\mathbb{E}_\mu$  的简化表达（特指  $(\Omega, \mathbb{P}_\mu)$  上对  $\mathbb{P}_\mu$  积分）。
4. 首次到达时刻：如无特殊说明，

$$T_x \stackrel{\text{def}}{=} \inf\{n \geq 1 : X_n = x\}, V_x \stackrel{\text{def}}{=} \inf\{n \geq 0 : X_n = x\}$$

还可以是集合：

$$T_A \stackrel{\text{def}}{=} \inf\{n \geq 1 : X_n \in A\}, V_A \stackrel{\text{def}}{=} \inf\{n \geq 0 : X_n \in A\}$$

第  $k$  次到达时刻：

$$T_y^0 = 0, T_y^k \stackrel{\text{def}}{=} \inf\{n > T_y^{k-1} : X_n = y\}$$

5. 途径次数：一般来说并不计入初始状态，

$$N(y) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} I_{\{X_n=y\}}$$

6. 有限时间内到达概率：

$$f_{xy} = \rho_{xy} \stackrel{\text{def}}{=} \mathbb{P}_x(T_y < +\infty)$$

7. 构造平稳测度：如无特殊说明，

$$\rho_x(y) = \mu_x(y) \stackrel{\text{def}}{=} \mathbb{E}_x \left[ \sum_{n=0}^{T_x-1} I_{\{X_n=y\}} \right]$$

## 第 3 章 Poisson 过程

从这一章起开始讨论连续时间的随机过程，Poisson 过程是一类连续时间离散状态的随机过程。

### 3.1 Poisson 过程的定义

#### 定义 3.1.1 (Poisson 过程的定义)

对于随机过程  $N = \{N_t, t \in \mathbb{R}, t \geq 0\}$ ，若满足：

(1). 独立增量： $\forall t_1 < s_1 < t_2 < s_2 < \cdots < t_n < s_n$ ，随机变量

$$N_{s_1-t_1}, N_{s_2-t_2}, \cdots, N_{s_n-t_n}$$

相互独立。

(2).  $N_0 = 0$ ， $\forall t < s$ ， $N_s - N_t \sim \text{Poisson}(\lambda(s-t))$ 。

则称  $N$  为参数/比率/速率 (rate) 为  $\lambda$  的泊松 (Poisson) 过程。

#### 推论 3.1.1

类似于马氏链的转移概率，我们可以计算出：对于  $s > t, m \geq n$ ，

$$\mathbb{P}(N_s = m | N_t = n) = \mathbb{P}(N_s - N_t = m - n) = e^{-\lambda(s-t)} \frac{(\lambda(s-t))^{m-n}}{(m-n)!}$$

倒过来也能求：

$$\mathbb{P}(N_t = n | N_s = m) = C_m^n \left(\frac{s}{t}\right)^n \left(\frac{t-s}{t}\right)^{m-n}$$

证明

$$\begin{aligned} \mathbb{P}(N_s = m | N_t = n) &= \frac{\mathbb{P}(N_s = m, N_t = n)}{\mathbb{P}(N_t = n)} \\ &= \frac{\mathbb{P}(N_s - N_t = m - n, N_t - N_0 = n)}{\mathbb{P}(N_t = n)} \\ &= \frac{\mathbb{P}(N_s - N_t = m - n) \mathbb{P}(N_t - N_0 = n)}{\mathbb{P}(N_t = n)} \\ &= \mathbb{P}(N_s - N_t = m - n) = e^{-\lambda(s-t)} \frac{(\lambda(s-t))^{m-n}}{(m-n)!} \end{aligned}$$

□

## 例 3.1.1.

假设进入超市中的顾客数量是一个参数为  $\lambda$  的 Poisson 过程, 记作  $N = \{N_t, t \geq 0\}$ , 已知顾客的性别与  $N$  独立, 是男性的概率为  $1/3$ , 记  $M_t$  为时刻  $t$  之前进入超市的男性顾客数量, 求证:  $M = \{M_t, t \geq 0\}$  是一个参数为  $\lambda/3$  的 Poisson 过程。

**证明**  $M_s - M_t$  只与  $N_s - N_t$  有关, 所以继承其独立增量性。下面我们求  $M_s - M_t$  的分布,

$$\begin{aligned} \mathbb{P}(M_s - M_t = n) &= \sum_{k=n}^{\infty} \mathbb{P}((s, t] \text{ 时间段内进入了 } k \text{ 名顾客, 其中有 } n \text{ 名男性}) \\ &= \sum_{k=n}^{\infty} \mathbb{P}((s, t] \text{ 时间段内进入了 } k \text{ 名顾客}) \mathbb{P}(k \text{ 名顾客中有 } n \text{ 名男性}) \\ &= \sum_{k=n}^{\infty} \mathbb{P}(N_s - N_t = k) \binom{k}{n} \left(\frac{1}{3}\right)^n \left(\frac{2}{3}\right)^{k-n} \\ &= \sum_{k=n}^{\infty} e^{-\lambda(s-t)} \frac{[\lambda(s-t)]^k}{k!} \frac{k!}{n!(k-n)!} \left(\frac{1}{3}\right)^n \left(\frac{2}{3}\right)^{k-n} \\ &= \left(\frac{1}{3}\right)^n e^{-\lambda(s-t)} \frac{[\lambda(s-t)]^n}{n!} \sum_{k=n}^{\infty} \frac{[\lambda(s-t)]^{k-n}}{(k-n)!} \left(\frac{2}{3}\right)^{k-n} \\ &= \left(\frac{1}{3}\right)^n e^{-\lambda(s-t)} \frac{[\lambda(s-t)]^n}{n!} \sum_{k=0}^{\infty} \frac{[\lambda(s-t)]^k}{k!} \left(\frac{2}{3}\right)^k \\ &= \left(\frac{1}{3}\right)^n e^{-\lambda(s-t)} \frac{[\lambda(s-t)]^n}{n!} e^{\frac{2}{3}\lambda(s-t)} \\ &= e^{-\frac{1}{3}\lambda(s-t)} \frac{[\frac{1}{3}\lambda(s-t)]^n}{n!} \end{aligned}$$

所以  $M_s - M_t \sim \text{Poisson}(\frac{1}{3}\lambda)$ . □

## 定义 3.1.2 (复合 Poisson 过程)

设  $N = \{N_t, t \geq 0\}$  是一个参数为  $\lambda$  的 Poisson 过程, 而  $Y_1, \dots, Y_n, \dots$  是独立同分布的随机变量, 且独立于  $N$ , 定义:

$$S_t = \sum_{i=1}^{N_t} Y_i = Y_1 + Y_2 + \dots + Y_{N_t}, \quad t \geq 0$$

则称  $\{S_t, t \geq 0\}$  为复合 Poisson 过程。 ✿

## 3.2 构造 Poisson 过程

## 定理 3.2.1

我们按照以下步骤构建出一个 Poisson 过程:

1. 一系列独立同分布的随机变量  $\xi_1, \xi_2, \dots$  服从参数为  $\lambda$  的指数分布, 代表对某个事件发生时间的独立重复测量。
2. 令  $T_0 = 0, T_n = \xi_1 + \dots + \xi_n$ , 代表第  $n$  次事件发生的时刻。
3. 对于  $t \geq 0$ , 令  $N_t = \sup\{n \geq 0 : T_n \leq t\}$ , 实际上

$$N_t = \sup\{n \geq 0 : T_n \leq t\} = \inf\{n \geq 1 : T_n > t\} = \sum_{n=1}^{\infty} I_{\{T_n \leq t\}}$$

即时间  $t$  前发生事件的次数。

那么,  $N = \{N_t, t \geq 0\}$  为 Poisson 过程。

**证明** 验证定义 3.1.1 的两条性质。根据构造过程可知,  $N_0 = 0$ ,  $T_n$  的密度函数为:

$$T_n \sim f_n(s) = \frac{\lambda^n s^{n-1} e^{-\lambda s}}{(n-1)!}, s \geq 0$$

然后我们来求  $N_t$  的分布,

$$\begin{aligned} \mathbb{P}(N_t = 0) &= \mathbb{P}(T_1 > t) = \mathbb{P}(\xi_1 > t) = e^{-\lambda t} \\ \mathbb{P}(N_t = n) &= \mathbb{P}(T_n \leq t < T_{n+1}) \\ &= \mathbb{P}(T_n \leq t < T_n + \xi_{n+1}) \\ &= \iint_{\{(s,u): s \leq t < s+u\}} \frac{\lambda^n s^{n-1} e^{-\lambda s}}{(n-1)!} \cdot \lambda e^{-\lambda u} ds du && \text{注意 } T_n \text{ 独立于 } \xi_{n+1} \\ &= \frac{e^{-\lambda t} (\lambda t)^n}{n!} \end{aligned}$$

接下来证明独立增量, 对于  $s > 0$ ,

$$\begin{aligned} \mathbb{P}(T_{n+1} - t \geq s | N_t = n) &= \frac{\mathbb{P}(T_{n+1} - t \geq s, N_t = n)}{\mathbb{P}(N_t = n)} \\ &= \frac{\mathbb{P}(T_{n+1} - t \geq s, T_n \leq t)}{\mathbb{P}(N_t = n)} \\ &= \frac{\mathbb{P}(\xi_{n+1} \geq s + t - T_n, T_n \leq t)}{\mathbb{P}(N_t = n)} \\ &= \frac{1}{\mathbb{P}(N_t = n)} \int_0^t dv \int_{t+s-v}^{+\infty} \frac{\lambda^n v^{n-1}}{(n-1)!} \lambda e^{-\lambda u} du \\ &= \frac{1}{\mathbb{P}(N_t = n)} e^{-\lambda(t+s)} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda s} \end{aligned}$$

令  $T'_1 = T_{N_t+1} - t$ ,  $T'_k = T_{N_t+k} - T_{N_t+k-1}$ ,  $k \geq 2$ , **Claim 1:**  $\{T'_k, k \geq 1\}$  独立同分布, 服从参数为  $\lambda$  的指数分布, 且与  $N_t$  独立。证明如下: 观察

$$\mathbb{P}(T_n \leq t, T_{n+1} - t \geq s, T_{n+2} - T_{n+1} \geq v_2, \dots, T_{n+m} - T_{n+m-1} \geq v_m)$$

实际上就是

$$\mathbb{P}(T_n \leq t, T_{n+1} - t \geq s, \xi_{n+2} \geq v_2, \dots, \xi_{n+m} \geq v_m)$$

根据独立性, 这等于

$$\mathbb{P}(T_n \leq t, T_{n+1} - t \geq s) e^{-\lambda(v_2 + \dots + v_m)}$$

于是

$$\begin{aligned} &\mathbb{P}(T_{n+1} - t \geq s, T_{n+2} - T_{n+1} \geq v_2, \dots, T_{n+m} - T_{n+m-1} \geq v_m | N_t = n) \\ &= \mathbb{P}(T_n \leq t, T_{n+1} - t \geq s, T_{n+2} - T_{n+1} \geq v_2, \dots, T_{n+m} - T_{n+m-1} \geq v_m) / \mathbb{P}(N_t = n) \\ &= e^{-\lambda(v_2 + \dots + v_m)} \mathbb{P}(T_{n+1} \geq s + t | N_t = n) \\ &= e^{-\lambda(v_2 + \dots + v_m)} e^{-\lambda s} \end{aligned}$$

所以

$$\begin{aligned} &\mathbb{P}(T'_1 \geq s, T'_2 \geq v_2, \dots, T'_m \geq v_m, N_t \leq l) \\ &= \sum_{n=0}^l \mathbb{P}(T'_1 \geq s, T'_2 \geq v_2, \dots, T'_m \geq v_m, N_t = n) \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=0}^l \mathbb{P}(T_{n+1} - t \geq s, T_{n+2} - T_{n+1} \geq v_2, \dots, T_{n+m} - T_{n+m-1} \geq v_m, N_t = n) \\
&= \sum_{n=0}^l e^{-\lambda(v_2 + \dots + v_m)} e^{-\lambda s} \mathbb{P}(N_t = n) \\
&= e^{-\lambda(v_2 + \dots + v_m)} e^{-\lambda s} \mathbb{P}(N_t \leq l) \\
&= \mathbb{P}(T'_1 \geq s) \mathbb{P}(T'_2 \geq v_2) \cdots \mathbb{P}(T'_m \geq v_m) \cdot \mathbb{P}(N_t \leq l)
\end{aligned}$$

于是 Claim 1 得证。Claim 2: 对于  $t_0 < t_1 < \dots < t_n$ ,

$$\mathbb{P}(N_{t_i} - N_{t_{i-1}} = k_i, i = 1, 2, \dots, n) = \prod_{i=1}^n e^{-\lambda(t_i - t_{i-1})} \frac{(\lambda(t_i - t_{i-1}))^{k_i}}{k_i!}$$

并且各个  $N_{t_i} - N_{t_{i-1}}$  是独立的。我们先考虑  $i = 2$  的简单情形, 一般情形的证明同理。令  $T''_1 = T_{N_{t_1}+1} - t_1$ ,  $T''_k = T_{N_{t_1}+k} - T_{N_{t_1}+k-1}$ ,  $k \geq 2$ , 则由 Claim 1,  $\{T''_k, k \geq 1\}$  独立同分布, 服从参数为  $\lambda$  的指数分布, 且与  $N_{t_1}$  独立, 又因为

$$\begin{aligned}
\{N_{t_2} - N_{t_1} = m\} &= \{T_{N_{t_1}+m} \leq t_2, T_{N_{t_1}+m+1} > t_2\} \\
&= \{T_{N_{t_1}+m} - t_1 \leq t_2 - t_1, T_{N_{t_1}+m+1} - t_1 > t_2 - t_1\} \\
&= \left\{ \sum_{k=1}^m T''_k \leq t_2 - t_1 < \sum_{k=1}^{m+1} T''_k \right\}
\end{aligned}$$

通过此式能看出  $N_{t_2} - N_{t_1}$  与  $N_{t_1}$  独立, 并且可以利用  $T''_k \sim \exp\{\lambda\}$  求出  $N_{t_2} - N_{t_1}$  的分布, 计算过程省略。□

从这个定理能看出 Poisson 过程在现实中的应用场景, 即“一段时间内某事件的发生次数”。

### 3.3 Poisson 的另一种定义 \*

#### 定义 3.3.1

对于随机过程  $N = \{N_t : t \geq 0\}$ , 若满足:

- (1).  $N_0 = 0$ .
- (2).  $\forall s < t, N_s \leq N_t$ .
- (3).

$$\mathbb{P}(N_{t+h} = m + n | N_t = n) = \begin{cases} 1 - \lambda h + o(h) & , m = 0 \\ \lambda h + o(h) & , m = 1 \\ o(h) & , m \geq 2 \end{cases}$$

这里  $o(h)$  代表  $h \rightarrow 0$  时的低阶无穷小:  $\lim_{h \rightarrow 0} o(h)/h = 0$ , 例如  $h^2$ .

- (4).  $s < t$  时,  $N_t - N_s$  与  $\{N_{s'}, \forall s' \in [0, s]\}$  独立。

则称  $N$  为强度为  $\lambda$  的泊松 (Poisson) 过程。

#### 定理 3.3.1

$N_t$  服从参数为  $\lambda t$  的 Poisson 分布。

**证明** 记  $p_j(t) = \mathbb{P}(N_t = j)$ , 那么  $j \geq 1$  时,

$$\begin{aligned} p_j(t+h) &= \mathbb{P}(N_{t+h} = j) \\ &= \sum_{i \leq j} \mathbb{P}(N_{t+h} = j, N_t = i) \\ &= \sum_{i \leq j} p_i(t) \cdot \mathbb{P}(N_{t+h} = j | N_t = i) \\ &= p_j(t) \cdot \mathbb{P}(N_{t+h} = j | N_t = j) + p_{j-1}(t) \cdot \mathbb{P}(N_{t+h} = j | N_t = j-1) + o(h) \\ &= p_j(t) \cdot (1 - \lambda h + o(h)) + p_{j-1}(t) \cdot (\lambda h + o(h)) + o(h) \\ &= p_j(t) \cdot (1 - \lambda h) + p_{j-1}(t) \cdot \lambda h + o(h) \end{aligned}$$

$j = 0$ , 则没有第二项。稍作变换可得

$$\frac{p_j(t+h) - p_j(t)}{h} = -\lambda p_j(t) + \lambda p_{j-1}(t) + \frac{o(h)}{h}$$

令  $h \rightarrow 0$  可得

$$\begin{aligned} \frac{d}{dt} p_j(t) &= -\lambda p_j(t) + \lambda p_{j-1}(t) \\ \frac{d}{dt} p_0(t) &= -\lambda p_0(t) \end{aligned}$$

同时也初值  $p_0(0) = 1, p_j(0) = 0, j \geq 1$ .

要解这个微分方程组, 既可以用递推的方法, 也可以用生成函数。具体过程不写了, 最终能解得

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}$$

□

接下来, 令

$$T_n = \inf\{t \geq 0 : N_t = n\}, n = 0, 1, \dots$$

为首次达到  $n$  的时刻,

$$X_n = T_n - T_{n-1}, n = 1, 2, \dots$$

为从  $n-1$  到  $n$  的等待时间。一个简单的推论是:

$$N_t = m \Leftrightarrow T_m \leq t < T_{m+1}$$

### 定理 3.3.2

$X_1, X_2, \dots$  独立同分布, 且服从参数为  $\lambda$  的指数分布。

**证明** 考虑  $X_1$  的分布,

$$\mathbb{P}(X_1 > t) = \mathbb{P}(N(t) = 0) = e^{-\lambda t}$$

所以  $X_1$  服从参数为  $\lambda$  的指数分布。

由独立增量可知

$$\mathbb{P}(X_2 > t | X_1 = t_1) = \mathbb{P}(N(t_1 + t) - N(t_1) = 0 | N(t_1) = 1, N(t) = 0, t < t_1) = \mathbb{P}(N(t_1 + t) - N(t_1) = 0) = e^{-\lambda t}$$

所以  $X_2$  和  $X_1$  独立同分布, 递推可证。□

### 推论 3.3.1 (Poisson 过程的马氏性)

设  $0 \leq t_1 < \dots < t_k < t < t + s$ ,  $0 \leq n_1 \leq \dots \leq n_k \leq n < n + m$ , 则

$$\mathbb{P}(N_{t+s} = n + m | N_t = n, N_{t_1} = n_1, \dots, N_{t_k} = n_k) = \mathbb{P}(N_{t+s} = n + m | N_t = n)$$



## 3.4 Poisson 过程的更多性质 \*

### 定理 3.4.1 (合并)

设  $\{X_t, t \geq 0\}$  和  $\{Y_t, t \geq 0\}$  是两个相互独立的 Poisson 过程, 参数分别为  $\lambda, \mu$ , 则  $\{X_t + Y_t, t \geq 0\}$  是参数为  $\lambda + \mu$  的 Poisson 过程。

**证明** 记  $N_t = X_t + Y_t$ , 我们来验证定义 3.3.1 中的四条要求, (1)(2)(4) 显然, 我们只说明 (3):

$$\begin{aligned} \mathbb{P}(N_{t+h} = n | N_t = n) &= \mathbb{P}(X_{t+h} - X_t = 0) \mathbb{P}(Y_{t+h} - Y_t = 0) \\ &= (1 - \lambda h + o(h))(1 - \mu h + o(h)) \\ &= 1 - (\lambda + \mu)h + o(h) \end{aligned}$$

$$\begin{aligned} \mathbb{P}(N_{t+h} = n + 1 | N_t = n) &= \mathbb{P}(X_{t+h} - X_t = 1) \mathbb{P}(Y_{t+h} - Y_t = 0) + \mathbb{P}(X_{t+h} - X_t = 0) \mathbb{P}(Y_{t+h} - Y_t = 1) \\ &= (\lambda h + o(h))(1 - \mu h + o(h)) + (\mu h + o(h))(1 - \lambda h + o(h)) \\ &= (\lambda + \mu)h + o(h) \end{aligned}$$

$$\begin{aligned} \mathbb{P}(N_{t+h} = n + 2 | N_t = n) &= \mathbb{P}(X_{t+h} - X_t = 1) \mathbb{P}(Y_{t+h} - Y_t = 1) + o(h) \\ &= (\lambda h + o(h))(\mu h + o(h)) + o(h) \\ &= o(h) \end{aligned}$$

□

### 定理 3.4.2 (细分)

$\{X_t, t \geq 0\}$  是参数为  $\lambda$  的 Poisson 过程,  $\varepsilon_1, \varepsilon_2, \dots$  独立同分布, 分布为 1, 0 的两点分布,  $\mathbb{P}(\varepsilon_1 = 1) = p$ , 则

$$\left\{ Y_t = \sum_{n=1}^{X_t} I_{\{\varepsilon_n=1\}}, t \geq 0 \right\}$$

与

$$\left\{ Z_t = \sum_{n=1}^{X_t} I_{\{\varepsilon_n=0\}}, t \geq 0 \right\}$$

是两个相互独立的 Poisson 过程, 参数分别为  $\lambda p$  和  $\lambda(1-p)$ .

### 定理 3.4.3

给定  $N(t) = n$  的条件下,  $(T_1, \dots, T_n)$  的联合密度为

$$f(t_1, t_2, \dots, t_n) = \frac{n!}{t^n} I_{\{0 < t_1 < t_2 < \dots < t_n < t\}}$$

相当于  $(0, t)$  均匀分布上的  $n$  个次序统计量。

上课讲的两个证明都不太严谨, 这里就省略了。从直观上理解, 就是因为等待时间独立同分布, 前  $n$  个等待时间之和  $\leq t$ , 所以这些时间点都是均匀分布在  $(0, t)$  上。

### 例 3.4.1.

假设乘客按照速率为  $\lambda$  的 Poisson 过程到达火车站, 且火车在时刻  $t$  离站, 求时间段  $(0, t]$  内乘客的总等待时间的期望。

**解答:** 用  $N_t$  代表  $t$  时乘客数量, 那么时间段  $(0, t]$  内就有  $N_t$  名乘客进行了等车。

第 1 位乘客在  $T_1$  时到达, 等待了  $t - T_1$ ; 第 2 位乘客在  $T_2$  时到达, 等待了  $t - T_2$ ; 以此类推, 总等待时间为

$$\sum_{n=1}^{N_t} (t - T_n)$$

先求条件期望:

$$\begin{aligned} \mathbb{E} \left[ \sum_{n=1}^{N_t} (t - T_n) \middle| N_t = k \right] &= kt - \mathbb{E} \left[ \sum_{n=1}^k T_n \middle| N_t = k \right] \\ &= kt - \mathbb{E} \left[ \sum_{n=1}^k U_n \middle| N_t = k \right] \\ &= kt - \frac{kt}{2} = \frac{kt}{2} \end{aligned}$$

其中  $U_n$  服从  $(0, t)$  上的均匀分布, 那么

$$\mathbb{E} \left[ \sum_{n=1}^{N_t} (t - T_n) \right] = \mathbb{E} \left[ \mathbb{E} \left[ \sum_{n=1}^{N_t} (t - T_n) \middle| N_t \right] \right] = \mathbb{E} \left[ \frac{tN_t}{2} \right] = \frac{\lambda t^2}{2}$$

## 3.5 生过程\*

### 定义 3.5.1 (生过程)

随机过程  $N = \{N(t), t \geq 0\}$  若满足:

- (1). 状态空间为  $S = \{0, 1, 2, \dots\}$ .
- (2).  $N(0) \geq 0, \forall s < t, N(s) < N(t)$ .
- (3).

$$\mathbb{P}(N(t+h) = n+m | N(t) = n) = \begin{cases} 1 - \lambda_n h + o(h) & , m = 0 \\ \lambda_n h + o(h) & , m = 1 \\ o(h) & , m \geq 2 \end{cases}$$

(4).  $s < t$  时, 给定  $N(s)$  的条件下,  $N(t) - N(s)$  与  $\{N(s'), \forall s' \in [0, s]\}$  独立。  
则称  $N$  为参数为  $\lambda_0, \lambda_1, \dots$  的生过程。

**例 3.5.1.**

当  $\lambda_n = \lambda$  为常数时, 生过程就是 (初值可能不为零的) Poisson 过程; 考虑这样一种情形: 每个个体以速率  $\lambda$  独立繁育后代, 即  $\lambda_n = \lambda \cdot n$ , 这称为简单生过程; 带迁移的简单生过程:  $\lambda_n = \lambda n + v$ , 其中常数  $v$  称为迁移速率。

**定义 3.5.2**

定义

$$T_n = \inf\{t \geq 0 : N(t) = n\}, \forall n \in \mathbb{N}$$

以及

$$T_\infty = \lim_{n \rightarrow \infty} T_n$$

如果  $N(t)$  在有限时间内 (概率为 1) 达到  $+\infty$ , 即:

$$\mathbb{P}(T_\infty = +\infty) = \mathbb{P}(\lim_{n \rightarrow +\infty} T_n = +\infty) = 1$$

则称  $N$  会爆破。否则, 则称其不会爆破。

**定理 3.5.1**

生过程  $N$  不会爆破  $\Leftrightarrow \sum \lambda_n^{-1} = +\infty$ .

**证明** 记  $X_n = T_n - T_{n-1} \sim \exp\{\lambda_{n-1}\}$ , 则

$$T_\infty = \lim_{m \rightarrow \infty} \sum_{n=1}^m X_n$$

如果  $\sum \lambda_n^{-1} < \infty$ , 则

$$\mathbb{E}[T_\infty] = \lim_{m \rightarrow \infty} \sum_{n=1}^m \mathbb{E}[X_n] = \sum_{n=1}^{\infty} \lambda_{n-1}^{-1} < +\infty$$

从而  $T_\infty < \infty$  a.s.  $\Rightarrow \mathbb{P}(T = \infty) = 0 < 1$ ,  $N$  不爆破。

如果  $\sum \lambda_n^{-1} = \infty$ , 则

$$\mathbb{E}[e^{-T_\infty}] = \lim_{m \rightarrow \infty} \prod_{i=1}^m \mathbb{E}[e^{-X_i}] = \lim_{m \rightarrow \infty} \prod_{i=1}^m \frac{1}{1 + \lambda_{i-1}^{-1}} \leq \lim_{m \rightarrow \infty} \frac{1}{\sum_{n=1}^m \lambda_{n-1}^{-1}} = 0$$

从而  $T_\infty = \infty$  a.s.,  $N$  会爆破。 □

**定理 3.5.2**

记

$$p_{ij}(t) = \mathbb{P}(N(s+t) = j | N(s) = i) = \mathbb{P}(N(t) = j | N(0) = i)$$

向前方程:  $\forall i, j \in \mathbb{N}_+$ ,

$$\begin{cases} \frac{d}{dt} p_{ij}(t) = \lambda_{j-1} p_{i,j-1}(t) - \lambda_j p_{ij}(t), \forall j \geq i \\ \lambda_{-1} = 0 \\ p_{ij}(0) = \delta_{ij} \end{cases}$$

向后方程:  $\forall i, j \in \mathbb{N}_+$ ,

$$\begin{cases} \frac{d}{dt} p_{ij}(t) = \lambda_i p_{i+1,j}(t) - \lambda_i p_{ij}(t), \forall j \geq i \\ p_{ij}(0) = \delta_{ij} \end{cases}$$

**证明** 边界条件由定义直接得出:  $p_{ij}(0) = \delta_{ij}$ .

任取  $j \geq i \geq 0, t \geq 0, h > 0$ , 有

$$\begin{aligned} p_{ij}(t+h) &= \mathbb{P}(N(t+h) = j | N(0) = i) \\ &= \sum_{i \leq k \leq j} \mathbb{P}(N(t+h) = j, N(t) = k | N(0) = i) \\ &= \sum_{i \leq k \leq j} \mathbb{P}(N(t+h) = j | N(t) = k, N(0) = i) \mathbb{P}(N(t) = k | N(0) = i) \\ &= \sum_{i \leq k \leq j} p_{kj}(h) p_{ik}(t) \end{aligned}$$

$j = i$  时,

$$\begin{aligned} p_{ii}(t+h) &= p_{ii}(h) p_{ii}(t) = (1 - \lambda_i h + o(h)) p_{ii}(t) \\ \Rightarrow \frac{d}{dt} p_{ii}(t) &= \lim_{h \rightarrow 0} \frac{p_{ii}(t+h) - p_{ii}(t)}{h} = \lim_{h \rightarrow 0} \frac{(-\lambda_i h + o(h)) p_{ii}(t)}{h} = -\lambda_i p_{ii}(t) \end{aligned}$$

$j = i + 1$  时,

$$\begin{aligned} p_{i,i+1}(t+h) &= p_{i,i+1}(h) p_{i,i}(t) + p_{i+1,i+1}(h) p_{i,i+1}(t) \\ &= (\lambda_i h + o(h)) p_{i,i}(t) + (1 - \lambda_{i+1} h + o(h)) p_{i,i+1}(t) \\ \Rightarrow \frac{d}{dt} p_{ii}(t) &= \lim_{h \rightarrow 0} \frac{p_{i,i+1}(t+h) - p_{i,i+1}(t)}{h} = \lambda_i p_{i,i}(t) - \lambda_{i+1} h \end{aligned}$$

$j \geq i + 2$  时,  $k$  可取值  $i, i + 1, \dots, j - 1, j$ , 注意到  $k \leq j - 2$  时,  $p_{kj}(h) = o(h)$ , 因此

$$\begin{aligned} p_{ij}(t+h) &= p_{jj}(h) p_{ij}(t) + p_{j-1,j}(h) p_{i,j-1}(t) + o(h) \\ &= (1 - \lambda_j h + o(h)) p_{ij}(t) + (\lambda_{j-1} h + o(h)) p_{i,j-1}(t) + o(h) \\ \Rightarrow \frac{d}{dt} p_{ij}(t) &= \lim_{h \rightarrow 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \lambda_{j-1} p_{i,j-1}(t) - \lambda_j p_{ij}(t) \end{aligned}$$

综上所述, 整理可得向前方程。

得到向前方程的思路可以大致表示为:  $p_{ij}(t+h) = \sum_k \mathbb{P}(i \xrightarrow{t} k \xrightarrow{h} j) = \sum_k p_{ik}(t) p_{kj}(h)$ , 那么我们考虑  $\sum_k \mathbb{P}(i \xrightarrow{h} k \xrightarrow{t} j)$ , 就可得到向后方程, 证明方法类似, 具体细节不再赘述。□

**定理 3.5.3**

向前方程有唯一解, 并且此解满足向后方程。

**定义 3.5.3**

随机过程  $M = \{M(t), t \geq 0\}$  的状态空间为可数集  $S$ , 如果

$$\forall \omega \in \Omega, \forall t \in [0, +\infty), \exists \varepsilon_{t,\omega} > 0 \text{ s.t. } M(t,\omega) = M(t+u,\omega), \forall 0 \leq u < \varepsilon_{t,\omega}$$

则称  $M$  右连续。

**定义 3.5.4**

称随机过程  $M = \{M(t), t \geq 0\}$  有平稳独立增量, 如果  $\forall 0 = t_0 < t_1 < \dots < t_n$ ,

$$M(t_1) - M(t_0), \dots, M(t_n) - M(t_{n-1})$$

相互独立, 且  $\forall 0 \leq s < t$ ,  $M(t) - M(s)$  与  $M(t-s) - M(0)$  同分布。

**例 3.5.2.**

随机过程  $M = \{M(t), t \geq 0\}$  非降、右连续、取值为整数、有平稳独立增量,  $M(0) = 0$ , 且  $M(t) - M(t-) \in \{0, 1\}$ , 则  $M$  为参数为  $\mathbb{E}[M(1)]$  的 Poisson 过程。

## 第 4 章 布朗运动

从这一章开始，使用的教材从 Durrett 换成了 Le Gall，后者的阅读体验真的很好，极力推荐。本章涉及内容为 Chapter 1,2.

### 4.1 高斯空间

这一小节为 Le Gall Chapter 1 的节选。主要内容是关于正态分布的一些知识，在以前的概率论课程中都有所介绍，所以阅读本章内容可以跳过本小节。但是笔者发现自己之前学的很粗糙很混乱，还是决定整理一下。

本节默认在概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上讨论。

#### 定义 4.1.1

一个实值的随机变量  $X$  被称为是高斯变量，如果其（Lebesgue 测度下的）密度函数有以下形式：

$$p_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

也就是我们熟知的正态分布： $X \sim N(\mu, \sigma^2)$ 。

随机向量的情形类似：如果  $X = (X_1, \dots, X_n) \sim N_n(\mu, \Sigma)$ ，称  $X$  是高斯向量。

#### 命题 4.1.1

设  $X_n \sim N(\mu_n, \sigma_n^2)$ ，如果  $X_n \xrightarrow{L^2} X$ ，那么

- (1).  $X \sim N(\mu, \sigma^2)$ ，其中  $\mu = \lim_{n \rightarrow \infty} \mu_n$ ， $\sigma = \lim_{n \rightarrow \infty} \sigma_n$ 。
- (2). 对于  $\forall 1 \leq p < \infty$ ，都有  $X_n \xrightarrow{L^p} X$ 。

**证明** 考虑

$$|\mathbb{E}[X_n] - \mathbb{E}[X]| \leq \mathbb{E}[|X_n - X|] \leq \mathbb{E}[|X_n - X|^2]^{\frac{1}{2}} \rightarrow 0$$

所以  $\mathbb{E}[X] = \mu$ ，同时

$$\begin{aligned} |\text{var}(X_n) - \text{var}(X)| &= |\mathbb{E}[X_n^2] - \mathbb{E}[X^2] + \mu_n^2 - \mu^2| \\ &\leq \mathbb{E}[|X_n - X| \cdot |X_n + X|] + |\mu_n^2 - \mu^2| \\ &\leq \mathbb{E}[|X_n - X|^2]^{\frac{1}{2}} \mathbb{E}[|X_n + X|^2]^{\frac{1}{2}} + |\mu_n^2 - \mu^2| \rightarrow 0 \end{aligned}$$

所以  $\text{var}(X_n) = \sigma_n^2$ . 然后我们考虑  $X$  的特征函数:

$$\mathbb{E}[e^{i\xi X}] = \lim_{n \rightarrow \infty} \mathbb{E}[e^{i\xi X_n}] = \lim_{n \rightarrow \infty} \exp(im_n \xi - \frac{\sigma_n^2}{2} \xi^2) = \exp(im\xi - \frac{\sigma^2}{2} \xi^2)$$

这说明  $X \sim N(\mu, \sigma^2)$ .

因为  $m_n, \sigma_n$  都是有界的, 所以

$$\sup_n \mathbb{E}[|X_n|^q] < +\infty, \forall q \geq 1$$

因此

$$\sup_n \mathbb{E}[|X_n - X|^q] < +\infty, \forall q \geq 1$$

取  $p \geq 1, Y_n = |X_n - X|^p$  依概率收敛到 0, 并且一致可积, 由定理 1.7.2 可得  $Y_n \xrightarrow{L^1} 0$ , 这就是我们要证的结论.  $\square$

**定义 4.1.2**

设  $H$  是  $L^2(\Omega, \mathcal{F}, \mathbb{P})$  的闭线性子空间, 并且  $H$  只包含中心高斯变量 (即服从期望为 0 的正态分布), 则称  $H$  是一个 (中心) 高斯空间.

**例 4.1.1.**

取  $X \sim N(0, 1), \varepsilon$  服从  $\pm$  两点分布并与  $X$  独立,  $Y = \varepsilon X$ , 那么  $X + Y$  不再符合正态分布, 因此不在同一个高斯空间里.

**推论 4.1.1**

如果  $(X_1, \dots, X_n) \sim N_n(\mathbf{0}, \Sigma)$ , 那么  $X_1, \dots, X_n$  张成的线性空间是一个高斯空间.

**定义 4.1.3**

如果随机过程  $X = \{X_t, t \in T\}$  中随机变量的任意有限线性组合都是中心高斯变量, 则称  $X$  是 (中心) 高斯过程.

**命题 4.1.2**

$X = \{X_t, t \in T\}$  是高斯过程, 则由  $X$  张成的闭线性空间是高斯空间.

**证明** 只需注意到命题 4.1.1 保证了可数线性组合也是中心高斯变量.  $\square$

**定理 4.1.1**

$H$  是高斯空间,  $H_i, i \in I$  是其线性子空间, 则  $H_i, i \in I$  两两正交当且仅当  $\sigma(H_i), i \in I$  相互独立.

**证明** ( $\Leftarrow$ ): 如果  $X \in H_i \subset \sigma(H_i)$  且  $Y \in H_j \subset \sigma(H_j), i \neq j$ , 由独立性可知  $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y] = 0$ , 所以  $H_i$  与  $H_j$  正交.

( $\Rightarrow$ ): 即证: 任取  $i_1, \dots, i_p \in I, \sigma(H_{i_j}), j \in \{1, \dots, p\}$  相互独立. 根据单调类定理, 任取

$$\xi_1^j, \xi_2^j, \dots, \xi_{n_j}^j \in H_{i_j}, j \in \{1, \dots, p\}$$

那么只需证明向量  $(\xi_1^j, \xi_2^j, \dots, \xi_{n_j}^j), j \in \{1, \dots, p\}$  相互独立即可. 对于任意一个线性子空间  $H_{i_j}, \xi_1^j, \xi_2^j, \dots, \xi_{n_j}^j$  张成了一个子空间, 取这个子空间上的正交基  $\eta_1^j, \dots, \eta_{m_j}^j$ , 这样就得到了一系列随机变量组成的向量:

$$(\eta_1^1, \dots, \eta_{m_1}^1, \dots, \eta_1^p, \dots, \eta_{m_p}^p)$$

因为分量都在  $H$  中, 这是一个高斯向量, 且分量之间不相关, 因此它们相互独立<sup>1</sup>, 于是定理得证.  $\square$

<sup>1</sup> “正态” + “不相关” 不足以推出独立性, 如例 4.1.1 中的  $X$  和  $Y$  就不相关, 但是显然不独立. 实际上 “正态” + “属于同一个高斯空间”  $\Leftrightarrow$  独立, 这是因为高斯向量联合分布的协方差矩阵  $\Sigma$  可分块  $\Leftrightarrow$  密度函数可分离导致的结果.

对于高斯过程  $X = \{X_t, t \in T\}$ , 我们可以定义协方差函数:

$$\Gamma : T \times T \rightarrow [0, +\infty), (s, t) \mapsto \text{cov}(X_s, X_t)$$

协方差函数  $\Gamma$  完全决定了  $X$  的有限维分布, 也就完全决定了  $X$ . 那么, 函数  $\Gamma$  需要满足什么条件才能保证存在相应的高斯过程?

#### 定理 4.1.2

函数  $\Gamma : T \times T \rightarrow [0, +\infty)$  若满足:

- (1). 对称性:  $\Gamma(s, t) = \Gamma(t, s)$ .
- (2). 正定性:  $c$  作为  $T$  上的实值函数, 有有限的支撑集, 那么

$$\sum_{s, t \in T} c(s)c(t)\Gamma(s, t) \geq 0$$

这一条保证的是  $X_t$  的有限线性组合的方差非负。

证明没办法在这里详细说明, 书上提及这是 Kolmogorov 延拓定理的推论。

#### 定义 4.1.4

$\mu$  是可测空间  $(E, \mathcal{G})$  上  $\sigma$ -有限测度,  $X$  是一个 (以下默认中心) 高斯空间, 如果

$$G : L^2(E, \mathcal{G}, \mu) \rightarrow X$$

是一个等距映射, 则称  $G$  是强度为  $\mu$  的高斯白噪声 (Gaussian White Noise)。

利用等距映射  $G$ , 可以计算:

$$\begin{aligned} \text{var}(G(f)) &= \mathbb{E}[G(f)^2] = \|G(f)\|_{L^2(\Omega, \mathcal{F}, \mathbb{P})}^2 = \|f\|_{L^2(E, \mathcal{G}, \mu)}^2 = \int f^2 d\mu \\ \text{cov}(G(f), G(g)) &= \langle G(f), G(g) \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P})} = \langle f, g \rangle_{L^2(E, \mathcal{G}, \mu)} = \int fg d\mu \end{aligned}$$

特别地, 取  $f = I_A, g = I_B, \mu(A), \mu(B) < +\infty$ , 那么

$$\text{cov}(G(f), G(g)) = \int I_{A \cap B} d\mu = \mu(A \cap B)$$

这说明, 如果  $A$  和  $B$  不相交, 则  $G(f)$  和  $G(g)$  不相关, 进而独立。

## 4.2 准布朗运动

从这一小节开始, 是 Chapter 2 的内容。  $\mathbb{R}_+$  代表  $[0, +\infty)$ 。

#### 定义 4.2.1

记  $\mathcal{B}_+$  为  $\mathbb{R}_+$  上的 Borel 集全体, 即全体开集生成的  $\sigma$  域, 那么 Lebesgue 测度  $m$  是  $(\mathbb{R}_+, \mathcal{B}_+)$  上的一个  $\sigma$  有限测度, 令  $G$  是  $(\mathbb{R}, \mathcal{B}_+)$  上强度为  $m$  的高斯白噪声, 定义

$$B_t = G(I_{[0, t]}), t \geq 0$$

那么, 随机过程  $B = (B_t)_{t \geq 0}$  称为一个准布朗运动 (pre-Brownian Motion)。

#### 命题 4.2.1 (准布朗运动的等价定义)

对于随机过程  $X = (X_t)_{t \geq 0}$ , 以下说法等价:

- (1).  $X$  是准布朗运动。
- (2).  $X$  是高斯过程, 且协方差函数是  $\Gamma(s, t) = s \wedge t$ .
- (3).  $X_0 = 0$  a.s.,  $\forall 0 \leq s < t, X_t - X_s \sim N(0, t - s)$ , 且与  $\sigma(X_r, r \leq s)$  独立。

(4).  $X_0 = 0$  a.s., 任取  $0 = t_0 < t_1 < \cdots < t_p$ ,  $X_{t_i} - X_{t_{i-1}} \sim N(0, t_i - t_{i-1})$ , 且相互独立。

**证明** (1)  $\Rightarrow$  (2): 根据定义可知  $(B_t)_{t \geq 0}$  同属一个高斯空间, 所以是高斯过程, 协方差函数:

$$\Gamma(s, t) = \mathbb{E}[B_t B_s] = \mathbb{E}[G(I_{[0,t]})G(I_{[0,s]})] = \int I_{[0,t]} I_{[0,s]} dm = s \wedge t$$

(2)  $\Rightarrow$  (3):  $\text{var}(X_0) = \Gamma(0, 0) = 0$ , 即  $X_0 \sim N(0, 0)$ , 说明  $X_0 = 0$  a.s., 考虑  $\forall s \geq 0$ , 记

$$H_s = \text{span}\{X_r, 0 \leq r \leq s\}$$

$$\tilde{H}_s = \text{span}\{X_{s+u} - X_s, u \geq 0\}$$

那么二者是正交的, 因为:

$$\mathbb{E}[X_r(X_{s+u} - X_s)] = r \wedge (s+u) - r \wedge s = r - r = 0$$

进而  $\sigma(H_s)$  和  $\sigma(\tilde{H}_s)$  是独立的。特别地取  $t > s$ , 则  $X_t \in \tilde{H}_s$  与  $\sigma(H_s) = \sigma(X_r, 0 \leq r \leq s)$  是独立的。

(3)  $\Rightarrow$  (4): 取  $t = t_p, s = t_{p-1}$ , 则可得  $X_{t_p} - X_{t_{p-1}}$  与  $\sigma(X_r, 0 \leq r \leq t_{p-1})$  独立, 进而与  $X_{t_{p-1}} - X_{t_{p-2}}, \cdots, X_{t_1} - X_{t_0}$  独立; 取  $t = t_{p-1}, s = t_{p-2}$ , 则可得  $X_{t_{p-1}} - X_{t_{p-2}}$  与  $X_{t_{p-2}} - X_{t_{p-3}}, \cdots, X_{t_1} - X_{t_0}$  独立; 以此类推。

(4)  $\Rightarrow$  (1):  $X_t - X_0 \sim N(0, t)$ , 所以  $X$  是高斯过程。对于阶梯函数

$$f = \sum_{i=1}^p \lambda_i I_{(t_{i-1}, t_i]}$$

定义

$$G(f) \stackrel{\text{def}}{=} \sum_{i=1}^p \lambda_i (X_{t_i} - X_{t_{i-1}})$$

容易验证  $G$  是一个等距映射:

$$\begin{aligned} \|G(f) - G(\tilde{f})\|^2 &= \mathbb{E} \left[ \left( \sum_{i=1}^p (\lambda_i - \tilde{\lambda}_i) (X_{t_i} - X_{t_{i-1}}) \right)^2 \right] \\ &= \sum_{i=1}^p (\lambda_i - \tilde{\lambda}_i)^2 \mathbb{E}[(X_{t_i} - X_{t_{i-1}})^2] \\ &= \sum_{i=1}^p (\lambda_i - \tilde{\lambda}_i)^2 (t_i - t_{i-1}) \\ \|f - \tilde{f}\|^2 &= \int \left| \sum_{i=1}^p (\lambda_i - \tilde{\lambda}_i) I_{(t_{i-1}, t_i]} \right|^2 dm \\ &= \sum_{i=1}^p (\lambda_i - \tilde{\lambda}_i)^2 \int I_{(t_{i-1}, t_i]} dm \\ &= \sum_{i=1}^p (\lambda_i - \tilde{\lambda}_i)^2 (t_i - t_{i-1}) \end{aligned}$$

因为阶梯函数的稠密性, 可以把  $G$  的定义域延拓到整个  $L^2(\mathbb{R}_+)$  上, 从而成为一个高斯白噪声, 且满足  $G(I_{[0,t]}) = X_t$ , 因此  $X$  是一个准布朗运动。□

#### 推论 4.2.1

$B = (B_t)_{t \geq 0}$  是准布朗运动, 则对于  $\forall 0 \leq t_0 < t_1 < \cdots < t_n$ , 向量  $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$  的概率密度函数为

$$f(x_1, x_2, \cdots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{t_1(t_2 - t_1) \cdots (t_n - t_{n-1})}} \exp \left\{ - \sum_{i=1}^n \frac{y_i^2}{2(t_i - t_{i-1})} \right\}$$

其中  $t_0 = 0$ .

### 命题 4.2.2

$B = (B_t)_{t \geq 0}$  是准布朗运动,

(1).  $B_t^- = -B_t$ , 则  $(B_t^-)_{t \geq 0}$  是准布朗运动.

(2).  $\forall \lambda > 0$ ,  $B_t^{(\lambda)} = \lambda^{-1} B_{\lambda^2 t}$ , 则  $(B_t^{(\lambda)})_{t \geq 0}$  是准布朗运动.

(3).  $\forall s \geq 0$ ,  $\tilde{B}_t = B_{s+t} - B_s$ , 则  $(\tilde{B}_t)_{t \geq 0}$  是准布朗运动, 且与  $\mathcal{F}_s = \sigma(B_r, r \leq s)$  独立.

(3) 也叫布朗运动的马氏性.

**证明** 验证命题 4.2.1(4) 即可, 不再赘述. □

## 4.3 连续轨道与布朗运动

一般情况下, 对于实值的随机过程  $X = (X_t)_{t \geq 0}$ , 我们一般的研究角度是从时刻入手, 比如“截止某个时刻  $X_t$  达到了什么值”, 或者“首次达到什么值的时刻”。本小节我们从样本空间入手, 固定  $\omega$ , 考虑映射  $t \mapsto X_t(\omega)$ , 这是一个我们很熟悉的  $\mathbb{R}_+ \rightarrow \mathbb{R}$  的实函数, 它被称为“样本轨道”。注意, 我们如果想要讨论样本轨道的连续性, 必须得要求随机变量在一个度量空间上取值。

### 定义 4.3.1 (样本轨道、布朗运动)

度量空间  $(E, d)$  取其 Borel  $\sigma$ -域作为可测空间, 随机过程  $(X_t)_{t \geq 0}$  在  $E$  上取值, 固定  $\omega \in \Omega$ , 得到映射:

$$[0, +\infty] \rightarrow \mathbb{E}$$

$$t \mapsto X_t(\omega)$$

这个映射就被称为样本轨道 (Sample Path).

对于一个准布朗运动  $B = (B_t)_{t \geq 0}$ , 如果  $\forall \omega \in \Omega$ ,  $t \mapsto X_t(\omega)$  处处连续, 则称  $B$  为布朗运动. ♣

事实上, 每一个准布朗运动都可以通过某些轻微的“修改”让它的轨道处处连续, 使它成为布朗运动, 这是我们本小节主要介绍的内容。

首先我们要定义什么情况下认为两个随机过程是相同的。

### 定义 4.3.2 (修改、不可分辨)

两个随机过程  $X = (X_t)$  和  $\tilde{X} = (\tilde{X}_t)$ , 如果

$$\forall t, \mathbb{P}(X_t = \tilde{X}_t)$$

则称  $\tilde{X}$  是  $X$  的一个修改 (modification).

如果存在一个零测集  $N \subset \Omega$ , 使得  $\forall \omega \in \Omega - N, \forall t, \tilde{X}_t(\omega) = X_t(\omega)$ , 则称  $\tilde{X}$  与  $X$  不可分辨 (indistinguishable), 也能形式地表述为

$$\mathbb{P}(\forall t, \tilde{X}_t = X_t) = 1$$

但是  $\{\forall t, \tilde{X}_t = X_t\}$  可能不是可测集, 所以并不合适. ♣

两个随机过程如果是不可分辨的, 则认为它们是同一个随机过程。

**推论 4.3.1**

取指标集为  $\mathbb{R}$  的某个区间, 如果  $X$  和  $\tilde{X}$  的轨道处处连续 a.s., 则  $\tilde{X}$  是  $X$  的修改  $\Leftrightarrow \tilde{X}$  与  $X$  不可分辨。把轨道处处连续的条件更换为左连续或者右连续, 结论也成立。



**定理 4.3.1 (Kolmogorov 引理)**

随机过程  $X = (X_t)_{t \in I}$ , 其中  $I$  是  $\mathbb{R}$  上的区间, 且  $X$  在完备度量空间  $(E, d)$  上取值, 后者取其 Borel  $\sigma$ -域作为可测空间。若存在  $q, \varepsilon, C > 0$  使得  $\forall s, t \in I$ ,

$$\mathbb{E}[d(X_s, X_t)^q] \leq C \cdot |t - s|^{1+\varepsilon}$$

则存在  $X$  的修改  $\tilde{X}$ , 其轨道满足指数为  $\alpha$  的 Holder 连续性, 即: 固定  $\omega \in \Omega$ ,  $\alpha \in (0, \frac{\varepsilon}{q})$ , 存在一个有限的、仅与  $\alpha, \omega$  有关的常数  $C_\alpha(\omega)$  使得

$$\forall s, t \in I, d(\tilde{X}_s(\omega), \tilde{X}_t(\omega)) \leq C_\alpha(\omega) |t - s|^\alpha$$

并且, 在不可分辨意义下,  $\tilde{X}$  这样的修改是唯一的。

**证明** 不妨取  $I = [0, 1]$ , 固定  $\alpha \in (0, \frac{\varepsilon}{q})$ , 由条件可知,  $\forall a > 0, s, t \in I$ , 有

$$\mathbb{P}(d(X_s, X_t) \geq a) \leq a^{-q} \mathbb{E}[d(X_s, X_t)^q] \leq C a^{-q} |t - s|^{1+\varepsilon}$$

取  $s = (i-1) \cdot 2^{-n}, t = i \cdot 2^{-n}, i \in \{1, 2, \dots, 2^n\}$ , 以及  $a = 2^{-n\alpha}$ , 则得到

$$\mathbb{P}(d(X_{(i-1) \cdot 2^{-n}}, X_{i \cdot 2^{-n}}) \geq 2^{-n\alpha}) \leq C \cdot 2^{nq\alpha} \cdot 2^{-(1+\varepsilon)n}$$

对  $i$  求和, 得到:

$$\mathbb{P}\left(\bigcup_{i=1}^{2^n} \{d(X_{(i-1) \cdot 2^{-n}}, X_{i \cdot 2^{-n}}) \geq 2^{-n\alpha}\}\right) \leq 2^n \cdot C \cdot 2^{nq\alpha - (1+\varepsilon)n} = C \cdot 2^{-n(\varepsilon - q\alpha)}$$

再对  $n$  求和, 得到

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\bigcup_{i=1}^{2^n} \{d(X_{(i-1) \cdot 2^{-n}}, X_{i \cdot 2^{-n}}) \geq 2^{-n\alpha}\}\right) < +\infty$$

根据 B-C 引理,

$$\mathbb{P}(\exists n_0 \text{ s.t. } \forall n > n_0, \forall i \in \{1, 2, \dots, 2^n\}, d(X_{(i-1) \cdot 2^{-n}}, X_{i \cdot 2^{-n}}) \leq 2^{-n\alpha}) = 1$$

于是我们定义

$$K_\alpha(\omega) \stackrel{\text{def}}{=} \sup_{n \geq 1} \left( \sup_{1 \leq i \leq 2^n} \frac{d(X_{(i-1) \cdot 2^{-n}}, X_{i \cdot 2^{-n}})}{2^{-n\alpha}} \right)$$

则  $K_\alpha$  a.s. 有限, 接下来, 记

$$D = \{i \cdot 2^{-n} : i = 0, 1, \dots, 2^n - 1\} \subset [0, 1]$$

**引理 4.3.1**

取映射  $f : D \rightarrow E$ , 如果存在  $\alpha > 0, K < +\infty$  使得  $\forall n \in \mathbb{N}_+, i \in \{1, 2, \dots, 2^n - 1\}$  都有

$$d(f((i-1)2^{-n}), f(i \cdot 2^{-n})) \leq K \cdot 2^{-n\alpha}$$

则  $\forall s, t \in D$ , 都有

$$d(f(s), f(t)) \leq \frac{2K}{1 - 2^{-\alpha}} |t - s|^\alpha$$



于是在  $\{K_\alpha < +\infty\}$  上,  $\forall s, t \in D, d(X_s, X_t) \leq C_\alpha(\omega) |t - s|^\alpha$ , 其中  $C_\alpha = 2(1 - 2^{-\alpha})^{-1} K_\alpha(\omega)$ , 所以映射  $t \mapsto X_t(\omega)$  在  $D$  上是 Holder 连续的, 进而是一致连续的。因为  $(E, d)$  完备, 该映射可以被唯一地延拓为  $[0, 1]$

上的连续映射<sup>2</sup>:

$$t \mapsto \tilde{X}_t(\omega) = \begin{cases} \lim_{s \rightarrow t, s \in D} X_s(\omega) & , K_\alpha(\omega) < +\infty \\ x_0 & , K_\alpha(\omega) = +\infty \end{cases}$$

那么  $\tilde{X}$  的轨道是 Holder 连续的, 只需证明  $\tilde{X}$  是  $X$  的修改。固定  $t \in [0, 1]$ ,

$$\mathbb{P}(\lim_{s \rightarrow t} X_s = X_t) = 1$$

而  $\tilde{X}_t$  同样是  $X_s$  的 a.s. 极限, 所以  $X_t = \tilde{X}_t$  a.s. □

**推论 4.3.2**

考虑布朗运动  $B = (B_t)_{t \geq 0}$ , 则  $B$  存在一个修改  $\tilde{B}$  满足其轨道是连续的, 进一步地, 还满足指数为  $\frac{1}{2} - \delta, \forall \delta \in (0, \frac{1}{2})$  的 Holder 连续性。 ♥

**证明** 取  $U \sim N(0, 1)$ , 则  $B_t - B_s \stackrel{d}{=} \sqrt{t-s}U$ , 于是  $\forall q > 0$ ,

$$\mathbb{E}[|B_t - B_s|^q] = (t-s)^{\frac{q}{2}} \mathbb{E}[|U|^q] = C_q(t-s)^{\frac{q}{2}}$$

其中  $C_q = \mathbb{E}[|U|^q] < +\infty$ , 取  $q > 2$ , 可知  $B$  存在修改  $\tilde{B}$ , 其轨道满足指数为  $\alpha, \forall \alpha < \frac{q-2}{2q}$  的局部 Holder 连续性, 取  $q$  充分大则得到结论。 □

在本章的剩余内容中, 我们默认一个随机过程只要满足准布朗运动的定义, 就把它视为修改过后得到的布朗运动。

### 4.4 布朗运动路径的性质

接下来, 记

$$\mathcal{F}_t = \sigma(B_s, s \leq t), \mathcal{F}_{0+} = \bigcap_{s>0} \mathcal{F}_s$$

**定理 4.4.1 (Blumenthal 0-1 律)**

$\forall A \in \mathcal{F}_{0+}, \mathbb{P}(A) = 1$  或  $0$ .

**证明** 取  $0 < t_1 < t_2 < \dots < t_k, g: \mathbb{R}^k \rightarrow \mathbb{R}$  为有界连续函数, 对于  $A \in \mathcal{F}_{0+}$ , 由连续性可知

$$\mathbb{E}[I_A \cdot g(B_{t_1}, \dots, B_{t_k})] = \lim_{\varepsilon \rightarrow 0+} \mathbb{E}[I_A \cdot g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)]$$

设  $0 < \varepsilon < t_1$ , 则  $B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon$  都与  $\mathcal{F}_\varepsilon$  独立, 进而与  $\mathcal{F}_{0+}$  独立, 于是

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0+} \mathbb{E}[I_A \cdot g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)] &= \lim_{\varepsilon \rightarrow 0+} \mathbb{P}(A) \cdot \mathbb{E}[g(B_{t_1} - B_\varepsilon, \dots, B_{t_k} - B_\varepsilon)] \\ &= \mathbb{P}(A) \cdot \mathbb{E}[I_A \cdot g(B_{t_1}, \dots, B_{t_k})] \end{aligned}$$

由  $t_1, \dots, t_k$  的任意性可知,  $\mathcal{F}_{0+}$  与  $\sigma(B_t, t > 0)$  独立, 又因为  $B_t \xrightarrow{P.W.} B_0$  as  $t \rightarrow 0, \sigma(B_t, t > 0) = \sigma(B_t, t \geq 0)$ , 而  $\mathcal{F}_{0+} \subset \sigma(B_t, t \geq 0)$ , 这说明  $\mathcal{F}_{0+}$  和自己独立, 那  $\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2 \Rightarrow \mathbb{P}(A) = 1$  或  $0$ . □

**命题 4.4.1**

对于布朗运动  $B = (B_t)_{t \geq 0}$ ,

- (1).  $\forall \varepsilon > 0, \sup_{0 \leq s \leq \varepsilon} B_s > 0, \inf_{0 \leq s \leq \varepsilon} B_s < 0, \text{ a.s.}$
- (2).  $\forall s \in \mathbb{R},$

$$\limsup_{t \rightarrow \infty} B_t = +\infty, \liminf_{t \rightarrow \infty} B_t = -\infty$$

<sup>2</sup>想到了实分析的 Tietze 延拓定理。

进而如果令  $T_a = \inf\{t \geq 0 : B_t = a\}$ , 则可得  $T_a < +\infty$  a.s.

**注** 这里关于  $\sup_{0 \leq s \leq \varepsilon} B_s$ 、 $\limsup_{t \rightarrow \infty} B_t$  无法确保可测性的问题, 书上原文说的不是很清楚。笔者在题目 4.2 的注记中尝试解释了这个问题, 建议读者一起阅读。

**证明** (1). 取一列严格单调递减到 0 的实数  $(\varepsilon_p)_{p \in \mathbb{N}_+}$ , 令

$$A = \bigcap_{p=1}^{\infty} \left\{ \sup_{0 \leq s \leq \varepsilon_p} B_s > 0 \right\}$$

则  $A \in \mathcal{F}_{0+}$ , 因为  $\forall s > 0$ , 可以取  $p_0$  使得  $\varepsilon_{p_0} < s$ , 那么

$$A = \bigcap_{p=p_0}^{\infty} \left\{ \sup_{0 \leq s \leq \varepsilon_p} B_s > 0 \right\} \in \mathcal{F}_s$$

由  $s$  任意性可得。而另一方面,

$$\mathbb{P}(A) = \lim_{p \rightarrow \infty} \mathbb{P} \left( \sup_{0 \leq s \leq \varepsilon_p} B_s > 0 \right)$$

而且  $\mathbb{P} \left( \sup_{0 \leq s \leq \varepsilon_p} B_s > 0 \right) \geq \mathbb{P}(B_{\varepsilon_p} > 0) = \frac{1}{2}$ , 根据 0-1 律 (定理 4.4.1), 只能  $\mathbb{P}(A) = 1$ , 注意右边是一个单调递减的极限, 因此每一项都等于 1, 也就是

$$\forall \varepsilon > 0, \sup_{0 \leq s \leq \varepsilon} B_s > 0 \text{ a.s.}$$

inf 的情况同理, 考虑  $-B$  即可。

(2). 根据上一题的结论, 我们得到:

$$1 = \mathbb{P} \left( \sup_{0 \leq s \leq 1} B_s > 0 \right) = \lim_{\delta \rightarrow 0+} \mathbb{P} \left( \sup_{0 \leq s \leq 1} B_s > \delta \right)$$

我们考虑用另一个布朗运动  $B'_s = M^{-1} \delta B_{M^2 \delta^{-2} s}$ , 替换最右侧中的  $B_s$ , 得到

$$\begin{aligned} 1 &= \lim_{\delta \rightarrow 0+} \mathbb{P} \left( \sup_{0 \leq s \leq 1} M^{-1} \delta B_{M^2 \delta^{-2} s} > \delta \right) \\ &= \lim_{\delta \rightarrow 0+} \mathbb{P} \left( \sup_{0 \leq s \leq M^2 \delta^{-2}} B_s > M \right) \\ &= \mathbb{P} \left( \sup_{s \geq 0} B_s > M \right) \end{aligned}$$

根据  $M$  任意性可知,  $\limsup_{t \rightarrow \infty} B_t = +\infty$  a.s.,  $\liminf$  的情形同理, 取  $-B$  即可。□

从直观上看, 这体现了布朗运动不稳定且混乱, 即不可能在某一处停留, 并且能到达所有位置。

#### 推论 4.4.1

映射  $r \mapsto B_t(w)$  在任何区间上不单调 a.s.

**证明**  $\forall q \in \mathbb{Q}_+, \forall \varepsilon > 0$ , 考虑布朗运动  $B_{q+t} - B_q$ , 应用命题 4.4.1, 可得

$$\sup_{q \leq t \leq q+\varepsilon} B_t > B_q, \quad \inf_{q \leq t \leq q+\varepsilon} B_t < B_q,$$

所以  $B_t$  在  $[q, q+\varepsilon]$  上不单调, 由  $q, \varepsilon$  任意性得证。□

#### 命题 4.4.2

$0 = t_0^n < t_1^n < \dots < t_{p_n}^n = t$  是  $[0, t]$  的一系列划分, 并且当  $n \rightarrow \infty$  时,

$$\sup_{1 \leq i \leq p_n} (t_i^n - t_{i-1}^n) \rightarrow 0$$

那么  $n \rightarrow \infty$  时,

$$\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 \xrightarrow{L^2} t$$

#### 推论 4.4.2

映射  $t \mapsto B_t(w)$  在任何区间上有无穷变差。

**证明** 不妨选取  $[0, t]$ , 取命题 4.4.2 中的划分,

$$\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 \leq \left( \sup_{1 \leq i \leq p_n} |B_{t_i^n} - B_{t_{i-1}^n}| \right) \times \sum_{i=1}^{p_n} |B_{t_i^n} - B_{t_{i-1}^n}|$$

根据轨道的连续性, 右式第一项  $\rightarrow 0$ , 但是左边  $\rightarrow t$ , 所以右式第二项一定有

$$\sum_{i=1}^{p_n} |B_{t_i^n} - B_{t_{i-1}^n}| \rightarrow +\infty$$

这正是  $t \mapsto B_t(w)$  在此划分下的变差。 □

无穷变差也意味着处处不可导, 这似乎说明布朗运动的“速度”是无穷大? 解决这一悖论的办法, 本章最后一小节有相关介绍。

## 4.5 布朗运动的强马氏性

回顾记号: 对于布朗运动  $B = (B_t)_{t \geq 0}$ ,

$$\mathcal{F}_t = \sigma(B_s, s \leq t), \quad \mathcal{F}_\infty = \sigma(B_s, s \geq 0)$$

#### 定义 4.5.1

随机变量  $T$  在  $[0, +\infty]$  上取值, 如果  $\forall t \geq 0, \{T \leq t\} \in \mathcal{F}_t$ , 则称  $T$  为关于  $B$  的停时。

#### 例 4.5.1.

如果令

$$T_a = \inf\{t \geq 0 : B_t = a\}$$

则  $T_a$  就是一个停时, 因为:

$$\{T_a \leq t\} = \{\exists s \in [0, t], B_s = a\} = \left\{ \inf_{s \in [0, t]} |B_s - a| = 0 \right\} \in \mathcal{F}_t$$

但是,  $T = \sup\{s \leq 1 : B_s = 0\}$  不是停时。

#### 定义 4.5.2

如果  $T$  是停时, 定义:

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}$$

称为  $T$  前  $\sigma$ -域。

## 例 4.5.2.

首先  $T$  本身就是  $\mathcal{F}_T$ -可测的, 因为:  $\forall s \geq 0, \forall t \geq 0,$

$$\{T \leq s\} \cap \{T \leq t\} = \{T \leq s \wedge t\} \in \mathcal{F}_{s \wedge t} \subset \mathcal{F}_t$$

那么,  $B_s I_{s \leq T}$  是  $\mathcal{F}_T$ -可测的, 因为:

$$\{B_s I_{s \leq T} \in A\} \cap \{T \leq t\} = \begin{cases} \emptyset & , t < s \\ \{B_s \in A\} \cap \{s \leq T \leq t\} & , t \geq s \end{cases} \in \mathcal{F}_t$$

从而  $I_{\{T < +\infty\}} B_T$  也是  $\mathcal{F}_T$ -可测的, 因为:

$$I_{\{T < +\infty\}} B_T = \lim_{n \rightarrow \infty} \sum_{i=0}^{\infty} I_{\{i \cdot 2^{-n} \leq T \leq (i+1) \cdot 2^{-n}\}} B_{i \cdot 2^{-n}}$$

## 定理 4.5.1 (强马氏性)

$T$  是停时, 且  $\mathbb{P}(T < +\infty) = 1$ , 令

$$B_t^{(T)} = I_{\{T < +\infty\}} (B_{T+t} - B_T)$$

则  $B^{(T)} = (B_t^{(T)})_{t \geq 0}$  是布朗运动, 并且和  $\mathcal{F}_T$  独立。

**证明** 取  $A \in \mathcal{F}_T, 0 \leq t_1 < \dots < t_p, F: \mathbb{R}^p \rightarrow \mathbb{R}_+$  有界连续, Claim:

$$\mathbb{E}[I_A \cdot F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)})] = \mathbb{P}(A) \cdot \mathbb{E}[F(B_{t_1}, \dots, B_{t_p})]$$

然后考虑  $A = \Omega$ , 说明  $B^{(T)}$  和  $B$  有着相同的有限维分布, 因此也是布朗运动。同时, 由  $t_1, \dots, t_p$  的任意性可知  $B^{(T)}$  和  $\mathcal{F}_T$  独立。

下面来证明 Claim, 对于  $\forall n \in \mathbb{N}_+, t \geq 0$ , 记

$$[t]_n = \min \{x = k \cdot 2^{-n} \geq t | k \in \mathbb{N}_+\}$$

即  $[t, +\infty)$  上最小的  $k \cdot 2^{-n}, k \in \mathbb{N}_+$ , 并定义  $[\infty]_n = \infty$ , 注意到

$$F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)}) = \lim_{n \rightarrow \infty} F(B_{[t_1]_n}^{(T)}, \dots, B_{[t_p]_n}^{(T)}) \text{ a.s.}$$

因此由 DCT 可得

$$\begin{aligned} & \mathbb{E}[I_A \cdot F(B_{t_1}^{(T)}, \dots, B_{t_p}^{(T)})] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}[I_A \cdot F(B_{[t_1]_n}^{(T)}, \dots, B_{[t_p]_n}^{(T)})] \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^{\infty} \mathbb{E}[I_A \cdot I_{\{(k-1)2^{-n} < T \leq k \cdot 2^{-n}\}} F(B_{k \cdot 2^{-n} + t_1} - B_{k \cdot 2^{-n}}, \dots, B_{k \cdot 2^{-n} + t_p} - B_{k \cdot 2^{-n}})] \end{aligned}$$

因为  $A \in \mathcal{F}_T$ , 所以

$$A \cap \{(k-1)2^{-n} < T \leq k \cdot 2^{-n}\} = (A \cap \{T \leq k \cdot 2^{-n}\}) \cap \{T \leq (k-1)2^{-n}\}^c \in \mathcal{F}_{k \cdot 2^{-n}}$$

因此

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}[I_{A \cap \{(k-1)2^{-n} < T \leq k \cdot 2^{-n}\}} \cdot F(B_{k \cdot 2^{-n} + t_1} - B_{k \cdot 2^{-n}}, \dots, B_{k \cdot 2^{-n} + t_p} - B_{k \cdot 2^{-n}})] \\ &= \mathbb{P}(A \cap \{(k-1)2^{-n} < T \leq k \cdot 2^{-n}\}) \cdot \mathbb{E}[F(B_{t_1}, \dots, B_{t_p})] \end{aligned}$$

对  $k$  求和即得到目标结论。  $\square$

## 定理 4.5.2 (反射原理)

$\forall t \geq 0$ , 令  $S_t = \sup_{s \leq t} B_s$ , 设常数  $a \geq 0$ ,  $b \leq a$ , 则

$$\mathbb{P}(S_t \geq a, B_t \leq b) = \mathbb{P}(B_t \geq 2a - b)$$

进而  $S_t$  与  $|B_t|$  有着相同的分布。

**证明** 取停时

$$T_a = \inf\{t \geq 0 : B_t = a\}$$

我们在命题 4.4.1 证明过  $T_a < +\infty$  a.s., 利用强马氏性,

$$\mathbb{P}(S_t \geq a, B_t \leq b) = \mathbb{P}(T_a \leq t, B_t \leq b) = \mathbb{P}(T_a \leq t, B_{t-T_a}^{(T_a)} \leq b - a) = (\star)$$

接下来, 我们记  $B' = B^{(T_a)}$ , 则  $B'$  也是布朗运动且与  $\mathcal{F}_{T_a}$  无关, 即与  $T_a$  独立, 而且因为  $B'$  与  $-B'$  同分布, 所以

$$(\star) = \mathbb{P}(T_a \leq t, -B_{t-T_a}^{(T_a)} \leq b - a) = \mathbb{P}(T_a \leq t, B_t \geq 2a - b) = \mathbb{P}(B_t \geq 2a - b)$$

因为  $B_t \geq 2a - b \geq a$ , 则在  $t$  之前就达到了  $a$ , 即  $T_a \leq t$ . 所以,

$$\mathbb{P}(S_t \geq a) = \mathbb{P}(S_t \geq a, B_t \geq a) + \mathbb{P}(S_t \geq a, B_t \leq a) = 2\mathbb{P}(B_t \geq a) = \mathbb{P}(|B_t| \geq a)$$

所以  $S_t$  与  $|B_t|$  有着相同的分布。 □

从直观上看,

$$\begin{aligned} S_t \geq a, B_t \leq b &\Leftrightarrow \text{在 } [0, t] \text{ 内曾达到了 } a, \text{ 而最终在 } t \text{ 时刻落回了 } b \text{ 的下方} \\ &\Leftrightarrow T_a \leq t, \text{ 在 } [T_a, t] \text{ 下降了至少 } a - b \\ &\Leftrightarrow T_a \leq t, \text{ 在 } [T_a, t] \text{ 上升了至少 } a - b \\ &\Leftrightarrow \text{在 } t \text{ 时刻达到了 } 2a - b \text{ 的上方, 即 } B_t \geq 2a - b \end{aligned}$$

即以  $T_a$  为分界点,  $T_a$  之后的轨迹上下翻转。

## 推论 4.5.1

$\forall a > 0$ ,  $T_a$  与  $a^2 B_1^{-2}$  同分布, 从而密度函数为:

$$f(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left\{-\frac{a^2}{2t}\right\} I_{\{t>0\}}$$

并且  $\mathbb{E}[T_a] = +\infty$ . ♥

**证明** 考虑

$$\mathbb{P}(T_a \leq t) = \mathbb{P}(S_t \geq a) = \mathbb{P}(|B_t| \geq a) = \mathbb{P}(B_t^2 \geq a^2) = \mathbb{P}(tB_1^2 \geq a^2) = \mathbb{P}\left(\frac{a^2}{B_1^2} \leq t\right)$$

然后利用  $B_1 \sim N(0, 1)$  即可得到结论。 □

## 4.6 习题

本小节用到的结论:

## 引理 4.6.1 (Fatou's Lemma)

$\{X_n\}$  是概率空间上的一列非负随机变量, 则有:

$$\mathbb{E}[\liminf_{n \rightarrow \infty} X_n] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n]$$



## 引理 4.6.2 (Borel-Cantelli Lemma)

概率空间上的一列事件  $\{A_n\}$  满足:

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < +\infty$$

那么

$$\mathbb{P}(\text{存在无数多个 } n, A_n \text{ 成立}) = 0$$



## 引理 4.6.3 (强大数定律)

$\{X_n, n \in \mathbb{N}_+\}$  独立同分布, 二阶矩有限, 期望为  $\mu$ , 则

$$\frac{1}{n} \sum_{k=1}^n X_k \rightarrow \mu \text{ a.s.}$$



## 引理 4.6.4 (Doob-Kolmogorov Inequality)

$\{X_n, n \in \mathbb{N}_+\}$  关于  $(\mathcal{F}_n)_{n \in \mathbb{N}_+}$  是鞅,  $\sup_{n \geq 1} \mathbb{E}[X_n^2] < M < +\infty$ , 则  $\forall \varepsilon > 0$ ,

$$\mathbb{P}(\max_{1 \leq i \leq n} |X_i| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}[X_n^2]$$



**证明** 由推论 1.2.1 知  $X_n^2$  是一个下鞅, 再利用 Doob 最大值不等式 (定理 1.6.1) 即可。题目 1.10 是一个更强的结论。□

下面这道题展示了如何证明一个过程是布朗运动。我们一般利用命题 4.2.1 来证明一个过程是准布朗运动, 而难点一般在于证明轨道连续性。

## 题目 4.1. Le gall (Exercise 2.25)

$B = (B_t)_{t \geq 0}$  是一个布朗运动, 设:

$$W_t = tB_{\frac{1}{t}}, t > 0$$

且  $W_0 = 0$ , 证明:  $(W_t)_{t \geq 0}$  在不可分辨意义下是一个布朗运动。

**注** 什么叫“在不可分辨意义下是一个布朗运动”? 我们曾提到过, 如果两个随机过程不可分辨, 我们就认为它们是同一个随机过程。题目这句话的意思我们只需要证明  $\mathbb{P}(W \text{ 的轨道处处连续}) = 1$  即可。

**证明** 不难看出,  $W_t \sim N(0, t)$ , 且  $\{W_t\}$  中任意元素的有限线性组合仍然服从正态分布, 因此是 (中心) 高斯过程, 并且  $\mathbb{E}[W_t W_0] = 0$ ,

$$\mathbb{E}[W_s W_t] = st \mathbb{E}[B_{\frac{1}{s}} B_{\frac{1}{t}}] = st \left( \frac{1}{s} \wedge \frac{1}{t} \right) = s \wedge t$$

由命题 4.2.1(2) 可知  $W$  是一个准布朗运动。

注意到  $W$  的轨道在  $t > 0$  是连续的 (因为  $B$  的轨道连续), 我们只需证明:

$$\lim_{t \rightarrow 0^+} W_t = \lim_{t \rightarrow \infty} \frac{B_t}{t} = 0 \text{ a.s.}$$

考虑鞅  $(B_{k+1} - B_k)_{k \in \mathbb{N}_+}$ , 这是一列独立同分布的随机变量, 由强大数定律可知

$$\frac{1}{n} \sum_{k=1}^n (B_{k+1} - B_k) = \frac{B_n}{n} \rightarrow 0 \text{ a.s. (as } n \rightarrow \infty)$$

取  $n, m \geq 0$ , 注意到  $\{X_k = B_{n+k \cdot 2^{-m}} - B_n, k = 0, 1, \dots, 2^m\}$  是一个鞅, 由 Doob-Kolmogorov 不等式,

$$\mathbb{P}\left(\max_{0 \leq k \leq 2^m} |B_{n+k \cdot 2^{-m}} - B_n| \geq n^{\frac{2}{3}}\right) \leq \frac{1}{n^{\frac{4}{3}}} \mathbb{E}[(B_{n+1} - B_n)^2] = \frac{1}{n^{\frac{4}{3}}}$$

令  $m \rightarrow \infty$ , 并由  $B$  的轨道连续性可知,

$$\mathbb{P}\left(\sup_{t \in [n, n+1]} |B_t - B_n| \leq n^{\frac{2}{3}}\right) \leq \frac{1}{n^{\frac{4}{3}}}$$

设  $A_n = \left\{ \sup_{t \in [n, n+1]} |B_t - B_n| \leq n^{\frac{2}{3}} \right\}$ , 则  $\sum \mathbb{P}(A_n) < +\infty$ , 由 B-C 引理可知,

$$\mathbb{P}(A_n \text{ 成立的 } n \text{ 的个数有限}) = 1$$

也就是存在  $N$ , 使得  $n > N$  时, 所有的  $A_n$  都不成立 a.s., 即

$$|B_t - B_n| \leq n^{\frac{2}{3}}, \forall t \in [n, n+1] \text{ a.s.}$$

得到

$$\left| \frac{B_t}{t} \right| \leq \frac{|B_n| + |B_n - B_t|}{t} \leq \frac{|B_n| + n^{\frac{2}{3}}}{n} = \frac{|B_n|}{n} + n^{-\frac{1}{3}} \rightarrow 0 \text{ a.s. (as } n \rightarrow \infty)$$

$t \rightarrow \infty$  时  $n \rightarrow \infty$ , 结论得证. □

下面这道题的思路类似于命题 4.4.1 的证明过程, 展示了 0-1 律的应用。

**题目 4.2. Le gall(Exercise 2.29)**

$B = (B_t)_{t \geq 0}$  是布朗运动, 证明: a.s. 成立

$$\limsup_{t \searrow 0} \frac{B_t}{\sqrt{t}} = +\infty$$

$$\liminf_{t \searrow 0} \frac{B_t}{\sqrt{t}} = -\infty$$

**注** 这里的  $\limsup_{t \searrow 0}$  不是可数运算, 不能保证可测性, 但是利用布朗运动轨道的连续性, 可知

$$\limsup_{t \searrow 0} \frac{B_t}{\sqrt{t}} = \lim_{\varepsilon \searrow 0} \sup_{t \in (0, \varepsilon]} \frac{B_t}{\sqrt{t}} = \lim_{\varepsilon \searrow 0} \sup_{t \in (0, \varepsilon] \cap \mathbb{Q}} \frac{B_t}{\sqrt{t}}$$

然后我们任取一列单调递减趋于 0 的实数  $\{\varepsilon_p, p \in \mathbb{N}_+\}$ , 并考虑

$$\lim_{p \rightarrow \infty} \sup_{t \in (0, \varepsilon_p] \cap \mathbb{Q}} \frac{B_t}{\sqrt{t}}$$

如果我们能够证明总是上式  $= +\infty$  a.s., 与  $\varepsilon_p$  的选取无关, 那么由  $\{\varepsilon_p\}$  的任意性即可得到结论, 这用到了函数极限里的小结论:

**引理 4.6.5**

任取单调递增的数列  $a_n$  且  $a_n \searrow a$ , 都有  $\lim_{n \rightarrow \infty} f(a_n) = y$ , 那么  $\lim_{x \rightarrow a^+} f(x) = y$ . ♥

反证法易证。

**证明** 取  $M > 0, \varepsilon > 0$ , 令

$$A_\varepsilon = \left\{ \sup_{t \in (0, \varepsilon]} \frac{B_t}{\sqrt{t}} \geq M \right\} = \left\{ \sup_{t \in (0, \varepsilon] \cap \mathbb{Q}} \frac{B_t}{\sqrt{t}} \geq M \right\} \in \mathcal{F}_\varepsilon$$

我们任取一列单调递减趋于 0 的实数  $\{\varepsilon_p, p \in \mathbb{N}_+\}$ , 则

$$A = \bigcap_{p=1}^{\infty} A_{\varepsilon_p} = \left\{ \limsup_{p \rightarrow \infty} \sup_{t \in (0, \varepsilon_p]} \frac{B_t}{\sqrt{t}} \geq M \right\} \in \mathcal{F}_{0+}$$

于是  $\mathbb{P}(A) = 0$  或 1, 我们希望证明  $\mathbb{P}(A) = 1$ , 所以要估计它的下界, 注意到

$$\limsup_{p \rightarrow \infty} \frac{B_{\varepsilon_p}}{\sqrt{\varepsilon_p}} \geq M = \lim_{p \rightarrow \infty} \sup_{k \geq p} \frac{B_{\varepsilon_k}}{\sqrt{\varepsilon_k}} \geq M \Rightarrow \lim_{p \rightarrow \infty} \sup_{t \in (0, \varepsilon_p]} \frac{B_t}{\sqrt{t}} \geq M$$

所以

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}\left(\limsup_{p \rightarrow \infty} \sup_{t \in (0, \varepsilon_p]} \frac{B_t}{\sqrt{t}} \geq M\right) \geq \mathbb{P}\left(\limsup_{p \rightarrow \infty} \frac{B_{\varepsilon_p}}{\sqrt{\varepsilon_p}} \geq M\right) \\ &= \mathbb{P}\left(\limsup_{p \rightarrow \infty} \left\{ \frac{B_{\varepsilon_p}}{\sqrt{\varepsilon_p}} \geq M \right\}\right) \\ &\geq \limsup_{p \rightarrow \infty} \mathbb{P}\left(\frac{B_{\varepsilon_p}}{\sqrt{\varepsilon_p}} \geq M\right) \quad \text{由 Fatou 引理} \\ &= \int_M^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx > 0 \end{aligned}$$

所以我们证明了:

$$\lim_{p \rightarrow \infty} \sup_{t \in (0, \varepsilon_p]} \frac{B_t}{\sqrt{t}} \geq M \text{ a.s.}$$

那么由  $M$  的任意性可知

$$\lim_{p \rightarrow \infty} \sup_{t \in (0, \varepsilon_p]} \frac{B_t}{\sqrt{t}} = +\infty \text{ a.s.}$$

□

下面这道题则考查的是命题 4.4.1 和强马氏性的应用。

### 题目 4.3. Le gall(Exercise 2.30)

$B = (B_t)_{t \geq 0}$  是布朗运动,  $H = \{t \in [0, 1] : B_t = 0\}$ , 证明以下事实 a.s. 成立:  $H$  是  $[0, 1]$  上的紧集、无孤立点、Lebesgue 测度为零。

**证明** 闭区间上连续函数的零点集一定是闭集, 所以  $H$  是 (有界) 闭集, 进而是紧集。

用  $m$  表示 Lebesgue 测度, 那么

$$\begin{aligned} \mathbb{E}[m(H)] &= \int_{\Omega} m(H) d\mathbb{P} \\ &= \int_{\Omega} \int_{[0, 1]} I_{t \in [0, 1]: B_t = 0} dm d\mathbb{P} \\ &= \int_{[0, 1]} \int_{\Omega} I_{t \in [0, 1], \omega \in \Omega: B_t(\omega) = 0} d\mathbb{P} dm \\ &= \int_{[0, 1]} \mathbb{P}(B_t = 0) dt = 0 \end{aligned}$$

所以  $m(H) = 0$  a.s.

最后来证明  $H$  没有孤立点, 我们的思路是这样的:

1. 对于  $q \in \mathbb{Q}$ , 定义  $T_q = \inf\{t \in [q, 1] : B_t = 0\}$ , 即 “ $q$  时刻开始的第一个零点”, 这是一个停时。
2. 证明  $\forall \varepsilon > 0$ ,  $T_q$  右侧  $\varepsilon$ -邻域内存在下一个零点, 这说明  $T_q$  不是孤立点。
3.  $T_q$  不一定是所有的零点, 任取一个其他的零点  $t$ , 取一列有理数  $q_n \nearrow t$ , 则  $q_n \leq T_{q_n} < t$ , 这说明  $t$  不是孤立点, 所以  $H$  没有孤立点。

回顾命题 4.4.1(1), 我们知道布朗运动的零点右侧任意小邻域内有正有负, 而轨迹是连续的, 所以肯定存在零点, 我们利用这一点来证明 2.

根据强马氏性,  $B_t^{(q)} = B_{T_q+t} - B_{T_q} = B_{T_q+t}$  是一个布朗运动, 根据命题 4.4.1(1),  $\forall \varepsilon \in (0, 1-q)$  有

$$\mathbb{P}\left(\sup_{t \in [0, \varepsilon]} B_{T_q+t} > 0, \inf_{t \in [0, \varepsilon]} B_{T_q+t} < 0\right) = 1$$

即

$$\mathbb{P}\left(\sup_{t \in [T_q, T_q+\varepsilon]} B_t > 0, \inf_{t \in [T_q, T_q+\varepsilon]} B_t < 0\right) = 1$$

即  $B_t$  在  $(T_q, T_q + \varepsilon]$  上有零点 (a.s.), 这说明  $T_q$  不是孤立点 (a.s.). □

最后放一道去年期末题, 比较简单。

#### 题目 4.4. 2023SPFinal.2

$B = (B_t)_{t \geq 0}$  是布朗运动, 求证:

$$\lim_{n \rightarrow \infty} \frac{\sup_{t \in [n, n+1]} B_t}{n} = 0 \text{ a.s.}$$

提示: 可能会用到

$$S_t = \sup_{0 \leq s \leq t} B_s$$

与  $|B_t|$  同分布。

**证明** 稍作变换:

$$\frac{\sup_{t \in [n, n+1]} B_t}{n} = \frac{\sup_{t \in [n, n+1]} B_t - B_n}{n} + \frac{B_n}{n} = \frac{\sup_{t \in [0, 1]} B_t}{n} + \frac{B_n}{n}$$

因为  $\sup_{t \in [0, 1]} B_t \stackrel{d}{=} |B_1|$ , 所以第一项  $\rightarrow 0$ , 第二项我们在题目 4.1 提到过, 大数定律得到  $\rightarrow 0$ . □

## 4.7 随机积分介绍 \*

### 4.7.1 定义

对于一个随机过程  $A = \{A_t, t \geq 0\}$ , 如果轨道  $t \mapsto A_t(\omega)$  a.s. 有界变差, 那么对于以下可测函数:

$$f : ([0, +\infty) \times \Omega, \mathcal{B}[0, +\infty) \otimes \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{R})$$

我们可以利用 Lebesgue 积分定义:

$$\left( \int_0^t f_s dA_s \right) (\omega) \stackrel{\text{def}}{=} \int_0^t f_\omega(s) dA_\omega(s)$$

但是, 我们在推论 4.4.2 曾提到过, 布朗运动的轨道没有有界变差, 因此不能直接采用 Lebesgue 积分。这里的记号  $f_t = f(t, \cdot)$ ,  $A_\omega = A(\cdot, \omega)$ , 实分析里乘积测度那一章节把这个叫做函数的截面。

我们将按照以下思路逐步定义出 (有限区间  $[0, t]$  上) 布朗运动的积分。

Step 1. 对于  $[0, t]$  上的阶梯函数: 取分割  $\Delta : 0 = t_0 < t_1 < t_2 < \cdots < t_n = t$ ,

$$f(s) = \sum_{j=1}^n x_{j-1} I_{\{t_{j-1} < s \leq t_j\}}$$

那么我们定义:

$$\int_0^t f(s) dB_s \stackrel{\text{def}}{=} \sum_{j=1}^n f_{j-1} (B_{t_j} - B_{t_{j-1}})$$

Step 2. 记  $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$ , 则

$$\int_0^t f(s) dB_s \in \mathcal{F}_t$$

Step 3. 可直接拆开计算验证以下两个式子成立:

$$\begin{aligned} \mathbb{E} \left[ \int_0^t f(s) dB_s \right] &= 0 \\ \mathbb{E} \left[ \left( \int_0^t f(s) dB_s \right)^2 \right] &= \int_0^t |f(s)|^2 ds \end{aligned}$$

Step 4. 如果  $g$  也是阶梯函数 (不必与  $f$  同分割), 则  $f - g$  也是一个阶梯函数, 从而

$$\mathbb{E} \left[ \left| \int_0^t f(s) dB_s - \int_0^t g(s) dB_s \right|^2 \right] = \int_0^t |f(s) - g(s)|^2 ds$$

Step 5. 阶梯函数在  $L^2[0, t]$  上是稠密的, 所以对于  $t_0 \in [0, t]$ , 任取  $h \in L^2[0, t_0]$ , 都存在一列阶梯函数  $f_n$  满足

$$\lim_{n \rightarrow \infty} \int_0^{t_0} |f(s) - f_n(s)|^2 ds = 0$$

Step 6. 由 Step 4, 可知

$$\mathbb{E} \left[ \left| \int_0^t f_n(s) dB_s - \int_0^t g_m(s) dB_s \right|^2 \right] = \int_0^t |f_n(s) - f_m(s)|^2 ds \rightarrow 0 \text{ as } n \rightarrow \infty$$

因此,

$$\left\{ \int_0^t f_n(s) dB_s, n \in \mathbb{N} \right\}$$

是  $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$  中的柯西列。因此, 存在平方可积的 r.v.  $U$  满足:

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left| \int_0^t f_n(s) dB_s - U \right|^2 \right] = 0$$

于是我们定义:

$$\int_0^t h(s) dB_s \stackrel{\text{def}}{=} U = \int_0^t f_n(s) dB_s \text{ 的 } L^2 \text{ 极限}$$

4.7.2  $L^2$  可积实函数对 B.M. 积分的性质

我们本小节先介绍对于  $f \in L^2[0, T]$  积分的性质。给定区间  $[0, T]$ , 对于  $t \in [0, T]$ , 记  $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$ ,

$$X_t \stackrel{\text{def}}{=} \int_0^t f(s) dB_s$$

## 命题 4.7.1

$$\mathbb{E}[|X_t|^2] = \int_0^t [f(s)]^2 ds$$

## 命题 4.7.2

对于  $\forall g \in L^2[0, T]$ ,

$$\mathbb{E} \left[ \int_0^t f(s) dB_s \int_0^t g(s) dB_s \right] = \int_0^t f(s)g(s) ds$$

## 命题 4.7.3

$$\int_0^t f(s) dB_s \sim N(0, \int_0^t [f(s)]^2 ds)$$

## 命题 4.7.4

$\{X_t, t \in [0, T]\}$  关于  $t$  连续。

## 命题 4.7.5

$\{X_t, t \in [0, T]\}$  关于  $\{\mathcal{F}_t, t \in [0, T]\}$  是连续平方可鞅, 从而  $\mathbb{E}[X_t] = \mathbb{E}[X_0] = 0$ 。

## 命题 4.7.6

任取分割  $0 = t_0 < t_1 < \cdots < t_n = T$ , 离散过程  $\{X_{t_i}, i = 0, \cdots, n\}$  关于  $\{\mathcal{F}_{t_i}\}$  为鞅。

## 4.7.3 随机过程对 B.M. 积分的定义

我们在第一小节中, 仅仅定义了  $h \in L^2[0, t]$  关于布朗运动  $(B_s)$  的积分, 但  $h$  是与  $\omega$  无关的实函数, 所以最后我们本小节进一步推广到满足以下条件的随机过程:

$$L_{C,t}^2 = \left\{ \varphi = (\varphi(s), s \geq 0) : \varphi \text{ 关于 } s \text{ 连续, } \varphi(s) \in \mathcal{F}_s = \sigma(B_u, 0 \leq u \leq s), \text{ 并且 } \mathbb{E} \left[ \int_0^t |\varphi(s)|^2 ds < +\infty \right] \right\}$$

而  $(B_s), (B_s^2), \cdots \in L_{C,T}^2$ , 根据多项式函数在连续函数中的  $L^2$  稠密性, 我们便可以定义形如以下形式的积分了:

$$\int_0^t f(B_s) dB_s$$

其中  $f: \mathbb{R} \rightarrow \mathbb{R}$  是连续函数。

4.7.3.1 定义  $\int_0^t B_s dB_s$ 

**Step 1.** 任取一个分割  $\Delta_n: 0 = t_0 < t_1 < \cdots < t_n = t$ , 定义如下过程 (即自变量是  $t$  和  $\omega$  的函数):

$$f_n(s) = \begin{cases} B_0 & , 0 \leq s \leq t_1 \\ B_{t_i} & , t_i < s \leq t_{i+1}, i \in \{1, 2, \cdots, n-1\} \end{cases}$$

并定义其积分为:

$$\int_0^t f_n(s)dB_s = \sum_{i=0}^{n-1} B_{t_i}(B_{t_{i+1}} - B_{t_i})$$

那么它有如下性质。

**命题 4.7.7**

(1).

$$\int_0^t f_n(s)dB_s \in \mathcal{F}_t$$

(2).

$$\mathbb{E} \left[ \int_0^t f_n(s)dB_s \right] = \sum_{i=1}^{n-1} \mathbb{E}[B_{t_i}(B_{t_{i+1}} - B_{t_i})] = 0$$

(3).

$$\mathbb{E} \left[ \left( \int_0^t f_n(s)dB_s \right)^2 \right] = \mathbb{E} \left[ \int_0^t [f_n(s)]^2 ds \right]$$

**Step 2.** 再取一个  $f_m$ , 其对应的分割是  $m$  段的  $\Delta_m$ , 则

$$\mathbb{E} \left[ \left| \int_0^t f_n(s)dB_s - \int_0^t f_m(s)dB_s \right|^2 \right] = \mathbb{E} \left[ \left| \int_0^t f_n(s) - f_m(s)dB_s \right|^2 \right] = \mathbb{E} \left[ \int_0^t |f_n(s) - f_m(s)|^2 ds \right]$$

**Step 3.** 记  $|\Delta_n|$  为分割  $\Delta_n$  中的最大区间长度, 考虑

$$\begin{aligned} \mathbb{E} \left[ \int_0^t |f_n(s) - B_s|^2 ds \right] &= \int_0^t \mathbb{E} \left[ |f_n(s) - B_s|^2 \right] ds \\ &= \int_0^t \mathbb{E} \left[ \left| \sum_{i=1}^n (B_{t_{i-1}} - B_s) I_{\{s \in [t_{i-1}, t_i]\}} \right|^2 \right] ds \\ &\leq \int_0^t \mathbb{E} \left[ \sum_{i=1}^n |B_{t_{i-1}} - B_s|^2 I_{\{s \in [t_{i-1}, t_i]\}} \right] ds \\ &\leq \int_0^t \mathbb{E} \left[ \sum_{i=1}^n |B_{t_{i-1}} - B_{t_i}|^2 I_{\{s \in [t_{i-1}, t_i]\}} \right] ds \\ &\leq \int_0^t \sum_{i=1}^n (t_i - t_{i-1}) I_{\{s \in [t_{i-1}, t_i]\}} ds \\ &= \sum_{i=1}^n (t_i - t_{i-1})^2 \rightarrow 0 \text{ as } |\Delta_n| \rightarrow 0 \end{aligned}$$

所以  $|\Delta_n| \rightarrow 0$  时,

$$\left\{ \int_0^t f_n(s)dB_s \right\}$$

是  $(\Omega, \mathcal{F}_t, \mathbb{P})$  中的柯西列, 其极限 (不依赖于具体  $\Delta_n$  的选取) 就定义为

$$\int_0^t B_s dB_s$$

### 4.7.3.2 一般情形

我们考虑过程  $\{f_t : t \geq 0\}$ , 其满足:

(A1)  $f_t \in \mathcal{F}_t$ .

(A2)  $t \mapsto f_t(\omega)$  连续。

(A3) 对于  $\forall t > 0$ , 任取分割  $\Delta^{(t)}: 0 = s_0 < s_1 < \dots < s_n = t$ ,

$$\lim_{|\Delta^{(t)}| \rightarrow 0} \sum_{i=0}^{n-1} \int_{s_i}^{s_{i+1}} \mathbb{E}[(f_u - f_{s_i})^2] du = 0$$

那么可以证明,  $\forall t \geq 0$ , 存在  $X_t$  使得

$$\lim_{|\Delta^{(t)}| \rightarrow 0} \mathbb{E} \left[ \left| \sum_{i=0}^{n-1} f_{s_i} \cdot (B_{s_{i+1}} - B_{s_i}) - X_t \right|^2 \right] = 0$$

于是我们定义:

$$X_t = \int_0^t f_s dB_s$$

下面介绍一些基本性质。

#### 命题 4.7.8

(1).  $\{X_t\}$  轨道连续。

(2). 线性:

$$\int_0^t a f_s + b g_s dB_s = a \int_0^t f_s dB_s + b \int_0^t g_s dB_s$$

(3).  $\mathbb{E}[\int_0^t f_s dB_s] = 0$ .

(4).  $\forall 0 \leq s \leq t$ ,

$$\mathbb{E} \left[ \int_0^s f_u dB_u \int_0^t g_u dB_u \right] = \int_0^s \mathbb{E}[f_u g_u] du$$

(5).  $X$  关于  $\{\mathcal{F}_t\}$  是鞅。

#### 例 4.7.1.

从定义出发, 证明  $\int_0^t B_s dB_s = \frac{1}{2}(B_t^2 - t)$ .

解答: 分析定义, 我们应当取  $f_t = B_t$ , 并且任取分割  $\Delta: 0 = s_0 < s_1 < \dots < s_n = t$ ,

$$\lim_{|\Delta| \rightarrow 0} \mathbb{E} \left[ \left| \sum_{i=0}^{n-1} B_{s_i} \cdot (B_{s_{i+1}} - B_{s_i}) - X_t \right|^2 \right] = 0$$

其中  $X_t = \frac{1}{2}(B_t^2 - t)$ , 注意到

$$\begin{aligned} \sum_{i=0}^{n-1} 2B_{s_i} \cdot (B_{s_{i+1}} - B_{s_i}) &= \sum_{i=0}^{n-1} [B_{s_{i+1}}^2 - B_{s_i}^2 - (B_{s_{i+1}} - B_{s_i})^2] \\ &= B_t^2 - \sum_{i=0}^{n-1} (B_{s_{i+1}} - B_{s_i})^2 \end{aligned}$$

所以我们只需证明:

$$\lim_{|\Delta| \rightarrow 0} \mathbb{E} \left[ \left| \sum_{i=0}^{n-1} (B_{s_{i+1}} - B_{s_i})^2 - t \right|^2 \right] = 0$$

(从这里开始是作业 15.1) 直接拆开算就行。

### 4.7.4 Itô 公式

我猜这一段讲义的意思是这样的。

首先, 对于一个随机过程  $\{X_t\}$ , 其微分具有以下形式

$$dX_t = u_t dt + f(t) dB_t$$

其中,  $u = \{u_s, s \geq 0\}$  轨道连续, 适应  $\mathcal{F}_s = \sigma(B_u, 0 \leq u \leq s)$ ,  $f = \{f_t, t \geq 0\}$  满足条件 (A1)-(A3)。然而并不知道  $X$  需要满足什么条件才能保证这样的  $u, f$  存在。

然后, 我们已经证明过了:

$$d(B_t^2) = 2B_t dB_t + dt$$

所以, 我们猜测存在这样一个公式:

$$d(g(B_t)) = g'(B_t)dB_t + \frac{1}{2}g''(B_t)dt$$

实际上还真就是如此。

#### 定理 4.7.1

$g: \mathbb{R} \rightarrow \mathbb{R}$  有界, 三阶连续可导, 所有导函数也有界, 记  $f(t) = B_t(\omega)$ , 分割  $\Delta: 0 = s_0 < s_1 < \cdots < s_n = t$ , 则有

$$\begin{aligned} g(f(t)) - g(f(0)) &= \sum_{i=0}^{n-1} [g(f(s_{i+1})) - g(f(s_i))] \\ &= \sum_{i=0}^{n-1} (g'(f(s_i))[f(s_{i+1}) - f(s_i)]) \\ &\quad + \sum_{i=0}^{n-1} \left( \frac{1}{2}g''(f(s_i))[f(s_{i+1}) - f(s_i)]^2 \right) \\ &\quad + \sum_{i=0}^{n-1} \left( \frac{1}{6}g'''(\xi_i)[f(s_{i+1}) - f(s_i)]^3 \right) \\ &= I_1^\Delta + I_2^\Delta + I_3^\Delta \end{aligned}$$

那么

$$\begin{aligned} \lim_{|\Delta| \rightarrow 0} I_1^\Delta &= \int_0^t g'(B_s)dB_s \\ \lim_{|\Delta| \rightarrow 0} I_2^\Delta &= \frac{1}{2} \int_0^t g''(B_s)dB_s \\ \lim_{|\Delta| \rightarrow 0} I_3^\Delta &= 0 \end{aligned}$$

因此,

$$g(B_t) - g(0) = \int_0^t g'(B_s)dB_s + \frac{1}{2} \int_0^t g''(B_s)dB_s$$

写成微分形式:

$$d(g(B_t)) = g'(B_t)dB_t + \frac{1}{2}g''(B_t)dt$$

然后, 我们承认以下事实:

$$(dB_t)^2 = dt$$

并且,  $(dt)^2$ 、 $(dt dB_t)$  相对于  $dt$  都是高阶无穷小。讲义里给出了直观理解方式:

$$\begin{aligned} \lim_{|\Delta| \rightarrow 0} \sum [t_{i+1} - t_i]^2 &= \sum (dt)^2 = 0 \\ \lim_{|\Delta| \rightarrow 0} \sum [B_{t_{i+1}} - B_{t_i}]^2 &= \sum (dB_t)^2 = t \end{aligned}$$

最后, 是一个推广的结论。

## 定理 4.7.2

$g(t, x) : [0, +\infty) \times \mathbb{R} \rightarrow \mathbb{R}$  二阶导函数连续, 令  $Y_t = g(t, X_t)$ , 则

$$Y_t = Y_0 + \int_0^t \partial_t g(s, X_s) ds + \int_0^t \partial_x g(s, X_s) dX_s + \frac{1}{2} \int_0^t \partial_{xx}^2 g(s, X_s) (dX_s)^2$$

其中,

$$(dX_t)^2 = (u_t dt)^2 + 2u_t dt \cdot f_t dB_t + (f_t dB_t)^2 = f_t^2 dt$$

完全展开之后的形式为:

$$Y_t = Y_0 + \int_0^t \partial_t g(s, X_s) ds + \int_0^t \partial_x g(s, X_s) \cdot u_s ds + \int_0^t \partial_x g(s, X_s) \cdot f_s dB_s + \frac{1}{2} \int_0^t \partial_{xx}^2 g(s, X_s) f_s^2 ds$$

以下是一些应用的例子。

## 例 4.7.2.

设  $X_t = B_t$ ,  $f(t, x) = \frac{x^2}{2}$ ,  $Y_t = f(X_t) = \frac{1}{2} B_t^2$ , 则

$$\begin{aligned} dY_t &= \frac{\partial f}{\partial t}(X_t) dt + \frac{\partial f}{\partial x}(X_t) dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(X_t) (dX_t)^2 \\ &= 0 + X_t dX_t + \frac{1}{2} dt \\ &= B_t dB_t + \frac{1}{2} dt \end{aligned}$$

所以

$$\frac{1}{2} B_t^2 = \int_0^t B_s dB_s + \frac{1}{2} t$$

## 例 4.7.3.

证明:

$$tB_t - \int_0^t B_s ds = \int_0^t s dB_s$$

解答: 思路: 取  $f(t, x) = tx$ ,  $Y_t = f(t, B_t)$ , 代入公式计算即可。

## 例 4.7.4.

设  $S_t$  满足下列随机微分方程:

$$dS_t = \sigma S_t dB_t + r S_t dt$$

其中  $\sigma, r > 0$ , 边界  $S_0 = 1$ , 利用 Itô 公式求解  $S_t$ , 并验证  $e^{-rt} S_t$  为鞅。

解答: 设存在二阶可导连续  $g(t, x)$  使得  $S_t = g(t, B_t)$ , 则

$$dS_t = \partial_t g(t, B_t) dt + \partial_x g(t, B_t) dB_t + \frac{1}{2} \partial_{xx}^2 g(t, B_t) dt$$

从而得到偏微分方程组:

$$\begin{cases} \partial_t g + \partial_{xx}^2 g = rg \\ \partial_x g = \sigma g \end{cases}$$

以及边界  $g(0, 0) = 1$ , 最后可解得

$$S_t = e^{(r - \frac{1}{2}\sigma^2)t + \sigma B_t}$$

## 第 5 章 连续时间鞅

本章内容主要来自 Le Gall chapter 3, 有些英文术语我也找不到很合适的翻译, 就按照自己顺口的习惯翻译了, 并且标注了原文。

### 5.1 滤流

开始这一章之前, 我们先阐明滤流和随机过程的详细定义。我们以下默认在概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上讨论, 并且取指标集  $I = [0, +\infty]$ 。

#### 定义 5.1.1 (滤流)

一族  $\sigma$ -域  $(\mathcal{F}_t)_{t \geq 0}$  满足:

$$\forall t \in [0, +\infty], \mathcal{F}_t \subset \mathcal{F}$$

$$\forall 0 \leq s \leq t, \mathcal{F}_0 \subset \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}_\infty$$

则称  $(\mathcal{F}_t)_{t \geq 0}$  是一个滤流 (filtration), 也称  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  为滤流概率空间 (filtered probability space). 

#### 例 5.1.1.

对于随机过程  $X = (X_t)_{t \geq 0}$ , 我们取

$$\mathcal{F}_t^X = \sigma(X_s, 0 \leq s \leq t)$$

且  $\mathcal{F}_\infty^X = \sigma(X_s, s \geq 0)$ , 那么  $(\mathcal{F}_t^X)_{t \geq 0}$  就是一个滤流, 被称为  $X$  的正规滤流 (canonical filtration).

#### 定义 5.1.2 (右连续的滤流)

对于滤流  $(\mathcal{F}_t)_{t \geq 0}$ , 定义:

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

如果  $\forall t \geq 0, \mathcal{F}_{t+} = \mathcal{F}_t$ , 称滤流  $(\mathcal{F}_t)_{t \geq 0}$  是右连续的。 

$(\mathcal{F}_{t+})_{t \geq 0}$  就是一个右连续的滤流。

**定义 5.1.3 (完备的滤流)**

对于滤流  $(\mathcal{F}_t)_{t \geq 0}$ , 记  $N$  为测度空间  $(\Omega, \mathcal{F}_\infty, \mathbb{P})$  上的所有零测集的子集, 即

$$N = \{A \subset \Omega : \exists A' \in \mathcal{F}_\infty \text{ s.t. } A \subset A', \mathbb{P}(A') = 0\}$$

如果  $N \subset \mathcal{F}_0$ , 则称  $(\mathcal{F}_t)_{t \geq 0}$  是完备的。

如果滤流  $(\mathcal{F}_t)_{t \geq 0}$  不完备, 我们可以定义  $\mathcal{F}'_t$  为包含  $\mathcal{F}_t$  和  $\sigma(N)$  的最小  $\sigma$ -域, 记作

$$\mathcal{F}'_t = \mathcal{F}_t \vee \sigma(N), \forall t \geq 0$$

则  $(\mathcal{F}'_t)_{t \geq 0}$  是完备的, 称为滤流  $(\mathcal{F}_t)_{t \geq 0}$  的完备化。

**定义 5.1.4 (可测、适应、循序)**

对于随机过程  $X = (X_t)_{t \geq 0}$ ,  $X_t$  在可测空间  $(E, \mathcal{E})$  上取值, 如果映射

$$\begin{aligned} f : (\Omega \times \mathbb{R}_+, \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)) &\rightarrow (E, \mathcal{E}) \\ (\omega, t) &\mapsto X_t(\omega) \end{aligned}$$

是可测的, 则称随机过程  $X$  是可测的 (measurable).

对于随机过程  $X = (X_t)_{t \geq 0}$ , 滤流  $(\mathcal{F}_t)_{t \geq 0}$ , 如果  $\forall t \geq 0$ ,  $X_t$  是  $\mathcal{F}_t$ -可测的, 那么称  $X$  是 (关于  $(\mathcal{F}_t)_{t \geq 0}$  的) 适应的 (adapted).

如果映射

$$\begin{aligned} g : (\Omega \times [0, t], \mathcal{F}_t \otimes \mathcal{B}([0, t])) &\rightarrow (E, \mathcal{E}) \\ (\omega, s) &\mapsto X_s(\omega) \end{aligned}$$

是可测的, 则称  $X$  是循序的 (progressive).

**命题 5.1.1**

随机过程  $X = (X_t)_{t \geq 0}$  在度量空间  $(E, d)$  (取 Borel  $\sigma$ -域得到可测空间) 上取值,  $X$  关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  适应, 并且  $X$  的样本轨道是右连续或者左连续的, 则  $X$  是循序的。

**证明** 只说明一下右连续的情况, 左连续类似。固定  $t > 0$ ,  $\forall n \geq 1, s \in [0, t]$ , 取 r.v.

$$X_s^n \stackrel{\text{def}}{=} X_{\frac{ks}{n}}, \text{ where } t \text{ s.t. } s \in \left[ \frac{(k-1)t}{n}, \frac{kt}{n} \right), k \in \{1, 2, \dots, n\}$$

并且  $X_t^n = X_t$ , 样本轨道右连续确保了:  $\forall s \in [0, t], \omega \in \Omega$ ,

$$X_s(\omega) = \lim_{n \rightarrow \infty} X_s^n(\omega)$$

那么任取 Borel 集  $A \subset E$ ,

$$\begin{aligned} &\{(\omega, s) \in \Omega \times [0, t] : X_s^n(\omega) \in A\} \\ &= (\{X_t \in A\} \times \{t\}) \cup \left( \bigcup_{k=1}^n \left( \{X_{\frac{ks}{n}} \in A\} \times \left[ \frac{(k-1)t}{n}, \frac{kt}{n} \right) \right) \right) \end{aligned}$$

后者属于  $\mathcal{F}_t \otimes \mathcal{B}([0, t])$ , 因此  $\forall n \geq 1$ , 映射:

$$\begin{aligned} (\Omega \times [0, t], \mathcal{F}_t \otimes \mathcal{B}([0, t])) &\rightarrow (E, \mathcal{B}(E)) \\ (\omega, s) &\mapsto X_s^n(\omega) \end{aligned}$$

是可测的。由于可测函数的逐点极限也是可测的, 这就说明  $X$  是渐进的。  $\square$

**定义 5.1.5 (循序  $\sigma$ -域)**

固定  $A \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$ , 并令  $X_t^A(\omega) = I_A(\omega, t)$ , 所有使得随机过程  $X = (X_t^A)_{t \geq 0}$  是循序过程的集合  $A$ , 组成了一个  $\sigma$ -域, 记作  $\mathcal{P}$ , 称之为循序  $\sigma$ -域。

**5.2 停时**

先来回顾一下停时相关的定义。

**定义 5.2.1 (停时)**

对于滤流  $(\mathcal{F}_t)_{t \geq 0}$ , 随机变量  $T$  在  $[0, +\infty]$  上取值, 如果  $\forall t \geq 0, \{T \leq t\} \in \mathcal{F}_t$ , 则称  $T$  是一个 (关于滤流  $(\mathcal{F}_t)_{t \geq 0}$ ) 的停时。



在本章的剩余内容中, 如无特殊说明, 默认停时  $T$  都是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的。

**推论 5.2.1**

$T$  是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的停时, 则有

$$\begin{aligned} \{T < t\} &= \bigcap_{q \in \mathbb{Q}, q < t} \{T \leq q\} \in \mathcal{F}_t \\ \{T < +\infty\} &= \bigcup_{q \in \mathbb{Q}, q \geq 0} \{T \leq q\} \in \mathcal{F}_\infty \end{aligned}$$

**定义 5.2.2 (停时前  $\sigma$ -域)**

$T$  是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的停时,

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\}$$

则  $\mathcal{F}_T$  是一个  $\sigma$ -域, 称为  $T$  前  $\sigma$ -域。

**定理 5.2.1**

$T: \Omega \rightarrow [0, +\infty]$  是随机变量, 则以下结论等价:

- (1).  $T$  关于  $(\mathcal{F}_{t+})_{t \geq 0}$  是停时。
- (2).  $\forall t > 0, \{T < t\} \in \mathcal{F}_t$ .
- (3).  $\forall t > 0, T \wedge t$  是  $\mathcal{F}_t$  可测的。

**证明** (1)  $\Rightarrow$  (2):  $\forall t \geq 0,$

$$\{T < t\} = \bigcup_{q \in \mathbb{Q}, q > 0} \{T \leq t - q\} \in \mathcal{F}_t$$

(2)  $\Rightarrow$  (1):  $\forall t \geq 0, \forall s > t,$

$$\{T > t\} = \bigcap_{q \in \mathbb{Q}, q > 0} \{T \geq t + q\} \in \mathcal{F}_{t+}$$

(2)  $\Rightarrow$  (3):  $\forall s < t,$

$$\{T \wedge t \leq s\} = \{T \leq s\} \cup \{t \leq s\} = \{T \leq s\} \in \mathcal{F}_s \subset \mathcal{F}_t$$

(3)  $\Rightarrow$  (2):  $\forall t \geq 0,$

$$\{T < t\} = \bigcap_{q \in [0, t] \cap \mathbb{Q}} \{T \leq t - q\} = \bigcap_{q \in [0, t] \cap \mathbb{Q}} \{T \wedge t \leq t - q\} \in \mathcal{F}_t$$

□

## 推论 5.2.2

$T$  是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的停时, 则  $T$  也是关于滤流  $(\mathcal{F}_{t+})_{t \geq 0}$  的停时。

## 定理 5.2.2

$T$  是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的停时, 则

$$\{A \in \mathcal{F}_\infty : \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_{t+}\} = \{A \in \mathcal{F}_\infty : \forall t \geq 0, A \cap \{T < t\} \in \mathcal{F}_t\}$$

我们将上述  $\sigma$ -域记作  $\mathcal{F}_{T+}$ .

**证明** 设  $A \in \mathcal{G}_T$ , 则

$$A \cap \{T < t\} = A \cap \left( \bigcup_{q \in \mathbb{Q}, q > 0} \{T \leq t - q\} \right) = \bigcup_{q \in \mathbb{Q}, q > 0} (A \cap \{T \leq t - q\}) \in \mathcal{F}_t$$

所以  $\mathcal{G}_T \subset \mathcal{F}_{T+}$ ; 反之, 设  $A \in \mathcal{F}_{T+}$ ,

$$A \cap \{T \leq t\} = A \cap \left( \bigcap_{q \in \mathbb{Q}, q > 0} \{T < t + q\} \right) = \bigcap_{q \in \mathbb{Q}, q > 0} (A \cap \{T < t + q\}) \in \mathcal{F}_{t+}$$

所以  $\mathcal{F}_{T+} \subset \mathcal{G}_T$ . □

## 命题 5.2.1 (停时的性质总结)

$T$  是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的停时, 下文中出现的停时如无特殊说明, 默认是关于滤流  $(\mathcal{F}_t)_{t \geq 0}$  的停时。

- (1).  $\mathcal{F}_T \subset \mathcal{F}_{T+}$ , 如果  $(\mathcal{F}_t)$  右连续, 则  $\mathcal{F}_{T+} = \mathcal{F}_T$ .
- (2). 如果  $T = t$  为常数, 则  $\mathcal{F}_T = \mathcal{F}_t$ , 且  $\mathcal{F}_{T+} = \mathcal{F}_{t+}$ .
- (3).  $T$  是  $\mathcal{F}_T$ -可测的。
- (4). 如果  $A \in \mathcal{F}_\infty$ , 令

$$T^A(\omega) = \begin{cases} T(\omega) & , \omega \in A \\ +\infty & , \omega \notin A \end{cases}$$

则  $A$  是  $\mathcal{F}_T$  可测的当且仅当  $T^A$  是停时。

- (5).  $S$  也是停时, 且  $S \leq T$ , 那么  $\mathcal{F}_S \subset \mathcal{F}_T$ , 且  $\mathcal{F}_{S+} \subset \mathcal{F}_{T+}$ .
- (6).  $S$  也是停时, 则  $S \wedge T$  和  $S \vee T$  都是停时, 并且  $\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$ ,  $\{S \leq T\} \in \mathcal{F}_{S \wedge T}$ ,  $\{S = T\} \in \mathcal{F}_{S \wedge T}$ .
- (7).  $\{S_n\}_{n \geq 1}$  是一列单调递增的停时, 则  $S = \lim_{n \rightarrow \infty} S_n$  也是停时。
- (8).  $\{S_n\}_{n \geq 1}$  是一列单调递减的停时, 则  $S = \lim_{n \rightarrow \infty} S_n$  是关于  $(\mathcal{F}_{t+})$  的停时, 并且

$$\mathcal{F}_{S+} = \bigcap_{n=1}^{\infty} \mathcal{F}_{S_n+}$$

- (9).  $\{S_n\}_{n \geq 1}$  是一列单调递减的停时,  $S = \lim_{n \rightarrow \infty} S_n$ , 而且  $\forall \omega$ , 都存在  $N(\omega)$  使得  $\forall n > N(\omega)$  有  $S_n(\omega) = S$ , 则

$$\mathcal{F}_S = \bigcap_{n=1}^{\infty} \mathcal{F}_{S_n}$$

- (10).  $(E, \mathcal{E})$  是一个可测空间, 对于映射

$$Y : \{T < +\infty\} \rightarrow E, \omega \mapsto Y(\omega)$$

定义

$$Y_t : \{T \leq t\} \rightarrow E, \omega \mapsto Y(\omega)$$

为  $Y$  在  $\{T \leq t\}$  上的限制。那么,  $Y$  是  $\mathcal{F}_T$ -可测的当且仅当  $\forall t \geq 0$ ,  $Y_t$  是  $\mathcal{F}_t$ -可测的。

**证明** 基本上都是验证定义以及利用定理 5.2.1 和定理 5.2.2 得到的。

(1). 我们熟知  $\mathcal{F}_t \subset \mathcal{F}_{t+}$ , 那么

$$\mathcal{F}_{T+} = \{A \in \mathcal{F}_\infty : \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_{t+}\} \supset \{A \in \mathcal{F}_\infty : \forall t \geq 0, A \cap \{T \leq t\} \in \mathcal{F}_t\} = \mathcal{F}_T$$

如果  $\mathcal{F}_{t+} = \mathcal{F}_t$ , 则上式中两个集合相等。

(2). 如果  $T = t$  是常数, 则  $A \in \mathcal{F}_T \Leftrightarrow \forall s \geq t, A \in \mathcal{F}_s \Leftrightarrow A \in \mathcal{F}_t$ , 所以  $\mathcal{F}_T = \mathcal{F}_t$ ;  $A \in \mathcal{F}_T \Leftrightarrow \forall s > t, A \in \mathcal{F}_s \Leftrightarrow A \in \mathcal{F}_{t+}$ .

(3).  $\forall s \geq 0, \forall t \geq 0$ ,

$$\{T \leq s\} \cap \{T \leq t\} = \{T \leq s \wedge t\} \in \mathcal{F}_t$$

所以  $\{T \leq s\} \in \mathcal{F}_T$ , 故  $T$  是  $\mathcal{F}_T$ -可测的。

(4). 注意到  $\{T^A \leq t\} = A \cap \{T \leq t\}$ .

(5). 注意到  $\{T \leq t\} \subset \{S \leq t\}$ , 所以

$$A \cap \{T \leq t\} = A \cap \{S \leq t\} \cap \{T \leq t\}$$

如果  $A \in \mathcal{F}_S$ , 则  $A \cap \{S \leq t\} \in \mathcal{F}_t, \{T \leq t\} \in \mathcal{F}_t$ ; 如果  $A \in \mathcal{F}_{S+}$ , 则  $A \cap \{S \leq t\} \in \mathcal{F}_{t+}, \{T \leq t\} \in \mathcal{F}_t \subset \mathcal{F}_{t+}$ .

(6).  $\forall t \geq 0$ ,

$$\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\} \in \mathcal{F}_t$$

$$\{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathcal{F}_t$$

所以这俩是停时; 由 (5) 可知  $\mathcal{F}_{S \wedge T} \subset \mathcal{F}_S \cap \mathcal{F}_T$ , 另一方面. 如果  $A \in \mathcal{F}_S \cap \mathcal{F}_T$ , 则

$$A \cap \{S \wedge T \leq t\} = (A \cap \{S \leq t\}) \cap (A \cap \{T \leq t\}) \in \mathcal{F}_t$$

所以  $\mathcal{F}_S \cap \mathcal{F}_T \subset \mathcal{F}_{S \wedge T}$ ; (待补充)

(7).  $S_n \nearrow S$ , 则

$$\{S \leq t\} = \bigcap_{n=1}^{\infty} \{S_n \leq t\} \in \mathcal{F}_t$$

(8).  $S_n \searrow S$ , 则

$$\{S < t\} = \bigcup_{n=1}^{\infty} \{S_n < t\} \in \mathcal{F}_t$$

所以  $S \in \mathcal{F}_{t+}$ ; 因为  $\mathcal{F}_{S+} \subset \mathcal{F}_{S_n+}$ , 所以

$$\mathcal{F}_{S+} \subset \bigcap_{n=1}^{\infty} \mathcal{F}_{S_n+}$$

另一方面, 如果  $A \in \mathcal{F}_{S_n+}, \forall n$ , 则  $\forall t$ ,

$$A \cap S < t = \bigcup_{n=1}^{\infty} (A \cap \{S_n < t\}) \in \mathcal{F}_t$$

(9). 如果  $A \in \mathcal{F}_S$ ,

$$\{S \leq t\} = \bigcup_{n=1}^{\infty} \{S_n \leq t\} \in \mathcal{F}_t$$

则  $S$  关于  $(\mathcal{F}_t)$  是停时;  $\mathcal{F}_S \subset \mathcal{F}_{S_n}$ , 所以

$$\mathcal{F}_S \subset \bigcap_{n=1}^{\infty} \mathcal{F}_{S_n}$$

反之, 如果  $A \in \mathcal{F}_{S_n}$ , 则

$$A \cap \{S \leq t\} = \bigcup_{n=1}^{\infty} (A \cap \{S_n \leq t\}) \in \mathcal{F}_t$$

(10). 假设  $Y$  限制在集合  $\{T \leq t\}$  上是关于  $\mathcal{F}_t$ -可测的, 则对于  $A \in \mathcal{B}(E)$ ,

$$\{Y \in A\} \cap \{T \leq t\} \in \mathcal{F}_t, \forall t \geq 0$$

此即  $\{Y \in A\} \in \mathcal{F}_T$ . 反之类似。 □

### 定理 5.2.3

$T$  是停时, 随机变量  $S: \Omega \rightarrow [0, +\infty]$  是  $\mathcal{F}_T$ -可测的, 且满足  $S \geq T$ , 则  $S$  也是停时。

**证明** 利用

$$\{S \leq t\} = \{S \leq t\} \cup \{T \leq t\}$$

其中  $\{S \leq t\} \in \mathcal{F}_T$ , 所以右式  $\in \mathcal{F}_t$ . □

### 推论 5.2.3

$T$  是停时, 那么  $\forall n \in \mathbb{N}_+$ ,

$$T_n = \sum_{k=0}^{\infty} \frac{k+1}{2^n} I_{\{k \cdot 2^{-n} < T \leq (k+1) \cdot 2^{-n}\}} + \infty \cdot I_{\{T = +\infty\}}$$

也是停时, 而且  $T_n \searrow T$ . ♥

### 例 5.2.1.

随机过程  $X = (X_t)_{t \geq 0}$  适应  $(\mathcal{F}_t)_{t \geq 0}$ , 在度量空间  $(E, d)$  上取值,  $X$  的轨道连续,  $F \subset E$  为闭集, 则

$$T_F = \inf\{t \geq 0 : X_t \in F\}$$

是一个停时。

**证明** 只需注意到

$$\begin{aligned} \{T_F \leq t\} &= \bigcup_{s \in [0, t]} \{X_s \in F\} \\ &= \left\{ \inf_{s \in [0, t]} d(X_s, F) = 0 \right\} \\ &= \left\{ \inf_{s \in [0, t] \cap \mathbb{Q}} d(X_s, F) = 0 \right\} \in \mathcal{F}_t \end{aligned}$$

这里利用了轨道和度量函数  $d$  的连续性, 将不可数并转化为了可数并。 □

### 例 5.2.2.

随机过程  $X = (X_t)_{t \geq 0}$  适应  $(\mathcal{F}_t)_{t \geq 0}$ , 在度量空间  $(E, d)$  上取值,  $X$  的轨道右连续,  $O \subset E$  为开集, 则

$$T_O = \inf\{t \geq 0 : X_t \in O\}$$

是一个关于  $(\mathcal{F}_{t+})$  的停时。

**证明** 只需注意到

$$\{T_O < t\} = \bigcup_{s \in [0, t) \cap \mathbb{Q}} \{X_s \in O\} \in \mathcal{F}_t$$

再利用定理 5.2.1 即可。 □

### 5.3 鞅的定义与基本性质

在本章的是剩余内容中，默认随机过程在  $\mathbb{R}$  上取值。

#### 定义 5.3.1

$X = (X_t)_{t \geq 0}$  适应  $(\mathcal{F}_t)_{t \geq 0}$ ，且  $\forall t, X_t \in L^1$ ，如果

$$\forall 0 \leq s < t, \mathbb{E}[X_t | \mathcal{F}_s] = X_s$$

则称  $X$  是 (连续时间) 鞅。如果把上式中的等号替换为小于等于号、大于等于号，则称之为上鞅、下鞅。

#### 例 5.3.1.

随机过程  $Z = (Z_t)_{t \geq 0}$  在  $\mathbb{R}$  上取值，适应  $(\mathcal{F}_t)$ ，并且有独立增量，即： $\forall 0 \leq s < t$ ， $Z_t - Z_s$  独立于  $\mathcal{F}_s$ 。那么

- (1). 如果  $\forall t \geq 0, Z_t \in L^1$ ，则  $\tilde{Z}_t = Z_t - \mathbb{E}[Z_t]$  是鞅。
- (2). 如果  $\forall t \geq 0, Z_t \in L^2$ ，则  $\tilde{Y}_t = \tilde{Z}_t^2 - \mathbb{E}[\tilde{Z}_t^2]$  是鞅。
- (3). 如果存在  $\theta \in \mathbb{R}$  使得  $\forall t \geq 0, \mathbb{E}[e^{\theta Z_t}] < +\infty$ ，则

$$X_t = \frac{e^{\theta Z_t}}{\mathbb{E}[e^{\theta Z_t}]}$$

是鞅。

**证明** 可积性和适应性显然，不再赘述，我们直接验证鞅的最后一条定义。对于  $\forall t \geq 0$ ，我们记  $\mu_t = \mathbb{E}[Z_t]$ ，那么对于  $\forall s > 0$ ，

$$\begin{aligned} \mathbb{E}[\tilde{Z}_{t+s} | \mathcal{F}_t] &= \mathbb{E}[Z_{t+s} | \mathcal{F}_t] - \mu_{t+s} \\ &= \mathbb{E}[Z_{t+s} - Z_t | \mathcal{F}_t] + \mathbb{E}[Z_t | \mathcal{F}_t] - \mu_{t+s} \\ &= \mathbb{E}[Z_{t+s} - Z_t] + Z_t - \mu_{t+s} \\ &= \mu_{t+s} - \mu_t + Z_t - \mu_{t+s} = Z_t - \mu_t = \tilde{Z}_t \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\tilde{Y}_{t+s} | \mathcal{F}_t] &= \mathbb{E}[(Z_{t+s} - \mu_{t+s})^2 | \mathcal{F}_t] - \mathbb{E}[(Z_{t+s} - \mu_{t+s})^2] \\ &= \mathbb{E}[(Z_{t+s} - Z_t + Z_t - \mu_{t+s})^2 | \mathcal{F}_t] - \mathbb{E}[Z_{t+s}^2] + \mu_{t+s}^2 \\ &= \mathbb{E}[(Z_{t+s} - Z_t)^2 + (Z_t - \mu_{t+s})^2 + 2(Z_{t+s} - Z_t)(Z_t - \mu_{t+s}) | \mathcal{F}_t] - \mathbb{E}[Z_{t+s}^2] + \mu_{t+s}^2 \\ &= \mathbb{E}[(Z_{t+s} - Z_t)^2] + (Z_t - \mu_{t+s})^2 + 2(Z_t - \mu_{t+s})\mathbb{E}[(Z_{t+s} - Z_t)] - \mathbb{E}[Z_{t+s}^2] + \mu_{t+s}^2 \\ &= \mathbb{E}[-2(Z_{t+s} - Z_t)Z_t - Z_t^2 + Z_{t+s}^2] + (Z_t - \mu_{t+s})^2 + 2(Z_t - \mu_{t+s})(\mu_{t+s} - \mu_t) - \mathbb{E}[Z_{t+s}^2] + \mu_{t+s}^2 \\ &= -2(\mu_{t+s} - \mu_t)\mu_t - \mathbb{E}[Z_t^2] + \mathbb{E}[Z_{t+s}^2] + (Z_t - \mu_{t+s})^2 + 2(Z_t - \mu_{t+s})(\mu_{t+s} - \mu_t) - \mathbb{E}[Z_{t+s}^2] + \mu_{t+s}^2 \\ &\quad (\text{全拆开之后消了很多项}) \\ &= Z_t^2 - 2\mu_t Z_t - \mathbb{E}[Z_t^2] \\ &= (Z_t - \mu_t)^2 + \mathbb{E}[Z_t^2] - \mu_t^2 = \tilde{Y}_t \end{aligned}$$

$$\begin{aligned} \mathbb{E}[X_{t+s} | \mathcal{F}_t] &= \frac{\mathbb{E}[e^{\theta Z_{t+s}} | \mathcal{F}_t]}{\mathbb{E}[e^{\theta Z_{t+s}}]} \\ &= \frac{\mathbb{E}[e^{\theta(Z_{t+s} - Z_t)} e^{\theta Z_t} | \mathcal{F}_t]}{\mathbb{E}[e^{\theta Z_{t+s}}]} \\ &= \frac{\mathbb{E}[e^{\theta(Z_{t+s} - Z_t)}] e^{\theta Z_t}}{\mathbb{E}[e^{\theta Z_{t+s}}]} \end{aligned}$$

$$\begin{aligned}
&= \frac{\mathbb{E}[e^{\theta(Z_{t+s}-Z_t)}]\mathbb{E}[e^{\theta Z_t}]e^{\theta Z_t}}{\mathbb{E}[e^{\theta Z_{t+s}}]\mathbb{E}[e^{\theta Z_t}]} \\
&= \frac{\mathbb{E}[e^{\theta Z_{t+s}}]e^{\theta Z_t}}{\mathbb{E}[e^{\theta Z_{t+s}}]\mathbb{E}[e^{\theta Z_t}]} \\
&= \frac{e^{\theta Z_t}}{\mathbb{E}[e^{\theta Z_t}]} = X_t
\end{aligned}$$

□

我们之前介绍过的 Poisson 过程和布朗运动就具有独立增量，因此我们得到了如下例子。

### 例 5.3.2. 由 Poisson 过程生成的鞅

$N = (N_t)_{t \geq 0}$  是一个参数为  $\lambda$  的 Poisson 过程，并设

$$\mathcal{F}_t = \sigma(N_s, 0 \leq s \leq t), t \geq 0$$

那么：

- (1).  $M_t = N_t - \lambda t$  是一个鞅。
- (2).  $Z_t = (N_t - \lambda t)^2 - \lambda t$  是一个鞅。
- (3). 对于  $\alpha > 0$ ，设  $\beta = (e^\alpha - 1)\lambda$ ，则

$$L_t = \exp\{\alpha N_t - \beta t\}, t \geq 0$$

是一个鞅。

证明

$$\mathbb{E}[e^{\alpha N_t}] = \sum_{k=0}^{\infty} e^{-\lambda} e^{\alpha k} \frac{\lambda^k}{k!} = e^{\lambda(e^\alpha - 1)}$$

□

### 例 5.3.3. 由布朗运动生成的鞅

$B = (B_t)_{t \geq 0}$  是一个布朗运动，并设

$$\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t), t \geq 0$$

那么  $\forall \theta \in \mathbb{R}$ ,

$$X_t = \exp\left\{\theta B_t - \frac{1}{2}\theta^2 t\right\}$$

是一个鞅。

证明

$$\mathbb{E}[e^{\theta B_t}] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{1}{2}\theta^2 t} e^{\frac{1}{2\pi t}(x-t\theta)^2} dx = e^{\frac{1}{2}\theta^2 t}$$

□

### 定理 5.3.1 (关于凸函数)

$f: \mathbb{R} \rightarrow \mathbb{R}$  是一个有界凸函数，则

- (1).  $X = (X_t)_{t \geq 0}$  是一个鞅，则  $f(X_t)$  是一个下鞅。
- (2).  $X = (X_t)_{t \geq 0}$  是一个下鞅，且  $f$  单调递增，则  $f(X_t)$  是一个下鞅。

证明 琴生不等式易证。

□

我们经常取凸函数  $f(x) = |x|^p$  和  $f(x) = x^+$ .

### 定理 5.3.2

$X = (X_t)_{t \geq 0}$  是一个下鞅, 那么  $\forall t \geq 0$ ,

$$\sup_{0 \leq s \leq t} \mathbb{E}[|X_s|] < +\infty$$

上鞅也有同样的结论。

**证明** 我们分别考虑  $X_s$  的正部和负部, 根据定理 5.3.1,  $(X_s^+)_{s \geq 0}$  是一个下鞅, 所以  $\forall s \leq t$ ,

$$\mathbb{E}[X_s^+] \leq \mathbb{E}[X_t^+]$$

另一方面,

$$\mathbb{E}[X_s^-] = \mathbb{E}[X_s^+] - \mathbb{E}[X_s] \leq \mathbb{E}[X_t^+] - \mathbb{E}[X_0]$$

于是

$$\sup_{0 \leq s \leq t} \mathbb{E}[|X_s|] = \sup_{0 \leq s \leq t} (\mathbb{E}[X_s^+] + \mathbb{E}[X_s^-]) \leq 2\mathbb{E}[X_t^+] - \mathbb{E}[X_0] < +\infty$$

□

### 定理 5.3.3 (两个不等式)

$X = (X_t)_{t \geq 0}$  为上鞅, 且其轨道右连续, 则我们有以下结论:

(1). 最大值不等式: 对于  $\forall \lambda > 0$ , 有

$$\mathbb{P}\left(\sup_{0 \leq s \leq t} |X_s| > \lambda\right) \leq \frac{1}{\lambda} (2\mathbb{E}[|X_t|] + \mathbb{E}[|X_0|])$$

(2). Doob 不等式: 对于  $p > 1$  和  $t > 0$ , 有

$$\mathbb{E}\left[\sup_{0 \leq s \leq t} |X_s|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_t|^p]$$

**证明** 我们的思路是用离散鞅来逼近连续鞅。首先, 固定  $t > 0$ , 取  $D$  是  $[0, t]$  的可数稠密子集, 并且满足  $0 \in D, t \in D$ , 设有一列集合  $D_m \nearrow D$ , 其中  $D_m$  具有以下形式:

$$D_m = \{0 = t_0^m < t_1^m < \cdots < t_m^m = t\}$$

固定  $m$ , 我们考虑  $\{Y_n = X_{t_n \wedge m}, n \in \mathbb{N}_+\}$ , 这是一个关于  $(\mathcal{G}_t = \mathcal{F}_{t_n \wedge m})$  的上鞅, 利用离散鞅的定理 1.6.1,

$$\lambda \mathbb{P}\left(\sup_{s \in D_m} |X_s| > \lambda\right) \leq \mathbb{E}[|X_0|] + 2\mathbb{E}[|X_t|]$$

由概率测度的上连续性,

$$\mathbb{P}\left(\sup_{s \in D} |X_s| > \lambda\right) = \lim_{m \rightarrow \infty} \mathbb{P}\left(\sup_{s \in D_m} |X_s| > \lambda\right) \leq \frac{1}{\lambda} (\mathbb{E}[|X_0|] + 2\mathbb{E}[|X_t|])$$

$X_s$  的轨道右连续, 所以

$$\mathbb{P}\left(\sup_{s \in [0, t]} |X_s| > \lambda\right) = \mathbb{P}\left(\sup_{s \in D} |X_s| > \lambda\right) \leq \frac{1}{\lambda} (\mathbb{E}[|X_0|] + 2\mathbb{E}[|X_t|])$$

第二个结论类似, 利用离散情形的定理 1.6.2 即可。

□

## 5.4 鞅的收敛性

### 5.4.1 轨道修正

在鞅的定义中，我们没有对轨道  $t \mapsto X_t(\omega)$  的连续性作任何要求，这导致我们无法从“离散”来逼近“连续”（在布朗运动那一章我们经常使用这个技巧）。本小节的目的在于介绍某些情况下可以对鞅的轨道进行修正，从而使得其具有良好的轨道性质。

#### 定义 5.4.1 (上穿次数)

对于函数  $f: I \rightarrow \mathbb{R}$ ，其中  $I \subset \mathbb{R}_+$ ， $a < b$  为常数，定义  $f$  在  $I$  上关于区间  $(a, b)$  的上穿次数：

$$M_{a,b}^f(I) \stackrel{\text{def}}{=} \sup\{k \in \mathbb{N}_+ : \exists\{s_1 < t_1 < s_2 < t_2 < \dots < s_k < t_k\} \subset I \text{ s.t. } f(s_i) \leq a, f(t_i) \geq b, \forall i \in [k]\}$$

如果右侧集合为空集，就取  $M_{a,b}^f(I) = 0$ ，即没有穿越出区间  $(a, b)$ 。

下面是一个分析的结论。

#### 引理 5.4.1

$D \stackrel{\text{dense}}{\subset} \mathbb{R}_+$ ， $f: D \rightarrow \mathbb{R}$ ，设对于任意的  $t \in D$ ，都有：

- 1°  $f$  在  $D \cap [0, t]$  上有界；
- 2° 任取有理数  $a < b$ ，都有：

$$M_{a,b}^f(D \cap [0, t]) < +\infty$$

那么如下左、右极限存在：

$$f(t-) = \lim_{D \ni s \nearrow t} f(s), \forall t > 0$$

$$f(t+) = \lim_{D \ni s \searrow t} f(s), \forall t \leq 0$$

进一步地，定义  $g(t) = f(t+)$ ，则  $g(t)$  右连续且左极限存在，简称右连左极 (càdlàg, RCLL)。

**证明** (反证) 假设对于某个  $t > 0$ ，左极限不存在，那么存在有理数  $a < b$  使得

$$\liminf_{D \ni s \nearrow t} f(s) < a < b < \limsup_{D \ni s \nearrow t} f(s)$$

这表明  $M_{a,b}^f(D \cap [0, t]) = \infty$ ，矛盾。

右极限的情形类似。

#### 定理 5.4.1

$X = (X_t)_{t \geq 0}$  为一个关于  $(\mathcal{F}_t)_{t \geq 0}$  的 (上) 鞅， $D \stackrel{\text{dense}}{\subset} \mathbb{R}_+$ ，那么

- (1). 以下极限 a.s. 存在：

$$f(t-) = \lim_{D \ni s \nearrow t} f(s), \forall t > 0$$

$$f(t+) = \lim_{D \ni s \searrow t} f(s), \forall t \leq 0$$

- (2). 对于  $t \geq 0$ ， $X_{t+} \in L^1$  且

$$X_t \geq \mathbb{E}[X_{t+} | \mathcal{F}_t]$$

其中等号成立当且仅当  $t \mapsto \mathbb{E}[X_t]$  右连续。

因此， $(X_{t+})_{t \geq 0}$  是一个关于  $(\mathcal{F}_{t+})_{t \geq 0}$  的 (上) 鞅。

**证明** (1). 固定  $t \in D$ , 由最大值不等式可得

$$\mathbb{P}\left(\sup_{s \in D \cap [0, t]} |X_s| > \lambda\right) \leq \frac{1}{\lambda} (2\mathbb{E}[|X_T|] + \mathbb{E}[|X_0|])$$

令  $\lambda \rightarrow \infty$  可得

$$\sup_{s \in D \cap [0, t]} |X_s| < +\infty$$

选择  $D$  的一列有限子集  $D_m$ , 满足:

$$0, t \in D_m, D_m \nearrow D, D \cap [0, t] = \bigcup_m D_m$$

而  $(X_t)$  限制在  $D_m$  上是离散时间鞅, 由离散时间鞅的上穿不等式 (定理 1.5.1), 有

$$\mathbb{E}[M_{a,b}^X(D_m)] \leq \frac{1}{b-a} \mathbb{E}[(X_T - a)^-]$$

令  $m \rightarrow \infty$ , 由 Fatou 引理可得

$$\mathbb{E}[M_{a,b}^X(D \cap [0, t])] \leq \liminf_{m \rightarrow \infty} \mathbb{E}[M_{a,b}^X(D_m)] \leq \frac{1}{b-a} \mathbb{E}[(X_T - a)^-] < +\infty$$

于是由引理 5.4.1 则得证。

(2). 定义:

$$X_{t+}(\omega) = \begin{cases} \lim_{D \ni s \searrow t} X_s(\omega) & , \text{如果此极限存在} \\ 0 & , \text{其他情况} \end{cases}$$

那么, 其他情况的集合 (是个零测集) 可以记作

$$N = \bigcup_{t \in D} \left( \left\{ \sup_{t \in D \cap [0, t]} |X_t| = +\infty \right\} \cup \left\{ \bigcup_{a < b, a, b \in \mathbb{Q}} \{M_{a,b}^X(D \cap [0, t]) = +\infty\} \right\} \right)$$

Step1. 选取  $D \ni t_n \searrow t, t_n > t$ , 由构造可知

$$X_{t+} = \lim_{n \rightarrow \infty} X_{t_n}$$

记  $Y_k = X_{t-k}, k \leq 0$ , 那么  $(Y_k)$  是一个向后鞅, 而且

$$\sup_{k \leq 0} \mathbb{E}[|Y_k|] = \sup_{k \geq 0} \mathbb{E}[|X_{t_k}|] < +\infty$$

那么

$$Y_k \xrightarrow{L^1} Y_{-\infty} = X_{t+} \in L^1$$

Step2. 由  $t_n > t$ ,

$$X_t \geq \mathbb{E}[X_{t_n} | \mathcal{F}_t]$$

令  $n \rightarrow \infty$ , 可得

$$X_t \geq \mathbb{E}[X_{t+} | \mathcal{F} + t]$$

我们希望证明等号成立, 那么只需证明:

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_{t+} | \mathcal{F} + t]] = \mathbb{E}[X_{t+}]$$

而

$$\mathbb{E}[X_{t+}] = \mathbb{E}[\lim_{n \rightarrow \infty} X_{t_n}] = \lim_{n \rightarrow \infty} \mathbb{E}[X_{t_n}] = \mathbb{E}[X_t]$$

Step3. 最后我们说明  $(X_{t+})$  是  $(\mathcal{F}_{t+})$  的上鞅, 对于  $s < t$ , 取  $D$  中序列  $s_n \searrow s, t_n \searrow t, s_n \leq t_n$ , 取

$$A \in \mathcal{F}_{s+} \bigcap_{s_n} \mathcal{F}_{s_n}$$

那么

$$A \in \mathcal{F}_{s_n}, \forall n \geq 1$$

于是

$$\begin{aligned} \mathbb{E}[X_{s+}I_A] &= \mathbb{E}\left[\lim_{n \rightarrow \infty} X_{s_n}I_A\right] = \lim_{n \rightarrow \infty} \mathbb{E}[X_{s_n}I_A] \\ &\geq \lim_{n \rightarrow \infty} \mathbb{E}[X_{t_n}I_A] \\ &= \mathbb{E}[X_{t+}I_A] \\ &= \mathbb{E}[\mathbb{E}[X_{t+}I_A|\mathcal{F}_{s+}]] = \mathbb{E}[\mathbb{E}[X_{t+}|\mathcal{F}_{s+}]I_A] \end{aligned}$$

由此可得：

$$\mathbb{E}[X_{t+}|\mathcal{F}_{s+}] \leq X_{s+}$$

鞅的情形类似。

这里利用了一个结论： $X, Y$  关于  $\mathcal{F}$  可测，那么  $X \leq Y$  当且仅当  $\mathbb{E}[XI_A] \leq \mathbb{E}[YI_A], \forall A \in \mathcal{F}$ . □

**定理 5.4.2 (轨道修正)**

$(\mathcal{F}_t)_{t \geq 0}$  右连续且完备， $X = (X_t)_{t \geq 0}$  是关于  $(\mathcal{F}_t)_{t \geq 0}$  的上鞅，且  $t \mapsto \mathbb{E}[X_t]$  右连续，那么存在一个  $(X_t)$  的修改  $(Y_t)$ ，其也是一个上鞅，且具有 RCLL 轨道。

**证明** 只需定义  $Y_t = X_{t+}(\omega)I_{\{\lim_{D \ni s \searrow t} X_s(\omega) \text{ 存在}\}}$ ，则  $(Y_t)$  是关于  $(\mathcal{F}_{t+}) = (\mathcal{F}_t)$  的上鞅，那么

$$X_t = \mathbb{E}[X_{t+}|\mathcal{F}_t] = \mathbb{E}[X_{t+}|\mathcal{F}_{t+}] = X_{t+} = Y_t \text{ a.s.}$$

□

### 5.4.2 右连续鞅的收敛性

**定理 5.4.3 (鞅收敛定理)**

$X$  是上鞅，且有右连续轨道，如果

$$\sup_{0 \leq t < \infty} \mathbb{E}[|X_t|] < +\infty$$

那么存在一个随机变量  $X_\infty \in L^1$ ，使得

$$\lim_{t \rightarrow \infty} X_t(\omega) = X_\infty(\omega) \text{ a.s.}$$

**证明** 设  $D$  是  $\mathbb{R}_+$  的一个可数稠密子集，对于任意的  $t \in D$ ，有理数  $a < b$ ，有

$$\mathbb{E}[M_{a,b}^X(D \cap [0, t])] \leq \frac{1}{b-a} \mathbb{E}[(X_T - a)^-] \leq \frac{1}{b-a} \left( \sup_{0 \leq t < \infty} \mathbb{E}[|X_t|] + a \right) < +\infty$$

令  $t \rightarrow \infty$  可得

$$\mathbb{E}[M_{a,b}^X(D)] < +\infty$$

从而可得

$$M_{a,b}^X(D) < +\infty \text{ a.s.}$$

所以存在

$$X_\infty = \lim_{D \ni t \rightarrow \infty} X_t \text{ a.s.}$$

由 Fatou 引理，

$$\mathbb{E}[|X_\infty|] = \mathbb{E}\left[\lim_{D \ni t \rightarrow \infty} |X_t|\right] = \mathbb{E}\left[\lim_{D \ni t \rightarrow \infty} |X_t|\right] \leq \liminf_{D \ni t \rightarrow \infty} \mathbb{E}[|X_t|] < +\infty$$

因此  $X_\infty \in L^1$ .

我们目前只说明了  $D$  上的收敛性, 下面将其扩充到  $\mathbb{R}_+$  上. 对于  $\forall \varepsilon > 0$ , 存在  $N$ , 当  $D \ni t \geq N$  时,

$$|X_t - X_\infty| \leq \varepsilon$$

那么对于任意的  $s \geq N$ , 存在一系列  $s_n \in D$  使得  $s_n \searrow s$ , 因此

$$|X_{s_n} - X_\infty| \leq \varepsilon$$

令  $n \rightarrow \infty$  得到

$$|X_s - X_\infty| \leq \varepsilon$$

□

注意, 相比于离散时间鞅的收敛定理 (定理 1.5.2), 此处多了一个结论:  $X_\infty \in L^1$ .

#### 定义 5.4.2

称一个鞅  $(X_t)_{t \geq 0}$  是闭的 (closed), 如果存在一个随机变量  $Z \in L^1$ , 使得

$$X_t = \mathbb{E}[Z | \mathcal{F}_t], \forall t \geq 0$$



#### 定理 5.4.4

$X = (X_t)_{t \geq 0}$  是一个鞅, 且具有右连续轨道, 那么以下结论等价:

- (1).  $X$  闭。
- (2).  $\{X_t, t \geq 0\}$  一致可积。
- (3).  $t \rightarrow \infty$  时,  $X_t$  几乎处处收敛, 且依  $L^1$  收敛。

**证明** (1)  $\Rightarrow$  (2): 由于

$$X_t = \mathbb{E}[Z | \mathcal{F}_t], t \geq 0$$

是  $Z$  生成的“条件期望列”, 所以是一致可积的, 见推论 1.7.2.

(2)  $\Rightarrow$  (3): 一致可积可得

$$\sup_{t \geq 0} \mathbb{E}[|X_t|] < +\infty$$

由定理 5.4.3 可知  $X_t \xrightarrow{\text{a.s.}} X_\infty$ , 结合一致可积性可知  $X_t \xrightarrow{L^1} X_\infty$

(3)  $\Rightarrow$  (1): 对于  $\forall s > t$ ,

$$X_t = \mathbb{E}[X_s | \mathcal{F}_t]$$

令  $s \rightarrow \infty$ ,

$$X_t = \mathbb{E}[X_\infty | \mathcal{F}_t]$$

这说明  $X$  闭。

□

## 5.5 鞅的择停定理

这一小节中，我们承认以下命题成立。

### 命题 5.5.1

$X = (X_t)_{t \geq 0}$  是循序过程，适应  $(\mathcal{F}_t)_{t \geq 0}$ ， $T$  是关于  $(\mathcal{F}_t)_{t \geq 0}$  的停时，则

$$X_T \cdot I_{T < +\infty}$$

是  $\mathcal{F}_T$ -可测的。

进一步地，如果  $\lim_{t \rightarrow \infty} X_t = X_\infty \in \mathcal{F}_\infty$  a.s.，定义

$$X_T(\omega) = X_\infty(\omega)I_{\{T(\omega)=+\infty\}} + X_{T(\omega)}(\omega)I_{\{T < +\infty\}}$$

那么， $X_T$  是  $\mathcal{F}_T$ -可测的。

本节内容中我们大部分情况下假设鞅是在实数（度量）空间上取值、轨道右连续的，而我们在命题 5.1.1 说明过，具有右连续轨道的适应过程是循序过程，所以我们可以直接使用  $X_T \in \mathcal{F}_T$  这个结论，以下不再作特殊说明。

Doob 择停定理首先给出了一致可积、右连续情形下的结论。

### 定理 5.5.1 (Doob 择停定理)

$X = (X_t)_{t \geq 0}$  是一致可积鞅，且具有右连续轨道，那么对于任意停时  $S \leq T$ ，有  $X_S, X_T \in L^1$ ，且

$$X_S = \mathbb{E}[X_T | \mathcal{F}_S]$$

特别地，取  $T = \infty$ ，那么

$$X_S = \mathbb{E}[X_\infty | \mathcal{F}_S]$$

$$\mathbb{E}[X_0] = \mathbb{E}[X_S] = \mathbb{E}[X_\infty]$$

**证明** 对于任意的  $n \in \mathbb{N}$ ，记

$$T_n = \sum_{k=0}^{\infty} \frac{k+1}{2^n} I_{\{\frac{k}{2^n} < T \leq \frac{k+1}{2^n}\}} + (+\infty) \cdot I_{\{T=+\infty\}}$$

$$S_n = \sum_{k=0}^{\infty} \frac{k+1}{2^n} I_{\{\frac{k}{2^n} < S \leq \frac{k+1}{2^n}\}} + (+\infty) \cdot I_{\{S=+\infty\}}$$

这是我们在推论 5.2.3 提到的构造， $T_n, S_n$  都是停时，并且  $T_n \searrow T, S_n \searrow S$ 。

记

$$Y_k^{(n)} = X_{k \cdot 2^{-n}}, k \geq 0$$

$$\mathcal{H}_k^{(n)} = \mathcal{F}_{k \cdot 2^{-n}}$$

那么  $\{Y_k^{(n)}\}$  是一致可积的，由离散鞅的择停定理：

$$X_{S_n} = T_{2^n \cdot S_n}^{(n)} = \mathbb{E}[Y_{2^n \cdot T_n}^{(n)} | \mathcal{H}_{2^n \cdot S_n}^{(n)}] = \mathbb{E}[X_{T_n} | \mathcal{F}_{S_n}]$$

任取  $A \in \mathcal{F}_S \subset \mathcal{F}_{S_n}$ ，由条件期望的定义，

$$\mathbb{E}[X_{S_n} I_A] = \mathbb{E}[X_{T_n} I_A]$$

结合轨道的右连续性，我们知道

$$X_S = \lim_{n \rightarrow \infty} X_{S_n}, X_T = \lim_{n \rightarrow \infty} X_{T_n} \text{ a.s.}$$

再结合一致可积性，可知上述收敛在  $L^1$  也成立，因此当  $n \rightarrow \infty$  时可得

$$\mathbb{E}[X_S I_A] = \mathbb{E}[X_T I_A], \forall A \in \mathcal{F}_S$$

所以  $X_S = \mathbb{E}[X_T | \mathcal{F}_S]$ , 然后我们接下来证明  $X_S, X_T \in L^1$ , 注意到

$$X_{S_n} = \mathbb{E}[X_\infty | \mathcal{F}_{S_n}], \forall n \in \mathbb{N}_+$$

所以  $X_{S_n} \in L^1$ , 因此  $X_S \in L^1$ . □

如果我们不假设一致可积性, 而假设停时有界, 也有一样的结论。

**定理 5.5.2**

$X = (X_t)_{t \geq 0}$  是一个鞅, 且具有右连续轨道, 如果  $S \leq T$  是两个有界的停时, 那么

$$\begin{aligned} X_S, X_T &\in L^1 \\ X_S &= \mathbb{E}[X_T | \mathcal{F}_S] \end{aligned}$$

**证明** 不妨设  $S \leq T \leq a$ , 考虑如下的鞅:  $Y_t = \mathbb{E}[X_a | \mathcal{F}_t] = X_{t \wedge a}$ , 所以  $Y_t$  是闭的, 进而是一致可积的, 可得

$$Y_S = \mathbb{E}[Y_T | \mathcal{F}_S] \Leftrightarrow X_{S \wedge a} = \mathbb{E}[X_{T \wedge a} | \mathcal{F}_S] \Leftrightarrow X_S = \mathbb{E}[X_T | \mathcal{F}_S]$$

同时, 由于  $Y_S, Y_T \in L^1$ , 所以  $X_S, X_T \in L^1$ . □

下面这个定理给出了停止鞅  $(X_{t \wedge T})$  和原来的鞅之间的关系。

**定理 5.5.3**

$(X_t)_{t \geq 0}$  是一个鞅, 且具有右连续轨道,  $T$  是停时, 那么

- (1).  $(X_{t \wedge T})_{t \geq 0}$  也是一个鞅。
- (2). 如果  $(X_t)$  一致可积, 那么  $(X_{t \wedge T})$  一致可积, 因此对于  $\forall t \geq 0$ ,

$$X_{t \wedge T} = \mathbb{E}[X_{\infty \wedge T} | \mathcal{F}_t] = \mathbb{E}[X_T | \mathcal{F}_t]$$

**证明** 我们先说明 (2), 希望证明:

$$X_{t \wedge T} = \mathbb{E}[X_T | \mathcal{F}_t]$$

对于停时  $T$ , 任取常数  $t \geq 0$ ,  $T \wedge t$  也是一个停时, 因此  $X_{t \wedge T} \in \mathcal{F}_{t \wedge T} \subset \mathcal{F}_t$ , 接下来只需证明对于  $\forall A \in \mathcal{F}_t$  都有

$$\mathbb{E}[X_{t \wedge T} I_A] = \mathbb{E}[X_T I_A]$$

注意到

$$\begin{aligned} \mathbb{E}[X_{t \wedge T} I_A] &= \mathbb{E}[X_{t \wedge T} I_{A \cap \{T \geq t\}}] + \mathbb{E}[X_{t \wedge T} I_{A \cap \{T < t\}}] \\ &= \mathbb{E}[X_{t \wedge T} I_{A \cap \{T \geq t\}}] + \mathbb{E}[X_T I_{A \cap \{T < t\}}] \end{aligned}$$

所以我们只需证明

$$\mathbb{E}[X_{t \wedge T} I_{A \cap \{T \geq t\}}] = \mathbb{E}[X_T I_{A \cap \{T \geq t\}}]$$

因为  $(X_t)_{t \geq 0}$  是一致可积鞅, 我们对  $t \wedge T \leq T$  这两个停时应用择停定理, 可得

$$X_{t \wedge T} = \mathbb{E}[X_T | \mathcal{F}_{w \wedge T}] \tag{*}$$

而对于  $A \in \mathcal{F}_t$ , 因为  $\{T \geq t\} \in \mathcal{F}_t$ , 因此  $A \cap \{T \geq t\} \in \mathcal{F}_t$ , 同时  $\forall s \geq 0$ ,

$$A \cap \{T \geq t\} \cap \{T \leq s\} = \begin{cases} \emptyset & , t > s \\ \underset{\in \mathcal{F}_t \subset \mathcal{F}_s}{A \cap \{T \geq t\} \cap \{T \leq s\}} & , t \leq s \end{cases} \in \mathcal{F}_s$$

因此  $A \cap \{T \geq t\} \in \mathcal{F}_T$ , 从而  $A \cap \{T \geq t\} \in \mathcal{F}_{t \wedge T}$ , 于是由 (\*) 式可知

$$\mathbb{E}[X_{t \wedge T} I_{A \cap \{T \geq t\}}] = \mathbb{E}[X_T I_{A \cap \{T \geq t\}}]$$

于是得证。

再来说明 (1), 任取常数  $a > 0$ , 那么  $X_{t \wedge a} = \mathbb{E}[X_a | \mathcal{F}_t]$  是一致可积鞅, 进而由 (2) 可知  $X_{t \wedge a \wedge T}$  是一致可积鞅, 也就是说, 在  $t \in [0, a]$  时,  $X_{t \wedge T}$  是鞅, 令  $a \rightarrow \infty$  即可. □

最后, 我们给出非负上鞅的择停定理, 它也不需要一致可积性的条件。

## 定理 5.5.4

$Z = (Z_t)_{t \geq 0}$  是非负上鞅, 且具有右连续轨道, 如果  $U \leq V$  是两个停时, 那么  $Z_U, Z_V \in L^1$ , 且

$$Z_U \geq \mathbb{E}[Z_V | \mathcal{F}_U]$$

**证明** 先假设  $U, V$  都是有界停时, 即  $U \leq V \leq P$ ,  $P$  为常数, 对于  $n \in \mathbb{N}$ , 定义

$$U_n = \sum_{k=0}^{\lfloor P \cdot 2^n - 1 \rfloor} \frac{k+1}{2^n} I_{\{\frac{k}{2^n} < U \leq \frac{k+1}{2^n}\}}$$

$$V_n = \sum_{k=0}^{\lfloor P \cdot 2^n - 1 \rfloor} \frac{k+1}{2^n} I_{\{\frac{k}{2^n} < V \leq \frac{k+1}{2^n}\}}$$

那么  $U_n \searrow U, V_n \searrow V$ , 结合  $(Z_t)$  的右连续性, 有

$$Z_{U_n} \rightarrow Z_U, Z_{V_n} \rightarrow Z_V$$

由离散时间上鞅的 Doob 择停定理 (有界停时),

$$Z_{U_{n+1}} \geq \mathbb{E}[Z_{U_n} | \mathcal{F}_{U_{n+1}}]$$

此时, 令  $Y_n = Z_{U_n}$ , 以及  $\mathcal{H}_n = \mathcal{F}_{U_n}$ , 那么  $(Y_n)$  是一个向后上鞅, 所以

$$\mathbb{E}[Y_n] = \mathbb{E}[Z_{U_n}] \leq \mathbb{E}[Z_0] < +\infty$$

所以  $\sup_n \mathbb{E}[Y_n] < +\infty$ , 结合鞅收敛定理, 可得  $Z_{U_n} \xrightarrow{L^1} Z_U$  和  $Z_{V_n} \xrightarrow{L^1} Z_V$ , 由离散时间上鞅的 Doob 择停定理,

$$\mathbb{E}[Z_{U_n}] \geq \mathbb{E}[Z_{V_n}]$$

令  $n \rightarrow \infty$  可得  $\mathbb{E}[Z_U] = \mathbb{E}[Z_V]$ .

现在考虑  $U, V$  不有界的情况, 对于  $P \geq 1$ , 考虑  $U \wedge P$  和  $V \wedge P$  是两个有界的停时, 那么

$$\mathbb{E}[Z_{U \wedge P}] \leq \mathbb{E}[Z_0], \mathbb{E}[Z_{V \wedge P}] \leq \mathbb{E}[Z_0]$$

令  $P \rightarrow \infty$ , 由 Fatou 引理, 可得

$$\mathbb{E}[Z_U], \mathbb{E}[Z_V] \leq \mathbb{E}[Z_0] < +\infty,$$

由此可知  $Z_U, Z_V \in L^1$ .

接下来我们证明  $Z_U \geq \mathbb{E}[Z_V | \mathcal{F}_U]$ , 其等价于证明  $\forall A \in \mathcal{F}_U$ ,

$$\mathbb{E}[Z_U I_A] \geq \mathbb{E}[Z_V I_A]$$

对  $A \in \mathcal{F}_U \subset \mathcal{F}_V$ , 定义:

$$U^A = \begin{cases} U, & \omega \in A \\ \infty, & \omega \in A^c \end{cases}, \quad V^A = \begin{cases} V, & \omega \in A \\ \infty, & \omega \in A^c \end{cases}$$

那么  $U^A$  和  $V^A$  都是停时且其满足  $U^A \leq V^A$ .

对任意的  $P \geq 1$ , 有

$$U^A \wedge P \leq V^A \wedge P$$

因此

$$\mathbb{E}[Z_{U^A \wedge P}] \geq \mathbb{E}[Z_{V^A \wedge P}]$$

事实上,

$$\begin{aligned} \text{LHS} &= \mathbb{E}[Z_{U^A \wedge P} I_{A^c}] + \mathbb{E}[Z_{U^A \wedge P} I_A] \\ &= \mathbb{E}[Z_P I_{A^c}] + \mathbb{E}[Z_U I_A I\{U \leq P\}] + \mathbb{E}[Z_P I_A I\{U > P\}] \end{aligned}$$

同理,

$$\begin{aligned} \text{RHS} &= E[Z_P I_{A^c}] + E[Z_{V \wedge P} I_A I\{U \leq P\}] \\ &\quad + E[Z_P I_A I\{U > P\}] \end{aligned}$$

所以有

$$E[Z_U I_A I\{U \leq P\}] \geq E[Z_{V \wedge P} I_A I\{U \leq P\}]$$

令  $p \rightarrow \infty$ , 由 Fatou 引理可得

$$\begin{aligned} E[Z_U I_A I\{U < \infty\}] &\geq \lim_{P \rightarrow \infty} E[Z_{V \wedge P} I_A I\{U \leq P\}] \\ &\geq E[\lim_{P \rightarrow \infty} Z_{V \wedge P} I_A I\{U \leq P\}] \\ &= E[Z_V I_A I\{U < \infty\}] \end{aligned}$$

另一方面, 由于  $U = \infty \implies V = \infty$ , 所以有

$$\begin{aligned} E[Z_U I_A I\{U = \infty\}] &= E[Z_V I_A I\{V = \infty\}] \\ &= E[Z_\infty I_A I\{U = \infty\}] \end{aligned}$$

即  $E[Z_U I_A] \geq E[Z_V I_A]$ , 其与之前的式子相加, 即可得到结论。 □

## 5.6 例子与习题

这一章的习题大概分为两类，一类是各种收敛性可积性的证明，另一类则是利用鞅的择停定理（总结：一致可积鞅的停止鞅也一致可积；右连续的前提下，一致可积鞅/停时有界/非负（上）鞅，可以使用择停定理，注意非负（上）鞅结论是不等号）的应用。后者往往技巧性较强，记一下例 5.3.1 以及其延伸出的两个例子里面鞅的构造就行了，更抽象的构造要是考了那确实没办法。

下面是一个上课讲过的例子。

### 例 5.6.1.

$B = (B_t)_{t \geq 0}$  是布朗运动，对于常数  $c$ ，记停时

$$T_c = \inf\{t \geq 0 : B_t = c\}$$

即首次到达  $c$  的时刻。设  $a < 0 < b$ ， $T = T_a \wedge T_b$ ，

- (1). 求  $\mathbb{P}(T_a < T_b)$ .
- (2). 求  $\mathbb{E}[T]$ .
- (3). 求  $\mathbb{E}[e^{-\lambda T_a}]$ ，其中  $\lambda > 0$ .
- (4). 设  $a + b = 0$ ，即

$$T = \inf\{t \geq 0 : |B_t| = b\}$$

求  $\mathbb{E}[T^2]$ .

解答：

- (1). 我们注意到鞅  $B_{t \wedge T} \in [a, b]$  有界，从而一致可积，由择停定理可知

$$0 = \mathbb{E}[B_{0 \wedge T}] = \mathbb{E}[B_{\infty \wedge T}] = \mathbb{E}[B_T] = a \cdot \mathbb{P}(T_a < T_b) + b \cdot \mathbb{P}(T_a > T_b)$$

而且  $\mathbb{P}(T_a < T_b) + \mathbb{P}(T_a > T_b) = 1$ ，解得

$$\mathbb{P}(T_a < T_b) = \frac{b}{b-a}, \quad \mathbb{P}(T_a > T_b) = \frac{-a}{b-a}$$

- (2). 注意到  $X_t = B_t^2 - t$  是一个鞅，那么  $X_{t \wedge T}$  是一个鞅，可知

$$0 = \mathbb{E}[X_{0 \wedge T}] = \mathbb{E}[X_{t \wedge T}]$$

所以

$$\mathbb{E}[t \wedge T] = \mathbb{E}[B_{t \wedge T}^2]$$

令  $t \rightarrow \infty$ ，因为  $B_{t \wedge T}^2 \in [0, a^2 \vee b^2]$  有界， $t \wedge T$  单调，利用控制收敛定理和单调收敛定理可知

$$\mathbb{E}[T] = \mathbb{E}[B_T^2] = a^2 \mathbb{P}(T_a < T_b) + b^2 \mathbb{P}(T_a > T_b) = -ab$$

- (3). 对于  $\theta > 0$ ，考虑如下的鞅：

$$N_t = \exp\left\{\theta B_t - \frac{\theta^2}{2}t\right\}$$

那么

$$N_{t \wedge T_a} \leq \exp\{\theta B_{t \wedge T_a}\} \leq \exp\{\theta a\}$$

一致有界  $\Rightarrow (N_{t \wedge T_a})$  一致可积，由 Doob 择停定理可知

$$1 = \mathbb{E}[N_0] = \mathbb{E}[N_{T_a}] = \mathbb{E}\left[\exp\left\{\theta a - \frac{\theta^2}{2}T_a\right\}\right]$$

即

$$\mathbb{E}\left[\exp\left\{-\frac{1}{2}\theta^2 T_a\right\}\right] = \exp\{-\theta a\}$$

取  $\theta = \sqrt{2\lambda}$  就得到

$$\mathbb{E}[e^{-\lambda T_a}] = e^{-a\sqrt{2\lambda}}$$

(4). 考虑如下的鞅:

$$K_t = B_t^4 - 6B_t^2t + 3t^2$$

具体过程省略。

### 题目 5.1. le gall(Exercise 3.26)

1.  $M = (M_t)_{t \geq 0}$  是非负鞅, 具有右连续轨道,  $M_0 = x$  a.s., 并且

$$\lim_{t \rightarrow \infty} M_t = 0 \text{ a.s.}$$

证明: 对于  $\forall y > x$ ,

$$\mathbb{P}(\sup_{t \geq 0} M_t \geq y) = \frac{x}{y}$$

2.  $B = (B_t)_{t \geq 0}$  是初值为  $x > 0$  的布朗运动,  $T_0 = \inf\{t \geq 0 : B_t = 0\}$ , 求

$$\sup_{t \leq T_0} B_t$$

的分布。

3.  $B = (B_t)_{t \geq 0}$  是初值为 0 的布朗运动,  $\mu > 0$ , 利用合适的指数鞅, 证明

$$\sup_{t \geq 0} (B_t - \mu t)$$

满足参数为  $2\mu$  的指数分布。

解答:

1. 设

$$T_y = \inf\{t \geq 0 : M_t \geq y\}$$

因为  $M_t$  具有(右)连续轨道, 所以  $(M_{t \wedge T_y})$  也是一个鞅, 那么

$$\mathbb{E}[M_{t \wedge T_y}] = \mathbb{E}[M_{0 \wedge T_y}] = x$$

而  $M_{t \wedge y} \in [0, y]$ , 所以由控制收敛定理可得

$$\mathbb{E}[M_{T_y} I_{\{T_y < +\infty\}} + M_\infty I_{\{T_y = \infty\}}] = \mathbb{E}[y I_{\{T_y < +\infty\}}] = x$$

从而

$$\mathbb{P}(\sup_{t \geq 0} M_t \geq y) = \mathbb{P}(T_y < +\infty) = \frac{x}{y}$$

2. 注意到  $(B_{t \wedge T_0})$  满足 1. 中条件, 所以  $\forall y \geq 0$ ,

$$\mathbb{P}(\sup_{t \leq T_0} B_t \geq y) = \mathbb{P}(\sup_{t \geq 0} B_{t \wedge T_0} \geq y) = \frac{x}{y}$$

3. 观察  $\sup_{t \geq 0} (B_t - \mu t)$ , 我们可以作以下变换:

$$\begin{aligned} \mathbb{P}(\sup_{t \geq 0} (B_t - \mu t) \geq y) &= \mathbb{P}(\sup_{t \geq 0} (B_{\frac{1}{4\mu^2}t} - \mu \cdot \frac{1}{4\mu^2}t) \geq y) \\ &= \mathbb{P}(\sup_{t \geq 0} (2\mu B_{\frac{1}{4\mu^2}t} - \frac{1}{2}t) \geq 2\mu y) \end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}(\sup_{t \geq 0} (B_t - \frac{1}{2}t) \geq 2\mu y) \\
&= \mathbb{P}(\sup_{t \geq 0} e^{B_t - \frac{1}{2}t} \geq e^{2\mu y})
\end{aligned}$$

其中  $N_t = e^{B_t - \frac{1}{2}t}$  是一个非负连续鞅, 并且  $N_t \xrightarrow{a.s.} 0$ ,  $N_0 = 1$ , 满足 1. 中条件, 所以  $\forall y > 0$ ,

$$\mathbb{P}(\sup_{t \geq 0} (B_t - \mu t) \geq y) = \mathbb{P}(\sup_{t \geq 0} N_t \geq e^{2\mu y}) = e^{-2\mu y}$$

$y \leq 0$  时显然概率为 1, 于是  $\sup_{t \geq 0} (B_t - \mu t)$  满足参数为  $2\mu$  的指数分布。

### 题目 5.2. le gall(Exercise 3.27)

$B = (B_t)_{t \geq 0}$  是初值为 0 的布朗运动,  $(\mathcal{F}_t = \sigma(B_s, s \in [0, t]))$ , 记

$$T_x = \inf\{t \geq 0 : B_t = x\}$$

并设  $T = T_a \wedge T_b$ , 其中  $a < 0 < b$  为常数, 对于  $\lambda > 0$ , 求:

1.  $\mathbb{E}[e^{-\lambda T}]$ .
2.  $\mathbb{E}[e^{-\lambda T} I_{\{T=T_a\}}]$ .
3.  $\mathbb{P}(T_a < T_b)$ .

解答:

1. 注意到

$$U_t = e^{\sqrt{2\lambda}B_t - \lambda t}, \quad V_t = e^{-\sqrt{2\lambda}B_t - \lambda t}$$

都是鞅, 那么对于  $\alpha > 0$  (待定),

$$M_t = e^{-\alpha\sqrt{2\lambda}}U_t + e^{\alpha\sqrt{2\lambda}}V_t$$

也是鞅。同时,

$$0 \leq U_{t \wedge T} \leq e^{b\sqrt{2\lambda}}, \quad 0 \leq V_{t \wedge T} \leq e^{-a\sqrt{2\lambda}}$$

说明  $(U_{t \wedge T})$  和  $(V_{t \wedge T})$  一致可积, 进而  $(M_{t \wedge T})$  也一致可积, 因此由 Doob 择停定理,

$$\begin{aligned}
e^{-\alpha\sqrt{2\lambda}} + e^{\alpha\sqrt{2\lambda}} &= \mathbb{E}[M_0] = \mathbb{E}[M_T] \\
&= e^{-\alpha\sqrt{2\lambda}}\mathbb{E}[U_T] + e^{\alpha\sqrt{2\lambda}}\mathbb{E}[V_T] \\
&= e^{-\alpha\sqrt{2\lambda}}\mathbb{E}[e^{\sqrt{2\lambda}\cdot a - \lambda T} I_{\{T_a < T_b\}} + e^{\sqrt{2\lambda}\cdot b - \lambda T} I_{\{T_a > T_b\}}] \\
&\quad + e^{\alpha\sqrt{2\lambda}}\mathbb{E}[e^{-\sqrt{2\lambda}\cdot a - \lambda T} I_{\{T_a < T_b\}} + e^{-\sqrt{2\lambda}\cdot b - \lambda T} I_{\{T_a > T_b\}}]
\end{aligned}$$

我们这里令  $\alpha = \frac{1}{2}(a+b)$ , 就得到

$$\begin{aligned}
\mathbb{E}[M_T] &= \mathbb{E}[e^{\sqrt{2\lambda}\cdot \frac{a+b}{2} - \lambda T} I_{\{T_a < T_b\}} + e^{\sqrt{2\lambda}\cdot \frac{b-a}{2} - \lambda T} I_{\{T_a > T_b\}}] \\
&\quad + \mathbb{E}[e^{\sqrt{2\lambda}\cdot \frac{b-a}{2} - \lambda T} I_{\{T_a < T_b\}} + e^{\sqrt{2\lambda}\cdot \frac{a+b}{2} - \lambda T} I_{\{T_a > T_b\}}] \\
&= (e^{\sqrt{2\lambda}\cdot \frac{a+b}{2}} + e^{\sqrt{2\lambda}\cdot \frac{b-a}{2}}) \mathbb{E}[e^{-\lambda T}]
\end{aligned}$$

从而

$$\mathbb{E}[e^{-\lambda T}] = \frac{e^{-\frac{a+b}{2}\sqrt{2\lambda}} + e^{\frac{a+b}{2}\sqrt{2\lambda}}}{e^{\sqrt{2\lambda}\cdot \frac{a-b}{2}} + e^{\sqrt{2\lambda}\cdot \frac{b-a}{2}}} = \frac{\cosh(\frac{a+b}{2}\sqrt{2\lambda})}{\cosh(\frac{b-a}{2}\sqrt{2\lambda})}$$

2. 令

$$N_t = e^{-\alpha\sqrt{2\lambda}}U_t - e^{\alpha\sqrt{2\lambda}}V_t$$

取  $\alpha = \frac{1}{2}(a+b)$ , 类似于上一问, 可求得

$$\mathbb{E}[e^{-\lambda T} I_{\{T=T_a\}}] = \frac{\sinh(b\sqrt{2\lambda})}{\sinh((b-a)\sqrt{2\lambda})}$$

3. 利用 2. 中的结论, 令  $\lambda \rightarrow 0^+$ , 由 DCT 可得

$$\mathbb{P}(T_a < T_b) = \lim_{\lambda \rightarrow 0^+} \frac{\sinh(b\sqrt{2\lambda})}{\sinh((b-a)\sqrt{2\lambda})} = \frac{b}{b-a}$$

**注** 这个构造方法多少有点逆天了, 看看就好。

### 题目 5.3. le gall(Exercise 3.28)

$B = (B_t)_{t \geq 0}$  是初值为 0 的布朗运动,  $(\mathcal{F}_t = \sigma(B_s, s \in [0, t]))$ ,  $a > 0$ , 设

$$\sigma_a = \inf\{t \geq 0 : B_t \leq t - a\}$$

1. 证明  $\sigma_a$  是停时, 且  $\sigma_a < +\infty$  a.s.
2. 利用合适的指数鞅, 证明:  $\forall \lambda \geq 0$ ,

$$\mathbb{E}[e^{-\lambda \sigma_a}] = \exp\{-a(\sqrt{1+2\lambda}-1)\}$$

(实际上, 对于  $\lambda \in [-\frac{1}{2}, 0]$  结论也成立, 证明涉及到分析的知识所以较为复杂, 此处不作要求, 但下一问仍可使用此结论。)

3. 对于  $\mu \in \mathbb{R}$ , 令  $M_t = \exp\{\mu B_t - \frac{1}{2}\mu^2 t\}$ , 证明停止鞅  $M_{\sigma_a \wedge t}$  是闭的当且仅当  $\mu \leq 1$ . (提示: 仿照定理 5.5.3 的过程, 证明  $\mathbb{E}[M_\sigma] = 1 \Rightarrow M_{\sigma_a \wedge t}$  闭)

**解答:**

1. 注意到  $\liminf_{t \rightarrow \infty} (B_t - t) = -\infty$  即可。
2. 先取  $\mu \in \mathbb{R}$  (待定), 考虑如下的鞅

$$M_t = e^{\mu B_t - \frac{1}{2}\mu^2 t}$$

那么

$$M_{t \wedge \sigma_a} \leq e^{\mu((t \wedge \sigma_a) - a) - \frac{1}{2}\mu^2(t \wedge \sigma_a)} = e^{-\mu a} e^{(-\frac{1}{2}\mu^2 + \mu)(t \wedge \sigma_a)}$$

于是我们后面取  $\mu$  时需要确保  $-\frac{1}{2}\mu^2 + \mu \leq 0$ , 这样的话就有

$$M_{t \wedge \sigma_a} \leq e^{-\mu a}$$

所以  $(M_{t \wedge \sigma_a})$  一致可积, 由 Doob 择停定理可得

$$1 = \mathbb{E}[M_0] = \mathbb{E}[M_{\sigma_a}] = \mathbb{E}[e^{\mu(\sigma_a - a) - \frac{1}{2}\mu^2 \sigma_a}]$$

这就说明

$$\mathbb{E}[e^{(-\frac{1}{2}\mu^2 + \mu)\sigma_a}] = e^{\mu a}$$

我们取  $\mu$  使得

$$-\frac{1}{2}\mu^2 + \mu = -\lambda \leq 0$$

则得到结论。

3. 考虑

$$\mathbb{E}[M_{\sigma_a}] = \mathbb{E}[e^{-(\frac{1}{2}\mu^2 - \mu)\sigma_a - \mu a}] = e^{-\mu a} \mathbb{E}[e^{-(\frac{1}{2}\mu^2 - \mu)\sigma_a}]$$

注意到其中的  $\frac{1}{2}\mu^2 - \mu \geq -\frac{1}{2}$ , 套用上一问结论可得

$$\mathbb{E}[M_{\sigma_a}] = e^{-\mu a} \cdot e^{-a(|\mu-1|-1)} = e^{-a(|\mu-1|-1+\mu)}$$

所以  $\mathbb{E}[M_{\sigma_a}] = 1 \Leftrightarrow \mu \leq 1$ . 如果  $M_{t \wedge \sigma_a}$  闭, 那么

$$1 = \mathbb{E}[M_{0 \wedge \sigma_a}] = \mathbb{E}[M_{\infty \wedge \sigma_a}] = \mathbb{E}[M_{\sigma_a}]$$

因此, 我们只剩下证明  $\mathbb{E}[M_{\sigma}] = 1 \Rightarrow M_{\sigma_a \wedge t}$  闭, 我们希望证明:  $\forall t \geq 0$ ,

$$M_{t \wedge \sigma_a} = \mathbb{E}[M_{\sigma_a} | \mathcal{F}_t]$$

首先,  $M_{t \wedge \sigma_a} \in \mathcal{F}_{t \wedge \sigma_a} \subset \mathcal{F}_t$ , 故只需证  $\forall A \in \mathcal{F}_t$ ,

$$\mathbb{E}[M_{t \wedge \sigma_a} I_A] = \mathbb{E}[M_{\sigma_a} I_A]$$

也就是

$$\mathbb{E}[M_{t \wedge \sigma_a} I_{A \cap \{\sigma_a \geq t\}}] = \mathbb{E}[M_{\sigma_a} I_{A \cap \{\sigma_a \geq t\}}] \quad (*)$$

然后我们考虑  $M_t$ , 它是一个非负鞅 (进而是非负上鞅), 套用非负上鞅的择停定理 (定理 5.5.4) 可得

$$M_{t \wedge \sigma_a} \geq \mathbb{E}[M_{\sigma_a} | \mathcal{F}_{t \wedge \sigma_a}]$$

但因为  $\mathbb{E}[M_{\sigma}] = 1$ , 我们发现

$$\mathbb{E}[M_{t \wedge \sigma_a}] = \mathbb{E}[M_{0 \wedge \sigma_a}] = 1 = \mathbb{E}[M_{\sigma}] = \mathbb{E}[\mathbb{E}[M_{\sigma_a} | \mathcal{F}_{t \wedge \sigma_a}]]$$

因此

$$M_{t \wedge \sigma_a} = \mathbb{E}[M_{\sigma_a} | \mathcal{F}_{t \wedge \sigma_a}]$$

容易验证  $A \cap \{\sigma_a \geq t\} \in \mathcal{F}_{t \wedge \sigma_a}$ , 根据条件期望的定义可得 (\*) 式成立。

#### 题目 5.4. 1e gall(Exercise 3.29.1-2)

$(Y_t)$  是一致可积鞅, 且具有右连续轨道,  $Y_0 = 0$  a.s., 设  $Y_{\infty} = \lim_{t \rightarrow \infty} Y_t$ , 固定  $p \geq 1$ , 我们称鞅  $Y$  满足性质 (P) 是指: 存在常数  $C$  使得对于任意停时  $T$  都有

$$\mathbb{E}[|Y_{\infty} - Y_T|^p | \mathcal{F}_T] \leq C$$

1. 证明如果  $Y_{\infty}$  有界, 则  $Y$  满足性质 (P).
2.  $B = (B_t)_{t \geq 0}$  是初值为 0 的布朗运动,  $(\mathcal{F}_t = \sigma(B_s, s \in [0, t]))$ , 证明鞅  $Y_t = B_{t \wedge 1}$  满足性质 (P).

**证明**

1. 不妨设  $|Y_{\infty}| < C$ ,  $(Y_t)$  一致可积, 于是  $Y_t = \mathbb{E}[Y_{\infty} | \mathcal{F}_t]$ , 那么

$$|Y_t| = |\mathbb{E}[Y_{\infty} | \mathcal{F}_t]| \leq \mathbb{E}[|Y_{\infty}| | \mathcal{F}_t] < C, \forall t \geq 0$$

从而  $|Y_T| < C$ . 那么我们就得到

$$\mathbb{E}[|Y_{\infty} - Y_T|^p | \mathcal{F}_T] \leq \mathbb{E}[(|Y_{\infty}| + |Y_T|)^p | \mathcal{F}_T] \leq (2C)^p$$

2. 先考虑  $p = 1$  的情形, 任取  $A \in \mathcal{F}_T$ ,

$$\mathbb{E}[\mathbb{E}[|Y_{\infty} - Y_T| | \mathcal{F}_T] I_A] = \mathbb{E}[|Y_{\infty} - Y_T| I_A] \leq \mathbb{E}[|Y_{\infty}| I_A] + \mathbb{E}[|Y_T| I_A]$$

同时由一致可积鞅的择停定理,  $Y_T = \mathbb{E}[Y_{\infty} | \mathcal{F}_T]$ , 因此

$$\begin{aligned} \mathbb{E}[|Y_T| I_A] &= \mathbb{E}[|\mathbb{E}[Y_{\infty} - Y_T]| I_A] \\ &\leq \mathbb{E}[\mathbb{E}[|Y_{\infty}| | \mathcal{F}_T] I_A] \\ &= \mathbb{E}[|Y_{\infty}| I_A] \end{aligned}$$

从而

$$\mathbb{E}[\mathbb{E}[|Y_{\infty} - Y_T| | \mathcal{F}_T] I_A] \leq 2\mathbb{E}[|Y_{\infty}| I_A] \leq 2\mathbb{E}[|Y_{\infty}|]$$

取  $C = 2\mathbb{E}[|Y_\infty|]$  即可。然后我们考虑  $p > 1$ , 根据  $L^p$  最大值不等式,

$$\mathbb{E}[\sup_{t \geq 0} |Y_t|^p] \leq \mathbb{E}[\sup_{0 \leq t \leq 1} |B_t|^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|B_1|^p]$$

因此  $\sup_{t \geq 0} |Y_t|^p \in L^p$ . 因此, 对于任意的  $A \in \mathcal{F}_T$ ,

$$\begin{aligned} \mathbb{E}[\mathbb{E}[|Y_\infty - Y_T|^p | \mathcal{F}_T] I_A] &= \mathbb{E}[|Y_\infty - Y_T|^p I_A] \\ &\leq \mathbb{E}[(|Y_\infty| + |Y_T|)^p I_A] \\ &= \mathbb{E}[(2 \sup_{t \geq 0} |Y_t|)^p I_A] \\ &= 2^p \mathbb{E}[(\sup_{t \geq 0} |Y_t|)^p I_A] \\ &\leq 2^p \mathbb{E}[\sup_{t \geq 0} |Y_t|^p] \leq 2^p \left(\frac{p}{p-1}\right)^p \mathbb{E}[|B_1|^p] < +\infty \end{aligned}$$

那么由  $A$  的任意性,

$$\mathbb{E}[|Y_\infty - Y_T|^p | \mathcal{F}_T] \leq 2^p \left(\frac{p}{p-1}\right)^p \mathbb{E}[|B_1|^p]$$

后者是一个固定的常数, 结论得证. □

最后放两道去年期末题。第一道的解法应该是类比于定理 5.4.3 的证明, 第二道则是对题目 5.3 的改编。

#### 题目 5.5. 2023SPFinal4.

$(X_t)$  是非负上鞅, 且具有右连续轨道, 证明: 存在  $X_\infty \in L^1$  使得  $X_t \xrightarrow{\text{a.s.}} X_\infty$ , 并且  $\mathbb{E}[X_\infty | \mathcal{F}_t] \leq X_t$ .

解答: 取  $\mathbb{R}_+$  的可数稠密子集  $D$ , 对于  $\forall t \in D$ , 有理数  $a < b$ , 由离散鞅的上穿次数估计不等式:

$$\mathbb{E}[M_{a,b}^X(D \cap [0, t])] \leq \frac{1}{b-a} \mathbb{E}[(X_t - a)^-]$$

由于  $X_t$  非负,  $\mathbb{E}[(X_t - a)^-] \leq \mathbb{E}[|a|] = |a|$ , 所以

$$\mathbb{E}[M_{a,b}^X(D \cap [0, t])] \leq \frac{|a|}{b-a} < +\infty$$

那么由 Fatou 引理,

$$\mathbb{E}[M_{a,b}^X(D)] \leq \liminf_{D \ni t \rightarrow \infty} \mathbb{E}[M_{a,b}^X(D \cap [0, t])] \leq \frac{|a|}{b-a} < +\infty$$

从而  $\lim_{D \ni t \rightarrow \infty} X_t$  a.s. 存在。现在将收敛性扩展到  $\mathbb{R}_+$  上, 对于  $\forall \varepsilon > 0$ , 存在  $N$  使得  $D \ni t \geq N$  时,

$$|X_t - X_\infty| \leq \varepsilon$$

那么, 对于任意的  $s \geq N$ , ( $D$  稠密) 取  $D \ni s_n \searrow s$ , 由右连续性得到

$$|X_s - X_\infty| = \lim_{n \rightarrow \infty} |X_{s_n} - X_\infty| \leq \varepsilon$$

这就说明了

$$X_\infty = \lim_{t \rightarrow \infty} X_t \text{ a.s. 存在}$$

由 Fatou 引理,

$$\mathbb{E}[|X_\infty|] \leq \liminf_{D \ni t \rightarrow \infty} \mathbb{E}[|X_t|] < +\infty$$

因此  $X_\infty \in L^1$ .

最后, 考虑常数  $t < s$ , 根据  $X$  是上鞅以及条件期望的 Fatou 引理,

$$\mathbb{E}[X_\infty | \mathcal{F}_t] \leq \liminf_{s \rightarrow \infty} \mathbb{E}[X_s | \mathcal{F}_t] \leq X_t$$

(或者考虑常数  $t$  和  $\infty$  作为停时利用非负上鞅的择停定理。)

## 题目 5.6. 2023SPFinal.5

$B = (B_t)_{t \geq 0}$  是布朗运动, 设停时  $T = \inf\{t \geq 0 : B_t \geq 1 - 2t\}$ ,

- (1). 证明  $T < +\infty$  a.s.
- (2). 求  $\mathbb{E}[e^{-\lambda T}]$ .

**证明** (1). 设  $T_1 = \inf\{t \geq 0 : B_t \geq 1\}$ , 则  $\{T_1 < +\infty\} \subset \{T < +\infty\}$ , 由于  $T_1 < +\infty$  a.s., 所以  $T < +\infty$  a.s.

(2). 对于任意的  $\theta \in \mathbb{R}$ , 令

$$X_t = \exp\left\{\theta B_t - \frac{1}{2}\theta^2 t\right\}$$

容易验证  $\{X_t, t \geq 0\}$  关于  $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$  是一个鞅, 我们令  $\theta > 0$ , 这确保了

$$0 \leq X_{t \wedge T} = \exp\left\{\theta(B_{t \wedge T}) - \frac{1}{2}\theta^2 T\right\} \leq \exp\left\{\theta(1 - 2T) - \frac{1}{2}\theta^2 T\right\} = \exp\left\{\theta - \left(\frac{1}{2}\theta^2 + 2\theta\right)T\right\} \leq \exp\{\theta\}$$

一致有界从而一致可积, 由择停定理可知

$$1 = \mathbb{E}[X_0] = \mathbb{E}[X_T] = e^\theta \mathbb{E}[e^{-(\frac{1}{2}\theta^2 + 2\theta)T}]$$

我们取  $\theta = \sqrt{2\lambda + 4} - 2 > 0$ , 就得到

$$\mathbb{E}[\exp\{-\lambda T\}] = e^{-\theta} = e^{2 - \sqrt{2\lambda + 4}}$$

□

## 第 6 章 连续时间马氏过程

本章节由两部分组成，不标注星号的内容来自研随课程，标注星号的内容来自应随课程。前者的内容标题应该叫做“连续马氏过程的一般理论”，用比较基础的语言叙述了连续情形的马氏过程，并以 Levy 过程等为例给出了详细介绍，能够帮助我们更好地引入后续内容。后者则着重分析了离散状态空间连续时间马氏链（跳跃过程）的性质。

### 6.1 马氏过程的定义

#### 6.1.1 转移核与转移半群

##### 定义 6.1.1 (转移核)

$(E, \mathcal{E})$  是一个可测空间，映射  $Q: E \times \mathcal{E} \rightarrow [0, 1]$  如果满足以下性质：

- (1).  $\forall x \in E$ ,  $Q(x, \cdot)$  作为  $\mathcal{E} \rightarrow [0, 1]$  的映射是一个  $(E, \mathcal{E})$  上的概率测度。
- (2).  $\forall A \in \mathcal{E}$ ,  $Q(\cdot, A)$  作为  $E \rightarrow [0, 1]$  的映射是  $\mathcal{E}$ -可测的。

那么就称  $Q$  是一个从  $E$  到  $E$  的马氏转移核 (Markovian transition kernel)，以下简称转移核。

转移核的概念类似于离散情形下的转移概率函数。

##### 定义 6.1.2 (转移半群)

$(Q_t)_{t \geq 0}$  是  $E$  上的转移核组成的集合，如果其满足：

- (1).  $\forall x \in E$ ,  $Q_0(x, dy) = \delta_x(dy)$ ，这里  $\delta_x$  是  $x$  处的独点分布，即

$$\int I_A Q_0(x, dy) = \begin{cases} 1 & , x \in A \\ 0 & , x \notin A \end{cases}$$

- (2).  $\forall s, t \geq 0$ ,  $A \in \mathcal{E}$ ,

$$Q_{t+s}(x, A) = \int_E Q_t(x, dy) Q_s(y, A)$$

Chapman-Kolmogorov identity

- (3).  $\forall A \in \mathcal{E}$ , 映射  $(t, x) \mapsto Q_t(x, A)$  是关于  $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$  可测的。

则称  $(Q_t)_{t \geq 0}$  是一个转移半群 (transition semigroup)。

转移半群的概念类似于离散情形下的转移概率矩阵。

**定义 6.1.3 (转移算子)**

$Q$  是一个转移核, 如果映射  $f: E \rightarrow \mathbb{R}$  有界可测 (或者非负可测), 定义映射  $Qf: E \rightarrow \mathbb{R}$  为

$$Qf(x) = \int Q(x, dy)f(y)$$

那么  $Qf$  依然是有界可测 (或者非负可测) 的。记  $B(E)$  为  $E$  上的有界可测实函数全体构成的向量空间, 并给予范数

$$\|f\| = \sup\{|f(x)| : x \in E\}$$

那么就可以把  $Q$  视为一个  $B(E) \rightarrow B(E)$  的线性算子。

**推论 6.1.1**

根据  $0 \leq Q(x, A) \leq 1$  可知,  $Q$  是一个压缩映射, 算子范数  $\leq 1 < +\infty$ , 从而是连续算子。

**命题 6.1.1**

从线性算子的观点来看, 由 Chapman-Kolmogorov identity 可以得到

$$Q_{s+t} = Q_t Q_s, \forall s, t \geq 0$$

这里  $Q_t Q_s$  是指  $Q_t \circ Q_s$ , 即算子的复合。

**证明**  $\forall f \in B(E), x \in E$ , 由 Chapman-Kolmogorov identity,

$$\begin{aligned} Q_{s+t}f(x) &= \int_E Q_{s+t}(x, dz)f(z) \\ &= \int_E \left( \int_E Q_t(x, dy)Q_s(y, dz) \right) f(z) \\ &\quad (\text{交换积分顺序}) \\ &= \int_E Q_t(x, dy) \int_E Q_s(y, dz)f(z) \\ &= \int_E Q_t(x, dy)Q_s f(y) \\ &= Q_t Q_s f(x) \end{aligned}$$

□

这也表明  $Q_t Q_s = Q_s Q_t$ , 即复合是可交换的。

**6.1.2 马氏过程****定义 6.1.4 (马氏过程)**

$(Q_t)_{t \geq 0}$  是  $E$  上的转移半群, 考虑滤流概率空间  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  上的一个在  $E$  上取值的  $(\mathcal{F}_t)$ -适应过程  $X = (X_t)_{t \geq 0}$ , 如果其满足  $\forall s, t \geq 0, f \in B(E)$ ,

$$\mathbb{E}[f(X_{s+t}) | \mathcal{F}_s] = Q_t f(X_s)$$

那么就称  $X$  是 (关于滤流  $(\mathcal{F}_t)$  的、转移半群为  $(Q_t)$  的连续时间) 马氏过程。

在本章的剩余内容中, 如无特殊说明, 默认马氏过程  $X$  是关于滤流  $(\mathcal{F}_t^X)$  的。

如果取  $f = I_A$ , 我们得到

$$\mathbb{P}(X_{s+t} \in A | \mathcal{F}_s) = Q_t(X_s, A)$$

特别地,

$$\mathbb{P}(X_{s+t} \in A | X_r, 0 \leq r \leq s) = Q_t(X_s, A)$$

这表明, 给定“过去” $(X_r, 0 \leq r \leq s)$ 的条件下,  $X_{s+t}$ 的条件分布只取决于“现在”的 $X_s$ , 并由 $Q_t(X_s, \cdot)$ 给出, 这就是马氏性。

### 定理 6.1.1 (马氏过程的有限维分布)

设 $X_0$ 的分布为 $\gamma(dy)$ , 对于任意的 $0 = t_0 < t_1 < t_2 < \cdots < t_p$ , 我们有

$$\mathbb{P}(X_{t_0} \in A_0, \cdots, X_{t_p} \in A_p) = \int_{A_0} \gamma(dx) \int_{A_1} Q_{t_1}(x, dx_1) \int_{A_2} Q_{t_2-t_1}(x_1, dx_2) \cdots \int_{A_p} Q_{t_p-t_{p-1}}(x_{p-1}, dx_p)$$

更一般地, 有:

$$\mathbb{E}[f_0(X_{t_0}) \cdots f_p(X_{t_p})] = \int_E f_0(x) \gamma(dx) \int_E f_1(x_1) Q_{t_1}(x, dx_1) \cdots \int_E f_p(x_p) Q_{t_p-t_{p-1}}(x_{p-1}, dx_p)$$

**证明** 归纳: 先考虑 $p=0$ 的情形, 即

$$\mathbb{E}[f_0(X_{t_0})] = \int_E f_0(x) \gamma(dx)$$

命题成立; 再假设 $p-1$ 时命题成立, 考虑 $p$ 时的情形:

$$\begin{aligned} & \mathbb{E}[f_0(X_{t_0}) \cdots f_{p-1}(X_{t_{p-1}}) f_p(X_{t_p})] \\ &= \mathbb{E}[\mathbb{E}[f_0(X_{t_0}) \cdots f_{p-1}(X_{t_{p-1}}) f_p(X_{t_p}) | \mathcal{F}_{t_{p-1}}]] \\ &= \mathbb{E}[f_0(X_{t_0}) \cdots f_{p-1}(X_{t_{p-1}}) \mathbb{E}[f_p(X_{t_p}) | \mathcal{F}_{t_{p-1}}]] \\ &= \mathbb{E}[f_0(X_{t_0}) \cdots f_{p-1}(X_{t_{p-1}}) Q_{t_p-t_{p-1}} f_p(X_{t_p})] \\ &= \int_E f_0(x) \gamma(dx) \int_E f_1(x_1) Q_{t_1}(x, dx_1) \cdots \int_E f_{p-1}(x_{p-1}) Q_{t_p-t_{p-1}} f_p(x_{p-1}) Q_{t_{p-1}-t_{p-2}, dx_{p-1}} \\ &= \int_E f_0(x) \gamma(dx) \int_E f_1(x_1) Q_{t_1}(x, dx_1) \cdots \int_E f_{p-1}(x_{p-1}) Q_{t_{p-1}-t_{p-2}, dx_{p-1}} \int_E Q_{t_p-t_{p-1}}(x_{p-1}, dx_p) f_p(x_p) \end{aligned}$$

成立。综上结论得证。  $\square$

这个定理表明, 马氏过程的有限维分布完全由其初始 $X_0$ 的分布和转移半群决定。那么, 给定转移半群 $(Q_t)_{t \geq 0}$ 之后, 我们取

$$X = \{(X_t)_{t \geq 0} : (X_t)_{t \geq 0} \text{是关于转移半群 } (Q_t)_{t \geq 0} \text{ 的马氏过程}\}$$

其中, 记初值为固定值 $X_0 = x$  a.s. 的过程为 $(X_t^x)_{t \geq 0}$ , 可以发现:

$$Q_t f(x) = Q_t f(X_0^x) = \mathbb{E}[f(X_t^x) | \mathcal{F}_0] = \mathbb{E}[f(X_t^x)] \stackrel{\text{def}}{=} \mathbb{E}_x[f(X_t)]$$

由此, 我们得到了转移算子的一个非常好用的表示方法。

注意, 我们接下来考虑马氏过程 $X = (X_t)_{t \geq 0}$ 的时候, 实际上考虑的是所有与转移半群 $(Q_t)_{t \geq 0}$ 有关的马氏过程, 也就是所有“给定初始分布和转移半群 $(Q_t)_{t \geq 0}$ 之后唯一确定的马氏过程”。如无特殊说明,  $X^x = (X_t^x)$ 就是指初值为固定值 $x$ 的马氏过程; 而如果在某个结论中没有指明初始分布, 而是表述为 $X = (X_t)$ , 就代表初始分布不影响此结论。

回忆离散马氏链时的情形, 当时我们也并没有给出一个具体的随机过程, 而是重点分析了转移概率和状态分类。这是因为, 时间齐次的马氏过程的关键不在于某一时刻的具体数值, 而是一个时间间隔内的行为。

### 例 6.1.1.

回顾布朗运动的定义, 我们当时规定了初值 $B_0 = 0$ , 但事实上, 初值可以换成任何一个固定的数值, 甚至是一个分布。因为布朗运动的核心并不在于初始, 而是“给定过去的分布, 一段时间后的变化”。

初值固定的布朗运动我们记作  $\{(B_t^x)_{t \geq 0} : x \in \mathbb{R}\}$ , 那么我们可以推导出布朗运动的转移半群为:

$$Q_t f(x) = \mathbb{E}[f(B_t^x)] = \mathbb{E}[f(B_t^0 + x)] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} f(y) dy$$

## 6.2 马氏过程的构造

事先给定一个马氏转移半群  $(Q_t)_{t \geq 0}$ , 记

$$\Omega^* = E^{\mathbb{R}^+} = \{\omega : \omega(\cdot) : [0, \infty) \rightarrow E\}$$

可以定义  $\Omega^*$  上的坐标过程 (Coordinate process)  $X = (X_t)$ , 定义为

$$X_t(\omega) = \omega(t) : \Omega^* \rightarrow E$$

记  $\mathcal{F}^* = \sigma(X_s, s \geq 0)$ ,  $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$ , 我们需要找到一个可测空间  $(\Omega^*, \mathcal{F}^*)$  上面的概率测度  $\mathbb{P}^*$ .

对任意的有限子集  $U = \{0 < t_1 < t_2 < \cdots < t_p\} \subset \mathbb{R}_+$ , 定义一个在  $E^U = \underbrace{E \times E \times \cdots \times E}_p$  上的概率测度, 如下:

$$\mu^U(A_1 \times A_2 \times \cdots \times A_p) = \int_E \gamma(dx) \int_{A_1} Q_{t_1}(x, dx_1) \cdots \int_{A_p} Q_{t_p - t_{p-1}}(x_{p-1}, dx_p)$$

记为  $\{\mu^U\}$ , 其中  $U$  是有限子集。

事实上,  $\{\mu^U\}$  具有相容性, 即: 对  $U \subset V$ , 定义

$$\pi_U^V : E^V \rightarrow E^U$$

其是一个投影 (或限制), 那么有

$$\mu^U = \mu^V \circ (\pi_U^V)^{-1}$$

结合相容性和 Kolmogorov 扩张定理, 一定存在一个  $(\Omega^*, \mathcal{F}^*)$  上的概率测度  $\mathbb{P}^*$ , 使得

$$\mathbb{P}^*(X_{t_1} \in A_1, \cdots, X_{t_p} \in A_p) = \mu^U(A_1 \times \cdots \times A_p)$$

在  $\mathbb{P}^*$  之下,  $(X_t)_{t \geq 0}$  是一个与  $(Q_t)$  相联系的马氏过程, 因为定理 6.1.1 中的等式成立, 而其可以推出马氏性。

## 6.3 马氏过程的其他要素

本小节我们主要介绍三个工具概念: 预解式、Feller 半群、生成元, 以及它们的性质与关系。用于进一步表述与研究马氏过程的性质。

### 6.3.1 预解式

#### 定义 6.3.1

设  $\lambda > 0$ , 半群  $(Q_t)_{t \geq 0}$  的  $\lambda$ -预解式 (resolvent) 是指线性算子  $R_\lambda : \mathcal{B}(E) \rightarrow \mathcal{B}(E)$ , 其满足

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} Q_t f(x) dt$$

预解式实际上是一个 Laplace 变换。

## 定理 6.3.1 (预解式的性质)

(1).  $\|R_\lambda f\| \leq \lambda^{-1} \|f\|$ , 这里的  $\|f\|$  为

$$\|f\| = \sup_{x \in E} |f(x)|$$

(2). 如果  $0 \leq f \leq 1$ , 则  $0 \leq \lambda R_\lambda f(x) \leq 1$ .(3). 预解恒等式:  $\lambda, \mu > 0$ , 则

$$R_\lambda - R_\mu + (\lambda - \mu)R_\lambda R_\mu = 0$$

**证明** 前两个由定义显然, 我们只说明 (3):  $\forall f \in B(E), x \in E$ , 不妨  $\lambda \neq \mu$ ,

$$\begin{aligned} R_\lambda R_\mu f(x) &= \mathbb{R}_\lambda(R_\mu f)(x) \\ &= \int_0^{+\infty} e^{-\lambda t} Q_t(R_\mu f)(x) dt \\ &= \int_0^{+\infty} e^{-\lambda t} \int_E Q_t(x, dy) R_\mu f(y) dt \\ &= \int_0^{+\infty} e^{-\lambda t} \int_E Q_t(x, dy) \int_0^{+\infty} e^{-\mu s} \int_E Q_s(y, dz) f(z) ds dt \end{aligned}$$

现在的积分顺序是  $dz \rightarrow ds \rightarrow dy \rightarrow dt$ , 我们交换  $ds$  和  $dy$  的顺序

$$\begin{aligned} &= \int_0^{+\infty} dt \int_0^{+\infty} ds \int_E Q_t(x, dy) \int_E Q_s(y, dz) f(z) e^{-\lambda t - \mu s} \\ &= \int_0^{+\infty} dt \int_0^{+\infty} ds Q_{s+t} f(x) e^{-\lambda t - \mu s} \end{aligned}$$

现在是我们熟悉的实数上的积分了, 换元  $r = s + t$ , 转化为区域  $t \in (0, r), r \in (0, +\infty)$  上的积分

$$\begin{aligned} &= \int_0^{+\infty} dr \int_0^r dt e^{-(\lambda - \mu)t} e^{-\mu r} Q_r f(x) \\ &= \int_0^{+\infty} \frac{e^{-\mu r} - e^{-\lambda r}}{\lambda - \mu} Q_r f(x) dr \\ &= \frac{R_\mu - R_\lambda}{\lambda - \mu} f(x) \end{aligned}$$

□

## 引理 6.3.1 (一个积分)

 $\lambda, \mu > 0$ ,

$$\int_0^\infty \exp\left\{-\lambda t - \frac{\mu}{t}\right\} \frac{1}{\sqrt{t}} dt = e^{-2\sqrt{\lambda\mu}} \sqrt{\frac{\pi}{\lambda}}$$

♥

**证明** 设原积分为  $I$ , 换元:

$$t = \frac{\sqrt{\mu}}{\sqrt{\lambda}} s, \quad dt = \frac{\sqrt{\mu}}{\sqrt{\lambda}} ds$$

则

$$I = \int_0^\infty \exp\left\{-\sqrt{\lambda\mu}(s + s^{-1})\right\} \frac{1}{\sqrt{s}} \sqrt{\frac{\sqrt{\mu}}{\sqrt{\lambda}}} ds$$

整理一下得到

$$\left(\frac{\lambda}{\mu}\right)^{\frac{1}{4}} I = \int_0^\infty \exp\left\{-\sqrt{\lambda\mu}(s + s^{-1})\right\} \frac{1}{\sqrt{s}} ds$$

换元:

$$s = u^2, \quad ds = 2udu$$

则

$$\frac{1}{2} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{4}} I = \int_0^{\infty} \exp \left\{ -\sqrt{\lambda\mu}(u^2 + u^{-2}) \right\} du$$

换元:

$$u = \frac{1}{v}, du = -\frac{1}{v^2} dv$$

则

$$\frac{1}{2} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{4}} I = \int_0^{\infty} \exp \left\{ -\sqrt{\lambda\mu}(v^2 + v^{-2}) \right\} \frac{1}{v^2} dv$$

相加可得

$$\begin{aligned} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{4}} I &= \int_0^{\infty} \exp \left\{ -\sqrt{\lambda\mu}(u^2 + u^{-2}) \right\} \left(1 + \frac{1}{u^2}\right) du \\ &= e^{-2\sqrt{\lambda\mu}} \int_0^{\infty} \exp \left\{ -\sqrt{\lambda\mu}(u - u^{-1})^2 \right\} \left(1 + \frac{1}{u^2}\right) du \end{aligned}$$

换元:

$$w = u - \frac{1}{u}, dw = \left(1 + \frac{1}{u^2}\right) du$$

则

$$\begin{aligned} e^{2\sqrt{\lambda\mu}} \left( \frac{\lambda}{\mu} \right)^{\frac{1}{4}} I &= \int_{-\infty}^{\infty} \exp \left\{ -\sqrt{\lambda\mu} w^2 \right\} dw \\ &= (\lambda\mu)^{-\frac{1}{4}} \sqrt{\pi} \end{aligned}$$

得到

$$I = e^{-2\sqrt{\lambda\mu}} \sqrt{\frac{\pi}{\lambda}}$$

□

### 例 6.3.1.

考虑布朗运动的转移半群:

$$Q_t(x, A) = \int_A \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{(y-x)^2}{2t} \right\} dy$$

$\lambda > 0$ , 求其预解式  $\mathbb{R}_\lambda$ .

**证明**

$$\begin{aligned} R_\lambda f(x) &= \int_0^{\infty} e^{-\lambda t} \int_E Q_t(x, dy) f(y) dt \\ &= \int_0^{\infty} e^{-\lambda t} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{(y-x)^2}{2t} \right\} f(y) dy dt \\ &\quad (\text{先对 } dt \text{ 积分}) \\ &= \int_{\mathbb{R}} f(y) dy \int_0^{\infty} e^{-\lambda t} \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{(y-x)^2}{2t} \right\} dt \\ &= \int_{\mathbb{R}} f(y) r_\lambda(y-x) dy \end{aligned}$$

由引理 6.3.1, 其中

$$r_\lambda(y-x) = \frac{1}{\sqrt{2\lambda}} \exp \{-|y-x|\sqrt{2\lambda}\}$$

□

**例 6.3.2.**

$(X_t)$  是关于半群  $(Q_t)$  的马氏过程, 如果  $h \in \mathcal{B}(E), h \geq 0, \lambda \geq 0$ , 则

$$M_t = e^{-\lambda t} R_\lambda h(X_t), t \geq 0$$

是一个上鞅。

**证明**  $M_t$  是关于  $X_t$  的函数, 所以  $M_t \in \mathcal{F}_t$ , 并且有界 (从而可积)。注意到  $\forall h \in \mathcal{B}(E), x \in E, s \geq 0$ ,

$$\begin{aligned} Q_s R_\lambda h(x) &= Q_s(R_\lambda h)(x) \\ &= \int_E Q_s(x, dy) R_\lambda h(y) \\ &= \int_E Q_s(x, dy) \int_0^\infty e^{-\lambda t} Q_t h(y) dt \\ &= \int_0^\infty e^{-\lambda t} Q_{s+t} h(x) dt \end{aligned}$$

从而可得

$$e^{-\lambda s} Q_s R_\lambda h(x) = \int_0^\infty e^{-\lambda(s+t)} Q_{s+t} h(x) dt = \int_s^\infty e^{-\lambda t} Q_t h(x) dt \leq R_\lambda h(x)$$

那么

$$\mathbb{E}[M_{s+t} | \mathcal{F}_t] = \mathbb{E}[e^{-\lambda(s+t)} R_\lambda h(X_{s+t}) | \mathcal{F}_t] = e^{-\lambda(s+t)} R_\lambda h(X_t) \leq e^{-\lambda t} R_\lambda h(X_t) = M_t$$

所以  $M_t$  是上鞅。 □

### 6.3.2 Feller 半群

现在我们假设  $E$  是一个 Polish 空间, 如果映射  $f : E \rightarrow \mathbb{R}$  满足:  $\forall \varepsilon > 0$ , 存在紧集  $K \subset E$  使得  $|f(x)| \leq \varepsilon, \forall x \in E \setminus K$ , 则称  $f$  在无穷远处趋于 0, 这样的  $f$  全体记作  $C_0(E)$ 。

取范数

$$\|f\| = \sup_{x \in E} |f(x)|$$

那么  $(C_0(E), \|\cdot\|)$  成为一个 Banach 空间。

**定义 6.3.2**

$(Q_t)_{t \geq 0}$  是转移半群, 称之为 Feller 半群, 如果:

- (1).  $\forall f \in C_0(E), Q_t f \in C_0(E)$ .
- (2).  $\forall f \in C_0(E), \|Q_t f - f\| \rightarrow 0$  as  $t \rightarrow 0$ .

如果一个马氏过程  $X$  的转移半群是 Feller 半群, 就称之为 Feller 过程。 ♣

**推论 6.3.1**

$(Q_t)_{t \geq 0}$  是 Feller 半群, 固定  $f \in C_0(E)$ , 映射

$$\mathbb{R}_+ \rightarrow C_0(E), t \mapsto Q_t f$$

是一致连续的。 ♥

**证明** 由定义 6.3.2(2),  $\forall s \geq 0$ , 因为  $Q_s$  是压缩映射,

$$\lim_{t \rightarrow 0^+} \|Q_{s+t} f - Q_s f\| = \lim_{t \rightarrow 0^+} \|Q_s(Q_t f - f)\| = 0$$

并注意到上述极限关于  $s \in \mathbb{R}_+$  一致收敛。 □

**推论 6.3.2**

$(Q_t)_{t \geq 0}$  是 Feller 半群, 其预解式满足:  $\forall \lambda > 0, f \in C_0(E)$ , 都有  $R_\lambda f \in C_0(E)$ .

**证明** 由定义 6.3.2(1) 和 DCT 即得. □

**定理 6.3.2**

对于 Feller 半群  $(Q_t)_{t \geq 0}$ , 任取  $\lambda > 0$ , 定义

$$R = \{R_\lambda f : f \in C_0(E)\}$$

则  $R$  不取决于  $\lambda$  的选取, 并且是  $C_0(E)$  的稠密子集.

**证明** 对于  $\lambda \neq \mu$ , 由预解恒等式,

$$R_\lambda f = R_\mu(f + (\mu - \lambda)R_\lambda f)$$

这说明  $R_\lambda f \in \{R_\mu f : f \in C_0(E)\}$ , 即  $R$  不取决于  $\lambda$  的选取.

$R$  是  $C_0(E)$  的线性子空间, 注意到

$$\lambda R_\lambda f = \lambda \int_0^\infty e^{-\lambda t} Q_t f dt = \int_0^\infty e^{-\lambda t} Q_{t\lambda^{-1}} f dt \rightarrow f \text{ as } \lambda \rightarrow \infty$$

所以稠密性得证. □

**6.3.3 生成元**

本小节我们默认在 Feller 半群  $(Q_t)_{t \geq 0}$  上讨论.

**定义 6.3.3**

令

$$D(L) = \{f \in C_0(E) : t \rightarrow 0^+ \text{ 时, } \frac{Q_t f - f}{t} \text{ 在 } C_0(E) \text{ 中收敛}\}$$

对于  $f \in D(L)$ , 则定义

$$Lf = \lim_{t \rightarrow 0^+} \frac{Q_t f - f}{t}$$

于是  $L$  成为一个线性算子, 称之为生成元(generator),  $D(L)$  称为  $L$  的定义域. ♣

**命题 6.3.1 (生成元与转移算子可交换)**

$f \in D(L)$ ,  $s > 0$ , 则  $Q_s f \in D(L)$ , 并且  $L(Q_s f) = Q_s(Lf)$ .

**证明** 由于  $Q_t Q_s$  可交换,

$$\frac{Q_t(Q_s f) - Q_s f}{t} = Q_s \left( \frac{Q_t f - f}{t} \right)$$

因为  $Q_s$  是一个连续算子, 所以

$$L(Q_s f) = \lim_{t \rightarrow 0^+} \frac{Q_t(Q_s f) - Q_s f}{t} = Q_s \left( \lim_{t \rightarrow 0^+} \frac{Q_t f - f}{t} \right) = Q_s(Lf)$$

□

**命题 6.3.2**

$f \in D(L)$ ,  $\forall t \geq 0$ , 都有

$$Q_t f = f + \int_0^t Q_s(Lf) ds = f + \int_0^t L(Q_s f) ds$$

证明

$$\begin{aligned}\lim_{s \rightarrow 0} \frac{Q_{t+s}f - Q_t f}{s} &= \lim_{s \rightarrow 0} Q_t \left( \frac{Q_s f - f}{s} \right) \\ &= Q_t(Lf)\end{aligned}$$

积分可得。 □**命题 6.3.3 (生成元与预解式的关系)** $\lambda > 0,$ (1).  $\forall g \in C_0(E)$ , 有  $R_\lambda g \in D(L)$  且  $(\lambda - L)R_\lambda g = g$ .(2).  $f \in D(L)$ , 则  $R_\lambda(\lambda - L)f = f$ .结合上述结论, 可知  $R = D(L)$ ,  $R_\lambda = (\lambda - L)^{-1}$ .

证明

(1). 对于  $g \in C_0(E)$ , 考虑

$$\begin{aligned}Q_s(R_\lambda g) - R_\lambda g &= Q_s \left( \int_0^\infty e^{-\lambda t} Q_t g dt \right) - R_\lambda g \\ &= \int_0^\infty e^{-\lambda t} Q_{s+t} g dt - R_\lambda g \\ &= e^{\lambda s} \int_0^\infty e^{-\lambda(s+t)} Q_{s+t} g dt - R_\lambda g \\ &= e^{\lambda s} \int_s^\infty e^{-\lambda t} Q_t g dt - R_\lambda g \\ &= e^{\lambda s} \left( R_\lambda g - \int_0^s e^{-\lambda t} Q_t g dt \right) - R_\lambda g \\ &= (e^{\lambda s} - 1)R_\lambda g - e^{\lambda s} \int_0^s e^{-\lambda t} Q_t g dt\end{aligned}$$

从而可得

$$\begin{aligned}L(R_\lambda g) &= \lim_{s \rightarrow 0^+} \frac{1}{s} (Q_s(R_\lambda g) - R_\lambda g) \\ &= \lim_{s \rightarrow 0^+} \left( \frac{e^{\lambda s} - 1}{s} R_\lambda g - e^{\lambda s} \frac{1}{s} \int_0^s e^{-\lambda t} Q_t g dt \right) \\ &= \lambda R_\lambda g - g\end{aligned}$$

(2). 对于  $f \in D(L)$ , 考虑

$$\begin{aligned}R_\lambda(Lf) &= \int_0^\infty e^{-\lambda t} Q_t(Lf) dt \\ &= \int_0^\infty e^{-\lambda t} L(Q_t f) dt \\ &= \int_0^\infty e^{-\lambda t} \lim_{s \rightarrow 0^+} \frac{Q_{s+t} f - Q_t f}{s} dt \\ &= \lim_{s \rightarrow 0^+} \frac{1}{s} \int_0^\infty e^{-\lambda t} (Q_{s+t} f - Q_t f) dt \\ &= \lim_{s \rightarrow 0^+} \frac{1}{s} \left( \int_0^\infty e^{-\lambda t} Q_{s+t} f dt - R_\lambda f \right) \\ &\text{(这里类似于 (1) 的证明)} \\ &= \lim_{s \rightarrow 0^+} \frac{1}{s} \left( e^{\lambda s} \left( R_\lambda f - \int_0^s e^{-\lambda t} Q_t f dt \right) - R_\lambda f \right) \\ &= \lim_{s \rightarrow 0^+} \frac{1}{s} \left( (e^{\lambda s} - 1)R_\lambda f - e^{\lambda s} \int_0^s e^{-\lambda t} Q_t f dt \right)\end{aligned}$$

$$= \lambda R_\lambda f - f$$

□

**例 6.3.3.**

布朗运动的转移半群由下式给出:

$$Q_t f(x) = \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy = \mathbb{E}[f(B_t + x)]$$

其中  $f \in C_0(E)$ , 并设  $f$  光滑, 证明:

$$Lf(x) = \frac{1}{2} f''(x), \quad x \in \mathbb{R}$$

**证明** 我们先计算:

$$\begin{aligned} Q_t f(x) - f(x) &= \mathbb{E}[f(B_t^x) - f(x)] \\ &= \int_{\mathbb{R}} [f(y) - f(x)] \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy \\ &= \int_{\mathbb{R}} \left[ f'(x)(y-x) + f''(x) \frac{(y-x)^2}{2} + f'''(\theta_{x,y}) \frac{(y-x)^3}{6} \right] \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy \\ &= f'(x) \mathbb{E}[B_t^0] + \frac{1}{2} f''(x) \mathbb{E}[(B_t^0)^2] + \int_{\mathbb{R}} \left[ f'''(\theta_{x,y}) \frac{(y-x)^3}{6} \right] \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy \\ &= \frac{1}{2} f''(x) t + \int_{\mathbb{R}} \left[ f'''(\theta_{x,y}) \frac{(y-x)^3}{6} \right] \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy \end{aligned}$$

因此,

$$\begin{aligned} Lf(x) - \frac{1}{2} f''(x) &= \lim_{t \rightarrow 0^+} \frac{Q_t f(x) - f(x) - \frac{1}{2} f''(x) t}{t} \\ &= \lim_{t \rightarrow 0^+} \frac{\int_{\mathbb{R}} \left[ f'''(\theta_{x,y}) \frac{(y-x)^3}{6} \right] \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy}{t} \end{aligned}$$

**引理 6.3.2**

$f \in C_0(E)$  且  $f$  光滑, 则  $f' \in C_0(E)$ , 进而  $f$  的任意阶导数都满足此性质。

**证明** 只需证明  $f'$  在无穷远处趋于零, 假设不然:  $\exists \varepsilon_0 > 0$  使得  $\forall M > 0$ , 存在  $x_0$  满足  $|x_0| > M$  且  $f'(x_0) > \varepsilon_0$ , 根据  $f'$  的连续性, 存在  $a > 0$  使得

$$f'(x) > \varepsilon_0, \quad x \in (x_0, x_0 + a)$$

从而得到

$$f(x_0 + a) - f(x_0) > \varepsilon_0 a$$

同时, 我们取  $\varepsilon < \frac{1}{2} a \varepsilon_0$ , 由于  $f$  在无穷远处趋于零, 所以存在  $M > 0$ , 使得

$$\sup_{|x| > M} |f(x)| < \varepsilon$$

从而

$$\sup_{|x|, |y| > M} |f(x) - f(y)| < 2\varepsilon < \varepsilon_0 a$$

矛盾。 □

由引理可见,  $f''' \in C_0(E)$ , 从而是有界的, 这表明

$$\left| \int_{\mathbb{R}} \left[ f'''(\theta_{x,y}) \frac{(y-x)^3}{6} \right] \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{(y-x)^2}{2t}\right\} dy \right| \leq C \mathbb{E}[|B_t^0|^3] = C \cdot t^{\frac{3}{2}}$$

这说明

$$\lim_{t \rightarrow 0^+} \left| \int_{\mathbb{R}} [f'''(\theta_{x,y}) \frac{(y-x)^3}{6}] \frac{1}{\sqrt{2\pi t}} \exp \left\{ -\frac{(y-x)^2}{2t} \right\} dy \right| \leq \lim_{t \rightarrow 0^+} Ct^{\frac{1}{2}} = 0$$

结论得证。  $\square$

### 定理 6.3.3

设  $h, g \in C_0(E)$ , 以下命题等价:

- (1).  $h \in D(L)$ ,  $Lh = g$ .
- (2).  $\forall x \in E$ ,

$$M_t^x = h(X_t^x) - \int_0^t g(X_s^x) ds, \quad t \geq 0$$

是一个鞅。

**证明** (1)  $\Rightarrow$  (2): 适应性和可积性不再赘述, 我们验证:

$$\mathbb{E}[M_{s+t}^x | \mathcal{F}_s] = M_s^x$$

观察左式,

$$\begin{aligned} \mathbb{E}[M_{s+t}^x | \mathcal{F}_s] &= \mathbb{E} \left[ h(X_{s+t}^x) - \int_0^{s+t} g(X_u^x) du \middle| \mathcal{F}_s \right] \\ &= \mathbb{E}[h(X_{s+t}^x) | \mathcal{F}_s] - \int_0^s g(X_u^x) du - \mathbb{E} \left[ \int_s^{s+t} g(X_u^x) du \middle| \mathcal{F}_s \right] \\ &= Q_t h(X_s^x) - \int_0^s g(X_u^x) du - \mathbb{E} \left[ \int_0^t g(X_{u+s}^x) du \middle| \mathcal{F}_s \right] \\ &= Q_t h(X_s^x) - \int_0^s g(X_u^x) du - \int_0^t \mathbb{E} [g(X_{u+s}^x) | \mathcal{F}_s] du \\ &= Q_t h(X_s^x) - \int_0^s g(X_u^x) du - \int_0^t Q_u(g(X_s^x)) du \end{aligned}$$

由命题 6.3.2 可得

$$\int_0^t Q_u(g(X_s^x)) du = \int_0^t Q_u(Lh(X_s^x)) du = Q_t h(X_s^x) - h(X_s^x)$$

这说明

$$\begin{aligned} \mathbb{E}[M_{s+t}^x | \mathcal{F}_s] &= Q_t h(X_s^x) - \int_0^s g(X_u^x) du - Q_t h(X_s^x) + h(X_s^x) \\ &= h(X_s^x) - \int_0^s g(X_u^x) du = M_s^x \end{aligned}$$

(2)  $\Rightarrow$  (1):  $M_t^x$  是鞅, 则

$$\mathbb{E}[M_t^x] = \mathbb{E}[M_0^x] = h(x)$$

另一方面,

$$\begin{aligned} \mathbb{E}[M_t^x] &= \mathbb{E}[h(X_t^x)] - \mathbb{E} \left[ \int_0^t g(X_s^x) ds \right] \\ &= Q_t h(x) - \int_0^t \mathbb{E}[g(X_s^x)] ds \\ &= Q_t h(x) - \int_0^t Q_s g(x) ds \end{aligned}$$

所以

$$Lh(x) = \lim_{t \rightarrow 0^+} \frac{Q_t h - h}{t}(x) = \lim_{t \rightarrow 0^+} \frac{1}{t} \int_0^t Q_s g(x) ds = g(x)$$

$\square$

## 6.4 马氏性与强马氏性

回顾一下之前的内容：考虑过程  $(X_t^x)$ ，半群：

$$Q_t f(x) = \mathbb{E}[f(X_t^x)], t \geq 0$$

预解式：

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} Q_t f(x) dt, \lambda > 0$$

生成元：

$$L f = \lim_{t \rightarrow 0^+} \frac{Q_t f - f}{t}$$

### 定理 6.4.1

$(X_t)_{t \geq 0}$  是马氏过程， $(Q_t)_{t \geq 0}$  是相应的 Feller 半群，那么一定存在  $X$  的一个修正  $X'$ ，使得  $(Q_t)_{t \geq 0}$  也是  $X'$  的 Feller 半群，且具有 RCLL 轨道。

我们省略详细证明<sup>1</sup>，以下均不妨假设马氏过程有 RCLL 轨道。记

$$D(E) = \{f(t) : [0, +\infty) \rightarrow E \text{ 是 RCLL 函数}\}$$

固定  $\omega \in \Omega$ ， $t \mapsto \omega(t) \stackrel{\text{def}}{=} X_t(\omega)$  是 RCLL 函数 a.s.，所以不妨认为  $D(E)$  就是  $\Omega$ ，其上的  $\sigma$ -域就是：

$$\mathcal{B}(D(E)) = \sigma(X_t(\omega) = \omega(t) \in D(E), 0 \leq t < +\infty)$$

### 定义 6.4.1 (时间平移算子)

对于随机过程  $(Y_t)_{t \geq 0}$ ，定义算子：

$$Y_t \circ \theta_s = Y_{t+s}, s \geq 0$$

### 定理 6.4.2 (马氏性的一般表述)

$(Q_t)_{t \geq 0}$  为转移半群， $X = (X_t)_{t \geq 0}$  是相应的马氏过程，且具有 RCLL 轨道。设  $\Phi : D(E) \rightarrow \mathbb{R}$  有界可测，那么对于  $\forall s \geq 0$ ，

$$\mathbb{E}[\Phi(X \circ \theta_s) | \mathcal{F}_s] = \mathbb{E}_{X_s}[\Phi(X)]$$

**注** 其中， $X$  视为一个把  $\omega \in \Omega$  映为样本轨道的映射，那么  $\Phi(X)$  就是一个随机变量，相当于：

$$\omega \mapsto \Phi(u \mapsto X_u(\omega))$$

不难看出，如果令

$$\Phi(X(\omega)) = \Phi(u \mapsto X_u(\omega)) = f(X_t(\omega))$$

其中  $t \geq 0$  是一个固定的常数， $f$  是一个固定的有界可测实函数，带入到本定理的结论中就得到

$$\mathbb{E}[f(X_{t+s}) | \mathcal{F}_s] = \mathbb{E}_{X_s}[f(X_t)] = Q_t f(X_s)$$

这正是我们在定义 6.1.4 中的表述，这说明本定理给出的表述是更一般化的马氏性表述。

**证明** 根据单调类定理，我们只需证明对以下形式的  $\Phi$  成立即可：

$$\Phi(X(\omega)) = \varphi_1(X_{t_1}) \varphi_2(X_{t_2}) \cdots \varphi_p(X_{t_p})$$

其中  $t_1 < t_2 < \cdots < t_p$ ， $\forall \varphi_i \in B(E)$ ，也就是证明：

$$\mathbb{E}[\varphi_1(X_{t_1+s}) \varphi_2(X_{t_2+s}) \cdots \varphi_p(X_{t_p+s}) | \mathcal{F}_s] = \mathbb{E}_{X_s}[\varphi_1(X_{t_1}) \varphi_2(X_{t_2}) \cdots \varphi_p(X_{t_p})]$$

<sup>1</sup>大致思路为：考虑上鞅  $\{e^{-\lambda t} g(X_t)\}$ ，对其进行轨道修正。

归纳： $p = 1$  时由定义即得，设  $p - 1$  时成立，考虑  $p$  的情形：

$$\begin{aligned}
 \text{LHS} &= \mathbb{E}[\varphi_1(X_{t_1+s})\varphi_2(X_{t_2+s})\cdots\varphi_p(X_{t_p+s})|\mathcal{F}_s] \\
 &= \mathbb{E}[\mathbb{E}[\varphi_1(X_{t_1+s})\varphi_2(X_{t_2+s})\cdots\varphi_p(X_{t_p+s})|\mathcal{F}_{t_{p-1}+s}]|\mathcal{F}_s] \\
 &= \mathbb{E}[\varphi_1(X_{t_1+s})\varphi_2(X_{t_2+s})\cdots\varphi_{p-1}(X_{t_{p-1}+s})\mathbb{E}[\varphi_p(X_{t_p+s})|\mathcal{F}_{t_{p-1}+s}]|\mathcal{F}_s] \\
 &= \mathbb{E}[\varphi_1(X_{t_1+s})\varphi_2(X_{t_2+s})\cdots\varphi_{p-1}(X_{t_{p-1}+s})Q_{t_p-t_{p-1}}\varphi_p(X_{t_{p-1}+s})|\mathcal{F}_s] \\
 &= \mathbb{E}_{X_s}[\varphi_1(X_{t_1})\varphi_2(X_{t_2})\cdots\varphi_{p-1}(X_{t_{p-1}})Q_{t_p-t_{p-1}}\varphi_p(X_{t_{p-1}})] \\
 \text{RHS} &= \mathbb{E}_{X_s}[\varphi_1(X_{t_1})\varphi_2(X_{t_2})\cdots\varphi_p(X_{t_p})] \\
 &= \mathbb{E}_{X_s}[\mathbb{E}[\varphi_1(X_{t_1})\varphi_2(X_{t_2})\cdots\varphi_p(X_{t_p})|\mathcal{F}_{t_{p-1}}]] \\
 &= \mathbb{E}_{X_s}[\varphi_1(X_{t_1})\varphi_2(X_{t_2})\cdots\varphi_{p-1}(X_{t_{p-1}})Q_{t_p-t_{p-1}}\varphi_p(X_{t_{p-1}})]
 \end{aligned}$$

二者相等。综上结论得证。  $\square$

### 定理 6.4.3 (强马氏性)

$(Q_t)_{t \geq 0}$  为 Feller 半群， $X = (X_t)_{t \geq 0}$  是相应的马氏过程，且具有 RCLL 轨道。设  $\Phi : D(E) \rightarrow \mathbb{R}$  有界可测， $T$  是停时，则

$$\mathbb{E}[\Phi(X \circ \theta_T) \cdot I_{\{T < +\infty\}} | \mathcal{F}_T] = I_{\{T < +\infty\}} \cdot \mathbb{E}_{X_T}[\Phi(X)]$$

**证明** 根据单调类定理，我们只需证明对以下形式的  $\Phi$  成立即可：

$$\Phi(X(\omega)) = \varphi_1(X_{t_1})\varphi_2(X_{t_2})\cdots\varphi_p(X_{t_p})$$

其中  $t_1 < t_2 < \cdots < t_p$ ,  $\forall \varphi_i \in C_0(E)$ ，也就是证明：

$$\mathbb{E}[\varphi_1(X_{t_1+T})\varphi_2(X_{t_2+T})\cdots\varphi_p(X_{t_p+T})I_{\{T < +\infty\}} | \mathcal{F}_T] = I_{\{T < +\infty\}} \mathbb{E}_{X_T}[\varphi_1(X_{t_1})\varphi_2(X_{t_2})\cdots\varphi_p(X_{t_p})]$$

由于  $I_{\{T < +\infty\}} \mathbb{E}_{X_T}[\Phi(X)]$  关于  $\mathcal{F}_T$  可测，由条件期望的定义，只需证明：对  $A \in \mathcal{F}_T$ ，有

$$E \left[ I_{A \cap \{T < +\infty\}} \Phi(X \circ \theta_T) \right] = E \left[ I_{A \cap \{T < +\infty\}} \mathbb{E}_{X_T}[\Phi(X)] \right]$$

归纳：先证明其对  $p = 1$  的情况成立，此时  $\Phi(X(\omega)) = \phi_1(X_{t_1}(\omega))$ ，

$$\text{LHS} = E \left[ I_{A \cap \{T < +\infty\}} \Phi(X \circ \theta_T) \right] = E \left[ I_{A \cap \{T < +\infty\}} \phi_1(X_{t_1+T}) \right]$$

定义

$$[T]_n = \frac{k}{2^n}, \text{ 如果 } \frac{k-1}{2^n} \leq T < \frac{k}{2^n}, k = 0, 1, \dots$$

那么有  $[T]_n \searrow T$ 。进而，

$$\begin{aligned}
 \text{LHS} &= \lim_{n \rightarrow \infty} E \left[ I_{A \cap \{T < +\infty\}} \phi_1(X_{t_1+[T]_n}) \right] \\
 &= \lim_{n \rightarrow \infty} \sum_{k=1}^n E \left[ I_{A \cap \{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\}} \phi_1(X_{t_1+\frac{k}{2^n}}) \right] \\
 &= \lim_{n \rightarrow \infty} \sum_{k=1}^n E \left[ I_{A \cap \{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\}} E \left[ \phi_1(X_{t_1+\frac{k}{2^n}}) \middle| \mathcal{F}_{\frac{k}{2^n}} \right] \right] \quad (A \in \mathcal{F}_T) \\
 &= \lim_{n \rightarrow \infty} \sum_{k=1}^n E \left[ I_{A \cap \{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\}} Q_{t_1} \phi_1(X_{\frac{k}{2^n}}) \right] \quad (\text{马氏性}) \\
 &= \lim_{n \rightarrow \infty} E \left[ I_{A \cap \{T < +\infty\}} Q_{t_1} \phi_1(X_{[T]_n}) \right] \\
 &= E \left[ I_{A \cap \{T < +\infty\}} Q_{t_1} \phi_1(X_T) \right] \quad (\text{Feller 性质})
 \end{aligned}$$

另一方面，

$$\text{RHS} = E \left[ I_{A \cap \{T < +\infty\}} \mathbb{E}_{X_T}[\phi_1(X_{t_1})] \right] = E \left[ I_{A \cap \{T < +\infty\}} Q_{t_1} \phi_1(X_{t_1}) \right]$$

二者相等，于是得证。 $p - 1$  到  $p$  的递推的证明和之前类似，不再赘述。  $\square$

## 6.5 两个例子：跳跃过程与 Levy 过程

### 6.5.1 跳跃过程

如果状态空间  $E$  是至多可数集，我们采用离散度量诱导的拓扑：

$$d(x, y) = \delta_{x, y}, \quad \forall x, y \in E$$

$E$  的子集全体作为其  $\sigma$ -域。容易验证任意  $\varphi: E \rightarrow \mathbb{R}$  都是连续的，所以  $C_0(E) = B(E)$ 。

用  $D(E)$  代表  $E$  上的 RCLL 函数全体，注意到  $\forall f \in D(E)$ ，右连续  $\Rightarrow f(0+) = f(0)$ ，故存在  $t_1 > 0$  使得

$$f(t) = f(0), \quad \forall t \in [0, t_1]$$

那么，我们就令

$$\gamma_1^f \stackrel{\text{def}}{=} \sup\{t_1 \geq 0 : f(t) = f(0), \forall t \in [0, t_1]\}$$

即  $f$  第一次改变初始状态的时刻（可以取  $+\infty$ ），如果  $\gamma_1^f < +\infty$ ，以此类推：

$$\gamma_2^f \stackrel{\text{def}}{=} \sup\{t_2 \geq \gamma_1^f : f(t_2) = f(\gamma_1^f), \forall t \in [\gamma_1^f, t_2]\}$$

$$\gamma_3^f \stackrel{\text{def}}{=} \sup\{t_3 \geq \gamma_2^f : f(t_3) = f(\gamma_2^f), \forall t \in [\gamma_2^f, t_3]\}$$

$\vdots$

一个简单的推论是： $\gamma_n^f \nearrow +\infty$ ，否则，假设  $\gamma_n^f \nearrow \gamma^f < +\infty$ ，则

$$\lim_{n \rightarrow \infty} f(\gamma_n^f) = f(\gamma^f -)$$

从而存在  $N$  使得  $n > N$  时  $f(\gamma_n^f) = f(\gamma^f -)$ ，矛盾。

现在我们考虑  $E$  上的 Feller 半群  $(Q_t)_{t \geq 0}$ ， $X$  是相应的（轨道 RCLL 的）马氏过程，我们记概率测度  $\mathbb{P}_x$  是满足  $\mathbb{P}_x(X_0 = x) = 1$  的测度， $\mathbb{E}_x$  是  $\mathbb{P}_x$  下的期望。那么，固定  $\omega \in \Omega$ ，得到 RCLL 的样本轨道  $t \mapsto X_t(\omega)$ ，根据前文中的分析，存在一系列

$$0 = T_0(\omega) \leq T_1(\omega) \leq T_2(\omega) \leq \cdots \leq +\infty$$

满足：如果  $T_i(\omega) < +\infty$ ，那么

$$X_t(\omega) = X_{T_i}(\omega), \quad \forall t \in [T_i(\omega), T_{i+1}(\omega))$$

不难验证  $T_i$  都是停时。

#### 定理 6.5.1 (指数时间)

对于固定的  $x \in E$ ，存在常数  $q(x) \geq 0$ ，使得  $\mathbb{P}_x$  测度下， $T_1$  服从参数为  $q(x)$  的指数分布<sup>a</sup>。如果  $q(x) > 0$ ，则  $T_1$  与  $X_{T_1}$  在  $\mathbb{P}_x$  测度下独立。

<sup>a</sup>参数为 0 的指数分布定义为  $= +\infty$  a.s.

**证明** 任取  $s, t \geq 0$ ,

$$\begin{aligned} \mathbb{P}_x(T_1 > s + t) &= \mathbb{E}_x[I_{\{T_1 > s+t\}}] \\ &= \mathbb{E}_x[I_{\{\omega: X_u(\omega) = X_0(\omega), \forall u \in [0, s+t]\}}] \\ &= \mathbb{E}_x[I_{\{\omega: X_u(\omega) = X_0(\omega), \forall u \in [0, s]\}} I_{\{\omega: X_u(\omega) = X_s(\omega), \forall u \in [s, s+t]\}}] \\ &\quad (\text{我们定义: } \Phi(f) = I_{\{f(u) = f(0), \forall u \in [0, t]\}}) \\ &= \mathbb{E}_x[I_{\{T_1 > s\}} \cdot \Phi(X \circ \theta_s)] \\ &= \mathbb{E}_x[\mathbb{E}[I_{\{T_1 > s\}} \cdot \Phi(X \circ \theta_s) | \mathcal{F}_s]] \\ &= \mathbb{E}_x[I_{\{T_1 > s\}} \cdot \mathbb{E}[\Phi(X \circ \theta_s) | \mathcal{F}_s]] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_x[I_{\{T_1 > s\}} \cdot \mathbb{E}_{X_s}[\Phi(X)]] \\
&= \mathbb{E}_x[I_{\{T_1 > s\}} \cdot \mathbb{E}_{X_s}[I_{\{T_1 > t\}}]] \\
&\quad (\text{注意 } T_1 > s \text{ 时 } X_s = x) \\
&= \mathbb{E}_x[I_{\{T_1 > s\}} \cdot \mathbb{E}_x[I_{\{T_1 > t\}}]] \\
&= \mathbb{P}_x(T_1 > s)\mathbb{P}_x(T_1 > t)
\end{aligned}$$

此即无记忆性，可得  $T_1$  服从  $\mathbb{P}_x$  下的指数分布。

如果  $q(x) > 0$ ，即  $T_1 < +\infty$   $\mathbb{P}_x$ -a.s.，于是  $\forall t \geq 0, y \in E$  有

$$\begin{aligned}
\mathbb{P}_x(T_1 > t, X_{T_1} = y) &= \mathbb{E}_x[I_{\{T_1 > t\}} I_{\{X_{T_1} = y\}}] \\
&\quad (\text{我们定义: } \Psi(f) = I_{\{\gamma_1^f < +\infty, f(\gamma_1^f) = y\}}) \\
&= \mathbb{E}_x[I_{\{T_1 > t\}} \Psi(X \circ \theta_t)] \\
&= \mathbb{E}_x[\mathbb{E}[I_{\{T_1 > t\}} \Psi(X \circ \theta_t) | \mathcal{F}_t]] \\
&= \mathbb{E}_x[I_{\{T_1 > t\}} \cdot \mathbb{E}[\Psi(X \circ \theta_t) | \mathcal{F}_t]] \\
&= \mathbb{E}_x[I_{\{T_1 > t\}} \cdot \mathbb{E}_{X_t}[\Psi(X)]] \\
&= \mathbb{E}_x[I_{\{T_1 > t\}} \cdot \mathbb{E}_x[I_{\{X_{T_1} = y\}}]] \\
&= \mathbb{P}_x(T_1 > t)\mathbb{P}_x(X_{T_1} = y)
\end{aligned}$$

独立性得证。 □

接下来，如果  $q(x) > 0$ ，记  $\pi(x, y) = \mathbb{P}_x(X_{T_1} = y)$ ，即从  $x$  出发第一次跳跃到  $y$ ，不难验证：

$$\sum_{y \in E} \pi(x, y) = 1, \pi(x, x) = 0$$

#### 定理 6.5.2 (生成元的表述)

$(Q_t)_{t \geq 0}$  的生成元为  $L$ ，定义域  $D(L) = C_0(E) = B(E)$ ，那么对于  $\varphi \in B(E)$ ，

$$L\varphi(x) = q(x) \sum_{y \neq x} \pi(x, y)(\varphi(y) - \varphi(x))$$

**证明** 如果  $q(x) = 0$ ，则

$$Q_t\varphi(x) = \mathbb{E}_x[\varphi(X_t)] = \mathbb{E}_x[\varphi(x)] = \varphi(x)$$

从而  $L\varphi(x) = 0$ 。

如果  $q(x) > 0$ ，

#### 引理 6.5.1

$\mathbb{P}_x(T_2 \leq t) = O(t^2)$ ，即存在  $C > 0$  使得  $\mathbb{P}_x(T_2 \leq t) \leq Ct^2$ 。

**证明**

$$\begin{aligned}
\mathbb{P}_x(T_2 \leq t) &\leq \mathbb{P}_x(T_1 \leq t, T_2 \leq T_1 + t) \\
&= \mathbb{P}_x(T_1 \leq t, T_1 + T_1 \circ \theta_{T_1} \leq T_1 + t) \\
&= \mathbb{P}_x(T_1 \leq t, T_1 \circ \theta_{T_1} \leq t) \\
&= \mathbb{E}_x[\mathbb{E}_x[I_{\{T_1 \leq t\}} I_{\{T_1 \circ \theta_{T_1} \leq t\}} | \mathcal{F}_{T_1}]] \\
&= \mathbb{E}_x[I_{\{T_1 \leq t\}} \mathbb{E}_{X_{T_1}}[I_{\{T_1 \leq t\}}]] \quad (\text{强马氏性}) \\
&\leq \mathbb{E}_x[I_{\{T_1 \leq t\}} \sup_{y \in E} \mathbb{P}_y(T_1 \leq t)] \\
&\leq \sup_{y \in E} \mathbb{P}_y(T_1 \leq t) \mathbb{P}_x(T_1 \leq t)
\end{aligned}$$

$$\begin{aligned}
&= \sup_{y \in E} (1 - e^{-q(y)t})(1 - e^{-q(x)t}) \\
&\leq Ct^2, \quad \text{当 } t \rightarrow 0
\end{aligned}$$

□ ♥

那么

$$\begin{aligned}
Q_t \varphi(x) &= \mathbb{E}_x[\varphi(X_t)] \\
&= \mathbb{E}_x[\varphi(X_t); t < T_1] + \mathbb{E}_x[\varphi(X_t); t \geq T_1] \\
&= \mathbb{E}_x[\varphi(X_t); t < T_1] + \mathbb{E}_x[\varphi(X_t); T_1 \leq t < T_2] + \mathbb{E}_x[\varphi(X_t); t > T_2] \\
&= \mathbb{E}_x[\varphi(x); t < T_1] + \mathbb{E}_x[\varphi(X_{T_1}); T_1 \leq t < T_2] + \mathbb{E}_x[\varphi(X_t); t > T_2] \\
&= \varphi(x)\mathbb{P}_x(t < T_1) + \mathbb{E}_x[\varphi(X_{T_1}); t \geq T_1] + O(t^2) \\
&= \varphi(x)e^{-q(x)t} + \mathbb{E}_x[\varphi(X_{T_1})]\mathbb{P}_x(t \geq T_1) + O(t^2) \\
&= \varphi(x)e^{-q(x)t} + \sum_{y \neq x} \pi(x, y)\varphi(y)(1 - e^{-q(x)t}) + O(t^2) \\
&= \varphi(x)e^{-q(x)t} + (1 - e^{-q(x)t}) \sum_{y \neq x} \pi(x, y)\varphi(y) + O(t^2)
\end{aligned}$$

因此

$$\frac{Q_t \varphi(x) - \varphi(x)}{t} = \frac{\varphi(x)(e^{-q(x)t} - 1) + (1 - e^{-q(x)t}) \sum_{y \neq x} \pi(x, y)\varphi(y) + O(t^2)}{t} \rightarrow q(x) \sum_{y \neq x} \pi(x, y)(\varphi(y) - \varphi(x))$$

□

**定理 6.5.3 (跳链)**

如果  $\forall q(y) > 0$ , 那么固定  $x \in E$ ,  $T_1(\omega) < T_2(\omega) < \dots$  都是  $\mathbb{P}_x$ -a.s. 有限的, 此时,  $\{X_{T_n}, n \in \mathbb{N}_+\}$  构成了一个 (离散时间、离散状态) 马氏链, 其一步转移概率  $p(x, y)$  就是  $\pi(x, y)$ .

证明

$$\begin{aligned}
\mathbb{P}_x(X_{T_1} = z_1, X_{T_2} = z_2) &= \mathbb{P}_x(X_{T_1} = z_1, X_{T_1+T_1 \circ \theta_{T_1}} = z_2) \\
&= \mathbb{P}_x(X_{T_1} = z_1, X_{T_1} \circ \theta_{T_1} = z_2) \\
&= \mathbb{P}_x(X_{T_1} = z_1, E_{X_{T_1}}[X_{T_1} = z_2]) && \text{(强马氏性)} \\
&= \mathbb{P}_x(X_{T_1} = z_1, E_{z_1}[X_{T_1} = z_2]) \\
&= \mathbb{P}_x(X_{T_1} = z_1)P_{z_1}(X_{T_1} = z_2) \\
&= \pi(x, z_1)\pi(z_1, z_2)
\end{aligned}$$

归纳可证

$$P_x(X_{T_1} = z_1, X_{T_2} = z_2, \dots, X_{T_n} = z_n) = \pi(x, z_1)\pi(z_1, z_2) \cdots \pi(z_{n-1}, z_n)$$

□

**6.5.2 Levy 过程****定义 6.5.1**

$Y = (Y_t)_{t \geq 0}$  是一个在  $\mathbb{R}$  上取值的随机过程, 如果:

- (1).  $Y_0 = 0$  a.s.
- (2).  $Y$  有独立平稳增量: 对于  $\forall s \leq t$ ,  $Y_t - Y_s$  与  $(Y_r, r \leq s)$  独立, 并且与  $Y_{t-s}$  同分布.

(3). 当  $t \rightarrow 0$  时,  $Y_t$  依概率收敛到 0.

### 定理 6.5.4

定义  $B(\mathbb{R})$  上的算子如下: 固定  $t \geq 0$ ,

$$Q_t f(x) = \mathbb{E}[f(Y_t + x)]$$

那么,  $(Q_t)_{t \geq 0}$  是  $C_0(\mathbb{R})$  上的一个 Feller 半群, 并且  $Y$  是关于  $(Q_t)_{t \geq 0}$  的马氏过程。

## 6.6 习题

### 题目 6.1. le gall(Exercise 6.26)

(Feynman-Kac formula) 设  $v \in C_0(E)$  非负, 对于  $\varphi \in B(E)$ , 我们定义

$$Q_t^* \varphi(x) = \mathbb{E}_x \left[ \varphi(X_t) \exp \left\{ - \int_0^t v(X_s) ds \right\} \right], \quad \forall x \in E, \forall t \geq 0$$

1. 证明: 对于  $\forall \varphi \in B(E), s, t \geq 0$  有  $Q_{t+s}^* \varphi = Q_t^*(Q_s^* \varphi)$ .
2. 根据下面这个恒等式:

$$1 - \exp \left\{ - \int_0^t v(X_s) ds \right\} = \int_0^t v(X_s) \exp \left\{ - \int_s^t v(X_r) dr \right\} ds$$

证明:  $\forall \varphi \in B(E)$ , 都有

$$Q_t \varphi - Q_t^* \varphi = \int_0^t Q_s(v Q_{t-s}^* \varphi) ds$$

3. 设  $\varphi \in D(L)$ , 证明:

$$\frac{d}{dt} Q_t^* \varphi|_{t=0} = L\varphi - v\varphi$$

解答: 本题没什么难度, 用的都是很简单的马氏性变换技巧。

1.

$$\begin{aligned} Q_{t+s}^* \varphi(x) &= \mathbb{E}_x \left[ \varphi(X_{t+s}) \exp \left\{ - \int_0^{t+s} v(X_u) du \right\} \right] \\ &= \mathbb{E}_x \left[ \mathbb{E} \left[ \varphi(X_{t+s}) \exp \left\{ - \int_0^{t+s} v(X_u) du \right\} \middle| \mathcal{F}_t \right] \right] \\ &= \mathbb{E}_x \left[ \exp \left\{ - \int_0^t v(X_u) du \right\} \cdot \mathbb{E} \left[ \varphi(X_{t+s}) \exp \left\{ - \int_t^{t+s} v(X_u) du \right\} \middle| \mathcal{F}_t \right] \right] \\ &= \mathbb{E}_x \left[ \exp \left\{ - \int_0^t v(X_u) du \right\} \cdot \mathbb{E}_{X_t} \left[ \varphi(X_s) \exp \left\{ - \int_0^s v(X_u) du \right\} \right] \right] \quad (\text{由马氏性}) \\ &= \mathbb{E}_x \left[ \exp \left\{ - \int_0^t v(X_u) du \right\} \cdot Q_s^*(X_t) \right] \\ &= Q_t^*(Q_s^*)(x) \end{aligned}$$

2.

$$\begin{aligned} Q_t \varphi(x) - Q_t^* \varphi(x) &= \mathbb{E}_x \left[ \varphi(X_t) \cdot \left( 1 - \exp \left\{ - \int_0^t v(X_s) ds \right\} \right) \right] \\ &= \mathbb{E}_x \left[ \varphi(X_t) \cdot \int_0^t v(X_s) \exp \left\{ - \int_s^t v(X_r) dr \right\} ds \right] \end{aligned}$$

$$\begin{aligned}
&= \int_0^t \mathbb{E}_x \left[ \varphi(X_t) \cdot v(X_s) \exp \left\{ - \int_s^t v(X_r) dr \right\} \right] ds \\
&= \int_0^t \mathbb{E}_x \left[ \mathbb{E} \left[ \varphi(X_t) \cdot v(X_s) \exp \left\{ - \int_s^t v(X_r) dr \right\} \middle| \mathcal{F}_s \right] \right] ds \\
&= \int_0^t \mathbb{E}_x \left[ v(X_s) \mathbb{E} \left[ \varphi(X_t) \cdot \exp \left\{ - \int_s^t v(X_r) dr \right\} \middle| \mathcal{F}_s \right] \right] ds \\
&= \int_0^t \mathbb{E}_x \left[ v(X_s) \mathbb{E}_{X_s} \left[ \varphi(X_{t-s}) \cdot \exp \left\{ - \int_0^{t-s} v(X_r) dr \right\} \right] \right] ds \\
&= \int_0^t \mathbb{E}_x [v(X_s) Q_{t-s}^* \varphi(X_s)] ds \\
&= Q_s(v(Q_{t-s}^* \varphi))(x)
\end{aligned}$$

3. 易证  $Q_0^* \varphi = \varphi$ , 所以

$$\frac{d}{dt} Q_t^* \varphi \Big|_{t=0} = \lim_{t \rightarrow 0^+} \frac{Q_t^* \varphi - \varphi}{t} = \lim_{t \rightarrow 0^+} \frac{Q_t^* \varphi - Q_t \varphi + Q_t \varphi - \varphi}{t} = L\varphi - \lim_{t \rightarrow 0^+} \frac{\int_0^t Q_s(vQ_{t-s}^* \varphi) ds}{t} = L\varphi - v\varphi$$

### 题目 6.2. 1e gall(Exercise 6.29.1-4)

1.  $g \in C_0(E)$ ,  $x \in E$ ,  $T$  为停时, 证明:

$$\mathbb{E}_x \left[ I_{\{T < +\infty\}} e^{-\lambda T} \int_{\mathbb{R}} e^{-\lambda t} g(X_{T+t}) dt \right] = \mathbb{E}_x [I_{\{T < +\infty\}} e^{-\lambda T} R_\lambda g(X_T)]$$

2. 证明:

$$R_\lambda g(x) = \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} g(X_t) dt \right] + \mathbb{E}_x [I_{\{T < +\infty\}} e^{-\lambda T} R_\lambda g(X_T)]$$

3. 如果  $f \in D(L)$ , 证明:

$$f(x) = \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} (\lambda f - Lf)(X_t) dt \right] + \mathbb{E}_x [I_{\{T < +\infty\}} e^{-\lambda T} f(X_T)]$$

4. 设  $\mathbb{E}_x[T] < +\infty$ , 利用 3. 中结论证明:

$$\mathbb{E}_x \left[ \int_0^T Lf(X_t) dt \right] = \mathbb{E}_x [f(X_T)] - f(x)$$

解答:

1.

$$\begin{aligned}
\text{LHS} &= \mathbb{E}_x \left[ I_{\{T < +\infty\}} e^{-\lambda T} \int_{\mathbb{R}} e^{-\lambda t} g(X_{T+t}) dt \right] \\
&= \mathbb{E}_x \left[ \mathbb{E} \left[ I_{\{T < +\infty\}} e^{-\lambda T} \int_{\mathbb{R}} e^{-\lambda t} g(X_{T+t}) dt \middle| \mathcal{F}_T \right] \right] \\
&= \mathbb{E}_x \left[ e^{-\lambda T} \mathbb{E} \left[ I_{\{T < +\infty\}} \int_{\mathbb{R}} e^{-\lambda t} g(X_{T+t}) dt \middle| \mathcal{F}_T \right] \right] \\
&= \mathbb{E}_x \left[ e^{-\lambda T} I_{\{T < +\infty\}} \mathbb{E}_{X_T} \left[ \int_{\mathbb{R}} e^{-\lambda t} g(X_t) dt \right] \right] \\
&= \mathbb{E}_x \left[ e^{-\lambda T} I_{\{T < +\infty\}} \int_{\mathbb{R}} \mathbb{E}_{X_T} [e^{-\lambda t} g(X_t)] dt \right] \\
&= \mathbb{E}_x \left[ e^{-\lambda T} I_{\{T < +\infty\}} \int_{\mathbb{R}} e^{-\lambda t} Q_t g(X_T) dt \right]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}_x [e^{-\lambda T} I_{\{T < +\infty\}} R_\lambda g(X_T)] \\
&= \text{RHS}
\end{aligned}$$

2.

$$\begin{aligned}
R_\lambda g(x) &= \int_0^\infty e^{-\lambda t} Q_t g(x) dt \\
&= \int_0^\infty e^{-\lambda t} \mathbb{E}_x [g(X_t)] dt \\
&= \mathbb{E}_x \left[ \int_0^\infty e^{-\lambda t} g(X_t) dt \right] \\
&= \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} g(X_t) dt \right] + \mathbb{E}_x \left[ \int_T^\infty e^{-\lambda t} g(X_t) dt \right] \\
&= \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} g(X_t) dt \right] + \mathbb{E}_x \left[ I_{\{T < +\infty\}} \int_T^\infty e^{-\lambda t} g(X_t) dt \right] \\
&= \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} g(X_t) dt \right] + \mathbb{E}_x \left[ e^{-\lambda T} I_{\{T < +\infty\}} \int_0^\infty e^{-\lambda t} g(X_{t+T}) dt \right] \\
&= \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} g(X_t) dt \right] + \mathbb{E}_x [I_{\{T < +\infty\}} e^{-\lambda T} R_\lambda g(X_T)] \quad (\text{由 1. 中结论})
\end{aligned}$$

3. 注意到  $R_\lambda = (\lambda - L)^{-1}$ , 考虑  $g = (\lambda - L)f$ , 代入 2. 中结论即可。4.  $\mathbb{E}_x[T] < +\infty$  说明  $T < +\infty$   $\mathbb{P}_x$ -a.s., 以下不再特殊说明。由于  $f, L(f)$  有界,  $\mathbb{E}_x[T] < +\infty$ , 令 3. 中的  $\lambda \rightarrow 0$ , 根据控制收敛定理,

$$\begin{aligned}
f(x) &= \lim_{\lambda \rightarrow 0} \mathbb{E}_x \left[ \int_0^T e^{-\lambda t} (\lambda f - Lf)(X_t) dt \right] + \lim_{\lambda \rightarrow 0} \mathbb{E}_x [I_{\{T < +\infty\}} e^{-\lambda T} f(X_T)] \\
&= \mathbb{E}_x \left[ \int_0^T \lim_{\lambda \rightarrow 0} e^{-\lambda t} (\lambda f - Lf)(X_t) dt \right] + \mathbb{E}_x [\lim_{\lambda \rightarrow 0} e^{-\lambda T} f(X_T)] \\
&= -\mathbb{E}_x \left[ \int_0^T Lf(X_t) dt \right] + \mathbb{E}_x [f(X_T)]
\end{aligned}$$

关于本题 4. 的结论, 有一种更直接的证法: 回顾定理 6.3.3, 可知

$$M_t = f(X_t) - \int_0^t Lf(X_s) ds$$

是一个鞅, 固定  $K > 0$ ,  $M_{t \wedge K}$  为一致可积鞅, 由择停定理可知

$$f(x) = \mathbb{E}_x [M_0] = \mathbb{E}_x [f(X_{T \wedge K}) - \int_0^{T \wedge K} Lf(X_s) ds]$$

由  $\mathbb{E}_x[T] < +\infty$  可知

$$\lim_{K \rightarrow \infty} f(X_{T \wedge K}) = f(X_T) \quad \mathbb{P}_x\text{-a.s.}$$

由控制收敛定理即得

$$f(x) = \mathbb{E}_x [f(X_T)] - \mathbb{E}_x \left[ \int_0^T Lf(X_t) dt \right]$$

最后放一道去年期末题, 比较简单, 甚至都不需要用到马氏性。

**题目 6.3. 2023SPFinal.6** $(X_t^x)$  为马氏过程,  $(Q_t)_{t \geq 0}$  为 Feller 半群, 生成元为  $L$ ,  $X_0^x = x$ ,  $v, h \in C_0(E)$ , 求证: 如果

$$M_t^x = e^{-\int_0^t v(X_s^x) ds} h(X_t^x)$$

是鞅, 则  $h \in D(L)$ , 并且  $Lh = vh$ .

解答: 任取  $t, u \geq 0$ ,  $\mathbb{E}[M_{t+u}^x | \mathcal{F}_t] = M_t^x$ , 左边等于

$$\mathbb{E} \left[ e^{-\int_0^{t+u} v(X_s^x) ds} h(X_{t+u}^x) \middle| \mathcal{F}_t \right] = e^{-\int_0^t v(X_s^x) ds} \mathbb{E} \left[ e^{-\int_t^{t+u} v(X_s^x) ds} h(X_{t+u}^x) \middle| \mathcal{F}_t \right]$$

因此

$$\mathbb{E} \left[ e^{-\int_t^{t+u} v(X_s^x) ds} h(X_{t+u}^x) \middle| \mathcal{F}_t \right] = h(M_t^x)$$

令  $t = 0$ , 得到

$$h(x) = \mathbb{E}_x \left[ e^{-\int_0^u v(X_s) ds} h(X_u) \right]$$

所以

$$\begin{aligned} Lh(x) &= \lim_{u \rightarrow 0^+} \frac{Q_u h(x) - h(x)}{u} \\ &= \lim_{u \rightarrow 0^+} \frac{\mathbb{E}_x [h(X_u)] - h(x)}{u} \\ &= \lim_{u \rightarrow 0^+} \frac{\mathbb{E}_x \left[ \left( 1 - e^{-\int_0^u v(X_s) ds} \right) h(X_u) \right]}{u} \\ &= \mathbb{E}_x \left[ \lim_{u \rightarrow 0^+} \frac{1 - e^{-\int_0^u v(X_s) ds}}{u} h(X_u) \right] \\ &= \mathbb{E}_x \left[ h(x) \lim_{u \rightarrow 0^+} \frac{\int_0^u v(X_s) \exp \left\{ -\int_s^u v(X_r) dr \right\} ds}{u} \right] \\ &= \mathbb{E}_x \left[ h(x) \lim_{u \rightarrow 0^+} v(X_0) \exp \left\{ -\int_0^u v(X_r) dr \right\} \right] \quad (\text{洛必达}) \\ &= \mathbb{E}_x [h(x)v(x)] \\ &= h(x)v(x) \end{aligned}$$

中间用到了题目 6.1.2 里面的那个恒等式。

## 6.7 关于跳跃过程的更多内容\*

此部分来源于应随课程，主要内容为离散状态、连续时间马氏过程（即跳跃过程，也称之为连续时间马氏链）的更多性质和应用。

### 6.7.1 基本概念

#### 定义 6.7.1 (连续时间马氏链)

概率空间  $(\Omega, \mathcal{F}, \mathbb{P})$  上的随机过程  $X = \{X(t), t \geq 0\}$  的状态空间  $S$  为至多可数集。如果  $X$  满足马氏性： $\forall j, i_1, \dots, i_{n-1} \in S, 0 \leq t_1 < \dots < t_n$ ，都有

$$\mathbb{P}(X(t_n) = j | X(t_{n-1}) = i_{n-1}, \dots, X(t_1) = i_1) = \mathbb{P}(X(t_n) = j | X(t_{n-1}) = i_{n-1})$$

并且  $X$  是右连续的，则称  $X$  为连续时间马氏链。

其转移概率定义为

$$p_{ij}(s, t) = \mathbb{P}(X(t) = j | X(s) = i)$$

如果  $p_{ij}(s, t) = p_{ij}(0, t - s)$ ，则称  $X$  为时间齐次的，以下我们总假设时间齐次，并把  $p_{ij}(0, t)$  简记为  $p_{ij}(t)$ 。记矩阵  $\mathbf{P}_t = (p_{ij}(t))_{|S| \times |S|}$ ， $\{\mathbf{P}_t, t \geq 0\}$  称为  $X$  的转移半群。

#### 定理 6.7.1

$\{\mathbf{P}_t, t \geq 0\}$  的性质：

- (1).  $\mathbf{P}_0 = I$ .
- (2).  $\forall p_{ij}(t) \geq 0, \sum_j p_{ij}(t) = 1$ .
- (3). CK 方程： $\mathbf{P}_{s+t} = \mathbf{P}_s + \mathbf{P}_t$ .

所以， $X$  的分布由  $X(0)$  和  $\{\mathbf{P}_t, t \geq 0\}$  决定。

#### 定义 6.7.2 (马氏性更一般化的定义)

对于  $A \in \mathcal{F}$ ，如果  $A \in \sigma(\{X(s), s < t\})$ ，则称  $A$  是  $t$ -historical；如果  $A \in \sigma(\{X(s), s > t\})$ ，则称  $A$  是  $t$ -future。

马氏性： $\forall t > 0$ ，设  $H$  是  $t$ -historical， $F$  是  $t$ -future，则

$$\mathbb{P}(F | X(t) = j, H) = \mathbb{P}(F | X(t) = j), \forall j \in S$$

#### 定义 6.7.3 (停时)

称随机变量  $T$  为关于  $X$  的停时，如果  $\forall t \geq 0, \{T \leq t\} \in \sigma(\{X(s), s \leq t\})$ 。

定义  $T_0 = 0$ ,

$$T_n = \inf\{t > T_{n-1} : X(t) \neq X(T_{n-1})\}, n \geq 1$$

代表  $X$  第  $n$  次跳跃的时刻，记  $U_m = T_{m+1} - T_m$  为  $X$  在第  $m$  个状态下的停留时间。这些都是停时。令

$$T_\infty \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} T_n$$

#### 定理 6.7.2 (强马氏性)

$T$  为关于  $X$  的停时，给定  $I = \{T < T_\infty\} \cap \{X(T) = i\}$  的条件下， $X^* = \{X^*(u) = X(T+u), u \geq 0\}$  与给定  $X(0) = i$  条件下的  $X$  同分布，且与  $\{X(s), s < T\}$  独立。

**定理 6.7.3 (指数时间)**

$U_0$  服从参数为  $g_i$  的指数分布,  $g_i$  只与初始状态  $X(0) = i$  有关。给定  $X(0) = i$  的条件下, 若  $g_i > 0$ , 则  $X(U_0)$  与  $U_0$  独立。

这是我们已经证明过的结论, 参见定理 6.5.1.

**定义 6.7.4 (跳链与生成元)**

对于连续时间马氏链  $X$ , 取离散时间马氏链  $Y = \{Y_n, n \in \mathbb{N}\}$ , 状态空间为  $S$ , 转移概率为:

$$y_{ij} = \begin{cases} \delta_{ij} & , g_i = 0 \\ \mathbb{P}(X(U_0) = j | X(0) = i) & , g_i > 0 \end{cases}$$

$Y$  称为  $X$  的跳链。

令

$$g_{ij} = g_i(y_{ij} - \delta_{ij}) = \begin{cases} g_i y_{ij} & , i \neq j \\ -g_i & , i = j \end{cases}$$

矩阵  $G = (g_{ij})_{|S| \times |S|}$ , 称为  $X$  的生成元。生成元的行之和为 0, 这点与转移矩阵略有不同。

一个连续时间马氏链  $X = \{X_t(\omega), t \geq 0\}$  可以由跳链  $Y$  和生成元  $G$  来描述: 注意到

$$\sum_{j \neq i} g_{ij} = g_i$$

我们取与  $Y$  独立的一系列随机变量  $U_0, U_1, \dots$ , 其中  $U_i$  服从参数为  $g_i$  的指数分布, 并设  $T_n = U_0 + \dots + U_{n-1}$ , 那么

$$X_t(\omega) \stackrel{\text{def}}{=} Y_n(\omega), \text{ where } t \in [T_n(\omega), T_{n+1}(\omega))$$

有一个问题: 如果  $\mathbb{P}(T_\infty < +\infty) > 0$ , 对于  $\omega \in \{T_\infty < +\infty\}$ ,  $t > T_\infty$  无定义。解决办法是: 令  $S' = S \cup \{+\infty\}$ , 且令  $X(t) = +\infty, t \geq T_\infty$ , 此时我们称  $X$  是 minimal markov process.

跳链是很重要的工具概念, 我们接下来会利用跳链对连续时间马氏链进行详细的分析。

**6.7.2 状态分类、状态之间的关系**

我们先考虑  $X$  的跳链  $Y$ .

**定理 6.7.4**

$X$  是连续时间马氏链,  $Y$  是  $X$  的跳链, 给定  $X(0) = i$  条件下,  $X$  不爆炸的充分条件为以下任意之一:

- (1).  $S$  有限。
- (2).  $\sup_j g_j < +\infty$ .
- (3).  $i$  是  $Y$  的常返态。

**证明** (1)  $\Rightarrow$  (2).

(2): 设  $r = \sup_j g_j < +\infty$ ,  $U_n \sim \exp\{g_{Y_n}\}$ , 如果存在某个  $g_{Y_n} = 0$ , 则  $U_n = +\infty \Rightarrow T_\infty = \infty$  a.s. 如果  $g_{Y_n} > 0$ , 则令  $V_n = g_{Y_n} \cdot U_n \sim \exp\{1\}$ , 于是

$$r \cdot T_\infty = \sum_{n=1}^{\infty} r U_n \geq \sum_{n=1}^{\infty} V_n = \infty \text{ a.s.}$$

(3):  $g_i = 0$  则不爆炸;  $g_i > 0$  时,  $i$  是  $Y$  的常返态,  $Y_0 = i$ , 说明存在  $0 = N_0 < N_1 < \dots$  使得  $Y_{N_j} = i$ , 从

而

$$g_i T_\infty = \sum_{j=0}^{\infty} g_i U_j \geq \sum_{j=0}^{\infty} g_i U_{N_j} = \infty \text{ a.s.}$$

□

### 例 6.7.1.

设离散时间马氏链  $Y = \{T_n, n \in \mathbb{N}\}$ , 状态空间为  $S$ , 转移矩阵  $\mathbf{P}_Y = (y_{ij})_{|S| \times |S|}$ , 并且  $\forall y_{ii} = 0$ , 设  $N$  是参数为  $\lambda$  的 Poisson 过程, 并设

$$T_0 = 0, T_n = \inf\{t \geq 0 : N(t) = n\}$$

然后, 我们定义  $X(t) = Y_n$ , 其中  $T_n \leq t < T_{n+1}$ , 则  $X$  是连续时间马氏链, 转移概率为:

$$\begin{aligned} p_{ij}(t) &= \mathbb{P}(X(t) = j | X(0) = i) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(X(t) = j, N(t) = n | X(0) = i) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(Y_n = j, N(t) = n | Y(0) = i) \\ &= \sum_{n=0}^{\infty} \mathbb{P}(N(t) = n) \cdot \mathbb{P}(Y_n = j | Y(0) = i) \\ &= \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n (\mathbf{P}_Y)_{ij}^n}{n!} \\ \Rightarrow \mathbf{P}_X(\mathbf{t}) &= (p_{ij}(t))_{|S| \times |S|} = e^{\lambda t (\mathbf{P}_Y - I)} \end{aligned}$$

这里矩阵指数是指:

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

### 定义 6.7.5

对于连续时间马氏链  $X$ , 如果  $\forall i, j \in S$ , 都存在  $t > 0$  使得  $p_{ij}(t) > 0$ , 则称  $X$  不可约。

### 推论 6.7.1

$|S| = 1$  则显然  $X$  不可约,  $|S| > 1$  时  $X$  不可约  $\Rightarrow \forall g_i > 0$ .

### 定理 6.7.5

连续时间马氏链  $X$  的跳链为  $Y$ , 以下命题等价:

- (1).  $X$  不可约。
- (2).  $\forall i, j \in S, \forall t > 0, p_{ij}(t) > 0$ .
- (3).  $Y$  不可约, 且  $\forall g_i > 0$ .

类似于离散时间马氏链, 我们可以定义连续情形下的常返态:

**定义 6.7.6 (常返与瞬时)**

对于连续时间马氏链  $X$ , 如果状态  $i$  满足

$$\mathbb{P}(\{t \geq 0 : X(t) = i\} \text{ 无界} | X(0) = i) = 1$$

则称  $i$  是  $X$  的常返态; 如果状态  $i$  满足

$$\mathbb{P}(\{t \geq 0 : X(t) = i\} \text{ 无界} | X(0) = i) = 0$$

则称  $i$  是  $X$  的瞬时态。



**定义 6.7.7 (正常返与零常返)**

对于连续时间马氏链  $X$  的常返态  $i$ , 定义最早返回时间:

$$R_i = \inf\{t > U_0 : X(t) = i\}$$

如果  $m_i = \mathbb{E}[R_i | X(0) = i] < +\infty$ , 则称状态  $i$  正常返, 否则称其零常返。



**定理 6.7.6**

连续时间马氏链  $X$  的跳链为  $Y$ ,

- (1). 若  $g_i = 0$ , 则  $i$  为  $X$  的常返态。
- (2). 若  $g_i > 0$ , 则  $i$  在  $X$ 、 $Y$  中的常返/瞬时状态一致。
- (3). 若  $X$  不可约, 则所有状态要么常返, 要么瞬时。

**6.7.3 Kolmogorov 方程**

回顾连续时间马氏链  $X$  和其跳链  $Y$  的关系, 得到转移概率的估计:

**定理 6.7.7**

连续时间马氏链  $X$  的生成元为  $G = (g_{ij})$ , 那么

$$p_{ij}(t+h) \stackrel{\text{def}}{=} \mathbb{P}(X(t+h) = j | X(t) = i) = \delta_{ij} + g_{ij}h + o(h) \quad (*)$$

**证明** 作业。 □

我们借此尝试推导 Kolmogorov 方程。

**定理 6.7.8 (Kolmogorov)**

连续时间马氏链  $X$  的状态空间  $S$  有限, 转移半群为  $\{\mathbf{P}_t, t \geq 0\}$ , 生成元为  $G = (g_{ij})$ , 则:

$$\begin{aligned} \frac{d}{dt} \mathbf{P}_t &= \mathbf{P}_t \cdot G && \text{向前方程} \\ \frac{d}{dt} \mathbf{P}_t &= G \cdot \mathbf{P}_t && \text{向后方程} \end{aligned}$$

边界条件:  $\mathbf{P}_0 = I$ .

**证明** 对于  $p_{ij}(t+h)$ , 考虑用  $i \xrightarrow{t} k \xrightarrow{h} j$  拆分:

$$\begin{aligned} p_{ij}(t+h) &= \mathbb{P}(X(t+h) = j | X(0) = i) \\ &= \sum_{k \in S} \mathbb{P}(X(t+h) = j | X(t) = k) \cdot \mathbb{P}(X(t) = k | X(0) = i) \\ &= \sum_{k \in S} p_{kj}(h) p_{ik}(t) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{k \in S} (\delta_{kj} + g_{kj}h + o(h))p_{ik}(t) \\
 &= p_{ij}(t) + \sum_{k \in S} p_{ik}(t)g_{kj}h + o(h)
 \end{aligned}$$

整理一下得到:

$$\frac{p_{ij}(t+h) - p_{ij}(t)}{h} = \sum_{k \in S} p_{ik}(t)g_{kj} + \frac{o(h)}{h}$$

令  $h \rightarrow 0$  则得到向前方程。类似的, 考虑用  $i \xrightarrow{h} k \xrightarrow{t} j$ , 就得到向后方程。 □

**推论 6.7.2**

连续时间马氏链  $X$  的状态空间  $S$  有限, 转移半群为  $\{\mathbf{P}_t, t \geq 0\}$ , 生成元为  $G = (g_{ij})$ , 则求解 Kolmogorov 方程, 得到唯一解:

$$\mathbf{P}_t = e^{tG} = \sum_{n=0}^{\infty} \frac{G^n}{n!} t^n$$

现在的问题是: 如果  $|S| = \infty$ ,  $\sum_{k \in S}$  变为无穷求和, 能否直接与  $\lim_{h \rightarrow 0}$  交换? 这会导致向后方程

**定理 6.7.9**

$|S| = \infty$ ,  $G$  为  $S$  上的生成元 (即满足非负、行和为 0 的矩阵), 令  $X$  为以  $G$  为生成元的 minimal markov process, 那么:

- (1).  $X$  的转移半群  $\{\mathbf{P}_t\}$  为向后方程的最小非负解。即任何一个满足向后方程的其他解  $\pi_{ij}(t)$ , 都有  $p_{ij}(t) \leq \pi_{ij}(t)$ .
- (2).  $X$  的转移半群  $\{\mathbf{P}_t\}$  也是向前方程的最小非负解。

**证明** 我们只说明 (1), (2) 见 (Norris 1997.p.100).

设  $T_1$  是  $X$  第一次改变初始状态的时刻, 即

$$T_1 = \inf\{t > 0 : X(t) \neq X(0)\}$$

那么,

$$\begin{aligned}
 p_{ij}(t) &= \mathbb{P}(T_1 > t, X(t) = j | X(0) = i) + \mathbb{P}(T_1 \leq t, X(t) = j | X(0) = i) \\
 &= \mathbb{P}(T_1 > t, X(t) = j | X(0) = i) + \sum_{k \neq i} \mathbb{P}(T_1 \leq t, X(t) = j, X(T_1) = k | X(0) = i) \\
 &= \delta_{ij}e^{-g_i t} + \sum_{k \neq i} \int_0^t g_i e^{-g_i s} y_{ik} p_{kj}(t-s) ds
 \end{aligned} \tag{*}$$

这里利用 Fubini 定理交换求和与积分顺序, 并换元  $u = t - s$ , 得到:

$$e^{g_i t} p_{ij}(t) = \delta_{ij} + \int_0^t \sum_{k \neq i} g_i e^{g_i u} y_{ik} p_{kj}(u) du$$

两边对  $t$  求导, 就得到

$$\frac{d}{dt} p_{ij}(t) = \sum_{k \in S} g_{ik} p_{kj}(t), \quad i \in S$$

这就验证了  $\mathbf{P}_t$  符合向后方程。下面证明它是最小非负解, 假设另一个非负解  $\pi_{ij}(t)$  满足向后方程, 它同样满足 (\*) 式。如果设  $T_n$  为第  $n$  次改变初始状态的时刻, 接下来我们利用归纳法证明:

$$\pi_{ij}(t) \geq \mathbb{P}(X(t) = j, T_n > t | X(0) = i), \quad \forall i, j \in S, t > 0, n \geq 1$$

当  $n = 1$  时,

$$\pi_{ij}(t) \geq \delta_{ij}e^{-g_i t} = \mathbb{P}(X(t) = j, T_1 > t | X(0) = i)$$

假设  $1 \leq n \leq N$  都成立, 考虑  $n = N + 1$  的情形,

$$\begin{aligned} & \mathbb{P}(X(t) = j, T_{N+1} > t | X(0) = i) \\ &= \mathbb{P}(X(t) = j, T_{N+1} > t, T_1 > t | X(0) = i) + \mathbb{P}(X(t) = j, T_{N+1} > t, T_1 \leq t | X(0) = i) \\ &= \mathbb{P}(X(t) = j, T_1 > t | X(0) = i) + \sum_{k \neq i} \mathbb{P}(X(t) = j, T_{N+1} > t, T_1 \leq t, X(T_1) = k | X(0) = i) \\ &= \delta_{ij}e^{-g_i t} + \sum_{k \neq i} \mathbb{P}(X(t) = j, T_{N+1} > t, T_1 \leq t | X(T_1) = k) \mathbb{P}(X(T_1) = k | X(0) = i) \\ &= \delta_{ij}e^{-g_i t} + \sum_{k \neq i} \int_0^t g_i e^{-g_i s} \mathbb{P}(X(t-s) = j, T_N > t | X(0) = k) y_{ik} ds \end{aligned}$$

从而由递推可得  $n = N + 1$  的情形成立。

那么, 令  $n \rightarrow +\infty$ ,  $T_n \rightarrow T_\infty$ ,

$$\pi_{ij}(t) \geq \lim_{n \rightarrow \infty} \mathbb{P}(X(t) = j, T_n > t | X(0) = i) = \mathbb{P}(X(t) = j, T_\infty > t | X(0) = i) = p_{ij}(t)$$

□

## 6.7.4 平稳分布

### 定义 6.7.8

类似于离散情形, 对于不可约、不爆炸的连续时间马氏链  $X$ , 其转移半群为  $\{\mathbf{P}_t\}$ , 若  $S$  上的测度  $\pi$  满足

$$\pi \mathbf{P}_t = \pi, \forall t \geq 0$$

则称  $\pi$  为  $X$  的平稳测度。上述  $\pi$  如果是概率测度, 则称其为平稳分布。



回忆: 关于常返态  $i$ , 记

$$R_i = \inf\{t > T_1 : X(t) = i\}$$

那么  $R_i | X(0) = i$  就是首次回到  $i$  的时刻, 其期望记作

$$m_i = \mathbb{E}[R_i | X(0) = i]$$

如果有限, 则称正常返。

### 定理 6.7.10

不可约的连续时间马氏链  $X$  的状态空间  $S$  满足  $|S| \geq 2$ , 若存在  $k \in S$  正常返, 则存在唯一平稳分布  $\pi$ , 并且  $\pi$  还是唯一满足  $\pi G = 0$  的概率测度。

若  $X$  不爆炸, 且存在概率测度  $\pi$  满足  $\pi G = 0$ , 则:

- (1). 所有状态都正常返;
- (2).  $\pi$  为平稳分布;
- (3).  $\pi_k^{-1} = m_k g_k$ .

### 引理 6.7.1

不可约的连续时间马氏链  $X$  的跳链为  $Y$ ,

- (1). 测度  $a$  满足  $aG = 0$  当且仅当测度  $v = (v_i = a_i g_i, i \in S)$  满足  $v \mathbf{P}^Y = v$ . 进一步地若  $X$  常返, 则  $a$  唯一。

(2). 若测度  $a$  满足  $aG = 0$ , 则  $\forall a_j > 0$ .

(3). 假设  $X$  常返, 任取  $k \in S$ , 则令

$$\mu_j(k) = \mathbb{E} \left[ \int_0^{R_k} I_{\{X(s)=j\}} ds \mid X(0) = k \right]$$

则  $\mu(k) = (\mu_j(k), j \in S)$  是  $X$  的平稳测度并且  $\mu(k)G = 0$ .



**例 6.7.2.**

连续时间马氏链  $X$  的状态空间为  $\mathbb{Z}$ , 生成元  $G = (g_{ij})$  为:

$$g_{m,m+1} = \begin{cases} 4^m & , m \geq 0 \\ \frac{1}{2}4^{-m} & , m \leq -1 \end{cases}, \quad g_{m,m-1} = \begin{cases} \frac{1}{2}4^m & , m \geq 1 \\ 4^{-m} & , m \leq 0 \end{cases}$$

$$g_{m,m} = -g_{m,m+1} - g_{m,m-1} = \begin{cases} -\frac{3}{2}4^{|m|} & , m \neq 0 \\ -1 & , m = 0 \end{cases}$$

其余为 0.

注意到  $\forall g_{m,m} < 0 \Rightarrow \forall g_m > 0$ ,  $X$  不可约。方程  $\pi G = 0$  有两组解:

$$\pi^{(1)}(m) = \frac{1}{3}2^{-|m|}$$

$$\pi^{(2)}(m) = \begin{cases} \frac{1}{3}4^{-m} & , m \geq 0 \\ \frac{1}{3} \frac{2^{-m+1}-1}{4^m} & , m \leq -1 \end{cases}$$

这说明  $X$  爆炸。(如果  $X$  不爆炸,  $\pi G = 0$  的解应当存在唯一。)

**定理 6.7.11**

$X$  是连续时间马氏链, 不可约、不爆炸,

(1). 若存在平稳分布  $\pi$ , 则唯一。且  $\forall i, j \in S$ , 有  $\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j$ .

(2). 若不存在平稳分布, 则  $\forall i, j \in S$ , 有  $\lim_{t \rightarrow \infty} p_{ij}(t) = 0$ .

**例 6.7.3.**

生灭过程: 状态空间为  $\mathbb{N}$ , 生成元为:

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & & \\ & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

其中  $\lambda, \mu > 0$ .

**定理 6.7.12 (Detailed Balance Condition)**

分布  $\pi$  满足  $\forall j \neq k$  有  $\pi_j g_{jk} = \pi_k g_{kj}$ , 则  $\pi$  是平稳分布。

## 6.7.5 终止分布问题 (Exit Distributions)

## 定理 6.7.13 (Exit Distributions, ED)

$X$  是连续时间马氏链,  $Z$  是其跳链, 状态空间为  $S$ ,  $A, B \subset S$ ,  $C = S - (A \cup B)$ , 令:

$$V_A = \inf\{t \geq 0 : X_t \in A\}$$

$$V_A^Z = \min\{n \geq 0 : Z_n \in A\}$$

$V_B, V_B^Z$  类似。令  $T = V_A \wedge V_B$ , 且  $\forall i \in C$ ,  $\mathbb{P}(T < +\infty | X_0 = i) > 0$ .

如果函数  $h : S \rightarrow [0, 1]$  满足:  $h(A) = 1$ ,  $h(B) = 0$ ,

$$\sum_{j \in S} g_{ij} h(j) = 0, \forall i \in C$$

则  $h(i) = \mathbb{P}(V_A < V_B | X_0 = i)$ .

## 定理 6.7.14 (Exit Times, ET)

$X$  是连续时间马氏链,  $Z$  是其跳链, 状态空间为  $S$ ,  $A \subset S$ ,  $C = S - A$  是有限集, 并且  $\forall i \in C$ ,  $\mathbb{P}(V_A < +\infty | X_0 = i) > 0$ , 如果函数  $h : S \rightarrow [0, 1]$  满足:  $h(A) = 0$ ,

$$\sum_{j \in S} g_{ij} h(j) = -1, \forall i \in C$$

则  $h(i) = \mathbb{E}[V_A | X_0 = i]$ .