布朗运动

应坚刚

2022 年 12 月 11 日

前言

布朗运动得名于英国生物学家 Robert Brown 在 19 世纪上半叶发表的论文, 他观察和分析了花粉在液体表面的细微的永不停止的无规则随机运动, 疑惑这样的运动究竟源自什么. 现在我们知道运动的原因是分子层面的碰撞. 布朗运动的数学模型在 20 世纪开始成型, 首先 Bachelier 在 1900 年尝试用它建立股票市场的数学模型, 物理学家 A. Einstein 证明其转移函数满足热方程, 因此布朗运动也可以看成是热扩散的机制. 最重要的是, N. Wiener 于 1923 年严格证明布朗运动作为连续函数空间上一个概率测度的存在性. 自此, 对布朗运动的研究一直是概率论的热点, 因为它是最重要的随机过程, 与物理, 金融, 生物以及数学的其它分支都有密切的联系.

布朗运动的存在性证明至今近 100 年, 现在布朗运动既是重要的模型也是重要的数学. 它的主要研究对象是样本轨道的概率性质. 例如, 对于离散时间, 是不是几乎所有样本轨道都会到达所有状态? 如果可以, 到达的平均时间有限还是无限? 对于连续时间的布朗运动, 问题就更多了, 例如, 样本轨道是不是连续的? 是的, 几乎所有轨道连续. 是不是可导的? 不, 几乎所有轨道无处可导. 它的水平集会是什么结构? 几乎所有的水平集是无处稠密但无孤立点的零测度闭集. 轨道会不会总是回来? 不一定, 这和空间维数有关.

本书作为复旦大学数学专业高年级学生之选修课的布朗运动课程的一个粗略的讲义, 共分七章. 第一章是预备知识, 主要是介绍单调类方法和条件期望, 这是进一步学习概率论的必需知识. 第二章是随机游动, 也就是独立同分布的 Bernoulli 随机序列的部分和, 它被看成是布朗运动的简化版, 或者说布朗运动是随机游动的极限. 这里我们只是简单地讨论研究随机游动的首中时以及相关的常返性, 目的是体会布朗运动的轨道性质, 也提前体会布朗运动的研究思想与方法. 第三章是讲布朗运动的存在性, 前面已经说到, 布朗运动存在性在历史上是一个困扰许多著名的数学家多

年的大问题, 也是概率论历史上的大问题, 甚至早于概率论的公理化. 我们在这一章循序渐进地把这个问题展现清楚, 顺便讲述 Kolmogorov 的相容性定理. 第四章主要讨论布朗运动的一些简单性质. 第五章介绍布朗运动的马氏性, 强马氏性, 以及位势性质, 其中包括常返暂留性质与 Dirichlet 问题的概率解. 第六章简单介绍关于布朗运动的鞅性质. 最后一章介绍 Itô 积分与 Itô 公式, 作为 Itô 公式的应用, 介绍布朗运动的局部时. 教师可以视课时情况从最后一章选取适当的内容. 讲义中基本上每个结论都给予证明或者提示留给学生完成. 总的来说, 课程内容差不多适用于一个 15 周每周 3 课时的课程. 课程兼顾直观和具体, 最终的目标是让学生对布朗运动轨道的随机与概率性质有较深入的理解, 体会到布朗运动数学的一面和物理的另一面.

应坚刚, 2021 年
复旦大学

目录

第一章　预备知识

可测空间 $(E,\mathcal{E})$ 是指一个集合与它的子集组成的一个 σ-代数, 可测空间上的一个概率测度 μ 也被称为是一个分布. 对于可测函数 f, $\mu(f)$ 表示 f 关于 μ 的积分

$$\mu(f)=\int_E f(x)\mu(dx).$$

这里可测集 $A\in\mathcal{E}$ 的测度 $\mu(A)=\mu(1_A)$, 其中 1_A 是 A 的示性函数.
概率论中有习惯用的名词和符号. 通常用 Ω 表示样本空间, 它的一个 σ-代数 $\mathcal{F}$ 被称为事件域, 其中的元素被称为事件, $(\Omega,\mathcal{F})$ 上的概率测度通常用 P 表示, 对于 $A\in\mathcal{F}$, $P(A)$ 是事件 A 的概率, Ω 上的有限值可测函数 X 被称为随机变量, 它关于 P 的积分 $P(X)$ 也习惯用符号 $E(X)$ 或者 $E[X]$, EX 表示, 被称为 X 的数学期望, 简称期望. 注意期望有平均的意思, 所以它仅用于关于概率测度的积分. 当然一样地 $P(A)=E[1_A]$. 随机变量在事件 A 上的积分 $E[X1_A]$ 也表示为 $E[X,A]$. 注意在本讲义中, 无需说明, E 总是对应于 P 的符号. 在讨论其它概率测度时, 我们仍然使用测度符号或者明确指定符号来表达对应的期望以免引起混淆. 三元组

$$(\Omega,\mathcal{F},P)$$

被称为概率空间. 这也是概率论的公理与研究概率问题的起点.
概率测度 P 可以通过可测映射诱导到任何可测空间上. 设 ξ 是 $(\Omega,\mathcal{F})$ 到 $(E,\mathcal{E})$ 的可测映射, 即对任何 $A\in\mathcal{E}$,

$$\xi^{-1}(A)\in\mathcal{F}.$$

这时我们定义

$$\mu(A):=P(\xi^{-1}(A)),$$

它是 $(E, \mathcal{E})$ 上的一个分布, 称为 P 在 ξ 下的像测度, 也记为 $P\circ\xi^{-1}$, 即

$$P\circ\xi^{-1}(A) = P(\xi^{-1}(A)).$$

对于 E 上的可测函数 f, $f\circ\xi$ 是 Ω 上的随机变量, 且 f 关于像测度的积分

$$P\circ\xi^{-1}(f) = P(f\circ\xi) = E[f\circ\xi]. \tag{1.1}$$

像测度是概率论的重要工具.

本章将介绍单调类方法与条件数学期望, 前者是一个基本方法, 后者是一个基本概念.

1.1　单调类方法

单调类方法是概率论一个不可缺少的工具, 在本课程中经常用到. 这部分内容理论性比较强, 初学的同学可能会觉得困难, 对于学习大学概率论也不是绝对的必要, 所以大学概率论不一定讲单调类方法, 但对于现在这个课程, 单调类方法却是必须的, 读者需用心学习.

在很多情况下, 特别是涉及概率测度时, 很难直接验证一个集类对可列并封闭, 但由于概率有可列可加性, 很容易得到对不交可列并的封闭性. 单调类方法是解决这个困境的一个有效方法. 它有好多不同的形式, 我们下面叙述的是单调类定理的 Dynkin 形式. 称一个子集类 $\mathcal{F}_0$ 是 π- 类, 如果对任何 $A, B \in \mathcal{F}_0$, 有 $A \cap B \in \mathcal{F}_0$. 而称一个子集类是 Dynkin 类 (或 λ- 类), 如果它包含有 $\varnothing, \Omega$ 且对于补集运算与不交可列并运算封闭. 显然, σ- 域是 Dynkin 类, 反之不对. 容易看出任意多个 Dynkin 类的交仍是 Dynkin 类, 因此对 Ω 的任何子集类 $\mathcal{A}$, 唯一存在一个包含 $\mathcal{A}$ 的最小 Dynkin 类, 记为 $\delta(\mathcal{A})$, 也类似地称为由 $\mathcal{A}$ 生成的 Dynkin 类. 下面的定理通常称为 Dynkin 引理，因为是他首次作为一个引理写在他的书上.

定理 1.1　设 $\mathcal{F}_0$ 是一个 π- 类, $\mathcal{F}$ 是一个 Dynkin 类且 $\mathcal{F} \supset \mathcal{F}_0$. 则 $\mathcal{F} \supset \sigma(\mathcal{F}_0)$.

证明. 因为 $\mathcal{F} \supset \delta(\mathcal{F}_0)$, 故证明 $\delta(\mathcal{F}_0)$ 是一个 σ- 域就够了. 由定义, 仅须验证 $\delta(\mathcal{F}_0)$ 对有限交运算封闭. 任取 $A \in \delta(\mathcal{F}_0)$, 定义

$$\kappa[A] := \{B \in \delta(\mathcal{F}_0) : \ A \cap B \in \delta(\mathcal{F}_0)\}.$$

只要证明对任何 $A \in \delta(\mathcal{F}_0)$, $\kappa[A] \supset \delta(\mathcal{F}_0)$.
先验证 $\kappa[A]$ 是一个 Dynkin 类. 事实上, 只需证明 $\kappa[A]$:

(1) 对补集运算封闭;

(2) 对不相交集列的可列并运算封闭.

对 (1), 取 $B \in \kappa[A]$, 则 $A, A^c, A \cap B \in \delta(\mathcal{F}_0)$, 因此 $A \cap B^c = [A^c \cup (A \cap B)]^c \in \delta(\mathcal{F}_0)$, 因此 $B^c \in \kappa[A]$. 为证 (2), 取 $\{B_n\} \subset \kappa[A]$ 是不交集列, 则显然 $\{A \cap B_n\}$ 是 $\delta(\mathcal{F}_0)$ 中不交集列, 因此 $A \bigcap (\bigcup_n B_n) \in \delta(\mathcal{F}_0)$, 推出 $\bigcup_n B_n \in \kappa[A]$.
因 $\mathcal{F}_0$ 是 π- 类, 故 $A \in \mathcal{F}_0$ 蕴含着 $\kappa[A] \supset \mathcal{F}_0$, 即 $\kappa[A] \supset \delta(\mathcal{F}_0)$, 或者说 $\mathcal{F}_0$ 中的集合与 $\delta(\mathcal{F}_0)$ 中的集合的交在 $\delta(\mathcal{F}_0)$ 中. 再换句话说, 当 $A \in \delta(\mathcal{F}_0)$ 时, $\kappa[A] \supset \mathcal{F}_0$. 因此 $\kappa[A] \supset \delta(\mathcal{F}_0)$, 即 $\delta(\mathcal{F}_0)$ 中元素对有限交运算封闭. □

比如, 如果读者愿意仔细地完成下面的命题的证明, 他们会发现单调收敛定理结合 Dynkin 引理是非常重要且强大的工具. 设 ξ, η 是两个 n 维随机变量, 那么下面断言是等价的:

(1) 对任何 $x \in \mathbf{R}^n$, $P(\xi \leqslant x) = P(\eta \leqslant x)$;

(2) 对任何 Borel 集 B, $P(\xi \in B) = P(\eta \in B)$;

(3) 对任何有界 (或非负) Borel 可测函数 f, $Ef(\xi) = Ef(\eta)$.

从 (1) 推出 (2) 是 Dynkin 引理, 从 (2) 推出 (3) 是单调收敛定理, 把这样两种方法结合起来用称为单调类方法. 它是分析特别是概率论的基本方法之一, 现在成为一种标准的证明程序. 为了展示单调类方法的威力, 让我们证明关于交换积分次序的著名的 Fubini 定理, 它是分析中一个非常有用的结果, 但在本书中几乎没有用到, 除了在最后的特征函数唯一性定理用到一点 Fubini 定理以外.

定理 1.2 **如果 ξ, η 独立, 则 Fubini 定理成立: 对任何 $\mathbf{R}^2$ 上非负或有界 Borel 可测函数 f, 有**

$$Ef(\xi, \eta) = E[Ef(x, \eta)|_{x=\xi}] = E[Ef(\xi, y)|_{y=\eta}], \tag{1.2}$$

右边的含义是先固定 x 求期望然后把 ξ 代入 x 后再求期望. 也就是说我们可分两步来求期望.

证明. 首先满足 Fubini 定理的可测函数全体作为 $\mathcal{L}$ 满足单调类定理的条件, 设 $f(x,y) = 1_{A\times B}(x,y) = 1_A(x)1_B(y)$, 其中 A, B 是 $\mathbf{R}$ 的区间. 那么由独立性推得

$$E(1_A(\xi)1_B(\eta)) = E1_A(\xi)E1_B(\eta) = E\{[E1_A(x)1_B(\eta)]\big|_{x=\xi}\},$$

因此 $\mathcal{L}$ 包含矩形全体 $\mathcal{F}_0$, 它是 π- 类, 由单调类定理推出定理结论. □

注意 $f(\xi) = f(x)|_{x=\xi}$ 这样的表示在本讲义后面经常使用. 如果 F 和 G 分别是 ξ 与 η 的分布函数, 那么 Fubini 定理可以写成积分形式:

$$\begin{aligned}\int_{\mathbf{R}^2} f(x,y)\mathrm{d}_2F(x)G(y) &= \int_{\mathbf{R}} \mathrm{d}F(x)\int_{\mathbf{R}} f(x,y)\mathrm{d}G(y) \\ &= \int_{\mathbf{R}} \mathrm{d}G(y)\int_{\mathbf{R}} f(x,y)\mathrm{d}F(x).\end{aligned} \tag{1.3}$$

同样的方法可以证明数学期望和 Riemann 积分是可以交换的. 如果 f 是非负 Borel 可测函数, 那么

$$E\left(\int_{\mathbf{R}} f(\xi,x)\mathrm{d}x\right) = \int_{\mathbf{R}} E[f(\xi,x)]\mathrm{d}x. \tag{1.4}$$

上面的三个公式是 Fubini 定理的不同表现形式. 用一个例子 (§1.5 的习题) 简单表现 Fubini 定理的重要.

例 1.1 如果 ξ 是非负随机变量, 那么由 Fubini 定理, 有

$$E\xi = E\int_0^{\xi} \mathrm{d}t = \int_0^{\infty} P(\xi > t)\mathrm{d}t.$$

类似地, 有

$$E\xi^n = \int_0^{\xi} nt^{n-1}\mathrm{d}t = \int_0^{\infty} nt^{n-1}P(\xi > t)\mathrm{d}t.$$

如果 F 是分布函数, 那么对任何 $a > 0$, 有

$$\int_{\mathbf{R}} (F(x+a) - F(x))\mathrm{d}x = a.$$

事实上, 有

$$\int_{\mathbf{R}} (F(x+a) - F(x))\mathrm{d}x = \int_{\mathbf{R}} P(x < \xi \leqslant x+a)\mathrm{d}x = E\int_{\xi-a}^{\xi} \mathrm{d}x = a.$$

想想不用 Fubini 定理应该怎么做呢? ▮

Euclid 空间上的阶梯函数是矩形区间上示性函数的有限线性组合. 连续函数是阶梯函数的极限, 所以对矩形区间成立的等式通常对连续函数也成立, 上面两个定理的证明都是这个思想. 反过来, 我们来验证 Euclid 空间的开集 (或闭集) 上的示性函数也可以表达为连续函数的递增 (或递减) 极限, 因此对连续函数成立的等式通常对开集或闭集上的示性函数也成立. 事实上, 设 F 是 $\mathbf{R}^d$ 的一个非空闭集, $d(x,F)$ 表示点 x 到集合 F 的距离, 它是 x 的连续函数, 那么 $F = \{x : d(x,F) = 0\}$. 定义 $\phi_n(x) := (n \cdot d(x,F)) \wedge 1$, 那么 ϕ_n 是连续函数列且点点递增收敛于 1_{F^c}, 因为在 F 上, $\phi_n = 0$, 而当 $x \in F^c$ 时, $d(x,F) > 0$, 故存在 n 使得 $n \cdot d(x,F) \geqslant 1$, 推出 $\phi_n(x) = 1$. 而闭集 F 上的示性函数表示为 $1_F = 1 - 1_{F^c}$.

定理 1.3 设 G 度量空间上的开集. 则存在连续函数函数列 ϕ_n 递增收敛于 1_G 且 $|\phi_n| \leqslant 1$.

练习 1.1 设 ξ,η 是两个随机变量. 证明: 如果对任何有界连续函数 f 有 $Ef(\xi) = Ef(\eta)$, 则 ξ, η 同分布.

1.2　条件期望

人无法预知随机事件的结果, 但总有一个预期, 这个预期就是期望. 预期随着时间的推移以及信息的增加会改变, 这就是条件期望. 条件期望是概率论最重要的概念之一, 是 Kolmogorov 在他 20 世纪 30 年代为概率论引入公理体系的那本著名的书中引入的. 下面我们定义随机变量关于 σ- 域的条件期望 (或条件数学期望).

首先我们说明为什么把事件域看成信息. 事件域类似于对样本空间的一个划分, 它体现出我们可以判断一个样本点在什么地方的信息. 比如说我们不知道一个人具体情况, 但知道他是哪个省的人就是一种信息. 再比如有一张 10000×10000 像素的照片要放在一个 1000×1000 像素的显示器上, 那么 10×10 个像素要合并成一个, 信息肯定会丢失. 所以我们可以把概率空间中抽象的事件域理解为信息, 事件域越大体现出的信息越多.

让我们使用 Hilbert 空间的一些概念和术语, Hilbert 空间是内积空间, 因此有正交或者垂直的概念. 考虑平方可积随机变量的空间 $L^2(\mathcal{F}) = L^2(\Omega, \mathcal{F}, P)$, 配以随机变量乘积的期望作为它的内积, 得

$$\langle \xi, \eta \rangle = E[\xi\eta]$$

是一个 Hilbert 空间. 首先一个随机变量 ξ 的数学期望就是 ξ 到子空间 $\mathbf{R}$ 的投影, 即

$$E(\xi - E\xi)^2 = \min\{E(\xi - x)^2 : x \in \mathbf{R}\}. \tag{1.5}$$

L^2 意义下的最近点唯一.

练习 1.2 证明 (1.5). 再证明: 设 ξ 的分布函数 F 连续, $E|\xi - m| = \min\{E|\xi - x| : x \in \mathbf{R}\}$, 点 m 称为 F 的中点, 则 $F(m) = 1/2$. L^1 意义下的最近点不一定唯一.

取一个子事件域 $\mathcal{G} \subset \mathcal{F}$, 那么 $\mathcal{G}$ 可测的平方可积随机变量空间 $L^2(\mathcal{G})$ 是一个闭子空间. 对任何平方可积随机变量 ξ, 它在 $L^2(\mathcal{G})$ 上的正交投影也是这个空间距离 ξ 的最近之处, 是存在且唯一的, 记为 $E(\xi|\mathcal{G})$. 用概率论的语言说, 它是基于已知信息 $\mathcal{G}$ 的条件下对随机变量 ξ 的最佳预测, 称为 ξ 关于 $\mathcal{G}$ 的条件期望.

我们将对可积随机变量定义条件数学期望.

定义 1.1 **设 ξ 是可积随机变量, $\mathcal{G}$ 是 $\mathcal{F}$ 的子事件域, 那么 $E(\xi|\mathcal{G})$, 称为 ξ 关于 $\mathcal{G}$ 的条件期望, 是满足下面条件的可积随机变量 η:**

(1) **η 是 $\mathcal{G}$ 可测的;**

(2) **对任何 $A \in \mathcal{G}$, $E(\xi 1_A) = E(\eta 1_A)$.**

当 $\xi = 1_A$ 时, $E[1_A|\mathcal{G}]$ 也写成为 $P(A|\mathcal{G})$.

定理 1.4 **满足上面两个条件的 η 唯一存在.**

证明. 先证明唯一性. 满足两个条件的 η 是 (a.s.) 唯一的. 实际上, 设 η 是 $\mathcal{G}$ 可测是且对任何 $A \in \mathcal{G}$ 有 $E[\eta 1_A] = 0$. 则

$$E[\eta 1_{\{\eta > 0\}}] = E[\eta 1_{\{\eta \leqslant 0\}}] = 0.$$

推出 $\eta 1_{\{\eta>0\}} = \eta 1_{\{\eta \leqslant 0\}} = 0$ a.s. 因此 $\eta = 0$ a.s.

再证明存在性. 实际上 η 是 $\mathcal{G}$ 上测度 $E[\xi 1_A]$ 关于测度 $P|_{\mathcal{G}}$ 的密度. 但是这个需要很多前序知识. 这里我们借用 Hilbert 空间的定理来证明. 用 L^1 表示可积随机变量全体, 装备 L^1 范数成为 Banach 空间, $E\xi^2 < \infty$ 的随机变量称为是平方可积的, 用 L^2 表示平方可积随机变量全体, 装备可内积化的 L^2 范数成为 Hilbert 空间. 由 Cauchy-Schwarz 不等式, $L^2 \subset L^1$, 但是 L^2 在 L^1 中稠. 用 $L^2(\mathcal{G})$ 表示 $\mathcal{G}$ 可测的平方可积随机变量全体, 它是 L^2 的闭子空间. 如果 $\xi \in L^2$, 那么 ξ 在 $L^2(\mathcal{G})$ 上的正

交投影 η 就是 $E(\xi|\mathcal{G})$. 用定义验证, 只要证对任何 $A \in \mathcal{G}$, $E(\xi; A) = E(\eta; A)$ 或者 $E[(\xi-\eta)1_A] = 0$, 而这由 $\xi - \eta$ 正交于 $L^2(\mathcal{G})$ 推出. 一般地, 如果 $\xi \in L^1$, 那么存在 $\{\xi_n\} \subset L^2$ 使得 $E|\xi_n - \xi| \to 0$. 那么 $\{E(\xi_n|\mathcal{G})\}$ 存在且是 L^1 中的 Cauchy 列, 不难验证其极限就是 $E(\xi|\mathcal{G})$. □

因为条件期望 $E(\xi|\mathcal{G})$ 只是在几乎处处相等的意义下唯一决定的, 故关于条件期望的等式或不等式都是在几乎处处的意义下叙述并证明的. 例如由定义容易验证, 如果 ξ 是 $\mathcal{G}$ 可测可积随机变量, 则 $E(\xi|\mathcal{G}) = \xi$ a.s.. 我们简单地写成 $E(\xi|\mathcal{G}) = \xi$.

先来看看离散事件域的情况. 固定概率空间 $(\Omega, \mathcal{F}, P)$. 设 ξ 是可积随机变量, 对任何正概率的事件 A, 定义

$$E(\xi|A) := \frac{E(\xi; A)}{P(A)},$$

称为 A 发生的条件下 ξ 的期望. 它是 ξ 在事件 A 上的平均, 表示 ξ 在概率 $P(\cdot|A)$ 下的期望. 给定一个离散 σ- 域 $\mathcal{G}$, 它总是由 Ω 的一个划分 $\{\Omega_n : n \geqslant 1\}$ 生成, 即 $\mathcal{G} = \sigma(\Omega_n : n \geqslant 1)$. 随机变量关于离散事件域的条件期望的存在唯一性由下面的定理确立.

定理 1.5 **设** ξ **可积. 如果** $\mathcal{G}$ **是离散事件域, 由划分** $\{\Omega_n : n \geqslant 1\}$ **生成且对任何** n, $P(\Omega_n) > 0$, **那么**

$$E(\xi|\mathcal{G}) = \sum_{n \geqslant 1} E(\xi|\Omega_n) 1_{\Omega_n}.$$

因此

$$E\xi = E(E(\xi|\mathcal{G})).$$

首先证明一个引理, 非空事件 $A \in \mathcal{G}$ 称为是 $\mathcal{G}$ 的原子, 如果 $B \subset A$ 且 $B \in \mathcal{G}$, 那么 $B = \varnothing$ 或 $B = A$, 即 A 除了 $\varnothing, A$ 外不包含其他事件.

引理 1.1 **若** ξ **是** $\mathcal{G}$ **可测的随机变量**, A **是** $\mathcal{G}$ **的原子, 那么** ξ **在** A **上是常数**.

证明. 设 a 是 ξ 在事件 A 的某个给定点上的值. 则 $\{\xi = a\} \in \mathcal{G}$ 且它与 A 的交不空, 由于 A 是原子, 故必有 $\{\xi = a\} \supset A$, 即 ξ 在 A 上恒等于 a. □

证明. (定理的证明) 因为 Ω_n 是 $\mathcal{G}$ 的原子, 由引理, $E(\xi|\mathcal{G})$ 在 Ω_n 上是常数, 记为 a_n, 再由条件期望的定义, $E(E(\xi|\mathcal{G}); \Omega_n) = E(\xi; \Omega_n)$. 左边等于 $a_n P(\Omega_n)$, 因此, 当

$P(\Omega_n) > 0$ 时, $E(\xi|\mathcal{G})$ 在 Ω_n 上是常数, 就是 $E(\xi|\Omega_n)$. 现在, 仿照全概率公式的证明, 有

$$E\xi = \sum_{n\geqslant 1} E(\xi;\Omega_n) = \sum_{n\geqslant 1} E(\xi|\Omega_n)P(\Omega_n) = E(E(\xi|\mathcal{G})).$$

最后的等式是来自离散随机变量期望公式. □

这个定理是说 $E(\xi|\mathcal{G})$ 在 Ω_n 上等于 ξ 在 Ω_n 上的期望或者积分平均. 电视里经常用的马赛克制作方法就是用的这个公式, 马赛克是把若干像素的颜色变成同一个颜色, 这个颜色最自然的选择就是用这些像素上的颜色的平均值. 这也是把一张高像素图片压缩为低像素图片的方法. 定理中第二个公式是全概率公式的推广.
随机变量关于一般 σ- 域的条件期望有下面的性质, 这些性质完全是由定义推导的.

定理 1.6 条件期望有下面的重要性质:

1. $\xi \mapsto E(\xi|\mathcal{G})$ **作为关于 $\mathcal{F}$ 可测的可积随机变量空间到关于事件域 $\mathcal{G}$ 可测的可积随机变量空间的映射是线性的, 保序的;**

2. **全概率公式的推广:** $E(E(\xi|\mathcal{G})) = E\xi$;

3. **如果 ξ, η 是随机变量, ξ 与 $\xi\eta$ 可积且 η 是关于事件域 $\mathcal{G}$ 可测的随机变量, 那么** $E(\xi\eta|\mathcal{G}) = \eta E(\xi|\mathcal{G})$;

4. **如果 ξ 与 $\mathcal{G}$ 独立, 那么** $E(\xi|\mathcal{G}) = E\xi$. **$\xi$ 与 $\mathcal{G}$ 独立是指对任何 $A \in \mathcal{G}$, ξ 与 1_A 独立;**

5. **设 $\mathcal{G}_1, \mathcal{G}_2$ 是子事件域, 且 $\mathcal{G}_1 \subset \mathcal{G}_2$. 那么**

$$E(E(\xi|\mathcal{G}_1)|\mathcal{G}_2) = E(E(\xi|\mathcal{G}_2)|\mathcal{G}_1) = E(\xi|\mathcal{G}_1);$$

6. $E(\xi|\{\varnothing, \Omega\}) = E\xi$, $E(\xi|\mathcal{F}) = \xi$.

证明. 1. 留作习题.

2. 按照条件期望的定义

$$E(E(\xi|\mathcal{G})) = E(E(\xi|\mathcal{G});\Omega) = E(\xi;\Omega) = E(\xi).$$

3. 首先对 $\eta = 1_B$, $B \in \mathcal{G}$, 结论显然, 再由单调类定理推出一般场合.

4. 对任何 $A \in \mathcal{G}$, $E(\xi; A) = E\xi 1_A = E\xi \cdot P(A) = E(E\xi; A)$.

5. 因为 $E(\xi|\mathcal{G}_1)$ 是 $\mathcal{G}_2$ 可测的, 故 $E(E(\xi|\mathcal{G}_1)|\mathcal{G}_2) = E(\xi|\mathcal{G}_1)$ 是显然的. 要证明另一个等式, 两者都是 $\mathcal{G}_1$ 可测的, 只需证明对任何 $A \in \mathcal{G}_1$, $E(\xi; A) = E(E(\xi|\mathcal{G}_2); A)$, 而这由于 $A \in \mathcal{G}_2$ 是显然的.

6. 由定义显然.

□

从条件期望保序性推出
$$|E[\xi|\mathcal{G}]| \leqslant E[|\xi||\mathcal{G}].$$

例 1.2 从全国公民中任选一人记其年收入为 ξ. 样本空间是公民全体, 概率是等可能概率. ξ 的分布是全国公民年收入的分布, 期望 $E\xi$ 是全国公民平均年收入. $\Omega_1, \cdots, \Omega_n$ 是被选取的人在每个省的事件, 自然 $\{\Omega_i : i \leqslant n\}$ 是样本空间的划分, 这个作为信息告诉我们所选的人在什么省, 因此 $E(\xi|\{\Omega_i\})$ 是一个随机变量, 在 Ω_i 上等于这个省的平均年收入 $P(\xi|\Omega_i)$. 也就是说在你得知此人属于哪个省的情况下, 你对其年收入的最佳预测是这个省的平均年收入. 公式 $E\xi = E(E(\xi|\{\Omega_i\}))$ 说明平均年收入等于各省平均年收入的加权平均. ▮

下面我们来看看怎么样具体计算随机变量关于随机向量的条件期望, 这是非常重要的. 如果 η 是离散随机变量 (或随机向量), 那么 $\{\{\eta = x\} : x \in R(\eta)\}$ 是 Ω 的划分. 由上面的定理 1.5, 有
$$E(\xi|\eta) = \sum_{y \in R(\eta)} E(\xi|\eta = y) 1_{\{\eta = y\}}.$$
换句话说, 定义 $f(x) := E(\xi|\eta = x)$, $x \in R(\eta)$, 那么 $E(\xi|\eta) = f(\eta)$. 设 (ξ, η) 是 2 维离散随机变量, $x \in R(\xi), y \in R(\eta)$, 记条件概率 $P(\xi = x|\eta = y)$ 为 $f_{\xi|\eta}(x|y)$. 那么条件数学期望有表达式 $P(\xi = x|\eta) = f_{\xi|\eta}(x|\eta)$, 其中右边理解为 η 代入 $f_{\xi|\eta}(x|y)$ 的 y 位置. 因为当 y 属于 η 的值域时, 有
$$P(\xi = x, \eta = y) = f_{\xi|\eta}(x, y) P(\eta = y) = E(f_{\xi|\eta}(x, \eta); \eta = y).$$

我们称 $x \mapsto f_{\xi|\eta}(x|y)$ 是 $\eta = y$ 时 ξ 的条件分布律, 并且条件期望就是条件分布律的期望, 即

$$E(\xi|\eta) = \sum_{x \in R(\xi)} xP(\xi = x|\eta).$$

例 1.3 从一个有 a 个白球, b 个黑球的袋中任取 n 个球, 放入另一个袋子中. 然后从中任取一球. 问如此方式取得白球和直接从原袋子中取一个球是白球的概率是否相同? 显然直接从原袋子中取一个球是白球的概率是 $\dfrac{a}{a+b}$. 设从原袋子中取的 n 个球中的白球数是 ξ, 放入另一个袋子, 从中取一个球是白球的事件为 A, 那么 $P(A|\xi = k) = k/n$. 因此 $P(A|\xi) = \xi/n$. 但是 $E\xi = \dfrac{na}{a+b}$, 因此

$$P(A) = E(P(A|\xi)) = \frac{E\xi}{n} = \frac{a}{a+b}.$$

两种方式取得白球概率相同. ∎

例 1.4 设 X_1, X_2 独立, 分别服从参数为 m, p 与 n, p 的二项分布, 那么 $X_1 + X_2$ 服从参数为 $n+m, p$ 的二项分布. 对任何 $0 \leqslant k \leqslant l$, 在给定 $X_1 + X_2 = l$ 的条件下, $X_1 = k$ 的概率为

$$P(X_1 = k|X_1 + X_2 = l) = \frac{P(X_1 = k, X_2 = l-k)}{P(X_1 + X_2 = l)} = \frac{\begin{pmatrix} m \\ k \end{pmatrix}\begin{pmatrix} n \\ l-k \end{pmatrix}}{\begin{pmatrix} m+n \\ l \end{pmatrix}},$$

即条件分布是超几何分布. 那么 $E(X_1|X_1 + X_2 = l) = \dfrac{lm}{n+m}$, 因此给定 $X_1 + X_2$, X_1 的条件期望为

$$E(X_1|X_1 + X_2) = \frac{m}{m+n}(X_1 + X_2).$$

类似地, 如果 X_1, X_2 独立且分别服从参数为 λ_1 与 λ_2 的 Poisson 分布, 那么 $X_1 + X_2$ 是参数为 $\lambda_1 + \lambda_2$ 的 Poisson 分布, 而给定 $X_1 + X_2 = n$ 的条件下, X_1 的分布是参数为 $n, \dfrac{\lambda_1}{\lambda_1 + \lambda_2}$ 的二项分布, 因此 $E(X_1|X_1 + X_2 = n) = \dfrac{\lambda_1}{\lambda_1 + \lambda_2}n$, 即推出

$$E(X_1|X_1 + X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}(X_1 + X_2).$$

∎

一般情况下, $\{\eta=y\}$ 的概率可能是零, 条件概率 $P(\xi=x|\eta=y)$ 没有意义. 因此为了讨论一般随机变量的问题, 我们先证明下面的引理.

引理 1.2　随机变量 $\xi'=E(\xi|\eta)$ 当且仅当 ξ' 关于 $\sigma(\eta)$ 可测且对任何 $y\in\mathbf{R}$, 有

$$E(\xi';\eta\leqslant y)=E(\xi;\eta\leqslant y).$$

证明. 必要性显然, 让我们证明充分性, 我们只需验证 $E(\xi';\eta\in B)=E(\xi;\eta\in B)$ 对任何 $B\in\mathcal{B}$ 成立就足够了. 如何从区间到 Borel 集? 我们要求助于 Dynkin 引理, 即定理 1.1. 这种方法是概率论的常用方法. 记 $\mathcal{F}$ 是使得上式成立的 Borel 集 B 的全体, 那么条件说明 $\mathcal{F}$ 包含 $\mathcal{B}_0:=\{(-\infty,x]:x\in\mathbf{R}\}$, 此集类对于交封闭, 现在的问题是我们无法简单地验证 $\mathcal{F}$ 是 σ- 域, 也就是无法直接说 $\mathcal{F}\subset\sigma(\mathcal{B}_0)$. 但是我们很容易验证稍弱点的结论, 就是 $\mathcal{F}$ 对于补集运算与不交可列并运算封闭, 因此是 Dynkin 类. □

注意 ξ' 关于 $\sigma(\eta)$ 可测当且仅当它是 η 的函数, 即存在 Borel 可测函数 ϕ 使得 $\xi'=\phi(\eta)$ (参考 §2.1 的习题), 我们只要把 ϕ 表达出来就可以了. 现在设 (ξ,η) 是 2 维随机变量, F 是它们的联合分布函数. 那么对任何 $y\in\mathbf{R}$, 有

$$E(\phi(\eta);\eta\leqslant y)=E(\xi;\eta\leqslant y).$$

左边可以用 η 的分布函数表达, 右边要用联合分布函数表达, 即

$$\int_{-\infty}^{y}\phi(t)dF_\eta(t)=\int_{\mathbf{R}}\int_{-\infty}^{y}s\,dF(s,t).$$

如果 F 有密度函数 f, 那么 ξ,η 有边缘密度 f_ξ 与 f_η, 因此上面的公式可以写成

$$\int_{-\infty}^{y}\phi(t)f_\eta(t)dt=\int_{\mathbf{R}}\int_{-\infty}^{y}sf(s,t)dsdt.$$

如果密度函数性质足够好, 比如分段连续, 那么两边对 y 求导得

$$\phi(y)f_\eta(y)=\int_{\mathbf{R}}sf(s,y)ds,$$

因此在密度正的地方, 也就是在 η 的值域范围内有

$$\phi(y)=\int_{\mathbf{R}}s\cdot\frac{f(s,y)}{f_\eta(y)}ds.$$

注意, 在固定 y 时, 函数

$$s \mapsto \frac{f(s,y)}{f_\eta(y)}$$

实际上是一个密度函数, 称为是给定 $\eta = y$ 的条件下 ξ 的密度, 或简单地称为条件密度函数, 通常记为 $f_{\xi|\eta}(\cdot|y)$, 而函数 $\phi(y)$ 是条件密度的期望, 形式上也用符号 $E(\xi|\eta = y)$ 表示. 尽管这个符号在 $P(\eta = y) = 0$ 时没有真实的意义, 但是它并不是毫无理由的, 因为实际上我们有下面的等式 (几乎处处的意义下):

$$E(\xi|\eta = y) = \lim_{\delta\to 0} E(\xi|\eta \in (y-\delta, y+\delta)).$$

因为

$$\begin{aligned} E(\xi|\eta \in (y-\delta, y+\delta)) &= \frac{\displaystyle\int_{\mathbf{R}} ds \int_{(y-\delta,y+\delta)} sf(s,t)dt}{\displaystyle\int_{(y-\delta,y+\delta)} f_\eta(t)dt} \\ &= \frac{\displaystyle\int_{\mathbf{R}} ds \frac{1}{2\delta}\int_{(y-\delta,y+\delta)} sf(s,t)dt}{\displaystyle\frac{1}{2\delta}\int_{(y-\delta,y+\delta)} f_\eta(t)dt}, \end{aligned}$$

当 δ 趋于零时, 其极限恰等于 $E(\xi|\eta = y)$. 这正是计算条件期望最常用的方法.

例 1.5 设 (ξ, η) 是二元正态分布的, 其边缘分布都是标准正态分布, 且相关系数是 r. 计算 $E(\xi|\eta)$. 先算条件密度, 有

$$\begin{aligned} f_{\xi|\eta}(x|\eta) &= \frac{1}{2\pi\sqrt{1-r^2}} e^{-\frac{x^2-2rx\eta+\eta^2}{2(1-r^2)}} \cdot \frac{1}{\frac{1}{\sqrt{2\pi}}e^{-\frac{\eta^2}{2}}} \\ &= \frac{1}{\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{x^2-2rx\eta+r^2\eta^2}{2(1-r^2)}\right) \\ &= \frac{1}{\sqrt{2\pi(1-r^2)}} \exp\left(-\frac{(x-r\eta)^2}{2(1-r^2)}\right), \end{aligned}$$

注意到这是正态分布 $N(r\eta, 1-r^2)$ 的密度函数, 因此容易推出

$$E(\xi|\eta) = \int_{\mathbf{R}} x f_{\xi|\eta}(x|\eta)dx = r\eta.$$

利用这个例子, 我们再介绍计算条件期望的另一个常用方法. 因为 $E(\xi|\eta)$ 是 η 可测的, 所以存在函数 f 使得 $f(\eta) = E(\xi|\eta)$. 那需要求 f. 由定义, 对任何 $x \in \mathbf{R}$,

$E(f(\eta);\eta\leqslant x)=E(\xi;\eta\leqslant x)$, 用密度写出来, 有

$$\int_{-\infty}^{x} f(u)\frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}du=\int_{-\infty}^{x}dv\int_{\mathbf{R}}\frac{1}{2\pi\sqrt{1-r^2}}e^{-\frac{u^2-2ruv+v^2}{2(1-r^2)}}udu.$$

两边对 x 求导数得

$$\begin{aligned}f(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}&=\int_{\mathbf{R}}\frac{1}{2\pi\sqrt{1-r^2}}e^{-\frac{u^2-2rux+x^2}{2(1-r^2)}}udu\\&=\int_{\mathbf{R}}\frac{1}{2\pi\sqrt{1-r^2}}e^{-\frac{(u-rx)^2+x^2(1-r^2)}{2(1-r^2)}}udu\\&=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}rx,\end{aligned}$$

因此 $f(x)=rx$, 两种方法本质上也是一样的. ∎

习　题

1. 子事件域 $\mathcal{G}_1$ 与 $\mathcal{G}_2$ 被称为在给定随机变量 η 下条件独立, 如果对任何 $A_1\in\mathcal{G}_1$, $A_2\in\mathcal{G}_2$ 有
$$P(A_1\cap A_2|\eta]=P(A_1|\eta)P(A_2|\eta).$$
证明: 随机变量 ξ 与 $\mathcal{G}$ 在给定 η 下条件独立当且仅当对任何 $x\in\mathbf{R}$, 有 $E(\xi\leqslant x|\eta,\mathcal{G})=E(\xi\leqslant x|\eta)$.
2. 设 ξ,η 是独立同分布的可积的离散随机变量. 证明: $E(\xi|\xi+\eta)=\dfrac{1}{2}(\xi+\eta)$.
3. 设 (X,Y) 服从区域 D 上均匀分布, 其中 D 是点 $(0,0),(1,0),(1,1)$ 所围成的三角形. 求条件期望 $E(Y|X)$.
4. 设 X 服从参数为 n, U 的二项分布, 其中 U 服从 $[0,1]$ 上均匀分布. 求 X 的分布.
5. 设连续型可积随机变量 X 的密度函数为 f, 写出 $E(X||X|)$.
6. 设 X,Y 是 $[0,1]$ 上两个独立均匀分布随机变量. 求 $E(X|\max(X,Y))$.
7. 设 ξ,η 可积, 独立且期望为零, 证明: $E|\xi+\eta|\geqslant E|\xi|$.
8. 设 ξ,η 是两个随机变量, 证明: 下面三个论断等价

(a) ξ, η 独立;

(b) 对任何 $x \in \mathbf{R}$ 有

$$\mathsf{E}[e^{ix\xi}|\eta] = \mathsf{E}[e^{ix\xi}];$$

(c) 对任何有界连续函数 f 有 $\mathsf{E}[f(\xi)|\eta] = \mathsf{E}[f(\xi)]$.

第二章　随机游动

布朗运动的离散模拟是整数格点上的简单随机游动. 在本质中, 我们将介绍简单随机游动的一些结果和研究方法, 以让大家对布朗运动产生直觉.

随机过程是以离散时间或者连续时间为指标的随机变量集合, 它主要研究样本轨道的概率性质. 设有随机序列 $X_1, X_2, \cdots, X_n, \cdots$, 对于任意的样本点 $\omega \in \Omega$,

$$X_1(\omega), X_2(\omega), \cdots, X_n(\omega), \cdots$$

是一个通常的数列, 成为是样本 Ω 的轨道, 简称为样本轨道, 一个样本轨道是一个序列, 因此随机序列可以这样理解. 如果时间是连续的, 随机过程是一族随机变量 $(X_t : t \geqslant 0)$, 对于任意 $\omega \in \Omega$,

$$t \mapsto X_t(\omega) = X(t, \omega)$$

是一个函数, 称为样本轨道, 因此随机过程也可以理解为随机函数.

在研究随机过程时, 经常以样本轨道的角度思考问题, 离散时间随机过程可以作为数列来考虑, 连续时间随机过程可以作为函数考虑. 直观地, 也把一个样本点称为一条样本轨道.

2.1　掷硬币

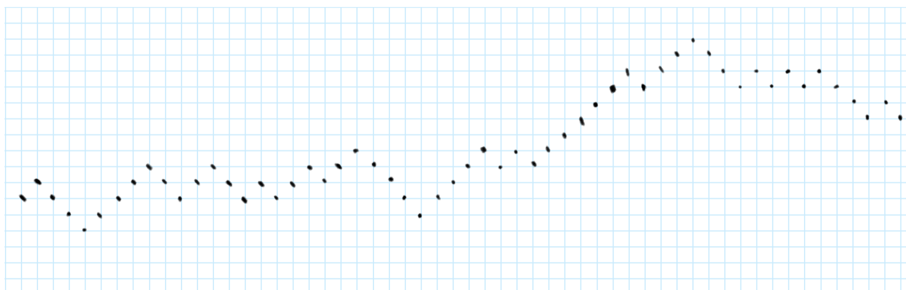
设 $\{\xi_n\}$ 是独立随机序列且

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

定义

$$S_0 = 0,\ S_n := \sum_{k=1}^{n} \xi_k,$$

那么 $\{S_n : n \geqslant 0\}$ 称为是直线上 (零点出发的简单对称) 随机游动. 随机游动是最早被研究的一个随机过程, 形象地被称为醉汉的脚印，可以用作很多随机现象的模型. 上面的随机游动实际上是从掷硬币这个简单的随机性实现的. 掷一个硬币, 这是最基本的随机现象, 有两个结果, 正面或者反面 (可以用不同的数字来表示, 例如 1 或者 0, 1 或者 -1), 概率一样, 都是 1/2. 这是通过实际掷 57 次硬币记录的一条样本轨道: $S_1(\omega), S_2(\omega), S_3(\omega), \cdots$



掷硬币是很平凡的, 但是重复不断地掷硬币就能得到不平凡的结果了. 数学上怎么描述重复不断地掷硬币呢? 用 Kolmogorov 概率公理的理论说, 存在一个概率空间 $(\Omega, \mathcal{F}, P)$ 和独立同分布随机序列 $\{X_n\}$, 其中 X_n 的分布是

$$P(X_n = H) = P(X_n = T) = \frac{1}{2},$$

注意这里 H 表示正面, T 表示反面. 有时候用数字表示更方便, 例如, 1 与 0 或者 1 与 -1. 称为对称的 Bernoulli 随机序列.

定理 2.1 **用 X_n 表示概率空间 $(\Omega, \mathcal{F}, P)$ 上随机变量 X 的二进制表示的第 n 位小数. 则 X 是 $[0,1]$ 上均匀分布当且仅当 $\{X_n : n \geqslant 1\}$ 是对称的 Bernoulli 随机序列.**

注意存在有二进位表示不唯一的数字, 但表示方法不唯一的数只有可列多个, 故 X 是二进位表示不唯一的概率是 0, 可以不必考虑. 那么 X_n 是随机变量, 取值 0 或 1, 计算 $(X_1, \cdots, X_n)$ 的联合分布, 事件 $\{X_1 = a_1, \cdots, X_n = a_n\}$ 等价于固定二进制的前 n 位小数, 这时 X 必定落在一个长度为 $1/2^n$ 的二分区间里, 因此

$$P(\{X_1 = a_1, \cdots, X_n = a_n\}) = \frac{1}{2^n},$$

故 $\{X_n\}$ 是独立随机变量序列, 且 $P(X_n = 1) = P(X_n = 0) = \dfrac{1}{2}$. 这个随机序列恰当地描述了重复地无限次掷硬币这个试验. 反之, 假设重复地无限次掷硬币这个试验

是可以实现的, 即存在概率空间 $(\Omega, \mathcal{F}, P)$ 和这样一个随机序列 $\{X_n\}$. 定义

$$X := \sum_{n=1}^{\infty} \frac{X_n}{2^n},$$

那么 X 是 $[0,1]$ 上均匀分布的随机变量. 为什么呢? 实际上, X 是一个二进制小数. $X \in [0,1/2]$ 当且仅当 $X_1 = 0$ (或者 $X_1 = 1, X_n = 0, n \geqslant 2$, 而这个概率是 0, 不必考虑). 因此 $P(X \in [0,1/2]) = 1/2$. 类似地

$$P(X \in [0, \frac{1}{2^n}]) = P(X_1 = 0, \cdots, X_n = 0) = \frac{1}{2^n},$$

一般地, 对任何 $n \geqslant 1$, $0 \leqslant k < 2^n$, $P(X \in [\frac{k}{2^n}, \frac{k+1}{2^n}]) = \frac{1}{2^n}$. 因此推出 X 是均匀分布的.

均匀分布的随机变量的存在性意味着什么呢? 设 X 是 $[0,1]$ 上均匀分布的. 首先 X 诱导出 $[0,1]$ 上的一个概率测度, 即对于 Borel 集 B, 定义 $\mu(B) = P(X \in B)$. 因为 P 是概率, 所以 μ 定义了 $[0,1]$ 上 Borel σ-代数上的一个测度, 称为是 X 的分布. 因为对 $x \in [0,1]$, 有 $P(X \leqslant x) = x$, 所以对任何区间 $I \subset [0,1]$, 有

$$P(X \leqslant I) = |I|.$$

因此 μ 是 Lebesgue 测度. 反过来, 如果 μ 是 $[0,1]$ 上 Lebesgue 测度, 那么

$$\Omega = [0,1], \mathcal{F} = \mathcal{B}[0,1], P = \mu$$

是个概率空间, 且其上的恒等函数 $X(x) = x$ 作为随机变量是均匀分布的. 因此, 均匀分布随机变量的存在性等价于 Lebesgue 测度的存在性.

Lebesgue 测度的存在性是数学史上极其不平凡的事件, 因此重复掷硬币的存在性也是一个极其不平凡的事件.

练习 2.1 给定分布函数 F, 在 $[0,1]$ 上存在可数个独立随机变量 $\{X_n, n \geqslant 1\}$ 使得它们的分布是 F.

2.1.1 大数定律

显然 S_n 是独立同分布 Bernoulli 随机变量的和, 所以它实际上是二项分布的, 不过因为 X_i 的取值是 -1 和 1, 所以 S_n 的取值是

$$-n, -n+2, -n+4, \cdots, n+2k, \cdots, -n+2n = n,$$

且 $S_n=-n+2k=-(n-k)+k$ 当且仅当 $X_1,\cdots,X_n$ 中有 k 个取 1 且 $n-k$ 个取 -1, 因此

$$P(S_n=n+2k)=\frac{C_n^k}{2^n}.$$

最大概率在中间达到, 当 n 是偶数时, 是 $P(S_n=0)$; 当 n 是奇数时是 $P(S_n=1)=P(S_n=-1)$. 当 n 趋于无穷时, 都趋于 0.

熟悉的是 $E[S_n]=0$, $D[S_n]=E[S_n^2]=n$. 由 Chebyshev 不等式得

$$P\left(\frac{|S_n|}{n}>\varepsilon\right)\leqslant\frac{E[S_n^2]}{n^2\varepsilon^2}=\frac{1}{n\varepsilon^2}.$$

推出 S_n/n 依概率收敛于 0, 即大数定律.

再应用 Chebyshev-Markov 不等式得

$$P\left(\frac{|S_n|}{n}>\varepsilon\right)\leqslant\frac{E[S_n^4]}{n^4\varepsilon^4}.$$

怎么计算 $E[S_n^4]$? 要计算它, 首先要能够展开

$$S_n^4=(X_1+X_2+\cdots+X_n)^4.$$

这是比二项展开 $(x_1+x_2)^k$ 的更一般的多项展开

$$(x_1+x_2+\cdots+x_n)^k,$$

在二项展开中, $x_1^ix_2^{k-i}$ 的系数是组合数 $\binom{k}{i}$, 而多项展开式中, 每一项 $x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$, 其中 $a_i\geqslant 0$ 且 $\sum_i a_i=k$, 它的系数是多项组合数

$$\binom{k}{a_1,a_2,\cdots,a_n}=\frac{k!}{a_1!a_2!\cdots a_n!}.$$

确切地,

$$(x_1+x_2+\cdots+x_n)^k=\sum_{a_i\geqslant 0,\sum_i a_i=k}\frac{k!}{a_1!a_2!\cdots a_n!}x_1^{a_1}x_2^{a_2}\cdots x_n^{a_n},$$

将此用于 $(X_1+\cdots+X_n)^4$, 如果 $a_1,\cdots a_n$ 有一个是奇数, 那么由独立性推出

$$E[X_1^{a_1}\cdots X_n^{a_n}]=0.$$

因此取期望之后, 只有那些 a_i 全部是偶数时, 期望

$$E[X_1^{a_1}\cdots X_n^{a_n}]=1.$$

那只有两种情况, 情况一是其中一个是 4 其余是 0, 这样共有 n 种不同选择, 情况二是其中两个都是 2 其余是 0, 这样共有 $\binom{n}{2}$ 种不同选择. 因此

$$E[S_n^4] = \frac{4!}{4!}n + \frac{4!}{2!2!}\binom{n}{2} = n + 3n(n-1).$$

$$P\left(\frac{|S_n|}{n} > \varepsilon\right) \leqslant \frac{E[S_n^4]}{n^4\varepsilon^4} \leqslant \frac{3}{n^2\varepsilon^4}.$$

由 Borel-Cantelli 引理, S_n/n 概率 1 收敛于 0, 这是强大数定律.

练习 2.2 证明 $E[S_n^{2k}]$ 是 n 的多项式并求其最高次数的系数.

利用这个结论可以证明: 对任何 $a > 1/2$, S_n/n^a 以概率 1 收敛于 0. 但是当 $a = 1/2$ 时不再以概率 1 收敛, 而是以分布收敛于标准正态分布

$$\lim_n P\left(\frac{S_n}{\sqrt{n}} \leqslant x\right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp(-y^2/2)dy.$$

这是 De Moivre 的中心极限定理.

2.2　随机游动的首中时

设 $X_1, \cdots, X_n, \cdots$ 是独立同分布随机变量服从 Bernoulli 分布

$$P(X_n = 1) = p,\ P(X_n = -1) = 1 - p = q,\ 0 < p < 1.$$

令 $S_0 = 0,\ S_n = \sum_{i=1}^n X_i$, 那么 $\{S_n : n \geqslant 0\}$ 被称为从原点出发的 1-维简单随机游动, 当 $p = 1/2$, 称为对称的, 否则称为非对称的. 随机游动的研究已经有几百年的历史, 几乎就是概率的历史, 围绕随机游动, 有很多有趣的经典问题, 例如上一节所述的大数定律, 但也仍然不断有各种新的问题, 当然, 发现问题的关键是看自己的想象力.

2.2.1　轨道遍历所有状态

在本章中, 我们关心一个简单的经典问题作为例子: 随机游动的常返性. 简单地说, 就是这样一个问题: 随机游动是不是以概率 1 能够回到任何它出发的状态? 对于现在的对称随机游动, 等价于问给定任何整数 a, 随机游动的几乎所有轨道都会碰到 a 吗? 虽然大数定律非常强大, 但是它不能给我们关于该问题的答案.

让我们用可以证明的数学语言来表达这个问题. 存在 $n > 0$ 使得 $S_n = a$ 就是事件 $\bigcup_{n>0}\{S_n = a\}$ 用数学语言来表达问题如下

$$P(\bigcup_{n>0}\{S_n = a\}) = P(\text{存在 } n > 0 \text{ 使得 } S_n = a) = 1?$$

这是第一种表述方式. 也可以定义轨道首次碰到 a 的时间

$$T_a = \inf\{n > 0 : S_n = a\},$$

称为首次通过时, 或者首中时. 显然, 当整个轨道都没碰到 a 时, $T_a = \infty$, 另外 $S_n = a$ 当且仅当 $T_a \leqslant n$, 因此

$$\bigcup_{n>0}\{S_n = a\} = \bigcup_{n>0}\{T_a \leqslant n\} = \bigcup_{n>0}\{T_a = n\} = \{T_a < \infty\},$$

即问题可表述如下

$$P(T_a < \infty) = 1?$$

这是第二种表述方式.

再因为随机游动 S_n 的步长是 1, 所以随机游动轨道碰到任何 a 当且仅当

$$\sup_n S_n = +\infty, \ \inf_n S_n = -\infty.$$

因此上面的问题也等价于

$$P(\sup_n S_n = +\infty, \ \inf_n S_n = -\infty) = 1?$$

这是第三种表述方式. 这样我们看到, 同一个问题可以有很多不同的表述方式. 本章中, 我们将应用各种不同的方法来思考并回答这个问题

首先考虑对称情况. 我们用初等概率论方法: 01-律和中心极限定理来证明

$$P(\sup_n S_n = +\infty, \ \inf_n S_n = -\infty) = 1.$$

事实上, 由于对称性, 只需要验证 $P(\sup_n S_n = \infty) = 1$ 就够了. 首先它属于 $\{\xi_n\}$ 的尾事件域, 因为 $\sup_n S_n$ 是否有限与 $\{\xi_n\}$ 的前有限个无关. 因此由 Kolmogorov 01 律, 它的概率非 0 即 1. 我们只需证明概率非零即可. 为此, 应用中心极限定理和 Fatou 引理, 得

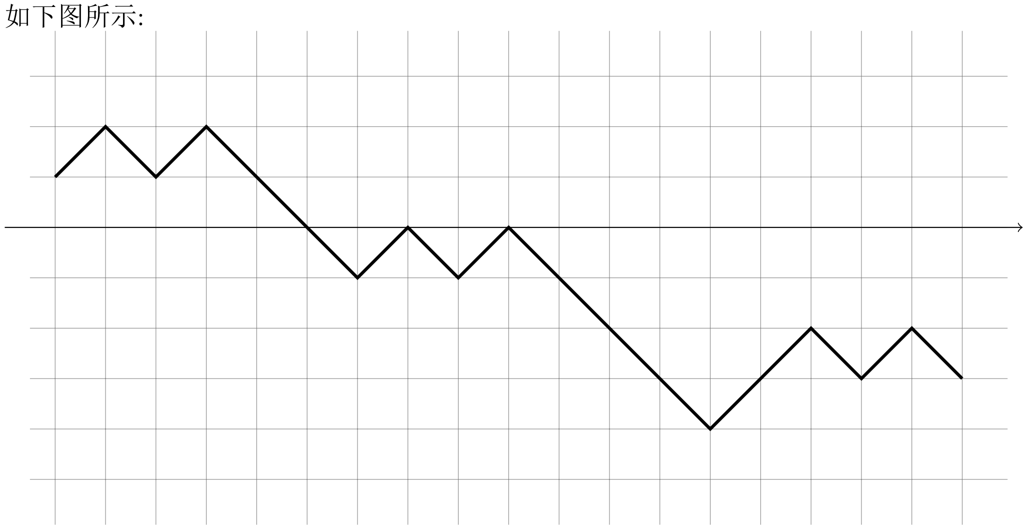
$$P(\sup_n S_n = \infty) \geqslant P(\limsup_n \{S_n > \sqrt{n}\})$$

$$
\begin{aligned}
&\geqslant \lim_n P(S_n > \sqrt{n}) \\
&= 1 - \Phi(1) > 0.
\end{aligned}
$$

2.2.2 反射原理

还是考虑对称情况. 第二个方法是反射原理, 其实也是对称性. 它把随机游动看成一个格点轨道组成的古典概率模型, 通过数数的办法来研究. 该方法局限性大, 缺乏推广性, 但想法奇妙, 值得介绍.

平面上的格点是指坐标都是整数的点, 把横轴看成时间轴, 纵轴看成状态轴. 一个格点轨道是指这样一条轨道, 它从任何一个格点出发, 每次向右上格点或者右下格点移动一格的轨道, 即在 (n, j) 处的点下一步只能走到 $(n+1, j+1)$ 或者 $(n+1, j-1)$. 如下图所示:



一个样本 $\omega \in \Omega$ 就是一次重复掷硬币的结果, 也就是一个取值为 -1 与 1 的序列: $x_n = X_n(\omega)$, $n \geqslant 1$. 由此诱导出随机游动的一个样本

$$
s_n = \sum_{i=1}^{n} x_i, n \geqslant 0,
$$

称为是随机游动的一个样本轨道. 随机游动的样本轨道很特殊, $s_{n+1} = s_n \pm 1$, 每次增加 1 或者减少 1. 这样的轨道组成一个格点轨道, 通常和时间 n 一起 (n, s_n) 画在平面上, 点之间用直线连接.

因为 $S_{n+1}=S_n+X_{n+1}$, 所以在每个节点 (n,S_n) 上, 下一刻向上移动和向下移动的机会是一样的. 因此从时间 1 到 n 随机游动 (时长 n) 可以走出 2^n 条不同的格点轨道, 每条轨道的机会都是 $1/2^n$, 等可能. 也就是说, 随机游动的问题可以通过数格点轨道个数的方法来解决.

设 $m<n$, 从点 (m,i) 到 (n,j) 有多少条格点轨道? 时间跨度为 $n-m$, 假设在这段时间掷了 r 个正面, s 个反面, 那么 (m,i) 可达 (n,j) 当且仅当存在非负整数 r,s 使得 $r+s=n-m$, 且 $r-s=j-i$, 即

$$\begin{cases} 2r=(n-m)+(j-i); \\ 2s=(n-m)-(j-i). \end{cases}$$

因此方程有解, 即 (m,i) 与 (n,j) 之间存在格点轨道连接的充分必要条件是 $n-m$ 与 $j-i$ 有相同的奇偶性且

$$|j-i|\leqslant n-m.$$

当上述条件满足时, (m,i) 到达 (n,j) 相当于掷 $n-m$ 次硬币得

$$\frac{(n-m)+(j-i)}{2}$$

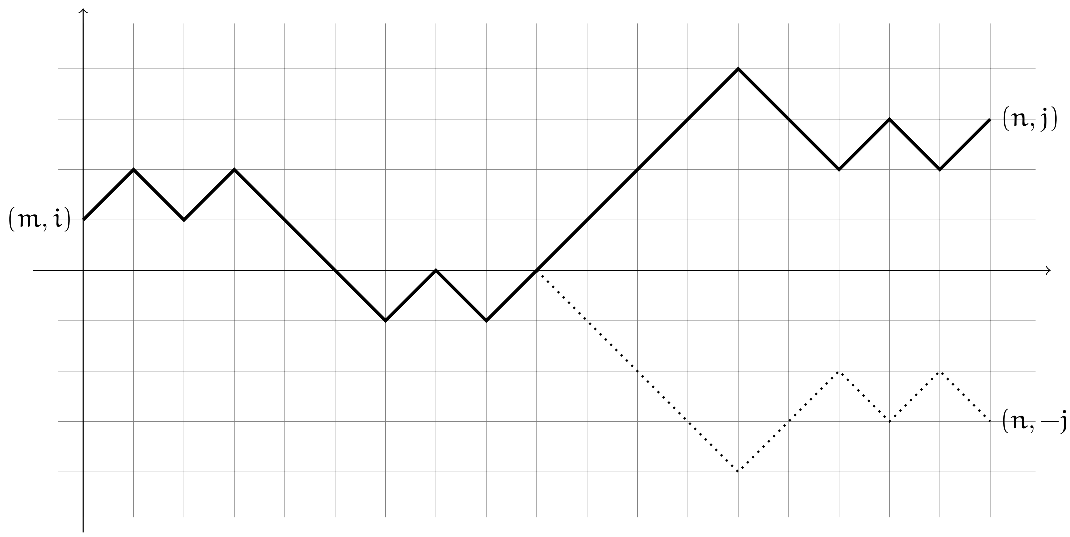
次正面, 用组合原理, 共有

$$\binom{n-m}{[(n-m)+(j-i)]/2}$$

种方法, 也就是从 (m,i) 到 (n,j) 的格点轨道数.

问题: 设 $i>0$, $j>0$, 问从 (m,i) 到 (n,j) 的轨道有多少条是碰到时间轴的? 即轨道上有一个点的纵坐标是零.

这个问题有一个非常漂亮的解决方法, 称为反射原理. 用 A 表示其中碰到时间轴的格点轨道全体, B 表示从 (m,i) 到 $(n,-j)$ (它是 (n,j) 关于时间轴的反射点) 的格点轨道全体. 建立 A 到 B 的一个变换: 任取 A 的一条轨道, 把其上最后一个纵坐标等于 0 的点之后的轨道关于时间轴反射下来, 之前保持不动, 得到新的轨道, 这条轨道的终点是 $(n,-j)$, 所以属于 B. 见下图



这个变换是 A 到 B 的一一对应 (请读者证明之), 即单 (不同的轨道的像不同) 且满 (B 中轨道必定是 A 中轨道反射来的). 因此 A 中的格点轨道数等于 B 中的格点轨道数, 即

$$|A| = |B| = \binom{n-m}{[(n-m)+(i+j)]/2}.$$

这个结论称为反射原理. 实际上时间轴可以换成任何横轴, 叙述如下.

定理 2.2 (反射原理) **设 $i, j < a$, $m < n$. 则从 (m,i) 到 (n,j) 途中 (位置) 到达过 a 的格点轨道数等于从 (m,i) 到 $(n, 2a-j)$ 的格点轨道数, 其中 $2a-j$ 是 j 关于 a 的对称点.**

例 2.1 (计票问题) 这可以解决经典的计票问题. 设有甲乙两个候选人, 在整个计票过程中, 获得 n 票的甲一直领先于获得 m 票的乙的概率是多少? 这里 $n > m$, 领先是指累积票数至少多一票. 一个计票过程可以看成一条格点轨道, 这个问题等价于计算从 $(0,0)$ 到 $(m+n, n-m)$ 的格点轨道中有多少不碰到时间轴 (除起点外) 的轨道, 等于 $(1,1)$ 到 $(m+n, n-m)$ 的格点轨道中有多少不碰到时间轴的轨道由反射原理, $(1,1)$ 到 $(m+n, n-m)$ 碰到时间轴的轨道数等于 $(1,-1)$ 到 $(m+n, n-m)$ 的轨道数, 即 $\binom{m+n-1}{n}$. 因此计票中甲一直领先的概率为

$$\frac{\binom{n+m-1}{n-1} - \binom{m+n-1}{n}}{\binom{m+n}{n}} = \frac{n-m}{n+m}.$$

类似的还有买票问题: 有 m 个人手持百元钞票一张, n 个人手持五十元钞票一张, 排队在一个没有任何零钱的卖电影票窗口买 50 元一张的电影票, 问在不能赊欠的条件下, 整个买票过程能够正常完成的概率是多少? ∎

2.2.3 首中时分布

继续对称情况. 为了表述到达过 a 这个事件, 我们定义首次通过时, 对整数 a, 定义 T_a 为随机游动首次碰到 a 的时间, 即

$$T_a = \inf\{n > 0 : S_n = a\},$$

称为 a 的首次通过时或者首中时, 约定空集的下确界为 $+\infty$. 因此 $T_a < \infty$ 当且仅当存在 n 使得 $S_n = a$, 即有限时间内到达 a. 另外 $T = T_0$ 是首次回归时.
用概率表示反射原理, 即对于 $a > 0$, $j < a$ 以及 $n \geqslant 1$ 有

$$P(T_a \leqslant n, S_n = j) = P(T_a \leqslant n, S_n = 2a - j) = P(S_n = 2a - j),$$

注意到 $S_n = 2a - j$ 蕴含有 $T_a \leqslant n$.
首次通过时 T_a 是一个随机时间, 第一个关心的问题是随机游动是否在有限时间内到达 a, 即

$$P(T_a < \infty) = 1?$$

让我们用反射原理来计算首次通过时 T_1 的分布.
显然 T_1 是个奇数, $T_1 = 2n + 1$ 意味着时间 $2n + 1$ 是随机游动首次到达 1 的时间, 因此我们要数从 $(0, 0)$ 出发到 $(2n + 1, 1)$ 不碰到状态 1 的轴的格点轨道数. 这相当于从 $(0, 0)$ 出发到 $(2n, 0)$ 不碰到状态 1 的轴的格点轨道数. 由反射原理, 这等于从 $(0, 0)$ 出发到 $(2n, 2)$ 的格点轨道数, 即

$$\binom{2n}{n-1} = \frac{(2n)!}{(n-1)!(n+1)!}.$$

因此 T_1 的分布为

$$P(T_1 = 2n + 1) = \frac{(2n)!}{(n-1)!(n+1)!2^{2n+1}}.$$

这样, T_1 有限的概率

$$P(T_1 < \infty) = \sum_{n \geqslant 0} \frac{(2n)!}{(n-1)!(n+1)!2^{2n+1}}.$$

仍然很难看出右边无穷和是不是 1, 需求助于 Taylor 展开.
换一个想法, 计算 $P(T_1 \leqslant n)$, 应用反射原理

$$\begin{aligned} P(T_1 \leqslant n) &= P(T_1 \leqslant n, S_n > 1) + P(T_1 \leqslant n, S_n = 1) + P(T_1 \leqslant n, S_n < 1) \\ &= 2P(S_n > 1) + P(S_n = 1) \\ &= 1 - P(S_n = 0) - P(S_n = 1), \end{aligned}$$

其中第二个等式

$$P(T_1 \leqslant n, S_n < 1) = P(S_n > 1)$$

是因为反射原理: 0 出发的随机游动在 n 时刻到达 1 之下且碰到 1 的轨道数与在 n 时刻到达 1 之上的轨道数一致, 即

$$\begin{aligned} P(T_1 \leqslant n, S_n < 1) &= \sum_{i<1} P(T_1 \leqslant n, S_n = i) \\ &= \sum_{i<1} P(T_1 \leqslant n, S_n = 2 - i) \\ &= \sum_{m=1}^{n} \sum_{j>1} P(T_1 = m, S_n = j) \\ &= P(T_1 \leqslant n, S_n > 1) = P(S_n > 1). \end{aligned}$$

因此

$$P(T_1 < \infty) = \lim_n P(T_1 \leqslant n) = 1.$$

这个结论说明, 随机游动以概率 1 可以到达它右边的邻居. 随机游动走到 1 再重新起步, 就可以看成一个从 1 出发的与原随机游动一样的随机游动, 所以它以概率 1 可以到达它右边的邻居, 也就是说原来的随机游动以概率 1 可以到达 2, 同样的理由说明它可以到达任意点 $a \neq 0$. 那么首次回归时 T 呢? 由全概率公式

$$P(T < \infty) = P(T < \infty | X_1 = 1)/2 + P(T < \infty | X_1 = -1)/2,$$

右边第一个实际上是从 1 出发的随机游动有限步到达 0, 第二个是从 -1 出发的随机游动有限步到达 0, 由对称性, 分别等价于从 0 出发有限步到达 -1 和 1, 概率都是 1, 所以 $P(T < \infty) = 1$.
再来看看具体的分布, T 应是偶数, T_1 是奇数, 故

$$P(T \leqslant 2n) = P(T_1 \leqslant 2n - 1)$$

$$= 1 - P(S_{2n-1} = 1) = 1 - P(S_{2n} = 0),$$

因此

$$P(T = 2n) = P(T \leqslant 2n) - P(T \leqslant 2n-2) = P(S_{2n-2} = 0) - P(S_{2n} = 0).$$

考虑非对称情形, 即 $p \neq 1/2$. 这时, 在 $S_{2n} = 0$ 的条件下, 随机游动从 0 到 $2n$ 的轨道是等可能的. 那么

$$P(T = 2n) = P(T = 2n | S_{2n} = 0) P(S_{2n} = 0),$$

其中条件概率 $P(T = 2n | S_{2n} = 0)$ 就是在 $2n$ 时刻回到 0 的条件下, $2n$ 是首次回到 0 的概率. 这实际上等于计票问题中, 获得 n 票的甲一直领先于获得 $n-1$ 票的乙的概率, 即 $\frac{1}{2n-1}$. 因此

$$P(T = 2n) = \frac{1}{2n-1} P(S_{2n} = 0).$$

最后

$$P(T < +\infty) = \sum_{n \geqslant 1} \frac{1}{2n-1} P(S_{2n} = 0).$$

怎么计算呢? 后面我们将看到这可以用幂级数展开来计算.
再看看回归时 T 的期望, 即平均回归时间. 因为 T 是偶数, 所以

$$E[T] = 2 \sum_{n \geqslant 0} P(T > 2n) = 2 \sum_{n \geqslant 0} P(S_{2n} = 0) = \infty.$$

最后的级数趋于无穷是由 Stirling 公式

$$n! \sim \sqrt{2\pi n}(n/e)^n$$

推出, 即

$$P(S_{2n} = 0) = \frac{(2n)!}{(n!)^2} 2^{-2n} \sim \frac{1}{\sqrt{\pi n}}.$$

同理, 对 $a > 0$,

$$P(T_a \leqslant n) = 2P(S_n > a) + P(S_n = a).$$

定义

$$D_n = \max(S_i : 1 \leqslant i \leqslant n), \ n \geqslant 0,$$

称为随机游动的极大游程. 显然 D_n 的轨道是递增的, 且对于任何整数 $a > 0$,

$$\{D_n \geqslant a\} = \{T_a \leqslant n\}.$$

因此

$$P(D_n \geqslant a) = P(T_a \leqslant n)$$

且 (S_n, D_n) 的联合分布

$$P(S_n = j, D_n \geqslant a) = P(S_n = j, T_a \leqslant n) = P(S_n = 2a - j),$$

其中 $a \geqslant j$.

2.3　随机游动的马氏性

这里要介绍的第三种方法是马氏性, 马氏性是随机过程中仅次于独立性的重要性质, 后面将详细地反复地进行讨论.

2.3.1　首中概率公式

直观地看, 如果随机游动现在 n 在 x 位置 $S_n = x$, 那么将来 (n 之后) 与过去 (n 之前) 是独立的, 确切地表示为

$$P(S_{n+1} = y | S_n = x, S_{n-1} = x_{n-1}, \cdots, S_1 = x_1) = P(S_{n+1} = y | S_n = x),$$

且右边的条件概率称为转移概率, 等于 $P(X_{n+1} = x)$, 与 n 无关, 称为时齐性. 这个性质是 150 年前由 Markov 自随机游动中抽象而来, 称为 Markov 性. 一个满足 Markov 性和时齐性的随机序列称为时齐 Markov 链. 随机游动自然是 Markov 链. 我们先来推导首中概率公式, 该公式对一般的 Markov 链成立, 因此我们不妨把 $(S_n : n \geqslant 0)$ 放在更一般的框架之下. 对于任何 $n \geqslant 0$, 设 S_n 的取值整数的随机变量, 满足上面的方程且当条件概率 $P(S_{n+1} = y | S_n = x)$ 与 n 无关, 称为转移概率, 记为 $p(x, y)$. 注意当我们谈论条件概率时, 总是在假设它有意义的情况下.

引理 2.1　**对任何** $j \geqslant 1$, **概率** $P(S_{n+j} = y | S_n = x)$ **与** n **无关.**

证明. 只需验证 $j=2$ 的情况.

$$
\begin{aligned}
P(S_{n+2}=y|S_n=x) &= \frac{P(S_{n+2}=y,S_n=x)}{P(S_n=x)} \\
&= \sum_z P(S_{n+2}=y|S_{n+1}=z,S_n=x)P(S_{n+1}=z|S_n=x) \\
&= \sum_z p(x,z)p(z,y),
\end{aligned}
$$

与 n 无关. □

记 $p^{(j)}(x,y)=P(S_{n+j}=y|S_n=x)$, 称为 j 步转移概率. 实际上, 可以用归纳法证明

$$
p^{(j+1)}(x,y)=\sum_z p^{(j)}(x,z)p(z,y)=\sum_z p(x,z)p^{(j)}(z,y).
$$

如果把 $p(x,y)$ 看成一个矩阵 P 的 x,y 处的元素, 那么 $p^{(j)}(x,y)$ 是矩阵 P^j 的 x,y 处的元素.

现在我们固定点 x, 设随机游动 (S_n) 从 x 点出发, T 是 x 的首中时. 令

$$
f_n=P(T=n),\ u_n=P(S_n=x).
$$

定理 2.3 对 $n\geqslant 1$, 有

$$
u_n=\sum_{k=1}^n f_k u_{n-k}.
$$

证明. 如果 $S_n=x$, 则它一定在 n 前包括 n 的某个时间 k 处首次回归 x, 然后从 k 时的 x 出发, 到 n 时再回到 x. 由此推出对 $n\geqslant 1$ 有

$$
\begin{aligned}
P(S_n=x) &= \sum_{k=1}^n P(S_n=x|T=k)P(T=k) \\
&= \sum_{k=1}^n P(T=k)P(S_n=x|S_k=x,S_j\not=x,\ 1\leqslant j<k) \\
&= \sum_{k=1}^n f_k P(S_n=x|S_k=x).
\end{aligned}
$$

这完成证明. □

利用首中概率公式推出, 当 $n \geqslant 1$ 时,

$$\begin{aligned} u_n &= \sum_{k=1}^{n} f_k u_{n-k} \\ &= f_1 u_{n-1} + f_2 u_{n-2} + \cdots + f_{n-1} u_1 + f_n. \end{aligned}$$

原则上, 当 $n = 1$ 时, $f_1 = u_1$. 当 $n = 2$ 时, 从 $u_2 = f_1 u_1 + f_2$ 解出 $f_2 = u_2 - (u_1)^2$, 然后原则上可以依次解出 f_3, f_4 以及所有的 f_n. 但实际上很难, 问题是即使写出 f_n 表达式, 还要算级数和, 也很难. 怎么办呢? 所幸我们的目标是计算级数和

$$\sum_{n=1}^{\infty} f_n,$$

而不是数列 f_n.

让我们把所有的方程列出来, u 的下标从小到大, f 的下标从大到小.

$$\begin{aligned} &u_1 = f_1; \\ &u_2 = f_2 + u_1 f_1; \\ &u_3 = f_3 + u_1 f_2 + u_2 f_1; \\ &u_4 = f_4 + u_1 f_3 + u_2 f_2 + u_3 f_1; \\ &u_5 = f_5 + u_1 f_4 + u_2 f_3 + u_3 f_2 + u_4 f_1; \\ &\cdots\cdots \end{aligned}$$

全部加起来, 右边每一列相加, 看出

$$\begin{aligned} \sum_{n=1}^{\infty} u_n &= \sum_{n=1}^{\infty} f_n + u_1 \sum_{n=1}^{\infty} f_n + u_2 \sum_{n=1}^{\infty} f_n + \cdots \\ &= (1 + u_1 + u_2 + \cdots) \sum_{n=1}^{\infty} f_n. \end{aligned}$$

由此推出, 如果级数 $\sum_n u_n < \infty$, 那么

$$P(T < \infty) = \sum_{n=1}^{\infty} f_n = \frac{\sum_{n \geqslant 1} u_n}{1 + \sum_{n \geqslant 1} u_n} < 1.$$

反之, 如果级数 $\sum_n u_n = \infty$, 那么上面的等式不会给我们有用的信息. 但我们有不等式

$$\sum_{j=1}^{n} u_j < \sum_{j=1}^{n} f_j \left(1 + \sum_{j=1}^{n} u_j \right),$$

它可以推出 $\sum_{j\geqslant 1} f_j = 1$. 这证明了下面定理.

定理 2.4　级数 $\sum_{n\geqslant 1} u_{2n} = \infty$ 当且仅当 $P(T < \infty) = 1$.

因此问题归结为判断级数 $\sum_{n\geqslant 1} u_n$ 是否收敛. 这个级数的直观意义是什么呢? 注意

$$\sum_{n\geqslant 1} u_n = \sum_{n\geqslant 1} P(S_n = x) = E\left(\sum_{n\geqslant 1} 1_{\{S_n = x\}}\right),$$

右边是 (S_n) 访问 x 点的次数的期望. 那么上面定理的概率含义是随机游动在有限时间内肯定返回 x 点当且仅当它访问 x 的平均次数是无穷. 其实还可以证明, 对任何 $n \geqslant 1$, 第 n 次访问 x 的时间 $T^{(n)}$ 也是以概率 1 有限的, 因此可以证明几乎所有轨道会访问 x 无穷多次, 留作习题.

练习 2.3　证明: 以概率 1 有

$$\sum_{n\geqslant 1} 1_{\{S_n = x\}} = +\infty.$$

回到零点出发的简单对称随机游动. 先计算 u_{2n}, 事件 $S_{2n} = 0$ 相当于说, 在 $2n$ 局游戏中, 硬币的正面次数与反面次数一样多, 所以

$$u_{2n} = P(S_{2n} = 0) = \frac{1}{2^{2n}}\binom{2n}{n}.$$

现在来看 $\sum_n u_{2n}$ 的收敛问题. 不难, 只是需要一个有关阶乘的估计和一点关于级数的基本常识, 称为 Stirling 公式

$$\lim_{n\to\infty} \frac{n!}{\sqrt{2\pi n}(ne^{-1})^n} = 1.$$

用它来估计 u_{2n},

$$u_{2n} = \frac{(2n)!}{(n!)^2}\frac{1}{2^{2n}} \sim \frac{\sqrt{2\pi 2n}(2ne^{-1})^{2n}}{(\sqrt{2\pi n}(ne^{-1})^n)^2} = \frac{1}{\sqrt{\pi n}}.$$

因此级数 $\sum_{n\geqslant 1} u_{2n} = +\infty$, 即推出

$$P(T < \infty) = 1.$$

练习 2.4　设 T_1 是首次抵达 1 的时间,

$$T = \inf\{n \geqslant 1 : S_n = 1\}.$$

证明: $P(T_1 < \infty) = 1$.

练习 2.5 将 X_n 的分布改为

$$P(X_n = 1) = q < 1/2,\ P(X_n = -1) = 1 - q,$$

问 $P(T < \infty)$ 是否等于 1?

T 是首次回归零的时间, 我们证明了

$$P(T < \infty) = 1,$$

即不管输赢, 在有限时间内, 总的输赢数一定回归于零. 我们想问的是, 平均回归时间会有多长? 也就是计算首次回归时间 T 的期望.

显然

$$E[T] = \sum_{n \geqslant 1} 2nP(T = 2n) = \sum_{n \geqslant 1} 2nf_{2n}.$$

在 §3.4.3 中我们有方程

$$u_{2n} = \sum_{k=1}^{n} f_{2k}u_{2n-2k}.$$

为了算期望, 我们需要新的想法. 在方程的两边乘 x^{2n}, 即

$$\begin{aligned}
u_2x^2 &= f_2x^2;\\
u_4x^4 &= f_4x^4 + u_2x^2f_2x^2;\\
u_6x^6 &= f_6x^6 + u_2x^2f_4x^4 + u_4x^4f_2x^2;\\
u_8x^8 &= f_8x^8 + u_2x^2f_6x^6 + u_4x^4f_4x^4 + u_6x^6f_2x^2;\\
u_{10}x^{10} &= f_{10}x^{10} + u_2x^2f_8x^8 + u_4x^4f_6x^6 + u_6x^6f_4x^4 + u_8x^8f_2x^2;\\
&\cdots\cdots
\end{aligned}$$

然后相加推出

$$\sum_{n \geqslant 1} u_{2n}x^{2n} = (1 + u_2x^2 + u_4x^4 + u_6x^6 + u_8x^8 + \cdots) \sum_{n \geqslant 1} f_{2n}x^{2n},$$

因此

$$F(x) = \frac{U(x)}{1 + U(x)},$$

其中, 我们记

$$F(x) = \sum_{n \geqslant 1} f_{2n}x^{2n},\ U(x) = \sum_{n \geqslant 1} u_{2n}x^{2n}.$$

现在的关键能不能把 $U(x)$ 表达式写出来. 如果能的话,

$$E[T]=\sum_{n\geqslant 1}2nf_{2n}=F'(1).$$

因为 $u_{2n}=\binom{2n}{n}2^{-2n}$, 所以

$$U(x)=\sum_{n\geqslant 1}\binom{2n}{n}2^{-2n}x^{2n}.$$

右边是 Taylor 展开, 因此我们要猜测什么样的函数会有这样的 Taylor 展开, 是个反问题.

考虑函数 $y=(1-x)^a$, 它的 Taylor 展开是

$$\begin{aligned}(1-x)^a&=1+a(-1)x+\frac{a(a-1)}{2}(-1)^2x^2+\cdots\\&\quad+\frac{a(a-1)\cdots(a-n+1)}{n!}(-1)^nx^n+\cdots.\end{aligned}$$

因此

$$(1-x^2)^a=1+\sum_{n\geqslant 1}\frac{(-a)(1-a)\cdots(n-a-1)}{n!}x^{2n},$$

现在比较两个级数的 x^{2n} 的系数来决定 a. 因为

$$u_{2n}=\frac{(2n)!}{n!n!}2^{-2n}=\frac{(2n-1)(2n-3)\cdots 3\cdot 1}{n!2^n}=\frac{(n-\frac{1}{2})\cdots\frac{3}{2}\cdot\frac{1}{2}}{n!},$$

所以容易看出当 $a=-1/2$ 时, 两者系数相同. 这时

$$1+U(x)=\frac{1}{\sqrt{1-x^2}},$$

推出 $F(x)$ 的表达式如下

$$F(x)=1-\sqrt{1-x^2},$$

它是以 $(0,1)$ 点为中心的单位圆的右下 $1/4$ 段. 从而

$$P(T<\infty)=\sum_{n\geqslant 1}f_{2n}=F(1)=1,\ \ E[T]=\sum_{n\geqslant 1}2nf_{2n}=F'(1)=+\infty.$$

也就是说在这样的赌博中, 输赢肯定会反转, 但等待这么一次反转的平均时间是无穷.

这个方法称为母函数方法, 其中函数

$$f(x) := \sum_{n \geqslant 0} a_n x^n,$$

称为是数列 $\{a_n : n \geqslant 0\}$ 的母函数. 母函数方法是处理数列问题的重要工具.

2.3.2 Polyá 定理

对称随机游动可以很直观地推广到任何图上. 设 $G = (V, E)$ 是一个无限连通但是每个顶点具有有限邻居的图, $y \sim x$ 是指 x, y 之间有一条边连接, $d(x)$ 表示 x 的邻居的个数. 一个在 G 上随机游动的蚂蚁是这样移动的: 从任何点 x 出发, 它等可能地挑选一个邻居跳过去, 然后等下一个时刻再等可能地挑选一个邻居跳过去. 这样的一个运动的存在性将在后面证明, 但它直观上是完全可以想象的. 一个图的结构被它上面的随机游动所刻画, 研究随机游动即等价于研究图结构.

考虑 d-维整数格点 $\mathbb{Z}^d$, 它可以自然地看成一个图, 点 x 的邻居是指与它距离等于 1 的点, 因此每个点 x 的度是 $2d$. $\mathbb{Z}^d$ 上的简单对称随机游动具体表述如下. 设有概率空间 $(\Omega, \mathcal{F}, P)$ 及其上独立同分布随机序列

$$X_1, X_2, \ldots, X_n, \ldots,$$

分布为

$$P(X_n = y) = \begin{cases} \frac{1}{2d}, & y \sim 0, \\ 0, & \text{其它}, \end{cases}$$

其中 $y \sim 0$ 表示 y 是 0 的邻居, 即与 0 的距离等于 1 的点. 再令

$$S_0 = 0,\ S_n = \sum_{k=1}^{n} X_k,\ n \geqslant 1.$$

定理 2.5 (Polyá, 1921)　**设 T 是 0 点的首次回归时, 则**

$$P(T < \infty) \begin{cases} = 1, & d \leqslant 2, \\ < 1, & d \geqslant 3. \end{cases}$$

练习 2.6　其中的关键是证明下面的结果: 令 $u_{2n}^{(d)} = P(S_{2n} = 0)$. 则

$$u_{2n}^{(d)} \sim Cn^{-d/2}.$$

提示: 考虑 S_{2n} 的 Fourier 变换, 再利用逆变换公式, 当 $d=3$ 时有

$$u_{2n}^{(3)} = \frac{1}{(2\pi)^3}\int_{[-\pi,\pi]^3}\left(\frac{\cos x_1+\cos x_2+\cos x_3}{3}\right)^6 dx_1dx_2dx_3.$$

2.4 随机游动的鞅方法

鞅是在二十世纪初引入的, 但五十年代 Doob 的工作给了鞅真正的生命力, 现在鞅是随机分析中的中心内容和不可缺少的工具. 在这一章中, 我们将简单介绍鞅及其在现代金融理论中的应用.

为了讨论鞅, 我们需要对条件期望等概念有更仔细的理解. 回忆可积随机变量 ξ 对子 σ-域 $\mathcal{A}$ 的条件期望 $E(\xi|\mathcal{A})$ 就是满足对任何 $A\in\mathcal{A}$, $E(\xi;A)=E(\eta;A)$ 的 $\mathcal{A}$ 可测的随机变量 η. 对随机变量集合 $\{\eta_i : i\in I\}$, 用 $\sigma(\eta_i : i\in I)$ 表示 $\{\eta_i^{-1}(B) : i\in I, B\in\mathcal{B}\}$ 生成的 σ-域, 注意它也是使得 $\{\eta_i : i\in I\}$ 都可测的最小 σ-域. 记 $E(\xi|\eta_i : i\in I) := E(\xi|\sigma(\eta_i : i\in I))$.

在随机过程理论中, 鞅是一个陌生的词, 直观地讲, 鞅是公平游戏的代名词. 拿随机游动为例, 设 $\{\xi_n\}$ 是独立同分布随机序列, 分布为 $P(\xi_n=1)=p$, $P(\xi_n=-1)=1-p$. 可以理解为简单的博弈游戏. 令 $X_n := \sum_{i=1}^n \xi_i$, 为 n 局游戏中总的输赢数. 显然当 $p=1/2$ 时, 游戏是公平的. 这时

$$E(X_n|X_{n-1},\cdots,X_1)=E(\xi_n|X_{n-1},\cdots,X_1)+X_{n-1}=X_{n-1},$$

也就是说, 如果我们以已知的前 $n-1$ 局游戏的结果来预测下一局输赢, 期望是 0.

设 $(\Omega,\mathcal{F},P)$ 是一个概率空间, 时间集 I 是整数集合的一个子集, 不妨设它是整数的一个区间, 其实这不是很重要, 如果不是区间, 我们不妨把 $n+1$ 理解为集 I 中整数 n 的下一个整数. 设 $\{X_n : n\in I\}$ 是随机序列. 为了方便, 对任何 $n\in I$, 定义

$$\mathcal{F}_n := \sigma(\{X_i : i\in I, i\leqslant n\}),$$

是时间 n 前的随机变量所生成的 σ-域, 当然 $\mathcal{F}_n$ 关于 n 是递增的, 集合 $\{\mathcal{F}_n : n\in I\}$ 被称为是随机序列决定的自然流.

定义 2.1　可积的实值随机序列 $\{X_n : n\in I\}$ 称为鞅序列, 如果对任何 $i,j\in I$ 且 $i<j$, 有

$$E(X_j|\mathcal{F}_i)=X_i.$$

另外 $\geqslant$ 号成立时, 称为下鞅, $\leqslant$ 号成立时, 称为上鞅.

鞅的定义等价于对任何 $n \in I$, $E(X_{n+1}|\mathcal{F}_n) = X_n$, 或者 $E(X_{n+1} - X_n|\mathcal{F}_n) = 0$. 一个鞅的期望与 n 无关, 因为 $EX_n = E(E(X_{n+1}|\mathcal{F}_n)) = EX_{n+1}$. 而下鞅的期望是 n 的递增序列.

注释 2.1 马氏性的定义与鞅的定义看上去类似, 但实际上完全不同. 马氏性是指现在位置给定时, 将来位置的与过去独立, 是一种独立性的判断. 而鞅是指无论过去 (包括现在) 怎么样, 未来对于现在的增量之期望等于零, 是一种价值的判断.

练习 2.7 证明:

1. 鞅全体是线性空间.
2. 鞅加下鞅还是下鞅.
3. 如果 (X_n) 是下鞅, 则 (X_n^+) 也是下鞅.

例 2.2 独立随机序列诱导的鞅: 设 $I = \mathbf{N}$, $\{\xi_n\}$ 是可积的独立随机序列, 且 $E\xi_n = 0$, 那么 $\{\xi_1 + \cdots + \xi_n\}_{n\geqslant 1}$ 是鞅序列. 如果 $\xi_n \geqslant 0$, 有界且 $E\xi_n = 1$, 那么 $\{\xi_1 \cdot \xi_2 \cdots \xi_n\}_{n\geqslant 1}$ 是鞅序列. ∎

例 2.3 (Wald 鞅) 设 $\{\xi_n\}$ 是上一节中定义的简单随机游动, 那么其中的 σ-域列 $\{\mathcal{B}_n\}$ 是递增的, 对 $\lambda > 0$, 令

$$Y_n := \lambda^{\xi_n},\ n \geqslant 0.$$

因 ξ_n 是 $\mathcal{B}_n$ 可测的, 故 Y_n 也是.

$$\begin{aligned} E^x(Y_{n+1}|\mathcal{B}_n) &= E^x(\lambda^{\xi_n + X_{n+1}}|\mathcal{B}_n) = \lambda^{\xi_n} \cdot E^x(\lambda^{X_{n+1}}|\mathcal{B}_n) \\ &= Y_n \cdot (\lambda p + \frac{q}{\lambda}). \end{aligned}$$

因此 $\{\lambda^{\xi_n}(\lambda p + \frac{q}{\lambda})^{-n}\}$ 是鞅序列, 称为 Wald 鞅. 当 $p \neq q$ 时, 取 $\lambda = \frac{q}{p}$, $\lambda p + \frac{q}{\lambda} = 1$, 因此 $\{\left(\frac{q}{p}\right)^{\xi_n}\}$ 是鞅序列. ∎

例 2.4 (Doob 鞅) 设 ξ 是 $(\Omega, \mathcal{F}, P)$ 上可积随机变量, $\{\mathcal{F}_n : n \in I\}$ 是 $\mathcal{F}$ 的一个关于 n 递增的子 σ-域的集合, 令

$$\xi_n := E(\xi|\mathcal{F}_n),$$

那么 $\{\xi_n\}$ 是一个关于流 $(\mathcal{F}_n)$ 的鞅. ∎

简单地设 I 是离散时间集 $\{0,1,2,\cdots\}$, 但只要重新编号, 结果显然对连续整数的集合也是成立的. 设 $(\Omega,\mathcal{F},P)$ 是概率空间, $(\mathcal{F}_n : n\in I)$ 是流. 一个随机序列 $\{H_n : n\geqslant 0\}$ 称为是可预料的, 如果 H_0 是 $\mathcal{F}_0$ 可测的且对任何 $n\geqslant 1$, H_n 是 $\mathcal{F}_{n-1}$ 可测的. 设 X 是适应过程, H_n 是可预料过程, 定义

$$\begin{aligned}&(H\bullet X)_0 := H_0X_0,\\&(H\bullet X)_n := (H\bullet X)_{n-1} + H_n(X_n - X_{n-1}),\ n\geqslant 1.\end{aligned}$$

称为是过程 H 关于 X 的随机积分, 它是一般随机积分的离散形式. 注意区别乘积与随机积分符号上的区别.

定理 2.6 设 X **是一个适应过程**, H **是可预料有界过程. 如果** X **是鞅, 那么过程** $H\bullet X$ **是鞅. 如果** X **是下鞅且** H **非负, 那么** $H\bullet X$ **是下鞅.**

证明. 显然 $(H\bullet X)_n$ 是可积的并且 $\mathcal{F}_n$ 可测的, 且对 $n\geqslant 1$,

$$\begin{aligned}E[(H\bullet X)_n - (H\bullet X)_{n-1}|\mathcal{F}_{n-1}] &= E(H_n(X_n - X_{n-1})|\mathcal{F}_{n-1})\\&= H_nE(X_n - X_{n-1}|\mathcal{F}_{n-1}),\end{aligned}$$

如果 X 是鞅, 则右边是零, 即 $H\bullet X$ 是鞅; 如果 X 是下鞅且 H 非负, 则右边也非负, 即 $H\bullet X$ 是下鞅. □

上面的定理是随机分析中非常本质的结果, 千百年来, 赌徒们总是想在赌桌上发现对自己有利的策略, 经验说明这是徒劳的. 以上定理从理论上满意地解释了这个经验. 设 X_n 是第 n 次赌博后某赌徒 A 的所有赌资, 则 $X_n - X_{n-1}$ 是 A 第 n 次赌博中输赢的数目, 另一个赌徒 B 赌 A 的运气, H_n 是乘子, 也就是 B 的策略. 但 B 也不可能预知下一局 A 的输赢, H_n 只能根据 $X_0, X_1, \cdots, X_{n-1}$ 的结果决定, 即 H_n 是 $\mathcal{F}_{n-1}$ 可测的, (从这个意义解释, 可预料也许应称为不可预料.) 定理指出 B 的运气不可能比 A 更好, 也不可能更坏.
在某个时间停止赌博是一种简单策略. 让我们引入停时的概念, 它是概率论中最重要的概念之一.

定义 2.2 **一个值域为** I **的随机变量** T **称为是 (相对于流** $(\mathcal{F}_n : n\in I)$ **的) 停时, 如果对任何** $n\in I$, $\{T=n\}\in\mathcal{F}_n$.

典型的停时是首中时, 如果 A 是 Borel 集, 定义 T 是序列 $\{X_n : n \in I\}$ 首次遇到 A 的时间, 即 $T := \inf\{n \in I : X_n \in A\}$, 那么 T 是停时, 理由是

$$\{T = n\} = \{X_n \in A\} \cap \{X_{n-1} \notin A\} \cap \cdots \cap \{X_0 \notin A\} \in \mathcal{F}_n.$$

因为 $\mathcal{F}_n$ 关于 n 递增, 故 T 是停时等价于对任何 $n \in I$, $\{T \leqslant n\} \in \mathcal{F}_n$. 如果 T 与 S 是停时, 那么 $\{T \wedge S \leqslant n\} = \{T \leqslant n\} \cup \{S \leqslant n\}$, 故有下面引理.

引理 2.2　如果 T 与 S 是停时, 那么 $T \wedge S$ 与 $T \vee S$ 也是停时.

对于随机序列 $\{X_n : n \in I\}$, 自然地在集合 $\{T = n\}$ 上定义 $X_T := X_n$, $n \in I$, 被称为 X 在停时 T 处的位置. 定义 T-停止序列

$$X_n^T(\omega) := X_{n \wedge T}(\omega),\ n \geqslant 0,$$

实际上, 它可以写为

$$\begin{aligned} X_n^T &= \sum_{k=0}^{n-1} X_k 1_{\{T=k\}} + X_n 1_{\{T \geqslant n\}} \\ &= \sum_{k=0}^{n-1} X_k (1_{\{T \geqslant k\}} - 1_{\{T \geqslant k+1\}}) + X_n 1_{\{T \geqslant n\}} \\ &= X_0 + \sum_{k=1}^{n} 1_{\{T \geqslant k\}} (X_k - X_{k-1}). \end{aligned}$$

而 $\{T \geqslant n\} = \{T < n\}^c \in \mathcal{F}_{n-1}$, 应用定理 2.6 得到有界停止定理.

定理 2.7 (Doob)　1. **设 X 是下鞅, S, T 是停时且 $S \leqslant T$, 则 $\{X_n^T - X_n^S : n \in I\}$ 是下鞅. 因此 $EX_n^S \leqslant EX_n^T$.**

2. **如果 $\{X_n : n \in I\}$ 是鞅, T 是停时, 那么鞅的停止序列 $\{X_n^T : n \in I\}$ 也是鞅. 进一步, 如果 T 是有界的, 那么 $EX_T = EX_0$.** (2)

证明. 只需证明 1. 由上面 X_n^T 的表达式, 结合条件 $S \leqslant T$ 得

$$X_n^T - X_n^S = \sum_{k=1}^{n} 1_{\{T \geqslant k\} \setminus \{S \geqslant k\}} (X_k - X_{k-1}).$$

由定理 2.6推出结论. □

定理中 (1) 的第二个结论通常称为 Doob 有界停止定理, 是非常有用的, 但问题是停时一般都不会是有界的, 所以研究什么情况下 $EX_T = EX_0$ 成立是非常有意义的问题. 首先举个例子说明, 结论一般是不对的.

例 2.5 设 $\{X_n : n \geqslant 0\}$ 是直线上 0 出发的的简单对称随机游动, 它是鞅. 定义 T 是点 1 的首中时, 那么 $X_T = 1$, 所以 $EX_T = 1 \neq 0 = EX_0$. ▮

但下面的定理说明在随机游动的情况下, T 的可积性能保证等式成立.

定理 2.8 (Wald) **设 $\{\xi_n : n \geqslant 1\}$ 是可积独立同分布随机序列且 $E\xi_1 = 0$, T 是可积停时, 则 $E\sum_{n=1}^{T} \xi_n = 0$.**

证明. 定义 $X_n := \sum_{i=1}^{n} \xi_i$, 那么 $\{X_n : n \geqslant 1\}$ 是鞅. 由 Doob 停止定理, 对任何 n, $EX_{T\wedge n} = 0$. 因此如果 T 有界, 定理结论成立.
再用 Doob 停止定理,

$$E\left(\sum_{i=1}^{T\wedge n} |\xi_i|\right) = E(T \wedge n) \cdot E|\xi_1|,$$

由单调收敛定理得

$$E|X_T| \leqslant E\sum_{i=1}^{T} |\xi_i| = \lim_{n\to\infty} E(T \wedge n) \cdot E|\xi_1| = ET \cdot E|\xi_1| < \infty.$$

因此 $\sum_{n=1}^{T} |\xi_n|$ 与 X_T 可积. 最后, $E(X_{T\wedge n}) = E(X_T; T \leqslant n) + E(X_n; T > n)$. 首先由控制收敛定理算右边第一项的极限, $\lim_n E(X_T; T \leqslant n) = EX_T$. 另一方面, $E(|X_n|; T > n) \leqslant E(\sum_{i=1}^{T} |\xi_i|; T > n)$, 因为 $\sum_{i=1}^{T} |\xi_i|$ 可积, 故再用控制收敛定理推出 $E(|X_n|; T > n) \longrightarrow 0$. 因此推出 $EX_T = 0$. □

下面两个例子显示鞅在解决经典问题时的应用. 回到 0 出发的随机游动

$$S_n = \xi_1 + \cdots + \xi_n,$$

其中 (ξ_n) 独立同分布 $P(\xi_n = 1) = p$, $P(\xi_n = -1) = q = 1 - p$.

例 2.6 **输光问题与持续时间:** 任取 $a, b \in \mathbf{Z}$, $a < 0 < b$, 令 $T := T_a \wedge T_b$, 即首次通过 a 或 b 其一的时间, 或 $\{a, b\}$ 的首中时, 也称为持续时间. 已知对任何 $P(T < \infty) = 1$. 显然 $\{T < \infty\} = \{T_a < T_b\} \cup \{T_a > T_b\}$, 概率

$$r = P(T_a < T_b)$$

是从 0 出发的随机游动, 到达点 a 在到达点 b 之前的概率. 形象地, 这相当于 A, B 两人各带 $-a, b$ 枚硬币参加一个简单随机游动形式的赌博游戏, 游戏至其中某人输光所有的硬币时结束, 因此 r 通常也称为输光概率, T 是游戏持续时间. 理论上, 这称为是一个具吸收壁 $\{a, b\}$ 的简单随机游动.

如果 $p = q = \frac{1}{2}$, 那么 $\{S_n\}$ 是一个鞅, $\{S_{n\wedge T}\}$ 也是鞅, 故

$$E[S_{n\wedge T}] = E[S_0] = 0,$$

由控制收敛定理得

$$E[S_T] = 0.$$

而

$$E[S_T] = E(S_T; T_a < T_b) + E(S_T; T_a > T_b) = aP(T_a < T_b) + bP(T_a > T_b),$$

因此

$$r = P(T_a < T_b) = \frac{b}{b-a}.$$

因为持续时间 T 是有限的, 怎么算持续时间 T 的期望呢? 这时我们需要另外一个鞅, 当 $p = 1/2$ 时, 容易验证

$$X_n = S_n^2 - n$$

是鞅. 这里有个问题, 这些鞅是怎么找到的? 实际上是从指数鞅那里找到的.

任意取 $x > 0$,

$$E[x^{S_{n+1}}|\mathcal{G}_n] = x^{S_n}E[x^{\xi_{n+1}}] = x^{S_n}(xp + x^{-1}q),$$

序列 (x^{S_n}) 不是鞅, 但可以看出对任何 $x > 0$,

$$Y_n := x^{S_n}(xp + x^{-1}q)^{-n}$$

组成的随机序列是一个鞅, 称为指数鞅.

现在说说怎么从指数鞅得到其它的鞅. 设 $p = 1/2$, 把指数鞅在 $x = 1$ 点 Taylor 展开

$$x^{S_n}(x/2 + x^{-1}/2)^{-n} = \sum_{k\geqslant 0} X_n^{(k)}\frac{(x-1)^k}{k!},$$

那么对于任何 k, 序列

$$(X_n^{(k)} : n \geqslant 1)$$

是鞅, 而且 $X_n^{(0)} = 1$, $X_n^{(1)} = S_n$.

练习 2.8 求 $X_n^{(2)}$ 与 $X_n^{(3)}$.

同样, $E[S_{n\wedge T}^2]=E[n\wedge T]$. 让 n 趋于无穷, 得

$$E[T]=E[S_T^2]=a^2\frac{b}{b-a}-b^2\frac{a}{b-a}=-ab.$$

再来算 T 的母函数, 这需要用指数鞅

$$Y_n=x^{S_n}(x/2+x^{-1}/2)^{-n},\ n\geqslant 0.$$

因为 $x/2+x^{-1}/2\geqslant 1$, 所以 Y_n 是有界鞅, 类似方法推出

$$E\left[x^{S_T}(x/2+x^{-1}/2)^{-T}\right]=1.$$

然后, 左边等于

$$E\left[x^a(x/2+x^{-1}/2)^{-T},T_a<T_b\right]+E\left[x^b(x/2+x^{-1}/2)^{-T},T_b<T_a\right].$$

令 $z=(x/2+x^{-1}/2)^{-1}$, 那么 $z\in(0,1]$ 且下面两个数都满足该方程

$$x_1=z^{-1}+\sqrt{z^{-2}-1}\geqslant 1,\ x_2=z^{-1}-\sqrt{z^{-2}-1}\leqslant 1.$$

这样我们得到两个等式

$$\begin{aligned}x_1^aE\left[z^T,T_a<T_b\right]+x_1^b(E\left[z^T\right]-E\left[z^T,T_a<T_b\right])=1\\ x_2^aE\left[z^T,T_a<T_b\right]+x_2^b(E\left[z^T\right]-E\left[z^T,T_a<T_b\right])=1\end{aligned}$$

这样得到

$$E[z^T]=\frac{x_2^a-x_2^b-x_1^a+x_1^b}{x_1^bx_2^a-x_2^bx_1^a}.$$

一个问题, T 是否是指数可积的? 即是否存在 $z>1$ 使得 $E[z^T]<\infty$? ∎

例 2.7 怎么计算 0 出发随机游动 (S_n) 的首中时 T_a 的母函数? 显然, $(S_n-n(p-q):n\geqslant 0)$ 是鞅, 但它不能解决我们的问题, 为什么不能解决呢? 当然需要你自己试试. 什么鞅能解决问题呢? 还是需要用指数鞅

$$Y_n:=x^{S_n}(xp+x^{-1}q)^{-n}.$$

因为 (Y_n) 是鞅, 所以 $(Y_{T_a\wedge n})$ 也是鞅. 因此有下面关键恒等式

$$E\left[x^{S_{T_a\wedge n}}(xp+x^{-1}q)^{-T_a\wedge n}\right]=E[Y_0]=1.$$

现在让 n 趋于无穷, 这时 $T_a \wedge n$ 趋于 T_a. 显然当 $T_a < \infty$ 时, $S_{T_a} = a$ 而当 $T_a = \infty$ 时, X_τ 是没有定义的. 什么时候极限与期望可以交换呢?
因为 $T_a \wedge n \leqslant T_a$, 所以 $S_{T_a \wedge n} \leqslant a$. 因此当 $x \geqslant 1$ 时,

$$x^{S_{T_a \wedge n}} \leqslant x^a.$$

当 $xp + x^{-1}q \geqslant 1$ 时

$$(xp + x^{-1}q)^{-T_a \wedge n} \leqslant 1.$$

当两个条件都满足时, $\{Y_n\}$ 被常数控制. 什么时候两个条件都满足呢? 分两种情况:

(1) $p \geqslant 1/2$. 这时 $x > 1$ 即保证两个条件满足;

(2) $p < 1/2$. 这时 $x > q/p$ 才能保证两个条件满足.

把上面的关键恒等式的左边分成两部分:

$$E\left[x^{S_{T_a \wedge n}}(xp + x^{-1}q)^{-T_a \wedge n} 1_{\{T_a < \infty\}}\right] \text{ 和 } E\left[x^{S_{T_a \wedge n}}(xp + x^{-1}q)^{-T_a \wedge n} 1_{\{T_a = \infty\}}\right].$$

在第一部分上, $\lim_n S_{T_a \wedge n} = a$, 在第二部分上

$$\lim_n (xp + x^{-1}q)^{-T_a \wedge n} = 0.$$

因此让 n 趋于无穷且应用控制收敛定理, 推出

$$x^a E\left[(xp + x^{-1}q)^{-T_a}\right] = 1.$$

现在我们来看两种不同情况, 在情况 (1) 时, 让 $x \downarrow\downarrow 1$, 那么

$$P(\tau < \infty) = 1.$$

在说明几乎所有轨道都会到达 a 点; 在情况 (2) 时, 让 $x \downarrow\downarrow q/p$, 那么

$$P(\tau < \infty) = (p/q)^a < 1.$$

这说明只有一部分轨道会达到 a 点, 而且 p 越小或者 a 越远, 概率越小. 这结论和直观符合.
在 $p \geqslant 1/2$ 时, $P(T_a < \infty) = 1$, 我们还可以写出 T_a 的母函数

$$z \mapsto E[z^{T_a}], \ z \in (0, 1).$$

令 $xp + x^{-1}q = z$, 注意 $x > 1$ 所以 $z > 1$, 解出

$$x = \frac{z + \sqrt{z^2 - 4pq}}{2p},$$

所以

$$E[z^{-T_a}] = \left(\frac{z + \sqrt{z^2 - 4pq}}{2p}\right)^{-a} = \left(\frac{z - \sqrt{z^2 - 4pq}}{2q}\right)^{a}.$$

写出随机变量的母函数和写出它的分布律在本质上是一样的. 比如说我们可以用母函数来算 T_a 的期望, 设 $a = 1$, 两边对 z 求导, 然后让 $z \downarrow 1$, 得

$$E[T_a] = -\frac{1}{2q}\left(1 - \frac{1}{\sqrt{1-4pq}}\right) = \frac{1}{2p-1},$$

当 $p > 1/2$ 时有限, 当 $p = 1/2$ 时无穷.
再来算首次回到出发点的首次回归时 T_0 有限的概率 $P(T_0 < \infty)$. 由全概率公式

$$P(T_0 < \infty) = pP(T_0 < \infty | X_1 = 1) + qP(\tau_0 < \infty | X_1 = -1) = 2(p \wedge q).$$

只有对称时, $P(T_0 < \infty) = 1$, 否则小于 1. ∎

习　题

1. 证明: $P(S_1 \neq 0, \ldots, S_{2n} \neq 0) = P(S_{2n} = 0)$.

2. 设 $L_{2n} := \sup\{i \leqslant 2n : S_i = 0\}$. 证明: $P(L_{2n} = 2k) = u_{2k}u_{2n-2k}$, $0 \leqslant k \leqslant n$.

3. 计算从 $(0,0)$ 到 $(2n, 0)$ 的恰有 $2k$ 条边在 x-轴上的格点轨道数, 证明: $n+1$ 整除 $\binom{2n}{n}$.

4. 设 $\{\xi_n\}$ 是简单随机游动, 如果 T 是停时, $T < \infty$, 定义 $\xi'_n = \xi_{n+T} - \xi_T$, $n \geqslant 0$, 证明: ξ' 是一个从 0 出发的简单随机游动. 形象地说, 随机游动从停时 T 重新开始.

5. 一个粒子在 $\{-N, -N+1, \cdots, N-1, N\}$ 上作简单随机游动 $(n \geqslant 1)$, 向右一步的概率为 p, 向左一步的概率为 $q = 1 - p$, $-N$ 与 N 是吸收状态. 设一个粒子从 0 点出发, 计算它在返回 0 前被吸收的概率.

6. A, B 两人玩一个骰子赌博, 两人共有 L 元钱, 如果骰子是偶数, 则互不给钱, 如果骰子是 1, 则 A 给 B 一元钱, 如果骰子是 3 或 5, 则 B 给 A 一元钱. 游戏一直到一人输光结束. 用 p_k 表示在开始时 A 有 k 元钱而最后赢的概率, 写出关于 $\{p_k\}$ 的差分方程.

7. 一个随机游动向前两格的概率是 p 向后一格的概率是 $q = 1 - p$, 对 $a > 0$, 用 $\pi(a)$ 表示这个随机游动有限步到达 0 的概率, 证明:
$$\pi(a) = p\pi(a+2) + q\pi(a-1),$$
当 $p \leqslant \frac{1}{3}$ 时, $\pi(a) = 1$.

8. 设 $\{X_n\}$ 是独立平方可积且期望等于零的随机序列, 验证: 随机序列
$$\left\{(\sum_{j=1}^{n} X_j)^2 - \sum_{j=1}^{n} E[X_j^2] : n \geqslant 1\right\}$$
是鞅.

9. 证明: 如果 X 是非负上鞅, 那么对任何固定 N,
$$P(\{X_N > 0, \inf_{n \leqslant N} X_n = 0\}) = 0.$$

10. (Wald 鞅) 设 $\{Y_n : n \geqslant 1\}$ 是独立同分布随机序列使得 $\phi(t) := E[e^{tY_n}]$ 对某个 $t \neq 0$ 有限. 令
$$X_n := \phi(t)^{-n} \exp[t(Y_1 + \cdots + Y_n)].$$
证明: $\{X_n : n \geqslant 1\}$ 是鞅.

11. 一个袋子中在时刻 0 有一个红球与一个白球. 随机地从袋子中取一个球, 然后将它放回并放入一个相同颜色的球, 无限地重复此过程. 记 X_n 为 n 次后袋中白球数与总球数之比. 证明: $\{X_n\}$ 是鞅.

12. 设 $\{Y_n : n \geqslant 1\}$ 是独立同分布随机序列, f_0, f_1 是两个概率密度函数, $f_0 > 0$. 令
$$X_n := \frac{f_1(Y_1)f_1(Y_2)\cdots f_1(Y_n)}{f_0(Y_1)f_0(Y_2)\cdots f_0(Y_n)}.$$
证明: 如果 f_0 是 Y_n 的密度函数, 那么 (X_n) 是鞅.

13. 设 $\{X_n : n \geqslant 0\}$ 是 $(\mathcal{F}_n)$ 适应的可积随机序列, 满足

$$E(X_{n+1}|\mathcal{F}_n) = \alpha X_n + \beta X_{n-1},\ n \geqslant 1,$$

其中 $\alpha > 0, \beta > 0, \alpha+\beta = 1$. 问 a 为何值时, 序列 $Y_0 := X_0,\ Y_n := aX_n+X_{n-1}$ 是 $(\mathcal{F}_n)$ 鞅.

14. (Doob 分解) 证明: 下鞅 (X_n) 可唯一分解为 $X_n = Y_n + Z_n$, 其中 (Y_n) 是鞅, (Z_n) 是从 0 出发的非负可预料的增过程: $0 = Z_1 \leqslant Z_2 \leqslant \cdots$.

15. 考虑 2- 维非负整数格点上的一个随机游动, 如果过程现在位于 (m, n), 那么下一步各以概率 $\frac{1}{2}$ 走到 $(m, n+1)$ 与 $(m+1, n)$. 设过程从 $(0, 0)$ 出发. Γ 是一条从 x 轴的某格点联结相邻格点到 y 轴某格点的折线. 记 Y_1, Y_2 分别是过程碰到 Γ 时向右与向上移动的步数. 证明: $EY_1 = EY_2$.

16. 设 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是独立同分布可积随机变量. 对 $k \geqslant 1$, 令

$$X_{-k} := \frac{1}{k}(\xi_1 + \cdots + \xi_k),$$
$$\mathcal{F}_{-k} := \sigma(X_{-k}, X_{-k-1}, \cdots).$$

证明: $\{X_k : k \leqslant -1\}$ 关于 $(\mathcal{F}_k : k \leqslant -1)$ 是鞅.

第三章　布朗运动的存在性

布朗运动实际上是对植物学家 R.Brown 所发现现象的一种数学模拟. 他发现悬浮在液体表面的花粉颗粒会无序地向各个方向运动, 无法预测. 布朗运动是最重要的一类随机过程. 可以说, 它是所有随机过程重要性质的交集. 如它是马氏过程, 是平稳独立增量的, 是连续的, 是鞅, 等等. 这样一个神奇的随机过程在数学公理下是不是存在? 这个问题实际上在当年也是困扰了数学家很多年, 如果从 A. Einstein 发现布朗运动的转移函数算起, 一直到 1923 年 Wiener 才证明布朗运动存在. 现在我们将讨论布朗运动的存在性, 实际上就是构造布朗运动. 为了做到这一点, 需要很多的准备工作, 但我们不可能从零开始, 还是需要一个合理的起点, 我们假设读者已经理解 Lebesgue 测度的构造, 也就是说理解 Caratheodary 的测度构造定理.

在这里我们稍微介绍一下关于实现的概念. 在实际问题中, 我们常常对一个问题有直观的描述, 比如它应该怎么样, 应该满足什么条件等等, 但想要把它纳入数学的框架, 必须在数学上证明它的存在性, 就是所谓的实现. 在概率论中, 我们证明一个分布函数是某个随机变量的分布函数就是典型的实现. 相比于随机变量, 随机过程更加复杂, 因此更加需要注意它的可实现性.

3.1　预备知识

首先介绍测度构造理论, 为了方便, 我们重复概率空间的定义.

定义 3.1　**设 Ω 是非空集合, 一个子集类 $\mathcal{F}$ 称为是 Ω 上的一个 σ-域 (或 σ-代数, 事件域等), 如果**

(1) $\varnothing, \Omega \in \mathcal{F}$;

(2) $A \in \mathcal{F}$ **蕴含** $A^c \in \mathcal{F}$;

(3) $A_n \in \mathcal{F}$, $n \geqslant 1$ 蕴含 $\bigcup_n A_n \in \mathcal{F}$.

称 $(\Omega, \mathcal{F})$ 是一个可测空间. $\mathcal{F}$ 上的非负集函数 P 称为是测度, 如果

(1) $P(\varnothing) = 0$;

(2) 对任何互不相容的集列 $\{A_n : n \geqslant 1\} \subset \mathcal{F}$, 有

$$P(\bigcup_{n \geqslant 1} A_n) = \sum_{n \geqslant 1} P(A_n).$$

这时, $(\Omega, \mathcal{F}, P)$ 称为是一个测度空间. 进一步, 如果 $P(\Omega) = 1$, 那么我们说 P 是概率测度, 简称概率.

概率测度无非是一种特殊的测度而已, 前面很多对概率测度叙述并证明的结果并没有用到 $P(\Omega) = 1$, 因此它们对于一般的测度也是对的, 需要的时候我们将特别地指出.

古典概率所涉及的概率空间通常是简单的, 我们甚至可以把样本空间列举出来, 但一般的概率空间是相当复杂的, 要列举其 σ-域和概率测度几乎是不可能的, 这里导致复杂的主要原因是可列运算和测度的可列可加性的要求. 所以我们先引入一个简单的子集类.

定义 3.2 一个子集类 $\mathcal{A}$ 称为是 Ω 上的域, 如果它包含空集和 Ω, 对补集运算封闭且对有限并运算封闭. 定义在域 $\mathcal{A}$ 上的非负集函数 μ 称为是可加的, 如果它满足对任何 $\mathcal{A}$ 中互斥的 A, B 有 $\mu(A \cup B) = \mu(A) + \mu(B)$.

域与 σ-域的定义看上去差别并不大, 一个对有限并封闭而另一个对可列并封闭, 但实际上域的结构要比 σ-域的结构简单许多. 因此我们的思路是在域上构造一个我们所期望的概率测度然后扩张到域生成的 σ-域上. 这本质上就是实变函数课程中介绍的所谓 Caratheodory 的测度构造理论.

定义 3.3 域 $\mathcal{A}$ 上的非负可加集函数被称为预测度, 如果它还满足: 对任何 $\mathcal{A}$ 中互不相容且并在 $\mathcal{A}$ 中的集列有可列可加性, 即若 $\{A_n : n \geqslant 1\} \subset \mathcal{A}$ 不相容且 $\bigcup_n A_n \in \mathcal{A}$, 有

$$\mu(\bigcup_n A_n) = \sum_n \mu(A_n).$$

显然 σ-域上的预测度就是测度.

引理 3.1 一个有限的 (即 $\mu(\Omega) < \infty$) 非负可加集函数成为预测度当且仅当它在 $\varnothing$ 处上连续.

定理 3.1 (Caratheodory) 设 μ 是域 $\mathcal{A}$ 上的预测度, 那么 μ 有一个扩张.

3.2 随机过程与构造

3.2.1 随机过程

在很多实际问题中, 我们需要研究随机过程, 也就是依照时间序记录的一族随机变量.

定义 3.4 设 $(\Omega, \mathcal{F})$ 和 $(E, \mathcal{E})$ 是可测空间, 一个映射 $\xi : \Omega \longrightarrow E$ 称为可测映射, 如果 $\xi^{-1}(\mathcal{E}) \subset \mathcal{F}$. 如果需要, 那么我们指明是 $\mathcal{F}/\mathcal{E}$-可测映射.

随机变量诱导象测度, 它很简单, 作用类似分布函数. 如果 $(\Omega, \mathcal{F})$ 上给定了一个概率测度 P, 那么利用可测映射 ξ, 很容易在 $(E, \mathcal{E})$ 上定义一个集函数 μ_ξ 如下:

$$\mu_\xi(A) := P(\xi^{-1}(A)) = P(\{\omega \in \Omega : \xi(\omega) \in A\}), \ A \in \mathcal{E}.$$

利用逆象的性质, 容易验证 μ_ξ 是一个概率测度, 它被称为 P 在 ξ 下的像测度, 或者 ξ 的分布, 也记为 ξP 或者 $P \circ \xi^{-1}$. 形象地说, ξ 把 P 诱导到 $(E, \mathcal{E})$ 上. 在 $(E, \mathcal{E}, \mu_\xi)$ 上恒等映射的分布就是 μ_ξ. 不难看出, 如果 η 是 E 到另一个可测空间 F 的可测映射, 那么

$$P \circ (\eta \circ \xi)^{-1} = (P \circ \xi^{-1}) \circ \eta^{-1}.$$

特别地, 设 ξ 是随机变量, 那么 μ_ξ 是 $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上的概率测度, 是 ξ 的分布, 它与 ξ 的分布函数 F_ξ 一一对应:

$$\mu_\xi = dF_\xi.$$

从数学上看, 在样本空间上研究随机变量的分布, 与在 $\mathbf{R}$ 研究 μ_ξ 是一样的. 注意一个随机变量的取值与取值的概率是两回事, 后者被称为随机变量的分布性质, 例如掷一个硬币, 正面反面是不可知的, 但它的分布性质是知道的.

定义 3.5 随机过程是一个概率空间 $(\Omega, \mathcal{F}, P)$ 到一个可测空间 $(E, \mathcal{E})$ 的一族有序排列的可测映射, $\{\xi_t : t \in T\}$, 其中 T 是一个全序集, 通常是 $\mathbf{R}$ 的子集. 这时候, E 称为随机过程的状态空间 (或相空间), T 称为是随机过程的时间或指标集. 当 T 是 $\mathbf{Z}$

的子集时, 通常称为随机序列或者离散时间随机过程. 当 T 是区间时, 随机过程称为是连续时间随机过程.

状态空间 E 通常理解是 $\mathbf{R}$, 这时随机过程被称为实值随机过程. 当 E 是拓扑空间时, $\mathcal{E}$ 取为其 Borel σ-域. 前面讨论的随机游动和鞅都是随机序列, 都是在一个给定概率空间上满足一定条件的随机序列.

例 3.1 设 F 是 $\mathbf{R}$ 上分布函数, 那么存在概率空间 $(\Omega, \mathcal{F}, P)$ 和独立随机序列 $\{\xi_n\}$ 且它们的分布函数都是 F. 进一步地, 定义 $S_1 = \xi_1, \cdots, S_n := \sum_{i=1}^n \xi_i$, 那么 $\{S_n\}$ 也是随机序列, 它也称为 $\mathbf{R}$ 上的一个随机游动. 如果 F 的期望等于 0, 则 $\{S_n\}$ 是一个鞅列. ∎

在一些简单情况下, 类似于 §2.2 中那样可以证明独立随机序列的存在性, 但在许多情况下的随机过程不能这样简单地利用独立随机序列来构成, 让我们先考察下面的模型.

例 3.2 蚂蚁迷宫: 设 $G = (V, E)$ 是一个无向图, V, E 分别是顶点和边的集合, 假设它是局部有限的, 即每个顶点只有有限多个邻居. 现在有一个蚂蚁从某个顶点出发在 G 上游走, 规则是这样的: (1) 它总是沿着边走; (2) 在到达某个顶点后, 它总是等可能地选取一个邻居走过去. 是不是有一个概率空间可以把这个问题描述清楚? 比如说, 让 ξ_n 表示 n 时刻 (我们总是把蚂蚁达到某个顶点的时间称为时刻) 蚂蚁的位置, 我们需要这样一个概率空间 $(\Omega, \mathcal{F}, P)$ 来安放随机序列 $\{\xi_n\}$, 使得它要满足

$$P(\xi_{n+1} = y | \xi_n = x) = \begin{cases} \frac{1}{c(x)}; & \text{如果 } y \text{ 是 } x \text{ 的邻居;} \\ 0, & \text{否则,} \end{cases}$$

这里 $c(x)$ 表示顶点 x 的邻居个数. ∎

例 3.3 随机带位: 设一个餐厅有编号为 $1, 2, 3, \cdots$ 的无穷个桌子, 某个桌子可以坐无穷多个客人, 餐厅经理总是坐在没有客人的桌子中号码最小的桌子上, 当下个客人来的时候, 餐厅按这样的规则带位: (1) 客人总是被带到有人的桌子上; (2) 客人被带到某个桌子的概率和桌子上已有的人数成比例. 我们是否可以用随机过程来描述这个模型? 例如让 ξ_n 表示第 n 个客人所带到的桌子号码, 同样地, 是否有一个概率空

间 $(\Omega,\mathcal{F},P)$ 使得 $\xi_1=1$,

$$P(\xi_{n+1}=i|\xi_1,\cdots,\xi_n)=\begin{cases}\frac{|\{k:\xi_k=i\}|}{n+1}; & \text{若 } i\leqslant\max(\xi_1,\cdots,\xi_n);\\ \frac{1}{n+1}; & \text{若 } i=\max(\xi_1,\cdots,\xi_n)+1;\\ 0; & \text{其它}.\end{cases}$$

∎

3.2.2 什么是布朗运动

一个布朗运动应该具有什么性质？在这一节中我们来仔细地讨论这个问题. 直观上, 布朗运动是生物学家 Robert Brown 在 1827 年所描述的花粉在液体表面移动的这种现象. 后来, 科学家认为这是因为花粉受到几乎无穷多粒子从各个方向不间断地无规律的撞击而产生的具有高度随机性的移动, 且在自然界有普遍性的现象. 在这个意义上, 布朗运动是生物学家或者物理学家的观察或者直觉, 我们的问题是, 数学上应该怎么描述布朗运动呢？对任何 $t\geqslant 0$, 用 B_t 是表示 t 时刻花粉位置, 为了简单, 我们只考虑 B_t 是实值的, 或者说是位置的 x-坐标, 是一个随机变量. 布朗运动的随机性规律由概率空间 $(\Omega,\mathcal{F},P)$ 刻画, 就是指随机过程 $B=(B_t:t\geqslant 0)$. 对于样本 $\omega\in\Omega$, $t\mapsto B_t(\omega)$ 是样本轨道, 也就是花粉移动的轨迹. 实际上这是一个比较容易做的实验, 如果你愿意花一个下午的时间去观察, 去记录, 你会发现它的轨迹看上去杂乱无章, 像一团乱麻, 如图.

(1) 花粉的轨迹应该是连续的, 用数学的术语说, 几乎所有的样本轨道是连续的. 确切地说, 存在 $N\in\mathcal{F}$ 使得 $P(N)=0$ 且对任何 $\omega\notin N$, 轨道 $t\mapsto B_t(\omega)$ 是连续的.

(2) 对于 $t>s\geqslant 0$, B_t-B_s 是时间段 $[s,t]$ 产生的位移 (或者增量), 是个随机变量. 由于花粉的运动是因为各个方向许多小粒子对它不间断的撞击, 也就是小扰动的近乎无限的叠加, 所以我们直观认为这导致位移 B_t-B_s 是期望 0 的正态分布, 期望零是因为从各个方向的撞击应该假设是均匀的. Einstein 以及其他科学家从理论与实验两个方面证实其方差是 $\sigma(t-s)$, 其中 σ 与介质有关, 假设等于 1. 因此, 我们说 $B_t-B_s\sim N(0,t-s)$.

(3) 还有一个直观性质, 就是位移 B_t-B_s 应该不依赖于位置 B_s, 或者说不依赖于 $[0,s]$ 时间段花粉的移动轨迹. 这个性质称为独立增量性. 更确切地说, B 满足下面

所述的独立增量性质: 对任何 n 与任何的 $0 < t_1 < t_2 < \cdots < t_n$, 位移

$$B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$$

是独立的. 该性质与下面两个性质的任意一个是等价的,

1. 对任何 n 与任何的 $0 < t_1 < t_2 < \cdots < t_n$, 位移 $B_{t_n} - B_{t_{n-1}}$ 与事件域 $\sigma(B_{t_1}, B_{t_2}, \cdots, B_{t_{n-1}})$ 独立.
2. 对任何 $t > s \geqslant 0$, 位移 $B_t - B_s$ 与事件域 $\sigma(B_u : u \leqslant s)$ 独立. 记该事件域为 $\mathcal{F}_t^B$, 它包含了花粉运动的过去 (s 时刻之前) 的信息.

练习 3.1 证明: 上面三个性质等价.

(4) 最后一个是假设, B_0 是花粉的初始位置, 它可以是一个随机变量, 但这里我们假设花粉从某一个位置 x 出发, 即 $B_0 = x$.

所谓随机过程, 实际上是一个概率空间上的一族随机变量. 当我们想要满足某些分布条件的一个随机过程时, 就涉及到按照条件来构造一个随机过程. 上面的例子基本上清楚地说明了我们要做的事情. 本章的主要目的就是证明布朗运动的存在性, 或者证明满足上面 4 个性质的随机过程是存在的.

3.2.3 轨道空间与随机过程的分布

随机变量的概率性质由它的分布刻画, 那么随机过程是不是也有分布? 为了谈论随机过程的分布, 首先我们需要知道随机过程是什么. 设 $\xi = (\xi_t : t \in T)$ 是 $(\Omega, \mathcal{F}, P)$ 上的随机过程. 取 $\omega \in \Omega$, $\xi_t(\omega)$ 当 t 变化时是 T 到 E 的映射, 称为 ω 的样本轨道. 所以随机过程可以看成是把样本点变成样本轨道的映射. 因此我们需要考虑样本轨道的空间, 也就是 T 到 E 的映射全体所组成的空间, 记为 E^T. 这个空间是乘积空间的推广, E 的 T 次方空间, 或者说是一个无穷维乘积空间.

简单介绍乘积可测空间的概念, 设 $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$ 是可测空间, 对 $A_1 \subset E_1$, $A_2 \subset E_2$, 记

$$A_1 \times A_2 := \{(x_1, x_2) : x_1 \in A_1, x_2 \in A_2\}.$$

它是乘积空间 $E_1 \times E_2$ 上的矩形, 如果 $A_1 \in \mathcal{E}_1, A_2 \in \mathcal{E}_2$, 这个矩形被称为可测矩形. 显然可测矩形全体对交封闭, 是 π 类, 但不是 σ-域. 那么我们自然地取以可测矩形

生成的 σ-域, 记为 $\mathcal{E}_1\times\mathcal{E}_2$, 作为 $E_1\times E_2$ 上的自然的 σ-域. 空间 $(E_1\times E_2,\mathcal{E}_1\times\mathcal{E}_2)$ 被称为乘积 (可测) 空间. 定义映射

$$\pi_i(x_1,x_2)=x_i,\ i=1,2,$$

它们分别称为是乘积空间到 E_1,E_2 上的投影. 乘积 σ-域可以如下刻画.

引理 3.2 σ-域 $\mathcal{E}_1\times\mathcal{E}_2$ **是** $E_1\times E_2$ **上使得两个投影可测的最小** σ-域.

特别地, 我们考虑空间的自乘积, 首先容易定义可测空间 $(E,\mathcal{E})$ 的 n 次自乘积可测空间, 让我们记它为 $(E^n,\mathcal{E}^n)$. 空间的一个点通常写为 $(x_1,\cdots,x_n)$, 也就是坐标表示, 我们也可以把它看成是集合 $[n]=\{1,2,\cdots,n\}$ 到 E 的一个映射 $x:i\mapsto x_i$. 反过来, 一个这样的映射也可以写成乘积空间的一个点, 换句话说 E^n 和 $[n]$ 到 E 的映射全体是一一对应的. 所以我们不区分乘积空间的一个点和它所对应的一个映射, 而乘积 σ-域恰是使得所有投影可测的最小 σ-域.

两种不同角度看同一个问题的好处是后者可以很容易被推广到无穷的情况. 上面的集合 $[n]$ 并非本质的, 它可以被任何一个元素个数为 n 的集合替代, 因为它只不过是用来标识空间中点的坐标顺序的.

现在到了探讨乘积空间 E^T 的时候了. 注意 T 实际上可以是任意集合. 当 T 是无限集, E^T 被称为无穷乘积空间. 在本讲义中, 我们只关心 T 是离散时间集合与连续时间集合两种情况. 怎么来定义它上面的 σ-域呢? 还是用可测矩形的方法或者说投影的方法.

对映射 $x:T\to E$, $t\in T$, 符号 $x(t)$ 或 x_t 是 x 在时刻 t 处的值或坐标. 任意给定 T 的有限子集 I, (后面写 $I\Subset T$ 时, 指 I 是有限子集.) 定义投影 $\pi_I:E^T\longrightarrow E^I$,

$$\pi_I(x):=(x(t):t\in I),\ x\in E^T.$$

简单地, $\pi_t:=\pi_{\{t\}}$. 定义 $\mathcal{E}^T$ 为 E^T 上使得所有 $\{\pi_t:t\in T\}$ 可测的最小 σ-域. 这个定义类似于前面两个空间乘积的情形. 这时 $(\pi_t:t\in T)$ 就是可测空间 $(E^T,\mathcal{E}^T)$ 上的一族可测映射, 即随机过程, 被称为轨道过程, 或者典则过程.

上面的定义说

$$\mathcal{E}^T=\sigma\left(\bigcup_{t\in T}\pi_t^{-1}(\mathcal{E})\right),$$

右边括号内的子集类不是一个域 (即代数), 也不是 π-类. 为了应用单调类方法, 我们需要 $\mathcal{E}^T$ 的一个等价刻画, 使得它由一个域生成.

定义 3.6 设 $A \subset E^T$, 如果存在 $I \Subset T$ 与 $H \in \mathcal{E}^I$ 使得

$$A = \pi_I^{-1}(H) = \{x \in E^T : (x(t), t \in I) \in H\},$$

那么 A 被称为是个柱集, I 是 A 的一个标签.

柱集这个名称的来源是平面上的一个圆在三维空间对其投影之下的逆像是一个圆柱体. 柱集全体是

$$\mathcal{E}_0^T := \bigcup_{I \Subset T} \pi_I^{-1}(\mathcal{E}^I).$$

显然柱集的表示不唯一, 标签也不唯一. 称 I 是柱集 A 最小标签, 如果 I 不含真子集标签.

练习 3.2 证明: (1) 柱集的最小标签是唯一的. (2) 设 $A \supset B$ 是两个柱集, 最小标签依次为 I, J, 则 $I \subset J$.

引理 3.3 柱集的性质

1. 如果 $I \subset J \Subset T$,
$$\pi_I^{-1}(\mathcal{E}^I) \subset \pi_J^{-1}(\mathcal{E}^J);$$
2. 柱集全体 $\mathcal{E}_0^T$ 是一个域;
3. $\mathcal{E}^T = \sigma(\mathcal{E}_0^T)$.

证明. 若 $s \notin I$, 则

$$\pi_I^{-1}(H) = \pi_I^{-1}(H) \cap \{x(s) \in E\},$$

后者的标签为 $I' = I \cup \{s\}$. 因此 $\pi_I^{-1}(\mathcal{E}^I) \subset \pi_{I'}(\mathcal{E}^{I'})$, 故性质 1 是显然的. 性质 1 可以推出任何两个柱集 A, B 总可以被看成是同一个指标集 I 的柱集, 即

$$A, B \in \pi_I^{-1}(\mathcal{E}^I).$$

因此性质 2 是显然的. 性质 3 也是显然的. □

这样我们得到了一个无穷乘积可测空间 $(E^T, \mathcal{E}^T)$ 与其上的轨道过程 $\{\pi_t : t \in T\}$.
下面我们来说明随机过程可以看成为一个可测映射, 其像测度就是随机过程的分布.
现在设给定概率空间 $(\Omega, \mathcal{F}, P)$ 上以 $(E, \mathcal{E})$ 为状态空间的随机过程 $\xi = \{\xi_t : t \in T\}$.
对于固定的样本点 $\omega \in \Omega$, $t \mapsto \xi_t(\omega)$ 是 T 到 E 的函数, 即 E^T 中的元素, 把它记

为 $\xi(\omega)$, 它是 ω 的 (样本) 轨道. 那么随机过程 ξ 实际上是 Ω 到 E^T 的映射, 从样本 ω 到样本轨道的映射. 现在 ξ 既是随机过程的符号也是样本到轨道这个映射的符号. 对于 $I \Subset T$, 定义 $\xi_I = (\xi_t, t \in I)$, 这是一个随机向量, 是 Ω 到 E^I 的映射, 是可测的, 且

$$\xi_I = \pi_I \circ \xi.$$

引理 3.4 上面定义的 ξ 是 Ω 到轨道空间 E^T 的可测映射.

实际上, 为了验证可测性, 只需证明对任何 $A \in \mathcal{E}_0^T$, 有 $\xi^{-1}(A) \in \mathcal{F}$. 柱集 A 可以表示为 $A = \pi_I^{-1}(H)$, $H \in \mathcal{E}^I$, 其中 $I \Subset T$. 因此

$$\xi^{-1}(A) = \xi^{-1}\left(\pi_I^{-1}(H)\right) = \{\xi_I^{-1} \in H\} \in \mathcal{F}.$$

也就是说, 随机变量是个可测映射, 随机过程也是, 两者的像空间不同. 用 μ^ξ 表示 P 在 ξ 下的像概率测度, 它是 $(E^T, \mathcal{E}^T)$ 上的概率测度.

定义 3.7 空间 $(E^T, \mathcal{E}^T)$ 上如上定义的像测度 μ^ξ 被称为是随机过程 $\xi = (\xi_t)_{t\in T}$ 的分布.

应该注意的是, 尽管没有提到, 随机过程的分布 μ^ξ 总是随机过程在概率 P 之下的分布.

3.2.4 有限维分布族与相容性

给定概率空间 $(\Omega, \mathcal{F}, P)$ 上的随机过程 $\{\xi_t : t \in T\}$ 和 $I \Subset T$ (我们总是假设其中元素按顺序排列), 定义 μ^I 为 ξ_I 的分布

$$\mu^I(A) = P(\xi_I \in A),\ A \in \mathcal{E}^I,$$

它也是 μ^ξ 在 π_I 下的像测度:

$$\mu^I = P\circ\xi_I^{-1} = (P\circ\xi^{-1})\circ\pi_I^{-1} = \mu^\xi\circ\pi_I^{-1}.$$

测度 μ^I 是 $(E^I, \mathcal{E}^I)$ 上的概率测度, 是 P 在 E^I 上的投影, 它体现了过程的部分信息, 被称为 P 的一个有限维分布.

定义 3.8

1. **上面定义的测度集 $\{\mu^I :\ I \Subset T\}$ 称为随机过程 ξ 的有限维分布族.**
2. **两个相同状态空间和相同时间集的随机过程 $\xi = \{\xi_t, t \in T\}$ 和 $\eta = \{\eta_t, t \in T\}$ 称为等价, 如果 $\mu^\xi = \mu^\eta$, 即同分布.**

3. 两个相同概率空间, 状态空间和相同时间集的随机过程 $\xi = \{\xi_t\}$ 和 $\eta = \{\eta_t\}$ 称为互为修正 (版本), 如果对任何 $t \in T$, 有 $P(\xi_t = \eta_t) = 1$.

4. 两个相同概率空间, 状态空间和相同时间集的随机过程 $\xi = \{\xi_t\}$ 和 $\eta = \{\eta_t\}$ 称为不可区分, 如果 $P(\xi_t = \eta_t, \forall t \in T) = 1$.

定理 3.2 1. 随机过程的分布 μ^ξ 由 ξ 的有限维分布族所唯一决定.

2. 两个随机过程等价当且仅当它们有同样的有限维分布族.

3. 两个互为修正的随机过程是等价的.

4. 两个互为修正的轨道右连续的随机过程是不可区分的.

练习 3.3 证明该定理.

上面的 1 等价于说随机过程的分布 μ^ξ 是使得轨道空间 $(E^T, \mathcal{E}^T)$ 上轨道过程 $(\pi_t : t \in T)$ 与随机过程 $\xi = (\xi_t : t \in T)$ 等价的唯一测度. 有限维分布族就是随机过程的分布 μ^ξ 在柱集 $\mathcal{E}_0^T$ 上的表现. 形象地说, 随机过程的分布就是一幅太大的画, 无法在一张画布上表达出来, 有限维分布族就象用来拼接成一幅大画的一块块小画布一样. 相当于 Lebesgue 测度在区间上的体现, 所以也就是随机变量的分布函数的角色. 再说得确切一点, 尽管我们知道 Lebesgue 测度是可测集的长度, 但, 除了区间外, 我们无法对一个可测集给出其长度的表达式. 也就是说, 本质上, Lebesgue 测度用区间描述, 随机过程用有限维分布描述.

如同分布函数有递增右连续性一样, 有限维分布也有一个重要的性质: 相容性. 直观地说, 不同画布上的画须有某种相容性才能拼成一幅完整的画. 设 $I \subset J \Subset T$, 定义映射 $\pi_I^J : E^J \longrightarrow E^I$ 为

$$\pi_I^J(x_t : t \in J) := (x_t : t \in I).$$

它是空间 E^J 到 E^I 的投影. 显然 $\pi_I^J \circ \pi_J = \pi_I$, 因此

$$\mu^J \circ (\pi_I^J)^{-1} = \mu^\xi \circ \pi_J^{-1} \circ (\pi_I^J)^{-1} = \mu^\xi \circ \pi_I^{-1} = \mu^I.$$

即 μ^I 等于 μ^J 在 E^I 上的投影, 这说明如果 I 是 J 的子集, μ^I 是完全由 μ^J 决定的. 这直观上描述一个随机向量的联合分布决定其中任意部分随机向量的联合分布.

引理 3.5 随机过程的有限维分布族 $\{\mu^I : I \Subset T\}$ 满足: 对任何 $I \subset J \Subset T$, 有 $\mu^J \circ (\pi_I^J)^{-1} = \mu^I$.

这个性质是重要的, 它是随机过程有限维分布族必须满足的必要条件, 由此我们可以定义相容的有限维分布族. 注意下面的定义只与状态空间 E 和时间集 T 有关, 不涉及概率空间. 这个定义类似于概率论中把随机变量的分布函数满足的性质拿来作为分布函数的定义.

定义 3.9 给定时间集 T 和可测空间 $(E, \mathcal{E})$.

1. 如果对任何 $I \Subset T$, μ^I 是 $(E^I, \mathcal{E}^I)$ 上的概率, 那么称概率测度的集合 $\{\mu^I: I \Subset T\}$ 是 (空间 E 和时间 T 上的) 一个有限维分布族.

2. 如果一个有限维分布族 $\{\mu^I: I \Subset T\}$ 满足引理 3.5中的性质

$$\mu^J \circ (\pi_I^J)^{-1} = \mu^I, \ I \subset J \Subset T,$$

   那么它被称为是相容的有限维分布族.

3. 给定有限维分布族 $\{\mu^I: I \Subset T\}$. 如果存在概率空间 $(\Omega, \mathcal{F}, P)$ 上以 E 为状态空间的随机序列 $\{\xi_t : t \in T\}$ 使得它的有限维分布族就是给定的那个, 那么就说这个有限维分布族是可以实现的, 且上面的概率空间和随机过程是它的一个实现.

3.2.5 相容性定理

综上所述, 概率空间 $(\Omega, \mathcal{F}, P)$ 上以 E 为状态空间的随机过程 $\{\xi_t : t \in T\}$ 的有限维分布族 $\{\mu^I: I \Subset T\}$ 是相容的. 也就是说, 一个有限维分布族可以实现的必要条件是相容性, 那么反过来, 一个相容的有限维分布族 $\{\mu^I: I \Subset T\}$ 是否一定可以实现呢? 回答这个问题是构造布朗运动的第一步.

定理 3.3 (Kolmogorov) **如果 E 是完备可分度量空间, T 是事件集, 那么一个相容的有限维分布族 $\{\mu^I: I \Subset T\}$ 是可以实现的.**

这是现代概率论中的最主要的定理之一. 在证明这个定理之前, 我们回顾一下分布函数构造分布的过程. 设有 $\mathbf{R}$ 上分布函数 F. 定义区间 $(a, b]$ 的测度为

$$\mu((a, b]) = F(b) - F(a).$$

这样的区间全体是一个 π-类, 不是代数, 但我们可以把定义扩张到这样区间的有限并全体, 记为 $\mathcal{A}$, $\mathcal{A}$ 是一个代数. 然后关键是证明 μ 在 $\mathcal{A}$ 上是预测度, 这需要用 $\mathbf{R}$

的拓扑性质. 现在我们要做差不多同样的事情. 上面的准备工作已经差不多了, 但还需要一个重要的分析定理.

定理 3.4 (Ulam) **设** E **是完备可分度量空间, 则** E **上任何概率测度** μ **是正则的: 对任何** $B \in \mathcal{E}$, **有**

$$\mu(B) = \sup\{\mu(K) : K \subset B, K \text{ 紧}\}.$$

证明. 不失一般性, 我们只需对 $B = E$ 的时候证明就可以了. 由可分性, 对任何 n, 存在半径为 $1/n$ 的可列个球 $\{A_{n,k} : k \geqslant 1\}$ 覆盖 E, 那么

$$\lim_{i\to\infty} \mu\left(\bigcup_{i\geqslant k\geqslant 1} A_{n,k}\right) = 1,$$

故存在 i_n 使得 $\mu\left(\bigcup_{k=1}^{i_n} A_{n,k}\right) > 1 - \varepsilon/2^n$. 令 $A = \bigcap_{n\geqslant 1}\bigcup_{k=1}^{i_n} A_{n,k}$, 那么 A 是完全有界集且

$$\mu(A^c) = \mu\left(\bigcup_{n\geqslant 1}\left(\bigcup_{k=1}^{i_n} A_{n,k}\right)^c\right) \leqslant \sum_{n\geqslant 1}\mu\left(\left(\bigcup_{k=1}^{i_n} A_{n,k}\right)^c\right) < \varepsilon,$$

用 K 表示 A 的闭包, 那么 K 是紧集且 $\mu(K) \geqslant \mu(A) > 1 - \varepsilon$. □

证明. (定理 3.3的证明) 我们打算在轨道空间 $(E^T, \mathcal{E}^T)$ 上来实现有限维分布族, 即在其上构造出所需要的概率测度 μ, 也就是说, 它要满足对 $I \Subset T$, $\mu\circ\pi_I^{-1} = \mu^I$. 也就是说, 对任何 $H \in \mathcal{E}^I$,

$$\mu(\pi_I^{-1}(H)) = \mu^I(H).$$

这立刻告诉我们 μ 在柱集的域 $\mathcal{E}_0^T$ 上应该这样来定义. 取 $A \in \mathcal{E}_0^T$, 那么存在 $I \Subset T$ 和 $H \in \mathcal{E}^I$ 使得 $A = \pi_I^{-1}(H)$. 定义

$$\mu(A) := \mu^I(H).$$

但这定义看上去与柱集 A 的表示有关, 因此为了定义的合理性, 必须证明 $\mu(A)$ 不依赖于 A 的表示. 换句话说, 如果有另外的 $J \Subset T$ 和 $K \in \mathcal{E}^J$ 使得 $A = \pi_J^{-1}(K)$, 那么 $\mu^J(K) = \mu^I(H)$. 事实上, 因为 $\pi_I^{-1}(\mathcal{E}^I)$ 关于 I 是递增的, 故不妨设 $I \subset J$.

练习 3.4　请仔细写下不妨设 $I \subset J$ 的理由.

因为 $\pi_I = \pi_I^J \circ \pi_J$, 故

$$\pi_J^{-1}((\pi_I^J)^{-1}(H)) = \pi_I^{-1}(H) = \pi_J^{-1}(K),$$

且因投影一定是满射, 推出 $(\pi_I^J)^{-1}(H) = K$. 然后相容性保证

$$\mu^I(H) = \mu^J((\pi_I^J)^{-1}(H)) = \mu^J(K).$$

这说明 μ 是域 $\mathcal{E}_0^T$ 上良定义的非负集函数. 它满足

(1) $\mu(E^T) = 1$;

(2) (有限可加性) 如果 $A, B \in \mathcal{E}_0^T$ 且 $A \cap B = \varnothing$, 那么 $\mu(A \cup B) = \mu(A) + \mu(B)$.

练习 3.5 验证 μ 满足这两个性质.

现在由定理 3.1, 我们只需要证明 μ 在空集处有上连续性就够了, 即我们要证明, 如果 $\{A_n\} \subset \mathcal{E}_0^T$ 且 $A_n \downarrow \varnothing$, 则 $\mu(A_n) \downarrow 0$.

当然 $\mu(A_n)$ 是递减的, 反设存在 $\varepsilon > 0$ 使得 $\mu(A_n) > \varepsilon$. 我们来证明 $\bigcap_n A_n$ 非空. 首先存在 $I_n \Subset T$, $H_n \in \mathcal{E}^{I_n}$ 使得 $A_n = \pi_{I_n}^{-1}(H_n)$. 还是因为 $\pi_I^{-1}(\mathcal{E}^I)$ 关于 I 递增, 所以不妨认为集列 I_n 关于 n 递增. 现在 $A_n = \pi_{I_n}^{-1}(H_n)$ 其中 $H_n \in \mathcal{E}^{I_n}$. 那么 $\mu(A_n) = \mu^{I_n}(H_n)$. 因为完备可分度量空间的乘积仍然是完备可分度量空间, 故存在紧集 $K_n \subset H_n$ 使得 $\mu^{I_n}(H_n \setminus K_n) < \varepsilon/2^n$. 令 $B_n = \pi_{I_n}^{-1}(K_n)$, 它是包含在 A_n 中的柱集但未必关于 n 递减. 因此再记 $C_n := \bigcap_{i=1}^n B_i$, 它是含在 A_n 中的柱集且关于 n 递减. 现在

$$\mu(A_n \setminus C_n) = \mu\left(\bigcup_{i=1}^n (A_n \setminus B_i)\right) \leqslant \sum_{i=1}^n \mu(A_i \setminus B_i) = \sum_{i=1}^n \mu^{I_i}(H_i \setminus K_i) < \varepsilon.$$

这蕴涵着对任何 n, $\mu(C_n) > 0$, 推出 C_n 非空. 取点 $x^{(n)} \in C_n$ 组成 E^T 的点列. 因为 C_n 递减, 故对任何 $m \geqslant 1$, 当 $m \leqslant n$ 时,

$$C_n \subset C_m \subset B_m = \pi_{I_m}^{-1}(K_m) \subset A_m,$$

这说明 $\pi_{I_m}(x^{(n)}) \in K_m$, 即 $\{(x^{(n)}(t) : t \in I_m) : n \geqslant m\} \subset K_m$. 因 K_m 紧, 故它有在 K_m 中收敛的子列. 结论是对每个 m 有一个收敛子列, 但我们需要一个共同的收敛子列.

现在应用对角线方法, 我们可以证明存在一个子列 k_n 使得对任意 m, $\{(x^{(k_n)}(t):t\in I_m)\}$ 收敛, 记

$$(x(t):t\in I_m)=\lim_n (x^{(k_n)}(t):t\in I_m)\in K_m.$$

因此我们可以得到 E 的一个点列 $\{x(t):t\in\bigcup_m I_m\}$. 随意取点 $e\in E$, 对 $t\notin T\setminus\bigcup_n I_n$, 定义 $x(t)=e$. 那么 $x=(x(t):t\in T)\in E^T$, 而上面的结论说明对任何 m, $\pi_{I_m}(x)\in K_m$ 或者说 $x\in\pi_{I_m}^{-1}(K_m)\subset A_m$, 这与假设矛盾. □

显然, Euclid 空间是完备可分度量空间. 另外赋以离散拓扑的可列集也是完备可分度量空间, 但是离散拓扑的不可列集就不是可分的. 怎么具体来给定一个相容的有限维分布族呢? 首先在时间离散 $T=\mathbf{N}$ 的情况下, 事情实际上简单得多, 令 $\mathcal{E}_1^{\mathbf{N}}:=\bigcup_n\pi_{[n]}^{-1}(\mathcal{E}^{[n]})$. 显然 $\mathcal{E}_1^{\mathbf{N}}\subset\mathcal{E}_0^{\mathbf{N}}$ 并且 $\mathcal{E}^{\mathbf{N}}=\sigma(\mathcal{E}_1^{\mathbf{N}})$, 即柱集 $\mathcal{E}_1^{\mathbf{N}}$ 就足够生成 $\mathcal{E}^{\mathbf{N}}$ 了, 另外当有限维分布族 $\{(E^I,\mathcal{E}^I,\mu^I):I\subset\mathbf{N}\}$ 相容时, 它由它的子集 $\{(E^{[n]},\mathcal{E}^{[n]},\mu^{[n]}):n\in\mathbf{N}\}$ (简单地称它为有限维分布族的主子列) 完全决定了, 因为任何有限子集 I 总是某个 $[n]$ 的子集.

概率空间列 $\{(E^{[n]},\mathcal{E}^{[n]},\mu^{[n]}):n\in\mathbf{N}\}$ 称为相容的, 如果对任何 $n\in\mathbf{N}$,

$$\mu^{[n+1]}\circ(\pi_{[n]}^{[n+1]})^{-1}=\mu^{[n]}.$$

更具体地说, 对任何 $A_1,\cdots,A_n\in\mathcal{E}$ 有

$$\mu^{[n+1]}(A_1\times A_2\times\cdots\times A_n\times E)=\mu^{[n]}(A_1\times A_2\times\cdots\times A_n).$$

很显然, 一个相容的有限维分布族中的主子列一定是相容的, 反过来, 从一个相容的概率空间列可以唯一构造一个相容的有限维分布族. 事实上, 对任何有限子集, 取 n 充分大使得 $I\subset[n]$, 定义 $\mu^I:=\mu^{[n]}\circ(\pi_I^{[n]})^{-1}$, 由相容性推出定义与 n 的取法无关, 容易验证这样定义的 μ^I 构成一个相容的有限维分布族. 也就是说当时间集离散时, 相容的有限维分布族可以用更容易构造的相容概率空间列等价表示.

再假设 E 是离散可列集. 给定一个随机序列 $\{\xi_n:n\geqslant 0\}$, 由乘积公式推知

$$\begin{aligned}&P(\xi_0=x_0,\cdots,\xi_{n-1}=x_{n-1},\xi_n=x_n)\\&\quad=P(\xi_0=x_0)\prod_{i=1}^n P(\xi_i=x_i|\xi_0=x_0,\cdots,\xi_{i-1}=x_{i-1}),\end{aligned}$$

这说明联合分布是由 ξ_0 的分布和条件概率决定的. 直观地, 我们会问过程在起始时刻在哪里? 当知道前 n 个时刻过程的行为后, 过程下一刻又会怎么走? 习惯地, 一个随机序列的初始时刻通常记为 0, 所以我们下面假设 $\mathbf{N} = \{n : n \geqslant 0\}$. 反过来, 如果我们知道过程在 0 时刻以概率 $\mu_0(x)$ 位于 x 点 ($x \in E$); 如果已知过程的 n 时刻前的行为 $\xi_i = x_i$, $0 \leqslant i < n$, 那么它在下一步以概率 $p_n((x_0, \cdots, x_{n-1}), x)$ 移动到 x 点 ($x \in E$), 那么这样应该足可以描述一个随机序列了. 让我们写下并证明这个结论. 下面说的分布律总是指一个和等于 1 的非负数列.

推论 3.1 设 E 是离散可列集, $\{\mu_0(x) : x \in E\}$ 是 E 上分布律, 对任何的 $n \geqslant 1$ 及 $\{x_i : 0 \leqslant i < n\} \subset E$, $\{p_n((x_0, \cdots, x_{n-1}), x) : x \in E\}$ 是 E 上分布律. 那么存在概率空间 $(\Omega, \mathcal{F}, P)$ 及其上随机序列 $\{\xi_n : n \in \mathbf{N}\}$ 使得 $P(\xi_0 = x) = \mu_0(x)$ 且 (当下面条件概率有意义时)

$$P(\xi_n = x | \xi_0 = x_0, \cdots, \xi_{n-1} = x_{n-1}) = p_n((x_0, \cdots, x_{n-1}), x),$$

它们分布称为过程的初始分布和转移概率.

证明. 定义 $\mu^{[0]} = \mu_0$, 对任何 $n \geqslant 1$, $x_i \in E$, $0 \leqslant i \leqslant n$, 归纳定义

$$\mu^{[n]}(x_0, \cdots, x_{n-1}, x_n) := \mu^{[n-1]}(x_0, \cdots, x_{n-1}) p_n((x_0, \cdots, x_{n-1}), x_n).$$

容易验证 $\mu^{[n]}$ 是概率测度且

$$\sum_{x \in E} \mu^{[n]}(x_0, \cdots, x_{n-1}, x) = \mu^{[n-1]}(x_0, \cdots, x_{n-1}).$$

因此 $\{\mu^{[n]} : n \in \mathbf{N}\}$ 是相容概率测度列. 由定理 3.3推出存在概率空间 $(\Omega, \mathcal{F}, P)$ 及其上随机序列 $\{\xi_n : n \in \mathbf{N}\}$ 满足

$$P(\xi_0 = x_0, \cdots, \xi_{n-1} = x_{n-1}, \xi_n = x_n) = \mu^{[n]}(x_0, \cdots, x_{n-1}, x_n).$$

由此容易推出上面的条件概率. □

再看例 3.2 和例 3.3, 随机过程的存在性是显然的. 对例 3.2, 任意固定 $a \in E$, 定义 $\mu_0(x) = 1_{\{x=a\}}$,

$$p_n((x_0, \cdots, x_{n-1}), x) = \frac{1_{\{x \sim x_{n-1}\}}}{c(x_{n-1})},$$

其中 $\sim$ 表示两者相邻. 对例 3.3, 时间从 1 开始, 定义 $\mu_1(x) = 1_{\{x=1\}}$,

$$p_{n+1}((x_1, \cdots, x_n), x) = \begin{cases} \frac{1}{n+1} \sum_{i=1}^n 1_{\{x_i = x\}}, & x \leqslant \max(x_1, \cdots, x_n); \\ \frac{1}{n+1}, & x = \max(x_1, \cdots, x_n) + 1; \\ 0, & x > \max(x_1, \cdots, x_n) + 1. \end{cases}$$

容易验证两者都满足推论的条件.

3.3　布朗运动的构造

3.3.1　布朗运动的有限维分布族

布朗运动是微粒由于大量分子的随机地无序地不间断地撞击所产生的随机移动, 看上去毫无规则. 从点 x 出发的布朗运动的数学模型是概率空间 $(\Omega, \mathcal{F}, P)$ 上一个连续时间状态空间为 $\mathbf{R}^d$ 的随机过程 $B = (B_t : t \geqslant 0)$, 根据 §3.2.2 中的分析, 我们总结出它应该满足下面的条件.

定义 3.10　概率空间 $(\Omega, \mathcal{F}, P)$ 上的一个随机过程 $B = (B_t)_{t \geqslant 0}$ 被称为是布朗运动, 如果它满足

(1) 轨道连续: 对几乎所有 ω, $t \mapsto B_t(\omega)$ 是连续的;

(2) 增量正态: 对任何 $t > s \geqslant 0$, 增量 $B_t - B_s$ 是一个期望零方差为 $t - s$ 的正态分布的独立随机向量;

(3) 独立增量: 对任何 $n \geqslant 1$, 以及 $0 < t_1 < \cdots < t_n$, 增量

$$B_{t_n} - B_{t_{n-1}}, \cdots, B_{t_2} - B_{t_1}, B_{t_1}$$

独立.

随机变量 B_0 是布朗运动的初始位置, 它的分布称为初始分布. 如果 $B_0 = x$ a.s., 那么我们说布朗运动从 x 出发. 一个从零点出发的布朗运动被称为标准布朗运动.

性质 (1) 是因为粒子运动直观上是连续的. 性质 (2)(3) 刻画布朗运动的分布性质. 这主要归功于 Albert Einstein 在 1905 年的发现.

满足 (1)(2)(3) 的随机过程是我们根据直观所预期存在的, 我们必须要证明它是真实存在的. 性质 (1) 是轨道性质, 留到最后. 性质 (2)(3) 决定布朗运动的分布性质, 它们一起唯一决定布朗运动的有限维分布族. 取 $I = \{0 < t_1 < t_2 < \cdots < t_n\}$, 下面我们就来计算 $(B_{t_i} : 1 \leqslant i \leqslant n)$ 的联合分布 μ^I. 假设 $B_0 = x$. 为了方便书写, 我们先把正态分布 $N(0,t)$ 的密度函数记为 $p(t,y)$, 即

$$p(t,y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{y^2}{2t}\right) \ t > 0, y \in \mathbf{R}.$$

这时, $B_t = B_t - B_0 + x \sim N(x,t)$, 密度函数为 $p(t, y-x)$. 因为独立增量性,

$$(B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}})$$

的联合密度为

$$p(t_1, x_1 - x)p(t_2 - t_1, x_2) \cdots p(t_n - t_{n-1}, x_n).$$

注释 3.1　让我们约定, 说一个随机向量 $(X_1, \cdots, X_n)$ 具有密度 $p(x_1, \cdots, x_n)$ 是指对任何非负或者有界可测函数 $f : \mathbf{R}^n \to \mathbf{R}$ 使得

$$E[f(X_1, X_2, \cdots, X_n)] = \int_{\mathbf{R}^n} f(x_1, \cdots, x_n)p(x_1, \cdots, x_n)dx_1 \cdots dx_n.$$

由此推出 $(B_{t_i} : 1 \leqslant i \leqslant n)$ 的联合密度是

$$p(t_1, x_1 - x)p(t_2 - t_1, x_2 - x_1) \cdots p(t_n - t_{n-1}, x_n - x_{n-1}),$$

也就是说 μ^I 是该密度对应的分布.

为了说明满足 (2)(3) 的随机过程真实存在, 只需要做两件事情:

1. 证明上面的有限维分布族 $D = \{\mu^I : I \Subset (0, \infty)\}$ 是相容的;

2. 应用 Kolmogorov 相容性定理, 有一个随机过程 $X = (X_t : t > 0)$ 具有有限维分布族 D. 然后证明 X 满足 (2)(3). 令 $X_0 = x$, 得到随机过程 $X = (X_t : t \geqslant 0)$, 它仍然满足 (2),(3).

这两件事情并不难, 只需按照定义验证, 留给读者完成.

练习 3.6　验证上面两个结论.

最困难的事情是证明满足 (2)(3) 的随机过程中至少可以找到一个满足 (1), 这样布朗运动的存在性证明就完成了. 实际上, 布朗运动的存在性证明的真正历史不是这样的. 历史上, 在 20 世纪初, 物理学家与数学家已经知道布朗运动的有限维分布族, 但一直没有办法证明 (1). 这个问题也许对于物理学家来说这不是很重要, 因为他们通常觉得在直觉上合理就可以了, 数学上的严密不是最重要的. 但对于数学家来说, 这个问题却不可能回避. 在 20 世纪初, 概率论还没有公理化 (概率公理化在 1933 年完成), 但由于 Lebesgue 1900 年的工作, 测度的思想已经很成熟了, 所以布朗运动的存在性问题是以连续函数空间上的构造测度的形式提出来的. 让我们把它简单地介绍一下, 取区间 $[0,1]$ 作为时间, 这个不是很重要, 在 $[0,+\infty)$ 上的布朗运动构造并没有本质差别, 令 $W = C[0,1]$, $[0,1]$ 上连续函数空间, 赋予一致范, 它是度量空间, 其 Borel σ-代数记为 $\mathcal{B}(W)$. 对任何 $t \in [0,1]$, 定义

$$\pi_t(w) := w(t),\ \ w \in W,$$

它是 W 上可测函数. 现在的问题归结于是否存在 $(W, \mathcal{B}(W))$ 上的测度 μ 使得

$$\mu(\{w \in W : (\pi_t(w) : t \in I) \in A\}) = \mu^I(A),$$

其中 $A \in \mathcal{B}(\mathbf{R}^I)$. 这个问题困扰了数学家很多年, 一直到 1923 年由数学家 Nobert Wiener 完成证明. 所以空间 $(W, \mathcal{B}(W), \mu)$ 称为 Wiener 空间, μ 称为 Wiener 测度, 这时 $(\pi_t : t \in [0,1])$ 是一个随机过程, 它满足 (1)(2)(3)(4), 是布朗运动, 也称为 Wiener 过程. 布朗运动和 Wiener 测度实际上代表了概率学者与分析学者看待布朗运动的两个角度, 概率学者的重点在于随机过程或者样本轨道, 分析学者的重点在于测度, 本质上一样. 例如当分析学者说研究 Wiener 测度的平移性质的时候, 是说 W 上 Wiener 测度 μ 通过平移算子 $T_y(x) = x + y$, $x, y \in W$, 所诱导的测度 $\mu \circ T_y^{-1}$ 与 μ 之间的关系 (Cameron-Martin 定理), 而概率学者是问布朗运动漂移之后 $B_t + y(t), t \geqslant 0$, 在一个什么概率下是布朗运动 (Girsanov 定理).

3.3.2　布朗运动的构造

我们要证明 X 有一个修正是具有连续轨道的, 通俗地说, X 有连续修正, 而修正与原过程具有相同的有限维分布族, 因此不改变过程的分布. 对比于 Wiener 原始的证明, 概率思想更浓, 可能更容易理解.

思想如下: 取时间集的一个可数稠密子集上的随机变量组成的随机过程, 证明它有连续扩张, 然后证明扩张是原过程的一个修正.

设 N 是任何给定的自然数, $\mathbf{D}$ 是 $[0,\infty)$ 上二分点全体, 即

$$\mathbf{D} := \{\frac{k}{2^m} : m, k = 0, 1, 2, \cdots\},$$

令 $\mathbf{D}_N := \mathbf{D} \cap [0, N]$. 形如 $\frac{k}{2^m}$ 的二分点称为 m 阶二分点.

引理 3.6 **设 f 是 $\mathbf{D}$ 上的函数, 如果存在常数 $\alpha > 0$, 对任何自然数 N, 存在常数 $C > 0$ 及自然数 u, 使得当 $m \geqslant u$ 时,**

$$\left|f(\frac{k+1}{2^m}) - f(\frac{k}{2^m})\right| \leqslant \frac{C}{2^{m\alpha}},\ k = 0, 1, \cdots, 2^m N - 1,$$

则 f 可以唯一地延拓成 $[0,\infty)$ 上的连续函数, 且是 α 阶局部 Hölder 连续的, 即在任何 $[0,\infty)$ 的有界区间上是 α 阶 Hölder 连续的.

证明. 固定自然数 N. 我们期望从条件推出存在常数 C', 只要 $t, s \in \mathbf{D}_N$, 且 $|s-t| \leqslant 1/2^u$ 就有

$$|f(s) - f(t)| \leqslant C'|s-t|^\alpha.$$

不妨设 $t > s$. 取 $m \geqslant u$, 使得 $1/2^{m+1} < |t-s| \leqslant 1/2^m$. 首先, s 落在某个区间 $[\frac{k}{2^m}, \frac{k+1}{2^m})$ 中, 我们来计算 $f(s)$ 与 f 在左端点的值 $f(\frac{k}{2^m})$ 之间的差, 因 $s \in \mathbf{D}_N$, 写出其二进表示, 形式为

$$s = \frac{k}{2^m} + \frac{b_1}{2^{m+1}} + \cdots + \frac{b_l}{2^{m+l}},$$

其中 b_j 取值为 0 或 1, 令

$$s_0 := \frac{k}{2^m},\ s_j := s_0 + \sum_{i=1}^{j} \frac{b_i}{2^{m+i}},\ j = 1, 2, \cdots, l.$$

显然 $\frac{k}{2^m} = s_0 \leqslant s_1 \leqslant \cdots \leqslant s_l = s$, 且 s_j 与 s_{j-1} 或相同或是 $m+j$ 阶二分点的相邻点, 由定理条件

$$|f(s_j) - f(s_{j-1})| \leqslant \frac{C}{2^{(m+j)\alpha}},\ j = 1, 2, \cdots, l.$$

因此

$$|f(s) - f(\frac{k}{2^m})| \leqslant \sum_{j=1}^{l} |f(s_j) - f(s_{j-1})| \leqslant \sum_{j=1}^{\infty} \frac{C}{2^{(m+j)\alpha}} = \frac{C}{2^{m\alpha}(2^\alpha - 1)},$$

右边与 k 无关.

其次, t 必定落在区间

$$[\frac{k}{2^m}, \frac{k+1}{2^m}),\ [\frac{k+1}{2^m}, \frac{k+2}{2^m})$$

之一, 那么 $f(t)$ 与 f 在 t 所在区间之左端点的值的差同样不大于 $\frac{C}{2^{m\alpha}(2^\alpha-1)}$, 因此无论哪种情况,

$$|f(s)-f(t)| \leqslant \frac{C}{2^{m\alpha}} + \frac{2C}{2^{m\alpha}(2^\alpha-1)} = \frac{C_1}{2^{m\alpha}},$$

其中常数 $C_1 = C + \frac{2C}{2^\alpha-1}$.

最后

$$|f(s)-f(t)| \leqslant C_1(1/2^m)^\alpha < C_1(2|s-t|)^\alpha = C'|s-t|^\alpha.$$

练习 3.7 证明: 存在常数 c 使得对任何 $s, t \in D_N$ 有

$$|f(s)-f(t)| \leqslant c|s-t|^\alpha.$$

这蕴含 f 在 $\mathbf{D}_N$ 上一致连续. 因此 f 可以唯一地延拓为 $[0,N]$ 上的连续函数, 记为 F_N. 满足对任何 $s, t \in [0,N]$ 且 $|s-t| \leqslant 1/2^u$ 有

$$|F_N(s)-F_N(t)| \leqslant C_1|s-t|^\alpha,$$

即 F_N 是 $[0,N]$ 上 α 阶 Hölder 连续的.

显然 $\{F_N : N \geqslant 1\}$ 是相容的, 即 F_{N+1} 限制在 $[0,N]$ 上与 F_N 是一致的, 对任何 $x \in [0,\infty)$, 如 $x < N$, 定义 $F(x) := F_N(x)$, 由相容性, 定义无歧义, 上面的讨论证明了 F 是 f 的唯一延拓, 且 F 是局部 α 阶 Hölder 连续的. □

设 $X = (X_t)$ 是上一节构造的满足定义 3.10(2)(3) 与 $X_0 = x$ 的随机过程.

定理 3.5 X **有一个连续修正.**

证明. 不失一般性, 设 $d=1$, $x=0$. 先让我们计算增量 $X_t - X_s$, $t > s \geqslant 0$ 的矩, 因热核是对称的, 奇数阶矩都是零, 对整数 $k \geqslant 1$,

$$E|X_t - X_s|^{2k} = \int_{\mathbf{R}} \frac{x^{2k}}{\sqrt{2\pi(t-s)}} e^{-\frac{x^2}{2(t-s)}} dx = |t-s|^k C(k),$$

而 $C(k) := E|X_1 - X_0|^{2k} = (2k-1)!!$, 特别地, $E|X_t - X_s|^4 = 3|t-s|^2$.

实际上，我们只需要考虑二分点时间上的那部分过程就可以了. 取 $a > 0$, 由 Chebyshev 不等式,

$$P\left(|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \geqslant \frac{1}{2^{ma}}\right) \leqslant 2^{4ma}E|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}|^4 = \frac{3}{2^{2m-4ma}},$$

然后, 对任何自然数 N,

$$P\left(\bigcup_{k=0}^{2^m N-1}\{|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \geqslant \frac{1}{2^{ma}}\}\right) \leqslant \sum_{k=0}^{2^m N-1}\frac{1}{2^{2m-4ma}} = \frac{3N}{2^{m-4ma}}.$$

让 $4a < 1$, 则

$$\sum_{m\geqslant 0} P\left(\bigcup_{k=0}^{2^m N-1}\{|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \geqslant \frac{1}{2^{ma}}\}\right) < \infty.$$

由 Borel-Cantelli 引理, 若记

$$\Omega_0^N := \liminf_m \bigcap_{k=0}^{2^m N-1}\{|X_{\frac{k+1}{2^m}} - X_{\frac{k}{2^m}}| \leqslant \frac{1}{2^{ma}}\},$$

则 $\Omega_0^N \in \mathcal{F}$ 且 $P(\Omega_0^N) = 1$, 再记 $\Omega_0 := \bigcap_N \Omega_0^N$, 则 $P(\Omega_0) = 1$.

对任何 $\omega \in \Omega_0$, 轨道函数 $t \mapsto X_t(\omega)$ 满足: 对任何自然数 N, 存在 $u = u(\omega, N)$, 使得当 $m \geqslant u$ 时,

$$|X_{\frac{k+1}{2^m}}(\omega) - X_{\frac{k}{2^m}}(\omega)| \leqslant \frac{1}{2^{ma}},\ k = 0, 1, 2, \cdots, 2^m N - 1.$$

由引理 3.6, $X_t(\omega), t \in \mathbf{D}$ 有唯一连续延拓, 记为 $B_t(\omega), t \in [0, \infty)$, 它在 $[0, \infty)$ 上连续, 对 $\omega \notin \Omega_0$, 定义 $B_t(\omega) = 0$, 显然对固定的 $t \geqslant 0$, 取点列 $\{t_n\} \subset \mathbf{D}$ 且 $t_n \to t$, 有

$$B_t = \lim_{n\to\infty} X_{t_n} \cdot 1_{\Omega_0},$$

因此 B_t 是随机变量, 故 $B = (B_t : t \geqslant 0)$ 是一个实值随机过程, 另外 B_t 是 $\{X_{t_n}\}$ 的几乎处处收敛的极限, 而 $E|X_{t_n} - X_t|^2 = |t_n - t|$, 即 X_t 是 $\{X_{t_n}\}$ 的 L^2 收敛的极限, 故 $B_t = X_t$ a.s., 连续修正 B 就是布朗运动. □

练习 3.8　利用

$$E|X_t - X_s|^{2k} = C(k)|t - s|^k,$$

证明: 对于任何 $a < 1/2$, B 的几乎所有轨道是 a-阶局部 Hölder 连续的. (但达不到 1/2-阶.)

练习 3.9 考虑 $\mathbf{R}^d$ 上的热传导方程

$$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u,$$

其中 $\triangle$ 是 Laplace 算子. 容易验证方程的基本解是

$$p_d(t,x) := \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2t}},\ t>0, x\in\mathbf{R}^d.$$

也就是方差为 t 的中心化协方差矩阵 tI 的 Gauss 分布的密度函数, 显然 $p(t,x)$ 就是 $p_1(t,x)$. 验证

$$p_d(t,\cdot) * p_d(s,\cdot) = p_d(t+s,\cdot).$$

两个密度函数的卷积是指分别具有两个密度的两独立随机变量和的密度函数.

把定义 3.11 中的 $\mathbf{R}$ 换成 $\mathbf{R}^d$, $p(t,x)$ 换成 $p_d(t,x)$, 所定义的随机过程称为 d-维标准布朗运动, 类似可以证明 d-维布朗运动的存在性. 现在设 $B=(B_t)$ 是 d-维标准布朗运动. 那么 $B_t=(B_1(t),\cdots,B_d(t))$, 下标是坐标编号. 对于点 $x=(x_1,\cdots,x_d)\in\mathbf{R}^d$, 在 P^x 之下, $B^i=(B_i(t))$ 是 x_i 出发的布朗运动且这 d 个布朗运动是独立的. 反之, 我们总是可以构造有 d 个这样的布朗运动, 那么它们拼在一起就是 d-维标准布朗运动.

练习 3.10 设 $B=((B_1(t),B_2(t)):t\geqslant 0)$ 是 2-维标准布朗运动. 证明在任何概率 P^x 下, $\sigma(B_1(t):t\geqslant 0)$ 与 $\sigma(B_2(t):t\geqslant 0)$ 是独立的. 注意说两个事件域 $\mathcal{G}_1,\mathcal{G}_2$ 独立是指任何两个事件 $A_1\in\mathcal{G}_1$, $A_2\in\mathcal{G}_2$ 独立.

3.3.3 Wiener 测度

我们已经证明了布朗运动的存在, 那么是否可以证明 Wiener 测度的存在呢? 设 $B=(B_t:t\geqslant 0)$ 是 $(\Omega,\mathcal{F},P)$ 上标准布朗运动, 对于 $\omega\in\Omega$, 如果 $t\mapsto B_t(\omega)$ 连续, 定义 $\xi(\omega)(t):=B_t(\omega)$, 否则定义 $\xi(\omega)\equiv 0$. W 是 $[0,\infty)$ 上的连续函数全体, $\mathcal{B}(W)$ 是 Borel σ-代数. 那么 $\xi(\Omega)\subset W$, 且 P 通过 ξ 在上诱导测度 μ, 但是 W 作为 $\mathbf{R}^{[0,\infty)}$ 的子集不是可测的, 但可以证明 W 的外测度等于 1. 因此 μ 限制在 W 上的测度是 Wiener 测度. 细节在下面的定理中完成. 下面我们近考虑时间 $[0,1]$ 的情况, 时间 $[0,\infty)$ 也是一样的, 不过这时 W 上需要重新定义度量.

定理 3.6 设 $E=\mathbf{R}$, $T=[0,1]$, μ 是 $B=(B_t)_{t\in T}$ 的分布 (参考定义 3.7), 证明:

1. $W \notin \mathcal{E}^T$,

2. $\mu^*(W) = 1$, μ^* 表示外测度, 定义为

$$\mu^*(W) = \inf\{\mu(B) : B \supset W, B \in \mathcal{E}^T\}.$$

3. $\mu_*(W) = 0$, μ_* 表示内测度, 定义为

$$\mu_*(A) = \sup\{\mu(B) : B \subset W, B \in \mathcal{E}^T\}.$$

4. $\mathcal{B}(W) = W \cap \mathcal{E}^T := \{B \cap W : B \in \mathcal{E}^T\}$.

5. 对于 $A \in \mathcal{B}(W)$, 存在 $B \in \mathcal{E}^T$ 使得 $A = W \cap B$, 定义

$$\mu'(A) := \mu(B).$$

则 μ' 在是良定义的, 且是 Wiener 测度 (仍然用符号 μ 表示).

证明. 1. $\mathcal{E}^T$ 是柱集全体生成的 σ-代数. 柱集是限定有限个时间处位置的集合, 我们来证明 $A \in \mathcal{E}^T$ 是限定可列个时间处位置的集合, 即满足称为可列决定的性质: 存在至多可数 $\{t_n : n \geqslant 1\} \subset T$ 使得只要 $x, y \in E^T$ 在所有 t_n 处相等, 则 x, y 或者都在 A 中或者都不在 A 中. 事实上, 如果 A 是柱集, 那么它一定满足这个性质. 另外容易验证满足可列决定性质的集合 A 全体是一个 σ-代数, 所以它必定包含 $\mathcal{E}^T$.

练习 3.11　验证满足可列决定性质的集合全体是一个 σ-代数.

但是 W 是连续函数空间, 它显然不是一个满足可列决定性质的集合, 因为不管什么样的可数集 $\{t_n\}$ 都能够找到一个 $x \in W$ 而 $y \notin W$ 使得 x, y 在 $\{t_n\}$ 上相等.
2. 为了证明 $\mu^*(W) = 1$, 只需证明对任何包含 W 的可测集 $B \in \mathcal{E}^T$, 有 $\mu(B) = 1$. 事实上, 因为 $\xi(\Omega) \subset W \subset B$, 所以 $\Omega = \xi^{-1}(B)$, 因此

$$\mu(B) = P \circ \xi^{-1}(B) = P(\Omega) = 1.$$

3. $\mu_*(W) = 0$ 是因为连续函数的任何非空子集都不可能是可测集.
4. 事实上, 只需验证 $W \cap \mathcal{E}_0^T$ 是 W 的一个拓扑基, 而这由 W 中的函数是一致连续的性质推出. 具体细节留给读者.

练习 3.12　验证 $\mathcal{B}(W) = W \cap \mathcal{E}^T := \{W \cap A : A \in \mathcal{E}^T\}$.

5. 假设有 $B, B' \in \mathcal{E}^T$ 满足 $W \cap B = W \cap B'$, 这时

$$B \triangle B' = (B \setminus B') \cup (B' \setminus B) \subset W^c,$$

因此左边事件的余集包含 W, 推出其概率为 1, 即 $\mu(B \triangle B') = 0$, 这蕴含着 $\mu(B') = \mu(B)$. 这说明 μ' 是无歧义的. □

Wiener 测度 μ 是 $(W, \mathcal{B}(W))$ 上的测度, 但是它只能通过有限维分布来表达, 如同 Lebesgue 测度本质上只能通过区间来表达一样. 设 $I = \{0 < t_1 < \cdots < t_n \leqslant 1\}$, $H \in \mathcal{B}(\mathbf{R}^I)$, 则

$$\begin{aligned}
&\mu(\{x \in W : (x(t_i), 1 \leqslant i \leqslant n) \in H\}) = \mu^I(H) \\
&\quad = P(\{(B_{t_1}, \cdots, B_{t_n}) \in H\}) \\
&\quad = \int_H \prod_{i=1}^n p(t_i - t_{i-1}, x_i - x_{i-1}) dx_1 \cdots dx_n,
\end{aligned}$$

其中 $t_0 = 0$, $x_0 = 0$.

对任何 $x \in \mathbf{R}$, 过程 $B + x = (B_t + x)$ 在 W 上的像测度记为 μ^x, 那么 $\mu^x(\{w \in W : w(0) = x\}) = 1$, 即在测度 μ^x 之下, 轨道过程是 x 出发的布朗运动. 因此下面定义的布朗运动是存在的.

定义 3.11 **样本空间 $(\Omega, \mathcal{F})$ 上的概率测度族 $(P^x : x \in \mathbf{R})$ 和随机过程 $B = (B_t)$ 被称为标准布朗运动, 如果**

1. **对任何 $x \in \mathbf{R}$, $n \geqslant 1$, $I = (0 < t_1 < \cdots t_n)$, 有**

$$\begin{aligned}
&P^x((B_{t_1}, \cdots, B_{t_n}) \in H) \\
&\quad = \int_H \prod_{i=1}^n p(t_i - t_{i-1}, x_i - x_{i-1}) dx_i, \ H \in \mathcal{B}(\mathbf{R}^I),
\end{aligned}$$

   其中 $t_0 = 0$, $x_0 = x$.

2. **连续: 存在 $\Omega_0 \in \mathcal{F}$, 对任何 x, $P^x(\Omega_0) = 1$, 使得对 $\omega \in \Omega_0$, $t \mapsto B_t(\omega)$ 是 $[0, +\infty)$ 到 $\mathbf{R}$ 的连续函数, 也就是说几乎所有样本轨道连续.**

注意, 这个是说在概率 P^x 之下, B 是 x 出发的布朗运动. 另外, 如果需要我们不妨假设布朗运动的所有轨道都是连续的, 如同在 Wiener 空间上一样.

布朗运动还可以用 Fourier 级数的方法产生, 这个方法极其精妙, 也归功于 Wiener. 这个构造方法适用性不广, 但不需要 Kolmogorov 相容性定理.

定义 3.12　一个随机过程 $X = (X_t)$ 被称为是 Gauss 过程, 如果它的任何有限维分布是 Gauss 分布 (即正态分布). 这时如果对任何 t, $EX_t = 0$, 则说 X 是中心化 Gauss 过程.

假设过程 $X = (X_t)$ 满足 §3.3.1 开头的性质 (2)(3), 且 $X_0 = 0$. 那么 X 是一个中心化的 Gauss 过程且有 $E[X_tX_s] = t \wedge s$. 反过来也成立, 我们把它作为一个练习. 证明的关键是两点, 一是 Gauss 分布随机向量的线性变换还是 Gauss 分布, 二是 Gauss 分布随机向量独立当且仅当协方差矩阵是对角阵.

练习 3.13　证明: 随机过程 $X = (X_t)$ 满足定义 3.10中的 (2)(3) 且 $X_0 = 0$ 当且仅当 X 是一个中心化的 Gauss 过程且对任何 $t, s \geqslant 0$, 有 $E[X_tX_s] = t \wedge s$.

设 $\{\xi_n : n \geqslant 0\}$ 是概率空间 $(\Omega, \mathcal{F}, P)$ 上的都服从标准正态分布的独立随机变量序列 (说明其存在性). 函数

$$\{1/\sqrt{\pi}, \sqrt{2/\pi}\cos x, \cdots, \sqrt{2/\pi}\cos nx, \cdots\}$$

是 $H = L^2([0,\pi]$ 的标准正交基, 依次记为 $\{e_n : n \geqslant 0\}$. 对任何 $f \in L^2([0,\pi])$, 令

$$H(f) := a_0\xi_0 + a_1\xi_1 + a_2\xi_2 + \cdots,$$

其中 $\{a_n\}$ 是 f 的 Fourier 系数: $a_n := \langle f, e_n\rangle$. 随机变量族 $\{H(f) : f \in H\}$ 是一个 Gauss 随机过程或者随机场, 即它的任何有限维分布是 Gauss 分布.

练习 3.14　证明: 对于任何 $f_1, \cdots, f_n \in H$, 随机向量 $(H(f_1), \cdots, H(f_n))$ 是 Gauss 分布的.

显然

$$E[H(f)^2] = \sum_{n\geqslant 0} a_n^2 = \int f^2(x)dx,$$

即 H 是 $L^2([0,\pi])$ 到 $L^2(\Omega, \mathcal{F}, P)$ 的一个等距嵌入. 对 $t \in [0,\pi]$, 显然 $1_{[0,t]}$ 的 Fourier 级数为

$$1_{[0,t]}(x) = \frac{t}{\pi} + \frac{2}{\pi}\sum_{n\geqslant 1}\frac{\sin nt}{n}\cdot\cos nx.$$

令

$$X_t := H(1_{[0,t]}) = \frac{t}{\sqrt{\pi}}\xi_0 + \sqrt{\frac{2}{\pi}}\sum_{n\geqslant 1}\frac{\sin nt}{n}\cdot\xi_n,$$

显然右边是 L^2 意义的极限.

练习 3.15 验证 $X=(X_t)$ 是中心化 Gauss 过程.

X 的协方差函数等于

$$E[X_tX_s]=E[1_{[0,t]}1_{[0,s]}]=t\wedge s.$$

怎么才能证明 X 有连续修正? 因为 X_t 右边是关于 t 的函数级数, 其部分和关于 t 连续, 所以我们需要证明部分和对几乎所有轨道有一致收敛的极限. 但是右边的级数本身并没有这种收敛, 奇妙的是, 重组之后的级数 (即部分和序列的一个子列) 有这样的收敛性.

定理 3.7 **将级数重组为下面的形式**

$$X_t=\frac{t}{\sqrt{\pi}}\xi_0+\sqrt{\frac{2}{\pi}}\sum_{n\geqslant 1}\left(\sum_{m=2^{n-1}}^{2^n-1}\frac{\sin mt}{m}\xi_m\right).$$

那么级数以概率 1 对于 $t\in[0,\pi]$ 一致收敛. 因此 (X_t) 有连续修正, 它在 $[0,\pi]$ 上是布朗运动.

证明. 证明参考 [4]. 设 $m<n$. 令

$$Y_{m,n}(t):=\sum_{m}^{n-1}\frac{\sin kt}{k}\xi_k,\ Z_{m,n}=:\max_{0\leqslant t\leqslant\pi}|Y_{m,n}(t)|.$$

我们的目的是证明

$$\sum E\left(Z_{2^{n-1},2^n}\right)<\infty.$$

因为这蕴含原级数几乎处处一致收敛.

用 Cauchy 不等式, 不妨估计 $E\left(Z^2_{m,n}\right)$, 注意理解下面每一步估计的理由, 其中还是有一些技巧值得学习,

$$\begin{aligned}
E\left(Z^2_{m,n}\right)&\leqslant E\left\{\max_t\left|\sum_m^{n-1}\frac{e^{ikt}}{k}\xi_k\right|^2\right\}\\
&\leqslant\sum_m^{n-1}\frac{1}{k^2}+2E\left(\sum_{l=1}^{n-m-1}\left|\sum_{j=m}^{n-l-1}\frac{\xi_j\xi_{j+l}}{j(j+l)}\right|\right)\\
&\leqslant\sum_m^{n-1}\frac{1}{k^2}+2\sum_{l=1}^{n-m-1}\left[E\left(\left|\sum_{j=m}^{n-l-1}\frac{\xi_j\xi_{j+l}}{j(j+l)}\right|^2\right)\right]^{\frac{1}{2}}
\end{aligned}$$

$$\leqslant \sum_{m}^{n-1}\frac{1}{k^2}+2\sum_{l=1}^{n-m-1}\left(\sum_{j=m}^{n-l-1}\frac{1}{j^2(j+l)^2}\right)^{\frac{1}{2}}$$
$$\leqslant \frac{n-m}{m^2}+2\frac{(n-m)^{3/2}}{m^2}\leqslant 3\frac{(n-m)^{3/2}}{m^2},$$

由此推出

$$E\left(\sum_{n\geqslant 1}Z_{2^{n-1},2^n}\right)\leqslant \sum_{n\geqslant 1}\left(E[Z^2_{2^{n-1},2^n}]\right)^{1/2}\leqslant C\sum_n 2^{-n/4}<+\infty.$$

□

练习 3.16　完成上面定理的证明, 严格地写出最后一步: (X_t) 有一个修正是连续的.

习　题

1. (Kolmogorov) 如果存在正常数 α, β, C 使得实值过程 X 满足对任何 $t, h > 0$,

$$E|X_{t+h}-X_t|^\alpha \leqslant C\cdot h^{1+\beta}.$$

证明: X 有连续修正.

2. 设 $y \in W$, μ 是 Wiener 测度. 定义 $T_y x := x + y$, $x \in W$. 证明: T_y 是 W 到自身的可测映射. 记 $\mu_y := \mu\circ T_y^{-1}$, 即 Wiener 测度在平移之下的像. 用随机过程的语言来说, μ 是布朗运动 $B=(B_t)$ 的分布, μ_y 是漂移布朗运动 $B_t^y = B_t + y(t)$, $t \geqslant 0$, 的分布.

   (a) 求 μ_y 的有限维分布, 也就是说对 $I = \{0 < t_1 < t_2 < \cdots < t_n\} \Subset [0,1]$, 随机向量 $(\pi_t : t \in T)$ 在 μ_y 下的分布, 记为 μ_y^I.

   提示:

$$\begin{aligned}\mu_y^I(f) &= \mu_y(f(\pi_t : t\in I)) = \mu(f(\pi_I\circ T_y))\\ &= \int f(x(t_1)+y(t_1),\cdots,x(t_n)+y(t_n))\mu(dx)\\ &= \int f(x_1+y(t_1),\cdots,x_n+y(t_n))\mu^I(dx_1\cdots dx_n),\end{aligned}$$

$$
\begin{aligned}
&= \int f(x_1,\cdots,x_n)h_y(x_1,\cdots,x_n)\mu^I(dx_1\cdots dx_n) \\
&= \mu^I(h_y\cdot f),
\end{aligned}
$$

其中 f 是有限乘积空间上非负可测函数. 这说明 $\mu_y^I \ll \mu^I$ 且密度函数 $\dfrac{d\mu_y^I}{d\mu^I} = h_y$.

(b) 证明 (*): 当且仅当 y 绝对连续且 $y' \in L^2[0,1]$ 时, $\mu_y \ll \mu$.

(这个问题说明即使所有有限维分布绝对连续, 也不一定有 $\mu \circ T_y^{-1} \ll \mu$.)

3. 设 $(W, \mathcal{B}(W))$ 是 $[0,1]$ 上连续函数全体及其柱集生成的 σ- 代数, 装备最大值范数. 证明: 这个拓扑生成的 Borel σ- 代数恰是 $\mathcal{B}(W)$. 令 x_n 是支撑在 $[0,1/n]$ 上但积分 $\int_0^1 x_n(t)dt$ 趋于无穷的连续函数列. 定义 W 上平移算子

$$
T_n x := x + x_n,\ x \in W,
$$

W 上任意给定概率测度 μ, 记 $\mu_n = \mu \circ T_n^{-1}$. 证明: μ_n 的任意有限维分布弱收敛于 μ 的对应有限维分布, 但是 μ_n 不弱收敛于 μ, 注意弱收敛是指对 W 上任意有界连续函数 f 有 $\mu_n(f) \to \mu(f)$. 提示: 考虑函数

$$
f(x) = \exp\left(-\left|\int_0^1 x(t)dt\right|\right).
$$

第四章　布朗运动的性质

4.1　性质

在本节中, $B = (B_t)$ 是标准布朗运动. 布朗运动有许多好的性质. 首先, 我们在前一节中实际上已经证明了下列刻画.

定理 4.1　**一个实值随机过程 $B = (B_t)$ 是标准布朗运动当且仅当它是连续的, 中心化的 Gauss 过程且 $E(B_t B_s) = t \wedge s$.**

由此验证下列关于布朗运动的性质, 详细的证明留给读者. 尤其要注意 (3) 中的 B' 之样本轨道在 $t = 0$ 的几乎处处连续性.

推论 4.1　**设 $B = (B_t : t \geqslant 0)$ 是一个布朗运动.**

(1) **马氏性**: **对** $s \geqslant 0$, $(B_{t+s} - B_s : \ t \geqslant 0)$ **是一个独立于** $\mathcal{F}_s^0$ **的标准 Brown 运动;**

(2) **自相似性**: **对任何** $c \neq 0$, $(\frac{1}{c} B_{c^2 t} : \ t \geqslant 0)$ **是一个 Brown 运动;**

(3) 0 **与** ∞ **的对称性**: **设** B **是标准的, 定义**

$$B'_t = \begin{cases} tB_{\frac{1}{t}}, & t > 0, \\ 0, & t = 0, \end{cases}$$

则 $B' = (B'_t : t \geqslant 0)$ **也是标准 Brown 运动.**

练习 4.1　证明推论中的所有结论. 注意 (3) 中要证明 B' 的几乎所有轨道在 $t = 0$ 时连续.

定理 4.2 设 $(\mathcal{F}_t)$ 是一个流, $X = (X_t)_{t\geqslant 0}$ 是一个实值连续适应过程且满足对任何 $t > s \geqslant 0$, $x \in \mathbf{R}$,

$$E[\exp(ix(X_t - X_s))|\mathcal{F}_s] = \exp(-\frac{1}{2}x^2(t-s)).$$

则 X 是布朗运动. 这时它被称为 $(\mathcal{F}_t)$-布朗运动.

练习 4.2 利用特征函数唯一性证明 $X_t - X_s$ 与 $\mathcal{F}_s$ 独立.

4.2　轨道粗糙性

在本节中, $B = (B_t)$ 是标准布朗运动. 虽然布朗运动的轨道是连续的, 但下面的定理说明它们不是那么光滑. 连续函数的光滑性是可以感觉的, 点点可导的函数是光滑的, 二次可导的函数更光滑, 等等. 往下是 Lipschitz 函数, 它可以在一些点不可导, 例如 $y = |x|$. 接着往下是有界变差函数, 它是几乎处处可导的再往下是 a-阶 Hölder 连续函数, $a \in (0,1)$, a 越小函数的光滑度越差. 前面我们已经知道布朗运动的几乎所有轨道是局部 a-阶 Hölder 连续的, 其中只要 $a < 1/2$. 另一方面, 函数的变差也是考察函数的光滑度或者粗糙度的指标. 函数的变差是衡量函数在一个区间变化的程度, 一个有界变差函数是几乎处处可导的, 一个 Lipschitz 函数是有界变差的. 但是有界变差的函数在某些点上可以很不光滑, 例如 $y = |x|^a$, $x \in [-1,1]$ 是有界变差的, 其中 $a \in (0,1)$. 对于 $a' > a$, 该函数在 0 点是 a-阶 Hölder 连续, 但不是 a'-阶 Hölder 连续的. 连续性是局部的, 有界变差性是整体的.

设 f 是区间 $[a,b]$ 上的连续函数, $p > 0$, D 是区间的一个划分, 那么

$$V_D^p f[a,b] := \sum_D |f(t_i) - f(t_{i-1})|^p$$

称为是 f 在 D 上的 p 次变差. 当 $p = 1$ 时是实变函数理论中大家熟知的变差. 注意一次变差的特殊性是三角不等式成立: $|x+y| \leqslant |x| + |y|$, 所以越细的 D 上得到的变差越大. 这时 f 在 $[a,b]$ 上的全变差定义为

$$V^1 f[a,b] := \sup_D V_D^1 f[a,b].$$

当 $V^1 f[a,b] < \infty$, f 被称为有界变差函数.

一个有界变差函数实际上是两个递增函数的差, 另外, 一个有界变差函数是几乎处处可导的. 有界变差函数 f 可以用来定义 Stieltjes 积分 $\int_a^b g df$, 发过来, 对于连续函数 g 以及 $[a, b]$ 的一个划分

$$D : a = t_0 < t_1 < \cdot < t_n = b,$$

定义

$$S_D f(g) := \sum_D g(t_k)(f(t_k) - f(t_{k-1})).$$

$D \to 0$ 是指 $|D| = \max_i |t_i - t_{i-1}| \to 0$.

练习 4.3 证明: 如果对任何连续函数 g, 当 D 趋于零时, $S_D f(g)$ 收敛, 则 f 是有界变差的.

显然

$$V_D^2 f[a, b] \leqslant \max_i |f(t_i) - f(t_{i-1})| V_D^1 f[a, b].$$

定理 4.3 **如果 f 连续, 且有界变差, 那么**

$$\lim_{D \to 0} V_D^2 f[a, b] = 0.$$

这个定理说明对于连续的 f, 如果有一个划分列 $D_n \to 0$ 使得 $V_{D_n}^2 f[a, b]$ 有正极限, 那么 f 不是有界变差的.

定理 4.4 **布朗运动的几乎所有轨道在任何有限区间上都不是有界变差的.**

证明. 设 $t > s \geqslant 0$, 则 $E[(B_t - B_s)^2] = t - s$ 且方差为

$$\begin{aligned} & E[(B_t - B_s)^2 - (t - s)]^2 \\ & \quad = E[(B_t - B_s)^4 - 2(t - s)(B_t - B_s)^2 + (t - s)^2] = 2(t - s)^2. \end{aligned}$$

另外, 设 $t_2 > s_2 \geqslant t_1 > s_1 \geqslant 0$, 则由独立增量性,

$$E[(B_{t_1} - B_{s_1})^2 - (t_1 - s_1)][(B_{t_2} - B_{s_2})^2 - (t_2 - s_2)] = 0.$$

现在, 固定 $t > s \geqslant 0$, 与 $[s, t]$ 的一个划分,

$$D := \{s = t_0 < t_1 < \cdots < t_n = t\}.$$

记 $V_D := \sum_j (B_{t_j} - B_{t_{j-1}})^2$. 其期望 $EV_D = t - s$ 与方差为

$$
\begin{aligned}
E\left(V_D - (t-s)\right)^2 &= E\left(\sum_{j=1}^{n}[(B_{t_j} - B_{t_{j-1}})^2 - (t_j - t_{j-1})]\right)^2 \\
&= \sum_{j=1}^{n} E\left((B_{t_j} - B_{t_{j-1}})^2 - (t_j - t_{j-1})\right)^2 \\
&= \sum_{j=1}^{n} 2(t_j - t_{j-1})^2 \leqslant 2(t-s)|D|,
\end{aligned}
$$

推出 L^2-意义下

$$\lim_{D\to 0} V_D = t - s.$$

因此存在划分列 $\{\Delta_n\}$, 几乎处处地 $V_{D_n} \longrightarrow t - s$. 则因为 B 连续, 使上式收敛到非零数的轨道在 $[s,t]$ 一定不是有界变差的. 从而对几乎所有轨道, 布朗运动在有理数端点的有限区间上都不是有界变差的. 推出定理结论. □

练习 4.4　设 $(B_t)_{t\geqslant 0}$ 是一维标准 Brown 运动. 那么对任何 $t > 0$, 当 n 趋于无穷时, 有

$$\sum_{j=1}^{2^n}\left(B_{\frac{j}{2^n}t} - B_{\frac{j-1}{2^n}t}\right)^2 \to t \quad \text{a.s.} \tag{4.1}$$

练习 4.5　前面已知, 当 $\alpha < 1/2$ 时, 布朗运动的轨道是 α-阶 Hölder 连续的. 证明: 当 $\alpha > 1/2$ 时, 布朗运动在任何区间上都不是 α-阶 Hölder 连续的. 请考虑 $\alpha = 1/2$ 的情况.

因为曲线 $\ell = \{(s, f(s)) : s \in [0,t]\}$ 的长度

$$L \geqslant \sum_i \sqrt{(t_i - t_{i-1})^2 + (f(t_i) - f(t_{i-1}))^2} \geqslant V_D^1 f[0,t],$$

其中 $D = \{t_i\}$ 是 $[0,t]$ 的划分, 所以区间上一个非有界变差的连续函数 f 的图像曲线一定是无穷长的, 这对于一个物理粒子运动是不可能的.

定义 4.1　**设 X 是连续的随机过程, 如果存在随机过程 $Y = (Y_t)$ 使得对任何 $t > 0$, 当 $[0,t]$ 上的划分 $D = \{t_i\}$ 趋于零时, 其二次变差**

$$V_D^2 X[0,t] = \sum_i (X_{t_i} - X_{t_{i-1}})^2$$

依概率收敛于 Y_t, 则说 Y 是 X 的二次变差过程, 记为 $\langle X\rangle$.

显然, 布朗运动的二次变差过程是 $\langle B\rangle_t = t$.

定理 4.5　布朗运动的几乎所有样本轨道是无处可导的.

证明. 证明在 $[0,1]$ 区间上无处可导就足够了. 一个 $[0,1]$ 上的连续函数 $y=f(t)$ 在某点 $0<t_0<1$ 可导是指

$$\lim_{h\to 0}\frac{f(t_0+h)-f(t_0)}{h}$$

存在, 如果是这样的话, 那么一定存在 $c>0$ 与 $\delta>0$, 使得当 $|h|<\delta$ 时有

$$|f(t_0+h)-f(t_0)|\leqslant c|h|.$$

那么当 n 充分大时, 存在 $1\leqslant j\leqslant n-3$ 使得

$$\frac{j}{n},\frac{j+1}{n},\frac{j+2}{n},\frac{j+3}{n}\in(t_0-\delta,t_0+\delta)\ \text{且}\ \frac{j+1}{n}\leqslant t_0<\frac{j+2}{n},$$

因此, 对 $i=j,j+1,j+2$ 有

$$\left|f\left(\frac{i+1}{n}\right)-f\left(\frac{i}{n}\right)\right|\leqslant\left|f\left(\frac{i+1}{n}\right)-f(t_0)\right|+\left|f(t_0)-f\left(\frac{i}{n}\right)\right|\leqslant\frac{4c}{n}.$$

用 D 表示使得轨道 $t\mapsto B_t(\omega)$ 在某个 $t\in[0,1]$ 处可导的样本 $\omega\in\Omega$ 的全体, 那么

$$D\subset\bigcup_{c=1}^{\infty}\bigcup_{l=1}^{\infty}\bigcap_{n\geqslant l}\bigcup_{j=1}^{n-3}\bigcap_{i=j}^{j+2}\left\{\left|B\left(\frac{i+1}{n}\right)-B\left(\frac{i}{n}\right)\right|\leqslant\frac{c}{n}\right\}$$

令

$$D_n=\bigcup_{j=1}^{n-3}\bigcap_{i=j}^{j+2}\left\{\left|B\left(\frac{i+1}{n}\right)-B\left(\frac{i}{n}\right)\right|\leqslant\frac{c}{n}\right\},$$

由布朗运动的性质得

$$\begin{aligned}P(D_n)&\leqslant n\left(P\left\{\left|B\left(\frac{1}{n}\right)\right|\leqslant\frac{c}{n}\right\}\right)^3=n\left(P\left\{B(1)\leqslant\frac{c}{\sqrt{n}}\right\}\right)^3\\&=n\left(\frac{2}{\sqrt{2\pi}}\int_0^{\frac{2c}{\sqrt{n}}}e^{-\frac{x^2}{2}}dx\right)^3<n\left(\frac{2c}{\sqrt{n}}\right)^3\longrightarrow 0.\end{aligned}$$

上面要在区间 $(t_0-\delta,t_0+\delta)$ 内找四个点形成三个区间的原因是为了让这个极限是零. 最后, 由 Fatou 引理,

$$P(D)\leqslant\sum_{c=1}^{\infty}P(\liminf_n D_n)\leqslant\sum_{c\geqslant 1}\liminf_n P(D_n)=0.$$

这说明几乎所有样本轨道无处可导. □

上面证明中, 在某个点可导的事件 D 是包含在一个零概率集中, 所以概率为零. 因为一个有界变差函数是几乎处处可导的, 所以点点不可导的函数肯定在任何区间上都不是有界变差的.

4.3 反射原理

反射原理是远比 Markov 性质更为经典的一个概念. 我们将应用这个原理来计算布朗运动的极大游程的分布. 在许多应用实例中, 特别在统计中, 我们需要估计随机过程的极大游程的分布. 对于布朗运动 $B = (B_t)_{t \geqslant 0}$ 来说, 极大游程 $\sup_{s \in [0,t]} B_s$ 的分布可以由反射原理的方法来导出. 设 $B = (B_t)$ 是标准布朗运动. 固定一个时间 $s > 0$, 让布朗运动在时间 s 之后关于 B_s 反射, 即令

$$B'_t = \begin{cases} B_t, & t \leqslant s, \\ 2B_s - B_t, & t > s, \end{cases}$$

那么 $B' = (B'_t)$ 仍然是连续的, 仍然是中心化 Gauss 过程且 $E[B'_u B'_v] = u \wedge v$. 因此 B' 也是标准布朗运动. 这个性质称为反射原理, 它对于停时也是成立的.

练习 4.6 验证 $E[B'_u B'_v] = u \wedge v$.

设 $b > 0$ 与 $b > a$, 并设

$$T_b = \inf\{t > 0 : B_t = b\},$$

那么 T_b 是停时, 称为 b 的首中时, 由轨道连续性知 $T_b \leqslant t$ 等价于 $\sup_{s \leqslant t} B_s \geqslant b$. 定义 T_b 处的反射

$$B'_t = \begin{cases} B_t, & t \leqslant T_b, \\ 2b - B_t, & t > T_b, \end{cases}$$

那么 $B' = (B'_t)$ 也是标准布朗运动 (证明待后). 用 T'_b 表示 B' 对于 b 的首中时, 实际上, 因 B 与 B' 在到达 b 之前是一样的, 故 $T'_b = T_b$. 因此

$$\begin{aligned} P\{T_b \leqslant t, B_t \leqslant a\} &= P\{T'_b \leqslant t, B'_t \leqslant a\} \\ &= P(T_b \leqslant t, 2b - B_t \leqslant a\} \\ &= P(T_b \leqslant t, B_t \geqslant 2b - a). \end{aligned}$$

该等式称为反射原理. 关于反射原理的严格证明要用到布朗运动的强 Markov 性, 将在后面补充证明.

由反射原理可得当 $b \geqslant a, b > 0$ 时

$$P(\sup_{s\in[0,t]} B_s \geqslant b, B_t \leqslant a) = P(T_b \leqslant t, B_t \leqslant a) = \frac{1}{\sqrt{2\pi t}}\int_{2b-a}^{+\infty} e^{-\frac{x^2}{2t}}dx\ ,$$

$b < a$ 的情况请读者自行推导. 它给出了布朗运动与其极大游程的联合分布.

特别地, 当 $a = b > 0$ 时, 有

$$\begin{aligned} P(T_b \leqslant t) &= 2P(B_t > b) = 2P\left(B_1 > \frac{b}{\sqrt{t}}\right) \\ &= \frac{2}{\sqrt{2\pi}}\int_{b/\sqrt{t}}^{+\infty} \exp\left(-\frac{x^2}{2}\right)dx, \end{aligned}$$

它的密度函数为

$$p(t) = \frac{b}{\sqrt{2\pi t^3}}\exp\left(-\frac{b^2}{2t}\right).$$

由此推出 $P(T_b < \infty) = 1$, 这意味着布朗运动肯定可以在有限时间内命中任何一点 b. 这在直观上是显然的, 因为它的轨道连续而且会在 $-\infty$ 到 $+\infty$ 之间震荡. 但是这个性质与维数有关, 后面将证明 2- 维以上的 Brown 运动不可能 (概率零) 还可以在有限时间内命中一个给定的点.

引理 4.1 T_b 的 Laplace 变换 $E[e^{-sT_b}] = e^{-\sqrt{2s}b}$.

这是一个有趣的积分计算

$$\begin{aligned} E[e^{-sT_b}] &= \int_0^\infty e^{-st}\frac{b}{\sqrt{2\pi t^3}}e^{-\frac{b^2}{2t}}dt \\ &= -\frac{2b}{\sqrt{2\pi}}\int_0^\infty e^{-st-\frac{b^2}{2t}}dt^{-1/2} \\ &= \frac{2b}{\sqrt{2\pi}}\int_0^\infty \exp\left(-\frac{s}{u^2}-\frac{b^2}{2}u^2\right)du \\ &= e^{-\sqrt{2s}b} \end{aligned}$$

最后一个等号还需要一点积分技巧, 留作习题.

4.4　占有时问题 (*)

用 L_{2n} 表示长度为 $2n$ 的格点轨道最后遇到 0 的时间, 即

$$L_{2n} := \sup\{k \leqslant 2n : \ S_k = 0\}$$

被称为 0 点的 (在时刻 $2n$ 前的) 末离时, 它必是偶数.

定理 4.6　$P(L_{2n} = 2k) = u_{2k} \cdot u_{2n-2k}$, $0 \leqslant k \leqslant n$.

证明. 令 $S'_j = \sum_{i=1}^{j} X_{2k+i}$. 显然

$$\begin{aligned}\{L_{2n} = 2k\} &= \{S_{2k} = 0, S_{2k+1} \neq 0, S_{2k+2} \neq 0, \cdots, S_{2n} \neq 0\}\\ &= \{S_{2k} = 0, S'_1 \neq 0, S'_2 \neq 0, \cdots, S'_{2n-2k} \neq 0\},\end{aligned}$$

因此由独立性和定理 2.2,

$$P(L_{2n} = 2k) = P(S_{2k} = 0)P(S'_1 \neq 0, S'_2 \neq 0, \cdots, S'_{2n-2k} \neq 0) = u_{2k} \cdot u_{2n-2k}.$$

完成证明. □

显然对任何 $n \geqslant 0$, 或者 S_{n-1} 与 S_n 都是非负的, 我们说第 n 个时段是正的; 或者 S_{n-1} 与 S_n 都是非正的, 说第 n 个时段是负的. 令 σ_{2n} 是 0 到 $2n$ 时段正的时段数, 即

$$\sigma_{2n} := \sum_{k=1}^{2n} 1_{\{S_{k-1} \geqslant 0, S_k \geqslant 0\}},$$

称为随机游动在正集上的逗留时, 自然 σ_{2n} 也必是偶数.

定理 4.7　$P(\sigma_{2n} = 2k) = u_{2k} \cdot u_{2n-2k}$, $0 \leqslant k \leqslant n$.

证明. 对 n 用归纳法. 当 $n = 1$ 时, 显然 $P(\sigma_2 = 0) = P(\sigma_2 = 2) = 1/2$, 结论成立. 另外由定理 2.2知 $P(\sigma_{2n} = 0) = P(\sigma_{2n} = 2n) = u_{2n}$, 结论成立, 故只需对 $0 < k < n$ 证明即可. 现在设 $P(\sigma_{2m} = 2k) = u_{2k} \cdot u_{2m-2k}$ 对任何 $m < n$ 和 $0 < k < m$ 成立. 用 τ 表示首次返回 0 的时间, 那么当 $0 < k < n$ 时, 由全概率公式和归纳假设,

$$P(\sigma_{2n} = 2k) = \sum_{r=1}^{n} P(\sigma_{2n} = 2k | \tau = 2r)P(\tau = 2r)$$

$$
\begin{aligned}
&= \sum_{r=1}^{n} \frac{1}{2}(P(\sigma_{2n-2r} = 2k-2r) + P(\sigma_{2n-2r} = 2k))P(\tau = 2r) \\
&= \frac{1}{2}\sum_{r=1}^{n}(u_{2k-2r}u_{2n-2k} + u_{2k}u_{2n-2r-2k})P(\tau = 2r) \\
&= \frac{1}{2}(u_{2k}u_{2n-2k} + u_{2k}u_{2n-2k}),
\end{aligned}
$$

由此推出结论, 而最后一个等号由定理 2.3得到. □

令

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}},\ x \in (0,1),$$

那么由 Stirling 公式 $u_{2k} \cdot u_{2n-2k} \sim \frac{1}{n} f(k/n)$. 即有

$$\sum_{k<xn} u_{2k} \cdot u_{2n-2k} \sim \sum_{k/n<x} \frac{1}{n} f(k/n) \longrightarrow \int_0^x f(y)dy = \frac{2}{\pi}\arcsin\sqrt{x},$$

由此推出著名的反正弦律, 也就是说 $L_{2n}/2n$ 和 $\sigma_{2n}/2n$ 的分布函数渐近地是上面的反正弦函数. 注意到密度函数 f 的形状, 我们发现 L_{2n} 与 σ_{2n} 分布都较集中在两端. 令人惊奇的是, 反正弦律在随机过程理论有不可思议的普遍性.

布朗运动有类似的时间. 设 $B = (B_t)$ 是实值的标准布朗运动, 令

$$\rho = \int_0^1 1_{\{B_t>0\}}dt$$

称为占有时, 则可以证明

$$P(\rho \leqslant x) = \lim_n P(\frac{\sigma_{2n}}{2n} \leqslant x) = \frac{2}{\pi}\arcsin\sqrt{x},\ 0 < x < 1.$$

设 $B = (B_t^1, B_t^2)$ 是 2-维标准布朗运动, 令

$$\rho(\pi/2) = \int_0^1 1_{\{B_t^1>0, B_t^2>0\}}dt.$$

公开问题: 这个随机变量是什么分布的? 更一般地, 用 $S(\theta)$ 表示顶点为原点角度为 θ 的扇形, 占有时

$$\rho(\theta) = \int_0^1 1_{\{B_t\in S(\theta)\}}dt$$

是什么分布?

习　题

1. 设 $r < s < t$. 求 $E^0(B_s|B_r, B_t)$.

2. (旋转不变性) 设 $B = (\Omega, \mathcal{F}, \{P^x : x \in \mathbf{R}^d\}, (B_t)_{t\in T})$ 是轨道空间上 d- 维布朗运动, O 是 $\mathbf{R}^d$ 上的正交变换. 定义 Ω 到自身的映射, 仍然用 O 表示:

$$(O\omega)(t) := O(\omega(t)),\ t \in T, \omega \in \Omega.$$

证明: O 是可测的且 $P^x \circ O^{-1} = P^{Ox}$, $x \in \mathbf{R}^d$. 因此 P^0 是正交变换下不变的.

3. 对 $r > 0$, 令 $T_r := \inf\{t : |B_t - B_0| \geqslant r\}$. 对 $x \in \mathbf{R}^d$, 证明:

   (a) $P^x(T_r < \infty) = 1$;

   (b) 分布 $P^x(B_{T_r} \in \cdot)$ 是球面 $\{y : |y - x| = r\}$ 上均匀分布.

4. 设 $B = (B_t)_{t\in T}$ 是 1- 维布朗运动. 证明:

$$P(\sup_{t>0} B_t = +\infty,\ \inf_{t>0} B_t = -\infty) = 1.$$

5. (重对数律 *) 设 $B = (B_t)_{t\in T}$ 是 1- 维布朗运动. 证明:

$$P\left[\limsup_{t\to 0} \frac{B_t}{\sqrt{2t\ln\ln\frac{1}{t}}} = 1\right] = 1;$$

$$P\left[\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\ln\ln t}} = 1\right] = 1.$$

6. 设 B 是 d- 维标准布朗运动. 证明: 对任何 $x \in \mathbf{R}^d$, $\|x\| = 1$, 过程 $(\langle x, B_t\rangle : t \geqslant 0)$ 是 1- 维标准布朗运动.

7. 设 $B = (B(t) : t \in T)$ 是 1- 维标准布朗运动, 令

$$X_t := e^{-t}B(e^{2t}),\ t \in \mathbf{R}.$$

证明: $X = (X_t)$ 是具有马氏性的平稳的中心化 Gauss 过程. 并求出其协方差函数. 过程 X 称为是 Ornstein-Ulenbeck 过程. 注意平稳过程与平稳增量过程的不同.

8. 设 (B_t) 是标准布朗运动, 计算

(a)

$$\mathsf{E}\left[\int_0^t \mathsf{B}_s \mathrm{d}s | \mathsf{B}_t\right];$$

(b)

$$\mathsf{E}\left[\exp\left(\int_0^t \mathsf{B}_s \mathrm{d}s\right)\right].$$

第五章　布朗运动与马氏性

上章所述的反射原理非常的直观, 但并没有真正的证明, 我们将在这一章中争取写清楚它的严格证明. 这不是很容易的事情, 需要很多的准备. 在本章中, 设 $(\Omega, \mathcal{F})$ 上有一族概率测度 $(P^x)_{\mathbf{R}}$ 和随机过程 $B = (B_t)$, 在概率 P^x 之下, B 是从 x 出发的布朗运动. 另外 $P = P^0$. 定义

$$\mathcal{G}_t := \sigma(B_s : s \leqslant t),\ \mathcal{G}_\infty = \sigma(B_s : s \geqslant 0),$$

是布朗运动的自然流.

当我们讨论条件期望 $Y = E[X|\mathcal{G}]$ 时, 是指几乎处处相等, 所以把 $\mathcal{G}$ 换成往 $\mathcal{G}$ 中加入零概率集之后再生成的 σ-代数 $\mathcal{G}'$, 等式仍然成立 $Y = E[X|\mathcal{G}']$.

5.1　简单马氏性

先来阐述布朗运动的简单马氏性. 尽管它对于布朗运动叙述, 但简单马氏性实际上非常一般地成立.

从布朗运动的有限维分布开始观察并思考. 应该牢记的是, 除了轨道连续性, 布朗运动的分布性质完全被其有限维分布族刻画, 而其有限维分布族被转移半群刻画, 所以让我们定义转移核

$$P_t(x, A) := \int_A p(t, y - x)dy = P^x(B_t \in A),\ x \in \mathbf{R}^d, A \in \mathcal{B}(\mathbf{R}^d).$$

固定 x, $P_t(x, \cdot)$ 是个概率测度 (期望为 x 方差为 t 的正态分布), 固定 A, $P_t(\cdot, A)$ 是个可测函数.

引理 5.1 Chapman-Kolmogorov 方程: 对任何 $x \in \mathbf{R}^d, A \in \mathcal{B}(\mathbf{R}^d)$, 有

$$\int_{\mathbf{R}^d} P_t(x, dy) P_s(y, A) = P_{t+s}(x, A). \tag{5.1}$$

按照布朗运动的定义, 对任何 $n \geqslant 1$, $0 = t_0 < t_1 < t_2 < \cdots < t_n$, 考察随机向量 $(B_{t_1}, \cdots, B_{t_n})$, 一个简单的观察是当固定 B_{t_j} 时, 之后的 $(B_{t_{j+1}}, \cdots, B_{t_n})$ 与 $(B_{t_1}, \cdots, B_{t_{j-1}})$ 独立. 怎么把它表达出来? $(B_{t_i} : 1 \leqslant i \leqslant n)$ 在概率 P^x 下的密度是

$$(x_i) \mapsto \prod_{i=1}^n p(t_i - t_{i-1}, x_i - x_{i-1}),$$

即对 $A_1, \cdots, A_n \in \mathcal{B}(\mathbf{R}^d)$,

$$\begin{aligned} & P^x(B_{t_n} \in A_n, \cdots, B_{t_1} \in A_1) \\ & \quad = \int_{\substack{x_i \in A_i, \\ 1 \leqslant i \leqslant n}} \left(\prod_{i=1}^n p(t_i - t_{i-1}, x_i - x_{i-1}) \right) dx_1 dx_2 \cdots dx_n, \end{aligned} \tag{5.2}$$

其中 $x_0 = x$, $t_0 = 0$. 已知联合密度, 可以算条件密度或者分布.
不妨设 t_{n-1} 是现在时刻 t, 容易验证

$$\begin{aligned} & P^x(B_{t_n} \in A_n, \{B_t \in A_{n-1}, \cdots, B_{t_1} \in A_1\}) \\ & \quad = \int_{\substack{x_i \in A_i, \\ 1 \leqslant i < n}} \left(\prod_{i=1}^{n-1} p(t_i - t_{i-1}, x_i - x_{i-1}) \right) P^{x_{n-1}}(B_{t_n - t} \in A_n) dx_1 dx_2 \cdots dx_{n-1} \\ & \quad = E^x \left[P^y(B_{t_n - t} \in A_n) \Big|_{y = B_t}, \{B_t \in A_{n-1}, \cdots, B_{t_1} \in A_1\} \right]. \end{aligned} \tag{5.3}$$

只要不引起混乱, 后面一般把符号 $P^y(\cdots)\big|_{y=B_t}$ 直接写成 $P^{B_t}(\cdots)$. 我们简单说, 在概率 P^x 之下, 给定 $(B_{t_{n-1}}, \cdots, B_{t_1})$, B_{t_n} 的条件分布等于在 P^{B_t} 之下, $B_{t_n - t}$ 的分布.
对于一般的随机过程 (X_t) 而言, X_{t_n} 的分布依赖于 $X_{t_1}, \cdots, X_{t_{n_1}}$. 而现在上面的表达式可以推出 B_{t_n} 的条件分布

$$P^x(B_{t_n} \in A_n | B_{t_{n-1}}, \cdots, B_{t_1})$$

只和离它最近的一个时刻的位置有关. 这就是马氏性.
为了方便表达, 引入推移算子 $\theta_t : \Omega \to \Omega$, 满足

$$B_s(\theta_t \omega) = B_{s+t}(\omega), \ \ \omega \in \Omega,$$

即推移之后的样本轨道上 s 时刻的位置是原样本轨道上 $s+t$ 时刻的位置. 因为布朗运动的样本空间总可以被认为是轨道空间, 也就是说 $\omega \in \Omega$ 是一个轨道 $\omega = (\omega(t): t \geqslant 0)$, 而 $B_t(\omega) = \omega(t)$, 那么对任何 $t \geqslant 0$, $\omega'(s) = \omega(s+t)$, $s \geqslant 0$ 仍然是一个轨道, 定义为 $\theta_t\omega$, 这样 θ_t 就是推移算子. 因此推移算子在轨道空间上总是存在的, 所以我们不妨假设它存在. 这样, 利用单调类方法, (5.3) 可以写成

$$E^x(1_{\{B_{t_n-t}\in A\}}\circ\theta_t|\mathcal{G}_t) = P^x(B_{t_n} \in A|\mathcal{G}_t) = P^{B_t}(B_{t_n-t} \in A). \tag{5.4}$$

类似地, 设 t_j 是现在时间 t, 联合分布 (5.2) 蕴含着

$$\begin{aligned}
&P^x(\{B_{t_n} \in A_n, \cdots, B_{t_{j+1}} \in A_{j+1}\}, \{B_{t_j} \in A_j, \cdots, B_{t_1} \in A_1\}) \\
&\quad = E^x\left[P^{B_t}(B_{t_{j+1}-t} \in A_{j+1}, \cdots, B_{t_n-t} \in A_n)1_{\{B_{t_j}\in A_j,\cdots,B_{t_1}\in A_1\}}\right].
\end{aligned}$$

等式的左边

$$\begin{aligned}
&P^x(B_{t_n} \in A_n, \cdots, B_{t_1} \in A_1) \\
&\quad = E^x\left[1_{\{B_{t_n-t}\in A_n,\cdots,B_{t_{j+1}-t}\in A_{j+1}\}}\circ\theta_t, B_{t_j} \in A_j, \cdots, B_{t_1} \in A_1\right].
\end{aligned}$$

令

$$\begin{aligned}
\Gamma &:= \{B_{t_n-t} \in A_n, \cdots, B_{t_{j+1}-t} \in A_{j+1}\}, \\
A &:= \{B_{t_j} \in A_j, \cdots, B_{t_1} \in A_1\}.
\end{aligned}$$

因此

$$E^x[1_\Gamma\circ\theta_t, A] = E^x[E^{B_t}[1_\Gamma], A]. \tag{5.5}$$

显然上面这样的 A 的全体是生成 $\mathcal{G}_t$ 的一个域, Γ 的全体是生成 $\mathcal{G}_\infty$ 的一个域, 因此, 由单调类方法推出下面的定理, 一个非常简单直观的表达式.

定理 5.1 布朗运动满足简单马氏性: 设 $t \geqslant 0$, Y 是 $\mathcal{G}_\infty$ 可测的有界或非负随机变量, 则 P^x-a.s.

$$E^x[Y\circ\theta_t|\mathcal{G}_t] = E^{B_t}[Y]. \tag{5.6}$$

练习 5.1 证明: 对于上面的 Γ, $P^x(\Gamma)$ 关于 x 连续. 对于上面的 Y, $E^x[Y]$ 关于 x 是 Borel 可测的.

其实, 定理证明并没有用到布朗运动的具体分布性质和轨道性质, 只要函数族 (P_t) 是转移半群即可, 因此该定理实际上可以推广到一般的转移半群情况.

练习 5.2 满足 C-K 方程的 (P_t) 称为转移函数或者转移半群. 实际上, 设 $(\nu_t, t \geqslant 0)$ 是 $\mathbf{R}^d$ 上的概率测度族且 $P_t(x, A) := \nu_t(A - x)$, 证明: (P_t) 是转移半群当且仅当 (ν_t) 满足: 对任何 $t, s \geqslant 0$, 有

$$\nu_t * \nu_s = \nu_{t+s}.$$

5.2 强马氏性

5.2.1 流与停时

本节简单介绍流与停时, 它们是随机过程理论的基础概念与语言. 设有概率空间 $(\Omega, \mathcal{F}, P)$ 和 $T = [0, \infty)$. 回忆 $\mathcal{F}$ 的子 σ- 代数族 $(\mathcal{F}_t : t \in T)$ 称为流, 如果对任何 $s < t$, 有 $\mathcal{F}_s \subset \mathcal{F}_t$. 定义

$$\mathcal{F}_\infty = \sigma(\bigcup_{t \geqslant 0} \mathcal{F}_t).$$

因为加入所有零概率集及其子集到每个 $\mathcal{F}_t$ 中不会影响条件期望, 所以我们总是假设每个 $\mathcal{F}_t$ 中含有零概率集及其子集. 对任何 $t \in T$, 定义 $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$, 那么 $(\mathcal{F}_{t+})$ 也是流. 连续情况下停时的定义类似于离散时间.

定义 5.1 给定概率空间 $(\Omega, \mathcal{F}, P)$ 及流 $(\mathcal{F}_t)$.

1. 映射 $\tau : \Omega \longrightarrow T \cup \{\infty\}$ 称为是一个 $(\mathcal{F}_t)$ 停时, 如果对任何 $t \in T$,

$$\{\omega \in \Omega : \tau(\omega) \leqslant t\} \in \mathcal{F}_t.$$

2. 设 τ 是 $(\mathcal{F}_t)$ 停时, 容易验证 $\mathcal{F}_\infty$ 中使得 $A \cap \{\tau \leqslant t\} \in \mathcal{F}_t$ 对所有 $t \in T$ 成立的 A 全体是一个 σ-代数, 记为 $\mathcal{F}_\tau$.

练习 5.3 证明: $\mathcal{F}_\tau$ 是 σ-代数.

停时是相对于一个流而言的, 但当上下文明确时, 就简单地说一个停时. 特别要注意 $(\mathcal{F}_t)$ 停时与 $(\mathcal{F}_{t+})$ 停时之间的不同. $(\mathcal{F}_t)$ 停时必是 $(\mathcal{F}_{t+})$ 停时. 设 τ 是 $(\mathcal{F}_{t+})$ 停时, 则 $\{\tau < t\} = \bigcup_n \{\tau \leqslant t - \frac{1}{n}\} \in \mathcal{F}_t$. 反之, 对任何 k, $\{\tau \leqslant t\} = \bigcap_{n>k} \{\tau < t + \frac{1}{n}\} \in \mathcal{F}_{t+\frac{1}{k}}$, 故 $\{\tau \leqslant t\} \in \mathcal{F}_{t+}$, 即 τ 是 $(\mathcal{F}_{t+})$ 停时. 因此有下面的引理中的结论 1.

引理 5.2 1. 随机时间 τ 是 $(\mathcal{F}_{t+})$ 停时当且仅当对任何 $t \in T$, $\{\tau < t\} \in \mathcal{F}_t$.

2. 设 τ 是常数 $u \geqslant 0$. 则 $\mathcal{F}_\tau = \mathcal{F}_u$, $\mathcal{F}_{\tau+} = \mathcal{F}_{u+}$.

3. 一个随机变量 ξ 是 $\mathcal{F}_\tau$ 可测当且仅当对任何 $t \in T$, $\xi \cdot 1_{\{\tau \leqslant t\}}$ 是 $\mathcal{F}_t$ 可测的.

4. 设 τ 是 $(\mathcal{F}_{t+})$ 停时, 则 $\mathcal{F}_{\tau+}$ 等于 $\mathcal{F}$ 中使得 $A \cap \{\tau < t\} \in \mathcal{F}_t$ 对所有 $t \in T$ 成立的 A 全体.

练习 5.4 证明上面的引理中的结论 2,3,4.

引理 5.3 设 τ, σ, τ_n 是 $(\mathcal{F}_t)$ 停时.

1. $\tau \vee \sigma$, $\tau \wedge \sigma$ 是停时;

2. 若 τ_n 递增, 则 $\lim \tau_n$ 是 $(\mathcal{F}_t)$ 停时;

3. 如果 $\sigma \leqslant \tau$, 则 $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$;

4. τ 是 $\mathcal{F}_\tau$ 可测的;

5. 如果 τ_n 递减, 则 $\lim \tau_n$ 是 $(\mathcal{F}_{t+})$ 停时, 且 $\bigcap_n \mathcal{F}_{\tau_n} \subset \mathcal{F}_{(\lim \tau_n)+}$;

证明. 5. 由 τ_n 递减推出

$$\{\lim \tau_n < t\} = \bigcup_n \{\tau_n < t\} \in \mathcal{F}_t,$$

故 $\lim \tau_n$ 是 $(\mathcal{F}_{t+})$ 停时. 设 $A \in \bigcap_n \mathcal{F}_{\tau_n}$, 则对 $t \in T$,

$$A \cap \{\lim_n \tau_n < t\} = \bigcup_n (A \cap \{\tau_n < t\}) \in \mathcal{F}_t.$$

即 $A \in \mathcal{F}_{(\lim \tau_n)+}$. □

练习 5.5 证明上面引理的 1,2,3,4.

随机过程中最重要的时间是首中时, 就是过程轨道首次碰到一个集合的时间. 前面我们看到, 离散时间情况下证明首中时是停时几乎是平凡的, 连续时间的情况要复杂得多, 下面的例子考虑开集或者闭集的首中时什么时候是停时.

例 5.1 设 E 是一个度量空间, X 是关于流 $(\mathcal{F}_t)$ 适应的以 E 为状态空间的随机过程. 对 $A\subset E$, $\omega\in\Omega$, 定义 A 的进入时和首中时如下

$$D_A(\omega):=\inf\{t\geqslant 0: X_t(\omega)\in A\};$$
$$T_A(\omega):=\inf\{t>0: X_t(\omega)\in A\}.$$

怎么把 $\{T_A\leqslant t\}$ 或者 $\{T_A<t\}$ 用 $X_s: s\leqslant t$ 中的可数多个来表示呢? 没有一定的轨道性质是做不到的.

我们先证明若 E 是具有度量 d 的度量空间, A 是闭集且 X 是连续过程, 则 D_A 是 $(\mathcal{F}_t)$ 停时. 事实上, 这时由轨道的连续性 $D_A\leqslant t$ 当且仅当

$$\inf_{s\in[0,t]\cap\mathbb{Q}} d(X_s,A)=d(\{X_s: s\in[0,t]\cap\mathbb{Q}\},A)=0.$$

但是 X 的轨道只是几乎处处连续的, 因此 $\{D_A\leqslant t\}$ 与 $\mathcal{F}_t$ 可测集

$$\left\{\inf_{s\in[0,t]\cap\mathbb{Q}} d(X_s,A)=0\right\}$$

只相差一个零概率集, 也就是说如果 $\mathcal{F}_t$ 包含了所有零概率集, 那么 $\{D_A\leqslant t\}\in\mathcal{F}_t$, 即 D_A 是停时. T_A 是不是停时呢? 上面的方法不适用于 T_A. 令

$$\tau_n:=\inf\{t\geqslant 1/n: X_t\in A\}.$$

则 τ_n 递减停时列, 极限是 T_A. 所以由上面引理的 (2) 推出 T_A 是 $(\mathcal{F}_{t+})$ 停时.

练习 5.6 举例说明适用于 D_A 的证明不适用于 T_A. 证明: $\tau_n\downarrow T_A$.

类似地, 若 A 是开集且只要求 X 是右连续的, 那么在右连续的轨道上, $T_A<t$ 当且仅当存在有理数 $s<t$ 使得 $X_s\in A$, 因此若 $\mathcal{F}_t$ 包含所有零概率集, 则 $\{T_A<t\}\in\mathcal{F}_t$, 即 T_A 是 $(\mathcal{F}_{t+})$ 停时, 但一般不是 $(\mathcal{F}_t)$ 停时.

直观地, X 的一条轨道 t 时刻 (包括 t) 前的信息不能告诉我们它是否将立刻进入一个开集. 因为只观察 $[0,t]$ 时段, 过程 X 可能和该开集没有交集, 而如果你再看过去一点, 就看到轨道进入该开集了, 即说明 $\{T_A\leqslant t\}\in\mathcal{F}_{t+}$. 也就是说在 $(\mathcal{F}_t)$ 包含所有零概率集及其子集且右连续的情况下, 我们可以证明闭集的进入时与首中时, 开集的首中时都是停时. 进入时与首中时的区别在于过程的初始位置, 若轨道的起始点不在 A 中, 则 $D_A=T_A$; 若轨道从 A 中的点出发, 则 $D_A=0$ 而 T_A 不一定. 因此, 不管轨道在什么点出发, $T_A<\infty$ 等价于轨道将再回到 A. ∎

从这个例子可以看出, 下面定义中的条件很重要.

定义 5.2 **一个流 $(\mathcal{F}_t)$ 满足通常条件是指它有右连续性且 $\mathcal{F}_0$ 包含有所有零概率集及其子集.**

在连续时间情况下, 假设流满足通常条件且轨道连续, 则开集与闭集的首中时是停时. 因此连续时间过程的首中时是否停时这件事情不是像离散时间时那么简单和平凡. 那么布朗运动的适应流是否满足通常条件呢?

5.2.2 流的强化

回到标准布朗运动 $B = (B_t)$, 自然流 $(\mathcal{G}_t)_{0\leqslant t\leqslant\infty}$. 对任何 $x \in \mathbf{R}$, $\mathcal{N}^x$ 表示 $\mathcal{G}_\infty$ 中的 P^x-零概率集及其子集, 令

$$\mathcal{F}_t^x := \sigma(\mathcal{G}_t \cup \mathcal{N}^x),\ \mathcal{F}_t := \bigcap_x \mathcal{F}_t^x,$$

其中 $t \in [0,\infty]$. 流 $(\mathcal{F}_t^x)$ 与 $(\mathcal{F}_t)$ 被称为自然流的强化. 这其中有很多细节, 我们不打算在这个课上逐一验证, 只需要知道强化就是把零概率集及其子集加入自然流重新生成的流. 前面的马氏性 (5.6) 对流 $(\mathcal{F}_t^x)$ 仍然成立:

$$E^x[Y\circ\theta_t|\mathcal{F}_t^x] = E^{B_t}[Y], \tag{5.7}$$

其中 Y 是 $\mathcal{F}_\infty$ 可测的有界或者非负随机变量.

5.2.3 强马氏性

为了方便, 我们不妨假设布朗运动的所有轨道是连续的. 设 T 是 $(\mathcal{G}_{t+})$ 停时, 即对任何 $t \geqslant 0$, 有 $\{T < t\} \in \mathcal{G}_t$. 回忆

$$\mathcal{G}_{T+} = \{A \in \mathcal{G}_\infty : \forall t \geqslant 0, A \cap \{T < t\} \in \mathcal{G}_t\}.$$

注意当 $T < \infty$ 时推移算子 θ_T 和过程 B_T 都有自然的定义, 当 $T = \infty$ 时, 它们没有定义. 当我们说 B_T 的可测性时, 是指在 $T < \infty$ 时. 下面我们证明简单马氏性 (5.6) 的时间 t 换成 T 之后仍然成立, 称为强马氏性.

定义 5.3 **布朗运动关于一个适应流 $(\mathcal{M}_t)$ 有强马氏性, 是指对任何 $\mathcal{G}_\infty$ 可测的有界或非负随机变量 Y, 对任何 $(\mathcal{M}_t)$ 停时 T, 任何 $x \in \mathbf{R}$, 有**

$$E^x[Y\circ\theta_T|\mathcal{M}_T]1_{\{T<\infty\}} = E^{B_T}[Y]1_{\{T<\infty\}},\ P^x\text{-a.s.}. \tag{5.8}$$

如果布朗运动关于流 $(\mathcal{G}_{t+})$ 有强马氏性, 则说布朗运动有强马氏性.

练习 5.7 取 $A \in \mathcal{B}(\mathbf{R})$, $Y = 1_A(B_t)$, 应用 (5.8), 推出对任何 A, $t \geqslant 0$, $x \in \mathbf{R}$, 有

$$E^x[1_A(B_{t+T})|\mathcal{M}_T]1_{\{T<\infty\}} = E^{B_T}[1_A(B_t)]1_{\{T<\infty\}}.$$

证明: 该性质可以推出布朗运动关于流 $(\mathcal{M}_t)$ 有强马氏性.

定理 5.2 布朗运动有强马氏性.

注意, 强马氏性即使在 T 是固定时间 t 时也比简单马氏性要强, 因为此时有

$$E^x[Y\circ\theta_t|\mathcal{G}_{t+}] = E^{B_t}[Y],\ P^x\text{-a.s.}, \tag{5.9}$$

其左边是关于 $\mathcal{G}_{t+}$ 的条件期望, 这比简单马氏性强一点的性质蕴含强化流的右连续性. 为什么可以强这一点呢? 因为布朗运动的轨道连续性和热核的连续性.

证明. 离散化 T, 对 $n \geqslant 0$, 定义

$$T_n := \begin{cases} k2^{-n}, & (k-1)2^{-n} \leqslant T < k2^{-n}, \\ \infty, & T = \infty. \end{cases}$$

那么 $\{T_n = k2^{-n}\} \in \mathcal{F}_{k2^{-n}}$, 因此 T_n 是停时. 根据下面的习题证明 B_T 是 $\mathcal{G}_{T+}$ 可测的.

练习 5.8 设离散化 T 为 T_n 如上, 则 $T_n \downarrow T$, 处处地有 $B(T_n) \longrightarrow B(T)$.

1. B_{T_n} 关于 $\mathcal{G}_{T_n}$ 可测 (这在离散时间时是显然的);
2. 对任何固定 n, B_T 是 $\mathcal{G}_{T_n}$ 可测的, 推出 B_T 是 $\bigcap_n \mathcal{G}_{T_n}$ 可测的.
3. 证明: $\bigcap_n \mathcal{G}_{T_n} \subset \mathcal{G}_{T+}$.

由此推出 $E^{B_T}[Y]1_{\{T<\infty\}}$ 是 $\mathcal{G}_{T+}$ 可测的, 然后仅需证明对任何 $A \in \mathcal{G}_{T+}$ 有

$$E^x\left[Y\circ\theta_T 1_{\{T<\infty\}\cap A}\right] = E^x\left[E^{B_T}[Y]1_{\{T<\infty\}\cap A}\right]. \tag{5.10}$$

因为 T 是 $(\mathcal{G}_{t+})$ 停时, 故

$$\{T_n = k2^{-n}\} \cap A \in \mathcal{G}_{k2^{-n}}.$$

应用简单马氏性, 以及

$$\{T<\infty\}=\bigcup_{k\geqslant 1}\{T_n=k2^{-n}\},$$

推出

$$\begin{aligned}E^x\left[Y\circ\theta_{T_n}1_{\{T<\infty\}\cap A}\right]&=\sum_k E^x\left[Y\circ\theta_{k2^{-n}}1_{\{T_n=k2^{-n}\}\cap A}\right]\\&=\sum_k E^x\left[E^{B_{k2^{-n}}}[Y]1_{\{T_n=k2^{-n}\}\cap A}\right]\\&=\sum_k E^x\left[E^{B_{T_n}}[Y]1_{\{T_n=k2^{-n}\}\cap A}\right]\\&=E^x\left[E^{B_{T_n}}[Y]\sum_k 1_{\{T_n=k2^{-n}\}\cap A}\right]\\&=E^x\left[E^{B_{T_n}}[Y]1_{\{T<\infty\}\cap A}\right].\end{aligned}$$

取特别的 Y, 例如

$$Y=f(B_{t_1},\cdots,B_{t_j}),$$

其中 $t_i\geqslant 0$, f 是有界连续函数. 现在让 n 趋于无穷, 显然 $T_n\downarrow T$. 因为布朗运动的轨道连续 (右连续即可), 所以 $B_{T_n}\to B_T$ a.s. 因此

$$Y\circ\theta_{T_n}=f(B(t_1+T_n),\cdots,B(t_j+T_n))\longrightarrow Y\circ\theta_T.$$

左边的收敛性没有问题, 右边为什么收敛? 这时需要证明 $E^x[Y]$ 关于 x 连续, 这是之前的一个练习. 对于一般的 Y, 利用单调类方法验证, 留作习题. $\square$

练习 5.9 完成定理证明中对一般 Y 的论证.

强马氏性等价于一个看起来更弱的性质.

练习 5.10 设 Y 是 $\mathcal{G}_\infty$ 可测的有界或非负随机变量. 定理 5.2 成立当且仅当对任何 $(\mathcal{G}_{t+})$ 停时 T, 有 a.s.

$$E^x[Y\circ\theta_T 1_{\{T<\infty\}}]=E^x[E^{B_T}[Y]1_{\{T<\infty\}}].\tag{5.11}$$

我们由此证明强化之后的流 $(\mathcal{F}_t^x)$ (关于概率 P^x) 满足通常条件.

定理 5.3 对任何 $x\in\mathbf{R}$, $t\geqslant 0$, $\mathcal{F}_{t+}^x=\mathcal{F}_t^x$,

证明. 把强马氏性用于固定时间 t 得

$$E^x[Y\circ\theta_t|\mathcal{G}_{t+}] = E^{B_t}[Y],\ \ P^x\text{-a.s.}$$

取 $Y = f_1(B_{t_1})\cdots f_n(B_{t_n})$, 其中 $f_1, \cdots, f_n$ 有界可测. 立刻推出

$$E^x[Y|\mathcal{G}_{t+}] = E^x[Y|\mathcal{G}_t],\ \ a.s.$$

然后由单调类方法推出对任何有界且关于 $\mathcal{G}_\infty$ 可测的 Y 成立. 特别地取 $A \in \mathcal{G}_{t+}$ 及 $Y = 1_A$, 有

$$1_A = E^x[1_A|\mathcal{G}_{t+}] = E^x[1_A|\mathcal{G}_t],\ \ P^x\text{-a.s.},$$

即对任何 x, A 与 $\mathcal{G}_t$ 中的一个可测集的对称差是 P^x-零概率集, 所以 $\mathcal{G}_{t+} \subset \mathcal{F}_t^x$, 因此 $\mathcal{F}_{t+}^x \subset \mathcal{F}_t^x$. □

练习 5.11 令 $\xi = E^x[1_A|\mathcal{G}_t]$, $B = \{\xi = 1\}$, 证明: $A\triangle B$ 是 P^x-零概率集.

因此流 $(\mathcal{F}_t^x)$ 关于测度 P^x 满足通常条件, 也推出流 $(\mathcal{F}_t)$ 是右连续的. 下面定理的证明留给读者.

定理 5.4 **布朗运动关于流 $(\mathcal{F}_t)$ 有强马氏性.**

现在设 T 是布朗运动对于任何一个开集或者闭集的首中时, 因为对任何 $x \in \mathbf{R}$, $(\mathcal{F}_t^x)$ 满足通常条件, 所以 T 是 $(\mathcal{F}_t^x)$ 停时, 即对任何 $t \geqslant 0$, $\{T \leqslant t\} \in \mathcal{F}_t^x$, 因此 T 也是 $(\mathcal{F}_t)$ 停时.

最后一个结论是关于 $\mathcal{F}_0$ 的. 因为 $P^x(B_0 = x) = 1$, 所以 B_0 在任何 P^x 之下是个常数, 即 $\mathcal{G}_0 = \sigma(B_0)$ 是平凡的 (其中事件的概率非 0 即 1). 现在设 $A \in \mathcal{F}_0$, 即对任何 x, A 与 $\mathcal{G}_0$ 中的某个事件相差一个 P^x-零事件, 因此有下面的 Blumenthal 0-1 律.

定理 5.5 (Blumenthal) **设 $A \in \mathcal{F}_0$, 则对任何 $x \in \mathbf{R}$, $P^x(A)$ 非 0 即 1.**

练习 5.12 设 B 是一般的布朗运动, 即几乎所有轨道连续. 这时对于自然流来说, 简单马氏性仍然成立, 但是强马氏性不再成立. 对于强化自然流才有强马氏性. 证明: 对任何 $x \in \mathbf{R}$, 任何 $(\mathcal{F}_{t+}^x)$ 停时 T, $\mathcal{G}_\infty$ 可测的有界或非负随机变量 Y 有 B_T 是 $\mathcal{F}_{T+}^x$ 可测, 且

$$E^x[Y\circ\theta_T|\mathcal{F}_{T+}^x]1_{\{T<\infty\}} = E^{B_T}[Y]1_{\{T<\infty\}},\ \ P^x\text{-a.s.}.$$

由此推出流 $(\mathcal{F}_t^x)$ 与 $(\mathcal{F}_t)$ 是右连续的且布朗运动关于 $(\mathcal{F}_t)$ 有强马氏性.

5.2.4 回顾反射原理

由定理 5.2 的强马氏性推出, 如果 T 是 $(\mathcal{F}_t)$-停时, 则对任何 $\mathcal{F}_\infty$ 可测的有界或非负随机变量 Y, 以及 $[0,\infty)$ 上有界或者非负 Borel 可测函数 f 有

$$E^x(Y\circ\theta_T\cdot f(T);T<\infty)=E^x(E^{B(T)}(Y)f(T);T<\infty). \tag{SM}$$

应用单调类方法推出: 如果 φ 是 $[0,\infty)\times\Omega$ 上关于 $\mathcal{B}[0,\infty)\times\mathcal{F}_\infty$ 可测的有界或非负函数, 即一个联合可测的随机过程, 那么

$$E^x\left(\varphi(T,\theta_T(\cdot));T<\infty\right)=E^x\left[\left(E^y\varphi(t,\cdot)\right)\Big|_{t=T,y=B(T)};T<\infty\right] \qquad (SM').$$

证明留作习题.

练习 5.13 说明为什么 (SM) 是对的? 怎么推出 (SM')?

强 Markov 性是非常直观且有用的性质, 下面我们给出它的一些应用. 首先我们将证明重要的反射原理. 下面我们记 $P=P^0$, T_b 为点 $b>0$ 的首中时.

定理 5.6 (Reflection principle) **设** $b>0$, $t>0$, **则**

$$P(T_b<t)=2P(B_t>b).$$

证明. 写布朗运动为 $B(t,\omega)=B_t(\omega)$. 怎么找到合适的随机过程使得我们可以应用强马氏性来得到反射原理呢? 令

$$\varphi(s,\omega):=1_{\{s<t,B(t-s,\omega)>b\}}.$$

那么

$$\varphi(T_b,\theta_{T_b}(\omega))=1_{\{T_b<t,B(t-T_b,\theta_{T_b}(\omega))>b\}}=1_{\{T_b<t,B_t>b\}}.$$

利用强 Markov 性 (SM')

$$\begin{aligned}P(B_t>b)&=P(T_b<t,B_t>b)\\&=E\left[E^y(\varphi(s,\cdot))\Big|_{s=T_b,y=B_{T_b}},T_b<t\right]\\&=E\left[P^b(B_{t-s}>b)\Big|_{s=T_b},T_b<t\right]\end{aligned}$$

$$= \frac{1}{2} P(T_b < t),$$

第三个等号是因为布朗运动的对称性推出 $P^b(B_{t-s} > b) = \frac{1}{2}$. 证毕. □

下面计算考虑 T_0. 在此之前, 我们需要首中时关于时间推移的性质, 后面会经常用到.

引理 5.4 **设** $T = T_A$ **是集合** A **的首中时, 则**

$$T \circ \theta_s + s = \inf\{t > s : B_t \in A\},$$

因此当 $T > s$ **时,** $T \circ \theta_s + s = T$ **且** $B_T \circ \theta_s = B_T$**, 另外, 当** $s \downarrow 0$ **时,** $T \circ \theta_s + s \downarrow T$.

结论很直观, 时间 s 后首次碰到 A 的时间等于从 s 时间开始算首次碰到 A 的时间加上 s. 所以如果 $[0, s]$ 区间内没有碰到 A, 那么时间 s 后首次碰到就是首次碰到. 如果在首次碰到 A 的时间之前开始看首次碰到的位置是不会变的. 上面的结论是对样本轨道而言的, 因此也就是相当于关于连续函数的结论.

证明. 由定义

$$\begin{aligned} T \circ \theta_s + s &= \inf\{t > 0 : B_{t+s} \in A\} + s \\ &= \inf\{t + s > s : B_{t+s} \in A\} \\ &= \inf\{t > s : B_t \in A\}. \end{aligned}$$

因此, 当 $T > s$ 时, $T \circ \theta_s + s = T$ 且

$$B_T \circ \theta_s = B_{T \circ \theta_s + s} = B_T.$$

完成证明. □

定理 5.7 $P(T = 0) = 1$

证明. 由 T_b 的 Laplace 变换推出对任何 $x \neq b$,

$$E^x e^{-sT_b} = e^{-\sqrt{2s}|x-b|}.$$

令 $T = T_0$. 对 $t > 0$, 用 Markov 性, 因为 $B_t \neq 0$ a.s.,

$$E\left[e^{-s(t+T\circ\theta_t)}\right] = e^{-st} E\left[E^{B_t} e^{-sT}\right] = e^{-st} E\left[e^{-\sqrt{2s}|B_t|}\right].$$

让 $t \downarrow 0$, 得 $E\left[e^{-sT}\right] = 1$, 因此 $P(T = 0) = 1$. □

另一个方法: 对任何 $t > 0$,

$$P(T_{[0,\infty)} \leqslant t) \geqslant P(B_t \geqslant 0) = \frac{1}{2}.$$

由 Blumenthal 0-1 律, $P(T_{[0,\infty)} = 0) = 1$. 同理 $P(T_{(-\infty,0]} = 0) = 1$. 由连续性推出 $P(T = 0) = 1$.

定理 5.8 设 B 是 1-维标准布朗运动. 对几乎所有的 ω, 轨道的零点集 $\{t : B_t(\omega) = 0\}$ 是一个 Lebesgue 测度为零的拓扑 Cantor 集, 精确地说, 它是无处稠的且无孤立点的闭集.

证明. 零点集是闭是显然的事实, 对任何 x,

$$E\big|\{t \geqslant 0 : B_t = x\}\big| = E\int_0^\infty 1_{\{B_t = x\}} dt = \int_0^\infty P(B_t = x) = 0,$$

推出零点集的 Lebesgue 测度为零 a.s., 因此它也一定是无处稠的 a.s.. 现在证明没有孤立的零点. 固定 $0 < s < t$, 令 $T_s := s + T\circ\theta_s$, 如果 B 在 (s,t) 间恰有一个零点, 那么 $t > T_s$ 表示 (s,t) 间至少有一个零点, 且 $T\circ\theta_{T_s} > 0$ 表示 T_s 后不能再马上命中 0, 因此若用 $N_{s,t}$ 表示事件: B 在 (s,t) 仅有一个零点, 则

$$P(N_{s,t}) \leqslant P(T_s < t, T\circ\theta_{T_s} > 0) = E(P^{B_{T_s}}(T > 0); T_s < t) \leqslant P(T > 0) = 0.$$

因为有孤立零点的轨道必定存在有理数端点的区间 (s,t) 使得其中只有唯一零点, 所以推出几乎所有轨道没有孤立零点. 证毕. □

练习 5.14 证明过程 $a \mapsto T_a$, $a \geqslant 0$, 是一个递增的左连续的平稳独立增量过程.

5.3 经典位势

我们证明了布朗运动是以热核作为转移函数的连续轨道的强马氏过程. 下面我们将研究布朗运动与位势理论间的关系.

5.3.1 暂留与常返

下面我们来讨论布朗运动的极集与常返集.

定义 5.4 设 A 是 Borel 集.

1. A 被称为极集, 如果对任何 $x \in E$, 有 $P^x(T_A < \infty) = 0$.

2. A 称为是常返的, 如果对任何 x 有 $P^x(T_A < \infty) = 1$; 否则称为暂留的.

练习 5.15 证明: 可数多个极集的并是极集.

另外一个有用的时间是末离时, 对任何 Borel 集, 定义末离时

$$L_A := \sup\{t \geqslant 0 : B_t \in A\}$$

(注意 $\sup\varnothing = 0$). 顾名思义, 它是布朗运动最后离开 A 的时间, 它不是一个首中时, 因为一个时间是不是末离要看之后的轨道, 而不是看之前的轨道. 显然 $L_A = \infty$ 当且仅当对任何 $t > 0$, $T_A \circ \theta_t < \infty$.

练习 5.16 证明: A 是常返的当且仅当 $P^x(L_A = \infty) = 1$.

也就是说极集是几乎所有轨道都不会遇到的集合, 而常返是几乎所有轨道一定会遇到的集合. 常返的概念类似于 Markov 链中所定义的. 首先上面 T_a 的 Laplace 变换的表达式说明 $P(T_a < \infty) = 1$, 即证明了一维布朗运动的单点集因此任何非空集都是常返的. 但是二维及以上的布朗运动的单点集却是极集. 这主要是因为

$$\int_0^1 p(t,0)dt = \int_0^1 \frac{1}{(2\pi t)^{\frac{d}{2}}}dt = \begin{cases} < \infty, & d = 1, \\ = \infty, & d \geqslant 2. \end{cases}$$

另外容易验证

$$\int_0^\infty p(t,x)dt = \int_0^\infty \frac{1}{(2\pi t)^{\frac{d}{2}}} e^{-\frac{|x|^2}{2t}}dt = \begin{cases} +\infty, & d = 1,2, \\ \frac{C(d)}{|x|^{d-2}}, & d \geqslant 3, \end{cases}$$

其中 $C(d)$ 是只与 d 有关的常数. 这决定了在 $d \leqslant 2$ 时紧集是常返的而 $d \geqslant 3$ 时不是常返的.

定理 5.9 设 B 是 d-维布朗运动.

(1) 当 $d \geqslant 2$ 时, 单点集是极集.

(2) 当 $d = 2$ 时, 不是极集的 Borel 集一定是常返的. 因此任何开集是常返的.

(3) 当 $d \geqslant 3$ 时, 紧集不是常返的.

证明. (1) 设 T_0 是 0 的首中时, 对 $M > s > 0$, 令 $T^{(s)} := s + T_0 \circ \theta_s$. 那么 $T^{(s)}$ 是停时. 用强 Markov 性, 对任何 $x \in \mathbf{R}^d$ 和以原点为中心 r 为半径的球 G_r,

$$\begin{aligned}
&E^x\left[\left(\int_0^1 1_{G_r}(B_t)dt\right)\circ\theta_{T^{(s)}}, T^{(s)} \leqslant M\right]\\
&= E^x\left(E^{B(T^{(s)})}\left(\int_0^1 1_{G_r}(B_t)dt\right), T^{(s)} \leqslant M\right)\\
&= P^x(T^{(s)} \leqslant M)\int_{G_r} dy \int_0^1 p(t,y)dt.
\end{aligned}$$

另一方面,

$$\begin{aligned}
&E^x\left[\left(\int_0^1 1_{G_r}(B_t)dt\right)\circ\theta_{T^{(s)}}, T^{(s)} \leqslant M\right]\\
&= E^x\left(\int_{T^{(s)}}^{1+T^{(s)}} 1_{G_r}(B_t)dt, T^{(s)} \leqslant M\right) \leqslant E^x\left(\int_s^{M+1} 1_{G_r}(B_t)dt\right)\\
&= \int_{G_r} dy \int_s^{M+1} p(t, y-x)dt \leqslant |G_r| \cdot \frac{M+1-s}{(2\pi s)^{\frac{d}{2}}},
\end{aligned}$$

其中 $|G_r|$ 表示 G_r 的体积. 因此对任何 $\varepsilon > 0$ 有

$$P^x(T^{(s)} \leqslant M)\cdot\frac{1}{|G_r|}\int_{G_r} dy\int_\varepsilon^1 p(t,y)dt \leqslant \frac{M+1-s}{(2\pi s)^{\frac{d}{2}}}.$$

先让 $r \to 0$, 得

$$P^x(T^{(s)} \leqslant M)\cdot\int_\varepsilon^1 p(t,0)dt \leqslant \frac{M+1-s}{(2\pi s)^{\frac{d}{2}}}.$$

因为 $\int_0^1 p(t,0)dt = \infty$, 故 $P^x(T^{(s)} \leqslant M) = 0$. 让 $M \to \infty$, $s \to 0$, 便推出 $P^x(T < \infty) = 0$.

(2) 设 $d \leqslant 2$. 设 $f(x) := P^x(T_A < \infty)$. 则 $P_t f = P^{\cdot}(T_A \circ \theta_t < \infty) \leqslant f$ 且 $P_t f \uparrow f$. 注意 $t > 0$ 时 $P_t f$ 连续. 对任何 $n \geqslant 1$, $t > s > 0$, $x \in \mathbf{R}^d$,

$$\begin{aligned}
\int_{\mathbf{R}^2}(P_s f(y) - P_t f(y))dy\int_0^n p(u, y-x)du &= \int_0^n P_u(P_s f - P_t f)(x)du\\
&= \int_s^{n+s} P_u f(x)du - \int_t^{n+t} P_u f(x)du\\
&= \int_s^t P_u f(x)du - \int_{n+s}^{n+t} P_u f(x)du \leqslant t - s,
\end{aligned}$$

让 $n \to \infty$, 因为当 $d \leqslant 2$ 时有 $\int_0^\infty p(u, y-x)du = +\infty$, 故点点地有 $P_s f = P_t f$. 因此推出 $P_t f = P_s f = f$. 另外, 对任何 $x, y \in \mathbf{R}^d$, 有

$$\begin{aligned}|f(x) - f(y)| &= |P_t f(x) - P_t f(y)| \\ &\leqslant \int |p(t, y-z) - p(t, x-z)| dz \\ &= \int \left| p(1, z) - p(1, z - \frac{x-y}{\sqrt{t}}) \right| dz.\end{aligned}$$

让 t 趋于无穷, 应用控制收敛定理推出上式趋于零 (严格证明留作习题), 因此 f 是常数.

练习 5.17 证明: 对任何 $x \in \mathbf{R}^d$,

$$\lim_{t\to\infty} \int_{\mathbf{R}^d} \left| p(1, z) - p(1, z - \frac{x}{\sqrt{t}}) \right| dz = 0.$$

现在对 $t > 0$, 由 Markov 性

$$\begin{aligned}P^y(t < T_A < \infty) &= P^y(T_A \circ \theta_t < \infty, t < T_A) \\ &= E^y(f(B_t), t < T_A) = P^y(t < T_A) \cdot f.\end{aligned}$$

让 $t \longrightarrow \infty$ 得 $0 = (1-f)f$, 因此 f 恒为 1 或恒为零, 即 A 是常返或是极集.

(3) 当 $d \geqslant 3$ 时, $g := \int_0^\infty p(t, \cdot)dt$ 是一个严格正局部可积函数. 取单位球 G 及紧集 K, 则由强 Markov 性, 对任何 $x \in \mathbf{R}^d$,

$$\begin{aligned}E^x \int_0^\infty 1_G(B_s)ds &\geqslant E^x \left(\int_{T_K}^\infty 1_G(B_s)ds, T_K < \infty \right) \\ &= E^x \left(E^{B(T_K)} \int_0^\infty 1_G(B_s)ds, T_K < \infty \right) \\ &\geqslant P^x(T_K < \infty) \inf_{z\in K} \int_{G-z} g(y)dy.\end{aligned}$$

函数 $\int_{G-z} g(y)dy$ 是正连续函数, 所以在紧集 K 上的最小值为正, 记为 c. 因此推出

$$\lim_{x\to\infty} P^x(T_K < \infty) \leqslant c^{-1} \lim_{x\to\infty} \int_{G-x} g(y)dy = 0,$$

即 K 不是常返的. 我们还可以得到另外一种表述. 将上面不等式两边的 x 代换为 B_t, 再取期望得

$$E^x \left[E^{B_t} \int_0^\infty 1_G(B_s)ds \right] \geqslant cE^x \left[P^{B_t}(T_k < \infty) \right].$$

应用马氏性推出

$$cP^x(T_K\circ\theta_t<\infty)\leqslant E^x\int_t^\infty 1_G(B_s)ds.$$

容易验证, 当 $t\to\infty$ 时, 右边因为暇积分收敛而趋于零, 所以

$$\lim_{t\to\infty}P^x(T_K\circ\theta_t<\infty)=0,$$

即 $P^x(L_K=\infty)=0$, 无论从任何点出发, 布朗运动从某个时间开始就会离开紧集不再回来. □

练习 5.18 证明:

$$\bigcap_n\{L_{G_n}<\infty\}\subset\{\{\lim_{t\to\infty}|B_t|=\infty\},$$

然后证明: 当 $d\geqslant 3$ 时, 对任何 $x\in\mathbf{R}^d$, P^x-a.s., $\lim_{t\uparrow+\infty}B_t=\infty$ a.s.

5.3.2 Dirichlet 问题的概率解

布朗运动实际上是经典的 Newton 位势理论的另一种表述. 下面我们给出经典的 Dirichlet 问题的概率解, 这是由 Kakutani(1944) 和 Doob(1954) 发现的.

对于一个开集或者闭集 A, 首中时 T_A 是停时, 推出 $\{T_A=0\}\in\mathcal{F}_0$, 因此,

$$P^x(T_A=0)=0 \text{ 或 } 1.$$

定义 5.5 如果 $P^x(T_A=0)=1$, 称 x 为 A 的正则点, 否则为非正则点. 如果 A 的所有点关于 A 是正则的, 称 A 是正则的.

由轨道连续性推出, 如果 A 是闭集, $x\notin A$, 那么 $P^x(T_A>0)=1$, 即 x 是 A 的非正则点, 但是开集 A^c 的正则点. 对于一维布朗运动, 0 是它自己 $\{0\}$ 的正则点, 二维以上就不是了.

定义 5.6 区域 $D\subset\mathbf{R}^d$ 上的函数 u 称为在 D 上调和, 如果 u 是局部可积且球面平均性质成立, 即对任何 $x\in D$ 存在 $\varepsilon>0$ 当 $r<\varepsilon$ 时有

$$u(x)=\int_{S_r(x)}u(y)\sigma_{x,r}(dy),$$

其中 $\sigma_{x,r}$ 是球面 $S_r(x)$ 上的均匀分布.

一个已知的事实是 u 在 D 上调和当且仅当 u 无穷次可微且 $\triangle u = 0$. 所谓的 Dirichlet 问题是这样的: 给定边界 ∂D 上的一个连续函数 f, 上面条件下存在一个在 $\overline{D}$ 上连续在 D 上调和的函数 u 使得 $u|_{\partial D} = f$? 当然 Dirichlet 问题在分析中已经有非常丰富深刻的结果. 我们在这里给出一个概率解只是为了说明布朗运动与位势之间的本质关系. 下面的引理的这个关系的重要表述.

引理 5.5 设 $\tau := \inf\{t > 0 : |B_t - B_0| \geqslant r\}$. **则** B_τ **在** P^x **下的分布是球面** $S_r(x)$ **上均匀分布**.

证明. 不妨设 $x = 0$. 用 O 表示 $\mathbf{R}^d$ 上正交变换, $\widetilde{B}_t = OB_t$. 那么 $\widetilde{B}$ 也是标准布朗运动. 显然 $\mathbf{R}^d$ 上一个零点出发的连续函数与经过旋转得到的函数达到单位球面的时间是一致的, 位置是旋转得到的. 因此 B 首次到达球面的时间与 $\widetilde{B}$ 首次到达的时间一致, 位置 $\widetilde{B}_\tau = O(B_\tau)$. 现在, 对任何 $f \in C(S_r(0))$,

$$E[f(B_\tau)] = E[f(\widetilde{B}_\tau)] = E[f(OB_\tau)],$$

设 B_τ 的分布是 σ, 则 $\sigma(f) = \sigma(fO)$, 即 $\sigma = \sigma \circ O^{-1}$, 说明 σ 是正交变换下不变的, 因此只能是均匀分布. □

定理 5.10 **设区域** $D \subset \mathbf{R}^d$ **是有界的且** D^c **是正则的. 则上述** Dirichlet **问题有唯一解**

$$u(x) = E^x f(B_T),\ x \in \overline{D}, \tag{5.12}$$

其中 T **是** D^c **的首中时.**

证明. 因 D 有界, 故 $P^x(T < \infty) = 1$. 由连续性推出 $B_T \in \partial D$, 故等式 (5.12) 右边是有意义的.

首先唯一性是容易的. 事实上, 因 u 在有界闭集 $\overline{D}$ 上连续, 故达到最大最小值. 而调和函数球面平均性质推出 u 不可能在 D 上达到最大最小值, 只可能在边界上达到. 因此唯一性成立.

下面分三步证明 (5.12) 中定义的函数 u 是 Dirichlet 问题的解.

(1) 验证 u 在 ∂D 上是 f. 因为 D^c 是正则的, 故对 $x \in \partial D$, $u(x) = E^x f(B_T) = E^x f(B_0) = f(x)$.

(2) 验证 u 在 D 上调和. 对任何 $x \in D$, 存在 $\varepsilon > 0$ 使得以 x 为中心, ε 为半径的闭球包含在 D 中. 对任何 $0 < r < \varepsilon$, 令

$$\tau := \inf\{t > 0 : |B_t - B_0| \geqslant r\}.$$

自然 P^x a.s. 有 $\tau < \infty$ 且 B_τ 在球面 $S_r(x)$ 上. 因为标准布朗运动是旋转不变的, 且球面是旋转不变的, 故推出在 P^x 之下, B_τ 是 $S_r(x)$ 上均匀分布的.
直观地, 从 x 点出发, 在遇到 D^c 前一定会先遇到 $S_r(x)^c$, 而且是 $T > \tau$ a.s. P^x, 因此推出 P^x a.s. 有 $T \circ \theta_\tau + \tau = T$. 由强 Markov 性,

$$\begin{aligned}\int_{S_r(x)} u(y)\sigma_{x,r}(dy) &= E^x u(B_\tau) = E^x E^{B_\tau} f(B_T) \\ &= E^x f(B_T)\circ\theta_\tau = E^x f(B(T\circ\theta_\tau + \tau)) = E^x f(B_T) = u(x).\end{aligned}$$

因此 u 在 D 上调和.
(3) 验证 u 在 $\overline{D}$ 上连续. 只需验证在 ∂D 上连续. 任取 $a \in \partial D$, 要证明

$$\lim_{x\to a, x\in D} E^x f(B_T) = f(a).$$

因为 f 是在 ∂D 上其实是一致连续的, 故只需验证对任何 $\varepsilon > 0$ 有

$$\lim_{x\to a} P^x(|B_T - a| > \varepsilon) = 0.$$

而这等同于 $\lim\limits_{x\to a} P^x(|B_T - x| > \varepsilon) = 0$. 任取 $t > 0$, 用 $\tau_\varepsilon(x)$ 表示布朗运动从 x 出发首次离开半径为 ε 的球的时间,

$$\begin{aligned}P^x(|B_T - x| > \varepsilon) &= P^x(|B_T - x| > \varepsilon, T < t) + P^x(|B_T - x| > \varepsilon, T \geqslant t) \\ &\leqslant P^x(\tau_\varepsilon(x) < t) + P^x(T \geqslant t) = P(\tau_\varepsilon(0) \leqslant t) + P^x(T \geqslant t).\end{aligned}$$

显然 $P(\tau_\varepsilon(0) = 0) = 0$, $P^a(T > 0) = 0$. 对任何固定的 $t > 0$, 只要 $x \mapsto P^x(T \geqslant t)$ 是上半连续的, 便推出 $\limsup_{x\to a} P^x(T \geqslant t) \leqslant P^a(T \geqslant t) = 0$, 再让 t 趋于零, 得

$$\limsup_{x\to a} P^x(|B_T - x| > \varepsilon) \leqslant \lim_{t\to 0} P(\tau_\varepsilon(0) < t) = P(\tau_\varepsilon(0) = 0) = 0.$$

故最后我们需要证明 $x \mapsto P^x(T \geqslant t)$ 上半连续. 事实上, 由马氏性,

$$P^x(T\circ\theta_{1/n} + 1/n \geqslant t) = E^x[g_n(B_{1/n})]$$

是关于 x 的连续函数, 其中 $g_n(x) = P^x(T \geqslant t - 1/n)$. 而当 $n \downarrow \infty$ 时,

$$P^x(T \circ \theta_{1/n} + 1/n \geqslant t) \downarrow P^x(T \geqslant t),$$

故 $P^\cdot(T \geqslant t)$ 是上半连续的. $\square$

练习 5.19 给定单位向量 $x \in \mathbf{R}^d$, $r > 0$ 以及 $\theta \in (0, \pi/2)$, 定义

$$C(x, r, \theta) := \{y \in \mathbf{R}^d : |y| < r, \ \angle xy < \theta\},$$

其中 $\angle xy$ 表示两个向量的夹角. 它被称为是中心轴为 x 且伸展角为 θ 的锥, 一个区域 D 满足锥条件, 如果对任何 $y \in \partial D$, 存在如上的 x, r, θ 使得

$$y + C(x, r, \theta) \subset D^c.$$

证明: 如果 D 是满足锥条件的有界区域, 则它是正则的.

习　题

1. 设 $b \geqslant a > 0$, $S_t := \sup_{s\in[0,t]} B_s$. 证明:

$$P(S_t > b, B_t < a) = P(B_t < a - 2b) = P^{2b}(B_t < a),$$

   并求 (B_t, S_t) 的联合密度.

2. 设 $(t, \omega) \mapsto X(t, \omega)$ 是 $[0, \infty) \times \Omega$ 上的联合可测函数, τ 是 Ω 上独立于 $\{X(t, \omega) : t \geqslant 0\}$ 的非负随机变量. 证明:

$$E[X(\tau, \cdot)] = E\left[E(X(t, \cdot))\Big|_{t=\tau}\right].$$

3. 设 $B_t = (B_t^{(1)}, B_t^{(2)})$ 是零点出发的标准 2-维布朗运动. 对 $a \geqslant 0$, 用 T_a^1 表示 $(B_t^{(1)})$ 首中 a 的时间, 证明: $X(a) := B^{(2)}(T_a^1)$, $a \geqslant 0$ 是独立增量过程. 求 X_a 的分布密度与特征函数.

4. (*) 设 B 是标准布朗运动. 证明: 过程 $(S_t - B_t)$ 与 $(|B_t|)$ 的分布相同.

5. 证明: 如果 $B = (B_t)$ 适应于流 $(\mathcal{M}_t)$, T 是 $(\mathcal{M}_t)$ 停时, 那么 B_T 关于 $\mathcal{M}_T$ 可测.

证明. 分三步证明. 首先证明任何 $t > 0$, $B = B(s, \omega)$ 是

$$([0, t] \times \Omega, \mathcal{B}[0, t] \times \mathcal{M}_t)$$

上可测函数. 令

$$B^{(n)}(s, \omega) := B(0, \omega)1_{\{0\}}(s) + \sum_{k=1}^{n} B(\frac{kt}{n}, \omega)1_{(\frac{(k-1)t}{n}, \frac{kt}{n}]}(s).$$

显然, 因为适应, 故 $B^{(n)}$ 关于上述乘积 σ-代数可测. 再由布朗运动的轨道连续性, $B^{(n)}$ 点点收敛于 B. 因此关于 B 的结论成立.

其次, 映射 $f : \omega \mapsto (T(\omega) \wedge t, \omega)$ 是 $(\Omega, \mathcal{M}_t)$ 到 $([0, t] \times \Omega, \mathcal{B}[0, t] \times \mathcal{M}_t)$ 的可测映射, 而

$$B_{T \wedge t} = B \circ f$$

作为两个可测映射的复合是 $(\Omega, \mathcal{M}_t)$ 上的可测函数.

最后, 要验证对任何 $x \in \mathbf{R}$, 有 $\{B_T \leqslant x\} \in \mathcal{M}_T$. 这等价于对任何 $t \geqslant 0$ 有

$$\{B_T 1_{\{T < \infty\}} \leqslant x\} \cap \{T \leqslant t\} \in \mathcal{M}_t.$$

事实上, 这几乎已经是显然的了

$$\{B_T 1_{\{T < \infty\}} \leqslant x\} \cap \{T \leqslant t\} = \{B_{T \wedge t} \leqslant x\} \cap \{T \leqslant t\} \in \mathcal{M}_t.$$

完成证明. □

注意, 上面的证明是假设布朗运动的所有轨道都连续的条件下. 如果布朗运动的轨道是几乎处处连续的, 那么我们需要在 $(\mathcal{G}_t)$ 中加入适当的零概率集才能保证 B_T 关于 $\mathcal{G}_{T+}$ 的可测性. 这也是前面所说的强化的必要性.

6. 证明: 过程 $\{T_a : a \geqslant 0\}$ 是左连续的, 有平稳独立增量在任何区间上都不连续的增过程. 如果定义 $T_{a+} := \lim_{b \downarrow a} T_b$, 证明: 对任何 $a \geqslant 0$, $T_a = T_{a+}$ a.s.

7. 证明:
$$\lim_{t \to +\infty} \sqrt{t} P(B_s \leqslant 1, \forall s \leqslant t) = \sqrt{\frac{2}{\pi}}.$$

8. 证明: 集合 A 的正则点一定是 A 的 Euclid 意义下的极限点. 对 1-维布朗运动, 反之亦然.

9. 设 T 是集合 A 的首中时. 证明: 对任何 $t \geqslant 0$, 当 $T > t$ 时有 $T \circ \theta_t + t = T$. 由此证明定理 5.10 中的用到的一个断言: $T \circ \theta_\tau + \tau = T$ a.s.

第六章　布朗运动与鞅

假设 $(\Omega, \mathcal{F}, P)$ 是概率空间, $(\mathcal{F}_t)$ 是 σ-代数流. $B = (B_t)$ 是 $(\mathcal{F}_t)$ 标准布朗运动, 即 B 是 $(\mathcal{F}_t)$-适应的连续实值随机过程, $B_0 = 0$, 且对任何 $t > s \geqslant 0$ 有 $B_t - B_s$ 独立于 $\mathcal{F}_s$ 且分布为 $N(0, t-s)$. 不妨假设流满足通常条件, 因为这总是可以做到的.

6.1　连续时间鞅

定义 6.1　设 $M = (M_t)$ 是一个可积的关于流 $(\mathcal{F}_t)$ 适应的右连续过程. 如果对任何 $t > s \geqslant 0$ 有

$$E[M_t|\mathcal{F}_s] = M_s,$$

那么称 M 是一个鞅; 如果等号用 $\geqslant$ (对应地, $\leqslant$) 代替, 称为下鞅 (对应地, 上鞅).

在概率 P 之下, 对任何 $t > s \geqslant 0$, 有 $B_t - B_s$ 与 $\mathcal{F}_s$ 独立. 因此, 不难验证, 下面的过程是鞅:

$$B_t,\ t \geqslant 0;$$
$$B_t^2 - t,\ t \geqslant 0;$$
$$e^{aB_t - a^2 t/2},\ t \geqslant 0,$$

其中 a 可以是任何复数. 第三个鞅称为布朗运动的指数鞅. 实际上, 其他两个鞅都来自于指数鞅. 令

$$f(a) = e^{aB_t - a^2 t/2} = \sum_{n \geqslant 0} \frac{f^{(n)}(0)}{n!} a^n.$$

那么 $f^{(0)}(0) = 1$, $f^{(1)}(0) = B_t$, $\cdots$.

6.1.1 一致可积条件

为了进一步阐述极限与期望的交换, 我们引入一致可积的条件, 比控制收敛定理更强大.

定义 6.2　一个可积随机变量族 $\{\xi_i : i \in I\}$ 称为是一致可积的, 若

$$\lim_{N\to\infty} \sup_I E(|\xi_i|; |\xi_i| \geqslant N) = 0.$$

若 $\{\xi_i : i \in I\}$ 被一个可积随机变量所控制, 则 $\{\xi_i\}$ 是一致可积的. 下面定理给出一致可积的一个等价条件, 先证明一个引理, 说明可积随机变量是绝对连续的.

引理 6.1　如果随机变量 ξ 可积, 则对任何 $\varepsilon > 0$, 存在 $\delta > 0$, 使得对任何 $A \in \mathcal{F}$, $P(A) < \delta$, 有 $E(|\xi|; A) < \varepsilon$.

证明. 因为

$$\begin{aligned} E(|\xi|; A) &= E(|\xi|; A, |\xi| \geqslant N) + E(|\xi|; A, |\xi| < N) \\ &\leqslant E(|\xi|; |\xi| \geqslant N) + N \cdot P(A), \end{aligned}$$

故对任何 $\varepsilon > 0$, 存在 N 使得 $E(|\xi|; |\xi| \geqslant N) < \frac{1}{2}\varepsilon$, 再取 $0 < \delta < \frac{1}{2N}\varepsilon$, 因此只要 $P(A) < \delta$, 就有 $E(|\xi|; A) < \varepsilon$. □

定理 6.1　设 $\{\xi_i : i \in I\}$ 是随机变量族. 则它是一致可积的充要条件是

(1) L^1- **有界**: $\sup_{i\in I} E|\xi_i| < \infty$;

(2) **一致绝对连续: 对任何 $\varepsilon > 0$, 存在 $\delta > 0$ 使得当 $A \in \mathcal{F}$, $P(A) < \delta$ 时,** $\sup_{i\in I} E(|\xi_i|; A) < \varepsilon$.

证明. 必要性. 对任意 $A \in \mathcal{F}$, $N > 0$, 有

$$\begin{aligned} E(|\xi_i|; A) &= E(|\xi_i|; A \cap \{|\xi_i| \geqslant N\}) + E(|\xi_i|; A \cap \{|\xi_i| < N\}) \\ &\leqslant E(|\xi_i|; \{|\xi_i| \geqslant N\}) + NP(A). \end{aligned}$$

运用一致可积性, 推出 $\{\xi_i\}$ 是一致绝对连续的. 再在上式令 $A = \Omega$, 得

$$E|\xi_i| \leqslant E(|\xi_i|; \{|\xi_i| \geqslant N\}) + N,$$

得到 $\{\xi_i\}$ 的 L^1- 有界性.

充分性. 设 $\{\xi_i\}$ 是一致绝对连续且 L^1- 有界的. 由 Chebyshev 不等式,

$$\sup_i P(|\xi_i| \geqslant N) \leqslant \frac{1}{N} \sup_i E|\xi_i|.$$

因此对任何 $\varepsilon > 0$, 存在 N 使得对任何 $i \in I$, $P(|\xi_i| \geqslant N) < \delta$. 然后由 $\{\xi_i\}$ 的一致绝对连续性得到, 对任何 $\varepsilon > 0$, 存在 $N > 0$, 使得 $E(|\xi_i|; \{|\xi_i| \geqslant N\}) \leqslant \varepsilon$ 对所有 $i \in I$ 成立, 即一致可积性. □

定理 6.2 **可积随机变量序列 $\{\xi_n\}$ L^1- 收敛于 ξ 的充要条件是 $\{\xi_n\}$ 是一致可积的且 $\xi_n \xrightarrow{p} \xi$.**

证明. 必要性. 首先 $\xi_n \xrightarrow{p} \xi$ 是显然的, $\{\xi_n\}$ 的 L^1- 有界性也是显然的. 只需验证 $\{\xi_n\}$ 的一致绝对连续性就够了, 而这与 $\{\xi_n - \xi\}$ 的一致可积性等价, 所以不妨设 $\xi_n \xrightarrow{L^1} 0$. 对任意 $A \in \mathcal{F}$,

$$E(|\xi_n|; A) \leqslant E(|\xi_n|).$$

对任何 $\varepsilon > 0$, 存在 N, 当 $n > N$ 时, $E(|\xi_n|) < \varepsilon$. 另外因 $\xi, \xi_1, \cdots, \xi_N$ 可积, 故存在 $\delta > 0$, 只要 $P(A) < \delta$ 便有对所有 $1 \leqslant i \leqslant N$ 有 $E(|\xi_i|; A) < \varepsilon$. 因此只要 $P(A) < \delta$, 就有 $\sup_n E(|\xi_n|; A) < \varepsilon$.

充分性. 首先要证明 ξ 是可积的, 因为依概率收敛, 故存在子列 $\{\xi_{k_n}\}$ 几乎处处收敛于 ξ. 由 Fatou 引理及 L^1- 有界性, $E|\xi| \leqslant \liminf_n E|\xi_n| < \infty$. 因此 ξ 可积. 现在不妨设 $\xi = 0$. 对任意 $\varepsilon > 0$, 由一致绝对连续性, 存在 $\delta > 0$, 只要 $P(A) < \delta$ 就有 $\sup_n E(|\xi_n|; A) < \varepsilon$. 再由于 $\xi_n \xrightarrow{p} 0$, 存在 N, 当 $n > N$ 时就有 $P(|\xi_n| > \varepsilon) < \delta$. 因此

$$\begin{aligned} E|\xi_n| &\leqslant E(|\xi_n|; |\xi_n| < \varepsilon) + E(|\xi_n|; |\xi_n| \geqslant \varepsilon) \\ &\leqslant \varepsilon + E(|\xi_n|; \{|\xi_n| \geqslant \varepsilon\}) < 2\varepsilon. \end{aligned}$$

证明了 $\xi_n \xrightarrow{L^1} 0$. □

6.1.2　负指标下鞅

定义 6.3 **如果 $X = (X_n)_{n \leqslant 0}$ 是关于流 $(\mathcal{F}_n)_{n \leqslant 0}$ 的下鞅, 则称之为负指标下鞅.**

定理 6.3 **设 X 是负指标下鞅且 $\inf_n EX_n > -\infty$, 则 X 是一致可积的;**

证明. 设 $n \leqslant 0$, 因 $EX_n \geqslant EX_{n-1}$, 故 $\inf_n EX_n > -\infty$ 蕴含着 $x = \lim_{n\to-\infty} EX_n$ 存在且有限. 对给定 $\varepsilon > 0$, 取 k 使得 $EX_k - x < \varepsilon$, 那么当 $n \leqslant k$ 时,

$$\begin{aligned}
E(|X_n| : |X_n| > \lambda) &= E(X_n : X_n > \lambda) - E(X_n : X_n < -\lambda) \\
&= E(X_n : X_n > \lambda) + E(X_n : X_n \geqslant -\lambda) - EX_n \\
&\leqslant E(X_k : X_n > \lambda) + E(X_k : X_n \geqslant -\lambda) - EX_k + \varepsilon \\
&\leqslant E(X_k : X_n > \lambda) + E(-X_k : X_n < -\lambda) + \varepsilon \\
&\leqslant E(|X_k| : |X_n| > \lambda) + \varepsilon.
\end{aligned}$$

另外

$$\begin{aligned}
P(|X_n| > \lambda) \leqslant \frac{1}{\lambda}E|X_n| &= \frac{1}{\lambda}E(2X_n^+ - X_n) \\
&= \frac{1}{\lambda}(2EX_n^+ - EX_n) \leqslant \frac{1}{\lambda}(2EX_0^+ - x),
\end{aligned}$$

由此推出 X 是一致可积的. □

注意, 因为 $\{X_n^+\}$ 也是下鞅, 所以 $0 \leqslant EX_n \leqslant EX_0$. 因此条件 $\inf_{n\geqslant 0} EX_n > -\infty$ 等价于 $\sup_{n\geqslant 0} EX_n^- < \infty$, 也等价于 $\sup_{n\geqslant 0} E|X_n| < \infty$.

6.2　Doob 停止定理

为了证明连续时间的 Doob 停止定理, 需要一些预备知识.

6.2.1　Doob 停止定理

定理 6.4　**设** X **是右连续** $(\mathcal{F}_t)$ **下鞅,** τ, σ **是有界停时且** $\sigma \leqslant \tau$, **则** X_σ, X_τ **是可积的且**

$$E(X_\sigma) \leqslant E(X_\tau).$$

证明. 首先, 上述结论在离散时间时成立, 而且在离散时间的情形下, 对任何有界停时 $\sigma \leqslant \tau$ 有下面的不等式

$$E(X_\tau|\mathcal{F}_\sigma) \geqslant X_\sigma \text{ a.s.}$$

这是我们下面需要的一个结论, 先证明它. 事实上, X_σ 是 $\mathcal{F}_\sigma$ 可测的, 对任何 $B \in \mathcal{F}_\sigma$, 令

$$\sigma_B := \sigma 1_B + N 1_{B^c},\ \tau_B := \tau 1_B + N 1_{B^c}.$$

容易验证 σ_B 是停时, 另外由于 $\sigma \leqslant \tau$, 故 τ_B 也是停时. 因此由 Doob 有界停止定理得

$$EX_{\sigma_B} \leqslant EX_{\tau_B},$$

也就是说

$$E(X_\sigma; B) + E(X_N; B^c) \leqslant E(X_\tau; B) + E(X_N; B^c),$$

推出 $E(X_\sigma; B) \leqslant E(X_\tau; B)$, 即 $E(X_\tau|\mathcal{F}_\sigma) \geqslant X_\sigma$ a.s.

回到连续时间情形, 对 $n \geqslant 0$, 令 $D_n := \{\frac{k}{2^n} : k = 0, 1, 2, \cdots\}$ 及

$$\sigma_n(\omega) := \inf\{t \in D_n :\ t \geqslant \sigma(\omega)\},\ \omega \in \Omega.$$

因 $D_n \subset D_{n+1}$, 故 σ_n 与 σ_{n+1} 是值域为 D_{n+1} 的关于流 $(\mathcal{F}_t :\ t \in D_{n+1})$ 的有界停时, 应用 Doob 离散时间有界停时定理于 $(\mathcal{F}_t :\ t \in D_{n+1})$ 下鞅 $(X_t :\ t \in D_{n+1})$, 得知 X_{σ_n} 与 $X_{\sigma_{n+1}}$ 是可积的且

$$X_{\sigma_{n+1}} \leqslant E(X_{\sigma_n}|\mathcal{F}_{\sigma_{n+1}}).$$

对 $n = 0, -1, -2, \cdots$, 令 $Y_n := X_{\sigma_{-n}}$, $\mathcal{G}_n := \mathcal{F}_{\sigma_{-n}}$, 则 $(Y_n : n = 0, -1, -2, \cdots)$ 是关于 $(\mathcal{G}_n : n = 0, -1, -2, \cdots)$ 的下鞅, 且对任何 $n \leqslant 0$, $EY_n = EX_{\sigma_{-n}} \geqslant EX_0$, 由定理 7.3, $(X_{\sigma_n} : n = 0, 1, 2, \cdots)$ 是一致可积的, 且因为 X 是右连续的, 故当 $n \to \infty$ 时, X_{σ_n} 几乎处处收敛于 X_σ, 因而也 L^1- 收敛于 X_σ, 因此 X_σ 是可积的, 同理 X_τ 也是可积的. 同样定义 $\tau_n := \inf\{t \in D_n :\ t \geqslant \tau\}$, 则 $\tau_n \geqslant \sigma_n$ 且都是有界停时, 对任何 $A \in \mathcal{F}_\sigma = \bigcap_n \mathcal{F}_{\sigma_n}$, 再应用 Doob 停止定理, 对任何 n,

$$E(X_{\tau_n}; A) \geqslant E(X_{\sigma_n}; A),$$

由 L^1 收敛性得 $E(X_\tau; A) \geqslant E(X_\sigma; A)$. □

6.2.2　应用

让我们利用布朗运动的鞅性质以及 Doob 定理来解决一些问题. 这在离散时间情形我们已经见到过. 总的来说, 需要找到一个合适的鞅 $\{M_n\}$, 应用 Doob 停止定理, 对

任何 $n>0$ 有

$$EM_{n\wedge T}=EM_0.$$

然后让 $n\to+\infty$. 当 $T<\infty$ 时, $M_{n\wedge T}$ 的极限是 M_T, 当 $T=\infty$ 时, $M_{n\wedge T}=M_n$ 却未必收敛. 由此我们需要考虑几件事情: 1. 极限是否存在? 2. 极限与期望能否交换? 3. 能否解决问题? 在连续时间情形, 问题是类似的.

例 6.1 设 $B=(B_t)$ 是 1-维标准布朗运动. 对 $a>0$, 定义 T_a 是 B 到点 a 的首中时, 那么它是停时, 前面用反射原理证明了布朗运动一定会到达 a, 即

$$P(T_a<\infty)=1.$$

我们也可以用鞅方法来解答这个问题. 更一般地, 我们问布朗运动会肯定碰到一条斜的直线 $x=kt-a$ 吗? 令 T 是布朗运动 B 首次碰到这条直线的时间, 即

$$T=\inf\{t>0:B_t=kt-a\}.$$

这也可以说是漂移布朗运动 (B_t-kt) 首次碰到 $-a$ 的时间, 求 $P(T<\infty)$. 不妨设 $a>0$, 当 $k=0$ 时, $T=T_{-a}$. 直观看, 当 $k>0$ 时, $P(T<\infty)=1$, 而当 $k<0$ 时未必.

由指数鞅性质, 对任何实数 z,

$$\exp\left(zB_t-\frac{z^2}{2}t\right),\ t\geqslant 0$$

是鞅. 由 Doob 定理,

$$E\left[\exp\left(zB_{t\wedge T}-\frac{z^2}{2}(t\wedge T)\right)\right]=1. \tag{6.1}$$

让 t 趋于无穷, 问题的关键是极限与期望是否可以交换. 当 $z<0$ 时,

$$\begin{aligned}zB_{t\wedge T}-\frac{z^2}{2}(t\wedge T)&\leqslant z(k(t\wedge T)-a)-\frac{z^2}{2}(t\wedge T)\\&=(zk-\frac{z^2}{2})(t\wedge T)-za.\end{aligned}$$

因此要保证 (6.1) 式中的指数关于 t,ω 有界, 必须 $zk-z^2/2\leqslant 0$. 需分两种情况, 一种是 $k\geqslant 0$, 这时只要 $z<0$; 另外一种情况是 $k<0$, 这时只要 $z<2k$. 无论哪种情况, 都可以应用有界收敛定理,

$$E\left[\exp\left(zB_T-\frac{z^2}{2}T\right);T<\infty\right]=1.$$

这时 $B_T = kT - a$, 因此

$$E\left[e^{(zk-\frac{z^2}{2})T}; T < \infty\right] = e^{za}.$$

在第一种情况下, 可以让 $z \uparrow 0$, 得 $P(T < \infty) = 1$; 在第二种情况下, 让 $z \uparrow 2k$, 得

$$P(T < \infty) = e^{2ka} < 1.$$

在第一种 $k \geqslant 0$ 情况下, 我们可以算出 T 的 Laplace 变换, 因为

$$E\left[e^{(zk-\frac{z^2}{2})T}\right] = e^{za},$$

令 $-s = zk - z^2/2$, $s > 0$, 得 $z = k - \sqrt{k^2 + 2s} < 0$, 因此

$$E[e^{-sT}] = e^{a(k-\sqrt{k^2+2s})}. \tag{6.2}$$

如果用 T_a 表示上面的 T, 那么 $(T_a : a \geqslant 0)$ 是一个平稳独立增量过程, 称为相对稳定过程. ∎

例 6.2　设 $a < 0 < b$, T_a, T_b 分别是原点出发的布朗运动首次碰到 a, b 的时间. 用鞅方法来求 $P(T_a < T_b)$, $E[T]$ 以及 T 的 Laplace 变换, 其中 $T = T_a \wedge T_b$. 由鞅的期望不变性, 对任何 $t > 0$,

$$EB_{T\wedge t} = 0.$$

当 $t \to \infty$ 时, $a \leqslant B_{T\wedge t} \leqslant b$, 因此由控制收敛定理得 $EB_T = 0$. 而

$$B_T = a1_{\{T_a<T_b\}} + b1_{\{T_b<T_a\}},$$

因此 $aP(T_a < T_b) + bP(T_b < T_a) = 0$, 即

$$P(T_a < T_b) = \frac{b}{b-a}.$$

要算 ET, 需要用鞅 $(B_t^2 - t)$, 还是由期望不变性,

$$E[B_{T\wedge t}^2] = E[T \wedge t].$$

然后让 $t \to \infty$, 左边应用控制收敛定理右边应用单调收敛定理得

$$E[T] = E[B_T^2] = a^2P(T_a < T_b) + b^2P(T_a > T_b) = -ab.$$

类似地, 要算 T 的 Laplace 变换, 应该用指数鞅 $(\exp(\lambda B_t - \lambda^2 t/2))$. 由期望不变性和控制收敛定理得

$$\mathsf{E}\left[\exp(\lambda B_T - \lambda^2 T/2)\right] = 1, \tag{6.3}$$

因此

$$e^{\lambda a}\mathsf{E}\left[e^{-\lambda^2 T/2}; T_a < T_b\right] + e^{\lambda b}\mathsf{E}\left[e^{-\lambda^2 T/2}; T_a > T_b\right] = 1. \tag{6.4}$$

这还不足以算出 $\mathsf{E}\left[e^{-\lambda^2 T/2}\right]$. 再考虑指数鞅 $(\exp(-\lambda B_t - \lambda^2 t/2))$, 类似地得

$$\mathsf{E}\left(\exp(-\lambda B_T - \lambda^2 T/2\right) = 1 \tag{6.5}$$

和

$$e^{-\lambda a}\mathsf{E}\left[e^{-\lambda^2 T/2}; T_a < T_b\right] + e^{-\lambda b}\mathsf{E}\left[e^{-\lambda^2 T/2}; T_a > T_b\right] = 1. \tag{6.6}$$

从上面两个方程推出

$$\mathsf{E}\left[e^{-\lambda^2 T/2}\right] = \frac{\sinh(\lambda a) - \sinh(\lambda b)}{\sinh(\lambda(a-b))}, \tag{6.7}$$

其中 $\sinh x = (e^x - e^{-x})/2$. 由此立刻得到 T 的 Laplace 变换. ∎

第七章 Itô 积分与 Itô 公式

7.1 随机积分

简单起见, 在概率空间 $(\Omega, \mathcal{F}, P)$ 上有流 $(\mathcal{F}_t)$, $B = (B_t)$ 是关于 $(\mathcal{F}_t)$ 的从 0 出发的布朗运动, 即轨道连续, $P(B_0 = 0) = 1$ 且对于任何 $A \in \mathcal{B}(\mathbf{R})$,

$$P(B_t - B_s \in A|\mathcal{F}_s) = \int_A p(t-s, x)dx.$$

流 $(\mathcal{F}_t)$ 可以简单地理解为布朗流.

7.1.1 Riemann 和

当我们说在区间 $[a, b]$ 上一个函数 g 关于另一个函数 f Riemann-Stieltjes 可积, 是指

$$\lim_{D \to 0} \sum_i g(\xi_i)[f(t_i) - f(t_{i-1})]$$

存在, 且独立于 ξ_i 在 $[t_{i-1}, t_i]$ 中的选取, 其中 $D = \{t_i\}$ 是 $[a, b]$ 的划分.

定理 7.1 **闭区间上连续函数 g 关于有界变差函数 f 是 Riemann-Stieltjes 可积的.**

因此一个连续的随机过程 X 关于一个有界变差过程 V 是 Riemann-Stieltjes 可积的, 其积分可以在几乎所有样本轨道上定义

$$\left(\int_0^t X dV\right)(\omega) := \int_0^t X_s(\omega) dV_s(\omega),$$

理解为通常的积分意义. 这对于布朗运动来说是不可能的, 例如考虑分别选取区间的右端点和左端点

$$\lim_{D \to 0} \left(\sum_i B_{t_i}[B_{t_i} - B_{t_{i-1}}] - \sum_i B_{t_{i-1}}[B_{t_i} - B_{t_{i-1}}] \right)$$

$$= \lim_{D\to 0}\left(\sum_i [B_{t_i} - B_{t_{i-1}}]^2\right) = b - a.$$

所以 Riemann-Stieltjes 意义下是不可积, 原因当然是布朗运动的轨道不是有界变差的.

7.1.2 Itô 积分

既然 Riemann-Stieltjes 积分是不可能的, 那么是否可以选定 ξ_i, 例如 $\xi_i = t_{i-1}$, 问上面的和是不是收敛? 因为是一个随机变量族, 所以应该问在什么意义下收敛? 由泛函分析中的结论, 在几乎处处收敛的意义下, 这也是不可能的. 但我们还是可以考虑在 L^2 收敛的意义下, 或者最弱的以概率收敛的意义下.

考虑区间 $[0,T]$, $D = \{0 = t_0 < t_1 < \cdots < t_n\}$ 是其上的划分, 定义

$$H_t := \sum_{i=1}^{n} h_{i-1} 1_{[t_{i-1},t_i)}(t),$$

其中 $\{h_i\}$ 是随机变量列, $H = (H_t)$ 称为阶梯过程. 容易验证下面两个断言

1. $H = (H_t)$ 适应于 $(\mathcal{F}_t)$ 当且仅当对任何 i, h_i 是 $\mathcal{F}_{t_i}$ 可测的.

2. 适应的阶梯过程全体是线性空间.

第一个断言很简单. 为了证明第二个断言, 注意到如果 D' 是比 D 更细的划分, 例如 $[t_{i-1}, t_i)$ 之间有一个分点 u, 那么

$$h_{i-1}1_{[t_{i-1},t_i)} = h_{i-1}1_{[t_{i-1},u)} + h_{i-1}1_{[u,t_i)},$$

显然 h_{i-1} 是 $\mathcal{F}_u$ 可测的, 因此不破坏右边的适应性. 这样两个适应的阶梯过程可以看成同一个划分上的适应阶梯过程, 它们的和自然是适应阶梯过程.

怎么定义关于布朗运动的积分? 先考虑简单的情况

$$\int_{t_{i-1}}^{t_i} dB_t = \int_0^T 1_{[t_{i-1},t_i)}(t)dB_t = B_{t_i} - B_{t_{i-1}},$$

那么阶梯过程 F 关于布朗运动的积分应该是

$$\int_0^T H_s dB_s = \sum_i h_{i-1}(B_{t_i} - B_{t_{i-1}}),$$

注意阶梯过程的表示不唯一, 但右边与它的表示无关. 用 $\mathcal{L}_0$ 表示有界的适应阶梯过程全体.

引理 7.1 Itô 等距: 对于 $H \in \mathcal{L}_0$,

$$E\left(\int_0^T H_t dB_t\right)^2 = \int_0^T E[H_t^2]dt = E\int_0^T H_t^2 dt.$$

证明. H 有界保证随机变量的可积性, 然后

$$\begin{aligned}E\left(\int_0^T H_t dB_t\right)^2 &= E\left(\sum_i h_{i-1}(B_{t_i} - B_{i_{i-1}})\right)^2 \\ &= E\left(\sum_i h_{i-1}^2(B_{t_i} - B_{i_{i-1}})^2\right) \\ &= \sum_i E[h_{i-1}^2]E(B_{t_i} - B_{i_{i-1}})^2 \\ &= \sum_i E[h_{i-1}^2](t_i - t_{i-1}) = E\int_0^T H_t^2 dt,\end{aligned}$$

其中展开时的交叉项因为独立增量性等于零. □

这个等式是定义随机积分的关键, 归功于 Itô, 称为 Itô 等距. 为什么称为 Itô 等距呢? 这个恒等式的两边提示我们应该使用什么样的距离. 左边是通常的 L^2 距离, 右边是被积过程空间的距离.

定义 7.1 设 $H = (H_t)$ 是随机过程, 作为 $[0,T] \times \Omega$ 上的函数是 $\mathcal{B}[0,T] \times \mathcal{F}$ 可测的. 范数 $\|H\|_2$ 是 L^2-空间

$$\mathbb{L}^2 = L^2([0,T] \times \Omega, \mathcal{B}[0,T] \times \mathcal{F}, dt \cdot P)$$

的 L^2-范数, 即

$$\|H\|_2 = \left(\int_0^T E[H_t^2]dt\right)^{1/2}.$$

这时, $\mathbb{L}^2$ 是 Hilbert 空间. $\mathcal{L}_0$ 在 $\mathbb{L}^2$ 中的闭包记为 $\mathcal{L}^2$.

对于 $H \in \mathcal{L}_0$, Itô 等距成立

$$\left\|\int_0^T H_t dB_t\right\|_{L^2} = \|H\|_2. \tag{7.1}$$

注意 T 是任意且固定的. 令

$$\mathrm{I}_{\mathsf{T}}(\mathsf{H}) := \int_0^{\mathsf{T}} \mathsf{H}_t \mathrm{d}\mathsf{B}_t,$$

即 I 是 $\mathcal{L}_0 \cap \mathcal{L}^2$ 到 $\mathsf{L}^2(\Omega, \mathcal{F}, \mathsf{P})$ 的等距线性算子. 唯一地延拓到 $\mathcal{L}^2$ 上.

定义 7.2　记

$$\mathrm{I}_{\mathsf{T}}(\mathsf{H}) = \int_0^{\mathsf{T}} \mathsf{H}_t \mathrm{d}\mathsf{B}_t,\ \mathsf{H} \in \mathcal{L}^2,$$

被称为 H 关于 B 在 $[0, T]$ 上的 Itô 积分.

注意, 这里的从阶梯过程到 $\mathcal{L}^2$ 中过程的积分收敛是 L^2 意义的. 另外, 完备化空间 $\mathcal{L}^2$ 具体包含哪些过程还不清楚.

设 $\mathsf{H} = (\mathsf{H}_t)$ 是连续适应过程且 $\mathsf{H} \in \mathbb{L}^2$. 定义

$$\mathsf{H}_t^{\mathsf{D}} := \sum_{i=1}^{n} \mathsf{H}_{t_{i-1}} 1_{[t_{i-1}, t_i)}(t),$$

它是一个适应阶梯过程, 称为 H 在 D 上取左端点的阶梯过程. 那么

$$\int_0^{\mathsf{T}} \mathsf{H}_t^{\mathsf{D}} \mathrm{d}\mathsf{B}_t = \sum_i \mathsf{H}_{t_{i-1}} (\mathsf{B}_{t_i} - \mathsf{B}_{i_{i-1}}),$$

右边从形式看是取左端点的 Riemann-Stieltjes 和, 简称为左 (端点) Riemann 和. 我们的目的是证明当 $\mathsf{D} \to 0$ 时, 左 Riemann 和在 L^2 意义下收敛.

定理 7.2　**设 $\mathsf{H} = (\mathsf{H}_t)$ 是连续适应过程且 $\|\mathsf{H}\|_2 < \infty$. 则 $\mathsf{H} \in \mathcal{L}^2$.**

证明. 先假设 H 有界, 即存在常数 $\mathsf{C} > 0$ 使得 $|\mathsf{H}_t(\omega)| \leqslant \mathsf{C}$. 取 D_n 是收敛于 0 的划分列, 则由轨道连续性, 对任何 $t \geqslant 0$, $\mathsf{H}_t^{\mathsf{D}_n} \longrightarrow \mathsf{H}_t$ a.s. 然后由控制收敛定理推出

$$\|\mathsf{H}^{\mathsf{D}_n} - \mathsf{H}\|_2 \longrightarrow 0,$$

即 $\mathsf{H} \in \mathcal{L}^2$. 对无界的 H, 取

$$\mathsf{H}^{(n)} := (-n) \vee \mathsf{H} \wedge n.$$

那么 $\mathsf{H}^{(n)} \in \mathcal{L}^2$, 且同样有 $\|\mathsf{H}^{(n)} - \mathsf{H}\|_2$ 趋于 0, 因此 $\mathsf{H} \in \mathcal{L}^2$.　□

因此, 假设 $H = (H_t)$ 是连续适应过程. 则对任何趋于零的划分列 D_n, 阶梯过程列 H^{D_n} 几乎处处趋于 H. 如果附加一定的条件, 例如有界性, 使得 $\|H^{D_n} - H\|_2$ 趋于零, 你们在 L^2 意义下左 Riemann 和收敛于 H 关于 B 的积分, 即

$$\lim_n \int_0^T H_t^{D_n} dB_t = \int_0^T H_t dB_t.$$

练习 7.1 **R** 上 Borel 可测函数 f 定义一个随机过程 $f(B_t)$, $t \geqslant 0$. 一个有趣的问题是: f 满足什么条件可以保证 $f(B) \in \mathcal{L}^2$? 这时首先要求

$$\|f(B)\|_2^2 = \int_0^T E[f(B_t)^2]dt = \int_0^T dt \int_{\mathbf{R}} p(1,x) f(\sqrt{t}x)^2 dx < \infty,$$

即 $f(\sqrt{t}x) \in L^2([0,T] \times \mathbf{R}, dt \cdot p(1,x)dx)$, 其中 $p(1,x)$ 是标准标准正态分布的密度, 这时我们说 f 满足 Itô 积分的可积性条件. 定义

$$\mathbb{H} = \{f \in \mathcal{B}(\mathbf{R}) : f(B) \in \mathcal{L}^2\},$$

它是装备范数 $\|f\|_{\mathbb{H}} = \|f(B)\|_2$ 的 Hilbert 空间. 容易看出, $\mathbb{H}$ 包含有界连续函数全体. 另外, 如果在 Lebesgue 测度下 $f = 0$ a.e., 那么 $\|f\|_{\mathbb{H}} = 0$. 这推出随机积分

$$\int_0^T f(B_t) dB_t = 0, \text{ a.s.}$$

因为对任意 T, 得到一个随机积分 $I_T(H)$, 所以我们实际上得到一个随机过程 $I(H) = (I_t(H) : t \geqslant 0)$, 其中

$$I_t(H) := \int_0^t H_s dB_s, t \geqslant 0,$$

其中每一个都是几乎处处确定的, 这个过程是不是有一个连续修正呢?

7.1.3 连续鞅

设 $H \in \mathcal{L}_0$, 表示 $H_s = \sum_i h_i 1_{[t_{i-1},t_i)}(s)$. 积分 $I(H)$ 是一个平方可积的随机变量. 任意固定 T, D 是 $[0,T]$ 的划分, 当 $t \in [0,T]$ 时, $D^t = D \wedge t$ 自然是 $[0,t]$ 的划分. 这时 H 自然可以看成是 $[0,t]$ 上的阶梯过程, 它的 Itô 积分是

$$\int_0^t H_s dB_s = \sum_i h_i (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}),$$

记为 $I_t(H)$.

引理 7.2 对于 $t \in [0, T]$,

$$E[I_T(H)|\mathcal{F}_t] = I_t(H),$$

即随机过程 $(I_t(H) : t \in [0, T]$ **是平方可积鞅, 记为** $I(H)$, **称为** Itô **积分** (**过程**).

当 $H \in \mathcal{L}_0$ 时, 由定义看出 $I_t(H)$ 是连续的过程, 所以 $I(H)$ 是连续的平方可积鞅. 为了证明当 $H \in \mathcal{L}^2$ 时, Itô 积分 $I(H)$ 也是连续平方可积鞅, 还需要一些预备工作. 首先是最基本的 Doob 不等式, 也是 Kolmogorov 不等式的推广.

定理 7.3 (Doob) **设** X **是一个鞅. 对任何** $\lambda > 0$ **及正整数** N,

$$P(\max_{0 \leqslant n \leqslant N} |X_n| \geqslant \lambda) \leqslant \frac{1}{\lambda^2} E[X_N^2].$$

证明. 令 $\tau := \min\{0 \leqslant n \leqslant N : |X_n| \geqslant \lambda\}$, 则 τ 是一个停时且 $\tau \leqslant N$, 应用 Doob 停止定理于下鞅 (X_n^2), 得

$$\begin{aligned} EX_N^2 &\geqslant EX_\tau^2 \\ &= E(X_\tau^2; \max_{0 \leqslant n \leqslant N} |X_n| \geqslant \lambda) + E(X_\tau^2; \max_{0 \leqslant n \leqslant N} |X_n| < \lambda) \\ &\geqslant \lambda^2 P(\max_{0 \leqslant n \leqslant N} |X_n| \geqslant \lambda). \end{aligned}$$

推出我们想要的不等式. □

怎么推广到连续时间鞅呢?

定理 7.4 **设** $X = (X_t : t \geqslant 0)$ **是连续的平方可积鞅. 则对任何** $\lambda > 0$ **及** $T > 0$,

$$P(\max_{0 \leqslant t \leqslant T} |X_t| \geqslant \lambda) \leqslant \frac{1}{\lambda^2} E[X_T^2].$$

证明. 取 $[0, T]$ 的一个可数稠子集 $D = \{t_i : t \geqslant 1\}$. 令

$$D_n := \{0, t_1, \cdots, t_n, T\}.$$

则 $(X_t : t \in D_n)$ 是个鞅, 因此

$$P(\max_{t \in D_n} |X_t| \geqslant \lambda) \leqslant \frac{1}{\lambda^2} E[X_T^2].$$

左边的最大值关于 n 递增, 推出

$$P(\max_{t \in D} |X_t| \geqslant \lambda) \leqslant \frac{1}{\lambda^2} E[X_T^2].$$

然后因为 X 连续, 所以 $\max_{t\in[0,T]}|X_t|$ 与 $\max_{t\in D}|X_t|$ 几乎处处相等, 因此推出定理结论. □

定理 7.5 **设** $H\in\mathcal{L}^2$. **则** $(I_t(H):t\in[0,T])$ **有一个修正是连续平方可积鞅**.

证明. 取 $H^{(n)}\in\mathcal{L}_0$ 以范数收敛于 H. 则显然有

$$E[(I_t(H^{(n)})-I_t(H))^2]\longrightarrow 0,$$

且 $E[I_T(H^{(n)})|\mathcal{F}_t]=I_t(H^{(n)})$. 因此推出 $(I_t(H):t\in[0,T])$ 是个平方可积鞅.
对于连续鞅 $I_t(H^{(n)})-I_t(H^{(m)}),t\in[0,T]$, 应用 Doob 不等式

$$\begin{aligned}P(\max_{t\in[0,T]}|I_t(H^{(n)})-I_t(H^{(m)})|>\lambda)&\leqslant\frac{1}{\lambda^2}E[(I_t(H^{(n)})-I_t(H^{(m)})^2]\\&=\frac{1}{\lambda^2}\|H^{(n)}-H^{(m)}\|_2^2.\end{aligned}$$

存在子列 n_k 使得

$$P\left(\max_{t\in[0,T]}|I_t(H^{(n_{k+1})})-I_t(H^{(n_k)})|>\frac{1}{k^2}\right)\leqslant\frac{k^4}{2^k}.$$

由 Borel-Cantelli 引理推出

$$P\left(\limsup_k\left\{\max_{t\in[0,T]}|I_t(H^{(n_{k+1})})-I_t(H^{(n_k)})|>\frac{1}{k^2}\right\}\right)=0.$$

几乎处处地有

$$\sum_k\max_{t\in[0,T]}\left|I_t(H^{(n_{k+1})})-I_t(H^{(n_k)})\right|<\infty.$$

存在 $X=(X_t:t\in[0,T])$ 使得几乎处处地 $I_t(H^{(n_k)})$ 在 $[0,T]$ 上一致地收敛于 X_t. 这推出两个结论: (1) X 连续; (2) 对任何 $t\in[0,T]$, $X_t=I_t(H)$ a.s. 证明了定理结论. □

定理说明, $\mathcal{L}^2$ 中的随机过程关于布朗运动的随机积分是可以被定义的, 其结果仍然可以看成是一个随机过程, 而且是一个平方可积的连续鞅.

练习 7.2 设 $H,K\in\mathcal{L}^2$, 证明:

$$E[I_T(H)I_T(K)]=\int_0^T E[H_tK_t]dt.$$

练习 7.3 若 $H \in \mathcal{L}_0$, 则

$$I_t(H)^2 - \int_0^t H_s^2 ds,\ t \in [0, T],$$

是一个连续鞅.

定理 7.6 **若** $H \in \mathcal{L}^2$, **则**

$$I_t(H)^2 - \int_0^t H_s^2 ds,\ t \in [0, T],$$

回忆 $\langle X \rangle$ 表示随机过程 X 的二次变差过程.

练习 7.4 若 $H \in \mathcal{L}_0$, 则

$$\langle I(H) \rangle_t = \int_0^t H_s^2 ds.$$

7.2 Itô 公式

微积分中最重要的公式是 Newton-Leibniz 公式, 或者称为微积分积分定理, 而随机积分最重要的公式是在其中扮演同样角色的 Itô 公式.

假设 f 是 $\mathbf{R}$ 上的二次连续可导的函数, 那么对任何 $x, y \in \mathbf{R}$, 有

$$f(y) - f(x) = f'(x) + \frac{1}{2} f''(\theta)(y - x)^2,$$

其中 θ 是位于 x, y 之间的某个点. 取 $[0, T]$ 的一个划分, 有

$$\begin{aligned} f(B_T) - f(B_0) &= \sim_i [f(B_{t_i}) - f(B_{i-1})] \\ &= \sum_i f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2} \sum_i f''(X_i)(B_{t_i} - B_{t_{i-1}})^2 \\ &= I_T(f'(B)^D) + \frac{1}{2} S_1^D. \end{aligned}$$

第一部分没有问题, 当 f' 满足 Itô 积分的可积性条件时, 它收敛于

$$I_T(f'(B)) = \int_0^T f'(B_t) dB_t,$$

第二部分麻烦一点. 先假设 f'' 有界: $|f''| \leqslant 1$. 令

$$S_2^D := \sum_i f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2.$$

那么

$$\begin{aligned}
E[|S_1^D - S_2^D|] &\leqslant E\sum_i |f''(X_i) - f''(B_{t_{i-1}})|(B_{t_i} - B_{t_{i-1}})^2 \\
&\leqslant E\left(\max_i |f''(X_i) - f''(B_{t_{i-1}})| \cdot \sum_i (B_{t_i} - B_{t_{i-1}})^2\right) \\
&\leqslant \left(E\max_i |f''(X_i) - f''(B_{t_{i-1}})|^2 \cdot E(\sum_i (B_{t_i} - B_{t_{i-1}})^2)\right)^{1/2} \\
&\leqslant \left(E\max_i |f''(X_i) - f''(B_{t_{i-1}})|^2 \cdot 3T^2\right)^{1/2}.
\end{aligned}$$

由轨道连续性知 $\max_i |f''(X_i) - f''(B_{t_{i-1}})|^2$ 有界且几乎处处趋于零, 因此

$$E[|S_1^D - S_2^D|] \longrightarrow 0.$$

再令

$$S_3^D := \sum_i f''(B_{t_{i-1}})(t_i - t_{i-1}).$$

$$\begin{aligned}
E[(S_2^D - S_3^D)^2] &= E\left(\sum_i f''(B_{t_{i-1}})[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]\right)^2 \\
&= E\left(\sum_i f''(B_{t_{i-1}})^2[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2\right) \\
&\leqslant E\left(\sum_i [(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})]^2\right) = 2\sum_i (t_i - t_{i-1})^2 \longrightarrow 0.
\end{aligned}$$

最后

$$S_3^D \longrightarrow \int_0^T f''(B_t)dt, \text{ a.s.}$$

这证明了 Itô 公式.

定理 7.7　如果 f 两次连续可导, f' 满足 Itô 积分的可积性条件, 且 f'' 有界, 则

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds, \text{ a.s.}.$$

与 Newton-Leibniz 公式比较, Itô 公式多了一项, 它是由于布朗运动的非零二次变差而导致的.

二阶导数 f'' 有界的条件是一个技术性条件, 读者可以自己想办法去掉.

练习 7.5 设 $V = (V_t)$ 是连续有界变差过程. 试推导 $f(B_t, V_t)$ 的 Itô 公式.

7.3 Tanaka 公式与局部时

布朗运动的几乎所有轨道的零点集是一个测度等于 0 的无处稠的完美集. 是否能够构造一个支撑在零点集上的测度呢? 如果 A 是一个正测度集, 那么

$$\alpha_t := \int_0^t 1_A(B_s)ds$$

是一个仅当布朗运动落在 A 上才会增加的增过程, 称为是 A 的占有时过程. 而当 $A = \{0\}$ 时, 这样定义的过程恒等于零, 所以需要有其它的思路. 下面介绍布朗运动的局部时及其性质, 这个概念的引入要归功于 Lévy.

定义 7.3 **一个点 x 称为是测度 μ 的一个支撑点, 如果对 x 的任何领域 U, 有 $\mu(U) > 0$. 所有支撑点的集合称为该测度的支撑, 是个闭集, 记为 $\operatorname{supp}\mu$.**

显然支撑依赖于拓扑.

练习 7.6 证明: (1) 如果 $\mu(G) = 0$ 并且 G 是开的, 则 $\operatorname{supp}\mu \subset G^c$. (2) $(\operatorname{supp}\mu)^c$ 是零测度开集中最大的.

定理 7.8 (Tanaka) 1. **存在一个连续增过程 $L = (L_t)$ 满足 $L_0 = 0$ 以及**

$$|B_t| = \int_0^t \operatorname{sgn}(B_s)dB_s + L_t.$$

2. **对任何 $t > 0$, 在 L^2 收敛的意义下, 且在几乎处处意义下对有界的 t 有**

$$L_t = \lim_{\delta\downarrow 0} \frac{1}{2\delta}\int_0^t 1_{(-\delta,+\delta)}(B_s)ds.$$

3. $\operatorname{supp} dL \subset Z$, **其中** $\operatorname{supp} dL$ **测度** dL_t **的支撑**, $Z = \{t \geqslant 0 : B_t = 0\}$.

证明. 作函数 h_δ 满足 $h_\delta(0) = h'_\delta(0) = 0$ 且

$$h''_\delta = \frac{1}{\delta}1_{(-\delta,+\delta)},$$

是个凸函数. 取 $g \in C_0^\infty$ 且 $\int g dx = 1$, 令 $g_n(x) = ng(nx)$, $f_n = g_n * h_\delta$. 那么点点地 $f_n \longrightarrow h_\delta$, 点点地 $f'_n \longrightarrow h'_\delta$, 而除了 $|x| = \delta$ 外, $f''_n(x) \longrightarrow h''_\delta(x)$. 由 Itô 公

式,

$$f_n(B_t) - f_n(B_0) = \int_0^t f_n'(B_s)dB_s + \frac{1}{2}\int_0^t f_n''(B_s)ds.$$

让 $n \longrightarrow \infty$. 首先对几乎所有轨道, $\{s \in [0,t] : |B_s| = \delta\}$ 的测度为零, 故对任何 $t > 0$, 几乎处处地

$$h_\delta(B_t) = \int_0^t h_\delta'(B_s)dB_s + \frac{1}{2}\int_0^t h_\delta''(B_s)ds.$$

因为两边对于 t 连续, 所以不可区分, 即等式几乎处处地对所有 t 成立. 而当 $\delta \longrightarrow 0$ 时, 左边和右边第一项对于固定 t 的 L^2 收敛是显然的. 因为

$$\max_x |h_\delta(x) - |x|| \leqslant \delta/2,$$

故 $h_\delta(B_t)$ 对于 $t \geqslant 0$ 在几乎所有轨道上一致收敛于 (几乎处处一致地) $|B_t|$. 因为

$$\begin{aligned}
&E\left(\int_0^t [h_\delta'(B_s) - \mathrm{sgn}(B_s)]dB_s\right)^2 \\
&= E\int_0^t [h_\delta'(B_s) - \mathrm{sgn}(B_s)]^2 ds \\
&\leqslant E\int_0^t 1_{[-\delta,+\delta]}(B_s)ds \\
&= \int_0^t P(|B_s| \leqslant \delta)ds \leqslant \int_0^t \frac{2\delta}{\sqrt{2\pi s}}ds \longrightarrow 0,
\end{aligned}$$

所以, 由 Doob 不等式, 几乎处处一致地

$$\int_0^t h_\delta'(B_s)dB_s \longrightarrow \int_0^t \mathrm{sgn}(B_s)dB_s.$$

完成了 1,2 的证明.

对于 3, 只需证明对于 2 中极限成立的那些 ω, 如果 $t \notin Z(\omega)$, 则 $t \notin \mathrm{supp}\,(dL_t(\omega))$. 下面证明这个断言. 为了方便, 省略 ω. 现在设 $B_t = a$, 存在 $s > 0$ 使得对任何 $u \in (t-s, t+s)$, $|B_u| > |a|/2$. 由 2 中的收敛性

$$L_{t+s} - L_{t-s} = \lim_{\delta \downarrow 0} \frac{1}{2\delta}\int_{t-s}^{t+s} 1_{(-\delta,\delta)}(B_u)du = 0.$$

推出 $t \notin \mathrm{supp}\,(dL)$. □

练习 7.7 令

$$\beta_t := \int_0^t \mathrm{sgn}(B_s)dB_s.$$

证明: $\beta = (\beta_t)$ 是标准布朗运动. 提示: 首先证明 $\mathrm{sgn}(B) = (\mathrm{sgn}(B_t)) \in \mathcal{L}^2$, 然后利用定义证明 β 是 Gauss 过程.

7.4 Skorohod 分解

下面证明 dL_t 的支撑 $\mathrm{supp}\, L$ 恰好是 $Z = \{t : B_t = 0\}$. 首先由 Tanaka 公式 $\mathrm{supp}\, L \subset Z$. 要证明 $\mathrm{supp}\, L \supset Z$, 需要一个引理.

引理 7.3 (Skorohod) **设** y **是** $[0,\infty)$ **上实值连续函数且** $y(0) \geqslant 0$. **则存在唯一的** $[0,\infty)$ **上连续函数对** (x, z) **满足:**

1. $z = x + y$;
2. z **是非负的;**
3. x **是递增连续函数,** $x(0) = 0$ **且** $dx(t)$ **支撑在** $\{t : z(t) = 0\}$ **上.**

实际上, $x(t) = \sup_{s\leqslant t}(-y(s))^+ = -\inf_{s\leqslant t} y(s)^-$.

证明. 验证 $x(t) := \sup_{s\leqslant t}(-y(s))^+$ 和 $z := x + y$ 满足条件 1,2,3 的工作读者可自行完成. 关键是证明唯一性. 设 $(\widehat{x}, \widehat{z})$ 也满足条件. 那么 $z - \widehat{z} = x - \widehat{x}$, 因 (3), 故

$$\begin{aligned}(z(t) - \widehat{z}(t))^2 &= 2\int_0^t (z(s) - \widehat{z}(s))d(x(s) - \widehat{x}(s)) \\ &= -2\int_0^t \widehat{z}(s)dx(s) - 2\int_0^t z(s)d\widehat{x}(s) \leqslant 0.\end{aligned}$$

因此 $z = \widehat{z}$, $x = \widehat{x}$. □

由局部时的性质, $|B_t| = \beta_t + L_t$, 其中

$$\beta_t = \int_0^t \mathrm{sgn}(B)dB$$

是一个标准布朗运动. 因此由上面的引理推出

$$L_t = \sup_{s\leqslant t}(-\beta_s)^+ = \sup_{s\leqslant t}(-\beta_s).$$

这得出两个推论 (1) 对任何 t, $L_t > 0$ a.s., 即 $0 \in \mathrm{supp}\, L$; (2) $\lim\limits_{t\to\infty} L_t = \infty$.

定理 7.9 局部时在且只在布朗运动的零点增加, **精确地** $\operatorname{supp} L = Z$ a.s.

证明. 只需验证 $\operatorname{supp} L \supset Z$. 对任何 t, 令 $r_t := \inf\{s > t : B_s = 0\}$. 则 r_t 是停时且由强 Markov 性 $s \mapsto B_{r_t+s}$ 是标准布朗运动. 它的局部时是 $s \mapsto L_{r_t+s} - L_{r_t}$. 故对任何 $s > 0$, $L_{r_t+s} > L_{r_t}$ a.s. 推出 $r_t \in \operatorname{supp} L$ a.s. 或对几乎所有 ω, 对所有有理数 t, $r_t(\omega) \in \operatorname{supp} L(\omega)$. 现在取这样的 ω, 如果 $a \notin \operatorname{supp} L$, 存在 a 的领域 $U_a \cap \operatorname{supp} L = \varnothing$. 取有理数 $t < a$ 且 $t \in U_a$. 因为 $r_t \in \operatorname{supp} L$, 故 $r_t \notin U_a$, 推出 $a \notin Z$. 因此有 $Z \subset \operatorname{supp} L$. □

设 (τ_t) 是 L_t 的右连续逆, 即 $\tau_t := \inf\{s > 0 : L_s > t\}$. 那么 τ_t 是停时且 $L_{\tau_t} = t$, $\tau_{L_t} = r_t$. 因 $\tau t- < \tau_t$ 当且仅当 L 在 $[\tau_{t-}, \tau_t]$ 上是常数, 而 $\operatorname{supp} L$ 实际上是 dL_t 测度为零的最大开集的补集, 故

$$(\operatorname{supp} L)^c = \bigcup_{t>0} (\tau_{t-}, \tau_t) = Z^c.$$

引理 7.4 (1) **设** T **是停时**, $t > 0$ **是正随机变量**. **则** $L_{T+t} = L_T + L_t \circ \theta_T$.
(2) **对任何** $t, s > 0$, $\tau_{t+s} = \tau_t + \tau_s \circ \theta_{\tau_t}$.

证明. 利用引理 7.8,

$$\begin{aligned}
L_{T+t} &= \lim_{\delta \to 0} \frac{1}{2\delta} \int_0^{T+t} 1_{(-\delta,+\delta)}(B_s) ds \\
&= L_T + \lim_{\delta \to 0} \frac{1}{2\delta} \int_T^{T+t} 1_{(-\delta,+\delta)}(B_s) ds \\
&= L_T + \lim_{\delta \to 0} \frac{1}{2\delta} \int_0^{t} 1_{(-\delta,+\delta)}(B_s \circ \theta_T) ds \\
&= L_T + L_t \circ \theta_T
\end{aligned}$$

利用上面的结论计算

$$\begin{aligned}
\tau_{t+s} &= \inf\{r > 0 : L_r > t + s\} = \inf\{r > 0 : r > \tau_t, L_r > t + s\} \\
&= \inf\{r > 0 : L_{r+\tau_t} - L_{\tau_t} > s\} + \tau_t \\
&= \inf\{r > 0 : L_r \circ \theta_{\tau_t} > s\} + \tau_t \\
&= \tau_s \circ \theta_{\tau_t} + \tau_t
\end{aligned}$$

□

习　　题

1. 设 $S_t := \sup_{s\leqslant t} B_s$. 证明: 过程 $\{(S_t - B_t, S_t) : t \geqslant 0\}$ 与 $\{(|B_t|, L_t) : t \geqslant 0\}$ 等价.

2. 证明 Itô-Tanaka 定理证明中所叙述的关于随机积分的 Fubini 定理.

3. 设 $f(x) = |x|$. 证明:

$$L_t = \lim_{s\to 0} \frac{1}{s}\int_0^t (P_s f(B_u) - f(B_u))du.$$

4. 证明: 局部时的逆 (τ_t) 是平稳独立增量过程 (独立增量且增量的分布只依赖与时间差).

5. 设 $B = (B_t^1, B_t^2)$ 是标准 2 维布朗运动. τ_t 是 B^1 的局部时的逆, 证明: $B^2_{\tau_t}$ 是对称 Cauchy 过程.

参考文献

[1] Chung, K.L., Lectures from Markov Processes to Brownian Motion, Springer-Verlag, New York Heidelberg Berlin, 1982

[2] Durrett, R., Brownian motion and Martingale in Analysis, Wadsworth Inc., 1985

[3] Feller, W., An Introduction to Probability Theory and its Applications, Vol. 1(1959), 2(1970), Wiley & Son

[4] Itô, K., Mckean Jr., H. P., Diffusion Processes and Their Sample Paths, Sringer-Verlag, 1965

[5] Karatzas, I., Shreve, S. E., Brownian Motion and Stochastic Calculus, Springer, 1991

[6] Port, S.C., Stone, C.J., Brownian Motion and Classical Potential Theory, Academic Press, 1978

[7] Revuz, D., Yor, M., Continuous Martingales and Brownian Motion, Springer, 1991

[8] 汪嘉冈, 现代概率论基础, 复旦大学出版社, 1988

[9] 应坚刚, 何萍, 概率论, 复旦大学出版社, 2005