随机分析

应坚刚

2016 年 12 月 14 日

目　录

第一章　预备知识 1

§1.1　为什么学随机分析

首先简单说说为什么要学随机分析. 金融行业本来与概率论是不相干的, 金融虽然要用数学, 但基本上都是加减乘除一类的基本运算, 中学数学的知识就足够了. 随着金融行业不断创新, 股票市场建立, 期货期权市场建立, 人们发现了解随机性对于了解金融市场的运作是非常重要的, 因为人们无法确定地掌控股票价格的走向. 在早期的金融行业, 随机行还比较简单, 概率论或者说大数定律就足够了, 但到后来对冲思想进入金融行业后, 人们发现概率论就不够了, 必须利用随机过程来描述股票价格, 而且期货期权背后有非常时刻的数学思想, 这个思想也正是日本数学家伊藤清 (K.Itô) 在 1940 年代建立的随机微积分中所表述的. 到如今, 要真正理解金融市场的运行, 就必须理解随机分析的思想, 或者说随机分析理论可以很好地解释金融市场背后的规则. 可以说, 随机分析之于金融就如同Riemann 几何之于广义相对论, 是数学解释自然与社会现象的典范. 这也就是金融专硕的学生为什么要学习随机分析的理由.

§1.2　事件域

虽然我们要学的是随机分析, 但是还是要从基本概念开始. 下面的同学数学基础不同, 有的我看上去认识, 是本院毕业的高才生, 有的不是数学专业毕业的, 虽然学过数学, 但没有经过严格的数学训练, 但是随机分析这个课程恰好需要非常好的数学基础才能很好的理解, 最好要学过公理概率论, 实变函数以及泛函分析等. 尽管随机分析的理论需要非常好的数学基础才能真正理解, 但它还有非常直观的一面, 它的所有定理结论都是从日常生活中抽象出来的, 有非常直观的解释, 对于数

学基础一般的同学是个好消息, 要试图从其直观的一面去了解其抽象的一面.

随机分析是解释和理解金融数学的基础, 也是深层次地解释一些金融现象的理论.

世界上有各种随机现象, 也就是结果是不确定的现象, 概率论是研究随机现象的主要方法. 这里所有的同学应该都学过概率论, 但是大多数同学只学了随机变量的分布期望等的计算, 并没有真正接触并了解概率论的基础, 概率论的基础是概率空间. 让我们先介绍并认真地审视概率空间, 究竟有哪些东西是我们以前学习的时候所忽略的? 直观地看, 一个随机现象会出现许多可能的结果, 称为样本空间, 用 Ω 表示, 然后我们会说什么事件发生 (或者出现) 的概率是多少. 比如说掷一个硬币, 正面出现的概率多少? 掷一个骰子, 偶数点出现的概率是多少? 所以我们可以用样本空间的子集来表示事件, 然后概率是赋予事件的一个 0 与 1 之间的一个数, 或者说概率是事件的函数. 另外直观地看, 概率应该有可加性: 两个 (互斥) 事件至少有一个发生的概率是两个事件概率的和. 这样, 把上面所说的抽象为数学语言, 样本空间 Ω 是一个非空集合, 这个似乎容易理解, $\mathscr{F}$ 是 Ω 的一些子集组成的集合, 满足:

1. $\emptyset, \Omega \in \mathscr{F}$;

2. 如果 $A \in \mathscr{F}$, 则 $A^c \in \mathscr{F}$;

3. 如果 $A_n \in \mathscr{F}$, $n \geq 1$, 则 $\bigcup_n A_n \in \mathscr{F}$.

满足以上条件的 $\mathscr{F}$ 称为是一个事件域. 事件是一个直观的概念, 但事件域却很不直观, 它是一种数学结构, 实际上是说事件的某些运算还是事件, 比如补, 并, 交, 可列并, 可列交, 等. 在概率论课程中, 事件域总是固定的, 所以不大提到事件域, 但在随机分析中, 事件域是动态的, 所以我们必须好好地了解事件域, 建立起直观.

事件域可以理解为一种分类, 这在样本空间 Ω 有限的时候特别地明显. 哪些是事件域? $\mathscr{F}=\{\emptyset,\Omega\}$ 是事件域, Ω 的全体子集是事件域, 设 $A\subset\Omega$, $\mathscr{F}=\{\emptyset,A,A^c,\Omega\}$ 也是事件域. 如果 Ω 是有限的, 一个事件域 $\mathscr{F}$ 必然给出一个分类, 也就是说, 它首先把样本空间分成为有限个原子事件的并, 原子事件是指它不包含非平凡事件, $\mathscr{F}$ 中的其它事件是若干原子事件的并, 这样 $\mathscr{F}$ 中事件的个数必然是 2^n 个, 其中 n 是原子事件的个数. 分类可以看作信息, 如果人按性别分类, 知道一个人在哪个类, 也就是知道性别, 就是信息; 如果人按年龄分类, 知道一个人在哪个类, 也就是知道他的年龄, 也是信息. 所以事件域可以看成为给出一定的信息. 事件域越大, 表示分类越细, 也表示信息越多. 例如知道一个人的具体年龄比知道这个人是儿童或者少年或者青年或者中年或者老年的信息要丰富得多. 但是身高的信息和年龄的信息是互相不可比较的. 这样我们大概可以建立起关于事件域的直观, 它因为给出分类而代表信息. 下面的命题很重要也很容易验证.

Proposition 1.2.1 任意多个事件域的交还是事件域.

随意取样本空间的一些子集组成集合 $\mathscr{A}$, 一般不是事件域, 但由这个命题, 可以取包含 $\mathscr{A}$ 的所有事件域的交, 它仍然是事件域, 是包含 $\mathscr{A}$ 的最小事件域或者由 $\mathscr{A}$ 生成的事件域, 记为 $\sigma(\mathscr{A})$. 不同的 $\mathscr{A}$ 可以生成相同的事件域, 例如,

$$\sigma(\{A\})=\sigma(\{\Omega,A\})=\sigma(\{A,A^c\})=\{,A,A^c,\Omega\}.$$

如果 $\Omega=\mathbf{R}$, 那么由开区间全体生成的事件域和闭区间全体生成的事件域是一样的, 称为 Borel 域, 记为 $\mathscr{B}(\mathbf{R})$. 它与单点集全体生成的事件域是不同的, 单点集全体生成的事件域是可列集以及余可列的集合全体, 远远小于 Borel 域.

§1.3 可测性及其解释

设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是一个概率空间. $\mathscr{F}$ 有很多的子事件域, 实际上是部分信息, 因为在随机分析中, 信息是动态的, 我们需要考虑子事件域.

定义 1.3.1 设 ξ 是 Ω 上的函数, $\mathscr{G}$ 是一个子事件域. 如果对任何 $x \in \mathbf{R}$ 有

$$\{\xi \le x\} = \{\omega \in \Omega : \xi(\omega) \le x\} \in \mathscr{G},$$

那么我们说 ξ 是关于 $\mathscr{G}$ 可测的 (随机变量). 如果 $\mathscr{G} = \mathscr{F}$, 我们就简单说是随机变量.

如同事件域一样, 可测又是一个非常抽象的概念, 理解它是学懂随机分析的第一步. 在概率论中, 我们一般就简单地说 ξ 是随机变量, 因为那时 $\mathscr{F}$ 是预先给定不变的, $\mathscr{G}$ 就取成 $\mathscr{F}$. 而现在在随机分析中, 必须注意到随机变量是依赖于事件域的, 因为事件域是在不断变化的. 实际上, 样本空间上的函数也直观地给出样本空间的一个分类, 比如 ξ 表示年龄, 那它给出关于年龄的分类, 如果 ξ 取有限个值 $x_1, \cdots, x_n$, 那么自然给出分类

$$\{\xi = x_1\}, \cdots, \{\xi = x_n\}.$$

当 ξ 是一般函数时, 事情要复杂的多. 无论如何, 我们可以认为样本空间上的函数会给出分类, 即给出信息. 它关于 $\mathscr{G}$ 可测是说它给出的这个分类要比 $\mathscr{G}$ 给出的分类更粗糙. 数学上怎么来表达清楚呢? 首先, 关于 $\mathscr{G}$ 可测的随机变量 ξ 在 $\mathscr{G}$ 的原子上是常数, 这大概的意思是: $\mathscr{G}$ 不能区分的样本 ξ 也不能区分. 精确地说, 设 ξ 是样本空间上的函数, 样本空间上的事件域有的使得 ξ 可测, 有的不可测, 把所有使得 ξ 可测的子事件域拿出来取它们的交, 记为 $\sigma(\xi)$, 它是一个子事件域, 它是由 ξ 唯一决定的, 是使得 ξ 可测的最小事件域, 也称为由 ξ 生成的子事件域, 是 ξ 给出的信息. ξ 是关于 $\mathscr{F}$ 可测的当且仅当

$$\sigma(\xi) \subset \mathscr{F},$$

即 ξ 给出的信息没有 $\mathscr{F}$ 给出的多. 不同的随机变量可能会给出相同的分类, 或者说给出相同的信息.

既然样本空间上的函数会给出信息, 那么我们可以定义 η 关于 ξ 可测这个概念. 设 ξ, η 是样本空间上两个函数, 如果 η 关于 $\sigma(\xi)$ 可测, 我们说 η 关于 ξ 可测. 这等价于说 $\sigma(\eta) \subset \sigma(\xi)$, η 的信息没有 ξ 多. 为了叙述下面的命题, 还要介绍一个概念. $\mathbf{R}$ 上的函数称为是可测函数, 如果它关于 Borel 域可测. 连续函数总是可测的.

Proposition 1.3.1 η 关于 ξ 可测当且仅当存在 $\mathbf{R}$ 上可测函数 g 使得

$$\eta = g(\xi).$$

这时, 也说 η 由 ξ 决定.

证明很简单, 只要能够验证

$$\sigma(\eta) = \{\{\eta \in B\} : \ B \in \mathscr{B}(\mathbf{R})\}$$

就可以了, 即 Borel 集的逆像全体. 由于 $\sigma(\eta) \subset \sigma(\xi)$, 对任何 $B \in \mathscr{B}(\mathbf{R})$ 存在 $A \in \mathscr{B}(\mathbf{R})$ 使得

$$1_B(\eta) = 1_A(\xi),$$

即说 η 的示性函数是 ξ 的函数. 因为 η 实际上可以有它的示性函数的线性组合组成的函数列来逼近, 即存在 $\mathbf{R}$ 上可测函数列 f_n 使得 $\eta = \lim f_n(\eta)$, 其中 f_n 是示性函数的线性组合. 这样, 存在 $\mathbf{R}$ 上可测函数 g_n 使得

$$f_n(\eta) = g_n(\xi).$$

令 $g = \lim g_n$, 则 $\eta = g(\xi)$.

这个命题说明, η 关于 ξ 可测就是 η 是 ξ 的函数, 由 ξ 决定. 这比可测性要容易理解得多, 这样一点点, 读者可以慢慢建立起对可测性的认识.

§1.4 概率空间

一个随机试验会带来一个概率空间, 什么是概率空间? 一个概率空间有三个要素: 样本空间 Ω, 事件域 $\mathscr{F}$ 与概率测度 $\mathbb{P}$. 概率测度 $\mathbb{P}$ 定义在事件域上, 满足以下条件:

1. $\mathbb{P}(A) \geq 0$;
2. $\mathbb{P}(\Omega) = 1$;
3. 如果 $A_n \in \mathscr{F}$, 且互斥, 那么

$$\mathbb{P}(\bigcup_n A_n) = \sum_n \mathbb{P}(A_n).$$

最后一条称为可列可加性. 如果对互斥的 $A, B \in \mathscr{F}$ 有

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B),$$

我们就说 $\mathbb{P}$ 有有限可加性. 因为概率定义在 $\mathscr{F}$ 上, 下面我们写 $\mathbb{P}(A)$ 时总是假设 $A \in \mathscr{F}$. 第三条称为是可列可加性, 是概率测度最重要的性质. 这些性质可以推出

1. $\mathbb{P}(\emptyset) = 0$;
2. 单调性: 如果 $A \subset B$, 那么 $\mathbb{P}(A) \leq \mathbb{P}(B)$;
3. 连续性: 如果 $A_n \uparrow A$, 那么 $\mathbb{P}(A_n) \uparrow \mathbb{P}(A)$.

这样的一个组合 $(\Omega, \mathscr{F}, \mathbb{P})$ 称为是一个概率空间.

习题

1. 证明: 事件域的交还是事件域. 举例说明事件域的并不一定还是事件域.

2. 证明: 上面列出的概率的三个性质.

3. 设 Ω 是自然数集合, 对任何 $A \subset \Omega$, 如果

$$\lim_n \frac{|A \cap \{1, 2, \cdots, n\}|}{n}$$

存在, 我们说 A 可度量, 并记极限为 $\mu(A)$, 证明: μ 有有限可加性, 没有可列可加性.

4. 设 η 是随机变量. 证明:

$$\sigma(\eta) = \{\{\eta \in B\} : \ B \in \mathscr{B}(\mathbf{R})\},$$

其中 $\mathscr{B}(\mathbf{R})$ 是 $\mathbf{R}$ 上的 Borel 域 (包含所有区间的最小事件域).

5. 设 $(\omega, \mathscr{F}, \mathbb{P})$ 是概率空间, 证明: 对任何 $\{A_n\} \subset \mathscr{F}$ 有

$$\mathbb{P}(\bigcup_n A_n) \leq \sum_n \mathbb{P}(A_n).$$

称为次可列可加.

6. 证明: 关于 $\mathscr{G}$ 可测的随机变量在 $\mathscr{G}$ 的原子上是常数.

第二章　预备知识 2

§2.1　分布函数

设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是个概率空间, ξ 是关于 $\mathscr{F}$ 可测的随机变量. 那么对任何 $x \in \mathbf{R}$, $\{\xi \le x\}$ 是个事件, 可以算它的概率, 定义

$$F_\xi(x) := \mathbb{P}(\{xi \le x\}),$$

它是 $\mathbf{R}$ 到 $[0,1]$ 的映射, 称为是 ξ 的分布函数. 顾名思义, 分布函数给出 ξ 在 $\mathbf{R}$ 上的概率大小的分布情况, 分布也是一个熟悉的词, 电视报纸上经常有收入分布, 年龄分布, 销售发布, 等等. ξ 在区间 $(a,b]$ 上的分布是

$$\mu_\xi(a,b] = \mathbb{P}(a < \xi \le b) = F_\xi(b) - F_\xi(a).$$

分布完全可以由分布函数来表达. 如果两个随机变量 ξ, η 有同样的分布函数, 我们说它们是同分布的. 这与两个随机变量相等是完全不同的概念, $\xi = \eta$ 是指它们必须是同一个样本空间上的函数且对所有 $\omega \in \Omega$ 有 $\xi(\omega) = \eta(\omega)$. 而同分布的随机变量甚至可以不是在同一个样本空间上定义. 相等的随机变量一定同分布, 但同分布随机变量不需要相等. 例如, 我掷两个骰子, 它们的点数是两个同分布随机变量, 但显然不相等, 因为点数不一定一样.

分布函数 $F = F_\xi$ 满足下列性质:

1. 它是递增且右连续的函数;

2. $F(+\infty) = 1$, $F(-\infty) = 0$.

常见的分布函数有两种类型, 一种是连续型的, 一种是离散型的. 如果存在函数 f 使得对任何 $x \in \mathbf{R}$ 有

$$F(x) = \int_{-\infty}^{x} f(y)dy,$$

那么我们说 F 是连续型的, f 是它的密度函数. 连续型的分布函数一定是连续的, 反之未必. 典型的连续型分布函数有均匀分布, 指数分布和正态分布. 如果存在 $\{x_n\} \subset \mathbf{R}$ 使得

$$\sum_n \mathbb{P}(\xi = x_n) = 1,$$

我们就说 ξ 的分布是离散的或者 ξ 是离散型分布的. 常见的离散型分布有二项分布, 几何分布和 Poisson 分布. 我们假设听众熟悉这些分布.

§2.2 期望和数字特征

随机变量的数学期望或期望, 用 $\mathbb{E}\xi$ 或者 $\mathbb{E}[\xi]$ 表示, 是随机变量的一个线性泛函, 它给出随机变量按照其分布的加权平均. 数学期望的严格定义是很抽象的, 不在这里展开, 但如果 ξ 是离散分布的, 那么

$$\mathbb{E}[\xi] = \sum_n x_n \mathbb{P}(\xi = x_n).$$

右边直观地表达了加权平均的意思, x_n 是值, $\mathbb{P}(\xi = x_n)$ 是取此值的权重. 类似地, 连续型随机变量的期望应该是

$$\mathbb{E}[\xi] = \int_{\mathbf{R}} x f(x) dx.$$

实际上, 一般的分布函数为 F 的随机变量 ξ 的数学期望是

$$\mathbb{E}[\xi] = \int_{\mathbf{R}} x dF(x).$$

注意到数学期望依赖于分布函数, 同分布随机变量一定有一样的期望. 上面右边的积分称为 Riemann-Stieltjes 积分, 它对于区间上的示性函数 $1_{(a,b]}$ 的积分为

$$\int 1_{(a,b]}(x) dF(x) = F(b) - F(a),$$

这个积分值决定了任意连续函数 g 的积分 $\int g(x)dF(x)$, 其中通常的 Riemann 积分是特殊情况 $F(x)=x$.

用单调类方法可以证明以下期望计算公式: 如果 ξ 的分布函数是 F, g 是 $\mathbf{R}$ 上连续函数, 那么

$$\mathbb{E}[g(\xi)] = \int g(x)dF(x),$$

没有表明积分上下限是指从负无穷到正无穷的积分.

期望是随机变量的一个数字特征, 或者说是分布的数字特征, 表示平均. 这是最有用的, 它是使得函数 $f(x)=\mathbb{E}[(\xi-x)^2]$ 达到最小值的地方, 通常两个随机变量 ξ,η 的均方距离定义为

$$\sqrt{\mathbb{E}[(\xi-\eta)^2]},$$

那么期望是所有常数中离 ξ 的均方距离最近的数字. 随机变量还有很多数字特征, 另外一个重要特征是方差, 方差是随机变量与其期望的均方距离, 即

$$D\xi := \mathbb{E}[(\xi-\mathbb{E}\xi)^2] = \mathbb{E}[\xi^2] - [\mathbb{E}\xi]^2.$$

方差的直观意义是随机变量的随机度, 或者说波动度. 比如说常值随机变量的方差是零, 如果 ξ 是区间 $[a,b]$ 上的均匀分布, 那么

$$D\xi = \frac{(b-a)^2}{12},$$

与区间长度平方成比例.

期望方差是随机变量的两个最著名最有用的数字特征. 另外一个与期望有类似意义但名气稍逊于期望的数字特征是中位数. 如果存在 x 使得 $F(x)=1/2$, 那么 x 被称为随机变量或者分布函数的中位数. 注意中位数不一定存在, 存在也不一定唯一. 如果 F 连续, 中位数一定存在. 中位数是使得 $f(x)=\mathbb{E}|\xi-x|$ 达到最小值的地方. 对某些分布, 期望与中位数一样, 比如均匀分布, 正态分布等. 顾名思义它是表示随机

变量取值的中间位置. 从统计的角度说, 如果有许多关于家庭收入的调查数据, 它们的平均就相当于期望, 它们依大小序排的中间那个数据就相当于中位数. 例如, 调查一个小区内的家庭收入, 如果马云搬进来了, 那么小区平均收入瞬间上升几百万级, 但中位数收入不会有太大变化.

最后我们要提到的一个数字特征是描述两个随机变量之间的关系度的. 设 ξ, η 是同一个样本空间上两个随机变量, 它们之间当然有很多中关系, 一种是确定性关系, 即一个随机变量是另外一个的函数; 另一种是独立关系, 互不影响, 定义为: 对任何 $x, y \in \mathbf{R}$ 有

$$\mathbb{P}(\xi \leq x, \eta \leq y) = \mathbb{P}(\xi \leq x)\mathbb{P}(\eta \leq y).$$

联合分布函数是边缘分布函数的 '积'. 其它就介于两者之间, 随机关系: 有关系但不确定. 相关系数定义为

$$\rho(\xi, \eta) := \frac{\mathbb{E}[\xi\eta] - \mathbb{E}\xi\mathbb{E}\eta}{\sqrt{D\xi \cdot D\eta}}.$$

当 $\rho(\xi, \eta) = 0$ 时, 称为两者不相关, 但不要把不相关理解为独立的意思, 两者差别很大. 例如, ξ 服从标准正态分布, $\eta = \xi^2$, 这是确定性关系, 但它们按定义是不相关的, 因为 $\mathbb{E}[\xi\eta] = \mathbb{E}\xi = 0$. 是不是很奇怪? 也不奇怪, 实际上相关系数只能从某种意义上描述两者的线性关系程度, 不能表达其他关系, 实际上, 当 $\rho(\xi, \eta) = 1$ 或 -1 时, 那么两者是确定的线性关系.

§2.3 特征函数

数字特征能够体现随机变量的分布的一些特征, 但不足以标识出分布, 因为分布是一个函数, 是无穷维的. 在这一节中, 我们要引入特征函数, 特征函数是一种方法, 是分析中的重要工具. 设随机变量 ξ 的分布函数是 F, 定义它的特征函数

$$\phi_\xi(x) := \mathbb{E}[e^{ix\xi}] = \int e^{ixy} dF(y), \ x \in \mathbf{R}.$$

特征函数也是分布函数的特征, 因为它实际上只依赖于分布函数, 同分布的随机变量有相同的特征函数. 因为 $e^{ix\xi}$ 是一个复值随机变量, 算它的期望就要算其实部和虚部的期望, 故特征函数是一个定义在实数域上的复值函数, 且有

$$|\phi_\xi(x)| = |\mathbb{E}[e^{ix\xi}| \leq 1.$$

请读者自己验证: 对于一个复值随机变量 X 也一样有以下不等式:

$$|\mathbb{E}[X]| \leq \mathbb{E}|X|.$$

特征函数有以下性质:

1. 特征函数连续;
2. 特征函数在零点处总是等于 1;
3. 如果 $\mathbb{E}|\xi|^n < \infty$, 那么其特征函数 n 次可导且
$$\phi_\xi^{(n)}(0) = i^n \mathbb{E}\xi^n;$$
4. 如果 ξ 与 η 是独立的两个随机变量, 那么它们和 $\xi+\eta$ 的特征函数等于特征函数乘积, 即
$$\phi_{\xi+\eta}(x) = \phi_\xi(x)\phi_\eta(x),\ x \in \mathbf{R}.$$

第四个性质特别重要, 是特征函数成为分析的重要工具的主要原因. 我们知道两个独立随机变量的和的分布函数是两个分布函数的卷积, 即

$$F_{\xi+\eta} = F_\xi * F_\eta.$$

怎么验证呢? 前面提到独立是指对任何 $x, y \in \mathbf{R}$ 有

$$\mathbb{P}(\xi \leq x, \eta \leq y) = \mathbb{P}(\xi \leq x)\mathbb{P}(\eta \leq y),$$

用期望的方式写出来

$$\mathbb{E}[1_{(-\infty,x]}(\xi)1_{(-\infty,y]}(\eta)] = \mathbb{E}[1_{(-\infty,x]}(\xi)]\mathbb{E}[1_{(-\infty,y]}(y)].$$

由单调类方法推出对任何连续函数 f, g 有

$$\mathbb{E}[f(\xi)g(\eta)] = \mathbb{E}[f(\xi)]\mathbb{E}[g(\eta)].$$

因此第四个性质显然是成立的, 因为只要取 $f(y) = e^{ixy}$ 与 $g(y) = e^{ixy}$ 就可以了.

特征函数有用的主要原因是它有两个性质: 唯一性和连续性. 唯一性是指特征函数可以唯一地标识分布函数, 连续性是指特征函数的收敛可以推出分布函数的收敛.

定理 2.3.1 特征函数唯一决定分布函数.

由于特征函数的性质四, 它通常比分布函数容易计算, 而分布函数被特征函数决定, 所以只要能算出特征函数也就够了. 例如, 参数为 λ 的 Poisson 分布的特征函数是

$$\phi(x) = \exp(-\lambda(1 - e^{ix})),$$

然后由性质四, 两个独立的参数分别为 λ_1, λ_2 的 Poisson 随机变量的和的特征函数是

$$\exp(-(\lambda_1 + \lambda_2)(1 - e^{ix})),$$

由唯一性推出其分布是参数为 $\lambda_1 + \lambda_2$ 的 Poisson 分布.

特征函数的收敛是指点点收敛, 但是特征函数列点点收敛的极限不一定仍然是特征函数. 例如, 如果 ϕ_n 是正态分布 $N(0, n^2)$ 的特征函数, 那么

$$\phi_n(x) = \exp(-x^2n^2/2),$$

显然它的点点收敛的极限是除 0 处等于 1 其它地方等于 0 的函数, 不连续, 所以不可能是特征函数. 分布函数列 F_n 的收敛是指它在 $\mathbf{R}$ 的

一个稠密子集上点点收敛, 称为是弱收敛, 它的极限也不一定是分布函数, 例如 $N(0,n^2)$ 的极限就不是分布函数. 弱收敛于一个分布函数的分布函数列对应的随机变量列称为是依分布收敛.

定理 2.3.2 如果特征函数列 ϕ_n 收敛于一个在零点连续的函数 ϕ, 那么 ϕ 是特征函数, 而且 ϕ_n 对应的分布函数列也收敛于 ϕ 对应的分布函数.

这个定理的证明这里不讲了, 读者可以在其他概率论教材中找到. 最后我们用它来证明中心极限定理.

定理 2.3.3 设 $\{\xi_n\}$ 是独立同分布平方克积随机序列, 其平均 $X_n := \sum_{j=1}^n \xi_j/n$ 的标准化依分布收敛于标准正态分布.

不失一般性, 设 ξ_n 的期望为零, 方差为 1, 那么 X_n 的标准化是 $X_n/\sqrt{n}$. 由以上连续性定理, 只需证明 $X_n/\sqrt{n}$ 的特征函数点点收敛于标准正态分布的特征函数. 而由特征函数性质的第四条, 独立和的特征函数是特征函数乘积, 因此 $X_n/\sqrt{n}$ 的特征函数是

$$\phi_n(x) = \mathbb{E}\{\exp[ix(\xi_1 + \cdots + \xi_n)/\sqrt{n}]\} = \phi(x\sqrt{n})^n,$$

其中 ϕ 是 ξ_n 的特征函数, 因为 ξ_n 的期望为 0 方差为 1, 所以可 Taylor 展开为

$$\phi(x) = 1 - \frac{1}{2}x^2 + g(x),$$

其中 $g(x)$ 是 x^2 的高价无穷小. 因此

$$\phi_n(x) = \left(1 - \frac{1}{2}\frac{x^2}{n} + g(x/\sqrt{n})\right)^n \to \exp(-\frac{1}{2}x^2),$$

而最右边恰好是标准正态分布的特征函数.

习题

1. 如果 ξ 的分布函数 F 是连续的, 证明: $F(\xi)$ 是均匀分布的.

2. 设 ξ 是连续型的, 可积, 且函数 $f(x)=\mathbb{E}|\xi-x|$ 在 $x=m$ 达到最小值. 证明:
$$\mathbb{P}(\xi<m)=1/2.$$

3. 计算参数为 α 的指数分布的期望和中位数.

4. 计算均匀分布与指数分布的特征函数.

5. 如果 ξ 与 $-\xi$ 同分布, 我们就说 ξ 是对称分布的. 证明: ξ 是对称分布的充分必要条件是其特征函数是实值函数.

第三章　条件期望 1

§3.1　极限与期望的交换

给定概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$, 样本空间上一个 $\mathscr{F}$ 可测的函数称为随机变量. 设 $\{\xi_n\}$ 是个收敛的随机 (变量) 序列, 是否有

$$\mathbb{E}[\lim_n \xi_n] = \lim_n \mathbb{E}\xi_n?$$

这个问题就是问极限和期望是否能够交换, 是一个重要的问题. 一般来说, 他们不能交换. 要说清楚这个问题, 要先介绍随机变量列的收敛. 随机变量是样本空间上的函数, 函数的收敛有很多种, 首先是点点收敛或者处处收敛, 即对任何 $\omega \in \Omega$ 有 $\xi_n(\omega) \to \xi(\omega)$. 但是这个要求太强, 我们更趋向于用几乎处处收敛的概念, 即收敛的概率是 1,

$$\mathbb{P}(\{\omega \in \Omega : \lim_n \xi_n(\omega) = \xi(\omega)\}) = 1.$$

另外还有两种经常使用的收敛是依概率收敛和距离收敛, 依概率收敛是指对任何 $\varepsilon > 0$ 有

$$\lim_n \mathbb{P}(|\xi_n - \xi| > \varepsilon) = 0.$$

依概率收敛要弱于几乎处处收敛, 即几乎处处收敛蕴含依概率收敛. 反过来不一定, 但可以由 Borel-Cantelli 引理推出一定有一个子列几乎处处收敛. 这两种收敛是最常用的. 距离收敛是指按照某种距离意义下收敛, 典型的距离是 L^1 与 L^2 距离, 两个可积随机变量 ξ, η 的 L^1 距离是 $\mathbb{E}|\xi - \eta|$, 如果它们平方可积, 它们的 L^2 距离是 $\{\mathbb{E}[(\xi - \eta)^2]\}^{1/2}$. 由 Cauchy-Schwarz 不等式推出 L^1 距离不会超过 L^2 距离. ξ_n L^1- 收敛于 ξ 是指 $\lim_n \mathbb{E}|\xi_n - \xi| = 0$, 类似地有 L^2- 收敛, 显然 L^2- 收敛蕴含 L^1- 收敛. 由 Chebyshev 不等式,

$$\mathbb{P}(|\xi_n - \xi| > \varepsilon) \le \frac{1}{\varepsilon^2}\mathbb{E}[(\xi_n - \xi)^2],$$

距离收敛可以推出依概率收敛, 但与几乎处处收敛之间没有蕴含关系.

这三种收敛就是常见的随机序列收敛, 我们来看它们能不能与期望交换, 首先, 因为

$$|\mathbb{E}\xi_n - \mathbb{E}\xi| \le \mathbb{E}|\xi_n - \xi|,$$

所以距离收敛是可以推出期望也是收敛的. 但是几乎处处收敛不一定能推出期望收敛, 例如设 η 是 $[0,1]$ 上均匀分布随机变量, 令

$$\xi_n = n1_{\{\eta \le 1/n\}},$$

那么 ξ_n 几乎处处收敛于 0, 但是 $\mathbb{E}\xi_n = 1$, 所以期望与极限不能交换.

用一个实际投资的例子看, 期望极限和极限期望差别是很大的. 取两个数, 算术平均大于 1, 几何平均小于 1, 例如 $x = 0.7$, $y = 1.4$, 那么 $xy = 0.98$, 令 $\xi_1, \cdots, \xi_n, \cdot$ 是独立同分布随机序列,

$$\mathbb{P}(\xi_1 = x) = \mathbb{P}(\xi_1 = y) = 1/2,$$

把 $\xi_n - 1$ 看成为投资收益, $X_n = \xi_1 \cdots \xi_n$ 看成为投资所得, 那么期望 $\mathbb{E}[X_n] = [\mathbb{E}\xi_1]^n$ 趋于无穷, 因为 $\mathbb{E}[\xi_1] = 1.05$. 而

$$X_n = \exp(\log \xi_1 + \cdots + \log \xi_n),$$

由大数定律 X_n 几乎处处趋于 0, 因为 $\mathbb{E}[\log \xi_1] < 0$.

那什么条件下两者可以交换呢? 先设 ξ_n 几乎处处收敛于 ξ, 满足下面两个条件之一就可以: (1) ξ_n 几乎处处非负单调, 即

$$\mathbb{P}(\xi_n \ge 0, \forall n \ge 1; \xi_n \uparrow \xi) = 1;$$

(2) ξ_n 被可积随机变量控制, 即存在可积的 η 使得

$$\mathbb{P}(|xi_n| \le \eta, \forall n) = 1.$$

第一个称为 Lévy 单调收敛定理, 第二个称为 Lebesgue 控制收敛定理. 其实中间还有一个 Fatou 引理: 如果 $\{\xi_n\}$ 是非负随机序列, 那么

$$\mathbb{E}[\liminf_n \xi_n] \le \liminf_n \mathbb{E}\xi_n,$$

其中随机序列的下极限是点点意义下而言.

单调收敛定理来自于概率的下连续性: 如果 $A_n \uparrow A$, 则 $\mathbb{P}(A_n) \uparrow \mathbb{P}(A)$. Fatou 引理来自来自单调收敛定理, 并推出 Lebesgue 控制收敛定理.

§3.2 单调类方法

两个随机变量 ξ, η 独立, 是指对任何 $x, y \in \mathbb{R}$,

$$\mathbb{P}(\xi \le x, \eta \le y) = \mathbb{P}(\xi \le x)\mathbb{P}(\eta \le y).$$

用期望来写

$$\mathbb{E}[1_{(-\infty,x]}(\xi)1_{(-\infty,y]}(\eta)] = \mathbb{E}[1_{(-\infty,x]}(\xi)]\mathbb{E}[1_{(-\infty,y]}(\eta)].$$

用 $\mathscr{G}$ 表示使得

$$\mathbb{E}[1_A(\xi)1_{(-\infty,y]}(\eta)] = \mathbb{E}[1_A(\xi)]\mathbb{E}[1_{(-\infty,y]}(\eta)]$$

成立的 Borel 子集 A 的全体, 那么 $\mathscr{G}$ 是一个包含所有区间 $(-\infty, x]$ 的单调类, 单调类定理告诉我们 $\mathscr{G}$ 包含由这些区间生成的事件域, 也就是 Borel 域. 然后由期望性质推出对任何简单函数 $f : \mathbb{R} \mapsto \mathbb{R}$ 有

$$\mathbb{E}[f(\xi)1_{(-\infty,y]}(\eta)] = \mathbb{E}[f(\xi)]\mathbb{E}[1_{(-\infty,y]}(\eta)].$$

因为任何非负可测函数都可以表示为递增的非负简单函数列的极限, 然后应用单调收敛定理得上式对所有非负 Borel 可测函数 f 成立, 实际上也对所有可测函数 f 成立. 然后固定 f, 对 η 做一样的事情, 可以将示性函数用可测函数 g 代替. 这个证明过程是一个程序, 通常称为单调类方法. 单调类方法是概率论的主要方法, 在本课程中会时常用到.

§3.3 条件期望的直观和定义

在这一节中, 我们将介绍随机变量关于部分信息的条件期望. 首先我们来看个特殊情况, 设 $(\Omega, \mathscr{F}, \mathbb{P})$ 是概率空间, 随机变量 ξ 平方可积, $\mathscr{F}$ 代表全部信息, 它的子事件域 $\mathscr{G}$ 代表部分信息. 直观上来说, ξ 关于 $\mathscr{G}$ 的条件期望是在已知部分信息 $\mathscr{G}$ 的条件下对 ξ 的最佳预测. 数学上如何表达最佳的意思呢? 就用距离最近的方式. 先看期望 $\mathbb{E}[\xi]$, 它是实数空间到 ξ 距离最近的一个数, 即函数

$$x \mapsto \mathbb{E}[(x-\xi)^2]$$

在 $x = \mathbb{E}\xi$ 时达到最小值.

让我们用 Hilbert 空间的眼光看条件期望的问题, 平方可积随机变量全体是一个 Hilbert 空间记为

$$L^2(\Omega, \mathscr{F}, \mathbb{P}),$$

其中两个随机变量 ξ, η 之间的内积定义为 $\mathbb{E}[\xi\eta]$, 而关于部分信息 $\mathscr{G}$ 可测的平方可积随机变量全体是以上空间的闭子空间, 记为 $L^2(\Omega, \mathscr{G}, \mathbb{P})$. 从数学上说, 条件期望就是空间 $L^2(\Omega, \mathscr{G}, \mathbb{P})$ 中距离 ξ 最近的那个随机变量, 由 Hilbert 空间理论, 因为子空间闭, 所以的确存在唯一的随机变量 η 使得

$$\mathbb{E}[(\xi-\eta)^2] = \min\{\mathbb{E}[(\xi-\zeta)^2] : \zeta \in L^2(\Omega, \mathscr{G}, \mathbb{P})\},$$

即最小值点是唯一存在的, 记为 $\mathbb{E}[\xi|\mathscr{G}]$.

虽然理论上唯一存在, 但具体怎么计算是另外一个问题. 我们现在给出一个更加容易计算的等价定义. 这需要从最近这个议题展开, 在 Hilbert 空间理论中, η 是子空间中离 ξ 最近的地方等价于说 η 是 ξ 在子空间上的正交投影, 即 $\xi - \eta$ 与子空间垂直或者说正交, 也就是和子

空间中所有向量正交, 那也就是说, 对任何 $\zeta \in L^2(\Omega, \mathscr{G}, \mathbb{P})$ 有

$$\mathbb{E}[(\xi - \eta)\zeta] = 0.$$

因为子集

$$\{1_A : A \in \mathscr{G}\}$$

张成的线性空间在 $L^2(\Omega, \mathscr{G}, \mathbb{P})$ 稠密, 所以正交性等价于对任何 $A \in \mathscr{G}$ 有

$$\mathbb{E}[\xi 1_A] = \mathbb{E}[\eta 1_A].$$

我们经常记 $\mathbb{E}[\xi 1_A]$ 为 $\mathbb{E}[\xi; A]$. 它可以作为条件期望的定义.

定义 3.3.1 设 ξ 是可积随机变量, $\mathscr{G}$ 是子事件域, ξ 关于 $\mathscr{G}$ 的条件期望是唯一的一个满足下面条件的随机变量 η:

(1) η 是关于 $\mathscr{G}$ 可测的;

(2) 对任何 $A \in \mathscr{G}$ 有

$$\mathbb{E}[\eta; A] = \mathbb{E}[\xi; A].$$

首先看最简单的一个例子, 还是上面的概率空间和可积随机变量 ξ, $\{\Omega_k : 1 \le k \le n\}$ 是 Ω 的可测分类, 即他们互斥且

$$\Omega_1 \cup \cdots \cup \Omega_n = \Omega.$$

$\mathscr{G}$ 是 $\{\Omega_k : 1 \le k \le n\}$ 生成的事件域. 那么他们都是 $\mathscr{G}$ 的原子, 我们来看条件期望 $\mathbb{E}[\xi|\mathscr{G}]$. 可以假设他们都是正概率的, 否则可以把概率 0 的事件扔掉. 定义的第一条推出 $\mathbb{E}[\xi|\mathscr{G}]$ 在原子 Ω_k 上是常数. 问题是这个常数是怎么确定的呢? 是由定义的第二条确定的, 令常数是 x, 那么由第二条

$$\mathbb{E}[\mathbb{E}[\xi|\mathscr{G}]; \Omega_k] = \mathbb{E}[\xi; \Omega_k],$$

所以

$$x\mathbb{P}(\Omega_k) = \mathbb{E}[\xi; \Omega_k],$$

推出

$$x = \frac{\mathbb{E}[\xi; \Omega_k]}{\mathbb{P}(\Omega_k)} = \mathbb{E}[\xi|\Omega_k],$$

最后的记号是仿照条件概率的记号, 因为

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{E}[1_A; B]}{\mathbb{P}(B)}.$$

这个值称为局部平均或者积分平均, 类似于函数 $y = f(x)$ 在区间 $[a, b]$ 上的积分平均

$$\frac{1}{b-a}\int_a^b f(x)dx.$$

因此,

$$\mathbb{E}[\xi|\mathscr{G}] = \sum_{k=1}^{n} \mathbb{E}[\xi|\Omega_k] 1_{\Omega_k},$$

条件期望在每个原子上是常熟, 等于在原子上的局部平均. 显然

$$\mathbb{E}[\mathbb{E}[\xi|\mathscr{G}]] = \mathbb{E}\xi,$$

即条件期望的期望就是期望.

§3.4 条件期望的性质 I

在本节中, 我们将列出条件期望的几个非常重要的性质.

(1) 条件期望 $\xi \mapsto \mathbb{E}[\xi|\mathscr{G}]$ 是线性算子, 即

$$\mathbb{E}[c_1\xi_1 + c_2\xi_2|\mathscr{G}] = c_1\mathbb{E}[\xi_1|\mathscr{G}] + c_2\mathbb{E}[\xi_2|\mathscr{G}].$$

这是因为条件期望作为算子是投影算子, 当然是线性算子.

(2) $\mathbb{E}[\xi|\{\emptyset, \Omega\}] = \mathbb{E}[\xi]$.

(3) 如果 ξ 是 $\mathscr{G}$ 可测的, 那么 $\mathbb{E}[\xi|\mathscr{G}] = \xi$. 这是因为投影算子在闭子空间自身上是恒等算子.
(4) 条件期望有保正性: 如果 $\xi \geq 0$ a.s., 那么 $\mathbb{E}[\xi|\mathscr{G}] \geq 0$ a.s. 事实上, 由定义的第二条推出, 对任何 $A \in \mathscr{G}$ 有 $\mathbb{E}[\mathbb{E}[\xi|\mathscr{G}]; A] \geq 0$, 因此 $\mathbb{E}[\xi|\mathscr{G}] \geq 0$.
(5) 塔性: 如果 $\mathscr{G}_1 \subset \mathscr{G}_2$ 是两个子事件域, 那么

$$\mathbb{E}[\mathbb{E}[\xi|\mathscr{G}_2]|\mathscr{G}_1] = \mathbb{E}[\xi|\mathscr{G}_1] = \mathbb{E}[\mathbb{E}[\xi|\mathscr{G}_1]|\mathscr{G}_2].$$

第一个等号实际上还是投影算子的性质, 因为 $L^2(\Omega, \mathscr{G}_1, \mathbb{P})$ 是 $L^2(\Omega, \mathscr{G}_2, \mathbb{P})$ 的闭子空间, ξ 在大的子空间上投影然后再投影到小的子空间上等于它直接投影到小的子空间上. 第二个等号从性质 (3) 推出. 这个公式的一个特殊情况是 $\mathscr{G}_1$ 平凡时

$$\mathbb{E}[\mathbb{E}[\xi|\mathscr{G}]] = \mathbb{E}\xi.$$

习题

1. 证明: 依概率收敛蕴含着依分布收敛.

2. (Borel-Cantelli) 证明: 如果 $\sum_n \mathbb{P}(A_n) < \infty$, 则 $\mathbb{P}(\limsup A_n) = 0$.

3. 设 X_n 服从期望为 0 方差为 n 的正态分布. 证明: 对任何 $a > 1/2$, 有 $n^{-a}X_n$ 几乎处处趋于 0. 提示: 利用 Borel-Cantelli 引理. 再证明: $\sqrt{n}X_n$ 不是依概率收敛的.

4. 设 η 是非负随机变量且 $\mathbb{E}\eta = 1$, 定义 $\tilde{\mathbb{P}}(A) := \mathbb{E}[\eta; A]$, $A \in \mathscr{F}$. 证明: (1) $\tilde{\mathbb{P}}$ 也是概率测度; (2) 如果 ξ 是非负随机变量, 那么

$$\tilde{\mathbb{E}}[\xi] = \mathbb{E}[\eta\xi],$$

其中 $\tilde{\mathbb{E}}$ 是概率 $\tilde{\mathbb{P}}$ 的期望. 提示: 应用单调类方法.

5. 证明: 单调收敛定理, Fatou 引理和 Lebesgue 控制收敛定理互相是等价的.

6. 证明条件期望的性质 (1): 线性.

7. 如果 ξ 是 $\mathscr{G}$ 可测的且对任何 $A \in \mathscr{G}$ 有 $\mathbb{E}[\xi; A] \geq 0$, 证明: $\mathbb{P}(\xi \geq 0) = 1$.

第四章　条件期望 2

§4.1　条件期望的性质 II

(6) 如果 η 是 $\mathscr{G}$ 可测的随机变量, 那么

$$\mathbb{E}[\xi\eta|\mathscr{G}] = \eta\mathbb{E}[\xi|\mathscr{G}].$$

要证明这个性质, 可以直接用定义. 因为右边是 $\mathscr{G}$ 可测的, 所以要证明右边是 $\xi\eta$ 关于 $\mathscr{G}$ 的条件期望, 只需要证明对任何 $A \in \mathscr{G}$ 有

$$\mathbb{E}[\eta\mathbb{E}[\xi|\mathscr{G}]1_A] = \mathbb{E}[\xi\eta 1_A]$$

就足够了. 这等价于

$$\mathbb{E}[(\xi - \mathbb{E}[\xi|\mathscr{G}])\eta 1_A] = 0.$$

而 $\eta 1_A$ 是 $\mathscr{G}$ 可测的, 即属于 $L^2(\Omega, \mathscr{G}, \mathbb{P})$, 所以与 $\xi - \mathbb{E}[\xi|\mathscr{G}]$ 垂直.
(7) 如果 ξ 与 $\mathscr{G}$ 独立, 那么 $\mathbb{E}[\xi|\mathscr{G}] = \mathbb{E}\xi$. 也就是说, $\mathscr{G}$ 代表的信息对于预测 ξ 没有任何作用. 随机变量 ξ 与 $\mathscr{G}$ 独立是指对任何 $x \in \mathbb{R}$, $A \in \mathscr{G}$ 有

$$\mathbb{P}(\xi \le x, A) = \mathbb{P}(\xi \le x)\mathbb{P}(A).$$

这蕴含着

$$\mathbb{E}[\xi 1_A] = \mathbb{E}\xi \cdot \mathbb{E}1_A = \mathbb{E}[\mathbb{E}\xi 1_A],$$

由定义, 这证明了右边是条件期望.

§4.2　条件期望的计算

我们来计算复杂一点的条件期望. 如果 ξ, η 是两个随机变量, 我们用 $\mathbb{E}[\xi|\eta]$ 表示条件期望 $\mathbb{E}[\xi|\sigma(\eta)]$. 注意定义的第一条告诉我们它关

于 $\sigma(\eta)$ 可测, 这当且仅当它是 η 的函数, 即存在Borel 可测函数 ϕ 使得 $\mathbb{E}[\xi|\eta] = \phi(\eta)$, 我们只要把 ϕ 表达式算出来就可以了. 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 $y \in \mathbb{R}$, 有

$$\mathbb{E}(\phi(\eta); \eta \le y) = \mathbb{E}(\xi; \eta \le y).$$

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

$$\int_{-\infty}^{y} \phi(t) \mathrm{d}F_\eta(t) = \int_{\mathbb{R}} \int_{-\infty}^{y} s \mathrm{d}F(s,t).$$

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

$$\int_{-\infty}^{y} \phi(t) f_\eta(t) \mathrm{d}t = \int_{\mathbb{R}} \int_{-\infty}^{y} s f(s,t) \mathrm{d}s \mathrm{d}t.$$

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

$$\phi(y) f_\eta(y) = \int_{\mathbb{R}} s f(s,y) \mathrm{d}s,$$

因此在密度正的地方, 也就是在 η 的值域范围内有

$$\phi(y) = \int_{\mathbb{R}} s \cdot \frac{f(s,y)}{f_\eta(y)} \mathrm{d}s.$$

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 $\mathbb{E}(\xi|\eta = y)$ 表示. 尽管这个符号在 $\mathbb{P}(\eta = y) = 0$ 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):

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

因为

$$\mathbb{E}(\xi|\eta\in(y-\delta,y+\delta))=\frac{\displaystyle\int_{\mathbb{R}}\mathrm{d}s\int_{(y-\delta,y+\delta)}sf(s,t)\mathrm{d}t}{\displaystyle\int_{(y-\delta,y+\delta)}f_\eta(t)\mathrm{d}t}$$
$$=\frac{\displaystyle\int_{\mathbb{R}}\mathrm{d}s\frac{1}{2\delta}\int_{(y-\delta,y+\delta)}sf(s,t)\mathrm{d}t}{\displaystyle\frac{1}{2\delta}\int_{(y-\delta,y+\delta)}f_\eta(t)\mathrm{d}t},$$

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

例 4.2.1 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 $\mathbb{E}(\xi|\eta)$. 先算条件密度, 有

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

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

$$\mathbb{E}(\xi|\eta)=\int_{\mathbb{R}}xf_{\xi|\eta}(x|\eta)\mathrm{d}x=r\eta.$$

对于联合正态分布随机变量, 它们独立等价于它们不相关, 显然 η 与 $\xi-r\eta$ 的协方差为零, 是不相关的, 也是独立的, 由条件期望性质得

$$\mathbb{E}[\xi|\eta]=\mathbb{E}[\xi-r\eta|\eta]+r\eta=r\eta.$$

这个方法当然很特殊, 只能由于正态. 对于一般二维正态分布随机变量 (ξ_1,ξ_2) 它们的期望分别是 μ_1,μ_2, 方差分别是 σ_1^2,σ_2^2, 相关系数还是 r,

那么

$$\mathbb{E}\left[\frac{\xi_1-\mu_1}{\sigma_1}|\frac{\xi_2-\mu_2}{\sigma_2}\right]=r\frac{\xi_2-\mu_2}{\sigma_2},$$

因为 $(\xi_2-\mu_2)/\sigma_2$ 与 ξ_2 可以互相表示, 再利用性质得

$$\mathbb{E}[\xi_1|\xi_2]=r\sigma_1\frac{\xi_2-\mu_2}{\sigma_2}+\mu_1.$$

由此看到, 学会应用条件期望性质很重要. ∎

§4.3 公平游戏与鞅

什么是随机过程? 随机过程是按时间记录的随机变量族. 通常分两类, 连续时间和离散时间. 连续时间是指 $\mathbb{R}$ 的一个区间, 离散时间是指一个连续整数集, 后者常称为随机序列. 固定概率空间 $(\Omega,\mathscr{F},\mathbb{P})$, 例如对任何 $t\geq 0$, X_t 是随机变量, 那么 $(X_t:t\geq 0)$ 称为是随机过程; 对任何整数 $n\geq 0$, X_n 是随机变量, 那么 $(X_n:n\geq 0$ 是随机序列. 在这一讲中, 我们先讨论随机序列. 对任何 $\omega\in\Omega$, $\{X_n(\omega):n\geq 0\}$ 是一个数列, 称为是样本点 ω 的轨道, 样本点可以等同于它的轨道, 称为样本轨道. 随机过程通常是讨论样本轨道的概率性质.

最简单的随机序列是随机游动, 设 $\xi_1,\xi_2,\cdots,\xi_n,\cdots$ 是独立随机序列 (即其中任何有限个随机变量是相互独立的), 令

$$X_0:=x, X_n:=X_{n-1}+\xi_n,\ n\geq 1,$$

它称为是 x 出发的随机游动. ξ_n 可以想象为某个赌徒在第 n 局赌博中的输赢结果, x 是他带来的赌资, 那么 (X_n) 记录了他的财富变化, 是非常直观的随机序列. 一个赌徒坐下来赌一天的财富变化序列就是一个样本轨道.

再简单一点, 设赌博就是输赢为一块钱的掷正面概率为 p 的硬币, 那么

$$\mathbb{P}(\xi_n=1)=p,\ \mathbb{P}(\xi_n=-1)=q=1-p.$$

对应的 (X_n) 称为是从 x 出发的简单随机游动. 显然 $\mathbb{E}\xi_n = p-q = 2p-1$, 所以由强大数定律, 当 $p > 1/2$ 时, X_n 几乎处处趋于无穷, 也就是说它的几乎所有样本轨道趋于正无穷, 同理当 $p < 1/2$, 它的几乎所有样本轨道趋于负无穷. 但是 $p = 1/2$ 时, 我们无法从强大数定律看出什么结论.

不妨设出发点是零. 我们来证明它的几乎所有样本轨道上下都是无界的, 即

$$\mathbb{P}(\sup_n X_n = +\infty, \inf_n X_n = -\infty) = 1.$$

只要证明

$$\mathbb{P}(\sup_n X_n = +\infty) = 1$$

也就够了. 由 Kolmogorov 0-1 律, 尾事件域

$$\bigcap_{n\geq 1} \sigma(\xi_k : k > n)$$

中的事件的概率不是 0 就是 1. 尾事件域中的事件实际上是与前面任意有限个都无关的事件, 而 $\{\sup_n X_n = +\infty\}$ 这个事件发生与否与 $(\xi_n : n \geq 1)$ 的前面任何有限个是没有关系的, 所以它的概率不是 0 就是 1, 我们只需证明概率不是 0 就好.

证明概率不是 0 需要用中心极限定理, 即以下结论

$$\lim \mathbb{P}(\frac{X_n}{\sqrt{n}} > 1) = \int_1^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt > 0.$$

而

$$\limsup_n \{X_n > \sqrt{n}\} \subset \{\sup_n X_n = +\infty\},$$

且由 Fatou 引理

$$\mathbb{P}(\limsup_n \{X_n > \sqrt{n}\}) \geq \lim_n \mathbb{P}(X_n > \sqrt{n}) > 0,$$

推出

$$\mathbb{P}(\sup_n X_n = \infty) > 0.$$

证明怎么一个简单的结论也不是很容易吧! 但是我们要介绍的鞅与鞅就给我们一个系统的方法处理这样的问题, 鞅理论是随机分析的基础. 鞅的概念是简单随机游动的抽象, 鞅表达的是公平游戏, 简单随机游动当 $p = 1/2$ 时是公平游戏, 这时

$$\mathbb{E}[X_{n+1} - X_n | X_1, \cdots, X_n] = 0.$$

定义 4.3.1 (可积) 随机序列 $(X_n : n \geq 0)$ 是鞅, 如果对任何 n 有

$$\mathbb{E}[X_{n+1} | X_1, \cdots, X_n] = X_n.$$

如果只是大于等于号成立, 成为下鞅; 如果只是小于等于号成立, 称为上鞅.

鞅的直观意思是对财富增量的 (给定信息下) 预期是零. 如果 (X_n) 是鞅, 那么定义式两边取期望, 推出 $\mathbb{E}[X_{n+1}] = \mathbb{E}[X_n]$, 所以鞅的期望是不变的. 类似地, 下鞅期望是递增的, 趋势变好; 上鞅期望是递减的, 趋势转坏.

为了方便, 我们介绍信息流的概念, 用信息流来定义鞅. 如果有时间的概念, 那么信息总是随着时间而增加的, 就是所谓的信息流. 一个递增的子事件域列 $(\mathscr{G}_n)$ 称为是信息流, 或者简单称为流. 一个随机序列 (X_n) 称为适应于流 $(\mathscr{G}_n)$, 如果对任何 n, X_n 关于 $\mathscr{G}_n$ 可测. 随机序列 (X_n) 本身诱导一个自然的流

$$\mathscr{G}_n^X := \sigma(X_1, \cdots, X_n),$$

序列 (X_n) 关于其自然流是适应的, 自然流是适应流中最小的. 直观地看, 四个人搓麻将, 每个人的财富序列关于自己输赢结果的流是适应的, 关于四个人总体的的输赢结果的流也是适应的, 但是每个赌徒收到的

信息绝对不止这些, 他们可能看到旁边一桌麻将的输赢, 他们也不时地看看电视, 看看微博等, 这些信息可以都汇总在一起组成一个流. 很多信息和他的财富是无关的.

现在给鞅一个升级版的定义.

定义 4.3.2 设有流 $(\mathscr{G}_n)$, (可积) 随机序列 $(X_n : n \geq 0)$ 称为关于 $(\mathscr{G}_n)$ 是鞅或者 $(\mathscr{G}_n)$ 鞅, 如果

(1) (X_n) 适应于流 $(\mathscr{G}_n)$;

(2) 对任何 n 有

$$\mathbb{E}[X_{n+1}|\mathscr{G}_n] = X_n.$$

因为 $\mathscr{G}_n^X \subset \mathscr{G}_n$, 所以在 (2) 的两边取关于 $\mathscr{G}_n^X$ 的条件期望就推出: 如果 (X_n) 关于一个流是鞅, 那么它一定是鞅. 或者说它是鞅当且仅当它关于某个信息流是鞅. 在这个定义下, 同一个信息流之下的鞅是线性空间.

习题

1. 设 ξ, η 独立且 $\mathbb{E}[\eta] = 0$. 证明:

$$\mathbb{E}|\xi| \leq \mathbb{E}|\xi + \eta|.$$

2. 设 $\mathbb{E}[X|Y] = Y$, $\mathbb{E}[Y|X] = X$. 证明: $X = Y$ a.s.

3. 设 X, Y 独立同分布. 证明: $\mathbb{E}[X|X+Y] = \mathbb{E}[Y|X+Y] = (X+Y)/2$.

4. 令 $X_n := \mathbb{E}[X|\mathscr{G}_n]$. 证明: (X_n) 是 $(\mathscr{G}_n)$ 鞅.

5. 设 $\xi_1, \cdots, \xi_n, \cdots$ 是独立同分布随机序列,

$$\mathbb{P}(\xi_1 = 1/2) = \mathbb{P}(\xi_1 = 3/2) = 1/2.$$

令 $X_n = \xi_1 \xi_2 \cdots \xi_n$. 证明: (X_n) 是鞅, 但是 X_n 几乎处处趋于 0. 提示: 利用强大数律.

第五章　鞅与 Doob 基本定理

§5.1　Jensen 不等式

再补充一个不等式, 称为 Jensen 不等式. $\mathbb{R}$ 上的一个函数 f 称为凸函数, 如果对任何 $x_1, x_2 \in \mathbb{R}$ 有

$$f\left(\frac{x_1+x_2}{2}\right) \le \frac{f(x_1)+f(x_2)}{2}.$$

常见的凸函数有 $y=x^2$, $y=|x|$, $y=x^+=x\vee 0$, $y=e^x$ 等等. 凸函数是说平均之后取函数值不会超过先取函数值再平均. 所以可以想象有下面的不等式: 如果 ξ 是随机变量, f 是凸函数, 那么

$$f(\mathbb{E}\xi) \le \mathbb{E}f(\xi).$$

称为 Jensen 不等式.

条件期望也有类似不等式,

$$f(\mathbb{E}[\xi|\mathscr{G}]) \le \mathbb{E}[f(\xi)|\mathscr{G}].$$

怎么证明呢? 有一个很巧妙的想法. 首先凸函数在任何点 $(x_0, f(x_0))$ 的切线肯定在函数以下, 即

$$f'(x_0)(x-x_0)+f(x_0) \le f(x).$$

以上不等式对任何 x_0, x 恒成立, 所以令 $x=\xi$, $x_0=\mathbb{E}[\xi|\mathscr{G}]$ 代入得

$$f'(\mathbb{E}[\xi|\mathscr{G}])(\xi-\mathbb{E}[\xi|\mathscr{G}])+f(\mathbb{E}[\xi|\mathscr{G}]) \le f(\xi),$$

两边取条件期望, 由条件期望的单调性和性质 (3) 与 (6) 推出 Jensen 不等式成立.

由 Jensen 不等式可以看出以下几个结论:

(1) 如果 (X_n) 是鞅, 且 f 是凸函数, 那么 $f(X_n)$ 是下鞅. 例如 (X_n^2) 与 $(|X_n|)$ 是下鞅;

(2) 如果 (X_n) 是下鞅, 且 f 是凸且递增的函数, 那么 $f(X_n)$ 也是下鞅. 例如 (X_n^+) 是下鞅.

§5.2 Doob 鞅基本定理

在赌博游戏中, 大家会碰到这样一个情况: 有四个人 A,B,C,D 搓麻将赌钱, 另外一个人 E 旁观, 他也想参与赌博, 那么他可以押在某个人的结果上, 比如说他押 A, 就是和 A 一样输赢, 他也可以选择以 A 的输赢乘一个系数, 比如 k 倍. 其中的关键是他作出的决定只能依赖于已经有的信息, 不能与下一赌局的信息有关, 比如他不能看到下一局的牌, 他也不能说我压下一局的赢者. 只要认同这样的规则, 这样的参与其他人不会觉得不公平, 所有人都明白这一点, 但是几乎没有人会去考虑其中的原理, J.L.Doob 观察并注意到这个现象, 把它用数学的语言表达出来并且加以证明, 这说的是不管你用什么策略投资于赌徒, 游戏还是公平的. 这个事实简单而直观, 是随机分析最基本的定理.

怎么把这个问题用数学表达出来呢? 首先假设赌徒 A 的财富过程是 (X_n), n 是现在时刻, $X_{n+1} - X_n$ 是下一局输赢结果, 这时 E 作出决策, 压 A 的输赢 H_n 倍, 即它的所得为

$$H_n(X_{n+1} - X_n),$$

注意他的决策只能基于过去的信息 $\mathscr{G}_n$, 把他的所得累积起来记为 (Y_n), 递归定义

$$Y_0 = y,\ Y_{n+1} := Y_n + H_n(X_{n+1} - X_n),\ n \geq 0.$$

这是典型的投资赢利过程.

再看一个例子, 设某投资者按周投资某个证券, 第 n 周证券的价格是 X_n, 构成一个随机序列. 投资者在第 n 周买入 H_n 份证券, H_n 可以

是随机变量, 但是它只能依赖于证券前 n 周的价格提供的信息 $\mathscr{G}_n^X$, 到下一周他的证券价值为 H_nX_{n+1}, 他这周的所得为

$$H_n(X_{n+1}-X_n),$$

与上面的过程类似.

这样通过 (X_n) 与 (H_n) 定义新的过程 (Y_n) 称为是 (H_n) 关于 (X_n) 的随机积分. 序列 (H_n) 可以看成是投资策略, (X_n) 是原始证券, (Y_n) 是投资所得.

定理 5.2.1 (Doob) 设 $(\mathscr{G}_n)$ 是流, (H_n) 适应于 $(\mathscr{G}_n)$.

(1) 如果 (X_n) 是 $(\mathscr{G}_n)$ 鞅, 那么 (Y_n) 也是;

(2) 如果 (X_n) 是 $(\mathscr{G}_n)$ 下鞅, (H_n) 非负, 那么 (Y_n) 也是 $(\mathscr{G}_n)$ 下鞅.

只要投资者没有超能力, 不能预知未来, 那么不管什么策略, 游戏还是公平的. 证明很简单, 由条件期望性质

$$\mathbb{E}[Y_{n+1}-Y_n|\mathscr{G}_n]=H_n\mathbb{E}[X_{n+1}-X_n|\mathscr{G}_n],$$

这样定理的两个结论是显然的.

§5.3 应用于首次通过时

让我们看看 Doob 的鞅基本定理是多么的强大. 先介绍随机时间的概念, 一个非负整数值 (可以等于无穷) 随机变量 τ 称为是随机时间. 随机时间有两种, 一种是随时可以根据当前的信息判断这个时间是否已经到达, 例如股票价格达到 100 元; 另外一种是无法判断的情况, 例如股票价格得到一年中最高点. 第一种随机时间称为停时, 其数学定义如下.

定义 5.3.1 随机时间 τ 称为停时, 如果对任何 n 有 $\{\tau=n\}\in\mathscr{G}_n$, 这等价于对任何 n 有 $\{\tau\le n\}\in\mathscr{G}_n$.

虽然我这里说得轻松, 但我们应该理解, 能把一个生活中的感觉转化为数学定义绝对是非常天才的事情. 停时是最重要的概率论概念, 直观有用.

停时的原型是首次通过时, 对 $x \in \mathbb{R}$, 定义

$$\tau_x = \inf\{n : X_n \geq x\},$$

首次通过 x 的时间. 因为

$$\{\tau_x = n\} = \{X_0 < x, \cdots, X_{n-1} < x, X_n \geq x\} \in \mathscr{G}_n,$$

所以它是停时. 通常的时间 n 也是停时. 两个停时 τ_1, τ_2 取小 $\tau_1 \wedge \tau_2$ 或者取大 $\tau_1 \vee \tau_2$ 都是停时 (习题).

随机时间可以代入到通常的时间中, 定义

$$X_\tau(\omega) := X_{\tau(\omega)}(\omega),\ \omega \in \Omega,$$

注意它只能在事件 $\{\tau < \infty\}$ 上定义. 它也是一个随机变量. 这样我们就有停止随机列

$$X_n^\tau = X_{\tau \wedge n},$$

它是说在时间 τ 之后就停止不动了. 例如赌徒会说: 我今天赢一百块就走了, 这是随机时间停止, 或者我今天到 12 点就走了, 这是固定时间停止. 这是两种常见的停止, 但一般不会说: 我今天在赢最多的时候走, 因为这件事情无法实现.

直观地想, 什么时候停止也是一种策略, 所以停止过程应该是某种策略关于 (X_n) 的随机积分. 实际上, 当 $n < \tau$ 时, 游戏继续, 反之游戏停止, 因此

$$X_{\tau \wedge (n+1)} - X_{\tau \wedge n} = 1_{\{\tau > n\}}(X_{n+1} - X_n).$$

定理 5.3.1 如果 (X_n) 是 $(\mathscr{G}_n)$ 鞅, τ 是停时, 那么 $(X_{\tau \wedge n})$ 也是 $(\mathscr{G}_n)$ 鞅.

因为 $(X_{\tau\wedge n})$ 是鞅, 所以有恒等式

$$\mathbb{E}[X_{\tau\wedge n}] = \mathbb{E}[X_0],$$

然后如果恰好极限期望可以交换的话就有

$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0],$$

这可以给出 τ 的信息. 但不是随便拿一个鞅都能解决问题, 关键是要找到有用的鞅, 这不是一件容易的事情.

让我们讨论从 0 出发的简单随机游动 (X_n), 它的 Bernoulli 序列 (ξ_n) 服从的分布是

$$\mathbb{P}(\xi_n = 1) = p,\ \mathbb{P}(\xi_n = -1) = 1 - p = q.$$

显然, $(X_n - n(p-q) : n \geq 0)$ 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. 任意取 $x > 0$,

$$\mathbb{E}[x^{X_{n+1}}|\mathscr{G}_n] = x^{X_n}\mathbb{E}[x^{\xi_{n+1}}] = x^{X_n}(xp + x^{-1}q),$$

序列 (x^{X_n}) 不是鞅, 但可以看出对任何 $x > 0$,

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

组成的随机序列是一个鞅. 对正整数 a 令

$$\tau_a := \inf\{n : X_n = a\}.$$

因为 (Y_n) 是鞅, 所以 $(Y_{\tau_a\wedge n})$ 也是鞅. 因此

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

现在让 n 趋于无穷, 这时 $\tau\wedge n$ 趋于 τ, 当 $\tau < \infty$ 时, $X_\tau = a$ 而当 $\tau = \infty$ 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 $\tau\wedge n \leq \tau$, 所以 $X_{\tau\wedge n} \leq a$. 因此当 $x \geq 1$ 时,

$$x^{X_{\tau\wedge n}} \leq x^a.$$

当 $xp+x^{-1}q\geq 1$ 时

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

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) $p\geq 1/2$. 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: $\{\tau<\infty\}$ 和 $\{\tau=\infty\}$. 在第一部分上, $\lim_n X_{\tau\wedge n}=a$, 在第二部分上

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

因此让 n 趋于无穷且应用控制收敛定理, 推出

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

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

$$\mathbb{P}(\tau<\infty)=1.$$

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

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

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 $p\geq 1/2$ 时, $\mathbb{P}(\tau<\infty)=1$, 我们还可以写出 τ 的母函数

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

令 $xp+x^{-1}q=z$, 即

$$x=\frac{z+\sqrt{z^2-4pq}}{2p}>1,$$

所以

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

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 τ 的期望, 设 $a=1$, 两边对 z 求导, 然后让 $z\uparrow 1$, 得

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

当 $p>1/2$ 时有限, 当 $p=1/2$ 时无穷.

§5.4 应用于金融: 模型无关

概率论公理体系和条件期望的概念诞生于 1930 年代, Itô 的随机分析理论诞生于 1940 年左右, Doob 的鞅论诞生于 1950-1960 这个时期, 所以 1930-1960 这个时期是随机分析的萌芽期, 这个理论在 1970 年代左右被日本和法国的概率论学者逐步完善. 同时, 芝加哥期权期货交易所诞生于 1970 年代初, 催生了衍生证券定价的数学理论, 第一篇关于欧式期权定价的论文发表于 1972 年的 Journal of Political Economics, 作者是 Black & Scholes, 他们用 Itô 公式在假设证券满足几何 Brown 运动的模型下给出了欧式期权价格的解析表达式, 之后期权定价理论蓬勃发展, 诞生了金融数学这个交叉学科, 其中最重要的就是与模型无关 (model free) 的衍生证券定价第一与第二基本定理, 它们大概诞生于 1980-1990 年代. 有此经历, 我们说随机分析是一个幸运的学科, 刚一孵化, 便见证了在金融领域中辉煌的应用.

随机分析最有意义的应用是在金融领域的期权定价问题, 让我们从离散时间说起. 一个金融市场上有两种东西: 风险证券和可以无风险存贷款的银行. 设 $(X_n: n\geq 0)$ 是风险证券的价格, 是个随机序列. 存贷款的收益由利率决定, 简单地假设存贷款利率是一样的, 都是 r, 即存款 x, 下个时刻的价值是 $(1+r)x$, $x>0$ 表示存款, $x<0$ 表示贷款.

在时刻 n, 投资者的财富为 Y_n, 他需要做一个决定: 买 H_n 份证券, 同样 $H_n>0$ 表示买入, $H_n<0$ 表示卖空 (借证券卖出), 剩下的钱存入

银行, 即他的财富如下分配:

$$Y_n = H_n X_n + (Y_n - H_n X_n).$$

到下个时刻, 他的财富 Y_{n+1} 中的风险部分变成 $H_n X_{n=1}$, 存款部分变成 $(1+r)(Y_n - H_n X_n)$, 因此

$$Y_{n+1} = H_n X_{n+1} + (1+r)(Y_n - H_n X_n).$$

令

$$\tilde{Y}_n = (1+r)^{-n} Y_n,\ \tilde{X}_n = (1+r)^{-n} X_n.$$

这是和利息相反的运算, 称为折现. 这两个过程称为折现后的财富和证券价格过程. 变型得

$$(\tilde{Y}_{n+1} - \tilde{Y}_n) = H_n(\tilde{X}_{n+1} - \tilde{X}_n).$$

这个等式是意思是折现后的财富是策略关于折现后证券价格的随机积分. 至于为什么要折现? 是因为货币是有时间价值的.

设市场是给定的, 即 (X_n) 和利率 r 是给定的. 设 Y_0 是常数, 表示初始财富, 也就是本钱. 那么 (Y_n) 是由初始财富 Y_0, 以及策略 (H_n) 决定的. 现在介绍一个套利的概念. 套利的直观意思是无风险的利润, 或者说不管市场任何都会赚钱的机会. 怎么用数学来表达呢?

定义 5.4.1 市场有套利是指存在一个策略 (H_n) 和一个时刻 N 使得 $Y_0 = 0$ 而

$$Y_N \geq 0,\ \mathbb{P}(Y_N > 0) > 0.$$

市场无套利就是不存在这样的策略, 也就是说, 对任何的策略 (H_n), 任何的时刻 N, 如果 $Y_0 = 0$ 且 $Y_N \geq 0$, 则必有 $Y_N = 0$ a.s.. 无套利市场也称为有效市场.

定义中 $Y_0 = 0$ 是不需要本钱, $Y_N \geq 0$ 表示不会亏本, $\mathbb{P}(Y_N > 0) > 0$ 表示盈利的可能性是正的, 两者合起来才是套利. 同学可以验证, 在 $Y_N \geq 0$ 的前提下, $\mathbb{P}(Y_N > 0) > 0$ 等价于 $\mathbb{E}[Y_N] > 0$. 直观上我们总是认为一个成熟的市场是不可能存在套利机会的. 这在数学上怎么表达呢? 其实证券和利率应该也是有点关系的, 通过两者投资的获利不应该相差太多, 太多就意味着套利. 在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 可以有其它概率测度, 例如如果 $\xi > 0$, 且 $\mathbb{E}\xi = 1$, 定义

$$\tilde{\mathbb{P}}(A) := \mathbb{E}[\xi 1_A], \ A \in \mathscr{F},$$

那么 $\tilde{\mathbb{P}}$ 也是概率测度. 可以看出来, 这时两个概率测度有相同的零概率事件, 即

$$\mathbb{P}(A) = 0 \Leftrightarrow \tilde{\mathbb{P}}(A) = 0.$$

如果两个概率测度有相同的零概率事件集, 我们说它们等价. 两个等价的概率测度会改变概率, 但不改变基本的随机特性, 不会把不可能变成可能, 也不会把可能变成不可能. 鞅的定义中概率测度是关键, 一个概率测度下的鞅在另一个测度下一般不是鞅.

定理 5.4.1 (第一基本定理) 市场有效当且仅当存在一个等价概率测度 $\tilde{\mathbb{P}}$, 在这个测度下, 折现后的证券价格过程 $(\tilde{X}_n)$ 是鞅. 这个测度通常称为等价鞅测度.

这个定理的充分性比较容易证明, 就是说, 如果存在等价鞅测度, 那么市场无套利. 因为如果 $(\tilde{X}_n)$ 在测度 $\tilde{\mathbb{P}}$ 是鞅, 则 $(\tilde{Y}_n)$ 在这个测度下也是鞅, 所以

$$\tilde{\mathbb{E}}[Y_0] = (1 + r)^{-N} \tilde{\mathbb{E}}[Y_N].$$

如果有个策略 (H_n) 和时刻 N 使得 $Y_0 = 0$ 且 $Y_N \geq 0$, 那么 $\tilde{\mathbb{P}}[Y_N] = 0$, 所以 $\tilde{\mathbb{P}}(Y_N = 0) = 1$. 由概率等价性推出 $\mathbb{P}(Y_N = 0) = 1$. 这里关于概率测度 $\tilde{\mathbb{P}}$ 的期望也用符号 $\tilde{\mathbb{E}}$ 表示. 因此市场不存在套利.

习题

1. 证明: 对任何 n, $\{\tau = n\} \in \mathscr{G}_n$ 当且仅当对任何 n, $\{\tau \le n\} \in \mathscr{G}_n$.

2. 设 $\xi \ge 0$. 证明: $\mathbb{P}(\xi > 0) > 0$ 当且仅当 $\mathbb{E}\xi > 0$.

3. 设 τ_1, τ_2 是停时. 证明: $\tau_1 + \tau_2$, $\tau_1 \wedge \tau_2$ 也是停时.

第六章　Doob 基本定理及其应用

§6.1　应用于金融: 模型无关 (续)

必要性的证明不是那么容易, 这里不去赘述, 只是简单解释一下. 考虑两个随机变量 X_0, X_1, 无套利假设蕴含着 $\xi := \tilde{X}_1 - X_0$ 或者恒等于 0 或者分布在 0 的两侧: 或者

$$\mathbb{P}(\xi > 0) \cdot \mathbb{P}(\xi < 0) > 0.$$

这时存在一个等价概率测度 $\tilde{\mathbb{P}}$ 使得 $\tilde{\mathbb{E}}[\xi] = 0$. 相当于说 $\{tX_0, tX_1\}$ 在概率测度 $\tilde{\mathbb{P}}$ 下是个鞅.

怎么证明呢? 假设 $\{\xi > 0\}$ 与 $\{\xi < 0\}$ 的概率都是正的. 如果对所有的 $x \in \mathbb{R}$ 有 $\mathbb{E}[e^{x\xi}] < \infty$, 那么令

$$f(x) = \mathbb{E}[e^{x\xi}],$$

这是一个光滑函数且由条件推出 (请验证)

$$\lim_{|x| \to \infty} f(x) = +\infty.$$

因此这个函数在某个点 x_0 处达到最小值, 必然有 $f'(x_0) = 0$, 即

$$\mathbb{E}[\xi e^{x_0 \xi}] = 0,$$

令

$$\tilde{\mathbb{P}}(A) := \frac{\mathbb{E}[e^{x_0\xi}; A]}{\mathbb{E}[e^{x_0\xi}]},$$

那么它是一个等价于 $\mathbb{P}$ 的概率测度且 $\tilde{\mathbb{E}}[\xi] = 0$.

如果上面这个条件不满足怎么办? 没关系, 稍微麻烦一点, 我们可以证明: 对任何 $x \in \mathbb{R}$,

$$\mathbb{E}\left[e^{-\xi^2 + x\xi}\right] < \infty.$$

然后差不多同样的程序可以证明之. 我们把它单独写成一个命题, 后面有用.

Proposition 6.1.1 存在等价概率测度 $\tilde{\mathbb{P}}$ 使得随机变量 ξ 的期望 $\tilde{\mathbb{E}}[\xi] = a$ 的充分必要条件是 $\xi = a$ 或者分布在 a 的两边, 即

$$\mathbb{P}(\xi > a) > 0,\ \mathbb{P}(\xi < a) > 0.$$

有效市场定理看起来简单, 但实际上意义伟大, 称为是衍生证券定价第一基本定理, 它把市场的性质和一个数学概念结合起来了. 什么是衍生证券? 顾名思义, 衍生证券是从证券衍生出来的, 它的价值由证券决定, 也就是关于证券价值可测, 或者说是证券的函数. 基本的衍生证券的例子是期权和期货, 最简单的期权是欧式期权, 是在未来某个敲定的时刻购买证券的权利, 可以放弃, 它在当前签约的时候是有价值的; 期货是在未来某个敲定的时刻购买 (或卖) 某种商品的约定, 必须履约, 它在当前签约的时候是没有价值的, 只有在签约到到期这个时间段内因为商品现货价格变化才会有价值. 期权和期货是金融市场最大的创新, 现在各种名目繁多的衍生证券充斥市场, 令人眼花缭乱, 有的衍生证券的复杂程度即使对于专业人士也很难了解. 因此衍生证券的定价便是一个重要的问题, 而随机分析理论正好可以解释其中的一些现象.

什么是衍生证券的数学意义呢? 通常来说, 期权是一个合约: 以某个敲定时刻 N 和某种敲定方式买卖证券的权利, 它的价值 V 依赖于证券价格, 但在时刻 N 时是明确的, 因此它是 $X_1, \cdots, X_N$ 的函数, 或者关于 $\sigma(X_1, \cdots, X_N)$ 可测. 例如欧式买入期权: 以敲定价格 K 购买证券的权利, 那么

$$V = (S_N - K)^+;$$

证券公司会设计各种各样的期权, 如美式, 亚式, 以及其它各种花式期权.

通常的随机产品是用期望来定价的, 比如市场上销售的彩票, 保险

公司的保险项目, 赌场里面的各种赌博项目等. 它背后的机制是大数定律, 定价仍然是有风险的, 经营的公司要靠样本数量来降低风险. 这里的期望指的是关于概率测度 $\mathbb{P}$ 的期望, 如果用等价鞅测度 $\tilde{\mathbb{P}}$ 的期望, 那么对应的定价称为无套利定价, 它想表达的意思是: 如果不按照这个期望定价就会产生套利机会. 但是金融里面有个著名的一价原理, 就是在将来某个时刻的价值一样的物品在当前的价格必须是一样的, 否则一定会有套利, 也就是说, 等价鞅测度需要有唯一性, 无套利定价才有意义.

另外一个重要的概念是无风险定价, 这与大数定律的定价完全不同. 无风险定价是应用对冲的思想把衍生证券具有的风险消除. 衍生证券可以看成是投资者对证券未来的价值直接下堵注, 原则上讲, 投资者可以通过对证券应用适当的初始投资和适当的投资策略来达到同样的目的, 这是对冲的思想.

回归到数学, V 是敲定时刻为 N 的衍生证券, 问题是: 是否存在一个初始投资 Y_0 和策略 (H_n) 使得 $Y_N = V$? 如果有, 我们说 V 是可以对冲的或者可以复制的, 这时, 因为

$$\tilde{Y}_n - \tilde{Y}_{n-1} = H_{n-1}(\tilde{X}_n - \tilde{X}_{n-1}), \tag{6.1.1}$$

故

$$Y_0 = \tilde{\mathbb{E}}[\tilde{Y}_N] = (1+r)^{-N}\tilde{\mathbb{E}}[V],$$

也就是说初始投资是衍生证券关于等价鞅测度下的期望, 它应该就是无风险定价: 卖出一份衍生证券的风险可以通过收取的费用 Y_0 以及一种投资策略被消除.

定义 6.1.1 如果所有的衍生证券都是可对冲的, 那么我们说市场是完备的.

有效市场和完备市场可以通过一个简单的例子解释. 一个古玩投资者去深山未开发的地区玩, 在当地一个土著家吃饭, 看到他使用的碗

是一个汉朝的古董, 市场上值一百万. 他问土著这个碗卖多少钱. 这时有两种情况: 一, 土著完全不了解外面的世界, 说十块钱你就拿走吧. 这是说市场不有效, 土著没有充分了解信息. 二, 土著通过电视或者以前的顾客了解这个碗在市场上值很多钱, 但是因为没有有钱人接手, 这个碗一直卖不出去, 所以他泄气了, 说一千块钱你就拿走吧. 这是说市场不完备, 值钱的东西不一定能卖出去.

对 (6.1.1) 两边的 n 从 1 到 N 求和, 得

$$(1+r)^{-N}V = \sum_{n=1}^{N} H_{n-1}(\tilde{X}_n - tX_{n-1}) + Y_0.$$

因为 $(\tilde{X}_n)$ 在概率 $\tilde{\mathbb{P}}$ 之下是鞅, 所以我们说衍生证券可以鞅表示, 这是对冲的数学意义.

定理 6.1.2 (第二基本定理) 市场完备的充分必要条件是等价鞅测度唯一.

§6.2 应用于金融: 二叉树模型

以上理论是与具体模型无关的, 第二基本定理很难证明. 下面考虑一个具体模型: 二叉树模型. 我们在此特别情况下看第二基本定理.

设 $\{X_n : n \geq 0\}$ 是风险证券价格, 通常价格总是正的, $(\mathscr{G}_n)$ 是它的自然流. 通常

$$\frac{X_n - X_{n-1}}{X_{n-1}}$$

理解为增长率或者涨幅, 记为 η_n. 假设涨幅是独立同发布的且令

$$\xi_n := 1 + \eta_n = X_n/X_{n-1} > 0,$$

那么

$$\xi_1, \xi_2, \cdots, \xi_n, \cdots$$

是独立同分布正随机变量序列, 利率还是设为 r. 这样的市场称为简单市场. 我们来讨论两个问题: 什么条件下市场是有效的? 什么条件下市场是完备的?

有效市场是指存在等价鞅测度. 因为

$$\tilde{\mathbb{E}}[\tilde{X}_n|\mathscr{G}_{n-1}] = \tilde{X}_{n-1}(1+r)^{-1}\tilde{\mathbb{E}}[\xi_n],$$

故存在等价鞅测度的意思就是存在等价概率测度 $\tilde{\mathbb{P}}$ 使得

$$\tilde{\mathbb{E}}[\xi_n] = 1+r.$$

因为 ξ_n 不是常数, 所以由上面的命题, 它必须分布在 $1+r$ 的两边. 直观地说, 它不能完全高于也不能完全低于利率收益, 否则就有套利.

假设 ξ_n 取两个可能的值 $u > d > 0$,

$$\mathbb{P}(\xi_n = u) = p > 0,\ \mathbb{P}(\xi_n = d) = q = 1-p > 0.$$

等价测度只会改变 p 的值, 设 $\tilde{\mathbb{P}}(\xi_n = u) = \tilde{p}$. 那么, 当 $u\tilde{p}+d(1-tildep) = 1+r$ 时, $(\tilde{X}_n)$ 在概率测度 $\tilde{\mathbb{P}}$ 之下是鞅. 上面方程当且仅当 $u > 1+r > d$ 时有解

$$\tilde{p} = \frac{1+r-d}{u-d},$$

也就是说市场有效的充分必要条件是

$$u > 1+r > d.$$

而且等价鞅测度是唯一的, 也就是说市场是完备的. 前面我们没有证明第二基本定理, 所以就这个例子可以看看鞅表示性质是不是成立?

不妨设 $r = 0$, 设敲定时间 $N = 1$, 衍生证券 $V = V(X_1) = V(X_0\xi_1)$. 是否存在 Y_0, H_0, 使得

$$V(X_0\xi_1) = H_0(X_1 - X_0) + Y_0 = H_0X_0(\xi_1 - 1) + Y_0? \tag{6.2.1}$$

因为 ξ_1 取两个值 u, d, 所以有两个方程:

$$V(X_0u) = H_0X_0(u-1) + Y_0, V(X_0d) \qquad = H_0X_0(d-1) + Y_0.$$

两个方程, 两个未知量, 正好有唯一解

$$H_0 = \frac{V(X_0u) - V(X_0d)}{X_0(u-d)}, Y_0 \quad = \frac{(u-1)V(X_0d) + (1-d)V(X_0u)}{u-d} = \tilde{\mathbb{E}}[V].$$

这是一个满足第二基本定理条件的例子.

其实有唯一解的情况是很特别的, 主要的原因是 ξ_n 取两个值, 称为二叉树模型. 如果 ξ_n 是取两个值 $x_1 < x_2 < x_3$, 那么存在等价鞅测度的充分必要条件是

$$x_1 < 1 + r < x_3.$$

因为方程

$$\tilde{\mathbb{E}}[\xi_n] = x_1\tilde{p}_1 + x_2\tilde{p}_2 + x_3(1 - \tilde{p}_1 - \tilde{p}_2) = 1 + r$$

有两个变量, 故有解但不唯一, 实际上有无穷多解, 因此这时市场有效但不完备. 实际上这时候, 因为 ξ_1 有三个状态, 方程 (6.2.1) 有三个方程, 两个变量, 因此解 Y_0, H_0 不存在. 这是一个不满足第二基本定理条件的例子.

§6.3 美式买入期权

敲定时间 N 敲定价格 K 的欧式买入期权必须在到期日执行, 不能提前执行, 所以它到期的价值是 $(S_N - K)^+$, 它在 0 时刻的价值是

$$\tilde{\mathbb{E}}[(1+r)^{-N}(S_N - K)^+],$$

其中 $\tilde{\mathbb{P}}$ 是等价鞅测度. 美式买入期权与欧式买入期权的区别是它可以提前执行. 如果在一个停时 $\tau \le N$ 时刻执行, 那么它在 N 时刻的价值是 $(1+r)^{N-\tau}(S_\tau - K)^+$, 它在 0 时刻的价格是

$$\tilde{\mathbb{E}}[(1+r)^{-\tau}(S_\tau - K)^+].$$

下面我们来证明

$$\tilde{\mathbb{E}}[(1+r)^{-\tau}(X_\tau - K)^+] \le \tilde{\mathbb{E}}[(1+r)^{-N}(X_N - K)^+], \tag{6.3.1}$$

也就是说, 美式买入期权提前执行不如在最后到期日执行.

为了证明这个, 我们需要一个引理. 设 σ, τ 是两个停时, 且 $\sigma \le \tau$, 那么

$$1_{\{\sigma < n \le \tau\}} = 1_{\{\tau \ge n\}} - 1_{\{\sigma \ge n\}}.$$

令

$$Y_n := X_{\tau \wedge n} - X_{\sigma \wedge n},$$

则

$$Y_n - Y_{n-1} = 1_{\{\tau \ge n > \sigma\}}(X_n - X_{n-1}).$$

由 Doob 的基本定理, 如果 (X_n) 是下鞅, 那么 (Y_n) 也是下鞅, 推出下面的定理.

定理 6.3.1 如果 (X_n) 是下鞅, 且 σ, τ 都是停时, 那么对任何 n,

$$\mathbb{E}[X_{\tau \wedge n}] \ge \mathbb{E}[X_{\sigma \wedge n}].$$

现在在等价测度 $\tilde{\mathbb{P}}$ 下, $(\tilde{X}_n)$ 是鞅, 所以

$$(1+r)^{-n}(X_n - K)^+ = (\tilde{X}_n - (1+r)^{-n}K)^+,$$

而 $(1+r)^{-n}K$ 递减, 因此 $(\tilde{X}_n - (1+r)^{-n}K)$ 是下鞅. 由因为 $x \mapsto x^+$ 是凸且递增的函数, 所以 $\{(1+r)^{-n}(X_n - K)^+ : n \ge 0\}$ 是下鞅. 但是 $\tau \le N$, 由以上定理推出 (6.3.1).

§6.4 连续时间随机过程

前面讨论的是随机序列, 是离散时间的随机过程, 下面介绍连续时间的随机过程. 给定概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$, 如果对任何 $t \ge 0$, X_t 是一个

随机变量, 那么我们说 $(X_t : t \geq 0)$ 是一个随机过程. 这时候, 对任何 $\omega \in \Omega$, $X_t(\omega)$ 是 $[0, \infty)$ 上的函数, 它被称为样本 ω 对应的样本轨道, 简单称为样本轨道. 如果一个性质对几乎所有 Ω 成立, 我们就说几乎所有样本轨道有这个性质. 例如几乎所有的样本轨道连续, 这时我们说随机过程是连续的. 有些概念和离散时间类似, 随机时间下同样有信息流的概念. 如果对任何 $t \geq 0$, $\mathscr{G}_t$ 是一个事件域且对任何 $0 \leq s < t$ 有 $\mathscr{G}_s \subset \mathscr{G}_t$, 那么我们说 $(\mathscr{G}_t)$ 是一个信息流, 或者简单称为流. 同样如果对任何 $t \geq 0$, X_t 关于 $\mathscr{G}_t$ 可测, 那么我们说随机过程 (X_t) 适应于 $(\mathscr{G}_t)$ 一个随机过程也有自然流, 自然流是随机过程所适应的最小流. 鞅的定义也类似.

定义 6.4.1 设 (X_t) 是一个适应于流 $(\mathscr{G}_t)$ 的随机过程. 如果

(1) 对任何 t, 有 $\mathbb{E}[|X_t|] < \infty$;

(2) 对任何 $s < t$ 有
$$\mathbb{E}[X_t|\mathscr{G}_s] = X_s,$$

我们说 (X_t) 是鞅. 上鞅下鞅的定义类似. 一个连续随机过程是鞅的话称为连续鞅.

一个随机时间 $\tau : \Omega \to [0, \infty)$ 称为是一个停时, 如果对任何 $t \geq 0$ 有
$$\{\tau \leq t\} \in \mathscr{G}_t.$$
对任何集合 $A \subset \mathbb{R}$, 定义首中时
$$\tau_A(\omega) = \inf\{t > 0 : X_t \in A\},$$
那么当 A 是开集或者闭集且 (X_t) 是连续过程时, τ_A 是停时, 直观上说, 只要看时间 t 之前的轨道就可以知道是不是已经到达 A, 这正是停时的直观意义. 下面的定理和离散时间的类似, 称为 Doob 停止定理, 直观但证明不容易, 省略.

定理 6.4.1 如果 (X_t) 是连续鞅, τ 是停时, 那么停止过程 $(X_{\tau\wedge t})$ 也是鞅.

习题

1. 证明命题 6.1.1.
2. 假设市场完备, 问市场上到期时间为 N 敲定价格为 K 的证券期货合同在时刻 n $(0 < n < N)$ 的价值几何?

第七章　Brown 运动及其性质

§7.1　Brown 运动 (1)

随机过程的重要例子是 Brown 运动, Brown 运动是英国生物学家 R.Brown 在 1850 年左右首先描述的运动, 它是指花粉在液体表面的无规则运动, 当然他的论文是描述性的并没有数学表达式, 因此没有引起科学家的注意. 但实际上, 这样的运动有普遍性. 1900 年, 法国数学家 Poincaré 的学生 Bachelier 的博士论文研究股票的运动, 得到股票价格的转移密度函数 $p_t(x,y)$ 满足热传导方程. 不知道什么原因, 他的这个结果直到 50 多年后才被人注意. 引起重视的是 A.Einstein 的工作, 他在 1905 年研究粒子运动时也得到同样的结果: 粒子运动作为一个随机过程, 其转移密度函数满足热传导方程. 这说明我们可以构造一个随机过程 (X_t), 它满足

$$\mathbb{P}(X_t \in A | X_s = x) = \int_A p_t(y-x)dy,$$

其中 $x \mapsto p_t(x)$ 称为转移密度函数, 满足

$$\frac{\partial p_t}{partialt} = c \cdot \frac{\partial^2 p_t}{\partial x^2}, \tag{7.1.1}$$

这里 c 是传导系数, 通常取为 1/2. 实际上满足以上方程的解是方差为 t 的正态分布密度函数

$$p_t(x) = \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t}).$$

Einstein 在做这个工作时也不知道 Brown 的论文, 后来数学家发现这样的运动在随机性上有共同性, 与 Brown 观察到的粒子运动相似, 所以统一称为 Brown 运动. 但是为了和实际模型符合, 数学上还有一道难题需要证明, 也就是说要证明这样的随机过程的几乎所有轨道是连续的, 这

个工作最为困难, 因为那时概率论公理体系尚未出现, 人们还在黑暗中摸索, 比如摸索怎么构造一个随机过程, 怎么才是证明了轨道的连续性, 等等问题. 一直到 1923 年天才 N.Wiener 出现, 他证明了轨道连续性, 要知道, 这早于公理体系出现 10 年.

如果你观察过花粉在液体表面的运动, 你大概心里会有感觉 Brown 运动是什么意思. 它是一种受到各个方向分子撞击所引起的运动, 所以毫无规律可言, 方向是完全随机的. 但是这无法给你什么数学感觉, 要得到数学上的感觉, 我们要回到简单对称随机游动 (X_n), 首先我们看到, 由中心极限定理 $X_n\sqrt{n}$ 依分布收敛于标准正态分布. 把 (X_n) 连续化, 对 $\omega\in\Omega$, 将 $(n,X_n(\omega))$ 与 $(n+1,X_{n+1}(\omega))$ 用直线连接起来, 这样得到一个 $[0,\infty)$ 上的函数: $X_t(\omega)$, 这样得到一个连续时间的随机过程 (X_t), 是个折线过程. 然后我离开, 离得越来越远, 在远处看这个折线过程, 它就像是 Brown 运动了, 具体地说, 需要对空间和时间进行细化, 令

$$Y_t^{(k)} := X_{kt}/\sqrt{k},\ k\geq 1.$$

空间和时间细化的比例是不同的, 时间间隔是 $1/k$, 空间间隔是 $1/\sqrt{k}$. 那么当 k 趋于无穷时, $(Y_t^{(k)})$ 的极限是 Brown 运动. 这个结果称为 Donsker 不变原理, 是中心极限定理的升级版, 证明很不容易, 涉及到无穷维空间的 Skorohod 拓扑.

尽管不能亲自证明这个定理, 但我们还是可以从中看到 Brown 运动的一点端倪, 下面的定义显得比较自然了.

定义 7.1.1 随机过程 $B=(B_t)$ 是 Brown 运动, 如果

(1) 对任何 $t>s\geq 0$, B_t-B_s 与 $\mathscr{G}_s$ 独立;

(2) 对任何 $t>s$, B_t-B_s 服从方差为 $t-s$ 期望为 0 的正态分布;

(3) 几乎所有轨道连续,

其中 $(\mathscr{G}_t)$ 是自然流. 如果再有 $\mathbb{P}(B_0=0)=1$, 那么称为标准 Brown 运动.

实际上, $(\mathscr{G}_t)$ 不一定非要求是自然流, 只要求是 (B_t) 的适应流就可以了, 这时, (B_t) 称为是 $(\mathscr{G}_t)$-Brown 运动. Brown 运动存在性的证明难在连续性的证明, 超出课程范围. 简单对称随机游动是最重要的随机游动, Brown 运动自然是最重要的随机过程. 下面给出标准 Brown 运动 (B_t) 的一些性质:

(1) $(-B_t)$ 也是标准 Brown 运动. 称为对称性;

(2) 对任何实数 c, $(c^{-1}B_{c^2t})$ 也是标准 Brown 运动. 称为分形性;

(3) $(tB_{t^{-1}})$ 也是标准 Brown 运动. 称为可逆性;

§7.2 反射原理

数学中有很多的反射原理, 这里的反射原理实际上是 Brown 运动的对称性之升级版. 设 (B_t) 是标准 Brown 运动, 把轨道作为函数 $x=B_t$ 画在时间 - 位置平面上, 对任何 $s>0$, 把 s 时间之后的轨道按高度为 B_s 且平行于时间轴的直线 $x=B_s$ 进行反射, 得到新的轨道, 那么新的轨道组成的随机过程也是标准 Brown 运动. 还是要回到数学, 定义

$$X_t := \begin{cases} B_t, & t\le s, \\ 2B_s-B_t, & t>s, \end{cases}$$

那么 (X_t) 也是标准 Brown 运动. 比如, 当 $u>s>t$ 时, 请验证: X_u-X_t 与 B_u-B_t 同分布. 然后再验证 (X_t) 的确是标准 Brown 运动.

这是在固定的时间反射, 我们也可以定义在固定位置反射. 设 $a>0$, 定义

$$\tau_a(\omega) := \inf\{t>0: B_t(\omega)=a\},$$

这是 a 的首中时. 定义

$$X_t(\omega) := \begin{cases} B_t(\omega), & t \le \tau_a(\omega), \\ 2a - B_t(\omega), & t > \tau_a(\omega), \end{cases}$$

那么 (X_t) 也是标准 Brown 运动. 这个的证明要比固定时间反射困难许多.

现在我们的目的是计算 τ_a 的分布. 由反射原理,

$$\begin{aligned} \mathbb{P}(\tau_a < t, B_t < a) &= \mathbb{P}(\tau_a < t, X_t < a) \\ &= \mathbb{P}(\tau_a < t, B_t > a), \end{aligned}$$

因此

$$\begin{aligned} \mathbb{P}(\tau_a < t) &= 2\mathbb{P}(\tau_a < t, B_t > a) \\ &= 2\mathbb{P}(B_t > a) \\ &= 2\int_{a/\sqrt{t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx, \end{aligned}$$

由此推出

$$\mathbb{P}(\tau_a < \infty) = 1,$$

也就是说, 标准 Brown 运动的几乎所有轨道在有限时间内都会碰到 a. 那么

$$\mathbb{P}(\sup_t B_t = \infty, \inf_t B_t = -\infty) = 1,$$

即几乎所有轨道振荡着趋于无穷.

实际上, 当 $a > 1/2$ 时, B_t/t^a 几乎处处趋于 0. 而 $a \le 1/2$ 时, B_t/t^a 几乎处处不收敛.

另外, 我们还可以证明: $\mathbb{P}(\tau_0 = 0) = 1$. 也就是说对几乎所有轨道, $t = 0$ 不是孤立零点. 实际上, 设 $Z = \{t : B_t = 0\}$, Brown 运动的零点集,

那么对于几乎所有轨道, Z 是一个测度等于 0 的没有孤立点的闭集, 也就是 Cantor 型集合. 事实上, 如果用 $|Z|$ 表示 Z 的 Lebesgue 测度, 那么

$$\mathbb{E}|Z| = \mathbb{E}\int_{\mathbb{R}} 1_{\{B_t=0\}}dt = \int_{\mathbb{R}} \mathbb{P}(B_t=0)dt = 0,$$

因此, $|Z| = 0$ a.s.

§7.3 Brown 运动与鞅

前面我们给出了鞅的定义, 但没有实际的例子, 现在 Brown 运动会给我们很多很多例子. 下面设 (B_t) 是关于流 $(\mathscr{G}_t)$ 的标准 Brown 运动. 因为 $B_t - B_s$ 与 $\mathscr{G}_s$ 独立, 所以

$$\mathbb{E}[B_t - B_s|\mathscr{G}_s] = \mathbb{E}[B_t - B_s] = 0,$$

即是说 (B_t) 本身是个鞅. 自然 $(B_t - B_s)^2$ 也与 $\mathscr{G}_s$ 独立, 所以

$$\mathbb{E}[(B_t - B_s)^2|\mathscr{G}_s] = \mathbb{E}[(B_t - B_s)^2] = t - s,$$

我们知道 (B_t^2) 是下鞅, 但是根据条件期望性质

$$\begin{aligned}\mathbb{E}[(B_t - B_s)^2|\mathscr{G}_s] &= \mathbb{E}[B_t^2 - 2(B_t - B_s)B_s - B_s^2|\mathscr{G}_s] \\ &= \mathbb{E}[B_t^2 - B_s^2|\mathscr{G}_s],\end{aligned}$$

因此随机过程 $(B_t^2 - t)$ 是鞅.

但最最重要的还是指数鞅. 对任何 $x \in \mathbb{R}$, 增量的函数 $e^{x(B_t - B_s)}$ 与 $\mathscr{G}_s$ 独立且它的期望为 $e^{x^2(t-s)/2}$, 所以

$$\mathbb{E}[e^{x(B_t - B_s)}|\mathscr{G}_s] = e^{x^2(t-s)/2},$$

或者

$$\mathbb{E}[e^{xB_t - x^2t/2}|\mathscr{G}_s] = e^{xB_s - x^2s/2},$$

因此随机过程

$$(e^{xB_t - x^2t/2} : t \geq 0)$$

是鞅. 说指数鞅是最重要的鞅的原因是它如同一个母鞅, 因为将它展开

$$\begin{aligned} e^{xB_t - x^2t/2} &= 1 + (xB_t - x^2t/2) + \frac{1}{2}(xB_t - x^2t/2)^2 + \cdots \\ &= 1 + xB_t + \frac{1}{2}x^2(B_t^2 - t) + \frac{1}{3!}x^3(B_t^3 - 3tB_t) + \cdots , \end{aligned}$$

从而看出 (B_t), $(B_t^2 - t)$, 还有 $(B_t^3 - 3tB_t)$ 都是鞅.

§7.4 Doob 停止定理的应用

首先看障碍问题, 设有 $a, b > 0$, (B_t) 是标准 Brown 运动, τ_{-a}, τ_b 分别是 $-a, b$ 的首中时, $\tau := \tau_{-a} \wedge \tau_b$, 它是首次到达集合 $\{-a, b\}$ 的时间. 显然

$$\mathbb{P}(\tau < \infty) = 1.$$

需要算先到达 b 的概率 $\mathbb{P}(\tau_b < \tau_{-a})$. 因为 (B_t) 是鞅, 所以 $\mathbb{E}[B_{t\wedge\tau}] = 0$. 由于 $t \wedge \tau \leq \tau$, 故

$$-a \leq B_{t\wedge\tau} \leq b,$$

因此由控制收敛定理推出 $\mathbb{E}[B_\tau] = 0$. 而

$$\begin{aligned} \mathbb{E}[B_\tau] &= \mathbb{E}[B_\tau; \tau_b < \tau_{-a}] + \mathbb{E}[B_\tau; \tau_b > \tau_{-a}] \\ &= b\mathbb{P}(\tau_b < \tau_{-a}) - a\mathbb{P}(\tau_b > \tau_{-a}), \end{aligned}$$

推出

$$\mathbb{P}(\tau_b < \tau_{-a}) = \frac{a}{a+b}.$$

再算 $\mathbb{E}[\tau]$. 这时利用 $(B_t^2 - t)$ 是鞅的结论, 有

$$\mathbb{E}[t \wedge \tau] = \mathbb{E}[B_{t\wedge\tau}^2].$$

再应用单调收敛定理和控制收敛定理得

$$\begin{aligned}\mathbb{E}[\tau] &= \mathbb{E}[B_\tau^2] \\ &= b^2\mathbb{P}(\tau_b < \tau_{-a}) + a^2\mathbb{P}(\tau_b > \tau_{-a}) \\ &= b^2\frac{a}{a+b} + a^2\frac{b}{a+b} = ab\end{aligned}$$

如果要算 τ 的分布, 这两个鞅都不行, 需要用指数鞅, 留作习题.

习题

1. 证明: 按固定时间反射的 Brown 运动还是 Brown 运动.
2. 证明:
$$\mathbb{P}(\inf_{t>0} B_t = -\infty, \sup_{t>0} B_t = +\infty) = 1.$$
3. 计算: $\mathbb{E}[B_tB_s]$
4. 求 $(B_{t_1}, \cdots, B_{t_n})$ 的协方差矩阵和联合密度函数, 其中 $0 < t_1 < t_2 < \cdots < t_n$.
5. 在 $s < t$ 和 $s > t$ 两种情况下求 $\mathbb{E}[B_t|B_s]$.
6. 问 $\int_0^t B_s ds$ 服从什么分布, 为什么?
7. 求 τ_a 的密度函数.
8. 当 $k > 1/2$ 时, 证明: $n^{-k}B_n$ 几乎处处趋于 0.
9. 证明: 对几乎所有轨道, $a \mapsto \tau_a$ 是左连续的.
10. 求 $\tau = \tau_{-a} \wedge \tau_b$ 的 Laplace 变换.

第八章　Brown 运动 (2)

§8.1　与直线的首次相交

下面我们来看另外一个有趣的例子, 算 $t \mapsto B_t$ 会碰到斜率为 k 的直线 $t \mapsto kt + a$ 的概率, 其中 $a > 0$, $k \in \mathbb{R}$.

设 $B = (B_t)$ 是 1 维标准 Brown 运动. 对 $a > 0$, 定义 τ_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 Brown 运动一定会到达 a, 即

$$\mathbb{P}(\tau_a < \infty) = 1.$$

我们也可以用鞅方法来解答这个问题. 更一般地, 我们问 Brown 运动会肯定碰到一条斜的直线 $x = kt + a$ 吗? 令 T 是 Brown 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 Brown 运动 $(B_t - kt)$ 首次碰到 a 的时间, 求 $\mathbb{P}(T < \infty)$. 不妨设 $a > 0$, 当 $k = 0$ 时, $T = \tau_a$. 直观看, 当 $k < 0$ 时, $\mathbb{P}(T < \infty) = 1$, 而当 $k > 0$ 时未必, 因为 (B_t) 的包络和 $\sqrt{t}$ 差不多.

由指数鞅性质, 对任何实数 z,

$$\exp\left(zB_t - \frac{z^2}{2}t\right),\ t \geq 0$$

是鞅. 那么由 Doob 定理,

$$\mathbb{E}\left[\exp\left(zB_{t\wedge T} - \frac{z^2}{2}(t \wedge T)\right)\right] = 1. \tag{8.1.1}$$

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, 因为 $B_{t\wedge T} \leq k(t \wedge T) + a$, 故

$$zB_{t\wedge T} - \frac{z^2}{2}(t \wedge T) \leq z(k(t \wedge T) + a) - \frac{z^2}{2}(t \wedge T)$$

$$= (zk - \frac{z^2}{2})(t \wedge T) + za.$$

因此要保证 (8.1.1) 中的指数关于 t, ω 有界, 必须 $zk - z^2/2 \leq 0$.

这是一个简单的二次函数, 要求 $z > 0$, 分两种情况:

(1) 当 $k \leq 0$ 时, $zk - z^2/2 < 0$;

(2) $k > 0$ 时, 只有当 $z > 2k$ 时才有 $zk - z^2/2 < 0$.

无论哪种情况, 只要 $zk - z^2/2 < 0$ 就可以应用有界收敛定理, 当 t 趋于无穷时,

$$\exp\left(zB_{t\wedge T} - \frac{z^2}{2}(t \wedge T)\right) \to \exp\left(zB_T - \frac{z^2}{2}T\right) 1_{\{T<\infty\}},$$

因此

$$\mathbb{E}\left[\exp\left(zB_T - \frac{z^2}{2}T\right); T < \infty\right] = 1.$$

当 $T < \infty$ 时, 有 $B_T = kT + a$, 所以

$$\mathbb{E}\left[\mathrm{e}^{(zk-\frac{z^2}{2})T}; T < \infty\right] = \mathrm{e}^{-za}.$$

在第一种情况下, 我们可以让 $z \uparrow 0$, 得 $\mathbb{P}(T < \infty) = 1$; 第二种情况下, 让 $z \uparrow 2k$, 得

$$\mathbb{P}(T < \infty) = \mathrm{e}^{-2ka} < 1.$$

在 $k \leq 0$ 情况下, 我们可以算出 T 的 Laplace 变换, 因为

$$\mathbb{E}\left[\mathrm{e}^{(zk-\frac{z^2}{2})T}\right] = \mathrm{e}^{-za},$$

令 $-s = zk - z^2/2$, $s > 0$, 得 $z = k + \sqrt{k^2 + 2s} > 0$, 因此

$$\mathbb{E}[\mathrm{e}^{-sT}] = \mathrm{e}^{-a(k+\sqrt{k^2+2s})}. \tag{8.1.2}$$

如果用 T_a 表示上面的 T 并且把 a 看成时间, 那么 $(T_a : a \geq 0)$ 也是随机过程, 可以证明它是一个平稳独立增量过程.

§8.2 粗糙轨道

什么是粗糙轨道? 其实只要看看股票市场的股票价格走势就知道了. 严格地说, 如果一个连续函数 $f:[0,\infty)\to\mathbb{R}$ 的图像曲线在任何一个区间上的长度都是无穷, 那么我们就说这个连续函数的轨道是粗糙的. 如果 f 在任何一段区间上是光滑 (连续可导) 的, 那么它在这一段上的长度必然是有限的, 所以轨道粗糙意味着它没有任何一段是光滑的. 我们将证明 Brown 运动的几乎所有的轨道都是粗糙的.

再严格地看粗糙的定义, 什么是曲线长度? 实际上长度并没有严格的定义, 假设 f 是 $[a,b]$ 上连续函数, 用 $l(f)$ 表示图像 $\{(x,f(x)):x\in[a,b]\}$ 的长度. 用 $D=\{a=t_0<\cdots<t_n=b\}$ 表示 $[a,b]$ 的一个划分, 由三角不等式

$$\begin{aligned}l(f)&\geq\sum_D[(t_i-t_{i-1})^2+(f(t_i)-f(t_{i-1}))^2]^{1/2}\\&\geq\sum_D|f(t_i)-f(t_{i-1})|,\end{aligned}$$

其中右边是 f 在 D 上的变差, 记为 $\mathrm{var}(f;D)$. 因此

$$l(f)\geq\mathrm{var}(f;D),$$

然后对所有的 D 取上确界, $\sup_D\mathrm{var}(f;D)$ 就是 f 的全变差, 记为 $\mathrm{var}(f)$, 有

$$\mathrm{var}(f)\leq l(f).$$

因为长度无法定义, 而全变差是有严格定义的, 所以我们把函数的粗糙理解为它在任何区间上的全变差是无穷.

现在, 因为 Brown 运动的几乎所有轨道是连续的, 所以我们不妨认为 $t\mapsto B_t$ 是连续的. 显然 (B_t) 在区间 $[a,b]$ 的划分 D 上变差为

$$\mathrm{var}(B;D)=\sum_D|B_{t_i}-B_{t_{i-1}}|,$$

它是一个随机变量, 我们要证明它几乎处处趋于无穷. 但直接算很不容易, 如果算期望倒是不难, 因为

$$\mathbb{E}|B_{t_i} - B_{t_{i-1}}| = \sqrt{t_i - t_{i-1}}.$$

推出

$$\mathbb{E}[\mathrm{var}(B)] = \infty.$$

但这也不能推出 $\mathrm{var}(B) = \infty$ a.s.

哪怎么办呢? 这里介绍一个非常聪明的想法, 引入函数的二次变差:

$$\mathrm{var}_2(f; D) := \sum_D (f(t_i) - f(t_{i-1}))^2.$$

显然

$$\mathrm{var}_2(f; D) \le \max_D |f(t_i) - f(t_{i-1})| \cdot \mathrm{var}(f; D) \le \max_D |f(t_i) - f(t_{i-1})| \mathrm{var}(f).$$

如果 $\mathrm{var}(f) < \infty$, 那么任取一个趋于 0 的划分列 $\{D_n\}$, 因为函数连续, 故有

$$\lim_n \max_{D_n} |f(t_i) - f(t_{i-1})| = 0,$$

因此推出

$$\mathrm{var}_2(f; D_n) \to 0.$$

我们用它的逆否命题: 如果存在一个趋于零的划分列 $\{D_n\}$ 有 $\lim_n \mathrm{var}_2(f; D_n) > 0$, 则 $\mathrm{var}(f) = \infty$.

我们把 (B_t) 在区间 $[a, b]$ 上关于划分 D 的二次变差写成为

$$\mathrm{var}_2(B; [a, b], D) := \sum_D (B(t_i) - B(t_{i-1}))^2,$$

其中 $B(t) = B_t$. 这是一个依赖于 D 的随机变量, 它的期望方差很容易算

$$\mathbb{E}[\mathrm{var}_2(B; [a, b], D)] = b - a,$$

$$\mathbb{E}[\mathrm{var}_2(B;[a,b],D)-(b-a)]^2 = 2\sum_D (t_i-t_{i-1})^2.$$

事实上, 这个等式由

$$\begin{aligned}&\mathbb{E}[(B(t_i)-B(t_{i-1}))^2-(t_i-t_{i-1})]^2\\&\quad=3(t_i-t_{i-1})^2-2(t_i-t_{i-1})^2+(t_i-t_{i-1})^2\\&\quad=(t_i-t_{i-1})^2,\end{aligned}$$

以及独立增量性推出. 因此, 当 D_n 是趋于零的划分列时, 有

$$\lim_n \mathbb{E}[\mathrm{var}_2(B;[a,b],D_n)-(b-a)]^2=0,$$

即二次变差 $\mathrm{var}_2(B;[a,b],D_n)$ 在 L^2 意义下的极限是 $b-a$. 尽管 L^2 收敛并不一定几乎处处收敛, 但是一定有一个子列是几乎处处收敛的, 也就是说, 对于几乎所有轨道, 二次变差沿一个趋于零的划分列的极限存在而且是正的. 精确地说, 存在零概率集 $N_{a,b}$ 使得当 $\omega\notin N_{a,b}$, 函数 $t\mapsto B_t(\omega)$ 在 $[a,b]$ 上的全变差无限. 取

$$N:=\bigcup_{a,b\in\mathbf{Q}} N_{a,b},$$

那么 $\mathbb{P}(N)=0$ 且对 $\omega\notin N$, 函数 $t\mapsto B_t(\omega)$ 在任何以有理数为端点的区间上都有无限的全变差, 所以也就是在任何区间上有无限的全变差.

定理 8.2.1 Brown 运动的几乎所有轨道在任何区间上都不是有界变差的.

这个事实让人惊讶, Brown 运动的轨道会是多么怪异的一个函数啊! 它其中的任何一段的长度都是无穷. 其实, 还可以证明: 几乎所有轨道在任何点都不可导.

§8.3 有界变差和 Stieltjes 积分

前面看到, 投资策略 (H_n) 关于证券价格 (X_n) 的随机积分是

$$Y_n = Y_0 + H_0(X_1 - X_0) + H_1(X_2 - X_1) + \cdot + H_{n-1}(X_n - X_{n-1}),$$

这个积分有明确的金融意义, 值得研究, 其连续时间的随机积分形式对应应该是

$$\int_0^t H_s dX_s,$$

所以我们需要研究这个随机积分怎么定义, 有什么性质, 怎么计算等问题.

一个自然想到的方法是按照轨道定义, 就是固定 $\omega \in \Omega$ 来看函数 $t \mapsto H_t(\omega)$ 关于函数 $t \mapsto X_s(\omega)$ 的积分, 即所谓的 Stieltjes 积分. 设有 $[0,\infty)$ 上的连续函数 $x = x(t)$ 和函数 $f(t)$, f 关于 x 的 Riemann-Stieltjes 积分 $\int_0^t f(s)dx(s)$ 是以下 R-S 和当划分趋于零时候的极限 (如果存在)

$$\sum_D f(\xi_i)(x(t_i) - x(t_{i-1})),$$

其中 D 是 $[0,t]$ 上的划分, $\xi_i \in [t_{i-1}, t_i]$. 这个定义和 Riemann 积分一致, 而且同样可以证明: 当 $x = x(t)$ 是有界变差函数, 且 f 连续的时候, 这个积分一定存在. 通常的 Riemann 积分是 $x(t) = t$ 时候的特殊情况. 实际上, 令人绝望的事实是这个积分对所有连续函数存在的充分必要条件就是 $x = x(t)$ 是有界变差函数, 因为 Brown 运动的轨道在任何区间上都不是有界变差的, 所以这让我们想按照轨道来定义关于 Brown 运动积分的希望完全破灭了.

§8.4 Brown 运动关于自身的积分

门虽然关上了, 但还会有窗户开着. 上面是说关于 Brown 运动的 Riemann-Stieltjes 和的按轨道的极限是不可能存在的, 但还有别的极限,

比如 L^2 极限或者更一般的依概率收敛的极限是不是会存在呢? Itô 在 1940 年代研究了这个问题并且找到了定义随机积分的思想, 这个思想的关键之处是 Brown 运动的二次变差存在.

让我们看一个简单的例子, 就是 Brown 运动关于自生的 R-S 和

$$I([a,b],D) := \sum_D B(t_{i-1})(B(t_i) - B(t_{i-1})),$$

其中 D 是 $[a,b]$ 的一个划分

$$a = t_0 < t_1 < \cdots < t_n = b.$$

因为其中的 $\xi_i = t_{i-1}$, 所以这个和称为是取左端点的 R-S 和, 对应的有取右端点的 R-S 和. 在 R-S 可积时, 这个极限不可能依赖于左端点或者右端点. 一个简单的 Abel 变换可以看出, 取右端点的和与取左端点的和有关系:

$$\begin{aligned}
\sum_{i=1}^n B(t_i)(B(t_i) - B(t_{i-1}) &= \sum_{i=1}^n B(t_i)^2 - \sum_{i=1}^n B(t_i)B(t_{i-1}) \\
&= \sum_{i=1}^n B(t_{i-1})^2 + B(b)^2 - B(a)^2 - \sum_{i=1}^n B(t_i)B(t_{i-1}) \\
&= B_b^2 - B_a^2 + \sum_{i=1}^n B(t_{i-1})(B(t_{i-1}) - B(t_i)) \\
&= B_b^2 - B_a^2 - I([a,b],D).
\end{aligned}$$

再来看二次变差

$$\begin{aligned}
\mathrm{var}_2(B;[a,b],D) &= \sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2 \\
&= \sum_{i=1}^n B(t_i)(B(t_i) - B(t_{i-1})) - I([a,b];D) \\
&= B_b^2 - B_a^2 - 2I([a,b],D),
\end{aligned}$$

也就是说

$$I([a,b],D)=\frac{1}{2}[B_b^2-B_a^2-\mathrm{var}_2(B;[a,b],D).$$

当划分列 D_n 趋于零时, 二次变差 $\mathrm{var}_2(B;[a,b],D_n)$ 在 L^2 意义下趋于 $b-a$, 因此我们有下面的引理.

引理 8.4.1 在 L^2 意义下, 取左端点的 R-S 和的极限存在

$$\lim_n I([a,b),D_n)=\frac{1}{2}[B_b^2-B_a^2-(b-a)].$$

由上面左端点和与右端点和的关系推出取右端点的 R-S 和的极限也存在, 但极限不同, 为

$$\frac{1}{2}[B_b^2-B_a^2+(b-a)].$$

这个结论给了我们一线曙光, 由此我们可以想象:

1. 尽管 R-S 和的极限在轨道意义下不一定存在, 但在 L^2 意义下存在, 而且与通常的情况不同, 取左端点的极限和取右端点的极限是不一样的;

2. 极限存在的关键是二次变差的极限存在.

按照习惯, 我们把取左端点和的极限写成积分的形式

$$\int_a^b B(t)dB(t),$$

但是我们也发现这与通常情况不同. 当函数 x 在 $[a,b]$ 上连续且有界变差的时候, 应该有下面的公式:

$$\int_a^b x(t)dx(t)=\frac{1}{2}[x(b)^2-x(a)^2].$$

这也说明, 尽管关于 Brown 运动的积分也是某种形式的 R-S 和的某种意义下的极限, 但它是一种完全新的积分, 与通常的 R-S 积分有本质的差别

§8.5 Itô 恒等式

现在我们来看怎么才能定义一个随机过程 $H = (H_t)$ 关于 Brown 运动 $B = (B_t)$ 的积分. 如果能够定义, 那么对任何 $t \geq 0$, 我们把 $[0, t]$ 上的积分记为

$$\int_0^t H_s dB_s$$

或者 $(H.B)_t$, 这形成一个新的随机过程 $H.B = ((H.B)_t)$, 称为 H 关于 B 的随机积分过程, 或者随机积分.

先粗略地讲讲思想. 我们想要证明 R-S 和

$$H^D.B := \sum_D H(t_{i-1})(B(t_i) - B(t_{i-1}))$$

的极限存在, 当然不是轨道意义下, 而是 L^2 收敛或者依概率收敛的意义下. 令

$$H^D(t) := \sum_D H(t_{i-1})1_{[t_{i-1},t_i)}(t),$$

它是一个阶梯型的随机过程, 在区间 $[t_{i-1}, t_i)$ 上是随机变量 $H(t_{i-1})$. 当 H 的轨道左连续的时候, 对几乎所有的轨道, H^D 点点收敛于 H. 直观地看, $H^D.B$ 就是 H^D 关于 B 在轨道意义上的积分, 即

$$H^D.B = \int_0^t H^D(s)dB(s).$$

回想我们是如何证明经典的 R-S 积分存在的, 实际上是利用 Darboux 上和与 Darboux 下和以及被积函数的一致连续性, 其实我们并不知道极限是什么. 现在的情况类似, 要证明上面的和 L^2 收敛, 我们只需要证明它是 L^2 空间上的一个 Cauchy 列.

为此需要算 R-S 和的 L^2 范数. 范数的计算主要依赖于独立增量性, 因为独立增量性, 当 $i < j$ 时, $i \leq j - 1$, 由条件期望的性质以及因为 Brown 运动的鞅性,

$$\mathbb{E}[H(t_{i-1})(B(t_i) - B(t_{i-1}))H(t_{j-1})(B(t_j) - B(t_{j-1}))]$$

$$= \mathbb{E}[H(t_{i-1})(B(t_i) - B(t_{i-1}))H(t_{j-1})\mathbb{E}[(B(t_j) - B(t_{j-1}))|\mathscr{G}_{t_{j-1}}]]$$
$$= 0.$$

再利用独立增量性, $H(t_{i-1})$ 与 $B(t_i) - B(t_{i-1})$ 独立, 因此有

$$\begin{aligned}\mathbb{E}[(H^D.B)^2] &= \sum_D \mathbb{E}[H(t_{i-1})^2(B(t_i) - B(t_{i-1}))^2] \\ &= \sum_D \mathbb{E}[H(t_{i-1})^2]\mathbb{E}[(B(t_i) - B(t_{i-1}))^2] \\ &= \sum_D \mathbb{E}[H(t_{i-1})^2](t_i - t_{i-1}).\end{aligned}$$

看最后一行, 它也是一个 Riemann 和的样子, 也就是说, 它是函数 $t \mapsto \mathbb{E}[H(t)^2]$ 的 Riemann 和. 如果 H 是连续有界的, 那么这个 Riemann 和是有极限的, 虽然这不能理论上保证 $H^D.B$ 有极限, 但至少是一个希望.

怎么办呢? 我们得到一个等式

$$\mathbb{E}[(H^D.B)^2] = \sum_D \mathbb{E}[H(t_{i-1})^2](t_i - t_{i-1}) = \mathbb{E}\int_0^t H^D(s)^2 ds.$$

其左边是 L^2 距离, 但是右边是 H^D 的一个函数, 也像一个 L^2 距离. 不妨给随机过程空间定义一个范数

$$\|H\|_2 := \left(\mathbb{E}\int_0^t H(s)^2 ds\right)^{1/2},$$

那么当 H 左连续的时候有

$$\lim_D \|H^D - H\|_2 = 0.$$

现在我们有一个漂亮的等式, 称为是 Itô 恒等式, 它是定义随机积分的关键.

定理 8.5.1 对 $[0,t]$ 的划分 D 有

$$\|H^D.B\|_{L^2} = \|H^D\|_2. \tag{8.5.1}$$

用算子的语言说, 这个等式是指算子

$$H^D \mapsto H^D.B$$

是一个某个线性空间到 L^2 空间的保范线性算子.

习题

1. 证明: 上面所说的

$$\sup_D \sum_D \sqrt{t_i - t_{i-1}} = \infty.$$

2. 证明上面所言: 如果 f 连续, 那么 $\lim_n \max_{D_n} |f(t_i) - f(t_{i-1})| = 0$.

3. 证明:

$$\mathbb{E}[\mathrm{var}(B; [a,b], D) - (b-a)]^2 = 2\sum_D (t_i - t_{i-1})^2.$$

4. 取二分点划分列 $D_n := \{k2^{-n} : 0 \le k \le 2^n\}$, 证明: $\mathrm{var}(B; [0,1], D_n)$ 几乎处处收敛于 1. 提示: 应用 Borel-Cantelli 引理.

5. 当函数 x 在 $[a,b]$ 上连续且有界变差的时候, 证明:

$$\int_a^b x(t)dx(t) = \frac{1}{2}[x(b)^2 - x(a)^2].$$

第九章　随机积分

§9.1　Itô 积分

因为 $\lim_D \left\|H^D - H\right\|_2 = 0$, 所以对任何 $\varepsilon > 0$, 存在 $\delta > 0$, 只要 $|D| < \delta$ 就有 $\left\|H^D - H\right\|_2 < \varepsilon$. 因此对任何两个划分 D, D' 只要 $|D|, |D'| < \delta$ 就有

$$\left\|H^D - H^{D'}\right\|_2 < 2\varepsilon.$$

由 Itô 恒等式推出

$$\left\|H^D.B - H^{D'}.B\right\|_{L^2} < 2\varepsilon.$$

这说明 $\{H^D.B :$ 划分 $D\}$ 是 L^2 空间中的 Cauchy 列, 由 L^2 空间的完备性, 推出 $H^D.B$ 的 L^2 极限存在唯一, 记为 $H.B$, 但是因为这个极限是区间 $[0,t]$ 上的划分得来的, 所以 $H^D.B$ 应该是 $(H^D.B)_t$ 且极限应该是 $(H.B)_t$, 即 $(H^D.B)_t$ 是 L^2 收敛于 $(H.B)_t$. 因此, 实际上 $H^D.B = \{(H^D.B)_t\}$ 和 $H.B = \{(H.B)_t\}$ 也是随机过程, 这个随机过程被称为是 H 关于 B 的 Itô 型随机积分, 和通常的积分类似, 记为

$$(H.B)_t = \int_0^t H_s dB_s.$$

当然, 也有相应的 Itô 恒等式

$$\mathbb{E}\left(\int_0^t H_s dB_s\right)^2 = \mathbb{E}\int_0^t H_s^2 ds. \tag{9.1.1}$$

上面的计算有效是依赖于以下三个关键点:

1. H 适应于 $(\mathscr{G}_t)$;

2. 几乎所有轨道左连续;

3. R-S 和是取 H 的左端点;

4. 对任何 $t \geq 0$, $\|H\|_2 < \infty$.

定义 9.1.1 如果有一个适应随机过程 H 满足当 D 趋于零时, $H^D.B$ 在依概率收敛的意义下有极限, 那么我们说 H 是 Itô 可积的.

如果对任何 $t \geq 0$, $\lim_D \left\|H^D - H\right\|_2 = 0$, 那么 $H^D.B$ 在 L^2 收敛的意义下有极限, 因此 H 是可积的. 例如当 H 满足上面四个条件时.

§9.2 随机积分的鞅性

对比 $H^D.B$ 和离散时间的随机积分, 不难发现两者是很类似的, 而由 Doob 基本定理, 当 X 是鞅时, 离散时间随机积分 $H.X$ 也是鞅, 所以我们可以期望 $H^D.B$ 也是鞅. 但是按照上面的定义显然是不行的, 因为定义 $(H^D.B)_t$ 时是依赖于划分 D 的, 所以为了看 $(H^D.B)_s$ 和 $(H^D.B)_t$ 的关系时, 我们需要用相同的划分.

现在我们设 D 是 $[0,t]$ 上的一个划分, $s \in [0,t]$. 我们来计算条件期望

$$\mathbb{E}[(H^D.B)_t|\mathscr{G}_s].$$

存在 k 使得 $t_{k-1} \leq s < t_k$, 把条件期望分成三部分

$$\begin{aligned}\mathbb{E}[(H^D.B)_t|\mathscr{G}_s] &= \mathbb{E}[\sum_{i=1}^{k-1} H(t_{i-1})(B(t_i) - B(t_{i-1}))|\mathscr{G}_s] \\ &\quad + \mathbb{E}[H(t_{k-1})(B(t_k) - B(t_{k-1}))|\mathscr{G}_s] \\ &\quad + \mathbb{E}[\sum_{i=k+1}^{n} H(t_{i-1})(B(t_i) - B(t_{i-1}))|\mathscr{G}_s].\end{aligned}$$

第一部分关于 $\mathscr{G}_s$ 可测, 第二部分因为 $H(t_{k-1})$ 是 $\mathscr{G}_s$ 可测的且由 Brown 运动的鞅性得它等于 $H(t_{k-1})(B(s) - B(t_{k-1}))$, 第三部分中的任何 i, $t_{i-1} > s$, 所以

$$\mathbb{E}[H(t_{i-1})(B(t_i) - B(t_{i-1}))|\mathscr{G}_s] = \mathbb{E}[\mathbb{E}[H(t_{i-1})(B(t_i) - B(t_{i-1}))|\mathscr{G}_{t_{i-1}}]|\mathscr{G}_s]$$

$$= \mathbb{E}[H(t_{i-1})\mathbb{E}[(B(t_i) - B(t_{i-1}))|\mathscr{G}_{t_{i-1}}]|\mathscr{G}_s] = 0.$$

因此, 应用取小 $\wedge$ 这个记号

$$\begin{aligned}\mathbb{E}[(H^D.B)_t|\mathscr{G}_s] &= \mathbb{E}[\sum_{i=1}^{k-1} H(t_{i-1})(B(t_i) - B(t_{i-1})) + H(t_{k-1})(B(s) - B(t_{k-1}))|\mathscr{G}_s] \\ &= \mathbb{E}[\sum_{i=1}^{n} H(t_{i-1})(B(t_i \wedge s) - B(t_{i-1} \wedge s))],\end{aligned}$$

因为 $\{t_i \wedge s : 0 \le i \le n\}$ 是 D 在 $[0,s]$ 上的投影, 是 $[0,s]$ 上的划分, 我们记

$$(H^D.B)_s = \sum_{i=1}^{n} H(t_{i-1})(B(t_i \wedge s) - B(t_{i-1} \wedge s)),\ s \in [0,t],$$

固定 t 以及 $[0,t]$ 的划分 D 之后, 它是时间 $[0,t]$ 上的随机过程. 什么的计算结果告诉我们, 它是一个鞅, 再看它的定义, 实际上它是连续的, 因为 $s \mapsto B(t_i \wedge s)$ 是连续的.

这样我们知道 $\{(H^D.B)_s : s \le t\}$ 是连续鞅, 而且当 D 趋于零时, 对任何 $s \in [0,t]$, $(H^D.B)_s$ 在 L^2 意义下收敛于 $(H.B)_s$. 由此推出

$$\mathbb{E}[(H.B)_t|\mathscr{G}_s] = (H.B)_s,$$

即极限 $H.B$ 是一个鞅. 因为 t 是任意的, 所以其实我们可以认为 $(H.B)_t$ 对任何 $t \ge 0$ 都有定义.

看 Brown 运动 (B_t). 它是鞅, $\mathbb{E}[B_t^2] = t$ 且 $(B_t^2 - t)$ 也是鞅, 那我们自然猜测随机过程

$$\left\{ \left(\int_0^t H_s dB_s \right)^2 - \int_0^t H_s^2 ds,\ t \ge 0 \right\}$$

也是鞅. 先用上面证明 Itô 恒等式差不多的方法证明

$$\mathbb{E}[((H^D.B)_t - (H^D.B)_s)^2|\mathscr{G}_s] = \mathbb{E}[\int_s^t H^D(u)^2 du|\mathscr{G}_s].$$

让 D 趋于零就得到下面的恒等式

$$\mathbb{E}[((H.B)_t-(H.B)_s)^2|\mathscr{G}_s]=\mathbb{E}[\int_s^t H(u)^2du|\mathscr{G}_s].$$

因为 $H.B$ 是鞅, 所以必然有

$$\mathbb{E}[((H.B)_t-(H.B)_s)^2|\mathscr{G}_s]=\mathbb{E}[(H.B)_t^2-(H.B)_s^2|\mathscr{G}_s],$$

因此 $\{(H.B)_t-\int_0^t H(s)^2ds\}$ 是鞅.

§9.3 连续性

在上一节中, 我们已经定义了随机过程 H 关于 Brown 运动 B 的随机积分 $H.B$, 它是一个鞅. 因为它是连续鞅 $H^D.B$ 的极限, 所以我们猜测 $H.B$ 也是一个连续鞅.

我们来证明一个一般性的命题. 设有 $X^{(n)}=(X_s^{(n)}:0\le s\le t)$ 是连续的平方可积鞅序列, 且 $X_t^{(n)}$ 在 L^2 意义下收敛于 X_t. 定义 $X_s:=\mathbb{E}[X_t|\mathscr{G}_s]$.

定理 9.3.1 $(X_s:0\le s\le t)$ 也是连续鞅 (有连续修正).

证明的关键是 Kolmogorov 不等式.

引理 9.3.1 如果 (X_t) 是右连续鞅, 那么对任何 $\lambda>0$ 有

$$\mathbb{P}(\max_{0\le s\le t}|X_s|>\lambda)\le\frac{1}{\lambda^2}\mathbb{E}[|X_t|^2].$$

实际上, Kolmogorov 证明了离散时间独立同分布随机序列和的情况, 用它来证明注明的 Kolmogorov 强大数定律, Doob 把它推广到鞅的情况. 把这个不等式应用于鞅 $X^{(n)}-X$, 得

$$\mathbb{P}(\max_{0\le s\le t}|X_s^{(n)}-X_s|>\lambda)\le\frac{1}{\lambda^2}\mathbb{E}[|X_t^{(n)}-X_t|^2],$$

由此推出 $\max|X_s^{(n)} - X_s|$ 依概率收敛于零, 这蕴含着存在子列按几乎所有轨道趋于零. 按一条样本轨道看就是 $X_s^{(n)}(\omega)$ 在区间 $[0,t]$ 上一致收敛于 $X_s(\omega)$, 因此 $s \mapsto X_s(\omega)$ 在 $[0,t]$ 上连续.

这样我们证明了 $H.B$ 不仅存在而且是连续的鞅. 这实际上证明了平方可积连续鞅的空间是一个完备的内积空间, 也就是 Hilbert 空间. 算子 $H \mapsto H.B$ 是两个空间之间的等距线性算子.

现在我们简单地说一下 Kolmogorov 不等式的证明. 令

$$\tau := \inf\{s > 0 : |X_s| > \lambda\} \wedge t.$$

那么由 Doob 定理, 因为 (X_t^2) 是下鞅, 且 $t \geq \tau$, 故

$$\begin{aligned}\mathbb{E}[X_t^2] &\geq \mathbb{E}[X_\tau^2] \\ &= \mathbb{E}[X_\tau^2; \max_{s\leq t}|X_s| > \lambda] \\ &\geq \lambda^2 \mathbb{P}(\max_{s\leq t}|X_s| > \lambda).\end{aligned}$$

§9.4 二次变差过程

在随机积分理论中, 二次变差是一个非常重要的概念, 也是定义随机积分的关键点之一. 对任何 t, Brown 运动 B 在 $[0,t]$ 上的二次变差是 t, 把它也看成是一个随机过程, 我们说 Brown 运动的二次变差过程是 t, 记为

$$\langle B, B\rangle_t = t.$$

其它过程是不是也有二次变差呢? 不管有没有, 我们先给个定义.

定义 9.4.1 设 $X = (X_t)$ 是一个随机过程, 如果存在一个随机过程 $A = (A_t)$ 使得对任何 $t \geq 0$ 以及 $[0,t]$ 上的划分 D, 当 D 趋于零时, 有

$$\sum_D (X(t_i) - X(t_{i-1}))^2$$

依概率收敛于 A_t, 那么我们说 A 是 X 的二次变差过程, 记为

$$\langle X,X\rangle = (\langle X,X\rangle_t),\ \langle X\rangle_t = A_t,$$

有时简单记为 $\langle X\rangle$.

为什么记为 $\langle X,X\rangle$? 这是为了一个更一般的概念准备的. 类似地, 如果 X,Y 是两个连续的随机过程, 那么

$$\sum_D (X(t_i)-X(t_{i-1}))(Y(t_i)-Y(t_{i-1}))$$

依概率收敛的极限存在的话就称为 X,Y 的协变差过程, 记为 $\langle X,Y\rangle$, 当然 $X=Y$ 时就是二次变差过程. 由平行四边形法则, 如果 X,Y 与 $X+Y$ 有二次变差过程, 那么 X,Y 有协变差, 且

$$\langle X,Y\rangle = \frac{1}{2}(\langle X+Y\rangle - \langle X\rangle - \langle Y\rangle).$$

如果 X,Y 至少有一个是有界变差过程, 即它的几乎所有轨道在任何有限区间上是有界变差的, 那么 $\langle X,Y\rangle = 0$. 事实上, 如果 X 是有界变差的, 那么

$$|\sum_D (X(t_i)-X(t_{i-1}))(Y(t_i)-Y(t_{i-1}))| \le \max_D |Y(t_i)-Y(t_{i-1})| \mathrm{var}(X,[0,t]),$$

右边一个趋于零, 另一个有限, 所以趋于零.

二次变差过程的意义基本上是衡量随机过程的样本轨道的粗糙程度的, 因为当随机过程是连续而且有有界变差时, 二次变差是零, 不会出现. 但是这个在经典分析中是个幽灵的概念在随机分析理论中却到处飘荡, 随处可见, 我们必须做好和它打交道的准备. 这里, 我们将证明下面的定理.

定理 9.4.1 随机积分 $H.B$ 有二次变差:

$$\langle H.B\rangle_t = \int_0^t H_s^2 ds.$$

首先对 $[0,t]$ 上的划分 D, 有

$$\begin{aligned}&\mathbb{E}\left[\left(\sum_D (H^D.B)_{t_i}-(H^D.B)_{t_{i-i}})^2-\int_{t_{i-1}}^{t_i}H^D(s)ds\right)^2\right]\\&=2\mathbb{E}\left[\sum_D H(t_{i-1})^4(t_i-t_{i-1})^2\right].\end{aligned}$$

假设 H 有界, D 趋于零, 便可以证明上面的定理. 数学好的读者可能看得出来, 其实这个证明不严格, 后面有时间我们再来严格地证明它.

如果有两个可积的随机过程 (H_t), (K_t), 那么 $H.B$ 与 $K.B$ 的协变差存在:

$$\langle H.B,K.B\rangle_t=\int_0^t H_sK_sds,$$

而且

$$\{(H.B)_t(K.B)_t-\int_0^t H_sK_sds:\ t\geq 0\}$$

是一个鞅.

一个随机过程 (M_t) 关于时间不变是指: 对任何 $t\geq 0$, $M_t\equiv M_0$ a.s. 不是关于时间不变的随机过程称为是非常值的. 一个随机积分是一个平方可积的连续鞅, 用 $\mathbf{M}^2$ 表示平方可积的连续鞅组成的空间, 我们来证明一个事实: 和 Brown 运动一样, 任何的平方可积的连续鞅都不是有界变差的, 除非它是关于时间不变的.

只需证明: 如果 M 在 $[0,T]$ 上有界变差的, 即它的几乎所有样本轨道 (在任何有限区间上) 是有界变差的, 那么对任何 $t\leq T$ 有 $\mathbb{E}[(M_t-M_0)^2]$.

$$\begin{aligned}\mathbb{E}[(M_t-M_0)^2]&=\mathbb{E}[M_t^2-M_0^2]\\&=\mathbb{E}\sum_D(M_{t_i}^2-M_{t_{i-1}}^2)\\&=\mathbb{E}\sum_D(M_{t_i}-M_{t_{i-1}})^2\end{aligned}$$

$$\leq \mathbb{E}[\max_D |M_{t_i} - M_{t_{i-1}}| \text{var}(M, [0,t])],$$

按样本轨道来看, 当 D 趋于零时, 因为 M 连续且有界变差, 所以

$$\lim_D \max_D |M_{t_i} - M_{t_{i-1}}| \text{var}(M, [0,t]) = 0.$$

所以, 只要保证极限和期望可以交换就可以推出

$$\mathbb{E}[(M_t - M_0)^2] = 0.$$

因此, 除非关于时间不变, 否则鞅的轨道总是粗糙的.

定理 9.4.2 一个非常值的平方可积的连续鞅不可能是有界变差的.

怎么保证极限和期望交换? 这里要介绍一种重要的方法: 局部化方法. 对 $n \geq 1$, 定义停时

$$\tau_n := \inf\{s > 0 : |M_s| > n, \text{ 或者 } \text{var}(M, [0,s]) > n\}.$$

那么 $\tau_n \uparrow +\infty$, 停止过程 M^{τ_n} 还是平方可积的连续鞅, 且它以及它的变差过程是被常数 n 控制的, 这样极限和期望就可以交换, 推出 M^{τ_n} 关于时间不变, 再让 n 趋于无穷即可. 随机分析中遇到的很多极限和期望交换的问题都可以用这个方法解决.

习题

1. 设 $M = (M_t)$ 是平方可积鞅, $\{t_n\}$ 是一个递增趋于无穷的数列, 定义
$$X_t := M_t^2 - \sum_{n \geq 1} (M(t_n \wedge t) - M(t_{n-1} \wedge t))^2.$$
证明: (X_t) 是鞅.

2. 设 $X = (X_t)$ 是连续随机过程, 定义 $\tau_n := \inf\{t : |X_t| > n\}$, 证明: (1) $\lim_n \tau_n = \infty$; (2) $|X(\tau_n)| = n$.

第十章　Itô 公式及其应用

§10.1　Itô 公式

前面我们定义了随机积分, $H = (H_t)$ 称为可积, 如果取左端点的 R-S 和 $H^D.B$ 的极限存在, 极限可以是依概率收敛的意义. 前面的结果给出一个可积的充分条件: H 适应且 $\|H^D - H\|_2$ 趋于零. 更明确的充分条件是 H 连续适应且

$$\mathbb{E}\int_0^t H_s^2 ds < \infty.$$

一个特殊情况是当 f 连续, $H_t = f(B_t)$ 时, 可积性条件转化为

$$\mathbb{E}\int_0^t f(B_s)^2 ds = \int_0^t ds \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} f(\sqrt{s}x)^2 e^{-x^2/2} dx < \infty.$$

当 f 是多项式函数时, 这个条件肯定满足; 当 $f(x) = e^{ax}$ 时, 这个条件也是满足的; 当 $f(x) = e^{ax^2}$ 且 $a < 0$ 时, 还是满足的, 但当 $a > 0$ 时, 不一定满足, 因为

$$\mathbb{E}\int_0^t f(B_s)^2 ds = \int_0^t ds \int \frac{1}{\sqrt{2\pi}} e^{(2as-1/2)x^2} dx,$$

所以当 $t > 1/(4a)$ 时, 这个可积性条件就不满足了.

在微积分中有一个基本定理, 它把微分和积分统一在一起. 设 f 可导, g 在 $[a,b]$ 是有界变差的, 那么

$$f(g(b)) - f(g(a)) = \int_a^b f'(g(t)) dg(t).$$

随机积分是不是也有这样的性质呢? 肯定没有, 因为前面有个例子, 算 Brown 运动关于自身的积分,

$$\int_0^t B_t dB_t = \frac{1}{2}(B_t^2 - B_0^2 - t),$$

因此

$$B_t^2 - B_0^2 = 2\int_0^t B_t dB_t + t,$$

前面部分和经典的基本定理一致, 但后面多了点东西, 是二次变差部分. 下面一般公式也是 Itô 给出的, 称为 Itô 公式.

定理 10.1.1 设 f 是二次连续可导的, 则

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)dB_s + \frac{1}{2}\int_0^t f''(B_s)ds. \tag{10.1.1}$$

看看微积分基本定理的证明, 它是利用微分中值定理

$$\begin{aligned} f(g(b)) - f(g(a)) &= \sum_D [f(g(t_i)) - f(g(t_{i-1}))] \\ &= \sum_D f'(\xi_i)(g(t_i) - g(t_{i-1})), \end{aligned}$$

其中 ξ_i 是 $g(t_i)$ 与 $g(t_{i-1})$ 之间的一个点. 因此右边趋于 R-S 积分

$$\int_a^b f'(g(t))dg(t).$$

但是这个证明对随机积分不适用, 因为随机积分必须是取左端点的 R-S 和的极限. 但是我们可以 Taylor 展开,

$$\begin{aligned} &f(B(t_i)) - f(B(t_{i-1})) \\ &\qquad = f'(B(t_{i-1}))(B(t_i) - B(t_{i-1})) + \frac{1}{2}f''(\xi_i)(B(t_i) - B(t_{i-1}))^2, \end{aligned}$$

其中 ξ_i 位于 $B(t_{i-1})$ 与 $B(t_i)$ 之间. 因此

$$\begin{aligned} &f(B(t)) - f(B(0)) \\ &\qquad = \sum_D f'(B(t_{i-1}))(B(t_i) - B(t_{i-1})) + \frac{1}{2}\sum_D f''(\xi_i)(B(t_i) - B(t_{i-1}))^2. \end{aligned}$$

右边的第一项没有问题, 极限是随机积分, 关键是第二项的极限. 我们来证明第二项的极限与和

$$S_3 := \sum_D f''(B(t_{i-1}))(t_i - t_{i-1})$$

的极限一样, 而这个极限是 Riemann 和的极限, 恰好是

$$\int_0^t f''(B_t)dt.$$

先来估计下面的和

$$S_1 := \sum_D [f''(\xi_i) - f''(B(t_{i-1}))](B(t_i) - B(t_{i-1}))^2.$$

用 M_i 表示 f'' 在 $B(t_{i-1})$ 与 $B(t_i)$ 之间的振幅, 那么

$$|S_1| \le \max_i M_i \cdot \mathrm{var}_2(B; [0,t], D).$$

因为 f'' 连续, 所以按照样本轨道看的话, $\lim_D |S_1| = 0$ a.s.. 再来估计下面的和

$$S_2 := \sum_D f''(B(t_{i-1}))[(B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1})].$$

用独立增量性得

$$\begin{aligned} \mathbb{E}[S_2^2] &= \mathbb{E}\sum_D f''(B(t_{i-1}))^2[(B(t_i) - B(t_{i-1}))^2 - (t_i - t_{i-1})]^2 \\ &= 2\mathbb{E}\sum_D f''(B(t_{i-1}))^2(t_i - t_{i-1})^2 \\ &\le 2|D|\mathbb{E}\left(\int_0^t f''(B^D(s))^2 ds\right), \end{aligned}$$

所以 $\lim_D |S_2| = 0$. 剩下的和就是 S_3.

上面的证明也不是完全严格, 因为在涉及到极限与期望交换的时候, 我们没有仔细检查是不是符合交换的条件. 但是, 如果读者想要一个严格的证明, 可以试着用前面介绍过的局部化方法.

由此推出

$$B_t^n - B_0^n = n\int_0^t B_s^{n-1} dB_s + \frac{1}{2}n(n-1)\int_0^t B_s^{n-2} ds.$$

§10.2 关于随机积分的积分

设 f 连续, g 是有界变差函数. 定义

$$q(t) := \int_0^t f(s)dg(s),$$

它也是一个有界变差的函数. 任何的连续函数 p 还可以关于它积分且有下面的公式

$$\int_0^t p(s)dq(s) = \int_0^t p(s)f(s)dg(s).$$

随机积分有没有类似的性质呢?

设 H 关于 Brown 运动 B 可积, 记 $M = H.B$. 现在看可积过程 K 关于 M 的取左端点的 R-S 和

$$\begin{aligned}(K^D.M)_t &= \sum_D K(t_{i-1})(M(t_i) - M(t_{i-1})) \\ &= \sum_D K(t_{i-1})\int_{t_{i-1}}^{t_i} [H(s) - H(t_{i-1})]dB(s) + [(HK)^D.B]_t.\end{aligned}$$

记右边第一个和为 S_t, 如果 K 有界, 即 $|K| \le C$, 那么

$$\begin{aligned}\mathbb{E}[S_t^2] &= \mathbb{E}\left[\sum_D K(t_{i-1})^2 \left(\int_{t_{i-1}}^{t_i} [H(s) - H(t_{i-1})]dB(s)\right)^2\right] \\ &= \mathbb{E}\left[\sum_D K(t_{i-1})^2 \int_{t_{i-1}}^{t_i} (H(s) - H(t_{i-1}))^2 ds\right]\end{aligned}$$

$$\leq C^2\|H-H^D\|_2^2.$$

因此 $K^D.M$ 与 $(KH)^D.B$ 的极限是一样的, 所以我们有下面的公式.

定理 10.2.1 如果 $M_t=\int_0^t H(s)dB(s)$, 且 HK 可积, 那么 K 关于 M 也可积且

$$\int_0^t K(s)dM(s)=\int_0^t K(s)H(s)dB(s),$$

形式地, $dM(s)=H(s)dB(s)$.

§10.3 Itô 半鞅

上面简单的关于 Brown 运动的 Itô 公式是不够的, 例如

$$e^{B_t}-1=\int_0^t e^{B_s}dB_s+\frac{1}{2}\int_0^t e^{B_s}ds,$$

也就是说我们知道 (e^{B_t}) 是下鞅, 但是 Itô 公式对 $\{\exp(B_t-\frac{1}{2}t)\}$ 就没有办法, 所以我们需要一个更强大的公式, 就是把关于 Brown 运动的 Itô 公式推广到半鞅.

所谓半鞅, 就是随机积分和连续有界变差过程的和. 定义 Itô 半鞅:

$$X_t:=X_0+\int_0^t H(s)dB(s)+V_t,$$

其中 H 可积的, V 是连续适应的有界变差过程且 $V_0=0$. 可以看出, X 的二次变差是存在的且

$$\langle X\rangle=\langle H.B\rangle+2\langle H.B,V\rangle+\langle V\rangle=\int_0^t H(s)^2ds,$$

也就是说 X 的二次变差仅由它的随机积分部分 (鞅部分) 产生. Itô 半鞅还不同于通常的半鞅, 是本讲义中杜撰的名词, 并没有在其它文献中看到, 在这里简称为半鞅. 半鞅的连续有界变差部分实际上可以写成为两个连续的递增过程的差.

注意关于 (V_t) 的积分总是可以按照轨道来定义的, 我们称之为通常积分, 以示与随机积分的区别. 另外, 一个过程关于连续的有界变差过程的积分还是连续的有界变差过程, 因此, 一个随机过程 (如果可积) 关于半鞅的积分还是半鞅.

类似于 Brown 运动的 Itô 公式的证明, 我们可以证明以下的 Itô 公式.

定理 10.3.1 设 X 是如上的 Itô 半鞅, f 是二次连续可导的函数, 则有

$$\begin{aligned} f(X_t) - f(X_0) &= \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d\langle X\rangle_s \\ &= \int_0^t f'(X_s)H_s dB_s + \int_0^t f'(X_s)dV_s + \frac{1}{2}\int_0^t f''(X_s)H_s^2 ds. \end{aligned}$$

定理中最后一行的第一项是关于 Brown 运动的随机积分, 所以是鞅, 第二第三项都是连续有界变差过程, 因此 $(f(X_t))$ 是半鞅.

现在我们来看 X 的指数什么时候是鞅. 上面的 Itô 公式告诉我们当且仅当

$$f'(X_s)dV_s + \frac{1}{2}f''(X_s)H_s^2 ds = 0$$

时, $(f(X_t))$ 是鞅. 如果 $f(x) = e^x$, 那么上式等价于

$$dV_s = -\frac{1}{2}H_s^2 ds,$$

因此

$$\exp\left(\int_0^t H_s dB_s - \frac{1}{2}\int_0^t H_s^2 ds\right),\ t \geq 0$$

是鞅, 简单地写, 如果 $M = H.B$, 那么

$$e^{M - \frac{1}{2}\langle M\rangle}$$

是鞅, 这个鞅称为是 M 的指数鞅.

§10.4 平方可积鞅

设 $(\mathscr{G}_t)$ 是流, (B_t) 是关于 $(\mathscr{G}_t)$ 的标准 Brown 运动, 对于任何的 $T>0$, 关于流 $(\mathscr{G}_t)$ 的连续平方可积鞅 $(M_t : t\in[0,T])$ 全体 $\mathbf{M}^2$ 是一个 Hilbert 空间, 对 $M,N\in\mathbf{M}^2$, 内积定义为

$$\langle M,N\rangle := \mathbb{E}[M_T N_T].$$

当 H 可积时, $H.B\in\mathbf{M}^2$. 一般的连续平方可积鞅不一定是这样的形式. 但是, 如果我们选 $(\mathscr{G}_t)$ 就是 (B_t) 生成的流, 那么对 $M\in\mathbf{M}^2$, 存在可积的 H 使得

$$M_t = M_0 + (H.B)_t.$$

设 $(\Omega,\mathscr{F},\mathbb{P})$ 是一个完备的概率空间, (B_t) 是关于某个适应流 $(\mathscr{G}_t)$ 的一个标准 Brown 运动, 用 $(\mathscr{F}_t^0)$ 表示 (B_t) 的自然流, 即

$$\mathscr{F}_t^0 := \sigma(B_s : s\le t).$$

将 $\mathscr{F}$ 中的所有零概率集加入到 $\mathscr{F}_t^0$ 之后重新生成事件域, 记为 $\mathscr{F}_t$, 那么 $(\mathscr{F}_t)$ 称为是 Brown 运动生成的强化流. 当然 (B_t) 关于 $(\mathscr{F}_t)$ 也是标准 Brown 运动.

定理 10.4.1 (K.Itô) 对任何 $T>0$, 如果 ξ 是一个关于 $\mathscr{F}_T$ 可测的平方可积随机变量, 那么存在一个唯一的可积随机过程 $(H_t : t\in[0,T])$ 使得

$$\xi = \mathbb{E}\xi + \int_0^T H(s)dB(s).$$

这个定理称为鞅表示定理, 是 Itô 首先在 1951 年证明的, 离期权定价公式还有 20 年. 定理非常重要, 为什么? 在离散的情形, 这相当于说: 如果 $\{X_n\}$ 是一个平方可积鞅, N 是个正整数, ξ 关于

$$\sigma(X_1,\cdots,X_N)$$

可测的随机变量, 那么存在 $H=(H_n)$ 使得

$$\xi=\mathbb{E}\xi+H_0(X_1-X_0)+H_1(X_2-X_1)+\cdots+H_{N-1}(X_N-X_{N-1}).$$

这实际上是说一个衍生证券是可以对冲的, 所以鞅表示定理实际上是在说以 Brown 运动作为证券模型的市场是完备的, 而且 H 正是所需的投资策略.

鞅表示定理的证明是很数学的, 需要 Hilbert 空间以及 Fourier 变换理论的一些知识. 首先我们不妨假设 $\mathbb{E}\xi=0$. 用 H_1 表示期望等于 0 的 $\mathscr{F}_T$ 可测的平方可积随机变量全体, H_2 表示可以写成为 $\int_0^T H(s)dB(s)$, 也就是说可以表示为随机积分的随机变量全体.

1. 显然, H_1 是 Hilbert 空间.

2. H_2 是 H_1 的闭子空间. 主要要证明闭性, 即证明如果 $\eta_n=\int_0^T H_n(s)dB(s)$ 在 L^2 意义下收敛于随机变量 η, 那么 η 也可以表示为随机积分. 由条件, $\{\eta_n\}$ 是一个 Cauchy 列, 由 Itô 恒等式

$$\|\eta_n-\eta_m\|_{L^2}=\|H_n(s)-H_m(s)\|_2$$

推出 $\{H_n\}$ 是 Cauchy 列, 它有极限 H 而且

$$\eta=\int_0^T H(s)dB(s).$$

3. 如果 $\eta\in H_1$ 与 H_2 正交, 那么 $\eta=0$. 也就是说 $H_2^{\perp}=\{0\}$, 这推出 $H_2=H_1$. 怎么证明呢? 取 $[0,T]$ 上 Riemann 可积函数 f. 令

$$H_t:=\exp\left(\int_0^t f(s)dB(s)-\frac{1}{2}\int_0^t f^2(s)ds\right).$$

由 Itô 公式

$$H_T-1=\int_0^T H_sf(s)dB(s)\in H_2.$$

因此

$$\mathbb{E}[\eta H_T]=\mathbb{E}[\eta(H_T-1)]=0.$$

因为 $\int_0^T f^2(s)ds$ 是个常数, 不是随机的, 所以对所有 Riemann 可积的 f 有

$$\mathbb{E}\left[\eta\cdot\exp\left(\int_0^T f(s)dB(s)\right)\right]=0.$$

现在, 对任何的 n, $0=t_0<t_1<\cdots<t_n=T$, $x_1,x_2,\cdots,x_n\in\mathbb{R}$, 令

$$f(s)=i[x_1 1_{[t_0,t_1)}(s)+x_2 1_{[t_1,t_2)}(s)+\cdots+x_n 1_{[t_{n-1},t_n)}(s)].$$

代入得

$$\mathbb{E}\left[\eta\cdot\exp\left(i\sum_j x_j(B(t_j)-B(t_{j-1}))\right)\right]=0.$$

令 $y_1=x_1,y_2=x_2-x_1,\cdots,y_n=x_n-x_{n-1}$, 那么由 $(x_1,\cdots,x_n)$ 的任意性得 $(y_1,\cdots,y_n)$ 也是任意的, 而且

$$\mathbb{E}\left[\eta\cdot\exp\left(i\sum_j y_jB(t_j)\right)\right]=0.$$

关于 $(B(t_1),\cdots,B(t_n))$ 取条件期望, 得

$$\mathbb{E}\left[\mathbb{E}[\eta|B(t_1),\cdots,B(t_n)]\cdot\exp\left(i\sum_j y_jB(t_j)\right)\right]=0.$$

由 Fourier 变换的性质, 对任何有界 Borel 可测函数 g 有

$$\mathbb{E}[\eta g(B(t_1),\cdots,B(t_n))]=0. \tag{10.4.1}$$

这个结论是非常重要的, 可以作为引理使用.

引理 10.4.1 设有两个随机变量 ξ,η, 如果对任何 $x\in\mathbb{R}$ 有 $\mathbb{E}[\eta e^{ix\xi}]=0$, 那么对任何有界 Borel 可测函数 g 有 $\mathbb{E}[\eta g(\xi)]=0$.

定义 $\mu(A):=\mathbb{E}[\eta 1_A(\xi)]$, 它是 $\mathbb{R}$ 上的一个有限符号测度. 由单调类方法

$$\int f(x)\mu(dx)=\mathbb{E}[\eta f(\xi)],$$

因此
$$\int e^{ixy}\mu(dy) = \mathbb{E}[\eta e^{ix\xi}] = 0.$$
由 Fourier 变换的唯一性, 得 $\mu = 0$.

记使得 $\mathbb{E}[\eta; A] = 0$ 成立的 $A \in \mathscr{F}_T$ 的全体为 $\mathscr{G}$, 那么 $\mathscr{G}$ 是一个单调类且包含 $\sigma\{(B(t_1), \cdots, B(t_n)\}$ 中的所有事件. 再因为 n 任意, $0 < t_1 < \cdots < t_n = T$ 任意, 所以由单调类方法推出 $\mathscr{G} = \mathscr{F}_T$. 也就是说, 对所有的 $A \in \mathscr{F}_T$, 有
$$\mathbb{E}[\eta; A] = 0,$$
这蕴含着 $\eta = 0$.

如果 $M = (M_t : t \in [0, T])$ 是一个关于流 $(\mathscr{F}_t)$ 的平方可积鞅, 那么 $M_T \in \mathscr{F}_T$ 且平方可积. 由 Itô 的表示定理知, 存在 $H = (H_t)$ 使得
$$M_T = \mathbb{E}M_T + \int_0^T H_s dB_s.$$
两边对 $\mathscr{F}_t$ 求条件期望得, 对任何 $t \leq T$ 有
$$M_t = \mathbb{E}M_0 + \int_0^t H_s dB_s,$$
即关于 Brown 运动流的平方可积鞅总是可以表示为一个过程关于 Brown 运动的随机积分.

习题

1. 如果 f 是连续函数, 问 $\int_0^t f(s)dB_s$ 服从什么分布? 为什么? 如果 f 连续可导, 证明:
$$\int_0^t f(s)dB_s = f(t)B_t - \int_0^t B_s f'(s)ds.$$

2. 证明:
$$\mathbb{E}[S_2^2] = 2\mathbb{E}\sum_D f''(B(t_{i-1}))^2(t_i - t_{i-1})^2.$$

3. 模仿 Brown 运动定义 H 关于连续平方可积鞅 $M=(M_t)$ 的随机积分

$$\int_0^t H_s dM_s,$$

说明定义的关键是 M 的二次变差过程存在, 并给出 H 关于 M 可积的充分条件.

4. 鞅表示定理的证明.

第十一章　等价测度变换

§11.1　关于平方可积鞅的随机积分

设 $M = (M_t : t \in [0,T])$ 是一个关于流 $(\mathscr{G}_t)$ 适应的连续的平方可积鞅. 从上面看到的那些结论我们看出, 我们可以定义 $H = (H_t)$ 关于 M 的随机积分. 仿照 M 是 Brown 运动的情况, 我们说 H 关于 M 可积是指当 D 趋于零时, 取左端点的 R-S 和

$$(H^D.M)_t = \sum_D H(t_{i-1})(M(t_i) - M(t_{i-1}))$$

有至少是依概率收敛的意义下的极限.

一个重要的结论是连续平方可积鞅 M 的二次变差过程 $\langle M \rangle$ 是存在的, 且

$$M_t^2 - \langle M \rangle_t,\ t \in [0,T]$$

也是鞅. 这个结论是随机积分最基本的定理, 证明较难, 这里省略不提.

定理 11.1.1　设 $M = (M_t)$ 是连续平方可积鞅, 那么存在唯一的连续递增适应随机过程, 记为 $\langle M \rangle = (\langle M \rangle_t)$, 使得 $(M_t^2 - \langle M \rangle_t)$ 是鞅, 且 $\langle M \rangle$ 恰好是 M 的二次变差过程.

由此推出, 如果 M, N 是两个平方可积鞅, 那么它们的协变差是有的

$$\langle M, N \rangle_t = \frac{1}{2}(\langle M + N \rangle_t - \langle M \rangle_t - \langle N \rangle_t),$$

它是一个连续有界变差过程且它是使得

$$(M_t N_t - V_t : t \geq 0)$$

是一个鞅的唯一的连续有界变差过程 V.

由此结论我们同样可以证明下面的 Itô 恒等式

$$\mathbb{E}[(H^D.M)_t^2] = \mathbb{E}\left(\int_0^t H^D(s)^2 d\langle M\rangle_s\right).$$

让 D 趋于零, 同样可以定义随机积分

$$(H.M)_t = \int_0^t H_s dM_s,$$

这里当 H 是适应且满足

$$\mathbb{E}\int_0^t [H^D(s) - H(s)]^2 d\langle M\rangle_s \longrightarrow 0,\ t \geq 0$$

时肯定是关于 M 可积的. 同样可以证明

$$(H.M)_t - \int_0^t H_s^2 dM_s,\ t \geq 0$$

是一个鞅, 因此随机积分的二次变差过程就是

$$\langle H.M\rangle_t = \int_0^t H_s^2 d\langle M\rangle_s.$$

关于连续平方可积鞅的随机积分也有 Itô 公式: 如果 f 二次连续可导, 那么

$$f(M_t) - f(M_0) = \int_0^t f'(M_s)dM_s + \frac{1}{2}\int_0^t f''(M_s)d\langle M\rangle_s.$$

只要有上面的定理, 关于连续平方可积鞅的随机积分理论和关于 Brown 运动的随机积分理论是平行的.

§11.2 连续半鞅

为了方便, 我们引入连续半鞅的概念. 一个随机过程 X 称为是连续半鞅, 如果它可以写成为一个连续平方可积鞅 M 和一个初值为零的连续适应的有界变差过程 V 的和

$$X_t = M_t + V_t.$$

M 和 V 分别称为是连续半鞅的鞅部分与有界变差部分. 要求 V 的初值为零是为了保证这个分解的唯一性.

随机过程 H 关于 X 可积是指它关于鞅 M 和关于有界变差过程 V 都可积, 这时

$$\int_0^t H_s dX_s := \int_0^t H_s dM_s + \int_0^t H_s dV_s,$$

分别称为随机积分和通常积分. 因为他们分别是鞅和有界变差的, 所以 $H.X$ 仍然是连续半鞅.

说到二次变差, 连续半鞅的二次变差是存在的, 而且就是其鞅部分的二次变差

$$\langle X \rangle = \langle M \rangle + 2\langle M, V \rangle + \langle V \rangle = \langle M \rangle.$$

另外, 如果 H 关于 X 可积, 且 Y 也是连续半鞅, 那么

$$\langle H.X, Y \rangle_t = \int_0^t H_s d\langle X, Y \rangle.$$

要证明这个, 思路如下. 先把 $(H^D.X)_t$ 写成为半鞅的线性组合

$$\begin{aligned}(H^D.X)_t &= \sum_D H(t_{i-1})(X(t \wedge t_i) - X(t \wedge t_{i-1})) \\ &= \sum_D H(t_{i-1})(X_t^{t_i} - X_t^{t_{i-1}}),\end{aligned}$$

其中 X^τ 表示停止过程: $X^\tau(t) = X(\tau \wedge t)$. 然后根据二次变差的线性性质得

$$\begin{aligned}\langle H^D.X, Y \rangle_t &= \sum_D H(t_{i-1})(\langle X^{t_i}, Y \rangle_t - \langle X^{t_{i-1}}, Y \rangle_t) \\ &= \sum_D H(t_{i-1})(\langle X, Y \rangle_t^{t_i} - \langle X, Y \rangle_t^{t_{i-1}}) \\ &= \int_0^t H^D(s) d\langle X, Y \rangle_s.\end{aligned}$$

由此推出, 如果 H 关于半鞅 X 可积, K 关于半鞅 Y 可积, 那么

$$\langle H.X, K.Y\rangle_t = H.\langle X, K.Y\rangle_t = (HK).\langle X, Y\rangle_t = \int_0^t H_s K_s d\langle X, Y\rangle_s.$$

关于连续半鞅 X 的积分也一样有 Itô 公式: 如果 f 二次连续可导, 那么

$$f(X_t) - f(X_0) = \int_0^t f'(X_s) dM_s + \frac{1}{2}\int_0^t f''(X_s) d\langle X\rangle_s.$$

当 $M_t = M_0$ 时, $\langle X\rangle = 0$, Itô 公式就是通常的 Newton-Leibniz 公式, 或者说微积分基本定理.

特别地, 令 $f(x) = x^2$, 有

$$X_t^2 - X_0^2 = 2\int_0^t X_s dX_s + \langle X\rangle_t.$$

这样如果 X, Y 都是半鞅, 那么

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y\rangle_t,$$

这个公式可以称为是随机积分的分部积分公式.

§11.3 Brown 运动的鞅刻画

连续平方可积鞅的很多性质类似 Brown 运动, 下面是 Lévy 的 Brown 运动鞅刻画.

定理 11.3.1 (P.Lévy) 设 M 是连续平方可积鞅. 如果 $\langle M\rangle_t = t$, 那么 M 就是 Brown 运动.

利用 Itô 公式一样可以证明, 当 M 是鞅时,

$$e^{M_t - \frac{1}{2}\langle M\rangle_t},\ t \geq 0$$

也是一个鞅, 严格地说是局部鞅, 当某种可积性条件比如 Novikov 条件满足时才是真正的鞅, 称为 M 的指数鞅. 自然对任何实数 x,

$$e^{ixM_t + \frac{1}{2}x^2\langle M\rangle_t}$$

是鞅. 由条件 $\langle M\rangle_t = t$ 推出

$$e^{ixM_t+\frac{1}{2}x^2t}$$

是鞅, 即对 $t>s$ 有

$$\mathbb{E}[e^{ixM_t+\frac{1}{2}x^2t}|\mathscr{G}_s] = e^{ixM_s+\frac{1}{2}x^2s}.$$

变型得

$$\mathbb{E}[e^{ix(M_t-M_s)}|\mathscr{G}_s] = e^{-\frac{1}{2}x^2(t-s)}.$$

从此式可以获得以下信息

1. $M_t - M_s$ 服从正态分布 $N(0, t-s)$;
2. $M_t - M_s$ 与 $\mathscr{G}_s$ 独立.

第二条性质来自于特征函数的唯一性, 有兴趣读者应该自己验证. 这两条性质再加上连续性假设就得到 (M_t) 就是 Brown 运动.

§11.4 测度变换

当 M 是连续的平方可积鞅时, 它的指数鞅

$$Z_t := \exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right)$$

当下面的可积性条件满足时是鞅:

$$\mathbb{E}\int_0^T Z_s^2 d\langle M\rangle_s < \infty.$$

另外还有当所谓的 Novikov 条件:

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M\rangle_T\right)\right] < \infty.$$

满足时, (Z_t) 也是鞅, 实际上是指 $(Z_s : s\in[0,T])$ 是鞅.

随机过程 $Z=(Z_t)$ 是严格正的, 因为是鞅, 所以 $\mathbb{E}Z_T=\mathbb{E}Z_0=1$. 我们可以用它来定义一个新的测度

$$\mathbb{Q}(A):=\mathbb{E}[Z_T;A],\ \ A\in\mathscr{G}_T.$$

容易验证 $\mathbb{Q}$ 是一个与 $\mathbb{P}$ 等价的概率测度, 所以称为等价测度变换.

先想想如果测度换了, 哪些东西会改变, 哪些不会改变? 鞅由条件期望定义, 它肯定改变了. 在 $\mathbb{P}$ 测度的 Brown 运动在新测度下不一定是 Brown 运动; 二次变差过程由极限定义, 只与零概率集有关, 不会改变; 随机积分也是只与零概率集有关, 不会改变.

下面我们给出几个引理.

引理 11.4.1 当 $A\in\mathscr{G}_t$ 时,

$$\mathbb{Q}(A)=\mathbb{E}[Z_t;A].$$

这很容易, 因为 $(Z_t:t\in[0,T])$ 是鞅, 所以

$$\begin{aligned}\mathbb{Q}(A)&=\mathbb{E}[Z_T1_A]=\mathbb{E}[\mathbb{E}[Z_T1_A|\mathscr{G}_t]]\\&=\mathbb{E}[1_A\mathbb{E}[Z_T|\mathscr{G}_t]]=\mathbb{E}[1_AZ_t].\end{aligned}$$

引理 11.4.2 $X=(X_t)$ 是 $\mathbb{Q}$- 鞅当且仅当 (Z_tX_t) 是 $\mathbb{P}$- 鞅.

设 (X_t) 是 $\mathbb{Q}$- 鞅. 验证 (Z_tX_t) 是 $\mathbb{P}$- 鞅, 只需证明对 $t>s$ 有

$$\mathbb{E}[Z_tX_t|\mathscr{G}_s]=Z_sX_s.$$

这等价于证明对 $A\in\mathscr{G}_s$ 有

$$\mathbb{E}[Z_tX_t1_A]=\mathbb{E}[Z_sX_s1_A].$$

事实上, 由前面的引理以及鞅定义 $\mathbb{E}^{\mathbb{Q}}[X_t|\mathscr{G}_s]=X_s$,

$$\mathbb{E}[Z_sX_s1_A]=\mathbb{E}^{\mathbb{Q}}[X_s1_A]=\mathbb{E}^{\mathbb{Q}}[X_t1_A]=\mathbb{E}[Z_tX_t1_A].$$

定理 11.4.1 (Girsanov) 如果 N 是 $\mathbb{P}$- 连续平方可积鞅, 那么

$$\tilde{N}_t = N_t - \langle M, N\rangle_t,\ T \geq t \geq 0$$

是 $\mathbb{Q}$- 鞅. 由此推出, $\mathbb{P}$- 连续半鞅类与 $\mathbb{Q}$- 连续半鞅的类是一样的.

利用上面的引理, 只需验证 $(\tilde{N}_t Z_t)$ 是 $\mathbb{P}$- 鞅就可以了. 因为两个过程都是 $\mathbb{P}$- 连续半鞅, 利用分部积分公式

$$\begin{aligned}\tilde{N}_t Z_t - \tilde{N}_0 Z_0 &= \int_0^t Z_s d\tilde{N}_s + \int_0^t \tilde{N}_s dZ_s + \langle \tilde{N}, Z\rangle_t \\ &= \int_0^t Z_s dN_s - \int_0^t Z_s d\langle M, N\rangle_s + \int_0^t \tilde{N}_s dZ_s + \langle N, Z\rangle_t \\ &= \int_0^t Z_s dN_s + \int_0^t \tilde{N}_s dZ_s,\end{aligned}$$

最后一个等号是因为

$$Z_t - Z_0 = \int_0^t Z_s dM_s,$$

故而

$$\langle N, Z\rangle_t = \int_0^t Z_s d\langle M, N\rangle_s.$$

特别地, 取

$$M_t = \int_0^t H_s dB_s.$$

那么 $B = (B_t)$ 在新概率测度下不是 Brown 运动, 但它是一个连续半鞅, 因为

$$\tilde{B}_t = B_t - \langle M, B\rangle_t = B_t - \int_0^t H_s ds$$

是一个 $\mathbb{Q}$- 连续平方可积鞅. 进一步, $\tilde{B}$ 的二次变差过程 $\langle \tilde{B}\rangle = \langle B\rangle_t = t$, 再由 Brown 运动的鞅刻画定理推出, $\tilde{B}$ 在新测度 $\mathbb{Q}$ 下是一个 Brown 运动. 这个结果在金融数学中很有用.

§11.5 泛函的观点看随机积分

在这一章的最后, 我们介绍看待随机积分的另外一种观点. 设时间区间为 $[0,T]$, $\mathscr{H}$ 是初值等于零的连续平方可积鞅全体, 对于空间中的两个元素 $M=(M_t)$ 与 $N=(N_t)$ 定义内积

$$\langle M,N\rangle_{\mathscr{H}} := \mathbb{E}[M_T N_T].$$

可以验证, 这是一个 Hilbert 空间. 固定 $M\in\mathscr{H}$, 与随机过程 $H=(H_t)$ 满足

$$\mathbb{E}\left[\int_0^T H_t^2 d\langle M\rangle_t\right] < \infty.$$

定义 $\mathscr{H}$ 上的泛函

$$\varphi(N) := \mathbb{E}\left[\int_0^T H_t d\langle M,N\rangle_t\right],\ N\in\mathscr{H}.$$

它是一个线性泛函且是有界的. 为什么有界? 让我们来证明著名的 Kunita-Watanabe 不等式: 设 $X,Y\in\mathscr{H}$, A,B 是两个随机过程, 那么

$$\left[\mathbb{E}\left(\int_0^T A_t B_t d\langle X,Y\rangle_t\right)\right]^2 \le \mathbb{E}\left(\int_0^T A_t^2 d\langle X\rangle_t\right)\cdot\mathbb{E}\left(\int_0^T B_t^2 d\langle Y\rangle_t\right).$$

事实上, 由二次变差和协变差过程的定义, 可以由 Cauchy 不等式验证对任何 $t>s$ 有

$$|\langle X,Y\rangle_t - \langle X,Y\rangle_s| \le [(\langle X\rangle_t - \langle X\rangle_s)(\langle Y\rangle_t - \langle Y\rangle_s)]^{1/2}.$$

因此, 再应用 Cauchy 不等式, 然后 R-S 和取极限, 推出

$$\int_0^T A_t B_t \langle X,Y\rangle_t \le \left(\int_0^T A_t^2 d\langle X\rangle_t \int_0^T B_t^2 d\langle Y\rangle_t\right)^{1/2},$$

两边取期望, 再次使用 Cauchy 不等式就得到 Kunita-Watanabe 不等式.

由这个不等式推出

$$|\varphi(N)| \le \left(\mathbb{E}\int_0^T H_t^2 d\langle M\rangle\right)^{1/2}\cdot\mathbb{E}[\langle N\rangle_T]^{1/2}.$$

但是 $N_t^2 - \langle N \rangle_t$ 是鞅,

$$\mathbb{E}[\langle N \rangle_T] = \mathbb{E}[N_T^2] = \|N\|_{\mathscr{H}}^2,$$

所以

$$|\varphi(N)| \le \left(\mathbb{E}\int_0^T H_t^2 d\langle M \rangle\right)^{1/2} \cdot \|N\|_{\mathscr{H}},$$

即 φ 是 $\mathscr{H}$ 上有界线性泛函. 由 Riesz 表示定理, 存在唯一的 $X \in \mathscr{H}$ 使得

$$\varphi(N) = \langle X, N \rangle_{\mathscr{H}}.$$

实际上, X 就是 H 关于 M 的随机积分. 为什么?

习题

1. 取 $[0, T]$ 上划分 D. 证明:

$$\langle H^D.M, N \rangle_t = \int_0^t H_s^D d\langle M, N \rangle_s.$$

2. 如果 $B^{(1)}$ 与 $B^{(2)}$ 是两个独立的标准 Brown 运动, 证明: $\langle B^{(1)}, B^{(2)} \rangle = 0$.

3. 设 $M \in \mathscr{H}$, $\langle M \rangle = 0$. 证明: 对任何 $t > 0$, $M_t = 0$.

第十二章　随机微分方程和 Black-Scholes 公式

§12.1　随机微分方程

直观地说, 随机微分方程是常微分方程加上一个由 Brown 驱动的随机扰动, 即

$$dX_t = f(X)dt + g(X)dB_t,$$

其中 $B = (B_t)$ 是 Brown 运动. 当 $g = 0$ 时就是一个常微分方程. 但是 f, g 对于 X 的依赖性的不同导致不同类型的随机微分方程, 最简单的一类是 Itô 型随机微分方程

$$dX_t = f(X_t)dt + g(X_t)dB_t,$$

其中 f, g 是状态空间上的函数. 这一类的随机微分方程的解一定有 Markov 性, 也就是说给定现在, 将来与过去是独立的. Markov 性是一种信念: 你是否相信市场是 Markov 的? 相信市场是 Markov 的, 也就是说, 市场现在的状态已经表达了市场的一切, 再研究其过去对预测其将来毫无帮助. Black-Scholes 模型就是相信市场是 Markov 的.

随机微分方程的解有不同的理解: 强解与弱解. 所谓强解, 就是随意给定一个概率空间以及概率空间上的 Brown 运动 $B = (B_t)$, 都存在一个随机过程 $X = (X_t)$ 满足以下方程

$$X_t - X_0 = \int_0^t f(X_s)ds + g(X_s)dB_s.$$

而所谓若解, 是指可以找到一个概率空间和概率空间上的两个随机过程 $X = (X_t)$ 与 $B = (B_t)$ 使得 B 是 Brown 运动, 且满足方程

$$X_t - X_0 = \int_0^t f(X_s)ds + g(X_s)dB_s.$$

注意前者是任意给定 B 都有解 X, 后者是同时存在 B 和 X 满足方程.

著名的例子是 Tanaka 方程:

$$dX_t = \mathrm{sgn}(X_t)dB_t.$$

这个方程没有强解, 但是有弱解. 存在概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 和上面的 Brown 运动 $X = (X_t)$, 定义

$$B_t := \int_0^t \mathrm{sgn}(X_s)dX_s.$$

那么由 Brown 运动的鞅刻画, 因为 $\langle B\rangle_t = \langle X\rangle_t = t$, 所以 (B_t) 是标准 Brown 运动且

$$X_t - X_0 = \int_0^t \mathrm{sgn}(X_s)^2 dX_s = \int_0^t \mathrm{sgn}(X_s)dB_s.$$

这里 $B = (B_t)$ 不是任意的 Brown 运动, 它是由 X 决定的.

和微分方程一样, 绝大多数随机微分方程的解是写不出解析表达式的, 只有一些存在唯一性的理论. 例如, 当上面的 f, g 满足整体 Lipschitz 条件和线性增长条件时, 给定初值的 Itô 型方程的强解存在唯一. 证明方法和常微分方程一样, 使用 Picard 迭代.

现在我们举两个简单方程. 第一个是金融领域著名的 Black-Scholes 方程

$$\frac{dX_t}{X_t} = \mu dt + \sigma dB_t.$$

方程左边代表瞬时增长率, 所以方程的直观意义是 X_t 的增长率是一个常数加上一个随机扰动. 如果 $\sigma = 0$, 那么方程是确定性的, 它的解为 $X_t = ce^{\mu t}$. 方程实际上可以写成

$$X_t - X_0 = \int_0^t \mu X_s ds + \sigma X_s dB_s.$$

因此, (X_t) 是连续半鞅, 其二次变差为

$$\langle X\rangle_t = \int_0^t \sigma^2 X_s^2 ds.$$

由 Itô 公式

$$\begin{aligned}\log X_t - \log X_0 &= \int_0^t \frac{dX_s}{X_s} - \frac{1}{2}\int_0^t \frac{d\langle X\rangle_s}{X_s^2} \\ &= \int_0^t \mu ds + \sigma dB_s - \frac{1}{2}\sigma^2 ds \\ &= (\mu - \frac{1}{2}\sigma^2)t + \sigma B_t,\end{aligned}$$

由此推出

$$X_t = X_0 \exp\left[(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\right],$$

本质上是一个几何 Brown 运动.

另外一个方程是

$$dX_t = dB_t + \beta X_t dt.$$

应用分部积分公式

$$\begin{aligned}d(e^{-\beta t}X_t) &= -\beta e^{-\beta t}X_t dt + e^{-\beta t}dX_t \\ &= e^{-\beta t}dB_t,\end{aligned}$$

因此有

$$e^{-\beta t}X_t - X_0 = \int_0^t e^{-\beta s}dB_s,$$

即

$$X_t = e^{-\beta t}\left(X_0 + \int_0^t e^{\beta(t-s)}dB_s\right).$$

这两个方程的解当然都是强解, 而且可以把解析表达式写出来, 尽管第二个例子的随机积分符号仍然存在. 下面我们用 Girsanov 定理对方程

$$dX_t = dB_t + b(X_t)dt$$

构造弱解. 找一个概率空间和 Brown 运动 $X = (X_t)$, 定义

$$N_t = \int_0^t b(X_t)dX_t,$$

需要保证它的指数鞅

$$Z_t = e^{N_t - \frac{1}{2}\langle N \rangle_t}$$

是真正的鞅. 利用 (Z_t) 作为密度构造一个新的概率 $\mathbb{Q}$, 在这个概率之下,

$$B_t := X_t - X_0 - \int_0^t b(X_t)dt$$

是一个标准 Brown 运动, 而且有

$$X_t - X_0 = B_t + \int_0^t b(X_t)dt,$$

即说 (X_t) 在新的概率 $\mathbb{Q}$ 之下是方程的解, 显然是个弱解.

§12.2 Black-Scholes 公式

Black-Scholes 公式是历史上第一个期权定价的显式表达式, 于 1972 年发布, 恰逢 Chicago 期权期货交易所开张的年代, 引起轰动, 使大家对随机分析在金融中的应用产生广泛的关注, 进入了一个辉煌的时代, 直到金融危机的爆发, 热度才慢慢下降, 反思的声音渐渐地多了起来.

如果 $S = (S_t)$ 是一个随机过程, 代表某个风险资产价格, r 代表连续复利的利率, $H = (H_t)$ 代表投资策略, $X = (X_t)$ 代表由初始投资 X_0 和投资策略 H 所诱导的自融资方式下的财富积累过程, 那么折现的财富过程是 H 关于折现的资产价格的随机积分, 即

$$\tilde{X}_t = X_0 + \int_0^t H_u d\tilde{S}_u,$$

其中 $\tilde{X}$ 和 $\tilde{S}$ 代表折现资产

$$\tilde{X}_t := e^{-rt}X_t,\ \tilde{S}_t := e^{-rt}S_t,\ t \geq 0.$$

到期日为 T 的衍生证券是

$$\mathscr{F}_T^S := \sigma(S_t : 0 \leq t \leq T)$$

可测的随机变量 V, 它有对冲是指存在 X_0 和 H 使得 $X_T = V$, 那么有

$$e^{-rT}V = X_0 + \int_0^T H_t d\tilde{S}_t.$$

对冲是指衍生证券所锁定的价值可以通过对资产自融资投资来实现. 问题归结为衍生证券是不是总是有对冲?

讨论在一般情况下是不是可以对冲是没有意义的, 必须考虑具体的模型. 假设 (S_t) 满足 Black-Scholes 方程

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t,$$

其中 μ 称为收益率, σ 称为波动率. (S_t) 有表达式

$$S_t = \exp\left((\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\right).$$

折现后

$$\begin{aligned}\tilde{S}_t &= \exp\left((\mu - r - \frac{1}{2}\sigma^2)t + \sigma B_t\right) \\ &= \exp\left(\sigma \tilde{B}_t - \frac{1}{2}\sigma^2 t\right),\end{aligned}$$

其中

$$\tilde{B}_t := B_t - (r - \mu)\sigma^{-1}t.$$

记 $\mu' := (r - \mu)\sigma^{-1}$. 由 Girsanov 定理, 存在新的测度

$$\mathbb{Q} := \exp(\mu' B_T - \frac{1}{2}\mu'^2 T) \cdot \mathbb{P}$$

使得 $\tilde{B}$ 是鞅, 但是 $\langle\tilde{B}\rangle_t = \langle B\rangle_t = t$, 所以 $\tilde{B}$ 在新测度下是标准 Brown 运动且 $(\tilde{S}_t)$ 是 $(\sigma\tilde{B}_t)$ 的指数鞅, 即

$$d\tilde{S}_t = \sigma\tilde{S}_t d\tilde{B}_t.$$

显然 $\mathscr{F}_T^S = \mathscr{F}_T^{\tilde{B}}$, Brown 运动有鞅表示性质, 所以在新测度空间下, 存在 (K_t) 使得

$$V = \tilde{\mathbb{E}}V + \int_0^T K_t d\tilde{B}_t = \tilde{\mathbb{E}}V + \int_0^T K_t(\sigma\tilde{S}_t)^{-1}d\tilde{S}_t,$$

其中 $\tilde{\mathbb{E}}$ 是关于新测度的期望, 因此推出

$$e^{-rT}V = e^{-rT}\tilde{\mathbb{E}}V + \int_0^T e^{-rT}K_t(\sigma\tilde{S}_t)^{-1}d\tilde{S}_t,$$

令 $H_t := e^{-rT}K_t(\sigma\tilde{S}_t)^{-1}$, 即说明了 V 是可以对冲的

$$e^{-rT}V = e^{-rT}\tilde{\mathbb{E}}V + \int_0^T H_t d\tilde{S}_t,$$

且定价表达式为

$$X_0 = e^{-rT}\tilde{\mathbb{E}}V,$$

直观地说, 是衍生证券在新测度下期望的折现值.

最后, 如果 V 是敲定价格 K 的欧式买入期权

$$V = (S_T - K)^+ = (e^{rT}\tilde{S}_T - K)^+ = (S_0 e^{rT-\sigma^2/2+\sigma\tilde{B}_T} - K)^+,$$

那么让我们来计算 $\tilde{\mathbb{E}}V$. 注意 $\tilde{B}_T$ 在新测度下的分布是 $N(0,T)$, 再引入一个记号

$$-d := \frac{\log\frac{K}{S_0} - (r-\sigma^2)T/2}{\sigma\sqrt{T}},$$

所以

$$\begin{aligned}
\tilde{\mathbb{E}}V &= \int_{\mathbb{R}} (S_0 e^{rT-\sigma^2 T/2+\sigma\sqrt{T}x} - K)^+ \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)dx \\
&= S_0 e^{rT}\int_{x>-d}\frac{1}{\sqrt{2\pi}}\exp[-\frac{1}{2}(x^2 - 2\sigma\sqrt{T}x + \sigma^2 T)]dx \\
&\quad - K\int_{x>-d}\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)dx \\
&= S_0 e^{rT}(1-\Phi(-d-\sigma\sqrt{T})) - K(1-\Phi(-d))
\end{aligned}$$

$$= S_0 e^{rT}\Phi(d+\sigma\sqrt{T}) - K\Phi(d),$$

其中 Φ 是标准正态分布函数.

这样, 我们得到下面的定理.

定理 12.2.1 (Black & Scholes) 欧式买入期权的无风险定价为

$$X_0 = e^{-rT}\tilde{\mathbb{E}}V = S_0\Phi(d+\sigma\sqrt{T}) - Ke^{-rT}\Phi(d),$$

此表达式称为是 Black-Scholes 公式.

要指出的一点是, 这个表达式仅与 r, σ 有关, 而与资产的收益率 μ 无关, 为什么? 这是一个值得深思的现象.

上面给出的方法是应用随机分析的理论, 通过这么一个简单的例子, 把随机分析怎么应用于金融定价理论的机制表达得非常清楚: 无风险定价存在的理由是鞅表示定理, 定价计算的方法是等价鞅测度, 概率方法的好处是直观, 容易理解, 也容易推导.

在他们原始的论文中, 随机分析的应用还没有到这个程度, Black 与 Scholes 研究期权在 $t < T$ 时刻的价值, 设它是时间 t 和状态 S_t 的函数 $V(t, S_t)$, 这样推出函数 $V(t, x)$ 满足一个具有终端条件的 PDE, 然后再把方程转化为初值问题的热方程, 而热方程的解就是 Brown 运动的转移密度.

第十三章　补充I: Brown 运动构造

在前面读者已经看到Brown 运动是一个神奇的随机过程, 无疑可以成为最神奇的随机过程, 有各种神奇的性质, 它有连续性, 有马氏性, 有鞅性, 有二次变差, 可以定义随机积分, 有鞅表示性, 它太神奇以至于初学者可能会怀疑它是否真的存在. 通常的构造方法比较麻烦, 要先讲述一般的测度构造定理, 然后证明Kolmogorov 的相容性定理, 就是利用有限维分布族在无穷维空间上构造所需的概率测度, 最后证明连续修正的存在性, 此方法繁复, 但有一般性. 但这个工程实在太庞大, 无法在这里细述. 好在Brown 运动的存在性有很多证明, 我们在这里讲述一个用Fourier 级数的纯粹分析性的构造方法, 也是Wiener 给出的, 缺点是它没有一般性, 除了Brown 运动, 不能用于其它过程的构造.

一个随机过程$X = (X_t)$ 称为是Gauss 过程, 如果对任何$0 < t_1 < \cdots < t_n$, $(X(t_1), \cdots, X(t_n))$ 服从正态分布. 它服从正态分布的充分必要条件是它的任何线性组合$k_1X(t_1) + \cdots + k_nX(t_n)$ 是正态分布的.

定理 13.0.2　一个连续随机过程$B = (B_t)$ 是标准Brown 运动的充分必要条件是它是一个Gauss 过程且

$$\mathbb{E}[B_tB_s] = t \wedge s.$$

看Brown 运动的定义主要有三点, 一是连续性, 条件里面已经有了; 二是独立增量性; 三是增量是服从期望为零(中心化)方差等于时间差的正态分布. 假设(B_t) 是Gauss 过程且$\mathbb{E}[B_tB_s] = t \wedge s$. 由多维正态分布性质, 对任何$0 < t_1 < \cdots < t_n$,

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

也是正态分布, 我们只需要验证他们的协方差是零. 事实上, 因为$\mathbb{E}[B_tB_s]=t\wedge s$, 所以

$$\mathbb{E}[B(t_i)-B(t_{i-1})]^2=t-2t_{i-1}+t_{i-1}=t_i-t_{i-1},$$

另外当$i<j$ 时,

$$\mathbb{E}[(B(t_j)-B(t_{j-1}))(B(t_i)-B(t_{i-1}))]=t_i-t_i-t_{i-1}-t_{i-1}=0.$$

因此, B 有独立增量性质且质量服从所需的正态分布.

让我们先来构造一个Gauss 过程$X=(X_t)$ 满足$\mathbb{E}[X_tX_s]=t\wedge s$. 一般地, 构造一个概率空间$(\Omega,\mathscr{F},\mathbb{P})$ 和独立且服从标准正态分布的随机序列$\{\xi_n:n\geq 0\}$. 再任取Hilbert 空间$(H,\langle\cdot,\cdot\rangle_H)$ 以及它的一个标准正交基$\{e_n:n\geq 1\}$, 对$h\in H$, 它有唯一的表示

$$h=\sum_{n\geq 0}\langle h,e_n\rangle_H e_n.$$

定义

$$X(h):=\sum_{n\geq 0}\langle h,e_n\rangle_H\xi_n,$$

那么$X(h)$ 是一个正态分布的随机变量且因为$\{\xi_n\}$ 在$L^2(\Omega)$ 上也是一个标准正交序列, 所以

$$\mathbb{E}[X(h_1)X(h_2)]=\langle h_1,h_2\rangle_H,\ h_1,h_2\in H,$$

它与具体的基的取法无关, 且说明X 是H 到$L^2(\Omega)$ 中的一个保内积不变的线性变换.

把H 作为指标集, $\{X(h):h\in H\}$ 称为一个随机场. 它实际上是一个Gauss 随机场, 即对于$h_1,\cdots,h_n\in H$,

$$(X(h_1),\cdots,X(h_n))$$

是联合Gauss (正态)分布的. 现在我们来找一个具体的Hilbert 空间H. 令$H = L^2[0,\pi]$,

$$e_0 = \frac{1}{\sqrt{\pi}},\ e_n(x) = \sqrt{\frac{2}{\pi}}\cos nx,\ x \in [0,\pi],\ n \geq 1.$$

那么$\{e_n : n \geq 0\}$ 是H 的一个标准正交基. 对任何$t \in [0,\pi]$, 定义

$$\begin{aligned} X_t := X(1_{[0,t]}) &= \sum_{n\geq 0} \langle 1_{[0,t]}, e_n\rangle_H \xi_n \\ &= \frac{t}{sqrt\pi}\xi_0 + \sum_{n\geq 1}\sqrt{\frac{2}{\pi}}\frac{\sin nt}{n}\xi_n. \end{aligned}$$

显然$\{X_t : t \in [0,\pi]\}$ 是一个Gauss 过程且

$$\mathbb{E}[X_t X_t] = \mathbb{E}[X(1_{[0,t]})X(1_{[0,s]})] = \langle 1_{[0,t]}, 1_{[0,s]}\rangle_H = t \wedge s.$$

现在就差证明(X_t) 的连续性, 这也是最困难最奇妙的地方.

上面的X_t 表示为一个级数, 这个级数是L^2 收敛的, 不是几乎处处的. 这个级数的部分和关于t 是连续函数, 所以如果我们能够证明这个部分和关于t 几乎肯定是一致收敛, 即

$$\sup_{t\in[0,\pi]} |X_t - \sum_{k=0}^{n} \langle 1_{[0,\pi]}, e_k\rangle_H \xi_k|$$

几乎处处收敛于0, 那么$t \mapsto X_t$ 的连续性就被证明了.

但实际上, 证明原级数几乎一致收敛是不可能的, 我们必须把级数重组为

$$\frac{t}{\sqrt{\pi}}\xi_0 + \sum_{n\geq 1}\sum_{k=2^{n-1}}^{2^n-1}\sqrt{\frac{2}{\pi}}\frac{\sin kt}{k}\xi_k,$$

证明它是几乎一致收敛的.

证明方法参考 [4] 的§1.5 中的Problem 3. 令

$$s_{m,n}(t) := \sum_{m}^{n-1}\frac{\sin kt}{k}\xi_k,\ t_{m,n} := \max_{0\leq t\leq\pi}|s_{m,n}(t)|;$$

那么

$$
\begin{aligned}
|s_{m,n}(t)|^2 &\le \max_t \left|\sum_m^{n-1} \frac{e^{ikt}}{k}\xi_k\right| \\
&= \sum_m^{n-1} \frac{\xi_k^2}{k^2} + 2\sum_{k=m}^{n-1}\sum_{j=k+1}^{n-1} \frac{e^{ijt-ikt}}{kj}\xi_k\xi_j \\
&= \sum_m^{n-1} \frac{\xi_k^2}{k^2} + 2\sum_{k=m}^{n-1}\sum_{j=1}^{n-k-1} \frac{e^{ijt}}{k(j+k)}\xi_k\xi_{j+k} \\
&= \sum_m^{n-1} \frac{\xi_k^2}{k^2} + 2\sum_{j=1}^{n-m-1} e^{ijt}\sum_{k=m}^{n-j-1} \frac{\xi_k\xi_{j+k}}{k(k+j)},
\end{aligned}
$$

因此

$$
\begin{aligned}
t_{m,n}^2 &= \max_t |s_{m,n}(t)| \\
&\le \sum_m^{n-1} \frac{\xi_k^2}{k^2} + 2\sum_{j=1}^{n-m-1} \left|\sum_{k=m}^{n-j-1} \frac{\xi_k\xi_{j+k}}{k(k+j)}\right|.
\end{aligned}
$$

取期望得

$$
\begin{aligned}
\mathbb{E}[t_{m,n}^2] &\le \sum_m^{n-1} \frac{1}{k^2} + 2\sum_{j=1}^{n-m-1} \mathbb{E}\left|\sum_{k=m}^{n-j-1} \frac{\xi_k\xi_{j+k}}{k(k+j)}\right| \\
&\le \sum_m^{n-1} \frac{1}{k^2} + 2\sum_{j=1}^{n-m-1} \left[\mathbb{E}\left|\sum_{k=m}^{n-j-1} \frac{\xi_k\xi_{j+k}}{k(k+j)}\right|^2\right]^{1/2} \\
&= \sum_m^{n-1} \frac{1}{k^2} + 2\sum_{j=1}^{n-m-1} \left[\sum_{k=m}^{n-j-1} \frac{1}{k^2(k+j)^2}\right]^{1/2} \\
&\le \frac{n-m}{m^2} + 2(n-m)\left(\frac{n-m}{m^4}\right)^{1/2} \\
&\le \frac{3(n-m)^{3/2}}{m^2}.
\end{aligned}
$$

因此

$$
\mathbb{E}\left[\sum_n t_{2^{n-1},2^n}\right] \le \sum_n [\mathbb{E}t_{2^{n-1},2^n}^2]^{1/2} \le \sum_n \frac{3}{2^{(n-1)/2}} < \infty.
$$

习题

1. 证明: $(X(t_1),\cdots,X(t_n))$ 服从正态分布的充分必要条件是它的任何线性组合$k_1X(t_1)+\cdots+k_nX(t_n)$ 是正态分布的. 提示: 利用特征函数方法.

2. 证明: 对于$h_1,\cdots,h_n\in H$,

$$(X(h_1),\cdots,X(h_n))$$

是联合Gauss (正态)分布的.

参考文献

[1] Chung, K.L., Williams, R.J., Introduction to Stochastic Integration, Birkhauser Boston, Inc., 1983

[2] Doob, J.L., Stochastic Processes, Wiley, New York, 1953

[3] Ikeda, N., Watanabe, S., Stochastic Differential Equations and Diffusion Processes, North-Holland, 1981

[4] Itô, K., Mckean Jr., H. P., Diffusion Processes and Their Sample paths, Sringer-Verlag, 1965

[5] Revuz, D., Yor, M., Continuous Martingales and Brownian Motion, Springer, 1991

[6] Wiener, N., Differential space", J. Math. Phys. 2, 132-174 (1924)

[7] Yosida, K., Functional Analysis, Springer-Verlag, 1980

[8] 汪嘉冈, 现代概率论基础, 复旦大学出版社, 1988

[9] 何声武, 汪嘉冈, 严加安, 半鞅与随机分析, 科学出版社, 1995

[10] 钱忠民, 应坚刚, 随机分析引论, 未出版讲义, 2016

[11] 严加安, 金融数学引论, 科学出版社, 2012

[12] 应坚刚, 金蒙伟, 随机过程基础, 复旦大学出版社, 2005