目录

第 1 章　预备知识

在正式介绍随机过程内容之前, 我们首先给出学习本书需要的基础知识.

1.1　概率与测度

本书讨论的对象主要是随机过程, 以及 20 世纪 40 年代以来建立的随机积分理论. 因此, 我们需要对随机过程建立严格的数学定义. 即需要在概率公理化的基础上给出随机过程需要的数学基础.

首先我们给出可测空间的定义.

设 Ω 是一个非空集合. 考虑 Ω 上的非空子集类.

定义 1.1.1　对于 Ω 上的非空子集类 $\mathcal{F}$, 若

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

(2) $A \in \mathcal{F},\ B \in \mathcal{F}$, 有 $A - B \in \mathcal{F}$,

(3) $A_i \in \mathcal{F},\ i = 1, 2, \cdots$, 有 $\bigcup\limits_{i=1}^{\infty} A_i \in \mathcal{F}$,

称 $\mathcal{F}$ 是 Ω 上的 $\sigma-$ 代数 ($\sigma-$ 域), 称 $(\Omega, \mathcal{F})$ 为可测空间.

事实上, 我们在实变函数课程中已经接触到如下可测空间;

例 1.1.1　令 Ω 为实数集 $\mathbb{R}$, $\mathbb{R}$ 上 Borel 集全体记为 $\mathcal{B}$, 则 $(\mathbb{R}, \mathcal{B})$ 是一个可测空间.

例 1.1.2　令 $\Omega = \mathbb{R}$, $\mathbb{R}$ 上 Lebesgue 可测集全体记为 $\mathcal{L}$, 则 $(\mathbb{R}, \mathcal{L})$ 是一个可测空间.

若 $\mathcal{A}$ 是 Ω 的一个子集类, $\sigma(\mathcal{A})$ 即为包含 $\mathcal{A}$ 的最小的 $\sigma-$ 代数, 称为 $\mathcal{A}$ 生成的 $\sigma-$ 代数.

若 Ω 是非空集合, ξ 是 Ω 到 $\mathbb{R}$ 的一个映射. 对于 $A \subset \mathbb{R}$, 定义

$$\xi^{-1}(A) = \{\omega \in \Omega : \xi(\omega) \in A\}.$$

若 $\mathcal{A}$ 是 $\mathbb{R}$ 的一个子集类, 令

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

定义 1.1.2　若 ξ 是可测空间 $(\Omega,\mathcal{F})$ 到 $(\mathbb{R},\mathcal{B})$ 的一个映射, 且 $\xi^{-1}(\mathcal{B})\subset\mathcal{F}$, 则称 ξ 是 $\mathcal{F}-$ 可测随机变量, 或简称随机变量.

在概率论中, 随机变量是最基本的讨论对象. 在可测空间 $(\Omega,\mathcal{F})$ 上, 若 $A\in\mathcal{F}$, 则称

$$\mathbb{1}_A(\omega)=\begin{cases}1, & \omega\in A,\\ 0, & \omega\notin A\end{cases}$$

为示性随机变量, 有限个示性函数的线性组合称为简单随机变量. 事实上, 随机变量在一定条件下可以由一列简单随机变量逼近. 例如, 对于非负随机变量 ξ, 可通过简单随机变量

$$\psi_n=\sum_{i=0}^{n2^n-1}\frac{i}{2^n}\mathbb{1}_{\{\frac{i}{2^n}\leqslant\xi<\frac{i+1}{2^n}\}}+n\mathbb{1}_{\{\xi\geqslant n\}}$$

逼近.

概率论作为一门数学分支, 研究可测空间上随机变量取值的可能性大小是其主要任务之一. 因此, 需要引入一个“量”, 来刻画随机变量取值的可能性大小. 这个“量”就是概率. 下面给出概率的公理化定义.

定义 1.1.3　在可测空间 $(\Omega,\mathcal{F})$ 上, P 是 $\mathcal{F}$ 上的一个集函数, 若 P 满足:

(1) 对于 $A\in\mathcal{F}$, $\mathsf{P}(A)\geqslant 0$,

(2) 对于 $\{A_n\}_{n\geqslant 1}\subset\mathcal{F}$, $A_n\cap A_m=\emptyset$, 有 $\mathsf{P}(\bigcup\limits_{n=1}^{\infty}A_n)=\sum\limits_{n=1}^{\infty}\mathsf{P}(A_n)$,

(3) $\mathsf{P}(\Omega)=1$,

则称 P 是 $(\Omega,\mathcal{F})$ 上的概率测度. $(\Omega,\mathcal{F},\mathsf{P})$ 称为概率空间.

定义 1.1.4　若 μ 是 $(\Omega,\mathcal{F})$ 上的关于 $\mathcal{F}$ 中集合的集函数, 满足定义 1.1.3 中的前 2 条, 则称 μ 是 $(\Omega,\mathcal{F})$ 上的测度. $(\Omega,\mathcal{F},\mu)$ 称为测度空间.

事实上, 如下例子就是常见的一个概率空间.

例 1.1.3　$([0,1],\mathcal{B}[0,1],m)$ 是一个概率空间, $\mathcal{B}[0,1]$ 是 $[0,1]$ 上的 Borel 集全体, m 是 Lebesgue 测度.

对于随机变量而言, 我们通过如下形式, 刻画随机变量取值的可能性大小.

定义 1.1.5　设 $(\Omega,\mathcal{F},\mathsf{P})$ 是一个概率空间, ξ 是 $(\Omega,\mathcal{F})$ 上的随机变量, 对于 $B\in\mathcal{B}$, 定义

$$\mathsf{P}\circ\xi^{-1}(B)=\mathsf{P}(\xi^{-1}(B))=\mathsf{P}(\{\omega:\xi(\omega)\in B\}).$$

此时, $\mathsf{P}\circ\xi^{-1}$ 是 $(\mathbb{R},\mathcal{B})$ 上的概率测度, 称 $\mathsf{P}\circ\xi^{-1}$ 为 ξ 的分布.

回忆之前学过的分布函数的定义: $\forall x\in\mathbb{R}$,

$$F(x)=\mathsf{P}(\xi\leqslant x)=\mathsf{P}(\{\omega:\xi(\omega)\in(-\infty,x]\}).$$

我们知道分布函数是实轴上的右连续的增函数. 对于实数 $a<b$, 定义

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

将上述集函数延拓至整个实轴上的 Borel 集上, 生成的 Lebesgue-Stieltjes 测度 μ 实际上与 $\mathsf{P}\circ\xi^{-1}$ 是一样的. 所以, 有时候我们会直接把分布函数对应为它生成的 Lebesgue-Stieltjes 测度, 直接称分布函数为分布. 常见的分布有以下几种:

例 1.1.4 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量 ξ, 若其分布函数为

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{x}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt,$$

这里 $-\infty<\mu<\infty$, $0<\sigma<\infty$, 则通常称 ξ 服从正态分布. 特别地, 若 $\mu=0$, $\sigma^2=1$, 则称 ξ 服从标准正态分布.

例 1.1.5 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量 ξ, 若其分布函数为

$$F(x)=\int_{-\infty}^{x}p(x)dx,$$

这里

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & x\in[a,b],\\ 0, & \text{其它},\end{cases}$$

且 $-\infty<a<b<\infty$, 则称 ξ 服从 $[a,b]$ 上的均匀分布.

例 1.1.6 设 ξ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量, 若其分布函数为

$$F(x)=\int_{-\infty}^{x}p(x)dx,$$

这里

$$p(x)=\begin{cases}\lambda e^{-\lambda x}, & x\geqslant 0,\\ 0, & x<0,\end{cases}$$

且 $\lambda>0$, 则称 ξ 服从指数分布.

例 1.1.7 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量 ξ, 对于非负整数 k, 若

$$\mathsf{P}(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda},$$

这里 $\lambda>0$, 则称 ξ 服从 Possion 分布.

读者已经在概率论中接触到很多分布, 这里不再赘述.

独立性是概率论中特有的一个概念. 设 $(\Omega, \mathcal{F}, \mathsf{P})$ 是概率空间, ξ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量. 令 $\sigma(\xi) = \xi^{-1}(\mathcal{B})$, $\mathcal{B}$ 为 $\mathbb{R}$ 上的 Borel 集类. $\sigma(\xi)$ 即为使 ξ 成为随机变量的最小 $\sigma-$ 代数, 称 $\sigma(\xi)$ 为由 ξ 生成的 $\sigma-$ 代数. 本质上, $\sigma(\xi)$ 是使 ξ 可测的最小 $\sigma-$ 代数.

设 $\{\xi_i, i \in I\}$ 是一族随机变量, 若 $\sigma(\xi_i, i \in I)$ 是使 $\{\xi_i : i \in I\}$ 都可测的最小 $\sigma-$ 代数, 则称其为 $\{\xi_i, i \in I\}$ 生成的 $\sigma-$ 代数.

定义 1.1.6 设 ξ 与 η 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量, $\sigma(\xi)$ 与 $\sigma(\eta)$ 是 ξ, η 生成的 $\sigma-$ 代数. 若对于 $A \in \sigma(\xi)$, $B \in \sigma(\eta)$, 均有 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$, 则称 ξ 与 η 独立.

在概率论中, 数学期望扮演着十分重要的角色. 我们利用测度论的语言, 可以有如下定义.

定义 1.1.7 在 $(\Omega, \mathcal{F}, \mathsf{P})$ 上, ξ 是非负随机变量, η 是非负简单随机变量, 设

$$\eta = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}, \quad A_i \in \mathcal{F}.$$

令 $\mathsf{E}[\eta] = \sum_{i=1}^{n} a_i \mathsf{P}(A_i)$, 定义

$$\mathsf{E}[\xi] = \sup\{\ \mathsf{E}[\eta] : 0 \leqslant \eta \leqslant \xi, \eta \text{是简单随机变量}\}.$$

若 ξ 是随机变量, $\xi = \xi^+ - \xi^-$, $\xi^+ = \max\{\xi, 0\}$, $\xi^- = \max\{-\xi, 0\}$, 定义 $\mathsf{E}[\xi] = \mathsf{E}[\xi^+] - \mathsf{E}[\xi^-]$. 若 $\mathsf{E}[\xi^+] < +\infty$, $\mathsf{E}[\xi^-] < +\infty$, 我们称 ξ 可积. 称 $\mathsf{E}[\xi]$ 是 ξ 的数学期望.

有时, 为强调概率测度 P, 会记数学期望为 $\mathsf{E}_{\mathsf{P}}[\xi]$.

关于数学期望的性质的讨论, 本质上与实变函数课程中 Lebesgue 可积函数的讨论是一样的. 读者可以假设概率空间为 $([0,1], \mathcal{B}[0,1], m)$, 然后按照实变函数课程中 Lebesgue 可积函数相关的讨论, 可以得到一系列关于数学期望的性质, 这里不再重复. 在后面遇到相关定理的证明时, 如果有必要, 会给予详细介绍. 事实上, 类似于实变函数中的讨论, 我们还可以得到如下的 Fubini 定理. 这里, 我们把证明略去.

对于集合 A, B, $A \times B = \{(a, b) : a \in A,\ b \in B\}$; 对于 $\sigma-$ 代数 $\mathcal{F}_1$, $\mathcal{F}_2$, $\mathcal{F}_1 \otimes \mathcal{F}_2$ 定义为包含所有形如 $A \times B$ 的最小 $\sigma-$ 代数, 其中 $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$.

定理 1.1.1 设 $(\Omega_1, \mathcal{F}_1, \mathsf{P})$, $(\Omega_2, \mathcal{F}_2, \mathsf{Q})$ 是两个概率空间, f 是 $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes$

$\mathcal{F}_2)$ 上的可测函数, 若 f 是非负或可积的, 则

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mathsf{P}\times\mathsf{Q} &= \int_{\Omega_1}\mathsf{P}(\,\mathrm{d}\omega_1)\int_{\Omega_2} f(\omega_1,\omega_2)\mathsf{Q}(\,\mathrm{d}\omega_2)\\ &= \int_{\Omega_2}\mathsf{Q}(\,\mathrm{d}\omega_2)\int_{\Omega_1} f(\omega_1,\omega_2)\mathsf{P}(\,\mathrm{d}\omega_1).\end{aligned}$$

在可测空间 $(\Omega,\mathcal{F})$ 上, 考虑两个概率测度 P,Q, 对于任意的 $A\in\mathcal{F}$, 若 $\mathsf{P}(A)=0$ 能推出 $\mathsf{Q}(A)=0$, 则称 Q 关于 P 绝对连续, 记为 $\mathsf{Q}\ll\mathsf{P}$. 下面, 我们给出十分重要的 Radon-Nikodym 定理. 其证明我们在附录中给出.

定理 1.1.2 (Radon-Nikodym 定理) 设 P 是 $(\Omega,\mathcal{F})$ 上的概率测度, Q 是 $(\Omega,\mathcal{F})$ 上的有限测度, 若 $\mathsf{Q}\ll\mathsf{P}$, 则存在唯一的 (几乎处处意义下) 可积随机变量 ξ, 使得对于任意 $A\in\mathcal{F}$,

$$\mathsf{Q}(A)=\int_A\xi\,\mathrm{d}\mathsf{P}=\mathsf{E}_{\mathsf{P}}[\xi\mathbb{1}_A].$$

上式中的 ξ 一般称为测度 Q 关于测度 P 的 Radon-Nikodym 导数. 我们在概率论中提到的连续型随机变量, 即其分布函数生成的 Lebesgue-Stieltjes 测度关于 Lebesgue 测度绝对连续. 所谓的密度函数, 实际上是分布函数生成的 Lebesgue-Stieltjes 测度关于 Lebesgue 测度的 Radon-Nikodym 导数.

证明: 为证明方便，我们假设 Q 也是概率测度. 令

$$\mathsf{P}^*=\frac{\mathsf{P}+\mathsf{Q}}{2}.$$

显然, 有 $\mathsf{Q}\ll\mathsf{P}^*$. 令 H 是满足 $\int X^2d\mathsf{P}^*<\infty$ 的随机变量全体, 并记 $(X,Y)=\int XYd\mathsf{P}^*$, 易得 (X,Y) 是 H 上的内积, 并容易证明 $(H,(\cdot,\cdot))$ 是一个 Hilbert 空间, 其上的范数为 $||X||_2=(\int X^2d\mathsf{P}^*)^{1/2}$. 令

$$L(Y)=\int Yd\mathsf{Q}.$$

易有 $L(\cdot)$ 是 $(H,(\cdot,\cdot))$ 上的线性泛函, 且注意到

$$\begin{aligned}|L(Y)| &\leqslant \int|Y|d\mathsf{P}+\int|Y|d\mathsf{Q}\\ &\leqslant 2\int|Y|d\mathsf{P}^*\leqslant 2(\int Y^2d\mathsf{P}^*)^{1/2}.\end{aligned}$$

因此, $L(\cdot)$ 是 $(H,(\cdot,\cdot))$ 上的有界线性泛函. 由 Riesz 表示定理，存在 $Z\in H$ 使得

$$L(Y)=(Y,Z)=\int YZd\mathsf{P}^*.$$

因此
$$\int Y d\mathsf{Q} = \int YZ d\mathsf{P}^*.$$
对于 $A \in \mathcal{F}$, 取 $Y = \mathbb{1}_A$, 因此
$$\mathsf{Q}(A) = \int_A Z d\mathsf{P}^*.$$
进一步有
$$0 \leqslant \frac{\mathsf{Q}(A)}{\mathsf{P}^*(A)} \leqslant 2.$$
由 A 的任意性, 我们得到 $0 \leqslant Z \leqslant 2$. 同时, 注意到
$$\begin{aligned}\int Y d\mathsf{Q} &= \int YZ d\mathsf{P}^* \\ &= \frac{1}{2}(\int YZ d\mathsf{P} + \int YZ d\mathsf{Q}).\end{aligned}$$
故
$$\int Y(1-\frac{Z}{2}) d\mathsf{Q} = \int \frac{YZ}{2} d\mathsf{P}.$$
取 $Y = \mathbb{1}_{\{Z=2\}}$, 我们有 $\mathsf{P}(Z=2)=0$. 因此
$$0 \leqslant Z < 2, a.s..$$
对于 $A \in \mathcal{F}$，取 $Y = (\frac{Z}{2})^n \mathbb{1}_A$, 有
$$\int_A (\frac{Z}{2})^n - (\frac{Z}{2})^{n+1} d\mathsf{Q} = \int_A (\frac{Z}{2})^{n+1} d\mathsf{P}.$$
将上式从 $n=0$ 加到 $n=N$，有
$$\int_A 1-(\frac{Z}{2})^{N+1} d\mathsf{Q} = \int_A \frac{Z}{2}\sum_{j=0}^{N}(\frac{Z}{2})^j d\mathsf{P}.$$
两边令 $N \to \infty$, 由单调收敛定理, 我们得到
$$\mathsf{Q}(A) = \int_A \frac{Z}{2-Z} d\mathsf{P}.$$
取 $\xi = \frac{Z}{2-Z}$, 唯一性的证明留给读者.

■

1.2 随机变量的收敛性

在这一节中, 我们介绍随机变量的收敛性. 这是我们在后面讨论中需要的. 首先, 我们给依概率收敛的定义.

定义 1.2.1 设 $\{\xi_n\}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的一列随机变量, ξ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量. 若对任意的 $\epsilon > 0$, $\lim\limits_{n\to\infty} \mathsf{P}(\{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}) = 0$. 称 $\{\xi_n\}$ 依概率收敛于 ξ, 记为 $\xi_n \xrightarrow{\mathsf{P}} \xi$.

除了依概率收敛, 几乎处处收敛也是一种十分重要的收敛. 下面, 我们给出几乎处处收敛的定义.

定义 1.2.2 设 $\{\xi_n\}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的一列随机变量, ξ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量. 若 $\mathsf{P}(\{\omega : \lim\limits_{n\to\infty} \xi_n(\omega) = \xi(\omega)\}) = 1$, 称 $\{\xi_n\}$ 几乎处处收敛于 ξ, 记为 $\xi_n \xrightarrow{a.s.} \xi$.

注意到, 如果对于任意的 $\epsilon > 0$, 有

$$\sum_{n=1}^{\infty} \mathsf{P}(\{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}) < \infty,$$

则会有 $\{\xi_n\}$ 几乎处处收敛于 ξ. 事实上, 考虑集合

$$\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}$$

的测度, 会有

$$\mathsf{P}(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}) \leqslant \sum_{n=k}^{\infty} \mathsf{P}(\{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}).$$

由级数收敛的性质,

$$\lim_{k\to\infty} \sum_{n=k}^{\infty} \mathsf{P}(\{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}) = 0,$$

故

$$\mathsf{P}(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{\omega : |\xi_n(\omega) - \xi(\omega)| \geqslant \epsilon\}) = 0.$$

因此 $\{\xi_n\}$ 几乎处处收敛于 ξ. 进一步, 我们可以通过几乎处处收敛刻画依概率收敛.

命题 1.2.1　设 $\{\xi_n\}$ 是概率空间 $(\Omega,\mathcal{F},\mathsf{P})$ 上的一列随机变量, ξ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量. $\{\xi_n\}$ 依概率收敛于 ξ 当且仅当对于 $\{\xi_n\}$ 的任意子列 $\{\xi_{n'}\}$, 存在子列 $\{\xi_{n'_k}\}$ 几乎处处收敛至 ξ.

证明: 命题的必要性由著名的 Riesz 定理可以得到. 这里主要证明充分性.

使用反证法, 如果 $\{\xi_n\}$ 不是依概率收敛于 ξ, 则存在 ϵ_0, 存在 δ_0, 对于任意的 k, 存在 n_k, 使得

$$\mathsf{P}(\{\omega : |\xi_{n_k}(\omega)-\xi(\omega)|\geqslant\epsilon_0\})>\delta_0.$$

显然, 此时子列 ξ_{n_k} 并不是几乎处处收敛于 ξ 的. 产生矛盾. ■

我们在概率论课程中涉及到的弱大数律和强大数律是关于依概率收敛和几乎处处收敛的, 而中心极限定理是关于依分布收敛展开讨论的. 下面给出依分布收敛的定义.

定义 1.2.3　设 $\{\xi_n\}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的一列随机变量, ξ_n 的分布函数 $F_n(x)$, ξ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量, ξ 的分布函数 $F(x)$. 对于 $F(x)$ 的任意连续点 u,

$$\lim_{n\to\infty}F_n(u)=F(u),$$

则称 ξ_n 依分布收敛于 ξ, 记为 $\xi_n\xrightarrow{d}\xi$.

依分布收敛本质上是一种弱收敛. 下面给出弱收敛的定义.

定义 1.2.4　设 $\{\xi_n\}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的一列随机变量, ξ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机变量. 若对任意有界连续函数 f,

$$\lim_{n\to\infty}\mathsf{E}[f(\xi_n)]=\mathsf{E}[f(\xi)],$$

则称 ξ_n 弱收敛于 ξ, 记为 $\xi_n\Longrightarrow\xi$.

随机变量列弱收敛于一个随机变量, 也就是我们通常所说的依分布收敛. 但是弱收敛的刻画形式更宽泛, 这里需要指出的是, 弱收敛与依分布收敛是等价的.

对于 $(\Omega,\mathcal{F},\mathsf{P})$ 上的一个随机变量 ξ, 若 ξ 可积, 不难证明, 对于任意的 $\epsilon>0$, 存在 $N>0$, 使得

$$\int_{\{|\xi|\geqslant N\}}\xi\,\mathrm{d}\mathsf{P}<\epsilon.$$

若对一族随机变量讨论类似于上式的性质, 可以引入一致可积的概念.

定义 1.2.5　设 I 是一个指标集, $(\Omega,\mathcal{F},\mathsf{P})$ 上有一个可积随机变量族 $\{\xi_\alpha,\alpha\in I\}$, 若对于任意的 $\epsilon>0$, 存在 $N>0$, 使得对于任意的 $\alpha\in I$,

$$\mathsf{E}[|\xi_\alpha|\mathbb{1}_{\{|\xi_\alpha|\geqslant N\}}]=\int_{\{|\xi_\alpha|\geqslant N\}}|\xi_\alpha|\mathrm{d}\mathsf{P}<\epsilon,$$

则称 $\{\xi_\alpha, \alpha \in I\}$ 为一致可积的.

定理 1.2.1 设 $\{\xi_\alpha, \alpha \in I\}$ 是可积随机变量族, 则 $\{\xi_\alpha, \alpha \in I\}$ 是一致可积的充要条件是

(i) 对于任意的 $\epsilon > 0$, 存在 $\delta > 0$, 使得当 $A \in \mathcal{F}$, $\mathsf{P}(A) < \delta$ 时, 对于任意的 $\alpha \in I$, $\mathsf{E}[|\xi_\alpha| \mathbb{1}_A] < \epsilon$ (一致绝对连续性);

(ii) $\sup\limits_{\alpha \in I} \mathsf{E}[|\xi_\alpha|] < +\infty$.

证明: 必要性: 对于任意的 $A \in \mathcal{F}$, $N > 0$,

$$\begin{aligned}\int_A |\xi_\alpha| \,\mathrm{d}\mathsf{P} &= \int_{A \cap \{|\xi_\alpha| \geqslant N\}} |\xi_\alpha| \,\mathrm{d}\mathsf{P} + \int_{A \cap \{|\xi_\alpha| < N\}} |\xi_\alpha| \,\mathrm{d}\mathsf{P} \\ &\leqslant \int_{\{|\xi_\alpha| \geqslant N\}} |\xi_\alpha| \,\mathrm{d}\mathsf{P} + N\mathsf{P}(A).\end{aligned}$$

由一致可积性可有一致绝对连续性. 取 $A = \Omega$, 有 $\sup\limits_{\alpha \in I} \mathsf{E}[|\xi_\alpha|] < +\infty$.

充分性:

$$\mathsf{P}(|\xi_\alpha| \geqslant N) \leqslant \frac{1}{N} \mathsf{E}[|\xi_\alpha|],$$

由于 $\sup\limits_{\alpha \in I} \mathsf{E}[|\xi_\alpha|] < +\infty$, 故由一致绝对连续性可得一致可积性. ■

一族随机变量如果一致可积, 再加上依概率收敛, 就会得到如下结果.

定理 1.2.2 设 $(\Omega, \mathcal{F}, \mathsf{P})$ 是概率空间, $\{\xi_n\}$ 是一列一致可积的随机变量列, 若存在一个可积随机变量 ξ, 使 $\xi_n \xrightarrow{\mathsf{P}} \xi$, 则

$$\lim_{n \to \infty} \mathsf{E}[|\xi_n - \xi|] = 0.$$

证明: 由于 $\{\xi_n\}$ 是一致可积的, ξ 是可积的, 故对于任意的 $\epsilon > 0$, 存在一个 $\delta > 0$, 使得对于任意的 $n \geqslant 0$, 当 $A \in \mathcal{F}$, $\mathsf{P}(A) < \delta$ 时

$$\int_A |\xi_n| \,\mathrm{d}\mathsf{P} < \epsilon, \qquad \int_A |\xi| \,\mathrm{d}\mathsf{P} < \epsilon.$$

由于 $\xi_n \xrightarrow{\mathsf{P}} \xi$, 故存在 $K > 0$, 使得 $n \geqslant K$ 时,

$$\mathsf{P}(|\xi_n - \xi| > \varepsilon) < \delta.$$

因此, 当 $n \geqslant K$ 时,

$$\begin{aligned}\int |\xi_n - \xi| \,\mathrm{d}\mathsf{P} &= \int_{\{|\xi_n - \xi| > \varepsilon\}} |\xi_n - \xi| \,\mathrm{d}\mathsf{P} + \int_{\{|\xi_n - \xi| \leqslant \varepsilon\}} |\xi_n - \xi| \,\mathrm{d}\mathsf{P} \\ &\leqslant \int_{\{|\xi_n - \xi| > \varepsilon\}} |\xi_n| \,\mathrm{d}\mathsf{P} + \int_{\{|\xi_n - \xi| > \varepsilon\}} |\xi| \,\mathrm{d}\mathsf{P} + \varepsilon \\ &\leqslant 3\varepsilon.\end{aligned}$$

故

$$\lim_{n\to\infty} \mathsf{E}[|\xi_n - \xi|] = 0.$$

■

今后, 为方便叙述, 对于概率空间 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量列 $\{\xi_n\}$, 对于 $p \geqslant 1$, 若存在随机变量 ξ,

$$\lim_{n\to\infty} \mathsf{E}[|\xi_n - \xi|^p] = 0,$$

则称 ξ_n L^p 收敛于 ξ.

1.3　条件数学期望

考虑 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量 X, 数学期望 $\mathsf{E}[X]$ 是对 X 的最佳预测. 如果考虑更加复杂的问题: $X_1, X_2, \cdots$ 是一列时间序列, 在时间 n, X_n 代表 n 时刻某支股票的价格, Y 代表另一支股票价格. 用 $\mathsf{E}(Y \mid X_1, X_2, \cdots, X_n)$ 代表给定 $X_1, X_2, \cdots, X_n$ 后对 Y 的最佳预测. 加一定的条件后, 我们再来考虑这个最佳预测问题. 令 $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$, 我们暂时使用符号 $\mathsf{E}[Y \mid \mathcal{F}_n]$ 来代表最佳预测结果 $\mathsf{E}[Y \mid X_1, X_2, \cdots, X_n]$. (下文将正式给出 $\mathsf{E}[Y \mid \mathcal{F}_n]$ 的定义). 注意到 $\mathsf{E}[Y \mid \mathcal{F}_n]$ 肯定与 $X_1, X_2, \cdots, X_n$ 有关, 因此可写为

$$\mathsf{E}[Y \mid \mathcal{F}_n] = \phi(X_1, X_2, \cdots, X_n).$$

事实上, 在特定的情况下, 我们可以进行如下讨论.

设 (X, Y) 具有联合密度 $f(x, y)$, $-\infty < x, y < \infty$, 边际密度为

$$f(x) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}y,\ g(y) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}x.$$

所谓条件密度为 $f(y \mid x) = \dfrac{f(x, y)}{f(x)}$, 我们设定

$$\mathsf{E}[Y \mid X = x] = \int_{-\infty}^{\infty} y f(y \mid x)\,\mathrm{d}y.$$

不严格地写, 在非常强的条件下, 会有

$$\mathsf{E}[Y \mid X] = \int_{-\infty}^{\infty} y f(y \mid X)\,\mathrm{d}y = \frac{\int_{-\infty}^{\infty} y f(X, y)\,\mathrm{d}y}{f(X)} = \phi(X).$$

进一步, 有

$$
\begin{aligned}
\mathsf{E}[\mathsf{E}[Y \mid X]] &= \int_{-\infty}^{\infty} \mathsf{E}[Y \mid X = x] f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} y f(y \mid x)\,\mathrm{d}y] f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y)\,\mathrm{d}y\,\mathrm{d}x = \mathsf{E}[Y].
\end{aligned}
$$

事实上, 如果进一步进行计算, 会有

$$\mathsf{E}[\mathsf{E}[Y \mid \mathcal{F}_n]] = \mathsf{E}[Y],$$

$$\mathsf{E}[\mathsf{E}[Y \mid \mathcal{F}_n]\mathbb{1}_A] = \mathsf{E}[Y\mathbb{1}_A].$$

上述例子提示我们, 在一定条件下的最佳预测应具有一定的性质. 事实上, 人们引入条件数学期望解决了这一问题. 关于条件数学期望, 有如下定义:

定义 1.3.1 设 $(\Omega, \mathcal{F}, \mathsf{P})$ 是一个概率空间, $\mathcal{A}$ 是 $\mathcal{F}$ 的子 $\sigma-$ 代数. ξ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的可积随机变量, ξ 关于 $\mathcal{A}$ 的条件数学期望, 记为 $\mathsf{E}[\xi \mid \mathcal{A}]$, 是满足以下条件的唯一的随机变量 (在几乎处处的意义下):

(i) $\mathsf{E}[\xi \mid \mathcal{A}]$ 关于 $\mathcal{A}$ 可测;

(ii) 对于任意 $B \in \mathcal{A}$,

$$\int_B \xi\,\mathrm{d}\mathsf{P} = \int_B \mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P}.$$

上述定义中的存在性及唯一性可由 Radon-Nikodym 定理得到. 在 $(\Omega, \mathcal{A})$ 上, $\mathsf{P}_{\mathcal{A}}$ 是将 P 限制在 $\mathcal{A}$ 上时测度. 令 $\mu(A) = \int_A \xi\,\mathrm{d}\mathsf{P}$, 由 Radon-Nikodym 定理可得存在性与唯一性.

下面给出关于条件数学期望的一些性质.

命题 1.3.1 设 ξ, η, $\{\xi_n\}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的可积随机变量.

(1) $\mathsf{E}[\xi \mid \mathcal{F}] = \xi$; 若 ξ 与 $\mathcal{A}$ 独立, $\mathsf{E}[\xi \mid \mathcal{A}] = \mathsf{E}[\xi]$; $\mathsf{E}[\xi \mid \{\Omega, \emptyset\}] = \mathsf{E}[\xi]$ $a.s.$.

(2) 当 $\xi = a$ 时, $\mathsf{E}[\xi \mid \mathcal{A}] = a$ $a.s.$.

(3) 设 a, b 是常数, 则 $\mathsf{E}[(a\xi + b\eta) \mid \mathcal{A}] = a\mathsf{E}[\xi \mid \mathcal{A}] + b\mathsf{E}[\eta \mid \mathcal{A}]$ $a.s.$.

(4) 若 $\xi \leqslant \eta$, 则 $\mathsf{E}[\xi \mid \mathcal{A}] \leqslant \mathsf{E}[\eta \mid \mathcal{A}]$ $a.s.$.

(5) $|\mathsf{E}[\xi \mid \mathcal{A}]| \leqslant \mathsf{E}[|\xi| \mid \mathcal{A}]$ $a.s.$.

(6) $\mathsf{E}[\mathsf{E}[\xi \mid \mathcal{A}]] = \mathsf{E}[\xi]$ $a.s.$.

(7) 若 η 是 $\mathcal{A}$ 可测的, 则 $\mathsf{E}[\xi\eta \mid \mathcal{A}] = \eta\mathsf{E}[\xi \mid \mathcal{A}]$ $a.s.$.

(8) 若 $\mathcal{A} \subset \mathcal{B}$, 则 $\mathsf{E}[\mathsf{E}[\xi \mid \mathcal{B}] \mid \mathcal{A}] = \mathsf{E}[\xi \mid \mathcal{A}]$ $a.s.$.

(9) 若 $\lim\limits_{n\to\infty} \xi_n = \xi$ $a.s.$ 且 $|\xi_n| \leqslant \eta$, 则 $\lim\limits_{n\to\infty} \mathsf{E}[\xi_n \mid \mathcal{A}] = \mathsf{E}[\xi \mid \mathcal{A}]$ $a.s.$.

(10) 若 $\{\xi_n\}$ 是一列非负可积随机变量, $\mathcal{A} \subset \mathcal{F}$, 则

$$\mathsf{E}[\liminf_{n\to\infty} \xi_n | \mathcal{A}] \leqslant \liminf_{n\to\infty} \mathsf{E}[\xi_n | \mathcal{A}] \quad a.s..$$

证明: (1) 由于 ξ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量, 故由 Radon-Nikodym 定理的唯一性, 可知 $\mathsf{E}[\xi \mid \mathcal{F}] = \xi$.

若 ξ 与 $\mathcal{A}$ 独立, 令 $A \in \mathcal{A}$, $\mathbb{1}_A$ 与 ξ 独立,

$$\int_A \mathsf{E}[\xi \mid \mathcal{A}] \, \mathrm{d}\mathsf{P} = \int_A \xi \, \mathrm{d}\mathsf{P} = \int \xi \mathbb{1}_A \, \mathrm{d}\mathsf{P} = \mathsf{P}(A)\mathsf{E}[\xi] = \int_A \mathsf{E}[\xi] \, \mathrm{d}\mathsf{P},$$

故 $\mathsf{E}[\xi \mid \mathcal{A}] = \mathsf{E}[\xi]$. 由于 $\{\Omega, \emptyset\}$ 与任何随机变量独立, 故 $\mathsf{E}[\xi \mid \{\Omega, \emptyset\}] = \mathsf{E}[\xi]$.

(2) 当 $\xi = a$ 时，对于任意的 $A \in \mathcal{A}$,

$$\int_A a \, \mathrm{d}\mathsf{P} = a\mathsf{P}(A) = \int_A \mathsf{E}[\xi \mid \mathcal{A}] \, \mathrm{d}\mathsf{P}.$$

故

$$\mathsf{E}[\xi \mid \mathcal{A}] = a \ a.s..$$

(3) 对于任意 $A \in \mathcal{A}$,

$$\begin{aligned}
\int_A \mathsf{E}[a\xi + b\eta \mid \mathcal{A}] \, \mathrm{d}\mathsf{P} &= \int_A a\xi + b\eta \, \mathrm{d}\mathsf{P} \\
&= a\int_A \xi \, \mathrm{d}\mathsf{P} + b\int_A \eta \, \mathrm{d}\mathsf{P} \\
&= a\int_A \mathsf{E}[\xi \mid \mathcal{A}] \, \mathrm{d}\mathsf{P} + b\int_A \mathsf{E}[\eta \mid \mathcal{A}] \, \mathrm{d}\mathsf{P}.
\end{aligned}$$

其中 a, b 是常数, 即线性关系成立.

(4) 对于任意 $A \in \mathcal{A}$, 有

$$\int_A \mathsf{E}[\xi \mid \mathcal{A}] \, \mathrm{d}\mathsf{P} = \int_A \xi \, \mathrm{d}\mathsf{P} \leqslant \int_A \eta \, \mathrm{d}\mathsf{P} = \int_A \mathsf{E}[\eta \mid \mathcal{A}] \, \mathrm{d}\mathsf{P}.$$

故,

$$\mathsf{E}[\xi \mid \mathcal{A}] \leqslant \mathsf{E}[\eta \mid \mathcal{A}] \ a.s..$$

(5) 由于 $-|\xi| \leqslant \xi \leqslant |\xi|$, 由 (4) 可得.

(6) 由定义, 取 $A = \Omega$, 则

$$\int_\Omega \mathsf{E}[\xi \mid \mathcal{A}] \, \mathrm{d}\mathsf{P} = \int_\Omega \xi \, \mathrm{d}\mathsf{P}.$$

(7) 只考虑非负情形. 下证对任意 $A \in \mathcal{A}$,

$$\int_A \mathsf{E}[\xi\eta \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_A \eta\mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P}.$$

设 $G \in \mathcal{A}$, 取 $\eta = \mathbb{1}_G$, 因此

$$\int_A \mathbb{1}_G \mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_{A\cap G} \mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_{A\cap G} \xi\,\mathrm{d}\mathsf{P},$$

$$\int_A \mathsf{E}[\xi\mathbb{1}_G \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_A \xi\mathbb{1}_G\,\mathrm{d}\mathsf{P} = \int_{A\cap G} \xi\,\mathrm{d}\mathsf{P}.$$

故 $\int_A \mathbb{1}_G \mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_A \mathsf{E}[\xi\mathbb{1}_G \mid \mathcal{A}]\,\mathrm{d}\mathsf{P}$. 再由条件数学期望的线性性质及单调类定理可得结论.

(8) 由于 $\mathcal{A} \subset \mathcal{B}$, 故 $\mathsf{E}[\xi \mid \mathcal{A}]$ 是 $\mathcal{B}-$ 可测的, 从而

$$\mathsf{E}[\mathsf{E}[\xi \mid \mathcal{A}] \mid \mathcal{B}] = \mathsf{E}[\xi \mid \mathcal{A}].$$

进一步, 对于 $A \in \mathcal{A}$,

$$\int_A \mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_A \xi\,\mathrm{d}\mathsf{P},$$

$$\int_A \mathsf{E}[\mathsf{E}[\xi \mid \mathcal{B}] \mid \mathcal{A}]\,\mathrm{d}\mathsf{P} = \int_A \mathsf{E}[\xi \mid \mathcal{B}]\,\mathrm{d}\mathsf{P}.$$

由于 $\mathcal{A} \subset \mathcal{B}$, 故 $A \in \mathcal{B}$, 故

$$\int_A \mathsf{E}[\xi \mid \mathcal{B}]\,\mathrm{d}\mathsf{P} = \int_A \xi\,\mathrm{d}\mathsf{P} = \int_A \mathsf{E}[\xi \mid \mathcal{A}]\,\mathrm{d}\mathsf{P}.$$

故

$$\mathsf{E}[\mathsf{E}[\xi \mid \mathcal{B}] \mid \mathcal{A}] = \mathsf{E}[\xi \mid \mathcal{A}]\ a.s..$$

(9) 令 $\psi_n = \sup\limits_{k\geqslant n} |\xi_k - \xi|$, $\psi_n \downarrow 0$, 且 $|\psi_n| \leqslant 2\eta$, 故 $\mathsf{E}\psi_n \downarrow 0$, 而

$$|\mathsf{E}[\xi_n \mid \mathcal{A}] - \mathsf{E}[\xi \mid \mathcal{A}]| \leqslant \mathsf{E}[\psi_n \mid \mathcal{A}],$$

且 $\mathsf{E}[\psi_n \mid \mathcal{A}]$ 是单调的, 记其极限为 ψ, 且 $\psi \geqslant 0$,

$$\mathsf{E}[\psi] = \mathsf{E}[\mathsf{E}[\psi \mid \mathcal{A}]] \leqslant \mathsf{E}[\mathsf{E}[\psi_n \mid \mathcal{A}]] = \mathsf{E}[\psi_n],$$

故 $\mathsf{E}[\psi] = 0$, 即 $\psi = 0\ a.s.$, 即有

$$\lim_{n\to\infty} \mathsf{E}[\xi_n \mid \mathcal{A}] = \mathsf{E}[\xi \mid \mathcal{A}]\ a.s..$$

(10) 对于 $A \in \mathcal{A}$, 记 $\eta_n = \inf_{k \geqslant n} \xi_k$, 由单调收敛定理,

$$\int_A \liminf_{n\to\infty} \xi_n \, \mathrm{d}P = \int_A \lim_{n\to\infty} \eta_n \, \mathrm{d}P = \lim_{n\to\infty} \int_A \eta_n \, \mathrm{d}P = \lim_{n\to\infty} \int_A \mathsf{E}[\eta_n|\mathcal{A}] \, \mathrm{d}P.$$

又由单调定理,

$$\lim_{n\to\infty} \int_A \mathsf{E}[\eta_n|\mathcal{A}] \, \mathrm{d}P = \int_A \lim_{n\to\infty} \mathsf{E}[\eta_n|\mathcal{A}] \, \mathrm{d}P.$$

因此

$$\mathsf{E}[\liminf_{n\to\infty} \xi_n|\mathcal{A}] = \lim_{n\to\infty} \mathsf{E}[\eta_n|\mathcal{A}] \quad a.s..$$

因为 $\eta_n \leqslant \xi_n$, 因此, $\mathsf{E}[\eta_n|\mathcal{A}] \leqslant \mathsf{E}[\xi_n|\mathcal{A}]$, $a.s.$. 于是

$$\lim_{n\to\infty} \mathsf{E}[\eta_n|\mathcal{A}] \leqslant \liminf_{n\to\infty} \mathsf{E}[\xi_n|\mathcal{A}] \quad a.s.,$$

即

$$\mathsf{E}[\liminf_{n\to\infty} \xi_n|\mathcal{A}] \leqslant \liminf_{n\to\infty} \mathsf{E}[\xi_n|\mathcal{A}] \quad a.s..$$

■

考虑给定的随机变量条件下的条件数学期望, 本质上是考虑随机变量生成的 $\sigma-$ 代数的条件下的条件数学期望.

例如, 若 X 与 Y 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量, $\mathsf{E}[Y \mid X]$ 即为 $\mathsf{E}[Y \mid \sigma(X)]$, 前面已经知道, $\sigma(X)$ 为 X 生成的 $\sigma-$ 代数, 是使得 X 可测的最小 $\sigma-$ 代数.

回到最佳预测问题, 有如下结论:

定理 1.3.1 设 ξ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的平方可积随机变量, $\mathcal{A} \subset \mathcal{F}$ 是 $\sigma-$ 代数, 则

$$\mathsf{E}[\xi|\mathcal{A}] = \underset{\{\text{关于 } \mathcal{A} \text{ 可测的平方可积随机变量 } \eta\}}{\arg\min} \mathsf{E}[(\xi - \eta)^2].$$ ①

证明: 对于任意的 $\mathcal{A}-$ 可测的平方可积随机变量 η,

$$\begin{aligned} \mathsf{E}[(\xi-\eta)^2] &= \mathsf{E}[(\xi - \mathsf{E}[\xi|\mathcal{A}] + \mathsf{E}[\xi|\mathcal{A}] - \eta)^2] \\ &= \mathsf{E}[(\xi - \mathsf{E}[\xi|\mathcal{A}])^2] + \mathsf{E}[(\mathsf{E}[\xi|\mathcal{A}] - \eta)^2] \\ &\quad + 2\mathsf{E}[(\xi - \mathsf{E}[\xi|\mathcal{A}])(\mathsf{E}[\xi|\mathcal{A}] - \eta)]. \end{aligned}$$

由于

$$\begin{aligned} & \mathsf{E}[(\xi - \mathsf{E}[\xi|\mathcal{A}])(\mathsf{E}[\xi|\mathcal{A}] - \eta)] = \mathsf{E}[\mathsf{E}[(\xi - \mathsf{E}[\xi|\mathcal{A}])(\mathsf{E}[\xi|\mathcal{A}] - \eta)|\mathcal{A}]] \\ = \; & \mathsf{E}[(\mathsf{E}[\xi|\mathcal{A}] - \eta)\mathsf{E}[(\xi - \mathsf{E}[\xi|\mathcal{A}])|\mathcal{A}]]] \\ = \; & 0. \end{aligned}$$

① argmin 表示使目标函数取最小值时得结果.

故
$$\mathsf{E}[(\xi-\eta)^2]=\mathsf{E}[(\xi-\mathsf{E}[\xi|\mathcal{A}])^2]+\mathsf{E}[(\mathsf{E}[\xi|\mathcal{A}]-\eta)^2],$$
即
$$\mathsf{E}[(\xi-\mathsf{E}[\xi|\mathcal{A}])^2]\leqslant\mathsf{E}[(\xi-\eta)^2].$$
得证. ■

事实上, 利用条件数学期望, 我们可以证明一些原本很难证明的结论. 下面列举几个例子.

例 1.3.1 *若 ξ, η 是独立的可积随机变量, $\mathsf{E}\xi=0$, $\mathsf{E}\eta=0$, 则 $\mathsf{E}|\xi|\leqslant\mathsf{E}|\xi+\eta|$.*

证明:
$$\mathsf{E}[\xi+\eta|\xi]=\xi+\mathsf{E}[\eta|\xi]=\xi,$$
故 $|\mathsf{E}[\xi+\eta|\xi]|=|\xi|$. 从而
$$\mathsf{E}|\xi|=\mathsf{E}|\mathsf{E}[\xi+\eta|\xi]|\leqslant\mathsf{E}[\mathsf{E}[|\xi+\eta|\mid\xi]]=\mathsf{E}[|\xi+\eta|].$$
■

例 1.3.2 *若 ξ, η 是独立同分布的可积随机变量, 求 $\mathsf{E}[\xi|\xi+\eta]$.*

解: 由于 ξ, η 是独立同分布的, 我们首先证明,
$$\mathsf{E}[\xi|\xi+\eta]=\mathsf{E}[\eta|\xi+\eta].$$
对于任意的实数 x, $A=\{\xi+\eta\leqslant x\}$, 我们需要证明
$$\int_A\xi d\mathsf{P}=\int_A\eta d\mathsf{P}.$$
事实上，对于任意的 Borel 集 B,
$$\int_A\mathbb{1}_B(\xi)d\mathsf{P}=\mathsf{P}(\xi\in B,\xi+\eta\leqslant x).$$
由于 ξ, η 是独立同分布, 因此
$$\int_A\mathbb{1}_B(\xi)d\mathsf{P}=\int_A\mathbb{1}_B(\eta)d\mathsf{P}.$$
由于随机变量都可以由简单随机变量逼近, 因此，我们得到了
$$\int_A\xi d\mathsf{P}=\int_A\eta d\mathsf{P}.$$

进一步,

$$\mathsf{E}[\xi+\eta|\xi+\eta]=\xi+\eta.$$

从而

$$\mathsf{E}[\xi|\xi+\eta]=\frac{\xi+\eta}{2}.$$

■

例 1.3.3 设 $\xi_1,\xi_2,\cdots,\xi_n$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的一列独立同分布的随机变量, $\mathsf{E}\xi_n=0$, $\mathsf{E}\xi_n^2=1$. 令 $S_n=\sum_{i=1}^{n}\xi_i$, $\mathcal{F}_n=\sigma(\xi_1,\xi_2,\cdots,\xi_n)$, 对于 $m<n$, 求 $\mathsf{E}[S_n^2|\mathcal{F}_m]$.

解:

$$\begin{aligned}\mathsf{E}[S_n^2|\mathcal{F}_m]&=\mathsf{E}[(S_n-S_m+S_m)^2|\mathcal{F}_m]\\&=\mathsf{E}[S_m^2+(S_n-S_m)^2+2(S_n-S_m)S_m|\mathcal{F}_m]\\&=S_m^2+2S_m\mathsf{E}[S_n-S_m|\mathcal{F}_m]+\mathsf{E}[(S_n-S_m)^2|\mathcal{F}_m]\\&=S_m^2+2S_m\mathsf{E}[S_n-S_m]+\mathsf{E}[(S_n-S_m)^2]\\&=S_m^2+(n-m).\end{aligned}$$

■

事实上, 我们注意到有 $\mathcal{F}_n\subseteq\mathcal{F}_{n+1}$, 我们称 $\{\mathcal{F}_n\}$ 为流. 最后, 我们给出两个定理, 在后面的内容中会用到.

定理 1.3.2 设 ξ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的可积随机变量, $\mathcal{A}\subset\mathcal{F}$ 是 $\sigma-$ 代数, Φ 是 $\mathbb{R}$ 上的凸函数, $\Phi(\xi)$ 是可积的, 则

$$\Phi(\mathsf{E}(\xi|\mathcal{A}))\leqslant\mathsf{E}[\Phi(\xi)|\mathcal{A}].$$

证明 [10]①: Φ 是凸函数, 故 Φ 左、右导数都存在. 设 $\phi'_+(x)$ 是其右导数, 且对 $x_0\in\mathbb{R}$,

$$\phi'_+(x_0)(x-x_0)+\Phi(x_0)\leqslant\Phi(x).$$

令 $x=\xi$, $x_0=\mathsf{E}[\xi|\mathcal{A}]$,

$$\phi'_+(\mathsf{E}[\xi|\mathcal{A}])(\xi-\mathsf{E}[\xi|\mathcal{A}])+\Phi(\mathsf{E}[\xi|\mathcal{A}])\leqslant\Phi(\xi).$$

由于 $\Phi(\xi)$ 可积, 若 $\mathsf{E}[\xi|\mathcal{A}]$ 有界, 则左边各项可积, 两边取条件数学期望, 得

$$\mathsf{E}[\phi'_+(\mathsf{E}[\xi|\mathcal{A}])(\xi-\mathsf{E}[\xi|\mathcal{A}])|\mathcal{A}]+\mathsf{E}[\Phi(\mathsf{E}[\xi|\mathcal{A}])|\mathcal{A}]\leqslant\mathsf{E}[\Phi(\xi)|\mathcal{A}],$$

① 表示证明过程或解答过程见参考文献 [7], 全书同.

即 $\Phi(\mathsf{E}[\xi|\mathcal{A}]) \leqslant \mathsf{E}[\Phi(\xi)|\mathcal{A}]$. 一般地, 令

$$G_n = \{\omega : \mathsf{E}[|\xi| \mid \mathcal{A}] \leqslant n\},\ G_n \in \mathcal{A},\ G_n \uparrow \Omega,$$

则,

$$\begin{aligned}\Phi(\mathsf{E}[\xi \mathbb{1}_{G_n}|\mathcal{A}]) &\leqslant \mathsf{E}[\Phi(\xi \mathbb{1}_{G_n})|\mathcal{A}] \\ &= \mathsf{E}[\mathbb{1}_{G_n}\Phi(\xi) + \mathbb{1}_{G_n^c}\Phi(0)|\mathcal{A}],\end{aligned}$$

其中 G_n^c 为 G_n 的补集. 由控制收敛定理得到本定理的结论. ■

上述定理的不等式称为 Jensen 不等式.

定理 1.3.3 设 X, Y 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的可积随机变量, $\mathcal{A} \subset \mathcal{F}$ 是 $\sigma-$ 代数, X 与 $\mathcal{A}$ 独立, Y 关于 $\mathcal{A}$ 可测, 则对于任意可测函数 f, 若 $\mathsf{E}[|f(X,Y)|] < \infty$, 则有

$$\mathsf{E}[f(X,Y)|\mathcal{A}] = \mathsf{E}[f(X,y)]\ |_{y=Y}\ .$$

证明: 我们只对任意非负有界可测函数 f 进行证明. 令 $g(y) = \mathsf{E}[f(X,y)]$, 下证对于任意非负 $\mathcal{A}-$ 可测随机变量 Z, 有

$$\mathsf{E}[f(X,Y)Z] = \mathsf{E}[g(Y)Z].$$

对于任意 Borel 集 B, 令 $\mu_X(B) = \mathsf{P}(\{\omega : X(\omega) \in B\})$.
对于 $\mathbb{R}^2$ 上的 Borel 集 A, 令 $\mu_{Y,Z}(A) = \mathsf{P}(\{\omega : (Y(\omega), Z(\omega)) \in A\})$.
于是,

$$g(y) = \int f(x,y)\mu_X(dx).$$

由于 X 与 (Y,Z) 独立, 故

$$\begin{aligned}\mathsf{E}[f(X,Y)Z] &= \int z f(x,y)\mu_X(dx)\mu_{Y,Z}(dy,dz) \\ &= \int z g(y)\mu_{Y,Z}(dy,dz) = \mathsf{E}[g(Y)Z].\end{aligned}$$

■

一般情况下，条件期望是计算不出的. 但是在特别的设定下，条件期望可以计算出. 下面给出几个具体的例子.

设 (X,Y) 具有联合密度函数 $p(x,y)$. 对于有界 Borel 可测函数 f, 有

$$\mathsf{E}[f(X,Y)] = \int\int f(x,y)p(x,y)dxdy.$$

设 $h(x)$ 是 Borel 可测函数, 且 $\mathsf{E}[|h(X)|] < \infty$, 下面计算 $\mathsf{E}[h(X)|Y]$.

记 $q(y) = \int p(x,y)dx$. 对于 Borel 可测函数 $g(y)$, 假设 $\mathsf{E}[|h(X)g(Y)|] < \infty$. 注意到 $q(y)=0$，意味着 $p(x,y)=0$ 关于 x 几乎处处成立. 此时有

$$\begin{aligned}\mathsf{E}[h(X)g(Y)] &= \int\int h(x)g(y)p(x,y)dxdy \\ &= \int\left(\int h(x)p(x,y)dx\right)g(y)\mathbb{1}_{q(y)>0}dy \\ &= \int\left(\frac{\int h(x)p(x,y)dx}{q(y)}\right)g(y)q(y)\mathbb{1}_{q(y)>0}dy\end{aligned}$$

记

$$\psi(y) = \begin{cases}\dfrac{\int h(x)p(x,y)dx}{q(y)}, & 若 \quad q(y)>0, \\ h(0), & 若 \quad q(y)=0.\end{cases}$$

因此

$$\begin{aligned}\mathsf{E}[h(X)g(Y)] &= \int \psi(y)g(y)q(y)\mathbb{1}_{q(y)>0}dy \\ &= \mathsf{E}[\psi(Y)g(Y)].\end{aligned}$$

若取 $g(x)=\mathbb{1}_A$, 由条件期望的定义, 可以知道

$$\mathsf{E}[h(X)|Y] = \psi(Y).$$

给定 y, 可以定义一个 $\mathbb{R}$ 上的测度:

记

$$\nu(y,dx) = \begin{cases}\dfrac{p(x,y)}{q(y)}dx, & 若 \quad q(y)>0, \\ \delta_0(dx), & 若 \quad q(y)=0,\end{cases}$$

这里 $\delta_0(A) = \begin{cases}1, & 若 \quad 0\in A, \\ 0, & 若 \quad 0\notin A.\end{cases}$ 此时, 会有

$$\mathsf{E}[h(X)|Y] = \int h(x)\nu(Y,dx).$$

特别地，若 $q(y>0)$, 此时

$$\mathsf{E}[h(X)|Y=y] = \int h(x)\nu(y,dx) = \frac{\int h(x)p(x,y)dx}{q(y)},$$

因此, $x \mapsto \frac{p(x,y)}{q(y)}$ 可以看作给定 Y 的条件下 X 的条件密度.

在很多场合里 $\nu(y,dx)$ 常被称为转移概率, 由条件期望的定义, 对于不同的 y, $\nu(y,dx)$ 不能保证是唯一的，事实上这和正则条件概率有关. 由 Kallenberg [7] 中定理 6.3, 可以保证对于不同的 y, $\nu(y,dx)$ 是唯一的. 因此，在初等概率论中引入条件密度也是有意义的.

特别地，在 $(\Omega,\mathcal{F},\mathsf{P})$ 上，$\mathcal{V}=\sigma(A_1,A_2,\cdots,A_n)$, 这里 $A_i\bigcap A_j=\emptyset$, $i\neq j$, $\bigcup_{i=1}^n A_i=\Omega$, 设 $\mathsf{P}(A_i)>0$, $i=1,\cdots,n$, X 是可积随机变量, 现在我们证明:

$$\mathsf{E}[X|\mathcal{V}]=\sum_{i=1}^n a_i\mathbb{1}_{A_i}, \tag{1.1}$$

这里 $a_i=\frac{1}{\mathsf{P}(A_i)}\mathsf{E}[X\mathbb{1}_{A_i}]$. 事实上 $\mathcal{V}$ 中的集合除了空集外，均可表示成 $\bigcup_{i=1}^m A_{n_i}$, 因此, 由条件期望的定义, 为了证明 (1.1), 只需证明对于任意的 A_i,

$$\int_{A_i}\sum_{i=1}^n a_i\mathbb{1}_{A_i}d\mathsf{P}=\int_{A_i}Xd\mathsf{P}.$$

事实上，由

$$\int_{A_i}\sum_{i=1}^n a_i\mathbb{1}_{A_i}d\mathsf{P}=\int_{A_i}a_i\mathbb{1}_{A_i}d\mathsf{P}=\frac{\mathsf{P}(A_i)}{\mathsf{P}(A_i)}\mathsf{E}[X\mathbb{1}_{A_i}]=\mathsf{E}[X\mathbb{1}_{A_i}].$$

因此, (1.1) 成立.

例 1.3.4 设 X 与 Y 独立，均服从标准正态分布. 求 $\mathsf{E}[X^2|XY]$.

解: 现在, 对 (X,Y) 进行极坐标变换:

$$(X,Y)=R\cdot(\cos\Theta,\sin\Theta).$$

设 (X,Y) 的联合密度函数是 $p(x,y)$, 则

$$\begin{aligned}\iint_{R\times R}p(x,y)dxdy&=\int_R\int_R\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}dxdy\\&=\frac{1}{2\pi}\int_0^\infty\int_0^{2\pi}re^{-\frac{r^2}{2}}drd\theta\\&=\int_0^\infty re^{-\frac{r^2}{2}}dr\int_0^{2\pi}\frac{d\theta}{2\pi}.\end{aligned}$$

则 R 和 Θ 的密度函数分别是:

$$p_R(r)=re^{-\frac{r^2}{2}}\mathbf{1}_{\{r>0\}}\quad p_\Theta(\theta)=\frac{\mathbf{1}_{\{0<\theta<2\pi\}}}{2\pi}.$$

注意到

$$\mathsf{E}\left(X^2 \mid XY\right)=\frac{1}{2}\mathsf{E}\left(X^2+Y^2 \mid XY\right)=\mathsf{E}(\frac{R^2}{2} \mid \frac{R^2}{2}\sin(2\Theta))=\mathsf{E}(S \mid SV).$$

这里 $S=\frac{R^2}{2}, V=\sin(2\Theta)$，$S$ 和 V 相互独立，且密度函数分别是

$$f_S(s)=e^{-s}\mathbf{1}_{s>0},\quad f_V(v)=\frac{\mathbf{1}_{|v|<1}}{\pi\sqrt{1-v^2}}.$$

令 $\mathsf{E}(S \mid SV)=g(SV)$, 则对任意有界可测函数 h ,

$$\mathsf{E}(Sh(SV))=\mathsf{E}(\mathsf{E}(Sh(SV) \mid SV))=\mathsf{E}(\mathsf{E}(S \mid SV)h(SV))=\mathsf{E}(g(SV)h(SV)).$$

即

$$\iint sh(sv)f_S(s)f_V(v)dsdv=\int g(w)h(w)f_{SV}(w)dw.$$

对上式左边进行变换，有

$$\iint sh(sv)f_S(s)f_V(v)dsdv=\iint h(w)f_S(s)f_V\left(s^{-1}w\right)dsdw.$$

因此,

$$g(w)f_{SV}(w)=\int f_S(s)f_V\left(s^{-1}w\right)ds.$$

即

$$g(w)\cdot\int_{|w|}^{\infty}e^{-s}\frac{ds}{\pi\sqrt{s^2-w^2}}=\int_{|w|}^{\infty}e^{-s}\frac{sds}{\pi\sqrt{s^2-w^2}}.$$

令 $s=|w|\cdot\cosh t$,

$$g(w)\cdot\int_0^{\infty}e^{-|w|\cosh t}dt=|w|\int_0^{\infty}e^{-|w|\cosh t}\cosh tdt.$$

上式可以用 Bessel 函数进行表示，

$$g(w)\cdot K_0(|w|)=|w|\cdot K_1(|w|),$$

这里

$$K_\alpha(x)=\int_0^{\infty}e^{-x\cosh t}\cosh\alpha tdt.$$

于是

$$\mathsf{E}\left(X^2 \mid XY\right)=g(XY)=|XY|\frac{K_1(|XY|)}{K_0(|XY|)}.$$

■

1.4 习题

1. 设 $X_1, X_2, \cdots$, 是独立同分布的随机变量, 且

$$\mathsf{P}(X_1 = 3) = \mathsf{P}(X_1 = -1) = \frac{1}{2}.$$

令 $S_n = \sum_{i=1}^{n} X_i$, $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$.

(1) 求 $\mathsf{E}[S_n], \mathsf{E}[S_n^2], \mathsf{E}_{[}S_n^3]$;

(2) 若 $m < n$, 求 $\mathsf{E}[S_n|\mathcal{F}_m], \mathsf{E}[S_n^2|\mathcal{F}_m], \mathsf{E}[S_n^3|\mathcal{F}_m]$;

(3) 求 $\mathsf{E}[X_1|S_n]$;

(4) 求 $\mathsf{E}[X_1|X_2, S_{10}]$;

2. 设 Y, Z 是独立的随机变量, 且均服从标准正态分布, 求 $\mathsf{E}[e^{YZ^2}|Z]$.

3. 若 ξ 是可积的随机变量, 请证明:

$$\mathsf{E}[|\xi|] = \int_0^{\infty} \mathsf{P}(|\xi| > t)\, \mathrm{d}t.$$

4. $\{\xi_n\}$ 是一列一致有界的非负随机变量列, ξ 是非负的有界的连续型随机变量, 若 $\xi_n \Longrightarrow \xi$, 请证明:

$$\lim_{n\to\infty} \mathsf{E}[\xi_n] = \mathsf{E}[\xi].$$

5. 设 $X_1, X_2, \cdots, X_n$ 是同分布的随机变量, 均服从标准正态分布, 请证明: 存在与 n 无关常数 $C > 0$, 使得

$$\mathsf{E}[\max_{1\leqslant i\leqslant n} |X_i|] \leqslant C\sqrt{\log n}.$$

6. 设 $X_1, X_2, \cdots, X_n$ 是独立同分布的随机变量, 均服从标准正态分布, 请证明: 存在与 n 无关常数 $C_1 > 0, C_2 > 0$, 使得

$$C_1\sqrt{\log n} \leqslant \mathsf{E}[\max_{1\leqslant i\leqslant n} |X_i|] \leqslant C_2\sqrt{\log n}.$$

7. 设 $\mathbf{X} = (X_1, X_2, X_3)^\top \sim N_3(0, \Sigma)$,

$$\Sigma = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

(1) 证明 $\mathbf{W} = (X_1, X_2)^\top$ 服从多元正态分布;

(2) 求 $Y = X_1 + X_2$ 的分布;

(3) 求 $\mathsf{E}[X_1|X_2, X_3]$;

(4) 求 $\mathsf{E}[e^{X_1X_3}|X_1]$;

第 2 章　离散时间鞅

本章介绍离散时间鞅的定义及性质.

2.1　定义及性质

在随机过程中, 鞅是重要的分支之一. 利用鞅的性质, 很多概率问题迎刃而解. 本书主要围绕鞅展开对随机过程的讨论. 在这一章中, 我们首先介绍离散时间鞅.

设 $(\Omega, \mathcal{F}, \mathsf{P})$ 是一个完备概率空间, 即 $\mathcal{F}$ 包含任意 $\mathsf{P}-$ 零集的全部子集. $\{\mathcal{F}_n\}_{n\geqslant 0}$ 是一列 $\mathcal{F}$ 的子 $\sigma-$ 代数, 且 $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$. 若无特别说明，假设 $\mathcal{F}_0$ 包含所有 $\mathsf{P}-$ 零测集, $\{\mathcal{F}_n\}_{n\geqslant 0}$ 称为流. $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 称为带流概率空间. 一个随机过程 $X = (X_n)_{n\geqslant 0}$, 即对于每个 n, X_n 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的随机变量.

定义 2.1.1　对于 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的随机过程 $X = (X_n)_{n\geqslant 0}$, 若 X_n 是 $\mathcal{F}_n-$ 可测的, 称 $X = (X_n)_{n\geqslant 0}$ 在 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上是 适应的.

例 2.1.1　$X = (X_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的一列随机变量, 设 $\mathcal{F}_n = \sigma(X_k; k \leqslant n)$, 显然 X 在 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上是适应的.

定义 2.1.2　$(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的随机过程 $X = (X_n)_{n\geqslant 0}$, 对于每个 n, $\mathsf{E}[|X_n|] < +\infty$, 且 X 是适应的, 若对所有 $n \geqslant 1$, $\mathsf{E}[X_n|\mathcal{F}_{n-1}] = X_{n-1}$, 则称 $X = (X_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的鞅. 若对所有 n, $\mathsf{E}[X_n|\mathcal{F}_{n-1}] \geqslant X_{n-1}$, 则称 $X = (X_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的下鞅. 若对所有 n, $\mathsf{E}[X_n|\mathcal{F}_{n-1}] \leqslant X_{n-1}$, 则称 $X = (X_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的上鞅.

例 2.1.2　设 $\{\xi_n\}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的一列相互独立的 Bernoulli 序列, $\mathsf{P}(\xi_n = 1) = \frac{1}{2}$, $\mathsf{P}(\xi_n = -1) = \frac{1}{2}$. 令 $X_0 = 0$, $X_n = \sum\limits_{i=1}^{n} \xi_i$, $\mathcal{F}_n = \sigma(\xi_1, \xi_2, \cdots, \xi_n)$, 又设 $H = (H_n)_{n\geqslant 1}$ 是一列随机变量, H_n 关于 $\mathcal{F}_{n-1}-$ 可测, 令 $Y_0 = 0$, 且当 $n \geqslant 1$ 时,

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

证明 $X = (X_n)_{n\geqslant 0}, Y = (Y_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的鞅.

证明: 注意到

$$\mathsf{E}[X_n|\mathcal{F}_{n-1}] = \mathsf{E}[\sum_{i=1}^{n-1} \xi_i + \xi_n|\mathcal{F}_{n-1}] = \sum_{i=1}^{n-1} \xi_i + \mathsf{E}[\xi_n] = X_{n-1}.$$

因此, $X=(X_n)_{n\geqslant 0}$ 是鞅. 进一步,

$$\mathsf{E}[Y_n|\mathcal{F}_{n-1}]=Y_{n-1}+H_n\mathsf{E}[X_n-X_{n-1}|\mathcal{F}_{n-1}]=Y_{n-1},$$

故 $Y=(Y_n)_{n\geqslant 0}$ 是鞅. ■

例 2.1.3 设 $(\Omega,\mathcal{F},\mathsf{P})$ 上有一列独立同分布的随机变量 $\{\xi_n\}$, $\mathsf{E}\xi_n=0$, $\mathsf{E}\xi_n^2=1$, $S_n=\sum\limits_{i=1}^{n}\xi_i$, $\mathcal{F}_n=\sigma(\xi_1,\xi_2,\cdots,\xi_n)$, 则 $(S_n^2-n)_{n\geqslant 0}$ 是鞅.

证明: 注意到当 $n>m$,

$$\mathsf{E}[S_n^2|\mathcal{F}_m]=\mathsf{E}[(S_n-S_m+S_m)^2|\mathcal{F}_m]=\mathsf{E}[(S_n-S_m)^2+S_m^2+2(S_n-S_m)S_m|\mathcal{F}_m],$$

这里 $\mathsf{E}[(S_n-S_m)^2|\mathcal{F}_m]=n-m$, $\mathsf{E}[2(S_n-S_m)S_m|\mathcal{F}_m]=0$, 故

$$\mathsf{E}[S_n^2-n|\mathcal{F}_m]=S_m^2-m$$

因此, $(S_n^2-n)_{n\geqslant 0}$ 是鞅. ■

例 2.1.4 设 $\{Y_n\}$ 是独立同分布的随机变量, f_0, f_1 是两个密度函数, $f_0>0$, 令

$$X_n=\frac{f_1(Y_1)f_1(Y_2)\cdots f_1(Y_n)}{f_0(Y_1)f_0(Y_2)\cdots f_0(Y_n)}\quad \mathcal{F}_n=\sigma(Y_1,Y_2,\cdots,Y_n),$$

证明: 若 f_0 是 Y_n 的密度函数, 则 $(X_n)_{n\geqslant 0}$ 是鞅.

证明: 由于

$$\mathsf{E}[X_{n+1}|\mathcal{F}_n]=X_n\mathsf{E}[\frac{f_1(Y_{n+1})}{f_0(Y_{n+1})}|\mathcal{F}_n]=X_n\mathsf{E}[\frac{f_1(Y_{n+1})}{f_0(Y_{n+1})}].$$

$$\mathsf{E}[\frac{f_1(Y_{n+1})}{f_0(Y_{n+1})}]=\int_{-\infty}^{+\infty}\frac{f_1(y)}{f_0(y)}f_0(y)\,\mathrm{d}y=\int_{-\infty}^{+\infty}f_1(y)\,\mathrm{d}y=1.$$

故则 $(X_n)_{n\geqslant 0}$ 是鞅. ■

例 2.1.5 设 $\{X_n\}$ 是一列 $\mathcal{F}_n-$ 适应的可积随机变量序列, 满足

$$\mathsf{E}[X_{n+1}|\mathcal{F}_n]=\alpha X_n+\beta X_{n-1},\ \ \alpha+\beta=1,\ \alpha>0,\ \beta>0,$$

定义 $Y_0=X_0,Y_n=aX_n+X_{n-1}$. 求 a, 使得 $(Y_n)_{n\geqslant 0}$ 是鞅.

解: 事实上

$$\begin{aligned}\mathsf{E}[Y_{n+1}|\mathcal{F}_n]&=\mathsf{E}[aX_{n+1}+X_n|\mathcal{F}_n]\\&=a\alpha X_n+a\beta X_{n-1}+X_n,\end{aligned}$$

即

$$\mathsf{E}[Y_{n+1}|\mathcal{F}_n] = (a\alpha+1)X_n + a\beta X_{n-1}.$$

由 $(Y_n)_{n\geqslant 0}$ 是鞅, 故 $\mathsf{E}(Y_{n+1}|\mathcal{F}_n) = Y_n = aX_n + X_{n-1}$. 即 $a\alpha+1=a$, $a\beta=1$, 有 $a=\dfrac{1}{\beta}$. ■

停时在随机分析的研究中扮演了十分重要的角色, 毫不夸张地说, 停时是随机分析这个学科中极具特色的概念. 我们首先给出停时的定义.

定义 2.1.3 在 $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geqslant 0},\mathsf{P})$ 上, τ 是一个可取值为 ∞ 的非负整数值广义随机变量[①], 且满足对于任意 n, $\{\omega:\tau(\omega)\leqslant n\}\in\mathcal{F}_n$, 称 τ 是停时或 $\mathcal{F}_n-$ 停时.

例 2.1.6 当 A 是 Borel 集时, $X=(X_n)_{n\geqslant 0}$ 是 $\mathcal{F}_n-$ 适应的, 定义

$$\tau(\omega) = \inf\{n : X_n(\omega)\in A\}.$$

由于

$$\{\tau\leqslant n\} = \bigcup_{k\leqslant n}\{X_k\in A\}\in\mathcal{F}_n.$$

因此, τ 是一个停时, 我们往往称 τ 为初遇 (Début).

与初遇相关的停时还有下面的例子.

例 2.1.7 当 A 是 Borel 集时, $X=(X_n)_{n\geqslant 0}$ 是 $\mathcal{F}_n-$ 适应的, S 是停时, 定义

$$T = \inf\{n > S : X_n\in A\}.$$

由于

$$\{T\leqslant n\} = \bigcup_{k\leqslant n}\{X_k\in A\}\cap\{S\leqslant k\}\in\mathcal{F}_n.$$

因此, T 是一个停时.

下面给出一个和停时有关的例子.

例 2.1.8 考虑 X 为游戏的结果, $X_i=1$ 代表第 i 次游戏获胜, $X_i=-1$ 代表第 i 次游戏失败. 设 $(\Omega,\mathcal{F},\mathsf{P})$ 上有一列独立同分布的随机变量 $\{X_n\}$, $\mathsf{P}(X_i=1)=\mathsf{P}(X_i=-1)=\frac{1}{2}$, 定义

$$W_n = \sum_{i=1}^{n} B_i X_i.$$

① 广义随机变量指可取值 ∞ 或者 $-\infty$ 的可测映射.

这里 $B_1 = 1$, $j > 1$ 时,

$$B_j = \begin{cases} 2^{j-1}, & X_1 = X_2 = \cdots = X_{j-1} = -1, \\ 0, & \text{其它}, \end{cases}$$

这里 $B_i \in \mathcal{F}_{i-1}$, $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$. 不难得到 $W = (W_n)_{n\geqslant 1}$ 是一个 $\mathcal{F}_n-$鞅. 这里很有意思的是, 当 $X_1 = X_2 = \cdots = X_n = -1$ 时,

$$W_n = -1 - 2 - 2^2 - \cdots - 2^{n-1} = -[2^n - 1].$$

其它情况, $W_n = 1$. 因此

$$\mathsf{E}[W_n] = 1 \cdot [1 - 2^{-n}] - [2^n - 1] \cdot 2^{-n} = 0.$$

所以, 无论过程如何, W_n 期望为 0. $X_n = 1$ 的时刻即为停时.

对于停时 T, X_T 是 T 时刻 X 所处的位置, 即

$$X_T(\omega) = X_{T(\omega)}(\omega).$$

定义 2.1.4 对于 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的随机过程 $X = (X_n)_{n\geqslant 0}$, 定义 $X^\tau = (X_n^\tau)_{n\geqslant 0}$, $X_n^\tau = X_{\tau\wedge n}$[①], 我们称 X^τ 为停止过程.

我们下面给出一个十分重要的结论- Doob 停止定理.

定理 2.1.1 (Doob 停止定理) 假设 T 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的停时, $M = (M_n)_{n\geqslant 0}$ 是鞅, 则 $M^T = (M_{n\wedge T})_{n\geqslant 0}$ 是鞅, 且对任意 n, $\mathsf{E}[M_{n\wedge T}] = \mathsf{E}[M_0]$.

证明:

$$\begin{aligned} M_n^T - M_{n-1}^T &= M_{n\wedge T} - M_{(n-1)\wedge T} \\ &= \mathbb{1}_{\{T\geqslant n\}}(M_n - M_{n-1}). \end{aligned}$$

这里 $\mathbb{1}_{\{T\geqslant n\}} = 1 - \mathbb{1}_{\{T<n\}} \in \mathcal{F}_{n-1}$. 故

$$\mathsf{E}[M_n^T - M_{n-1}^T | \mathcal{F}_{n-1}] = \mathbb{1}_{\{T\geqslant n\}} \mathsf{E}[M_n - M_{n-1} | \mathcal{F}_{n-1}] = 0.$$

故 M^T 是鞅, 且 $\mathsf{E}[M_{n\wedge T}] = \mathsf{E}[M_0]$. ■

利用控制收敛定理, 有

定理 2.1.2 若 T 是有界停时, $\mathsf{E}[M_T] = \mathsf{E}[M_0]$.

定理 2.1.2 中停时的有界性, 是非常强的条件, 可以做如下推广:

① $\tau \wedge n = \min\{\tau, n\}$.

定理 2.1.3　设 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上, $M = (M_n)_{n\geqslant 0}$ 是鞅, T 是停时, $\mathsf{P}(T < +\infty) = 1$, $\mathsf{E}[|M_T|] < +\infty$. 且

$$\lim_{n\to\infty} \mathsf{E}[|M_n| \mathbb{1}_{\{T>n\}}] = 0.$$

则 $\mathsf{E}[M_T] = \mathsf{E}[M_0]$.

证明: 由定理 2.1.1, M^T 是鞅, $\mathsf{E}[M_{n\wedge T}] = \mathsf{E}[M_0]$,

$$\mathsf{E}[M_{n\wedge T}] = \mathsf{E}[M_T] + \mathsf{E}[M_{n\wedge T} - M_T],$$

$$M_{n\wedge T} - M_T = \mathbb{1}_{\{T>n\}}(M_n - M_T).$$

由于 $\mathsf{E}[|M_T|] < +\infty$, 故 $\lim\limits_{n\to\infty} \mathsf{E}[M_T \mathbb{1}_{\{T>n\}}] = 0$.
由 $\lim\limits_{n\to\infty} \mathsf{E}[|M_n| \mathbb{1}_{\{T>n\}}] = 0$ 知,

$$\lim_{n\to\infty} \mathsf{E}[M_{n\wedge T} - M_T] = 0,$$

故 $\mathsf{E}[M_T] = \mathsf{E}[M_0]$. ■

利用上述定理，我们可以讨论一个这样的问题：设 $X_1, X_2, \cdots$, 是独立同分布的随机变量，且

$$\mathsf{P}(X_1 = 1) = \frac{1}{2},\ \mathsf{P}(X_1 = -1) = \frac{1}{2}.$$

令 $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$, $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$. 设 $\tau = \inf\{n : X_n = 1\}$. 则有 $\mathsf{E}[S_\tau] = 0$.

首先, 我们说明 $\mathsf{P}(\tau < +\infty) = 1$.

事实上,

$$\mathsf{P}(\tau = +\infty) = \mathsf{P}\left(\bigcap_{k=1}^{\infty} \{X_k = -1\}\right) = \lim_{N\to\infty} \prod_{k=1}^{N} \frac{1}{2} = 0.$$

其次,

$$\begin{aligned}
\mathsf{E}\big[|S_\tau|\big] &= \mathsf{E}\left[\left|\sum_{i=1}^{\tau} X_i\right|\right] = \mathsf{E}\left[\sum_{k=1}^{\infty} 1_{\{\tau=k\}} \left|\sum_{i=1}^{k} X_i\right|\right] \\
&= \sum_{k=1}^{\infty} \mathsf{E}\Big[1_{\{\tau=k\}} \big| -(k-1)+1 \big|\Big] \\
&\leqslant \sum_{k=1}^{\infty} k \cdot \mathsf{P}(\tau = k) = \sum_{k=1}^{\infty} \frac{k}{2^k} < \infty.
\end{aligned}$$

于是 $\mathsf{E}[|S_T|] < +\infty$. 进一步,

$$\begin{aligned}&\mathsf{E}\Big[|S_n|1_{\{\tau>n\}}\Big]\\ \leqslant\ &\mathsf{E}\Big[n\sum_{k=n+1}^{\infty}1_{\{\tau=k\}}\Big]=n\sum_{k=n+1}^{\infty}\frac{1}{2^k}=\frac{n}{2^n}\to 0.\end{aligned}$$

因此，$\mathsf{E}[S_\tau]=0$.

事实上，关于随机游动，我们还有如下进一步讨论.

例 2.1.9　设 $X_1, X_2, \cdots$, 是独立同分布的随机变量，且

$$\mathsf{P}(X_1=1)=\frac{1}{2},\ \mathsf{P}(X_1=-1)=\frac{1}{2}.$$

令 $S_0=0$, $S_n=\sum_{i=1}^n X_i$, $\mathcal{F}_0=\{\emptyset,\Omega\}$, $\mathcal{F}_n=\sigma(X_1,X_2,\cdots,X_n)$. 令

$$T=\inf\{n: S_n=-a \text{ 或 } S_n=b\},$$

这里 $a>0$, $b>0$. 由计算可知 $(S_n)_{n\geqslant 0}$ 是 $\{\mathcal{F}_n\}_{n\geqslant 0}-$ 鞅, T 是停时, 由 Doob 停止定理, $(S_{n\wedge T})_{n\geqslant 1}$ 是 $\mathcal{F}_n-$ 鞅, 因此

$$\mathsf{E}(S_{n\wedge T})=\mathsf{E}(S_{0\wedge T})=0.$$

由于 $(S_{n\wedge T})_{n\geqslant 1}$ 是 $\{\mathcal{F}_n\}_{n\geqslant 1}-$ 鞅, $(S_{n\wedge T}^2-n\wedge T)_{n\geqslant 1}$ 是 $\{\mathcal{F}_n\}_{n\geqslant 1}-$ 鞅. 因此

$$\mathsf{E}(S_{n\wedge T}^2-n\wedge T)=0.$$

由单调性,

$$\lim_{n\to\infty}\mathsf{E}(n\wedge T)=\mathsf{E}(T).$$

注意到 $|S_{n\wedge T}|\leqslant\max\{a,b\}$，因此 $\mathsf{E}(T)<\max\{a^2,b^2\}$, 于是 $\mathsf{P}(T<\infty)=1$.

由于 $|S_{n\wedge T}|\leqslant\max\{a,b\}$，注意到

$$\lim_{n\to\infty}S_{n\wedge T}=S_T.$$

因此

$$\lim_{n\to\infty}\mathsf{E}(S_{n\wedge T})=\mathsf{E}(S_T)=0.$$

即

$$\mathsf{P}(S_T=-a)(-a)+\mathsf{P}(S_T=b)b=0,$$

$$\mathsf{P}(S_T=-a)+\mathsf{P}(S_T=b)=1,$$

因此

$$\mathsf{P}(S_T=-a)=\frac{b}{a+b},\ \mathsf{P}(S_T=b)=\frac{a}{a+b}.$$

同时,

$$\begin{aligned}\lim_{n\to\infty}\mathsf{E}(S_{n\wedge T}^2) &= \mathsf{E}(S_T^2)\\ &= \mathsf{P}(S_T=-a)a^2+\mathsf{P}(S_T=b)b^2\\ &= ab.\end{aligned}$$

令

$$U=\inf\{n: S_n=-a\},$$

这里 $a>0$. 我们来讨论 U 的分布. 首先, 我们先说明

$$\mathsf{P}(U<\infty)=1.$$

令

$$H_b=\inf\{n: S_n=b\},$$

$-a<0<b$.

$$\mathsf{P}(U<\infty)=\lim_{b\to+\infty}\mathsf{P}(U<H_b)$$

取 $\tau=U\wedge H_b$. 由上面的讨论,

$$\lim_{b\to+\infty}\mathsf{P}(U<H_b)=\lim_{b\to+\infty}\mathsf{P}(S_\tau=-a)=\lim_{b\to+\infty}\frac{b}{a+b}=1.$$

首先研究 U 的母函数. 设 $r\in(0,1)$, 令 $Z_n=r^n\rho^{S_n}$.

$$\begin{aligned}\mathsf{E}(Z_n|\mathcal{F}_{n-1}) &= \mathsf{E}(r^n\rho^{S_n}|\mathcal{F}_{n-1})\\ &= r^n\rho^{S_{n-1}}\mathsf{E}(\rho^{X_n})\\ &= r^{n-1}\rho^{S_{n-1}}\frac{1}{2}(r\rho+\frac{r}{\rho})\end{aligned}$$

当 $\frac{1}{2}(r\rho+\frac{r}{\rho})=1$ 时, $(Z_n)_{n\geqslant1}$ 是 $\mathcal{F}_n-$ 鞅, 此时

$$\rho_1=\frac{1-\sqrt{1-r^2}}{r},\quad \rho_2=\frac{1+\sqrt{1-r^2}}{r}.$$

注意到 $0<\rho_1<1<\rho_2$, 因此

$$0\leqslant Z_{n\wedge U}\leqslant\rho_1^{-a}.$$

由控制收敛定理,

$$1=\mathsf{E}(Z_0)=\lim_{n\to\infty}\mathsf{E}(Z_{n\wedge U})=\mathsf{E}(Z_U).$$

因此

$$\mathsf{E}(r^U\rho_1^{-a})=1,$$

即

$$\mathsf{E}(r^U)=(\frac{1-\sqrt{1-r^2}}{r})^a.$$

这样, 便计算出了随机变量 U 的母函数, 可利用 $(\frac{1-\sqrt{1-r^2}}{r})^a$ 泰勒展开及

$$\mathsf{E}(r^U)=\sum_{k=0}^{\infty}\mathsf{P}(U=k)r^k$$

计算 $\mathsf{P}(U=k)$.

接下来考虑更一般的情况:

设 $X_1,X_2,\cdots,$ 是独立同分布的随机变量，且

$$\mathsf{P}(X_1=1)=p,\ \mathsf{P}(X_1=-1)=q.$$

$p+q=1$, 且 $\frac{1}{2}<p<1$. 令 $S_0=0$, $S_n=\sum_{i=1}^n X_i$, $\mathcal{F}_0=\{\emptyset,\Omega\}$, $\mathcal{F}_n=\sigma(X_1,X_2,\cdots,X_n)$. 令

$$T_x=\inf\{n:S_n=x\},$$

考虑 $x<0<y$. 下面讨论 $\mathsf{P}(T_x<T_y)$.

注意到 $((\frac{q}{p})^{S_n})_{n\geqslant 0}$ 是 $\mathcal{F}_n-$ 鞅, 令 $\tau=T_x\wedge T_y$, $((\frac{q}{p})^{S_{n\wedge\tau}})_{n\geqslant 0}$ 是一致有界鞅, 由控制收敛定理,

$$1=\mathsf{E}((\frac{q}{p})^{S_0})=\lim_{n\to\infty}\mathsf{E}((\frac{q}{p})^{S_{n\wedge\tau}})=\mathsf{E}((\frac{q}{p})^{S_\tau})$$

因此

$$\mathsf{P}(T_x<T_y)(\frac{q}{p})^x+\mathsf{P}(T_x>T_y)(\frac{q}{p})^y=1,$$

注意到

$$\mathsf{P}(T_x<T_y)+\mathsf{P}(T_x>T_y)=1.$$

因此

$$\mathsf{P}(T_x<T_y)=\frac{(\frac{q}{p})^y-1}{(\frac{q}{p})^y-(\frac{q}{p})^x},$$

$$\mathsf{P}(T_x > T_y) = \frac{1-(\frac{q}{p})^x}{(\frac{q}{p})^y - (\frac{q}{p})^x}.$$

下面计算 $\mathsf{P}(T_x < +\infty)$ 与 $\mathsf{P}(T_y < +\infty)$.

事实上, $T_x < +\infty$ 当且仅当 $T_x < T_y$ 对任意足够大的 $y > 0$ 成立,

$$\mathsf{P}(T_x < +\infty) = \lim_{y\to+\infty} \frac{(\frac{q}{p})^y - 1}{(\frac{q}{p})^y - (\frac{q}{p})^x} = (\frac{p}{q})^x.$$

类似,

$$\mathsf{P}(T_y < +\infty) = \lim_{x\to-\infty} \frac{1-(\frac{q}{p})^x}{(\frac{q}{p})^y - (\frac{q}{p})^x} = 1.$$

因此, 可以尝试计算 $\mathsf{E}(T_y)$.

由停止定理

$$\mathsf{E}[S_{T_y\wedge n} - (p-q)(T_y \wedge n)] = 0$$

由单调收敛性

$$\lim_{n\to\infty} \mathsf{E}[(p-q)(T_y\wedge n)] = (p-q)\mathsf{E}(T_y).$$

而

$$\inf_n S_n \leqslant S_{T_y\wedge n} \leqslant y,$$

$$\mathsf{P}(\inf_n S_n < x) = \mathsf{P}(T_x < +\infty) = (\frac{p}{q})^x,$$

因此

$$-\infty < \mathsf{E}(\inf_n S_n) \leqslant y.$$

由控制收敛定理,

$$\lim_{n\to\infty} \mathsf{E}[S_{T_y\wedge n}] = E[S_{T_y}] = y.$$

故

$$\mathsf{E}(T_y) = \frac{y}{p-q}.$$

关于停时, 事实上有所谓 T 前事件 $\sigma-$ 域的概念. 下面给出 T 前事件 $\sigma-$ 域的定义，并给出一些应用.

定义 2.1.5 设 T 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的停时, 称

$$\mathcal{F}_T = \{A \in \mathcal{F} : \forall n \geqslant 0, \{T \leqslant n\} \bigcap A \in \mathcal{F}_n\}$$

为 T 前事件 $\sigma-$ 域.

定理 2.1.4 设 $(X_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的适应过程, ξ 是随机变量, T 为停时, 令 $X_\infty = \xi$, $X_T(\omega) = X_{T(\omega)}(\omega)$, 则 X_T 是 $\mathcal{F}_T$-可测的.

证明: 对于 Borel 集 B, 由于

$$\{X_T \in B\} = \bigcup_{k\in\mathcal{N}} \{X_k \in B\} \cap \{T = k\} \in \mathcal{F},$$

$$\{X_T \in B\} \cap \{T = n\} = \{X_n \in B\} \cap \{T = n\} \in \mathcal{F}_n,$$

所以 $\{X_T \in B\} \in \mathcal{F}_T$. 故 X_T 是 $\mathcal{F}_T$-可测的. ■

定理 2.1.5 设 T 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的停时, 且 $\mathsf{P}(T < \infty) = 1$, 则

$$\mathcal{F}_T = \sigma(X_T : X\text{是}(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})\text{上的适应过程}).$$

证明: 由定理 2.1.4, 令 $\mathcal{G} = \sigma(X_T : X\text{是适应序列})$, 知 $\mathcal{G} \subset \mathcal{F}_T$.
若 $A \in \mathcal{F}_T$, 令

$$X_n = \mathbb{1}_A \mathbb{1}_{\{T\leqslant n\}},$$

由于 $A \in \mathcal{F}_T$, $\{T \leqslant n\} \in \mathcal{F}_n$, 故

$$A \cap \{T \leqslant n\} \in \mathcal{F}_n.$$

从而 X_n 是适应的. 注意到 $X_T = \mathbb{1}_A$, 故 $A \in \mathcal{G}$. 于是, $\mathcal{F}_T = \mathcal{G}$. ■

本节的最后, 列出关于停时的两个重要结论.

定理 2.1.6 设 S 和 T 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的有界停时, 且 $S \leqslant T$, X 是一个鞅, 则 $\mathsf{E}[X_T|\mathcal{F}_S] = X_S$.

证明: X 是鞅, 令

$$H_n = \mathbb{1}_{[n\leqslant T]} - \mathbb{1}_{[n\leqslant S]}.$$

设 $S \leqslant T \leqslant M$, 当 $n > M$ 时,

$$\begin{aligned}(H\cdot X)_n &= \sum_{i=1}^{n} (X_{T\wedge i} - X_{T\wedge(i-1)}) - \sum_{i=1}^{n} (X_{S\wedge i} - X_{S\wedge(i-1)}) \\ &= X_T - X_S,\end{aligned}$$

且 $(H\cdot X)$ 是鞅, 故 $\mathsf{E}[(H\cdot X)_n] = \mathsf{E}[(H\cdot X)_0] = 0$, 即 $\mathsf{E}(X_T) = \mathsf{E}(X_S)$.

对于 $B \in \mathcal{F}_S$, 取 $S^B = S\mathbb{1}_B + M\mathbb{1}_{B^c}$, $T^B = T\mathbb{1}_B + M\mathbb{1}_{B^c}$ 代入上式.

$$\mathsf{E}[X_T\mathbb{1}_B + X_M\mathbb{1}_{B^c}] = \mathsf{E}[X_S\mathbb{1}_B + X_M\mathbb{1}_{B^c}],$$

故 $\mathsf{E}[X_T|\mathcal{F}_S] = X_S$.

■

下面的定理提供了一个刻画离散时间鞅的充分必要条件.

定理 2.1.7 若 $X = (X_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的适应过程, 且对于任意的 $n \geqslant 0$, $\mathsf{E}[|X_n|] < +\infty$, 则 X 是鞅当且仅当对于任意有界停时 S 和 T, $\mathsf{E}[X_T] = \mathsf{E}[X_S]$.

证明: 只需证明充分性, 若对于任意有界停时 S 和 T, 均有 $\mathsf{E}[X_T] = \mathsf{E}[X_S]$. 设 $S \leqslant T \leqslant M$, $l < k$ 均为小于 M 正整数. 对于 $B \in \mathcal{F}_l$, 令 $T = k\mathbb{1}_B + M\mathbb{1}_{B^c}$, $S = l\mathbb{1}_B + M\mathbb{1}_{B^c}$, 故有 $\mathsf{E}[X_l|\mathcal{F}_k] = X_k$. 故 X 是一个鞅.

■

2.2 Doob 不等式与鞅的收敛定理

鞅的收敛定理在研究随机过程中起了十分重要的作用. Doob 不等式及上穿不等式在研究鞅的收敛定理中起了至关重要的作用. 我们首先介绍著名的 Doob 不等式.

定理 2.2.1 设 $M = (M_n)_{0\leqslant n\leqslant N}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的鞅, $p > 1$, 对于任意 $0 \leqslant n \leqslant N$, $\mathsf{E}[|M_n|^p] < \infty$, 那么对 $x > 0$,

$$\mathsf{P}(\max_{0\leqslant n\leqslant N} |M_n| \geqslant x) \leqslant \frac{\mathsf{E}[|M_N|^p]}{x^p},$$

$$\mathsf{E}[\max_{0\leqslant n\leqslant N} |M_n|^p] \leqslant (\frac{p}{p-1})^p \mathsf{E}[|M_N|^p].$$

证明: 若集合 $\{0 \leqslant n \leqslant N : M_n \geqslant x\}$ 非空, 定义 $T = \inf\{0 \leqslant n \leqslant N : M_n \geqslant x\}$. 若集合 $\{0 \leqslant n \leqslant N : M_n \geqslant x\}$ 为空集, 定义 $T = N$. T 是一个停时. 记 $M^* = \max\limits_{0\leqslant n\leqslant N} |M_n|$, 由停止定理, 由于 $|M|^p$ 是下鞅,

$$\begin{aligned}
\mathsf{E}[|M_N|^p] &\geqslant \mathsf{E}[|M_T|^p] = \mathsf{E}[|M_T|^p \mathbb{1}_{\{M^*\geqslant x\}}] + \mathsf{E}[|M_T|^p \mathbb{1}_{\{M^*<x\}}] \\
&\geqslant x^p \mathsf{P}(\max_{0\leqslant n\leqslant N} |M_n| \geqslant x) + \mathsf{E}[|M_N|^p \mathbb{1}_{\{M^*<x\}}].
\end{aligned}$$

于是

$$\mathsf{P}(\max_{0\leqslant n\leqslant N} |M_n| \geqslant x) \leqslant \frac{\mathsf{E}[|M_N|^p \mathbb{1}_{\{M^*>x\}}]}{x^p} \leqslant \frac{\mathsf{E}[|M_N|^p]}{x^p}.$$

对于常数 $\eta > 0$,

$$\begin{aligned}\mathsf{E}[|M^*\wedge\eta|^p] &= \mathsf{E}[\int_0^{M^*\wedge\eta} px^{p-1}\,\mathrm{d}x] = \mathsf{E}[\int_0^{\eta} px^{p-1}\mathbb{1}_{\{M^*\geqslant x\}}\,\mathrm{d}x] \\ &= \int_0^{\eta} px^{p-1}\mathsf{P}(M^*\geqslant x)\,\mathrm{d}x \leqslant \int_0^{\eta} px^{p-2}\mathsf{E}[|M_N|\mathbb{1}_{\{M^*\geqslant x\}}]\,\mathrm{d}x \\ &= p\mathsf{E}[|M_N|\int_0^{M^*\wedge\eta} x^{p-2}\,\mathrm{d}x] = \frac{p}{p-1}\mathsf{E}[|M_N|(M^*\wedge\eta)^{p-1}].\end{aligned}$$

由 Hölder 不等式

$$\mathsf{E}[|M^*\wedge\eta|^p] \leqslant \frac{p}{p-1}\mathsf{E}[|M^*\wedge\eta|^p]^{\frac{p-1}{p}}\mathsf{E}[|M_N|^p]^{\frac{1}{p}}.$$

令 $\eta\to\infty$, 有

$$\mathsf{E}[\max_{0\leqslant n\leqslant N}|M_n|^p] \leqslant (\frac{p}{p-1})^p\mathsf{E}[|M_N|^p].$$

■

Doob 不等式在证明鞅的强大数律时会起到重要作用, 此外我们在第 4 章也会用到该不等式. 这一节中, 我们主要讨论鞅的收敛定理, 要用到下面的上穿不等式.

定理 2.2.2 (上穿不等式) 设 $M=(M_n)_{n\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geqslant 0},\mathsf{P})$ 上的鞅, 且存在常数 C, 使 $\mathsf{E}[|M_n|]\leqslant C$. 假设 a,b $(a<b)$ 是实数, 令 U_n 是从 0 到 n 时刻, $M=(M_n)_{n\geqslant 0}$ 从 a 下上穿至 b 上的次数, 则

$$\mathsf{E}[U_n] \leqslant \frac{|a|+C}{b-a}.$$

证明: 定义:

$$S_1=\inf\{n: M_n\leqslant a\},\ T_1=\inf\{n>S_1: M_n\geqslant b\}.$$

当 $j>1$ 时

$$S_j=\inf\{n>T_{j-1}: M_n\leqslant a\},\ T_j=\inf\{n>S_j: M_n\geqslant b\}.$$

另外, 定义

$$W_n=\sum_{k=1}^{n} B_k(M_k-M_{k-1}).$$

这里,

$$B_n = 0, \quad 若 \quad n-1 < S_1,$$
$$B_n = 1, \quad 若 \quad S_j \leqslant n-1 < T_j,$$
$$B_n = 0, \quad 若 \quad T_j \leqslant n-1 < S_{j+1}.$$

观察到当 $T_j < n \leqslant T_{j+1}$ 时, $U_n = j$. 不难得到

$$W_n \geqslant U_n(b-a) - |M_n - a|.$$

由于 W_n 是鞅, $\mathsf{E}[W_n] = \mathsf{E}[W_0] = 0$. 故

$$\mathsf{E}[U_n(b-a)] \leqslant \mathsf{E}[|M_n - a|],$$

即 $\mathsf{E}[U_n] \leqslant \dfrac{|a|+C}{b-a}$. ■

利用上穿不等式, 得到鞅的收敛定理.

定理 2.2.3　给定 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的鞅 $M = (M_n)_{n\geqslant 0}$, 若对任意 n, $\mathsf{E}[|M_n|] \leqslant C$, $C > 0$, 则存在 M_∞, 使

$$\lim_{n\to\infty} M_n = M_\infty \quad a.s..$$

证明: 设 $M^* = \limsup\limits_{n\to\infty} M_n$, $M_* = \liminf\limits_{n\to\infty} M_n$,

$$\{M^* > M_*\} = \bigcup_{a,b\in\mathsf{Q}} \{M_* < a < b < M^*\}.$$

下证 $\mathsf{P}(\{M_* < a < b < M^*\}) = 0$.

由于 $\{M_* < a < b < M^*\} \subset \{U_\infty = +\infty\}$, 且

$$\mathsf{E}[U_\infty] \leqslant \frac{|a|+C}{b-a} < +\infty,$$

故 $U_\infty < +\infty \quad a.s.$ 从而 $\mathsf{P}(\{M_* < a < b < M^*\}) = 0$. 因此 $\mathsf{P}(\{M_* < M^*\}) = 0$, 即 $\lim\limits_{n\to\infty} M_n = M_\infty \quad a.s..$ ■

下面给出一个十分重要的定理, 该定理刻画了一致可积鞅的形式.

定理 2.2.4　若 $M = (M_n)_{n\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geqslant 0}, \mathsf{P})$ 上的一致可积鞅, 则

$$\lim_{n\to\infty} E[|M_n - M_\infty|] = 0,$$

且 $M_n = \mathsf{E}[M_\infty|\mathcal{F}_n]$.

证明: 由于 $\{M_n\}$ 是一致可积的. 由定理 2.2.3知, $M=(M_n)_{n\geqslant 0}$ 是 L^1 收敛至 M_∞. $m>n$ 时, $\mathsf{E}[M_m|\mathcal{F}_n]=M_n$, 对于 $B\in\mathcal{F}_n$,

$$\int_B \mathsf{E}[M_m|\mathcal{F}_n]\,\mathrm{d}\mathsf{P}=\int_B M_m\,\mathrm{d}\mathsf{P}.$$

由于

$$\lim_{n\to\infty} E[|M_n-M_\infty|]=0,$$

故

$$\int_B M_m\,\mathrm{d}\mathsf{P}\longrightarrow\int_B M_\infty\,\mathrm{d}\mathsf{P}.$$

而由条件数学期望,

$$\int_B M_\infty\,\mathrm{d}\mathsf{P}=\int_B \mathsf{E}[M_\infty|\mathcal{F}_n]\,\mathrm{d}\mathsf{P}.$$

从而

$$\lim_{m\to\infty}\int_B \mathsf{E}[M_m|\mathcal{F}_n]\,\mathrm{d}\mathsf{P}=\int_B \mathsf{E}[M_\infty|\mathcal{F}_n]\,\mathrm{d}\mathsf{P}.$$

由于 $\forall m>n,\ \mathsf{E}[M_m|\mathcal{F}_n]=M_n$, 故

$$\int_B M_n\,\mathrm{d}\mathsf{P}=\int_B \mathsf{E}[M_\infty|\mathcal{F}_n]\,\mathrm{d}\mathsf{P}.$$

因此 $\mathsf{E}[M_\infty|\mathcal{F}_n]=M_n$. ■

利用上述一致可积鞅的结论, 我们可以得到停止定理的加强版本, 证明留给读者.

定理 2.2.5 若 $M=(M_n)_{n\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n\geqslant 0},\mathsf{P})$ 上的一致可积鞅, S 和 T 是两个停时, 且 $S\leqslant T$, 则 M_T 和 M_S 是可积的, 且 $\mathsf{E}[M_T|\mathcal{F}_S]=M_S$.

下面的例子生动刻画了上述收敛定理的作用.

例 2.2.1 Pólya 罐子模型.

假设有一个罐子, 罐子中有红球和白球. $n=0$ 时, 有一个红球, 一个白球. 设 R_n 为 n 时刻红球的个数, G_n 为 n 时刻白球的个数, 即

$$R_0=1,G_0=1.$$

在罐子中随机摸一个球, 看颜色, 将球放回, 并将一个相同颜色的球同时放入. 即有

$$R_n+G_n=2+n.$$

令 $M_n = \dfrac{R_n}{R_n + G_n} = \dfrac{R_n}{2+n}$ 为红球比例. 令

$$\mathcal{F}_n = \sigma(M_1, M_2, \cdots, M_n) = \sigma(R_1, R_2, \cdots, R_n).$$

事实上, 这里不难看出有 $\mathsf{E}[R_{n+1}|\mathcal{F}_n] = \mathsf{E}[R_{n+1}|R_n]$.

下证, $(M_n)_{n\geqslant 0}$ 是一个鞅.

由于 $\mathsf{E}[M_{n+1}|\mathcal{F}_n] = \mathsf{E}[M_{n+1}|M_n]$,

$$\mathsf{E}[M_{n+1}|M_n] = \mathsf{E}[\frac{R_{n+1}}{n+3}|R_n],$$

$$\mathsf{P}(R_{n+1} = k+1|R_n = k) = \frac{k}{n+2},$$
$$\mathsf{P}(R_{n+1} = k|R_n = k) = \frac{n+2-k}{n+2},$$

故

$$\begin{aligned}\mathsf{E}[R_{n+1}|R_n] &= (\frac{R_n}{n+2}) \cdot (R_n + 1) + (1 - \frac{R_n}{n+2}) \cdot R_n \\ &= R_n \cdot \frac{R_n}{n+2} + \frac{R_n}{n+2} + R_n - \frac{R_n}{n+2} \cdot R_n \\ &= R_n + \frac{R_n}{n+2}.\end{aligned}$$

于是

$$\mathsf{E}[M_{n+1}|M_n] = \frac{1}{n+3}\mathsf{E}[R_{n+1}|R_n] = \frac{1}{n+3}(R_n + \frac{R_n}{n+2}) = \frac{R_n}{n+2} = M_n.$$

$(M_n)_{n\geqslant 0}$ 是鞅. 注意到

$$\mathsf{E}[|M_n|] = \mathsf{E}[M_n] = \mathsf{E}[M_0] = \frac{1}{2},$$

因此, 存在 M_∞, 使 $\lim\limits_{n\to\infty} M_n = M_\infty \quad a.s.$.

下面, 我们来考虑 M_∞ 的分布. 首先考虑 M_n 的分布.

当 $n=1$ 时, $M_1 = \frac{R_1}{3}$, $R_1 = 1$ 的概率为 $\frac{1}{2}$, $R_1 = 2$ 的概率为 $\frac{1}{2}$. 故

$$\mathsf{P}(M_1 = \frac{1}{3}) = \frac{1}{2},\ \mathsf{P}(M_1 = \frac{2}{3}) = \frac{1}{2}.$$

当 $n=2$ 时, $M_2 = \frac{R_2}{4}$,

$R_2 = 1$ 的概率为 $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$,

$R_2 = 2$ 的概率为 $\frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{3}$,

$R_2 = 3$ 的概率为 $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$, 故

$$\mathsf{P}(M_2 = \frac{1}{4}) = \frac{1}{3},\ \mathsf{P}(M_2 = \frac{2}{4}) = \frac{1}{3},\ \mathsf{P}(M_2 = \frac{3}{4}) = \frac{1}{3}.$$

设 $n = k$ 时, M_k 服从 $\{\dfrac{1}{k+2}, \dfrac{2}{k+2}, \cdots, \dfrac{k+1}{k+2}\}$ 上的均匀分布.
当 $n = k+1$ 时,
$R_{k+1} = 1$ 的概率 $\frac{1}{k+1} \cdot \frac{k+1}{k+2} = \frac{1}{k+2}$,
$R_{k+1} = 2$ 的概率 $\frac{1}{k+1} \cdot (1 - \frac{k+1}{k+2}) + \frac{1}{k+1} \cdot \frac{k}{k+2} = \frac{1}{k+2}$,
$R_{k+1} = 3$ 的概率 $\frac{1}{k+1} \cdot \frac{2}{k+2} + \frac{1}{k+1} \cdot \frac{k-1}{k+2} = \frac{k+1}{(k+1)(k+2)} = \frac{1}{k+2}$,
$R_{k+1} = 4$ 的概率 $\frac{1}{k+1} \cdot \frac{3}{k+2} + \frac{1}{k+1} \cdot \frac{k-2}{k+2} = \frac{1}{k+2}$,
故 $n = k+1$ 时, M_{k+1} 服从 $\{\dfrac{1}{k+3}, \dfrac{2}{k+3}, \cdots, \dfrac{k+2}{k+3}\}$ 上的均匀分布.

2.3　习题

1. 设 $X_1, X_2, \cdots,$ 是独立同分布的随机变量, 且

$$\mathsf{P}(X_1 = 1) = 1 - \mathsf{P}(X_1 = -1) = q.$$

这里 $\frac{1}{2} < q < 1$, 令 $S_0 = 0$, $S_n = \sum_{i=1}^{n} X_i$, $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$.

(1) 请问 S_n 是 $\{\mathcal{F}_n\}-$ 鞅吗?

(2) 求 r, 使 $M_n = S_n - rn$ 是 $\{\mathcal{F}_n\}-$ 鞅.

(3) 令 $\theta = (1-q)/q$, $M_n = \theta^{S_n}$. 证明: $M = (M_n)_{n \geqslant 0}$ 是 $\{\mathcal{F}_n\}-$ 鞅.

(4) 令 a, b 是个正整数, 令

$$T_{a,b} = \min\{j : S_j = b, \quad \text{或} \quad S_j = -a\},$$

求 $\mathsf{P}(S_{T_{a,b}} = b)$.

(5) 求 $\mathsf{P}(T_{a,\infty} < \infty)$

2. 设 $\{\xi_n\}$ 是一个独立同分布的序列.

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

令 $X_0 = 0,\ X_n = \sum_{i=1}^{n} \xi_i$, 对于 $a > 0$, 记 $T_a = \inf\{n > 0 : X_n = a\}$, 请证明 $p > q$ 时, $\mathsf{P}(T_a < \infty) = 1$.

3. 设 $X_1, X_2, \cdots$, 是独立同分布的随机变量, 且

$$\mathsf{P}(X_1 = 1) = \mathsf{P}(X_1 = -1) = \frac{1}{2}.$$

令 $S_0 = 0$, $S_n = \sum_{i=1}^{n} X_i$, $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$,

$$T = \min\{n : |S_n| = K\},$$

这里 $K > 0$.

(1) 证明: $\mathsf{P}(T \leqslant j + K | T > j) \geqslant 2^{-K}$.

(2) 证明存在 $c < \infty$, $\alpha > 0$, 使得对于所有的 j, $\mathsf{P}(T > j) \leqslant ce^{-\alpha j}$, 且对任意 $r > 0$, $E[T^r] < \infty$.

(3) 令 $M_n = S_n^2 - n$, 证明存在 $C < \infty$, 使得对于所有的 n,

$$\mathsf{E}[M_{T\wedge n}^2] \leqslant C.$$

4. 设 $X_1, X_2, \cdots$, 是独立的随机变量, 对于任意的 j, $\mathsf{E}[X_j] = 0$, $Var[X_j] = \sigma_j^2$, 令 $S_0 = 0$, $S_n = \sum_{i=1}^{n} X_i$, $\mathcal{F}_n = \sigma(X_1, X_2, \cdots, X_n)$, 且 $\sum_{j=1}^{\infty} \sigma_j^2 < \infty$.

(1) 证明以概率 1 有

$$\lim_{n\to\infty} S_n = S_\infty.$$

(2) 求 $\mathsf{E}[S_\infty], Var[S_\infty]$.

5. 设 $X_1, X_2, \cdots$, 是 $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geqslant 1}, \mathsf{P})$ 上的一列随机变量, X_n 关于 $\mathcal{F}_n$ 可测, 且 $\mathsf{E}[X_n|\mathcal{F}_{n-1}] = 0$. 令 $S_n = \sum_{i=1}^{n} X_i$, $S_0 = 0$. T 是一个停时, 且 $\mathsf{E}[T] < +\infty$, 在 $\{T \geqslant n+1\}$ 上,

$$\mathsf{E}[|X_{n+1}||\mathcal{F}_n] \leqslant C,$$

请证明 $\mathsf{E}[|S_T|] < \infty$, 且 $\mathsf{E}[S_T] = 0$.

第 3 章　连续鞅与布朗运动

从这一章开始, 我们开始进入本书的主要部分：连续时间随机过程.

3.1　流与停时, 连续时间鞅

我们所有讨论的对象均定义在带流概率空间 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathsf{P})$ 上, 这里 $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的流, 即 $\mathcal{F}_t$ 是 $\mathcal{F}$ 的子 $\sigma-$ 代数. 不仅如此, 我们还要求当 $s < t$ 时,$\mathcal{F}_s \subseteq \mathcal{F}_t, \mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$[①], 且 $(\Omega, \mathcal{F}, \mathsf{P})$ 是完备的, 即对于任意的 $t \in \mathbb{R}_+, \mathcal{F}_t$ 均包含所有 $\mathsf{P}-$ 零测集.

此外, $\Omega\times\mathbb{R}_+$ 的子集, 我们称之为随机集. 我们考虑的随机过程 $X = (X_t)_{t\in\mathbb{R}_+}$, 即 $\Omega \times \mathbb{R}_+$ 到 $\mathbb{R}$ 上的一个映射. 不仅如此, 对于随机过程 $X = (X_t)_{t\in\mathbb{R}_+}$, 对于每个 $t \geqslant 0$, X_t 都是随机变量.

此处, 列出几个定义, 在读者学习和研究随机过程与随机分析的过程中会用到.

定义 3.1.1　设 X 是一个随机过程,T 是 $\Omega \to \overline{\mathbb{R}}_+$ 的一个映射, 定义 $X_t^T = X_{t\wedge T}$ 为 X 的停止过程 (process stopped at time T).

定义 3.1.2　设 A 是一个随机集, 若 $\mathsf{P}(\{\omega : \exists t \in \mathbb{R}_+, (\omega, t) \in A\}) = 0$, 则称 A 是不足道的 (evanescent).

定义 3.1.3　若 X 与 Y 是两个随机过程, 随机集 $\{X \neq Y\} = \{(\omega, t) : X_t(\omega) \neq Y_t(\omega)\}$ 是不足道的, 则称 X 与 Y 是不可区分的 (indistinguishable).

定义 3.1.4　若 X 与 Y 是两个随机过程, 若对任意 $t \in \mathbb{R}_+, \mathsf{P}(\{\omega : X_t \neq Y_t\}) = 0$, 则称 X 与 Y 互为修正 (modification).

通常, 我们考虑的随机过程 $X = (X_t)_{t\in\mathbb{R}_+}$ 是所谓的右连左极过程 (càdlág) 即对于固定的 ω 而言,$(X_t(\cdot))_{t\in\mathbb{R}_+}$ 作为 t 的函数是右连续且左极限存在的. 对于固定的 ω 而言, 我们往往称 $(X_t(\cdot))_{t\in\mathbb{R}_+}$ 是 X 的一条轨道. 我们记 $X_{t-} = \lim_{s\uparrow t} X_s, \Delta X_t = X_t - X_{t-}$. 若对所有的 t, ω, $\Delta X_t = 0$, 则称 X 轨道连续, 或者称 X 为连续过程.

① 此条件通常被称为流是右连续的

类似与离散时间的情形, 连续时间框架下, 也有停时等概念.

定义 3.1.5　给定带流概率空间 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathsf{P})$, T 是 $\Omega \to \bar{\mathbb{R}}_+$ 上的一个映射, 若 T 满足, 对任意 t, $\{T \leqslant t\} \in \mathcal{F}_t$, 称 T 为停时.

定义 3.1.6　$X = (X_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathsf{P})$, 上的随机过程, 若对任意 $t \in \mathbb{R}_+, X_t$ 是 $\mathcal{F}_t-$ 可测的, 则称 X 是适应的.

除此之外, 我们还需要如下概念:

定义 3.1.7　$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathsf{P})$ 上的适应过程 X, 若对任意 t, $0 \leqslant s \leqslant t$ 时, $X(s,\omega)$ 可看作 $[0,t]\times\Omega \to \mathbb{R}$ 上的映射, 且关于 $\mathcal{B}[0,t]\otimes\mathcal{F}_t$ 可测, 则称 X 是循序可测过程.

类似于离散时间框架的情形, 有如下定义:

定义 3.1.8　若 T 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathsf{P})$ 上的停时, 称

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leqslant t\} \in \mathcal{F}_t, \quad \forall t \geqslant 0\}$$

为 T 前事件 $\sigma-$ 域.

下面的命题给出了 T 前事件 $\sigma-$ 域的表示.

命题 3.1.1　若 T 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}_+}, \mathsf{P})$ 上的停时, $\mathsf{P}(T < \infty) = 1$,

$$\mathcal{F}_T = \sigma(X_T : X \text{ 是适应的右连左极过程}).$$

证明: 令

$$\mathcal{G} = \sigma(X_T : X \text{ 是适应的右连左极过程}).$$

设 T 是一个停时, 且 $T<\infty$, 定义 T_n 如下:

$$\text{在 } \{\frac{k-1}{2^n} \leqslant T < \frac{k}{2^n}\} \text{ 上, 令 } T_n = \frac{k}{2^n},$$

显然 T_n 是停时, 且 $T_n \downarrow T$. 对于 Borel 集 B

$$\{X_{T_n} \in B\}\bigcap\{T_n \leqslant t\} = \bigcup_{k\in N, \frac{k}{2^n}\leqslant t} [\{X_{\frac{k}{2^n}} \in B\}\bigcap\{T_n = \frac{k}{2^n}\}].$$

显然, 上面的集合在 $\mathcal{F}_t$ 中, 故 X_{T_n} 是 $\mathcal{F}_{T_n}-$ 可测的. 由于 X 是右连续的, 故 X_{T_n} 收敛至 X_T . 由于 $\bigcap_{n=1}^{\infty}\mathcal{F}_{T_n} = \mathcal{F}_T$ 故 X_T 是 $\mathcal{F}_T-$ 可测的. 故 $\mathcal{G} \subset \mathcal{F}_T$.

即 $\mathcal{G} = \mathcal{F}_T$. ■

在连续时间框架下, 鞅过程定义如下:

定义 3.1.9　设 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathsf{P})$ 是带流概率空间, $X = (X_t)_{t\in\mathbb{R}_+}$ 是其上的适应过程, 且是右连左极的. 若对任意 t, $\mathsf{E}[|X_t|] < +\infty$, 且对于 $s < t$, $\mathsf{E}[X_t|\mathcal{F}_s] = X_s \ \ a.s.$, 则称 X 是鞅.

通常, 关于鞅的定义中并没有右连左极的假设, 但是可以证明在流是右连续的前提下, 连续时间鞅存在右连左极的修正. 证明参见 Revuz D. 和 Yor M. [12]. 为讨论方便, 这里直接假设所有的连续时间鞅是右连左极的.

事实上, 在连续时间框架, 人们往往首先从一致可积鞅出发研究鞅的性质. 下面, 给出一致可积鞅的定义.

定义 3.1.10 给定 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in\mathbb{R}_+},\mathsf{P})$ 上的鞅 X, 若 $(X_t)_{t\geqslant 0}$ 是一致可积的, 则称 X 是一致可积鞅.

例 3.1.1 ξ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in\mathbb{R}_+},\mathsf{P})$ 上的可积随机变量, 令 $X_t=\mathsf{E}[\xi|\mathcal{F}_t]$, $X=(X_t)_{t\in\mathbb{R}_+}$ 是一致可积鞅.

对于大部分研究而言, 由于一致可积条件很难达到, 研究对象往往是更广的一类过程-局部鞅.

定义 3.1.11 若存在一列单调递增的停时 $(T_n)_{n\geqslant 0}$, $\lim\limits_{n\to\infty} T_n=+\infty$, 使得对于适应过程 X, $X^{T_n}\mathbb{1}_{\{T_n>0\}}$ 是一致可积鞅, 则称 X 是局部鞅.

例 3.1.2 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in\mathbb{R}_+},\mathsf{P})$ 上, 考虑 ξ 是 $\mathcal{F}_1-$ 可测的, 但是 $\mathsf{E}[|\xi|]=\infty$, $\mathsf{P}(\eta=1)=\mathsf{P}(\eta=-1)=\frac{1}{2}$, η 是 $\mathcal{F}_2-$ 可测的, 但是与 $\vee_{s<2}\mathcal{F}_s$ 独立. 令 $M_t=\mathbb{1}_{\{t\geqslant 2\}}\eta\xi$. 令

$$T_n=\begin{cases} n, & \xi\leqslant n,\\ 1, & \xi>n.\end{cases}$$

可知 $T_n\uparrow\infty$, M^{T_n} 是鞅, 然而 $\mathsf{E}[|M_2|]=\mathsf{E}[|\xi|]=\infty$, 因此, $(M_t)_{t\geqslant 0}$ 是局部鞅, 但不是鞅.

这里, 需要注意几点:

(1). 鞅往往是局部鞅, 而局部鞅不一定是鞅, 人们往往经常需要讨论, 在何时, 某种局部鞅是鞅;

(2). 鞅在有限闭区间上是一致可积的;

(3). 非负局部鞅是上鞅.

(4). 令 $Y_t=X_t-X_0$, 局部鞅的定义中, $X^{T_n}\mathbb{1}_{\{T_n>0\}}$ 也可替换为 Y^{T_n}.

一致可积鞅是我们讨论的重点, 有如下定理.

定理 3.1.1 若 X 是适应的右连左极过程, 且 X_t 几乎处处收敛于 X_∞, 若对于停时 T, X_T 可积且 $\mathsf{E}(X_T)=\mathsf{E}(X_\infty)$, 则 X 是一致可积鞅.

证明: 若 X_t 几乎处处收敛于 X_∞, X_∞ 可积, 对于 $t\in\mathbb{R}_+$, $A\in\mathcal{F}_t$, 定义

$$T=\begin{cases} t, & w\in A,\\ \infty, & w\notin A,\end{cases}$$

则

$$\mathsf{E}[X_T]=\mathsf{E}[X_t\mathbb{1}_A]+\mathsf{E}[X_\infty\mathbb{1}_{A^c}],\mathsf{E}[X_\infty]=\mathsf{E}[X_\infty\mathbb{1}_A]+\mathsf{E}[X_\infty\mathbb{1}_{A^c}].$$

由假设 $\mathsf{E}[X_T]=\mathsf{E}[X_\infty]$, 则 $\mathsf{E}[X_t\mathbb{1}_A]=\mathsf{E}[X_\infty\mathbb{1}_A]$, 故 $X_t=\mathsf{E}[X_\infty|\mathcal{F}_t]$. 显然, X 是一致可积鞅. ■

在连续时间情形, 也有停止定理.

定理 3.1.2 若 X 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续的一致可积鞅, 若 S, T 是两个停时, 且 $S\leqslant T$, 则 X_S 和 X_T 可积, 且

$$\mathsf{E}[X_T|\mathcal{F}_S]=X_S \quad a.s..$$

证明 [6]: 对于非负整数 n, 定义

$$T_n=\sum_{k=0}^{+\infty}\frac{k+1}{2^n}\mathbb{1}_{\{k2^{-n}<T\leqslant(k+1)2^{-n}\}}+\infty\mathbb{1}_{\{T=\infty\}},$$

$$S_n=\sum_{k=0}^{+\infty}\frac{k+1}{2^n}\mathbb{1}_{\{k2^{-n}<S\leqslant(k+1)2^{-n}\}}+\infty\mathbb{1}_{\{S=\infty\}},$$

显然 T_n, S_n 单调递减收敛至 T 和 S, 且对 $n\geqslant 0$, $S_n\leqslant T_n$.

固定 n, 令 $\mathcal{H}_k^{(n)}=\mathcal{F}_{\frac{k}{2^n}}$, $M_k^{(n)}=X_{\frac{k}{2^n}}$. 这里 $(M_k^{(n)})_{k\geqslant 0}$ 是关于流 $(\mathcal{H}_k^{(n)})_{k\geqslant 0}$ 的一致可积鞅. 由定理 2.2.5, X_{S_n}, X_{T_n} 可积, 且

$$X_{S_n}=M_{2^nS_n}^{(n)}=\mathsf{E}[M_{2^nT_n}^{(n)}|\mathcal{H}_{2^nS_n}^{(n)}]=\mathsf{E}[X_{T_n}|\mathcal{F}_{S_n}].$$

对于 $A\in\mathcal{F}_S$, 由于 $S\leqslant S_n$, 故 $A\in\mathcal{F}_{S_n}$, 于是

$$\mathsf{E}[X_{S_n}\mathbb{1}_A]=\mathsf{E}[X_{T_n}\mathbb{1}_A].$$

由于 X 是连续的一致可积鞅, 不难证明 X_S 和 X_T 可积, 且

$$\mathsf{E}[X_S\mathbb{1}_A]=\mathsf{E}[X_T\mathbb{1}_A].$$

即

$$\mathsf{E}[X_T|\mathcal{F}_S]=X_S \quad a.s..$$

■

进一步, 我们有

定理 3.1.3 若 X 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续鞅, 若 T 是停时, 则 X^T 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅.

证明: 我们下面证明对于 $t > s$,

$$\mathsf{E}[X_t^T|\mathcal{F}_s] = X_s^T.$$

考虑 $T_1 = T \wedge t$, $T_2 = T \wedge s$, T_1, T_2 均为停时, 且 $T_1 \geqslant T_2$. 我们在 $[0, t]$ 上讨论 X, 此时 X 具备一致可积性.

$$\begin{aligned}
\mathsf{E}[X_t^T|\mathcal{F}_s] &= \mathsf{E}[X_t^T 1_{T\leqslant s} + X_t^T 1_{T>s}|\mathcal{F}_s] \\
&= \mathsf{E}[X_T \mathbb{1}_{T\leqslant s} + X_t^T \mathbb{1}_{T>s}|\mathcal{F}_s] \\
&= \mathsf{E}[X_s^T \mathbb{1}_{T\leqslant s} + X_t^T \mathbb{1}_{T>s}|\mathcal{F}_s].
\end{aligned}$$

注意到 $X_s^T \mathbb{1}_{T\leqslant s} \in \mathcal{F}_s$, 故

$$\mathsf{E}[X_s^T \mathbb{1}_{T\leqslant s}|\mathcal{F}_s] = X_s^T \mathbb{1}_{T\leqslant s}.$$

考虑 $A \in \mathcal{F}_s$,

$$A \cap \{T > s\} \in \mathcal{F}_s \cap \mathcal{F}_T \subset \mathcal{F}_{s\wedge T}.$$

因此

$$\begin{aligned}
&\mathsf{E}[X_t^T 1_{T>s} \mathbb{1}_A] \\
=\ &\mathsf{E}[\mathsf{E}[X_t^T 1_{T>s} \mathbb{1}_A|\mathcal{F}_{s\wedge T}]] \\
=\ &\mathsf{E}[1_{T>s} \mathbb{1}_A \mathsf{E}[X_t^T|\mathcal{F}_{s\wedge T}]] \\
=\ &\mathsf{E}[1_{T>s} \mathbb{1}_A \mathsf{E}[X_{t\wedge T}|\mathcal{F}_{s\wedge T}]] \\
=\ &\mathsf{E}[1_{T>s} \mathbb{1}_A X_{s\wedge T}] = \mathsf{E}[1_{T>s} \mathbb{1}_A X_s^T].
\end{aligned}$$

因此

$$\mathsf{E}[X_t^T \mathbb{1}_{T>s}|\mathcal{F}_s] = X_s^T \mathbb{1}_{T>s},$$

$$\mathsf{E}[X_t^T|\mathcal{F}_s] = X_s^T.$$

X^T 是 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathsf{P})$ 上的鞅.

■

上述定理的证明可以看到，在离散时间框架下很容易得到的结论, 在连续时间框架下不一定很容易得到.

在这一节的最后, 我们给出连续鞅的 Doob 不等式.

定理 3.1.4 若 X 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续鞅, $p>1$ 是常数, 若对于 $t>0$, $\mathsf{E}[|X_t|^p]<\infty$, 则

$$\mathsf{E}[\sup_{0\leqslant s\leqslant t}|X_s|^p]\leqslant(\frac{p}{p-1})^p\mathsf{E}[|X_t|^p].$$

证明: 设 $\widetilde{D}$ 是 $\mathbb{R}$ 上的可列稠密集, $D=\widetilde{D}\bigcap[0,t]$. 令 $D=\bigcup\limits_{m=1}^{\infty}D_m$,$D_m$ 是一列单调递增的集合, 且 $D_m=\{t_0^m,\cdots,t_m^m\}$, 这里 $0=t_0^m<t_1^m<\cdots<t_m^m=t$. 由离散时间鞅的 Doob 不等式,

$$\mathsf{E}[\sup_{s\in D_m}|X_s|^p]\leqslant(\frac{p}{p-1})^p\mathsf{E}[|X_t|^p].$$

令 $m\uparrow\infty$, 有

$$\mathsf{E}[\sup_{s\in D}|X_s|^p]\leqslant(\frac{p}{p-1})^p\mathsf{E}[|X_t|^p].$$

由于 X 轨道连续,

$$\sup_{s\in D}|X_s|=\sup_{0\leqslant s\leqslant t}|X_s|,$$

故

$$\mathsf{E}[\sup_{0\leqslant s\leqslant t}|X_s|^p]\leqslant(\frac{p}{p-1})^p\mathsf{E}[|X_t|^p].$$

■

3.2 布朗运动的定义与性质

布朗运动是重要的连续鞅的例子, 我们下面暂时不谈论鞅, 从另外的角度讨论布朗运动.

首先从中心极限定理出发, 设 $\{X_n\}$ 是一列独立同分布的随机变量, $\mathsf{E}[X_1]=0,\mathsf{E}[X_1^2]=1$. 此时, 由中心极限定理可知,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i\xrightarrow{d}N(0,1).$$

考虑特殊的情形,

$$\mathsf{P}(X_1=1)=\mathsf{P}(X_1=-1)=\frac{1}{2},$$

$$S_n=X_1+X_2+\cdots+X_n.$$

这里, S_n 是零点出发的随机游动. 假设在上述过程中, 时间间隔为 1, 位移也为 1,$\Delta t=1$, $\Delta x=1$. 现在将时间间隔变为 $\dfrac{1}{N}$, 位移变为 $\sqrt{\dfrac{1}{N}}$, 即 $\Delta x=\sqrt{\Delta t}$.

在时刻 1, 有

$$W_1^{(N)}=\Delta x(X_1+\cdots+X_N),$$

$$\mathsf{E}[W_1^{(N)}]=0,$$

$$\mathsf{E}[(W_1^{(N)})^2]=(\Delta x)^2(\mathsf{E}[X_1^2]+\cdots+\mathsf{E}[X_N^2])=(\Delta x)^2N=1.$$

由中心极限定理

$$\frac{X_1+\cdots+X_N}{\sqrt{N}}\xrightarrow{d}N(0,1).$$

我们可以这样理解布朗运动：上述随机游动的轨迹的极限 ($\Delta t\to 0$). 事实上, 可以考虑下式极限.

$$W_t^{(n)}(\omega)=\frac{1}{\sqrt{n}}S_{[nt]}(\omega)+(nt-[nt])\frac{1}{\sqrt{n}}X_{[nt]+1}(\omega).$$

当然, 这个极限的意义是在弱收敛的意义下. 历史上著名的 Donsker 不变原理即是关于这个结论的详细表述. 下面我们给出布朗运动的定义.

定义 3.2.1　布朗运动 $B=(B_t)_{t\geqslant 0}$ 是定义在 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机过程, 满足:

(1) $B_0=0$;

(2) 对于 $t>s$, B_t-B_s 与 B_r 独立, 其中 $0\leqslant r\leqslant s$;

(3) B_t-B_s 与 B_{t-s} 同分布, 且服从 $N(m(t-s),\sigma^2(t-s))$;

(4) $(B_t)_{t\geqslant 0}$ 的几乎所有的样本轨道连续.

称 B 是布朗运动. 特别地, $m=0,\sigma=1$ 时, 称 B 是标准布朗运动.

这里 $B_0=0$ 不是必需的, 如果不考虑这个条件, 可以对 B_t-B_0 进行考虑.

历史上, 关于布朗运动的存在性定理是很多数学家关心的内容. 下面介绍著名的 Kolomogorov 连续性定理, 该定理和布朗运动的存在性有关.

定理 3.2.1　(Kolomogorov 连续性定理) $X=(X_t)_{t\in I}$ 是一个随机过程, I 是一个 $\mathbb{R}_+$ 上的有界区间, 存在常数 $q,\varepsilon,C>0$, 使得对于 $s,t\in I$,

$$\mathsf{E}[|X_t-X_s|^q]\leqslant C|t-s|^{1+\varepsilon}.$$

则, 存在 X 的修正 $\tilde{X}$, 使得对于 $\omega\in\Omega$, $\alpha\in(0,\frac{\varepsilon}{q})$, 存在 $C_\alpha(\varepsilon)$, 使得 $s,t\in I$,

$$|\tilde{X}_t(\omega)-\tilde{X}_s(\omega)|\leqslant C_\alpha(\omega)|t-s|^\alpha.$$

上述定理的证明我们这里略去，感兴趣的读者参考 Le Gall, F. [6] 中的定理 2.9 的证明，由上述定理，对于随机过程 $X=(X_t)_{t\geqslant 0}$, 可以分为可列个闭区间依次讨论，因此会存在 $\mathbb{R}_+$ 上局部 Hölder 连续的过程的修正. 由布朗运动的分布性质，可以利用 Kolomogorov 连续性定理得到布朗运动有连续的修正. 本书直接假设布朗运动是存在的，且拥有连续轨道.

设 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, B_t-B_s 服从 $N(0,t-s)$, 可离散地对布朗运动进行采样, 取得 $B_0,B_{\Delta t},B_{2\Delta t},\cdots,B_{k\Delta t}$, 且有

$$B_{(k+1)\Delta t}-B_{k\Delta t}=\sqrt{\Delta t}\;N_k,\qquad N_k\sim N(0,1).$$

因此, 布朗运动利用随机模拟的方法, 可以模拟出其轨道的走向. 但是理论上来讲, 布朗运动的轨道有十分著名的性质: 布朗运动虽然处处连续, 但是处处不可微. 下面, 我们来证明这个性质.

定理 3.2.2 *在 $(\Omega,\mathcal{F},\mathsf{P})$ 上, 存在 $N\in\mathcal{F}$, $\mathsf{P}(N)=0$. 在 $\Omega-N$ 上,$B=(B_t)_{t\geqslant 0}$ 是处处不可微的.*

证明 [4]: 令

$$Y_{k,n}=\max\{|B_{\frac{k+1}{n}}-B_{\frac{k}{n}}|,|B_{\frac{k+2}{n}}-B_{\frac{k+1}{n}}|,|B_{\frac{k+3}{n}}-B_{\frac{k+2}{n}}|\},$$

$$Y_n=\min\{Y_{k,n}:k=0,1,\cdots,n-1\}.$$

考虑 M 是正整数, $A_M=\{\omega\in\Omega:存在足够大的\ n,\ Y_n\leqslant\dfrac{M}{n}\}$. 由概率的基本运算,

$$\mathsf{P}(Y_n\leqslant\frac{M}{n})\leqslant\sum_{k=0}^{n-1}\mathsf{P}(Y_{k,n}\leqslant\frac{M}{n}).$$

由独立增量性

$$\begin{aligned}\mathsf{P}(Y_{k,n}\leqslant\frac{M}{n})&=[\mathsf{P}(|B_{\frac{1}{n}}|\leqslant\frac{M}{n})]^3\\&=[\mathsf{P}(\frac{1}{\sqrt{n}}|B_1|\leqslant\frac{M}{n})]^3\\&=[\int_{|x|\leqslant\frac{M}{\sqrt{n}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\mathrm{d}x]^3\leqslant\frac{M^3}{n^{3/2}}.\end{aligned}$$

故

$$\mathsf{P}(Y_n\leqslant\frac{M}{n})\leqslant\frac{M^3}{\sqrt{n}}.$$

于是 $\mathsf{P}(A_M)=0$, 进一步, 有 $\mathsf{P}(\bigcup_{M=1}^{\infty} A_M)=0$.

若 $B=(B_t)_{t\geqslant 0}$ 在 t_0 处可微, 则存在 δ, $|s-t_0|<\delta$ 时,

$$|B_s-B_{t_0}|\leqslant(|r|+1)|s-t_0|,\qquad r \text{ 为 } B_t \text{ 在 } t_0\text{处的导数}.$$

故

$$\{\omega: B\text{在某处可微}\}\subset\bigcup_{M=1}^{\infty}A_M.$$

令 $N=\bigcup_{M=1}^{\infty}A_M$, 故 $B=(B_t)_{t\geqslant 0}$ 几乎处处不可微. ■

利用布朗运动的性质, 可以考虑下面几个例子.

例 3.2.1 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, $\mathcal{F}_s=\sigma(B_r,0\leqslant r\leqslant s)$, 求 $\mathsf{E}[e^{B_1B_2}\mid\mathcal{F}_1]$.

解: 考虑

$$\begin{aligned}\mathsf{E}[e^{B_1B_2}\mid\mathcal{F}_1]&=\mathsf{E}[e^{(B_2-B_1)B_1+B_1^2}\mid\mathcal{F}_1]\\&=e^{B_1^2}\mathsf{E}[e^{(B_2-B_1)B_1}\mid\mathcal{F}_1].\end{aligned}$$

事实上

$$\mathsf{E}[e^{(B_2-B_1)B_1}\mid\mathcal{F}_1]=\mathsf{E}[e^{(B_2-B_1)x}]|_{x=B_1}=e^{\frac{B_1^2}{2}}.$$

故

$$\mathsf{E}[e^{B_1B_2}\mid\mathcal{F}_1]=e^{B_1^2+\frac{B_1^2}{2}}=e^{\frac{3B_1^2}{2}}.$$

■

例 3.2.2 求 $\mathsf{E}[\int_0^t B_r^2\,\mathrm{d}r]$, 这里 $B=(B_r)_{r\geqslant 0}$ 是标准布朗运动.

解:

$$\mathsf{E}[\int_0^t B_r^2\,\mathrm{d}r]=\int_0^t\mathsf{E}[B_r^2]\,\mathrm{d}r=\int_0^t r\,\mathrm{d}r=\frac{1}{2}t^2.$$

■

例 3.2.3 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 求 $\int_0^1 B_r\,\mathrm{d}r$ 的分布.

解: 注意到 $\int_0^1 B_r\,\mathrm{d}r$ 可以写为 $\frac{1}{n}\sum_{i=1}^n B_{i/n}$ 的极限. 由于 B 是高斯过程, $(B_{\frac{1}{n}},B_{\frac{2}{n}},\cdots,B_1)$ 服从多元正态分布, 因此 $\frac{1}{n}\sum_{i=1}^n B_{i/n}$ 服从正态分布. 利用特

征函数, 可以知道一列正态随机变量依分布收敛的极限依然服从正态分布. 因此 $\int_0^1 B_r\,\mathrm{d}r$ 服从正态分布. 注意到

$$\mathsf{E}[\int_0^1 B_r\,\mathrm{d}r] = \int_0^1 \mathsf{E}[B_r]\,\mathrm{d}r = 0,$$

同时,

$$\begin{aligned}\mathsf{E}[(\int_0^1 B_r\,\mathrm{d}r)^2] &= \mathsf{E}[\int_0^1 B_r\,\mathrm{d}r \int_0^1 B_s\,\mathrm{d}s] \\ &= \mathsf{E}[\int_0^1\int_0^1 B_rB_s\,\mathrm{d}r\,\mathrm{d}s] \\ &= \int_0^1\int_0^1 \mathsf{E}[B_rB_s]\,\mathrm{d}r\,\mathrm{d}s \\ &= \int_0^1\int_0^1 r\wedge s\,\mathrm{d}r\,\mathrm{d}s = \frac{1}{3}.\end{aligned}$$

因此

$$\int_0^1 B_r\,\mathrm{d}r \sim N(0,\frac{1}{3}).$$

■

例 3.2.4 *若* $B=(B_t)_{t\geqslant 0}$ *是* $(\Omega,\mathcal{F},\mathsf{P})$ *上的标准布朗运动, 求* $\mathsf{E}[B_1|B_2]$.

解: 由于 B 是高斯过程, (B_1,B_2) 服从二元正态分布, 注意到 $\mathsf{E}[(B_1-\frac{1}{2}B_2)B_2]=0$，因此 $B_1-\frac{1}{2}B_2$ 与 B_2 独立, 故

$$\mathsf{E}[(B_1-\frac{1}{2}B_2)+\frac{1}{2}B_2|B_2] = \frac{1}{2}B_2.$$

■

在近现代概率论的发展历史中, Markov 性 (简称马氏性) 曾经起了重要的作用. 满足马氏性的过程称为 Markov 过程 (简称马氏过程). 马氏过程的研究起源于 Markov 链, 最早是由俄国数学家 Markov 1907 年提出来. 马氏过程有很强的应用背景, 一百多年来, 在随机过程的研究中扮演了非常重要的角色. 简言之, 马氏过程是描述一种条件独立的过程. 即在已知“现在”的条件下,“将来”与“过去”无关. 若 $X=(X_t)_{t\geqslant 0}$ 是一个马氏过程, 考虑 s 时刻及 s 时刻之前的信息的条件下, $f(X_{t+s})$ 的条件数学期望应与只考虑 s 时刻的条件下, $f(X_{t+s})$ 的条件数学期望相同. 我们下面来证明布朗运动具备这样的性质.

设 B 是标准布朗运动. 定义

$$p(t,x) = \frac{1}{(2\pi t)^{1/2}}e^{-\frac{x^2}{2t}}, \qquad x\in\mathbb{R},\ t>0.$$

设 $p(t,x,y)=p(t,y-x)$, 对于 R 上的有界连续函数 f, 定义

$$T_t\ f(x)=\int_R f(y)\ p(t,x,y)\ \mathrm{d}y=\mathsf{E}[f(x+B_t)].$$

由于

$$p(t+s,x,y)=\int_R p(t,x,z)\ p(s,z,y)\ \mathrm{d}z,$$

故有

$$T_{t+s}=T_tT_s.$$

命题 3.2.1 设 $t,s>0$, f 是有界的 Borel 可测函数, 那么

$$\mathsf{E}[f(B_{t+s})\mid\mathcal{F}_s]=T_t\ f(B_s)\qquad a.s..$$

这里 $\mathcal{F}_s=\sigma(B_r:0\leqslant r\leqslant s)$, 特别地

$$\mathsf{E}[f(B_{t+s})\mid\mathcal{F}_s]=\mathsf{E}[f(B_{t+s})\mid B_s].$$

证明: 由于 $B_{t+s}-B_s$ 与 $\mathcal{F}_s$ 独立, 其密度是 $p(t,x)$, 故

$$\begin{aligned}\mathsf{E}[f(B_{t+s})\mid\mathcal{F}_s]&=\mathsf{E}[f(B_{t+s}-B_s+B_s)\mid\mathcal{F}_s]\\&=\mathsf{E}[f(B_{t+s}-B_s+x)]|_{x=B_s}\\&=\int f(y+x)\ p(t,y)\ \mathrm{d}y|_{x=B_s}\\&=\int f(y)\ p(t,B_s,y)\ \mathrm{d}y\\&=T_tf(B_s).\end{aligned}$$

上式两端对 B_s 取条件数学期望, 有

$$\mathsf{E}[f(B_{t+s})\mid B_s]=T_tf(B_s),$$

故

$$\mathsf{E}[f(B_{t+s})\mid B_s]=\mathsf{E}[f(B_{t+s})\mid\mathcal{F}_s].$$

■

值得一提的是, 上面提到的一族算子 $(T_t)_{t\geqslant0}$, 我们往往称为转移半群. 与转移半群相关联的是下面给出的无穷小生成元.

对于转移半群 $(T_t)_{t\geqslant 0}$, 对于任意的可测函数 f, 算子

$$\mathcal{L}f = \lim_{t\to 0}\frac{T_t f - f}{t}$$

称为半群 $(T_t)_{t\geqslant 0}$ 的**无穷小生成元**.

事实上，很多情形下, 需要对 f 加以限制, 有时候会要求 f 具有可微性. 在下文的叙述中, 会在相应位置给出 f 的具体要求. 无穷小生成元与我们后面要介绍的 Itô 公式是密切相关的. 我们会详细讨论一些过程的无穷小生成元.

命题 3.2.2 设 f 的二阶导数有界且一致连续时, 对于标准布朗运动 B, 其生成元为

$$\mathcal{L}f = \frac{1}{2}f''(x).$$

证明:

$$\begin{aligned}
T_t f(x) &= \mathsf{E}[f(x+B_t)] \\
&= \int_{\mathbb{R}} f(y)\ p(t,x,y)\ \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} f(y)e^{-\frac{(x-y)^2}{2t}}\ \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x+y\sqrt{t})e^{-\frac{y^2}{2}}\ \mathrm{d}y.
\end{aligned}$$

将 $f(x+y\sqrt{t})$ 在 x 泰勒展开,

$$f(x+y\sqrt{t}) = f(x) + f'(x)y\sqrt{t} + \frac{1}{2}f''(x)y^2 t + o(t).$$

故

$$\mathcal{L}f = \frac{1}{2}f''(x).$$

■

以下推导中, $||\cdot||$ 记为算子的范数或者函数的范数, 当 $s>0$,

$$\begin{aligned}
&||\frac{1}{s}(T_sT_tf - T_tf) - T_t\mathcal{L}f|| \\
\leqslant\ & ||T_t(\frac{1}{s}(T_sf - f) - \mathcal{L}f)|| \\
\leqslant\ & ||\frac{1}{s}(T_sf - f) - \mathcal{L}f||
\end{aligned}$$

因此

$$\lim_{s\to 0+}\frac{T_{s+t}f - T_tf}{s} = \lim_{s\to 0+}\frac{T_tT_sf - T_tf}{s} = T_t\mathcal{L}f.$$

进一步, $s>0$ 时

$$\begin{aligned}
& ||(-\frac{1}{s})(T_{t-s}f-T_tf)-T_t\mathcal{L}f|| \\
\leqslant\ & ||(-\frac{1}{s})(T_{t-s}f-T_tf)-T_{t-s}\mathcal{L}f||+||T_{t-s}\mathcal{L}f-T_t\mathcal{L}f|| \\
\leqslant\ & ||\frac{1}{s}(T_sf-f)-\mathcal{L}f||+||\mathcal{L}f-T_s\mathcal{L}f||.
\end{aligned}$$

因此

$$\lim_{s\to 0+}\frac{T_tf-T_{t-s}f}{s}=T_t\mathcal{L}f.$$

于是

$$T_tf=f+\int_0^t T_s\mathcal{L}fds.$$

命题 3.2.3 设 f 的二阶导数有界且一致连续时, 对于标准布朗运动 B, 令

$$M_t=f(B_t)-\frac{1}{2}\int_0^t f''(B_s)ds,$$

$$\mathcal{F}_t=\sigma(B_r,0\leqslant r\leqslant t).$$

则 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅.

证明 [11]: 对于 $u\geqslant 0$,

$$\begin{aligned}
\mathsf{E}[M_{t+u}|\mathcal{F}_t] &= \mathsf{E}[f(B_{t+u})|\mathcal{F}_t]-\mathsf{E}[\int_t^{t+u}\mathcal{L}f(B_s)ds|\mathcal{F}_t]-\frac{1}{2}\int_0^t f''(B_s)ds \\
&= \int_{\mathbb{R}} f(y)p(u,B_t,y)dy-\int_t^{t+u}\int_{\mathbb{R}}\mathcal{L}f(y)p(s-t,B_t,y)dyds \\
&\quad -\frac{1}{2}\int_0^t f''(B_s)ds \\
&= T_uf(B_t)-\int_0^u T_s\mathcal{L}f(B_t)ds-\frac{1}{2}\int_0^t f''(B_s)ds \\
&= f(B_t)-\frac{1}{2}\int_0^t f''(B_s)ds.
\end{aligned}$$

因此, $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅. ■

接下来, 我们给出一个更强的结论. 该定理阐述了布朗运动具备所谓的强马氏性.

定理 3.2.3 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上标准布朗运动, $\mathcal{F}_s=\sigma(B_r,0\leqslant r\leqslant s)$. 若 T 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的停时, 且 $\mathsf{P}(T<+\infty)=1$, 令 $Y_t=B_{T+t}-B_T$, $Y=(Y_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 且 Y 与 $\mathcal{F}_T$ 独立.

证明 [6]: 由 Y 的定义, Y 是零点出发的轨道连续的随机过程. 对于 $A \in \mathcal{F}_T$, 下面证明对于任意的有界连续函数 g, $0 \leqslant t_1 < t_2 < \cdots < t_p$,

$$\mathsf{E}[\mathbb{1}_A g(Y_{t_1}, \cdots, Y_{t_p})] = \mathsf{P}(A)\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})].$$

若能证明上式, 便能证明 Y 与 $\mathcal{F}_T$ 独立, 且可以得到

$$\mathsf{E}[g(Y_{t_1}, \cdots, Y_{t_p})|\mathcal{F}_T] = \mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})].$$

取 $A = \Omega$, 便能得到 Y 的有限维分布与 B 相同, 同时, 独立平稳增量性质也能得到. 因此, 可以得到 $Y = (Y_t)_{t \geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动. 下面我们证明对于 $A \in \mathcal{F}_T$, 有界连续函数 g, $0 \leqslant t_1 < t_2 < \cdots < t_p$,

$$\mathsf{E}[\mathbb{1}_A g(Y_{t_1}, \cdots, Y_{t_p})] = \mathsf{P}(A)\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})].$$

首先考虑 T 只可取可列个离散值的情形. 设 T 只取可列值 $r_1 < r_2 < \cdots < r_n < \cdots$. 注意到 $\{T = r_n\} \in \mathcal{F}_{r_n}$, 定义 $A_n := A\mathbb{1}_{\{T=r_n\}} \in \mathcal{F}_{r_n}$. 因此

$$\begin{aligned}
& \mathsf{E}[\mathbb{1}_A g(Y_{t_1}, \cdots, Y_{t_p})] \\
= \ & \mathsf{E}[\mathbb{1}_A \sum_{n=1}^{\infty} \mathbb{1}_{\{T=r_n\}} g(Y_{t_1}, \cdots, Y_{t_p})] \\
= \ & \sum_{n=1}^{\infty} \mathsf{E}[\mathbb{1}_{A_n} g(B_{r_n+t_1} - B_{r_n}, \cdots, B_{r_n+t_p} - B_{r_n})] \\
= \ & \sum_{n=1}^{\infty} \mathsf{E}[\mathbb{1}_{A_n}]\mathsf{E}[g(B_{r_n+t_1} - B_{r_n}, \cdots, B_{r_n+t_p} - B_{r_n})] \\
= \ & \sum_{n=1}^{\infty} \mathsf{E}[\mathbb{1}_{A_n}]\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})] \\
= \ & \mathsf{E}[\mathbb{1}_A]\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})] \\
= \ & \mathsf{P}(A)\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})].
\end{aligned}$$

考虑一般情况. 令

$$T_n = \sum_{k=1}^{\infty} \frac{k}{2^n} \mathbb{1}_{\{\frac{k-1}{2^n} < T \leqslant \frac{k}{2^n}\}}.$$

易知, T_n 是停时, 同时可有 $T_n \downarrow T$, 由控制收敛定理,

$$\begin{aligned}
& \mathsf{E}[\mathbb{1}_A g(Y_{t_1}, \cdots, Y_{t_p})] \\
= & \lim_{n\to\infty} \mathsf{E}[\mathbb{1}_A g(B_{T_n+t_1} - B_{T_n}, \cdots, B_{T_n+t_p} - B_{T_n})] \\
= & \lim_{n\to\infty} \mathsf{P}(A)\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})] \\
= & \mathsf{P}(A)\mathsf{E}[g(B_{t_1}, \cdots, B_{t_p})].
\end{aligned}$$

■

$B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动, 定义

$$\begin{aligned}
T_a &= \begin{cases} \inf\{t \geqslant 0 : B_t \geqslant a\}, & a > 0, \\ \inf\{t \geqslant 0 : B_t \leqslant a\}, & a < 0, \end{cases} \\
&= \inf\{t \geqslant 0 : B_t = a\}.
\end{aligned}$$

定义

$$\widetilde{B} = \begin{cases} B_t, & t < T_a, \\ 2a - B_t, & t \geqslant T_a, \end{cases}$$

下面给出布朗运动的反射原理.

定理 3.2.4　$\widetilde{B} = (\widetilde{B}_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动,

证明 [11]:

我们首先证明 $\mathsf{P}(T_a < \infty) = 1$. 只考虑 $a > 0$ 的情形, 注意到

$$\{\limsup_{n\to\infty} B_n = +\infty\} \in \bigcap_{n=1}^{\infty} \sigma(\xi_k : k \geqslant n),$$

这里 $\xi_k = B_k - B_{k-1}$. $\{\xi_k\}_{k\geqslant 1}$ 是一列独立同分布的随机变量, 注意到 $\sigma(\xi_k : k \geqslant n+1)$ 与 $\sigma(\xi_n)$ 独立, $\bigcap_{n=1}^{\infty} \sigma(\xi_k : k \geqslant n) \subset \sigma(\xi_k : k \geqslant n+1)$, 因此 $\sigma(\xi_n)$ 与 $\bigcap_{n=1}^{\infty} \sigma(\xi_k : k \geqslant n)$ 独立. 同时, $\bigcap_{n=1}^{\infty} \sigma(\xi_k : k \geqslant n) \subset \sigma(\xi_n)$, 因此 $\bigcap_{n=1}^{\infty} \sigma(\xi_k : k \geqslant n)$ 与自己独立. 故 $\bigcap_{n=1}^{\infty} \sigma(\xi_k : k \geqslant n)$ 中集合的概率测度或为 1 或为 0, 而

$$\begin{aligned}
& P(\limsup_{n\to\infty} B_n \geqslant a) \\
\geqslant & P(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{B_n \geqslant a\}) \geqslant \limsup_{n\to\infty} P(B_n \geqslant a) \\
> & 0.
\end{aligned}$$

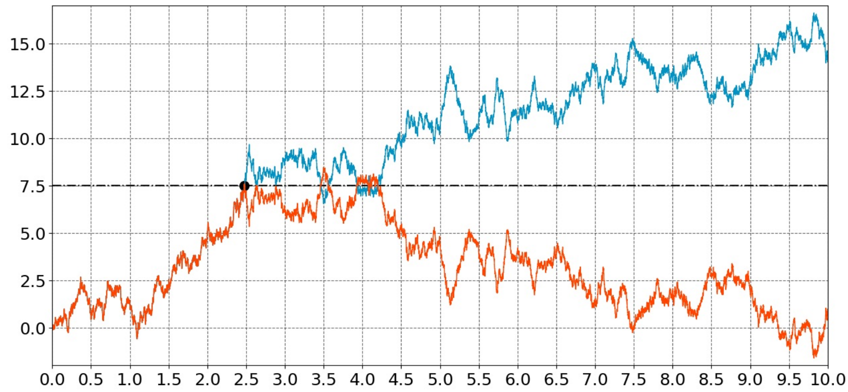

图 3.1 布朗运动与反射布朗运动

因此

$$P(\limsup_{n\to\infty} B_n = +\infty) \geqslant P(\limsup_{n\to\infty} B_n \geqslant a) = 1.$$

于是 $\mathsf{P}(T_a < \infty) = 1$. 令 $Y_t = B_{t\wedge T_a}$, $Z_t = B_{T_a+t} - B_{T_a}$, 由布朗运动的强马氏性, $Z = (Z_t)_{t\geqslant 0}$ 是标准布朗运动, 且与 $Y = (Y_t)_{t\geqslant 0}$ 独立. 由布朗运动的性质, $-Z$ 是标准布朗运动, 且与 Y 独立. 因此 (Y, Z) 与 $(Y, -Z)$ 同分布.

$$\phi : (Y, Z) \to (Y_t \mathbb{1}_{\{t\leqslant T_a\}} + (a + Z_{t-T_a})\mathbb{1}_{\{t>T_a\}})_{t\geqslant 0}.$$

$\phi(Y, Z) = B$, $\phi(Y, -Z) = \widetilde{B}$, 由于 (Y, Z) 与 $(Y, -Z)$ 同分布，ϕ 是可测映射, 因此 $\widetilde{B} = (\widetilde{B}_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动.

■

利用反射原理, 我们可以得到如下结果.

定理 3.2.5 *若 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动, 任意给定 $t > 0$, 令 $M_t = \sup\limits_{0<s\leqslant t} B_s$. 则*

$$M_t \overset{d}{=} |B_t|.$$

证明: 由于 $B_0 = 0$, 故 $M_t \geqslant 0$. 任意给定 $x > 0$, 故

$$\mathsf{P}(M_t \geqslant x) = \mathsf{P}(M_t \geqslant x, B_t \geqslant x) + \mathsf{P}(M_t \geqslant x, B_t < x).$$

显然

$$\mathsf{P}(M_t \geqslant x, B_t \geqslant x) = \mathsf{P}(B_t \geqslant x).$$

下证

$$\mathsf{P}(M_t \geqslant x, B_t < x) = \mathsf{P}(M_t \geqslant x, B_t \geqslant x).$$

令

$$T_x = \inf\{t \geqslant 0 : B_t = x\}.$$

定义

$$\widetilde{W} = \begin{cases} B_t, & t < T_x, \\ 2x - B_t, & t \geqslant T_x, \end{cases}$$

由反射原理, $\widetilde{W}$ 是布朗运动. 定义 $\widetilde{M}_t = \sup\limits_{0<s\leqslant t} \widetilde{W}_s$. 考虑 $x > y \geqslant 0$. 注意到

$$\mathsf{P}(M_t \geqslant x, B_t < x - y) = \mathsf{P}(\widetilde{M}_t \geqslant x, \widetilde{W}_t < x - y).$$

同时

$$\{\widetilde{M}_t \geqslant x, \widetilde{W}_t < x - y\} = \{M_t \geqslant x, B_t > x + y\}.$$

因此

$$\mathsf{P}(M_t \geqslant x, B_t < x - y) = \mathsf{P}(M_t \geqslant x, B_t > x + y).$$

取 $y = 0$, 我们得到, 且注意到 $\mathsf{P}(B_t = x) = 0$,

$$\mathsf{P}(M_t \geqslant x, B_t < x) = \mathsf{P}(M_t \geqslant x, B_t \geqslant x).$$

因此

$$\mathsf{P}(M_t \geqslant x) = 2\mathsf{P}(B_t \geqslant x).$$

由 B 的对称性

$$M_t \overset{d}{=} |B_t|.$$

■

在上述讨论中，我们还可以注意到, 对于 $x, y \geqslant 0$,

$$\{M_t > x, B_t > x + y\} = \{B_t > x + y\}.$$

因此

$$\mathsf{P}(M_t > x, B_t < x - y) = \mathsf{P}(B_t > x + y).$$

令 $a=x$, $b=x-y$, 带入上式, 我们可得对于 $a>0$, $a>b\geqslant 0$,

$$\mathsf{P}(M_t>a,B_t<b)=\mathsf{P}(B_t>2a-b).$$

在上面的证明过程中, 用到了停时 T_x, T_x 的概率性质是十分有趣的. 事实上, 会有一个十分有趣的结论: $|x|$ 无论多大, 从 0 出发的布朗运动总能在有限时间内达到 $|x|$, 但平均到达时间为 ∞. 下面给出的命题即阐述了上述事实.

命题 3.2.4 $\mathsf{E}[T_x]=\infty$.

证明: 设 $f_x(t)$ 是 T_x 的密度函数. 对 $x>0,t>0$,

$$\mathsf{P}(T_x\leqslant t)=\mathsf{P}(M_t\geqslant x)=\sqrt{\frac{2}{\pi}}\int_x^\infty \frac{1}{t^{1/2}}e^{-\widetilde{x}^2/(2t)}\mathrm{d}\widetilde{x}.$$

对 t 求导, 有

$$\begin{aligned}f_x(t)&=\frac{1}{\sqrt{2\pi}}\int_x^\infty\left(-\frac{1}{t^{3/2}}+\frac{\widetilde{x}^2}{t^{5/2}}\right)e^{-\widetilde{x}^2/(2t)}\,\mathrm{d}\widetilde{x}\\&=\frac{1}{\sqrt{2\pi}}\int_x^\infty -\frac{1}{t^{3/2}}e^{-\widetilde{x}^2/(2t)}\,\mathrm{d}\widetilde{x}+\frac{1}{\sqrt{2\pi}}\int_x^\infty\frac{\widetilde{x}^2}{t^{5/2}}e^{-\widetilde{x}^2/(2t)}\,\mathrm{d}\widetilde{x}.\end{aligned}$$

利用分部积分, 有

$$\int_x^\infty -\frac{1}{t^{3/2}}e^{-\widetilde{x}^2/(2t)}\,\mathrm{d}\widetilde{x}=\frac{x}{t^{3/2}}e^{-x^2/(2t)}-\int_x^\infty\frac{\widetilde{x}^2}{t^{5/2}}e^{-\widetilde{x}^2/(2t)}\,\mathrm{d}\widetilde{x}.$$

故

$$f_x(t)=\frac{x}{\sqrt{2\pi}t^{3/2}}e^{-x^2/(2t)}.$$

当 $x<0$ 时可有类似结果. 故

$$f_x(t)=\frac{|x|}{\sqrt{2\pi}t^{3/2}}e^{-x^2/(2t)}\qquad t>0.$$

不妨设 $x>0$, 我们研究 T_x. 事实上

$$\begin{aligned}\mathsf{E}[T_x]&=\int_0^\infty tf_x(t)\,\mathrm{d}t\\&=\frac{x}{\sqrt{2\pi}}\int_0^\infty\frac{1}{t^{1/2}}e^{-x^2/(2t)}\,\mathrm{d}t\\&=\infty.\end{aligned}$$

■

利用反射原理, 我们可以考虑下面这个例子.

例 3.2.5 若 $X=(X_t)_{t\geqslant 0}, Y=(Y_t)_{t\geqslant 0}$ 是独立的标准布朗运动. 求 P(对所有$0\leqslant t\leqslant 2, X_t<Y_t+1$).

解: 令 B 是标准布朗运动, 由布朗运动的性质,

$$\begin{aligned}\mathsf{P}(\text{对所有}0\leqslant t\leqslant 2, X_t<Y_t+1) &= \mathsf{P}(\text{对所有}0\leqslant t\leqslant 2, \sqrt{2}B_t<1,)\\ &= 1-\mathsf{P}(\text{存在 } s\in[0,2], \text{使得 } B_s\geqslant \frac{1}{\sqrt{2}},)\\ &= 1-2\mathsf{P}(B_2\geqslant \frac{1}{\sqrt{2}}).\end{aligned}$$

■

关于布朗运动的刻画, 有很多方法, 下面我们给出其中一个比较常见的方法.

定理 3.2.6 令 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的随机过程, 则以下推断等价:

(1) B 是标准布朗运动;

(2) B 是连续轨道的中心化高斯过程(任取有限个点 $t_1,t_2,\cdots,t_n$, $(B_{t_1},B_{t_2},\cdots,B_{t_n})$ 服从高斯分布), 且对于 $s,t>0$, $Cov[B_s,B_t]=s\wedge t$.

证明:

(1) $\Longrightarrow$ (2): 假设 $t>s$, 则

$$Cov[B_t-B_s+B_s,B_s]=Cov[B_t-B_s,B_s]+Cov[B_s,B_s]=s.$$

(2) $\Longrightarrow$ (1): 假设 $t>s$, 则

$$Cov[B_t,B_s]=Cov[B_t-B_s,B_s]+Cov[B_s,B_s].$$

由于

$$Cov[B_t-B_s,B_s]=0.$$

故 B_t-B_s与B_s 独立. 此时

$$\begin{aligned}\mathsf{E}[B_t-B_s]^2 &= Cov[B_t-B_s,B_t-B_s]\\ &= Cov[B_t,B_t]-Cov[B_t,B_s]-Cov[B_s,B_t]\\ &\quad +Cov[B_s,B_s]\\ &= t-s-s+s=t-s.\end{aligned}$$

从而

$$B_t-B_s\sim N(0,t-s),$$

即 B_t-B_s 与 B_{t-s} 同分布. 即 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动. ■

利用上面的定理, 有

定理 3.2.7 若 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 对于 $\lambda>0$, 令 $Y_t=\dfrac{B_{\lambda t}}{\sqrt{\lambda}}$, 则 $Y=(Y_t)_{t\geqslant 0}$ 是标准布朗运动.

证明: 显然 $Y=(Y_t)_{t\geqslant 0}$ 是中心化的高斯过程. 且

$$\begin{aligned}Cov(Y_t,Y_s)&=Cov(\frac{B_{\lambda t}}{\sqrt{\lambda}},\frac{B_{\lambda s}}{\sqrt{\lambda}})=\frac{1}{\lambda}Cov[B_{\lambda t},B_{\lambda s}]\\&=\frac{1}{\lambda}(\lambda s)\wedge(\lambda t)=s\wedge t.\end{aligned}$$

故 $Y=(Y_t)_{t\geqslant 0}$ 是标准布朗运动. ■

我们利用上述定理, 可以考虑下面这个例子.

例 3.2.6 令 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 则

$$W_t=\begin{cases}tB_{1/t}, & t>0,\\ 0, & t=0,\end{cases}$$

是标准布朗运动.

证明 [4]: 当 $t>0$ 时, 显然 W_t 是连续的中心化的高斯过程.

$$Cov(W_t,W_s)=tsCov(B_{1/t},B_{1/s})=ts(\frac{1}{t}\wedge\frac{1}{s})=t\wedge s.$$

为证 W 是标准布朗运动, 只需证明 W 在 $t=0$ 处连续. 事实上

$$\lim_{t\to 0}W_t=\lim_{t\to+\infty}\frac{1}{t}B_t\leqslant\lim_{n\to\infty}\frac{1}{n}B_n+\lim_{n\to\infty}\frac{1}{n}\sup_{t\in[n,n+1]}(B_t-B_n).$$

这里

$$B_n=\sum_{i=1}^{n}(B_i-B_{i-1})\quad B_i-B_{i-1}\sim N(0,1).$$

由强大数律, 我们有

$$\lim_{n\to\infty}\frac{1}{n}B_n=0.$$

注意到

$$\begin{aligned}\mathsf{P}(\sup_{t\in[n,n+1]}(B_t-B_n)\geqslant x)&=\mathsf{P}(\sup_{t\in[0,1]}B_t\geqslant x)\\&=2\mathsf{P}(B_1\geqslant x)=\frac{2}{\sqrt{2\pi}}\int_x^{\infty}e^{-t^2/2}\,\mathrm{d}t,\end{aligned}$$

而

$$\frac{2}{\sqrt{2\pi}}\int_x^{\infty} e^{-t^2/2}\,\mathrm{d}t \leqslant \int_x^{\infty}\frac{t}{x}e^{-t^2/2}\,\mathrm{d}t$$
$$= \frac{1}{x}\int_x^{\infty} e^{-t^2/2}\,\mathrm{d}\tfrac{t^2}{2} = \frac{1}{x}e^{-x^2/2},$$

于是, 对任意 $\varepsilon > 0$,

$$\sum_{n=1}^{\infty}\mathsf{P}(\sup_{t\in[n,n+1]}(B_t - B_n) \geqslant n\varepsilon) \leqslant \sum_{n=1}^{\infty}\frac{1}{n\varepsilon}e^{-(n\varepsilon)^2/2} < \infty.$$

因此,

$$\lim_{n\to\infty}\frac{1}{n}\sup_{t\in[n,n+1]}(B_t - B_n) = 0 \quad a.s..$$

即

$$\lim_{t\to 0} tB_{1/t} = 0,$$

故 W_t 在 $t = 0$ 处连续. ■

前面我们提到过, 布朗运动是一种重要的鞅. 下面, 讨论与布朗运动相关的鞅的性质.

定理 3.2.8 一维标准布朗运动是连续鞅.

证明: 若 $B = (B_t)_{t\geqslant 0}$ 是 $(\varOmega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, 故

$$\mathsf{E}[B_t] = 0, \qquad \mathsf{E}[B_t^2] = t,$$

故平方可积. 进一步, 当 $t > s$ 时,

$$\mathsf{E}[B_t \mid \mathcal{F}_s] = \mathsf{E}[B_t - B_s + B_s \mid \mathcal{F}_s] = B_s + \mathsf{E}[B_t - B_s] = B_s.$$

故 B 是连续鞅. ■

例 3.2.7 若 $B = (B_t)_{t\geqslant 0}$ 是 $(\varOmega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, Z 是一个不可积的随机变量, 且关于 $\mathcal{F}_0$ 可测, $M_t = B_t + Z$, $M = (M_t)_{t\geqslant 0}$ 是一个局部鞅, 但不是鞅.

例 3.2.8 若 $B = (B_t)_{t\geqslant 0}$ 是 $(\varOmega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, $a < 0 < b$ 是常数, 令 $\tau = \inf\{t : B_t \geqslant b \text{ 或 } B_t \leqslant a\}$, 求 $\mathsf{P}(B_\tau = a)$ 及 $\mathsf{P}(B_\tau = b)$.

解: 由于布朗运动从零点出发, 由定理 3.2.4 证明中的讨论, 布朗运动总会在有限时刻内到达 a 或者 b, 且在集合 $\{t \leqslant \tau\}$ 上 $|B_t| \leqslant |a| + b$.

由假设, $\mathsf{P}(B_\tau = a) + \mathsf{P}(B_\tau = b) = 1$ 且 $(B_{\tau\wedge t})$ 是鞅. 由控制收敛定理,

$$\lim_{t\to\infty} \mathsf{E}[B_{\tau\wedge t}] = \mathsf{E}[B_\tau] = 0,$$

故 $a\mathsf{P}(B_\tau = a) + b\mathsf{P}(B_\tau = b) = 0$. 于是

$$\mathsf{P}(B_\tau = a) = \frac{b}{|a| + b},$$

$$\mathsf{P}(B_\tau = b) = \frac{|a|}{|a| + b}.$$

■

除此之外, 布朗运动还有如下性质.

定理 3.2.9 *若 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, 则 $(B_t^2 - t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的鞅.*

证明: 记 $t > s$. 考虑

$$\begin{aligned}
\mathsf{E}[B_t^2 - t \mid \mathcal{F}_s] &= \mathsf{E}[B_t^2 \mid \mathcal{F}_s] - t \\
&= \mathsf{E}[(B_t - B_s)^2 + 2(B_t - B_s)B_s + B_s^2 \mid \mathcal{F}_s] - t \\
&= B_s^2 + t - s - t + 2B_s\mathsf{E}[(B_t - B_s) \mid \mathcal{F}_s] \\
&= B_s^2 - s,
\end{aligned}$$

故 $(B_t^2 - t)_{t\geqslant 0}$ 是鞅. ■

定理 3.2.10 *若 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, 则 $M_t = \exp\{\lambda B_t - \frac{\lambda^2}{2}t\}$ 是鞅, 这里 $\lambda \in \mathbb{R}$.*

证明: 令 $t > s$, 考虑

$$\mathsf{E}[M_t \mid \mathcal{F}_s] = \frac{\mathsf{E}[\exp\{\lambda B_t\} \mid \mathcal{F}_s]}{e^{\frac{\lambda^2}{2}t}}.$$

注意到

$$\begin{aligned}
\mathsf{E}[\exp\{\lambda B_t\} \mid \mathcal{F}_s] &= \mathsf{E}[\exp\{\lambda(B_t - B_s)\exp\{\lambda B_s\} \mid \mathcal{F}_s] \\
&= \exp\{\lambda B_s\}\mathsf{E}[\exp\{\lambda(B_t - B_s)\} \mid \mathcal{F}_s] \\
&= \exp\{\lambda B_s\}e^{\frac{\lambda^2}{2}(t-s)}.
\end{aligned}$$

故

$$\mathsf{E}[M_t \mid \mathcal{F}_s] = \exp\{\lambda B_s - \frac{\lambda^2}{2}s\}.$$

■

事实上, 上述 λ 用虚数 iλ 代替, 我们会得到 $\widetilde{M}_t = \exp\{\mathrm{i}\lambda B_t + \dfrac{\lambda^2}{2}t\}$ 是鞅. 更加令人惊奇的是, 这是关于布朗运动的一个充分必要条件.

定理 3.2.11 *若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的适应过程, 则下列论述等价:*

(1)B 是标准布朗运动;

(2)B 是连续的鞅, 且对任意的实数 λ, $\exp\{\mathrm{i}\lambda B_t + \dfrac{\lambda^2}{2}t\}$ 是鞅.

证明 [6]: 上面得到了 (1) $\Longrightarrow$ (2). 下面证明 (2) $\Longrightarrow$ (1). 由 (2) 可知, $t>s$ 时

$$\mathsf{E}[\exp\{\mathrm{i}\lambda B_t + \frac{\lambda^2}{2}t\} \mid \mathcal{F}_s] = \exp\{\mathrm{i}\lambda B_s + \frac{\lambda^2}{2}s\}.$$

于是

$$\mathsf{E}[\exp\{\mathrm{i}\lambda(B_t - B_s)\} \mid \mathcal{F}_s] = \exp\{-\frac{\lambda^2}{2}(t-s)\}.$$

故对于 $A\in\mathcal{F}_s$

$$\mathsf{E}[\mathbb{1}_A \exp\{\mathrm{i}\lambda(B_t - B_s)\}] = \mathsf{P}(A)\exp\{-\frac{\lambda^2}{2}(t-s)\}.$$

取 $A=\Omega$ 得

$$B_t - B_s \sim N(0, t-s).$$

此外, 取 $A\in\mathcal{F}_s$, $\mathsf{P}(A)>0$, 令

$$\mathsf{Q}(\cdot) = \frac{\mathsf{P}(\cdot\bigcap A)}{\mathsf{P}(A)},$$

则

$$\mathsf{E}_{\mathsf{Q}}[\exp\{\mathrm{i}\lambda(B_t - B_s)\}] = \exp\{-\frac{\lambda^2}{2}(t-s)\}.$$

即 $B=(B_t)_{t\geqslant 0}$ 有 $B_t - B_s$ 在 Q 下的分布与在 P 下的分布相同, 与 A 无关.
故对任意可测函数 f

$$\mathsf{E}_{\mathsf{Q}}[f(B_t - B_s)] = \mathsf{E}[f(B_t - B_s)].$$

即

$$\mathsf{E}[\mathbb{1}_A f(B_t - B_s)] = \mathsf{P}(A)\mathsf{E}[f(B_t - B_s)],$$

即 $B_t - B_s$ 与 $\mathcal{F}_s$ 独立, 从而证明了 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动. ■

鞅的性质往往与停止定理密切相关, 下面的结论往往被称为 Wald 等式.

定理 3.2.12 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动, τ 是停时, 且 $\mathsf{E}[\tau]<\infty$, 则 $\mathsf{E}[B_\tau]=0$, $\mathsf{E}[B_\tau^2]=\mathsf{E}[\tau]$.

证明: 由定理 3.1.3, $(B_{\tau\wedge t})_{t\geqslant 0}$ 是鞅, 因此 $\mathsf{E}[B_{\tau\wedge t}]=0$, 同时 $(B_{\tau\wedge t}^2-\tau\wedge t)_{t\geqslant 0}$ 是鞅, $\mathsf{E}[B_{\tau\wedge t}^2]=\mathsf{E}[\tau\wedge t]$. 注意到对于 $n\geqslant 1$,

$$\mathsf{E}[B_{\tau\wedge n}^2]=\mathsf{E}[\tau\wedge n]\leqslant\mathsf{E}[\tau]<\infty.$$

因此 $(B_{\tau\wedge n})_{n\geqslant 1}$ 是一致可积的, 故

$$0=\lim_{n\to\infty}\mathsf{E}[B_{\tau\wedge n}]=\mathsf{E}[B_\tau].$$

同时，利用定理 3.1.2,

$$\mathsf{E}[B_\tau|\mathcal{F}_{\tau\wedge n}]=B_{\tau\wedge n}.$$

因此, 由 Jensen 不等式,

$$\begin{aligned}\mathsf{E}[(B_{\tau\wedge n})^2] &= \mathsf{E}[(\mathsf{E}[B_\tau|\mathcal{F}_{\tau\wedge n}])^2]\\ &\leqslant \mathsf{E}[\mathsf{E}[(B_\tau)^2|\mathcal{F}_{\tau\wedge n}]]\\ &= \mathsf{E}[B_\tau^2].\end{aligned}$$

因此

$$\mathsf{E}[\tau\wedge n]=\mathsf{E}[(B_{\tau\wedge n})^2]\leqslant\mathsf{E}[B_\tau^2].$$

由单调收敛定理,

$$\mathsf{E}[\tau]\leqslant\mathsf{E}[B_\tau^2].$$

因此，$\mathsf{E}[B_\tau^2]=\mathsf{E}[\tau]$. ■

本节的最后, 我们来证明关于布朗运动离散化的结论. 在现代金融统计的研究中, 下面的结论引发了关于已实现波动率的研究.

设 $D=\{0=t_0<t_1<\cdots<t_n=t\}$ 是区间 $[0,t]$ 的有限划分, 且设

$$V_D=\sum_{l=1}^n(B_{t_l}-B_{t_{l-1}})^2,\quad m(D)=\max_l|t_l-t_{l-1}|.$$

定理 3.2.13 若 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 则对任意 t,

$$V_D\xrightarrow{\mathsf{P}}t.$$

证明: 事实上,

$$\mathsf{E}[V_D] = \sum_{l=1}^{n} \mathsf{E}[B_{t_l} - B_{t_{l-1}}]^2 = \sum_{l=1}^{n}(t_l - t_{l-1}) = t.$$

于是, 要证明对于任意的 $\varepsilon > 0$

$$\mathsf{P}(|V_D - \mathsf{E}[V_D]| \geqslant \varepsilon) \to 0.$$

注意到

$$\mathsf{P}(|V_D - \mathsf{E}[V_D]| \geqslant \varepsilon) \leqslant \frac{\mathsf{E}|V_D - \mathsf{E}[V_D]|^2}{\varepsilon^2},$$

故只需要证明 $\mathsf{E}|V_D - \mathsf{E}[V_D]|^2 \to 0$. 我们来看 $\mathsf{E}|V_D - \mathsf{E}[V_D]|^2$:

$$\begin{aligned}
& \mathsf{E}[(V_D - \mathsf{E}[V_D])^2] \\
= \ & \mathsf{E}[(\sum_{l=1}^{n} |B_{t_l} - B_{t_{l-1}}|^2 - t)^2] \\
= \ & \mathsf{E}[(\sum_{l=1}^{n}(B_{t_l} - B_{t_{l-1}})^2 - \sum_{l=1}^{n}(t_l - t_{l-1}))^2] \\
= \ & \sum_{l=1}^{n} \mathsf{E}[((B_{t_l} - B_{t_{l-1}})^2 - (t_l - t_{l-1}))^2] \\
& + \sum_{k\neq l}^{n} \mathsf{E}[((B_{t_l} - B_{t_{l-1}})^2 - (t_l - t_{l-1}))((B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1}))].
\end{aligned}$$

由于布朗运动具有独立增量性, 且 $(B_t^2 - t)_{t\geqslant 0}$ 是鞅, 故

$$\sum_{k\neq l}^{n} \mathsf{E}[((B_{t_l} - B_{t_{l-1}})^2 - (t_l - t_{l-1}))((B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1}))] = 0.$$

于是

$$\begin{aligned}
\mathsf{E}(V_D - \mathsf{E}[V_D])^2 &= \sum_{l=1}^{n} \mathsf{E}[(B_{t_l} - B_{t_{l-1}})^4 - 2(t_l - t_{l-1})(B_{t_l} - B_{t_{l-1}})^2 + (t_l - t_{l-1})^2] \\
&= \sum_{l=1}^{n} \mathsf{E}[(B_{t_l} - B_{t_{l-1}})^4 - 2(t_l - t_{l-1})^2 + (t_l - t_{l-1})^2] \\
&= 2\sum_{l=1}^{n}(t_l - t_{l-1})^2.
\end{aligned}$$

因此

$$\mathsf{E}[(V_D - \mathsf{E}[V_D])^2] \leqslant 2m(D)\sum_{l=1}^{n}(t_l - t_{l-1}) = 2tm(D) \to 0.$$

■

3.3 布朗运动与嵌入定理

在第二节的介绍中, 我们发现布朗运动是随机游动某种意义下的极限. 在这一节中, 从另一个角度, 来探究随机游动与布朗运动之间的关系.

设 $\{X_n\}$ 是一列独立同分布的随机变量, $\mathsf{P}(X_1 = 1) = \mathsf{P}(X_1 = -1) = \frac{1}{2}$,

$$S_n = X_1 + X_2 + \cdots + X_n.$$

这里, S_n 是从零点出发的随机游动. 这一节的任务是考虑 $(S_n)_{n\geqslant 0}$ 与布朗运动 $(B_t)_{t\geqslant 0}$ 之间的关系.

为了讨论得更充分, 不再假设 $\mathsf{P}(X_1 = 1) = \mathsf{P}(X_1 = -1) = \frac{1}{2}$, 只考虑 $\{X_n\}$ 是一列独立同分布的随机变量, 满足 $\mathsf{E}[X_1] = 0, \mathsf{E}[X_1^2] = 1$.

研究 $(S_n)_{n\geqslant 0}$ 与布朗运动 $(B_t)_{t\geqslant 0}$ 之间的关系的方式有多种. 人们熟知的是中心极限定理:

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1).$$

中心极限定理中, 仅仅是布朗运动的分布与 S_n 之间的关系. 这一节将介绍嵌入定理, 该定理可以将随机游动与布朗运动通过停时直接联系起来. 嵌入定理是 20 世纪概率论的杰出成果, 很多深刻优美的数学结论可以通过嵌入定理得到.

首先, 我们给出一个命题, 这是证明嵌入定理的关键.

命题 3.3.1 *设 $F(x)$ 是一个分布函数, 且 $\int_{\mathbb{R}} x\,\mathrm{d}F(x) = 0$, $0 < \int_{\mathbb{R}} x^2\,\mathrm{d}F(x) < +\infty$, 则存在布朗运动 B, 及停时 τ, 使得*

$$\mathsf{P}(B_\tau < x) = F(x), \qquad \mathsf{E}[\tau] = \int_{\mathbb{R}} x^2\,\mathrm{d}F(x).$$

证明: 设 B 是概率空间 $(\Omega_1, \mathcal{F}_1, \mathsf{P}_1)$ 上的标准布朗运动. Y 和 Z 是概率空间 $(\Omega_2, \mathcal{F}_2, \mathsf{P}_2)$ 上的随机变量, 使得 $Y \leqslant 0 \leqslant Z$, 且 (Y, Z) 的联合概率分布为:

$$F(y,z) = \begin{cases} \dfrac{1}{C}(z-y)dF(y)dF(z), & y \leqslant 0 \leqslant z, \\ 0, & \text{其他}, \end{cases}$$

其中 C 是正规化参数. 考虑乘积概率空间 $(\Omega,\mathcal{F},\mathsf{P})=(\Omega_1\times\Omega_2,\mathcal{F}_1\otimes\mathcal{F}_2,\mathsf{P}_1\otimes\mathsf{P}_2)$. 故 B 与 (Y,Z) 独立. 对 $t\geqslant 0$, 令

$$\mathcal{G}_t=\sigma(B_r,Y,Z;0\leqslant r\leqslant t),$$

$$\tau=\inf\{t:B_t\notin(Y,Z)\}.$$

由 $(\Omega,\mathcal{F},\mathsf{P})$ 的构造, τ 是停时, 对于 $x>0$,

$$\mathsf{P}(B_\tau>x)=\mathsf{E}[\mathsf{P}(B_\tau>x|Y,Z)].$$

下面计算 $\mathsf{P}(B_\tau>x|Y=y,Z=z)$ 由例 3.2.6, 当 $z<x$ 时, $\mathsf{P}(B_\tau>x|Y=y,Z=z)=0$. 当 $z\geqslant x$ 时,

$$\mathsf{P}(B_\tau>x|Y=y,Z=z)=\frac{-y}{z-y}.$$

故

$$\mathsf{P}(B_\tau>x)=\int_{-\infty}^{0}\int_{x}^{\infty}\frac{-y}{z-y}\,\mathrm{d}F(y,z)=1-F(x).$$

同理可得 $x<0$ 时, $\mathsf{P}(B_\tau<x)=F(x)$. 同时, 注意到 $(B_{\tau\wedge t}^2-(\tau\wedge t))_{t\geqslant 0}$ 是鞅, 利用控制收敛定理可得

$$\mathsf{E}[\tau]=\int_{\mathbb{R}}x^2\,\mathrm{d}F(x).$$

■

利用上面这个命题, 可以得到著名的嵌入定理.

定理 3.3.1　设 $\{X_n\}$ 是一列独立同分布的随机变量, 满足 $\mathsf{E}[X_1]=0,\mathsf{E}[X_1^2]=1$. 令 $S_0=0$, $S_n=X_1+X_2+\cdots+X_n$. 则存在标准布朗运动 $B=(B_t)_{t\geqslant 0}$, 及与 B 独立的正值随机变量 $(\tau_n)_{n\geqslant 1}$ 使得

(1) $(\tau_n)_{n\geqslant 1}$ 是独立同分布的随机变量;

(2) $(B_{\sum_{i=1}^n\tau_i})_{n\geqslant 1}$ 与 $(S_n)_{n\geqslant 1}$ 同分布.

证明: 设 B 是概率空间 $(\Omega_0,\mathcal{F}_0,\mathsf{P}_0)$ 上的标准布朗运动. 对于任意的 $n\geqslant 1$, (Y_n,Z_n) 是概率空间 $(\Omega_n,\mathcal{F}_n,\mathsf{P}_n)$ 上的随机向量, 使得 $Y_n\leqslant 0\leqslant Z_n$, 且对于任意的 $n\geqslant 1$, (Y_n,Z_n) 的联合概率分布均为 $F(y,z)$.

$$F(y,z)=\begin{cases}\dfrac{1}{C}(z-y)dF(y)dF(z), & y\leqslant 0\leqslant z,\\ 0, & \text{其他},\end{cases}$$

这里 $F(y) = \mathsf{P}(X_1 \leqslant y)$. 定义概率空间

$$(\Omega, \mathcal{F}, \mathsf{P}) = (\Omega_0 \times \Omega_1 \times \Omega_2 \times \cdots, \mathcal{F}_0 \otimes \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes \cdots, \mathsf{P}_0 \otimes \mathsf{P}_1 \otimes \mathsf{P}_2 \otimes \cdots).$$

在新的概率空间上, B 与 (Y_n, Z_n) 是相互独立的, 这里 $n \geqslant 1$.

由前面命题, 存在停时 τ_1, 使得 B_{τ_1} 与 X_1 同分布. 考虑 $(B_{\tau_1+t} - B_{\tau_1})_{t\geqslant 0}$, 由布朗运动的强马氏性,$(B_{\tau_1+t} - B_{\tau_1})_{t\geqslant 0}$ 是布朗运动, 进一步存在停时 τ_2, 使得 $B_{\tau_1+\tau_2} - B_{\tau_1}$ 与 X_2 同分布. 这样依次做下去, 存在一列停时 $(\tau_n)_{n\geqslant 1}$ 使得 $(\tau_n)_{n\geqslant 1}$ 是独立同分布的随机变量, 且

$$B_{\tau_1}, B_{\tau_1+\tau_2} - B_{\tau_1}, \cdots, B_{\sum_{i=1}^n \tau_i} - B_{\sum_{i=1}^{n-1} \tau_i}, \cdots,$$

与

$$X_1, X_2, \cdots, X_n, \cdots$$

同分布. 故 $(B_{\sum_{i=1}^n \tau_i})_{n\geqslant 1}$ 与 $(S_n)_{n\geqslant 1}$ 同分布. ■

利用嵌入定理, Strassen 在 1964 年首次得到了关于独立随机变量和的强不变原理, 匈牙利数学家 Komlós , Major, Tusnády 在 1975 年和 1976 年两个工作中得到了强不变原理的最佳收敛速度. 我们下面给出强不变原理, 其证明读者可参考 Csörgö 和 Révész [1] 或林正炎, 陆传荣, 苏中根 [5].

定理 3.3.2 (强不变原理) 设 $\{X_n\}$ 是一列独立同分布的随机变量, 满足 $\mathsf{E}[X_1] = 0, \mathsf{E}[X_1^2] = 1$. 令 $S_0 = 0$, $S_n = X_1 + X_2 + \cdots + X_n$. 可构造新的概率空间, 其上存在一个标准布朗运动 B 和一个新的随机变量序列 $(\widetilde{S}_n)_{n\geqslant 1}$, 使得

(1) $(\widetilde{S}_n)_{n\geqslant 1}$ 与 $(S_n)_{n\geqslant 1}$ 同分布;

(2)

$$\lim_{n\to\infty} \frac{|\widetilde{S}_n - B_n|}{\sqrt{n \log\log n}} = 0 \quad a.s..$$

相对于强不变原理, 弱不变原理讨论的对象是不一样的, 而且也是在弱收敛意义下进行讨论的. 设 $(X_n)_{n\geqslant 1}$ 是一系列独立同分布的随机变量, 满足 $\mathsf{E}[X_1] = 0, \mathsf{E}[X_1^2] = 1$. 令 $S_0 = 0$, $S_n = X_1 + X_2 + \cdots + X_n$.

$$W_t^{(n)}(\omega) = \frac{1}{\sqrt{n}} S_{[nt]}(\omega) + (nt - [nt]) \frac{1}{\sqrt{n}} X_{[nt]+1}(\omega).$$

考虑 W^n 作为概率空间 $(\Omega, \mathcal{F}, \mathsf{P})$ 到 $\mathbb{C}[0,1]$ ($[0,1]$ 上连续函数全体在一致距离下的完备可分空间) 上的随机元. 弱不变原理告诉我们, 在弱收敛的意义下, W^n 收敛至布朗运动, 本质上讲, 该收敛是无穷维空间上的弱收敛结果. 与前面的强不变原理有本质区别. 关于弱不变原理, 感兴趣的读者可参考 Jacod 和 Shiryaev [3] 或林正炎, 陆传荣, 苏中根 [5].

3.4 习题

1. 设 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动.

(1) 求 $\mathsf{P}(B_3>\frac{1}{2})$;

(2) 求 $\mathsf{P}(B_1<\frac{1}{2},B_3>B_1+2)$;

(3) 求 $t=10$ 之前, 布朗运动的轨道始终在 6 以下的概率;

(4) 求 $\mathsf{P}(B_4\leqslant 0|B_2\geqslant 0)$.

2. 设 $B=(B_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动, 令 $\mathcal{F}_t=\sigma(B_r,0\leqslant r\leqslant t)$. 若 $s<t$,

(1) 求 $\mathsf{E}[B_t^2|\mathcal{F}_s]$;

(2) 求 $\mathsf{E}[e^{4B_t-2}|\mathcal{F}_s]$;

(3) 求 $\mathsf{E}[e^{B_tB_s}|\mathcal{F}_s]$.

3. 假定一只股票价格按初值为发行价的标准布朗运动 $B=(B_t)_{t\geqslant 0}$. 某人以价格 $b+c$ 买入, 而现在的价格恰好为 b, 这里 $c>0$. 若此人计划在股票价格回到 $b+c$ 或最多再观望 T 时间后将股票抛出, 问此人不能重新获得买入价的概率.

4. 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 求 $\mathsf{P}(B_2>0\mid B_1>0)$.

5. 在甲, 乙两人的自行车比赛中, 以 $B=(B_t)_{t\geqslant 0}$ 表示当完成 $100t\%$ 的比赛路程时甲领先的时间 (以秒记). 假设 B 是原点出发的标准布朗运动, 若甲以 1 秒领先赢得比赛, 问甲在路程中点领先的概率是多大?

6. 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 求 $\mathsf{E}[B_2|B_1,B_3]$.

7. 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 求 $\mathsf{E}[\int_0^2 B_r\,\mathrm{d}r|B_1]$.

8. 设 $a<0<b$, T_a,T_b 分别是原点出发的布朗运动首次到达 a,b 的时间, 令 $T=T_a\wedge T_b$. 求:

(1) $\mathsf{P}(T_a<T_b)$;

(2) $\mathsf{E}T$.

9. 若 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 请证明

$$\sum_{j=1}^{2^n}[B_{\frac{j}{2^n}t}-B_{\frac{j-1}{2^n}t}]^2\to t \quad a.s..$$

10. 若 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 请证明 $(B_t^4-6tB_t^2+3t^2)_{t\geqslant 0}$ 是鞅.

11. 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, $a>0$,

$$\tau=\inf\{t\geqslant 0,|B_t|=a\}.$$

请证明: $\mathsf{E}[\tau^2] = 5a^4/3$.

12. 若 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, τ 是停时, 且 $\mathsf{E}[\tau] < \infty$, 请证明:$\mathsf{E}[\tau^2] \leqslant 4\mathsf{E}[B_\tau^4]$.

第 4 章　Itô 积分

本章将介绍 Itô 积分相关知识.

4.1　引论

1944 年, 日本数学家伊藤清 (Itô Kiyoshi) 率先针对布朗运动给出了 Itô 积分的定义. 1951 年, Itô 给出了 Itô 公式. Itô 开启了随机分析这一随机过程的分支. 近八十年以来, 随机分析学科迅猛发展, 同时, 在金融资产定价的研究中扮演了十分关键的角色. 从这一章开始, 我们开始介绍 Itô 型随机积分 (以下简称 Itô 积分).

在上一章的讨论中, 我们已经知道, 对于标准布朗运动 $B = (B_t)_{t\geqslant 0}$ 而言, 其轨道是几乎处处不可微的. 因此, 不可能针对布朗运动建立类似于 Lebesgue-Stieltjes 型轨道积分. 事实上, 任何一个非平凡的连续鞅都是具有无界变差轨道的.

命题 4.1.1　*若 M 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的连续鞅, 若 M 是具有有限变差的轨道, 则 $M \equiv M_0$ a.s..*

证明: 不妨在 $M_0 \equiv 0$ 下考虑, 下证 $M \equiv 0$ a.s..

设

$$V_t = \sup_n \sum_{l=1}^{n} |M_{t_l} - M_{t_{l-1}}|,$$

$$M_0 = 0,$$

$$D = \{0 = t_0 < t_1 < \cdots < t_n = t\}.$$

若 M 具有有限变差, 即 $V = (V_t)_{t\geqslant 0}$ 是一个随机过程. 令

$$S_n = \inf\{s : V_s \geqslant n\},$$

考虑 M^{S_n} 是具有有界变差的, 若我们能对 M^{S_n} 证明其是常数, 令 $n \to \infty$, 便能证明 M 是常数, 因此只考虑 M 的全变差被常数 K 控制的情形.

事实上, 由于 M 是鞅, 故

$$\mathsf{E}[M_t^2]=\mathsf{E}[\sum_{i=1}^{n}(M_{t_i}^2-M_{t_{i-1}}^2)]=\mathsf{E}[\sum_{i=1}^{n}(M_{t_i}-M_{t_{i-1}})^2].$$

由于 M 的全变差被常数 K 控制, 且 M 是连续的,

$$\begin{aligned}\mathsf{E}[M_t^2]&\leqslant \mathsf{E}[V_t\sup_i|M_{t_i}-M_{t_{i-1}}|]\\&\leqslant K\mathsf{E}[\sup_i|M_{t_i}-M_{t_{i-1}}|]\to 0,\end{aligned}$$

故 $M_t\equiv 0$, 从而 $M\equiv 0$ a.s..

■

因此, 针对非平凡的连续鞅不可能建立 Lebesgue-Stieltjes 型轨道积分.

4.2 二次变差过程

根据上一节的讨论，我们知道非平凡的连续鞅不可能具备有限变差. 具体来讲，若 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的非平凡的连续鞅, 对于 $t>0$, 若 $0=t_0^n<t_1^n<\cdots<t_{p_n}^n=t$ 且 $\max\limits_i|t_i^n-t_{i-1}^n|\to 0$, 则

$$\lim_{n\to\infty}\sum_{i=1}^{p_n}|M_{t_i^n}-M_{t_{i-1}^n}|=\infty.$$

如果考虑更深一点的问题，

$$\sum_{i=1}^{p_n}(M_{t_i^n}-M_{t_{i-1}^n})^2$$

的极限会是什么?

在这一节中，我们考虑这个问题. 上式的极限事实上就是所谓的二次变差过程. 我们对更广的一类过程进行讨论.

首先引入一类过程: 平方可积鞅. 若 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅. 若 $\sup\limits_t\mathsf{E}M_t^2<+\infty$, 则称 M 是平方可积鞅. 若存在一列停时 $(T_n)_{n\geqslant 0}$ 且 $\lim\limits_{n\to\infty}T_n=\infty$, 使 M^{T_n} 是平方可积鞅, 则称 M 是局部平方可积鞅. 局部平方可积鞅全体记为 $\mathscr{M}_{loc}^2$, 平方可积鞅全体记为 $\mathscr{M}^2$.

下面给出一个基础性结论.

定理 4.2.1 令 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续局部鞅, 则存在一个唯一的连续增过程 $\langle M,M\rangle$, 使 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是一个连续局部鞅. 进一步, 对于 $t>0$, 若 $0=t_0^n<t_1^n<\cdots<t_{p_n}^n=t$ 且 $\max\limits_i|t_i^n-t_{i-1}^n|\to 0$, 则有

$$\sum_{i=1}^{p_n}(M_{t_i^n}-M_{t_{i-1}^n})^2\xrightarrow{\mathsf{P}}\langle M,M\rangle_t.$$

证明 [12]: 首先证明唯一性. 若存在两个连续增过程 $\langle M,M\rangle,\langle M,M\rangle'$, 使得 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是连续鞅, $(M_t^2-\langle M,M\rangle_t')_{t\geqslant 0}$ 是连续鞅. 故 $(\langle M,M\rangle_t-\langle M,M\rangle_t')_{t\geqslant 0}$ 是连续鞅, 且 $\langle M,M\rangle_0-\langle M,M\rangle_0'=0$. 此时, $\langle M,M\rangle_t'-\langle M,M\rangle_t$ 是连续的鞅且具有有限变差的轨道. 故

$$\langle M,M\rangle_t'-\langle M,M\rangle_t=0.$$

于是 $\langle M,M\rangle$ 的唯一性得证.

下面证明存在性. 首先假设 $M=(M_t)_{t\geqslant 0}$ 是一个连续的一致有界鞅.

设 Δ 是直线上的一组分点:

$$\Delta:0=t_0<t_1<\cdots<t_n<\cdots.$$

当 $t_k\leqslant t<t_{k+1}$ 时，定义

$$T_t^\Delta=\sum_{i=1}^{k-1}(M_{t_i}-M_{t_{i-1}})^2+(M_t-M_{t_k})^2.$$

考虑 $t_k<s<t<t_{k+1}$,

$$\begin{aligned}\mathsf{E}[T_t^\Delta-T_s^\Delta|\mathcal{F}_s] &= \mathsf{E}[(M_t-M_{t_k})^2-(M_s-M_{t_k})^2|\mathcal{F}_s]\\ &= \mathsf{E}[M_t^2-M_s^2-2M_tM_{t_k}+2M_sM_{t_k}|\mathcal{F}_s],\end{aligned}$$

注意到

$$\mathsf{E}[M_tM_{t_k}|\mathcal{F}_s]=M_{t_k}\mathsf{E}[M_t|\mathcal{F}_s]=M_sM_{t_k}.$$

因此

$$\mathsf{E}[T_t^\Delta-T_s^\Delta|\mathcal{F}_s]=\mathsf{E}[M_t^2-M_s^2|\mathcal{F}_s].$$

考虑 $t_{k-1} < s < t_k < t < t_{k+1}$,

$$\begin{aligned}
& \mathsf{E}[T_t^\Delta - T_s^\Delta | \mathcal{F}_s] \\
= \ & \mathsf{E}[(M_t - M_{t_k})^2 + (M_{t_k} - M_{t_{k-1}})^2 - (M_s - M_{t_{k-1}})^2 | \mathcal{F}_s] \\
= \ & \mathsf{E}[M_t^2 - 2M_t M_{t_k} + M_{t_k}^2 + M_{t_k}^2 - 2M_{t_k} M_{t_{k-1}} + M_{t_{k-1}}^2 | \mathcal{F}_s] \\
& - \mathsf{E}[M_s^2 - 2M_s M_{t_{k-1}} + M_{t_{k-1}}^2 | \mathcal{F}_s]
\end{aligned}$$

注意到

$$\mathsf{E}[M_{t_k} M_{t_{k-1}} | \mathcal{F}_s] = M_{t_{k-1}} \mathsf{E}[M_{t_k} | \mathcal{F}_s] = M_s M_{t_{k-1}},$$

$$\begin{aligned}
\mathsf{E}[M_t M_{t_k} | \mathcal{F}_s] &= \mathsf{E}[\mathsf{E}[M_t M_{t_k} | \mathcal{F}_{t_k}] | \mathcal{F}_s] \\
&= \mathsf{E}[M_{t_k} E[M_t | \mathcal{F}_{t_k}] | \mathcal{F}_s] \\
&= \mathsf{E}[M_{t_k}^2 | \mathcal{F}_s].
\end{aligned}$$

因此

$$\mathsf{E}[T_t^\Delta - T_s^\Delta | \mathcal{F}_s] = \mathsf{E}[M_t^2 - M_s^2 | \mathcal{F}_s].$$

我们可以得到 $M^2 - T^\Delta$ 是鞅. 记

$$|\Delta| = \sup_{i \geqslant 1} |t_i - t_{i-1}|.$$

我们接下来证明当 $|\Delta| \to 0$, T^Δ 在 L^2 意义下有极限.

设 Δ, Δ' 是两组直线上的分点，$\Delta\Delta'$ 是两组分点合并之后的分点. 记 $X = T^\Delta - T^{\Delta'}$, 对于 $a > 0$,

$$T_a^{\Delta\Delta'}(X) = \sum_{i=1}^{l-1} (X_{a_i} - X_{a_{i-1}})^2 + (X_a - X_{a_l})^2.$$

这里

$$\Delta\Delta' : 0 = a_0 < \cdots < a_l < \cdots .$$

由上面的讨论可知

$$(T^\Delta - T^{\Delta'})^2 - T^{\Delta\Delta'}(X)$$

是鞅. 因此, 对 $a \geqslant 0$,

$$\mathsf{E}[(T_a^\Delta - T_a^{\Delta'})^2] = \mathsf{E}[T_a^{\Delta\Delta'}(X)].$$

为证明当 $|\Delta| \to 0$, T^{Δ} 在 L^2 意义下有极限, 利用柯西准则, 我们可以证明: 当 $|\Delta| \to 0, |\Delta'| \to 0$,

$$\mathsf{E}[(T_a^{\Delta} - T_a^{\Delta'})^2] \to 0.$$

即证明:

$$\mathsf{E}[T_a^{\Delta\Delta'}(X)] \to 0.$$

注意到:

$$\begin{aligned}
T_a^{\Delta\Delta'}(X) &= \sum_{i=1}^{l-1}(X_{a_i} - X_{a_{i-1}})^2 + (X_a - X_{a_l})^2 \\
&= \sum_{i=1}^{l-1}(T_{a_i}^{\Delta} - T_{a_i}^{\Delta'} - T_{a_{i-1}}^{\Delta} + T_{a_{i-1}}^{\Delta'})^2 \\
&\quad +(T_a^{\Delta} - T_a^{\Delta'} - T_{a_l}^{\Delta} + T_{a_l}^{\Delta'})^2 \\
&\leqslant 2[\sum_{i=1}^{l-1}(T_{a_i}^{\Delta} - T_{a_{i-1}}^{\Delta})^2 + (T_a^{\Delta} - T_{a_l}^{\Delta})^2] \\
&\quad +2[\sum_{i=1}^{l-1}(T_{a_i}^{\Delta'} - T_{a_{i-1}}^{\Delta'})^2 + (T_a^{\Delta'} - T_{a_l}^{\Delta'})^2]
\end{aligned}$$

即

$$T_a^{\Delta\Delta'}(X) \leqslant 2T_a^{\Delta\Delta'}(T^{\Delta}) + 2T_a^{\Delta\Delta'}(T^{\Delta'}).$$

我们只需证明

$$\mathsf{E}[T_a^{\Delta\Delta'}(T^{\Delta})] \to 0.$$

设最靠近 a_j 的 Δ 的分点为 t_n, 满足 $t_n \leqslant a_i < a_{i+1} \leqslant t_{n+1}$,
此时

$$\begin{aligned}
T_{a_{i+1}}^{\Delta} - T_{a_i}^{\Delta} &= (M_{a_{i+1}} - M_{t_n})^2 - (M_{a_i} - M_{t_n})^2 \\
&= (M_{a_{i+1}} + M_{a_i} - 2M_{t_n})(M_{a_{i+1}} - M_{a_i})
\end{aligned}$$

因此

$$T_a^{\Delta\Delta'}(T^{\Delta}) \leqslant \sup_i |M_{a_{i+1}} + M_{a_i} - 2M_{t_n}|^2 T_a^{\Delta\Delta'}.$$

由 Schwarz 不等式

$$\mathsf{E}[T_a^{\Delta\Delta'}(T^{\Delta})] \leqslant (\mathsf{E}[\sup_i |M_{a_{i+1}} + M_{a_i} - 2M_{t_n}|^4])^{1/2}(\mathsf{E}[(T_a^{\Delta\Delta'})^2])^{1/2}.$$

由于 M 是连续的,且一致有界,我们设 $|M_t| \leqslant C$, 因此,只需要证明 $\mathsf{E}[(T_a^{\Delta\Delta'})^2]$ 有界，且界与 Δ 无关, 我们便可证明 $\mathsf{E}[T_a^{\Delta\Delta'}(T^{\Delta})] \to 0$.

对于 Δ, 为方便起见, 假设 $a = t_n$,

$$\begin{aligned}(T_a^{\Delta})^2 &= (\sum_{i=1}^{n}(M_{t_i} - M_{t_{i-1}})^2)^2 \\ &= \sum_{i=1}^{n}(M_{t_i} - M_{t_{i-1}})^4 + 2\sum_{i=1}^{n}(T_{t_n}^{\Delta} - T_{t_i}^{\Delta})(T_{t_i}^{\Delta} - T_{t_{i-1}}^{\Delta}).\end{aligned}$$

注意到

$$\mathsf{E}[(T_{t_n}^{\Delta} - T_{t_i}^{\Delta})|\mathcal{F}_{t_i}] = \mathsf{E}[(M_{t_n} - M_{t_i})^2|\mathcal{F}_{t_i}],$$

因此

$$\begin{aligned}\mathsf{E}[(T_a^{\Delta})^2] &= \sum_{i=1}^{n}\mathsf{E}[(M_{t_i} - M_{t_{i-1}})^4] + 2\sum_{i=1}^{n}\mathsf{E}[(T_{t_n}^{\Delta} - T_{t_i}^{\Delta})(T_{t_i}^{\Delta} - T_{t_{i-1}}^{\Delta})] \\ &= \sum_{i=1}^{n}\mathsf{E}[(M_{t_i} - M_{t_{i-1}})^4] \\ &\quad +2\sum_{i=1}^{n}\mathsf{E}[\mathsf{E}(T_{t_n}^{\Delta} - T_{t_i}^{\Delta})(T_{t_i}^{\Delta} - T_{t_{i-1}}^{\Delta})|\mathcal{F}_{t_i}]] \\ &= \sum_{i=1}^{n}\mathsf{E}[(M_{t_i} - M_{t_{i-1}})^4] \\ &\quad +2\sum_{i=1}^{n}\mathsf{E}[(M_{t_n} - M_{t_i})^2(T_{t_i}^{\Delta} - T_{t_{i-1}}^{\Delta})]\end{aligned}$$

$$\begin{aligned}\mathsf{E}[(T_a^{\Delta})^2] &\leqslant \mathsf{E}[(\sup_k |M_{t_n} - M_{t_k}|^2 + \sup_k |M_{t_k} - M_{t_{k-1}}|^2)T_a^{\Delta}] \\ &\leqslant 12C^2\mathsf{E}[T_a^{\Delta}] \leqslant 48C^4.\end{aligned}$$

至此, 我们便证明了对于一列 $\{\Delta_n\}$, 若 $|\Delta_n| \to 0$, $T_a^{\Delta_n}$ 在 L^2 意义下收敛到一个极限，此极限我们记为 $\langle M, M\rangle_a$. 即

$$\mathsf{E}[|T_a^{\Delta_n} - \langle M, M\rangle_a|^2] \to 0.$$

我们现在说明：$\langle M, M\rangle_a$ 具有一系列性质.

我们设 Δ_n 不断到加细, 即 Δ_n 的分点仍然在 Δ_{n+1} 中. 同时, 我们还要求 $\bigcup_n \Delta_n$ 在 $[0,a]$ 中稠密.

$T^{\Delta_n} - T^{\Delta_m}$ 是鞅, 由 Doob 不等式

$$\mathsf{E}[\sup_{t\leqslant a}|T_t^{\Delta_n} - T_t^{\Delta_m}|^2] \leqslant \mathsf{E}[|T_a^{\Delta_n} - T_a^{\Delta_m}|^2].$$

存在 Δ_n 的子列，使得 T^{Δ_n} 几乎处处收敛，同时由上式, 可知此收敛在 $[0,a]$ 上一致收敛, 因此可得 $\langle M,M\rangle$ 的连续性.

取 $s<t$, $s,t\in\bigcup_n \Delta_n$, 存在 n_0, $n\geqslant n_0$, $s,t\in\Delta_n$, 此时 $T_s^{\Delta_n}\leqslant T_t^{\Delta_n}$，因此 $\langle M,M\rangle_s \leqslant \langle M,M\rangle_t$. 由于 $\bigcup_n \Delta_n$ 在 $[0,a]$ 中稠密, 又由于 $\langle M,M\rangle$ 是连续的, 因此 $\langle M,M\rangle$ 单调递增.

由于 $M^2 - T^{\Delta_n}$ 是鞅, 考虑 $s<t$, 对于 $A\in\mathcal{F}_s$,

$$\int_A M_t^2 - T_t^{\Delta_n} dP = \int_A M_s^2 - T_s^{\Delta_n} d\mathsf{P},$$

由于

$$\mathsf{E}[|T_a^{\Delta_n} - \langle M,M\rangle_a|^2] \to 0,$$

$$\int_A T_t^{\Delta_n} d\mathsf{P} \to \int_A \langle M,M\rangle_t d\mathsf{P},$$
$$\int_A T_s^{\Delta_n} d\mathsf{P} \to \int_A \langle M,M\rangle_s d\mathsf{P},$$

因此

$$\int_A M_t^2 - \langle M,M\rangle_t d\mathsf{P} = \int_A M_s^2 - \langle M,M\rangle_s d\mathsf{P}.$$

即 $M^2 - \langle M,M\rangle$ 是鞅.

现在不再假设 M 的有界性, 考虑一般的情况.

令

$$T_n = \inf\{t : |M_t| \geqslant n\}.$$

利用上面的讨论, 存在连续的增过程 $Q^{(n)} = \langle M^{T_n}, M^{T_n}\rangle$, 使得 $(M^{T_n})^2 - Q^{(n)}$ 是鞅. 在集合 $\{(t,\omega) : t < T_n(\omega)\}$ 上令 $\langle M,M\rangle = Q^{(n)}$, 对于 $t>0$, $\lim\limits_{n\to\infty}\mathsf{P}(t < T_n) = 1$, 结合 $Q^{(n)}$ 的唯一性, 我们最终证明结论. ■

这里值得注意的是, 上述证明最后的部分被称为局部化技巧. 至此, 我们得到, 对于任一连续局部鞅 M, 存在一个唯一的连续增过程 $\langle M,M\rangle$, 使得 $M^2 - \langle M,M\rangle$ 是局部鞅.

定义 4.2.1 若 M 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的连续局部鞅, 存在一个唯一的连续增过程 $\langle M, M\rangle$, 使得 $M^2 - \langle M, M\rangle$ 是局部鞅, 我们称 $\langle M, M\rangle$ 为 M 的二次变差过程.

对于两个连续局部鞅 M, N,

$$MN = \frac{1}{4}((M+N)^2 - (M-N)^2),$$

由于

$$\frac{1}{4}((M+N)^2 - (M-N)^2) - \frac{1}{4}(\langle M+N, M+N\rangle - \langle M-N, M-N\rangle)$$

是局部鞅, 故

$$MN - \frac{1}{4}(\langle M+N, M+N\rangle - \langle M-N, M-N\rangle)$$

是局部鞅. 因此我们定义

$$\langle M, N\rangle = \frac{1}{4}(\langle M+N, M+N\rangle - \langle M-N, M-N\rangle).$$

此时, $MN - \langle M, N\rangle$ 是局部鞅.

这里不加证明地给出下列定理.

定理 4.2.2 若 M, N 是连续的局部鞅, 存在唯一的连续有限变差过程 $\langle M, N\rangle$, 使 $MN - \langle M, N\rangle$ 是局部鞅. 进一步, 对于 $[0,t]$, 若划分 $D:\ 0 = t_0 < t_1 < \cdots < t_n = t$, 且 $\max\limits_i |t_i - t_{i-1}| \to 0$ (当 $n \to +\infty$), 则

$$\sum_{i=1}^{n}(M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i}) \xrightarrow{\mathsf{P}} \langle M, N\rangle_t.$$

进一步, 若考虑停止过程, 有如下结果.

命题 4.2.1 若 T 是停时, 则对连续局部鞅 M, N, 有

$$\langle M^T, N^T\rangle = \langle M, N^T\rangle = \langle M, N\rangle^T = \langle M^T, N\rangle,$$

证明: 由于 $MN - \langle M, N\rangle$ 是局部鞅, 故 $M^T N^T - \langle M, N\rangle^T$ 是局部鞅. 故

$$\langle M^T, N^T\rangle = \langle M, N\rangle^T.$$

其次, 由于 $N^T(M - M^T) = \mathbb{1}_{\{T\leqslant t\}} N_T(M_t - M_T)$ 是局部鞅. 故 $N^T M - \langle M, N\rangle^T = N^T M - N^T M^T + N^T M^T - \langle M, N\rangle^T$ 是局部鞅. 由二次变差过程的唯一性

$$\langle M, N^T\rangle = \langle M, N\rangle^T.$$

其余类似可得. ■

接下来的几个命题给出了二次变差过程的一些性质.

命题 4.2.2 若 $M=(M_t)_{t\geqslant 0}$ 是连续的局部鞅, $M_0=0$, 若 $\langle M,M\rangle=0$, 则 $M\equiv 0,\ \ a.s.$

证明: 若 $\langle M,M\rangle=0$, 则 $\mathsf{E}[M_t^2]=0$, 即 $M_t\equiv 0,\ \ a.s.$. ■

命题 4.2.3 若 M,N,Y 是连续的局部鞅, 则

$$\langle M+N,Y\rangle=\langle M,Y\rangle+\langle N,Y\rangle.$$

证明: 考虑 $(M+N)Y=MY+NY$, $MY-\langle M,Y\rangle$ 是局部鞅, $NY-\langle N,Y\rangle$ 是局部鞅, 则 $(M+N)Y-\langle M,Y\rangle-\langle N,Y\rangle$ 是局部鞅, 则

$$\langle M+N,Y\rangle=\langle M,Y\rangle+\langle N,Y\rangle.$$

■

在随机分析中, 半鞅经常被研究. 下面我们给出半鞅的定义.

定义 4.2.2 若 S 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的适应过程. $S=M+A$, M 是局部鞅, A 是适应的有限变差过程, 则称 S 称半鞅.

事实上,可以证明,若 S 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续半鞅,则有 $S=M+A$, 其中 M 是连续的局部鞅, A 是连续的有限变差过程.

关于有限变差过程, 有

命题 4.2.4 若 A 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的适应的连续的有限变差过程, 则 $\langle A,A\rangle_t=0,\ \ a.s.$.

证明: 对于 $[0,t]$ 考虑划分, $0=t_0<t_1<\cdots<t_n=t$. 利用局部化技巧，不妨假设 A 是有界的, 注意到

$$\sum_{i=1}^{n}(A_{t_i}-A_{t_{i-1}})^2\leqslant \sup_i|A_{t_i}-A_{t_{i-1}}|\sum_{i=1}^{n}|A_{t_i}-A_{t_{i-1}}|,$$

由于 A 是有限变差的, 且连续的, 故上式极限为 0, 即 $\langle A,A\rangle_t=0,\ \ a.s.$. ■

类似可有, 若 A, B 是连续的适应的有限变差过程, 则 $\langle A,B\rangle=0$, 且易得以下结论.

命题 4.2.5 若 S 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续半鞅, 若 $S=M+A$, M 是连续局部鞅, A 是连续的有限变差过程, 则 $\langle S,S\rangle=\langle M,M\rangle$.

在连续时间框架下，我们经常讨论局部鞅. 一个过程如果是局部鞅，很可能不是鞅，下面，我们给出一个命题，给出了局部鞅和鞅的关系.

命题 4.2.6 若 M 是连续的局部鞅, $M_0=0$, 则以下论断等价:

(1). 对任意的 $t>0$, $\mathsf{E}[\langle M,M\rangle_t]<+\infty$;

(2). M 是鞅, 且对任意的 $t>0$, $\mathsf{E}[M_t^2]<+\infty$.

特别地, 若上述条件成立, $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是鞅. 若 M 和 N 是连续的局部鞅, 对任意的 $t>0$, $\mathsf{E}[\langle M,M\rangle_t]<+\infty$, $\mathsf{E}[\langle N,N\rangle_t]<+\infty$, 则 $(M_tN_t-\langle M,N\rangle_t)_{t\geqslant 0}$ 是鞅.

证明: 我们首先证明在较强的条件下的情形: 若 M 是连续的局部鞅, $M_0=0$, 我们证明以下两个论断等价:

(3). $\mathsf{E}[\langle M,M\rangle_\infty]<+\infty$;

(4). M 是鞅, 且 $\sup_{t\geqslant 0}\mathsf{E}[M_t^2]<+\infty$.

首先证明 (4)⇒(3).

对于 $K>0$, 由 Doob 不等式

$$\mathsf{E}[\sup_{0\leqslant t\leqslant K} M_t^2]\leqslant 4\mathsf{E}[M_K^2].$$

令 $K\to+\infty$, 则有

$$\mathsf{E}[\sup_{t\geqslant 0} M_t^2]\leqslant 4\sup_{t\geqslant 0}\mathsf{E}[M_t^2]<+\infty.$$

令

$$S_n=\inf\{t>0:\langle M,M\rangle_t\geqslant n\},$$

则连续局部鞅 $M_{t\wedge S_n}^2-\langle M,M\rangle_{t\wedge S_n}$ 被 $\sup\limits_{t\geqslant 0} M_t^2+n$ 控制. 由 $\sup\limits_{t\geqslant 0} M_t^2$ 的可积性, $M_{t\wedge S_n}^2-\langle M,M\rangle_{t\wedge S_n}$ 是一致可积鞅. 且

$$\mathsf{E}[\langle M,M\rangle_{t\wedge S_n}]=\mathsf{E}[M_{t\wedge S_n}^2]\leqslant \mathsf{E}[\sup_{t\geqslant 0} M_t^2].$$

令 $n\to+\infty$, $t\to+\infty$, 有

$$\mathsf{E}[\langle M,M\rangle_\infty]\leqslant \mathsf{E}[\sup_{t\geqslant 0} M_t^2]<+\infty.$$

下证 (3)⇒(4).

假设 $\mathsf{E}[\langle M,M\rangle_\infty]<+\infty$, 令

$$T_n=\inf\{t\geqslant 0;|M_t|\geqslant n\}.$$

此时, $M_{t\wedge T_n}^2-\langle M,M\rangle_{t\wedge T_n}$ 被 $n^2+\langle M,M\rangle_\infty$ 控制. 故 $(M_{t\wedge T_n}^2-\langle M,M\rangle_{t\wedge T_n})$ 是一致可积鞅. 于是

$$\mathsf{E}[M_{t\wedge T_n}^2]=\mathsf{E}[\langle M,M\rangle_{t\wedge T_n}]\leqslant \mathsf{E}[\langle M,M\rangle_\infty]<+\infty.$$

令 $n\to+\infty$,

$$\mathsf{E}[M_t^2]\leqslant \mathsf{E}[\langle M,M\rangle_\infty]<+\infty.$$

故 $\sup\limits_{t\geqslant 0}\mathsf{E}[M_t^2]<+\infty$.

由于

$$\mathsf{E}[M_{t\wedge T_n}^2]\leqslant \mathsf{E}[\langle M,M\rangle_\infty]<+\infty,$$

故 $(M_{t\wedge T_n})_{n\geqslant 0}$ 是一致可积的. 故 $(M_{t\wedge T_n})_{n\geqslant 1}$ 是 $\mathbf{L}^1$ 收敛至 M_t. 当 $t\geqslant s$ 时,

$$\mathsf{E}[M_{t\wedge T_n}\mid \mathcal{F}_s]=M_{s\wedge T_n}.$$

$\mathbf{L}^1$ 收敛意味着, 对于 $A\in\mathcal{F}_s$,

$$\int_A M_{t\wedge T_n}\,\mathrm{d}\mathsf{P}=\int_A M_{s\wedge T_n}\,\mathrm{d}\mathsf{P}.$$

令 $n\to+\infty$,

$$\int_A M_t\,\mathrm{d}\mathsf{P}=\int_A M_s\,\mathrm{d}\mathsf{P}.$$

故 $\mathsf{E}[M_t\mid\mathcal{F}_s]=M_s$. M 是鞅.

上面证明了条件 (3) 与条件 (4) 是等价的. 对于给定的 $a\geqslant 0$, 我们考虑 $(M_{t\wedge a})_{t\geqslant 0}$. 对随机过程 $(M_{t\wedge a})_{t\geqslant 0}$ 条件 (3) 与条件 (4) 是等价的, 我们便得到对随机过程 M. 条件 (1) 与条件 (2) 是等价的.

接下来, 我们证明在条件 (3) 下, 即 $\mathsf{E}[\langle M,M\rangle_\infty]<+\infty$ 时, $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是一致可积鞅.

注意到

$$|M_t^2-\langle M,M\rangle_t|\leqslant \sup_{t\geqslant 0}M_t^2+\langle M,M\rangle_\infty,$$

有条件 (3) 及 $\mathsf{E}[\sup_{t\geqslant 0}M_t^2]\leqslant 4\sup_{t\geqslant 0}\mathsf{E}[M_t^2]<+\infty$. 我们可以知道 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是一致可积的.

由于 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是连续局部鞅, 存在一列停时 $T_n\uparrow\infty$, 使得对于 $t>s$,

$$\mathsf{E}[M_{T_n\wedge t}^2-\langle M,M\rangle_{T_n\wedge t}|\mathcal{F}_s]=M_{T_n\wedge s}^2-\langle M,M\rangle_{T_n\wedge s}.$$

由于 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是一致可积的, $M_{T_n\wedge t}^2-\langle M,M\rangle_{T_n\wedge t}$ 是 $\mathbf{L}^1$ 收敛至 $M_t^2-\langle M,M\rangle_t$, 类似于上面的讨论，我们可知 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是一致可积鞅.

在条件 (1) 下, 即对任意的 $t>0$, $\mathsf{E}[\langle M,M\rangle_t]<+\infty$ 时, 类似于条件 (3) 与条件 (4) 到条件 (1) 与条件 (2) 的转换, 我们得到 $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是鞅.

对于 MN 的情形, 由于 $MN=\frac{1}{4}[(M+N)^2-(M-N)^2]$, 注意到 $\mathsf{E}[\langle M,M\rangle_t]<+\infty$, $\mathsf{E}[\langle N,N\rangle_t]<+\infty$, 可以推出 $\mathsf{E}[\langle M+N,M+N\rangle_t]<+\infty$, $\mathsf{E}[\langle M-N,M-N\rangle_t]<+\infty$, 因此, $\mathsf{E}[\langle N,N\rangle_t]<+\infty$, 则 $(M_tN_t-\langle M,N\rangle_t)_{t\geqslant 0}$ 是鞅.

■

若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动. 因为 $(B_t^2-t)_{t\geqslant 0}$ 是鞅，故 $\langle B,B\rangle_t=t$.

下面考虑稍微复杂一点的问题.

定义 4.2.3 设 $0=t_0<t_1<\cdots<t_k\to+\infty$, f 是关于 $\mathcal{F}_0$ 可测的有界随机变量, f_i 是关于 $\mathcal{F}_{t_i}$ 可测的有界随机变量, 令

$$F_t=f\mathbb{1}_{\{0\}}(t)+\sum_{i=0}^{\infty}f_i\mathbb{1}_{(t_i,t_{i+1}]}(t).$$

则 $F=(F_t)_{t\geqslant 0}$ 称为简单过程. 简单过程全体记为 $\mathscr{L}_0$.

显然, $F=(F_t)_{t\geqslant 0}$ 是一个适应的左连续过程. 对于布朗运动 $B=(B_t)_{t\geqslant 0}$, 定义过程 $(\mathcal{J}(F)_t)$ 如下:

$$\mathcal{J}(F)_t=\sum_{i=0}^{\infty}f_i(B_{t\wedge t_{i+1}}-B_{t\wedge t_i}).$$

事实上, $\mathcal{J}(F)_t$ 的表达式中只有有限多项是非零的. 并且有下面的命题.

命题 4.2.7 $(\mathcal{J}(F)_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅.

证明: 下面证明对于 $t>s$, 有 $\mathsf{E}[\mathcal{J}(F)_t\mid\mathcal{F}_s]=\mathcal{J}(F)_s$. 设 $t_j<t\leqslant t_{j+1},t_k<s\leqslant t_{k+1},k\leqslant j$, 注意到:

$$\mathcal{J}(F)_t=\sum_{i=0}^{j-1}f_i(B_{t_{i+1}}-B_{t_i})+f_j(B_t-B_{t_j}),$$

$$\mathcal{J}(F)_s=\sum_{i=0}^{k-1}f_i(B_{t_{i+1}}-B_{t_i})+f_k(B_s-B_{t_k}).$$

当 $k<j-1$ 时,

$$\mathcal{J}(F)_t-\mathcal{J}(F)_s=\sum_{i=k+1}^{j-1}f_i(B_{t_{i+1}}-B_{t_i})+f_j(B_t-B_{t_j})+f_k(B_{t_{k+1}}-B_s).$$

若 $k+1 \leqslant i \leqslant j-1$, 那么 $s < t_i$, 于是 $\mathcal{F}_s \subset \mathcal{F}_t$, 则

$$\begin{aligned}\mathsf{E}[f_i(B_{t_{i+1}} - B_{t_i}) \mid \mathcal{F}_s] &= \mathsf{E}[\mathsf{E}[f_i(B_{t_{i+1}} - B_{t_i}) \mid \mathcal{F}_{t_i}] \mid \mathcal{F}_s] \\ &= \mathsf{E}[f_i \mathsf{E}[(B_{t_{i+1}} - B_{t_i}) \mid \mathcal{F}_{t_i}] \mid \mathcal{F}_s] = 0.\end{aligned}$$

因此

$$\mathsf{E}[\sum_{i=k+1}^{j-1} f_i(B_{t_{i+1}} - B_{t_i}) \mid \mathcal{F}_s] = 0.$$

对于 $f_j(B_t - B_{t_j})$,

$$\begin{aligned}\mathsf{E}[f_j(B_t - B_{t_j}) \mid \mathcal{F}_s] &= \mathsf{E}[\mathsf{E}[f_j(B_t - B_{t_j}) \mid \mathcal{F}_{t_j}] \mid \mathcal{F}_s] \\ &= \mathsf{E}[f_j \mathsf{E}[(B_t - B_{t_j}) \mid \mathcal{F}_{t_j}] \mid \mathcal{F}_s] = 0,\end{aligned}$$

$$\mathsf{E}[f_k(B_{t_{k+1}} - B_s) \mid \mathcal{F}_s] = f_k \mathsf{E}[(B_{t_{k+1}} - B_s) \mid \mathcal{F}_s] = 0,$$

故当 $k < j-1$ 时,

$$\mathsf{E}[\mathcal{J}(F)_t - \mathcal{J}(F)_s \mid \mathcal{F}_s] = 0.$$

当 $k = j-1$ 时, $t_{j-1} < s \leqslant t_j < t \leqslant t_{j+1}$, 此时

$$\mathcal{J}(F)_t - \mathcal{J}(F)_s = f_{j-1}(B_{t_j} - B_{t_s}) + f_j(B_t - B_{t_j}).$$

$$\begin{aligned}\mathsf{E}[\mathcal{J}(F)_t - \mathcal{J}(F)_s \mid \mathcal{F}_s] &= \mathsf{E}[f_{j-1}(B_{t_j} - B_{t_s}) \mid \mathcal{F}_s] + \mathsf{E}[f_j(B_t - B_{t_j}) \mid \mathcal{F}_s] \\ &= f_{j-1} \mathsf{E}[(B_{t_j} - B_{t_s}) \mid \mathcal{F}_s] \\ &\quad + \mathsf{E}[\mathsf{E}[f_j(B_t - B_{t_j}) \mid \mathcal{F}_{t_j}] \mid \mathcal{F}_s] \\ &= 0 + \mathsf{E}[f_j \mathsf{E}[(B_t - B_{t_j}) \mid \mathcal{F}_{t_j}] \mid \mathcal{F}_s] \\ &= 0.\end{aligned}$$

当 $t_j < s < t < t_{j+1}$ 时, 易证.

故 $(\mathcal{J}(F)_t)_{t \geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0}, \mathsf{P})$ 上的鞅. ■

进一步，有

命题 4.2.8 $(\mathcal{J}(F)_t^2 - \int_0^t F_s^2 \, \mathrm{d}s)_{t \geqslant 0}$ 是鞅.

证明 [10]: 下面证明 $t \geqslant s$ 时,

$$\mathsf{E}[\mathcal{J}(F)_t^2 - \int_0^t F_u^2 \, \mathrm{d}u \mid \mathcal{F}_s] = \mathcal{J}(F)_s^2 - \int_0^s F_u^2 \, \mathrm{d}u.$$

考虑

$$\begin{aligned}\mathcal{J}(F)_t^2-\mathcal{J}(F)_s^2&=(\mathcal{J}(F)_t-\mathcal{J}(F)_s+\mathcal{J}(F)_s)^2-\mathcal{J}(F)_s^2\\&=(\mathcal{J}(F)_t-\mathcal{J}(F)_s)^2+2(\mathcal{J}(F)_t-\mathcal{J}(F)_s)\mathcal{J}(F)_s,\end{aligned}$$

$$\mathsf{E}[2(\mathcal{J}(F)_t-\mathcal{J}(F)_s)\mathcal{J}(F)_s)\mid\mathcal{F}_s]=2\mathcal{J}(F)_s\mathsf{E}[\mathcal{J}(F)_t-\mathcal{J}(F)_s\mid\mathcal{F}_s]=0.$$

进一步,

$$(\mathcal{J}(F)_t-\mathcal{J}(F)_s)^2=(\sum_{i=0}^{\infty}f_i(B_{t_{i+1}\wedge t}-B_{t_i\wedge t})-\sum_{i=0}^{\infty}f_i(B_{t_{i+1}\wedge s}-B_{t_i\wedge s}))^2.$$

设 $t_j<t\leqslant t_{j+1}$，$t_k<s\leqslant t_{k+1}$，若 $k<j-1$, 此时

$$\begin{aligned}(\mathcal{J}(F)_t-\mathcal{J}(F)_s)^2&=(\sum_{i=k+1}^{j-1}f_i(B_{t_{i+1}}-B_{t_i})+f_j(B_t-B_{t_j})+f_k(B_{t_{k+1}}-B_s))^2\\&=f_j^2(B_t-B_{t_j})^2+f_k^2(B_{t_{k+1}}-B_s)^2\\&\quad+2f_kf_j(B_t-B_{t_j})(B_{t_{k+1}}-B_s)+(\sum_{i=k+1}^{j-1}f_i(B_{t_{i+1}}-B_{t_i}))^2\\&\quad+2\sum_{i=k+1}^{j-1}f_if_j(B_{t_{i+1}}-B_{t_i})(B_t-B_{t_j})\\&\quad+2\sum_{i=k+1}^{s-1}f_if_k(B_{t_{k+1}}-B_s)(B_{t_{i+1}}-B_{t_i}).\end{aligned}$$

这时

$$\begin{aligned}\mathsf{E}[f_j^2(B_t-B_{t_j})^2\mid\mathcal{F}_s]&=\mathsf{E}[\mathsf{E}[f_j^2(B_t-B_{t_j})^2\mid\mathcal{F}_{t_j}]\mid\mathcal{F}_s]\\&=\mathsf{E}[f_j^2(t-t_j)\mid\mathcal{F}_s].\end{aligned}$$

$$\mathsf{E}[f_k^2(B_{t_{k+1}}-B_s)^2\mid\mathcal{F}_s]=\mathsf{E}[f_k^2(t_{k+1}-s)\mid\mathcal{F}_s].$$

当 $j-1\geqslant i\geqslant k+1$ 时,

$$\begin{aligned}&\mathsf{E}[f_if_j(B_{t_{i+1}}-B_{t_i})(B_t-B_{t_j})\mid\mathcal{F}_s]\\&=\mathsf{E}[\mathsf{E}[f_if_j(B_{t_{i+1}}-B_{t_i})(B_t-B_{t_j})\mid\mathcal{F}_{t_j}]\mid\mathcal{F}_s]=0.\end{aligned}$$

同理

$$\mathsf{E}(f_if_k(B_{t_{k+1}}-B_s)(B_{t_{i+1}}-B_{t_i})\mid\mathcal{F}_s)=0.$$

接下来可以考虑

$$(\sum_{i=k+1}^{j-1} f_i(B_{t_{i+1}} - B_{t_i}))^2 = \sum_{i=k+1}^{j-1} f_i^2(B_{t_{i+1}} - B_{t_i})^2 + 2\sum_{i=k+1}^{j-1}\sum_{l=k+1}^{j-1} f_{t_i} f_{t_l}(B_{t_{i+1}} - B_t)(B_{t_{l+1}} - B_{t_l}).$$

类似于上面的讨论,

$$\mathsf{E}[(\sum_{i=k+1}^{j-1} f_i(B_{t_{i+1}} - B_{t_i}))^2 \mid \mathcal{F}_s] = \mathsf{E}[\sum_{i=k+1}^{j-1} f_i^2(t_{i+1} - t_i) \mid \mathcal{F}_s],$$

$$\mathsf{E}[\sum_{i=k+1}^{j-1}\sum_{l=k+1}^{j-1} f_{t_i} f_{t_l}(B_{t_{i+1}} - B_t)(B_{t_{l+1}} - B_{t_l}) \mid \mathcal{F}_s] = 0.$$

综上

$$\begin{aligned}&\mathsf{E}[(\mathcal{J}(F)_t - \mathcal{J}(F)_s)^2 \mid \mathcal{F}_s]\\ =\ &\mathsf{E}(\sum_{i=k+1}^{j-1} f_i^2(t_{i+1} - t_i) + f_j^2(t - t_j) + f_k^2(t_{k+1} - s) \mid \mathcal{F}_s),\end{aligned}$$

于是

$$\mathsf{E}[\mathcal{J}(F)_t^2 - \int_0^t F_u^2\,\mathrm{d}u \mid \mathcal{F}_s] = 0.$$

即有 $(\mathcal{J}(F)_t^2 - \int_0^t F_s^2\,\mathrm{d}s)_{t\geqslant 0}$ 是鞅. ■

由前面的讨论我们知道

$$\langle \mathcal{J}(F), \mathcal{J}(F)\rangle_t = \int_0^t F_s^2\,\mathrm{d}s.$$

4.3 连续局部鞅的 Itô 积分

在上一节的介绍中，我们讨论了形如

$$\mathcal{J}(F)_t = \sum_{i=0}^{\infty} f_i(B_{t\wedge t_{i+1}} - B_{t\wedge t_i})$$

的性质. 由于 $\mathcal{J}(F)$ 很像之前学过的黎曼和. 事实上, $\mathcal{J}(F)$ 就是接下来要讲的 Itô 积分. (但要注意的是, Itô 积分绝不是黎曼积分.) 接下来, 我们考虑一个问题:

如果 $F=(F_s)_{s\geqslant 0}$ 是循序可测过程, 满足 $\int_0^{+\infty}F_s^2\,\mathrm{d}s<+\infty$, 我们能否找到一个连续鞅 M, 使得 $\langle M,M\rangle_t=\int_0^t F_s^2\,\mathrm{d}s$?

为了回答这一问题，我们引入连续局部鞅的 Itô 积分. 首先需要一些记号. 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上, 记 $\mathcal{H}^2$ 为满足 $\sup\limits_t\ \mathsf{E}[M_t^2]<+\infty$ 的连续鞅全体, 记 $\mathcal{H}_0^2$ 为 $\mathcal{H}^2$ 中满足 $M_0=0$ 的连续鞅全体. 在 $\mathcal{H}^2$ 中引入范数:

$$||M||_{\mathcal{H}^2}=(\mathsf{E}[M_\infty^2])^{1/2}=\lim_{t\to+\infty}(\mathsf{E}[M_t^2])^{1/2}.$$

事实上, 由于 $\mathcal{H}^2$ 中的连续鞅满足 $\sup\limits_t\ \mathsf{E}[M_t^2]<+\infty$, 显然 $\mathcal{H}^2$ 中的连续鞅是一致可积的, 因此有 $M_t=\mathsf{E}[M_\infty\mid\mathcal{F}_t]$. 容易验证 $\mathcal{H}^2$ 在引入范数 $||M||_{\mathcal{H}^2}$ 后, 定义内积 $(M,N)=\mathsf{E}[M_\infty\ N_\infty]$ 可作成一个 Hilbert 空间.

设 $M\in\mathcal{H}^2$, $\mathbf{L}^2\ (M)$ 记满足

$$||K||_M^2=\mathsf{E}[\int_0^\infty K_s^2\,\mathrm{d}\langle M,M\rangle_s]<+\infty$$

的循序可测的 K 的全体.

介绍 Itô 积分之前，我们首先给出两个关键的不等式.

命题 4.3.1 *若 M, N 是连续的局部鞅, K, H 是循序可测过程, 则*

$$\int_0^t|H_s||K_s|\,|\mathrm{d}\langle M,N\rangle|_s\leqslant(\int_0^t H_s^2\,\mathrm{d}\langle M,M\rangle_s)^{1/2}(\int_0^t K_s^2\,\mathrm{d}\langle N,N\rangle_s)^{1/2}.$$

证明 [12]: 由单调类方法，只需考虑

$$H=H_0\mathbb{1}_{\{0\}}+H_1\mathbb{1}_{(0,t_1]}+\cdots+H_n\mathbb{1}_{(t_{n-1},t_n]}$$

及

$$K=K_0\mathbb{1}_{\{0\}}+K_1\mathbb{1}_{(0,t_1]}+\cdots+K_n\mathbb{1}_{(t_{n-1},t_n]}$$

的情形. 这里 $K_0,K_1,\cdots K_n$, $H_0,H_1,\cdots,H_n$ 是有界随机变量. 记

$$\langle M,N\rangle_s^t=\langle M,N\rangle_t-\langle M,N\rangle_s,$$

故

$$|\int_0^t H_sK_s\,\mathrm{d}\langle M,N\rangle_s|=\sum_i|H_iK_i||\langle M,N\rangle_{t_i}^{t_{i+1}}|.$$

考虑对于 $s<t$,

$$\langle M,M\rangle_s^t+2r\langle M,N\rangle_s^t+r^2\langle N,N\rangle_s^t=\langle M+rN,M+rN\rangle_s^t\geqslant 0,$$

故

$$|\langle M,N\rangle_s^t| \leqslant (\langle M,M\rangle_s^t)^{1/2}(\langle N,N\rangle_s^t)^{1/2}.$$

于是

$$|\langle M,N\rangle_{t_i}^{t_{i+1}}| \leqslant (\langle M,M\rangle_{t_i}^{t_{i+1}})^{1/2}(\langle N,N\rangle_{t_i}^{t_{i+1}})^{1/2}.$$

可以得到

$$|\int_0^t H_s K_s \,\mathrm{d}\langle M,N\rangle_s| \leqslant \sum_i |H_i||K_i|(\langle M,M\rangle_{t_i}^{t_{i+1}})^{1/2}(\langle N,N\rangle_{t_i}^{t_{i+1}})^{1/2}.$$

进一步

$$\begin{aligned}
&\sum_i |H_i||K_i|(\langle M,M\rangle_{t_i}^{t_{i+1}})^{1/2}(\langle N,N\rangle_{t_i}^{t_{i+1}})^{1/2}\\
\leqslant&\ (\sum_i H_i^2\langle M,M\rangle_{t_i}^{t_{i+1}})^{1/2}(\sum_i K_i^2\langle N,N\rangle_{t_i}^{t_{i+1}})^{1/2}\\
=&\ (\int_0^t H_s^2 \,\mathrm{d}\langle M,M\rangle_s)^{1/2}(\int_0^t K_s^2 \,\mathrm{d}\langle N,N\rangle_s)^{1/2}.
\end{aligned}$$

而由 $|\langle M,N\rangle_s^t| \leqslant (\langle M,M\rangle_s^t)^{1/2}(\langle N,N\rangle_s^t)^{1/2}$, 可得

$$|\int_0^t 1 \,\mathrm{d}\langle M,N\rangle_u| \leqslant (\langle M,M\rangle_0^t)^{1/2}(\langle N,N\rangle_0^t)^{1/2}.$$

于是

$$\int_0^t |H_s||K_s|\,|\mathrm{d}\langle M,N\rangle|_s \leqslant (\int_0^t H_s^2 \,\mathrm{d}\langle M,M\rangle_s)^{1/2}(\int_0^t K_s^2 \,\mathrm{d}\langle N,N\rangle_s)^{1/2}.$$

■

利用命题 4.3.1 及 Hölder 不等式即可得以下的 Kunita-Watanabe 不等式.

命题 4.3.2 对于 $p \geqslant 1$, $\frac{1}{p}+\frac{1}{q}=1$, H,K,M,N 满足命题 4.3.1 的条件, 则

$$\mathsf{E}[\int_0^\infty |H_s||K_s|\,|\mathrm{d}\langle M,N\rangle|_s] \leqslant ||(\int_0^\infty H_s^2 \,\mathrm{d}\langle M,M\rangle_s)^{1/2}||_p||(\int_0^\infty K_s^2 \,\mathrm{d}\langle N,N\rangle_s)^{1/2}||_q.$$

下面，给出一个关键的定理，根据这个定理, 可以给出 Itô 积分的定义.

定理 4.3.1 令 $M \in \mathcal{H}^2$, 对于 $K \in \mathbf{L}^2\,(M)$, 存在唯一一个 $\mathcal{H}_0^2$ 中的元素, 记为 $K\cdot M$, 使对于任意 $N \in \mathcal{H}^2$,

$$\langle K\cdot M, N\rangle = K\cdot\langle M,N\rangle,$$

且 $K \to K\cdot M$ 是 $\mathbf{L}^2\,(M)$ 到 $\mathcal{H}_0$ 的一个等距映射.

证明 [12]: 首先证明唯一性. 记 L, L' 均为 $\mathcal{H}_0^2$ 中满足条件的鞅, 使 $\langle L, N\rangle = \langle L', N\rangle$, 即 $\langle L-L', N\rangle = 0$. 令 $N = L - L'$, 故 $\langle L-L', L-L'\rangle = 0$, 即 $L = L'$.

下证存在性. 对于任意 $N \in \mathcal{H}_0^2$, 令

$$f(N) = \mathsf{E}[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s].$$

若 $N_1 \in \mathcal{H}_0^2$, $N_2 \in \mathcal{H}_0^2$,

$$\begin{aligned} f(N_1+N_2) &= \mathsf{E}[\int_0^\infty K_s \,\mathrm{d}\langle M, N_1\rangle_s] + \mathsf{E}[\int_0^\infty K_s \,\mathrm{d}\langle M, N_2\rangle_s] \\ &= f(N_1) + f(N_2). \end{aligned}$$

易得 f 是 $\mathcal{H}_0^2$ 上的线性泛函. 进一步, 若 $N \in \mathcal{H}_0^2$, 则

$$||N||_{\mathcal{H}^2} = (\mathsf{E}[N_\infty^2])^{1/2} = (\mathsf{E}\langle N, N\rangle_\infty)^{1/2},$$

$$||K||_{\mathcal{M}} = (\mathsf{E}[\int_0^\infty K_s^2 \,\mathrm{d}\langle M, M\rangle_s])^{1/2}.$$

由 Kunita-Watanabe 不等式

$$\begin{aligned} &\mathsf{E}[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s] \\ \leqslant &||(\int_0^\infty K_s^2 \,\mathrm{d}\langle M, M\rangle_s)^{1/2}||_2 \cdot ||(\int_0^\infty 1 \,\mathrm{d}\langle N, N\rangle_s)^{1/2}||_2 \\ = &(\mathsf{E}\int_0^\infty K_s^2 \,\mathrm{d}\langle M, M\rangle_s)^{1/2} \cdot (\mathsf{E}\langle N, N\rangle_\infty)^{1/2} \\ = &||N||_{\mathcal{H}^2} \cdot ||K||_{\mathcal{M}}. \end{aligned}$$

故 $f(N)$ 是 $\mathcal{H}_0^2$ 上的连续线性泛函. 由 Riesz 表示定理, 存在唯一元素 $K \cdot M \in \mathcal{H}_0^2$, 使

$$\mathsf{E}[\int_0^\infty K_s \,\mathrm{d}\langle M, N\rangle_s] = \mathsf{E}[K \cdot M_\infty N_\infty].$$

由于 $K \cdot M$ 是一致可积鞅. 对于任意停时 T, $\mathsf{E}[(K\cdot M)_\infty \mid \mathcal{F}_T] = (K \cdot M)_T$. 于是

$$\begin{aligned} \mathsf{E}[(K\cdot M)_T N_T] &= \mathsf{E}[\mathsf{E}[(K\cdot M)_\infty \mid \mathcal{F}_T] N_T] \\ &= \mathsf{E}[\mathsf{E}[(K\cdot M)_\infty N_T \mid \mathcal{F}_T]] \\ &= \mathsf{E}[(K\cdot M)_\infty N_T] = \mathsf{E}[(K\cdot M)_\infty N_\infty^T] \\ &= \mathsf{E}[(K\cdot \langle M, N^T\rangle)_\infty] = \mathsf{E}[(K\cdot \langle M, N\rangle^T)_\infty] \\ &= \mathsf{E}[K\cdot \langle M, N\rangle_T]. \end{aligned}$$

从而 $\mathsf{E}[(K\cdot M)_T N_T]=\mathsf{E}[K\cdot\langle M,N\rangle_T]$, 即 $(K\cdot M)N-K\cdot\langle M,N\rangle$ 是鞅, 亦即 $\langle K\cdot M,N\rangle=K\cdot\langle M,N\rangle$. 进一步,

$$||K\cdot M||^2_{\mathcal{H}^2}=\mathsf{E}[(K\cdot M)^2_\infty]=\mathsf{E}[(K^2\cdot\langle M,M\rangle_\infty)]=||K||^2_{\mathcal{M}},$$

故 $K\to K\cdot M$ 是 $\mathbf{L}^2(M)$ 到 $\mathcal{H}^2_0$ 的一个等距映射. ■

定义 4.3.1 定理 4.3.1 中得到的 $K\cdot M$ 被称为 K 关于 M 的 Itô 积分.

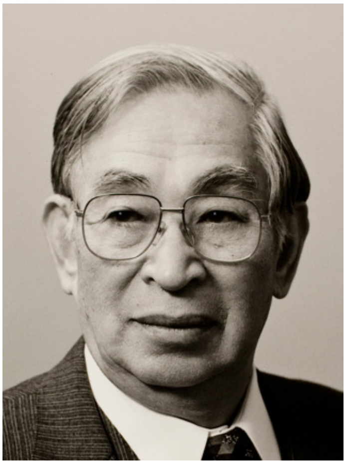

图 4.1 Itô Kiyoshi

伊藤清 (Itô Kiyoshi 1915 年 9 月 7 日—2008 年 11 月 10 日)，日本数学家，日本学士院院士，生于日本三重县. 西方文献中他的姓氏常写为 Itô. 伊藤清因在概率论方面的奠基性工作而获得了 1987 年的沃尔夫奖，并于 1998 年获得京都奖，2006 年获得首届高斯奖. 伊藤清的工作集中于概率论，他在 1944 年和 1946 年的两份著作立下随机积分和随机微分方程的理论基础，被誉为“现代随机分析之父”，因他命名的理论有 Itô 引理、Itô 积分、Itô 过程等. 他的理论被应用于很多不同领域，包括自然科学和经济学，例如金融数学中期权定价用的 Black-Scholes 模型. 诺贝尔经济学奖获奖者 Myron Scholes 遇到伊藤时，向他握手，称赞他的理论. 从这个插曲可以明白伊藤的成就不仅对数学，对社会科学也带来很大影响. 他是少有的在世的时候看到自己的理论研究被应用到现实生活中的数学家之一.

$K\cdot M$ 具体是一个什么形式呢？接下来对简单过程考虑这个问题. 不妨设

$$K=K\mathbb{1}_{\{0\}}+\sum_{i=0}^{\infty}K_i\mathbb{1}_{(t_i,t_{i+1}]}.$$

这里 $0=t_0<t_1<t_2<\cdots$, $\lim\limits_{i\to\infty}t_i=\infty$, K_i 关于 $\mathcal{F}_{t_i}$ 可测. 考虑

$$(\mathcal{J}_M(K))_t=\sum_{i=0}^{n-1} K_i(M_{t_{i+1}}-M_{t_i})+K_n(M_t-M_{t_n}),$$

这里 $\mathcal{J}_M(K)\in\mathcal{H}_0^2$, $M\in\mathcal{H}^2$. 对任意 $N\in\mathcal{H}_0^2$, 我们来求 $\langle\mathcal{J}_M(K),N\rangle$. 考虑

$$(\mathcal{J}_M(K))_tN_t=(\sum_{i=0}^{n-1} K_i(M_{t_{i+1}}-M_{t_i}))N_t+K_n(M_t-M_{t_n})N_t.$$

此时, 若 $t_n<t<t_{n+1}$, 则

$$\begin{aligned}\int_0^t K_s\,\mathrm{d}\langle M,N\rangle_s=&\sum_{i=0}^{n-1}K_i(\langle M,N\rangle_{t_{i+1}}-\langle M,N\rangle_{t_i})\\&+K_n(\langle M,N\rangle_t-\langle M,N\rangle_{t_n}).\end{aligned}$$

当 $t_n<s<t<t_{n+1}$ 时,

$$\begin{aligned}&\mathsf{E}[(K\cdot M)_tN_t-\int_0^t K_u\,\mathrm{d}\langle M,N\rangle_u\mid\mathcal{F}_s]\\=&\sum_{i=0}^{n-1}K_i(M_{t_{i+1}}-M_{t_i})N_s+\mathsf{E}[K_n(M_t-M_{t_n})N_t\mid\mathcal{F}_s]\\&-\sum_{i=0}^{n-1}K_i(\langle M,N\rangle_{t_{i+1}}-\langle M,N\rangle_{t_i})\\&-\mathsf{E}[K_n(\langle M,N\rangle_t-\langle M,N\rangle_{t_n})\mid\mathcal{F}_s],\end{aligned}$$

这里

$$\begin{aligned}&\mathsf{E}[K_n(M_t-M_{t_n})N_t\mid\mathcal{F}_s]-\mathsf{E}[K_n(\langle M,N\rangle_t-\langle M,N\rangle_{t_n})\mid\mathcal{F}_s]\\=&K_n\mathsf{E}[M_tN_t-\langle M,N\rangle_t\mid\mathcal{F}_s]-K_nM_{t_n}N_s+K_n\langle M,N\rangle_{t_n}\\=&K_n(M_sN_s-\langle M,N\rangle_s)-K_nM_{t_n}N_s+K_n\langle M,N\rangle_{t_n}\\=&K_n(M_s-M_{t_n})N_s-K_n(\langle M,N\rangle_s-\langle M,N\rangle_{t_n}).\end{aligned}$$

故

$$\begin{aligned}&\mathsf{E}[(\mathcal{J}_M(K))_tN_t-\int_0^t K_n\,\mathrm{d}\langle M,N\rangle_n\mid\mathcal{F}_s]\\=&(\mathcal{J}_M(K))_sN_s-\int_0^s K_n\,\mathrm{d}\langle M,N\rangle_n.\end{aligned}$$

当 $t_{n-1} < s < t_n < t < t_{n+1}$ 时,

$$\begin{aligned}&\mathsf{E}[(\mathcal{J}_M(K))_t N_t - \int_0^t K_u \,\mathrm{d}\langle M, N\rangle_u \mid \mathcal{F}_s]\\ =&\,\mathsf{E}[\mathsf{E}[(\mathcal{J}_M(K))_t N_t - \int_0^t K_u \,\mathrm{d}\langle M, N\rangle_u \mid \mathcal{F}_{t_n}] \mid \mathcal{F}_s].\end{aligned}$$

事实上,

$$\begin{aligned}&\mathsf{E}[(\mathcal{J}_M(K))_t N_t - \int_0^t K_n \,\mathrm{d}\langle M, N\rangle_n \mid \mathcal{F}_{t_n}]\\ =&\sum_{i=0}^{n-1} K_i(M_{t_{i+1}} - M_{t_i})N_{t_n} + \mathsf{E}[K_n(M_t - M_{t_n})N_t \mid \mathcal{F}_{t_n}]\\ &-\sum_{i=0}^{n-1} K_i(\langle M, N\rangle_{t_{i+1}} - \langle M, N\rangle_{t_i}) - \mathsf{E}[K_n(\langle M, N\rangle_t - \langle M, N\rangle_{t_n}) \mid \mathcal{F}_{t_n}].\end{aligned}$$

由于

$$\begin{aligned}&\mathsf{E}[K_n(M_t - M_{t_n})N_t \mid \mathcal{F}_{t_n}] - \mathsf{E}[K_n(\langle M, N\rangle_t - \langle M, N\rangle_{t_n}) \mid \mathcal{F}_{t_n}]\\ =&K_n\mathsf{E}[(M_tN_t - \langle M, N\rangle_t) \mid \mathcal{F}_{t_n}] - K_nM_{t_n}N_{t_n} + K_n\langle M, N\rangle_{t_n}\\ =&K_n(M_{t_n}N_{t_n} - \langle M, N\rangle_{t_n}) - K_nM_{t_n}N_{t_n} + K_n\langle M, N\rangle_{t_n}\\ =&0.\end{aligned}$$

故

$$\begin{aligned}&\mathsf{E}[(\mathcal{J}_M(K))_t N_t - \int_0^t K_u \,\mathrm{d}\langle M, N\rangle_u \mid \mathcal{F}_{t_n}]\\ =&(\mathcal{J}_M(K))_{t_n} N_{t_n} - \int_0^{t_n} K_u \,\mathrm{d}\langle M, N\rangle_u.\end{aligned}$$

于是

$$\begin{aligned}&\mathsf{E}[(\mathcal{J}_M(K))_t N_t - \int_0^t K_u \,\mathrm{d}\langle M, N\rangle_u \mid \mathcal{F}_s]\\ =&\mathsf{E}[(\mathcal{J}_M(K))_{t_n} N_{t_n} - \int_0^{t_n} K_u \,\mathrm{d}\langle M, N\rangle_u \mid \mathcal{F}_s] \ \text{(类似于上面讨论, 可有)}\\ =&(K\cdot M)_s N_s - \int_0^s K_u \,\mathrm{d}\langle M, N\rangle_u.\end{aligned}$$

其他情况类似可得, 因此，得到 $\langle \mathcal{J}_M(K), N\rangle = K\cdot\langle M, N\rangle$. 故

$$\mathcal{J}_M(K) = K\cdot M.$$

进一步，可以得到一些关于 Itô 积分的性质.

命题 4.3.3 若 $K \in \mathbf{L}^2(M), H \in \mathbf{L}^2(K\cdot M)$, 则 $HK \in \mathbf{L}^2(M)$, 且 $(HK)\cdot M = H \cdot (K \cdot M)$.

证明: 由于

$$\langle (HK) \cdot M, N\rangle = HK \cdot \langle M, N\rangle = H \cdot (K \cdot \langle M, N\rangle),$$

同时

$$\langle H \cdot (K \cdot M), N\rangle = H \cdot \langle K \cdot M, N\rangle = HK \cdot \langle M, N\rangle.$$

故

$$(HK) \cdot M = H \cdot (K \cdot M).$$

■

命题 4.3.4 若 T 是停时, $K \in \mathbf{L}^2(M)$, $H \in \mathbf{L}^2(K \cdot M)$, 则

$$K \cdot M^T = K\mathbb{1}_{[0,T]} \cdot M = (K \cdot M)^T.$$

证明: 注意到 $M^T = \mathbb{1}_{[0,T]} \cdot M$, 事实上对于 $N \in \mathcal{H}^2$,

$$\langle M^T \cdot N\rangle = \langle M, N\rangle^T = \mathbb{1}_{[0,T]} \cdot \langle M, N\rangle = \langle \mathbb{1}_{[0,T]} \cdot M, N\rangle.$$

于是

$$K \cdot M^T = K \cdot (\mathbb{1}_{[0,T]} \cdot M) = K\mathbb{1}_{[0,T]} \cdot M,$$

$$(K \cdot M)^T = \mathbb{1}_{[0,T]} \cdot (K \cdot M) = \mathbb{1}_{[0,T]}K \cdot M.$$

从而

$$K \cdot M^T = K\mathbb{1}_{[0,T]} \cdot M = (K \cdot M)^T.$$

■

上面的命题还是限制在 $\mathcal{H}^2$ 上, 下面进行推广.

定义 4.3.2 若 M 是连续局部鞅, 对于循序可测过程 K, 若存在一个增长到 ∞ 的停时列 $(T_n)_{n\geqslant 0}$, 使 $\mathsf{E}[\int_0^{T_n} K_s^2 \,\mathrm{d}\langle M, M\rangle_s] < \infty$. 称 $K \in \mathbf{L}^2_{loc}M$.

定理 4.3.2 对于 $K \in \mathbf{L}^2_{loc}M$, M 是给定的连续的局部鞅, 则存在唯一的连续局部鞅, 记为 $K \cdot M$, 使得对任意连续局部鞅 N,

$$\langle K \cdot M, N\rangle = K \cdot \langle M, N\rangle.$$

证明: 由于 M 是连续的局部鞅, 存在停时列 $(T_n)_{n\geqslant 0}$, 使得 $M^{T_n}\in\mathcal{H}^2$, 且 $K^{T_n}\in\mathbf{L}^2(M^{T_n})$. 于是, 存在 $X^{(n)}=K^{T_n}\cdot M^{T_n}$, 对于 $X^{(n+1)}$ 在 $[0,T_n]$ 上与 $X^{(n)}$ 重合 (唯一性). 故在 $[0,T_n]$ 上定义 $K\cdot M$ 为 $X^{(n)}$ 即可. ■

利用上述性质, 当 a, b 是实数时, 可得

$$(aK+bH)\cdot M=aK\cdot M+bH\cdot M.$$

今后，我们也可以用之前学过的积分符号 $\int_0^t K_s\,\mathrm{d}M_s$ 表示 $K\cdot M_t$

若 $t>r>0$, 我们定义

$$\int_r^t K_s\,\mathrm{d}M_s=\int_0^t K_s\,\mathrm{d}M_s-\int_0^r K_s\,\mathrm{d}M_s.$$

控制收敛定理在实变函数与测度论中起了十分重要的作用, 类似地，也有在 Itô 积分意义下的相关结论.

命题 4.3.5 设 $M=(M_t)_{t\geqslant 0}$ 是连续局部鞅. $(H^n)_{n\geqslant 1}$, H 是局部有界的循序可测过程, K 是非负的循序可测过程, 且下列条件成立:

(1) 对于任意 $s\in[0,t]$,

$$H_s^n\longrightarrow H_s;$$

(2) 对于任意 $s\in[0,t]$, 对于任意 n, $|H_s^n|\leqslant K_s$;

(3) $\int_0^t(K_s)^2\,\mathrm{d}\langle M,M\rangle_s<+\infty$. 则

$$\int_0^t H_s^n\,\mathrm{d}M_s\xrightarrow{\mathsf{P}}\int_0^t H_s\,\mathrm{d}M_s.$$

证明: 令 $T_p=\inf\{r:\ \int_0^r(K_s)^2\,\mathrm{d}\langle M,M\rangle_s\geqslant p\}\ \wedge\ t$. 由于

$$\mathsf{E}[(\int_0^{T_p}H_s^n\,\mathrm{d}M_s-\int_0^{T_p}H_s\,\mathrm{d}M_s)^2]=\mathsf{E}[\int_0^{T_p}(H_s^n-H_s)^2\,\mathrm{d}\langle M,M\rangle_s],$$

利用假设 (1), (2), 及 $\int_0^{T_p}(K_s)^2\,\mathrm{d}\langle M,M\rangle_s\leqslant p$, 由控制收敛定理可知

$$\mathsf{E}[\int_0^{T_p}(H_s^n-H_s)^2\,\mathrm{d}\langle M,M\rangle_s]\to 0.$$

又由

$$\lim_{p\to\infty}\mathsf{P}(T_p=t)=1.$$

可知命题结论成立. ■

感兴趣的读者可以把上述命题的证明细节补齐. 利用上面的命题，有如下命题.

命题 4.3.6 设 $M=(M_t)_{t\geqslant 0}$ 是局部连续鞅, H 为适应的连续过程. 则对任意 $t>0$, 当

$$0=t_0^n<t_1^n<\cdots<t_{p_n}^n=t,\ \lim_{n\to\infty}\max_i|t_i^n-t_{i-1}^n|=0$$

时,

$$\sum_{i=0}^{p_n-1}H_{t_i^n}(M_{t_{i+1}^n}-M_{t_i^n})\xrightarrow{\mathsf{P}}\int_0^t H_s\,\mathrm{d}M_s.$$

证明: 对于 $n\geqslant 1$, 令

$$H^n=\begin{cases}H_{t_i^n}, & t_i^n<s\leqslant t_{i+1}^n,\ i\in\{0,1,\cdots,p_n-1\}\\ H_0, & s=0\\ 0, & s>t\end{cases}$$

令 $K_s=\max\limits_{0\leqslant r\leqslant s}|H_s|$, 由命题 4.3.5, $\int_0^t H_s^n\,\mathrm{d}M_s\xrightarrow{\mathsf{P}}\int_0^t H_s\,\mathrm{d}M_s$, 这里注意到 $\int_0^t H_s^n\,\mathrm{d}M_s=\sum\limits_{i=0}^{p_n-1}H_{t_i^n}(M_{t_{i+1}^n}-M_{t_i^n})$ 即可. ■

下面看几个例子.

例 4.3.1 设 $M=(M_t)_{t\geqslant 0}$ 是连续局部鞅, 则对于任意的 $t>0$,

$$\sum_{i=0}^{p_n-1}M_{t_i^n}(M_{t_{i+1}^n}-M_{t_i^n})\xrightarrow{\mathsf{P}}\int_0^t M_s\,\mathrm{d}M_s.$$

注意到

$$\sum_{i=0}^{p_n-1}M_{t_i^n}(M_{t_{i+1}^n}-M_{t_i^n})+\sum_{i=0}^{p_n-1}(M_{t_{i+1}^n}-M_{t_i^n})^2=\sum_{i=0}^{p_n-1}M_{t_{i+1}^n}(M_{t_{i+1}^n}-M_{t_i^n}).$$

可得

$$\sum_{i=0}^{p_n-1}M_{t_{i+1}^n}(M_{t_{i+1}^n}-M_{t_i^n})\xrightarrow{\mathsf{P}}\int_0^t M_s\,\mathrm{d}M_s+\langle M,M\rangle_t.$$

而

$$\sum_{i=0}^{p_n-1}M_{t_i^n}(M_{t_{i+1}^n}-M_{t_i^n})\xrightarrow{\mathsf{P}}\int_0^t M_s\,\mathrm{d}M_s,$$

二式相加

$$\lim_{n\to\infty}\sum_{i=0}^{p_n-1}[(M_{t_{i+1}^n})^2-(M_{t_i^n})^2]=2\int_0^t M_s\,\mathrm{d}M_s+\langle M,M\rangle_t.$$

即 $(M_t)^2-(M_0)^2=2\int_0^t M_s\,\mathrm{d}M_s+\langle M,M\rangle_t$. 这是下一章中 Itô 公式的基本例子.

例 4.3.2 设 M 是连续局部鞅, 求 $\langle M^2, M^2\rangle$.

解: $M_t^2 = M_0^2 + 2\int_0^t M_s\,\mathrm{d}M_s + \langle M, M\rangle_t$, 不妨设 $M_0 = 0$. 因此

$$\begin{aligned}\langle M^2, M^2\rangle_t =&\langle 2\int_0^t M_s\,\mathrm{d}M_s + \langle M, M\rangle, 2\int_0^t M_s\,\mathrm{d}M_s + \langle M, M\rangle\rangle_t\\ =&4\int_0^t M_s^2\,\mathrm{d}\langle M, M\rangle_s.\end{aligned}$$

■

例 4.3.3 若 M 是连续局部鞅, 求 M^n $(n \geqslant 2)$.

解: 由于

$$M_t^2 = M_0^2 + 2\int_0^t M_s\,\mathrm{d}M_s + \langle M, M\rangle_t.$$

若 M, N 是连续局部鞅, 由 $MN = \frac{1}{4}((M+N)^2 - (M-N)^2)$ 知

$$M_tN_t = M_0N_0 + \int_0^t M_s\,\mathrm{d}N_s + \int_0^t N_s\,\mathrm{d}M_s + \langle M, N\rangle.$$

下面求 M^3. 由于 M^3 是 M^2 与 M 的乘积, 不妨设 $M_0 = 0$, 事实上,

$$M_t^3 = M_t^2 M_t = \int_0^t M_s^2\,\mathrm{d}M_s + \int_0^t M_s\,\mathrm{d}M_s^2 + \langle M^2, M\rangle,$$

而

$$\int_0^t M_s\,\mathrm{d}M_s^2 = 2\int_0^t M_s^2\,\mathrm{d}M_s + \int_0^t M_s\,\mathrm{d}\langle M, M\rangle_s,$$

$$\langle M^2, M\rangle = 2\int_0^t M_s\,\mathrm{d}\langle M, M\rangle_s.$$

故

$$M_t^3 = 3\int_0^t M_s^2\,\mathrm{d}M_s + 3\int_0^t M_s\,\mathrm{d}\langle M, M\rangle_s.$$

利用数学归纳法, 有

$$M_t^n = M_0^n + n\int_0^t M_s^{n-1}\,\mathrm{d}M_s + \frac{1}{2}n(n-1)\int_0^t M_s^{n-2}\,\mathrm{d}\langle M, M\rangle_s.$$

■

例 4.3.4 设 X 是一个连续平方可积鞅且有独立平稳增量①, $X_0 = 0$. 证明:

$$\langle X, X\rangle_t = \mathsf{E}[X_1^2]t.$$

① 随机过程 X 具有独立平稳增量是指: 对于 $t > s$, $X_t - X_s$ 与 $\sigma(X_r : 0 \leqslant r \leqslant s)$ 独立, $X_t - X_s$ 与 $X_{t-s} - X_0$ 同分布.

证明: 若 $t > s$, 则

$$\begin{aligned}
&\mathsf{E}[X_t^2 - \mathsf{E}[X_1^2]t \mid \mathcal{F}_s] \\
=&\mathsf{E}[(X_t - X_s + X_s)^2 - \mathsf{E}[X_1^2]t \mid \mathcal{F}_s] \\
=&\mathsf{E}[(X_t - X_s)^2 + X_s^2 + 2(X_t - X_s)X_s - \mathsf{E}[X_1^2]t \mid \mathcal{F}_s] \\
=&\mathsf{E}[(X_t - X_s)^2] + X_s^2 + 2\mathsf{E}[(X_t - X_s)X_s \mid \mathcal{F}_s] - \mathsf{E}[X_1^2]t.
\end{aligned}$$

为了证明结论，需要讨论 $\mathsf{E}[X_t^2]$. 设 $f(t) = \mathsf{E}[X_t^2]$. 首先考虑

$$\mathsf{E}[X_n^2] = \mathsf{E}[\sum_{i=1}^n (X_i^2 - X_{i-1}^2)] = \sum_{i=1}^n \mathsf{E}[X_i^2 - X_{i-1}^2].$$

事实上,

$$\begin{aligned}
\mathsf{E}[(X_i - X_{i-1})^2] =&\mathsf{E}[\mathsf{E}[(X_i - X_{i-1})^2 \mid \mathcal{F}_{i-1}]] \\
=&\mathsf{E}[\mathsf{E}[(X_i^2 + X_{i-1}^2 - 2X_iX_{i-1}) \mid \mathcal{F}_{i-1}]] \\
=&\mathsf{E}[\mathsf{E}[X_i^2 - X_{i-1}^2 \mid \mathcal{F}_{i-1}]] \\
=&\mathsf{E}[X_i^2 - X_{i-1}^2].
\end{aligned}$$

故

$$\mathsf{E}[X_n^2] = \sum_{i=1}^n \mathsf{E}[(X_i - X_{i-1})^2] = n\mathsf{E}[X_1^2].$$

因此 $f(n) = n\mathsf{E}X_1^2$. 对于任意 $t > 0,\ s > 0$, $f(t+s) = \mathsf{E}[X_{t+s}^2]$, $f(s) = \mathsf{E}[X_s^2]$, 由于

$$\begin{aligned}
\mathsf{E}[X_{t+s}^2 - X_s^2] =&\mathsf{E}[\mathsf{E}[(X_{t+s}^2 - X_s^2) \mid \mathcal{F}_s]] \\
=&\mathsf{E}[\mathsf{E}[(X_{t+s} - X_s)^2 \mid \mathcal{F}_s]] \\
=&\mathsf{E}[(X_{t+s} - X_s)^2] \\
=&\mathsf{E}[X_t^2] = f(t),
\end{aligned}$$

故 $f(t+s) = f(t) + f(s)$. 易知 $f(t) = \mathsf{E}[X_1^2]t$，于是 $\mathsf{E}[(X_t - X_s)^2] = (t-s)\mathsf{E}[X_1^2]$. 从而

$$\mathsf{E}(X_t^2 - \mathsf{E}[X_1^2]t \mid \mathcal{F}_s) = X_s^2 - \mathsf{E}[X_1^2]s.$$

于是 $\langle X, X\rangle_t = \mathsf{E}[X_1^2]t$. ■

上面的例子实际上说明了，如果对于连续局部鞅 M, 如果能找到一个连续增过程 A_t, 使得 $(M_t^2 - A_t)_{t\geqslant 0}$ 是局部鞅，那么 $\langle M, M\rangle = A$. 进一步，如果对于连

续局部鞅 M, N, 如果能找到一个连续增过程 A_t, 使得 $(M_tN_t - A_t)_{t\geqslant 0}$ 是局部鞅, 那么 $\langle M, N\rangle = A$. 如果 M, N 其中一个变成连续的适应增过程, 结论将会有很大的变化. 请看下面的定理.

定理 4.3.3 设 Z 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上有界随机变量, A 是从 0 出发的适应的一致有界连续增过程. 则

$$\mathsf{E}[ZA_\infty] = \mathsf{E}[\int_0^\infty \mathsf{E}[Z \mid \mathcal{F}_t]\,\mathrm{d}A_t].$$

证明 [3]: 设 $M_t = \mathsf{E}[Z \mid \mathcal{F}_t]$, 易知 $M = (M_t)_{t\geqslant 0}$ 是有界鞅, 有收敛的极限, 且 $M_\infty = Z$. 故 $\mathsf{E}[M_\infty A_\infty] = \mathsf{E}[ZA_\infty]$, 于是

$$\int_0^\infty \mathsf{E}[Z \mid \mathcal{F}_t]\,\mathrm{d}A_t = \int_0^\infty M_t\,\mathrm{d}A_t.$$

设 $C_t = \inf\{s : A_s \geqslant t\}$, 显然 $\{C_t \leqslant s\} = \{A_s \geqslant t\}$. 首先令 $H_s(w) = \mathbb{1}_{[0,t]}(s)$, 有

$$\int_0^\infty H_s(w)\,\mathrm{d}A_s(w) = \int_0^\infty H_{C_s(w)}\mathbb{1}_{\{C_{s(w)}<\infty\}}\,\mathrm{d}s.$$

由单调类定理, 知当 H 是有界可测过程时也成立. 对于任意停时 T,

$$\begin{aligned}
\mathsf{E}[M_TA_T] =& \mathsf{E}[\int_0^\infty M_T\mathbb{1}_{\{s\leqslant T\}}\,\mathrm{d}A_s] \quad (\text{令}H_s = M_T\mathbb{1}_{\{s\leqslant T\}})\\
=& \mathsf{E}[\int_0^\infty M_T\mathbb{1}_{\{C_s<+\infty\}}\mathbb{1}_{\{C_s\leqslant T\}}\,\mathrm{d}s]\\
=& \int_0^\infty \mathsf{E}[M_T\mathbb{1}_{\{C_s<+\infty\}}\mathbb{1}_{\{C_s\leqslant T\}}]\,\mathrm{d}s\\
=& \int_0^\infty \mathsf{E}[M_{C_s}\mathbb{1}_{\{C_s<+\infty\}}\mathbb{1}_{\{C_s\leqslant T\}}]\,\mathrm{d}s\\
=& \mathsf{E}[\int_0^\infty M_{C_s}\mathbb{1}_{\{C_s<+\infty\}}\mathbb{1}_{\{C_s\leqslant T\}}\,\mathrm{d}s]\\
=& \mathsf{E}[\int_0^\infty M_s\mathbb{1}_{\{s\leqslant T\}}\,\mathrm{d}A_s] \quad (\text{令}H_s = M_s\mathbb{1}_{\{s\leqslant T\}})\\
=& \mathsf{E}[\int_0^T M_s\,\mathrm{d}A_s].
\end{aligned}$$

从而, $\mathsf{E}[M_TA_T] = \mathsf{E}[\int_0^T M_s\,\mathrm{d}A_s]$ 对任意停时 T 都成立. 故 $(A_tM_t - \int_0^t M_s\,\mathrm{d}A_s)$ 是鞅. 因此

$$\mathsf{E}[A_tM_t - \int_0^t M_s\,\mathrm{d}A_s] = \mathsf{E}[A_0M_0 - \int_0^0 M_s\,\mathrm{d}A_s] = 0.$$

由 Z 与 A 的有界性, 可知结论成立. ■

4.4 习题

1. 设 $M=(M_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathsf{P})$ 上的平方可积鞅. 若 $s<t$, 请证明 $\mathsf{E}[(M_t-M_s)^2]=\mathsf{E}[M_t^2-M_s^2]$.

2. 若 X 满足

$$dX_t=0.5X_tdt+3X_tdB_t,$$

这里 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 若 $X_0=1/2$, 求 $\mathsf{P}(X_2<2)$.

3. 设 $B=(B_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动, 设 $H(x)=\mathbb{1}_{\{x>0\}}-\mathbb{1}_{\{x<0\}}$, 求 $\int_0^1 H(B_s)dB_s$ 的分布.

4. 设 $B=(B_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动, 求 $\int_0^1 \sin s dB_s$ 的分布.

5. 设 $B=(B_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动, 设 $M_t=\int_0^t e^{B_s}dB_s$, 求 $\mathsf{E}[M_1M_2]$.

6. 设 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\},\mathsf{P})$ 上的标准布朗运动. 定义 $(A_t)_{t\geqslant 0}$ 如下: 当 $0\leqslant t\leqslant 1$ 时, $A_t=1$; 当 $t>1$ 时, $A_t=\begin{cases}1, & B_1>0\\ B_1, & B_1<0\end{cases}$, 令 $Z_t=\int_0^t A_s\,\mathrm{d}B_s$.

(1) 求 $\mathsf{E}[Z_3]$, $\mathsf{E}[Z_3^2]$;

(2) 求 $\langle Z,Z\rangle_3$, $\mathsf{E}[(Z_3-Z_1)^2\mid\mathcal{F}_1]$.

7. B 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 定义

$$X_t=\int_0^t(1-s)dB_s, Y_t=\int_0^t(1+s)dB_s.$$

求 t 为何值时, X_t 与 Y_t 独立?

8. B 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动, 定义

$$X_t=\int_0^t sgn(B_s)dB_s.$$

求 $Cov(X_t,B_t)$.

第 5 章　Itô 公式及其应用

本章将介绍 Itô 公式及其相关应用.

5.1　Itô 公式

我们在前面的讨论中已经知道, 若 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, 那么 B 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的鞅. 现在考虑 B^2. 我们知道, B^2 已经不是鞅了, $(B_t^2 - t)_{t\geqslant 0}$ 是鞅. 由前面的讨论, 我们也知道, 对于 B^3, $(B_t^3 - 3\int_0^t B_s \,\mathrm{d}s = 3\int_0^t B_s^2 \,\mathrm{d}B_s)_{t\geqslant 0}$ 是局部鞅. 如果考虑 $\sin B$, $\cos B$ 会有什么结论? Itô 公式将告诉我们答案.

Itô 公式是随机分析中十分重要的内容. 我们首先给出 Itô 公式及其证明. 为简单起见, 我们首先给出关于布朗运动的 Itô 公式.

定理 5.1.1　设 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, f 是 $\mathbb{R}$ 上的二阶连续可微函数, 则对任意 $t \geqslant 0$,

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s$$

证明: 考虑 $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$, 且当 $n \to \infty$ 时, 假设 $\max\limits_{1\leqslant i\leqslant n} |t_i^n - t_{i-1}^n| \to 0$.

$$f(B_t) = f(B_0) + \sum_{i=0}^{p_n-1} (f(B_{t_{i+1}^n}) - f(B_{t_i^n})).$$

由 Talyor 公式,

$$\begin{aligned} f(B_{t_{i+1}^n}) - f(B_{t_i^n}) =& f'(B_{t_i^n})(B_{t_{i+1}^n} - B_{t_i^n}) \\ &+ \frac{1}{2} f''(B_{t_i^n} + c(B_{t_{i+1}^n} - B_{t_i^n}))(B_{t_{i+1}^n} - B_{t_i^n})^2 \end{aligned}$$

$c \in [0, 1]$. 由命题 4.3.6 ,

$$\sum_{i=0}^{p_n-1} f'(B_{t_i^n})(B_{t_{i+1}^n} - B_{t_i^n}) \xrightarrow{\mathsf{P}} \int_0^t f'(B_s)\,\mathrm{d}B_s.$$

设 $g_{n,i}=f''(B_{t_i^n}+c(B_{t_{i+1}^n}-B_{t_i^n}))$, 有

$$\max_{0\leqslant i\leqslant p_n-1}|g_{n,i}-f''(B_{t_i^n})|\leqslant \sup_{0\leqslant i\leqslant p_n-1}\Big(\sup_{x\in[B_{t_i^n}\wedge B_{t_{i+1}^n},B_{t_i^n}\vee B_{t_{i+1}^n}]}|f''(x)-f''(B_{t_i^n})|\Big).$$

由于 f'' 在一个闭区间上一致连续, 且 $\sum\limits_{i=0}^{p_n-1}(B_{t_{i+1}^n}-B_{t_i^n})^2\xrightarrow{\mathsf{P}}t$, 故

$$\Big|\sum_{i=0}^{p_n-1}g_{n,i}(B_{t_{i+1}^n}-B_{t_i^n})^2-\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_{t_{i+1}^n}-B_{t_i^n})^2\Big|\xrightarrow{\mathsf{P}}0.$$

由第四章例 4.3.1,

$$(B_{t_{i+1}^n}-B_{t_i^n})^2=2\int_{t_i^n}^{t_{i+1}^n}(B_s-B_{t_i^n})\,\mathrm{d}B_s+(t_{i+1}^n-t_i^n),$$

注意到

$$\begin{aligned}&\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_{t_{i+1}^n}-B_{t_i^n})^2-\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(t_{i+1}^n-t_i^n)\\=\;&2\sum_{i=0}^{p_n-1}f''(B_{t_i^n})\int_{t_i^n}^{t_{i+1}^n}(B_s-B_{t_i^n})\,\mathrm{d}B_s\\=\;&2\int_0^t\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_s-B_{t_i^n})\mathbb{1}_{(t_i^n,t_{i+1}^n]}(s)\,\mathrm{d}B_s,\end{aligned}$$

又由于

$$\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_s-B_{t_i^n})\mathbb{1}_{(t_i^n,t_{i+1}^n]}(s)\to 0,$$

且

$$\Big|\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_s-B_{t_i^n})\mathbb{1}_{(t_i^n,t_{i+1}^n]}\Big|\leqslant 2\sup_s|f''(B_s)B_s|,$$

由命题 4.3.6,

$$\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_{t_{i+1}^n}-B_{t_i^n})^2-\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(t_{i+1}^n-t_i^n)\xrightarrow{\mathsf{P}}0.$$

因此

$$\frac{1}{2}\sum_{i=0}^{p_n-1}f''(B_{t_i^n})(B_{t_{i+1}^n}-B_{t_i^n})^2\xrightarrow{\mathsf{P}}\frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s.$$

得证. ■

当对一般的连续半鞅考虑上述定理的推广时，证明会稍微复杂一些, 我们这里列出对一般连续半鞅的证明, 希望读者对比两种方法，体会其中的不同. 事实上, 这两个定理证明的前半部分是一样的.

定理 5.1.2　设 $X^1,\cdots,X^p$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上 p 个连续半鞅, F 是 $\mathbb{R}^p$ 上的二阶连续可微函数, 则对任意 $t\geqslant 0$,

$$\begin{aligned}F(X_t^1,\cdots,X_t^p)=&F(X_0^1,\cdots,X_0^p)+\sum_{i=1}^p\int_0^t\frac{\partial F}{\partial x^i}(X_s^1,\cdots,X_s^p)\,\mathrm{d}X_s^i\\&+\frac{1}{2}\sum_{i,j=1}^p\int_0^t\frac{\partial^2 F}{\partial x^i\partial x^j}(X_s^1,\cdots,X_s^p)\,\mathrm{d}\langle X^i,X^j\rangle_s.\end{aligned}$$

证明 [6]: 首先考虑 $p=1$ 的情形. 为简单起见, 记 $X=X^1$. 令 $t>0$, 设 $0=t_0^n<t_1^n<\cdots<t_{p_n}^n=t$, 且当 $n\to\infty$ 时, 假设 $\max\limits_{1\leqslant i\leqslant n}|t_i^n-t_{i-1}^n|\to 0$, 注意到

$$F(X_t)=F(X_0)+\sum_{i=0}^{p_n-1}(F(X_{t_{i+1}^n})-F(X_{t_i^n})).$$

由 Talyor 公式,

$$\begin{aligned}F(X_{t_{i+1}^n})-F(X_{t_i^n})=&F'(X_{t_i^n})(X_{t_{i+1}^n}-X_{t_i^n})\\&+\frac{1}{2}F''(X_{t_i^n}+c(X_{t_{i+1}^n}-X_{t_i^n}))(X_{t_{i+1}^n}-X_{t_i^n})^2\end{aligned}$$

$c\in[0,1]$. 由命题 4.3.6 ,

$$\sum_{i=0}^{p_n-1}F'(X_{t_i^n})(X_{t_{i+1}^n}-X_{t_i^n})\xrightarrow{\mathsf{P}}\int_0^t F'(X_s)\,\mathrm{d}X_s.$$

设 $f_{n,i}=F''(X_{t_i^n}+c(X_{t_{i+1}^n}-X_{t_i^n}))$, 有

$$\max_{0\leqslant i\leqslant p_n-1}|f_{n,i}-F''(X_{t_i^n})|\leqslant\sup_{0\leqslant i\leqslant p_n-1}(\sup_{x\in[X_{t_i^n}\wedge X_{t_{i+1}^n},X_{t_i^n}\vee X_{t_{i+1}^n}]}|F''(x)-F''(X_{t_i^n})|).$$

由于 F'' 在一个闭区间上一致连续, 且 $\sum\limits_{i=0}^{p_n-1}(X_{t_{i+1}^n}-X_{t_i^n})^2\xrightarrow{\mathsf{P}}\langle X,X\rangle_t$, 故

$$\Big|\sum_{i=0}^{p_n-1}f_{n,i}(X_{t_{i+1}^n}-X_{t_i^n})^2-\sum_{i=0}^{p_n-1}F''(X_{t_i^n})(X_{t_{i+1}^n}-X_{t_i^n})^2\Big|\xrightarrow{\mathsf{P}}0.$$

令 $\mu_n(\mathrm{d}r)=\sum_{i=0}^{p_n-1}(X_{t_{i+1}^n}-X_{t_i^n})^2\delta_{t_i^n}(\mathrm{d}r)$, 由于 $\mu_n([0,r])\xrightarrow{\mathsf{P}}\langle X,X\rangle_r$, 故存在子列 $\{n_k\}$, 使得 $\mu_{n_k}([0,r])\xrightarrow{a.s.}\langle X,X\rangle_r$.

若令 $\mu(\mathrm{d}r)=\mathbb{1}_{[0,t]}(r)\,\mathrm{d}\langle X,X\rangle_r$, 对于固定的 ω, 事实上会有 $\mu_{n_k}(\mathrm{d}r)$ 依分布收敛于 $\mu(\mathrm{d}r)$, 于是,对于任意有界连续函数 f, $\int_{[0,t]}f(r)\mu_{n_k}(\mathrm{d}r)\to\int_{[0,t]}f(r)\mu(\mathrm{d}r)$, 便有

$$\int_{[0,t]}F''(X_s)\mu_{n_k}(\mathrm{d}s)\xrightarrow{a.s.}\int_{[0,t]}F''(X_s)\mu(\mathrm{d}s).$$

故

$$\int_{[0,t]}F''(X_s)\mu_n(\mathrm{d}s)\xrightarrow{\mathsf{P}}\int_{[0,t]}F''(X_s)\mu(\mathrm{d}s).$$

即

$$\sum_{0}^{p_n-1}f_{n,i}(X_{t_{i+1}^n}-X_{t_i^n})^2\xrightarrow{\mathsf{P}}\int_0^t F''(X_s)\,\mathrm{d}\langle X,X\rangle_s.$$

于是

$$F(X_t)=F(X_0)+\int_0^t F'(X_s)\,\mathrm{d}X_s+\frac{1}{2}\int_0^t F''(X_s)\,\mathrm{d}\langle X,X\rangle_s.$$

下面考虑 $p-$ 维情形. 注意到

$$\begin{aligned}&F(X_{t_{i+1}^n}^1,\cdots,X_{t_{i+1}^n}^p)-F(X_{t_i^n}^1,\cdots,X_{t_i^n}^p)\\=&\sum_{k=1}^p\frac{\partial F}{\partial x^k}(X_{t_i^n}^1,\cdots,X_{t_i^n}^p)(X_{t_{i+1}^n}^k-X_{t_i^n}^k)+\sum_{k,l=1}^p\frac{f_{n,i}^{k,l}}{2}(X_{t_{i+1}^n}^k-X_{t_i^n}^k)(X_{t_{i+1}^n}^l-X_{t_i^n}^l),\end{aligned}$$

这里

$$f_{n,i}^{k,l}=\frac{\partial^2F}{\partial x_k\partial x_l}(X_{t_i^n}+c(X_{t_{i+1}^n}-X_{t_i^n}),\ c\in[0,1],\ X_t=(X_t^1,\cdots,X_t^p).$$

类似上面的证明, 即可得结论. ■

下面, 我们给出一些关于 Itô 公式的应用. 为叙述方便，记 $\mathbf{C}^2$ 函数是二阶连续可微函数. 若 f 是 $\mathbb{R}^2$ 上的 $\mathbf{C}^2$ 函数, B 是标准布朗运动, 则

$$\begin{aligned}f(t,B_t)=&f(0,B_0)+\int_0^t\partial_xf(s,B_s)\,\mathrm{d}B_s\\&+\int_0^t\partial_sf(s,B_s)\,\mathrm{d}s+\frac{1}{2}\int_0^t\partial_{xx}f(s,B_s)\,\mathrm{d}s.\end{aligned}$$

例 5.1.1 若 $B=(B_t)_{t\geqslant0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant0},\mathsf{P})$ 上的标准布朗运动. 计算 $\int_0^t s\,\mathrm{d}B_s$.

解:

令 $f(t,B_t)=tB_t$, 注意到 $\partial_t f(t,B_t)=B_t$, $\partial_x f(t,B_t)=t$, $\partial_{xx} f(t,B_t)=0$. 故

$$tB_t=\int_0^t B_s\,\mathrm{d}s+\int_0^t s\,\mathrm{d}B_s,$$

于是

$$\int_0^t s\,\mathrm{d}B_s=tB_t-\int_0^t B_s\,\mathrm{d}s.$$

■

利用 Itô 公式，我们可以计算 $e^{\sigma B_t}$, 这里 $\sigma>0$. 事实上，$f(x)=e^{\sigma x}$, 经简单计算可以得到

$$e^{\sigma B_t}=e^{\sigma B_0}+\int_0^t \sigma e^{\sigma B_s}\,\mathrm{d}B_s+\frac{1}{2}\int_0^t \sigma^2 e^{\sigma B_s}\,\mathrm{d}s,$$

若借用方程形式, 令 $X_t=e^{\sigma B_t}$, 我们可以用

$$\mathrm{d}X_t=\sigma X_t\,\mathrm{d}B_t+\frac{\sigma^2}{2}X_t\,\mathrm{d}t$$

表示 X. 这里要注意 $\mathrm{d}B_t$ 与 $\mathrm{d}t$ 的意义是不一样的.

进一步，如果考虑 e^{at+bB_t}, 令 $f(t,x)=e^{at+bx}$, 由 Itô 公式,

$$e^{at+bB_t}=1+\int_0^t ae^{as+bB_s}\,\mathrm{d}s+\int_0^t be^{as+bB_s}\,\mathrm{d}B_s+\frac{1}{2}\int_0^t b^2e^{as+bB_s}\,\mathrm{d}s.$$

令 $X_t=e^{at+bB_t}$, 可以用

$$\mathrm{d}X_t=(a+\frac{b^2}{2})X_t\,\mathrm{d}t+bX_t\,\mathrm{d}B_t$$

表示 X_t.

反过来，注意到一个半鞅 X, 若满足 $\mathrm{d}X_t=mX_t\,\mathrm{d}t+\sigma X_t\,\mathrm{d}B_t$, 则

$$X_t=X_0\exp\{(m-\frac{\sigma^2}{2})t+\sigma B_t\}.$$

往往称 $X=(X_t)_{t\geqslant 0}$ 是几何布朗运动.

例 5.1.2　若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动. 令 $X_t=e^{\frac{1}{2}t}\sin B_t$, $Y_t=e^{\frac{1}{2}t}\cos B_t$. 利用 Itô 公式计算 $\mathsf{E}[X_t]$, $\mathsf{E}[Y_t]$.

解: 由 Itô 公式, 令 $f(t,x)=e^{\frac{1}{2}t}\sin x$.

$$
\begin{aligned}
X_t = f(t,B_t) =& f(0,B_0)+\int_0^t \frac{1}{2}e^{\frac{1}{2}s}\sin B_s\,\mathrm{d}s\\
&+\int_0^t e^{\frac{1}{2}s}\cos B_s\,\mathrm{d}B_s-\int_0^t \frac{1}{2}e^{\frac{1}{2}s}\sin B_s\,\mathrm{d}s\\
=&\int_0^t e^{\frac{1}{2}s}\cos B_s\,\mathrm{d}B_s.
\end{aligned}
$$

利用命题 4.2.6, $\mathsf{E}[\int_0^t e^s\cos^2 B_s\,\mathrm{d}s]<\infty$，$\int_0^t e^{\frac{1}{2}s}\cos B_s\,\mathrm{d}B_s$ 是鞅, 因此 $\mathsf{E}[X_t]=0$.

令 $g(t,x)=e^{\frac{1}{2}t}\cos x$.

$$
\begin{aligned}
Y_t = g(t,B_t) =& g(0,B_0)+\frac{1}{2}\int_0^t e^{\frac{1}{2}s}\cos B_s\,\mathrm{d}s\\
&-\int_0^t e^{\frac{1}{2}s}\sin B_s\,\mathrm{d}B_s-\frac{1}{2}\int_0^t e^{\frac{1}{2}s}\cos B_s\,\mathrm{d}s\\
=&1-\int_0^t e^{\frac{1}{2}s}\sin B_s\,\mathrm{d}B_s.
\end{aligned}
$$

利用命题 4.2.6, $\int_0^t e^{\frac{1}{2}s}\sin B_s\,\mathrm{d}B_s$ 是鞅, 因此 $\mathsf{E}[Y_t]=1$. ■

例 5.1.3 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动. $B_0=0$, 设 a,b 都是非负实数,

$$
T=\inf\{t>0: B_t=-a \text{ 或 } B_t=b\}.
$$

利用 Itô 公式计算 $\mathsf{E}[TB_T]$.

解: 由 Itô 公式, 令 $f(t,x)=tx+g(x)$.

$$
\begin{aligned}
X_t = f(t,B_t) =& f(0,B_0)+\int_0^t B_s\,\mathrm{d}s\\
&+\int_0^t (s+g'(B_s))\,\mathrm{d}B_s+\int_0^t \frac{1}{2}g''(B_s)\,\mathrm{d}s
\end{aligned}
$$

当 $g(x)=-\frac{1}{3}x^3$ 时, $X=(X_t)_{t\geqslant 0}$ 是局部鞅. 利用命题 4.2.6, 可以验证 $X=(X_t)_{t\geqslant 0}$ 是鞅. 再利用控制收敛定理，$\mathsf{E}[TB_T-\frac{1}{3}B_T^3]=0$. 因此

$$
\mathsf{E}[TB_T]=\mathsf{E}[\frac{1}{3}B_T^3]=\frac{ab}{3}(b-a).
$$

■

对于 Markov 过程, 参考下面的例子.

例 5.1.4 设 $X=(X_t)_{t\geqslant 0}$ 满足如下关系:

$$\mathrm{d}X_t=\frac{a}{X_t}\,\mathrm{d}t+\mathrm{dB}_t,\ X_0=x_0>0,\ a>0.$$

令 $T=\inf\{t:X_t=0\}$, 考虑何种条件下 $\mathsf{P}(T=\infty)=1$.

设 $0<r<x<R<+\infty$, 令 $\tau=\inf\{t:X_t=r \text{ 或 } X_t=R\}$, $\phi(x)=\mathsf{P}(X_\tau=R\mid X_0=x)$, 故 $\phi(r)=0,\ \phi(R)=1$.

令 $J=\mathbb{1}_{\{X_\tau=R\}}$, $M_t=\mathsf{E}[J\mid\mathcal{F}_t]$, 故 $M=(M_t)_{t\geqslant 0}$ 是一个鞅. 由 X 是 Markov 过程, 于是

$$\mathsf{E}[J\mid\mathcal{F}_t]=\mathsf{E}[J\mid X_{t\wedge\tau}]=\phi(X_{t\wedge\tau}).$$

假设 ϕ 是二阶连续可微的. 由 Itô 公式

$$\begin{aligned}\phi(X_{t\wedge\tau})=&\phi(X_0)+\int_0^{t\wedge\tau}\phi'(X_s)\,\mathrm{d}X_s+\frac{1}{2}\int_0^{t\wedge\tau}\phi''(X_s)\,\mathrm{d}\langle X,X\rangle_s\\ =&\phi(X_0)+\int_0^{t\wedge\tau}\phi'(X_s)\,\mathrm{d}B_s+\int_0^{t\wedge\tau}\frac{a\phi'(X_s)}{X_s}+\frac{1}{2}\phi''(X_s)\,\mathrm{d}s.\end{aligned}$$

由于 $\phi(X_{t\wedge\tau})$ 是鞅. 故 $x\phi''(x)+2a\phi'(x)=0$.

由 $\phi(r)=0,\ \phi(R)=1$, 故 $a\neq\frac{1}{2}$ 时,

$$\phi(x)=\frac{x^{1-2a}-r^{1-2a}}{R^{1-2a}-r^{1-2a}}.$$

当 $a=\frac{1}{2}$ 时,

$$\phi(x)=\frac{\log x-\log r}{\log R-\log r}.$$

从而当 $a\geqslant\frac{1}{2}$ 时, $\lim\limits_{r\to 0}\phi(x)=1$, 即 $X=\{X_t\}$ 先到达 R 处. 因此到达 0 的概率为 0, 即 $\mathsf{P}(T=\infty)=1$.

当 $a<\frac{1}{2}$ 时

$$\lim_{r\to 0}\phi(x)=(\frac{x}{R})^{1-2a}.$$

若 $T<+\infty$, 必有 $R=\infty$, 于是

$$\mathsf{P}(T<+\infty)=\lim_{R\to+\infty}1-(\frac{x}{R})^{1-2a}=1.$$

总结: $a\geqslant\frac{1}{2}$ 时, $\mathsf{P}(T=\infty)=1$. $a<\frac{1}{2}$ 时, $\mathsf{P}(T<\infty)=1$.

5.2 Doléans-Dade 指数与鞅表示定理

前面我们已经讨论了连续局部鞅的一些性质, 已经知道连续的 Itô 积分是一个局部鞅. 我们或许会有一个疑问: 若 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geqslant 0},\mathsf{P})$ 上的连续局部鞅, $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动, 在什么条件下，会存在一个循序可测过程 K, 使得

$$M_t=\int_0^t K_s\,\mathrm{d}B_s\ ?$$

事实上，鞅表示定理即可部分回答这一问题. 这一节中, 就鞅表示定理展开讨论. 我们首先介绍 Doléans-Dade 指数. Doléans-Dade 指数在鞅表示定理及测度变换的讨论中起了关键作用.

令 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的连续局部鞅. 考虑 $X_t=\exp\{M_t-\frac{1}{2}\langle M,M\rangle_t\}$. (不妨假设 $X_0=1$, $M_0=0$), 令 $f(x,y)=e^{x-y}$, 由 Itô 公式

$$\begin{aligned}f(M_t,\frac{1}{2}\langle M,M\rangle_t)=&f(0,0)+\int_0^t e^{M_s-\frac{1}{2}\langle M,M\rangle_s}\,\mathrm{d}M_s\\&-\frac{1}{2}\int_0^t e^{M_s-\frac{1}{2}\langle M,M\rangle_s}\,\mathrm{d}\langle M,M\rangle_s\\&+\frac{1}{2}\int_0^t e^{M_s-\frac{1}{2}\langle M,M\rangle_s}\,\mathrm{d}\langle M,M\rangle_s.\end{aligned}$$

故 $X_t=1+\int_0^t X_s\,\mathrm{d}M_s$, 即可写为

$$\mathrm{d}X_t=X_t\,\mathrm{d}M_t.$$

这与轨道积分有很大的不同. 事实上, 若令 $A=(A_t)$ 是一个连续的有限变差过程时, 考虑 $\mathrm{d}X_t=X_t\,\mathrm{d}A_t$, 会有 $X_t=X_0e^{A_t}$.

通过 Doléans-Dade 指数可以解决很多问题. 首先, 回顾之前关于布朗运动刻画的结果, 我们发现可以通过布朗运动的 Doléans-Dade 指数得到更直接的结论.

定理 5.2.1 设 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的适应过程, 则下列论述等价:

(1) B 是标准布朗运动;

(2) B 是连续局部鞅, 且 $(B_t^2-t)_{t\geqslant 0}$ 是鞅.

证明: 我们只需要证明 (2)⇒(1).

若 B 是连续局部鞅, 且 $(B_t^2-t)_{t\geqslant 0}$ 是鞅. 考虑

$$M_t=\mathrm{i}\lambda B_t,\ \langle M,M\rangle_t=-\lambda^2 t,$$

由上面的讨论, 即有,

$$\exp\{\mathrm{i}\lambda B_t+\frac{1}{2}\lambda^2 t\}=1+\int_0^t \exp\{\mathrm{i}\lambda B_s+\frac{1}{2}\lambda^2 s\}\,\mathrm{d}(\mathrm{i}\lambda B_s),$$

分别考虑实部与虚部, 即可得到 $\exp\{\mathrm{i}\lambda B_t+\frac{1}{2}\lambda^2 t\}$ 是鞅.

由定理 3.2.11, B 是标准布朗运动当且仅当 B 是连续鞅, 且 $(\exp\{\mathrm{i}\lambda B_t+\frac{1}{2}\lambda^2 t\})_{t\geqslant 0}$ 是鞅. 因此 B 是标准布朗运动. ■

我们可以考虑下面这个例子:

例 5.2.1 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上, B 与 B' 都是标准布朗运动, 且相互独立, 则 $\langle B,B'\rangle_t=0$.

证明: 考虑 $X_t=\frac{1}{\sqrt{2}}(B_t+B'_t)$. 由于 B 与 B' 独立, 可知 $X=(X_t)_{t\geqslant 0}$ 是一个起点为 0, 轨道连续的中心化高斯过程. 对于 s, t,

$$\begin{aligned}Cov(X_s,X_t)=&Cov(\frac{1}{\sqrt{2}}(B_s+B'_s),\frac{1}{\sqrt{2}}(B_t+B'_t))\\=&\frac{1}{2}Cov(B_s+B'_s,B_t+B'_t)\\=&\frac{1}{2}Cov(B_s,B_t)+\frac{1}{2}Cov(B'_s,B'_t),\end{aligned}$$

故 $X=(X_t)$ 是标准布朗运动. 于是 $\langle X,X\rangle_t=t$. 即

$$\frac{1}{2}\langle B+B',B+B'\rangle=t.$$

我们得到 $\langle B,B'\rangle_t=0$. ■

下面, 利用 Doléans-Dade 指数讨论鞅表示定理.

设

$$f=\sum_{j=1}^{n}\lambda_j\mathbb{1}_{(t_{j-1},t_j]},\ \lambda_j\in\mathbb{R},\ 0=t_0<t_1<\cdots<t_n<\cdots$$

f 是 $[0,\infty)$ 上具有紧支撑的函数, 记所有具备这种形式的函数全体为 Γ, $\mathcal{E}_t^f=\exp\{\int_0^t f_s\,\mathrm{d}B_s-\frac{1}{2}\int_0^t f_s^2\,\mathrm{d}s\}$.

首先, 我们给出一个命题.

命题 5.2.1 $\{\mathcal{E}_\infty^f,\ f\in\Gamma\}$ 张成的线性空间在 $\mathbf{L}^2(\Omega,\mathcal{F}_\infty,\mathsf{P})$ 中稠密, 这里 $\mathcal{F}_\infty=\sigma(B_t,\ 0\leqslant t<+\infty)$.

证明 [6]: 我们只需证明: 对于非负随机变量 $Y\in\mathbf{L}^2(\Omega,\mathcal{F}_\infty,\mathsf{P})$, 若 Y 与任意的 $\mathcal{E}_\infty^f$, $f\in\Gamma$ 正交, 则通过 $\int_A Y\,\mathrm{d}\mathsf{P}=\mu(A)$ 定义的测度 μ 是零测度.

由 Γ 中函数的形式, 事实上只需证明对于任意给定的 $(t_1,\cdots,t_n)$, $\mu(\cdot)$ 在 $\sigma(B_{t_1},\cdots,B_{t_n})$ 上是零测度. 我们考虑

$$\widetilde{\mu}(F)=\mathsf{E}[Y\mathbb{1}_F(B_{t_1},B_{t_2}-B_{t_1},\cdots,B_{t_n}-B_{t_{n-1}})].$$

若 $\widetilde{\mu}(\cdot)$ 为零测度，必有 $\mu(\cdot)$ 是零测度. 因此, 我们只需证明 $\widetilde{\mu}(\cdot)$ 为零测度.

对于复数 z_i, $1\leqslant i\leqslant n$, 记

$$\varphi(z_1,\cdots,z_n)=\mathsf{E}[\exp\{\sum_{j=1}^n z_j(B_{t_j}-B_{t_{j-1}})\}Y].$$

这里 $\varphi(z_1,\cdots,z_n)$ 是复平面上的解析函数. 由于 Y 与任意的 $\mathcal{E}_\infty^f$, $f\in\Gamma$ 正交, 故对于任意的 $\lambda_i\in\mathbb{R}$, 记 $\Xi=\sum_{i=1}^n\lambda_i^2(t_i-t_{i-1})$, 有

$$\exp\{\frac{1}{2}\Xi\}\mathsf{E}[\exp\{\sum_{j=1}^n\lambda_j(B_{t_j}-B_{t_{j-1}})\}\cdot Y]=0.$$

即 $\varphi(\lambda_1,\cdots,\lambda_n)=0$. 说明 $\varphi(z_1,\cdots,z_n)$ 在实数轴上为 0. 由解析函数的解析延拓，$\varphi(z_1,\cdots,z_n)$ 在整个复平面上为 0. 此时, $\widetilde{\mu}(\cdot)$ 的傅立叶变换为 0, 故 $\widetilde{\mu}(\cdot)$ 为零测度. 得证. ■

进一步，我们有如下命题:

命题 5.2.2 对于 $F\in\mathbf{L}^2(\Omega,\mathcal{F}_\infty,\mathsf{P})$, $\mathcal{F}_\infty=\sigma(B_t:0\leqslant t<+\infty)$, 则存在唯一的可测过程 h, $\mathsf{E}[\int_0^\infty h_s^2\,\mathrm{d}s]<+\infty$, 使

$$F=\mathsf{E}[F]+\int_0^\infty h_s\,\mathrm{d}B_s.$$

证明 [12]: 首先证明唯一性. 若存在两个这样的过程: h_s', h_s'', 使

$$F=\mathsf{E}[F]+\int_0^\infty h_s'\,\mathrm{d}B_s=\mathsf{E}(F)+\int_0^\infty h_s''\,\mathrm{d}B_s,$$

则

$$\mathsf{E}[\int_0^\infty(h_s'-h_s'')^2\,\mathrm{d}s]=\mathsf{E}[(\int_0^\infty h_s'\,\mathrm{d}B_s-\int_0^\infty h_s''\,\mathrm{d}B_s)^2]=0.$$

故唯一性得证.

设满足命题结论的随机变量全体为 $\mathcal{H}$, (F^n) 是 $\mathcal{H}$ 中一个 $\mathbf{L}^2$ 意义下的柯西列, 记

$$F^n=\mathsf{E}(F^n)+\int_0^\infty h_s^n\,\mathrm{d}B_s.$$

由于 $\mathsf{E}[(F^n - F^m)^2] \to 0$, 故 $\mathsf{E}[\int_0^\infty (h_s^n - h_s^m)^2\,\mathrm{d}s] \to 0$, 故由 $\mathbf{L}^2(B)$(使 $\mathsf{E}[\int_0^\infty h_s^2\,\mathrm{d}s] < +\infty$ 成立的 h 全体) 性质, 存在 h, 使 h^n 在 $\mathbf{L}^2$ 意义下收敛于 h. 由于 $\mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$ 的完备性, F^n 在 $\mathbf{L}^2$ 意义下收敛至 F, h^n 在 $\mathbf{L}^2(B)$ 意义下收敛至 h, 故

$$F = \mathsf{E}[F] + \int_0^\infty h_s\,\mathrm{d}B_s.$$

故 $\mathcal{H}$ 是 $\mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$ 的闭子空间.

由 Doléans-Dade 指数性质

$$\mathcal{E}_\infty^f = 1 + \int_0^\infty \mathcal{E}_s^f f(s)\,\mathrm{d}B_s.$$

即 $\{\mathcal{E}_\infty^f,\ f \in \Gamma\} \in \mathcal{H}$, 由于 $\{\mathcal{E}_\infty^f,\ f \in \Gamma\}$ 张成的线性空间在 $\mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$ 中稠密，故 $\mathcal{H} = \mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$. ■

综上, 我们可以得到重要的鞅表示定理.

定理 5.2.2 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, $\mathcal{F}_t = \sigma(B_r, 0 \leqslant r \leqslant t)$, 则 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的任意局部鞅 M 可写为

$$M_t = M_0 + \int_0^t H_s\,\mathrm{d}B_s.$$

证明: 首先设 M 满足 $\sup\limits_t \mathsf{E}M_t^2 < +\infty$, 则存在 $M_\infty \in \mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$,

$$M_\infty = \mathsf{E}[M_\infty] + \int_0^\infty H_s\,\mathrm{d}B_s.$$

由条件数学期望性质,

$$\begin{aligned}\mathsf{E}[M_\infty \mid \mathcal{F}_t] =& \mathsf{E}[M_\infty] + \mathsf{E}[\int_0^\infty H_s\,\mathrm{d}B_s \mid \mathcal{F}_t] \\ =& \mathsf{E}[M_\infty] + \int_0^t H_s\,\mathrm{d}B_s.\end{aligned}$$

由于 $\sup\limits_t \mathsf{E}M_t^2 < +\infty$, 故 $M = (M_t)_{t\geqslant 0}$ 是一致可积的. 因此

$$M_t = \mathsf{E}[M_\infty \mid \mathcal{F}_t] = \mathsf{E}[M_\infty] + \int_0^t H_s\,\mathrm{d}B_s.$$

若 $M = (M_t)_{t\geqslant 0}$ 是一致可积的, 由于 $\mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$ 在 $\mathbf{L}^1(\Omega, \mathcal{F}_\infty, \mathsf{P})$ 中稠密, 故存在 $M_\infty^n \in \mathbf{L}^2(\Omega, \mathcal{F}_\infty, \mathsf{P})$, 使 $\mathsf{E}[|M_\infty - M_\infty^n|] \to 0$. 通过

$$\mathsf{P}(\sup_t |M_t - M_t^n| > \lambda) \leqslant \frac{1}{\lambda}\mathsf{E}[|M_\infty - M_\infty^n|]$$

易知 M 具有连续轨道. 若 M 是局部鞅, 则通过局部化技巧即可得. ■

定理 5.2.2 说明若带流概率空间的流是布朗运动生成的自然流，则这个带流概率空间上的局部鞅都是连续的. 因此可以体会到，一个随机过程的鞅的性质与带流概率空间上的概率测度与流密切相关.

鞅表示定理是随机分析的重要内容, 如果定理 5.2.2 中的流发生改变, $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上存在局部鞅 M，则在一定条件下，任意局部鞅 X, 会有

$$X = X_0 + K \cdot M + M^{\perp},$$

这里 $M^{\perp}$ 是局部鞅，且满足 $M^{\perp}M$ 是局部鞅，往往称 $M^{\perp}$ 为与 M 的正交的局部鞅. 上述结论的细节，读者可参考 Jacod 和 Shiryaev [3].

这里值得注意的是，鞅表示定理是关于局部鞅过程的. 如果考虑单个随机变量的相关性质，要根据单个随机变量的各自性质了. 这与鞅表示定理不矛盾, 这也是命题 5.2.2 反映的事实.

例 5.2.2 $B = (B_t)_{t\geqslant 0}$ 是标准布朗运动. 对于固定的 T, 计算 e^{B_T} 在命题 5.2.2 下的表示.

解: 由 Itô 公式, 考虑 $e^{B_t - \frac{1}{2}t}$, 令 $f(t,x) = e^{x-\frac{1}{2}t}$, 有

$$\begin{aligned}
&\partial_t f(t,x) = -\frac{1}{2}e^{x-\frac{1}{2}t},\ \partial_x f(t,x) = e^{x-\frac{1}{2}t},\\
&\partial_{xx} f(t,x) = e^{x-\frac{1}{2}t},\\
&e^{B_t-\frac{1}{2}t} = 1 + \int_0^t e^{B_s - \frac{1}{2}s}\,\mathrm{d}B_s,
\end{aligned}$$

对于固定的 T,

$$e^{B_T} = e^{\frac{1}{2}T} + e^{\frac{1}{2}T}\int_0^T e^{B_s-\frac{1}{2}s}\,\mathrm{d}B_s.$$

于是

$$e^{B_T} = \mathsf{E}e^{B_T} + \int_0^T e^{B_s + \frac{1}{2}(T-s)}\,\mathrm{d}B_s.$$

■

例 5.2.3 $B = (B_t)_{t\geqslant 0}$ 是标准布朗运动. 对于固定的 T, 求 $\cos B_T$ 在命题 5.2.2 下的表示.

解: 由 Itô 公式

$$\begin{aligned} e^{\frac{t}{2}}\cos B_t =& 1+\frac{1}{2}\int_0^t e^{\frac{s}{2}}\cos B_s\,\mathrm{d}s-\int_0^t e^{\frac{s}{2}}\sin B_s\,\mathrm{d}B_s \\ &-\frac{1}{2}\int_0^t e^{\frac{s}{2}}\cos B_s\,\mathrm{d}s. \end{aligned}$$

故

$$e^{\frac{t}{2}}\cos B_t = 1-\int_0^t e^{\frac{s}{2}}\sin B_s\,\mathrm{d}B_s,$$

于是

$$\begin{aligned} \cos B_T =& e^{-\frac{T}{2}}-\int_0^T e^{\frac{s-T}{2}}\sin B_s\,\mathrm{d}B_s \\ =& \mathsf{E}\cos B_T-\int_0^T e^{\frac{s-T}{2}}\sin B_s\,\mathrm{d}B_s. \end{aligned}$$

■

5.3　测度变换

首先从一个简单问题开始讨论. 当随机变量 X 在概率空间 $(\Omega,\mathcal{F},\mathsf{P})$ 上服从正态分布时, 我们设定 $X\sim N(\mu,\sigma^2)$, 我们可以变换测度，使得在新的概率空间上，X 的数学期望发生改变.

对于任意的 $A\in\mathcal{F}$, 考虑

$$\mathsf{Q}(A)=\mathsf{E}_{\mathsf{P}}[\frac{e^{aX}}{\mathsf{E}_{\mathsf{P}}[e^{aX}]}1_A].$$

通过计算可知, 在 $(\Omega,\mathcal{F},\mathsf{Q})$ 上, X 的数学期望为

$$\mathsf{E}_{\mathsf{Q}}[X]=\int_{-\infty}^{\infty}x\frac{e^{ax}}{e^{\sigma^2a^2/2+\mu a}}\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}dx=\int_{-\infty}^{\infty}x\frac{e^{-\frac{(x-\mu-a\sigma^2)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma}dx.$$

取 $a=-\frac{\mu}{\sigma^2}$, 在 $(\Omega,\mathcal{F},\mathsf{Q})$ 上, X 的数学期望为 0.

统计学中, 如果需要估计一个小概率事件, 我们往往需要用计算机去进行模拟验证. 若 $Y\sim N(6,1)$, 我们需要估计 $\mathsf{P}(Y<0)$. 事实上, $\mathsf{P}(Y<0)\approx 10^{-10}$ 是非常小的. 我们利用经验分布进行作模拟验证:

$$\mathsf{P}(Y<0)\approx\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{\{x_i<0\}},$$

这里 x_i 取自 $N(6,1)$ 总体的样本值. 由于 $\mathsf{P}(Y<0)\approx 10^{-10}$ 很小, 即使作 $n=10^6$ 次试验, 很有可能不能较准确估计出 $\mathsf{P}(Y<0)$.

若使用测度变换, 这个问题可以这样考虑. 设 P 是 $\mathbb{R}$ 上的概率测度. 对于 $A\in\mathcal{B}$,

$$\mathsf{P}(A)=\int_A \varphi(x)\,\mathrm{d}x,$$

这里 $\varphi(x)$ 为 $N(6,1)$ 的密度函数. 令 $\Lambda(x)=e^{6x-18}$. P 是 $\mathbb{R}$ 上的概率测度: 对 $A\in\mathcal{B}$,

$$\mathsf{Q}(A)=\int_A f(x)\,\mathrm{d}x,$$

这里 $f(x)$ 为 $N(0,1)$ 的密度函数. 事实上, $\mathsf{P}(A)=\int_A \Lambda(x)f(x)\,\mathrm{d}x$.

$$\begin{aligned}\mathsf{P}(Y<0)&\approx\frac{1}{n}\sum_{i=1}^{n}e^{6x_i-18}\mathbb{1}_{\{x_i<0\}}\\&=\frac{e^{-12}}{n}\sum_{i=1}^{n}e^{6(x_i-1)}\mathbb{1}_{\{x_i<0\}},\end{aligned}$$

这里 x_i 是取自总体 $N(0,1)$ 的样本值. 这样, 不需要进行大量试验，即可较准确估计出 $\mathsf{P}(Y<0)$.

当然上面的问题本质上是不同概率分布之间关系的问题. 在考虑随机过程的问题时, 我们经常会遇到一些随机过程，不是鞅过程. 例如, 若 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的标准布朗运动, 几何布朗运动

$$\mathrm{d}X_t=mX_t\,\mathrm{d}t+\sigma X_t\,\mathrm{d}B_t$$

是一个半鞅. 我们感兴趣的问题是：可否找到一个测度 Q, 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{Q})$ 上 X 是一个局部鞅. 这一节的测度变换即可解决这一问题.

为叙述方便, 这一节中我们考虑在时间上加以限制. 给定常数 $T>0$, 考虑在 $[0,T]$ 上的问题. 我们首先从一个布朗运动开始考虑.

若在概率空间 $(\Omega,\mathcal{F},\mathsf{P})$ 上, B 是标准布朗运动. 在 $[0,T]$ 上考虑

$$\widetilde{B}_t=B_t+\mu t$$

这里, $\mu\neq 0$. 希望构造一个新的测度空间，在新的测度空间上，$\widetilde{B}$ 在 $[0,T]$ 上是标准布朗运动. 由上面的启发, 对于任意的 $A\in\mathcal{F}$, 考虑

$$\mathsf{Q}(A)=\mathsf{E}_{\mathsf{P}}[\frac{e^{a\widetilde{B}_T}}{\mathsf{E}_{\mathsf{P}}[e^{a\widetilde{B}_T}]}1_A].$$

现在求 a. 由上面的讨论，当 $a=-\mu$ 时, $\widetilde{B}_T$ 的期望才为 0. 因此

$$\mathsf{Q}(A)=\mathsf{E}_{\mathsf{P}}[\exp\{-\mu\widetilde{B}_T+\frac{\mu^2}{2}T\}1_A].$$

接下来, 我们讨论在 $(\Omega,\mathcal{F},\mathsf{Q})$, 且有限区间 $[0,T]$ 上, $\widetilde{B}$ 是标准布朗运动. 这里, 需要注意的是, 我们只在有限区间上考虑此问题, 无穷时间区间上的情况会复杂些.

令 $0=t_0<t_1<\cdots<t_n=T$, H_i 是 Borel 集, $i=1,\cdots,n$. 令

$$A=\{\widetilde{B}_{t_1}\in H_1,\widetilde{B}_{t_2}-\widetilde{B}_{t_1}\in H_2,\cdots,\widetilde{B}_{t_n}-\widetilde{B}_{t_{n-1}}\in H_n\}.$$

于是

$$\begin{aligned}\mathsf{Q}(A) &= \mathsf{E}_{\mathsf{P}}[\exp\{-\mu\widetilde{B}_T+\frac{\mu^2}{2}T\}1_A]\\ &= \mathsf{E}_{\mathsf{P}}[\exp\{-\mu\sum_{i=1}^n(\widetilde{B}_{t_i}-\widetilde{B}_{t_{i-1}})+\frac{\mu^2}{2}\sum_{i=1}^n(t_i-t_{i-1})\}1_A]\\ &= \int_{H_1}\cdots\int_{H_n}e^{-\sum_{i=1}^n(\mu y_i-\frac{\mu^2}{2}(t_i-t_{i-1}))}\prod_{i=1}^n\frac{e^{-\frac{(y_i-\mu(t_i-t_{i-1}))^2}{2(t_i-t_{i-1})}}}{\sqrt{2\pi(t_i-t_{i-1})}}dy_1\cdots dy_n\\ &= \int_{H_1}\cdots\int_{H_n}\prod_{i=1}^n\frac{e^{-\frac{y_i^2}{2(t_i-t_{i-1})}}}{\sqrt{2\pi(t_i-t_{i-1})}}dy_1\cdots dy_n.\end{aligned}$$

上式说明了 $\widetilde{B}$ 在 $(\Omega,\mathcal{F},\mathsf{Q})$, 且有限区间 $[0,T]$ 上具有独立平稳增量性质, 而且是高斯过程. 容易得到 $\widetilde{B}$ 是标准布朗运动.

如果我们利用鞅的性质进行讨论，会得到更一般的结果. 我们下面从鞅的角度出发考虑此问题. 带流概率空间 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上, $\Lambda=(\Lambda_t)_{0\leqslant t\leqslant T}$ 是一个鞅, 且严格大于 0. 且 $\mathsf{E}[\Lambda_T]=1$ 时, 定义 Q 如下: 对于 $A\in\mathcal{F}$,

$$\mathsf{Q}(A)=\int_A\Lambda_T\,\mathrm{d}\mathsf{P}.$$

若 ξ 是在 Q 意义下可积的随机变量, 易知有 $\mathsf{E}_{\mathsf{Q}}[\xi]=\mathsf{E}_{\mathsf{P}}[\Lambda_T\xi]$. 当 $0\leqslant t\leqslant T$ 时, 我们有

$$\mathsf{E}_{\mathsf{Q}}[\xi\mid\mathcal{F}_t]=\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_t}\xi\mid\mathcal{F}_t].$$

下面, 我们证明上式. 由条件数学期望的性质, 对于任意有界 $\mathcal{F}_t-$ 可测随机变量 η,

$$\mathsf{E}_{\mathsf{Q}}[\xi\eta]=\mathsf{E}_{\mathsf{Q}}[\eta\ \mathsf{E}_{\mathsf{Q}}[\xi\mid\mathcal{F}_t]].$$

事实上, 由于 $\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_t}\xi \mid \mathcal{F}_t]$ 关于 $\mathcal{F}_t-$ 可测,

$$\begin{aligned}\mathsf{E}_{\mathsf{Q}}[\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_t}\xi \mid \mathcal{F}_t]\eta] &= \mathsf{E}_{\mathsf{P}}[\Lambda_T\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_t}\xi \mid \mathcal{F}_t]\eta] \\ &= \mathsf{E}_{\mathsf{P}}[\mathsf{E}_{\mathsf{P}}[\Lambda_T \mid \mathcal{F}_t]\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_t}\xi \mid \mathcal{F}_t]\eta] \\ &= \mathsf{E}_{\mathsf{P}}[\mathsf{E}_{\mathsf{P}}[\Lambda_T\xi \mid \mathcal{F}_t]\eta] \\ &= \mathsf{E}_{\mathsf{P}}[\Lambda_T\xi\eta] = \mathsf{E}_{\mathsf{Q}}[\xi\eta].\end{aligned}$$

因此，我们得到

$$\mathsf{E}_{\mathsf{Q}}[\xi \mid \mathcal{F}_t] = \mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_t}\xi \mid \mathcal{F}_t].$$

进一步, 若 $s \leqslant t$, ξ 是 $\mathcal{F}_t-$ 可测的，我们有

$$\mathsf{E}_{\mathsf{Q}}[\xi \mid \mathcal{F}_s] = \mathsf{E}_{\mathsf{P}}[\frac{\Lambda_t}{\Lambda_s}\xi \mid \mathcal{F}_s].$$

事实上,

$$\begin{aligned}\mathsf{E}_{\mathsf{Q}}[\xi \mid \mathcal{F}_s] &= \mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_s}\xi \mid \mathcal{F}_s] = \mathsf{E}_{\mathsf{P}}[\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_T}{\Lambda_s}\xi \mid \mathcal{F}_t] \mid \mathcal{F}_s] \\ &= \mathsf{E}_{\mathsf{P}}[\frac{\xi}{\Lambda_s}\mathsf{E}_{\mathsf{P}}[\Lambda_T \mid \mathcal{F}_t]\mathcal{F}_s] = \mathsf{E}_{\mathsf{P}}[\frac{\Lambda_t}{\Lambda_s}\xi \mid \mathcal{F}_s].\end{aligned}$$

由上面的讨论, 我们得到如下命题.

命题 5.3.1 $M = (M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{Q})$ 上的鞅, 当且仅当 $(\Lambda_t M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的鞅.

证明: 首先，若 $(\Lambda_t M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的鞅，因此

$$\begin{aligned}\mathsf{E}_{\mathsf{Q}}[M_t \mid \mathcal{F}_s] &= \mathsf{E}_{\mathsf{P}}[\frac{\Lambda_t}{\Lambda_s}M_t \mid \mathcal{F}_s] = \frac{1}{\Lambda_s}\mathsf{E}_{\mathsf{P}}[\Lambda_t M_t \mid \mathcal{F}_s] \\ &= \frac{1}{\Lambda_s}\Lambda_s M_s = M_s.\end{aligned}$$

故 $M = (M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{Q})$ 上的鞅.

反之, 若 $M = (M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{Q})$ 上的鞅,

$$\begin{aligned}\mathsf{E}_{\mathsf{P}}[\Lambda_t M_t \mid \mathcal{F}_s] &= \Lambda_s\mathsf{E}_{\mathsf{P}}[\frac{\Lambda_t}{\Lambda_s}M_t \mid \mathcal{F}_s] \\ &= \Lambda_s\mathsf{E}_{\mathsf{Q}}[M_t \mid \mathcal{F}_s] = \Lambda_s M_s.\end{aligned}$$

命题得证. ■

进一步, 有

命题 5.3.2　设 $\Lambda=(\Lambda_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅且 $\Lambda>0$, $\mathsf{E}_{\mathsf{P}}[\Lambda_T]=1$. 令

$$\mathsf{Q}(A)=\int_A \Lambda_T\,\mathrm{d}\mathsf{P},\ \ A\in\mathcal{F},$$

则 $\frac{1}{\Lambda_t}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{Q})$ 上的鞅.

结合前面的讨论, 我们总结出如下的结论，也是本节列出的第一个测度变换定理.

定理 5.3.1　(Cameron-Martin-Girsanov 定理 (I)) 设 B 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上的标准布朗运动. 令 $0\leqslant t\leqslant T$, μ 是常数, $W_t=B_t+\mu t$. 定义 Q 如下:

$$\mathsf{Q}(A)=\int_A e^{-\mu B_T-\frac{1}{2}\mu^2 T}\,\mathrm{d}\mathsf{P},$$

则在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{Q})$ 上, $W=(W_t)_{0\leqslant t\leqslant T}$ 是标准布朗运动.

我们可以注意到, 在做测度变换时，构造指数鞅十分关键. 下面的定理称为 Novikov 定理, 给出了判断指数鞅为鞅的准则.

定理 5.3.2　令 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上, $B=(B_t)_{0\leqslant t\leqslant T}$ 是标准布朗运动. $h=(h_t)_{0\leqslant t\leqslant T}$ 是循序可测过程, 使得 $X=h\cdot B=(X_t)_{0\leqslant t\leqslant T}$, $X_t=\int_0^t h_s\,\mathrm{d}B_s$ 可以被定义. 记 $\mathcal{E}(X)$ 为 X 的 Doléans-Dade 指数:

$$\mathcal{E}(X)_t=\exp\{\int_0^t h_s\,\mathrm{d}B_s-\frac{1}{2}\int_0^t h_s^2\,\mathrm{d}s\}.$$

若

$$\mathsf{E}[\exp\{\frac{1}{2}\int_0^T h_s^2 ds\}]<\infty,$$

则 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{0\leqslant t\leqslant T}$ 是鞅.

证明: 由于

$$\mathrm{d}\mathcal{E}(X)_t=\mathcal{E}(X)_t\,\mathrm{d}X_t=h_t\mathcal{E}(X)_t\,\mathrm{d}B_t,$$

因此 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{t\geqslant 0}$ 是非负局部鞅. 同时, 可以注意到 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{t\geqslant 0}$ 是非负上鞅. 事实上，由于 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{t\geqslant 0}$ 是局部鞅, 存在停时 $T_n\to\infty$, 使得 $(\mathcal{E}(X)_{t\wedge T_n})_{t\geqslant 0}$ 是一致可积鞅. 因此对于 $t>s$,

$$\mathsf{E}[\mathcal{E}(X)_{t\wedge T_n}|\mathcal{F}_s]=\mathcal{E}(X)_{s\wedge T_n}.$$

利用 Fatou 引理,

$$\begin{aligned}\mathsf{E}[\mathcal{E}(X)_t|\mathcal{F}_s] &= \mathsf{E}[\lim_{n\to\infty}\mathcal{E}(X)_{t\wedge T_n}|\mathcal{F}_s]\\ &\leqslant \lim_{n\to\infty}\mathsf{E}[\mathcal{E}(X)_{t\wedge T_n}|\mathcal{F}_s]=\lim_{n\to\infty}\mathcal{E}(X)_{s\wedge T_n}\\ &= \mathcal{E}(X)_s.\end{aligned}$$

因此 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{t\geqslant 0}$ 是非负上鞅. 令 $0<a<1$,

$$\begin{aligned}\mathcal{E}(ah\cdot B)_t &= \exp\{\int_0^t ah_s\,\mathrm{d}B_s-\frac{a^2}{2}\int_0^t h_s^2\,\mathrm{d}s\}\\ &= \mathcal{E}(h\cdot B)_t^a\exp\{\frac{a(1-a)}{2}\int_0^t h_s^2\,\mathrm{d}s\}.\end{aligned}$$

对于任意停时 $\tau\leqslant T$, $A\in\mathcal{F}_\tau$, 由于 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{t\geqslant 0}$ 是非负上鞅,

$$\begin{aligned}\mathsf{E}[\mathcal{E}(ah\cdot B)_\tau 1_A] &\leqslant (\mathsf{E}[\mathcal{E}(h\cdot B)_\tau])^a(\mathsf{E}[\exp\{\frac{a}{2}\int_0^\tau h_s^2\,\mathrm{d}s\}1_A])^{1-a}\\ &\leqslant (\mathsf{E}[\exp\{\frac{1}{2}\int_0^\tau h_s^2\,\mathrm{d}s\}1_A])^{1-a}.\end{aligned}$$

因此 $(\mathcal{E}(ah\cdot B)_t)_{0\leqslant t\leqslant T}$ 是一致可积的. 在不等式 $\mathsf{E}[\mathcal{E}(ah\cdot B)_t 1_A]\leqslant(\mathsf{E}[\mathcal{E}(h\cdot B)_t])^a(\mathsf{E}[\exp\{\frac{a}{2}\int_0^t h_s^2\,\mathrm{d}s\}1_A])^{1-a}$ 中令 a 趋向于 1, A 取全集, $\tau=T$, 有 $\mathsf{E}[\mathcal{E}(h\cdot B)_T]=1$. 进而可以得到 $\mathcal{E}(X)=(\mathcal{E}(X)_t)_{0\leqslant t\leqslant T}$ 是鞅.

■

根据上述结论，给出本节第二个结论.

定理 5.3.3 (Cameron-Martin-Girsanov **定理** (II)) 令 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上, $B=(B_t)_{0\leqslant t\leqslant T}$ 是标准布朗运动. $\mathbb{H}=(\mathbb{H}_t)_{0\leqslant t\leqslant T}$ 是循序可测过程, 使 $X=(X_t)_{0\leqslant t\leqslant T}$, $X_t=-\int_0^t\mathbb{H}_s\,\mathrm{d}B_s$ 可以被定义, 且满足 Novikov 定理的条件. 记 $\mathcal{E}(X)$ 为 X 的随机指数:

$$\mathcal{E}(X)_T=\exp\{-\int_0^T\mathbb{H}_s\,\mathrm{d}B_s-\frac{1}{2}\int_0^T\mathbb{H}_s^2\,\mathrm{d}s\}.$$

定义 Q 如下, 对于 $A\in\mathcal{F}$,

$$\mathsf{Q}(A)=\int_A\mathcal{E}(X)_T\,\mathrm{d}\mathsf{P}.$$

则 $W=(W_t)_{0\leqslant t\leqslant T}$, $W_t=B_t+\int_0^t\mathbb{H}_s\,\mathrm{d}s$ 是 Q 下的标准布朗运动.

证明: 首先注意到，若 Q 关于 P 绝对连续, 则在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上连续过程 W 的二次变差过程与 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{Q})$ 上 W 的二次变差过程相等. 如需证明 W 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{Q})$ 上是标准布朗运动, 只需证明 W 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{Q})$ 上是鞅. 由前讨论, 只需证 $\mathcal{E}(X)_tW_t$ 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上是鞅. 对于本定理的所叙述的问题, 考虑局部化技巧，只需要说明 $\mathcal{E}(X)_tW_t$ 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{0\leqslant t\leqslant T},\mathsf{P})$ 上是局部鞅. 注意到

$$\mathrm{d}\mathcal{E}(X)_t=\mathcal{E}(X)_t\,\mathrm{d}X_t=-\mathbb{H}_t\mathcal{E}(X)_t\,\mathrm{d}B_t,$$

故由分部积分公式,

$$\begin{aligned}\mathcal{E}(X)_t W_t = &\mathcal{E}(X)_0 W_0 + \int_0^t \mathcal{E}(X)_s \,\mathrm{d}W_s \\ &+ \int_0^t W_s \,\mathrm{d}\mathcal{E}(X)_s + < W, \mathcal{E}(X) >_t \\ = &\mathcal{E}(X)_0 W_0 + \int_0^t \mathcal{E}(X)_s \,\mathrm{d}B_s + \int_0^t \mathcal{E}(X)_s \mathbb{H}_s \,\mathrm{d}s \\ &- \int_0^t \mathbb{H}_s \mathcal{E}(X)_s W_s \,\mathrm{d}B_s - \int_0^t \mathcal{E}(X)_s \mathbb{H}_s \,\mathrm{d}s.\end{aligned}$$

故 $\mathcal{E}(X)W$ 在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{P})$ 上是局部鞅. ■

例 5.3.1 设 $B = (B_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{P})$ 上的标准布朗运动令

$$X_t = \int_0^t b_s \,\mathrm{d}s + \int_0^t h_s \,\mathrm{d}B_s,$$

这里 $h_s > 0$, 对任意 $0 \leqslant s \leqslant T$ 成立. 构造测度 Q, 使 $X = (X_t)_{0\leqslant t\leqslant T}$ 在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{Q})$ 上是局部鞅.

解: 定义 $W = (W_t)_{0\leqslant t\leqslant T}$, $W_t = B_t + \int_0^t \frac{b_s}{h_s}\,\mathrm{d}s$. 下面定义 Q, 使在 Q 下,$W = (W_t)_{0\leqslant t\leqslant T}$ 是标准布朗运动. 由测度变换定理, 构造

$$Y_t = -\int_0^t \frac{b_s}{h_s}\,\mathrm{d}B_s,$$

$$\mathcal{E}(Y)_T = \exp\{-\int_0^T \frac{b_s}{h_s}\,\mathrm{d}B_s - \frac{1}{2}\int_0^T \frac{b_s^2}{h_s^2}\,\mathrm{d}s\},$$

令

$$\mathsf{Q}(A) = \int_A \mathcal{E}(Y)_T \mathrm{d}\mathsf{P},\ A \in \mathcal{F}.$$

即可.

■

关于连续的局部鞅, 我们有如下定理.

定理 5.3.4 (Cameron-Martin-Girsanov **定理** (III)) $(M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{P})$ 上的连续局部鞅, 设 $\Lambda = (\Lambda_t)_{0\leqslant t\leqslant T}$ 是连续正鞅, $\Lambda_0 = 1$. 令

$$\mathsf{Q}(A) = \int_A \Lambda_T \,\mathrm{d}\mathsf{P},\ A \in \mathcal{F}.$$

则在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{Q})$ 上,

$$X = (X_t)_{0\leqslant t\leqslant T},\ X_t = M_t - \int_0^t \frac{1}{\Lambda_s}\,\mathrm{d}\langle M, \Lambda\rangle_s$$

是连续局部鞅.

证明: 类似前面的证明, 我们只需证明 ΛX 在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{P})$ 上是局部鞅. 由分部积分公式

$$\begin{aligned}\Lambda_t X_t &= \Lambda_0 X_0 + \int_0^t \Lambda_s \,\mathrm{d}X_s + \int_0^t X_s \,\mathrm{d}\Lambda_s + \langle X, \Lambda\rangle_t \\ &= \Lambda_0 X_0 + \int_0^t \Lambda_s \,\mathrm{d}M_s - \int_0^t \mathrm{d}\langle M, \Lambda\rangle_s + \int_0^t X_s \,\mathrm{d}\Lambda_s + \langle X, \Lambda\rangle_t \\ &= \Lambda_0 X_0 + \int_0^t \Lambda_s \,\mathrm{d}M_s + \int_0^t X_s \,\mathrm{d}\Lambda_s.\end{aligned}$$

故 ΛX 在 P 下是局部鞅. ■

考虑定理 5.3.4 中的正鞅 Λ, 对 $\log\Lambda$ 使用 Itô 公式.

$$\log\Lambda_t = \log\Lambda_0 + \int_0^t \frac{1}{\Lambda_s}\,\mathrm{d}\Lambda_s - \frac{1}{2}\int_0^t \frac{1}{\Lambda_s^2}\,\mathrm{d}\langle\Lambda, \Lambda\rangle_s.$$

令 $N_t = \int_0^t \frac{1}{\Lambda_s}\,\mathrm{d}\Lambda_s$, 故 $\Lambda = \mathcal{E}(N)$, 即 $\mathrm{d}\Lambda_t = \Lambda_t\,\mathrm{d}N_t$, 且 $\langle M, \Lambda\rangle = \Lambda\cdot\langle M, N\rangle$.

故定理 5.3.4 可改写如下:

定理 5.3.5 设 $M = (M_t)_{0\leqslant t\leqslant T}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{P})$ 上的连续局部鞅, $N = (N_t)_{0\leqslant t\leqslant T}$ 是初值为 0 的连续局部鞅. 令

$$\mathsf{Q}(A) = \int_A \mathcal{E}(N)_T\,\mathrm{d}\mathsf{P},$$

则在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{Q})$ 上, $X_t = M_t - \langle M, N\rangle_t$ 是连续局部鞅.

利用测度变换, 可以求解不少问题. 考虑布朗运动极值的分布问题. 设 $B = (B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上标准布朗运动. 令

$$X_t = B_t - bt,\ t\geqslant 0,\ b\in\mathbb{R}.$$

我们考虑 $\sup\limits_{0\leqslant s\leqslant t} X_s$ 的分布.

对于 $a\in\mathbb{R}$, 令

$$T_{a,b} = \begin{cases}\inf\{t\geqslant 0:\quad X_t\geqslant a\}, & a>0,\\ \inf\{t\geqslant 0:\quad X_t\leqslant a\}, & a<0.\end{cases}$$

显然对于 $a>0$, 有

$$\mathsf{P}(\sup_{0\leqslant s\leqslant t} X_s\geqslant a) = \mathsf{P}(T_{a,b}\leqslant t).$$

事实上 $T_{a,b}$ 是一个连续型随机变量，我们下面给出 $T_{a,b}$ 的密度函数.

由于 $X_t = B_t - bt$, 令 $A \in \mathcal{F}$,

$$\mathsf{Q}(A) = \int_A e^{bB_T - \frac{b^2 T}{2}}\,\mathrm{d}\mathsf{P}.$$

$T > t > 0$ 是常数. 在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{Q})$ 上，$X = (X_t)_{0\leqslant t\leqslant T}$ 是标准布朗运动. 故对任意可积的随机变量 W(关于 Q),

$$\mathsf{E}_{\mathsf{Q}}[W] = \mathsf{E}_{\mathsf{P}}[We^{bB_T - \frac{b^2 T}{2}}].$$

取

$$W = e^{-bB_T + \frac{b^2 T}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)},$$

因此

$$\mathsf{E}_{\mathsf{Q}}[e^{-bB_T + \frac{b^2 T}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}] = \mathsf{E}_{\mathsf{P}}[\mathbb{1}_{(T_{a,b}\leqslant t)}] = \mathsf{P}(T_{a,b} \leqslant t),$$

而

$$\mathsf{E}_{\mathsf{Q}}[e^{-bB_T + \frac{b^2 T}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}] = \mathsf{E}_{\mathsf{Q}}[e^{-bX_T - \frac{b^2 T}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}].$$

故

$$\mathsf{P}(T_{a,b} \leqslant t) = \mathsf{E}_{\mathsf{Q}}[e^{-bX_T - \frac{b^2 T}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}].$$

在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathsf{Q})$ 上 $\{e^{-bX_t - \frac{b^2 t}{2}}\}_{0\leqslant t\leqslant T}$ 是鞅. 因此

$$\mathsf{E}_{\mathsf{Q}}[e^{-bX_T - \frac{b^2 T}{2}} \mid \mathcal{F}_{t\wedge T_{a,b}}] = e^{-bX_{T_{a,b}\wedge t} - \frac{b^2 (t\wedge T_{a,b})}{2}}.$$

由于 $\{T_{a,b} \leqslant t\} \in \mathcal{F}_{t\wedge T_{a,b}}$, 故

$$\begin{aligned}
\mathsf{E}_{\mathsf{Q}}[e^{-bX_T - \frac{b^2 T}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}] &= \mathsf{E}_{\mathsf{Q}}[e^{-bX_{T_{a,b}\wedge t} - \frac{b^2 (t\wedge T_{a,b})}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}] \\
&= \mathsf{E}_{\mathsf{Q}}[e^{-bX_{T_{a,b}} - \frac{b^2 T_{a,b}}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}] \\
&= e^{-ab}\mathsf{E}_{\mathsf{Q}}[e^{-\frac{b^2 T_{a,b}}{2}} \mathbb{1}_{(T_{a,b}\leqslant t)}].
\end{aligned}$$

在测度 Q 下, $T_{a,b}$ 实际上是标准布朗运动的首中时.

$$\begin{aligned}
\mathsf{Q}(T_{a,b} \leqslant t) &= \mathsf{Q}(\max_{0\leqslant s\leqslant t} X_s \geqslant a) \\
&= 2\mathsf{Q}(X_t \geqslant a) \\
&= \sqrt{\frac{2}{\pi}} \int_a^\infty \frac{1}{t^{\frac{1}{2}}} e^{-\frac{x^2}{2t}}\,\mathrm{d}x.
\end{aligned}$$

对 t 求导, 有 $T_{a,b}$ 在 Q 下的密度为 $\frac{|a|}{\sqrt{2\pi}t^{\frac{3}{2}}}e^{-\frac{a^2}{2t}}$. 故

$$\begin{aligned}&e^{-ab}\mathsf{E}_{\mathsf{Q}}[e^{-\frac{b^2T_{a,b}}{2}}\mathbb{1}_{(T_{a,b}\leqslant t)}]\\=&e^{-ab}\int_0^t e^{-\frac{b^2s}{2}}\frac{|a|}{\sqrt{2\pi}s^{\frac{3}{2}}}e^{-\frac{a^2}{2s}}\,\mathrm{d}s\\=&\int_0^t\frac{|a|}{\sqrt{2\pi}s^{\frac{3}{2}}}e^{-\frac{(a+bs)^2}{2s}}\,\mathrm{d}s.\end{aligned}$$

对其关于 t 求导. 可知 $T_{a,b}$ 在测度 P 下的密度函数为 $\frac{|a|}{\sqrt{2\pi}t^{\frac{3}{2}}}e^{-\frac{(a+bt)^2}{2t}}$.

5.4 Black-Scholes 公式

我们这一节考虑使用 Itô 公式研究金融中的问题. 我们首先介绍 Feymann-Kac 公式. 假设 X 满足:

$$\mathrm{d}X_t=mX_t\,\mathrm{d}t+\sigma X_t\,\mathrm{d}B_t,$$

令

$$F(x)=(x-s)_+=\max\{x-s,0\}.$$

很多金融数学的问题是求解以下条件数学期望:

$$\phi(t,x)=\mathsf{E}[e^{-r(T-t)}F(X_T)\mid X_t=x].$$

直接去求这个条件数学期望是很困难的. 下面考虑使用 Itô 公式求解这个问题. 假设 $\phi(t,x)$ 是足够光滑的函数. 更一般地, 令 $\mathrm{d}X_t=m(t,X_t)\,\mathrm{d}t+\sigma(t,X_t)\,\mathrm{d}B_t$, 通货膨胀率为 $r(t,x)$. 则

$$\phi(t,x)=\mathsf{E}[\exp\{-\int_t^T r(s,X_s)\,\mathrm{d}s\}F(X_T)\mid X_t=x].$$

令 $R_T=R_0\exp\{\int_0^T r(s,X_s)\,\mathrm{d}s\}$, $M_t=\mathsf{E}[R_T^{-1}F(X_T)\mid\mathcal{F}_t]$, 于是 $M=(M_t)_{t\geqslant 0}$ 是鞅, 且

$$\begin{aligned}M_t=&R_t^{-1}\mathsf{E}[\exp\{-\int_t^T r(s,X_s)\,\mathrm{d}s\}F(X_T)\mid\mathcal{F}_t]\\=&R_t^{-1}\mathsf{E}[\exp\{-\int_t^T r(s,X_s)\,\mathrm{d}s\}F(X_T)\mid X_t]=R_t^{-1}\phi(t,X_t).\end{aligned}$$

由 Itô 公式

$$\mathrm{d}M_t = R_t^{-1}\,\mathrm{d}\phi(t,X_t) + \phi(t,X_t)\,\mathrm{d}R_t^{-1}.$$

这里

$$\begin{aligned}\mathrm{d}\phi(t,X_t) =&\partial_t\phi(t,X_t)\,\mathrm{d}t + \partial_x\phi(t,X_t)\,\mathrm{d}X_t + \frac{1}{2}\partial_{xx}\phi(t,X_t)\,\mathrm{d}\langle X,X\rangle_t\\ =&[\partial_t\phi(t,X_t) + \partial_x\phi(t,X_t)m(t,X_t) + \frac{1}{2}\partial_{xx}\phi(t,X_t)\sigma^2(t,X_t)]\,\mathrm{d}t\\ &+\sigma(t,X_t)\partial_x\phi(t,X_t)\,\mathrm{d}B_t,\\ \mathrm{d}R_t^{-1} =& -r(t,X_t)R_t^{-1}\,\mathrm{d}t.\end{aligned}$$

因此

$$\begin{aligned}\mathrm{d}M_t =&R_t^{-1}(-r(t,X_t)\phi(t,X_t) + \partial_t\phi(t,X_t)\\ &+\partial_x\phi(t,X_t)m(t,X_t) + \frac{1}{2}\partial_{xx}\phi(t,X_t)\sigma^2(t,X_t))\,\mathrm{d}t\\ &+\sigma(t,X_t)\partial_x\phi(t,X_t))\,\mathrm{d}B_t.\end{aligned}$$

由上讨论可知只有满足

$$-r(t,x)\phi(t,x) + \partial_t\phi(t,x) + m(t,x)\partial_x\phi(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\phi(t,x) = 0$$

$M = (M_t)_{t\geqslant 0}$ 才可能是鞅. 结合边界条件,

$$\phi(T,x) = F(x),$$

我们只要求解上面的偏微分方程即可. 这就是著名的 Feymann-Kac 公式.

下面, 我们讨论金融数学中期权定价的问题. 考虑欧式看涨期权的定价问题. 设股票价格 S 满足

$$\mathrm{d}S_t = S_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t],$$

这里 $m>0, \sigma\neq 0$, $S_0 = s_0$ 为初始价格. 现在考虑欧式看涨期权 (European call option). 此期权允许购买者在时刻 T 以价格 K 购买股票. 即购买者拥有在成交时刻 (strike time) 以成交价格 (stike price) K 买入此股票的权利. (购买者也可以在成交时间到来之际不购买此股票，购买者拥有最终是否购买的选择权). 现在金融数学问题是如何给这个期权定价.

设银行的无风险利率是 $r>0$, 因此在 t 时刻，若股票价格为 S_t, 则欧式期权的期望收益为

$$f(t,x) = \mathsf{E}[e^{-r(T-t)}F(S_T)|S_t = x],$$

这里

$$F(S_T) = (S_T - K)_+ = \begin{cases} S_T - K, & S_T > K, \\ 0, & S_T \leqslant K. \end{cases}$$

利用 Feymann-Kac 公式的求解方法，我们可以给出欧式期权的价格. 即可通过求解如下方程来获得定价方法:

$$-rf(t,x) + \partial_t f(t,x) + m\partial_x f(t,x) + \frac{1}{2}\sigma^2 \partial_{xx} f(t,x) = 0.$$

然而，如果以此计算方法为欧式期权定价，无法避免其他金融主体利用机会获得无风险的收益. 金融市场中的无风险收益往往被称为套利. 为了避免其他金融主体获得套利机会，因此需要通过风险调控，利用无风险存款和股票的持有量进行动态调整，从而达到与套期保值相同的效果. 套期保值又称对冲交易，原本指交易人在买进（或卖出）实际货物的同时，在期货交易所卖出（或买进）同等数量的期货交易合同作为保值。

Black-Scholes 的定价方法是将 $f(t,x)$ 看作无风险存款和股票的持有量的组合:

$$f(t,S_t) = V_t = a_t S_t + b_t R_t, \tag{5.1}$$

这里 $V = (V_t)_{0\leqslant t\leqslant T}$ 称为财富过程, R 为无风险存款的收益, 满足

$$\mathrm{d}R_t = rR_t\,\mathrm{d}t.$$

这里 a_t 与 b_t 是两个随机过程, 同时我们期望 $V_T = (S_T - K)_+$.

如果投资人在投资组合的运行途中不再投入更多的资金或者取出现金，投资组合的收益变换仅与股票与存款的收益变换有关，因此 V_t 可以写为:

$$\mathrm{d}V_t = a_t\,\mathrm{d}S_t + b_t\,\mathrm{d}R_t \tag{5.2}$$

我们把这种情况称之为自融资 (self-financing), 此时，结合考虑 (5.1) 和 (5.2) 有

$$\begin{aligned}\mathrm{d}V_t &= a_t S_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t] + b_t r R_t\,\mathrm{d}t \\ &= a_t S_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t] + r[V_t - a_t S_t]\,\mathrm{d}t \\ &= [m a_t S_t + r(V_t - a_t S_t)]\,\mathrm{d}t + \sigma a_t S_t\,\mathrm{d}B_t.\end{aligned}$$

同时, 利用 Itô 公式,

$$\begin{aligned}\mathrm{d}V_t &= \mathrm{d}f(t,S_t)\\ &= [\partial_t f(t,S_t) + mS_t\partial_x f(t,S_t) + \frac{\sigma^2 S_t^2}{2}\partial_{xx} f(t,S_t)]\,\mathrm{d}t\\ &\quad + \sigma S_t \partial_x f(t,S_t)\,\mathrm{d}B_t.\end{aligned}$$

以上两式相等, 可知:

$$a_t = \partial_x f(t,S_t).$$

由 (5.1),

$$b_t = \frac{V_t - a_t S_t}{R_t}.$$

又由 $V_t = f(t,S_t)$,

$$\begin{aligned}&\partial_t f(t,S_t) + mS_t a_t + \frac{\sigma^2 S_t^2}{2}\partial_{xx} f(t,S_t)\\ &= m a_t S_t + r f(t,S_t) - r\partial_x f(t,S_t) S_t.\end{aligned}$$

于是

$$\partial_t f(t,x) = r f(t,x) - r x \partial_x f(t,x) - \frac{\sigma^2 x^2}{2}\partial_{xx} f(t,x). \tag{5.3}$$

在这个式子中已经不再出现 m 了, 通过求解这个方程的来寻求定价的方法即为 Black-Scholes 的定价方法，(5.3) 被称为 Black-Scholes 方程.

此时, 我们也可以按照如下思路考虑这个问题：在原来的概率空间上, 股票价格 S 满足

$$\mathrm{d}S_t = S_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t],$$

我们通过测度变换，定义

$$Q(A) = \mathsf{E}_{\mathsf{P}}[\exp\{\frac{(r-m)}{\sigma}B_T - \frac{(r-m)^2 T}{2\sigma^2}\}1_A],$$

在新的测度空间 $(\Omega, \mathcal{F}, \mathsf{Q})$ 上, S 满足

$$\mathrm{d}S_t = S_t[r\,\mathrm{d}t + \sigma\,\mathrm{d}W_t].$$

也就是在 Q 下, 风险资产的期望收益率等于无风险收益率，因此 Q 也被称为风险中性测度.

因此，我们也可以按照如下公式理解期权定价问题.

$$
\begin{aligned}
f(t,x) &= \mathsf{E}_{\mathsf{Q}}[e^{-r(T-t)}F(S_T)|S_t = x] \\
\mathrm{d}S_t &= S_t[r\,\mathrm{d}t + \sigma\,\mathrm{d}W_t],
\end{aligned}
$$

W 是标准布朗运动. 利用分部积分公式, 我们可以观察到, 若考察 $\widetilde{S}_t = e^{-rt}S_t$, 会有

$$
\mathrm{d}\widetilde{S}_t = \sigma\widetilde{S}_t\,\mathrm{d}W_t.
$$

令 $\widetilde{V}_t = e^{-rt}V_t$, 同时注意到 $F(S_T) = (S_T - K)_+$, 我们可以看到

$$
\begin{aligned}
\widetilde{V}_t = e^{-rt}f(t,S_t) &= \mathsf{E}_{\mathsf{Q}}[e^{T}F(S_T)|\mathcal{F}_t] \\
&= \mathsf{E}_{\mathsf{Q}}[e^{T}F(S_T)|S_t].
\end{aligned}
$$

由几何布朗运动的性质,

$$
\begin{aligned}
\widetilde{S}_T &= \widetilde{S}_t \exp\{\int_t^T \sigma\,\mathrm{d}W_s - \frac{1}{2}\int_t^T \sigma^2\,\mathrm{d}s\} \\
&= \widetilde{S}_t \exp\{\sigma(W_T - W_t) - \sigma^2(T-t)/2\}.
\end{aligned}
$$

由布朗运动的独立增量性, 我们可以知道给定 $\widetilde{S}_t = h$ 的条件下，$\widetilde{S}_T$ 的分布函数是 $Z = \exp\{aN + y\}$ 的分布, 这里

$$
a = \sigma\sqrt{T-t},\ N \sim N(0,1),\ y = \log h - \frac{a^2}{2}.
$$

令

$$
\begin{aligned}
G(z) &= \mathsf{Q}(Z \leqslant z) \\
&= \mathsf{Q}(e^{aN+y} \leqslant z) \\
&= \mathsf{Q}(N \leqslant (\log z - y)/a),
\end{aligned}
$$

进一步,

$$
g(z) = G'(z) = \frac{1}{az}\phi(\frac{-y+\log z}{a}), \quad \phi\text{是标准正态分布的密度函数}.
$$

令 $\widetilde{K} = e^{-rT}K$, 由于 $\widetilde{V}_T = e^{-rT}(S_T - K)_+ = (\widetilde{S}_T - \widetilde{K})_+$, 再次利用布朗运动的独立增量性,

$$
\widetilde{V}_t = \mathsf{E}_{\mathsf{Q}}[(R_T^{-1}F(S_T)]|_{h=\widetilde{S}_t} = \int_{\widetilde{K}}^{\infty}(z-\widetilde{K})g(z)\,\mathrm{d}z|_{h=\widetilde{S}_t}.
$$

注意到

$$\int_{\widetilde{x}}^{\infty}(z-x)g(z)\,\mathrm{d}z=e^{y+\frac{a^2}{2}}\Phi(\frac{y-\log x+a^2}{a})-x\Phi(\frac{y-\log x}{a}),$$

$\Phi(x)$ 是标准正态分布的分布函数. 因此,

$$\widetilde{V}_t=\widetilde{S}_t\Phi(\frac{\log(\widetilde{S}_t/\widetilde{K})+\frac{a^2}{2}}{a})-\widetilde{K}\Phi(\frac{\log(\widetilde{S}_t/\widetilde{K})-\frac{a^2}{2}}{a}),$$

即

$$\begin{aligned}V_t=e^{rt}\widetilde{V}_t=S_t\Phi(&\frac{\log(S_t/K)+rs+\frac{a^2}{2}}{a})\\&-e^{-rs}K\Phi(\frac{\log(S_t/K)+rs-\frac{a^2}{2}}{a}),\end{aligned}$$

这里 $s=T-t$. 将 $a=\sigma\sqrt{T-t}$ 带入, 我们可知在 $S_t=x$ 的条件下,

$$f(t,x)=x\Phi(\frac{\log(x/K)+(r+\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}})-e^{-r(T-t)}K\Phi(\frac{\log(x/K)+(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}).$$

以上就是著名的 Black-Scholes 公式, 也被称为 Black-Scholes-Merton 公式.

5.5 随机微分方程与倒向随机微分方程

正倒向随机微分方程理论是随机分析中的重要内容之一, 它们在偏微分方程，随机控制，金融数学等领域有着广泛应用. 本节介绍 (正向) 随机微分方程和倒向随机微分方程的基本内容. 简单来说, 正向随机微分方程是将今天的确定性状态变为明天的一般是不确定的状态，以研究统计规律. 倒向随机微分方程的解是将明天的目标变为今天的确定状态, 以制定今天的决策.

设 $B=(B_t)_{t\geqslant 0}$ 是概率空间 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的 d-维标准布朗运动. 考虑时间区间 $[0,T]$ 上的随机微分方程:

$$\begin{cases}dX_t=b(t,X_t)dt+\sigma(t,X_t)dB_t,\\X_0=\eta\end{cases}$$

其中 η 为取值 $\mathbb{R}^n$ 中的 $\mathcal{F}_0$-可测随机变量, $b(t,x):[0,T]\times\Omega\times\mathbb{R}^n\to\mathbb{R}^n,\sigma(t,x):[0,T]\times\Omega\times\mathbb{R}^n\to\mathbb{R}^{n\times d}$ 为循序可测过程. 当 $b(t,x),\sigma(t,x)$ 为确定性函数时，称其为马氏随机微分方程. 由于布朗运动不可微，它的数学含义为

$$X_t=\eta+\int_0^t b(s,X_s)ds+\int_0^t\sigma(s,X_s)dB_s.\tag{5.4}$$

在工程技术和科学研究中，随机微分方程可用来描述不确定性环境中的系统演化行为.

为了便于读者理解, 我们首先讨论一类线性随机微分方程.

例 5.5.1 ([17]) 考虑线性随机微分方程, 其中 $n = d = 1$,

$$X_t = \eta + \int_0^t (b_s^1 X_s + b_s^0) ds + \int_0^t (\sigma_s^1 X_s + \sigma_s^0) dB_s.$$

为了求解这个方程, 需要引入辅助过程:

$$\Gamma_t = \exp\left(\int_0^t \beta_s dB_s + \int_0^t (\alpha_s - \frac{1}{2}\beta_s^2) ds \right),$$

可以验证:

$$d\Gamma_t = \alpha_t \Gamma_t dt + \beta_t \Gamma_t dB_t.$$

利用 Itô 公式,

$$d(X_t\Gamma_t) = \Gamma_t \big[[b_t^1 + \alpha_t + \beta_t \sigma_t^1] X_t + [b_t^0 + \beta_t \sigma_t^0] \big] dt + \Gamma_t \big[[\sigma_t^1 + \beta_t] X_t + \sigma_t^0 \big] dB_t.$$

取

$$\beta := -\sigma^1, \ \alpha := -b^1 + |\sigma^1|^2,$$

可得:

$$X_t = (\Gamma_t)^{-1} \left[\eta + \int_0^t \Gamma_s [b_s^0 - \sigma_s^0 \sigma_s^1] ds + \int_0^t \Gamma_s \sigma_s^0 dB_s \right].$$

下面, 介绍随机微分方程解的存在唯一性. 为方便叙述, 引入如下符号:

- $\mathcal{S}^2(\mathbb{R}^n)$ 表示 $[0,T]$ 上取值于 $\mathbb{R}^n$ 中满足

$$\mathsf{E}\left[\sup_{t \in [0,T]} |y_t|^2 \right] < \infty$$

  的所有的连续适应过程 y, $\mathcal{S}^2_{[t,s]}(\mathbb{R}^n)$ 表示 $[t,s]$ 上满足上式的所有的连续适应过程 y,

- $\mathcal{H}^2(\mathbb{R}^n)$ 表示 $[0,T]$ 上取值于 $\mathbb{R}^n$ 中满足

$$\mathsf{E}\left[\int_0^T |z_t|^2 dt \right] < \infty$$

  的所有的循序可测过程 z, $\mathcal{H}^2_{[t,s]}(\mathbb{R}^n)$ 表示 $[t,s]$ 上满足上式的所有的循序可测过程 z.

下面给出本节的主要结果.

定理 5.5.1　假设随机微分方程的初值与系数满足:

$$\mathsf{E}\left[|\eta|^2+\left(\int_0^T|b(t,0)|dt\right)^2\right]<\infty,\quad \mathsf{E}\left[\int_0^T|\sigma^2(t,0)|dt\right]<\infty,$$

及 Lipschitz 条件:

$$|b(t,x)-b(t,y)|+|\sigma(t,x)-\sigma(t,y)|\leqslant L|x-y|,\quad 0\leqslant t\leqslant T.$$

则随机微分方程 (5.4) 存在唯一的解 $X\in\mathcal{S}^2(\mathbb{R}^n)$.

注 5.5.1　若随机微分方程(5.4)存在解 $X\in\mathcal{H}^2(\mathbb{R}^n)$, 则 $X\in\mathcal{S}^2(\mathbb{R}^n)$.

事实上，注意到

$$X_t=\eta+\int_0^t b(s,X_s)ds+\int_0^t\sigma(s,X_s)dB_s.$$

因此

$$\begin{aligned}\mathsf{E}\left[\sup_{t\in[0,T]}|X_t|^2\right]\leqslant&3\mathsf{E}[\eta^2]+3\mathsf{E}\left[\left(\int_0^T|b(s,X_s)|ds\right)^2\right]\\&+3\mathsf{E}\left[\left(\sup_{t\in[0,T]}\left|\int_0^t\sigma(s,X_s)dB_s\right|\right)^2\right].\end{aligned}$$

然而

$$\mathsf{E}\left[\left(\int_0^T|b(s,X_s)|ds\right)^2\right]\leqslant\mathsf{E}\left[\left(\int_0^T(|b(s,0)|+L|X_s|)ds\right)^2\right],$$

$$\mathsf{E}\left[\left(\sup_{t\in[0,T]}\left|\int_0^t\sigma(s,X_s)dB_s\right|\right)^2\right]\leqslant 4\mathsf{E}\left[\left|\int_0^T(|\sigma(s,0)|+L|X_s|)^2ds\right|\right].$$

所以, 若随机微分方程(5.4) 存在解 $X\in\mathcal{H}^2(\mathbb{R}^n)$, 则 $X\in\mathcal{S}^2(\mathbb{R}^n)$.

证明: 我们首先证明方程具有局部解. 任给 $0\leqslant t<T$, 考虑区间 $[t,t+\delta]$ 上的随机微分方程

$$X_s=\zeta+\int_t^s b(r,X_r)dr+\int_t^s\sigma(r,X_r)dB_r,\tag{5.5}$$

其中 $\zeta\in L^2(\mathcal{F}_t)$. 我们将利用不动点原理证明存在 $\delta>0$, 使得方程(5.5)在区间 $[t,t+\delta]$ 上有唯一解, 并且常数 δ 只依赖于 L.

任给 $U \in \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n)$, 构造映射

$$I(U) := X^U.$$

其中

$$X^U_s = \zeta + \int_t^s b(r, U_r)dr + \int_t^s \sigma(r, U_r)dB_r.$$

显然 $X^U \in \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n)$. 因此,

$$I(U) : \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n) \to \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n).$$

接下来, 我们证明映射 $U \mapsto I(U)$ 具有压缩性: 任给 $U^1, U^2 \in \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n)$, 存在常数 $\rho(\delta) < 1$，使得

$$\mathsf{E}[\sup_{s\in[t,t+\delta]} |X^1_s - X^2_s|^2]^{\frac{1}{2}} \leqslant \rho(\delta)\mathsf{E}[\sup_{s\in[t,t+\delta]} |U^1_s - U^2_s|^2]^{\frac{1}{2}}.$$

注意到 $\mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n)$ 是 Banach 空间, 从而可以使用不动点方法. 因此，如果上式成立则方程(5.5)的局部解存在.

我们往证上式成立. 由于

$$X^1_s - X^2_s = \int_t^s [b(r, U^1_r) - b(r, U^2_r)]dr + \int_t^s [\sigma(r, U^1_r) - \sigma(r, U^2_r)]dB_r.$$

因此

$$\begin{aligned}
\mathsf{E}\left[\sup_{s\in[t,t+\delta]} |X^1_s - X^2_s|^2\right] \leqslant& 2\mathsf{E}\left[\left(\int_t^{t+\delta} |b(s, U^1_s) - b(s, U^2_s)|ds\right)^2\right] \\
&+ 2\mathsf{E}\left[\left(\sup_{s\in[t,t+\delta]} \left|\int_t^s [\sigma(r, X^1_r) - \sigma(r, X^2_r)]dB_r\right|\right)^2\right].
\end{aligned}$$

注意到

$$\begin{aligned}
\mathsf{E}\left[\left(\int_t^{t+\delta} |b(s, U^1_s) - b(s, U^2_s)|ds\right)^2\right] \leqslant& \delta\mathsf{E}\left[\int_t^{t+\delta} L^2|U^1_s - U^2_s|^2 ds\right] \\
\leqslant& L^2\delta^2\mathsf{E}\left[\sup_{s\in[t,t+\delta]} |U^1_s - U^2_s|^2\right],
\end{aligned}$$

同时

$$\begin{aligned}
\mathsf{E}\left[\left(\sup_{s\in[t,t+\delta]} \left|\int_t^s [\sigma(r, U^1_r) - \sigma(r, U^2_r)]dB_r\right|\right)^2\right] \leqslant& 4\mathsf{E}\left[\left(\left|\int_t^{t+\delta} [\sigma(s, U^1_s) - \sigma(s, U^2_s)]dB_s\right|\right)^2\right] \\
=& 4\mathsf{E}\left[\int_t^{t+\delta} |\sigma(s, U^1_s) - \sigma(s, U^2_s)|^2 ds\right] \\
\leqslant& 4L^2\delta\mathsf{E}\left[\sup_{s\in[t,t+\delta]} |U^1_s - U^2_s|^2\right].
\end{aligned}$$

因此

$$\mathsf{E}[\sup_{s\in[t,t+\delta]}|X_s^1-X_s^2|^2]^{\frac{1}{2}}\leqslant\rho(\delta)\mathsf{E}[\sup_{s\in[t,t+\delta]}|U_s^1-U_s^2|^2]^{\frac{1}{2}},$$

这里

$$\rho(\delta)=(2L^2\delta^2+8L^2\delta)^{\frac{1}{2}}.$$

显然可以取足够小的 $\delta>0$ 使得 $\rho(\delta)<1$. 因此，随机微分方程(5.5)在区间 $[t,t+\delta]$ 上存在唯一的 $\mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n)$-解.

最后证明方程存在唯一的 $\mathcal{S}^2(\mathbb{R}^n)$-全局解. 分割区间 $[0,T]$ 使得

$$0<\delta<2\delta<\cdots<(n-1)\delta<T\leqslant n\delta.$$

利用上述局部解结果, 显然可以从前往后迭代, 求解区间 $[(i-1)\delta,i\delta]$ 上的随机微分方程:

$$X_s=X_{(i-1)\delta}+\int_{(i-1)\delta}^s b(r,X_r)dr+\int_{(i-1)\delta}^s\sigma(r,X_r)dB_r,$$

从而可以得到区间 $[0,T]$ 上的全局解. 由于全局解也是局部解，则由每一小区间上局部解的唯一性可得整体解的唯一性. ■

需要强调的是，马氏随机微分方程是一类重要的马氏过程. 特别地, 它与线性偏微分方程有着密切联系，这就是著名的 Feynman-Kac 公式.

定理 5.5.2　假设定理 5.5.1 中的条件成立且 $b(t,x),\sigma(t,x)$ 为连续函数. 若 $u(t,x)\in C_b^{1,2}(\mathbb{R}^+\times\mathbb{R}^n)$ 是如下线性偏微分方程的解

$$\begin{cases}\partial_t u(t,x)-b(t,x)\partial_x u(t,x)-\frac{1}{2}tr[\partial_{xx}^2u(t,x)\sigma(t,x)\sigma^\top(t,x)]=0,(t,x)\in\mathbb{R}^+\times\mathbb{R}^n,\\ u(0,x)=\varphi(x),\end{cases}$$

则

$$u(t,x)=\mathsf{E}[\varphi(X_t^x)],\ (t,x)\in\mathbb{R}^+\times\mathbb{R}^n,$$

其中 X^x 为初值取 $x\in\mathbb{R}^n$ 随机微分方程(5.4)的解.

证明: 应用 Itô 公式可得,

$$du(t-s,X_s^x)=\partial_x u(t-s,X_s^x)\sigma(s,X_s^x)dB_s,$$

则 $(u(t-s,X_s^x))_{0\leqslant s\leqslant t}$ 是一鞅. 因此,

$$u(t,x)=\mathsf{E}[u(0,X_t^x)]=\mathsf{E}[\varphi(X_t^x)].$$

■

接下来, 我们介绍倒向随机微分方程. 1991 年, Pardoux 和 Peng 建立了非线性倒向随机微分方程理论，在随机控制, 金融数学等领域产生了巨大影响.

假设 $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的 d-维标准布朗运动. 与正向情形不同, 这里 $\{\mathcal{F}_t\}_{t\geqslant 0}$ 是布朗运动的自然域流

$$\mathcal{F}_t=\sigma\{\mathcal{N},B_s,0\leqslant s\leqslant t\}.$$

$\mathcal{N}$ 是所有的 P-零测集及其子集. 这样, 保证了 $\{\mathcal{F}_t\}_{t\geqslant 0}$ 是右连续的.

考虑倒向随机微分方程

$$\begin{cases} dY_t=-g(t,Y_t,Z_t)dt+Z_tdB_t,\\ Y_T=\xi, \end{cases}$$

其中 ξ 为取值 $\mathbb{R}^n$ 中的 $\mathcal{F}_T$-可测随机变量, $g(t,y,z):[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}^n$ 为循序可测过程. 由于布朗运动不可微，它的数学含义为

$$Y_t=\xi+\int_t^T g(s,Y_s,Z_s)ds-\int_t^T Z_sdB_s. \tag{5.6}$$

下面的定理给出了倒向随机微分方程解的存在唯一性.

定理 5.5.3　假设倒向随机微分方程的初值与系数满足:

$$\mathsf{E}\left[\xi^2+\left(\int_0^T|g(t,0,0)|dt\right)^2\right]<\infty,$$

及 Lipschitz 条件,

$$|g(t,y^1,z^1)-g(t,y^2,z^2)|\leqslant L(|y^1-y^2|+|z^1-z^2|),\ \forall 0\leqslant t\leqslant T.$$

则倒向随机微分方程 (5.6) 存在唯一的解 $(Y,Z)\in\mathcal{S}^2(\mathbb{R}^n)\times\mathcal{H}^2(\mathbb{R}^{n\times d})$.

注 5.5.2　若倒向随机微分方程(5.6)存在解 $(Y,Z)\in\mathcal{H}^2(\mathbb{R}^n)\times\mathcal{H}^2(\mathbb{R}^{n\times d})$, 则 $Y\in\mathcal{S}^2(\mathbb{R}^n)$.

事实上,

$$Y_t=\xi+\int_t^T g(s,Y_s,Z_s)ds-\int_t^T Z_sdB_s.$$

由于

$$\begin{aligned}\mathsf{E}\left[\sup_{t\in[0,T]}|Y_t|^2\right]\leqslant&3\mathsf{E}[\xi^2]+3\mathsf{E}\left[\left(\int_0^T|g(s,Y_s,Z_s)|ds\right)^2\right]\\&+3\mathsf{E}\left[\left(\sup_{t\in[0,T]}\left|\int_0^T Z_sdB_s-\int_0^t Z_sdB_s\right|\right)^2\right],\end{aligned}$$

同时注意到

$$\mathsf{E}\left[\left(\int_0^T |g(s,Y_s,Z_s)|ds\right)^2\right] \leqslant \mathsf{E}\left[\left(\int_0^T (|g(s,0,0)|+L|Y_s|+L|Z_s|)ds\right)^2\right],$$

$$\mathsf{E}\left[\left(\sup_{t\in[0,T]}\left|\int_0^t Z_s dB_s\right|\right)^2\right] \leqslant 4\mathsf{E}\left[\left(\left|\int_0^T Z_s dB_s\right|\right)^2\right] = 4\mathsf{E}\left[\int_0^T |Z_s|^2 ds\right],$$

因此 $Y \in \mathcal{S}^2(\mathbb{R}^n)$.

证明: 我们仍然采用不动点的方法证明结果. 首先考虑局部解, 任给 $0 \leqslant t < T$, 考虑区间 $[t,t+\delta]$ 上的倒向随机微分方程

$$Y_s = \zeta + \int_s^{t+\delta} g(r,Y_r,Z_r)dr - \int_s^{t+\delta} Z_r dB_r, \tag{5.7}$$

其中 $\zeta \in L^2(\mathcal{F}_{t+\delta})$. 我们将证明存在 $\delta > 0$ 使得方程(5.7)在区间 $[t,t+\delta]$ 上有唯一的 $\mathcal{S}^2(\mathbb{R}^n)\times\mathcal{H}^2(\mathbb{R}^{n\times d})$-解, 并且常数 δ 只依赖于 T,L.

任给 $(U,V)\in\mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n)\times\mathcal{H}^2_{[t,t+\delta]}(\mathbb{R}^{n\times d})$, 构造映射

$$I(U,V) := (Y,Z),$$

其中

$$Y_s = \zeta + \int_s^{t+\delta} g(r,U_r,V_r)dr - \int_s^{t+\delta} Z_r dB_r.$$

首先验证映射 I 良定义. 对

$$\zeta + \int_t^{t+\delta} g(r,U_r,V_r)dr \in L^2(\mathcal{F}_{t+\delta})$$

使用鞅表示定理可得,

$$\zeta + \int_t^{t+\delta} g(r,U_r,V_r)dr = \mathsf{E}\left[\zeta + \int_t^{t+\delta} g(r,U_r,U_r)dr\right] + \int_0^{t+\delta} Z_r dB_r.$$

其中 $Z \in \mathcal{H}^2_{[0,t+\delta]}(\mathbb{R}^{n\times d})$. 容易验证,

$$\zeta + \int_t^{t+\delta} g(r,U_r,V_r)dr = \mathsf{E}\left[\zeta + \int_t^{t+\delta} g(r,U_r,V_r)dr|\mathcal{F}_s\right] + \int_s^{t+\delta} Z_r dB_r.$$

则当 $t \leqslant s \leqslant t+\delta$ 时,

$$\zeta + \int_s^{t+\delta} g(r,U_r,V_r)dr - \int_s^{t+\delta} Z_r dB_r = \mathsf{E}\left[\zeta + \int_s^{t+\delta} g(r,U_r,V_r)dr|\mathcal{F}_s\right].$$

则有

$$\begin{aligned} Y_s &:= \mathsf{E}\Big[\zeta + \int_s^{t+\delta} g(r, U_r, V_r)dr | \mathcal{F}_s\Big] \\ &= \zeta + \int_s^{t+\delta} g(r, U_r, V_r)dr - \int_s^{t+\delta} Z_r dB_r \in \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n). \end{aligned}$$

接下来证明映射

$$I(U, V) : \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n) \times \mathcal{H}^2_{[t,t+\delta]}(\mathbb{R}^{n\times d}) \to \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n) \times \mathcal{H}^2_{[t,t+\delta]}(\mathbb{R}^{n\times d})$$

是压缩的. 任给 $(U^1, V^1), (U^2, V^2) \in \mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n) \times \mathcal{H}^2_{[t,t+\delta]}(\mathbb{R}^{n\times d})$, 需要证明存在常数 $\rho(\delta) < 1$ 使得

$$\begin{aligned} &\mathsf{E}\Big[\sup_{s\in[t,t+\delta]} |Y^1_s - Y^2_s|^2 + \int_t^{t+\delta} |Z^1_s - Z^2_s|^2 ds\Big]^{\frac{1}{2}} \\ \leqslant\ &\rho(\delta)\mathsf{E}\Big[\sup_{s\in[t,t+\delta]} |U^1_s - U^2_s|^2 + \int_t^{t+\delta} |V^1_s - V^2_s|^2 ds\Big]^{\frac{1}{2}}. \end{aligned}$$

由于 $\mathcal{S}^2_{[t,t+\delta]}(\mathbb{R}^n) \times \mathcal{H}^2_{[t,t+\delta]}(\mathbb{R}^{n\times d})$ 为一 Banach 空间，如果上式成立则局部解存在.

事实上,

$$Y^1_s - Y^2_s = \int_s^{t+\delta} [g(r, U^1_r, V^1_r) - g(r, U^2_r, V^2_r)]dr - \int_s^{t+\delta} [Z^1_r - Z^2_r]dB_r.$$

从而

$$Y^1_s - Y^2_s = \mathsf{E}\Big[\int_s^{t+\delta} [g(r, U^1_r, V^1_r) - g(r, U^2_r, V^2_r)]dr | \mathcal{F}_s\Big].$$

于是

$$\begin{aligned} \sup_{s\in[t,t+\delta]} |Y^1_s - Y^2_s| &\leqslant \sup_{s\in[t,t+\delta]} \mathsf{E}\Big[\int_s^{t+\delta} |g(r, U^1_r, V^1_r) - g(r, U^2_r, V^2_r)|dr | \mathcal{F}_s\Big] \\ &\leqslant \sup_{s\in[t,t+\delta]} \mathsf{E}\Big[\int_t^{t+\delta} |g(r, U^1_r, V^1_r) - g(r, U^2_r, V^2_r)|dr | \mathcal{F}_s\Big]. \end{aligned}$$

因此

$$\mathsf{E}\Big[\sup_{s\in[t,t+\delta]} |Y^1_s - Y^2_s|^2\Big] \leqslant 4\mathsf{E}\Big[\Big(\int_t^{t+\delta} |g(s, U^1_s, V^1_s) - g(s, U^2_s, V^2_s)|ds\Big)^2\Big].$$

利用 Lipschitz 条件可得

$$
\begin{aligned}
&\mathsf{E}\left[\left(\int_t^{t+\delta}|g(s,U_s^1,V_s^1)-g(s,U_s^2,V_s^2)|ds\right)^2\right]\\
\leqslant&\mathsf{E}\left[\left(\int_t^{t+\delta}(L|U_s^1-U_s^2|+L|V_s^1-V_s^2|)ds\right)^2\right]\\
\leqslant&2L^2\delta\mathsf{E}\left[\int_t^{t+\delta}(|U_s^1-U_s^2|^2+|V_s^1-V_s^2|^2)ds\right]\\
\leqslant&2L^2\delta\mathsf{E}\left[\delta\sup_{s\in[t,t+\delta]}|U_s^1-U_s^2|^2+\int_t^{t+\delta}|V_s^1-V_s^2|^2ds\right].
\end{aligned}
$$

因此

$$
\mathsf{E}\left[\sup_{s\in[t,t+\delta]}|Y_s^1-Y_s^2|^2\right]\leqslant 8L^2\delta\mathsf{E}\left[\delta\sup_{s\in[t,t+\delta]}|U_s^1-U_s^2|^2+\int_t^{t+\delta}|V_s^1-V_s^2|^2ds\right].
$$

注意到

$$
\int_t^{t+\delta}[g(s,U_s^1,V_s^1)-g(s,U_s^2,V_s^2)]ds-(Y_t^1-Y_t^2)=\int_t^{t+\delta}(Z_s^1-Z_s^2)dB_s.
$$

则由 Itô 等距公式,

$$
\begin{aligned}
&\mathsf{E}\left[\int_t^{t+\delta}|Z_s^1-Z_s^2|^2ds\right]=\mathsf{E}\left[\left|\int_t^{t+\delta}(Z_s^1-Z_s^2)dB_s\right|^2\right]\\
\leqslant&2\mathsf{E}\left[|Y_t^1-Y_t^2|^2+\left(\int_t^{t+\delta}|g(s,U_s^1,V_s^1)-g(s,U_s^2,V_s^2)|ds\right)^2\right]\\
\leqslant&20L^2\delta\mathsf{E}\left[\delta\sup_{s\in[t,t+\delta]}|U_s^1-U_s^2|^2+\int_t^{t+\delta}|V_s^1-V_s^2|^2ds\right].
\end{aligned}
$$

因此

$$
\begin{aligned}
&\mathsf{E}\left[\sup_{s\in[t,t+\delta]}|Y_s^1-Y_s^2|^2+\int_t^{t+\delta}|Z_s^1-Z_s^2|^2ds\right]^{\frac{1}{2}}\\
\leqslant& L(28(1+T)\delta)^{\frac{1}{2}}\mathsf{E}\left[\sup_{s\in[t,t+\delta]}|U_s^1-U_s^2|^2+\int_t^{t+\delta}|V_s^1-V_s^2|^2ds\right]^{\frac{1}{2}}.
\end{aligned}
$$

故

$$
\rho(\delta)=L(28(1+T)\delta)^{\frac{1}{2}}.
$$

于是局部解存在唯一性得证.

最后, 分割区间 $[0,T]$ 使得

$$0<\delta<2\delta<\cdots<(n-1)\delta<T\leqslant n\delta.$$

利用上述局部解结果, 显然可以从后往前迭代, 求解区间 $[(i-1)\delta,i\delta]$ 上倒向随机微分方程

$$Y_s=Y_{i\delta}+\int_s^{i\delta}g(r,Y_r,Z_r)dr-\int_s^{i\delta}Z_rdB_r$$

得到全局解的存在性. 类似正向情形，唯一性则由每一小区间上局部解的唯一性保证. ■

倒向随机微分方程和非线性偏微分方程存在密切关系,这就是非线性 Feynman-Kac 公式, 它给出了一大类半线性偏微分方程解的概率表示. 以下定理的详细解释可参考 Peng [8].

定理 5.5.4 假设定理 5.5.1 条件成立且 $b(t,x),\sigma(t,x)$ 为连续函数. 此外, $u(t,x)\in C_b^{1,2}([0,T]\times\mathbb{R}^n)$ 是如下非线性偏微分方程的解

$$\begin{cases}\partial_tu+b\partial_xu+\frac{1}{2}tr[\partial_{xx}^2u\sigma\sigma^\top]+f(t,x,u,\sigma^\top\partial_xu)=0,(t,x)\in[0,T]\times\mathbb{R}^n,\\ u(0,x)=\varphi(x),\end{cases}$$

其中连续函数 $f(t,x,y,z)$ 满足

$$|f(t,x^1,y^1,z^1)-f(t,x^2,y^2,z^2)|\leqslant L(|x^1-x^2|+|y^1-y^2|+|z^1-z^2|),\ \forall 0\leqslant t\leqslant T.$$

则

$$u(t,x)=Y_t^{t,x},\quad (t,x)\in[0,T]\times\mathbb{R}^n,$$

其中 $(X^{t,x},Y^{t,x},Z^{t,x})\in\mathcal{S}^2_{[t,T]}(\mathbb{R}^n)\times\mathcal{S}^2_{[t,T]}(\mathbb{R})\times\mathcal{H}^2_{[t,T]}(\mathbb{R}^d)$ 为区间 $[t,T]$ 上的正倒向随机微分方程系统

$$\begin{cases}X_s^{t,x}=x+\int_t^sb(r,X_r^{t,x})dr+\int_t^s\sigma(r,X_r^{t,x})dB_r,\\ Y_s^{t,x}=\varphi(X_T^{t,x})+\int_s^Tf(r,X_r^{t,s},Y_r^{t,x},Z_r^{t,x})dr-\int_s^TZ_r^{t,x}dB_r\end{cases}$$

的解.

证明: 应用 Itô 公式, 可得

$$\begin{aligned}u(s,X_s^{t,x})=&\varphi(X_T^{t,x})+\int_s^Tf(r,X_r^{t,x},u(r,X_r^{t,x}),\sigma^\top(r,X_r^{t,x})\partial_xu(r,X_r^{t,x}))dr\\&-\int_s^T\sigma^\top(r,X_r^{t,x})\partial_xu(r,X_r^{t,x})dB_r.\end{aligned}$$

由倒向随机微分方程解的存在唯一性，可得

$$(Y_s^{t,x}, Z_s^{t,x}) = \big(u(s, X_s^{t,x}), \sigma^\top(s, X_s^{t,x})\partial_x u(s, X_s^{t,x})\big).$$

取 $s=t$ 可得证明成立. ■

本节的最后, 简要介绍倒向随机微分方程在金融数学中的应用. 特别地, 我们将利用倒向随机微分方程理论回顾下欧式期权定价问题. 简单来说, 我们需要寻找一种满足自融资条件的投资组合

$$V_t = a_t S_t + b_t R_t,$$

使得 $V_T = (S_T - K)_+$. 那么 V_0 就是价格, 这里

$$\begin{cases} dS_t = S_t[mdt + \sigma dB_t], \\ dR_t = rR_t dt. \end{cases}$$

根据自融资假设，$\mathrm{d}V_t = a_t\,\mathrm{d}S_t + b_t\,\mathrm{d}R_t$. 因此,

$$\begin{aligned} \mathrm{d}V_t &= a_t S_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t] + b_t r R_t\,\mathrm{d}t \\ &= a_t S_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t] + r[V_t - a_t S_t]\,\mathrm{d}t \\ &= \left(rV_t + \frac{(m-r)}{\sigma}\sigma a_t S_t\right)\mathrm{d}t + \sigma a_t S_t\,\mathrm{d}B_t. \end{aligned}$$

则 $(V_t, \sigma a_t S_t)$ 为如下线性倒向随机微分方程的解

$$V_t = (S_T - K)^+ - \int_t^T \left(rV_s + \frac{(m-r)}{\sigma}Z_s\right)\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}B_s.$$

接下来，我们引入风险中性测度

$$Q(A) = \mathsf{E}_\mathsf{P}\left[\exp\{\frac{(r-m)}{\sigma}B_T - \frac{(r-m)^2 T}{2\sigma^2}\}1_A\right].$$

在测度空间 $(\Omega, \mathcal{F}, \mathsf{Q})$ 上, $B_t^\mathsf{Q} = B_t + \frac{(m-r)}{\sigma}t$ 为一布朗运动, 且上述倒向随机微分方程变为

$$V_t = (S_T - K)^+ - \int_t^T rV_s ds - \int_t^T Z_s\,\mathrm{d}B_s^\mathsf{Q}.$$

容易验证

$$V_0 = \mathsf{E}_\mathsf{Q}[e^{-rT}(S_T - K)^+)],$$

这与 Black-Scholes 公式是一致的.

事实上，倒向随机微分方程提供了一个有效的资产定价理论. 特别地, 它可用于研究各种不同情景下的定价问题. 譬如, 一般来说市场中的存贷款利率是不一样的. 假若存款利率为 r, 贷款利率为 $l > r$, 这意味着自融资投资组合 V 满足

$$\begin{aligned}\mathrm{d}V_t &= a_tS_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t] + (V_t - a_tS_t)^+ r\,\mathrm{d}t - (V_t - a_tS_t)^- l\,\mathrm{d}t \\ &= a_tS_t[m\,\mathrm{d}t + \sigma\,\mathrm{d}B_t] + r[V_t - a_tS_t]\,\mathrm{d}t - (l-r)[V_t - a_tS_t]^-\,\mathrm{d}t \\ &= \left(rV_t + \frac{(m-r)}{\sigma}\sigma a_tS_t - (l-r)(V_t - a_tS_t)^-\right)\mathrm{d}t + \sigma a_tS_t\,\mathrm{d}B_t.\end{aligned}$$

于是 $(V_t, \sigma a_tS_t)$ 为如下非线性倒向随机微分方程的解:

$$V_t = (S_T - K)_+ - \int_t^T \left(rV_s + \frac{(m-r)}{\sigma}Z_s - (l-r)(V_s - \frac{Z_s}{\sigma})^-\right)\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}B_s.$$

关于倒向随机微分方程理论的更多应用请参阅 Peng [9], Zhang [17].

5.6 习题

1. $X = (X_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的 O - U 过程, $X_0 = 1$,

$$dX_t = mX_tdt + dB_t,$$

这里 $B = (B_t)_{t\geqslant 0}$ 是标准布朗运动. $R > 1$ 是常数.

$$T = \inf\{t : X_t = R \text{ or } X_t = 0\}.$$

(1) 若存在一个函数 F, 满足 $F(0) = 0$, $x > 0$ 时 $F(x) > 0$, 且使 $F(X_{t\wedge T})$ 是 $\mathcal{F}_t-$ 局部鞅. 求 F;

(2) 求 $\mathsf{P}(X_T = 0)$;

(3) 请问 m 满足什么条件时会有

$$\lim_{R\to\infty} \mathsf{P}(X_T = R) = 0.$$

2. $X = (X_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上扩散过程, 满足

$$dX_t = -X_t^2dt + dB_t,$$

(1) 令 $T = \inf\{t : X_t = 0 \text{ 或 } X_t = 2\}$, $p(x) = \mathsf{P}(X_T = 0|X_0 = x)$, $0 < x < 2$, 设 $p(x)$ 是一个 C^2 函数, 请给出 $p(x)$ 满足的方程;

(2) 若 $M_t = X_t \exp\{-\int_0^t g(X_s)ds\}$, 求 $g(x)$ 使得 $M=(M_t)_{t\geqslant 0}$ 是 $\mathcal{F}_t-$ 局部鞅.

3. $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上随机过程, 满足

$$dM_t = M_t B_t dB_t,$$

这里 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动.

(1) 令 $T = \inf\{t : B_t = 2 \quad \text{or} \quad B_t = -1\}$, 假设 $P(T < +\infty) = 1$, 求 $\mathsf{P}(B_T = 2)$;

(2) 若

$$\frac{d\mathsf{Q}|_{\mathcal{F}_t}}{d\mathsf{P}|_{\mathcal{F}_t}} = M_t,$$

请给出 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{Q})$ 上 $B=(B_t)_{t\geqslant 0}$ 的形式;

(3) 假设 $\mathsf{Q}(T<+\infty)=1$, $\mathsf{Q}(B_T=2)$.

4. $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的布朗运动, $B_0 = x > 0$. 求概率测度 Q 使得在 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{Q})$ 上, $(B_t^3)_{t\geqslant 0}$, $(\cos B_t)_{t\geqslant 0}$ 是局部鞅.

5. B 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动, 定义

$$T = \inf\{t \geqslant 0 : B_t = -3 \text{ 或 } B_t = 4\}.$$

令 $f(t,x) = tx + g(x)$. 求 $g(x)$, 使得 $(f(t,B_t))_{t\geqslant 0}$ 有可能是鞅, 并利用此过程求得 $\mathsf{E}[TB_T]$.

6. 设 $B=(B_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathsf{P})$ 上的标准布朗运动, 请给出

$$\mathsf{E}[e^{-\sigma\int_t^T B_s ds} | B_t = x]$$

的计算方法.

7. $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的标准布朗运动. $b>0$. 对于 $t\in[0,T]$,

$$Z_t = \exp\{-b(B_t^2 - t) - 2b^2 \int_0^t B_s^2 ds\}.$$

令 $\{\mathcal{F}_t\}_{t\geqslant 0} = \mathcal{F}_t = \sigma\{\mathcal{N}, B_s, 0\leqslant s\leqslant t\}$, $\mathcal{N}$ 是所有的 P-零测集及其子集.

(1). 证明 $Z=(Z_t)_{t\in[0,T]}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathsf{P})$ 上的鞅.

(2). 令

$$\widetilde{B}_t = B_t + b\int_0^t B_s ds.$$

构造 $\widetilde{\mathsf{P}}$, 使得 $\widetilde{B}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\widetilde{\mathsf{P}})$ 上的布朗运动.

8. $B=(B_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\mathsf{P})$ 上的标准布朗运动. 对于 $t\in[0,T]$, $X_t=e^{-mB_t^2}$

$$M_t=X_t\exp\{\int_0^t g(X_s)ds\}.$$

令 $\{\mathcal{F}_t\}_{t\geqslant 0}=\mathcal{F}_t=\sigma\{\mathcal{N},B_s,0\leqslant s\leqslant t\}$, $\mathcal{N}$ 是所有的 P-零测集及其子集.

(1). 求 g 使得 $M=(M_t)_{t\geqslant 0}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的局部鞅.

(2). 代入上式的 g, 请证明 M 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的鞅. 令

$$\widetilde{\mathsf{P}}(A)=\mathsf{E}_{\mathsf{P}}[M_T 1_A].$$

求 B 在 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\widetilde{\mathsf{P}})$ 上的形式.

9. B 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的一维标准布朗运动，$B_0=0$. 求解如下倒向随机微分方程的解:

(1). $Y_t=\exp(\frac{T^2}{2})-\int_t^T(s|Y_s|+s\exp(\frac{s}{2})|Z_s|)ds-\int_t^T Z_s dB_s$;

(2). $Y_t=\xi-\int_t^T(|Y_s|+Z_s)ds-\int_t^T Z_s dB_s$, 其中 $\xi\geqslant 0$.

10. B 是 $(\Omega,\mathcal{F},\{(\mathcal{F}_t\}_{t\geqslant 0},\mathsf{P})$ 上的一维标准布朗运动，$B_0=0$. 对每一 $\xi\in L^2(\mathcal{F}_T)$, 定义 $\mathcal{E}[\xi]:=Y_0^\xi$, 其中

$$Y_t^\xi=\xi+\int_t^T|Z_s^\xi|ds-\int_t^T Z_s^\xi dB_s.$$

证明:

$$\mathcal{E}[\xi]=\sup_{\theta_s\text{为一适应过程且}|\theta_s|\leqslant 1}\mathsf{E}_{\mathsf{P}_\theta}[\xi],$$

其中

$$\frac{d\mathsf{P}_\theta}{d\mathsf{P}}:=\exp(\int_0^T\theta_s dB_s-\frac{1}{2}\int_0^T\theta_s^2 ds).$$

第 6 章　Lévy 过程初步

前面几章，我们讨论了连续鞅的性质, 以及和连续鞅相关的一系列过程. 在讨论的过程中, 轨道的连续性起了很重要的作用。在这一章中, 我们介绍轨道不连续的过程. 我们着眼于介绍一大类重要的随机过程 —— Lévy 过程的随机分析性质. 作为 Lévy 过程的重要例子, Possion 过程，复合 Possion 过程也有很多应用. 当然，由于篇幅有限, 我们不可能详细很深入地介绍 Lévy 过程的随机分析性质. 特别是与 Lévy 过程密切相关的半鞅的 Lévy 刻画. 感兴趣的读者可以参考 Jacod 和 Shiryaev [3].

6.1　Lévy 过程的定义

我们首先给出 Lévy 过程的定义.

定义 6.1.1　$(\Omega, \mathcal{F}, \mathsf{P})$ 是一个概率空间. $X = (X_t)_{t\geqslant 0}$ 是其上一个随机过程. 若对 $s, t > 0$, $X_{s+t} - X_s$ 与 $\sigma(X_r, r \leqslant s)$ 独立, 且 $X_{s+t} - X_s$ 与 $X_t - X_0$ 同分布, 称 X 为 **Lévy** 过程.

我们之前所学的布朗运动即是一个典型的 Lévy 过程. 下面再给出一些 Lévy 过程的例子.

例 6.1.1　设 $B = (B_t)_{t\geqslant 0}$ 是标准布朗运动. 设 $T_s = \inf\{t:\ B_t = s\}$. 由于布朗运动具有强马氏性, 故对于 $T = (T_s)_{s\geqslant 0}$, 具有独立平稳增量, 因此 T 是 Lévy 过程. 注意到这个 Lévy 过程并不是高斯过程. 前面利用反射原理, 我们可知 T_s 的密度函数为

$$f_s(x) = \frac{s}{x^{\frac{3}{2}}\sqrt{2\pi}} e^{-\frac{s^2}{2x}},\ 0 < x < +\infty,$$

由于 Lévy 过程仅仅是对过程的增量进行限制, 因此其很多性质并不是显然的.

命题 6.1.1　对于一个右连左极的 Lévy 过程 $X = (X_t)_{t\geqslant 0}$, 对于任意 t, $\mathsf{P}(\Delta X_t \neq 0) = 0$. 这里 $\Delta X_t = X_t - X_{t-}$, $X_{t-} = \lim\limits_{s\uparrow t} X_s$.

证明: 由于 $\Delta X_t = \lim\limits_{s\uparrow t}(X_t - X_s)$, 由于 Lévy 过程具有平稳增量性质, ΔX_t 的分布与 t 无关. 故 $\mathsf{P}(\Delta X_t \neq 0)$ 的值与 t 无关. 由于一个右连左极过程至多有

可列个间断点, 故只可能 $\mathsf{P}(\Delta X_t \neq 0) = 0$, 否则会产生矛盾. ■

正是由于 Lévy 过程的增量性质, 还可以有如下讨论.

设 $X = (X_t)_{t\geqslant 0}$ 是一个 Lévy 过程. 此时, $X_1 = \sum_{j=1}^{n} Y_{j,n}$, $Y_{j,n} = X_{\frac{j}{n}} - X_{\frac{j-1}{n}}$. 由于 X 具有独立平稳增量, 故 $\{Y_{j,n}\}_{j\geqslant 1}$ 是独立同分布的. 往往, 我们把 X_1 具有的分布称为无穷可分分布. 事实上，由于无穷可分性，Lévy 过程的分布具有特殊的形式. 我们不加证明地给出下面的定理. 该定理刻画了 Lévy 过程的分布.

定理 6.1.1 设 $X = (X_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的 Lévy 过程，则 X 的特征函数为

$$\mathsf{E}[\exp\{\mathrm{i}uX_t\}] = \exp\{\mathrm{i}ub_t - \frac{u^2}{2}c_t + \int_{\mathbb{R}} (\exp\{\mathrm{i}ux\} - 1 - \mathrm{i}uh(x))\,\mathrm{d}F_t(x)\}.$$

这里 $(b_t)_{t\geqslant 0}, (c_t)_{t\geqslant 0}$ 是关于 t 的确定性函数, F_t 是关于 t 的测度, $h(x) = x\mathbb{1}_{|x|\leqslant 1}$.

6.2 Possion 过程

下面介绍一类十分重要的 Lévy 过程 —— Possion 过程.

假设有一个过程 S, 令 $\mathbb{P}(s)$ 为 S 在 $[t, t+s]$ 内至少有一次跳的概率. 同时, 我们要求该过程是 Lévy 过程, 且每次跳的幅度为 1. 由于 S 是 Lévy 过程, 因此 $\mathbb{P}(s)$ 应与 t 无关, 我们假设

$$\mathbb{P}(\Delta t) = \lambda \Delta t + o(\Delta t),\ \Delta t \downarrow 0.$$

令 X_t 为到时刻 t, 发生跳的次数. 令 $T = \inf\{t:\ X_t = 1\}$, 则

$$\begin{aligned}
\mathsf{P}(T > t) &= \lim_{n\to\infty} \prod_{j=1}^{n} \mathsf{P}\ (S\text{在}[\frac{(j-1)t}{n}, \frac{jt}{n}]\text{中无跳发生}) \\
&= \lim_{n\to\infty} [1 - \mathbb{P}(\frac{t}{n})]^n \\
&= \lim_{n\to\infty} [1 - \frac{\lambda t}{n} + o(\frac{\lambda t}{n})]^n = e^{-\lambda t},
\end{aligned}$$

这说明 T 具有指数分布 (参数为 λ).

我们令 $T_1, T_2, T_3, \cdots$ 为一列独立同分布的随机变量, 且 T_1 服从参数为 λ 的指数分布. 令 $\tau_n = T_1 + \cdots + T_n$. 当 $\tau_n \leqslant t < \tau_{n+1}$, 令 $X_t = n$. 通过上述方式定义的 $X = (X_t)_{t\geqslant 0}$, 有 $X_t = X_{t+} = \lim_{s\downarrow t} X_s$, $X_{t-} = \lim_{s\uparrow t} X_s$. 若 $t = \tau_n$, $X_t = X_{t-} + 1$. 同时, 可以看到 $X_{t+s} - X_s$ 具有 Possion 分布, 且

$$\mathsf{P}(X_{t+s} - X_s = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

记

$$g_k(t) = \mathsf{P}(X_t = k).$$

在 $[t, t+\Delta t]$ 上, 有一次跳的概率是 $\lambda\Delta t + o(\Delta t)$. 即

$$\mathsf{P}(X_{t+\Delta t} = k) = \mathsf{P}(X_t = k-1)(\lambda\Delta t + o(\Delta t)) + \mathsf{P}(X_t = k)(1 - \lambda\Delta t - o(\Delta t)).$$

因此

$$g_k(t+\Delta t) - g_k(t) = \lambda\Delta t[g_{k-1}(t) - g_k(t)] + o(\Delta t),$$

于是

$$\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} = \lambda[g_{k-1}(t) - g_k(t)].$$

解这个方程, 便有

$$g_k(t) = e^{-\lambda t}\frac{(\lambda t)^k}{k!}.$$

从上面的讨论, 我们可以看到 X 本质上即是一开始讨论的 Lévy 过程. 我们称这样的过程为 Possion 过程或 Possion 点过程. 通常习惯用 N 来表示.

定义 6.2.1　设 $N = (N_t)_{t\geqslant 0}$ 是 $(\Omega, \mathcal{F}, \mathsf{P})$ 上的一个随机过程, 满足:

(i) $N_0 = 0$, $N_t \geqslant 0$ 且 N_t 只可能取非负整数值,

(ii) $N_t - N_s$ 与 N_s 独立, $t > s$,

(iii) $N_t - N_s$ 与 N_{t-s} 同分布, 且 $\mathsf{P}(N_t = k) = \frac{(\lambda t)^k}{k!}e^{-\lambda t}$,

我们称 $N = (N_t)_{t\geqslant 0}$ 是一个 **Possion** 过程, 或参数为 λ 的 Possion 过程.

通过 Possion 分布的数字特征, 有

命题 6.2.1　$\mathsf{E}[N_t] = \lambda t$,　　$\mathsf{E}[(N_t - \mathsf{E}N_t)^2] = \lambda t$.

不仅如此, 还可以看到

命题 6.2.2　$\mathsf{P}(N_{t+\Delta t} - N_t \geqslant 2) = o(\Delta t)$.

证明: 由平稳增量性,

$$\begin{aligned}\mathsf{P}(N_{t+\Delta t} - N_t \geqslant 2) &= \mathsf{P}(N_{\Delta t} \geqslant 2)\\ &= \sum_{n=2}^{\infty} e^{-\lambda\Delta t}\frac{(-\lambda\Delta t)^n}{n!}\\ &= o(\Delta t).\end{aligned}$$

■

命题 6.2.3　$(N_t - \lambda t)_{t\geqslant 0}$ 是鞅.

证明: 设 $\mathcal{F}_t = \sigma(N_r:\ 0 \leqslant r \leqslant t)$. 对于 $t > s$,

$$\begin{aligned} \mathsf{E}[N_t - \lambda t \mid \mathcal{F}_s] &= \mathsf{E}[N_t - N_s + N_s - \lambda t \mid \mathcal{F}_s] \\ &= \mathsf{E}[N_t - N_s] + N_s - \lambda t \\ &= N_s - \lambda s, \end{aligned}$$

故 $(N_t - \lambda t)_{t \geqslant 0}$ 是 $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geqslant 0}, \mathsf{P})$ 上的鞅. ∎

Possion 过程往往可以描述一个顾客服务系统. 前面的讨论中 T_n 实际上可认为是第 $n-1$ 个顾客和第 n 个顾客到达时刻的时间间隔. τ_n 即第 n 个顾客的到达时刻. 其密度函数为

$$f_{\tau_n}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t},\ t > 0.$$

事实上, 由于 $\tau_n > t$ 可有 $N_t \leqslant n-1$. 因此

$$\begin{aligned} \mathsf{P}(\tau_n > t) &= \mathsf{P}(N_t \leqslant n-1) \\ &= \sum_{K=0}^{n-1} \mathsf{P}(N_t = K) \\ &= \sum_{K=0}^{n-1} \frac{(\lambda t)^K}{K!} e^{-\lambda t}. \end{aligned}$$

故

$$\mathsf{P}(\tau_n \leqslant t) = 1 - \sum_{K=0}^{n-1} \frac{(\lambda t)^K}{K!} e^{-\lambda t}.$$

通过求导可有

$$f_{\tau_n}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t},\ t > 0.$$

由于 Possion 过程具有独立平稳增量性质. 显然, 这是一个马氏过程. 马氏过程往往可以通过生成元去刻画. 回顾之前的介绍, 对函数 f, 算子 L 为无穷小生成元:

$$Lf(x) = \lim_{t \downarrow 0} \frac{\mathsf{E}[f(N_{t+s}) | N_s = x] - f(x)}{t}.$$

当然, 这里 x 只取非负整值. 当 t 足够小时,

$$\mathsf{P}(N_{t+s} = N_s + 1) = 1 - \mathsf{P}(N_{t+s} = N_s) = \lambda t.$$

因此

$$\mathsf{E}[f(N_{t+s}) \mid N_s = x] = \lambda t f(x+1) + [1 - \lambda t] f(x) + o(t).$$

于是

$$Lf(x) = \lambda[f(x+1) - f(x)].$$

这里没有给出对 f 的限制, 在下文的叙述中，关于 f 的具体要求可参考 Lawler G. [4].

6.3 复合 Possion 过程

令 $T_1, T_2, \cdots$ 是一列独立同分布的随机变量且服从参数为 λ 的指数分布. 令 $N_t = n$. 当 $T_1 + \cdots + T_n \leqslant t < T_1 + \cdots + T_{n+1}$, $N = (N_t)_{t \geqslant 0}$ 是参数为 λ 的 Possion 过程. 令 $Y_1, Y_2, \cdots$ 是独立同分布的随机变量, 且 $\{Y_1, Y_2, \cdots\}$ 与 $\{T_1, T_2, \cdots\}$ 相互独立. 设 Y_1 的分布函数为 F. 令 $S_n = Y_1 + Y_2 + \cdots + Y_n$.

定义 6.3.1 令 $X_t = S_{N_t}$, 则 $X = (X_t)_{t \geqslant 0}$ 称为复合 **Possion** 过程.

Possion 过程与复合 Possion 过程之间的联系可通过随机测度去体现. 令

$$\mu(\mathrm{d}t, \mathrm{d}x) = \sum_{n \geqslant 1}^{\infty} \varepsilon_{(T_n, 1)}(\mathrm{d}t, \mathrm{d}x),$$

这里, $\{T_n - T_{n-1}\}$ 是一列独立的服从参数为 λ 的指数分布的随机变量. 如果 T_n 代表的是跳跃时刻, 那么 $T_i - T_{i-1}$ 独立且共同服从参数为 λ 的指数分布,

$$\int_0^t \int_{\mathbb{R}} x \, \mu(\mathrm{d}s, \mathrm{d}x) = \int_0^t \int_{\mathbb{R}} \sum_{n \geqslant 1}^{\infty} x \ \varepsilon_{(T_n, 1)}(\mathrm{d}s, \mathrm{d}x),$$

$\mathcal{E}_{(T_n, 1)}(\mathrm{d}t, \mathrm{d}x)$ 为 δ-测度. 此时 $\int_0^t \int_{\mathbb{R}} x \ \mu(\mathrm{d}s, \mathrm{d}x)$ 为一个 Possion 过程.
定义随机测度

$$\widetilde{\mu}(\mathrm{d}t, \mathrm{d}x) = \sum_{n \geqslant 1}^{\infty} \varepsilon_{(T_n, Y_n)}(\mathrm{d}t, \mathrm{d}x),$$

$(T_n)_{n \geqslant 1}$ 同上, $(Y_n)_{n \geqslant 1}$ 是一个独立同分布的随机变量, 分布函数为 $F(x)$, $(Y_n)_{n \geqslant 1}$ 与 $(T_n)_{n \geqslant 1}$ 独立. 这时

$$\int_0^t \int_{\mathbb{R}} x \ \widetilde{\mu}(\mathrm{d}s, \mathrm{d}x) = S_{N_t} = X_t.$$

我们能从随机测度 μ 与 $\widetilde{\mu}$ 的区别看出 Possion 过程与复合 Possion 过程之间的联系.

下面介绍 $X = (X_t)_{t \geqslant 0}$ 的一个非常重要的指标 —— Lévy 测度. 令 $B \in \mathcal{B}$, $\mathsf{P}(Y_1 \in B) = \nu^{\#}(B)$, 显然 $\nu^{\#}$ 是 $\mathbb{R}$ 上的一个概率测度. 令

$$\nu = \lambda \nu^{\#}, \qquad \nu(\mathbb{R}) = \lambda,$$

这里称 ν 复合 Possion 过程的 Lévy 测度.

我们知道, 对于连续局部鞅 M, 对于函数 $f \in \mathbf{C}^2$, $f(M)$ 不一定是局部鞅，但是 Itô 公式告诉我们

$$f(M) - \frac{1}{2} f''(M) \cdot \langle M, M \rangle$$

是一个局部鞅. 前面我们已经知道, 若 N 是参数为 λ 的 Possion 过程，$\widetilde{N} = (N_t - \lambda t)$ 是鞅, 人们也想知道 $f(\widetilde{N})$ 是不是鞅? $f(\widetilde{N})$ 减去一个怎么样的过程是局部鞅? 这里需要涉及带跳局部鞅的 Itô 公式，所以就需要定义带跳局部鞅的 Itô 积分, 有关这部分的内容是非常复杂的. 本书不打算涉及这一部分. 事实上，如果 M 是马氏过程, Itô 公式与马氏过程的生成元密切相关. 我们尝试使用生成元来解释 $f(\widetilde{N})$ 减去一个怎么样的过程会变成局部鞅. 在这个过程中，离不开 Lévy 测度.

我们先讨论一些关于 Lévy 测度的性质. 这里涉及到特征函数, 引入如下记号.

若 X 是随机变量, 设 $\mathsf{E}[e^{\mathrm{i}sX}] = e^{\Phi(s)}$, $\Phi(0) = 0$. 对于 $\Phi(s)$, 有时为强调 X, 记为 $\Phi_X(s)$. 若 X 与 Y 独立, 则有 $\Phi_{X+Y} = \Phi_X + \Phi_Y$.

若 $X \sim N(m, \sigma^2)$, 则 $\Phi_X(s) = \mathrm{i}ms - \frac{\sigma^2}{2}s^2$.

若 Y 服从参数为 λ 的 Possion 分布,

$$\begin{aligned} \mathsf{E}[e^{\mathrm{i}sY}] &= \sum_{n=0}^{\infty} e^{\mathrm{i}sn} \mathsf{P}(Y = n) \\ &= \sum_{n=0}^{\infty} \frac{(\lambda e^{\mathrm{i}s})^n}{n!} e^{-\lambda} = e^{\lambda e^{\mathrm{i}s}} e^{-\lambda}. \end{aligned}$$

所以

$$\Phi_Y(s) = \lambda[e^{\mathrm{i}s} - 1].$$

命题 6.3.1 设 $X = (X_t)_{t\geqslant 0}$ 是 Lévy 测度为 ν 的复合 Possion 过程, 则对于 X_1, $\Phi_{X_1}(s) = \int_{-\infty}^{\infty} [e^{\mathrm{i}sx} - 1] \,\mathrm{d}\nu(x)$.

证明 [4]: 令

$$\phi(s) = \mathsf{E}[e^{\mathrm{i}sY_j}] = \int_{-\infty}^{\infty} e^{\mathrm{i}sx} \,\mathrm{d}\nu^{\#}(x).$$

对于 X_t,

$$\mathsf{E}[e^{\mathrm{i}sX_t}] = \sum_{n=0}^{\infty} \mathsf{P}(N_t = n) \mathsf{E}[e^{\mathrm{i}sX_t} \mid N_t = n].$$

注意到

$$\mathsf{E}[e^{\mathrm{i}sX_t} | N_t = n] = \phi(s)^n,$$

故

$$\begin{aligned}\mathsf{E}[e^{\mathrm{i}sX_t}] &= \sum_{n=0}^{\infty} e^{-t\lambda}\frac{(t\lambda)^n}{n!}\phi(s)^n \\ &= e^{-t\lambda}\sum_{n=0}^{\infty}\frac{(t\lambda\phi(s))^n}{n!} = \exp\{t\lambda[\phi(s)-1]\} \\ &= \exp\{t\lambda\int_{-\infty}^{\infty}[e^{\mathrm{i}sx}-1]\,\mathrm{d}\nu^{\#}(x)\} \\ &= \exp\{t\int_{-\infty}^{\infty}[e^{\mathrm{i}sx}-1]\,\mathrm{d}\nu(x)\}.\end{aligned}$$

■

下面我们求复合 Possion 过程的生成元.

定理 6.3.1　设 $X=(X_t)_{t\geqslant 0}$ 是 Lévy 测度为 ν 的复合 Possion 过程. $X_0=0$. 则 X 的生成元为

$$Lf(x)=\int_{-\infty}^{\infty}[f(x+y)-f(x)]\,\mathrm{d}\nu(y).$$

进一步, 设 $\sigma^2=\int x^2\,\mathrm{d}\nu(x)$, $m=\int x\,\mathrm{d}\nu(x)$. 令 $M_t=X_t-tm$, 且 $M=(M_t)_{t\geqslant 0}$ 是平方可积鞅. 且 $Var[M_t]=t\sigma^2$.

证明 [4]**:** 由于 $X=(X_t)_{t\geqslant 0}$ 是复合 Possion 过程. 设 $X_t=S_{N_t}$. $\{N_t\}_{t\geqslant 0}$ 是参数为 λ 的 Possion 过程. 由于 Possion 过程在一个小区间内,

$$\mathsf{P}(N_{t+\Delta t}-N_t\geqslant 2)=o(\Delta t).$$

因此, 由马氏性,

$$\begin{aligned}&\mathsf{E}[f(X_{\Delta t})|X_0=x] \\ &= (1-\lambda\Delta t)f(x)+\lambda\Delta t\int_{-\infty}^{\infty}f(x+y)\,\mathrm{d}\nu^{\#}(y)+o(\Delta t) \\ &= (1-\lambda\Delta t)f(x)+\Delta t\int_{-\infty}^{\infty}f(x+y)\,\mathrm{d}\nu(y)+o(\Delta t) \\ &= f(x)+\Delta t\int_{-\infty}^{\infty}[f(x+y)-f(x)]\,\mathrm{d}\nu(y)+o(\Delta t).\end{aligned}$$

于是

$$\begin{aligned}Lf(x) &= \lim_{\Delta t\to 0}\frac{\mathsf{E}[f(X_{\Delta t})\mid X_0=x]-f(x)}{\Delta t} \\ &= \int_{-\infty}^{\infty}[f(x+y)-f(x)]\,\mathrm{d}\nu(y).\end{aligned}$$

设 $\mathcal{F}_s = \sigma(X_r:\ 0 \leqslant r \leqslant s)$. 下证 $\mathsf{E}[M_t \mid \mathcal{F}_s] = M_s,\ t > s$.

$$\mathsf{E}[X_t \mid \mathcal{F}_s] = \mathsf{E}[X_t - X_s + X_s \mid \mathcal{F}_s] = X_s + \mathsf{E}[X_t - X_s].$$

因此, 需计算 $\mathsf{E}[X_t - X_s]$. 由平稳增量性, 仅计算 $\mathsf{E}[X_t]$.

设

$$\phi_{X_t}(s) = \mathsf{E}[e^{\mathrm{i}sX_t}] = \exp\{t\Phi_{X_1}(s)\}.$$

注意到

$$\Phi_{X_1}(0) = 0,$$
$$\phi'_{X_t}(0) = \mathrm{i}\mathsf{E}[X_t] = t\Phi'_{X_1}(0).$$

故

$$\mathsf{E}[X_t] = -\mathrm{i}t\Phi'_{X_1}(0),\quad \Phi'_{X_1}(s) = \mathrm{i}\int_{-\infty}^{\infty} e^{\mathrm{i}sx} x\,\mathrm{d}\nu(x).$$

注意到

$$\Phi'_{X_1}(0) = \mathrm{i}\int_{-\infty}^{\infty} x\,\mathrm{d}\nu(x).$$

即

$$\mathsf{E}[X_t] = tm, \qquad \mathsf{E}[X_s] = sm.$$

即 $M_t = X_t - tm$ 时,$M = (M_t)_{t\geqslant 0}$ 是鞅. 进一步,

$$Var[M_t] = Var[X_t] = \mathsf{E}[X_t - \mathsf{E}X_t]^2 = \mathsf{E}X_t^2 - (\mathsf{E}X_t)^2.$$

现在计算 $\mathsf{E}[X_t^2]$. 由于

$$\phi''_{X_t}(s) = [t\Phi'_{X_1}(s)]^2 \cdot \exp\{t\Phi_{X_1}(s)\} + t\Phi''_{X_1}(s) \cdot \exp\{t\Phi_{X_1}(s)\}.$$

于是

$$\phi''_{X_t}(0) = t\Phi''_{X_1}(0) + [t\Phi'_{X_1}(0)]^2 = -\mathsf{E}[X_t^2],$$
$$\Phi''_{X_1}(0) = -\int_{-\infty}^{\infty} x^2\,\mathrm{d}\mu(x) = -\sigma^2.$$

故

$$\mathsf{E}[X_t^2] = t\sigma^2 + t^2m^2.$$

注意到 $\mathsf{E}[X_t] = tm$, 于是

$$Var(X_t) = t\sigma^2.$$

■

在上面的讨论中, 我们注意到 $(X_t - tm)_{t\geqslant 0}$ 是鞅. 这里 tm 是确定性的一个函数.

定义 6.3.2 若 X 是一个复合 Possion 过程, 其 Lévy 测度为 ν. 记 $\int x\,\nu(\mathrm{d}x) < +\infty$, 且 $m=\int x\,\nu(\mathrm{d}x)$. 称 $(X_t-mt)_{t\geqslant 0}$ 是与 X 相关的补偿复合 **Possion** 过程. 对于一个半鞅 $S=(S_t)_{t\geqslant 0}$, 若存在一个可料的有限变差过程 $(A_t)_{t\geqslant 0}$ 使 $(S_t-A_t)_{t\geqslant 0}$ 是局部鞅. 称 $A=(A_t)_{t\geqslant 0}$ 为 S 的补偿子.

例 6.3.1 $B=(B_t)_{t\geqslant 0}$ 是标准布朗运动. $(B_t^2-t)_{t\geqslant 0}$ 是鞅. 此时 t 是 B^2 的补偿子.

例 6.3.2 $M=(M_t)_{t\geqslant 0}$ 是连续局部鞅. $(M_t^2-\langle M,M\rangle_t)_{t\geqslant 0}$ 是局部鞅. 称 $\langle M,M\rangle$ 为 M^2 的补偿子.

例 6.3.3 N 是 Possion 过程, 参数为 λ, 则 $(N_t-\lambda t)_{t\geqslant 0}$ 是鞅. 称 λt 为 $(N_t)_{t\geqslant 0}$ 的补偿子.

例 6.3.4 X 是 Lévy 测度为 ν 的复合 Possion 过程, 则 $(X_t-\int_{-\infty}^{\infty} x\,\nu(\mathrm{d}x)t)_{t\geqslant 0}$ 是鞅, 称 $\int_{-\infty}^{\infty} x\,\nu(\mathrm{d}x)t$ 是 X 的补偿子.

下面讨论上面提到的问题: 对于复合 Possion 过程 X, $f(x)$ 是连续函数, $f(X)$ 减去什么过程会变成一个鞅? 换句话说, $f(X)$ 的补偿子是什么?

定理 6.3.2 设 $X=(X_t)_{t\geqslant 0}$ 是 Lévy 测度为 ν 的复合 Possion 过程. 设 f 是连续函数, 且对任意 x 与 t, $\mathsf{E}[f(X_t)^2\mid X_0=x]<+\infty$. 则 $(f(X_t))_{t\geqslant 0}$ 的补偿子为 $(\int_0^t Lf(X_s)\,\mathrm{d}s)_{t\geqslant 0}$. 即 $f(X_t)-\int_0^t Lf(X_s)\,\mathrm{d}s$ 是鞅.

证明 [4]: 考虑 $X_t=\sum\limits_{i=1}^{N_t}\xi_i$, $X_t-X_s=\sum\limits_{i=N_s+1}^{N_t}\xi_i$. 由于 N 是独立平稳增量的, 又由 N 与 $\{\xi_1,\xi_2,\cdots\}$ 独立. 故 X_t-X_s 与 $\sigma(X_r:\ 0\leqslant r\leqslant s)$ 独立, 且 X_t-X_s 与 X_{t-s} 同分布. 故当 $t_2>t_1$ 时,

$$
\begin{aligned}
&\mathsf{E}[f(X_{t_2})-\int_0^{t_2} Lf(X_s)\,\mathrm{d}s\mid \mathcal{F}_{t_1}]\\
&=\mathsf{E}[f(X_{t_2})-f(X_{t_1})+f(X_{t_1})-\int_0^{t_1} Lf(X_s)\,\mathrm{d}s-\int_{t_1}^{t_2} Lf(X_s)\,\mathrm{d}s\mid \mathcal{F}_{t_1}]\\
&=f(X_{t_1})-\int_0^{t_1} Lf(X_s)\,\mathrm{d}s+\mathsf{E}[f(X_{t_2})-f(X_{t_1})-\int_{t_1}^{t_2} Lf(X_s)\,\mathrm{d}s].
\end{aligned}
$$

下面证明对于任意 t,

$$
\mathsf{E}[f(X_t)-f(X_0)-\int_0^t Lf(X_s)\,\mathrm{d}s]=0.
$$

事实上, 只考虑 $t=1$ 即可. 由于

$$
\begin{aligned}
& f(X_1)-f(X_0)-\int_0^1 Lf(X_s)\,\mathrm{d}s \\
=& \sum_{j=1}^n [f(X_{\frac{j}{n}})-f(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s] \\
=& \sum_{j=1}^n [f(X_{\frac{j}{n}})-f(X_{\frac{j-1}{n}})-\frac{1}{n}Lf(X_{\frac{j-1}{n}})] \\
& +\sum_{j=1}^n [\frac{1}{n}Lf(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s].
\end{aligned}
$$

事实上, 由生成元的定义可知

$$\mathsf{E}[f(X_{t+\Delta t}) \mid X_t]=f(X_t)+Lf(X_t)\Delta t+o(\Delta t).$$

故

$$\mathsf{E}[f(X_{t+\Delta t})-f(X_t)-Lf(X_t)\Delta t]=o(\Delta t).$$

因此

$$\mathsf{E}[\sum_{j=1}^n [f(X_{\frac{j}{n}})-f(X_{\frac{j-1}{n}})-\frac{1}{n}Lf(X_{\frac{j-1}{n}})]]=\mathsf{E}[\sum_{j=1}^n o(\frac{1}{n})]\to 0.$$

下面证明

$$\lim_{n\to\infty}\sum_{j=1}^n \mathsf{E}[\frac{1}{n}Lf(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s]=0.$$

考虑

$$\mathsf{E}[\frac{1}{n}Lf(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s].$$

当在 $[\frac{j-1}{n},\frac{j}{n}]$ 中发生跳时, 上式期望不为 0. 此事件发生的概率为 $O(\frac{1}{n})$. 这种情形下,$\frac{1}{n}Lf(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s$ 在局部化后可认为 $O(\frac{1}{n})$. 故

$$\mathsf{E}[\frac{1}{n}Lf(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s]=O(\frac{1}{n^2}).$$

因此

$$\sum_{j=1}^n \mathsf{E}[\frac{1}{n}Lf(X_{\frac{j-1}{n}})-\int_{\frac{j-1}{n}}^{\frac{j}{n}} Lf(X_s)\,\mathrm{d}s]=o(1).$$

于是

$$\mathsf{E}[f(X_1)-f(X_0)-\int_0^1 Lf(X_s)\,\mathrm{d}s]=0.$$

结合独立增量性, $(f(X_t)-\int_0^t Lf(X_s)\,\mathrm{d}s)_{t\geqslant 0}$ 是鞅, $(\int_0^t Lf(X_s)\,\mathrm{d}s)_{t\geqslant 0}$ 是 $(f(X_t))_{t\geqslant 0}$ 的补偿子. ■

例 6.3.5 设 $X=(X_t)_{t\geqslant 0}$ 是复合 Possion 过程. $X_t=\sum_{i=1}^{N_t}\xi_i$, N_t 是参数为 2 的 Possion 过程, $\xi_i\sim N(0,2)$, $X_0=0$.

(1) 求 X 的 Lévy 测度.

(2) 求 $\mathsf{E}[X_t]$.

(3) 求 $\mathsf{E}[X_t^2]$.

(4) 求 $\mathsf{E}[e^{X_t}]$.

解: (1). ξ_i 的分布 $\nu^{\#}$ 为 $\mu^{\#}(A)=\int_A \frac{1}{\sqrt{4\pi}}e^{-\frac{x^2}{4}}\,\mathrm{d}x$, $A\in\mathcal{B}$. 故 X 的 Lévy 测度 μ 定义如下: 对于 $A\in\mathcal{B}$,

$$\nu(A)=\int_A \frac{2}{\sqrt{4\pi}}e^{-\frac{x^2}{4}}\,\mathrm{d}x=\frac{1}{\sqrt{\pi}}\int_A e^{-\frac{x^2}{4}}\,\mathrm{d}x.$$

(2).

$$\mathsf{E}[X_t]=t\int_{-\infty}^{\infty}x\,\nu(\mathrm{d}x)=t\int_{-\infty}^{\infty}x\frac{1}{\sqrt{\pi}}e^{-\frac{x^2}{4}}\,\mathrm{d}x=0.$$

■

(3).

$$\mathsf{E}[X_t^2]=t\int_{-\infty}^{\infty}x^2\,\nu(\mathrm{d}x)=t\int_{-\infty}^{\infty}x^2\frac{1}{\sqrt{\pi}}e^{-\frac{x^2}{4}}\,\mathrm{d}x=4t.$$

(4). 令 $Z_t=e^{X_t}$, 可有

$$\mathsf{E}[Z_{t+s}\mid\mathcal{F}_t]=\mathsf{E}[Z_t\exp\{X_{t+s}-X_t\}\mid\mathcal{F}_t]=Z_t\mathsf{E}[\exp\{X_{t+s}-X_t\}].$$

于是

$$\mathsf{E}[Z_{t+s}]=\mathsf{E}[Z_t]\mathsf{E}[Z_s].$$

由于 $X_0=0$, 故 $Z_0=1$. 有 $\mathsf{E}[Z_t]=e^{gt}$. 下面求 g. 令 $g(x)=e^x$, 由 $\mathsf{E}[Z_t]$ 的形式

$$g=\lim_{s\downarrow 0}\frac{\mathsf{E}[g(X_s)]-\mathsf{E}[g(X_0)]}{s}=Lg(0).$$

X 的生成元为

$$Lg(x)=\int_{-\infty}^{\infty}[g(x+y)-g(x)]\,\nu(\mathrm{d}y).$$

故

$$
\begin{aligned}
Lg(0) &= \int_{-\infty}^{\infty} [e^y - 1]\,\nu(\mathrm{d}y) \\
&= \int_{-\infty}^{\infty} [e^y - 1]\frac{1}{\pi}e^{-\frac{y^2}{4}}\,\mathrm{d}y = 2[e-1].
\end{aligned}
$$

因此

$$\mathsf{E}[Z_t] = e^{2[e-1]}t.$$

■

上面的讨论中, 关于复合 Possion 过程的讨论主要围绕鞅展开. 下面, 关于补偿复合 Possion 过程的鞅性, 有如下结论.

定理 6.3.3 设 X 是复合 Possion 过程, ν 是 X 的 Lévy 测度, 且

$$\sigma^2 = \int x^2\,\nu(\mathrm{d}x) < +\infty, \qquad m = \int x\,\nu(\mathrm{d}x).$$

设 $M_t = X_t - mt$. $A = (A_t)_{t\geqslant 0}$ 是关于流 $\{\mathcal{F}_t\}_{t\geqslant 0}$ 适应的随机过程. $\mathcal{F}_t = \sigma(X_s : 0 \leqslant s \leqslant t)$, 且 $A_t = A_{t-}$, $\int_0^\infty \mathsf{E}[A_s^2]\,\mathrm{d}s < +\infty$. 关于 M 作 Lebesgue-Stieltjes 积分 $Z_t = \int_0^t A_s\,\mathrm{d}M_s$, 则 $Z = (Z_t)_{t\geqslant 0}$ 是鞅, 且 $\mathsf{E}[Z_t^2] = \sigma^2 \int_0^t \mathsf{E}[A_s^2]\,\mathrm{d}s$.

证明: 首先假设 $A = (A_t)_{t\geqslant 0}$ 是简单过程:

$$A = \sum_{i=0}^{\infty} Y_{\frac{i}{n}} \mathbb{1}_{(\frac{i}{n},\frac{i+1}{n}]},\ Y_{\frac{i}{n}} \text{关于} \mathcal{F}_{\frac{i}{n}} \text{可测}.$$

因此, 对于所有 t, 当 $\frac{j}{n} \leqslant t < \frac{j+1}{n}$ 时,

$$Z_t = \sum_{i=0}^{j-1} A_{\frac{i}{n}}(M_{\frac{i+1}{n}} - M_{\frac{i}{n}}) + A_{\frac{j}{n}}(M_t - M_{\frac{j}{n}}).$$

显然, $Z = (Z_t)_{t\geqslant 0}$ 是鞅. 且 $\mathsf{E}[Z_t^2] = \mathsf{E}[\int_0^t A_s^2\,\mathrm{d}s]\sigma^2$. 对于左连续过程 A, 存在一列简单过程 A^n, 使 A_t^n 收敛至 A_t. 利用局部化技巧及控制收敛定理, 可知 $\int_0^t A_s^n\,\mathrm{d}M_s$ 依概率收敛于 $\int_0^t A_s\,\mathrm{d}M_s$. $(\int_0^t A_s\,\mathrm{d}M_s)_{t\geqslant 0}$ 的鞅性与方差的计算可由 $\mathbf{L}^2$ 收敛保证.

■

6.4 习题

1. 请给出参数为 λ 的 Possion 过程的 Lévy 测度.
2. 设 $N = (N_t)_{t\geqslant 0}$ 是参数为 3 的 Possion 过程,

(1) 求 $\mathsf{P}(N_3 > N_1 + 3 | N_1 = 2)$;

(2) 求 $\mathsf{P}(N_1 = 1 | N_3 = 4)$.

3. 设 $N = (N_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$ 上的参数为 2 的 Possion 过程, $\xi_1, \xi_2, \cdots$, 是一列独立同分布的随机变量, 且

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

定义 X 如下:

$$X_0 = 0, \qquad X_t = \sum_{i=1}^{N_t} \xi_i.$$

(1) 求 X 的 Lévy 测度;

(2) 求 X 的生成元;

(3) 求关于 X 的补偿 Possion 过程;

(4) $Z_t = \exp\{X_t\}$, 求 $Z = (Z_t)_{t\geqslant 0}$ 的补偿子.

4. 设 $N = (N_t)_{t\geqslant 0}$ 是带流概率空间 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathsf{P})$ 上的参数为 2 的 Possion 过程, $\xi_1, \xi_2, \cdots$, 是一列独立同分布的随机变量, 且服从 $[-\frac{1}{2}, \frac{1}{2}]$ 上的均匀分布.

定义 X 如下:

$$X_0 = 0, \qquad X_t = \sum_{i=1}^{N_t} \xi_i.$$

(1) 求 X 的 Lévy 测度;

(2) 求 $\mathsf{E}[X_t^2]$;

(3) 求 X_1 的特征函数;

(4) 求 $S = (S_t)_{t\geqslant 0}$, 使得 $(X_t^2 - S_t)_{t\geqslant 0}$ 是鞅;

(5) 令 $Z_t = \exp\{X_t\}$, 求 $A = (A_t)_{t\geqslant 0}$, 使得 $(Z_t A_t)_{t\geqslant 0}$ 是鞅.

参考文献

[1] Csörgö M., Révész P. Strong Approximations in Probability and Statistics. Academic Press, (1981).

[2] Halmos R. Measure Theory. Springer, (1998).

[3] Jacod J., Shiryaev A. Limit Theorems for Stochastic Processes (Second edition). Springer, (2003).

[4] Lawler G. Introduction to Stochastic Calculus with Applications. Chapman and Hall/CRC, (2021).

[5] 林正炎, 陆传荣, 苏中根. 概率极限理论基础 (第二版). 北京: 高等教育出版社, (2015).

[6] Le Gall, F. Brownian Motion, Martingales, and Stochastic Calculus.Springer International Publishing, (2013).

[7] Kallenberg, O. Foundations of Modern Probability. Probability and its Applications. 2nd ed. Springer, (2002).

[8] Peng, S. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. Stochastics Stochastics Rep.37, no.1-2, 61-74, (1991).

[9] Peng, S. Backward stochastic differential equation, nonlinear expectation and their applications. Proceedings of the International Congress of Mathematicians. Volume I, 393-432, Hindustan Book Agency, New Delhi, (2010).

[10] 钱忠民, 应坚刚. 随机分析引论. 上海: 复旦大学出版社, (2017).

[11] 任佳刚. 随机过程教程. 北京: 科学出版社, (2022).

[12] Revuz D.,Yor M. Continuous Martingales and Brownian Motion(Third Edition).Springer, (2010).

[13] 汪嘉冈. 现代概率论基础 (第二版). 上海：复旦大学出版社, (2005).

[14] 严加安. 测度论讲义 (第二版). 北京: 科学出版社, (2004).

[15] 严士健, 刘秀芳. 测度与概率 (第二版). 北京：北京师范大学出版社, (2014).

[16] 应坚刚, 金蒙伟. 随机过程基础 (第二版). 上海：复旦大学出版社, (2020).

[17] Zhang J. Backward Stochastic Differential Equations. Springer, (2017).