

随机分析引论

Introduction to Stochastic Analysis

作者: 尹俊淇 Derek Franklin

组织: School of Economics, Zhejiang University; Faculty of Science, National University of Singapore

时间: July 13, 2024

版本: 3.14



我想说的话

随机分析作为现代数学的重要分支之一,近年来在诸多领域中得到了广泛的应用,尤其是在金融数学中展现出了其独特的魅力和强大功能.《随机分析引论》一书旨在帮助读者打下一个系统的、全面的随机分析基础,为读者打开随即分析的门.同时渗透一些在金融数学中的实际应用.参考本资料需要有一定的概率论、随机过程、实变函数、常微分方程、线性代数基础.

本书共分为六章内容:第一章概率基础:介绍概率论的基础内容,包括随机变量、数字特征、收敛定理,在附录中还提到了测度变换.理解这些基础知识是掌握随机分析其他内容的关键.第二章马尔科夫链:探讨离散时间和连续时间马尔科夫链的理论,研究其平稳分布和长期行为.这部分内容在各种随机过程的分析中起着重要作用.第三章鞅:介绍鞅的定义和基本性质,讨论其在随机过程中的重要应用.鞅理论是理解更复杂随机过程的基础工具之一.第四章布朗运动:布朗运动作为随机过程的经典例子,具有丰富的理论和应用;同时是随即分析的核心.第五章随机积分和随机分析:介绍伊藤积分和伊藤公式,探讨随机微积分的基本原理和风险中性定价.理解随机积分是进一步研究随机微分方程的前提.第六章随机微分方程和偏微分方程:讲解随机微分方程和偏微分方程的关系.

在每一章中,我们不仅关注理论的严谨性和完整性,还特别注重应用实例和直观解释,力求使读者不仅能够理解随机分析的理论基础,还能掌握其实际应用技巧.特别是书中将结合金融数学的思想,通过具体案例分析,展示随机分析在金融工程、风险管理和衍生品定价等方面的重要应用.

在此,鄙人深知个人学识有限,虽尽己所能编写此书,但难免会有疏漏和不足之处. 恳请读者在阅读过程中,不吝指正,以期共同进步. 感谢所有对本书编写提供帮助和支持的同学、朋友、师兄师长及家人,特别感谢读者对本书的关注和厚爱. 愿本文能够成为您学习和研究随机分析的有力助手,并为您的专业成长提供有益的启迪.

目录

第1草	预备知识	1
1.1	概率空间	1
1.2	随机变量及其数字特征	2
1.3	随机变量的独立性、联合分布	7
1.4	条件数学期望	9
1.5	收敛的概念	11
1.6	特征函数	16
1.7	附录	24
第2章		31
2.1	离散、有限时间的马尔可夫链	
2.2	不变分布	
2.3	常返性与瞬时性	37
2.4	极限分布与有限马尔可夫链的遍历定理	43
2.5	连续时间离散状态的马尔可夫链	45
	2.5.1 基本概念	45
	2.5.2 嵌入式链和不可约性	48
	2.5.3 Q 过程的遍历定理	49
2.6	连续时间连续状态的马氏过程介绍	50
第3章	岩 血	52
3.1	************************************	
3.1	停时原理和离散鞅收敛定理	
3.3	连续时间鞅介绍	
3.3	任练时间软灯组	39
第4章	布朗运动	61
4.1	高斯过程再讨论	61
4.2	随机游动	61
4.3	布朗运动	67
	4.3.1 布朗运动的分布	69
	4.3.2 布朗运动的 Karhunen-Loeve 展开	71
	4.3.3 布朗运动的鞅性质	72
4.4	二次变差与几何布朗运动	74
	4.4.1 一阶变差	74
	4.4.2 二次变差	75
		77
		78
4.5		80
4.6	=	81
7.0		81
		83
	4.6.3 有零点的概率和反正弦定律	04

		录
4.7	布朗运动的变式	36
	4.7.1 布朗桥	36
	4.7.2 漂移布朗运动	37
第5章	随机积分与随机分析初步 8	39
5.1	关于随机游动的随机积分	39
5.2	简单被积函数的 Ito 积分	90
	5.2.1 简单函数 Ito 积分的构造	90
	5.2.2 简单函数 Ito 积分的性质	91
5.3	一般被积函数的 Ito 积分	93
	5.3.1 一般被积函数 Ito 积分的构造 9	93
	5.3.2 Ito 积分的性质	94
5.4	Ito 公式	96
5.5	Ito 积分过程)0
	5.5.1 Ito 过程	00
	5.5.2 相关金融模型)4
5.6	测度变换)6
5.7	Black-Scholes-Merton 方程)8
	5.7.1 Black-Scholes-Merton 模型的设定)8
	5.7.2 Delta 对冲法则)9
	5.7.3 Black-Scholes-Merton 方程的解	10
	5.7.4 希腊字母	12
	5.7.5 看跌、看涨平价公式	14
5.8	多元随机分析初步	14
	5.8.1 多元布朗运动	14
	5.8.2 多个过程的 Ito 公式	15
5.9	布朗桥再讨论 11	18
5.10	附录	21
	5.10.1 资产定价前置定理	21
	5.10.2 资产定价基本定理 12	23
第6章	随机微分方程与偏微分方程 12	24
6.1	随机微分方程	
6.2	费曼-卡茨定理	
	一些随机微分方程示例	

第1章 预备知识

内容提要

- □ 概率空间
- □ 随机变量
- □ 条件数学期望

- □ 收敛的概念
- □ 特征函数
- □ 附录

1.1 概率空间

定义 1.1 (样本空间)

样本空间 Ω 是随机试验所有可能结果的集合. 每个元素 $\omega \in \Omega$ 被称为样本点.

*

定义 1.2 (σ 代数)

一个 σ -代数 (或称 σ -域) F 是满足如下条件的 Ω 的子集所构成的集族.

- $\Omega \in \mathcal{F}$:
- $\not\equiv A_1, A_2, \dots \in \mathcal{F}, \ y \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.$

每一个集合 $A \in \mathcal{F}$ 被称为一个事件. 令 \mathcal{B} 为一个 Ω 的子集的集族. 称 $\sigma(\mathcal{B})$ 为由 \mathcal{B} 中集合生成的最小 σ -代数. 具有如上性质的 (Ω, \mathcal{F}) 是一个可测空间.

例题 1.1 很简单的, 从事件与样本的角度来说, 随机投郑一次股子是作为一种随机试验, 它简单满足三个条件:

- 1. 可以被重复;
- 2. 具有多种可能结果并均明确可知;
- 3. 试验结果不可预先获知.

而我们将骰子可能出现的每一种结果称为样本点,样本点是一种单例,换言之,它的结果是单一而不可再分解的,每一个样本点都彼此不相容,因而它是构成试验结果的最基础原子,故而样本点也被称为基本事件.而所有样本点 (也即所有的可能结果)的集合就被称为样本空间,亦或者可以说所有的基本事件组成的集合称为基本事件空间,记作:

$$\Omega = \{\omega\} = \{1, 2, 3, 4, 5, 6\}$$

接着,我们可以将一次试验可能出现的结果称为随机事件,例如例子中的骰子为偶数点数就是一个随机事件,可使用大写字母记为:

$$A = \{$$
股子点数为偶数 $\} = \{2,4,6\}$

定义 1.3 (概率测度)

概率测度 $\mathbb{P}: \mathcal{F} \to [0,1]$ 是定义在 \mathcal{F} 上的集函数, 满足:

- 规范性: $\mathbb{P}(\emptyset) = 0, \mathbb{P}(\Omega) = 1;$
- 可列可加性: $A_1, A_2, \dots \in \mathcal{F}$ 两两不相交, 即 $A_i \cap A_j = \emptyset$ 若 $i \neq j$, 则

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}\left(A_n\right)$$

定义 1.4 (条件概率)

令 A, B ∈ F 并假设 $\mathbb{P}(B) \neq 0$. A 关于给定的 B 的条件概率被定义为:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

既然根据定义有 $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\mathbb{P}(B)$, 所以我们能得到:

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \mid B \cap C)\mathbb{P}(B \mid C)\mathbb{P}(C),$$

并且以此类推,就得到下面的结论:

命题 1.1 (链式法则)

$$\mathbb{P}(A_1 A_2 \cdots A_m) = \mathbb{P}(A_1) \mathbb{P}(A_2 | A_1) \mathbb{P}(A_3 | A_2 A_1) \cdots \mathbb{P}(A_m | A_{m-1} A_{m-2} \cdots A_1)$$

初等概率论课程中详细讲过全概率公式和贝叶斯公式,它是一些统计学习算法的基础,也是贝叶斯统计的基础.

定理 1.1 (朴素贝叶斯)

如果 A_1, A_2, \cdots 互不相交, 且 $\bigcup_{i=1}^{\infty} A_i = \Omega$ (此时称诸事件为样本空间的一个分割), 则有

$$\mathbb{P}\left(A_j \mid B\right) = \frac{\mathbb{P}\left(A_j\right)\mathbb{P}\left(B \mid A_j\right)}{\sum_{n=1}^{\infty}\mathbb{P}\left(A_n\right)\mathbb{P}\left(B \mid A_n\right)} \quad \text{ $\forall \text{ $\not = \mathbb{N}$.}}$$

这在贝叶斯统计中很有用, A_j 对应于原假设, $\mathbb{P}(A_j)$ 是假设 A_j 的**先验概率**. 条件概率 $\mathbb{P}(A_j \mid B)$ 就是当某事件 B 发生时, A_j 的**后验概率**.

1.2 随机变量及其数字特征

有些时候 Ω 并不是必须可数的, 或者说其构成的元素不能直接进行数学分析. 让我们从随机变量的定义开始. 令 \mathcal{R} 表示 \mathbb{R} 上的博雷尔 σ -代数, 即包含所有开区间的最小 σ -代数.

定义 1.5 (随机变量)

一个随机变量 X 是一个 F 可测的 实值函数 $X:\Omega\to\mathbb{R}$; 即对任意 $B\in\mathcal{R},X^{-1}(B)\in\mathcal{F}$.

Ŷ 笔记

- 1. $B ∈ \mathcal{R}$ 意思是 B 为一个实数域上的 Borel 集;
- 2. 随机变量是一个函数 (映射), 且满足 $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$ (是一系列样本点构成的集合). 简而言之, 随机变量把事件域映射到实数域, 使得可以用区间来表示事件, 进而进行定量分析.

定义 1.6 (分布)

随机变量 X 的分布是一个 \mathbb{R} 上的概率测度 μ , 对任意 $B \in \mathcal{R}$, 我们定义

$$\mu(B) = \mathbb{P}(X \in B) = \mathbb{P} \circ X^{-1}(B).$$

特别地, 定义分布函数为 $F(x) = \mathbb{P}(X \leq x)$ 当 $B = (-\infty, x]$

◈ 笙记

- 1. 不可以直接把概率测度 \mathbb{P} 作用在 B 上, 因为概率定义在事件域 \mathcal{F} 上, 所以随机变量落在某 Borel 集上的概率是一个复合: $\mathbb{P}(X \in B) = \mathbb{P} \circ X^{-1}(B)$.
- 2. 分布依旧是一个测度, 它定义在 R 上.

3. 分布函数 F(x) 具有单调、有界、右连续性质, 且 $F(-\infty) = 0$, $F(+\infty) = 1$.

测度空间 $(\{X(\omega)\}, \mathcal{R}, m)$ 对应概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ (其中 m 为 Lebesgue 测度). 而 $\mu(B): \mathcal{R} \to [0, 1]$ 是一个广义测度, 因为其满足:

- 1. $\mu(\emptyset) = 0$;
- 2. $\forall B = \bigcup_{i=1}^{\infty} B_i, B_i \in \mathcal{R}(\{B_i\} 两两不相交), \mu(B) = \sum_{i=1}^{\infty} \mu(B_i).$
- 3. μ是绝对连续函数.

由 Radon-Nikodym 定理, 存在导函数 $\rho(x)$, 使得

$$\mu(B) = \int_{B} \rho(x) dm (\forall B \in \mathcal{R})$$

且记 $\rho = \frac{d\mu}{dm}$, 称之为随机变量 X 的概率密度函数 (probability density function).



笔记 连续随机变量取值为 \mathbb{R} 上一个区间或多个区间, 并存在一个密度函数 p(x) 使得

$$F(x) = \int_{-\infty}^{x} p(s)ds, \quad x \in \mathbb{R}.$$

该分布函数为绝对连续函数, 其导数为 p(x)

例题 1.2 正态分布

连续型随机变量 X 如果有如下形式的密度函数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (\mu \in R, \sigma > 0)$$

则称 X 服从参数为 (μ, σ^2) 的正态分布 (normal distribution), 记为 $X \sim N(\mu, \sigma^2)$.

初等概率论中已经详细证明过正态分布的两个参数就为其期望和方差,接下来我们用概率测度的角度来对期望进行定义.

令 X 是定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量. 若 Ω 是有限的, 则 X 的期望可以简单表示为:

$$\mathbb{E}X = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\omega)$$

如果 Ω 是可列但无限的, 即其元素可以列为 $\omega_1, \omega_2, \omega_3, \cdots$, 则我们也可以用级数来定义 $\mathbb{E} X$:

$$\mathbb{E}X = \sum_{k=1}^{\infty} X(\omega_k) \mathbb{P}(\omega_k).$$

然而, 如果 Ω 是不可计数的无穷大, 就会出现困难, 因为不可计数的和不能被定义. 此时, 我们必须从积分的角度来考虑问题.

首先回顾黎曼积分: 如果 f(x) 是在封闭区间 [a,b] 上对所有 x 定义的连续函数, 我们定义黎曼积分 $\int_a^b f(x)dx$ 如下: 首先分割 [a,b] 为一系列子区间 $[x_0,x_1]$, $[x_1,x_2]$, \cdots , $[x_{n-1},x_n]$, 其中 $a=x_0 < x_1 < \cdots < x_n = b$. 我们用 $\Pi = \{x_0,x_1,\cdots,x_n\}$ 表示划分点的集合, 用

$$\|\Pi\| = \max_{1 \le k \le n} (x_k - x_{k-1})$$

表示分区中最长子区间的长度 (即细度). 对于每个子区间 $[x_{k-1},x_k]$, 我们设 $M_k=\max_{x_{k-1}\leq x\leq x_k}f(x)$ 和 $m_k=\min_{x_{k-1}\leq x\leq x_k}f(x)$. 于是, 上黎曼和为

$$RS_{\Pi}^{+}(f) = \sum_{k=1}^{n} M_k (x_k - x_{k-1})$$

下黎曼和是

$$RS_{\Pi}^{-}(f) = \sum_{k=1}^{n} m_k (x_k - x_{k-1}).$$

当 $\|\Pi\|$ 收敛于零时 (即当我们放入越来越多的分割点, 分划中的子区间变得越来越短), 上黎曼和 $\mathrm{RS}_\Pi^+(f)$ 和下黎曼和 $\mathrm{RS}_\Pi^-(f)$ 收敛于同一极限, 我们就称之为黎曼积分 $\int_a^b f(x)dx$.

我们在模仿这个过程来定义期望时遇到的问题是: 随机变量 X 与上一段中的函数 f 不同, 它是 $\omega \in \Omega$ 的函

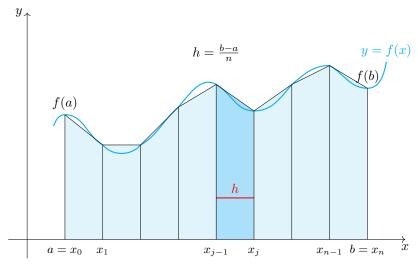


图 1.1: 黎曼积分示意图

数, 而 Ω 通常不是 \mathbb{R} 的子集. 即下图中的 "x 轴"不是实数轴, 而是一些<u>抽象的空间</u> Ω . 没有一种自然的方法像我们在上面划分 [a,b] 那样划分 Ω 集合. 因此, 我们将以下图中的 y 轴进行划分.

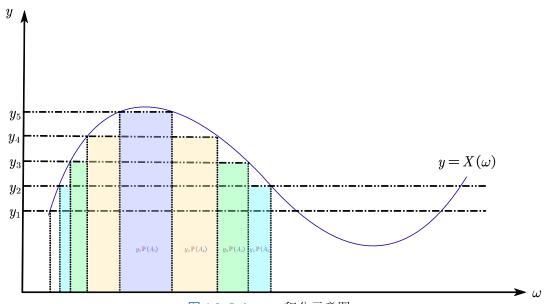


图 1.2: Lebesgue 积分示意图

首先假设每个 $\omega \in \Omega$ 对应 $0 \le X(\omega) < \infty$, 令分割点的点集为 $\Pi = \{y_0, y_1, y_2, \cdots\}$, 其中 $0 = y_0 < y_1 < y_2 < \cdots$. 对于每个子区间 $[y_k, y_{k+1}]$, 我们设

$$A_k = \{ \omega \in \Omega; y_k \le X(\omega) < y_{k+1} \}$$

定义下勒贝格和为

$$\mathrm{LS}_{\Pi}^{-}(X) = \sum_{k=1}^{\infty} y_k \mathbb{P}(A_k)$$

当 y_k 划分点之间的细度 $\|\Pi\|$ 趋近于零时,这个较低的和收敛,我们将这个极限定义为勒贝格积分 $\|\Pi\|$

$$\int_{\Omega} X(\omega) d\mathbb{P}(\omega)$$

¹这个方式是模仿实分析中的勒贝格积分进行定义,严格来讲这里应该称为依概率测度定义的积分

简记为 $\int_{\Omega} X d\mathbb{P}$. 勒贝格积分可以是 ∞ , 因为我们没有对 X 的值有多大做任何假设. 如果 $\mathbb{P}\{\omega; X(\omega) \geq 0\} = 1$ 但 $\mathbb{P}\{\omega; X(\omega) = \infty\} > 0$, 那么我们定义 $\int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \infty$.

最后, 我们需要考虑可以取正负的随机变量 X. 和实分析中类似, 定义正负部:

$$X^{+}(\omega) = \max\{X(\omega), 0\}, \quad X^{-}(\omega) = \max\{-X(\omega), 0\}.$$

 X^+ 和 X^- 都是非负随机变量, 且满足 $X=X^+-X^-$ 和 $|X|=X^++X^-$. $\int_\Omega X^+(\omega)d\mathbb{P}(\omega)$ 和 $\int_\Omega X^-(\omega)d\mathbb{P}(\omega)$ 都是由上述的过程定义的, 只要它们不是 ∞ , 我们就可以定义

$$\int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int_{\Omega} X^{+}(\omega) d\mathbb{P}(\omega) - \int_{\Omega} X^{-}(\omega) d\mathbb{P}(\omega).$$

如果 $\int_{\Omega} X^+(\omega) d\mathbb{P}(\omega)$ 和 $\int_{\Omega} X^-(\omega) d\mathbb{P}(\omega)$ 都是有限的, 我们就说 X 是**可积的**, 且 $\int_{\Omega} X(\omega) d\mathbb{P}(\omega)$ 也是有限的. 如果 $\int_{\Omega} X^+(\omega) d\mathbb{P}(\omega) = \infty$, $\int_{\Omega} X^-(\omega) d\mathbb{P}(\omega)$ 是有限的, 那么 $\int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \infty$. 如果 $\int_{\Omega} X^+(\omega) d\mathbb{P}(\omega)$ 是有限的, $\int_{\Omega} X^-(\omega) d\mathbb{P}(\omega) = \infty$, 那么 $\int_{\Omega} X(\omega) d\mathbb{P}(\omega) = -\infty$. 如果 $\int_{\Omega} X^+(\omega) d\mathbb{P}(\omega) = \infty$ 和 $\int_{\Omega} X^-(\omega) d\mathbb{P}(\omega) = \infty$ 同时出现,则出现" $\infty - \infty$ "的情况,此时认为 $\int_{\Omega} X(\omega) d\mathbb{P}(\omega)$ 没有定义.

定义 1.7 (数学期望)

令 X 是定义在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量. X 的期望定义为

$$\mathbb{E}X = \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

期望有意义当且仅当

$$\mathbb{E}|X| = \int_{\Omega} |X(\omega)| d\mathbb{P}(\omega) < \infty$$

或 $X \ge 0$, a.s.. 在后一种情况, $\mathbb{E}X$ 可能为 ∞ .

定理 1.2 (数学期望的性质)

令 X 是定义在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量.

1. 若 X 取离散值 $x_0, x_1, x_2, \dots, 则$

$$\mathbb{E}X = \sum_{k=0}^{\infty} x_k \mathbb{P}\left\{X = x_k\right\}$$

2. 若 X 是连续随机变量,则期望可以表示为

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \cdot p(x) dx$$

其中 p(x) 是 X 的密度函数;

3. (可积性) X 是可积的当且仅当

$$\mathbb{E}|X| < \infty$$

令 Y 是另一个定义在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量, 则有有如下结论:

4. (可比较性) 若 $X \leq Y$ a.s. 且 X,Y 都是可积的或几乎必然非负,则

$$\mathbb{E}X < \mathbb{E}Y$$
.

特别地,如果 X = Y 几乎必然成立,且其中一个随机变量是可积的或几乎必然非负的,那么它们分别是可积的或几乎肯定非负的,并且

$$\mathbb{E}X = \mathbb{E}Y$$
.

5. (线性性) 若 α 和 β 为实常数, X 和 Y 可积, 则

$$\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}X + \beta \mathbb{E}Y.$$

6. (琴生不等式) 若 φ 是实值凸函数, 且若 $\mathbb{E}|X| < \infty$, 则

$$\varphi(\mathbb{E}X) \leq \mathbb{E}\varphi(X)$$

 \Diamond

定义 1.8 (数字特征)

• X 的方差 (variance) 定义为:

$$Var(X) = \mathbb{E}(X - \mathbb{E}X)^2$$
.

• 对两个随机变量 X and Y, 定义二者之间的协方差 (covariance) 为:

$$Cov(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y).$$

如果 Cov(X,Y) = 0 就称 X,Y 不相关.

 $\stackrel{\diamondsuit}{\mathbf{C}}$ 笔记 以上所有的定义都可以推广到向量的情况, $\mathbf{X} = (X_1, X_2, \cdots, X_d)^{\top} \in \mathbb{R}^d$ 是一个随机向量, 每个分量 X_k 是一个随机变量. 在这种情况下, \mathbf{X} 的协方差矩阵 (亦称随机向量的方差) 定义为

$$Cov(X) = \mathbb{E}(X - \mathbb{E}X)(X - \mathbb{E}X)^{\top}$$

注意, 在讨论随机变量数学期望有无意义的时候, 我们考虑了 图|X| 是否有限, 接下来我们进行一些推广.

定义 **1.9** (L^p 空间)

对任意 $p \ge 1$, $L^p(\Omega)$ 空间 (或记为 $L^p(\Omega)$) 包含了所有 p 阶矩有限的随机变量:

$$L^p(\Omega) = \{ \boldsymbol{X}(\omega) : \mathbb{E} | \boldsymbol{X} | p < \infty \}.$$

对 $X \in L^p(\Omega)$, 令

$$\|X\|_p = (\mathbb{E}|X|^p)^{1/p}, \quad p \ge 1.$$

称为随机变量 (向量)X 的 p 范数.

定理 1.3 (两个不等式)

1. Minkowski 不等式.

$$\|X + Y\|_p \le \|X\|_p + \|Y\|_p, \quad p \ge 1, \quad X, Y \in L^p(\Omega)$$

2. Hölder 不等式.

$$\mathbb{E}|(\boldsymbol{X},\boldsymbol{Y})| \leq \|\boldsymbol{X}\|_p \|\boldsymbol{Y}\|_q, \quad p>1, \quad 1/p+1/q=1, \quad \boldsymbol{X} \in L^p(\Omega), \quad \boldsymbol{Y} \in L^q(\Omega),$$
 其中 $(\boldsymbol{X},\boldsymbol{Y})$ 表示两个随机向量在 \mathbb{R}^d 中的内积.

证明 首先证明对于任意正数 a 和 b,且正数 p 和 q 满足 $\frac{1}{n} + \frac{1}{a} = 1$,必有

$$\frac{1}{p}a^p + \frac{1}{q}b^q \ge ab$$

当且仅当 $a^p=b^q$ 时等号成立. 该式的证明, 只需构造函数 $g(a)=\frac{1}{p}a^p+\frac{1}{q}b^q-ab$, 该函数最小值在 $a^{p-1}=b$ 时取到 0, 因此上式成立. 再将 $a=\frac{|X|}{\left[\mathbb{E}\left(|X|^p\right)\right]^{1/p}}$ 和 $b=\frac{|Y|}{\left[\mathbb{E}\left(|Y|^q\right)\right]^{1/q}}$ 代入, 再对两边同时取期望, 即可证得 Hölder 不等式.

取满足 $\frac{1}{p} + \frac{1}{q} = 1$ 的 q, 有

$$\begin{split} \mathbb{E} \left(|X+Y|^p \right) = & \mathbb{E} \left(|X+Y||X+Y|^{p-1} \right) \\ \leq & \mathbb{E} \left(|X||X+Y|^{p-1} \right) + \mathbb{E} \left(|Y||X+Y|^{p-1} \right) \\ \leq & [\mathbb{E} \left(|X|^p \right)]^{1/p} \left[\mathbb{E} \left(|X+Y|^{(p-1)q} \right) \right]^{1/q} \\ & + \left[\mathbb{E} \left(|Y|^p \right) \right]^{1/p} \left[\mathbb{E} \left(|X+Y|^{(p-1)q} \right) \right]^{1/q} \\ = & \left\{ \left[\mathbb{E} \left(|X|^p \right) \right]^{1/p} + \left[\mathbb{E} \left(|Y|^p \right) \right]^{1/p} \right\} \left[\mathbb{E} \left(|X+Y|^p \right) \right]^{1/q} \end{split}$$

其中,第一个不等式用到了三角不等式 $|X\pm Y|\leq |X|+|Y|$,第二个不等式则是直接使用 Hölder 不等式. 最后两

边同除以 $[\mathbb{E}(|X+Y|^p)]^{1/q}$, 即得证.

笔记 Cauchy-Schwarz 不等式是 Hölder 不等式的特例:

$$(\mathbb{E}|XY|)^2 \le \mathbb{E}X^2\mathbb{E}Y^2$$

命题 1.2

随机变量 (向量)X 的 p 范数是一个范数 (norm).

证明 只需验证范数定义里的几个性质.

- 1. 非负性,显然满足;
- 2. 非退化性: $||X||_p = 0 \leftrightarrow X = 0$;
- 3. 齐次性: $\|\alpha X\|_p = |\alpha| \cdot \|X\|_p$;
- 4. 三角不等式:(Minkowski 不等式).

事实上, 可以进一步证明 $L^p(\Omega)$ 是一个 Banach 空间以及 $L^2(\Omega)$ 是 Hilbert 空间 (完备的、定义了内积的赋范 线性空间), 其内积定义为

$$(oldsymbol{X},oldsymbol{Y})_{L^2_{\omega}}=\mathbb{E}(oldsymbol{X},oldsymbol{Y})$$

定理 1.4 (L^p 是完备赋范线性空间)

 $(L^p, \|\cdot\|_p)$ 是完备赋范线性空间 (Banach Space). 具体地说, 假设 $(X_n, n \ge 1)$ 在 $\|\cdot\|_p$ 下是 Cauchy 序列, 即

$$\lim_{n \ge m \to \infty} \|X_n - X_m\|_p = 0,$$

那么一定存在一个随机变量 $X \in L^p$ 使得

$$\lim_{n \to \infty} \|X_n - X\|_p = 0$$

1.3 随机变量的独立性、联合分布

定义 1.10 (独立性)

两个随机变量 X and Y 被称为独立的, 如果对任意两个 Borel 集 A 和 B, 都有 $X^{-1}(A)$ 和 $Y^{-1}(B)$ 独立; 即:

$$\mathbb{P}\left(X^{-1}(A)\cap Y^{-1}(B)\right) = \mathbb{P}\left(X^{-1}(A)\right)\mathbb{P}\left(Y^{-1}(B)\right)$$

一族随机变量 X_1, \dots, X_n 被称为独立的, 如果对任意一族 Borel 集 $B_i \in \mathcal{R}$, 都有

$$\mathbb{P}\left(\bigcap_{j=1}^{n} X_{j}^{-1}\left(B_{j}\right)\right) = \prod_{j=1}^{n} \mathbb{P}\left(X_{j}^{-1}\left(B_{j}\right)\right)$$

笔记 令 $(\Omega, \mathcal{F}, \mathbb{P})$ 为一概率空间, \mathcal{G} 和 \mathcal{H} 为 \mathcal{F} 的子 σ -代数. 称两个子 σ - 代数独立, 如果:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$
 for all $A \in \mathcal{G}, B \in \mathcal{H}$.

若 X,Y 是概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 上的两个随机变量, 则显然, 诸二随机变量生成的 $\sigma(X)$ 和 $\sigma(Y)$ 如果独立, 就 有 X 和 Y 独立.

如果两个子事件域是独立的,则分别从二者里任意各自抽取的事件独立. 由定义可知, 若随机变量 X,Y 相互 独立, 当且仅当对任意的 $C, D \in \mathcal{B}(\mathbb{R})$, 有

$$\mathbb{P}(X \in C \mathbb{H} Y \in D) = \mathbb{P}(X \in C) \mathbb{P}(Y \in D)$$

定理 1.5

设 X,Y 是独立的随机变量, f,g 是 \mathbb{R} 上的 Borel 可测函数. 则 f(X) 和 g(Y) 也是独立的随机变量.

 \Diamond

证明 设 $A \in \sigma(f(X))$. 任意 $A \in \sigma(f(X))$ 都形如 $\{\omega \in \Omega : f(X(\omega)) \in C\}$, 其中 $C \in \mathcal{B}(\mathbb{R})$. 定义 $D = \{x \in \mathbb{R} : f(x) \in C\}$, 则有

$$A = \{\omega \in \Omega : f(X(\omega)) \in C\} = \{\omega \in \Omega : X(\omega) \in D\} \subset \sigma(X)$$

所以 $A \in \sigma(X)$. 同理设 $B \in \sigma(f(Y))$, 则可得 $B \in \sigma(Y)$. 由于 X, Y 独立, 则 $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. 由于 A, B 是 任意取的, 则说明 f(X) 和 g(Y) 也是独立的随机变量.

定义 1.11 (联合分布测度)

设X,Y是随机变量,随机向量(X,Y)在平面 \mathbb{R}^2 上取值,定义(X,Y)的联合分布测度为

$$\mu_{X,Y} = \mathbb{P}\{(X,Y) \in C\}, \forall C \in \mathcal{B}(\mathbb{R}^2)$$



Ŷ 笔记

- 1. 这是一个概率测度, 对 \mathbb{R}^2 上的每一个 Borel 子集都指定了一个 0 和 1 之间的数; 并且满足 $\mu_{X,Y}(\mathbb{R}^2)=1$ 以及可数可加性.
- 2. (X,Y) 也可定义联合累积分布函数

$$F_{X,Y}(x,y) = \mu_{X,Y}((-\infty, x] \times (-\infty, y]) = \mathbb{P}\{X \le x, Y \le y\}$$

非负 Borel 可测函数 $f_{X,Y}(x,y)$ 称为随机向量 (X,Y) 的**联合密度函数**, 如果:

$$\mu_{X,Y}(C) = \iint_{\mathbb{R}^2} \mathbb{I}_{(C)}(x,y) f_{X,Y}(x,y) dx dy, \forall C \in \mathcal{B}(\mathbb{R}^2)$$

且该条件等价于

$$F_{X,Y}(x,y) = \int_{-\infty}^{s} \int_{-\infty}^{t} f_{X,Y}(s,t) ds dt, \forall x \in \mathbb{R}, y \in \mathbb{R}$$

定义 1.12 (边际分布测度)

定义随机变量 X 和 Y 的边际分布测度为:

$$\mu_X(A) = \mathbb{P}(X \in A) = \mu_{X,Y}(A \times \mathbb{R}), \forall A \in \mathcal{B}(\mathbb{R})$$

$$\mu_Y(B) = \mathbb{P}(Y \in B) = \mu_{X,Y}(\mathbb{R} \times B), \forall B \in \mathcal{B}(\mathbb{R})$$

同样地,简化起见有边际分布函数:

$$F_X(a) = \mu_X(-\infty, a] = \mathbb{P}\{X \le a\} \forall a \in \mathbb{R},$$

$$F_Y(b) = \mu_Y(-\infty, b] = \mathbb{P}\{Y \le b\} \forall b \in \mathbb{R}.$$

如果联合密度 $f_{X,Y}$ 存在,则边际密度存在且由下式给出:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$$
$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx.$$

边际密度函数是非负 Borel 可测函数,满足:

$$\mu_X(A) = \int_A f_X(x) dx \forall A \in ; B(\mathbb{R}),$$

$$\mu_Y(B) = \int_B f_Y(y) dy \forall B \in ; B(\mathbb{R}).$$

这些条件成立当且仅当

$$F_X(a) = \int_{-\infty}^a f_X(x) dx \forall a \in \mathbb{R},$$

$$F_Y(b) = \int_{-\infty}^b f_Y(y) dy \forall b \in \mathbb{R}$$

定理 1.6 (独立性的等价条件)

设X,Y是独立的随机变量,则

1. (充要) 联合分布测度可以因子分解:

$$\mu_{X,Y}(A \times B) = \mu_X(A) \times \mu_Y(B), \forall A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{B}(\mathbb{R}).$$

- 2. (充要) 联合累积分布函数可以因子分解;
- 3. (充要) 联合密度可以因子分解:

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

4. (充要) 特征函数^a(或生成函数) 可以因子分解:

$$f_{X+Y}(\xi) = f_X(\xi)f_Y(\xi)$$

独立性和上述条件蕴含下面的条件但不等价:

5. 期望可以因子分解:

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

a见 1.6 节

Ç

1.4 条件数学期望

令 G 为一 F 的子 σ -代数. 随机变量 X 满足 $\mathbb{E}|X| < \infty$.

定义 1.13 (条件数学期望)

给定G的条件下,X的条件期望Z通过如下条件定义:

- 1. Z 是 G 可测的;
- 2. 对任意 $A \in \mathcal{G}$,

$$\int_{A} Z(\omega) d\mathbb{P} = \int_{A} X(\omega) d\mathbb{P}$$

称 $Z = \mathbb{E}(X|G)$ 为给定 G 的条件下,X 的条件期望.

🖹 笔记

- 1. 条件期望是一个随机变量.
- 2. 2 中

$$\int_A X d\mathbb{P} =: \mathbb{E}(X \cdot \mathbb{I}_A)$$

对于 $A \in \mathcal{G}$, 条件期望 $\mathbb{E}(X|A)$ 的定义事实上是

$$\mathbb{E}(X|A) = \int_{\Omega} X d\mathbb{P}(\cdot|A) = \int_{\Omega} X d\left(\frac{\mathbb{P}(\cdot \cap A)}{\mathbb{P}(A)}\right) = \frac{1}{\mathbb{P}(A)} \int_{A} X d\mathbb{P} = \frac{1}{\mathbb{P}(A)} \mathbb{E}(X \cdot \mathbb{I}_{A})$$

假设 $(A_m, m \ge 1)$ 是一列互不相交的事件, 且满足 $\mathbb{P}(A_m) > 0$, $\Omega = \sum_{m=1}^{\infty} A_m$. 令 $\mathcal{G} = \sigma\{A_m, m \ge 1\}$ 为由 $(A_m, m \ge 1)$ 所生成的最小 σ — 域. 定义 X 关于 \mathcal{G} 的条件期望为

$$\mathbb{E}(X|\mathcal{G}) = \sum_{m=1}^{\infty} \mathbb{E}(X|A_m) \mathbb{I}_{A_m}$$

换句话说, $\mathbb{E}(X|\mathcal{G})$ 是关于 σ — 域 \mathcal{G} 的可测随机变量, 且满足: 当 $\omega \in A_i$ 时,

$$\mathbb{E}(X|\mathcal{G})(\omega) = \mathbb{E}(X|A_i)$$

一般地, 假设 $G \subset F$ 是任意的子 σ — 域, 定义 $\mathbb{E}(X|G)$ 是关于 G 的可测随机变量, 并且当 $A \in G$ 时, 有

$$\int_{A} \mathbb{E}(X|\mathcal{G})(\omega) d\mathbb{P} = \int_{A} X(\omega) d\mathbb{P}$$

事实上就是在 A 上对 $\mathbb{E}(X|\mathcal{G})(\omega) = \mathbb{E}(X|A_i)$ 两边关于 ω 积分.

定理 1.7 (条件期望的性质)

假设 X, Y 满足 $\mathbb{E}|X|, \mathbb{E}|Y| < \infty$. 令 $a, b \in \mathbb{R}$. 我们有:

- 1. $\mathbb{E}(aX + bY \mid \mathcal{G}) = a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}).$
- 2. $\mathbb{E}(\mathbb{E}(X \mid \mathcal{G})) = \mathbb{E}(X)$.
- 3. $\mathbb{E}(X \mid \mathcal{G}) = X 若 X 是 \mathcal{G}$ -可测的.
- 5. $\mathbb{E}(XY \mid \mathcal{G}) = Y\mathbb{E}(X \mid \mathcal{G})$ 若 $Y \in \mathcal{G}$ -可测的.
- 6. 如果 \mathcal{H} 是 \mathcal{G} 的一个子 σ -代数, 则 $\mathbb{E}(X \mid \mathcal{H}) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{H}) \mid \mathcal{G})$.
- 7. (琴生不等式) ϕ : ℝ → ℝ 是一个凸函数且满足 $\mathbb{E}|\phi(X)|$ < ∞. 则有

$$\mathbb{E}(\phi(X) \mid \mathcal{G}) \ge \phi(\mathbb{E}(X \mid \mathcal{G}))$$

特别地,

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

下面我们考察给定另一个随机变量 Y 取值的条件下,X 的条件期望. 自然的想法就是, 给定 Y 取值的 X 的条件期望将由 $\mathcal{G} = \sigma(Y)$ 来定义:

$$\mathcal{G} = \sigma(Y) := \{Y^{-1}(B), B \in \mathcal{R}\}.$$

"给定 Y 的取值"诱导我们令 $\Omega_i = \{\omega : Y(\omega) = j\}$ 并且

$$\Omega = \bigcup_{j=1}^{n} \Omega_j$$

则 σ -代数 $\mathcal G$ 就是 Ω_j 所有可能的并集. 由于 $\mathbb E(X|Y)$ 在每个 Ω_i 上都是常数, 我们就可以通过定义

$$\int_{\Omega_j} \mathbb{E}(X|Y)(\omega) d\mathbb{P} = \int_{\Omega_j} X(\omega) d\mathbb{P}$$

推出

$$\mathbb{E}(X|Y) = \frac{1}{\mathbb{P}\left(\Omega_j\right)} \int_{\Omega_j} X(\omega) d\mathbb{P}$$

$$\mathbb{E}(X|Y) = \lim_{h \to 0} \frac{\displaystyle \int_G X(\omega) d\mathbb{P}}{\mathbb{P}(G)} = \lim_{h \to 0} \frac{\displaystyle \int \left(\frac{1}{h} \int_y^{y+h} x p(x,y) dy\right) dx}{\displaystyle \frac{1}{h} \int_y^{y+h} p_Y(y) dy} \xrightarrow{\frac{\text{$\sharp x \ni \psi \text{ if $\sharp \in \mathbb{Z}$}}}{}} \frac{\displaystyle \int x p(x,y) dx}{\displaystyle p_Y(y)} = \int x \frac{p(x,y)}{\displaystyle p_Y(y)} dx$$

这说明基于测度论的条件期望理论与初等概率论中所讲的如出一辙.

最后我们给出条件期望的另一个重要性质: 均方误差意义下的最优估计.

命题 1.3

令 q 为一可测函数. 于是

$$\mathbb{E}(X - \mathbb{E}(X|Y))^2 \le \mathbb{E}(X - g(Y))^2.$$

1.5 收敛的概念

设 $\{X_n\}_{n=1}^{\infty}$ 为定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量序列, 设 μ_n 为 X_n 的分布. 设 X 为分布为 μ 的随机变量. 我们将讨论收敛的四个概念:几乎确定收敛、按概率收敛、按分布收敛和按 L^p 收敛.

定义 1.14 (收敛的概念)

1. 依概率 1 收敛 (Almost sure convergence). X_n 依概率 1 收敛到 X, 如果

$$\mathbb{P}\left(\omega: X_n(\omega) \to X(\omega)\right) = 1.$$

记作 $X_n \to X$, a.s.

2. 依概率收敛 (Convergence in probability). X_n 依概率收敛到 X, 如果对任意 $\varepsilon > 0$, 有

$$\mathbb{P}(\omega:|X_n(\omega)-X(\omega)|>\varepsilon)\to 0,\quad n\to\infty$$

记作 $X_n \stackrel{P}{\to} X$.

- 3. 依分布收敛 (Convergence in distribution). 设 F 是一分布函数, $\{F_n, n \geq 1\}$ 是一列分布函数, 如果对 F 的每个连续点 $x \in \mathbb{R}$, 都有 $F_n(x) \to F(x)(n \to \infty)$, 则称 F_n 弱收敛 (weak convergence) 于 F ,记 作 $F_n \stackrel{w}{\to} F$. 设 X 为一随机变量, $\{X_n, n \geq 1\}$ 是一列随机变量, 如果 X_n 的分布函数弱收敛于 X 的分布函数,则称 X_n 依分布收敛于 X,记作 $X_n \stackrel{d}{\to} X$.
- 4. 矩收敛 (Convergence in L^p). X_n 以 L^p 收敛到 X (0 如果

$$\mathbb{E}|X_n - X|^p \to 0$$

若p=1, 就称为以均值收敛; 若p=2, 就称为均方收敛.



笔记

1. 如果存在一个零概率事件 (零测度集) Ω_0 , 使得对任意 $\omega \in \Omega - \Omega_0$, 当 $n \to \infty$ 时,

$$X_n(\omega) \to X(\omega)$$

则称 X_n 几乎处处收敛 到 X.

这是最强的收敛性,是函数逐点收敛的推广形式:

$$\mathbb{P}\left(\lim_{n\to\infty}X_n(\omega_i)=X(\omega_i)\right)=1$$

排除一个"不收敛的可数个少到几乎没有的确定的点名单", 存在 N, 当 n>N 时, X_n 和 X 的差不超过 ε 对其它所有 ω_i 都成立.

- 2. 依概率收敛是比几乎处处收敛弱的收敛性. 不管某一个确定的点是不是收敛; 只需要最后在大局上不收敛的点数几乎没有.
- 3. 分布函数的弱极限未必是分布函数. 例如:

$$F_n(x) = \begin{cases} 0, & x \le -n, \\ \frac{x+n}{2n}, & -n < x \le n, \\ 1, & x > n, \end{cases}$$
$$F(x) = \frac{1}{2}.$$

如果分布函数的弱极限确实是分布函数,则可以定义依分布收敛.

接下来的一些定理将给出收敛性的判别方法.

定理 1.8 (Helly 第一定理 (Helly's selection theorem))

设 $\{F_n, n \ge 1\}$ 是一列定义在 \mathbb{R} 上的分布函数,则存在一列子列 $\{F_{n(k)}, k \ge 1\}$,它有弱极限 F,F 右连续,非降.



证明 采用对角线方法证明该定理. 由于有理数集可数,不妨记 $\mathbb{Q} = \{q_1, q_2, \cdots\}$. 下面取 $m_0 = \mathbb{N}^+, m_1$ 为 m_0 的 一个子列, m_2 为 m_1 的一个子列 · · · 以此类推. 由于对任意的正整数 k 和 m, 有 $F_m(q_k) \in [0,1]$. 故对任意的 k, $\{F_{m_k}(q_k)\}$ 必有收敛子列 $\{F_{m_k(i)}(q_k)\}$.

记 $G(q_k) = \lim_{k \to \infty} F_{m_k(i)}(q_k)$, 此时根据构造, 对所有有理数 q 都有 $F_{n(k)}(q) \to G(q)$. 函数 G 可能不是右连续 的, 但 $F(x) = \inf\{G(q): q \in \mathbb{Q}, q > x\}$ 是右连续的, 因为

$$\lim_{x_n \downarrow x} F(x_n) = \inf \left\{ G(q) : q \in \mathbb{Q}, q > x_n \text{ for some } n \right\}$$

$$=\inf\{G(q):q\in\mathbb{Q},q>x\}=F(x)$$

为了完成证明, 令 x 作为 F 的一个连续点. 取有理数 $r_1, r_2, s(r_1 < r_2 < x < s)$, 使得

$$F(x) - \varepsilon < F(r_1) \le F(r_2) \le F(x) \le F(s) < F(x) + \varepsilon$$

由于 $F_{n(k)}(r_2) \to G(r_2) \ge F(r_1)$ 以及 $F_{n(k)}(s) \to G(s) \le F(s)$, 因此, 如果 k 很大, 就有

$$F(x) - \varepsilon < F_{n(k)}(r_2) \le F_{n(k)}(x) \le F_{n(k)}(s) < F(x) + \varepsilon$$

这就是我们想要的结论.

定理 1.9 (Helly 第二定理)

(1) 设 F 是一分布函数, $\{F_n, n \ge 1\}$ 是一列分布函数, $F_n \stackrel{w}{\to} F$. 如果 g(x) 是 \mathbb{R} 上的有界连续函数, 则

$$\int_{-\infty}^{\infty} g(x)dF_n(x) \to \int_{-\infty}^{\infty} g(x)dF(x).$$

(2) 设 F, F_n 是单调不减右连续函数 (不一定是分布函数), 并且对 F 的任一连续点 x 有 $F_n(x) \to F(x)$. 如 果 a < b 是 F 的连续点, g(x) 是 [a,b] 上的连续函数, 则

$$\int_{a}^{b} g(x)dF_{n}(x) \to \int_{a}^{b} g(x)dF(x).$$

证明 我们只证第一部分结论,对于第二部分,只要重新如下定义 F,F_n 和g就可以归到

$$F^*(x) = \begin{cases} 0, & x < a, \\ \frac{F(x) - F(a)}{F(b) - F(a)}, & a \le x < b, \quad F_n^*(x) = \begin{cases} 0, & x < a, \\ \frac{F_n(x) - F_n(a)}{F_n(b) - F_n(a)}, & a \le x < b, \\ 1, & x \ge b; \end{cases}$$

$$g^*(x) = \begin{cases} g(a), & x < a; \\ g^*(x) = g(x), & a \le x < b; \\ g^*(x) = g(b), & x \ge b. \end{cases}$$

$$g^*(x) = \begin{cases} g(a), & x < a; \\ g^*(x) = g(x), & a \le x < b; \\ g^*(x) = g(b), & x \ge b. \end{cases}$$

接下来证明 (1). 因 g 是有界函数, 必存在 c>0, 使得 $|g(x)|< c, x\in\mathbb{R}$. 因 F 的所有连续点构成 \mathbb{R} 上的稠密集, 又 $F(-\infty) = 0, F(\infty) = 1$, 故对任意给定的 $\varepsilon > 0$, 可以选取 a > 0, 使得 $\pm a$ 是 F 的连续点, 并且

$$F(-a) < \frac{\varepsilon}{12c}, \quad 1 - F(a) < \frac{\varepsilon}{12c}.$$

由于 $F_n \xrightarrow{w} F$, 存在 $N_1(\varepsilon)$, 使得当 $n \geq N_1(\varepsilon)$ 时,

$$|F_n(-a) - F(-a)| < \frac{\varepsilon}{12c}, \quad |1 - F_n(a) - (1 - F(a))| < \frac{\varepsilon}{12c}.$$

这样我们有

$$\left| \int_{-\infty}^{-a} g(x) dF_n(x) - \int_{-\infty}^{-a} g(x) dF(x) + \int_{a}^{\infty} g(x) dF_n(x) - \int_{a}^{\infty} g(x) dF(x) \right|$$

$$\leq c \left[F_n(-a) + F(-a) + 1 - F_n(a) + 1 - F(a) \right]$$

$$\leq c \left[|F_n(-a) - F(-a)| + 2F(-a) + |1 - F_n(a) - (1 - F(a))| + 2(1 - F(a)) \right]$$

$$< \frac{\varepsilon}{2}.$$

下面考虑 $\left| \int_{-a}^{a} g(x) dF_n(x) - \int_{-a}^{a} g(x) dF(x) \right|$. 由于 g(x) 在闭区间 [-a,a] 上一致连续, 可以选取 $-a = x_0 < x_1 < \cdots < x_m = a$ 使得所有 x_i 是 F 的连续点, 且 $\max_{x_{i-1} < x < x_i} |g(x) - g(x_i)| < \varepsilon/8$. 于是

$$\left| \int_{-a}^{a} g(x) dF_{n}(x) - \int_{-a}^{a} g(x) dF(x) \right|$$

$$= \left| \sum_{i=1}^{m} \int_{x_{i-1}}^{x_{i}} g(x) dF_{n}(x) - \sum_{i=1}^{m} \int_{x_{i-1}}^{x_{i}} g(x) dF(x) \right|$$

$$\leq \sum_{i=1}^{m} \int_{x_{i-1}}^{x_{i}} |g(x) - g(x_{i})| dF_{n}(x) + \sum_{i=1}^{m} \int_{x_{i-1}}^{x_{i}} |g(x) - g(x_{i})| dF(x) + \sum_{i=1}^{m} |g(x_{i})| \left| \int_{x_{i-1}}^{x_{i}} dF_{n}(x) - \int_{x_{i-1}}^{x_{i}} dF(x) \right|$$

$$\leq \frac{\varepsilon}{8} \sum_{i=1}^{m} \left\{ F_{n}(x_{i}) - F_{n}(x_{i-1}) + F(x_{i}) - F(x_{i-1}) \right\} + 2c \sum_{i=0}^{m} |F_{n}(x_{i}) - F(x_{i})|$$

$$= \frac{\varepsilon}{8} \left(F_{n}(a) - F_{n}(-a) + F(a) - F(-a) \right) + 2c \sum_{i=0}^{m} |F_{n}(x_{i}) - F(x_{i})| \quad (*).$$

由于 $F_n(a) - F_n(-a) \le 1$, $F(a) - F(-a) \le 1$, 再选择 $N_2(\varepsilon)$ 使得当 $n \ge N_2(\varepsilon)$ 时,

$$|F_n(x_i) - F(x_i)| < \frac{\varepsilon}{8mc}, \quad i = 0, 1, \dots, m.$$

故 (*) 不超过 $\varepsilon/2$. 因此, 当 $n \ge \max(N_1(\varepsilon), N_2(\varepsilon))$ 时,

$$\left| \int_{-\infty}^{\infty} g(x) dF_n(x) - \int_{-\infty}^{\infty} g(x) dF(x) \right| < \varepsilon.$$

定理证毕.

定义 1.15 (Infinitely Often)

给定事件序列 $\{A_n\}_{n=1}^{\infty}$,按照如下定义一个事件:

$$\{A_n \text{ i.o. }\} = \{\omega : \omega \in A_n \text{ i.o. }\} = \limsup_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

 $\stackrel{ extbf{Q}}{ extbf{Q}}$ 笔记 实分析中集合序列的上极限与下极限:对于一个集合序列 $\{A_n\}_{n=1}^\infty$,我们定义它的上极限与下极限分别为

$$\limsup_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k, \quad \liminf_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

分别称其为 $\{A_n\}$ 的上极限和下极限. 显然有

$$\limsup_{n\to\infty} A_n = \left\{ \omega \mid \omega \text{ 属于无穷多个} A_n \right\},$$

$$\liminf_{n\to\infty} A_n = \left\{ \omega \mid \omega \text{ 至多不属于有限多个} A_n \right\},$$

用符号表述也就是

$$\limsup_{n \to \infty} A_n = \{ \omega \mid \forall N \in \mathbb{N}, \exists n > N, \omega \in A_n \}$$
$$\liminf_{n \to \infty} A_n = \{ \omega \mid \exists N_\omega \in \mathbb{N}, \forall n > N_\omega, \omega \in A_n \}$$

在概率中, 所谓的 A_n i.o.(infinitely often) 就是表示集合序列 $\{A_n\}$ 的上极限 $\limsup_{n \to \infty} A_n$.

引理 1.1 (Borel-Cantelli 定理)

- 1. $\not\equiv \sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, $\mathbb{M} \mathbb{P}(\{A_n \text{ i.o. }\}) = 0$.
- 2. 若 $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ 以及 A_n 相互独立, 则 $\mathbb{P}(\{A_n \text{ i.o. }\}) = 1$.

证明 (1) 对任意 n, 我们有

$$\mathbb{P}\left(\left\{\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_{k}\right\}\right)\leq\mathbb{P}\left(\left\{\bigcup_{k=n}^{\infty}A_{k}\right\}\right)\leq\sum_{k=n}^{\infty}\mathbb{P}\left(A_{k}\right)$$

最右边一项趋于 0 因为 $\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty$.

(2) 利用独立性,有

$$\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) = 1 - \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k^c\right) = 1 - \prod_{k=n}^{\infty} \mathbb{P}\left(A_k^c\right) = 1 - \prod_{k=n}^{\infty} \left(1 - \mathbb{P}\left(A_k\right)\right).$$

利用基本不等式 $1-x \le e^{-x}$, 有

$$\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) \ge 1 - \prod_{k=n}^{\infty} e^{-\mathbb{P}(A_k)} = 1 - e^{-\sum_{k=n}^{\infty} \mathbb{P}(A_k)} = 1.$$

证毕.

定理 1.10 (依概率 1 收敛的判别准则)

设 X 和 $\{X_n, n \ge 1\}$ 是定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量和随机变量序列. $X_n \to X$ a.s. 当且仅当对任意的 $\varepsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}\left(\bigcup_{k \ge n} \{|X_k - X| \ge \varepsilon\}\right) = 0,$$

或者等价地

$$\lim_{n \to \infty} \mathbb{P}\left(\sup_{k \ge n} |X_k - X| \ge \varepsilon\right) = 0.$$

 \odot

证明 记事件 $A_n^{\varepsilon} = \{|X_n - X| \ge \varepsilon\}, \, \text{则 } A^{\varepsilon} = \{A_n^{\varepsilon} \text{ i.o.}\} = \bigcap_{n=1}^{\infty} \bigcup_{k > n} A_k^{\varepsilon}. \, \text{则可知}$

$${X_n \nrightarrow X} = \bigcup_{m=1}^{\infty} A^{1/m}.$$

由概率的连续性得

$$\mathbb{P}\left(A^{\varepsilon}\right)=\lim_{n\rightarrow\infty}\mathbb{P}\left(\bigcup_{k\geq n}A_{k}^{\varepsilon}\right).$$

则下列关系式成立:

$$0 = \mathbb{P}(X_n \nrightarrow X) \Leftrightarrow \mathbb{P}\left(\bigcup_{m=1}^{\infty} A^{1/m}\right) = 0$$

$$\Leftrightarrow \mathbb{P}\left(A^{1/m}\right) = 0, \forall m \ge 1$$

$$\Leftrightarrow \mathbb{P}\left(\bigcup_{k \ge n} A_k^{1/m}\right) \to 0, \forall m \ge 1$$

$$\Leftrightarrow \mathbb{P}\left(\bigcup_{k \ge n} \left\{|X_k - X| \ge 1/m\right\}\right) \to 0, \quad \forall m \ge 1$$

$$\Leftrightarrow \mathbb{P}\left(\bigcup_{k \ge n} \left\{|X_k - X| \ge \varepsilon\right\}\right) \to 0, \quad \forall \varepsilon > 0.$$

又

$$\mathbb{P}\left(\sup_{k\geq n}|X_k-X|\geq\varepsilon\right)=\mathbb{P}\left(\bigcup_{k\geq n}\left\{|X_k-X|\geq\varepsilon\right\}\right),$$

得证.

由上述定理和 Borel-Cantelli 定理可以立即得到如下判别定理:

推论 1.1

如果对任意的 $\varepsilon>0,\sum_{n=1}^{\infty}\mathbb{P}\left(|X_n-X|\geq\varepsilon\right)<\infty$, 则 $X_n\to X$ a.s.

C

下面的定理揭示几种概率收敛之间的关系.

定理 1.11

- 1. 几乎处处收敛蕴含依概率收敛.
- 2. 依概率收敛可以推出存在一个子序列是几乎处处收敛的.
- 3. 若p < q, 则依 L^q 收敛可以推出依 L^p 收敛.
- 4. 依 Lp 收敛蕴含依概率收敛.
- 5. 依概率收敛蕴含依分布收敛.
- 6. 依分布收敛于一个常数可以推出也依概率分布到此常数.

证明 (i) 若 $X_n \to X$ a.s., 则

$$\lim_{n \to \infty} \mathbb{P}\left(\sup_{k \ge n} |X_k - X| \ge \varepsilon\right) = 0.$$

所以

$$\lim_{n \to \infty} \mathbb{P}\left(|X_n - X| \ge \varepsilon\right) \le \lim_{n \to \infty} \mathbb{P}\left(\sup_{k > n} |X_k - X| \ge \varepsilon\right) = 0,$$

 $\mathbb{P} X_n \stackrel{P}{\to} X$.

(ii) 不是一般性, 假设 X=0. 根据依概率收敛的条件我们可以取随 k 递增的 $n_k\in\mathbb{N}$ 使得对任意 $k\in\mathbb{N}$ 有:

$$\mathbb{P}\left(|X_{n_k}| \ge \frac{1}{k}\right) \le \frac{1}{2^k}$$

对任意固定的 ε , 存在正整数 k_0 使得 $k_0^{-1} \le \varepsilon$, 我们就有

$$\sum_{k=1}^{\infty} \mathbb{P}\left(|X_{n_k}| \ge \varepsilon\right) \le \sum_{k=1}^{k_0} \mathbb{P}\left(|X_{n_k}| \ge \varepsilon\right) + \sum_{k=k_0+1}^{\infty} \mathbb{P}\left(|X_{n_k}| \ge \frac{1}{k}\right)$$

$$\le \sum_{k=1}^{k_0} \mathbb{P}\left(|X_{n_k}| \ge \varepsilon\right) + \sum_{k=k_0+1}^{\infty} \frac{1}{2^k} < \infty$$

根据 Borel-Cantelli 定理的推论得到 $X_{n_k} \rightarrow 0$ a.s..

(iii) 根据 Hölder 不等式 $(\frac{p}{q} + \frac{q-p}{p} = 1)$:

$$\mathbb{E}\left|X_{n}-X\right|^{p}\cdot 1 \leq \left(\mathbb{E}\left(\left|X_{n}-X\right|^{p}\right)^{\frac{q}{p}}\right)^{\frac{p}{q}} = \left(\mathbb{E}\left|X_{n}-X\right|^{q}\right)^{\frac{p}{q}}, \quad p < q.$$

(iv) 由 Chebyshev 不等式可以立即得到对任意 $\varepsilon > 0$ 有:

$$\mathbb{P}\left(\omega: |X_n(\omega) - X(\omega)| > \varepsilon\right) \le \frac{1}{\varepsilon^p} \mathbb{E}\left|X_n - X\right|^p$$

(v) 设 F 和 F_n 分别是 X 和 X_n 的分布函数, x 为 F 的连续点. 任意给定 $\varepsilon > 0$, 有

$$\{X \le x - \varepsilon\} = \{X \le x - \varepsilon, X_n \le x\} \cup \{X \le x - \varepsilon, X_n > x\}$$
$$\subset \{X_n \le x\} \cup \{X_n - X \ge \varepsilon\},$$

所以

$$F(x-\varepsilon) < F_n(x) + \mathbb{P}(X_n - X > \varepsilon)$$
.

因为 $X_n \stackrel{P}{\to} X$. 所以

$$\mathbb{P}(X_n - X \ge \varepsilon) \le \mathbb{P}(|X_n - X| \ge \varepsilon) \to 0,$$

从而

$$F(x-\varepsilon) \le \liminf_{n \to \infty} F_n(x).$$

类似地,

$$\{X_n \le x\} = \{X_n \le x, X \le x + \varepsilon\} \cup \{X_n \le x, X > x + \varepsilon\}$$
$$\subset \{X \le x + \varepsilon\} \cup \{X - X_n \ge \varepsilon\},$$

从而

$$F_n(x) \le F(x+\varepsilon) + \mathbb{P}(X - X_n \ge \varepsilon)$$
.

因此

$$\limsup_{n \to \infty} F_n(x) \le F(x + \varepsilon).$$

结合上述可知对任意的 $\varepsilon > 0$,

$$F(x-\varepsilon) \le \liminf_{n \to \infty} F_n(x) \le \limsup_{n \to \infty} F_n(x) \le F(x+\varepsilon)$$

因此有 $\lim_{x\to\infty} F_n(x) = F(x)$, 即 X_n 依分布收敛到 X.

(vi) 如果 $X_n \stackrel{d}{\rightarrow} c$, 则当 $x \neq c$ 时

$$\lim_{n \to \infty} F_n(x) = \begin{cases} 0, & x < c, \\ 1, & x > c. \end{cases}$$

因此对任意的 $\varepsilon > 0$ 有

$$\mathbb{P}(|X_n - c| \ge \varepsilon) = \mathbb{P}(X_n \ge c + \varepsilon) + \mathbb{P}(X_n \le c - \varepsilon)$$
$$= 1 - \mathbb{P}(X_n < c + \varepsilon) + \mathbb{P}(X_n \le c - \varepsilon)$$
$$= 1 - F_n(c + \varepsilon - 0) + F_n(c - \varepsilon) \to 0.$$

1.6 特征函数

定义 1.16 (特征函数)

随机变量 X 的特征函数如下定义:

$$f(t) = \mathbb{E} e^{itX} = \int_{\mathbb{R}} e^{itx} d\mu \xrightarrow{\underline{\text{i.i.g.}}} \int_{\mathbb{R}} p(x) \cdot e^{itx} dx.$$

命题 1.4

特征函数具有如下基本性质:

- 1. $\forall t \in \mathbb{R}, |f(t)| \le 1, f(t) = \overline{f(-t)}, f(0) = 1;$
- 2. f 在 \mathbb{R} 上一致连续.

证明 证明第二部分:

$$|f(t_1) - f(t_2)| = \left| \mathbb{E} \left(e^{it_1 X} - e^{it_2 X} \right) \right| = \left| \mathbb{E} \left(e^{it_1 X} \left(1 - e^{i(t_2 - t_1)X} \right) \right) \right|$$

$$\leq \mathbb{E} \left| 1 - e^{i(t_2 - t_1)X} \right|.$$

因为 $|1-e^{i(t_2-t_1)X}|$ 只由 t_1 和 t_2 之间的差决定,并且当差趋于 0 时这部分也趋于 0,则根据控制收敛定理可以立即得到一致连续性.

例题 1.3 下面是一些典型分布的特征函数:

1. Bernoulli 分布.

$$f(t) = q + pe^{it}$$

2. 二项分布 B(n, p).

$$f(t) = \left(q + pe^{it}\right)^n.$$

3. 泊松分布 $\mathcal{P}(\lambda)$.

$$f(t) = e^{\lambda \left(e^{it} - 1\right)}.$$

4. 指数分布 $\mathcal{E}(\lambda)$.

$$f(t) = (1 - \lambda^{-1}it)^{-1}$$
.

5. 正态分布 $N(\mu, \sigma^2)$.

$$f(t) = \exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right).$$

6. 多元正态分布 $N(\mu, \Sigma)$.

$$f(\boldsymbol{t}) = \exp\left(i\boldsymbol{\mu}^T\boldsymbol{t} - \frac{1}{2}\boldsymbol{t}^T\boldsymbol{\Sigma}\boldsymbol{t}\right).$$

接下来是一个核心定理, 诠释依分布收敛中最重要的结论.

定理 1.12 (Levy 连续定理)

设 F 是一分布函数, $\{F_n, n \ge 1\}$ 是一列分布函数. 如果 $F_n \stackrel{w}{\to} F$. 则相应的特征函数 $\{f_n(t), n \ge 1\}$ 收敛于 F 的特征函数 f(t), 且在 t 的任一有限区间内收敛是一致的.

证明 在 Helly 第二定理中取 $g_t(x) = e^{itx}$ 证得收敛性. 任意有限区间内的一致收敛性的证明类似 Helly 第二定理的证明.

下面给出其余重点性质:

定义 1.17 (半正定函数)

函数 f 称为半正定函数, 如果对任意有限集 $\{t_1,\cdots,t_n\}$ $n\in\mathbb{N}$, 矩阵 $(f(t_i-t_j))_{i,j=1}^n$

$$M = \begin{pmatrix} f(t_1 - t_1) & f(t_1 - t_2) & \cdots & f(t_1 - t_n) \\ f(t_2 - t_1) & f(t_2 - t_2) & \cdots & f(t_2 - t_n) \\ \vdots & \vdots & \ddots & \vdots \\ f(t_n - t_1) & f(t_n - t_2) & \cdots & f(t_n - t_n) \end{pmatrix}$$

是半正定的, 即对任意复数 $v_1, v_2, \cdots, v_n \in \mathbb{C}$:

$$\sum_{i,j} f(t_i - t_j) v_i \bar{v}_j \ge 0$$

定理 1.13 (Bochner's theorem)

函数 f 是一个随机变量的特征函数, 当且仅当它是半正定连续函数且 f(0) = 1.

证明 仅给出必要性的证明.

$$\sum_{k=1}^{n} \sum_{j=1}^{n} f(t_k - t_j) c_k \overline{c_j} = \sum_{k=1}^{n} \sum_{j=1}^{n} c_k \overline{c_j} \int_{-\infty}^{\infty} e^{i(t_k - t_j)x} p(x) dx$$

$$= \int_{-\infty}^{\infty} \sum_{k=1}^{n} \sum_{j=1}^{n} c_k \overline{c_j} e^{i(t_k - t_j)x} p(x) dx$$

$$= \int_{-\infty}^{\infty} \left(\sum_{k=1}^{n} c_k e^{it_k x} \right) \left(\sum_{j=1}^{n} \overline{c_j} e^{-it_j x} \right) p(x) dx$$

$$= \int_{-\infty}^{\infty} \left| \sum_{k=1}^{n} c_k e^{it_k x} \right|^2 p(x) dx \ge 0$$

定理 1.14 (特征函数与各阶矩的关系)

若 $\mathbb{E}(X^l)$ 存在,则 X 的特征函数 f(t) 可 l 次求导,且对 $1 \le k \le l$,有

$$f^{(k)}(0) = i^k \mathbb{E}\left(X^k\right)$$

如果特征函数是 f(t), 则可以展开如下:

$$f(t) = 1 + (it)\mathbb{E}X + \frac{(it)^2}{2!}\mathbb{E}(X^2) + \dots + \frac{(it)^n}{n!}\mathbb{E}(X^n) + o(t^n)$$

证明 由于

$$\int_{-\infty}^{\infty} \left| \frac{d^k}{dt^k} e^{itx} \right| dF(x) = \int_{-\infty}^{\infty} \left| i^k x^k e^{itx} \right| dF(x)$$
$$= \int_{-\infty}^{\infty} \left| x^k \right| dF(x) = \mathbb{E} \left| X^k \right| < \infty,$$

因此 $\int_{-\infty}^{\infty} \frac{d^k}{dt^k} e^{itx} dF(x)$ 对 t 一致收敛, 故 $f^{(k)}(t)$ 存在, 且可以交换求导和积分的顺序

$$f^{(k)}(t) = \frac{d^k}{dt^k} \int_{-\infty}^{\infty} e^{itx} dF(x)$$
$$= \int_{-\infty}^{\infty} \frac{d^k}{dt^k} e^{itx} dF(x)$$
$$= i^k \int_{-\infty}^{\infty} x^k e^{itx} dF(x).$$

所以有

$$f^{(k)}(0) = i^k \int_{-\infty}^{\infty} x^k dF(x) = i^k \mathbb{E} X^k$$

命题 1.5

设 Z = aX + b, 其中 a, b 为任意常数, 则

$$f_Z(t) = e^{ibt} f_X(at)$$

室记此性质显而易见;利用此性质可以计算出一些非标准分布的特征函数.

根据特征函数的定义我们知道,随机变量的分布函数唯一确定它的特征函数,反之,特征函数是否可以唯一确定相应的分布函数?接下来将讨论这个问题.

引理 1.2 (狄利克雷积分)

给出一个分析的结果: 有

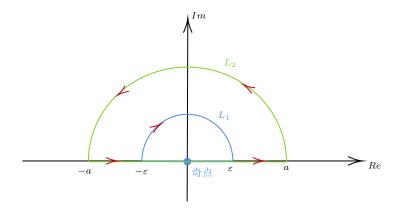
$$\lim_{x \to \infty} \int_{0}^{x} \frac{\sin au}{u} du = \frac{\pi}{2} \operatorname{sgn}\{a\},\,$$

其中 $sgn{a}$ 是 a 的符号函数.

证明 设积分值为 I, 根据对称性有 $2I = \int_{-\infty}^{+\infty} \frac{\sin ax}{x} dx$. 研究函数 $f(z) = \frac{e^{iaz}}{z}$, a > 0, 此函数在实轴上有孤立奇点 z = 0. 根据 Cauchy 积分定理,有 $\int_{-R}^{-\varepsilon} \frac{e^{iax}}{x} dx + \int_{L_1} f(z) dz + \int_{\varepsilon}^{R} \frac{e^{iax}}{x} dx + \int_{L_2} f(z) dz = 0$, 令 $R \to \infty$, $\varepsilon \to 0$, 有

$$\int_{-\infty}^{0} \frac{e^{iax}}{x} dx + \int_{L_1} f(z) dz + \int_{0}^{+\infty} \frac{e^{iax}}{x} dx + \int_{L_2} f(z) dz = 0$$

又在周线 L_2 上, 当 $R \to \infty$ 时有



根据留数定理,

$$\int_{L_1} f(z)dz = i(0 - \pi) \lim_{z \to 0} z f(z) = -i\pi$$

所以

$$\int_{\mathbb{R}} \frac{e^{iax}}{x} dx = \int_{\mathbb{R}} \frac{\cos ax + i \sin ax}{x} dx = i\pi$$

对应虚部得

$$2I = \pi, \quad I = \frac{\pi}{2}$$

若 a < 0, 则 L_1, L_2 在下半平面构造. 最终可得

$$I = \frac{\pi}{2}\operatorname{sgn}\{a\}.$$

引理 1.3

设 $x_1 < x_2$,

$$g(T, x, x_1, x_2) = \frac{1}{\pi} \int_0^T \left[\frac{\sin t (x - x_1)}{t} - \frac{\sin t (x - x_2)}{t} \right] dt$$

则

$$\lim_{T \to \infty} g(T, x, x_1, x_2) \triangleq h(x, x_1, x_2)$$

$$= \begin{cases} 0, & x < x_1 \ & \exists x > x_2, \\ \frac{1}{2}, & x = x_1 \ & \exists x = x_2, \\ 1, & x_1 < x < x_2. \end{cases}$$

证明 证: 由狄利克雷积分知

$$D(a) = \frac{1}{\pi} \lim_{T \to \infty} \int_0^T \frac{\sin at}{t} dt = \frac{1}{2} \operatorname{sgn}\{a\} = \begin{cases} 1/2, & a > 0 \\ 0, & a = 0 \\ -1/2, & a < 0 \end{cases}$$

而

$$\lim_{T \to \infty} g(T, x, x_1, x_2) = D(x - x_1) - D(x - x_2)$$
$$= \frac{1}{2} \left[\operatorname{sgn} \left\{ x - x_1 \right\} - \operatorname{sgn} \left\{ x - x_2 \right\} \right],$$

易知引理成立.

定理 1.15 (逆转公式)

设分布函数 F(x) 的特征函数为 f(t), 又 x_1, x_2 是 F(x) 的两个连续点, 则

$$F(x_2) - F(x_1) = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx_1} - e^{-itx_2}}{it} f(t) dt.$$

证明 不妨设 $x_1 < x_2$. 记

$$I_{T} = \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx_{1}} - e^{-itx_{2}}}{it} f(t)dt$$
$$= \frac{1}{2\pi} \int_{-T}^{T} \int_{-\infty}^{\infty} \frac{e^{-itx_{1}} - e^{-itx_{2}}}{it} e^{itx} dF(x)dt.$$

为证被积函数的有界性,用到不等式

$$\left| e^{i\alpha} - 1 \right| \le |\alpha|.$$

事实上, 对 $\alpha \ge 0$

$$\left|e^{i\alpha}-1\right| = \left|\int_0^\alpha e^{ix}dx\right| \le \int_0^\alpha \left|e^{ix}\right|dx = \alpha,$$

对 α <0,取共轭即知不等式也成立.因此,

$$\left| \frac{e^{-itx_1} - e^{-itx_2}}{it} e^{itx} \right| \le x_2 - x_1.$$

交换上述二次积分顺序得到

$$\begin{split} I_T &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-T}^{T} \frac{e^{-itx_1} - e^{-itx_2}}{it} e^{itx} dt \right] dF(x) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{0}^{T} \frac{e^{it(x-x_1)} - e^{-it(x-x_1)}}{it} - \frac{e^{it(x-x_2)} - e^{-it(x-x_2)}}{it} dt \right] dF(x) \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \left[\int_{0}^{T} \left(\frac{\sin t (x - x_1)}{t} - \frac{\sin t (x - x_2)}{t} \right) dt \right] dF(x) \\ &= \int_{-\infty}^{\infty} g (T, x, x_1, x_2) dF(x) \\ &= \mathbb{E}g (T, X, x_1, x_2) \,. \end{split}$$

对任意的 $x \in \mathbb{R}$, $|g(T, x, x_1, x_2)|$ 有界. 因此由 Lebesgue 控制收敛定理并利用引理的结果可得 (同时注意到 x_1 和 x_2 为 F(x) 的连续点):

$$\lim_{T \to \infty} I_T = \lim_{T \to \infty} \mathbb{E}g (T, X, x_1, x_2)$$

$$= \mathbb{E} \lim_{T \to \infty} g (T, X, x_1, x_2)$$

$$= \mathbb{E}h (X, x_1, x_2)$$

$$= 0 \cdot \mathbb{P} (X < x_1 \text{ or } X > x_2) + \frac{1}{2} \cdot \mathbb{P} (X = x_1 \text{ or } X = x_2) + 1 \cdot \mathbb{P} (x_1 < X < x_2)$$

$$= \mathbb{P} (x_1 < X < x_2)$$

$$= F (x_2 - 0) - F (x_1) = F (x_2) - F (x_1).$$

至此引出重要定理如下:

定理 1.16 (唯一性定理)

分布函数可由特征函数唯一确定.

 \Diamond

证明 由逆转公式1.15, 令 $y=x_1$ 沿 F(x) 的连续点趋向 $-\infty$, 令 $x=x_2$, 则

$$F(x) = \lim_{y \to -\infty} \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ity} - e^{-itx}}{it} f(t) dt.$$

于是在F(x) 的连续点,由f(t) 决定了F(x). 至于在F(x) 的间断点上,由于F(x) 的右连续性,只需沿连续点取右极限,就唯一确定了在间断点处的F(x) 值.

Ŷ 笔记

- 1. 结合此定理和特征函数的定义可知: 分布函数和特征函数相互唯一确定;
- 2. 用唯一性定理和逆转公式计算分布函数是困难的, 它的意义主要是理论上的.

定理 1.17 (逆 Fourier 变换)

设 f(t) 是特征函数, 且 $\int_{-\infty}^{\infty} |f(t)| dt < \infty$, 则分布函数 F(x) 的导数存在且连续, 此时

$$F'(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t) dt.$$

 \Diamond

证明 先证明 F(x) 在 \mathbb{R} 上是连续的. 任取一点 $x \in \mathbb{R}$, 选择 $\delta > 0$ 使得 $x \pm \delta$ 都是分布函数的连续点, 那么根据逆转公式可得

$$F(x+\delta) - F(x-\delta) = \frac{1}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} \frac{e^{-it(x-\delta)} - e^{-it(x+\delta)}}{it} f(t) dt.$$

利用不等式 $|e^{i\alpha}-1| \leq |\alpha|$, 可得

$$\left| \frac{e^{-it(x-\delta)} - e^{-it(x+\delta)}}{it} \right| \le 2\delta.$$

根据假设 $\int_{-\infty}^{\infty} |f(t)| dt < \infty$, 可得

$$F(x+\delta) - F(x-\delta) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-it(x-\delta)} - e^{-it(x+\delta)}}{it} f(t) dt$$

上式右边关于变量 δ 是连续的. 令 $\delta \to 0$, 同时使得 $x \pm \delta$ 都是 F(x) 的连续点, 那么由控制收敛定理得

$$\lim_{\delta \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-it(x-\delta)} - e^{-it(x+\delta)}}{it} f(t) dt = 0.$$

所以

$$F(x) = F(x+0) = F(x-0).$$

所以 F(x) 在 \mathbb{R} 上是连续的.

接下来, 由逆转公式, 对任意的 $x, x + \Delta x \in \mathbb{R}$, 有

$$\frac{F(x+\Delta x)-F(x)}{\Delta x}=\lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itx}-e^{-it(x+\Delta x)}}{it\Delta x}f(t)dt.$$

再次利用不等式 $|e^{i\alpha}-1| \leq |\alpha|$ 以及 f(t) 的绝对可积性, 可得

$$\frac{F(x+\Delta x)-F(x)}{\Delta x}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-itx}-e^{-it(x+\Delta x)}}{it\Delta x}f(t)dt.$$

利用控制收敛定理得到

$$F'(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{\Delta x \to 0} \frac{e^{-itx} - e^{-it(x + \Delta x)}}{it\Delta x} f(t) dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} f(t) dt.$$

因此 p(x) = F'(x) 存在且有界. 再次利用控制收敛定理可得

$$\lim_{h \to 0} \left[F'(x+h) - F'(x) \right] = \lim_{h \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(e^{-it(x+h)} - e^{-itx} \right) f(t) dt$$
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \lim_{h \to 0} \left(e^{-it(x+h)} - e^{-itx} \right) f(t) dt$$
$$= 0$$

所以 F'(x) 是连续的.

輸完 此定理说明,如果特征函数绝对可积,对应的随机变量必为连续型,且特征函数公式和该公式互为一对傅里叶变换.

对于离散型随机变量亦有类似结果: 假设 X 是取非负整数值的随机变量. 分布列为

$$P(X = k) = p_k, \quad k = 0, 1, 2, \dots,$$

那么其特征函数为

$$f(t) = \sum_{k=0}^{\infty} p_k e^{itk}$$

由

$$\int_0^{2\pi} e^{int} dt = \begin{cases} 2\pi, & n = 0, \\ 0, & n \neq 0. \end{cases}$$

那么,我们有

$$p_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-itk} f(t) dt.$$

至此我们讨论完毕特征函数与分布函数相互唯一确定, 所以我们可以得到如下定理2:

定理 1.18 (Levy 逆极限定理)

设 $\{f_n(t), n \ge 1\}$ 是分布函数 $\{F_n(x), n \ge 1\}$ 的特征函数, 如果对每一 $t, f_n(t) \to f(t)$, 且 f(t) 在 t = 0 处 连续, 则 f(t) 一定是某个分布函数 F 的特征函数, 且 $F_n \stackrel{w}{\to} F$.

命题 1.6 (依分布收敛的重要性质)

设X为随机变量, $\{X_n\}$ 为随机变量序列, 则下述命题等价:

- 1. $X_n \stackrel{d}{\longrightarrow} X$;
- 2. 对任一有界连续函数 g(x), 有 $\mathbb{E}g(X_n) \to \mathbb{E}g(X)^a$;
- 3. 对任一有界一致连续函数 g(x), 有 $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$;
- 4. 对任一有界连续并且任意阶导数也有界的函数 g(x), 有 $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$;
- 5. 对任意实数 $t, f_n(t) \to f(t)$, 其中 $f(t), f_n(t)$ 分别是 X, X_n 的特征函数.

a这一条在某些教材中被直接认定为依分布收敛的定义;也可认为此条是逆 Helly 定理

证明 对第二点进行粗略说明: $\Diamond f_n(t)$ 和 f(t) 分别为 X_n 和 X 的特征函数. 根据所给条件 $\mathbb{E}g(X_n) \to \mathbb{E}g(X)$, 我们可以得到

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} g(x) dF_n(x) = \int_{-\infty}^{\infty} g(x) dF(x)$$

由逆傅里叶变换,

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}g(x)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}f_n(t)dt\right]dx=\int_{-\infty}^{\infty}g(x)\left[\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}f(t)dt\right]dx$$

²细节见附录

交换积分的次序,有

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}g(x)e^{-itx}dx\right]f_n(t)dt=\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}g(x)e^{-itx}dx\right]f(t)dt$$

内层积分是 g(x) 的傅里叶变换, 表示为 $\hat{g}(t)$, 上式改写为

$$\lim_{n \to \infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{g}(t) f_n(t) dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{g}(t) f(t) dt$$

由于特征函数是一致连续,任意有限区间一致收敛的,所以交换极限和积分次序,可以得到

$$f_n(t) \to f(t)$$

又由逆极限定理, 可得 $X_n \to X$.

1.7 附录

测度与概率 测度 (Measure) 是集合函数的一种,主要用于度量集合的大小,常用于度量长度、面积、体积以及更一般的概念.

给定一个集合 Ω , 一个 σ - 代数 $F \in \Omega$ 的某个子集族, 如果它满足:

- 1. 空集属于 *F*;
- 2. 如果 A 属于 \mathcal{F} , 那么 A 的补集也属于 \mathcal{F} ;
- 3. 如果 A_i 是 \mathcal{F} 中的一列集合, 那么它们的并集 $\bigcup_{i=1}^{\infty} A_i$ 也属于 \mathcal{F} .

那么, 有定义在 σ - 代数 F 上的测度 μ 需要满足以下条件:

- 1. 非负性: 对于所有的集合 A 属于 \mathcal{F} , 都有 $\mu(A) \geq 0$;
- 2. 空集的测度是 0: $\mu(\emptyset) = 0$;
- 3. 可数可加性: 如果 A_1 , A_2 , A_3 , ... 是 F 上的两两不相交的集合序列, 那么

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu\left(A_i\right)$$

满足上述条件的函数 μ 称为 F 上的一个测度

博雷尔 σ - 代数, 随机变量和分布

- Borel σ 代数的定义: 在实数轴 \mathbb{R} 上, 所有开区间的集合形成的 σ 代数就是 Borel σ 代数, 记作 $\mathcal{B}(R)$. 也就是说, Borel σ 代数是对实数集合进行最精细划分的一个代数结构, 它包含所有开区间, 也包含这些开区间的补集, 以及这些集合的可数并与可数交.
- 随机变量的定义: 随机变量实际上是**定义在样本空间上的函数**, 对于每一个随机试验的基本事件, 随机变量 都**给出一个实数作为标志**. 更形式化的说, 设 (Ω, \mathcal{F}, P) 是一个概率空间, 随机变量 X 是定义在 Ω 上的可测 实值函数, 即对于任何 Borel 集 $B \in \mathcal{B}(R)$, 其逆像 $X^{-1}(B) \in \mathcal{F}$.
- 分布测度的定义: 设 X 是定义在 (Ω, \mathcal{F}) 上的随机变量, 对任意的 $B \in \mathcal{B}(R)$, 定义 $\mu_X(B) = \mu(X \in B) = \mu(\{\omega : X(\omega) \in B\})$, 则映射 $\mu_X : \mathcal{B}(R) \to [0,1]$ 被称为随机变量 X 的分布测度.
- 累积分布函数: 用 $F(x) = \mathbb{P}(X \le x)$ 描述随机变量的分布.

拉东-尼克蒂姆定理

- 级数的收敛与发散: 令 $\{a_n\}_{n\in\mathbb{N}}$ 为一实数列, 考虑级数 $\sum_{i=1}^{\infty}a_i$ 的部分和数列 $\{S_n=\sum_{i=1}^na_i\}_{n\in\mathbb{N}}$:
 - (1) 称级数 $\sum_{i=1}^{\infty} a_i$ 收敛到 S (convergent), 当且仅当 $\lim_{n\to\infty} S_n = S$, 即: $\forall \varepsilon > 0, \exists N \in \mathbb{N}$, s.t. $\forall n \geq N, |S_n S| < \varepsilon$;
 - (2) 称级数 $\sum_{i=1}^{\infty} a_i$ 绝对收敛 (converge absolutely), 当且仅当级数 $\sum_{i=1}^{\infty} |a_i|$ 收敛;
 - (3) 称级数 $\sum_{i=1}^{\infty} a_i$ 发散 (divergent), 当且仅当级数 $\sum_{i=1}^{\infty} a_i$ 不收敛;
 - (4) 称级数 $\sum_{i=1}^{\infty} a_i$ 定向发散到 ∞ (properly diverge to ∞), 当且仅当 $\lim_{n\to\infty} S_n = \infty$, 即, $\forall M \in \mathbb{R}$, $\exists N \in \mathbb{N}$, s.t. $\forall n \geq N, S_n > M$; 称级数 $\sum_{i=1}^{\infty} a_i$ 定向发散到 $-\infty$ (properly diverge to $-\infty$), 当且仅当 $\lim_{n\to\infty} S_n = -\infty$, 即, $\forall M \in \mathbb{R}$, $\exists N \in \mathbb{N}$, s.t. $\forall n \geq N, S_n < M$;
- 令 \mathcal{F} 是一个 σ -代数. 一个带符号测度 (广义测度) 被定义成一个集合函数 $\mu: \mathcal{F} \to (-\infty, \infty]$, 满足 $\mu(\varnothing) = 0$ 且对于 \mathcal{F} 中两两不相交的集合列 $\{A_i\}_{i\in\mathbb{N}}$ 有 $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$. 且若 $\mu(\bigcup_{i=1}^\infty A_i)$ 是有限的, 我们要求 $\sum_{i=1}^\infty \mu(A_i)$ 是绝对收敛的; 若 $\mu(\bigcup_{i=1}^\infty A_i)$ 是无穷大, 则我们要求 $\sum_{i=1}^\infty \mu(A_i)$ 是定向发散的.
- Radon-Nikodym 定理

定理 1.19 (Radon-Nikodym)

 (X, \mathcal{F}, μ) 是一个有限测度空间, $\mu(X) < +\infty, \lambda : \mathcal{F} \to (-\infty, +\infty]$ 是 (X, \mathcal{F}) 上的一个符号测度,则存在导函数 f,使得 $\lambda(A) = \int_A f d\mu (\forall A \in \mathcal{F})$ 当且仅当 λ 是绝对连续的,记 $f = \frac{d\lambda}{d\mu}$.

信息与 σ - 代数

- 1. 设 Ω 是非空集合. 设 T 是固定的某正数. 对于每一个 $t \in [0,T]$, 都有一个 σ 代数 $\mathcal{F}(t)$ (按时间标序). 进一步假定: 如果 $s \leq t$, 则 $\mathcal{F}(s)$ 中的所有事件都在 $\mathcal{F}(t)$ 中, 则称 σ 代数族 $\mathcal{F}(t)$, $0 \leq t \leq T$ 是一个域流 (filtration).
- 2. 设 X 是定义在非空样本空间 Ω 上的随机变量. 由 X 生成的 σ 代数 (记为 $\sigma(X)$) 是所有形如

$$\{\omega \in \Omega; X(\omega) \in B\} =: \{X \in B\}$$

的子集族, 其中 $B \in \mathbb{R}$ 的博雷尔子集.

3. 设 X 是定义在非空样本空间 Ω 上的一个随机变量. G 是其子集的 σ - 代数. 如果 $\sigma(X)$ 中的所有集合都在 G 中, 就称 X 是 G- 可测的.(G- measurable)

1. $\sigma(X)$ 是 σ — 代数, 其元素是集合, 这些集合都是样本空间 Ω 的子集, 元素是样本点 ω , 这些样本点满足通过 X 映射后落在 $\mathbb R$ 中的 Borel 集上, 这个 Borel 集是随机变量 X 的可取值范围. 比如一个描述是否下雨的随机变量 X 可取值 0,1, 则 $\sigma(X)$ 就是

$$\{\emptyset, \Omega, \{\mathsf{Fn}\}, \{\mathsf{AFn}\}\}$$

由随机变量 X 生成的 σ 代数, 简而言之, 如果 X 是个离散的话, 那就是 X 的取值集合所对应的样本点集合的所有子集组成的; 这个 σ 代数有 2^n 个集合.

2. 随机变量 $X \neq G$ 可测的的含义是当且仅当 G 中的信息足以确定 X. 且对任意 Borel 可测函数 f, f(X) 也 是 G 可测的.

期望收敛定理

定理 1.20 (Fatou 引理)

 X_1, X_2, \cdots 是定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的正值随机变量, 有

$$\mathbb{E}(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} \mathbb{E}(X_n)$$

实变函数中,相对应地,有:

若 $\{f_n\}_{n>0}$ 是一个实值可测正函数列,则有:

$$\int_{-\infty}^{+\infty} \lim_{n \to \infty} f_n(x) dx \le \lim_{n \to \infty} \int_{-\infty}^{\infty} f_n(x) dx$$

定理 1.21 (单调收敛定理 (Monotone convergence))

令 X_1, X_2, X_3, \cdots 是一系列依概率 1 收敛到 X 的随机变量. 若

$$0 \le X_1 \le X_2 \le X_3 \le \cdots$$
 a.s.

则

$$\lim_{n \to \infty} \mathbb{E} X_n = \mathbb{E} X$$

在实变函数中,相对应的:

令 f_1, f_2, f_3, \cdots 是一列定义在 $\mathbb R$ 上的博雷尔可测函数,且逐点收敛到 f.(i.e. $\forall x \in X, f(x) = \lim_{k \to \infty} f_k(x)$). 若

$$\forall x \in X, k \in N, f_k(x) \leq f_{k+1}(x)$$

则

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} f_n(x) dx = \int_{-\infty}^{\infty} f(x) dx$$

C

\$

笔记 关于 almost surely 和 almost everywhere: almost surely 就是指的是不满足条件的样本点集合是零概率事件; almost everywhere 就是不满足条件的自变量 x 构成的集合是勒贝格零测度集.

定理 1.22 (控制收敛定理 (Dominated convergence))

令 X_1, X_2, \cdots 是一系列依概率 1 收敛到 X 的随机变量. 如果存在另一个可积随机变量 Y 即 $\mathbb{E}Y < \infty$ 且几乎处处有 $|X_n| < Y$ 则

$$\lim_{n \to \infty} \mathbb{E} X_n = \mathbb{E} X$$

实变函数中,相对应的有:

令 f_1,f_2,\cdots 是一列定义在 $\mathbb R$ 上,逐点收敛到 f 的博雷尔可测函数. 如果存在另一函数 g 满足 $\int_{-\infty}^\infty g(x)dx<\infty$ 以及 $|f_n|\leq g$,则

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} f_n(x) dx = \int_{-\infty}^{\infty} f(x) dx$$

 \Diamond

逆极限定理的证明

引理 1.4

设 f(t) 为分布函数 F(x) 的特征函数,则对任意 $\lambda > 0$,有

$$\int_{|x| \ge 2\lambda} dF(x) \le 2\lambda \int_0^{1/\lambda} [1 - \text{Re}(f(t))] dt.$$

0

证明

$$\begin{split} \lambda \int_0^{1/\lambda} [1 - \operatorname{Re}(f(t))] dt &= \lambda \int_0^{1/\lambda} \int_{-\infty}^{\infty} [1 - \cos tx] dF(x) dt \\ &= \int_{-\infty}^{\infty} \lambda \int_0^{1/\lambda} [1 - \cos tx] dt dF(x) \\ &= \int_{-\infty}^{\infty} \left[1 - \frac{\sin(x/\lambda)}{x/\lambda} \right] dF(x) \\ &\geq \int_{|x/\lambda| \geq 2} \left[1 - \frac{\sin(x/\lambda)}{x/\lambda} \right] dF(x) \\ &\geq \frac{1}{2} \int_{|x| > 2\lambda} dF(x). \end{split}$$

下面证明逆极限定理. 对 $\{F_n\}$ 的任一子列 $\{F_{n'}\}$, 由海莱第一定理, 存在一个子子列 $\{F_{n''}\}\subset \{F_{n'}\}$ 和一个单调不减的右连续函数 F(x), 使得对 F 的任意连续点 x

$$F_{n''}(x) \to F(x)$$
.

下面只需证明 F 是分布函数. 因为如果 F 是分布函数,则由连续性定理, $f_{n''}(t) \to f_F(t)$,其中 $f_F(t)$ 为 F 的特征函数. 由己知条件 $f_{n'}(t) \to f(t)$,必有 $f_F(t) = f(t)$,故 f(t) 为特征函数,唯一确定了极限分布函数 F.这样, $\{F_n\}$ 的任意子列都有弱收敛的子子列,并且他们的极限均是由 f(t) 确定的同一个分布函数 F,因此 $F_n \xrightarrow{w} F$.

 $F_{n''}(x) \to F(x)$ 已然蕴含 $F \in [0,1]$, 下面证明 $\lim_{a \to \infty} F(a) - F(-a) = 1$. 当 $\pm a$ 是 F 的连续点时,

$$\int_{|x| \le a} dF_{n''}(x) \to F(a) - F(-a), \quad n'' \to \infty.$$

因此只需证明

$$\lim_{a \to \infty} \lim_{n'' \to \infty} \int_{|x| > a} dF_{n''}(x) = 0$$

只需证明

$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} \int_{|x| > \lambda} dF_n(x) = 0.$$

当上式成立时, 称 $\{F_n\}$ 胎紧 (tight). 由引理以及 f(0)=1, f(t) 在 t=0 处连续, 有

$$\lim_{\lambda \to \infty} \limsup_{n \to \infty} \int_{|x| \ge 2\lambda} dF_n(x) \le \lim_{\lambda \to \infty} \limsup_{n \to \infty} 2\lambda \int_0^{1/\lambda} \left[1 - \operatorname{Re}\left(f_n(t)\right) \right] dt$$

$$= \lim_{\lambda \to \infty} 2\lambda \int_0^{1/\lambda} \left[1 - \operatorname{Re}(f(t)) \right] dt$$

$$\le 2 \lim_{\lambda \to \infty} \sup_{0 \le t \le 1/\lambda} \left[1 - \operatorname{Re}(f(t)) \right] = 0.$$

证毕.

测度变换

先来看一个例子.

例题 1.4 (正态随机变量的测度变换)

初等概率论讲过, 对于一个正态随机变量 $X \sim N(\mu, \sigma^2)$, 可以通过变换随机变量的方式得到标准正态分布, 即令 $Y = \frac{X-\mu}{\sigma}$, 则 $Y \sim N(0,1)$.

在概率测度 ℙ下.

$$\mathbb{P}(Y \le y) = \int_{-\infty}^{y} \phi(t)dt, \forall y \in \mathbb{R}$$

现在, 我们考虑, 变换到 Ω 上的一个新的概率测度 $\widetilde{\mathbb{P}}$, 使得在 $\widetilde{\mathbb{P}}$ 下是标准正态随机变量 3 (即希望 $\widetilde{\mathbb{E}}(X)=0$, $\widetilde{\mathrm{VAR}}=1$). 为此我们不对 X 做变换, 而是对 $X(\omega)$ 取特定值的 ω 指定不同的概率.

例如 $X \sim N(\theta, 1), \theta > 0$. 则我们想对使得 $X(\omega)$ 取得正数的 ω 指定较小的概率, 对其取负数的 ω 指定较大的概率. 我们的目的是在不改变随机变量 X 本身的条件下改变 X 的分布.

定义随机变量

$$Z(\omega) = \exp\left(-\theta Y(\omega) - \frac{1}{2}\theta^2\right) \quad \forall \omega \in \Omega$$

该随机变量具有两个性质:

- 1. $\forall \omega \in \Omega, Z(\omega) > 0$;
- 2. $\mathbb{E}(Z) = 1$

对性质 2 的证明:

$$\mathbb{E}Z = \int_{-\infty}^{\infty} \exp\left\{-\theta y - \frac{1}{2}\theta^2\right\} \varphi(y) dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} \left(y^2 + 2\theta y + \theta^2\right)\right\} dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} (y + \theta)^2\right\} dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2} x^2\right\} dx,$$

我们利用随机变量 Z,来调节 Ω 中事件的概率,以产生新的概率测度 $\widetilde{\mathbb{P}}$. 为此,定义

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega) = \mathbb{E}(Z(\omega)\mathbb{I}_A), \quad \forall A \in \mathcal{F}$$

³在线性代数中, 前者类似于对向量本身做线性变换, 基不变, 而后者是向量不变, 改变基, 向量在新基下有新坐标, 就类似于这里在新概率测度 下有新密度函数

继而有

$$\begin{split} \widetilde{\mathbb{P}}\{X \leq b\} &= \int_{\{\omega; X(\omega) \leq b\}} Z(\omega) d\mathbb{P}(\omega) \\ &= \int_{\Omega} \mathbb{I}_{\{X(\omega) \leq b\}} Z(\omega) d\mathbb{P}(\omega) \\ &= \int_{\Omega} \mathbb{I}_{\{Y(\omega) \leq b - \theta\}} \exp\left\{-\theta Y(\omega) - \frac{1}{2}\theta^2\right\} d\mathbb{P}(\omega) \\ &= \int_{-\infty}^{\infty} \mathbb{I}_{\{y \leq b - \theta\}} e^{-\theta y - \frac{1}{2}\theta^2} \varphi(y) dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b - \theta} e^{-\theta y - \frac{1}{2}\theta^2} e^{-\frac{y^2}{2}} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b - \theta} e^{-\frac{1}{2}(y + \theta)^2} dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{b} e^{-\frac{1}{2}x^2} dx \end{split}$$

这就表明了在新概率测度 ℙ下,X 是标准正态变量.

定理 1.23 (测度变换定理)

设 $(\Omega, \mathcal{F}, \mathbb{P})$ 为一个概率空间,Z 为一个几乎处处非负的,且 $\mathbb{E}Z=1$ 的随机变量. 对于 $A\in\mathcal{F}$,定义

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega).$$

那么 $\widetilde{\mathbb{P}}$ 就是一个概率测度.此外,如果X是一个非负随机变量,则

$$\widetilde{\mathbb{E}}X = \mathbb{E}[XZ].$$

如果Z几乎处处严格为正,对于每个非负随机变量Y,我们也有

$$\mathbb{E}Y = \tilde{\mathbb{E}}\left[\frac{Y}{Z}\right]$$

证明 为了验证 \tilde{P} 是概率测度, 必须证明 $\tilde{P}(\Omega) = 1$ 以及 \tilde{P} 具有可数可加性质. 由假设, 我们有:

$$\tilde{\mathbb{P}}(\Omega) = \int_{\Omega} Z(\omega) d\mathbb{P}(\omega) = \mathbb{E}Z = 1$$

为证可数可加性, 设 A_1,A_2,\cdots 是 F 中一列互不相交的集合, 并定义 $B_n=\bigcup_{n=1}^n A_k, B_\infty=\bigcup_{k=1}^\infty A_k$. 由于

$$\mathbb{I}_{B_1} \leq \mathbb{I}_{B_2} \leq \mathbb{I}_{B_3} \leq \cdots$$

以及 $\lim_{n\to\infty} \mathbb{I}_{B_n} = \mathbb{I}_{B_\infty}$, 利用单调收敛定理, 可得:

$$\tilde{\mathbb{P}}\left(B_{\infty}\right) = \int_{\Omega} \mathbb{I}_{B_{\infty}}(\omega) Z(\omega) d\mathbb{P}(\omega) = \lim_{n \to \infty} \int_{\Omega} \mathbb{I}_{B_{n}}(\omega) Z(\omega) d\mathbb{P}(\omega)$$

但是, $\mathbb{I}_{B_n}(\omega) = \sum_{k=1}^n \mathbb{I}_{A_k}(\omega)$, 因此:

$$\int_{\Omega} \mathbb{I}_{B_n}(\omega) Z(\omega) d\mathbb{P}(\omega) = \sum_{k=1}^n \int_{\Omega} \mathbb{I}_{A_k}(\omega) Z(\omega) d\mathbb{P}(\omega) = \sum_{k=1}^n \tilde{\mathbb{P}}(A_k)$$

由上述两式, 我们得到可数可加性质:

$$\widetilde{\mathbb{P}}\left(\bigcup_{k=1}^{\infty} A_k\right) = \lim_{n \to \infty} \sum_{k=1}^{n} \widetilde{\mathbb{P}}\left(A_k\right) = \sum_{k=1}^{\infty} \widetilde{\mathbb{P}}\left(A_k\right)$$

现在

(1) 假设 X 是非负随机变量; 如果 X 是指示函数 $X = \mathbb{I}_A$, 则:

$$\widetilde{\mathbb{E}}X = \widetilde{\mathbb{P}}(A) = \int_{\Omega} \mathbb{I}_{A}(\omega)Z(\omega)d\mathbb{P}(\omega) = \mathbb{E}\left[\mathbb{I}_{A}Z\right] = \mathbb{E}(XZ)$$

(2) 进一步假设 X 是非负简单函数 $X=\sum_{k=1}^n\alpha_k\mathbb{I}_{A_k}(\omega)$, 其中 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是非负常数. 由积分的线性性

得

$$\widetilde{\mathbb{E}}(X) = \sum_{k=1}^{n} \alpha_k \widetilde{\mathbb{P}}(A_k) = \sum_{k=1}^{n} \alpha_k \int_{\Omega} \mathbb{I}_{A_k}(\omega) Z(\omega) d\mathbb{P}(\omega) = \int_{\Omega} Z(\omega) \sum_{k=1}^{n} \alpha_k \mathbb{I}_{A_k} d\mathbb{P}(\omega) = \mathbb{E}(XZ)$$

(3) 进一步假设 X 是非负可测函数. 对<u>每一正整数</u>n,定义集合

$$B_{k,n} = \left\{ \omega : \frac{k}{2^n} \le X(\omega) < \frac{k+1}{2^n} \right\}, \quad k = 0, 1, 2, \dots, 4^n - 1$$

对每一固定的 n, 集合 $B_{0,n},\cdots,B_{4^n-1,n}$ 对应于分划

$$0 < \frac{1}{2^n} < \frac{2}{2^n} < \dots < \frac{4^n}{2^n} = 2^n$$

并且在接下来的第n+1次分划中包含了第n次分划的所有分点,且新的分点都是相邻原分点之间的中点.因此,简单函数

$$X_n(\omega) = \sum_{k=0}^{4^n - 1} \frac{k}{2^n} \mathbb{I}_{B_{k,n}}(\omega)$$

满足 $0 \le X_1 \le X_2 \le \cdots \le X$,且随着 n 的增大,它们越来越精确地逼近函数 X,即 $\lim_{n\to\infty} X_n(\omega) = X(\omega)$,a.s. 由第二步可知,对每一个固定的 n 都有

$$\widetilde{\mathbb{E}}(X_n) = \mathbb{E}(X_n Z)$$

由单调收敛定理,得

$$\lim_{n \to \infty} \widetilde{\mathbb{E}}(X_n) = \lim_{n \to \infty} \mathbb{E}(X_n Z) = \mathbb{E}(X Z)$$



- 1. 在 $\widetilde{\mathbb{E}}X = \mathbb{E}[XZ]$ 中, 用 Y/Z 替代 X, 就得到 $\mathbb{E}Y = \widetilde{\mathbb{E}}\left[\frac{Y}{Z}\right]$.
- 2. 随机变量 X 可正可负, 对其正部和负部 $X^+ = \max\{X,0\}, X^- = \max\{-X,0\}$ 分别运用在 $\widetilde{\mathbb{E}}X = \mathbb{E}[XZ]$, 然后两式相减, 只要不出现 $\infty \infty$ 的情形, 就能说明在 $\widetilde{\mathbb{E}}X = \mathbb{E}[XZ]$ 对这样的 X 仍然成立.

定义 1.18 (等价概率测度)

设 Ω 是一个非空集合, \mathcal{F} 是 Ω 的子集的 σ -代数. 如果 (Ω,\mathcal{F}) 上的两个概率测度 \mathbb{P} 和 $\widetilde{\mathbb{P}}$ 在该 σ -代数中拥有相同的零概率集合. 则认为它们是等价的.

在定理 1.23的假设以及几乎处处 Z>0 的假设下, \mathbb{P} 和 $\widetilde{\mathbb{P}}$ 是等价的. 假设 $A\in\mathcal{F}$ 和 $\mathbb{P}(A)=0$, 那么意味着

$$\widetilde{\mathbb{P}}(A) = \int_{\Omega} \mathbb{I}_A(\omega) Z(\omega) d\mathbb{P}(\omega) = 0.$$

另一方面, 假设 $B \in \mathcal{F}$ 满足 $\tilde{\mathbb{P}}(B) = 0$. 于是 $\frac{1}{2}\mathbb{I}_B = 0$ 在 $\tilde{\mathbb{P}}$ 下几乎处处成立, 所以

$$\widetilde{\mathbb{E}}\left[\frac{1}{Z}\mathbb{I}_{B}\right]=0.$$

这意味着 $\mathbb{P}(B) = \mathbb{E}\mathbb{I}_B = 0$. 这表明 \mathbb{P} 和 \mathbb{P} 有一致的概率为零的集合. 因为概率为 1 的集合是概率为 0 的集合的补集, 所以 \mathbb{P} 和 \mathbb{P} 也在概率为 1 的集合上一致. 因为 \mathbb{P} 和 \mathbb{P} 是等价的, 所以一个事件几乎肯定会发生时, 我们不需要特定指定哪一种测度.

定义 1.19 (拉东-尼克蒂姆导数)

 $(\Omega, \mathcal{F}, \mathbb{P})$ 是一个概率空间, $\widetilde{\mathbb{P}}$ 是另一个在 (Ω, \mathcal{F}) 上的, 与 \mathbb{P} 等价的概率测度, 令 \mathbb{Z} 为一几乎处处为正的随机变量, 通过如下方式将 \mathbb{P} 沟通起来:

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega)$$

这样,Z 被称为 $\widetilde{\mathbb{P}}$ 对 \mathbb{P} 的 Radon-Nikodým 导数,记作

$$Z=\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$$

定理 1.24 (Radon-Nikodým)

设 \mathbb{P} 和 $\widetilde{\mathbb{P}}$ 是 (Ω,\mathcal{F}) 上等价的概率测度,则存在几乎必然为正的随机变量Z,满足 $\mathbb{E}(Z)=1$ 并且

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega) d\mathbb{P}(\omega), \quad \forall A \in \mathcal{F}$$

C

第2章 马尔可夫链

内容提要

- □ 离散时间离散状态的马尔可夫链
- □ 不变分布
- □ 常返性和瞬时性

- □ 极限分布和遍历定理
- □ 连续时间离散状态的马尔可夫链
- □ 连续时间连续状态的马氏过程介绍

定义 2.1 (随机过程)

假设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是概率空间, T 为指标集, \mathcal{E} 为点集. 称一族随机变量 $X(\omega, t): \Omega \mapsto \mathcal{E}, t \in T$ 为随机过程, 记作 $\mathbf{X} = (X(t), t \in T)$, 其中称 T 为时间参数空间, \mathcal{E} 为状态空间. $\{X(t) = a\}$ 意思是随机过程 t 时刻处于状态 a. 有时将时间 t 作为下标, 如 $X_t(\omega)$; 有时为简洁起见, 略去 ω , 仅写作 X(t) 或 X_t .

 $\widehat{\mathbf{Y}}$ 笔记 常见的时间参数空间 T 包括 $(-\infty,\infty),[0,\infty),[0,1],\{\cdots,-2,-1,0,1,2,\cdots\},\{0,1,2,\cdots\}$ 或 $\{1,2,\cdots\}$. 这 些 T 有一个重要特点: 即自然的"序"关系.

在随机过程研究中, 有一个全新的概念一样本曲线 (或样本轨道). 任给一个 $\omega \in \Omega$, 称 $X(\omega, \cdot): T \to \mathcal{E}$ 为样本曲线.

常用以下函数来表示随机过程的数字特征:

1. 均值函数

假设对每一个 $t \in T$, $\mathbb{E}|X(t)| < \infty$, 称

$$\mu_X(t) = \mathbb{E}X(t), \quad t \in T$$

为X的均值函数.

2. 自协方差函数

假设对每一个 $t \in T$, $\mathbb{E}(X(t)^2) < \infty$, 称

$$\sigma_X^2(t) = \operatorname{Var}(X(t)), \quad t \in T$$

为 X 的方差函数. 定义

$$r_X(s,t) = \mathbb{E}(X(s)X(t)), \quad s,t \in T,$$

称 $r_X(s,t)$ 为 X 的自相关函数. 定义自协方差函数为

$$K(s,t) = \operatorname{Cov}(X(s), X(t)) = r_X(s,t) - \mu_X(s)\mu_X(t), \quad s, t \in T.$$

3. 互相关函数

假设 $X=(X(t),t\in T)$ 和 $Y=(Y(t),t\in T)$ 是两个随机过程, 并且对每一个 $t\in T$, $\mathbb{E}\left(X(t)^2\right)<\infty$, $\mathbb{E}\left(Y(t)^2\right)<\infty$. 定义

$$r_{X,Y}(s,t) = \mathbb{E}(X(s)Y(t)), \quad s,t \in T,$$

称 $r_{X,Y}(s,t)$ 为 X 和 Y 的互相关函数.

定义 2.2 (平稳过程)

令 $X = (X(t), t \in T)$ 是一个随机过程, 对每一个 $t \in T$, $\mathbb{E}(X(t)^2) < \infty$. 如果

• 均值函数为常数,即存在一个常数 μ 使得

$$\mu_X(t) \equiv \mu, \quad t \in T$$

• 自相关函数 $r_X(s,t)$ 仅与时间差 s-t 有关,即存在一个函数 $r_X: \mathbb{R} \to \mathbb{R}$ 使得

$$r_X(s,t) = \tau_X(s-t), \quad s,t \in T,$$

称 X 为弱平稳过程, 有时也称为宽平稳过程.

令 $X = (X(t), t \in T)$ 是一个随机过程. 如果对任意 $k \ge 1, t_1, t_2, \dots, t_k \in T$ 以及 $t \in T$ 都有

$$(X(t_1+t), X(t_2+t), \cdots, X(t_k+t)) \stackrel{d}{=} (X(t_1), X(t_2), \cdots, X(t_k)),$$

称X为强平稳过程,有时也称为严平稳过程.

命题 2.1

一般来说, 平稳随机过程的强弱性并没有直接联系. 如果强平稳过程的二阶矩存在且有限, 那么它一定是弱平稳过程.

定义 2.3 (独立增量过程)

令 $\pmb{X} = (X(t), t \in T)$ 是一个随机过程. 对任意 s < t, 称 X(t) - X(s) 为过程增量. 如果 X(t) - X(s) 的分布仅依赖于时间差 t-s, 而与 s 和 t 无关, 称 \pmb{X} 是平稳增量过程. 如果对任意 $k \ge 1$ 和 $t_1 < t_2 < \cdots < t_k$, 增量

$$X(t_1), X(t_2) - X(t_1), \dots, X(t_k) - X(t_{k-1})$$

是相互独立的,则称 X 是独立增量过程.

例题 2.1 令 $X = (X(t), t \in T)$ 是一个随机过程. 如果其任意有限维分布为联合正态分布, 则称 X 为正态过程, 有时也称为 Gauss 过程.

注: 对任何 k 个时刻 $t_1, \dots, t_k, (X(t_1), \dots, X(t_k))$ 服从正态分布当且仅当其任意线性组合

$$\sum_{i=1}^{k} a_i X\left(t_i\right)$$

服从正态分布, 其中 a_1, \dots, a_k 为任意实数.

命题 2.2

由于联合正态向量的分布由均值向量和协方差矩阵所唯一确定, 所以弱平稳正态过程一定是强平稳正态过程.

2.1 离散、有限时间的马尔可夫链

令 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间. 马尔可夫链是一系列拥有所谓的马尔可夫性质的随机变量 $\{X_t\}_{t\in \mathbf{T}}$ 组成的随机过程, 其中 \mathbf{T} 是指标集. 当 $\mathbf{T}=\mathbb{N}$, 我们就记 $\{X_n\}_{n\in \mathbb{N}}$, 这就是离散时间马尔可夫链. 假设 X_n 在状态空间 S 上取值.

例题 2.2 对称随机游动 (Symmetric random walk)

考虑随机变量序列 $\{\xi_j\}_{j=1}^{\infty}$, 其中 $\{\xi_j\}$ 独立同分布且 $\xi_j=\pm 1$ 拥有概率 1/2. 令

$$X_n = \sum_{j=1}^n \xi_j.$$

则 $\{X_n\}_{n\in\mathbb{N}}$ 是整数集 \mathbb{Z} 上的对称随机游动. 给定 $X_n=i$, 我们有

$$\mathbb{P}(X_{n+1} = i \pm 1 \mid X_n = i) = \mathbb{P}(\xi_{n+1} = \pm 1) = \frac{1}{2}$$

以及 $\mathbb{P}(X_{n+1} = \text{anything else} \mid X_n = i) = 0$. 我们看出, 如果知道了 X_n , 那么 X_{n+1} 的分布就完完全全被确定了,

即

$$\mathbb{P}(X_{n+1} = i_{n+1} \mid \{X_m = i_m\}_{m=0}^n) = \mathbb{P}(X_{n+1} = i_{n+1} \mid X_n = i_n);$$

也就是说 X_{n+1} 关于整个历史序列 $\{X_m\}_{m=0}^n$ 的条件概率, 等于 X_{n+1} 仅仅关于前一期 X_n 的条件概率. 满足这种性质的过程或序列 $\{X_n\}_{n\in\mathbb{N}}$ 就被称为 **马尔可夫链**.

不失一般性, 令状态空间 $S = \{1, 2, \dots, I\}$, 开始问题的探讨.

定义 2.4 (一步状态转移矩阵)

定义一步转移概率为过程基于时间n在状态k的条件下,在时间n+1位于状态j的概率.即从n到n+1,状态从k转移到j的概率.

$$p_{n:kj} = \mathbb{P}\left(X_{n+1} = j \mid X_n = k\right).$$

称马尔可夫链是平稳的, 如果 $p_{n:kj}$ 和时间 n 无关. 从现在开始讨论的都是平稳马尔可夫链, 也叫时间齐次的马尔可夫链, 令 $p_{kj}=p_{n:kj}$. 定义一步状态转移矩阵为 $\mathbf{P}=(p_{ij})_{i,i\in S}$:

$$\boldsymbol{P} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1I} \\ p_{21} & p_{22} & \cdots & p_{2I} \\ \vdots & \vdots & & \vdots \\ p_{I1} & p_{I2} & \cdots & p_{II} \end{pmatrix}$$

P 是随机矩阵, 满足

$$p_{ij} \ge 0, \quad \sum_{j \in S} p_{ij} = 1.$$

马尔可夫链的一个基本性质是:

命题 2.3 (Chapman-Kolmogorov 等式)

$$\mathbb{P}(X_n = j \mid X_0 = i)$$

$$= \sum_{k \in S} \mathbb{P}(X_n = j \mid X_m = k) \mathbb{P}(X_m = k \mid X_0 = i), \quad 1 \le m \le n - 1.$$



笔记 给定一个 Markov 链 $X = (X_n, n \ge 0)$, 如何确定它的分布呢? 令

$$p_0(i) = \mathbb{P}(X_0 = i),$$

 $p_0 = (p_0(1), p_0(2), \dots, p_0(I)),$

称 p_0 为 X 的初始分布, 描述随机系统的初始状态分布规律. 通常情况下, 它是可选择、可控制的. 初始分布 p_0 和转移概率矩阵 P 完全决定着 Markov 链的分布. 首先, 计算 1 维分布. 令

$$p_n(j) = \mathbb{P}(X_n = j), \quad j = 1, 2, \dots, I,$$

 $p_n = (p_n(1), p_n(2), \dots, p_n(I))$

表示 Markov 链在 n 时刻处于各个状态的概率分布. 利用全概率公式得

$$p_1(j) = \sum_{i=1}^{I} \mathbb{P}(X_1 = j \mid X_0 = i) \mathbb{P}(X_0 = i)$$
$$= \sum_{i=1}^{I} p_0(i) p_{ij},$$

即

$$p_1 = p_0 P$$
.

类似地,2时刻的分布

$$p_{2}(j) = \mathbb{P}(X_{2} = j) = \sum_{i=1}^{I} \mathbb{P}(X_{2} = j \mid X_{1} = i) \mathbb{P}(X_{1} = i)$$
$$= \sum_{i=1}^{I} p_{1}(i)p_{ij},$$

即

$$\boldsymbol{p}_2 = \boldsymbol{p}_1 \boldsymbol{P} = \boldsymbol{p}_0 \boldsymbol{P}^2$$

依此运用归纳法得

$$p_n = p_0 P^n$$
.

下面计算任意有限维联合分布. 对任意给定的 $i_0,i_1,\cdots,i_n\in S$, 利用条件概率的链式法则得

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n)$$

$$= \mathbb{P}(X_0 = i_0) \mathbb{P}(X_1 = i_1 \mid X_0 = i_0) \mathbb{P}(X_2 = i_2 \mid X_1 = i_1, X_0 = i_0) \cdots$$

$$\mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1}, \cdots, X_0 = i_0).$$

进而,

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n)$$

= $p_0(i_0) p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n}.$

回忆一下, 定义 2.4 给出了在 n 时刻处于状态 k 的条件下, n+1 时刻处于状态 j 的概率. 从时间角度看, 这是一步转移概率. 人们自然会问, 2 步转移概率

$$p_{ij}^{(2)} := \mathbb{P}(X_{n+2} = j \mid X_n = i)$$

为多少? 由条件概率运算法则知

$$p_{ij}^{(2)} = \sum_{k=1}^{I} \mathbb{P}(X_{n+2} = j \mid X_{n+1} = k, X_n = i) \mathbb{P}(X_{n+1} = k \mid X_n = i)$$
$$= \sum_{k=1}^{I} p_{kj} p_{ik}.$$

用矩阵表示如下:

$$oldsymbol{P}^{(2)} = \left(p_{ij}^{(2)}
ight)_{I imes I} = oldsymbol{P}^2$$

类似地, 三步转移概率用矩阵表示如下:

$$P^{(3)} = (p_{ij}^{(3)})_{I \times I} = P^{(2)}P = P^3.$$

更一般地, 今

$$p_{ij}^{(m)} = \mathbb{P}(X_{n+m} = j \mid X_n = i), \quad \mathbf{P}^{(m)} = (p_{ij}^{(m)})_{I \times I},$$

有

$$\mathbf{P}^{(m)} = \mathbf{P}^m$$

因此,对任意 n 和 m 成立

$$\boldsymbol{P}^{(n+m)} = \boldsymbol{P}^{(n)} \boldsymbol{P}^{(m)},$$

即

$$p_{ij}^{(n+m)} = \sum_{k=1}^{I} p_{ik}^{(n)} p_{kj}^{(m)}, \quad i, j = 1, 2, \dots, I.$$

这个等式即为 Chapman-Kolmogorov 方程, 将在后面讨论中反复使用.

2.2 不变分布

定义 2.5 (不变分布)

给定马尔可夫链的初始分布 μ_0 ,则 X_n 的分布就是

$$\mu_n = \mu_0 P^n$$
.

称一个分布 π 是 不变分布或平稳分布, 如果:

$$\pi = \pi P$$

例题 2.3 状态转移矩阵 P 如果是:

1.

$$\mathbf{P} = \left(\begin{array}{rrr} 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{array} \right)$$

则有两个线性独立的不变分布 $\pi_1 = (1,0,0)$ a 以及 $\pi_2 = (0,0,1)$.

2.

$$\boldsymbol{P} = \left(\begin{array}{ccc} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \end{array} \right)$$

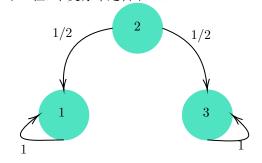
则只有一个不变分布 $\pi = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4}).$

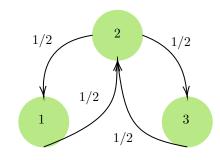
基于这两个简单的例子, 人们自然会考虑不变分布的存在性以及唯一性 (existence and uniqueness).

1. 存在性: 是否存在不变分布. 相当于问是否存在一个特征值为1的 P 的非负左特征向量.

$$\pi = \pi P \Longleftrightarrow \pi (I - 1 \cdot P) = 0$$

2. 唯一性: 不变分布是否唯一?





问题的答案与矩阵 P 的性质有关.

引理 2.1 (最大特征值为 1)

矩阵 P 的谱半径等于 1.

$$\rho(\mathbf{P}) = \max_{\lambda} |\lambda| = 1$$

最大值通过遍历矩阵 P 的所有特征值取得.

证明 我们已然知道 1 是一个 P 的特征根. 为了说明它是最大特征值, 令 u 为 P 特征值 λ 的左特征向量. 于是

$$\lambda u_i = \sum_{j \in S} u_j p_{ji}$$

这说明

$$|\lambda| \sum_{i \in S} |u_i| = \sum_{i \in S} \left| \sum_{j \in S} u_j p_{ji} \right| \le \sum_{i,j \in S} |u_j| p_{ji} = \sum_{j \in S} |u_j|.$$

所以 $|\lambda| \leq 1$.

定义 2.6 (矩阵的不可约性 (Irreducibility))

如果存在一个置换矩阵 (permutation matrix: 重新排列过的单位矩阵)R, 使得

$$\boldsymbol{R}^T \boldsymbol{P} \boldsymbol{R} = \left(\begin{array}{cc} \boldsymbol{A}_1 & \boldsymbol{B} \\ 0 & \boldsymbol{A}_2 \end{array} \right)$$

则称 P 是可约的. 否则是不可约的.

除了上面的考察方法,还可以通过考虑各个状态之间的可通达性 (communication).

定义 2.7 (可通达和互达)

我们说状态 i 是自 i 可通达的 (accessible), 如果 $\exists s > 1$ 以及状态 $(k_1, k_2, \dots, k_{s-1})$, 满足

$$p_{k_0k_1}p_{k_1k_2}\cdots p_{k_{s-1}k_s} > 0, \quad k_0 = i, k_s = j$$

如果状态i也是自j可通达的,就称状态(i,j) 互达 (pair (i,j) communicates). 可以根据状态节点的互通性,把马尔可夫链的状态空间分为若干等价类. 在统一等价类内的状态彼此相通,在不同等价类中的状态不可能彼此相通.



- 对称性: 如果 $i \leftrightarrow j$, 那么 $j \leftrightarrow i$;
- 传递性: 如果 $i \leftrightarrow j, j \leftrightarrow k$, 那么 $i \leftrightarrow k$;
- 自反性: $i \leftrightarrow i$.

利用等价关系就可以把状态空间 S 分成若干个等价类.

引理 2.2

转移矩阵 P 不可约, 等价于状态空间中的每对节点互达.

下面的重要定理给出了在不可约性条件下不变分布的存在唯一性的肯定答案.

定理 2.1 (Perron-Frobenius)

令 A 为一个不可约、非负矩阵,则

1. 存在 A 的正、右和左特征向量 x, y, 使得

$$\mathbf{A}\mathbf{x} = \rho(\mathbf{A})\mathbf{x}, \quad x_i > 0; \quad \mathbf{y}^{\mathsf{T}}\mathbf{A} = \rho(\mathbf{A})\mathbf{y}^{\mathsf{T}}, \quad y_i > 0.$$

2. $\lambda = \rho(\mathbf{A})$ 是重数为 1 的特征值.

笔记从马尔可夫链的有向图表示来看,置换矩阵的作用对应于状态节点的重新标记.不可约性意味着不存在包含一些节点、与其他所有节点都不相通的子组(即一马尔可夫链的所有状态属于同一等价类).违反不可约性可能导致不变分布的非唯一性,如之前例子所示.然而,不可约性却不需要不变分布的唯一性.

例题 2.4 考虑一个拥有如下转移矩阵的马尔可夫链

$$\boldsymbol{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0.3 & 0.4 & 0.3 & 0 & 0 \\ 0.3 & 0 & 0.4 & 0.3 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

矩阵可约,但你也可以发现此马尔可夫链拥有唯一不变分布

$$\pi = (0, 0, 0, 0.5, 0.5)$$

2.3 常返性与瞬时性

定义 2.8 (周期)

假设状态 $i \in S$. 令

$$d_i = \gcd\left\{n \ge 1 : p_{ii}^{(n)} > 0\right\},\,$$

其中 gcd 表示最大公因子, 称 d_i 为状态 i 的周期. 如果 $d_i = 1$, 称状态 i 是非周期的.

命题 2.4

 $d_i = d_j$, if $i \leftrightarrow j$.

证明 d_i 和 d_j 都是自然数, 只要证明 d_i 和 d_j 互相整除即可. 不失一般性, 假设 $i\neq j$. 既然 $i\leftrightarrow j$, 所以一定存在 $n,m\geq 1$ 使得 $p_{ij}^{(n)}>0,p_{ji}^{(m)}>0$. 这样,

$$p_{jj}^{(n+m)} \ge p_{ji}^{(m)} p_{ij}^{(n)} > 0.$$

因此

$$d_j \mid n+m$$

另外, 假设 $s \ge 1$ 使得 $p_{ii}^{(s)} > 0$, 那么

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

因此

$$d_j \mid n+s+m$$

结合起来得

$$d_j \mid s$$
.

由于s是使得 $p_{ii}^{(s)} > 0$ 的任意自然数,所以

$$d_i \mid d_i$$
.

类似地,可以证明

$$d_i \mid d_j$$

所以 $d_i = d_i$. 证毕.

定义 2.9 (首到时、各种常返状态)

给定状态 $j \in S$, 令

$$T_j = \inf \{ n \ge 1 : X_n = j \}.$$

约定 $\inf \emptyset = \infty$, 即如果不存在 $n \ge 1$ 使得 $X_n = j$, 则令 $T_i = \infty$.

当 $X_0=j$ 时, T_j 表示 Markov 链首次回到状态 j 的时刻; 当 $X_0=i\neq j$ 时, T_j 表示 Markov 链首次到达状态 j 的时刻.

如果

$$\mathbb{P}\left(T_{j} < \infty \mid X_{0} = j\right) = 1$$

称 j 是常返状态. 如果 $\mathbb{E}(T_j \mid X_0 = j) < \infty$, 称 j 为正常返状态; 如果 $\mathbb{E}(T_j \mid X_0 = j) = \infty$, 称 j 为零常返状态. 如果

$$\mathbb{P}\left(T_{j} < \infty \mid X_{0} = j\right) < 1$$

则称j是瞬时状态.

命题 2.5

如果 $i \leftrightarrow j$, 那么

- 1. i是瞬时状态当且仅当j是瞬时状态;
- 2. i是常返状态当且仅当j是常返状态;
- 3. i是正常返状态当且仅当j是正常返状态.

该性质表明常返性和瞬时性是一个"类性质",同一等价类中的各个元素具有相同属性.

令

$$f_{jj}^{(n)} = \mathbb{P}(T_j = n \mid X_0 = j)$$

= $\mathbb{P}(X_1 \neq j, \dots, X_{n-1} \neq j, X_n = j \mid X_0 = j).$

则 $f_{jj}^{(n)}$ 表示从状态 j 出发后在 n 时刻首次返回到 j 的概率, 而 $p_{jj}^{(n)}$ 表示从状态 j 出发后在 n 时刻处于 j 的概率. 按定义, j 是常返状态当且仅当

$$\sum_{n=1}^{\infty} f_{jj}^{(n)} = 1$$

j 是瞬时状态当且仅当

$$\sum_{n=1}^{\infty} f_{jj}^{(n)} < 1$$

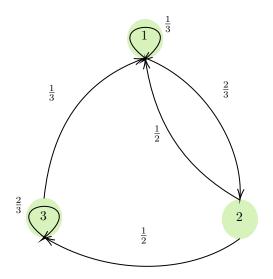
另外, 常返状态 i 的**平均返回时间**为

$$\tau_j = \mathbb{E}(T_j \mid X_0 = j) = \sum_{n=1}^{\infty} n f_{jj}^{(n)}.$$

例题 2.5 假设 $X = (X_n, n \ge 0)$ 是 Markov 链, 状态空间为 $S = \{1, 2, 3\}$, 转移概率矩阵为

$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0\\ \frac{1}{2} & 0 & \frac{1}{2}\\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$$

有向图表示为



容易看出所有状态互通,且具有周期 1. 下面考察状态 1 的常返性问题. 不难得到概率:

$$f_{11}^{(1)} = \mathbb{P} (T_1 = 1 \mid X_0 = 1) = \frac{1}{3},$$

$$f_{11}^{(2)} = \mathbb{P} (T_1 = 2 \mid X_0 = 1) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3},$$

$$f_{11}^{(3)} = \mathbb{P} (T_1 = 3 \mid X_0 = 1) = \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{9}.$$

继续下去, 当 $n \ge 3$ 时,

$$f_{11}^{(n)} = \mathbb{P}(T_1 = n \mid X_0 = 1) = \frac{2}{3} \cdot \frac{1}{2} \cdot \left(\frac{2}{3}\right)^{n-3} \cdot \frac{1}{3} = \frac{1}{9} \cdot \left(\frac{2}{3}\right)^{n-3}.$$

因此,

$$\mathbb{P}(T_1 < \infty \mid X_0 = 1) = \sum_{n=1}^{\infty} f_{11}^{(n)} = f_{11}^{(1)} + f_{11}^{(2)} + \sum_{n=3}^{\infty} f_{11}^{(n)}$$
$$= \frac{1}{3} + \frac{1}{3} + \sum_{n=3}^{\infty} \frac{1}{9} \cdot \left(\frac{2}{3}\right)^{n-3} = 1.$$

这样,状态1是常返的,并且

$$\tau_1 = \mathbb{E} \left(T_1 < \infty \mid X_0 = 1 \right) = \sum_{n=1}^{\infty} n f_{11}^{(n)}$$

$$= 1 \cdot f_{11}^{(1)} + 2 \cdot f_{11}^{(2)} + \sum_{n=3}^{\infty} n f_{11}^{(n)}$$

$$= \frac{1}{3} + \frac{2}{3} + \frac{1}{9} \cdot \sum_{n=3}^{\infty} n \left(\frac{2}{3} \right)^{n-3}$$

$$= \frac{8}{3}$$

1 是正常返的, 平均常返时间为 $\frac{8}{3}$. 可以类似讨论其他状态. 每一次都求解 $f_{jj}^{(n)}$ 是很繁琐的, 我们希望通过转移概率就能判断常返性.

引理 **2.3** $(p_{jj}^{(n)}$ 和 $f_{jj}^{(n)}$ 的关系)

$$p_{jj}^{(n)} = \sum_{k=1}^{n} f_{jj}^{(k)} p_{jj}^{(n-k)}$$

证明

$$p_{jj}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = j)$$

$$= \sum_{k=1}^n \mathbb{P}(T_j = k, X_n = j \mid X_0 = j)$$

$$= \sum_{k=1}^n \mathbb{P}(T_j = k \mid X_0 = j) \mathbb{P}(X_n = j \mid T_j = k, X_0 = j)$$

$$= \sum_{k=1}^n f_{jj}^{(k)} p_{jj}^{(n-k)}.$$

定理 2.2

1. j 是常返状态当且仅当

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty;$$

2. j 是瞬时状态当且仅当

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty.$$

证明 由引理,

$$\sum_{n=1}^{\infty} p_{jj}^{(n)} = \sum_{n=1}^{\infty} \sum_{k=1}^{n} f_{jj}^{(k)} p_{jj}^{(n-k)}$$

交换求和次序得

$$\begin{split} \sum_{n=1}^{\infty} p_{jj}^{(n)} &= \sum_{k=1}^{\infty} f_{jj}^{(k)} \sum_{n=k}^{\infty} p_{jj}^{(n-k)} \\ &= \sum_{k=1}^{\infty} f_{jj}^{(k)} \sum_{n=0}^{\infty} p_{jj}^{(n)} \\ &= \sum_{k=1}^{\infty} f_{jj}^{(k)} \left(1 + \sum_{n=1}^{\infty} p_{jj}^{(n)} \right) \end{split}$$

假设 $\sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty$, 从上式可以直接得到

$$\sum_{k=1}^{\infty} f_{jj}^{(k)} = \frac{\sum_{n=1}^{\infty} p_{jj}^{(n)}}{1 + \sum_{n=1}^{\infty} p_{jj}^{(n)}} < 1$$

假设 $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$, 对任意 $M \ge 1$,

$$\sum_{n=1}^{M} p_{jj}^{(n)} = \sum_{n=1}^{M} \sum_{k=1}^{n} f_{jj}^{(k)} p_{jj}^{(n-k)}$$

交换求和次序得

$$\sum_{n=1}^{M} p_{jj}^{(n)} = \sum_{k=1}^{M} f_{jj}^{(k)} \sum_{n=k}^{M} p_{jj}^{(n-k)}$$

$$= \sum_{k=1}^{M} f_{jj}^{(k)} \sum_{n=0}^{M-k} p_{jj}^{(n)}$$

$$\leq \sum_{k=1}^{M} f_{jj}^{(k)} \left(1 + \sum_{n=1}^{M} p_{jj}^{(n)} \right)$$

即

$$\sum_{k=1}^{M} f_{jj}^{(k)} \ge \frac{\sum_{n=1}^{M} p_{jj}^{(n)}}{1 + \sum_{n=1}^{M} p_{jj}^{(n)}}.$$

两边取极限, 令 $M \to \infty$, 得 $\sum_{k=1}^{\infty} f_{jj}^{(k)} \ge 1$. 因此,

$$\sum_{k=1}^{\infty} f_{jj}^{(k)} = 1$$

定理 2.3

1. 如果j是瞬时状态,那么

$$\lim_{n \to \infty} p_{jj}^{(n)} = 0$$

2. 如果j是零常返状态,那么

$$\lim_{n \to \infty} p_{jj}^{(n)} = 0;$$

3. 如果j是非周期正常返状态,那么

$$\lim_{n \to \infty} p_{jj}^{(n)} = \frac{1}{\tau_j}$$

4. 如果j是周期为 d_i 的正常返状态,那么

$$\lim_{n \to \infty} p_{jj}^{(nd_j)} = \frac{d_j}{\tau_i}$$

其中 $\tau_j = \mathbb{E}(T_j \mid X_0 = j) < \infty.$

 \odot

推论 2.1

1. 如果i是零常返或瞬时状态,那么对任意状态i,

$$\lim_{n \to \infty} p_{ij}^{(n)} = 0;$$

2. 如果j是非周期正常返状态,那么对任意状态i,

$$\lim_{n \to \infty} p_{ij}^{(n)} = \frac{f_{ij}}{\tau_i},$$

其中 f_{ij} 表示从 i 出发可到达 j 的概率.

C

定理 2.4

有限状态 Markov 链一定存在正常返状态.

证明 采用反证法. 假设每个状态都是瞬时的或零常返的, 那么 $\lim_{n\to\infty}p_{ij}^{(n)}=0$ 成立. 另一方面, 对任意 $n\geq 1$ 和任意状态 i,j,

$$\sum_{j=1}^{I} p_{ij}^{(n)} = 1.$$

由于 $I < \infty$, 令 $n \to \infty$, 交换极限号和求和号次序得

$$1 = \lim_{n \to \infty} \sum_{i=1}^{I} p_{ij}^{(n)} = \sum_{i=1}^{I} \lim_{n \to \infty} p_{ij}^{(n)} = 0,$$

矛盾. 证明完毕.

2 M27 + A

笔记 若一个马尔可夫链的状态空间是有限的,那么其所有常返状态均为正常返;

若一个不可约的马尔可夫链的状态空间是有限的,那么其所有状态均为正常返.

定义 2.10 (状态停留次数)

任意给定状态 1,令

$$M_j = \sharp \{ n \geq 0 : X_n = j \}$$
.

 M_i 表示 Markov 链运行过程中处于状态 j 的总次数.

下列定理表明, 如果 j 是常返状态, 那么 Markov 链一定有无穷多个时刻处于状态 j; 如果 j 是瞬时状态, 那么 Markov 链处于状态 j 的平均次数是有限的.

定理 2.5

(i) 如果 j 是常返状态, 那么

$$\mathbb{P}\left(M_{j}=\infty\mid X_{0}=j\right)=1;$$

(ii) 如果 i 是瞬时状态, 那么

$$\mathbb{P}\left(M_{i} < \infty \mid X_{0} = j\right) = 1.$$

进而,令

$$\rho = \mathbb{P}(T_j < \infty \mid X_0 = j) = \sum_{k=1}^{\infty} f_{jj}^{(k)} < 1,$$

那么 M_i 服从参数为 $1-\rho$ 的几何分布,

$$\mathbb{P}(M_j = n \mid X_0 = j) = (1 - \rho)\rho^{n-1}, \quad n \ge 1.$$

证明 对任意 $m \ge 1$,

$$\mathbb{P}(M_{j} \geq m \mid X_{0} = j) = \sum_{k=1}^{\infty} \mathbb{P}(T_{j} = k \mid X_{0} = j) \mathbb{P}(M_{j} \geq m \mid T_{j} = k)$$

$$= \sum_{k=1}^{\infty} \mathbb{P}(T_{j} = k \mid X_{0} = j) \mathbb{P}(M_{j}(\mathbb{K} k \mapsto \mathbb{N}) + \mathbb{N}) \geq m - 1 \mid X_{k} = j)$$
(由 Markov 性)
$$= \rho \mathbb{P}(M_{j} \geq m - 1 \mid X_{0} = j).$$

由此类推,

$$\mathbb{P}(M_j \ge m \mid X_0 = j) = \rho^{m-1} \mathbb{P}(M_j \ge 1 \mid X_0 = j)$$

= ρ^{m-1} .

因此,

$$\mathbb{P}(M_j = \infty \mid X_0 = j) = \lim_{m \to \infty} \mathbb{P}(M_j \ge m \mid X_0 = j) = \lim_{m \to \infty} \rho^{m-1}$$
$$= \begin{cases} 1, & \rho = 1, \\ 0, & \rho < 1. \end{cases}$$

即当 ρ <1时 M_i 服从参数为 $1-\rho$ 的几何分布.

下面的定理把常返性、不变分布联系在一起.

定理 2.6 (不变分布与正常返)

假设 $X = (X_n, n \ge 0)$ 是非周期、不可约 Markov 链, 转移概率矩阵为 P. 那么该 Markov 链存在不变分布 当且仅当每个状态都是正常返的, 并且

$$\pi_j = \frac{1}{\tau_j}, \quad j \in S,$$

其中 $\tau_i = \mathbb{E}T_i$ 为状态 j 的平均常返时间.

C

2.4 极限分布与有限马尔可夫链的遍历定理

定义 2.11 (极限分布)

假设 $\boldsymbol{X}=(X_n,n\geq 0)$ 是 Markov 链, 状态空间为 $S=\{1,2,\cdots,I\}$, 转移概率矩阵为 \boldsymbol{P} , 初始分布为 \boldsymbol{p}_0 . 如果存在 S 上一个概率分布 $\boldsymbol{\mu}=(\mu_1,\mu_2,\cdots,\mu_I)$ 使得对任意 $j\in S$,

$$\lim_{n\to\infty} p_n(j) = \mu_j$$

其中 $p_n(j)$ 表示 n 时刻处于状态 j 的概率. 称 μ 是该 Markov 链的极限分布.

自然的问题是, 什么条件下 Markov 链存在极限分布? 如何计算极限分布? 注意到,

$$p_n(j) = \sum_{i=1}^{I} p_0(i) p_{ij}^{(n)}.$$

如果对每个 $i \in S$, 存在 S 上一个概率分布 $\nu_i = (\nu_{i1}, \nu_{i2}, \dots, \nu_{iI})$ 使得

$$\lim_{n \to \infty} p_{ij}^{(n)} = \nu_{ij}, \quad j \in S,$$

那么由控制收敛定理得

$$\lim_{n \to \infty} p_n(j) = \lim_{n \to \infty} \sum_{i=1}^{I} p_0(i) p_{ij}^{(n)}$$

$$= \sum_{i=1}^{I} p_0(i) \lim_{n \to \infty} p_{ij}^{(n)}$$

$$= \sum_{i=1}^{I} p_0(i) \nu_{ij}.$$

令

$$\mu_j = \sum_{i=1}^{I} p_0(i)\nu_{ij}, \quad j \in S,$$

那么

$$\sum_{j=1}^{I} \mu_j = \sum_{i,j=1}^{I} p_0(i)\nu_{ij} = 1.$$

这样, $\mu = (\mu_1, \mu_2, \dots, \mu_N)$ 是该 Markov 链的极限分布.

根据上述分析, 问题归结于计算 $\lim_{n\to\infty} \mathbf{P}^{(n)}$. 由于

$$P^{(n)} = P^n$$

因此只需要计算

$$\lim_{n\to\infty} P^n.$$

例题 2.6 考虑 3 状态 Markov 链, 转移概率矩阵为

$$\boldsymbol{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{8} & \frac{3}{8} \\ 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

有特征分解

$$P = C \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{-1+3i}{8} & 0 \\ 0 & 0 & \frac{-1-3i}{8} \end{pmatrix} C^{-1}$$

其中

$$C = \begin{pmatrix} 1 & -7 - 9i & -7 + 9i \\ 1 & -16 + 24i & -16 - 24i \\ 1 & 26 & 26 \end{pmatrix},$$

$$C^{-1} = \frac{1}{780} \begin{pmatrix} 416 & 156 & 208 \\ -8 + 14i & -3 - 11i & 11 - 3i \\ -8 - 14i & -3 - 11i & 11 - 3i \end{pmatrix}.$$

容易看出

$$m{P}^n = m{C} \left(egin{array}{ccc} 1 & 0 & 0 \ 0 & \left(rac{-1+3i}{8}
ight)^n & 0 \ 0 & 0 & \left(rac{-1-3i}{8}
ight)^n \end{array}
ight) m{C}^{-1}$$

进一步取极限得

$$\lim_{n \to \infty} \mathbf{P}^n = \mathbf{C} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{C}^{-1}$$
$$= \begin{pmatrix} \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \\ \frac{8}{15} & \frac{1}{5} & \frac{4}{15} \end{pmatrix}$$

因此极限分布为 $\mu = (\frac{8}{15}, \frac{1}{5}, \frac{4}{15}).$

定理 2.7

假设 $X=(X_n,n\geq 0)$ 是非周期不可约 Markov 链, 转移概率矩阵为 P. 那么该 Markov 链存在极限分布当且仅当存在不变分布, 并且二者相等.

\$

笔记 对于有限马尔可夫链, 非周期、不可约等价于存在 s>0, 使得对任意两个状态 (i,j), 都有 $\mathbf{P}^s_{ij}>0$. 这是因为一个有限马尔可夫链如果非周期和不可约, 则存在某个正整数 s, 使得从任意状态 i 出发, 在 s 步以内可以到达任意状态 j, 即 $\mathbf{P}^s_{ij}>0$; 反之, 如果存在 s>0, 使得对于所有状态 i,j, 都有 $\mathbf{P}^s_{ij}>0$, 则该马尔可夫链是非周期的和不可约的. 这是因为如果存在某个状态 i, 使得从 i 出发不能在有限步内到达某个其他状态 j, 则 $\mathbf{P}^s_{ij}=0$, 与假设矛盾.

证明 给出两个初始分布 μ_0 以及 $\tilde{\mu}_0$, 定义总变差为

$$d(\boldsymbol{\mu}_{0}, \tilde{\boldsymbol{\mu}}_{0}) = \frac{1}{2} \sum_{i \in S} |\mu_{0,i} - \tilde{\mu}_{0,i}|.$$

由于

$$0 = \sum_{i \in S} (\mu_{0,i} - \tilde{\mu}_{0,i}) = \sum_{i \in S} (\mu_{0,i} - \tilde{\mu}_{0,i})^{+} - \sum_{i \in S} (\mu_{0,i} - \tilde{\mu}_{0,i})^{-},$$

其中 $a^+ = \max(a, 0)$ and $a^- = \max(-a, 0)$, 则也有

$$d(\boldsymbol{\mu}_{0}, \tilde{\boldsymbol{\mu}}_{0}) = \frac{1}{2} \sum_{i \in S} (\mu_{0,i} - \tilde{\mu}_{0,i})^{+} + \frac{1}{2} \sum_{i \in S} (\mu_{0,i} - \tilde{\mu}_{0,i})^{-}$$
$$= \sum_{i \in S} (\mu_{0,i} - \tilde{\mu}_{0,i})^{+} \leq 1.$$

于是, 令 $\mu_s = \mu_0 P^s$, $\tilde{\mu}_s = \tilde{\mu}_0 P^s$ 并考虑 $d(\mu_s, \tilde{\mu}_s)$. 我们就有

$$d\left(\boldsymbol{\mu}_{s}, \tilde{\boldsymbol{\mu}}_{s}\right) = \sum_{i \in S} \left[\sum_{j \in S} \left(\mu_{0,j} \left(\boldsymbol{P}^{s}\right)_{ji} - \tilde{\mu}_{0,j} \left(\boldsymbol{P}^{s}\right)_{ji} \right) \right]^{+}$$

$$\leq \sum_{j \in S} \left(\mu_{0,j} - \tilde{\mu}_{0,j} \right)^{+} \sum_{i \in B_{+}} \left(\boldsymbol{P}^{s}\right)_{ji}$$

其中 B_+ 是 S 的一个子集, 其元素 i 满足 $\sum_{j \in S} (\mu_{0,j} - \tilde{\mu}_{0,j}) (\mathbf{P}^s)_{ji} > 0$. B_+ 不可能包含 S 中所有元素, 否则对任意 i 有 $(\mu_0 \mathbf{P}^s)_i > (\tilde{\mu}_0 \mathbf{P}^s)_i$ 以及

$$\sum_{i \in S} (\boldsymbol{\mu}_0 \boldsymbol{P}^s)_i > \sum_{i \in S} (\tilde{\boldsymbol{\mu}}_0 \boldsymbol{P}^s)_i$$

这是不可能的因为两边都是 1. 因此至少存在一个不属于 B_+ 的状态 l. 根据假设, 存在 s>0 和 $\alpha\in(0,1)$ 使得对任意状态 (i,j) 有 $(\mathbf{P}^s)_{ij}\geq\alpha$. 则在第 j 行对于该状态 l 就有 $(\mathbf{P}^s)_{jl}\geq\alpha$, 在第 j 行对其余从状态 j 可到达的属于 B_+ 的状态 i 相加 (按列相加), 就有 $\sum_{i\in B_+} (\mathbf{P}^s)_{ii}\leq (1-\alpha)<1$. 于是

$$d(\boldsymbol{\mu}_s, \tilde{\boldsymbol{\mu}}_s) \le d(\boldsymbol{\mu}_0, \tilde{\boldsymbol{\mu}}_0) (1 - \alpha);$$

即这个马尔可夫链每走s步,两个初始分布之间的变差在缩小.类似的对于任意 $m \ge 0$,有

$$d\left(\boldsymbol{\mu}_{n},\boldsymbol{\mu}_{n+m}\right) \leq d\left(\boldsymbol{\mu}_{n-sk},\boldsymbol{\mu}_{n+m-sk}\right) (1-\alpha)^{k} \leq (1-\alpha)^{k},$$

其中 k 是满足 $n-sk\geq 0$ 的最大整数. 如果 n 充分大, 右边就会充分小. 这说明序列 $\{\mu_n\}_{n=0}^\infty$ 是 Cauchy 列. 因此 其有极限 π , 满足

$$oldsymbol{\pi} = \lim_{n o \infty} oldsymbol{\mu}_0 oldsymbol{P}^{n+1} = \lim_{n o \infty} \left(oldsymbol{\mu}_0 oldsymbol{P}^n
ight) oldsymbol{P} = oldsymbol{\pi} oldsymbol{P}.$$

 π 也是唯一的. 如果存在不同的 $\pi^{(1)}$ 以及 $\pi^{(2)}$, 则

$$d\left(\boldsymbol{\pi}^{(1)},\boldsymbol{\pi}^{(2)}\right) = d\left(\boldsymbol{\pi}^{(1)}\boldsymbol{P}^{s},\boldsymbol{\pi}^{(2)}\boldsymbol{P}^{s}\right) < d\left(\boldsymbol{\pi}^{(1)},\boldsymbol{\pi}^{(2)}\right).$$

这是矛盾的, 所以 $\pi^{(1)} = \pi^{(2)}$.

定理 2.8 (Ergodic theorem)

令 $X=(X_n,n\geq 0)$ 是非周期不可约 Markov 链, 具有唯一不变分布 $\pi(x)$; 令 f 是一个有界函数 $f:S\to\mathbb{R}$, 则当 $N\to\infty$ 时,

$$\frac{1}{N} \sum_{n=0}^{N-1} f(X_n) \to \langle f \rangle_{\pi}^{a} = \sum_{i \in S} f(i) \pi_i, \quad \text{a.s.}$$

 $a\langle f\rangle$ 表示在某概率测度下 f 的均值

(

笔记 该定理是"均值遍历定理",其内涵就是"时间平均等于样本平均".

2.5 连续时间离散状态的马尔可夫链

2.5.1 基本概念

接下来讨论的是,时间连续但状态空间依然是离散的一种随机过程,称为 0 过程.

定义 2.12 (Q 过程)

假设 $X = (X(t), t \ge 0)$ 是连续时间随机过程, 状态空间为 $S = \{1, 2, \dots, I\}$. 如果对任意 $0 \le s < t$,

$$\mathbb{P}(X(t) = j \mid X(s) = i, X(u) = i_u, u < s) = \mathbb{P}(X(t) = j \mid X(s) = i),$$

其中 $i, j, i_u \in S$, 称 X 为连续时间 Markov 链. 等价地可以写成: 对任意 $t_0 < t_1 < \cdots < t_n < t_{n+1}$,

$$\mathbb{P}(X(t_{n+1}) = j \mid X(t_n) = i, X(t_{n-1}) = i_{n-1}, \dots, X(t_0) = i_0)$$

$$= \mathbb{P}\left(X\left(t_{n+1}\right) = j \mid X\left(t_{n}\right) = i\right),\,$$

其中 $i, j, i_{n-1}, \dots, i_0 \in S$.

假设轨道 $\{X_t\}_{t\in\mathbb{R}}$ 是右连续的; 状态空间 $S=\{1,2,\cdots,I\}$. 令转移概率

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

同样地, 我们讨论的是时间齐次的马尔可夫链, 即转移概率只与时间差有关, 而与 s 无关. 根据定义我们有

$$p_{ij}(t) \ge 0, \quad \sum_{j \in S} p_{ij}(t) = 1.$$

进一步地, 我们要求, 当 $h \to 0^+$ 时, 满足如下几个等式:

$$p_{ii}(h) = 1 - \lambda_i h + o(h), \quad \lambda_i > 0,$$

$$p_{ij}(h) = \lambda_{ij} h + o(h), \quad j \neq i.$$

$$p_{ij}(0) = 1.$$

第一个等式说明的是**时间规范性 (regularity in time)**, 第二个等式表示 λ_{ij} 为状态跳跃率: 如果随机过程在时间 t 时位于状态 j, 然后 t 到 t+h 之间有变化发生, 则随机过程必然是跳跃到 $i \neq j$ 的状态. 同时, 我们有

$$\lambda_{ij} \ge 0, \quad \sum_{j \in S, j \ne i} \lambda_{ij} = \lambda_i.$$

命题 2.6 (Chapman-Kolmogorov 方程)

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

矩阵形式为 P(t+s) = P(t)P(s) = P(s)P(t), 其中 $P(t) = (p_{ij}(t))_{i,j \in S}$.

现在定义

$$Q = \lim_{h \to 0+} \frac{1}{h} (P(h) - I)$$

于是就有

$$q_{ii} = -\lambda_i, \quad q_{ij} = \lambda_{ij} \quad (i \neq j), \quad \sum_{j \in S} q_{ij} = 0$$

称 Q 为马尔可夫链的转移速率矩阵 (generator). 既然

$$\frac{\boldsymbol{P}(t+h) - \boldsymbol{P}(t)}{h} = \frac{\boldsymbol{P}(h) - \boldsymbol{I}}{h} \boldsymbol{P}(t)$$

以及 P(h), P(t) 乘积可交换, 则有

定理 2.9 (Kolmogorov 前后方程)

1. Kolmogorov 向后方程

$$\boldsymbol{P}'(t) = \boldsymbol{Q}\boldsymbol{P}(t);$$

2. Kolmogorov 向前方程

$$P'(t) = P(t)Q$$

3.

$$\mathbf{P}(t) = e^{t\mathbf{Q}} = \mathbf{I} + \sum_{n=1}^{\infty} \frac{t^n \mathbf{Q}^n}{n!}$$

 \Diamond

\$

笔记 (1) $\lim_{t\to 0} \frac{1-p_{ii}(t)}{t} = q_{ii} \le \infty$;

 $(2) \lim_{t \to 0} \frac{p_{ij}(t)}{t} = q_{ij} < \infty.$

称 q_{ij} 为从状态 i 转移到 j 的转移速率. 这个速率的分子是很短时间内转移到 j 的概率, 分母是时间区间长度; 而 q_{ii} 是从状态 i 离开的速率.

例题 2.7 Poisson 过程.

假设 N = (N(t), t > 0) 是参数为 λ 的 Poisson 过程, 那么 $S = \mathbb{Z}_+$, 并且

$$p_{0} = (1, 0, 0, \cdots),$$

$$p_{ij}(t) = \mathbb{P}(N(t+s) = j \mid N(s) = i)$$

$$= \mathbb{P}(N(t) = j - i)$$

$$= \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}, \quad j \ge i.$$

不难验证, Q-矩阵为

$$Q = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & 0 \\ 0 & -\lambda & \lambda & \cdots & 0 \\ 0 & 0 & -\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \end{pmatrix}$$

Poisson 过程是最简单的连续时间 Markov 链.

定义 2.13 (首次转移时间 (waiting time))

任意给定 $i \in S$, 定义

$$\tau_i = \inf\{t > 0 : X(t) \neq i\}.$$

如果 X(0)=i, 那么 τ_i 表示 X 首次实现转移的时刻, 也可看作 X 在状态 i 处的停留时间. 并定义首次转移时间 τ 的概率分布为 $\nu(t)=\mathbb{P}(\tau\geq t)$.

1

定理 2.10

$$\mathbb{P}\left(\tau_i > t \mid X(0) = i\right) = e^{-\alpha_i t}, \quad t > 0,$$

其中

$$\alpha_i = -q_{ii} = \sum_{j \neq i} q_{ij}.$$

 \sim

同样地需要考察一下不变分布. 假设存在 S 上概率分布 π 使得

$$\boldsymbol{\pi}\boldsymbol{P}(t) = \boldsymbol{\pi}, \quad t > 0,$$

那么通过选择 $p_0 = \pi, X$ 便成为平稳 Markov 链. 称 π 为 X 的不变分布.

定理 2.11

如果某概率分布 π 满足下列方程

$$\pi Q = 0$$
,

那么 π 为X的不变分布.

 \odot

证明 注意, $P(t) = \mathbb{E}^{tQ}$. 那么

$$\boldsymbol{\pi}(\boldsymbol{P}(t)-\boldsymbol{I}) = \boldsymbol{\pi} \sum_{k=1}^{\infty} \frac{1}{k!} \boldsymbol{Q}^k t^k = \sum_{k=1}^{\infty} \frac{1}{k!} \boldsymbol{\pi} \boldsymbol{Q}^k t^k = 0.$$

因此,对任意 t > 0,

$$\pi P(t) = \pi.$$

即 π 为 Markov 链的不变分布.

2.5.2 嵌入式链和不可约性

定义 2.14 (嵌入式链 (Embedded Chain))

设一个 Q 过程的转移速率矩阵为 Q. 该过程与一个离散时间的马尔可夫链紧密联系, 其状态转移矩阵 $ilde{Q}=(ilde{q}_{ij})$ 被定义为

$$\tilde{q}_{ij} = \begin{cases} q_{ij}/q_{ii}, & \text{if } i \neq j \text{ and } q_{ii} > 0, \\ 0, & \text{if } i \neq j \text{ and } q_{ii} = 0, \end{cases}$$

$$\tilde{q}_{ii} = \begin{cases} 0, & \text{if } q_{ii} > 0, \\ 1, & \text{if } q_{ii} = 0 \end{cases}$$

其中 $q_{ii} = \sum_{j \neq i} q_{ij}$.

状态转移矩阵 $\tilde{\mathbf{Q}} = (\tilde{q}_{ij})$ 被称为跳跃阵,与之相联系的马尔可夫链称为原 Q 过程的嵌入式链或跳跃链.

定义跳跃时间为

$$J_0 = 0$$
, $J_{n+1} = \inf \{ t : t \ge J_n, X_t \ne X_{J_n} \}$, $n \in \mathbb{N}$,

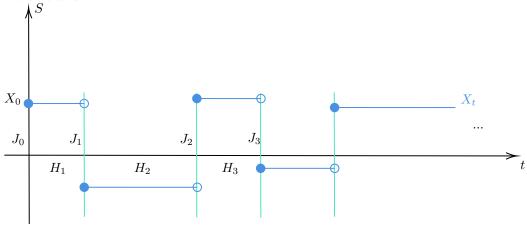
在这里我们依照习俗规定 $\inf \emptyset = \infty$. 对 $n = 1, 2, \dots$, 定义持续时间为

$$H_n = \begin{cases} J_n - J_{n-1}, & \text{if } J_{n-1} < \infty, \\ \infty, & \text{otherwise,} \end{cases}$$

令 $X_{\infty} = X_{J_n}$ 如果 $J_{n+1} = \infty$. 则由 $\{X_t\}$ 生成的跳跃链就被定义为

$$Y_n = X_{J_n}, \quad n \in \mathbb{N}$$

这些概念的示意图如下所示:



下面的结果提供了嵌入式链和原始的 Q 过程之间的关系.

定理 2.12

跳跃链 $\{Y_n\}$ 是一个以 \tilde{Q} 为状态转移矩阵的马尔可夫链, 并且所有的持续时间 H_1,H_2,\cdots 是独立的指数分布随机变量, 且其参数分别为 $q_{Y_0,Y_0},q_{Y_1,Y_1},\cdots$

证明 首先由马尔可夫性 (无记忆性) 知诸 H_i 应服从指数分布; 设 H_i 的速率参数为 ν_i , 则

$$\mathbb{P}\left(H_i \le t\right) = 1 - e^{-\nu_i t}.$$

但

$${H_i \le t} \iff {X(t+s) \ne i \mid X(s) = i}$$

所以

$$q_{ii} = \lim_{t \to 0} \frac{1 - p_{ii}(t)}{t}$$

$$= \lim_{t \to 0} \frac{1}{t} \mathbb{P}(X(t+s) \neq i \mid X(s) = i)$$

$$= \lim_{t \to 0} \frac{1}{t} \mathbb{P}(H_i \leq t) = \lim_{t \to 0} \frac{1}{t} \left(1 - e^{-\nu_i t}\right)$$

$$= \nu_i.$$

Q过程的一些性质和其嵌入式链息息相关.

定理 2.13 (嵌入式链和 Q 过程不可约性的等价性)

O 过程 X 是不可约的, 当且仅当其嵌入式链 Y 是不可约的.

证明 首先注意到,对任意 $i \neq j$,如果 $q_{ij} > 0$,则有

 $\begin{aligned} p_{ij}(t) &\geq \mathbb{P}\left(J_{1} \leq t, Y_{1} = j, H_{2} > t \middle| X_{0} = Y_{0} = i\right) = \int_{0}^{t} \exp\left(-q_{i}u\right) q_{i} \frac{q_{ij}}{q_{i}} du \exp\left(-q_{j}t\right) \\ &= \left(1 - e^{-q_{i}t}\right) \frac{q_{ij}}{q_{i}} e^{-q_{j}t} > 0 \end{aligned}, \forall t > 0$

如果 Y 是不可约的,则对任意 $i \neq q$,都存在 $s \geq 1$ 以及 (k_1, \dots, k_{s-1}) 使得

$$\tilde{q}_{k_0 k_1} \tilde{q}_{k_1 k_2} \cdots \tilde{q}_{k_{s-1} k_s} > 0, \quad k_0 = i, k_s = j.$$

因此,我们有

$$q_{k,i,k,j+1} > 0 (j = 0, \dots, s-1)$$

以及

$$p_{ij}(t) \ge p_{k_0k_1}(t/s)p_{k_1k_2}(t/s)\cdots p_{k_{s-1}k_s}(t/s) > 0$$

反过来, 假设对某 t 成立 $p_{ij}(t) > 0$. 则由

$$p_{ij}(t) = \left(e^{\mathbf{Q}t}\right)_{ij} = \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} \mathbf{Q}^n\right)_{ij} > 0$$

知存在 $n_0 \ge 1$ 以及 k_1, \dots, k_{n_0-1} 使得

$$q_{k_0k_1}q_{k_1k_2}\cdots q_{k_{n_0-1}k_{n_0}} > 0, \quad k_0 = i, k_{n_0} = j$$

这说明对于Y而言j是自i可达.证毕.

2.5.3 Q过程的遍历定理

不可约性已经足以保证极限分布存在且等于不变分布.

定理 2.14

设Q是不可约矩阵,则对于任意初始分布 μ_0 ,都有

$$\mu(t) = \mu_0 P(t) \to \pi, \quad t \to \infty$$

证明 根据不可约性和定理2.13, 知对任意 $i \neq j, t > 0, p_{ij}(t) > 0$. 显然对任意 $i \in S, t > 0$ 有 $p_{ii}(t) \geq \mathbb{P}(J_1 > t | X_0 = i) = e^{-q_i t} > 0$. 据此, 我们推断对任意 (i, j) 存在 $t_0 > 0, \alpha \in (0, 1)$ 使得 $p_{ij}(t_0) \geq \alpha$. 和之前离散时间该定理的证明类似, 对任意两个不同的初始分布 $\mu(0), \tilde{\mu}(0)$,

$$d(\boldsymbol{\mu}(t), \tilde{\boldsymbol{\mu}}(t)) \le d(\boldsymbol{\mu}(t - kt_0), \tilde{\boldsymbol{\mu}}(t - kt_0))(1 - \alpha)^k, \quad t \ge kt_0, k \in \mathbb{N}.$$

由压缩映射原理,存在唯一的不变分布.

定理 2.15 (Ergodic theorem)

设 Q 不可约. 则对于任意有界函数 $f: S \to \mathbb{R}$, 有

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f\left(X_s\right) ds = \langle f \rangle_\pi = \sum_{i \in S} f(i) \pi_i, \quad \text{ a.s.}$$

2.6 连续时间连续状态的马氏过程介绍

定义 2.15 (适应随机过程)

 Ω 是一个包含域流 F(t), $0 \le t \le T$ 的非空样本空间. 设 X(t) 是一个由 $t \in [0,T]$ 索引的随机过程; 如果对于每个 t, 随机变量 X(t) 是 F(t) -可测的, 我们就说 X(t) 是一个适应的随机过程.

定义 2.16 (马氏过程)

设 $\{X(t), \mathcal{F}(t), t > 0\}$ 是概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的适应过程, 如果对任意有界 Borel 函数 g(x), 有

$$\mathbb{E}(g(X(t)) \mid \mathcal{F}(s)) = \mathbb{E}(g(X(t)) \mid X(s)), \forall 0 \le s < t,$$

则称 $\{X(t), \mathcal{F}(t), t \geq 0\}$ 为马氏过程.

笔记 回忆域流、适应随机过程的定义; 马尔可夫性说明, f(X(t)) 在时刻 s 的估计 (预测) 只依赖于过程在时刻 s 的 值 X(s), 而不依赖于时刻 s 之前的路径.

定义 2.17 (转移函数)

令 $B ∈ \mathcal{B}(\mathbb{R})$, 函数

$$P(B, t|x, s) = \mathbb{P}(X_t \in B|X_s = x), \quad t \ge s$$

称为转移函数.

转移函数 $P:\mathcal{B}(\mathbb{R})\times[0,\infty)\mid\mathbb{R}\times[0,\infty)\to\mathbb{R}^+$ 具有如下性质:

- 1. 对任意 $t \ge s, x \in \mathbb{R}, P(\cdot, t|x, s)$ 都是博雷尔域 $\mathcal{B}(\mathbb{R})$ 上的概率测度.
- 2. 对任意 $t \geq s, B \in \mathcal{B}(\mathbb{R}), P(B,t|\cdot,s)$ 都是实数域上的可测函数.
- 3. (C-K 方程)

$$P(B, t|x, s) = \int_{\mathbb{R}} P(B, t|y, u) P(dy, u|x, s), s \le u \le t$$

$$F(s, t; x, y) = \mathbb{P}(X(t) \le y \mid X(s) = x), 0 \le s < t, x, y \in \mathbb{R}$$

称为过程的转移概率分布; 如果存在非负函数 f(s,t;x,y) 使得

$$F(s,t;x,y) = \int_{-\infty}^{y} f(s,t;x,u)du$$

就称 f 为过程的转移概率密度. 如果 F(s,t;x,y)=F(0,t-s;x,y) ,则称过程是时齐的,记 $F(0,\tau;x,y)$ 为 $F(\tau;x,y)$,对时齐马氏过程,若转移概率密度存在则记为 $f(\tau;x,y)$.

第3章 鞅

内容提要

□ 离散时间鞅

□ 连续时间鞅介绍

□ 停时原理和离散鞅收敛定理

3.1 离散时间鞅

假设一个赌博者正在进行一系列赌博, 每次赌博输赢的概率都是 $\frac{1}{2}$. 令 $\{Y_n, n=1,2,\cdots\}$, 是一列独立同分布的随机变量, 表示每次赌博的结果

$$\mathbb{P}\{Y_n = 1\} = \mathbb{P}\{Y_n = -1\} = \frac{1}{2}$$

这里 $\{Y_n = 1\}$ ($\{Y_n = -1\}$) 表示赌博者在第 n 次赌博时的赢 (输). 如果赌博者采用的赌博策略 (即所下赌注) 依赖于前面的赌博结果, 那么他的赌博可以用下面的随机变量序列

$$b_n = b_n (Y_1, \dots, Y_{n-1}), n = 2, 3, \dots$$

描述, 其中 $b_n < \infty$ 是第 n 次的赌注, 若赌赢则获利 b_n , 否则输掉 b_n . 设 X_0 是该赌博者的初始赌资, 则

$$X_n = X_0 + \sum_{i=1}^n b_i Y_i$$

是他在第n次赌博后的赌资. 我们可以断言

$$\mathbb{E}\left[X_{n+1}\mid Y_1,\cdots,Y_n\right]=X_n$$

证明 事实上,可以得到

$$X_{n+1} = X_n + b_{n+1}Y_{n+1},$$

因此

$$\mathbb{E}[X_{n+1} \mid Y_1, \cdots, Y_n] = \mathbb{E}[X_n \mid Y_1, \cdots, Y_n] + \mathbb{E}[b_{n+1}Y_{n+1} \mid Y_1, \cdots, Y_n]$$

$$= X_n + b_{n+1}\mathbb{E}[Y_{n+1} \mid Y_1, \cdots, Y_n]$$

$$(因为X_n 与b_{n+1} 由Y_1, \cdots, Y_n 确定)$$

$$= X_n + b_{n+1}\mathbb{E}[Y_{n+1}]$$

$$(因为\{Y_n\} 是独立随机变量序列)$$

$$= X_n \quad (因为\mathbb{E}[Y_{n+1}] = 0, \forall n \ge 0)$$

这证明了,如果每次赌博的输赢机会是均等的,并且赌博策略是依赖于前面的赌博结果,则赌博是"公平的".因此任何赌博者都不可能将公平的赌博通过改变赌博策略使得赌博变成有利于自己的赌博.

定义 3.1 (鞅 (Martingale))

假设 $X=(X_n,n\geq 0)$ 是一列随机变量, $(F_n,n\geq 0)$ 是 F 的一列单调不减子 σ -域. 如果满足

- 1. 对每一个 $n \geq 0, X_n$ 是 \mathcal{F}_{n-} 可测的;
- 2. 对每一个 n > 0, $\mathbb{E}|X_n| < \infty$;
- 3. 对每一个 n > 1.

$$\mathbb{E}(X_n \mid \mathcal{F}_{n-1}) = X_{n-1} \text{ a.s.,}$$

称 $(X_n, \mathcal{F}_n, n \geq 0)$ 是鞅.

类似地,如果对每一个n > 1,

$$\mathbb{E}\left(X_n\mid\mathcal{F}_{n-1}\right)\geq X_{n-1}\quad\text{ a.s.,}$$

 $(X_n, \mathcal{F}_n, n \geq 0)$ 是下鞅; 如果对每一个 $n \geq 1$,

$$\mathbb{E}\left(X_n\mid\mathcal{F}_{n-1}\right)\leq X_{n-1} \text{ a.s.,}$$

称 $(X_n, \mathcal{F}_n, n \geq 0)$ 是上鞅.

当 $\mathcal{F}_n = \sigma \{X_0, X_1, \dots, X_n\}$ 时, 简称 $X = (X_n, n \ge 0)$ 是鞅 (下鞅、上鞅).

例题 3.1 随机利率折现.

考虑按随机利率增长的资产构成的鞅. 设存款本金为 X_0 , 按日固定利率 r 增值, 则

$$X_{n+1} = (1+r)X_n = X_0(1+r)^{n+1}, n = 0, 1, 2, \cdots$$

如果利率为连续复利 σ . 则不管时间连续还是离散. 有

$$X_t = e^{\sigma t} X_0, \quad X_t = e^{\sigma(t-s)} X_s, 0 \le s < t.$$

设 $\{\mathcal{F}_n, n=0,1,2,\cdots\}$ 为子 σ 代数流, 随机连续复利 (对数收益率) 序列 $\{\sigma_n\}$ 关于 $\{\mathcal{F}_n\}$ 适应, 若资金 X_n 满足条件增值关系:

$$\mathbb{E}\left[X_{n+1} \mid \mathcal{F}_n\right] = e^{\sigma_n} X_n,$$

其中 X_0 关于 \mathcal{F}_0 可测; 将 X_n 折现到 0 时刻得

$$Y_n = \exp\left(-\left[\sigma_0 + \sigma_1 + \dots + \sigma_{n-1}\right]\right) X_n,$$

则 $\{Y_n, \mathcal{F}_n, n \geq 0\}$ 为鞅.

事实上, 易见 Y_n 关于 \mathcal{F}_n 可测, 且

$$\mathbb{E}(Y_{n+1} \mid \mathcal{F}_n)$$

$$= \mathbb{E}\left\{\exp\left(-\left[\sigma_0 + \sigma_1 + \dots + \sigma_n\right]\right) X_{n+1} \mid \mathcal{F}_n\right\}$$

$$= \exp\left(-\left[\sigma_0 + \sigma_1 + \dots + \sigma_n\right]\right) \cdot \mathbb{E}\left\{e^{\sigma_n} X_n \mid \mathcal{F}_n\right\}$$

$$= \exp\left(-\left[\sigma_0 + \sigma_1 + \dots + \sigma_n\right]\right) \cdot e^{\sigma_n} X_n$$

$$= \exp\left(-\left[\sigma_0 + \sigma_1 + \dots + \sigma_{n-1}\right]\right) X_n = Y_n$$

例题 3.2 假设 X 是随机变量, $\mathbb{E}|X|<\infty$. 又假设 $(\mathcal{F}_n, n\geq 0)$ 是一列单调不减的子 σ -域, 定义

$$X_n = \mathbb{E}(X \mid \mathcal{F}_n), \quad n \ge 0,$$

那么 $(X_n, \mathcal{F}_n, n \geq 0)$ 是鞅.

定理 3.1 (鞅的基本性质)

假设 $(X_n, \mathcal{F}_n, n \geq 0)$ 是鞅, 那么

- 1. $\mathbb{E}X_n = \mathbb{E}X_0$, $n \ge 0$;
- 2. $\mathbb{E}(X_{n+m} \mid \mathcal{F}_{n-1}) = X_{n-1}$ a.s., $n \ge 1$, $m \ge 0$;
- 3. 如果 $\varphi: \mathbb{R} \to \mathbb{R}$ 是凸函数, 那么 $(\varphi(X_n), \mathcal{F}_n, n > 0)$ 是下鞅.

证明 对第三点进行说明. 当 $\{X_n\}$ 是鞅且 φ 是凸函数时, 由条件 Jensen 不等式,

$$\mathbb{E}\left(\phi\left(X_{n+1}\right)\mid\mathcal{F}_{n}\right) \geq \phi\left[\mathbb{E}\left(X_{n+1}\mid\mathcal{F}_{n}\right)\right] = \phi\left(X_{n}\right),$$

即 $\{\phi(X_n), \mathcal{F}_n, n \geq 0\}$ 是下鞅.

推论 3.1

假设 $(X_n, \mathcal{F}_n, n \ge 0)$ 是下鞅, 如果 $\varphi : \mathbb{R} \to \mathbb{R}$ 是非降凸函数, 那么 $(\varphi(X_n), \mathcal{F}_n, n \ge 0)$ 是下鞅.

定理 3.2 (Doob 下鞅分解定理)

设 $(X_n, \mathcal{F}_n, n \geq 0)$ 是下鞅且 $\mathbb{E}|X_n| < \infty, \forall n$. 则存在关于 $\{\mathcal{F}_n, n \geq 0\}$ 的鞅 $\{M_n\}$ 和随机变量序列 $\{A_n\}$, 满足 $A_0 = 0, A_n \leq A_{n+1}, A_n$ 关于 \mathcal{F}_{n-1} 可测且 $\mathbb{E}(A_n) < \infty, \forall n$, 使得

$$X_n = M_n + A_n, n = 0, 1, 2, \cdots$$

证明 证明: 取 $A_0 = 0$,

$$A_{n+1} = A_n + \left[\mathbb{E} \left(X_{n+1} \mid \mathcal{F}_n \right) - X_n \right]$$
$$= \sum_{k=0}^n \left[\mathbb{E} \left(X_{k+1} \mid \mathcal{F}_k \right) - X_k \right],$$

则易见 A_n 非负、单调增, 关于 \mathcal{F}_{n-1} 可测, 且 $\mathbb{E}|A_n| < \infty$. 因为

$$M_{n+1} = X_{n+1} - A_{n+1} = X_{n+1} - \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) + X_n - A_n,$$

有

$$\mathbb{E}(M_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) - \mathbb{E}(X_{n+1} \mid \mathcal{F}_n) + X_n - A_n$$
$$= X_n - A_n = M_n.$$

在分解中, A_n 关于 \mathcal{F}_{n-1} 可测, 称 $\{A_n, \mathcal{F}_n\}$ 为可料 (predictable) 随机序列.

定理 3.3 (Doob 下鞅最大值不等式)

假设 $(X_n, \mathcal{F}_n, n \ge 0)$ 是非负下鞅, 那么对任意 x > 0,

$$x\mathbb{P}\left(\max_{1\leq k\leq n}X_k>x\right)\leq \mathbb{E}X_n$$

0

证明 令

$$A_1 = \{X_1 > x\}, \quad A_k = \{X_1 \le x, \dots, X_{k-1} \le x, X_k > x\}, \quad 1 \le k \le n.$$

那么 $A_k \subset \mathcal{F}_k, k = 1, 2, \cdots, n$ 且互不相交, 并且

$$\left\{ \max_{1 \le k \le n} X_k > x \right\} = \bigcup_{k=1}^n A_k.$$

由 $X_k = \mathbb{E}(X_k|A_k) > x$ 易知:

$$\int_{A_k} X_k d\mathbb{P} = \mathbb{E}(XA_k) = \mathbb{P}(A_k)\mathbb{E}(X_k|A_k) = \mathbb{P}(A_k)X_k > x\mathbb{P}(A_k)$$

因此

$$x\mathbb{P}\left(\max_{1\leq k\leq n}X_{k}>x\right)=x\sum_{k=1}^{n}\mathbb{P}\left(A_{k}\right)$$

$$\leq\sum_{k=1}^{n}\int_{A_{k}}X_{k}d\mathbb{P}$$

$$\leq\sum_{k=1}^{n}\int_{A_{k}}X_{n}d\mathbb{P}\left(\text{下鞅性质}\right)$$

$$=\int_{\max_{1\leq k\leq n}X_{k}>x}X_{n}d\mathbb{P}$$

$$\leq\mathbb{E}X_{n}.$$

证毕.

定义 3.2 (鞅差序列)

假设 $d = (d_n, n \ge 0)$ 是一列随机变量, $(\mathcal{F}_n, n \ge 0)$ 是 \mathcal{F} 的一列单调不减子 σ — 域. 如果满足

- 1. 对每一个 $n \geq 0$, $d_n \in \mathcal{F}_{n-}$ 可测的;
- 2. 对每一个 $n \geq 0$, $\mathbb{E} |d_n| < \infty$;
- 3. 对每一个 $n \geq 1$,

$$\mathbb{E}\left(d_n\mid\mathcal{F}_{n-1}\right)=0\text{ a.s.,}$$

称 $(d_n, \mathcal{F}_n, n \geq 0)$ 是鞅差序列.

Ŷ 笔记

• 显然, 当 $(X_n, \mathcal{F}_n, n \geq 0)$ 是鞅时, 定义

$$d_n = X_n - X_{n-1}, \quad n \ge 1,$$

那么 $(d_n, \mathcal{F}_n, n \geq 0)$ 是鞅差序列.

• 反过来, 如果 $(d_n, \mathcal{F}_n, n \geq 0)$ 是鞅差序列, 定义

$$X_n = \sum_{k=0}^n d_k, \quad n \ge 0$$

那么 $(X_n, \mathcal{F}_n, n \geq 0)$ 是鞅.

• 更一般地, 假设 $(d_n, \mathcal{F}_n, n \ge 1)$ 是鞅差, 令 $(C_n, n \ge 1)$ 是一列随机变量, $C_n \in \mathcal{F}_{n-1}$. 定义

$$X_0 = x_0, \quad X_n = X_{n-1} + C_n d_n, \quad n \ge 1,$$

那么 $(X_n, \mathcal{F}_n, n \geq 0)$ 是鞅.

定理 3.4 (总收益为鞅定理)

若 $X_n, n \ge 0$ 为 \mathcal{F}_n 上的鞅, H_n 为可料的一列有界随机变量, 则 $(H \cdot X)_n = \sum_{m=1}^n H_m \cdot d_m$ 为 \mathcal{F}_n 上的鞅.

拿 笔记 若用 H_n 表示在时刻为 n 时买入的股票份额, 其只与 n 之前的时刻股价表现有关. (这一决策自然是基于之前的股价表现来决定的) X_n 表示时刻为 n 时的股价. 自然, 总收益可用 $(H \cdot X)_n = \sum_{m=1}^n H_m \cdot d_m$ 表示.

上述定理中, 若加上 $H_n \geq 0$, 则若 X_n 为下鞅, 那么 $(H \cdot X)_n$ 仍为下鞅.

证明 由于对任意的 $n, (H \cdot X)_n \in \mathcal{F}_n, H_n \in \mathcal{F}_{n-1}$.

$$\mathbb{E}\left((H \cdot X)_{n+1} \mid \mathcal{F}_n\right) = (H \cdot X)_n + \mathbb{E}\left(H_{n+1} \left(X_{n+1} - X_n\right) \mid \mathcal{F}_n\right)$$
$$= (H \cdot X)_n + H_{n+1}\mathbb{E}\left(X_{n+1} - X_n \mid \mathcal{F}_n\right)$$
$$= (H \cdot X)_n + H_{n+1}\left(X_n - X_n\right) = (H \cdot X)_n$$

3.2 停时原理和离散鞅收敛定理

定义 3.3 (停时)

假设 $(\mathcal{F}_n, n \ge 0)$ 是一列单调不减子 σ — 域, $T: \Omega \mapsto \mathbb{Z}_+ \bigcup \{\infty\}$ 是一个映射. 如果对每一个 $n \ge 0$,

$$\{T \le n\} \in \mathcal{F}_n$$

称 T 关于 $(\mathcal{F}_n, n \geq 0)$ 是停时.

 $\widehat{\Sigma}$ 笔记 显然, 每个非负整数都是停时; 另外, T 关于 $(\mathcal{F}_n, n \geq 0)$ 是停时当且仅当 $\{T=n\} \in \mathcal{F}_n$; 停时不一定是随机变量, 因为停时可以取 ∞ .

例题 3.3 假设 $X=(X_n,n\geq 0)$ 是离散时间、离散有限状态的 Markov 链. 给定状态 j, 定义

$$T_i = \inf \{ n \ge 1 : X_n = j \}.$$

如果 $X_0 = j$, 那么 T_j 表示 Markov 链从 j 出发以后首次返回到 j 的时刻; 如果 $X_0 \neq j$, 那么 T_j 表示 Markov 链出 发后首次到达 j 的时刻. 令 $\mathcal{F}_n = \sigma \{X_0, \cdots, X_n\}$, 那么 T_j 关于 $(\mathcal{F}_n, n \geq 0)$ 是停时.

命题 3.1

设T是取值于 $\{0,1,2,\cdots,\infty\}$ 的随机变量,则下述三者等价

- 1. $\{T = n\} \in \sigma(X_0, X_1, \cdots, X_n);$
- 2. $\{T \le n\} \in \sigma(X_0, X_1, \cdots, X_n);$
- 3. $\{T > n\} \in \sigma(X_0, X_1, \dots, X_n)$.

证明 只要注意到如下关系,即可证明等价性.

$$\{T \le n\} = \bigcup_{k=0}^{n} \{T = k\}$$

$$\{T > n\} = \Omega - \{T \le n\}$$

$$\{T = n\} = \{T \le n\} - \{T \le n - 1\}$$

引理 3.1

如果 T, S 关于 $(\mathcal{F}_n, n \ge 0)$ 是停时, 那么

$$\min\{T,S\},\max\{T,S\},T+S$$

都关于 $(\mathcal{F}_n, n \geq 0)$ 是停时.

证明 对任意 $n \ge 0$, 根据 $(\mathcal{F}_n, n \ge 0)$ 的单调性,

$$\{T+S=n\} = \sum_{k=0}^{n} \{T=k\} \cap \{S=n-k\} \in \mathcal{F}_n,$$

所以T+S关于 $(\mathcal{F}_n, n \geq 0)$ 是停时. 其余类似证明.

假设 $(X_n, n \ge 0)$ 是一列随机变量, T 关于 $(\mathcal{F}_n, n \ge 0)$ 是有限停时. 定义

$$X_T(\omega) = X_{T(\omega)}(\omega), \quad \omega \in \Omega.$$

注意到, 对任意 Borel 集 B,

$$\{\omega: X_T(\omega) \in B\} = \sum_{n=0}^{\infty} \{\omega: X_T(\omega) \in B, T(\omega) = n\} \in \mathcal{F},$$

因此 X_T 是随机变量.

定理 3.5 (停时定理)

假设 $(X_n, \mathcal{F}_n, n \ge 0)$ 是鞅, T 关于 $(\mathcal{F}_n, n \ge 0)$ 是停时.

- 1. 定义 $Y_n = X_{T \wedge n}$, 那么 $(Y_n, \mathcal{F}_n, n \geq 0)$ 是鞅;
- 2. 如果T是有界停时,那么

$$\mathbb{E}X_T = \mathbb{E}X_0$$

3. 如果 T 是有限停时, $(X_n, \mathcal{F}_n, n \geq 0)$ 是有界鞅, 那么

$$\mathbb{E}X_T = \mathbb{E}X_0$$

C

证明 由于 $\{T > n\} \in \mathcal{F}_n$, 所以

$$\{Y_n \in B\} = \bigcup_{k=0}^n \{X_k \in B, T = k\} \bigcup \{X_n \in B, T > n\} \in \mathcal{F}_n$$

再证明 $\mathbb{E}|Y_n|<\infty$. 事实上,

$$\mathbb{E} |Y_n| = \mathbb{E} |X_{T \wedge n}|$$

$$= \sum_{k=0}^{n-1} \mathbb{E} |X_k| \mathbf{1}_{(T=k)} + \mathbb{E} |X_n| \mathbf{1}_{(T \geq n)}$$

$$\leq \sum_{k=0}^{n} \mathbb{E} |X_k| < \infty$$

往下验证

$$\mathbb{E}\left(Y_n \mid \mathcal{F}_{n-1}\right) = Y_{n-1} \quad \text{a.s.}$$

任意给定 $A \in \mathcal{F}_{n-1}$, 由于 $\{T \ge n\} \in \mathcal{F}_{n-1}$, 那么

$$\begin{split} \int_A Y_n d\mathbb{P} &= \int_A X_{T \wedge n} d\mathbb{P} \\ &= \sum_{k=0}^{n-1} \int_{A,T=k} X_k d\mathbb{P} + \int_{A,T \geq n} X_n d\mathbb{P} \\ &= \sum_{k=0}^{n-1} \int_{A,T=k} X_k d\mathbb{P} + \int_{A,T \geq n} X_{n-1} d\mathbb{P} \quad (\text{由于 } X_n \text{ 是鞅}) \\ &= \int_A Y_{n-1} d\mathbb{P} = Y_{n-1}. \end{split}$$

因此 $(Y_n, \mathcal{F}_n, n \geq 0)$ 是鞅.

如果 T 是有界停时, 那么存在 $m \ge 0$ 使得 $T \le m$ a.s., 由 (1) 得

$$\mathbb{E}X_T = \mathbb{E}X_{T \wedge m} = \mathbb{E}X_0.$$

如果 X_n 是有界鞅, 那么存在 M>0 使得 $|X_n|\leq M$ a.s., 由 (1) 和控制收敛定理得

$$\mathbb{E}X_T = \mathbb{E}\lim_{n \to \infty} X_{T \wedge n} = \lim_{n \to \infty} \mathbb{E}X_{T \wedge n} = \mathbb{E}X_0$$

定理证毕.

定理 3.6 (离散鞅收敛定理)

假设 $(X_n, \mathcal{F}_n, n \geq 0)$ 是下鞅. 如果:

$$\sup_{n\geq 0} \mathbb{E} \left| X_n \right| < \infty$$

那么 $X_{\infty} = \lim_{n \to \infty} X_n$ a.s. 存在且有限.

详细地论证这个定理,需要一些前置概念,

定义 3.4 (上穿越数)

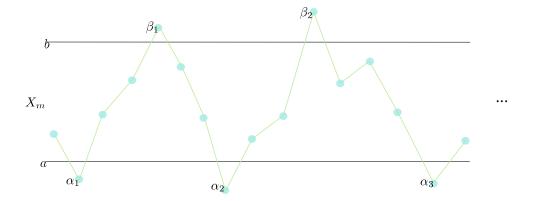
对鞅 $\{X_m, m \ge 0\}$, 给定一个区间 (a, b), 令 $\beta_0 = 0$. 对任意 $k \ge 1$, 令

$$\alpha_k = \inf \{ m > \beta_{k-1} \mid X_m \le a \}, \quad \beta_k = \inf \{ m > \alpha_k \mid X_m \ge b \}$$

称 $U_n = \sup \{k \mid \beta_k \le n\}$ 为到时刻 n 为止区间 (a,b) 的上穿越数.

💡 筆记 下面的图非常直观地给出了上穿越数的含义.

理解:假如我们希望股价为a时买入,b时卖出,如果一次买入一次卖出当作一次投资周期,则上穿越数就表示投资周期的次数.



定理 3.7 (上穿越不等式)

对下鞅 $\{X_m, \mathcal{F}_m\}, m \geq 0$, 区间 (a,b) 的上穿越数 U_n 满足如下不等式:

$$(b-a)\mathbb{E}U_n \le \mathbb{E}(X_n-a)^+ - \mathbb{E}(X_0-a)^+$$

证明 需要用到定理3.4. 首先要构造一个可料随机变量列. 由于

$$\{\alpha_1 = n\} = \bigcap_{m=1}^{n-1} \{X_m > a\} \cap \{X_n \le a\} \in \mathcal{F}_n$$

$$\{\beta_1 = n\} = \bigcap_{m=1}^{n-1} \{\alpha_1 = m, X_k < b, k = m+1, \dots, n\} \cap \{X_n \ge b\} \in \mathcal{F}_n.$$

以此类推, 归纳可得, 对任意的 k, α_k , β_k 均为停时.

对任意的 n, 当存在 k 使得 $\alpha_k < n \le \beta_k$ 时令 $H_n = 1$. 由于

$$\{\alpha_k < n \le \beta_k\} = \{\alpha_k \le n - 1\} \cap \{\beta_k \le n - 1\}^c \in \mathcal{F}_{n-1}$$

故 H_n 是可料的. 令 $Y_m = a + (X_m - a)^+$. 易知 Y_m 为下鞅 (由于凸函数).

下面证明 $(b-a)U_n < (H \cdot Y)_n$, 其中 $(H \cdot Y)_n = \sum_{m=1}^n H_m \cdot (Y_m - Y_{m-1})$. 由定义,

$$H_n = \sum_{k=1}^{U_n} \mathbf{1}_{\{\alpha_k < n \le \beta_k\}}$$

则

$$(H \cdot Y)_n = \sum_{m=1}^n H_m (Y_m - Y_{m-1})$$

$$= \sum_{m=1}^n \sum_{k=1}^{U_n} \mathbf{1}_{\{\alpha_k < m \le \beta_k\}} (Y_m - Y_{m-1})$$

$$= \sum_{k=1}^{U_n} \sum_{m=1}^n \mathbf{1}_{\{\alpha_k < m \le \beta_k\}} (Y_m - Y_{m-1})$$

$$= \sum_{k=1}^{U_n} (Y_{\beta_k} - Y_{\alpha_k}) \ge (b - a)U_n$$

令 $K_n = 1 - H_n$, 其为可料的随机变量列, 由于

$$(H \cdot Y)_n + (K \cdot Y)_n = \sum_{n=1}^{n} (Y_m - Y_{m-1}) = Y_n - Y_0$$

由定理3.4, $(K \cdot Y)_n$ 为下鞅, 由下鞅的性质,

$$\mathbb{E}(K \cdot Y)_n \geq \mathbb{E}(K \cdot Y)_0 = 0;$$

故 $\mathbb{E}(H \cdot Y)_n \leq \mathbb{E}(Y_n - Y_0)$. 结合上面证得的 $(b - a)U_n \leq (H \cdot Y)_n$ 与 Y_n 的定义知定理成立.

下面给出离散鞅收敛定理的证明.

证明 不难看出对任意的随机变量 $X, (X-a)^+ \le X^+ + |a|$. 由上穿越不等式, 对任意区间 (a,b), 其上穿数 U_n 满足

$$(b-a)\mathbb{E}U_n \le \mathbb{E}(X_n-a)^+ - \mathbb{E}(X_0-a)^+ \le \mathbb{E}(X_n-a)^+ \le \mathbb{E}X_n^+ + |a|$$

由单调收敛定理, 上穿越数 U_n 必然几乎处处收敛到一个随机变量 U. 由 $\sup_n \mathbb{E} X_n^+ < \infty$, 可知 $\mathbb{E} U = \sup_n \mathbb{E} U_n < \infty$, 故 $U < \infty$ 几乎处处成立.

根据 U 的定义, $U < \infty$ 相当于当 n 充分大时, X_n 要么都落在 $(-\infty,b)$ 中, 要么都落在 (a,∞) 中. (即是说 X_n 不会 "穿过" (a,b) 区间); 于是对任意的 a,b, \mathbb{P} (lim inf $X_n < a < b < \limsup X_n$) = $\mathbb{P}(U = \infty) = 0$. 故

$$\mathbb{P}\left(\bigcup_{a,b\in\mathbb{Q}} \{\liminf X_n < a < b < \limsup X_n\}\right) = 0$$

于是 $\liminf X_n = \limsup X_n$ 几乎处处成立, 故 $\lim X_n$ 几乎处处存在, 将其记为 X.

最后说明 $\mathbb{E}X < \infty$,由 Fatou 引理, $\mathbb{E}X^+ \leq \liminf \mathbb{E}X_n^+ < \infty$ 另一方面,由于 X_n 为下鞅, $\mathbb{E}X_n^- = \mathbb{E}X_n^+ - \mathbb{E}X_n \leq \mathbb{E}X_n^+ - \mathbb{E}X_0$. 故再由 Fatou 引理有 $\mathbb{E}X^- \leq \liminf \mathbb{E}X_n^- \leq \sup \mathbb{E}X_n^+ - \mathbb{E}X_0 < \infty$. 于是 $\mathbb{E}X < \infty$.

推论 3.2

有随机过程 $(X_n, \mathcal{F}_n, n \geq 0)$:

- 1. 如果为下鞅, 且 $X_n \le 0, n \ge 0$;
- 2. 如果为上鞅, 且 $X_n \ge 0, n \ge 0$;

则 $X_{\infty} = \lim_{n \to \infty} X_n$, a.s. 存在且有限.

 \Diamond

🕏 笔记 鞅收敛定理类比"单调有界数列收敛";有上界的下鞅收敛,有下界的上鞅收敛.

3.3 连续时间鞅介绍

定义 3.5 (连续鞅)

设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是概率空间, T 是固定的正数, $\mathcal{F}(t)$, $0 \le t \le T$ 是 \mathcal{F} 的子 σ — 代数的域流. 考虑一个适应的随机过程 M(t), $0 \le t \le T$:

1. 如果:

$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] = M(s), \quad \forall 0 \le s \le t \le T$$

则我们称 M(t) 是一个鞅——鞅没有上升或者下降的趋势;

2. 如果:

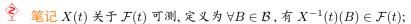
$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] \ge M(s), \quad \forall 0 \le s \le t \le T$$

则我们称 M(t) 是一个下鞅——下鞅没有下降的趋势, 但可能有上升趋势;

3. 如果:

$$\mathbb{E}[M(t) \mid \mathcal{F}(s)] \le M(s), \quad \forall 0 \le s \le t \le T$$

则我们称 M(t) 是一个上鞅——上鞅没有上升的趋势, 但可能有下降趋势.



 $\sigma(X(u), 0 \le u \le t)$ 表示由 $\{X(u), 0 \le u \le t\}$ 生成的 σ 代数,即包含一切形如 $\{X(u) \le x\}$ ($0 \le u \le t, x \in \mathbb{R}$) 的事件的最小 σ 代数. 这个定义可以推广到X(t) 取值于 \mathbb{R}^d 的情形.

例题 3.4 设 $\{Y_t, t \geq 0\}$ 是零初值具有平稳独立增量的随机过程. 令

$$X_t = X_0 e^{Y_t},$$

其中 X_0 为一常数. 若 $\mathbb{E}\left[e^{Y_t}\right]=1,$ 则 $\{X_t,t\geq 0\}$ 是一个鞅. 证明 事实上

 $\mathbb{E}\left[|X_t|\right] = |X_0| \,\mathbb{E}\left[e^{Y_t}\right] = |X_0| < \infty,$

再对 $0 \le s < t$, 有

$$\begin{split} \mathbb{E}\left[X_t \mid X_r, 0 \leq r \leq s\right] &= \mathbb{E}\left[X_s e^{Y_t - Y_s} \mid X_r, 0 \leq r \leq s\right] \\ &= X_s \mathbb{E}\left[e^{Y_t - Y_s}\right] \\ &= X_s \mathbb{E}\left[e^{Y_t - s}\right] = X_s, \text{ a.s.} \end{split}$$

注: Y(t) 的一个取法是取为带漂移的布朗运动

$$Y(t) = -\frac{1}{2}t + W(t).$$

最后不加证明给出极限定理.

定理 3.8 (鞅极限定理)

设 $\{X_t, t \geq 0\}$ 是一个鞅并且 $X_t \geq 0$, $\forall t \geq 0$, 则存在几乎处处收敛的有限极限,即有

$$\lim_{t\to\infty}X_t=X_\infty<\infty, \text{ a.s.}$$

\

第4章 布朗运动

内容提要

- □ 高斯过程
- □ 随机游动
- □ 布朗运动
- □ 二次变差

- □ 马尔可夫性
- □ 首达时间分布
- □ 反射原理

4.1 高斯过程再讨论

定义 4.1 (高斯过程)

随机过程 $\{X_t\}_{t \geq 0}$ 被称为高斯过程, 如果对任意 $0 \leq t_1 < t_2 < \cdots < t_k$, 其有限维分布 μ_{t_1,\dots,t_k} 都为联合正态分布.

命题 4.1

一旦其均值函数 $\mu_X(t)$ 和协方差函数 K(s,t) 确定,则高斯过程就被唯一确定.

正态随机向量 $\mathbf{X} = (X_1, X_2, \dots, X_n)^{\mathsf{T}}$ 完全由它的一阶矩和二阶矩决定,

$$m = \mathbb{E}X, \quad K = \mathbb{E}(X - m)(X - m)^{\top},$$

或者 $m_i = \mathbb{E}X_i$ 和 $K_{ij} = \mathbb{E}(X_i - m_i)(X_j - m_j)$. 通过 m 和 K, 可以写出 X 的特征函数:

$$\mathbb{E}e^{it\cdot X} = e^{it\cdot m - \frac{1}{2}t^{\top}Kt}.$$

因此, 给定任何 $0 \le t_1 < t_2 < \dots < t_k$, 高斯过程的有限维分布 μ_{t_1,\dots,t_k} 是唯一由均值 $\boldsymbol{m} = (m(t_1),\dots,m(t_k))$ 和协方差矩阵 $\boldsymbol{K} = (K(t_i,t_j))_{i,j=1}^k$ 决定的.

4.2 随机游动

接连抛掷一枚均匀地硬币, 出现正面、背面的结果均为 $\frac{1}{2}$. 结果记为 $\omega_1, \omega_2, \cdots$, 设

$$X_j = \begin{cases} 1, & \text{如果}\omega_j = \mathbb{E}\mathbb{I}\\ -1, & \text{如果}\omega_j = \mathbb{G}\mathbb{I} \end{cases}$$

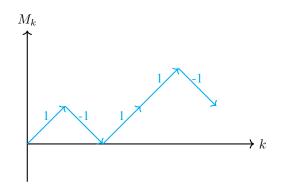
定义

$$M_0 = 0, \quad M_k = \sum_{j=1}^k X_j, k = 1, 2, \cdots$$

则过程 $\{M_k\}_{k\in\mathbb{N}}$ 是一个**对称随机游动**¹. 对应于每一次抛掷, 它上升或者下降一个单位, 并且两者的可能性相同. 下图展示一个可能的趋势图. N 步以后, 可能取值为 $-N, -N+2, \cdots, N-2, N,$ 令 $W(m,N)=\mathbb{P}(M_N=m),$ 则 W(m,N) 服从二项分布:

$$W(m,N) = \frac{N!}{\left(\frac{N+m}{2}\right)! \left(\frac{N-m}{2}\right)!} \left(\frac{1}{2}\right)^N$$

 $^{^{1}}$ 一个粒子在数轴上从原点开始做运动, 每单位时间只能走一步, 且分别以二分之一的概率向左或向右. 则 M_{k} 表示 k 时粒子所处的坐标位置



定义 4.2 (扩散系数 (Diffusion Coefficient))

扩散系数被定义为

$$D = \frac{\left\langle \left(M_N - M_0 \right)^2 \right\rangle}{2N}$$

一维对称随机游动的扩散系数为 1/2.

命题 4.2

对称随机游动是独立增量过程: 对于非负整数 $0 = k_0 < k_1 < \cdots < k_m$, 随机变量

$$M_{k_1} = (M_{k_1} - M_{k_0}), (M_{k_2} - M_{k_1}), \cdots, (M_{k_m} - M_{k_{m-1}})$$

是两两独立的.

每个增量 $M_{k_{i+1}} - M_{k_i}$ 具有期望值 0 和方差 $k_{i+1} - k_i$. 对称随机游动在单位时间内累积的方差为 1 ,于是对任何非负整数 k < t,从 k 到 t 时间区间内增量的方差为 t - k.

证明

$$\operatorname{Var}\left(M_{k_{i+1}} - M_{k_i}\right) = \sum_{j=k_i+1}^{k_{i+1}} \operatorname{Var}\left(X_j\right) = \sum_{j=k_i+1}^{k_{i+1}} 1 = k_{i+1} - k_i$$

命题 4.3

对称随机游动是鞅.

证明 对任何非负整数 k < t, 有

$$\begin{split} \mathbb{E}\left[M_t \mid \mathcal{F}_k\right] &= \mathbb{E}\left[(M_t - M_k) + M_k \mid \mathcal{F}_k\right] \\ &= \mathbb{E}\left[M_t - M_k \mid \mathcal{F}_k\right] + \mathbb{E}\left[M_k \mid \mathcal{F}_k\right] \\ &= \mathbb{E}\left[M_t - M_k \mid \mathcal{F}_k\right] + M_k \\ &= \mathbb{E}\left[M_t - M_k\right] + M_k = M_k \end{split}$$

命题 4.4

对称随机游动的二次变差等于其方差.

$$[M, M]_k = \sum_{j=1}^k (M_j - M_{j-1})^2 = k = \text{Var}(M_k)$$

接下来讨论随机游动在合适缩放下的连续极限.

定义 4.3 (按比例缩小型随机游动)

固定正整数 n, 定义

$$W^{(n)}(t) = \frac{1}{\sqrt{n}} M_{nt}$$

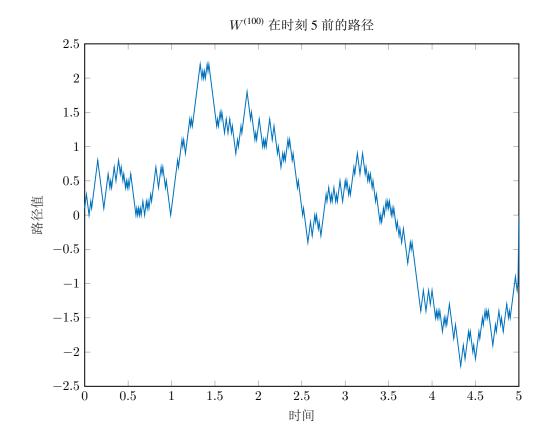
其中nt 是整数. 如果nt 不是整数,取t 左右两边最邻近的s 和u 使得ns 和nu 是整数,然后利用s 和u 处的值,通过插值定义 $W^{(n)}(t)$,

$$W^{(n)}(t) = \frac{1}{\sqrt{n}}(\theta M_k + (1 - \theta)M_{k+1})$$

其中 $nt \in (k, k+1], k = 0, 1, \dots, n-1, \theta = \lceil nt \rceil^{a} - nt$.

" $[\cdot]$: 天花板取整. 对 $m \in \mathbb{Z}$, 若 $x \in [m, m+1)$, [x] = m+1. $[\cdot]$: 地板取整. 对 $m \in \mathbb{Z}$, 若 $x \in [m, m+1)$, [x] = m

利用 Matlab 代码模拟的 $W^{(100)}$ 在前 5 个时刻的路径如下:



可以验证按比例缩小型随机游动依然具有独立增量性、鞅性质. 重要的是接下来的定理.

定理 4.1 (中心极限定理)

固定 $t \ge 0$, 按比例缩小型随机游动 $W^{(n)}(t)$ 在时刻 t 取值的分布当 $n \to \infty$ 时收敛于正态分布 N(0,t).

 \odot

证明 N(0,t) 的矩母函数是

$$\varphi(u) = \exp\left(\frac{1}{2}u^2t\right)$$

如果 t 使得 nt 为整数,则 $W^{(n)}(t)$ 的矩母函数为:

$$\varphi_n(u) = \mathbb{E} \exp\left\{uW^{(n)}(t)\right\} = \mathbb{E} \exp\left\{\frac{u}{\sqrt{n}}M_{nt}\right\} = \mathbb{E} \exp\left\{\frac{u}{\sqrt{n}}\sum_{j=1}^{nt}X_j\right\} = \mathbb{E} \prod_{j=1}^{nt} \exp\left\{\frac{u}{\sqrt{n}}X_j\right\}$$

根据独立性,有

$$\mathbb{E} \prod_{j=1}^{nt} \exp\left\{\frac{u}{\sqrt{n}} X_j\right\} = \prod_{j=1}^{nt} \mathbb{E} e^{\frac{u}{\sqrt{n}} X_j} = \prod_{j=1}^{nt} \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}}\right) = \left(\frac{1}{2} e^{\frac{u}{\sqrt{n}}} + \frac{1}{2} e^{-\frac{u}{\sqrt{n}}}\right)^{nt}$$

下面证明

$$\lim_{n \to \infty} \varphi_n(u) = \phi(u)$$

为此,取对数,并做变换 $x=\frac{1}{\sqrt{n}}$.于是

$$\lim_{n \to \infty} \log \varphi_n(u) = t \lim_{x \downarrow 0} \frac{\log \left(\frac{1}{2} e^{ux} + \frac{1}{2} e^{-ux}\right)}{x^2} \xrightarrow{\underline{L'Hopital}} \frac{t}{2} \lim_{x \downarrow 0} \frac{\frac{u}{2} e^{ux} - \frac{u}{2} e^{-ux}}{x} = \frac{1}{2} u^2 t$$

证毕.

基于对称随机游动的渐近分析,可以推导出本章中要介绍的布朗运动的许多性质.一个典型的例子是布朗轨迹的反正弦定律.

引理 4.1 (近似方程)

当 $n \gg 1, m \ll n$ 时,

$$W(m,n) = \frac{n!}{\left(\frac{n+m}{2}\right)! \left(\frac{n-m}{2}\right)!} \left(\frac{1}{2}\right)^n \approx \left(\frac{2}{\pi n}\right)^{\frac{1}{2}} \exp\left(-\frac{m^2}{2n}\right)$$

证明 由 Stirling 公式, 对 $n \gg 1$, 有

$$\log n! = \left(n + \frac{1}{2}\right) \log n - n + \frac{1}{2} \log 2\pi + O\left(n^{-1}\right)$$

因此有:

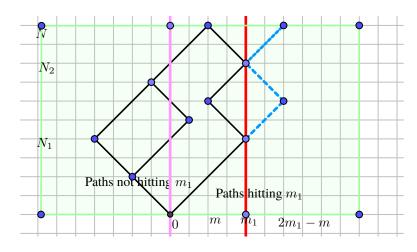
$$\begin{split} \log W(m,n) &\approx \left(n + \frac{1}{2}\right) \log n - \frac{1}{2}(n+m+1) \log \left[\frac{n}{2}\left(1 + \frac{m}{n}\right)\right] \\ &- \frac{1}{2}(n-m+1) \log \left[\frac{n}{2}\left(1 - \frac{m}{n}\right)\right] - \frac{1}{2} \log 2\pi - n \log 2. \end{split}$$

又因为 $m \ll n$, 利用 $\log(1+x) = x - \frac{1}{2}x^2 + O(x^3), x \ll 1$, 有:

$$\log W(m, n) \approx -\frac{1}{2} \log n + \log 2 - \frac{1}{2} \log 2\pi - \frac{m^2}{2n}$$

- * 带有反射和吸收壁的随机游动
 - 在 $m=m_1>0$ 处拥有反射壁

记 $W(m, N; m_1)$ 表示粒子在 N 步时抵达 $m(\leq m_1)$ 的概率. 我们可以通过下图分析问题. 从上图看出, 实际



样本轨道是实线部分 (包括反射路径), 跨过 m_1 的不受限制的随机游动路径由蓝色虚线表示. 所有路径分为

两类, 一类是没有撞击 m_1 最终抵达 m 的路径, 另一类是在 N 之前撞击 m_1 然后最终抵达 m(或那些不受限制的路径就是抵达 $2m_1-m$). 我们有如下结论 (反射原理):

- 1. 在不受限制场合, 所有样本曲线拥有同等概率 🗓
- 2. 撞击 m_1 的有反射路径的概率等于在不受限制场合中, 经过 m_1 并且最终抵达 m 和 $2m_1 m$ 的概率之 和.

$$W_r(m, N; m_1) = W(m, N) + W(2m_1 - m, N)$$

- 3. 撞击 m_1 最终到 m 的路径个数等于撞击 m_1 最终到 $2m_1 m$ 的路径个数.
- 在 $m = m_1 > 0$ 处拥有吸收壁 吸收的含义是抵达 m_1 后路径结束. 类似地分析, 我们有:

$$W_a(m, N; m_1) = W(m, N) - W(2m_1 - m, N)$$

定义首次撞击概率 $a(m_1, N) = \mathbb{P}(X_N = m, X_n < m_1, \forall n < N)$, 我们有²:

$$a(m_1, N) = \frac{1}{2}W_a(m_1 - 1, N - 1; m_1) = \frac{m_1}{N}W(m_1, N)$$

定义 $P_{2k,2n}$ 为粒子在区间 [0,2n] 内有 2k 单位时间在正侧的概率³. 则有

引理 4.2

$$P_{2k,2n} = \mathbb{P}(M_{2k} = 0)\mathbb{P}(M_{2n-2k} = 0) \stackrel{\triangle}{=} u_{2k} \cdot u_{2n-2k}.$$

证明 定义首到时为 $\sigma_{2n}=\min\{1\leq k\leq 2n:M_k=0\}$, 并且对于 $1\leq k\leq 2n$ 时若 $M_k\neq 0$ 则令 $\sigma_{2n}=+\infty$. 定义

$$f_{2k} = \mathbb{P}(\sigma_{2n} = 2k)$$

显然

$$u_{2k} = \binom{2k}{k} 2^{-2k}$$

根据反射原理,有

$$f_{2k} = 2 \cdot \frac{1}{2} \cdot \frac{1}{2k-1} W(1, 2k-1) = \frac{1}{2k} u_{2k-2} = u_{2k-2} - u_{2k}$$

往证引理. 首先 k=0 时必然成立. 假设拥有一个路径满足 $M_{2n}=0$, $\min_{0\leq k\leq 2n}M_k=-m(m>0)$. 定义 $l=\min\{k|M_k=-m\}$. 我们能将这条路径映射到一条只位于正边的路径上: 对路径 $\{M_k\}_{0\leq k\leq l}$ 做关于 t=l 的反射处理, 定义新路径 $\{\widetilde{M_k}\}_{0\leq k\leq l}$ 满足 $\widetilde{M_k}=2m+M_{l-k}$. 调整新路径的起点, 把 $\widetilde{M_0}$ 和点 (2n,0) 连接起来. 通过这样的操作, 我们得到一个全部在正侧并且最终抵达 (2n,2m) 的路径.

接下来证明如下结论:

$$u_{2k} = \sum_{r=1}^{k} f_{2r} u_{2(k-r)}$$

既然 $\{M_{2k} = 0\} \subset \{\sigma_{2n} \leq 2k\}$, 则我们有

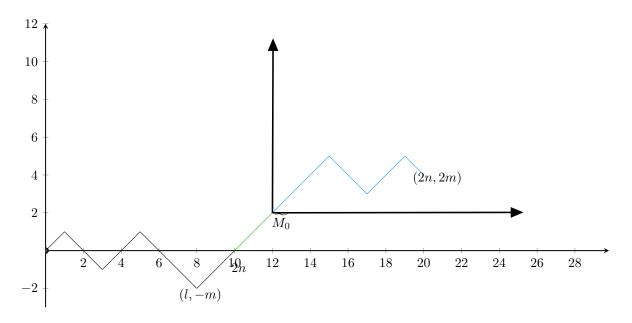
$$\{M_{2k} = 0\} = \{M_{2k} = 0\} \cap \{\sigma_{2n} \le 2k\} = \sum_{n=1}^{k} \{M_{2k} = 0\} \cap \{\sigma_{2n} = 2r\}$$

因此

$$u_{2k} = \mathbb{P}(M_{2k} = 0) = \sum_{r=1}^{k} \mathbb{P}(M_{2k} = 0, \sigma_{2n} = 2r)$$
$$= \sum_{r=1}^{k} \mathbb{P}(M_{2k} = 0 \mid \sigma_{2n} = 2r) \mathbb{P}(\sigma_{2n} = 2r)$$

²注意到 $W(m,N) = \frac{N}{N+m}W(m-1,N-1)$

 $^{^{3}}$ 如果 M_{n-1}, M_{n} 至少有一个为正, 就称粒子在 [n-1, n] 上处于正侧



然而

$$\mathbb{P}(M_{2k} = 0 \mid \sigma_{2n} = 2r) = \mathbb{P}(M_{2k} = 0 \mid M_1 \neq 0, \dots, M_{2r-1} \neq 0, M_{2r} = 0)$$

$$= \mathbb{P}(M_{2r} + (\xi_{2r+1} + \dots + \xi_{2k}) = 0 \mid M_1 \neq 0, \dots, M_{2r-1} \neq 0, M_{2r} = 0)$$

$$= \mathbb{P}(M_{2r} + (\xi_{2r+1} + \dots + \xi_{2k}) = 0 \mid M_{2r} = 0)$$

$$= \mathbb{P}(\xi_{2r+1} + \dots + \xi_{2k} = 0) = \mathbb{P}(M_{2(k-r)} = 0)$$

综合上述式就可以得到 $u_{2k} = \sum_{r=1}^{k} f_{2r} u_{2(k-r)}$.

接下来利用归纳法证明引理结论. 令 $1 \le k \le n-1$. 如果粒子恰好有 2k 时刻处于正侧,则路径必然穿越 0. 令 2r 为粒子首次穿越 0 的时间. 注意, 这有两种可能性:

$$M_k \ge 0, k \le 2r, \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ M_k \le 0, k \le 2r.$$

第一种可能性包含的路径总数有

$$\left(2^{2r} \cdot \frac{1}{2} f_{2r}\right) \cdot \left(2^{2(n-r)} \cdot P_{2(k-r),2(n-r)}\right) = \frac{1}{2} 2^{2n} f_{2r} P_{2(k-r),2(n-r)}$$
$$\frac{1}{2} 2^{2n} f_{2r} P_{2k,2(n-r)}.$$

因此

第二种,有

$$P_{2k,2n} = \frac{1}{2} \sum_{r=1}^{k} f_{2r} P_{2(k-r),2(n-r)} + \frac{1}{2} \sum_{r=1}^{k} f_{2r} P_{2k,2(n-r)}$$

假设 $P_{2k,2m} = u_{2k} \cdot u_{2m-2k}$ 对 $m = k, k+1, \cdots, n-1$ 成立, 则有

$$P_{2k,2n} = \frac{1}{2}u_{2n-2k} \sum_{r=1}^{k} f_{2r}u_{2k-2r} + \frac{1}{2}u_{2k} \sum_{r=1}^{k} f_{2r}u_{2n-2k-2r}$$
$$= \frac{1}{2}u_{2n-2k}u_{2k} + \frac{1}{2}u_{2k}u_{2n-2k} = u_{2k}u_{2n-2k}.$$

记 $\gamma(2n)$ 为在区间[0,2n] 上处于正侧的总时间(次数).则有:

$$\mathbb{P}\left(\frac{1}{2} < \frac{\gamma(2n)}{2n} \le x\right) = \sum_{k,1/2 < 2k/2n \le x} P_{2k,2n}$$

由近似方程4.1, 当 $k, n-k \to \infty$ 时, 我们有:

$$u_{2k} \sim \frac{1}{\sqrt{\pi k}}, \quad P_{2k,2n} \sim \frac{1}{\pi \sqrt{k(n-k)}}$$

因此

$$\sum_{\{k,1/2 < 2k/2n \le x\}} P_{2k,2n} - \sum_{k,1/2 < 2k/2n \le x} \frac{1}{\pi n} \cdot \left[\frac{k}{n} \left(1 - \frac{k}{n} \right) \right]^{-\frac{1}{2}} \to 0, \quad n \to \infty$$

显然得到

$$\sum_{\{k,1/2<2k/2n\leq x\}} P_{2k,2n} \to \frac{1}{\pi} \int_{\frac{1}{2}}^{x} \frac{dt}{\sqrt{t(1-t)}}$$

总结为如下定理:

定理 4.2 (反正弦定律)

对称随机游动的粒子在正侧的时间比率不超过x的概率趋向于反正弦函数. 即当 $n \to \infty$ 时,

$$\sum_{\{k,k/n \le x\}} P_{2k,2n} \to \frac{2}{\pi} \arcsin \sqrt{x}$$

 \Diamond

笔记 这个定理告诉了我们关于对称随机游走的一个有点违反直觉的事实. 这表明我们将观察到, 在一场公平的投币游戏中, 一方持续在获胜一方(正侧), 而另一方在失败一方(负侧), 而不是频繁地交换获胜状态.

4.3 布朗运动

引入布朗运动的定义之前, 先极限化随机游动⁴. 回到 1905 年 Einstein 论文, 假定 (花粉) 粒子的运动是由分子的热运动引起的不规则运动, Einstein 做了如下的一些假定:

- 1. 每一个粒子所进行的运动, 同其他一切粒子的运动都无关;
- 2. 同一个粒子在各个不同时间间隔中的运动,都必须被看作是相互独立的过程,只要我们设想选择的这些时间间隔不要太小.即粒子运动是独立增量过程. 做这样的假定是基于观测时间间隔在宏观尺度上小,但在微观尺度大,因此可以忽略掉一些在非常小尺度上的物理细节.
- 3. 假设液体中共有 n 个粒子, 经过时间间隔 τ , 单个粒子的 X 坐标将要增加 Δ , 此处 Δ 对于每个粒子都有一个不同的值 (可正可负), 对于 Δ , 它满足某种分布定律. 利用统计学方法, 要估计某个粒子在 τ 后处于 Δ 和 Δ + d Δ 之间的概率, 就用处于该区间的粒子数和总粒子数的比来估计. 在时间间隔 τ 内经历了处于 Δ 和 Δ + d Δ 之间位移的粒子数 dn, 可由如下形式的方程表示

$$\frac{\mathrm{d}n}{n} = \varphi(\Delta)\mathrm{d}\Delta, \quad \text{ if } \int_{-\infty}^{+\infty} \varphi(\Delta)\mathrm{d}\Delta = 1.$$

而 φ 只是对非常小的 Δ 值才不是零, 并且满足条件 $\varphi(\Delta) = \varphi(-\Delta)$.

等式右边的 $\varphi(\Delta)d\Delta$ 表示一个粒子在经历 τ 后处于 Δ 到 $\Delta+d\Delta$ 之间的概率, 那么应该等在时间间隔 τ 内经历了处于 Δ 和 $\Delta+d\Delta$ 之间位移的粒子数占总粒子数的比例 (即等式左边), 而这个比例正比于区间间隔 $d\Delta$, 比例系数与区间的位置有关.

4. 假设单位体积的粒子数 ν 只同 x 和 t 有关, 设

$$\nu = f(x, t).$$

也就是说处于 x 位置, 在时刻 t 粒子数的密度是 f(x,t).

⁴只考虑一个粒子的随机游动,但为了分析问题,引入 n 个粒子进行统计分析,把概率转化为比率

我们要从粒子在 t 时刻分布计算出 $t+\tau$ 时刻的分布. 由 $\varphi(\Delta)$ 的定义可以计算出 $t+\tau$ 时刻位于 X 轴上 $[x,x+\mathrm{d}x]$ 之间的粒子数,

$$f(x, t + \tau) dx = \int_{-\infty}^{+\infty} f(x - \Delta, t) dx \varphi(\Delta) d\Delta$$
$$= \int_{-\infty}^{+\infty} f(x - \Delta, t) \varphi(\Delta) d\Delta dx$$
$$= \int_{-\infty}^{+\infty} f(x + \Delta, t) \varphi(\Delta) d\Delta dx$$

利用泰勒展开, 7 很小, 因此可以设

$$f(x, t + \tau) = f(x, t) + \tau \frac{\partial f}{\partial t}$$
.

对 $f(x + \Delta, t)$ 关于 x 变量用泰勒展开

$$f(x+\Delta,t) = f(x,t) + \Delta \frac{\partial f(x,t)}{\partial x} + \frac{\Delta^2}{2} \frac{\partial^2 f(x,t)}{\partial x^2} + \cdots$$
$$f(x,t) + \frac{\partial f}{\partial t} \tau$$
$$= f(x,t) \cdot \int_{-\infty}^{+\infty} \varphi(\Delta) d\Delta + \frac{\partial f}{\partial x} \int_{-\infty}^{+\infty} \Delta \varphi(\Delta) d\Delta + \frac{\partial^2 f}{\partial x^2} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \varphi(\Delta) d\Delta + \cdots \quad (A)$$

既然 $\varphi(\Delta) = \varphi(-\Delta)$, 则上面偶数项为 0. 因为只有很小的 Δ 值才对积分有贡献, 所以第 5, 7, · · · 项都比第 1 项小很多. 设

$$\frac{1}{\tau} \int_{-\infty}^{+\infty} \frac{\Delta^2}{2} \varphi(\Delta) \mathrm{d}\Delta = D^5.$$

只考虑(A)式右端中的第1项和第3项,有

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2}$$

假定在 t=0 时刻, n 个粒子都集中在 x=0 点, 即

$$f(x,0) = \begin{cases} 0, & x \neq 0\\ \int_{-\infty}^{+\infty} f(x,0) dx = n, & x = 0. \end{cases}$$

求偏微分方程,因此有

$$f(x,t) = \frac{n}{\sqrt{4\pi Dt}} \exp\left\{-\frac{x^2}{4Dt}\right\}.$$

f(x,t)dx 表示 t 时刻 [x,x+dx] 内粒子数, $\frac{f(x,t)$ d $x}{n}=\frac{1}{\sqrt{4\pi Dt}}\exp\left\{-\frac{x^2}{4Dt}\right\}$ dx 表示 t 时刻在 [x,x+dx] 内粒子数占总粒子数的比例; 概率论的解释就是在 t 时刻在 [x,x+dx] 内发现某个粒子的概率, 这就对应了布朗运动.

进一步计算,一个粒子在 X 轴方向上平均经历的路程 λ_X ,或者比较准确的说就是在 X 轴方向上这些位移平方的算术平均的平方根 $\lambda_X=\sqrt{\overline{X^2}}=\sqrt{2Dt}$. (平均自由程)

从概率论的角度上面这段话可以理解为, 在利用时间 [0,t] 内一个粒子的平均位移替代多个粒子位移的算术平均, 利用一个粒子位移平方的平均替代多个粒子位移平方的算术平均, 即

$$\overline{X} = \int x f(x, t) dx = 0, \quad (均值)$$

$$\sqrt{\overline{X^2}} = \left(\int x^2 f(x, t) dx \right)^{\frac{1}{2}} = \sqrt{2Dt}. \quad (标准差)$$

$$D = \frac{1}{2t} \left\langle (X_t - X_0)^2 \right\rangle$$

在这里, $t=\tau$, 粒子在 τ 时处于位置 Δ , 因此在 τ 时该粒子的位移平方的期望就是 $\int \Delta^2 \varphi(\Delta) d\Delta$

⁵D 就是扩散系数. 对于一般的随机过程, 扩散系数被定义为:

利用 $\sqrt{X^2}$ 来表示平均路程的大小. 因此平均路程与时间的平方根成正比, $\sqrt{2Dt} \sim \sqrt{t}$. 这点很重要 6 由此我们可以抽象出布朗运动的定义:

- 1. 粒子轨道连续;
- 2. $\forall 0 < t_1 < t_2 < \dots < t_n$, 在 $[0, t_1]$, $(t_1, t_2]$, \dots , $(t_{n-1}, t_n]$ 之间过程独立;
- 3. 在 $[0,\tau]$ 和 $[t,t+\tau]$ 间运动规律相同;
- 4. 粒子运动规律关于 0 点对称.

4.3.1 布朗运动的分布

定义 4.4 (标准布朗运动)

设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是概率空间. 对每个 $\omega \in \Omega$, 假设存在依赖于 ω 的, 满足 W(0)=0 的连续函数 $W(t)(t\geq 0)$. 则 $W(t)(t\geq 0)$ 是一个布朗运动, 如果对所有 $0=t_0 < t_1 < \cdots < t_m$, 增量

$$W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1})$$

两两独立,每个增量服从正态分布,并且

$$\mathbb{E}[W(t_{i+1}) - W(t_i)] = 0$$

$$Var[W(t_{i+1}) - W(t_i)] = t_{i+1} - t_i$$



- 1. 布朗运动首先是一个高斯过程;
- 2. 标准布朗运动的均值函数为 m(t) = 0, 自协方差函数为 $K(s,t) = \min(s,t)$;

证明 由于增量

$$W(t_1) = W(t_1) - W(t_0), W(t_2) - W(t_1), \dots, W(t_m) - W(t_{m-1})$$

是独立的正态随机变量, 因此, $W(t_1)$, $W(t_2)$, \cdots , $W(t_m)$ 是联合正态随机变量. 联合正态随机变量的联合分布可由它们的均值和协方差确定. 每一个随机变量 $W(t_i)$ 的均值为 0. 对任意两个时刻 $0 \le s < t$, W(s) 和 W(t) 的协方差为:

$$\mathbb{E}[W(s)W(t)] = \mathbb{E}\left[W(s)(W(t) - W(s)) + W^{2}(s)\right]$$
$$= \mathbb{E}[W(s)] \cdot \mathbb{E}[W(t) - W(s)] + \mathbb{E}\left[W^{2}(s)\right]$$
$$= 0 + \text{Var}[W(s)] = s$$

基于此性质,可以构造布朗运动的协方差矩阵

$$\begin{bmatrix} & \mathbb{E}\left[W^{2}\left(t_{1}\right)\right] & \mathbb{E}\left[W\left(t_{1}\right)W\left(t_{2}\right)\right] & \cdots & \mathbb{E}\left[W\left(t_{1}\right)W\left(t_{m}\right)\right] \\ & \mathbb{E}\left[W\left(t_{2}\right)W\left(t_{1}\right)\right] & \mathbb{E}\left[W^{2}\left(t_{2}\right)\right] & \cdots & \mathbb{E}\left[W\left(t_{2}\right)W\left(t_{m}\right)\right] \\ & \vdots & \vdots & & \vdots \\ & \mathbb{E}\left[W\left(t_{m}\right)W\left(t_{1}\right) & \mathbb{E}\left[W\left(t_{m}\right)W\left(t_{2}\right)\right] & \cdots & \mathbb{E}\left[W^{2}\left(t_{m}\right)\right] \end{bmatrix} \\ = \begin{bmatrix} t_{1} & t_{1} & \cdots & t_{1} \\ t_{1} & t_{2} & \cdots & t_{2} \\ \vdots & \vdots & & \vdots \\ t_{1} & t_{2} & \cdots & t_{m} \end{bmatrix}$$

$$X(t+\tau) - X(t) \sim \sqrt{\tau}, \quad \frac{X(t+\tau) - X(t)}{\tau} \approx (\sqrt{\tau})' = \frac{1}{2\sqrt{\tau}}.$$

因此,当时间间隔很小时,花粉粒子运动方向变化极快速度无穷大,这就是观测到所谓的"运动的不规则性".从数学上看就是花粉粒子的运动轨迹每一点都不可微.

⁶花粉粒子的位移大致满足

布朗运动的矩母函数为:

$$\varphi(u_1, u_2, \dots, u_m)$$

$$= \mathbb{E} \exp \{u_m W(t_m) + u_{m-1} W(t_{m-1}) + \dots + u_1 W(t_1)\}$$

$$= \exp \left\{ \frac{1}{2} (u_1 + u_2 + \dots + u_m)^2 t_1 + \frac{1}{2} (u_2 + u_3 + \dots + u_m)^2 (t_2 - t_1) + \dots + \frac{1}{2} (u_{m-1} + u_m)^2 (t_{m-1} - t_{m-2}) + \frac{1}{2} u_m^2 (t_m - t_{m-1}) \right\}$$

例题 4.1 求概率

$$\mathbb{P}\left\{\int_0^1 W(t)dt > \frac{2}{\sqrt{3}}\right\}$$

首先需要指出的是, 布朗运动具有连续路径, 所以对每个路径来说, 黎曼 (Riemann) 积分 $\int_0^1 W(t)dt$ 存在. 我们只需找出 $\int_0^1 W(t)dt$ 的分布. 由黎曼积分的定义, 我们可以从逼近和

$$\sum_{i=1}^{n} W(t_i) \, \Delta t_i$$

的极限分布得到. 这里 $0=t_0 < t_1 < \cdots < t_n = 1$ 是 [0,1] 的分点, $\Delta t_i = t_i - t_{i-1}$, 例如取 $t_i = \frac{i}{n}$. 当 n=4 时, 逼近和即为例中的随机变量. 由于正态分布的线性组合仍为正态分布, 所以逼近和的分布都是零均值的正态分布. 方差为

$$\operatorname{Var}\left(\sum_{i=1}^{n} W\left(t_{i}\right) \Delta t_{i}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \operatorname{Cov}\left(W\left(t_{i}\right), W\left(t_{j}\right)\right) \Delta t_{i} \Delta t_{j}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \min\left(t_{i}, t_{j}\right) \Delta t_{i} \Delta t_{j}$$

$$\to \int_{0}^{1} \int_{0}^{1} \min(s, t) ds dt$$

$$= \frac{1}{3}$$

逼近和的极限 $\int_0^1 W(t)dt$ 仍服从正态分布 $N\left(0,\frac{1}{3}\right)$. 于是所求概率为

$$\mathbb{P}\left\{ \int_{0}^{1} W(t)dt > \frac{2}{\sqrt{3}} \right\} = P\left\{ \sqrt{3} \int_{0}^{1} W(t)dt > 2 \right\} = 1 - \Phi(2).$$

也可以形式地计算:

$$\mathbb{E} \int_0^1 W(t)dt = \int_0^1 \mathbb{E}[W(t)]dt = 0$$

$$\operatorname{Var} \left(\int_0^1 W(t)dt \right) = \mathbb{E} \left(\int_0^1 W(t)dt \cdot \int_0^1 W(s)ds \right)$$

$$= \int_0^1 \int_0^1 \mathbb{E}[W(t)W(s)]dsdt$$

$$= \int_0^1 \int_0^1 \min(t,s)dsdt$$

$$= \frac{1}{3}.$$

引理 4.3 (交换积分次序)

设 $\{W(t), t \geq 0\}$ 为标准布朗运动, g(x) 为 $\mathbb R$ 上的连续函数, 对 $0 \leq s < t$, 设 $\int_s^t \mathbb E[|g(W(u))|]du < \infty$, 则 $\int_s^t g(W(u))du$ 存在且

$$\mathbb{E} \int_{0}^{t} g(W(u))du = \int_{0}^{t} \mathbb{E}[g(W(u))]du.$$

0

证明 因为 W(t) 轨道连续, $g(\cdot)$ 连续故 g(W(t)) 轨道连续, $\int_s^t g(W(u))du$ 有定义, 且 $g(W(t,\omega))$ 是 $([0,\infty) \times \Omega, \mathcal{B}([0,\infty)) \times \mathcal{F})$ 到 $\mathcal{B}(\mathbb{R})$ 的可测变换.

因为

$$\int_{s}^{t} \mathbb{E}[|g(W(u))|] du = \int_{s}^{t} \int_{\Omega} |g(W(u,\omega))| d\mathbb{P}(\omega) du < \infty,$$

由 Fubini 定理即知

$$\begin{split} &\mathbb{E} \int_{s}^{t} g(W(u)) du = \int_{\Omega} \int_{s}^{t} g(W(u, \omega)) du d\mathbb{P}(\omega) \\ &= \int_{s}^{t} \int_{\Omega} g(W(u, \omega)) d\mathbb{P}(\omega) du \\ &= \int_{s}^{t} \mathbb{E}[g(W(u))] du. \end{split}$$

4.3.2 布朗运动的 Karhunen-Loeve 展开

先从离散随机过程考虑,假设随机过程 $X = \{X_n\}$ 的协方差矩阵为 B; 设 Y 是一个标准随机过程,即均值为 0,协方差矩阵为单位矩阵;设 X = AY,则根据随机向量方差的性质,有

$$\operatorname{Var}(\boldsymbol{A}\boldsymbol{Y}) = \boldsymbol{A}\operatorname{Var}(\boldsymbol{Y})\boldsymbol{A}^{\top} = \boldsymbol{A}\boldsymbol{A}^{\top} = \boldsymbol{B}$$

协方差矩阵 \boldsymbol{B} 有正交分解 $\boldsymbol{O}\boldsymbol{D}\boldsymbol{O}^{\top}$, 其中 $\boldsymbol{D}=\mathrm{diag}(\lambda_1,\cdots,\lambda_n), \boldsymbol{O}=(\boldsymbol{o}_1,\boldsymbol{o}_2,\cdots,\boldsymbol{o}_n), \boldsymbol{O}\boldsymbol{O}^{\top}=\boldsymbol{I}$, 每一个 \boldsymbol{o}_i 都满足 $|\boldsymbol{o}_i|=1$;

则有:

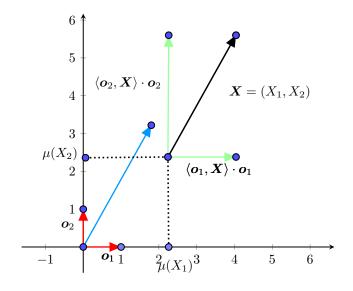
$$oldsymbol{A}oldsymbol{A}^{ op} = oldsymbol{O}oldsymbol{D}^{ op} = oldsymbol{O}oldsymbol{D}^{rac{1}{2}}oldsymbol{D}^{rac{1}{2}}oldsymbol{O}^{ op}$$

得到 $A^{\top} = D^{\frac{1}{2}} O^{\top}$, 于是 $Y = A^{-1} X = D^{-\frac{1}{2}} O^{\top} X$. 于是通过对 X 进行如下的线性变换就可得到标准随机变量.

$$Y_i = \frac{1}{\sqrt{\lambda_i}} \langle \boldsymbol{o}_i, \boldsymbol{X} \rangle$$

于是利用投影,可以把 X_n 展开为

$$\boldsymbol{X} = (\{X_j\}_{j=1}^n) = \mathbb{E}(X_n) + \sum_{i=1}^n \langle \boldsymbol{o}_i, X \rangle \boldsymbol{o}_i = \mathbb{E}(X_n) + \sum_{i=1}^n \sqrt{\lambda_i} Y_i \cdot \boldsymbol{o}_i$$



把离散的问题连续化,得到如下引理:

引理 4.4

随机过程 $X(t;\omega)$, $t\in\mathcal{T}=[0,T]$, $\omega\in\Omega$ 有均值函数 $m(t)=\mathbb{E}(X_t(\omega))$, 自协方差函数 $K(s,t)=\mathbb{E}[(X_t(\omega)-m(t))(X_s(\omega)-m(s))]$. K(s,t) 诱导出自共轭紧算子 $H:L^2(\mathcal{T})\to L^2(\mathcal{T})$:

$$H(\varphi(t)) = \int_{\mathcal{T}} K(s,t)\varphi(s)ds$$

利用 $H\varphi_i=\lambda_i\varphi_i(t)$ 计算出算子 H 的特征值 λ_i 和特征函数 φ_i . 进一步, 定义均值为 0, 方差为 1 的独立同分布的随机变量

$$Z_{i}(\omega) = \frac{1}{\sqrt{\lambda_{i}}} \int_{\mathcal{T}} (X(t) - m(t)) \varphi_{i}(t) dt$$

则如下的 K-L 展开成立:

$$X(t) = m(t) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} Z_i \varphi_i(t)$$

定理 4.3 (布朗运动 K-L 展开)

 $W(t) = \sum_{i=1}^{\infty} \frac{2\sqrt{2}}{(2i-1)\pi} \sin\left[\left(i-\frac{1}{2}\right)\pi t\right] \xi_i \,, \, \mbox{其中}\xi_i \,\, \mbox{为相互独立的标准正态分布随机变量}$

证明 设 $\mathcal{T} = [0,1]$, 布朗运动的均值函数为 0, 协方差函数 $K(s,t) = \min(s,t)$ 因而我们有

$$\int_0^1 \min(s, t) \varphi(t) dt = \lambda_i \varphi(s) \quad (*)$$

通过二次求导得到微分方程,求解得到

$$\varphi_i(t) = A \sin\left(\frac{t}{\sqrt{\lambda_i}}\right) + B \cos\left(\frac{t}{\lambda_i}\right) \quad (**)$$

在 (*) 中令 s=0, 得 $\varphi_i(0)=0$, 因而 B=0 将 (**) 代入 (*) 中, 计算得到

$$A = \sqrt{2}, \lambda_i = \frac{4}{(2i-1)^2 \pi^2}$$

故

$$W(t) = \sum_{1}^{\infty} \frac{2\sqrt{2}}{(2i-1)\pi} \sin\left[\left(i - \frac{1}{2}\right)\pi t\right] \xi_i$$

其中 ξ_i 为相互独立的标准正态分布随机变量.

下图是利用 K-L 展开模拟的一条布朗运动.

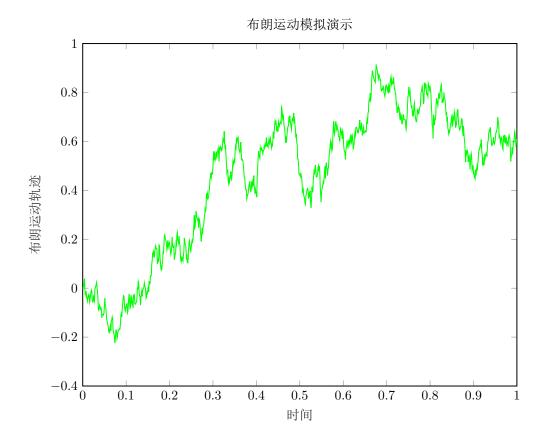
4.3.3 布朗运动的鞅性质

定义 4.5 (布朗运动的域流)

设 $W(t)(t \ge 0)$ 是定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的标准布朗运动. 布朗运动的域流是满足下列条件的一族 σ - 代数 $\mathcal{F}(t)(t \ge 0)$:

- 1. (信息累积) 对于 $0 \le s < t, \mathcal{F}(s) \subset \mathcal{F}(t)$.
- 2. (适应性) 对每个 $t \ge 0$, 在时刻 t, 布朗运动 W(t) 是 $\mathcal{F}(t)$ -可测的. 即在时刻 t 所获得的信息足以确定 布朗运动 W(t) 在该时刻的值.
- 3. (未来增量的独立性) 对于 $0 \le t < u$, 增量 W(u) W(t) 独立于 $\mathcal{F}(t)$. 即时刻 t 以后布朗运动的任何 增量都与时刻 t 所获得的信息无关.

设 $\Delta(t)(t\geq 0)$ 是随机过程. 如果对每个 $t\geq 0$, 随机变量 $\Delta(t)$ 是 $\mathcal{F}(t)$ — 可测的, 则称 $\Delta(t)$ 适应于域流 $\mathcal{F}(t)$.



定理 4.4 (布朗运动是鞅)

对于标准布朗运动 $(W_t)_{t>0}$, 设 $\mathcal{F}_t = \sigma(W_s, s \leq t)$, 则

1.
$$(W_t)_{t>0}$$
 是 $(\mathcal{F}_t)_{t>0}$ 鞅;

2.
$$(W_t^2 - t)_{t>0} \not\in (\mathcal{F}_t)_{t\geq 0} \not\ni$$
;

3.
$$\left(z_t = \exp\left\{CW_t - \frac{C^2t}{2}\right\}\right)_{t\geq 0} \not\in (\mathcal{F}_t)_{t\geq 0} \not\in \mathcal{F}_t$$

证明 (1)

$$\mathbb{E}[W_t \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s + W_s \mid \mathcal{F}_s]$$
$$= \mathbb{E}(W_t - W_s) + W_s = W_s$$

(2)
$$\mathbb{E}\left[W_{t}^{2} - t \mid \mathcal{F}_{s}\right] = \mathbb{E}\left[\left(W_{t} - W_{s}\right)^{2} + W_{s}^{2} + 2\left(W_{t} - W_{s}\right)W_{s} \mid \mathcal{F}_{s}\right] - t$$

$$= \mathbb{E}\left[\left(W_{t} - W_{s}\right)^{2} \mid \mathcal{F}_{s}\right] + \mathbb{E}\left[W_{s}^{2} \mid \mathcal{F}_{s}\right] + 2\mathbb{E}\left[\left(W_{t} - W_{s}\right)W_{s} \mid \mathcal{F}_{s}\right] - t$$

$$= \mathbb{E}\left(W_{t} - W_{s}\right)^{2} + W_{s}^{2} + 2W_{s}\mathbb{E}\left(W_{t} - W_{s}\right) - t = W_{s}^{2} - s$$

(3)
$$\mathbb{E}\left[z_{t} \mid \mathcal{F}_{s}\right] = \mathbb{E}\left[z_{s} \frac{z_{t}}{z_{s}} \mid \mathcal{F}_{s}\right]$$

$$= z_{s} \mathbb{E}\left[\frac{z_{t}}{z_{s}} \mid \mathcal{F}_{s}\right]$$

$$= z_{s} \mathbb{E}\left[e^{c(W_{t} - W_{s}) - \frac{c^{2}(t - s)}{2}} \mid \mathcal{F}_{s}\right]$$

$$= z_{s} \mathbb{E}\left(e^{c(W_{t} - W_{s}) - \frac{c^{2}(t - s)}{2}}\right) = z_{s}$$

4.4 二次变差与几何布朗运动

对于布朗运动, 没有自然的步幅. 如果给定 T > 0, 我们可以简单地选取 $\frac{T}{n}$ (n 充分大) 作为步幅, 计算截至时刻 T 的二次变差. 即计算:

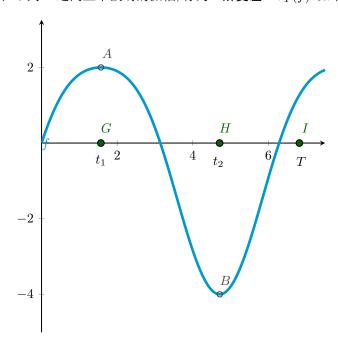
$$f(T) = \sum_{j=0}^{n-1} \left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right) \right]^2$$

$$\tag{4.1}$$

我们的兴趣在于步幅很小的情形, 再令 $n \to \infty$ 取极限. 最终得到 T.

4.4.1 一阶变差

考虑函数 f(t), 该函数在 0 到 T 之间上下波动的振幅, 称为**一阶变差** $FV_T(f)$. 如下图所示的一个函数, 有:



$$FV_T(f) = [f(t_1) - f(0)] - [f(t_2) - f(t_1)] + [f(T) - f(t_2)]$$

$$= \int_0^{t_1} f'(t)dt + \int_{t_1}^{t_2} (-f'(t)) dt + \int_{t_2}^T f'(t)dt$$

$$= \int_0^T |f'(t)| dt$$

一般地,为了计算函数截至时刻T的一阶变差,先取[0,T]的分划 $\Pi=\{t_0,t_1,\cdots,t_n\}$,其中的分点

$$0 = t_0 < t_1 < \dots < t_n = T$$

将被用来确定步幅 (并不要求这些分点是等分点). 该分划最大的步幅记为 $\|\Pi\|=\max_{j=0,\cdots,n-1}(t_{j+1}-t_j)$. 然后定义:

$$FV_T(f) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} |f(t_{j+1}) - f(t_j)|$$

在上式中, 当分点的数目 $n \to \infty$ 并且最大子区间的长度 $t_{j+1} - t_j \to 0$ 时取极限. 我们首先验证定义式与关于上图中的函数 f(t) 所得到的公式是一致的. 利用微分中值定理, 在每个子区间 $[t_j, t_{j+1}]$ 内存在一点 t_i^* 使得:

$$\frac{f\left(t_{j+1}\right) - f\left(t_{j}\right)}{t_{j+1} - t_{j}} = f'\left(t_{j}^{*}\right)$$

则有 $f(t_{j+1}) - f(t_j) = f'(t_j^*)(t_{j+1} - t_j)$, 于是

$$FV_T(f) = \lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} |f'(t_j^*)| (t_{j+1} - t_j) = \int_0^T |f'(t)| dt$$

4.4.2 二次变差

定义 4.6 (二次变差)

设 f(t) 是 $t \in [0,T]$ 上的函数, 截止时刻 T,f 的二次变差为:

$$[f, f](T) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} [f(t_{j+1}) - f(t_j)]^2$$

其中 $\Pi = \{t_0, t_1, \dots, t_n\}$ 并且 $0 = t_0 < t_1 < \dots < t_n = T$.

命题 4.5

具有连续导数的函数二次变差为 0.

证明

$$\sum_{j=0}^{n-1} \left[f\left(t_{j+1}\right) - f\left(t_{j}\right) \right]^{2} = \sum_{j=0}^{n-1} \left| f'\left(t_{j}^{*}\right) \right|^{2} \left(t_{j+1} - t_{j}\right)^{2}$$

$$\leq \left\| \Pi \right\| \cdot \sum_{j=0}^{n-1} \left| f'\left(t_{j}^{*}\right) \right|^{2} \left(t_{j+1} - t_{j}\right)$$

于是:

$$[f, f](T) \leq \lim_{\|\Pi\| \to 0} \left[\|\Pi\| \cdot \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j) \right]$$

$$= \lim_{\|\Pi\| \to 0} \|\Pi\| \cdot \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} |f'(t_j^*)|^2 (t_{j+1} - t_j)$$

$$= \lim_{\|\Pi\| \to 0} \|\Pi\| \cdot \int_0^T |f'(t)|^2 dt = 0$$

命题 4.6 (布朗运动的轨道分析性质)

布朗运动 W(t) 关于时间 t 处处连续, 但处处不可导. 对任意的 $t \ge 0$, 有

$$P\left(\lim_{h\to 0}\sup\left|\frac{W(t+h)-W(t)}{h}\right|=+\infty\right)=1$$

引理 4.5

设 $\{W(t), t \geq 0\}$ 为标准布朗运动, $\{t_i\}$ 为 [s, t] 的分割, $\delta_n = \max_i (t_{i+1} - t_i)$, 则

$$\mathbb{E} \left| \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2 - (t-s) \right|^2 \le 2\delta_n(t-s).$$

证明 证明记

$$\Delta t_i = t_{i+1} - t_i$$
, $\Delta W(t_i) = W(t_{i+1}) - W(t_i)$,

则

$$\mathbb{E} \left| \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2 - (t - s) \right|^2$$

$$= \mathbb{E} \left| \sum_{i=0}^{n-1} \left[(\Delta W(t_i))^2 - \Delta t_i \right] \right|^2$$

$$= \sum_{i=0}^{n-1} \mathbb{E} \left[(\Delta W(t_i))^2 - \Delta t_i \right]^2$$

$$+ \sum_{i \neq j} \mathbb{E} \left\{ \left[(\Delta W(t_i))^2 - \Delta t_i \right] \left[(\Delta W(t_j))^2 - \Delta t_j \right] \right\}$$

因为 $i \neq j$ 时 $\Delta W(t_i)$ 和 $\Delta W(t_j)$ 相互独立, 而 $\mathbb{E}\left[\left(\Delta W(t_i)\right)^2\right] = \Delta t_i$, 所以上式右侧的第二项为零. 利用 $\Delta W(t_i) \sim \mathcal{N}(0, \Delta t_i)$ 以及正态分布的四阶矩, 可得

$$\mathbb{E} \left| \sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2 - (t-s) \right|^2$$

$$= \sum_{i=0}^{n-1} \left[\mathbb{E} \left[(\Delta W(t_i))^4 \right] + (\Delta t_i)^2 - 2 (\Delta t_i) \mathbb{E} \left[(\Delta W(t_i))^2 \right] \right]$$

$$= \sum_{i=0}^{n-1} \left[3 (\Delta t_i)^2 + (\Delta t_i)^2 - 2 (\Delta t_i)^2 \right]$$

$$= 2 \sum_{i=0}^{n-1} (\Delta t_i)^2 \le 2 \delta_n \sum_{i=0}^{n-1} (\Delta t_i)$$

$$= 2 \delta_n (t-s).$$

下面的定理就是我们想要的结果.

定理 4.5 (布朗运动二次变差的性质)

设 $\{W(t), t \geq 0\}$ 为标准布朗运动, $\{t_i\}$ 是 [s,t] 的分割且 $\delta_n = \max_i (t_{i+1} - t_i) \to 0$, 则

$$\mathbb{E}\left|\sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))^2 - (t-s)\right|^2 \to 0, n \to \infty,$$
 从而 $[W,W]([s,t]) = t-s, [W,W](T) = T.$

证明 由引理可知极限为 0, 从而 $\sum_{i=0}^{n-1} (W(t_{i+1}) - W(t_i))$ 均方收敛到 t-s, 从而依概率收敛到 t-s.

笔记 因为布朗运动在 [0,t] 上的二次变差为 t, 即在依概率收敛的意义下

$$\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} \left[W(t_{i+1}) - W(t_i) \right]^2 = t$$

其中 $\delta_n = \max_{0 \le i \le n-1} (t_{i+1} - t_i)$, 上式从形式上可表示为

$$\int_{0}^{t} [dW(s)]^{2} = \int_{0}^{t} ds = t$$

或

$$[dW(t)]^2 = dt$$

下述定理说明了布朗运动的二次变差性质具有充要性.

定理 4.6 (Levy 布朗运动识别定理)

设 M(t) 是相应于域流 \mathcal{F}_t 的鞅. 若 M(0) = 0, M(t) 具有连续路径且对所有 $t \ge 0$ 都有

$$[M, M](t) = t$$

则 M(t) 是布朗运动.

 \Diamond

该定理的证明需要用到 Ito 公式.

4.4.3 布朗运动不是有界变差的

一阶变差有限的函数称为有界变差函数. 闭区间上连续可微函数必定为有界变差函数. 接下来我们将说明布朗运动的轨道不是有界变差的.

引理 4.6 (二次变差具有更强的收敛性)

设 $\{W(t), t \ge 0\}$ 为标准布朗运动, $0 \le s < t$, 取 $s = t_0 < t_1 < \dots < t_{2^n} = t$ 为 [s, t] 的 2^n 等分点, 则

$$\sum_{i=0}^{2^{n}-1} \left(\Delta W\left(t_{i}\right)\right)^{2} \rightarrow t-s, \text{ a.s., } (n \rightarrow \infty)$$

 \sim

证明 记

$$S_n = \sum_{i=0}^{2^n - 1} (\Delta W(t_i))^2$$

则

$$\mathbb{E}S_n = \sum_{i=0}^{2^n - 1} (t_{i+1} - t_i) = t - s,$$

由二次变差引理,

$$\mathbb{E}|S_n - (t - s)|^2 \le 2\delta_n(t - s) = 2(t - s)^2 2^{-n},$$

于是

$$\sum_{n=1}^{\infty} \mathbb{E} \left| S_n - (t - s) \right|^2 < \infty$$

由单调收敛定理有

$$\mathbb{E}\sum_{n=1}^{\infty} |S_n - (t - s)|^2 = \sum_{n=1}^{\infty} \mathbb{E} |S_n - (t - s)|^2 < \infty$$

所以

$$\sum_{n=1}^{\infty} |S_n - (t - s)|^2 < \infty, \text{ a.s. },$$
$$|S_n - (t - s)|^2 \to 0, \text{ a.s.}$$

引理得证.

定理 4.7 (布朗运动一阶变差的性质)

设 $\{W(t), t \ge 0\}$ 为标准布朗运动, $0 \le s < t$, 取 $s = t_0 < t_1 < \dots < t_{2^n} = t$ 为 [s, t] 的 2^n 等分点, 则

$$\sum_{i=0}^{2^{n}-1} |\Delta W(t_i)| \to \infty, \text{ a.s., } (n \to \infty).$$

 \Diamond

证明 记

$$M_n(\omega) = \max_{i} |W(t_{i+1}, \omega) - W(t_i, \omega)|,$$

当 $n\to\infty$ 时 $t_{i+1}-t_i=(t-s)2^{-n}$,由 $W(t,\omega)$ 对固定的 ω 关于 t 连续以及闭区间连续函数一致连续,可知 $M_n(\omega)\to 0, n\to\infty$. 而

$$\sum_{i=0}^{2^{n}-1} |\Delta W(t_i)|^2 \le M_n(\omega) \sum_{i=0}^{2^{n}-1} |\Delta W(t_i)|,$$

由引理知 $\sum_{i=0}^{2^{n}-1}\left|\Delta W\left(t_{i}\right)\right|^{2}
ightarrow t-s>0$, a.s., 于是

$$\sum_{i=0}^{2^{n}-1} \left| \Delta W\left(t_{i}\right) \right| \geq \frac{\sum_{i=0}^{2^{n}-1} \left| \Delta W\left(t_{i}\right) \right|^{2}}{M_{n}(\omega)} \to \infty, n \to \infty.$$

推论 4.1

布朗运动的轨道都不是有界变差的.

 \Diamond

下面总结布朗运动的路径性质. 从时刻 0 到时刻 T 对布朗运动的一次观察称为布朗运动在区间 [0,T] 上的一个路径 (轨道) 或一个实现. 布朗运动的几乎所有样本路径 $W(t), 0 \le t \le T$ 都具有下述性质:

- 1. 是t的连续函数;
- 2. 在任意区间 (无论区间多么小) 上都不是单调的;
- 3. 在任意点都不是可微的;
- 4. 在任意区间(无论区间多么小)上都不是有界变差的;
- 5. 对任意 $0 \le s < t$, 在 [s,t] 上的二次变差等于 t s.

🕏 笔记 关于性质 1

$$\lim_{\Delta t \to 0} \mathbb{E}|W(t + \Delta t) - W(t)|^2 = \lim_{\Delta t \to 0} \Delta t = 0.$$

关于性质 2: 将 $[t,t+\Delta t]$ 等分为 n 分,则每一份上的增量独立同 $\mathcal{N}\left(0,\frac{\Delta t}{n}\right)$ 分布,这 n 个增量同为正号或者同为负号的概率为

$$2 \times \left(\frac{1}{2}\right)^n = 2^{-(n-1)} \to 0,$$

所以不可能单调.

关于性质 3:

$$\begin{split} \mathbb{E} \left| \frac{W(t + \Delta t) - W(t)}{\Delta t} \right|^2 &= \frac{\Delta t}{\Delta t^2} \to \infty, \\ \mathbb{E} \left| \frac{W(t + \Delta t) - W(t)}{\Delta t} \right| &= \frac{\sqrt{\Delta t} \mathbb{E}|Z|}{\Delta t} \to \infty, \end{split}$$

(其中 Z 表示标准正态分布随机变量) 这提示 $\lim_{\Delta t \to 0} \frac{W(t+\Delta t)-W(t)}{\Delta t}$ 不存在.

4.4.4 几何布朗运动

在金融市场中,人们经常假定股票的价格按照几何布朗运动变化.

定义 4.7 (几何布朗运动)

设 α 和 $\sigma > 0$ 是常数,定义几何布朗运动为

$$S(t) = S(0) \exp\left\{\sigma W(t) + \left(\alpha - \frac{1}{2}\sigma^2\right)t\right\}$$

称

$$S(t) = \exp\{W(t)\}\$$

为标准几何布朗运动.

由定理4.4知几何布朗运动是鞅. 另外还有:

命题 4.7

标准几何布朗运动的期望和方差为

$$\begin{split} \mathbb{E}[S(t)] &= \mathbb{E}\left[e^{W(t)}\right] = e^{t/2}; \\ \operatorname{Var}[S(t)] &= \mathbb{E}\left[S^2(t)\right] - (\mathbb{E}[S(t)])^2 \\ &= \mathbb{E}\left[e^{2W(t)}\right] - e^t \\ &= e^{2t} - e^t \end{split}$$

几何布朗运动的边缘分布为对数正态分布.

引理 4.7 (交互变差)

$$\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} (W(t_{j+1}) - W(t_j)) (t_{j+1} - t_j) = 0$$

证明 注意到

$$|(W(t_{j+1}) - W(t_j))(t_{j+1} - t_j)| \le \max_{0 \le k \le n-1} |W(t_{k+1}) - W(t_k)|(t_{j+1} - t_j)$$

于是:

$$\left| \sum_{j=0}^{n-1} \left(W\left(t_{j+1} \right) - W\left(t_{j} \right) \right) \left(t_{j+1} - t_{j} \right) \right| \leq \max_{0 \leq k \leq n-1} \left| W\left(t_{k+1} \right) - W\left(t_{k} \right) \right| \cdot T$$

由于 W 是连续的, 当最大子区间的长度 $\|\Pi\| \to 0$ 时, $\max_{0 \le k \le n-1} \|W(t_{k+1}) - W(t_k)\|$ 的极限为 0.

我们将利用布朗运动的二次变差和交互变差, 通过几何布朗运动确定其波动率 σ . 给定 $0 \le T_1 < T_2$, 取区间 $[T_1, T_2]$ 的分划 $T_1 = t_0 < t_1 < \cdots < t_m = T_2$, 观察对数回报⁷:

$$\log \frac{S(t_{j+1})}{S(t_{i})} = \sigma \left(W(t_{j+1}) - W(t_{j}) \right) + \left(\alpha - \frac{1}{2} \sigma^{2} \right) (t_{j+1} - t_{j})$$

对数回报的平方和(实现波动率)为

$$\begin{split} &\sum_{j=0}^{m-1} \left(\log \frac{S\left(t_{j+1}\right)}{S\left(t_{j}\right)}\right)^{2} \\ &= &\sigma^{2} \sum_{j=0}^{m-1} \left(W\left(t_{j+1}\right) - W\left(t_{j}\right)\right)^{2} + \left(\alpha - \frac{1}{2}\sigma^{2}\right)^{2} \sum_{j=0}^{m-1} \left(t_{j+1} - t_{j}\right)^{2} \\ &+ 2\sigma \left(\alpha - \frac{1}{2}\sigma^{2}\right) \sum_{j=0}^{m-1} \left(W\left(t_{j+1}\right) - W\left(t_{j}\right)\right) \left(t_{j+1} - t_{j}\right) \end{split}$$

 $^{^7}$ 回报率为 $r = \frac{S(t_{j+1}) - S(t_j)}{S(t_j)} = \frac{S(t_{j+1})}{S(t_j)} - 1$,所以 $r \to 0$ 时, $\log(1+r) \sim r$,近似认为对数回报衡量了回报率

利用布朗运动二次变差和交互变差引理, 我们得到

$$\lim_{||\Pi|| \to 0} \sum_{j=0}^{m-1} \left(\log \frac{S(t_{j+1})}{S(t_j)} \right)^2 = \sigma^2(T_2 - T_1) + 0 + 0$$

所以得到:

定理 4.8

股票价格的平均实现波动率由几何布朗运动的参数 σ^2 近似.

$$\frac{1}{T_2 - T_1} \sum_{j=0}^{m-1} \left(\log \frac{S\left(t_{j+1}\right)}{S\left(t_{j}\right)} \right)^2 \approx \sigma^2$$

4.5 布朗运动的马尔可夫性

首先回顾马氏过程: 设 $\{X(t),\mathcal{F}(t),t\geq 0\}$ 是概率空间 (Ω,\mathcal{F},P) 上的适应过程, 如果对任意有界 Borel 函数 g(x) , 有

$$\mathbb{E}(g(X(t)) \mid \mathcal{F}(s)) = \mathbb{E}(g(X(t)) \mid X(s)), \forall 0 \le s < t,$$

则称 $\{X(t), \mathcal{F}(t), t \geq 0\}$ 为马氏过程.

上述定义是利用将来的状态的函数的条件期望;也可以用将来的条件分布进行定义:

定义 4.8 (马氏过程)

设 $\{X(t), t \ge 0\}$ 是一个连续时间的随机过程, 如果 $\forall t, s > 0$, 有

$$\mathbb{P}\left\{X(t+s) \leq y \mid \mathcal{F}_t\right\} = \mathbb{P}\left\{X(t+s) \leq y \mid X(t)\right\}, \text{ a.s.}$$

则称 $\{X(t)\}$ 为马氏过程, 这里 $\mathcal{F}_t = \sigma\{X(u), 0 \le u \le t\}$. 满足上式的性质称为马氏性.

定理 4.9

布朗运动 $\{W(t)\}$ 具有马尔可夫性.

证明 注意到 W(t+s) - W(t) 与 \mathcal{F}_t 独立, 用矩母函数方法可以证明 W(t+s) 在给定条件 \mathcal{F}_t 下的分布与在给定条件 W(t) 下的分布是一致的. 事实上, 由定理4.4第 (3) 条, $e^{-\frac{u^2}{2}t}e^{uW(t)}$ 是鞅, 可得

$$\mathbb{E}\left[e^{uW(t+s)} \mid \mathcal{F}_t\right]$$

$$=e^{\frac{u^2}{2}(t+s)}\mathbb{E}\left[e^{-\frac{u^2}{2}(t+s)}e^{uW(t+s)}\right| \mathcal{F}_t\right]$$

$$=e^{\frac{u^2}{2}(t+s)}e^{-\frac{u^2}{2}t}e^{uW(t)}$$

$$=e^{\frac{u^2}{2}s}e^{uW(t)}.$$

而

$$\mathbb{E}\left[e^{uW(t+s)} \mid W(t)\right]$$

$$=\mathbb{E}\left[e^{uW(t)}e^{u(W(t+s)-W(t))} \mid W(t)\right]$$

$$=e^{uW(t)}\mathbb{E}\left[e^{u(W(t+s)-W(t))}\right]$$

$$=e^{uW(t)}e^{\frac{u^2}{2}s}$$

$$=\mathbb{E}\left[e^{uW(t+s)} \mid \mathcal{F}_t\right],$$

所以 W(t+s) 在给定条件 \mathcal{F}_t 下的分布与在给定条件 W(t) 下的分布是一致的,即 $\{W(t)\}$ 具有马氏性.

等于在 X(t) 条件下的联合条件分布; 进一步地, X(t+s), $s \ge 0$ 在 F_t 下的分布, 等于在 X(t) 条件下的分布.

连续状态的马氏过程 $\{X(t)\}$ 的转移概率分布定义为在时刻 s 过程处于状态 x 的条件下, 过程在时刻 t 的分布函数

$$F(y, t, x, s) = \mathbb{P}\{X(t) \le y \mid X(s) = x\},\$$

在布朗运动的情况下,这一分布函数是正态的

$$F(y, t, x, s) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi(t-s)}} \exp\left\{-\frac{(u-x)^2}{2(t-s)}\right\} du.$$

布朗运动的转移概率函数满足方程 F(s,t;x,y) = F(0,t-s,x,y), 这说明布朗运动是时齐的马氏过程. 这等价于

$$\mathbb{P}\{W(t) \le y \mid W(s) = x\} = \mathbb{P}\{W(t - s) \le y \mid W(0) = x\}$$

当 s=0 时, F(0,t;x,y) 具有密度函数

$$p_t(x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(y-x)^2}{2t}},$$

即布朗运动的转移概率密度.

4.6 布朗运动的最大值变量及反正弦律

4.6.1 首达时

考虑从 0 出发的标准布朗运动 $\{W(t), t \geq 0\}$, 记 T_x 为到达 x 的首达时,

$$T_x = \inf\{t > 0 : W(t) = x\}.$$

因为布朗运动的轨道连续, 所以对 $x \neq 0, 0 < T_x \leq +\infty$. 当 x > 0 时, 为计算 $\mathbb{P}\{T_x \leq t\}$, 我们考虑 $\mathbb{P}\{W(t) \geq x\}$. 由全概率公式

$$\mathbb{P}\{W(t) \ge x\}$$

$$= \mathbb{P}\{W(t) \ge x \mid T_x \le t\} \mathbb{P}\{T_x \le t\}$$

$$+ \mathbb{P}\{W(t) \ge x \mid T_x > t\} \mathbb{P}\{T_x > t\},$$

右侧的第二项 $T_x > t$ 表示截止到 t 时刻未达到 x > 0 处, 因为 W(0) = 0 以及轨道连续性, 这时必有 W(t) < x 成立, 即 $\mathbb{P}\{W(t) \ge x \mid T_x > t\} = 0$, 所以

$$\mathbb{P}\{W(t) \ge x\} = \mathbb{P}\{W(t) \ge x \mid T_x \le t\} \mathbb{P}\{T_x \le t\}.$$

若 $T_x \leq t$,则 W(t) 在 [0,t] 中的某个时刻击中 $x,W(T_x) = x$,由布朗运动的马氏性,可以认为布朗运动从 $W(T_x) = x$ 重新开始,上行与下行概率各 $\frac{1}{9}$,即于是 $\mathbb{P}(W(t) \geq x) = \mathbb{P}(W(t) \leq x) = \frac{1}{9}$,

$$\begin{split} \mathbb{P}\{W(t) \geq x\} &= \frac{1}{2} \mathbb{P}\left\{T_x \leq t\right\}, \\ \mathbb{P}\left(T_x \leq t\right) &= 2 \mathbb{P}(W(t) \geq x) \\ &= 2 \mathbb{P}(\sqrt{t}Z \geq x) \\ &= 2 \left[1 - \Phi\left(\frac{x}{\sqrt{t}}\right)\right] \\ &= 2 \Phi\left(-\frac{x}{\sqrt{t}}\right) \\ &= 2 \int_{\frac{x}{-\infty}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv, x > 0, t > 0. \end{split}$$

对 x < 0, 同理可得以上结果. 于是 T_x 的分布函数为

$$\mathbb{P}(T_x \le t) = 2\left[1 - \Phi\left(\frac{|x|}{\sqrt{t}}\right)\right]$$
$$= 2\int_{\frac{|x|}{\sqrt{t}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv, t > 0.$$

由此可见

$$\mathbb{P}\left\{T_x < \infty\right\} = \lim_{t \to \infty} \mathbb{P}\left\{T_x \le t\right\} = 1.$$

求导得 T_x 的密度函数为

$$f_{T_x}(t) = |x|t^{-\frac{3}{2}}\phi\left(-\frac{|x|}{\sqrt{t}}\right)$$
$$= \frac{|x|}{t\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}, t > 0.$$

定理 4.10

对于 $x \in \mathbb{R}$, 布朗运动关于水平x 的首达时几乎必然有限, 并且其分布的拉普拉斯 (Laplace) 变换为:

$$\mathbb{E}\exp\left\{-\alpha T_x\right\} = \exp\left\{-|x|\sqrt{2\alpha}\right\}, \quad \forall \alpha > 0$$

证明 定义几何布朗运动

$$Z(t) = \exp\left\{\sigma W(t) - \frac{1}{2}\sigma^2 t\right\}$$

由其鞅的性质3.1,有

$$1 = Z(0) = \mathbb{E}Z(\min\{t, T_x\}) = \mathbb{E}\left[\exp\left\{\sigma W(\min\{t, T_x\}) - \frac{1}{2}\sigma^2(\min\{t, T_x\})\right\}\right]$$

假设 $\sigma > 0, x > 0$, 此情形下, 若 $t \le T_x$, 则布朗运动总是不超过水平 x, 于是:

$$0 \le e^{\sigma W(t)} \le e^{\sigma x}$$

则如果 $T_x = \infty$, 就有这一项有界.

另一方面, 我们有:

$$\lim_{t \to \infty} \exp\left\{-\frac{1}{2}\sigma^2(\min\{t, T_x\})\right\} = \mathbb{I}_{T_x < \infty} \exp\left\{-\frac{1}{2}\sigma^2 T_x\right\}$$

综合地,有

$$\lim_{t \to \infty} \exp \left\{ \sigma W \left(t \wedge T_x \right) - \frac{1}{2} \sigma^2 \left(t \wedge T_x \right) \right\} = \mathbb{I}_{\left\{ T_x < \infty \right\}} \exp \left\{ \sigma x - \frac{1}{2} \sigma^2 T_x \right\}$$

因此我们由控制收敛定理,

$$1 = \mathbb{E}\left[\mathbb{I}_{\{T_x < \infty\}} \exp\left\{\sigma x - \frac{1}{2}\sigma^2 T_x\right\}\right]$$

由于 T_x 以概率1有限,所以

$$\mathbb{E}\left[\exp\left\{-\frac{1}{2}\sigma^2T_x\right\}\right] = e^{-\sigma x}$$

设 $\alpha > 0$, 令 $\sigma = \sqrt{2\alpha}$, 这就得到定理;

如果 x 为负数, 由于布朗运动是对称的, 所以首达时 T_x 和 $T_{|x|}$ 具有同分布, 从而得到关于负数 x 的等式. 证 毕.

推论 4.2 (首达时以概率 1 有限, 但其期望无限)

$$\mathbb{E}(T_x) = \infty$$

证明 两边关于 α 求导数,得到:

$$\mathbb{E}\left[T_x \exp\left\{-\alpha T_x\right\}\right] = \frac{|x|}{\sqrt{2\alpha}} \exp\left\{-|x|\sqrt{2\alpha}\right\}, \quad \forall \alpha > 0$$

 $\phi \alpha \rightarrow 0^+$, 只要 $x \neq 0$, 就得到

$$\mathbb{E}(T_x) = \infty$$

笔记 T_x 虽然几乎必然是有限的, 但有无穷的期望. 直观地看, 就是布朗运动以概率 1 会击中 x , 但它的平均时间是无穷的. 性质 $\mathbb{P}\{T_x<\infty\}=1$ 称为布朗运动的常返性. 由于始于点 a 的布朗运动与 $\{a+W(t)\}$ 是相同的, 这里 $\{W(t)\}$ 是始于 0 的布朗运动,所以

$$\mathbb{P}_a \left\{ T_x < \infty \right\} = \mathbb{P}_0 \left\{ T_{x-a} < \infty \right\} = 1,$$

即布朗运动从任意点出发, 击中x的概率都是1.

4.6.2 反射原理、最大值和最小值的分布

定理 4.11 (反射原理)

设 $\{W(t), t \ge 0\}$ 是标准布朗运动, 固定时间 t > 0, 令

$$W(t)' = \begin{cases} W(t), & t \le s \\ 2W(s) - W(t), & t > s \end{cases}$$

则 W(t)' 依然是标准布朗运动, 并满足:

$$\mathbb{E}[W'(u)W'(v)] = \min\{u, v\}$$

证明 对 t > 0,令

$$Y(t) = W(t)\mathbb{I}_{t < T_b}, Z(t) = W(t + T_b) - b.$$

由 Markov 性, $Z = \{Z(t), t \ge 0\}$ 是独立于 $Y = \{Y(t), t \ge 0\}$ 的 Brown 运动. 所以 $-Z = \{-Z(t), t \ge 0\}$ 是独立于 Y 的 Brown 运动; (Y, Z) 与 (Y, -Z) 有相同有限维分布. 定义

$$\varphi: (Y, Z) \to \{Y(t)\mathbb{I}_{\{t < T_b\}} + (b + Z(t - T_b))\mathbb{I}_{\{t > T_b\}}, t \ge 0\}$$

生成一个连续过程, $\varphi(Y, -Z)$ 也是一个连续过程, 且二者具有相同有限维分布. 而

$$\varphi(Y,Z) = X, \varphi(Y,-Z) = \hat{X},$$

所以 \hat{X} 也是Brown运动.

 $\diamondsuit s = T_x$, 有:

$$\mathbb{P}(t > T_x, W(t) < w) = \mathbb{P}(W(t)' > 2x - w)$$

上式被称为反射等式.

另一个有趣的随机变量是布朗运动在 [0,t] 中达到的最大值

$$M(t) = \max_{0 \le u \le t} W(u).$$

由轨道连续性, M(t) 有定义且取有限值. 对 x>0,事件 $M(t)\geq x$ 等价于 $T_x\leq t$. 事实上, 当 $M(t)\geq x$,由轨道连续性以及 W(0)=0,[0,t] 中一定存在某个 t' 使得 W(t')=x,所以 $T_x\leq t$;另一方面,如果 $T_x\leq t$,则 [0,t] 中的某个 W(t')=x,所以 $M(t)\geq x$. 于是

$$\mathbb{P}(M(t) \ge x) = \mathbb{P}(T_x \le t) = 2\left[1 - \Phi\left(\frac{x}{\sqrt{t}}\right)\right].$$

令 $x \to 0$ 得 $\mathbb{P}(M(t) \ge 0) = 1$. 求导得 M(t) 的密度函数为

$$f_{M(t)}(x) = \frac{2}{\sqrt{t}}\phi\left(-\frac{x}{\sqrt{t}}\right) = \frac{2}{\sqrt{2\pi t}}e^{-\frac{x^2}{2t}}, x > 0.$$

易见 M(t) 与 |W(t)| 同分布.

笔记 这一随机过程用于定价障碍期权; 对于正数 x, M(t) > x 当且仅当 $T_x < t$. 于是, 可将反射等式改写为:

$$\mathbb{P}\{M(t) \ge x, W(t) \le w\} = \mathbb{P}\{W(t) \ge 2x - w\}, \quad w \le x, m > 0$$

由此, 我们可以得到 W(t) 和 M(t) 的联合分布:

$$f_{M(t),W(t)}(x,w) = \frac{2(2x-w)}{t\sqrt{2\pi t}}e^{-\frac{(2x-w)^2}{2t}}, \quad w \le x, m > 0$$

推论 4.3 (最大值条件分布)

给定 W(t) = w, M(t) 具有密度函数:

$$f_{M|W}(x|w) = \frac{2(2x-w)}{t}e^{-\frac{2x(x-w)}{t}}, \quad w \le x, x > 0$$

证明 条件密度就是联合密度除以条件随机变量的边际密度.

$$f_{M|W}(x|w) = \frac{f_{M,W}(x,w)}{f_{W}(w)} = \frac{2(2x-w)}{t\sqrt{2\pi t}}\sqrt{2\pi t}\exp\left\{-\frac{(2x-w)^2}{2t} + \frac{w^2}{2t}\right\} = \frac{2(2x-w)}{t}\exp\left\{-\frac{2x(x-w)}{t}\right\}$$

命题 4.8

$$\mathbb{E}[M^2(t)] = t$$

证明

$$\mathbb{E}\left[M^{2}(t)\right] = \int_{0}^{\infty} \frac{2x^{2}}{\sqrt{2\pi t}} e^{-\frac{x^{2}}{2t}} dt$$
$$= \int_{-\infty}^{\infty} \frac{x^{2}}{\sqrt{2\pi t}} e^{-\frac{x^{2}}{2t}} dt$$
$$= t.$$

最后的等式是对 N(0,t) 分布求方差.

考虑最小值的分布,类似可得

$$m(t) = \min_{0 \le u \le t} W(u)$$

的分布函数为

$$\mathbb{P}(m(t) \le x) = 2\Phi\left(-\frac{|x|}{\sqrt{t}}\right), x < 0.$$

可得 $P(m(t) \le 0) = 1$, 以及 |m(t)| = -m(t) 与 M(t) 同分布. 令

$$H(t) = \max_{0 \le u \le t} |W(u)|.$$

关于绝对值的最大值暂未得到简单的分布,但是可以有控制

$$H(t) \leq M(t) + |m(t)|,$$

 $\mathbb{E}\left[H^2(t)\right] \leq 4t.$

4.6.3 有零点的概率和反正弦定律

如果有时刻 τ 使得 $W(\tau)=0$,就称 τ 为布朗运动的零点.

定理 4.12 (至少有一个零点的概率)

设 $\{W^x(t)\}$ 为始于 x 的布朗运动,则其在 (0,t) 上至少有一个零点的概率为

$$\frac{|x|}{\sqrt{2\pi}} \int_0^t u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} du$$

证明 如果 x < 0,则由 $\{W^x(t)\}$ 的连续性得

$$\mathbb{P}\left\{W^x \ \text{在}(0,t) \ \text{中至少有一个零点}\right\} = \mathbb{P}\left\{\max_{0 \leq s \leq t} W^x(s) \geq 0\right\},$$
因为 $W^x(t) = W(t) + x$, 我们有
$$\mathbb{P}\left\{W^x \ \text{在}(0,t) \ \text{中至少有一个零点}\right\}$$

$$= \mathbb{P}\left\{\max_{0 \leq s \leq t} W^x(s) \geq 0\right\}$$

$$= \mathbb{P}\left\{\max_{0 \leq s \leq t} W(s) + x \geq 0\right\} = \mathbb{P}\left\{\max_{0 \leq s \leq t} W(s) \geq -x\right\}$$

$$= \mathbb{P}\left\{T_{-x} \leq t\right\} = \mathbb{P}\left\{T_x \leq t\right\}$$

$$= \int_0^t f_{T_x}(u) du = \frac{|x|}{\sqrt{2\pi}} \int_0^t u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} du.$$

对于x > 0的情况的证明类似,结果相同.

定理 4.13

W(t) 在 $t \in (a,b)$ 中至少有一个零点的概率为

$$\frac{2}{\pi} \arccos \sqrt{\frac{a}{b}}$$

က

证明 记

$$h(x) = \mathbb{P}\{W \ \text{在}(a,b) \ \text{中至少有一个零点} \ | \ W(a) = x\},$$

由对称性可知 h(-x) = h(x). 由马氏性,

$$h(x) = \mathbb{P}\left\{W^x \ \text{在}(0, b - a) \ \text{中至少有一个零点}\right\}$$

$$= \frac{|x|}{\sqrt{2\pi}} \int_0^{b-a} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} du.$$

由全期望公式

$$\begin{split} & \mathbb{P}\{W \; \dot{a}(a,b) \; \dot{\mathbf{P}} \, \underline{\mathbf{F}} \, \dot{\mathbf{F}} \, \dot{\mathbf{F}} - \dot{\mathbf{F}} \, \underline{\mathbf{F}} \, \big| \, W(a) = x\} f_{W(a)}(x) dx \\ & = \int_{-\infty}^{\infty} h(x) \frac{1}{\sqrt{2\pi a}} e^{-\frac{x^2}{2a}} dx \\ & = \frac{2}{\sqrt{2\pi a}} \int_{0}^{\infty} h(x) e^{-\frac{x^2}{2a}} dx \\ & = \frac{2}{\sqrt{2\pi a}} \int_{0}^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2a}} \int_{0}^{b-a} u^{-\frac{3}{2}} e^{-\frac{x^2}{2u}} du dx \\ & = \frac{1}{\pi \sqrt{a}} \int_{0}^{b-a} u^{-\frac{3}{2}} \int_{0}^{\infty} x e^{-x^2 \left(\frac{1}{2u} + \frac{1}{2a}\right)} dx du \\ & = \frac{1}{\pi \sqrt{a}} \int_{0}^{b-a} u^{-\frac{3}{2}} \frac{au}{a+u} du = \frac{\sqrt{a}}{\pi} \int_{0}^{b-a} \frac{u^{-\frac{1}{2}}}{a+u} du \\ & = \frac{2\sqrt{a}}{\pi} \int_{0}^{b-a} \frac{d\sqrt{u}}{a+u} = \frac{2\sqrt{a}}{\pi} \int_{0}^{\sqrt{b-a}} \frac{dv}{a+v^2} \end{split}$$

$$\begin{split} &=\frac{2\sqrt{a}}{\pi}\int_0^{\frac{\sqrt{b-a}}{\sqrt{a}}}\frac{\sqrt{a}dw}{a\left(1+w^2\right)} = \frac{2}{\pi}\int_0^{\frac{\sqrt{b-a}}{\sqrt{a}}}\frac{dw}{1+w^2}\\ &=\frac{2}{\pi}\arctan\frac{\sqrt{b-a}}{\sqrt{a}} = \frac{2}{\pi}\arccos\sqrt{\frac{a}{b}}. \end{split}$$

这里利用了 $\int \frac{1}{1+w^2} dw = \arctan w$,以及若 $y = \arctan \sqrt{\frac{b-a}{a}}$,则 $1 + \tan y^2 = \frac{1}{\cos^2 y} = \frac{b}{a}$ 从而 $y = \arccos \sqrt{\frac{a}{b}}$

定理 4.14 (反正弦定律)

设 $\{W(t), t \ge 0\}$ 是布朗运动,则

$$\mathbb{P}\{W(t)\ \texttt{在}(a,b)\ \texttt{中没有零点}\} = \frac{2}{\pi}\arcsin\sqrt{\frac{a}{b}}.$$

下面介绍布朗运动在时刻 t 之前最后一个零点以及在 t 之后的第一个零点的分布情况. 令

$$\zeta_t = \sup\{s \le t, W(s) = 0\} = t$$
 之前最后一个零点,
 $\beta_t = \inf\{s \ge t, W(s) = 0\} = t$ 之后第一个零点,

由反正弦律有

$$\mathbb{P}\left\{\zeta_{t} \leq x\right\} = \mathbb{P}\left\{W \ \text{在}(x,t) \ \text{中没有零点}\right\} = \frac{2}{\pi}\arcsin\sqrt{\frac{x}{t}},$$

$$\mathbb{P}\left\{\beta_{t} \geq y\right\} = \mathbb{P}\left\{W \ \text{E}(t,y) \ \text{中没有零点}\right\} = \frac{2}{\pi}\arcsin\sqrt{\frac{t}{y}},$$

$$\mathbb{P}\left\{\zeta_{t} \leq x, \beta_{t} \geq y\right\} = \mathbb{P}\left\{W \ \text{E}(x,y) \ \text{中没有零点}\right\} = \frac{2}{\pi}\arcsin\sqrt{\frac{x}{y}}.$$

由 ζ_t 的分布得

$$\mathbb{P}\left(\zeta_t > 0\right) = 1 - \frac{2}{\pi} \arcsin\sqrt{\frac{0}{t}}, = 1, \forall t > 0.$$

这说明布朗运动 0 时刻从状态 0 出发, 在任意的 (0,t) 区间内都存在零点. 这说明布朗运动一旦达到零点, 会瞬时无穷次达到零点. 这从布朗运动的随机游动逼近可以理解.

4.7 布朗运动的变式

4.7.1 布朗桥

定义 4.9 (布朗桥)

设 $\{W(t), t \geq 0\}$ 是布朗运动. 令

$$W^*(t) = W(t) - tW(1), 0 \le t \le 1,$$

则称随机过程 $\{W^*(t), 0 \le t \le 1\}$ 为布朗桥.

布朗桥的性质

1. 布朗桥是高斯过程.

$$\begin{split} \forall 0 \leq s \leq t \leq 1 \; , \, \bar{\pi} \\ & \mathbb{E}\left[W^*(t)\right] = 0, \\ & \mathbb{E}\left[W^*(s)W^*(t)\right] = \mathbb{E}[(W(s) - sW(1))(W(t) - tW(1))] \\ & = \mathbb{E}\left[W(s)W(t) - tW(s)W(1) - sW(t)W(1) + tsW^2(1)\right] \\ & = s - ts - ts + ts = s(1 - t) \\ & = \min(s,t)[1 - \max(s,t)]. \end{split}$$

定理 4.15

布朗桥 $\{W^*(t), 0 \le t \le 1\}$ 的分布与 $\{W(t), 0 \le t \le 1 | W(0) = W(1) = 0\}$ 的分布相同.

2. 布朗桥也可以用来给出条件分布. 设 $0 \le a < b < c$, 则(W(a), W(c), W(b)) 服从多元正态分布

$$\mathcal{N}\left(\left(\begin{array}{c}0\\0\\0\end{array}\right),\left(\begin{array}{ccc}a&a&a\\a&c&b\\a&b&b\end{array}\right)\right).$$

在 W(a) = x, W(c) = z 条件下, 由联合正态分布性质可知 W(b) 服从条件正态分布, 条件均值为

$$\mathbb{E}(W(b) \mid (W(a), W(c)) = (x, z))$$

$$=0+(a,b)\begin{pmatrix} a & a \\ a & c \end{pmatrix}^{-1}(x,z)^{\top}$$
$$=\frac{c-b}{c-a}x+\frac{b-a}{c-a}z.$$

条件方差为

$$Var(W(b) \mid (W(a), W(c)) = (x, z))$$

$$=b-(a,b)\begin{pmatrix} a & a \\ a & c \end{pmatrix}^{-1}(a,b)^{\top}$$
$$=\frac{b(a+c-b)-ac}{c-a}.$$

如果 $b = \frac{a+c}{2}$,则条件分布为 $N\left(\frac{x+z}{2}, \frac{c-a}{4}\right)$. 比如取 $a = 0, b = \frac{1}{2}, c = 1$,则条件分布为 $N\left(\frac{x+z}{2}, \frac{1}{4}\right)$. 利用上述条件分布,可以用来生成 [0,1] 中的布朗运动的模拟轨道. 先抽取 W(0), W(1),利用条件分布抽取 $W\left(\frac{1}{2}\right)$,然后抽取 $W\left(\frac{1}{4}\right)$ 和 $W\left(\frac{3}{4}\right)$,然后抽取 $W\left(\frac{1}{8}\right)$, $W\left(\frac{5}{8}\right)$, $W\left(\frac{5}{8}\right)$,如此重复可以得到抽样间隔越来越密的布朗运动轨道的离散化值.

4.7.2 漂移布朗运动

定义 4.10 (漂移布朗运动)

设 $\{W(t), t > 0\}$ 为标准布朗运动, $\mu \in \mathbb{R}, \sigma > 0$,

$$X(t) = X(0) + \mu t + \sigma W(t), t \ge 0,$$

称 $\{W(t), t \geq 0\}$ 为有漂移的布朗运动, 漂移系数为 μ , 方差参数为 σ^2 , 记为 $WM\left(\mu, \sigma^2\right)$. 等价地, 若随机过程 $\{X(t), t \geq 0\}$ 为平稳独立增量过程, $X(t) - X(0) \sim \mathcal{N}\left(\mu t, \sigma^2 t\right)$, 则称 $\{X(t), t \geq 0\}$ 为漂移布朗运动.

定理 4.16

设随机过程 $\{X(t), t \geq 0\}$ 满足:

- 1. $\forall 0 < t_1 < \dots < t_n X(0), X(t_1) X(0), \dots, X(t_n) X(t_{n-1})$ 相互独立;
- 2. $\forall s \ge 0, t \ge 0, X(t) X(0)$ 与 X(t+s) X(s) 同分布 (没有要求正态分布);
- 3. 对于每一轨道 $\omega \in \Omega, X(t,\omega)$ 是 $t \in [0,\infty)$ 的连续函数,
- 4. $\forall t \geq 0, X(t) X(0)$ 分布关于 0 对称.

则 $\{X(t)\}$ 是布朗运动 (不要求从 0 出发), 或者 X(t)-X(0) 恒等于 0 . 如果仅要求 (1)-(3) 成立, 则必存在标准布朗运动 $\{W(t)\}$ 使得

$$X(t) = X(0) + \mu t + \sigma W(t), t \ge 0.$$

即这时 $\{X(t)\}$ 为漂移布朗运动.



第5章 随机积分与随机分析初步

内容提要

- □ 关于随机游动的随机积分
- □ 简单、一般被积函数的 Ito 积分
- □ Ito 过程及相关模型

- □ Ito 公式
- □ Black-Scholes 方程
- □ 多元随机分析

5.1 关于随机游动的随机积分

我们从讨论关于简单的随机游动的积分开始. 设 X_1, X_2, \ldots 相互独立, 都以各自 $\frac{1}{5}$ 概率分别取 +1 和 -1 值,

$$S_n = \sum_{k=1}^n X_k$$

是对称随机游动. $S_0 = 0$, $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. 令 B_n 是 \mathcal{F}_{n-1} 可测的随机变量序列, 比如它表示第 n 次赌博时所下赌注, 则它只能利用第 n-1 次及以前的信息, 而不能利用第 n 次赌博的结果. 于是到时刻 n 的收益 Z_n 为

$$Z_n = \sum_{i=1}^n B_i X_i = \sum_{i=1}^n B_i (S_i - S_{i-1}) = \sum_{i=1}^n B_i \Delta S_i,$$

这里 $\Delta S_i = S_i - S_{i-1}$, 我们称 Z_n 为 B_n 关于 S_n 的**随机积分**.

命题 **5.1** (Z_n 的性质)

- 1. $\{Z_n\}$ 是关于 \mathcal{F}_n 的鞅;
- 2. $\mathbb{E}[Z_n] = \mathbb{E}[Z_0] = 0$;
- 3. 如果假定 $\mathbb{E}[B_n^2] < \infty$,则

$$\operatorname{Var}\left[Z_{n}\right] = \mathbb{E}\left[Z_{n}^{2}\right] = \sum_{i=1}^{n} \mathbb{E}\left[B_{i}^{2}\right].$$

证明 容易看出 $\{Z_n\}$ 是关于 \mathcal{F}_n 的鞅. 事实上,

$$\mathbb{E}(Z_n \mid \mathcal{F}_{n-1}) = \mathbb{E}\left(\sum_{i=1}^{n-1} B_i X_i \mid \mathcal{F}_{n-1}\right) + \mathbb{E}(B_n X_n \mid \mathcal{F}_{n-1})$$

$$= \sum_{i=1}^{n-1} B_i X_i + B_n \mathbb{E}(X_n \mid \mathcal{F}_{n-1})$$

$$= \sum_{i=1}^{n-1} B_i X_i + B_n \mathbb{E}(X_n)$$

$$= \sum_{i=1}^{n-1} B_i X_i + 0$$

$$= Z_{n-1}$$

易见对 $0 \le m < n$ 有

$$\mathbb{E} [Z_n \mid \mathcal{F}_m]$$

$$= \mathbb{E} [\mathbb{E} (Z_n \mid \mathcal{F}_{n-1}) \mid \mathcal{F}_m] = \mathbb{E} [Z_{n-1} \mid \mathcal{F}_m]]$$

$$= \dots = \mathbb{E} [Z_{m+1} \mid \mathcal{F}_m]] = Z_m$$

下一步,

$$Z_n^2 = \sum_{i=1}^n B_i^2 X_i^2 + 2 \sum_{1 \le i < j \le n} B_i B_j X_i X_j$$

注意到 $X_i^2 = 1$, 如果 i < j, 则 B_i, X_i, B_j 都是 \mathcal{F}_{i-1} 可测的, 且 X_i 独立于 \mathcal{F}_{i-1} , 于是

$$\mathbb{E} [B_i B_j X_i X_j]$$

$$= \mathbb{E} \{ \mathbb{E} [B_i B_j X_i X_j \mid \mathcal{F}_{j-1}] \}$$

$$= \mathbb{E} \{ B_i B_j X_i \mathbb{E} [X_j \mid \mathcal{F}_{j-1}] \}$$

$$= \mathbb{E} \{ B_i B_j X_i \mathbb{E} [X_j] \} = 0.$$

5.2 简单被积函数的 Ito¹积分

给定正数 T, 我们来定义积分

$$\int_0^T \Delta(t)dW(t) \tag{5.1}$$

拿 笔记

- 1. 这里基本的要素是布朗运动 W(t), t > 0 以及相应的域流 $\mathcal{F}(t), t > 0$.
- 2. 要求被积函数 $\Delta(t)$ 是适应的随机过程 (即要求 $\Delta(t)$ 是适应过程意味着要求对每个 $t \geq 0$, $\Delta(t)$ 是 F(t)-可测的. 在时刻 t 能够获得的信息足以确定该时刻 $\Delta(t)$ 的值), 这是因为 $\Delta(t)$ 最终将是在时刻 t 我们持有资产的头寸, 通常依赖于截至时刻 t 的资产价格路径²;
- 3. 时刻 t 之后布朗运动的增量独立于 F(t), 而由于 $\Delta(t)$ 是 F(t)-可测的, 故 $\Delta(t)$ 也独立于未来的布朗运动增量. 我们所持有的资产头寸可以依赖于该资产价格的历史, 但是必须独立于驱动价格过程的布朗运动的未来增量.

想必大家都注意到了, 在定义 Ito 积分(5.1)时, 无法变成对时间求黎曼积分, 因为布朗运动路径不能对时间求 微分. 下面我们来看一看 Ito 是如何克服这个问题的.

5.2.1 简单函数 Ito 积分的构造

定义 5.1 (简单过程)

设 $\Pi = \{t_0, t_1, \cdots, t_n\}$ 是 [0, T] 的一个分划, 即

$$0 = t_0 < t_1 < \dots < t_n = T$$

假定 $\Delta(t)$ 在每个子区间 $[t_j,t_{j+1})$ 中是常量. 这样的过程 $\Delta(t)$ 称为简单过程.

设想简单过程 $\Delta(t)$ 与布朗运动 W(t) 按以下方式相互作用:将 W(t) 看作每份资产在时刻 t 的价格³设 t_0, t_1, \dots, t_{n-1} 是资产的交易日, $\Delta(t_0), \Delta(t_1), \dots, \Delta(t_{n-1})$ 是从每个交易日持有至下一个交易日的资产份额

¹伊藤清(英語:Kiyoshi Ito、1915 年 9 月 7 日-2008 年 11 月 10 日)、男、三重県北勢町生まれ、日本の著名な数学者の故、博士号、東京帝国大学(現東京大学)卒業. 1935 年から 1938 年まで東京帝国大学数学科で学んだ.1939 年から 1943 年まで政府統計局に勤務した.1943 年から 1952 年まで名古屋帝国大学(現名古屋大学)で助教授を務めた. 期間は 1945 年に東京帝国大学理学博士号を取得.1952 年から 1979 年まで京都大学で教授を務めた. 日本学士院院士に入選し、ウルフ数学賞、ガウス賞、京都賞、朝日賞、藤原賞を受賞した. 西洋の文献では彼の姓はよく It 舵と書かれている. ブラウン運動などの偶然性を伴う自然現象を説明するために、Ito 清は Ito 公式を提案し、これはこの数学の新しい分岐をランダムに分析する基礎定理となった.Ito 氏の成果は 1980 年代以降金融分野で広く使われており、「ウォール街で最も有名な日本人」と呼ばれている.

 $^{^2}$ 设 t 严格为正, 在初始时刻 $0, \Delta(t)$ 是未知的, 它是随机变量. 到了时刻 t, 我们就有足够的信息确定 $\Delta(t)$ 的值, 其随机性得以化解.

³由于布朗运动既取正值也取负值, 因此对于股票之类有限责任资产的价格, 它并非一个好的模型. 这里只是为了给出一个说明, 我们忽略这一点.

(即头寸). 因此在每个时刻 t 的收益为:

$$\begin{split} I(t) &= \Delta \left(t_0 \right) \left[W(t) - W \left(t_0 \right) \right] = \Delta (0) W(t), \quad 0 \leq t \leq t_1 \\ I(t) &= \Delta (0) W \left(t_1 \right) + \Delta \left(t_1 \right) \left[W(t) - W \left(t_1 \right) \right], \quad t_1 \leq t \leq t_2 \\ I(t) &= \Delta (0) W \left(t_1 \right) + \Delta \left(t_1 \right) \left[W \left(t_2 \right) - W \left(t_1 \right) \right] + \Delta \left(t_2 \right) \left[W(t) - W \left(t_2 \right) \right], \quad t_2 \leq t \leq t_3 \end{split}$$

等等. 一般地, 如果 $t_k < t < t_{k+1}$, 则:

$$I(t) = \sum_{j=0}^{k-1} \Delta(t_j) \left[W(t_{j+1}) - W(t_j) \right] + \Delta(t_k) \left[W(t) - W(t_k) \right]$$
(5.2)

上式中的过程 I(t) 就是简单过程 $\Delta(t)$ 的 Ito 积分, 记为:

$$I(t) = \int_0^t \Delta(u)dW(u)$$

5.2.2 简单函数 Ito 积分的性质

命题 5.2 (积分的性质)

- 1. 由式(5.2)定义的积分是一个鞅.
- 2. Ito 等距: 由式(5.2)定义的 Ito 积分满足:

$$\mathbb{E}I^2(t) = \mathbb{E}\int_0^t \Delta^2(u)du$$

3. 二次变差: 截止时刻 t,Ito 积分(5.2)累积的二次变差为:

$$[I,I](t) = \int_0^t \Delta^2(u) du$$

证明 下面分别证明积分性质 1-3.

1. 给定 $0 \le s \le t \le T$. 我们假定 s 和 t 位于分划 Π 的不同的子区间内 (即存在分点 t_l 和 t_k , 使得 $t_l < t_k, s \in [t_l, t_{l+1})$, 且 $t \in [t_k, t_{k+1})$). 如果 s 和 t 在同一子区间中,则以下证明可以简化. 将等式(5.2)改写为:

$$I(t) = \sum_{j=0}^{t-1} \Delta(t_j) [W(t_{j-1}) - W(t_j)] + \Delta(t_l) [W(t_{l+1}) - W(t_l)]$$
$$+ \sum_{j=l+1}^{k-1} \Delta(t_j) [W(t_{j-1}) - W(t_j)] + \Delta(t_k) [W(t) - W(t_k)]$$

即证明 $[I(t) \mid \mathcal{F}(s)] = I(s)$. 关于上式右端四项中的每一项取条件期望. 第一项和式 $\sum_{j=0}^{l-1} \Delta(t_j) [W(t_{j+1}) - W(t_j)]$ 中所有随机变量都是 $\mathcal{F}(s)$ 可测的, 因此:

$$\mathbb{E}\left[\sum_{j=0}^{t-1} \Delta\left(t_{j}\right) \left[W\left(t_{j+1}\right) - W\left(t_{j}\right)\right] \mid \mathcal{F}(s)\right] = \sum_{j=0}^{t-1} \Delta\left(t_{j}\right) \left[W\left(t_{j+1}\right) - W\left(t_{j}\right)\right]$$

右端第二项写为:

$$\mathbb{E}\left[\Delta\left(t_{l}\right)\left(W\left(t_{l+1}\right)-W\left(t_{l}\right)\right)\mid\mathcal{F}(s)\right] = \Delta\left(t_{l}\right)\left(\mathbb{E}\left[W\left(t_{l+1}\right)\mid\mathcal{F}(s)\right]-W\left(t_{l}\right)\right)$$
$$= \Delta\left(t_{l}\right)\left(W(s)-W\left(t_{l}\right)\right)$$

第二个等号是来自于鞅的性质. 接下来证明右端第三项和第四项条件期望为 0. 利用重期望公式、鞅的性质,有:

$$\mathbb{E}\left[\Delta\left(t_{j}\right)\left(W\left(t_{j+1}\right)-W\left(t_{j}\right)\right)\mid\mathcal{F}(s)\right]$$

$$=\mathbb{E}\left[\mathbb{E}\left[\Delta\left(t_{j}\right)\left(W\left(t_{j+1}\right)-W\left(t_{j}\right)\right)\mid\mathcal{F}\left(t_{j}\right)\right]\mid\mathcal{F}(s)\right]$$

$$=\mathbb{E}\left[\Delta\left(t_{j}\right)\left(\mathbb{E}\left[W\left(t_{j+1}\right)\mid\mathcal{F}\left(t_{j}\right)\right]-W\left(t_{j}\right)\right)\mid\mathcal{F}(s)\right]$$

$$=\mathbb{E}\left[\Delta\left(t_{j}\right)\left(W\left(t_{j}\right)-W\left(t_{j}\right)\right)\mid\mathcal{F}(s)\right]=0$$

则第三项条件期望为0;第四项类似.把四项条件期望相加即得证.

2. 第二个性质其实就是 Ito 积分的方差. 首先注意到 $\mathbb{E}I(t) = 0$, $\text{Var}\,I(t) = \mathbb{E}I^2(t)$. 令 $D_j = W(t_{j+1}) - W(t_j)$, $j = 0, \dots, k-1$ 以及 $D_k = W(t) - W(t_k)$, 于是式(5.2)可以写成 $I(t) = \sum_{j=0}^k \Delta(t_j) D_j$ 并且

$$I^{2}(t) = \sum_{j=0}^{k} \Delta^{2}(t_{j}) D_{j}^{2} + 2 \sum_{0 \leq i < j \leq k} \Delta(t_{i}) \Delta(t_{j}) D_{i} D_{j}$$

首先证明每一交叉项的期望值为 0. 对于 i < j, 随机变量 $\Delta(t_i)\Delta(t_j)D_i$ 为 $\mathcal{F}(t_j)$ -可测, 而布朗运动增量 D_i 独立于 $\mathcal{F}(t_i)$, 并且 $ED_i = 0$. 因此:

$$\mathbb{E}\left[\Delta\left(t_{i}\right)\Delta\left(t_{j}\right)D_{i}D_{j}\right] = \mathbb{E}\left[\Delta\left(t_{i}\right)\Delta\left(t_{j}\right)D_{i}\right] \cdot \mathbb{E}D_{j} = \mathbb{E}\left[\Delta\left(t_{i}\right)\Delta\left(t_{j}\right)D_{i}\right] \cdot 0 = 0$$

然后考虑平方项 $\Delta^2(t_j)$ D_j^2 . 随机变量 $\Delta^2(t_j)$ 为 $\mathcal{F}(t_j)$ -可测, 而布朗运动增量的平方 D_j^2 独立于 $\mathcal{F}(t_j)$, 并且 $\mathbb{E}D_i^2=t_{j+1}-t_j$, $j=0,\cdots,k-1$ 以及 $\mathbb{E}D_k^2=t-t_k$ (布朗运动的定义). 因此:

$$\mathbb{E}I^{2}(t) = \sum_{j=0}^{k} \mathbb{E}\left[\Delta^{2}(t_{j}) D_{j}^{2}\right] = \sum_{j=0}^{k} \mathbb{E}\Delta^{2}(t_{j}) \cdot \mathbb{E}D_{j}^{2} = \sum_{j=0}^{k-1} \mathbb{E}\Delta^{2}(t_{j})(t_{j+1} - t_{j}) + \mathbb{E}\Delta^{2}(t_{k})(t - t_{k})$$

由于在区间 $[t_j, t_{j+1})$ 上 $\Delta(t_j)$ 是常数, 因此 $\Delta^2(t_j)(t_{j+1} - t_j) = \int_{t_j}^{t_{j+1}} \Delta^2(u) du$. 类似地, $\Delta^2(t_k)(t - t_k) = \int_{t_k}^t \Delta^2(u) du$. 于是可得:

$$\mathbb{E}I^{2}(t) = \sum_{j=0}^{k-1} \mathbb{E} \int_{t_{j}}^{t_{j+1}} \Delta^{2}(u) du + \mathbb{E} \int_{t_{k}}^{t} \Delta^{2}(u) du$$
$$= \mathbb{E} \left[\sum_{j=0}^{k-1} \int_{t_{j}}^{t_{j+1}} \Delta^{2}(u) du + \int_{t_{k}}^{t} \Delta^{2}(u) du \right] = \mathbb{E} \int_{0}^{t} \Delta^{2}(u) du$$

3. 先计算 Ito 积分在子区间 $[t_j,t_{j+1}]$ 上 (此时 $\Delta(u)$ 在 $[t_j,t_{j-1})$ 上是常量) 累积的二次变差. 为此, 取分点

$$t_i = s_0 < s_1 < \dots < s_m = t_{i+1}$$

并考虑

$$\sum_{i=0}^{m-1} \left[I\left(s_{i+1}\right) - I\left(s_{i}\right) \right]^{2} = \sum_{i=0}^{m-1} \left[\Delta\left(t_{j}\right) \left(W\left(s_{i+1}\right) - W\left(s_{i}\right) \right) \right]^{2}$$
$$= \Delta^{2}\left(t_{j}\right) \sum_{i=0}^{m-1} \left(W\left(s_{i+1}\right) - W\left(s_{i}\right) \right)^{2}$$

当 $m \to \infty$ 并且最大步幅 $\max_{i=0,\dots,m-1} (s_{i+1}-s_i)$ 趋于 0 时, $\sum_{i=0}^{m-1} (W(s_{i+1})-W(s_i))^2$ 收敛于布朗运动在时刻 t_j 与 t_{j+1} 之间累积的二次变差 $t_{j+1}-t_j$. 因此,上式的极限,即 Ito 积分在时刻 t_j 与 t_{j+1} 之间累积的二次变差为:

$$\Delta^{2}(t_{j})(t_{j+1} - t_{j}) = \int_{t_{j}}^{t_{j+1}} \Delta^{2}(u) du$$

类似地,Ito 积分在时刻 t_k 与 t 之间累积的二次变差为 $\int_{t_k}^t \Delta^2(u) du$, 把所有的子区间相加则得.

🕏 笔记

- 由性质 2-3, 我们可以看出一个过程的二次变差和方差是不同的.
- 二次变差结果依赖于路径. 二次变差可以被视为风险的度量, 他依赖于我们选择头寸的大小.
- 方差是二次变差关于所有可能路径的平均值.
- 回顾布朗运动的二次变差:[W,W](t)=t,即布朗运动在单位时间内累积二次变差的速率为 1; 而 Ito 积分在单位时间内累积二次变差的速率为 $\Delta^2(t)$.
- 另外,Ito 积分还可以写作:

$$dI(t) = \Delta(t)dW(t)$$

这被称为 Ito 积分的微分形式.

5.3 一般被积函数的 Ito 积分

假定 $\Delta(t), t \ge 0$ 适应于域流 $\mathcal{F}(t), t \ge 0$, 并且满足平方可积条件⁴:

$$\mathbb{E} \int_0^T \Delta^2(t) dt < \infty$$

我们把满足上述条件的随机过程组成的集合记为 $\mathcal{H} = \{\Delta(t): 满足平方可积、适应性\}$

5.3.1 一般被积函数 Ito 积分的构造

我们用简单过程逼近 $\Delta(t)$, 这与实变函数中的思想一致. 取分划 $0 = t_0 < t_1 < t_2 < \cdots$, 在每一点 t_j 令简单过程等于 $\Delta(t_j)$, 并且在子区间 $[t_j, t_{j+1})$ 上保持不变. 当分划的细度趋近于 0 时, 简单过程将越来越趋近于连续变化的被积函数.

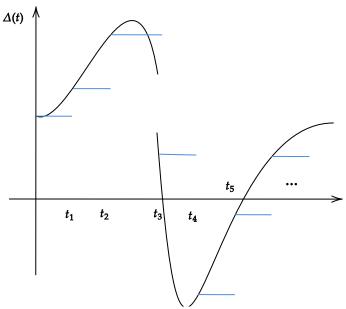


图 5.1: 简单过程逼近连续变化的被积函数

引理 5.1

任给 $\Delta(t) \in \mathcal{H}$, 存在一列简单过程 Δ_n , 使得当 $n \to \infty$ 时,

$$||\Delta_n - \Delta||_{\mathcal{H}} \to 0$$

其中
$$||\Delta||_{\mathcal{H}}^2 = \mathbb{E} \int_0^t \Delta^2(u) du$$

定义 5.2 (Ito 积分)

由上述引理保证, 可以选取一列简单过程 $\Delta_n(t)$, 当 $n \to \infty$ 时收敛于连续变化的 $\Delta(t)$:

$$\lim_{n \to \infty} \mathbb{E} \int_0^T \left| \Delta_n(t) - \Delta(t) \right|^2 dt = 0$$

对于每个 $\Delta_n(t)$, Ito 积分 $\int_0^t \Delta_n(u)dW(u)$ 关于 $0 \le t \le T$ 已经都有定义. 连续变化的被积函数 $\Delta(t)$ 的 Ito 积分定义为:

$$\int_0^t \Delta(u)dW(u) = \lim_{n \to \infty} \int_0^t \Delta_n(u)dW(u), \quad 0 \le t \le T$$

⁴注意,没有假定连续变化,可以有跳跃

我们有必要说明这个定义式中, 右边极限的存在性. 根据引理可知 Δ_n 是 Cauchy 列 (因为简单过程也属于 \mathcal{H}), 于是有:

$$\lim_{n>m\to\infty} ||\Delta_n - \Delta_m||_{\mathcal{H}} = 0$$

根据等距性, 我们有: $||I_n(t)||_2^2 = ||\Delta_n||_{\mathcal{H}}^{2.5}$, 所以根据引理有

$$||I_n(t) - I_m(t)||_2 \to 0$$

所以对每个 $t,I_n(t)=\int_0^t \Delta_n(u)dW(u)$ 是 $L_2(\Omega,\mathcal{F},\mathbb{P})$ 中的 Cauchy 列. 因此极限存在.

5.3.2 Ito 积分的性质

由于是用简单过程逼近的, 所以根据简单函数 Ito 积分的形式, 我们有如下结论:

命题 5.3 (Ito 积分的极限表达式)

设 $\Delta(t)$ 关于域流F(t)是适应的,且满足平方可积性,则

$$\int_{0}^{t} \Delta(W(u))dW(u) = \lim_{n \to \infty} \sum_{i=1}^{n} \Delta(W(t_{i-1}))(W(t_{i}) - W(t_{i-1})); \tag{5.3}$$

$$\int_{0}^{t} \Delta(u)dW(u) = \lim_{n \to \infty} \sum_{i=1}^{n} \Delta(t_{i}^{*})(W(t_{i}) - W(t_{i-1})).$$
(5.4)

其中 $0 < t_1 < \dots < t_n = t$, 且 $\max_{1 \le i \le n} (t_i - t_{i-1}) \to 0, t_i^* \in [t_{i-1}, t_i]$.

这样定义的 Ito 积分继承了简单过程 Ito 积分的许多性质, 总结为如下定理:

定理 5.1 (Ito 积分的性质)

- 1. (连续性) 作为积分上限 t 的函数, I(t) 的路径连续.
- 2. (适应性) 对每个 t, I(t) 为 F(t)-可测.

3. (线性性) 如果
$$I(t) = \int_0^t \Delta(u)dW(u), J(t) = \int_0^t \Gamma(u)dW(u), 则 I(t) \pm J(t) = \int_0^t (\Delta(u) \pm \Gamma(t))dW(u);$$
 对任意常数 $c,cI(t) = \int_0^t c\Delta(u)dW(u)$.

- 4. (鞅性质) *I(t)* 是鞅.
- 5. (Ito 等距) $\mathbb{E}I^2(t) = \mathbb{E}\int_0^t \Delta^2(u)du$.
- 6. (二次变差) $[I,I](t) = \int_0^t \Delta^2(u) du$.

例题 5.1 计算积分 $\int_0^T W(t)dW(t)$.

取被积函数 $\Delta(t) = W(t)$ 的如下逼近:

$$\Delta_n(t) = \begin{cases} W(0) = 0, & 0 \le t < \frac{T}{n} \\ W\left(\frac{T}{n}\right), & \frac{T}{n} \le t < \frac{2T}{n} \end{cases}$$
$$\vdots$$
$$W\left(\frac{(n-1)T}{n}\right), & \frac{(n-1)T}{n} \le t < T$$

 $[|]I_n(t)||_2^2 = \mathbb{E}(I_n^2(t))$

于是 $\lim_{n\to\infty} \mathbb{E} \int_0^T \left| \Delta_n(t) - W(t) \right|^2 dt = 0$. 由定义知:

$$\int_0^T W(t)dW(t) = \lim_{n \to \infty} \int_0^T \Delta_n(t)dW(t)$$
(5.5)

$$= \lim_{n \to \infty} \sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right) \left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right) \right]$$
 (5.6)

记 $W_j = W\left(\frac{jT}{n}\right)$. 我们有:

$$\begin{split} \frac{1}{2} \sum_{j=0}^{n-1} \left(W_{j+1} - W_j \right)^2 &= \frac{1}{2} \sum_{j=0}^{n-1} W_{j+1}^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} \sum_{k=1}^{n} W_k^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} W_n^2 + \frac{1}{2} \sum_{k=0}^{n-1} W_k^2 - \sum_{j=0}^{n-1} W_j W_{j+1} + \frac{1}{2} \sum_{j=0}^{n-1} W_j^2 \\ &= \frac{1}{2} W_n^2 + \sum_{j=0}^{n-1} W_j^2 - \sum_{j=0}^{n-1} W_j W_{j+1} \\ &= \frac{1}{2} W_n^2 + \sum_{j=0}^{n-1} W_j \left(W_j - W_{j+1} \right) \end{split}$$

由上式可得:

$$\sum_{j=0}^{n-1} W_j (W_{j+1} - W_j) = \frac{1}{2} W_n^2 - \frac{1}{2} \sum_{j=0}^{n-1} (W_{j+1} - W_j)^2$$

采用原先的记号,即有:

$$\begin{split} &\sum_{j=0}^{n-1} W\left(\frac{jT}{n}\right) \left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right] \\ &= \frac{1}{2} W^2(T) - \frac{1}{2} \sum_{j=0}^{n-1} \left[W\left(\frac{(j+1)T}{n}\right) - W\left(\frac{jT}{n}\right)\right]^2 \end{split}$$

令 $n \to \infty$, 我们得到:

$$\int_0^T W(t)dW(t) = \frac{1}{2}W^2(T) - \frac{1}{2}[W,W](T) = \frac{1}{2}W^2(T) - \frac{1}{2}T$$

$$\int_0^T g(t)dg(t) = \int_0^T g(t)g'(t)dt = \frac{1}{2}g^2(t)\bigg|_0^T = \frac{1}{2}g^2(T)$$

额外的项 $-\frac{1}{2}T$ 来自布朗运动的非零二次变差以及 Ito 积分的构造: 积分和式中, 被积函数总是在子区间左端点取值 (见式子(5.6)以及5.3). 如果改为在子区间中点取值,则(5.6)右端变为:

$$\lim_{n \to \infty} \sum_{j=0}^{n-1} W \left[\frac{\left(j + \frac{1}{2}\right)T}{n} \right] \left[W \left(\frac{(j+1)T}{n} \right) - W \left(\frac{jT}{n} \right) \right]$$

就不会得到这一项6.

⁶这样得到的积分被称为 Stratonovich 积分, 并且通常微积分的运算法则适用. 但是, 这种积分不适合金融建模. 在金融中, 被积函数是我们持有资产的头寸. 我们不可能在下午 13:00 决定上午 9:00 所取的头寸. 我们必须在每个时间区间的起始点就决定持有资产的头寸, 这样的资产交易收益的极限就是 Ito 积分

5.4 Ito 公式

对于普通的微积分, 我们有微分链式法则:

$$df(g(t)) = f'(g(t))dg(t) = f'(g(t))g'(t),$$

和积分变量替换法则:

$$\int_0^t f'(g(s))dg(s)$$

$$= \int_{g(0)}^{g(t)} f'(u)du$$

$$= f(g(t)) - f(g(0)).$$

对上式求导,则得到

$$f'(g(t))dg(t) = d[f(g(t))]$$

如果是 Ito 积分, 是否有:

$$df(W(t)) = f'(W(t))dW(t),$$

或者

$$\int_0^t f'(W(s))dW(s) = f(W(t)) - f(W(0))?$$

答案是否定的, 因为 W(t) 不可微. 现给出一个重要引理7.

引理 5.2

设
$$g(x)$$
 是 \mathbb{R} 上的连续函数, 满足 $\int_0^T \mathbb{E}\left[g^2(W(u))\right] du < \infty, \{t_i\}$ 是 $[0,t]$ 的分割, 有
$$\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} g\left(W(t_i)\right) \left[W\left(t_{i+1}\right) - W\left(t_i\right)\right]^2 = \int_0^t g(W(s)) ds$$

证明 按黎曼积分定义有

$$\sum_{i=0}^{n-1} g(W(t_i)) (t_{i+1} - t_i) \to \int_0^t g(W(s)) ds, \text{ a.s.}$$

存在有限. 这个极限是 a.s. 收敛, 也依概率收敛. 下证

$$\left\| \sum_{i=0}^{n-1} g(W(t_i)) (W(t_{i+1}) - W(t_i))^2 - \sum_{i=0}^{n-1} g(W(t_i)) (t_{i+1} - t_i) \right\|^2 \to 0 (n \to \infty)$$

由布朗运动的独立增量性,有:

$$\sum_{i=0}^{n-1} \mathbb{E}\{g(B(t_i))[(W(t_{i+1}) - W(t_i))^2 - (t_{i+1} - t_i)]\} = \sum_{i=0}^{n-1} \mathbb{E}g(W(t_i))\mathbb{E}[(W(t_{i+1}) - W(t_i))^2 - (t_{i+1} - t_i)] = 0$$

 $^{^{7}}$ 实际上就是 $\lim_{\delta_{n}\to 0}\sum_{i=0}^{n-1}\left[W\left(t_{i+1}\right)-W\left(t_{i}\right)\right]^{2}=t$ 即4.6的推广形式

以及

$$\operatorname{Var} \left\{ \sum_{i=0}^{n-1} g\left(W\left(t_{i}\right)\right) \left[\left(W\left(t_{i+1}\right) - W\left(t_{i}\right)\right)^{2} - \left(t_{i+1} - t_{i}\right) \right] \right\}$$

$$= \sum_{i=0}^{n-1} \mathbb{E} \left[g\left(W\left(t_{i}\right)\right) \right]^{2} \mathbb{E} \left[\left(W\left(t_{i+1}\right) - W\left(t_{i}\right)\right)^{2} - \left(t_{i+1} - t_{i}\right) \right]^{2}$$

$$= 2 \sum_{i=0}^{n-1} \mathbb{E} \left[g\left(W\left(t_{i}\right)\right) \right]^{2} \left(t_{i+1} - t_{i}\right)^{2}$$

$$\leq \frac{2t^{2}}{n} \sup_{x \in \mathbb{R}} |g(x)| \to 0, \quad n \to \infty.$$

综合起来就得,在均方收敛意义下,有

$$\lim_{\delta_n \to 0} \sum_{i=0}^{n-1} g(W(t_i)) [W(t_{i+1}) - W(t_i)]^2 = \int_0^t g(W(s)) ds$$

定理 5.2 (Ito 公式)

$$f(W(t)) - f(W(0)) = \int_0^t f'(W(u))dW(u) + \frac{1}{2} \int_0^t f''(W(u))du$$
 (5.7)

注: 可以简写成

$$df(W(t)) = f'(W(t))dW(t) + \frac{1}{2}W''(t)dt$$

证明 假设存在 M>0 使得当 |x|>M 时, f(x)=0. 任给 $n\geq 1$, 将 [0,t] 进行 n 等分: $t_i=\frac{it}{n}, 0\leq i\leq n$. 显然就有

$$f(W(t)) - f(W(0)) = \sum_{i=1}^{n} [f(W(t_i)) - f(W(t_{i-1}))], i = 1, 2, \dots, n$$

由于 f(x) 具有二次连续导数,因此对任意 x, y 有8

$$f(y) - f(x) = f'(x)(y - x) + \frac{1}{2}f''(x)(y - x)^{2} - \int_{-\infty}^{y} (y - u)(F''(x) - F''(u))du$$

其中 $\left|\int_x^y (f''(x)-f''(u))du\right| \le (y-x)^2 G(x,y)$, G(x,x)=0 且 G(x,y) 是一致连续有界函数. 且由控制收敛定理: $\mathbb{E}(G(t_i,t_{i-1}))^2 \to 0$.

令
$$x = W(t_{i-1}), y = W(t_i), 则第一项:$$

$$\sum_{i=1}^{n} f'(W(t_{i-1}))(W(t_i) - W(t_{i-1})) \to \int_{0}^{t} f'(W(s))dW(s)$$

由引理,第二项满足

$$\sum_{i=1}^{n} f''(W(t_{i-1}))(W(t_i) - W(t_{i-1}))^2 \to \int_0^t f''(W(s))ds$$

往证:

$$\sum_{i=1}^{n} (W(t_i) - W(t_{i-1}))^2 G(t_i, t_{i-1}) \to 0$$

由 Cauchy-Schwarz 不等式,有

 $\mathbb{E}\left[(W(t_i) - W(t_{i-1}))^2 G(t_i, t_{i-1})\right] \leq \left[\mathbb{E}(W(t_i) - W(t_{i-1}))^4 \mathbb{E}(G(t_i, t_{i-1}))^2\right]^{\frac{1}{2}} \leq \frac{\sqrt{3}t}{n} \max_{i \leq i \leq n} \left[\mathbb{E}(G(t_i, t_{i-1}))^2\right]^{\frac{1}{2}} \to 0$ 三部分和在一块证得. 对于一般的 f 采用截尾逼近的方式即可.

⁸泰勒展开的积分型余项

推论 5.1 (Ito-Döeblin 公式)

设函数 f(t,x) 的偏导数 $f_t(t,x)$, $f_x(t,x)$ 和 $f_{xx}(t,x)$ 都有定义并且连续, W(t) 是布朗运动, 则对于每个 $T\geq 0$, 有:

$$f(T, W(T)) = f(0, W(0)) + \int_0^T f_t(t, W(t))dt + \int_0^T f_x(t, W(t))dW(t) + \frac{1}{2} \int_0^T f_{xx}(t, W(t))dt$$

证明 如果取时间变量 t 以及变量 x 的函数 f(t,x), 则由泰勒定理得:

$$f(t_{j+1}, x_{j+1}) - f(t_j, x_j)$$

$$= f_t(t_j, x_j) (t_{j+1} - t_j) + f_x(t_j, x_j) (x_{j+1} - x_j)$$

$$+ \frac{1}{2} f_{xx} (t_j, x_j) (x_{j+1} - x_j)^2 + f_{tx} (t_j, x_j) (t_{j+1} - t_j) (x_{j+1} - x_j)$$

$$+ \frac{1}{2} f_{tt} (t_j, x_j) (t_{j+1} - t_j)^2 + \text{高阶项}$$
在上式中令 $x_j = W(t_j), x_{j+1} = W(t_{j+1}),$ 然后相加:
$$f(T, W(T)) - f(0, W(0))$$

$$= \sum_{j=0}^{n-1} [f(t_{j+1}, W(t_{j+1})) - f(t_j, W(t_j))]$$

$$= \sum_{j=0}^{n-1} f_t (t_j, W(t_j)) (t_{j+1} - t_j) + \sum_{j=0}^{n-1} f_x (t_j, W(t_j)) (W(t_{j+1}) - W(t_j))$$

$$+ \frac{1}{2} \sum_{j=0}^{n-1} f_{xx} (t_j, W(t_j)) (W(t_{j+1}) - W(t_j))^2$$

$$+ \sum_{j=0}^{n-1} f_{tx} (t_j, W(t_j)) (t_{j+1} - t_j) (W(t_{j+1}) - W(t_j))$$

$$+ \frac{1}{2} \sum_{j=0}^{n-1} f_{tt} (t_j, W(t_j)) (t_{j+1} - t_j)^2 + \text{高阶项}$$

右边取极限有

$$\lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} f_t(t_j, W(t_j)) (t_{j+1} - t_j) = \int_0^T f_t(t, W(t)) dt$$

以及第二项收敛于 Ito 积分 $\int_0^T f_x(t,W(t))dW(t)$. 第三项由引理, 收敛到积分 $\frac{1}{2}\int_0^T f_{xx}(t,W(t))dt$. 第四项:

$$\lim_{\|\Pi\|} \left| \sum_{j=0}^{n-1} f_{tx} \left(t_{j}, W \left(t_{j} \right) \right) \left(t_{j+1} - t_{j} \right) \left(W \left(t_{j+1} \right) - W \left(t_{j} \right) \right) \right| \\
\leq \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \left| f_{tx} \left(t_{j}, W \left(t_{j} \right) \right) \right| \cdot \left(t_{j+1} - t_{j} \right) \cdot \left| W \left(t_{j+1} \right) - W \left(t_{j} \right) \right| \\
\leq \lim_{\|\Pi\| \to 0} \max_{0 \leq k \leq n-1} \left| W \left(t_{k+1} \right) - W \left(t_{k} \right) \right| \cdot \lim_{\|\|\| \to 0} \sum_{j=0}^{n-1} \left| f_{tx} \left(t_{j}, W \left(t_{j} \right) \right) \right| \left(t_{j+1} - t_{j} \right) \\
= 0 \cdot \int_{0}^{T} \left| f_{tx} (t, W(t)) \right| dt = 0$$

第五项也可以同样推导:

$$\lim_{\|\Pi\| \to 0} \left| \frac{1}{2} \sum_{j=0}^{n-1} f_{tt} (t_j, W (t_j)) (t_{j+1} - t_j)^2 \right| \\
\leq \lim_{\|\Pi\| \to 0} \frac{1}{2} \sum_{j=0}^{n-1} |f_{tt} (t_j, W (t_j))| \cdot (t_{j+1} - t_j)^2 \\
\leq \frac{1}{2} \lim_{\|\Pi\| \to 0} \max_{0 \le k \le n-1} (t_{k+1} - t_k) \cdot \lim_{\|\Pi \to 0} \sum_{j=0}^{n-1} |f_{tt} (t_j, W (t_j))| (t_{j+1} - t_j) \\
= \frac{1}{2} \cdot 0 \cdot \int_{0}^{T} |f_{tt} (t, W (t))| dt = 0$$

类似地, 所有高阶项的极限均为0.

🕏 笔记 在证明中用到了交互变差4.7极限为 0. 类似地, 还有

$$\lim_{\|\Pi\| \to 0} \sum_{i=0}^{n-1} (t_{i+1} - t_i)^2 = 0$$

我们可以把二次变差、交互变差、时间二次变差形式地记为:

$$dW(t)dW(t) = dt$$

$$dtdW(t) = dW(t)dt = 0$$

$$dtdt = 0$$
(5.8)

因此微分形式

$$\begin{split} df(t, W(t)) &= f_t(t, W(t))dt + f_x(t, W(t))dW(t) + \frac{1}{2}f_{xx}(t, W(t))dW(t)dW(t) \\ &+ f_{tx}(t, W(t))dtdW(t) + \frac{1}{2}f_{tt}(t, W(t))dtdt \end{split}$$

可以简写为:

$$df(t, W(t)) = f_t(t, W(t))dt + f_x(t, W(t))dW(t) + \frac{1}{2}f_{xx}(t, W(t))dt$$
(5.9)

Ŷ 笔记 在一般微积分中,从

$$\frac{1}{2} \sum_{j=0}^{n-1} f_{xx}(t_j, x_j) (x_{j+1} - x_j)^2$$

这一项开始极限就为 0 了, 因为连续可导函数二次变差为 0, 而对于布朗运动则不然. 因为被积函数在左端点取值, 所以才得到二次变差项, 如果不是在左端点取值, 就会得到其它值. 这也是 Ito 积分的特殊之处.

例题 5.2 计算 $\int_0^t W(u)dW(u)$.

在上一节用逼近给出了积分结果. 下面用 Itô 公式计算. 这里 f'(x)=x , 从而 $f(x)=\frac{1}{2}x^2$, f''(x)=1 , 则由 Ito 公式

$$\begin{split} \frac{1}{2}W^2(t) &= f(W(t)) \\ &= f(0) + \int_0^t f'(W(s))dW(s) + \frac{1}{2} \int_0^t f''(W(s))ds \\ &= 0 + \int_0^t W(s)dW(s) + \frac{1}{2} \int_0^t 1ds \\ &= \int_0^t W(s)dW(s) + \frac{t}{2}. \end{split}$$

即有

$$\int_0^t W(s)dW(s) = \frac{1}{2}W^2(t) - \frac{t}{2}.$$

例题 **5.3** 定理**4.6**的证明. 定理中叙述的过程 M(t) 也是一个具有非 0 二次变差的过程, 因此 Ito 公式(5.9)也同样适用. 对任何存在连续导数的函数 f(t,x), 有:

$$df(t, M(t)) = f_t(t, M(t)) + f_x(t, M(t))dM(t) + \frac{1}{2}f_{xx}(t, M(t))dt$$

上式的积分形式是:

$$f(t, M(t)) = f(0, M(0)) + \int_0^t \left(f_t(s, M(s)) + \frac{1}{2} f_{xx}(s, M(s)) \right) ds + \int_0^t f_x(s, M(s)) dM(s)$$

可证明 $^9\int_0^t f_x(s,M(s))dM(s)$ 是鞅. 上式两边取期望, 得到:

$$\mathbb{E}f(t, M(t)) = f(0, M(0)) + \mathbb{E}\int_{0}^{t} \left(f_{t}(s, M(s)) + \frac{1}{2} f_{xx}(s, M(s)) \right) ds$$

固定 u, 定义 $f(t,x) = \exp\left\{ux - \frac{1}{2}u^2t\right\}$, 则:

$$f_t(t,x) = -\frac{1}{2}u^2 f(t,x), f_x(t,x) = uf(t,x), f_{xx}(t,x) = u^2 f(t,x)$$

于是有 $f_t + \frac{1}{2}f_{xx} = 0$. 进而有:

$$\mathbb{E}f(t, M(t)) = \mathbb{E}\exp\left\{uM(t) - \frac{1}{2}u^2t\right\} = 1$$

因此得到 M(t) 的矩母函数为

$$\mathbb{E}e^{uM(t)} = e^{\frac{1}{2}u^2t}$$

因此 $M(t) \sim \mathcal{N}(0,t)$.

5.5 Ito 积分过程

随机分析中主要考虑的是以下定义的 Ito 过程.

5.5.1 Ito 过程

定义 5.3 (Ito 过程)

设 $W(t), t \ge 0$ 是布朗运动, $F(t), t \ge 0$ 是相应的域流; Ito 过程是以下形式的随机过程:

$$X(t) = X(0) + \int_0^t \Delta(u)dW(u) + \int_0^t \Theta(u)du$$

其中 X(0) 非随机, $\Delta(u)$ 和 $\Theta(u)$ 是适应的随机过程.

上述定义的 Ito 过程的简写形式

$$dX(t) = \Delta(t)dW(t) + \Theta(t)dt$$

被称为随机微分方程.

上述定义其实就是 Ito 积分加上一个初始常值和一般黎曼积分; 常数和黎曼积分对于累积二次变差是没用的, 所以有:

⁹先用简单函数证明,一般情形用简单函数逼近

引理 5.3

Ito 过程的二次变差为:

$$[X,X](t) = \int_0^t \Delta^2(u)du$$

定义 5.4 (关于 Ito 过程的积分)

设 $X(t), t \ge 0$ 是定义5.3中描述的 Ito 过程, $\Gamma(t), t \ge 0$ 是一个适应过程. 关于 Ito 过程的积分定义为

$$\int_0^t \Gamma(u) dX(u) = \int_0^t \Gamma(u) \Delta(u) dW(u) + \int_0^t \Gamma(u) \Theta(u) du$$

定理 5.3 (Ito 过程的 Ito 公式)

设 $X(t), t \ge 0$ 是 Ito 过程, 函数 f(t,x) 的偏导数 $f_t(t,x), f_x(t,x)$ 和 $f_{xx}(t,x)$ 都有定义并且连续, 则对于每个 T > 0, 有:

$$\begin{split} f(T,X(T)) = & f(0,X(0)) + \int_0^T f_t(t,X(t))dt + \int_0^T f_x(t,X(t))dX(t) \\ & + \frac{1}{2} \int_0^T f_{xx}(t,X(t))d[X,X](t) \\ = & f(0,X(0)) + \int_0^T f_t(t,X(t))dt + \int_0^T f_x(t,X(t))\Delta(t)dW(t) \\ & + \int_0^T f_x(t,X(t))\Theta(t)dt + \frac{1}{2} \int_0^T f_{xx}(t,X(t))\Delta^2(t)dt \end{split}$$

证明 类似推论5.1的证明,有

$$f(T, X(T)) - f(0, X(0))$$

$$= \sum_{j=0}^{n-1} f_t(t_j, X(t_j)) (t_{j+1} - t_j) + \sum_{j=0}^{n-1} f_x(t_j, X(t_j)) (X(t_{j+1}) - X(t_j))$$

$$+ \frac{1}{2} \sum_{j=0}^{n-1} f_{xx}(t_j, X(t_j)) (X(t_{j+1}) - X(t_j))^2$$

$$+ \sum_{j=0}^{n-1} f_{tx}(t_j, X(t_j)) (t_{j+1} - t_j) (X(t_{j+1}) - X(t_j))$$

$$+ \frac{1}{2} \sum_{j=0}^{n-1} f_{tt}(t_j, X(t_j)) (t_{j+1} - t_j)^2 + \bar{\beta} \bar{\beta} \bar{\beta} \bar{\beta}$$

上式最后两项当 $\|\Pi\| \to 0$ 时极限为 0. 类似地, 高阶项的极限也为 0. 右端第一项的极限为 $\int_0^T f_t(t,X(t))dt$, 第二项的极限为:

$$\int_0^T f_x(t,X(t))dX(t) = \int_0^T f_x(t,X(t))\Delta(t)dW(t) + \int_0^T f_x(t,X(t))\Theta(t)dt$$

由于伊藤过程 X(t) 单位时间内累积二次变差的速率为 $\Delta^2(u)$, 右端第三项的极限为:

$$\frac{1}{2} \int_0^T f_{xx}(t, X(t)) d[X, X](t) = \frac{1}{2} \int_0^T f_{xx}(t, X(t)) \Delta^2(t) dt$$

笔记采用微分记法,则更便于记忆和使用该定理结果.改写为:

$$df(t,X(t)) = f_t(t,X(t))dt + f_x(t,X(t))dX(t) + \frac{1}{2}f_{xx}(t,X(t))dX(t)dX(t)$$
(5.10)

利用伊藤过程的微分形式 [即 $dX(t) = \Delta(t)dW(t) + \Theta(t)dt$] 以及 X(t) 累积二次变差的速率公式 [即 dX(t)dX(t) =

 $\Delta^2(t)dt$, 我们可以将上式化为仅含 dt 和 dW(t) 的表达式. 利用乘法表(5.8)我们得到:

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))\Delta(t)dW(t)$$
$$+ f_x(t, X(t))\Theta(t)dt + \frac{1}{2}f_{xx}(t, X(t))\Delta^2(t)dt$$

推论 5.2 (Ito 扩散过程)

如果满足随机微分方程

$$dX(t) = \underbrace{a(X(t))dt}_{\text{ \& \$r}_{\overline{\eta}}} + \underbrace{b(X(t))dW(t)}_{\text{ \& r}_{\overline{\eta}}}$$

则称 X(t) 为 Ito 扩散过程, 具有漂移系数 a 和扩散系数 b. 对于 Ito 扩散过程, 有:

$$df(t, X(t)) = f_t(t, X(t))dt + f_x(t, X(t))a(t, X(t))dt + f_x(t, X(t))b(t, X(t))dW(t) + \frac{1}{2}f_{xx}(t, X(t))b^2(t, X(t))dt$$

当我们谈论 Ito 扩散过程时, 其实是在描述随机变量随时间变化的方式. 想象一下你在湖面上放一片叶子, 它会随着水流的变化而漂动. 在数学上, Ito 扩散过程就像是这片叶子的运动轨迹, 其中有两个关键因素: 扩散系数和漂移系数.

- 1. 扩散系数描述了扩散过程中的随机波动或者说不确定性的大小. 在叶子的例子中, 扩散系数就代表了湖水的湍流程度. 如果湖水湍流很大, 叶子就会快速随机地漂动; 反之则漂动较慢.
- 2. 漂移系数描述了随机变量的平均趋势或者说确定性的部分. 在叶子的情景中, 漂移系数可以看作是湖水整体流动的速度和方向. 如果湖水整体朝向某个方向流动, 叶子会带着这个流向有所移动.

X(t) 一定是 t, W(t) 的函数, 利用 Ito 公式(5.9)可以求解简单的随机微分方程¹⁰.

例题 5.4 在金融应用中, 股票的价格 S(t) 是用随机微分方程

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t)$$

进行描述的. 其中 $\mu \in \mathbb{R}$, $\sigma > 0$. 给定初值 $S(0) = s_0 > 0$. 把上述微分方程和(5.9)对比, 令 S(t) = f(t, W(t)), 知:

$$\frac{\partial f(t,x)}{\partial x} = \sigma f(t,x)$$
$$\frac{\partial f(t,x)}{\partial t} + \frac{1}{2} \frac{\partial^2 f(t,x)}{\partial x^2} = \mu f(t,x)$$

观察偏微分方程形式,解应该具有变量可分离性质. 第一式有通解 $S(t)=f(t,x)=Ce^{\sigma W(t)};$ 把第一个式子带入第二个式子,带入两遍得

$$\frac{1}{2}\sigma^2 f(t,x) + \frac{\partial f(t,x)}{\partial t} = \mu f(t,x)$$

即

$$\left(\mu - \frac{1}{2}\sigma^2\right)S(t) = S'(t)$$

解得:

$$S(t) = f(t, W(t)) = \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t\right\}$$

组合两部分的解,得到几何布朗运动:

$$S(t) = s_0 \exp\left\{ \left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W(t) \right\}$$

例题 5.5 定义伊藤过程

$$X(t) = \int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right)ds$$

¹⁰随机微分方程和偏微分方程关系密切,将在后面的章节讨论

则:

$$dX(t) = \sigma(t)dW(t) + \left(\alpha(t) - \frac{1}{2}\sigma^{2}(t)\right)dt$$
$$dX(t)dX(t) = \sigma^{2}(t)dW(t)dW(t) = \sigma^{2}(t)dt$$

考虑由下式给出的资产价格过程(广义几何布朗运动)

$$S(t) = S(0)e^{X(t)} = S(0) \exp\left\{ \int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right)ds \right\}$$

其中 S(0) 是非随机的正数. 上式可以写成 S(t) = f(X(t)), 其中 $f(x) = S(0)e^x$, $f'(x) = S(0)e^x$, $f''(x) = S(0)e^x$. 根据关于伊藤过程的伊藤公式(5.10)得:

$$\begin{split} dS(t) &= df(X(t)) \\ &= f'(X(t))dX(t) + \frac{1}{2}f''(X(t))dX(t)dX(t) \\ &= S(0)e^{X(t)}dX(t) + \frac{1}{2}S(0)e^{X(t)}dX(t)dX(t) \\ &= S(t)dX(t) + \frac{1}{2}S(t)dX(t)dX(t) \\ &= S(t)\left(\sigma(t)dW(t) + \left(\alpha(t) - \frac{1}{2}\sigma^2(t)\right)dt\right) + \frac{1}{2}S(t)\sigma^2(t)dt \\ &= \alpha(t)S(t)dt + \sigma(t)S(t)dW(t) \end{split}$$

资产价格 S(t) 具有瞬时平均回报率 $\alpha(t)$ 和波动率 $\sigma(t)^{11}$. 当它们都是常数时退化为一般几何布朗运动. 如果伊藤积分中被积函数是非随机的, 即不是随机过程而只是时间变量的函数, 我们有如下结论:

定理 5.4 (非随机被积函数的 Ito 积分)

设 $W(s), s \ge 0$ 是布朗运动,g(s) 是关于时间的非随机函数. 则对于每个 $t \ge 0$, 有:

$$I(t) = \int_0^t g(s)dW(s) \sim \mathcal{N}\left(0, \int_0^t g^2(s)ds\right)$$

且 I(t) 是高斯过程.

证明 由于 I(t) 是鞅, 并且 I(0) = 0, 我们必有 $\mathbb{E}I(t) = I(0) = 0$. 由伊藤等距, 我们有:

$$\operatorname{Var} I(t) = \mathbb{E}I^{2}(t) = \int_{0}^{t} g^{2}(s)ds$$

由于 g(s) 非随机,上式右端的 $\int_0^t g^2(s)ds$ 不需要取期望值.

关键是证明 I(t) 服从正态分布. 我们将证明: I(t) 具有均值为 0 、方差为 $\int_0^t g^2(s)ds$ 的正态随机变量的矩母函数:

$$\mathbb{E}e^{u(t)} = \exp\left\{\frac{1}{2}u^2 \int_0^t g^2(s)ds\right\}, \quad \forall u \in \mathbb{R}$$

由于 g(s) 非随机, 等价于:

$$\mathbb{E}\exp\left\{uI(t) - \frac{1}{2}u^2 \int_0^t g^2(s)ds\right\} = 1$$

或者改写成:

$$\mathbb{E}\exp\left\{\int_0^t ug(s)dW(s) - \frac{1}{2}\int_0^t (ug(s))^2 ds\right\} = 1$$

注意到期望内部的过程

$$\exp\left\{\int_0^t ug(s)dW(s) - \frac{1}{2}\int_0^t (ug(s))^2 ds\right\}$$

是鞅. 事实上, 它是平均回报率 $\alpha=0$ 的广义几何布朗运动, 其中 $\sigma(s)=ug(s)$. 而且该过程在 t=0 时取值为 1,

 $^{^{11}}$ 过程 $S(0)\exp\left(\sigma W(t)-rac{1}{2}\sigma^2 t
ight)$ 是鞅, 即整体趋势稳定, 如果在指数中加上 lpha t 则得到趋势 lpha,它衡量了平均回报率

因而期望值恒为1.这就得证.

接下来证明 I(t) 是高斯过程, 即必须说明, 对 $0 < t_1 < t_2 < \cdots < t_n$, 随机变量 $I(t_1), I(t_2), \cdots, I(t_n)$ 服从联合正态分布. 固定 u. 令

$$M_u(t) = e^{uI(t) - \frac{1}{2}u^2 \int_0^t \Delta^2(s)ds}$$

下面证明: 对于 $0 < t_1 < t_2$, 随机增量 $I(t_1) - I(t_0) = I(t_1)$ 和 $I(t_2) - I(t_1)$ 相互独立并且服从正态分布. 累次利用这里提供的证法, 可以证明任意多个增量的情形结果也成立. 对于固定的 $u_2 \in \mathbb{R}$, 由 M_{u_2} 的鞅性质可知:

$$M_{w_2}(t_1) = \mathbb{E}\left[M_{w_2}(t_2) \mid \mathcal{F}(t_1)\right]$$

再固定 $u_1 \in \mathbb{R}$. 由于 $\frac{M_{u_1}(t_1)}{M_{u_2}(t_1)}$ 为 $\mathcal{F}(t_1)$ -可测, 以此同乘上式两边, 可得:

$$M_{u_1}(t_1) = \mathbb{E}\left[\frac{M_{u_1}(t_1) M_{u_2}(t_2)}{M_{u_2}(t_1)} \middle| \mathcal{F}(t_1)\right]$$

$$= \mathbb{E}\left[\exp\left\{u_1(t_1) + u_2(I(t_2) - I(t_1)) - \frac{1}{2}u_1^2 \int_0^{t_1} \Delta^2(s) ds - \frac{1}{2}u_2^2 \int_{t_1}^{t_2} \Delta^2(s) ds\right\} \middle| \mathcal{F}(t_1)\right]$$

两边取期望,得到:

$$1 = M_{u_1}(0) = \mathbb{E}M_{u_1}(t_1) = \mathbb{E}\left[\exp\left\{u_1I\left(t_1\right) + u_2\left(I\left(t_2\right) - I\left(t_1\right)\right) - \frac{1}{2}u_1^2 \int_0^{t_1} \Delta^2(s)ds - \frac{1}{2}u_2^2 \int_{t_1}^{t_2} \Delta^2(s)ds\right\}\right]$$

$$= \mathbb{E}\left[\exp\left\{u_1I\left(t_1\right) + u_2\left(I\left(t_2\right) - I\left(t_1\right)\right)\right\}\right] \cdot \exp\left\{-\frac{1}{2}u_1^2 \int_0^{t_1} \Delta^2(s)ds - \frac{1}{2}u_2^2 \int_{t_1}^{t_2} \Delta^2(s)ds\right\}$$

在上式右端, 由于 $\Delta^2(s)$ 非随机这一事实, 所以把 $\Delta^2(s)$ 的积分提出期望外. 由上, 得到矩母函数公式:

$$\mathbb{E}\left[\exp\left\{u_1I(t_1) + u_2(I(t_2) - I(t_1))\right\}\right] = \exp\left\{\frac{1}{2}u_1^2 \int_0^{t_1} \Delta^2(s)ds\right\} \exp\left\{\frac{1}{2}u_2^2 \int_{t_1}^{t_2} \Delta^2(s)ds\right\}$$

上式右端是均值为 0、方差为 $\int_0^{t_1} \Delta^2(s) ds$ 的正态随机变量矩母函数以及均值为 0、方差为 $\int_{t_1}^{t_2} \Delta^2(s) ds$ 的正态随机变量矩母函数的乘积. $I(t_1)$ 和 $I(t_2) - I(t_1)$ 服从这些分布,并由于联合矩母函数可分解为各自矩母函数的乘积,故 $I(t_1)$ 和 $I(t_2) - I(t_1)$ 相互独立.

5.5.2 相关金融模型

Vasicek Model 瓦西塞克模型 Vasicek 模型是一种用于描述短期利率变化的随机过程模型. 由捷克经济学家 Oldřich Vašíček 于 1977 年提出, 在 Vasicek 模型下, 可以推导出零息债券的价格公式, 从而用于债券定价. Vasicek 模型广泛 用于各种利率衍生品的定价, 包括利率期权、互换等. Vasicek 模型假设短期利率 r_t 的动态服从以下随机微分方程:

$$dr_t = a(b - r_t)dt + \sigma dW(t)$$

其中:

- r_t : 时间 t 时的短期利率.
- a: 速度参数, 表示利率回归到长期均值的速度.
- b: 长期均值, 即利率趋向的平均水平12.
- σ: 利率的波动率,表示利率的随机波动强度.

Vasicek 模型的一个优点是其具有闭式解,这使得许多金融工具的定价可以通过解析方法而不是数值方法来实现. 但缺点也很明显,比如利率在 Vasicek 模型下可以为负值,这在实际情况中是不合理的. 模型假设波动率是恒定的,而实际市场中的波动率可能是随时间变化的.

通过求解 Vasicek 模型的随机微分方程, 可以得到短期利率的显式解:

$$r_t = r_0 e^{-at} + b(1 - e^{-at}) + \sigma \int_0^t e^{-a(t-s)} dW(s)$$

 $^{^{12}}$ Vasicek 模型的一个关键特性是均值回复, 即利率有回归到长期均值 b 的趋势. 参数 a 决定了这个回复的速度. 如果当前利率高于长期均值, 那么 d t 公趋向负值, 推动利率下降, 反之亦然

其中, r_0 是初始时刻的短期利率.

使用积分因子法. 首先, 确定积分因子 I(t):

$$I(t) = e^{at}$$

将其乘以原方程的两边:

$$e^{at}dr_t = e^{at}a(b - r_t)dt + e^{at}\sigma dW(t)$$

这可以写成:

$$e^{at}dr_t + ae^{at}r_tdt = abe^{at}dt + \sigma e^{at}dW(t)$$

注意到左侧可以写成一个全微分:

$$d\left(e^{at}r_t\right) = abe^{at}dt + \sigma e^{at}dW(t)$$

对整个方程在0到t的区间上积分:

$$e^{at}r_t - r_0 = ab \int_0^t e^{as} ds + \sigma \int_0^t e^{as} dW(t)$$

其中 r_0 是初始条件, 即 $r_0 = r_0$. 求解黎曼积分并整理, 得到:

$$r_t = r_0 e^{-at} + b \left(1 - e^{-at} \right) + \sigma e^{-at} \int_0^t e^{as} dW(t)$$

由定理**5.4**知, r_t 是均值为 $e^{-at}r_0 + b(1 - e^{-at})$, 方差为 $\frac{\sigma^2}{2a}(1 - e^{-2at})$ 的正态变量. 因此总有概率为负, 这就是之前提到的局限性.

关于均值回复性, 作如下解释: 当 $r_t = b$ 时, 漂移项 (dt 项) 为 0. 当 $r_t > b$ 时, 这一项是负的, 迫使 r_t 回向 b. 当 $r_t < b$ 时, 这一项是正的, 仍迫使 r_t 回向 b. 如果 $r_0 = b$, 则对所有 t > 0, $\mathbb{E}r_t = b$. 如果 $r_0 \neq b$, 则 $\lim \mathbb{E}r_t = b$.

Cox-Ingersoll-Ross Model (CIR 模型) Cox-Ingersoll-Ross 模型是用于解决 Vasicek 模型出现负利率问题的单因素模型. CIR 模型基于如下微分方程:

$$dr_t = a(b - r_t) dt + \sigma \sqrt{r_t} dW_t$$

与瓦西塞克模型不同, CIR 方程不再具有闭形式的解; CIR 模型中, 利率不会取负值: 如果 r_t 达到 0, 则 dW(t) 前面的因子为 0, 从而方程中正的漂移项 abdt 驱动利率回到取正值的范围; 另外,CIR 模型也具有均值回复性质.

虽然其不再具有显式解, 但可以窥探其分布性质. 看一看其期望和方差: 为此, 令 $f(t,x)=e^{\alpha t}x$, 由伊藤公式 计算

$$d\left(e^{at}r_t\right) = df(t, r_t)$$

$$= f_t(t, r_t)dt + f_x(t, r_t)dr_t + \frac{1}{2}f_{xx}(t, r_t)dr_t dr_t$$

$$= ae^{at}r_t dt + e^{at}(ab - ar_t)dt + e^{at}\sigma\sqrt{r_t}dW(t)$$

$$= abe^{at}dt + \sigma e^{at}\sqrt{r_t}dW(t)$$

两边同时积分,可得:

$$e^{at}r_t = r_0 + ab \int_0^t e^{au} du + \sigma \int_0^t e^{au} \sqrt{r_u} dW(u)$$
$$= r_0 + b \left(e^{at} - 1\right) + \sigma \int_0^t e^{au} \sqrt{r_u} dW(u)$$

由于伊藤积分的期望值为0,我们有:

$$e^{at}\mathbb{E}r_t = r_0 + b\left(e^{at} - 1\right)$$

即:

$$\mathbb{E}(r_t) = b + (r_0 - b)e^{-at}$$

为了计算 r_t 的方差, 令 $X(t) = e^{at}r_t$, 我们已经算得

$$dX(t) = abe^{at}dt + \sigma e^{at}\sqrt{r_t}dW(t) = abe^{at}dt + \sigma e^{\frac{at}{2}}\sqrt{X(t)}dW(t)$$

以及 $\mathbb{E}X(t) = r_0 + b(e^{at} - 1)$. 根据 Ito 过程的 Ito 公式, 其中 $f(x) = x^2$, 有

$$d(X^{2}(t)) = 2X(t)dX(t) + dX(t)dX(t)$$
$$= 2abe^{at}X(t)dt + 2\sigma e^{\frac{at}{2}}X^{\frac{a}{2}}(t)dW(t) + \sigma^{2}e^{at}X(t)dt$$

两边同时积分,可得:

$$X^{2}(t) = X^{2}(0) + \left(2ab + \sigma^{2}\right) \int_{0}^{t} e^{au} X(u) du + 2\sigma \int_{0}^{t} e^{\frac{au}{2}} X^{\frac{a}{2}}(u) dW(u)$$

两边取期望,利用伊藤积分的期望值为0这一事实以及关于 X(t) 已经推得的公式,我们得到:

$$\mathbb{E}(X^{2}(t)) = r_{0}^{2} + (2ab + \sigma^{2}) \int_{0}^{t} e^{au} (r_{0} + b (e^{au} - 1)) du$$

$$= r_{0}^{2} + \frac{2ab + \sigma^{2}}{a} (r_{0} - b) (e^{at} - 1) + \frac{2ab + \sigma^{2}}{2a} \cdot b (e^{2at} - 1)$$

因此:

$$\mathbb{E}r_t^2 = e^{-2at}\mathbb{E}X^2(t)$$

$$= e^{-2at}r_0^2 + \frac{2ab + \sigma^2}{a}(r_0 - b)\left(e^{-at} - e^{-2at}\right) + \frac{b\left(2ab + \sigma^2\right)}{2a}\left(1 - e^{-2at}\right)$$

最后:

$$Var(r_t) = \mathbb{E}r_t^2 - (\mathbb{E}(r_t))^2$$

$$= e^{-2at}r_0^2 + \frac{2ab + \sigma^2}{a} (r_0 - b) (e^{-at} - e^{-2at})$$

$$+ \frac{b(2ab + \sigma^2)}{2a} (1 - e^{-2at}) - e^{-2at}r_0^2$$

$$- 2br_0 (e^{-at} - e^{-2at}) - \frac{b^2}{a} (1 - e^{-at})^2$$

$$= \frac{\sigma^2}{a}r_0 (e^{-at} - e^{-2at}) + \frac{b\sigma^2}{2a} (1 - 2e^{-at} + e^{-2at})$$

特别地:

$$\lim_{t \to \infty} \operatorname{Var}(r_t) = \frac{b\sigma^2}{2a}$$

5.6 测度变换

在第一章的附录中曾简单介绍过测度变换. 在考虑随机过程问题时, 我们经常遇到一些随机过程不是鞅的情况, 这时我们需要对其进行测度变换, 使其成为鞅. 例如, 几何布朗运动 $S(t) = S(0) \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right\}$ 的漂移项 μ 是常数, 而不是随机变量, 因此 S(t) 不是鞅. 我们的问题是, 能否找到一个新的测度 $\mathbb Q$, 使得在概论空间 $(\Omega, \mathcal F, \mathbb Q)$ 下, 几何布朗运动 S(t) 是鞅?

定义 5.5 (测度变换)

 $\Gamma = (\Gamma_t)_{0 \le t \le T}$ 是定义在带流概率空间 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P})$ 上的一个鞅, 且严格大于 0. 当 $\mathbb{E}(\Gamma_T) = 1$ 时, 定义一个新的概率测度 \mathbb{Q} , 使得对于任意 $A \in \mathcal{F}_t$, 有

$$\mathbb{Q}(A) = \mathbb{E}\left(\Gamma_T \mathbb{I}_A\right) = \int_A \Gamma_T d\mathbb{P}$$

则称 ℚ 是 ℙ的一个测度变换.

•

 \Diamond

命题 5.4

设 € 是在 ② 意义下可积的随机变量,则

$$\mathbb{E}_{\mathbb{Q}}(\xi) = \mathbb{E}_{\mathbb{P}}(\Gamma_T \xi)$$

以及当s < t时,

$$\mathbb{E}_{\mathbb{Q}}(\xi \mid \mathcal{F}_s) = \mathbb{E}_{\mathbb{P}}\left(\frac{\Gamma_t}{\Gamma_s} \xi \mid \mathcal{F}_s\right)$$

定理 5.5

 $M=(M_t)_{0\leq t\leq T}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\},\mathbb{Q})$ 上的鞅, 当且仅当 $(M_t\Gamma_t)_{0\leq t\leq T}$ 是 $(\Omega,\mathcal{F},\{\mathcal{F}_t\},\mathbb{P})$ 上的鞅.

证明

1. 必要性: 由于 $M \in \mathbb{Q}$ -鞅,则对于任意 $0 \le s \le t \le T$,有

$$\mathbb{E}_{\mathbb{P}}(\Gamma_t M_t | \mathcal{F}_s) = \Gamma_s \mathbb{E}_{\mathbb{P}}\left(\frac{\Gamma_t}{\Gamma_s} M_t | \mathcal{F}_s\right) = \Gamma_s \mathbb{E}_{\mathbb{Q}}(M_t | \mathcal{F}_s) = \Gamma_s M_s$$

即 $(M_t\Gamma_t)_{0\leq t\leq T}$ 是 \mathbb{P} -鞅.

2. 充分性: 由于 $(M_t\Gamma_t)_{0 \le t \le T}$ 是 \mathbb{P} -鞅, 则对于任意 $0 \le s \le t \le T$, 有

$$\mathbb{E}_{\mathbb{Q}}(M_t|\mathcal{F}_s) = \mathbb{E}_{\mathbb{P}}\left(\frac{\Gamma_t}{\Gamma_s}M_t \mid \mathcal{F}_s\right) = \frac{1}{\Gamma_s}\mathbb{E}_{\mathbb{P}}(\Gamma_t M_t|\mathcal{F}_s) = \frac{\Gamma_s M_s}{\Gamma_s} = M_s$$

即 M 是 \mathbb{Q} -鞅.

定理 5.6 (布朗运动的测度变换定理)

设 $B \neq (\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ 上的标准布朗运动, 定义 $W_t = B_t + \frac{\mu t}{\sigma}$ 以及

$$\begin{split} \Gamma_t &= \exp\left\{-\frac{\mu B_t}{\sigma} - \frac{\mu^2 t}{2\sigma^2}\right\} \\ \mathbb{Q}(A) &= \int_A \Gamma_T d\mathbb{P} \end{split}$$

则 \mathbb{Q} 是 \mathbb{P} 的一个测度变换, 使得 W_t 是 \mathbb{Q} -标准布朗运动.

证明 首先证明 $W \neq \mathbb{Q}$ -鞅. 由上述定理, 只需要证明 $W_t\Gamma_t \neq \mathbb{P}$ -鞅. 考虑 $0 \leq s < t \leq T$, 有:

$$\mathbb{E}_{\mathbb{P}}(W_{t}\Gamma_{t}|\mathcal{F}_{s}) = \mathbb{E}_{\mathbb{P}}\left(\left(B_{t} + \frac{\mu t}{\sigma}\right) \exp\left\{-\frac{\mu B_{t}}{\sigma} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mid \mathcal{F}_{s}\right)$$

$$= \mathbb{E}_{\mathbb{P}}\left(\left(B_{t} - B_{s} + B_{s} + \frac{\mu t}{\sigma}\right) \exp\left\{-\frac{\mu B_{t}}{\sigma} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mid \mathcal{F}_{s}\right)$$

$$= \mathbb{E}_{\mathbb{P}}\left(\left(B_{t} - B_{s}\right) \exp\left\{-\frac{\mu B_{t}}{\sigma} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mid \mathcal{F}_{s}\right) + \left(B_{s} + \frac{\mu t}{\sigma}\right) \exp\left\{-\frac{\mu B_{s}}{\sigma} - \frac{\mu^{2} s}{2\sigma^{2}}\right\}$$

其中:

$$\mathbb{E}_{\mathbb{P}}\left((B_{t} - B_{s}) \exp\left\{-\frac{\mu B_{t}}{\sigma} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mid \mathcal{F}_{s}\right) = \mathbb{E}_{\mathbb{P}}\left((B_{t} - B_{s}) \exp\left\{-\frac{\mu (B_{t} - B_{s})}{\sigma} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \exp\left\{-\frac{\mu}{\sigma} B_{s}\right\} \mid F_{s}\right)$$

$$= \exp\left\{-\frac{\mu}{\sigma} B_{s}\right\} \mathbb{E}_{\mathbb{P}}\left((B_{t} - B_{s}) \exp\left\{-\frac{\mu (B_{t} - B_{s})}{\sigma} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mid F_{s}\right)$$

$$= \exp\left\{-\frac{\mu}{\sigma} B_{s} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mathbb{E}_{\mathbb{P}}\left((B_{t} - B_{s}) \exp\left\{-\frac{\mu (B_{t} - B_{s})}{\sigma}\right\} \mid F_{s}\right)$$

$$= \exp\left\{-\frac{\mu}{\sigma} B_{s} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \mathbb{E}_{\mathbb{P}}\left((B_{t} - B_{s}) \exp\left\{-\frac{\mu (B_{t} - B_{s})}{\sigma}\right\}\right)$$

$$= \exp\left\{-\frac{\mu}{\sigma} B_{s} - \frac{\mu^{2} t}{2\sigma^{2}}\right\} \exp\left\{-\frac{\mu^{2} (s - t)}{2\sigma^{2}}\right\} \left[-\frac{\mu}{\sigma} (t - s)\right]$$

$$= \exp\left\{-\frac{\mu}{\sigma}B_s - \frac{\mu^2 s}{2\sigma^2}\right\} \left[-\frac{\mu}{\sigma}(t-s)\right]$$

因此:

$$\mathbb{E}_{\mathbb{P}}(W_t \Gamma_t | \mathcal{F}_s) = \exp\left\{-\frac{\mu}{\sigma} B_s - \frac{\mu^2 s}{2\sigma^2}\right\} \left[-\frac{\mu}{\sigma} (t-s)\right] + \left(B_s + \frac{\mu t}{\sigma}\right) \exp\left\{-\frac{\mu B_s}{\sigma} - \frac{\mu^2 s}{2\sigma^2}\right\}$$
$$= \left(B_s + \frac{\mu s}{\sigma}\right) \exp\left\{-\frac{\mu B_s}{\sigma} - \frac{\mu^2 s}{2\sigma^2}\right\} = W_s \Gamma_s$$

因此鞅性得到了证明, 注意到 W 的二次变差 $[W,W](t) = [B + \frac{\mu}{\sigma}t, B + \frac{\mu}{\sigma}t](t) = t$, 因而由定理**4.6**知 W 是布朗运动.

5.7 Black-Scholes-Merton 方程

5.7.1 Black-Scholes-Merton 模型的设定

BSM 方程的思想与二叉树差不多, 都是旨在确定为了完全对冲期权空头所需要的初始资本. 设每个时刻 t, 投资者资产组合的价值为 X(t), 这一组合投资于支付常数利率 r 的货币市场账户以及股票市场, 股票价格过程服从几何布朗运动

$$dS(t) = \alpha S(t)dt + \sigma S(t)dW(t)$$

假设在时刻 t 投资者持有股票的份额为 $\Delta(t)$, 头寸 $\Delta(t)$ 可以随机, 但是必须与相应于布朗运动 W(t), $t \ge 0$ 的域流相适应. 资产组合中余下部分 $X(t) - \Delta(t)S(t)$ 投资于货币市场账户.

投资者在时刻 t 资产组合价值的微分 dX(t) 由两个因素引起: 关于股票头寸的资本增值 $\Delta(t)dS(t)$ 以及关于现金头寸的利息收入 $r(X(t)-\Delta(t)S(t))dt$. 换言之:

$$\begin{split} dX(t) &= \Delta(t)dS(t) + r[X(t) - \Delta(t)S(t)]dt \\ &= \Delta(t)(\alpha S(t)dt + \sigma S(t)dW(t)) + r(X(t) - \Delta(t)S(t))dt \\ &= rX(t)dt + \Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) \end{split}$$

等 筆记 上式最后一行可以如下理解:

- 1. 资产组合平均的保底回报率r, 由rX(t)dt体现.
- 2. 股票的平均回报率为 α , 而投资者的回报率为 r, 则 $\Delta(t)(\alpha-r)S(t)dt$ 体现了投资者的风险溢价.
- 3. 股票的波动率为 σ , 则 $\Delta(t)\sigma S(t)dW(t)$ 体现了投资者的风险.

离散时间下, 我们有:

$$X_{n+1} - X_n = \Delta_n (S_{n+1} - S_n) + r(X_n - \Delta_n S_n)$$

它类似于连续时间模型的第一行.

考虑贴现股价 $V(t) = e^{-rt}S(t)$, 以及贴现资产组合价值 $Y(t) = e^{-rt}X(t)$, 根据 Ito 公式, 贴现股价的微分是:

$$\begin{split} dV(t) &= e^{-rt}dS(t) - re^{-rt}S(t)dt \\ &= e^{-rt}(\alpha S(t)dt + \sigma S(t)dW(t)) - re^{-rt}S(t)dt \\ &= e^{-rt}(\alpha - r)S(t)dt + e^{-rt}\sigma S(t)dW(t) \end{split}$$

贴现资产组合价值的微分是:

$$\begin{split} dY(t) &= e^{-rt} dX(t) - re^{-rt} X(t) dt \\ &= e^{-rt} r X(t) dt + e^{-rt} \Delta(t) (\alpha - r) S(t) dt + e^{-rt} \Delta(t) \sigma S(t) dW(t) - re^{-rt} X(t) dt \\ &= e^{-rt} \Delta(t) \sigma S(t) dW(t) + e^{-rt} \Delta(t) (\alpha - r) S(t) dt \\ &= \Delta(t) d(V(t)) \end{split}$$

掌记上式最后一行可以如下理解:

- 1. 贴现股价的微分 $e^{-rt}(\alpha-r)S(t)dt+e^{-rt}\sigma S(t)dW(t)$ 体现了股票的风险: 股价的贴现使得平均回报率减少r.
- 2. 贴现资产组合价值的微分 $e^{-rt}\Delta(t)(\alpha-r)S(t)dt+e^{-rt}\Delta(t)\sigma S(t)dW(t)$ 体现了投资者的风险: 因为失去了保底回报项 rX(t)dt, 资产组合价值的改变完全由股票的风险和风险溢价决定.

5.7.2 Delta 对冲法则

期权价值的演化 考虑一个欧式看涨期权 13 , 其价格为C(t, S(t)), 则其价值的微分为:

$$\begin{split} dC(t,S(t)) &= \frac{\partial C}{\partial t} dt + \frac{\partial C}{\partial S} dS(t) + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} dS(t) dS(t) \\ &= \frac{\partial C}{\partial t} dt + \frac{\partial C}{\partial S} dS(t) + \frac{1}{2} \frac{\partial^2 C}{\partial S^2} \sigma^2 S^2(t) dt \\ &= \left(\frac{\partial C}{\partial t} + \frac{1}{2} \sigma^2 S^2(t) \frac{\partial^2 C}{\partial S^2} \right) dt + \frac{\partial C}{\partial S} dS(t) \\ &= \left(\frac{\partial C}{\partial t} + \alpha \frac{\partial C}{\partial S} S(t) + \frac{1}{2} \sigma^2 S^2(t) \frac{\partial^2 C}{\partial S^2} \right) dt + \frac{\partial C}{\partial S} \sigma S(t) dW(t) \end{split}$$

接下来计算贴现期权价值 $e^{-rt}C(t,S(t))$, 根据 Ito 公式, 令 $f(t,x) = e^{-rt}x$, 则贴现期权价值的微分为:

$$\begin{split} d\left(e^{-rt}C(t,S(t))\right) &= e^{-rt}dC(t,S(t)) - re^{-rt}C(t,S(t))dt \\ &= e^{-rt}\left(\frac{\partial C}{\partial t} + \alpha\frac{\partial C}{\partial S}S(t) + \frac{1}{2}\sigma^2S^2(t)\frac{\partial^2 C}{\partial S^2}\right)dt + e^{-rt}\frac{\partial C}{\partial S}\sigma S(t)dW(t) - re^{-rt}C(t,S(t))dt \\ &= e^{-rt}\left(\frac{\partial C}{\partial t} + \alpha\frac{\partial C}{\partial S}S(t) + \frac{1}{2}\sigma^2S^2(t)\frac{\partial^2 C}{\partial S^2} - rC(t,S(t))\right)dt + e^{-rt}\frac{\partial C}{\partial S}\sigma S(t)dW(t) \end{split}$$

Delta 对冲法则 一开始用什么样的投资组合,才能**复制**期权价值的演化,即对冲期权空头呢?

定义 5.6 (对冲组合)

若投资者以初始资本 X(0) 投资于股票和货币市场, 使得在每个时刻 t, 资产组合价值 X(t) 满足:

$$X(t) = C(t, S(t))$$

其中 C(t, S(t)) 是期权的价格, 则称 X(t) 是期权的对冲组合.

命题 5.5

找到对冲组合的等价方式是确定

$$d(e^{-rt}X(t)) = d(e^{-rt}C(t,S(t))), \quad \forall t \in [0,T)$$

$$\mathrm{s.t.}X(0) = C(0,S(0))$$

比较期权的贴现价值微分与贴现资产组合价值微分, 我们得:

$$\begin{split} &\Delta(t)(\alpha - r)S(t)dt + \Delta(t)\sigma S(t)dW(t) \\ &= \left(\frac{\partial C}{\partial t} + \alpha \frac{\partial C}{\partial S}S(t) + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 C}{\partial S^2} - rC(t,S(t))\right)dt \\ &+ \sigma S(t)\frac{\partial C}{\partial S}dW(t) \end{split}$$

由于 dW(t) 的系数相等, 我们有:

$$\Delta(t) = \frac{\partial C}{\partial S}, \quad \forall t \in [0, T)$$

这就是 Delta 对冲法则.



笙记

 $^{^{13}}$ 在时刻 $^{\prime}$ 7 的支付是 $^{\prime}$ 8 $^{\prime}$ 6 $^{\prime}$ 7,敲定价格 $^{\prime}$ 7 是个非负常数. Black 等人论证了看涨期权在任何时刻的价值都依赖于时间和该时刻的股价

- 1. Delta 对冲法则告诉我们, 为了对冲期权空头, 投资者应该持有的股票份额 $\Delta(t)$ 等于期权价格 C(t,S(t)) 对股票价格 S(t) 的偏导数.
- 2. Delta 对冲法则的直观解释是: 投资者持有的股票份额 $\Delta(t)$ 应该与股票价格 S(t) 的变化成比例, 比例系数 就是期权价格 C(t,S(t)) 对股票价格 S(t) 的偏导数.

利用 Delta 对冲法则, 再令 dt 项相等, 我们有:

$$\frac{\partial C}{\partial S}(\alpha - r)S(t) = \frac{\partial C}{\partial t} + \alpha \frac{\partial C}{\partial S}S(t) + \frac{1}{2}\sigma^2 S^2(t) \frac{\partial^2 C}{\partial S^2} - rC(t, S(t))$$

即

$$\frac{\partial C}{\partial t} + rS(t)\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2S^2(t)\frac{\partial^2C}{\partial S^2} - rC(t,S(t)) = 0$$

称上式为 **Black-Scholes-Merton** 方程. 我们的目标是寻求期权价格函数 C(t, S(t)) 满足 Black-Scholes-Merton 方程, 以及终止条件 $C(T, S(T)) = (S(T) - K)^+$.

5.7.3 Black-Scholes-Merton 方程的解

定理 5.7 (Black-Scholes-Merton 公式 (BSM 方程的解))

Black-Scholes-Merton 方程

$$\frac{\partial C}{\partial t} + rS(t)\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2(t)\frac{\partial^2 C}{\partial S^2} - rC(t, S(t)) = 0$$

的解为

$$C(t, S(t)) = S(t)N(d_1) - Ke^{-r(T-t)}N(d_2)$$

其中

$$d_1 = \frac{1}{\sigma\sqrt{T-t}} \left[\ln\left(\frac{S(t)}{K}\right) + \left(r + \frac{\sigma^2}{2}\right) (T-t) \right]$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

N(x) 是标准正态分布函数

$$N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{u^2}{2}} du$$

我们接下来说明这个解是怎么来的14

证明 令 $Z=\exp\left\{-\frac{-\nu^2T}{2\sigma^2}+\frac{\nu W(T)}{\sigma}\right\}$, 其中 $\nu=r-\alpha$. 定义一个新的概率测度

$$\tilde{\mathbb{P}}(A) = \mathbb{E}[Z\mathbb{I}_A]$$

令

$$\tilde{W}(t) = W(t) - \frac{\nu t}{\sigma}$$

则 $\tilde{W}(t)$ 是 $\tilde{\mathbb{P}}$ 下的标准布朗运动. 我们有:

$$\begin{split} d(e^{-rt}X(t)) &= e^{-rt}dX(t) - rX(t)dt \\ &= e^{-rt}\Delta(t)dS(t) + e^{-rt}r(X(t) - \Delta(t)S(t))dt - rX(t)dt \\ &= e^{-rt}\Delta(t)(\alpha S(t)dt + \sigma S(t)dW(t)) + e^{-rt}r(X(t) - \Delta(t)S(t))dt - rX(t)dt \\ &= e^{-rt}\Delta(t)(\alpha - r)S(t)dt + e^{-rt}\Delta(t)\sigma S(t)dW(t) \\ &= e^{-rt}\Delta(t)\sigma S(t)d\tilde{W}(t) \end{split}$$

 $^{^{14}}$ 事实上, 还应满足如下边界条件: C(t,0)=0, $C(t,\infty)=S(t)-Ke^{-r(T-t)}$, 后者表明期权价格和股票价格, 在股价很大时增长速率相同; 该说明只是以测度变换的方式, 通过对冲法则给出的定价公式, 要想说明定价公式符合偏微分方程的解, 还需费曼卡茨定理的支持, 见第六章.

即:

$$e^{-rt}X(t) = X(0) + \int_0^t e^{-rs}\Delta(s)\sigma S(s)d\tilde{W}(s)$$

这样, $(e^{-rt}X(t), 0 \le t \le T)$ 在概率测度 $\tilde{\mathbb{P}}$ 下关于 $(\mathcal{F}, 0 \le t \le T)$ 是一个鞅. 于是

$$\mathbb{E}^{\tilde{\mathbb{P}}}[e^{-rT}X(T)|\mathcal{F}_t] = e^{-rt}X(t)$$
$$= X(0) + \int_0^t e^{-rs}\Delta(s)\sigma S(s)d\tilde{W}(s)$$

把终止条件 $X(T) = C(T, S(T)) = (S(T) - K)^+$ 代入上式, 我们有:

$$e^{r(T-t)}X(t) = \mathbb{E}^{\tilde{\mathbb{P}}}\left[\left(\mathrm{S}(T) - K\right)^{+} | \mathcal{F}_{t}\right] = \mathbb{E}^{\tilde{\mathbb{P}}}\left[\left(\exp\left\{\left(\alpha - \frac{\sigma^{2}}{2}\right)T + \sigma W(T)\right\} - K\right)^{+} | \mathcal{F}_{t}\right]$$

$$= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\left(\exp\left\{\left(r - \frac{\sigma^{2}}{2}\right)T + \sigma \tilde{W}(T)\right\} - K\right)^{+} | \mathcal{F}_{t}\right]$$

$$= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\left(\exp\left\{\left(r - \frac{\sigma^{2}}{2}\right)T + \sigma \tilde{W}(t) + \sigma(\tilde{W}(T) - \tilde{W}(t))\right\} - K\right)^{+} | \mathcal{F}_{t}\right]$$

$$= \mathbb{E}^{\tilde{\mathbb{P}}}\left[\left(\exp\left\{\left(r - \frac{\sigma^{2}}{2}\right)T + \sigma \tilde{W}(t)\right\} \exp\left\{\sigma(\tilde{W}(T) - \tilde{W}(t))\right\} - K\right)^{+} | \mathcal{F}_{t}\right]$$

由于 $\exp\left\{\sigma(\tilde{W}(T)-\tilde{W}(t))\right\}$ 与 \mathcal{F}_t 是独立的, 令 $x=(\tilde{W}(T)-\tilde{W}(t))/\left(\sqrt{T-t}\right)$, 则 x 是一个标准正态分布; 记 $\phi(\cdot)$ 是标准正态密度函数. 因此由随机变量函数的期望公式, 得:

$$\begin{split} e^{r(T-t)}X(t) &= \int_{\mathbb{R}} \left(\exp\left\{ \left(r - \frac{\sigma^2}{2}\right)T + \sigma \tilde{W}(t) \right\} \exp\left\{ \sigma \sqrt{T - t}x \right\} - K \right)^+ \phi\left(x\right) dx \\ &= \exp\left\{ \left(r - \frac{\sigma^2}{2}\right)T + \sigma \tilde{W}(t) \right\} \int_{e^{(r - \sigma^2/2)T + \sigma \tilde{W}(t) + \sigma \sqrt{T - t}x} > K} e^{\sigma \sqrt{T - t}x} \phi\left(x\right) dx \\ &- K \int_{e^{(r - \sigma^2/2)T + \sigma \tilde{W}(t) + \sigma \sqrt{T - t}x} > K} \phi\left(x\right) dx \end{split}$$

注意到 $S(t) = \exp\left\{\left(r - \frac{\sigma^2}{2}\right)t + \sigma \tilde{W}(t)\right\}$, 因此

$$\exp\left\{\left(r-\frac{\sigma^2}{2}\right)T+\sigma \tilde{W}(t)\right\}=S(t)\exp\left\{\left(r-\frac{\sigma^2}{2}\right)(T-t)\right\}$$

处理积分条件:

$$e^{(r-\sigma^2/2)T+\sigma \tilde{W}(t)+\sigma \sqrt{T-t}x} > K \longleftrightarrow$$

$$x > \frac{1}{\sigma\sqrt{T-t}} \left\{ \log \left(\frac{K}{S(t)} \right) - \left(r - \frac{\sigma^2}{2} \right) (T-t) \right\}, \quad \exists \exists x > m_1$$

则:

第一项 =
$$S(t) \exp\left\{\left(r - \frac{\sigma^2}{2}\right)(T - t)\right\} \int_{m_1}^{\infty} e^{\sigma\sqrt{T - t}x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

= $S(t)e^{r(T - t)} \int_{m_1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x - \sigma\sqrt{T - t})^2} dx$
= $S(t)e^{r(T - t)} \int_{m_1 - \sigma\sqrt{T - t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$
= $S(t)e^{r(T - t)} (1 - N(m_1 - \sigma\sqrt{T - t}))$
= $S(t)e^{r(T - t)} N(-m_1 + \sigma\sqrt{T - t})$

而

$$-m_1 + \sigma\sqrt{T - t} = -\frac{1}{\sigma\sqrt{T - t}} \left\{ \log\left(\frac{K}{S(t)}\right) - \left(r - \frac{\sigma^2}{2}\right)(T - t) \right\} + \sigma\sqrt{T - t}$$

$$= \frac{1}{\sigma\sqrt{T - t}} \left\{ \left(r - \frac{\sigma^2}{2}\right)(T - t) - \log\left(\frac{K}{S(t)}\right) \right\}$$

$$= \frac{1}{\sigma\sqrt{T - t}} \left\{ \log\left(\frac{S(t)}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T - t) \right\}$$

$$= d_1$$

因此第一项就是 $S(t)e^{r(T-t)}N(d_1)$. 又:

第二项 =
$$-K \int_{m_1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

= $-K(1 - N(m_1))$
= $-KN(-m_1)$
= $-KN(d_2)$

因此:

$$e^{r(T-t)}X(t) = S(t)e^{r(T-t)}N(d_1) - KN(d_2)$$

两边约掉 $e^{r(T-t)}$, 我们就得到了 BSM 公式:

$$X(t) = C(t, S(t)) = S(t)N(d_1) - Ke^{-r(T-t)}N(d_2)$$

把满足上面等式的期权定价函数 C(t, S(t)) 称为 BSM 函数, 有如下定义:

定义 5.7 (BSM 函数)

记

$$BSM(\tau, x, K, r, \sigma) = xN(d_{+}) - Ke^{-r\tau}N(d_{-})$$

为 BSM 函数, 其中

$$d_{\pm} = \frac{1}{\sigma\sqrt{\tau}} \left[\ln\left(\frac{x}{K}\right) + \left(r \pm \frac{\sigma^2}{2}\right) \tau \right]$$

 τ 是离期权到期时间尚余的时间 T-t,x 是股票价格,K 是敲定价格,r 是无风险利率, σ 是波动率.



<mark>笔记 BSM 公式告诉我们, 欧式看涨期权的价格等于股票价格 S(t) 与一个折现敲定价格 K 的差额的期望值.</mark>

5.7.4 希腊字母

定义 5.8 (希腊字母)

- 1. Delta: $\Delta(t) = \frac{\partial C}{\partial S}$, 表示期权价格对股票价格的敏感程度.
- 2. Gamma: $\Gamma(t) = \frac{\partial^2 C}{\partial S^2}$, 表示 Delta 对股票价格的敏感程度.
 3. Theta: $\Theta(t) = \frac{\partial C}{\partial t}$, 表示期权价格对时间的敏感程度.
 4. Vega: $\nu(t) = \frac{\partial C}{\partial \sigma}$, 表示期权价格对波动率的敏感程度.

- 5. Rho: $\rho(t) = \frac{\partial C}{\partial r}$, 表示期权价格对无风险利率的敏感程度.

命题 5.6

BSM 函数的希腊字母为:

$$\begin{split} &\Delta(t) = N(d_+) \\ &\Gamma(t) = \frac{1}{\sigma S(t) \sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}d_+^2} \\ &\Theta(t) = -\frac{S(t)\sigma}{2\sqrt{T-t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}d_+^2} - rKe^{-r(T-t)} N(d_-) \\ &\nu(t) = S(t) \sqrt{T-t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}d_+^2} \\ &\rho(t) = K(T-t) e^{-r(T-t)} N(d_-) \end{split}$$

定押58

看涨期权的空头对冲需要借贷资金,看涨期权的多头对冲应持有股票空头,并且投资于货币市场账户.

证明 若时刻 t 的股价为 x, 则空头对冲法则要求我们持有 $\Delta(t) = N(d_+)$ 份股票, 价值为 $xN(d_+)$, 因此需要借贷 $Ke^{-r(T-t)}N(d_-)$ 的资金.

定义 5.9 (多头 Gamma)

能从 C(t,S(t)) 的凸性中获益的资产组合被称为多头 Gamma.

以价格 $C(t,x_1)$ 买入期权, 卖空 $C_x(t,x_1)$ 份股票, 收入为 $x_1C_x(t,x_1)$, 并将差额

$$M = x_1 C_x \left(t, x_1 \right) - C \left(t, x_1 \right)$$

投资于货币市场账户. 考虑资产组合关于股价变动的敏感性, 我们的资产组合含三种成分: 期权多头、股票空头和货币市场账户多头. 在我们建立头寸的时刻 t, 初始资产组合价值 $C(t,x_1)-x_1C_x(t,x_1)+M$ 为 0. 如果股价下

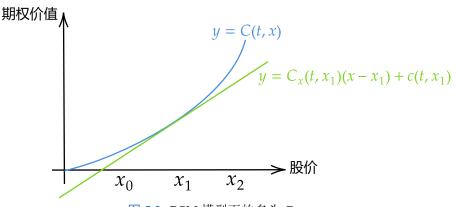


图 5.2: BSM 模型下的多头 Gamma

跌到 x_0 ,而我们的股票和货币市场账户头寸保持不变,则资产组合价值为:

$$C(t,x_0) - x_0C_x(t,x_0) + M = C(t,x_0) - x_0C_x(t,x_0) + x_1C_x(t,x_1) - C(t,x_1)$$

这就是图中 x_0 处曲线 y = C(t,x) 与直线 $y = C_x(t,x_1)(x-x_1) + C(t,x_1)$ 的纵坐标之差. 它是一个正值, 我们的资产组合在股价瞬间下跌时获益.

同样地,如果股价上涨到 x2,则资产组合价值为:

$$C(t, x_2) - x_2C_x(t, x_2) + M = C(t, x_2) - x_2C_x(t, x_2) + x_1C_x(t, x_1) - C(t, x_1)$$

这就是图中 x_2 处曲线 y = C(t,x) 与直线 $y = C_x(t,x_1)(x-x_1) + C(t,x_1)$ 的纵坐标之差. 它是一个正值, 我们的资产组合在股价瞬间上涨时获益. 因此我们说我们因为凸性 (即 Gamma) 而获益.

图中直线与曲线相切, 我们称这样的资产组合为 **Delta-中性的**. 对于股价的微小变动, 该直线是一个很好的近似.

Ŷ 笔记

- 1. 如果在起始点 x_1 处的直线比曲线更陡峭,则对应于空头 **Delta**, 股价的上涨使资产组合价值亏损,因为股票空头引起的负债增长快于期权多头的资产增长.
- 2. Theta 是负的, 因此随着时间增加, 曲线 y = C(t,x) 将往下移. 一般而言, 曲线往下移动要快于货币市场投资和 Gamma 头寸收益的增加, 所以资产组合有可能亏损.
- 3. 股票的平均回报率 α 不在 BSM 方程中, 因此 Delta-中性的资产组合同时对冲了涨跌两种可能, 股价的涨跌 无关紧要. 而股价的波动率才是最重要的, 波动率越大, 越能从多头 Gamma 中获益.BSM 的 Vega 是正的, 波动率的增加会使期权价格增加.

5.7.5 看跌、看涨平价公式

定理 5.9 (看跌、看涨平价公式)

令看跌期权的价格为 P(t,S(t)), 则看涨期权的价格为 C(t,S(t)), 则在 BSM 模型下, 看涨期权和看跌期权的价格之间有如下关系:

$$C(t, S(t)) - P(t, S(t)) = S(t) - Ke^{-r(T-t)}$$

用 BSM 函数表示为:

$$P(t, S(t)) = Ke^{-r(T-t)}N(-d_2) - S(t)N(-d_1)$$

看跌看涨平价公式告诉我们,看涨期权和看跌期权的价格之间的差额等于股票价格与折现敲定价格之间的差额,这实际上就是**远期合约**¹⁵的价值.

5.8 多元随机分析初步

5.8.1 多元布朗运动

定义 5.10 (多元布朗运动)

设 $W(t) = (W_1(t), W_2(t), \dots, W_n(t))$ 是 n 个独立的一维布朗运动组成的随机向量,则 W(t) 是 n 维布朗运动.

相应于 n 维布朗运动, 我们有域流 \mathcal{F}_t , 使得以下性质成立:

- 1. W(t) 是 \mathcal{F}_t 可测的.
- 2. 对于 0 < s < t, 有 $W(t) W(s) \sim N(0, \Sigma(t-s))$, 其中 $\Sigma(t-s)$ 是 $n \times n$ 的对称半正定矩阵.
- 3. (未来增量的独立性): 对于 $0 \le t < u$, 增量向量 W(u) W(t) 独立于 \mathcal{F}_t .

同样地, 我们有分量的二次变差公式: $dW_i(t)dW_i(t) = dt$, 以及: 样本交互变差为 0:

命题 5.7

设 $\Pi = \{t_0, t_1, \cdots, t_n\}$ 是 [0, T] 上的一个分划,对于 $i \neq j, W_i, W_j$ 在 [0, T] 上的样本交互变差定义为 $C_\Pi = \sum_{k=0}^{n-1} [W_i(t_{k+1}) - W_i(t_k)][W_j(t_{k+1}) - W_j(t_k)].$

- 1. $\mathbb{E} C_{\Pi} = 0$;
- 2. $Var(C_{\Pi}) \rightarrow 0$

 $^{^{15}}$ 交割价格为 K 的远期合约持有人在到期日 T 必须支付 K 用以购买一份股票

3. $C_{\Pi} \to \mathbb{E}C_{\Pi} = 0$

证明

1. 由于 W_i, W_j 是独立的, 所以 $[W_i(t_{k+1}) - W_i(t_k)]$ 和 $[W_j(t_{k+1}) - W_j(t_k)]$ 是独立的, 所以 $\mathbb{E}C_{\Pi} = 0$.

$$\operatorname{Var}(C_{\Pi}) = \mathbb{E}C_{\Pi}^{2} = \mathbb{E}\sum_{k=0}^{n-1} [W_{i}(t_{k+1}) - W_{i}(t_{k})][W_{j}(t_{k+1}) - W_{j}(t_{k})]^{2}$$

$$= \sum_{k=0}^{n-1} \mathbb{E}[W_{i}(t_{k+1}) - W_{i}(t_{k})]^{2} \mathbb{E}[W_{j}(t_{k+1}) - W_{j}(t_{k})]^{2}$$

$$= \sum_{k=0}^{n-1} (t_{k+1} - t_{k})^{2} \le ||\Pi|| \cdot \sum_{k=0}^{n-1} (t_{k+1} - t_{k}) = ||\Pi||T$$

$$\xrightarrow{||\Pi|| \to 0} 0$$

3. 均方收敛.

5.8.2 多个过程的 Ito 公式

定义 5.11 (多维 Ito 过程)

式

$$X = \Theta dt + \Sigma dW$$

表示一个多维 Ito 过程, 其中:

- $X \in \mathbb{Z}$ 4 $\mathbb{Z} = (X_1, X_2, \cdots, X_n)^{\mathsf{T}};$
- Θ 是一个 n 维随机过程: $\Theta = (\theta_1, \theta_2, \cdots, \theta_n)^{\mathsf{T}};$
- Σ 是一个 $n \times m$ 的矩阵:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nm} \end{pmatrix}$$

• W 是一个 m 维布朗运动.

上述定义写成积分形式是:

$$X_{i}(t) = X_{i}(0) + \int_{0}^{t} \theta_{i}(s)ds + \sum_{i=1}^{m} \int_{0}^{t} \sigma_{ij}(s)dW_{j}(s)$$

定理 5.10 (多维 Ito 公式)

设 $X = (X_1, X_2, \dots, X_n)^{\top}$ 是一个 n 维 Ito 过程, 设 Σ 的行向量分别是 $\sigma_1, \dots, \sigma_n$, 设 $f: \mathbb{R}^n \times [0, T] \to \mathbb{R}$ 是一个光滑函数, 则 f(X, t) 是一个 Ito 过程, 并且有如下 Ito 公式:

$$df(\boldsymbol{X},t) = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_{i}}dX_{i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^{2} f}{\partial x_{i} \partial x_{j}}dX_{i}dX_{j}$$

$$df(\boldsymbol{X},t) = \left(\frac{\partial f}{\partial t} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \langle \boldsymbol{\sigma_i}, \boldsymbol{\sigma_j} \rangle \frac{\partial^2 f}{\partial x_i \partial x_j}\right) dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dX_i$$

其中 $<\sigma_i,\sigma_j>$ 表示行向量 σ_i 和 σ_j 的内积.

C

证明 证明概要. 注意以下乘法规则:

$$dtdt = 0$$
, $dW_i dW_i = dt$, $dtdW_i = dW_i dt = 0$, $dW_i dW_i = 0$

则有 $d\mathbf{W}d\mathbf{W}^{\top} = dt\mathbf{I}_{m \times m}$. 因此有:

$$d\mathbf{X}d\mathbf{X}^{\top} = \mathbf{\Sigma}d\mathbf{W}d\mathbf{W}^{\top}\mathbf{\Sigma}^{\top} = \mathbf{\Sigma}\mathbf{\Sigma}^{\top}dt$$

展开写其中一个 $X_i(t)$ 有 $X_i(t) = X_i(0) + \int_0^t \theta_i(s) ds + \sum_{j=1}^m \int_0^t \sigma_{ij}(s) dW_j(s)$. 后面每一个伊藤积分 $\int_0^t \sigma_{ij}(u) dW_j(u)$ 在单位时间内累计二次变差的速率为 $\sigma_{ij}^2(u)$, 因此过程 $X_i(t)$ 累计二次变差的速率为 $\sum_{j=1}^m \sigma_{ij}^2(u)$, 这就是 $\Sigma\Sigma^{\mathsf{T}}$ 的对角线元素. 而对于 $i \neq j$, 把 dX_i 和 dX_j 相乘, 就能看出其累计二次变差的速率就是行向量 σ_i 和 σ_j 的内积:

$$dX_i(t) = \theta_i(t)dt + \sigma_{i1}dW_1(t) + \sigma_{i2}dW_2(t) + ...$$

$$\Delta X_j(t) = \theta_j(t)dt + \sigma_{j1}dW_1(t) + \sigma_{j2}dW_2(t) + ...$$

$$\sigma_{ik}\sigma_{ik}dW_k(t)dW_k(t) = \sigma_{ik}\sigma_{ik}dt$$

图 5.3: 多维 Ito 公式证明示意图

因此 X 的累计二次变差的速率为 $\Sigma\Sigma^{\mathsf{T}}$. 对 f(X,t) 进行泰勒展开,有:

$$df = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dX_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}dX_i dX_j +$$
高阶项

可证明高阶项都为 0. 因此第一式证得. 把 $dX_i dX_i$ 代入, 得第二式.

推论 5.3 (Ito 乘积法则)

设X(t),Y(t)是两个Ito 过程,则:

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + dX(t)dY(t)$$

证明 令 f(x,y,t)=xy, 则有 $f_t=0$, $f_x=y$, $f_y=x$, $f_{xx}=0$, $f_{xy}=1$, $f_{yy}=0$. 代入二维 Ito 公式即可.

笔记 与一般微积分下的乘法公式相比, 这里多出来了一个 dX(t)dY(t) 项, 这也是由于布朗运动二次变差不为 0 导致的.

定理 5.11 (多维 Levy 布朗运动识别定理)

设 $\boldsymbol{X}(t) = (X_1(t), X_2(t), \cdots, X_n(t))^{\top}$ 是一个相应于域流 \mathcal{F}_t 的 n 维鞅, 则 $\boldsymbol{X}(t)$ 是一个 n 维布朗运动的充要条件是:

- 1. X 是一个连续过程;
- 2. 对于 $i \neq j$, 有 $[X_i, X_j](t) = 0$;
- 3. 对于 $i = 1, 2, \dots, n$, 有 $[X_i, X_i](t) = t$.

证明 由多维 Ito 公式以及已知条件,有:

$$df(\boldsymbol{X},t) = \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dX_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}dX_i dX_j$$
$$= \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dX_i + \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}dt$$

两边同时积分,得到:

$$f(\boldsymbol{X},t) = f(\boldsymbol{0},0) + \int_0^t \left(\frac{\partial f}{\partial t}(\boldsymbol{X}(s),s) + \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(\boldsymbol{X}(s),s) \right) ds + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i} dX_i(s)$$

最后一个和式是鞅(证明和证明伊藤积分是鞅如出一辙).于是其期望值为0.因此:

$$\mathbb{E}f(\boldsymbol{X},t) = f(\boldsymbol{0},0) + \mathbb{E}\int_0^t \left(\frac{\partial f}{\partial t}(\boldsymbol{X}(s),s) + \frac{1}{2}\sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(\boldsymbol{X}(s),s)\right) ds$$

固定数 u_1, u_2, \cdots, u_n , 定义:

$$f(\boldsymbol{x},t) = \exp\left\{\boldsymbol{u}^{\top}\boldsymbol{x} - \frac{1}{2}\boldsymbol{u}^{\top}\boldsymbol{u}t\right\}$$

则: $f_t = -\frac{1}{2} \boldsymbol{u}^{\top} \boldsymbol{u} f, f_{x_i} = u_i f, f_{x_i x_i} = u_i^2 f$ 因此对此函数,

$$\mathbb{E}f(\boldsymbol{X},t) = f(\boldsymbol{0},0) + \mathbb{E}\int_0^t \left(\frac{\partial f}{\partial t}(\boldsymbol{X}(s),s) + \frac{1}{2}\sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(\boldsymbol{X}(s),s)\right) ds$$
$$= f(\boldsymbol{0},0) = 1$$

因而:

$$\mathbb{E} \exp \left\{ \boldsymbol{u}^{\top} \boldsymbol{X}(t) - \frac{1}{2} \boldsymbol{u}^{\top} \boldsymbol{u} t \right\} = 1$$

从而给出矩母函数:

$$\mathbb{E}\exp\left\{\sum_{i=1}^{n}u_{i}X_{i}(t)\right\} = \prod_{i=1}^{n}e^{\frac{1}{2}u_{i}^{2}t}$$

联合矩母函数可以因子分解,说明各个分量都是独立的一维布朗运动,因此 X(t) 是多维布朗运动.

例题 5.6 相关股价金融模型.

设

$$\frac{dS_1(t)}{S_1(t)} = \alpha_1 dt + \sigma_1 dW_1(t) \frac{dS_2(t)}{S_2(t)} = \alpha_2 dt + \sigma_2 \left[\rho dW_1(t) + \sqrt{1 - \rho^2} dW_2(t) \right]$$

其中 $W_1(t)$ 和 $W_2(t)$ 是相互独立的布朗运动, $\sigma_1 > 0$, $\sigma_2 > 0$ 和 $-1 \le \rho \le 1$ 是常数. 为了分析第二个股价过程, 我们定义:

$$W_3(t) = \rho W_1(t) + \sqrt{1 - \rho^2} W_2(t)$$

于是, $W_3(t)$ 是连续鞅, 满足 $W_3(0) = 0$, 以及

$$dW_3(t)dW_3(t) = \rho^2 dW_1(t)dW_1(t) + 2\rho\sqrt{1-\rho^2}dW_1(t)dW_2(t) + (1-\rho^2)dW_2(t)dW_2(t)$$
$$= \rho^2 dt + (1-\rho^2)dt = dt$$

换言之, $[W_3, W_3](t) = t$. 根据一维情形的 Levy 定理4.6, $W_3(t)$ 是布朗运动. 由于我们可以将 $S_2(t)$ 的微分写为:

$$\frac{dS_2(t)}{S_2(t)} = \alpha_2 dt + \sigma_2 dW_3(t)$$

故可知 $S_2(t)$ 是平均回报率为 α_2 、波动率为 σ_2 的几何布朗运动.

布朗运动 $W_1(t)$ 和 $W_3(t)$ 是相关的. 根据伊藤乘积法则有:

$$d(W_1(t)W_3(t)) = W_1(t)dW_3(t) + W_3(t)dW_1(t) + dW_1(t)dW_3(t)$$

= $W_1(t)dW_3(t) + W_3(t)dW_1(t) + \rho dt$

两边积分可得:

$$W_1(t)W_3(t) = \int_0^t W_1(s)dW_3(s) + \int_0^t W_3(s)dW_1(s) + \rho t$$

上式右端的伊藤积分期望值为 0, 于是 $W_1(t)$ 和 $W_3(t)$ 的协方差为:

$$E[W_1(t)W_3(t)] = \rho t$$

由于 $W_1(t)$ 和 $W_3(t)$ 的标准差均为 \sqrt{t} , 数 ρ 是 $W_1(t)$ 与 $W_3(t)$ 的相关性系数.

5.9 布朗桥再讨论

在第四章介绍过布朗桥, 见定义4.9. 这个定义是一个时间从 0 到 1 的布朗桥, 且状态从 0 到 0. 现在我们来讨论一个更一般的布朗桥.

定义 5.12 (布朗桥)

设W(t)是一个布朗运动,对于T>0,定义[0,T]上的从0到0的布朗桥X(t)为:

$$X(t) = W(t) - \frac{t}{T}W(T), t \in [0, T]$$



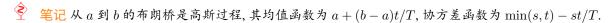
- 1. 作为 t 的函数,tW(T)/T 是连接 (0,0) 和 (T,W(T)) 的直线,由布朗运动减去该直线,得到的过程 X(t) 是一个从 0 到 0 的布朗桥,即 X(0)=X(T)=0.
- 2. 布朗桥并不适应于由 W(t) 生成的域流 F_t .
- 3. 从 0 到 0 的布朗桥是高斯过程, 其均值函数为 0, 协方差函数为 $\min(s,t) st/T$.

定义 5.13 (不水平布朗桥)

设W(t)是一个布朗运动,对于T>0,定义[0,T]上的从a到b的布朗桥X(t)为:

$$X^{a \to b}(t) = a + (b - a)\frac{t}{T} + X(t), t \in [0, T]$$

其中 $X(t) = X^{0 \to 0}$ 是一个从 0 到 0 的布朗桥.



定理 5.12 (布朗桥作为按比例缩放型随机积分)

定义 [0,T] 上的连续过程:

$$Y(t) = \begin{cases} (T-t) \int_0^t \frac{1}{T-u} dW(u), & 0 \le t \le T \\ 0, & t = T \end{cases}$$

则 Y(t) 是 [0,T] 上的从 0 到 0 的布朗桥.

证明 根据定理5.4, 积分 $\int_0^t \frac{1}{T-u} dW(u)$ 是一个正态过程. 只需要 Y(t) 的均值函数和协方差函数和布朗桥一致即可, 因为正态过程的二阶矩完全决定其分布.

均值为 0, 下面计算协方差函数. 假设 $0 \le s \le t < T$, 令 $I(t) = \int_0^t \frac{1}{T-u} dW(u)$, 则:

$$\begin{split} \mathbb{E}[Y(s)Y(t)] &= \mathbb{E}[(T-s)I(s)(T-t)I(t)] \\ &= (T-s)(T-t)\mathbb{E}[I(s)I(t)] \\ &= (T-s)(T-t)\mathbb{E}\left[\int_0^s \frac{1}{T-u}dW(u)\int_0^t \frac{1}{T-v}dW(v)\right] \\ &= (T-s)(T-t)\mathbb{E}\left[\int_0^s \frac{1}{T-u}dW(u)\left(\int_0^s \frac{1}{T-v}dW(v) + \int_s^t \frac{1}{T-v}dW(v)\right)\right] \\ &\xrightarrow{\underline{I(s)} \ni I(t) - I(s) \stackrel{\text{def}}{=}}} (T-s)(T-t)\mathbb{E}\left(\int_0^s \frac{1}{T-u}dW(u)\right)^2 \\ &\xrightarrow{\underline{\#} \stackrel{\text{def}}{=}}} (T-s)(T-t)\int_0^s \frac{1}{(T-u)^2}du \end{split}$$

$$= (T - s)(T - t)\left(\frac{1}{T - s} - \frac{1}{T}\right)$$
$$= s - \frac{st}{T}$$

证毕.

◈ 箕记

- 如上定义的 Y(t) 适应于布朗运动生成的域流 F_t .
- 当 $t \uparrow T$ 时, 过程 Y(t) 几乎必然收敛与 0.

通过方差的计算, 我们可以看出 Y(t) 的方差是 t(1-t/T), 当 $t \uparrow T$ 时, 方差趋于 0. 计算 Y(t) 的随机微分, 有:

$$dY(t) = dW(t) - \frac{Y(t)}{T - t}dt$$

当 $t \uparrow T$ 时, 如果 Y(t) 是正的, 则漂移项是负的, 驱使 Y 趋向 0, 如果 Y(t) 是负的, 则漂移项是正的, 也驱使 Y 趋向 0.

布朗桥作为条件布朗运动*

为了说明布朗桥其实就是条件布朗运动, 首先研究布朗桥的多维分布. 给定 $a \in \mathbb{R}$ 和 $b \in \mathbb{R}$, 设 $X^{a \to b}(t)$ 为 [0,T] 上从 a 到 b 的不水平布朗桥. 固定 $0 = t_0 < t_1 < t_2 < \cdots < t_n < T$. 我们计算 $X^{a \to b}(t_1)$, \cdots , $X^{a \to b}(t_n)$ 的联合密度.

从a到b的布朗桥具有均值函数

$$m^{a\to b}(t)=a+\frac{(b-a)t}{T}=\frac{(T-t)a}{T}+\frac{bt}{T}$$

和协方差函数

$$c(s,t) = s - \frac{st}{T} = \frac{s(T-t)}{T}, \quad 0 \le s \le t \le T$$

$$Z_{j} = \frac{X^{a \to b}(t_{j})}{\tau_{i}} - \frac{X^{a \to b}(t_{j-1})}{\tau_{j-1}}$$

由于 $X^{a\to b}(t_1), \cdots, X^{a\to b}(t_n)$ 为联合正态, 故 Z_1, \cdots, Z_n 也联合正态. 计算可得:

$$\mathbb{E}Z_{j} = \frac{1}{\tau_{j}} \mathbb{E}X^{a \to b}(t_{j}) - \frac{1}{\tau_{j-1}} \mathbb{E}X^{a \to b}(t_{j-1})$$

$$= \frac{a}{T} + \frac{bt_{j}}{T_{\tau_{j}}} - \frac{a}{T} - \frac{bt_{j-1}}{T_{\tau_{j-1}}}$$

$$= \frac{bt_{j}(T - t_{j-1}) - bt_{j-1}(T - t_{j})}{T_{\tau_{j}\tau_{j-1}}}$$

$$= \frac{b(t_{j} - t_{j-1})}{\tau_{j}\tau_{j-1}}$$

并且

$$\begin{aligned} \operatorname{Var}\left(Z_{j}\right) &= \frac{1}{\tau_{j}^{2}} \operatorname{Var}\left(X^{a \to b}\left(t_{j}\right)\right) - \frac{2}{\tau_{j}\tau_{j-1}} \operatorname{Cov}\left(X^{a \to b}\left(t_{j}\right), X^{a \to b}\left(t_{j-1}\right)\right) + \frac{1}{\tau_{j-1}^{2}} \operatorname{Var}\left(X^{a-b}\left(t_{j-1}\right)\right) \\ &= \frac{1}{\tau_{j}^{2}} c\left(t_{j}, t_{j}\right) - \frac{2}{\tau_{j}\tau_{j-1}} c\left(t_{j}, t_{j-1}\right) + \frac{1}{\tau_{j-1}^{2}} c\left(t_{j-1}, t_{j-1}\right) \\ &= \frac{t_{j}}{T_{\tau_{j}}} - \frac{2t_{j-1}}{T_{\tau_{j-1}}} + \frac{t_{j-1}}{T_{\tau_{j-1}}} \\ &= \frac{t_{j}(T - t_{j-1}) - 2t_{j-1}(T - t_{j}) + t_{j-1}(T - t_{j})}{T\tau_{j}\tau_{j-1}} \\ &= \frac{t_{j} - t_{j-1}}{\tau_{j}\tau_{j-1}} \end{aligned}$$

最后, 当 i < j 时, 计算 Z_i 与 Z_j 的协方差. 我们得到:

$$\operatorname{Cov}(Z_{i}, Z_{j}) = \frac{1}{\tau_{i}\tau_{j}}c(t_{i}, t_{j}) - \frac{1}{\tau_{i}\tau_{j-1}}c(t_{i}, t_{j-1}) - \frac{1}{\tau_{i-1}\tau_{j}}c(t_{i-1}, t_{j}) + \frac{1}{\tau_{i-1}\tau_{j-1}}c(t_{i-1}, t_{j-1})
= \frac{t_{i}(T - t_{j})}{T_{\tau_{i}\tau_{j}}} - \frac{t_{i}(T - t_{j-1})}{T_{\tau_{i}\tau_{j-1}}} - \frac{t_{i-1}(T - t_{j})}{T_{\tau_{i-1}-1}\tau_{j}} + \frac{t_{i-1}(T - t_{j-1})}{T_{\tau_{i-1}}\tau_{j-1}}
= 0$$

由此可知, 正态随机变量 Z_1, \dots, Z_n 是两两独立的, 它们的联合密度可以表示为:

$$f_{Z(t_1),Z(t_2),\cdots,Z(t_n)}(z_1,\cdots,z_n)$$

$$= \prod_{j=1}^n \frac{1}{\sqrt{2\pi \frac{t_j-t_{j-1}}{\tau_j\tau_{j-1}}}} \exp\left\{-\frac{1}{2} \cdot \frac{\left(z_j - \frac{b(t_j-t_{j-1})}{\tau_j\tau_{j-1}}\right)^2}{\frac{t_j-t_{j-1}}{\tau_j\tau_{j-1}}}\right\}$$

$$= \exp\left\{-\frac{1}{2} \cdot \sum_{j=1}^n \frac{\left(z_j - \frac{b(t_j-t_{j-1})}{\tau_j\tau_{j-1}}\right)^2}{\frac{t_j-t_{j-1}}{\tau_j\tau_{j-1}}}\right\} \cdot \prod_{j=1}^n \frac{1}{\sqrt{2\pi \frac{t_j-t_{j-1}}{\tau_j\tau_{j-1}}}}$$

作变量代换:

$$z_j = \frac{x_j}{\tau_j} - \frac{x_{j-1}}{\tau_{j-1}}, \quad j = 1, \dots, n$$

雅可比矩阵为

$$J = \begin{pmatrix} \frac{1}{\tau_1} & 0 & \cdots & 0 \\ -\frac{1}{\tau_1} & \frac{1}{\tau_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\tau_n} \end{pmatrix}$$

因此我们得到 $X^{a-b}(t_1), \dots, X^{a\to b}(t_n)$ 的联合密度:

$$\begin{split} & f_{x^{a \to b}(t_1), \cdots, x^{a \to b}(t_n)} \left(x_1, \cdots, x_n \right) \\ & = \prod_{j=1}^n \frac{1}{\sqrt{2\pi \left(t_j - t_{j-1} \right)}} \sqrt{\frac{\tau_{j-1}}{\tau_j}} \cdot \exp\left\{ -\frac{1}{2} \cdot \sum_{j=1}^n \frac{\left(x_j - x_{j-1} \right)^2}{t_j - t_{j-1}} - \frac{\left(b - x_n \right)^2}{2 \left(T - t_n \right)} + \frac{\left(b - a \right)^2}{2T} \right\} \\ & = \sqrt{\frac{T}{T - t_n}} \cdot \prod_{j=1}^n \frac{1}{\sqrt{2\pi \left(t_j - t_{j-1} \right)}} \cdot \exp\left\{ -\frac{1}{2} \cdot \sum_{j=1}^n \frac{\left(x_j - x_{j-1} \right)^2}{t_j - t_{j-1}} - \frac{\left(b - x_n \right)^2}{2 \left(T - t_n \right)} + \frac{\left(b - a \right)^2}{2T} \right\} \\ & = \frac{p \left(T - t_n, x_n, b \right)}{p \left(T, a, b \right)} \prod_{j=1}^n p \left(t_j - t_{j-1}, x_{j-1}, x_j \right) \end{split}$$

其中:

$$p(\tau, x, y) = \frac{1}{\sqrt{2\pi\tau}} \exp\left\{-\frac{(y-x)^2}{2\tau}\right\}$$

是布朗运动的转移密度.

定理 5.13 (布朗桥就是条件布朗运动)

从 a 到 b 的不水平布朗桥的分布和在时间区间 [0,T] 上以 W(0)=a,W(T)=b 为条件的布朗运动的分布一致.

证明 给定 $0 = t_0 < t_1 < t_2 < \cdots < t_n < T, W(t_1), \cdots, W(t_n), W(T)$ 的联合密度为:

$$f_{W(t_1),\dots,W(t_n),W(T)}(x_1,\dots,x_n,b) = p(T-t_n,x_n,b) \prod_{i=1}^n p(t_j-t_{j-1},x_{j-1},x_j)$$

其中 $W(0)=x_0=a$. 这是因为 $p(t_1-t_0,x_0,x_1)=p(t_1,a,x_1)$ 是布朗运动在 t=0 与 $t=t_1$ 之间从 W(0)=a

 \Diamond

到 $W(t_1) = x_1$ 的密度. 类似地, $p(t_2 - t_1, x_1, x_2)$ 是在 $t = t_1$ 与 $t = t_2$ 之间从 $W(t_1) = x_1$ 到 $W(t_2) = x_2$ 的密度. 于是由独立性, $W(t_1)$ 和 $W(t_2)$ 的联合密度就是乘积

$$p(t_1, a, x_1) p(t_2 - t_1, x_1, x_2)$$

如此继续, 我们就得到给出的布朗运动联合密度式. W(T) 的边际密度为 p(T,a,b). 因此以 W(T)=b 为条件, $W(t_1), \dots, W(t_n)$ 的密度为:

$$\frac{p(T - t_n, x_n, b)}{p(T, a, b)} \prod_{i=1}^{n} p(t_j - t_{j-1}, x_{j-1}, x_j)$$

与不水平布朗桥联合密度一致.

推论 5.4 (布朗桥最大值的分布)

定义 $M^{a\to b}(T) = \max_{0 < t < T} X^{a\to b}(t)$, 其密度为:

$$f_M(y) = \frac{2(2y - b - a)}{T} \exp\left\{-\frac{2}{T}(y - a)(y - b)\right\}, \quad y > \max\{a, b\}$$

证明 由于 [0,T] 上从 0 到 w 的布朗桥是以 W(T) = w 为条件的布朗运动, [0,T] 上 $X^{0\to\omega}(T)$ 的最大值就是 [0,T] 上以 W(T) = w 为条件 W 的最大值. 因此, $M^{0\to w}(T)$ 的密度可由推论4.3算得:

$$f_{M^{0 \to w}(T)}(x) = \frac{2(2x - w)}{T}e^{-2x(x - w)}, \quad w < x, x > 0$$

密度 $f_{M^{a\to b}(T)}(y)$ 的证明, 以 y-a 取代 x, 并以 b-a 取代 w 即可.

5.10 附录

5.10.1 资产定价前置定理

定理 5.14 (Girsanov 定理)

设 W(t) 是一个布朗运动, 对于任意 t > 0, 定义 F_t 上的概率测度 \mathbb{Q} 如下:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left\{-\int_0^t \theta(s)dW(s) - \frac{1}{2}\int_0^t \theta^2(s)ds\right\}$$

其中 $\theta(t)$ 是 F_t 可测的随机过程. 则 \mathbb{P} 下的布朗运动:

$$\widetilde{W}(t) = W(t) + \int_0^t \theta(s)ds$$

在 ℚ 下是一个标准布朗运动.

本定理可视为定理5.6的推广.

证明 首先证明 $\widetilde{W}(t)$ 在 \mathbb{Q} 下是一个鞅. 首先验证 $Z(t) = \frac{d\mathbb{Q}}{d\mathbb{P}}$ 是一个鞅. 令 $X(t) = -\int_0^t \theta(u)dW(u) - \frac{1}{2}\int_0^t \theta^2(u)du$, 以及 $f(x) = e^x$, 因此:

$$dZ(t) = Z(t)dX(t) + \frac{1}{2}Z(t)d[X,X](t) = -Z(t)\theta(t)dW(t)$$

积分得到: $Z(t) = Z(0) - \int_0^t Z(s)\theta(s)dW(s)$, 因此 Z(t) 是一个鞅.

验证 $\widetilde{W}(t)Z(t)$ 在 \mathbb{P} 下是鞅.

$$\begin{split} d(\widetilde{W}(t)Z(t)) &= Z(t)d\widetilde{W}(t) + \widetilde{W}(t)dZ(t) + d[\widetilde{W},Z](t) \\ &= Z(t)dW(t) + \theta(t)Z(t)dt + \widetilde{W}(t)(-Z(t)\theta(t)dW(t)) + (dW(t) + \theta(t)dt)(-\theta(t)Z(t)dW(t)) \\ &= (-\widetilde{W}(t)\theta(t) + 1)Z(t)dW(t) \end{split}$$

表达式不含 dt, 因此 $\widetilde{W}(t)Z(t)$ 是一个鞅. 则

$$\mathbb{E}_{\mathbb{Q}}[\widetilde{W}(t)|\mathcal{F}_s] = \frac{1}{Z(s)} \mathbb{E}[\widetilde{W}(t)Z(t)|\mathcal{F}_s] = \frac{1}{Z(s)} \widetilde{W}(s)Z(s) = \widetilde{W}(s), \quad 0 \le s \le t$$

证毕.

定理 5.15 (风险中性测度)

定义适当的 $\theta(t)$, 就称如 Girsanov 定理所述的测度 $\mathbb Q$ 为风险中性测度, 在此测度下, 资产组合的预期收益率等于无风险利率 (即货币市场账户的收益率).

为了说明这一点, 定义股价过程为 $dS(t) = \mu(t)S(t)dt + \sigma(t)S(t)dW(t)$, 其中 $\mu(t)$ 是股价的预期收益率, $\sigma(t)$ 是股价的波动率. 定义无风险利率过程为 r(t), 定义贴现过程为 $D(t) = e^{-\int_0^t r(s)ds}$ (即 dD(t)=-r(t)D(t)dt), 则贴现股价过程为:

$$\begin{split} dD(t)S(t) &= D(t)dS(t) + S(t)dD(t) + dD(t)dS(t) \\ &= D(t)\mu(t)S(t)dt + D(t)\sigma(t)S(t)dW(t) + S(t)(-r(t)D(t)dt) \\ &= D(t)S(t)(\mu(t) - r(t))dt + D(t)\sigma(t)S(t)dW(t) \\ &= \sigma(t)D(t)S(t)[\theta(t)dt + dW(t)] \end{split}$$

其中 $\theta(t) = \frac{\mu(t) - r(t)}{\sigma(t)}$ 称为风险的市场价格.

用 Girsanov 定理中的概率测度 Q, 可以把上式改写为

$$d(D(t)S(t)) = D(t)S(t)\sigma(t)d\widetilde{W}(t)$$

这说明, 在 \bigcirc 下, 贴现股价过程是一个鞅. 而在股价过程中用 \widetilde{W} 替代, 可以得到:

$$dS(t) = r(t)S(t)dt + \sigma(t)S(t)d\widetilde{W}(t)$$

这就说明在风险中性测度下,股价的平均回报率就等于无风险利率.回顾 BSM 模型的推导过程,我们当时做出的变换其实就是将股价过程转换到了风险中性测度下.

而对于资产组合,有:

$$\begin{split} dX(t) &= \Delta(t)dS(t) + r(t)(X(t) - \Delta(t)S(t))dt \\ &= \Delta(t)(\mu(t)S(t)dt + \sigma(t)S(t)dW(t)) + r(t)(X(t) - \Delta(t)S(t))dt \\ &= r(t)X(t)dt + \Delta(t)\sigma(t)S(t)d\widetilde{W}(t) \end{split}$$

亦说明资产组合的平均回报率等于无风险利率.

计算资产组合的贴现过程,有:

$$\begin{split} dD(t)X(t) &= D(t)dX(t) + X(t)dD(t) + dD(t)dX(t) \\ &= D(t)r(t)X(t)dt + D(t)\Delta(t)\sigma(t)S(t)d\widetilde{W}(t) + X(t)(-r(t)D(t)dt) \\ &= D(t)\Delta(t)\sigma(t)S(t)d\widetilde{W}(t) \end{split}$$

亦说明资产组合的贴现过程是一个鞅.

定义 5.14 (风险中性定价公式)

设 V(T) 是 \mathcal{F}_T -可测的随机变量, 表示在时刻 T 某金融衍生品的定价 (支付). 投资者希望选择 X(0) 以及资产组合策略 $\Delta(t)$ 使能够对冲金融衍生品空头, 即 X(T)=V(T). 并有如下公式a:

$$V(t) = \mathbb{E}_{\mathbb{Q}} \left[\exp \left\{ - \int_{t}^{T} r(u) du \right\} V(T) \mid \mathcal{F}_{t} \right]$$

a蕴含 X(t) = V(t)

回顾 BSM 模型的推导过程, 我们就是在算 $C(t,S(t)) = \mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)(S(T)-K)^+|\mathcal{F}_t}].$

上述风险中性定价公式成立的前提假设是投资者能够选择这种对冲策略,也就是对冲资产组合的存在性问题.该问题依赖于如下定理:

定理 5.16 (鞅表示定理)

设 $W(t), 0 \le t \le T$ 是概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的布朗运动, $\mathcal{F}(t), 0 \le t \le T$ 是由该布朗运动生成的域流. 设 $M(t), 0 \le t \le T$ 关于这一域流是一个鞅, 则存在一个适应过程 $\Gamma(u), 0 \le u \le T$, 使得:

$$M(t) = M(0) + \int_0^t \Gamma(u)dW(u), \quad 0 \le t \le T$$

 \Diamond

Ŷ 笔记

- 1. 该定理意思是, 当域流中仅含截至时刻 t 观测布朗运动所获得的信息, 则关于这个域流的每一个鞅都可以表示为初始条件和关于该布朗运动的一个 Ito 积分.
- 2. 模型中唯一不确定源是布朗运动, 因此只有一个不确定性源需要通过对冲消除. 接下来我们来解释这个定理. 由风险中性定价公式决定 V(t), 就有:

$$D(t)V(t) = \mathbb{E}_{\mathbb{O}}[D(T)V(T)|\mathcal{F}_t]$$

由重期望公式可以证明 D(t)V(t) 是 \mathbb{Q} – 鞅. 因此由鞅表示定理, 存在一个适应过程 $\widetilde{\Gamma}(t)$, 使得:

$$D(t)V(t) = V(0) + \int_0^t \widetilde{\Gamma}(u)d\widetilde{W}(u), \quad 0 \le t \le T$$

另一方面,

$$D(t)X(t) = X(0) + \int_0^t \Delta(u)\sigma(u)D(u)S(u)d\widetilde{W}(u)$$

因此投资者只需选取

$$\Delta(t) = \frac{\widetilde{\Gamma}(t)}{\sigma(t)D(t)S(t)}$$

就可以完全对冲由定价公式给出的支付为V(T)的金融衍生品的空头.

5.10.2 资产定价基本定理

在具有多个股票的多维金融市场下, 我们把上一节讨论的问题总结如下.

- 1. 风险中性: 称概率测度 $\widetilde{\mathbb{P}}$ 是风险中性的, 如果: $\widetilde{\mathbb{P}}$ 与 \mathbb{P} 等价 (即对每个 $A \in \mathcal{F}, \mathbb{P}(A) = 0$ 当且仅当 $\widetilde{\mathbb{P}}(A) = 0$); 且在 \widetilde{P} 下, 贴现股价 $D(t)S_i(t), i = 1, \cdots, m$ 是鞅.
- 2. 风险的市场价格方程:

$$\mu_i(t) - r(t) = \sum_{j=1}^{d} \sigma_{ij}(t)\theta_j(t), \quad i = 1, \dots, m$$

- 3. 在风险中性测度下,贴现资产组合价值是鞅。
- 4. **套利**: 套利是满足以下条件的资产组合价值过程 X(t): X(0) = 0, 并且存在某个时刻 T > 0,

$$\mathbb{P}{X(T) > 0} = 1, \quad \mathbb{P}{X(T) > 0} > 0$$

定理 5.17 (资产定价第一定理)

具有风险中性概率测度的市场模型中不存在套利.

\odot

定理 5.18 (资产定价第二定理)

在具有风险中性概率测度的市场模型中,任何衍生品都能得到对冲当且仅当风险中性测度是唯一的.

第6章 随机微分方程与偏微分方程

6.1 随机微分方程

定义 6.1 (随机微分方程)

设 W(t) 是一个布朗运动, 对于给定的函数 f(t,x) 和 q(t,x), 随机微分方程 (SDE) 为:

$$dX(t) = \beta(t, X(t))dt + \gamma(t, X(t))dW(t)$$

其中 β , γ 是给定的函数,分别称为飘移项和扩散项.

给定初始条件 $X(0) = x_0$, 我们希望求解 X(t) 的解析解: X(t) 的解析解具有形式:

$$X(t) = x_0 + \int_0^t \beta(s, X(s))ds + \int_0^t \gamma(s, X(s))dW(s)$$

随机微分方程一般难以求解,但是一维线性随机微分方程

$$dX(t) = (a(t) + b(t)X(t))dt + (\gamma(t) + \sigma(t)X(t))dW(t)$$

有显式解, 其中 a, b, γ, σ 都是非随机函数.

6.2 费曼-卡茨定理

考虑由6.1定义的随机微分方程. 设 $h(\cdot)$ 是 Borel 可测函数, 对于满足初始条件 X(t) = x 的解 X(T), 记

$$u(t,x) = \mathbb{E}[h(X(T))|X(t) = x]$$

定理 6.1 (随机微分方程的解具有马尔可夫性)

设 $X(u), u \ge 0$ 是随机微分方程满足时刻 0 的初始条件的解. 则对于任意 $0 \le t \le T$, 有

$$\mathbb{E}[h(X(T))|\mathcal{F}_t] = u(t, X(t))$$

即随机微分方程的解是马尔可夫过程.

证明 证明轮廓: 对r > t,

$$X(r) = X(t) + \int_{t}^{r} \beta(s, X(s))ds + \int_{t}^{r} \gamma(s, X(s))dW(s)$$

X(r) - X(t) 与 \mathcal{F}_t 独立, 因此:

$$\mathbb{P}(X(r) \le \pi | \mathcal{F}_t) = \mathbb{P}(X(r) - X(t) \le \pi - x, X(t) = x | \mathcal{F}_t)$$
$$= \mathbb{P}(X(r) - X(t) \le \pi - x, X(t) = x)$$
$$= \mathbb{P}(X(r) \le \pi | X(t) = x)$$

因此说明了马尔可夫性.

推论 6.1

设X(u)是随机微分方程满足时刻0的初始条件的解,随机过程

$$g(t, X(t)), \quad 0 \le t \le T$$

是鞅.

证明 给定 $0 \le s \le t \le T$,

$$\mathbb{E}[h(X(T)) \mid \mathcal{F}(s)] = g(s, X(s))$$
$$\mathbb{E}[h(X(T)) \mid \mathcal{F}(t)] = g(t, X(t))$$

第二个等式两边取条件期望,利用累次条件期望以及第一个等式,我们得到:

$$\mathbb{E}[g(t, X(t)) \mid \mathcal{F}(s)] = \mathbb{E}[\mathbb{E}[h(X(T)) \mid \mathcal{F}(t)] \mid \mathcal{F}(s)]$$
$$= \mathbb{E}[h(X(T)) \mid \mathcal{F}(s)]$$
$$= g(s, X(s))$$

定理 6.2 (费曼-卡茨定理)

考虑由定义6.1给出的随机微分方程. 设 $h(\cdot)$ 是 Borel 可测函数. 给定 $T>0, t\in[0,T]$, 定义函数 $g(t,x)=\mathbb{E}[h(X(T))|X(t)=x]$, 则 g(t,x) 满足偏微分方程

$$g_t(t,x) + \beta(t,x)g_x(t,x) + \frac{1}{2}\gamma^2(t,x)g_{xx}(t,x) = 0$$

以及终值条件

$$g(T,x) = h(x), \quad \forall x$$

证明 证明概要:

$$dg(t, X(t)) = g_t dt + g_x dX + \frac{1}{2} g_{xx} dX dX$$
$$= g_t dt + \beta g_x dt + \gamma g_x dW + \frac{1}{2} \gamma^2 g_{xx} dt$$
$$= \left[g_t + \beta g_x + \frac{1}{2} \gamma^2 g_{xx} \right] dt + \gamma g_x dW$$

由于 g(t, X(t)) 是鞅, 所以上面微分式中 dt 项应该为 0. 则:

$$g_t(t, X(t)) + \beta(t, X(t))g_x(t, X(t)) + \frac{1}{2}\gamma^2(t, X(t))g_{xx}(t, X(t)) = 0$$

沿着 X 的任何路径成立. 因此:

$$g_t(t,x) + \beta(t,x)g_x(t,x) + \frac{1}{2}\gamma^2(t,x)g_{xx}(t,x) = 0$$

定理 6.3 (含贴现的费曼-卡茨定理)

承上设定、定义r为利率、函数

$$f(t,x) = \mathbb{E}[e^{-r(T-t)}h(X(T))|X(t) = x]$$

则 f(t,x) 满足偏微分方程:

$$f_t(t,x) + \beta(t,x)f_x(t,x) + \frac{1}{2}\gamma^2(t,x)f_{xx}(t,x) = rf(t,x)$$

以及终值条件

$$f(T,x) = h(x), \quad \forall x$$

证明 注意到 $e^{-rt}f(t,X(t)) = \mathbb{E}[e^{-rT}h(X(T))|\mathcal{F}_t]$, 于是可说明 $e^{-rt}f(t,X(t))$ 是鞅.

$$\begin{split} d\left(e^{-rt}f(t,X(t))\right) &= e^{-rt}\left[-rfdt + f_tdt + f_xdX + \frac{1}{2}f_{xx}dXdX\right] \\ &= e^{-rt}\left[-rf + f_t + \beta f_x + \frac{1}{2}\gamma^2 f_{xx}\right]dt + e^{-rt}\gamma f_xdW \end{split}$$

在上式令 dt 项为 0 即可.

笔记 费曼-卡茨定理说明了随机微分方程和偏微分方程的联系.

例题 6.1 BSM 方程的解.

股票价格符合:

$$dS(u) = rS(u)du + \sigma S(u)d\widetilde{W}(u)$$

即在这里, $\beta = rS(u)$, $\gamma = \sigma S(u)$.

则偏微分方程 (BSM)

$$\frac{\partial C}{\partial t} + rS(t)\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2S^2(t)\frac{\partial^2C}{\partial S^2} - rC(t,S(t)) = 0$$

的解应当具有形式:

$$C(t, S(t)) = \widetilde{\mathbb{E}}[e^{-r(T-t)}h(S(T))|\mathcal{F}_t]$$

这里 h(S(T)) 就是支付 $(S(T)-K)^+$. 因此我们通过费曼卡茨定理论证了风险中性定价公式就是 BSM 偏微分方程的解.

6.3 一些随机微分方程示例

例题 **6.2** 选择 X(t) = W(t), 它为 1 维的布朗运动,

$$f(t,x) = e^{ix} = (\cos x, \sin x) \in \mathbb{R}^2, \quad \forall x \in \mathbb{R},$$

那么由 Ito 公式可得

$$Y(t) = f(t, X(t)) = e^{iW(t)} = (\cos W(t), \sin W(t)),$$

又是一个 Ito 过程. 它的坐标分量, Y_1, Y_2 满足

$$dY_1(t) = -\sin(W(t)) dW(t) - \frac{1}{2}\cos(W(t)) dt,$$

$$dY_2(t) = \cos(W(t)) dW(t) - \frac{1}{2}\sin(W(t)) dt,$$

因此对过程 $Y = (Y_1, Y_2)$, 称之为单位圆上的布朗运动, 它是下面的随机微分方程组的解:

$$\begin{cases} dY_1(t) = -\frac{1}{2}Y_1dt - Y_2dW(t) \\ dY_2 = -\frac{1}{2}Y_2dt + Y_1dW(t) \end{cases}$$

或用矩阵形式有

例题 6.3 几何均值回复过程

X(t) 定义为如下随机微分方程

$$dX(t) = k \left(\alpha - \log X(t)\right) X(t)dt + \sigma X(t)dW(t), \quad X_0 = x > 0$$

的解. 这里 k, α, σ, x 都为正常数. 这个过程被 Tvedt(1995) 用于模拟海运中的现场运货率.

a) 解为

$$X(t) = \exp\left(e^{-kt}\ln x + \left(\alpha - \frac{\sigma^2}{2k}\right)\left(1 - e^{kt}\right) + \sigma e^{-kt}\int_0^t e^{ks}dW(s)\right).$$

把 $Y_t = \log X(t)$ 代入到微分方程中化为关于 Y_t 的线性方程.

b) 期望满足:

$$\mathbb{E}\left[X(t)\right] = \exp\left(e^{-kt}\ln x + \left(\alpha - \frac{\sigma^2}{2k}\right)\left(1 - e^{-kt}\right) + \frac{\sigma^2\left(1 - e^{-2kt}\right)}{2k}\right).$$

Bibliography

- [1] 王汉超, 于志勇. 应用随机分析 [M]. 高等教育出版社,2022.
- [2] 弗里德曼. 随机微分方程及其应用 [M]. 科学出版社,1983.
- [3] 胡适耕, 黄乘明, 吴付科著. 随机微分方程 [M]. 科学出版社,2008.
- [4] 林正炎, 陆传荣, 苏中根. 概率极限理论基础 [M]. 高等教育出版社,1999.
- [5] 林正炎, 苏中根, 张立新. 概率论: Probability theory[M]. 浙江大学出版社,2014.
- [6] 苏中根. 随机过程 [M]. 高等教育出版社,2016.
- [7] 钱忠民, 应坚刚.-. 随机分析引论 [M]. 复旦大学出版社有限公司,2017.
- [8] Durrett R .Probability. Theory and examples[J]. 2010.DOI:10.1017/CBO9780511779398.
- [9] Ee, Weinan, Li Tiejun, Vanden-Eijnden, Eric. Applied Stochastic Analysis[M]. 2019. DOI: 10.1090/gsm/199.
- [10] Shreve S E .Stochastic Calculus for Finance II[M]. 世界图书出版公司,2007.
- [11] 日伊藤清 -. 伊藤清概率论 [M]. 人民邮电出版社,2021.