

随机过程

杨敬轩

2024 年 9 月 12 日

目录

第1章 概率论与随机过程的基本概念	1
1.1 概率	1
1.2 随机变量、分布函数及数字特征	4
1.2.1 随机变量与分布函数	4
1.2.2 Riemann-Stieltjes 积分	5
1.2.3 数字特征	7
1.2.4 常用随机变量的分布	8
1.2.5 示性函数的线性组合	9
1.3 矩母函数、特征函数和拉普拉斯变换	10
1.3.1 矩母函数	10
1.3.2 特征函数	11
1.3.3 Laplace-Stieltjes 变换	11
1.4 条件数学期望	11
1.4.1 离散型随机变量	11
1.4.2 连续型随机变量	12
1.4.3 一般随机变量	13
1.4.4 条件概率与条件分布函数	14
1.4.5 条件数学期望的基本性质	15
1.4.6 多元随机变量的条件数学期望	15
1.4.7 条件乘法公式与条件独立性	16
1.4.8 全概率公式	17
1.5 随机过程的概念	17
1.5.1 随机过程的定义	17
1.5.2 随机过程的分布	18
1.5.3 随机过程的分类	18
第2章 泊松过程	21
2.1 定义及其背景	21
2.2 相邻事件的时间间隔	22
2.3 剩余寿命与年龄	24
2.4 到达时间的条件分布	24
2.5 复合泊松过程	26

2.6	更新过程	26
2.7	若干极限定理与基本更新定理	26
第3章	马尔可夫链	27
3.1	定义与例子	27
3.2	转移概率矩阵	28
3.3	状态的分类	29
3.4	状态空间的分解	32
3.5	转移矩阵的极限性态	33
3.5.1	转移矩阵的极限性态	33
3.5.2	平稳分布	34
3.5.3	极限分布	35
第4章	离散鞅引论	37
4.1	定义与例子	37
4.2	上下鞅及分解定理	38
4.3	停时与停时定理	39
4.4	鞅收敛定理	41
4.5	连续参数鞅	44
第5章	布朗运动	46
5.1	随机游动与布朗运动的定义	46
5.1.1	布朗运动矩母函数	46
5.1.2	布朗运动联合概率密度	47
5.1.3	布朗运动马尔可夫性	47
5.1.4	布朗运动与正态过程	47
5.1.5	布朗运动的鞅性	48
5.2	布朗运动轨道的性质	49
5.3	首达时与最大值	52
5.3.1	首达时的分布	52
5.3.2	首达时的期望	54
5.4	布朗桥	55
5.5	布朗运动的各种变形与推广	55
5.5.1	在某点被吸收的布朗运动	55
5.5.2	在点反射的布朗运动	55
5.5.3	几何布朗运动	56
5.5.4	布朗运动的积分	56
5.5.5	布朗运动的形式导数	56
5.6	带有漂移的布朗运动	56

5.7 多维布朗运动	57
第6章 连续时间参数马尔可夫链	58
6.1 定义与若干基本概念	58
6.2 转移率矩阵及其概率意义	61
6.3 Kolmogorov 向前向后微分方程	63
6.4 生灭过程	64
6.5 强马尔可夫性与嵌入马尔可夫链	64
第7章 随机微分方程	65
7.1 H 空间和均方收敛	65
7.2 均方分析	67
7.2.1 均方连续性	67
7.2.2 均方可导	68
7.2.3 均方积分	69
7.3 Itô 随机积分	70
7.4 Itô 过程与 Itô 公式	74
7.5 Itô 随机微分方程	78
7.5.1 常系数的线性随机微分方程	78
7.5.2 简单的线性齐次随机微分方程	78
7.5.3 一般的线性非齐次随机微分方程	79
7.6 解的存在性和唯一性问题	83

第1章 概率论与随机过程的基本概念

1.1 概率

概率论的一个基本概念是随机试验. 一个试验或观察, 若它的结果预先无法确定, 则称之为随机试验, 简称为试验. 所有试验的可能结果组成的集合, 称为样本空间, 记作 Ω . Ω 中的元素则称为样本点, 用 ω 表示. 由 Ω 的某些样本点构成的子集合, 常用大写字母 A, B, C 等表示, 由 Ω 中的若干子集构成的集合称为集类, 用花写字母 $\mathcal{A}, \mathcal{B}, \mathcal{F}$ 等表示.

由于并不是在所有的 Ω 的子集上都能方便地定义概率, 一般只限制在满足一定条件的集类 \mathcal{F} 上研究概率性质, 为此引入 σ 域的概念.

定义: σ 域与可测空间

设 \mathcal{F} 为由 Ω 的某些子集构成的非空集类, 若满足

(1) 若 $A \in \mathcal{F}$, 则 $A^c \in \mathcal{F}$, A^c 是 A 的补集, 即 $A^c = \bar{A} = \Omega - A$

(2) 若 $A_n \in \mathcal{F}$, $n \in \mathbb{N}$, 则

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F} \quad (1.1)$$

则称 \mathcal{F} 为 σ 域, 或 σ 代数, 称 (Ω, \mathcal{F}) 为可测空间.

σ 域性质

若 \mathcal{F} 为 σ 域, 则 \mathcal{F} 对可列次交、并、差等运算封闭, 即 \mathcal{F} 中的任何元素经可列次运算后仍属于 \mathcal{F} .

定义: 最小 σ 域

设 \mathcal{A} 为由 Ω 的某些子集构成的集类, 一切包含 \mathcal{A} 的 σ 域的交, 记为 $\sigma(\mathcal{A})$, 称 $\sigma(\mathcal{A})$ 为由 \mathcal{A} 生成的 σ 域, 或称为包含 \mathcal{A} 的最小 σ 域. 一维 Borel σ 域即为包含 \mathbb{R} 上所有形如集合 $(-\infty, a]$ 的最小 σ 域, 记为 $\mathcal{B}_{\mathbb{R}} = \sigma\{(-\infty, a], \forall a \in \mathbb{R}\}$.

定义: 概率测度与概率空间

设 (Ω, \mathcal{F}) 为可测空间, $\mathbb{P}: \mathcal{F} \rightarrow \mathbb{R}$ 是一个定义在 \mathcal{F} 上的函数, 若满足

(1) 非负性: $\mathbb{P}(A) \geq 0, \forall A \in \mathcal{F}$

(2) 归一性: $\mathbb{P}(\Omega) = 1$

(3) 可列可加性: 若 $A_i \in \mathcal{F}$, $i = 1, 2, \dots$, 且 $A_i A_j = \emptyset, \forall i \neq j$ 有

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) \quad (1.2)$$

则称 \mathbb{P} 为可测空间 (Ω, \mathcal{F}) 上的一个概率测度, 简称概率. 称 $(\Omega, \mathcal{F}, \mathbb{P})$ 为概率空间, 称 \mathcal{F} 为事件域. 若 $A \in \mathcal{F}$, 则称 A 为随机事件, 简称为事件, 称 $\mathbb{P}(A)$ 为事件 A 的概率.

概率的基本性质

补集概率: $\mathbb{P}(\emptyset) = 0, \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

有限可加性: 若 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$ 互不相容则

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) \quad (1.3)$$

交集和并集概率:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(AB) \quad (1.4)$$

$$\mathbb{P}(A - B) = \mathbb{P}(A) - \mathbb{P}(AB) \quad (1.5)$$

概率比较大小: 若 $A \subset B$, 则 $\mathbb{P}(A) \leq \mathbb{P}(B)$

次可列可加性: 若 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 则

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i) \quad (1.6)$$

可列交并等式: 若 $A_i \in \mathcal{F}, i = 1, 2, \dots, n$, 则

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n \left[(-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) \right] \quad (1.7)$$

定义: 事件列的单调性与极限

一事件列 $\{A_n : n \geq 1\}$ 称为单调增序列, 若 $A_n \subset A_{n+1}, n \geq 1$; 称为单调减序列, 若 $A_n \supset A_{n+1}, n \geq 1$.

如果 $\{A_n : n \geq 1\}$ 是单调增序列, 定义事件列的极限

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{i=1}^{\infty} A_i \quad (1.8)$$

如果 $\{A_n : n \geq 1\}$ 是单调减序列, 定义事件列的极限

$$\lim_{n \rightarrow \infty} A_n = \bigcap_{i=1}^{\infty} A_i \quad (1.9)$$

命题: 连续性定理

若 $\{A_n : n \geq 1\}$ 是单调增或单调减序列, 则极限与概率测度可换序

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}\left(\lim_{n \rightarrow \infty} A_n\right) \quad (1.10)$$

命题: Borel-Cantelli 引理

设 $\{A_n : n \geq 1\}$ 是一事件序列, 若

$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty, \quad (1.11)$$

则

$$\mathbb{P}\left(\limsup_{i \rightarrow \infty} A_i\right) = 0, \quad (1.12)$$

其中

$$\limsup_{i \rightarrow \infty} A_i := \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i. \quad (1.13)$$

定义：事件独立

两个事件 $A, B \in \mathcal{F}$, 若满足

$$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B), \quad (1.14)$$

则称 A 与 B 相互独立.

容易证明下列命题等价:

- (1) A 与 B 独立;
- (2) A 与 B^c 独立;
- (3) $\mathbb{P}(A|B) = \mathbb{P}(A)$;
- (4) $\mathbb{P}(A|B^c) = \mathbb{P}(A)$.

三个事件 $A, B, C \in \mathcal{F}$, 若满足

$$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B), \mathbb{P}(AC) = \mathbb{P}(A)\mathbb{P}(C), \mathbb{P}(BC) = \mathbb{P}(B)\mathbb{P}(C) \quad (1.15)$$

及

$$\mathbb{P}(ABC) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) \quad (1.16)$$

则称 A, B, C 相互独立.

若 A, B, C 相互独立, 则 $A \cup B$ 与 C , AB 与 C , $A - B$ 与 C 相互独立.

n 个事件 $A_1, A_2, \dots, A_n \in \mathcal{F}$, 若对其中任意 k ($2 \leq k \leq n$) 个事件 $A_{i_1}, A_{i_2}, \dots, A_{i_k}$, 其中 $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$, 有

$$\mathbb{P}(A_{i_1}A_{i_2} \cdots A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k}), \quad (1.17)$$

则称 A_1, A_2, \dots, A_n 相互独立.

若 A_1, A_2, \dots, A_n 相互独立, 取 $1 \leq m < n$, 记

$$\mathcal{F}_1 = \sigma(A_k, 1 \leq k \leq m), \mathcal{F}_2 = \sigma(A_k, m+1 \leq k \leq n), \quad (1.18)$$

任取 $B_1 \in \mathcal{F}_1, B_2 \in \mathcal{F}_2$, 则 B_1 与 B_2 独立.

命题：独立事件 Borel-Cantelli 引理

若 $\{A_n : n \geq 1\}$ 是相互独立的事件序列, 且

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty, \quad (1.19)$$

则有

$$\mathbb{P}\left(\limsup_{i \rightarrow \infty} A_i\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{i=n}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \mathbb{P}\left(\sup_{i \geq n} A_i\right) = 1. \quad (1.20)$$

1.2 随机变量、分布函数及数字特征

1.2.1 随机变量与分布函数

定义：随机变量

设 $(\Omega, \mathcal{F}, \mathbb{P})$ 是一概率空间, $X(\omega) : \Omega \rightarrow \mathbb{R}$ 是定义在 Ω 上的单值实函数, 如果对 $\forall a \in \mathbb{R}$, 有

$$\{\omega : X(\omega) \leq a\} \in \mathcal{F}, \quad (1.21)$$

则称 $X(\omega)$ 为随机变量.

常简记 $\{\omega : X(\omega) \leq a\} = \{X \leq a\} = \{X \in (-\infty, a]\}$, $X^{-1}(B) = \{\omega : X(\omega) \in B\}$, 且若 $X(\omega)$ 满足 $\{\omega : X(\omega) \leq a\} \in \mathcal{F}$, 则 $\forall a, b \in \mathbb{R}$, 有 $\{X > a\}, \{X < a\}, \{X = a\}, \{a < X \leq b\}, \{a \leq X < b\}, \{a < X < b\}, \{A \leq X \leq b\} \in \mathcal{F}$.

定义：分布函数

设 X 为 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量, 对 $\forall x \in \mathbb{R}$, 定义

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X \in (-\infty, x]), \quad (1.22)$$

称 $F(x)$ 为 X 的分布函数.

定义：连续型与离散型随机变量

(1) 离散型随机变量: 分布函数为分段右连续的阶梯函数

(2) 连续型随机变量: 分布函数为几乎处处连续的函数

定义：概率密度函数

对随机变量 X 的分布函数 $F(x)$, 若存在一非负函数 $f(x)$, 对 $\forall x \in \mathbb{R}$, 有

$$F(x) = \int_{-\infty}^x f(u) du \quad (1.23)$$

则称 $f(x)$ 为随机变量 X 的概率密度函数.

若 $f(x)$ 连续, 则

$$\frac{dF(x)}{dx} = f(x) \quad (1.24)$$

即

$$\lim_{h \rightarrow 0} \frac{\mathbb{P}(x < X \leq x + h)}{h} = f(x), \quad (1.25)$$

或

$$\mathbb{P}(x < X \leq x + h) = f(x)h + o(h). \quad (1.26)$$

定义：联合分布

二维随机变量 (X, Y) 的联合分布函数 $F(x, y)$ 定义为

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y). \quad (1.27)$$

X 和 Y 的边缘分布分别定义为

$$F_X(x) = \mathbb{P}(X \leq x) = \lim_{y \rightarrow \infty} F(x, y) = F(x, \infty), \quad (1.28)$$

$$F_Y(y) = \mathbb{P}(Y \leq y) = \lim_{x \rightarrow \infty} F(x, y) = F(\infty, y). \quad (1.29)$$

若存在一非负函数 $f(x, y)$, 对 $\forall(x, y) \in \mathbb{R}^2$ 有

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv, \quad (1.30)$$

则称 $f(x, y)$ 为 (X, Y) 的联合概率密度函数.

n 维随机向量 (X_1, X_2, \dots, X_n) 的联合分布函数定义为

$$F(x_1, x_2, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n). \quad (1.31)$$

定理: 分布函数与事件独立的关系

称随机变量 X 与 Y 相互独立, 若对 $\forall(x, y) \in \mathbb{R}^2$, 有

$$F(x, y) = F_X(x)F_Y(y). \quad (1.32)$$

若对 $\forall(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ 有

$$F(x_1, x_2, \dots, x_n) = F_1(x_1)F_2(x_2) \cdots F_n(x_n), \quad (1.33)$$

则称 X_1, X_2, \dots, X_n 相互独立, 这里 $F_i(x_i) = F_{X_i}(x_i) = \mathbb{P}(X_i \leq x_i)$, $\forall i = 1, 2, \dots, n$.

若 X, Y, Z 相互独立, 则 $X \pm Y$ 与 Z 独立, $X \cdot Y$ 与 Z 独立, $X/Y (Y \neq 0)$ 与 Z 独立, 更一般有 $g_1(X, Y)$ 与 $g_2(Z)$ 独立, 其中 $g_1(X, Y)$, $g_2(Z)$ 可以是逐段单调函数或逐段连续函数.

1.2.2 Riemann-Stieltjes 积分

定义: Riemann-Stieltjes 积分

设 $F(x)$ 为 $(-\infty, \infty)$ 上的单调不减右连续函数, $g(x)$ 为 $(-\infty, \infty)$ 上的单值实函数, $\forall a < b$, 任取分点 $a = x_0 < x_1 < x_2 < \cdots < x_{i-1} < x_i < \cdots < x_n = b$, $\forall u_i \in [x_{i-1}, x_i]$, 作积分和式

$$\sum_{i=1}^n g(u_i) \Delta F(x_i) = \sum_{i=1}^n g(u_i) [F(x_i) - F(x_{i-1})]. \quad (1.34)$$

令 $\lambda = \max_{1 \leq i \leq n} \Delta x_i = \max_{1 \leq i \leq n} (x_i - x_{i-1})$, 若极限

$$J(a, b) = \lim_{\lambda \rightarrow 0} \sum_{i=1}^n g(u_i) \Delta F(x_i) \quad (1.35)$$

存在, 则记

$$J(a, b) = \int_a^b g(x) dF(x), \quad (1.36)$$

称极限 $J(a, b)$ 为 $g(x)$ 关于 $F(x)$ 在 $[a, b]$ 上的 Riemann-Stieltjes 积分.

当取 $F(x) = x$ 时, Riemann-Stieltjes 积分化为原来的 Riemann 积分, 所以 Riemann-Stieltjes 积分是 Riemann 积分的推广.

当 $a \rightarrow -\infty, b \rightarrow \infty$ 时, 若极限

$$J(-\infty, \infty) = \lim_{a \rightarrow -\infty, b \rightarrow \infty} \int_a^b g(x) dF(x) \quad (1.37)$$

存在, 则称

$$J(-\infty, \infty) = \int_{-\infty}^{\infty} g(x) dF(x) \quad (1.38)$$

为 $g(x)$ 关于 $F(x)$ 在 $(-\infty, \infty)$ 上的 Riemann-Stieltjes 积分.

Riemann-Stieltjes 积分的基本性质

分段相加: 当 $a = c_0 < c_1 < \cdots < c_n < c_{n+1} = b$ 时,

$$\int_a^b g(x) dF(x) = \sum_{i=0}^n \int_{c_i}^{c_{i+1}} g(x) dF(x) \quad (1.39)$$

对被积函数线性:

$$\int_a^b \sum_{i=1}^n g_i(x) dF(x) = \sum_{i=1}^n \int_a^b g_i(x) dF(x) \quad (1.40)$$

非负性: 若 $g(x) \geq 0$ 且 $a < b$ 则

$$\int_a^b g(x) dF(x) \geq 0 \quad (1.41)$$

对被微函数线性: 若 $F_1(x), F_2(x)$ 为两个分布函数, $c_1, c_2 > 0$ 为常数, 则

$$\int_a^b g(x) d[c_1 F_1(x) + c_2 F_2(x)] = c_1 \int_a^b g(x) dF_1(x) + c_2 \int_a^b g(x) dF_2(x) \quad (1.42)$$

被积函数为1: 若 $g(x) = 1$, 则

$$\int_a^b g(x) dF(x) = \int_a^b 1 dF(x) = F(b) - F(a) = \mathbb{P}(a < X \leq b) \quad (1.43)$$

离散型随机变量: 若 X 为离散型随机变量, 即 $\mathbb{P}(X = c_i) = p_i, i = 1, 2, \dots$, 则

$$F(x) = \sum_{c_i \leq x} p_i \quad (1.44)$$

是一跳跃型分布函数, 即 $F(x)$ 的变化只在 c_1, c_2, \dots 这些点且其跃度为 p_i , 则 Riemann-Stieltjes 积分

$$\int_{-\infty}^{\infty} g(x) dF(x) = \sum_{n=1}^{\infty} g(c_n) [F(c_n + 0) - F(c_n - 0)] = \sum_{n=1}^{\infty} g(c_n) p_n \quad (1.45)$$

化成了一个级数.

1.2.3 数字特征

定义：数学期望

设 X 为随机变量, $F(x)$ 为 X 的分布函数, 若

$$\int_{-\infty}^{\infty} |x| dF(x) \quad (1.46)$$

存在, 则称

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF(x) \quad (1.47)$$

为随机变量 X 的数学期望.

若 X 为非负随机变量, 则有

$$\begin{aligned} \mathbb{E}X &= \int_0^{\infty} x dF(x) \\ &= \int_0^{\infty} \int_0^x 1 dt dF(x) \\ &= \int_0^{\infty} \int_t^{\infty} 1 dF(x) dt \\ &= \int_0^{\infty} \mathbb{P}(X > t) dt \end{aligned} \quad (1.48)$$

同理可知, 对一般随机变量 X 有

$$\begin{aligned} \mathbb{E}X &= \int_0^{\infty} \mathbb{P}(X > t) dt + \int_{-\infty}^0 \mathbb{P}(X \leq t) dt \\ &= \int_0^{\infty} \mathbb{P}(X > t) dt - \int_0^{\infty} \mathbb{P}(X \leq -t) dt \end{aligned} \quad (1.49)$$

期望性质

线性:

$$\mathbb{E} \left(\sum_{i=1}^n c_i X_i \right) = \sum_{i=1}^n c_i \mathbb{E}X_i \quad (1.50)$$

函数复合随机变量:

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) dF(x) \quad (1.51)$$

离散型随机变量: $\mathbb{P}(X = x_n) = p_n, n \in \mathbb{N}$,

$$\mathbb{E}X = \sum_{n=1}^{\infty} x_n p_n \quad (1.52)$$

连续型随机变量: 概率密度函数 $f(x)$

$$\mathbb{E}X = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x) dx \quad (1.53)$$

定义：方差

随机变量 X 的方差为

$$\text{Var}(X) = \sigma_X^2 = DX := \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}X^2 - (\mathbb{E}X)^2. \quad (1.54)$$

定义：协方差

对两个随机变量 (X, Y) 定义协方差为

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)] = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y). \quad (1.55)$$

若 X 与 Y 独立, 则

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y) = (\mathbb{E}X)(\mathbb{E}Y) - (\mathbb{E}X)(\mathbb{E}Y) = 0. \quad (1.56)$$

定义：相关系数

若 $0 < DX = \sigma_X^2 < \infty$, $0 < DY = \sigma_Y^2 < \infty$, 则称

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\text{Cov}(X, Y)}{\sqrt{(DX)(DY)}} \quad (1.57)$$

$\rho(X, Y)$ 刻画了 X, Y 之间线性关系的密切程度, 若 $\rho = 0$, 则称 X, Y 不相关.

定义：矩

随机变量 X 的 $k \geq 1$ 阶矩定义为

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k dF_X(x) \quad (1.58)$$

数字特征性质

方差平方线性且相关:

$$D\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i^2 DX_i + 2 \sum_{i < j} c_i c_j \text{Cov}(X_i, X_j) \quad (1.59)$$

Schwartz 不等式: 若随机变量 X, Y 的二阶矩存在, 则

$$|\mathbb{E}(XY)|^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2) \quad (1.60)$$

1.2.4 常用随机变量的分布

二项分布: $X \sim B(n, p)$, $\mathbb{E}X = np$, $DX = np(1-p)$,

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n \quad (1.61)$$

泊松分布: $X \sim \mathbb{P}(\lambda)$, $\mathbb{E}X = \lambda$, $DX = \lambda$,

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = 0, 1, 2, \dots \quad (1.62)$$

几何分布: $X \sim G(p)$, $\mathbb{E}X = 1/p$, $DX = (1-p)/p^2$,

$$\mathbb{P}(X = k) = (1-p)^{k-1} p, \quad k = 1, 2, \dots \quad (1.63)$$

均匀分布: $X \sim U(a, b)$, $\mathbb{E}X = (a + b)/2$, $DX = (b - a)^2/12$,

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a < x < b \\ 0, & \text{o.w.} \end{cases} \quad (1.64)$$

正态分布: $X \sim N(\mu, \sigma^2)$, $\mathbb{E}X = \mu$, $DX = \sigma^2$, $\mathbb{E}(X - \mu)^{2k} = (2k - 1)!! \sigma^{2k}$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\} \quad (1.65)$$

指数分布: $X \sim E(\lambda)$, $\mathbb{E}X = 1/\lambda$, $DX = 1/\lambda^2$,

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{o.w.} \end{cases} \quad (1.66)$$

gamma 分布: $X \sim \Gamma(\alpha, \lambda)$, $\mathbb{E}X = \alpha/\lambda$, $DX = \alpha/\lambda^2$,

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & \text{if } \alpha > 0 \\ 0, & \text{o.w.} \end{cases} \quad (1.67)$$

这里 gamma 函数为

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1} dt, \quad \forall \alpha > 0 \quad (1.68)$$

beta 分布: $X \sim \mathcal{B}(\alpha, \beta)$, $\mathbb{E}X = \alpha/(\alpha + \beta)$, $DX = \alpha\beta/[(\alpha + \beta + 1)(\alpha + \beta)^2]$,

$$f(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, & \text{if } 0 < x < 1 \\ 0, & \text{o.w.} \end{cases} \quad (1.69)$$

这里 beta 函数为

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx = \frac{\Gamma(\alpha) \cdot \Gamma(\beta)}{\Gamma(\alpha + \beta)} \quad (1.70)$$

1.2.5 示性函数的线性组合

非负随机变量可由事件示性函数的线性组合表示: 设 $X(\omega)$ 为非负随机变量, $\mathbb{P}(X < \infty) = 1$, 令

$$X_n(\omega) = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{I}_{\{\frac{k}{2^n} \leq X < \frac{k+1}{2^n}\}}(\omega) + n \mathbb{I}_{\{X \geq n\}}(\omega) \quad (1.71)$$

则 $X_n(\omega)$ 是随机变量, 且 $\forall \omega \in \Omega$, 有

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \quad (1.72)$$

一般随机变量可由事件示性函数的线性组合表示: 设 $X(\omega)$ 为一般的随机变量, 令

$$X^+ = X \vee 0 = \max(X, 0), \quad X^- = -(X \wedge 0) = -\min(X, 0). \quad (1.73)$$

显然 $X^+, X^- \geq 0$. 由上面的结论, 对 X^+, X^- 存在 $X_n^+ \uparrow X^+, X_n^- \uparrow X^-$, 若令

$$X_n = X_n^+ - X_n^-, \quad (1.74)$$

则 $X_n \uparrow X$.

1.3 矩母函数、特征函数和拉普拉斯变换

1.3.1 矩母函数

定义: 矩母函数或生成函数

随机变量 X 的矩母函数或称生成函数定义为

$$\varphi(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} dF_X(x). \quad (1.75)$$

显然, 如 X 的 k 阶矩存在, 则

$$\mathbb{E}(X^k) = \varphi^{(k)}(0), \quad (1.76)$$

矩母函数由此得名, 可以证明矩母函数与分布函数是一一对应的.

对于取值非负整数的随机变量 X , 即

$$\mathbb{P}(X = k) = p_k \geq 0, \quad k \geq 0, \quad \sum_{k=0}^{\infty} p_k = 1, \quad (1.77)$$

则 X 的矩母函数记为

$$g(s) = \mathbb{E}(s^X) = \sum_{k=0}^{\infty} p_k s^k. \quad (1.78)$$

显然

$$p_k = \frac{g^{(k)}(0)}{k!}, \quad (1.79)$$

且有

$$\mathbb{E}[X(X-1)\cdots(X-k+1)] = g^{(k)}(1). \quad (1.80)$$

特别地,

$$\mathbb{E}(X) = g'(1), \quad \mathbb{E}(X(X-1)) = \mathbb{E}X^2 - \mathbb{E}X = g''(1), \quad (1.81)$$

则有

$$D(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = g''(1) + g'(1) - [g'(1)]^2. \quad (1.82)$$

若 X_1, X_2 相互独立, 其矩母函数分别记为 $g_1(s), g_2(s)$, 则 $X_1 + X_2$ 的矩母函数为

$$g_{X_1+X_2}(s) = g_1(s)g_2(s). \quad (1.83)$$

1.3.2 特征函数

定义：特征函数

随机变量 X 的特征函数定义为

$$\phi(x) := \mathbb{E}[e^{itX}] = \int_{-\infty}^{\infty} e^{itx} dF_X(x). \quad (1.84)$$

1.3.3 Laplace-Stieltjes 变换

定义：Laplace-Stieltjes 变换

设非负随机变量 X , 分布函数 $F_X(x)$, $s = a + ib$, 这里 $a > 0$, b 是实数, 称

$$\hat{F}_X(s) = \int_0^{\infty} e^{-sx} dF_X(x) \quad (1.85)$$

为 $F_X(x)$ 的 Laplace-Stieltjes 变换, 或称随机变量 X 的 Laplace-Stieltjes 变换, 简记 L-S 变换.

$\hat{F}_X(s)$ 与 $F_X(x)$ 也有一一对应关系, 且对 $X_1, X_2 > 0$ 相互独立, 有

$$\hat{F}_{X_1+X_2}(s) = \hat{F}_{X_1}(s)\hat{F}_{X_2}(s). \quad (1.86)$$

1.4 条件数学期望

1.4.1 离散型随机变量

设 (X, Y) 为两个离散型随机变量, 其联合分布律为

$$\mathbb{P}(X = x_i, Y = y_j) = p_{ij} \geq 0, \sum_{ij} p_{ij} = 1, \quad (1.87)$$

若

$$\mathbb{P}(Y = y_j) = \sum_i \mathbb{P}(X = x_i, Y = y_j) = \sum_i p_{ij} > 0, \quad (1.88)$$

则称

$$\mathbb{P}(X = x_i | Y = y_j) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)} \quad (1.89)$$

为给定 $Y = y_j$ 时, X 的条件分布律.

称

$$\mathbb{E}(X | Y = y_j) := \sum_i x_i \mathbb{P}(X = x_i | Y = y_j) \quad (1.90)$$

为给定 $Y = y_j$ 时, X 的条件数学期望.

记

$$\mathbb{E}(X | Y) := \sum_j \mathbb{I}_{\{Y=y_j\}}(\omega) \mathbb{E}(X | Y = y_j), \quad (1.91)$$

称 $\mathbb{E}(X | Y)$ 为 X 关于 Y 的条件数学期望.

随机变量 $\mathbb{E}(X|Y)$ 是随机变量 Y 的函数, 当 $\omega \in \{\omega : Y(\omega) = y_j\}$ 时, $\mathbb{E}(X|Y)$ 的取值为 $\mathbb{E}(X|Y = y_j)$, 其数学期望应为

$$\begin{aligned}
 \mathbb{E}[\mathbb{E}(X|Y)] &= \sum_j \mathbb{E}(X|Y = y_j) \mathbb{P}(Y = y_j) \\
 &= \sum_{i,j} x_i \mathbb{P}(X = x_i | Y = y_j) \mathbb{P}(Y = y_j) \\
 &= \sum_{i,j} x_i \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)} \mathbb{P}(Y = y_j) \\
 &= \sum_{i,j} x_i \mathbb{P}(X = x_i, Y = y_j) \\
 &= \sum_i x_i \mathbb{P}(X = x_i) \\
 &= \mathbb{E}X
 \end{aligned} \tag{1.92}$$

1.4.2 连续型随机变量

设 (X, Y) 的联合概率密度函数为 $f(x, y)$, 若 Y 的概率密度函数

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx > 0 \tag{1.93}$$

则称

$$f_{X|Y=y}(x|y) = \frac{f(x, y)}{f_Y(y)} \tag{1.94}$$

为给定 $Y = y$ 时, X 的条件概率密度函数.

条件分布函数为

$$F_{X|Y=y}(x|y) = \mathbb{P}(X \leq x | Y = y) = \int_{-\infty}^x \frac{f(u, y)}{f_Y(y)} du, \tag{1.95}$$

条件数学期望为

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{\infty} x f_{X|Y=y}(x|y) dx = \int_{-\infty}^{\infty} x \frac{f(x, y)}{f_Y(y)} dx \tag{1.96}$$

考虑 $D \in \mathcal{B}_{\mathbb{R}}$, 给定 $Y \in D$, 若 $\mathbb{P}(Y \in D) > 0$, X 的条件分布函数为

$$\begin{aligned}
 F_{X|Y \in D}(x|D) &= \mathbb{P}(X \leq x | Y \in D) \\
 &= \frac{\mathbb{P}(X \leq x, Y \in D)}{\mathbb{P}(Y \in D)} \\
 &= \frac{\int_{-\infty}^x \int_{y \in D} f(x, y) dy dx}{\int_{y \in D} f_Y(y) dy}
 \end{aligned} \tag{1.97}$$

所以给定 $Y \in D$, X 的条件概率密度函数为

$$f_{X|Y \in D}(x|D) = \frac{\int_{y \in D} f(x, y) dy}{\int_{y \in D} f_Y(y) dy} = \frac{\int_{y \in D} f(x, y) dy}{\mathbb{P}(Y \in D)} \quad (1.98)$$

于是给定 $Y \in D$, X 的条件数学期望为

$$\begin{aligned} \mathbb{E}(X|Y \in D) &= \int_{-\infty}^{\infty} x f_{X|Y \in D}(x|D) dx \\ &= \int_{-\infty}^{\infty} x \frac{\int_{y \in D} f(x, y) dy}{\mathbb{P}(Y \in D)} dx \\ &= \frac{1}{\mathbb{P}(Y \in D)} \int_{-\infty}^{\infty} \int_{y \in D} x f(x, y) dy dx \\ &= \frac{1}{\mathbb{P}(Y \in D)} \int_{y \in D} \int_{-\infty}^{\infty} \frac{x f(x, y)}{f_Y(y)} f_Y(y) dx dy \\ &= \frac{1}{\mathbb{P}(Y \in D)} \int_{y \in D} \mathbb{E}(X|Y = y) f_Y(y) dy \\ &= \mathbb{E}[\mathbb{E}(X|Y)|Y \in D] \end{aligned} \quad (1.99)$$

若取 $D = \mathbb{R}$, 则有

$$\mathbb{E}(X|Y \in \mathbb{R}) = \mathbb{E}(X) = \mathbb{E}[\mathbb{E}(X|Y)|Y \in \mathbb{R}] = \mathbb{E}[\mathbb{E}(X|Y)] \quad (1.100)$$

此即是所谓全期望公式.

上面两点分别是对条件期望取单点值和区间值的要求, 基于此即可定义连续型随机变量条件数学期望.

定义: 连续型随机变量条件数学期望

设 (X, Y) 具有联合概率密度函数 $f(x, y)$, Y 的概率密度函数为 $f_Y(y) > 0$, $\mathbb{E}|X| < \infty$, 若随机变量 $\mathbb{E}(X|Y)$ 满足

$\mathbb{E}(X|Y)$ 是随机变量 Y 的函数, 当 $Y = y$ 时, 它的取值为 $\mathbb{E}(X|Y = y)$

对任意 $D \in \mathcal{B}_{\mathbb{R}}$, 有

$$\mathbb{E}[\mathbb{E}(X|Y)|Y \in D] = \mathbb{E}[X|Y \in D] \quad (1.101)$$

则称随机变量 $\mathbb{E}(X|Y)$ 为 X 关于 Y 的条件数学期望.

1.4.3 一般随机变量

设 (X, Y) 为一般随机变量, 其联合分布函数为 $\mathbb{P}(X \leq x, Y \leq y)$. 以下假设 $\mathbb{E}|X| < \infty$, 分两种情况讨论.

定义: 一般随机变量条件数学期望

设 $D \in \mathcal{B}_{\mathbb{R}}$, $\mathbb{P}(Y \in D) > 0$. $\forall x \in \mathbb{R}$, 称

$$\mathbb{P}(X \leq x | Y \in D) = \frac{\mathbb{P}(X \leq x, Y \in D)}{\mathbb{P}(Y \in D)} \quad (1.102)$$

为 X 关于事件 $\{\omega : Y(\omega) \in D\}$ 的条件分布函数.

称

$$\mathbb{E}(X | Y \in D) = \int_{-\infty}^{\infty} x d\mathbb{P}(X \leq x | Y \in D) \quad (1.103)$$

为 X 关于事件 $\{\omega : Y(\omega) \in D\}$ 的条件数学期望.

在许多问题中常常需要考虑 D 为单点集 $\{y\}$ 的情形. 若 $\mathbb{P}(Y = y) > 0$, 这时定义条件分布同上. 当 $\mathbb{P}(Y = y) = 0$ 时, 定义 $\mathbb{P}(X \leq x | Y = y)$ 如下.

定义: 一般随机变量单点概率为零时条件数学期望

设 $(x, y) \in \mathbb{R}^2$, 对充分小的 $h > 0$, 有 $\mathbb{P}(y < Y \leq y + h) > 0$. 若

$$\mathbb{P}(X \leq x | Y = y) := \lim_{h \rightarrow 0} \mathbb{P}(X \leq x | y < Y \leq y + h) \quad (1.104)$$

存在, 则称 $\mathbb{P}(X \leq x | Y = y)$ 为 X 关于事件 $\{\omega : Y(\omega) = y\}$ 的条件分布函数.

称

$$\mathbb{E}(X | Y = y) = \int_{-\infty}^{\infty} x d\mathbb{P}(X \leq x | Y = y) \quad (1.105)$$

为 X 关于事件 $\{\omega : Y(\omega) = y\}$ 的条件数学期望.

定义: 条件数学期望

若随机变量 $\mathbb{E}(X | Y)$ 满足

(1) $\mathbb{E}(X | Y)$ 是随机变量 Y 的函数, 当 $Y = y$ 时, 它的取值为 $\mathbb{E}(X | Y = y)$;

(2) 对任意 $D \in \mathcal{B}_{\mathbb{R}}$, 有

$$\mathbb{E}[\mathbb{E}(X | Y) | Y \in D] = \mathbb{E}[X | Y \in D]. \quad (1.106)$$

则称随机变量 $\mathbb{E}(X | Y)$ 为 X 关于 Y 的条件数学期望.

定理: 全期望公式

设随机变量 $\mathbb{E}(X | Y)$ 为 X 关于 Y 的条件数学期望, 则有

$$\mathbb{E}X = \mathbb{E}[\mathbb{E}(X | Y)] = \int_{-\infty}^{\infty} \mathbb{E}(X | Y = y) d\mathbb{P}(Y \leq y). \quad (1.107)$$

上式可看作是数学期望形式的全概率公式, 即全期望公式.

1.4.4 条件概率与条件分布函数

条件概率, 条件分布函数均可用条件数学期望的概念及性质来处理.

由示性函数的定义有

$$\mathbb{P}(B) = \mathbb{E}[\mathbb{I}_B(\omega)], \quad (1.108)$$

称

$$\mathbb{E}[\mathbb{I}_B(\omega) | Y] = \mathbb{P}(B | Y) \quad (1.109)$$

为事件 B 关于随机变量 Y 的条件概率, 此时 $\mathbb{P}(B|Y)$ 是随机变量且是 Y 的函数.

对于 $\forall x \in \mathbb{R}$, 取 $B = (\omega : X(\omega) \leq x)$, 称

$$F(x|Y) := \mathbb{P}(X \leq x|Y) = \mathbb{E}[\mathbb{I}_{\{X(\omega) \leq x\}}(\omega)|Y] \quad (1.110)$$

为 X 关于 Y 的条件分布函数.

1.4.5 条件数学期望的基本性质

两个随机变量 Z_1, Z_2 , 如果 $\mathbb{P}(Z_1 = Z_2) = 1$, 则称 Z_1, Z_2 几乎处处 (almost everywhere) 或几乎必然 (almost surely) 相等, 记作 $Z_1 = Z_2$ a.e. 或 a.s.

定理: 条件数学期望的基本性质

设 $X, Y, X_i, 1 \leq i \leq n$ 为随机变量, $g(x), h(y)$ 为一般函数, 且 $\mathbb{E}|X| < \infty, \mathbb{E}|X_i| < \infty, 1 \leq i \leq n, \mathbb{E}|g(X)h(Y)| < \infty, \mathbb{E}|g(X)| < \infty$. 则有

全期望公式:

$$\mathbb{E}X = \mathbb{E}[\mathbb{E}(X|Y)] \quad (1.111)$$

线性: $\forall \alpha_i \in \mathbb{R}, 1 \leq i \leq n$, 有

$$\mathbb{E}\left(\sum_{i=1}^n \alpha_i X_i | Y\right) = \sum_{i=1}^n \alpha_i \mathbb{E}(X_i | Y) \quad (1.112)$$

条件期望平滑性:

$$\mathbb{E}[g(X)h(Y)|Y] = h(Y)\mathbb{E}[g(X)|Y] \quad (1.113)$$

特别地, 有

$$\mathbb{E}(X|X) = X \quad (1.114)$$

$$\mathbb{E}[g(X)h(Y)] = \mathbb{E}\{\mathbb{E}[g(X)h(Y)|Y]\} = \mathbb{E}\{h(Y)\mathbb{E}[g(X)|Y]\} \quad (1.115)$$

独立性: 若 X 与 Y 相互独立, 则

$$\mathbb{E}(X|Y) = \mathbb{E}X \quad (1.116)$$

1.4.6 多元随机变量的条件数学期望

离散型随机变量

设三个随机变量 (X, Y, Z) , 其中 (Y, Z) 为离散型随机变量, 称随机变量 $\mathbb{E}(X|Y, Z)$ 是 X 关于 Y, Z 的条件数学期望, 若它满足

(1) $\mathbb{E}(X|Y, Z)$ 是 (Y, Z) 的二元函数, 当 $Y = y_j, Z = z_k$ 时, $\mathbb{E}(X|Y, Z)$ 的取值为 $\mathbb{E}(X|Y = y_j, Z = z_k)$

(2) 对任意 $D_j \in \mathcal{B}_{\mathbb{R}}, D_k \in \mathcal{B}_{\mathbb{R}}$, 有

$$\mathbb{E}[\mathbb{E}(X|Y, Z)|Y \in D_j, Z \in D_k] = \mathbb{E}(X|Y \in D_j, Z \in D_k) \quad (1.117)$$

用示性函数表示, 即

$$\mathbb{E}(X|Y, Z) := \sum_{j,k} \mathbb{I}_{\{Y(\omega)=y_j, Z(\omega)=z_k\}}(\omega) \mathbb{E}(X|Y=y_j, Z=z_k) \quad (1.118)$$

当 $\mathbb{E}|X| < \infty$ 时, 由对 Z 的全期望公式和条件期望平滑性有

$$\mathbb{E}[\mathbb{E}(X|Y, Z)|Y] = \mathbb{E}(X|Y) = \mathbb{E}[\mathbb{E}(X|Y)|Y, Z] \quad (1.119)$$

连续型随机变量

如 (X, Y, Z) 为连续型随机变量, 联合概率密度函数为 $f(x, y, z)$, (Y, Z) 的联合概率密度函数为 $f_{Y,Z}(y, z)$, X 关于 $Y = y, Z = z$ 的条件概率密度函数为

$$f_{X|(Y,Z)=(y,z)}(x|y, z) = \frac{f(x, y, z)}{f_{Y,Z}(y, z)}, \quad (1.120)$$

设 $\mathbb{E}|X| < \infty$, $f_{Y,Z}(y, z) > 0$, 若随机变量 $\mathbb{E}(X|Y, Z)$ 满足

$\mathbb{E}(X|Y, Z)$ 是 (Y, Z) 的二元函数, 当 $Y = y, Z = z$ 时, $\mathbb{E}(X|Y, Z)$ 的取值为 $\mathbb{E}(X|Y = y, Z = z)$

对任意 $D_1 \in \mathcal{B}_{\mathbb{R}}, D_2 \in \mathcal{B}_{\mathbb{R}}$, 有

$$\mathbb{E}[\mathbb{E}(X|Y, Z)|Y \in D_1, Z \in D_2] = \mathbb{E}(X|Y \in D_1, Z \in D_2) \quad (1.121)$$

称随机变量 $\mathbb{E}(X|Y, Z)$ 是 X 关于 Y, Z 的条件数学期望.

定理: 多元随机变量条件期望的性质

设以下所涉及的期望全都有有限, 则有

全期望公式:

$$\mathbb{E}X = \mathbb{E}[\mathbb{E}(X|Y_1, Y_2, \dots, Y_n)] \quad (1.122)$$

线性: $\forall \alpha_i \in \mathbb{R}, 1 \leq i \leq n$, 有

$$\mathbb{E}\left(\sum_{i=1}^n \alpha_i X_i \middle| Y_1, Y_2, \dots, Y_n\right) = \sum_{i=1}^n \alpha_i \mathbb{E}(X_i|Y_1, Y_2, \dots, Y_n) \quad (1.123)$$

条件期望平滑性:

$$\mathbb{E}[g(X)h(Y_1, Y_2, \dots, Y_n)|Y_1, Y_2, \dots, Y_n] = h(Y_1, Y_2, \dots, Y_n)\mathbb{E}[g(X)|Y_1, Y_2, \dots, Y_n] \quad (1.124)$$

独立性: 若 X 与 Y_1, Y_2, \dots, Y_n 独立, 则

$$\mathbb{E}(X|Y_1, Y_2, \dots, Y_n) = \mathbb{E}X \quad (1.125)$$

1.4.7 条件乘法公式与条件独立性

下面在条件概率测度下推广原本的概率乘法公式和事件独立性.

条件概率的乘法公式

设 A, B 为两个随机事件, 由条件概率的定义可知

$$\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B|A) \quad (1.126)$$

与上面的概率乘法公式类似, 条件概率测度 $\mathbb{P}(\cdot|A)$ 的乘法公式如下.

定义: 条件概率测度下的乘法公式

设 A, B, C 为随机事件, 则

$$\mathbb{P}(BC|A) = \mathbb{P}(B|A)\mathbb{P}(C|AB) \quad (1.127)$$

条件独立性

当两个随机事件 A, B 独立时, 有 $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$, 即 $\mathbb{P}(A|B) = \mathbb{P}(A)$. 与上面的独立性概念类似, 条件独立性的定义如下.

定义: 条件独立

设 A, B, C 为随机事件, 称事件 A, B 关于事件 C 条件独立, 即在条件概率测度 $\mathbb{P}(\cdot|C)$ 下独立, 若满足

$$\mathbb{P}(A|BC) = \mathbb{P}(A|C). \quad (1.128)$$

命题: 条件独立的充要条件

设 A, B, C 为随机事件, 则事件 A, B 关于事件 C 条件独立的充要条件为

$$\mathbb{P}(AB|C) = \mathbb{P}(A|C)\mathbb{P}(B|C), \quad (1.129)$$

此即事件独立定义在条件概率测度 $\mathbb{P}(\cdot|C)$ 下的自然推广.

1.4.8 全概率公式

设 $\{B_k\}$, $k = 1, 2, \dots$ 为 Ω 的一个分割, 则有全概率公式

$$\mathbb{P}(A) = \sum_{k=1}^{\infty} \mathbb{P}(B_k)\mathbb{P}(A|B_k). \quad (1.130)$$

1.5 随机过程的概念

在概率论中, 研究了随机变量, n 维随机向量. 在极限定理中, 涉及到了无穷多个随机变量, 但局限在它们之间是相互独立的情形. 将上述情形加以推广, 即研究一族无穷多个、相互有关的随机变量, 这就是随机过程.

1.5.1 随机过程的定义

定义: 随机过程

设 T 是一个指标集, 如 $T = \mathbb{Z}, \mathbb{Z}^+, \mathbb{R}, \mathbb{R}^+, [0, t]$ 等. $\forall t \in T$, X_t 是定义在概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上取值与 S 的随机变量, 则称 $X = \{X_t : t \in T\}$ 是定义在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上以 S 为状态空间的随机过程.

当 $T = \mathbb{Z}, \mathbb{Z}^+$ 或其子集时, 称 X 是离散参数随机过程, 当 $T = \mathbb{R}, \mathbb{R}^+$ 或其子区间时, 称 X 是连续参数随机过程.

若 S 是有限集或可列无穷集时, 称 X 是离散状态的, 若 S 是连续流, 则称 X 是连续状态的.

有时记 $X_t(\omega) = X(t, \omega) : T \times \Omega \rightarrow S$, 即 $X(\cdot, \cdot)$ 为 $T \times \Omega$ 上的二元单值函数.

固定 $t \in T$, 函数 $X_t(\omega) : \Omega \rightarrow S$ 是定义在样本空间 Ω 上的函数, 即为一随机变量.

固定 $\omega \in \Omega$, 函数 $X_t(\omega) : T \rightarrow S$ 称为 X 的一条样本轨道.

1.5.2 随机过程的分布

随机过程的概率特性由其有限维分布族决定. 设 $S = \mathbb{R}, \forall n \geq 1, t_1, t_2, \dots, t_n \in T$, 记

$$F(t_1, t_2, \dots, t_n; x_1, x_2, \dots, x_n) = \mathbb{P}(X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n) \quad (1.131)$$

为 $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ 的联合分布函数, 其全体

$$\{F(t_1, t_2, \dots, t_n; x_1, x_2, \dots, x_n) : t_1, t_2, \dots, t_n \in T, n \geq 1\} \quad (1.132)$$

称为 $X = \{X_t : t \in T\}$ 的有限维分布族, 具有以下性质.

对称性: 对 $(1, 2, \dots, n)$ 的任一排列 (j_1, j_2, \dots, j_n) 有

$$F(t_{j_1}, t_{j_2}, \dots, t_{j_n}; x_{j_1}, x_{j_2}, \dots, x_{j_n}) = F(t_1, t_2, \dots, t_n; x_1, x_2, \dots, x_n) \quad (1.133)$$

即任意排列均不会改变分布函数

相容性: 对 $\forall m < n$ 有边缘分布

$$F(t_1, \dots, t_m, t_{m+1}, \dots, t_n; x_1, \dots, x_m, \infty, \dots, \infty) = F(t_1, t_2, \dots, t_m; x_1, x_2, \dots, x_m) \quad (1.134)$$

即有限维分布族也可像普通分布函数那样求边缘分布

1.5.3 随机过程的分类

设 $X = \{X_t : t \in T\}$ 为随机过程, 按其概率特征分类如下.

点过程或计数过程一个随机过程 $\{N(A), A \subset T\}$ 是点过程, 若 $N(A)$ 表示在集合 A 中事件发生的总数, 即它满足

对 $\forall A \subset T, N(A)$ 是一取值非负整数的随机变量, $N(\emptyset) = 0$

对 $\forall A_1, A_2 \subset T$, 若 $A_1 A_2 = \emptyset$, 则对每一个样本有

$$N(A_1 \cup A_2) = N(A_1) + N(A_2) \quad (1.135)$$

参数集 T 可以是 \mathbb{R}^n , 也可以是任意一抽象非空集

泊松过程是简单的点过程

独立与平稳增量过程

对 $t_1 < t_2 < \cdots < t_n$, $t_i \in T$, $1 \leq i \leq n$, 若增量

$$X_{t_1}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}} \quad (1.136)$$

相互独立, 则称 X 为独立增量过程

若 $\forall 0 \leq s < t$, 增量 $X_t - X_s$ 的分布只依赖于 $t - s$, 则称 X 有平稳增量

有平稳增量的独立增量过程简称为独立平稳增量过程

泊松过程和维纳过程或称布朗运动是两个最简单也是最重要的独立平稳增量过程

马尔可夫过程与扩散过程

对 $t_1 < t_2 < \cdots < t_n < t$, $t_i \in T$, x_i , $1 \leq i \leq n$, 及 $A \subset \mathbb{R}$, 若总有

$$\mathbb{P}(X_t \in A | X_{t_1} = x_1, X_{t_2} = x_2, \cdots, X_{t_n} = x_n) = \mathbb{P}(X_t \in A | X_{t_n} = x_n) \quad (1.137)$$

则称此过程为马尔可夫过程

离散状态马尔可夫过程称为马尔可夫链

连续时间参数且连续状态的马尔可夫过程称为扩散过程

泊松过程是一个最简单连续时间参数马尔可夫链

布朗运动是一个最简单的扩散过程

鞅

若 $\forall t \in T$, $\mathbb{E}|X_t| < \infty$, 且对 $\forall t_1 < t_2 < \cdots < t_n < t_{n+1}$ 有

$$\mathbb{E}(X_{t_{n+1}} | X_{t_1}, X_{t_2}, \cdots, X_{t_n}) = X_{t_n} \quad (1.138)$$

则称 X 为鞅.

布朗运动关于自身是鞅

Itô 积分关于布朗运动是鞅

二阶矩过程

若对 $\forall t \in T$, $\mathbb{E}|X_t|^2 < \infty$, 则称 X 为二阶矩过程

均方分析中针对的均为二阶矩过程

更新过程

设 $\{X_k : k \geq 1\}$ 为独立同分布的正的随机变量序列, 对 $t > 0$, 令 $S_0 = 0, S_n = \sum_{k=1}^n X_k$, 并定义

$$N_t = \max\{n : n \geq 0, S_n \leq t\}, \quad (1.139)$$

称 $\{N_t : t \geq 0\}$ 为更新过程

更新过程可以解释为 $[0, t]$ 内更换零件的个数, 或系统来的信号数, 或服务站的顾客数等

宽平稳过程或协方差平稳过程

若对 $\forall \tau, t \in Y$, $DX_t < \infty$, 且

$$\mathbb{E}X_t = m, \quad \text{Cov}(X_t, X_{t+\tau}) = R(\tau) \quad (1.140)$$

仅依赖 τ , 则称 X 为宽平稳过程, 即它的协方差不随时间推移而改变
 严平稳过程

若对 $\forall t_1, t_2, \dots, t_n \in T$, 及 $h > 0$,

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}) \quad (1.141)$$

则称该过程为严平稳过程

严平稳过程的一切有限维分布对时间的推移保持不变

特别地, (X_t, X_s) 的二维分布只依赖于 $(t - s)$

第2章 泊松过程

2.1 定义及其背景

定义：时齐泊松过程

随机过程 $\{N_t : t \geq 0\}$ 称为时齐泊松过程, 若满足
计数过程, 且初值为 0

$$N_0 = 0 \quad (2.1)$$

独立增量过程, 即 $\forall 0 < t_1 < t_2 < \cdots < t_n$,

$$N_{t_1}, N_{t_2} - N_{t_1}, \cdots, N_{t_n} - N_{t_{n-1}} \quad (2.2)$$

相互独立

增量平稳性, 即 $\forall s, t \geq 0, n \geq 0$,

$$\mathbb{P}(N_{s+t} - N_s = n) = \mathbb{P}(N_t = n) \quad (2.3)$$

增 1 同 λ 阶

$$\begin{aligned} \mathbb{P}(N_{t+\Delta t} - N_t = 1) &= \lambda \Delta t + o(\Delta t) \\ \mathbb{P}(N_{t+\Delta t} - N_t = 2) &= o(\Delta t) \end{aligned} \quad (2.4)$$

定理：泊松过程的分布

若 $\{N_t : t \geq 0\}$ 为泊松过程, 则 $\forall s, t \geq 0$, 有

$$\mathbb{P}(N_{s+t} - N_s = k) = \mathbb{P}(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad (2.5)$$

即 $N_{s+t} - N_s \stackrel{d}{=} N_t \sim \mathbb{P}(\lambda t)$.

定义：泊松过程

计数过程 $\{N_t : t \geq 0\}$ 称为参数是 λ 的泊松过程, 若满足
初值为 0

$$N_0 = 0 \quad (2.6)$$

独立增量过程

增量是参数为 λt 的泊松分布, 即 $\forall s, t \geq 0$,

$$N_{s+t} - N_s \sim \mathbb{P}(\lambda t) \quad (2.7)$$

2.2 相邻事件的时间间隔

设 $\{N_t : t \geq 0\}$ 是计数过程, N_t 是 $[0, t]$ 内事件发生次数. 令 $S_0 = 0$, 第 n 次事件发生的时刻为

$$S_n = \inf_{t \geq 0} \{t : N_t \geq n\} = \inf \{t : t > S_{n-1}, N_t = n\}, \quad (2.8)$$

第 $n-1$ 个事件与第 n 个事件发生的时间间隔为

$$X_n = S_n - S_{n-1}, \quad n \geq 1. \quad (2.9)$$

固定 $\omega \in \Omega$, 观察随机过程的一条轨道, 即一条时间轴, 可知此时存在事件等价

$$\{t : N_t \geq n\} = \{t : t \geq S_n\} = \{t : S_n \leq t\} \quad (2.10)$$

$$\{t : N_t = n\} = \{t : S_n \leq t < S_{n+1}\} = \{t : S_n \leq t\} \setminus \{t : S_{n+1} \leq t\} \quad (2.11)$$

因此 S_n 的分布函数为

$$\mathbb{P}(S_n \leq t) = \mathbb{P}(N_t \geq n) = 1 - \mathbb{P}(N_t \leq n-1) = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad t \geq 0 \quad (2.12)$$

则 S_n 的概率密度函数为

$$\begin{aligned} f_{S_n}(t) &= -\sum_{k=1}^{n-1} \frac{k\lambda(\lambda t)^{k-1}}{k!} e^{-\lambda t} + \lambda \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t} \\ &= \left[\sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} - \sum_{k=0}^{n-2} \frac{(\lambda t)^k}{k!} \right] \lambda e^{-\lambda t} \\ &= \frac{\lambda(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}, \quad t \geq 0 \end{aligned} \quad (2.13)$$

即

$$S_n \sim \Gamma(n, \lambda). \quad (2.14)$$

定理: 计数过程 $\{N_t : t \geq 0\}$ 是泊松过程的充分必要条件是 $\{X_n : n \geq 1\}$ 是独立且参数同 λ 的指数分布.

例

求 (S_1, S_2) 的联合概率密度, 并证明 $S_1, S_2 - S_1$ 独立.

解: 令 $0 < t_1 < t_2$, 取充分小的 $h > 0$, 使得

$$t_1 - \frac{h}{2} < t_1 < t_1 + \frac{h}{2} < t_2 - \frac{h}{2} < t_2 < t_2 + \frac{h}{2}, \quad (2.15)$$

由

$$\begin{aligned} &\left\{ t_1 - \frac{h}{2} < S_1 \leq t_1 + \frac{h}{2} < t_2 - \frac{h}{2} < S_2 \leq t_2 + \frac{h}{2} \right\} \\ &= \left\{ N\left(t_1 - \frac{h}{2}\right) = 0, N\left(t_1 + \frac{h}{2}\right) - N\left(t_1 - \frac{h}{2}\right) = 1, \right. \\ &\quad \left. N\left(t_2 - \frac{h}{2}\right) - N\left(t_1 + \frac{h}{2}\right) = 0, N\left(t_2 + \frac{h}{2}\right) - N\left(t_2 - \frac{h}{2}\right) = 1 \right\} \cup H_n, \end{aligned} \quad (2.16)$$

其中

$$H_n = \left\{ N\left(t_1 - \frac{h}{2}\right) = 0, N\left(t_1 + \frac{h}{2}\right) - N\left(t_1 - \frac{h}{2}\right) = 1, \right. \\ \left. N\left(t_2 - \frac{h}{2}\right) - N\left(t_1 + \frac{h}{2}\right) = 0, N\left(t_2 + \frac{h}{2}\right) - N\left(t_2 - \frac{h}{2}\right) \geq 2 \right\}, \quad (2.17)$$

得

$$\mathbb{P}\left(t_1 - \frac{h}{2} < S_1 \leq t_1 + \frac{h}{2} < t_2 - \frac{h}{2} < S_2 \leq t_2 + \frac{h}{2}\right) = (\lambda h)^2 e^{-\lambda(t_2 + h/2)} + o(h^2) \\ = \lambda^2 h^2 e^{-\lambda t_2} + o(h^2), \quad (2.18)$$

所以 (S_1, S_2) 的联合概率密度为

$$f_{(S_1, S_2)}(t_1, t_2) = \frac{\partial^2 P}{\partial h^2} = \begin{cases} \lambda^2 e^{-\lambda t_2}, & \text{if } 0 < t_1 < t_2, \\ 0, & \text{o.w.} \end{cases} \quad (2.19)$$

为证明 X_1, X_2 的独立性, 下面求 (X_1, X_2) 的联合概率密度.

注意到 $X_1 = S_1$, $X_2 = S_2 - S_1$, 令 $x_1 = t_1$, $x_2 = t_2 - t_1$, 则此变换的雅可比行列式为

$$J = \left| \frac{\partial[t_1(x_1, x_2), t_2(x_1, x_2)]}{\partial(x_1, x_2)} \right| = \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} = 1, \quad (2.20)$$

此处可以直观地理解为全微分相等

$$g_{(X_1, X_2)}(x_1, x_2) \partial(x_1, x_2) = f_{(S_1, S_2)}(t_1, t_2) \partial(t_1, t_2), \quad (2.21)$$

则相互转换需要乘以雅可比行列式作为转换系数, 即

$$g_{(X_1, X_2)}(x_1, x_2) = f_{(S_1, S_2)}(t_1, t_2) \frac{\partial(t_1, t_2)}{\partial(x_1, x_2)}. \quad (2.22)$$

于是 X_1, X_2 的联合概率密度为

$$g_{(X_1, X_2)}(x_1, x_2) = f_{(S_1, S_2)}[t_1(x_1, x_2), t_2(x_1, x_2)] \cdot J \\ = \begin{cases} \lambda^2 e^{-\lambda(x_1 + x_2)}, & \text{if } x_1, x_2 \geq 0, \\ 0, & \text{o.w.} \end{cases} \quad (2.23)$$

对联合概率密度求边缘分布可知 X_1, X_2 的概率密度分别为

$$g_{X_1}(x_1) = \lambda e^{-\lambda x_1}, \quad g_{X_2}(x_2) = \lambda e^{-\lambda x_2}, \quad (2.24)$$

所以

$$g_{(X_1, X_2)}(x_1, x_2) = g_{X_1}(x_1) g_{X_2}(x_2), \quad (2.25)$$

故 X_1, X_2 独立, 即 $S_1, S_2 - S_1$ 独立.

2.3 剩余寿命与年龄

设 N_t 表示在 $[0, t]$ 上事件发生的个数, S_n 表示第 n 个事件发生的时刻, 那么

S_{N_t} 表示在 t 时刻前最后一个事件发生的时刻

S_{N_t+1} 表示 t 时刻后首次事件发生的时刻

令

$$W_t = S_{N_t+1} - t, \quad V_t = t - S_{N_t}, \quad (2.26)$$

则

年龄: $V_t = t - S_{N_t} \in [0, t]$

剩余寿命: $W_t = S_{N_t+1} - t \geq 0$

定理: 剩余寿命与年龄的分布

设 $\{N_t : t \geq 0\} \sim PP(\lambda)$ 是参数为 λ 的泊松过程, 则

剩余寿命分布: W_t 与 $\{X_n : n \geq 1\}$ 同分布, 即

$$\begin{aligned} \mathbb{P}(W_t \leq x) &= 1 - \mathbb{P}(W_t > x) \\ &= 1 - \mathbb{P}(N_{t+x} - N_t = 0) \\ &= 1 - \mathbb{P}(N_x = 0) \\ &= 1 - \frac{(\lambda x)^0}{0!} e^{-\lambda x} \\ &= 1 - e^{-\lambda x}, \quad x \geq 0 \end{aligned} \quad (2.27)$$

年龄分布: V_t 是截尾的指数分布, 即对 $x < t$ 有

$$\begin{aligned} \mathbb{P}(V_t \leq x) &= 1 - \mathbb{P}(V_t > x) \\ &= 1 - \mathbb{P}(N_t - N_{t-x} = 0) \\ &= 1 - \mathbb{P}(N_x = 0) \\ &= 1 - \frac{(\lambda x)^0}{0!} e^{-\lambda x} \\ &= 1 - e^{-\lambda x}, \quad 0 \leq x < t \end{aligned} \quad (2.28)$$

故

$$\mathbb{P}(V_t \leq x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } 0 \leq x < t \\ 1, & \text{if } x \geq t \end{cases} \quad (2.29)$$

由于 V_t 取单点 t 的概率大于零, 即分布函数有跳跃, 则其是混合型随机变量.

定理: 若 $\{X_n : n \geq 1\}$ 独立同分布, 又对 $\forall t \geq 0$, W_t 与 $\{X_n : n \geq 1\}$ 同分布, 分布函数为 $F(x)$, 且 $F(0) = 0$, 则 $\{N_t : t \geq 0\}$ 为泊松过程.

2.4 到达时间的条件分布

本节讨论在给定 $N_t = n$ 的条件下, S_1, S_2, \dots, S_n 的条件分布, 有关性质及其应用.

定理: 设 $\{N_t : t \geq 0\}$ 是泊松过程, 则对 $\forall 0 < s < t$, 有

$$\mathbb{P}(X_1 \leq s | N_t = 1) = \frac{s}{t}. \quad (2.30)$$

证明:

$$\begin{aligned} \mathbb{P}(X_1 \leq s | N_t = 1) &= \frac{\mathbb{P}(X_1 \leq s, N_t = 1)}{\mathbb{P}(N_t = 1)} \\ &= \frac{\mathbb{P}(N_s = 1, N_t - N_s = 0)}{\mathbb{P}(N_t = 1)} \\ &= \frac{\mathbb{P}(N_s = 1)\mathbb{P}(N_t - N_s = 0)}{\mathbb{P}(N_t = 1)} \\ &= \frac{\frac{(\lambda s)^1}{1!} e^{-\lambda s} \frac{(\lambda(t-s))^0}{0!} e^{-\lambda(t-s)}}{\frac{(\lambda t)^1}{1!} e^{-\lambda t}} \\ &= \frac{s}{t}. \end{aligned} \quad (2.31)$$

独立同分布非负随机变量的顺序统计量的联合概率密度为

$$f(y_1, y_2, \dots, y_n) = \begin{cases} n! \prod_{i=1}^n f(y_i), & \text{if } 0 < y_1 < y_2 < \dots < y_n, \\ 0, & \text{o.w.} \end{cases} \quad (2.32)$$

若为均匀分布, 则

$$f(y_1, y_2, \dots, y_n) = \begin{cases} \frac{n!}{t^n}, & \text{if } 0 < y_1 < y_2 < \dots < y_n \leq t, \\ 0, & \text{o.w.} \end{cases} \quad (2.33)$$

定理: 设 $\{N_t : t \geq 0\}$ 是泊松过程, 在给定 $N_t = n$ 时, 事件相继发生的时间 S_1, S_2, \dots, S_n 的条件概率密度为

$$f(t_1, t_2, \dots, t_n) = \begin{cases} \frac{n!}{t^n}, & \text{if } 0 < t_1 < t_2 < \dots < t_n \leq t, \\ 0, & \text{o.w.} \end{cases} \quad (2.34)$$

本定理说明在给定 $N_t = n$ 的条件下, S_1, S_2, \dots, S_n 的条件分布函数与 n 个在 $[0, t]$ 上相互独立同均匀分布的顺序统计量的分布函数相同.

定理: 设 $\{N_t : t \geq 0\}$ 是计数过程, X_n 为第 n 个时间与第 $n-1$ 个时间的时间间隔, $\{X_n : n \geq 1\}$ 独立同分布且

$$F(x) = \mathbb{P}(X_n \leq x), \quad (2.35)$$

若 $F(0) = 0$, 且对 $\forall 0 \leq s \leq t$, 有

$$\mathbb{P}(X_1 \leq s | N_t = 1) = \frac{s}{t}, \quad (2.36)$$

则 $\{N_t : t \geq 0\}$ 是泊松过程.

定理: 设 $\{N_t : t \geq 0\}$ 是计数过程, X_n 为第 n 个时间与第 $n-1$ 个时间的时间间隔, $\{X_n : n \geq 1\}$ 独立同分布且

$$F(x) = \mathbb{P}(X_n \leq x), \quad (2.37)$$

若 $\mathbb{E}X_n < \infty$, $F(0) = 0$, 且对 $\forall 0 \leq s \leq t$, $n \geq 1$, 有

$$\mathbb{P}(S_n \leq s | N_t = n) = \left(\frac{s}{t}\right)^n, \quad (2.38)$$

则 $\{N_t : t \geq 0\}$ 是泊松过程.

定理: 设 $\{N_t : t \geq 0\} \sim PP(\lambda)$ 是参数为 λ 的泊松过程, S_k , $k \geq 1$ 为其到达时刻, 则对任意 $[0, \infty)$ 上的可积函数 f 有

$$\mathbb{E}[f(S_n)] = \lambda \int_0^\infty f(t) dt. \quad (2.39)$$

2.5 复合泊松过程

2.6 更新过程

2.7 若干极限定理与基本更新定理

第3章 马尔可夫链

3.1 定义与例子

定义： 马尔可夫链

若随机序列 $\{X_n : n \geq 0\}$ 对任意 $i_0, i_1, \dots, i_n, i_{n+1} \in S$, $n \in \mathbb{N}_0$ 及 $\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) > 0$, 有

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n) \quad (3.1)$$

则称其为马尔可夫链 (Markov chain).

注释：

马尔可夫链的定义式揭示了马尔可夫链的特性, 即马尔可夫性或无后效性: 将来 X_{n+1} 与过去 $X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}$ 关于现在 $X_n = i_n$ 条件独立.

定义： 一步转移概率与转移矩阵

$\forall i, j \in S$, $\mathbb{P}(X_{n+1} = j | X_n = i) := p_{ij}(n)$ 称为时刻 n 从 i 到 j 的一步转移概率. 若 $\forall i, j \in S$, $p_{ij}(n) = p_{ij}$ 不依赖于 n , 则称 $X = \{X_n, n \geq 0\}$ 为时齐马氏链 (HMC, Homogeneous Markov Chain). 记 $\mathbf{P} = (p_{ij})_{i,j \in S}$, 则称 \mathbf{P} 为 X 的一步转移概率矩阵, 简称为转移矩阵 (transition matrix).

定义： 转移图

转移图是一个有向图 $G = (V, E)$, $V = S$, $E = \{\vec{ij} \mid i, j \in S, p_{ij} > 0\}$.

例： 随机游动

独立同分布随机变量序列 $\{\xi_n, n \geq 1\}$ 取整数值, 整数值随机变量 X_0 与 $\{\xi_n, n \geq 1\}$ 独立,

$$X_n = X_0 + \sum_{k=1}^n \xi_k, \quad \forall n \geq 1 \quad (3.2)$$

则 $\{X_n, n \geq 0\}$ 是时齐马氏链.

证明: $\forall n \in \mathbb{N}$, $i_0, i_1, \dots, i_{n-1}, i, j \in \mathbb{Z}$,

$$\begin{aligned} & \mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= \mathbb{P}(X_n + \xi_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= \mathbb{P}(i + \xi_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= \mathbb{P}(\xi_{n+1} = j - i) \end{aligned} \quad (3.3)$$

另一方面

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(i + \xi_{n+1} = j | X_n = i) = \mathbb{P}(\xi_{n+1} = j - i) \quad (3.4)$$

则

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i) \quad (3.5)$$

所以 $\{X_n, n \geq 0\}$ 是时齐马氏链.

定理: 由当前事件和独立序列生成的马氏链

独立同分布随机变量序列 $\{\xi_n, n \geq 1\}$ 均取值于 S , $f: S \times S \rightarrow S$, $X_n = f(X_{n-1}, \xi_n)$, X_0 与 $\{\xi_n, n \geq 1\}$ 相互独立, 则 $X = \{X_n, n \geq 0\}$ 是时齐马氏链, 转移概率 $p_{ij} = \mathbb{P}(f(i, \xi_1) = j)$.

证明: $\forall i_0, i_1, \dots, i_{n-1}, i, j \in S$,

$$\begin{aligned} & \mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= \mathbb{P}(f(X_n, \xi_{n+1}) = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= \mathbb{P}(f(i, \xi_{n+1}) = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) \\ &= \mathbb{P}(f(i, \xi_{n+1}) = j) \\ &= \mathbb{P}(f(i, \xi_1) = j) \end{aligned} \quad (3.6)$$

另一方面

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(f(i, \xi_{n+1}) = j | X_n = i) = \mathbb{P}(f(i, \xi_{n+1}) = j) \quad (3.7)$$

则

$$\mathbb{P}(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = j | X_n = i) \quad (3.8)$$

所以 $\{X_n, n \geq 0\}$ 是时齐马氏链.

3.2 转移概率矩阵

设 $X = \{X_n, n \geq 0\}$ 是时齐马氏链, 一步转移概率矩阵 $\mathbf{P} = (p_{ij})_{i,j \in S}$, $p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i)$, 则显然有 $p_{ij} \geq 0$, $\sum_{j \in S} p_{ij} = 1$, 这两条来自条件概率测度 $\mathbb{P}(\cdot | X_n = i)$ 的基本性质.

定义: 转移矩阵

称矩阵 $\mathbf{A} = (a_{ij})_{i,j \in S}$ 为转移矩阵, 若 $a_{ij} \geq 0$, 且

$$\sum_{j \in S} a_{ij} = 1. \quad (3.9)$$

记 $\pi_i(n) = \mathbb{P}(X_n = i)$,

$$\boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \dots, \pi_i(n), \dots) \quad (3.10)$$

表示 n 时刻 X_n 的概率分布向量, $X_0 \sim \boldsymbol{\pi}(0)$ 称为 X 的初始分布.

定理： 概率分布的全概率公式

$$\pi(n+1) = \pi(n)\mathbf{P}, \quad \pi(n) = \pi(0)\mathbf{P}^n \quad (3.11)$$

证明：由全概率公式可知

$$\begin{aligned} \mathbb{P}(X_{n+1} = j) &= \sum_{i \in S} \mathbb{P}(X_n = i) \mathbb{P}(X_{n+1} = j | X_n = i) \\ &= \sum_{i \in S} \pi_i(n) p_{ij} \end{aligned} \quad (3.12)$$

写成向量形式为

$$\pi(n+1) = \pi(n)\mathbf{P} \quad (3.13)$$

反复应用则有

$$\pi(n) = \pi(0)\mathbf{P}^n \quad (3.14)$$

注释：

一个马尔可夫链的特性完全由它的一步转移概率矩阵 \mathbf{P} 及初始分布向量 $\pi(0)$ 决定.

定理： Kolmogorov-Chapman 方程多步转移的全概率公式

记 $p_{ij}^{(m)} = \mathbb{P}(X_{n+m} = j | X_n = i)$ 为 m 步转移概率, $\mathbf{P}^{(m)} = \{p_{ij}^{(m)}\}$ 为 m 步转移概率矩阵, 则

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}, \quad \mathbf{P}^{(m+n)} = \mathbf{P}^{(m)} \mathbf{P}^{(n)}, \quad \mathbf{P}^{(n)} = \mathbf{P}^n \quad (3.15)$$

证明：由条件概率形式的全概率公式可知

$$\begin{aligned} p_{ij}^{(m+n)} &= \mathbb{P}(X_{m+n} = j | X_0 = i) \\ &= \sum_{k \in S} \mathbb{P}(X_m = k | X_0 = i) \mathbb{P}(X_{m+n} = j | X_0 = i, X_m = k) \\ &= \sum_{k \in S} \mathbb{P}(X_m = k | X_0 = i) \mathbb{P}(X_{m+n} = j | X_m = k) \\ &= \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)} \end{aligned} \quad (3.16)$$

写成向量形式为

$$\mathbf{P}^{(m+n)} = \mathbf{P}^{(m)} \mathbf{P}^{(n)} \quad (3.17)$$

注释：

一个马尔可夫链运动规律的概率特性取决于它的转移概率矩阵特性.

3.3 状态的分类

定义： 吸收态可达与互通

吸收态: 称 $i \in S$ 为吸收态, 若 $p_{ii} = 1$, 即到达 i 之后就不再移动了

可达: 若 $i, j \in S$, 存在 $n \geq 0$ 使得 $p_{ij}^{(n)} > 0$, 则称 i 可达 j , 记为 $i \rightarrow j$, 即转移图中存在从 i 到 j 的通路

互通: 若 $i \rightarrow j$ 且 $j \rightarrow i$, 则称 i 与 j 互通, 记为 $i \leftrightarrow j$

定理: 互通为等价关系

状态相通关系为等价关系因其满足

自反性: $i \leftrightarrow i$

对称性: 若 $i \leftrightarrow j$, 则 $j \leftrightarrow i$

传递性: 若 $i \leftrightarrow j$ 且 $j \leftrightarrow k$, 则 $i \leftrightarrow k$

注释:

利用等价关系, 可以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. 然而, 从某一类出发以正的概率到达另一类的情形是可能的. 如一马尔可夫链的所有状态属于同一等价类, 则称它是不可约链.

定义: 首达时间与首达概率

首达时间: 从 i 出发首次到达 j 的时间定义为

$$T_{ij} = \min_{n \geq 1} \{n : X_n = j, X_0 = i\} \quad (3.18)$$

若右侧为空集, 则令 $T_{ij} = \infty$

首达概率: 从 i 出发经 n 步首次到达 j 的概率定义为

$$f_{ij}^{(n)} = \mathbb{P}(T_{ij} = n | X_0 = i) = \mathbb{P}(X_n = j, X_k \neq j, 1 \leq k \leq n-1 | X_0 = i) \quad (3.19)$$

有限步首达概率: 从 i 出发经有限步首次到达 j 的概率定义为

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} \quad (3.20)$$

定义: 常返与非常返

常返状态: 称状态 i 为常返状态, 若 $f_{ii} = 1$

非常返状态: 称状态 i 为非常返状态, 若 $f_{ii} < 1$

定义: 正频率常返与零频率常返

当 $f_{ii} = 1$ 时, $\{f_{ii}^{(n)} : n \geq 1\}$ 是一个概率分布. 记 μ_i 表示从 i 出发再回到 i 的平均回转时间 $\mathbb{E}T_{ii}$, 即

$$\mu_i = \mathbb{E}T_{ii} = \sum_{n=1}^{\infty} n \mathbb{P}(T_{ii} = n | X_0 = i) = \sum_{n=1}^{\infty} n f_{ii}^{(n)}. \quad (3.21)$$

正常返状态: 称状态 i 为正常返状态, 若 $\mu_i = \mathbb{E}T_{ii} < \infty$, 此时平均返回频率

$$\nu_i = \frac{1}{\mu_i} = \frac{1}{\mathbb{E}T_{ii}} > 0 \quad (3.22)$$

零常返状态: 称状态 i 为零常返状态, 若 $\mu_i = \mathbb{E}T_{ii} = \infty$, 此时平均返回频率

$$\nu_i = \frac{1}{\mu_i} = \frac{1}{\mathbb{E}T_{ii}} = 0 \quad (3.23)$$

定义：周期与遍历

周期：若返回步数集合

$$\{n : n \geq 1, p_{ii}^{(n)} > 0\} \neq \emptyset, \quad (3.24)$$

则称该集合的最大公约数 $d(i)$ 为状态 i 的周期

周期状态: $d(i) > 1$

非周期状态: $d(i) = 1$

遍历状态: 状态 i 为正常返状态且非周期

定理：首达全概率公式

对 $\forall i, j \in S, n \geq 1$, 有

n 步转移概率

$$p_{ij}^{(n)} = \sum_{l=1}^n f_{ij}^{(l)} p_{jj}^{(n-l)} \quad (3.25)$$

n 步首达概率

$$f_{ij}^{(n)} = \sum_{k \neq j} p_{ik} f_{kj}^{(n-1)} \mathbb{I}_{\{n>1\}} + p_{ij} \mathbb{I}_{\{n=1\}} \quad (3.26)$$

可达与互通的有限步首达概率表示

$$i \rightarrow j \Leftrightarrow f_{ij} > 0, \quad i \leftrightarrow j \Leftrightarrow f_{ij} f_{ji} > 0 \quad (3.27)$$

定理：常返与非常返的充要条件

状态 i 为常返状态, $f_{ii} = 1$, 当且仅当

$$G_{ii} := \sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty, \quad (3.28)$$

状态 i 为非常返状态, $f_{ii} < 1$, 当且仅当

$$G_{ii} = \sum_{n=0}^{\infty} p_{ii}^{(n)} = \frac{1}{1 - f_{ii}} < \infty. \quad (3.29)$$

证明: 约定 $p_{ii}^{(0)} = 1, f_{ii}^{(0)} = 0$. 记生成函数

$$\mathbb{P}(\rho) = \sum_{n=0}^{\infty} p_{ii}^{(n)} \rho^n, \quad F(\rho) = \sum_{n=0}^{\infty} f_{ii}^{(n)} \rho^n, \quad (3.30)$$

通过交换求和次序可知

$$\mathbb{P}(\rho) - 1 = \mathbb{P}(\rho) F(\rho), \quad (3.31)$$

即

$$\mathbb{P}(\rho) = \frac{1}{1 - F(\rho)}, \quad 0 \leq \rho < 1, \quad (3.32)$$

两边令 $\rho \uparrow 1$ 可得

$$G_{ii} = \frac{1}{1 - f_{ii}}, \quad (3.33)$$

故由常返和非常返定义可得结论.

定理: 正常返与零常返的充要条件

设 i 为常返状态, 则有

状态 i 为零常返状态, $\mu_i = \mathbb{E}T_{ii} = \infty$, 当且仅当

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0, \quad (3.34)$$

状态 i 为正常返状态且非周期, 即遍历态, $\mu_i = \mathbb{E}T_{ii} < \infty$, 当且仅当

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = \frac{1}{\mu_i} > 0. \quad (3.35)$$

定理: 常返可达必常返

如果 i 为常返状态, 且 $i \rightarrow j$, 则 j 必为常返状态, 且

$$f_{ji} = 1. \quad (3.36)$$

定理: 状态互通必相同

若 $i \leftrightarrow j$, 则

i 与 j 同为常返状态或非常返状态. 若为常返状态, 则它们同为正常返状态或同为零常返状态

i 与 j 或有相同的周期, 或同为非周期

3.4 状态空间的分解

定义: 闭集

设 $C \subset S$, 若对 $\forall i \in C, j \notin C$, 有

$$p_{ij} = 0, \quad (3.37)$$

则称 C 为闭集, 即闭集之内不可能走出去. 若闭集 C 的状态相通, 则称 C 为不可约的.

引理: 闭集的充要条件

C 是闭集的充要条件为: 对 $\forall i \in C, j \notin C, n \geq 1$, 有

$$p_{ij}^{(n)} = 0. \quad (3.38)$$

注释:

吸收态为单点集构成一个闭集

整个状态空间也构成一个闭集

所有常返状态构成闭集, 由常返可达必常返易得

不可约马尔可夫链或者没有非常返状态或者没有常返状态

定理: 常返闭集分解定理

设所有常返状态构成的闭集 $C \neq \emptyset$, 则它可以分解为若干个互不相交的闭集 $\{C_n\}$, 使得

$$C = C_1 \cup C_2 \cup \dots \quad (3.39)$$

且 $\{C_n\}$ 均为互不相通不可约闭集.

推论: 状态空间分解定理

状态空间 S 可分解为

$$S = T \cup C = T \cup C_1 \cup C_2 \cup \dots, \quad (3.40)$$

其中 $\{C_n\}$ 为基本常返闭集, 非常返状态构成的集合 T 不一定是闭集.

注释: 有限状态空间情形

若状态空间有限, 则非常返状态构成的集合 T 一定不是闭集

有限不可约马尔可夫链的状态都是常返状态

引理: 闭集生成的转移矩阵

设 $C_h \subset S$ 为闭集, 只考虑 C_h 上所得的 m 步转移矩阵

$$P^{(m)} = [p_{ij}^{(m)}], \quad i, j \in C_h, \quad (3.41)$$

所以它们也是转移矩阵.

3.5 转移矩阵的极限性态

3.5.1 转移矩阵的极限性态

定理: 非常返或零常返的极限性态

若 j 为非常返或零常返状态, 则对 $\forall i \in S$, 有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0. \quad (3.42)$$

推论:

有限马尔可夫链没有零常返状态

不可约的有限马尔可夫链的状态都是正常返状态

若马尔可夫链有一零常返状态, 则必有无限多个零常返状态

定理: 正常返的极限性态

若 j 为正常返状态, 则对 $\forall i \in S$ 及 $0 \leq r \leq d-1$, 有

$$\lim_{n \rightarrow \infty} p_{ij}^{(nd+r)} = f_{ij}(r) \frac{d}{\mu_j}, \quad f_{ij}(r) = \sum_{m=0}^{\infty} f_{ij}^{(md+r)}. \quad (3.43)$$

推论:

若 j 为遍历状态, 则对 $\forall i \in S$, 有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{f_{ij}}{\mu_j} \quad (3.44)$$

不可约遍历链对 $\forall i, j \in S$, 有

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \frac{1}{\mu_j} \quad (3.45)$$

若 j 为常返状态, 则对 $\forall i \in S$, 有

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n p_{ij}^{(l)} = \frac{f_{ij}}{\mu_j}. \quad (3.46)$$

不可约常返链对 $\forall i, j \in S$, 有

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n p_{ij}^{(l)} = \frac{1}{\mu_j}. \quad (3.47)$$

定理:

对不可约遍历链, $\left\{ \pi_i = \frac{1}{\mu_i} \right\}$ 是方程组

$$x_j = \sum_{i \in S} x_i p_{ij} \quad (3.48)$$

满足条件 $x_j \geq 0, j \in S, \sum_{j \in S} x_j = 1$ 的唯一解.

3.5.2 平稳分布

定义: 平稳分布

一个定义在 S 上的概率分布 $\pi = \{\pi_1, \pi_2, \dots\}$ 称为马尔可夫链的平稳分布, 若

$$\pi = \pi P, \quad (3.49)$$

即 $\forall j \in S$, 有

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}. \quad (3.50)$$

定理: 平稳过程的充要条件

设 $\{X_n : n \geq 0\}$ 是马尔可夫链, 则其是平稳过程的充要条件是

$$\pi(0) = (\pi_i(0), i \in S) \quad (3.51)$$

是平稳分布, 即

$$\pi(0) = \pi(0)P. \quad (3.52)$$

定理: 不可约遍历链平稳分布

不可约遍历链恒有唯一的平稳分布 $\left\{ \pi_i = \frac{1}{\mu_i} \right\}$, 且

$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)}. \quad (3.53)$$

定理: 平稳分布存在定理令 C^+ 为马尔可夫链中全体正常返状态构成的集合, 则有

平稳分布不存在的充要条件为 $C^+ = \emptyset$

平稳分布唯一存在的充要条件为只有一个基本正常返闭集 $C_a = C^+$

有限状态马尔可夫链的平稳分布总存在

有限不可约非周期马尔可夫链存在唯一的平稳分布

3.5.3 极限分布

定义：极限分布

若马尔可夫链概率分布的极限

$$\lim_{n \rightarrow \infty} \pi_j(n) = \pi_j^*, \quad j \in S \quad (3.54)$$

存在, 则称

$$\boldsymbol{\pi}^* = \{\pi_1^*, \pi_2^*, \dots\} \quad (3.55)$$

为马尔可夫链的极限分布.

定理：

非周期不可约链是正常返的充要条件是它存在平稳分布, 且此时平稳分布就是极限分布.

例：平稳分布与极限分布

设 X 为时齐 Markov 链, 其一步转移概率为

$$\mathbf{P} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix} \quad (3.56)$$

求不变分布 $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)$

设

$$\mathbb{P}(X_0 = 1) = \frac{1}{3}, \mathbb{P}(X_0 = 2) = \frac{1}{3}, \mathbb{P}(X_0 = 3) = \frac{1}{3}, \mathbb{P}(X_0 = 4) = 0 \quad (3.57)$$

求 $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = i), i = 1, 2, 3, 4$

解:

由 $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$ 可得

$$(\pi_1, \pi_2, \pi_3, \pi_4) = (\pi_1, \pi_2, \pi_3, \pi_4) \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/3 & 0 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 & 1/3 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}, \quad (3.58)$$

又

$$\sum_{i=1}^4 \pi_i = 1, \quad (3.59)$$

所以可求得

$$\boldsymbol{\pi} = \left(\frac{4}{14}, \frac{3}{14}, \frac{3}{14}, \frac{4}{14} \right). \quad (3.60)$$

故有

$$\lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{\pi} \\ \boldsymbol{\pi} \\ \boldsymbol{\pi} \end{bmatrix}, \quad (3.61)$$

又由题给条件有

$$\pi(0) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right), \quad (3.62)$$

所以

$$\begin{aligned} \lim_{n \rightarrow \infty} \pi(n) &= \lim_{n \rightarrow \infty} \pi(0) \mathbf{P}^n \\ &= \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0\right) \begin{bmatrix} 4/14 & 3/14 & 3/14 & 4/14 \\ 4/14 & 3/14 & 3/14 & 4/14 \\ 4/14 & 3/14 & 3/14 & 4/14 \\ 4/14 & 3/14 & 3/14 & 4/14 \end{bmatrix} \\ &= \left(\frac{4}{14}, \frac{3}{14}, \frac{3}{14}, \frac{4}{14}\right), \end{aligned} \quad (3.63)$$

故

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = i) = \pi_i, \quad i = 1, 2, 3, 4. \quad (3.64)$$

例：首达概率与正常返

取值非负整数的 Markov 链 X , 其一步转移概率为

$$p_{n,n+1} = p \in (0, 1), \quad p_{n,0} = 1 - p, \quad n = 0, 1, 2, \dots, \quad (3.65)$$

其他元素为 0, 设

$$T_0 = \inf\{n > 0 : X_n = 0\}. \quad (3.66)$$

求 $\mathbb{P}(T_0 = n | X_0 = 0)$

该马氏链是否为正常返链, 说明理由

解: 由转移图可知

$$\mathbb{P}(T_0 = n | X_0 = 0) = f_{00}^{(n)} = p^{n-1}(1-p), \quad (3.67)$$

所以 $T_0 |_{X_0=0}$ 服从几何分布, 即

$$T_0 |_{X_0=0} \sim G(1-p) \quad (3.68)$$

由几何分布的相关公式可得

$$f_{00} = \sum_{n=1}^{\infty} f_{00}^{(n)} = \sum_{n=1}^{\infty} p^{n-1}(1-p) = 1, \quad (3.69)$$

且

$$\mu_0 = \sum_{n=1}^{\infty} n f_{00}^{(n)} = \sum_{n=1}^{\infty} n p^{n-1}(1-p) = \frac{1}{1-p} < \infty, \quad (3.70)$$

所以 0 状态正常返, 又此马氏链是不可约的, 故整个马氏链正常返.

第4章 离散鞅引论

4.1 定义与例子

定义：关于自己是鞅

随机过程 $\{X_n : n \geq 0\}$ 是鞅, 若 $\forall n \geq 0$, 有

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

$$\mathbb{E}(X_{n+1}|X_0, X_1, \dots, X_n) = X_n$$

定义：关于另一个过程是鞅

设有两个过程, $\{X_n : n \geq 0\}$ 和 $\{Y_n : n \geq 0\}$, 称 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, 若

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

$$\mathbb{E}(X_{n+1}|Y_0, Y_1, \dots, Y_n) = X_n$$

由全期望公式易知鞅在任何时刻的期望均相等,

$$\mathbb{E}X_{n+1} = \mathbb{E}[\mathbb{E}(X_{n+1}|Y_0, Y_1, \dots, Y_n)] = \mathbb{E}X_n = \mathbb{E}X_0. \quad (4.1)$$

例：似然比构成的鞅

设 $Y_0, Y_1, \dots, Y_n, \dots$ 是独立同分布随机变量序列, f_0 和 f_1 是概率密度函数, 令

$$X_n = \frac{f_1(Y_0)f_1(Y_1) \cdots f_1(Y_n)}{f_0(Y_0)f_0(Y_1) \cdots f_0(Y_n)}, \quad n \geq 0. \quad (4.2)$$

假设 $\forall y \in \mathbb{R}, f_0(y) > 0$, 且 Y_n 的概率密度函数为 f_0 , 则 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅.

证明: 首先计算单个比值的期望, $\forall i = 0, 1, \dots$, 有

$$\mathbb{E} \left[\frac{f_1(Y_i)}{f_0(Y_i)} \right] = \int_{-\infty}^{\infty} \frac{f_1(y)}{f_0(y)} f_0(y) dy = 1, \quad (4.3)$$

所以同理由独立同分布可证连乘的绝对值期望有限

$$\mathbb{E}|X_n| = \mathbb{E} \left[\frac{f_1(Y_0)f_1(Y_1) \cdots f_1(Y_n)}{f_0(Y_0)f_0(Y_1) \cdots f_0(Y_n)} \right] = \prod_{i=0}^n \mathbb{E} \left[\frac{f_1(Y_i)}{f_0(Y_i)} \right] = 1 < \infty, \quad (4.4)$$

其次有

$$\begin{aligned} \mathbb{E}(X_{n+1}|Y_0, Y_1, \dots, Y_n) &= \mathbb{E} \left[X_n \frac{f_1(Y_{n+1})}{f_0(Y_{n+1})} | Y_0, Y_1, \dots, Y_n \right] \\ &= X_n \mathbb{E} \left[\frac{f_1(Y_{n+1})}{f_0(Y_{n+1})} \right] \\ &= X_n, \end{aligned} \quad (4.5)$$

因此 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅.

例: Doob 鞅过程

设 $Y_0, Y_1, \dots, Y_n, \dots$ 是任一随机变量序列, 有随机变量 X 满足 $\mathbb{E}|X| < \infty$, 令

$$X_n = \mathbb{E}(X|Y_0, Y_1, \dots, Y_n), \quad (4.6)$$

$\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, 并称之为 Doob 过程.

证明: 易知

$$\mathbb{E}|X_n| = \mathbb{E}[\mathbb{E}(X|Y_0, Y_1, \dots, Y_n)] \leq \mathbb{E}[\mathbb{E}(|X||Y_0, Y_1, \dots, Y_n)] = \mathbb{E}|X| < \infty, \quad (4.7)$$

又由多元随机变量的全期望公式可得

$$\begin{aligned} \mathbb{E}(X_{n+1}|Y_0, Y_1, \dots, Y_n) &= \mathbb{E}[\mathbb{E}(X|Y_0, Y_1, \dots, Y_{n+1})|Y_0, Y_1, \dots, Y_n] \\ &= \mathbb{E}(X|Y_0, Y_1, \dots, Y_n) \\ &= X_n, \end{aligned} \quad (4.8)$$

因此 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅.

4.2 上下鞅及分解定理

定义: 上鞅

设 $\{X_n : n \geq 0\}$ 与 $\{Y_n : n \geq 0\}$ 是随机过程, 称 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是上鞅, 若

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

$$\mathbb{E}(X_{n+1}|Y_0, Y_1, \dots, Y_n) \leq X_n$$

X_n 是 Y_0, Y_1, \dots, Y_n 的函数

定义: 下鞅

设 $\{X_n : n \geq 0\}$ 与 $\{Y_n : n \geq 0\}$ 是随机过程, 称 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是下鞅, 若

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

$$\mathbb{E}(X_{n+1}|Y_0, Y_1, \dots, Y_n) \geq X_n$$

X_n 是 Y_0, Y_1, \dots, Y_n 的函数

定理: 期望 Jensen 不等式

设 $\phi(x)$ 为凸函数, 则由 Jensen 不等式可知

$$\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}X). \quad (4.9)$$

引理: 凸函数构造下鞅

若 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, $\phi(x)$ 为凸函数且 $\forall n \geq 0, \mathbb{E}|\phi(X_n)| < \infty$, 则 $\{\phi(X_n) : n \geq 0\}$ 是关于 $\{Y_n : n \geq 0\}$ 的下鞅.

推论:

若 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, 且 $\forall n \geq 0, \mathbb{E}|X_n|^2 < \infty$, 则 $\{|X_n| : n \geq 0\}$ 与 $\{X_n^2 : n \geq 0\}$ 是关于 $\{Y_n : n \geq 0\}$ 的下鞅.

上下鞅的基本性质

若 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是上鞅, 则

$$\mathbb{E}(X_{n+k} | Y_0, Y_1, \dots, Y_n) \leq X_n, \quad \forall k \geq 0 \quad (4.10)$$

若 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是上鞅, 则

$$\mathbb{E}X_n \leq \mathbb{E}X_k \leq \mathbb{E}X_0, \quad \forall 0 \leq k \leq n \quad (4.11)$$

若 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是上鞅, g 是关于 Y_0, Y_1, \dots, Y_n 的非负函数, 则

$$\mathbb{E}[g(Y_0, Y_1, \dots, Y_n)X_{n+k} | Y_0, Y_1, \dots, Y_n] \leq g(Y_0, Y_1, \dots, Y_n)X_n, \quad \forall k \geq 0 \quad (4.12)$$

定理: 鞅分解定理

若 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是下鞅, 则必存在过程 $\{M_n : n \geq 0\}$ 与 $\{Z_n : n \geq 0\}$, 使得 $\{M_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅 Z_n 是 Y_0, Y_1, \dots, Y_{n-1} 的函数, 且 $Z_0 = 0, Z_n \leq Z_{n+1}, \mathbb{E}Z_n < \infty, \forall n \geq 1$

$$X_n = M_n + Z_n, \quad \forall n \geq 0$$

由本定理可知, 一个下鞅总可分解为一个鞅与一增过程之和.

4.3 停时与停时定理

定义: 停时

设取值为非负整数的随机变量 T 及随机序列 $\{Y_n : n \geq 0\}$, $\mathcal{F}_n = \sigma(Y_0, Y_1, \dots, Y_n)$, 若对 $\forall n \geq 0$, 有

$$\{\omega : T(\omega) = n\} \in \mathcal{F}_n \quad (4.13)$$

则称 T 是关于 $\{Y_n : n \geq 0\}$ 的停时.

注释: 停时

设取值为非负整数的随机变量 T 及随机序列 $\{Y_n : n \geq 0\}$, 若对 $\forall n \geq 0$, 事件 $\{\omega : T(\omega) = n\}$ 的示性函数 $\mathbb{I}_{\{T(\omega)=n\}}(\omega)$ 仅是 Y_0, Y_1, \dots, Y_n 的函数, 则称 T 是关于 $\{Y_n : n \geq 0\}$ 的停时.

定义中要求的事件是 $\{T = n\}$, 实际上也可以用 $\{T \leq n\}, \{T < n\}, \{T \geq n\}, \{T > n\}$ 代替.

$\mathbb{I}_{\{T(\omega)=n\}}(\omega)$ 仅是 Y_0, Y_1, \dots, Y_n 的函数则在计算条件期望的时候可以直接提出.

停时性质

$T = k, k \in \mathbb{N}$ 是停时, 即常数是停时

设 T_1, T_2 是关于 $\{Y_n : n \geq 0\}$ 的两个停时, 则

$$T_1 + T_2, \quad T_1 \wedge T_2, \quad T_1 \vee T_2 \quad (4.14)$$

均是停时

引理: 降序号

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, 则 $\forall n \geq k$, 有

$$\mathbb{E}(X_n \mathbb{I}_{\{T=k\}}) = \mathbb{E}(X_k \mathbb{I}_{\{T=k\}}). \quad (4.15)$$

引理: 取小停时的期望

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, 则 $\forall n \geq 1$, 有

$$\mathbb{E}X_0 = \mathbb{E}X_{T \wedge n} = \mathbb{E}X_n. \quad (4.16)$$

引理:

设 X 是随机变量满足 $\mathbb{E}|X| < \infty$, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, 且 $\mathbb{P}(T < \infty) = 1$, 则

$$\lim_{n \rightarrow \infty} \mathbb{E}(X \mathbb{I}_{\{T > n\}}) = 0, \quad \lim_{n \rightarrow \infty} \mathbb{E}(X \mathbb{I}_{\{T \leq n\}}) = 1 \quad (4.17)$$

定理: 停时定理

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, $\mathbb{P}(T < \infty) = 1$, 且

$$\mathbb{E} \left(\sup_{n \geq 0} |X_{T \wedge n}| \right) < \infty, \quad (4.18)$$

所以

$$\mathbb{E}X_T = \mathbb{E}X_0. \quad (4.19)$$

推论: 停时定理

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, 若 $\mathbb{E}T < \infty$, 且存在常数 $b < \infty$ 满足对 $\forall n < T$, 有

$$\mathbb{E}(|X_{n+1} - X_n| | Y_0, Y_1, \dots, Y_n) \leq b, \quad (4.20)$$

所以

$$\mathbb{E}X_T = \mathbb{E}X_0. \quad (4.21)$$

定理: 停时定理

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, 若

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

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

$$\lim_{n \rightarrow \infty} \mathbb{E}|X_n \mathbb{I}_{\{T > n\}}| = 0 \text{ 所以}$$

$$\mathbb{E}X_T = \mathbb{E}X_0. \quad (4.22)$$

推论： 停时定理

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, T 是关于 $\{Y_n : n \geq 0\}$ 的停时, 若

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

对某个 $k < \infty$, $\forall n \geq 0$, $\mathbb{E}(X_{T \wedge n}^2) \leq k$ 所以

$$\mathbb{E}X_T = \mathbb{E}X_0. \quad (4.23)$$

推论： 停时定理

设 $Y_0 = 0$, $\{Y_k : k \geq 1\}$ 独立同分布, $\mathbb{E}Y_k = \mu$, $DY_k = \sigma^2 < \infty$, $S_0 = 0$, $S_n = \sum_{k=1}^n Y_k$, $X_n = S_n - n\mu$. 若 T 为停时, $\mathbb{E}T < \infty$, 则 $\mathbb{E}|X_T| < \infty$, 且

$$\mathbb{E}X_T = \mathbb{E}S_T - \mu\mathbb{E}T = 0. \quad (4.24)$$

4.4 鞅收敛定理

引理： 上穿不等式

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是下鞅, $V^{(n)}(a, b)$ 表示 $\{X_k : 0 \leq k < n\}$ 上穿区间 (a, b) 的次数, $a < b$, 则

$$\mathbb{E}[V^{(n)}(a, b)] \leq \frac{\mathbb{E}(X_n - a)^+ - \mathbb{E}(X_0 - a)^+}{b - a} \leq \frac{\mathbb{E}X_n^+ + |a|}{b - a}, \quad (4.25)$$

这里记 $a^+ = \max\{a, 0\} = a \vee 0$.

定理： 鞅收敛定理

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是下鞅,

$$\sup_n \mathbb{E}|X_n| < \infty, \quad (4.26)$$

所以存在随机变量 X_∞ , 使得

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = X_\infty\right) = 1, \quad (4.27)$$

且 $\mathbb{E}|X_\infty| < \infty$.

证明: 由于

$$\mathbb{E}X_n^+ \leq \mathbb{E}|X_n| \leq 2\mathbb{E}X_n^+ - \mathbb{E}X_n, \quad (4.28)$$

故

$$\sup_n \mathbb{E}|X_n| < \infty \Leftrightarrow \sup_n \mathbb{E}X_n^+ < \infty. \quad (4.29)$$

当 $n \rightarrow \infty$ 时, $V^{(n)}(a, b) \rightarrow V(a, b)$, 即 X_n 上穿 (a, b) 的次数, 故

$$\begin{aligned} \mathbb{E}[V(a, b)] &= \mathbb{E} \left[\lim_{n \rightarrow \infty} V^{(n)}(a, b) \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} [V^{(n)}(a, b)] \\ &\leq \lim_{n \rightarrow \infty} \frac{\mathbb{E} X_n^+ + |a|}{b - a} \\ &\leq \frac{\sup_n \mathbb{E} X_n^+ + |a|}{b - a} \\ &< \infty. \end{aligned} \quad (4.30)$$

因此

$$\mathbb{P}(V(a, b) < \infty) = 1, \quad (4.31)$$

即当 $n \rightarrow \infty$ 时, $X_n(\omega)$ 以概率1存在极限, 设

$$\lim_{n \rightarrow \infty} X_n = X_\infty, \quad (4.32)$$

所以

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} X_n = X_\infty \right) = 1. \quad (4.33)$$

另外, 由 Fatou 引理

$$\mathbb{E}|X_\infty| = \mathbb{E} \left(\lim_{n \rightarrow \infty} |X_n| \right) \leq \lim_{n \rightarrow \infty} \mathbb{E}|X_n| \leq \sup_n \mathbb{E}|X_n| < \infty, \quad (4.34)$$

即

$$\mathbb{E}|X_\infty| < \infty. \quad (4.35)$$

定理: Chebyshev 不等式与 Kolmogorov 不等式

设 $\{Y_k : k \geq 0\}$ 独立同分布, $\mathbb{E}Y_k = 0$, $\mathbb{E}Y_k^2 = \sigma^2 < \infty$, $X_0 = 0$, $X_n = \sum_{k=1}^n Y_k$, 则由 Chebyshev 不等式有

$$\varepsilon^2 \mathbb{P}(|X_n| > \varepsilon) \leq n\sigma^2, \quad (4.36)$$

由 Kolmogorov 不等式有

$$\varepsilon^2 \mathbb{P} \left(\max_{0 \leq k \leq n} |X_k| > \varepsilon \right) \leq n\sigma^2. \quad (4.37)$$

由

$$\mathbb{E}X_n^2 = \mathbb{E} \left(\sum_{k=1}^n Y_k \right)^2 = \sum_{k=1}^n \mathbb{E}Y_k^2 = n\sigma^2 \quad (4.38)$$

可知

$$\varepsilon^2 \mathbb{P} \left(\max_{0 \leq k \leq n} |X_k| > \varepsilon \right) = \varepsilon^2 \mathbb{P} \left(\max_{0 \leq k \leq n} X_k^2 > \varepsilon^2 \right) \leq \mathbb{E}X_n^2. \quad (4.39)$$

定理: 最大值不等式

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是下鞅, 且 $\forall n \geq 0$ 有 $X_n \geq 0$, 则对 $\forall \lambda > 0$ 有

$$\lambda \mathbb{P} \left(\max_{0 \leq k \leq n} X_k > \lambda \right) \leq \mathbb{E} X_n. \quad (4.40)$$

推论:

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, 则对 $\forall \lambda > 0$ 有

$$\lambda \mathbb{P} \left(\max_{0 \leq k \leq n} |X_k| > \lambda \right) \leq \mathbb{E} |X_n|. \quad (4.41)$$

证明: 易知 $\forall n \geq 0$ 有 $|X_n| \geq 0$ 且是下鞅, 则由最大值不等式立即可得.

定理:

设 $\{X_n : n \geq 0\}$ 关于 $\{Y_n : n \geq 0\}$ 是鞅, 且存在常数 k 使得 $\forall n \geq 0$ 有 $\mathbb{E} X_n^2 \leq k < \infty$, 则存在有限随机变量 X_∞ , 使得 $\{X_n : n \geq 0\}$ 不仅依概率收敛

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} X_n = X_\infty \right) = 1, \quad (4.42)$$

而且均方收敛

$$\lim_{n \rightarrow \infty} \mathbb{E} |X_n - X_\infty|^2 = 0. \quad (4.43)$$

例: 最大值不等式

设 $\forall n \geq 1$, $\mathbb{E} X_n^2 \leq k < \infty$. 令 $S_n = \sum_{k=1}^n X_k$, 已知 $\{S_n : n \geq 1\}$ 是鞅, 证明 $\forall \varepsilon > 0$, 有

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{S_n}{n} \right| > \varepsilon \right) = 0. \quad (4.44)$$

证明: 易知 $X_i = S_i - S_{i-1}$, 则对 $\forall i < j$ 由全期望公式, 条件期望平滑性以及鞅的降序性有

$$\begin{aligned} \mathbb{E}(X_i X_j) &= \mathbb{E}[(S_i - S_{i-1})(S_j - S_{j-1})] \\ &= \mathbb{E}(S_i S_j - S_{i-1} S_j - S_i S_{j-1} + S_{i-1} S_{j-1}) \\ &= \mathbb{E}[\mathbb{E}(S_i S_j - S_{i-1} S_j - S_i S_{j-1} + S_{i-1} S_{j-1} | S_1, S_2, \dots, S_i)] \\ &= \mathbb{E}(S_i^2 - S_{i-1} S_i - S_i^2 + S_{i-1} S_i) \\ &= 0, \end{aligned} \quad (4.45)$$

故由题给不等式有

$$\mathbb{E} S_n^2 = \mathbb{E} \left(\sum_{k=1}^n X_k \right)^2 = \sum_{k=1}^n \mathbb{E} X_k^2 + 2 \sum_{i < j} \mathbb{E}(X_i X_j) \leq nk. \quad (4.46)$$

所以由最大值不等式可得

$$\mathbb{P} \left(\left| \frac{S_n}{n} \right| > \varepsilon \right) = \mathbb{P}(|S_n| > n\varepsilon) = \mathbb{P}(S_n^2 > n^2 \varepsilon^2) \leq \frac{\mathbb{E} S_n^2}{n^2 \varepsilon^2} \leq \frac{nk}{n^2 \varepsilon^2} = \frac{k}{n \varepsilon^2}, \quad (4.47)$$

所以

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{S_n}{n} \right| > \varepsilon \right) \leq \lim_{n \rightarrow \infty} \frac{k}{n \varepsilon^2} = 0, \quad (4.48)$$

故

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{S_n}{n} \right| > \varepsilon \right) = 0. \quad (4.49)$$

例：最大值不等式

设 $\{X_n : n \geq 0\}$ 是鞅, 且对某一 $\alpha > 1$, $\mathbb{E}|X_n|^\alpha < \infty$, $\forall n \geq 0$. 证明

$$\mathbb{E} \left(\max_{0 \leq k \leq n} |X_k| \right) \leq \frac{\alpha}{\alpha - 1} (\mathbb{E}|X_n|^\alpha)^{1/\alpha}. \quad (4.50)$$

证明: 易知 $|X_n|^\alpha$ 是下鞅, 则由非负随机变量的期望计算方法, 概率值恒小于等于1以及最大值不等式有

$$\begin{aligned} \mathbb{E} \left(\max_{0 \leq k \leq n} |X_k| \right) &= \int_0^\infty \mathbb{P} \left(\max_{0 \leq k \leq n} |X_k| > t \right) dt \\ &= \int_0^\infty \mathbb{P} \left(\max_{0 \leq k \leq n} |X_k|^\alpha > t^\alpha \right) dt \\ &\leq \int_0^{(\mathbb{E}|X_n|^\alpha)^{1/\alpha}} 1 dt + \int_{(\mathbb{E}|X_n|^\alpha)^{1/\alpha}}^\infty \mathbb{P} \left(\max_{0 \leq k \leq n} |X_k|^\alpha > t^\alpha \right) dt \\ &\leq (\mathbb{E}|X_n|^\alpha)^{1/\alpha} + \int_{(\mathbb{E}|X_n|^\alpha)^{1/\alpha}}^\infty \frac{\mathbb{E}|X_n|^\alpha}{t^\alpha} dt \\ &= (\mathbb{E}|X_n|^\alpha)^{1/\alpha} + \frac{1}{-\alpha + 1} \mathbb{E}|X_n|^\alpha t^{-\alpha+1} \Big|_{(\mathbb{E}|X_n|^\alpha)^{1/\alpha}}^\infty \\ &= (\mathbb{E}|X_n|^\alpha)^{1/\alpha} + \frac{1}{\alpha - 1} (\mathbb{E}|X_n|^\alpha)^{1/\alpha} \\ &= \frac{\alpha}{\alpha - 1} (\mathbb{E}|X_n|^\alpha)^{1/\alpha}. \end{aligned} \quad (4.51)$$

4.5 连续参数鞅

定义：鞅

随机过程 $\{X_t : t \geq 0\}$ 是鞅, 若

$\forall t \geq 0$, 有 $\mathbb{E}|X_t| < \infty$

$\forall 0 \leq t_1 < \cdots < t_n < t_{n+1}$, 有

$$\mathbb{E}(X_{t_{n+1}} | X_{t_1}, X_{t_2}, \dots, X_{t_n}) = X_{t_n} \quad (4.52)$$

定义：停时

对随机过程 $X = \{X_t : t \geq 0\}$, 若取值于 $[0, \infty]$ 上的随机变量 T 满足 $\forall t \geq 0$, $\{T \leq t\}$ 由 $\{X_s : 0 \leq s \leq t\}$ 决定, 则称 T 关于 X 是停时.

定理：停时定理

设 $\{X_t : t \geq 0\}$ 是鞅, T 是停时, 若

$$\mathbb{P}(T < \infty) = 1, \quad (4.53)$$

且

$$\mathbb{E} \left(\sup_{t \geq 0} |X_{T \wedge t}| \right) < \infty, \quad (4.54)$$

所以

$$\mathbb{E}X_T = \mathbb{E}X_0. \quad (4.55)$$

第5章 布朗运动

布朗运动作为具有连续时间参数和连续状态空间的一个随机过程, 是一个最基本最简单同时又是最重要的随机过程, 许多其他的随机过程常常可以看作是它的泛函或某种意义下的推广. 它又是迄今了解得最清楚, 性质最丰富多彩的随机过程之一. 今天, 布朗运动及其推广已广泛地出现在许多纯科学领域中, 如物理, 经济, 通信理论, 生物, 管理科学与数理统计等, 同时, 由于布朗运动与微分方程如热传导方程等有密切的联系, 它又成为概率与分析联系的重要渠道.

5.1 随机游动与布朗运动的定义

考虑在一直线上的简单的, 对称的随机游动. 设质点每经过 Δt 时间, 随机地以概率 $p = 1/2$ 向右移 $\Delta x > 0$, 以概率 $q = 1/2$ 向左移动一个 Δx , 且每次移动相互独立. 若 X_t 表示 t 时刻质点的位置, 且有 $\Delta t \rightarrow 0$ 时 $\Delta x = c\sqrt{\Delta t}$, 则

$$X_t \sim N(0, c^2 t) \quad (5.1)$$

定义: 布朗运动

若一个随机过程 $\{X_t : t \geq 0\}$ 满足

X_t 是独立增量过程

增量平稳且服从期望为0, 方差为 $c^2 t$ 的正态分布, 即

$$X_{s+t} - X_s \sim N(0, c^2 t) \quad (5.2)$$

X_t 关于 t 是连续函数则称 $\{X_t : t \geq 0\}$ 是布朗运动或维纳过程.

当 $c = 1$ 时, 称 $\{X_t : t \geq 0\}$ 为标准布朗运动, 此时若 $X_0 = 0$, 则称为零初值标准布朗运动, 此时

$$X_t \sim N(0, t). \quad (5.3)$$

本章仅讨论标准布朗运动, 记为 $\{B_t : t \geq 0\}$, 其在 t 时刻的概率密度为

$$p(x; t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right), \quad (5.4)$$

即固定 t 时, 布朗运动 X_t 就是一个遵循正态分布 $N(0, t)$ 的随机变量.

5.1.1 布朗运动矩母函数

B_t 的矩母函数为

$$\phi(s) = \mathbb{E}(e^{sB_t}) = e^{s^2 t/2}. \quad (5.5)$$

5.1.2 布朗运动联合概率密度

定理: 布朗运动联合概率密度

设 $\{B_t : t \geq 0\}$ 为零初值标准布朗运动, 令 $x_0 = 0, t_0 = 0$, 对 $\forall 0 < t_1 < t_2 < \cdots < t_n$, $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$ 的联合概率密度函数为

$$g(x_1, x_2, \cdots, x_n; t_1, t_2, \cdots, t_n) = \prod_{i=1}^n p(x_i - x_{i-1}; t_i - t_{i-1}), \quad (5.6)$$

其中

$$p(x; t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right). \quad (5.7)$$

证明: 布朗运动增量独立且遵循正态分布, 则增量联合概率密度等于他们概率密度函数的直接累乘. 布朗运动等于增量求和的变换行列式为1, 则可直接带入得出布朗运动的联合概率密度函数.

给定初始条件 $B_{t_0} = x_0$, 对于任意的 $t > 0$, 布朗运动在 $t_0 + t$ 时刻的位置高于或低于初始位置的概率相等, 均为 $1/2$, 此即布朗运动的对称性.

5.1.3 布朗运动马尔可夫性

正向马尔可夫性 $\forall t_1 < t_2 < \cdots < t_n$, 在给定 $B_{t_1}, B_{t_2}, \cdots, B_{t_{n-1}}$ 下, B_{t_n} 的条件概率密度函数与只给定 $B_{t_{n-1}}$ 下 B_{t_n} 的条件概率密度相同

逆向马尔可夫性 $\forall t_1 > t_2 > \cdots > t_n$, 在给定 $B_{t_1}, B_{t_2}, \cdots, B_{t_{n-1}}$ 下, B_{t_n} 的条件概率密度函数与只给定 $B_{t_{n-1}}$ 下 B_{t_n} 的条件概率密度相同

中间关于两边的马尔可夫性 $\forall t_1 < t_2 < \cdots < t_n$, 在给定 $B_{t_1}, \cdots, B_{t_{i-1}}, B_{t_{i+1}}, \cdots, B_{t_n}$ 下, B_{t_i} 的条件概率密度函数与只给定 $B_{t_{i-1}}, B_{t_{i+1}}$ 下 B_{t_i} 的条件概率密度相同

由以上三种马尔可夫性可知, 布朗运动的条件概率密度函数只和与其刚好相邻的事件有关.

定理: 给定两边求中间分布

对 $0 \leq t_1 < t < t_2$, 给定 $B_{t_1} = a, B_{t_2} = b, B_0 = 0$, 则 B_t 的条件概率密度是正态密度, 其均值为

$$a + \frac{(b-a)(t-t_1)}{t_2-t_1}, \quad (5.8)$$

方差为

$$\frac{(t_2-t)(t-t_1)}{t_2-t_1}. \quad (5.9)$$

证明: 条件概率密度函数即为 (B_{t_1}, B_t, B_{t_2}) 联合密度与 (B_{t_1}, B_{t_2}) 联合密度的直接相除, 通过化简和配方即可得到期望与方差.

5.1.4 布朗运动与正态过程

定义: 正态过程

若随机过程 $X = \{X_t : t \in T\}$ 对任意 $t_i \in T, i = 1, 2, \dots, n$, 有 $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ 的联合分布为 n 维正态分布, 则称 X 为正态过程.

定理: 布朗运动的充要条件

设 $\{B_t : t \geq 0\}$ 为正态过程, 轨道连续, $B_0 = 0, \forall s, t > 0$, 有

$$\mathbb{E}B_t = 0, \quad \mathbb{E}[B_s B_t] = s \wedge t, \quad (5.10)$$

则 $\{B_t : t \geq 0\}$ 为布朗运动, 反之亦然.

定理: 由正态过程导出的布朗运动

设 $\{B_t : t \geq 0\}$ 为布朗运动, 则

$$\begin{aligned} & \{B_{t+\tau} - B_\tau : t \geq 0\}, \forall \tau \geq 0 \\ & \left\{ \frac{1}{\sqrt{\lambda}} B_{\lambda t} : t \geq 0 \right\}, \lambda > 0 \\ & \left\{ tB\left(\frac{1}{t}\right) : t \geq 0 \right\}, \text{ 其中 } tB\left(\frac{1}{t}\right) \Big|_{t=0} := 0 \\ & \{B_{t_0+s} - B_{t_0} : 0 \leq s \leq t_0\}, t_0 > 0 \end{aligned}$$

仍为布朗运动.

证明: 由正态过程, 轨道连续, 期望和协方差可知这些过程满足布朗运动的充要条件.

5.1.5 布朗运动的鞅性

定理: 布朗运动的鞅性

设 $\{B_t : t \geq 0\}$ 为布朗运动, 则

$$\begin{aligned} & \{B_t : t \geq 0\} \\ & \{B_t^2 - t : t \geq 0\} \\ & \{\exp(\lambda B_t - \lambda^2 t/2) : t \geq 0\} \\ & \{\exp(i\lambda B_t + \lambda^2 t/2) : t \geq 0\} \end{aligned}$$

关于布朗运动 $\{B_t : t \geq 0\}$ 均为鞅.

证明: 利用增量独立性, 作如下变换 $B_{t_{n+1}} = (B_{t_{n+1}} - B_{t_n}) + B_{t_n}$, 之后 $(B_{t_{n+1}} - B_{t_n})$ 相关项保留期望符号并计算相应的期望值, B_{t_n} 相关项直接提出到期望之外, 即得结论.

$$\{B_t^2 - t : t \geq 0\}$$

首先验证绝对值的期望有限,

$$\mathbb{E}|B_t^2 - t| \leq \mathbb{E}|B_t|^2 + t = 2t < \infty. \quad (5.11)$$

对于 $\forall 0 \leq t_1 < \dots < t_n < t_{n+1}$, 有

$$\begin{aligned} & \mathbb{E}(B_{t_{n+1}}^2 - t_{n+1} | B_{t_1}, B_{t_2}, \dots, B_{t_n}) \\ &= \mathbb{E}[(B_{t_{n+1}} - B_{t_n})^2 - t_{n+1} + 2B_{t_{n+1}}B_{t_n} - B_{t_n}^2 | B_{t_1}, B_{t_2}, \dots, B_{t_n}] \\ &= (t_{n+1} - t_n) - t_{n+1} + 2B_{t_n} \mathbb{E}[B_{t_{n+1}} - B_{t_n} + B_{t_n} | B_{t_1}, B_{t_2}, \dots, B_{t_n}] - B_{t_n}^2 \\ &= 2B_{t_n}^2 - B_{t_n}^2 - t_n \\ &= B_{t_n}^2 - t_n \end{aligned} \quad (5.12)$$

所以 $\{B_t^2 - t : t \geq 0\}$ 关于布朗运动 $\{B_t : t \geq 0\}$ 是鞅.

$\{\exp(\lambda B_t - \lambda^2 t/2) : t \geq 0\}$ 注意到, 令 $\{X_t = B_t + \mu t : t \geq 0\}$ 是漂移系数为 μ 的布朗运动, 令 $\lambda = -2\mu$, 则 $V_t = e^{-2\mu X_t}$ 是鞅, 故其在用鞅的停时定理求取漂移布朗运动停时大小概率时颇为有用.

首先验证绝对值的期望有限,

$$\begin{aligned}
 \mathbb{E}|\exp(\lambda B_t - \lambda^2 t/2)| &= \exp(-\lambda^2 t/2) \mathbb{E}[\exp(\lambda B_t)] \\
 &= \exp(-\lambda^2 t/2) \int_{-\infty}^{\infty} e^{\lambda x} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) dx \\
 &= \exp(-\lambda^2 t/2) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x - t\lambda)^2 - t^2 \lambda^2}{2t}\right) dx \\
 &= \exp(-\lambda^2 t/2) \exp(\lambda^2 t/2) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x - t\lambda)^2}{2t}\right) dx \\
 &= 1 < \infty.
 \end{aligned} \tag{5.13}$$

对于 $\forall 0 \leq t_0 < t_1 < \cdots < t_n < t_{n+1}$, 有

$$\begin{aligned}
 &\mathbb{E}(\exp(\lambda B_{t_{n+1}} - \lambda^2 t_{n+1}/2) | B_{t_1}, B_{t_2}, \dots, B_{t_n}) \\
 &= \mathbb{E}(\exp(\lambda(B_{t_{n+1}} - B_{t_n}) + \lambda B_{t_n} - \lambda^2 t_{n+1}/2) | B_{t_1}, B_{t_2}, \dots, B_{t_n}) \\
 &= \exp(\lambda B_{t_n} - \lambda^2 t_{n+1}/2) \mathbb{E}[\exp(\lambda(B_{t_{n+1}} - B_{t_n}))] \\
 &= \exp(\lambda B_{t_n} - \lambda^2 t_{n+1}/2) \mathbb{E}[\exp(\lambda B_{t_{n+1}-t_n})] \\
 &= \exp(\lambda B_{t_n} - \lambda^2 t_{n+1}/2) \exp(\lambda^2(t_{n+1} - t_n)/2) \\
 &= \exp(\lambda B_{t_n} - \lambda^2 t_n/2)
 \end{aligned} \tag{5.14}$$

所以 $\{\exp(\lambda B_t - \lambda^2 t/2) : t \geq 0\}$ 关于布朗运动 $\{B_t : t \geq 0\}$ 是鞅.

由以上结论可知, 布朗运动本身既是马尔可夫过程, 又是连续参数鞅. 这个结果很别致, 但并不奇怪. 因为已讨论的泊松过程, 马尔可夫链, 鞅, 布朗运动等随机过程, 不过是对一些随机过程某些方面的特殊性质进行了专门的、分类的讨论, 并不排斥这些性质可以交叉, 可以共存于一个随机过程中. 在介绍这些概念时只能串行讲述, 但实际上要能够并行应用, 融会贯通.

5.2 布朗运动轨道的性质

定理:

对给定的 $t > 0$, 有

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \sum_{k=1}^{2^n} \left(B_{\frac{kt}{2^n}} - B_{\frac{(k-1)t}{2^n}}\right)^2 = t\right) = 1. \tag{5.15}$$

引理: 令

$$Y_n = \max_{1 \leq k \leq 2^n} \left| B_{\frac{kt}{2^n}} - B_{\frac{(k-1)t}{2^n}} \right|, \tag{5.16}$$

所以

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} Y_n = 0\right) = 1. \quad (5.17)$$

定理： 布朗运动非有限变差

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \sum_{k=1}^{2^n} \left| B_{\frac{kt}{2^n}} - B_{\frac{(k-1)t}{2^n}} \right| = \infty\right) = 1. \quad (5.18)$$

定理： 均方收敛

固定 $t > 0$, 设 $0 = t_0 < t_1 < \cdots < t_n = t$, 记

$$\lambda = \max_{1 \leq k \leq n} (t_k - t_{k-1}), \quad (5.19)$$

则布朗运动差值平方和均方收敛到 t

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 \stackrel{\text{m.s.}}{=} t, \quad (5.20)$$

即两端差值平方的期望极限为 0

$$\lim_{\lambda \rightarrow 0} \mathbb{E} \left[\sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 - t \right]^2 = 0. \quad (5.21)$$

证明: 记 $Y_k = B_{t_k} - B_{t_{k-1}}$, $1 \leq k \leq n$, 则 $Y_k \sim N(0, t_k - t_{k-1})$. 由正态分布的性质可知

$$\mathbb{E}Y_k^2 = t_k - t_{k-1}, \quad \mathbb{E}Y_k^4 = 3(t_k - t_{k-1})^2, \quad (5.22)$$

又 Y_k^2 相互独立, 故当 $k \neq l$ 时, 有

$$\mathbb{E}(Y_k^2 Y_l^2) = \mathbb{E}Y_k^2 \mathbb{E}Y_l^2 = (t_k - t_{k-1})(t_l - t_{l-1}). \quad (5.23)$$

因此

$$\begin{aligned}
& \mathbb{E} \left[\sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 - t \right]^2 \\
&= \mathbb{E} \left(\sum_{k=1}^n Y_k^2 - t \right)^2 \\
&= \mathbb{E} \left(\sum_{k=1}^n Y_k^2 \right)^2 - 2t \mathbb{E} \left(\sum_{k=1}^n Y_k^2 \right) + t^2 \\
&= \sum_{k=1}^n \mathbb{E} Y_k^4 + 2 \sum_{k < l} \mathbb{E} (Y_k^2 Y_l^2) - 2t \sum_{k=1}^n \mathbb{E} Y_k^2 + t^2 \\
&= 3 \sum_{k=1}^n (t_k - t_{k-1})^2 + 2 \sum_{k < l} (t_k - t_{k-1})(t_l - t_{l-1}) - 2t \sum_{k=1}^n (t_k - t_{k-1}) + t^2 \quad (5.24) \\
&= 2 \sum_{k=1}^n (t_k - t_{k-1})^2 + \left(\sum_{k=1}^n (t_k - t_{k-1}) \right)^2 - 2t^2 + t^2 \\
&= 2 \sum_{k=1}^n (t_k - t_{k-1})^2 + t^2 - t^2 \\
&= 2 \sum_{k=1}^n (t_k - t_{k-1})^2 \\
&\leq 2\lambda \sum_{k=1}^n (t_k - t_{k-1}) \\
&= 2\lambda t \rightarrow 0 \quad (\lambda \rightarrow 0)
\end{aligned}$$

其中用到了求和平方的逆运算

$$\left(\sum_{k=1}^n (t_k - t_{k-1}) \right)^2 = \sum_{k=1}^n (t_k - t_{k-1})^2 + 2 \sum_{k < l} (t_k - t_{k-1})(t_l - t_{l-1}) \quad (5.25)$$

故

$$\lim_{\lambda \rightarrow 0} \mathbb{E} \left[\sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 - t \right]^2 = 0, \quad (5.26)$$

所以

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2 \stackrel{\text{m.s.}}{=} t. \quad (5.27)$$

定理:

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n B_{t_{k-1}} (B_{t_k} - B_{t_{k-1}}) \stackrel{\text{m.s.}}{=} \frac{B_t^2 - t}{2}, \quad (5.28)$$

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n B_{t_k} (B_{t_k} - B_{t_{k-1}}) \stackrel{\text{m.s.}}{=} \frac{B_t^2 + t}{2}, \quad (5.29)$$

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n B_{t_k + \theta(t_k - t_{k-1})} (B_{t_k} - B_{t_{k-1}}) \stackrel{\text{m.s.}}{=} \frac{B_t^2 - t + 2\theta t}{2}, \quad 0 \leq \theta \leq 1. \quad (5.30)$$

证明: 令

$$A_n = \sum_{k=1}^n B_{t_{k-1}} (B_{t_k} - B_{t_{k-1}}), \quad C_n = \sum_{k=1}^n B_{t_k} (B_{t_k} - B_{t_{k-1}}), \quad (5.31)$$

所以

$$A_n + C_n = B_t^2, \quad C_n - A_n \stackrel{\text{m.s.}}{\rightarrow} t, \quad (5.32)$$

故

$$A_n \stackrel{\text{m.s.}}{\rightarrow} \frac{B_t^2 - t}{2}, \quad C_n \stackrel{\text{m.s.}}{\rightarrow} \frac{B_t^2 + t}{2}. \quad (5.33)$$

由于我们只知道区间差平方的均方收敛情况, 故需先将原式化为区间差的平方的组合

$$\begin{aligned} & \sum_{k=1}^n B_{t_k + \theta(t_k - t_{k-1})} (B_{t_k} - B_{t_{k-1}}) \\ &= \frac{1}{2} \sum_{k=1}^n \left[B_{t_k}^2 - B_{t_{k-1}}^2 + (B_{t_k + \theta(t_k - t_{k-1})} - B_{t_{k-1}})^2 - (B_{t_k} - B_{t_k + \theta(t_k - t_{k-1})})^2 \right] \\ & \stackrel{\text{m.s.}}{\rightarrow} \frac{1}{2} [B_t^2 + \theta t - (1 - \theta)t] \\ &= \frac{B_t^2 - t + 2\theta t}{2}. \end{aligned} \quad (5.34)$$

定理: 布朗运动轨道不存在有限导数

设 $\{B_t : t \geq 0\}$ 为标准布朗运动, 则对任意固定的 $t \geq 0$ 和 $h > 0$, 有

$$\mathbb{P} \left(\limsup_{h \downarrow 0} \frac{B_{t+h} - B_t}{h} = +\infty \right) = 1, \quad (5.35)$$

$$\mathbb{P} \left(\liminf_{h \downarrow 0} \frac{B_{t+h} - B_t}{h} = -\infty \right) = 1, \quad (5.36)$$

可知布朗运动对几乎所有轨道 ω 都没有有限的导数.

5.3 首次时与最大值

5.3.1 首次时的分布

设 $\{B_t : t \geq 0\}$ 为零初值标准布朗运动, 令

$$T_a = \inf_{t > 0} \{t : B_t = a\}, \quad (5.37)$$

则 T_a 表示首次到达 a 的时间.

对 $\forall t > 0$,

$$M_t = \max_{0 \leq u \leq t} B_u \geq 0 \quad (5.38)$$

表示 $[0, t]$ 上布朗运动达到的最大值.

当 $a > 0$ 时, 有下列事件等价关系

$$\{T_a \leq t\} = \{M_t \geq a\}, \quad (5.39)$$

即到达 a 的时间若想要小于等于 t , 则布朗运动在 $[0, t]$ 上到达的最大值必须大于等于 a .

因此有 $\mathbb{P}(T_a \leq t) = \mathbb{P}(M_t \geq a)$, 又由全概率公式有

$$\mathbb{P}(B_t \geq a) = \mathbb{P}(T_a \leq t)\mathbb{P}(B_t \geq a|T_a \leq t) + \mathbb{P}(T_a > t)\mathbb{P}(B_t \geq a|T_a > t), \quad (5.40)$$

显然可知 $\mathbb{P}(B_t \geq a|T_a > t) = 0$, 又由布朗运动的对称性可知, 在 $T_a \leq t$ 的条件下, 即 $B_{T_a} = a$ 时, $\{B_t \geq a\}$ 与 $\{B_t < a\}$ 是等可能的, 即

$$\mathbb{P}(B_t \geq a|T_a \leq t) = \mathbb{P}(B_t < a|T_a \leq t) = \frac{1}{2}, \quad (5.41)$$

故 $\mathbb{P}(T_a \leq t) = 2\mathbb{P}(B_t \geq a)$. 于是 $a > 0$ 时有

$$\begin{aligned} \mathbb{P}(T_a \leq t) &= 2\mathbb{P}(B_t \geq a) \\ &= 2 \frac{1}{\sqrt{2\pi t}} \int_a^\infty \exp\left(-\frac{u^2}{2t}\right) du \\ &= 2 \left[1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right] \end{aligned} \quad (5.42)$$

因此也有

$$\begin{aligned} \mathbb{P}(M_t \leq a) &= 1 - \mathbb{P}(M_t > a) \\ &= 1 - 2 \left[1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right] \\ &= 2\Phi\left(\frac{a}{\sqrt{t}}\right) - 1 \end{aligned} \quad (5.43)$$

所以 M_t 分布函数为

$$F(a) = \begin{cases} 2 \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^a \exp\left(-\frac{u^2}{2t}\right) du - 1, & \text{if } a > 0 \\ 0, & \text{if } a \leq 0 \end{cases} \quad (5.44)$$

可知 $F(a)$ 连续, 且除 $a = 0$ 外其导数存在且连续, 故 M_t 为连续型随机变量, 其概率密度函数为

$$p(a) = \frac{2}{\sqrt{2\pi t}} \exp\left(-\frac{a^2}{2t}\right) \mathbb{I}_{\{a>0\}}. \quad (5.45)$$

当 $a < 0$ 时, 由布朗运动的对称性有 $\mathbb{P}(T_{-a} \leq t) = \mathbb{P}(T_a \leq t)$, 所以对一般的 $a \in \mathbb{R}$ 有

$$\mathbb{P}(T_a \leq t) = 2 \left[1 - \Phi\left(\frac{|a|}{\sqrt{t}}\right)\right]. \quad (5.46)$$

T_a 几乎处处有限, 因

$$\mathbb{P}(T_a < \infty) = \lim_{t \rightarrow \infty} \mathbb{P}(T_a \leq t) = 2[1 - \Phi(0)] = 1 \quad (5.47)$$

T_a 期望为无穷, 通过累次积分换序并截断积分进行放缩有

$$\begin{aligned} \mathbb{E}T_a &= \int_0^\infty \mathbb{P}(T_a > t) dt \\ &= \frac{2}{\sqrt{2\pi}} \int_0^\infty \int_0^{|a|/\sqrt{t}} \exp\left(-\frac{x^2}{2}\right) dx dt \\ &= \frac{2}{\sqrt{2\pi}} \int_0^\infty \int_0^{a^2/x^2} 1 dt \exp\left(-\frac{x^2}{2}\right) dx \\ &= \frac{2}{\sqrt{2\pi}} \int_0^\infty \frac{a^2}{x^2} \exp\left(-\frac{x^2}{2}\right) dx \\ &\geq \frac{2a^2}{\sqrt{2\pi}} \int_0^1 \frac{1}{x^2} \cdot e^{-1/2} dx = \infty \end{aligned} \quad (5.48)$$

5.3.2 首达时的期望

令 $T = \inf\{t : B_t \notin (-r, 2r)\}$, $r > 0$, 即

$$T = \inf\{t : B_t = -r \text{ or } B_t = 2r\}, \quad (5.49)$$

由于 T 是停时, 故可用鞅 $\{Z_t = B_t^2 - t : t \geq 0\}$ 的停时定理来求取 $\mathbb{E}T$.

记 $\forall a \in \mathbb{R}$, $T_a = \inf\{t : B_t = a\}$, 由 $T_{2r} \subset T$ 有

$$1 \geq \mathbb{P}(T < \infty) \geq \mathbb{P}(T_{2r} < \infty) = 1, \quad (5.50)$$

故 $\mathbb{P}(T < \infty) = 1$, 所以 $\sup_{t \geq 0} |T \wedge t| < \infty$, 又 $|B_{T \wedge t}| \leq 2r$, 则

$$\mathbb{E}\left(\sup_{t \geq 0} |Z_{T \wedge t}|\right) = \mathbb{E}\left(\sup_{t \geq 0} |B_{T \wedge t}^2 - T \wedge t|\right) \leq 4r^2 + \sup_{t \geq 0} |T \wedge t| < \infty, \quad (5.51)$$

故由连续时间参数鞅停时定理有

$$\mathbb{E}Z_T = \mathbb{E}B_T^2 - \mathbb{E}T = \mathbb{E}Z_0 = 0, \quad (5.52)$$

所以

$$\begin{aligned} \mathbb{E}T &= \mathbb{E}B_T^2 \\ &= (-r)^2 \mathbb{P}(T_{-r} < T_{2r}) + (2r)^2 \mathbb{P}(T_{2r} \leq T_{-r}) \\ &= r^2 \cdot \frac{2}{3} + 4r^2 \cdot \frac{1}{3} \\ &= 2r^2. \end{aligned} \quad (5.53)$$

其中, $\mathbb{P}(T_{-a} < T_b)$ 这种概率可由5.6节定理取 $\mu \rightarrow 0$ 极限得到, 也可由鞅的停时定理求得. 令 $T = \inf\{t : B_t = -a \text{ or } B_t = b\}$, $a, b > 0$, 易知 T 关于 B_t 是停时, $\mathbb{P}(T < \infty) = 1$, 且

$$\mathbb{E}\left(\sup_{t \geq 0} |B_{T \wedge t}|\right) \leq \max\{a, b\} < \infty, \quad (5.54)$$

所以由连续时间参数鞅停时定理有

$$\mathbb{E}B_T = \mathbb{E}B_0 = 0, \quad (5.55)$$

即

$$\mathbb{E}B_T = (-a)\mathbb{P}(T_{-a} < T_b) + b[1 - \mathbb{P}(T_{-a} < T_b)], \quad (5.56)$$

所以

$$\mathbb{P}(T_{-a} < T_b) = \frac{b}{a+b}. \quad (5.57)$$

5.4 布朗桥

5.5 布朗运动的各种变形与推广

5.5.1 在某点被吸收的布朗运动

设

$$Z_t = \begin{cases} B_t, & \text{if } t < T_x, \\ x & \text{if } t \geq T_x, \end{cases} \quad (5.58)$$

其中 $T_x = \min_{t>0}\{t : B_t = x\}$. $\{Z_t : t \geq 0\}$ 表示一旦随机过程第一次到达 x 后即被吸收停留在 x , 称为在 x 点被吸收的布朗运动, 其为混合型随机变量.

其分布函数为

$$\begin{cases} \mathbb{P}(Z_t \leq y) = 1, & \text{if } y > x, \\ \mathbb{P}(Z_t = y) = \mathbb{P}(T_y \leq t), & \text{if } y = x, \\ \mathbb{P}(Z_t \leq y) = \mathbb{P}(B_t \leq y, M_t < x), & \text{if } y < x, \end{cases} \quad (5.59)$$

其中

$$\begin{aligned} \mathbb{P}(B_t \leq y, M_t < x) &= \mathbb{P}(B_t \leq y) - \mathbb{P}(M_t \geq x, B_t \leq y) \\ &= \mathbb{P}(B_t \leq y) - \mathbb{P}(B_t \geq x + x - y) \\ &= \mathbb{P}(B_t \leq y) - \mathbb{P}(B_t \leq y - 2x) \\ &= \mathbb{P}(y - 2x \leq B_t \leq y). \end{aligned} \quad (5.60)$$

5.5.2 在原点反射的布朗运动

令 $Y_t = |B_t|$, 则称 $\{Y_t : t \geq 0\}$ 为在原点反射的布朗运动. 当 $y < 0$ 时, $\mathbb{P}(Y_t \leq y) = 0$. 当 $y \geq 0$ 时,

$$\mathbb{P}(Y_t \leq y) = \mathbb{P}(|B_t| \leq y) = \mathbb{P}(-y \leq B_t \leq y) = 2\mathbb{P}(B_t \leq y) - 1. \quad (5.61)$$

5.5.3 几何布朗运动

令 $W_t = e^{B_t}$, 则称 $\{W_t : t \geq 0\}$ 为几何布朗运动, 且由布朗运动矩母函数可直接得

$$\mathbb{E}W_t = e^{t/2}, \quad \mathbb{E}W_t = e^{2t} - e^t. \quad (5.62)$$

5.5.4 布朗运动的积分

令 $S_t = \int_0^t B_u du$, 则称 $\{S_t : t \geq 0\}$ 为布朗运动的积分.

$$\mathbb{E}S_t = 0, \quad \mathbb{E}S_t^2 = \int_0^t (t-s)^2 ds = \frac{t^3}{3}. \quad (5.63)$$

5.5.5 布朗运动的形式导数

令 $D_t = \frac{\Delta B_t}{\Delta t}$, 则称 $\{D_t : t \geq 0\}$ 为布朗运动的形式导数, 是正态过程, 且

$$D_t \sim N(0, 1/\Delta t). \quad (5.64)$$

5.6 带有漂移的布朗运动

定义: 带有漂移的布朗运动

设 $\{B_t : t \geq 0\}$ 为标准布朗运动, 记 $X_t = B_t + \mu t$, $\mu \in \mathbb{R}$, 称 $\{X_t : t \geq 0\}$ 为带有漂移系数 μ 的布朗运动.

定理: 首达时大小关系

设 $\{X_t = B_t + \mu t : t \geq 0\}$ 是漂移系数为 μ 的布朗运动, 对 $a, b > 0$, $-b < x < a$, 令

$$T_a = \min\{t : t > 0, X_t = a\}, \quad T_{-b} = \min\{t : t > 0, X_t = -b\} \quad (5.65)$$

有

$$f(x) = \mathbb{P}(T_a < T_{-b} < \infty | X_0 = x) = \frac{e^{2\mu b} - e^{-2\mu x}}{e^{2\mu b} - e^{-2\mu a}}, \quad (5.66)$$

特别地, 当 $\mu = 0$ 时, 有

$$f(x) = \mathbb{P}(T_a < T_{-b} < \infty | X_0 = x) = \frac{2\mu b + 2\mu x}{2\mu b + 2\mu a} = \frac{b+x}{b+a}. \quad (5.67)$$

当 $x = 0$, $\mu = 0$ 时, 有

$$\mathbb{P}(T_a < \infty | X_0 = 0) = \frac{b}{b+a}, \quad (5.68)$$

上式与用鞅的停时定理求出的概率相同.

推论:

设 $\{X_t = B_t + \mu t : t \geq 0\}$ 是漂移系数为 μ 的布朗运动, 若 $\mu < 0$, 则

$$\mathbb{P}(T_a < \infty | X_0 = x) = \lim_{b \rightarrow \infty} \frac{e^{2\mu b} - e^{-2\mu x}}{e^{2\mu b} - e^{-2\mu a}} = \frac{0 - e^{-2\mu x}}{0 - e^{-2\mu a}} = e^{2\mu(a-x)}. \quad (5.69)$$

特别地, 当 $x = 0$ 时, 有

$$\mathbb{P}(T_a < \infty | X_0 = 0) = e^{2\mu a}. \quad (5.70)$$

当 $x = 0, \mu = 0$ 时, 有

$$\mathbb{P}(T_a < \infty | X_0 = 0) = 1, \quad (5.71)$$

上式再次验证了 T_a 几乎处处有限的事实.

5.7 多维布朗运动

第6章 连续时间参数马尔可夫链

在第3章中, 曾详细地讨论了离散参数马尔可夫链的有关问题, 本章将着重研究连续参数可列状态空间的马尔可夫过程.

6.1 定义与若干基本概念

仍记状态空间为 $S = \{0, 1, 2, \dots\}$.

定义: 连续参数马氏链

设随机过程 $X = \{X_t : t \geq 0\}$ 对 $\forall 0 \leq t_0 < t_1 < \dots < t_n < t_{n+1}$, $i_k \in S$, $0 \leq k \leq n+1$, 若

$$\mathbb{P}(X_{t_0} = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) > 0, \quad (6.1)$$

有

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, X_{t_1} = i_1, \dots, X_{t_n} = i_n) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n), \quad (6.2)$$

则称 X 为连续时间参数马氏链. 若对 $\forall s, t \geq 0$, $i, j \in S$, 有

$$\mathbb{P}(X_{s+t} = j | X_s = i) = \mathbb{P}(X_t = j | X_0 = i) := P_{ij}(t), \quad (6.3)$$

即增量转移概率只与时间增量有关, 则称 X 为时齐马氏链. 称

$$\mathbf{P}(t) = [P_{ij}(t)], \quad i, j \in S \quad (6.4)$$

为转移概率矩阵, 易知它满足

非负性:

$$P_{ij}(t) \geq 0, \quad \forall i, j \in S \quad (6.5)$$

归一性: 条件概率测度下的概率归一性

$$\sum_{j \in S} P_{ij}(t) = 1, \quad \forall i \in S \quad (6.6)$$

Kolmogorov-Chapman 方程: 注意到状态空间的可列性则有

$$P_{ij}(s+t) = \sum_{k \in S} P_{ik}(s)P_{kj}(t), \quad \forall s, t \geq 0, \quad i, j \in S \quad (6.7)$$

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t) \quad (6.8)$$

初值条件: 使用 Kronecker δ 符号有

$$P_{ij}(0) = \delta_{ij}, \quad \forall i, j \in S \quad (6.9)$$

$$\mathbf{P}(0) = \mathbf{I} \quad (6.10)$$

连续性条件: $P_{ij}(t)$ 在原点连续

$$\lim_{t \rightarrow 0} P_{ij}(t) = P_{ij}(0) = \delta_{ij}, \quad \forall i, j \in S \quad (6.11)$$

$$\lim_{t \rightarrow 0} \mathbf{P}(t) = \mathbf{P}(0) = \mathbf{I} \quad (6.12)$$

命题: 一致连续性

若 $\mathbf{P}(t) = [P_{ij}(t)]$, $i, j \in S$ 为标准转移概率矩阵, 则

对 $\forall i \in S$, $P_{ij}(t)$ 在 $[0, \infty)$ 上一致连续, 且此时一致性对 j 亦成立

$\forall t \geq 0, i \in S$,

$$P_{ii}(t) > 0 \quad (6.13)$$

定义: 概率分布

记 $\pi_i(t) = \mathbb{P}(X_t = i)$, $\forall t \geq 0, i \in S$, 称

$$\boldsymbol{\pi}(t) = (\pi_i(t), i \in S) \quad (6.14)$$

为马尔可夫链在 t 时刻的分布, 称 $\boldsymbol{\pi}(0)$ 为初始分布, 且有

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\mathbf{P}(t) \quad (6.15)$$

定义: 离散骨架

对于连续时间马尔可夫链 $X = \{X_t : t \geq 0\}$, 任取 $h > 0$, 定义

$$X_n(h) = X(nh), \quad n \geq 0, \quad (6.16)$$

由马尔可夫性可知, $\{X_n(h) : n \geq 0\}$ 是一个离散时间的马尔可夫链, 称其为以 h 为步长的 h -离散骨架, 简称 h 骨架, 它的 n 步转移概率矩阵为 $\mathbf{P}(nh)$.

对于连续参数马尔可夫链与离散参数马尔可夫链, 由于它们都具有马尔可夫性, 且状态空间均为可数集或有限集, 因而许多概念和性质有相同或相似之处, 例如状态相通, 状态分类, 不可约链, 平稳分布与极限分布等.

定义: 可达与互通

可达: 若 $\exists t > 0, P_{ij}(t) > 0$, 则称由状态 i 可达状态 j , 记为 $i \rightarrow j$

不可达: 若 $\forall t > 0, P_{ij}(t) = 0$, 则称由状态 i 不可达状态 j , 记为 $i \nrightarrow j$

互通: 若 $i \rightarrow j$ 且 $j \rightarrow i$, 则称状态 i 与 j 相通, 记为 $i \leftrightarrow j$

由 $\forall i \in S, P_{ii}(t) > 0$, 即 $i \leftrightarrow i$, 可知相通关系具有自反性、对称性、传递性, 故相通关系是等价关系, 从而可以按相通关系给状态分类, 相通的状态组成一个状态类. 若整个状态空间是一个状态类, 则称该马尔可夫链是不可约的.

对于连续时间的马尔可夫链, 对所有 $h > 0$ 及正整数 $n, \forall i \in S, P_{ii}(nh) > 0$, 这意味着对每一个离散的骨架 $X_n(h)$, 每一个状态立都是非周期的, 故可知对 $\forall j \in S, \forall h > 0$,

$$\lim_{n \rightarrow \infty} P_{ij}(nh) = \pi_{ij} \quad (6.17)$$

总存在. 所以对连续时间的马尔可夫链就无需引入周期的概念, 而且利用 $P_{ij}(t)$ 在 $[0, \infty)$ 上一致连续性及 $\lim_{n \rightarrow \infty} P_{ij}(nh) = \pi_{ij}$ 总存在, 可以证明 $P_{ij}(t)$ 在 $t \rightarrow \infty$ 时极限总存在.

命题

$\forall i, j \in S$, 下述极限总存在

$$\lim_{t \rightarrow \infty} P_{ij}(t) = \pi_{ij}. \quad (6.18)$$

定义: 常返与平稳分布

常返: 若

$$\int_0^\infty P_{ii}(t) dt = \infty, \quad (6.19)$$

则称状态 i 为常返状态, 否则称为非常返状态

正常返: 若

$$\lim_{t \rightarrow \infty} P_{ii}(t) > 0 \quad (6.20)$$

则称状态 i 为正常返状态, 否则若

$$\lim_{t \rightarrow \infty} P_{ii}(t) = 0 \quad (6.21)$$

则称状态 i 为零常返状态

平稳分布: 若概率分布 $\pi = (\pi_i, i \in S)$ 满足

$$\pi = \pi P(t), \quad \forall t \geq 0, \quad (6.22)$$

则称 π 为马氏链的平稳分布

极限分布: 若对 $\forall i \in S$,

$$\lim_{n \rightarrow \infty} \pi_i(t) = \pi_i^* \quad (6.23)$$

存在, 则称

$$\pi^* := \{\pi_i^* : i \in S\} \quad (6.24)$$

为马氏链的极限分布

可总结离散参数马氏链与连续参数马氏链常返与正常返相关判断依据如下.

定理: 平稳分布与极限分布

不可约链是正常返的充要条件是它存在平稳分布, 且此时平稳分布就等于极限分布.

状态	非常返	常返	正常返	零常返
离散 MC	$f_{ii} < 1$	$f_{ii} = 1$	$\mu_i = \infty$	$\mu_i < \infty$
离散 MC	$G_{ii} = \frac{1}{1-f_{ii}} < \infty$	$G_{ii} = \infty$	$p_{ii}^{(n)} \rightarrow \frac{1}{\mu_i} > 0$	$p_{ii}^{(n)} \rightarrow 0$
连续 MC	$\int_0^\infty P_{ii}(t)dt < \infty$	$\int_0^\infty P_{ii}(t)dt = \infty$	$P_{ii}(t) \rightarrow C > 0$	$P_{ii}(t) \rightarrow 0$

6.2 转移率矩阵及其概率意义

在离散参数马尔可夫链中, 我们知道由一步转移概率矩阵 $\mathbf{P} = [P_{ij}]$ 可以完全确定 n 步转移矩阵, 即有

$$\mathbf{P}^{(n)} = \mathbf{P}^n = \mathbf{e}^{n \ln \mathbf{P}}. \quad (6.25)$$

那么对连续参数马尔可夫链, 是否有类似的表达式, 即

$$\mathbf{P}(t) = \mathbf{e}^{t\mathbf{Q}} \quad (6.26)$$

呢? 其中 \mathbf{Q} 为与 t 无关的实数矩阵, 假如上式存在, 则应有

$$\mathbf{P}'(0) = \lim_{t \rightarrow 0} \frac{\mathbf{P}(t) - \mathbf{P}(0)}{t} = \lim_{t \rightarrow 0} \frac{\mathbf{e}^{t\mathbf{Q}} - \mathbf{I}}{t} = \mathbf{Q}. \quad (6.27)$$

直观来说, 即

$$\begin{aligned} \mathbf{Q} &= \left(\begin{bmatrix} P_{00}(t) & P_{01}(t) & P_{02}(t) & \cdots \\ P_{10}(t) & P_{11}(t) & P_{12}(t) & \cdots \\ P_{00}(t) & P_{01}(t) & P_{02}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \right)'_t \\ &= \begin{bmatrix} P_{00}(t) - 1 & P_{01}(t) & P_{02}(t) & \cdots \\ P_{10}(t) & P_{11}(t) - 1 & P_{12}(t) & \cdots \\ P_{00}(t) & P_{01}(t) & P_{02}(t) - 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}'_t \\ &:= \begin{bmatrix} -q_0 := q_{00} & q_{01} & q_{02} & \cdots \\ q_{10} & -q_1 := q_{11} & q_{12} & \cdots \\ q_{20} & q_{21} & -q_2 := q_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \end{aligned} \quad (6.28)$$

这就提示我们先要研究 $\mathbf{P}(t)$ 在 $t = 0$ 的导数即变化率是否存在的问题.

定理:

对 $\forall i \in S$, 极限

$$q_i = -q_{ii} := \lim_{t \rightarrow 0} \frac{1 - P_{ii}(t)}{t} \quad (6.29)$$

存在, 但可能是无限.

定理:

对 $\forall i, j \in S, j \neq i$, 极限

$$q_{ij} := P'_{ij}(0) = \lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} \quad (6.30)$$

存在且有限.

推论:

对 $\forall i \in S$,

$$q_i \geq \sum_{j \neq i} q_{ij} \geq 0. \quad (6.31)$$

推论:

当 S 为有限状态空间时, $\forall i \in S$, 有

$$0 \leq q_i = \sum_{j \neq i} q_{ij} < \infty. \quad (6.32)$$

定义: Q 矩阵

记 $Q = [q_{ij}]$ $i, j \in S$, 称 Q 为 $X = \{X_t : t \geq 0\}$ 的转移率矩阵. 若转移率矩阵满足 $\forall i \in S$,

$$q_i = \sum_{j \neq i} q_{ij} < \infty, \quad (6.33)$$

则称 Q 为保守 Q 矩阵. 易知当 S 为有限状态空间时, Q 必保守.

定义: 逗留时间令 τ_1 表示马氏链逗留在初始状态的时间, 即

$$\tau_1 = \inf_{t > 0} \{t : X_t \neq X_0\}. \quad (6.34)$$

定理: 逗留时间服从指数分布

设马尔可夫链 $X = \{X_t : t \geq 0\}$ 轨道右连续, 则对 $\forall i \in S, t \geq 0$, 有

$$\mathbb{P}(\tau_1 > t | X_0 = i) = e^{-q_i t}, \quad (6.35)$$

这说明系统逗留在 $X_0 = i$ 状态的时间 τ_1 服从参数为 q_i 的指数分布, 其期望为

$$\mathbb{E}[\tau_1 | X_0 = i] = \int_0^\infty \mathbb{P}(\tau_1 > t | X_0 = i) dt = \int_0^\infty e^{-q_i t} dt = \frac{1}{q_i}. \quad (6.36)$$

定义: 吸收状态, 瞬时状态与逗留状态

吸收状态: $q_i = 0$, 即从 i 出发, 过程以概率 1 永远停留在 i 状态

瞬时状态: $q_i = \infty$, 这说明 X 在 i 状态几乎不停留立即跳到别的状态

逗留状态: $0 < q_i < \infty$, 过程停留在状态 i , 若干时间后跳到其他状态, 且停留时间服从指数分布

定理:

设马尔可夫链 $X = \{X_t : t \geq 0\}$ 轨道右连续, 且 $0 < q_i < \infty$, 则对 $\forall t \geq 0, j \neq i$, 有

$$\mathbb{P}(\tau_1 \leq t, X_{\tau_1} = j | X_0 = i) = \frac{q_{ij}}{q_i}(1 - e^{-q_i t}), \quad (6.37)$$

$$\mathbb{P}(X_{\tau_1} = j | X_0 = i) = \frac{q_{ij}}{q_i}. \quad (6.38)$$

推论:

设马尔可夫链 $X = \{X_t : t \geq 0\}$ 轨道右连续, \mathbf{Q} 为保守 Q 矩阵, $0 < q_i < \infty$, 则 $\forall i \in S, X_{\tau_1}$ 与 τ_1 关于 $X_0 = i$ 条件独立, 换句话说, 已知当前状态, 逗留时间与下一步跳到哪里去是条件独立的.

6.3 Kolmogorov 向前向后微分方程

定理: Kolmogorov 向前向后微分方程

设马尔可夫链 $X = \{X_t : t \geq 0\}$, $\mathbf{P}(t) = [P_{ij}(t)]$, $\mathbf{Q} = [q_{ij}]$, $i, j \in S$, 当 S 为有限状态空间时, 有

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}, \quad (6.39)$$

$$\mathbf{P}'(t) = \mathbf{Q}\mathbf{P}(t). \quad (6.40)$$

其中 $\mathbf{P}(t)$ 在前面称为向前微分方程, $\mathbf{P}(t)$ 在后面称为向后微分方程.

定理: Fokker-Planck 方程

当 S 为有限状态空间时, 成立下列 Fokker-Planck 方程

$$\pi'(t) = \pi(0)\mathbf{P}'(t) = \pi(0)\mathbf{P}(t)\mathbf{Q} = \pi(t)\mathbf{Q} \quad (6.41)$$

例: 解 Kolmogorov 向前向后微分方程

设有连续时间参数 Markov 链 X , 其轨道右连续,

$$\tau_1 = \inf_{t>0} \{t : X_t \neq X_0\}, \quad (6.42)$$

已知

$$\mathbb{P}(\tau_1 > t | X_0 = 0) = e^{-\lambda t}, \quad \mathbb{P}(\tau_1 > t | X_0 = 1) = e^{-\mu t} \quad (6.43)$$

写出关于 $p_{00}(t)$, $p_{11}(t)$ 的 Kolmogorov 向前微分方程, 并求解得到表达式.

解: 由题意可知 $q_0 = \lambda$, $q_1 = \mu$, 即 $q_{00} = -\lambda$, $q_{11} = -\mu$, 故 \mathbf{Q} 矩阵为

$$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}, \quad (6.44)$$

所以由 Kolmogorov 向前微分方程

$$\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q} \quad (6.45)$$

可得

$$\begin{bmatrix} p'_{00}(t) & p'_{01}(t) \\ p'_{10}(t) & p'_{11}(t) \end{bmatrix} = \begin{bmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{bmatrix} \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix} \quad (6.46)$$

即

$$\begin{bmatrix} p'_{00}(t) & p'_{01}(t) \\ p'_{10}(t) & p'_{11}(t) \end{bmatrix} = \begin{bmatrix} -\lambda p_{00}(t) + \mu p_{01}(t) & \lambda p_{00}(t) - \mu p_{01}(t) \\ -\lambda p_{10}(t) + \mu p_{11}(t) & \lambda p_{10}(t) - \mu p_{11}(t) \end{bmatrix} \quad (6.47)$$

又

$$p_{00}(t) + p_{01}(t) = 1, \quad p_{10}(t) + p_{11}(t) = 1, \quad (6.48)$$

所以关于 $p_{00}(t)$, $p_{11}(t)$ 的 Kolmogorov 向前微分方程为

$$\begin{aligned} p'_{00}(t) &= -\lambda p_{00}(t) + \mu[1 - p_{00}(t)] = -(\lambda + \mu)p_{00}(t) + \mu \\ p'_{11}(t) &= \lambda[1 - p_{11}(t)] - \mu p_{11}(t) = -(\lambda + \mu)p_{11}(t) + \lambda \end{aligned} \quad (6.49)$$

即

$$\begin{aligned} p'_{00}(t) + (\lambda + \mu)p_{00}(t) &= \mu \\ p'_{11}(t) + (\lambda + \mu)p_{11}(t) &= \lambda \end{aligned} \quad (6.50)$$

两边乘以积分因子

$$\begin{aligned} \exp[(\lambda + \mu)t][p'_{00}(t) + (\lambda + \mu)p_{00}(t)] &= \mu \exp[(\lambda + \mu)t] \\ \exp[(\lambda + \mu)t][p'_{11}(t) + (\lambda + \mu)p_{11}(t)] &= \lambda \exp[(\lambda + \mu)t] \end{aligned} \quad (6.51)$$

即

$$\begin{aligned} \{\exp[(\lambda + \mu)t]p_{00}(t)\}'_t &= \mu \exp[(\lambda + \mu)t] \\ \{\exp[(\lambda + \mu)t]p_{11}(t)\}'_t &= \lambda \exp[(\lambda + \mu)t] \end{aligned} \quad (6.52)$$

两边从 0 到 t 积分并利用初始条件 $p_{00}(0) = p_{11}(0) = 1$ 可得

$$\begin{aligned} \exp[(\lambda + \mu)t]p_{00}(t) - 1 &= \frac{\mu}{\lambda + \mu} \exp[(\lambda + \mu)t] - \frac{\mu}{\lambda + \mu} \\ \exp[(\lambda + \mu)t]p_{11}(t) - 1 &= \frac{\lambda}{\lambda + \mu} \exp[(\lambda + \mu)t] - \frac{\lambda}{\lambda + \mu} \end{aligned} \quad (6.53)$$

故

$$\begin{aligned} p_{00}(t) &= \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp[-(\lambda + \mu)t] \\ p_{11}(t) &= \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} \exp[-(\lambda + \mu)t] \end{aligned} \quad (6.54)$$

6.4 生灭过程

6.5 强马尔可夫性与嵌入马尔可夫链

第7章 随机微分方程

7.1 H 空间和均方收敛

在许多情况下,人们关心的是一个过程的一阶矩、二阶矩特征,这是比较容易得到的随机变量的外部数字特征.因此,研究二阶矩存在的随机变量是一个重要的方向,通过深入分析来探知这类随机变量有哪些共同的性质

为了对存在二阶矩的随机变量的全体进行统一考察,引出 H 空间的概念,为研究这一类随机变量提供数学框架与几何直观解释

定义: H 空间

$H := \{X : \mathbb{E}|X|^2 < \infty\}$, 即 H 是由二阶矩存在的随机变量全体构成的集合,称作 H 空间.

H 空间性质

H 空间是线性空间,即 $\forall X_1, X_2 \in H$ 及常数 $\alpha_1, \alpha_2 \in \mathbb{R}$, 都有 $\alpha_1 X_1 + \alpha_2 X_2 \in H$ 成立.

证明: 由 Cauchy-Schwartz 公式可得

$$\begin{aligned}\mathbb{E}|\alpha_1 X_1 + \alpha_2 X_2|^2 &= \mathbb{E}|\alpha_1^2 X_1^2 + 2\alpha_1 \alpha_2 X_1 X_2 + \alpha_2^2 X_2^2| \\ &\leq \alpha_1^2 \mathbb{E}|X_1|^2 + 2|\alpha_1 \alpha_2| \mathbb{E}|X_1 X_2| + \alpha_2^2 \mathbb{E}|X_2|^2 \\ &\leq \alpha_1^2 \mathbb{E}|X_1|^2 + 2|\alpha_1 \alpha_2| \sqrt{\mathbb{E}|X_1|^2 \mathbb{E}|X_2|^2} + \alpha_2^2 \mathbb{E}|X_2|^2 \\ &< +\infty\end{aligned}\tag{7.1}$$

所以 $\alpha_1 X_1 + \alpha_2 X_2 \in H$

定义: 内积

$\forall X, Y \in H$, 定义内积为 $\langle X, Y \rangle = \mathbb{E}(XY)$.

内积性质

共轭对称: $\langle Y, X \rangle = \langle X, Y \rangle$

齐次性: $\langle cX, Y \rangle = c\langle X, Y \rangle$, $c \in \mathbb{R}$

线性: $\langle X_1 + X_2, Y \rangle = \langle X_1, Y \rangle + \langle X_2, Y \rangle$

非负性: $\langle X, X \rangle = \mathbb{E}|X|^2 \geq 0$, $\langle X, X \rangle = 0 \Leftrightarrow X = 0$ (a.e.)

定义: 正交

若 $\langle X, Y \rangle = 0$, 则称 X 与 Y 正交, 记为 $X \perp Y$.

定义: 范数

对 $\forall X \in H$, 定义范数为 $\|X\| = \sqrt{\langle X, X \rangle} = \sqrt{\mathbb{E}|X|^2}$.

范数性质

非负性: $\|X\| \geq 0$, $\|X\| = 0 \Leftrightarrow X = 0$ (a.e.)

齐次性: $\|cX\| = |c|\|X\|$, $c \in \mathbb{R}$

三角不等式: $\|X_1 + X_2\| \leq \|X_1\| + \|X_2\|$

不等式: $|\mathbb{E}X| \leq \mathbb{E}|X| \leq \|X\|$, 由凸函数性质和 Cauchy-Schwartz 公式可得

定义: 距离

对 $\forall X, Y \in H$, 定义距离为 $d(X, Y) = \|X - Y\| = \sqrt{\langle X - Y, X - Y \rangle} = \sqrt{\mathbb{E}|X - Y|^2}$.

距离性质

非负性: $d(X, Y) \geq 0$, $d(X, Y) = 0 \Leftrightarrow X = Y$ (a.e.)

对称性: $d(X, Y) = d(Y, X)$

三角不等式: $d(X, Z) \leq d(X, Y) + d(Y, Z)$

定义: 极限与收敛

设 $X, X_n \in H$, $n \geq 1$, 若

$$\lim_{n \rightarrow \infty} d(X_n, X) = \lim_{n \rightarrow \infty} \|X_n - X\| = \lim_{n \rightarrow \infty} \sqrt{\mathbb{E}|X_n - X|^2} = 0, \quad (7.2)$$

则称 X 为序列 $\{X_n, n \geq 1\}$ 的均方极限, 记作

$$\lim_{n \rightarrow \infty} X_n \stackrel{\text{m.s.}}{=} X, \quad (7.3)$$

其中 m.s. 表示 mean square, 简记为

$$\lim_{n \rightarrow \infty} X_n = X, \quad X_n \xrightarrow{\text{m.s.}} X, \quad (7.4)$$

即序列 $\{X_n, n \geq 1\}$ 均方收敛于 X .

定义: Cauchy 序列

设 $X_n \in H$, $n \geq 1$, 若

$$\lim_{m, n \rightarrow \infty} d(X_m, X_n) = 0, \quad (7.5)$$

则称 $\{X_n, n \geq 1\}$ 是 Cauchy 序列.

命题: 完备的赋范线性空间

设 $\{X_n, n \geq 1\}$ 为 H 空间中的 Cauchy 序列, 则必存在随机变量 $X \in H$ 使得

$$X_n \xrightarrow{\text{m.s.}} X. \quad (7.6)$$

命题: 均方极限的运算法则

设 $X_n, X, Y_n, Y \in H$, $X_n \xrightarrow{\text{m.s.}} X$, $Y_n \xrightarrow{\text{m.s.}} Y$, 则

极限与期望可交换:

$$\lim_{n \rightarrow \infty} \mathbb{E}X_n = \mathbb{E}X = \mathbb{E} \lim_{n \rightarrow \infty} X_n, \quad \lim_{n \rightarrow \infty} \mathbb{E}|X_n|^2 = \mathbb{E}|X|^2 \quad (7.7)$$

极限与内积可交换:

$$\lim_{m, n \rightarrow \infty} \langle X_m, Y_n \rangle = \langle X, Y \rangle \quad (7.8)$$

线性:

$$\lim_{n \rightarrow \infty} (\alpha X_n + \beta Y_n) = \alpha X + \beta Y, \quad \forall \alpha, \beta \in \mathbb{R} \quad (7.9)$$

命题：收敛的充要条件

$$\lim_{n \rightarrow \infty} X_n = X \Leftrightarrow \lim_{m, n \rightarrow \infty} \langle X_m, X_n \rangle = C, C \in \mathbb{R} \quad (7.10)$$

且此时

$$\lim_{m, n \rightarrow \infty} \langle X_m, X_n \rangle = C = \mathbb{E}|X|^2 = \|X\|^2 \quad (7.11)$$

命题：均方收敛与依概率、分布收敛的关系

若序列 $\{X_n, n \geq 1\}$ 均方收敛

$$\lim_{n \rightarrow \infty} X_n \stackrel{\text{m.s.}}{=} X \quad (7.12)$$

则

依概率收敛:

$$\lim_{n \rightarrow \infty} X_n \stackrel{P}{=} X \quad (7.13)$$

依分布收敛:

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x) \quad (7.14)$$

其中 $x \in \mathbb{R}$ 是 $\mathbb{P}(X \leq x)$ 的连续点, 即

$$X_n \xrightarrow{d} X (n \rightarrow \infty) \quad (7.15)$$

7.2 均方分析

在引出均方极限之后, 就可以对过程的均方分析特性展开深入的讨论. 类似于数学分析中的概念, 本节介绍均方分析, 亦称为随机微积分, 即均方连续性, 均方可导, 均方积分

定义：二阶矩过程

设 $\{X_t : t \geq 0\}$ 为一随机过程, 若 $\forall t \geq 0, X_t \in H$, 则称 $\{X_t : t \geq 0\}$ 为二阶矩过程.

7.2.1 均方连续性

定义：均方连续

设二阶矩过程 $\{X_t : t \in T\}$ 对 $\forall t_0 \geq 0$ 有

$$\lim_{t \rightarrow t_0} X_t \stackrel{\text{m.s.}}{=} X_{t_0}, \quad (7.16)$$

即

$$\lim_{t \rightarrow t_0} X_t - X_{t_0} \stackrel{\text{m.s.}}{=} 0 \quad (7.17)$$

则称 X_t 在 t_0 点均方连续.

若 X_t 对 $\forall t \in T$ 都均方连续, 则称 $\{X_t : t \in T\}$ 在 T 上均方连续.

定理：均方连续的充要条件

记 $R(s, t) = \mathbb{E}(X_s X_t) = \langle X_s, X_t \rangle$, 则 $\{X_t : t \geq 0\}$ 在 t_0 点均方连续的充要条件为 $R(s, t)$ 在 (t_0, t_0) 点连续.

推论: 全定义域均方连续的充要条件

$\{X_t : t \in T\}$ 在 T 上均方连续的充要条件为 $R(s, t)$ 在 $\{(t, t) : t \in T\}$ 上二元连续.

推论: 全区域均方连续与对角线均方连续等价

若 $R(s, t)$ 在 $\{(t, t) : t \in T\}$ 上连续, 则它在 $T \times T$ 上连续. 即对协方差 $R(s, t)$ 而言, 它在整个区域 $T \times T$ 上连续与它在 $T \times T$ 的对角线上连续是等价的.

7.2.2 均方可导

定义: 均方导数与均方微分

称二阶矩过程 $\{X_t : t \in T\}$ 在 $t_0 \in T$ 点上均方可导, 若

$$\lim_{h \rightarrow 0} \frac{X_{t_0+h} - X_{t_0}}{h} \stackrel{\text{m.s.}}{=} X'_{t_0} \quad (7.18)$$

存在, 此时称 X'_{t_0} 与 $X'_{t_0} dt$ 分别为 X_t 在 t_0 点的均方导数与均方微分.

若 X_t 对 $\forall t \in T$ 均方可导, 则称 $\{X_t : t \in T\}$ 在 T 上是均方可导的, 此时记

$$\lim_{h \rightarrow 0} \frac{X_{t+h} - X_t}{h} \stackrel{\text{m.s.}}{=} X'_t \quad (7.19)$$

X'_t 与 $X'_t dt$ 分别称为 X_t 在 T 上的均方导数与均方微分.

定理: 均方可导判定准则

二阶矩过程 $\{X_t : t \in T\}$ 在 t 点均方可导的充要条件是

$$\lim_{h, l \rightarrow 0} \frac{R(t+h, t+l) - R(t+h, t) - R(t, t+l) + R(t, t)}{hl} \quad (7.20)$$

存在, 即要求函数 $R(s, t)$ 在 (t, t) 点广义二次可微.

定理: 均方可导的性质

若 $\{X_t : t \in T\}, \{X_{1,t}, X_{2,t} : t \in T\}$ 在 T 上均方可导, $f(t)$ 为一般函数且在 T 上可导, 则

可导必连续: $X_t, X_{1,t}, X_{2,t}$ 在 T 上均方连续

线性:

$$(c_1 X_{1,t} + c_2 X_{2,t})' = c_1 X'_{1,t} + c_2 X'_{2,t}, \quad \forall c_1, c_2 \in \mathbb{R} \quad (7.21)$$

乘积求导公式:

$$[f(t)X_t]' = f'(t)X_t + f(t)X'_t \quad (7.22)$$

期望与求导换序:

$$\mathbb{E}X'_t = (\mathbb{E}X_t)' \quad (7.23)$$

内积与求导换序:

$$\langle X'_s, X'_t \rangle = \frac{\partial^2 R(s, t)}{\partial s \partial t} \quad (7.24)$$

7.2.3 均方积分

定义: Riemann 均方积分

设 $\{X_t : t \in T\}$ 为二阶矩过程, $f(t)$, $t \in T$ 为定义在 T 上的函数, $[a, b] \in T$, 任取 $a = t_0 < t_1 < \cdots < t_n = b$, $\Delta t_k = t_k - t_{k-1}$, 令 $\lambda = \max_{1 \leq k \leq n} \Delta t_k$, 任取 $u_k \in [t_{k-1}, t_k]$. 若

$$\lim_{\lambda \rightarrow 0} \sum_{k=1}^n f(u_k) X_{u_k} \Delta t_k \stackrel{\text{m.s.}}{=} \int_a^b f(t) X_t dt \quad (7.25)$$

均方极限存在, 则称 $\int_a^b f(t) X_t dt$ 为 $f(t) X_t$ 在 $[a, b]$ 上的 Riemann 均方积分.

若

$$\lim_{m, n \rightarrow \infty} \int_a^b f(t) X_t dt \stackrel{\text{m.s.}}{=} \int_{-\infty}^{\infty} f(t) X_t dt \quad (7.26)$$

均方极限存在, 则称上述极限为 $f(t) X_t$ 在 $(-\infty, \infty)$ 上的 Riemann 均方积分.

定理: 均方可积的充分条件

若

$$\int_a^b \int_a^b f(s) f(t) R(s, t) ds dt \quad (7.27)$$

存在, 则

$$\int_a^b f(t) X_t dt \quad (7.28)$$

也存在.

推论: 无穷区间均方可积的充分条件

若

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(s) f(t) R(s, t) ds dt \quad (7.29)$$

存在, 则

$$\int_{-\infty}^{\infty} f(t) X_t dt \quad (7.30)$$

也存在.

定理: 均方积分的性质

设 $f(t) X_t$, $f_k(t) X_{k,t}$ 在 $[a, b]$ 上均方可积, 则

期望与积分换序:

$$\mathbb{E} \left[\int_a^b f(t) X_t dt \right] = \int_a^b f(t) \mathbb{E} X_t dt \quad (7.31)$$

$$\mathbb{E} \left[\int_a^b f(s) X_s ds \int_a^b f(t) X_t dt \right] = \int_a^b \int_a^b f(s) f(t) R(s, t) ds dt \quad (7.32)$$

求和与积分换序, 线性:

$$\int_a^b \sum_k c_k f_k(t) X_{k,t} dt = \sum_k \int_a^b c_k f_k(t) X_{k,t} dt, \quad \forall c_k \in \mathbb{R} \quad (7.33)$$

分段积分:

$$\int_a^c f(t)X_t dt + \int_c^b f(t)X_t dt = \int_a^b f(t)X_t dt, \quad \forall c \in [a, b] \quad (7.34)$$

定理: 均方连续与均方积分的关系

设 X_t 在 $[a, b]$ 上均方连续, 则

连续必可积: X_t 在 $[a, b]$ 上均方可积, 即 $\int_a^b f(t)X_t dt$ 存在

积分不等式:

$$\left\| \int_a^b f(t)X_t dt \right\| \leq \int_a^b f(s)\|X_s\| ds \quad (7.35)$$

变上限积分均方连续可导: 记

$$Y_t = \int_a^t f(u)X_u du \quad (7.36)$$

则 $\{Y_t : t \geq 0\}$ 在 $[a, b]$ 上均方连续, 均方可导且 $Y'_t = X_t$

推论: 均方积分 Newton-Leibniz 公式

设 X'_t 在 $[a, b]$ 上均方连续, 则

$$X_t - X_a = \int_a^t X'_s ds, \quad \forall t \in [a, b] \quad (7.37)$$

7.3 Itô 随机积分

定义: σ 域流

若可测空间 (Ω, \mathcal{F}) 上子 σ 域族 $\{\mathcal{F}_t : t \geq 0\}$ 使得 $\forall 0 \leq s < t < \infty$, 若 $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$, 则称 $\{\mathcal{F}_t : t \geq 0\}$ 是 (Ω, \mathcal{F}) 上的 σ 域流.

注释:

对给定的随机过程 $X = \{X_t : t \geq 0\}$, 其 σ 域流为 $\mathcal{F}_t^X = \sigma(X_s : 0 \leq s \leq t) = \sigma(\cup_{0 \leq s \leq t} X_s^{-1}\mathcal{B}_{\mathbb{R}})$, 其中 $X_s^{-1}\mathcal{B}_{\mathbb{R}} = \{\{\omega \in \Omega : X_s(\omega) \in A\} | A \in \mathcal{B}_{\mathbb{R}}\} \subset \mathcal{F}$.

σ 域流可看作信息流, 即 t 时刻前随机过程的全部信息.

定义: 关于 σ 域流适应

若可测空间 (Ω, \mathcal{F}) 上随机过程 $X = \{X_t : t \geq 0\}$, σ 域流 $\{\mathcal{F}_t : t \geq 0\}$ 使得 $\forall t \geq 0$, X_t 关于 \mathcal{F}_t 可测, 即 $\forall x \in \mathbb{R}$, $\{\omega \in \Omega : X_t(\omega) \in (-\infty, x]\} \subset \mathcal{F}_t$, 记为 $X_t \in \mathcal{F}_t$, 则称 X 是 $\{\mathcal{F}_t\}$ 适应的.

定义: 布朗运动

若概率空间 $(\Omega, \mathcal{F}, \mathbb{P})$ 上连续轨道适应过程 $\{B_t, \mathcal{F}_t : t \geq 0\}$ 满足

$$B_0 = 0$$

$\forall 0 \leq s < t$, $B_t - B_s$ 与 \mathcal{F}_s 独立, 即 $\forall A \in \mathcal{F}_s$, \mathbb{I}_A 与 $B_t - B_s$ 独立, 且

$$B_t - B_s \sim N(0, t - s) \quad (7.38)$$

则称 $B = \{B_t : t \geq 0\}$ 为 $\{\mathcal{F}_t\}$ 布朗运动.

注释:

$$\forall t \geq 0, \mathcal{F}_t^B = \sigma(B_s : \forall 0 \leq s \leq t) \subset \mathcal{F}_t$$

定义: 乘积空间与乘积 σ 域流

乘积空间定义为

$$[0, T] \times \Omega = \{(t, \omega) : 0 \leq t \leq T, \omega \in \Omega\}, \quad (7.39)$$

乘积 σ 域流定义为

$$\mathcal{B}_{[0, T]} \times \mathcal{F}_t = \sigma(\{B \times A : B \in \mathcal{B}_{[0, T]}, A \in \mathcal{F}_t\}) \quad (7.40)$$

事实上, 乘积 σ 域流为包含所有可测矩形的 σ 域流.

定义: 循序可测过程

若 $\{\mathcal{F}_t : t \geq 0\}$ 是 (Ω, \mathcal{F}) 上 σ 域流, (Ω, \mathcal{F}) 上随机过程 $X = \{X_t : t \geq 0\}$ 使得 $\forall t > 0$,

$$X : ([0, t] \times \Omega, \mathcal{B}_{[0, T]} \times \mathcal{F}_t) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}}) \quad (7.41)$$

可测, 则称 X 是 $\{\mathcal{F}_t\}$ 循序可测过程.

定义: 平方可积循序过程空间

$\forall T > 0$, 记 $\mathcal{L}_T^2 := \left\{ (\Omega, \mathcal{F}) \text{ 上循序可测过程 } g(t, \omega) : \int_0^T \mathbb{E}[g^2(t, \omega)] dt < \infty \right\} = L^2([0, T] \times \Omega, \mathcal{P}_g, \lambda \times \mathbb{P})$, $\lambda \times \mathbb{P}(B \times A) = \lambda(B) \times \mathbb{P}(A)$, λ 为 Lebesgue 测度, 定义里的积分也可写为

$$\int_0^T \mathbb{E}[g^2(t, \omega)] dt = \int_{\Omega} \int_0^T g^2(t, \omega) dt dP \quad (7.42)$$

记 $\mathcal{L}_2 := \bigcap_{T>0} \mathcal{L}_T^2 = \left\{ (\Omega, \mathcal{F}) \text{ 上循序可测过程 } g(t, \omega) : \forall T > 0, \int_0^T \mathbb{E}[g^2(t, \omega)] dt < \infty \right\}$.

定义: 准范数

$\forall g \in \mathcal{L}_2$, 定义准范数

$$\|g\|_2 := \frac{\|g\|_{2,n} \wedge 1}{2^n} \quad (7.43)$$

其中

$$\|g\|_{2,n}^2 = \int_0^n \mathbb{E}[g^2(t, \omega)] dt \quad (7.44)$$

定义: L_0

定义:

$$\mathcal{L}_0 := \left\{ g(t, \omega) = g_0(t, \omega) \mathbb{I}_{\{0\}}(t) + \sum_{k=0}^{\infty} g_k(t, \omega) \mathbb{I}_{(t_k, t_{k+1}]}(t) \right. \\ \left. : 0 = t_0 < t_1 < \cdots < t_n \uparrow \infty, g_0 \in \mathcal{F}_0, g_k \in \mathcal{F}_{t_k}, \mathbb{E}g_0^2 < \infty, \mathbb{E}g_k^2 < \infty \right\} \quad (7.45)$$

\mathcal{L}_0 在 $\|\cdot\|_2$ 下是 \mathcal{L}_2 线性稠密子集.

定义：简单过程的随机积分

设 $B = \{B_t, \mathcal{F}_t : t \geq 0\}$ 为布朗运动, $\forall g \in \mathcal{L}_0$,

$$g(t, \omega) = g_0(t, \omega) \mathbb{I}_{\{0\}}(t) + \sum_{k=0}^{\infty} g_k(t, \omega) \mathbb{I}_{(t_k, t_{k+1}]}(t) \quad (7.46)$$

定义 Itô 随机积分为

$$I_g(t) := \int_0^t g_s dB_s = \sum_{k=0}^{\infty} g_k(B_{t_{k+1} \wedge t} - B_{t_k \wedge t}) \quad (7.47)$$

给定 $t > 0$, 不妨将 t 加入序列 $\{t_k\}$, 则 $\forall 0 = t_0 < t_1 < \cdots < t_n = t, \forall 0 \leq s \leq t$,

$$g(s, \omega) = g_0(s, \omega) \mathbb{I}_{\{0\}}(s) + \sum_{k=0}^{n-1} g_k(s, \omega) \mathbb{I}_{(t_k, t_{k+1}]}(s), \quad (7.48)$$

则 Itô 随机积分

$$I_g(t) = \int_0^t g_s dB_s = \sum_{k=0}^{n-1} g_k(B_{t_{k+1}} - B_{t_k}) \quad (7.49)$$

Itô 积分性质-L0

设 $g \in \mathcal{L}_0$, 则有

线性: $\forall g_1, g_2 \in \mathcal{L}_0, \alpha_1, \alpha_2 \in \mathbb{R}$, 有

$$\int_0^t (\alpha_1 g_{1,s} + \alpha_2 g_{2,s}) dB_s = \alpha_1 \int_0^t g_{1,s} dB_s + \alpha_2 \int_0^t g_{2,s} dB_s \quad (7.50)$$

期望: $\mathbb{E}I_g(t) = 0$

二阶矩:

$$\mathbb{E}I_g^2(t) = \int_0^t \mathbb{E}g_s^2 ds \quad (7.51)$$

鞅: $\left\{ I_g(t) := \int_0^t g_s dB_s : t \geq 0 \right\}$ 关于 $\{\mathcal{F}_t\}$ 是鞅

定义: L2 随机积分

$\forall g \in \mathcal{L}_2$, 因 \mathcal{L}_0 在 $\|\cdot\|_2$ 下是 \mathcal{L}_2 线性稠密子集, 则 $\exists g_n \in \mathcal{L}_0, n \geq 1$, 使得

$$\lim_{n \rightarrow \infty} \|g_n - g\|_2 = 0 \quad (7.52)$$

从而

$$\lim_{m, n \rightarrow \infty} \|g_n - g_m\|_2 = 0 \quad (7.53)$$

而 $g_n - g_m \in \mathcal{L}_0$,

$$\begin{aligned} \|g_n - g_m\|_{2,t}^2 &= \int_0^t \mathbb{E}|g_{n,u} - g_{m,u}|^2 du \\ &= \mathbb{E} \left[\int_0^t (g_{n,u} - g_{m,u}) dB_u \right]^2 \\ &= \mathbb{E} \left(\int_0^t g_{n,u} dB_u - \int_0^t g_{m,u} dB_u \right)^2 \end{aligned} \quad (7.54)$$

可见 $\forall t \geq 0$, $\{I_{g_n}(t) : n \geq 0\}$ 是 $H = L_2(\Omega, \mathcal{F}, \mathbb{P})$ 中的 Cauchy 列, $\exists I_g(t)$ 使得

$$I_g(t) = L_2\text{-}\lim_{n \rightarrow \infty} I_{g_n}(t) = L_2\text{-}\lim_{n \rightarrow \infty} \int_0^t g_{n,u} dB_u \quad (7.55)$$

Itô 积分性质-L2

设 $g \in \mathcal{L}_2$, 则有

线性: $\forall g_1, g_2 \in \mathcal{L}_2$, $\alpha_1, \alpha_2 \in \mathbb{R}$, 有

$$\int_0^t (\alpha_1 g_{1,s} + \alpha_2 g_{2,s}) dB_s = \alpha_1 \int_0^t g_{1,s} dB_s + \alpha_2 \int_0^t g_{2,s} dB_s \quad (7.56)$$

期望: $\mathbb{E}I_g(t) = 0$

二阶矩:

$$\mathbb{E}I_g^2(t) = \int_0^t \mathbb{E}g_s^2 ds \quad (7.57)$$

内积:

$$\langle I_{g_1}(t), I_{g_2}(t) \rangle = \mathbb{E} \left[\int_0^t g_{1,u} dB_u \int_0^t g_{2,u} dB_u \right] = \int_0^t \mathbb{E}(g_{1,u} g_{2,u}) du \quad (7.58)$$

$$\langle I_g(s), I_g(t) \rangle = \mathbb{E} \left[\int_0^s g_u dB_u \int_0^t g_u dB_u \right] = \int_0^s \mathbb{E}g_u^2 du, \quad \forall 0 \leq s \leq t \quad (7.59)$$

分段积分:

$$\int_0^t g_u dB_u = \int_0^s g_u dB_u + \int_s^t g_u dB_u, \quad \forall s \in [0, t] \quad (7.60)$$

鞅性: $\left\{ I_g(t) := \int_0^t g_s dB_s : t \geq 0 \right\}$ 关于 $\{\mathcal{F}_t\}$ 是鞅

Doob 极大值不等式: 记 $X_t = \int_0^t g_u dB_u$, 则 $\forall t \geq 0$, $\lambda > 0, p \geq 1$ 有

$$\mathbb{P} \left(\max_{0 \leq u \leq t} |X_u| > \lambda \right) \leq \frac{\mathbb{E}|X_t|^p}{\lambda^p} \quad (7.61)$$

定理: 鞅表示定理

若 $\{X_t : t \geq 0\}$ 关于 $\{\mathcal{F}_t\}$ 是鞅, $X_t \in H$, 则必存在唯一适应过程 $\{g_t : t \geq 0\} \in \mathcal{L}_T^2$, 满足 $0 \leq t \leq T$,

$$X_t - X_0 = \int_0^t g_s dB_s. \quad (7.62)$$

例: 均方连续循序可测过程的 Itô 积分

设 $\{\phi_t : t \geq 0\}$ 均方连续, 关于 $\{\mathcal{F}_t\}$ 适应且循序可测, 定义

$$\phi_t^{(n)} := \sum_{k=0}^{n-1} \phi_{t_k^{(n)}} \mathbb{I}_{(t_k^{(n)}, t_{k+1}^{(n)}]}(t) + \phi_0 \mathbb{I}_{\{0\}}(t) \xrightarrow{\mathcal{L}_2} \phi_t \quad (n \rightarrow \infty) \quad (7.63)$$

其中 $\{t_k^{(n)} : 0 \leq k \leq n\}$ 是 $[0, t]$ 的分割, 令

$$\lambda_n := \max_{0 \leq k \leq n-1} \{t_{k+1}^{(n)} - t_k^{(n)}\} \rightarrow 0 \quad (n \rightarrow \infty) \quad (7.64)$$

从而

$$\int_0^t \phi_u dB_u = L_2\text{-}\lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \phi_{t_k^{(n)}} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}}). \quad (7.65)$$

上式也可以理解为 Itö 积分的定义.

7.4 Itö 过程与 Itö 公式

定义: Itö 过程

设随机过程 $X = \{X_t : t \geq 0\}$ 满足如下的 Itö 随机积分方程: 对 $\forall 0 \leq t_0 < t < T$, 有

$$X_t - X_{t_0} = \int_{t_0}^t b(s, X_s) ds + \int_{t_0}^t \sigma(s, X_s) dB_s, \quad (7.66)$$

或等价地写作 Itö 随机微分方程

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad (7.67)$$

其中 $b(t, X_t)$, $\sigma(t, X_t)$ 是二元连续函数, 且对 $\forall x \in \mathbb{R}$,

$$b(t, X_t) \in \mathcal{L}_T^1, \quad \sigma(t, X_t) \in \mathcal{L}_T^2, \quad (7.68)$$

则称 X 为 Itö 过程.

注释: Itö 过程与扩散过程

Itö 过程在 $b(t, X_t)$, $\sigma(t, X_t)$ 具有比较好的性质的情况下也是扩散过程, 其中 $b(t, X_t)$ 称为漂移系数, $\sigma(t, X_t)$ 称为扩散系数.

定理: Itö 公式

设随机过程 $X = \{X_t : t \geq 0\}$ 是 Itö 过程, $y = f(t, x)$ 是二元函数, 且具有连续偏导数

$$\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x^2}. \quad (7.69)$$

令 $Y_t := f(t, X_t)$, 则过程 $Y = \{Y_t : t \geq 0\}$ 也是随机过程, 且对 $\forall 0 \leq t_0 < t$ 满足如下的 Itö 积分方程

$$Y_t - Y_{t_0} = \int_{t_0}^t \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} b + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 \right] (s, X_s) ds + \int_{t_0}^t \left(\frac{\partial f}{\partial x} \sigma \right) (s, X_s) dB_s, \quad (7.70)$$

或等价的 Itö 微分方程

$$dY_t = \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} b + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2 \right] (t, X_t) dt + \left[\frac{\partial f}{\partial x} \sigma \right] (t, X_t) dB_t \quad (\text{a.e.}). \quad (7.71)$$

定理: 重对数律

$$\limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = 1 \quad (\text{a.s.}) \quad (7.72)$$

$$\liminf_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \ln \ln t}} = -1 \quad (\text{a.s.}) \quad (7.73)$$

$$\limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln(1/t)}} = 1 \quad (\text{a.s.}) \quad (7.74)$$

$$\liminf_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \ln \ln(1/t)}} = -1 \quad (\text{a.s.}) \quad (7.75)$$

粗略地说, $B_{s+\Delta t} - B_s \sim \sqrt{\Delta t}$, 即 $dB_t \sim \sqrt{dt}$.

Itô 公式理解

由 Itô 过程定义可知

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad (7.76)$$

令 $Y_t := f(t, X_t)$, 欲求其关于 dt, dB_t 的全微分, 由 $dB_t \sim \sqrt{dt}$ 可知 $f(t, x)$ 对 x 的偏导需要展开到二阶来获取那个与 dt 同阶的 dB_t^2 项.

简记 $f = f(t, X_t)$, 则 Itô 公式微分形式可以粗略理解为

$$\begin{aligned} dY_t &= df(t, X_t) \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}dX_t^2 \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(bdt + \sigma dB_t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(bdt + \sigma dB_t)^2 \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}bdt + \frac{\partial f}{\partial x}\sigma dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(\sigma^2 dt + o(dt)) \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}bdt + \frac{\partial f}{\partial x}\sigma dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}\sigma^2 dt \\ &= \left(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}b + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}\sigma^2 \right) dt + \left(\frac{\partial f}{\partial x}\sigma \right) dB_t \\ &= \left[\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}b + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}\sigma^2 \right] (t, X_t)dt + \left[\frac{\partial f}{\partial x}\sigma \right] (t, X_t)dB_t \end{aligned} \quad (7.77)$$

例: 使用 Itô 公式求解 Itô 积分

求证

$$\int_0^t B_s^2 dB_s = \frac{B_t^3}{3} - \int_0^t B_s ds. \quad (7.78)$$

证明: 与该积分方程对应的微分方程为

$$B_t^2 dB_t = d\left(\frac{B_t^3}{3}\right) - B_t dt \quad (7.79)$$

即

$$d\left(\frac{B_t^3}{3}\right) = B_t dt + B_t^2 dB_t \quad (7.80)$$

对比上式与 Itô 公式可猜测函数 $f(t, x) = x^3/3$, 下面证明之.

布朗运动本身即是 Itô 过程, 满足 Itô 微分方程

$$dB_t = 0 \times dt + 1 \times dB_t \quad (7.81)$$

令 $f(t, x) = x^3/3$, 则 $Y_t = f(t, B_t) = B_t^3/3$, 所以由 Itô 公式可得

$$\begin{aligned} dY_t &= d\left(\frac{B_t^3}{3}\right) \\ &= \left(0 + B_t^2 \cdot 0 + \frac{1}{2} 2B_t \cdot 1^2\right) dt + (B_t^2 \cdot 1) dB_t \\ &= B_t dt + B_t^2 dB_t \end{aligned} \quad (7.82)$$

对上式从 0 到 t 积分可得

$$\frac{B_t^3}{3} = \int_0^t B_s ds + \int_0^t B_s^2 dB_s \quad (7.83)$$

例: 求解布朗运动自身的 n 次 Itô 积分

$\forall n \geq 1$, 求下面积分

$$\int_0^t B_s^n dB_s. \quad (7.84)$$

解: 布朗运动本身即是 Itô 过程, 满足 Itô 微分方程

$$dB_t = 0 \times dt + 1 \times dB_t \quad (7.85)$$

令 $f(t, x) = x^{n+1}/(n+1)$, 则 $Y_t = f(t, B_t) = B_t^{n+1}/(n+1)$, 所以由 Itô 公式可得

$$\begin{aligned} dY_t &= d\left(\frac{B_t^{n+1}}{n+1}\right) \\ &= \left(0 + B_t^n \cdot 0 + \frac{1}{2} n B_t^{n-1} \cdot 1^2\right) dt + (B_t^n \cdot 1) dB_t \\ &= \frac{n}{2} B_t^{n-1} dt + B_t^n dB_t \end{aligned} \quad (7.86)$$

对上式从 0 到 t 积分可得

$$\frac{B_t^{n+1}}{n+1} = \int_0^t \frac{n}{2} B_s^{n-1} ds + \int_0^t B_s^n dB_s \quad (7.87)$$

即

$$\int_0^t B_s^n dB_s = \frac{B_t^{n+1}}{n+1} - \int_0^t \frac{n}{2} B_s^{n-1} ds. \quad (7.88)$$

例: 求随机过程满足的 Itô 微分方程

利用 Itô 公式, 求以下随机过程满足的 Itô 微分方程:

$$X_t = \exp(ut + \alpha B_t), \quad Y_t = \exp(t/2) \cos B_t. \quad (7.89)$$

解: 令 $f(t, x) = \exp(ut + \alpha x)$, $g(t, x) = \exp(t/2) \cos x$, 它们的相应偏导数为

$$\frac{\partial f}{\partial t} = u \exp(ut + \alpha x), \quad \frac{\partial f}{\partial x} = \alpha \exp(ut + \alpha x), \quad \frac{\partial^2 f}{\partial x^2} = \alpha^2 \exp(ut + \alpha x) \quad (7.90)$$

$$\frac{\partial g}{\partial t} = \frac{1}{2} \exp(t/2) \cos x, \quad \frac{\partial g}{\partial x} = -\exp(t/2) \sin x, \quad \frac{\partial^2 g}{\partial x^2} = -\exp(t/2) \cos x \quad (7.91)$$

所以由 Itô 公式可知

$$\begin{aligned} dX_t &= \left[u \exp(ut + \alpha B_t) + 0 + \frac{1}{2} \alpha^2 \exp(ut + \alpha B_t) \right] dt + \alpha \exp(ut + \alpha B_t) dB_t \\ &= \exp(ut + \alpha B_t) \left[\left(u + \frac{1}{2} \alpha^2 \right) dt + \alpha dB_t \right] \\ &= \left(u + \frac{1}{2} \alpha^2 \right) X_t dt + \alpha X_t dB_t \end{aligned} \quad (7.92)$$

$$\begin{aligned} dY_t &= \left[\frac{1}{2} \exp(t/2) \cos B_t + 0 - \frac{1}{2} \exp(t/2) \cos B_t \right] dt - \exp(t/2) \sin B_t dB_t \\ &= -\exp(t/2) \sin B_t dB_t \end{aligned} \quad (7.93)$$

定义: 多元 Itô 过程

设随机过程 $\mathbf{X} = \{\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{n,t}) : t \geq 0\}$ 满足如下的多元 Itô 随机微分方程

$$d\mathbf{X}_t = \mathbf{b}(t, \mathbf{X}_t)dt + \boldsymbol{\Sigma}(t, \mathbf{X}_t)d\mathbf{B}_t, \quad (7.94)$$

其中

$$\mathbf{b}(t, \mathbf{X}_t) = [b_1(t, \mathbf{X}_t), b_2(t, \mathbf{X}_t), \dots, b_n(t, \mathbf{X}_t)]^T, \quad (7.95)$$

$$\boldsymbol{\Sigma}(t, \mathbf{X}_t) = [\sigma_{ij}(t, \mathbf{X}_t)]_{n \times m}, \quad (7.96)$$

且 $b_i(t, \mathbf{X}_t) \in \mathcal{L}_T^1$, $\sigma_{ij}(t, \mathbf{X}_t) \in \mathcal{L}_T^2$,

$$\mathbf{B}_t = [B_{1,t}, B_{2,t}, \dots, B_{m,t}]^T, \quad (7.97)$$

则称 \mathbf{X} 为 n 维 Itô 过程.

定理: 多元 Itô 公式

设随机过程 $\mathbf{X} = \{\mathbf{X}_t : t \geq 0\}$ 是 n 维 Itô 过程,

$$\mathbf{f}(t, \mathbf{x}) = [f_1(t, \mathbf{x}), f_2(t, \mathbf{x}), \dots, f_d(t, \mathbf{x})]^T \quad (7.98)$$

是 $[0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}^d$ 上的函数, 其中 $\mathbf{x} = (x_1, x_2, \dots, x_n)$, 且具有连续偏导数

$$\frac{\partial f_k}{\partial t}, \frac{\partial f_k}{\partial x_i}, \frac{\partial^2 f_k}{\partial x_i \partial x_j}. \quad (7.99)$$

令 $\mathbf{Y}_t := \mathbf{f}(t, \mathbf{X}_t)$, 则过程 $\mathbf{Y} = \{\mathbf{Y}_t : t \geq 0\}$ 是 d 维 Itô 过程, 且对 $\forall 1 \leq k \leq d$ 满足如下的多元 Itô 微分方程

$$d\mathbf{Y}_{k,t} = \left[\frac{\partial f_k}{\partial t} + \frac{\partial f_k}{\partial x_i} b_i + \frac{1}{2} \frac{\partial^2 f_k}{\partial x_i \partial x_j} \sigma_{il} \sigma_{jl} \right] (t, \mathbf{X}_t) dt + \left[\frac{\partial f_k}{\partial x_i} \sigma_{il} \right] (t, \mathbf{X}_t) dB_{l,t}. \quad (7.100)$$

需要指出的是上式使用了 Einstein 求和约定, 即重复指标自动求和.

7.5 Itô 随机微分方程

7.5.1 常系数的线性随机微分方程

例: Ornstein-Uhlenbeck 方程

设 $\{X_t : t \geq 0\}$ 满足 Ornstein-Uhlenbeck 方程

$$dX_t = -bX_t dt + \sigma dB_t, \quad b, \sigma \in \mathbb{R}, \quad (7.101)$$

求解 X_t .

解: 方程两边同乘积分因子 e^{bt} 可得

$$e^{bt} dX_t + b e^{bt} X_t dt = \sigma e^{bt} dB_t, \quad (7.102)$$

令 $f(t, x) = e^{bt}x$, 由 Itô 公式可知

$$\begin{aligned} d(e^{bt} X_t) &= b e^{bt} X_t dt + e^{bt} dX_t \\ &= \sigma e^{bt} dB_t, \end{aligned} \quad (7.103)$$

两边从 t_0 到 t 积分可得

$$e^{bt} X_t - e^{bt_0} X_{t_0} = \int_{t_0}^t \sigma e^{bs} dB_s, \quad (7.104)$$

于是

$$X_t = e^{-b(t-t_0)} X_{t_0} + \int_{t_0}^t \sigma e^{-b(t-s)} dB_s. \quad (7.105)$$

7.5.2 简单的线性齐次随机微分方程

例: Black-Scholes 模型

设 X_t 为 t 时刻某种股票的价格, 在某些条件下它满足以下随机微分方程

$$dX_t = bX_t dt + \sigma X_t dB_t, \quad b, \sigma \in \mathbb{R}, \quad (7.106)$$

求解 X_t .

解: 方程两边同除 X_t 可得

$$\frac{dX_t}{X_t} = bdt + \sigma dB_t. \quad (7.107)$$

遂令 $f(t, x) = \ln x$, 则 $Y_t = \ln X_t$, 由 Itô 公式可知

$$\begin{aligned} dY_t &= d(\ln X_t) \\ &= \left(0 + \frac{1}{X_t} bX_t - \frac{1}{2} \frac{1}{X_t^2} \sigma^2 X_t^2\right) dt + \frac{1}{X_t} \sigma X_t dB_t \\ &= \left(b - \frac{1}{2} \sigma^2\right) dt + \sigma dB_t, \end{aligned} \quad (7.108)$$

两边从 0 到 t 积分可得

$$\begin{aligned} \ln X_t - \ln X_0 &= \int_0^t \left(b - \frac{1}{2} \sigma^2\right) dt + \int_0^t \sigma dB_t \\ &= \left(b - \frac{1}{2} \sigma^2\right) t + \sigma B_t, \end{aligned} \quad (7.109)$$

于是

$$X_t = X_0 \exp \left[\left(b - \frac{1}{2} \sigma^2\right) t + \sigma B_t \right]. \quad (7.110)$$

7.5.3 一般的线性非齐次随机微分方程

设 $\{X_t : t \geq 0\}$ 满足

$$dX_t = [b_1(t)X_t + b_2(t)]dt + [\sigma_1(t)X_t + \sigma_2(t)]dB_t, \quad (7.111)$$

求解 X_t .

解: 在求解一般线性非齐次随机微分方程之前, 先求解一类稍简单的狭义随机微分方程. 与线性常微分方程求解方法类似, 由简入繁, 先求齐次解, 再求非齐次解, 最大的不同在于随机微分方程中复合函数的求导法则由 Itô 公式决定, 多了二阶偏导项, 作微分运算时尤其需要谨慎.

狭义随机微分方程

先求解以下简单齐次线性随机微分方程

$$dX_t = b_1(t)X_t dt, \quad (7.112)$$

由 Black-Scholes 模型求解结果易得

$$X_t = X_{t_0} \exp \left(\int_{t_0}^t b_1(t) dt \right). \quad (7.113)$$

再求解以下狭义随机微分方程

$$dX_t = [b_1(t)X_t + b_2(t)]dt + \sigma_2(t)dB_t, \quad (7.114)$$

移项变形可得

$$dX_t - b_1(t)X_t dt = b_2(t)dt + \sigma_2(t)dB_t, \quad (7.115)$$

由齐次解推测, 若两边同乘以下积分因子

$$\rho_{t_0}^{-1}(t) = \exp\left(-\int_{t_0}^t b_1(t)dt\right) \quad (7.116)$$

则可以通过 Itô 公式解上述微分方程, 即

$$\rho_{t_0}^{-1}(t)[dX_t - b_1(t)X_t]dt = \rho_{t_0}^{-1}(t)[b_2(t)dt + \sigma_2(t)dB_t]. \quad (7.117)$$

令 $f(t, x) = \rho_{t_0}^{-1}(t)x$, 则由 Itô 公式可知

$$\begin{aligned} d[\rho_{t_0}^{-1}(t)X_t] &= \rho_{t_0}^{-1}(t)[dX_t - b_1(t)X_t]dt \\ &= \rho_{t_0}^{-1}(t)[b_2(t)dt + \sigma_2(t)dB_t], \end{aligned} \quad (7.118)$$

两边从 t_0 到 t 积分可得

$$\begin{aligned} &\rho_{t_0}^{-1}(t)X_t - \rho_{t_0}^{-1}(t_0)X_{t_0} \\ &= \rho_{t_0}^{-1}(t)X_t - X_{t_0} \\ &= \int_{t_0}^t \rho_{t_0}^{-1}(s)b_2(s)ds + \int_{t_0}^t \rho_{t_0}^{-1}(s)\sigma_2(s)dB_s, \end{aligned} \quad (7.119)$$

故

$$X_t = \rho_{t_0}(t) \left[X_{t_0} + \int_{t_0}^t \rho_{t_0}^{-1}(s)b_2(s)ds + \int_{t_0}^t \rho_{t_0}^{-1}(s)\sigma_2(s)dB_s \right]. \quad (7.120)$$

一般线性非齐次随机微分方程

先求解以下齐次线性随机微分方程

$$dX_t = b_1(t)X_t dt + \sigma_1(t)X_t dB_t. \quad (7.121)$$

由 Black-Scholes 模型求解结果易得

$$X_t = X_{t_0} \exp \left[\int_{t_0}^t \left(b_1(t) - \frac{1}{2}\sigma_1^2(t) \right) dt + \int_{t_0}^t \sigma_1(t)dB_t \right]. \quad (7.122)$$

与前述求解过程类似, 下面仍然使用积分因子的思想求解随机微分方程. 记积分因子

$$\rho_{t_0}(t) = \exp \left[\int_{t_0}^t \left(b_1(t) - \frac{1}{2}\sigma_1^2(t) \right) dt + \int_{t_0}^t \sigma_1(t)dB_t \right], \quad (7.123)$$

则积分因子满足以下 Itô 微分方程

$$d\rho_{t_0}(t) = b_1(t)\rho_{t_0}(t)dt + \sigma_1(t)\rho_{t_0}(t)dB_t. \quad (7.124)$$

令 $f(t, x) = 1/x$, 则由 Itô 公式可知 $\rho_{t_0}^{-1}(t)$ 满足以下 Itô 微分方程

$$\begin{aligned} d\rho_{t_0}^{-1}(t) &= \left[0 - \frac{1}{\rho_{t_0}^2(t)}b_1(t)\rho_{t_0}(t) + \frac{1}{2}\frac{2}{\rho_{t_0}^3(t)}\sigma_1^2(t)\rho_{t_0}^2(t) \right] dt - \frac{1}{\rho_{t_0}^2(t)}\sigma_1(t)\rho_{t_0}(t)dB_t \\ &= \rho_{t_0}^{-1}(t)[\sigma_1^2(t) - b_1(t)]dt - \rho_{t_0}^{-1}(t)\sigma_1(t)dB_t. \end{aligned} \quad (7.125)$$

又

$$dX_t = [b_1(t)X_t + b_2(t)]dt + [\sigma_1(t)X_t + \sigma_2(t)]dB_t, \quad (7.126)$$

令 $g(t, x_1, x_2) = x_1x_2$, 则由二维 Itô 公式可知

$$\begin{aligned} d[\rho_{t_0}^{-1}(t)X_t] &= \rho_{t_0}^{-1}(t) \left\{ X_t[\sigma_1^2(t) - b_1(t)] + [b_1(t)X_t + b_2(t)] - \sigma_1(t)[\sigma_1(t)X_t + \sigma_2(t)] \right\} dt \\ &\quad + \left\{ X_t[-\rho_{t_0}^{-1}(t)\sigma_1(t)] + \rho_{t_0}^{-1}(t)[\sigma_1(t)X_t + \sigma_2(t)] \right\} dB_t \\ &= \rho_{t_0}^{-1}(t)[b_2(t) - \sigma_1(t)\sigma_2(t)]dt + \rho_{t_0}^{-1}(t)\sigma_2(t)dB_t. \end{aligned} \quad (7.127)$$

两边从 t_0 到 t 积分可得

$$\begin{aligned} \rho_{t_0}^{-1}(t)X_t - \rho_{t_0}^{-1}(t_0)X_{t_0} &= \rho_{t_0}^{-1}(t)X_t - X_{t_0} \\ &= \int_{t_0}^t \rho_{t_0}^{-1}(s)[b_2(s) - \sigma_1(s)\sigma_2(s)]ds + \int_{t_0}^t \rho_{t_0}^{-1}(s)\sigma_2(s)dB_s, \end{aligned} \quad (7.128)$$

故

$$X_t = \rho_{t_0}(t) \left[X_{t_0} + \int_{t_0}^t \rho_{t_0}^{-1}(s)[b_2(s) - \sigma_1(s)\sigma_2(s)]ds + \int_{t_0}^t \rho_{t_0}^{-1}(s)\sigma_2(s)dB_s \right]. \quad (7.129)$$

例：求解 Itô 微分方程

通用公式

设 $\{X_t : t \geq 0\}$ 满足

$$dX_t = [b_1(t)X_t + b_2(t)]dt + [\sigma_1(t)X_t + \sigma_2(t)]dB_t, \quad (7.130)$$

则

$$\rho_{t_0}(t) = \exp \left[\int_{t_0}^t \left(b_1(t) - \frac{1}{2}\sigma_1^2(t) \right) dt + \int_{t_0}^t \sigma_1(t)dB_t \right], \quad (7.131)$$

$$X_t = \rho_{t_0}(t) \left[X_{t_0} + \int_{t_0}^t \rho_{t_0}^{-1}(s)[b_2(s) - \sigma_1(s)\sigma_2(s)]ds + \int_{t_0}^t \rho_{t_0}^{-1}(s)\sigma_2(s)dB_s \right]. \quad (7.132)$$

$$(1) \quad dX_t = -X_t dt + e^{-t} dB_t$$

解: $b_1(t) = -1$, $b_2(t) = 0$, $\sigma_1(t) = 0$, $\sigma_2(t) = e^{-t}$, 则

$$\rho_{t_0}(t) = \exp \left[\int_{t_0}^t (-1 - 0) dt + \int_{t_0}^t 0 dB_t \right] = e^{-(t-t_0)}, \quad (7.133)$$

$$\begin{aligned} X_t &= e^{-(t-t_0)} \left[X_{t_0} + \int_{t_0}^t e^{(s-t_0)}[0 - 0]ds + \int_{t_0}^t e^{(s-t_0)}e^{-s}dB_s \right] \\ &= e^{-(t-t_0)} [X_{t_0} + e^{-t_0}(B_t - B_{t_0})] \\ &= e^{-(t-t_0)}X_{t_0} + e^{-t}(B_t - B_{t_0}) \end{aligned} \quad (7.134)$$

$$(2) \quad dX_t = \gamma dt + \alpha X_t dB_t, \quad \gamma, \alpha \in \mathbb{R}$$

解: $b_1(t) = 0$, $b_2(t) = \gamma$, $\sigma_1(t) = \alpha$, $\sigma_2(t) = 0$, 则

$$\rho_{t_0}(t) = \exp \left[\int_{t_0}^t \left(0 - \frac{1}{2} \alpha^2 \right) dt + \int_{t_0}^t \alpha dB_t \right] = \exp \left[-\frac{1}{2} \alpha^2 (t - t_0) + \alpha (B_t - B_{t_0}) \right], \quad (7.135)$$

$$\begin{aligned} X_t &= \rho_{t_0}(t) \left[X_{t_0} + \int_{t_0}^t \gamma \rho_{t_0}^{-1}(s) ds \right] \\ &= \exp \left[-\frac{1}{2} \alpha^2 (t - t_0) + \alpha (B_t - B_{t_0}) \right] \left\{ X_{t_0} + \int_{t_0}^t \gamma \exp \left[\frac{1}{2} \alpha^2 (s - t_0) - \alpha (B_s - B_{t_0}) \right] ds \right\} \\ &= \exp \left(-\frac{1}{2} \alpha^2 t + \alpha B_t \right) \left[\exp \left(\frac{1}{2} \alpha^2 t_0 - \alpha B_{t_0} \right) X_{t_0} + \gamma \int_{t_0}^t \exp \left(\frac{1}{2} \alpha^2 s - \alpha B_s \right) ds \right] \end{aligned} \quad (7.136)$$

(3) $dX_t = (e^{-t} + X_t)dt + \sigma X_t dB_t$, $\sigma > 0$

解: $b_1(t) = 1$, $b_2(t) = e^{-t}$, $\sigma_1(t) = \sigma$, $\sigma_2(t) = 0$, 则

$$\rho_{t_0}(t) = \exp \left[\int_{t_0}^t \left(1 - \frac{1}{2} \sigma^2 \right) dt + \int_{t_0}^t \sigma dB_t \right] = \exp \left[\left(1 - \frac{1}{2} \sigma^2 \right) (t - t_0) + \sigma (B_t - B_{t_0}) \right], \quad (7.137)$$

$$\begin{aligned} X_t &= \rho_{t_0}(t) \left[X_{t_0} + \int_{t_0}^t e^{-s} \rho_{t_0}^{-1}(s) ds \right] \\ &= \rho_{t_0}(t) \left\{ X_{t_0} + \int_{t_0}^t e^{-s} \exp \left[-\left(1 - \frac{1}{2} \sigma^2 \right) (s - t_0) - \sigma (B_s - B_{t_0}) \right] ds \right\} \\ &= \rho_{t_0}(t) \left\{ X_{t_0} + \exp \left[\left(1 - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right] \int_{t_0}^t \exp \left[\left(\frac{1}{2} \sigma^2 - 2 \right) s - \sigma B_s \right] ds \right\} \end{aligned} \quad (7.138)$$

(4) $dX_t = f(t)B_t dt$

解: 布朗运动本身即为 Itô 过程, 有 Itô 微分方程

$$dB_t = 0 \times dt + 1 \times dB_t. \quad (7.139)$$

函数 f 在 $[0, +\infty)$ 上连续则可积, 令其原函数为

$$F(t) := \int_0^t f(s) ds. \quad (7.140)$$

令 $h(t, x) = F(t)x$, 其相应偏导数为

$$\frac{\partial h}{\partial t} = f(t)x, \quad \frac{\partial h}{\partial x} = F(t), \quad \frac{\partial^2 h}{\partial x^2} = 0, \quad (7.141)$$

令 $W_t = F(t)B_t$, 由 Itô 公式可知

$$\begin{aligned} dW_t &= \left[f(t)B_t + F(t) \cdot 0 + \frac{1}{2} \cdot 0 \cdot 1^2 \right] dt + [F(t) \cdot 1] dB_t \\ &= f(t)B_t dt + F(t) dB_t. \end{aligned} \quad (7.142)$$

注意到 B_t 零初值即 $B_0 = 0$, 那么上式两边从 0 到 t 积分可得

$$\begin{aligned} W_t - W_0 &= F(t)B_t - 0 \\ &= \int_0^t f(s)B_s ds + \int_0^t F(s)dB_s, \end{aligned} \quad (7.143)$$

故由 Itô 积分的一阶矩和二阶矩性质可知

$$\begin{aligned} X_t &= \int_0^t f(s)B_s ds \\ &= F(t)B_t - \int_0^t F(s)dB_s \\ &= \int_0^t [F(t) - F(s)]dB_s \\ &\sim N\left(0, \int_0^t [F(t) - F(s)]^2 ds\right), \end{aligned} \quad (7.144)$$

即 X_t 的概率分布是期望为 0, 方差为 $\int_0^t [F(t) - F(s)]^2 ds$ 的正态分布.

$$(5) \quad dY_t = f(t)B_t dB_t$$

解: 布朗运动本身即为 Itô 过程, 有 Itô 微分方程

$$dB_t = 0 \times dt + 1 \times dB_t. \quad (7.145)$$

令 $g(t, x) = f(t)x^2/2$, 其相应偏导数为

$$\frac{\partial g}{\partial t} = f'(t)\frac{x^2}{2}, \quad \frac{\partial g}{\partial x} = f(t)x, \quad \frac{\partial^2 g}{\partial x^2} = f(t), \quad (7.146)$$

令 $Z_t = f(t)B_t^2/2$, 由 Itô 公式可知

$$\begin{aligned} dZ_t &= \left[f'(t)\frac{B_t^2}{2} + f(t)B_t \cdot 0 + \frac{1}{2}f(t) \cdot 1^2 \right] dt + [f(t)B_t \cdot 1]dB_t \\ &= \frac{1}{2}[f'(t)B_t^2 + f(t)]dt + f(t)B_t dB_t. \end{aligned} \quad (7.147)$$

注意到 B_t 零初值即 $B_0 = 0$, 那么上式两边从 0 到 t 积分可得

$$\begin{aligned} Z_t - Z_0 &= \frac{1}{2}f(t)B_t^2 - 0 \\ &= \int_0^t \frac{1}{2}[f'(s)B_s^2 + f(s)]ds + \int_0^t f(s)B_s dB_s, \end{aligned} \quad (7.148)$$

故

$$Y_t = \int_0^t f(s)B_s dB_s = \frac{1}{2} \left\{ f(t)B_t^2 - \int_0^t [f(s) + f'(s)B_s^2] ds \right\}. \quad (7.149)$$

7.6 解的存在性和唯一性问题

定理: 随机微分方程解的存在和唯一性

设对 $\forall t \geq 0, x, y \in \mathbb{R}, K > 0, b(t, x), \sigma(t, x)$ 满足整体 Lipschitz 条件

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|, \quad (7.150)$$

和线性增长条件

$$|b(t, x)| + |\sigma(t, x)| \leq K(1 + |x|). \quad (7.151)$$

$B = \{B_t, \mathcal{F}_t^B : t \geq 0\}$ 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的布朗运动, ξ 是 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的随机变量且与布朗运动独立, $\mathbb{E}|\xi|^2 < \infty$, 则随机微分方程

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = \xi, \quad (7.152)$$

存在唯一强解, 而且 $\exists C > 0, C$ 可以依赖于 (K, T) , 使得 $\forall 0 \leq t \leq T$, 有

$$\mathbb{E}|X_t|^2 \leq C(1 + \mathbb{E}|\xi|^2)e^{Ct}. \quad (7.153)$$