

西安交通大学数学研究生教学丛书

随机逼近及自适应算法

聂赞坎 徐宗本 编著

科学出版社

北 京

内 容 简 介

本书系统地介绍了随机逼近理论,着重介绍了几类自适应算法以及随机逼近理论对它们的应用.作为一本研究生教材,本书难度适中,注重基础知识的讲述,深入浅出,易于自学.

读者对象:高等院校数学系高年级学生,信息科学、计算机科学、运筹学、应用数学等专业的研究生.

图书在版编目(CIP)数据

随机逼近及自适应算法/聂赞坎,徐宗本编著.—北京:科学出版社,2003.1

(西安交通大学数学研究生教学丛书)

ISBN 7-03-010534-6

I .随… II .①聂… ②徐… III .①随机逼近-研究生-教材②自适应控制-算法理论-研究生-教材 IV .①O174.41 ②TP273

中国版本图书馆 CIP 数据核字(2002)第 047935 号

责任编辑:林 鹏 杨 波/责任校对:宣 慧

责任印制:安春生/封面设计:高海英 王 浩

科学出版社发行 各地新华书店经销

*

2003 年 1 月第 一 版 开本: B5 (720×1000)

2003 年 1 月第一次印刷 印张: 9 1/2

印数: 1—3 000 字数: 176 000

定价: 19.00 元

(如有印装质量问题,我社负责调换〈环伟〉)

前 言

随机逼近算法可以理解为是利用观测值估计未知函数的极值或未知方程解的自适应问题求解技术.它起源于 20 世纪 50 年代初 Ribbins 和 Monro 所提出的求未知函数零点的一个递推算法(称之为 RM 算法)以及此后 Kiefer 和 Wolfowitz 等人有关未知函数极值问题的研究.随机逼近理论(即随机逼近算法的理论基础与分析)近几十年来得到了飞速发展,已成为数理统计与人工智能的交叉分支,并已广泛地应用于系统辨识、自适应控制、模式识别、自适应滤波、神经网络等领域.有关随机逼近理论的研究,近几年出现了 Albert Benveniste、Lennart Ljung、陈翰馥等国内外著名学者的专著,它们对随机逼近理论的方方面面作了系统、深入的阐述.这些著作虽各有深入讨论的侧重面,但由于篇幅过大,论题过多,对初学者来说,不易掌握.本书的编写目的在于向读者(特别是数学系研究生)提供一本难度适中、内容基本、易于自学的教学用书.因此,本书的前一部分(第一章~第四章)着重介绍随机逼近基础理论中的核心问题(随机逼近算法的收敛性)和基本的收敛性结果,以及这些基本结果在一些具体自适应性算法中的应用.本书后一部分集中介绍近几年随高性能计算机发展而出现的几类自适应算法(遗传算法、模拟退火算法、主成分分析神经网络算法)以及随机逼近理论对它们的应用.

遗传算法是模拟自然界生物进化过程与机制求解极值问题的一类自适应算法.它的产生归功于美国 Michigan 大学的 Holland 在 20 世纪 60 年代末 70 年代初的开创性工作.他不仅设计了遗传算法的模拟与操作原理,而且更重要的是运用统计决策理论对遗传算法的搜索机理进行了理论分析,从而为遗传算法的发展奠定了基础.近十几年来,遗传算法无论是在应用上,还是在基础理论上都取得了长足发展,已成为信息科学、计算机科学、运筹学和应用数学等诸多学科所共同关注的热点研究领域.但相对卓有成效的、广泛的实际应用而言,遗传算法的数学基础理论研究还相对滞后,还不能说很完善或很深入.近期研究主要是围绕如何提高算法效率和建立算法理论基础方面.本书第五章和第六章专门讨论遗传算法的随机过程分析及其收敛性分析,其中大部分内容是西安交通大学理学院信息与系统科学研究所师生们近期的科研成果.

模拟退火算法源于对固体退火过程的模拟,是解全局优化问题(特别是组合优化问题)的一种通用自适应方法.该类方法以随机方法为基础,结合一系列相关迭代算法,既克服了迭代算法的一些本质缺陷(如收敛的局部性,通常只收敛到问题的局部极小),而又保持了它们的突出优点,如快速局部收敛和

所获得的解不依赖于初始状态等.本书第七章专门讨论马尔可夫链模型的模拟退火算法,介绍齐次算法和非齐次算法的一些渐近收敛性结果.

主成分分析是一种经典的统计技术,它用于分析多变量统计观察的方差结构,它与估计理论中的最小方差技术,时间序列分析中的 Karhunen Loeve 变换和数值分析中的奇异值分解密切相关.这些方法在信号处理、图像编码、信息压缩等领域有十分基本而重要的应用.近年来,人们利用神经网络技术求解主成分显示出极大的优越性,许多用于求解主成分的神经网络算法问世.这些主成分分析神经网络算法的突出优点是它们的自组织性和自适应性,而且易于硬件实现.本书最后一章介绍与此相关的一些简单、通用算法,并且作为随机逼近理论的应用,研究这些算法的收敛性质.

全书共分八章.第一章给出本书必需的概率论方面的预备知识.第二章介绍研究随机逼近算法的三种常用方法.第三章讨论形式非常一般的具有局部有界矩随机逼近算法的几乎必然收敛性.第四章将第三章的收敛性结果应用于一些具体的适应性算法,获得这些算法的几乎必然收敛性.第五章介绍遗传算法的基本概念、过程分析和遗传算法的 Markov 链模型.第六章对抽象的(算子型)遗传算法在各种意义下的收敛性进行了讨论,并对两类特殊类型遗传算法的收敛性进行了细致分析.第七章建立模拟退火算法的马尔可夫链模型,介绍齐次算法和非齐次算法的一些渐近收敛性结果.第八章介绍一些简单、通用的主成分分析神经网络算法,并利用随机逼近理论研究这些算法的收敛性.

本书的主要内容曾经在西安交通大学理学院应用数学系硕士生的课程上讲授过,有不少内容是直接由文献改写的.限于作者自身水平,难免有遗漏与不妥之处,恳请读者批评与指正.

编 著

2002 年 8 月

于西安交通大学

目 录

第一章 预备知识	1
§ 1.1 概率论的若干基本概念	1
1.1.1 随机变量及其分布	1
1.1.2 随机变量列的收敛性	1
1.1.3 随机变量的期望和条件期望	2
1.1.4 条件期望的基本性质	4
§ 1.2 离散参数鞅	8
1.2.1 停时	8
1.2.2 鞅	9
1.2.3 离散鞅的基本不等式	10
1.2.4 离散鞅的收敛定理	14
1.2.5 Doob 停时定理	15
§ 1.3 马尔可夫链	17
1.3.1 马尔可夫链的定义及其转移概率	17
1.3.2 状态的分类	18
1.3.3 状态空间的分类	22
1.3.4 $p_{ij}^{(n)}$ 的渐近性质与平稳分布	23
1.3.5 离散时间连续状态的马尔可夫链	26
第二章 随机逼近算法的分析方法	29
§ 2.1 随机逼近算法	29
§ 2.2 鞅方法	30
§ 2.3 常微分方程方法	33
§ 2.4 Lyapunov 函数方法	36
第三章 具有局部有界矩随机逼近算法的几乎必然收敛性	41
§ 3.1 一般算法的引进	41
3.1.1 算法模型	41
3.1.2 例子	42
3.1.3 关于 H, ρ_n 和 Π 的一般假设	43
3.1.4 例子(续 3.1.2)	44
§ 3.2 一般算法的分解	46
§ 3.3 L^2 估计	49
§ 3.4 通过常微分方程的解作算法的逼近	54

§ 3.5 算法的渐近分析	57
§ 3.6 收敛定理的另一种叙述	60
§ 3.7 一个全局收敛性定理	61
§ 3.8 一些算法的 L^2 收敛速度	65
3.8.1 Robbins-Monro 算法	65
3.8.2 一般算法的局部 L^2 上界	67
第四章 应用	70
§ 4.1 马尔可夫链的几何遍历性	70
4.1.1 预备引理	70
4.1.2 不变概率与 Poisson 方程的解	71
4.1.3 由 $L_i(p)$ 到 $L_i(p)$ 的连续转移函数 P 的情形	73
§ 4.2 依赖于参数 θ 的马尔可夫链	74
4.2.1 v_θ 关于 θ 的 Holder 正则性	74
4.2.2 定理 4.2.1 的意义	75
4.2.3 定理 4.2.1 的证明	75
4.2.4 v_θ 关于 θ 为 Lipschitz 的情形	77
4.2.5 定理 4.2.2 的证明	78
4.2.6 转移概率 P_θ 不依赖于 θ 情形	79
§ 4.3 线性动力系统	80
4.3.1 假设和记号	80
4.3.2 预备结果	81
4.3.3 P_θ° 的性质	82
4.3.4 验证假设 (A.4)	84
§ 4.4 例子	85
4.4.1 接收信号的马尔可夫表示	85
4.4.2 横向均衡器, 学习阶段	86
4.4.3 最小二乘算法	86
第五章 遗传算法	89
§ 5.1 基本概念	89
§ 5.2 遗传算子及其性质	91
§ 5.3 遗传机制的过程分析	94
§ 5.4 遗传算法的马氏链模型	100
第六章 抽象遗传算法及其收敛性的一般理论	104
§ 6.1 演化算子及其特征数	104
6.1.1 选择算子	105
6.1.2 变异算子	107
6.1.3 杂交算子	107

§ 6.2 遗传算法收敛性的一般理论	108
6.2.1	110
6.2.2	113
§ 6.3 两类特殊类型遗传算法的收敛性	115
第七章 模拟退火算法	119
§ 7.1 模拟退火算法的数学模型	119
7.1.1 算法介绍	119
7.1.2 算法的数学模型	120
§ 7.2 齐次算法的渐近收敛性	121
7.2.1 平稳分布的存在性	122
7.2.2 平稳分布的收敛性	123
§ 7.3 非齐次算法的渐近收敛性	126
第八章 主成分分析神经网络算法	130
§ 8.1 主成分分析	130
§ 8.2 主成分分析神经网络算法	132
8.2.1 Hebbian 规则	133
8.2.2 对于单个主成分的 Oja 规则	134
8.2.3 广义的 Hebbian 算法 (GHA)	138
8.2.4 多分量的子空间规则	139
参考文献	142

第一章 预备知识

本章给出阅读本书所必需的若干预备知识,包括概率论、离散参数鞅和马尔可夫链等领域中与本书有关的一些事实与概念.

§ 1.1 概率论的若干基本概念

1.1.1 随机变量及其分布

设 $(\Omega, \mathcal{F}, P)$ 为一概率空间, X 是在 Ω 上定义的实值函数,记 $X^{-1}(B) = \{\omega; X(\omega) \in B\}$, $B \subset R$, $X^{-1}(B)$ 称为集合 B 的原像.

定义 1.1.1 定义在可测空间 $(\Omega, \mathcal{F})$ 上的实函数 X ,如果对数直线上任意 Borel 集 B , $X^{-1}(B) \in \mathcal{F}$, 则称 X 为 $\mathcal{F}$ 可测, 概率空间上的可测函数也称为随机变量.

对可测函数 X , 集系 $\{X^{-1}(B); B \in \mathcal{B}(R)\}$ 构成一个 σ 域, 称为由 X 生成的 σ 域, 记为 $\sigma(X)$, 其中 $\mathcal{B}(R)$ 是 R 上的 Borel 域. 对于随机变量 X , 定义函数

$$F(x) = P\{\omega; X(\omega) \leq x\}, \quad x \in R,$$

$F(x)$ 称为 X 的分布函数.

1.1.2 随机变量列的收敛性

设 $\{X, X_n\}$ 为一列随机变量.

(1) 如果对于任意正数 ϵ , 有 $\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$, 则称 $\{X_n\}$ 依概率收敛到 X , 记作 $X_n \xrightarrow{P} X, n \rightarrow \infty$.

(2) 如果 $P(\lim_{n \rightarrow \infty} X_n = X) = 1$, 则称随机变量列 X_n 概率 1 收敛到 X , 或称随机变量列 X_n 几乎处处收敛到 X , 记作 $X_n \xrightarrow{a.s} X, n \rightarrow \infty$.

(3) 如果 $\lim_{n \rightarrow \infty} E(|X_n - X|^2) = 0$, 则称随机变量列 X_n 均方收敛到 X , 记作 $X_n \xrightarrow{L^2} X, n \rightarrow \infty$.

上面各种收敛性有如下关系

$$X_n \xrightarrow{a.s} X, n \rightarrow \infty \Rightarrow X_n \xrightarrow{P} X, n \rightarrow \infty \Leftarrow X_n \xrightarrow{L^2} X, n \rightarrow \infty.$$

$$X_n \xrightarrow{P} X, n \rightarrow \infty \Rightarrow \text{存在子列 } X_{n_k} \xrightarrow{a.s.} X, k \rightarrow \infty.$$

1.1.3 随机变量的期望和条件期望

定义 1.1.2 设随机变量 $X \geq 0$, a.s., 记

$$A_{ni} = \{ \omega : i2^{-n} < X(\omega) \leq (i+1)2^{-n} \}, \quad n, i = 1, 2, \dots$$

X 的数学期望 EX 定义为

$$EX = \int_{\Omega} X dP = \lim_{n \rightarrow \infty} \left[\sum_{i=0}^{n2^n-1} i2^{-n} P(A_{ni}) + nP(X > n) \right].$$

对不是非负的随机变量 X , 可表为正部 $X^+ = \max(X, 0)$ 与负部 $X^- = \max(-X, 0)$ 的差 $X = X^+ - X^-$, 如果 $EX^+ < \infty$ 或 $EX^- < \infty$, 则定义 X 的数学期望 $EX = EX^+ - EX^-$; 当 $EX^+ < \infty, EX^- < \infty$ 时, 称 X 是可积的; 当 EX 有定义时, 它还可表为 Lebesgue-Stieltjes 积分

$$EX = \int_{-\infty}^{\infty} x dF(x).$$

由初等概率论知, 在事件 $A (P(A) > 0)$ 发生的条件下, 事件 B 发生的条件概率定义为

$$P(B|A) = P(AB)/P(A), \quad (1.1.1)$$

如果 $P(A) = 0$, 就规定 $P(B|A) = 0$.

固定 $A (P(A) > 0)$, $P(\cdot|A)$ 是 $\mathcal{F}$ 上的概率测度, (1.1.1) 式可写成

$$\int I_B(\omega) P(d\omega|A) = \frac{1}{P(A)} \int_A I_B(\omega) P(d\omega),$$

其中 I_B 表示集 B 的示性函数. 因此随机变量 X 关于 A 的条件数学期望可定义为

$$E(X|A) = \int_{\Omega} X(\omega) P(d\omega|A) = \frac{1}{P(A)} \int_A X(\omega) P(d\omega). \quad (1.1.2)$$

如果 $P(A) = 0$, 则规定 $E(X|A) = 0$. 特别当 $X(\omega) = I_B(\omega)$ 时, (1.1.2) 就化为 (1.1.1), 故我们只须考虑条件期望.

将 $E(X|A)$ 推广如下: 考虑 Ω 的某一可测分割 $A_1, A_2, \dots$, 即对一切 $i \geq 1, A_i \in \mathcal{F}, A_i A_j = \emptyset, i \neq j, \sum_{i=1}^{\infty} A_i = \Omega$. 为了从整体上考查 $E(X|A_i), i = 1, 2, \dots$, 定义随机变量 $Y(\omega)$ 如下

$$Y(\omega) = \sum_{i=1}^{\infty} E(X|A_i) I_{A_i}(\omega), \quad (1.1.3)$$

于是当 $E(|X|) < \infty, P(A_i) > 0$ 时, 由 (1.1.2), (1.1.3) 知

$$\int_{A_i} X(\omega) P(d\omega) = P(A_i) E(X | A_i) = \int_{A_i} Y(\omega) P(d\omega), \quad (1.1.4)$$

如果 $P(A_i)=0$, 上式两边都等于零, 故 (1.1.4) 仍然成立. 易见, 当取 $\{A, A^c\}$, $P(A)>0$, 为 Ω 的可测分割时, 在 (1.1.4) 中以 A 代替 A_i 即得 (1.1.2). 令 $\mathcal{S}=\sigma(A_i; i=1, 2, \dots)$ 是含一切 $\{A_i; i=1, 2, \dots\}$ 的最小 σ 域. 由于 $\mathcal{S}$ 中元素 (除空集外) 都是某些 A_i 的并集, 故由 (1.1.4) 知, 对一切 $B \in \mathcal{S}$, 有

$$\int_B X(\omega) P(d\omega) = \int_B Y(\omega) P(d\omega), \quad (1.1.5)$$

记 $Y = E(X | \mathcal{S})$, 则 $E(X | \mathcal{S})$ 具有性质:

$$(1) E(X | \mathcal{S}) \text{ 是 } \mathcal{S} \text{ 可测}. \quad (1.1.6)$$

$$(2) \forall B \in \mathcal{S}, \int_B X(\omega) P(d\omega) = \int_B E(X | \mathcal{S})(\omega) P(d\omega). \quad (1.1.7)$$

根据上面对特殊的 σ 域 $\mathcal{S}$ 讨论的启示, 我们将条件期望推广到一般的 σ 域.

定义 1.1.3 给定随机变量 X 及 $\mathcal{F}$ 的子 σ 域 $\mathcal{S}$, $E(|X|) < \infty$, 称随机变量 $E(X | \mathcal{S})$ 为 X 关于 $\mathcal{S}$ 的条件数学期望, 如果它满足 (1.1.6) 和 (1.1.7).

为使上述定义有意义, 必须保证满足 (1) 和 (2) 的随机变量存在. 为此考查集函数

$$\varphi(B) = \int_B X dP, \quad \forall B \in \mathcal{S},$$

易知, $\varphi(\cdot)$ 是 $\mathcal{S}$ 上广义测度, 而且如果 $P(B)=0$, 则有 $\varphi(B)=0$. 即 φ 在 $\mathcal{S}$ 上关于测度 P 绝对连续. 根据 Radon-Nikodym 定理, 在关于测度 P 几乎处处相等的意义下, 满足 (1) 和 (2) 的随机变量 $E(X | \mathcal{S})$ 惟一存在. 我们说 X 关于 $\mathcal{S}$ 的条件数学期望时, 指的是上述等价类中的一个代表.

例 1.1.4 若随机变量 X 关于 $\mathcal{S}$ 独立, 则有

$$E(X | \mathcal{S}) = EX. \quad (1.1.8)$$

证明 因为 EX 关于 $\mathcal{S}$ 可测, 且 $\forall B \in \mathcal{S}$,

$$\int_B X dP = E(X I_B) = E(X) P(B) = \int_B EX dP.$$

例 1.1.5 若 $\mathcal{S}=\mathcal{F}$ 或 X 为 $\mathcal{S}$ 可测, 则有

$$E(X | \mathcal{S}) = X, \quad a.s. \quad (1.1.9)$$

特别 $\forall B \in \mathcal{S}$ 有

$$P(B | \mathcal{S}) = I_B, \quad a.s. \quad (1.1.10)$$

例 1.1.6 假设 X 与 Y 同时为连续型随机变量, $E(|Y|) < \infty$, 而且 (X, Y) 有联合概率密度 $f(x, y)$, $x, y \in R$, 如果 $\mathcal{S}=\sigma(X)$, 即 $\mathcal{S}$ 是包含一切 $\{(X \in B), B \in \mathcal{B}(R)\}$ 的最小 σ 域, 则 $E(Y | \mathcal{S})$ 与概率论中定义的 Y 关于 X 的条件期望一致. 事实上, Y 在 $X=x$ 条件下的条件概率密度定义为

$$f(y|x) = f(x, y)/p(x), \quad p(x) > 0, \quad (1.1.11)$$

其中 $p(x) = \int_{-\infty}^{\infty} f(x, y) dy$, 当 $p(x) = 0$ 时, 规定 (1.1.11) 左方值为零. 于是 Y 在 $X=x$ 条件下的条件期望为

$$E(Y|X=x) = \int_{-\infty}^{\infty} yf(y|x) dy = g(x), \quad (1.1.12)$$

由 (1.1.11) 和 (1.1.12) 知右方为 x 的可测函数, 因此得到可测函数

$$g(x) = E(Y|X=x), \quad x \in R,$$

可见, Y 关于 X 的条件期望 $E(Y|X) = g(X)$ 为 $\sigma(X)$ 可测; 往证 $\forall B \in \sigma(X)$, 有

$$\int_B g(X) dP = \int_B Y dP, \quad (1.1.13)$$

假设 $h(x)$ 为任意有界可测函数, $A = \{x: p(x) > 0\}$, 则有

$$\begin{aligned} E(h(X)Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x) y f(x, y) dx dy \\ &= \int_A \left[\int_{-\infty}^{\infty} y \frac{f(x, y)}{p(x)} dy \right] p(x) h(x) dx \\ &= \int_A g(x) p(x) h(x) dp \\ &= \int_{\Omega} g(X) h(X) dP. \end{aligned}$$

如果在上式中令 $h(x) = I_{\Gamma}(x)$, $\Gamma \in \mathcal{R}(R)$ 则有

$$\int_{(X \in \Gamma)} Y dp = \int_{(X \in \Gamma)} g(X) dp,$$

即 (1.1.13) 成立, 可见 $g(X) = E(Y|\mathcal{S})$, a.s..

1.1.4 条件期望的基本性质

以下提到的随机变量 X, X_n 均假设是可积的. 不等式, 等式及极限关系都是几乎处处成立, 不另再声明.

性质 1.1.7 对于任意实数 $c_i, i=1, 2$, 有

$$E\left(\sum_{i=1}^2 c_i X_i | \mathcal{S}\right) = \sum_{i=1}^2 c_i E(X_i | \mathcal{S}).$$

证明 由定义知 $\sum_{i=1}^2 c_i E(X_i | \mathcal{S})$ 为 $\mathcal{S}$ 可测, 而且 $\forall B \in \mathcal{S}$, 有

$$\begin{aligned} \int_B \left[\sum_{i=1}^2 c_i E(X_i | \mathcal{S}) \right] dp &= \sum_{i=1}^2 c_i \int_B E(X_i | \mathcal{S}) dp \\ &= \sum_{i=1}^2 c_i \int_B X_i dp = \int_B \left(\sum_{i=1}^2 c_i X_i \right) dp. \end{aligned}$$

性质 1.1.8

(1) 如果 $X \geq 0$, 则有 $E(X|\mathcal{G}) \geq 0$.

(2) 如果 $X_1 \geq X_2$, 则有 $E(X_1|\mathcal{G}) \geq E(X_2|\mathcal{G})$.

证明 设 $X \geq 0$, 令

$$B = \{\omega; E(X|\mathcal{G})(\omega) < 0\},$$

$$B_m = \{\omega; E(X|\mathcal{G})(\omega) < -1/m\}, \quad m = 1, 2, \dots,$$

易见 $B = \bigcup_{m=1}^{\infty} B_m$, 而且

$$-\frac{1}{m}P(B_m) \geq \int_{B_m} E(X|\mathcal{G})dP = \int_{B_m} XdP \geq 0,$$

所以 $P(B_m) = 0, P(B) = 0$, (1) 成立; 由 (1) 及性质 1.1.7 推得 (2) 成立.

性质 1.1.9 $|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$.

证明 由性质 1.1.7 和 1.1.8 直接推得.

性质 1.1.10 (单调收敛定理) 假设 $0 \leq X_n \uparrow X, n \rightarrow \infty$, 则有

$$\lim_{n \rightarrow \infty} E(X_n|\mathcal{G}) = E(X|\mathcal{G}).$$

证明 由 $0 \leq X_n \leq X_{n+1} \leq X$ 得

$$0 \leq E(X_1|\mathcal{G}) \leq E(X_2|\mathcal{G}) \leq \dots \leq E(X|\mathcal{G}),$$

故对几乎所有 ω , $\lim_{n \rightarrow \infty} E(X_n|\mathcal{G})(\omega)$ 存在, 对极限不存在的 ω 定义极限值为零. 如此规定后 $\lim_{n \rightarrow \infty} E(X_n|\mathcal{G})(\omega)$ 为 $\mathcal{G}$ 可测. 往证它等于 $E(X|\mathcal{G})$. 任取 $B \in \mathcal{G}$, 根据积分的单调收敛定理得

$$\begin{aligned} \int_B \lim_{n \rightarrow \infty} E(X_n|\mathcal{G})dP &= \lim_{n \rightarrow \infty} \int_B E(X_n|\mathcal{G})dP \\ &= \lim_{n \rightarrow \infty} \int_B X_n dP = \int_B X dP, \end{aligned}$$

所以

$$\lim_{n \rightarrow \infty} E(X_n|\mathcal{G}) = E(X|\mathcal{G}).$$

性质 1.1.11 (Fatou 引理) 设随机变量列 $\{X_n; n=1, \dots\}$ 的每个 X_n 的期望存在.

(1) 若存在随机变量 X , 使 $EX > -\infty$, 且 $\forall n \geq 1, X_n \geq X, a.s.$, 则 $\lim_{n \rightarrow \infty} X_n$ 的期望存在, 且有

$$E[\lim_{n \rightarrow \infty} X_n|\mathcal{G}] \leq \lim_{n \rightarrow \infty} E[X_n|\mathcal{G}].$$

(2) 若存在随机变量 X , 使 $EX < \infty$, 且 $\forall n \geq 1, X_n \leq X, a.s.$, 则 $\overline{\lim_{n \rightarrow \infty} X_n}$ 的期望存在, 且有

$$E[\overline{\lim}_{n \rightarrow \infty} X_n | \mathcal{G}] \geq \overline{\lim}_{n \rightarrow \infty} E[X_n | \mathcal{G}].$$

证明 因为

$$0 \leq \inf_{k \geq n} X_k - X \uparrow \lim_{n \rightarrow \infty} X_n - X, \quad n \rightarrow \infty,$$

由单调收敛定理得

$$\begin{aligned} E[\lim_{n \rightarrow \infty} X_n | \mathcal{G}] - E[X | \mathcal{G}] &= E[\lim_{n \rightarrow \infty} X_n - X | \mathcal{G}] \\ &= \lim_{n \rightarrow \infty} E[(\inf_{k \geq n} X_k - X) | \mathcal{G}] \\ &= \lim_{n \rightarrow \infty} E[X_n - X | \mathcal{G}] \\ &= \lim_{n \rightarrow \infty} E[X_n | \mathcal{G}] - E[X | \mathcal{G}]. \end{aligned}$$

从而(1)得证,类似地可以证明(2).

性质 1.1.12 (控制收敛定理) 假设 $|X_n| \leq Y$, Y 可积, 而且 $\lim_{n \rightarrow \infty} X_n = X$, 则有

$$\lim_{n \rightarrow \infty} E(X_n | \mathcal{G}) = E(X | \mathcal{G}).$$

证明 令

$$S_n = \sup_{k \geq n} X_k, \quad I_n = \inf_{k \geq n} X_k,$$

因为

$$\begin{aligned} |X_n| &\leq Y, \quad S_n \downarrow X, \quad I_n \uparrow X, \quad n \rightarrow \infty, \\ 0 &\leq Y - S_n \downarrow Y - X, \quad 0 \leq Y + I_n \uparrow Y + X, \quad n \rightarrow \infty, \end{aligned}$$

由单调收敛定理得

$$\begin{aligned} E(Y - S_n | \mathcal{G}) &\downarrow E(Y - X | \mathcal{G}), \quad n \rightarrow \infty, \\ E(Y + I_n | \mathcal{G}) &\uparrow E(Y + X | \mathcal{G}), \quad n \rightarrow \infty, \end{aligned}$$

从而

$$\begin{aligned} E(X | \mathcal{G}) &= \lim_{n \rightarrow \infty} E(I_n | \mathcal{G}) \\ &\leq \lim_{n \rightarrow \infty} E(X_n | \mathcal{G}) \\ &\leq \lim_{n \rightarrow \infty} E(S_n | \mathcal{G}) \\ &= E(X | \mathcal{G}). \end{aligned}$$

性质 1.1.13 如果随机变量 X 是 $\mathcal{G}$ 可测的, 而且 $E(|XY|) < \infty$, 则有

$$E(XY | \mathcal{G}) = XE(Y | \mathcal{G}). \quad (1.1.14)$$

证明 令

$$\begin{aligned} L &= \{X : E(|XY|) < \infty\}, \\ \mathbf{L} &= \{X : X \text{ 使 (1.1.14) 满足}\}, \end{aligned}$$

由性质 1.1.7 和 1.1.10 容易证明 L 是 $\mathbf{L}$ 系, 另外当 $X = I_A$, $A \in \mathcal{G}$ 时, X 是 $\mathcal{G}$

可测的,而且对于任意 $B \in \mathcal{S}$, 有

$$\int_B XY dP = \int_{AB} Y dP = \int_{AB} E(Y|\mathcal{S}) dP = \int_B X E(Y|\mathcal{S}) dP,$$

故 $I_A \in L$, 根据 L 系方法 (参见文献 [2] 463 页引理 4) 知 L 包含一切属于 L 且关于 $\mathcal{S}$ 可测的随机变量 X .

性质 1.1.14

(1) 全数学期望公式: $E(E(X|\mathcal{S})) = EX$.

(2) 重条件期望公式: 当 $\mathcal{H} \subset \mathcal{S}$ 时, 有

$$E(E(X|\mathcal{H})|\mathcal{S}) = E(X|\mathcal{S}) = E(E(X|\mathcal{S})|\mathcal{H}).$$

证明 (1) 在 (1.1.7) 中令 $B = \Omega$, 即得到 (1).

(2) 因为 $\mathcal{H} \subset \mathcal{S}$, 故 $E(X|\mathcal{H})$ 为 $\mathcal{S}$ 可测, 从而由例 1.1.5 的结论得第 1 个等式, 为证明第 2 个等式, 只需注意, 如果 $B \in \mathcal{H}$, 则 $B \in \mathcal{S}$, 故有

$$\int_B X dP = \int_B E(X|\mathcal{S}) dP.$$

性质 1.1.15 (Holder 不等式) 设 $1 < p, q < \infty, \frac{1}{p} + \frac{1}{q} = 1$, 则有

$$E[|XY||\mathcal{S}] \leq (E[|X|^p|\mathcal{S}])^{1/p} (E[|Y|^q|\mathcal{S}])^{1/q},$$

特别当 $p = q = 2$ 时, 有 Cauchy-Schwarz 不等式

$$E[|XY||\mathcal{S}] \leq E(X^2|\mathcal{S}) E(Y^2|\mathcal{S}).$$

证明 当 $E[|X|^p|\mathcal{S}] \neq 0, E[|Y|^q|\mathcal{S}] \neq 0$ 时, 在不等式

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

中置

$$a = X / E[|X|^p|\mathcal{S}] \neq 0,$$

$$b = Y / E[|Y|^q|\mathcal{S}] \neq 0,$$

利用条件期望的线性性, 单调性和性质 1.1.11, 对所得不等式稍加整理, 即得所需结果.

为了便于证明下面 Jensen 不等式, 我们介绍关于凸函数的一个性质. 称函数 $f(x), x \in R$ 为凸函数, 如果对于任意 $x, y \in R, 0 \leq \alpha \leq 1$ 有

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad (1.1.15)$$

对任意固定的 $x < y$, 令 $z = \alpha x + (1 - \alpha)y$, 由 (1.1.15) 得

$$\begin{aligned} \frac{f(z) - f(x)}{z - x} &\leq \frac{\alpha f(x) + (1 - \alpha)f(y) - f(x)}{y - x} \\ &= \frac{f(y) - f(x)}{y - x}. \end{aligned}$$

令 $\alpha \uparrow 1$ (这等价于 $z \downarrow x$) 得 $f(x)$ 的右导数 $f'_+(x)$ 存在, 且满足下面不等式

$$\begin{aligned} f'_+(x) &\leq \frac{f(y)-f(x)}{y-x}, \\ f'_+(x)(y-x) &\leq f(y)-f(x), \end{aligned} \quad (1.1.16)$$

显然当 $y < x$ 时, 此不等式仍然成立.

性质 1.1.16 (Jensen 不等式) 设 $f(x), x \in R$ 为凸函数, 随机变量 X 是可积的, $\mathcal{G}$ 是 $\mathcal{A}$ 的任意子 σ 域则有

$$f[E(X|\mathcal{G})] \leq E[f(X)|\mathcal{G}], \quad a.s.$$

证明 在不等式 (1.1.16) 中, 令 $y = X, x = E(X|\mathcal{G})$ 得

$$f'_+(E(X|\mathcal{G}))(X - E(X|\mathcal{G})) \leq f(X) - f(E(X|\mathcal{G}))$$

两边对 $\mathcal{G}$ 求条件期望得

$$f[E(X|\mathcal{G})] \leq E[f(X)|\mathcal{G}], \quad a.s.$$

§ 1.2 离散参数鞅

1.2.1 停时

定义 1.2.1 定义在可测空间 $(\Omega, \mathcal{A})$ 上取值于 $\{0, 1, \dots, +\infty\}$ 的可测函数 $\tau(\omega)$, 称为关于 $\mathcal{F}$ 的子 σ 域流 $\{\mathcal{F}_n; n=0, 1, \dots\}$ 是停时, 如果对任意非负整数 $n, (\tau \leq n) \in \mathcal{F}_n$, 这等价于 $(\tau = n) \in \mathcal{F}_n$.

若 τ_1, τ_2 为停时, 根据定义不难验证 $\tau_1 + \tau_2, \tau_1 \vee \tau_2 \equiv \max(\tau_1, \tau_2), \tau_1 \wedge \tau_2 \equiv \min(\tau_1, \tau_2)$ 都是停时. 进一步若假设 $\{\tau_n\}$ 为一列停时, 容易证明 $\sup_{n \geq 1} \{\tau_n\}, \inf_{n \geq 1} \{\tau_n\}, \limsup_{n \rightarrow \infty} \{\tau_n\}, \liminf_{n \rightarrow \infty} \{\tau_n\}$ 都是停时.

定义 1.2.2 设 τ 是 $\{\mathcal{F}_n; n=0, 1, \dots\}$ 停时, 令

$$\mathcal{F}_\tau = \{A : A \in \mathcal{F}_\infty, A(\tau \leq n) \in \mathcal{F}_n, \forall n \geq 0\}, \quad (1.2.1)$$

称 $\mathcal{F}_\tau$ 为 τ 前 σ 域.

其中 $\mathcal{F}_\infty = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n)$; 易知当 $\tau \equiv n$ 时, $\mathcal{F}_\tau = \mathcal{F}_n$.

性质 1.2.3 设 τ, τ_1, τ_2 是 $\{\mathcal{F}_n; n=0, 1, \dots\}$ 停时, 则有

- (1) τ 为 $\mathcal{F}_\tau$ 可测.
- (2) 当 $\tau_1 \leq \tau_2$ 时, $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

证明 (1) 对任意非负整数 m, n , 令 $A = (\tau \leq m)$, 则

$$A(\tau \leq n) = (\tau \leq n \wedge m) \in \mathcal{F}_{n \wedge m} \subset \mathcal{F}_n \Rightarrow A \in \mathcal{F}_\tau.$$

故 τ 为 $\mathcal{F}_\tau$ 可测.

(2) $\forall A \in \mathcal{F}_{\tau_1}$, 由定义知,

$$\begin{aligned} \forall n \geq 0, \quad A(\tau_1 \leq n) &\in \mathcal{F}_n, \\ A(\tau_2 \leq n) &= A(\tau_1 \leq n)(\tau_2 \leq n) \in \mathcal{F}_n, \end{aligned}$$

故 $A \in \mathcal{F}_{\tau_2}, \mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

性质 1.2.4 设 $\tau_1 \geq \tau_2 \geq \dots \geq \tau_n \geq \dots$ 是一列 $\{\mathcal{F}_n: n=0, 1, \dots\}$ 停时, 则 $\tau = \inf_{n \geq 1} \{\tau_n\}$ 是停时, 而且 $\mathcal{F}_\tau = \bigcap_{n \geq 1} \mathcal{F}_{\tau_n}$.

证明 由性质 1.2.3 知, 对任意正整数 $n, \mathcal{F}_\tau \subset \mathcal{F}_{\tau_n} \Rightarrow \mathcal{F}_\tau \subset \bigcap_{n \geq 1} \mathcal{F}_{\tau_n}$, 反之任取 $A \in \bigcap_{n \geq 1} \mathcal{F}_{\tau_n}$, 有

$$A(\tau \leq n) = \bigcup_{k \geq 1} [A(\tau_k \leq n)] \in \mathcal{F}_n \Rightarrow A \in \mathcal{F}_\tau.$$

1.2.2 鞅

定义 1.2.5 称过程 $X = \{X_n: n=0, 1, \dots\}$ 为 $\{\mathcal{F}_n: n=0, 1, \dots\}$ 鞅 (相应地上鞅, 下鞅), 如果

(1) X 是 $\{\mathcal{F}_n: n=0, 1, \dots\}$ 适应的, 即对每个 n, X_n 是 $\mathcal{F}_n$ 可测.

(2) $E(|X_n|) < \infty, n=0, 1, \dots$.

(3) 对任意非负整数 $m < n, E(X_n | \mathcal{F}_m) = X_m$ (相应地 $\leq X_m, \geq X_m$).

例 1.2.6 假设 $\{Y_n: n=0, 1, \dots\}$ 是独立的随机变量列, $E(|Y_n|^2) < \infty, E(Y_n) = 0, g_k$ 是 k 元 Borel 可测函数, $k=1, 2, \dots$ 令

$$b_k = g_k(Y_0, Y_1, \dots, Y_{k-1}),$$

$$X_n = X_0 + \sum_{k=1}^n b_k Y_k,$$

$\mathcal{F}_n = \sigma(Y_0, Y_1, \dots, Y_n), n=1, 2, \dots$. 其中 X_0 为常量, 假定 $E(|b_k|^2) < \infty, k=1, 2, \dots$, 则 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅.

事实上, 显然对任意 $n=1, 2, \dots, E(|X_n|) < \infty$; 又因为

$$\begin{aligned} X_{n+1} &= X_n + b_{n+1} Y_{n+1}, \\ E(X_{n+1} | \mathcal{F}_n) &= E(X_n | \mathcal{F}_n) + E(b_{n+1} Y_{n+1} | \mathcal{F}_n) \\ &= X_n + b_{n+1} E(Y_{n+1} | \mathcal{F}_n) \\ &= X_n + b_{n+1} EY_{n+1} = X_n. \end{aligned}$$

定理 1.2.7 假设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 和 $\{Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅 (或下鞅), 则

(1) $\{X_n + Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅 (或下鞅).

(2) $\{X_n \vee Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅.

证明 (1) $X_n + Y_n$ 的适应性和可积性是显然的, 只需验证定义 1.2.5 中的条件 (3), 对任意非负整数 $m < n$,

$$\begin{aligned} E(X_n + Y_n | \mathcal{F}_m) &= E(X_n | \mathcal{F}_m) + E(Y_n | \mathcal{F}_m) \\ &= X_m + Y_m \quad (\text{相应地 } \geq X_m + Y_m). \end{aligned}$$

(2) 因为 $X_n \vee Y_n \geq X_n, X_n \vee Y_n \geq Y_n$, 故有

$$E(X_n \vee Y_n | \mathcal{F}_m) \geq E(X_n | \mathcal{F}_m) \vee E(Y_n | \mathcal{F}_m) \geq X_m \vee Y_m.$$

由定理 1.2.7(2) 知, 如果 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 和 $\{Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是上鞅, 则 $\{X_n \wedge Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 也是上鞅.

定理 1.2.8 (1) 假设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅, f 是定义在 R_1 上的凸函数. 如果, $\forall n \geq 0; E|f(X_n)| < \infty$, 则 $\{f(X_n), \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅.

(2) 假设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅(上鞅), f 是定义在 R_1 上的非降凸函数. 如果, $\forall n \geq 0, E|f(X_n)| < \infty$, 则 $\{f(X_n), \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅(上鞅).

(3) 假设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是上鞅, f 是定义在 R_1 上的非降凹函数. 如果, $\forall n \geq 0, E|f(X_n)| < \infty$, 则 $\{f(X_n), \mathcal{F}_n, n=0, 1, \dots\}$ 是上鞅.

证明 我们只证明(2), 其余的证明是类似的. 对任意非负整数 $m < n$, 因为函数 f 是非降的, 而且 $E(X_n | \mathcal{F}_m) \geq X_m$, 所以 $f(E(X_n | \mathcal{F}_m)) \geq f(X_m)$, 由 Jensen 不等式有

$$E(f(X_n) | \mathcal{F}_m) \geq f(E(X_n | \mathcal{F}_m)) \geq f(X_m).$$

推论 1.2.9 假设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅(或非负下鞅), $\lambda \geq 1$ 为一常数. 如果对 $n \geq 0, |X_n|^\lambda$ 是可积的, 则 $\{|X_n|^\lambda, \mathcal{F}_n, n \geq 0\}$ 是下鞅.

推论 1.2.10 假设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅, 则 $\{X_n^+, \mathcal{F}_n, n=0, 1, \dots\}$ 也是下鞅.

1.2.3 离散鞅的基本不等式

设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅(或下鞅), 则对于 $n \geq m$ 有 $E(X_n | \mathcal{F}_m) = X_m$ (或 $\geq X_m$), 如果将 n, m 换成停时后, 上面式子是否仍然成立? 下面的定理指出, 当停时有界时, 回答是肯定的.

定理 1.2.11 (Doob 有界停时定理) 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅, $\sigma \leq \tau$ 是一对有界停时, 则有

$$(1) E|X_\sigma| < \infty, E|X_\tau| < \infty.$$

$$(2) E(X_\tau | \mathcal{F}_\sigma) \geq X_\sigma. \quad (1.2.2)$$

如果 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅, 则(1.2.2)中不等号改为等号.

证明 设 $\tau \leq N$, 则有 $|X_\tau| \leq \sum_{i=0}^N |X_i| \in L^1$, 同理 $|X_\sigma| \in L^1$. 设 $A \in \mathcal{F}_\sigma$, 则有

$$A(\sigma = j)(\tau > j) \in \mathcal{F}_j. \quad (1.2.3)$$

如果 $\tau - \sigma \leq 1$, 那么由下鞅性及(1.2.3), 对 $A \in \mathcal{F}_\sigma$ 有

$$\int_A (X_\sigma - X_\tau) dP = \sum_{j=0}^N \int_{A \cap (\sigma=j)(\tau>j)} (X_j - X_{j+1}) dP \leq 0, \quad (1.2.4)$$

对一般情况, 令 $\alpha_j = \tau \wedge (\sigma + j)$, $j=1, 2, \dots, N$, 则 α_j 是停时, 且

$$\sigma \leq \alpha_1 \leq \dots \leq \alpha_N = \tau, \quad \alpha_1 - \sigma \leq 1, \quad \alpha_j - \alpha_{j+1} \leq 1, \\ j = 1, 2, \dots, N. \quad (1.2.5)$$

于是对任意 $A \in \mathcal{F}_\sigma$, 有 $A \in \mathcal{F}_{\alpha_j}$, $1 \leq j \leq N$, 从而

$$\int_A X_\sigma dP \leq \int_A X_{\alpha_1} dP \leq \dots \leq \int_A X_\tau dP,$$

又因为 X_σ 是 $\mathcal{F}_\sigma$ 可测, 故 (1.2.2) 成立.

推论 1.2.12 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅 (鞅), σ, τ 是一对有界停时, 则有

$$E(X_\tau | \mathcal{F}_\sigma) \geq X_{\sigma \wedge \tau} \quad (= X_{\sigma \wedge \tau}).$$

证明 因为

$$X_\sigma = X_\sigma I_{\sigma \geq \tau} + X_\sigma I_{\sigma < \tau} \\ = X_\sigma I_{\sigma \geq \tau} + X_{\sigma \wedge \tau} I_{\sigma > \tau}, \\ (\sigma < \tau) \in \mathcal{F}_{\sigma \wedge \tau} \subset \mathcal{F}_\tau,$$

故有

$$E(X_{\sigma \wedge \tau} I_{\sigma < \tau} | \mathcal{F}_\tau) = X_{\sigma \wedge \tau} I_{\sigma < \tau}, \\ E(X_\sigma I_{\sigma \geq \tau} | \mathcal{F}_\tau) = I_{\sigma \geq \tau} E(X_\sigma | \mathcal{F}_{\sigma \wedge \tau}) \geq I_{\sigma \geq \tau} X_{\sigma \wedge \tau}, \\ E(X_\sigma | \mathcal{F}_\tau) \geq I_{\sigma \geq \tau} X_{\tau \wedge \tau} + I_{\sigma < \tau} X_{\sigma \wedge \tau} = X_{\sigma \wedge \tau}.$$

定理 1.2.13 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅, 则 $\forall \lambda > 0, n \geq 0$, 有

$$(1) \lambda P\{\max_{k \leq n} X_k \geq \lambda\} \leq \int_{\{\max_{k \leq n} X_k \geq \lambda\}} X_n dP \leq E|X_n|.$$

$$(2) \lambda P\{\min_{k \leq n} X_k \leq -\lambda\} \leq -EX_0 + \int_{\{\min_{k \leq n} X_k > -\lambda\}} X_n dP \leq EX_n^+ - EX_0 \leq$$

$$E|X_0| + E|X_n|.$$

$$(3) \lambda P\{\max_{k \leq n} |X_k| \geq \lambda\} \leq 2EX_n^+ - EX_0 \leq E|X_0| + 2E|X_n|.$$

证明 (1) 令

$$\tau = \begin{cases} \min\{k: k \leq n, x_k \geq \lambda\}, \\ n, \quad \text{如果上面集合是空集}, \end{cases}$$

则 τ 是停时, 而且 $\tau \leq n$, 记 $M = \{\max_{k \leq n} X_k \geq \lambda\}$, 则在 M 上, $x_\tau \geq \lambda$, 而且 $M \in \mathcal{F}_\tau$, 实际上 $\forall j \leq n, M \cap (\tau = j) = (X_k < \lambda, 0 \leq k \leq j-1, X_j \geq \lambda) \in \mathcal{F}_j$. 因此对 τ, n 应用 Doob 有界停时定理得

$$\lambda P(M) \leq \int_M X_\tau dP \leq \int_M X_n dP,$$

(1)的第一个不等式得证;(1)的第二,三个不等式是显然的.

(2)类似地令

$$\tau = \begin{cases} \min\{k: k \leq n, X_k \leq -\lambda\}, \\ n, \quad \text{如果上面集合是空集}, \end{cases}$$

记 $M = \{\min_{k \leq n} X_k \leq -\lambda\}$, 因为 $\tau \leq n$, 对 $0, \tau$ 应用 Doob 有界停时定理得

$$EX_0 \leq EX_\tau = \int_M X_\tau dP + \int_{M^c} X_\tau dP \leq -\lambda P(M) + \int_{M^c} X_\tau dP.$$

即(2)的第一个不等式成立,(2)的第二,三个不等式是显然的.

推论 1.2.14 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅, 且 $E|X_n|^p < \infty, p \geq 1, n \geq 0$, 则 $\forall n \geq 1$,

$$P\{\max_{k \leq n} |X_k| \geq \lambda\} \leq E|X_n|^p / \lambda^p.$$

证明 由 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是鞅知 $\{|X_n|^p\}$ 是下鞅, 由定理 1.2.13 (1)得

$$\begin{aligned} \lambda^p P\{\max_{k \leq n} |X_k|^p \geq \lambda^p\} &\leq E|X_n|^p, \\ P\{\max_{k \leq n} |X_k| \geq \lambda\} &= P\{\max_{k \leq n} |X_k|^p \geq \lambda^p\} \leq E|X_n|^p / \lambda^p. \end{aligned}$$

定理 1.2.15 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是非负下鞅, $p > 1$, 则

$$E\{\max_{k \leq n} X_k^p\} \leq \left(\frac{p}{p-1}\right)^p E|X_n|^p. \quad (1.2.6)$$

证明 令 $Y = \max_{k \leq n} X_k$, 由定理 1.2.13(1)知

$$\lambda p(Y \geq \lambda) \leq \int_\Omega I_{Y \geq \lambda} X_n dP, \quad (1.2.7)$$

由(1.2.7)及 Fubini 定理得

$$\begin{aligned} EY^p &= \int_\Omega dP \int_0^Y p\lambda^{p-1} d\lambda = \int_\Omega dP \int_0^\infty I_{(Y \geq \lambda)} p\lambda^{p-1} d\lambda \\ &= p \int_0^\infty \lambda^{p-1} P(Y \geq \lambda) d\lambda \leq p \int_0^\infty \lambda^{p-1} \left(\frac{1}{\lambda} \int_{(Y \geq \lambda)} X_n dP\right) d\lambda \\ &= \int_\Omega X_n \left(\int_0^Y p\lambda^{p-1} d\lambda\right) dP = \frac{p}{p-1} E(X_n Y^{p-1}), \end{aligned} \quad (1.2.8)$$

再由(1.2.8)及 Holder 不等式得

$$EY^p \leq q \|X_n\|_p [EY^{(p-1)q}]^{1/p} = q \|X_n\|_p \|Y\|_p^{p/q},$$

如果 $EY^p \neq 0$, 两边除以 $\|Y\|_p^{p/q}$ 得 $\|Y\|_p \leq q \|X_n\|_p$, (1.2.6)得证; 如果 $EY^p = 0$, 则 $EX_n^p = 0$, (1.2.6)仍成立.

推论 1.2.16 (Doob 不等式) 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是非负下鞅, $p > 1$, 则

$$[E\{\sup_{k \leq n} X_k\}^p]^{1/p} \leq q \sup_{n \geq 0} \|X_n\|_p. \quad (1.2.9)$$

证明 由 (1.2.6) 得

$$[E\{\max_{k \leq n} X_k^p\}]^{1/p} \leq q \|X_n\|_p \leq q \sup_{n \geq 0} \|X_n\|_p,$$

令 $n \rightarrow \infty$ 得 (1.2.9).

下面证明关于半鞅的上穿不等式, 先介绍上穿概念. 给定闭区间 $[a, b]$ 及 M 个数 $x_1, x_2, \dots, x_M$, 如果从 x_1 起顺次到 x_M , 自 $[a, b]$ 的左方到其右方共 v 次, 就说数列 $x_1, x_2, \dots, x_M$ 上穿 $[a, b]$ v 次, 确切的定义如下, 令

$$\begin{aligned} \tau_1 &= \begin{cases} \min\{n: 0 \leq n \leq M, x_n \leq a\}, \\ M+1, & \text{如果上面集合是空集,} \end{cases} \\ \sigma_1 &= \begin{cases} \min\{n: \tau_1 \leq n \leq M, x_n \geq b\}, \\ M+1, & \text{如果上面集合是空集,} \end{cases} \\ \tau_k &= \begin{cases} \min\{n: \sigma_{k-1} \leq n \leq M, x_n \leq a\}, \\ M+1, & \text{如果上面集合是空集,} \end{cases} \\ \sigma_k &= \begin{cases} \min\{n: \tau_k \leq n \leq M, x_n \geq b\}, \\ M+1, & \text{如果上面集合是空集,} \end{cases} \quad k=1, 2, \dots. \end{aligned}$$

定义 1.2.17 使 $\sigma_k \leq M$ 的最大的 k 称为数列 $x_1, x_2, \dots, x_M$ 上穿 $[a, b]$ 的次数, 记为 ν_a^b (如果 $\sigma_1 = M+1$, 则 $\nu_a^b = 0$).

定理 1.2.18 (上穿不等式) 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是下鞅, ν_a^b 是它上穿 $[a, b]$ 的次数, 则有

$$E(\nu_a^b) \leq \frac{1}{b-a} [E(X_M - a)^+ - E(X_0 - a)^+] \leq \frac{1}{b-a} [EX_M^+ + |a|]. \quad (1.2.10)$$

证明 令 $Y_n = (X_n - a)^+$, $n=0, 1, \dots, M$, 由推论 1.2.10 知 $\{Y_n\}$ 是下鞅, 显然 $\{Y_n\}$ 非负. 而且 $\{Y_n\}$ 上穿 $[0, b-a]$ 的次数等于 ν_a^b . 在 τ_k, σ_k 的定义中将 a, b, X_n 分别换成 $0, b-a, Y_n$, 并补充定义 $\sigma_0 = 0, \tau_{M+1} = M+1, Y_{M+1} = Y_M$, 则有

$$Y_M - Y_0 = \sum_{k=1}^M (Y_{\sigma_k} - Y_{\tau_k}) + \sum_{k=1}^{M+1} (Y_{\tau_k} - Y_{\sigma_{k-1}}), \quad (1.2.11)$$

如果 $\nu_a^b = r > 0$, 则有 $Y_{\sigma_k} - Y_{\tau_k} \geq b-a, k=1, \dots, r$, 由于当 $k > r$ 时, $Y_{\sigma_k} - Y_{\tau_k} \geq 0$, 故

$$\sum_{k=1}^M (Y_{\sigma_k} - Y_{\tau_k}) \geq (b-a)r = (b-a)\nu_a^b,$$

当 $r=0$ 时,上面式子仍然成立,于是有

$$E\left[\sum_{k=1}^M(Y_{\tau_k}-Y_{\sigma_{k-1}})\right]\geqslant(b-a)E\nu_a^b, \quad (1.2.12)$$

其次由于 τ_k, σ_{k-1} 都是有界停时,而且 $\tau_k \geqslant \sigma_{k-1}$,由定理 1.2.11 得 $EY_{\tau_k} \geqslant EY_{\sigma_{k-1}}$,从而

$$E\left[\sum_{k=1}^{M+1}(Y_{\tau_k}-Y_{\sigma_{k-1}})\right]=\sum_{k=1}^{M+1}(EY_{\tau_k}-EY_{\sigma_{k-1}})\geqslant 0, \quad (1.2.13)$$

由(1.2.11),(1.2.12),(1.2.13)得 $EY_M-EY_0 \geqslant (b-a)E\nu_a^b$,即

$$E\nu_a^b \leqslant \frac{1}{b-a}[E(X_M-a)^+-E(X_0-a)^+].$$

1.2.4 离散鞅的收敛定理

定理 1.2.19 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots, M\}$ 是下鞅,满足

$$\sup_{n \leqslant M} E|X_n| < \infty, \quad (1.2.14)$$

则存在 $\mathcal{F}_\infty$ 可测的随机变量 X_∞ , 使 $E|X_\infty| < \infty$, 而且

$$p(\lim_{n \rightarrow \infty} X_n = X_\infty) = 1, \quad (1.2.15)$$

其中 $\mathcal{F}_\infty = \sigma(\bigcup_{n=0}^\infty \mathcal{F}_n)$.

证明 令

$$A = \{\omega; \lim_{n \rightarrow \infty} \sup X_n(\omega) > \lim_{n \rightarrow \infty} \inf X_n(\omega)\},$$

$$A(a, b) = \{\omega; \lim_{n \rightarrow \infty} \inf X_n(\omega) < a < b < \lim_{n \rightarrow \infty} \sup X_n(\omega)\},$$

则有 $A, A(a, b) \in \mathcal{F}_\infty$, 而且

$$A = \bigcup_{a < b, a, b \in Q} A(a, b), \quad (1.2.16)$$

其中 Q 为全体有理数.下面证明 $P(A)=0$, 令 $\nu_a^b(M)$ 表示数列 $X_0, X_1, \dots, X_M$ 上穿 $[a, b]$ 的次数, ν_a^b 表示数列 $X_0, X_1, \dots$ 上穿 $[a, b]$ 的次数.显然 $\nu_a^b(M)$ 关于 M 不减, 而且有 $\nu_a^b = \lim_{M \rightarrow \infty} \nu_a^b(M)$, 由上穿不等式得

$$\begin{aligned} E\nu_a^b(M) &\leqslant \frac{1}{b-a}[EX_M^+ + |a|] \\ &\leqslant \frac{1}{b-a}[\sup_{M \geqslant 0} E|X_M| + |a|], \end{aligned} \quad (1.2.17)$$

由(1.2.14)和(1.2.17)知 $E\nu_a^b < \infty$, 从而

$$p(\nu_a^b < \infty) = 1, \quad (1.2.18)$$

由于 $A(a, b) \subset \{\omega; \nu_a^b(\omega) = +\infty\}$, 再由(1.2.18)知 $p\{A(a, b)\} = 0$, 从而 $p(A) = 0$, 即 $\lim_{n \rightarrow \infty} X_n$ 几乎处处存在. 令 $X_\infty = \lim_{n \rightarrow \infty} \inf X_n$, 则 X_∞ 是 $\mathcal{F}_\infty$ 可测, 而

且 $X_\infty = \lim_{n \rightarrow \infty} X_n, a.s.$, 再用 Fatou 引理, 和 (1.2.15) 式, 得

$$E|X_\infty| \leq \liminf_{n \rightarrow \infty} E|X_n| \leq \sup_{n \geq 0} E|X_n| < \infty.$$

注 1.2.20 由于 $E|X_n| = 2EX_n^+ - EX_n \leq 2EX_n^+ - EX_1$, 可见对于下鞅条件 (1.2.14) 等价于形式上较弱的条件 $\sup_{n \geq 0} EX_n^+ < \infty$, 对于上鞅, 由于

$$E|X_n| = 2EX_n^- + EX_n \leq 2EX_n^- + EX_1,$$

故 (1.2.14) 等价于条件 $\sup_{n \geq 0} EX_n^- < \infty$.

1.2.5 Doob 停时定理

定义 1.2.21 称鞅(或下鞅) $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是右闭的, 如果存在 $\mathcal{F}_\infty$ 可测, 且可积的随机变量 X_∞ , 使得对任意 $n \geq 0$ 有

$$E(X_\infty | \mathcal{F}_n) = X_n \text{ (或 } \leq X_n \text{)}, \quad (1.2.19)$$

显然鞅(或下鞅) $\{X_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是右闭的充分必要条件是存在可积的 Y , 对任意 $n \geq 0$ 有

$$E(Y | \mathcal{F}_n) = X_n \text{ (或 } \geq X_n \text{)},$$

实际上, 只需令 $X_\infty = E(Y | \mathcal{F}_\infty)$, 那么 X_∞ 满足右闭定义的要求.

下面证明离散参数时的 Doob 停时定理.

定理 1.2.22 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots, M\}$ 是右闭鞅(或下鞅), 停时 $\sigma \leq \tau$, 则有

(1) X_σ, X_τ 都可积, 且 $E(X_\infty | \mathcal{F}_\tau) = X_\tau$ (或 $\geq X_\tau$).

(2) $E(X_\tau | \mathcal{F}_\sigma) = X_\sigma$ (或 $\geq X_\sigma$).

证明 对任意 Borel 集 Γ 及 $n \geq 0$ 有

$$(X_\tau \in \Gamma)(\tau \leq n) = \bigcup_{k=0}^n (X_k \in \Gamma)(\tau = k) \in \mathcal{F}_n,$$

故 $X_\tau \in \mathcal{F}_\tau$, 同理 $X_\sigma \in \mathcal{F}_\sigma$

(一) 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots, M\}$ 是右闭鞅, 由 (1) 及条件期望的性质, 有

$$\begin{aligned} \int_{\Omega} |X_\tau| dp &= \sum_{n=0}^{\infty} \int_{(\tau=n)} |X_n| dp \\ &\leq \sum_{n=0}^{\infty} \int_{(\tau=n)} |X_\infty| dp = E|X_\infty| < \infty, \end{aligned} \quad (1.2.20)$$

从而 X_τ 可积, 同理 X_σ 可积. 对任意 $A \in \mathcal{F}_\tau$, 由 (1.2.19) 得

$$\begin{aligned} \int_A |X_\tau| dp &= \sum_{n=0}^{\infty} \int_{A(\tau=n)} |X_n| dp \\ &\leq \sum_{n=0}^{\infty} \int_{A(\tau=n)} |X_\infty| dp = \int_A X_\infty dp, \end{aligned} \quad (1.2.21)$$

即 $E(X_\infty | \mathcal{F}_\tau) = X_\tau$, 同理 $E(X_\infty | \mathcal{F}_\sigma) = X_\sigma$, 故(1)成立. 再来证明(2), 由(1)及 $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$ 可得

$$E(X_\tau | \mathcal{F}_\sigma) = E(E(X_\infty | \mathcal{F}_\sigma)) = E(X_\infty | \mathcal{F}_\sigma) = X_\sigma.$$

(二) 设 $\{X_n, \mathcal{F}_n, n=0, 1, \dots, M\}$ 是右闭下鞅

先将 X_n 分解为

$$X_n = E(X_\infty | \mathcal{F}_n) - Y_n, \quad 0 \leq n \leq \infty, \quad (1.2.22)$$

其中

$$Y_n = E(X_\infty | \mathcal{F}_n) - X_n,$$

易知 $\{Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是非负上鞅而且

$$Y_\infty = E(X_\infty | \mathcal{F}_\infty) - X_\infty = 0,$$

因为 Y_n 非负, 所以 $E(0 | \mathcal{F}_n) = 0 \leq Y_n, n=0, 1, \dots$, 这表明 $\{Y_n\}$ 是右闭的. 记

$$\pi_n = E(X_\infty | \mathcal{F}_n), \quad n = 0, 1, \dots,$$

则 $\{\pi_n, \mathcal{F}_n, n=0, 1, \dots\}$ 是右闭鞅, 根据(一)的结果, $\{\pi_n\}$ 满足(1)和(2), 因此只需证明 $\{Y_n, \mathcal{F}_n, n=0, 1, \dots\}$ 满足相应的(1)和(2). 因为 $Y_\infty = 0, a.s.$, 故有

$$Y_\tau = Y_\tau I_{(\tau < \infty)}, \quad a.s.$$

$$EY_\tau = E \lim_{n \rightarrow \infty} Y_{\tau \wedge n} I_{(\tau < n)} \leq E \lim_{n \rightarrow \infty} Y_{\tau \wedge n}, \quad (1.2.23)$$

由 Fatou 引理得 $EY_\tau \leq \liminf_{n \rightarrow \infty} EY_{\tau \wedge n}$, 因为 0 和 $\tau \wedge n$ 是有界停时, 由定理 1.2.12 知 (1.2.23) 的右方不大于 EY_0 , 从而 $0 \leq EY_\tau < \infty$, 显然 $0 = E(Y_\infty | \mathcal{F}_\tau) \leq Y_\tau$, 故(1)成立. 现来证明(2), 令 $\tau_k = \tau \wedge k, \sigma_k = \sigma \wedge k, k=0, 1, \dots$. 则 τ_k, σ_k 都是有界停时, 且 $\tau_k \leq \sigma_k$, 任取 $A \in \mathcal{F}_\sigma$, 由于 $A(\sigma \leq k) \in \mathcal{F}_{\sigma_k}$, 由定理 1.2.21 得

$$\int_{A(\tau \leq k)} Y_{\tau_k} dp \leq \int_{A(\sigma \leq k)} Y_{\sigma_k} dp, \quad (1.2.24)$$

但是 $(\tau \leq k) \subset (\sigma \leq k)$, 又因为 $Y_n \geq 0, a.s.$ 故有

$$\int_{A(\tau \leq k)} Y_{\tau_k} dp \leq \int_{A(\sigma \leq k)} Y_{\sigma_k} dp, \quad (1.2.25)$$

注意在 $(\tau \leq k)$ 上 $\tau = \tau_k$, 在 $(\sigma \leq k)$ 上 $\sigma = \sigma_k$, 于是由 (1.2.25) 得

$$\int_{A(\tau \leq k)} Y_\tau dp \leq \int_{A(\sigma \leq k)} Y_\sigma dp,$$

令 $k \rightarrow \infty$ 得

$$\int_{A(\tau < \infty)} Y_\tau dp \leq \int_{A(\sigma < \infty)} Y_\sigma dp, \quad (1.2.26)$$

注意到 $Y_\infty = 0, a.s.$, 由 (1.2.26) 得

$$\int_A Y_\tau dp \leq \int_A Y_\sigma dp,$$

即 $Y_\sigma \geq E(Y_\tau | \mathcal{F}_\sigma)$, 于是由 (2.6.4) 得

$$E(X_\tau | \mathcal{F}_\sigma) = E(\pi_\tau | \mathcal{F}_\sigma) - E(Y_\tau | \mathcal{F}_\sigma) \geq \pi_\sigma - Y_\sigma = X_\sigma.$$

§ 1.3 马尔可夫链

本节介绍离散时间参数马尔可夫链 (Markov chain) 的基本理论.

1.3.1 马尔可夫链的定义及其转移概率

定义 1.3.1 称随机序列 $\{X_n: n=0, 1, \dots\}$ 为离散参数马尔可夫链, 如果它满足

(1) $\{X_n: n=0, 1, \dots\}$ 的状态空间 E 为至多可数集.

(2) 对任意非负整数 n 及状态 $i_0, i_1, \dots, i_{n+1}$ 有

$$p(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = p(X_{n+1} = i_{n+1} | X_n = i_n), \quad (1.3.1)$$

如果对任意非负整数 m, n 及任意状态 i 和 j 有

$$p(X_{n+1} = j | X_n = i) = p(X_{m+1} = j | X_m = i), \quad (1.3.2)$$

则称马氏链 $\{X_n: n=0, 1, \dots\}$ 是齐次的. 在齐次情形, 记 $p_{ij} = p(X_{n+1} = j | X_n = i)$, 称 p_{ij} 为 $\{X_n: n=0, 1, \dots\}$ 的转移概率, 以 $p_{ij}, i, j \in E$ 为元素形成的矩阵 $\mathbf{P} = (p_{ij})_{i, j \in E}$ 称为 $\{X_n: n=0, 1, \dots\}$ 的转移矩阵, 本节只介绍齐次马氏链.

例 1.3.2 简单随机游动, 假设 $X_0 = a, a$ 是常数, $Z_1, Z_2, \dots$ 是独立同分布的随机变量列, 而且 $P(Z_n = 1) = p, P(Z_n = 0) = r, P(Z_n = -1) = q, p + q$

$+ r = 1$, 令 $X_n = \sum_{k=0}^n Z_k$, 则 $\{X_n: n=0, 1, \dots\}$ 为齐次马氏链. 其状态空间 $E = \{a \pm n, n=0, 1, \dots\}$, 转移概率为 $p_{i, j+1} = p, p_{i, j-1} = q, p_{i, j} = r$, 其余的 $p_{ij} = 0$, 此马氏链称为简单随机游动. 当 $p + q = 1/2$ 时, 就称 $\{X_n: n=0, 1, \dots\}$ 为对称随机游动.

例 1.3.3 带吸收壁的随机游动. 假设简单随机游动 $\{X_n: n=0, 1, \dots\}$ 的状态空间为 $\{0, 1, \dots, b\}$, 如果质点移动到状态 0 或 b 后就永远停留在该状态, 即 $p_{0,0} = p_{b,b} = 1$, 其余的 $p_{ij} = 0$, 就称 $\{X_n: n=0, 1, \dots\}$ 为带吸收壁 0 和 b 的随机游动.

例 1.3.4 如果例 1.3.3 中的质点移动到状态 0 或 b 后, 下一次移动必返回, 即 $p_{0,1} = p_{b,b-1} = 1$, 就称 $\{X_n: n=0, 1, \dots\}$ 为带反射壁 0 和 b 的随机游动.

定理 1.3.5 齐次马氏链的联合分布由初始分布 $p(x_0 = i) = p_i, i \in E$ 及一步转移概率矩阵 $P = (p_{ij})$ 完全确定.

证明 对任意 $n=0, 1, \dots$ 和任意状态 $i_0, i_1, \dots, i_n \in E$, 由马氏链的定义有

$$\begin{aligned} & p(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) \\ &= p(X_0 = i_0) p(X_1 = i_1 | X_0 = i_0) p(X_2 = i_2 | X_0 = i_0, X_1 = i_1) \\ & \quad \cdots p(X_n = i_n | X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) \\ &= p_{i_0} p_{i_0, i_1} p_{i_1, i_2} \cdots p_{i_{n-1}, i_n}, \end{aligned} \quad (1.3.3)$$

从而 $\{X_n; n=0, 1, \dots\}$ 的任意有限维分布由初始分布及一步转移概率矩阵 P 完全确定, 再由 Kolmogorov 扩张定理知定理 1.3.6 的结论成立.

定理 1.3.6 对于齐次马氏链及任意 $m, n=0, 1, \dots$ 任意 $i, j \in E$ 有

$$P_{ij}^{(n+m)} = \sum_{k \in E} P_{ik}^{(n)} P_{kj}^{(m)},$$

其中

$$P_{ij}^{n+m} = p(X_{n+m} = j | X_0 = i).$$

证明

$$\begin{aligned} P_{ij}^{(n+m)} &= p(X_{n+m} = j | X_0 = i) \\ &= \sum_{k \in E} p(X_n = k, X_{n+m} = j | X_0 = i) \\ &= \sum_{k \in E} p(X_n = k | X_0 = i) p(X_{n+m} = j | X_0 = i, X_n = k) \\ &= \sum_{k \in E} P_{ik}^{(n)} P_{kj}^{(m)}, \end{aligned} \quad (1.3.4)$$

(1.3.4) 式称为 Chapman-Kolmogorov 方程.

1.3.2 状态的分类

定义 1.3.7 如果集合 $\{n; n \geq 1, P_{ij}^{(n)} > 0\}$ 不空, 该数集的最大公约数 $d = d(i)$ 称为状态 i 的周期; 如果 $d > 1$, 称 i 为周期的状态; 如果 $d = 1$, 称 i 为非周期的.

现在我们引入常返性概念, 记

$$f_{ij}^{(n)} = P(X_{m+n} \neq j, 1 \leq \nu \leq n-1, X_{n+m} = j | X_m = i), \quad n \geq 1, \quad (1.3.5)$$

$$f_{ij}^{(0)} = 0,$$

易知, 上式右方等于

$$\sum_{s_k \neq j, 1 \leq k \leq n-1} p_{i, s_1} p_{s_1, s_2} \cdots p_{s_{n-1}, j}, \quad (1.3.6)$$

与 m 无关,它表示质点由 i 出发,经 n 步首次到达 j 的概率.记 $f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$,它表示质点由 i 出发,经有限步终于到达 j 的概率.

定义 1.3.8 如果 $f_{ij}=1$,称状态 i 为常返的;如果 $f_{ij}<1$,称状态 i 为非常返的.

对常返状态 $i, \{f_{ij}^{(n)}, n=1,2,\dots\}$ 形成一概率分布,此分布的期望值 $\mu_i = \sum_{n=1}^{\infty} n f_{ij}^{(n)}$,表示质点由 i 出发再返回到 i 的平均回转时间,于是可将状态 i 细分如下.

定义 1.3.9 称常返状态 i 为正常返的,如果 $\mu_i < \infty$.称常返状态 i 为零常返的,如果 $\mu_i = \infty$.非周期的正常返状态称为遍历状态.

$f_{ij}^{(n)}$ 与 $p_{ij}^{(n)}$ 有如下关系.

定理 1.3.10 对任意状态 i, j 及 $n=1,2,\dots$,有

$$p_{ij}^{(n)} = \sum_{k=1}^n f_{ij}^{(k)} p_{jj}^{(n-k)} = \sum_{k=0}^{n-1} p_{ij}^{(k)} f_{jj}^{(n-k)}. \quad (1.3.7)$$

证明

$$\begin{aligned} p_{ij}^{(n)} &= p(X_n = j | X_0 = i) \\ &= \sum_{k=1}^n p(X_v \neq j, 1 \leq v \leq k-1, X_k = j, X_n = j | X_0 = i) \\ &= \sum_{k=1}^n p(X_v \neq j, 1 \leq v \leq k-1, X_k = j | X_0 = i), \\ p(X_n = j | X_0 = i, X_v \neq j, 1 \leq v \leq k-1, X_k = j) \\ &= \sum_{k=1}^n f_{ij}^{(k)} p_{jj}^{(n-k)}. \end{aligned}$$

类似地可以证明第二个等式.

下面研究常返性的判别及其性质.假设 $a_0, a_1, a_2, \dots$ 为一实数列,我们称幂级数 $A(s) = \sum_{n=0}^{\infty} a_n s^n$ 为数列 $\{a_n\}$ 的母函数.易知,如果 $\{a_n\}$ 有界,则 $A(s)$ 对一切 $|s| < 1$ 收敛.假设 $\{a_n\}$ 与 $\{b_n\}$ 的母函数分别为 $A(s)$ 与 $B(s)$ 且对一切 $|s| < 1$ 收敛,那么

$$c_n = \sum_{k=0}^n a_k b_{n-k}, \quad n = 0, 1, \dots, \quad (1.3.8)$$

的母函数

$$C(s) = A(s)B(s), \quad |s| < 1. \quad (1.3.9)$$

定理 1.3.11 状态 i 是常返的当且仅当 $\sum_{n=0}^{\infty} p_{ii}^{(n)} = \infty$; 如果状态 i 是非

常返的,则有

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

证明 记 $\{p_{ii}^{(n)}\}$ 和 $\{f_{ii}^{(n)}\}$ 的母函数分别为 $P(s)$ 和 $F(s)$, 并且约定 $p_{ii}^{(0)} = 1, f_{ii}^{(0)} = 0$, 由 (1.3.7) 得

$$P(s) - 1 = P(s)F(s), \quad (1.3.10)$$

当 $0 \leq s < 1$ 时, $F(s) < f_{ii} \leq 1$. 故由 (1.3.10) 得

$$P(s) = \frac{1}{1 - F(s)}, \quad 0 \leq s < 1, \quad (1.3.11)$$

又因为对任意 $0 \leq s < 1$ 及任意正整数 N , 有

$$\sum_{n=0}^N p_{ii}^{(n)} s^n \leq P(s) \leq \sum_{n=0}^{\infty} p_{ii}^{(n)}, \quad (1.3.12)$$

在 (1.3.12) 中令 $s \uparrow 1, N \rightarrow \infty$, 得

$$\lim_{s \uparrow 1} P(s) = \sum_{n=0}^{\infty} p_{ii}^{(n)}. \quad (1.3.13)$$

同理

$$\lim_{s \uparrow 1} F(s) = \sum_{n=0}^{\infty} f_{ii}^{(n)} = f_{ii}, \quad (1.3.14)$$

在 (1.3.11) 中令 $s \uparrow 1$, 由 (1.3.13), (1.3.14) 得到定理的证明.

$$\text{令 } g_{ij} = p_i(\text{有无限多个 } n \text{ 使 } X_n = j) = p_i\left\{\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} (X_n = j)\right\}. \quad (1.3.15)$$

定理 1.3.12 对任意状态 i , 有

$$g_{ij} = \begin{cases} f_{ij}, & \text{当 } f \text{ 常返,} \\ 0, & \text{否则.} \end{cases} \quad (1.3.16)$$

证明 令

$$A_k = \{\omega; \text{至少有 } k \text{ 个 } n \text{ 使 } X_n(\omega) = j\},$$

显然 $A_{k+1} \subset A_k$, 且有

$$\lim_{k \rightarrow \infty} p_i(A_k) = g_{ij}.$$

另一方面,

$$\begin{aligned} p_i(A_{k+1}) &= p_i\left\{\bigcup_{m=1}^{\infty} (X_v \neq j, 1 \leq v \leq m-1, X_m = j, \right. \\ &\quad \left. \text{且至少有 } k \text{ 个 } n \text{ 使 } X_{n+m} = j)\right\} \\ &= \sum_{m=1}^{\infty} p_i\{(X_v \neq j, 1 \leq v \leq m-1, X_m = j)p_j \\ &\quad (\text{至少有 } k \text{ 个 } n \text{ 使 } X_n = j)\} \end{aligned}$$

$$= \sum_{m=1}^{\infty} f_{ij}^{(m)} p_j(A_k) = f_{ij} p_j(A_k). \quad (1.3.17)$$

反复迭代(1.3.17),并注意 $p_j(A_1) = f_{jj}$, 我们得到

$$p_i(A_{k+1}) = f_{ij} f_{jj} p_j(A_{k-1}) = \cdots = f_{ij} (f_{jj})^n, \quad (1.3.18)$$

令 $k \rightarrow \infty$, 如果 $f_{jj} = 1$, 则得到 $g_{ij} = f_{ij}$, 如果 $f_{jj} < 1$, 则得到 $g_{ij} = 0$, 定理得证.

推论 1.3.13 状态 i 是常返的当且仅当 $g_{ii} = 1$, 如果状态 i 是非常返的, 则有 $g_{ii} = 0$.

证明 直接由定理 1.3.12 得到.

定理 1.3.14 假设状态 i 是常返的, 且有周期为 d , 则有

$$\lim_{n \rightarrow \infty} p_{ii}^{(nd)} = d / \mu_i. \quad (1.3.19)$$

证明 参见[2]64 页定理 2.

定理 1.3.15 假设状态 i 是常返的, 则有

(1) 状态 i 是零常返的当且仅当 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$.

(2) 状态 i 是遍历的当且仅当 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 1 / \mu_i > 0$.

证明 (1) 如果状态 i 是零常返, 由(1.3.19)知 $\lim_{n \rightarrow \infty} p_{ii}^{(nd)} = 0$, 但当 $n \neq 0 \pmod{d}$ 时 $p_{ii}^{(n)} = 0$, 故有 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$; 反之如果 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$, 而且 i 是正常返的, 则由(1.3.19)得 $\lim_{n \rightarrow \infty} p_{ii}^{(nd)} > 0$, 与假设矛盾.

(2) 假设 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 1 / \mu_i > 0$, 这说明状态 i 是正常返的, 且 $\lim_{n \rightarrow \infty} p_{ii}^{(nd)} = 1 / \mu_i$, 与(1.3.19)比较得到 $d = 1$, 故状态 i 是遍历的. 反之由定理 1.3.14 知结论是显然的.

我们称自状态 i 可达状态 j , 并记为 $i \rightarrow j$, 如果存在 $n > 0$ 使 $p_{ij}^{(n)} > 0$, 称状态 i 与状态 j 互通, 并记为 $i \leftrightarrow j$, 如果 $i \rightarrow j, j \rightarrow i$. 由 C-K 方程易知, 如果 $i \rightarrow j, j \rightarrow k$, 则 $i \rightarrow k$, 即可达关系具有传递性; 当 i 是常返状态时, 有 $i \rightarrow i$, 即状态 i 是自返的; 下面证明对常返状态可达关系还具有对称性.

定理 1.3.16 若状态 i 是常返的, 而且 $i \rightarrow j$, 则状态 j 是常返的, 且有 $f_{ji} = 1$.

证明 记

$$\begin{aligned} {}_j p_{ik}^{(n)} &= p_i(X_v \neq j, 1 \leq v \leq n-1, X_n = k), \\ {}_j f_{ik}^{(n)} &= p_i(X_v \notin \{j, k\}, 1 \leq v \leq n-1, X_n = k). \end{aligned}$$

因为 $i \rightarrow j$, 故 $f_{ij} > 0$, 但是

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)} = \sum_{n=1}^{\infty} \left[{}_i f_{ij}^{(n)} + \sum_{r=1}^{n-1} {}_i p_{ii}^{(r)} \cdot {}_i f_{ij}^{(n-r)} \right],$$

故存在 N , 使

$${}_i f_{ij}^{(N)} > 0. \quad (1.3.20)$$

另外

$$0 = 1 - f_{ii} = \sum_{k \neq i} p_{ik}^{(N)} (1 - f_{ki}) \geqslant {}_i p_{ij}^{(N)} (1 - f_{ji}), \quad (1.3.21)$$

所以由(1.3.20), (1.3.21)知 $(1 - f_{ji}) = 0, j \rightarrow i$, 假设 $p_{ji}^{(r)} = \alpha > 0, p_{ij}^{(s)} = \beta > 0$, 则由 C-K 方程知, 对任意正整数 n , 有

$$p_{jj}^{(r+n+s)} \geqslant p_{ji}^{(r)} p_{ii}^{(n)} p_{ij}^{(s)} = \alpha \beta p_{ii}^{(r)}, \quad (1.3.22)$$

由定理 1.3.11 知 $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$, 从而 $\sum_{n=1}^{\infty} p_{jj}^{(n)} = \infty$, 状态 j 是常返的.

下一定理指出互通的状态是同一类型的.

定理 1.3.17 如果 $i \leftrightarrow j$, 则

(1) i 与 j 同为常返状态或非常返状态; 如果同为常返状态, 则它们同为正常返状态或零常返状态.

(2) i 与 j 有相同的周期.

证明 (1) 的前一部分是定理 1.3.16 的直接推论. 现假设 j 为零常返状态, 根据定理 1.3.15 $\lim_{n \rightarrow \infty} p_{jj}^{(n)} = 0$, 于是由(1.3.22)得 $\lim_{n \rightarrow \infty} p_{ii}^{(n)} = 0$, 故 i 也是零常返状态.

(2) 设 i 的周期为 d, j 的周期为 t . 由(1.3.22)知, 对任意使得 $p_{ii}^{(n)} > 0$ 的 n , 必有 $p_{jj}^{(n+r+s)} > 0$, 从而 t 可以除尽 $n+r+s$, 但 $p_{jj}^{(r+s)} \geqslant p_{ji}^{(r)} p_{ij}^{(s)} = \beta \alpha > 0$, 所以 t 也可以除尽 $r+s$, 可见 t 除尽 n . 这表明 $t \leqslant d$, 对称地可得 $t \geqslant d$, 故 $t = d$.

1.3.3 状态空间的分类

定义 1.3.18 状态空间 E 的子集 C 称为(随机)闭集, 如果对任意 $i \in C$ 及 $k \notin C$ 都有 $p_{ik} = 0$. 闭集 C 称为不可分的, 如果 C 的所有状态是互通的. 马氏链 $\{X_n; n=0, 1, \dots\}$ 称为是不可分的, 如果状态空间 E 是不可分的.

定义 1.3.19 称状态 i 为吸收的, 如果 $p_{ii} = 1$.

定理 1.3.16 指出自常返状态只能到达常返状态, 因此 E 中全体常返状态组成一个闭集 C . 在 C 中可达关系具有自返性, 对称性和传递性, 于是可将 C 按互通关系分解, 而得到 E 的如下分解定理.

定理 1.3.20 任意马氏链的状态空间 E , 可惟一地分解成有限个或可列个互不相交的子集 $D, C_1, C_2, \dots$ 之和, 使得

(1) 每一个 C_n 是常返状态组成的不可分的闭集.

(2) C_n 中的所有状态是同类型的, 或全是正常返, 或全是零常返. 它们有相同的周期, 且 $f_{jk} = 1, j, k \in C_n$.

(3) D 由全体非常返状态组成. 自 C_n 中的状态不能到达 D 中的状态.

证明 记 C 为全体常返状态组成的集合, $D = E - C$ 为非常返状态全体, 将 C 按互通关系分解, 状态空间 E 可分解成 $E = D \cup C_1 \cup C_2 \cup \dots$, 其中, 每一个 C_n 是常返状态组成的不可分的闭集, 且由定理 1.3.18 知 C_n 中的所有状态是同类型的. 显然自 C_n 中的状态不能到达 D 中的状态.

1.3.4 $p_{ij}^{(n)}$ 的渐近性质与平稳分布

本段研究 $p_{ij}^{(n)}$ 的渐近性质, 讨论两个问题.

(1) $\lim_{n \rightarrow \infty} p_{ij}^{(n)}$ 是否存在.

(2) 如果此极限存在, 它是否与 i 有关.

定理 1.3.21 如果 j 是非常返或零常返, 则对任意 $i \in E$ 有 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

证明 由 (1.3.7) 式, 我们有

$$p_{ij}^{(n)} = \sum_{k=1}^n f_{ij}^{(k)} p_{jj}^{(n-k)} = \sum_{k=1}^{\infty} I_{[0, n]}(k) f_{ij}^{(k)} p_{jj}^{(n-k)}, \quad (1.3.23)$$

因为 $\sum_{k=1}^{\infty} f_{ij}^{(k)} \leq 1$, 并且根据定理 1.3.11 和定理 1.3.15 知 $\lim_{n \rightarrow \infty} p_{jj}^{(n)} = 0$, 在 (1.3.23) 中, 令 $n \rightarrow \infty$, 根据控制收敛定理, 得到 (1.3.23) 的右边趋于零, 故 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

推论 1.3.22 如果马氏链的状态个数有限, 则不可能全是非常返状态, 也不可能含有零常返状态. 从而不可分的有限马氏链必是正常返的.

证明 设 $E = \{0, 1, \dots, N\}$, 如果全是非常返状态, 则对任意 $i, j \in E$, 由定理 1.3.21 知 $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$, 故有

$$1 = \sum_{j=0}^N p_{ij}^{(n)} \rightarrow 0, \quad n \rightarrow \infty,$$

这就导出矛盾; 其次, 如果 E 含有零常返状态 i , 则 $C = \{j: i \rightarrow j\}$ 是由有限个零常返状态组成的闭集, 由定理 1.3.21 得

$$1 = \sum_{j \in C} p_{ij}^{(n)} \rightarrow 0, \quad n \rightarrow \infty, \quad (1.3.24)$$

得到同样的矛盾.

下面讨论 j 是正常返状态情形, 记

$$f_{ij}(r) = \sum_{m=0}^{\infty} f_{ij}^{(md+r)}, \quad 0 \leq r \leq d-1, \quad (1.3.25)$$

显然

$$\sum_{r=0}^{d-1} f_{ij}(r) = \sum_{m=0}^{\infty} f_{ij}^{(m)} = f_{ij}. \quad (1.3.26)$$