中国科学: 数学 2017 年 第 47 卷 第 10 期: 1223～1254
SCIENTIA SINICA Mathematica

《中国科学》杂志社
SCIENCE CHINA PRESS

综 述

非线性期望的理论、方法及意义

献给李大潜教授 80 华诞

彭实戈[1,2]

1. 山东大学数学学院, 济南 250100;
2. 山东大学高等研究院, 济南 250100
E-mail: peng@sdu.edu.cn

收稿日期: 2016-11-11; 接受日期: 2017-02-14; 网络出版日期: 2017-07-19
国家自然科学基金 (批准号: L1624032, 11526205 和 11626247)、国家外专局 111 研究计划、中国科学院和国家自然科学基金委员会交叉学科战略发展研究资助项目

摘要 本文是非线性期望理论进展的一个综述, 首先给出非线性期望空间的基本定义, 并通过非线性期望的表示定理和几个典型的非线性独立同分布 (i.i.d.) 的例子来说明为什么这个新框架可以广泛地用来分析和计算现实世界 (高维) 数据背后隐藏的概率和统计分布的不确定性; 进而介绍次线性期望空间中两个最重要的统计分布—非线性正态分布和最大分布, 以及相应的非线性大数定律和中心极限定理, 是新领域的基础性和关键性的突破, 其典型的应用就是对于现实的 (高维) 样本数据的非常简单而深刻的 φ-max-mean 算法. 本文还介绍一个最重要的连续时间随机过程—非线性 Brown 运动及相关随机分析, 包括随机积分、随机微分方程和非线性鞅理论. 新的理论框架实质性地推广了 Kolmogorov 于 1933 年建立的、以概率测度为核心的概率论公理体系 $(\Omega,\mathcal{F},P)$. 其关键不同的是, 其核心概念是 (非线性) 期望 $\hat{\mathbb{E}}$, 期望为线性的特殊情形对应着概率论公理体系. 正是这种非线性使人们能够对于现实世界中无处不在的概率模型本身的不确定性也能进行定量的分析和计算. 从而实质性地放宽了概率统计理论中对于现实世界的随机数据的统计假设要求, 本文也因而获得了实际样本数据的非线性分布的 φ-max-mean 算法, 它是一种新的非线性 Monté-Carlo 算法.

关键词 非线性数学期望 非线性正态分布 非线性独立同分布 非线性大数定理和中心极限定理 非线性 Brown 运动及其随机分析 非线性鞅 非线性 Monté-Carlo 算法 φ-max-mean 算法

MSC (2010) 主题分类 60A, 60H10, 60H05, 60H30, 60E05, 60E07, 62-00, 62C05, 62D05, 62F10, 62J05, 65C05, 93E03, 35J, 35K

1 引言

这个综述也可作为非线性期望理论的一个导读, 很多详细的论述和证明可以参考作者放在 arXiv 上的讲义 [115].

英文引用格式: Peng S G. Theory, methods and meaning of nonlinear expectation theory (in Chinese). Sci Sin Math, 2017, 47: 1223–1254, doi: 10.1360/N012016-00209

非线性期望理论发展非常迅速, 这与当前对于很多风险的稳健的量化分析和计算的迫切需求密切相关. 事实上, 随着互联网、计算机科学以及数据和信息技术的飞速发展, 我们所面临的现实世界的动态特性都明显地变得越来越强, 而系统各部分间的相互作用也越来越紧密和复杂, 经常要处理大量高维的随机数据, 并且其不确定性变得越来越大, 从而可能产生的风险也变得越来越难以预料和控制. 非常需要一个新的更基础更深刻的理论和相应的稳健、可靠的分析和计算工具; 需要一个像非线性期望这样的, 能够在概率模型本身还无法确定的情形下, 就可以对各种风险进行稳健的定量分析和计算的数学理论.

自 20 世纪初以来, 概率理论获得了革命性的进展, 并在数理统计、随机控制、信息科学和经济学等很多领域获得了深刻和令世人瞩目的重要应用.

事实上, 数学大师 Kolmogorov 在 1933 年引入的概率论公理体系[75] 在基础理论研究和应用方面都获得了巨大的成功, 堪称概率论王国中的 Euclid 公理体系. 而现在科学家和工程师在处理物理、化学、生命科学、信息科学和工程等方面所涉及的不确定变量的分析和计算问题时, 首先想到的就是应用概率统计或者随机分析的方法. 但是概率理论主要擅长处理那些其相应的概率模型能够通过数理统计方法和数据分析予以确定的情形. 而我们所面对的现实世界中的大部分与人的活动密切相关的那些数据却很难满足这类理想的条件1). 其根本原因是, 相应地, 由概率测度积分所导出的数学期望是线性的. 从线性推进到非线性理论, 其跨度之大不亚于从 Euclid 几何到非 Euclid 几何.

其实早在 1921 年, 芝加哥经济学派的创始人 Frank Knight 就在他的名著《风险、不确定性和利润》中明确指出, 在经济和管理领域中, 概率统计模型本身的不确定性是本质的, 不能消除掉的. 通常称这种概率本身的不确定性为 Knight 不确定性 (或 Ambiguity). 事实上, 不仅在经济和金融领域, 现实世界中绝大多数的不确定量 (随机变量) 都具有程度不同的 Knight 不确定性. 如何分析和计算 Knight 不确定性下的金融和经济问题已经成为当前重要的关注点.

事实上, 能否利用非线性期望理论来分析和计算具有高度动态复杂特性的金融风险, 曾经是长期悬而未决的问题. 此领域的一个特别重要的里程碑是 20 世纪 50 年代初, 法国科学院院士 Choquet 将 Lebesgue 积分的概念推广应用于非可加测度, 获得了一个很重要的非线性期望—Choquet 期望[19]. 但是这种 Choquet 期望不具备动态相容性, 难以适用于我们的瞬息万变的现实世界. 因而在理论上也难以获得可与 Kolmogorov 相当的非线性理论体系.

文献 [103] 发现, 通过引入一个典型的概率空间 $(\Omega,\mathcal{F},P)$ (特别是 Wiener 空间) 上的倒向随机微分方程 (BSDE), 可以构造出一大类动态相容的非线性期望. 这个发现为我们在其后 20 多年的研究中建立一个非线性期望的理论体系提供了关键性的启示. 由于这种期望运算被倒向随机微分方程的生成函数 g 所唯一地确定, 故称其为 g- 期望 (小写的 g- 期望), 而相应的动态风险度量则称为 g- 风险度量 (参见文献 [24,106,125]). 当然, 这种非线性数学期望是定义在线性的 (无穷维的) Wiener 期望空间中的, 故还不是 "内蕴" 的, 相当于在一个 Euclid 空间中来定义一个 Riemann 流形.

有趣的是, 长期以来, 经济学界也正在探索和寻找这样的数学工具, 以刻画消费者对于 Knight 不确定性厌恶的效用函数. 著名的 Allais 经济学实验 (1953) 导致的悖论 (参见文献 [1]) 对作为现代经济学的基石的 von Neumann-Morgenstern 期望效用最大化公理化体系提出了严重的挑战, 而 1961 年由 Ellsberg[36] 设计的系统而精致的经济学实验则直接指出, 实际上人们在做决策时会对 Knight 不确定性有非常显著的厌恶 (aversion), 从而进一步揭示了以非线性来代替线性期望效用公理化体系的必要性. 针对这个重要问题, 1989 年, Schmeidler[126] 提出了非线性的 Choquet 期望效用, 接着, Gilboa 和

1)在概率统计模型中, 通常假设样本数据满足或间接地满足独立同分布条件.

Schmeidler [47] 又提出了基于最小期望效用的最大化原则. 但这些方法只能解决静态情形, 怎样解决现实世界所面对的动态的、特别是连续时间的情形下的问题? 很巧的是, 经济学专家 Epstein 在访问山东大学时, 与当时还在攻读博士学位的陈增敬合作发现, 用一种特殊的次线性的生成函数 g 构成的 g-期望, 就可以很好地解决这个经济学领域中的基础性问题. 他们的研究结果发表在了著名经济学杂志 Econometrica 上 (参见文献 [16]), 并且在经济学界获得了重要影响, 引起了诸如像诺贝尔经济学奖获得者 Hansen 和 Sargent 在内的一批国际知名经济学家的关注, 可参见文献 [2, 14, 49, 50] 等.

原则上讲, 这种 g- 期望的概念可以应用于处理所有的被一个参考概率 P 所控制的不确定概率集合 $\{P_\theta\}_{\theta\in\Theta}$. 但是, 对于那些可能出现 $P(A)=0$ 但 $P_\theta(A)>0$ 的情形, 即所谓奇异情形, g- 期望却无能为力. 在很多领域, 这种奇异情形是非常普遍的, 一个著名的例子就是波动率不确定性. 在 g- 期望提出多年之后, 我们终于领悟到了解决这个困难的关键 [104, 108, 109] 是, 需要跳出概率空间 $(\Omega,\mathcal{F},P)$ 的框架, 从头开始, 直接建立起非线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 的理论框架, 特别地, 通过倒向分析, 我们引入了一种更一般的非线性期望—G- 期望 (大 G 期望).

最近, 在 G- 期望理论进展的直接推动下, 经济学界在这个方向也获得了非常重要的进展, 例如, 经济学专家 Epstein 和 Ji [37, 38], 不仅高度评价 G- 期望理论, 且进一步基于此理论, 在波动率不确定情形下, 推导出了 Radner 均衡下的资本资产定价公式. 一些重要的经济学家对于用 G- 期望解决稳健经济学中的其他的一些关键问题也抱有很大的期望[2]).

我们知道, 基于一个概率空间 $(\Omega,\mathcal{F},P)$ 能够全面系统地获得概率论、随机分析和概率统计理论中的所有重要结论. 相应地, 在非线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 的理论框架中, 一切也都可以从公理体系的基本公设出发, 进一步在概率模型不确定的框架下给出一个全新的理解. 例如, 随机变量的分布、独立性、相关性、平稳性、Markov 过程、非线性扩散过程和非线性半鞅等. 特别地, 我们有相应的非线性 Brown 运动, 而其在实际问题中适用的范围则要比经典的 Brown 运动宽得多. 我们还建立了非线性 Brown 运动驱动下的随机积分理论, 及相应的随机分析理论, 它也是经典的 Itô 随机分析理论的实质性的推广, 并且其结果在原来的线性期望空间下仍然成立. 反之则不然, 一般不能把那些经典概率论和随机分析的结果直接搬到非线性的框架中, 这主要是由于所涉及的不确定概率是可以相互奇异的, 忽视这一点常常会造成根本性的错误.

经典的概率理论主要适用于概率测度能够被确定或者近似地确定的情形, 而非线性概率理论则可以适用于概率测度还无法确定的情形, 适用于非常广泛情形下的定量分析. 一个极端的情形是, 我们只知道未知概率分布的支撑集合落在一个区间 $[\underline{\mu},\overline{\mu}]$ 中, 此时其分布是一个称为最大分布 $M_{[\underline{\mu},\overline{\mu}]}$ 的全新的非线性分布 (见定义 12).

非线性期望理论研究中的一个意外而重要的收获是发现了随机变量的非线性分布与非线性偏微分方程之间的本质关系. 事实上, 在作为非线性期望理论基础的大数定律和中心极限定理这个层次上, 抛物型全非线性偏微分方程 (G- 热方程) 就已经自然而深刻地蕴含于其中. 特别地, 一个随机变量的期望值就是相应的非线性偏微分方程的路径解, 而同时又是一个非线性的鞅. 因此, 我们可以从两个截然不同的角度去理解、分析和计算同一个问题: 从期望的观点, 或从 PDE 的观点. 所得的结果既是概率统计理论的 (非线性) 推广又是非线性 PDE 的在路径解意义上的非平凡推广.

简要地讲, 由于一个概率测度 P 与相应的数学期望 $\mathrm{E}_P[\cdot]$ 有着一一对应的关系, 所以, 遵循 “往最坏处打算” 的原则, 由一个不确定的概率集合 $\{P_\theta\}_{\theta\in\Theta}$ 所产生的期望的风险可以等价地通过

$$\hat{\mathbb{E}}[\cdot]:=\sup_{\theta\in\Theta}\mathrm{E}_{P_\theta}[\cdot]$$

[2]) Hansen L P, Sargent T J. Prices of macroeconomic uncertainties. Private communication, 2016.

来分析和计算, 而这恰恰就是非线性期望理论中最典型的、起着特别作用的次线性期望. 而把 Kolmogorov 概率公理体系的概率空间 $(\Omega,\mathcal{F},P)$ 换为非线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$, 可使我们获得一个处处考虑到了 Knight 不确定性的统一的理论体系. 事实上, 在很多无法忽略概率分布的不确定性的情形下, 可以代之以相应的非线性期望的分析和计算.

非线性期望理论的一个显著特点是它的直观性. 读者也不需要学过、熟悉概率统计理论. 一切从非常基础的非线性期望的公理体系 (见定义 1–3) 出发.

本文结构如下: 第 2 节首先从公理系统的水平建立非线性期望的基本理论框架, 并通过其表示定理说明为什么这个框架可以自然地用来分析概率和统计分布的模型不确定性. 第 3 节则介绍次线性期望空间中两个最典型的分布, 以及这个理论的最基础的定理—大数定律和中心极限定理. 基于此极限定理, 可以直接导出简单而实用的 max-mean 计算方法, 并应用于对大量的不满足经典意义下的 i.i.d. 条件的实际数据的分析和计算. 第 4 和 5 节将介绍一个最重要、最基础的随机过程—非线性 Brown 运动及其相关的随机分析, 以及相应的非线性鞅理论. 最后, 第 6 节介绍当前非线性随机分析的一些进展.

非线性期望理论的发展体现了数学之精妙: 只从非常简单的定义 1 和 2 出发, 即收获这样丰富多彩和意想不到的深刻结果. 作者深信: 其实序幕才刚刚拉开.

2 非线性期望和非线性分布

2.1 非线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$

非线性期望理论的出发点是直接对于不确定量—随机变量来定义其非线性期望泛函. 下面将在公理体系的基础水平上来引入非线性期望的概念.

定义 1 设 Ω 是一个给定的集合[3)], 而其上的一个向量格 $\mathcal{H}$ 是定义在 Ω 上的实值函数所组成的一个线性空间, 且满足以下条件:

(1) 每一个实值的常数 c 都在 $\mathcal{H}$ 中;

(2) 如果 $X(\cdot)\in\mathcal{H}$, 则也有 $|X(\cdot)|\in\mathcal{H}$.

我们把 $\mathcal{H}$ 中的函数称为随机变量, 而称二元组 $(\Omega,\mathcal{H})$ 为随机变量空间.

我们将会在例 8 和 9 中给出例子. 注意到一个 $\mathcal{H}$ 中的随机变量 $X(\omega)$ 是关于自变量 ω 的函数. 它之所以随机是因为我们无法知道哪一个 ω 要发生. 概率论和期望的方法都是, 退而求其次地计算 X 的数学期望, 或者 $\varphi(X)$ 的期望, 其中 φ 是一个给定的函数.

以下为了叙述简便, 不失一般性, 假设 $\mathcal{H}$ 还满足: 如果 $X\in\mathcal{H}$, 则对于每一个 $\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^d)$ 都有 $\varphi(X)\in\mathcal{H}$, 其中 $C_{\mathrm{Lip}}(\mathbb{R}^n)$ 是定义在 $\mathbb{R}^n$ 上的一致 Lipschitz 连续函数全体构成的线性空间. 如果还是有界函数, 则将 Lip 换成 b.Lip.

定义 2 一个非线性期望 $\hat{\mathbb{E}}$ 是定义在随机变量空间 $\mathcal{H}$ 上的满足以下两个性质的 (非线性) 泛函 $\hat{\mathbb{E}}:\mathcal{H}\mapsto\mathbb{R}$:

(1) 单调性, 即对于所有满足 $X(\omega)\geqslant Y(\omega)$ $(\forall\,\omega\in\Omega)$ 的随机变量 $X,Y\in\mathcal{H}$, 都有 $\hat{\mathbb{E}}[X]\geqslant\hat{\mathbb{E}}[Y]$;

(2) 保常数性, 即 $\hat{\mathbb{E}}[c]=c$,

并且称三元组 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 为一个非线性期望空间. 称 $\hat{\mathbb{E}}$ 为一个次线性期望, 如果它还满足

[3)] 这里的 Ω 起着与概率论公理体系中的基本事件的集合 Ω 同样的作用. 但我们无须马上就为其配上一个 σ 代数构成的可测空间 $(\Omega,\mathcal{F})$, 我们将在其后的表示定理中看到, 其实相应的数学结构中会 “自动地” 蕴含一个 (唯一的) σ 代数.

(3) 次线性, 即 $\hat{\mathbb{E}}[X+Y] \leqslant \hat{\mathbb{E}}[X]+\hat{\mathbb{E}}[Y]$, $\hat{\mathbb{E}}[\lambda X]=\lambda\hat{\mathbb{E}}[X]$, $\forall X, Y \in \mathcal{H}$, $\lambda \geqslant 0$;

(4) 如果对于 $X \in \mathcal{H}$, 这个次线性期望还满足 $\hat{\mathbb{E}}[-X]=-\hat{\mathbb{E}}[X]$, 则称 $\hat{\mathbb{E}}$ 为一个线性期望.

定义 3 称一个定义在 $(\Omega, \mathcal{H})$ 上的次线性期望 $\hat{\mathbb{E}}$ 是正则的, 如果对于 $\mathcal{H}$ 中的每一个单调下降的且在每一点 $\omega \in \Omega$ 都有 $\lim_{i\to\infty} X_i(\omega)=0$ 的随机变量序列 $\{X_i\}_{i=1}^{\infty}$, 我们都有 $\hat{\mathbb{E}}[X_i] \downarrow 0$.

次线性期望在非线性期望框架中起着非常重要的作用. 在统计学和经济学中, 这类泛函曾由 Williams [141] 和 Huber [68, 69], 继而又由 Walley [138] 进行了更系统的研究, 而其称谓也各有不同, 如 "indeterminate probabilities"、"上期望" 和 "coherent previsions". 最近一个在定量金融研究中特别引人注目的金融风险的一致风险度量 (coherent risk measure), 实质上也是这样的非线性泛函, 参见文献 [3, 39, 136, 152]. 事实上, 次线性期望等价于关于一族线性期望取上期望, 它是我们熟知的 Hahn-Banach 定理的直接推论, 曾经以不同的形式和条件出现在很多文献中, 如文献 [3, 11, 23, 39, 40, 68, 94].

定理 4 设 $\hat{\mathbb{E}}$ 为一个定义在 $(\Omega, \mathcal{H})$ 上的次线性期望, 则存在定义在 $(\Omega, \mathcal{H})$ 上的一族线性期望 $\{\mathrm{E}_\theta : \theta \in \Theta\}$ 满足

$$\hat{\mathbb{E}}[X]=\max_{\theta\in\Theta} \mathrm{E}_\theta[X]. \tag{2.1}$$

若 $\hat{\mathbb{E}}$ 还是正则的, 则对于每一个 $\theta \in \Theta$, 存在唯一的定义在可测空间 $(\Omega, \sigma(\mathcal{H}))$ 上的概率测度 P_θ 使得

$$\mathrm{E}_\theta[X]=\int_\Omega X(\omega)dP_\theta(\omega), \quad \forall X \in \mathcal{H}, \tag{2.2}$$

其中 $\sigma(\mathcal{H})$ 是 $\mathcal{H}$ 中的随机变量全体生成的最小 σ 代数.

证明概要 我们只给出证明的思想, 详细证明参见文献 [115]. 表示 (2.1) 可以利用 Hahn-Banach 定理直接推出. 如果 $\hat{\mathbb{E}}$ 还是正则的, 则由 (2.1) 立刻有, 对于每一个给定的 $\theta \in \Theta$, $\mathrm{E}_\theta[X_n] \downarrow 0$. 这样我们就可以利用著名的 Daniell-Stone 定理 (参见文献 [151, 定理 3.6.8]) 来证明, 此时存在唯一的定义在可测空间 $(\Omega, \sigma(\mathcal{H}))$ 上的满足 (2.2) 的概率测度 P_θ. □

在遇到一个实际问题中的随机变量, 如明天的温度、明天某股票的价格时, 我们经常想到的是要找到隐藏在其背后的概率. 但是经济学家早就警告我们, 那种指望能够找出其概率来定量地处理经济和金融中的问题的想法常常是不现实的. 而真正经常会遇到的则是一个概率集合 $\{P_\theta\}_{\theta\in\Theta}$, 但是, 原则上, 我们无法确定出其中哪一个概率是 "真的". 我们称这样的 $\{P_\theta\}_{\theta\in\Theta}$ 为一个不确定概率族. 定理 4 恰恰告诉我们可以利用次线性期望 $\hat{\mathbb{E}}$ 来等价地刻画一个不确定概率族 $\{P_\theta\}_{\theta\in\Theta}$. 本文以下所讨论的次线性期望都假设是正则的.

还需提醒的是, 在我们从 worst-case 的角度去定义稳健的次线性期望 $\hat{\mathbb{E}}[X]$ 时, 随机变量 $X(\omega)$ 通常是一个可能发生的损失的随机量. 如果它代表的是收益, 则相应的稳健期望是超线性的,

$$-\hat{\mathbb{E}}[-X]=\min_{\theta\in\Theta} \mathrm{E}_\theta[X].$$

还有一些人是在追求风险的 (risk-seeking). 非线性期望理论也适用于这些问题的分析和计算, 需注意的是要用对方向.

2.2 随机变量的 (非线性) 分布与独立性

概率分布和独立性在概率理论中是至关重要的概念. 而始于 20 世纪 50 年代的非线性期望理论之所以在这么长的时间越来越滞后于概率论的飞速发展, 一个主要原因是, 未能从期望的观点恰当地提出这两个关键概念. 现在介绍我们提出的非线性分布与独立的概念 (参见文献 [109–111]). 它是后面

要介绍的非线性大数定律、中心极限定理和 Brown 运动的随机分析理论的基础, 也充分地表征了现实世界中大量的不确定性的特点. 设 $X=(X_1,\dots,X_n)$ 为非线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个 n- 维随机向量, 记

$$\hat{\mathbb{F}}_X[\varphi]:=\hat{\mathbb{E}}[\varphi(X)]:\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^n)\mapsto\mathbb{R}.$$

由此又导出一个三元组 $(\mathbb{R}^n,C_{\mathrm{Lip}}(\mathbb{R}^n),\hat{\mathbb{F}}_X[\cdot])$, 它显然构成了一个非线性期望空间. 我们称 $\hat{\mathbb{F}}_X$ 为随机向量 X 的分布. 显然, 如果 $\hat{\mathbb{E}}$ 是次线性 (线性) 的, 那么 $\hat{\mathbb{F}}_X$ 也是次线性 (线性) 的. 由定理 4 知, $\hat{\mathbb{F}}_X$ 也有如下表示: 存在定义在 $(\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$ 上的概率测度族 $\{F_X(\theta,\cdot)\}_{\theta\in\Theta}$ 满足

$$\hat{\mathbb{F}}_X[\varphi]=\sup_{\theta\in\Theta}\int_{\mathbb{R}^n}\varphi(x)F_X(\theta,dx),\quad \text{对于每一个 } \varphi\in C_{\mathrm{Lip}}(\mathbb{R}^n).$$

也就是说, $\hat{\mathbb{F}}_X$ 的确可以用来刻画随机向量 X 的分布不确定性.

有了非线性期望下的分布概念, 我们就可以定义多个随机变量的同分布的概念, 其适用范围就广的多了:

定义 5 称定义在次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的两个 n- 维随机向量 X 和 Y 为分布相同的, 记为 $X\overset{d}{=}Y$, 如果

$$\hat{\mathbb{E}}[\varphi(X)]=\hat{\mathbb{E}}[\varphi(Y)],\quad \forall\,\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^n),$$

则称 Y 为 X 的一个复制. 在次线性情形下, 显然, $X\overset{d}{=}Y$是指这两个随机变量有同样的分布不确定性.

我们也称 X 的分布强于 Y (记为 $X\overset{d}{\geqslant}Y$), 或者说 Y 的分布被 X 的所覆盖, 如果

$$\hat{\mathbb{E}}[\varphi(X)]\geqslant\hat{\mathbb{E}}[\varphi(Y)],\quad \forall\,\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^n).$$

如有 $X\overset{d}{\geqslant}Y$, 则显然 X 的分布不确定性比 Y 的大, 我们此时常可取更稳健的次线性期望 $\hat{\mathbb{E}}^*$, 使有 $X\overset{d}{=}Y$. 这也是非线性期望的一个非常关键的优点. 换一种说法, 如果有 $\{F_\theta^X\}_{\theta\in\Theta}\supset\{F_\theta^Y\}_{\theta\in\Theta}$, 由于不确定集合具有主观特性, 有时为了计算和分析的方便, 我们有时可以扩大后者, 而假定两者有相同的分布不确定性. 这样做会使得分析和计算更稳健.

另外, 由于我们考虑的仅是 X 和 Y 的分布, 所以, 即使 X 和 Y 不在一个期望空间中, 仍然可以比较它们的分布.

这种非线性期望下的同分布和分布覆盖的概念也可以推广到比较两个随机变量序列, 以及随机过程的有限维分布.

随机变量 $X\in\mathcal{H}$ 的分布有两个非常典型的参数: 上均值 $\bar{\mu}:=\hat{\mathbb{E}}[X]$ 和下均值 $\underline{\mu}:=-\hat{\mathbb{E}}[-X]$. 如果 $\bar{\mu}=\underline{\mu}$, 则称 X 是均值确定的.

我们还要引入下面推广的独立性的概念:

定义 6 称非线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个随机向量 $Y=(Y_1,\dots,Y_n)$ $(Y_i\in\mathcal{H})$ 独立于另一个随机向量 $X=(X_1,\dots,X_m)$ $(X_i\in\mathcal{H})$, 如果对于每一个试验函数 $\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^m\times\mathbb{R}^n)$, 都有

$$\hat{\mathbb{E}}[\varphi(X,Y)]=\hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x,Y)]_{x=X}].$$

在一个次线性期望 $\hat{\mathbb{E}}$ 下, Y 独立于 X 的意义是, Y 的分布的不确定性不会由于 X的值实现为 $X=x$ $(x\in\mathbb{R}^n)$ 而改变. 特别要注意, Y 独立于 X 并不自动导出 X 也独立于 Y, 即独立性并不对称. 事实上, 现实世界中的随机变量之间的独立性也并非一定是对称的, 不对称的其实反而居多. 这一点特别反映在与很多时间顺序有关的随机变量上. 我们知道时间是有方向的, 已经发生的随机变量与将

要发生的随机变量是非常不对称的. 我们将看到, 对于很多非常重要的结论 (如大数定律和中心极限定理等), 其实单方向的独立性假设已经足够了. 这方面的讨论可参见文献 [115], 而满足这种对称性的一个简明而深刻的充分必要条件是由文献 [58] 获得的.

定义 7　称非线性期望空间 $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ 中的一个 d- 维随机向量序列 $\{\eta_i\}_{i=1}^{\infty}$ 按分布收敛, 如果对于每一个试验函数 $\varphi \in C_{b.\mathrm{Lip}}(\mathbb{R}^n)$, 序列 $\{\hat{\mathbb{E}}[\varphi(\eta_i)]\}_{i=1}^{\infty}$ 都收敛, 其中 $C_{b.\mathrm{Lip}}(\mathbb{R}^n)$ 为 $C_{\mathrm{Lip}}(\mathbb{R}^n)$ 中的所有有界元素.

此时按定义易证, 由极限

$$\hat{\mathbb{F}}[\varphi] := \lim_{i\to\infty} \hat{\mathbb{E}}[\varphi(\eta_i)], \quad \varphi \in C_{b.\mathrm{Lip}}(\mathbb{R}^n)$$

定义的泛函是 $(\mathbb{R}^n, C_{b.\mathrm{Lip}}(\mathbb{R}^n))$ 上的一个非线性期望. 如果 $\hat{\mathbb{E}}$ 是次线性 (线性) 的, 则 $\hat{\mathbb{F}}$ 也是次线性 (线性) 的.

2.3　非线性 i.i.d. 序列的典型和具有普遍意义的例子

以下的三个例子告诉我们, 前面所给出的非线性期望及其分布的看起来比较抽象的概念其实经常会发生在我们的现实生活和工作中.

例 8　一个特别典型的例子就是在一个装着 100 个同样大小和质量、颜色为黑色或白色球的箱子里完全随机地摸出一个球, 但是我们并不知道箱子里具体有多少白球, 只知道白球的数目在 40 至 60 之间. 如摸到白球则赢一元, 否则输一元. 此时的收益是一个随机变量

$$\xi_1(\omega) = \mathbf{1}_{\{\text{摸到白球}\}} - \mathbf{1}_{\{\text{摸到黑球}\}}.$$

一次次地重复这个博弈, 而每次我们都不知道白球数目 W 和黑球数目 (B_t) 的变化情形, 只知道 $W+B = 100$ 且 $40 \leqslant W \leqslant 60$. 记第 i 次的收益为 ξ_i, 则 $\xi_1(\omega), \xi_2(\omega), \ldots$ 构成了一列随机变量. 基于收益 ξ_i 的任何一个合同 $\varphi(\xi_i)$ 的期望值是

$$\hat{\mathbb{E}}[\varphi(\xi_i)] = \hat{\mathbb{E}}[\varphi(\xi_1)] = \max_{p\in[0.4,0.6]}[p\varphi(1) + (1-p)\varphi(-1)], \quad i = 1, 2, \ldots,$$

从而 ξ_i 与 ξ_1 同分布. 同样可以验证, 在次线性期望 $\hat{\mathbb{E}}$ 下, ξ_i 关于 $\xi_1, \ldots, \xi_{i-1}$ 是独立的. 这样就可以分析和计算任何一个金融产品 $X(\omega) = \varphi(\xi_1, \ldots, \xi_i)$ 的损失风险的非线性期望,

$$\hat{\mathbb{E}}[X] = \hat{\mathbb{E}}[\varphi(\xi_1, \ldots, \xi_{i-1}, \xi_i)] = \hat{\mathbb{E}}[\hat{\mathbb{E}}[\varphi(x_1, \ldots, x_{i-1}, \xi_i)]_{\{\xi_j = x_j, 1\leqslant j\leqslant i-1\}}],$$

即 ξ_i 独立于 $\xi_1, \ldots, \xi_{i-1}$. 因此不能通过前 $i-1$ 次的数据结果来获得并改善对于 ξ_i 的分布的不确定性.

例 9　另一个与上例十分相似但却又有实质性的不同的例子则是在一个装着 100 个同样大小和质量而颜色为黑色、白色和黄色球的箱子里完全随机地摸出一个球, 我们还是不知道箱子里各色球的具体比例, 只知道白球数与黑球数相同 $(W = B)$, 黄球数 Y 在 20 到 40 之间. 游戏规则是, 如果摸出的是白球则赢一元, 黑球则输一元 , 而摸出黄球则赢 0 元, 那么这时的随机变量是

$$\xi_1(\omega) = \mathbf{1}_{\{\text{摸到白球}\}} - \mathbf{1}_{\{\text{摸到黑球}\}}.$$

一次次地重复这个博弈, 每次都不知道白球数目 W 和黑球数目 (B_t) 的变化, 只知道 $2W = 2B = 100 - Y$ 且 $20 \leqslant Y \leqslant 40$. 各次博弈的盈亏就构成了一列随机变量 $\xi_1(\omega), \xi_2(\omega), \ldots$ 这时, 我们第 i 次摸球

的收益和第 1 次摸球的收益具有完全相同的分布不确定性:

$$\hat{\mathbb{E}}[\varphi(\xi_i)] = \hat{\mathbb{E}}[\varphi(\xi_1)] = \max_{p\in[0.2,0.4]}\left[\left(1-\frac{p}{2}\right)(\varphi(1)+\varphi(-1))+p\varphi(0)\right], \quad i=1,2,\dots$$

而此时也容易验证, $\{\xi_i\}_{i=1}^{\infty}$ 在次线性期望 $\hat{\mathbb{E}}$ 下是独立同分布的.

例 10 更一般的例子是将上面的 “黑箱” 换成一个随机变量发生器, 每次完全随机地产生一个 d-维的随机向量 ξ_i, 但是其概率分布是不确定的, 记此不确定分布集合为 $\{F_\theta\}_{\theta\in\Theta}$, 如前例一样, 每次我们都完全不知道 θ 会取 Θ 中的哪一个值, 由此产生的 $\{\xi_i\}_{i=1}^{\infty}$ 构成一个分布不确定的 i.i.d. 随机变量序列, 并且

$$\hat{\mathbb{E}}[\varphi(\xi_i)] = \max_{\theta\in\Theta}\int_{\mathbb{R}^d}\varphi(x)F_\theta(dx).$$

现实世界中可以接触到的其实主要是这一类的随机变量数序列.

如何通过现实的数据来获得隐藏于其后次线性分布? 下一节介绍一个非常简单而深刻的 max-mean 算法, 见第 3.3 小节.

有些情形下我们甚至完全不能确定 $\{F_\theta\}_{\theta\in\Theta}$, 而只能确定 $\xi_i(\omega)$ 的取值范围. 这种不确定性构成了一个非常典型的最大分布, 将在下一节专门讨论.

我们将看到, 例 8 和 9 的随机变量序列 $\{\eta_i\}_{i=1}^{\infty}$ 和 $\{\xi_i\}_{i=1}^{\infty}$ 的各项和都可以构成典型的非线性随机游走, 它们都是经典的 Bernoulli 序列的非线性推广, 可以由新的非线性大数定律和中心极限定理来描述其非线性的极限分布 (见下节), 也可通过它们逼近两种非常不同的非线性 Brown 运动. 这是著名的 Donsker 不变原理[30] 的非线性推广 (参见文献 [149]).

3 次线性期望空间中的典型分布、大数定律和中心极限定理

3.1 次线性期望下的正态分布和最大分布

次线性期望空间中当然可以包含所有经典的线性分布, 而特别重要的是, 它还包含大量的新的非线性分布. 首先来介绍其中的一个典型的非线性正态分布, 它相当于概率统计理论中的正态分布. 我们知道, 概率空间 $(\Omega,\mathcal{F},P)$ 中的一个均值为零方差为 σ^2 的随机变量 X 是正态分布的 ($X\overset{d}{=}N(0,\sigma^2)$), 即

$$\mathrm{E}[\varphi(X)] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\varphi(x)\exp\left\{-\frac{x^2}{2\sigma^2}\right\}dx,$$

等价于

$$X+\bar{X}\overset{d}{=}\sqrt{2}X,$$

其中 $\bar{X}$ 是 X 的一个独立复制. 与此完全一样的非线性关系同样也刻画了次线性期望空间中的正态分布随机变量:

定义 11 称次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个 d- 维随机向量 $X=(X_1,\dots,X_d)$ 为零均值正态分布, 如果

$$X+\bar{X}\overset{d}{=}\sqrt{2}X, \quad \forall\, a,b\geqslant 0, \tag{3.1}$$

其中 $\bar{X}$ 是 X 的一个独立复制.

容易验证, 如果 X 满足 (3.1), 则 X 各分量的任何线性组合也会满足 (3.1). 由

$$\hat{\mathbb{E}}[X_i+\bar{X}_i] = 2\hat{\mathbb{E}}[X_i]$$

和

$$\hat{\mathbb{E}}[X_i+\bar{X}_i]=\hat{\mathbb{E}}[\sqrt{2}X_i]=\sqrt{2}\hat{\mathbb{E}}[X_i]$$

得到 $\hat{\mathbb{E}}[X_i]=0$. 同理可得 $\hat{\mathbb{E}}[-X_i]=0$, $i=1,\dots,d$.

记 $\mathbb{S}(d)$ 为所有 $d\times d$ 对称矩阵构成的线性空间, $\mathbb{S}_+(d)$ 为 $\mathbb{S}(d)$ 中的所有 $A\in\mathbb{S}(d)$ 满足 $(Ax,x)\geqslant 0$, $x\in\mathbb{R}^d$. 事实上, 一个正态分布的随机变量 X 的分布完全由以下函数:

$$G(A):=G_X(A):=\frac{1}{2}\hat{\mathbb{E}}[\langle AX,X\rangle],\quad A\in\mathbb{S}(d)\tag{3.2}$$

所唯一地确定. 容易验证, $G:\mathbb{S}(d)\mapsto\mathbb{R}$ 是定义在 $\mathbb{S}(d)$ 上的次线性单调函数, 从而存在一个 $\mathbb{S}_+(d)$ 中的有界的闭子集 Θ, 使得

$$\frac{1}{2}\hat{\mathbb{E}}[\langle AX,X\rangle]=\hat{G}(A)=\frac{1}{2}\max_{Q\in\Theta}\mathrm{tr}[AQ],\quad A\in\mathbb{S}(d).\tag{3.3}$$

若 Θ 只是一个单点集合, 即 $\Theta=\{\sigma^2\}$, 则 X 就是一个经典的均值为零而协方差矩阵为 σ^2 的正态分布. 在一般次线性情形下, Θ 刻画了 X 的协方差不确定性, 我们记为 $X\overset{d}{=}N(\{0\}\times\Theta)$.

次线性期望空间中还有一个在例 10 中讨论到的非常独特且常常会用到的重要分布:

定义 12 称次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个 d- 维随机向量 $Y=(Y_1,\dots,Y_d)$ 是最大分布, 如果存在 $\mathbb{R}^d$ 中的一个有界的凸闭子集合 $\bar{\Theta}$, 使得

$$\hat{\mathbb{F}}_Y[\varphi]=\hat{\mathbb{E}}[\varphi(Y)]=\max_{v\in\bar{\Theta}}\varphi(v),\quad \varphi\in C_{\mathrm{Lip}}(\mathbb{R}^d),$$

记为 $Y\overset{d}{=}M(\bar{\Theta})$. 如果 $\bar{\Theta}$ 是非凸集合, 则称其为非凸的最大分布.

事实上, 最大分布与非线性正态分布的数学表达 (3.1) 有着非常优美的对称性: 根据文献 [115], Y 为最大分布等价于

$$Y+\bar{Y}\overset{d}{=}2Y,\tag{3.4}$$

其中 $\bar{Y}$ 是 Y 的一个独立复制. 以上两个等价条件可以通过一个次线性函数 $g=g_Y(p):\mathbb{R}^d\mapsto\mathbb{R}$ 来联系:

$$g_Y(p):=\hat{\mathbb{E}}[\langle p,Y\rangle].\tag{3.5}$$

事实上, 对于 $\mathbb{R}^d$ 上的任何一个次线性函数 g, 存在 $\mathbb{R}^d$ 中的唯一的有界凸闭子集合 $\bar{\Theta}\in\mathbb{R}^d$ 满足

$$g(p)=\sup_{q\in\bar{\Theta}}\langle p,q\rangle,\quad p\in\mathbb{R}^d.\tag{3.6}$$

性质 13 如果次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个 d- 维随机向量 Y 满足条件 (3.4), 则它具有下述类型的最大分布:

$$\hat{\mathbb{F}}_Y[\varphi]=\hat{\mathbb{E}}[\varphi(Y)]=\max_{v\in\bar{\Theta}}\varphi(v),\quad \varphi\in C_{\mathrm{Lip}}(\mathbb{R}^d),$$

其中集合 $\bar{\Theta}$ 由 (3.5) 和 (3.6) 来确定.

其实人们常会碰到这样的随机变量 X, 其概率分布完全未知, 而仅能确定 $\underline{\mu}\leqslant X\leqslant\overline{\mu}$. 此时很容易误将 X 的分布取为均匀分布 $U_{[\underline{\mu},\overline{\mu}]}$, 即在区间 $[\underline{\mu},\overline{\mu}]$ 上的等概率分布, 而这种错误常会导致很大的风险, 所以要特别注意等概率分布 $U_{[\underline{\mu},\overline{\mu}]}$ 与分布为同等程度的未知 $M_{[\underline{\mu},\overline{\mu}]}$ 这两个概念本质的不同. 显然有

$$M_{[\underline{\mu},\overline{\mu}]}\overset{d}{\geqslant}U_{[\underline{\mu},\overline{\mu}]}.$$

注 14 更一般地, 我们称一对 d- 维随机向量 (X,Y) 为 G- 分布, 如果

$$(X+\bar{X}, Y+\bar{Y}) \overset{d}{=} (\sqrt{2}X, 2Y), \tag{3.7}$$

其中 $(\bar{X},\bar{Y})$ 为 (X,Y) 的独立复制. 此时可以验证 X 是 G- 正态分布, Y 是最大分布, 而 (X,Y) 的分布由以下函数刻画:

$$G(p,A) := \hat{\mathbb{E}}\left[\frac{1}{2}\langle AX, X\rangle + \langle p, Y\rangle\right], \quad (p,A)\in\mathbb{R}^d\times\mathbb{S}(d). \tag{3.8}$$

容易验证 $G:\mathbb{R}^d\times\mathbb{S}(d)\mapsto\mathbb{R}$ 是一个关于 A 单调的次线性函数. 显然 G 也是一个连续函数. 从而存在一个凸闭集合 $\Gamma\subset\mathbb{R}^d\times\mathbb{S}_+(d)$ 满足

$$G(p,A) = \sup_{(q,Q)\in\Gamma}\left[\frac{1}{2}\mathrm{tr}[AQ] + \langle p, q\rangle\right], \quad \forall\,(p,A)\in\mathbb{R}^d\times\mathbb{S}(d). \tag{3.9}$$

以下结果表明, 每一个满足 (3.9) 的函数 G 都唯一地确定了一个 G- 正态分布.

性质 15 (参见文献 [112, 命题 4.2]) 设 $G:(p,A)\in\mathbb{R}^d\times\mathbb{S}(d)\mapsto\mathbb{R}$ 是次线性的, 且为关于 A 单调的函数 (从而 G 可由 (3.9) 来表示), 则存在一个次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的满足 (3.7) 和 (3.8) 的一对 d- 维随机向量 (X,Y), 而 (X,Y) 的分布唯一地被函数 G 确定. 更具体地, 给定任意的函数 $\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^d)$, 由下式定义的函数 u:

$$u(t,x,y) := \hat{\mathbb{E}}[\varphi(x+\sqrt{t}X, y+tY)], \quad (t,x,y)\in[0,\infty)\times\mathbb{R}^d\times\mathbb{R}^d \tag{3.10}$$

是以下抛物型偏微分方程的唯一的 (黏性) 解:

$$\partial_t u - G(D_y u, D_x^2 u) = 0, \quad u\,|_{t=0} = \varphi, \tag{3.11}$$

其中

$$D_y = (\partial_{y_i})_{i=1}^d, \quad D_x^2 = (\partial^2_{x_i,x_j})_{i,j=1}^d.$$

注 16 一般地, 我们可以用黏性解的概念来刻画偏微分方程 (3.11) 的解[4)], 而 $G(p,A)$ 的次线性和关于 A 的单调性保证了上述 PDE (3.11) 的解的存在唯一性. 不仅如此, 我们还证明了 u^φ 关于初始条件 φ 也是次线性的, 而这一条是非常关键的.

我们还注意到 PDE (3.11) 的黏性解理论的一个特点, 即方程可以是部分乃至完全退化的, 这里的完全退化即 G 不含 A. 如果 (3.11) 是非退化的, 则方程存在唯一的光滑的解 $u\in C^{1+\frac{\alpha}{2},2+\alpha}((0,\infty)\times\mathbb{R}^d)$, 参见文献 [13,76,139] 以及其推荐的相关参考书目. 当 $d=1$ 且 G 仅依赖于二阶导数项 $D_x^2 u$ 时, G- 热方程就是 Baronblatt- 方程, 参见文献 [5,87].

如果 (X,Y) 和 $(\bar{X},\bar{Y})$ 都是同一个函数 G 的 G- 正态分布,

$$\begin{aligned}G(p,A) :=&\, \hat{\mathbb{E}}\left[\frac{1}{2}\langle AX, X\rangle + \langle p, Y\rangle\right]\\ =&\, \hat{\mathbb{E}}\left[\frac{1}{2}\langle A\bar{X}, \bar{X}\rangle + \langle p, \bar{Y}\rangle\right], \quad \forall\,(p,A)\in\mathbb{S}(d)\times\mathbb{R}^d,\end{aligned}$$

[4)] 黏性解是一个非常重要的概念, 由 Crandall 和 Lions [21] 引入, 可以参见文献 [20] 中对于一般 PDE 的黏性解理论系统的介绍, 也可参见文献 [112,115] 的附录中对于针对此类方程的专门介绍.

则由方程 (3.11) 的解的唯一性可以证明 $(X,Y)\overset{d}{=}(\bar{X},\bar{Y})$. 特别地, $X\overset{d}{=}-X$.

若 (X,Y) 为 G- 正态分布, 对于每一个 $\psi\in C_{\mathrm{Lip}}(\mathbb{R}^d)$, 定义一个函数

$$v(t,x):=\hat{\mathbb{E}}[\psi(x+\sqrt{t}X+tY)],\quad (t,x)\in[0,\infty)\times\mathbb{R}^d,$$

则 v 是以下抛物型偏微分方程:

$$\partial_t v-G(D_x v,D_x^2 v)=0,\quad v\,|_{t=0}=\psi \tag{3.12}$$

的唯一解. 还可以验证 $v(t,x+y)\equiv u(t,x,y)$, 其中 u 是偏微分方程 (3.11) 的初始条件为 $u(t,x,y)\,|_{t=0}=\psi(x+y)$ 的唯一的解.

以下性质表明, 任何有界的随机序列都可以被次线性期望下的 i.i.d. 序列的分布所覆盖, 也就是说, 次线性期望下的 i.i.d. 假设可以应用于非常多的实际的随机数据.

性质 17　假设 $\{Y_i\}_{i=1}^{\infty}$ 是次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的随机变量序列, 且有

$$Y_1\overset{d}{=}M_{[\underline{\mu},\overline{\mu}]},$$

$\{X_i\}_{i=1}^{\infty}$ 是一个概率空间 $(\Omega,\mathcal{F},\mathbb{P})$ 中的实值随机变量空间序列且 $\underline{\mu}\leqslant X_i(\omega)\leqslant\overline{\mu}$, $\mathbb{P}$-a.s., 则 $\{X_i\}_{i=1}^{\infty}$ 的有限维分布被 $\{Y_i\}_{i=1}^{\infty}$ 的所覆盖, 即对于任何的正整数 N 和任何的 $\varphi\in C_b(\mathbb{R}^N)$, 都有

$$\hat{\mathbb{E}}[\varphi(Y_1,\dots,Y_N)]\geqslant \mathrm{E}_{\mathbb{P}}[\varphi(X_1,\dots,X_N)].$$

这个性质的证明非常容易, 且很容易推广到高维随机向量序列的情形.

3.2　非线性大数定律和中心极限定理

在概率和统计理论体系中, 著名的大数定律和中心极限定理显然占据中心和基础性的重要位置. 原因是我们经常能够接触到的很多现象是很多微小的随机因素积累起来, 共同作用所产生的效果, 这使我们可以利用这些极限定理来对于很多随机量进行定量的分析和计算.

另一方面, 我们的现实世界中积累了巨量的与人们的各种活动密切相关的数据, 但是, 正像前面所论述的, 这些数据有不可忽视的概率分布模型的不确定性, 而金融数据就是非常典型的例子: 数据很难被假设其为经典意义下的 i.i.d. 但是正如前面所论述的, 我们完全可以将其中的很多数据稳健地假设为在非线性期望意义下的 i.i.d. 在这种情形下, 如何获得非线性期望理论框架下的大数定律和中心极限定理也自然成了一个基础性的重大问题.

定义 18　我们称非线性期望空间中的 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个 $\mathbb{R}^d$- 值的随机变量序列 $\{X_i\}_{i=1}^{\infty}$ 为独立同分布, 简记为 i.i.d, 如果对于每一个 $i=1,2,\dots$, 有 $X_{i+1}\overset{d}{=}X_i$ 且 X_{i+1} 独立于 $\{X_1,\dots,X_i\}$.

注 19　注意到我们的非线性期望下 i.i.d. 条件只需要假设一个方向的独立性—倒向独立性. 其实我们所获得的与人文、管理、经济和金融有关的大部分数据都很难满足概率论意义下的 i.i.d. 条件, 而第 2.3 小节表明, 它们可满足, 或近似地满足, 由定义 18 所给出的次线性 i.i.d. 条件.

定理 20 (大数定理[110, 115])　设 $\{X_i\}_{i=1}^{\infty}$ 为 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的一个 $\mathbb{R}^d$- 值的 i.i.d. 的随机变量序列, 且满足

$$\lim_{c\to\infty}\hat{\mathbb{E}}[(|X_1|-c)^+]=0, \tag{3.13}$$

则序列 $\{\sum_{i=1}^n \frac{X_i}{n}\}_{n=1}^\infty$ 按分布收敛到一个最大分布

$$\lim_{n\to\infty}\hat{\mathbb{E}}\left[\varphi\left(\sum_{i=1}^n \frac{X_i}{n}\right)\right]=\max_{\mu\in\Gamma}[\varphi(\mu)], \tag{3.14}$$

对于所有的满足线性增长条件的函数 $\varphi\in C(\mathbb{R}^d)$, Γ 是由下式决定的 $\mathbb{R}^d$ 中的有界凸闭集:

$$\max_{\mu\in\Gamma}\langle\mu,p\rangle=\hat{\mathbb{E}}[\langle p,X_1\rangle],\quad p\in\mathbb{R}^d.$$

特别地, 当 $d=1$ 时, 有 $\Gamma=[\underline{\mu},\overline{\mu}]$, 其中

$$\overline{\mu}=\hat{\mathbb{E}}[X_1],\quad \underline{\mu}=-\hat{\mathbb{E}}[-X_1].$$

定理 21 (大数定理 - 中心极限定理[110, 115]) 设 $\{(X_i,Y_i)\}_{i=1}^\infty$ 为 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的 $\mathbb{R}^{2d}$- 值的 i.i.d. 随机变量序列. 我们还假设 $\hat{\mathbb{E}}[Y_1]=\hat{\mathbb{E}}[-Y_1]=0$, 及 (3.13) 和

$$\lim_{c\to\infty}\hat{\mathbb{E}}[(|Y_1|^2-c)^+]=0. \tag{3.15}$$

记

$$\bar{S}_n:=\sum_{i=1}^n\left(\frac{X_i}{n}+\frac{Y_i}{\sqrt{n}}\right),$$

则随机向量序列 $\{\bar{S}_n\}_{n=1}^\infty$ 按分布收敛:

$$\lim_{n\to\infty}\hat{\mathbb{E}}[\varphi(\bar{S}_n)]=\hat{\mathbb{E}}[\varphi(\xi+\zeta)],\quad \forall\,\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^d), \tag{3.16}$$

其中随机变量对 (ξ,ζ) 为 G- 正态分布随机向量且相应的次线性函数 $G:\mathbb{R}^d\times\mathbb{S}(d)\mapsto\mathbb{R}$ 由下式给出:

$$G(p,A):=\hat{\mathbb{E}}\left[\langle p,X_1\rangle+\frac{1}{2}\langle AY_1,Y_1\rangle\right],\quad p\in\mathbb{R}^d,\quad A\in\mathbb{S}(d).$$

推论 22 以上的定理包含了两个重要定理作为特殊情形:

(1) $\sum_{i=1}^n\frac{Y_i}{\sqrt{n}}$ 按分布收敛于非线性正态分布 $N(\{0\}\times\hat{\Theta})$, 其中子集合 $\hat{\Theta}\subset\mathbb{S}_+(d)$ 由 (3.3) 定义,

$$\hat{G}(A)=G(0,A);$$

(2) $\sum_{i=1}^n\frac{X_i}{n}$ 按分布收敛于最大分布 $M(\bar{\Theta})$, 其中集合 $\bar{\Theta}\subset\mathbb{R}^d$ 由 (3.6) 定义,

$$\bar{G}(p)=G(p,0),\quad p\in\mathbb{R}^d.$$

特别地, 如果取

$$\varphi(y)=d_{\bar{\Theta}}(y)=\inf\{|x-y|:x\in\bar{\Theta}\},$$

我们有以下推广的大数定律:

$$\lim_{n\to\infty}\hat{\mathbb{E}}\left[d_{\bar{\Theta}}\left(\sum_{i=1}^n\frac{X_i}{n}\right)\right]=\sup_{\theta\in\bar{\Theta}}d_{\bar{\Theta}}(\theta)=0. \tag{3.17}$$

如果 X_i 有确定的均值, 即 $\bar{\Theta}$ 是一个单点集 $\bar{\Theta}=\{\bar{\theta}\}$, 则 (3.17) 成为强收敛:

$$\lim_{n\to\infty}\hat{\mathbb{E}}\left[\left|\sum_{i=1}^n\frac{X_i}{n}-\bar{\theta}\right|\right]=0.$$

但可以证明, 当其没有确定均值时不可能有强收敛, 此时非线性的大数定律的按分布收敛是实质性的结果.

注 23 以上的定理的证明中采用了与传统证明有很大不同的新方法 (参见文献 [110,115]): 证明借用了对全非线性偏微分方程解的非常深刻的估计 (参见文献 [13,76,139]), 从而使得对这个深刻的定理的证明直观而简单. 也进一步揭示了不确定量的积累与相应的非线性偏微分方程的深刻而本质的内在联系.

非线性期望框架下的大数定律和中心极限定理表明了极大分布和 G- 正态分布是普适性很强的分布. 它还提供了一个研究模型不确定的统计推断和误差分析等的基本工具. 上面的两个定理由文献 [110, 115] 首先在关于 X_1 和 Y_1 的较强的矩条件下获得, 而上述证明方法对于更弱的条件 (3.13) 和 (3.15) 仍然适用. 条件 (3.13) 和 (3.15) 等价于

$$\lim_{n\to\infty}\bar{\mathbb{E}}[|X_1|\mathbf{1}_{|X_1|>n}]=0,$$
$$\lim_{n\to\infty}\bar{\mathbb{E}}[|Y_1|^2\mathbf{1}_{|Y_1|>n}]=0.$$

本文没有采用以上我们常用的条件, 而采用了张立新提出的条件 (3.13), (3.15) 是由于后者更适于非线性期望形式的表述. 文献 [116] 进一步引进和发展了非线性期望框架下胎紧和弱收敛技术, 据此给出了中心极限定理新的证明方法. 这方面的研究也可参见文献 [29]. Chen [15] (也见 Chen 等 [17]) 在此基础上通过容度理论证明了非线性期望框架下的强大数定律. 此处的强大数定律并不意味着定理 20 中的 $(X_1+\cdots+X_n)/n$ $(n=1,2,\ldots)$ 是在 q.s. (quasi surely) 或按照容度是强收敛的, 因为非线性大数定律的按分布收敛于最大分布是实质性的, 只有当上均值等于下均值时才会变为强收敛; Zhang [148–150] 及 Lin 和 Zhang [85] 进一步系统地发展了非线性期望框架下的泛函中心极限定理和重对数律等极限理论和 Donsker 不变原理 (参见文献 [30]), 并且获得了相应的自正则化 (self-normalized) 的收敛结果. 另一方面, 文献 [60] 通过非线性偏微分积分方程定量分析了非线性 Lévy 分布, 据此, Bayraktar 和 Munk [8] 研究了非线性框架下的稳定分布和广义中心极限定理, 这些研究结果极大地促进了非线性极限理论的发展. 关于这一方面的更多研究, 参见文献 [67, 78, 142] 等相关文献.

注意到此前经济学界已有在容度意义下极限理论的研究 (参见文献 [88, 89]), 但其结果从条件、结论和证明方法上都与我们所获得的有着实质的不同.

3.3 关于实际样本数据的非线性分布的 φ-max-mean 算法

在实际问题中我们所碰到的将来要发生的不确定量, 都可以视作一个随机变量 X, 而在大多数情形下, 我们能做的其实就是根据历史数据 (样本数据) 来获得关于我们所关心的量 $\varphi(X)$ 的期望. 这个函数 φ 对于不同的情形有着不同的意义. 例如, $\varphi(X)$ 可以是基于 X 的一个金融合同, 例如, 一个看跌期权 $\varphi(x)=\max\{0,x-k\}$, 或者一个消费函数、损益函数, 或者一个最优控制系统中的代价函数等.

在经典的概率空间 $(\Omega,\mathcal{F},P)$ 中, 从理论上讲, X 的线性分布可以通过大数定律通过下式来获得:

$$\mathrm{E}[\varphi(X)]=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\varphi(x_i),$$

其中 $\{x_i\}_{i=1}^{\infty}$ 是随机变量 X 的 i.i.d. 的样本序列. 就是说我们可以通过以下均值算子:

$$\mathbb{M}[\varphi(X)]:=\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\varphi(x_i):C_{\mathrm{Lip}}(\mathbb{R}^d)\mapsto\mathbb{R}$$

来获得 X 的分布. 实际上这就是 Monté-Carlo 逼近算法, 而算法的合理性则是基于经典的强大数定理. 但事实上, Monté-Carlo 方法所使用的数据绝大多数是通过计算机来产生的, 可充分接近经典 i.i.d 的

数据. 但是如果样本是实际数据, 则常常无法保证其满足经典的 i.i.d. 条件. 但大部分实际数据却可以满足, 或近似地满足, 我们提出的非线性的 i.i.d. 条件 (见第 2.3 小节) .

下面介绍如何应用非线性大数定理, 通过数据 $\{x_i\}_{i=1}^{nm}$ 来获得更稳健的次线性分布下的期望值. 考虑一个次线性期望空间 $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$, 设 x_i $(i=1,2,\ldots,n\times m)$ 为一个 i.i.d. 的样本. 这时可引入下面的 φ-max-mean 计算程序来获得关于 $\overline{\mu}_{\varphi(X)}=\hat{\mathbb{E}}[\varphi(X)]$ 估计:

$$\hat{\mathbb{M}}[\varphi](X_1,\ldots,X_{mn})=\hat{\mathbb{M}}[\varphi]=\max\{Y_n^k:k=1,\ldots,m\},$$

其中

$$Y_n^k=\frac{1}{n}\sum_{i=1}^{n}\varphi(x_{n(k-1)+i}),\quad \varphi\in C_{\mathrm{Lip}}(\mathbb{R}^d).$$

事实上, 根据非线性大数定律 (定理 20), 当 n 足够大时, $\{Y_n^k\}_{k=1}^m$, 按分布收敛于 i.i.d 的 $\{Y^k\}_{k=1}^m$, 并且 Y^k 又都满足最大分布 $M([\underline{\mu}_{\varphi(X)},\overline{\mu}_{\varphi(X)}])$, 而这样, $\max\{Y^k:k=1,\ldots,m\}$ 就给了我们关于

$$\overline{\mu}_{\varphi(X)}=\hat{\mathbb{E}}[\varphi(X)]\quad 和 \quad \underline{\mu}_{\varphi(X)}=-\hat{\mathbb{E}}[-\varphi(X)]$$

的渐进无偏最优估计. 我们的研究表明这个方法可用来改善金融衍生证券的风险控制中对风险度量的计算[5)]. 其无偏估计的最优性则基于下面的一个 (最近获得的) 非常简单而又基础的定理:

定理 24 [74] 设 $Y^1,\ldots,Y^m$ 为取自最大分布的 i.i.d. 样本:

$$Y^i\overset{d}{=}M_{[\underline{\mu},\overline{\mu}]},\quad i=1,\ldots,m,$$

其中 $\underline{\mu}\leqslant\overline{\mu}$ 是两个未知参数, 则 q.s. (即拟必然的, 或对于每一个 $\theta\in\Theta$ 都 P_θ- 几乎必然的),

$$\underline{\mu}\leqslant\min\{Y^1(\omega),\ldots,Y^n(\omega)\}\leqslant\max\{Y^1(\omega),\ldots,Y^n(\omega)\}\leqslant\overline{\mu},$$

并且

$$\widehat{\overline{\mu}}_n=\max\{Y^1,\ldots,Y^n\}$$

为关于上均值 $\overline{\mu}$ 的最大无偏估计,

$$\widehat{\underline{\mu}}_n=\min\{Y^1,\ldots,Y^n\}$$

为关于下均值 $\underline{\mu}$ 的最小无偏估计.

有读者可能会问: 如果要用 max-mean 方法来获得关于实际数据的 i.i.d 样本 $\{X_i\}_{i=1}^N$ 的随机变量 X 的非线性分布 $\hat{\mathbb{E}}[\varphi(X)]$, 岂不是要取遍所有的函数 φ? 回答是, 这种做法当然是不现实的, 其实正如经典的 Monté-Carlo 算法一样, 在处理实际问题中, 只需计算其所关心的函数 φ 所对应的非线性期望值 $\hat{\mathbb{E}}[\varphi(X)]$, 这种函数的实际例子很多, 如均值、方差、协方差和效用函数, 以及各种各样的合同所对应的不同的损益函数等.

另外, 上面的对于数据的分组方式有很多种, 而我们只选了其中最简单的一种介绍, 数据分组的设计实际上非常重要, 与当事者所处的环境和要做出的判断的实际情形等都有密切的关系.

以上方法也可以用于 X 是很高维数的随机向量的情形, 而计算量的增加则只与维数成正比.

另一方面, 还有许多随机变量遵守一些典型的非线性分布模型, 如非线性正态分布和非线性 Poisson 分布等, 例如, 我们可以通过中心极限定理来确定 X 的分布为 $N(\mu,[\underline{\sigma}^2,\overline{\sigma}^2])$. 此时我们可以对其

5) Peng S, Yang S. Empirical tested of G-VaR. Preprint.

i.i.d. 样本运用 max-mean 算法来估计参数 μ、$\underline{\sigma}^2$ 和 $\overline{\sigma}^2$. 当 μ 已经被确定时, 我们可以获得关于下方差 $\underline{\sigma}^2$ 和上方差 $\underline{\sigma}^2$ 的最优渐进无偏估计:

$$\hat{\underline{\sigma}}^2 := \min_{1\leqslant k\leqslant m} \sigma_k^2, \quad \hat{\overline{\sigma}}^2 := \max_{1\leqslant k\leqslant m} \sigma_k^2,$$

其中

$$\sigma_k^2 := \frac{1}{n}\sum_{j=1}^{n}(x_{n(k-1)+j} - \mu)^2.$$

实际上一直有统计学专家在研究和探索样本数据的参数估计, 典型的研究结果参见文献 [77].

4 非线性 Brown 运动及相应的随机分析

我们知道, 概率论在 20 世纪上半叶获得了巨大的革命性进展, 其中非常突出的是关于连续时间随机过程的随机分析理论, 特别是 Brown 运动 (也称为 Wiener 过程) 的数学理论以及相应的 Itô 随机分析理论的建立. 由此形成了概率论, 随机分析在 20 世纪下半叶的蓬勃发展, 对于科学技术发展产生了巨大的影响. 一个重要的挑战性问题是, 我们是否能够在非线性期望的理论体系中建立起非线性 Brown 运动的概念及其相应的数学理论呢? 这个问题在文献 [109] 中获得了解决. 本节将给以简要介绍 (参见文献 [109, 113, 115]).

4.1 次线性期望下的 Brown 运动

定义 25 称非线性期望空间 $(\Omega, \mathcal{H}, \mathbb{E})$ 上的一个映射 $\eta_t(\omega) : \Omega\times[0,\infty) \to \mathbb{R}^d$ 为一个 d- 维随机过程, 如果对每一个 $t\in[0,\infty)$, 都有 $\eta_t(\cdot)\in\mathcal{H}^d$.

本节介绍次线性期望空间中的 Brown 运动, 也称为 G- 期望下的 Brown 运动, 或 G-Brown 运动. 注意这里用的是大写的 G, 它是 g- 期望 (小写的 g) 理论的完全非线性的推广. 与其相应的鞅也是一个全非线性的 PDE 的路径解.

我们以下引入非线性期望下的 Brown 运动, 为叙述简单, 我们仅考虑一维 G-Brown 运动, 高维情形可参见文献 [113, 115]. 注意到, 非线性 Brown 运动的每一个增量的分布虽然都是 G- 正态的, 但是 Brown 运动本身不再是一个非线性的 Gauss 过程[118].

定义 26 称定义在一个次线性期望 $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ 空间中的随机过程 $B_t(\omega)$ $(t\geqslant 0)$ 为关于 $\hat{\mathbb{E}}$ 的 Brown 运动, 如果对每一个 $n\in\mathbb{N}$ 和 $0\leqslant t_1,\ldots,t_n<\infty$, 都有

(1) $B_0(\omega)=0$;

(2) (B_t) 的增量平稳且独立, 即对于每一个 $t,s\geqslant 0$, $B_{t+s}-B_t \overset{d}{=} B_s$, 并且对于所有 $n\in\mathbb{N}$ 和 $t_1,\ldots,t_n\in[0,t]$, $B_{t+s}-B_t$ 独立于 $(B_{t_1}, B_{t_2},\ldots,B_{t_n})$;

(3) $|B_t|^3\in\mathcal{H}$ 且 $\lim_{t\downarrow 0}\hat{\mathbb{E}}[|B_t|^3]/t\to 0$.

称 (B_t) 为一个对称的 Brown 运动, 如果它还满足 $\hat{\mathbb{E}}[B_t] = -\hat{\mathbb{E}}[-B_t]=0$.

若存在一个定义在 $(\Omega,\mathcal{H})$ 上的按下述意义被次线性期望 $\hat{\mathbb{E}}$ 控制的另一个非线性期望 $\widetilde{\mathbb{E}}$:

$$\widetilde{\mathbb{E}}[X]-\widetilde{\mathbb{E}}[Y]\leqslant\hat{\mathbb{E}}[X-Y], \quad X,Y\in\mathcal{H},$$

并且 (B_t) 在 $\widetilde{\mathbb{E}}$ 下仍然是增量独立和平稳的 (即 (2) 在 $\widetilde{\mathbb{E}}$ 下仍然成立), 则称 (B_t) 为 $\widetilde{\mathbb{E}}$ 下的一个 Brown 运动. 注意 $\widetilde{\mathbb{E}}$ 不必一定是次线性的, 从而可以应用于更一般的情形.

条件 (3) 保证了 Brown 运动 (B_t) 会具有连续的轨道. 而这个条件如果不满足, 则 (B_t) 成为一个更一般的非线性 Lévy 过程 (参见文献 [60]).

定理 27 设 $(B_t)_{t\geqslant 0}$ 为定义在次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 上的一个对称的一维 Brown 运动, 则

$$\frac{B_t}{\sqrt{t}}\overset{d}{=}N(0,[\underline{\sigma}^2,\overline{\sigma}^2]),$$

其中 $\overline{\sigma}^2=\hat{\mathbb{E}}[\tilde{B}_1^2]$, $\underline{\sigma}^2=-\hat{\mathbb{E}}[-\tilde{B}_1^2]$. 如果还有 $\underline{\sigma}^2=\overline{\sigma}^2>0$, 则由 $(B_t/\overline{\sigma}^2)_{t\geqslant 0}$ 定义的随机过程的有限维分布与经典的一维标准 Brown 运动完全一致.

我们将一个次线性期望空间中的 Brown 运动 $(B_t)_{t\geqslant 0}$ 称为 G-Brown 运动. 这时 B_t 满足 G- 正态分布, 其生成函数 G 由下式定义:

$$G(\alpha):=\frac{1}{2}\hat{\mathbb{E}}[\alpha B_1^2],\quad \alpha\in\mathbb{R}.$$

容易验证, 对于每一个 $\lambda>0$, $t_0>0$, $(\lambda^{-\frac{1}{2}}B_{\lambda t})_{t\geqslant 0}$ 和 $(B_{t+t_0}-B_{t_0})_{t\geqslant 0}$ 都是对称的具有相同生成函数 G 的 Brown 运动, 即一个对称 G-Brown 运动满足与经典情形相同的伸缩和平移不变性.

与对称 Brown 运动完全不同的是具有有限变差的 Brown 运动 (b_t), 它满足 $b_{s+t}-b_s\overset{d}{=}M_{[\underline{\mu},\overline{\mu}]}$, 注意到这导致了

$$\underline{\mu}t\leqslant b_{s+t}(\omega)-b_s(\omega)\leqslant\overline{\mu}t,$$

从而, (b_t) 的每一条轨线 $b_t(\omega)$ 都是速度限制在 $[\underline{\mu},\overline{\mu}]$ 之间的关于 t 的 Lipschitz 函数. 一个特别的例子是 $\overline{\mu}=c$, $\underline{\mu}=-c$, 此处 c 是光速. 我们知道现实世界中的任何粒子的速度都是小于光速, 所以, 它们可以用 (b_t), 而不宜用对称 Brown 运动 (包括经典的标准 Brown 运动) 来描述. 还可以证明, 这样的随机过程 (b_t) 的不确定分布恰好覆盖了速度限制在 $[\underline{\mu},\overline{\mu}]$ 以内的全部可能的随机过程的 (线性) 分布.

4.2 *G*-Brown 运动的构造—存在性定理

经典 Brown 运动的构造是 Wiener 在 1923 年获得的 (参见文献 [140]). 我们知道构造一个经典的标准 Brown 运动实际上可以通过构造连续轨道空间 $\Omega=C([0,\infty),\mathbb{R})$ 上的一个适当的概率测度 -Wiener 测度 P 来实现. 也就是说用这个 P 来测度 $\omega\in\Omega$, 则全体过程 ω 就构成了一个标准 Brown 运动. 构造一个非线性期望下的 Brown 运动的逻辑也是同样的, 不同的是, 这里需要构造的是一个适当的非线性期望 $\hat{\mathbb{E}}$, 用它来测度 $\omega\in\Omega$, 则全体过程 ω 就构成了一个非线性的 Brown 运动. 而最自然的仍是采用 Wiener 和 Kolmogorov 都使用过的方法: 先建立 G-Brown 运动的有限维 (次线性) 分布所对应的次线性期望, 进而将其所在的空间完备化.

下面简要介绍如何构造一个对称 G-Brown 运动, 更系统的研究可参见文献 [114, 115, 117]. 记 $\Omega=C([0,\infty),\mathbb{R})$ 为所有定义在 $[0,\infty)$ 上的实值的且初值为零的连续函数全体. 对于每一个给定的 $T\geqslant 0$, 考虑以下随机变量空间 $(\Omega_T,\mathcal{H}_T)=(\Omega_T,\mathrm{Lip}(\Omega_T))$:

$$\Omega_T:=\{\omega(t\wedge T),t\in[0,\infty),\omega\in\Omega\},\quad t\wedge T:=\min\{t,T\},\tag{4.1}$$

$$\mathcal{H}_T:=\{X(\omega)=\varphi(\omega(t_1\wedge T),\dots,\omega(t_m\wedge T)),\forall\, m\geqslant 1,\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^m)\}.\tag{4.2}$$

显然有 $\mathrm{Lip}(\Omega_t)\subseteq\mathrm{Lip}(\Omega_T)$, $t\leqslant T$. 记

$$\mathcal{H}=\mathrm{Lip}(\Omega):=\bigcup_{n=1}^{\infty}\mathrm{Lip}(\Omega_n).$$

我们将用 $[0,\infty)$ 上的连续函数全体 Ω 作为 Brown 运动的轨道空间, 记 $B_t(\omega)=\omega(t)$, $t\in[0,\infty)$, $\omega\in\Omega$. 由于 G-Brown 运动 (B_t) 是增量独立的, 且每一个增量都是 G- 正态分布, 从而, 我们只需按照下面的方法来引进一个定义在 $(\Omega,\mathcal{H})$ 上的次线性期望 $\hat{\mathbb{E}}$, 使得典则过程 $B_t(\omega)$ 关于 $\hat{\mathbb{E}}$ 满足 G-Brown 运动的定义 26, 其中 G 是一个给定的次线性函数,

$$G(a)=\frac{1}{2}(\underline{\sigma}^2a^+-\overline{\sigma}^2a^-),\quad a\in\mathbb{R}.$$

设 $\{\xi_i\}_{i=1}^{\infty}$ 为一个定义在某个次线性期望空间 $(\widetilde{\Omega},\widetilde{\mathcal{H}},\widetilde{\mathbb{E}})$ 的 i.i.d. 的 G- 正态分布随机变量序列. 从而,

$$\xi_i\overset{d}{=}N(\{0\}\times[\underline{\sigma}^2,\overline{\sigma}^2]),\tag{4.3}$$

且对于每一个 $i=1,2,\ldots$, ξ_{i+1} 都独立于 $(\xi_1,\ldots,\xi_i)$. 对于每一个下面形式的随机变量:

$$X=\varphi(B_{t_1}-B_{t_0},B_{t_2}-B_{t_1},\ldots,B_{t_m}-B_{t_{m-1}})\in\mathcal{H},$$
$$\varphi\in C_{\mathrm{Lip}}(\mathbb{R}^m),\quad 0=t_0<t_1<\cdots<t_m<\infty,$$

我们通过

$$\hat{\mathbb{E}}[X]=\widetilde{\mathbb{E}}[\varphi(\sqrt{t_1-t_0}\xi_1,\ldots,\sqrt{t_m-t_{m-1}}\xi_m)]$$

来定义随机变量 X 的期望值 $\hat{\mathbb{E}}[X]$, 而对于任何的 t_k, 通过

$$\hat{\mathbb{E}}_{t_k}[X]=\Phi(B_{t_1},\ldots,B_{t_k}-B_{t_{k-1}}),$$

其中

$$\Phi(x_1,\ldots,x_k)=\widetilde{\mathbb{E}}[\varphi(x_1,\ldots,x_k,\sqrt{t_{k+1}-t_k}\xi_{k+1},\ldots,\sqrt{t_m-t_{m-1}}\xi_m)]$$

来定义 X 的条件期望值 $\hat{\mathbb{E}}_{t_k}[X]$ (由此也定义了关于所有时间 t 的条件期望). 容易验证 $\hat{\mathbb{E}}:\mathcal{H}\mapsto\mathbb{R}$ 定义了随机变量空间 $(\Omega,\mathcal{H})$ 上的一个次线性期望, 并且 $(B_t)_{t\geqslant0}$ 是在期望 $\hat{\mathbb{E}}$ 下的对称 Brown运动 (定义 26), 且可以验证 $G(a)=\frac{1}{2}\hat{\mathbb{E}}[aB_1^2]$. 从而, (B_t) 就是对应于给定次线性函数 G 的次线性期望 $\hat{\mathbb{E}}$ 下的 Brown 运动—G-Brown 运动. 注意到以上的 t_k 其实是可以设在 $[0,\infty)$ 上的任意点, 这样我们还定义出了已知 Ω_t 的条件期望 $\hat{\mathbb{E}}_t:\mathcal{H}\mapsto\mathcal{H}_t$, $t\geqslant0$, 满足

(1) 若 $X\geqslant Y$, 则 $\hat{\mathbb{E}}_t[X]\geqslant\hat{\mathbb{E}}_t[Y]$;

(2) $\hat{\mathbb{E}}_t[\eta]=\eta$, 对于任意的 $t\in[0,\infty)$, $\eta\in\mathrm{Lip}(\Omega_t)$;

(3) $\hat{\mathbb{E}}_t[X]+\hat{\mathbb{E}}_t[Y]\leqslant\hat{\mathbb{E}}_t[X+Y]$;

(4) 对于每一个 $\eta\in\mathrm{Lip}(\Omega_t)$, 有 $\hat{\mathbb{E}}_t[\eta X]=\eta^+\hat{\mathbb{E}}_t[X]+\eta^-\hat{\mathbb{E}}_t[-X]$, 我们还有

$$\hat{\mathbb{E}}_t[\hat{\mathbb{E}}_s[X]]=\hat{\mathbb{E}}_{t\wedge s}[X],$$

特别地, 有

$$\hat{\mathbb{E}}[\hat{\mathbb{E}}_t[X]]=\hat{\mathbb{E}}[X].$$

例 28 作为例子, 上述方法可用来定义线性情形下的经典的标准 Brown 运动, 即 Wiener 过程. 此时我们只需对于所涉及的 i.i.d 的 G- 正态随机变量列 $\{\xi_i\}_{i=1}^{\infty}$ 需要满足的 (4.3) 中的上下方差参量设定为 $\underline{\sigma}=\overline{\sigma}=1$, 从而 $B_1\overset{d}{=}\xi_i$ 成为了标准的正态分布 $N(0,1)$. 以上构造所引入的表示和计算都变得非常直观, 非常简便.

可能有的读者会诧异为什么这样简单的构造方法就获得了概率理论非常深刻的标准 Brown 运动的存在性定理, 甚至进而获得了作为其非平凡推广的非线性空间中的 Brown 运动. 事实上, 对应于经典意义下的 Brown 运动的构造尚未完成, 还需要一小步, 即下面要讲的的一个空间完备化的步骤, 但是总的来讲, 整个构造还是变得更简单直观了. 这对于 Brown 运动的初学者提供了一个非常好的入门途径.

4.3 完备化空间中的非线性 Brown 运动

上面用直接的方法简洁地构造出了一个次线性期望空间 $(\Omega,\mathcal{H},\hat{\mathbb{E}})$ 中的 Brown 运动—G-Brown 运动. 但与经典 Brown 运动相比, 我们的非线性期望 $\hat{\mathbb{E}}$ 只是定义在了一个较小的随机变量空间 $\mathcal{H}$. 这对建立新的 Brown 运动的随机分析理论带来诸多不便. 解决此问题的简单而自然的方法是, 直接用我们已定义好的次线性期望 $\hat{\mathbb{E}}[\cdot]$ 来建立随机变量空间 $\mathcal{H}_T=\mathrm{Lip}(\Omega_T)$ 和 $\mathcal{H}=\mathrm{Lip}(\Omega)$ 上的范数, 并在此范数下对随机变量空间 $\mathcal{H}_T$ 和 $\mathcal{H}$ 进行完备化, 同时也将次线性期望 $\hat{\mathbb{E}}$ 连续地扩张到此完备化空间.

事实上, 此完备化也很简单: 对于每一个 $p\geqslant 1$ 和 $X\in\mathrm{Lip}(\Omega)$, 显然,

$$\|X\|_p:=\hat{\mathbb{E}}[|X|^p]^{\frac{1}{p}}$$

构成了随机变量空间 $\mathrm{Lip}(\Omega)$ 上的一个范数, 我们只需在此范数下将 $\mathrm{Lip}(\Omega)$ 和 $\mathrm{Lip}(\Omega_T)$ 扩张到其完备化的 Banach 空间, 记为 $L_G^p(\Omega)$ 和 $L_G^p(\Omega_T)$.

另外, 对每一个 $0\leqslant t\leqslant T<\infty$, 有

$$L_G^p(\Omega_t)\subseteq L_G^p(\Omega_T)\subset L_G^p(\Omega).$$

容易验证 $\hat{\mathbb{E}}[\cdot]$ 仍然是 $(\Omega,L_G^p(\Omega))$ 的次线性期望. 对于每一个 $t\geqslant 0$, $\hat{\mathbb{E}}_t[\cdot]$ 也同样可扩张为连续映射 $\hat{\mathbb{E}}_t[\cdot]:L_G^1(\Omega)\mapsto L_G^1(\Omega_t)$. 它仍满足上一小节的条件 (1)–(4).

事实上, 引入 $L_G^p(\Omega)$ 的方法有两种, 第一种就是前面介绍的首先通过 PDE 方法来获得非线性 Brown 运动的有限维分布, 这种方法被 Peng [108] 用于定义比非线性 Brown 运动更一般的非线性 Markov 过程. 接着, Peng [109, 113] 系统地使用此方法来定义非线性的 G-Brown 运动, 并由此获得了相应推广了的 Itô 积分和 Itô 公式. 据此思想, Ibragimov [70] 建立了无穷维非线性 Brown 运动理论; Hu 和 Peng [60] 研究了更一般的非线性独立平稳增量过程, 特别是包含了跳过程的非线性 Lévy 过程, 并巧妙地应用 Daniell-Stone 定理得到了没有小跳的 G-Lévy 过程分布的 Lévy-Khintchine 公式以及与之对应的偏微分积分方程 (关于这方面的研究可进一步参见文献 [90]). 第二种则是通过随机最优控制的方法或者鞅测度的方法来获得被那些被 G- 期望控制的概率测度族, 进而求其上确界. 其中鞅测度方法由 Denis 和 Martini [26] 所引入, 而随机控制方法由 Peng [104] 提出, 这两个方法的等价性则在文献 [25] 中揭示, 并对此理论进行了系统发展.

4.4 非线性 Brown 运动的拟必然分析理论

G- 期望理论在范数完备化的框架下显得非常完美, 但是还需要给出完备化后的随机变量空间 $L_G^p(\Omega)$ 中的随机变量 $X(\omega)$ 的性质. 非控条件下连续过程的拟必然分析首先由文献 [26] 引入, 在文献 [26, 104, 109] 的启发下, Denis 等 [25] 用随机控制的方法得到了 G- 期望的表示定理, 并首次实现了用相应的 G- 容度 (代替概率) 对空间 $L_G^p(\Omega)$ 中的随机变量的一个本质性刻画.

我们在连续路径空间 $\Omega=C_0([0,\infty),\mathbb{R}^d)$ 上引入一个自然的距离:

$$d(\omega_1,\omega_2):=\sum_{n=1}^{\infty}2^{-n}\Big(\max_{t\in[0,n]}|\omega_1(t)-\omega_2(t)|\wedge 1\Big),\quad \omega_1,\omega_2\in\Omega.$$

注意 (Ω,d) 是一个完备可分的距离空间. 记 $L^0(\Omega)$ 为定义在 Ω 上的 Borel 可测函数全体.

定理 29 (1) 存在一个定义在 $(\Omega,\mathcal{B}(\Omega))$ 的 σ- 可加的相对弱紧的概率测度族 $\mathcal{P}_G$, 使得

$$\hat{\mathbb{E}}[X]=\sup_{P\in\mathcal{P}_G}\mathrm{E}_P[X]=\sup_{P\in\mathcal{P}_G}\int_\Omega X(\omega)dP,\quad \forall\, X\in\mathrm{Lip}(\Omega).$$

获得此概率族后, 我们可以引入随机变量空间

$$\mathbb{L}_G^1(\Omega):=\Big\{X\in L^0(\Omega):\sup_{P\in\mathcal{P}_G}\mathrm{E}_P[|X|]<\infty\Big\}.$$

在其上可以将次线性期望 $\hat{\mathbb{E}}$ 的定义扩张到 $\mathbb{L}_G^1(\Omega)$:

$$\hat{\mathbb{E}}[X]:=\sup_{P\in\mathcal{P}_G}\mathrm{E}_P[X],\quad X\in\mathbb{L}_G^1(\Omega).$$

由此可引入对应于此概率族 $\mathcal{P}_G$ 的 Choquet 容度[19,27]:

$$\hat{c}(A)=\hat{\mathbb{E}}[\mathbf{1}_A]=\sup_{P\in\mathcal{P}_G}P(A),\quad A\in\mathcal{B}(\Omega).$$

(2) 设 $C_b(\Omega)$ 为定义于 Ω 上的所有有界的连续函数所构成的空间, 并记

$$\mathbb{L}^p(\Omega):=\Big\{X\in L^0(\Omega):\sup_{P\in\mathcal{P}_G}\mathrm{E}_P[|X|^p]<\infty\Big\},\quad p\geqslant 1.$$

我们有 $\mathbb{L}^p(\Omega)\supset L_G^p(\Omega)\supset C_b(\Omega)$. 对每一个 $X\in L_G^p(\Omega)$ 都存在一个在容度 $\hat{c}$ 下的拟必然连续的版本, 即存在 $Y\in L_G^p(\Omega)$ 使得 $X=Y$ q.s., 且对于任何 $\varepsilon>0$, 都存在开集 $O\subset\Omega$ 满足 $\hat{c}(O)<\varepsilon$ 使得 $Y\,|_{O^c}$ 是连续的.

上面 $L_G^p(\Omega)$ 中的一个关系 (如 $X(\omega)\leqslant Y(\omega)$) 的成立就是去掉 Ω 的一个容度为零的集合 A ($\hat{c}(A)=0$) 之后成立, 简称 $\hat{c}$-q.s. 成立.

以下是由 Denis 等[25] 获得的对于 $L_G^p(\Omega)$ 中随机变量的非常重要的一个刻画:

$$L_G^p(\Omega)=\Big\{X\in\mathbb{L}^p(\Omega):X\ \text{有一个}\ \hat{c}\text{- 拟连续的版本且}\ \lim_{n\to\infty}\hat{\mathbb{E}}[|X|^p\mathbf{1}_{\{|X|>n\}}]=0\Big\}.$$

著名的 Ergrov 定理告诉我们, 可测函数是 P- 拟必然连续的. 但是, 一个 Borel- 可测的函数在 G- 容度下可以不是拟连续的, 而非线性期望理论中的很多深刻的公开问题与此有紧密的联系.

此后, 文献 [59] 在此基础上应用非线性期望下的 Kolmogorov 方法得到了 G- 期望的表示定理, 这一方法更加简单直接; 通过随机控制的方法, Nutz [93] 引入了更一般的随机 G- 期望, 其中的生成函数 G 可以是随机的: $G=G(t,\omega,A)$; Nutz 和 Van Handel [95] 还将 G- 期望推广到了比 $\mathbb{L}^p(\Omega)$ 还要大的解析集上去. Ren [124] 得到了没有小跳的 G-Lévy 过程轨道的容度刻画定理和相应的表示定理.

需要注意的是, Ω 上的简单函数不全都有拟连续修正. 事实上, 在非线性期望框架下控制收敛定理不再自动成立, 从而产生一些极具挑战性的且有实际背景的研究问题. 为此, 文献 [64] 通过偏微分方程技术给出了 Ω 上的一大类拟连续的简单函数, 这说明 G- 期望空间具有充分多的元素和丰富的结构.

5 非线性 Brown 运动的 Itô 随机分析理论[115]

对称 Brown 运动是处处不可微、处处变差无界的, 能否对这样过程建立微积分理论曾是一个非常重大的数学问题. 1942 年, 在几乎孤立于所有的概率学家的情形下, Itô [71] 建立了随机微积分的理论, 被誉为 "随机王国中的 Newton 定律". 非线性 Brown 运动同样是处处不可微、处处变差无界的, 我们能否对其建立相应的微积分理论? 事实上, 文献 [109] 在构造非线性 Brown 运动之后, 就在同一文章中解决这一基础性的理论问题.

5.1 *G*-Brown 运动的 Itô 随机积分

应该指出, 如果不确定的概率全体都可以被一个参考概率测度所控制, 则没有必要重新去定义一个新的 Itô 随机积分, 因为在这个参考概率下定义的随机积分也适用于其他的那些概率测度, 因而我们可以在一个 Wiener 概率空间中通过倒向随机微分方程去定义各种不同的 g- 期望. 但是这种作法不适用于 G- 期望空间. 因为其中所涉及的那些不确定的概率是互相奇异的. 因而有必要重新研究和定义随机积分的概念. 我们知道在随机分析的研究中, 具有连续路径的 Brown 运动处于特别重要的中心位置.

对应于 G-Brown 运动 Itô 随机积分的定义与经典的随机积分很相近, 但是在 "$\hat{c}$- 拟必然", 或等价地, 在 L_G^2- 范数意义下, 以下随机积分的定义是由文献 [26, 109] 分别独立地获得的. 关于这一空间的进一步研究可参见文献 [64].

对于每一个 $T \geqslant 0$, 以及区间 $[0,T]$ 的一个有限子集合 $\Delta = \{t_1, \dots, t_N\}$, $0 = t_0 < t_1 < \cdots < t_N = T$. 对给定的 $p \geqslant 1$, 考虑以下形式的简单过程: 对于一个给定的分割 $\{t_0, \dots, t_N\} = \Delta$, 记

$$\eta_t(\omega) = \sum_{j=0}^{N-1} \xi_j(\omega) \mathbf{I}_{[t_j, t_{j+1})}(t),$$

其中 $\xi_i \in L_G^p(\Omega_{t_i})$ $(i = 0, 1, 2, \dots, N-1)$ 是给定的. 这类简单过程的集合记为 $M_G^{p,0}(0,T)$.

按照 Itô 积分定义的思想, 我们先对每一个简单过程 $\eta_\cdot \in M_G^{2,0}(0,T)$ 来定义 Itô 积分:

$$I(\eta) = \int_0^T \eta_s dB_s := \sum_{j=0}^{N-1} \xi_j (B_{t_{j+1}} - B_{t_j}).$$

容易验证 I: $M_G^{2,0}(0,T) \mapsto L_G^2(\Omega_T)$ 构成了一个连续的线性映射, 从而此映射可以连续地扩张到 I : $M_G^2(0,T) \mapsto L_G^2(\Omega_T)$. 而这个扩张后的映射 I 满足以下性质:

$$\hat{\mathbb{E}}[I] = 0 \quad \text{和} \quad \hat{\mathbb{E}}[I^2] \leqslant \overline{\sigma}^2 \int_0^T \hat{\mathbb{E}}[(\eta_t)^2] dt, \quad \eta_\cdot \in M_G^2(0,T).$$

因而, 对于每一个 $\eta_\cdot \in M_G^2(0,T)$, 可以定义其随机积分 $\int_0^T \eta_s dB_s =: I(\eta)$.

我们列出这个新的相对于 G-Brown 运动的 Itô 型积分重要的性质. 对于 $0 \leqslant s \leqslant t \leqslant T$, 记 $\int_s^t \eta_u dB_u := \int_0^T \mathbf{I}_{[s,t]}(u) \eta_u dB_u$.

定理 30 设 $\eta, \theta \in M_G^2(0,T)$ 以及 $0 \leqslant s \leqslant r \leqslant t \leqslant T$, 则有

(1) $\int_s^t \eta_u dB_u = \int_s^r \eta_u dB_u + \int_r^t \eta_u dB_u$;

(2) $\int_s^t (\alpha \eta_u + \theta_u) dB_u = \alpha \int_s^t \eta_u dB_u + \int_s^t \theta_u dB_u$, 若 $\alpha \in L_G^1(\Omega_s)$ 且有界;

(3) $\hat{\mathbb{E}}_t[X + \int_t^T \eta_u dB_u] = \hat{\mathbb{E}}_t[X]$, $\forall X \in L_G^1(\Omega)$.

5.2 G-Brown 运动平方变差过程 $\langle B\rangle$

G-Brown 运动的平方变差过程是非线性期望理论中的一个特别重要、并且特别有个性的随机过程. 虽然我们已经获得了关于这个过程的很多有趣的性质, 但是仍有许多有待于进一步深刻认识的问题. 而其定义其实很像经典情形: 设 π_t^N $(N=1,2,\ldots)$ 为一个区间 $[0,t]$ 的一个满足 $|\pi_t^N|\to 0$ 的分割的序列, 很容易证明

$$\langle B\rangle_t=\lim_{\mu(\pi_t^N)\to 0}\sum_{j=0}^{N-1}(B_{t_{j+1}^N}-B_{t_j^N})^2=B_t^2-2\int_0^t B_s dB_s.$$

$(\langle B\rangle_t)_{t\geqslant 0}$ 是一个增过程且 $\langle B\rangle_0=0$. 我们称 $\langle B\rangle$ 为 G-Brown 运动 (B_t) 的平方变差过程. 此过程深刻地刻画了 G-Brown 运动 (B_t) 的分布不确定性. 非常重要的一点是除了当 $\underline{\sigma}^2=\overline{\sigma}^2$ 的特殊情形以外, $\langle B\rangle$ 不会是一个确定性过程. 相反, $\langle B\rangle$ 的很多性质却很像 G-Brown 运动 (B_t) 自己: $\langle B\rangle$ 任意的增量 $\langle B\rangle_{t+s}-\langle B\rangle_t$ 是独立于 $\langle B\rangle$ 在时刻 t 以前的历史 $(\langle B\rangle_{t_1},\ldots,\langle B\rangle_{t_n},\, t_1,\ldots,t_n\in[0,t])$ 的, 而且也有增量分布的平稳性 $\langle B\rangle_{t+s}-\langle B\rangle_t\overset{d}{=}\langle B\rangle_s$. 另外, 由 $\hat{\mathbb{E}}[|\langle B\rangle_t|^3]\leqslant Ct^3$ 可以证明平方变差过程 $\langle B\rangle$ 也是个连续过程, 从而按照非线性 Brown 运动的定义, $\langle B\rangle$ 本身也构成了一个新的非常特别的 Brown 运动.

我们有以下推广的 Itô 同构公式:

$$\hat{\mathbb{E}}\bigg[\bigg(\int_0^T\eta_s dB_s\bigg)^2\bigg]=\hat{\mathbb{E}}\bigg[\int_0^T\eta^2(s)d\langle B\rangle_s\bigg],\quad \eta.\in M_G^2(0,T).$$

另外, $\langle B\rangle_t$ 的分布还可以通过 $\hat{\mathbb{E}}[\varphi(\langle B\rangle_t)]=\max_{v\in[\underline{\sigma}^2,\overline{\sigma}^2]}\varphi(vt)$ 来计算. 此外, 还可以证明, $\hat{c}$- 拟必然地 $\underline{\sigma}^2t\leqslant\langle B\rangle_{t+s}-\langle B\rangle_s\leqslant\overline{\sigma}^2t$. 特别地,

$$\hat{\mathbb{E}}[|\langle B\rangle_{s+t}-\langle B\rangle_s|^2]=\sup_{P\in\mathcal{P}_G}\mathrm{E}_P[|\langle B\rangle_{s+t}-\langle B\rangle_s|^2]=\max_{v\in[\underline{\sigma}^2,\overline{\sigma}^2]}|vt|^2=\overline{\sigma}^4t^2.$$

5.3 G- 期望空间中的 Itô 公式和随机微分方程

在此基础上, 文献 [115] 系统地建立了非线性随机分析理论. 下面的 Itô 公式首先由文献 [109] 获得, 并由文献 [42, 147] 进行了改进. 以下的结果的主要思想来自文献 [80], 它显著地改进了前面的结果. 注意到, 与经典的 Itô 过程非常重要的一个不同是, 在 $(\Omega,L_G^p(\Omega),\hat{\mathbb{E}})$ 中的一个完备的 Itô 过程是由下面的三种积分组成:

$$X_t^\nu=X_0^\nu+\int_0^t\alpha_s^\nu ds+\int_0^t\eta_s^\nu d\langle B\rangle_s+\int_0^t\beta_s^\nu dB_s.$$

Song [133] 告诉我们, 在很一般的假设下就会有 $X_t^\nu\equiv 0$ 等价于 $\alpha_t^\nu\equiv 0$, $\eta_t^\nu\equiv 0$, $\eta_t^\nu\equiv 0$.

定理 31　设 α^ν, $\eta^\nu\in M_G^1(0,T)$ 和 $\beta^\nu\in M_G^2(0,T)$, $\nu=1,\ldots,n$, 则对于每一个 $t\in[0,T]$ 和函数 $\Phi\in C^{1,2}([0,T]\times\mathbb{R}^n)$, 有

$$\begin{aligned}\Phi(t,X_t)-\Phi(s,X_s)=&\sum_{\nu=1}^n\int_s^t\partial_{x^\nu}\Phi(u,X_u)\beta_u^\nu dB_u+\int_s^t[\partial_u\Phi(u,X_u)+\partial_{x_u}\Phi(u,X_u)\alpha_u^\nu]du\\&+\int_s^t\bigg\{\sum_{\nu=1}^n\partial_{x^\nu}\Phi(u,X_u)\eta_u^\nu+\frac{1}{2}\sum_{\nu,\mu=1}^n\partial^2_{x^\mu x^\nu}\Phi(u,X_u)\beta_u^\mu\beta_u^\nu\bigg\}d\langle B\rangle_u.\end{aligned}$$

文献 [80] 的结果比上面的条件更弱: 过程 α^ν、β^ν 和 η^ν 可以在一个更大的空间 $M_\omega^2(0,T)$ 中.

考虑如下形式的 G-Brown 运动驱动的随机微分方程 (SDE):

$$X_t = X_0 + \int_0^t b(X_s)ds + \int_0^t h(X_s)d\langle B\rangle_s + \int_0^t \sigma(X_s)dB_s, \quad t \in [0,T], \tag{5.1}$$

其中初始条件 $X_0 \in \mathbb{R}^n$ 且 $b, h, \sigma : \mathbb{R}^n \mapsto \mathbb{R}^n$ 为给定的 Lipschitz 函数. 此处时间区间 $[0,T]$ 可以设为 $[0,\infty)$. 该方程的存在唯一性由文献 [111] 获得.

注 32 以上的结果实际上是在更深刻、更精细的层次上的 Itô 分析. 设 $\underline{\sigma} < 1 < \overline{\sigma}$, 则使得 (B_t) 为经典的标准 Brown 运动的 Wiener 概率测度 P 被 $\hat{\mathbb{E}}[\cdot]$ 所控制, 即 $\mathrm{E}_P[\cdot] \leqslant \hat{\mathbb{E}}[\cdot]$, 从而 E_P 是范数 $\|\cdot\|_{L_G^1}$ 下的连续泛函. 但此时我们有 $\langle B\rangle_t - t \equiv 0$, P-a.s., 即 Wiener- 测度无法分辨开 $\langle B\rangle_t$ 和 t, 这样 Itô 过程 (5.1) 中的前两个积分项就只能合并, 从而只剩下了 dt 和 dB_t 两个积分项. 另外, 在相应的 Wiener 空间 $(\Omega, \mathcal{F}, P)$ 中不存在一个等价的测度变换 Q 使得过程 (λB_t) 也是一个标准 Brown 运动, 而这在次线性空间 $(\Omega, L_G^1, \hat{\mathbb{E}})$ 中则是很容易的事情.

随后, Gao [42] 考虑了 G-SDE 解的同胚性质; Guo 等 [48] 提出了非线性期望框架下的鞅问题, 这与经典鞅问题有本质的区别, 原因在于线性情形下偏微分算子就是扩散过程的无穷小生成元, 而在非线性情形下不存在这种对应关系, 故而, 他们应用偏微分方程技术解决了这一问题, 建立了非线性随机微分方程的弱解理论; Geng 等 [46,123] 通过 Lyons 建立的 Rough-Path 理论发展了 G-SDE 的轨道分析理论; Lin [84] 引入了非线性期望框架下的关于增过程的随机积分, 得到了 G-Brown 运动驱动的反射随机微分方程解的适定性理论; 关于这方面的进一步研究可参见文献 [6,79,82,86] 等.

5.4 非线性期望下的 Brown 运动和鞅

我们可以在次线性或更一般的非线性期望空间中定义一个非对称的 G-Brown 运动. 设 $G(p,A) : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$ 为一个给定的形如 (3.9) 的关于 A 单调的次线性函数. Peng [115] (第 3.7 和 3.8 小节) 证明了, 存在取值于 $\mathbb{R}^{2d}$ 的 Brown 运动 $(B_t, b_t)_{t\geqslant 0}$ 使得 (B_1, b_1) 为 G- 分布, 其中的非线性期望空间为 $\Omega = C([0,\infty), \mathbb{R}^{2d})$, 并且 $(B_t(\omega), b_t(\omega))$ 为其典则过程, 而其完备化的随机变量空间就是 $(\Omega, L_G^1(\Omega))$. (B_t) 是对称的, 而 (b_t) 是非对称的 Brown 运动. 此时在次线性期望 $\hat{\mathbb{E}}$ 下, B_t 是正态分布而 b_t 则是最大分布. 对于每一个给定的非线性函数 $\tilde{G}(p,A) : \mathbb{R}^d \times \mathbb{S}(d) \mapsto \mathbb{R}$, 只要它满足以下的 G- 控制条件:

$$\tilde{G}(p,A) - \tilde{G}(p',A') \leqslant G(p-p', A-A'), \quad p, p \in \mathbb{R}, \quad A, A' \in \mathbb{S}(d),$$

我们都能在空间 $(\Omega, L_G^1(\Omega))$ 中构造出一个非线性期望 $\mathbb{E}_{\widetilde{G}}$ 使得

$$\mathbb{E}_{\widetilde{G}}[X] - \mathbb{E}_{\widetilde{G}}[Y] \leqslant \hat{\mathbb{E}}[X-Y], \quad X, Y \in L_G^1(\Omega),$$

且过程 $(B_t, b_t)_{t\geqslant 0}$ 就是取值于 $\mathbb{R}^{2d}$、关于非线性期望 $\mathbb{E}_{\widetilde{G}}$ 的 Brown 运动. 特别地, 我们有

$$\tilde{G}(p,A) = \mathbb{E}_{\widetilde{G}}\left[\langle b_1, p\rangle + \frac{1}{2}\langle AB_1, B_1\rangle\right], \quad p \in \mathbb{R}^d, \quad A \in \mathbb{S}(d).$$

我们实际上构造了一个非线性期望的变换, 它将同一个典则过程 (B,b) 从一个 G-Brown 运动变换到另一个 $\tilde{G}$-Brown 运动. 它是相应的概率理论中测度变换定理的重要的非平凡推广.

不仅如此, $\mathbb{E}_{\widetilde{G}}$ 还具有动态相容的条件期望 $\mathbb{E}_{\widetilde{G}}[\cdot \mid \Omega_t] : L_G^p(\Omega) \mapsto L_G^p(\Omega_t)$, 并且其条件期望依然也同样是受控的:

$$\mathbb{E}_{\widetilde{G}}[X \mid \Omega_t] - \mathbb{E}_{\widetilde{G}}[Y \mid \Omega_t] \leqslant \hat{\mathbb{E}}[X - Y \mid \Omega_t].$$

此外, $\widetilde{\mathbb{E}}_t[\cdot]$ 还满足以下性质:

(1) $\mathbb{E}_{\widetilde{G}}[X \mid \Omega_t] \geqslant \mathbb{E}_{\widetilde{G}}[Y \mid \Omega_t]$, 如果 $X \geqslant Y$;

(2) $\mathbb{E}_{\widetilde{G}}[X+\eta \mid \Omega_t] = \mathbb{E}_{\widetilde{G}}[X \mid \Omega_t] + \eta$, 若 $\eta. \in L_G^p(\Omega_t)$;

(3) $\mathbb{E}_{\widetilde{G}}[\mathbb{E}_{\widetilde{G}}[X \mid \Omega_s] \mid \Omega_t] = \mathbb{E}_{\widetilde{G}}[X \mid \Omega_{s\wedge t}]$, 特别地, $\mathbb{E}_{\widetilde{G}}[\mathbb{E}_{\widetilde{G}}[X \mid \Omega_s]] = \mathbb{E}_{\widetilde{G}}[X]$;

(4) 若 $X \in L_G^p(\Omega^t)$, 则 $\mathbb{E}_{\widetilde{G}}[X \mid \Omega_s] = \mathbb{E}_{\widetilde{G}}[X]$.

特别地, 次线性期望 $\hat{\mathbb{E}}$ 下的条件期望 $\hat{\mathbb{E}}_t : L_G^p(\Omega) \mapsto L_G^p(\Omega_t)$ 仍然是次线性的:

(5) $\hat{\mathbb{E}}[X \mid \Omega_t] - \hat{\mathbb{E}}[Y \mid \Omega_t] \leqslant \hat{\mathbb{E}}[X-Y \mid \Omega_t]$;

(6) $\hat{\mathbb{E}}[\eta X \mid \Omega_t] = \eta^+\hat{\mathbb{E}}[X \mid \Omega_t] + \eta^-\hat{\mathbb{E}}[-X \mid \Omega_t]$, 其中 $\eta \in L_G^1(\Omega_t)$ 且有界.

我们称一个过程 $(M_t)_{t\geqslant 0}$ 为一个 $\tilde{G}$- 鞅 ($\tilde{G}$- 上鞅; $\tilde{G}$- 下鞅), 如对于每一个 $t \in [0,\infty)$, $M_t \in L_G^1(\Omega_t)$ 且对于每一个 $s \in [0,t]$, 都有

$$\mathbb{E}_{\widetilde{G}}[M_t \mid \Omega_s] = M_s \quad (\leqslant M_s; \geqslant M_s).$$

显然, 对于每一个随机变量 $X \in L_G^1(\Omega)$, $M_t := \mathbb{E}_{\widetilde{G}}[X \mid \Omega_t]$ 是一个 $\tilde{G}$- 鞅. 特别地, 如果 φ 是一个 $\mathbb{R}^d$ 上的有界的 Lipschitz 连续函数, 对于 $X = \varphi(b_T + B_T)$, 随机过程

$$M_t = \widetilde{\mathbb{E}}[X \mid \Omega_t] = u(t, b_t + B_t)$$

是一个 $\tilde{G}$- 鞅, 其中 u 满足以下全非线性 PDE [4,7]:

$$\partial_t u + \tilde{G}(D_x u, D_{xx}^2 u) = 0, \quad t \in (0,T), \quad x \in \mathbb{R}^d,$$

终端条件为 $u\mid_{t=T} = \varphi$ 的唯一的黏性解. 一般而言, $\tilde{G}$- 鞅实际上就是完全非线性上述 PDE 的一个路径解. 反之, 一个 PDE 的黏性解则是一个状态依赖的 $\tilde{G}$- 鞅.

5.5 *G*-Sobolev 范数下的偏微分方程

事实上, 以上的 Brown 运动所对应的次线性期望框架为 PDE 理论提供了一个非常新且非常重要的 Sobolev 范数:

$$\|\varphi\|_{L_G^p} := \hat{\mathbb{E}}[|\varphi(B_T)|^p]^{1/p},$$

即构成了定义在 $\mathbb{R}^d$ 上的实函数 φ 非平凡的范数. 一般地, 我们可以用一个定义在 $C_b(\mathbb{R}^n)$ 上单调的非线性 (也称为 Nisio 半群 [91,92], 或非线性 Markov 半群) $Q_t(\cdot)$ 来定义一个范数:

$$\|\varphi\|_Q = (P_t(|\varphi|^p))^{1/p}.$$

如果 $G = G(p,A)$ 是一个关于 (p,A) 次线性的且关于 A 单调的函数, 则由如下形式的非线性抛物型的 PDE:

$$\partial_t u^\varphi(t,x) - G(Du^\varphi(t,x), D^2u^\varphi(t,x)) = 0, \quad u^\varphi(0,x) = \varphi(x)$$

的解 (如黏性解或经典) 构成的映射 $u^\varphi(1,0) : C_{b,\mathrm{Lip}}(\mathbb{R}^n) \to \mathbb{R}$ 会构成一个次线性的单调泛函 $F_G(\varphi) := u^\varphi(1,0)$, 从而构成了一个范数

$$\|\varphi\|_{W_G^{0,2}} := F_G(|\varphi|^2)^{1/2},$$

并且可以进一步构成相应的 Sobolev 范数:

$$\|\varphi\|_{W_G^{1,2}} := \|\varphi\|_{W_G^{0,2}} + \sum_{i=1}^d \|\partial_{x_i}\varphi\|_{W_G^{0,2}},$$

以及更一般地定义

$$\|\varphi\|_{W_G^{k+1,2}} := \|\varphi\|_{W_G^{0,2}} + \sum_{i=1}^{d} \|\partial_{x_i}\varphi\|_{W_G^{k,2}}, \quad k = 1, 2, \ldots$$

稍微将这个方法扩大一下, 就可以用来相应地解决路径型的 PDE. 也会使我们更加看清一件事情: 整个的非线性的 G- 鞅分析实际上也可以等价地解释为全非线性偏微分方程的 G-Sobolev 范数下的路径型的解, 关于这方面的研究可参见文献 [119, 120, 135].

5.6 G- 鞅和 G-BSDE

利用文献 [111] 提出的方法, 容易证明, 对给定的 $Z \in M_G^2(0,T)$ 和 $p, q \in M_G^1(0,T)$, 过程

$$Y_t = Y_0 + \int_0^t Z_s dB_s + \int_0^t p_s db_s + \int_0^t q_s d\langle B\rangle_s - \int_0^t \tilde{G}(p_s, q_s) ds, \quad t \in [0,T] \tag{5.2}$$

定义了一个非线性的 $\tilde{G}$- 鞅. 而此问题的逆问题则是, 求找一个适当的子空间 $\mathcal{M} \in L_G^1(\Omega_T)$, 使得对于其中的每一给定的随机变量 $X \in \mathcal{M}$, 相应的非线性鞅 $Y_t := \widetilde{\mathbb{E}}_t[X]$ 都有表示 (5.2). 事实上, 这等价于四元过程 $(Y, Z, p, q) \in M_G^2(0,T)$ 满足以下形式的新的倒向随机微分方程 (BSDE)[6]:

$$-dY_t = \tilde{G}(p_t, q_t)dt - Z_t dB_t - p_t db_t - q_t d\langle B\rangle_t, \quad Y_T = X. \tag{5.3}$$

这个问题是文献 [111, 115] 对形如 $G = G(A)$ 的次线性函数 ($b_t \equiv 0$) 的情形提出的. 在此情形下可以证明

$$Y_t = Y_0 + \int_0^t Z_s dB_s + K_t,$$

其中 K 为一个连续的递减的 G- 鞅且 $K_0 = 0$. 这个重要结果首先是由 Soner 等[127] 在条件

$$\hat{\mathbb{E}}\Big[\sup_{t\in[0,T]} \mathbb{E}_t[|X|^2]\Big] < \infty$$

下获得. 之后, 文献 [127, 131] 各自独立地将范数条件减弱到了 $\mathbb{E}[|X|^{2+\delta}] < \infty$. 事实上, 文献 [131] 将条件减弱到了 $\mathbb{E}[|X|^{\beta}] < \infty$, 其中的常数 β 只需满足 $\beta > 1$. 由于此分解的唯一性已获得证明, 从而问题就集中到了证明

$$K_T = \int_0^T G(q_s)ds - \int_0^T q_s d\langle B\rangle_s.$$

为此, 文献 [66] 引入了一个关于过程 q 的先验估计, 从而获得关于 q 的一个唯一性结果. 另一方面, 文献 [132] 巧妙地引入了一个范数, 成功地处理了非对称 G- 鞅的唯一性问题. 特别重要地, 文献 [132] 成功地解决了一个重要的公开问题: 将 G-Itô 过程中的三个积分项 dt、dB_t 和 $d\langle B\rangle_t$ 清楚地区分开来. 最近, 文献 [121] 结合其所引入的范数以及偏微分方程的技术, 基本上证明了关于 G- 鞅表示定理的猜想, 也可参见文献 [128].

另外, 对于一个给定的 $X \in L_G^1(\Omega_T)$, 一个终端条件为 $Y_T = X$ 的 $\tilde{G}$- 鞅 Y_t 显然就是 $Y_t = \widetilde{\mathbb{E}}_t[X]$. 这实际上就是完全非线性的倒向随机微分方程的一个典型结果. 而对于完全非线性的倒向随机微分方程组, 则可利用文献 [107] 引入的控制方法 (domination approach), 用 Picard 迭代 (或压缩不动点原理) 来证明以下类型的多维全非线性 BSDE:

$$Y_t^i = \widetilde{\mathbb{E}}_t^i\Big[X^i + \int_t^T f^i(s, Y_s)ds\Big], \quad i = 1, \ldots, m, \quad Y = (Y^1, \ldots, Y^m) \tag{5.4}$$

[6] 经典 BSDE 理论由文献 [99] 所建立.

的解 Y 的存在唯一性. 与 $\tilde{G}$- 期望类似, 对于每一个 $i=1,\dots,m$, $\widetilde{\mathbb{E}}^i$ 是一个 $\tilde{G}_i$- 期望且 $\tilde{G}_i$ 是一个定义在 $\mathbb{R}^d\times\mathbb{S}(d)$ 上的被 G 控制的函数. 可以证明, 如果对于每一个 i, $f^i(\cdot,y)\in M_G^1(0,T)$, $y\in\mathbb{R}^d$, 且是关于 y 的 Lipschitz 连续函数, 那么对于每一个给定的终端条件 $X=(X^1,\dots,X^m)\in L_G^1(\Omega_T,\mathbb{R}^m)$, BSDE (5.4) 存在唯一的解 $Y\in M_G^1(0,T,\mathbb{R}^m)$. 在不同框架下, 文献 [18] 应用 PDE 的方法也独立地获得了一类被称为 2BSDE 的完全非线性 BSDE 的解的存在唯一性.

6 非线性期望下的随机分析的进展

非线性期望提供了一个研究 Knight 不确定性问题的统一的理论体系. 许多专家学者进一步研究了非线性随机分析中的相关问题. 很多经典概率理论体系中原来所难以解决的概率和分布的不确定性的问题, 往往都可代之以基于非线性期望的运算, 使问题迎刃而解.

基于非线性随机微分方程理论的发展, Gao 和 Jiang [44] 研究了 G- 随机微分方程的大偏差问题, 而 Gao 和 Xu [45] 给出了次线性期望框架下相对熵的概念, 从而建立了次线性期望下的独立随机变量列经验测度的大偏差原理. 随后, Gao [43] 和 Osuka [97] 建立了 G-Brown 运动泛函的变分表示, 给出了 G-Brown 运动泛函的大偏差原理. 这些工作为进一步研究非线性期望下极限定理提供了有力的工具.

众所周知, Brown 运动的 Lévy 鞅刻画定理对研究鞅问题和弱解理论非常重要, 因此, 一个很有趣的问题是如何研究非线性 Brown 运动的 Lévy 鞅刻画. Xu 和 Zhang [144,145] 通过非线性对称鞅的随机计算方法证明了 G-Brown 运动的 Lévy 鞅刻画定理. 在此基础上, Lin [83] 研究了在非线性期望空间中对称鞅的随机积分表示. 此后, Song [133] 系统地研究了非线性独立平稳增量过程的基本性质, 给出了一个不同形式 G-Brown 运动鞅刻画准则, 建立了一般非线性 Brown 运动的表示定理.

接下来一个重要问题是建立 G-Brown 运动驱动的倒向随机微分方程 (G-BSDE) 理论. 经典 Brown 运动下的 BSDE 已经获得很大的进展并对非线性期望的发展产生了决定性的影响 (参见文献 [12, 41, 101–106]). 在 G- 期望理论中, 由于非对称鞅的存在, 使得这一问题比经典 BSDE 更加复杂. 此时 G-BSDE 的解为三元组 (Y,Z,K) 使得

$$Y_t=\xi+\int_t^T g(s,Y_s,Z_s)ds+\int_t^T f(s,Y_s,Z_s)d\langle B\rangle_t-\int_t^T Z_s dB_s-(K_T-K_t),$$

其中 K 为一个连续的减 G- 鞅且 $K_0=0$[7]. 需要注意的是, 经典的 Picard 迭代方法不再适用于处理非线性期望框架下的倒向随机微分方程. 为此, 文献 [52] 通过偏微分方程方法和 Galerkin 逼近技术建立了 G-BSDE 的适定性理论. 在此基础上, 文献 [53] 获得了 G-BSDE 的比较定理以及 Markov 情形下的全非线性 Feynman-Kac 公式[8], 并建立了非线性 Brown 运动的 Girsanov 变换[9]. 关于平方增长系数下的 GBSDE 的研究参见文献 [65]. 此后, 文献 [120] 提出了 G-Sobolev 空间理论, 通过 G-BSDE 得到了全非线轨道偏微分方程 Sobolev 解的存在唯一性, 提供了新的思路来研究倒向随机微分方程理论和偏微分方程理论. 注意到, 关于拟线性路径偏微分方程的光滑解的研究结果是由文献 [122] 通过 BSDE 方法获得. 对此类方程的黏性解的研究参见文献 [32].

基于非线性框架下的倒向随机微分方程理论, 文献 [57, 63] 研究了非线性期望框架下的遍历倒向随机微分方程理论; Hu 和 Ji [55] 研究了 G- 期望框架下的随机递归最优控制问题, 他们引入了一种全新的 "隐式分解" 方法, 证明了动态规划原理. 最近, 文献 [54] 巧妙地利用线性化方法和弱收敛技术得

[7] 对于这一方程的研究也可参见文献 [129], 但在他们的框架中方程的解不一定在 G- 期望空间中.

[8] 非线性 Feynman-Kac 公式由文献 [101, 102] 建立

[9] 文献 [98, 143] 分别研究了一维情形和多维情形下非线性 Girsanov 变换.

到了非线性期望框架下的随机最大值原理, 从而实质地推广了 Peng [100] 的一般随机最大值原理. 关于这一方面的研究, 也可参见文献 [10,56,62,146] 等.

目前, 非线性期望理论仍处于发展中, 仍有许多重要的问题有待研究和解决. 例如, 在非线性期望框架下的停时理论就有很大的挑战性, 这是由于停止过程不一定具有拟连续性. Song [130] 证明了一类 G- 鞅的停止过程仍然为拟连续过程且为 G- 鞅; 结合这个技术, 文献 [134] 利用耦合方法、全非线性 Feynman-Kac 公式和 Girsanov 变换公式首次给出了全非线性偏微分方程解的梯度估计. 文献 [61] 研究了条件 G- 期望的推广, 并且得到了关于一类停时的 G- 鞅停时定理. 另一方面, 文献 [80] 在一扩张的空间中建立了非线性随机分析的局部化理论, 但怎样在这一空间中定义条件 G- 期望的问题仍有很大的挑战性, 关于这一方向的研究也可参见文献 [61,79,96] 中的非常系统的研究.

此外, 当前利用非线性 G- 期望理论来研究 Knight 不确定性资产定价和风险管理理论是经济金融界的热点问题, 例如, Beissner 和 Riedel [9] 发现在非线性期望框架下 Arrow-Debreu 均衡不能通过连续交易永久证券来实现成为 Radner 均衡; Epstein 和 Ji [37,38] 研究了带有均值 - 波动率不确定性的递归效用问题, 解决了一类波动率不确定情形下的资产定价与风险度量问题, 关于这方面的研究也可参见文献 [137] 等. 由于考虑了模型不确定性的实质性的影响, 这些研究成果给出了更加完备的资产定价理论.

非线性的 G- 期望理论已经获得了很多重要的进展, 可参见文献 [28,33,34,51,81]. 另外, G- 期望理论的提出和发展与我们 1997 年提出的 g- 期望理论 [103] 并从中获得了关键性的启发是分不开的 (参见文献 [22,35,72,73,104,106]).

7 结论

本文给出非线性期望理论进展的一个综述, 首先介绍如何建立一个非线性期望的基本理论框架, 说明了为什么这个框架可以自然和稳健地用来分析和计算我们现实世界中经常面对的概率和分布本身的不确定性. 我们介绍了非线性 i.i.d. 的概念, 并通过例子说明为什么非线性 i.i.d. 假设对于现实世界的样本数据有非常好的普适性.

本文介绍了次线性期望空间中两个最典型的新的统计分布—非线性正态分布和最大分布与作为这个理论的基础的大数定律和中心极限定理. 其特别典型的应用则是关于现实样本数据的非线性期望的 φ-max-mean 计算方法. 我们还介绍一个最重要的随机过程—非线性 Brown 运动及其相关的随机分析, 以及相应的非线性鞅理论. 关于这个新领域还有大量的有趣而富有挑战性的公开问题, 我们会在随后的文章中讨论. 我们也特别注意到, 最近 E 等 [31] 应用倒向随机微分方程 (BSDE) 理论获得了计算非常高维数的拟线性抛物型偏微分方程组的算法. 相信怎样通过 G-Brown 运动驱动的 BSDE 来计算全非线性偏微分方程的问题将会引起很多学者的关注.

非线性期望理论非平凡地推广了 Kolmogorov 于 1933 年建立的概率论公理体系, 那里的最核心的概念是概率测度 P. 我们的理论的关键不同是, 最核心的概念是 (非线性) 期望 $\hat{\mathbb{E}}$, 而不是很多人所预期的非线性 (非可加) 概率. 期望为线性的, 而正是这种内蕴形式的非线性[10)] 为我们提供了一个高倍的望远镜和显微镜, 使我们能够对于现实世界中很难发现而又无处不在的概率和统计分布的不确定性进行的定量分析和计算.

我们以一个经典的概率理论框架与作为其推广的非线性期望理论框架的对照表 1 来结束本文.

10) 参见文献 [109] 和 [46] 的介绍.

表 1　非线性期望框架与经典概率理论框架的对照表

经典概率空间	非线性期望空间
$(\Omega, \mathcal{F}, P)$	非线性 $(\Omega, \mathcal{H}, \mathbb{E})$: (sublinear is basic)
概率同分布 $X \overset{d}{=} Y$	非线性期望同分布 $X \overset{d}{=} Y$
随机变量 Y 独立于随机变量 X	Y 非线性地独立于 X (非对称独立)
大数定律和中心极限定理	非线性大数定律和中心极限定理
正态分布	非线性正态分布
Brown 运动 $B_t(\omega) = \omega_t$	非线性 Brown 运动 $B_t(\omega) = \omega_t$
平方变差过程 $\langle B \rangle_t = t$	$\langle B \rangle_t$: 仍然为非线性 Brown 运动
Lévy 过程	非线性 Lévy 过程
Brown 运动的 Itô 随机分析	非线性 Brown 运动对应于更精的随机分析
SDE $dx_t = b(x_t)dt + \sigma(x_t)dB_t$	$dx_t = b(x_t)dt + \sigma(x_t)dB_t + \beta(x_t)d\langle B \rangle_t$
扩散: $\partial_t u - \mathcal{L}u = 0$	非线性 $\partial_t u - G(t, x, u, Du, D^2 u) = 0$
Markov 过程和 Markov 半群	非线性 Markov 过程和非线性 Markov 半群
鞅和半鞅	非线性鞅和半鞅
$\mathrm{E}[X \mid \mathcal{F}_t] = \mathrm{E}[X] + \int_0^T z_s dB_s$	$\mathbb{E}[X \mid \mathcal{F}_t] = \mathbb{E}[X] + \int_0^t z_s dB_s + K_t$ $K_t = \int_0^t \eta_s d\langle B \rangle_s - \int_0^t 2G(\eta_s) ds$
样本数据的统计平均算法和 Monté-Carlo 算法	φ-max-mean 算法和非线性 Monté-Carlo 算法

致谢　感谢陈增敬、胡明尚、稽少林、李海梁、王法磊、王汉超、吴臻和杨淑振等及审稿人在本文形成过程中提出的宝贵建议和意见. 特别由衷地感谢我国数学界对我在非线性期望理论研究方面研究所给予的长期强力的支持. 仅举一例: 2003 年, 严加安和马志明专门在五所大楼组织了为期两天的研讨班, 让作者详细地报告了非线性期望理论的一个研究进展.

参考文献

1 Allais M. Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. Econometrica, 1953, 21: 503–546

2 Anderson E W, Hansen L P, Sargent T J. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. J Eur Econ Assoc, 2003, 1: 68–123

3 Artzner P, Delbaen F, Eber J, et al. Coherent Measures of Risk. Math Finance, 1999, 9: 203–228

4 Atlan M. Localizing volatilities. ArXiv:0604316, 2006

5 Avellaneda M, Levy A, Paras A. Pricing and hedging derivative securities in markets with uncertain volatilities. Appl Math Finance, 1995, 2: 73–88

6 Bai X, Lin Y. On the existence and uniqueness of solutions to stochastic differential equations driven by G-Brownian motion with integral-Lipschitz coefficients. Acta Math Appl Sin Engl Ser, 2014, 30: 589–610

7 Barenblatt G I. Similarity, Self-Similarity and Intermediate Asymptotics. New York: Consultants Bureau, 1979

8 Bayraktar E, Munk A. An α-Stable limit theorem under sublinear expectation. Bernoulli, 2014, 22: 2548–2578

9 Beissner P, Riedel F. Non-implementability of arrow-debreu equilibria by continuous trading under Knightian uncertainty. ArXiv:1409.6940, 2014

10 Biagini F, Meyer-Brandis T, Øksendal B, et al. Optimal control with delayed information flow of systems driven by G-Brownian motion. ArXiv:1402.3139, 2014

11 Bion-Nadal J, Kervarec M. Risk measuring under model uncertainty. Ann Appl Probab, 2012, 22: 213–238

12 Bismut J M. Conjugate convex functions in optimal stochastic control. J Math Anal Appl, 1973, 44: 384–404

13 Cabre X, Caffarelli L A. Fully Nonlinear Elliptic Partial Differential Equations. Providence: Amer Math Soc, 1997

14 Cagetti M, Hansen L P, Sargent T J. Robustness and pricing with uncertain growth. Rev Financ Stud, 2002, 15: 363–404
15 Chen Z. Strong laws of large numbers for sub-linear expectations. Sci China Math, 2016, 59: 945–954
16 Chen Z, Epstein L. Ambiguity, risk and asset returns in continuous time. Econometrica, 2002, 70: 1403–1443
17 Chen Z J, Wu P Y, Li B M. A strong law of large numbers for non-additive probabilities. Internat J Approx Reason, 2013, 54: 365–377
18 Cheridito P, Soner H M, Touzi N, et al. Second order backward stochastic differential equations and fully non-linear parabolic PDEs. Comm Pure Appl Math, 2007, 60: 1081–1110
19 Choquet G. Theory of capacities. Ann Inst Fourier (Grenoble), 1953, 5: 131–295
20 Crandall M G, Ishii H, Lions P L. User's guide to viscosity solutions of second order partial differential equations. Bull Amer Math Soc (NS), 1992, 27: 1–67
21 Crandall M G, Lions P L. Condition d'unicité pour les solutions generalisées des équations de Hamilton-Jacobi du premier ordre. C R Acad Sci Paris Sér I Math, 1981, 292: 183–186
22 Cohen S N. Quasi-sure analysis, aggregation and dual representations of sublinear expectations in general spaces. Electron J Probab, 2012, 17: 1–15
23 Delbaen F. Coherent Risk Measures (Lectures given at the Cattedra Galileiana at the Scuola Normale di Pisa, March 2000). Pisa: the Scuola Normale di Pisa, 2002
24 Delbaen F, Peng S, Rosazza Gianin E. Representation of the penalty term of dynamic concave utilities. Finance Stoch, 2010, 14: 449–472
25 Denis L, Hu M, Peng S. Function spaces and capacity related to a sublinear expectation: Application to G-Brownian motion paths. Potential Anal, 2011, 34: 139–161
26 Denis L, Martini C. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. Ann Appl Probab, 2006, 16: 827–852
27 Denneberg D. Non-Additive Measure and Integral. New York: Springer, 1994
28 Dolinsky Y. Numerical schemes for G-expectations. Electron J Probab, 2012, 17: 1–15
29 Dolinsky Y, Nutz M, Soner H M. Weak approximation of G-expectations. Stochastic Process Appl, 2012, 122: 664–675
30 Donsker M. An invariance principle for certain probability limit theorems. In: Memoirs of the American Mathematical Society, vol. 6. Providence: Amer Math Soc, 1951
31 E W N, Hutzenthaler M, Jentzen A, et al. On full history recursive multilevel Picard approximations and numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. ArXiv:1607.03295, 2017
32 Ekren I, Keller C, Touzi N, et al. On viscosity solutions of path dependent PDEs. Ann Probab, 2014, 42: 204–236
33 Ekren I, Touzi N, Zhang J. Optimal stopping under nonlinear expectation. Stochastic Process Appl, 2014, 124: 3277–3311
34 Ekren I, Touzi N, Zhang J. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. Ann Probab, 2016, 44: 1212–1253
35 El Karoui N, Peng S, Quenez M C. A dynamic maximum principle for the optimization of recursive utilities under constraints. Ann Appl Probab, 2001, 11: 664–693
36 Ellsberg D. Risk, ambiguity, and the Savage axiom. Q J Econ, 1961, 7: 643–669
37 Epstein L, Ji S. Ambiguous volatility and asset pricing in continuous time. Rev Financ Stud, 2013, 26: 1740–1786
38 Epstein L, Ji S. Ambiguous volatility, possibility and utility in continuous time. J Math Econom, 2014, 50: 269–282
39 Föllmer H, Schied A. Statistic Finance: An Introduction in Discrete Time, 2nd ed. Berlin: Walter de Gruyter, 2004
40 Frittelli M, Rossaza Gianin E. Putting order in risk measures. J Bank Finance, 2002, 26: 1473–1486
41 Frittelli M, Rossaza Gianin E. Dynamic convex risk measures. In: New Risk Measures for the 21st Century. New York: John Wiley & Sons, 2004, 227–247
42 Gao F. Pathwise properties and homeomorphic flows for stochastic differential equations driven by G-Brownian motion. Stochastic Process Appl, 2009, 119: 3356–3382
43 Gao F. A variational representation and large deviations for functionals of G-Brownian motion. ArXiv:1204.4525, 2012
44 Gao F, Jiang H. Large deviations for stochastic differential equations driven by G-Brownian motion. Stochastic Process Appl, 2010, 120: 2212–2240
45 Gao F, Xu M. Relative entropy and large deviations under sublinear expectation. Acta Math Sci Ser B Engl Ed, 2012, 32: 1826–1834

46 Geng X, Qian Z, Yang D. G-Brownian motion as rough paths and differential equations driven by G-Brownian motion. In: Séminaire de Probabilités, XLVI. Lecture Notes in Mathematics, vol. 2123. Switzerland: Springer, 2014, 125–193
47 Gilboa I, Schmeidler D. Maxmin expected utility with non-unique prior. J Math Econ, 1989, 18: 141–153
48 Guo X, Pan C, Peng S. Martingale problem under nonlinear expectations. ArXiv:1211.2869, 2012
49 Hansen L P, Sargent T J. Robust control and model uncertainty. Amer Econ Rev, 2001, 91: 60–66
50 Hansen L P, Sargent T J. Prices of macroeconomic uncertainties with tenuous beliefs. Https://ssrn.com/abstract =2888511, 2016
51 Hu M. Explicit solutions of the G-heat equation for a class of initial conditions. Nonlinear Anal, 2012, 75: 6588–6595
52 Hu M, Ji S, Peng S, et al. Backward stochastic differential equations driven by G-Brownian motion. Stochastic Process Appl, 2014, 124: 759–784
53 Hu M, Ji S, Peng S, et al. Comparison theorem, Feynman-Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion. Stochastic Process Appl, 2014, 124: 1170–1195
54 Hu M, Ji S. Stochastic maximum principle for stochastic recursive optimal control problem under volatility ambiguity. SIAM J Control Optim, 2016, 54: 918–945
55 Hu M, Ji S. Dynamic programming principle for stochastic recursive optimal control problem under G-framework. Stochastic Process Appl, in press, 2016
56 Hu M, Ji S, Yang S. A stochastic recursive optimal control problem under the G-expectation framework. Appl Math Optim, 2014, 70: 253–278
57 Hu M, Li H, Wang F, et al. Invariant and ergodic nonlinear expectations for G-diffusion processes. Electron Commun Probab, 2015, 20: 1–15
58 Hu M, Li X. Independence under the G-expectation framework. J Theoret Probab, 2014, 27: 1011–1020
59 Hu M, Peng S. On representation theorem of G-expectations and paths of G-Brownian motion. Acta Math Appl Sin Engl Ser, 2009, 25: 539–546
60 Hu M, Peng S. G-Lévy processes under sublinear expectations. ArXiv:0911.3533, 2009
61 Hu M, Peng S. Extended conditional G-expectations and related stopping times. ArXiv:1309.3829, 2013
62 Hu M, Peng S, Song Y. Stein type characterization for G-normal distributions. ArXiv:1603.04611, 2014
63 Hu M, Wang F. Ergodic BSDEs driven by G-Brownian motion and their applications. ArXiv:1407.6210, 2014
64 Hu M, Wang F, Zheng G. Quasi-continuous random variables and processes under the G-expectation framework. Stochastic Process Appl, 2016, 126: 2367–2387
65 Hu Y, Lin Y, Abdoulaye S H. Quadratic backward stochastic differential equations driven by G-Brownian motion: Discrete solutions and approximation. ArXiv:1603.03637, 2016
66 Hu Y, Peng S. Some estimates for martingale representation under G-expectation. ArXiv:1004.1098, 2010
67 Hu Z, Zhou L. Multi-dimensional central limit theorems and laws of large numbers under sublinear expectations. Acta Math Sin Engl Ser, 2015, 31: 305–318
68 Huber P. Robust Statistics. New York: John Wiley & Sons, 1981
69 Huber P, Strassen V. Minimax tests and the Neyman-Pearson lemma for capacities. Ann Statist, 1973, 1: 251–263
70 Ibragimov A. G-expectations in infinite dimensional spaces and related PDEs. ArXiv:1306.5272, 2013
71 Itô K. Differential equations determining a Markov process (in Japanese). J Pan-Japan Math Colloq, 1942, 1077: 1352–1400
72 Ji S L, Zhou X Y. A generalized Neyman-Pearson lemma for g-probabilities. Probab Theory Related Fields, 2010, 148: 645–669
73 Jiang L. Convexity, translation invariance and subadditivity for g-expectations and related risk mea-sures. Ann Appl Probab, 2008, 18: 245–285
74 Jin H, Peng S. Optimal unbiased estimation for maximal distribution. ArXiv:1611.07994, 2016
75 Kolmogorov A N. Foundations of the Theory of Probability. Oxford: Chelsea Publishing, 1956
76 Krylov N V. Nonlinear Parabolic and Elliptic Equations of the Second Order. Dordrecht: Reidel Publishing, 1987
77 Song Y, Lin L. Sublinear expectation nonlinear regression for the financial risk measurement and management. Discrete Dyn Nat Soc, 2013, 2013: 398750
78 Li X. A central limit theorem for m-dependent random variables under sublinear expectations. Acta Math Appl Sin Engl Ser, 2015, 31: 435–444
79 Li X, Lin X, Lin Y. Lyapunov-type conditions and stochastic differential equations driven by G-Brownian motion. J Math Anal Appl, 2016, 439: 235–255
80 Li X, Peng S. Stopping times and related Itô calculus with G-Brownian motion. Stochastic Process Appl, 2009, 121: 1492–1508

81 Lin Q. Local time and Tanaka formula for the G-Brownian motion. J Math Anal Appl, 2013, 398: 315–334
82 Lin Q. Some properties of stochastic differential equations driven by the G-Brownian motion. Acta Math Sin Engl Ser, 2013, 29: 923–942
83 Lin Q. General martingale characterization of G-Brownian motion. Stochastic Anal Appl, 2013, 31: 1024–1048
84 Lin Y. Stochastic differential equations driven by G-Brownian motion with reflecting boundary. Electron J Probab, 2013, 18: 1–23
85 Lin Z, Zhang L. Convergence to a self-normalized G-Brownian motion. In: Probability, Uncertainty and Quantitative Risk. New York: Springer, 2017, 2–4
86 Luo P, Wang F. Stochastic differential equations driven by G-Brownian motion and ordinary differential equations. Stochastic Process Appl, 2014, 124: 3869–3885
87 Lyons T. Uncertain volatility and the risk free synthesis of derivatives. Appl Math Finance, 1995, 2: 7–133
88 Maccheroni F, Marinacci M. A strong law of large numbers for capacities. Ann Probab, 2005, 33: 1171–1178
89 Marinacci M. Limit laws for non-additive probabilities and their frequentist interpretation. J Econom Theory, 1999, 84: 145–195
90 Neufeld A, Nutz M. Nonlinear Lévy processes and their characteristics. Trans Amer Math Soc, 2017, 369: 69–95
91 Nisio M. On a nonlinear semigroup attached to optimal stochastic control. Publ Res Inst Math Sci, 1976, 13: 513–537
92 Nisio M. On stochastic optimal controls and envelope of Markovian semi-groups. In: Proceedings of the International Symposium on Stochastic Differential Equations. Kyoto: John Wiley & Sons, 1976, 297–325
93 Nutz M. Random G-expectations. Ann Appl Probab, 2013, 23: 1755–1777
94 Nutz M, Soner H M. Superhedging and dynamic risk measures under volatility uncertainty. SIAM J Control Optim, 2010, 50: 2065–2089
95 Nutz M, Van Handel R. Constructing sublinear expectations on path space. Stochastic Process Appl, 2013, 123: 3100–3121
96 Nutz M, Zhang J. Optimal stopping under adverse nonlinear expectation and related games. Ann Appl Probab, 2015, 25: 2503–2534
97 Osuka E. A variational representation for G-Brownian functionals. ArXiv:1204.4077, 2012
98 Osuka E. Girsanov's formula for G-Brownian motion. Stochastic Process Appl, 2013, 123: 1301–1318
99 Pardoux E, Peng S. Adapted solutions of a backward stochastic differential equation. Systems Control Lett, 1990, 14: 55–61
100 Peng S. A general stochastic maximum principle for optimal control problems. SIAM J Control Optim, 1990, 28: 966–979
101 Peng S. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. Stochastics, 1991, 37: 61–74
102 Peng S. A nonlinear Feynman-Kac formula and applications. In: Proceedings of Symposium of System Sciences and Control Theory. Singapore: World Scientific, 1992, 173–184
103 Peng S. Backward SDE and related g-expectations. In: Backward Stochastic Differential Equations. Pitman Research Notes in Mathematics Series, No. 364. Paris: El Karoui Mazliak Edit, 1997, 141–159
104 Peng S. Filtration consistent nonlinear expectations and evaluations of contingent claims. Acta Math Appl Sin Engl Ser, 2004, 20: 1–24
105 Peng S. Dynamical evaluations. C R Math Acad Sci Paris, 2004, 339: 585–589
106 Peng S. Nonlinear expectation, nonlinear evaluations and risk measurs. In: Stochastic Methods in Finance Lectures. New York: Springer-Verlag, 2004, 143–217
107 Peng S. Dynamically consistent nonlinear evaluations and expectations. ArXiv:0501415, 2005
108 Peng S. Nonlinear expectations and nonlinear Markov chains. Chin Ann Math Ser B, 2005, 26: 159–184
109 Peng S. G-expectation, G-Brownian motion and related stochastic calculus of Itô's type. ArXiv:0601035, 2006
110 Peng S. Law of large numbers and central limit theorem under nonlinear expectations. ArXiv:0702358, 2007
111 Peng S. Lecture notes: G-Brownian motion and dynamic risk measure under volatility uncertainty. ArXiv:0711.2834, 2007
112 Peng S. A new central limit theorem under sublinear expectations. ArXiv:0803.2656, 2008
113 Peng S. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. Stochastic Process Appl, 2008, 118: 2223–2253
114 Peng S. Survey on normal distributions, central limit theorem, Brownian motion and the related stochastic calculus under sublinear expectations. Sci China Ser A, 2009, 52: 1391–1411
115 Peng S. Nonlinear expectations and stochastic calculus under uncertainty. ArXiv:1002.4546, 2010

116 Peng S. Tightness, weak compactness of nonlinear expectations and application to CLT. ArXiv:1006.2541, 2010
117 Peng S. Backward stochastic differential equation, nonlinear expectation and their applications. In: Proceedings of the International Congress of Mathematicians. Singapore: World Scientific, 2010, 393–432
118 Peng S. G-Gaussian processes under sublinear expectations and q-Brownian motion in quantum mechanics. ArXiv:1105.1055, 2011
119 Peng S. Note on viscosity solution of path-dependent PDE and G-martingales. ArXiv:1106.1144, 2011
120 Peng S, Song Y. G-expectation weighted Sobolev spaces, backward SDE and path dependent PDE. J Math Soc Japan, 2015, 67: 1725–1757
121 Peng S, Song Y, Zhang J. A complete representation theorem for G-martingales. Stochastics, 2014, 86: 609–631
122 Peng S, Wang F. BSDE, path-dependent PDE and nonlinear Feynman-Kac formula. Sci China Math, 2016, 59: 1–18
123 Peng S, Zhang H. Stochastic calculus with respect to G-Brownian motion viewed through rough paths. Sci China Math, 2017, 60: 1–20
124 Ren L. On representation theorem of sublinear expectation related to G-Lévy process and paths of G-Lévy process. Statist Probab Lett, 2013, 83: 1301–1310
125 Rosazza Gianin E. Risk measures via g-expectations. Insurance Math Econom, 2006, 39: 19–34
126 Schmeidler D. Subjective probability and expected utility without additivity. Econometrica, 1989, 57: 571–587
127 Soner H M, Touzi N, Zhang J. Martingale representation theorem under G-expectation. Stochastic Process Appl, 2011, 121: 265–287
128 Soner H M, Touzi N, Zhang J. Quasi-sure stochastic analysis through aggregation. Electron J Probab, 2011, 16: 1844–1879
129 Soner H M, Touzi N, Zhang J. Wellposedness of second order backward SDEs. Probab Theory Related Fields, 2012, 153: 149–190
130 Song Y. Properties of hitting times for G-martingales and their applications. Stochastic Process Appl, 2001, 121: 1770–1784
131 Song Y. Some properties on G-evaluation and its applications to G-martingale decomposition. Sci China Math, 2011, 54: 287–300
132 Song Y. Uniqueness of the representation for G-martingales with finite variation. Electron J Probab, 2012, 17: 1–15
133 Song Y. Characterizations of processes with stationary and independent increments under G-expectation. Ann Inst H Poincaré Probab Statist, 2013, 49: 252–269
134 Song Y. Gradient estimates for nonlinear diffusion semigroups by coupling Methods. ArXiv:1407.5426, 2014
135 Song Y. Properties of G-martingales with finite variation and the application to G-Sobolev spaces. ArXiv:1607.00616, 2016
136 Szegö G. Risk Measures for the 21st Century. New York: Wiley-Finance, 2004
137 Vorbrink J. Financial markets with volatility uncertainty. J Math Econom, 2014, 53: 64–78
138 Walley P. Statistical Reasoning with Imprecise Probabilities. London-New York: Chapman and Hall, 1991
139 Wang L. On the regularity of fully nonlinear parabolic equations, II. Comm Pure Appl Math, 1992, 45: 141–178
140 Wiener N. Differential space. J Math Phys, 1923, 2: 131–174
141 Williams P M. Indeterminate probabilities. In: Formal Methods in the Methodology of Empirical Sciences. Dordrecht: Springer, 1976, 229–246
142 Wu P, Chen Z. Invariance principles for the law of the iterated logarithm under G-framework. Sci China Math, 2015, 58: 1251–1264
143 Xu J, Shang H, Zhang B. A Girsanov type theorem under G-framework. Stoch Anal Appl, 2011, 29: 386–406
144 Xu J, Zhang B. Martingale characterization of G-Brownian motion. Stoch Process Appl, 2009, 119: 232–248
145 Xu J, Zhang B. Martingale property and capacity under G-framework. Electron J Probab, 2010, 67: 2041–2068
146 Xu Y. Stochastic maximum principle for optimal control with multiple priors. System Control Lett, 2014, 64: 114–118
147 Zhang B, Xu J, Kannan D. Extension and application of Itô's formula under G-framework. Stoch Anal Appl, 2010, 28: 322–349
148 Zhang L. Exponential inequalities under sublinear expectations with applications to laws of the iterated logarithm. ArXiv:1409.0285, 2014
149 Zhang L. Donsker's invariance principle under the sub-linear expectation with an application to Chung's law of the iterated logarithm. Comm Math Statist, 2015, 3: 187–214
150 Zhang L. Rosenthal's inequalities for independent and negatively dependent random variables under sub-linear expectations with applications. Sci China Math, 2016, 59: 751–768
151 严加安. 测度论讲义. 第二版. 北京: 科学出版社, 2004

152 严加安. 金融数学引论. 北京: 科学出版社, 2012

Theory, methods and meaning of nonlinear expectation theory

PENG ShiGe

Abstract This is a survey on the research developments of nonlinear expectation theory. We first recall the basic definition of a space of nonlinear expectation, and then, through the representation theorem and some examples of nonlinear i.i.d. (independent and identically distributed), to explain why this new framework can be applied to calculate and quantitatively analyze probabilistic and distributional uncertainty hidden behind a real world (possibly high-dimensional) data sequence. Then we introduce two fundamentally important nonlinear normal distribution and maxima distribution and the corresponding nonlinear law of large numbers and nonlinear central limit theorem, which are crucial and fundamental breakthroughs of this new research domain. A typical application is a basic algorithm named "φ-max-mean". We also present a basic continuous-time stochastic process—nonlinear Brownian motion and its stochastic calculus, including stochastic integral, stochastic differential equations, and the corresponding nonlinear martingale theory. This new theoretical framework has generalized the axiomatical probability theory founded by Kolmogorov (1933). The key difference is the notion of nonlinear expectation $\hat{\mathbb{E}}$, whose special linear case corresponds a probability space $(\Omega, \mathcal{F}, P)$. It is the nonlinearity that allows us to quantitatively measure the uncertainty of probabilities and probabilistic distributions inhabited in our real world.

Keywords **nonlinear expectation, nonlinear normal distribution, nonlinear i.i.d., nonlinear large numbers theorem and central limit theorem, nonlinear Brownian motion and its stochastic calculus, nonlinear martingale theory, nonlinear Monté-Carlo method, φ-max-mean algorithm**

MSC(2010) **60A, 60H10, 60H05, 60H30, 60E05, 60E07, 62-00, 62C05, 62D05, 62F10, 62J05, 65C05, 93E03 35J, 35K**

doi: 10.1360/N012016-00209