值域更新问题

献给钱敏教授 90 华诞

谢践生

复旦大学数学科学学院, 上海 200433

E-mail: jsxie@fudan.edu.cn

收稿日期: 2017-05-24; 接受日期: 2017-08-16; 网络出版日期: 2017-11-28

国家自然科学基金 (批准号: 11271077) 资助项目

摘要 本文是有关“值域更新”问题的综述. 首先简要回顾了随机游动 (及 Brown 运动) 有关值域问题研究的历史脉络. 本文给出独立同分布 (i.i.d.)/连分数/简单对称随机游动三类模型下的“值域更新”结构定理, 并提出一些公开问题与猜测.

关键词 值域更新 连分数 简单对称随机游动

MSC (2010) 主题分类 11K50, 60G50, 60F15, 60F05, 60F17

1 引言

2010 年冬, 本文作者获知美国一所大学概率论方向的研究生入学考试中的一道考题如下¹⁾: 假设有 n 个独立同分布 (i.i.d.) 的简单样本, 其共同分布是一个离散型分布, 记 R_n 为样本取不同值的个数, 求证: 随着样本量 n 的增大, $\frac{R_n}{n}$ 几乎处处趋于 0. 这个简单问题的遍历论解法 (借鉴文献 [1] 中方法) 和概率解法 (借鉴文献 [2] 中方法) 激发了作者对类似问题的研究兴趣.

任意给定一个取可列无穷多的离散值的随机过程, 我们可以问许多与上面类似的问题. 例如, 可以考虑时间区间 $[0, n]$ 内过程访问的不同状态的数目 R_n 随着 n 变化的规律 (如大数律、中心极限定理、重对数律和大偏差等); 借助随机网络图研究领域的思想²⁾ (参见文献 [3, 4]), 还可以考察时间区间 $[0, n]$ 内过程访问频次恰为 k 的状态的数目 $R_{n,k}$ (特别是比值 $\frac{R_{n,k}}{R_n}$) 随着 n 变化的规律; 凡此等等.

上述类型的几个问题我们统称为值域更新问题. 这些问题实际上起源于随机游动 (含 Brown 运动) 这一概率理论中久盛不衰且影响重大的课题. 有关历史脉络, 我们简单梳理如下:

(1) 历史上, 最先开始研究的是“值域元素个数” R_n 这个量, 相应的背景过程通常为 (简单、对称) 随机游动. 1950 年前后, Dvoretzky 和 Erdős^[2] 研究了 $d \geq 2$ 维的整数格点上的简单对称随机游动前 n 步历经的格点数目 (“值域元素个数”), 得到了相应的强大数律结果 (其中 2 维的结果论证中略有瑕疵,

¹⁾在此感谢应坚刚教授的无私分享与交流讨论.

²⁾在此感谢刘兴武、卜东波同志在 2007 年前后关于随机网络方向的无私分享与讨论交流, 使得作者首次获知相关文献与 Power Law 等现象.

英文引用格式: Xie J S. Range-renewal problems (in Chinese). Sci Sin Math, 2017, 47: 1761–1770, doi: 10.1360/N012017-00109

后被 Jain 和 Pruitt^[5] 修正); 后续 Jain 和 Pruitt^[6,7] 及 Le Gall^[8] 研究了这个量的中心极限定理; 进一步, Jain 和 Pruitt^[9] 及 Bass 和 Kumagai^[10] 研究了 R_n 的重对数律等.

(2) 后来有学者进一步研究了随机游动或 Brown 运动的轨迹; 这方面有诸多相关的数学量 (或物理量) 被研究. Brown 运动轨迹等的 Hausdorff 维数问题研究由 Taylor^[11-13] 及 Erdős 和 Taylor^[14] 等首开先河. Dvoretzky 等^[15]、Dvoretzky 等^[16] 和 Taylor^[17] 则较早开展了 Brown 运动的轨迹中的双重点 (double points)、三重点 (triple points) 和多重点 (multiple points) 的研究; Erdős 和 Taylor^[18,19] 同时研究了随机游动轨迹的高频访问点 (frequently visited sites, favorite sites) 和相交性质 (intersection properties) 等.

上述有关随机游动和 Brown 运动的研究方向自 20 世纪中叶兴起、深化、拓广并延续到 21 世纪, 大部分的研究成果被收录在 Révész 编写的专著 [20]³⁾ 中; 众多华人学者 (如 Shao^[21] 和 Shi^[22] 等) 也有突出贡献被专著 [20] 引用. 专著 [20] 版本的多次更新, 页码的不断增多, 从一个侧面强有力地说明了此领域研究所具有的丰富成果、旺盛生命力和深刻影响力等.

有关随机游动的“轨迹”问题 (或广义的“值域”问题) 的研究, 最近 20 多年来有一些新的进展. 围绕有关随机游动的无穷次相遇或有限次相遇的问题, Krishnapur 和 Peres^[23] 发现了一类具有有限相交性质的常返图, 由此引发了国内学者 (参见文献 [24-28]) 对此类现象或相关性质的研究.

本文只讨论狭义的“值域”问题. 更具体地, 本文拟介绍本文作者与合作者在值域更新问题研究中, 结合前人已有结果逐渐形成的部分理论 (我们称之为“值域更新结构”理论) 及其有关公开问题. 下文中将按照 i.i.d. 模型、连分数模型、随机游动模型的顺序分别进行介绍.

2 i.i.d. 模型下的值域更新结构理论与公开问题

考虑引言中提及的 i.i.d. 模型与相关问题. 假定简单样本 $\{X_n : n \geq 1\}$ 来源于一个离散分布 π ; 由于我们不关心样本的具体取值, 因而不妨假设样本取值于正整数集合 $\mathbb{N}$, 且为避免平凡性, 不妨假设概率分布列 $\{\pi_x := \mathbb{P}(X_1 = x)\}_{x=1}^{\infty}$ 是单调递减的正数列. 记 $R_n := \#\{X_k : 1 \leq k \leq n\}$ 为前 n 个样本的不同取值个数, 其他记号 $R_{n,k}$ 等按照引言中解释理解. 在本节, 样本编号同时也作时间理解.

关于 R_n , 容易证明下面的结果, 它详细回答了引言第一段的问题.

定理 2.1^[29] $\mathbb{E}R_n = \sum_x [1 - (1 - \pi_x)^n]$, $\text{Var}(R_n) \leq \mathbb{E}R_n$. 于是几乎处处有

$$\lim_{n \rightarrow +\infty} \frac{R_n}{\mathbb{E}R_n} = 1.$$

进而

$$\lim_{n \rightarrow +\infty} \frac{R_n}{n} = 0.$$

在进一步假定分布 π 的尾部具有一定规则性的基础上, 可以探讨时间 $[0, n]$ 内 (即前 n 个样本中) 恰被访问 k 次的状态数目

$$R_{n,k} := \sum_x 1_{\{N_n(x)=k\}}, \quad (2.1)$$

其中

$$N_n(x) := \sum_{k \in [0, n]} 1_{\{X_k=x\}},$$

³⁾ 目前该专著出版了三个版本: 1990 年第一版, 前言等 15 页, 正文 332 页; 2005 年第二版, 前言等 16 页, 正文 380 页; 2013 年第三版, 前言等 18 页, 正文 402 页.

被访问至少 k 次的状态数目 $R_{n,k+} := \sum_{\ell \geq k} R_{n,\ell}$. 定义

$$\varphi(x) := \frac{1}{\pi_x}, \quad x \in \mathbb{N};$$

我们假定 φ 有反函数 (记作 φ^{-1})⁴⁾ 此处所谓的“分布 π 的尾部具有一定规则性”, 其含义是函数 φ^{-1} 在 Karamata 的意义下关于指数 γ 正则 (regular), 或称 φ^{-1} 以指数 γ 规则变化, 即 $\varphi^{-1}(x) = x^\gamma \cdot L(x)$, 其中 L 为缓慢变化函数; 由于论证技术的需要, 我们通常还要求这个缓慢变化函数有适当控制, 此处忽略更具体的定义式, 感兴趣的读者可以参见文献 [30] 中有关规则变化函数的章节及文献 [29, 第 5–7 页]. 显然, 由于 π 是一个概率分布列, 应有 $\gamma \in [0, 1]$; $\gamma = 0$ 及 $\gamma = 1$ 是临界情形. 举一个简单例子, 如果分布 π 满足

$$\pi_x = \frac{C + o(1)}{x^\alpha}, \quad x \rightarrow \infty$$

其中 $\alpha > 1$, $C > 0$ 为常数, 那么 $\gamma = 1/\alpha \in (0, 1)$.

为了突出主要思想和结果, 本文只陈述非临界情形的结果; 临界情形的处理参见文献 [29].

定理 2.2 ^[29] 设分布 π 规则, 且具有指数 $\gamma \in (0, 1)$, 则

$$\mathbb{E}R_n = \Gamma(1 - \gamma) \cdot \varphi^{-1}(n) \cdot (1 + o(1)).$$

对任意自然数 $\ell \geq 1$ 几乎处处有

$$\lim_{n \rightarrow +\infty} \frac{R_{n,\ell}}{R_n} = r_\ell(\gamma) := \frac{\gamma \cdot \Gamma(\ell - \gamma)}{\ell \Gamma(1 - \gamma)}.$$

特别地, 记 $R_{n,\ell+} := \sum_{k \geq \ell} R_{n,k}$, 则

$$\lim_{n \rightarrow +\infty} \frac{R_{n,\ell}}{R_{n,\ell+}} = \frac{\gamma}{\ell}.$$

序列 $\{r_\ell(\gamma) : \ell \in \mathbb{N}\}$ 是一个指数为 $1 + \gamma$ 的 Power Law 序列⁵⁾.

由上述定理可以知道如下一个有意思的事实: 当 $\frac{1}{1+\gamma} = \gamma$, 即 $\gamma = 0.618\dots$ 为黄金分割数时, 取 $\pi := \{\pi_x = r_x(\gamma)\}_{x=1}^\infty$, 则分布 π 诱导的“值域更新结构”中的分布律仍然是 π .

上述定理 2.2 与随机游动模型的暂留情形的结果相呼应, 见本文定理 4.1. 文献 [31] 进一步讨论了 i.i.d. 情形下一般分布的相应结果.

平行类比 Jain、Pruitt、Le Gall、Bass、Kumagai 等工作, 有如下很自然的问题:

问题 2.3 对于当前的 i.i.d. 模型, 是否能建立 R_n 的中心极限定理, 乃至重对数律?

3 连分数模型下的值域更新结构理论与公开问题

在上一节结果的基础上, 一个很自然的观点是, 对于正常返模型, 上述 i.i.d. 模型下的理论结果很可能具有普遍性, 只要这个模型具有较强的渐近独立性 (混合性), 并且相应平稳不变分布具有一定规则性. 机缘巧合之下我们找到连分数模型这一个研究切入点. 注意到连分数系统在配上 Gauss 测

⁴⁾ 此处只要 φ^{-1} 对充分大的数有定义即可.

⁵⁾ 在随机网络图领域中通常意义的 Power Law 现象是关于图形中节点的边数 (degree) 的频率的极限分布, 且 Power Law 指数通常 > 2 (也有人认为 > 3). 本文中考虑的则是关于点的访问频率的极限分布. 或许两处 Power Law 指数取值范围的差别在于我们的 i.i.d. 模型可以视为已经成熟的网络, 而通常随机网络图领域的研究对象则是规模很大的、发展中 (容许删除节点或连边) 的网络模型.

度 $\mu^{(6)}$ 后形成一个保测系统. 具体来说, 可以定义 $\mathbb{X} := (0, 1) \setminus \mathbb{Q}$ 为 $(0, 1)$ 中的无理数全体, 并定义 Gauss 映射

$$T(x) := \frac{1}{x} \pmod{1}.$$

于是, Gauss 测度 μ 是 T 不变概率测度, 并且保测系统 $(\mathbb{X}, \mu, T)$ 是指数速度混合的 (参见文献 [32, 第 43–51 页]). 对于任意无理数 $x \in \mathbb{X}$, 记其无穷连分数展式为

$$x = [a_1, a_2, \dots] := \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}.$$

于是有符号空间 $\Omega := \mathbb{N}^{\mathbb{N}}$ 到 $\mathbb{X}$ 的自然投射

$$p: \omega = (a_1, a_2, \dots) \mapsto x = [a_1, a_2, \dots].$$

上述投射是一一的. 在符号空间 Ω 上有自然的左推移映射 σ 以及概率测度 $\mathbb{P} := \mu \circ p$, 使得下图交换:

$$\begin{array}{ccc} (\Omega, \mathbb{P}) & \xrightarrow{\sigma} & (\Omega, \mathbb{P}) \\ p \downarrow & & \downarrow p \\ (\mathbb{X}, \mu) & \xrightarrow{T} & (\mathbb{X}, \mu) \end{array}$$

Ω 上坐标过程 $\{a_n(\omega) : n \geq 1\}$ 在概率测度 $\mathbb{P}$ 下是一个平稳过程, 并且具有很好的渐近独立性质, 它满足对于充分大 $x \in \mathbb{N}$, 有

$$\pi_x := \mathbb{P}(a_1 = x) = \frac{1}{(\log 2) \cdot (x+1)^2} + O\left(\frac{1}{x^4}\right). \quad (3.1)$$

我们同样可以考察关于坐标过程 $\{a_n(\omega) : n \geq 1\}$ 的 R_n 和 $R_{n,k}$ 等量的性质. 由此出发, 借助上一节对 i.i.d. 模型已经建立的理论, 文献 [33] 证明了如下结果:

定理 3.1 对于 Lebesgue 几乎处处的点 $\omega \in (0, 1)$, 其连分数展开式得到的序列 $\{a_n(\omega) : n \geq 1\}$ 满足

$$\lim_{n \rightarrow \infty} \frac{R_n(\omega)}{\sqrt{n}} = \sqrt{\frac{\pi}{\log 2}}. \quad (3.2)$$

对任意 $k \geq 1$, 有

$$\lim_{n \rightarrow \infty} \frac{R_{n,k}(\omega)}{R_n(\omega)} = \frac{C_{2k}^k}{(2k-1) \cdot 4^k} =: r_k, \quad \lim_{n \rightarrow \infty} \frac{R_{n,k}(\omega)}{R_{n,k+}(\omega)} = \frac{1}{2k}, \quad (3.3)$$

其中

$$C_n^m := \frac{n!}{m!(n-m)!}, \quad \sum_{k=1}^{\infty} r_k = 1.$$

序列 $\{r_k\}_1^{\infty}$ 是一个指数为 $\frac{3}{2}$ 的 Power Law 序列.

定理 3.2 对任意 $\beta \geq 0, c > 0$, 定义

$$\begin{aligned} E(\beta, c) &:= \left\{ \omega \in [0, 1) : \lim_{n \rightarrow \infty} \frac{R_n(\omega)}{n^\beta} = c \right\}, \\ F(c) &:= \left\{ \omega \in [0, 1) : \lim_{n \rightarrow \infty} \frac{R_n(\omega)}{n} = c \right\}. \end{aligned}$$

⁶⁾注意, Gauss 测度 μ 是 $(0, 1)$ 上具有密度函数 $\frac{1}{(\log 2) \cdot (1+x)}$ 的概率测度, 显然不是概率论中通常谈的 Gauss 分布.

有如下结论:

(1) 如果 $\beta = 0$, 那么

$$\dim_H \bigcup_{c \in \mathbb{N}} E(\beta, c) = 1, \quad \lim_{\mathbb{N} \ni c \rightarrow \infty} \dim_H E(\beta, c) = 1; \quad (3.4)$$

(2) 如果 $\beta \in (0, 1)$ 且 $c > 0$, 那么

$$\dim_H E(\beta, c) = 1; \quad (3.5)$$

(3) 如果 $c \in (0, 1]$, 那么

$$\dim_H F(c) = \frac{1}{2}. \quad (3.6)$$

文献 [34, 35] 讨论了 $E(0, c)$ 的 Hausdorff 维数的计算公式, 此处从略; 本质上维数结论 (3.4) 来源于文献 [34, 35].

本节一个很自然的公开问题同样是如下问题:

问题 3.3 对于当前的连分数模型, 是否能建立 R_n 的中心极限定理, 乃至重对数律?

从已有结果论证路线看, 上面这个问题的解答可能有赖于问题 2.3 的解决.

4 随机游动模型下的值域更新结构理论与公开问题

关于随机游动的常返暂留性的讨论源于 Pólya 的工作 [36], 他论证了 $\mathbb{Z}^d$ 上的简单对称随机游动 $\{S_n : n \geq 0\}$ 在 $d \geq 3$ 时是暂留的, 具有逃逸概率 $\gamma_d \in (0, 1)$; 在 $d = 1$ 或 $d = 2$ 时是零常返的, 此时逃逸概率 $\gamma_d = 0$.

在本节中, $R_n := \#\{S_k : 0 \leq k \leq n\}$ 表示时间区间 $[0, n]$ 内过程访问的不同节点数目. $R_{n,k}$ 等作类似理解 (即 (2.1) 中 X 替换为 S).

4.1 暂留情形: $d \geq 3$ 维简单对称随机游动

定理 4.1 对于 $\mathbb{Z}^d$ ($d \geq 3$) 上的简单对称随机游动 $\{S_n : n \geq 0\}$, 下式成立:

$$\lim_{n \rightarrow \infty} \frac{R_n(\omega)}{n} = \gamma_d, \quad (4.1)$$

其中 γ_d 为过程的逃逸概率. 对任意 $k \geq 1$, 有

$$\lim_{n \rightarrow \infty} \frac{R_{n,k}(\omega)}{R_n(\omega)} = \gamma_d(1 - \gamma_d)^{k-1}, \quad \lim_{n \rightarrow \infty} \frac{R_{n,k}(\omega)}{R_{n,k+}(\omega)} = \gamma_d. \quad (4.2)$$

关于 k 而言, 上述序列 $\{\gamma_d(1 - \gamma_d)^{k-1}\}_1^\infty$ 是几何分布序列.

上述结果中 (4.1) 对于 $d \geq 1$ 均成立, 只不过 $\gamma_1 = \gamma_2 = 0$, $\gamma_d > 0$, $\forall d \geq 3$. (4.2) 本质上可由文献 [20] 中结论导出, 但该形式首次出现于文献 [37], 该文献作者利用次可加遍历定理把该结论推广到一般的离散群上暂留的简单随机游动框架.

如引言中所述, 对于 $d \geq 3$ 情形下的简单对称随机游动, R_n 的中心极限定理、重对数律等都被前人解决, 我们略去相应结果的准确陈述, 感兴趣的读者可以参见引言中提及的相关文献.

显然, 本情形下的所谓“值域更新结构”理论是本文提及的所有模型中最为成熟的.

4.2 零常返情形 (A): 一维简单对称随机游动

我们通常的印象中, 一维的简单对称随机游动是所有随机游动中最简单的模型, 那么关于它的所谓“值域更新问题”的解答似乎应该很简单且完整; 但事实上很少有文献完整论及于此. 当然关于 R_n 的极限律是比较简单的; 它有如下的中心极限定理成立:

定理 4.2 $\frac{R_n}{\sqrt{n}} \xrightarrow{d} \sup\{B_t - B_s : s, t \in [0, 1]\}$, 其中 $\{B_t : t \geq 0\}$ 是一维标准 Brown 运动.

上述定理可以通过 Donsker 不变原理 (或称泛函中心极限定理) 来快速论证. 事实上, 由于

$$\left\{ \frac{S_{[nt]}}{\sqrt{n}} : t \in [0, 1] \right\} \xrightarrow{d} \{B_t : t \in [0, 1]\}, \quad (4.3)$$

此处 $[\cdot]$ 表示取整函数, 而显然,

$$\frac{R_n - 1}{\sqrt{n}} = \frac{1}{\sqrt{n}} \cdot \sup\{S_{[nt]} - S_{[ns]} : s, t \in [0, 1]\},$$

其中涉及的轨道泛函 $g(\omega) := \sup\{\omega_t - \omega_s : s, t \in [0, 1]\}$ 是连续的, 因而定理 4.2 成立.

很自然会问:

问题 4.3 对于 $\mathbb{Z}$ 上简单对称随机游动, $R_{n,k}$ 的极限律如何?

观察 (4.3). 我们可以想象如下模型: 有人拿着一只出墨水速度均匀的画笔, 按照随机游动的规律对 $\mathbb{Z}$ 上的点 (反复) 染色; 于是由于访问频率的差别, 诸点颜色有差别; 再按照 (4.3) 中左端规律作空间的尺度变换, 得到连续区间上的染色. 于是 $\frac{R_{n,k}}{\sqrt{n}}$ 可以解释为具有染色水平 $\frac{k}{\sqrt{n}}$ 的点的 Lebesgue 测度. 特别地, 具有染色水平 $\frac{k}{\sqrt{n}}$ 的点也可以解释为在时间区间 $[0, 1]$ 内具有局部时水平 $\frac{k}{\sqrt{n}}$; 这是因为在上述过程中, $\frac{1}{n}$ 和 $\frac{1}{\sqrt{n}}$ 分别是时间尺度和空间尺度的单位, $k \cdot \frac{1}{n}$ 代表了离散过程访问某空间点 k 次所消耗的实际时间 (“占位时”), 因而 $\frac{k}{\sqrt{n}} = k \cdot \frac{1}{n} / \frac{1}{\sqrt{n}}$ 代表了被离散过程访问了恰好 k 次的那些实际空间点相应的占位时测度值的空间密度 (“占位时密度”), 亦即所谓的局部时水平. 于是有下面的猜测:

猜测 4.4 对于 $\mathbb{Z}$ 上简单对称随机游动,

$$\frac{1}{\sqrt{n}} \sum_{k: k \leq \sqrt{n}y} R_{n,k} \xrightarrow{d} \text{Leb}(\{x : L_1^x(\omega) \leq y\}), \quad (4.4)$$

其中 $y > 0$, $\text{Leb}(\cdot)$ 表示 Lebesgue 测度, $L_t^x(\omega)$ 表示标准 Brown 运动 $\{B_s : s \geq 0\}$ 在 $[0, t]$ 时间内 x 位置的局部时. 特别地, 可以猜测样本累积分布函数的极限

$$F_n(y) := \sum_{k: k \leq \sqrt{n}y} \frac{R_{n,k}}{R_n} \xrightarrow{d} \frac{\text{Leb}(\{x : L_1^x(\omega) \leq y\})}{\sup\{B_t - B_s : s, t \in [0, 1]\}} =: F_\omega(y).$$

上式中后者是一个随机的 (依赖于 Brown 运动轨道的) 分布函数. 针对这个模型, 我们有如下通过模拟一条随机游动轨道得到的模拟结果的两张柱状图 (见图 1), 其中横坐标为 $\frac{k}{\sqrt{n}}$, 纵坐标为对应的频率或累积频率, 柱子的水平位置区间总取为区间 $[\frac{k-0.5}{\sqrt{n}}, \frac{k+0.5}{\sqrt{n}}]$. 上图是 $\frac{R_{n,k}}{R_n}$ 关于 $\frac{k}{\sqrt{n}}$ 的柱状图 (类似于频率直方图). 下图是样本分布函数 $F_n(\frac{k}{\sqrt{n}})$ 关于 $\frac{k}{\sqrt{n}}$ 的柱状图 (类似于累积频率的柱状图); 并且对于 $\frac{R_{n,k}}{R_n} = 0$ 的 k , 对应位置不做图 (这对应于真正的样本分布函数图像中的水平折线), 这也是此图中产生竖条纹空隙乃至空白的原因 (采用这种做图方法是为了观察样本分布函数的支撑集是否具有分形现象).

图 1 模拟柱状图, 其中 $n = 10^6$

关于 Brown 运动的局部时有下面著名的 Tanaka 公式:

$$L_t^x(\omega) = |B_t - x| - |B_0 - x| - \int_0^t \operatorname{sgn}(B_s - x) dB_s, \quad (4.5)$$

其中符号函数 $\operatorname{sgn}(x)$ 对于非零实数 x 根据其正负情形总取值于 ± 1 , $\operatorname{sgn}(0) = 0$.

上述猜测的思想与如下一些等式对之间是相容的:

$$\sum_k k R_{n,k} = n + 1, \quad \int_0^1 L_1^x(\omega) dx = 1, \quad (4.6)$$

$$\sum_k R_{n,k} = R_n, \quad \operatorname{Leb}(\{x : L_1^x(\omega) > 0\}) = \sup\{B_t - B_s : s, t \in [0, 1]\}. \quad (4.7)$$

因此我们倾向于猜测 1 成立.

借助上面的模式, 还可以探讨 $\max\{R_{n,k} : k \geq 1\}$ 及达到这个极大值的 k 值 K_n (即它满足 $R_{n,K_n} = \max\{R_{n,k} : k \geq 1\}$) 的极限律. 假如 (4.4) 左端作为 y 的单调函数具有连续的密度 $\rho(y, \omega)$, 那么可以猜测

$$\max\left\{\frac{R_{n,k}}{\sqrt{n}} : k \geq 1\right\} \xrightarrow{d} \max_y \rho(y, \omega), \quad \frac{K_n}{\sqrt{n}} \xrightarrow{d} \arg \max_y \rho(y, \omega).$$

然而上述众多猜测无法直接使用 Donsker 不变原理来论证, 其原因在于这个泛函中心极限定理要求涉及的泛函必须是关于轨道的连续泛函 — 而我们猜测中涉及的泛函在一致范数下不连续.

撇开这些技术困难, 我们由上述讨论也看到一个关于 Brown 运动局部时的有趣问题:

问题 4.5 (4.4) 右端作为 y 的单调函数是否几乎处处有密度? 如果有密度, 其连续性等分析性质如何?

从模拟图给人的感觉来看, 上述问题 4.5 的答案可能是否定的. 由此, 关于局部时 $L_t^x(\omega)$ 的 (关于空间变量 x 的) 水平集或上面提及的分布函数 $F_\omega(\cdot)$ 的支撑集的分形研究也是非常有价值的. 有兴趣的读者或许可以参见文献 [38] 及其他有关工作.

另外, 由于一维简单对称随机游动可以视作 Bernoulli 实验结果的泛函过程, 也可以视作 $(0, 1)$ 上的二进制“无理数”的展开对应的泛函过程. 于是类似我们在连分数的框架下的结果, 可以探讨:

问题 4.6 R_n 的水平集的 Hausdorff 维数问题.

上面这个问题用于粗略理解有“多少”样本点 ω 使得 $R_n(\omega) \sim c \cdot n^\beta$ 成立. 这是分形维数理论中的典型问题, 也可视作数论中的分形问题, 它与 $[0, 1)$ 上扩张映射 $T(x) = 2x \pmod{1}$ 有深刻联系.

4.3 零常返情形 (B): 2 维简单对称随机游动

在 Révész 的专著 [20] 中, $\mathbb{Z}^2$ 上简单对称随机游动的“值域更新”方面的结论可以罗列为如下强大数律 (参见文献 [20, 定理 20.11]): 对任意 $k \geq 1$,

$$\lim_{n \rightarrow \infty} \frac{R_n}{\pi n / \log n} = 1 \text{ a.s.}, \quad (4.8)$$

$$\lim_{n \rightarrow \infty} \frac{R_{n,k}}{\pi^2 n / \log^2 n} = 1 \text{ a.s.}, \quad (4.9)$$

这将导致

$$\lim_{n \rightarrow \infty} \frac{R_{n,k}}{R_n} \cdot \frac{\log n}{\pi} = 1 \text{ a.s.}$$

对于 2 维 Brown 运动, 局部时的概念无法照搬一维情形的定义 (参见文献 [38–40] 等文献中的相关讨论), 此时自然也无法照搬上一小节中一维模型下的那些讨论来进行研究; 但类似问题仍然值得探索.

5 总结与致谢

我们在前几节罗列了 i.i.d.、连分数、 $\mathbb{Z}^d$ 上的简单对称随机游动三类模型下的值域更新问题的相关结果及一些公开问题. 按照随机游动理论发展的规律, 很自然也可以考虑更一般的图 (如规则树等) 或随机图上的随机游动等模型的“值域更新问题”; 从已有的文献与研究结果来看, 常返模型的值域更新问题通常难度更大.

本文第 4.2 小节中的频率直方图与累积频率柱状图是余景辉同学通过 Python 软件模拟计算绘制的, 特此致谢. 作者也感谢审稿人针对本文初稿给出的细致修改意见.

本文作者在研究过程中感受到, 问题驱动式的研究方法让人受益匪浅 — 而这也是李大潜院士等提倡的. 钱敏先生在我们一众师兄弟等的历年讨论班上也是一直言传身教地向我们极力倡导、辐射影响这种理念和研究思路; 他与钱敏平教授等还积极提倡把简单、具体、形象的模型吃透再深挖、抽象、拔高. 本文作者记得本科同学曾焰所选修毕业论文讨论班的指导老师就是钱敏先生; 曾焰课后与本文作者论及先生让讨论班的同学把泛函 I、II 中的一些抽象化的定理返回到有限维欧氏空间来看其原型, 由此理解、品味高深数学理论的具象与抽象, 从平凡现象与事实中挖掘不平凡, 大发感触, 既叹先生眼界学识、又仰先生风骨⁷⁾ — 此为本文作者首次与先生间接结缘, 顿生钦慕之心; 后有幸在钱敏平教授

⁷⁾当时先生已退休, 坚持讨论班等学术育人活动多年, 俱为义务劳动.

指导下硕博连读, 长年受教于先生. 本文中谈及的 i.i.d. 模型无疑简单易懂, 问题也容易理解、易于下手, 为美国高校学生习题, 容易被忽视. 正是钱敏先生、钱敏平教授等师长的影响与鼓励, 本文作者方能以他们为榜样在此简单模型的研究成果的基础上, 敝帚自珍并逐步深挖. 适值先生 90 华诞, 借此机会作者感谢钱敏先生、钱敏平教授等师长多年培养、学术道路上一路帮扶、生活上细致关心的恩情, 并祝钱敏先生与家人身体健康、福寿绵延.

参考文献

- 1 Derriennic Y. Entropie, théorèmes limite et marches aléatoires (French) [Entropy, limit theorems and random walks] Probability measures on groups, VIII (Oberwolfach, 1985). Lecture Notes in Mathematics, vol. 1210. Berlin: Springer, 1986, 241–284
- 2 Dvoretzky A, Erdős P. Some problems on random walk in space. In: Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability (1950). Berkeley-Los Angeles: University of California Press, 1951, 353–367
- 3 Barabási A L, Albert R. Emergence of scaling in random networks. Science, 1999, 286: 509
- 4 Barabási A L. Scale-free networks: A decade and beyond. Science, 2009, 325: 412
- 5 Jain N C, Pruitt W E. The range of random walk. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, vol. 3. Berkeley: University of California Press, 1972, 31–50
- 6 Jain N C, Pruitt W E. The range of transient random walk. J Anal Math, 1971, 24: 369–393
- 7 Jain N C, Pruitt W E. Further limit theorems for the range of random walk. J Anal Math, 1974, 27: 94–117
- 8 Le Gall J F. Propriétés d'intersection des marches aléatoires, I: Convergence vers le temps local d'intersection. Comm Math Phys, 1986, 104: 471–507
- 9 Jain N C, Pruitt W E. The law of the iterated logarithm for the range of random walk. Ann Math Statist, 1972, 43: 1692–1697
- 10 Bass R F, Kumagai T. Laws of the iterated logarithm for the range of random walks in two and three dimensions. Ann Probab, 2002, 30: 1369–1396
- 11 Taylor S J. The Hausdorff α -dimensional measure of Brownian paths in n -space. Math Proc Cambridge Philos Soc, 1953, 49: 31–39
- 12 Taylor S J. The α -dimensional measure of the graph and set of zeros of a Brownian path. Math Proc Cambridge Philos Soc, 1955, 51: 265–274
- 13 Taylor S J. The exact Hausdorff measure of the sample path for planar Brownian motion. Math Proc Cambridge Philos Soc, 1964, 60: 253–258
- 14 Erdős P, Taylor S J. On the Hausdorff measure of Brownian paths in the plane. Math Proc Cambridge Philos Soc, 1961, 57: 209–222
- 15 Dvoretzky A, Erdős P, Kakutani S. Double points of paths of Brownian motion in n -space. Acta Sci Math (Szeged), 1950, 12: 75–81
- 16 Dvoretzky A, Erdős P, Kakutani S, et al. Triple points of Brownian paths in 3-space. Math Proc Cambridge Philos Soc, 1957, 53: 856–862
- 17 Taylor S J. Multiple points for the sample paths of the symmetric stable process. Z Wahrscheinlichkeitstheorie Verw Gebiete, 1966, 5: 247–264
- 18 Erdős P, Taylor S J. Some problems concerning the structure of random walk paths. Acta Math Acad Sci Hungar, 1960, 11: 137–162
- 19 Erdős P, Taylor S J. Some intersection properties of random walk paths. Acta Math Acad Sci Hungar, 1960, 11: 231–248
- 20 Révész P. Random Walk in Random and Non-Random Environments, 3rd ed. Hackensack: World Scientific Publishing, 2013
- 21 Shao Q M. On a conjecture of Revesz. Proc Amer Math Soc, 1995, 123: 575–582
- 22 Shi Z, Toth B. Favourite sites of simple random walks. Period Math Hungar, 2000, 41: 237–249
- 23 Krishnapur M, Peres Y. Recurrent graphs where two independent random walks collide finitely often. Electron Comm Probab, 2004, 9: 72–81
- 24 Chen D Y, Wei B, Zhang F X. A note on the finite collision property of random walks. Statist Probab Lett, 2008, 78: 1742–1747
- 25 Chen X X, Chen D Y. Two random walks on the open cluster of $\mathbb{Z}^2$ meet infinitely often. Sci China Math, 2010, 53:

- 1971–1978
- 26 Chen X X, Chen D Y. Some sufficient conditions for infinite collisions of simple random walks on a wedge comb. *Electron J Probab*, 2011, 16: 1341–1355
 - 27 Chen X X. Gaussian bounds and collisions of variable speed random walks on lattices with power law conductances. *Stochastic Process Appl*, 2016, 126: 3041–3064
 - 28 陈大岳, 章复熹. 随机游动的常返和相遇问题. *数学进展*, 2014, 43: 175–182
 - 29 Chen X X, Xie J S, Ying J G. Range-Renewal processes: SLLN, power law and beyonds. *ArXiv:1305.1829*, 2013
 - 30 Feller W. *An Introduction to Probability Theory and Its Applications: Volume II*, 2nd ed. New York: John Wiley & Sons, 2008
 - 31 Wang L H, Xie J S. Range-renewal structure of i.i.d. samplings. *ArXiv:1310.4259*, 2013
 - 32 Iosifescu M, Kraaikamp C. *Metrical Theory of Continued Fractions*. Mathematics and its Applications, 547. Dordrecht: Kluwer Academic Publishers, 2002
 - 33 Wu J, Xie J S. Range-renewal structure in continued fractions. *Ergodic Theory Dynam Systems*, 2017, 37: 1323–1344
 - 34 Good I J. The fractional dimensional theory of continued fractions. *Math Proc Cambridge Philos Soc*, 1941, 37: 199–228
 - 35 Jarník V. Zur metrischen Theorie der diophantischen Approximationen. *Práce Mat-Fiz*, 1928-9, 36: 91–106
 - 36 Pólya G. Über eine aufgabe der wahrscheinlichkeitsrechnung betreffend die irrfahrt im strassennetz (German). *Math Ann*, 1921, 84: 149–160
 - 37 Xie J S, Xu Y Y. Range-renewal structure of transient simple random walk. *Statist Probab Lett*, 2013, 83: 2220–2221
 - 38 Dembo A, Peres Y, Rosen J, et al. Thick points for spatial Brownian motion: Multifractal analysis of occupation measure. *Ann Probab*, 2000, 28: 1–35
 - 39 Hu X, Taylor S J. The multifractal structure of stable occupation measure. *Stochastic Process Appl*, 66: 1997, 283–299
 - 40 Taylor S J. Regularity of irregularities on a Brownian path. *Ann Inst Fourier (Grenoble)*, 1974, 39: 195–203

Range-renewal problems

XIE JianSheng

Abstract In this paper, we give a simple survey on the so-called range-renewal problems. We first briefly review the history of related researches on random walks. Then we present range-renewal structure theorems for i.i.d./continued fractions/simple symmetrical random walks models and also some open problems and conjectures.

Keywords range-renewal, continued fractions, symmetrical simple random walks

MSC(2010) 11K50, 60G50, 60F15, 60F05, 60F17

doi: 10.1360/N012017-00109