

令和元年6月12日現在

機関番号：32612

研究種目：基盤研究(C) (一般)

研究期間：2014～2018

課題番号：26400146

研究課題名(和文) 確率順位付けモデルとウェブランキングへの応用

研究課題名(英文) Stochastic ranking process and its applications to web ranking

研究代表者

服部 哲弥 (HATTORI, Tetsuya)

慶應義塾大学・経済学部(日吉)・教授

研究者番号：10180902

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：ウェブで見られるランキングの時間発展のモデルとなる確率順位付けモデルの位置強度結合経験分布の大数の法則(流体力学極限)と軌道についてのカオスの伝搬の証明を、強度(先頭に跳ぶ確率を与える点過程を定める関数)が位置依存性を持つ場合に一般化した。この一般化は、極限を記述する点過程の強度が直前の到着時刻に依存する非独立増分であるなど、強度が位置依存性を持たない場合の非自明な摂動問題である。確率順位付けモデルはネット書店の売上順位の時間変化のモデルで、ロングテール構造の数理解析の基礎モデルとなる。強度の位置依存性は好順位の注目効果を意味する。

研究成果の学術的意義や社会的意義

急速な発達によって時代を特徴付けるに至った計算機とネットワーク環境によって、インターネット小売業のような大規模なアイテムの人気度の全順位がリアルタイムで更新される現象が可視化されオンライン中古市場のように経済的な意味を持つ時代に、実際のデータと現象の対応する数学的な方程式を「細いがまっすぐな道でつなぐ」こと(末節をそぎ落とした論理的中核の抽出)に成功した。世界的に例のない独自の研究である。最終的に到達した強度の位置依存性効果は、ランキングデータ自体の宣伝効果を与えるもので、数学的に精密で非自明な結果であり、将来技術が追いついてきたときに前もって筋の良い理屈を用意できていることを意味する。

研究成果の概要(英文)：We generalized a proof of existence of hydrodynamic limits and propagation of chaos of the stochastic ranking processes, to the cases of spatially dependent intensities for the point processes which determine the move-to-front (jump-to-top) random times. The processes are particle systems which mathematically model the rankings on the web, such as sales ranks of online bookstores, and the models give mathematical basis for the analysis of so called long-tail structure of online retails. The spatial dependence of intensities correspond to appeal effect of top sales. From mathematical point of view, introduction of spatial dependence on the intensity functions is a non-trivial perturbation, leading to continuum limits described in terms of the point processes with last-arrival-time dependent intensities.

研究分野：数物系科学

キーワード：確率論 数理科学 確率過程論 大数の法則 流体力学極限 粒子系 確率順位付けモデル

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

(1) 本研究の研究代表者による前年度までの科研費基盤研究(C)「くりこみ群に基づく確率解析の基盤研究」(2009~2013年度)において、数学の外にある現象や科学的知見の中から既存の数理モデルとは異なる数学的広がりを持つクラスを発見しその数学的性質を研究するという方法論の具体例として、確率順位付け模型の研究を本格的に開始した。この模型の原型は、1列に並ぶ粒子が先頭に跳ぶ規則と呼ばれる特徴的な規則に従い、先頭に跳ぶ時刻をポワソン過程に基づく確率法則で与えるという単純な1次元粒子系(多成分確率過程)である。先頭に跳ぶ規則は20世紀半ばに既に知られていたが、大規模な順位付けの時間変化が即時的に可視化されるようになったのは21世紀の変わり目頃からの、インターネットと個人所有のネット対応機器の社会への急速な浸透に伴うオンライン小売業の急拡大と定着による。そのウェブページに商品の売り上げ順位をリアルタイムで表示する宣伝が定着した。大規模な順位付けの時間発展を記述することが社会的な意味を持つようになったのが今世紀の特徴であることから、確率順位付け模型の無限粒子極限(流体力学極限)の研究が各粒子の軌道の追跡(カオスの伝搬)を含めて応用面からも今日的な意義を持つこととなった。

(2) 確率順位付け模型を重要と考えてその研究に深入りする理論上の理由はくりこみ群の理論で言う普遍性にある。流行度に従って順位付けする際、流行からもっとも遠い最下位側の順位付けは先頭に跳ぶ規則以外にないことが直感できる。すなわち、どのようなアルゴリズムを用いて上位側の順位を定義しようとも、下位の側は先頭に跳ぶ規則、したがって、確率順位付け模型が良い数理モデルのはずであり、下位が大きな部分を占める通常の世界現象に由来する強度分布の下では適切な極限で流行度の粒子系模型はすべて確率順位付け模型に漸近するはずである。平たく言えば、流行度の順位付けという今日的な現象のもっとも単純で典型的な数理模型としての基本的な役割を確率順位付け模型が持つと確信する。普遍性という考え方はくりこみ群と呼ばれる20世紀の理論物理学の概念であるが、この数学的意味を探る研究は上記科研費のもう1つ前に頂いた科研費基盤研究(B)「くりこみ群に基づく新しい確率解析の構築のための基盤研究」(2005~2008年度)およびそれ以前に頂いてきた一連の科研費基盤研究の長い研究の上に着想したものである。書き方を変えると、確率順位付け模型のように数学的に単純化した模型によって得られる数学的抽象度の高い繊細な性質は、現実の現象では複雑な要因に攪乱されて観測困難なことが通例だが、普遍性によってそのような単純な模型の性質が現実の現象でも観測される確信を持った。

(3) 先に流行度の順位付けという応用上の特徴を書いたが、数学的には非対称マルコフ過程の典型例という側面がある。自明でない確率過程の典型的な例として対称マルコフ過程が知られる。ディリクレ形式と呼ばれる手法が扱う系の複雑さに関係なく適用可能で、精密な結論が得られる。他方、確率順位付け模型を統計力学の言葉で書くと、1次元的な流れの各点から蒸発が起こり、蒸発した分だけ上流から下流への流れが起きる系(最上流は蒸発で失われた分がそのまま供給されることで粒子量が保存されるのが先頭に跳ぶ規則)なので、詳細釣り合いの原理が成り立たず、したがって非対称マルコフ過程であり、マルコフ過程という数学的に簡単な確率過程なのにディリクレ形式が適用できず、現代の確率論ではモデル(方程式)固有の手法でしか詳細な性質が得られない。このように、確率順位付け模型は粒子系としては単純であるいっぽう、数学についても応用についても研究を進める現代的な意義がある基礎的な確率模型であると考えられる。

(4) 前述の前年度までの科研費研究課題において、先頭に跳ぶ時刻を定める点過程の強度が位置依存性を持たない場合は無限粒子極限を証明していた。流れが他の粒子の蒸発に基づくので粒子間には確率過程として従属性があるが、先頭に跳ぶ時刻は定義により独立なので、空間的な分布関数への変数変換によって独立確率過程の大数の強法則の問題に帰着する。先述のとおり、非対称マルコフ過程の例として固有の方法によって切り抜ける状況であるが、単純な内容である。強度関数が時刻依存性を持って以上の場合も同様である。より詳細な研究に深入りする価値があるとの見通しの下に、自然な方向として先頭に跳ぶ確率法則についての強度が(時刻依存性だけでなく)位置依存性を持つ場合への拡張を考える。このとき、先頭に跳ぶ確率がその時の順位を通じて他の粒子に従属する確率過程となるので、位置依存性を持たない場合に対する非自明な摂動である。これが本科研費研究課題の入口である。時刻依存性と位置依存性を両方含めるといふ素朴な動機他に、位置依存性は、順位が上位にあることが宣伝効果を持つかという現実的に自然な問をこの模型で分析するために必要となる。

(5) 強度が位置依存性を持つ場合もすでに直前の科研費研究課題の最後に取りかかっていたが、強度の粒子についての分布が有界などの限定された条件の下で、かつ、位置強度結合経験分布の収束の位相も自然な弱収束位相よりも強い分布の変動ノルムによるもので、たとえば、応用上の出発点であるウェブのランキングに基づくオンライン小売りのロングテール構造の解析ではパレート分布という強度が非有界な連続分布が自然な対象だが、この枠組には入らない。数学的にも、強度分布の有界性が必要な本質的理由は見当たらず、また、強い位相は確率的揺らぎが同一種類の強度を持つ粒子間のみで消える大数の強法則を意味するので、異なる種類の粒子の揺らぎが打ち消すという、確率順位付け模型の本来の(強度が位置依存性を持たない場合に証明できている)描像よりも狭い原理の証明に過ぎなかった。この不自然な仮定を取り除くことが目標となる。

2. 研究の目的

先頭へ跳ぶ確率法則を定める点過程の強度関数が時刻だけでなく位置依存性も持つ確率順位付けモデルの流体力学極限、すなわち、位置と強度関数の結合経験分布の粒子数についての大数の法則とカオスの伝搬すなわち指定した粒子の組の極限が独立な確率過程に収束することの証明および極限確率過程の決定を、強度分布と経験分布の収束の位相について自然な仮定の下で行う。ここで自然な仮定とは、位置依存性を持たない場合に証明できていたクラス、特に、強度が上に有界でない場合で、強度分布の位相は、大数の法則として異なる強度関数の粒子間の確率的揺らぎの打ち消しを許す、弱収束位相を考える。

3. 研究の方法

(1) 確率順位付けモデルは、粒子1つに1つの確率過程が対応する多粒子系で、順位をその値とし、各粒子に対して独立なポワソン過程に基づいて先頭に跳ぶ(確率過程の値が1になる)ことと、各粒子は自分より下位の粒子が先頭に跳ぶときに順位を1ずつ下げる(値が1増える)ことだけで値が時間的に変化する。ポワソン過程の強度は時刻依存性を許し位置依存性を持たない場合は粒子間で独立である。粒子系としては自分より下位の粒子が先頭に跳ぶ場合のみ順位を下げることで従属性を持つ。確率積分を用いることで、強度が位置依存性を持つ場合も定式化可能で、この場合は先頭に跳ぶ確率過程が強度関数と順位を表す確率変数の合成関数がある。(前もって答を見通していたわけではなく、研究の結果わかったことだが、)強度の位置依存性は、位置依存性を持たない場合に比べて本質的な摂動であり、極限は位置依存性を持たない確率順位付けモデルの極限の簡単な書き換えではない。そこで、まず極限を記述する確率過程を決定する。

(2) 証明上の中間的な補助モデルとして、極限で現れる確率過程に従って時間発展する「位置依存性を持たない」確率順位付けモデルを構想できる。これを流れが定める強度に従う確率順位付けモデルと名付ける。より具体的には、粒子の軌道を固定し、その軌道と位置依存性を持つ強度の合成関数を強度とする確率順位付けモデルを考える(世間で知られる流体力学極限は対称マルコフ過程なので非対称マルコフ過程である確率順位付けモデルとは数学的に全く異なる証明だが、比喩としては)局所平衡に対応する中間モデルである。流れが定める強度に従う確率順位付けモデルは位置依存性を持たない場合と同様の意味で独立確率変数(過程)列の大数の強法則が成り立つことが期待できる。ただし、これを本来のモデルと極限において比較する必要があることから、強い収束を証明できるほど最後の評価が容易になることに注意する。

(3) 流れが定める強度に従う確率順位付けモデルでは、名称のとおり、流れの全体像が1つのパラメータである。元のモデルと同じ極限となる流れを見出す。

(4) 上で見出した正しい流れに従う確率順位付けモデルと本来の確率順位付けモデルが同じ極限に収束することの証明を行う。具体的には、Gronwall型不等式の多変数多段階への一般化を目指す。

4. 研究成果

研究目的の項で掲げた問題を研究の方法の項で掲げた手順で解決した。すなわち、現実的なロングテールの分析において自然なクラスでもある、強度の経験分布が非有界な連続分布に弱収束位相で収束する場合について、位置に依存する強度を持つ確率順位付けモデルの流体力学極限、すなわち、位置強度結合経験分布の収束とカオスの伝搬を証明した。この一般化は、極限を記述する点過程の強度が直前の到着時刻に依存する非独立増分であるなど、強度が位置依存性を持たない場合の非自明な摂動問題である。背景の項で記したとおり、確率順位付けモデルはネット書店の売上順位の時間変化のモデルで、ロングテール構造の数理解析の基礎モデルとなる。強度の位置依存性は好順位の注目効果を意味する。

(1) 極限粒子の時間発展を記述するのは強度が直前の到着時刻に依存する点過程と名付けた、ポワソン過程の非自明な独立増分性を持たない一般化である。その分布は級数でのみ書かれる複雑だがあらわな公式を持ち、樋に海水を流す塩田の流れの方程式というべき、位置依存性を持つ蒸発率を持つ混合液体の蒸発だけで生じる1次元流に従う非局所項(積分項)を持つ偏微分方程式で特徴付けられる。ポワソン過程の非独立増分だが良い性質を持ちうる一般化として将来役割を持つことを期待する。

(2) 流れに従う確率順位付けモデルについての独立確率過程の大数の強法則は、関数値独立確率変数列の2重に一樣な大数の完全法則と名付けた強い収束定理の形に一般化して解決した。バナッハ空間値独立確率変数の大数の強法則などの研究につながる可能性を期待する。

(3) 流れが定める強度に従う確率順位付けモデルで元のモデルと同じ極限となる流れは、パラメータの流れから得られるこのモデルの時間発展が定める流れを対応させることを、流れの集合上の写像とみると、縮小写像であることが証明できた。その不動点として唯一決まる流れが求めるものである。

(4) 研究の一部は前回の科研費研究期間と本研究課題の次の科研費研究期間にまたがる。着想が形になるまでの試行錯誤や証明が完成してから論文採択までの時間差等があるので、これはむしろ自然である。本研究は良く言えば高度に特色があって世界的に類似の進展を目指す研究のない、他の追従を許さない独創的研究と自負する。逆に言えば、世界に基礎研究に対する余裕が小さい時代には、選択や集中から漏れる側かもしれない。そのような研究を完成度を高め

て後世に残すことも時代に巡り合わせた研究者の役割であると考え、論文を採択してくれる学術誌が見つかりにくいことがこのような研究の障害となるが、自分のウェブページに置いたプレプリントを逐次更新し、また、ウェブページに置いた確率順位付け模型についての一般的な解説に最新の研究成果の概説を和英両文で追加するなど、掲載に前向きな学術誌を見出す間も成果開示は進めてきた。この他、適宜確率論シンポジウム等で成果を発表してきた。

5. 主な発表論文等

〔雑誌論文〕(計 2 件)

T. Hattori, Point process with last-arrival-time dependent intensity and 1-dimensional incompressible fluid system with evaporation, *Funkcialaj Ekvacioj*, 査読有, **60** (2017) 171-212. DOI: 10.1619/fesi.60.171

T. Hattori, Open problems to an infinite system of quasi-linear partial differential equations with non-local terms, *Symposium on Probability Theory, Kyoto*, 2014/12/16-19, *RIMS Kokyuroku*, 査読無, **1952** (2015) 9-16. URL: <http://www.kurims.kyoto-u.ac.jp/~kyodo/kokyuroku/contents/1952.html>

〔学会発表〕(計 5 件)

服部哲弥, Khintchine の不等式・大数の完全法則・Glivenko-Cantelli の定理, 岡山-広島 解析・確率論セミナー 2018, 2018 年.

服部哲弥, 位置依存確率順位付け模型, 岡山確率論セミナー, 2016 年.

服部哲弥, 独立確率過程の大数の強法則と流れが定める確率順位付け模型, OR 学会待ち行列研究部会例会, 2015 年.

服部哲弥, 独立確率過程の大数の強法則について, 東京確率論セミナー, 2015 年.

服部哲弥, 非局所項を持つ 1 階準線形偏微分方程式の解の独立増分でない点過程による表現, 確率論シンポジウム, 2014 年.

〔その他〕

(研究代表者日本語ホーム)

<http://web.econ.keio.ac.jp/staff/hattori/hattori.htm>

(研究代表者英語ホーム, 含論文へのリンク)

<http://web.econ.keio.ac.jp/staff/hattori/research.htm>

(確率順位付け模型初等解説)

<http://web.econ.keio.ac.jp/staff/hattori/amazonj.htm>

(確率順位付け模型英語初等解説)

<http://web.econ.keio.ac.jp/staff/hattori/amazone.htm>

(研究代表者確率順位付け模型関連原著学術論文へのリンク一覧)

<http://web.econ.keio.ac.jp/staff/hattori/liamzn.htm>