

令和 6 年 6 月 14 日現在

機関番号：62603

研究種目：挑戦的研究（萌芽）

研究期間：2021～2023

課題番号：21K18589

研究課題名（和文）スケーラブルなベイズ計算法の解析

研究課題名（英文）Analysis for scalable Bayesian calculations

研究代表者

鎌谷 研吾（Kamatani, Kengo）

統計数理研究所・統計基盤数理研究系・教授

研究者番号：00569767

交付決定額（研究期間全体）：（直接経費） 3,000,000円

研究成果の概要（和文）：確率過程を用いた手法である、区分確定的マルコフ過程を用いた新しい頑健な手法の実装技術の開発と、理論的な性質を研究し、研究は継続中である。また、この離散時間版ともいうべき手法の研究も行った。近年、多くの手法が局所的な情報を利用した決定論的提案を用いるが、頑健性に欠ける。一方で、頑健な既存手法は局所的な情報を入れにくい。本研究では、両者の良い部分を取った手法、Haar-Weave-Metropolisカーネルを作成し、数値実験で有効サンプルサイズと平均二乗ジャンプ距離で他の方法に優れていることを示した。

研究成果の学術的意義や社会的意義

本研究課題では、近年注目されている確率過程の生成を用いる手法の技術的課題であった確率過程の効率的な生成を目的としていた。国際共同研究により、技術的困難を軽減することができた。ベイズ統計学の分野では、マルコフ連鎖モンテカルロ法が依然として主流であり、この手法が苦手とする部分は、そのままベイズ統計学の適用が困難な部分でもあった。本研究の技術的な発展により、従来手法で困難とされていた領域をさらに狭めることができると期待している。

研究成果の概要（英文）：We are developing implementation techniques for a new robust method using piecewise deterministic Markov processes and studying its theoretical properties, with research still ongoing. Meanwhile, we also explored the Haar-Weave-Metropolis method, which can be considered a discrete-time version. Recently, many methods have used deterministic proposals based on local information but lack robustness. Conversely, existing robust methods are difficult to incorporate local information into. In our study, we developed the Haar-Weave-Metropolis kernel, which combines the strengths of both approaches, and demonstrated its superiority in terms of effective sample size and mean squared jump distance in numerical experiments.

研究分野：ベイズ計算

キーワード：マルコフ連鎖 モンテカルロ法 ベイズ統計学 確率過程 ハミルトニアン

1. 研究開始当初の背景

本研究の背景を述べる。

- Amazon や Google のような企業は、**情報戦略**で大規模データを解析し、価値を創出している。
- スタンフォード大学のブラッドリー・エフロン教授は**情報科学の二つの側面**を指摘している。産業界は統計計算（アルゴリズム）開発に注力し、統計モデルの解釈は軽視され、結論が不明確になるリスクがある。
- **ベイズ統計学**は明瞭な結果を提供するが、統計計算の構成が困難。1990年代に統計計算の革新である、**マルコフ連鎖モンテカルロ法**で進展したベイズ統計学であるが、現代のデータ潮流に取り残されつつある。
- データ解釈が重要視される中、本研究はスケーラブルな統計計算を探求する。ベイズ統計でもスケーラブルなアルゴリズムが注目され、伝統的統計計算から脱却しようとしているが、代わりに解釈の明瞭さを犠牲にしていた。しかし、最近の**逐次確定的マルコフ過程**や**拡散過程**の手法により、解釈の明瞭さを犠牲にしないスケーラブルな計算法が実現され、注目されている。

2. 研究の目的

本研究の目的を述べる。

- 明瞭な意味を保つスケーラブルアルゴリズムは今後のトレンドだろう。しかし、重大な問題が一つある。構成に**確率過程の乱数の生成**を用いており、**技術的に困難**なことだ。
- 本研究では技術的困難の解消を目的とする。確率過程の生成部分がスケーラブルな手法の**足かせ**になっている。確率過程生成の難しさを解消しなければ、明瞭な意味を保つスケーラブルな手法の発展はない。

3. 研究の方法

- 本研究では**確率過程の技術が鍵**となる。オンラインによる国内外の交流によって技術を学び、技術開発につなげる。
- しかし、確率過程やベイズ統計学の枠にとらわれる必要はない。**数値最適化**、**微分方程式の数値解法**と広い意味の**確率論**の研究者とも積極的な意見交換を行いたい。今後につながる国内でのネットワークの作成をする。
- コロナウィルス下でも、国際研究の重要性は変わらない。国際共同研究を行いたい。

4. 研究成果

- **現在進行中の研究の離散時間版** Haar-Weave-Metropolis kernel [1]: 最近、ハミルトンモンテカルロ法に触発された決定論的提案を用いた多くのマルコフ連鎖モンテカルロ (MCMC) 法が開発されている。これらの決定論的提案は、目標分布の局所情報 (勾配など) を利用することで、しばしば効率的なサンプリングを可能にする。しかし、エルゴード理論が示すように、これらの決定論的提案法はロバスト性に欠け、特に重い裾を持つ目標分布に対しては収束が悪いことが多い。一方、ハール測度を用いたマルコフカーネルは、目標分布に関するグローバルな情報を学習するため、比較的ロバストである

が、提案定常分布の密度保存条件が必要であり、多くの決定論的提案はこの条件を破る。この論文では、提案定常分布の密度関数の値を維持する決定論的変換を慎重に選択し、決定論的変換を使用してマルコフカーネル “Weave-Metropolis カーネル” を作成した。また、ハール測度と組み合わせることで “Haar-Weave-Metropolis カーネル” を導入した。このようにして、マルコフカーネルは決定論的提案を用いて目標分布の局所情報を利用し、ハール測度によって目標分布のグローバル情報も利用できる。数値実験においても、提案手法が他の方法と比較して有効サンプルサイズと平均二乗ジャンプ距離の面で優れていることを示した。

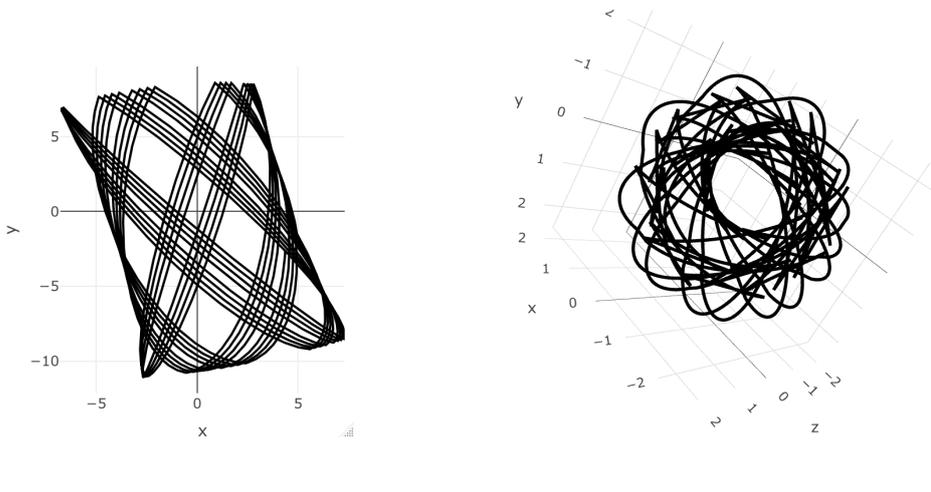


図1 左と右の両方のパネルは、ステップ数 $L = 40$ の、“Weave-Metropolis カーネル” で使われる特徴的な決定論的変換の軌跡を示す。左のパネルは2次元のスチューデント t 分布を対象としており、右のパネルは3次元のスチューデント t 分布を対象としている。

- **確率過程の乱数生成の研究：現在進行中** Haar Boomerang sampler : [1] の区分確定的マルコフ過程版とも捉えられる本論文では、様々な区分確定的マルコフ過程の実装の工夫を研究した。適合的なサンプリングを実装することで、確率過程の生成の困難さを軽減し、実用に耐えうる技術を開発した。また、理論面においては、従来の区分確定的マルコフ過程が裾の重い分布に対してうまく機能しなかったのに対し、Haar Boomerang sampler は広い確率分布のクラスで指数エルゴード性を持つ。国際共同研究ですすめている。
- **ネットワーク生成** 2023年からセミナーシリーズを開始し、2023年度は国内外の研究者合わせて9名の発表者があった。

<https://sites.google.com/view/bayescmpjp/>

参考文献

- [1] Kengo KAMATANI and Xiaolin SONG. Haar-weave-metropolis kernel. *Bulletin of informatics and cybernetics*, 54:1–31, 2022.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 KAMATANI Kengo, SONG Xiaolin	4. 巻 54
2. 論文標題 HAAR-WEAVE-METROPOLIS KERNEL	5. 発行年 2022年
3. 雑誌名 Bulletin of informatics and cybernetics	6. 最初と最後の頁 1~31
掲載論文のDOI（デジタルオブジェクト識別子） 10.5109/4755997	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計5件（うち招待講演 3件/うち国際学会 5件）

1. 発表者名 Kengo Kamatani
2. 発表標題 Non-reversible guided Metropolis kernel
3. 学会等名 Unification algorithmique d'analyses statistiques multiples Computational methods for unifying multiple statistical analyses (Fusion)（招待講演）（国際学会）
4. 発表年 2022年

1. 発表者名 Kengo Kamatani
2. 発表標題 Scaling limit of Markov chain/process Monte Carlo methods
3. 学会等名 IASC-ARS 2022（招待講演）（国際学会）
4. 発表年 2022年

1. 発表者名 Kengo Kamatani
2. 発表標題 High-dimensional asymptotics using multiscale analysis
3. 学会等名 MCM 2021（国際学会）
4. 発表年 2021年

1. 発表者名 Kengo Kamatani
2. 発表標題 Haar-Weave-Metropolis Kernel
3. 学会等名 ISI World Statistics Congress, 2023 (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 Kengo Kamatani
2. 発表標題 Scaling of Piecewise Deterministic Monte Carlo for Anisotropic Targets
3. 学会等名 CMStatistics 2023 (国際学会)
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

Kengo Kamatani https://sites.google.com/view/kengokamatani/home Seminar website https://sites.google.com/view/bayescompjp/

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
連携研究者	仲北 祥悟 (Nakakita Shogo) (80855114)	東京大学・大学院総合文化研究科・特任助教 (12601)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
フランス	Paris Dauphine University			