

科学研究費助成事業 研究成果報告書

平成 28 年 10 月 19 日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2012～2015

課題番号：24740062

研究課題名(和文) 隠れマルコフモデルに対する統計計算手法の漸近的解析

研究課題名(英文) Asymptotic analysis of statistical computation methods for hidden Markov models

研究代表者

鎌谷 研吾 (Kamatani, Kengo)

大阪大学・基礎工学研究科・講師

研究者番号：00569767

交付決定額(研究期間全体)：(直接経費) 2,700,000円

研究成果の概要(和文)：隠れマルコフモデルのための逐次モンテカルロ法の解析と、高次元で有効なマルコフ連鎖モンテカルロ(MCMC)法の解析を行った。前者については、近年のホットトピックである高次元で有効な手法の提案、およびマルチレベルモンテカルロ法との組み合わせを行った。後者では高次元解析、とくに複雑なベイズモデルでの推定を可能にするモンテカルロ法を提案した。従来手法を大幅に改善することを数値的にも理論的にも示すことができた。

研究成果の概要(英文)：For the project, I performed analysis on (a) sequential monte carlo methods and (b) markov chain monte carlo (MCMC) methods for high-dimensional complicated models. For (a), we proposed (a-1) efficient strategy for high-dimensional state space models, and (a-2) ensemble strategy with multi-level monte carlo. For (b) we proposed a scale-free MCMC and analysed its performance via high-dimensional asymptotic and ergodicity analysis.

研究分野：統計学

キーワード：ベイズ統計 モンテカルロ 漸近理論 高次元解析 複雑モデル 大規模データ

1. 研究開始当初の背景

高次元や高頻度の所謂大規模データを扱える環境が整い、今までにない困難な状況での統計解析が必要になる場面が増えた。尤度が陽にかけない場合や、パラメータの次元が極端に多い場合がその例である。モンテカルロ法もそれに合わせて新しいアイデアが次々生み出されている。逐次モンテカルロ法とマルコフ連鎖モンテカルロ法の組み合わせや

Approximate Bayesian Computationを利用した手法など、モンテカルロ法はより複雑に発展している。理論的な道具も計算機の知識もより発展的内容が必要になってきている。

2. 研究の目的

高次元や複雑なモデルでのモンテカルロ法の解析を行う。とくに隠れマルコフモデルに対する統計計算手法の解析を行う。隠れマルコフモデルは応用例に富むが、ノイズの影響による非正則性が、統計問題を解析的にも計算的にも複雑にしている。計算手法の漸近理論を見るとという私の既存研究を拡張して、隠れマルコフモデルの計算手法の解析を行う。あわせて数値的解析も行う必要がある。応用として遺伝子解析におけるコピーナンバー多型の解析に適用する。

3. 研究の方法

マルコフ連鎖モンテカルロ法の解析については、従来の繰り返し回数の極限を見るエルゴード性解析に加えて、次元などキーとなる統計量の極限を見る解析を行う。こうした解析は、古典的な確率過程における極限理論の一つの応用例とみなすことができるため、技術的道具は豊富である。逐次モンテカルロ法の解析、とくに実際の計算については研究協力者の協力が不可欠である。本分野は計算機環境の発展や大規模データのトレンドで大きく変化しうる。積極的に国際学会へ参加し、発表、情報収集を行う。

4. 研究成果

- (1) 逐次モンテカルロ法の解析は本研究課題の第一のテーマである。一つ目の

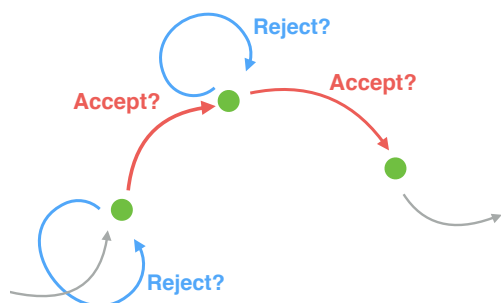
結果として、近年のホットトピックである、高次元で有効な手法の提案を行った。通常の逐次モンテカルロ法は高次元を苦手とするが、本研究で提案した手法は、通常のステップに加え、高次元空間を細分して逐次モンテカルロ法を行う手法である。そのパフォーマンスのよさは数値計算で調べられ、また理論的な保証も行った。

- (2) 逐次モンテカルロ法の二つ目としてマルチレベルモンテカルロ法との組み合わせを行い、新しい手法を提案した。マルチレベルモンテカルロ法は、計算したい対象を層に分け、最初に大雑把な計算を大規模に行い、残りの精密計算の回数を少なくすることで効率的に積分計算を行う手法である。マルチレベルモンテカルロ法と逐次モンテカルロ法(フィルタリング)の組み合わせは難問であったが特徴的なカップリングで達成した。後者は最終年度の結果である。これらの結果はシンガポール、イギリス、アメリカの研究者との共同研究で得ることができた。

- (3) 高次元で有効なマルコフ連鎖モンテカルロ(MCMC)法の発見は本研究課題の主要な結果の一つである。通常のモンテカルロ法は高次元では著しくそのパフォーマンスを落とす。とくに裾が重い場合など複雑な分布の近似問題ではその傾向は顕著であった。本研究では、MCMC法の提案カーネルの不変分布に着目し、不変分布が事後分布よりも裾が重くなるように工夫を施した。

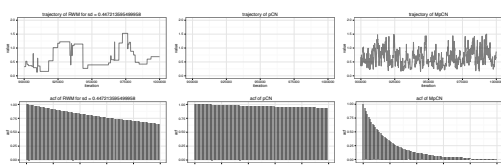
MCMC法の多くは、下記のように点列を提案し、それを採択(Accept)もし

くは棄却(Reject)をする手続きを踏む。本研究の提案手法では、現時点の状況を見ながら、大域的(なるべく今の地点より遠くに)な提案をする提案カーネルをデザインした。



この工夫によって頑健な手法を構成できた。手法のよさはシミュレーションだけでなく、高次元漸近論で示すことができた。加えて最終年度にはエルゴード性でも他の手法を凌駕する収束性を持つことを示した。本手法はまだ未解明な部分もあり、海外の研究者と協力しつつ、今後も引き続き研究を行う予定である。

以下、後者のMCMCの研究に関して、提案手法と従来手法の比較として、離散観測された確率過程のベイズ統計量の計算の例を紹介する。左の二列の図は従来手法、右一列が提案手法である。各図の上部の図はMCMC法によるマルコフ連鎖の経路を表し、下部の図はその自己相関である。自己相関はx軸の値(lagと呼ばれる)が大きくなるに従って早く0へ収束する方が良い。提案手法のパフォーマンスの良さが際立っている。なお、中央列のマルコフ連鎖の経路は欄外にはみ出している。



これらの研究は当初想定していた以上に発展した。とくにMCMC法の研究は新手法の導入に加え、二種類の理論的整備を行えた。一方で遺伝統計学への実応用についてはある程度の進展を見たが、残念ながら大きな成果を残すことはできなかった。遺伝統計学への応用については引き続き大阪大学医学系研究科や理化学研究所等と連携し研究を進めていきたい。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 9 件)

- ① Kengo Kamatani, Masayuki Uchida “Hybrid multi-step estimators for stochastic differential equations based on sampled data”, *Statistical Inference for Stochastic Processes*, 査読有, 18, 2015, 177-204, 10.1007/s11203-014-9107-4

[学会発表] (計 39 件)

- ① Kengo Kamatani, High-dimensional asymptotic properties of Markov chain Monte Carlo methods for heavy-tailed target distributions, *Statistics for Stochastic Processes and Analysis of High Frequency Data V*, 2016年3月23日, University Pierre and Marie Curie (Paris 6), Paris, France
- ② 鎌谷研吾, マルコフ連鎖のエルゴード性と regular variation, 日本統計学会春季集会, 2016年3月5日, 東北大学川内南キャンパス, 仙台, 宮城
- ③ Kengo Kamatani, Markov chain

Monte Carlo in high-dimension
with heavy-tailed target
probability distributions, MCMski
V, 2016年1月6日, Lenzerheide,
Switzerland

- ④ Kengo Kamatani, Efficient strategy
for the Markov chain Monte Carlo
in high-dimension and its
implementation, IASC-ARS, 2015
年12月17日, Singapore
- ⑤ Kengo Kamatani, On some ergodic
properties of the MpCN algorithm,
ERCIM 2015, 2015年12月13日,
University of London, London

[その他]

<http://www.sigmath.es.osaka-u.ac.jp/~kamatani/research/>

6. 研究組織

(1) 研究代表者

鎌谷 研吾 (KAMATANI KENGO)

大阪大学 ・ 大学院基礎工学研究科 ・ 講師

研究者番号 : 00569767