

令和 6 年 6 月 15 日現在

機関番号：32704

研究種目：若手研究

研究期間：2021～2023

課題番号：21K13273

研究課題名（和文）時間情報を持つ大規模空間データのための高速統計解析

研究課題名（英文）Fast statistical analysis for large spatial datasets with temporal information

研究代表者

平野 敏弘（Hirano, Toshihiro）

関東学院大学・経済学部・准教授

研究者番号：10816010

交付決定額（研究期間全体）：（直接経費） 1,200,000円

研究成果の概要（和文）：本研究課題では、主に以下の2点の成果を得た。
第一に、大規模時空間データに対するカルマンフィルタに前回採択された科研費研究課題で提案した「線形射影を用いた多重解像度近似」を適用し高速化した「線形射影を用いた多重解像フィルタ」を提案した。さらに、提案手法を非線形・非ガウス状態空間モデルの場合に拡張した。第二に、大規模空間データに対して「線形射影を用いた多重解像度近似」で高速に作成した予測平面に不連続性が発生した場合について、covariance taperingを用いた解消方法を提案し、実データ分析による評価を行った。
また、その他の空間統計学に関連する研究も行った。

研究成果の学術的意義や社会的意義

本研究課題において主に得られた研究成果は、大規模空間データや大規模時空間データに対する新しい高速統計解析手法を提案したものであり、そのうちの一部は統計学における国際学術誌に投稿中である。特に、1つ目の研究成果である「線形射影を用いた多重解像フィルタ」は、時空間データが非ガウス性や非定常性を持つ場合にも適用可能であるだけでなく、大規模空間データが時々刻々と観測される状況でもリアルタイムで高速に統計解析を実行できる。提案手法は地価、交通量、人工衛星から得られる大気中の水蒸気量といった時空間データの統計解析に有用であり、不動産市場の分析、交通渋滞の削減、気候変動問題といった応用と結びついている。

研究成果の概要（英文）：In this program, I mainly conducted the following two research topics in the fast statistical analysis for large spatial or spatio-temporal datasets. First, I proposed a new fast computation method of the Kalman filter for large spatio-temporal datasets, which was referred to as a multi-resolution filter via linear projection, by using a multi-resolution approximation via linear projection which was proposed in the previous Grant-in-Aid for Early-Career Scientists. Furthermore, I extended the proposed method to nonlinear and non-Gaussian state-space models. Second, I considered a modification of the multi-resolution approximation via linear projection by using the covariance tapering, which can resolve the artificiality in the prediction surface. I evaluated the modified method through the real data analysis. I also conducted other researches related to spatial statistics.

研究分野：空間統計学

キーワード：確率場 大規模空間データ 多重解像度近似 時空間データ 状態空間モデル カルマンフィルタ データサイエンス

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

全球測位衛星システムなどの計測技術の発展により、センサーを通じて位置情報を伴う大規模な空間データが時々刻々と観測されている。たとえば、交通網では道路に設置されたセンサーから単位時間あたりの車両数を表す交通量が得られている。このようなデータに対して空間統計学・時空間統計学における統計解析手法を適用することで、空間相関構造の推定、未観測地点における空間データの予測、各時点における予測平面の作成を実行することができ、渋滞予測や極端な天候現象の予測が可能となる。

2. 研究の目的

空間データに対して最尤法、クリギング、マルコフ連鎖モンテカルロ法(MCMC)などの統計解析手法を適用した場合、計算途中で現れる逆行列に関連する計算はサンプルサイズの3乗のオーダーの計算量となる。そのため、空間データが大量に得られた場合、計算負荷は極めて大きくなる。また、時空間データに対してカルマンフィルタを適用した場合、各時点における空間データのサンプルサイズが非常に大きいときにも同様の問題が発生する。このような問題に対処するため、本研究課題では大規模な空間データが得られた場合の高速統計解析手法の提案を目的とする。

3. 研究の方法

実際に観測される大規模空間データは時間情報、高次元性、非定常性、非正規性を持つ場合が多い。このような性質を持つ空間データに対して、先行研究で提案された大規模空間データに対する高速統計解析を適用した場合の推定・予測精度はまだ改善の余地がある。本研究は、応募者が過去に執筆した学術論文や学会発表での成果を拡張することで、時間情報、高次元性、非定常性、非正規性を持つ大規模空間データに対する新しい高速統計解析手法の提案とその理論的性質の導出を目標とする。

4. 研究成果

本研究課題では、以下にあげる研究成果が得られた。

(1) 線形射影を用いた多重解像フィルタの提案

本研究では、Hirano (2021)で提案された線形射影を用いた多重解像度近似を大規模時空間データに対するカルマンフィルタに適用して高速計算する線形射影を用いた多重解像フィルタを提案した。具体的には、カルマンフィルタにおける予測共分散行列に線形射影を用いた多重解像度近似を適用し、フィルタリング平均を高速計算するアルゴリズムを導出した。さらに、テイラー展開による1次近似とラプラス近似を用いることで、提案アルゴリズムを非線形・非ガウス状態空間モデルの場合に拡張した。提案アルゴリズムは、グラフ理論の性質よりアルゴリズム中の特定の行列のブロックスパース構造が時間を通じて保持されているため、高速化前のアルゴリズムより計算量が小さくなっている。様々な設定の下でシミュレーションを行い、フィルタリング平均の近似精度や計算時間に基づいて最近の先行研究や代表的な手法であるアンサンブルカルマンフィルタと比較した。結果として、提案したアルゴリズムは線形射影に関連する部分の計算が必要であるため、計算時間が先行研究より若干かかるが、フィルタリング平均の近似精度は改善している。また、大気中の水蒸気量を用いた実データ分析において、提案したアルゴリズムはシミュレーションと同様に先行研究より良好な結果を示した。これは、提案手法は先行研究よりチューニングパラメータが多い分、柔軟な調整が可能となるため、計算時間を削減しつつ高い近似精度が実現できるからである。得られた研究成果について、国内学会・国際学会・国内研究会で発表するとともに、Hirano and Ishihara (2024)で論文にまとめて国際学術誌に投稿した。投稿論文は現在査読中である。

(2) 線形射影と covariance tapering を用いた多重解像度近似の提案

本研究では、Hirano (2021)で提案された線形射影を用いた多重解像度近似に対して、線形射影だけでなく covariance tapering も適用することにより、大規模空間データに対して高速に統計解析が可能であるだけでなく、線形射影を用いた多重解像度近似において発生する可能性がある予測平面の不連続性を解消できるアルゴリズムを提案した。具体的には、提案アルゴリズムは、正規確率場における共分散関数のパラメータの最尤推定量と未観測地点の予測分布を高速計算できる。提案したアルゴリズムの理論的性質を導出するとともに、実データ分析において性能評価を行った。結果として、線形射影を用いた多重解像度近似により作成した予測平面は分割領域の境界付近で不自然な変化を示したが、提案アルゴリズムはこれを解消するとともに線

形射影を用いた多重解像度近似と同等の推定精度・予測精度・計算効率性を示した。当該研究成果について国内学会・国際学会で発表した。現在、得られた研究成果を論文としてまとめて投稿予定である。

(3) 非正規確率場における covariance tapering を用いた最良線形不偏予測量の漸近有効性

ここでの研究成果は、「5. 主な発表論文等〔雑誌論文〕」における1つ目の論文に対応している。本研究では、Hirano and Yajima (2013)における理論的結果を、ある設定の下では弱い条件で証明できることを確認した。具体的には、エルミート多項式を用いた変換によって表現される非正規確率場において、変換前の確率場が指数型共分散関数を持つ場合、Hirano and Yajima (2013)より弱い条件で、高速計算可能な covariance tapering を用いた最良線形不偏予測量がオリジナルの最良線形不偏予測量だけでなく最適な非線形予測量に関しても漸近有効であることを証明した。

<参考文献>

Hirano, T. (2021). A multi-resolution approximation via linear projection for large spatial datasets. *Japanese Journal of Statistics and Data Science*, 4, 215-256.

Hirano, T. and Ishihara, T. (2024). Multi-resolution filters via linear projection for large spatio-temporal datasets. arXiv:2401.05315.

Hirano, T. and Yajima, Y. (2013). Covariance tapering for prediction of large spatial data sets in transformed random fields. *Annals of the Institute of Statistical Mathematics*, 65, 913-939.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 平野 敏弘	4. 巻 285
2. 論文標題 非正規確率場におけるCovariance Taperingを用いた最良線形不偏予測量の漸近有効性	5. 発行年 2022年
3. 雑誌名 経済系	6. 最初と最後の頁 38-46
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Hirano Toshihiro	4. 巻 4
2. 論文標題 A multi-resolution approximation via linear projection for large spatial datasets	5. 発行年 2021年
3. 雑誌名 Japanese Journal of Statistics and Data Science	6. 最初と最後の頁 215 ~ 256
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s42081-020-00092-x	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計8件（うち招待講演 1件 / うち国際学会 5件）

1. 発表者名 平野敏弘
2. 発表標題 非線形・非ガウス状態空間モデルに基づく大規模時空間データに対する高速フィルタリング
3. 学会等名 2023年度 統計関連学会連合大会
4. 発表年 2023年

1. 発表者名 Toshihiro Hirano
2. 発表標題 Multi-resolution filters via linear projection for large spatio-temporal datasets
3. 学会等名 Spatial Statistics 2023 (国際学会)
4. 発表年 2023年

1. 発表者名 Toshihiro Hirano
2. 発表標題 Multi-resolution filters via linear projection for large spatio-temporal datasets
3. 学会等名 ISM Symposium on Environmental Statistics 2023 (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 Toshihiro Hirano
2. 発表標題 A multi-resolution approximation by linear projection and covariance tapering for large spatial datasets
3. 学会等名 The 15th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 平野 敏弘
2. 発表標題 大規模時空間データに対するLinear Projectionを用いた多重解像フィルタ
3. 学会等名 2022年度 統計関連学会連合大会
4. 発表年 2022年

1. 発表者名 Toshihiro Hirano
2. 発表標題 A multi-resolution approximation via linear projection for large spatial datasets
3. 学会等名 The 14th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 平野 敏弘
2. 発表標題 Covariance TaperingとLinear Projectionを用いた多重解像度近似について
3. 学会等名 2021年度 統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 Toshihiro Hirano
2. 発表標題 A multi-resolution approximation via linear projection for large spatial datasets
3. 学会等名 The XV World Conference of the Spatial Econometrics Association (SEA 2021) (国際学会)
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
---------------------------	-----------------------	----

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
---------	---------