

令和 6 年 5 月 13 日現在

機関番号：12601

研究種目：若手研究

研究期間：2020～2023

課題番号：20K13468

研究課題名（和文）確率過程に対する一様・高次元正規近似法の開発と変数誤差モデルへの応用

研究課題名（英文）Development of uniform and high-dimensional Gaussian approximation for stochastic processes and their applications to errors-in-variable models

研究代表者

栗栖 大輔 (Kurisu, Daisuke)

東京大学・空間情報科学研究センター・准教授

研究者番号：70825835

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究課題を通じて、（1）非定常な空間データ・関数データに対するノンパラメトリック回帰、（2）高次元空間データ・時空間データのための統計的推測方法の開発（3）極値統計学を用いたノンパラメトリック分位点回帰、（4）観測誤差を伴って観測される変数が存在する場合のノンパラメトリックな確率密度の推定法開発

について取り組んだ。以上の研究成果は何れも国際ジャーナルに採択され、特に研究テーマ（1）、（2）については研究成果の一部は統計分野のトップジャーナルに掲載されており、今後は関連テーマへの応用やより発展的な課題について研究を行う予定である。

研究成果の学術的意義や社会的意義

本研究課題では研究テーマとして（1）時系列データ、（2）空間データ、（3）時空間データ、（4）関数時系列データ、（5）変数誤差モデルに関するノンパラメトリックな統計分析手法の開発に取り組んだ。（1）～（5）の各テーマにおいては、研究代表者の前研究課題において得られた一連の理論解析手法が上記の各テーマにおける重要な問題に適用可能であることが予想されていた。実際、本研究課題において提案したデータ分析手法の理論解析のアプローチは各テーマで共通する部分が多く、統一的な視点でこれらの問題の解決策を与えることに成功した。

研究成果の概要（英文）：In this research project, we worked on the following research topics: (1) Nonparametric regression for non-stationary spatial data and functional data, (2) Development of statistical inference methods for high-dimensional spatial data and spatio-temporal data, (3) Nonparametric quantile regression using extreme value theory, and (4) Development of nonparametric density estimation methods for variables observed with measurement errors. All these research outcomes have been accepted in international journals, with some parts of the results, particularly in research themes (1) and (2), published in top journals in the field of statistics. I am planning to further explore applications to related topics and tackle more advanced issues.

研究分野：時空間統計

キーワード：空間データ 時空間データ 関数データ ノンパラメトリック回帰 分位点回帰 変数誤差モデル 高次元中心極限定理 ブートストラップ法

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、Z - 19 (共通)

1. 研究開始当初の背景

研究開始当初の課題その背景は以下のとおりである。

研究 , : 変数誤差モデルは実社会における様々なデータ分析に用いられ、医療分野では複数の測定方法で得られたデータから患者の状態を正確に診断するためデータの互換性の評価に利用される。さらに、変数誤差モデルの数学的な解析手法がファイナンスや損害保険のリスク評価のモデルとして利用されるレヴィ過程と共通している部分が多いことから研究代表者の前研究課題の成果が利用できる。

研究 : 研究代表者の過去の研究(Kurisu(2019))では本研究と同様のデータ観測の枠組みで空間回帰モデルに対して平均関数と分散関数のノンパラメトリックな推定量を構成し、それらの信頼区間の構成法を与えている。高次元の空間過程に対してデータから計算される統計量の高次元ガウス近似の結果が与えられれば、Chernozhukov et al.(2017)で研究されているような高次元の中心極限定理を利用した漸近的な一様信頼バンドを構成することができ、Kurisu(2019)の結果を改良することができる。

2. 研究の目的

研究 , 変数誤差(measurement error, 以下 ME)モデルにおいて興味の対象となる変数(X)の確率分布のノンパラメトリックな推定量を構成し、その収束レートと一様信頼バンドを構成することを目指す。特に本研究では興味のある確率変数が観測誤差を伴って複数回観測される状況を考える。例えばある地域の住民の所得などの経済学的には景気を反映していると考えられるデータを集めるアンケート調査を行う際、虚偽の申告やデータの改ざんがないかといった統計データの信頼性の評価を行うため複数回アンケート調査を行う場合が該当する。理論的には観測誤差が対称な確率分布に従うか否かで興味のある変数(例えば所得)の確率分布の推定の難しさに違いが生じる。このような社会科学におけるサーベイデータの評価の例にとどまらず、複数回観測の ME モデルが適用可能な状況では調査によって得られたデータがモデルの理論的仮定に適合しているか、データが実際に母集団の特徴を正しく反映しているか客観的に評価することは実用上非常に重要な課題である。

研究 : Lahiri and Zhu(2006)などで考えられている、定常な空間過程(より一般には確率場)に対して、空間過程を空間上のデータの観測領域はデータが増えるにしたがって広がり、観測点は密になっていく漸近論の枠組みで考える。このような状況は各地域の降水量や地価のデータに代表されるような、不等間隔に配置された多くの地点でデータが利用可能な場合における各地点の間の関係性を統計的に分析する際によく利用されるデータ観測の枠組みである。

研究 では空間過程が α -mixing と呼ばれる観測値同士の従属性の仮定を満たすという仮定の下で、各観測地点で複数の変数(その地点における気温・風速・降水量など)が利用可能な場合の各変数の特徴量に対する信頼区間を同時に構成することを目指す。これにより例えばある地域における平均気温・平均風速などの信頼区間の構成や地域間の降水量の比較など、その地域の特徴を反映する複数の変数の地域間での比較を同時に行うことができるようになる。

3. 研究の方法

研究 1 : 具体的にまず(i)複数回観測される場合の ME モデルにおける X の確率密度を、観測誤差が対称・非対称に関わらず与えられた区間上で一様に推定する場合の収束レートを導出し、(ii)その結果を利用して X の累積分布関数の一様信頼バンドを構成する方法を新たに提案することを目指す。(i)については、レヴィ過程の特徴量であるレヴィ測度を特性関数を利用して(高頻度ではない)離散観測データから推定する方法を提案した Neumann and Reiss(2009)で用いられている経験過程の解析手法を拡張すれば一様収束レートの導出が可能である。この方法により(i)の先行研究である Li and Vuong(1998), Bonhomme and Robin(2010)よりも良い収束レートを導出した。 X の確率密度の推定量の推定レートの改善や最適な収束レートの導出は計量経済学において理論的にも重要な未解決問題の一つである。(ii)については本研究の設定のもとで X の分布関数の一様信頼バンドの構成方法を与えた先行研究は知られていない。従って(ii)により、理論的・実用的にも計量経済学における重要な問題の(一部)解決策を与えることができた。

さらに(ii)の累積分布関数に対応する結果を X の密度関数に対しても与えた ((iii))。これは統計分析の目的に応じて分布関数と密度関数のどちらを使うのが望ましいかが異なるためである。密度関数に対する結果については先行研究として Kato and Sasaki(2018)があるが、先行研究では複数観測の回数がデータの観測数よりも多い状況を考えており、本研究の状況(複数観測は有限回)とは異なる。(i)の研究をより発展させて複数回観測される場合の ME モデルにおける X の確率密度の一様信頼バンドを構成する方法を開発したことで当該分野においてより重要な貢献となった。

研究 2 : 理論的には統計量が(漸近的に)線形である場合を想定し、まず経験過程の理論解析の分野で近年開発された結果(Chernozhukov et al.(2017))と β -mixing の空間過程に対する独立なデータのブロックを構成する方法(Yu(1994))を利用して、各地点の変数の数(p)がデータの観測数(n)よりも多い高次元の場合でも利用可能な中間的ガウス近似を証明した。次にガウス近似の結果を利用し、空間的な従属構造を捉えられるような block bootstrap 法または Shao(2010)で提案されているような dependent wild bootstrap 法を用いて実際に興味のある変数の信頼区間を(n に対して p が非常に大きい高次元の場合であっても)同時に構成することが可能になった。Lahiri and Zhu(2006)のように p が有限の場合はいくつか先行研究が知られているが、高次元の場合も想定した研究は知られていない。一方、情報技術の進歩により、近年ではビッグデータと呼ばれる膨大な数のデータの統計分析が可能になってきている。例えば多数の金融資産からなる(高次元)ポートフォリオのリスク管理などでは従来の時系列モデルに加え、モデルのパラメータ節約のために空間過程を利用した高次元のモデルが実際に利用されるようになってきている。研究 2 の目的が達成されたことで、既存の空間過程のモデルに対する結果の改善のみならず、ビッグデータ解析のためのより高次元の確率過程モデルの正当性に関して理論的保証を与えることができた。

4. 研究成果

本研究課題では

(1) 非定常な空間データ・関数データに対するノンパラメトリック回帰,

- (2) 高次元空間データ・時空間データのための統計的推測方法の開発
- (3) 極値統計学を用いたノンパラメトリック分位点回帰，
- (4) 観測誤差を伴って観測される変数が存在する場合のノンパラメトリックな確率密度の推定法開発

について取り組んだ．結果的には当初の研究課題(研究テーマ(2),(4))に加え，関連テーマとして空間データ・関数データのノンパラメトリック分析に関する研究，さらには研究課題の期間中に新たなテーマとしてノンパラメトリックな分位点回帰についての研究を行った．以上の研究は，

テーマ(1)については Kurisu (2022a), Kurisu (2022b), Kurisu and Matsuda (2024)

テーマ(2)については Kurisu, Kato and Shao (2024)

テーマ(3)については Kurisu and Otsu (2024)

テーマ(4)については Kurisu and Otsu (2022a), Kurisu and Otsu (2022b)

としてそれぞれ国際ジャーナルに掲載済み，または掲載が決定している．

【参考文献】

1. Bonhomme, S. and Robin, J.-M. (2010) Generalized non-parametric deconvolution with an application to earnings dynamics. *Review of Economic Studies*. 77, 491-533.
2. Chernozhukov, V., Chetverikov, D. and Kato, K. (2017) Central limit theorems and bootstrap in high dimensions. *Annals of Probability*. 45, 2309-2352.
3. Kato, K. and Sasaki, Y. (2018) Uniform confidence bands in deconvolution with unknown error distribution. *Journal of Econometrics*. 207, 129-161.
4. Kurisu, D. (2019) On nonparametric inference for spatial regression models under domain expanding and infill asymptotics. *Statistics and Probability Letters*. 154, 108543.
5. Kurisu, D. (2022a). Nonparametric regression for locally stationary random fields under stochastic sampling design. *Bernoulli*, 28, 1250-1275.
6. Kurisu, D. (2022b) Nonparametric regression for locally stationary functional time series. *Electronic Journal of Statistics*, 16, 3973-3995.
7. Kurisu, D., Kato, K., and Shao, X. (2024) Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. Forthcoming in *Journal of the American Statistical Association*.
8. Kurisu, D. and Matsuda, Y. (2024). Local polynomial trend regression for spatial data on R^d . Forthcoming in *Bernoulli*.
9. Kurisu, D. and Otsu, T. (2022a). On the uniform convergence of deconvolution estimators from repeated measurements. *Econometric Theory*, 38, 172-193.
10. Kurisu, D. and Otsu, T. (2022b) On linearization of nonparametric deconvolution estimators for repeated measurements model. *Journal of Multivariate Analysis*, 189, 104921.
11. Kurisu, D. and Otsu, T. (2024) Subsampling inference for nonparametric extremal conditional quantiles. Forthcoming in *Econometric Theory*.
12. Lahiri, S.N. and Zhu, J. (2006) Resampling methods for spatial regression models under a class of stochastic designs. *Annals of Statistics*. 34, 1774-1813.
13. Li, T. and Vuong, Q. (1998) Nonparametric estimation of the measurement error model using multiple indicators. *Journal of Multivariate Analysis*. 65, 139-165.
14. Neumann, M. and Reiss, M. (2009) Nonparametric estimation for Levy processes from low-frequency observations. *Bernoulli* 15, 223-24.
15. Shao, X. (2010) The dependent wild bootstrap. *Journal of American Statistical Association*.

105, 218-235.

16. Yu, B. (1994) Rates of convergence for empirical processes of stationary mixing sequences. *Annals of Probability*. 22, 94-116.

5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 9件/うち国際共著 9件/うちオープンアクセス 2件）

1. 著者名 Kurusu Daisuke	4. 巻 16
2. 論文標題 Nonparametric regression for locally stationary functional time series	5. 発行年 2022年
3. 雑誌名 Electronic Journal of Statistics	6. 最初と最後の頁 3973-3995
掲載論文のDOI（デジタルオブジェクト識別子） 10.1214/22-EJS2041	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 該当する
1. 著者名 Kurusu Daisuke, Otsu Taisuke	4. 巻 38
2. 論文標題 On the uniform convergence of deconvolution estimators from repeated measurements	5. 発行年 2021年
3. 雑誌名 Econometric Theory	6. 最初と最後の頁 172 ~ 193
掲載論文のDOI（デジタルオブジェクト識別子） 10.1017/S0266466620000572	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 該当する
1. 著者名 Kurusu Daisuke, Otsu Taisuke	4. 巻 189
2. 論文標題 On linearization of nonparametric deconvolution estimators for repeated measurements model	5. 発行年 2022年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 104921 ~ 104921
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jmva.2021.104921	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 該当する
1. 著者名 Kurusu Daisuke	4. 巻 28
2. 論文標題 Nonparametric regression for locally stationary random fields under stochastic sampling design	5. 発行年 2022年
3. 雑誌名 Bernoulli	6. 最初と最後の頁 1250 ~ 1275
掲載論文のDOI（デジタルオブジェクト識別子） 10.3150/21-BEJ1385	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 該当する

1. 著者名 Daisuke Kurisu, Taisuke OTsu	4. 巻 -
2. 論文標題 On the uniform convergence of deconvolution estimators from repeated measurements.	5. 発行年 2021年
3. 雑誌名 Econometric Theory	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1017/S0266466620000572	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 該当する

1. 著者名 Naoto Kunitomo, Daisuke Kurisu	4. 巻 -
2. 論文標題 Detecting factors of quadratic variation in the presence of market microstructure noise	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 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s42081-020-00104-w	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 該当する

1. 著者名 Daisuke Kurisu, Yasumasa Matsuda	4. 巻 -
2. 論文標題 Local polynomial trend regression for spatial data on R^d	5. 発行年 2024年
3. 雑誌名 Bernoulli	6. 最初と最後の頁 -
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 該当する

1. 著者名 Kurisu Daisuke, Otsu Taisuke	4. 巻 -
2. 論文標題 Subsampling inference for nonparametric extremal conditional quantiles	5. 発行年 2024年
3. 雑誌名 Econometric Theory	6. 最初と最後の頁 1~15
掲載論文のDOI (デジタルオブジェクト識別子) 10.1017/S0266466623000336	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 該当する

1. 著者名 Kurusu Daisuke, Kato Kengo, Shao Xiaofeng	4. 巻 -
2. 論文標題 Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data	5. 発行年 2024年
3. 雑誌名 Journal of the American Statistical Association	6. 最初と最後の頁 1~13
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/01621459.2023.2218578	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 該当する

[学会発表] 計25件 (うち招待講演 11件 / うち国際学会 14件)

1. 発表者名 Daisuke Kurisu
2. 発表標題 Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data.
3. 学会等名 EcoSta2022 (招待講演) (国際学会)
4. 発表年 2022年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Nonparametric regression for locally stationary random fields on R^d .
3. 学会等名 3rd Tohoku-ISM-UUIm Joint Workshop (国際学会)
4. 発表年 2022年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Adaptive deep learning for nonparametric time series regression.
3. 学会等名 CMStatistics2022 (国際学会)
4. 発表年 2022年

1. 発表者名 栗栖大輔
2. 発表標題 Adaptive deep learning for nonlinear time series.
3. 学会等名 統計関連学会連合大会 (国際学会)
4. 発表年 2022年

1. 発表者名 栗栖大輔
2. 発表標題 深層学習による時系列データの適応的推定
3. 学会等名 科研費シンポジウム「データサイエンスと周辺領域の双方向的理解への挑戦」
4. 発表年 2022年

1. 発表者名 栗栖大輔
2. 発表標題 スパース制約DNNによる時系列データの適応的推定
3. 学会等名 科研費シンポジウム「大規模データ解析の統計的方法論の展開」
4. 発表年 2022年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Spatially dependent wild bootstrap for high-dimensional spatial data
3. 学会等名 XV World Conference of the Spatial Econometrics Association (国際学会)
4. 発表年 2021年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Gaussian approximation and bootstrap for high-dimensional spatial data
3. 学会等名 63rd ISI World Statistics Congress 2021 (国際学会)
4. 発表年 2021年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Wild bootstrap for high-dimensional spatial data
3. 学会等名 Bernoulli-IMS 10th World Congress (国際学会)
4. 発表年 2021年

1. 発表者名 Daisuke Kurisu
2. 発表標題 On the estimation of nonstationary functional time series
3. 学会等名 CSA-KSS-JSS joint international session (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Daisuke Kurisu
2. 発表標題 On the estimation of nonstationary functional data
3. 学会等名 CMStatistics2021 (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 栗栖大輔
2. 発表標題 Spatially dependent wild bootstrap
3. 学会等名 横浜国立大学国際社会科学府セミナー（招待講演）
4. 発表年 2021年

1. 発表者名 栗栖大輔
2. 発表標題 スペクトルアプローチによる確率過程のジャンプ分析
3. 学会等名 第8回 統計数理研究所 リスク解析戦略研究センター 金融シンポジウム「金融が直面する新環境への対応と方法論」（招待講演）
4. 発表年 2021年

1. 発表者名 栗栖大輔
2. 発表標題 非定常な関数時系列データの統計分析
3. 学会等名 シンポジウム「多様な分野における統計科学に関する理論と方法論の革新的展開」（国際学会）
4. 発表年 2021年

1. 発表者名 栗栖大輔
2. 発表標題 非定常な関数時系列データの特徴量推定
3. 学会等名 統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 栗栖大輔
2. 発表標題 局所線形極値分位点回帰
3. 学会等名 JAFEE大会
4. 発表年 2022年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Wild bootstrap for spatio-temporal data.
3. 学会等名 CMStatistics2020 (国際学会)
4. 発表年 2020年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Spatially dependent wild bootstrap for high-dimensional spatial data.
3. 学会等名 University of Alberta Statistics Seminar (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 栗栖大輔
2. 発表標題 確率場に対する高次元正規近似.
3. 学会等名 慶應義塾大学, 計量経済学ワークショップ (招待講演)
4. 発表年 2020年

1. 発表者名 栗栖大輔
2. 発表標題 点過程アプローチによる条件付き極値分位点のノンパラメトリック推定.
3. 学会等名 研究集会「極値理論の工学への応用」
4. 発表年 2020年

1. 発表者名 栗栖大輔
2. 発表標題 Nonparametric regression for locally stationary random fields.
3. 学会等名 統計関連学会連合大会
4. 発表年 2020年

1. 発表者名 栗栖大輔
2. 発表標題 Bootstrap for spatio-temporal data.
3. 学会等名 東京大学, 応用統計ワークショップ (招待講演)
4. 発表年 2020年

1. 発表者名 栗栖大輔
2. 発表標題 Inference on extremal conditional quantiles.
3. 学会等名 東北大学, Data Science Workshop (招待講演)
4. 発表年 2020年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Subsampling inference for nonparametric extremal conditional quantiles
3. 学会等名 ICIAM2023 (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 Daisuke Kurisu
2. 発表標題 Spatially dependent wild bootstrap for high-dimensional spatial data.
3. 学会等名 IMS-APRM2024 (招待講演) (国際学会)
4. 発表年 2024年

〔図書〕 計1件

1. 著者名 国友 直人, 栗栖 大輔	4. 発行年 2021年
2. 出版社 朝倉書店	5. 総ページ数 413
3. 書名 極値現象の統計分析: 裾の重い分布のモデリング	

〔産業財産権〕

〔その他〕

<p>研究代表者ホームページ https://sites.google.com/site/daisukekurisu/home</p>

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			
米国	Cornell University	University of Illinois (UIUC)		
英国	London School of Economics			