

令和 3 年 6 月 7 日現在

機関番号：12501

研究種目：若手研究

研究期間：2018～2020

課題番号：18K12754

研究課題名(和文)形状に制約のあるベイズ関数回帰モデルと経済データへの応用

研究課題名(英文) Bayesian shape-restricted functional regression with application to economic data

研究代表者

小林 弦矢 (Kobayashi, Genya)

千葉大学・大学院社会科学研究院・准教授

研究者番号：00725103

交付決定額(研究期間全体)：(直接経費) 4,800,000円

研究成果の概要(和文)：本研究では形状に制約のある関数モデルのベイズ分析を取り上げ、特に非線形分位点回帰モデル、状態空間SIRモデル、ロレンツ曲線を取り扱い、女性の労働、所得不平等などの経済学的重要な課題及びCOVID-19の感染拡大予測といった社会的に喫緊の課題に関する応用を行った。形状に制約のあるモデルの取り扱いは様々で、ガウス過程、微分方程式、パラメトリック関数などに基づいたものを取り扱い、それらを中心に状態空間モデル、階層モデル、ノンパラメトリックベイズを構築することで制約を満たしつつ柔軟なモデルを提案した。関連して、所得分布の空間モデル、柔軟性と推定の安定性を両立するような混合モデルについても扱った。

研究成果の学術的意義や社会的意義

関数の形状に制約がある事前情報がある場合、それを適切にモデルに反映させることで柔軟性を保ちつつより効率的にモデルの推定を行うことができるようになった。本研究の意義は回帰モデルにおける非線形関数の推定にとどまらず、ロレンツ曲線やSIRモデルなどといった関数あるいは方程式の解に形状制約があるものに対して状態空間モデルを構築することで、現実のデータのばらつきを許容・捕捉キャプチャーできるようになり、また時点感・空間上で情報を借り合うことでモデルをより安定的に推定し、将来の予測を行えるようになった。

研究成果の概要(英文)：This project considered Bayesian analysis for shape-restricted functional models, particularly, nonlinear quantile regression models, state space SIR model and Lorenz curves, and applied the proposed approaches to some important economic problems, such as female labour, income inequality, and urgent social and economic problem of the spread of COVID-19 and its prediction. Various approaches to introducing shape restriction were considered, for example, Gaussian process, differential equations, parametric families of functions. Centred on those shape-restricted functions, flexible models based on state space models, hierarchical models, nonparametric Bayes models were considered. As related studies, the model for spatial income distributions and mixture model with both flexibility and estimation stability were also considered.

研究分野：ベイズ統計学

キーワード：ベイズ統計学 状態空間モデル 分位点回帰モデル

1. 研究開始当初の背景

単純な回帰分析では被説明変数と共変量の間には線形な関係を仮定して統計分析を行うが、線形性の仮定は関心のある重要な共変量の被説明変数に対する非線形な効果を捉えきることができない。関数回帰では被説明変数を共変量の未知関数に回帰するモデルを考え、データから被説明変数と共変量の関数的な関係について推測を行う。滑らかな関数をデータから推定するために関数を種々の互いに独立な既知関数(基底関数)の線形結合によって表し、線形結合の係数パラメータをデータから推定することで関数の推定を行う。近年では応用研究においてより高度な統計的分析手法の適用が必要とされてきているが、本研究で扱うような応用研究において重要とされている分析手法において未知関数を柔軟にかつ安定的に分析するための制約付関数回帰の手法については皆無であるのが現状であり、新たな統計的分析手法を提案する必要がある。さらに、制約付関数回帰を実際の経済データに適用するにあたって、制約を課さない場合と課した場合とで推定の精度・安定性や得られる政策的な示唆がどれほど異なるのか、課された形状制約が本当に妥当なものであるか、といった点についてもまだ明らかにされていないため、本研究において追求していく必要がある。

2. 研究の目的

制約付関数回帰モデルの新たなベイズ計量モデルの開発：本研究では近年経済データ分析において広く使用される重要な統計的分析手法に対して制約付関数回帰に基づいた新たなベイズ計量モデルの開発を行う。

新たな計量モデルに対する統計的分析手法の開発：本研究で提案するモデルは潜在変数を多く含む階層モデルであったりするため、尤度関数が複雑となり解析的に解けないことが予想される。そのため、シミュレーションによる推定方法(主にマルコフ連鎖モンテカルロ法、MCMC法)による推定方法の開発を目指す。またシミュレーションによる推定方法を用いることで、未知関数に対する区間推定も容易に行うことができる。さらに、分析対象のデータにおいてどの形状制約が適切であるかを評価するために、いくつかのモデルの比較基準の適用も試みる。

現実の経済問題への応用：本研究で提案する統計的分析手法の有用性を示すためにそれぞれの分析手法において実際の経済データの分析を行う。ここでは制約を課さない場合と課した場合とで推定結果の安定性や得られる政策的な示唆がどれほど異なるのか、課された形状制約が本当に妥当なものであるかについて検証する。

3. 研究の方法

分位点に関するモデル：分位点関数回帰を考えた場合、中央値など分布の中央での回帰に比べて、分布の裾における回帰では推定結果が非常に不安定になりやすく、分位点の推定値が単調性をより一層満たしにくくなってしまふ。分位点回帰モデルでは関数の形状が分位点ごとに異なる可能性があるため、分位点回帰モデルでは様々な形状を表現できるガウス過程に基づいたアプローチを検討する。数値実験によって、提案する推定方法が効率的であるかを確認するとともに、複数の分位点において関数の形状に制約を課さない場合に比べて形状制約を課した場合では、関数の推定精度がどれだけ改善されるかを様々なシナリオのもとで調べる。さらに、各分位点において関数の形状選択のための情報量基準などの基準がどれだけうまく機能するかについても調べる。そして女性の労働参加や医療支出に関するデータへ適用し、どの形状制約がデータに対して適切なのか、形状制約によって労働経験年数が労働時間に与える効果や家計所得が医療支出に与える効果についての政策的な示唆を明らかにする。

可変・時変係数モデル：可変・時変係数モデルは、ある共変量の被説明変数に対する効果が個人の特性や時間によって非線形に変化するようなモデルで、個人や時間による様々な異質性を柔軟にモデルに反映することができる。モデルの構造上潜在変数が多く含まれていたりするため、関数の推定結果が不安定となってしまうことがある。まずは被説明変数の平均に関する形状制約付可変・時変係数モデルをとりあげ、さらに近年可変・時変係数モデルへの拡張が注目されている分位点回帰モデルへの適用も行う。

ローレンツ曲線：ローレンツ曲線は所得格差の分析において重要な増加凸関数であり、分位点関数とも関係を持っている。既存のパラメトリックなローレンツ曲線は柔軟性や所得分布としての解釈性に欠けるため、分位点関数に対しての凸関数への適用を行うことでローレンツ曲線のモデルの再定式化を試みる。これにより尤度関数の定式化が可能となり、ベイズ推定を行うことが可能となると予想される。数値実験において推定方法の妥当性の確認やローレンツ曲線の推定精度の比較を行う。

4. 研究成果

分位点に関するモデル：ベイズ分位点回帰モデルにおいて柔軟性と関数の形状制約による安定性を両立する方法を次のように提案した。この研究ではガウス過程 2 乗の 1 次および 2 次導関数に制約を導入することで、単調関数、単調凸・凹関数、U・S 字関数を実現する方法を取り上げ

た。またベイズ分位点回帰モデルの誤差項に対して一般化非対称ラプラス分布の尺度パラメータと形状パラメータに対するディリクレ過程混合モデルを考えることで、制約的であった従来の非対称ラプラス誤差分布の裾の形状の柔軟性の改善を試みた。数値実験の結果から、正しい形状制約を課したモデルは特に被説明変数の条件付き分布の裾にあたる分位点を推定する場合に推定精度や安定性が高いことが確認された。また、モデル選択規準に基づいた形状制約の選択についても、正しい形状制約が課したモデルが選ばれる頻度が一番高くなるという結果も得られた。また提案するディリクレ過程混合モデルと通常的一般化非対称ラプラス分布の比較においては、提案するモデルのほうが非説明変数の分布の裾の情報をより良く捉えることができ、結果として未知関数の推定精度がよいことが確認された。

可変・時変係数モデル：本研究では前述の分位点回帰モデルを、パネルデータを分析するために階層モデルへの拡張を行った。具体的にはガウス過程に基づく基底係数に階層構造を導入し、パネルデータのユニット間あるいは時点間における未知関数の異質性を補足することができるようなモデルの提案を行った。上述の分位点回帰モデルと合わせてこの研究は韓国との共同で行い、研究成果は統計学の国際査読誌に掲載された。

ローレンツ曲線：当初は Bernstein 多項式の基底係数に制約を導入することでセミパラメトリックなローレンツ曲線を表現することを考えていたが、日本の家計調査のデータのようにアクセスしやすいデータにおいては少ない数（例えば5）の所得階級からなるグループデータが一般的であり、精度良くセミパラメトリックローレンツ曲線を推定することは試験的な数値実験から極めて困難であることがわかった。そこでパラメトリックローレンツ曲線に対する推定の改善を行う方針で研究を進めた。まず、所得分布のローレンツ曲線とそれに付随する不平等指標をグループデータから安定的にベイズ推定する方法として近似ベイズ計算法（ABC）を取り上げ、逐次モンテカルロサンプラー（SMC サンプラー）に基づく推定方法を提案した。数値実験から、提案する方法は一般化ベータ分布といったパラメータ数の多い分布も既存手法よりも精度良く推定することが確認された。ここまでの研究は国際査読誌に掲載された。続いて、時系列グループデータにおいて時点間の情報を借り合うことで、ローレンツ曲線に関する推定の安定性の改善と将来予測を可能にするために状態空間モデルを構築した。既存研究では一カ国・一時点に関する分析が独立に行われていたのに対し、本研究で複数時点のデータを同時に取り扱う。観測方程式にはディリクレ分布あるいはその一般化を用い、ローレンツ曲線のパラメータの変換が自己回帰過程などに従うという状態方程式を考えた。数値実験の結果から、提案する手法は時点ごとに独立に推定するよりも不確実性を大幅に減らすことができることが確認された。また日本の家計調査のデータの分析結果から、2010年代では上位の所得階級のシェアとジニ係数が少しずつ低下してきていることもわかった。本研究成果は国際学会で報告され、国際査読誌に掲載された。

COVID-19 の感染予測：2020年はじめより COVID-19 の流行が社会経済的に大きな問題になってきており、今後の感染拡大の度合いを評価するために本研究では状態空間 SIR モデルを取り上げ2020年6月以降の感染状況の予測を行った。疫学でよく利用される SIR モデルは、微分方程式の解が決定的に与えられるので、大きなランダムネスがある現実のデータに対しては当てはまりがよくなく、また予測区間などといった不確実性の評価を行うことができない。本研究では、状態方程式に SIR モデルの解を取り入れ、各コンパートメントの人口に対する割合がディリクレ分布に従い、観測された感染者の割合がベータ分布に従うモデルを取り上げた。さらに緊急事態宣言後には流動人口の減少から感染率が下がることが考えられるため、ある未知の時点以降感染率パラメータが変化するという変化点モデルを考えた。日本におけるデータの分析の結果、緊急事態宣言などの介入によって流動人口の低下が長期的に続き、介入終了後の感染率の水準が介入前よりも十分に小さい状態であれば国内の感染は比較的早期に収束するという予測結果となった。また感染率パラメータの値は2020年4月の緊急事態宣言発令後に発令前よりも小さな値として推定された。本研究成果は国際査読誌に掲載され、毎日新聞などの記事でも取り上げられることになった。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 4件／うち国際共著 1件／うちオープンアクセス 0件）

1. 著者名 Sugasawa Shonosuke, Kobayashi Genya, Kawakubo Yuki	4. 巻 145
2. 論文標題 Estimation and inference for area-wise spatial income distributions from grouped data	5. 発行年 2020年
3. 雑誌名 Computational Statistics & Data Analysis	6. 最初と最後の頁 106904 ~ 106904
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.csda.2019.106904	査読の有無 無
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 Kobayashi Genya, Kakamu Kazuhiko	4. 巻 34
2. 論文標題 Approximate Bayesian computation for Lorenz curves from grouped data	5. 発行年 2018年
3. 雑誌名 Computational Statistics	6. 最初と最後の頁 253 ~ 279
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s00180-018-0831-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 -
1. 著者名 Sugasawa Shonosuke, Kobayashi Genya, Kawakubo Yuki	4. 巻 29
2. 論文標題 Latent mixture modeling for clustered data	5. 発行年 2018年
3. 雑誌名 Statistics and Computing	6. 最初と最後の頁 537 ~ 548
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s11222-018-9821-7	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 Kobayashi Genya, Yamauchi Yuta, Kakamu Kazuhiko, Kawakubo Yuki, Sugasawa Shonosuke	4. 巻 -
2. 論文標題 Bayesian Approach to Lorenz Curve Using Time Series Grouped Data	5. 発行年 2021年
3. 雑誌名 Journal of Business & Economic Statistics	6. 最初と最後の頁 1 ~ 16
掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/07350015.2021.1883438	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 Kobayashi Genya, Roh Taeyoung, Lee Jangwon, Choi Taeryon	4. 巻 -
2. 論文標題 Flexible Bayesian quantile curve fitting with shape restrictions under the Dirichlet process mixture of the generalized asymmetric Laplace distribution	5. 発行年 2020年
3. 雑誌名 Canadian Journal of Statistics	6. 最初と最後の頁 1
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/cjs.11582	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 Kobayashi Genya, Sugasawa Shonosuke, Tamae Hiromasa, Ozu Takayuki	4. 巻 14
2. 論文標題 Predicting intervention effect for COVID-19 in Japan: state space modeling approach	5. 発行年 2020年
3. 雑誌名 BioScience Trends	6. 最初と最後の頁 174 ~ 181
掲載論文のDOI (デジタルオブジェクト識別子) 10.5582/bst.2020.03133	査読の有無 無
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6件 (うち招待講演 5件 / うち国際学会 5件)

1. 発表者名 Genya Kobayashi
2. 発表標題 Bayesian approach to Lorenz curve using time series grouped data
3. 学会等名 IISA 2019 Conference (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Genya Kobayashi
2. 発表標題 Bayesian approach to Lorenz curve using time series grouped data
3. 学会等名 Workshop on Bayesian Modelling of Income and Wealth (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Genya Kobayashi
2. 発表標題 Estimation and inference for area-wise spatial income distributions from grouped data
3. 学会等名 SAE 2019 (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Genya Kobayashi
2. 発表標題 Flexible Bayesian quantile curve fitting with shape restrictions based on Gaussian process and generalised asymmetric Laplace distribution
3. 学会等名 ISBA World Meeting 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Genya Kobayashi
2. 発表標題 Flexible Bayesian quantile curve fitting with shape restrictions based on Gaussian process and generalised asymmetric Laplace distribution
3. 学会等名 ISBA EAC 2018 (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Genya Kobayashi
2. 発表標題 Shrinking semiparametric Bayesian quantile regression with transferring functions
3. 学会等名 ベイズ計量経済分析研究集会 (招待講演)
4. 発表年 2018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			
韓国	Korea University			