

令和元年5月22日現在

機関番号：12501

研究種目：若手研究(B)

研究期間：2016～2018

課題番号：16K17101

研究課題名(和文) 混合効果モデルの変数選択問題と小地域推定への応用

研究課題名(英文) Variable selection problem for mixed effects model and its application to small area estimation

研究代表者

川久保 友超 (Kawakubo, Yuki)

千葉大学・大学院社会科学研究院・准教授

研究者番号：80771881

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)：本研究課題では、小地域推定で用いられる線形混合モデルの変数選択問題において、主に以下の2点の研究において成果をあげた。

1点目は、conditional AICという線形混合モデルの変数選択規準において、推定に用いる共変量の値と予測に用いる共変量の値が異なる状況(共変量シフト)を考え、従来のcAICを改良した変数選択規準を構成するという研究を行った。2点目は、予測量の平均二乗誤差が最小となるようにモデルの未知パラメータを推定するobserved best prediction (OBP) 法を拡張し、同一の観点からモデル選択規準を構築する研究を行った。その他の小地域推定に関連した研究も行った。

研究成果の学術的意義や社会的意義

本研究課題であげた研究成果は、線形混合モデルの変数選択における新しい手法を提案したものであり、また提案手法の理論的な性質についても議論している。ともに統計学分野の国際査読誌に掲載され、一定の学術的意義を認められたと言える。

またここであげた研究成果は、線形混合モデルを用いた統計手法に広く応用可能であるが、その中でも特に小地域推定と呼ばれる官庁統計における重要な課題への応用を強く意識して研究を行ってきた。小地域推定における実務での有用性という観点からも、社会的意義のある研究成果だと言える。

研究成果の概要(英文)：In this program, I mainly conducted the following two pieces of research in the variable selection problem for linear mixed model, which is used in small area estimation.

First, I considered the covariate shift situation, which is when the values of covariates in the model for prediction differ from those in the model for observed data. Under the covariate shift situation, I constructed a variable selection criterion for linear mixed model based on the conditional AIC. Second, I considered variable selection and estimation simultaneously to minimize the total mean squared prediction errors of the predictor. The derived method can be regarded as an extension of the observed best prediction (OBP) method.

I also conducted research related to small area estimation or mixed effects model.

研究分野：経済統計

キーワード：小地域推定 情報量規準

## 様式 C-19、F-19-1、Z-19、CK-19（共通）

### 1. 研究開始当初の背景

#### (1) 線形混合モデルを用いた小地域推定とは

まず小地域推定とはどのような研究分野であるかを簡単に説明する。ある標本調査にもとづいて各市区町村における何らかの特徴量（平均所得・貧困率・疾病リスクなど）を測ろうとする際、サンプルサイズが小さかったりゼロであったりする市区町村（小地域）においては、標本平均のようにその地域の標本調査のデータのみにもとづいたものは、分散が大きくなり安定しない。そこで、線形混合モデルと呼ばれる統計モデルを適切に仮定し、小地域の特徴量の予測量として経験最良予測量（empirical best predictor, EBP）をモデルから導出すると、この予測量は安定する。その理由は、EBPは補助情報を与える変数（補助変数）や他の似た地域の情報（地理情報）を利用しており、標本調査では十分な情報を得ることができなかった地域において他の情報を補填した予測量となっているからである。

#### (2) 線形混合モデルの変数選択問題

線形混合モデルを用いた小地域推定問題において、関心のある小地域の特徴量を推定するために補助情報がいくつか利用可能な場合がある。その中から、線形混合モデルの説明変数としてどの変数を入れるべきか、統計学的な規準が必要となる。モデル選択および変数選択と呼ばれるこの問題に対しては、赤池情報量規準（AIC）が最も広く用いられているが、線形混合モデルの変数選択においては、変量効果を積分消去した周辺尤度にもとづいて構成する従来のAICは、特定の変量効果（地域効果など）の予測に関心がある際には適切でない。そこで Vaida and Blanchard (2005) は、変量効果を所与とした条件付尤度にもとづいた Kullback-Leibler (KL) ダイバージェンスで予測モデルを評価した conditional AIC (cAIC) を提案した。

### 2. 研究の目的

線形混合モデルを用いた小地域推定は、官庁統計の分野のみならず多くの応用分野で近年ますます関心が寄せられてきている。一方で、線形混合モデルの変数選択問題が議論されるようになったのはごく最近からであり、多くの課題が残っている。線形混合モデルの変数選択における1つの大きな技術革新は、Vaida and Blanchard (2005) による cAIC の開発である。同論文は、従来のAICを線形混合モデルに適用する際の問題点を指摘した上で、従来のAICを優越した規準 cAIC を提案した。理論と応用の両面から cAIC は大きな注目を集め、盛んに研究がなされている。そこで本研究課題では、cAIC の理論的課題と、その小地域推定問題への応用の研究を行うという目的で、研究を行ってきた。

### 3. 研究の方法

本研究課題においては、統計学における一般理論と応用の双方にとって有用な研究であることが求められる。そこで、小地域推定という具体的な応用問題に着目したうえで、一般理論の研究で得られた線形混合モデルの変数選択規準の手法を適用しようとした場合の問題点を探るところから始めた。その中で見つけた具体的な問題が、「(1) 共変量シフト下での条件付 AIC」および「(2) OBP 法にもとづいた変数選択」の2点であった。これらの問題に対して、統計理論を用いた解決法を提案するとともに、提案手法のパフォーマンスをシミュレーションや実データへの応用をとおして調べていった。

### 4. 研究成果

研究成果の概要であげた2点の研究について、以下で詳しく述べる。

#### (1) 共変量シフト下での条件付 AIC

ここで述べる研究成果は、「5. 主な発表論文等〔雑誌論文〕」における③の国際査読誌に掲載されている。本研究では、共変量シフト下での線形混合モデルにおける固定効果に関する説明変数の選択問題を考えた。共変量シフトとは、予測のためのモデルにおける共変量の値が、観測データに対するモデルの共変量の値と異なる状況のことを言う。この問題設定において、Vaida and Blanchard (2005) で導入された conditional Akaike information にもとづいた変数選択規準を構築した。特に、小地域推定における共変量シフトに着目し、提案規準の有用性を示した。また、シミュレーションをとおして、提案手法の数値的なパフォーマンスも調べた。シミュレーションのうちの1つは、日本の公示地価の実データにもとづいたデザインベースのシミュレーションであり、統計モデルで仮定した確率構造が必ずしも正しくない状況においても、提案手法が既存手法よりも良いパフォーマンスを示した。

#### (2) OBP 法にもとづいた変数選択

ここで述べる研究成果は、「5. 主な発表論文等〔雑誌論文〕」における①の国際査読誌に掲載されている。小地域推定の方法論においては通常、適切な共変量の選択と、選択されたモデルにおける推定が、別々に考えられてきた。本研究では、小地域平均の予測量の平均予測二乗誤

差 (mean squared prediction error, MSPE) の合計値を最小にするように、変数選択と推定を同時にすることを考えた。導出した手法を observed best selective prediction (OBSP) 法と呼ぶが、これは Jiang et al. (2011) で提案された observed best prediction (OBP) 法の拡張とみなすことができる。真のモデルが候補モデルのクラスに含まれている場合、OBSP 推定量は真のモデルにおける未知パラメータの一致推定量となる。漸近論にもとづいて、予測量の MSPE を推定するパラメトリックブートストラップ法を用いた数値的な手法を開発した。数値実験をとおして、OBSP 法の有限標本におけるパフォーマンスを調べた。対抗手法として、AIC と BIC で変数選択した OBP 予測量、および利用可能なすべての説明変数をモデルに取り入れた OBP 予測量を考えた。さらに、提案手法を日本の家計調査のデータに応用し、実務における有用性も示した。

#### <引用文献>

Jiang, J., Nguyen, T. and Rao, J. S. (2011). Best selective small area estimation, *Journal of the American Statistical Association*, 107, 732-745.

Vaida, F. and Blanchard, S. (2005). Conditional Akaike information for mixed-effects models, *Biometrika*, 92, 351-370.

#### 5. 主な発表論文等

##### [雑誌論文] (計 4 件)

① Sugasawa, S., Kawakubo, Y. and Datta, G. S. (2019). "Observed best selective prediction in small area estimation". *Journal of Multivariate Analysis*, 173, 383-392. DOI: 10.1016/j.jmva.2019.04 (査読あり)

② Sugasawa, S., Kobayashi, G. and Kawakubo, Y. (2019). "Latent mixture modeling for clustered data". *Statistics and Computing*, 29, 537-548. DOI: 10.1007/s11222-018-9821-7 (査読あり)

③ Kawakubo, Y., Sugasawa, S. and Kubokawa, T. (2018). "Conditional Akaike information under covariate shift with application to small area estimation". *Canadian Journal of Statistics*, 46, 316-335. DOI: 10.1002/cjs.11354 (査読あり)

④ Kawakubo, Y., Kubokawa, T. and Srivastava, M. S. (2018). "A variant of AIC based on the Bayesian marginal likelihood". *Sankhya B*, 80, 60-84. DOI: 10.1007/s13571-018-0152-7 (査読あり)

##### [学会発表] (計 10 件)

① 川久保友超 "小地域推定モデルの変数選択". 研究集会「公的大規模データの利用におけるプライバシー保護の理論と応用」, 2018 年, 統計数理研究所

② 川久保友超 "observed best prediction 法による小地域推定". ベイズ計量経済分析研究集会, 2018 年, 関西学院大学大阪梅田キャンパス

③ 川久保友超 "度数分布にもとづいた小地域推定". 計量経済学ワークショップ@慶應義塾大学経済研究所, 2018 年, 慶應義塾大学

④ Yuki Kawakubo "Small area estimation for grouped data". Colloquium at Department of Statistics, University of Georgia, 2018, University of Georgia, USA.

⑤ Yuki Kawakubo "Small area estimation for grouped data". Eastern Asia Chapter of ISBA, 2018, Seoul National University, South Korea.

⑥ Yuki Kawakubo "Small area estimation for grouped data". ISBA World Meeting, 2018, The University of Edinburgh, UK.

⑦ Yuki Kawakubo "Small area estimation for grouped data". International Workshop on Bayesian Econometric Analysis, 2017, University of Tokyo, Japan.

⑧ 川久保友超 "Small area estimation for grouped data". 統計関連学会連合大会, 2017 年, 南山大学

⑨Yuki Kawakubo "Small area estimation for grouped data". Small Area Estimation 2017, 2017, Pierre and Marie Curie University, France.

⑩川久保友超 "Conditional Akaike information under covariate shift with application to small area prediction". 統計関連学会連合大会, 2016年, 金沢大学

〔図書〕 (計 0 件)

〔産業財産権〕

○出願状況 (計 0 件)

○取得状況 (計 0 件)

〔その他〕

ホームページ等

Yuki Kawakubo Home Page

<https://sites.google.com/site/ykawakubostat/>

## 6. 研究組織

(1) 研究分担者

無

(2) 研究協力者

無

※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、研究者個人に帰属されます。