

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 31 日現在

機関番号：12601
研究種目：基盤研究(C) (一般)
研究期間：2013～2016
課題番号：25330035
研究課題名(和文) 縮小型事前分布に基づくBayes統計理論の研究

研究課題名(英文) Bayesian statistics based on shrinkage prior

研究代表者
丸山 祐造 (Maruyama, Yuzo)

東京大学・空間情報科学研究センター・准教授

研究者番号：30304728
交付決定額(研究期間全体)：(直接経費) 3,800,000円

研究成果の概要(和文)：この研究では平均と分散の両方が未知の場合の予測分布の構成を扱った。まずカルバックライブラ損失のもとで通常の予測分布が有限標本のもとで非許容的であることを示した。興味深いのは平均ベクトルの次元が3未満であっても、このスタイン現象が生じることである。これは分散が未知の場合の設定との比較において驚くべき理論的結果である。また、予測問題とパラメータ推定問題の関係についても理解が進んだ。

研究成果の概要(英文)：This work treats the problem of estimating the predictive density of a random vector when both the mean vector and the variance are unknown. We prove that the density of reference in this context is inadmissible under the Kullback-Leibler loss in a nonasymptotic framework. Our result holds even when the dimension of the vector is strictly lower than three, which is surprising compared to the known variance setting. Finally, we discuss the relationship between the prediction and the estimation problems.

研究分野：統計科学

キーワード：ベイズ統計学 統計的決定理論

1. 研究開始当初の背景

線形回帰モデルは、多変量統計解析を行う上での最も基本的かつ重要なモデルである。もちろん複雑なデータ構造に対処するために様々な非線形な統計モデルが使われるが、Bスプラインのように、いくつかの非線形関数の線形結合として表現される例は多い。このことから、計算機が高度に発達した現代においても、線形回帰モデルは依然として重要な位置を占めている。またモデルの解釈の容易さの観点から、線形回帰モデルの優位性は動かない。本研究では、理論的立場から線形回帰モデルにおける統計的推測問題を考える。

2. 研究の目的

(1) 線形回帰モデルにおける統計的推測問題において、特にその重要性・有効性が明らかになってきている縮小型事前分布をベイズ統計的推測手法に照準を絞る。もちろん回帰分析の文脈では、リッジ回帰分析が縮小型推測手法として、まず想定されるが、必ずしもリッジ回帰分析だけに焦点を絞らず、縮小型推測手法一般について、いわゆる統計的決定理論の観点から理論的性質を解明していくことが本研究の大きな目的である。

本研究では主に、誤差項に正規性を仮定した正規線形回帰モデルを想定し、統計的推測として、推定、予測、変数選択を扱う。(ただし、行った一部の研究ではより広い分布のクライスである一般球面对称分布のもとでのベイズ推測手法が、正規分布のもとでのそれと一致するというロバスト性の観点から興味深い結果も得られる。)具体的には、有限個のデータが得られたもとで、以下の典型的問題

「推定」回帰係数ベクトルの関数の推定、誤差分散の推定

「予測」新たな有限個の説明変数が得られたときに、対応する目的変数に関する予測分布の構成

「変数選択」無駄な変数を含む利用可能な全ての説明変数が与えられていると想定した場合の、データに照らして妥当な説明変数の組み合わせの選択

において、理論的にも実用的にも優れた推測方法の提案を目指す。

3. 研究の方法

過去に受領した科研費では、本研究テーマと深く関連した研究テーマを扱った。その研究費を使って、アメリカの複数の研究者を訪問して、共同研究を行ってきた。近年の業績の多くがその共同研究の成果である。本研究においても、そのスタイルでの共同研究を継続する。そのため研究費のかなりの部分を旅費に当てる。また、海外の学会、研究会で研究成果を発表し、同分野の専門家と議論することにより、さらに研究を進展させる。その旅費にも研究費を充当する。

4. 研究成果

(1) 平成 25 年度は、ミニマクス性とスパース性を併せ持つ推定量の構成に集中して研究した。具体的な問題は、 d 変量正規分布の平均ベクトルの二乗損失関数のもとでの推定問題である。 $d \geq 3$ のときには、Stein 現象が生じて、最尤推定量は非許容的になる。このとき James-Stein positive-part 推定量 (JSPP) は 1 つの改良型推定量として知られている。JSPP 推定量をモデル選択の枠組みで考えるとき、null model か full model の二択になっていることが欠点だと考えられる。Zhou and Hwang (2005, Annals) は、縮小関数を L_2 norm の関数でなく L_p norm の関数とすることによって 2 つの候補からのモデル選択を可能にし、また同時にミニマクス性を持つ縮小型推定量を提案した。私は、Zhou and Hwang の結果を拡張して、彼らが p に課していた制約を除き、任意の正なる p を用いた L_p norm の関数で推定量を構成できることを示した。James-Stein positive-part 推定量は、経験ベイズ推定量として解釈できる。Zhou and Hwang も彼らの推定量がある種のベイズ推定量として解釈できることを示したが、不完全である。 L_p norm を縮小関数とする縮小型推定量のベイズ的解釈を与えることは今後の課題としたい。なお、上に述べた平成 25 年度に得られた結果は、テクニカルレポートとして Arxiv に投稿した。

(2) 平成 26 年度は、正規線形回帰モデルにおけるベイズ統計学に基づくモデル選択問題を主に研究した。この問題において典型的に想定される事前分布として、Zellner の g -prior がある。この事前分布は、その事前分散共分散行列が説明変数行列に依存するという性質を持つ。結果として、周辺尤度が解析的に計算可能であり、その計算可能性は最終的な選択規準である Bayes Factor にも引き継がれるという長所を持つ。例えば、Maruyama and George (Annals of Statistics, 2011) はその性質を最大限に利用している。Bayes Factor の理論的な Justification として、サンプル数 n が十分大きい場合の一致性がある。Maruyama and George (2011) では、説明変数の数 p を固定した下で一致性を議論したが、 p が n とともに増えるような状況における一致性については必ずしも結果が知られていなかった。今回、Michigan Technological University の Wang 氏との共同研究として、 p も大きい状況における一致性に関する結果を導出した。特に p が n の線形関数である場合は、本質的に Analysis of Variance において、水準数が大きい場合に対応する。この場合には、モデル間のある種の距離が小さい場合に一致性を持たない場合が生じることが興味深い。一連の研究結果は雑誌 Bernoulli に掲載されることになっている。

(3) 平成 27 年度は、分散未知の場合の多変量正規分布の平均ベクトルの推定問題に注力した。この設定においても、3 次元以上の場合にスタイン現象が生じることが知られており、自然な推定量を改良するミニマクス推定量のクラスが知られている。一方、(特にミニマクス推定量であり、かつ)一般化ベイズ推定量の許容性に関する結果は知られていない。これは分散既知の場合の一般化ベイズ推定量の許容性について、完全な結果が得られていることとは対照的であり、より現実的な設定である分散未知の設定における、一般化ベイズ推定量が許容的であるための十分条件に興味を持たれてきた。この問題に対して、考え得る一つのアプローチとして、スタインのリスクの不偏推定量に基づく方法があり、少なくとも許容的であるためのある種の必要条件が得られる。我々はその必要条件を満たすことを quasi-admissibility と定義して、quasi-admissible となる一般化ベイズ推定量に関する様々な性質を得ている。もちろん quasi-admissibility は許容性よりも弱い良さの基準ではあるが、分散既知の場合に quasi-admissible でかつ一般化ベイズであれば、許容的であることが示されるため、見込みのあるアプローチであると考えている。近いうちにこれまで得られた結果をまとめて学術雑誌に投稿予定である。また同時に Arxiv にも置く予定である。この研究はアメリカ・Rutgers 大学の Strawderman 教授との共同研究であり、平成 27 年度においては 2 度先方を訪問して共同研究を遂行した。

(4) 最終年度である平成 28 年度は、平成 27 年度に引き続き分散未知の場合の多変量正規分布の平均ベクトルの推定問題を考えた。平成 25 年度から継続してきた研究テーマであるが、平成 28 年度に不完全であった quasi-(in)admissibility の境界部分をようやく解明した。論文としてまとめて Arxiv に投稿し、また同時に査読付き国際学術誌にも投稿した。なお、この研究はアメリカ・Rutgers 大学の Strawderman 教授との共同研究であり、平成 28 年度においては 2 度先方を訪問して共同研究を遂行した。

(5) 最終年度である平成 28 年度のもう一つの研究成果として、主に平成 25 年度に研究していた l_p norm に関わる James-Stein 推定量の研究の進展があげられる。古典的なミニマクス性は非常に保守的であり、結果として sparse 性を追求すると、ミニマクス性を失ってしまうケースが多い。これに対して、より弱いミニマクス性の規準であるアンサンブルミニマクス性を持つ l_p norm James-Stein 推定量を提案した。アンサンブルミニマクス性は、ペンシルバニア大学の Brown 教授のグループによって別の文脈で推奨されていたベイズ的な規準である。ある種の対称性があるために、 l_p norm

James-Stein 推定量の性質を議論する場合において親和性が高い。この研究は引き続き研究中であり、平成 29 年度中に Arxiv に論文を投稿したいと考えている。

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 5 件)

- ① Min Wang and Yuzo Maruyama (2015). Consistency of Bayes factor for nonnested model selection when the model dimension grows. *Bernoulli*, 22, 2080-2100. 査読有 doi:10.3150/15-BEJ720
- ② Yuzo Maruyama and William E. Strawderman (2014). Robust Bayesian variable selection in linear models with spherically symmetric errors. *Biometrika*, 101, 992-998. 査読有 doi:10.1093/biomet/asu039
- ③ Aurélie Boisbunon and Yuzo Maruyama (2014). Inadmissibility of the best equivariant predictive density in the unknown variance case. *Biometrika*, 101, 733-740. 査読有 doi:10.1093/biomet/asu024
- ④ Edward, I. George and Yuzo Maruyama (2014). Posterior odds with a generalized hyper-g prior. *Econometric Reviews*, 33, 251-269. 査読有 doi:10.1080/07474938.2013.807181
- ⑤ Yuzo Maruyama and William E. Strawderman (2013). Improved robust Bayes estimators of the error variance in linear models. *Journal of Statistical Planning and Inference*, 143, 1091-1097. 査読有 doi:10.1016/j.jspi.2013.01.007

[学会発表] (計 5 件)

- ① Yuzo Maruyama, Harmonic Bayesian prediction under alpha-divergence. 9th International Conference of the ERCIMWG on Computational and Methodological Statistics (CMStatistics 2016), University of Seville, Spain. 2016 年 12 月 10 日
- ② Yuzo Maruyama, Harmonic Bayesian prediction under alpha-divergence. International Conference on Statistical Distributions and Applications (ICOSDA 2016), Crowne Plaza, Niagara Falls, Canada. 2016 年 10 月 15 日
- ③ Yuzo Maruyama, Inadmissibility of the best equivariant predictive density in the unknown variance case. 2015 ICSA

China Statistics Conference, Shanghai Jianguo Hotel, China. 2015年7月7日

④ Yuzo Maruyama, L_p -norm based James-Stein estimation with minimaxity and sparsity. 3rd Meeting on Statistics, Athens University, Greece. 2015年6月25日

⑤ Yuzo Maruyama, Posterior inference and model selection of Bayesian probit regression. Statistics department seminar, Università Cattolica del Sacro Cuore, Italy. 2013年9月17日

[図書] (計 0 件)

[産業財産権]

○出願状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号:

出願年月日:

国内外の別:

○取得状況 (計 0 件)

名称:

発明者:

権利者:

種類:

番号:

取得年月日:

国内外の別:

[その他]

ホームページ等

<http://home.csis.u-tokyo.ac.jp/~maruyama>

6. 研究組織

(1) 研究代表者

丸山 祐造 (MARUYAMA, Yuzo)

東京大学・空間情報科学研究センター・准教授

研究者番号: 30304728

(2) 研究分担者

()

研究者番号:

(3) 連携研究者

()

研究者番号:

(4) 研究協力者

()