

博士論文の内容と審査結果の要旨

計算機システムと計測・測定技術の高度な発展は、自然科学、社会科学、産業界のあらゆる分野で大規模・高次元データの獲得と蓄積を可能とし、多様な形式でデータベース化されつつある。観測・測定データに内包される有用な情報とパターン抽出に当たっては、対象となる現象のモデリングが不可欠である。一般に、回帰モデリングとは、現象の結果とそれを引き起こす様々な要因を結びつけて、複雑な自然現象や社会現象の解明と予測・制御を目的とした分析手法である。このモデリングのプロセスにおいて、現象を捉える最適なモデルの推定と評価・選択が適切な推論や新たな知識発見のための鍵を握る。

しかしながら、多数の要因を含む高次元データを分析するためのモデルは必然的に大規模なものとなり、従来の分析手法によるモデルの推定、評価・選択のプロセスは有効に機能せず、現象の解明と予測には繋がらない。この問題の解決に向けて一つの方向性を示したのが、Tibshirani (1996, *J. Roy. Stat. Soc. Ser. B*) によって提唱された **lasso** であった。その特徴は、回帰係数の一部を完全に 0 と推定することから、モデルの推定と変数選択を同時に実行できる点にあった。

Lasso は、損失関数にパラメータの絶対値の和を正則化項として課した目的関数の最小化による正則化推定法で、この研究に端を発し、生命科学、システム工学、医用画像工学など科学の様々な分野で応用が試みられてきた。適用分野の広がりや問題提起へと繋がり、パラメータの L_1 ノルム正則化項を損失関数に課するという **lasso** の基本的な考え方を拡張、発展させ、各分野のデータの特長、分析目的に合わせて様々な L_1 型正則化回帰モデリングの提唱と推定アルゴリズムの開発研究が集中的に行われてきた。

本論文では、 L_1 型正則化回帰モデリングに対して、データからの情報に加えて、蓄積された知識を事前情報としてベイズアプローチによってモデルに同化させるベイズモデリングについて、理論的・数値的両側面に渡って研究し、新たな計算アルゴリズム、モデリング手法を提唱した。

第 2 章では、説明変数の個数がデータ数を超える線形回帰モデルの推定と変数選択を同時に実行可能な様々な L_1 型正則化法の性質と特徴を理論的、実際の側面から述べている。第 3 章では、本論文の主要なテーマである L_1 型正則化法をベイズの枠組みで捉えるために必要な事前分布、事後分布の定式化と推定アルゴリズムについて述べ、問題点を提起している。代表的な研究が、 L_1 ノルム正則化法に対してラプラス分布型のパラメータの絶対値に基づく制約を課した確率分布を導入した **Bayesian lasso** (Park and Casella, 2008; *J. Amer. Stat. Assoc.*) と呼ばれる手法である。ベイズアプローチによる **lasso** 推定は、最大事後確率 (MAP; maximum a posterior) 推定と捉えることができ、階層型ベイズモデルを構

成し、事後分布の推定に **Gibbs sampler** を用いることで汎用性の高い手法となった。これによって、マルコフ連鎖モンテカルロ法によるモデルの推定が可能となったが、**Bayesian lasso** は、スパース推定の最大の特徴である回帰係数を正確に 0 と推定し、不要な変数をモデルの中から除去する機能を有していないことを指摘し、新たな計算アルゴリズムの開発研究の必要性が生じた。

第 4 章では、この問題に取り組み、**Bayesian lasso** に対する新たなスパースアルゴリズムを提唱し、線形回帰モデルの回帰係数の推定値の一部を正確に 0 と推定することに成功した。提案した計算アルゴリズムは、シミュレーション実験によって予測や推定に与える影響等について検証し、さらに **Bayesian elastic net** 等、様々な L_1 型正則化推定法への適用可能性について検討した。

一般に、 L_1 型正則化推定法は、損失関数に正則化項を課した目的関数の最小化に基づく推定法であるが、0 と推定する回帰係数の個数は損失関数と正則化項のウェイトを調整するパラメータの値に大きく依存する。このため調整パラメータの値の決定法がモデリングの過程で重要な役割を果たす。本論文では、ベイズ型予測分布の良さを測るためのモデル評価基準である予測情報量規準を適用することによって、調整パラメータの値を決定する方法を提唱した。その特徴は、**Laplace** 分布を回帰係数の事前分布とする **Bayesian lasso** の適用上の問題点を、**Kullback-Leibler** 情報量で **Laplace** 分布を正規近似することで克服し、一つの決定法を提示したといえる。

ベイズアプローチによる様々な L_1 型正則化推定は、最大事後確率推定が基本となり、その効率的かつ安定した推定法を組み込むことが重要となる。この問題に対して、本論文では積分計算を離散近似し、ニュートン・ラフソン法による数値計算と融合した計算アルゴリズムを提唱し、シミュレーション実験を通してその有効性を立証した。

一連の研究は、現在、集中的に研究が行われつつあるスパースモデリングの研究進展に貢献したことから、2011 年度統計関連学会連合大会コンペティション最優秀報告賞、2012 年度中央大学学会会長賞、2013 年度日本計算機統計学会奨励賞、2014 年度第 8 回日本統計学会春季集会優秀発表賞を受賞した。

以上のように、本博士論文は、データからの情報に加えて蓄積された知識を L_1 ノルム事前分布としてベイズアプローチによってモデルに同化させた **Bayesian lasso** と呼ばれる推定法について、理論的・数値的両側面から研究に取り組み、有効な推定法を提唱した価値ある業績と認められる。

よって、本論文は博士（理学）の学位を授与するに十分なものであると認める。