

令和元年6月6日現在

機関番号：17102

研究種目：若手研究(B)

研究期間：2015～2018

課題番号：15K15949

研究課題名（和文）スパース正則化法に基づく探索的構造方程式モデリング

研究課題名（英文）Exploratory structural equation modeling via sparse regularization

研究代表者

廣瀬 慧 (Hirose, Kei)

九州大学・マス・フォア・インダストリ研究所・准教授

研究者番号：40609806

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）：本研究課題の成果は、主にスパース正則化法による因子分析の新たな手法の提案である。まず、因子分析における正則化が、因子回転の一般化であることを理論的に示した。その結果、Lassoタイプの正則化と因子回転の関係性の考察、Prenet (Product elastic net) 正則化などの新たな正則化法の提案、さらには因子回帰モデルへの拡張など、様々な研究へと発展した。

研究成果の学術的意義や社会的意義

因子回転は50年以上前から使われている古い手法であるが、それを一般化した正則化法を使うことにより、高次元データのスパース推定や完全単純構造によるクラスタリングなど、今までできなかったような解析ができるようになった。また、本研究課題でRパッケージを作成して公開したことにより、誰でも容易に正則化因子分析をできるようになった。

研究成果の概要（英文）：The main contribution based on this research fund is that we have proposed various methodologies via sparse estimation in factor analysis. Specifically, we have introduced a variety of regularization methods, such as Prenet (product elastic net) regularization. These methods are constructed by our theory that the regularization in factor analysis is a generalization of factor rotation.

研究分野：統計科学

キーワード：スパース正則化 Prenet正則化 ロバスト推定

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

多様な構造を呈するデータから有効な情報を抽出するためには、複雑・高度な統計モデリングが必要である。このことを達成するための解析法として、きわめて柔軟・多様なモデリングを可能とする構造方程式モデリングに着目した。構造方程式モデリングは、回帰分析や因子分析を含む柔軟なモデルで、データから変数間の複雑な関係性をモデル化することができる。しかし、この方法は、モデルがあまりに複雑なため、分析者の主観が必要となる。そこで、申請者は、分析者の主観をできるだけ必要のない新たな解析法の開発が必要だと考えた。

2. 研究の目的

本研究の目的は、できるだけ複雑なモデルを分析者の主観を極力入力しないで推定する手法を提案することであった。そこで、仮説を立てずにモデルを自動的に探索することのできる、探索的構造方程式モデリングを提案することを考えた。

3. 研究の方法

データから客観的にパラメータを0と推定することのできる、 L_1 正則化法を適用することを考えた。まずは、構造方程式モデリングの特殊な場合における因子分析モデルにおいて、 L_1 正則化法が識別性の問題を対処できるのか、どのくらいスパース推定できるのかなど、試してみることから考えた。また、パラメータを推定する計算アルゴリズムとして、MMアルゴリズムや座標降下法など、因子分析とLassoのアルゴリズムを組み合わせることを考えた。

4. 研究成果

(1) lassoに基づく因子分析の正則化法は、2011年に提案された、既存の手法であったが、実際に動かしてみると、lassoでは正しく因子を推定できないことが多いことがわかった。そこで、まずはMinimax Concave正則化(MCP)などの非凸正則化法を用いることを提案した。数値実験を行ったところ、MCPはlassoよりもスパースに推定できた。この内容は、計算機統計の分野で一流雑誌であるStatistics and Computingに掲載された(雑誌論文[10])。

さらに、正則化と因子回転の関係性を調べるために、因子回転の基準に基づいた正則化法の提案を試みた。まず、Varimax回転などのスタンダードな回転の一般化を試みたが、たしかに一般化できるものの、通常のVarimaxとほぼ同じ結果が得られただけで、特に特筆すべき結果は得られなかった。しかし、パラメータの積に基づくペナルティが有効であることを発見し、Prenet(Product elastic net)と呼ばれるペナルティを提案した。本ペナルティは、因子回転で用いられるQuartimin法を拡張したものである。実際、この方法は、適当な正則条件のもとでQuartimin回転による最適値に漸近的に一致することを示した。また、Prenet正則化が完全単純構造を持つことを示し、それがk平均法の一般化であることを示した。このように、因子分析とクラスタリングの関係性を明らかにした論文は、申請者の知る限りほとんど存在しなかった。

Prenet正則化法の有効性を検証するため、fMRIデータと大規模アンケートデータへ適用した。その結果、fMRIデータに対して、従来法であるWard法よりも良い結果が得られることを数値的に示した。また、大規模アンケートデータに対しては、従来のQuartimin回転よりも単純完全構造に近く、解釈しやすい因子負荷行列が推定でき、さらにはあてはまりも比較的良好であることを確認できた。

(2) 因子分析モデルを拡張させたモデルとして因子回帰モデルがある。これは、複数の説明変数を一つの因子としてまとめ、その因子を使って回帰を行うという方法である。本研究では、因子回帰モデルにおける正則化法を提案し、その理論的性質を調べた。回帰モデルでは、予測が重要になるが、予測の観点から推定量の性質を調べた文献は申請者の知る限り存在しなかった。そこで、予測の観点から推定量の性質を調べたところ、lassoの推定量のバイアスが予測には大きく影響せず、かつ多重共線性に対処することができるという知見が得られた。本研究内容は、Statistical Papersに掲載された(雑誌論文[2])。

(3) ガウシアングラフィカルモデルは、共分散逆行列のスパース性を用いてネットワークを推定するモデルである。現実世界では、スケールフリーネットワークと呼ばれる、「ハブ」を持つネットワークが多いが、このようなネットワークを普通のlassoで推定することは難しい。そこで、ハブを持つ場合にうまく推定できる新たなペナルティを考案した。本研究内容は、Journal of the Japanese Society of Computational Statisticsに掲載された(雑誌論文[9])。

(4) ガンマダイバージェンスに基づくロバストかつスパースなグラフィカルモデルを考案し、その性質を調べた。まず、提案法が、再下降性と呼ばれる、ロバスト推定の望ましい性質を満たすことがわかった。また、先行研究と比べ、パフォーマンスが圧倒的に良いことを数値的に示した。さらに、実データ解析においても、提案法が、適切に外れ値に対処できることがわかった。本研究内容は、Journal of Multivariate Analysisに掲載された(雑誌論文[4])。また、国内外の学会で講演した。

(5) ロバストな相対誤差推定に関する研究を提案した。ロバスト推定を実現するため、ガン

マ尤度を用いた推定を行った。推定値を求めるために、MM アルゴリズムに基づく推定を行った。提案手法を電力消費量データに適用したところ、電力消費量が極端に小さな値を取るとき、従来の相対誤差推定だとその値に引っ張られて不安定になるが、提案法を使うと、その影響が取り除かれることを確認した。また、交差検証法を適用することにより、適切なガンマの値を推定できることがわかった。この内容は、国際誌 *Entropy* (オープンジャーナル) に掲載された (雑誌論文[1])。さらに、国内外の学会・シンポジウムでこの内容を講演した。

5 . 主な発表論文等

〔雑誌論文〕(計 10 件, すべて査読あり)

- [1] Hirose, K. and Masuda H. Robust relative error estimation. *Entropy*, 20(9), 632, 1-24, 2018.
- [2] Hirose, K. and Imada, M. Sparse factor regression via penalized maximum likelihood estimation. *Statistical Papers*, 59(2), 633–662, 2018.
- [3] Dolinský, J., Hirose, K. and Konishi, S. Readouts for Echo-State Networks Built using Locally Regularized Orthogonal Forward Regression. *Journal of Applied Statistics*, 45(4), 740-762, 2018.
- [4] Hirose, K., Fujisawa, H. and Sese, J. Robust sparse Gaussian graphical modeling. *Journal of Multivariate Analysis*, 161, 172-190, 2017.
- [5] Imada, M., Hirose, K., Yoshida, M., Sunyong, K., Toyozumi, N., Lopez, G. and Kano, Y. An Interpersonal Sentiment Quantification method applied to Work Relationship Prediction. *NTT Technical Review* 15(3), 2017.
- [6] Yamamoto, M., Hirose, K., Nagata, H. Graphical tool of sparse factor analysis. *Behaviormetrika*, 44(1), 229-250, 2017.
- [7] 廣瀬慧. スパースモデリングとモデル選択. *電子情報通信学会誌*, 99 巻, 5 号, 392-399 項, 2016 年.
- [8] Hirose, K., Kim, S., Kano, Y., Imada, M., Yoshida, M., and Matsuo, M. Full information maximum likelihood estimation in factor analysis with a large number of missing values. *Journal of Statistical Computation and Simulation*, 86(1), 91-104, 2016.
- [9] Hirose, K., Ogura, Y. and Shimodaira, H. Estimating Scale-Free Networks via the Exponentiation of Minimax Concave Penalty. *Journal of the Japanese Society of Computational Statistics*, 28, 139-154, 2015.
- [10] Hirose, K. and Yamamoto, M. Sparse estimation via nonconcave penalized likelihood in a factor analysis model. *Statistics and Computing*, 25(5), 863-875. 2015.

〔学会発表〕(計 20 件)

- [1] 廣瀬慧, 増田弘毅. 相対誤差に基づく回帰モデルのロバスト推定. 2018 年度統計関連学会連合大会. 中央大学. 2018 年 9 月 9 日～13 日.
- [2] 廣瀬慧, 寺田吉彦. 因子分析における単純構造推定のための正則化法とその応用. 日本行動計量学会 第 46 回大会. 慶應義塾大学. 2018 年 9 月 3 日～6 日.
- [3] 廣瀬慧. Prenet 正則化法による単純構造推定. 日本地球惑星科学連合 2018 年大会. 幕張メッセ. 2018 年 5 月 20 日～24 日.
- [4] Hirose, K. and Terada, Y. Prenet Penalization in Factor Analysis and its Applications. International Conference on Advances in Interdisciplinary Statistics and Combinatorics. University of North Carolina at Greensboro, USA October 5-7, 2018.
- [5] Hirose, K. and Terada, Y. Simple structure estimation via prenet penalization in factor analysis model. The 2nd International Conference on Econometrics and Statistics (EcoSta 2018). City University of Hong Kong, Hong Kong 19-21 June 2018.
- [6] Hirose, K., and Terada, Y. Estimation of well-clustered structure via penalized maximum likelihood method in factor analysis model. 10th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2017). University of London, London, UK. December, 2017.
- [7] 山口尚哉, 廣瀬慧, 堀磨伊也, 出口喜也. 電力取引市場における電力調達の最適化. 統計関連学会連合大会. 南山大学. 2017 年 9 月 5 日～8 日.
- [8] 小家亜斗吏, 廣瀬慧. 群の数が多い場合における多群線形判別. 統計関連学会連合大会. 南山大学. 2017 年 9 月 5 日～8 日.
- [9] 廣瀬慧, 寺田吉彦, 因子分析における単純構造推定のための正則化法. 統計関連学会連合大会. 南山大学. 2017 年 9 月 5 日～8 日.
- [10] Hirose, K. Perfect simple structure estimation via extension of quartimin criterion. Conference of the International Federation of Classification Societies (IFCS 2017), Tokai University, Japan. August, 2017.
- [11] Hirose, K., Fujisawa, H. Robust estimation for high-dimensional Gaussian graphical models.

The 1st International Conference on Econometrics and Statistics (EcoSta 2017), Hong Kong University of Science and Technology, Hong Kong, June, 2017.

[12] 廣瀬慧. 超高次元データの統計解析における最適化問題. OR九州支部. 2017年3月27日.

[13] 廣瀬慧, 正則化法によるスパース推定とその応用. 電気学会. 全国大会. 富山大学. 2017年3月17日.

[14] 廣瀬慧. 進化計算研究者のための統計解析及び機械学習. 第12回進化計算学会研究会. 九州大学. 2017年3月13日~3月14日

[15] 廣瀬慧. スパース推定法による高次元データ解析. 統計科学研究会. 2016年12月2日.

[16] Hirose, K. Robust estimation for high-dimensional graphical models. Forum "Math-for-Industry" 2016. Science and Engineering Centre, Garden's Point Campus, Queensland University of Technology, Brisbane, Australia. November 21-23, 2016

[17] 廣瀬慧. Rパッケージ fanc --グラフィカルツールを用いた新たな因子分析-- . 数学協働プログラム「確率的グラフィカルモデルの産業界への応用」. 慶應義塾大学日吉キャンパス. 2016年11月10日~11月11日

[18] 廣瀬慧. Pernet ペナルティによる因子分析の単純構造の推定. 大規模統計モデリングと計算統計 III. 東京大学. 2016年9月27日~28日.

[19] Hirose, K. Sparse factor model via regularization and its extension to regression analysis. 日本計算機統計学会第30回大会. ハートピア京都. 2016年5月19日

[20] Hirose, K., Ogura, Y. and Shimodaira, H. Estimation of Scale-Free Networks with the Exponentiation of Minimax Concave Penalty. The 9th Conference of the Asian Regional Section of the IASC (IASC-ARS 2015), National University of Singapore, Singapore. December, 2015.

〔図書〕(計1件)

[1] 川野秀一, 松井秀俊, 廣瀬慧 (2018)
“スパース推定法による統計モデリング”
共立出版. 168 ページ.

〔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:
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:
国内外の別:

〔その他〕

ホームページ等

<http://www.keihirose.com>

6. 研究組織

(1) 研究分担者

研究分担者氏名:

ローマ字氏名:

所属研究機関名:

部局名:

職名：

研究者番号（8桁）：

(2)研究協力者

研究協力者氏名：山本 倫生，永田 晴久

ローマ字氏名：Michio Yamamoto, Haruhisa Nagata

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。