

令和 6 年 6 月 14 日現在

機関番号：62603

研究種目：研究活動スタート支援

研究期間：2022～2023

課題番号：22K20178

研究課題名（和文）関係データに関するベイズ縮小推定

研究課題名（英文）Bayesian shrinkage estimation for relational data

研究代表者

湯浅 良太（Yuasa, Ryota）

統計数理研究所・統計思考院・助教

研究者番号：90964487

交付決定額（研究期間全体）：（直接経費） 1,600,000円

研究成果の概要（和文）：関係データは行列やテンソルの形で表すことが出来る。まず、関係データを表す最も簡単な形である行列データに対して、ベイズ縮小推定量に関する研究と重み付き縮小推定量に関する研究を行った。また、人口動態データでは都道府県ごとに男女別の流入流出者数が何年にも渡って得られており、その分析を行うためにはテンソルデータの時系列モデルが必要となる。そこで、テンソルデータでの時系列構造を活かしたベイズモデリングに関する研究を行った。

研究成果の学術的意義や社会的意義

ベクトルの場合にはベイズ縮小推定量の研究が多くなされているが、行列データの場合には複雑になることもあり得られている結果は限定される。行列データの場合にシミュレーションに基づく数値実験によって評価するだけのものが多かったが、理論的に評価を行ったことで、行列データの分析に際して一定の保証を持った方法を提案している。

テンソルデータに関して、パラメータの解釈については考慮しないようなモデルがほとんどであったが、パラメータの一意性に関して考慮することで、予測や推定の精度向上のためだけでなく、解釈のためにも用いることができるようになった。

研究成果の概要（英文）：Relational data can be represented in the form of matrices or tensors. First, research was conducted on Bayesian shrinkage estimators and weighted shrinkage estimators for matrix data, which is the simplest form to represent relational data. Additionally, demographic data include the number of incoming and outgoing individuals by gender for each prefecture over many years. We need a time series statistical model for tensor data to analyze it. Therefore, research was carried out on Bayesian modeling that leverages the time series structure in tensor data.

研究分野：数理統計

キーワード：縮小推定 ベイズ統計

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

各都道府県間の人口移動がそれぞれ何人であるかを都道府県間の関係とみなす事ができる。そうした関係に着目したデータを関係データと呼ぶ。月毎に都道府県間の人口移動が公表されている場合、1 か月分のデータは行列データとして、複数月分のデータはテンソルデータとして自然にみなせる。こうしたデータを要素毎にばらばらに分析するのではなく、行列やテンソルの構造を活かした統計モデルを用いて分析する事で、関係データが持つ情報を活かせる。特に、ベイズモデルを用いて人口移動の傾向をモデルに取り入れる事により、共通する傾向を持つデータを相補的に活かす縮小推定が可能となる。しかし、ベクトルデータに対する解析手法は多く研究されているのに対し、行列やテンソルに対する研究は複雑になってしまうこともあり、限定的である。特に、推定精度はシミュレーションによるだけであったり、予測や推定の精度を向上させても統計モデルに現れるパラメータの解釈が不可能であったり、といった、分析者にとっての問題がある。

2. 研究の目的

本研究では、行列やテンソルの構造を活かした統計モデルを用いて分析する事で、関係データが持つ情報を活かすことを考える。特に、推定精度について理論的保証を持ったベイズ縮小推定量の提案や、人口移動を含む経済統計で生じる時系列構造を想定したテンソルデータに対する、パラメータの一意性が保証されたベイズモデルの提案を行う。

3. 研究の方法

ベクトルの場合から行列への拡張のために、行列式を用いる。行列正規分布に対しては、平均行列には行列式を用いた事前分布、分散共分散行列にはウィシャート分布による事前分布を用いる。統計的決定理論の枠組みでリスクの評価を行うことで推定精度の評価を行う。

また、構造を十分に活かすために、ベクトル型の縮小推定量と行列型の縮小推定量の重み付きの和を考え、その重みがデータによって選択されるような推定量を提案し、そのリスク評価を行う。

時系列テンソルデータに対しては、Tucker 分解と呼ばれるテンソル分解を用いて、ベイズモデルを考える。中でも HOSVD と呼ばれる分解を用いることで、一意性を考慮したモデルを用いることで解釈性を保ったベイズモデルを提案する。

4. 研究成果

(1) 行列データのベイズ縮小推定

行列データに対する縮小推定量を考える際には、分散共分散行列を既知として扱うことが多いが、これは理論を簡単にするためであり、分散共分散行列を推定するデータが少ないときには望ましくない。そこで、分散共分散行列を未知とした上で、平均行列には行列式を用いた事前分布、分散共分散行列にはウィシャート分布による事前分布を用いて、ベイズ推定量を導出した。さらに、推定精度を理論的に示すために、リスクの不偏推定量を導出し、評価を行うことにより、推定精度の改善を示した。

(2) 重み付きの縮小推定量

行列構造が明確に意味を持つ場合には、行列型の縮小推定量を用いることがよいが、構造の意味が不明確なときにはベクトル型の縮小推定量を用いるべきか、行列型の縮小推定量を用いるべきか明らかでない。そこで、データの持つ構造を十分に活かすために、ベクトル型の縮小推定量と行列型の縮小推定量の重み付きの和を考えた。重みの決定のためには、リスクの不偏推定量の最小化による方法と、ベクトル型の事前分布と行列型の前分布の混合分布を用いた経験ベイズ推定量を用いる方法の 2 つの方法を提案した。それぞれの方法に対し、推定精度の理論的な改善を示した。特に、分散共分散行列が既知の場合と未知の場合のどちらも考えており、分散共分散行列が未知の場合には、分散共分散行列を推定するためのデータが分散共分散行列の次元より少ない場合も含めて改善を示した。

(3) 時系列テンソルデータに対するベイズモデル

時系列構造を持ったテンソルデータに対して、Tucker 分解と呼ばれる分解を用いて、状態空間モデルのような時系列構造をコアテンソルや列直交行列に対して課したベイズモデルを考え

た. Tucker 分解の中でも HOSVD と呼ばれる一意性を持った分解を用いることで、因子の解釈が可能となっている. しかし HOSVD ではコアテンソルに対する順序制約や直交制約が課されることで、サンプリングをする際の空間が制約され、サンプリングが困難になる. そこで、SMC サンプラーと呼ばれる方法を用いて、制約が弱い状態からサンプリングを始めることで、サンプリングを可能とした.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Yuasa Ryota, Kubokawa Tatsuya	4. 巻 222
2. 論文標題 Generalized Bayes estimators with closed forms for the normal mean and covariance matrices	5. 発行年 2023年
3. 雑誌名 Journal of Statistical Planning and Inference	6. 最初と最後の頁 182 ~ 194
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.jspi.2022.06.007	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 Yuasa Ryota, Kubokawa Tatsuya	4. 巻 194
2. 論文標題 Weighted shrinkage estimators of normal mean matrices and dominance properties	5. 発行年 2023年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 105138 ~ 105138
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.jmva.2022.105138	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 -

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 湯浅 良太
2. 発表標題 行列正規分布の縮小推定とその適用
3. 学会等名 2022年度 統計関連学会連合大会
4. 発表年 2022年

1. 発表者名 Ryota Yuasa
2. 発表標題 Bayesian Tucker decomposition model with time varying factor matrices
3. 学会等名 The 16th International Conference of the ERCIM WG on Computational and Methodological Statistics
4. 発表年 2023年

1. 発表者名 湯浅 良太
2. 発表標題 制約下での行列平均の特異値縮小推定
3. 学会等名 2023年度 統計関連学会連合大会
4. 発表年 2023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
---------	---------