

令和 2 年 6 月 11 日現在

機関番号：20103

研究種目：基盤研究(C) (一般)

研究期間：2016～2019

課題番号：16K00051

研究課題名(和文) 一般化ダイバージェンスを用いた非確率・拡張モデルの推定理論とその応用

研究課題名(英文) Estimation with extended models using generalized divergences and its applications

研究代表者

竹之内 高志 (Takenouchi, Takashi)

公立はこだて未来大学・システム情報科学部・准教授

研究者番号：50403340

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：離散空間における確率モデルは、確率であることを要請するために正規化項の計算が必須であるが、高次元離散空間においては正規化項の計算はしばしば指数オーダーの計算量が必要となる。本研究では確率モデルではなく、正規化されていない非確率モデル(拡張モデルと呼ぶ)を用い、経験局所化、同次ダイバージェンスと組み合わせることで、正規化項の計算をすることなくクラメル・ラオの下限を達成する推定量を構成することが可能であることを示した。また、一般化ダイバージェンスを用いることで多値判別手法を扱うための統一的な枠組みや、ノイズに対して頑健な非負値行列因子分解法を構築した。

研究成果の学術的意義や社会的意義

提案した推定量は正規化項の計算が不要であるうえに、漸近的にクラメル・ラオの下限を達成するため、通常の推定量と比較して数十-数百分の一の計算コストで最尤推定量に匹敵する性能を達成可能である。また、多値判別手法を扱うための統一的な枠組みは多くの従来手法を特殊ケースとして含むため、性能に関する理論的な考察や比較が用意になった。非負値行列因子分解法については、再下降性と呼ばれる性質を手法に付与することが出来たため、ノイズに対して強力な頑健性をもたせることが可能となった。

研究成果の概要(英文)：The estimation for probabilistic models in a discrete space requires the computation of the normalization term to be a probability. However, in high-dimensional discrete spaces, the computation of the normalization term often requires a computational complexity of exponential order. In this study, we use an un-normalized non-probabilistic model (called the extended model) instead of a probabilistic model, and proposed an estimator by combining the extended model with the technique of empirical localization and homogeneous divergence, which can be constructed without the normalization term and asymptotically achieves the Cramer-Rao bound. We also construct a unified framework for handling multi-class classification methods and robust non-negative matrix factorization algorithm, by using generalized divergences.

研究分野：統計的機械学習

キーワード：一般化ダイバージェンス 離散確率モデル 非負値行列因子分解

様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

確率モデルは様々な現象・データを表現するための有効な手段であり、広く用いられている。確率モデルは、正規分布をはじめとする連続値を表現するための確率モデルと、ベルヌーイ分布や(隠れ変数を持つ)ボルツマンマシンなどに代表される離散的な現象を表現するための離散確率モデルに大別されるが、本研究では特に離散確率モデルに着目した。離散確率モデルは、パターン認識問題や、近年様々な分野で目覚ましい成果を上げているディープニューラルネットワークで用いられる制約付きボルツマンマシンなどで頻繁に用いられている。

データを適切に記述するパラメータを推定するための手段として、現在提案されている多くの統計的手法、機械学習アルゴリズムは最尤推定法、あるいは尤度関数に事前確率を組み込んだベイズ推定法の枠組みを用いている。これらの推定法はカルバック・ライブラー・ダイバージェンスの最小化問題として記述することができ、一致性や有効性など、好ましい統計的性質を持つことが知られている。しかし、これらの方法で離散確率モデルのパラメータを推定する際にしばしば計算量が問題となる。最尤推定法に類する方法では陽にパラメータを求めることはできないため、勾配法等を用いてパラメータを推定することになるが、パラメータの更新を行う度に指数オーダーの計算量を必要とする正規化項の計算が必要となる。この計算量の問題を解決するために、現状では MCMC サンプリングなどを用いた近似計算が用いられることが多い。その一方で、離散的現象を記述する際に、確率モデルではなく正規化されていない拡張モデルを用いるアプローチが存在する。主たる例として、判別問題において用いられるブースティング法がある。ブースティング法は計算量を抑えつつ一致性を持つ判別器を構成することが可能なため、様々な応用分野で大きな成功を収めている。ブースティング法は、ある一般化ダイバージェンスとそれに対応する拡張モデルを用いて定式化することが可能であり、様々な拡張(計算量の削減、ノイズに対する頑健性)を行うことが可能であることが知られていた。別の例としては、ボルツマンマシンを含む特定のクラスの離散確率モデルに対して、一般化ダイバージェンスの一種である板倉・斉藤ダイバージェンスと指数型の拡張モデルを用いることで一致性を持つ推定量を構成することが可能であることがわかっていた。確率モデルではなく拡張モデルを用いることの最大のメリットは、パラメータを推定する際に正規化項の計算を行う必要がないため、必要となる計算量が指数オーダーからサンプル数オーダーに削減可能となるという点であった。

### 2. 研究の目的

多様な離散的現象の予測・推論のために、確率モデルの代替として、正規化されていない拡張モデルを用いて効率的なパラメータ推定を可能とする枠組みを確立し、その統計的性質を明らかにする。

### 3. 研究の方法

上述の目的を達成するために、拡張モデルを用いた離散確率モデルのパラメータ推定に対して、特に有用な性質を持つことが明らかになっている板倉・斉藤ダイバージェンスを含むクラスの一般化ダイバージェンスや、漸近効率の面で最尤推定量と同様の効率を見込むことが可能なアルファ・ダイバージェンス、ノイズに対して特に頑健であるガンマ・ダイバージェンスに着目する。これらのダイバージェンスを用いて、一致性、あるいは、効率性や頑健性などの好ましい性質を持つ、拡張モデルによるパラメータ推定量を構成する。

#### 4. 研究成果

(a)高次元離散空間上の確率モデルに対する推定量を構築するために、本研究では確率モデルではなく、正規化されていない非確率モデル(拡張モデルと呼ぶ)を用いるアプローチを採用した。通常のカルバック・ライブラーダイバージェンスを用いた推定は拡張モデルの適用を想定していないため、同次性を持つダイバージェンスを用いた推定法について検討を行い、考案した推定量が、特定の条件下で凸性などの最適化する際に有用な性質を持つことを示した。また漸近的な解析により、拡張モデルと同次性を持つダイバージェンスを組み合わせることでクラメール・ラオの下限を達成する推定量を構成することが可能であることを示した。数十-数百分の一の計算コストで最尤推定量に匹敵する性能を達成可能であることを人工データを用いた小規模な実験で確認している。また、カルバック・ライブラーダイバージェンスを含む一般的なクラスであるプレグマンダイバージェンスを更に変形した変形プレグマンダイバージェンスと拡張モデルを組み合わせ推定を行うための枠組みを考案し、予備的な解析を行った。提案した推定量がフィッシャー一致性などの好ましい性質を持つことを示した。

(b)多値判別問題は機械学習における重要な問題の一つであるが、対象とするクラス数が大きくなると計算量が増大するという問題点を抱えている。そこで比較的簡単に構成することができる2値判別器を複数組み合わせることで多値判別器を構成するアンサンブル学習に基づくアプローチが採用されることが多い。本研究では、一般化ダイバージェンスを用いて2値判別器から構成される非確率モデルを統合するためのアンサンブル手法を提案し、その理論的性質についての考察を行った。一般化ダイバージェンスの採用により、多くの既存の2値判別器統合法を統一的に扱うことが可能となった。また、多くの既存の2値判別器統合法は多値ロジスティックモデルに特殊な制約を課したケースに該当することが明らかとなった。

(c)非負値行列因子分解法は特徴抽出のための代表的な手法である。一般的に普及しているアルゴリズムは行列とモデルの要素値同士の2乗誤差関数を最小化することで導出され、これは正規ノイズを仮定しカルバック・ライブラーダイバージェンスの最小化問題を考えることと等価である。このアルゴリズムは簡易な更新則が導出できる一方で、データに含まれるノイズに対する頑健性を持たないという欠点があった。この問題を解決するため、手法を頑健化するためのアプローチが種々提案されている。代表的なものは、行列の要素値とモデル値の間の距離尺度として一般化ダイバージェンスの一種であるbetaダイバージェンスやgammaダイバージェンスを直接用いる手法である。これに対し、本研究では行列の要素値とモデル値の誤差を統計モデルとして表現し、統計モデルの同定のために一般化ダイバージェンスを用いることで、従来の頑健化された手法よりもノイズに対して頑健な手法を提案することができた。理論的な考察を行うことで、提案した手法が再下降性と呼ばれる好ましい性質を持つことが明らかとなり、ある程度大きなノイズに対してもノイズ成分を無視して頑健な推定を行うことが可能であることが示された。また、従来の要素値の差異を直接測る手法は再下降性を持たないことも明らかになったため、統計モデルと一般化ダイバージェンスを組み合わせる提案法の方がより頑健な手法であることが示された。上記のような事柄に対し数値的な検証もを行い、理論によって示された性質が正しいことを確認した。

## 5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 K. Machida, and T. Takenouchi	4. 巻 2
2. 論文標題 Statistical modeling of robust non-negative matrix factorization based on $\alpha$ -divergence and its applications.	5. 発行年 2019年
3. 雑誌名 Japanese Journal of Statistics and Data Science	6. 最初と最後の頁 441-464
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s42081-019-00041-3	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 T. Takenouchi, and S. Ishii.	4. 巻 273(17)
2. 論文標題 Binary classifiers ensemble based on Bregman divergence for multi-class classification	5. 発行年 2018年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 424-434
掲載論文のDOI（デジタルオブジェクト識別子） <a href="https://doi.org/10.1016/j.neucom.2017.08.004">https://doi.org/10.1016/j.neucom.2017.08.004</a>	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 Takashi Takenouchi, and Takafumi Kanamori	4. 巻 18
2. 論文標題 Statistical Inference with Unnormalized Discrete Models and Localized Homogeneous Divergences.	5. 発行年 2017年
3. 雑誌名 The Journal of Machine Learning Research	6. 最初と最後の頁 1-26
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 T. Kanamori, and T. Takenouchi	4. 巻 95
2. 論文標題 Graph-based Composite Local Bregman Divergences on Discrete Sample Spaces.	5. 発行年 2017年
3. 雑誌名 Neural Networks	6. 最初と最後の頁 44-56
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.neunet.2017.06.005	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13件（うち招待講演 1件 / うち国際学会 3件）

1. 発表者名 竹之内高志
2. 発表標題 非正規化モデルを用いた推定量とその性質
3. 学会等名 応用統計学会（招待講演）
4. 発表年 2019年

1. 発表者名 谷本 啓, 坂井 智哉, 竹之内 高志, 鹿島 久嗣
2. 発表標題 Causal Outcome Prediction on Combinatorial Action Spaces
3. 学会等名 情報論的学習理論ワークショップ(IBIS2019)
4. 発表年 2019年

1. 発表者名 佐々木 博昭, 竹之内 高志, R. Monti, A. Hyvarinen
2. 発表標題 Robust contrastive learning and nonlinear ICA in the presence of outliers
3. 学会等名 情報論的学習理論ワークショップ(IBIS2019)
4. 発表年 2019年

1. 発表者名 M. Ishii, T. Takenouchi, and M. Sugiyama
2. 発表標題 Zero-shot Domain Adaptation Based on Attribute Information
3. 学会等名 The Eleventh Asian Conference on Machine Learning（国際学会）
4. 発表年 2019年

1. 発表者名 T. Takenouchi.
2. 発表標題 Parameter Estimation with Generalized Empirical Localization.
3. 学会等名 Geometric Science of Information (国際学会)
4. 発表年 2019年

1. 発表者名 竹之内高志
2. 発表標題 非正規化モデルを用いた推定法.
3. 学会等名 大規模複雑データの理論と方法総合的研究
4. 発表年 2018年

1. 発表者名 谷本啓, 山田聡, 竹之内高志, 鹿島久嗣.
2. 発表標題 Relaxing Imbalance with Positiveness
3. 学会等名 第 21 回情報論的学習理論ワークショップ (IBIS2018)
4. 発表年 2018年

1. 発表者名 江口真透, 竹之内高志
2. 発表標題 異なるモデルの尤度関数の結合
3. 学会等名 統計関連学会連合大会
4. 発表年 2018年

1. 発表者名 竹之内高志
2. 発表標題 Parameter Estimation with Deformed Bregman Divergence
3. 学会等名 Information Geometry and its Applications IV (国際学会)
4. 発表年 2016年

1. 発表者名 竹之内高志
2. 発表標題 変形ブレグマンダイバージェンスを用いたパラメーター推定
3. 学会等名 2016年度 統計関連学会連合大会
4. 発表年 2016年

1. 発表者名 金森 敬文, 竹之内高志
2. 発表標題 離散空間上のグラフ構造に基づく局所ブレグマンダイバージェンス
3. 学会等名 2016年度 統計関連学会連合大会
4. 発表年 2016年

1. 発表者名 竹之内高志
2. 発表標題 変形ブレグマン擬距離とその応用
3. 学会等名 第19回情報論的学習理論ワークショップ
4. 発表年 2016年

1. 発表者名 金森 敬文, 竹之内高志
2. 発表標題 グラフ上の局所プレグマンダイバージェンスによる統計的推論
3. 学会等名 第19回情報論的学習理論ワークショップ
4. 発表年 2016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
--	---------------------------	-----------------------	----