

## 科学研究費助成事業 研究成果報告書

令和 3 年 6 月 15 日現在

機関番号：62603

研究種目：基盤研究(C) (一般)

研究期間：2018～2020

課題番号：18K11208

研究課題名(和文)高次元母数の推定技法としての深層学習の理解と展開

研究課題名(英文) Understanding and developing deep learning as estimation procedures of the high-dimensional parameter

研究代表者

柳本 武美 (Yanagimoto, Takemi)

統計数理研究所・名誉教授

研究者番号：40000195

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：深層学習の成功は多様な改良の可能性を示している。本研究では尤度比統計量を基礎にして損失関数・ベイズモデルからの改良を企図した。活性化関数としての softmax, ReLU の選択は有効な母数の推定方法の選択と見なすことができる。その上で、これまで高次元母数モデルの推定技法を適用した。二点での展開を計った。まず活性化関数を ramp関数と見なして一般化させる点にある。Ramp関数の理解は分野により異なり、spline関数の分野と機械学習で異なる。更に、分布関数の裾の重さの視点を加えた。別の一点は、損失関数についてである。深層学習は分類問題を多項分布の出現確率の推定に還元させるからである。

研究成果の学術的意義や社会的意義

データサイエンスへの期待が高まる中で、新し手法としての深層学習の実用性が認められた。その適用範囲は従来の統計手法が及ばない領域を広く含んでいる。また、その基本的構造は従来の回帰分析の自然な拡張である。その意味でも多様な研究が求められる研究テーマである。その構造の解明と理解を深める研究が求められている

研究成果の概要(英文)：The success of the deep learning indicates its possible improvements in various ways. Improvements through the loss function and Bayesian models based on the likelihood ratio statistics are primary targets in the present study. The selection of activation functions, such as the softmax and the ReLU functions, is regarded as that of inferential procedures of parameters. The techniques of the estimation of a high-dimensional parameter are applied. Two approaches are employed. One is to generalize activation functions by regarding them as ramp functions. The understanding of the ramp function varies with different disciplines, such as the spline function and the machine learning. We add a view of the heaviness of tail of the distribution. Another approach pertains the loss function. The loss and the risk are the keys in the deep learning, since it is broken down into the estimation problem of the multinomial distribution.

研究分野：理論統計学・生物統計学

キーワード：Activation function LRT statistic Regression analysis

## 1. 研究開始当初の背景

(1) 統計的方法と深層学習には通常に考えられている以上に類似点がある。データから有用な情報を抽出しようとするために、確率モデルを仮定して機械的な操作により推論を行う機械学習は、本来的に統計的方法である。過去に細かい違いが妙に強調されたり、確率モデルが仮定されなかったりしてきた経緯とは大きく異なる。一方で、既に画期的な性能が示されているにも拘わらず、なお懐疑的な関連研究者が多い。特徴量の抽出とかデータの粗性のような分かり易さに欠けるとの批判がある。しかし、手続きの分かり易さよりも仮定と推論形式の明瞭さの方が重要な要素である。基礎的な研究が必要とされた。

## 2. 研究の目的

(1) 本研究では従来の高次元母数の推定法と深層学習と間に見られる共通点に着目して、その延長上に深層学習の推論構造を捉えてより深い理解を求める。そのための一つの重要な視点が、解析を支える学習データ集合の育成に着眼する。データそのものの検討に加えて、解析結果をも評価してデータ集合の育成を図る工夫と努力が欠かせない。深層学習がこれ迄手が着かなかった新しい形態のデータに強力な性能を示すことが明瞭になった。一方で意表を突くアイデアにと言うより、従来の手法の愚直な拡張である点を明確にすることが、この分野の発展に不可欠であることを示すことが大きい目的であった。

## 3. 研究の方法

(1) Softmax関数とReLU関数の役割を中心に研究し、学習データ集合の育成と結びつけて研究する。これらの活性化関数は母数の推定としての機能があり、統計的推定として現れる。Softmax関数の利用の前提として、分類問題を多項分布の出現確率の推定問題への還元がある。データ解析で最も基本的なモデルである回帰モデルの延長線上に捉えられる。

(2) 深層学習を含めて統計的手法は、解析手法と解析手法に対応したデータの収集法が密接に繋がっている。今日の凄まじいデータの変容に拘わらず、アカデミックの研究者の関心は専ら解析手法の開発に目が向いている。従って、解析結果に基づいた学習データ集合の育成について、汎化誤差の低減の観点から研究が必要である。深層学習では汎化誤差は常に意識されるが、過剰適合の排除が目的であって推定理論におけるリスクの減少である。汎化誤差には、学習データ集合と試験データとの乖離への配慮も重要である。

(3) 研究を進める過程で、損失関数を従来より深く考えるようになった。元より高次元母数を扱う研究では、統計量を導出するための損失関数と、導出した推定量を評価するための損失関数を柔軟に使い分ける必要性は認められていた。深層学習では母数の次元が巨大な上に Wasserstein 距離の役割が明瞭になってきたので、新しい視点からの研究が求められることになった。

## 4. 研究成果

(1) 活性化関数は回帰分析の非線形化の具体化である。同じ用語が用いられているが、中間層での活性化関数と最終層での活性化関数は異なる。この違いを明示することから研究を始めた。中間層での活性化関数は今日では ReLU 関数が広く用いられている。この関数のより深い理解

とその改良を当面の目標とした。

(2) ReLU 関数の理解に関しては幸い過去に関連研究がある (引用文献)。重要な違いは折れ線回帰関数の折れ点の重要性である。自然科学では自然の滑らかな変化を想定する。しかし、人間の認識は滑らかな自然を簡略化して理解する。滑らかなままでは扱えない。その中で、logistic 分布と Laplace 分布の役割の違いに注目することになった。Laplace 分布が従来看過された理由は密度関数に折れ点が存在することにあった。別の視点では回帰モデルの非線形化を対数尤度比との視点で捉えることであった。対数尤度比はモデルの比較の基本量であるから、情報の伝達過程での非線形化は対応していると期待される。具体的には

- 1) 対数尤度比から活性化関数を捉えて一般化を図る (Manuscript)
- 2) Laplace 分布の認識過程での役割を整理して開設する (雑誌論文 3)

を通じて理解を深め発展を試みた。この研究の基礎的裏付けとしての研究については a) 活性化関数と回帰関数 について口頭発表 (2018 年度統計関連学会連合大会)をおこない、b) Ramp 関数の性質を調べてその一部と口頭発表 (DSSV2019) した。また、Wasserstein 距離からの考察については内部の研究会で議論した。

(3) 本研究では解析方法とデータの収集についての俯瞰的な視点がある。文献上ではデータの爆発が強調されるが、成功事例を吟味すると解析に手法にとって好都合で質の高いデータが得られていることが見て取れる。形式的で発展性のない典型例とされる臨床評価を例にして深層学習と無作為化比較試験の共通点を指摘した (2019 年度統計関連学会連合大会)。深層学習が日常的な臨床行為の支援システムへの寄与が期待差されることを論じた。また、ホッケー・スティック法のような折れ線回帰モデルが、自然現象の記述ではなく人間の認知システムの記述として捉えることとしてその有用性を示した。

(4) 活性化関数としての softmax 関数は多項分布における自然母数から平均母数への変換に他ならない。言い換えると一般化線形モデルへの埋め込みである。この分野ではベイズ法の有用性が明瞭になってきている。深層学習の分野でも高い次元の母数を扱うので結局はベイズ法による解釈が見込まれている。これに関して

- 3) 複数ポアソン母数のベイズ推定に改良を加えた (雑誌論文 3)

Softmax 関数と直接関連する多項分布についての研究発表も行った (Bayes on the Beach 2019)。現状ではベイズ法の視点は必ずしも深層学習の手法の完全に取り入れられていない。力づくのトライアンドエラーの積み重ねである。超高次元母数モデルの下での母数推定推定の性能を示す基礎的な結果である。明示的な成果を得るには至らなかったが、Wasserstein 距離からの研究も行った。この距離は二つの同一測度空間の積空間上の同時測度を構築することがその基本問題である。この問題は近年注目を浴びている分布の相関構造の記述の特殊な場合である。この視点からの研究を見込んでいる。

#### <引用文献>

Yanagimoto, T. and Yamamoto, E. Estimation of safe doses: Critical review of the hockey stick regression method. *Env. Health Persp.* 32, 1979, 193-199,

## 5. 主な発表論文等

〔雑誌論文〕 計3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 2件）

1. 著者名 Yanagimoto, T. and Ohnishi, T.	4. 巻 2
2. 論文標題 A characterization of Jeffreys' prior with its implications to likelihood inference.	5. 発行年 2021年
3. 雑誌名 Pioneering Works on Distribution Theory: In Honor of Masaaki Sibuya, eds by N. Hoshino et al.	6. 最初と最後の頁 103-121
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 Ogura, T. and Yanagimoto, T	4. 巻 1
2. 論文標題 Bayesian estimator of multiple Poisson means assuming two different priors.	5. 発行年 2022年
3. 雑誌名 Commun. In Statist. - Simul. Comput,	6. 最初と最後の頁 0
掲載論文のDOI（デジタルオブジェクト識別子） 10.1080/03610918.2020.1861465	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 柳本 武美	4. 巻 2157
2. 論文標題 Laplace分布の再評価：ベイズ法と活性化関数から	5. 発行年 2020年
3. 雑誌名 RIMS Kokyuroku	6. 最初と最後の頁 83-96
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 無
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15件（うち招待講演 2件/うち国際学会 5件）

1. 発表者名 T. Yanagimoto and T. Ogura
2. 発表標題 Use of two different priors in an empirical Bayes estimator: Case of multiple Poisson means
3. 学会等名 4th EAC- ISBA（招待講演）（国際学会）
4. 発表年 2019年

1. 発表者名 T. Ogura and T. Yanagimoto
2. 発表標題 Innovative conjugate analysis of the unknown dimensional multinomial probabilities
3. 学会等名 4th EAC- ISBA ( 国際学会 )
4. 発表年 2019年

1. 発表者名 T. Yanagimoto and K. Ohkusa
2. 発表標題 Properties of the ramp function as an activation function in deep neural network
3. 学会等名 DSSV2019 ( 国際学会 )
4. 発表年 2019年

1. 発表者名 柳本 武美
2. 発表標題 DNN と RCT の共通点に見る統計的推測の要点
3. 学会等名 2019年度統計関連学会連合大会
4. 発表年 2019年

1. 発表者名 K. Tahata, R. Takami and T. Yanagimoto
2. 発表標題 On estimators of multinomial parameters using bayesian approach
3. 学会等名 Bayes on the Beach ( 国際学会 )
4. 発表年 2019年

1. 発表者名 柳本 武美
2. 発表標題 ベイズモデルの母数：事前分布と estimand
3. 学会等名 科研費シンポ秋田
4. 発表年 2019年

1. 発表者名 柳本 武美
2. 発表標題 Laplace 分布の再評価：ベイズ法と活性化関数から
3. 学会等名 RIMS共同研究
4. 発表年 2020年

1. 発表者名 柳本武美
2. 発表標題 活性化関数と回帰関数の性能と人の認知からの評価
3. 学会等名 2018年度統計関連学会連合大会
4. 発表年 2018年

1. 発表者名 小椋透 (三重大学)* 柳本武美
2. 発表標題 複数のポアソン分布の平均値の経験ベイズ推定における対数変換を用いた工夫
3. 学会等名 2018年度統計関連学会連合大会
4. 発表年 2018年

1. 発表者名 高見遼太（東京理科大学），柳本武美，田畑耕治（東京理科大学）
2. 発表標題 多項分布における自然母数の事後平均
3. 学会等名 2018 年度日本分類学会シンポジウム
4. 発表年 2018年

1. 発表者名 柳本武美
2. 発表標題 RCT と DNN が医療水準の向上を駆動する
3. 学会等名 科研費研究集会「多変量データ解析法における理論と応用」
4. 発表年 2018年

1. 発表者名 Yanagimoto, T. * and Ohnishi, T. (Kyushu University)
2. 発表標題 Conjugate analysis under Jeffreys' prior with its implications to likelihood inference
3. 学会等名 Pioneering Workshop on Extreme Value and Distribution Theories in honor of Prof. Sibuya (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 作村 建紀・柳本 武美
2. 発表標題 von Mises 分布における自然母数の事後平均のバイアスについて
3. 学会等名 2020年度統計関連学会連合大会
4. 発表年 2020年

1. 発表者名 小椋 透・柳本 武美
2. 発表標題 Zeta事前分布を用いた多項分布におけるパラメータ推定
3. 学会等名 2020年度統計関連学会連合大会
4. 発表年 2020年

1. 発表者名 柳本 武美
2. 発表標題 ベイズ型対数尤度に基づくモデルの信用集合
3. 学会等名 科研費シンポジウム「大規模複雑データの理論と方法論」
4. 発表年 2020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