

令和 6 年 6 月 17 日現在

機関番号：13904

研究種目：基盤研究(C) (一般)

研究期間：2019～2023

課題番号：19K11825

研究課題名(和文) 一般化ベイズ学習法の情報論的解釈と設計

研究課題名(英文) Information Theoretic Interpretation and Design of Generalized Bayesian Learning Methods

研究代表者

渡辺 一帆 (Watanabe, Kazuho)

豊橋技術科学大学・工学(系)研究科(研究院)・准教授

研究者番号：10506744

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：機械学習法の性質を情報理論における圧縮限界から特徴づけ、学習法の設計論を構築するための部分課題として、逐次時系列予測、レート歪み関数を達成する最適再構成分布、損失関数における不感応性の推定、疎性に基づく推定における正則化パラメータの推定を研究し、それぞれ、二値データの変動生成確率の予測に対する近似最適性予測法の開発、板倉・斎藤歪み尺度に対してレート歪み関数を達成する再構成分布の解明、不感応パラメータの推定誤差の解析、効率的な近似推定法の構築などの成果が得られた。

研究成果の学術的意義や社会的意義

歪み有りデータ圧縮の限界であるレート歪み関数を達成する再構成分布を特徴づける例を追加しており、既存の結果との対比に用いることができる。再構成分布の最適化に対応する、ベイズの定理に基づく学習法は一般に計算困難性を伴う。具体的な時系列解析問題において効率的な近似法を構築し、その性質が実験的に、または一部理論的に明らかにされた。

研究成果の概要(英文)：To establish a fundamental theory for the information theoretically principled design of statistical learning methods, we studied online time-series prediction, the optimal reconstruction distribution achieving the rate-distortion function, the estimation of the insensitivity in loss functions and the estimation of the regularization parameter in sparse estimation methods. For respective subproblems, we extended a minimax optimal prediction method for real valued data to the prediction of the time-varying probabilities from binary inputs, characterized the optimal reconstruction distribution achieving the rate-distortion function for Itakura-Saito distortion measure, analyzed the generalization errors in learning of the insensitive parameter, and developed an efficient approximate estimation method for the regularization parameter of L1 trend filtering, which extracts piece-wise linear trend from time-series data.

研究分野：統計的学習理論

キーワード：レート歪み関数 板倉・斎藤距離 オンライン予測 局所変分近似 不感応パラメータ 正則化パラメータ L1トレンドフィルタ

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、研究者個人に帰属します。

1. 研究開始当初の背景

統計的機械学習法は、データに内在する構造を、潜在変数を用いて抽出する学習モデル(潜在変数モデル)の推測と捉えることができる。その中でもベイズ推測に基づく学習法であるベイズ学習法は、学習モデルの自由度をデータから推定するための指標が自然に導出されるといった利点を持つだけでなく、潜在変数モデルに対し有効であることを示す理論的な成果も得られてきている。ベイズ学習による機械学習法の性能は、特にその事前分布の設計や事後分布の実現に大きく依存するにも関わらず、潜在変数モデルにおいては、統計学や情報理論分野における既存の結果からその性能を解析することができない。このため事前分布の設計は、応用毎に解析者の多くの試行錯誤が必要とされている。また、実用的な潜在変数モデルの事後分布は非共役性などの原因により一般に計算困難となり、その実現自体に工夫が必要とされる。

2. 研究の目的

ベイズ学習による機械学習法の性能を情報理論におけるデータの圧縮限界から特徴づけ、理論的限界を実現もしくは近似する方法論として、その設計原理を導くことを目的とする。

3. 研究の方法

歪み有りデータ圧縮の限界であるレート歪み関数を達成する再構成分布の性質を解析する。再構成分布の最適化に対応するベイズ学習法の近似について、その性質を調査する。具体的な時系列解析問題において効率的な近似法を構築し、その性質を実験的かつ理論的に解析する。また、不感応性を持つ損失関数に基づく学習法について、その汎化性能を解析する。

4. 研究成果

ベイズ学習による系列の予測は一定の最適性を持つが、観測データ系列の出方によっては、全てのデータを観測した際に最適なオフライン予測との乖離が大きくなる可能性がある。実数値データのトレンド成分の予測問題において開発された、この乖離に関する最適性を持つ予測法を、二値データからの生成確率の変動を予測する方法に拡張し、ベイズ学習で用いられる局所変分近似法が有効であることを示し、近似の下で最適性を持つ予測法を開発した(図1)。オフライン予測との乖離の系列の出方に対する最大値の系列長に対する振る舞いを解析し、実数値データからの予測問題と同様な振る舞いとなることを示した。

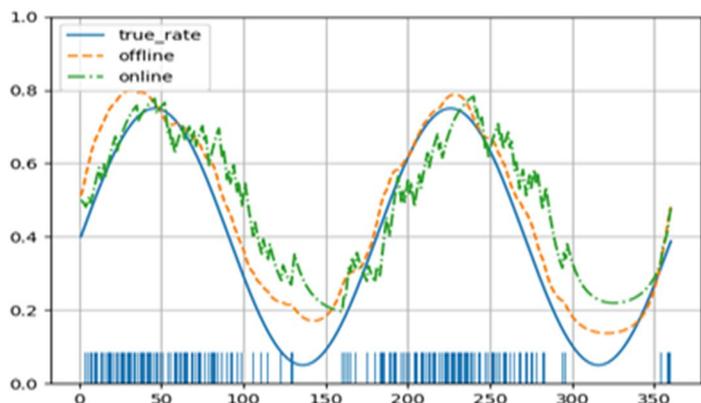


図1：変動二値情報源のオンライン予測 (実線：真の変動確率、鎖線：オフライン予測、一点鎖線：オンライン予測、下部：二値系列データ)

音声信号処理や通信網理論に用いられる板倉・斎藤損失について、レート歪み関数を達成する最適再構成分布を解析した。情報源がガンマ分布の場合以外一般に、情報理論の創始者であるシャノンによるレート歪み関数の下界より、レート歪み関数が真に大きくなり、それを達成する再構成分布は離散分布となることを示した。特に情報源が一様分布の場合に離散分布を数値的に最適化することにより、レート歪み関数を求め(図2) 既知の結果である二乗損失の場合では点が分割され新たな点が生成されるのと対照的に、新しい点が離れた位置に発生するように離散点の数が増加し、再構成分布が変化していくことが明らかにされた。

データ解析において幅広い応用を持つサポートベクトル回帰は 不感応損失関数により学習データに関する疎性や学習データに含まれる外れ値への頑健性を実現している。不感応パラメータ のデータからの推定法を構築し、その統計的推定精度を評価する汎化誤差を解析した(図3) その結果、構築した推定法は が正の範囲では、統計的正則モデルでの推定と同様に扱うことができるが、 が0においてだけ正則モデルとは異なる性質を示すことが明らかにされた。人工データを用いて赤池情報量規準やベイズ情報量規準の適用可能性を検証した。非線形な特

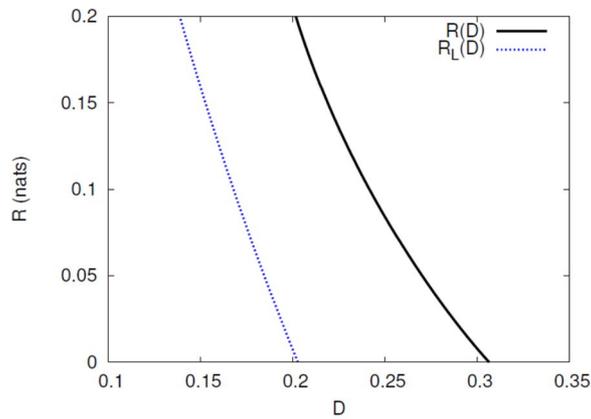


図2：板倉・斎藤距離に対するレート歪み関数とそのシャノン下界（情報源：一様分布）。

徴変換を効率的に計算するカーネル法と組み合わせた場合にも、従来のモデル選択規準の適用可能性と改変による精度向上が確認された。

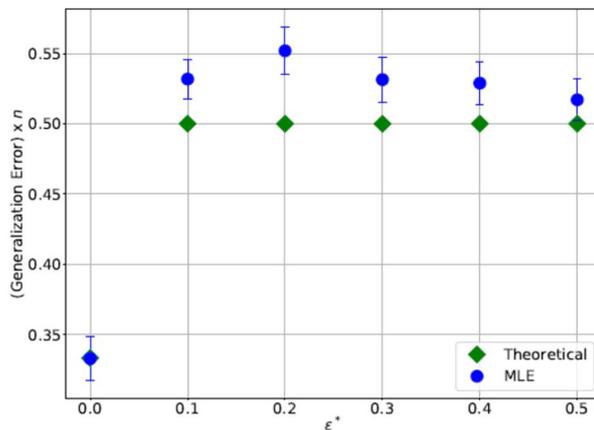


図3：不感応パラメータの学習における汎化誤差。

の真値（横軸）が0の場合のみ他の場合に比べて異なる振る舞いを示す。

疎性に基づく推定法（スパース推定法）においては、疎性の度合いを決定する正則化パラメータを適切に設定する必要がある。経験ベイズ法による正則化パラメータの推定は、ベイズ法における事前分布の最適化問題と捉えることができるが、スパース推定法においては共役性が成り立たず計算が困難となる。時系列データから区分線形なトレンド成分を抽出するスパース推定法であるL1トレンドフィルタリングについて、局所変分近似による近似法を導出し、正則化パラメータの推定法を構築した（図4）。トレンド推定精度、変化点検出精度に関して、系列長から決まる一定値に基づく従来の推定法と数値実験による比較を行い、導出された推定法の有効性が示された。また、時系列の滑らかさの仮定を一般化し、カルマンフィルタに基づく効率的な推定法を正則化パラメータとの同時推定の場合に拡張できることを示した。

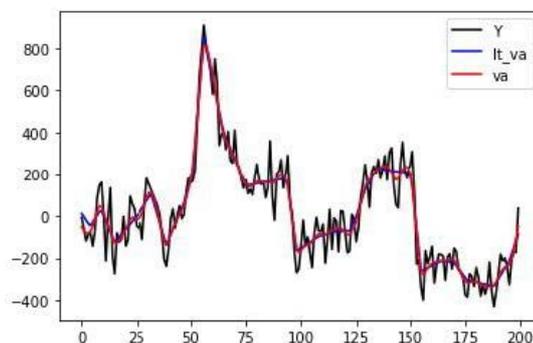


図4：L1トレンドフィルタリングの脳波データへの適用例。

変化点数も含め区分線形なトレンド成分がデータから抽出される。

5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 9件/うち国際共著 0件/うちオープンアクセス 3件）

1. 著者名 Masahiro Kobayashi, Kazuho Watanabe	4. 巻 -
2. 論文標題 Unbiased estimating equation on inverse divergence and its conditions	5. 発行年 2024年
3. 雑誌名 Proc. of 2024 IEEE International Symposium on Information Theory (ISIT)	6. 最初と最後の頁 1~1
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 Masahiro Kobayashi, Kazuho Watanabe	4. 巻 -
2. 論文標題 Unbiased Estimating Equation and Latent Bias under f-Separable Bregman Distortion Measures	5. 発行年 2024年
3. 雑誌名 IEEE Transactions on Information Theory	6. 最初と最後の頁 1~1
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TIT.2024.3366538	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 Akiharu Omae, Kazuho Watanabe	4. 巻 1
2. 論文標題 Approximate Empirical Bayes Estimation of the Regularization Parameter in l1 Trend Filtering	5. 発行年 2022年
3. 雑誌名 Proc. of 2022 IEEE International Symposium on Information Theory (ISIT)	6. 最初と最後の頁 462-467
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ISIT50566.2022.9834623	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 Masahiro Kobayashi, Kazuho Watanabe	4. 巻 458
2. 論文標題 Generalized Dirichlet-process-means for f-separable distortion measures	5. 発行年 2021年
3. 雑誌名 Neurocomputing	6. 最初と最後の頁 667~689
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.neucom.2020.03.123	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 Masahiro Kobayashi, Kazuho Watanabe	4. 巻 -
2. 論文標題 Unbiased Estimation Equation under f-Separable Bregman Distortion Measures	5. 発行年 2021年
3. 雑誌名 Proc. of 2020 IEEE Information Theory Workshop	6. 最初と最後の頁 311 ~ 315
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ITW46852.2021.9457678	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 Kazuho Watanabe	4. 巻 -
2. 論文標題 Statistical Learning of the Insensitive Parameter in Support Vector Models	5. 発行年 2021年
3. 雑誌名 Proc. of IEEE International Symposium on Information Theory	6. 最初と最後の頁 2501 ~ 2506
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ISIT45174.2021.9518182	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 Kazuho Watanabe	4. 巻 -
2. 論文標題 Discrete Optimal Reconstruction Distributions for Itakura-Saito Distortion Measure	5. 発行年 2020年
3. 雑誌名 Proc. of IEEE International Symposium on Information Theory	6. 最初と最後の頁 2399 ~ 2404
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ISIT44484.2020.9174352	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 Daisuke Kaji, Kazuho Watanabe, Masahiro Kobayashi	4. 巻 -
2. 論文標題 Multi-Decoder RNN Autoencoder Based on Variational Bayes Method	5. 発行年 2020年
3. 雑誌名 Proc. of 2020 International Joint Conference on Neural Networks (IJCNN)	6. 最初と最後の頁 1 ~ 1
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/IJCNN48605.2020.9206686	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 Kenta Konagayoshi, Kazuho Watanabe	4. 巻 101
2. 論文標題 Minimax Online Prediction of Varying Bernoulli Process under Variational Approximation	5. 発行年 2019年
3. 雑誌名 Proc. of the 11th Asian Conference on Machine Learning (ACML 2019)	6. 最初と最後の頁 141 - 156
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 -

〔学会発表〕 計17件 (うち招待講演 2件 / うち国際学会 9件)

1. 発表者名 Kazuho Watanabe
2. 発表標題 Rate-distortion theoretical views of Bayesian learning coefficients
3. 学会等名 IMSI Workshop on Bayesian Statistics and Statistical Learning (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 望月翔太, 渡辺一帆
2. 発表標題 不感応パラメータ推定サポートベクトル回帰における最適な次数選択
3. 学会等名 電気・電子・情報関係学会東海支部連合大会
4. 発表年 2023年

1. 発表者名 小林真佐大, 渡辺一帆
2. 発表標題 逆ガウスモデルにおける推定方程式の不偏性とその一般化
3. 学会等名 電気・電子・情報関係学会東海支部連合大会
4. 発表年 2023年

1. 発表者名 小林真佐大, 渡辺一帆
2. 発表標題 逆数ダイバージェンスにおける推定方程式の不偏性とその多次元拡張
3. 学会等名 第46回情報理論とその応用シンポジウム
4. 発表年 2023年

1. 発表者名 Akiharu Omae, Kazuho Watanabe
2. 発表標題 Approximate Empirical Bayes Estimation of the Regularization Parameter in l_1 Trend Filtering
3. 学会等名 IEEE International Symposium on Information Theory (ISIT2022) (国際学会)
4. 発表年 2022年

1. 発表者名 大前昭晴, 渡辺一帆
2. 発表標題 変分近似による L_1 トレンドフィルタリングの超パラメータ推定
3. 学会等名 第44回情報理論とその応用シンポジウム
4. 発表年 2021年

1. 発表者名 Masahiro Kobayashi, Kazuho Watanabe
2. 発表標題 Unbiased Estimation Equation under f -Separable Bregman Distortion Measures
3. 学会等名 IEEE Information Theory Workshop (ITW2020) (国際学会)
4. 発表年 2021年

1. 発表者名 Kazuho Watanabe
2. 発表標題 Statistical Learning of the Insensitive Parameter in Support Vector Models
3. 学会等名 IEEE International Symposium on Information Theory (ISIT2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Kazuho Watanabe
2. 発表標題 Discrete Optimal Reconstruction Distributions for Itakura-Saito Distortion Measure
3. 学会等名 IEEE International Symposium on Information Theory (ISIT2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Daisuke Kaji, Kazuho Watanabe, Masahiro Kobayashi
2. 発表標題 Multi-Decoder RNN Autoencoder Based on Variational Bayes Method
3. 学会等名 IEEE International Joint Conference on Neural Networks (IJCNN2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Kazuho Watanabe
2. 発表標題 Rate-distortion theoretic interpretation of Bayesian learning coefficients
3. 学会等名 Information Theory and Applications Workshop (招待講演) (国際学会)
4. 発表年 2020年

1. 発表者名 Kenta Konagayoshi, Kazuho Watanabe
2. 発表標題 Minimax Online Prediction of Varying Bernoulli Process under Variational Approximation
3. 学会等名 Asian Conference on Machine Learning (ACML2019) (国際学会)
4. 発表年 2019年

1. 発表者名 小林真佐大, 渡辺一帆
2. 発表標題 f分離可能Bregman歪み尺度における推定方程式の不偏性と推定量の性質
3. 学会等名 電子情報通信学会情報理論研究会
4. 発表年 2019年

1. 発表者名 Masahiro Kobayashi, Kazuho Watanabe
2. 発表標題 Unbiased Estimation Equation Under f-separable Extension of Squared and Itakura-Saito Distances
3. 学会等名 Data Science, Statistics and Visualization (DSSV) (国際学会)
4. 発表年 2019年

1. 発表者名 沖本卓也, 渡辺一帆
2. 発表標題 不感応損失関数におけるパラメータの推定量とその性質
3. 学会等名 電気・電子・情報関係学会東海支部連合大会
4. 発表年 2019年

1. 発表者名 渡辺一帆
2. 発表標題 Discreteness of Optimal Reconstruction Distributions in Lossy Compression under Itakura-Saito Distortion
3. 学会等名 第42回情報理論とその応用シンポジウム
4. 発表年 2019年

1. 発表者名 小林真佐大, 渡辺一帆
2. 発表標題 f分離可能ブレグマン歪み尺度に基づくロバストな非負値行列分解
3. 学会等名 第42回情報理論とその応用シンポジウム
4. 発表年 2019年

〔図書〕 計1件

1. 著者名 Shinichi Nakajima, Kazuho Watanabe, Masashi Sugiyama	4. 発行年 2019年
2. 出版社 Cambridge University Press	5. 総ページ数 558
3. 書名 Variational Bayesian learning theory	

〔産業財産権〕

〔その他〕

豊橋技術科学大学学習推論システム研究室ホームページ http://www.lisl.cs.tut.ac.jp/
--

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
---------	---------