

令和元年6月21日現在

機関番号：13904

研究種目：若手研究(B)

研究期間：2015～2018

課題番号：15K16050

研究課題名（和文）最適潜在分布を用いた学習推論モデルの設計原理

研究課題名（英文）Design Principles of Learning and Inference Models with Optimal Latent Distributions

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）・講師

研究者番号：10506744

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）：機械学習においてその有用性が知られているイプシロン不感応損失やカーネル関数に基づく歪み尺度に対して、歪み有りデータ圧縮の限界を示すレート歪み関数の評価を与えた。また、データの潜在構造を抽出するためのクラスタリング手法やベイズ推論における事前分布の設計に関し、潜在変数分布の最適化としての解釈を与え、手法の拡張や近似法の開発および解析を行った。データ可視化に基づくデータ解析法のために潜在変数を持つ学習モデルを改良し、実データ解析での性能を調査した。

研究成果の学術的意義や社会的意義

レート歪み関数は歪み有りデータ圧縮の限界を示すため、イプシロン不感応損失やカーネル関数を用いた損失などの有用性が知られている歪み尺度において、圧縮法の性能評価や改良を与える際の基準が得られた。また、クラスタリング手法やベイズ推論における事前分布の設計において得られた拡張や近似法により、既存手法の個々の問題におけるより柔軟な適用や効率的な計算が可能になった。

研究成果の概要（英文）：We evaluated the rate-distortion functions of practical loss functions such as epsilon-insensitive distortion measures and those defined with kernel functions, which demonstrate the performance of optimal lossy data compression systems under these distortion measures. We interpreted a recent clustering method to extract latent structure of data and the design of the prior distribution in Bayesian inference through the optimization of latent distributions, and provided their extensions and approximations. We developed a modified latent variable model for an interactive data analysis method based on data visualization, and examined its performance through its applications to some real datasets.

研究分野：統計的学習理論

キーワード：レート歪み関数 再構成分布 イプシロン不感応損失 ディリクレ過程平均法 ミニマックス予測

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

### 1. 研究開始当初の背景

パターン認識やデータマイニングにおいて広く用いられている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。機械学習法の性能は、主に損失関数の選択に代表される、その設計に大きく依存するにも関わらず、統計学や情報理論分野における既存の結果から解析することができず、学習モデルの設計は、応用毎に解析者の多くの試行錯誤が必要とされていた。

### 2. 研究の目的

本研究では、情報理論の一分野であり歪み有りデータ圧縮の限界を明かにするレート歪み理論の発展と、それを通じて学習モデルの推論・予測法の性能や限界を明かにし、学習モデルの設計論を構築することを目的とした。

### 3. 研究の方法

以下の三つの課題を設定した。

- (1) 実データの生成過程に近い情報源のモデルや実用的な学習モデルに対応する複雑な歪み尺度に対する、レート歪み関数の評価。
- (2) レート歪み理論の知見による学習・予測法の理論的限界の解明と設計論の構築。
- (3) 学習モデルの実際の応用に対する学習法の導出・改良および性能解析。

### 4. 研究成果

上記の三課題について、以下のような成果が得られた。

#### (1) レート歪み関数の評価

音声信号処理や通信網理論に用いられるガンマ情報源、および対数変換を伴う歪み尺度について、レート歪み関数の導出を行い、最適潜在変数分布の変化を詳細に解明した[論文 ]。また、外れ値に対する頑健性を持たせるための工夫としてサポートベクトルマシンなどで用いられている

不感応損失や、効率的に非線形処理を実現するカーネル法における損失関数について、レート歪み関数の評価を与えた(図1)[論文 , , ]。さらに、二乗損失について効率的な計算量でレート歪み関数を達成することが証明されているスパース線形回帰符号を、最適潜在変数分布の知見に基づき、絶対損失の場合に拡張した[学会発表 ]。

#### (2) 学習モデルの設計論の構築

大規模データ解析向けに開発されたディリクレ過程平均法と呼ばれるクラスタリング手法は、歪み尺度として最大歪み基準によりクラスタ数を推定していることを示し、この知見に基づき、多様な損失関数に対し、計算量のオーダーを保持した拡張を与えた[論文 , 学会発表 ]。

逐次的にデータ系列を予測するオンライン学習において、データセットの出方に対する最悪符号長により測られる予測誤差を最小化する最適な予測法は、一般に膨大な計算量が必要とされるため、潜在変数モデルを用いた効率的な予測法が考えられてきた。予測法が十分な精度を保証するための潜在変数分布の条件を特徴付け、効率的かつ精度の良い近似法を構成した(図2)[論文 ]。

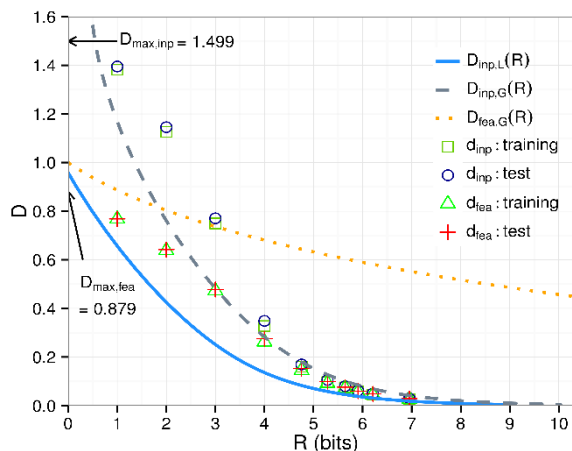


図1: レート歪み関数と圧縮法の比較  
(カーネル法の損失関数の場合[論文 ])

実用的な潜在変数モデルについて、その有効性が知られているベイズ推論に対し、レート歪み理論による解釈を与え、潜在変数を持つ学習モデルの性能を示す学習係数と、レート歪み関数から求められるレート歪み次元との間の関係を明らかにした[学会発表]。また、疎性に基づく学習・推定に用いられる潜在変数分布に対し、その決定法である経験ベイズ法の解析解を導出し、近似法の精度を定量的に明らかにした[論文]。

(3) 実データ解析における学習法の応用および性能解析

潜在変数モデルによるデータ解析と多変量データの可視化と組み合わせたインタラクティブな解析環境を提案し(図3)[論文, 学会発表]、天文データ解析への応用において、専門家の知見を半自動的なデータ解析により再現できることを示した[論文]。さらに、近赤外分光断層撮影、関係データ解析などに応用される実用的な潜在変数モデルについて、変分ベイズ法による学習法の導出や近似精度および潜在変数分布への依存性の解析を行った[図書, 論文]。

5. 主な発表論文等

[雑誌論文](計16件)[全て査読有]

T. Yoshida, K. Watanabe, "Empirical Bayes estimation for L<sub>1</sub> regularization: a detailed analysis in the one-parameter lasso model," *IEICE Transactions on Fundamentals*, vol.E101-A, no.12, pp.2184-2192, 2018.

DOI: 10.1587/transfun.E101.A.2184

I. Fujishiro, N. Sawada, M. Nakayama, H-Y Wu, K. Watanabe, S. Takahashi, M. Uemura, "TimeTubes: visual exploration of observed blazar datasets," *Journal of Physics: Conference Series*, vol.1036, 012011, 2018.

DOI: 10.1088/1742-6596/1036/1/012011

小林真佐大, 渡辺一帆, "ディリクレ過程平均法のレートひずみ理論による解釈," *電子情報通信学会論文誌A*, vol.J100-A, no.12, pp.475-486, 2017.

<http://id.nii.ac.jp/1117/00002012/>

K. Watanabe, "Rate-distortion bounds for kernel-based distortion measures," *Entropy*, vol.19, no.7, article-id: 336, 2017.

DOI: 10.3390/e19070336

K. Watanabe, "Projection to mixture families and rate-distortion bounds with power distortion measures," *Entropy*, vol.19, no.6, article-id: 262, 2017.

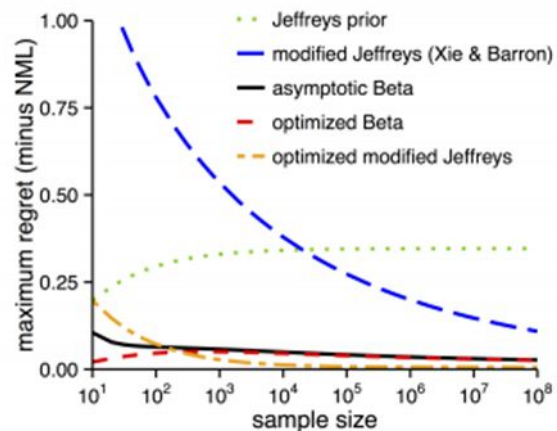


図2: データ数に対する各手法の予測誤差 (上部2つが既存手法, 下部3つが提案法)

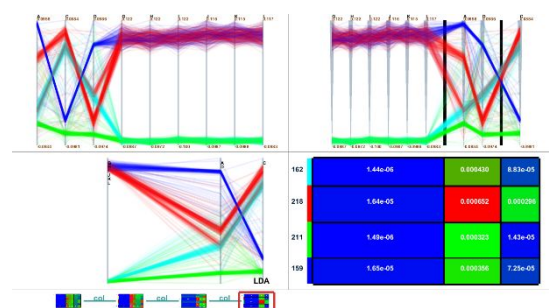


図3: 可視化を用いたデータ解析環境

DOI: 10.3390/e19060262

K. Watanabe, S. Ikeda, "Rate-distortion functions for gamma-type sources under absolute-log distortion measure," *IEEE Transactions on Information Theory*, vol.62, no.10, pp.5496-5502, 2016.

DOI: 10.1109/TIT.2016.2602100

M. Uemura, R. Itoh, L. Xu, M. Nakayama, H-Y. Wu, K. Watanabe, S. Takahashi, I. Fujishiro, "TimeTubes: Visualization of polarization variations in blazars," *Galaxies*, vol. 4, no. 3, article-id: 23, 2016.

DOI: 10.3390/galaxies4030023

藤代一成, 高橋成雄, 渡辺一帆, Hsiang-Yun Wu, "スパースモデリングと情報可視化," 電子情報通信学会誌, vol.99, No.5, pp.466-470, 2016 .

[http://www.journal.ieice.org/summary.php?id=k99\\_5\\_466&year=2016&lang=J](http://www.journal.ieice.org/summary.php?id=k99_5_466&year=2016&lang=J)

K. Watanabe, H-Y. Wu, S. Takahashi, I. Fujishiro, "Asymmetric biclustering with constrained von Mises-Fisher models," *Journal of Physics: Conference Series*, vol.699, 012018, 2016.

DOI: 10.1088/1742-6596/699/1/012018

M. Uemura, K. S. Kawabata, S. Ikeda, K. Maeda, H-Y. Wu, K. Watanabe, S. Takahashi, I. Fujishiro, "Data-driven approach to Type Ia supernovae: variable selection on the peak luminosity and clustering in visual analytics," *Journal of Physics: Conference Series*, vol.699, 012009, 2016.

DOI: 10.1088/1742-6596/699/1/012009

K. Watanabe, "Rate-distortion bounds for epsilon-insensitive distortion measures," *IEICE Transactions on Fundamentals*, vol.E99-A, no.1, pp.370-377, 2016.

DOI: 10.1587/transfun.E99.A.370

K. Watanabe, T. Roos, "Achievability of asymptotic minimax regret by horizon-dependent and horizon-independent strategies," *Journal of Machine Learning Research*, vol.16, pp.2357-2375, 2015.

<http://jmlr.org/papers/volume16/watanabe15a/watanabe15a.pdf>

K. Watanabe, "Vector quantization based on epsilon-insensitive mixture models," *Neurocomputing*, vol.165, pp.32-37, 2015.

DOI: 10.1016/j.neucom.2015.01.081

T. Konishi, T. Kubo, K. Watanabe, K. Ikeda, "Variational Bayesian inference algorithms for infinite relational model of network data," *IEEE Transactions on Neural Networks and Learning Systems*, vol.26, no.9, pp.2176-2181, 2015.

DOI: 10.1109/TNNLS.2014.2362012

能野琴, 吳湘筠, 渡辺一帆, 高橋成雄, 藤代一成, "グラフスペクトル解析を用いた平行座標系の軸縮約," 画像電子学会誌, vol.44, no.3, pp.447-456, 2015 .

<http://www.iieej.org/gakkaishi.html>

A. Miyamoto, K. Watanabe, K. Ikeda, M. Sato, "Variational inference with ARD prior for NIRS diffuse optical tomography," *IEEE Transactions on Neural Networks and Learning Systems*, vol.26, no.5, pp.1109-1114, 2015.

DOI: 10.1109/TNNLS.2014.2328576

〔学会発表〕(計 10 件)

M. Kobayashi, K. Watanabe, "Generalized Dirichlet-process-means for robust and maximum distortion criteria," International Symposium on Information Theory and its Applications, 2018.

T. Yoshida, T. Moriya, K. Watanabe, Y. Shinohara, Y. Yamaguchi, Y. Aono, "Automatic DNN node pruning using mixture distribution-based group regularization," Interspeech, 2018.

R. Konabe, K. Watanabe, "Sparse regression code with sparse dictionary for absolute error criterion," IEEE International Symposium on Information Theory, 2018.

K. Watanabe, "Rate-distortion dimension and Bayesian learning coefficient," Workshop on Information Theoretic Methods in Science and Engineering, 2017.

K. Watanabe, "Rate-distortion tradeoffs under kernel-based distortion measures," IEEE International Symposium on Information Theory, 2017.

H-Y. Wu, Y. Niibe, K. Watanabe, S. Takahashi, M. Uemura, I. Fujishiro, "Making many-to-many parallel coordinate plots scalable by asymmetric biclustering," IEEE Pacific Visualization Symposium, 2017.

K. Watanabe, "Rate-distortion theoretic views of learning problems," Workshop on Information Theoretic Methods in Science and Engineering, 2016.

K. Watanabe, "Constant-width rate-distortion bounds for power distortion measures," IEEE Information Theory Workshop, 2016.

L. Xu, M. Nakayama, H-Y. Wu, K. Watanabe, S. Takahashi, M. Uemura, I. Fujishiro, "TimeTubes: Design of a visualization tool for time-dependent, multivariate blazar datasets," NICOGRAPH International, 2016.

K. Watanabe, "Rate-distortion analysis for kernel-based distortion measures," Workshop on Information Theoretic Methods in Science and Engineering, 2015.

〔図書〕(計 1 件)

S. Nakajima, K. Watanabe, M. Sugiyama, *Variational Bayesian Learning Theory*, Cambridge University Press, 560 pages, 2019.

〔その他〕

豊橋技術科学大学情報・知能工学系学習推論システム研究室ホームページ  
<http://www.lisl.cs.tut.ac.jp/>

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。