

## 科学研究費助成事業 研究成果報告書

平成 27 年 6 月 18 日現在

機関番号：13904

研究種目：若手研究(B)

研究期間：2011～2014

課題番号：23700175

研究課題名(和文)カーネル法を用いた学習アルゴリズムの情報論的解釈と発展

研究課題名(英文) Information-Theoretic View and Development of Kernel-Based 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)・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)：混合モデルはクラスタリングや確率密度推定に用いられる学習モデルである。既存の凸クラスタリング法やカーネルベクトル量子化を含む統一的な学習法を導出し、歪み有りデータ圧縮の限界を示すレート歪み関数を求める最適化問題と捉えられることを示した。これまで評価が与えられてこなかった複雑な歪み尺度に対するレート歪み関数に対する評価を与えた。また、最小の符号長からの損失分で測られる逐次予測の精度について、混合モデルにより最適解を効率良く近似する手法を与えた。

研究成果の概要(英文)：We developed a learning method of mixture models, which unifies existing methods such as convex clustering and kernel vector quantization. We provided an interpretation of this method as an optimization problem for the evaluation of the rate-distortion function. The rate-distortion function indicates the performance of optimal lossy data compression systems. We evaluated rate-distortion functions of practical loss functions such as absolute-log distortion and epsilon-insensitive distortion measures. We also evaluated the performance of an online prediction algorithm to provide an efficient method to approximate the optimal prediction algorithm.

研究分野：統計的学習理論、機械学習

キーワード：混合モデル レート歪み関数 再構成分布 イプシロン不感応損失 漸近的ミニマックス性

### 1. 研究開始当初の背景

パターン認識、データマイニングの諸問題は、多量のデータの背後にあるデータの生成過程を学習する問題であり、その解決には計算統計・機械学習において開発された種々の学習アルゴリズムが用いられている。カーネル法は簡単な計算により複雑な非線形データ処理を可能にする手法として様々な学習アルゴリズムと組み合わせて用いられる。しかしながら、カーネル法を用いた学習アルゴリズムの性質を調べる理論研究はあまり発展しておらず、学習モデルの設計やアルゴリズムの設計には、理論的な背景や性能保証が与えられることは少なかった。

### 2. 研究の目的

本研究では、既存の学習アルゴリズムの背後にある原理を説明するとともに、学習アルゴリズムの汎化性能の解析につなげるため、カーネル法を用いた学習アルゴリズムを確率推論として定式化し、ベイズ推定や情報理論における符号長最小化としての理解を通じて、学習アルゴリズムの性能の評価及び設計の指針を与えることを目的とする。

### 3. 研究の方法

既に確率推論として定式化されている学習アルゴリズムを中心に典型的な汎化性能の解析を行う。また他の学習アルゴリズムについて確率推論としての定式化を行い汎化性能を調べる。情報論的な特徴づけを用いて、確率推論としての解釈が困難なモデルに対する解釈を与えると同時に、確率推論として与えられるモデルに対する拡張を与える。典型的な汎化性能評価及び情報論的な特徴により、学習モデルの選択法やカーネルパラメータの設計法を構築する。

### 4. 研究成果

(1) 混合モデルはクラスタリングや確率密度推定に用いられる学習モデルである。有限混合分布の学習においては混合数を適切に決めることが必要となるが、混合数自体もデータから推定するアプローチの一つとして、混合分布を構成する重み分布に連続分布を含む一般の分布を仮定したノンパラメトリック推定法がある。最尤推定を用いたこの枠組みでは、最適な重み分布が離散分布で与えられること、およびそのサポートの数が学習データ数以下になることが示されている。しかしながら、最尤推定の性質上、与えられたデータに過学習することが起こりやすい。

この研究課題では、最尤推定における対数尤度関数を一般化した目的関数を構成し、その最小化による推定法を提案した。この目的関数は一つのパラメータを持っており、パラメータが無限大の極限ではカーネルベクトル量子化法の目的関数に一致する。この統一

的な目的関数について、離散分布の最適性を示し、それを求める学習アルゴリズムを構成した。また、レニーダイバージェンスやガンマダイバージェンスなどの確率分布間のダイバージェンスの最小化との関係を考察した。最尤推定からの偏りと推定の分散を考察し、平均汎化誤差の意味で最適なパラメータの学習データ数に関するオーダーを評価した。人工データを用いた数値実験により、パラメータを調整することで最尤推定の場合 ( $\beta=0$ ) に比べて汎化誤差を改善できることが示された(図1)。さらに、この学習法がレート歪み関数を求める最適化問題と捉えられることを示した。

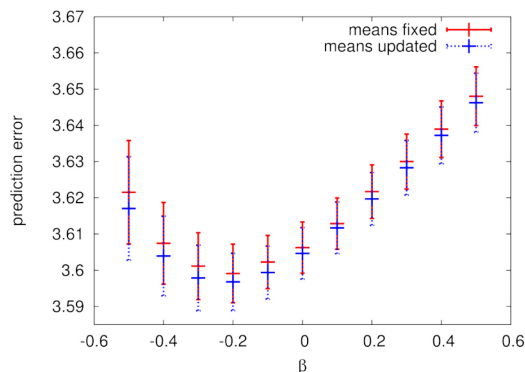


図1：混合分布の汎化誤差

(2) レート歪み関数は歪み有りデータ圧縮において、許容する歪みの量と必要な情報量とのトレードオフを示す関数である。ガンマ情報源と対数絶対損失による歪み尺度に対するレート歪み関数の評価を行った。一般にレート歪み関数の評価にはシャノンの導出した下界が重要な役割を果たすが、下界と等しくならない場合の厳密な評価はごく限られた例に対してのみ与えられていた。この研究課題では、対数絶対損失に対するレート歪み関数を求める問題が、その中の一例に帰着できることを示し、ガンマ情報源に対してレート歪み関数の厳密な評価を与えた。ガンマ分布は神経科学におけるスパイク間隔や音声信号処理におけるパワースペクトルなどの分布として広く用いられている。ガンマ情報源に対するレート歪み関数の導出過程から、対数絶対損失による歪み有りデータ圧縮については、最適な再構成分布が離散成分と連続成分からなることが示された。

(3) 外れ値に対する頑健性のために用いられるイプシロン不感応損失関数についてレート歪み関数の下界および上界評価を与えた。また、いくつかの特別な情報源の例において、上下界の近似精度を示した。イプシロン不感応損失の重み付け和を最小化する方法を構成し、対応する雑音モデルの混合モデルの学習法を導出し、レート歪み関数の近似計算法を導出した(図2)。

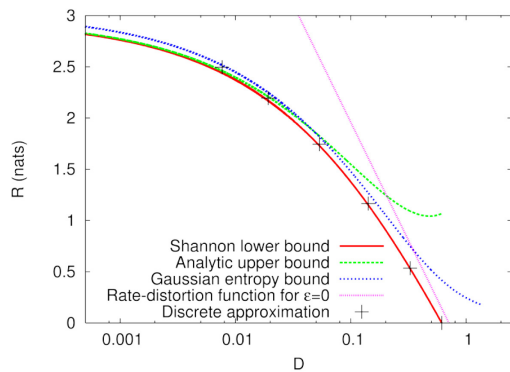


図 2 : レート歪み関数の近似計算

(4) 逐次的に与えられるデータに対し適応的に予測を変化させていく枠組みであるオンライン学習について、データセットの出方についての損失の最悪値によって評価される精度の最適値を効率よく近似する漸近的ミニマックス性の達成条件を考察した。ヘルシンキ大学計算機科学部に短期滞在し情報交換を行い、離散データの生成モデルである多項モデルにおいては、漸近的ミニマックス性を達成するためにはデータ数の知識が必要であることを示した。また、データ数に対する単純な依存性を持つ共役事前分布を用いたベイズ予測により、漸近的ミニマックス性が達成できることを示した。

(5) 分類問題で広く用いられるロジスティック回帰モデルなどのベイズ学習においては、凸関数の性質を用いた近似計算法が提案され、これは局所変分法として広く用いられている。局所変分法の目的関数がカルバック情報量とブレグマン情報量の期待値の和に分解できることを示し情報論的な見通しを与えた。潜在変数を持つ学習モデルに対するベイズ学習の近似法として用いられる変分ベイズ法について、局所変分法としての解釈を与え、その近似精度が、学習モデルのカルバック情報量への対数和不等式による上界によって記述できることを示した。また、変分ベイズ法における汎化誤差と目的関数の間の関係を導出し、数値実験による検証を行った。さらに、データの背後にある潜在的な要因を疎に推定する問題として、近赤外拡散光断層撮影において関連度自動決定を用いた場合や、Web などの文書解析や購買データの解析法として用いられる潜在ディリクレ配分法に関して、変分ベイズ法の目的関数の振る舞いを詳細に解析し、疎な推定結果が得られる事前分布の設定を明らかにした。

## 5. 主な発表論文等

[雑誌論文] (計 8 件)

① K. Watanabe, M. Okada, K. Ikeda, ``Divergence measures and a general framework for local variational approximation,`` *Neural Networks*, vol. 24, pp. 1102-1109, 査読有, 2011.  
DOI:10.1016/j.neunet.2011.06.004

② K. Watanabe, M. Okada, ``Approximate Bayesian estimation of varying binomial process,`` *IEICE Transactions on Fundamentals*, vol. E94A, pp. 2879-2885, 査読有, 2011.

DOI: 10.1587/transfun.E94.A.2879

③ K. Watanabe, ``An alternative view of variational Bayes and asymptotic approximations of free energy,`` *Machine Learning*, vol. 86, pp. 273-293, 査読有, 2012.  
DOI: 10.1007/s10994-011-5264-5

④ K. Watanabe, S. Ikeda, ``Entropic risk minimization for nonparametric estimation of mixing distributions,`` *Machine Learning*, vol. 99, no. 1, pp. 119-136, 査読有, 2015.

DOI:10.1007/s10994-014-5467-7

⑤ 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, 査読有, 2015.

DOI:10.1109/TNNLS.2014.2328576

⑥ K. Watanabe, ``Vector Quantization Based on Epsilon-Insensitive Mixture Models,`` *Neurocomputing*, vol. 165, pp. 32-37, 査読有, 2015.

DOI: 10.1016/j.neucom.2015.01.081

⑦ 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*, 印刷中, 査読有.

DOI:10.1109/TNNLS.2014.2362012

⑧ K. Watanabe, T. Roos, ``Achievability of Asymptotic Minimax Regret by Horizon-Dependent and Horizon-Independent Strategies,`` *Journal of Machine Learning Research*, 印刷中, 査読有.

[学会発表] (計 22 件)

① 宮本敦史, 渡辺一帆, 池田和司, 佐藤雅昭, ``NIRS-DOT のための変分ベイズによる逆問題解法,`` 第 55 回システム制御情報学会研究発表講演会, 大阪大学 (大阪府吹田市), 平成 23 年 5 月 19 日.

② 渡辺一帆, 岡田真人, 池田和司, ``ブレグマン情報量を用いたベイズ学習の局所変分近似,`` 電子情報通信学会ニューロコンピューティング研究会, 平成 23 年 7 月 25 日, 神戸大学 (兵庫県神戸市).

③ A. Miyamoto, K. Watanabe, K. Ikeda, M. Sato, ``Phase diagrams of a variational Bayesian approach with ARD prior in NIRS-DOT,`` *International Joint Conference on Neural Networks*, Aug 1st, 2011, San Jose (USA).

④ K. Watanabe, ``Minimum variational stochastic complexity and average generalization error in latent variable models,’’ Workshop on Information Theoretic Methods in Science and Engineering (招待講演), Aug 10th, 2011, Helsinki (Finland).

⑤ 宮本敦史, 渡辺一帆, 池田和司, ``付加的計測を用いたネットワークトモグラフィ,’’ 電子情報通信学会ソサイエティ大会, 平成 23 年 9 月 15 日, 北海道大学 (北海道札幌市).

⑥ 宮本敦史, 渡辺一帆, 池田和司, ``アクティブ計測とパッシブ計測を用いたネットワークトモグラフィ,’’ 電子情報通信学会情報論の学習理論と機械学習研究会, 平成 23 年 11 月 9 日, 奈良女子大学 (奈良県奈良市).

⑦ 渡辺一帆, 池田思朗, ``混合分布のノンパラメトリック推定と最小ダイバージェンス法,’’ 電子情報通信学会情報論の学習理論と機械学習研究会, 平成 23 年 11 月 10 日, 奈良女子大学 (奈良県奈良市).

⑧ 宮本敦史, 渡辺一帆, 池田和司, ``アクティブ計測とパッシブ計測を用いたパケットロス率推定法,’’ 電子情報通信学会情報論の学習理論と機械学習研究会, 平成 24 年 3 月 13 日, 統計数理研究所 (東京都立川市).

⑨ K. Watanabe, S. Ikeda, ``Convex formulation for nonparametric estimation of mixing distribution,’’ Workshop on Information Theoretic Methods in Science and Engineering (招待講演), Aug 30th, 2012, Amsterdam (Netherlands).

⑩ A. Miyamoto, K. Watanabe, K. Ikeda, ``Packet Loss Rate Estimation with Active and Passive Measurements,’’ Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Dec 4th, 2012, Hollywood (USA).

⑪ 山崎 啓介, 渡辺一帆, 梶 大介, ``自由エネルギーによる潜在変数推定精度の計算法,’’ 情報論の学習理論ワークショップ, 筑波大学 (東京都文京区), 平成 24 年 11 月 7 日.

⑫ 渡辺一帆, 池田思朗, ``Rate-distortion function for gamma sources under absolute-log distortion,’’ 情報理論とその応用シンポジウム, 大分県別府市, 平成 24 年 12 月 13 日.

⑬ 渡辺一帆, テームルース, ペトリミリマキ ``漸近的ミニマックス性とデータ数の知識の関係について,’’ 電子情報通信学会情報論の学習理論と機械学習研究会, 平成 25 年 3 月 5 日, 名古屋工業大学 (愛知県名古屋市).

⑭ K. Watanabe, S. Ikeda, ``Rate-Distortion Function for Gamma Sources under Absolute-Log Distortion Measure,’’ IEEE International Symposium on Information Theory, Jul 11th, 2013,

Istanbul (Turkey).

⑮ K. Watanabe, T. Roos, P. Myllymäki, ``Achievability of Asymptotic Minimax Optimality in Online and Batch Coding,’’ Workshop on Information Theoretic Methods in Science and Engineering (招待講演), Aug 28th, 2013, 東京大学 (東京都文京区).

⑯ K. Watanabe, ``Rate-Distortion Bounds for an Epsilon-Insensitive Distortion Measure,’’ IEEE Information Theory Workshop, Sep 13th, 2013, Sevilla (Spain).

⑰ K. Watanabe, ``Vector Quantization Using Mixture of Epsilon-Insensitive Components,’’ International Conference on Neural Information Processing, Nov 5th 2013, Daegu (Korea).

⑱ K. Watanabe, T. Roos, P. Myllymäki, ``Achievability of Asymptotic Minimax Regret in Online and Batch Prediction,’’ Asian Conference on Machine Learning, Nov 15th, 2013, Canberra (Australia).

⑲ A. R. Barron, T. Roos, K. Watanabe, ``Bayesian Properties of Normalized Maximum Likelihood and its Fast Computation,’’ IEEE International Symposium on Information Theory, Jul 2nd, 2014, Honolulu (USA).

⑳ 中島伸一, 佐藤一誠, 杉山将, 渡辺一帆, 小林寛子, ``Asymptotic Analysis of Variational Bayesian Latent Dirichlet Allocation,’’ 電子情報通信学会情報論の学習理論と機械学習研究会, 平成 26 年 11 月 18 日, 名古屋大学 (愛知県名古屋市).

㉑ S. Nakajima, I. Sato, M. Sugiyama, K. Watanabe, H. Kobayashi, ``Analysis of variational Bayesian latent Dirichlet allocation: weaker sparsity than MAP,’’ Neural Information Processing Systems, Dec 11th, Montreal (Canada).

㉒ 渡辺一帆, ``レート歪み関数と測度の最適化,’’ 第 37 回情報理論とその応用シンポジウムワークショップ, 平成 26 年 12 月 10 日, 富山県黒部市.

[図書] (計 1 件)

日本シミュレーション学会編「シミュレーション辞典」ビジュアルデータマイニングの基礎理論：機械学習担当

[その他]

ホームページ:

<http://www.lisl.cs.tut.ac.jp/wkazuho/index.html>

## 6. 研究組織

### (1) 研究代表者

渡辺 一帆 (WATANABE, Kazuho)

豊橋技術科学大学・大学院工学研究科・講師

研究者番号：10506744