

自己評価報告書

平成23年 5月11日現在

機関番号：13901

研究種目：若手研究(B)

研究期間：2008～2011

課題番号：20700251

研究課題名(和文)

機械学習と最適化理論の横断的研究

研究課題名(英文)

Transversal study of machine learning and optimization

研究代表者

金森 敬文 (TAKAFUMI KANAMORI)

名古屋大学・情報科学研究科・准教授

研究者番号：60334546

研究分野：総合領域

科研費の分科・細目：情報学・統計科学

キーワード：統計的学習理論

1. 研究計画の概要

本研究では、機械学習と最適化理論・数値計算法を横断する理論的な枠組を構築することを計画している。研究計画は主に3つのテーマからなる：(1)さまざまなリスク尺度やパラメトリック最適化の機械学習への応用、(2)不確実性のもとでの最適化、(3)機械学習と最適化のゲーム理論的融合。いずれの研究課題においても、理論だけではなく数値的に効率的な学習アルゴリズムの構築と実装を目指す。

2. 研究の進捗状況

研究計画(1)に対しては、数理ファイナンスにおける条件付きバリューアットリスク尺度を機械学習アルゴリズムに応用する研究をおこなった。とくにデータに含まれるノイズをうまく処理できるように、リスク尺度にロバスト性をもたせた定式化を行った。いくつかの実データに適用し、良好な結果を得た。研究計画(2)に対しては、不確実性を含む最適化問題をロバスト最適化として定式化し、統計的サンプリング手法を用いて近似解を計算する方法について提案した。そのとき重要となるのはサンプリング数の上限の評価である。その問題に対して機械学習の精度評価の方法を応用し、理論的な結果を得た。さらに、不確実性が存在する状況で得られた結果の信頼性評価について、ブートストラップ法などの統計的な方法を導入した。これにより、不確実性のもとでの問題解決のための計算アルゴリズムだけでなく、統計的な信頼性評価まで可能になった。また、確率密度の比(密度比)を推定する問題に対して、データにノイズが含まれる状況での計算効率の評価を、ランダム行列の理論を援用して解析し、不確実

性が存在する場合における、最も効率的な計算方法を導出した。研究計画(3)に対しては、「ゲーム理論的融合」を最適化における双対理論の応用と解釈して、いくつかの最適化アルゴリズムを理論的に解析した。機械学習で重要な概念であるブレグマン擬距離を用いて、最適化アルゴリズムの計算効率や数値安定性などを議論した。以上の成果はそれぞれ、学術論文や国際会議の予稿としてまとめられ公表されている。

3. 現在までの達成度

研究はおおむね順調に進展している。本研究課題を開始した当初は、あまり重要性を認識していなかった密度比推定の問題が、研究を進めるにつれて理論的にも応用上も非常に重要な統計的問題であることが明らかになってきた。密度比推定の問題を、不確実性のもとでの最適化問題として定式化することで、多くの成果が得られた。

4. 今後の研究の推進方策

密度比推定の問題を通して、研究計画(1)と研究計画(2)の関連が明らかになりつつあるので、今後はその方向の研究をすすめる方針である。また研究計画(3)については、最適化や情報幾何における双対理論を基礎として、さまざまな学習アルゴリズムの解析を行う方針である。ある最適化問題に対して幾何学的双対理論を応用することで、計算アルゴリズムの理論的な性質が明らかになることが分かってきた。今後はさらに、さまざまな計算アルゴリズムに対して、情報幾何的な視点から、統一的な理論を構築することを目指す。

5. 代表的な研究成果

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 11 件)

① Takafumi Kanamori,
Deformation of Log-Likelihood Loss Function for Multiclass Boosting.

Neural Networks, 査読有 vol. 23, pp. 843--864, 2010.

② Takafumi Kanamori, Shohei Hido, Masashi Sugiyama,

A Least-squares Approach to Direct Importance Estimation.

Journal of Machine Learning Research. 査読有 10(Jul):1391--1445, 2009

③ Takeda, A., Kanamori, T.,

A Robust Approach Based on Conditional Value-at-Risk Measure to Statistical Learning Problems.

European Journal of Operational Research, 査読有, 198, pp. 287--296, 2009

④ Kanamori, T., Suzuki, T., Sugiyama,
Theoretical Analysis of Density Ratio Estimation.

IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. E93-A, no. 4, pp. 787--798, 2010

⑤ Takenouchi, T., Eguchi, S., Murata, N., Kanamori, T.

Robust Boosting Algorithm against Mislabelling in Multi-Class Problems.

Neural Computation, vol. 20, num. 6, pp. 1596--1630, 2008

[学会発表] (計 13 件)

① Kanamori T. and Ohara Atsumi

A Bregman extension of quasi-Newton updates .

Information Geometry and its Applications, Germany, August. 2010

② Shimodaira H. Kanamori T., Masayoshi A., Kouta Mine

Multiscale Bagging with Applications to Classification and Active Learning .

The 2nd Asian Conference on Machine Learning, November. 2010, Tokyo, Japan.

③ Takafumi Kanamori, Masashi Sugiyama, and Shohei Hido

Efficient Direct Density Ratio Estimation for Non-stationarity Adaptation and Outlier Detection

NIPS, 2008, *Vancouver, Canada, December, . 2008*

④ T. Kanamori, T. Suzuki, M. Sugiyama,
Condition Number Analysis of Kernel-based Density Ratio Estimation .

ICML workshop on Numerical Mathematics in Machine Learning, Montreal Canada, June 2009

⑤ T. Kanamori

Efficient direct importance estimation for covariate shift adaptation and outlier detection .

The 1st Institute of Mathematical Statistics, Asia Pacific Rim Meeting, Seoul, June 28-July 1, 2009

[図書] (計 2 件)

① T. Kanamori, H. Shimodaira,
[Chapter contribution] **Geometry of Covariate Shift with Applications to Active Learning,**

Dataset Shift in Machine Learning, MIT Press, 2008, pp.87-105.

② 金森敬文, 竹之内高志, 村田昇, パターン認識 (R で学ぶデータサイエンス 5), 共立出版, 2009 年, 総ページ 274 ページ. 約 90 ページ執筆.

[産業財産権]

○出願状況 (計 0 件)

名称 :

発明者 :

権利者 :

種類 :

番号 :

出願年月日 :

国内外の別 :

○取得状況 (計 0 件)

名称 :

発明者 :

権利者 :

種類 :

番号 :

取得年月日 :

国内外の別 :

[その他]