

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 5 日現在

機関番号：12601

研究種目：若手研究（B）

研究期間：2010 ～ 2012

課題番号：22700289

研究課題名（和文） データ横断的機械学習手法の理論と応用

研究課題名（英文） Theory and applications of cross-data-type machine learning methods

研究代表者

鈴木 大慈（SUZUKI TAIJI）

東京大学・大学院情報理工学系研究科・助教

研究者番号：60551372

研究成果の概要（和文）：

MKL (Multiple Kernel Learning) は Lasso や Group Lasso を含む広いクラスの正則化学習法である。画像認識など様々な応用で用いられている。本研究では、MKL の統計的性質を一般的な枠組みで解明し、どのような正則化を用いれば最適性を有するかについて調べた。また、MKL のベイジックの変種を考察し、それが緩い条件である種の最適性を満たすことを示した。さらに、大量データにおける構造的な正則化学習に有用なオンライン型の Alternating Direction Multiplier Method を提案し、その収束に関する理論的な正当性を与えた。

研究成果の概要（英文）：

We have investigated statistical convergence properties of Multiple Kernel Learning (MKL) with various types of regularizations. Moreover, we proposed a Bayesian variant of MKL and showed that it has optimality without strong assumptions on the design which are assumed in conventional theoretical analysis of MKL. We also proposed an online optimization method that is useful for structured regularization. It was shown that the proposed algorithm converges in the mini-max optimal rate.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010 年度	1000,000 円	300,000 円	1300,000 円
2011 年度	900,000 円	270,000 円	1170,000 円
2012 年度	900,000 円	270,000 円	1170,000 円
年度			
年度			
総計	2800,000 円	840,000 円	3640,000 円

研究分野：総合領域

科研費の分科・細目：統計科学

キーワード：統計的学習理論

1. 研究開始当初の背景

計算機やインターネットの発達に伴い、我々が扱うデータは多様化の一途をたどっている。検索エンジンではあらゆる種類のテキストデータが調べられるのみでなく、画像

や動画といったコンテンツにも手軽にアクセスでき、バイオインフォマティクスにおいても実験から得られるデータの種別は日々増大しつつある。例えばあるウェブページの内容をカテゴリ分類することを考えてみる。

従来ならばテキストの内容だけからその記事のカテゴリを推定することが行われてきたが、そのウェブページに関する他の情報として掲載されている画像やそのページ周辺のリンク構造やリンク先のデータなど、様々な付加情報を用いることにより推定精度の向上やアプリケーションの拡大が望まれる。このように多様なデータを用いて、多様な要求にこたえるべく、柔軟に情報を統合整理し提示するニーズが高まっている。

2. 研究の目的

機械学習を含めた情報科学の諸分野（自然言語処理や画像認識、バイオインフォマティクス等）において、近年、様々なデータソース（画像やテキスト、音声等）の情報を同時に扱い、いくつもの似たようなタスクを同時に解くような、関連のある事物をまとめて扱う問題設定が多なる重要性を帯びてきている。本研究は、増大しつつある多種多様なデータソースやタスクを総括的に扱う学習手法の統計的観点からの厳密な性能解析を軸とし、理論的な知見をもとにした新しい手法の導出およびその応用までを体系的に行うのが目的である。

3. 研究の方法

データソースの多様性については、スパース性の度合いを数量化し、様々な状況を含む包括的な理論的枠組みを構築する。次いで統計的理論で得た知見をもとに統計的性質のよいスパース学習手法の構築を行う。次いで、タスクの多様性に対応するため、マルチタスク学習についての研究を進める。マルチタスク学習に現れる状況設定を理論的に整理しその概念を拡張する形でノンパラメトリックな性能評価を与える。

さらに、年次を進めるにつれこれらの研究を統合する形で研究を進める。すなわち、マルチタスク学習におけるスパース学習を拡張し、多種多様なデータソースから情報を取捨選択し多種多様なタスクに柔軟に適用する手法を考案し、その統計的理論を各トピックを統合する形で構築してゆく。また、MKLをマルチタスク学習の問題設定に拡張し、柔軟さと効率性を求めてさらなる手法の洗練を最適化技法を駆使しながら実現してゆく。

4. 研究成果

(1) MKL (Multiple Kernel Learning) は LassoやGroup Lassoを含む広い枠組みであり、L1ノルムを正則化項とした正則化学習法であるとみなすことができる。拡張としてLpノルムやElasticnet型の正則化を用いた手法が提案されてきたが、興味深い事実として、近年多くの数値実験により、 $p>1$ における

Lp-MKLやElasticnet-MKLが普通のL1-MKLを精度の面で上回ることが知られるようになってきた。この事実は、Corinna CortesがICML 2009の招待講演で紹介し、広く機械学習業界に知られるようになった。「なぜ密な学習方法が疎なL1正則化を上回るのか」という問いは非常に重要な問題である。本研究では、この問いに対する理論的解答を与えるため、一般的な枠組み構築して包括的な考察を行った。この枠組みは全ての単調増加な混合ノルム型正則化に適用できるという点で特徴的である。我々の学習レートは局所化の技術を使っており、そのため大域的なバウンドよりもタイトである。我々の一般的な枠組みを具体例にあてはめることにより具体的な収束レートを導出することが可能である。また、その枠組みによって導出されるレートがミニマクスレートを達成することも示した。そして、「なぜL1正則化を密な正則化が上回るのか」という問いへの解答として、RKHSの“複雑さ”が一樣に等しいわけではないとき、密な正則化が疎なL1正則化を上回ることを示した。“複雑さ”が不均一という状況は実問題では非常に自然な設定である。

上では真が密な場合を想定していたが、真が疎な場合の解析も行った。ここではElasticnet型の正則化項に注目し、L1と比べてどのような違いがあるかを調べた。結論として、Elasticnet型正則化を使うことにより真の関数の滑らかさに応じて速い収束レートを達成できることが示された。一方、L1正則化を用いると、推定量がより疎になるため真の非ゼロ要素数に関するオーダーが小さくなる。よって滑らかさと疎性の間にトレードオフが生じ、もし真の関数が滑らかならElasticnetが、滑らかでない場合はL1が良いということが示唆された。

(2) MKLのベイズ的変種を提案し、その統計的性質をPAC-ベイズという手法を用いて理論的に評価した。これまでのMKLの標準的な理論解析においてはデザインにrestricted eigenvalue conditionといった強い仮定を課していた。しかし、この仮定は実際には満たされないことが多い。そこで、本研究ではMKLのベイズ的変種を考察し、それがデザインに仮定を置かずに最適収束レートを達成することを証明した。我々の証明は基本的にPAC-ベイズ法とガウシアンプロセス回帰の理論を組み合わせでなされる。既存のガウシアンプロセス回帰の統計的収束に関する理論においては、主にSobolev空間をモデルの空間として考えてきたが、我々の理論はinterpolation spaceの理論を用いてそれを一般化した。その意味で我々の手法は

既存のガウシアンプロセス回帰の理論を包含するものである。

(3) 大量データにおける構造的正則化学習に有用なオンライン型の Alternating Direction Multiplier Method (ADMM) を提案し、その収束に関する理論的な正当性を与えた。提案方法は、基本的に ADMM と確率的最適化技法との組み合わせである。確率的最適化手法の中でも特に広く用いられているオンライン型近接勾配法と双対平均化法に注目し、それら二つの手法の ADMM 型変種を考察した。提案手法は更新式が陽に書き下せるため計算が軽く、実装も容易である。また、提案手法の期待リスクはミニマクス最適レートを達成することを示した。計算機実験による既存手法との比較により提案手法の有用性を確かめた。

5. 主な発表論文等

(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 14 件)

① Taiji Suzuki, and Masashi Sugiyama: Fast learning rate of multiple kernel learning: trade-off between sparsity and smoothness. The Annals of Statistics, 2013.
<http://arxiv.org/abs/1203.0565>

② Makoto Yamada, Taiji Suzuki, Takafumi Kanamori, Hirotaka Hachiya, Masashi Sugiyama, Relative Density-Ratio Estimation for Robust Distribution Comparison. Neural Computation,

③ Takafumi Kanamori, Taiji Suzuki, and Masashi Sugiyama: Computational complexity of kernel-based density-ratio estimation: A condition number analysis. Machine Learning, vol. 90, pp. 431-460, 2013.

④ Taiji Suzuki, and Masashi Sugiyama: Sufficient dimension reduction via squared-loss mutual information estimation. Neural Computation, vol. 25, pp. 725-758, 2013.

⑤ Taiji Suzuki and Ryota Tomioka: SpicyMKL: A Fast Algorithm for Multiple Kernel Learning with Thousands of Kernels. Machine Learning, vol. 85, issue 1, pp. 77-108, 2011.

⑥ Taiji Suzuki, and Fumiyasu Komaki: On prior selection and covariate shift of

beta-Bayesian prediction under alpha-divergence risk. Communications in Statistics --- Theory and Methods, 39(8) (2010), 1655-1673.

[学会発表] (計 23 件)

① Taiji Suzuki: Some convergence results on multiple kernel additive models. Nonparametric and High-dimensional Statistics, 2012年12月18日, Luminy, France.

② Taiji Suzuki: PAC-Bayesian Bound for Gaussian Process Regression and Multiple Kernel Additive Model. Conference on Learning Theory (COLT2012), 2012年06月25日, Edinburgh, Scotland.

③ Taiji Suzuki and Masashi Sugiyama: Fast Learning Rate of Multiple Kernel Learning: Trade-off between Sparsity and Smoothness. Fifteenth International Conference on Artificial Intelligence and Statistics (AISTATS2012), 2012年04月21日, La Palma, Canary Islands.

④ 鈴木 大慈: 統計的学習理論チュートリアル: 基礎から応用まで. 第15回情報論的学習理論ワークショップ (IBIS2012), 2012年11月07日, 筑波大学東京キャンパス文京校舎.

⑤ Taiji Suzuki: Unifying Framework for Fast Learning Rate of Non-Sparse Multiple Kernel Learning. Advances in Neural Information Processing Systems 24 (NIPS2011), 2011年12月12日, Granada, Spain.

[図書] (計 1 件)

① Masashi Sugiyama, Taiji Suzuki, and Takafumi Kanamori: Density Ratio Estimation in Machine Learning. Cambridge University Press, 2012. 329 pages.

[産業財産権]

○出願状況 (計 0 件)

○取得状況 (計 0 件)

[その他]

ホームページ等

<http://www.simplex.t.u-tokyo.ac.jp/~s-taiji/>

6. 研究組織

(1) 研究代表者

鈴木 大慈 (SUZUKI TAIJI)
東京大学・大学院情報理工学系研究科・
助教
研究者番号：60551372