九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Robust and Efficient Prediction with Structural Constraints

森富, 賢一郎

https://doi.org/10.15017/1931929

出版情報：Kyushu University, 2017, 博士（情報科学）, 課程博士
バージョン：
権利関係：

KYUSHU UNIVERSITY

（別紙様式 2）

氏　　名　：森富　賢一郎

論 文 名　：Robust and Efficient Prediction with Structural Constraints
（構造的制約の下でのロバストかつ効率的な予測手法に関する研究）

区　　分　：甲

論　文　内　容　の　要　旨

本論文では，統計学習およびオンライン予測と呼ばれる２つの理論的な学習モデルを扱う．オンライン予測とは，アルゴリズムによる予測（意思決定）とフィードバックとしての環境からの情報の開示が繰り返される意思決定過程のモデルである．アルゴリズムの目標は，全ての試行が終了した後に判明する，最適な意思決定過程に匹敵する予測を行う事である．ただし，環境は敵対的な振る舞いをすると考えることで，最悪の場合の性能評価を行う．一方，統計学習では，アルゴリズムは入力として，何らかの確率分布に従って生成された，ラベル付きの事例集合が与えられる．各々のラベルは事例の正しい分類を表す．アルゴリズムは仮説として，事例からラベルへの関数を出力する．アルゴリズムの目標は，できるだけ汎化誤差の小さい仮説を出力することである．ただし，確率分布は任意かつ未知であるとし，最悪の分布における性能評価を行う．この学習モデルは統計学習の中でも特に PAC 学習として知られている．

本論文ではこれらの学習モデルの下で，構造的な制約下での予測・学習として定式化される問題を扱う．ここで構造的な制約とは，代数的あるいは組合せ論的な制約を指す．機械学習の問題の多くは，自然にこれらの構造的制約下での最適化問題として定式化することができる．

例えば，遅延時間を最小化するルーティングの問題は，グラフの全域木や s-t 路を逐次的に（オンラインで）予測する問題として定式化することができるが，これらは，組合せ最適化問題のオンライン予測版の問題と捉えることができる．またランキングに基づく推薦システムでは，アルゴリズムの出力する解をアイテム上の順列に制約していると考えられる．

代数的な制約の例として，行列補完問題がある．行列補完とは，複数のアイテムに複数のユーザが付与した評価値からなる行列を，一部の評価値から推定することである．この問題に対し，低ランク制約や行列ノルムの制約などの代数的な制約付きの最適化問題として定式化する手法が有効であることが知られている．

本論文では，2 つの理論的な学習モデルのもと，構造的制約の下での予測・学習として定式化される問題に対し，最悪時でも成立するロバストな理論保証と効率的な計算量の保証を持つアルゴリズムの設計と，その解析手法の開発を行った．以下に本論文の 4 つの主結果の概要を述べる．

1 つめは，行列のオンライン予測問題に関するものである．オンライン予測問題では，正則化項と呼ばれる凸関数をパラメータとするアルゴリズムの設計と解析手法がある程度確立されている．しかし，行列のオンライン予測問題に対して，従来手法では最適性を示すことができなかった．本研究では，正則化項の凸性に対する新しい概念を導入することにより解析手法を改善し，オンライン行列補完問題を含む様々な行列のオンライン予測問題に対して，最適で効率的なアルゴリズムを与えた．

2 つめは，オンライン二値行列補完問題に関するものである．この問題は，同一のアイテム集合に関する二値分類問題を複数のユーザに対して同時に解くものと見なせる．提案アルゴリズムは，アイテム集合が最適な特徴空間に写像され，各ユーザが最適な線形分類器で予測したと仮定したときの予測性能に匹敵する性能を，特徴写像や線形分類器を陽に求めることなく達成する．また，その予測性能が最適であることも示した. さらに, この手法を統計学習に応用して汎化性能を評価し，特徴写像が陽に与えられた場合のサポートベクトルマシン（SVM）の汎化性能に匹敵することを示した.

3 つめは，統計学習の枠組みのもとでの行列補完問題に関するものである．従来では，現実のユーザの評価は少数人の典型的な仮想ユーザの評価の線形結合で表せるという考えを，仮説となる行列の低ランク制約としてモデル化していた．本研究では，仮説となる行列に対し，低ランク制約に加え，行列のノルムに関する制約を置くことで，汎化性能が改善されることを示した．また，いくつかの実データ上で提案手法が有効であること示す実験結果を得た．この結果は，これまで理論評価が行われていなかった非負値行列分解に基づくアルゴリズムについて，その性能の根拠を示唆するものである.

4 つめは，組合せ集合上のメトリカルタスクシステム（MTS）問題に関するものである． MTS 問題とは，オンライン予測問題と同様に，情報の開示とアルゴリズムによる意思決定を繰り返すプロトコルに従う問題であるが，決定の変更に伴う損失も考慮する点が異なる．本研究では MTS 問題を組合せ集合上のサンプリング問題に帰着する手法を示した．この帰着手法と，組合せ集合上の効率的なサンプリング手法を同時に用いることにより，全域木や s-t 路など様々な組合せ集合上の MTS 問題に対し，世界初の効率的なアルゴリズムを構築することに成功した．また，提案アルゴリズムの予測性能が，ある組合せ集合上の MTS 問題に対して最適であることを示した.