

Statistical Multi-Task Learning for Robust Task Clustering via Fused Regularization

岡崎, 彰良

<https://hdl.handle.net/2324/7363603>

出版情報 : Kyushu University, 2024, 博士 (数理学), 課程博士
バージョン :
権利関係 :



氏 名	岡崎 彰良			
論 文 名	Statistical Multi-Task Learning for Robust Task Clustering via Fused Regularization (連結正則化法による頑健なタスク分類のための統計的マルチタスク学習)			
論文調査委員	主 査	九州大学	教授	川野 秀一
	副 査	九州大学	教授	廣瀬 慧
	副 査	九州大学	准教授	佃 康司
	副 査	広島大学	教授	柳原 宏和

論 文 審 査 の 結 果 の 要 旨

マルチタスク学習とは、複数のタスクが存在するときに、各タスクに対してモデルを独立に構築するのではなく、複数のタスクを同時に扱うことにより、タスク間の情報を共有してモデルの予測性能を向上させる統計的方法論である。たとえば、複数のデータ集合が同じ解析目的で取得された場合、これらを複数のタスクとして扱い、データ集合間に内在する共通構造を抽出して情報を共有する手法は、マルチタスク学習の一つのアプローチとして見なすことができる。マルチタスク学習の研究において、回帰モデルに基づく方法論の開発は基礎的かつ重要である。タスク間の情報共有を行うアプローチの一つとして、回帰モデルの枠組みでは、連結正則化法に基づく方法がある。この方法では、各タスクに対して回帰モデルを設定し、タスク間の回帰係数にスパース性を誘導する正則化項を導入することで、類似タスク間の回帰係数を統合し、情報共有を実現する。しかし、従来の連結正則化法に基づく方法では、タスク集合内に異質なタスクが混在する状況には十分対応できておらず、そのために推定精度や予測精度の低下が生じることが知られている。

本研究者は、異質なタスクが混在する状況から生じる問題に対応する、新たなマルチタスク学習を統計モデルの観点から提案している。二つの提案を行っており、この二つの研究成果を纏めたものが本博士論文である。本論文は全 6 章より構成され、第 4 章と第 5 章に本研究者が行った研究が述べられている。以下、本論文の構成に沿って、各章を要約する。

第 1 章は導入であり、本研究の目的、先行研究、得られた研究成果を要約するとともに、本論文の構成について述べている。

第 2 章は一般化線形モデルの枠組みの下、マルチタスク学習のこれまでの研究を紹介している。まず、一般的な問題設定下でマルチタスク学習における最適化問題を定式化した後、回帰係数に低ランク性を仮定した方法、スパース推定による方法、クラスタリングによる方法、回帰係数行列を和の形に分解する方法について述べ、最後にこれら既存手法間の関係性について議論している。

第 3 章は凸クラスタリングに関するこれまでの研究を紹介している。凸クラスタリングとは、従来のクラスタリング手法における非凸最適化問題を、凸緩和することで容易に解けるように定式化した手法である。凸クラスタリングの最適化問題を定式化した後、交互方向乗数法に基づく推定アルゴリズムについて述べている。第 2 章および第 3 章は、第 4 章以降の研究成果の内容を理解するために必要な基礎知識を網羅的に解説したものであり、博士論文として十分な量の解説と認められる。

第4章では、本研究者によって得られた成果のうち、過剰に統合される回帰係数に対処したマルチタスク学習に関する研究成果を述べている。従来の連結正則化法に基づく方法では、類似性が低いタスク間の回帰係数を過剰に統合する傾向にあり、これにより推定精度や予測精度が悪化することがある。そこで、タスクのクラスターの重心を表すパラメータを新たに導入し、回帰係数ではなく重心パラメータを統合することにより、各タスクの回帰係数が類似性の低いタスクから受ける影響を軽減する方法を提案している。本手法は従来のマルチタスク学習に凸クラスタリングの考えを援用することにより定式化され、ブロック座標降下法に基づいた推定アルゴリズムを提案している。モンテカルロ・シミュレーションや実データへの適用を通して、提案手法の有用性を検証している。なお、本研究成果は本研究者を筆頭として、査読付きの国際学術雑誌に公刊されている。

第5章では、独自性が大きい外れ値タスクを検出するマルチタスク学習の研究成果を述べている。独自性が大きい外れ値タスクが存在している状況においては、第4章の提案手法を用いたとしても、その外れ値タスクの影響を受けてしまい、各タスクの回帰係数の推定精度や予測精度が悪化する。そこで、各タスクに対して外れ値パラメータを新たに導入し、外れ値パラメータの値をグループ正則化法で取捨選択することにより、外れ値タスクに頑健な手法を提案している。具体的には、凸クラスタリングにおいて外れ値を扱うことができるロバスト凸クラスタリングを非凸に拡張し、その非凸化した方法と第4章で提案した方法とを組み合わせることにより、新たな手法を提案している。提案手法はロバスト統計の枠組みで捉えることができ、その関係性についても議論している。推定アルゴリズムを交互方向乗数法に基づき構築し、その計算量を導出するとともに、アルゴリズムの収束性を証明している。人工データや実データへの適用を通して、既存手法と比較検討することにより、提案手法の有用性を数値的に検証している。なお、本研究成果も本研究者を筆頭として、査読付きの国際学術雑誌に公刊されている。

第6章はまとめであり、本研究で得られた成果の要約と今後の研究の展望を述べている。

以上のように本研究者は、マルチタスク学習の開発研究において顕著な研究成果を上げており、これらは数理統計学の分野において価値ある業績と認められる。

よって、本研究者は博士（数理学）の学位を受ける資格があるものと認める。