

特別研究報告書

共通のサポートベクターをもつ
ベクトル値 SVM とその数値解法

指導教員 山下信雄 教授
福田エレン秀美 准教授

京都大学工学部情報学科
数理工学コース
平成 29 年 4 月入学

羽原 圭吾

令和 3 年 2 月 12 日提出

摘要

教師あり学習とは、機械学習の手法の一つである。教師データと呼ばれる入力とラベルの組を利用して学習し、未知の入力に対してそのラベルを予測することを目的とする。サポートベクターマシンは、教師あり学習のモデルの一つである。入力がある非線形写像を用いて特徴空間へと写し、その特徴空間における線形な関数をモデルとする。サポートベクターマシンによって未知の入力に対してそのラベルを予測するとき、教師データの一部と、未知の入力とのカーネルを計算する必要がある。カーネルとは、特徴空間における内積のことである。このような教師データの一部をサポートベクターと呼ぶ。

サポートベクターの割合を小さくすることは、カーネルの計算コストを抑えるために重要である。Schölkopfらは、教師データの大きさに対して、サポートベクターの割合が事前に決めたパラメータに漸近するようなモデルを提案した。しかしラベルがベクトル値であるような教師あり学習に対して、このモデルを適用することは難しい。最も単純な方法は、ラベルの次元ごとに別々のサポートベクターマシンで学習する方法であるが、この分割的なアプローチではサポートベクターの割合を正確に制御することはできない。

本研究で提案するモデルは、ベクトル値をまとめて扱うことでサポートベクターの割合の制御を可能とした。特に実数値ラベルを予測する回帰問題に対して、サポートベクターの割合の漸近性を示した。また、提案モデルの効率的な解法を提案した。凸最適化の高速なアルゴリズムを用いた数値解法を提案し、さらに問題特有の構造を利用することで計算コストを削減した。最後に数値実験を行い、提案モデルが上述の分割的なアプローチに比べてサポートベクターの割合を良く制御できていることを確認した。また、計算時間、反復回数、予測の性能といった観点から、提案した数値解法の性能を検証した。