

Frank-Wolfe-based Boosting Algorithms and Its Extensions

三星, 諒太郎

<https://hdl.handle.net/2324/7329498>

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



氏 名 : 三星 諒太郎

論文名 : Frank-Wolfe-based Boosting Algorithms and Its Extensions
(フランク・ウルフ法に基づくブースティング手法とその拡張)

区 分 : 甲

論文内容の要旨

ブースティングは機械学習の枠組みの一つで、必ずしも性能が高くない学習アルゴリズム(以下、弱学習器)が出力する複数の仮説を統合することで、精度の高い仮説を構築する手法の総称である。任意の弱学習器と組み合わせることができる柔軟性と、その有用性から、ブースティングは機械学習の標準的なツールとして広く用いられている。これまで様々なブースティング手法が提案されてきたが、その多くは、非常に多くの(ときには無限個の)制約を持つ最適化問題の効率的な解法とみなせることが知られている。本論文は、これらのブースティング手法が、フランク・ウルフ法と呼ばれる汎用の凸最適化アルゴリズムとして統一的に捉えることができることを示した上で、フランク・ウルフ法の設計と解析の技術に基づき、ブースティング手法の改善、およびその拡張に関する4つのテーマについて取り組んだものである。以下にその概要を述べる。

1 つめは、ソフトマージン最適化問題として定式化されたブースティング問題に対する既存アルゴリズムの解析、およびその効率化に関するものである。ソフトマージン最適化に基づくブースティング手法は、経験誤差最小化に基づく手法より優れた汎化性能を持つことが知られている。しかしながら、これまでに提案されたソフトマージン最適化手法はいずれも、実用的には極めて高速だが理論的な収束保証がないか、あるいは、理論的な収束保証を持つが実的に遅いという問題があった。本研究では、まず、理論的な性能保証を持つ既存手法が、いずれもフランク・ウルフ法とみなせることを示し、フランク・ウルフ法の解析技法に基づいて、その収束性の証明を非常に単純化することに成功した。また、フランク・ウルフ法の各繰り返し過程において、理論保証を持つ解の更新規則だけでなく任意の更新規則 A も同時に適用してみて、目的関数値をより減少させる方を採用することとしても収束性の保証は変わらないことに注目し、更新規則 A として実用的に高速なブースティング手法の更新規則を用いることにより、既存手法と同じ収束保証を持ち、かつ実用面でも高速なブースティング手法を得ることに成功した。

2 つめは、回帰問題に対する効率的なブースティング手法の開発である。回帰問題に対するブースティングに関する既存手法として **XGBoost** や **LightGBM** の回帰版があるが、理論的な収束保証を得るには、各繰り返し過程において弱学習器がある意味で最適な仮説を出力しなければならないという、非現実的な仮定をおく必要がある。また、平均二乗誤差最小化に基づく回帰問題ではなく、いわば回帰版のソフトマージン最適化として定式化された回帰ブースティング問題が提案されているが、収束保証を持つ解法は知られていない。本研究では、回帰版のソフトマージン最適化問題に対し、フランク・ウルフ法を適用することにより、収束保証を持つ世界初のブースティング手法を得ることに成功した。

3 つめは、拡張定式化の自動生成と、ソフトマージン最適化問題への応用に関するものである。

0-1 係数の線形制約集合は、係数行列を集合族とみなすことにより、NZDD と呼ばれる有向グラフ型のデータ構造を用いて圧縮表現することができる。本研究では、その NZDD から、元の線形制約の拡張定式化、すなわち、元の線形制約と等価でサイズの小さい線形制約集合を生成するアルゴリズムを構築した。したがって、0-1 係数の線形制約を大量に含む最適化問題に対し、本手法を最適化ソルバーの前処理として用いることにより、その効率化に寄与することが期待できる。実際、本研究では、種々の整数計画問題や混合整数計画問題に対し、本手法の有効性を実験的に示した。さらに、本研究では、この手法をソフトマージン最適化問題に適用し、訓練データの圧縮表現上に基づく新しい定式化を与えた上で、これを解く高速なブースティング手法を開発した。

4 つめは、オンライン組合せ線形最適化問題の効率化に関するものである。この問題は、 n 次元ベクトルの有限集合 K を実行可能解集合とする逐次的な意思決定問題である。通常の（オフラインの）問題が、損失ベクトル $\mathbf{l} \in \mathbb{R}^n$ が与えられたとき内積 $\mathbf{x} \cdot \mathbf{l}$ を最小とする解 $\mathbf{x} \in K$ を求める問題として定義されるのに対し、オンライン問題では、各ラウンド $t = 1, 2, \dots, T$ に対し、アルゴリズムが解 $\mathbf{x}_t \in K$ を出力した後で損失ベクトル $\mathbf{l}_t \in \mathbb{R}^n$ が与えられる。アルゴリズムの目標は、累積損失 $\sum_{t=1}^T \mathbf{x}_t \cdot \mathbf{l}_t$ を最小化することである。この問題の標準的な解法では、 K の凸包上の解 $\hat{\mathbf{x}}_t$ を、ある条件を満たす K の解 $\mathbf{x}_t$ に変換する「メタラウンディング」アルゴリズムの存在を仮定している。オフライン問題が LP 緩和法による近似アルゴリズムを持つとき、メタラウンディングアルゴリズムが存在することは知られているが、既存のメタラウンディングアルゴリズムには、その効率に問題があった。本研究では、メタラウンディングがハードマージン最適化に類似した最適化問題として定式化されることに注目し、フランク・ウルフ法を適用することで、計算量が大幅に小さいメタラウンディングアルゴリズムを構築することに成功した。