

研 究 報 告 書

「オンライン劣モジュラ最適化に対する
効率的かつ汎用的なアルゴリズムの開発」

研究期間: 2017 年 9 月 ~ 2019 年 3 月

研究者番号: 50144

研 究 者: 相馬 輔 (東京大学情報理工学系研究科 助教)

1. 研究のねらい

近年,劣モジュラ関数を利用したモデルが,機械学習,圧縮センシング,ネットワーク科学といった分野で注目されている.劣モジュラ最適化にもとづくアプローチは,モデリングの汎用性が高く,効率的なアルゴリズムが存在することもある.しかしながら,既存の劣モジュラ最適化によるアプローチは,最適化すべき劣モジュラ関数が既知である必要がある.機械学習などでは,学習対象は本質的に不確実性を持っており,既存の劣モジュラ最適化によるアプローチでは不確実性をもつ最適化問題に対処できない.そこで本申請課題では,劣モジュラ最適化を不確実性を考慮した最適化へ拡張する.具体的には,不確実性を考慮した最適化であるオンライン最適化を,劣モジュラ最適化に対して適用し,理論保証付きの効率的かつ汎用的なアルゴリズムの開発を目指す.

(未来ビジョン) 本課題では,オンライン最適化の技法を通じて,「不確実性を考慮した劣モジュラ最適化」という枠組みを確立することを目指している.既に,劣モジュラ最適化は幅広い問題に応用されているため,「不確実性を考慮した劣モジュラ最適化」は,実問題に即した設定にも適用できる強力な枠組みとなる.また,劣モジュラ最適化の最大の利点は,高速なアルゴリズムに確固たる理論保証があることである.単なるヒューリスティックではなく,理論保証付きの高速なアルゴリズムを提供することは,機械学習の研究コミュニティだけでなく,機械学習を利用したプロダクト開発を行うユーザーにも大きなインパクトがある.

2. 研究成果

(1) 概要

オンライン劣モジュラ最適化に関して,下記の2つの研究成果が得られた.

第一に,非単調劣モジュラ最大化や k 劣モジュラ最大化の拡張である,オンライン k 劣モジュラ最大化に対して,劣線形の $1/2$ リグレットを達成する多項式時間アルゴリズムを得た.本成果は,オンライン非単調劣モジュラ最適化に対する既存アルゴリズムを k 劣モジュラ最大化に拡張しただけでなく,既存アルゴリズムで場当たりのに解かれていた部分を整理して,より簡潔な証明を与えたものである.本成果は機械学習国際会議 AISTATS 2019 に単著論文として採択された.

第二に,辞書選択と呼ばれる機械学習の問題に対して,劣モジュラ最適化の技法を応用した高速な貪欲アルゴリズムを与えた.辞書選択は,圧縮センシングやスパースモデリングで必要となる基底集合をデータから学習するための手法である.既存手法では,難しい非凸最適化問題を解く必要があったり,簡単に解ける荒い近似をしたために実際の性能が劣化するなどの欠

点があった。本研究では、二段階劣モジユラ最適化の技法を圧縮センシングの技法と組合せて、辞書選択に対する高速な貪欲アルゴリズムを提案した。提案アルゴリズムは、理論的な近似比保証があるだけでなく、非凸最適化に基づく手法で得られる解と同程度の品質の解を 10～100 倍高速に求めることができ、理論的にも実用的にも有効なアルゴリズムである。また、提案アルゴリズムは複雑な疎性制約やオンライン辞書選択にも拡張できることを示した。本研究は、機械学習トップ会議 NeurIPS 2018 の spotlight(採択論文中上位 20%論文)として採択された。

(2) 詳細

研究テーマ A「オンライン k 劣モジユラ最大化に対する no-regret アルゴリズム」

非単調劣モジユラ最大化は、与えられた非負劣モジユラ関数を最大にする集合を求める問題である。オフラインの場合(関数が明示的に与えられる場合)に対する $1/2$ 近似アルゴリズムが存在し、 $1/2$ 近似を破ることは多項式回の関数値クエリでは不可能であることがすでに分かっている。ごく最近、これをオンライン(関数が逐次的に与えられる場合)に拡張したオンライン劣モジユラ最大化に対して、劣線形の $1/2$ リグレットを持つアルゴリズムが Roughgarden-Wang (2018)によって提案された。

本研究では、これをオンライン k 劣モジユラ最大化に拡張した。 k 劣モジユラ関数は、劣モジユラ関数や双劣モジユラ関数など組合せ最適化・機械学習で重要な関数の一般化である。例えば、 $k=1$ の場合は劣モジユラ関数に対応し、 $k=2$ の場合は双劣モジユラ関数に対応する。オフラインの k 劣モジユラ最大化に対しては、やはり $1/2$ 近似アルゴリズムが Iwata-Tanigawa-Yoshida (2016)によって得られている。本研究では、オンライン k 劣モジユラ関数最大化に対して劣線形の $1/2$ リグレットを達成する多項式時間アルゴリズムを与えた。特に、劣モジユラ関数($k=1$)の場合には、Roughgarden-Wang の既存アルゴリズムと同様のリグレット上界を達成する。

さらに、得られたアルゴリズムは Roughgarden-Wang を単純に拡張したわけではなく、より簡潔な解析を与えるものになっている。具体的には、彼らのアルゴリズムでは、サブルーチンとして、エキスパート問題に対して強い形のリグレット保証をもつオンラインアルゴリズムが必要であった。しかしながら、Roughgarden-Wang のサブルーチンは場当たり的に設計されており、正当性の証明はもちろんあるとはいえ、どうしてこのアルゴリズムで所望の強いリグレット保証が得られるのかが不明確であった。また、 k 劣モジユラ関数のような、より複雑な状況に対してどのように拡張すべきか明らかではなかった。本研究では、Blackwell の定理というオンライン学習の結果を利用し、Roughgarden-Wangとは異なるアプローチで、同様のリグレット保証をもつサブルーチンを設計した。これにより、証明が簡潔になると同時に、 k 劣モジユラ最大化への拡張も容易に行えることがわかった。

本研究は単著論文として機械学習国際会議 AISTATS 2019 に採択された。

研究テーマ B「辞書選択に対する高速な貪欲アルゴリズム」

圧縮センシングやスパースモデリングといった機械学習手法においては、与えられたデータをスパースに(疎に)表現するための基底集合が必要である。例えば、通常の写真などであれば、フーリエ基底やウェーブレット基底といった基底集合が用いられる。このような基底集合は、かつては各分野の膨大な知識の蓄積のもとに構築されていた。一方で、そのような知識の蓄積がないドメインに対しても、スパース表現を与える基底を得たいというのは自然である。辞書学習・辞書選

択は、データからスパース表現の基底を求める手法である。

辞書学習はデータのみでスパース表現を学習できるが、難しい非凸最適化問題を解く必要があり、理論的に保証をつけるのが比較的難しく、また実際に運用するとかなりの時間がかかってしまう。辞書選択は、すでに知られている多数の基底集合から基底ベクトルを選び出すことで、スパース表現の基底を求める手法である。データの他に、基底集合の候補が必要になるものの、組合せ最適化の技法で非常に高速に解くことができる。一方で、既存の辞書選択アルゴリズムでは、最終的に得られる解は辞書学習の解に対して品質が劣ることが多かった。

本研究では、二段階劣モジュラ最適化という劣モジュラ最適化の枠組みで用いられた手法を、Orthogonal Matching Pursuit という圧縮センシングの技法と組合せて、辞書選択に対する新たな貪欲アルゴリズムを提案した。提案アルゴリズムは、辞書学習で得られる解とほぼ同程度の解を10～100倍高速に得られる。また、アルゴリズムの出力に対して近似比保証を与えることにも成功した。さらに、提案アルゴリズムの別の利点として、これまで扱うことが難しかった複雑な疎性制約（平均疎性制約やブロック疎性制約）を扱うことも可能になった。辞書学習で必要となるデータは膨大であり、データをすべて保持することなく逐次的に処理できるオンライン最適化は重要である。提案アルゴリズムは自然にオンライン最適化へと拡張できることもあわせて示した。

上記の成果は東京大学の藤井海斗氏との共同研究で行われたものである。成果は共著論文として、機械学習系トップ会議 NeurIPS 2018 の spotlight paper (投稿論文中上位 20%論文) に採択された。

3. 今後の展開

オンライン k 劣モジュラ最大化のオンラインアルゴリズムは非常に簡潔であるため、オンライン劣モジュラ最大化に帰着される機械学習タスクでは、ベースライン手法として利用されることが期待できる。また、辞書選択の新しいアルゴリズムは、高速に高品質な解を求められるため、既存の辞書学習アルゴリズムの代用、あるいはホットスタート用の前処理として活用できる。

4. 自己評価

- 研究目的の達成状況: NeurIPS 2018 には spotlight 論文が採択、AISTATS 2019 には単著論文が採択され、想定以上の結果が得られた。
- 研究の進め方(研究実施体制及び研究費執行状況): ACT-I 採択後にオンライン非単調劣モジュラ最大化が海外のグループに先取されるハプニングがあったものの、最終的には共著者の協力を得て2つの成果を取りまとめることができた。また、研究期間中には、ISMP, ICALP, NeurIPS などの国際会議に参加し、自身の成果発表と同時に、先端のアルゴリズム技法の情報を収集した。さらに、Max Planck Institute for Mathematics in Sciences に短期訪問し、機械学習の最新の成果について情報交換を行った。
- 研究成果の科学技術及び学術・産業・社会・文化への波及効果(今後の見込みも重視してください。): 学術・産業に対する直接の波及効果は「今後の展開で」で述べた通りである。また、劣モジュラ最適化を ACT-I に参加する幅広い分野の研究者に広めることができた。
- 研究課題の独創性・挑戦性: 既存の劣モジュラ最適化では、不確実性のある状況設定はあまり考えられてこなかった。この点、不確実性を考慮した劣モジュラ最適化の観点は新しい。また、オンライン最適化のような難しい設定でも理論保証付きのアルゴリズムを与える

ことは、理論上興味深い。

5. 主な研究成果リスト

(1) 論文(原著論文)発表

1. K. Fujii and T. Soma, "Fast Greedy Algorithms for Dictionary Selection with Generalized Sparsity Constraints," Advances in Neural Information Processing Systems (NeurIPS), pp. 4744–4753, 2018. [Spotlight: 採択論文の中でトップ 20% 論文](#)
2. T. Soma "No regret algorithms for online k-submodular maximization," Proceedings of the The 22nd International Conference on Artificial Intelligence and Statistics (AISTATS), 2019, to appear.
3. T. Soma and Y. Yoshida, "A New Approximation Guarantee for Monotone Submodular Function Maximization via Discrete Convexity," Proceedings of the 45th International Colloquium on Automata, Languages, and Programming (ICALP), 2018.

(2) 特許出願

なし

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

学会発表

1. 相馬輔, 「No-regret algorithms for online k-submodular maximization」, 応用数理学会連合研究発表会, 口頭発表, つくば, 2019.
2. K. Fujii and T. Soma, "Fast Greedy Algorithms for Dictionary Selection with Generalized Sparsity Constraints," Advances in Neural Information Processing Systems (NeurIPS), ポスター発表, モントリオール, 2018.
3. 相馬輔, 「No-regret algorithms for online k-submodular maximization」, 本部 SSOR, 口頭発表, 水上温泉, 2018.
4. T. Soma and Y. Yoshida, "A New Approximation Guarantee for Monotone Submodular Function Maximization via Discrete Convexity," the 45th International Colloquium on Automata, Languages, and Programming (ICALP), 口頭発表, プラハ, 2018.
5. T. Soma and Y. Yoshida, "A New Approximation Guarantee for Monotone Submodular Function Maximization via Discrete Convexity," the 23rd International Symposium on Mathematical Programming, 口頭発表, ボルドー, 2018.

アウトリーチ活動

1. 相馬輔, 「劣モジュラ最適化入門」, 第3回 ACT-I 応用数学研究会, 講師, 2019.
2. 相馬輔, 「劣モジュラ最適化入門」, 自動車制御とモデル研究部門委員会, 依頼講演, 2018