様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科研費
KAKENHI

科学研究費助成事業　　研究成果報告書

平成 ３０ 年　 ５ 月 ２１ 日現在

機関番号：１２６０１

研究種目：研究活動スタート支援

研究期間：2016～2017

課題番号：１６Ｈ０６６７６

研究課題名（和文）劣モジュラ最適化による機械学習モデルとアルゴリズムの研究

研究課題名（英文）Studies of models and algorithms in machine learning via submodular optimization

研究代表者

相馬　輔（Soma, Tasuku）

東京大学・大学院情報理工学系研究科・助教

研究者番号：９０７８４８２７

交付決定額（研究期間全体）：（直接経費）　　2,300,000円

研究成果の概要（和文）：近年，劣モジュラ最適化という組合せ最適化の手法が，機械学習などの分野で注目されている．本研究では，劣モジュラ最適化と機械学習への応用について研究し，以下のような成果を得た．
1．これまでの「曲率」と呼ばれる指標にもとづく近似比評価を改善し，離散凸性にもとづいた新たな近似比評価指標とアルゴリズムを開発した．
2．辞書学習と呼ばれる圧縮センシング・機械学習の問題に対して，新しい組合せ的なアルゴリズムを与えた．

研究成果の概要（英文）：Recently, submodular optimization -- a branch of combinatorial optimization -- has attracted interests in the machine learning community. In this project, we studied submodular optimization and its applications to machine learning, and obtained the following results:
1. We developed new approximation guarantee based on the concept of discrete convexity, which improves previous approaches based on the concept of curvature.
2. For dictionary learning (a problem studied in compressed sensing and machine learning), we devise a new combinatorial algorithm.

研究分野：組合せ最適化

キーワード：組合せ最適化　機械学習

1　研究開始当初の背景

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．**劣モジュラ関数**とは，任意の $X, Y \subseteq S$ に対して $f(X)+f(Y) \geq f(X \cup Y)+f(X \cap Y)$ を満たす集合関数 f のことである．2000 年代になって，機械学習に現れる集合関数が自然に劣モジュラ性を持つことが認識されてきた．特に，単調な劣モジュラ関数を様々な制約のもとで最大化する問題は，**単調劣モジュラ関数最大化**と呼ばれ，文書要約，センサー配置問題，感染型マーケティングなど多くの問題で共通に現れる定式化として知られている．単調劣モジュラ関数最大化は，最も単純なサイズ制約でも NP 困難だが，**貪欲アルゴリズム**と呼ばれる近似アルゴリズムが知られており，近似比の上限である $1-1/\mathrm{e}$ 近似を達成できる．貪欲アルゴリズムは実用的には大変良い解を出力することが多くの事例で観察されており，最適値の 90% 以上の値の解を出力することも多い．

定式化の汎用性と効率的なアルゴリズムの存在から，劣モジュラ関数にもとづく機械学習手法は非常に注目を集めている．実際，**機械学習のトップ会議である ICML では 2009 年，2013 年の 2 回にわたり劣モジュラ最適化のチュートリアル講演があり，同じく機械学習トップ会議の NIPS でも多くの劣モジュラ最適化の研究者が基調講演を行っている．**

2　研究の目的

このように単調劣モジュラ関数最大化は，機械学習と理論計算機科学の両分野で様々なモデルやアルゴリズムが開発されているが，既存研究で分かっていないこと・解決すべき点として以下が挙げられる．

1．複雑な制約に対する高速な近似アルゴリズム　2000 年代から，マトロイド制約など従来より複雑な制約について近似アルゴリズムが提案されているが，計算量が大きいため機械学習で実際に使われるには至っていない．**実問題に現れる重要な制約に対して，実用的なアルゴリズムを開発することは依然として大きな課題の 1 つである．**本研究では，**連続貪欲法**と呼ばれる近年提唱された広範な枠組みを高速化することで，複雑な制約に対して実用的なアルゴリズムを与えることを目指す．

2．理論評価と実性能とのギャップを埋める新たな近似比評価指標　実データへアルゴリズムを適用したときに，実際にどの程度の解が得られるのかは，ユーザーが一番気になる点である．実は，多くのアルゴリズムで，理論的な近似比評価と実際に観測される近似比の間には大きなギャップがある．**なぜ実データでは理論的な近似比よりずっと良い解が得られるのか，満足のいく理論的な説明は未だに存在しない．**本研究では，離散凸性にもとづいて，現実の挙動をより正確に予測できる新たな近似比評価指標を開発する．

3．単調劣モジュラ関数最大化にもとづく新たな機械学習モデル　単調劣モジュラ関数最大化を利用した既存モデルは，被覆関数や相互情報量の劣モジュラ性にもとづいたものがほとんどで，似通った現象にしかモデルが提案されていない．**機械学習に現れる問題に対して，単調劣モジュラ関数最大化にもとづく新たなモデルを構築することも重要な課題である．**本研究では，単調劣モジュラ関数最大化を用いた新たなモデルとして，**圧縮センシング**に対する新たなモデルとアルゴリズムを提案する．

3　研究の方法

研究目標「**単調劣モジュラ関数最大化理論の深化と高速な近似アルゴリズムの開発を通じて，劣モジュラ性にもとづく新たな機械学習モデルの開発**」を達成するため，以下の具体的な研究計画を提案する．

- **連続貪欲法の効率的な丸めアルゴリズムの開発** ... 連続貪欲法のボトルネックを高速化することで，複雑な制約に対する実用的なアルゴリズムを与える．

- **離散凸性による新たな近似比評価指標の開発** … これまでの曲率にもとづく近似比評価を一般化し，離散凸性にもとづく新たな近似比評価指標を開発する．
- **単調劣モジュラ関数最大化を用いた疎性モデルの開発** … 近年，機械学習などで注目されている圧縮センシング・スパースモデリングに対し，劣モジュラ性を用いた統一的な説明を目指す．

4 研究成果

「実際の性能を予測できる新たな近似比評価指標」と「複雑な制約に対する高速な近似アルゴリズム」については以下の成果が得られた．単調劣モジュラ関数と M 凹関数の和で書かれる劣モジュラ関数の最大化について，前者の項については $1-1/e$ 近似，後者の項については最適値を達成するアルゴリズムを開発した．これは既存研究で扱われてきた曲率と呼ばれる指標に基づく近似比評価を一般化したものであり，曲率では捉えられなかった関数について近似比評価を改善するものである．提案アルゴリズムは連続貪欲法と呼ばれる手法に M 凹関数の凹閉包を組合せたものだが，連続緩和解を離散解に丸める過程で，既存のスワップ丸めを M 凹関数制約に一般化した丸めアルゴリズムも得られた．本成果は，"A New Approximation Guarantee for Monotone Submodular Function Maximization via Discrete Convexity" として，**理論計算機科学系国際会議 ICALP 2018 に採択された**．

「単調劣モジュラ関数最大化を用いた疎性モデルの開発」について，辞書選択 (dictionary selection) と呼ばれる圧縮センシング・機械学習の問題に対して，ロス関数の制限強凸性を利用した組合せ的近似アルゴリズムを開発した．辞書選択における目的関数は劣モジュラ関数とは限らない．従来，非劣モジュラな目的関数に対しては，劣モジュラ比 (submodularity ratio) と呼ばれる指標で近似比評価がなされてきた．対して，本手法では制限強凸性に注目することで，劣モジュラ比を用いずに近似アルゴリズムを与えた点が新しい．また，これらの結果をオンライン最適化の設定に拡張することにも成功した．本成果は機械学習系国際会議に投稿中である．

また，本研究計画の直接の内容ではないものの，劣モジュラ最適化と機械学習に関連した成果として，整数格子点上の非単調劣モジュラ関数最大化に対する 1/2 近似アルゴリズムと，多目的単調劣モジュラ最大化に対するリグレット比最小化アルゴリズムを得た．これらの成果は機械学習系国際会議 AAAI 2017 に採択され，2017 年 2 月に口頭発表を行った．

5 主な発表論文等

■〔雑誌論文〕（計 4 件）

1. T. Soma and Y. Yoshida, "Regret Ratio Minimization in Multi-objective Submodular Function Maximization," *Proceedings of the 31st AAAI Conference on Artificial Inteligence*, 2017.
2. T. Soma and Y. Yoshida, "Non-monotone DR-Submodular Function Maximization," *Proceedings of the 31st AAAI Conference on Artificial Inteligence*, 2017.
3. 相馬 輔,『整数格子点上の劣モジュラ最大化と近似アルゴリズム』, オペレーションズ・リサーチ, 63(3), pp.36–42, 2018.
4. 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, to appear.

■〔学会発表〕（計 7 件）

1. T. Soma and Y. Yoshida, "Regret Ratio Minimization in Multi-objective Submodular Function Maximization," The 31st AAAI Conference on Artificial Inteligence, 2017.

2. T. Soma and Y. Yoshida, "Non-monotone DR-Submodular Function Maximization," The 31st AAAI Conference on Artificial Inteligence, 2017.
3. T. Soma and Y. Yoshida, "Regret Ratio Minimization in Multi-objective Submodular Function Maximization," The 10th Japanese-Hungarian Symposium on Discrete Mathematics and Its Applications, 2017.
4. 相馬 輔, 『整数格子点上の劣モジュラ最大化と近似アルゴリズム』[**招待有り**], RAMP シンポジウム, 2017.
5. 相馬 輔, 吉田悠一,『多目的劣モジュラ最大化に対するリグレット比最小化』, OR 学会春季研究発表会, 2017.
6. 相馬輔, 吉田悠一, 『離散凸性による劣モジュラ最大化の近似比保証』, Japanese Center for Combinatorics and its Applications Japanese Conference on Combinatorics and its Applications, 2017.
7. 相馬輔, 吉田悠一, 『離散凸性による劣モジュラ最大化の近似比保証』, 日本オペレーションズ・リサーチ学会 「離散アルゴリズムの応用と理論」研究部会, 2018.

6 研究組織

(1) 研究代表者
相馬 輔（東京大学大学院　情報理工学系研究科）, 研究者番号： 90784827
(2) 研究分担者
なし
(3) 連携研究者
なし
(4) 研究協力者
なし