

数理・情報のフロンティア
2020 年度採択研究者

2021 年度 年次報告書

黒木 祐子

東京大学 大学院情報理工学系研究科
助教

限られた情報に基づく統計的機械学習と数理最適化アルゴリズムの開発

§ 1. 研究成果の概要

本研究課題では限られた情報に基づく統計的機械学習の理論構築に取り組み、オンライン意思決定問題の一つであるバンディット問題を対象に、数理最適化理論および統計学的観点の両側面からの解決法の提案と理論解明を行った。

バンディット問題とは、実際に試行した行動に対してのみ情報が得られるという制限下で、最適な意思決定を目指すオンライン学習問題である。その解法は、推薦システム、ネット広告表示の最適化、薬の治験など、様々な重要課題に応用され、有効性が示されている。更に複雑な問題である、推薦システムにおける類似商品の集合、通信ネットワークにおける接続形態、道路ネットワークにおける経路などの最適化は、組合せ的な構造により特徴付けられており、このような一般化されたバンディット問題は、組合せバンディットと呼ばれている。

今年度は、昨年度に引き続きこの確率的組合せバンディット問題に関する数理課題について取り組んだ。組合せ構造を含むバンディット問題では、報酬関数が非線形である場合既存のアルゴリズムは適用が難しかった。本研究では、非線形関数の新しいクラスとして、選択した行動を構成する要素のうち最も報酬が小さいもので全体の報酬が定義される **bottleneck reward** という重要な設定に対して、初となる多項式時間アルゴリズムと統計的理論保証を持つアルゴリズムを設計した。これらの結果をまとめた論文は機械学習の国際会議 **NeurIPS2021** に採択された 2)。さらなる拡張として、報酬関数が再生核ヒルベルト空間における正定値カーネルにより定義される場合で、エージェントは複数いる場合のバンディットアルゴリズムを提案した 3)。上記の研究と並行して、限られた観測に基づく組合せバンディット問題における解説記事の執筆も行った 1)。

来年度は今年度の成果をまとめるとともに、データの背後にある確率分布が時刻によって変化する非定常環境下でも頑健なアルゴリズムの開発に取り組む予定である。

【代表的な原著論文情報】

- 1) Combinatorial pure exploration with full-bandit feedback and beyond: Solving combinatorial optimization under uncertainty with limited observation, Yuko Kuroki, Junya Honda, Masashi Sugiyama, arXiv preprint arXiv:2012.15584
- 2) Combinatorial Pure Exploration with Bottleneck Reward Function, Yihan Du, Yuko Kuroki, Wei Chen. In Proc. of Advances in Neural Information Processing Systems 2021, pages 23956–23967, 2021.
- 3) Collaborative Pure Exploration in Kernel Bandit, Yihan Du, Wei Chen, Yuko Kuroki, Longbo Huang. arXiv preprint arXiv:2110.15771