

# 研究報告書

## 「部分的フィードバックに基づくオンライン凸最適化」

研究期間：2020年4月～2022年3月  
研究者番号：50242  
研究者：伊藤 伸志

### 1. 研究のねらい

評価指標が不確定ななかで意思決定を繰り返し逐次的に戦略を改善するための数理モデルとして、オンライン最適化の枠組みが知られている。この枠組みでは、複数のラウンドにわたって解の出力と目的関数の観測を繰り返しながら、それらの総和の最小化を目指す。ここで目的関数の情報は解を出力した直後に与えられる。

オンライン最適化の技術は、オンライン予測問題やオンラインポートフォリオ選択問題など、様々な逐次的意思決定問題において有用であるが、目的関数の情報について部分的なフィードバックしか得られない実問題には直接適用できない。そこで本研究課題では、部分的なフィードバックのもとでのオンライン凸最適化問題を考察し、アルゴリズムを開発するとともに、問題の複雑性を解析することで、ある意味での最適なアルゴリズムの開発と最適性の証明を目指す。これによって、制限された観測のもとでのオンライン予測問題などを含む、多くの逐次的意思決定問題に対する統一的なアプローチを提供する。

研究提案者はこれまでの研究を通して、線形の問題設定に対しては少情報性・計算効率性・最適性のトレードオフの問題を解決し、非線形の問題設定に対しても特定のクラスや仮定を加えた状況を考えることで、いくつかの進展を得た。その一方で、より一般の非線形の問題設定や、異なる評価指標での解析など、多くの重要な未解決問題が残っている。例えば、一般のバンディット凸最適化に対しては、計算効率性を無視した場合でも、達成可能な性能の限界が分かっておらず、大きな改善の余地を残している。また、損失関数が一定の確率分布に振る舞うかそうでないか判断できない場合には、どちら状況に対しても有効なハイブリッド型アルゴリズムを考えることが実用上重要であるが、現状知られるハイブリッド型アルゴリズムは非常に単純な問題クラスに限られている。本研究ではこれらの課題の解決に取り組む。

### 2. 研究成果

#### (1) 概要

本研究では、オンライン最適化の理論の探求を目指しておもに「ハイブリッド型アルゴリズムの開発」、「敵対的外乱に対する頑健性の考察」、「問題設定の探索・拡張と解析」の3つの研究テーマに取り組んだ。オンライン最適化の研究領域においては、目的関数が定常的にふるまう確率的モデルや非定常的にふるまう敵対的モデルなどのさまざまな環境モデルが考察され、それぞれのモデルに適したアルゴリズムが提案されてきた。一方で、これらの結果を実問題に応用するとき、適切なアルゴリズムの選定がむずかしいという課題があった。「ハイブリッド型アルゴリズムの開発」においては、さまざまな環境モデルに自動的に適応し、複数の異なるアルゴリズムの強みを両立するアルゴリズムを提案した。また、「敵対的

外乱に対する頑健性の考察」においては、定常性と非定常性が混合されている状況を考え、その問題設定における最適な性能保証の特徴付けをあたえた。「問題設定の探索・拡張と解析」においては、フィードバック遅延がある状況や、通常と異なる性能指標の性質を考察し、新たなアルゴリズムの提案や性能の解析を与えた。

上記のそれぞれのテーマにおいて複数の論文を執筆し、そのうち3件の主著論文 [1, 2, 3] が機械学習の最難関国際会議 NeurIPS2020 の spotlight presentation 枠(全投稿論文中上位約 4%の評価)で採択され、発表を行った。また、1件の単著論文 [4] を機械学習理論の国際会議 COLT2021 で発表した。加えて、2 件の単著論文 [5, 6] が国際会議 NeurIPS2021 に採択された。

## (2) 詳細

### 研究テーマ A「ハイブリッド型アルゴリズムの開発」

テーマ A の研究活動では、さまざまな環境に自動的に適応し、複数の異なるアルゴリズムの強みを両立するアルゴリズムを考察した。オンライン最適化の研究領域においては、目的関数が定常的にふるまう確率的モデルや非定常的にふるまう敵対的モデルなどのさまざまな環境モデルが考察され、それぞれのモデルに適したアルゴリズムが提案されてきた(図 1)。とくに、敵対的モデルに対しては、ばらつきの小さいデータ列や滑らかに変化するデータ列など、特定の構造をもつ目的関数列に対してより有効的に動作するアルゴリズムが構成されている。このような、特定の構造に対する有効性の理論保証はデータ依存リグレット上界とよばれる。一方で、これらの結果を実問題に応用するとき、適切なアルゴリズムの選定がむずかしいという課題があった。たとえば、部分的フィードバックに基づくオンライン最適化の枠組みにおいて、基本的な問題設定である多腕バンディット問題に対しては、(1)確率的モデル用アルゴリズム(2)敵対的モデル用アルゴリズムにくわえて、複数の(3)データ依存リグレット上界保証ありアルゴリズムが構成されており、それぞれ異なる強みをもつが、実問題への適用のさい、どのアルゴリズムが最も有効的であるかを事前に判断するのは困難である。既存研究では、(1)と(2)のそれぞれの長所を両立する、つまり確率的モデルにおいても敵対的モデルにおいてもほぼ最適に動作するアルゴリズムが提案されていたが、(3)のほうがより有効である可能性は残っていた。本研究では、この結果を拡張し、(1)(2)(3)すべての長所を両立する初のアルゴリズムを提案し、その有効性を数理的に証明した(図 2)。この結果をまとめた論文 [4] は機械学習理論の国際会議 COLT2021 に採択された。また、同様の結果をより一般的かつ複雑な問題設定である組合せセミバンディット問題、線形バンディット問題に拡張し、それらの結果をまとめた論文 [2, 6] はそれぞれ機械学習の国際会議 NeurIPS2020, NeurIPS2021 に採択された。

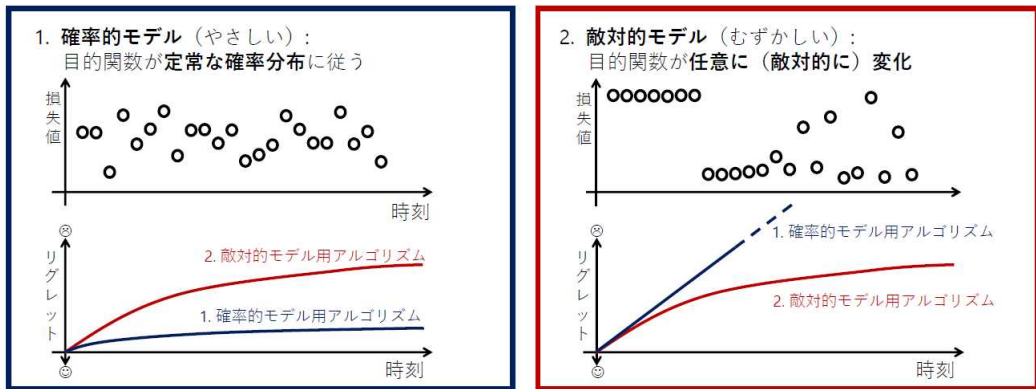


図 1. 二つの環境モデルの例とアルゴリズムの性能の傾向.

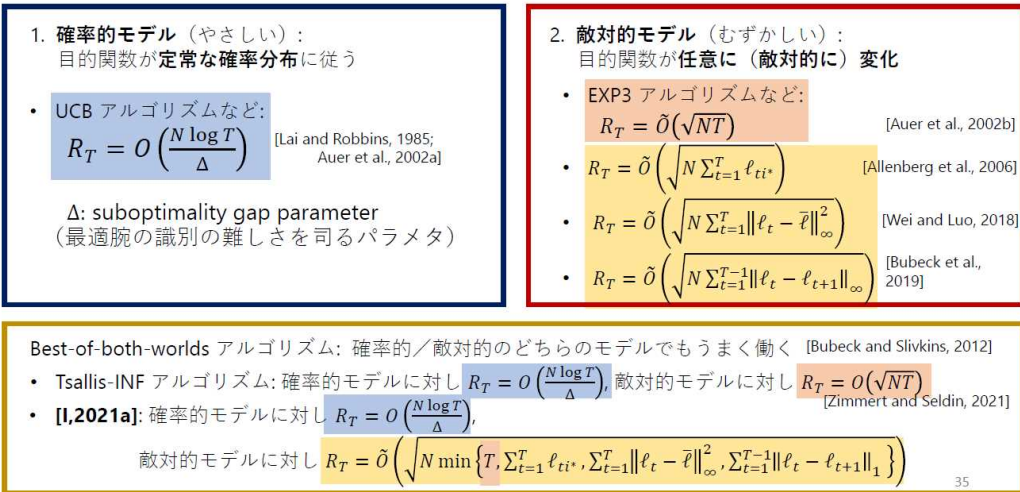


図 2. 確率的モデル用アルゴリズム (薄青), 敵対的モデル用アルゴリズム (薄赤) データ依存リグレット上界 (薄期黄) と提案アルゴリズム ([I, 2021a]).

### 研究テーマ B「敵対的外乱に対する頑健性の考察」

テーマ B の研究活動では、定常性と非定常性が混合されている状況を考え、その問題設定における最適な性能保証の特徴付けをあたえた。テーマ A のハイブリッドアルゴリズムに関する取り組みにおいては、定常的な確率的モデルと非定常的な敵対的モデルのそれぞれに適応する方法を考察したが、現実には、完全に定常的な状況や完全に非定常的な状況は珍しく、むしろこのふたつのモデルの中間的な状況が多く見られる。本研究では、そのような中間的な状況の具体例として、定常的なモデルから生成された目的関数列に非定常的な外乱が加わるモデルを考察した。一般に、このモデルに対してのアルゴリズムの性能を解析することで、そのアルゴリズムの訓練データの改ざんに対する頑健性を評価できる。本研究では、エキスパート問題、多腕バンディット問題、組合せセミバンディット問題などさまざまなオンライン最適化問題において、敵対的改変付き確率的モデルを考察し、あるクラスのア

ゴリズムが最適な性能保証をもつことを証明した。主要な結果をまとめた論文 [5] が国際会議 NeurIPS2021 に採択された。論文 [6] も敵対的改変付き確率的モデルに対する性能解析の結果を含む。

#### 研究テーマ C「問題設定の探索・拡張と解析」

テーマ C の研究活動では、フィードバック遅延がある状況や、アルゴリズムの性能指標の性質を考察し、新たなアルゴリズムの提案や性能の解析を与えた。オンライン最適化の標準的な枠組みでは、決定した行動とその時点の目的関数についてのフィードバックが行動の直後に得られる状況が想定されるが、実問題においてはそのような即時的なフィードバックが得られるとは限らない。本研究では、そのような状況に対応できるよう遅延フィードバックの問題設定を考察した。とくに、遅延フィードバックの敵対的線形バンディット問題を考察し、この問題設定に対するアルゴリズムを提案、その最適性を数理的に証明した。この結果をまとめた論文 [3] は国際会議 NeurIPS2020 に採択された。また、swap regret と呼ばれるアルゴリズムの性能指標の性質を解析した。この指標は、複数の意思決定主体がオンライン最適化の枠組みに基づいて意思決定を繰り返したときの振る舞いに関連しており、アルゴリズム的ゲーム理論の領域においても研究されてきたが、この指標を最小化する問題の達成可能限界は明らかにされていなかった。本研究では、この問題の複雑性解析を改善し、達成可能限界を初めて明らかにした。

## 2. 今後の展開

これまでの研究を通して、凸性や線形性の構造をもつ一部のオンライン最適化問題に対してハイブリッド型オンライン最適化アルゴリズムや、外乱に対して頑健なアルゴリズムを構成することができた。その一方で、一般の非線形の問題設定や、各時刻での意思決定がそれ以降の意思決定の制約に影響する階層的な意思決定問題などで、同様のアルゴリズムを構成可能かという問題ははまだ未解決である。また、最適行動が時刻変化するような任意の環境変化への追従を実現するには、標準的な解析技法を超えた新たな技法が必要と考えられているが、その考察対象は現時点では非常に単純な問題クラスに限られている。今後の研究ではこれらの課題の解決に取り組む。並行して、オンライン最適化ソフトウェアの開発を推進する。現状では、理論に留まっているオンライン最適化の研究成果が多くあり、実世界での応用例や利用しやすいソフトウェアはあまり知られていない。今後の研究活動では、この課題の解決も目指す。

## 4. 自己評価

研究の達成状況を振り返ると、ハイブリッド型アルゴリズムの開発については当初の計画を超えた進展があり、満足のいく結果が得られたと考える。オンライン最適化の理論研究コミュニティにおいては、特定の環境下における最適なアルゴリズムが何かという問題意識が重視され、環境に特化したアルゴリズムが主な考察対象になっていたが、その産業応用や社会実装を見据えると、環境の識別・有効的なアルゴリズムの選定が重要な課題である。この課題に注目した研究は世界的にも珍しいものであるが、本研究の成果はこの課題を解決し、応用へのハードルを下げると考えている。また、敵対的外乱への頑健性の研究成果は、最適な頑健性をもつアルゴリズムとその数理的構造を明らかにした点から学術的なインパクトがあるだけでなく、機械

学習システムの信頼性への解析を与えた点から、産業応用の促進にも一定の波及効果があると見込んでいる。

研究の進め方については、研究費を活用してほかの研究者との交流の機会をもち、新たな共同研究を開始できた点はよかったと思う。しかし、理論研究に集中して論文等の成果はだせた一方で、ソフトウェアの開発など実応用に向けた活動の範囲が限定的であった点は反省点である。今後、自身の専門性から離れた人との協働などを通して、視野を広げるとともに実応用の機会を探索し、自身の研究成果の波及効果をより大きくしていきたいと考える。

## 5. 主な研究成果リスト

### (1) 論文(原著論文)発表

1. Shinji Ito. A Tight Lower Bound and Efficient Reduction for Swap Regret. *Advances in Neural Information Processing Systems (NeurIPS)*. 2020, 33, 18550–18559, selected for spotlight presentation.
2. Shinji Ito, Shuichi Hirahara, Tasuku Soma, Yuichi Yoshida. Tight First- and Second-Order Regret Bounds for Adversarial Linear Bandits. *Advances in Neural Information Processing Systems (NeurIPS)*. 2020, 33, 2028–2038, selected for spotlight presentation.
3. Shinji Ito, Daisuke Hatano, Hanna Sumita, Kei Takemura, Takuro Fukunaga, Naonori Kakimura, Ken-ichi Kawarabayashi. Delay and Cooperation in Nonstochastic Linear Bandits. *Advances in Neural Information Processing Systems (NeurIPS)*. 2020, 33, 4872–4883, selected for spotlight presentation.
4. Shinji Ito. Parameter-Free Multi-Armed Bandit Algorithms with Hybrid Data-Dependent Regret Bounds. *Proceedings of Thirty Fourth Conference on Learning Theory (COLT2021)*, 2021, 2552–2583.
5. Shinji Ito. On Optimal Robustness to Adversarial Corruption in Online Decision Problems. *Advances in Neural Information Processing Systems (NeurIPS)*. 2021, 34, in press.
6. Shinji Ito. Hybrid Regret Bounds for Combinatorial Semi-Bandits and Adversarial Linear Bandits. *Advances in Neural Information Processing Systems (NeurIPS)*. 2021, 34, in press.

### (2) 特許出願

研究期間累積件数: 5 件