

革新的コンピューティング技術の開拓
2019 年度採択研究者

2021 年度 年次報告書

坂田綾香

統計数理研究所
准教授

求解軌道のマクロ表現によるアルゴリズム制御理論の創出

§ 1. 研究成果の概要

本年度は、学習において用いるデータの効率的な取得法に関する研究を行なった。その背景には、2020年度から行なってきたグループテストの研究がある。この研究では、推定される変数が離散変数であることを生かしてベイズ予測分布を評価し、そのエントロピーを最大にするデータを取得することで、学習効率が上がることを示した。本年度は、この方法を連続変数に拡張する研究を行なった。

この目的に合った問題として、**query by committee** と呼ぶ学習を扱った。この方法では、すでに取得済みデータのもとで構成した事後分布に従い、**committee member** と呼ぶ確率変数の実現値を得て、ベイズ予測分布を経験分布により近似する方法である。従って、**committee member** 数が無限大となる極限では、ベイズ予測分布に基づくデータ取得と見なすことができる。本年度は、この問題に対して確率伝搬法を適用してアルゴリズムを構成した。

一般に、確率伝搬法はランダムデータに対して効率的であることが知られている。一方で、データ間になんらかの相関が含まれる場合は、収束しにくいことが経験的に知られている。**Query by committee** は、ベイズ予測分布に従ってデータを取得するため、データに相関を生じさせる。したがって、アルゴリズムの収束性の悪さが予想された。

実際に **query by committee** に対してアルゴリズムの性能を観察したところ、予想通り収束の悪さが確認された。具体的には、データサンプル数に対して、収束時間が指数関数的に増加することが示された。しかし、本年度の研究結果から、アルゴリズムの初期条件の構成法を工夫することで、指数関数的な収束時間の増加が抑えられることがわかった。この発見により、連続変数に対して能動的にデータを取得する学習法に対するアルゴリズムの実現可能性が示された。

アルゴリズムの初期条件依存性のメカニズムについては未知の点が多いため、2022年度は、2021年度の発見を理論的に裏付けるべく、マクロ変数を用いたアルゴリズム挙動の研究を行なう。

【代表的な原著論文情報】

- 1) Ayaka Sakata and Tomoyuki Obuchi
"Perfect reconstruction of sparse signals with piecewise continuous nonconvex penalties and nonconvexity control"
Journal of Statistical Mechanics: Theory and Experiment, vol. 2021(9), 093401 (2021).
- 2) A. Sakata and Y. Kabashima
"Decision Theoretic Cutoff and ROC Analysis for Bayesian Optimal Group Testing"
arXiv:2110.10877 (2021). (投稿中)