

数理・情報のフロンティア
2021 年度採択研究者

2021 年度 年次報告書

藤澤 将広

東京大学 大学院新領域創成科学研究科
大学院生(博士課程)

ロバスト性と汎化性能を両立する機械学習法の確立

§ 1. 研究成果の概要

初年度は、「人間が機械学習モデル側に意図せず与えてしまう誤情報」を汚染された情報の一種として注目し、これを適応的に修正しつつ高い汎化性能を獲得する学習アルゴリズムの開発を試みた。PAC-Bayes 汎化学習アルゴリズムでは「事前分布」と呼ばれる事前情報を与える必要があるが、それが不適切に設定された場合、汎化性能が著しく損なわれる。この問題は、不適切な事前分布が、汎化誤差上界に現れるパラメータ複雑度の値を増大させることで生じる。これを回避するために、小さいパラメータ複雑度を達成する事前分布をデータから推定する方法が多く提案されているが、期待経験誤差とのトレードオフ関係から、汎化性能は必ずしも保証されるわけではない。

本年度の研究では、「パラメータ複雑度を大きくすることも、小さくしすぎることも良くない」ことに注目し、パラメータ複雑度を「あまり変化させない」方向に事前分布を適応的に修正する方法を探究した。まず、ロバスト統計で用いられる影響関数を応用し、事前分布の変化がパラメータ複雑度を与える影響を定式化した。そして、影響関数を小さくする確率分布、すなわちパラメータ複雑度があまり変化しないような事前分布のパラメータが、指数型分布族の性質の元では解析的に得られることを明らかにし、それを用いて適応的に事前分布修正を行うアルゴリズムを提案した。さらに、提案法によって修正された事前分布は、簡単なベイズ線形回帰モデルにおいて、期待経験誤差を小さくすることを理論的に証明した。PAC-Bayes における汎化誤差上界が(定数項を除けば)「期待経験誤差+パラメータ複雑度」で表せることから、パラメータ複雑度をあまり変化させずに期待経験誤差を小さくする提案法は、より良い汎化性能の獲得を期待させる。実際、ベイズ深層学習モデルを用いた実験を通して、提案アルゴリズムがより良い汎化性能を獲得することが経験的に示された。