

数理・情報のフロンティア
2019 年度採択研究代表者

2020 年度 年次報告書

二見 太

NTT コミュニケーション科学基礎研究所
研究員

相互作用系を活用したサンプリング手法の開発

§ 1. 研究成果の概要

本年度は1)昨年度の研究からの課題であった、Langevin dynamics と相互作用を使ったサンプリングにおける、相互作用の構成の仕方、およびハミルトンモンテカルロ系への拡張、2)サンプリングや最適化によって得られた複数のモデルから統合して予測を行う際に、モデルのばらつきが汎化能力にあたえる影響の理論解析、の2つを中心に研究を進めた。

設定として複数のモデルを並列に学習させる場合に、互いに独立に学習をさせるのではなく、図1の右側のように互いに相互作用をさせることでより多様なモデルを作り出すことを目指すというものである。昨年度の研究ではこの設定の下、サンプリングの場合に、相互作用を入れることでサンプリングの収束の速さを向上させることに関し研究を行った。本年度はそれを受けて、どのような相互作用が理論的に望ましいのかということ、目的のモデルのポテンシャル関数に注目することで、収束の加速が発生するための必要十分条件を導き、その条件を満たす相互作用を構成するアルゴリズムの開発を行った。またより実用面で優れた、ハミルトンモンテカルロの手法と組み合わせた並列アルゴリズムの開発を行った。

次に我々は、例えばサンプリングなどで複数のモデルが得られたとしたときに、それらの結果を統合して予測をする問題において、モデルのばらつきが学習の汎化性能にどのような影響を持つのかに関して研究を行った。特に我々はモデルのばらつきに関する情報を自然に取り込んだ、2次の Jensen の不等式を新たに導出した。この不等式を基に我々はばらつきの情報を含んだ形で新たな PAC ベイズの汎化誤差を導出した。これにより我々は、汎化誤差を効果的に縮めるためには、各モデルが訓練データによくフィッティングされることに加え、モデル間のばらつきを大きくする、というトレードオフを制御する必要があることを発見した。