

新しい社会システムデザインに向けた情報基盤技術の創出
2018 年度採択研究者

2020 年度 年次報告書

今泉允聡

東京大学 総合文化研究科
准教授

深層学習の高速化にむけた適応ネットワークの数学的発見と学習法開発

§ 1. 研究成果の概要

本研究の目的は、深層学習が精度を発揮するのに必要な「適応ネットワーク」という概念を数学的に発見し、それを用いて深層学習の高速化を行うことである。今年度はその「適応ネットワーク」の理論的性質について以下のような成果を得た。

1. データの固有次元に関する適応性

実データは見かけが高次元でも、その本質的な次元(固有次元)は低次元であることが知られている。本研究は、深層ニューラルネットワークがデータの低次元性に適応し、その汎化誤差が速く収束することを証明した。また、他手法に対して、深層学習がフラクタルなどの複雑な形状の固有次元に適応することを示した。

2. ニューラルネットワークの不変性に関する適応性

点群データ解析などに用いられるニューラルネットワークは、有限群に対する不変性を持つことが知られている。本研究では、このようなニューラルネットワークのもとでの汎化誤差が、不変性の度合い(例えば点の数)に対して指数的に減少することを示した。これにより、モデルに対して不変性の構造を導入することが高い精度を達成することを理論的に示した。

3. 深層モデルの大パラメータ極限の漸近リスク解析

いくつかの統計モデルでパラメータの数を無限大にすると、汎化誤差の分散が減少することが知られている(例: 二重降下減少)。しかし、この現象が理論的に示されるのはランダム特徴量モデルなどの線形モデルに限られていた。本研究では、深層モデルを含む広いクラスのニューラルネットワークが、一定の正則条件のもとでこの現象に従うことを示した。具体的には、並列化ニューラルネットワークや ResNet などの特有の構造が、大パラメータ極限での誤差の減少を保証することを明らかにした。

【代表的な原著論文情報】(最大 5 件)

- 1) R.Nakada, M.Imaizumi (2020), "Adaptive Approximation and Generalization of Deep Neural Network with Intrinsic Dimensionality", Journal of Machine Learning Research 21(174).
- 2) A.Sannai, M.Imaizumi, M.Kawano, (2021), "Improved Generalization Bound of Group Invariant / Equivariant Deep Networks via Quotient Feature Space", PMLR: Uncertainty on Artificial Intelligence 2021.
- 3) R. Nakada, M.Imaizumi, "Asymptotic Risk of Overparameterized Likelihood Models: Double Descent Theory for Deep Neural Networks". arXiv preprint.