

2023 年度年次報告書
次世代 AI を築く数理・情報科学の革新
2023 年度採択研究代表者

大古 一聡

東京大学 大学院情報理工学系研究科
大学院生

構造を持つデータの学習可能性解析による巨大深層モデルの理解

研究成果の概要

本研究課題では高次元データの持つ構造と学習の効率性の関係を調べる。現実のデータに見られる性質を定式化して具体的なケーススタディに落とし込み、それらにおいて学習ダイナミクスの解析を行う。今期取り組んだテーマは、(1)非等方入力スパース parity, (2)ridge combination の2つの設定である。

(1) 現実の高次元データはある種の低次元性を持つことが知られている。その1つには①目的変数が入力の一部の方向のみに依存するスパース性があり、もう1つには②入力が各方向に一様ではなくしばしば目的関数の依存する方向に揃っているという非等方性がある。非等方入力スパース parity は classification の文脈において①②を両立する重要なケーススタディであり、本研究では①②の性質がニューラルネットワークとカーネル法の学習の効率性に及ぼす影響を調べた。両手法とも非等方性から恩恵を受けスパース parity をより効率的に学習できるが、その中でも、特徴学習能力を持つニューラルネットワークに比べ、基底が固定されたカーネル法は学習に必要なサンプル数がより多くなることが分かった。本研究成果は特徴学習能力の有無がデータの持つ構造への適応力の差に繋がることを明らかにした重要な結果と言える。

(2) 巨大深層モデルの事前学習においては、多様なソースから集めた大規模なデータセットが用いられ、モデルは多数の特徴を同時に学習している。Ridge combination は1方向のみに依存する関数の和を学習する問題であり、そのような状況の自然な定式化である。ニューラルネットワークによる ridge combination の学習ダイナミクスの解析を行い、同時にこの問題が持つ本質的困難性を statistical query というフレームワークを用いて議論した。

【代表的な原著論文情報】

- 1) Atsushi Nitanda, Kazusato Oko, Taiji Suzuki, Denny Wu (alphabetical order). Improved statistical and computational complexity of the mean-field Langevin dynamics under structured data. The Twelfth International Conference on Learning Representations (ICLR 2024).
- 2) Juno Kim, Kakei Yamamoto, Kazusato Oko, Zhuoran Yang, Taiji Suzuki. Symmetric Mean-field Langevin Dynamics for Distributional Minimax Problems. The Twelfth International Conference on Learning Representations (ICLR 2024).