

数学と情報科学で解き明かす多様な対象の数理構造と活用
2019 年度採択研究者

2020 年度 年次報告書

二反田 篤史

東京大学大学院情報理工学系研究科
助教

深層学習の潜在的な正則構造の理解に基づく学習法の安定化と高速化

§ 1. 研究成果の概要

二層ニューラルネットワークの学習ダイナミクスに関して以下の成果を得た。

- (1) ニューラルネットワークの最適化問題はそのパラメータについて非凸最適化であり非常に困難な問題であるが、ニューラルネットワークそのものを変数として、その最適化ダイナミクスを適切に捉えることで目的関数の凸性を活かすことができる。そのための理論的枠組みの一つにニューラルタンジェントカーネル (NTK) がある。NTK 理論ではニューラルネットワークの学習ダイナミクスを NTK に付随する再生核ヒルベルト空間で記述する。それにより高次元ニューラルネットワークの最適化をカーネル法の理論を援用して議論できる。実際、NTK を用いた既存研究においてニューラルネットワークの最適化法に対する大域収束性および汎化性が示されたが、その理論には改善の余地があった。本研究では二乗損失を用いた回帰問題において NTK の固有値減衰を考慮にいたれた精緻な汎化誤差解析を与えた。それにより平均化確率的勾配降下法が NTK 理論を適用できる条件下でミニマックス最適な汎化誤差を達成することを示した。
- (2) ラベルの低ノイズ条件下でランダム特徴を用いた SGD によって期待識別誤差が線形収束することを示した。これは無限次元再生核ヒルベルト空間を仮説空間とする場合の既存研究のランダム特徴近似を用いた場合への拡張である。ランダム特徴は計算効率の観点で有用性が知られていたが、本設定ではその有用性がより際立つことを示した。またランダム特徴近似を用いたカーネル法は実質的には二層ニューラルネットワークの出力層パラメータの学習とみなせるため、成果(1)と(2)はともに二層ニューラルネットワークの研究に貢献するものである。

【代表的な原著論文情報】

- 1) Atsushi Nitanda and Taiji Suzuki. Optimal Rates for Averaged Stochastic Gradient Descent under Neural Tangent Kernel Regime. The 9th International Conference on Learning Representations (ICLR), 2021.
- 2) Shun-ichi Amari, Jimmy Ba, Roger Grosse, Xuechen Li, Atsushi Nitanda, Taiji Suzuki, Denny Wu, and Ji Xu. When Does Preconditioning Help or Hurt Generalization?. The 9th International Conference on Learning Representations (ICLR), 2021.
- 3) Shingo Yashima, Atsushi Nitanda, and Taiji Suzuki. Exponential Convergence Rates of Classification Errors on Learning with SGD and Random Features. The 24th International Conference on Artificial Intelligence and Statistics (AISTATS), *Proceedings of Machine Learning Research*, 130:1954–1962, 2021.
- 4) Atsushi Nitanda and Taiji Suzuki. Functional Gradient Boosting for Learning Residual-like Networks with Statistical Guarantees. The 23rd International Conference on Artificial Intelligence and Statistics (AISTATS), *Proceedings of Machine Learning Research*, 108:2981–2991, 2020.
- 5) Atsushi Nitanda, Tomoya Murata, and Taiji Suzuki. Sharp Characterization of Optimal Minibatch Size for Stochastic Finite Sum Convex Optimization. 2019 IEEE International

Conference on Data Mining (ICDM2019), pp. 488–497. 2019