

2023 年度年次報告書

数学・数理科学と情報科学の連携・融合による情報活用基盤の創出と社会課題解決に向けた展開

2020 年度採択研究代表者

福水 健次

情報・システム研究機構 統計数理研究所
教授

数理知能表現による深層構造学習モデルの革新

主たる共同研究者:

鈴木 大慈 (東京大学 大学院情報理工学系研究科 准教授)

原田 達也 (東京大学 大学院情報理工学系研究科 教授)

研究成果の概要

数理深層モデリング G では、まず、群作用によって遷移するペア／系列データが多数与えられた状況での表現学習に関して研究し、データからの学習に基づく新しい非線形フーリエ変換の方法論を提案した¹⁾。また、カーネル法と深層学習を融合した、条件付カーネル平均埋め込みのスケラブルな計算法を提案した²⁾。さらに、常微分方程式による生成モデルであるフローマッチングに関して、条件付確率からの有効性の高い生成方法を新たに開発するとともに、フローマッチング学習が確率分布推定に関する最適収束レートを持つことを示した。

数理表現・最適化 G では、まず深層生成モデルの統計的最適性の評価を行った。拡散モデルは深層学習によってスコア関数を高い精度で推定し、最終的な生成モデルが真の分布をミニマックス最適な精度で推定できることを示し³⁾、Transformer は重要なトークンを入力系列に応じて選択することで次元の呪いを回避できることを示した⁴⁾。また、二層ニューラルネットワークの勾配法による学習ダイナミクスを解析し、データに内在する実質的な低次元性を捉えることで k-parity 問題や single index model で情報理論的下限の推定精度を多項式時間で達成することを示した^{5,6)}。

実践的学習 G では、教師無しドメイン適応問題に取り組んだ。従来手法では、ソース領域における分類誤差と、ソース領域とターゲット領域の周辺分布間の距離を同時に最小化することで実現されているが、クラス情報を加味した結合誤差はその難解さゆえに無視されてきた。本研究では、特に領域間のギャップが大きい場合に、結合誤差の考慮が不可欠であることを示し、この問題に対処するために、結合誤差の上限に関する新しい目的関数を提案し、この目的関数を最小化させることで、画像分類精度等の性能向上が得られることを示した⁷⁾。

【代表的な原著論文情報】

- 1) Koyama, Fukumizu, Hayashi, Miyato: Neural Fourier Transform: A General Approach to Equivariant Representation Learning. 12th Inter. Conf. Learning Representations (ICLR 2024)
- 2) Shimizu, Fukumizu, Sejdinovic: Neural-Kernel Conditional Mean Embeddings. 41st Intern. Conf. Machine Learning (ICML2024)
- 3) Oko, Akiyama, Suzuki: Diffusion Models are Minimax Optimal Distribution Estimators. Proc. 40th Intern. Conf. Machine Learning (ICML2023), PMLR 202:26517--26582, 2023.
- 4) Takakura, Suzuki: Approximation and Estimation Ability of Transformers for Sequence-to-Sequence Functions with Infinite Dimensional Input. Proc. 40th Intern. Conf. Machine Learning (ICML2023), PMLR 202:33416--33447, 2023.
- 5) Suzuki, Wu, Oko, Nitanda: Feature learning via mean-field Langevin dynamics: classifying sparse parities and beyond. 37th Conf. Neural Information Processing Systems (NeurIPS2023).
- 6) Ba, Erdogdu, Suzuki, Wang, Wu: Learning in the Presence of Low-dimensional Structure: A Spiked Random Matrix Perspective. 37th Conf. Neural Information Processing Systems (NeurIPS2023).
- 7) Zhang, Harada. Unsupervised Domain Adaptation via Minimized Joint Error. Trans. Machine Learning Research (TMLR), 2023. <https://openreview.net/forum?id=kiPsMct7vL>