

2024 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 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、主に Flow Matching に関する研究と、シミュレーションに基づく微分方程式のパラメータ推定に関する研究を行った。連続条件値を任意に転移できる最適輸送ベースの Flow Matching を用いた条件転移法を提案し、分子改変の問題で良好な結果を得た。また、Flow Matching の汎化誤差が訓練データ数に対して minimax 最適性を持つことを理論的に示した。さらに、シミュレーションに基づく微分方程式のパラメータ推定に対し、Markov 性の仮定の下での拡散モデルのスコア分解を用いた効率的 Bayes 推論法を提案した。

数理表現・最適化 G では、主にニューラルネットワークの高次元特徴学習に関する最適化理論と文脈内学習の理論を展開した。前者に関しては、平均場ランジュバン動力学による学習で情報理論的下限のサンプル複雑度を達成すること、およびバッチ再利用によってサンプル/計算複雑度ともに情報理論的下限を達成することを示した。後者に関しては、Transformer による文脈内学習のミニマックス最適性、および勾配法による事前学習によって特徴学習が進みテスト時のサンプル複雑度が大きく改善することを証明した。

実践的学習 G では、オープンセットドメイン適応 (ODA) 問題に取り組んだ。ODA は、ターゲットデータにソースデータに存在しない未知クラスが含まれるという、より現実的なオープンセットドメインシフト問題において、学習アルゴリズムの汎化性能向上を目的とする。既存の多くのアルゴリズムは、ヒューリスティックに未知クラスを分離する手法に依存し、汎化誤差を厳密に制限することはできない。そこで、本研究では、Positive-Unlabeled 学習理論とドメイン適応における結合誤差に基づき、ターゲットタスク全体のリスクを厳密に制限するエンドツーエンドのアルゴリズムを提案した。

【代表的な原著論文情報】

- 1) Fukumizu, K., Suzuki, T., Isobe, N., Oko, K., Koyama, M. Flow matching achieves almost minimax optimal convergence. *13th Intern. Conf. on Learning Representations (ICLR) 2025*.
- 2) Gloeckler, M., Toyota, S., Fukumizu, K. and Macke, J.H. Compositional simulation-based inference for time series. *13th Intern. Conf. on Learning Representations (ICLR) 2025*.
- 3) Kim, J., Nakamaki, T., Suzuki, T. Transformers are Minimax Optimal Nonparametric In-Context Learners. *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*. pp. 106667--106713, 2024. (Won Best Paper Award in ICML 2024 Workshop on Theoretical Foundations of Foundation Models)
- 4) Lee, J.D., Oko, K., Suzuki, T., Wu, D. Neural network learns low-dimensional polynomials with SGD near the information-theoretic limit. *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*. pp. 58716--58756, 2024.
- 5) Zhang, D., Westfechtel, T., Harada, T. Open-set Domain Adaptation via Joint Error based Multi-class Positive and Unlabeled Learning. *European Conference on Computer Vision (ECCV)*, pp.105-120, 2024.