

東北大学大学院情報科学研究科
純粋・応用数学研究センター

情報数理談話会のお知らせ

日 時： 2025 年 4 月 30 日 (火) 16:30 より 17:30 まで

場 所： 東北大学大学院情報科学研究科棟 2 階大講義室

講演者： Guan Xin 氏 (東北大学大学院情報科学研究科)

題 目： Sparse kernel k-means clustering for high-dimensional data

[概 要] 従来のスパース k-means 法は、特徴量の選択とクラスタリングを同時に実行するが、線形なクラスタ境界を持つ単純なクラスタ構造しか考慮できないという欠点がある。一方、カーネル k-means 法は非線形なクラスタ構造を持つデータに対して有効な方法である。しかし、この方法はノイズ変数を含む高次元データに対しては上手く機能しない。本研究では非線形クラスタリングと特徴選択を同時に実行できるスパースカーネル k-means 法を提案した。提案手法は各特徴量が選択されるかどうかを表すインディケーターを導入し、このインディケーターに関する制約を考えた。それで、特徴選択により、ノイズ変数の影響を受けることなく、高次元データに対してもカーネル k-means 法の長所を活かすことができる。本研究では理論分析を通じて、適切な条件の下で提案手法のクラスタリング一致性と特徴選択一致性を示した。また、数値実験および実データ解析を通じて、提案手法が既存の高次元クラスタリング法の性能を大幅に改善することを示した。

ホームページ：<https://www.math.is.tohoku.ac.jp/research/colloquium.html>