

大自由度モデルのための新しい高次元統計と深層ニューラルネットワークの動力学理論

概要：

現代的データ科学のひとつである深層学習・人工知能技術が大きく発展し、これらの効率的に制御・発展させるための数理的な理解が求められている。本講演では、このトピックに関連する研究を紹介する。(I) ひとつ目は、過剰パラメータモデルのための高次元統計である。従来の高次元統計は余剰次元を削減する方法論を発展させてきたが、近年の大自由度モデルは明示的な余剰次元を持たないため、近年は異なる方法論が発展している。本講演では、この方法論をより実践的な統計モデルに応用した研究成果を紹介する。(II) ふたつ目は、過剰パラメータなモデルの典型例である深層ニューラルネットワークの学習動力学の理論である。深層モデルをアルゴリズムで更新する過程は非常に複雑で解析が困難だが、パラメータ分布のマクロ的な構造を捉えることでその動力学を構成できる。本講演ではこれに関連する複数の試みを紹介する。(III) 時間に余裕があれば、ChatGPT などの基盤モデルという人工知能技術を説明する、トランスフォーマーおよび関連する統計理論の解析である。本研究では、基盤モデルに用いられるトランスフォーマーの観測分布を効率的に扱えるという性質により、一定条件下で効率的な学習を実現できることを議論する。