

2023 年度
創発的研究支援事業 年次報告書

研究担当者	唐木田 亮
研究機関名	産業技術総合研究所
所属部署名	人工知能研究センター
役職名	主任研究員
研究課題名	学習力学を数理基盤とした革新的ニューラルネットワークの開拓
研究実施期間	2023 年 4 月 1 日～2024 年 3 月 31 日

研究成果の概要

本年度の主要な活動成果は以下のとおりである。

・深層学習は一般に無数の解を持ち、モデルの性能は学習アルゴリズムの力学によって選択される解に依存することが明らかになりつつある。本課題ではその数理的解析と理解を大目標のひとつとしている。これに関し、1 次勾配という極めて基本的な量の正則化がのぞましい解への陰的バイアスを持ち、性能向上につながることを発見・発表した [Karakida, Takase, Hayase & Osawa, ICML 2023]。この勾配正則化は深層学習の文脈でのぞましいとされている平坦な解の探索とつながりがあることまで明らかにし、今後の学習力学解析の土台となる成果のひとつとなった。

・大規模化した深層学習では、数理的に望ましい最適化手法を厳密に計算することが、計算量の観点から困難であることが多い。そこでは層ごとのブロック近似に代表される発見的近似が行われる。これに関して、創発リサーチアシスタントとの共同研究を実施し採択された [Ishikawa&Karakida, ICLR 2024]。無限幅極限で特徴学習を促進するパラメータ表示を、2 次最適化およびその層ごとの近似を包含する形で定式化し、学習力学の観点から数理的に解析した。理論と実験から提案法の有用性を明らかにした。

・現代的な深層学習アーキテクチャは、入力の構成要素であるトークン次元と特徴次元からなる行列が信号伝播する構造を持つ。中でも MLP-Mixer は両方の次元が単純な MLP で処理される形になっており、現代のアーキテクチャが高い性能を発揮するメカニズムを探る観点から基本のモデルといえよう。我々は MLP-Mixer が、結合にスパース性を持つ全結合 MLP に変形できることに注目した。MLP-Mixer はスパースかつ幅が大きい MLP として振る舞うことを、いくつかの理論解析および実験から検証し採択された [Hayase&Karakida, ICML 2024]。また、性能の観点から、数理的にのぞましいスパース度合いに定量的な知見を与え、アーキテクチャ設計への示唆まで与えている。スパース性は生体神経回路の研究やニューラルネットに関する様々な文脈でその有用性が議論されてきた。MLP-Mixer はそのスパース性を、現代的な行列積ベースの計算の枠組みで陰的に実現していると解釈でき、今後のアーキテクチャ開拓の基盤となる役割が期待できる。