

画像データを用いた深層学習における 長時間ダイナミックス

サイバーメディアセンター 吉野 元

目的 現実の画像データを学習する深層ニューラルネットワーク(DNN)について統計力学的な解析を行い、(1) SGDによってある種の熱平衡化が起こるのか (2) 理論的に期待されている学習の空間的不均一性は実在するのかを明らかにする。

内容 層間を全結合でつないだ最も単純な長方形DNNに公開画像データセットMNISTを学習させるシミュレーションを行った。ダイナミックスは確率勾配法(SGD)にAdam、バッチ正則化を組み合わせたものである。具体的なタスクとしては、標準的な分類問題とノイズ除去を行わせた。同じデータで学習する2つのマシン(レプリカ)の「重なり」を定義し、これを観測した。

結果 数千エポックの長時間ダイナミックスを観測し、(1) SGDによって初期条件によらない定常状態が実現すること(2) 理論的に予言された学習の空間的不均一性(H. Yoshino (2020), (2023))を強く示唆する結果を得た。本研究は大阪大学大学院理学研究科物理学専攻 織戸悠輔氏(M2)との共同研究である。

利用した計算機

OCTOPUS GPUノード群

ノード時間 400時間

使用メモリ 32GB

