

深層ニューラルネットワークの適応的最適化手法 における汎化性・頑健性の分析

An Investigation of Generalization and Robustness in Adaptive Optimization Methods for Deep Neural Networks

学籍番号：201921649

氏名：本川 哲哉

Motokawa Tetsuya

深層ニューラルネットワーク (Deep Neural Network, DNN) の学習において, Adam をはじめとする適応的最適化手法は確率的勾配降下法 (Stochastic Gradient Descent, SGD) よりも高速に収束することで知られており, 近年様々な深層学習タスクでよく利用される. その反面, 特に画像分類タスクにおいて, SGD に比べて最終的に得られる DNN モデルの汎化性が確保されにくいという報告も見られる. しかしながらその原因解明はまだ進んでいない.

近年ヘッセ行列の固有値の分析をすることで, DNN モデルにおける学習ダイナミクスを解釈する研究は盛んに行われているが, この分析を用いて適応的最適化手法の特性を分析した研究は行われていない. そこで本研究では損失関数におけるヘッセ行列の固有値を分析することで, 異なる最適化手法によって学習された DNN モデルの重みパラメータの汎化性, 頑健性を分析した.

具体的には, いくつかの DNN モデルと画像分類タスクを用いて大きく 3 つの分析を行った.

まず学習後の DNN モデルの重みパラメータ空間において, 損失関数のヘッセ行列の固有値分布を解析した. 全ての重みパラメータを用いた損失関数形状の比較には, 比較的小さな重みパラメータ数の DNN モデルである C3, LeNet を使用した. プラクティカルな DNN モデルとしてよく利用される ResNet-34, DenseNet-121 に対しては最終層の重みパラメータのみを用いた. その結果, 適応的最適化手法の方が SGD に比べて局所的に急峻な形状に収束することを確認した.

汎化性を測る指標の 1 つとして, ヘッセ行列と勾配共分散行列を用いて計算される竹内情報量基準 (Takeuchi Information Criterion, TIC) を用いて, SGD と適応的最適化手法における汎化性の比較を行った. 実験の結果, 今回用いた 2 つのデータセット (CIFAR-10, Fashion-MNIST), モデルについては SGD に比べて適応的最適化手法によって得られる重みパラメータに関するヘッセ行列の固有値の方が絶対値の大きな値に分布しやすく, また pseudo TIC の値が大きくなることを確認した.

さらに, Fast Gradient Sign Method (FGSM) によって生成された敵対的摂動を用いて, SGD と適応的最適化手法によって学習されたモデルの騙されやすさ (頑健性) の比較も行った. C3 on CIFAR-10 においては FGSM によって生成された敵対的摂動に対して, SGD に比べて適応的最適化手法によって学習されたモデルの方が頑健性が確保されにくいことを確認した.

以上の実験結果より, 今回用いた画像分類タスクにおける DNN モデルの学習結果において, SGD に比べて適応的最適化手法によって得られる最終的な重みパラメータの汎化性, 頑健性は確保されにくいことが示された. 今後の課題としては, 自然言語処理や音声処理などの画像分類以外の深層学習タスクにおいても同様の分析を行う必要がある. また本研究で得られた最適化手法ごとのヘッセ行列の固有値に関する知見を利用して, 汎化性・頑健性を考慮した最適化手法の開発も行う予定である.

研究指導教員：若林 啓

副研究指導教員：加藤 誠