



Title: Performance of Neural Network Models Without Backpropagation

(誤差逆伝播法を用いないニューラルネットワークの性能)

Authors: Hiroaki Kurokawa

(黒川弘章 (東京工科大 工学部 教授))

Journal: Mechanisms and Machine Science 177 (2025) pp. 388 - 396

掲載年月: 2025 年 4 月

研究概要: ニューラルネットワークは、近年の人工知能応用を支える基盤技術となっています。ディープラーニングは、ニューラルネットワークの 50 年にわたる歴史の中で大きなブレイクスルーをもたらし、畳み込みニューラルネットワーク (CNN) はその代表例です。CNN は多層のニューラルネットワーク構造に基づいて構築され、学習には誤差逆伝播法を用いています。一方で、誤差逆伝播法に依存しないニューラルネットワークモデルも開発されており、代表的なものに動径基底関数 (RBF) ネットワークやエクストリームラーニングマシンがあります。これらの代替モデルは、小規模な回帰タスクにおいて高い性能を発揮します。これらのモデルの主な利点は、計算時間が非常に短くて済む点にあり、多大な計算時間を要する誤差逆伝播法の欠点を克服する手段となっています。本研究では、RBF ネットワークと誤差逆伝播法を用いる一般的な多層ニューラルネットワークとの性能差について検証しています。具体的には、遺伝子制御ネットワークの推定において、両手法の推定精度および学習時間を比較しています。さらに、この課題に対する推定精度を向上させる手法を提案し、ディープラーニングに依存しない性能向上のための代替アプローチについても検討します。

研究背景: 近年、人工知能は社会に深く浸透し、これまでにない高度なサービスを提供するようになってきました。こうした現代の AI サービスを可能にしている基盤技術がニューラルネットワークです。ニューラルネットワークの起源は 1950 年代にさかのぼりますが、誤差逆伝播学習の提案により 1980 年代の大ブームを引き起こしました。さらに最近では計算能力の向上やビッグデータの利用可能性により、性能が飛躍的に向上しています。近年の人工知能に用いられる深層学習でも、学習の中心となるのは誤差逆伝播法です。生成的敵対ネットワーク (GAN) やリカレントニューラルネットワーク (RNN)、トランスフォーマーといった、最近の人工知能の中心をなす新しいニューラルネットワークモデルの学習にも用いられています。誤差逆伝播法は、多層ニューラルネットワークにおける標準的な学習手法ですが、反復計算に依存するため、比較的高い計算コストを必要とします。一方で反復計算によらない学習を行うニューラルネットワークモデルも存在します。代表的なモデルとして動径基底関数ネットワークがあり誤差逆伝播法を用いる多層ニューラルネットワークに比べて非常に軽い計算コストで学習を完了させることができます。このようなモデルを用いて問題を解決することはエネルギー問題の解決に繋がり、サステイナブルな社会の実現にも大きく貢献します。

研究成果: 本研究では、誤差逆伝播法を用いないニューラルネットワークモデルとして動径基底関数ネットワークを扱い、現実的な問題に適用した際の性能評価を行いました。計算量は圧倒的に少ないのですが、得られる結果の精度に難があります。そこで、新たにさまざまな条件で出した答えの中から多数決で最終的な答えを決定するアンサンブル学習の考え方を導入し誤差逆伝播法を用いた場合を遜色のない結果を圧倒的に少ない計算コストで実現しました。

社会への影響: 近年の人工知能ブームを牽引するニューラルネットワークの技術は誤差逆伝播法と呼ばれる計算コストの高いアルゴリズムが基本となっています。ChatGPT に代表される現在の生成系 AI では計算機を動かすために必要なエネルギーの膨大さが無視できないほどに膨れ上がっています。本研究で評価したアルゴリズムは驚くほど少ない計算量で動作しますが、計算結果の良さに難点が生じます。これをうまく利用して性能を向上させるための提案を含めた評価をしています。

専門用語:

多層ニューラルネットワーク: 単純な計算ノードを重み付きリンクで結合したネットワークで、人工知能を支える基本的な技術です。学習により重みを最適化し、任意の入力に対して正しい答えを出力するように調整することができる仕組みです。

誤差逆伝播法: 多層ニューラルネットワークに適用される典型的な学習法で、出力の誤差を最小化するためにニューラルネットワークを

構成するパラメータを調整するアルゴリズムです。少しずつ最適解に近づけていく反復法を用いるため、ニューラルネットワークの規模が大きくなると計算コストが膨大になる問題点があります。

動径基底関数ネットワーク：多層ニューラルネットワークの一種で、計算ノードの中で行う計算が一般的な多層ニューラルネットワークと異なるものです。ノード内の計算で動径基底関数を用いるためこのような名前が付いています。任意の入力に対して正しい答えを出力するように調整することができる仕組みであることは多層ニューラルネットワークと変わりません。

アンサンブル学習：複数の機械学習モデルを組み合わせ、単体のモデルよりも高い精度や汎化性能を目指す手法のことです。本研究では同一のモデルの初期条件を変えて複数の答えを出し、それらの多数決を取るシンプルな方法を適用しています。