



| | |
|------------------------|---|
| Title | ニューラルネットワークの効果的な訓練のための探索と収束の制御 [論文内容及び審査の要旨] |
| Author(s) | 高瀬, 朝海 |
| Citation | 北海道大学. 博士(情報科学) 甲第13298号 |
| Issue Date | 2018-09-25 |
| Doc URL | http://hdl.handle.net/2115/71804 |
| Rights(URL) | https://creativecommons.org/licenses/by-nc-sa/4.0/ |
| Type | theses (doctoral - abstract and summary of review) |
| Additional Information | There are other files related to this item in HUSCAP. Check the above URL. |
| File Information | Tomoumi_Takase_review.pdf (審査の要旨) |



[Instructions for use](#)

学位論文審査の要旨

博士の専攻分野の名称 博士 (情報科学) 氏名 高瀬 朝海

審査担当者 主査 教授 栗原 正仁
副査 教授 山本 雅人
副査 教授 川村 秀憲
副査 教授 小野 哲雄
副査 准教授 小山 聡

学位論文題名

ニューラルネットワークの効果的な訓練のための探索と収束の制御
(Controlling Exploration and Convergence for Effective Training of Neural Networks)

ニューラルネットワークは、人工知能研究の一分野として活発に研究されてきたネットワーク型のシステムである。生物の神経回路網に模して、変数をニューロンに模したノードに対応させ、それらを重み付きのリンクによって結合することにより、変数間の関係を表現する。特に、階層型のネットワークは、数学的には、重みのベクトルをパラメータとし、入力ベクトルと出力ベクトルを関連付ける関数を表現する。

重みの値を適切に設定するためには、学習あるいは訓練と呼ばれる処理が必要である。典型的には、望ましい入出力関係を満たす訓練データを多数用意し、関数がそれらにできるだけ良く適合するように重みを設定する。そのアルゴリズムの基本的な考え方は、望ましい入出力関係と実際の入出力関係の違いを誤差関数 (あるいは損失関数) と呼ばれる関数として定義し、その値を最小化する重みを探索することである。実用上の観点から、この探索は最急降下法を基礎として、誤差関数の勾配が減少する方向に、適当なステップ幅で重みを反復的に更新していく方法で行われる。近年、この方法は大きく発展し、振る舞いに確率的な変化を持たせる確率的勾配降下法や、ステップの方向や幅を精密に制御できる AdaGrad などの方法が開発されてきたが、そのさらなる発展が期待されている。

本研究では、ニューラルネットワークの訓練に用いられる探索技術をさらに発展させることを目的として、訓練過程をその前半と後半に分けて、それぞれにおいてアルゴリズムが効果的に振る舞い、最終的に望ましい解に安定して収束するように探索の制御を行う方法について議論している。具体的には、訓練の前半は誤差の大きい局所解を避けて広い範囲を探索し、後半は誤差の小さな解に向けて安定した収束を行わせるという戦略に基づき、それを具体化する2つの方法を考案し、議論することによって、次のような成果を得ている。

(1) ニューラルネットワークのハイパーパラメータの一部である学習率 (重みの更新ステップ幅を制御する値) に着目し、その値を訓練中に動的かつ適応的に増減させる方法 (適応的学習率調整法) を提案し、この方法が上記の戦略に沿っていること、及びその有効性を実証している。

(2) ハイパーパラメータの一部であるミニバッチ (訓練データの集合の部分集合。重みの更新を一括して行うための処理単位) のサイズに着目し、その値を漸増させる方法 (バッチサイズ増加法) を提案し、この方法が上記の戦略に沿っていること、及びその有効性を実証している。

本論文は次のように構成されている。

第1章では、研究背景、研究目的、および本論文の構成について述べている。

第2章では、本論文を理解するために必要な背景知識を説明している。まず、多層パーセプトロン及び畳み込みニューラルネットワークについて紹介し、その学習アルゴリズムの基礎となる勾配降下法及びそれを発展させた確率的勾配降下法や AdaGrad などの方法を説明している。その後、誤差関数の非凸性を考慮して、訓練過程をその前半と後半に分けた訓練戦略について議論している。また、関連手法であるシミュレーテッドアニーリング (焼きなまし法) は、ニューラルネットワークには効果的には適用できないことを説明している。

第3章では、学習率の動的変化と誤差の小さな解への収束との関係について論じ、提案手法である適応的学習率調整法について説明している。さらに、実験結果に基づき、この方法が前章で述べた戦略に沿っていること、及びその有効性について議論するとともに、この方法の問題点について述べている。

第4章では、ミニバッチのサイズを動的に増加させることと、探索がフラットな局所解に収束することとの関係について論じ、提案手法であるバッチサイズ増加法について説明している。まず、非凸最適化問題の関連研究で示されていた、大きなバッチサイズを用いたときに汎化性能が悪化する結果について紹介し、その理由を誤差関数の非定常的な特性に基づいて説明し、その議論を発展させて、効果的な訓練を行うバッチサイズ増加法を導いている。その後、実験結果に基づき、この方法が第3章で述べた戦略に沿っていること、及びその有効性について議論している。

第5章では、本論文のまとめ及び今後の課題について述べている。

これを要するに、著者は、ニューラルネットワークの訓練過程をその前半と後半に分け、それぞれにおいてアルゴリズムが効果的に振る舞い、最終的に望ましい解に安定して収束するように探索の制御を行う方法を考案することにより、ニューラルネットワークの学習技術をさらに発展させるための知見を得たものであり、人工知能学に貢献するところ大なるものがある。よって著者は、北海道大学博士 (情報科学) の学位を授与される資格あるものと認める。