

令和 6 年 6 月 13 日現在

機関番号：32706

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K11979

研究課題名（和文）モーメント準ニュートン法と確率的分散縮小法の融合による強非線形ビッグデータの学習

研究課題名（英文）Learning of Strongly Nonlinear Big Data by Momentum Quasi-Newton Method combined with Stochastic Variance Reduction

研究代表者

二宮 洋（Ninomiya, Hiroshi）

湘南工科大学・情報学部・教授

研究者番号：60308335

交付決定額（研究期間全体）：（直接経費） 2,700,000円

研究成果の概要（和文）：本研究では、より複雑（強非線形）かつ膨大なデータ（ビッグデータ）の処理を可能とするニューラルネットワークに対する新たな学習法を開発した。この過程で、モーメント法、準ニュートン法、ネステロフの加速勾配法、および、不動点加速法との関連性に対する検討を行い、高速化に対する各アルゴリズムのメカニズムの検証やその融合による収束性に関する検討を行った。また、様々な検証用例題や実問題に対するコンピュータシミュレーションを通して、その有効性および問題点を明らかにすることができた。これらの研究により、従来では実現不可能であった複雑さを持つ学習データに対するニューラルネットワークを用いた解決に関する知見を得た。

研究成果の学術的意義や社会的意義

IoTの発展により、あらゆる場面でデータが蓄積され、これまで全く無関係であると考えられてきたデータを同時に扱うことが必要な時代となった。従って、今後はデータ量が増大するだけでなく、より複雑な関係性を内包する大規模データの扱いが必要となってきた。一方、大規模データの学習には、データの一部を用いた勾配による確率的勾配法が必要であるが、強非線形データの学習では十分に高精度な解を得られないといった問題があった。本研究では、この両者の特徴を持つ強非線形ビッグデータの学習に対する高精度かつ高速な手法の可能性を探り、従来よりも強力な学習アルゴリズムの開発に成功したことに学術的および社会的意義がある。

研究成果の概要（英文）：In this research, we developed a new learning method for neural networks that enables processing of more complex (strongly non-linear) and massive data (big data). In this process, we examined the relationship with the momentum method, quasi-Newton method, Nesterov's accelerated gradient method, and fixed-point acceleration method, and verified the mechanisms of each algorithm for acceleration and discussed the convergence by integrating them. Furthermore, through computer simulations on various test problems and real problems, we were able to clarify their effectiveness and issues. Through this research, we gained insights into solving problems using neural networks for learning data with a complexity that was previously impossible to achieve.

研究分野：人工知能

キーワード：ニューラルネットワーク 学習アルゴリズム 準ニュートン法 モーメント法 不動点加速法 ネステロフの加速勾配法

1. 研究開始当初の背景

近年の人工知能 (AI) と IoT の発展により、多様な分野の膨大なデータの蓄積が可能となった。その実用の為に、強非線形なビッグデータの高精度かつ高速なニューラルネットワーク (NN) による処理に注目が集まっている。特に、AI のコア技術として深層学習 (DL) に端を発した大規模な深層 NN の利用である。その中心に位置する研究の一つが NN の学習アルゴリズムである。現在、そのほとんどが計算量の観点から、勾配のみを用いる 1 次近似勾配法を中心に研究されている。しかし、これらの手法では強非線形ビッグデータを高精度かつ高速に学習することは困難であった。一方で、ハードウェアの発達により、計算量は多くなるが高精度な学習が可能である準ニュートン法 (QN) に代表される、勾配と曲率情報を用いる 2 次近似勾配法に対する関心が高まっている。正確には、QN は、2 次近似をさらに近似しており、収束は超 1 次特性となるが、便宜上、2 次近似手法と表現する。従って、強非線形ビッグデータの高精度かつ高速な学習を可能とする QN に基づく新たな学習法を確立することで、これまでよりも飛躍的に複雑で大規模なデータ処理が可能となり、データに内包される非線形性を NN がより詳細に理解することも可能となることが予測され、AI の応用範囲の飛躍的な拡大につながる。

一方で、1 次近似手法の高速化にはモーメント法や勾配の導出に慣性項を用いたネステロフの加速勾配法が応用されており、近年、これに加えて学習係数を適応的に変化させる手法、Adam 等が多く用いられている。しかし、強非線形性を内包する学習データの高精度な学習に対しては、これらの 1 次近似手法では現実的な時間で学習することが不可能で、QN のような高精度な手法が必要であった。この様な状況の中、QN の強力な収束特性を最大限に引き出すために、1 次近似手法と同様に慣性項を QN に導入することで得られる“慣性付 2 次近似勾配モデル”を用いた“ネステロフの加速勾配準ニュートン法 (NAQ)”が提案され、慣性項の導入による QN の高速化が行われた。しかしながら、この手法は 1 度の反復にすべての学習データを用いるバッチ学習であった。一方、“ビッグデータ”の学習には、データの一部を用いた勾配による確率的勾配法 (ミニバッチ学習) が必要であるが、“強非線形データ”の学習には確率的手法を導入した QN では十分に高精度な解を得られないことが多い。従って、本研究課題では、この両者の特徴を持つ、強非線形ビッグデータの学習に対する高精度かつ高速な学習アルゴリズムの確立を目指す。

2. 研究の目的

本研究の目的は、学習データの複雑・大規模化を予想し、これを高速かつ高精度に学習を可能とする新たなアルゴリズムの開発である。この実現のために必要な基礎技術が、QN 等の 2 次近似手法に基づく学習法である。しかし、QN を 1 次近似手法と同様にモーメント法で加速させるアイデアは世界的にもほとんど研究されておらず、NN の学習に限っては、上述した NAQ のみである。従って、NAQ をさらに改良し高速・高精度かつ安定なアルゴリズムの構築を目指すことで、他にはない学習アルゴリズムの開発を行うことを目的としている。また、これまでの研究では、データが複雑になる場合、NN の規模を大きくする、例えば深層化することで、その表現能力を高め、1 次近似学習法で解決してきた。しかしながら、深層 NN は表現力が高い一方でパラメータ空間の次元が高く、その振る舞いも複雑である。例えば局所解や鞍点、関数表現の冗長性に依存したパラメータの変化に対してモデルの感度が低い領域の存在等の問題がある。従って、データの複雑かつ大規模化に対応しつつも、深層化のみに頼らず、関数の複雑さに影響されにくい高速なアルゴリズム、つまり、強非線形ビッグデータを学習可能な新たなアルゴリズムの開発は未解決な問題であり、この未解決問題を解決する学習法の開発が本研究の最終目標である。

3. 研究の方法

本研究は、以下の 3 点に関して研究を進めることで実現する。

(1) モーメント準ニュートン法と不動点加速法の融合による NN に対する学習アルゴリズムの高速化及び安定化手法の確立

まず、強非線形データを安定して高精度かつ高速に学習できる QN に基づくアルゴリズムの確立を目指す。申請者はこれまでに、NAQ、および、その問題点 (1 反復に 2 度の勾配計算が必要な点) を克服した“モーメント準ニュートン法 (MoQ)”を提案してきた。これらは、QN と比較して、学習時間を、NAQ は $2/3 \sim 1/3$ に、MoQ は $1/4 \sim 1/9$ へと大幅に削減することに成功した。一方で、これらのアルゴリズムには、モーメント係数 (ハイパーパラメータ) の値によって収束が不安定になってしまう問題があった。本研究では、この問題を解決しつつ、さらなる高速化を目指す。具体的には、MoQ で用いる勾配を一般化することで得られる勾配に着目する。結果として、ここでは、MoQ の加速性能を向上させ、学習効率を向上させる。さらには、パラメータを適応的に求める手法を用いることで、提案手法の解析がアルゴリズムの収束性に対する問題として明確にする。最終的に、QN の収束性に関するこれまでの多くの知見を利用することで、提案手法のメカニズムを解析する。

(2) 確率的分散縮小法の導入によるビックデータの複雑化(強非線形化)を考慮した学習アルゴリズムの開発

次に、強非線形“ビック”データに対応したアルゴリズムへと発展させる。具体的には(1)で開発した手法に確率的処理を導入する。これまでのQNやNAQに対する確率的処理の導入では、同じ例題に対しては、1次近似手法と比較して高速な収束性や局所解や鞍点等の問題解決能力を持つことは示されてきた。しかし、強非線形データに対する有効性や高精度な解は申請者の実験においても示されていない。この原因は強非線形データ特有の学習データのばらつきにあると考えられる。また、QNでは、勾配を用いて曲率情報(近似ヘッセ行列)を更新する必要があり、確率的勾配を用いることで曲率情報も学習データの複雑さに応じてばらつきが大きくなり、収束への影響はさらに強くなる。従って、反復毎の勾配の分散を縮小させる仕組みをアルゴリズムに導入する必要がある。本研究では、この仕組みを取り入れることが可能な手法として、近年、注目されている確率的分散縮小勾配法(SVRG)に着目し、MoQと融合させることを考える。具体的には、SVRGが2重ループ構造を用いている点に着目し、MoQの曲率情報の更新には外ループでバッチ処理の勾配を用い、内ループの反復では、確率的勾配を用い、その分散を縮小させながらデータを更新するハイブリッド手法を考える。これにより、全学習データを用いるバッチ学習の強非線形データに対する強みを保ちつつ、反復のほとんどは確率的処理を用いるため、分散を縮小しつつも実用的で効率的な学習が可能となる。従って、ここで提案するアルゴリズムは、強非線形性による確率的勾配の分散を縮小しつつ、QNの高精度・高速な解を得る能力を最大限に用いることで、強非線形ビックデータの学習を可能にする。

(3) 強非線形・大規模データを持つ実問題への提案手法に対する有効性の検証

次に、実問題に対する提案手法の有効性を検証する。実問題の1つとしては、アナログ高周波回路設計の自動化に関する基礎研究を想定している。申請者はこれまで、強非線形データを用いた応用例として、高周波回路設計のためのNNによるモデリングを行ってきた。これは、物理データと回路パラメータを入力とし、周波数応答を出力とする非常に複雑な問題であり、NAQやMoQでなければ学習できなかった。これらの例題に対する有効性を検証する。また、アナログ高周波回路のモデリング以外にもNNを用いた様々な実問題が想定される。本研究は学習アルゴリズムの提案であり、NNを用いるあらゆる実問題への応用の可能性がある。従って、提案手法の応用範囲の特定やその可能性を探るために、様々な実問題を想定した研究も想定している。結果として、AIの発展に大きく貢献することが予想され、その新たな可能性の探究へとつながっていく。

4. 研究成果

研究成果としては、本研究の基となる、MoQに関する収束性、及び、ハイパーパラメータの解析的な導出によるロバスト性の検証の基礎研究を行い、モーメント係数(ハイパーパラメータ)の適応的な決定方法を導入するメカニズムの解明を目指した。その結果、NNの学習の初期値に対するロバスト性を、パラメータを固定した場合に比べて飛躍的に改善できることを示した。さらに、確率的勾配法や記憶制限法を導入した大規模データ学習を考慮したアルゴリズムへの改良に対する研究を行った。これまでのQNやNAQに対する確率的処理の導入では、同じ例題に対しては、1次近似手法と比較して高速な収束性や局所解や鞍点等の問題解決能力を持つことは示されてきたが、強非線形データに対する有効性や高精度な解は示されていない。特に、QNでは、勾配を用いて曲率情報を更新する必要があり、確率的勾配を用いることで曲率情報も学習データの複雑さに応じてばらつきが大きくなる。また、収束への影響はさらに強くなることを示されており、実験により確認も行った。以上より、反復毎の勾配の分散を縮小させる仕組みをアルゴリズムに導入した、確率的分散縮小勾配法(SVRG)に着目した。つまり、SVRGが2重ループ構造を用いている点に着目し、MoQの曲率情報の更新には外ループでバッチ処理の勾配を用い、内ループの反復では、確率的勾配を用い、その分散を縮小させながらデータを更新するハイブリッドアルゴリズムに関して、その実現可能性を検討した。以上の知見を基に、確率的分散縮小法の導入によるビックデータの複雑化(強非線形化)を考慮した学習アルゴリズムの開発に着手した。結果として、提案するアルゴリズムは全学習データを用いるバッチ学習の強非線形データに対する強みを保ちつつ、反復のほとんどは確率的処理を用いるため、分散を縮小しつつも実用的で効率的な学習が可能となることをいくつかの実問題に対する実験により確認した。さらに、提案アルゴリズムの有効性を示すために、強非線形学習データを持つ様々な実問題への応用を検討した。この1つとして、近年、ニューラルネットワークの応用として注目を集めている画像認識や生成AIを新たな実問題として取り上げ、それらの問題のデータの非線形性や問題規模などの調査を行うとともに、それらの問題を本研究の提案手法で扱うための基礎研究に着手した。これにより、今後、提案手法の有効性に対する検証の幅を広げられることが期待でき、AIの発展に寄与できる研究につながっていくことができると予想される。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 2件）

1. 著者名 Yasuda Sota, Indrapriyadarsini S., Ninomiya Hiroshi, Kamio Takeshi, Asai Hideki	4. 巻 13
2. 論文標題 addHessian: Combining quasi-Newton method with first-order method for neural network training	5. 発行年 2022年
3. 雑誌名 Nonlinear Theory and Its Applications, IEICE	6. 最初と最後の頁 361 ~ 366
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/nolta.13.361	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Sendilkumaar, Indrapriyadarsini, Mahboubi, Shahrzad and Ninomiya, Hiroshi, and Asai, Hideki	4. 巻 vol.11, issue 4
2. 論文標題 aSNAQ : An Adaptive Stochastic Nesterov 's Accelerated Quasi-Newton Method for Training RNNs	5. 発行年 2020年
3. 雑誌名 IEICE NOLTA	6. 最初と最後の頁 409-421
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -
1. 著者名 Sendilkumaar, Indrapriyadarsini, Mahboubi, Shahrzad and Ninomiya, Hiroshi, Takeshi Kamio and Asai, Hideki	4. 巻 N/A
2. 論文標題 A Nesterov's Accelerated quasi-Newton method for Global Routing using Deep Reinforcement Learning	5. 発行年 2020年
3. 雑誌名 Proc. NOLTA 2020	6. 最初と最後の頁 251-254
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 R. Yamatomi, S. Mahboubi and H. Ninomiya	4. 巻 14325
2. 論文標題 Generative Model of Suitable Meme Sentences for Images Using AutoEncoder	5. 発行年 2023年
3. 雑誌名 Proc. PRICAI 2023	6. 最初と最後の頁 237-248
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計7件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 マハブービ シェヘラザード, 二宮 洋
2. 発表標題 慣性項による確率的重み差分伝播法の高速化に関する研究
3. 学会等名 電子情報通信学会 信学技報 非線形問題研究会
4. 発表年 2022年

1. 発表者名 山富龍, マハブービ シェヘラザード, 二宮洋
2. 発表標題 2次情報を用いた二重適応縮約型学習アルゴリズムにおけるネステロフの加速勾配の効果に関する研究
3. 学会等名 2022 NOLTAソサイエティ大会
4. 発表年 2022年

1. 発表者名 山富 龍, マハブービ シェヘラザード, 二宮 洋
2. 発表標題 大喜利生成AIの実装に関する研究
3. 学会等名 情報処理学会第85回全国大会, 6V-07
4. 発表年 2023年

1. 発表者名 山崎 太郎, 山富 龍, マハブービ シェヘラザード, 二宮 洋
2. 発表標題 道路標識検出用SSDのためのUnityを用いたアノテーション自動化に関する研究
3. 学会等名 情報処理学会第85回全国大会, 6Q-04
4. 発表年 2023年

1. 発表者名 Mahboubi Shahrzad and Ninomiya Hiroshi
2. 発表標題 Weight Difference Propagation for Stochastic Gradient Descent Learning
3. 学会等名 The Eighteenth International Multi-Conference on Computing in the Global Information Technology, ICCGI 2023 (国際学会)
4. 発表年 2023年

1. 発表者名 安田 壮太, S. Indrapriyadarsini, 二宮 洋, 神尾 武司, 浅井 秀樹
2. 発表標題 ニューラルネットワーク学習のための適応的な分散低減二次手法
3. 学会等名 NOLTAソサイエティ大会
4. 発表年 2021年

1. 発表者名 田中和真, マハブービ・シャヘラザード, 二宮 洋
2. 発表標題 慣性項付き記憶制限準ニュートン法を用いた深層強化学習のTensorflowへの実装
3. 学会等名 IEICE 総合大会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------