

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 15 日現在

機関番号：32706

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26330287

研究課題名(和文) 複雑な非線形特性を内包する大規模データを対象にした勾配法に基づく学習アルゴリズム

研究課題名(英文) A Training Algorithm based on Gradient Method for Big Data Including High-nonlinearity

研究代表者

二宮 洋(Ninomiya, Hiroshi)

湘南工科大学・工学部・教授

研究者番号：60308335

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：本研究課題では、複雑な非線形特性を持ち、その特性を表現するために必要な学習データ数が大規模になる非線形システムのニューラルネットワークによる近似モデルを実現する。このため、以下の2点に関して研究する。1つ目は、『統計的手法を用いた複雑な非線形特性を内包する大規模データの分散化』、2つ目は『誤差関数の凸化による学習アルゴリズムのロバスト性の向上とその分散化』である。これらより、従来では実現不可能であった複雑さと規模の学習問題を解決するアルゴリズムの開発を目指す。さらに、回路シミュレータ等への応用、特に高周波回路や新たなデバイスの詳細な近似モデルのNNによる実現を目的とした研究へと発展させる。

研究成果の概要(英文)：In this research, it is a purpose to enable the approximation model by the feedforward neural networks for the function or the system with the highly nonlinear behavior and huge data by the following studies. Specifically, "Distribution of large scale data including highly nonlinear characteristics using statistical method", and "Improvement of robustness of training algorithm by convexity of error function and its decentralization". Aimed at the development of the proposed algorithm to solve the complexity and scale of the training problem was not feasible with conventional methods. Furthermore, this approach is useful for the circuit modeling for the design and optimization, where analytical formulas are not available or original model is computationally too expensive. A neural model is trained once, and can be used again and again. This avoids repetitive circuit simulations where a change in the physical dimension requires a re-simulation of the circuit structure.

研究分野：学習アルゴリズム

キーワード：ニューラルネットワーク 学習アルゴリズム 勾配法 準ニュートン法 大規模データ

1. 研究開始当初の背景

従来、ニューラルネットワーク(以下、NN)による非線形システムの高精度で簡便な近似モデルの実現に関しては様々な研究がなされてきた。しかしながら、これまでの研究では複雑な非線形特性(以下、強非線形と呼ぶ)の近似モデルを実現することができるのは、次元が小さく、学習サンプル数がそれほど多くない問題に限られていた。一方、近年のクラウドコンピューティングの発展により、膨大なデータの蓄積が可能になり、様々な分野でビッグデータの処理の必要性が増してきた。

これらの実現のためには、2つの問題を解決する必要がある。1つは大規模データを扱うために、分散処理システムを活用したデータの分散化とそれに伴う新たなNNの学習アルゴリズムの開発である。2つ目は、NNの勾配学習において勾配ベクトルがゼロとなり学習が継続できなくなる問題、つまり、局所解やNNの構造自体に起因するプラトーなどの問題を解決する必要がある。ここで、本研究課題の対象とする強非線形システムは、ノイズのような振る舞いに見える学習サンプルであっても、モデル化が必要な学習データを持つシステムである。従って、バッチ学習が必要となる。その上で、汎化能力を高めるためにパラメータの数を少なく抑える必要がある。

現在、クラウドコンピュータ上に分散したデータを統一的に処理する研究が盛んである。これは、機械学習においても反復パラメータ混合法(IPM)として盛んに研究されている。IPMは、各サーバが個別に保有する学習サンプルの部分集合を用いてパラメータの更新を行い、それらを1つにまとめて混合する手法である。従って、パラメータの通信コストのみでバッチ学習が可能となる。しかし、これまでの研究ではクラウド上の各サーバが保有すべき学習サンプル数に関する議論はほとんどされておらず、1回のパラメータの更新に必要な学習サンプル数に関する研究もほとんどない。これは、これらの研究の問題設定が凸計画問題を対象としており、1次の勾配法で解を求めることができたためである。一方、NNの学習は非線形計画問題であり準ニュートン法等の2次の勾配法が必要である。また、準ニュートン法にIPMを適用することで、1回の更新に必要な学習サンプル数によって結果が大きく変化することが報告されている。これは、各サーバが持つ学習サンプルの評価関数から得られた勾配を用いてパラメータを更新した場合、その更新が全学習データを用いた評価関数を減少させているかどうかの判断ができないことに起因している。そこで、本研究課題では、全学習データを減少させることができる学習サンプル数を、統計的手法を用いて決定することを試みる。申請者はこれまで、オンライン勾配学習法において、学習データ

の部分集合を用いた1回のパラメータ更新で、全データの誤差関数を減少させることが可能な学習サンプル数に関して検討を行い、統計的手法を用いて解析的に導出する手法を提案した。従って、この手法を発展させることで、分散システムにおいても、各サーバに必要な学習サンプル数を決定可能であると予想された。

一方、局所解問題を回避するための手法としては、申請者によってオンライン学習法とバッチ学習法を組合せた手法が提案されていた。しかしながら、これらの手法は理想的には無限の時間を必要とする手法であり、実際にコンピュータを用いて解く場合、問題の規模によっては現実的な時間では解を得ることができなくなってしまう。このため、本研究のように大規模データを扱う場合には、勾配法に基づくアルゴリズムのみで局所解問題を回避する必要がある。そこで、本研究で注目する手法が、誤差関数を凸化させることで局所解問題を回避する手法である。この手法は勾配法を用いて最適化できるため、計算時間の点ではMHMと比較して優位である。しかしながら、凸化誤差関数はバッチ学習を前提に検討されているため、分散システムによる学習には適さないアルゴリズムであった。

2. 研究の目的

本研究課題では、複雑な非線形特性を持ち、かつ、その特性を表現するために必要な学習データ数が大規模になる非線形システムのNNによる近似モデルを実現することを目的とする。この目的の実現のために、まず、以下の2点に関して研究する。

- (ア) 統計的手法を用いた複雑な非線形特性を内包する大規模データの分散化
- (イ) 誤差関数の凸化による学習アルゴリズムのロバスト性の向上とその分散化

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と規模のNNによる学習問題を解決するアルゴリズムの開発を目指す。さらに、回路シミュレータ等への応用、特に高周波回路や新たなデバイスの詳細な近似モデルのNNによる実現を目的とした研究へと発展させる。

3. 研究の方法

以下の手順で研究を進める。

(1) 複雑な非線形特性を持つ大規模データの統計的手法による分割方法の確立

本研究において対象としている大規模学習データはバッチ処理が必要なデータであることがほとんどである。従って、分散システムでバッチ処理を実現するためには、反復パラメータ混合法(IPM)が現実的な手法である。IPMは各サーバが持つ学習サンプルを用いてパラメータを更新し、このパラメータのみを1反復毎に通信し、混合する手法で

ある。しかしながら、従来のIPMは1次の勾配法を対象としているため、各サーバに必要なサンプル数に関しては厳密に求める必要はなかった。しかし、NNの学習、特に強非線形特性のモデル化では2次の勾配法(準ニュートン法等)が必要となる。以上より、各サーバが持つ学習サンプルから得られる誤差関数の勾配による更新が、全学習サンプルの誤差関数を減少させているかに関する議論は非常に重要となる。従って、本研究の分散システムにおいては、各サーバに割り当てる学習サンプル数を統計的に全学習データによる誤差関数を減少させるように求める。

(2) 分散処理された大規模データを用いた準ニュートン法に基づくバッチ学習法

さらに、(1)によって分割処理された大規模データの部分集合を用いたIPMを検討する。勾配法には準ニュートン法を用いることで、分散システムにおける学習にもかかわらず、準ニュートン法に基づくバッチ学習法の強力な収束特性を持つアルゴリズムが構築可能であると考えられる。

(3) NNの誤差関数の凸化による学習のロバスト性と準ニュートン法に対する有効性

本研究課題において解決しなければならない問題に局所解問題がある。一般的に、NNの学習は、対象となるシステムの特徴が複雑になるほど局所解問題は顕著になる。近年、誤差関数を疑似的に凸化することで、局所解の問題を回避する手法が注目を集めている。この関数はリスク回避誤差関数と呼ばれており、最適制御分野で盛んに研究されている手法である。また、微分可能な関数であるため勾配法を用いて最適化することができる。従って、局所解を回避する手法として一般的に知られているメタ・ヒューリスティック手法に頼る必要がなくなり、実用的な計算時間で解を得ることができる。本研究では、この凸化誤差関数に着目し、そのロバスト性の検証と準ニュートン法や記憶制限準ニュートン法に対する有効性を調査する。

(4) 分散処理に適した凸化誤差関数の提案とオリジナル凸化誤差関数の上界に対する検討

(3)での検討をもとに、凸化誤差関数を分散処理可能な関数へと変形することを考える。凸化誤差関数は、通常のNNの学習で用いられる平均二乗誤差をlog-sum-exp関数を用いて凸化させている。ここで、sum関数は学習サンプルに対する和である。従って、学習サンプルに対して分割することができない。つまり、バッチ学習を前提にした手法であった。そこで、本研究では、log関数の凸性を利用して、オリジナルの凸化誤差関数を学習サンプルごとのsum関数に変更することを考える。また、提案する分散処理に適した凸化誤差関数をオリジナルの凸化関数

の上界として定義することを試みる。さらに、log関数の凸性を利用して、その定義された上界の最適解が、オリジナルの凸化誤差関数の最適解と一致することを解析的に示すことも検討する。これらの研究により、局所解の問題を回避することが可能となる凸化誤差関数を、最適解を保証しながら、分散システム上の各サーバに分割することが可能となる。

(5) 実際の大規模データを用いた実験及び提案アルゴリズムの有効性の検証

本研究で対象とする実問題の1つである、高周波回路のNNによるモデル化に対する基礎研究を行う。高周波回路の周波数応答は非常に複雑な非線形性を内包しており、NNによるモデル化は従来のモデルをはるかにしのぐ柔軟で有効なモデルが実現できた。しかし、これまでの手法では回路の大規模化及び微細化に伴う、学習データの大規模化に対応できなかった。この技術を確立するための基礎研究を行う。

4. 研究成果

本研究では、それぞれの目的に対して、以下の研究成果を得た。

(1) 複雑な非線形特性を持つ大規模データの統計的手法による分割方法の確立

(1)に関しては、すべての学習データを最小化に導く部分集合に必要な学習データ数を、統計的情報を用いて、導出する手法を解析的に確立し、これを応用した準ニュートン法に基づくアルゴリズムを実現した。この手法により、準ニュートン法におけるミニバッチ手法とバッチ手法の関係を検証することができた。また、従来のミニバッチ処理からバッチ処理に変更するアルゴリズムに対して適応した結果、その有効性を示すことができた。

(2) 分散処理された大規模データを用いた準ニュートン法に基づくバッチ学習法

(2)に関しては、準ニュートン法にIPM処理を試みた結果の検証、及び、その分割数が結果にどのように影響するかを検証した。基本的には準ニュートン法に対して、IPM処理の有効性を見出すことができなかった。つまり、部分集合の学習データ数に関しては、学習誤差への影響にあるものの、分割することでほとんどの場合、良好な(テスト誤差を小さくする)学習誤差を得ることはできなかった。

(3) NNの誤差関数の凸化による学習のロバスト性と準ニュートン法に対する有効性

(3)に関しては、誤差関数の凸化と、NNの学習に対する有効性を検証した。また、準ニュートン法に対しての有効性についても検証した。

(4) 分散処理に適した凸化誤差関数の提案とオリジナル凸化誤差関数の上界に対する検討

本研究課題の本質的なアルゴリズムの部分である。凸化誤差関数を分散処理に適した形にするために、log 関数の凸性を利用することで実現することができた。また、log 関数だけではなく、Products of Sigmoids Bound を用いた分散処理に適した凸化誤差関数を実現した。それぞれの誤差関数に関しては、上界を数学的に示し、学習アルゴリズムにおいてはその上界を最小化することでNNの学習を実現している。また、このアルゴリズムを実装し、NNの学習において、分散処理に対する手法の有効性、及び、NNの学習に対するロバスト性をシミュレーションにより示すことができた。また、複雑な非線形特性を持つ大規模データの学習に関してもシミュレーションを行いその有効性を示した。

(5) 実際の大規模データを用いた実験及び提案アルゴリズムの有効性の検証

(5)に関しては、現在進行中である。これまでに、小さい規模のベンチマーク問題に対しては検証を行い、その有効性を確認した。今後は、より大規模な問題に対して検証を行う予定である。

5. 主な発表論文等

[雑誌論文](計 4件)

- 1 Ninomiya, H.: "Neural Network Training based on quasi-Newton Method using Nesterov's Accelerated Gradient", *Proc. IEEE TENCON 2016*, pp.51-54, Nov., 2016.
- 2 Ninomiya, H.: "A Novel quasi-Newton-based Training using Nesterov's Accelerated Gradient for Neural Networks", *Proc. 2016 International Conference on Artificial Neural Networks (ICANN'16)*, Part II, LNCS 9887, pp.540, Sep., 2016.
- 3 Ninomiya, H.: "Distributed Robust Training of Multilayer Neural Networks Using Normalized Risk-Averting Error", *Proc. 2014 IEEE Symposium Series on Computational Intelligence (IEEE/SSCI'14, IEEE/CCMB'14)*, pp.134-140, Dec., 2014.
- 4 Kobayashi, M., Ninomiya, H., Miura, Y., and Watanabe S.: "Reconfigurable Dynamic Logic Circuit Generating t-Term Boolean Functions Based on Double-Gate CNTFETs", *IEICE Trans. on Fundamentals.*, vol.E97-A, no.5, pp.1051-1058, May, 2014.

[学会発表](計 10件)

- 1 Shahrzad Mahboubi, 二宮 洋: "ネステロフの加速勾配を用いた Quickprop 学習法の高速度化", *2017 年 電子情報通信学会 総合大会, N-1-23*, p.309, 2017 年 3 月. 「名城大学 (愛知県・名古屋市)」
- 2 三宅あかり, 越森恵莉菜, 二宮 洋: "ネステロフの加速勾配を用いた準ニュートン学習法に関する研究", *2016 年 電子情報通信学会 ソサイエティ大会, N-1-12*, 2016 年 9 月. 「北海道大学 (北海道・札幌市)」
- 3 吉本昌史, 小林 学, 二宮 洋: "潜在クラスモデルのニューラルネットワークによる学習", *2016 年 電子情報通信学会 ソサイエティ大会, N-1-11*, 2016 年 9 月. 「北海道大学 (北海道・札幌市)」
- 4 二宮 洋: "ネステロフの加速準ニュートン法による学習アルゴリズムの提案", *電子情報通信学会 信学技報 非線形問題研究会, vol.115, no.425, NLP2015-141*, pp.87-92, 2016 年 1 月. 「九州工業大学 (福岡県・北九州市)」
- 5 嘉藤淳紀, 渡辺重佳, 二宮 洋, 小林学, 三浦康之: "再構成可能論理回路の設計法と各種方式の比較", *電子情報通信学会 信学技報, VLD2015-77, CPSY2015-109, RECONF2015-59*, pp.1-6, 2016 年 1 月. 「慶應義塾大学 (神奈川県・横浜市)」
- 6 佐伯 誠, 坂下善彦, 二宮 洋: "粒子群最適化におけるローカル化及び寿命付きリーダーの有効性に関する研究", *情報処理学会 第 77 回 全国大会, 1*, pp.259-260, 2015 年 3 月「京都大学 (京都府・京都市)」
- 7 嘉藤淳紀, 渡辺重佳, 二宮 洋, 小林学, 三浦康之: "再構成可能論理回路の設計と評価", *電子情報通信学会 信学技報, VLD2014-119, CPSY2014-128, RECONF2014-52*, pp.35-40, 2015 年 1 月. 「慶應義塾大学 (神奈川県・横浜市)」
- 8 嘉藤淳紀, 渡辺重佳, 二宮 洋, 小林学, 三浦康之: "ダブルゲート型トランジスタを用いた再構成可能論理回路の設計法", *電子情報通信学会 信学技報, CPM2014-126, ICD2014-69*, pp.21-26, 2014 年 11 月. 「別府国際コンベンションセンター (大分県・別府市)」
- 9 佐伯 誠, 坂下善彦, 二宮 洋: "ローカル粒子群最適化における寿命付きリーダーの有効性", *電子情報通信学会 基礎・境界ソサイエティ大会, A-2-4*, 2014 年 9 月「徳島大学 (徳島県・徳島市)」

- 10 嘉藤淳紀, 渡辺重佳, 二宮 洋, 小林学, 三浦康之:"ダブルゲート MOS 型トランジスタを用いた論理回路の検討", 電子情報通信学会 信学技報, **SDM2014-78, ICD2014-47**, pp.87-92, 2014 年 8 月.「北海道大学(北海道・札幌市)」

〔その他〕

<http://www.info.shonan-it.ac.jp/ninomiya-lab/ninomiya.html>

6. 研究組織

(1) 研究代表者

二宮 洋 (NINOMIYA, Hiroshi)

湘南工科大学・工学部・教授

研究者番号：60308335