

(続紙 1)

京都大学	博士 (情報学)	氏名	中井彩乃
論文題目	Scalable Estimation on Linear and Nonlinear Regression Models via Decentralized Processing: Adaptive LMS Filter and Gaussian Process Regression (分散処理による線形・非線形回帰モデルでのスケーラブルな推定：適応LMSフィルタとガウス過程回帰)		
(論文内容の要旨)			
<p>Today large amounts of data are being collected from numerous sensors via communication networks. Such situations pose difficulties in processing those data in terms of computational and/or communication costs, as well as robustness against failures in the systems. Distributed processing is motivated by these difficulties. This thesis summarizes contributions to two instances of distributed processing of data: One is regarding improvements of the diffusion least-mean-square (D-LMS) algorithm, and the other is on extension of an approximation method of the Gaussian process regression (GPR). This thesis comprises six chapters, and the main contents are divided in two parts. Part I on the D-LMS algorithm consists of Chapters 2, 3, and 4, and Part II on the GPR consists of Chapter 5. The contents of the chapters are summarized as follows.</p>			
<p>Chapter 1 is the introduction of this thesis. It argues the demand for scalable processing of data collected over networks, which places the background of this thesis. It then moves on to discussing the main objective of this thesis, that is to enhance performance of distributed processing, and showing an overview of the contributions of this thesis.</p>			
<p>Chapter 2 is the first chapter of Part I of this thesis. It reviews the D-LMS algorithm as an example of distributed signal processing over networks. It argues that several existing variants of the D-LMS algorithm can be characterized by choices of the combination weights used in the combination step in the D-LMS algorithm. It also points out a connection between the combination step and the average consensus protocol, which lays the basis for the developments in Chapters 3 and 4. Convergence analysis and the steady-state mean squared deviation (MSD) analysis are reviewed for the generic D-LMS algorithm. This chapter also reviews the consensus propagation (CP), which offers an essential idea to the following developments.</p>			
<p>Chapter 3 proposes two variants of the D-LMS algorithm via applying the idea of CP to the combination step of the D-LMS algorithm. One of the two, called CP-LMS, is based on application of the exact CP to achieve a perfect consensus over a spanning tree extracted from the network. The other, called LCP-LMS, is based on application of the loopy CP as the combination step. It is then argued that restricting the number of iterations of the loopy CP in LCP-LMS to one allows analytical treatment of the steady-state MSD. Convergence analysis and computational complexity of the proposed algorithms are presented. Optimization of the combination weights in LCP-LMS is also discussed on the basis of an upper bound of the steady-state MSD. Via numerical simulations, the proposed algorithms are shown to be superior, in terms of the trade off between convergence speed and the steady-state MSD, to the existing variants of</p>			

the D-LMS algorithms, and in larger networks choosing an appropriate number of CP iterations in LCP-LMS is shown to further improve performance.

Chapter 4 proposes a variant of sparse diffusion LMS (SD-LMS) algorithm via combination of the SD-LMS with the loopy CP, similarly to LCP-LMS in Chapter 3. Under the condition that the combination step consists of only one iteration of the loopy CP, convergence analysis and computational complexity of the proposed algorithm are presented. Optimization of the combination coefficients is also discussed on the basis of an upper bound of the steady-state MSD. Efficiency of the proposed algorithm in estimating sparse vectors is confirmed via numerical simulations.

Chapter 5 discusses scalable approximation of GPR. It focuses on a recently proposed aggregation method named nested pointwise aggregation of experts (NPAE), and provides its reinterpretation in terms of sketching. On the basis of the sketching-based reinterpretation, it proposes an extended framework of NPAE, called nested aggregation of experts using inducing points (NAE-IP), which offers a more flexible trade off between accuracy and computational cost. Efficiency of the proposed method is investigated via numerical simulations.

Chapter 6 concludes this thesis by summarizing the contributions as well as discussing possible future directions of the research.

(論文審査の結果の要旨)

本論文は、ネットワークを介して計測、収集されるデータに対する分散処理手法のなかで、特に代表的なネットワーク上の分散アルゴリズムである拡散LMS (D-LMS) アルゴリズムと、非線形回帰を行う代表的な手法であるガウス過程回帰 (GPR) の近似的分散処理アルゴリズムとを取り上げ、それぞれのアルゴリズムの改良に関する研究を行った成果を取りまとめたものである。得られた主要な成果は以下の通りである。

1. D-LMSアルゴリズムにおいて近傍ノードの情報を統合する処理に注目し、ネットワーク上の平均合意形成との関連性にもとづいて、代表的なメッセージ伝搬による平均合意アルゴリズムである合意伝搬 (CP) をD-LMSアルゴリズムに組み合わせることにより新たな分散LMSアルゴリズムを提案している。第一の提案はネットワークから抽出された全域木上で動作するアルゴリズムで、反復ごとに完全な合意形成を行う。第二の提案はもとのネットワーク上で動作するアルゴリズムで、CPの有限回反復により近傍ノードの情報統合を行う。後者は、全域木の抽出が不要であり、アルゴリズムの効率もよいと期待される一方で、ネットワークがサイクルを含む場合にはアルゴリズムの振る舞いが明らかでない。本研究では、第二の提案においてCPの反復回数を1回に制限したものについて詳しく解析し、CPのパラメータの最適値を定常MSDの上界の最適化により決定できることを示している。また、定常MSDと収束速度とのトレードオフ関係に着目した場合の二つの提案の有効性を、数値実験によって示している。さらに、第二の提案においてCPの反復回数を適切に選択することにより、大規模なネットワークではさらに性能改善が期待できることを実験的に示している。

2. 推定対象となるベクトルのスパースさを活用する分散LMSアルゴリズムとして、スパース拡散LMS (SD-LMS) アルゴリズムが提案されている。本研究では、上記と同様にCPをSD-LMSアルゴリズムに組み合わせることにより、推定対象ベクトルのスパースさを活用する新たなSD-LMSアルゴリズムを提案している。提案手法においてCPの反復回数を1回に制限したものについて詳しく解析し、CPのパラメータの最適値を定常MSDの上界の最適化により決定できることを示している。定常MSDと収束速度とのトレードオフ関係に着目した場合の提案手法の有効性を、数値実験によって示している。

3. GPRは、非線形回帰を容易に行うことのできる手法として注目を集めているが、大規模なデータを扱う際にはその計算量の多さが問題とされ、計算量を低減させられる様々な近似手法が提案されている。それらのなかで、本研究では最近提案された nested pointwise aggregation of experts (NPAE) と呼ばれる近似的分散処理手法を取り上げている。データ解析を容易にするために高次元データの低次元表現を活用する「スケッチング」の概念にもとづいてNPAEを拡張できることを示し、誘導点を含めてスケッチング行列の設計を行えるより一般的な手法として nested aggregation of experts using inducing points (NAE-IP) を提案している。ある条件下で、提案手法が一致性を有することを示すとともに、誘導点のいくつかの選び方に関して提案手法と既存手法との性能比較を人工データ、実データにもとづいて行い、提案手法の有効性を示している。

以上のように、本研究は、D-LMSアルゴリズムおよびGPRに対する近似アルゴリズムに着目し、それぞれに対して新たな着眼点から新規な手法を提案した結果をまとめ

たものであり、情報学分野における高度な学術を含み、当該研究分野の今後の発展に大きく寄与しうる内容を含むものと認められる。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、令和3年10月14日に行った論文内容とそれに関連した事項についての試問の結果、合格と認めた。また、本論文は、京都大学学位規程第14条第2項に該当するものと判断し、公表に際しては、当該論文の第5章を除く部分については全文公表とするが、第5章については全文に代えてその内容を要約したものとすることを認める。

要旨公開可能日： _____ 年 _____ 月 _____ 日以降