

(続紙 1)

京都大学	博士 (情報学)	氏名	TANKARIA HARDIK
論文題目	The Utilization of Second-order Information for Large-Scale Unconstrained Optimization Problems (大規模な制約なし最適化問題における2次の情報の活用)		
(論文内容の要旨)			
<p>制約なし最適化問題は数理最適化の基礎的な問題の一つである。特に大規模な問題は機械学習や信号処理などに現れ、その効率的な解法の開発が活発になされている。大規模な問題には、一般には、目的関数値とその1次の微分情報のみを用いた最急降下法(SD)やその確率版である確率勾配降下法(SGD)が用いられる。しかし、1次の情報だけでは目的関数を十分に近似できず、ニュートン法などの目的関数の2次の微分情報を用いた手法と比べて速い収束は期待できない。一方、大規模な問題では、ヘッセ行列やその逆行列の計算コストが高いためニュートン法は利用できない。</p> <p>本論文の目的は、SDやSGDの高速化のために、少ない計算コストで2次の微分情報を活用する方法を提案し、その理論的な収束性を解明し、さらに数値実験によってその妥当性を確認することである。本論文は、以下の6章から成っている。</p> <p>第1章は序論であり、大規模な制約なし最適化問題に対する代表的なアルゴリズムとその性質を紹介し、論文全体の構成を概説している。第2章では、本論文で用いる数学的事項をまとめている。</p> <p>第3章では、正則化記憶制限付きBFGS法(RL-BFGS法)を考え、その実装上の工夫をいくつか提案している。RL-BFGS法は2次の情報を利用したL-BFGS法と大域的収束性を保証する信頼領域法を組み合わせた手法であり、理論的には優れた性質をもつ。しかしながら、性質が悪い問題に対しては、各反復での更新量が小さくなり、収束が遅くなるという欠点があった。そのような欠点を克服するために、更新量を大きくするウルフのルールに基づく直接探索を組み合わせた手法を提案している。数値実験によって、提案手法は、既存のL-BFGS法やRL-BFGS法よりも、安定的に解を求めることを確認している。この工夫はL-BFGS法だけでなく、信頼領域法に基づく多くの最適化手法でも適用できる。</p> <p>第4章では、SGDにおける2次の情報の活用法を提案している。SGDでは探索方向を確率的に決定する。探索方向の期待値は目的関数の勾配となるが、その実現値は勾配と一致しない。そこで、実現値と真の勾配との差が小さくなるように、分散減少確率勾配降下法による探索方向に2次の情報を活用した探索方向を提案している。まず、2次の情報を活用する一般的な枠組みを与え、それによるSGDが一次収束するための十分条件を与えている。次に、具体的な手法として Barzilai-Borwein法による2次の情報を組み込む方法を提案し、それが1次収束の十分条件を満たすことを示している。さらに、数値実験によって提案手法の有効性を確認している。</p> <p>第5章では、機械学習などに現れるヘッセ行列の部分行列を効率よく計算できる問題に対して、近似ヘッセ行列をNystrom近似によって構成する方法を提案している。Nystrom近似はカーネル行列の近似によく用いられるが、最適化手法に使われることはあまりなかった。これは、Nystrom近似による行列が正定値行列にならないためで</p>			

ある．そこで，ヘッセ行列が正定値とならない非凸な問題においては，部分行列の一部を正則化することによってNystrom近似行列を半正定値とする手法を提案している．さらに半正定値とした近似行列に対して正定値対角行列を加えることで，少ない計算量で正定値となる近似ヘッセ行列およびその逆行列を構成する手法を与えている．この近似ヘッセ行列を用いた降下法を提案し，その大域的収束性を示している．さらにデータ解析に現れる応用問題に対して，提案手法を適用し，その有効性を確認している．

第6章は結論であり，本論文のまとめと今後の課題を述べている．

(続紙 2)

(論文審査の結果の要旨)

本論文は、大規模な制約なし最適化問題に対する最急降下法や確率的勾配降下法(SGD)の高速化のために、少ない計算コストで目的関数の2次の微分情報を活用する方法を提案し、その理論的な収束性を解明し、さらに数値実験によってその妥当性を示しており、得られた結果は以下のとおりである。

1. 正則化記憶制限付きBFGS法(RL-BFGS法)を考え、その実装上の工夫をいくつか提案している。RL-BFGS法は2次の情報を利用した記憶制限付きBFGS法(L-BFGS法)と大域的収束性を保証する信頼領域法を組み合わせた手法である。性質が悪い問題に対しては、各反復での更新量が小さくなり、収束が遅くなるという欠点があった。その欠点を克服するために、更新量を大きくするウルフのルールに基づく直接探索を組み合わせた手法を提案している。大域的収束することを理論的に示すとともに、数値実験によって実際に安定的に解が求まることを確認している。この工夫はL-BFGS法だけでなく、信頼領域法に基づく多くの最適化手法でも適用できる。
2. SGDにおける2次の情報の活用法を提案している。SGDの探索方向の期待値は目的関数の勾配となるが、その実現値は勾配と一致しない。そこで、実現値と真の勾配との差が小さくなるように、分散減少確率勾配降下法による探索方向に2次の情報を活用した探索方向を提案している。まず、2次の情報を活用する一般的な枠組みを与え、それによるSGDが一次収束するための十分条件を与えている。次に、具体的な手法として Barzilai-Borwein法による2次の情報を組み込む方法を提案し、それが1次収束のための十分条件を満たすことを示している。さらに、数値実験によって提案手法の有効性を確認している。
3. ヘッセ行列の部分行列を効率よく計算できる問題に対して、Nystrom近似を用いて近似ヘッセ行列を構成する方法を提案している。Nystrom近似は最適化手法に使われることはあまりなかった。これはNystrom近似による行列が正定値にならないためである。そこで、ヘッセ行列が正定値とならない非凸な問題においては、部分行列の一部を正則化することによってNystrom近似による行列を半正定値とする手法を提案している。さらに半正定値とした近似行列に対して正定値対角行列を加えることで、少ない計算量で正定値行列となる近似ヘッセ行列およびその逆行列を構成する手法を与えている。この近似ヘッセ行列を用いた降下法を提案し、その大域的収束性を示している。さらにデータ解析に現れる応用問題に対して、提案手法を適用し、その有効性を確認している。

以上のように、本論文では、大規模な制約なし最適化問題に対する最急降下法や確率的勾配降下法の高速化のために、目的関数の2次の微分情報を少ない計算コストを提案し、その性質を数値実験によって検証しており、得られた成果は学術上および応用上極めて優れている。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6年2月27日に実施した論文内容とそれに関連する内容についての試問の結果、合格と認めた。また、令和7年1月31日以降の本論文のイ

インターネットでの全文公表についても支障がないことを確認した。