日本オペレーションズ・リサーチ学会
2022年 秋季研究発表会

Levenberg–Marquardt 法のリーマン多様体への拡張と大域及び局所的収束性解析

東京大学大学院 情報理工学系研究科 *足立 勝 ADACHI Sho
02006704 成蹊大学/理化学研究所 奥野 貴之 OKUNO Takayuki
01308490 東京大学/理化学研究所 武田 朗子 TAKEDA Akiko

1. はじめに

従来の連続最適化手法はユークリッド空間上を中心として研究が進められてきたが, 近年ではユークリッド空間を一般化した空間のクラスであるリーマン多様体上の最適化の研究が注目されている. 本研究では, リーマン多様体上の非線形最小二乗問題を解くために, ユークリッド空間上でよく知られた手法であるLevenberg-Marquardt（LM）法を拡張する. リーマン多様体上のLM法は [2] で枠組みが既に提案されているものの, 理論的収束性については考察されていない. 本研究の貢献は, 次の理論保証付きのリーマン多様体上のLM法を具体的に与えたことである: (i) 大域的収束性及び反復計算量の導出, (ii) 零残差問題に対する局所的な2次収束性及び非零残差問題に対する局所的な1次収束性. 更に, 数値実験を通して, 十分残差が小さいときに提案手法が局所的に早い収束を実現することを実証した.

2. 本研究の対象となる最適化問題

$(\mathcal{M}, \langle\cdot,\cdot\rangle)$ を n 次元完備リーマン多様体として, 以下の最適化問題を考える:

$$\min_{x\in\mathcal{M}} \quad f(x) := \frac{1}{2}\|F(x)\|^2 (= \frac{1}{2}\sum_{i=1}^{m} F_i(x)^2). \quad (1)$$

ここで $F_i\colon \mathcal{M} \to \mathbb{R}$ $(i = 1,\dots,m)$ は C^1-級で $F:\mathcal{M}\to\mathbb{R}^m$ は $F = (F_1, F_2, \dots, F_m)^T$ により定められる. (1) はユークリッド空間上の非線形二乗問題をリーマン多様体上に拡張したものである. この問題は, テンソル分解, 低ランク行列補完, 測地線回帰など多くの応用例をもつ.

3. 提案アルゴリズム

まず, $\mathcal{M} = \mathbb{R}^n$ の場合の (1) に対する代表的手法であるLM法の概略を述べる. LM法では暫定解 x_k と正値パラメータ λ_k が与えられたときに, 以下の部分問題を解くことで次の解 x_{k+1} を求める. (2) において, $J(x_k) \in \mathbb{R}^{m\times n}$ は x_k における F のヤコビ行列である.

$$\begin{aligned}\min_{x\in\mathbb{R}^n} \quad & \frac{1}{2}\|F(x_k) + J(x_k)(x - x_k)\|^2 \\ & + \frac{1}{2}\lambda_k\|x - x_k\|^2\end{aligned} \quad (2)$$

上式の第一項は F の線形近似により目的関数 f を近似したもので, 第二項は λ_k によりステップ長を調節する機能を持っている. この λ_k は減衰パラメータと呼ばれ, ユークリッド空間上のLM法では減衰パラメータの更新手法とそのときの理論保証などについて様々な研究が行われてきた.

$\mathcal{M}$ が一般のリーマン多様体の場合に解くべき部分問題を考える. λ_k は所与とし, 与え方は後述する. まず, (2) のように $x - x_k$ を考えることはできないので, x_k における接空間 $T_{x_k}\mathcal{M}$ の接ベクトルに関する最適化として定式化することになる. 点 x_k における F のヤコビ作用素を $J_k\colon T_{x_k}\mathcal{M} \to \mathbb{R}^n$ と書くと, (2) に対応する問題は以下のようになる.

$$\min_{s\in T_{x_k}\mathcal{M}} \quad \theta^k(s) := \|F(x_k) + J_k s\|^2 + \lambda_k\|s\|_{x_k}^2 \quad (3)$$

ここで $\|\cdot\|_{x_k} := \sqrt{\langle\cdot,\cdot\rangle_{x_k}}$ である. (3) の解 s_k は, J_k の随伴作用素 J_k^* 及び $T_{x_k}\mathcal{M}$ 上の恒等作用素 I_k を用いると, 以下の線形方程式の解である.

$$({J_k}^* J_k + \lambda_k I_k)s_k \quad = \quad -\mathrm{grad} f(x_k) \quad (4)$$

(4) の解 s_k が得られれば, 通常のリーマン多様体上の最適化と同様, レトラクション $R\colon T\mathcal{M} \to \mathcal{M}$ で, $x_{k+1} := R_{x_k}(s_k)$ と点を更新する.

次に減衰パラメータ λ_k の更新手法について述べる. ユークリッド空間におけるLM法では, 信頼領域法における信頼半径の更新方法に類似した手法がとられており (例えば [1]), 本研究でも類似した

手法を採用する．詳細は下記記載のアルゴリズムの通りである．

Algorithm 1 RLM method

Input: $x_0 \in \mathcal{M}, \eta \in (0,1), \mu_{\min} > 0, \beta > 1$

Output: stationary point of (1)

$\mu_0 \leftarrow \mu_{\min}$, $\bar{\mu} \leftarrow \mu_0$
$k \leftarrow 0$
while not convergence **do**
  compute $F(x_k), J_k$
  compute s_k by solving (4) with $\lambda_k = \mu_k \|F(x_k)\|^2$.
  compute $\rho_k := \frac{f(x_k) - f(R_{x_k}(s_k))}{\frac{1}{2}(\theta^k(0_{x_k}) - \theta^k(s_k))}$
  if $\rho_k \geq \eta$ **then**
    $x_{k+1} \leftarrow R_{x_k}(s_k)$, $\mu_{k+1} \in \left[\max\left(\mu_{\min}, \frac{\bar{\mu}}{\beta}\right), \bar{\mu}\right]$, $\bar{\mu} = \mu_k$
  else
    $x_{k+1} \leftarrow x_k$, $\mu_{k+1} \leftarrow \beta \mu_k$
  end if
  $k \leftarrow k+1$
end while
return x_k

4. 大域収束性及び反復計算量

以下が大域収束性に関する主結果である．

仮定 4.1 F のヤコビ作用素 $J : T\mathcal{M} \to \mathbb{R}^m$ とその随伴 J^* は $\mathcal{L}(x_0) := \{x \in \mathcal{M} \mid f(x) \leq f(x_0)\}$ 上で有界とする．即ち，ある定数 $M > 0$ が存在し，$\max\{\|J(x)\|, \|J(x)^*\|\} \leq M$ が任意の $x \in \mathcal{L}(x_0)$ に対して成立する．

定理 4.2 仮定 *4.1* の下で $\liminf_{k \to \infty} \|\mathrm{grad} f(x_k)\| = 0$.

系 4.3 仮定 *4.1* が成立し，$\{\mu_k\}_k$ が有界ならば $\lim_{k \to \infty} \|\mathrm{grad} f(x_k)\| = 0$.

次に反復計算量の解析について．仮定 4.1 に加えて以下の仮定を置く．

仮定 4.4 $\mathcal{L}(x_0)$ 上で $\mathrm{grad} f$ が L リプシッツ連続である．

定理 4.5 仮定 *4.1,4.4* の下で，*Algorithm 1* が $\|\mathrm{grad} f(x)\| < \epsilon$ または $f(x) < \epsilon$ を満たす解を得るのに必要な反復数は $O\left(\log(\epsilon^{-1})\epsilon^{-3}\right)$.

5. 局所収束性

まずいくつかの記法を準備する．X^* を残差が一定の (局所) 最適解の集合とする．即ち，f^* を定数として $X^* := \{x \in \mathcal{M} \mid \mathrm{grad} f(x) = 0, f(x) = f^*\}$. また $\mathrm{dist}(x,y)$ を $x, y \in \mathcal{M}$ 間のリーマン距離として，$B(x,b) := \{y \in \mathcal{M} \mid \mathrm{dist}(x,y) \leq b\}$ とする．最後に，$x \in \mathcal{M}$ に対し，$\bar{x} \in X^*$ を $\mathrm{dist}(x, \bar{x}) = \inf_{y \in X^*} \mathrm{dist}(x,y)$ なる点とする．

ユークリッド空間上の LM 法の局所収束性の解析では，[3] にある局所エラーバウンドによる解析が重要である．本研究でも，そのような仮定を置いて解析を進めた．具体的には以下の通りである．

仮定 5.1 ある $x^* \in X^*$ で次の条件を満たすものが存在する*:*

(a) ある $b \in (0,\infty)$ と $c_1 \in (0,\infty)$ が存在し，$\|J(y)R_y^{-1}(x) - (F(x) - F(y))\| \leq c_1 \|R_y^{-1}(x)\|^2$ が任意の $x, y \in B(x^*, b)$ に対し成立する．

(b) ある $c_2 > 0$ が存在し，$c_2 \|R_x^{-1}(\bar{x})\| \leq \|F(x) - F(\bar{x})\|$ が任意の $x \in B(x^*, b)$ に対し成立する．

ユークリッド空間上の LM 法では，$f^* = 0$ のときに局所的に X^* に 2 次収束し，$f^* > 0$ のときに局所的に X^* に 1 次収束することが知られている．提案アルゴリズムにおいても，同様の結果が仮定 5.1 といくつかの条件下で得られることを示した．

6. 数値実験

本研究では提案アルゴリズムを (1) テンソル分解と (2) 低ランク行列補完の 2 種類の問題に適用し，性能評価を行った．数値実験結果については当日報告する．

参考文献

[1] E. H. Bergou, Y. Diouane, and V. Kungurtsev : *Convergence and Complexity Analysis of a Levenberg–Marquardt Algorithm for Inverse Problems* : Journal of Optimization Theory and Applications volume 185, pp. 927–944 , 2020.

[2] R. L. M. Peeters, *On a Riemannian version of the Levenberg-Marquardt algorithm*, `https://EconPapers.repec.org/RePEc:vua:wpaper:1993-11`, 1993.

[3] N. Yamashita and M. Fukushima : *On the rate of convergence of the Levenberg– Marquardt method* : In G. Alefeld and X. Chen, editors, Topics in Numerical Analysis, Vienna, Springer Vienna, pp. 239–249, 2001.