Applied Mathematics and Computation 249 (2014) 453–467

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Extending the applicability of Gauss–Newton method for convex composite optimization on Riemannian manifolds

Ioannis K. Argyros [a], Ángel Alberto Magreñán [b,*]

[a] Cameron University, Department of Mathematics Sciences, Lawton, OK 73505, USA
[b] Universidad de La Rioja, Departamento de Matemáticas y Computación, 26002 Logroño, La Rioja, Spain

ARTICLE INFO

Keywords:
Gauss–Newton method
Riemannian manifold
Convex composite optimization
Semi-local convergence
Quasi-regularity

ABSTRACT

We present a semi-local convergence analysis of the Gauss–Newton method for solving convex composite optimization problems in Riemannian manifolds using the notion of quasi-regularity for an initial point. Using a combination the L-average Lipszhitz condition and the center L_0-average Lipschitz condition we introduce majorizing sequences for the Gauss–Newton method that are more precise than in earlier studies. Consequently, our semi-local convergence analysis for the Gauss–Newton method has the following advantages under the same computational cost: weaker sufficient convergence conditions; more precise estimates on the distances involved and an at least as precise information on the location of the solution.

1. Introduction

In this study we are concerned with the convex composite optimization problem on a Riemannian manifold. Applications can be found in computer vision [14], machine learning [15], mathematical programming problems such as convex inclusion, minimax problems, penalization methods, goal programming and constrained optimization (see, e.g., [3,8,11–21,24]). These problems can be formulated like composite optimization problems.

Recently, in the elegant study by Wang, Yao and Li in [23], the notion of quasi-regularity for an initial point $x_0 \in \mathbb{R}^l$ with respect to inclusion problem was used to present a semilocal convergence analysis of the Gauss–Newton method on Riemannian manifolds. This notion generalizes the case of regularity studied in the seminal paper by Burke and Ferris (see, e.g.,[9]) as well as the case when $d \longrightarrow DF(x_0)d - \mathcal{C}$ is surjective. The regularity condition was inaugurated by Robinson in [17] (see also, e.g., [18,21,22]).

In this paper, motivated by the works in [10,12,23] and optimization considerations, we present a convergence analysis of Gauss–Newton method on Riemannian manifolds (defined by Algorithm (GNAR) in Section 2). In [23], the convergence of (GNAR) is based on the L-average Lipschitz conditions inaugurated by Wang [22] (to be precised in Section 2). Using a combination of L-average Lipszhitz condition and a center L_0-average Lipschitz condition which is a special case of the L-average Lipschitz condition and a more precise condition to use than the L-average Lipschitz for the computation of the upper bounds on the norm of the inverses involved we presented. In the present paper a finer convergence analysis is given, with the advantages ($\boldsymbol{\mathcal{A}}$): tighter error estimates on the distances involved and the information on the location of the solution is

* Corresponding author.
E-mail addresses: iargyros@cameron.edu (I.K. Argyros), alberto.magrenan@gmail.com (Á.A Magreñán).

http://dx.doi.org/10.1016/j.amc.2014.09.119

at least as precise. These advantages were obtained (under the same computational cost) using the same or weaker hypotheses.

The study is organized as follows: Section 2 contains the notions of generalized Lipschitz conditions and the majorizing sequences for (GNAR). Semilocal convergence analysis of (GNAR) using L-average conditions is presented in Section 3. In Section 4, numerical example to illustrate our theoretical results and favorable comparisons to earlier studies (see, e.g., [12,13,23]) are presented.

2. Generalized Lipschitz conditions and majorizing sequences

2.1. Gauss–Newton algorithms

The purpose of this paper is to study the convex composite optimization problem

$$\min_{x\in\mathbb{R}^l} f(x) := h(F(x)), \tag{2.1}$$

where $h:\ \mathbb{R}^m \longrightarrow \mathbb{R}$ is convex, $F:\ \mathbb{R}^l \longrightarrow \mathbb{R}^m$ is Fréchet-differentiable operator and $m, l \in \mathbb{N}^\star$. The study of (2.1) is very important. On the one hand the study of (2.1) provides a unified framework for the development and analysis of algorithmic method and on the other hand it is a powerful tool for the study of first and second-order optimality conditions in constrained optimization (see, e.g., [3,7,11–21,24]). We assume that the minimum h_{min} of the function h is attained. Problem (2.1) is related to

$$F(x) \in \mathcal{C}, \tag{2.2}$$

where

$$\mathcal{C} = \operatorname{argmin} h \tag{2.3}$$

is the set of all minimum points of h.

A semilocal convergence analysis for Gauss–Newton method (GNA) was presented using the popular algorithm (see, e.g., [5,13]):

Algorithm (GNA) : (ξ, Δ, x_0)

Let $\xi \in [1, \infty[, \Delta \in]0, \infty]$ and for each $x \in \mathbb{R}^l$, define $\mathcal{D}_\Delta(x)$ by

$$\begin{aligned}\mathcal{D}_\Delta(x) \quad = \quad & \{d \in \mathbb{R}^l : \|d\| \leqslant \Delta,\ h(F(x) + DF(x)\,d) \leqslant h(F(x) + DF(x)\,d') \\ & \text{for all } d' \in \mathbb{R}^l \text{ with } \|d'\| \leqslant \Delta\}.\end{aligned} \tag{2.4}$$

Let also $x_0 \in \mathbb{R}^l$ be given. Having $x_0, x_1, \ldots, x_k$ $(k \geqslant 0)$, determine x_{k+1} by:

If $0 \in \mathcal{D}_\Delta(x_k)$, then STOP;

If $0 \notin \mathcal{D}_\Delta(x_k)$, choose d_k such that $d_k \in \mathcal{D}_\Delta(x_k)$ and

$$\|d_k\| \leqslant \xi\, d(0, \mathcal{D}_\Delta(x_k)). \tag{2.5}$$

Then, set $x_{k+1} = x_k + d_k$.

Here, $d(x, W)$ denotes the distance from x to W in the finite dimensional Banach space containing W. Note that the set $\mathcal{D}_\Delta(x)$ $(x \in \mathbb{R}^l)$ is nonempty and is the solution of the following convex optimization problem

$$\min_{d\in\mathbb{R}^l,\ \|d\|\leqslant\Delta} h(F(x) + DF(x)\,d), \tag{2.6}$$

which can be solved by well known methods such as the subgradient or cutting plane or bundle methods (see, e.g., [18]). Many popular iterative methods such as trust region method, conjugate gradient method, steepest descent method, Newton-like methods, has been extended from a Banach space to a Riemannian manifold setting. Recently, Wang et al. [23] extended the Gauss–Newton method (GNA) to Riemmannian manifold to solve the convex composite optimization on Riemannian manifold $\mathcal{M}$ which is formulated as follows:

$$\min_{p\in\mathcal{M}} f(p) := h(F(p)), \tag{2.7}$$

where h is same as defined above and F is a differentiable mapping from $\mathcal{M}$ to $\mathbb{R}^l$. As mentioned before, the study of (2.7) naturally relates to the convex inclusion problem

$$F(p) \in C, \tag{2.8}$$

where $C = \operatorname{argmin} h$, the set of all minimum points of h. The extended Gauss–Newton method for convex composite optimization problem on Riemannian manifold (2.7) (GNAR) is defined as follows.