Journal of Complexity 27 (2011) 111–125

Contents lists available at ScienceDirect

Journal of Complexity

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

Local convergence analysis of the Gauss–Newton method under a majorant condition

O.P. Ferreira [a], M.L.N. Gonçalves [b,*], P.R. Oliveira [b]

[a] IME/UFG, Campus II- Caixa Postal 131, CEP 74001-970 - Goiânia, GO, Brazil

[b] COPPE-Sistemas, Universidade Federal do Rio de Janeiro, 21945-970 Rio de Janeiro, RJ, Brazil

ARTICLE INFO

Article history:
Received 2 April 2010
Accepted 8 September 2010
Available online 18 September 2010

Keywords:
Nonlinear least squares problems
Gauss–Newton method
Majorant condition
Local convergence

ABSTRACT

The Gauss–Newton method for solving nonlinear least squares problems is studied in this paper. Under the hypothesis that the derivative of the function associated with the least square problem satisfies a majorant condition, a local convergence analysis is presented. This analysis allows us to obtain the optimal convergence radius and the biggest range for the uniqueness of stationary point, and to unify two previous and unrelated results.

1. Introduction

We consider the *nonlinear least squares* problem

$$\min_{x\in\Omega} \|F(x)\|^2, \tag{1}$$

where $\Omega \subseteq \mathbb{X}$ is an open set and $F : \Omega \to \mathbb{Y}$ is a continuously differentiable nonlinear function, with $\mathbb{X}$ and $\mathbb{Y}$ being real or complex Hilbert spaces. The interest in this problems arises in data fitting, when $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{Y} = \mathbb{R}^m$, where m is the number of observations and n is the number of parameters; see [11,19].

Denote by $F'(x)$ the derivative of F at a point $x \in \Omega$. When the derivative $F'(x)$ is injective, the problem of finding a stationary point of problem (1), that is, a solution of the nonlinear equation

$$F'(x)^*F(x) = 0,$$

* Corresponding author.
E-mail addresses: orizon@mat.ufg.br (O.P. Ferreira), maxlng@cos.ufrj.br (M.L.N. Gonçalves), poliveir@cos.ufrj.br (P.R. Oliveira).

0885-064X/$ – see front matter
doi:10.1016/j.jco.2010.09.001

where A^* denotes the adjoint of the operator A, is equivalent to finding a least squares solution of the overdetermined nonlinear equation

$$F(x) = 0. \tag{2}$$

This problem has been extensively studied by Dedieu and Kim [8] and Dedieu and Shub [10] for analytic functions, and in the Riemannian context by Adler et al. [1].

When $F'(x)$ is injective and has a closed image for all $x \in \Omega$, the Gauss–Newton method finds stationary points of the above problem. Formally, the Gauss–Newton method is described as follows. Given an initial point $x_0 \in \Omega$, define

$$x_{k+1} = x_k - [F'(x_k)^* F'(x_k)]^{-1} F'(x_k)^* F(x_k), \quad k = 0, 1, \ldots.$$

If the above method converges to $x_* \in \Omega$, then x_* is a stationary point of problem (1), but we cannot conclude that x_* is a solution of (1) or $F(x_*) = 0$. In order to ensure that a stationary point x_* is a solution of (1), we have to apply optimality conditions. It is worth pointing that, if $F'(x)$ is invertible for all $x \in \Omega$, then the Gauss–Newton method becomes the Newton method. Early works dealing with the convergence of the Newton and Gauss–Newton methods include [1–3,5–10,12–14,17,18,20,21,24].

It is well known that the Gauss–Newton method may fail or even fail to be well defined; see the example on p. 225 of [11] and the example in [10]. To ensure that the method is well defined and converges to a stationary point of (1), some conditions must be imposed. For instance, classical convergence analysis (see [11,19]) requires that F' satisfies the Lipschitz condition and that the initial iterate is 'close enough' to the solution, but it cannot make us clearly see how big the convergence radius of the ball is. For the analytic function, Dedieu and Shub [10] have given an estimate of the convergence radius and a criterion for convergence of the Gauss–Newton method.

Our aim in this paper is to present a new local convergence analysis for the Gauss–Newton method under a majorant condition as introduced by Kantorovich [16], and successfully used by Ferreira [12], Ferreira and Gonçalves [13] and Ferreira and Svaiter [14] for studying the Newton method. In our analysis, the classical Lipschitz condition is relaxed using a majorant function. It is worth pointing out that this condition is equivalent to Wang's condition as introduced in [24] and used by Chen, Li [6,7] and Li et al. [17,18] for studying the Gauss–Newton and Newton methods. The convergence analysis presented provides a clear relationship between the majorant function, which relaxes the Lipschitz continuity of the derivative, and the function associated with the nonlinear least square problem (see, for example, Lemmas 13–15). Besides, the results presented here have made the conditions and the proof of convergence simpler. They also allow us to obtain the biggest range for the uniqueness of the stationary point and the optimal convergence radius for the method as regards the majorant function. Moreover, two previous and unrelated results pertaining to the Gauss–Newton method are unified: namely, the result for analytic functions the appeared in Dedieu and Shub [10] and the classical one for functions with Lipschitz derivative (see, for example, [11] and [19]).

The organization of the paper is as follows. In Section 1.1, we list some notation and basic results used in our presentation. In Section 2, the main result is stated, and in Section 2.1, some properties involving the majorant function are established. In Section 2.2, we present the relationships between the majorant function and the nonlinear function F, and in Section 2.3, the optimal ball of convergence and the uniqueness of the stationary point are established. In Section 2.4, the main result is proved, and some applications of this result are given in Section 3.

1.1. Notation and auxiliary results

The following notation and results are used throughout our presentation. Let $\mathbb{X}$ and $\mathbb{Y}$ be Hilbert spaces. The open and closed balls at $a \in \mathbb{X}$ and radius $\delta > 0$ are denoted, respectively, by

$$B(a, \delta) := \{x \in \mathbb{X};\ \|x - a\| < \delta\}, \qquad B[a, \delta] := \{x \in \mathbb{X};\ \|x - a\| \leqslant \delta\}.$$

The set $\Omega \subseteq \mathbb{X}$ is an open set, the function $F : \Omega \to \mathbb{Y}$ is continuously differentiable, and $F'(x)$ has a closed image in Ω.

Let $A : \mathbb{X} \to \mathbb{Y}$ be a continuous and injective linear operator with closed image. The Moore–Penrose inverse $A^\dagger : \mathbb{Y} \to \mathbb{X}$ of A is defined by