Journal of Computational and Applied Mathematics 236 (2012) 2487–2498

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

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

Local convergence analysis of inexact Gauss–Newton like methods under majorant condition

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

[a] IME/UFG, Campus II- Caixa Postal 131, 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 19 July 2011

Keywords:
Nonlinear least squares problems
Inexact Gauss–Newton like methods
Majorant condition
Local convergence

ABSTRACT

In this paper, we present a local convergence analysis of inexact Gauss–Newton like methods for solving nonlinear least squares problems. Under the hypothesis that the derivative of the function associated with the least squares problem satisfies a majorant condition, we obtain that the method is well-defined and converges. Our analysis provides a clear relationship between the majorant function and the function associated with the least squares problem. It also allows us to obtain an estimate of convergence ball for inexact Gauss–Newton like methods and some important, special cases.

1. Introduction

Let $\mathbb{X}$ and $\mathbb{Y}$ be real or complex Hilbert spaces. Let $\Omega \subseteq \mathbb{X}$ be an open set, and $F : \Omega \to \mathbb{Y}$ a continuously differentiable nonlinear function. Consider the following *nonlinear least squares* problems

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

The interest in this problem arises in data fitting, when $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{Y} = \mathbb{R}^m$ and m is the number of observations and n is the number of parameters, see for example [1]. A solution $x_* \in \Omega$ of (1) is also called a least-squares solution of nonlinear equation $F(x) = 0$.

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

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

where A^* denotes the adjoint of the operator A. It is worth pointing out that if x_* is a solution of (1), $F(x_*) = 0$ and $F'(x_*)$ is invertible, then the theories of the Gauss–Newton method merge into the theories of Newton method. Early works dealing with the convergence of the Newton and Gauss–Newton methods include [2–18].

The inexact Gauss–Newton process is described as follows: Given an initial point $x_0 \in \Omega$, define

$$x_{k+1} = x_k + S_k, \quad k = 0, 1, \ldots,$$

where $B_k : \mathbb{X} \to \mathbb{Y}$ is a linear operator and S_k is any approximated solution of the linear system

$$B_kS_k = -F'(x_k)^*F(x_k) + r_k,$$

* Corresponding author. Tel.: +55 62 35211208; fax: +55 62 35211180.
E-mail addresses: orizon@mat.ufg.br (O.P. Ferreira), maxlng@cos.ufrj.br, maxlng@mat.ufg.b (M.L.N. Gonçalves), poliveir@cos.ufrj.br (P.R. Oliveira).

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.12.008

for a suitable residual $r_k \in \mathbb{Y}$. In particular, the above process is *inexact Gauss–Newton method* if $B_k = F'(x_k)^T F'(x_k)$, the process is *inexact modified Gauss–Newton method* if $B_k = F'(x_0)^T F'(x_0)$, and it represents *inexact Gauss–Newton like method* if B_k is an approximation of $F'(x_k)^T F'(x_k)$.

For inexact Newton methods, as shown in [19], if $\|r_k\| \le \theta_k \|F(x_k)\|$ for $k = 0, 1, \ldots$ and $\{\theta_k\}$ is a sequence of forcing terms such that $0 \le \theta_k < 1$ then there exists $\epsilon > 0$ such that the sequence $\{x_k\}$, for any initial point $x_0 \in B(x_*, \epsilon) = \{x \in \mathbb{R}^n : \|x_* - x\| < \epsilon\}$, is well defined and converges linearly to x_* in the norm $\|y\|_* = \|F'(x_*)y\|$, where $\| \ \|$ is any norm in $\mathbb{R}^n$. As pointed out by [20] (see also [21]) the result of [19] is difficult to apply due to a dependence of the norm $\| \ \|_*$, which is not computable.

Formally, the inexact Gauss–Newton like methods for solving (1), which we will consider, are described as follows: Given an initial point $x_0 \in \Omega$, define

$$x_{k+1} = x_k + S_k, \qquad B(x_k)S_k = -F'(x_k)^* F(x_k) + r_k, \quad k = 0, 1, \ldots,$$

where $B(x_k)$ is a suitable invertible approximation of the derivative $F'(x_k)^* F'(x_k)$ and the residual tolerance r_k and the preconditioning invertible matrix P_k (considered for the first time in [21]) for the linear system defining the step S_k satisfy

$$\|P_k r_k\| \le \theta_k \|P_k F'(x_k)^* F(x_k)\|,$$

for suitable forcing number θ_k. Note that, if the forcing sequence vanishes, i.e., $\theta_k = 0$ for all k, the inexact Gauss–Newton methods include the class of Gauss–Newton iterative methods. Hence, the theories of inexact Gauss–Newton methods merge into the theories of Gauss–Newton methods.

The classical local convergence analysis for the inexact Newton methods (see [19,21]) requires, among other hypotheses, that F' satisfies the Lipschitz condition. In the last years, there have been papers dealing with the issue of convergence of the Newton method and inexact Newton method, including the Gauss–Newton method and the inexact Gauss–Newton method, by relaxing the assumption of Lipschitz continuity of the derivative (see for example: [5,7,10–12,15,18,22–24]). One of the main conditions that relaxes the condition of the Lipschitz continuity of the derivative is the majorant condition, which we will use, and Wang's condition, introduced in [18] and used in [5,6,14,15,22,23] to study the Gauss–Newton and Newton methods. In fact, it can be shown that these conditions are equivalent. But the formulation as a majorant condition is in some sense better than Wang's condition, as it provides a clear relationship between the majorant function and the nonlinear function under consideration. Besides, the majorant condition provides a simpler proof of convergence.

In the present paper, we are interested in the local convergence analysis, i.e., based on the information in a neighborhood of a stationary point of (1) we determine the convergence ball of the method. Following the ideas of [10–12,24], we will present a new local convergence analysis for inexact Gauss–Newton like methods under majorant condition. The convergence analysis presented provides a clear relationship between the majorant function, which relaxes the Lipschitz continuity of the derivate, and the function associated with the nonlinear least squares problem (see for example: Lemmas 12–14). Besides, the results presented here have the conditions and the proof of convergence in quite a simple manner. Moreover, two unrelated previous results pertaining to inexact Gauss–Newton like methods are unified, namely, the result for analytical functions and the classical one for functions with Lipschitz derivative.

The organization of the paper is as follows. In Section 1.1, we list some notations 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 non-linear function F. In Section 2.3 the main result is proven and some applications of this result are given in Section 3. Some final remarks are offered in Section 4.

1.1. Notation and auxiliary results

The following notations and results are used throughout our presentation. Let $\mathbb{X}$ and $\mathbb{Y}$ be Hilbert spaces. The open and closed ball 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 and 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

$$A^\dagger := (A^* A)^{-1} A^*,$$

where A^* denotes the adjoint of the linear operator A.

Lemma 1 (*Banach's Lemma*)**.** *Let $B : \mathbb{X} \to \mathbb{X}$ be a continuous linear operator, and $I : \mathbb{X} \to \mathbb{X}$ the identity operator. If $\|B - I\| < 1$, then B is invertible and $\|B^{-1}\| \le 1/(1 - \|B - I\|)$.*

Proof. See the proof of Lemma 1, p. 189 of Smale [25] with $A = I$ and $c = \|B - I\|$. □

Lemma 2. *Let $A, B : \mathbb{X} \to \mathbb{Y}$ be a continuous linear operator with closed image. If A is injective, $E = B - A$ and $\|EA^\dagger\| < 1$, then B is injective.*