Computers and Mathematics with Applications 66 (2013) 490–499

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

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

Local convergence of the Gauss–Newton method for injective-overdetermined systems of equations under a majorant condition

M.L.N. Gonçalves *

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

ARTICLE INFO

Article history:
Received 4 May 2012
Received in revised form 11 January 2013
Accepted 27 May 2013

Keywords:
Injective-overdetermined systems of equations
Gauss–Newton method
Majorant condition
Local convergence

ABSTRACT

A local convergence analysis of the Gauss–Newton method for solving injective-overdetermined systems of nonlinear equations under a majorant condition is provided. The convergence as well as results on its rate are established without a convexity hypothesis on the derivative of the majorant function. The optimal convergence radius, the biggest range for uniqueness of the solution along with some other special cases are also obtained.

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 *systems of nonlinear equations*

$$F(x) = 0. \tag{1}$$

If $F'(x)$ is invertible, the Newton method and its variants (see [1–4]) are the most efficient methods known for solving such systems. However, $F'(x)$ may not even be a square matrix. One simple example arises when $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{Y} = \mathbb{R}^m$, with $n \neq m$. In this case, $F'(x)$ is not invertible and (1) becomes an overdetermined system ($n < m$) or an underdetermined system ($n > m$). In general, if $F'(x)$ is injective or surjective, we say (1) is an injective-overdetermined or surjective-underdetermined system of equations, respectively.

If $F'(x)$ is not necessarily invertible, a generalized Newton method called the Gauss–Newton method can be used (see [5,6]). It is defined by

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

where $F'(x_k)^{\dagger}$ denotes the Moore–Penrose inverse of the linear operator $F'(x_k)$. This algorithm finds least-squares solutions of (1). These least-squares solutions, which may or may not be solutions of the original problem (1), are related to the nonlinear least squares problem

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

* Tel.: +55 6235211208; fax: +55 6235211180.
E-mail address: maxlng@mat.ufg.br.

0898-1221/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.camwa.2013.05.019

that is, they are stationary points of $H(x) = \|F(x)\|^2$. This paper is focused on the case in which the least-squares solutions of (1) also solve (1). In the theory of nonlinear least squares problems, this case is called the zero-residual case.

Regarding the local and semi-local convergence analysis of the Newton and Gauss–Newton methods, in the last years there has been much work attempting to alleviate the assumption of Lipschitz continuity on the operator F', see for example [1–15]. The main conditions that relax the Lipschitz continuity on the derivative is the majorant condition, used for example in [1–6], and the generalized Lipschitz condition according to X. Wang, used for example in [7–10,12–15]. In fact, as proved in [2], if the majorant function has convex derivative, these conditions are equivalent. Otherwise, the Wang's condition can be seen as a particular case of the majorant condition. Moreover, the majorant formulation provides a clear relationship between the majorant function and the nonlinear operator under consideration, simplifying the proof of convergence substantially.

Our aim in this paper is to present a new local convergence analysis of the Gauss–Newton method for solving injective-overdetermined systems of equations under a majorant condition. The convergence, uniqueness, superlinear rate and an estimate of the best possible convergence radius will be established without a convexity hypothesis on the derivative of the majorant function, which was assumed in [5]. In addition to the special cases obtained in [5], the lack of convexity of the derivative of the majorant function in this analysis, allows us to obtain two new important special cases, namely, the convergence can be ensured under Hölder-like and generalized Lipschitz conditions. In the latter case, the results are obtained without assuming that the function that defines the condition is nondecreasing, thus generalizing Corollary 6.3 in [8]. Moreover, it is worth to mention that, similar to the convergence analysis of the Newton method (see [2]), the hypothesis of convex derivative of the majorant function or nondecreasing of the function which defines the generalized Lipschitz condition, are needed only to obtain quadratic convergence rate.

The organization of the paper is as follows. In Section 1.1, we list some notations and one basic result used in our presentation. In Section 2, we state the main result and in Section 2.1 some properties of the majorant function are established and the main relationships between the majorant function and the nonlinear function F are presented. The optimal ball of convergence and the uniqueness of the solution are also discussed in Section 2.1. In Section 2.2 our main result is proven and some applications of this result are obtained 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 $a \in \mathbb{X}$ with 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.

Some properties related to the Moore–Penrose inverse will be needed. More details about the Moore–Penrose inverse can be found in [16,17].

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

$$\|B^\dagger\| \le \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|A - B\|}.$$

2. Local analysis for the Gauss–Newton method

Our goal is to state and prove a local theorem for the Gauss–Newton method, which generalizes the Corollary 8 of [5], as well as Theorem 2 of [2]. First, we prove some results regarding the scalar majorant function, which relaxes the Lipschitz condition. Then, we establish the main relationships between the majorant function and the nonlinear function F. We also obtain the optimal ball of convergence and the uniqueness of the solution in a suitable region. Finally, we show well definedness and convergence, along with results on the convergence rates. The statement of the theorem is:

Theorem 2. *Let $\mathbb{X}$ and $\mathbb{Y}$ be Hilbert spaces, $\Omega \subseteq \mathbb{X}$ be an open set and $F : \Omega \to \mathbb{Y}$ be a continuously differentiable function such that F' has a closed image in Ω. Let $x_* \in \Omega$, $R > 0$, $\beta := \|F'(x_*)^\dagger\|$ and $\kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}$. Suppose that $F(x_*) = 0$, $F'(x_*)$ is injective and there exists a $f : [0,\ R) \to \mathbb{R}$ continuously differentiable such that*

$$\beta \left\|F'(x) - F'(x_* + \tau(x - x_*))\right\| \le f'\left(\|x - x_*\|\right) - f'\left(\tau\|x - x_*\|\right), \tag{2}$$