J. Nonlinear Funct. Anal. 2016 (2016), Article ID 1

MATHRES

Journal of Nonlinear Functional Analysis

Available online at http://jnfa.mathres.org

ON THE CONVERGENCE OF INEXACT GAUSS-NEWTON METHOD FOR SOLVING SINGULAR EQUATIONS

IOANNIS K. ARGYROS[1,*], SANTHOSH GEORGE[2]

[1]Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA

[2]Department of Mathematical and Computational Sciences, NIT Karnataka, 575025, India

Abstract. We present a new semi-local convergence analysis of the inexact Gauss-Newton method for solving singular equations. The convergence analysis is based on a combination of a center-majorant, majorant function and restricted convergence domains which are more precise than in the studies using only the majorant function leading to the extension of the applicability of Gauss-Newton method under the same computational cost as in earlier studies such as [5, 7, 12–43]. In particular, the advantages are: the error estimates on the distances involved are tighter and the convergence ball is at least as large. Numerical examples are also provided in this study.

Keywords. Gauss-Newton method; Local convergence; Majorant function; Center-majorant function; Convergence ball.

1. Introduction

A lot of problems such as convex inclusion, minimax problems, penalization methods, goal programming, constrained optimization and other problems can be formulated like

$$F(x) = 0, \tag{1.1}$$

where D is open and convex and $F : D \subset \mathbb{R}^j \to \mathbb{R}^m$ is a nonlinear operator with its Fréchet derivative denoted by F'. The solutions of equation (1.1) can rarely be found in closed form.

*Corresponding author.

E-mail addresses: iargyros@cameron.edu (I.K. Argyros), sgeorge@nitk.ac.in (S. George)

Received December 4, 2015

That is why the solution methods for these equations are usually iterative. In particular, the practice of numerical analysis for finding such solutions is essentially connected to Newton-type methods [2, 6–9, 11, 12, 18, 19, 27, 28, 34, 36, 37]. The study about convergence matter of iterative procedures is usually centered on two types: semilocal and local convergence analysis. The semilocal convergence matter is, based on the information around an initial point, to give criteria ensuring the convergence of iterative procedures; while the local one is, based on the information around a solution, to find estimates of the radii of convergence balls. A plethora of sufficient conditions for the local as well as the semilocal convergence of Newton-type methods as well as an error analysis for such methods can be found in [1–43]. In the case $m = j$, the inexact Newton method was defined in [6] by:

$$x_{n+1} = x_n + s_n, \quad F'(x_n)s_n = -F(x_n) + r_n \quad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{1.2}$$

where x_0 is an initial point, the residual control r_n satisfy

$$\|r_n\| \le \lambda_n \|F(x_n)\| \quad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{1.3}$$

and $\{\lambda_n\}$ is a sequence of forcing terms such that $0 \le \lambda_n < 1$. Let x^* be a solution of (1.1) such that $F'(x^*)$ is invertible. As shown in [6], if $\lambda_n \le \lambda < 1$, then, there exists $r > 0$ such that for any initial guess $x_0 \in U(x^*, r) := \{x \in \mathbb{R}^j : \|x - x^*\| < r\}$, the sequence $\{x_n\}$ is well defined and converges to a solution x^* in the norm $\|y\|_* := \|F'(x^*)y\|$, where $\|\cdot\|$ is any norm in $\mathbb{R}^j$. Moreover, the rate of convergence of $\{x_n\}$ to x^* is characterized by the rate of convergence of $\{\lambda_n\}$ to 0. It is worth noting that, in [6], no Lipschitz condition is assumed on the derivative F' to prove that $\{x_n\}$ is well defined and linearly converging. However, no estimate of the convergence radius r is provided. A pointed out by [15] the result of [6] is difficult to apply due to dependence of the norm $\|\cdot\|_*$, which is not computable.

The residual control (1.3) is non-affine invariant. The advantages of affine versus non-affine invariant forms have been explained in [19]. That is why, Ypma used in [40] the affine invariant condition of residual control in the form:

$$\|F'(x_n)^{-1} r_n\| \le \lambda_n \|F'(x_n)^{-1} F(x_n)\| \quad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{1.4}$$

to study the local convergence of inexact Newton method (1.2). And the radius of convergent result are also obtained.

To study the local convergence of inexact Newton method and inexact Newton-like method (called inexact methods for short below), Morini presented in [31] the following variation for the residual controls:

$$\|P_n r_n\| \leq \lambda_n \|P_n F(x_n)\| \quad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{1.5}$$

where $\{P_n\}$ is a sequence of invertible operator from $\mathbb{R}^j$ to $\mathbb{R}^j$ and $\{\lambda_n\}$ is the forcing term. If $P_n = I$ and $P_n = F'(x_n)$ for each n, (1.5) reduces to (1.3) and (1.4), respectively. These methods are linearly convergent under Lipschitz Condition. It is worth nothing that the residual controls (1.5) are used in iterative methods if preconditioning is applied and lead to a relaxation on the forcing terms. But we also note that the results obtained in [31] do not provide an estimate of the radius of convergence. This is why Chen and Li [15] obtained the local convergence properties of inexact methods for (1.1) under a weak Lipschitz condition, which was first introduced by Wang in [37] to study the local convergence behaviour of Newton's method. The result in [15] easily provides an estimate of convergence ball for the inexact methods. Furthermore, Ferreira and Gonçalves presented in [22] a new local convergence analysis for inexact Newton-like under so-called majorant condition.

Recent attentions are focused on the study of finding zeros of singular nonlinear systems by Gauss-Newton's method, which is defined by

$$x_{n+1} = x_n - F'(x_n)^{\dagger} F(x_n) \quad \text{for each} \quad n = 0, 1, 2, \ldots, \tag{1.6}$$

where $x_0 \in D$ is an initial point and $F'(x_n)^{\dagger}$ denotes the Moore-Penrose inverse of the linear operator (of matrix) $F'(x_n)$. Shub and Smale extended in [35] the Smale point estimate theory (includes α-theory and γ-theory) to Gauss-Newton's methods for underdetermined analytic systems with surjective derivatives. For overdetermined systems, Dedieu and Schub studied in [17] the local linear convergence properties of Gauss-Newton's for analytic systems with injective derivatives and provided estimates of the radius of convergence balls for Gauss-Newton's method. Dedieu and Kim in [16] generalized both the results of the undetermined case and the overdetermined case to such case where $F'(x)$ is of constant rank (not necessary full rank), which has been improved by some authors in [1, 11–14, 19, 20].

Recently, several authors have studied the convergence behaviour of inexact versions of Gauss-Newton's method for singular nonlinear systems. For example, Chen [14] employed the ideas of [37] to study the local convergence properties of several inexact Gauss-Newton type methods where a scaled relative residual control is performed at each iteration under weak Lipschitz conditions. Ferreira, Gonçalves and Oliveira presented in their recent paper [25] a local convergence analysis of an inexact version of Gauss-Newton's method for solving nonlinear least squares problems. Moreover, the radius of the convergence balls under the corresponding conditions were estimated in these two papers. The preceding results were improved by Argyros et al. [2–11] using the concept of the center Lipschitz condition (see also (2.11) and the numerical examples) under the same computational cost on the parameters and functions involved.

In the present study, we are motivated by the elegant work in [41, 42] and optimization considerations. Using more precise majorant condition and functions, and more precise convergence domains, we provide a new local convergence analysis for Gauss-Newton method under the same computational cost and the following advantages: larger radius of convergence; tighter error estimates on the distances $\|x_n - x^*\|$ for each $n = 0, 1, \ldots$ and a clearer relationship between the majorant function (see (2.8) and the associated least squares problems (1.1). These advantages are obtained because we use a center-type majorant condition (see (2.11)) for the computation of inverses involved which is more precise that the majorant condition used in [20–25, 29, 30, 38–42]. Moreover, these advantages are obtained under the same computational cost, since as we will see in Section 3 and Section 4, the computation of the majorant function requires the computation of the center-majorant function. Furthermore, these advantages are very important in computational mathematics, since we have a wider choice of initial guesses x_0 and fewer computations to obtain a desired error tolerance on the distances $\|x_n - x^*\|$ for each $n = 0, 1, 2, \ldots$.

The rest of this study is organized as follows. In Section 2, we introduce some preliminary notions and properties of the majorizing function. The main result about the local convergence are stated in section 3. In Section 4, we prove the local convergence results given in Section 3. Section 5 contains the numerical examples and Section 6 the conclusion of this study.

2. **Preliminaries**

We present some standard results to make the study as selfcontained as possible. More results can be found in [12, 27, 34].

Let $A : \mathbb{R}^j \to \mathbb{R}^m$ be a linear operator (or an $m \times j$ matrix). Recall that an operator (or $j \times m$ matrix) $A^\dagger : \mathbb{R}^m \to \mathbb{R}^j$ is the Moore-Penrose inverse of A if it satisfies the following four equations:

$$A^\dagger A A^\dagger = A^\dagger; \quad AA^\dagger A = A; \quad (AA^\dagger)^* = AA^\dagger; \quad (A^\dagger A) = A^\dagger A,$$

where A^* denotes the adjoint of A. Let $kerA$ and imA denote the kernel and image of A, respectively. For a subspace E of $\mathbb{R}^j$, we use Π_E to denote the projection onto E. Clearly, we have that

$$A^\dagger A = \Pi_{kerA^\perp} \quad \text{and} \quad AA^\dagger = \Pi_{imA}.$$

In particular, in the case when A is full row rank (or equivalently, when A is surjective), $AA^\dagger = I_{\mathbb{R}^m}$; when A is full column rank (or equivalently, when A is injective), $A^\dagger A = I_{\mathbb{R}^j}$.

The following lemma gives a Banach-type perturbation bound for Moore-Penrose inverse, which is stated in [24].

Lemma 2.1. *([24, Corollary 7. 1. 1 & Corollary 7. 1. 2]). Let A and B be $m \times j$ matrices and let $r \leq \min\{m, j\}$. Suppose that $rankA = r$, $1 \leq rankB \leq A$ and $\|A^\dagger\|\|B - A\| < 1$. Then, $rankB = r$ and*

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

Also, we need the following useful lemma about elementary convex analysis.

Lemma 2.2. *([24, Proposition 1.3]). Let $R > 0$. If $\varphi : [0, R] \to \mathbb{R}$ is continuously differentiable and convex, then, the following assertions hold:*

(a) $\dfrac{\varphi(t) - \varphi(\varsigma t)}{t} \leq (1 - \varsigma)\varphi'(t)$ *for each $t \in (0, R)$ and $\varsigma \in [0, 1]$.*

(b) $\dfrac{\varphi(u) - \varphi(\varsigma u)}{u} \leq \dfrac{\varphi(v) - \varphi(\varsigma v)}{v}$ *for each $u, v \in [0, R)$, $u < v$ and $0 \leq \varsigma \leq 1$.*

From now on we suppose that the (I) conditions listed below hold.

For a positive real $R \in \mathbb{R}^+$, let

$$\psi : [0,R] \times [0,1) \times [0,1) \to \mathbb{R}$$

be a continuous differentiable function of three of its arguments and satisfy the following properties:

(i) $\psi(0,\lambda,\theta) = 0$ *and* $\frac{\partial}{\partial t}\psi(t,\lambda,\theta)\Big|_{t=0} = -(1+\lambda+\theta)$.

(ii) $\frac{\partial}{\partial t}\psi(t,\lambda,\theta)$ *is convex and strictly increasing with respect to the argument* t.

For fixed $\lambda, \theta \in [0,1)$, we write $h_{\lambda,\theta}(t) \triangleq \psi(t,\lambda,\theta)$ for short below. Then the above two properties can be restated as follows.

(iii) $h_{\lambda,\theta}(0) = 0$ *and* $h'_{\lambda,\theta}(0) = -(1+\lambda+\theta)$.

(iv) $h'_{\lambda,\theta}(t)$ *is convex and strictly increasing.*

(v) $g : [0,R] \to \mathbb{R}$ *is strictly increasing with* $g(0) = 0$.

(vi) g' *is convex and strictly increasing with* $g'(0) = -1$.

(vii) $g(t) \le h_{\lambda,\theta}(t)$, $g'(t) \le h'_{\lambda,\theta}(t)$ *for each* $t \in [0,R)$, $\lambda,\ \theta \in [0,1]$.

Define

$$\zeta_0 := \sup\{t \in [0,R) : h'_{0,0}(t) < 0\}, \quad \zeta := \sup\{t \in [0,+\infty) : g'(t) < 0\}, \tag{2.1}$$

$$\rho_0 := \sup\left\{t \in [0,\zeta_0) : \left|\frac{h_{\lambda,\theta}(t)}{h'_{0,0}(t)} - t\right| < t\right\}, \quad \rho = \sup\left\{t \in [0,\zeta) : \left|\frac{h_{\lambda,\theta}(t) - th'_{0,0}(t)}{g'(t)}\right| < t\right\} \tag{2.2}$$

$$\sigma := \sup\{t \in [0,R) : U(x^*,t) \subset D\}. \tag{2.3}$$

The next two lemmas show that the constants ζ and ρ defined in (2.1) and (2.2), respectively, are positive.

Lemma 2.3. *The constant ζ defined in (2.1) is positive and $\frac{th'_{0,0}(t) - h_{\lambda,\theta}(t)}{g'(t)} < 0$ for each $t \in (0,\zeta)$.*

Proof. Since $g'(0) = -1$, there exists $\delta > 0$ such that $g'(t) < 0$ for each $t \in (0,\delta)$. Then, we get $\zeta \ge \delta (> 0)$. We must show that $\frac{th'_{0,0}(t) - h_{\lambda,\theta}(t)}{g'(t)} < 0$ for each $t \in (0,\zeta)$. By hypothesis, functions $h'_{\lambda,\theta}(t)$, $g'(t)$ are strictly increasing, then functions $h_{\lambda,\theta}$, $g'(t)$ are strictly convex. It follows from Lemma 2.2 (i) and hypothesis (vii) that

$$\frac{h_{\lambda,\theta}(t)-h_{\lambda,\theta}(0)}{t}<h'_{\lambda,\theta}(t), \quad t\in(0,R).$$

In view of $h_{\lambda,\theta}(0)=0$ and $g'(t)<0$ for all $t\in(0,\zeta)$. This together with the last inequality yields the desired inequality.

Lemma 2.4. *The constant ρ defined in (2.2) is positive. Consequently,* $\left|\frac{th'_{0,0}(t)-h_{\lambda,\theta}(t)}{g'(t)}\right|<t$ *for each* $t\in(0,\rho)$.

Proof. Firstly, by Lemma 2.3, it is clear that $\left(\frac{h_{\lambda,\theta}(t)}{th'_{0,0}(t)}-1\right)\frac{h'_{0,0}(t)}{g'(t)}>0$ for $t\in(0,\zeta)$. Secondly, we get from Lemma 2.2 (i) that

$$\lim_{t\to 0}\left(\frac{h_{\lambda,\theta}(t)}{th'_{0,0}(t)}-1\right)\frac{h'_{0,0}(t)}{g'(t)}=0.$$

Hence, there exists a $\delta>0$ such that

$$0<\left(\frac{h_{\lambda,\theta}(t)}{th'_{0,0}(t)}-1\right)\frac{h'_{0,0}(t)}{g'(t)}<1, \quad t\in(0,\zeta).$$

That is ρ is positive.

Define

$$r:=\min\{\rho,\delta\}, \tag{2.4}$$

where ρ and δ are given in (2.2) and (2.3), respectively. For any starting point $x_0\in U(x^*,r)\backslash\{x^*\}$, let $\{t_n\}$ be a sequence defined by:

$$t_0=\|x_0-x^*\|, \quad t_{n+1}=\left|\left(t_n-\frac{h_{\lambda,\theta}(t_n)}{h'_{0,0}(t_n)}\right)\frac{h'_{0,0}(t_n)}{g'(t_n)}\right| \quad \text{for each} \quad n=0,1,2,\ldots \tag{2.5}$$

Lemma 2.5. *The sequence $\{t_n\}$ given by (2.5) is well defined, strictly decreasing, remains in $(0,\rho)$ for each $n=0,1,2,\ldots$ and converges to* 0.

Proof. Since $0 < t_0 = \|x_0 - x^*\| < r \leq \rho$, using Lemma 2.4, we have that $\{t_n\}$ is well defined, strictly decreasing and remains in $[0, \rho)$ for each $n = 0, 1, 2, \ldots$ Hence, there exists $t^* \in [0, \rho)$ such that $\lim_{n \to +\infty} t_n = t^*$. That is, we have

$$0 \leq t^* = \left(\frac{h_{\lambda,\theta}(t^*)}{h'_{0,0}(t^*)} - t^* \right) \frac{h'_{0,0}(t^*)}{g'(t^*)} < \rho.$$

If $t^* \neq 0$, it follows from Lemma 2.4 that

$$\left(\frac{h_{\lambda,\theta}(t^*)}{h'_{0,0}(t^*)} - t^* \right) \frac{h'_{0,0}(t^*)}{g'(t^*)} < t^*,$$

which is a contradiction. Hence, we conclude that $t_n \to 0$ as $n \to +\infty$.

If $g(t) = h_{\lambda,\theta}(t)$, then Lemmas 2.3–2.5 reduce to the corresponding ones in [41, 42]. Otherwise, i. e., if $g(t) < h_{\lambda,\theta}(t)$, then our results are better, since

$$\zeta_0 < \zeta \quad \text{and} \quad \rho_0 < \rho.$$

Moreover, the scalar sequence used in [41, 42] is defined by

$$u_0 = \|x_0 - x^*\|, \quad u_{n+1} = \left| u_n - \frac{h_{\lambda,\theta}(u_n)}{h'_{0,0}(u_n)} \right| \quad \text{for each} \quad n = 0, 1, 2, \ldots \tag{2.6}$$

Using the properties of the functions $h_{\lambda,\theta}$, g, (2.5), (2.6) and a simple inductive argument we get that

$$t_0 = u_0, \quad t_1 = u_1, \quad t_n < u_n, \quad t_{n+1} - t_n < u_{n+1} - u_n \quad \text{for each} \quad n = 1, 2, \ldots$$

and

$$t^* \leq u^* = \lim_{n \to +\infty} u_n,$$

which justify the advantages of our approach as claimed in the introduction of this study.

In Section 3 we shall show that $\{t_n\}$ is a majorizing sequence for $\{x_n\}$.

We state the following modified majorant condition for the convergence of various Newton-type methods in [9–12].

Definition 2.6. *Let $r > 0$ be such that $U(x^*, r) \subset D$. Then, F' is said to satisfy the majorant condition on $U(x^*, r)$, if*

$$\|F'(x^*)^\dagger [F'(x) - F'(x^* + \varsigma(x - x^*))]\| \leq \bar{h}'_{\lambda,\theta}(\|x - x^*\|) - \bar{h}'_{\lambda,\theta}(\varsigma\|x - x^*\|) \tag{2.7}$$

for any $x \in U(x^*, r)$ *and* $\varsigma \in [0,1]$, *where function* $\bar{h}_{\lambda,\theta}$ *satisfies the same conditions as function* $h_{\lambda,\theta}$.

In the case when $F'(x^*)$ is not surjective, the information on $imF'(x^*)^{\perp}$ may be lost. This is why the above notion was modified in [41, 42] to suit the case, when $F'(x^*)$ is not surjective as follows:

Definition 2.7. *Let* $r > 0$ *be such that* $U(x^*, r) \subset D$. *Then,* F' *is said to satisfy the modified majorant condition on* $U(x^*, r)$, *if*

$$\|F'(x^*)^{\dagger}\|\|F'(x) - F'(x^* + \varsigma(x - x^*))\| \leq \bar{h}'_{\lambda,\theta}(\|x - x^*\|) - \bar{h}'_{\lambda,\theta}(\varsigma\|x - x^*\|) \tag{2.8}$$

for any $x \in U(x^*, r)$ *and* $\varsigma \in [0,1]$,

If $\varsigma = 0$, condition (2.8) reduces to

$$\|F'(x^*)^{\dagger}\|\|F'(x) - F'(x^*)\| \leq \bar{h}'_{\lambda,\theta}(\|x - x^*\|) - \bar{h}'_{\lambda,\theta}(0). \tag{2.9}$$

In particular, for $\lambda = \theta = 0$, condition (2.9) reduces to

$$\|F'(x^*)^{\dagger}\|\|F'(x) - F'(x^*)\| \leq \bar{h}'_{0,0}(\|x - x^*\|) - \bar{h}'_{0,0}(0). \tag{2.10}$$

Condition (2.10) is used to produce the Banach-type perturbation Lemmas in [41, 42] for the computation of the upper bounds on the norms $\|F'(x)^{\dagger}\|$. In this study we use a more flexible function g than function $\bar{h}_{\lambda,\theta}$ for the same purpose. This way the advantages as stated in the Introduction of this study can be obtained.

In order to achieve these advantages we introduce the following notion [2–11].

Definition 2.8. *Let* $r > 0$ *be such that* $U(x^*, r) \subset D$. *Then* g' *is said to satisfy the center-majorant condition on* $U(x^*, r)$, *if*

$$\|F'(x^*)^{\dagger}\|\|F'(x) - F'(x^*)\| \leq g'(\|x - x^*\|) - g'(0). \tag{2.11}$$

Clearly,

$$g'(t) \leq \bar{h}'_{\lambda,\theta}(t) \quad \text{for any} \quad t \in [0, R], \quad \lambda, \theta \in [0,1] \tag{2.12}$$

holds in general and $\frac{\bar{h}'_{\lambda,\theta}(t)}{g'(t)}$ can be arbitrarily large [2–11].

It is worth noticing that (2.11) is not an additional condition to (2.8) since in practice the computation of function $\bar{h}_{\lambda,\theta}$ requires the computation of g as a special case (see also the numerical examples).

We have noticed that the iterates x_n lie in the more precise domain $U(x^*,r)\cap U(x^*,\varsigma)$ than $U(x^*,r)$. That is why, we introduce two more definitions not possible before, since Definition 2.8 was not used (i.e., function g was not known).

Definition 2.9. *Let $r>0$ be such that $U(x^*,r)\subset D$. Then, F' is said to satisfy the restricted majorant condition on $U(x^*,r)\cap U(x^*,\varsigma)$, if*

$$\|F'(x^*)^{\dagger}[F'(x)-F'(x^*+\varsigma(x-x^*))]\|\leq h'_{\lambda,\theta}(\|x-x^*\|)-h'_{\lambda,\theta}(\varsigma\|x-x^*\|) \tag{2.13}$$

for any $x\in U(x^*,r)\cap U(x^*,\varsigma)$ and $\varsigma\in[0,1]$.

Definition 2.10. *Let $r>0$ be such that $U(x^*,r)\subset D$. Then, F' is said to satisfy the restricted modified majorant condition on $U(x^*,r)\cap U(x^*,\varsigma)$, if*

$$\|F'(x^*)^{\dagger}\|\|F'(x)-F'(x^*+\varsigma(x-x^*))\|\leq h'_{\lambda,\theta}(\|x-x^*\|)-h'_{\lambda,\theta}(\varsigma\|x-x^*\|) \tag{2.14}$$

for any $x\in U(x^*,r)\cap U(x^*,\varsigma)$ and $\varsigma\in[0,1]$.

Then, again we have that

$$h'_{\lambda,\theta}(t)\leq\bar{h}_{\lambda,\theta}(t) \tag{2.15}$$

for any $t\in[0,R],\lambda,\theta\in[0,1]$.

3. **Local Convergence**

In this section, we present local convergence for inexact Newton method (1.2). Equation (1.1) is a surjective-undetermined (resp. injective-overdetermined) system if the number of equations is less (resp. greater) than the number of knowns and $F'(x)$ is of full rank for each $x\in D$. It is well known that, for surjective-underdetermined systems, the fixed points of the Newton operator $N_F(x) := x - F'(x)^{\dagger}F(x)$ are the zeros of F, while for injective-overdetermined systems, the fixed points of N_F are the least square solutions of (1.1), which, in general, are not necessarily the zeros of F.

We shall use the notation $D_0 = U(x^*, \varsigma)$ and set $D = U(x^*, R)$ and $D_1 = D_0 \cap U(x^*, r)$.

Next, we present the local convergence properties of inexact Newton method for general singular systems with constant rank derivatives.

Theorem 3.1. *Let $F : D \subset \mathbb{R}^j \to \mathbb{R}^m$ be continuously Fréchet differentiable nonlinear operator, D is open and convex. Suppose that $F(x^*) = 0$, $F'(x^*) \neq 0$ and that F' satisfies the restricted modified majorant condition (2.14) on D_1 and the center-majorant condition (2.11) on D, where r is given in (2.4). In addition, we assume that $rankF'(x) \leq rankF'(x^*)$ for any $x \in D_1$ and that*

$$\|[I_{\mathbb{R}^j} - F'(x)^{\dagger}F'(x)](x - x^*)\| \leq \theta\|x - x^*\|, \text{ for all } \quad x \in D_1, \tag{3.1}$$

where the constant θ satisfies $0 \leq \theta < 1$. Let sequence $\{x_n\}$ be generated by inexact Gauss-Newton method with any initial point $x_0 \in D_1 \backslash \{x^*\}$ and the conditions for the residual r_n and the forcing term λ_n:

$$\|r_n\| \leq \lambda_n\|F(x_n)\|, \quad 0 \leq \lambda_n F'(x_k) \leq \lambda \quad \textit{for each} \quad n = 0, 1, 2, \ldots \tag{3.2}$$

Then, $\{x_n\}$ converges to a zero x^* of $F'(\cdot)^{\dagger}F(\cdot)$ in $\bar{D}_1$. Moreover, we have the following estimate:

$$\|x_{n+1} - x^*\| \leq \frac{t_{n+1}}{t_n}\|x_n - x^*\| \quad \textit{for each} \quad n = 0, 1, 2, \ldots, \tag{3.3}$$

where the sequence $\{t_n\}$ is defined by (2.5).

Remark 3.2. *(a) If $g(t) = h_{\lambda,\theta}(t)$ and $\varsigma = R$ then the results obtained in Theorem 3.1 reduce to the ones given in [41, 42].*

(b) If $g(t)$ and $h_{\lambda,\theta}(t)$ are

$$g(t) = h_{\lambda,\theta}(t) = -(1 + \lambda + \theta)t + \int_0^t L(u)(t - u)\, du, \quad t \in [0, R], \tag{3.4}$$

then the results obtained in Theorem 3.1 reduce to the one given in [24]. Moreover, if taking $\lambda = 0$ (in this case $\lambda_n = 0$ and $r_n = 0$) in Theorem 3.1, we obtain the local convergence of Newton's method for solving the singular systems, which has been studied by Dedieu and Kim in [16] for analytic singular systems with constant rank derivatives and Li, Xu in [38] and Wang in [37] for some special singular systems with constant rank derivatives.

(c) If $g(t) < h_{\lambda,\theta}(t)$ then the improvements as mentioned in the Introduction of this study we obtained (see also the discussion above and below Definition 2.6)

If $F'(x)$ is full column rank for every $x \in D_1$, then we have $F'(x)^\dagger F'(x) = I_{\mathbb{R}^j}$. Thus,

$$\|[I_{\mathbb{R}^m} - F'(x)^\dagger F'(x)](x - x^*)\| = 0,$$

i. e., $\theta = 0$. We immediately have the following corollary:

Corolary 3.3. *Suppose that $rankF'(x) \leq rankF'(x^*)$ and that*

$$\|[I_{\mathbb{R}^m} - F'(x)^\dagger F'(x)](x - x^*)\| = 0,$$

for any $x \in D_1$. Suppose that $F(x^*) = 0$, $F'(x^*) \neq 0$ and that F' satisfies the restricted majorant condition (2.14) on D_1 and the center-majorant condition (2.11) on D. Let sequence $\{x_n\}$ be generated by inexact Gauss-Newton method with any initial point $x_0 \in D_1 \backslash \{x^*\}$ and the condition (3.2) for the residual r_n and the forcing term λ_n. Then, $\{x_n\}$ converges to a zero x^* of $F'(\cdot)^\dagger F(\cdot)$ in $\overline{D_1}$. Moreover, we have the following estimate:

$$\|x_{n+1} - x^*\| \leq \frac{t_{n+1}}{t_n}\|x_n - x^*\| \quad \textit{for each} \quad n = 0, 1, 2, \ldots, \tag{3.5}$$

where the sequence $\{t_n\}$ is defined by (2.5) for $\theta = 0$.

In the case when $F'(x^*)$ is full row rank, the modified majorant condition (2.8) can be replaced by the majorant condition (2.13).

Theorem 3.4. *Suppose that $F(x^*) = 0$, $F'(x^*)$ is full row rank, and that F' satisfies the restricted majorant condition (2.13) on D_1 and the center-majorant condition (2.11) on D, where r is given in (2.4). In addition, we assume that $rankF'(x) \leq ranfF'(x^*)$ for any $x \in$*

D_1 and that condition (3.1) holds. Let sequence $\{x_n\}$ be generated by inexact Gauss-Newton method with any initial point $x_0 \in D_1 \setminus \{x^*\}$ and the conditions for the residual r_n and the forcing term λ_n:

$$\|F'(x^*)^\dagger r_n\| \leq \lambda_n \|F'(x^*)^\dagger F(x_n)\|, 0 \leq \lambda_n F'(x^*)^\dagger F'(x_n) \leq \lambda \text{ for each } n = 0,1,2,\dots. \tag{3.6}$$

Then, $\{x_n\}$ converges to a zero x^* of $F'(\cdot)^\dagger F(.)$ in $\overline{D_1}$. Moreover, we have the following estimate:

$$\|x_{n+1} - x^*\| \leq \frac{t_{n+1}}{t_n} \|x_n - x^*\| \quad \text{for each} \quad n = 0,1,2,\dots,$$

where the sentence $\{t_n\}$ is defined by (2.5).

Remark 3.5. *Comments as in Remark 3.2 can follow for this case.*

Theorem 3.6. *Suppose that $F(x^*) = 0$, $F'(x^*)$ is full row rank, and that F' satisfies the restricted majorant condition (2.13) on D_1 and the center-majorant condition on D, where r is given in (2.4). In addition, we assume that rank$F'(x) \leq$ rank$F'(x^*)$ for any $x \in D_1$ and that condition (3.1) holds. Let sequence $\{x_n\}$ sequence generated by inexact Gauss-Newton method with any initial point $x_0 \in D_1 \setminus \{x^*\}$ and the conditions for the control residual r_n and the forcing term λ_n:*

$$\|F'(x_n)^\dagger r_n\| \leq \lambda_n \|F'(x_n)^\dagger F(x_n)\|, \quad 0 \leq \lambda_n F'(x_n) \leq \lambda \quad \text{for each} \quad n = 0,1,2,\dots \tag{3.7}$$

Then, $\{x_n\}$ converges to a zero x^* of $F'(\cdot)^\dagger F'(.)F$ in $\overline{D_1}$. Moreover, we have the following estimate:

$$\|x_{n+1} - x^*\| \leq \frac{t_{n+1}}{t_n} \|x_k - x^*\| \quad \text{for each} \quad n = 0,1,2,\dots,$$

where sequence $\{t_n\}$ is defined by (2.5).

Remark 3.7. *(a) In the case when $F'(x^*)$ is invertible in Theorem 3.6, $h_{\lambda,\theta}$ is given by (3.4) and $g(t) = -1 + \int_0^t L_0(t)(t-u)\,du$ for each $t \in [0,R]$, we obtain the local convergence results of inexact Gauss-Newton method for nonsingular systems, and the convergence ball r is this case satisfies*

$$\frac{\int_0^r L(u)u\,du}{r((1-\lambda) - \int_0^r L_0(u)\,du)} \leq 1, \quad \lambda \in [0,1). \tag{3.8}$$

In particular, if taking $\lambda = 0$, the convergence ball r determined in (3.8) reduces to the one given in [37] by Wang and the value r is the optimal radius of the convergence ball when the equality holds. That is our radius is r larger than the one obtained in [37], if $L_0 < L$ (see also the numerical examples). Notice that L is used in [37] for the estimate (3.8). Then, we can conclude that vanishing residuals, Theorem 3.6 merges into the theory of Newton's method.

(b) Clearly, all the preceding results hold if function $\bar{h}_{\lambda,\theta}$ replaces function $h_{\lambda,\theta}$.

4. **Proofs**

In this section, we prove our main results of local convergence for inexact Gauss-Newton method (1.2) given in Section 3.

4.1. **Proof of Theorem 3.1.**

Lemma 4.1. *Suppose that F' satisfies the center majorant condition on D_0 and that $\|x^* - x\| < \min\{\rho, \delta\}$, where r, ρ, δ and x^* are defined in (2.4), (2.2) and (2.1), respectively. Then, $rankF'(x) = rankF'(x^*)$ and*

$$\|F'(x)^\dagger\| \leq -\frac{\|F'(x^*)^\dagger\|}{g'(\|x - x^*\|)}.$$

Proof. Since $g'(0) = -1$, we have

$$\|F'(x^*)^\dagger\|\|F'(x) - F'(x^*)\| \leq g'(\|x - x^*\|) - g'(0) < -g'(0) = 1.$$

It follows from Lemma (2.1) that $rankF'(x) = rankF'(x^*)$ and

$$\|F'(x)^\dagger\| \leq \frac{\|F'(x^*)^\dagger\|}{1 - (g'(\|x - x^*\|) - g'(0))} = -\frac{\|F'(x^*)^\dagger\|}{g'(\|x - x^*\|)}.$$

Proof of Theorem 3.1. We shall prove by mathematical induction on n that $\{t_n\}$ is the majorizing sequence for $\{x_n\}$, i. e.,

$$\|x^* - x_j\| \leq t_j \quad \text{for each} \quad j = 0, 1, 2, \ldots \tag{4.1}$$

Because $t_0 = \|x_0 - x^*\|$, thus (4.1) holds for $j = 0$. Suppose that $\|x^* - x_j\| \le t_j$ for some $j = n \in \mathbb{N}$. For the case $j = n+1$, we first have that,

$$\begin{aligned} x_{n+1} - x^* &= x_n - x^* - F'(x_n)^\dagger[F(x_n) - F(x^*)] + F'(x_n)^\dagger r_n \\ &= F'(x_n)^\dagger[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)] + F'(x_n)^\dagger r_n \\ &\quad + [I_{\mathbb{R}^j} - F'(x_n)^\dagger F'(x_n)](x_n - x^*) \\ &= F'(x_n)^\dagger \int_0^1 [F'(x_n) - F'(x^* + \varsigma(x_n - x^*))](x_n - x^*)\, d\varsigma \\ &\quad + F'(x_n)^\dagger r_n + [I_{\mathbb{R}^j} - F'(x_n)^\dagger F'(x_n)](x_n - \zeta). \end{aligned} \tag{4.2}$$

By using the restricted modified majorant condition (2.14), Lemma 2.4, the inductive hypothesis (4.1) and Lemma 2.2, we obtain in turn that

$$\begin{aligned} &\left\| F'(x_n)^\dagger \int_0^1 [F'(x_n) - F'(x^* + \varsigma(x_n - x^*))](x_n - x^*)\, d\varsigma \right\| \\ \le\ & -\frac{1}{g'(\|x_n - x^*)\|} \int_0^1 \|F'(x^*)^\dagger\| \|F'(x_n) - F'(x^* + \varsigma(x_n - x^*))\| \|x_n - x^*\|\, d\varsigma \\ =\ & -\frac{1}{g'(\|x_n - x^*\|)} \int_0^1 \frac{h'_{\lambda,0}(\|x_n - x^*\|) - h'_{\lambda,0}(\varsigma\|x_n - x^*\|)}{\|x_n - x^*\|}\, d\varsigma \cdot \|x_n - x^*\|^2 \\ \le\ & -\frac{1}{g'(t_n)} \int_0^1 \frac{h'_{\lambda,0}(t_n) - h_{\lambda,0}(\varsigma t_n)}{t_n}\, d\varsigma \cdot \|x_n - x^*\|^2 \\ =\ & -\frac{1}{g'(t_n)} (t_n h'_{\lambda,0}(t_n) - h_{\lambda,0}(t_n)) \frac{\|x_n - x^*\|^2}{t_n^2}. \end{aligned}$$

In view of (3.2),

$$\|F'(x_n)^\dagger r_n\| \le \|F'(x_n)^\dagger\| \|r_n\| \le \lambda_n \|F'(x_n)^\dagger\| \|F(x_n)\|. \tag{4.3}$$

We have that

$$\begin{aligned} -F(x_n) &= F(x^*) - F(x_n) - F'(x_n)(x^* - x_n) + F'(x_n)(x^* - x_n) \\ &= \int_0^1 [F'(x_n) - F'(x^* + \varsigma(x_n - x^*))](x_n - x^*)\, d\varsigma \\ &\quad + F'(x_n)(x^* - x_n). \end{aligned} \tag{4.4}$$

Then, combining Lemma 2.2, Lemma 4.1, the restricted majorant condition (2.14), the inductive hypothesis (4.1) and the condition (3.2), we obtain in turn that

$$\begin{aligned}
& \lambda_n\|F'(x_n)^{\dagger}\|\|F(x_n)\| \\
\leq\ & \lambda_n\|F'(x_n)^{\dagger}\|\int_0^1 \|F'(x_n)-F'(x^*+\varsigma(x_n-x^*))\|\|x_n-x^*\|\,d\varsigma \\
& +\lambda_n\|F'(x_n)^{\dagger}\|\|F'(x_n)\|\|x_n-x^*\| \\
\leq\ & -\frac{\lambda}{g'(t_n)}(t_nh'_{\lambda,0}(t_n)-h_{\lambda,0}(t_n))\frac{\|x_n-x^*\|^2}{t_n^2}+\lambda t_n\frac{\|x_n-x^*\|}{t_n} \\
\leq\ & \lambda\frac{\lambda t_n+h_{\lambda,0}(t_n)}{g'(t_n)}\frac{\|x_n-x^*\|}{t_n}. \qquad (4.5)
\end{aligned}$$

Combining (3.1), (4.3), (4.3) and (4.5), we get that

$$\begin{aligned}
\|x_{n+1}-x^*\| &\leq \left[-\frac{t_nh'_{\lambda,0}(t_n)-h_{\lambda,0}(t_n)}{g'(t_n)}+\lambda\frac{\lambda t_n+h_{\lambda,0}(t_n)}{g'(t_n)}+\theta t_n\right]\frac{\|x_n-x^*\|}{t_n} \\
&= \left[-t_n+(1+\lambda)\left(\frac{\lambda t_n}{g'(t_n)}+\frac{h_{\lambda,0}(t_n)}{g'(t_n)}\right)+\theta t_n\right]\frac{\|x_n-x^*\|}{t_n}.
\end{aligned}$$

But, we have that $-1<g'(t)<0$ for any $t\in(0,\rho)$, so

$$(1+\lambda)\left(\frac{\lambda t_n}{g'(t_n)}+\frac{h_{\lambda,0}(t_n)}{g'(t_n)}\right)+\theta t_n\leq\frac{h_{\lambda,0}(t_n)}{g'(t_n)}+\theta_n\leq\frac{h_{\lambda,0}(t_n)-\theta t_n}{g'(t_n)}=\frac{h_{\lambda,\theta}(t_n)}{g'(t_n)}.$$

Using the definition of $\{t_n\}$ given in (2.5), we get that

$$\|x_{n+1}-x^*\|\leq\frac{t_{n+1}}{t_n}\|x_n-x^*\|,$$

so we deduce that $\|x_{n+1}-x^*\|\leq t_{n+1}$, which completes the induction. In view of the fact that $\{t_n\}$ converges to 0 (by Lemma 2.5), it follows from (4.1) that $\{x_n\}$ converges to x^* and the estimate (3.3) holds for all $n\geq 0$.

4.2. **Proof of Theorem 3.4.**

Lemma 4.2. *Suppose that $F(x^*)=0$, $F'(x^*)$ is full row rank and that F' satisfies the center condition (2.7) on D_0. Then, for each $x\in D_0$, we have $rankF'(x)=rankF'(x^*)$ and*

$$\|[I_{\mathbb{R}^j}-F'(x^*)^{\dagger}(F'(x^*)-F'(x))]^{-1}\|\leq-\frac{1}{g'(\|x-x^*\|)}.$$

Proof. Since $g'(0) = -1$, we have

$$\|F'(x^*)^\dagger [F'(x) - F'(x^*)]\| \leq g'(\|x - x^*\|) - g'(0) < -g'(0) = 1.$$

It follows from Banach lemma that $[I_{\mathbb{R}^j} - F'(x^*)^\dagger (F'(x^*) - F'(x))]^{-1}$ exists and

$$\|[I_{\mathbb{R}^j} - F'(x^*)^\dagger (F'(x^*) - F'(x))]^{-1}\| \leq -\frac{1}{g'(\|x - x^*\|)}.$$

Since $F'(x^*)$ is full row rank, we have $F'(x^*)F'(x^*)^\dagger = I_{\mathbb{R}^m}$ and

$$F'(x) = F'(x^*)[I_{\mathbb{R}^j} - F'(z^*)^\dagger (F'(x^*) - F'(x))],$$

which implies that $F'(x)$ is full row, i. e., $rankF'(x) = rankF'(x^*)$.

Proof of Theorem 3.4. Let $\widehat{F} : U(x^*, r) \to \mathbb{R}^m$ be defined by

$$\widehat{F}(x) = F'(x^*)^\dagger \widehat{F}(x), \quad x \in U(x^*, r),$$

with residual $\widehat{r}_k = F'(x^*)^\dagger r_n$. In view of

$$\widehat{F}'(x)^\dagger = [F'(x^*)^\dagger F'(x)]^\dagger = F'(x)^\dagger F'(x^*), \quad x \in U(x^*, r),$$

we have that $\{x_n\}$ coincides with the sequence generated by inexact Gauss-Newton method (1.2) for $\widehat{F}$. Moreover, we get that

$$\widehat{F}'(x^*)^\dagger = (F'(x^*)^\dagger F'(x^*))^\dagger = F'(x^*)^\dagger F'(x^*).$$

Consequently,

$$\|\widehat{F}'(x^*)^\dagger \widehat{F}'(x^*)\| = \|F'(x^*)^\dagger F'(x^*) F'(x^*)^\dagger F(x^*)\| = \|F'(x^*)^\dagger F(x^*)\|.$$

Because $\|F'(x^*)^\dagger F(x^*)\| = \|\Pi_{kerF'(x^*)^\perp}\| = 1$, thus, we have

$$\|\widehat{F}'(x^*)^\dagger\| = \|\widehat{F}'(x^*)^\dagger \widehat{F}'(x^*)\| = 1.$$

Therefore, by (2.7), we can obtain that

$$\begin{aligned} \|\widehat{F}'(x^*)^\dagger\| \|\widehat{F}'(x) - \widehat{F}'(x^* + \varsigma(x - x^*))\| &= \|F'(x^*)^\dagger (F'(x) - F'(x^* + \varsigma(x - x^*)))\| \\ &\leq h'_{\lambda,\theta}(\|x - x^*\|) - h_{\lambda,\theta}(\varsigma\|x - x^*\|). \end{aligned}$$

Hence, $\widehat{F}$ satisfies the restricted majorant condition (2.14) on D_1. Then, Theorem 3.1 is applicable and $\{x_k\}$ converges to x^* follows. Note that, $\widehat{F}'(\cdot)^\dagger\widehat{F}(\cdot) = F'(\cdot)^\dagger F(\cdot)$ and $F(\cdot) = F'(\cdot)F'(\cdot)^\dagger F(\cdot)$. Hence, we conclude that x^* is a zero of F.

4.3. **Proof of Theorem 3.6.**

Lemma 4.3. *Suppose that $F(x^*) = 0$, $F'(x^*)$ is full row rank and that F' satisfies the center majorant condition (2.7) on D. Then, we have*

$$\|F'(x)^\dagger F'(x^*)\| \leq -\frac{1}{g'(\|x-x^*\|)} \quad \text{for each} \quad x \in D.$$

Proof. Since $F'(x^*)$ is full row rank, we have $F'(x^*)F'(x^*)^\dagger = I_{R^m}$. Then, we get that

$$F'(x)^\dagger F'(x^*)(I_{\mathbb{R}^j} - F'(x^*)^\dagger(F'(x^*) - F'(x^*))) = F'(x)^\dagger F'(x), \quad x \in D_1.$$

By Lemma 4.2, $I_{\mathbb{R}^j} - F'(x^*)^\dagger(F'(x^*) - F'(x))$ is invertible for any $x \in D_1$. Thus, in view of the equality $A^\dagger A = \Pi_{kerA^\perp}$ for any $m \times j$ matrix A, we obtain that

$$F'(x)^\dagger F'(x^*) = \Pi_{kerF'(x)^\perp}[I_{\mathbb{R}^j} - F'(x^*)^\dagger(F'(x^*) - F'(x))]^{-1}.$$

Therefore, by Lemma 4.2 we deduce that

$$\|F'(x)^\dagger F'(x^*)\| \leq \|\Pi_{kerF'(x)^\perp}\|\|[I_{\mathbb{R}^j} - F'(x^*)^\dagger(F'(x^*) - F'(x))]^{-1}\| \leq -\frac{1}{g'(\|x-x^*\|)}.$$

Proof of Theorem 3.6 Using Lemma 4.3, majorant condition (2.7) and the residual condition (3.7), respectively, instead of Lemma 4.1, restricted majorant condition (2.14) and condition (3.2), one can complete the proof of Theorem 3.6 in an analogous way to the proof of Theorem 3.1.

Remark 4.4. *Our results improve the results in [42, 43], since $D_1 \subset U(x^*, r)$ leading to an at least tight function $h'_{\lambda,\theta}$ than the one used in [42, 43].*

5. **Numerical Examples**

We present some numerical examples, where

$$g(t) < h_{\lambda,\theta}(t), \; g'(t) < h'_{\lambda,\theta}(t) \tag{5.1}$$

and

$$g(t) < \bar{h}_{\lambda,\theta}(t), g'(t) < \bar{h}_{\lambda,\theta}(t). \tag{5.2}$$

For simplicity we take $F'(x)^{\dagger} = F'(x)^{-1}$ for each $x \in D$.

Example 5.1. Let $X = Y = (-\infty, +\infty)$ and define function $F : X \to Y$ by

$$F(x) = d_0 x - d_1 \sin(1) + d_1 \sin(e^{d_2 x})$$

where d_0, d_1, d_2 are given real numbers. Then $x^* = 0$. Define functions g and $h_{\lambda,\theta}$ by $g(t) = \frac{L_0}{2}t^2 - t$ and $h_{\lambda,\theta}(t) = \frac{L}{2}t^2 - t$. Then, it can easily be seen that for d_2 sufficiently large and d_1 sufficiently small $\frac{L}{L_0}$ can be arbitrarily large. Hence, (5.1) and (5.2) hold.

Example 5.2. Let $F(x,y,z) = 0$ be a nonlinear system, where $F : D := U(0,1) \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ and $F(x,y,z) = (x, \dfrac{e-1}{2}y^2 + y, \mathrm{e}^z - 1)^T$. It is obvious that $\bar{x}^* = (0,0,0)^T$ is a solution of the system.We have from the definition of operator F that its Fréchet derivative is given by

$$F'(\bar{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (e-1)y & 0 \\ 0 & 0 & \mathrm{e}^z \end{pmatrix} \quad \text{and} \quad F'(\bar{x}^*) = \mathrm{diag}\{1,1,1\},$$

where $\bar{x} = (x,y,z)^T$. Hence, $[F'(\bar{x}^*)]^{-1} = \mathrm{diag}\{1,1,1\}$. Moreover, we can define for $L_0 = e-1 < L = e$, $g(t) = \dfrac{e-1}{2}t^2 - t$, $\bar{h}_{\lambda,\theta}(t) = \frac{e}{2}t^2 - t$ and $h_{\lambda,\theta}(t) = \dfrac{e^{\frac{1}{L_0}}}{2}t^2 - t$. Then, again (5.1) and (5.2) hold, since $L_0 < e^{\frac{1}{L_0}} = 1.78957239 < L$. Hence, the present results improve the ones in [42, 43].

Other examples, where (5.1) and (5.2) are satisfied can be found in [2, 5, 8, 9, 11, 12].

6. Conclusion

We expanded the applicability of inexact Gauss-Newton method under a restricted majorant and a center-majorant condition. The advantages of our analysis over earlier works such as [5, 7, 12–42] are also shown under the same computational cost for the functions and constants involved. These advantages include: a large radius of convergence and more precise error estimates on the distances $\|x_{n+1} - x^*\|$ for each $n = 0, 1, 2, \ldots$, leading to a wider choice of initial guesses and computation of less iterates x_n in order to obtain a desired error tolerance. Numerical examples show that the center-function can be smaller than the majorant function.

REFERENCES

[1] S. Amat, S. Busquier, and J. M.,Gutiérrez, Geometric constructions of iterative functions to solve nonlinear equations, J. Comput. Appl. Math. 157 (2003), 197–205.

[2] I. K. Argyros, Computational theory of iterative methods, Studies in Computational Mathematics, 15, Editors: K. Chui and L. Wuytack. Elsevier, New York, (2007).

[3] I. K. Argyros, Concerning the convergence of Newton's method and quadratic majorants, J. Appl. Math. Comput. 29 (2009), 391–400.

[4] I. K. Argyros, and S. Hilout, On the Gauss-Newton method, J. Appl. Math. 2010 (2010), 1–14.

[5] I. K. Argyros, and S. Hilout, Extending the applicability of the Gauss-Newton method under average Lipschitz-conditions, Numer. Alg. 58 (2011), 23–52.

[6] I. K. Argyros, A semilocal convergence analysis for directional Newton methods, Math. Comput. AMS, 80 (2011), 327–343.

[7] I. K. Argyros, and S. Hilout, On the solution of systems of equations with constant rank derivatives, Numer. Algor. 57 (2011), 235–253.

[8] I. K. Argyros, Y. J. Cho, and S. Hilout, Numerical methods for equations and its applications, CRC Press/Taylor and Francis Group, New-York (2012).

[9] I. K. Argyros and S. Hilout, Improved local convergence of Newton's method under weaker majorant condition, J. Comput. Appl. Math. 236 (2012), 1892–1902.

[10] I. K. Argyros, and S. Hilout, Weaker conditions for the convergence of Newton's method, J. Complexity 28 (2012), 364–387.

[11] I. K. Argyros and S. Hilout, Computational Methods in Nonlinear Analysis, World Scientific Publ. Comput. New Jersey (2013).

[12] I. K. Argyros and A. A. Magrenan, Local convergence analysis of proximal Gauss-Newton method for penalized least-square problems, Appl. Math. Comput. 241 (2014), 401–408.

[13] A. Ben-Israel and T. N. E. Greville, Generalized inverses. CMS Books in Mathematics/Ouvrages de Mathematiques de la SMC, 15. Springer-Verlag, New York, second edition, Theory and Applications (2003).

[14] J. V. Burke and M. C. Ferris, A Gauss-Newton method for convex composite optimization, Math. Programming Ser A. 71 (1995), 179–194.

[15] J. Chen, The convergence analysis of inexact Gauss-Newton, Comput. Optim. Appl. 40 (2008), 97–118.

[16] J. Chen and W. Li, Convergence behaviour of inexact Newton methods under weak Lipschitz condition, J. Comput. Appl. Math. 191 (2006), 143–164.

[17] J. P. Dedieu and M. H. Kim, Newton's method for analytic systems of equations with constant rank derivatives, J. Complexity 18 (2002), 187–209.

[18] J. P. Dedieu and M. Shub, Newton's method for overdetermined systems of equations, Math. Comput. 69 (2000), 1099–1115.

[19] R. S. Dembo, S.C. Eisenstat and T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (1982), 400–408.

[20] P. Deuflhard and G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods, SIAM J. Numer. Anal. 16 (1979), 1–10.

[21] O. P. Ferreira, Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle, IMA J. Numer. Anal. 29 (2009), 746–759.

[22] O. P. Ferreira, Local convergence of Newton's method under majorant condition, J. Comput. Appl. Math. 235 (2011), 1515–1522.

[23] O. P. Ferreira and M. L. N. Gonçalves, Local convergence analysis of inexact Newton-like methods under majorant condition, Comput. Optim. Appl. 48 (2011), 1–21.

[24] O. P. Ferreira, M. L. N. Gonçalves and P. R. Oliveira, Local convergence analysis of the Gauss-Newton method under a majorant condition, J. Complexity 27 (2011), 111–125.

[25] O. P. Ferreira, M. L. N. Gonçalves and P. R. Oliveira, Local convergence analysis of inexact Gauss-Newton like method under majorant condition, J. Comput. Appl. Math. 236 (2012), 2487–2498.

[26] M. L. N. Gonçalves and P. R. Oliveira, Convergence of the Gauss-Newton method for a special class of systems of equations under a majorant condition, Optim. http://dx.doi.org/10.1080/02331934.2013.778854.

[27] W. M. Häussler, A Kantorovich-type convergence analysis for the Gauss-Newton method. Numer. Math. 48 (1986), 119–125.

[28] J. B. Hiriart-Urruty and C. Lemaréchal, Convex analysis and minimization algorithms (two volumes). I. Fundamentals. II. Advanced theory and bundle methods, 305 and 306, Springer–Verlag, Berlin (1993).

[29] L. V. Kantorovich and G. P. Akilov, Functional Analysis, Pergamon Press, Oxford (1982).

[30] C. Li and K. F. Ng, Majorizing functions and convergence of the Gauss-Newton method for convex composite optimization, SIAM J. Optim. 18 (2007), 613–692.

[31] C. Li, N. Hu and J. Wang, Convergence behaviour of Gauss-Newton's method and extensions of Smale point estimate theory, J. Complexity 26 (2010), 268–295.

[32] B. Morini, Convergence behaviour of inexact Newton methods, Math. Comput. 68 (1999), 1605–1613.

[33] F. A. Potra and V. Pták, Nondiscrete induction and iterative processes, Pitman, (1994).

[34] S. M. Robinson, Extension of Newton's method to nonlinear functions with values in a cone, Numer. Math. 19 (1972), 341–347.

[35] R. T. Rockafellar, Convex analysis, Princeton Mathematical Series, 28, Princeton University Press, Princeton, N. J. (1970).

[36] M. Shub and S. Smale, Complexity of Bézout's theorem iv: Probability of success extensions, SIAM J. Numer. Anal. 33 (1996), 128–148.

[37] S. Smale, Newton's method estimates from data at one point. The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985), 185-196, Springer, New York (1986).

[38] W. Wang, Convergence of Newton's method and uniqueness of the solution of equations in Banach space, IMA J. Numer. Anal. 20 (2000), 123–134.

[39] X. Xu and C. Li, Convergence of Newton's method for systems of equations with constant rank derivatives, J. Comput. Math. 25 (2007), 705–718.

[40] X. Xu and C. Li, Convergence criterion of Newton's method for singular systems with constant rank derivatives, J. Math. Anal. Appl. 345 (2008), 689–701.

[41] T. J. Ypma, Local convergence of inexact Newton's methods, SIAM J. Numer. Anal., 21 (1984) 583–590.

[42] F. Zhou, An analysis on local convergence of inexact Newton-Gauss method solving singular systems of equations, Sci. World J. 2014 (2014), Article ID 752673.

[43] F. Zhou, On local convergence analysis of inexact Newton method for singular systems of equations under majorant condition, Sci. World J. 2014 (2014), Article ID 498016.