



LOCAL CONVERGENCE ANALYSIS OF INEXACT GAUSS-NEWTON METHOD FOR SINGULAR SYSTEMS OF EQUATIONS UNDER RESTRICTED CONVERGENCE DOMAINS

IOANNIS K. ARGYROS^{1,*}, SANTHOSH GEORGE²

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

²Department of Mathematical and Computational Sciences,
National Institute of Technology, Karnataka, 575025, India

Abstract. A new semi-local convergence analysis of the Gauss-Newton method for solving convex composite optimization problems is presented using restricted convergence domains. The results extend the applicability of the Gauss-Newton method under the same computational cost as in earlier studies. 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; Restricted convergence domain; Majorant function; Center-majorant function.

1. Introduction

In this study, we are concerned with the problem of approximating a solution of the equation

$$F(x) = 0, \quad (1.1)$$

where D is open and convex and $F : D \subset \mathbb{R}^j \rightarrow \mathbb{R}^m$ is a nonlinear operator with its Fréchet derivative denoted by F' . In the case $m = j$, the inexact Newton method was defined in [7] by:

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

*Corresponding author.

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

Received March 25, 2016

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

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

and $\{\lambda_n\}$ is a sequence of forcing terms such that $0 \leq \lambda_n < 1$. Let x^* be a solution of (1.1) such that $F'(x^*)$ is invertible. As shown in [7], if $\lambda_n \leq \lambda < 1$, 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 [7], 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. As pointed out by [16] the result of [7] is difficult to apply due to dependence of the norm $\|\cdot\|_*$, which is not computable.

In [41], Ypma used the affine invariant condition of residual control in the form:

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

instead of (1.3) to study the local convergence of inexact Newton method (1.2). And the radius of convergent result are also obtained. Morini in [32] presented the following variation for the residual controls:

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

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.

Recently, several authors have studied the convergence behaviour of singular nonlinear systems by Gauss-Newton's method, which is defined by

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

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) $F'(x_n)$ [1, 12, 14, 15, 17, 18, 20, 21, 36].

In the present study, using the idea of restricted 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 $n = 0, 1, \dots$ and a clearer relationship between the majorant function (see (2.7) and the associated least squares problems (1.1). These advantages are obtained because we use a center-type majorant condition (see (2.8)) for the computation of inverses involved which is more precise than the majorant condition used in [21–26, 30, 31, 39–43]. 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 $n = 0, 1, 2, \dots$

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 [13, 28, 35].

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

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

where A^* denotes the adjoint of A . Let $\ker A$ and $\operatorname{im} A$ 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

$$A^\dagger A = \Pi_{\ker A^\perp} \quad \text{and} \quad A A^\dagger = \Pi_{\operatorname{im} A}.$$

In particular, in the case when A is full row rank (or equivalently, when A is surjective), $A A^\dagger = I_{\mathbb{R}^m}$; 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 [25].

Lemma 2.1. ([25, 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 $\text{rank}A = r$, $1 \leq \text{rank}B \leq A$ and $\|A^\dagger\| \|B - A\| < 1$. Then, $\text{rank}B = r$ 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. ([25, Proposition 1.3]). *Let $R > 0$. If $\varphi : [0, R] \rightarrow \mathbb{R}$ is continuously differentiable and convex, then, the following assertions hold:*

- (a) $\frac{\varphi(t) - \varphi(\tau t)}{t} \leq (1 - \tau)\varphi'(t)$ for each $t \in (0, R)$ and $\tau \in [0, 1]$.
- (b) $\frac{\varphi(u) - \varphi(\tau u)}{u} \leq \frac{\varphi(v) - \varphi(\tau v)}{v}$ for each $u, v \in [0, R]$, $u < v$ and $0 \leq \tau \leq 1$.

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

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

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

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

- (i) $\psi(0, \lambda, \theta) = 0$ and $\left. \frac{\partial}{\partial t} \psi(t, \lambda, \theta) \right|_{t=0} = -(1 + \lambda + \theta)$.
- (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) $g : [0, R] \rightarrow \mathbb{R}$ is strictly increasing with $g(0) = 0$.
- (vi) g' is convex and strictly increasing with $g'(0) = -1$.
- (vii) $g(t) \leq h_{\lambda, \theta}(t)$, $g'(t) \leq h'_{\lambda, \theta}(t)$ for each $t \in [0, R)$, $\lambda, \theta \in [0, 1]$.

Define

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

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

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

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 $\zeta \geq \delta (> 0)$. 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 $h'_{\lambda,\theta}$, $g'(t)$ 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,*

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

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 2.3 that

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

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

$$r := \min\{\rho, \delta\}, \quad (2.4)$$

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

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

Lemma 2.5. *The sequence $\{t_n\}$ given by (2.5) is well defined, strictly decreasing, remains in $(0, \rho)$ for each $n = 0, 1, 2, \dots$ 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 $n = 0, 1, 2, \dots$. Hence, there exists $t^* \in [0, \rho)$ such that $\lim_{n \rightarrow +\infty} t_n = t^*$. 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 $t_n \rightarrow 0$ as $n \rightarrow +\infty$.

If $g(t) = h_{\lambda, \theta}(t)$, then Lemmas 2.3, 2.4 and 2.5 reduce to the corresponding ones in [42, 43]. 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 [42, 43] is defined by

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

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

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

and

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

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 [10–13].

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

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

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

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 [42, 43] 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

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

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

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

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

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

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

Condition (2.10) is used to produce the Banach-type perturbation Lemmas in [42, 43] for the computation of the upper bounds on the norms $\|F'(x)^\dagger\|$. In this study we use a more flexible function g than $h_{\lambda, \theta}$ function 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, 3, 5–12].

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

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

Clearly,

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

holds in general and $\frac{h'_{\lambda, \theta}(t)}{g'(t)}$ can be arbitrarily large [2, 3, 5–12].

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

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 $D_0 = U(x^*, \xi)$ and $D = U(x^*, R)$ and set $D_1 = D_0 \cap U(x^*, r)$.

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

Theorem 3.1. *Let $F : D \subset \mathbb{R}^j \rightarrow \mathbb{R}^m$ 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 modified majorant condition (2.7) on D_1 and the center-majorant condition (2.8) on D , where r is given in (2.4). In addition, we assume that $\text{rank}F'(x) \leq \text{rank}F'(x^*)$ for any $x \in U(x^*, r)$ and that*

$$\|[I_{\mathbb{R}^j} - F'(x)^\dagger F'(x)](x - x^*)\| \leq \theta \|x - x^*\|, \quad x \in U(x^*, r), \quad (3.1)$$

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

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

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

$$\|x_{n+1} - x^*\| \leq \frac{t_{n+1}}{t_n} \|x_n - x^*\| \quad \text{for each } n = 0, 1, 2, \dots, \quad (3.3)$$

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

Remark 3.2.

(a) If $g(t) = h_{\lambda, \theta}(t)$, then the results obtained in Theorem 3.1 reduce to the ones given in [42, 43].

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

$$(0.1) \quad g(t) = h_{\lambda, \theta}(t) = -(1 + \lambda + \theta)t + \int_0^t L(u)(t - u) du, \quad t \in [0, R],$$

then the results obtained in Theorem 3.1 reduce to the one given in [25]. 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 [17] for analytic singular systems with constant rank derivatives and Li, Xu in [39] and Wang in [38] 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 $x \in U(x^*, r)$, 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.

Corollary 3.3. Suppose that $\text{rank} F'(x) \leq \text{rank} F'(x^*)$ and that

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

for any $x \in U(x^*, r)$. Suppose that $F(x^*) = 0$, $F'(x^*) \neq 0$ and that F' satisfies the modified majorant condition (2.7) on D_1 and the center-majorant condition (2.8) on D . Let sequence $\{x_n\}$ be generated by inexact Gauss-Newton method with any initial point $x_0 \in U(x^*, r) \setminus \{x^*\}$ 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 $\overline{U}(x^*, r)$. Moreover, we have the following estimate:

$$\|x_{n+1} - x^*\| \leq \frac{t_{n+1}}{t_n} \|x_n - x^*\| \quad \text{for each } n = 0, 1, 2, \dots, \quad (3.5)$$

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 (2.7) can be replaced by the majorant condition (2.6).

Theorem 3.4. *Suppose that $F(x^*) = 0$, $F'(x^*)$ is full row rank, and that F' satisfies the majorant condition (2.6) on D_1 and the center-majorant condition (2.8) on D , where r is given in (2.4). In addition, we assume that $\text{rank}F'(x) \leq \text{rank}F'(x^*)$ for any $x \in U(x^*, r)$ and that condition (3.1) holds. Let sequence $\{x_n\}$ be generated by inexact Gauss-Newton method with any initial point $x_0 \in U(x^*, r) \setminus \{x^*\}$ and the conditions for the residual r_n and the forcing term λ_n :*

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

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

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

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 majorant condition (2.6) on D_1 and the center-majorant condition on D , where r is given in (2.4). In addition, we assume that $\text{rank}F'(x) \leq \text{rank}F'(x^*)$ for any $x \in U(x^*, r)$ and that condition (3.1) holds. Let sequence $\{x_n\}$ be generated by inexact Gauss-Newton method with any initial point $x_0 \in U(x^*, r) \setminus \{x^*\}$ and the conditions for the control residual r_n and the forcing term λ_n :*

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

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

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

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

Remark 3.7. In the case when $F'(x^*)$ is invertible in Theorem 3.6, $h_{\lambda, \theta}$ is given by 3.4 and $g(t) = -1 + \int_0^t L_0(u)(t-u) du$ 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 in this case satisfies

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

In particular, if taking $\lambda = 0$, the convergence ball r determined in (3.8) reduces to the one given in [38] by Wang and the value r is the optimal radius of the convergence ball when the equality holds. That is our radius r is larger than the one obtained in [38], if $L_0 < L$ (see also the numerical examples). Notice that L is used in [38] for the estimate (3.8). Then, we can conclude that vanishing residuals, Theorem 3.6 merges into the theory of Newton's method.

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 modified majorant condition on $U(x^*, r)$ and that $\|x^* - x\| < \min\{\rho, x^*\}$, where r , ρ and x^* are defined in (2.4), (2.2) and (2.1), respectively. Then, $\text{rank}F'(x) = \text{rank}F'(x^*)$ 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 $\text{rank}F'(x) = \text{rank}F'(x^*)$ 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.,

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

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

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

By using the modified majorant condition (2.8), Lemma 2.4, the inductive hypothesis (4.1) and Lemma 2.2, we obtain in turn that

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

In view of (3.2),

$$\|F'(x_n)^\dagger r_n\| \leq \|F'(x_n)^\dagger\| \|r_n\| \leq \lambda_n \|F'(x_n)^\dagger\| \|F(x_n)\|, \quad (4.3)$$

we have that

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

Then, combining Lemma 2.2, Lemma 4.1, the modified majorant condition (2.8), the inductive hypothesis (4.1) and the condition (3.2), we obtain in turn that

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

Combining (3.1), (4.3), (4.4) and (4.5), we get that

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

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

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

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.

Lema 4.2. *Suppose that $F(x^*) = 0$, $F'(x^*)$ is full row rank and that F' satisfies the majorant condition (2.7) on D_1 . Then, for each $x \in U(x^*, r)$, we have $\text{rank}F'(x) = \text{rank}F'(x^*)$ 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

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

which implies that $F'(x)$ is full row, i. e., $\text{rank}F'(x) = \text{rank}F'(x^*)$.

Proof of Theorem 3.4. Let $\widehat{F} : U(x^*, r) \rightarrow \mathbb{R}^m$ 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 $\|F'(x^*)^\dagger F'(x^*)\| = \|\Pi_{\ker F'(x^*)^\perp}\| = 1$, thus, we have

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

Therefore, by (2.7), we can obtain that

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

Hence, \widehat{F} satisfies the modified majorant condition (2.8) 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 majorant condition (2.7) on D_1 . Then, we have*

$$\|F'(x)^\dagger F'(x^*)\| \leq \frac{1}{g'(\|x - x^*\|)} \quad \text{for each } x \in D_1.$$

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 $A^\dagger A = \Pi_{\ker A^\perp}$ for any $m \times j$ matrix A , we obtain that

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

Therefore, by Lemma 4.2 we deduce that

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

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, modified majorant condition (2.8) 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. The results in [4] improved the corresponding ones in [42, 43]. In the present study, we improved the results in [4], since $D_1 \subset U(x^*, r)$ leading to an at least as tight function $h'_{\lambda, \theta}$ than the one used in [4] (see also Example 5.2).

5. Numerical examples

We present some numerical examples, where

$$g(t) < h_{\lambda,\theta}(t) \quad (5.1)$$

and

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

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 $F : X \rightarrow Y$ by

$$F(x) = d_0x - d_1 \sin(1) + d_1 \sin(e^{d_2x})$$

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 $F : D = U(0, 1) \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and $F(x, y, z) = (x, \frac{e-1}{2}y^2 + y, e^z - 1)$. It is obvious that $(0, 0, 0) = \bar{x}^*$ is a solution of the system.

From F , we deduce

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

where $\bar{x} = (x, y, z)$. Hence, $[F'(\bar{x}^*)]^{-1} = \text{diag}\{1, 1, 1\}$. Moreover, we can define for $L_0 = e - 1 < L = 1.78957239$, $g(t) = \frac{e-1}{2}t^2 - t$ and $h_{\lambda,\theta}(t) = \frac{1.78957239}{2}t^2 - t$. Then, again (5.1) and (5.2) hold. Notice also that in [4] we used $L = e$ and $\bar{h}_{\lambda,\theta}(t) = \frac{e}{2}t^2 - t > h_{\lambda,\theta}(t)$. Hence, the present results improve the ones in [4].

Other examples where (5.1) and (5.2) are satisfied can be found in [2, 6, 9, 10, 12].

6. Conclusions

We expanded the applicability of inexact Gauss-Newton method under a majorant and a center-majorant condition. The advantages of our analysis over earlier works such as [6, 8, 13–43] 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 $n = 0, 1, 2, \dots$, 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, 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 (2007) New York, U.S.A.
- [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, D. González, Local convergence analysis of inexact Gauss-Newton method for singular systems of equations under majorant and center-majorant condition, *SeMA*, 69 (2015), 37–51.
- [5] I.K. Argyros, S. Hilout, On the Gauss-Newton method, *J. Appl. Math.* 2010 (2010), 1–14.
- [6] I.K. Argyros, S. Hilout, Extending the applicability of the Gauss-Newton method under average Lipschitz conditions, *Numer. Alg.* 58 (2011), 23–52.
- [7] I.K. Argyros, A semilocal convergence analysis for directional Newton methods, *Math. Comput.* 80 (2011), 327–343.
- [8] I.K. Argyros, S. Hilout, On the solution of systems of equations with constant rank derivatives, *Numer. Algor.* 57 (2011), 235–253.
- [9] I.K. Argyros, Y.J. Cho, S. Hilout, Numerical methods for equations and its applications, CRC Press/Taylor and Francis Group, New-York (2012).
- [10] I.K. Argyros, S. Hilout, Improved local convergence of Newton’s method under weaker majorant condition, *J. Comput. Appl. Math.* 236 (2012,) 1892–1902.
- [11] I.K. Argyros, S. Hilout, Weaker conditions for the convergence of Newton’s method, *J. Complexity* 28 (2012), 364–387.
- [12] I.K. Argyros, S. Hilout, Computational Methods in Nonlinear Analysis, World Scientific Publ. Comp., New Jersey (2013).

- [13] A. Ben-Israel, T.N.E. Greville, Generalized inverses. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 15. Springer-Verlag, New York, second edition, Theory and Applications (2003).
- [14] J.V. Burke, M.C. Ferris, A Gauss-Newton method for convex composite optimization, *Math. Program.* 71 (1995), 179–194.
- [15] J. Chen, The convergence analysis of inexact Gauss-Newton, *Comput. Optim. Appl.* 40 (2008), 97–118.
- [16] J. Chen, W. Li, Convergence behaviour of inexact Newton methods under weak Lipschitz condition, *J. Comput. Appl. Math.* 191 (2006), 143–164.
- [17] J.P. Dedieu, 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, M. Shub, Newton’s method for overdetermined systems of equations, *Math. Comput.* 69 (2000), 1099–1115.
- [19] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, *SIAM J. Numer. Anal.* 19 (1982), 400–408.
- [20] P. Deuffhard, 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, 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, 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, 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, P.R. Oliveira, Convergence of the Gauss-Newton method for a special class of systems of equations under a majorant condition, *Optimization*, <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, 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, G.P. Akilov, *Functional Analysis*, Pergamon Press, Oxford (1982).
- [30] C. Li, 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, 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, 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, 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, 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, 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.