JOURNAL OF NUMERICAL ANALYSIS AND APPROXIMATION THEORY
J. Numer. Anal. Approx. Theory, vol. 45 (2016) no. 1, pp. 3–13
ictp.acad.ro/jnaat

EXPANDING THE APPLICABILITY OF THE GAUSS-NEWTON METHOD FOR A CERTAIN CLASS OF SYSTEMS OF EQUATIONS

IOANNIS K. ARGYROS* and SANTHOSH GEORGE†

Abstract. We present a new semi-local convergence analysis of the Gauss-Newton method in order to solve a certain class of systems of equations under a majorant condition. Using a center majorant function as well as a majorant function and under the same computational cost as in earlier studies such as [11]-[13], we present a semilocal convergence analysis with the following advantages: weaker sufficient convergence conditions; tighter error estimates on the distances involved and an at least as precise information on the location of the solution. Special cases and applications complete this study.

MSC 2010. 65H10, 65G99, 65K10, 47H17, 49M15.
Keywords. Gauss-Newton method, Newton's method, semilocal convergence, least squares problem.

1. INTRODUCTION

Let $D \subseteq \mathbb{R}^n$ be open. Let $F : D \to \mathbb{R}^m$ be continuously Fréchet- differentiable. We are concerned with the problem of finding least squares solutions x^* of the nonlinear least squares problem

$$\min_{x\in D} \|F(x)\|^2. \tag{1.1}$$

The least squares solutions of (1.1) are stationary points of $G(x) = \|F(x)\|^2$. Diversified problems arising in applied sciences and in engineering can be expressed in a form like (1.1). For example in data fitting n is the number of parameters and m is the number of observations. Other examples can be found in [5, 15, 18] and the references therein. The famous Gauss-Newton method defined by

$$x_{k+1} = x_k - F'(x_k)^\dagger F(x_k), \quad k = 0, 1, \dots, \tag{1.2}$$

where x_0 is an initial point and $F'(x_k)^\dagger$ the Moore-Penrose inverse of the linear operator $F'(x_k)$ has been used extensively to generate a sequence $\{x_k\}$ converging to x^* [1]–[5], [7, 9, 19, 13, 14, 16].

[1]Department of Mathematicsal Sciences, Cameron University, Lawton, OK 73505, USA, Email: ioannisa@cameron.edu.
[2]Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, India-575 025, Email: sgeorge@nitk.ac.in.

In the present paper, we are motivated by the work of Goncalves and Oliveira in [13] (see also [11], [12]) and optimization considerations. They provided a semilocal convergence analysis for the Gauss-Newton method (1.2) for systems of nonlinear equations where the function F satisfies

$$\|F'(y)^{\dagger}(I_{\mathbb{R}^m} - F'(x)F'(x)^{\dagger})F(x)\| \leq k\|x-y\|, \quad \forall\, x, y \in D,$$

where $k \in [0,1)$ and $I_{\mathbb{R}^m}$ denotes the identity operator on $\mathbb{R}^m$. Their semilocal convergence analysis is based on the construction of a majorant function (see (h_3)). Their results unify the classical results for functions involving Lipschitz derivative [5, 6, 15, 17] with results for analytical functions (α-theory or γ-theory) [8, 10, 14, 16, 18, 19].

We introduce a center majorant function (see (h_3)) which is a special case of the majorant function that can provide more precise estimates on the distances $\|F'(x)^{\dagger}\|$. This modification leads to: weaker sufficient convergence conditions; more precise error estimates on the distances $\|x_{k+1} - x_k\|$, $\|x_k - x^*\|$ and an at least as precise information on the location of the solution.

The paper is organized as follows: Section 2 contains standard information on Moore-Penrose inverses added so that the paper should be as self contained as possible. The semilocal convergence analysis of the Gauss-Newton method is presented in Section 3. Special cases and applications are given in the concluding Section 4.

2. BACKGROUND

Let $U(x,r)$ denote the open ball with center $x \in D$ and radius $r > 0$. Let also $\overline{U(x,r)}$ denote its closure. The Moore-Penrose inverse of a linear operator $M : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is the linear operator $M^{\dagger} : \mathbb{R}^m \longrightarrow \mathbb{R}^n$ which satisfies

$$MM^{\dagger}M = M,\ M^{\dagger}MM^{\dagger} = M,\ (MM^{\dagger})^* = MM^{\dagger},\ (M^{\dagger}M)^* = M^{\dagger}M, \tag{2.1}$$

where M^* denotes the adjoint of M. We denote by $\mathrm{Ker}(M)$ and $\mathrm{Im}(M)$ the kernel and image of M. It follows from (2.1) that

$$MM^{\dagger} = \prod_{\mathrm{Ker}(M)^{\perp}}, \qquad M^{\dagger}M = \prod_{\mathrm{Im}(M),}$$

where $\prod_S$ denote the projection of $\mathbb{R}^n$ onto the subspace S. Moreover, if M is surjective, we have

$$M^{\dagger} = M^*(MM^*)^{-1}, \quad MM^{\dagger} = I_{\mathbb{R}^m}, \quad (MM^{\dagger})^{\dagger} = MM^{\dagger}.$$

We refer the reader to [7] for more properties and results on Moore-Penrose inverses.

Next, we list a number of usefull auxiliary results, starting with Banach's lemma on invertible operators [5, 7, 15].

Lemma 2.1. [7, 13] *Let $M : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be linear and continuous. Suppose that $\|M - I\| < 1$ then M is invertible and $\|M^{-1}\| \leq \frac{1}{1-\|M-I_{\mathbb{R}^n}\|}$.*

LEMMA 2.2. [7, 13] *Let* $M_1, M_2 : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be linear and continuous. Suppose that* $1 \le \operatorname{rank}(M_2) \le \operatorname{rank}(M_1)$ *and* $\|M_1^\dagger\|\|M_1 - M_2\| < 1$. *Then the following hold*

$$\operatorname{rank}(M_1) = \operatorname{rank}(M_2) \text{ and } \|M_2^\dagger\| \le \frac{\|M_1^\dagger\|}{1 - \|M_1^\dagger\|\|M_1 - M_2\|}.$$

LEMMA 2.3. [10, 13] *Let* $R > 0$. *Suppose* $\varphi : [0, R) \longrightarrow \mathbb{R}$ *is convex. Then, the following holds*

$$D^+\varphi(0) = \lim_{u \longrightarrow 0^+} \frac{\varphi(u) - \varphi(0)}{u} = \inf_{u>0} \frac{\varphi(u) - \varphi(0)}{u}.$$

LEMMA 2.4. [10, 13] *Let* $R > 0$ *and* $\theta \in [0, 1]$. *Suppose* $\varphi : [0, R) \longrightarrow \mathbb{R}$ *is convex. Then,* $\varphi_0 : [0, R) \longrightarrow \mathbb{R}$ *defined by* $\varphi_0(t) = \frac{\varphi(t) - \varphi(\theta t)}{t}$ *is increasing.*

3. SEMI-LOCAL CONVERGENCE ANALYSIS

In this section we present the semi-local convergence analysis of the Gauss-Newton method. We shall use the hypotheses given by:
(H) Let $D \subseteq \mathbb{R}^n$ be open and $F : D \to \mathbb{R}^m$ be continuously Fréchet-differentiable.
(h_1) Suppose that

$$\|F'(y)^\dagger (I_{\mathbb{R}^m} - F'(x)F'(x)^\dagger)F(x)\| \le \kappa\|x - y\|, \quad \forall x, y \in D,$$

where $\kappa \in [0, 1)$. Let $x_0 \in D$ and set $\beta = \|F'(x_0)^\dagger F(x_0)\| > 0,\ F'(x_0) \ne 0$.
(h_2) Suppose that

$$\operatorname{rank}(F'(x)) \le \operatorname{rank}(F'(x_0)) \ne 0, \quad \forall x \in D.$$

(h_3) Suppose that there exist $R > 0$ and $f_0, f : [0, R) \to \mathbb{R}$ such that $U(x_0, R) \subseteq D$,

$$\|F'(x_0)^\dagger\|\|F'(x) - F'(x_0)\| \le f_0'(\|x - x_0\|) - f_0'(0), \quad \forall x \in D$$

and

$$\|F'(x_0)^\dagger\|\|F'(y) - F'(x)\| \le f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|) \quad \forall x, y \in U(x_0, R)$$

with $\|y - x\| + \|x - x_0\| < R$.
(h_4)

$$f_0(0) = f(0) = 0,\ f'(0) = f_0'(0) = -1$$
$$f_0(t) \le f(t) \text{ and } f_0'(t) \le f'(t), \quad \forall t \in [0, R).$$

(h_5) f_0', f' are convex and strictly increasing.

Let $\lambda \ge 0$ be such that $\lambda \ge -\kappa f'(\beta)$ and define $h_{\beta,\lambda} : [0, R) \to \mathbb{R}$ by

$$h_{\beta,\lambda} = \beta + \lambda t + f(t).$$

(h_6) $h_{\beta,\lambda} = 0$ for some $t \in [0, R)$.
(h_7) For each $s, t, u \in [0, R)$ with $s \leq t \leq u$

$$t + \frac{h_{\beta,\lambda}(u)}{f_0'(u)} \leq u + \frac{h_{\beta,\lambda}(t) - h_{\beta,\lambda}(s) - h'_{\beta,\lambda}(s)(t-s)}{f_0'(t)}$$

Remark 3.1. If $f_0 = f$, then hypotheses (H) and our results reduce to the ones given in [13]. Notice that the second hypotheses in (h_3) implies the first one but not necessarily vice versa. That is the first hypotheses is a special case of the second. From the computational point of view, the computation of function f involves the computation of function f_0. Hence, the results in this study are given under the same computational cost as in [13] (since the rest of the (H) hypotheses are the same as in [13]). From now on we assume the (H) conditions hold.

The majorizing iteration $\{s_k\}$ for $\{x_k\}$ is given by

$$s_0 = 0, \quad s_{k+1} = s_k - \frac{h_{\beta,\lambda}(s_k)}{f_0'(s_k)}. \tag{3.1}$$

The corresponding iteration $\{t_n\}$ used in [13] (for $f_0 = f$) is given by

$$t_0 = 0, \quad t_{k+1} = t_k - \frac{h_{\beta,\lambda}(t_k)}{f'(t_k)}. \tag{3.2}$$

The proofs in this study also for each $k = 0, 1, 2, \ldots$ (see Lemma 11 in [13] and our Lemma 3.8) shall show that the following iterations $\{r_k\}$ and $\{q_k\}$ are also majorizing for $\{x_k\}$:

$$\begin{aligned} & r_0 = 0,\ r_1 = \beta, \\ & r_{k+2} = r_{k+1} - \frac{h_{\beta,\lambda}(r_{k+1}) - h_{\beta,\lambda}(r_k) - h'_{\beta,\lambda}(r_k)(r_{k+1} - r_k)}{f_0'(r_{k+1})}, \quad k = 0, 1, 2, \ldots, \end{aligned} \tag{3.3}$$

and for $g_{\beta,\lambda} = \beta + \lambda t + f_0(t)$,

$$\begin{aligned} & q_0 = 0,\ q_1 = \beta,\ q_2 = q_1 - \frac{g_{\beta,0}(q_1) - g_{\beta,0}(q_0) - g'_{\beta,0}(q_0)(q_1 - q_0)}{f_0'(q_1)} \\ & q_{k+2} = q_{k+1} - \frac{h_{\beta,\lambda}(q_{k+1}) - h_{\beta,\lambda}(q_k) - h'_{\beta,\lambda}(q_k)(q_{k+1} - q_k)}{f_0'(q_{k+1})}, \quad k = 1, 2, \ldots. \end{aligned} \tag{3.4}$$

New majorizing sequence $\{s_k\}, \{r_k\}, \{q_k\}$ can not only be tighter than $\{t_k\}$ but their sufficient convergence conditions can be weaker than the corresponding ones for $\{t_k\}$. $\square$

Next, we first study the convergence of iteration $\{s_k\}$ and the properties of majorizing functions. The proofs are analogous to the corresponding ones in [13]. We simply replace $f'(t)$ by $f_0'(t)$ in the proofs and use $f_0'(t) \leq f'(t)$. However, there are some differences in the proofs which are not immediate to the reader. Therefore, in what follows we will point out those differences. As already noted above for the rest of the differences, we refer the reader to [13].

Proposition 3.2. [13] *The following hold*

(i) $h_{\beta,\lambda}(0) = \beta,\ h'_{\beta,\lambda}(0) = \lambda^{-1}$;
(ii) $h'_{\beta,\lambda}$ *is convex and strictly increasing.*

Proposition 3.3. *The function $h_{\beta,\lambda}$ has a smallest zero $s^* \in [0, R)$, is strictly convex, and the following hold:*

$$h_{\beta,\lambda}(t) > 0,\ f_0'(t) < 0,\ t < t - \frac{h_{\beta,\lambda}(t)}{f_0'(t)} < t^*, \quad \forall t \in [0, t^*). \tag{3.5}$$

Moreover, $f_0'(t^*) \le h_{\beta,\lambda}(t^*) = f'(t^*) \le 0.$

Proof. The proof until the first estimate in (3.5) have been given in [13]. It was also shown that $f'(t) = h'_{\beta,0}(t) < 0$. But by hypotheses (h_4) $f_0'(t) \le f'(t)$. Hence, we get $f_0'(t) < 0$. By convexity of $h_{\beta,\lambda}$, we have that

$$0 = h_{\beta,\lambda}(t^*) > h_{\beta,\lambda}(t) + h'_{\beta,\lambda}(t)(t^* - t), \quad \forall t \in [0, R),\, t \ne t^*.$$

Then, the preceeding inequality can be written as

$$t - \frac{h_{\beta,\lambda}(t)}{h'_{\beta,\lambda}(t)} < t^*, \quad \forall t \in [0, t^*),$$

(since $-h'_{\beta,\lambda}(t) > 0$). We also have that

$$0 < -h'_{\beta,\lambda}(t) \le -h'_{\beta,0}(t) = -f'(t) \le -f_0'(t),$$

which together with the preceeding estimates shows the third estimate in (3.5). Moreover, we have that $h_{\beta,\lambda}(t) > 0$ for each $t \in [0, t^*)$ and $h_{\beta,\lambda}(t^*) = 0$. Then, we must have $h'_{\beta,\lambda}(t) \le 0$ for each $t \in [0, t^*)$. Hence, the last estimate follows from

$$0 \ge h'_{\beta,\lambda}(t^*) = \lambda + h'_{\beta,0}(t^*) = \lambda + f'(t^*) \ge f_0'(t^*).$$

□

It follows from the second estimate in (3.5) that the map $\psi_{h_{\beta,\lambda}}$:

$$\psi_{h_{\beta,\lambda}} : [0, t^*) \to \mathbb{R}, \quad t \to t - \frac{\psi_{h_{\beta,\lambda}}}{f_0'(t)}$$

is well defined. Notice that if $f_0 = f$ this map reduces to the one used in [13]. Moreover, if $\lambda = 0$ this map reduces to the one given in [11].

We can define sequence $\{s_k\}$ given in (3.1) by

$$s_0 = 0,\ s_{k+1} = \psi_{h_{\beta,\lambda}}(s_k), \quad k = 0, 1, 2, \ldots. \tag{3.6}$$

Proposition 3.4. *The following hold*

(a) $\beta \le \psi_{h_{\beta,\lambda}}(t) < t^*, \forall t \in [0, t^*)$.

(b) *The map* $\psi_{h_{\beta,\lambda}}$ *maps* $[0,t^*)$ *in* $[0,t^*)$ *and* $t < \psi_{h_{\beta,\lambda}}(t)$, for each $t \in [0,t^*)$. *If* $\lambda = 0$ *or* $\lambda = 0$ *and* $f_0'(t^*) < 0$, *then, the following hold, respectively*

$$\begin{aligned} t^* - \psi_{h_{\beta,\lambda}}(t) &\leq \tfrac{1}{2}(t^* - t), \\ t^* - \psi_{h_{\beta,\lambda}}(t) &\leq \frac{D^- f_0'(t^*)}{-2f_0'(t^*)}(t^* - t)^2, \text{ for each } t \in [0,t^*) \end{aligned}$$

where D^- *stands for the left directional derivative* [5, 13].

(c) *The sequence* $\{s_k\}$ *is: well defined; strictly increasing; contained in* $[0,t^*)$ *and converges to* t^*.

If $\lambda = 0$ *or* $\lambda = 0$ *and* $f_0'(t^*) < 0$, *the sequence* $\{s_k\}$ *converges Q-linearly or Q-quadratically to* t^*, *respectively and*

$$\begin{aligned} t^* - s_{k+1} &\leq \tfrac{1}{2}(t^* - s_k), \\ s^* - s_{k+1} &\leq \frac{D^- f_0'(t^*)}{-2f_0'(t^*)}(t^* - s_k)^2, \text{ for each } k = 0,1,2,\ldots. \end{aligned}$$

In what follows we prove the convergence of Gauss-Newton method (1.2). We need the following Banach type Lemma:

Proposition 3.5. *Suppose that* $x \in U(x_0, t)$ *for* $t \in [0,t^*)$. *Then, the following hold* $\operatorname{rank}(F'(x)) = \operatorname{rank}(F'(x_0)) \geq 1$ *and*

$$\|F'(x)^\dagger\| \leq -\frac{\|F'(x_0)^\dagger\|}{f_0'(t)}. \tag{3.7}$$

Proof. Using (h_2), (h_4), (h_5), the first hypothesis in (h_3) and the second hypothesis in (3.5), we obtain in turn that

$$\begin{aligned} \|F'(x_0)^\dagger\|\|F'(x) - F'(x_0)\| &\leq f_0'(\|x - x_0\|) - f_0'(0) \\ &\leq f_0'(t) + 1 = h_{\beta,0}'(t) + 1 < 1. \end{aligned}$$

In view of the preceeding estimate and Lemma 2.2 we deduce that $\operatorname{rank}(F'(x)) = \operatorname{rank}(F'(x_0)) \geq 1$ and

$$\|F'(x)^\dagger\| \leq -\frac{\|F'(x_0)^\dagger\|}{1 - (f_0'(t) + 1)} = -\frac{\|F'(x_0)^\dagger\|}{f_0'(t)}.$$

$\square$

Remark 3.6. It is worth noticing that the second inequality in (h_3) is used in [13] to obtain instead of (3.7) that

$$\|F'(x)^\dagger\| \leq -\frac{\|F'(x_0)^\dagger\|}{f'(t)} \tag{3.8}$$

which is more expensive to arrive at and less precise than (3.7) if $f_0'(t) < f'(t)$ for each $t \in [0,t^*)$. This important observation makes the main difference in our approach over the one used in [13] and the earlier studies [11, 12, 14, 16, 19].

As in earlier studies of this type [4, 5, 6, 9, 10, 11, 13] we study the linearization error of F at each point in D by defining

$$E_F(x,y) := F(y) - [F(x) + F'(x)(y-x)], \quad \forall x, y \in D \tag{3.9}$$

and the error in the linearization on majorant function f by

$$e_f(t,u) := f(u) - [f(t) + f'(t)(u-t)], \quad \forall t, u \in [0,R). \tag{3.10}$$

□

Lemma 3.7. [11] *Let $x, y \in U(x_0,R)$ and $0 \le t < r < R$. Suppose that $\|x - x_0\| \le t$ and $\|y - x\| \le r - t$. Then,*

$$\|F'(x_0)^\dagger\| \|E_F(x,y)\| \le e_f(t,u) \frac{\|y-x\|^2}{(r-t)^2}. \tag{3.11}$$

It follows from Proposition 3.5 that Gauss-Newton map G_F for F :

$$\begin{aligned} G_F : U(x_0,t^*) &\longrightarrow \mathbb{R}^n \\ x &\longrightarrow x - F'(x)^\dagger F(x) \end{aligned} \tag{3.12}$$

is well-defined. In order to guarantee that the Gauss-Newton iteration can be repeated indefinitely we need to introduce certain sets:

$$K_0(t) := \left\{ x \in D : \|x - x_0\| \le t,\ \|F'(x)^\dagger F(x)\| \le -\frac{h_{\beta,\lambda}(t)}{f_0'(t)} \right\} \tag{3.13}$$

$$K_0 = \bigcup_{t \in [0,t^*)} K_0(t), \tag{3.14}$$

$$K(t) := \left\{ x \in D : \|x - x_0\| \le t,\ \|F'(x)^\dagger F(x)\| \le -\frac{h_{\beta,\lambda}(t)}{f'(t)} \right\} \tag{3.15}$$

and

$$K = \bigcup_{t \in [0,t^*)} K(t). \tag{3.16}$$

The sets $K(t)$ and K were defined in [11, 13]. It follows from the estimate $f_0'(t) \le f'(t)$ for each $t \in [0,t^*)$ and definitions (3.13)-(3.16) that

$$K(t) \subset K_0(t) \tag{3.17}$$

and

$$K \subset K_0. \tag{3.18}$$

In view of (3.17) and (3.18) the next two results improve the corresponding ones in [13]. The proofs are given in exactly the same way but replacing $K(t), K, \{t_n\}, h'_{\beta,0}$ by $K_0(t), K_0, \{s_n\}, f_0'$, respectively.

Lemma 3.8. *The following hold for each $t \in [0,t^*)$:*

(i) $K_0(t) \subset U(x_0,t^*)$;

(ii) $\|G_F(G_F(x)) - G_F(x)\| \le -\frac{h_{\beta,\lambda}(\psi_{h_{\beta,\lambda}}(t))}{f_0'(\psi_{h_{\beta,\lambda}}(t))} \left(\frac{\|G_F(x)-x\|}{\psi_{h_{\beta,\lambda}}(t)-t}\right)^2$, $\forall x \in K_0(t)$

(iii) $G_F(K_0(t)) \subset K_0(\psi_{h_{\beta,\lambda}}(t))$
(iv) $K_0 \subset U(x_0, t^*)$ *and* $G_F(K_0) \subset K_0$.

The Gauss-Newton method (1.2) can be written as

$$x_{k+1} = G_F(x_k), \ k = 0, 1, 2, \ldots. \tag{3.19}$$

Next, the main semi-local convergence result for the Gauss-Newton method is presented.

Theorem 3.9. *Suppose that the* (H) *conditions hold. Then, the following hold:*
$h_{\beta,\lambda}(t)$ *has a smallest zero* $t^* \in (0, R)$, *the sequences* $\{s_k\}$ *and* $\{x_k\}$ *for solving* $h_{\beta,\lambda}(t) = 0$ *and* $F(x) = 0$, *with starting point* $t_0 = 0$ *and* x_0, *respectively given by* (3.2) *and* (3.19) *are well defined,* $\{s_k\}$ *is strictly increasing, remains in* $[0, t^*)$, *and converges to* t^*, $\{x_k\}$ *remains in* $U(x_0, t^*)$, *converges to a point* $x^* \in U(x_0, t^*)$ *such that* $F'(x^*)^\dagger F(x^*) = 0$. *Moreover, the following estimates hold:*

$$\|x_{k+1} - x_k\| \leq s_{k+1} - s_k, \quad k = 0, 1, 2, \ldots,$$
$$\|x^* - x_k\| \leq t^* - s_k, \quad k = 0, 1, 2, \ldots,$$

and

$$\|x_{k+1} - x_k\| \leq \frac{s_{k+1} - s_k}{(s_k - s_{k-1})^2} \|x_k - x_{k-1}\|^2, \quad k = 0, 1, 2, \ldots.$$

Furthermore, if $\lambda = 0$ *(*$\lambda = 0$ *and* $f_0'(t^*) < 0$*), the sequence* $\{s_k\}$, $\{x_k\}$ *converge Q-linearly and R-linearly (Q-quadratically and R-quadratically) to* t^* *and* x^*, *respectively.*

Remark 3.10. (i) As noted in [13] the best choice for λ is given by $\lambda = -\kappa f'(\kappa)$.

(ii) Sequence $\{r_k\}$ appears in the proof of Lemma 11 in [13] or our Lemma 3.8 (see also Theorem 4.1).

(iii) Sequence $\{q_k\}$ is derived from the proof of Lemma 11 in [13] or our Lemma 3.8 if we simply use $x = x_0$ and notice that the first instead of the second hypothesis in (h_3) can be used for the corresponding estimate. Notice that $\{q_k\}$ is the tightest among $\{s_k\}$ and $\{r_k\}$. Moreover $\{s_k\}$, $\{t_k\}$ and $\{q_k\}$ converges under the (H) conditions. □

4. SPECIAL CASES AND APPLICATIONS

In this section, we present some special cases and applications. Then, we arrive at:

Theorem 4.1. *Under the* (H) *hypotheses, the conclusions of Theorem* 3.9 *hold. Moreover, the following estimates hold*

$$\|x_{k+1} - x_k\| \leq q_{k+1} - q_k \leq r_{k+1} - r_k \leq s_{k+1} - s_k < t_{k+1} - t_k, \quad k = 0, 1, 2, \ldots \tag{4.1}$$

Proof. The conclusions of Theorem 3.9 hold under the (H) conditions. Then (4.1) follows using the definition of sequences $\{q_k\}$, $\{r_k\}$, $\{q_k\}$, $\{t_k\}$, (h_7) a simple induction argument and the discussion in Remark 3.10 (ii) and (iii). Indeed, we have that $s_0 = r_0$ and $s_1 = r_1 = \beta$. Then it follows from (3.1) and (3.3) for $\kappa = 0$ and (h_7) that

$$\begin{aligned} r_2 &= r_1 - \frac{h_{\beta,\lambda}(r_1) - h_{\beta,\lambda}(r_0) - h'_{\beta,\lambda}(r_0)(r_1 - r_0)}{f'_0(r_1)} \\ &= s_1 - \frac{h_{\beta,\lambda}(s_1) - h_{\beta,\lambda}(s_0) - h'_{\beta,\lambda}(s_0)(s_1 - s_0)}{f'_0(s_1)} \\ &\le s_1 - \frac{h_{\beta,\lambda}(s_1)}{f'_0(s_1)} = s_2 \end{aligned}$$

and

$$r_2 - r_1 = -\frac{h_{\beta,\lambda}(s_1) - h_{\beta,\lambda}(s_0) - h'_{\beta,\lambda}(s_0)(s_1 - s_0)}{f'_0(s_1)} \le s_2 - s_1.$$

Suppose that $r_i \le s_i$ and $r_{i+1} - r_i \le s_{i+1} - s_i$ for $i = 0, 1, 2, \cdots k$. We shall show that $r_{k+2} \le s_{k+2}$ and $r_{k+2} - r_{k+1} \le s_{k+2} - s_{k+1}$. We have from the induction hypotheses and (h_7) that

$$\begin{aligned} r_{k+2} &= r_{k+1} - \frac{h_{\beta,\lambda}(r_{k+1}) - h_{\beta,\lambda}(r_k) - h'_{\beta,\lambda}(r_k)(r_{k+1} - r_k)}{f'_0(r_{k+1})} \\ &\le s_{k+1} - \frac{h_{\beta,\lambda}(s_{k+1})}{f'_0(s_{k+1})} = s_{k+2} \end{aligned}$$

and

$$r_{k+2} - r_{k+1} \le s_{k+2} - s_{k+1}.$$

Similarly, we show that $q_k \le r_k$ and $q_{k+1} - q_k \le r_{k+1} - r_k$. □

So, far we have shown that under the (H) and (h_7) hypotheses more precise majorizing sequences can be obtained for $\{x_k\}$. At this point we are wondering if a direct study of the convergence of majorizing sequences $\{r_k\}, \{s_k\}, \{q_k\}$ lead to weaker sufficient convergence conditions than (h_6). Let us present this in the interesting case of Newton's method

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k), \quad k = 0, 1, 2, \ldots \tag{4.2}$$

for solving nonlinear equation

$$F(x) = 0. \tag{4.3}$$

That is $\lambda = \kappa = 0$. Let us define f_0 and f on $[0, R)$ by $f_0(t) = \frac{L_0}{2} t^2 - t$ and $f(t) = \frac{L}{2} t^2 - t$ for $L_0 > 0$ and $L > 0$. Then, sequences $\{t_k\}, \{s_k\}, \{r_k\}$ and

$\{q_k\}$ reduce to

$$\begin{aligned}
t_0 &= 0,\ t_{k+1} = t_k - \frac{\frac{L}{2}t_k^2 - t_k + \beta}{L_0 t_k - 1} \\
&= t_k - \frac{L(t_k - t_{k-1})^2}{2(Lt_k - 1)}, \\
s_0 &= 0, s_1 = \beta, s_{k+1} = s_k - \frac{\frac{L}{2}s_k^2 - s_k + \beta}{L_0 s_k - 1} \\
r_0 &= 0, r_1 = \beta, \\
r_{k+2} &= r_{k+1} - \frac{\frac{L}{2}r_{k+1}^2 - r_{k+1} - (\frac{L}{2}r_k^2 - r_k) - (Lr_k - 1)(r_{k+1} - r_k)}{L_0 r_{k+1} - 1} \\
q_0 &= 0, q_1 = \beta,\ q_2 = q_1 - \frac{L_0(q_1 - q_0)^2}{2(L_0 q_1 - 1)}, \\
q_{k+2} &= q_{k+1} - \frac{L(q_{k+1} - q_k)^2}{2(L_0 q_{k+1} - 1)}.
\end{aligned}$$

Then, according to (h_3), and Theorem 4.1 sequences $\{t_k\}, \{s_k\}, \{r_k\}, \{q_k\}$ converge, if the famous for its simplicity and clarity Kantorovich hypothesis [5, 15]

$$h = L\beta \le \tfrac{1}{2}$$

is satisfied. We also have that

$$t^* = \frac{1 - \sqrt{1 - 2h}}{L}.$$

However, a direct study of sequence $\{q_k\}$ (see [5, 6]) shows that this sequence converges provided that

$$h_0 = \bar{L}\beta \le \tfrac{1}{2},$$

where

$$\bar{L} = \tfrac{1}{8}\Big(4L + \sqrt{8L^2 + L_0 L} + \sqrt{L_0 L}\Big)$$

Notice that since $\bar{L} \le L$,

$$h \le \tfrac{1}{2} \Rightarrow h_0 \le \tfrac{1}{2}$$

but not necessarily vice versa unless if $L_0 = L$. Moreover, we have that

$$\frac{h_0}{h} \to 0, \quad \text{as } \frac{L}{L_0} \to 0. \tag{4.4}$$

Implication (4.4) shows by how many times (at most) the applicability of Newton's method (4.2) is extended under our approach. Notice also that if $L_0 < L$

$$q_k < t_k, \quad \text{for each } k = 2, 3, \dots,$$

$$q_{k+1} - q_k < t_{k+1} - t_k, \quad \text{for each } k = 2, 3, \dots,$$

and $q^* = \lim_{k\to\infty} q_k \le t^*$. Examples, where $L_0 < L$ can be found in [5, 6].

REFERENCES

[1] I. Argyros, *On the semilocal convergence of the Gauss-Newton method*, Adv. Nonlinear Var. Inequal., **8** (2005) 2, pp. 93–99.

[2] I. Argyros and S. Hilout, *On the local convergence of the Gauss-Newton method*, Punjab Univ. J. Math., **41** (2009), pp. 23–33.

[3] I. Argyros and S. Hilout, *On the Gauss-Newton method*, J. Appl. Math. Comput., **35** (2011), pp. 537–550.

[4] I. Argyros and S. Hilout, *Extending the applicability of the Gauss-Newton method under average Lipschitz-type conditions*, Numer. Algor., **58** (2011) 1, pp. 23–52.

[5] I. Argyros and S. Hilout, *Improved local convergence of Newton's method under weak majorant condition*, J. Comp. Appl. Math., **236** (2012) 7, pp. 1892–1902.

[6] I. Argyros and S. Hilout, *Numerical methods in nonlinear analysis*, World Scientific Publ. Comp. New Jersey, USA, 2013.

[7] 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, 2nd ed., Theory and Applications, 2003.

[8] E. Cătinaş, *The inexact, inexact perturbed, and quasi-Newton methods are equivalent models*, Math. Comput., **74** (2005) 249, pp. 291–301.

[9] J.P. Dedieu and M.H. Kim, *Newton's method for analytic systems of equations with constant rank derivatives*, J. Complexity, **18**, (2002) 1, pp. 187–209.

[10] O.P. Ferreira, M.L.N. Gonçalves and P.R. Oliveira, *Local convergence analysis of inexact Gauss-Newton like methods under majorant condition*, J. Complexity, **27** (2011) 1, pp. 111–125.

[11] O.P. Ferreira and B.F. Svaiter, *Kantorovich's majorants principle for Newton's method*, Comput. Optim. Appl., **42**(2009) 2, pp. 213–229.

[12] W.M. Häussler, *A Kantorovich-type convergence analysis for the Gauss-Newton-method*, Numer. Math., **48** (1986) 1, pp. 119–125.

[13] 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*, Optimization, **64** (2015) 3, pp. 577–594.

[14] N. Hu, W. Shen and C. Li, *Kantorovich's type theorems for systems of equations with constant rank derivatives*, J. Comput. Appl. Math., **219** (2008) 1, pp. 110–122.

[15] L.V. Kantorovich and G.P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.

[16] C. Li, N. Hu and J. Wang, *Convergence behavior of Gauss-Newton's method and extensions of the Smale point estimate theory.* J. Complexity, **26** (2010) 3, pp. 268–295.

[17] F.A. Potra and V. Ptak, *Nondiscrete induction and iterative processes*, Research notes in Mathematics, 103, Pitman (Advanced Publishing Program), Boston, MA, 1984.

[18] 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), pp. 185–196, Springer, New York, 1986.

[19] X.H. Wang, *Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces*, IMA J. Numer. Anal., **20** (2000), pp. 123–134.

Received by the editors: November 28, 2013.