Nonlinear Functional Analysis and Applications
Vol. 22, No. 1 (2017), pp. 197-207
ISSN: 1229-1595(print), 2466-0973(online)

http://nfaa.kyungnam.ac.kr/journal-nfaa

EXPANDING THE APPLICABILITY OF THE GAUSS-NEWTON METHOD FOR CONVEX OPTIMIZATION UNDER RESTRICTED CONVERGENCE DOMAINS AND MAJORANT CONDITIONS

Ioannis K. Argyros[1] and Santhosh George[2]

[1]Department of Mathematicsal Sciences
Cameron University, Lawton, OK 73505, USA
e-mail: `ioannisa@cameron.edu`

[2]Department of Mathematical and Computational Sciences
National Institute of Technology Karnataka, India
e-mail: `sgeorge@nitk.ac.in`

Abstract. Using our new idea of restricted convergent domains, new semi-local convergence analysis of the Gauss-Newton method for solving convex composite optimization problems is presented. Our convergence analysis is based on a combination of a center-majorant and majorant function. The results extend the applicability of the Gauss-Newton method under the same computational cost as in earlier studies using a majorant function or Wang's condition or Lipchitz condition. The special cases and applications include regular starting points, Robinson's conditions, Smale's or Wang's theory.

1. Introduction

In this study we are concerned with the convex composite optimizations problem. This work is mainly motivated by the work in [15, 23]. We present a convergence analysis of Gauss–Newton method (defined by Algorithm (GNA) in Sect. 2). The convergence of GNA is based on the majorant function in [15] (to be precised in Sect. 2).

They follow the same formulation using the majorant function provided in [23](see, [21, 23, 28, 29]). In [3, 7, 8], a convergence analysis in a Banach space

[0]Received April 22, 2016. Revised July 18, 2016.
[0]2010 Mathematics Subject Classification: 90C30, 65G99, 65K10.
[0]Keywords: Gauss-Newton method, convex composite optimization problem, semi-local convergence, restricted convergeny domains, majorant function, center-majorant function.

setting was given for (GNM) defined by

$$x_{k+1} = x_k - \left[F'(x_k)^+ F'(x_k)\right]^{-1} F'(x_k)^+ F(x_k) \quad \text{for each} \quad k = 0, 1, 2, \cdots,$$

where x_0 is an initial point and $F'(x)^+$ in the Moore-Penrose inverse [11, 12, 13, 19, 26] of operator $F'(x)$ with $F : \mathbb{R}^n \to \mathbb{R}^m$ being continuously differentiable. In [23], a semilocal convergence analysis using a combination of a majorant and a center majorant function was given with the advantages $(\mathcal{A})$: tighter error estimates on the distances involved and the information on the location of the solution is at least as precise. These advantages were obtained (under the same computational cost) using same or weaker sufficient convergence hypotheses. Here, we extend the same advantages $(\mathcal{A})$ but to hold for GNA.

The study is organized as follows: Section 2 contains the definition of GNA. In order for us to make the paper as self contained as possible, the notion of quasi-regularity is also re-introduced (see, e.g., [12, 15, 21]). The semilocal convergence analysis of GNA is presented in Section 3. Numerical examples and applications of our theoretical results and favorable comparisons to earlier studies (see, e.g., [12, 15, 18, 21, 22]) are presented in Section 4.

2. Gauss-Newton Algorithm and Quasi-Regularity condition

2.1. Gauss-Newton Algorithm GNA. Using the idea of restricted convergence domains, we study the convex composite optimization problem

$$\min_{x \in \mathbb{R}^n} p(x) := h(F(x)), \tag{2.1}$$

where $h : \mathbb{R}^m \longrightarrow \mathbb{R}$ is convex, $F : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is Fréchet-differentiable operator and $m, l \in \mathbb{N}^\star$. The importants of (2.1) can be found in [2, 10, 12, 19, 21, 22, 23, 25, 27]. We assume that the minimum h_{min} of the function h is attained. Problem (2.1) is related to

$$F(x) \in \mathcal{C}, \tag{2.2}$$

where

$$\mathcal{C} = \operatorname{argmin} h \tag{2.3}$$

is the set of all minimum points of h.

Let $\xi \in [1, \infty[$, $\Delta \in]0, \infty]$ and for each $x \in \mathbb{R}^n$, define $\mathcal{D}_\Delta(x)$ by

$$\begin{aligned}\mathcal{D}_\Delta(x) = \{d \in \mathbb{R}^n :& \| d \| \leq \Delta,\ h(F(x) + F'(x)\, d) \leq h(F(x) + F'(x)\, d') \\ &\text{for all } d' \in \mathbb{R}^n \text{ with } \| d' \| \leq \Delta\}.\end{aligned} \tag{2.4}$$

Let $x_0 \in \mathbb{R}^n$ be an initial point. The Gauss-Newton algorithm GNA associated with (ξ, Δ, x_0) as defined in [12] (see also [15]) is as follows:

Algorithm GNA : (ξ, Δ, x_0)

INICIALIZATION. Take $\xi \in [1, \infty)$, $\Delta \in (0, \infty]$ and $x_0 \in \mathbb{R}^n$, set $k = 0$.
STOP CRITERION. Compute $D_\Delta(x_k)$. If $0 \in \mathcal{D}_\Delta(x_k)$, STOP. Otherwise.
ITERATIVE STEP. Compute d_k satisfying $d_k \in D_\Delta(x_k)$,

$$\|d_k\| \leq \xi d(0, D\Delta(x_k)),$$

Then, set $x_{k+1} = x_k + d_k$, $k = k + 1$ and GO TO STOP CRITERION.

Here, $d(x, W)$ denotes the distance from x to W in the finite dimensional Banach space containing W. Note that the set $\mathcal{D}_\Delta(x)$ $(x \in \mathbb{R}^n)$ is nonempty and is the solution of the following convex optimization problem

$$\min_{d \in \mathbb{R}^n, \, \|d\| \leq \Delta} h(F(x) + F'(x)\, d), \tag{2.5}$$

which can be solved by well known methods such as the subgradient or cutting plane or bundle methods (see, e.g., [12, 19, 25, 26, 27]).

Let $U(x, r)$ denote the open ball in $\mathbb{R}^n$ (or $\mathbb{R}^m$) centered at x and of radius $r > 0$. By $\overline{U}(x, r)$ we denote its closure. Let W be a closed convex subset of $\mathbb{R}^n$ (or $\mathbb{R}^m$). The negative polar of W denoted by $W^\ominus$ is defined as

$$W^\ominus = \{z \, :< z, w > \leq 0 \quad \text{for each} \quad w \in W\}. \tag{2.6}$$

2.2. **Quasi Regularity.** In this section, we mention some concepts and results on regularities which can be found in [12] (see also, e.g., [15, 21, 22, 23, 25]). For a set-valued mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and for a set A in $\mathbb{R}^n$ or $\mathbb{R}^m$, we denote by

$$D(T) = \{x \in \mathbb{R}^n \, : \, Tx \neq \emptyset\}, \quad R(T) = \bigcup_{x \in D(T)} Tx, \tag{2.7}$$

$$T^{-1}y = \{x \in \mathbb{R}^n \, : \, y \in Tx\} \quad \text{and} \quad \parallel A \parallel = \inf_{a \in A} \parallel a \parallel .$$

Consider the inclusion

$$F(x) \in C, \tag{2.8}$$

where C is a closed convex set in $\mathbb{R}^m$. Let $x \in \mathbb{R}^n$ and

$$\mathcal{D}(x) = \{d \in \mathbb{R}^n \, : \, F(x) + F'(x)\, d \in C\}. \tag{2.9}$$

Definition 2.1. Let $x_0 \in \mathbb{R}^n$.

(a) x_0 is quasi-regular point of (2.8) if there exist $R_0 \in]0, +\infty[$ and an increasing positive function β on $[0, R_0[$ such that

$$\mathcal{D}(x) \neq \emptyset \text{ and } d(0, \mathcal{D}(x)) \leq \beta(\parallel x - x_0 \parallel)\, d(F(x), C) \tag{2.10}$$

for all $x \in U(x_0, R_0)$, where $\beta(\parallel x - x_0 \parallel)$ is an "error bound" in determining how for the origin is away from the solution set of (2.8).

(b) x_0 is a regular point of (2.8) if

$$ker(F'(x_0)^T) \cap (C - F(x_0))^{\ominus} = \{0\}. \tag{2.11}$$

Proposition 2.2. (see, e.g., [12, 15, 21, 25]) *Let x_0 be a regular point of* (2.8). *Then, there are constants $R_0 > 0$ and $\beta > 0$ such that* (2.10) *holds for R_0 and $\beta(\cdot) = \beta$. Therefore, x_0 is a quasi-regular point with the quasi-regular radius $R_{x_0} \geq R_0$ and the quasi-regular bound function $\beta_{x_0} \leq \beta$ on $[0, R_0]$.*

Remark 2.3. (a) $\mathcal{D}(x)$ can be considered as the solution set of the linearized problem associated to (2.8)

$$F(x) + F'(x)\,d \in C. \tag{2.12}$$

(b) If C defined in (2.8) is the set of all minimum points of h and if there exists $d_0 \in \mathcal{D}(x)$ with $\| d_0 \| \leq \Delta$, then $d_0 \in \mathcal{D}_\Delta(x)$ and for each $d \in \mathbb{R}^n$, we have the following equivalence

$$d \in \mathcal{D}_\Delta(x) \Longleftrightarrow d \in \mathcal{D}(x) \Longleftrightarrow d \in \mathcal{D}_\infty(x). \tag{2.13}$$

(c) Let R_{x_0} denote the supremum of R_0 such that (2.10) holds for some function β defined in Definition 2.1. Let $R_0 \in [0, R_{x_0}]$ and $\mathcal{B}_R(x_0)$ denotes the set of function β defined on $[0, R_0)$ such that (2.10) holds. Define

$$\beta_{x_0}(t) = \inf\{\beta(t) \,:\, \beta \in \mathcal{B}_{R_{x_0}}(x_0)\} \text{ for each } t \in [0, R_{x_0}). \tag{2.14}$$

All function $\beta \in \mathcal{B}_R(x_0)$ with $\lim\limits_{t \to R^-} \beta(t) < +\infty$ can be extended to an element of $\mathcal{B}_{R_{x_0}}(x_0)$ and we have that

$$\beta_{x_0}(t) = \inf\{\beta(t) \,:\, \beta \in \mathcal{B}_R(x_0)\} \text{ for each } t \in [0, R_0). \tag{2.15}$$

R_{x_0} and β_{x_0} are called the quasi-regular radius and the quasi-regular function of the quasi-regular point x_0, respectively.

Definition 2.4. (a) A set-valued mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is convex if the following items hold

(i) $Tx + Ty \subseteq T(x + y)$ for all $x, y \in \mathbb{R}^n$.
(ii) $T\lambda x = \lambda Tx$ for all $\lambda > 0$ and $x \in \mathbb{R}^n$.
(iii) $0 \in T0$.

3. Semi-local convergence

In this section we present the semi-local convergence of GNA. First, we study the convergence of majorizing sequences for GNA. Then, we study the convergence of GNA. We need the definition of the center-majorant function and the definition of the majorant function for F.

Definition 3.1. Let $R > 0$, $x_0 \in \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ be continuously Fréchet-differentiable. A twice-differentiable function $f_0 : [0, R) \to \mathbb{R}$ is called a center-majorant function for F on $U(x_0, R)$, if for each $x \in U(x_0, R)$,

(h_0^0) $\|F'(x) - F'(x_0)\| \leq f_0'(\|x - x_0\|) - f_0'(0)$;
(h_1^0) $f_0(0) = 0$, $f_0'(0) = -1$; and
(h_2^0) f_0' is convex and strictly increasing.

Definition 3.2. ([7, 8, 15]) Let $x_0 \in \mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^m$ be continuously differentiable. Define $R_0 = \sup\{t \in [0, R) : f_0'(t) < 0\}$. A twice-differentiable function $f : [0, R_0) \to \mathbb{R}$ is called a majorant function for F on $U(x_0, R_0)$, if for each $x, y \in U(x_0, R_0)$, $\|x - x_0\| + \|y - x\| < R_0$,

(h_0) $\|F'(y) - F'(x)\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|)$;
(h_1) $f(0) = 0$, $f'(0) = -1$;
and
(h_2) f' is convex and strictly increasing.

Moreover, assume that

(h_3) $f_0(t) \leq f(t)$ and $f_0'(t) \leq f'(t)$ for each $t \in [0, R_0)$.

Remark 3.3. Suppose that $R_0 < R$. If $R_0 \geq R$, then we do not need to introduce Definition 3.2.

In Section 4, we present examples where hypothesis (h_3) is satisfied. Let $\xi > 0$ and $\alpha > 0$ be fixed and define auxiliary function $\varphi : [0, R_0) \to \mathbb{R}$ by

$$\varphi(t) = \xi + (\alpha - 1)t + \alpha f(t). \tag{3.1}$$

We shall use the following hypotheses

(h_4) there exists $s^* \in (0, R)$ such that for each $t \in (0, s^*)$, $\varphi(t) > 0$ and $\varphi(s^*) = 0$;
(h_5) $\varphi(s^*) < 0$.

From now on we assume the hypotheses $(h_0) - (h_4)$ and $(h_0^0) - (h_2^0)$ which will be called the hypotheses (H).

Next, we present the main semi-local convergence result of the Gauss-Newton method generated by the Algorithm GNA for solving (2.1).

Theorem 3.4. *Suppose that the (H) conditions are satisfied. Then,*

(i) *sequence $\{s_k\}$ generated by the Gauss-Newton method for $s_0 = 0, s_{k+1} = s_k - \frac{\varphi(s_k)}{\varphi'(s_k)}$ for solving equation $\psi(t) = 0$ is well defined, strictly increasing, remains in $[0, s^*)$ and converges Q-linearly to s^*.*

Let $\eta \in [1, \infty]$, $\Delta \in (0, \infty]$ and $h : \mathbb{R}^m \to \mathbb{R}$ be real-valued convex with minimizer set C such that $C \neq \emptyset$.

(ii) *Suppose that* $x_0 \in \mathbb{R}^n$ *is a quasi-regular point of the inclusion*

$$F(x) \in C,$$

with the quasi-regular radius r_{x_0} *and the quasi-regular bound function* β_{x_0} *defined by* (2.14) *and* (2.15)*, respectively. If* $d(F(x_0), C) > 0$, $s^* \le r_{x_0}$, $\Delta \ge \xi \ge \eta_{\beta_{x_0}}(0) d(F(x_0), C)$,

$$\alpha \ge \sup \left\{ \frac{\eta_{\beta_{x_0}}(t)}{\eta_{\beta_{x_0}}(t)(1 + f'(t)) + 1} : \xi \le t < s^* \right\},$$

then, sequence $\{x_k\}$ *generated by GNA is well defined, remains in* $U(x_0, s^*)$ *for each* $k = 0, 1, 2, \cdots$*, such that*

$$F(x_k) + F'(x_k)(x_{k+1} - x_k) \in C \quad \text{for each} \quad k = 0, 1, 2 \cdots . \tag{3.2}$$

Moreover, the following estimates hold

$$\|x_{k+1} - x_k\| \le s_{k+1} - s_k, \tag{3.3}$$

$$\|x_{k+1} - x_k\| \le \frac{s_{k+1} - s_k}{(s_k - s_{k-1})^2} \|x_k - x_{k-1}\|^2, \tag{3.4}$$

for each $k = 0, 1, 2 \cdots$, *and* $k = 1, 2, \cdots$*, respectively and converges to a point* $x^* \in U(x_0, s^*)$ *satisfying* $F(x^*) \in C$ *and*

$$\|x^* - x_k\| \le t^* - s_k \quad \text{for each} \quad k = 0, 1, 2, \cdots . \tag{3.5}$$

The convergence is R-linear. If hypothesis (h_5) *holds, then the sequences* $\{s_k\}$ *and* $\{x_k\}$ *converge Q-quadratically and R-quadratically to* s^* *and* x^**, respectively. Furthermore, if*

$$\alpha > \overline{\alpha} := \sup \left\{ \frac{\eta_{\beta_{x_0}}(t)}{\eta_{\beta_{x_0}}(t)(1 + f'(t)) + 1} : \xi \le t < s^* \right\},$$

then, the sequence $\{x_k\}$ *converges R-quadratically to* x^*.

Proof. Simply replace function g in [23] (see also [15]) by function f in the proof, where g is a majorant function for F on $U(x_0, R)$. That is we have instead of (h_0):

$$(h_0') \qquad \|F'(y) - F'(x)\| \le g'(\|y - x\| + \|x - x_0\|) - g'(\|x - x_0\|) \tag{3.6}$$

for each $x, y \in U(x_0, R)$ with $\|x - x_0\| + \|y - x\| < R$. The iterates $\{x_n\}$ lie in $U(x_0, R_0)$ which is a more precise location than $U(x_0, R)$. □

Remark 3.5. (a) If $f(t) = g(t) = f_0(t)$ for each $t \in [0, R_0)$ and $R_0 = R$, then Theorem 2.1 reduces to the corresponding Theorem in [15]. Moreover, if $f_0(t) \leq f(t) = g(t)$ we obtain the results in [23]. Notice that, we have that

$$f_0'(t) \leq g'(t) \quad \text{for each} \quad t \in [0, R) \tag{3.7}$$

and

$$f'(t) \leq g'(t) \quad \text{for each} \quad t \in [0, R_0). \tag{3.8}$$

Therefore, if

$$f_0'(t) \leq f'(t) < g'(t) \quad \text{for each} \quad t \in [0, R_0), \tag{3.9}$$

then the following advantages denoted by $(\mathcal{A})$ are obtained: weaker sufficient convergence criteria, tighter error bounds on the distances $\|x_n - x^*\|, \|x_{n+1} - x_n\|$ and an at least as precise information on the location of the solution x^*. These advantages are obtained using less computational cost, since in practice the computation of function g requires the computation of functions f_0 and f as special cases. It is also worth noticing that under (h_0^0) function f_0' is defined and therefore R_0 which is at least as small as R. Therefore the majorant function to satisfy (*i.e.*, f') is at least as small as the majorant function satisfying (h_0') (*i.e.*, g') leading to the advantages of the new approach over the approach in [15] or [23]. Indeed, we have that if function ψ has a solution t^*, then, since $\varphi(t^*) \leq \psi(t^*) = 0$ and $\varphi(0) = \psi(0) = \xi > 0$, we get that function φ has a solution r^* such that

$$r^* \leq t^*, \tag{3.10}$$

but not necessarily vice versa. If also follows from (3.10) that the new information about the location of the solution x^* is at least as precise as the one given in [18].

(b) Let us specialize conditions (2.8)–(2.10) even further in the case when L_0, K and L are constant functions and $\alpha = 1$. Then, function corresponding to (3.1) reduce to

$$\psi(t) = \frac{L}{2}t^2 - t + \xi \tag{3.11}$$

[15, 23] and

$$\varphi(t) = \frac{K}{2}t^2 - t + \xi, \tag{3.12}$$

respectively. In this case the convergence criteria become, respectively

$$h = L\xi \leq \frac{1}{2} \tag{3.13}$$

and

$$h_1 = K\xi \leq \frac{1}{2}. \tag{3.14}$$

Notice that

$$h \leq \frac{1}{2} \quad \Longrightarrow \quad h_1 \leq \frac{1}{2}, \tag{3.15}$$

but not necessarily vice versa. Unless, if $K = L$. Criterion (3.13) is the famous for its simplicity and clarity Kantorovich hypotheses for the semilocal convergence of Newton's method to a solution x^* of equation $F(x) = 0$ [20]. In the case of Wang's condition [29] we have

$$\begin{aligned} \varphi(t) &= \frac{\gamma t^2}{1-\gamma t} - t + \xi, \\ \psi(t) &= \frac{\beta t^2}{1-\beta t} - t + \xi, \\ L(u) &= \frac{2\gamma}{(1-\gamma u)^3}, \quad \gamma > 0, \quad 0 \leq t \leq \frac{1}{\gamma} \end{aligned}$$

and

$$K(u) = \frac{2\beta}{(1-\beta u)^3}, \quad \beta > 0, \quad 0 \leq t \leq \frac{1}{\beta}$$

with convergence criteria, given respectively by

$$H = \gamma\xi \leq 3 - 2\sqrt{2}, \tag{3.16}$$

$$H_1 = \beta\xi \leq 3 - 2\sqrt{2}. \tag{3.17}$$

Then, again we have that

$$H \leq 3 - 2\sqrt{2} \quad \Longrightarrow \quad H_1 \leq 3 - 2\sqrt{2},$$

but not necessarily vice versa, unless, if $\beta = \gamma$.

(c) Concerning the error bounds and the limit of majorizing sequence, let us define majorizing sequence $\{r_{\alpha,k}\}$ by

$$r_{\alpha,0} = 0; r_{\alpha,k+1} = r_{\alpha,k} - \frac{\varphi(r_{\alpha,k})}{\varphi'_{\alpha,0}(r_{\alpha,k})}$$

for each $k = 0, 1, 2, \cdots$, where

$$\varphi_{\alpha,0}(t) = \xi - t + \alpha \int_0^t L_0(t)(t-u)du.$$

Suppose that

$$-\frac{\varphi(r)}{\varphi'_{\alpha,0}(r)} \leq -\frac{\varphi(s)}{\varphi'(s)}$$

for each $r, s \in [0, R_0]$ with $r \leq s$. According to the proof of Theorem 3.1 sequence $\{r_{\alpha,k}\}$ is also a majorizing sequence for GNA.

Moreover, a simple inductive argument shows that

$$r_k \leq s_k,$$

$$r_{k+1} - r_k \leq s_{k+1} - s_k$$

and

$$r^* = \lim_{k \longrightarrow \infty} r_k \leq s^*.$$

Furthermore, the first two preceding inequality are strict for $n \geq 2$, if

$$L_0(u) < K(u) \quad \text{for each} \quad u \in [0, R_0].$$

Similarly, suppose that

$$-\frac{\varphi(s)}{\varphi'(s)} \leq -\frac{\psi(t)}{\psi'(t)} \tag{3.18}$$

for each $s, t \in [0, R_0]$ with $s \leq t$. Then, we have that

$$s_{\alpha,k} \leq t_{\alpha,k},$$

$$s_{\alpha,k+1} - s_{\alpha,k} \leq t_{\alpha,k+1} - t_{\alpha,k}$$

and

$$s^* \leq t^*.$$

The first two preceding inequalities are also strict for $k \geq 2$, if strict inequality holds in (3.18).

4. Numerical examples

Specializations of Theorem 3.3 to some interesting cases such as Smale's $\alpha-$theory (see also Wang's $\alpha-$theory) and Kantorovich theory have been reported in [15, 23, 25], if $f_0'(t) = f'(t) = g'(t)$ for each $t \in [0, R)$ with $R_0 = R$ and in [23], if $f_0'(t) < f'(t) = g'(t)$ for each $t \in [0, R_0)$. Next, we present examples where $f_0'(t) < f'(t) < g'(t)$ for each $t \in [0, R_0)$ to show the advantages of the new approach over the ones in [15, 24, 26]. We choose for simplicity $m = n = \alpha = 1$.

Example 4.1. Let $x_0 = 1, D = U(1, 1-q)$, $q \in [0, \frac{1}{2})$ and define function F on D by

$$F(x) = x^3 - q. \tag{4.1}$$

Then, we have that $\xi = \frac{1}{3}(1-q), L_0 = 3 - q, L = 2(2-q)$ and $K = 2(1 + \frac{1}{L_0})$. The Newton-Kantorovich condition (3.13) is not satisfied, since

$$h > \frac{1}{2} \quad \text{for each} \quad q \in [0, \frac{1}{2}). \tag{4.2}$$

Hence, there is no guarantee by the Newton-Kantorovich Theorem [15] that Newton's method (2.1) converges to a zero of operator F. Let us see what gives: We have by (3.14) that

$$h_1 \leq \frac{1}{2}, \quad \text{if} \quad 0.461983163 < q < \frac{1}{2}. \tag{4.3}$$

Hence, we have demonstrated the improvements using this example.

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, 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, *A semilocal convergence analysis for directional Newton methods*, Mathematics of Computation, AMS, **80** (2011), 327–343.

[5] 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.

[6] I.K. Argyros and S. Hilout, *On the Gauss–Newton method*, J. Appl. Math. Comput., **35**(1) (2011), 537–550.

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

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

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

[10] I.K. Argyros and S. Hilout, *Computational Methods in Nonlinear Analysis*, World Scientific Publ. Comp., New Jersey, 2013.

[11] 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.

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

[13] J.E. Jr. Dennis and R.B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Classics in Appl. Math., **16**, SIAM, Philadelphia, PA, 1996.

[14] 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**(1) (2011), 111–125.

[15] O.P. Ferreira, M.L.N. Gonçalves and P.R. Oliveira, *Convergence of the Gauss-Newton method for convex composite optimization under a majorant condition*, SIAM J. Optim., **23**(3) (2013), 1757-1783.

[16] O.P. Ferreira and B.F. Svaiter, *Kantorovich's theorem on Newton's method in Riemmanian manifolds*, J. Complexity, **18**(1) (2002), 304–329.

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

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

[19] 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.

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

[21] C. Li and K.F. Ng, *Majorizing functions and convergence of the Gauss-Newton method for convex composite optimization*, SIAM J. Optim., **18**(2) (2007), 613–692.

[22] C. Li and X.H. Wang, *On convergence of the Gauss-Newton method for convex composite optimization*, Math. Program. Ser A. **91** (2002), 349–356.

[23] A. Magrenán and I.K. Argyros, *Expanding the applicability of the Gauss-Newton method for convex optimization under a majorant condition*, SeMA, **65** (2014), 37–56.

[24] F.A. Potra, *Sharp error bounds for a class of Newton-like methods*, Libertas Mathematica, **5** (1985), 71–84.

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

[26] R.T. Rockafellar, *Monotone proccesses of convex and concave type*, Memoirs of the American Mathematical Society, **77**. American Mathematical Society, Providence, R.I., 1967.

[27] R.T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, 28, Princeton University Press, Princeton, N.J. 1970.

[28] 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.

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