Journal of Mathematical Sciences and Modelling, 1 (2) (2018) 80-85

Journal of Mathematical Sciences and Modelling

Journal Homepage: www.dergipark.gov.tr/jmsm
ISSN 2636-8692

Improved semi-local convergence of the Gauss-Newton method for systems of equations

Ioannis K. Argyros[a] and Santhosh George[b*]

[a]*Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA*
[b]*Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, India-575 025*
[*]*Corresponding author E-mail: sgeorge@nitk.ac.in*

Article Info

Keywords: *Gauss- Newton method, Newton's method, Semi-local convergence, Least squares problem*
2010 AMS: *65H10, 65G99, 65K10, 47H17, 49M15*
Received: *8 June 2018*
Accepted: *13 September 2018*
Available online: *30 September 2018*

Abstract

Our new technique of restricted convergence domains is employed to provide a finer convergence analysis of the Gauss-Newton method in order to solve a certain class of systems of equations under a majorant condition. The advantages are obtained under the same computational cost as in earlier studies such as [5, 14]. Special cases and a numerical example are also given in this study.

1. Introduction

Let $\Omega \subseteq \mathbb{R}^n$ be open. Let $F : \Omega \to \mathbb{R}^m$ be continuously Fréchet- differentiable. The problem of approximating least squares solutions x^* of the nonlinear problem

$$\min_{x\in\Omega} \|F(x)\|^2, \tag{1.1}$$

is very important in computational mathematics. The least squares solutions of (1.1) are stationary points of $Q(x) = \|F(x)\|^2$. A lot of 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 [6, 16, 19] and the references therein. The famous Gauss-Newton method defined by

$$x_{k+1} = x_k - F'(x_k)^{\dagger} F(x_k), \text{ for each } k = 0, 1, \cdots, \tag{1.2}$$

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]–[6], [8, 10, 20, 14, 15, 17].
In the present paper, we are motivated by the work of Goncalves and Oliveira in [14] (see also [12], [13]) and our works in [1, 2, 3, 4, 6, 7, 8]. These authors presented a semi-local convergence analysis for the Gauss-Newton method (1.2) for systems of nonlinear equations where the function F satisfies

$$\|F'(y)^{\dagger}(I_{\mathbb{R}^m} - F'(x)F'(x)^{\dagger})F(x)\| \le k\|x - y\| \text{ for each } x \text{ and } y \in \Omega,$$

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 condition (h_3)). Their results unify the classical results for functions involving Lipschitz derivative [6, 7, 16, 18] with results for analytical functions ($\alpha-$theory or $\gamma-$theory) [9, 11, 15, 17, 19, 20].
We introduce a center majorant function (see (c_3)) which is a special case of the majorant function that can provide more precise estimates on the distances $\|F'(x)^{\dagger}\|$. Then, we find a domain where the iterates lie which is more precise than in the aforementioned studies. This leads to "smaller" majorant functions yielding 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 rest of the paper is organized as follows: The semi-local convergence analysis of the Gauss-Newton method is presented in Section 2. Special cases and numerical examples are given in the concluding Section 3.

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

2. Semi-local convergence analysis

In this section we present the semi-local convergence analysis of the Gauss-Newton method. Let $R > 0$. Denote by $B(x_0,R), \bar{B}(x_0,R)$ the open and closed balls in $\mathbb{R}^n$, respectively with center $x_0 \in \mathbb{R}^n$ and radius R. We shall use the hypotheses denoted by $(\mathscr{C})$.
(c_0) Let $B(x_0,R) \subseteq \mathbb{R}^n$ and $F : B(x_0,R) \to \mathbb{R}^m$ be continuously Fréchet- differentiable.
(c_1) continuously differentiable functions $f_0 : [0,R) \longrightarrow \mathbb{R}$, $f : [0,R^*) \to \mathbb{R}$

$$\|F'(x_0)^\dagger\|\|F'(x) - F'(x_0)\| \leq f_0'(\|x - x_0\|) - f_0'(0) \text{ for each } x \in B(x_0,R)$$

and

$$\|F'(x_0)^\dagger\|\|F'(y) - F'(x)\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|) \text{ for each } x,y \in B(x_0,R^*)$$

with $\|y - x\| + \|x - x_0\| < R^*$ where $R_0 := \sup\{t \in [0,R] : f_0'(t) < 0\}$. Set

$$R^* := \min\{R_0, R\}.$$

(c_2)

$$\|F'(y)^\dagger(I_{\mathbb{R}^m} - F'(x)F'(x)^\dagger)F(x)\| \leq \kappa\|x - y\| \text{ for each } x \text{ and } y \in B(x_0,R^*),$$

where $\kappa \in [0,1)$.
(c_3) Set $\eta = \|F'(x_0)^\dagger F(x_0)\| > 0,\ F'(x_0) \neq 0.$

$$rank(F'(x)) \leq rank(F'(x_0)) \neq 0 \text{ for each } x \in B(x_0,R^*).$$

(c_4)

$$f_0(0) = f(0) = 0,\ f'(0) = f_0'(0) = -1$$
$$f_0(t) \leq f(t) \text{ and } f_0'(t) \leq f'(t) \text{ for each } t \in [0,R^*).$$

(c_5) f_0', f' are convex and strictly increasing.
Let $\mu \geq 0$ be such that $\mu \geq -\kappa f'(\eta)$ and define $\varphi_{\eta,\mu} : [0,R^*) \to \mathbb{R}$ by

$$\varphi_{\eta,\mu}(t) = \eta + \mu t + f(t).$$

(c_6) $\varphi_{\eta,\mu}(t) = 0$ for some $t \in [0,R^*)$.
(c_7) For each $s,t,u \in [0,R^*)$ with $s \leq t \leq u$

$$t + \frac{\varphi_{\eta,\mu}(u)}{f_0'(u)} \leq u + \frac{\varphi_{\eta,\mu}(t) - \varphi_{\eta,\mu}(s) - \varphi_{\eta,\mu}'(s)(t-s)}{f_0'(t)}$$

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

$$r_0 = 0,\ r_{k+1} = r_k - \frac{\varphi_{\eta,\mu}(r_k)}{f_0'(r_k)}. \tag{2.1}$$

The corresponding iteration $\{t_n\}$ used in [14] is given by

$$t_0 = 0,\ t_{k+1} = t_k - \frac{\bar{\varphi}_{\eta,\mu}(t_k)}{g'(t_k)}, \tag{2.2}$$

where $\bar{\varphi}_{\eta,\mu}(t) = \eta + \mu t + g(t)$, continuously differentiable function $g : [0,R) \longrightarrow \mathbb{R}$ is such that

$$\|F'(x_0)^+\|\|F'(x) - F'(y)\| \leq g'(\|y - x\| + \|x - x_0\|) - g'(\|x - x_0\|)$$

for each $x,y \in B(x_0,R)$. Moreover, define iterations $\{s_k\}$ by

$$s_0 = 0, s_{k+1} \quad = \quad s_k - \frac{\varphi_{\eta,\mu}(s_k)}{f_0'(s_k)}.$$

This iteration was used by us in [5]. In view of these conditions, we have

$$f_0'(t) \leq g'(t) \tag{2.3}$$

and

$$f'(t) \leq g'(t) \tag{2.4}$$

for each $t \in [0,R^*)$. Next, the main semi-local convergence result for the Gauss-Newton method is presented.

Theorem 2.1. *Suppose that the $(\mathscr{C})$ conditions hold and $f_0'(t) \leq f'(t)$ for each $t \in [0,R^*]$. Then, the following hold: $\varphi_{\eta,\mu}(t)$ has a smallest zero $r^* \in (0,R^*)$, the sequences $\{r_k\}$ and $\{x_k\}$ for solving $\varphi_{\eta,\mu}(t)=0$ and $F(x)=0$, with starting point $t_0=0$ and x_0, respectively given by (1.2) and (2.3) are well defined, $\{r_k\}$ is strictly increasing, remains in $[0,r^*)$, and converges to r^*, $\{x_k\}$ remains in $B(x_0,r^*)$, converges to a point $x^* \in B(x_0,r^*)$ such that $F'(x^*)^{\dagger}F(x^*)=0$. Moreover, the following estimates hold:*

$$\|x_{k+1}-x_k\| \leq r_{k+1}-r_k \text{ for each } k=0,1,2,\cdots,$$

$$\|x^*-x_k\| \leq r^*-r_k \text{ for each } k=0,1,2,\cdots,$$

and

$$\|x_{k+1}-x_k\| \leq \frac{r_{k+1}-r_k}{(r_k-r_{k-1})^2}\|x_k-x_{k-1}\|^2 \text{ for each } k=0,1,2,\cdots.$$

Furthermore, if $\mu=0$($\mu=0$ and $f_0'(r^*)<0$), the sequence $\{r_k\}$, $\{x_k\}$ converge $Q-$linearly and $R-$linearly ($Q-$ quadratically and $R-$ quadratically) to r^* and x^*, respectively.

Proof. Simply repeat the proof of Theorem 3.9 in [5] (or the proof in [14]) with f replacing g. Notice also that the iterates x_n remain in $B(x_0,R_0)$ which is a more precise location than $B(x_0,R^*)$ used in [5, 14].

Remark 2.2. *(i) As noted in [14] the best choice for μ is given by $\mu=-\kappa f'(\kappa)$.*

(ii) 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 [8]. Moreover, if $f_0'(t) \leq f'(t) = g'(t)$ we obtain the results in [5]. If

$$f_0'(t) \leq f'(t) \leq g'(t) \text{ for each } t \in [0,R^*) \tag{2.5}$$

then the following advantages denoted by $(\mathscr{A})$ 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 (c_1) function f_0' is defined and therefore R^* which is at least as small as R.

We have that, if function $\bar{\varphi}_{\eta,\mu}$ has a solution t^*, then, since $\varphi_{\eta,\mu}(t^*) \leq \bar{\varphi}_{\eta,\mu}(t^*)=0$ and $\varphi_{\eta,\mu}(0)=\bar{\varphi}_{\eta,\mu}(0)=\eta>0$, we get that function $\varphi_{\eta,\mu}$ has a solution r^* such that

$$r^* \leq t^* \tag{2.6}$$

but not necessarily vice versa. It also follows from (2.6) that the new information about the location of the solution x^* is at least as precise as the one given in [14, 5].

Let us specialize conditions $(\mathscr{C})$ even further in the case when f_0, f and g are constant functions L_0, K, L, respectively. Then, (for $\mu=0$) we have that:

$$\bar{\varphi}_{\eta,\mu}(t)=\frac{L}{2}t^2-t+\eta \tag{2.7}$$

and

$$\varphi_{\eta,\mu}(t)=\frac{K}{2}t^2-t+\eta, \tag{2.8}$$

respectively. In this case the convergence criteria become, respectively

$$h=L\eta \leq \frac{1}{2}$$

and

$$h_1=K\eta \leq \frac{1}{2}.$$

Notice that

$$h \leq \frac{1}{2} \Longrightarrow h_1 \leq \frac{1}{2}$$

but not vice versa unless, $K=L$. Criterion (2.8) is famous for its simplicity and clarity Kantorovich hypothesis for the semilocal convergence of Newton's method to a solution x^* of nonlinear equation $F(x)=0$ [7, 16]. In the case of Wang's conditions [20] we have for $\mu=0$:

$$g(t)=\frac{\gamma t^2}{1-\gamma t}-t,\ f(t)=\frac{\beta t^2}{1-\beta t}-t,\ f_0(t)=\frac{\gamma_0 t^2}{1-\gamma_0 t}-t,$$

$$\bar{\varphi}_{\eta,\mu}(t)=\frac{\gamma t^2}{1-\gamma t}-t+\eta, \tag{2.9}$$

$$\varphi_{\eta,\mu}(t)=\frac{\beta t^2}{1-\beta t}-t+\eta \tag{2.10}$$

with convergence criteria, given respectively by

$$H=\gamma\eta \leq 3-2\sqrt{2} \tag{2.11}$$

$$H_1=\beta\eta \leq 3-2\sqrt{2}. \tag{2.12}$$

Then, again we have that

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

but not necessarily vice versa, unless if $\beta = \gamma$.
Concerning the error bounds and the limit of majorizing sequence, suppose that

$$-\frac{\varphi_{\eta,\mu}(r)}{f_0'(r)} \leq -\frac{\varphi_{\eta,\mu}(s)}{f_0'(s)}$$

for each $r,s \in [0,R^*]$ with $r \leq s$. According to the proof of Theorem 2.1, sequence $\{r_n\}$ is also a majorizing sequence for (1.2). Moreover, a simple induction argument shows that

$$r_n \leq s_n, r_{n+1}-r_n \leq s_{n+1}-s_n$$

and

$$r^* = \lim_{n\longrightarrow\infty} r_n \leq s^*.$$

Furthermore, the first two preceding inequalities are strict, for $n \geq 2$ if $f_0'(t) < f'(t)$ for each $t \in [0,R^*]$. Similarly, suppose that

$$-\frac{\varphi_{\eta,\mu}(s)}{f_0'(s)} \leq -\frac{\varphi_{\eta,\mu}(t)}{f_0'(t)}$$

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

$$s_n \leq t_n, s_{n+1}-s_n \leq t_{n+1}-t_n.$$

The first two preceding inequalities are also strict for $n \geq 2$, if strict inequality holds in (2.12).

Finally, the rest of the results in [5, 14] can be improved along the same lines by also using K instead of L. We leave the details to the motivated reader.

3. Numerical examples

We present a simple example where we show that Wang's condition (2.11) [20] is violated but our condition (2.12) is satisfied. More examples can be found in [7] where $L_0 \leq K \leq L$ are satisfied as strict inequalities (therefore the new advantages apply) (or see also [19]).

Example 3.1. *Let $\mu = 0$, $p \in (0,1), x_0 = 1$, $\Omega = B(x_0, \frac{1}{2-p})$ and define functions on Ω by*

$$f(x) = \frac{x^4}{4} - px, \ F(x) = x^3 - p. \tag{3.1}$$

Define $\Omega^* = B(x_0, 1-p)$. Then, we have

$$\Omega^* \subseteq \Omega, \text{ if } p \in [0.381966, 1). \tag{3.2}$$

Let $L_0 = 3-p$ and $L = 2(2-p)$. Then, Argyros showed in [8] that for each $x,y \in \Omega$

$$|F'(x_0)^{-1}(F'(x)-F'(x_0))| \leq L_0|x-x_0| \tag{3.3}$$

and

$$|F'(x_0)^{-1}(F'(x)-F'(y))| \leq L|x-y|. \tag{3.4}$$

Consider the conditions

$$\|F'(x_0)^{-1}F''(x)\| \leq \frac{2\gamma}{(1-\gamma\|x-x_0\|)^3} \tag{3.5}$$

for each $x \in \Omega$,

$$\|F'(x_0)^{-1}(F'(x)-F'(x_0))\| \leq \frac{1}{(1-\gamma_0\|x-x_0\|)^2} - 1 \tag{3.6}$$

for each $x \in \Omega$ and

$$\|F'(x_0)^{-1}F''(x)\| \leq \frac{2\beta}{(1-\beta\|x-x_0\|)^3} \tag{3.7}$$

for each $x \in \Omega^*$. Notice that functions $\bar{\varphi}_{\eta,0}$, $\varphi_{\eta,0}$ satisfy these conditions, respectively. In view of (3.4) and (3.5), we have $L \leq 2\gamma$, so we choose $\gamma = 2-p$. Then, since $\eta = \frac{1}{3}(1-p)$, condition (2.11) is satisfied, if

$$0.6255179 \leq p < 1. \tag{3.8}$$

We must have

$$B(x_0, (1-\frac{1}{\sqrt{2}})\frac{1}{\gamma}) \subseteq B(x_0, 1-p),$$

which is true for

$$0 < p \leq 0.7631871. \tag{3.9}$$

It follows from (3.8) and (3.9) that

$$0.6255179 < p \leq 0.7631871. \tag{3.10}$$

Set $y = \gamma_0|x-x_0|$ *and* $L_0 = d\gamma_0, d > 0, \gamma_0 > 0$. *Using (3.6) and (3.3), we must have*

$$L_0|x-x_0| \leq \frac{1}{(1-\gamma_0|x-x_0|)^2} - 1$$

or

$$d(1-y)^2 \leq 2-y$$

or

$$dy^2 + (1-2d)y + d - 2 \leq 0. \tag{3.11}$$

Let e.g. $d = 2$, *then* $\gamma_0 = \frac{L_0}{2} = \frac{3-p}{2}$ *and (3.11) becomes* $(p-3)(p-1) \leq 3$ *or* $p(p-4) \leq 0$, *which is true. We must show* $(1-\frac{1}{\sqrt{2}})\frac{1}{\gamma_0} \leq 1-p$ *or* $p^2 - 4p + 1 + \sqrt{2} \geq 0$, *which is true for*

$$0 < p \leq 0.7407199. \tag{3.12}$$

Notice that $\Omega_0 \subset \Omega$, *since* $(1-\frac{1}{\sqrt{2}})\frac{1}{\gamma_0} < \frac{1}{\gamma}$ *or* $p \leq 3+\sqrt{2}$, *which is true, so*

$$\Omega \cap \Omega_0 = \Omega_0. \tag{3.13}$$

Then, for $x \in \Omega_0$

$$\begin{aligned} |F'(x_0)^{-1}F''(x)| &= 2|x| \leq 2(|x-x_0| + |x_0|) \\ &\leq 2((1-\frac{1}{\sqrt{2}})\frac{2}{3-p} + 1) \end{aligned}$$

must be smaller than 2β, *so we can choose*

$$\beta = 1 + (1-\frac{1}{\sqrt{2}})\frac{2}{3-p} = 1 + \frac{2-\sqrt{2}}{3-p}.$$

Notice that $\beta < \gamma$, *if (3.12) holds. We also have that* $\gamma_0 < \beta$, *if*

$$\frac{3-p}{2} < 1 + \frac{2-\sqrt{2}}{3-p}$$

or if

$$p^2 - 4p - 1 + 2\sqrt{2} < 0$$

or, if

$$0.5263741 < p < 1. \tag{3.14}$$

We also must have

$$(1-\frac{1}{\sqrt{2}})\frac{1}{\beta} \leq 1-p$$

or

$$2p^2 + (\sqrt{2}-10)p + 4 + \sqrt{2} \leq 0,$$

which is true for

$$p \leq 0.767996. \tag{3.15}$$

Then, notice that

$$1-p \leq \frac{1}{\gamma},$$

if $p^2 - 3p + 1 \leq 0$, *which is true for*

$$0.381966 \leq p < 1. \tag{3.16}$$

Then, we have that $\alpha_0 \leq 3 - 2\sqrt{2} = q$, *if* $(1+\frac{2-\sqrt{2}}{3-p})\frac{1}{3}(1-p) \leq q$ *or if*

$$p^2 + (\sqrt{2}-6+3q)p + 5 - \sqrt{2} - 9q \leq 0,$$

which is true for

$$0.5857931 \leq p < 1. \tag{3.17}$$

In view of (3.12), (3.14), (3.15) and (3.17) we must have

$$0.5857931 \leq p \leq 0.7407199. \tag{3.18}$$

Define intervals I and I_1 by

$$I = [0.5857931, 0.6255179) \tag{3.19}$$

and

$$I_1 = (0.7407199, 0.7631871]. \tag{3.20}$$

In view of (3.10), (3.19) and (3.20), we see that for $p \in I$ [20] cannot guarantee the convergence of x_n to $x^* = \sqrt[3]{p}$. However, our Theorem 2.1 guarantees the convergence of x_n to x^*. Notice that, if $p \in I_1$, then we can set $\beta = \gamma = \gamma_0$.
Next, we compare the error bounds. Choose $p = 0.623$. Then, we have the following comparison table, which shows that the new error bounds are more precise than the ones in [20].

n	$r_{n+1} - r_n$	$t_{n+1} - t_n$
1	0.1257	0.1257
2	0.0268	0.0333
3	0.0013	0.0027
4	3.3384e-06	1.8199e-05
5	2.0876e-11	8.2197e-10

Table 1: Comparison table.

References

[1] Argyros, I.: On the semilocal convergence of the Gauss-Newton method. Adv. Nonlinear Var.Inequal. 8(2), 93–99, 2005.
[2] Argyros, I., Hilout, S.: On the local convergence of the Gauss-Newton method. Punjab Univ. J.Math. 41, 23–33, 2009.
[3] Argyros, I., Hilout, S.: On the Gauss-Newton method. J. Appl. Math. Comput. 1–14, 2010.
[4] Argyros, I. K, Hilout, S.: Extending the applicability of the Gauss-Newton method under average Lipschitz-type conditions. Numer. Algorithms 58(1), 23–52, 2011.
[5] Argyros, I. K, S. George, Expanding the applicability of the Gauss-Newton method for a certain class of systems of equations, J. Numer.Anal. Approx. Theory, vol., 45(1), 3–13, (2016).
[6] Argyros, I. K., Hilout, S.: Improved local convergence of Newton's method under weak majorant condition, Journal of Computational and Applied Mathematics, 236(7), 1892–1902, 2012.
[7] Argyros, I., Hilout, S.: Numerical methods in nonlinear analysis, World Scientific Publ. Comp. New Jersey, USA, 2013.
[8] Ben-Israel, A., Greville, T.N.E.: Generalized inverses. CMS Books in Mathematics/Ouvrages de Mathematiques de la SMC, 15. Springer-Verlag, New York, second edition, Theory and Applications, 2003.
[9] Catinas, E.: The inexact, inexact perturbed, and quasi-Newton methods are equivalent models, Math. Comput. 74, 249, (2005), 291-301.
[10] Dedieu, J.P., Kim, M.H.: Newton's method for analytic systems of equations with constant rank derivatives. J. Complexity, 18(1): 187-209, 2002.
[11] Ferreira, O.P., Gonçalves, M.L.N, Oliveira, P.R.:, Local convergence analysis of inexact Gauss–Newton like methods under majorant condition, J. Complexity, 27(1), 111-125, 2011.
[12] Ferreira, O.P., Svaiter, B.F.: Kantorovich's majorants principle for Newton's method. Comput.Optim. Appl. 42(2), 213–229, 2009.
[13] Häussler, W.M.: A Kantorovich-type convergence analysis for the Gauss-Newton-method. Numer. Math. 48(1), 119–125, 1986.
[14] Gonçalves, M.L.N, Oliveira, P.R.:, Convergence of the Gauss- Newton method for a special class of systems of equations under a majorant condition, Optimization, 64, 3(2015), 577–594.
[15] Hu, N., Shen, W. and Li, C.: Kantorovich's type theorems for systems of equations with constant rank derivatives, J. Comput. Appl.Math., 219(1): 110-122, 2008.
[16] Kantorovich, L.V., Akilov, G.P., Functional Analysis, Pergamon Press, Oxford, 1982.
[17] Li, C., Hu, N., Wang, J.: Convergence behavior of Gauss-Newton's method and extensions of the Smale point estimate theory. J. Complex. 26(3), 268–295, 2010.
[18] Potra, F.A., Ptak, V.: Nondiscrete induction and iterative processes. Research notes in Mathematics, 103, Pitman(Aóvanced Publishing Program), Boston, MA, 1984.
[19] Smale, S., 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.
[20] Wang, X.H., Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces, IMA J. Numer. Anal., 20 , 123–134, 2000.