MATHEMATICA, Tome 51 (74), N° 2, 2009, pp. 99–110

AN IMPROVED CONVERGENCE ANALYSIS OF NEWTON'S METHOD FOR SYSTEMS OF EQUATIONS WITH CONSTANT RANK DERIVATIVES

IOANNIS K. ARGYROS and SAÏD HILOUT

Abstract. We use Newton's method to solve systems of equations with constant rank derivatives. Motivated by optimization considerations, and using more precise estimates, we provide a convergence analysis for Newton's method with the following advantages over the work in [11]: larger convergence domain; finer error estimates on the distances involved, and an at least as precise information on the location of the solution. These improvements are obtained under the same hypotheses and computational cost as in [11]. Kantorovich-type as well as Smale-type point estimate applications are also provided.

MSC 2000. 65F20, 65H10, 49M15.

Key words. Newton's method, constant rank derivatives, semilocal convergence, Lipschitz condition with L average.

1. INTRODUCTION

In this study we are concerned with the problem of approximating a solution $x^\star$ of equation

$$F'(x)^+F(x) = 0, \tag{1}$$

where F is a Fréchet-differentiable operator defined on a convex subset $\mathcal{D}$ of an Euclidean space $\mathcal{X}$ into an Euclidean space $\mathcal{Y}$. We assume that $m = \dim\,\mathcal{X} \leq n = \dim\,\mathcal{Y}$.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time–invariant system is driven by the equation $\dot{x} = T(x)$, for some suitable operator T, where x is the state. Then the equilibrium states are determined by solving equation (1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative–when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

Let $\mathcal{M} : \mathcal{X} \longrightarrow \mathcal{Y}$ be an $m \times n$ matrix. Then the $m \times n$ matrix $\mathcal{M}^+ : \mathcal{Y} \longrightarrow \mathcal{X}$ is the Moore–Penrose inverse of $\mathcal{M}$ if it satisfies

$$\mathcal{M}\mathcal{M}^+\mathcal{M} = \mathcal{M}, \quad \mathcal{M}^+\mathcal{M}\mathcal{M}^+ = \mathcal{M}^+,$$
$$(\mathcal{M}\mathcal{M}^+)^\star = \mathcal{M}\mathcal{M}^+, \quad (\mathcal{M}^+\mathcal{M})^\star = \mathcal{M}^+\mathcal{M},$$

where $\mathcal{M}^\star$ is the adjoint of $\mathcal{M}$. Let Ker $\mathcal{M}$ and Im $\mathcal{M}$ denote the kernel and image of $\mathcal{M}$, respectively. For a subspace S of $\mathcal{X}$, we denote by $\mathcal{P}_S$ the projection onto S. We then have:

$$\mathcal{M}^+\mathcal{M} = \mathcal{P}_{\mathrm{Ker}\mathcal{M}^\perp} \quad \text{and} \quad \mathcal{M}\mathcal{M}^+ = \mathcal{P}_{\mathrm{Im}\mathcal{M}}.$$

The most popular method for generating a sequence $\{x_k\}$ $(k \geq 0)$ approximating $x^\star$ is undoubtedly Newton's method

$$x_{k+1} = x_k - F'(x_k)^+ F(x_k) \quad (x_0 \in \mathcal{D}), \quad (k \geq 0). \tag{2}$$

Here $F'(x)$ denote the Fréchet–derivative of operator F [3], [4].
In case $F'(x_k)$ is an isomorphism $(k \geq 0)$, then (2) reduces to the classical Newton's method:

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k) \quad (x_0 \in \mathcal{D}), \quad (k \geq 0). \tag{3}$$

There is an extensive literature on the local as well as the semilocal convergence analysis of methods (2) and (3) [1]–[13].

In particular, there are two ways of studying the convergence of methods (2) and (3). The first way one is the Kantorovich approach using information from the bound of F'' in a neighborhood of the initial guess x_0 [1]–[4], [7], [8]. Refinements of this theory were provided by Argyros in [1]–[5], where weaker sufficient conditions have extended the applicability of these methods. The new idea is to use a combination of Lipschitz and center-Lipschitz condition instead of only the Lipschitz condition in the error analysis of this methods. The second way is the Smale's point estimate theory based on the invariant (see [12])

$$\gamma(F, x_0) = \sup_{i \geq 2} \left\| \frac{F'(x_0)^{-1} F^{(i)}(x_0)}{i\,!} \right\|^{\frac{1}{i-1}}. \tag{4}$$

In the case when $F'(x)$ is not an isomorphism, Dedieu and Kim [6], Li et al. [10], [11], Argyros [3], [4], have provided sufficient convergence conditions for Newton's method (2) under various conditions.

In this paper, we are motivated by optimization considerations and the elegant paper by Li and Xu [11], where Wang's theory [13] was extended using Lipschitz conditions with L average to study Newton's method (2). Here we also use some ideas introduced by us in [5] to enlarge the radius of convergence of Newton's method provided by Huang in [9]. In particular, using more precise estimates than in [11], and under the same computational cost, we provide a semilocal convergence analysis for method (2) with the following advantages: weaker sufficient convergence conditions (i.e., the convergence domain is extended); finer estimates on the distances $\| x_k - x^\star \|$, and an at least as precise information on the location of solution $x^\star$.

The paper ends with some applications of Kantorovich-type and Smale-type results.

2. SEMILOCAL CONVERGENCE ANALYSIS OF NEWTON'S METHOD (2)

To make the paper as self–contained as possible, we provide two lemmas concerning the perturbation bounds for the Moore–Penrose inverse:

LEMMA 1. [11] *Let A and B be $m \times n$ matrices and let $r \leq \min\{m, n\}$. Suppose that* $\operatorname{rank}(A) = r$, $\operatorname{rank}(A + B) \leq r$, *and* $\| A^+ \| \| B \| < 1$. *Then, the following hold:*

$$\operatorname{rank}(A + B) = r \quad \textit{and} \quad \| (A + B)^+ \| \leq \frac{\| A^+ \|}{1 - \| A^+ \| \| B \|}.$$

LEMMA 2. [11] *Let A and B be $m \times n$ matrices and let $r \leq \min\{m, n\}$. Suppose that* $\operatorname{rank}(A) = r$, $\operatorname{rank}(B) \leq r$, *and* $\| A^+ \| \| B - A \| < 1$. *Then, the following hold:*

$$\| B^+ - A^+ \| \leq c_r \frac{\| A^+ \|^2 \| B - A \|}{1 - \| A^+ \| \| B - A \|},$$

where,

$$c_r = \begin{cases} \dfrac{1+\sqrt{5}}{2} & \textit{if} \quad r < \min\{m, n\} \\ \sqrt{2} & \textit{if} \quad r = \min\{m, n\}, m \neq n \\ 1 & \textit{if} \quad r = m = n. \end{cases}$$

We need the following terminology.

Let $\mathcal{I}_{\mathcal{X}}$ denote the identity on $\mathcal{X}$ and $K(A) = \| A^+ \| \| A \|$ denote the condition number of a linear operator $A : \mathcal{X} \longrightarrow \mathcal{Y}$. Denote by $U(x_0, R)$, ($x_0 \in \mathcal{D}$, $R > 0$) the open ball in $\mathcal{X}$ with center x_0 and radius R.

Let $L_0(t)$ and $L(t)$ be positive non–decreasing functions defined on $[0, \infty)$. We need the following notions of Lipschitz conditions with L average.

DEFINITION 1. Let $\mathcal{D} \subseteq \mathcal{X}$ and $y \in \mathcal{D}$. Then $F' : \mathcal{D} \longrightarrow \mathcal{Y}$ is said to satisfy:

(i) The center-Lipschitz condition with L_0 average at y on $\mathcal{D}$ if

$$\| F'(x) - F'(y) \| \leq \int_0^{\|x-y\|} L_0(t) \mathrm{d}t, \quad x \in \mathcal{D}; \tag{5}$$

(ii) The radius Lipschitz condition with L average at y on $\mathcal{D}$ if

(6)
$$\| F'(x) - F'(x_\tau) \| \leq \int_{\tau\|x-y\|}^{\|x-y\|} L(t)\mathrm{d}t,\ x \in \mathcal{D},\ 0 \leq \tau \leq 1,\ x_\tau = y + \tau(x-y).$$

Remark 1. If $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ satisfies the radius Lipschitz condition L with average at y on $\mathcal{D}$, then it also satisfies the center-Lipschitz condition with L average at y on $\mathcal{D}$. Note also that in this case

$$L_0(t) \leq L(t) \quad t \in [0, \infty) \tag{7}$$

holds in general and $\dfrac{L(t)}{L_0(t)}$ can be arbitrarily large [1], [3], [4].

Let x, $y \in \mathcal{D}$, we set

$$\theta_0 = \theta_0(L_0, y, x) = \| F'(y)^+ \| \int_0^{\|x-y\|} L_0(t)\mathrm{d}t. \tag{8}$$

If $L_0(t) = L(t)$ on $[0, \infty)$, then our Definition 1 reduces to the one in [11] (see [4], [9], [13]). If $L_0(t) < L(t)$ on $[0, \infty)$, then we have:

$$\theta_0 = \theta_0(L_0, y, x) < \theta = \theta(L, y, x) = \| F'(y)^+ \| \int_0^{\|x-y\|} L(t)\mathrm{d}t. \tag{9}$$

It turns out more precise θ_0 can replace θ in all the subsequent results in [11, starting on page 707]. Note that the computation of θ_0 is also cheaper that of parameter θ. This observation leads to the advantages of our approach over the corresponding ones in [11] (as stated at the end of the introduction of our paper).

The following four lemmas on the properties of F' and F^+ under center-Lipschitz condition and the theorem that follows reduce to the corresponding ones in [11] if $L = L_0$. It will become clear from the proof that θ_0 can replace θ in all the results in [11].

Therefore, to avoid repetitions we will only provide the proof of Lemma 4 (see also Lemma 3.1 in [11]). For the rest of the following results, simply replace θ by θ_0 in [11].

Lemma 3. *Let x, $y \in \mathcal{D}$ be such that* rank $F'(x) \leq$ rank $F'(y) = r$, *and* $\theta_0(L_0, y, x) < 1$. *Suppose that F' satisfies the center-Lipschitz condition with L_0 average at y on $\{y, x\}$. Then the following hold:*

$$\mathrm{rank} F'(x) = r, \tag{10}$$

$$\| F'(x) \| \leq \| F'(y)^+ \|^{-1} \left(K(F'(y)) + \theta_0\right), \tag{11}$$

$$\| F'(x)^+ \| \leq \frac{\| F'(y)^+ \|}{1 - \theta_0}, \tag{12}$$

and

$$\| F'(x)^+ - F'(y)^+ \| \leq c_r \frac{\| F'(y)^+ \| \theta_0}{1-\theta_0}. \tag{13}$$

Proof. Simply replace θ by θ_0 in Lemma 3.1 in [11]. □

LEMMA 4. *Let* $x \in \mathcal{D}$, *and* $y \in \mathcal{Z} = \{y \in \mathcal{D} : F'(y)^+ F(y) = 0\}$ *be such that* rank $F'(x) \leq$ rank $F'(y) = r$, *and* $\theta_0(L_0, y, x) < 1$. *Suppose that* F' *satisfies the center-Lipschitz condition with* L_0 *average at* y *on* $\{y, x\}$, *and the radius Lipschitz condition with* L *average at* y *on the line segment* $\{y + \tau(x-y) : 0 \leq \tau \leq 1\}$. *Then for* $\mathcal{N}_F(x) = x - F'(x)^+ F(x)$ $(x \in \mathcal{D})$, *the following hold:*

$$\begin{aligned}\| \mathcal{N}_F(x) - y \| \leq & \| \mathcal{P}_{\mathrm{Ker}F'(y)}(x-y) \| + \theta_0 \| x-y \| \\ & + c_r \left(K(F'(y)) + \theta_0\right) \frac{\theta_0 \| x-y \|}{1-\theta_0} \\ & + \frac{\| F'(y)^+ \| \int_0^{\|x-y\|} tL(t)\mathrm{d}t}{1-\theta_0} + c_r \frac{\| F'(y)^+ \| \| F(y) \| \theta_0}{1-\theta_0}.\end{aligned} \tag{14}$$

Moreover, if $F'(x)$ *is full rank, then the following hold:*

$$\| \mathcal{N}_F(x) - y \| \leq \frac{\| F'(y)^+ \| \int_0^{\|x-y\|} tL(t)\mathrm{d}t}{1-\theta_0} + c_r \frac{\| F'(y)^+ \| \| F(y) \| \theta_0}{1-\theta_0}. \tag{15}$$

Proof. We shall use the identity

$$\begin{aligned}\mathcal{N}_F(x) - y &= \mathcal{P}_{\mathrm{Ker}F'(x)}(x-y) + F'(x)^+ F'(x)(x-y) - F'(x)^+ F(x) \\ &= \mathcal{P}_{\mathrm{Ker}F'(x)}(x-y) + F'(x)^+[F'(x)(x-y) - F(x) + F(y)] \\ &\quad - F'(x)^+ F(y).\end{aligned} \tag{16}$$

We need estimates on

$$E_1 = \| \mathcal{P}_{\mathrm{Ker}F'(x)}(x-y) \|, \quad E_2 = \| F'(x)^+[F'(x)(x-y) - F(x) + F(y)] \|,$$

and $E_3 = \| F'(x)^+ F(y) \|$.

In view of (16), we have:

$$\| \mathcal{N}_F(x) - y \| \leq E_1 + E_2 + E_3. \tag{17}$$

We shall first estimate E_1. We can have:

$$\begin{aligned}\mathcal{P}_{\mathrm{Ker}F'(x)} &= \mathcal{I}_{\mathcal{X}} - F'(y)^+ F'(y) + F'(y)^+ F'(y) - F'(x)^+ F'(x) \\ &= \mathcal{P}_{\mathrm{Ker}F'(x)} + F'(y)^+(F'(y) - F'(x)) + (F'(y)^+ - F'(x)^+)F'(x),\end{aligned} \tag{18}$$

and consequently by (5), (11)–(13), we obtain in turn:

(19)

$$\begin{aligned} E_1 \leq & \| \mathcal{P}_{\mathrm{Ker} F'(y)}(x-y) \| + (\| F'(y)^+ \| \| F'(y) - F'(x) \| \\ & + \| F'(y)^+ - F'(x)^+ \| \| F'(x) \|) \| x-y \| \\ \leq & \| \mathcal{P}_{\mathrm{Ker} F'(y)}(x-y) \| + \| F'(y)^+ \| \int_0^{\|x-y\|} L_0(t) \mathrm{d}t \| x-y \| \\ & + c_r \frac{\| F'(y)^+ \| \theta_0}{1-\theta_0} \| F'(y)^+ \|^{-1} (K(F'(y)) + \theta_0) \| x-y \| \\ = & \| \mathcal{P}_{\mathrm{Ker} F'(y)}(x-y) \| + \theta_0 \| x-y \| + c_r (K(F'(y)) + \theta_0) \frac{\theta_0 \| x-y \|}{1-\theta_0}. \end{aligned}$$

We also estimate E_2. By (6), we have:

$$\begin{aligned} & \| F'(x)(x-y) - F(x) + F(y) \| \\ = & \| \int_0^1 [F'(x) - F'(y+\tau(x-y))](x-y) \mathrm{d}\tau \| \\ \leq & \int_0^1 \int_{\tau\|x-y\|}^{\|x-y\|} L(t) \mathrm{d}t \| x-y \| \mathrm{d}\tau = \int_0^{\|x-y\|} t L(t) \mathrm{d}t. \end{aligned} \tag{20}$$

It then follows from (12) and (20):

$$E_2 \leq \frac{\| F'(y)^+ \| \int_0^{\|x-y\|} t L(t) \mathrm{d}t}{1-\theta_0}. \tag{21}$$

Finally, using (13), we estimate E_3:

$$E_3 \leq \| F'(x)^+ - F'(y)^+ \| \| F(y) \| \leq c_r \frac{\| F'(y)^+ \| \| F(y) \| \theta_0}{1-\theta_0}. \tag{22}$$

Estimate (14) now follows from (17), (19), (21) and (22). Moreover, if $F'(x)$ is of full rank, Ker $F'(x) = \{0\}$. That is $E_1 = 0$. Hence (15) follows from (14). That completes the proof of Lemma 4. □

Lemma 5. (Lemma 3.2 in [11]) *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be twice Fréchet-differentiable and assume $y \in \mathcal{Z} = \{y \in \mathcal{D} : F'(y)^+ F(y) = 0\}$. Then the following holds:*

$$(F'^+ F)'(y) w = \mathcal{P}_{\mathrm{Ker} F'(y)} w + (F'(y)^\star F'(y))^+ (F''(y) w)^\star F(y), \quad w \in \mathcal{D}. \tag{23}$$

Set

$$a(y) = \| F'(y)^+ \|^2 \| F(y) \| \| F''(y) \|, \quad y \in \mathcal{Z}. \tag{24}$$

Lemma 6. (Lemma 3.3 in [11]) *Let $F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ be twice Fréchet-differentiable. Let $x \in \mathcal{D}$, $y \in \mathcal{Z}$, and $a(y) < 1$, where $\mathcal{Z}$ and a are defined in* *the previous lemma and by* (24) *respectively. Then the following hold:*

$$\| \mathcal{P}_{\mathrm{Ker} F'(y)}(x-y) \| \leq \| \mathcal{P}_{\mathrm{Ker}(F'+F)'(y)}(x-y) \| + \left(c_r \frac{a(y)(1+a(y))}{1-a(y)} \right) \| x-y \| . \tag{25}$$

Remark 2. It is convenient to define for $y \in \mathcal{Z}$, $x \in \mathcal{D}$, with $x \neq y$, the parameters P_0, P, Q_0, and Q by:

$$P_0 = P_0(L_0, y, x) = c_r \frac{\| F'(y)^+ \| \| F(y) \| \theta_0}{(1-\theta_0) \| x-y \|} + c_r \frac{a(y)(1+a(y))}{1-a(y)} + a(y), \tag{26}$$

(27)

$$\begin{aligned} Q_0 &= Q_0(L_0, y, x) \\ &= \frac{\| F'(y)^+ \| \displaystyle\int_0^{\|x-y\|} tL(t)\mathrm{d}t}{(1-\theta_0) \| x-y \|^2} + \frac{\theta_0}{\| x-y \|} + c_r \frac{(K(F'(y))+\theta_0)\theta_0}{(1-\theta_0) \| x-y \|} \end{aligned}$$

and P, Q as P_0 and Q_0 respectively by replacing θ_0 by θ.

By convention, we set $P_0(L_0, y, y) = \lim_{x \to y} P_0(L_0, y, x)$, and $Q_0(L_0, y, y) = \lim_{x \to y} Q_0(L_0, y, x)$. Similarly, we define $P(L, y, y)$ and $Q(L, y, y)$.

In view of the above definitions we have:

$$P_0 \leq P, \text{ and} \tag{28}$$

$$Q_0 \leq Q. \tag{29}$$

Moreover, if $L_0(t) < L(t)$ on $[0, \infty)$, then strict inequality holds in (28) and (29).

We can state the main convergence theorem for Newton's method (2):

Theorem 1. *Assume:*
$F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ *is twice Fréchet-differentiable;*
$x^\star \in \mathcal{Z}$, *and* $R > 0$ *such that*

$$\| F'(x^\star)^+ \| \int_0^R L_0(t)\mathrm{d}t < 1, \tag{30}$$

$$\sup\{a(y) : y \in \mathcal{Z} \cap U(x^\star, R)\} < 1, \text{ and} \tag{31}$$

$$v_0 = \sup\{Q_0 \| x-y \| + P_0 : y \in \mathcal{Z} \cap U(x^\star, R), x \in U(x^\star, R)\} < 1; \tag{32}$$

the set $\mathcal{Z} \cap U(x^\star, R)$ *is a smooth submanifold in* $\mathcal{X}$, *with rank* $F'(x) \leq$ *rank* $F'(x^\star)$, $x \in U(x^\star, R)$*;*
the operator F' *satisfies the center-Lipschitz condition with* L_0 *average on*

$\{x^\star, x\}$, and the radius Lipschitz condition with L average at each $y \in \mathcal{Z} \cap U(x^\star, R)$ on $U(x^\star, R)$.

Let

$$R_0^0 = R \min\{1, \frac{1 - v_0}{2v_0}\}. \tag{33}$$

Further, assume $x_0 \in U(x^\star, R_0)$ is such that $x^\star$ is the projection of x_0 onto $\mathcal{Z}$. Then, sequence $\{x_k\}$ generated by Newton's method (2) is well defined, remains in $U(x^\star, R)$ for all $k \geq 0$, and converges to a point in $\mathcal{Z}$.

Moreover, the following estimates hold:

$$d(x_k, \mathcal{Z}) \leq v_0 d(x_{k-1}, \mathcal{Z}), \quad k \geq 1, \tag{34}$$

$$d(x_k, \mathcal{Z}) \leq q_0 d(x_{k-1}, \mathcal{Z})^2 + p_0 d(x_{k-1}, \mathcal{Z}), \quad k \geq 1, \tag{35}$$

where, $d(x_k, \mathcal{Z})$ is the distance from x_k to $\mathcal{Z}$ ($k \geq 0$), and p_0, q_0 are defined by:

$$p_0 = \sup\{P_0 \,:\, y \in \mathcal{Z} \cap U(x^\star, R), x \in U(x^\star, R)\}, \tag{36}$$

and

$$q_0 = \sup\{Q_0 \,:\, y \in \mathcal{Z} \cap U(x^\star, R), x \in U(x^\star, R)\}. \tag{37}$$

Denote also by v, R_0, p, and q the parameters (used in [11]) obtained from v_0, R_0^0, p_0 and q_0 by simply replacing P_0, Q_0 by P, and Q, respectively.

Proof. (Proof of Theorem 1) Simply replace v, R_0, p, q by v_0, R_0^0, p_0, q_0 in the proof of Theorem 3.1 in [11]. □

Remark 3. If $L_0(t) = L(t)$ on $[0, \infty)$, our Theorem 1, reduces to Theorem 3.1 in [11]. Otherwise (i.e., if $L_0(t) < L(t)$), Theorem 1 constitutes an improvement with advantages as already stated in the introduction. Note also that in this case

$$v_0 < v, \tag{38}$$

$$R_0 \leq R_0^0, \tag{39}$$

$$p_0 < p, \tag{40}$$

and

$$q_0 < q. \tag{41}$$

Corollary 1. *Assume:*
$F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ is twice Fréchet-differentiable;
the operator F has zeros;
there exist $x^\star \in \mathcal{Z}$, and $R > 0$ such that

$$\| F'(x^\star)^+ \| \int_0^R L_0(t)\mathrm{d}t < 1, \quad \text{and} \quad \lambda_0 = q_0 R < 1; \tag{42}$$

rank $F'(x) \leq$ rank $F'(x^\star)$ *for $x \in U(x^\star, R)$;*

the operator F' satisfies the center-Lipschitz condition with L_0 average on $\{x^\star, x\}$, and the radius Lipschitz condition with L average at each $y \in \mathcal{Z} \cap U(x^\star, R)$ on $U(x^\star, R)$.

Let

$$R_0^0 = R \min\left\{1, \frac{1-\lambda_0^2}{2\lambda_0}\right\},$$

and $x_0 \in U(x^\star, R)$ be such that $x^\star$ is the projection of x_0 onto $\mathcal{Z}$.

Then, sequence $\{x_k\}$ generated by Newton's method (2) is well defined, remains in $U(x^\star, R)$ for all $k \geq 0$, and converges to a zero of F.

Moreover, the following estimates hold:

$$d(x_k, \mathcal{Z}) \leq q_0 d(x_{k-1}, \mathcal{Z})^2 \leq \lambda_0^{2^k - 1} d(x_0, \mathcal{Z}), \quad k \geq 1. \tag{43}$$

Denote by λ, R_0, the parameters obtained from λ_0, R_0^0 for L_0 replaced by L.

Proof. (Proof of Corollary 1) Simply replace λ, R_0 by λ_0, R_0^0 in the proof of Corollary 3.1 in [11]. □

Remark 4. It follows by (37), (42) and the definition of R_0 and R_0^0 that

$$\lambda_0 \leq \lambda, \tag{44}$$

and

$$R_0 \leq R_0^0. \tag{45}$$

Note also that strict inequality holds in (44) if so does in (7).

3. APPLICATIONS

We provide two applications:

Case 1 (The Kantorovich-type). Let us assume

$$L_0(t) = L_0, \qquad t \in [0, \infty) \tag{46}$$

and

$$L(t) = L, \qquad t \in [0, \infty) \tag{47}$$

for some $L_0 > 0$ and $L > 0$.
We can have the following corollaries:

Corollary 2. *Assume:*
$F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ is twice Fréchet-differentiable;
there exist $x^\star \in \mathcal{Z}$, $R > 0$, $L_0 > 0$ such that

$$L_0 R \parallel F'(x^\star)^+ \parallel < 1, \quad v_0 < 1, \tag{48}$$

and

$$\sup\{a(y) \,:\, y \in \mathcal{Z} \cap U(x^\star, R), x \in U(x^\star, R)\} < 1; \tag{49}$$

$\mathcal{Z} \cap U(x^\star, R)$ is a smooth submanifold in $\mathcal{X}$;
rank $F'(x) \leq$ rank $F'(x^\star)$, $x \in U(x^\star, R)$;
the operator F' satisfies the center-Lipschitz condition, and Lipschitz condition

$$\| F'(x) - F'(x^\star) \| \leq L_0 \| x - x^\star \| \tag{50}$$

and

$$\| F'(x) - F'(y) \| \leq L \| x - y \| \tag{51}$$

for $x \in U(x^\star, R)$ and $y \in \mathcal{Z} \cap U(x^\star, R)$.

Let

$$R_0^0 = R \min \{1, \frac{1 - v_0}{2v_0}\},$$

and let $x_0 \in U(x^\star, R_0^0)$ be such that $x^\star$ is the projection of x_0 onto $\mathcal{Z}$.

Then, sequence $\{x_k\}$ generated by Newton's method (2) is well defined, remains in $U(x^\star, R)$ for all $k \geq 0$, and converges to a point in $\mathcal{Z}$.

Moreover, the following estimates hold:

$$d(x_k, \mathcal{Z}) \leq v_0 d(x_{k-1}, \mathcal{Z}) \leq v_0^k d(x_0, \mathcal{Z}) \quad (k \geq 1), \tag{52}$$

$$d(x_k, \mathcal{Z}) \leq q_0 d(x_{k-1}, \mathcal{Z})^2 + p_0 d(x_{k-1}, \mathcal{Z}) \quad (k \geq 1). \tag{53}$$

Similarly, we can have a corollary corresponding to Corollary 1 in this case.

Case 2 (The Smale point estimate-type). Let

$$L_0(t) = 2b\gamma_0(1 - \gamma_0 t)^{-3} \tag{54}$$

and

$$L(t) = 2b\gamma(1 - \gamma t)^{-3} \tag{55}$$

where

$$b = \frac{1}{\sup\{\| F'(y)^+ \| : \, y \in \mathcal{Z}\}},$$

$\gamma_0 > 0$, and $\gamma > 0$ with

$$\gamma_0 \leq \gamma. \tag{56}$$

We use standard notations:

$$u = \gamma_0 \| x - y \|, \quad \psi(u) = 1 - 4u + 2u^2, \quad \alpha = \| F'(y)^+ \| \| F(y) \| \gamma,$$

and assume

$$\| F'(y)^+ \| \| F''(y) \| \leq 2\gamma, \quad y \in \mathcal{Z}, \tag{57}$$

which leads to

$$a(y) \leq 2 \| F'(y)^+ \| \| F(y) \| \gamma = 2\alpha. \tag{58}$$

Let $R > 0$, $x^\star \in \mathcal{Z}$, and $x \in U(x^\star, R)$.

Then, $u \leq \gamma_0 R$, and

$$\theta_0 < 1, \qquad \text{if} \quad 0 \leq R\gamma_0 < 1 - \frac{\sqrt{2}}{2}, \tag{59}$$

where

$$\theta_0 = b \parallel F'(y^+) \parallel \left(\frac{1}{(1-\gamma_0 \parallel x-y \parallel)^2} - 1 \right), \tag{60}$$

$$\parallel F'(x) - F'(y) \parallel \leq \left(\frac{1}{(1-\gamma_0 \parallel x-y \parallel)^2} - 1 \right) b, \tag{61}$$

$$\parallel F'(x) - F'(x_\tau) \parallel \leq \left(\frac{1}{(1-\gamma \parallel x-y \parallel)^2} - \frac{1}{(1-2\gamma \parallel x-y \parallel)^2} \right) b. \tag{62}$$

As in [6], [11], define for $0 \leq u < 1 - \frac{\sqrt{2}}{2}$, $K \geq 0$, and $\alpha \geq 0$:

$$A_0 = A_0(u,K) = \frac{1}{\psi(u)} + \frac{2-u}{(1-u)^2} + c_r \frac{2-u}{\psi(u)} \left(K + \frac{2u-u^2}{(1-u)^2} \right),$$

$$B_0 = B_0(u,\alpha) = c_r \frac{2-u}{\psi(u)^2} + 2 \left(1 + c_r \frac{1+2\alpha}{1-2\alpha} \right),$$

which lead to

$$P_0 \leq \alpha B_0, \quad Q_0 \leq \gamma_0 A_0, \tag{63}$$

$$q_0 \leq \gamma A_R = \gamma \sup\{A_0(u, K(F'(y))) \,:\, x,\, y \in U(x^\star, R)\}, \tag{64}$$

and

(65)

$$\begin{aligned} v_0 &\leq \Lambda_0 \\ &= \sup\{uA_0(u,K(F'(y))) + \alpha B_0(u,\alpha) \,:\, y \in \mathcal{Z} \cap U(x^\star,R), x \in U(x^\star,R)\}. \end{aligned}$$

It follows from (54)–(65) that Theorem 1 leads to:

Corollary 3. *Assume:*
$F : \mathcal{D} \subseteq \mathcal{X} \longrightarrow \mathcal{Y}$ *is twice Fréchet-differentiable;*
the condition (57) *is satisfied;*
there exist $x^\star \in \mathcal{Z}$, $R > 0$, $\gamma_0 > 0$, $\gamma > 0$ *such that* $U(x^\star,R) \subseteq \mathcal{D}$,

$$0 < R\gamma_0 < 1 - \frac{\sqrt{2}}{2}, \quad v_0 < 1 \quad \text{and} \quad \sup_{y \in \mathcal{Z} \cap U(x^\star,R)} \alpha < \frac{1}{2};$$

and the rest of the hypotheses of Corollary 2 *hold.*

Then the conclusions of Corollary 2 *hold true.*

Remark 5. If $\gamma_0 = \gamma$, the results of this section reduce to corresponding ones in section 4 in [11]. However, if $\gamma_0 < \gamma$, then our results constitute an improvement of the ones in [11] (which in turn improved the ones in [6]) with benefits as stated in the introduction of this paper.

REFERENCES

[1] Argyros, I.K., *On the Newton–Kantorovich hypothesis for solving equations*, Comput. Appl. Math., **169** (2004), 315–332.

[2] Argyros, I.K., *A convergence analysis for Newton–like methods for singular equations using outer or generalized inverses*, Appl. Math. (Warsaw), **32** (2005), 37–49.

[3] Argyros, I.K., *A unifying local–semilocal convergence analysis and applications for two–point Newton–like methods in Banach space*, J. Math. Anal. Appl., **298** (2004), 374–397.

[4] Argyros, I.K., *Computational theory of iterative methods*, Studies in Computational Mathematics, **15**, Elsevier, 2007, New York, U.S.A.

[5] Argyros, I.K., *On the radius of convergence of Newton's method under average mild differentiability conditions*, Nonlinear Funct. Anal. Appl., **13** (2008), 409–415.

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

[7] Gutiérrez, J.M., *A new semilocal convergence theorem for Newton's method*, J. Comput. Appl. Math., **79** (1997), 131–145.

[8] Huang, Z., *A note on the Kantorovich theorem for Newton iteration*, J. Comput. Appl. Math., **47** (1993), 211–217.

[9] Huang, Z., *The convergence ball of Newton's method and the uniqueness ball of equations under Hölder-type continuous derivatives*, Comput. Math. Appl., **47** (2004), 247–251.

[10] Li, C., Zhang, W. and Jin, X., *Convergence and uniqueness properties of the Gauss–Newton's method*, Comput. Math. Appl., **47** (2004), 1057–1067.

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

[12] Smale, S., *Newton's method estimates from data at a point*, The merging of disciplines: New directions in Pure, Applied and Computational Mathematics, R. Ewing, K. Gross and C. Martin eds, New–York, Springer Verlag Publ., (1986), 185–196.

[13] Wang, X., *Convergence of Newton's method and inverse function in Banach space*, Math. Comp., **68** (1999), 169–186.

Received March 28, 2008
Accepted June 9, 2009

Cameron University
Department of Mathematics Sciences
Lawton, OK 73505, USA
E-mail: iargyros@cameron.edu

Poitiers University
Laboratoire de Mathématiques et Applications
Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179
86962 Futuroscope Chasseneuil Cedex, France
E-mail: said.hilout@math.univ-poitiers.fr