Journal of Computational Mathematics, Vol.25, No.6, 2007, 705–718.

CONVERGENCE OF NEWTON'S METHOD FOR SYSTEMS OF EQUATIONS WITH CONSTANT RANK DERIVATIVES*

Xiubin Xu
(*Department of Mathematics, Zhejiang Normal University, Jinhua 321004, China*
Email: xxu@zjnu.cn)
Chong Li
(*Department of Mathematics, Zhejiang University, Hangzhou 310027, China*
Email: cli@zju.edu.cn)

Abstract

The convergence properties of Newton's method for systems of equations with constant rank derivatives are studied under the hypothesis that the derivatives satisfy some weak Lipschitz conditions. The unified convergence results, which include Kantorovich type theorems and Smale's point estimate theorems as special cases, are obtained.

Mathematics subject classification: 49M15, 65F20, 65H10.
Key words: Newton's method, Overdetermined system, Lipschitz condition with L average, Convergence, Rank.

1. Introduction

Let $\mathbb{X}$ and $\mathbb{Y}$ be Euclidean spaces or more generally Banach spaces, and let

$$f : \mathbb{X} \to \mathbb{Y}$$

be a Fréchet differentiable function. Consider the system of nonlinear equations

$$f(x) = 0. \tag{1.1}$$

Such a system is widely used in both theoretical and applied areas of mathematics. Newton's method is the most efficient method for solving such systems. In the case when $f'(x)$ is an isomorphism, there are two points of view to analyze the convergence properties for Newton's method: the Kantorovich type theorem and Smale's point estimate theory. The Kantorovich theorem gives the convergence criteria based on the boundary of f'' in a neighborhood of the initial point x_0, see Ortega and Rheinboldt [10] or Ostrowski [11]; while Smale's point estimate theory gives that based on the invariant

$$\gamma(f, x_0) = \sup_{k\geq 2} \left\| f'(x_0)^{-1} \frac{f^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}, \tag{1.2}$$

see, e.g., Kim [8], Smale [13], Shub and Smale [12]. The convergence criteria based on the radii around a solution of (1.1) were given independently by Traub and Wozniakowski in [14] and Wang in [16]. Since then, there have been many extensions of the above results, see, e.g., [5, 7, 17–20]. In particular, a great progress was made by Wang [17, 18], where the notions of

* Received June 28, 2005; final revised June 29, 2006; accepted June 29, 2006.

Lipschitz conditions with L average were introduced and Kantorovich's and Smale's convergence criteria were unified, see also [21].

Recent attentions have been focused on the study of convergence properties of Newton's method for the case when $f'(x)$ is not an isomorphism. For example, Dedieu and Shub in [3] (resp. [4]) developed the convergence properties for underdetermined (resp. overdetermined) systems with surjective (resp. injective) derivatives under the hypothesis that f is analytic; Li et al. [9] achieved the convergence for overdetermined systems with injective derivatives under the hypothesis that f' satisfies the Lipschitz conditions with L average.

Dedieu and Kim [2] studied the convergence properties of Newton's iteration for analytic systems of equations with constant rank derivatives. They considered an analytic function $f : \mathbb{X} \to \mathbb{Y}$ between two Euclidean spaces and obtained the convergence theorems for solutions and the least square solutions of $f = 0$, respectively. This case generalizes both the surjective-underdetermined case ($\mathrm{rank} f'(x) = \dim \mathbb{Y}$) and the injective-overdetermined case ($\mathrm{rank} f'(x) = \dim \mathbb{X}$).

In this paper, we will investigate the convergence properties of Newton's method for systems of equations with constant rank derivatives under the hypothesis that the derivatives satisfy Lipschitz conditions with L average. The unified convergence properties are obtained. Our results extend and improve those in [2]. We end this section by briefly describing the organization of this paper. The notion of Lipschitz condition with L average and several preliminary results are given in Section 2. The main convergence theorem is stated and proved in Section 3. In Section 4, we discuss the convergence for two special cases. These discussions result in the Kantorovich type results and Smale's point estimate results, respectively. The latter improves the results in [2].

2. Notions and Preliminary Results

In this section, we give some properties related to the Moore-Penrose inverse, which will be used in the next section. Let $A : \mathbb{X} \to \mathbb{Y}$ be a linear operator (or an $m \times n$ matrix). Recall that an operator (or an $n \times m$ matrix) $A^\dagger : \mathbb{Y} \to \mathbb{X}$ is the Moore-Penrose inverse of A if it satisfies the following four equations,

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

where A^* denotes the adjoint of A. Let $\ker A$ and $\mathrm{im} A$ denote the kernel and image of A, respectively. For a subspace E of $\mathbb{X}$, we use Π_E to denote the projection onto E. Then it is clear that

$$A^\dagger A = \Pi_{\ker A^\perp} \text{ and } AA^\dagger = \Pi_{\mathrm{im} A}. \tag{2.1}$$

The following two lemmas give some perturbation bounds for the Moore-Penrose inverse. The first one is stated in [15, Corollary 7.1.1 (2)] and [15, Corollary 7.1.2], while the second one is a direct consequence of [15, Corollary 7.1.1 (2)] and [15, Corollary 7.1.4].

Lemma 2.1. *Let A and B be $m \times n$ matrices and let $r \le \min\{m, n\}$. Suppose that* $\mathrm{rank} A = r$, $\mathrm{rank}(A + B) \le r$ *and* $\|A^\dagger\|\|B\| < 1$. *Then*

$$\mathrm{rank}(A + B) = r \quad \text{and} \quad \|(A + B)^\dagger\| \le \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|B\|}.$$

Lemma 2.2. *Let A and B be $m \times n$ matrices and let $r \leq \min\{m, n\}$. Suppose that* $\mathrm{rank}A = r$, $\mathrm{rank}B \leq r$ *and* $\|A^\dagger\|\|B - A\| < 1$. *Then*

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

where

$$c_r = \begin{cases} \dfrac{1+\sqrt{5}}{2} & if\ r < \min\{m, n\}, \\ \sqrt{2} & if\ r = \min\{m, n\}\ (m \neq n), \\ 1 & if\ r = m = n. \end{cases} \tag{2.2}$$

In the rest of this paper, we assume that $\mathbb{X}$ and $\mathbb{Y}$ are two Euclidean spaces with

$$m \overset{def}{=} \dim\mathbb{X} \leq n \overset{def}{=} \dim\mathbb{Y}$$

and that f is a continuously Fréchet differentiable function from an open subset $U \subseteq \mathbb{X}$ to $\mathbb{Y}$. We also need the following notations. Let $\mathrm{I}_{\mathbb{X}}$ denote the identity on $\mathbb{X}$ and $K(A) = \|A^\dagger\|\|A\|$ the condition number of a linear operator $A : \mathbb{X} \to \mathbb{Y}$. For $\xi_0 \in \mathbb{X}$, $R > 0$, let $\mathbf{B}(\xi_0, R)$ denote the open ball with radius R and center ξ_0.

Let $L(\mu)$ be a positive nondecreasing function defined on $[0, \infty)$. The following notions of Lipschitz conditions with L average were introduced in [17].

Definition 2.1. *Let $V \subseteq \mathbb{X}$ and $\xi \in V$. Let f' be the Fréchet derivative of a function $f : U \subseteq \mathbb{X} \to \mathbb{Y}$. Then f' is said to satisfy*
(i) *the center Lipschitz condition with L average at ξ on V if*

$$\|f'(x) - f'(\xi)\| \leq \int_0^{\|x-\xi\|} L(\mu)d\mu, \quad x \in V; \tag{2.3}$$

(ii) *the radius Lipschitz condition with L average at ξ on V if*

$$\|f'(x) - f'(x^\tau)\| \leq \int_{\tau\|x-\xi\|}^{\|x-\xi\|} L(\mu)d\mu, \qquad x \in V, \quad 0 \leq \tau \leq 1, \tag{2.4}$$

where $x^\tau = \xi + \tau(x - \xi)$.

Remark 2.1. If $f : \mathbb{X} \to \mathbb{Y}$ satisfies the radius Lipschitz condition with L average at ξ on V, then it also satisfies the center Lipschitz condition with L average at ξ on V.

Letting $\xi, x \in \mathbb{X}$, we write, for simplicity,

$$\theta(\xi, x) = \|f'(\xi)^\dagger\| \int_0^{\|x-\xi\|} L(\mu)d\mu. \tag{2.5}$$

The following lemma contains some properties of f' and $f^\dagger$ under the center Lipschitz condition with L average.

Lemma 2.3. *Let $x, \xi \in \mathbb{X}$ be such that* $\mathrm{rank}f'(x) \leq \mathrm{rank}f'(\xi) = r$ *and* $\theta(\xi, x) < 1$. *Suppose that f' satisfies the center Lipschitz condition with L average at ξ on $\{\xi, x\}$. Then*

$$\mathrm{rank}f'(x) = r, \tag{2.6}$$

$$\|f'(x)\| \leq \|f'(\xi)^\dagger\|^{-1}(K(f'(\xi)) + \theta(\xi, x)), \tag{2.7}$$

$$\|f'(x)^\dagger\| \leq \frac{\|f'(\xi)^\dagger\|}{1 - \theta(\xi, x)}, \tag{2.8}$$

$$\|f'(x)^\dagger - f'(\xi)^\dagger\| \leq c_r \frac{\|f'(\xi)^\dagger\|\theta(\xi, x)}{1 - \theta(\xi, x)}. \tag{2.9}$$

Proof. Note that

$$\Pi_{\ker f'(\xi)} + f'(\xi)^\dagger f'(x) = \mathrm{I}_{\mathbb{X}} - f'(\xi)^\dagger \left(f'(\xi) - f'(x)\right).$$

It follows that $\Pi_{\ker f'(\xi)} + f'(\xi)^\dagger f'(x)$ is nonsingular because

$$\|f'(\xi)^\dagger(f'(\xi) - f'(x))\| \le \|f'(\xi)^\dagger\|\|f'(\xi) - f'(x)\| \le \theta(\xi, x) < 1. \tag{2.10}$$

Since, by (2.1),

$$\begin{aligned}\Pi_{\mathrm{im} f'(\xi)} f'(x) &= f'(\xi)f'(\xi)^\dagger f'(x) + f'(\xi)\Pi_{\ker f'(\xi)} \\ &= f'(\xi)\left(f'(\xi)^\dagger f'(x) + \Pi_{\ker f'(\xi)}\right),\end{aligned}$$

we have that

$$\mathrm{rank}\left(\Pi_{\mathrm{im} f'(\xi)} f'(x)\right) = \mathrm{rank} f'(\xi) = r;$$

hence

$$\mathrm{rank} f'(x) \ge \mathrm{rank}\left(\Pi_{\mathrm{im} f'(\xi)} f'(x)\right) = r,$$

which together with the assumptions implies (2.6). The estimate (2.7) follows from the following observation:

$$\|f'(x)\| \le \|f'(\xi)\| + \|f'(x) - f'(\xi)\| \le \|f'(\xi)\| + \int_0^{\|x-\xi\|} L(\mu)d\mu.$$

As for (2.8), we let $A = f'(\xi)$ and $B = f'(x) - f'(\xi)$. Then

$$\mathrm{rank} A = r, \quad \|A^\dagger\|\|B\| \le \theta(\xi, x) < 1,$$

by the assumptions and (2.10). Thus Lemma 2.1 is applicable to conclude that

$$\|f'(x)^\dagger\| = \|(A+B)^\dagger\| \le \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|B\|} \le \frac{\|f'(\xi)^\dagger\|}{1 - \theta(\xi, x)}.$$

So (2.8) is proved. To prove (2.9), let $A = f'(\xi)$ and $B = f'(x)$. Then Lemma 2.2 is applicable. Therefore,

$$\|f'(x)^\dagger - f'(\xi)^\dagger\| \le c_r \frac{\|f'(\xi)^\dagger\|^2\|f'(x) - f'(\xi)\|}{1 - \|f'(\xi)^\dagger\|\|f'(x) - f'(\xi)\|} \le c_r \frac{\|f'(\xi)^\dagger\|\theta(\xi, x)}{1 - \theta(\xi, x)}.$$

This completes the proof of Lemma 2.3. □

3. Convergence Theorem

Recall that f is a continuously Fréchet differentiable function from an open subset $U \subseteq \mathbb{X}$ to $\mathbb{Y}$. Newton's method for f is defined as follows.

$$x_{n+1} = x_n - f'(x_n)^\dagger f(x_n), \quad n = 0, 1, \cdots, \tag{3.1}$$

where $x_0 \in \mathbb{X}$ is an initial value. Note that, in the case when each $f'(x_n)$ is an isomorphism, (3.1) reduces to the classical Newton's method defined by

$$x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \quad n = 0, 1, \cdots. \tag{3.2}$$

Define the Newton operator $N_f : U \longrightarrow \mathbb{X}$ by

$$N_f(x) = x - f'(x)^\dagger f(x), \quad x \in U. \tag{3.3}$$

Then Newton's method (3.1) can be rewritten as

$$x_{n+1} = N_f(x_n), \quad n = 0, 1, \cdots. \tag{3.4}$$

Remark 3.1. Recall that the system (1.1) is a surjective-underdetermined (resp. injective-overdetermined) system if the number of equations is less (resp. greater) than the number of unknowns and $f'(x)$ is of full rank for each $x \in U$. Note that, for surjective-underdetermined systems, the fixed points of N_f are the zeros of f, while for injective-overdetermined systems, the fixed points of N_f are the least square solutions of $f(x) = 0$, which, in general, are not necessarily the zeros of f.

The main results of this paper are given in Theorem 3.1 below. To prepare for the proof, we begin with some lemmas. Let Z denote the set of all the least square solutions of $f = 0$. Then

$$Z = \{\xi \in \mathbb{X} : f'(\xi)^\dagger f(\xi) = 0\}. \tag{3.5}$$

Recall that $\theta(\xi, x)$ is defined in (2.5).

Lemma 3.1. *Let $x \in \mathbb{X}$ and $\xi \in Z$ be such that* $\operatorname{rank} f'(x) \leq \operatorname{rank} f'(\xi)$ *and* $\theta(\xi, x) < 1$. *Suppose that f' satisfies the radius Lipschitz condition with L average at ξ on the line segment $\{\xi + \tau(x - \xi) : 0 \leq \tau \leq 1\}$. Then*

$$\begin{aligned}\|N_f(x) - \xi\| \leq{}& \|\Pi_{\ker f'(\xi)}(x - \xi)\| \\ &+ \theta(\xi, x)\|x - \xi\| + c_r\Big(K(f'(\xi)) + \theta(\xi, x)\Big)\frac{\theta(\xi, x)\|x - \xi\|}{1 - \theta(\xi, x)} \\ &+ \frac{\|f'(\xi)^\dagger\| \int_0^{\|x-\xi\|} \mu L(\mu) d\mu}{1 - \theta(\xi, x)} + c_r \frac{\|f'(\xi)^\dagger\|\|f(\xi)\|\theta(\xi, x)}{1 - \theta(\xi, x)}.\end{aligned} \tag{3.6}$$

Moreover, if $f'(x)$ is additionally of full rank, then

$$\|N_f(x) - \xi\| \leq \frac{\|f'(\xi)^\dagger\| \int_0^{\|x-\xi\|} \mu L(\mu) d\mu}{1 - \theta(\xi, x)} + c_r \frac{\|f'(\xi)^\dagger\|\|f(\xi)\|\theta(\xi, x)}{1 - \theta(\xi, x)}. \tag{3.7}$$

Proof. It follows from (2.1) that

$$\begin{aligned}N_f(x) - \xi &= \Pi_{\ker f'(x)}(x - \xi) + f'(x)^\dagger f'(x)(x - \xi) - f'(x)^\dagger f(x) \\ &= \Pi_{\ker f'(x)}(x - \xi) + f'(x)^\dagger \left(f'(x)(x - \xi) - f(x) + f(\xi)\right) - f'(x)^\dagger f(\xi).\end{aligned} \tag{3.8}$$

Set

$$\begin{aligned}\Delta_1 &= \|\Pi_{\ker f'(x)}(x - \xi)\|, \\ \Delta_2 &= \|f'(x)^\dagger \left(f'(x)(x - \xi) - f(x) + f(\xi)\right)\|, \\ \Delta_3 &= \|f'(x)^\dagger f(\xi)\|.\end{aligned}$$

Then from (3.8) we have

$$\|N_f(x) - \xi\| \leq \Delta_1 + \Delta_2 + \Delta_3. \tag{3.9}$$

Below we will estimate the bounds of Δ_1, Δ_2 and Δ_3. From the assumption $\xi\in Z$ and (2.9), it is easy to see that

$$\begin{aligned}\Delta_3 &= \|f'(x)^\dagger f(\xi)-f'(\xi)^\dagger f(\xi)\|\\ &\le \|f'(x)^\dagger - f'(\xi)^\dagger\|\|f(\xi)\| \le c_r\frac{\|f'(\xi)^\dagger\|\|f(\xi)\|\theta(\xi,x)}{1-\theta(\xi,x)}.\end{aligned}$$

As for Δ_2, since f' satisfies the radius Lipschitz condition with L average at ξ on the line segment $\{\xi+\tau(x-\xi):\ 0\le\tau\le1\}$, we have

$$\begin{aligned}&\|f'(x)(x-\xi)-f(x)+f(\xi)\| = \left\|\int_0^1 \left(f'(x)-f'(\xi+\tau(x-\xi))\right)(x-\xi)d\tau\right\|\\ &\le \int_0^1\int_{\tau\|x-\xi\|}^{\|x-\xi\|} L(\mu)d\mu\|x-\xi\|d\tau = \int_0^{\|x-\xi\|}\mu L(\mu)d\mu.\end{aligned}$$

This, together with (2.8), yields

$$\Delta_2 \le \frac{\|f'(\xi)^\dagger\|\int_0^{\|x-\xi\|}\mu L(\mu)d\mu}{1-\theta(\xi,x)}.$$

It remains to estimate Δ_1. To this end, note that, by (2.1),

$$\begin{aligned}\Pi_{\ker f'(x)} &= \mathrm{I}_{\mathbb{X}} - f'(\xi)^\dagger f'(\xi) + f'(\xi)^\dagger f'(\xi) - f'(x)^\dagger f'(x)\\ &= \Pi_{\ker f'(\xi)} + f'(\xi)^\dagger(f'(\xi)-f'(x)) + (f'(\xi)^\dagger - f'(x)^\dagger)f'(x).\end{aligned}$$

Therefore,

$$\begin{aligned}\Delta_1 &\le \|\Pi_{\ker f'(\xi)}(x-\xi)\|\\ &\quad + \left\{\|f'(\xi)^\dagger\|\|f'(\xi)-f'(x)\| + \|f'(\xi)^\dagger - f'(x)^\dagger\|\|f'(x)\|\right\}\|x-\xi\|\\ &\le \|\Pi_{\ker f'(\xi)}(x-\xi)\| + \|f'(\xi)^\dagger\|\int_0^{\|x-\xi\|}L(\mu)d\mu\|x-\xi\|\\ &\quad + c_r\frac{\|f'(\xi)^\dagger\|\theta(\xi,x)}{1-\theta(\xi,x)}\|f'(\xi)^\dagger\|^{-1}\big(K(f'(\xi))+\theta(\xi,x)\big)\|x-\xi\|\\ &= \|\Pi_{\ker f'(\xi)}(x-\xi)\| + \theta(\xi,x)\|x-\xi\| + c_r\Big(K(f'(\xi))+\theta(\xi,x)\Big)\frac{\theta(\xi,x)\|x-\xi\|}{1-\theta(\xi,x)},\end{aligned}$$

where we have used (2.7)-(2.9) and the assumption that f' satisfies Lipschitz condition with L average. Thus (3.6) holds due to (3.9) and the bounds of Δ_i $(i=1,2,3)$.

Finally, if $f'(x)$ is of full rank, then $\ker f'(x)=\{0\}$. Thus $\Delta_1=0$ and (3.6) holds due to (3.9) and the bounds of Δ_2 and Δ_3. The proof of the lemma is complete. □

The following lemma was given in [2, Lemma 12] where the result was stated for an analytic function f.

Lemma 3.2. *Suppose that f has a second Fréchet derivative on U. Let $\xi\in Z$. Then*

$$(f'^\dagger f)'(\xi)\ \hat{x} = \Pi_{\ker f'(\xi)^\perp}\ \hat{x} + (f'(\xi)^*f'(\xi))^\dagger(f''(\xi)\ \hat{x})^* f(\xi),\quad \hat{x}\in\mathbb{X}. \tag{3.10}$$

For simplicity, set

$$a(\xi) = \|f'(\xi)^\dagger\|^2\|f(\xi)\|\|f''(\xi)\|,\quad \xi\in Z. \tag{3.11}$$

Lemma 3.3. *Suppose that f has a second Fréchet derivative on U. Let $x \in \mathbb{X}$ and $\xi \in Z$ be such that $a(\xi) < 1$. Then*

$$\begin{aligned} &\|\Pi_{\ker f'(\xi)}(x-\xi)\| \\ &\le \|\Pi_{\ker (f'^{\dagger} f)'(\xi)}(x-\xi)\| + \left(c_r \frac{a(\xi)\,(1+a(\xi))}{1-a(\xi)} + a(\xi)\right)\|x-\xi\|. \end{aligned} \tag{3.12}$$

Proof. Define operators A and B respectively by

$$A\hat{x} = \Pi_{\ker f'(\xi)^{\perp}}\hat{x}, \quad \hat{x} \in \mathbb{X}, \tag{3.13}$$

$$B\hat{x} = (f'(\xi)^* f'(\xi))^{\dagger}(f''(\xi)\ \hat{x})^* f(\xi), \quad \hat{x} \in \mathbb{X}. \tag{3.14}$$

Then by Lemma 3.2, we have

$$(f'^{\dagger} f)'(\xi) = A + B. \tag{3.15}$$

From (3.13), we get $\ker A = \ker f'(\xi)$. Consequently,

$$\begin{aligned} \Pi_{\ker f'(\xi)} &= \Pi_{\ker (A+B)} + (\Pi_{\ker A} - \Pi_{\ker (A+B)}) \\ &= \Pi_{\ker (f'^{\dagger} f)'(\xi)} + (\Pi_{\ker A} - \Pi_{\ker (A+B)}). \end{aligned} \tag{3.16}$$

Therefore, to complete the proof of this lemma, it suffices to prove

$$\|\Pi_{\ker A} - \Pi_{\ker (A+B)}\| \le c_r \frac{a(\xi)\,(1+a(\xi))}{1-a(\xi)} + a(\xi). \tag{3.17}$$

Note that, by (3.14),

$$B\hat{x} \in \mathrm{im}(f'(\xi)^* f'(\xi))^{\dagger} = (\ker f'(\xi)^* f'(\xi))^{\perp} = \ker f'(\xi)^{\perp}, \quad \hat{x} \in \mathbb{X}.$$

This together with (3.13) yields $\mathrm{im}(A+B) \subset \ker f'(\xi)^{\perp}$. It follows that

$$\begin{aligned} \mathrm{rank}(A+B) &= \dim(\mathrm{im}(A+B)) \\ &\le \dim(\ker f'(\xi)^{\perp}) = \mathrm{rank}\Pi_{\ker f'(\xi)^{\perp}} = \mathrm{rank} A. \end{aligned}$$

Since

$$\|A^{\dagger}\|\|B\| = \|B\| \le \|(f'(\xi)^* f'(\xi))^{\dagger}\|\|f''(\xi)\|\|f(\xi)\| \le a(\xi) < 1, \tag{3.18}$$

Lemma 2.2 is applicable to get

$$\|(A+B)^{\dagger} - A^{\dagger}\| \le c_r \frac{\|A^{\dagger}\|^2\|B\|}{1-\|A^{\dagger}\|\|B\|}. \tag{3.19}$$

Noting that $\Pi_{\ker A} = \mathrm{I}_{\mathbb{X}} - \Pi_{\ker A^{\perp}}$, it follows from (2.1) that

$$\begin{aligned} \|\Pi_{\ker A} - \Pi_{\ker (A+B)}\| &= \|\Pi_{\ker (A+B)^{\perp}} - \Pi_{\ker A^{\perp}}\| \\ &= \|(A+B)^{\dagger}(A+B) - A^{\dagger}A\| \\ &= \|((A+B)^{\dagger} - A^{\dagger})(A+B) + A^{\dagger}B\| \\ &\le \|(A+B)^{\dagger} - A^{\dagger}\|(\|A\| + \|B\|) + \|A^{\dagger}\|\|B\|. \end{aligned} \tag{3.20}$$

Combining this with (3.18) and (3.19), we have

$$\begin{aligned} \|\Pi_{\ker A} - \Pi_{\ker (A+B)}\| &\le c_r \frac{\|A^{\dagger}\|^2\|B\|}{1-\|A^{\dagger}\|\|B\|}(\|A\| + \|B\|) + \|A^{\dagger}\|\|B\| \\ &\le c_r \frac{a(\xi)\,(1+a(\xi))}{1-a(\xi)} + a(\xi). \end{aligned}$$

Hence (3.17) holds.

In order to state our main results, we introduce more notations. We first recall that c_r, $\theta(\xi, x)$ and $a(\xi)$ are defined in (2.2), (2.5) and (3.11), respectively. For $\xi \in Z$ and $x \in \mathbb{X}$, let

$$P(\xi, x) = c_r \frac{\|f'(\xi)^\dagger\| \|f(\xi)\| \theta(\xi, x)}{\|x - \xi\|(1 - \theta(\xi, x))} + c_r \frac{a(\xi)\,(1 + a(\xi))}{1 - a(\xi)} + a(\xi) \tag{3.21}$$

and

$$Q(\xi, x) = \frac{\|f'(\xi)^\dagger\| \int_0^{\|x-\xi\|} \mu L(\mu) d\mu}{\|x - \xi\|^2 (1 - \theta(\xi, x))} + \frac{\theta(\xi, x)}{\|x - \xi\|} + c_r \frac{\Big(K(f'(\xi)) + \theta(\xi, x)\Big)\theta(\xi, x)}{\|x - \xi\|(1 - \theta(\xi, x))}, \tag{3.22}$$

where we use the convention that

$$P(\xi, \xi) = \lim_{x \to \xi} P(\xi, x), \quad Q(\xi, \xi) = \lim_{x \to \xi} Q(\xi, x).$$

Theorem 3.1. *Suppose that f has the second Fréchet derivative on U. Let $\xi_0 \in Z$ and $R > 0$ be such that*

$$\|f'(\xi_0)^\dagger\| \int_0^R L(\mu) d\mu < 1, \tag{3.23}$$

$$\sup\{a(\xi) : \ \xi \in Z \cap \mathbf{B}(\xi_0, R)\} < 1 \tag{3.24}$$

and

$$\nu \overset{def}{=} \sup\{Q(\xi, x)\|x - \xi\| + P(\xi, x) : \ \xi \in Z \cap \mathbf{B}(\xi_0, R),\ x \in \mathbf{B}(\xi_0, R)\} < 1. \tag{3.25}$$

Suppose that $Z \cap \mathbf{B}(\xi_0, R)$ is a smooth submanifold in $\mathbb{X}$, that $\mathrm{rank} f'(x) \le \mathrm{rank} f'(\xi_0)$ for $x \in \mathbf{B}(\xi_0, R)$ and that f' satisfies the radius Lipschitz condition with L average at each $\xi \in Z \cap \mathbf{B}(\xi_0, R)$ on $\mathbf{B}(\xi_0, R)$. Let $R_0 = \min\{1, \frac{1-\nu}{2\nu}\}R$ and let $x_0 \in \mathbf{B}(\xi_0, R_0)$ be such that ξ_0 is the projection of x_0 onto Z. Then Newton's sequence $\{x_n\}$ generated by (3.1) is contained in $\mathbf{B}(\xi_0, R)$ and converges to a point in Z. Moreover,

$$d(x_n, Z) \le \nu d(x_{n-1}, Z), \quad n = 1, 2, \cdots, \tag{3.26}$$

$$d(x_n, Z) \le q\ d(x_{n-1}, Z)^2 + p\ d(x_{n-1}, Z), \quad n = 1, 2, \cdots, \tag{3.27}$$

where $d(x_n, Z)$ is the distance from x_n to Z, $n = 0, 1, 2, \cdots$,

$$p = \sup\{P(\xi, x) : \ \xi \in Z \cap \mathbf{B}(\xi_0, R),\ x \in \mathbf{B}(\xi_0, R)\}, \tag{3.28}$$

$$q = \sup\{Q(\xi, x) : \ \xi \in Z \cap \mathbf{B}(\xi_0, R),\ x \in \mathbf{B}(\xi_0, R)\}. \tag{3.29}$$

Proof. Let $n = 0, 1, \cdots$. Recall from (3.4) that $x_n = N_f(x_{n-1})$. Let ξ_n be the projection of x_n onto Z. Since $Z \cap \mathbf{B}(\xi_0, R)$ is a smooth submanifold in $\mathbb{X}$, we have

$$\Pi_{\ker(f'^\dagger f)'(\xi_n)}(x_n - \xi_n) = 0.$$

Below we will use induction to verify the following assertions:

$$\left.\begin{array}{l} x_{n+1},\ \xi_{n+1} \in \mathbf{B}(\xi_0, R) \\ \|x_n - \xi_{n-1}\| \le \nu d(x_{n-1}, Z) \\ \|x_n - \xi_{n-1}\| \le q\ d(x_{n-1}, Z)^2 + p\ d(x_{n-1}, Z) \end{array}\right\}, \quad n = 1, 2, \cdots. \tag{3.30}$$

By Lemma 3.1 and 3.3, one gets

$$\|x_1 - \xi_0\| = \|N_f(x_0) - \xi_0\| \le Q(\xi_0, x_0)\|x_0 - \xi_0\|^2 + P(\xi_0, x_0)\|x_0 - \xi_0\|. \tag{3.31}$$

This implies that

$$\|x_1 - \xi_0\| \le q\ d(x_0, Z)^2 + p\ d(x_0, Z)$$

and

$$\|x_1 - \xi_0\| \le \nu\|x_0 - \xi_0\| = \nu\ d(x_0, Z) < R.$$

Hence, $x_1 \in \mathbf{B}(\xi_0, R)$. Moreover, $\xi_1 \in \mathbf{B}(\xi_0, R)$ because

$$\|\xi_1 - \xi_0\| \le \|\xi_1 - x_1\| + \|x_1 - \xi_0\| \le 2\|x_1 - \xi_0\| \le 2\nu\|x_0 - \xi_0\| < R.$$

Hence, (3.30) holds for $n = 1$.

We now proceed by induction. Let k be a positive integer. Assume that (3.30) holds for $n \le k$. Then, similar arguments as in the case when $n = 1$ show that

$$\|x_{k+1} - \xi_k\| \le q\ d(x_k, Z)^2 + p\ d(x_k, Z)$$

and

$$\|x_{k+1} - \xi_k\| \le \nu\|x_k - \xi_k\| = \nu\ d(x_k, Z).$$

Hence, for $n = 1, 2, \cdots, k+1$,

$$d(x_n, Z) = \|x_n - \xi_n\| \le \|x_n - \xi_{n-1}\| \le \nu\|x_{n-1} - \xi_{n-1}\| = \nu d(x_{n-1}, Z),$$

which implies that

$$\|x_n - \xi_{n-1}\| \le \nu^n d(x_0, Z), \quad n = 1, 2, \cdots, k+1.$$

Therefore $x_{k+1}, \xi_{k+1} \in \mathbf{B}(\xi_0, R)$ because

$$\begin{aligned} &\|x_{k+1} - \xi_0\| \\ &\le \|x_{k+1} - \xi_k\| + \|\xi_k - \xi_0\| \le 2\sum_{j=0}^{k} \|x_{j+1} - \xi_j\| \\ &\le 2(\sum_{j=0}^{k} \nu^{j+1})\|x_0 - \xi_0\| \le \frac{2\nu}{1-\nu}\|x_0 - \xi_0\| < R, \end{aligned} \tag{3.32}$$

and

$$\|\xi_{k+1} - \xi_0\| \le \sum_{j=0}^{k} \|\xi_{j+1} - \xi_j\| \le 2\sum_{j=0}^{k} \|x_{j+1} - \xi_j\| < R.$$

We have now proved that (3.30) holds. This together with

$$d(x_n, Z) = \|x_n - \xi_n\| \le \|x_n - \xi_{n-1}\|, \quad n = 1, 2, \cdots$$

gives (3.26) and (3.27). From (3.26), it is not difficult to see that $\{x_n\}$ converges to a point in Z. The proof is complete. □

Corollary 3.1. *Suppose that f has a second Fréchet derivative on U and that $f = 0$ has solutions. Let $\xi_0 \in Z$ and $R > 0$ be such that*

$$\|f'(\xi_0)^\dagger\| \int_0^R L(\mu)d\mu < 1, \quad \lambda \stackrel{def}{=} qR < 1. \tag{3.33}$$

Suppose that $\mathrm{rank} f'(x) \le \mathrm{rank} f'(\xi_0)$ *for $x \in \mathbf{B}(\xi_0, R)$ and that f' satisfies the radius Lipschitz condition with L average at each $\xi \in Z \cap \mathbf{B}(\xi_0, R)$ on $\mathbf{B}(\xi_0, R)$. Let $R_0 = \min\{1, \frac{1-\lambda^2}{2\lambda}\}R$ and let $x_0 \in \mathbf{B}(\xi_0, R_0)$ be such that ξ_0 is the projection of x_0 onto Z. Then Newton's sequence $\{x_n\}$ generated by* (3.1) *is contained in $\mathbf{B}(\xi_0, R)$ and converges to a zero of f. Moreover,*

$$d(x_n, Z) \le q\ d(x_{n-1}, Z)^2 \le \lambda^{2^n-1} d(x_0, Z), \quad n = 1, 2, \cdots. \tag{3.34}$$

Proof. By a result in [6, Chap. I, Sect. 15.2], $Z \cap \mathbf{B}(\xi_0, R)$ is a smooth submanifold. Note that $a(\xi) = 0$ for each $\xi \in Z$ and $p = 0$. Therefore, the proof of this corollary is a simpler version of the proof of Theorem 3.1 except for the following: For $k > 1$, if

$$\|x_n - \xi_n\| \le \lambda^{2^n-1}\|x_0 - \xi_0\|, \quad n = 0, 1, \cdots, k$$

and

$$\|x_{k+1} - \xi_{k+1}\| \le q\|x_k - \xi_k\|^2,$$

then

$$\|x_{k+1} - \xi_{k+1}\| \le q\|x_k - \xi_k\|^2 \le q(\lambda^{2^k-1}\|x_0 - \xi_0\|)^2 \le \lambda^{2^{k+1}-1}\|x_0 - \xi_0\|.$$

While the estimate for $\|x_{k+1} - \xi_0\|$ in (3.32) becomes

$$\|x_{k+1} - \xi_0\| \le 2\sum_{j=0}^{k} \|x_{j+1} - \xi_j\|$$

$$\le 2(\sum_{j=0}^{k} \lambda^{2^{j+1}-1})\|x_0 - \xi_0\| \le \frac{2\lambda}{1-\lambda^2}\|x_0 - \xi_0\| < R.$$

4. Applications

In this section, we will apply the obtained results to some concrete cases. First, we take $L(\mu)$ as a constant $L > 0$. Then the weak Lipschitz conditions in Definition 2.1 become the classical Lipschitz condition. The expressions in (2.5), (3.21) and (3.22) take the forms of

$$\begin{aligned}
&\theta(\xi, x) = L\|f'(\xi)^\dagger\|\|x - \xi\|,\\
&P(\xi, x) = CL\frac{\|f'(\xi)^\dagger\|^2\|f(\xi)\|}{1 - L\|f'(\xi)^\dagger\|\|x - \xi\|} + c_r\frac{a(\xi)\,(1 + a(\xi))}{1 - a(\xi)} + a(\xi),\\
&Q(\xi, x) = L\|f'(\xi)^\dagger\| + CL\frac{\Big(K(f'(\xi)) + L\|f'(\xi)^\dagger\|\|x - \xi\|\Big)\|f'(\xi)^\dagger\|}{1 - L\|f'(\xi)^\dagger\|\|x - \xi\|}\\
&\qquad\qquad + \frac{L}{2}\ \frac{\|f'(\xi)^\dagger\|}{1 - L\|f'(\xi)^\dagger\|\|x - \xi\|},
\end{aligned}$$

respectively. By recalling the definitions of Z, p, q, ν and λ in (3.5), (3.28), (3.29), (3.25) and (3.33), respectively, we immediately obtain the convergence results under the classical Lipschitz condition.

Corollary 4.1. *Suppose that f has a second Fréchet derivative on U. Let $\xi_0 \in Z$, $R > 0$ and $L > 0$ be such that*

$$LR\|f'(\xi_0)^\dagger\| < 1, \quad \nu < 1,$$

and

$$\sup\{a(\xi) : \xi \in Z \cap \mathbf{B}(\xi_0, R)\} < 1.$$

Suppose that $Z \cap \mathbf{B}(\xi_0, R)$ is a smooth submanifold in $\mathbb{X}$, that $\mathrm{rank} f'(x) \leq \mathrm{rank} f'(\xi_0)$ for $x \in \mathbf{B}(\xi_0, R)$ and that f' satisfies the classical Lipschitz condition

$$\|f'(x) - f'(\xi)\| \leq L\|x - \xi\|, \quad x \in \mathbf{B}(\xi_0, R),\ \xi \in Z \cap \mathbf{B}(\xi_0, R).$$

Let $R_0 = \min\{1, \frac{1-\nu}{2\nu}\}R$ and let $x_0 \in \mathbf{B}(\xi_0, R_0)$ be such that ξ_0 is the projection of x_0 onto Z. Then Newton's sequence $\{x_n\}$ generated by (3.1) is contained in $\mathbf{B}(\xi_0, R)$ and converges to a point in Z. Moreover

$$d(x_n, Z) \leq \nu d(x_{n-1}, Z) \leq \nu^n d(x_0, Z), \quad n = 1, 2, \cdots,$$

and

$$d(x_n, Z) \leq q\ d(x_{n-1}, Z)^2 + p\ d(x_{n-1}, Z), \quad n = 1, 2, \cdots.$$

Corollary 4.2. *Suppose that f has a second Fréchet derivative on U and that $f = 0$ has solutions. Let $\xi_0 \in Z$, $R > 0$ and $L > 0$ be such that*

$$LR\|f'(\xi_0)^\dagger\| < 1, \quad \lambda < 1. \tag{4.1}$$

Suppose that $\mathrm{rank} f'(x) \leq \mathrm{rank} f'(\xi_0)$ for $x \in \mathbf{B}(\xi_0, R)$ and that f' satisfies the classical Lipschitz condition

$$\|f'(x) - f'(\xi)\| \leq L\|x - \xi\|, \quad x \in \mathbf{B}(\xi_0, R), \quad \xi \in Z \cap \mathbf{B}(\xi_0, R).$$

Let $R_0 = \min\{1, \frac{1-\lambda^2}{2\lambda}\}R$ and let $x_0 \in \mathbf{B}(\xi_0, R_0)$ be such that ξ_0 is the projection of x_0 onto Z. Then Newton's sequence $\{x_n\}$ generated by (3.1) is contained in $\mathbf{B}(\xi_0, R)$ and converges to a zero of f. Moreover,

$$d(x_n, Z) \leq q\ d(x_{n-1}, Z)^2 \leq \lambda^{2^n-1} d(x_0, Z), \quad n = 1, 2, \cdots.$$

Next, taking $L(\mu) = 2b\gamma(1 - \gamma\mu)^{-3}$, where $b = 1/\sup\{\|f'(\xi)^\dagger\| : \xi \in Z\}$ and $\gamma > 0$ is a real number, we can deduce the improved ones of [2, Theorem 5] and [2, Theorem 6]. In this case, the center and radius Lipschitz conditions defined in (2.3) and (2.4) are given by

$$\|f'(x) - f'(\xi)\| \leq \frac{b}{(1 - \gamma\|x - \xi\|)^2} - b,$$

and

$$\|f'(x) - f'(x^\tau)\| \leq \frac{b}{(1 - \gamma\|x - \xi\|)^2} - \frac{b}{(1 - \tau\gamma\|x - \xi\|)^2}, \tag{4.2}$$

respectively.

For convenience, we adopt some notations used in [2]. Let $v = \gamma\|x - \xi\|$, $\psi(v) = 1 - 4v + 2v^2$ and $\alpha = \|f'(\xi)^\dagger\|\|f(\xi)\|\gamma$ (one should note that v is a function of x and ξ, while α is a function of ξ). Recall that $a(\xi)$ is defined by (3.11). Assuming

$$\|f'(\xi)^\dagger\|\|f''(\xi)\| \leq 2\gamma, \quad \xi \in Z, \tag{4.3}$$

we get

$$a(\xi) \le 2\|f'(\xi)^\dagger\|\|f(\xi)\|\gamma = 2\alpha, \quad \xi \in Z. \tag{4.4}$$

From (2.5), we have

$$\theta(\xi, x) = \|f'(\xi)^\dagger\| \left(\frac{b}{(1-\gamma\|x-\xi\|)^2} - b \right) \le \frac{2v - v^2}{(1-v)^2}.$$

Let $R > 0$, $\xi_0 \in Z$ and $x \in \mathbf{B}(\xi_0, R)$. Then $v \le R\gamma$ and

$$\theta(\xi_0, x) < 1, \quad \text{if } 0 \le R\gamma < 1 - \frac{\sqrt{2}}{2}. \tag{4.5}$$

As in [2], we define the following functions for $0 \le v < 1 - \frac{\sqrt{2}}{2}$, $K \ge 0$ and $\alpha \ge 0$:

$$\begin{aligned} A(v, K) &= \frac{1}{\psi(v)} + \frac{2-v}{(1-v)^2} + c_r \frac{(2-v)}{\psi(v)} \Big(K + \frac{2v - v^2}{(1-v)^2} \Big), \\ B(v, \alpha) &= c_r \frac{2-v}{\psi(v)^2} + 2 \left(1 + c_r \frac{1+2\alpha}{1-2\alpha} \right). \end{aligned}$$

Then it follows from (3.21), (4.4), (3.22), (3.29) and (3.25) that, for $\xi \in Z$ and $x \in \mathbb{X}$,

$$\begin{aligned} P(\xi, x) &\le c_r \alpha \frac{2-v}{\psi(v)^2} + \alpha \left(2c_r \frac{1+2\alpha}{1-2\alpha} + 2 \right) = \alpha B(v, \alpha), \\ Q(\xi, x) &\le \frac{\gamma}{\psi(v)} + \gamma \frac{2-v}{(1-v)^2} + c_r \frac{\gamma(2-v)}{\psi(v)} \Big(K(f'(\xi)) + \frac{2v - v^2}{(1-v)^2} \Big) \\ &= \gamma A(v, K(f'(\xi))), \\ q &\le \gamma A_R \overset{def}{=} \gamma \sup\left\{ A(v, K(f'(\xi))) : \xi \in Z \cap \mathbf{B}(\xi_0, R), \, x \in \mathbf{B}(\xi_0, R) \right\}, && (4.6) \\ \nu &\le \Lambda \overset{def}{=} \sup\left\{ vA(v, K(f'(\xi))) + \alpha B(v, \alpha) : \xi \in Z \cap \mathbf{B}(\xi_0, R), \, x \in \mathbf{B}(\xi_0, R) \right\}. && (4.7) \end{aligned}$$

By (4.3)-(4.5), it is not difficult to see that Theorem 3.1 yields the following corollary.

Corollary 4.3. *Suppose that f has the second Fréchet derivative on U such that (4.3) is satisfied. Let $\xi_0 \in Z$, $R > 0$ and $\gamma > 0$ be such that*

$$0 < R\gamma < 1 - \frac{\sqrt{2}}{2}, \quad \nu < 1 \quad \text{and} \quad \sup_{\xi \in Z \cap \mathbf{B}(\xi_0, R)} \alpha < \frac{1}{2}.$$

Suppose that $Z \cap \mathbf{B}(\xi_0, R)$ is a smooth submanifold in $\mathbb{X}$, that $\operatorname{rank} f'(x) \le \operatorname{rank} f'(\xi_0)$ for $x \in \mathbf{B}(\xi_0, R)$ and that f' satisfies the radius Lipschitz condition (4.2) at each $\xi \in Z \cap \mathbf{B}(\xi_0, R)$ on $\mathbf{B}(\xi_0, R)$. Let $R_0 = \min\{1, \frac{1-\nu}{2\nu}\}R$ and let $x_0 \in \mathbf{B}(\xi_0, R_0)$ be such that ξ_0 is the projection of x_0 onto Z. Then Newton's sequence $\{x_n\}$ generated by (3.1) is contained in $\mathbf{B}(\xi_0, R)$ and converges to a point in Z. Moreover,

$$d(x_n, Z) \le \nu d(x_{n-1}, Z) \le \nu^n d(x_0, Z), \quad n = 1, 2, \cdots,$$

and

$$d(x_n, Z) \le q \, d(x_{n-1}, Z)^2 + p \, d(x_{n-1}, Z), \quad n = 1, 2, \cdots.$$

Remark 4.1. Because $\nu \le \Lambda$ due to (4.7), and the fact that (4.3) is satisfied by the definition of γ in [2], one can see that the above corollary is an improvement of Theorem 6 in [2]. We note that γ and α here are corresponding to γ_R and α_1 in [2], respectively. Moreover, there is a typo in [2, Theorem 6], where the condition $0 < R < 1 - \sqrt{2}/2$ should be $0 < R\gamma_R < 1 - \sqrt{2}/2$.

Recall that $\lambda = qR$, where q is defined in (3.29). Since $\lambda \le \gamma A_R R$ due to (4.6), one can verify that the following corollary of Corollary 3.1 is an improvement of Theorem 5 in [2].

Corollary 4.4. *Suppose that f has a second Fréchet derivative on U and that $f = 0$ has solutions. Let $\xi_0 \in Z$, $R > 0$ and $\gamma > 0$ be such that*

$$0 < R\gamma < 1 - \frac{\sqrt{2}}{2} \quad \text{and} \quad \lambda < 1.$$

Suppose that $\mathrm{rank} f'(x) \le \mathrm{rank} f'(\xi_0)$ *for $x \in \mathbf{B}(\xi_0, R)$ and that f' satisfies the radius Lipschitz condition* (4.2) *at each $\xi \in Z \cap \mathbf{B}(\xi_0, R)$ on $\mathbf{B}(\xi_0, R)$. Let $R_0 = \min\{1, \frac{1-\lambda^2}{2\lambda}\}R$ and let $x_0 \in \mathbf{B}(\xi_0, R_0)$ be such that ξ_0 is the projection of x_0 onto Z. Then Newton's sequence $\{x_n\}$ generated by* (3.1) *is contained in $\mathbf{B}(\xi_0, R)$ and converges to a zero of f. Moreover,*

$$d(x_n, Z) \le q\ d(x_{n-1}, Z)^2 \le \lambda^{2^n - 1} d(x_0, Z), \quad n = 1, 2, \cdots.$$

In fact, the condition $RA_R\gamma \le \frac{1}{2}$ in [2, Theorem 5] implies that $\lambda = qR \le \gamma A_R R \le \frac{1}{2}$ and $R\gamma < 1 - \frac{\sqrt{2}}{2}$ due to $A_R \ge 3$. While the conditions $x_0 \in \mathbf{B}(\xi_0, \frac{3}{4}R)$ in [2, Theorem 5] (misprinted as $x_0 \in \mathbf{B}(\xi_0, \frac{4}{3}R)$ there) and $\lambda \le \frac{1}{2}$ yield that $x_0 \in \mathbf{B}(\xi_0, R_0)$.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China (Grant No. 10671175) and Program for New Century Excellent Talents in Universities. The first author was also supported in part by the Education Ministry of Zhejiang Province (Grant No. 20060492).

References

[1] J.P. Dedieu, Estimations for the separation number of a polynomial system, *J. Symbolic Comput.*, **24** (1997), 683-693.

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

[3] J.P. Dedieu and M. Shub, Multihomogeneous Newton methods, *Math. Comput.*, **69** (2000), 1071-1098.

[4] J.P. Dedieu and M. Shub, Newton's method for overdetermined systems of equations, *Math. Comput.*, **69** (2000), 1099-1115.

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

[6] S. Helgason, Differential Geometry, Lie Groups, and Symmetric Spaces, Academic Press, San Diego, 1979.

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

[8] M.H. Kim, Computational Complexity of the Euler Type Algorithms for the Roots of Polynomials, Ph.D. Thesis, CUNY, January 1986.

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

[10] J. Ortega and V. Rheinboldt, Numerical Solutions of Nonlinear Problems, SIAM, Philadelphia, 1968.

[11] A. Ostrowksi, Solutions of Equations in Euclidean and Banach Spaces, Academic Press, New York, 1976.

[12] M. Shub and S. Smale, Complexity of Bézout's theorem I. Geometric aspects, *J. Amer. Math. Soc.*, **6** (1993), 459-501.

[13] S. Smale, Newton's method estimates from data at one point, in The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics, Eds. R. Ewing, K. Gross, and C. Maring, Springer-Verlag, New York/Berlin, 1986, 185-196.

[14] J. Traub and H. Wozniakowski, Convergence and complexity of Newton iteration for operator equation, *J. Assoc. Comput. Math.*, **29** (1979), 250-258.

[15] G. Wang, Y. Wei and S. Qiao, Generalized Inverse: Theory and Computations, Science Press, Beijing/New York, 2004.

[16] X. Wang, The convergence on Newton's method, *KeXue TongBao, (A Special Issue of Mathematics, Physics and Chemistry)*, **25** (1980), 36-37.

[17] X. Wang, Convergence of Newton's method and inverse function theorem, *Math. Comp.*, **68** (1999), 169-186.

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

[19] X. Wang and D. Han, On dominating sequence method in the point estimate and Smale theorem, *Scientia Sinica Ser. A*, **33** (1990), 135-144.

[20] X. Wang and D. Han, Criteria α and Newton's method, *Chinese J. Numer. and Appl. Math.*, **19** (1997), 96-105.

[21] X. Wang and C. Li, Local and global behavior for algorithms of solving equations, *Chinese Science Bulletin*, **46** (2001), 441-448.

[22] X. Wang and C. Li, Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces II, *Act. Math. Sinica, New Ser.*, **18** (2003), 405-412.