



Contents lists available at ScienceDirect

J. Math. Anal. Appl.

www.elsevier.com/locate/jmaa

Convergence criterion of Newton's method for singular systems with constant rank derivatives [☆]

Xiubin Xu ^{a,*}, Chong Li ^b

^a Department of Mathematics, Zhejiang Normal University, Jinhua 321004, China

^b Department of Mathematics, Zhejiang University, Hangzhou 310027, China

ARTICLE INFO

Article history:

Received 14 August 2007

Available online 10 April 2008

Submitted by T.D. Benavides

Keywords:

Newton's method

Singular system

Lipschitz condition with L -average

Moore–Penrose inverse

Convergence criterion

ABSTRACT

The present paper is concerned with the convergence problem of Newton's method to solve singular systems of equations with constant rank derivatives. Under the hypothesis that the derivatives satisfy a type of weak Lipschitz condition, a convergence criterion based on the information around the initial point is established for Newton's method for singular systems of equations with constant rank derivatives. Applications to two special and important cases: the classical Lipschitz condition and the Smale's assumption, are provided; the latter, in particular, extends and improves the corresponding result due to Dedieu and Kim in [J.P. Dedieu, M. Kim, Newton's method for analytic systems of equations with constant rank derivatives, *J. Complexity* 18 (2002) 187–209].

© 2008 Elsevier Inc. All rights reserved.

1. Introduction

Let \mathbb{X} and \mathbb{Y} be Euclidean spaces. Let G be an open convex subset of \mathbb{X} , and

$$f : G \subset \mathbb{X} \rightarrow \mathbb{Y}$$

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

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

As it is well known, solving such a system is a very general subject which is widely used in both theoretical and applied areas of mathematics. Newton's method with initial point x_0 is defined by

$$x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \quad n = 0, 1, \dots, \tag{1.2}$$

which is the most efficient method known for solving such systems.

There are three types of convergence issues about Newton's method: local, semi-local and global convergence analysis. The first is to determine the convergence ball based on the information in a neighborhood of the solution of $f(x) = 0$, the second is the convergence criterion based on the information only in a neighborhood of the initial point x_0 , and the last is the convergence analysis based on the information on the whole domain of f . In the present paper, we are interested

[☆] Supported in part by the Education Ministry of Zhejiang Province (Grant No. 20060492) for the first author, and the National Natural Science Foundation of China (Grant Nos. 10671175 and 10731060) and Program for New Century Excellent Talents in University for the second author.

* Corresponding author.

E-mail addresses: xxu@zjnu.cn (X. Xu), cli@zju.edu.cn (C. Li).

in the semi-local convergence analysis. As it is well known, there are mainly two points of view to analyze the semi-local convergence: Kantorovich like theorems and Smale’s point estimate theorems. The first kind gives the convergence criterion in terms of the value of the function at the initial point x_0 and the behavior of f'' or f' in a neighborhood of the initial point x_0 with an assumption that f'' is bounded or f' satisfies the Lipschitz conditions, see for example, Ortega and Rheinboldt [11], Ostrowski [12], Kantorovich and Akilov [8]. The second one assumes that f is analytic at the initial point x_0 , and gives the convergence criterion in terms of the following invariants:

$$\left. \begin{aligned} \alpha(f, x_0) &= \beta(f, x_0)\gamma(f, x_0), \\ \beta(f, x_0) &= \|f'(x_0)^{-1}f(x_0)\|, \\ \gamma(f, x_0) &= \sup_{k \geq 2} \left\| f'(x_0)^{-1} \frac{f^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}, \end{aligned} \right\} \tag{1.3}$$

see for example, [9,13,14]. Motivated by the works above, there are many other authors who studied the semi-local convergence of Newton’s method under various of conditions, see [4,17–19]. Gutiérrez in [4] assumed $f''(x)$ satisfies a kind of Lipschitz condition in a neighborhood of the initial point, which is a generalization of Kantorovich like condition. Wang and Han in [19] discussed α criteria under some “weak condition” and generalized Smale’s point estimate theory. In particular, Wang in [17,18] introduced some weak Lipschitz conditions called Lipschitz conditions with L -average, under which Kantorovich like convergence criteria and Smale’s point estimate theory can be put together to be investigated. For a survey on the convergence analysis of Newton’s method, the reader is referred to [20].

All the above mentioned studies are based on the invertibility of f' , which sometimes may fail, that is, f' is singular. One typical example is the case when \mathbb{X} and \mathbb{Y} are two Euclidean spaces with $\dim \mathbb{X} \neq \dim \mathbb{Y}$. Clearly, in this case, f' is not invertible and (1.1) becomes an overdetermined system (i.e. $\dim \mathbb{X} < \dim \mathbb{Y}$) or an underdetermined system (i.e. $\dim \mathbb{X} > \dim \mathbb{Y}$), for which the convergence analysis of Newton’s method has been extensively studied, see for example [3,5–7,13]. In particular, Dedieu and Shub in [3] established Smale’s point estimate theory for Newton’s method for the overdetermined system such that $f'(x)$ is of full rank. Dedieu and Kim in [2] generalized the results in [3] to such case where $f'(x)$ is of constant rank (not necessary full rank). Recently, Li et al. in [10], and Xu and Li in [21] extended respectively the local convergence results in [3] and [2] to the case when the derivative satisfies Lipschitz conditions with L -average.

In the present paper, under the hypothesis that the derivatives satisfy the center Lipschitz condition in the inscribed sphere with L -average introduced in [17], we will investigate the semi-local convergence of Newton’s method for singular systems (not necessary $\dim \mathbb{X} < \dim \mathbb{Y}$) with constant rank derivatives. In Section 2, we introduce some preliminary notions and results. The convergence criterion is established in Section 3. In the last section, applications to two special and important cases: the classical Lipschitz condition and the Smale’s assumption, are provided, and the corresponding convergence result due to Dedieu and Kim in [2] is improved.

2. Notions and preliminary results

In the rest of this paper, \mathbb{X} and \mathbb{Y} denote two Euclidean spaces with $m \stackrel{\text{def}}{=} \dim \mathbb{X}$ and $l \stackrel{\text{def}}{=} \dim \mathbb{Y}$. Let $f : G \subset \mathbb{X} \rightarrow \mathbb{Y}$ be a continuously Fréchet differentiable system with $\text{rank } f'(x) \leq r$ for any $x \in G$, where G is an open convex subset and $r \leq \min\{m, l\}$ is a positive integer. We use the following conventions: $I_{\mathbb{X}}$ denotes the identity on \mathbb{X} and Π_E denotes the orthogonal projection onto a subspace $E \subset \mathbb{X}$. For $\xi_0 \in \mathbb{X}$ and $R > 0$, we use $B(\xi_0, R)$ to denote the open ball with radius R and center ξ_0 .

To give Newton’s method for the case when f' is not of full rank, we need the notion and some properties of Moore–Penrose inverse. For a detailed description of the Moore–Penrose inverse, one can refer to [1,15,16]. Let A be an $l \times m$ matrix (or equivalently, a linear operator $A : \mathbb{X} \rightarrow \mathbb{Y}$). If another $m \times l$ matrix A^\dagger (or equivalently, a linear operator $A^\dagger : \mathbb{Y} \rightarrow \mathbb{X}$), satisfies the following four equalities:

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

where A^* is the conjugate of A , then A^\dagger is called the Moore–Penrose inverse of A . Let $\ker A$ and $\text{im } A$ denote the kernel and image of A , respectively. Then the following properties hold:

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

The next two lemmas are on the perturbation of Moore–Penrose inverse. The first one can be obtained by Corollaries 7.1.1(2) and 7.1.4 in [16], and the two results of the second one are stated in Corollaries 7.1.1(2) and 7.1.2 in [16], respectively.

Lemma 2.1. *Let A and B be two $m \times l$ matrices with $\text{rank } B = \text{rank } A = r$ and $\|A^\dagger\| \|B - A\| < 1$. Then*

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

where

$$C = \begin{cases} \frac{1+\sqrt{5}}{2} & \text{if } r < \min\{m, l\}, \\ \sqrt{2} & \text{if } r = \min\{m, l\} \ (m \neq l), \\ 1 & \text{if } r = m = l. \end{cases} \tag{2.2}$$

Remark 2.1. In the rest of the paper, we only focus on such singular cases when $r < \min\{m, l\}$. Then $C = \frac{1+\sqrt{5}}{2}$ in Lemma 2.1. The full rank case, $r = \min\{m, l\}$, can be studied similarly.

Lemma 2.2. Let A and B be two $m \times l$ matrices with $\text{rank}(A + B) \leq \text{rank } A = r$ and $\|A^\dagger\| \|B\| < 1$. Then

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

Let $L(\mu)$ be a positive nondecreasing function defined on $[0, \infty)$. In order to estimate $\|f'(x_0)^{-1}(f'(x) - f'(y))\|$, Wang in [17] introduced the following concepts of center Lipschitz condition with L -average and center Lipschitz condition in the inscribed sphere with L -average, which play a key role in the study of Newton’s method for nonsingular systems.

Definition 2.1. Suppose that $m = l$ and that f is a continuously Fréchet differentiable function from \mathbb{X} to \mathbb{Y} . Let $R > 0$ and let $x_0 \in \mathbb{X}$ be such that $f'(x_0)^{-1}$ exists.

(i) If

$$\|f'(x_0)^{-1}(f'(x) - f'(x_0))\| \leq \int_0^{\|x-x_0\|} L(\mu) \, d\mu, \quad x \in \mathbf{B}(x_0, R), \tag{2.3}$$

then f' is said to satisfy center Lipschitz condition with L -average in $\mathbf{B}(x_0, R)$.

(ii) If

$$\|f'(x_0)^{-1}(f'(x) - f'(y))\| \leq \int_{\|x-x_0\|}^{\|x-x_0\|+\|y-x\|} L(\mu) \, d\mu \tag{2.4}$$

holds for any $x \in \mathbf{B}(x_0, R)$ and $y \in \mathbf{B}(x, R - \|x - x_0\|)$, then f' is said to satisfy center Lipschitz condition in the inscribed sphere with L -average in $\mathbf{B}(x_0, R)$.

In the following lines, we want to modify this notion to suit to the singular systems considered here. Since $f'(x_0)^{-1}$ does not exist, we will replace $f'(x_0)^{-1}$ by $f'(x_0)^\dagger$. On the other hand, noting that $f'(x_0)^\dagger = B\Pi_{\text{im } f'(x_0)}$, where B is the right inverse of $f'(x_0)$ with the domain $\text{im } f'(x_0)$ and the image $\ker f'(x_0)^\perp$, we lose the information about the component of $f'(x)$ on $\text{im } f'(x_0)^\perp$ via $f'(x_0)^\dagger$. Thus, we give the modification of Definition 2.1 as follows.

Definition 2.2. Suppose that f is a continuously Fréchet differentiable function from \mathbb{X} to \mathbb{Y} . Let $x_0 \in \mathbb{X}$ and $R > 0$.

(i) If

$$\|f'(x_0)^\dagger\| \|f'(x) - f'(x_0)\| \leq \int_0^{\|x-x_0\|} L(\mu) \, d\mu, \quad x \in \mathbf{B}(x_0, R), \tag{2.5}$$

then f' is said to satisfy center Lipschitz condition with L -average in $\mathbf{B}(x_0, R)$.

(ii) If

$$\|f'(x_0)^\dagger\| \|f'(x) - f'(y)\| \leq \int_{\|x-x_0\|}^{\|x-x_0\|+\|y-x\|} L(\mu) \, d\mu \tag{2.6}$$

holds for any $x \in \mathbf{B}(x_0, R)$ and $y \in \mathbf{B}(x, R - \|x - x_0\|)$, then f' is said to satisfy center Lipschitz condition in the inscribed sphere with L -average in $\mathbf{B}(x_0, R)$.

Remark 2.2. From the definitions, we know that if f' satisfies center Lipschitz condition in the inscribed sphere with L -average in $\mathbf{B}(x_0, R)$, then it satisfies center Lipschitz condition with L -average there.

The following lemmas are useful when we prove the convergence theorem in the next section.

Lemma 2.3. *Suppose that f' satisfies center Lipschitz condition with L -average in $\mathbf{B}(x_0, R)$ and that $x \in \mathbf{B}(x_0, R)$ satisfies $\text{rank } f'(x) \leq \text{rank } f'(x_0) = r$ and $\int_0^{\|x-x_0\|} L(\mu) d\mu < 1$. Then the following assertions hold.*

- (i) $\text{rank } f'(x) = r$.
- (ii) $\|f'(x)\| \leq \|f'(x_0)\| + \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|x-x_0\|} L(\mu) d\mu$.
- (iii) $\|f'(x)^\dagger\| \leq \frac{\|f'(x_0)^\dagger\|}{1 - \int_0^{\|x-x_0\|} L(\mu) d\mu}$.

Proof. (i) Noting that

$$\Pi_{\ker f'(x_0)} + f'(x_0)^\dagger f'(x) = I_{\mathbb{X}} - f'(x_0)^\dagger (f'(x_0) - f'(x))$$

and

$$\|f'(x_0)^\dagger\| \|f'(x_0) - f'(x)\| \leq \int_0^{\|x-x_0\|} L(\mu) d\mu < 1, \tag{2.7}$$

we know that $\Pi_{\ker f'(x_0)} + f'(x_0)^\dagger f'(x)$ is nonsingular. By (2.1),

$$\Pi_{\text{im } f'(x_0)} f'(x) = f'(x_0) f'(x_0)^\dagger f'(x) + f'(x_0) \Pi_{\ker f'(x_0)} = f'(x_0) (f'(x_0)^\dagger f'(x) + \Pi_{\ker f'(x_0)}).$$

Hence

$$\text{rank}(\Pi_{\text{im } f'(x_0)} f'(x)) = \text{rank } f'(x_0) = r.$$

Thus

$$\text{rank } f'(x) \geq \text{rank}(\Pi_{\text{im } f'(x_0)} f'(x)) = r.$$

This together with the assumed condition $\text{rank } f'(x) \leq \text{rank } f'(x_0) = r$ implies that (i) holds.

(ii) This assertion follows from

$$\|f'(x)\| \leq \|f'(x_0)\| + \|f'(x) - f'(x_0)\| \leq \|f'(x_0)\| + \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|x-x_0\|} L(\mu) d\mu.$$

(iii) Set $A = f'(x_0)$ and $B = f'(x) - f'(x_0)$. Then by the assumptions and (2.7), $\text{rank } A = r$ and $\|A^\dagger\| \|B\| \leq \int_0^{\|x-x_0\|} L(\mu) d\mu < 1$. Thus, Lemma 2.2 is applicable to concluding that

$$\|f'(x)^\dagger\| = \|(A + B)^\dagger\| \leq \frac{\|A^\dagger\|}{1 - \|A^\dagger\| \|B\|} \leq \frac{\|f'(x_0)^\dagger\|}{1 - \int_0^{\|x-x_0\|} L(\mu) d\mu}.$$

The proof of Lemma 2.3 is complete. \square

Lemma 2.4. *Suppose that f' satisfies center Lipschitz condition in the inscribed sphere with L -average in $\mathbf{B}(x_0, R)$ and $\int_0^R L(\mu) d\mu < 1$. Let $x \in \mathbf{B}(x_0, R)$ and $y \in \mathbf{B}(x, R - \|x - x_0\|)$ be such that $\text{rank } f'(x) \leq \text{rank } f'(x_0) = r$ and $\text{rank } f'(y) \leq \text{rank } f'(x_0) = r$. Then*

$$\|f'(y)^\dagger - f'(x)^\dagger\| \leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x)^\dagger\|^2 \|f'(x_0)^\dagger\|^{-1} \int_0^{\|y-x\|} L(\|x-x_0\| + \mu) d\mu}{1 - \|f'(x)^\dagger\| \|f'(x_0)^\dagger\|^{-1} \int_0^{\|y-x\|} L(\|x-x_0\| + \mu) d\mu}.$$

Proof. Set $A = f'(x)$ and $B = f'(y)$. Then by Lemma 2.3(i) and (iii), we have

$$\text{rank } A = \text{rank } B = \text{rank } f'(x_0) = r$$

and

$$\|A^\dagger\| \leq \frac{\|f'(x_0)^\dagger\|}{1 - \int_0^{\|x-x_0\|} L(\mu) d\mu}. \tag{2.8}$$

By the assumed center Lipschitz condition, one has that

$$\begin{aligned} \|f'(x_0)^\dagger\| \|B - A\| &\leq \int_{\|x-x_0\|}^{\|x-x_0\|+\|y-x\|} L(\mu) d\mu \\ &= \int_0^{\|x-x_0\|+\|y-x\|} L(\mu) d\mu - \int_0^{\|x-x_0\|} L(\mu) d\mu \\ &< \int_0^R L(\mu) d\mu - \int_0^{\|x-x_0\|} L(\mu) d\mu \\ &< 1 - \int_0^{\|x-x_0\|} L(\mu) d\mu. \end{aligned}$$

Combining this with (2.8) yields that $\|A^\dagger\| \|B - A\| < 1$. Therefore Lemma 2.1 is applicable and so

$$\begin{aligned} \|f'(y)^\dagger - f'(x)^\dagger\| &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x)^\dagger\|^2 \|f'(y) - f'(x)\|}{1 - \|f'(x)^\dagger\| \|f'(y) - f'(x)\|} \\ &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x)^\dagger\|^2 \|f'(x_0)^\dagger\|^{-1} \int_{\|x-x_0\|}^{\|x-x_0\|+\|y-x\|} L(\mu) d\mu}{1 - \|f'(x)^\dagger\| \|f'(x_0)^\dagger\|^{-1} \int_{\|x-x_0\|}^{\|x-x_0\|+\|y-x\|} L(\mu) d\mu} \\ &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x)^\dagger\|^2 \|f'(x_0)^\dagger\|^{-1} \int_0^{\|y-x\|} L(\|x-x_0\| + \mu) d\mu}{1 - \|f'(x)^\dagger\| \|f'(x_0)^\dagger\|^{-1} \int_0^{\|y-x\|} L(\|x-x_0\| + \mu) d\mu}, \end{aligned}$$

which completes the proof. \square

The following lemma is a direct consequence of the known results [10, Lemma 2.3] and [17, Proposition 3.2].

Lemma 2.5. Let $\rho \geq 0$ and define the functions ψ_1 and ψ_2 as follows:

$$\psi_1(t) := \frac{1}{t} \int_0^t L(\mu) d\mu, \quad t \in (0, +\infty),$$

and

$$\psi_2(t) := \frac{1}{t} \int_0^t (t - \mu)L(\rho + \mu) d\mu, \quad t \in (0, +\infty).$$

Then ψ_1 and ψ_2 are positive nondecreasing on $(0, +\infty)$.

Lemma 2.6. Suppose that f' satisfies center Lipschitz condition in the inscribed sphere with L -average in $\mathbf{B}(x_0, R)$ and let $x, y \in \mathbf{B}(x_0, R)$ be such that $\|x - x_0\| + \|y - x\| < R$. Then the following assertions hold.

- (i) $\|f'(x)(y - x) + f(x) - f(y)\| \leq \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|y-x\|} (\|y - x\| - \mu)L(\|x - x_0\| + \mu) d\mu.$
- (ii) $\|f(y) - f(x)\| \leq \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|y-x\|} (\|y - x\| - \mu)L(\|x - x_0\| + \mu) d\mu + \|f'(x)\| \|y - x\|.$

Proof. By the assumed center Lipschitz condition, one has that

$$\|f'(x)(y - x) + f(x) - f(y)\| = \left\| \int_0^1 [f'(x) - f'(x + \tau(y - x))](y - x) d\tau \right\|$$

$$\begin{aligned} &\leq \frac{1}{\|f'(x_0)^\dagger\|} \int_0^1 \int_{\|x-x_0\|}^{\|x-x_0\|+\tau\|y-x\|} L(\mu) d\mu \|y-x\| d\tau \\ &= \frac{1}{\|f'(x_0)^\dagger\|} \int_{\|x-x_0\|}^{\|x-x_0\|+\|y-x\|} (\|y-x\| + \|x-x_0\| - \mu)L(\mu) d\mu \\ &= \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|y-x\|} (\|y-x\| - \mu)L(\|x-x_0\| + \mu) d\mu. \end{aligned}$$

This proves (i).

Since

$$\|f(y) - f(x)\| \leq \|f'(x)(y - x) + f(x) - f(y)\| + \|f'(x)\| \|y - x\|,$$

(ii) now follows from (i) and the proof is complete. \square

3. Convergence criterion

In the rest of the paper, we assume that \mathbb{X} and \mathbb{Y} are two Euclidean spaces with finite dimensions, and f is a continuously Fréchet differentiable function from an open subset G of \mathbb{X} to \mathbb{Y} . Let $x_0 \in G$. Newton’s method for f with initial point x_0 is defined by

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

We note that when $f'(x_n)$ is an isomorphism, (3.1) becomes the classical Newton’s method (1.2).

Remark 3.1. Let

$$Z := \{\xi \in \mathbb{X}: f'(\xi)^\dagger f(\xi) = 0\}. \tag{3.2}$$

In general, when f is a singular system, Newton’s method (3.1) may converge to a point in Z rather than a solution of the equation $f = 0$.

Suppose that $L(\mu)$ is a positive nondecreasing function defined on $[0, \infty)$. For simplicity, we introduce some notations. Let

$$K = \|f'(x_0)^\dagger\| \|f'(x_0)\|, \tag{3.3}$$

$$\beta = \|f'(x_0)^\dagger\| \|f(x_0)\|, \tag{3.4}$$

$$\delta_1 = \int_0^\beta L(2\beta + \mu) d\mu, \tag{3.5}$$

$$\delta_2 = \int_0^{2\beta} L(\mu) d\mu, \tag{3.6}$$

$$\Delta_1 = \int_0^\beta (\beta - \mu)L(2\beta + \mu) d\mu, \tag{3.7}$$

$$\Delta_2 = \int_0^{2\beta} (2\beta - \mu)L(2\beta + \mu) d\mu, \tag{3.8}$$

$$p = \frac{\Delta_1}{\beta(1 - \delta_2)} + \frac{1 + \sqrt{5}}{2} \frac{\delta_1(\beta + \beta\delta_2 + 3\beta K + \Delta_1 + \Delta_2)}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)}. \tag{3.9}$$

Our main result is as follows.

Theorem 3.1. Suppose that f' satisfies center Lipschitz condition in the inscribed sphere with L -average in $\mathbf{B}(x_0, 2\beta)$ and suppose that $\text{rank } f'(x) \leq \text{rank } f'(x_0)$ for each $x \in \mathbf{B}(x_0, 2\beta)$. If

$$\delta_1 + \delta_2 < 1 \quad \text{and} \quad p \leq \frac{1}{2}, \tag{3.10}$$

then the Newton's sequence $\{x_n\}$ defined by (3.1) converges to a point ξ in Z , which is defined by (3.2), and the following assertions hold:

$$\|x_{n+1} - x_n\| \leq \left(\frac{1}{2}\right)^n \|x_1 - x_0\|, \quad n = 1, 2, \dots, \tag{3.11}$$

$$\|x_0 - \xi\| \leq 2\|x_1 - x_0\|. \tag{3.12}$$

Proof. To prove (3.11), it suffices to show

$$\|x_{k+1} - x_k\| \leq \frac{1}{2} \|x_k - x_{k-1}\|, \quad k = 1, 2, \dots \tag{3.13}$$

We use mathematical induction to prove (3.13). Since $\delta_1 + \delta_2 < 1$, we have

$$\int_0^{\|x-x_0\|} L(\mu) d\mu \leq \int_0^{2\beta} L(\mu) d\mu = \delta_2 < 1, \quad \forall x \in \mathbf{B}(x_0, 2\beta). \tag{3.14}$$

It follows from Lemma 2.3(i) that $\text{rank } f'(x) = \text{rank } f'(x_0)$. Noting that

$$x_1 - x_0 = -f'(x_0)^\dagger f(x_0) \in \ker f'(x_0)^\perp$$

and

$$f'(x_0)^\dagger f'(x_0) = \Pi_{\ker f'(x_0)^\perp},$$

we have

$$\begin{aligned} x_2 - x_1 &= f'(x_0)^\dagger f'(x_0)(x_1 - x_0) + f'(x_0)^\dagger f(x_0) - f'(x_1)^\dagger f(x_1) \\ &= f'(x_0)^\dagger [f'(x_0)(x_1 - x_0) + f(x_0) - f(x_1)] + [f'(x_0)^\dagger - f'(x_1)^\dagger][f(x_1) - f(x_0)] + [f'(x_0)^\dagger - f'(x_1)^\dagger]f(x_0). \end{aligned}$$

Since $\|x_1 - x_0\| \leq \beta$ and $L(\mu)$ is nondecreasing, it follows from Lemmas 2.4, 2.5 and 2.6 that

$$\begin{aligned} \|x_2 - x_1\| &\leq \int_0^{\|x_1-x_0\|} (\|x_1 - x_0\| - \mu)L(\|x_1 - x_0\| + \mu) d\mu \\ &\quad + \frac{1 + \sqrt{5}}{2} \frac{\int_0^{\|x_1-x_0\|} L(\mu) d\mu}{1 - \int_0^{\|x_1-x_0\|} L(\mu) d\mu} \left(\int_0^{\|x_1-x_0\|} (\|x_1 - x_0\| - \mu)L(\|x_1 - x_0\| + \mu) d\mu + K\|x_1 - x_0\| \right) \\ &\quad + \frac{1 + \sqrt{5}}{2} \frac{\beta \int_0^{\|x_1-x_0\|} L(\mu) d\mu}{1 - \int_0^{\|x_1-x_0\|} L(\mu) d\mu} \\ &\leq \left\{ \frac{\Delta_1}{\beta} + \frac{1 + \sqrt{5}}{2} \frac{\delta_1}{1 - \delta_1} \left(\frac{\Delta_1}{\beta} + K \right) + \frac{1 + \sqrt{5}}{2} \frac{\delta_1}{1 - \delta_1} \right\} \|x_1 - x_0\| \\ &= \left\{ \frac{\Delta_1}{\beta} + \frac{1 + \sqrt{5}}{2} \frac{\delta_1}{\beta(1 - \delta_1)} (\beta + \beta K + \Delta_1) \right\} \|x_1 - x_0\| \\ &\leq p \|x_1 - x_0\| \leq \frac{1}{2} \|x_1 - x_0\|. \end{aligned}$$

This shows that (3.13) holds for $k = 1$.

Suppose that $n > 1$ and (3.13) holds for $k = 2, 3, \dots, n$. With the same argument as estimating $\|x_2 - x_1\|$, we get

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq \|f'(x_{n-1})^\dagger\| \|f'(x_{n-1})(x_n - x_{n-1}) + f(x_{n-1}) - f(x_n)\| + \|f'(x_{n-1})^\dagger - f'(x_n)^\dagger\| \|f(x_n) - f(x_{n-1})\| \\ &\quad + \|f'(x_{n-1})^\dagger - f'(x_n)^\dagger\| \|f(x_{n-1})\| \\ &\stackrel{\text{def}}{=} T_1 + T_2 + T_3. \end{aligned}$$

Below we will estimate T_1 , T_2 and T_3 , respectively. Since (3.13) holds for $k = 2, 3, \dots, n$, it follows that

$$\|x_i - x_0\| \leq \sum_{k=1}^n \|x_k - x_{k-1}\| \leq 2\|x_1 - x_0\| \leq 2\beta, \quad i = n - 1, n. \tag{3.15}$$

Thus, applying Lemma 2.3(iii), one has that

$$\|f'(x_{n-1})^\dagger\| \leq \frac{\|f'(x_0)^\dagger\|}{1 - \int_0^{\|x_{n-1}-x_0\|} L(\mu) d\mu} \leq \frac{\|f'(x_0)^\dagger\|}{1 - \int_0^{2\beta} L(\mu) d\mu} = \frac{\|f'(x_0)^\dagger\|}{1 - \delta_2}. \tag{3.16}$$

By (3.15),

$$L(\|x_n - x_0\| + \mu) \leq L(2\beta + \mu), \quad \forall \mu \geq 0. \tag{3.17}$$

Combining this with the induction hypothesis ($\|x_n - x_{n-1}\| \leq \|x_1 - x_0\| \leq \beta$), we apply Lemma 2.5 to get that

$$\begin{aligned} \int_0^{\|x_n-x_{n-1}\|} (\|x_n - x_{n-1}\| - \mu)L(\|x_n - x_0\| + \mu) d\mu &\leq \frac{1}{\|x_n - x_{n-1}\|} \int_0^{\|x_n-x_{n-1}\|} (\|x_n - x_{n-1}\| - \mu)L(2\beta + \mu) d\mu \|x_n - x_{n-1}\| \\ &\leq \frac{\Delta_1}{\beta} \|x_n - x_{n-1}\|. \end{aligned}$$

This, together with Lemma 2.6(i), yields the following bound of T_1 :

$$T_1 \leq \frac{\|f'(x_{n-1})^\dagger\|}{\|f'(x_0)^\dagger\|} \int_0^{\|x_n-x_{n-1}\|} (\|x_n - x_{n-1}\| - \mu)L(\|x_n - x_0\| + \mu) d\mu \leq \frac{\Delta_1}{\beta(1 - \delta_2)} \|x_n - x_{n-1}\|.$$

To obtain some bounds for T_2 and T_3 , we first estimate their common factor, $\|f'(x_{n-1})^\dagger - f'(x_n)^\dagger\|$. By (3.17), the induction hypothesis and Lemma 2.5, we have

$$\int_0^{\|x_n-x_{n-1}\|} L(\|x_{n-1} - x_0\| + \mu) d\mu \leq \frac{1}{\|x_1 - x_0\|} \int_0^{\|x_1-x_0\|} L(2\beta + \mu) d\mu \|x_n - x_{n-1}\| \leq \frac{\delta_1}{\beta} \|x_n - x_{n-1}\|.$$

From Lemma 2.4, we obtain

$$\begin{aligned} \|f'(x_{n-1})^\dagger - f'(x_n)^\dagger\| &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x_{n-1})^\dagger\|^2 \|f'(x_0)^\dagger\|^{-1} \int_0^{\|x_n-x_{n-1}\|} L(\|x_{n-1} - x_0\| + \mu) d\mu}{1 - \|f'(x_{n-1})^\dagger\| \|f'(x_0)^\dagger\|^{-1} \int_0^{\|x_n-x_{n-1}\|} L(\|x_{n-1} - x_0\| + \mu) d\mu} \\ &\leq \frac{1 + \sqrt{5}}{2} \times \frac{\|f'(x_0)^\dagger\|}{(1 - \delta_2)^2} \times \frac{\delta_1}{\beta} \times \frac{\|x_n - x_{n-1}\|}{1 - \frac{\delta_1}{\beta(1-\delta_2)} \|x_n - x_{n-1}\|} \\ &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x_0)^\dagger\| \delta_1}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)} \|x_n - x_{n-1}\|. \end{aligned} \tag{3.18}$$

For the other factor of T_2 , it follows from Lemma 2.6(ii) and Lemma 2.3(ii) that

$$\begin{aligned} \|f(x_n) - f(x_{n-1})\| &\leq \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|x_n-x_{n-1}\|} (\|x_n - x_{n-1}\| - \mu)L(\|x_{n-1} - x_0\| + \mu) d\mu + \|f'(x_{n-1})\| \|x_n - x_{n-1}\| \\ &\leq \frac{\Delta_1}{\|f'(x_0)^\dagger\|} + \frac{1}{\|f'(x_0)^\dagger\|} \left(K + \int_0^{\|x_{n-1}-x_0\|} L(\mu) d\mu \right) \beta \\ &\leq \frac{\beta\delta_2 + \beta K + \Delta_1}{\|f'(x_0)^\dagger\|}. \end{aligned}$$

This, together with (3.18), yields that

$$\begin{aligned} T_2 &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x_0)^\dagger\| \delta_1}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)} \|x_n - x_{n-1}\| \times \frac{\beta\delta_2 + \beta K + \Delta_1}{\|f'(x_0)^\dagger\|} \\ &= \frac{1 + \sqrt{5}}{2} \frac{\delta_1(\beta\delta_2 + \beta K + \Delta_1)}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)} \|x_n - x_{n-1}\|. \end{aligned}$$

In order to get a bound for T_3 , it suffices to estimate $\|f(x_{n-1})\|$. By Lemma 2.6(i), we have

$$\begin{aligned} \|f(x_{n-1})\| &\leq \|f(x_0)\| + \|f'(x_0)\| \|x_{n-1} - x_0\| + \|f(x_{n-1}) - f(x_0) - f'(x_0)(x_{n-1} - x_0)\| \\ &\leq \|f(x_0)\| + 2\beta \|f'(x_0)\| + \frac{1}{\|f'(x_0)^\dagger\|} \int_0^{\|x_{n-1} - x_0\|} (\|x_{n-1} - x_0\| - \mu)L(\mu) d\mu \\ &\leq \|f(x_0)\| + 2\beta \|f'(x_0)\| + \frac{\Delta_2}{\|f'(x_0)^\dagger\|}. \end{aligned}$$

Hence it follows from (3.18) that

$$\begin{aligned} T_3 &\leq \frac{1 + \sqrt{5}}{2} \frac{\|f'(x_0)^\dagger\| \delta_1}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)} \|x_n - x_{n-1}\| \left(\|f(x_0)\| + 2\beta \|f'(x_0)\| + \frac{\Delta_2}{\|f'(x_0)^\dagger\|} \right) \\ &\leq \frac{1 + \sqrt{5}}{2} \frac{\delta_1(\beta + 2\beta K + \Delta_2)}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)} \|x_n - x_{n-1}\|. \end{aligned}$$

Therefore

$$\begin{aligned} \|x_{n+1} - x_n\| &\leq T_1 + T_2 + T_3 \\ &\leq \left\{ \frac{\Delta_1}{\beta(1 - \delta_2)} + \frac{1 + \sqrt{5}}{2} \frac{\delta_1(\beta + \beta\delta_2 + 3\beta K + \Delta_1 + \Delta_2)}{\beta(1 - \delta_2)(1 - \delta_1 - \delta_2)} \right\} \|x_n - x_{n-1}\| \\ &= p \|x_n - x_{n-1}\| \\ &\leq \frac{1}{2} \|x_n - x_{n-1}\|. \end{aligned}$$

Hence (3.13) holds by mathematical induction.

We turn to prove the rest of the theorem. From (3.11), it is not difficult to see that $\{x_n\}$ is a Cauchy sequence, thus, it converges to a point in \mathbb{X} , say ξ . Letting $n \rightarrow \infty$ in

$$x_{n+1} = x_n - f'(x_n)^\dagger f(x_n),$$

we get $f'(\xi)^\dagger f(\xi) = 0$, that is, $\xi \in Z$. Since

$$\|x_{n+1} - x_0\| \leq \sum_{k=1}^{n+1} \|x_k - x_{k-1}\| \leq \left(\sum_{k=1}^{n+1} \left(\frac{1}{2}\right)^{k-1} \right) \|x_1 - x_0\| \leq 2\|x_1 - x_0\|,$$

taking $n \rightarrow \infty$ gives (3.12). The proof is complete. \square

4. Applications

In this section, we will apply the obtained results in the previous section to two concrete cases, one is the case when $L(\mu) \equiv L$ is a constant function and the other case when L is defined by

$$L(\mu) := 2\gamma/(1 - \gamma\mu)^3, \quad \forall \mu \in \left[0, \frac{1}{\gamma}\right), \tag{4.1}$$

where $\gamma > 0$ is a constant. For the first case, Theorem 3.1 gives the Kantorovich like convergence result; and for the second case, it gives Smale's point estimate result. The latter one improves the corresponding convergence result in [2].

When $L(\mu) \equiv L$ is a constant function, the center Lipschitz condition in the inscribed sphere with L -average (2.6) becomes

$$\|f'(x_0)^\dagger\| \|f'(x) - f'(y)\| \leq L\|y - x\|, \quad \forall x \in \mathbf{B}(x_0, R), \quad y \in \mathbf{B}(x, R - \|x - x_0\|). \tag{4.2}$$

Therefore, if f satisfies the classical Lipschitz condition in $\mathbf{B}(x_0, R)$, then f satisfies center Lipschitz condition in the inscribed sphere with constant L -average in $\mathbf{B}(x_0, R)$. In this case, the expressions (3.5)–(3.8) and (3.9) take the following forms, respectively,

$$\begin{aligned} \delta_1 &= L\beta = L\|f'(x_0)^\dagger\| \|f(x_0)\|, \\ \delta_2 &= 2L\beta = 2\delta_1, \\ \Delta_1 &= \frac{L}{2}\beta^2 = \frac{1}{2}\delta_1\beta, \\ \Delta_2 &= 2L\beta^2 = 2\delta_1\beta, \end{aligned}$$

and

$$p = \frac{\delta_1}{2(1-2\delta_1)} \left(1 + \frac{1 + \sqrt{5} 2 + 6K + 9\delta_1}{2(1-3\delta_1)} \right),$$

where K and β are defined by (3.3) and (3.4), respectively. Thus, by Theorem 3.1, the following corollary is immediate.

Corollary 4.1. *Suppose that f' satisfies (4.2) in the ball $\mathbf{B}(x_0, 2\beta)$, and that $\text{rank } f'(x) \leq \text{rank } f'(x_0)$ for each $x \in \mathbf{B}(x_0, 2\beta)$. If*

$$\delta_1 = L \|f'(x_0)^\dagger\| \|f(x_0)\| < \frac{1}{3}$$

and

$$p = \frac{\delta_1}{2(1-2\delta_1)} \left(1 + \frac{1 + \sqrt{5} 2 + 6K + 9\delta_1}{2(1-3\delta_1)} \right) \leq \frac{1}{2},$$

then Newton's sequence $\{x_n\}$ defined by (3.1) converges to a point $\xi \in Z$, and (3.11) and (3.12) hold, where Z is defined by (3.2).

We can use a sufficient condition of $p \leq \frac{1}{2}$ to simplify the conditions of Corollary 4.1. Let

$$g(d) = \frac{d}{2(1-2d)} \left(1 + \frac{1 + \sqrt{5} 8 + 9d}{2(1-3d)} \right).$$

We can prove that $g(d)$ is increasing on $[0, \frac{1}{3})$. Since the root of $g(d) - \frac{1}{2} = 0$ in $[0, \frac{1}{3})$ is $d_0 := 0.0519926\dots$, we have $g(d) \leq \frac{1}{2}$ whenever $0 \leq d \leq d_0$.

Corollary 4.2. *Suppose that f' satisfies (4.2) in the ball $\mathbf{B}(x_0, 2\beta)$, and that $\text{rank } f'(x) \leq \text{rank } f'(x_0)$ for each $x \in \mathbf{B}(x_0, 2\beta)$. If*

$$K\delta_1 = L \|f'(x_0)^\dagger\|^2 \|f'(x_0)\| \|f(x_0)\| \leq d_0 = 0.0519926\dots,$$

then Newton's sequence $\{x_n\}$ defined by (3.1) converges to a point $\xi \in Z$, and (3.11) and (3.12) hold, where Z is defined by (3.2).

Proof. Suppose that $K\delta_1 \leq d_0$. Since $K \geq 1$, it follows that $\delta_1 < \frac{1}{3}$ and

$$p \leq \frac{1}{2(1-2K\delta_1)} \left(K\delta_1 + \frac{1 + \sqrt{5} 8K\delta_1 + 9(K\delta_1)^2}{2(1-3K\delta_1)} \right) = g(K\delta_1) \leq g(d_0) = \frac{1}{2}.$$

Thus the conditions in Corollary 4.1 are satisfied. Hence the conclusion holds by Corollary 4.1 and we complete the proof. \square

Next, taking L to be the function defined by (4.1), then the center Lipschitz condition in the inscribed sphere with L -average (2.6) becomes

$$\|f'(x_0)^\dagger\| \|f'(x) - f'(y)\| \leq \frac{1}{(1 - \gamma \|x - x_0\| - \gamma \|y - x\|)^2} - \frac{1}{(1 - \gamma \|x - x_0\|)^2} \tag{4.3}$$

for each $x \in \mathbf{B}(x_0, \frac{1}{\gamma})$ and $y \in \mathbf{B}(x, \frac{1}{\gamma} - \|x - x_0\|)$. We adopt the traditional notation used in Smale's point estimate theory, i.e., $\alpha = \beta\gamma$. Assume that $\alpha < \frac{1}{2}$. Then we have from (3.5)–(3.8) and (3.9) that

$$\delta_1 = \frac{\alpha(2-5\alpha)}{(1-2\alpha)^2(1-3\alpha)^2} \leq \frac{K\alpha(2-5K\alpha)}{(1-2K\alpha)^2(1-3K\alpha)^2} \quad \text{if } K\alpha < \frac{1}{5}, \tag{4.4}$$

$$\delta_2 = \frac{4\alpha(1-\alpha)}{(1-2\alpha)^2} \leq \frac{4K\alpha(1-K\alpha)}{(1-2K\alpha)^2} \quad \text{if } K\alpha < \frac{1}{2}, \tag{4.5}$$

$$\Delta_1 = \frac{\beta\alpha}{(1-2\alpha)^2(1-3\alpha)} \leq \frac{\beta K\alpha}{(1-2K\alpha)^2(1-3K\alpha)} \quad \text{if } K\alpha < \frac{1}{3}, \tag{4.6}$$

$$\Delta_2 = \frac{4\beta\alpha}{(1-2\alpha)^2(1-4\alpha)} \leq \frac{4\beta K\alpha}{(1-2K\alpha)^2(1-4K\alpha)} \quad \text{if } K\alpha < \frac{1}{4}, \tag{4.7}$$

and

$$p = \frac{\alpha}{(1-3\alpha)(1-8\alpha+8\alpha^2)} + \frac{1+\sqrt{5}}{2} \frac{3K\alpha(2-5\alpha)}{(1-8\alpha+8\alpha^2)(1-12\alpha+18\alpha^2)} + \frac{1+\sqrt{5}}{2} \frac{\alpha(2-5\alpha)(1-2\alpha-4\alpha^2)}{(1-3\alpha)(1-4\alpha)(1-2\alpha)^2(1-8\alpha+8\alpha^2)(1-12\alpha+18\alpha^2)}. \tag{4.8}$$

Noting that $2\beta \leq \frac{1}{\gamma}$ if $2\alpha \leq 1$, we obtain the following corollary from Theorem 3.1.

Corollary 4.3. Suppose that f' satisfies (4.3) for any $x \in \mathbf{B}(x_0, 2\beta)$ and $y \in \mathbf{B}(x, 2\beta - \|x - x_0\|)$, and that $\text{rank } f'(x) \leq \text{rank } f'(x_0)$ for any $x \in \mathbf{B}(x_0, 2\beta)$. Assume $2\alpha \leq 1$. If

$$\delta_1 + \delta_2 < 1 \quad \text{and} \quad p \leq \frac{1}{2},$$

where δ_1, δ_2 and p are expressed in (4.4), (4.5) and (4.8), respectively, then Newton's sequence $\{x_n\}$ defined in (3.1) converges to a point $\xi \in Z$, and (3.11) and (3.12) hold, where Z is defined by (3.2).

The conditions in the above corollary can be simplified. Let the function h be defined by

$$h(d) = \frac{d}{(1-3d)(1-8d+8d^2)} + \frac{1+\sqrt{5}}{2} \frac{3d(2-5d)}{(1-8d+8d^2)(1-12d+18d^2)} + \frac{1+\sqrt{5}}{2} \frac{d(2-5d)(1-2d-4d^2)}{(1-3d)(1-4d)(1-2d)^2(1-8d+8d^2)(1-12d+18d^2)}$$

for each $d \in [0, \frac{2-\sqrt{2}}{6})$. Since h' is positive on $(0, \frac{2-\sqrt{2}}{6})$, $h(d)$ is increasing on $[0, \frac{2-\sqrt{2}}{6})$. Let \bar{d}_0 be the root of $h(d) - \frac{1}{2} = 0$ in $[0, \frac{2-\sqrt{2}}{6})$. Then $\bar{d}_0 = 0.0223063\dots$, and $h(d) \leq \frac{1}{2}$ whenever $0 \leq d \leq \bar{d}_0$.

Corollary 4.4. Suppose that f' satisfies (4.3) for any $x \in \mathbf{B}(x_0, 2\beta)$ and $y \in \mathbf{B}(x, 2\beta - \|x - x_0\|)$, and that $\text{rank } f'(x) \leq \text{rank } f'(x_0)$ for any $x \in \mathbf{B}(x_0, 2\beta)$. If

$$K\alpha \leq \bar{d}_0 = 0.0223063\dots, \tag{4.9}$$

then Newton's sequence $\{x_n\}$ defined in (3.1) converges to a point $\xi \in Z$, and (3.11) and (3.12) hold, where Z is defined by (3.2).

Proof. Suppose that (4.9) holds. Then $K\alpha < \frac{1}{5}$. Hence the inequalities (4.4)–(4.7) hold and

$$\delta_1 K = \frac{K\alpha(2-5\alpha)}{(1-2\alpha)^2(1-3\alpha)^2} \leq \frac{K\alpha(2-5K\alpha)}{(1-2K\alpha)^2(1-3K\alpha)^2}. \tag{4.10}$$

By (3.9),

$$p = \frac{\Delta_1}{\beta(1-\delta_2)} + \frac{1+\sqrt{5}}{2} \frac{\delta_1(\beta + \beta\delta_2 + 3\beta K + \Delta_1 + \Delta_2)}{\beta(1-\delta_2)(1-\delta_1-\delta_2)}.$$

It follows from (4.4)–(4.7) and (4.10) that $p \leq h(K\alpha)$. Consequently, when $K\alpha \leq \bar{d}_0$, we have $p \leq \frac{1}{2}$, and $\delta_1 + \delta_2 < 1$ (actually, $\delta_1 + \delta_2 < 1$ holds if $\alpha < 0.09$). Thus the conditions in the previous corollary are satisfied, and the current corollary holds. \square

Remark 4.1. Suppose that f is analytic at x_0 , and let

$$\gamma = \sup_{k \geq 2} \left(\|f'(x_0)^\dagger\| \left\| \frac{f^{(k)}(x_0)}{k!} \right\| \right)^{\frac{1}{k-1}} < \infty. \tag{4.11}$$

In [2, Theorem 7], Dediou and Kim proved the conclusion of the above corollary under the condition $K\alpha < \frac{1}{48} \approx 0.0208333$. In this case, similar arguments to that in [17, p. 178], we can show that f' satisfies (4.3) on $\mathbf{B}(x_0, \frac{1}{\gamma})$ with γ defined by (4.11). Therefore Corollary 4.4 is an improvement of [2, Theorem 7].

We end this paper with two examples to illustrate our theoretical results.

Example 4.1. Let \mathbb{R}^2 be endowed with the l_1 -norm and define

$$f(x) := \left(t - s, \frac{1}{2}(t - s)^2 \right)^T, \quad x = (t, s)^T \in \mathbb{R}^2.$$

Then f is a continuously Fréchet differentiable function from \mathbb{R}^2 to \mathbb{R}^2 , and

$$f'(x) = \begin{pmatrix} 1 & -1 \\ t-s & s-t \end{pmatrix}, \quad x = (t, s)^T \in \mathbb{R}^2.$$

Hence, $\text{rank } f'(x) = 1$ for each $x \in \mathbb{R}^2$, and its Moore–Penrose inverse is

$$f'(x)^\dagger = \frac{1}{2[1+(t-s)^2]} \begin{pmatrix} 1 & t-s \\ -1 & s-t \end{pmatrix}, \quad x = (t, s)^T \in \mathbb{R}^2.$$

Let $x_0 = (\frac{27}{50}, \frac{1}{2})^T$. Then $\|f'(x_0)^\dagger\| = \frac{625}{626}$. Furthermore, since for any $x = (t_1, s_1)^T, y = (t_2, s_2)^T \in \mathbb{R}^2$,

$$\|f'(x_0)^\dagger\| \|f'(x) - f'(y)\| \leq \frac{625}{626} (|t_1 - t_2| + |s_1 - s_2|) \leq \frac{625}{626} \|y - x\|,$$

one sees that f' satisfies (4.2) in \mathbb{R}^2 with $L = \frac{625}{626}$. Note that

$$\delta_1 = L \|f'(x_0)^\dagger\| \|f(x_0)\| = \left(\frac{625}{626}\right)^2 \times \frac{51}{1250} = \frac{31875}{783752}$$

and

$$K = \|f'(x_0)^\dagger\| \|f'(x_0)\| = \frac{625}{626} \times \frac{26}{25} = \frac{325}{313}.$$

It follows that

$$K\delta_1 = \frac{325}{313} \times \frac{31875}{783752} = \frac{10359375}{245314376} \approx 0.042229 < d_0.$$

Hence the assumptions in Corollary 4.2 are satisfied and Newton's sequence $\{x_n\}$ defined in (3.1) with $x_0 = (\frac{27}{50}, \frac{1}{2})^T$ converges to a point $\xi \in Z$ satisfying (3.11) and (3.12).

Example 4.2. Let \mathbb{R}^2 be endowed with the l_1 -norm, $x_0 = (0, 0)^T$ and $0 < a < \bar{d}_0$, where $\bar{d}_0 = 0.0223063\dots$ as given in Corollary 4.4. Define

$$f(x) := (a + t - s, a + \ln(1 + t - s))^T, \quad x = (t, s)^T \in \mathbf{B}(x_0, 1).$$

Then f is analytic on $\mathbf{B}(x_0, 1)$, and

$$f'(x) = \begin{pmatrix} 1 & -1 \\ \frac{1}{1+t-s} & -\frac{1}{1+t-s} \end{pmatrix}, \quad x = (t, s)^T \in \mathbf{B}(x_0, 1).$$

Hence, $\text{rank } f'(x) = 1$ for each $x \in \mathbf{B}(x_0, 1)$. In particular, $f'(x_0) = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ and its Moore–Penrose inverse is

$$f'(x_0)^\dagger = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

Moreover, by mathematical induction, we can easily get that, for each $x = (t, s)^T \in \mathbf{B}(x_0, 1)$,

$$f^{(k)}(x)u_1u_2 \cdots u_k = (-1)^{k+1}(k-1)! \prod_{i=1}^k (u_i^1 - u_i^2) \begin{pmatrix} 0 \\ \frac{1}{(1+t-s)^k} \end{pmatrix},$$

where $u_i = (u_i^1, u_i^2)^T \in \mathbb{R}^2$ for each $i = 1, 2, \dots, k$. Consequently,

$$\|f'(x_0)^\dagger\| = \frac{1}{2} \quad \text{and} \quad \|f^{(k)}(x_0)\| = (k-1)!.$$

This implies that

$$\gamma := \sup_{k \geq 2} \left(\|f'(x_0)^\dagger\| \left\| \frac{f^{(k)}(x_0)}{k!} \right\| \right)^{\frac{1}{k-1}} = 1.$$

By Remark 4.1, f' satisfies (4.3) on $\mathbf{B}(x_0, \frac{1}{\gamma})$ with $\gamma = 1$. Note that

$$\alpha = \beta\gamma = \|f'(x_0)^\dagger\| \|f(x_0)\| = a \quad \text{and} \quad K = \|f'(x_0)^\dagger\| \|f'(x_0)\| = 1.$$

It follows that

$$K\alpha = a < \bar{d}_0.$$

Hence the assumptions in Corollary 4.4 are satisfied and Newton's sequence $\{x_n\}$ defined in (3.1) with $x_0 = (0, 0)^T$ converges to a point $\xi \in Z$ satisfying (3.11) and (3.12).

References

- [1] A. Ben-Israel, T.N.E. Greville, *Generalized Inverses: Theory and Applications*, John Wiley, New York, 1974; second ed., Springer-Verlag, New York, 2003.
- [2] J.P. Dedieu, M. Kim, Newton's method for analytic systems of equations with constant rank derivatives, *J. Complexity* 18 (2002) 187–209.
- [3] J.P. Dedieu, M. Shub, Newton's method for overdetermined systems of equations, *Math. Comput.* 69 (2000) 1099–1115.
- [4] J.M. Gutiérrez, A new semilocal convergence theorem for Newton method, *J. Comput. Appl. Math.* 79 (1997) 131–145.
- [5] W.M. Häußler, A Kantorovich-type convergence analysis for the Gauss–Newton-method, *Numer. Math.* 48 (1986) 119–125.
- [6] J.S. He, J.H. Wang, C. Li, Newton's method for underdetermined systems of equations under the modified γ -condition, *Numer. Funct. Anal. Optim.* 28 (2007) 663–679.
- [7] N.C. Hu, W.P. Shen, C. Li, Kantorovich's type theorems for systems of equations with constant rank derivatives, *J. Comput. Appl. Math.* (2007), doi:10.1016/j.cam.2007.07.006.
- [8] L.V. Kantorovich, G.P. Akilov, *Functional Analysis*, translated from the Russian by Howard L. Silcock, second ed., Pergamon Press, Oxford/Elmsford, 1982.
- [9] M.H. Kim, Computational complexity of the Euler type algorithms for the roots of polynomials, thesis, CUNY, January 1986.
- [10] C. Li, W.H. Zhang, X.Q. Jin, Convergence and uniqueness properties of Gauss–Newton's method, *Comput. Math. Appl.* 47 (2004) 1057–1067.
- [11] J. Ortega, V. Rheinboldt, *Numerical Solutions of Nonlinear Problems*, SIAM, Philadelphia, 1968.
- [12] A. Ostrowski, *Solutions of Equations in Euclidean and Banach Spaces*, Academic Press, New York, 1976.
- [13] M. Shub, S. Smale, Complexity of Bézout's theorem I. Geometric aspects, *J. Amer. Math. Soc.* 6 (1993) 459–501.
- [14] S. Smale, Newton's method estimates from data at one point, in: R. Ewing, K. Gross, C. Maring (Eds.), *The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics*, Springer-Verlag, New York, 1986, pp. 185–196.
- [15] G.W. Stewart, J.G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
- [16] G.R. Wang, Y.M. Wei, S. Qiao, *Generalized Inverse: Theory and Computations*, Science Press, Beijing, 2004.
- [17] X.H. Wang, Convergence of Newton's method and inverse function theorem in Banach spaces, *Math. Comp.* 68 (1999) 169–186.
- [18] X.H. Wang, Convergence of Newton's method and uniqueness of the solution of equations in Banach spaces, *IMA J. Numer. Anal.* 20 (2000) 123–134.
- [19] X.H. Wang, D.H. Han, Criteria α and Newton's method, *Chinese J. Numer. Math. Appl.* 19 (1997) 96–105.
- [20] X.H. Wang, C. Li, Local and global behavior for algorithms of solving equations, *Chinese Sci. Bull.* 46 (2001) 444–451.
- [21] X.B. Xu, C. Li, Convergence of Newton's method for systems of equations with constant rank derivatives, *J. Comput. Math.* 25 (2007) 705–718.