Kantorovich’s type theorems for systems of equations with constant rank derivatives

Nuchun Hu, Weiping Shen, Chong Li*  
Department of Mathematics, Zhejiang University, Hangzhou 310027, PR China  
Received 26 March 2007

Abstract  
The famous Newton–Kantorovich hypothesis has been used for a long time as a sufficient condition for the convergence of Newton’s method to a solution of an equation. Here we present a “Kantorovich type” convergence analysis for the Gauss–Newton’s method which improves the result in [W.M. Häußler, A Kantorovich-type convergence analysis for the Gauss–Newton-method, Numer. Math. 48 (1986) 119–125.] and extends the main theorem in [I.K. Argyros, On the Newton-Kantorovich hypothesis for solving equations, J. Comput. Appl. Math. 169 (2004) 315–332]. Furthermore, the radius of convergence ball is also obtained. © 2007 Published by Elsevier B.V.

Keywords: Gauss–Newton’s method; Majorizing sequence; Semilocal convergence; Local convergence; Lipschitz condition

1. Introduction

Let $F : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a nonlinear operator with its Frechét derivative denoted by $F'$. Finding solutions of a nonlinear operator equation

$$F(x) = 0$$

(1.1)

is a very general subject which is widely studied in both theoretical and applied areas of mathematics. In the case when $m = n$ and $F'(x)$ is invertible for each $x \in D$, the most important method to find an approximation solution is Newton’s method, which, with initial point $x_0 \in D$, is defined by

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k) \quad \text{for each } k = 0, 1, 2, \ldots$$

(1.2)

One of the most famous results on Newton’s method is the well-known Kantorovich theorem (cf. [14]) which provides a simple and clear convergence criterion of Newton’s method based on the data around the initial point for functions having the bounded second derivative $F''$ (or the Lipschitz continuous first derivative). Another important result concerning Newton’s method is Smale’s point estimate theory, which gives a convergence criterion of Newton’s method only based on the information at the initial point for analytic functions (cf. [3,17–19]).

There are a lot of works on the weakness and/or extension of the hypothesis made on the functions, see for example, [1,8–12,22] and references therein. In particular, Wang introduced in [22] the notions of Lipschitz conditions with...
L average to unify both Kantorovich’s and Smale’s convergence criteria. While, Argyros in [1] used simultaneously the center Lipschitz condition (1.3) and the Lipschitz condition (1.4) below to improve Kantorovich’s convergence criterion:

\[ \| F'(x_0)^{-1}(F'(x) - F'(x_0)) \| \leq K_0 \| x - x_0 \| \quad \text{for each } x \in D \]  

and

\[ \| F'(x_0)^{-1}(F'(x) - F'(y)) \| \leq K \| x - y \| \quad \text{for each } x, y \in D. \]  

Other results such as estimates of the radii of convergence balls of Newton’s method are referred to [20,21,23,24].

Recent attentions are focused on finding zeros of singular nonlinear systems by Gauss–Newton’s method (abbrev. GNM), which is defined as follows (cf. [4]):

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

where \( x_0 \in D \) is an initial point and \( F'(x_k)^\dagger \) is the Moore–Penrose inverse of the linear operator (or matrix) \( F'(x_k) \). For example, Shub and Smale in [16] (resp. Dedieu and Shub in [7]) developed the convergence properties of GNM for underdetermined (resp. overdetermined) analytic systems with surjective (resp. injective) derivatives. Dedieu and Kim in [6] studied the convergence properties of GNM for analytic systems of equations with constant rank derivatives. In spirit of Wang’s idea of the Lipschitz conditions with \( L \) average, Li et al. established in [15] an unified convergence theorem for overdetermined systems with injective derivatives; while Xu and Li extended and improved in [25] the corresponding results in [6]. However, almost all the results above are local, that is, the convergence properties are closely dependent on the information around the least square solution of \( F \); and there has been little work on Kantorovich’s type convergence criterion of GNM in terms of the information around the initial point. Häußler considered in [13] a special class of singular nonlinear systems \( F \) together with the derivative \( F' \) satisfying

\[ \| F'(y)^\dagger(I - F'(x)F'(x)^\dagger)F(x) \| \leq \bar{\kappa} \| x - y \| \quad \text{for each } x, y \in D, \]  

where \( 0 \leq \bar{\kappa} < 1 \), and established a Kantorovich’s type convergence criterion under the Lipschitz continuity of the first derivative \( F' \) on \( D \). In the present paper, we will incorporate the center Lipschitz continuity in the study of the convergence of GNM for the class of singular systems satisfying (1.6) and, with a different technique, establish a Kantorovich’s type convergence criterion. In particular, the convergence criterion produces a sharper one than that in [13] under the same hypothesis, which is also illustrated by an example; while, in the underdetermined case with surjective derivatives, it extends the corresponding result in [1, Theorem 1] for nonsingular system. Furthermore, as applications, an estimate of the radius of the convergence ball, which seems new, is presented in Section 4.

We end this introduction with a short remark that, following the technique in [13], Argyros in [2] used the center Lipschitz continuity to give a convergence criterion of GNM for singular system satisfying (1.6). However, our convergence criterion in the present paper is clearer than that in [2]; in particular, it is sharper in the special case when \( K = K_0 \) as shown in Remark 3.1.

2. Preliminaries

Let \( \alpha > 0, p > 0 \) and \( 1 \geq q > 0 \). We begin with the majorizing function \( \varphi_q \) defined by

\[ \varphi_q(t) = \frac{p}{2} t^2 - qt + \alpha \quad \text{for each } t \geq 0. \]  

Clearly, if

\[ \alpha \leq \frac{q^2}{2p}, \]  

then the function \( \varphi_q \) has two zeros:

\[
\begin{align*}
 t^* & = \frac{q + \sqrt{q^2 - 2p\alpha}}{p} \\
 t^{**} & = \frac{q - \sqrt{q^2 - 2p\alpha}}{p}.
\end{align*}
\]
Let \( \{t_k\} \) be the sequence generated by
\[
t_0 = 0, \quad t_{k+1} = t_k - \frac{\varphi_q(t_k)}{\varphi_q'(t_k)} \quad \text{for each } k = 0, 1, \ldots .
\]
(2.4)

In particular, in the case when \( q = 1 \), (2.4) reduces to Newton’s sequence. The convergence property of the sequence \( \{t_k\} \) is described in the following lemma, which is crucial for the convergence analysis of the GNM.

**Lemma 2.1.** The sequence \( \{t_k\} \) is increasingly convergent to \( t^* \) if and only if (2.2) holds. In particular, in the case when \( q = 1 \), the following estimate holds:
\[
t^* - t_k = \frac{\xi^{2^k-1}}{\sum_{j=0}^{2^k-1} \xi^j} t^* \quad \text{for each } k = 0, 1, \ldots .
\]
where
\[
\xi = \frac{1 - \sqrt{1 - 2zp}}{1 + \sqrt{1 - 2zp}}.
\]
(2.6)

**Proof.** We first prove that for each \( k \in \mathbb{N} \),
\[
t_{k-1} < t_k < t^*.
\]
(2.7)

Granting this, one sees that \( \{t_k\} \) is increasing and bounded, and consequently \( \{t_k\} \) is increasingly convergent to \( t^* \).

To show (2.7), note that \( 0 = t_0 < t_1 = x < t^* \), which means (2.7) holds for \( k = 1 \). Assume that \( t_0 < t_1 < \cdots < t_k < t^* \). Then one has \( \varphi_q(t_k) > 0 \) and
\[
-\varphi_q'(t_k) = 1 - pt_k > 1 - pt^* = (1 - q) + \sqrt{q^2 - 2zp} > 0.
\]
It follows that
\[
t_{k+1} = t_k - \frac{\varphi_q(t_k)}{\varphi_q'(t_k)} > t_k.
\]
(2.8)

Note that the function \( N_q \) defined by \( N_q(t) := t - \varphi_q(t)/\varphi_q'(t) \) for each \( t \in [0, t^*] \) has positive derivative on \([0, t^*] \) (Note: \( \varphi_q'(t^*) < 0 \), unless \( q = 1 \) and \( q^2 - 2zp = 0 \), in this case \( t^* = 1/p \), and, by L’Hospital’s rule, \( \varphi_q(t^*)/\varphi_q'(t^*) = 0 \).

One has that
\[
t_{k+1} = N_q(t_k) < N_q(t^*) = t^*.
\]
(2.9)

This together with (2.8) implies that (2.7) holds for \( k + 1 \) and the claim (2.7) is complete by mathematical induction. On the other hand, it is clear that the sequence \( \{t_k\} \) converging implies (2.1) having solution, and consequently (2.2) holds. Thus the proof of the first assertion is complete. The second assertion is well known, see for example [22]. \( \square \)

We conclude this section with some properties related to Moore-Penrose inverse, which are known in textbooks, see for example [5].

Let \( A : \mathbb{R}^m \rightarrow \mathbb{R}^n \) be a linear operator (or an \( m \times n \) matrix). Recall that an operator (or an \( n \times m \) matrix) \( A^\dagger : \mathbb{R}^m \rightarrow \mathbb{R}^n \) 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 \( \im A \) denote the kernel and image of \( A \), respectively. For a subspace \( E \) of \( \mathbb{R}^n \), we use \( \Pi_E \) to denote the projection onto \( E \). Then it is clear that
\[
A^\dagger A = \Pi_{(\ker A)^\perp} \quad \text{and} \quad AA^\dagger = \Pi_{\im A}.
\]
(2.10)

In particular, in the case when \( A \) is full row rank, \( AA^\dagger = I_{\mathbb{R}^m} \).
The following proposition gives a perturbation bound for Moore–Penrose inverse, which will be useful.

**Proposition 2.1.** Let $A$ and $B$ be $m \times n$ matrices. Assume
\[
\text{rank}(A) \leq \text{rank}(B) = l \geq 1 \quad \text{and} \quad \|A - B\| B^\dagger < 1.
\]
Then
\[
\text{rank}(A) = l \quad \text{and} \quad \|A^\dagger\| \leq \frac{\|B^\dagger\|}{1 - \|B^\dagger\| A - B}.
\]

3. Semilocal convergence analysis of the GNM

Let $B(x, r)$ and $\overline{B}(x, r)$ stand, respectively, for the open and closed ball in $\mathbb{R}^n$ with center $x$ and radius $r > 0$. Let $F : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a Frechét differentiable operator, where $D$ is a convex set. Let $x_0 \in D$ be such that $F'(x_0) \neq 0$, or equivalently, rank$(F'(x_0)) \geq 1$. Let $\tilde{r} > 0$ be such that $B(x_0, \tilde{r}) \subseteq D$. Throughout the whole section, we will always assume that rank$(F'(x_0)) \leq \text{rank}(F'(x))$ for each $x \in B(x_0, \tilde{r})$,
\[
\|F'(x) - F'(y)\| \leq K\|x - y\| \quad \text{for each } x, y \in B(x_0, \tilde{r})
\]
and
\[
\|F'(x) - F'(x_0)\| \leq K_0\|x - x_0\| \quad \text{for each } x \in B(x_0, \tilde{r}).
\]
Clearly, (3.1) implies that (3.2) holds for some $0 \leq K_0 \leq K$. Furthermore, we will also assume that
\[
\|F'(y)^\dagger (I - F'(x)F'(x)^\dagger)F(x)\| \leq \bar{k}\|x - y\| \quad \text{for each } x, y \in B(x_0, \tilde{r})
\]
with $0 \leq \bar{k} < 1$. For convenience, we write
\[
\Delta := \frac{(1 - \bar{k})^2}{(\bar{k}^2 - \bar{k} + 1) + \sqrt{2\bar{k}^2 - 2\bar{k} + 1}}.
\]
Before verifying the main theorem, we need a simple lemma. For this purpose, let
\[
\alpha_F := \|F'(x_0)^\dagger F(x_0)\| \quad \text{and} \quad \beta_F := \|F'(x_0)^\dagger\|.
\]
**Lemma 3.1.** Suppose that $0 < r \leq \min(\tilde{r}, 1/(\beta_F K_0))$. Then, for each $x \in B(x_0, r)$, rank$(F'(x)) = \text{rank}(F'(x_0))$ and
\[
\|F'(x)^\dagger\| \leq \frac{\beta_F}{1 - \beta_F K_0\|x - x_0\|}.
\]
**Proof.** Let $x \in B(x_0, r)$. Then $\beta_F K_0\|x - x_0\| < \beta_F K_0 r \leq 1$. Hence, by (3.2), one has that
\[
\|F'(x_0)^\dagger\|\|F'(x) - F'(x_0)\| \leq \beta_F K_0\|x - x_0\| < 1.
\]
Thus Proposition 2.1 is applicable to complete the proof.

Set
\[
p = \frac{\beta_F K}{1 + (K - K_0)\alpha_F \beta_F} \quad \text{and} \quad q = 1 - \left(\frac{1 - \alpha_F \beta_F K_0}{1 + (K - K_0)\alpha_F \beta_F}\right)\bar{k}.
\]
Let $t^*$ be defined by (2.3) and $\{t_k\}$ the sequence generated by (2.4) with $\alpha = \alpha_F$. Then the main theorem of the present paper can be stated as follows.

**Theorem 3.1.** Suppose that
\[
\alpha_F \beta_F \leq \frac{\Delta}{K - \Delta(K - K_0)} \quad \text{and} \quad t^* \leq \tilde{r}.
\]
Let \( \{x_k\} \) be the sequence generated by GNM (1.5) with initial point \( x_0 \). Then \( \{x_k\} \) converges to a zero \( x^* \) of \( F'(\cdot)^\top F(\cdot) \) in \( B(x_0, t^*) \) and the following estimate holds:
\[
\|x_k - x^*\| \leq t^* - t_k \quad \text{for each } k \geq 0.
\] (3.9)

**Proof.** Recall that \( p \) and \( q \) are given by (3.7). Simple calculation shows that the first inequality of (3.8) implies
\[
\alpha_F \leq \frac{q^2}{2p}.
\] (3.10)

Thus, by Lemma 2.1, \( \{t_k\} \) is strictly increasingly convergent to \( t^* \) and
\[
t^* \leq \frac{1 + (K - K_0)\alpha_F}{\beta_F K}.
\] (3.11)

Since \( \alpha_F = t_1 \leq t^* \), it follows from (3.11) that \( \beta_F K t^* \leq 1 + (K - K_0)\beta_F t^* \) and hence
\[
\beta_F K_0 t^* \leq 1.
\] (3.12)

Define
\[
G(x) := x - F'(x)^\top F(x) \quad \text{for each } x \in D.
\]

Let \( x \in B(x_0, t^*) \) be such that \( G(x) \in B(x_0, t^*) \). Then
\[
\|G^2(x) - G(x)\| \leq \frac{\beta_F K}{2(1 - \beta_F K_0\|G(x) - x_0\|)} \|G(x) - x\|^2 + \bar{\kappa}\|G(x) - x\|.
\] (3.13)

To see this, by (3.12), Lemma 3.1 is applicable to getting that
\[
\|F'(G(x))^\top\| \leq \frac{\beta_F}{1 - \beta_F K_0\|G(x) - x_0\|}.
\] (3.14)

Hence
\[
\|G^2(x) - G(x)\| \leq \left\| F'(G(x))^\top \int_0^1 \{F'(x + t(G(x) - x)) - F'(x)\} (G(x) - x) \, dt \right\| \\
+ \|F'(G(x))^\top (I - F'(x)F'(x)^\top) F(x)\|
\leq \|F'(G(x))^\top \int_0^1 \|F'(x + t(G(x) - x)) - F'(x)\| \|G(x) - x\| \, dt \\
+ \|F'(G(x))^\top (I - F'(x)F'(x)^\top) F(x)\|
\leq \frac{\beta_F K}{2(1 - \beta_F K_0\|G(x) - x_0\|)} \|G(x) - x\|^2 + \bar{\kappa}\|G(x) - x\|,
\]
where the last inequality holds because of (3.1), (3.3) and (3.14).

Below we shall verify that
\[
\|x_k - x_{k-1}\| \leq t_k - t_{k-1}
\] (3.15)
holds for each \( k = 1, 2, \ldots \) by mathematical induction.

It is clear that \( \|x_1 - x_0\| \leq \alpha_F = t_1 - t_0 \) which means (3.15) holds for \( k = 1 \). Assume that (3.15) holds for all \( k \leq j \). It follows that
\[
\|x_k - x_0\| \leq \sum_{i=1}^k \|x_i - x_{i-1}\| \leq t_k < t^* < \bar{\rho} \quad \text{for each } k = 1, 2, \ldots, j
\] (3.16)
thanks to (3.8). In particular, \( x_{j-1}, x_j \in B(x_0, t^* \leq B(x_0, \tilde{r}) \). Noting that \( x_k = G(x_{k-1}) \) for each \( k = 1, 2, \ldots \), we get from (3.13) that

\[
\|x_{j+1} - x_j\| \leq \frac{\beta_F K}{2(1 - \beta_F K_0 \|x_j - x_0\|)} \|x_j - x_{j-1}\|^2 + \tilde{k}\|x_j - x_{j-1}\|.
\]  

(3.17)

Consequently,

\[
\|x_{j+1} - x_j\| \leq \frac{\beta_F K(t_j - t_{j-1})^2}{2(1 - \beta_F K_0 t_j)} + \tilde{k}(t_j - t_{j-1}).
\]  

(3.18)

Since \( \alpha_F = t_1 \leq t_j \), we have

\[
1 - \frac{\beta_F K t_j}{1 + (K - K_0)\alpha_F \beta_F} \leq 1 - \frac{\alpha_F \beta_F K}{1 + (K - K_0)\alpha_F \beta_F}
\]  

(3.19)

and

\[
\frac{1}{1 - \beta_F K t_j / (1 + (K - K_0)\alpha_F \beta_F)} \leq \frac{1}{1 - \beta_F K_0 t_j + (K - K_0)\beta_F (\alpha_F - t_j)} \geq \frac{1}{1 - \beta_F K_0 t_j}.
\]  

(3.20)

Recalling definitions of \( p \) and \( q \) in (3.7), it follows from (3.18) to (3.20) that

\[
\|x_{j+1} - x_j\|
\leq \frac{1}{1 - \beta_F K t_j / (1 + (K - K_0)\alpha_F \beta_F)} \left( \frac{\beta_F K(t_j - t_{j-1})^2}{2(1 + (K - K_0)\alpha_F \beta_F)} 
+ \left( 1 - \frac{\alpha_F \beta_F K}{1 + (K - K_0)\alpha_F \beta_F} \right) \tilde{k}(t_j - t_{j-1}) \right)
\leq \frac{1}{1 - pt_j} \left( \frac{p}{2}(t_j - t_{j-1})^2 + (1 - q)(t_j - t_{j-1}) \right)
= \frac{1}{\varphi_1(t_j)} (\varphi_q(t_j) - \varphi_q(t_{j-1}) - \varphi'(t_{j-1})(t_j - t_{j-1}))
= t_{j+1} - t_j.
\]  

(3.21)

This means that (3.15) holds for \( k = j + 1 \) and so for each \( k = 1, 2, \ldots \). Consequently, Lemma 2.1 is applicable to concluding that \( \{x_k\} \) converges to some point \( x^* \in B(x_0, t^* \)\). Since

\[
\|F'(x^*)^\dagger F(x_k)\| \leq \|F'(x^*)^\dagger (I - F'(x_k)F'(x_k)^\dagger)F(x_k)\|
+ \|F'(x^*)^\dagger \| \cdot \|F'(x_k)F'(x_k)^\dagger F(x_k)\|
\leq \tilde{k}\|x_k - x^*\| + \|F'(x^*)^\dagger\| \|F'(x_k)\| \|x_{k+1} - x_k\|.
\]  

(3.22)

one sees that \( x^* \) is a zero of \( F'(\cdot)^\dagger F(\cdot) \) and the proof is complete. \( \square \)

**Remark 3.1.** In [2, Theorem 2], Argyros gave the following convergence criterion for GNM (1.5): there exists \( \delta \in [\tilde{k}, 1) \) such that for all \( n \geq 0 \),

\[
\left( \frac{1}{2} (1 - \delta) \delta^n K + \delta (1 - \delta^{n+1}) K_0 \right) \alpha_F \beta_F + (\tilde{k} - \delta)(1 - \delta) \leq 0,
\]  

(3.23)

\[
\frac{\alpha_F \beta_F K_0}{1 - \delta} (1 - \delta^n) < 1 \quad \text{and} \quad s^* \leq \tilde{r},
\]  

(3.24)
where \( s^* \) is the limit of the majorizing sequence \( \{s_k\} \) defined by
\[
    s_0 = 0, \quad s_1 = x_F, \quad s_{k+1} = s_k + \frac{1}{1 - \beta_F K s_k} \left( \frac{1}{2} \beta_F K (s_k - s_{k-1})^2 + \bar{k}(s_k - s_{k-1}) \right).
\]

Below we shall show that this convergence criterion is stronger than (3.8) in the case when \( K = K_0 \). In fact, in this case, sequence \( \{s_k\} \) reduces to
\[
    s_{k+1} = s_k + \frac{1}{1 - \beta_F K s_k} \left( \frac{1}{2} \beta_F K (s_k - s_{k-1})^2 + \bar{k}(s_k - s_{k-1}) \right),
\]
where \( s_0 = 0 \) and \( s_1 = x_F \). Note that the sequence \( \{t_k\} \) generated by (2.4) can be rewritten as (thanks to (3.21))
\[
    t_{k+1} = t_k + \frac{1}{1 - \beta_F K t_k} \left( \frac{1}{2} \beta_F K (t_k - t_{k-1})^2 + (1 - x_F \beta_F K) \bar{k}(t_k - t_{k-1}) \right),
\]
where \( t_0 = 0 \) and \( t_1 = x_F \). Hence
\[
    t^* \leq s^* \quad \text{and} \quad t_k \leq s_k \quad \text{for all} \quad k \geq 0.
\]
This implies that \( \{t_k\} \) is convergent and hence (3.8) holds thanks to Lemma 2.1.

In the special case when \( \bar{k} = 0 \), \( \Delta = \frac{1}{2} \) and \( q = 1 \). Therefore the following corollary is a direct consequence of Theorem 3.1 together with Lemma 2.1.

**Corollary 3.1.** Suppose that
\[
    x_F \beta_F (K + K_0) \leq 1, \quad t_1^* \leq \bar{r},
\]
and that
\[
    \|F'(y)^\dagger (I - F'(x)^\dagger F'(x)^\dagger) F(x)\| = 0 \quad \text{for each} \quad x, y \in B(x_0, \bar{r}).
\]
Let \( \{x_k\} \) be the sequence generated by GNM (1.5) with initial point \( x_0 \). Then \( \{x_k\} \) converges to a zero \( x^* \) of \( F'(\cdot)^\dagger F(\cdot) \) in \( B(x_0, t_1^*) \) and the following estimate holds:
\[
    \|x_k - x^*\| \leq \frac{\xi_1 x_k - 1}{\sum_{j=0}^{k-1} \xi_j} t_1^* \quad \text{for each} \quad k = 0, 1, \ldots,
\]
where \( t_1^* \) and \( \xi_1 \) are, respectively, defined by
\[
    t_1^* = 1 + (K - K_0) x_F \beta_F - \sqrt{(1 - (K + K_0) x_F \beta_F)(1 + (K - K_0) x_F \beta_F)} \quad \text{and}
\]
\[
    \xi_1 = \frac{1 - K_0 x_F \beta_F - \sqrt{(1 - (K + K_0) x_F \beta_F)(1 + (K - K_0) x_F \beta_F)}}{x_F \beta_F K}.
\]

In the special case when \( F'(x_0) \) is invertible, Argyros used in [1] the following Lipschitz conditions to analyze the convergence of Newton’s method.
\[
    \|F'(x_0)^\dagger (F'(x) - F'(y))\| \leq K \|x - y\| \quad \text{for each} \quad x, y \in B(x_0, \bar{r})
\]
and
\[
    \|F'(x_0)^\dagger (F'(x) - F'(x_0))\| \leq K_0 \|x - x_0\| \quad \text{for each} \quad x \in B(x_0, \bar{r}).
\]
It was proved in [1, Theorem 3.1] that if

there exists \( \delta \in [0, 1] \) such that \((K + \delta K_0) x_F \leq \delta\) and \( s^{**} \leq \tilde{r}\), \hspace{1cm} (3.33)

where \( s^{**} = 2x_F / (2 - \delta) \), then Newton’s method with initial point \( x_0 \) is convergent. Let \( \delta \in [0, 1] \) such that \((K + \delta K_0) x_F \leq \delta\). Then

\[
K_0 x_F \leq 1 \quad \text{and} \quad \frac{K x_F}{1 - x_F K_0} \leq \delta.
\]

The first inequality in (3.34) implies that

\[
x_F (K + K_0) = (K + \delta K_0) x_F + (1 - \delta) K_0 x_F \leq 1.
\]

Note that

\[
1 - (K + K_0) x_F \leq \frac{1}{1 - (K - K_0) x_F}.
\]

This together with the second inequality in (3.34) implies that

\[
(1 - \delta)^2 \leq \left( \frac{1 - K x_F}{1 - K_0 x_F} \right)^2 = \left( \frac{1 - (K + K_0) x_F}{1 - K_0 x_F} \right)^2 \leq \frac{1 - (K + K_0) x_F}{1 + (K - K_0) x_F}.
\]

On the other hand,

\[
\hat{r}^* = \frac{1 + (K - K_0) x_F - \sqrt{(1 - (K + K_0) x_F)(1 + (K - K_0) x_F)}}{K} = \frac{2x_F (1 + (K - K_0) x_F)}{(1 + (K - K_0) x_F) + \sqrt{(1 - (K + K_0) x_F)(1 + (K - K_0) x_F)}} = \frac{2x_F}{1 + \sqrt{(1 - (K + K_0) x_F)/(1 + (K - K_0) x_F)}}.
\]

Combining this with (3.36) gives that \( \hat{r}^* \leq s^{**} \). Therefore, (3.33) implies (3.37) below thanks to (3.35). Thus Corollary 3.2 below is an extension and improvement of [1, Theorem 1], in particular, a closed form of the estimate for \( \|x_k - x^*\| \) is presented in this corollary.

**Corollary 3.2.** Suppose that the Lipschitz conditions (3.31) and (3.32) hold. Let \( x_0 \in D \) be such that \( F'(x_0) \) is full row rank. Suppose that

\[
x_F (K + K_0) \leq 1 \quad \text{and} \quad \hat{r}^* \leq \tilde{r},
\]

where

\[
\hat{r}^* = \frac{1 + (K - K_0) x_F - \sqrt{(1 - (K + K_0) x_F)(1 + (K - K_0) x_F)}}{K},
\]

\[
\xi_1 = \frac{1 - K_0 x_F - \sqrt{(1 - (K + K_0) x_F)(1 + (K - K_0) x_F)}}{K x_F}.
\]

**Proof.** Define \( \tilde{F} = F'(x_0) \hat{F} \). We shall apply Corollary 3.1 to \( \tilde{F} \). For this end, take \( \tilde{r} = \hat{r}^* \) in Corollary 3.1. Then (3.31) and (3.32) imply that (3.1) and (3.2) are satisfied by \( \tilde{F} \). We claim that \( F'(x) \) is full row rank for each \( x \in B(x_0, \tilde{r}) \). In fact, since

\[
\tilde{r} = \hat{r}^* = \frac{1 + (K - K_0) x_F - \sqrt{(1 - (K + K_0) x_F)(1 + (K - K_0) x_F)}}{K} \leq \frac{1 + (K - K_0) x_F}{K}
\]

(3.40)
and \( \alpha_F = t_1 \leq \bar{r} \), it follows that
\[
K\bar{r} \leq 1 + (K - K_0) \alpha_F \leq 1 + (K - K_0)\bar{r},
\]
and consequently \( K_0\bar{r} \leq 1 \). Therefore, together with (3.32) it follows that, for each \( x \in \mathbf{B}(x_0, \bar{r}) \),
\[
\|F'(x_0) (F(x) - F'(x_0))\| \leq K_0 \|x - x_0\| < K_0\bar{r} \leq 1.
\]
By Banach Lemma, \( (I_{\mathbb{R}^n} - F'(x_0) (F'(x) - F'(x_0)))^{-1} \) exists. Noting that \( F'(x_0) \) is full row rank, we have that \( F'(x_0) F'(x_0)^\dagger = I_{\mathbb{R}^n} \) and
\[
F'(x) = F'(x_0)(I_{\mathbb{R}^n} - F'(x_0) (F'(x) - F'(x_0))).
\]
This implies that \( F'(x) \) is full row rank because \( I_{\mathbb{R}^n} - F'(x_0) (F'(x) - F'(x_0)) \) is invertible; hence the claim stands. Thus, in view of the definition of the Moore-Penrose inverse, one sees that
\[
(F'(x))^\dagger = (F'(x_0))^\dagger F'(x) = F'(x)^\dagger F'(x_0)
\]
for each \( x \in \mathbf{B}(x_0, \bar{r}) \). (3.42)
This implies that (3.27) is satisfied by \( \widetilde{F} \) and that \( \{x_k\} \) coincides with the sequence generated by GNM (1.5) with initial point \( x_0 \) for \( \widetilde{F} \). Furthermore, since by (3.42)
\[
(F'(x_0))^\dagger = (F'(x_0)^\dagger F'(x_0))^\dagger = F'(x_0)^\dagger F'(x_0),
\]
(3.43)


it follows that
\[
\alpha_F = \|(F'(x_0)^\dagger F'(x_0))^\dagger F'(x_0)\| = \|F'(x_0)^\dagger F'(x_0)\| = \alpha_F
\]
and
\[
\beta_F = \|(F'(x_0)^\dagger F'(x_0))\| = \|I_{(\ker F'(x_0))^\perp}\| = 1.
\]
(3.45)
Hence (3.26) is satisfied thanks to (3.37). Therefore, Corollary 3.1 is applicable to \( \widetilde{F} \) and \( \{x_k\} \) converges to a zero \( x^* \) of \( F'(\cdot)^\dagger \widetilde{F}(\cdot) \). Noting that \( F'(\cdot)^\dagger \widetilde{F}(\cdot) = F'(\cdot)^\dagger F(\cdot) \) and \( F(\cdot) = F'(\cdot)(F'(\cdot)^\dagger F(\cdot)) \), it follows that \( x^* \) is a zero of \( F(\cdot) \). The proof is complete. \( \square \)

In [13], Häußler took \( \mathbf{K} = K_0 \) and proved that if
\[
\alpha_F \beta_F K \leq \frac{(1 - \bar{k})^2}{2} \quad \text{and} \quad s^* \leq \bar{r},
\]
(3.46)
where \( s^* = ((1 - \bar{k}) - \sqrt{(1 - \bar{k})^2 - 2s_F \beta_F K})/\beta_F K \), then GNM (1.5) with initial point \( x_0 \) converges to a zero \( x^* \) of \( F'(\cdot)^\dagger F(\cdot) \) in \( \mathbf{B}(x_0, s^*) \). Set
\[
\bar{r}^* = \frac{1 - (1 - \alpha_F \beta_F K)\bar{k} - \sqrt{(1 - (1 - \alpha_F \beta_F K)\bar{k})^2 - 2s_F \beta_F K}}{\beta_F K}.
\]
(3.47)
Clearly, \( (1 - \bar{k})^2/2 \leq \bar{r}^* \). Note that the function \( t \mapsto 1 - t - \sqrt{(1 - t)^2 - a} \) with \( a = 2s_F \beta_F K \) is increasing on \([0, \bar{k}]\). It is seen that \( \bar{r}^* \leq s^* \). Therefore the following corollary improves [13, Theorem 2.4].

**Corollary 3.3.** Let \( x_0 \in D \) be such that \( F'(x_0) \neq 0 \). Suppose that \( \text{rank } (F'(x)) \leq \text{rank } (F'(x_0)) \) for each \( x \in D \) and that (3.1) holds. Let \( \alpha_F \) and \( \beta_F \) be defined by (3.5). If
\[
\alpha_F \beta_F K \leq \bar{r}^* \quad \text{and} \quad \bar{r}^* \leq \bar{r},
\]
(3.48)
then GNM (1.5) with initial point \( x_0 \) converges to a zero \( x^* \) of \( F'(\cdot)^\dagger F(\cdot) \) in \( \mathbf{B}(x_0, \bar{r}^*) \) and
\[
\|x_k - x^*\| \leq \bar{r}^* - t_k \quad \text{for each } k \geq 0.
\]
(3.49)

We now give an example for which Corollary 3.3 is applicable but neither [13, Theorem 2.4] nor [2, Theorem 2].
Example 3.1. Let $n = m = 2$ and let $\mathbb{R}^2$ be endowed with the $l_1$-norm. Let $D = \{ x = (\xi_1, \xi_2)^T : -1 < \xi_i < 1, \ i = 1, 2 \} \subseteq \mathbb{R}^2$, $x_0 = (\xi_0^1, \xi_0^2)^T = (\frac{1}{4}, 0)^T$, and $\bar{r} = \frac{18}{25}$. Define $F : D \to \mathbb{R}^2$ by

$$F(x) := \left( \frac{1}{2}(\xi_1 - \xi_2), 1 \right)^T$$

for each $x = (\xi_1, \xi_2)^T \in D$.

Then, for each $x = (\xi_1, \xi_2)^T \in D$,

$$F'(x) = \begin{pmatrix} 1 & -1 \\ \xi_1 - \xi_2 & -(\xi_1 - \xi_2) \end{pmatrix}$$

and

$$F'(x)^\dagger = \frac{1}{2(1 + (\xi_1 - \xi_2)^2)} \begin{pmatrix} 1 & \xi_1 - \xi_2 \\ -1 & -(\xi_1 - \xi_2) \end{pmatrix}.$$ 

Hence, for $x = (\xi_1, \xi_2)^T, y = (\zeta_1, \zeta_2)^T \in D$,

$$\|F'(x) - F'(y)\| = |(\xi_1 - \zeta_1) - (\xi_2 - \zeta_2)| \leq \|x - y\|.$$

This means that $K = K_0 = 1$. Since, for $x = (\xi_1, \xi_2)^T, y = (\zeta_1, \zeta_2)^T \in D$,

$$\|F'(y)^\dagger (I - F'(x)F'(x)^\dagger)F(x)\|$$

$$= \frac{1}{2(1 + (\xi_1 - \xi_2)^2)} \frac{2(\xi_1 - \xi_2)^2}{2(1 + (\xi_1 - \xi_2)^2)} \| (\xi_1 - \zeta_1) - (\xi_2 - \zeta_2) \|$$

$$= \frac{1}{2(1 + (\xi_1 - \xi_2)^2)} \frac{2(\xi_1 - \zeta_2)^2}{2(1 + (\xi_1 - \xi_2)^2)} \| (\xi_1 - \zeta_1) - (\xi_2 - \zeta_2) \|,$$

it follows that

$$\|F'(y)^\dagger (I - F'(x)F'(x)^\dagger)F(x)\| \leq \frac{2(\xi_1 - \xi_2)^2}{2(1 + (\xi_1 - \xi_2)^2)} \| x - y \| \leq \frac{2}{5} \| x - y \|$$

because

$$\frac{(\xi_1 - \xi_2)^2}{2(1 + (\xi_1 - \xi_2)^2)} = \frac{1}{2} \left( 1 - \frac{1}{1 + (\xi_1 - \xi_2)^2} \right) \leq \frac{2}{5},$$

hence $\bar{k} = \frac{2}{5}$. Moreover,

$$\alpha_F = \|F'(x_0)^\dagger F(x_0)\| = \left\| \frac{8}{17} \begin{pmatrix} 1 & 4/4 \\ -1 & -4/4 \end{pmatrix} \right\| = \frac{33}{136} \quad (3.50)$$

and

$$\beta_F = \|F'(x_0)^\dagger\| = \left\| \frac{8}{17} \begin{pmatrix} 1 & 4/4 \\ -1 & -4/4 \end{pmatrix} \right\| = \frac{16}{17}. \quad (3.51)$$

Since

$$\alpha_F \beta_F K \leq \frac{66}{17^2} \geq \frac{9}{50} = \frac{(1 - \bar{k})^2}{2},$$
Theorem 2.4 in [13] is not applicable. On the other hand, Theorem 2 in [2] is not applicable too. In fact, since
\[
\frac{16}{17} \cdot \frac{33}{136} \delta + \left( \frac{2}{5} - \delta \right) (1 - \delta) \leq 0
\]
has no solutions, it follows that there does not exist \( \delta \in [0, 1) \) with \( \bar{\kappa} \leq \delta \) such that (3.23) satisfying for all \( n \geq 0 \). However, since
\[
\alpha_F \beta_F K = \frac{66}{172} \leq \frac{9}{19 + 5\sqrt{13}} = \frac{(1 - \bar{\kappa})^2}{(\bar{\kappa}^2 - \bar{\kappa} + 1) + \sqrt{2\bar{\kappa}^2 - 2\bar{\kappa} + 1}}
\]
and
\[
\tilde{r}^* = \frac{999 - \sqrt{44301}}{1530} \leq \frac{18}{25} = \bar{r},
\]
Corollary 3.3 is applicable.

We end this section with an example for which condition (3.27) in Corollary 3.1 is satisfied but \( F'(x) \) is not full row rank.

**Example 3.2.** Let \( F : \mathbb{R}^2 \to \mathbb{R}^2 \) be defined by
\[
F(x) := \left( \frac{1}{2}(\xi_1 + \xi_2)^2, \frac{1}{2}(\xi_1 + \xi_2)^2 - 1 \right)^T.
\]
Then
\[
F'(x) = (\xi_1 + \xi_2) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}
\]
and
\[
F'(x)^\dagger = \frac{1}{4(\xi_1 + \xi_2)} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.
\]
Let \( \mathbb{R}^2 \) be endowed with the \( l_1 \)-norm. Therefore, for \( x = (\xi_1, \xi_2)^T, y = (\bar{\xi}_1, \bar{\xi}_2)^T \),
\[
\| F'(y)^\dagger (I - F'(x) F'(x)^\dagger) F(x) \| = \left\| \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\| = 0.
\]
Thus, for \( \bar{\kappa} = 0 \), we have
\[
\| F'(y)^\dagger (I - F'(x) F'(x)^\dagger) F(x) \| \leq \bar{\kappa} \| x - y \| \quad \text{for all } x, y \in \mathbb{R}^2.
\]

4. **Local convergence analysis of the GNM**

In this section, let \( x^* \in D \) be such that \( F(x^*) = 0 \) and \( F'(x^*) \neq 0 \). We shall assume that \( \text{rank}(F'(x)) \leq \text{rank}(F'(x^*)) \) for each \( x \in D \),
\[
\| F'(x) - F'(y) \| \leq K \| x - y \| \quad \text{for each } x, y \in D \tag{4.1}
\]
and
\[
\| F'(y)^\dagger (I - F'(x) F'(x)^\dagger) F(x) \| \leq \bar{\kappa} \| x - y \| \quad \text{for each } x, y \in D \tag{4.2}
\]
with \( 0 \leq \bar{\kappa} < 1 \). Let \( \beta^* = \| F'(x^*)^\dagger \| \) and recall that \( \Delta \) is defined by (3.4). Then the local convergence result for GNM (1.5) is stated in the following theorem.
Theorem 4.1. Let
\[ r = \frac{1 - 1/\sqrt{2A + 1}}{\beta^* K}. \]

Suppose that \( \mathcal{B}(x^*, 1/(\beta^* K)) \subseteq D. \) \hfill (4.4)

Then, for each \( x_0 \in \mathcal{B}(x^*, r) \), the sequence \( \{x_k\} \) generated by GNM (1.5) with initial point \( x_0 \) converges to a zero of \( F'(\cdot)F(\cdot) \).

Proof. Let \( x_0 \in \mathcal{B}(x^*, r) \). Then Lemma 3.1 implies that \( \text{rank}(F'(x_0)) = \text{rank}(F'(x^*)) \) and
\[ \beta_F = \|F'(x_0)\|^2 \leq \frac{\beta^*}{1 - \beta^* K \|x^* - x_0\|^2}. \]

Hence \( \text{rank}(F'(x)) \leq \text{rank}(F'(x_0)) \) for each \( x \in D \). Let \( \bar{r} = 1/(\beta^* K) - \|x_0 - x^*\| \). Then, \( \mathcal{B}(x_0, \bar{r}) \subseteq D \) thanks to (4.4). By Corollary 3.3, it suffices to show that (3.48) holds. Note that
\[ -F'(x_0)^\top F(x_0) = F'(x_0)^\top (F(x^*) - F(x_0)) = F'(x_0)(x^* - x_0) \]
\[ = F'(x_0) \int_0^1 (F'(x_0 + \theta(x^* - x_0)) - F'(x_0))(x^* - x_0) \, d\theta + \Pi_{(\ker F'(x_0))} (x^* - x_0). \]

It follows from (4.1) and (4.5) that
\[ \alpha_F = \|F'(x_0)\|^2 \]
\[ \leq \frac{\beta^*}{1 - \beta^* K \|x^* - x_0\|^2} \frac{1}{2} K \|x^* - x_0\|^2 + \|x^* - x_0\| \]
\[ = \frac{2 - \beta^* K \|x^* - x_0\|}{2(1 - \beta^* K \|x^* - x_0\|)^2} \|x^* - x_0\|. \]

Combining this with (4.5) gives that
\[ \alpha_F \beta_F K \leq \frac{2 - \beta^* K \|x^* - x_0\|}{2(1 - \beta^* K \|x^* - x_0\|)^2} \beta^* K \|x^* - x_0\| \leq \Delta, \]
where the inequality holds because \( \beta^* K \|x^* - x_0\| \leq 1 - 1/\sqrt{2A + 1} \) and the function \( t \mapsto ((2 - t)/(2(1 - t)^2))t \) is increasing on \((0, 1)\). Hence the first inequality in (3.48) holds. On the other hand,
\[ \bar{r}^* = \frac{1 - (1 - \alpha_F \beta_F K) \bar{r}}{\beta_F K} \]
\[ = \frac{1 - (1 - \alpha_F \beta_F K) \bar{r} + \sqrt{(1 - (1 - \alpha_F \beta_F K) \bar{r})^2 - 2 \alpha_F \beta_F K \bar{r}}}{2 \alpha_F} \]
\[ \leq \frac{2 \alpha_F}{1 - (1 - \alpha_F \beta_F K) \bar{r} + \sqrt{(1 - (1 - \alpha_F \beta_F K) \bar{r})^2 - 2 \alpha_F \beta_F K / (1 - \beta^* K \|x_0 - x^*\|)}} \]
\[ = \frac{1 - (1 - \alpha_F \beta_F K) \bar{r} - \sqrt{(1 - (1 - \alpha_F \beta_F K) \bar{r})^2 - 2 \alpha_F \beta^* K / (1 - \beta^* K \|x_0 - x^*\|)}}{\beta^* K / (1 - \beta^* K \|x_0 - x^*\|)} \]
\[ \leq \frac{1}{\beta^* K} - \|x^* - x_0\|. \]
where the first inequality holds because of (4.5). Therefore \( \tilde{t}^* \leq \bar{r} \), which together with (4.7) completes the proof of (3.48). The proof is complete. \( \square \)

In the case when \( F'(x^*) \) is full row rank, we can take \( \bar{r} = 0 \), and hence, \( \Delta = \frac{1}{2} \). Then, using a similar proof of Theorem 4.1, Corollary 3.2 yields the following result.

Corollary 4.1. Let \( x^* \in D \) be such that \( F(x^*) = 0 \) and \( F'(x^*) \) is full row rank. Suppose that
\[
\|F'(x^*)^{-1}(F'(x) - F'(y))\| \leq K\|x - y\| \quad \text{for all } x, y \in D
\]
and \( B(x^*, 1/K) \subseteq D \). Let \( r = (1 - 1/\sqrt{2})/K \) and let \( x_0 \in B(x^*, r) \). Then GNM (1.5) with initial point \( x_0 \) converges to a zero of \( F(x) = 0 \).

Acknowledgment

This project was supported in part by the National Natural Science Foundation of China (Grant 10671175) and Program for New Century Excellent Talents in University.

References