



Weaker Kantorovich type criteria for inexact Newton methods



I.K. Argyros^a, S.K. Khattri^{b,*}

^a Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA

^b Department of Engineering, Stord Haugesund University College, Norway

ARTICLE INFO

Article history:

Received 19 November 2012

Received in revised form 12 September 2013

MSC:

65H10

65J20

65G99

65B05

65N30

Keywords:

Inexact Newton method

Banach space

Kantorovich-theory

Semilocal convergence

Fréchet derivative

Center-Lipschitz condition

ABSTRACT

We develop a tighter semilocal convergence analysis for the Inexact Newton Method (INM) than in earlier studies such as Shen and Li (2009, 2010), Guo (2007), Smale (1986), Morini (1999), Argyros (1999, 1999, 2007, 2011), Argyros and Hilout (2010, 2012) and Argyros et al. (2012). Our approach is based on the center-Lipschitz condition instead of the Lipschitz condition for computing the inverses of the linear operators involved. Moreover, we expand the applicability of the method by providing weaker sufficient convergence criteria under the same computational cost. Numerical examples where the old convergence criteria are not satisfied but the new convergence criteria hold are also provided in this study. In particular we solve a two-point boundary value problem appearing in magnetohydrodynamics.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In this study, we are concerned with the problem of approximating a locally unique solution x^* of nonlinear equation

$$\mathcal{F}(x) = 0 \quad (1.1)$$

where \mathcal{F} is a twice-Fréchet continuously differentiable operator defined on a nonempty convex subset \mathcal{D} of a Banach space \mathbf{X} with values in a Banach space \mathbf{Y} . Many problems from computational sciences and other disciplines can be expressed in the form of Eq. (1.1) using Mathematical Modeling [1–6]. The solution of these equations can rarely be found in closed form. That is why the most solution methods for these equations are iterative. The study about convergence analysis of iterative procedures are usually divided into two categories: semilocal and local convergence analysis. The semilocal convergence analysis derive convergence criteria from the information around an initial point while the local analysis find estimates of the radii of convergence balls from the information around a solution. There is ample literature on local as well as semilocal convergence results under various conditions. Interested readers are referred to [7–9, 1, 10–12, 2, 13–15, 3, 16, 4, 5, 17, 18, 6, 19–22, and references therein].

* Corresponding author. Tel.: +47 95220848.

E-mail addresses: iargyros@cameron.edu (I.K. Argyros), sanjay.khattri@hsh.no (S.K. Khattri).

The most popular iterative method for solving Eq. (1.1) is undoubtedly the Newton's method defined by

$$x_{n+1} = x_n - \mathcal{F}'(x_n)^{-1} \mathcal{F}(x_n) \quad \text{for each } n = 0, 1, 2, \dots \quad (1.2)$$

where x_0 is an initial point. And the most famous among the semilocal convergence results is the Kantorovich's theorem for solving nonlinear equations. This theorem provides a simple yet transparent convergence criterion for operators with bounded second derivatives \mathcal{F}'' or the Lipschitz continuous first derivatives [2–4]. Another important theorem was inaugurated by Professor Smale at the International Conference of Mathematicians (cf. [6]). In this theorem, the concept of an approximate zero was proposed and the convergence criteria were provided to determine an approximate zero for analytic functions depending on the information at the initial point. Wang et al. [21] generalized Smale's result by introducing the γ -condition. According to (1.2), Newton's method requires to solve equation

$$\mathcal{F}'(x_n)w_n = -\mathcal{F}(x_n) \quad \text{for each } n = 0, 1, 2, \dots \quad (1.3)$$

However, in practice it may not be feasible to solve the preceding equation exactly especially when $\mathcal{F}'(x_n)$ is too large and dense. That is why, as a suitable alternative to solving Eq. (1.3) [4,5,22], linear iterative methods have been extensively developed to approximate the solution of (1.3). These methods are called Inexact Newton Methods (**INM**) and are defined by the Algorithm 1 [8,9,12,15,16,4,17,18,22].

Algorithm 1 Inexact Newton Method

1: Given the residual control z_n and iteration x_n , find the step w_n satisfying

$$\mathcal{F}'(x_n)w_n = -\mathcal{F}(x_n) + z_n.$$

2: Set

$$x_{n+1} = x_n + w_n.$$

3: Set $n = n + 1$ and go to step 1.

Note that $\{z_n\} \in \mathbf{Y}$ and in general depends on $\{x_n\}$. There are many convergence results for the (**INM**). We refer the reader to [3,16,4,5,22, and references therein]. These results depend either on the Kantorovich or the Smale (α , γ)-theory.

Let $U(x, R)$, $\bar{U}(x, R)$ stand, respectively, for the open and closed balls in \mathbf{X} with center x and radius $R > 0$. Let also $\mathbb{L}(\mathbf{X}, \mathbf{Y})$ be the space of bounded linear operators from \mathbf{X} into \mathbf{Y} . The following Kantorovich-type conditions were used by Shen and Li (see [17]) to prove the semilocal convergence of (**INM**):

\mathcal{C}_1 : there exists $x_0 \in \mathcal{D}$ such that $\mathcal{F}'(x_0)^{-1} \in \mathbb{L}(\mathbf{Y}, \mathbf{X})$ and $\|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_0)\| \leq a$;

\mathcal{C}_2 : $\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x) - \mathcal{F}'(y))\| \leq \mathcal{L}\|x - y\|$ for each x and y in \mathcal{D} ;

\mathcal{C}_3 : $\|\mathcal{F}'(x_0)^{-1}z_n\| \leq \eta_n \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_n)\|^{1+\phi}$;

\mathcal{C}_4 : $\eta := \sup_n \eta_n < 1$.

Now onwards the conditions \mathcal{C}_1 – \mathcal{C}_4 would be referred to as (**C**). The computation of the upper bounds on the norms $\|\mathcal{F}'(x)^{-1} \mathcal{F}'(x_0)\|$ using (\mathcal{C}_2) and the Banach Lemma on invertible operators [1–4] leads to the estimate

$$\|\mathcal{F}'(x)^{-1} \mathcal{F}'(x_0)\| \leq \frac{1}{1 - \mathcal{L}\|x - x_0\|} \quad \text{for each } x \in U\left(x_0, \frac{1}{\mathcal{L}}\right) \cap \mathcal{D}. \quad (1.4)$$

In view of (\mathcal{C}_2) there exists \mathcal{L}_0 such that the center-Lipschitz condition

\mathcal{C}'_2 : $\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x) - \mathcal{F}'(x_0))\| \leq \mathcal{L}_0\|x - x_0\|$ for each $x \in \mathcal{D}$.

Note that in general

$$\mathcal{L}_0 \leq \mathcal{L} \quad (1.5)$$

holds and $\mathcal{L}_0/\mathcal{L}$ can be arbitrarily small [8,16]. Using (\mathcal{C}'_2) instead of (\mathcal{C}_2) we obtain

$$\|\mathcal{F}'(x)^{-1} \mathcal{F}'(x_0)\| \leq \frac{1}{1 - \mathcal{L}_0\|x - x_0\|} \quad \text{for each } x \in U\left(x_0, \frac{1}{\mathcal{L}_0}\right) \cap \mathcal{D}_0. \quad (1.6)$$

Notice that

$$U\left(x_0, \frac{1}{\mathcal{L}}\right) \subset U\left(x_0, \frac{1}{\mathcal{L}_0}\right). \quad (1.7)$$

If $\mathcal{L}_0 < \mathcal{L}$, then the estimate (1.6) is tighter than the estimate (1.4). This observation leads to tighter majorizing sequences for the (**INM**). Then it is reasonable to expect that these majorizing sequences converge under weaker sufficient convergence

criteria. Results of these type have been provided by us in [8,9,1,10–12,2] mainly for the Newton-type methods. In the present paper, we are motivated by the work of Shen and Li (see [17]) (which improved the results by Guo [15]) and also by optimization consideration and estimate (1.7). We prove that under the (C) conditions the following advantages can be obtained:

1. weaker sufficient convergence criteria;
2. tighter error estimates on the distances $\|x_{n+1} - x_n\|$, $\|x_n - x^*\|$;
3. at least as precise information on the location of the solution x^* .

These advantages are obtained under the same computational cost as before (cf. [17]), since in practice the computation of \mathcal{L} requires computing \mathcal{L}_0 thus (\mathcal{C}'_2) is not any additional computational burden to the (\mathcal{C}_2) hypothesis.

The rest of the paper is organized as follows. In Section 2, we present results on the convergence of majorizing sequences for $\{x_n\}$. Section 3 develops the semilocal convergence analysis of the (INM) using the (C) conditions. Finally numerical examples are provided in the concluding Section 4.

2. Majorizing sequences and their convergence

We present scalar sequences and they are shown to be majorizing for (INM). Let $\eta \geq 0$, $\mathcal{L}_0 > 0$, $\mathcal{L} > 0$ and $\phi \in [0, 1]$. Set

$$\mathcal{K}_0 = \mathcal{L}_0 (1 + \eta^{1/(1+\phi)}), \quad \mathcal{K} = \mathcal{L} (1 + \eta^{1/(1+\phi)}) \quad \text{and} \quad \mathcal{M} = 2\eta (1 + \eta^{1/(1+\phi)})^{-\phi}. \quad (2.1)$$

We can prove the following auxiliary results on majorizing sequences for (INM) using the above notation.

Lemma 2.1. Suppose for some $s > 0$ that

$$s_n < \frac{1}{\mathcal{L}_0} \quad \text{for each } n = 1, 2, 3, \dots, \quad (2.2)$$

where s_n is a scalar sequence defined by

$$s_0 = 0, \quad s_1 = s, \quad s_{n+2} = s_{n+1} + \frac{\mathcal{K}(s_{n+1} - s_n) + \mathcal{M}(s_{n+1} - s_n)^\phi}{2(1 - \mathcal{L}_0 s_{n+1})} (s_{n+1} - s_n). \quad (2.3)$$

Then, sequence $\{s_n\}$ is well defined, increasing and converges to its unique least upper bound s^* which satisfies

$$s < s^* \leq \frac{1}{\mathcal{L}_0}. \quad (2.4)$$

Proof. It follows from (2.2) and (2.5) that sequence $\{s_n\}$ is increasing and bounded above by $1/\mathcal{L}_0$ and as such it converges to s^* which satisfies (2.4). That completes the proof of the Lemma. \square

Remark 2.1. The verification of (2.2) is possible, especially if there exists an integer N such that $s_N = s_{N+n}$ for each $n = 0, 1, 2, \dots$. In this case: $s^* = s_N$. Next we provide several, stronger than (2.2) sufficient convergent criteria which imply (2.2).

Lemma 2.2. Let sequences $\{f_n\}$, $\{g_n\}$, $\{h_n\}$, $\{\alpha_n\}$ for each $n = 1, 2, 3, \dots$, be defined on $(0, 1)$ by

$$f_n(t) = \frac{\mathcal{K}}{2} st^n + \frac{\mathcal{M}}{2} s^\phi t^{\phi n} + \mathcal{L}_0 st \frac{1 - t^{n+1}}{1 - t} - t, \quad (2.5)$$

$$g_n(t) = \frac{\mathcal{K}}{2} st^{n+1} - \frac{\mathcal{K}}{2} st^n + \frac{\mathcal{M}}{2} s^\phi t^{\phi(n+1)} - \frac{\mathcal{M}}{2} s^\phi t^{\phi n} + \mathcal{L}_0 st^{n+2}, \quad (2.6)$$

$$h_n(t) = \frac{\mathcal{K}}{2} s(t-1)^2 t^n + \frac{\mathcal{M}}{2} s^\phi (t^\phi - 1)^2 t^{\phi n} + \mathcal{L}_0 s(t-1)t^{n+2}, \quad (2.7)$$

$$\alpha_n = \frac{\mathcal{K}(s_{n+1} - s_n) + \mathcal{M}(s_{n+1} - s_n)^\phi}{2(1 - \mathcal{L}_0 s_{n+1})}, \quad (2.8)$$

where $\mathcal{L}_0 > 0$, $\mathcal{L} > 0$, $\mathcal{K}_0 > 0$, $\mathcal{K} > 0$, $\phi \in [0, 1]$, $s \in (0, 1/\mathcal{L}_0)$ and $\mathcal{M} \geq 0$ are given in (2.1). Suppose that either of the following two criteria is satisfied:

(1) there exists $\alpha \in (0, 1)$ such that

$$0 < \alpha_0 = \frac{\mathcal{K}s + \mathcal{M}s^\phi}{2(1 - \mathcal{L}_0s)} \leq \alpha \leq 1 - \mathcal{L}_0s \quad (2.9)$$

and

$$h_n(\alpha) \geq 0; \quad (2.10)$$

or

(2) there exists $\alpha \in (0, 1)$ such that

$$\alpha_0 \leq \alpha; \quad (2.11)$$

$$h_n(\alpha) \leq 0 \quad (2.12)$$

and

$$f_1(\alpha) \leq 0. \quad (2.13)$$

Then, sequence $\{s_n\}$ – given by (2.3) – is well defined, increasing bounded from above by $s^{**} = s/(1 - \alpha)$ and converges to its unique least upper bound s^* which satisfies

$$s \leq s^* \leq s^{**}. \quad (2.14)$$

Moreover, the following estimates

$$0 < s_{n+1} - s_n \leq \alpha^n s \quad (2.15)$$

and

$$s^* - s_n \leq \frac{\alpha^n s}{1 - \alpha} \quad (2.16)$$

hold for each $n = 1, 2, 3, \dots$

Proof. We shall prove using mathematical induction that

$$0 < \alpha_k \leq \alpha. \quad (2.17)$$

Then estimate (2.15) follows from (2.3), (2.9) and (2.11). Estimate (2.17) holds for $k = 0$ by (2.3) for $n = 0$ and (2.9). Let us assume that (2.17) and with it (2.15) hold for all $k \leq n$. Then we have by (2.3), (2.8), (2.15) and (2.17) that

$$0 < s_{k+1} - s_k \leq \alpha^k s \quad \text{and} \quad s_{k+1} \leq \frac{1 - \alpha^{1+k}}{1 - \alpha} s < s^*. \quad (2.18)$$

Evidently, estimate (2.17) holds if

$$\frac{\mathcal{K}}{2}s\alpha^k + \frac{\mathcal{M}}{2}s^\phi\alpha^{\phi k} + \mathcal{L}_0s\alpha\frac{1 - \alpha^{k+1}}{1 - \alpha} - \alpha \leq 0. \quad (2.19)$$

The above estimate motivates us to introduce recurrent functions f_k . Then we must show

$$f_k(\alpha) \leq 0. \quad (2.20)$$

From (2.5)–(2.7) we obtain

$$f_{k+1}(\alpha) = f_k(\alpha) + g_k(\alpha), \quad (2.21)$$

where

$$g_{k+1}(\alpha) = g_k(\alpha) + h_k(\alpha). \quad (2.22)$$

If (2.10) holds then we have that

$$f_k(\alpha) \leq f_{k+1}(\alpha). \quad (2.23)$$

We define function f_∞ on $(0, 1)$ as follows

$$f_\infty(t) = \lim_{k \rightarrow \infty} f_k(t). \quad (2.24)$$

We get by (2.21) and (2.22) that

$$f_k(\alpha) \leq f_\infty(\alpha). \quad (2.25)$$

Then, instead of (2.20) we can show

$$f_{\infty}(\alpha) \leq 0. \quad (2.26)$$

But we have by (2.5), (2.19) and (2.24) that

$$f_{\infty}(\alpha) = \alpha \left(\frac{\mathcal{L}_0}{1-\alpha} s - 1 \right). \quad (2.27)$$

Recalling that $\alpha \leq 1 - \mathcal{L}_0 s$, we deduce by (2.27) that (2.26) is true. If (2.12) holds then, we have

$$f_{k+1}(\alpha) \leq f_k(\alpha). \quad (2.28)$$

Hence, we must show instead of (2.20) that $f_1(\alpha) \leq 0$ which is hypothesis (2.13). The induction for (2.17) is complete. Hence, sequence $\{s_n\}$ is increasing, bounded from above by s^{**} and as such it converges to s^* . The proof of the lemma is complete. \square

In the case when $0 \leq \phi < 1$ we can have easier to verify criteria than the ones in the Lemma 2.3.

Lemma 2.3. Let $0 \leq \phi \leq 1$. Suppose that (2.9) is satisfied for

$$\alpha = \frac{(\mathcal{K}s + \mathcal{M}s^{\phi})}{1 - \mathcal{L}_0 s + \frac{\mathcal{K}}{2}s + \frac{\mathcal{M}}{2}s^{\phi} + \sqrt{\left(1 - \mathcal{L}_0 s + \frac{\mathcal{K}}{2}s + \frac{\mathcal{M}}{2}s^{\phi}\right)^2 - 2(\mathcal{K}s + \mathcal{M}s^{\phi})}} \quad (2.29)$$

and

$$\mathcal{K}s + \mathcal{M}s^{\phi} \leq 2(1 + \mathcal{L}_0 s) \quad (2.30)$$

hold, where $\mathcal{L}_0 > 0$, $\mathcal{L} > 0$, $\mathcal{K}_0 > 0$, $\mathcal{K} > 0$, $\phi \in [0, 1]$, $s \in (0, 1/\mathcal{L}_0)$ and $\mathcal{M} \geq 0$ are given in (2.1) and sequence $\{s_n\}$ is defined by (2.3). Then, the conclusions of the Lemma 2.3 hold.

Proof. Define function q on $[0, 1/\mathcal{L}_0]$ by

$$q(t) = \left(1 - \mathcal{L}_0 t + \frac{\mathcal{K}}{2}t + \frac{\mathcal{M}}{2}t^{\phi}\right)^2 - 2(\mathcal{K}t + \mathcal{M}t^{\phi}).$$

Using (2.9) and (2.30), we may notice that $q(0) = 1 > 0$ and $q(1/\mathcal{L}_0) < 0$. Hence, the discriminant of the quadratic polynomial

$$p(t) = t^2 - \left(1 - \mathcal{L}_0 s + \frac{\mathcal{K}}{2}s + \frac{\mathcal{M}}{2}s^{\phi}\right)t + \frac{\mathcal{K}}{2}s + \frac{\mathcal{M}}{2}s^{\phi} \quad (2.31)$$

is non-negative and α given by (2.29) satisfies $p(\alpha) = 0$. According to the proof of the Lemma 2.3, we must show (2.17) or the stronger inequality

$$0 < \frac{\mathcal{K}s + \mathcal{M}s^{\phi}}{2(1 - \mathcal{L}_0 s_{k+1})} \leq \alpha. \quad (2.32)$$

Estimate (2.32) is true for $k = 0$ by (2.9). Then, as in the proof of the Lemma 2.3, we show that

$$f_k^1(\alpha) \leq 0, \quad (2.33)$$

where

$$f_k(t)^1 = t\mathcal{L}_0 s \left(\frac{1 - t^{k+1}}{1 - t} \right) - t + \frac{\mathcal{K}}{2}s + \frac{\mathcal{M}}{2}s^{\phi}. \quad (2.34)$$

We have

$$f_{k+1}^1(\alpha) = f_k^1(\alpha) + \mathcal{L}_0 s \alpha^{k+2} \geq f_k^1(\alpha).$$

Define f_{∞}^1 on $[0, 1)$ by

$$f_{\infty}^1(t) = \lim_{k \rightarrow \infty} f_k(t).$$

Then, we must show that

$$f_{\infty}^1(t) \leq 0.$$

But

$$f_{\infty}^1(\alpha) = \frac{\alpha \mathcal{L}_0 s}{1 - \alpha} - \alpha + \frac{\mathcal{K}}{2}s + \frac{\mathcal{M}}{2}s^{\phi}$$

and which can be non positive by the choice of α given in (2.29). The proof of the lemma is complete. \square

In the case when $\phi = 1$ or $\eta = 0$, let $\mathcal{L}_1 = \mathcal{K} + \mathcal{M}$. Then, Lemma 2.3 reduces to the following lemma (see [10, Lemma 2.1])

Lemma 2.4. Assume that there exist constants $\mathcal{L}_0 > 0$, $\mathcal{L}_1 > 0$ and $s > 0$ such that:

$$h_2 = \overline{\mathcal{L}} s \begin{cases} \leq \frac{1}{2}, & \text{if } \mathcal{L}_0 \neq 0 \\ < \frac{1}{2}, & \text{if } \mathcal{L}_0 = 0, \end{cases} \quad (2.35)$$

where,

$$\overline{\mathcal{L}} = \frac{1}{8} \left(\mathcal{L}_1 + 4 \mathcal{L}_0 + \sqrt{\mathcal{L}_1^2 + 8 \mathcal{L}_0 \mathcal{L}_1} \right). \quad (2.36)$$

Then, sequence $\{s_k\}$ ($k \geq 0$) given by

$$s_0 = 0, \quad s_1 = s, \quad s_{k+1} = s_k + \frac{\mathcal{L}_1 (s_k - s_{k-1})^2}{2 (1 - \mathcal{L}_0 s_k)} \quad (k \geq 1), \quad (2.37)$$

is well defined, non decreasing, bounded above by s^{**} , and converges to its unique least upper bound $s^{**} \in [0, s^{**}]$, where

$$s^{**} = \frac{s}{1 - \alpha}, \quad (2.38)$$

$$\frac{1}{2} \leq \alpha = \frac{2 \mathcal{L}_1}{\mathcal{L}_1 + \sqrt{\mathcal{L}_1^2 + 8 \mathcal{L}_0 \mathcal{L}_1}} < 1 \quad \text{for } \mathcal{L}_0 \neq 0. \quad (2.39)$$

Moreover, the following estimates hold:

$$\mathcal{L}_0 s^{*} \leq 1, \quad (2.40)$$

$$0 \leq s_{k+1} - s_k \leq \alpha (s_k - s_{k-1}) \leq \dots \leq \alpha^k s, \quad (k \geq 1), \quad (2.41)$$

$$s_{k+1} - s_k \leq \alpha^k (2 h_2)^{2^k - 1} s, \quad (k \geq 0), \quad (2.42)$$

$$0 \leq s^{*} - s_k \leq \alpha^k \frac{(2 h_2)^{2^k - 1} s}{1 - (2 h_2)^{2^k}}, \quad (2 h_2 < 1), \quad (k \geq 0). \quad (2.43)$$

Definition 2.2. Define scalar sequence $\{r_n\}$ as follows

$$\begin{aligned} r_0 &= 0, \quad r_1 = s, \quad r_2 = r_1 + \frac{\mathcal{K}_0(r_1 - r_0) + \mathcal{M}(r_1 - r_0)^{\phi}}{2(1 - \mathcal{L}_0 r_1)}(r_1 - r_0), \\ r_{n+2} &= r_{n+1} + \frac{\mathcal{K}(r_{n+1} - r_n) + \mathcal{M}(r_{n+1} - r_n)^{\phi}}{2(1 - \mathcal{L}_0 r_{n+1})}(r_{n+1} - r_n) \end{aligned} \quad (2.44)$$

for each $n = 1, 2, \dots$. Here, $\mathcal{K}_0 = \mathcal{L}_0(1 + \eta^{1/(1+\phi)})$.

A simple inductive argument shows that

$$r_n \leq s_n, \quad (2.45)$$

$$r_{n+1} - r_n \leq s_{n+1} - s_n \quad (2.46)$$

and

$$r^{*} \leq s^{*} = \lim_{n \rightarrow \infty} s_n. \quad (2.47)$$

Estimates (2.45) and (2.46) hold as strict inequalities if $\mathcal{K}_0 < \mathcal{K}$. Clearly $\{r_n\}$ can replace $\{r_n\}$ in all of the preceding lemmas. However, a more direct study, for the interesting case when $\eta = 0$, leads to the following convergence results for sequence $\{r_n\}$ [12,2].

Lemma 2.5. Let $\mathcal{L}_1 > 0$, $\mathcal{L}_0 > 0$ and $s \geq 0$ be given constants such that

$$h_1 = \overline{\mathcal{L}} s \leq \frac{1}{2} \quad (2.48)$$

holds, where

$$\overline{\mathcal{L}} = \frac{1}{8} \left(4\mathcal{L}_0 + \sqrt{\mathcal{L}_0 \mathcal{L}_1} + \sqrt{\mathcal{L}_1 \mathcal{L}_0 + 8\mathcal{L}_0^2} \right). \quad (2.49)$$

Then, scalar sequence $\{r_n\}$ ($n \geq 0$) given by

$$\begin{aligned} r_0 &= 0, & r_1 &= s, & r_2 &= s + \frac{\mathcal{L}_0 s^2}{2(1 - \mathcal{L}_0 s)}, \\ r_{n+2} &= r_{n+1} + \frac{\mathcal{L}_1 (r_{n+1} - r_n)^2}{2(1 - \mathcal{L}_0 r_{n+1})} \quad (n \geq 1) \end{aligned} \quad (2.50)$$

is well defined, increasing, bounded from above by

$$r^{**} = s + \frac{\mathcal{L}_0 s^2}{2(1 - \alpha)(1 - \mathcal{L}_0 s)} \quad (2.51)$$

and converges to its unique least upper bound r^* satisfying

$$0 \leq r^* \leq r^{**}. \quad (2.52)$$

Moreover, the following estimates hold:

$$0 < r_{n+2} - r_{n+1} \leq \alpha^n \frac{\mathcal{L}_0 s^2}{2(1 - \mathcal{L}_0 \eta)} \quad (n \geq 1). \quad (2.53)$$

Lemma 2.6. Let $\mathcal{L}_1 > 0$, $\mathcal{L}_0 > 0$ and $s \geq 0$ be given constants with $\mathcal{L}_1 \geq \mathcal{L}_0$. Assume there exists a minimum integer $N > 1$ such that iterates r_i ($i = 0, 1, \dots, N - 1$) given by (2.50) are well defined,

$$r_i < r_{i+1} < \frac{1}{\mathcal{L}_0}, \quad i = 0, 1, \dots, N - 2 \quad (2.54)$$

and

$$r_N \leq \frac{1}{\mathcal{L}_0} (1 - (1 - \mathcal{L}_0 r_{N-1}) \alpha). \quad (2.55)$$

Then, the following assertions hold

$$\mathcal{L}_0 r_N < 1, \quad (2.56)$$

$$r_{N+1} \leq \frac{1}{\mathcal{L}_0} (1 - (1 - \mathcal{L}_0 r_N) \alpha), \quad (2.57)$$

$$\alpha_{N-1} \leq \alpha \leq 1 - \frac{\mathcal{L}_0 (r_{N+1} - r_N)}{1 - \mathcal{L}_0 r_N}, \quad (2.58)$$

sequence $\{r_n\}$ ($n \geq 0$) given by (2.50) is well defined, increasing, bounded from above by

$$r^{**} = r_{N-1} + \frac{1}{1 - \alpha} (r_N - r_{N-1}) \quad (2.59)$$

and converges to its unique least upper bound r^* satisfying

$$0 \leq r^* \leq r^{**}, \quad (2.60)$$

where α is given by (2.39) and

$$\alpha_n = \frac{\mathcal{L} (r_{n+2} - r_{n+1})}{2(1 - \mathcal{L}_0 r_{n+2})}. \quad (2.61)$$

Moreover, the following estimates hold:

$$0 < r_{N+n} - r_{N+n-1} \leq \alpha^{n-1} (r_{N+1} - r_N) \quad (n \geq 1). \quad (2.62)$$

Remark 2.3. If $N = 2$, we must have

$$r_2 = s + \frac{\mathcal{L}_0 s}{2(1 - \mathcal{L}_0 s)} \leq \frac{\mathcal{L}_1 s + 2\alpha}{\mathcal{L}_1 + 2\alpha \mathcal{L}_0}$$

which is (2.48). When $N > 2$, we do not have closed form inequalities (solved for n) any more given by

$$c_0 s \leq c_1, \quad (2.63)$$

where c_0, c_1 may depend on \mathcal{L}_0 and \mathcal{L}_1 (like (2.35) or (2.48)). However, the corresponding inequalities can also be verified, since only computations involving s, \mathcal{L}_0 and \mathcal{L}_1 are carried out (see also Section 4).

Clearly, the sufficient convergence conditions of the form (2.55) becomes weaker as N increases. Note also that (2.56) and (2.57) imply that $\{r_n\}$ is increasing, bounded from above by $1/\mathcal{L}_0$ and as such it converges to some limit point $r_\infty \in [0, 1/\mathcal{L}_0]$. However, we use (2.58) to arrive at r^{**} , which provides a better information than r_∞ on the upper bounds for sequence $\{r_n\}$.

Definition 2.4. Let there be a positive constant $a > 0$. We define the following parameters and functions

$$\begin{aligned} \sigma &= \frac{\mathcal{L}(1 + \eta^{1/(1+\phi)})}{(1 + a\mathcal{L}\eta(1 + \eta))(1 + \eta(a^\phi - 1))}, & \theta &= \frac{\eta}{1 + \eta(a^\phi - 1)}, \\ \lambda &= \begin{cases} a(1 + \eta), & \text{if } \phi = 0, \\ a(1 + \eta^{1/(1+\phi)})(1 + \theta), & \text{if } \phi > 0, \end{cases} \\ \varphi_\phi(t) &= \frac{\sigma}{2}t^2 + 2^{1-\phi}\theta t^{1+\phi} - (1 + \theta)t + \lambda, & \psi_\phi(t) &= \frac{\sigma}{2}t^2 + \theta t^{1+\phi} - (1 + \theta)t + \lambda, \\ b &= \frac{\sigma}{2}t^2 + \phi 2^{1-\phi}\theta \mu^{1+\lambda}, \end{aligned}$$

where μ is the unique positive zero of the function φ'_ϕ . Furthermore, we define the iteration

$$t_0 = 0, \quad t_{n+1} = t_n - \frac{\varphi_\phi(t_n)}{\psi'_\phi(t_n)} \quad (2.64)$$

for each $n = 0, 1, 2, \dots$

In order to make the paper as self contained as possible, the subsequent definition and lemma are borrowed from the work of Shen and Li (see [17]).

Lemma 2.7. The function φ_ϕ is strictly decreasing on $[0, \mu]$ and satisfies $\varphi_\phi(\mu) = \lambda - b$. Moreover, if

$$\lambda \leq b \quad (2.65)$$

then the following assertions hold

\mathcal{A}_1 : φ_ϕ has a unique zero t^* in $[0, \mu]$

\mathcal{A}_2 : $t_n < t_{n+1} < t^* = \lim_{n \rightarrow \infty} t_n$ for each $n = 0, 1, 2, \dots$,

\mathcal{A}_3 : $t_{n+1} - t_n \leq \frac{\lambda}{1+\theta}$ for $\phi > 0$ and for each $n = 0, 1, 2, 3, \dots$

Remark 2.5. Convergence criteria (2.65) can also be replaced by the Kantorovich-type (see [17])

$$a \leq \begin{cases} \frac{(1-\eta)^2}{\mathcal{L}(1+\eta)(2(1+\eta)-\eta(1-\eta)^2)}, & \phi = 0, \\ \min \left\{ \frac{\sigma \mu^2 + 2\phi 2^{1-\phi} \theta \mu^{1+\phi}}{2(1+\theta)(1+\eta^{1/(1+\phi)})}, \eta^{-1/(1+\phi)} \right\}, & \phi > 0, \end{cases} \quad (2.66)$$

or

$$a \leq \min \left\{ \frac{2^{(1-1/\phi)}}{\delta^{1/\phi} \delta_0^{(1+1/\phi)}} - \frac{2^{(1-1/\phi)} \eta}{\delta^{(1+1/\phi)} \delta_0^{(2+1/\phi)}} - \frac{2^{(1-2/\phi)} \mathcal{L}}{\delta^{2/\phi} \delta_0^{2/\phi}}, \eta^{-1/(1+\phi)} \right\}, \quad \phi > 0, \quad (2.67)$$

where

$$\delta = \mathcal{L} + (1 + \phi)\eta \quad \text{and} \quad \delta_0 = 1 + \eta^{1/(1+\phi)}.$$

Remark 2.6. It was shown by Shen and Li (see [17, Theorem 3.1]) that if (2.65) holds then $\{t_n\}$ is a majorizing sequence for the (INM). According to their proof

$$s_n \leq t_n, \quad (2.68)$$

$$s_{n+1} - s_n \leq t_{n+1} - t_n, \quad (2.69)$$

$$s^* \leq t^*. \quad (2.70)$$

Hence, (2.65) is another sufficient convergence condition for our sequences $\{s_n\}$ and $\{r_n\}$.

So far we have proved that the sequence $\{r_n\}$ is tighter than $\{s_n\}$ which is tighter than $\{t_n\}$ and all of these sequences converge under the condition (2.65). Sequences $\{s_n\}$ and $\{r_n\}$ can also converge under weaker criteria than (2.65) provided the lemmas before Definition 2.10 hold. Clearly (2.2) is the weakest criterion among all the criteria presented in this section. Let us compare these conditions.

In the interesting case when $\eta = 0$ and $\phi = 1$ then (2.65) reduces to the famous, for its simplicity and clarity, Kantorovich criterion

$$h = \mathcal{L} a \leq \frac{1}{2}, \quad (2.71)$$

(see [1–4]). Note that

$$h \leq \frac{1}{2} \implies h_1 \leq \frac{1}{2} \implies h_2 \leq \frac{1}{2} \quad (2.72)$$

but not necessarily vice versa unless if $\mathcal{L}_0 = \mathcal{L}$. We have that

$$\frac{h_1}{h} \longrightarrow \frac{1}{4}, \quad \frac{h_2}{h} \longrightarrow 0, \quad \frac{h_2}{h_1} \longrightarrow 0 \quad \text{as} \quad \frac{\mathcal{L}_0}{\mathcal{L}} \longrightarrow 0. \quad (2.73)$$

Hence, we conclude that $\{s_n\}$, $\{r_n\}$ are tighter than $\{t_n\}$ and, in some interesting cases, $\{s_n\}$, $\{r_n\}$ also converge under weaker hypotheses.

In the next section, we show that these sequences are indeed majorizing sequences for $\{x_n\}$.

3. Semilocal convergence of (INM)

Using majorizing sequence $\{s_n\}$, defined in (2.3), we develop the main semilocal convergence result for (INM).

Theorem 3.1. Suppose that (2.2),

$$a\eta^{1/(1+\phi)} \leq 1 \quad (3.1)$$

and conditions (C) hold for

$$\text{each } x \in \bar{U}\left(x_0, \frac{1}{\mathcal{L}_0}\right) \cap \mathcal{D}. \quad (3.2)$$

Then, sequence $\{x_n\}$ generated by the Algorithm 1 is well defined, remains in $\bar{U}(x_0, 1/\mathcal{L}_0)$ for all $n \geq 0$ and converges to a solution x^* of equation $\mathcal{F}(x) = 0$. Moreover, the following estimate hold

$$\|x_n - x^*\| \leq s^* - s_n \quad \text{for each } n = 0, 1, 2, \dots, \quad (3.3)$$

where $s_1 = s = a\left(1 + \eta^{1/(1+\phi)}\right)$.

Proof. We use mathematical induction to prove that

$$\frac{1 + \eta^{1/(1+\phi)}}{1 - \mathcal{L}_0 s_k} \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_k)\| \leq s_{k+1} - s_k, \quad (3.4)$$

$$\|x_{k+1} - x_k\| \leq s_{k+1} - s_k \quad (3.5)$$

and

$$\bar{U}(x_{k+1}, s^* - s_{k+1}) \subseteq \bar{U}(x_k, s^* - s_k). \quad (3.6)$$

If $\{x_n\}$ is well defined then

$$\|\mathcal{F}'(x_0)^{-1} z_k\| \leq \eta_k \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_k)\|^{1+\phi} \leq \eta \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_k)\|^{1+\phi}. \quad (3.7)$$

By the definitions of λ , a , s_1 and the Algorithm I, the inequality (3.7), we obtain

$$\begin{aligned}\|x_1 - x_0\| &= \|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_0) + \mathcal{F}'(x_0)^{-1}z_0\| \\ &\leq \|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_0)\| + \|\mathcal{F}'(x_0)^{-1}z_0\| \\ &\leq a + \eta a^{1+\phi} \leq a(1 + \eta^{1/(1+\phi)}) = s_1 - s_0,\end{aligned}$$

since by (3.1), $\eta a^\phi \leq \eta^{1/(1+\phi)}$. Hence (3.5) holds for $k = 0$. Suppose that (3.4) and (3.5) hold for all natural integers $k \leq n$. Then, we have that

$$\|x_{k+1} - x_0\| \leq \sum_{i=1}^{k+1} \|x_i - x_{i-1}\| \leq \sum_{i=1}^{k+1} (s_i - s_{i-1}) = s_{k+1} - s_0 = s_{k+1}$$

and

$$\|x_k + \tau(x_{k+1} - x_k) - x_0\| \leq s_k + \tau(s_{k+1} - s_k) \leq s^*$$

for each $\tau \in [0, 1]$. Using (\mathcal{C}'_2) for $x = x_{k+1}$ and the induction hypotheses, we get

$$\|\mathcal{F}'(x_0)^{-1}(\mathcal{F}'(x_{k+1} - \mathcal{F}'(x_0)))\| \leq \mathcal{L}_0 \|x_{k+1} - x_0\| \leq \mathcal{L}_0 s_{k+1} < 1. \quad (3.8)$$

It follows from (3.8) and the Banach Lemma on invertible operators [1–4] that

$$\begin{aligned}\mathcal{F}'(x_{k+1})^{-1} &\in \mathbb{L}(\mathbf{Y}, \mathbf{X}) \\ \|\mathcal{F}'(x_{k+1})^{-1}\mathcal{F}'(x_0)\| &\leq \frac{1}{1 - \mathcal{L}_0 \|x_{k+1} - x_0\|} \leq \frac{1}{1 - \mathcal{L}_0 s_{k+1}}.\end{aligned} \quad (3.9)$$

Using Algorithm I, we obtain that

$$\begin{aligned}\mathcal{F}(x_{k+1}) &= \mathcal{F}(x_{k+1}) - \mathcal{F}(x_k) - \mathcal{F}'(x_k)(x_{k+1} - x_k) + z_k \\ &= \int_0^1 [\mathcal{F}'(x_k^\tau) - \mathcal{F}'(x_k)] d\tau (x_{k+1} - x_k) + z_k\end{aligned} \quad (3.10)$$

where $x_k^\tau = x_k + \tau(x_{k+1} - x_k)$ for each $\tau \in [0, 1]$. Then, we have by the foregoing equation

$$\begin{aligned}\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1})\| &\leq \|\mathcal{F}'(x_0)^{-1} \int_0^1 [\mathcal{F}'(x_k^\tau) - \mathcal{F}'(x_k)] d\tau (x_{k+1} - x_k)\| + \|\mathcal{F}'(x_0)^{-1}z_k\| \\ &\leq \mathcal{B}_1 + \mathcal{B}_2.\end{aligned} \quad (3.11)$$

In view of (\mathcal{C}_2) , we get that

$$\begin{aligned}\mathcal{B}_1 &\leq \int_0^1 \mathcal{L} \|x_k^\tau - x_k\| \|x_{k+1} - x_k\| d\tau \\ &= \frac{\mathcal{L}}{2} \|x_{k+1} - x_k\|^2 \leq \frac{\mathcal{L}}{2} (s_{k+1} - s_k)^2.\end{aligned} \quad (3.12)$$

Moreover by (3.4), we have

$$\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_k)\| \leq \left(1 + \eta^{1/(1+\phi)}\right)^{-1} (s_{k+1} - s_k)$$

and by (3.7)

$$\begin{aligned}\mathcal{B}_2 &= \|\mathcal{F}'(x_0)^{-1}z_k\| \leq \eta \|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_k)\|^{1+\phi} \\ &\leq \eta \left(1 + \eta^{1/(1+\phi)}\right)^{-1-\phi} (s_{k+1} - s_k)^{1+\phi}.\end{aligned} \quad (3.13)$$

In view of (3.11)–(3.13)

$$\|\mathcal{F}'(x_0)^{-1}\mathcal{F}(x_{k+1})\| \leq \frac{\mathcal{L}_2}{2} (s_{k+1} - s_k)^2 + \eta \left(1 + \eta^{1/(1+\phi)}\right)^{-1-\phi} (s_{k+1} - s_k)^{1+\phi}. \quad (3.14)$$

So,

$$\begin{aligned} \frac{1 + \eta^{1/(1+\phi)}}{1 - \mathcal{L}_0 s_{k+1}} \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\| &\leq \frac{\mathcal{L}_2 (1 + \eta^{1/(1+\phi)})}{2(1 - \mathcal{L}_0 s_{k+1})} (s_{k+1} - s_k)^2 + \frac{\eta (1 + \eta^{1/(1+\phi)})^{-\phi}}{(1 - \mathcal{L}_0 s_{k+1})} (s_{k+1} - s_k)^{1+\phi} \\ &\leq \frac{\mathcal{K}_1 (s_{k+1} - x_k)^2 + \mathcal{M} (s_{k+1} - x_k)^{1+\phi}}{2(1 - \mathcal{L}_0 s_{k+1})}, \end{aligned} \quad (3.15)$$

where

$$\mathcal{K}_1 = \begin{cases} \mathcal{K}_0, & \mathcal{K} = 0 \\ \mathcal{K}, & \mathcal{K} > 0 \end{cases} \quad \text{and} \quad \mathcal{L}_2 = \begin{cases} \mathcal{L}_0, & \mathcal{K} = 0 \\ \mathcal{L}, & \mathcal{K} > 0. \end{cases}$$

We claim that

$$\eta \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\|^{1+\phi} \leq \eta^{1/(1+\phi)} \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\|. \quad (3.16)$$

Estimate (3.16) is true as equality when $\phi = 0$. If $\phi > 0$, we have in turn that

$$\begin{aligned} \eta^{\phi/(1+\phi)} \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\|^\phi &\leq \frac{\eta^{\phi/(1+\phi)} (1 - \mathcal{L}_0 s_{k+1})^\phi}{(1 + \eta^{1/(1+\phi)})^\phi} (s_{k+2} - s_{k+1})^\phi \\ &\leq \frac{\eta^{\phi/(1+\phi)} \lambda^\phi}{(1 + \eta^{1/(1+\phi)}) (1 + \theta)^\phi} = a^\phi \eta^{\phi/(1+\phi)} \leq 1 \end{aligned} \quad (3.17)$$

by the definition of λ and (3.1). Hence, (3.16) is true. Moreover, using (3.3), (3.9), (3.14) and (3.16), we obtain

$$\begin{aligned} \|x_{k+2} - x_{k+1}\| &\leq \|\mathcal{F}'(x_{k+1})^{-1} \mathcal{F}(x_0)\| \left(\|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\| + \|\mathcal{F}'(x_0)^{-1} z_{k+1}\| \right) \\ &\leq \frac{1}{1 - \mathcal{L}_0 s_{k+1}} \left(\|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\| + \eta \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\|^{1+\phi} \right) \\ &\leq \frac{(1 + \eta^{1/(1+\phi)})}{1 - \mathcal{L}_0 s_{k+1}} \|\mathcal{F}'(x_0)^{-1} \mathcal{F}(x_{k+1})\| \\ &\leq \frac{\mathcal{K}_1 (1 + \eta^{1/(1+\phi)}) (s_{k+1} - s_k)^2 + \mathcal{M} (s_{k+1} - s_k)^{1+\phi}}{2(1 - \mathcal{L}_0 s_{k+1})} \\ &\leq s_{k+2} - s_{k+1}, \end{aligned} \quad (3.18)$$

which completes the induction for (3.4) and (3.5). Furthermore, let $z \in \overline{U}(x_{k+2}, s^* - s_{k+2})$. Then, we get that

$$\begin{aligned} \|z - x_{k+1}\| &\leq \|z - x_{k+2}\| + \|x_{k+2} - x_{k+1}\| \\ &\leq s^* - s_{k+2} + s_{k+2} - s_{k+1} = s^* - s_{k+1} \leq s^*, \end{aligned}$$

which implies that $z \in \overline{U}(x_{k+1}, s^* - s_{k+1})$ and it completes the induction for (3.6). It follows from the Lemma 2.1 that sequence $\{s_n\}$ is complete. Hence, sequence $\{x_n\}$ is complete in a Banach space \mathbf{X} and as such it converges to some $x^* \in \overline{U}(x_0, s^*)$ (since $\overline{U}(x_0, s^*)$ is a closed set). By letting $k \rightarrow \infty$ in (3.14), we deduce that $\mathcal{F}(x^*) = 0$. Finally, estimate (3.3) follows from (3.5) by using standard majorization techniques [1–4]. The proof of the theorem is complete. \square

Remark 3.2. Clearly criterion (2.2) of Lemma 2.1 can be replaced in Theorem 3.1 by the stronger but more practical convergence criteria given in the rest of lemmas of Section 2.

4. Numerical examples

We present two numerical examples. The first example is academic and is used to show that earlier sufficient convergence conditions are not satisfied but the new conditions are satisfied. The second example is a nonlinear two point boundary value problem appearing in magnetohydrodynamics.

Example 4.1. Let $\mathbf{X} = \mathbf{Y} = \mathbb{R}^2$ be equipped with the max-norm, $x_0 = (1, 1)^T$, $\mathcal{D} = \{x : \|x - x_0\| < 1 - \epsilon\}$, $\epsilon \in (0, 1/2)$ and define the function \mathcal{F} on \mathcal{D} by

$$\mathcal{F}(x) = \left(\xi_1^3 - \epsilon, \xi_2^3 - \epsilon \right)^T, \quad x = (\xi_1, \xi_2)^T. \quad (4.1)$$

Table 1Comparison among the sequences $\{s_n\}$, $\{r_n\}$ and $\{t_n\}$ for $\phi = 1$, $\eta = 0$ and $\epsilon = 0.7$.

n	r_n	s_n	t_n	$r_{n+1} - r_n$	$s_{n+1} - s_n$	$t_{n+1} - t_n$
0	$0.000 \times 10^{+00}$	$0.000 \times 10^{+00}$	$0.000 \times 10^{+00}$	1.000×10^{-01}	1.000×10^{-01}	1.000×10^{-01}
1	1.000×10^{-01}	1.000×10^{-01}	1.000×10^{-01}	1.494×10^{-02}	1.688×10^{-02}	1.757×10^{-02}
2	1.149×10^{-01}	1.169×10^{-01}	1.176×10^{-01}	3.942×10^{-04}	5.068×10^{-04}	5.778×10^{-04}
3	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	2.749×10^{-07}	4.574×10^{-07}	6.265×10^{-07}
4	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	1.337×10^{-13}	3.725×10^{-13}	7.365×10^{-13}
5	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	3.163×10^{-26}	2.472×10^{-25}	1.018×10^{-24}
6	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	1.771×10^{-51}	1.088×10^{-49}	1.944×10^{-48}
7	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	5.547×10^{-102}	2.108×10^{-98}	7.091×10^{-96}
8	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	5.443×10^{-203}	7.910×10^{-196}	9.436×10^{-191}
9	1.153×10^{-01}	1.174×10^{-01}	1.181×10^{-01}	5.243×10^{-405}	1.114×10^{-390}	1.671×10^{-380}

Then, the Fréchet derivative of the operator \mathcal{F} is given as

$$\mathcal{F}'(x) = \begin{bmatrix} 3\xi_1^2 & 0 \\ 0 & 3\xi_2^2 \end{bmatrix}. \quad (4.2)$$

From the conditions (C) and (\mathcal{C}'_2) , we get

$$a = \frac{(1-\epsilon)}{3}, \quad \mathcal{L}_0 = 3 - \epsilon, \quad \mathcal{L} = 2(2 - \epsilon). \quad (4.3)$$

For simplicity let us assume that $\eta = 0$ and $\phi = 1$. Then sequences $\{r_n\}$, $\{s_n\}$ and $\{t_n\}$ become

$$\left. \begin{aligned} r_0 &= 0, & r_1 &= a, & r_2 &= r_1 + \frac{\mathcal{L}_0(r_1 - r_0)^2}{2(1 - \mathcal{L}_0 r_1)}, \\ r_{n+2} &= r_{n+1} + \frac{\mathcal{L}(r_{n+1} - r_n)^2}{2(1 - \mathcal{L}_0 r_{n+1})}, & n &= 1, 2, 3, \dots \\ s_0 &= 0, & s_1 &= a, & s_{n+2} &= s_{n+1} + \frac{\mathcal{L}(s_{n+1} - s_n)^2}{2(1 - \mathcal{L}_0 s_{n+1})}, & n &= 0, 1, 2, 3, \dots \\ t_0 &= 0, & t_1 &= a, & t_{n+2} &= t_{n+1} + \frac{\mathcal{L}(t_{n+1} - t_n)^2}{2(1 - \mathcal{L} t_{n+1})}, & n &= 0, 1, 2, 3, \dots \end{aligned} \right\}.$$

Then from (2.71), we obtain

$$h = \frac{2(2-\epsilon)(1-\epsilon)}{3} > \frac{1}{2} \quad \text{for all } \epsilon \in \left(0, \frac{1}{2}\right),$$

thus the Kantorovich criterion (2.71) (see [3]) is not satisfied. Hence, there is no guarantee that sequence $\{x_n\}$ starting at x_0 converges to x^* . On the other hand, our criteria (2.35) and (2.48) are satisfied since

$$\left\{ \begin{aligned} h_2 &\leq \frac{1}{2} & \text{for all } \epsilon \in [0.450339002, 0.5), \\ h_1 &\leq \frac{1}{2} & \text{for all } \epsilon \in [0.4271907643, 0.5). \end{aligned} \right.$$

Consequently, our approach expands the applicability of the method in this case. Let us now compare the sequences through the following two interesting cases.

Case I Let us now consider $\epsilon \in (0, 1)$. In order for us to compare sequences $\{r_n\}$, $\{s_n\}$ and $\{t_n\}$, let us choose $\epsilon = 0.7$, $\phi = 1$ and $\eta = 0$. We have: $a = 0.1$, $\mathcal{K}_0 = \mathcal{L}_0 = 2.3$, $\mathcal{K} = \mathcal{L} = \mathcal{L}_1 = 2.6$, $\mathcal{M} = 0$, $\sigma = 2.6$, $\theta = 0$, $\lambda = 0.1$ and $s = \lambda = 0.1$. We may check that all the h criteria, given by (2.35), (2.48) and (2.71), are satisfied

$$h = 0.2600000000 < \frac{1}{2}, \quad h_2 = 0.2398647660 < \frac{1}{2}, \quad h_1 = 0.2324403000 < \frac{1}{2}.$$

Table 1 presents a comparison among the sequences. In the Table 1, we observe that the sequence r_n is converging the fastest.

Case II Let us consider the case when $\eta = 0.1$, $\phi = 0$ and $\epsilon = 0.7$. Then, we have $a = 0.1$, $\mathcal{L} = 2.6$, $\mathcal{L}_0 = 2.3$, $s = 0.11$, $\mathcal{K} = 2.86$, $\mathcal{M} = 2 \times 0.1 \times (1 + 0.1)^0 = 0.2$, $\mathcal{L}_1 = 3.06$, $\mathcal{K}_0 = 2.53$, $\theta = 0.1$, $\lambda = 0.11$, $\sigma = 2.78048$ and $s = \lambda$. The convergence criteria (2.66) is satisfied since $0.1 < 0.133655867$. Therefore the all the sequences should converge.

Table 2 presents a comparison among the sequences. We observe that the sequence $\{r_n\}$ converges the fastest.

Table 2Comparison among the sequences $\{s_n\}$, $\{r_n\}$ and $\{t_n\}$ for $\phi = 0$, $\eta = 0.1$ and $\epsilon = 0.7$.

n	r_n	s_n	t_n	$r_{n+1} - r_n$	$s_{n+1} - s_n$	$t_{n+1} - t_n$
0	$0.000 \times 10^{+00}$	$0.000 \times 10^{+00}$	$0.000 \times 10^{+00}$	1.100×10^{-01}	1.100×10^{-01}	1.000×10^{-01}
1	1.100×10^{-01}	1.100×10^{-01}	1.000×10^{-01}	3.522×10^{-02}	3.789×10^{-02}	4.696×10^{-02}
2	1.452×10^{-01}	1.479×10^{-01}	1.470×10^{-01}	7.951×10^{-03}	8.853×10^{-03}	1.312×10^{-02}
3	1.532×10^{-01}	1.567×10^{-01}	1.601×10^{-01}	1.367×10^{-03}	1.560×10^{-03}	2.797×10^{-03}
4	1.545×10^{-01}	1.583×10^{-01}	1.629×10^{-01}	2.162×10^{-04}	2.507×10^{-04}	5.311×10^{-04}
5	1.547×10^{-01}	1.586×10^{-01}	1.634×10^{-01}	3.368×10^{-05}	3.961×10^{-05}	9.805×10^{-05}
6	1.548×10^{-01}	1.586×10^{-01}	1.635×10^{-01}	5.232×10^{-06}	6.238×10^{-06}	1.800×10^{-05}
7	1.548×10^{-01}	1.586×10^{-01}	1.635×10^{-01}	8.125×10^{-07}	9.822×10^{-07}	3.302×10^{-06}
8	1.548×10^{-01}	1.586×10^{-01}	1.635×10^{-01}	1.262×10^{-07}	1.546×10^{-07}	6.056×10^{-07}
9	1.548×10^{-01}	1.586×10^{-01}	1.635×10^{-01}	1.959×10^{-08}	2.434×10^{-08}	1.111×10^{-07}

Example 4.2. Consider the nonlinear two-point boundary value problem

$$x''(t) - \exp(x(t)) = 0, \quad t \in (0, 1); \quad x(0) = x(1) = 0. \quad (4.4)$$

The above boundary value problem arises in magnetohydrodynamics [1,2,4]. A unique solution of the above equation is

$$x(t) = -\ln(2) + 2 \ln\left(\frac{c}{\cos(c/2(t - 1/2))}\right), \quad t \in [0, 1]$$

where $c = 1.33605569490611 \dots$ is the root of the equation

$$c - \sqrt{2} \cos(c/4) = 0.$$

The boundary value problem (4.4) may be reformulated as the following nonlinear Hammerstein integral equation

$$x(t) = \int_0^1 g(t, s) \exp(x(s)) ds, \quad s, t \in [0, 1] \quad (4.5)$$

where the kernel

$$g(t, s) = \begin{cases} -s(1-t), & s \leq t, \\ -t(1-s), & s > t, \end{cases}$$

is the Green's function for the homogeneous problem

$$x''(t) = 0, \quad t \in (0, 1); \quad x(0) = x(1) = 0.$$

To solve the nonlinear integral equation (4.5), we divide the interval $(s, t \in [0, 1])$ into n -points and approximate the integral part through an n -point Gauss–Legendre quadrature. Let these n -points be ξ_i with $i = 1, 2, \dots, n$. Thus we obtain

$$x(\xi_j) = \int_0^1 g(\xi_j, s) \exp(x(s)) ds \approx \sum_{i=1}^n \omega_i G(\xi_j, \xi_i) \exp(x(\xi_i)) \quad (4.6)$$

where the nodes ξ_i and weights w_i are given as

$$\xi_i = \frac{1}{2}z_i + \frac{1}{2}, \quad \omega_i = \frac{2}{(1 - z_i^2)(\mathcal{P}'_n(z_i))^2}$$

where z_i (also known as i th Gauss-node) are the i th zeros of the normalized Legendre, i.e. $\mathcal{P}_n(1) = 1$, polynomial $\mathcal{P}_n(z)$

$$\mathcal{P}_n(z) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n].$$

From (4.6), we get the nonlinear-system $\mathcal{F}: \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$\mathcal{F}(\mathbf{x}) \equiv \mathbf{x} - \mathbf{A} \mathbf{v}_x = 0 \quad (4.7)$$

where

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T, \quad \mathbf{A} = [a_{ij}]_{i,j=1}^n, \quad \mathbf{v}_x = [\exp(x_1), \exp(x_2), \dots, \exp(x_n)]^T$$

where $a_{ij} = \omega_i g(\xi_j, \xi_i)$. Moreover, $\mathcal{F}'(\mathbf{x}) = \mathbf{I} - \mathbf{A} \mathbf{D}(\mathbf{x})$ where

$$\mathbf{D}(\mathbf{x}) = \text{diag}\{\exp(x_1), \exp(x_2), \dots, \exp(x_n)\}.$$

Table 3

Numerical solution of (4.5) – after 6 nonlinear iterations – at Gauss–Legendre points.

i	ξ_i	$x(\xi_i)$ (NM)	$x(\xi_i)$ (INM)	Exact solution
1	0.9965642995925470	−0.0015924739746710	−0.0015924739432730	−0.0015870037923855
2	0.9819859636389570	−0.0082214593992380	−0.0082214592345480	−0.0081900880182086
3	0.9561172141256630	−0.0194635089301890	−0.0194635085284890	−0.0193889850411605
4	0.9195584859111090	−0.0342276895763190	−0.0342276888386920	−0.0340980438789252
5	0.8731659532300750	−0.0510965252916960	−0.0510965241291770	−0.0509060219070535
6	0.8180268403632580	−0.0684755287362770	−0.0684755270812490	−0.0682246435843263
7	0.7554335009754140	−0.0847497147034950	−0.0847497125276870	−0.0844443701704358
8	0.6868530443577100	−0.0984318402491070	−0.0984318375843290	−0.0980822368132514
9	0.6138929255708230	−0.1082899854534830	−0.1082899824069280	−0.1079093417676064
10	0.5382632605667490	−0.1134468259055490	−0.1134468226569600	−0.1130502244868677
11	0.4617367394332510	−0.1134468259055490	−0.1134468226771700	−0.1130502244868677
12	0.3861070744291770	−0.1082899854534830	−0.1082899824607560	−0.1079093417676064
13	0.3131469556422900	−0.0984318402491070	−0.0984318376554380	−0.0980822368132514
14	0.2445664990245860	−0.0847497147034950	−0.0847497125984820	−0.0844443701704358
15	0.1819731596367420	−0.0684755287362770	−0.0684755271392680	−0.0682246435843263
16	0.1268340467699250	−0.0510965252916960	−0.0510965241695290	−0.0509060219070535
17	0.0804415140888910	−0.0342276895763190	−0.0342276888628560	−0.0340980438789252
18	0.0438827858743370	−0.0194635089301890	−0.0194635085408410	−0.0193889850411605
19	0.0180140363610430	−0.0082214593992380	−0.0082214592394210	−0.0081900880182086
20	0.0034357004074530	−0.0015924739746710	−0.0015924739441910	−0.0015870037923855

Table 4Performance of Inexact Newton Method and Newton Method. Here, n is the nonlinear iteration.

n	INM			Itr	NM		
	$\frac{\ x_n - x_{n-1}\ }{\ x_1 - x_0\ }$	$\frac{\ \mathcal{F}(x_n)\ }{\ \mathcal{F}(x_1)\ }$			$\frac{\ x_n - x_{n-1}\ }{\ x_1 - x_0\ }$	$\frac{\ \mathcal{F}(x_n)\ }{\ \mathcal{F}(x_1)\ }$	Itr
1	1.0000000000000000	1.0000000000000000	6		1.0000000000000000	1.0000000000000000	30
2	0.496366634997387	0.416139697284216	4		0.552304885261726	0.386703193410162	30
3	0.475899499741994	0.120660155092460	4		0.453416350986069	0.105578937478914	30
4	0.246304688610884	0.008818732595683	4		0.215113173639512	0.006173051813673	30
5	0.020572725608867	0.000031055906855	4		0.014012128224901	0.000014267100206	30
6	0.000072947179990	0.00000000369562	6		0.000032533964849	0.00000000073751	30

To solve the nonlinear integral equation (4.5), we divide the interval through a 20-point Gauss–Legendre quadrature rule which result in nonlinear system of 20 equations. And, to solve the nonlinear system – by the Inexact Newton Method and Newton Method – we employ the *Algorithm 1*. In the *Algorithm 1*, for the Inexact Newton Method the residual control z_n is selected as $\|\mathcal{F}(x_n)\| \times 10^{-n}$ while for Newton method the residual is fixed at 10^{-20} . Here, n represent nonlinear iteration number. We employ 6 nonlinear iterations of Inexact Newton Method and Newton Method. To solve the linear systems (step 1 in the *Algorithm 1*), Gauss–Seidel iterative procedure is employed. For both – Inexact Newton Method and Newton Method – the maximum allowed Gauss–Seidel iterations are 30.

Tables 3 and 4 report output of our numerical work. Solution – at the 20-Gauss–Legendre points – is reported in the Table 3 after 6 nonlinear iterations of Inexact Newton Method and Newton Method. Table 4 compares Inexact Newton Method and Newton Method. Here, n is the nonlinear iteration while Itr are the number of linear iterations (Gauss–Seidel iterations) required to solve the linear system (step 1 in *Algorithm 1*). In the Table 4, we note that after 6 nonlinear iterations Inexact Newton Method produces

$$\|x_6 - x_5\| \approx 7.3 \times 10^{-5} \|x_1 - x_0\| \quad \text{and} \quad \|\mathcal{F}(x_n)\| \approx 3.7 \times 10^{-10} \|\mathcal{F}(x_1)\|$$

while the Newton Method produces

$$\|x_6 - x_5\| \approx 3.3 \times 10^{-5} \|x_1 - x_0\| \quad \text{and} \quad \|\mathcal{F}(x_n)\| \approx 7.4 \times 10^{-11} \|\mathcal{F}(x_1)\|.$$

To achieve the above error, Inexact Newton Method uses 28 linear iterations while Newton Method uses 180 linear iterations.

5. Conclusions

In this work a semilocal convergence analysis of the Inexact Newton Method is developed for approximating a locally unique solution of a nonlinear equation in a Banach space setting. For evaluating the inverses of linear operators, we employ the concept of center-Lipschitz condition instead of the prevalent Lipschitz condition. Weaker sufficient convergence criteria are also provided which broaden the applicability of the method. Numerical examples illustrating the theoretical results are also provided in this study.

References

- [1] I.K. Argyros, in: C.K. Chui, L. Wuytack (Eds.), *Computational Theory of Iterative Methods*, in: Series: Studies in Computational Mathematics, vol. 15, Elsevier Publ. Co., New York, USA, 2007.
- [2] I.K. Argyros, Y.J. Cho, S. Hilout, *Numerical Methods for Equations and its Applications*, CRC Press/Taylor and Francis Group, NY, 2012.
- [3] L.V. Kantorovich, G.P. Akilov, *Functional Analysis*, Pergamon Press, Oxford, 1982.
- [4] L.M. Ortega, W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
- [5] P.D. Proinov, New general convergence theory for iterative processes and its applications to Newton–Kantorovich type theorems, *J. Complexity* 26 (2010) 3–42.
- [6] S. Smale, Newton's method estimates from data at one point, in: R. Ewing, K. Gross, C. Martin (Eds.), *The Merging of Disciplines: New Directions*, in: *Pure Appl. Comput. Math.*, Springer-Verlag, New York, 1986, pp. 185–196.
- [7] S. Amat, S. Busquier, M. Negra, Adaptive approximation of nonlinear operators, *Numer. Funct. Anal. Optim.* 25 (2004) 397–405.
- [8] I.K. Argyros, Relation between forcing sequences and inexact Newton iterates in Banach space, *Computing* 63 (1999) 134–144.
- [9] I.K. Argyros, A new convergence theorem for the inexact Newton methods based on assumptions involving the second Frechét derivative, *Comput. Math. Appl.* 37 (1999) 109–115.
- [10] I.K. Argyros, A semilocal convergence analysis for directional Newton methods, *Math. Comp.* 80 (2011) 327–343.
- [11] I.K. Argyros, S. Hilout, Extending the Newton–Kantorovich hypothesis for solving equations, *J. Comput. Appl. Math.* 234 (2010) 2993–3006.
- [12] I.K. Argyros, S. Hilout, Weaker conditions for the convergence of Newton's method, *J. Complexity* 28 (2012) 364–387.
- [13] R.S. Dembo, S.C. Eisenstant, T. Steihaug, Inexact Newton methods, *SIAM J. Numer. Anal.* 19 (1982) 400–408.
- [14] J.A. Ezquerro, J.M. Gutiérrez, M.A. Hernández, N. Romero, M.J. Rubio, The Newton method: from Newton to Kantorovich, *Gac. R. Soc. Mat. Esp.* 13 (2010) 53–76 (in Spanish).
- [15] X.P. Guo, On semilocal convergence of inexact Newton methods, *J. Comput. Math.* 25 (2007) 231–242.
- [16] B. Morini, Convergence behaviour of inexact Newton methods, *Math. Comp.* 68 (1999) 1605–1613.
- [17] W.P. Shen, C. Li, Kantorovich-type convergence criterion for inexact Newton methods, *Appl. Numer. Math.* 59 (2009) 1599–1611.
- [18] W.P. Shen, C. Li, Smale's α -theory for inexact Newton methods under the γ -conditions, *J. Math. Anal. Appl.* 369 (2010) 29–42.
- [19] R.S. Dembo, T. Steihaug, Truncated-newton algorithms for large-scale unconstrained optimization, *Math. Program.* 26 (1983) 190–212.
- [20] T. Steihaug, *Quasi-Newton methods for large scale nonlinear problems*, Ph.D. Dissertation, Yale University, New Haven, CT, 1980.
- [21] X.H. Wang, Convergence of Newton's method and inverse function theorem in Banach space, *Math. Comp.* 68 (1999) 169–186.
- [22] T.J. Ypma, Local convergence of inexact Newton methods, *SIAM J. Numer. Anal.* 21 (1984) 583–590.