Accepted Manuscript

Kantorovich-Type Convergence Criterion for Inexact Newton Methods

Weiping Shen, Chong Li

PII: S0168-9274(08)00195-5
DOI: 10.1016/j.apnum.2008.11.002
Reference: APNUM 2199

To appear in: Applied Numerical Mathematics

Received date: 2 September 2008
Revised date: 23 October 2008
Accepted date: 4 November 2008

Please cite this article as: W. Shen, C. Li, Kantorovich-Type Convergence Criterion for Inexact Newton Methods, Applied Numerical Mathematics (2008), doi: 10.1016/j.apnum.2008.11.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Kantorovich-Type Convergence Criterion for Inexact Newton Methods

Weiping Shen† and Chong Li‡
† Department of Mathematics, Zhejiang Normal University
Jinhua, 321004, P R China
‡ Department of Mathematics, Zhejiang University
Hangzhou, 310027, P R China

Abstract: Assuming that the first derivative of an operator satisfies the Lipschitz condition, a Kantorovich-type convergence criterion for inexact Newton methods is established, which includes the well-known Kantorovich’s theorem as a special case. Comparisons and a numerical example are presented to illustrate that our results obtained in the present paper improve and extend some recent results in [10, 15].

MSC: 65H10; 65J15; 47H30

Keywords: Nonlinear equation; inexact Newton methods; Lipschitz condition; Kantorovich-type theorem;

1 Introduction

Let $X$ and $Y$ be (real or complex) Banach spaces, $D \subseteq X$ be an open subset and let $F : D \subseteq X \rightarrow Y$ be a nonlinear operator with the continuous Fréchet derivative denoted by $F'$. Finding solutions of the nonlinear operator equation

$$F(x) = 0$$

in Banach spaces is a very general subject which is widely used in both theoretical and applied areas of mathematics. The most important method to find an approximation of a solution of (1.1) is Newton’s method which takes the following form:

$$x_{n+1} = x_n - F'(x_n)^{-1}F(x_n), \quad n = 0, 1, 2, \cdots .$$

*This project was supported in part by the National Natural Science Foundations of China (Grant No. 10671175; 10731060).
†email: shenweiping@zjnu.cn
‡email: cli@zju.edu.cn
2 Convergence Criterion for Inexact Newton Methods

The convergence issue of Newton’s method has been studied extensively, see for example [5, 7, 8, 11, 16, 17, 18, 21, 22, 23, 24]. These results can be distinguished into two classes. One is about local convergence discussing the properties of Newton’s method through the information depending on the solution \( x^* \) of (1.1) (cf. [22, 23, 24]), and the other is about semilocal convergence which only deals with the initial point \( x_0 \) (cf. [5, 7, 8, 11, 16, 17, 18, 21]). Among the semilocal convergence results on Newton’s method, one of the famous results is the well-known Kantorovich’s theorem (cf. [11]) which, under the very mild condition that the first derivative \( F' \) is Lipschitz continuous, provides the following convergence criterion of Newton’s method:

\[
0 < \alpha \gamma \leq \frac{1}{2}
\]

where \( \gamma \) is the Lipschitz constant and \( \alpha \geq \| F'(x_0)^{-1} F'(x_0) \| \).

As expressed in (1.2), we have to solve the following Newton equation at each Newton step:

\[
F'(x_n)s_n = -F(x_n).
\]

(1.4)

This sometimes makes Newton’s method inefficient from the point of view of practical calculations especially when \( F'(x_n) \) is large and sparse. While using linear iterative methods to approximate the solution of (1.4) instead of solving it exactly can reduce some of the costs of Newton’s method which was studied extensively and applied in [1, 2, 3, 6, 10, 14, 19, 25] (such a variant is the so-called inexact Newton method). In general, the inexact Newton method has the following general form:

**Algorithm A** \([r_n; x_0] \)

For \( n = 0 \) and a given initial guess \( x_0 \) until convergence do

1. For the residual \( r_n \) and the iteration \( x_n \), find the step \( s_n \) satisfying

\[
F'(x_n)s_n = -F(x_n) + r_n.
\]

(1.5)

2. \( x_{n+1} = x_n + s_n \).

3. set \( n = n + 1 \) and turn to step 1.

Here \( \{r_n\} \) is a sequence of elements in \( Y \) (depending on \( \{x_n\} \) in general).

As is well known, the convergence behavior of the inexact Newton method depends on the residual controls of \( \{r_n\} \). Several authors (cf. [6, 19]) have analyzed the local convergence behavior in some manner such that the stopping relative residuals \( \{r_n\} \) satisfy \( \|r_n\| \leq \eta_n \| F(x_n) \| \). Ypma used in [25] the affine invariant condition \( \| F'(x_n)^{-1} r_n \| \leq \eta_n \| F'(x_n)^{-1} F(x_n) \| \), which makes the method become an affine invariant one, and analyzed also the local convergence property of the inexact Newton method.
The above results concern mainly with the local convergence of the inexact Newton method. In the spirit of Kantorovich’s theorem, the semi-local convergence analysis for the inexact Newton method was studied recently, see [1, 2, 10] for example. As in the case of the local convergence analysis, different residual controls were used. For example, the residual controls \( \| r_n \| \leq \eta_n \| F(x_n) \| \) was adopted in [2]; while in [1], Argyros considered the residual controls \( \| r_n \| \leq d_n h_n c_n \) where \( \{ d_n \}, \{ h_n \} \) and \( \{ c_n \} \) are real sequences satisfying some suitable conditions.

In particular, Guo considered in her recent paper [10] the residual controls \( \| F'(x_0)^{-1} r_n \| \leq \eta_n \| F'(x_0)^{-1} F(x_n) \| \) for each \( n \in \mathbb{N} \),

\[
(1.6)
\]
and, under the Lipschitz continuity assumption on \( F' \), established the following convergence criterion for the inexact Newton method (cf. [10, Theorem 2.6]):

\[
\alpha \gamma \leq \frac{\sqrt{(4\eta + 5)^2 - 2\eta^2 - 14\eta - 11}}{(1 + \eta)(1 - \eta)^2},
\]

\[
(1.7)
\]
where \( \alpha, \gamma \) are the same as in (1.3) and \( \eta := \sup_{n \geq 0} \eta_n \). However, the criterion (1.7) is not satisfactory; in particular, the convergence criterion in the case when \( \eta_n \equiv 0 \) is much worse than the criterion (1.3).

Motivated by the ideas of the inexact Newton-like method for the inverse eigenvalue problem (cf. [4]), the authors of the present paper presented in [12] the following residual controls:

\[
\| P_n r_n \| \leq \eta_n \| P_n F(x_n) \|^{1+\beta}
\]

for each \( n \in \mathbb{N} \),

\[
(1.8)
\]
where \( \{ P_n \} \) is a sequence of invertible operators from \( Y \) to \( X \) and \( 0 \leq \beta \leq 1 \), and established the local convergence of order \( 1+\beta \) for inexact Newton-like methods. A semi-local convergence analysis for inexact Newton-like methods under the \((L, p)\)-Hölder condition on \( F'(x_0)^{-1} F' \) was done in [15] for the residual controls (1.8) with \( \beta > 0 \). Obviously, in the special case when \( P_n = F'(x_0)^{-1} \), (1.8) reduces to

\[
\| F'(x_0)^{-1} r_n \| \leq \eta_n \| F'(x_0)^{-1} F(x_n) \|^{1+\beta}
\]

for each \( n \in \mathbb{N} \).

In the present paper, by considering the residual controls (1.9), we continue our study on the convergence criterion of the inexact Newton method. Under the assumption of the Lipschitz continuity of \( F'(x_0)^{-1} F' \), we establish a Kantorovich-type convergence criterion for the inexact Newton method Algorithm A\( [r_n; x_0] \). Note that in the special case when \( \beta = 0 \), (1.9) reduces to Guo’s controls (1.6) and our result extends the corresponding one in [10]. In particular, we get the following convergence criterion:

\[
\alpha \gamma \leq \frac{(1 - \eta)^2}{(1 + \eta)(2(1 + \eta) - \eta(1 - \eta)^2)},
\]

\[
(1.10)
\]
which is better than (1.7) for \( \eta \in [0, \frac{1}{2}] \) as showed in the last section. For the case when \( \beta = 1 \), our result in the present paper is better than the corresponding one in paper [15]. Moreover,
for the case when $0 < \beta < 1$, our main result is also sharper than the corresponding one in paper [15] for the case when $\gamma \geq 1$. It is worth to be remarked that our results obtained in the present paper includes Kantorovich’s theorem as a special case. Our approach is based on the construction of the majorizing function, which is completely different from that used in [10].

2 Preliminaries

Let $X$ and $Y$ be Banach spaces. Throughout the whole paper, we use $B(x, r)$ to stand for the open ball in $X$ with center $x$ and radius $r > 0$, and assume that $0 \leq \beta \leq 1$. Let $\sigma$, $\lambda$ and $\theta$ be positive constants. We define two real-valued functions $\varphi_\beta$ and $\psi_\beta$ respectively by

\[
\varphi_\beta(t) = \frac{\sigma}{2} t^2 + 2^{1-\beta} \theta t^{1+\beta} - (1 + \theta) t + \lambda \quad \text{for each} \quad t \geq 0
\]

and

\[
\psi_\beta(t) = \frac{\sigma}{2} t^2 + \theta t^{1+\beta} - (1 + \theta) t + \lambda \quad \text{for each} \quad t \geq 0.
\]

The derivatives of $\varphi_\beta$ and $\psi_\beta$ are respectively

\[
\varphi'_\beta(t) = \sigma t + 2^{1-\beta}(1 + \beta) \theta t^\beta - (1 + \theta) \quad \text{for each} \quad t \geq 0
\]

and

\[
\psi'_\beta(t) = \sigma t + (1 + \beta) \theta t^\beta - (1 + \theta) \quad \text{for each} \quad t \geq 0.
\]

Hence

\[
-\psi'_\beta(t) \geq -\varphi'_\beta(t) \quad \text{for each} \quad t \geq 0.
\]

Furthermore, we assume that $\theta < 1$ in the case when $\beta = 0$. By (2.3), the derivative $\varphi'_\beta$ is strictly increasing and has the values $\varphi'_\beta(0) < 0$ and $\varphi'_\beta(+\infty) = +\infty$. It follows that the equation $\varphi'_\beta(t) = 0$ has a unique positive solution. In the remainder, we denote the solution by $r_*$, that is, $r_*$ satisfies

\[
\varphi'_\beta(r_*) = \sigma r_* + 2^{1-\beta}(1 + \beta) \theta r_*^\beta - (1 + \theta) = 0.
\]

Write

\[
b := \frac{\sigma}{2} r_*^2 + \beta 2^{1-\beta} \theta r_*^{1+\beta}.
\]

Thus, we are ready to prove the following lemma which describes some useful properties of the function $\varphi_\beta$.

**Lemma 2.1.** Let $\varphi_\beta$ be defined by (2.1). Then $\varphi_\beta$ is strictly decreasing on $[0, r_*]$ and satisfies that

\[
\varphi_\beta(r_*) = \lambda - b.
\]

Moreover, if

\[
\lambda \leq b,
\]

then $\varphi_\beta$ has a unique zero $t^*$ in $[0, r_*]$. 


Proof. By (2.3) and (2.6), $\varphi'_\beta$ is negative on $[0, r_*)$; hence $\varphi_\beta$ is strictly decreasing on $[0, r_*)$.

Since, by (2.1),

$$\varphi_\beta(r_*) = \frac{\sigma}{2} r_*^2 + 2^{1-\beta} \theta r_*^{1+\beta} - (1 + \theta) r_* + \lambda$$

it follows from (2.6) and (2.7) that (2.8) holds.

Now assume that (2.9) holds. Then, $\varphi_\beta(r_*) \leq 0$ by (2.8). Note that $\varphi_\beta$ is strictly decreasing on $[0, r_*)$ and that $\varphi_\beta(0) > 0$. Therefore, $\varphi_\beta$ has a unique zero $t^*$ in $[0, r_*)$, and the proof is complete. \qed

Let $t_0 = 0$ and let $\{t_n\}$ denote the sequence generated by

$$t_{n+1} = t_n - \frac{\varphi_\beta(t_n)}{\psi'_\beta(t_n)} \quad \text{for each } n = 0, 1, \ldots \tag{2.11}$$

The convergence property of the sequence $\{t_n\}$, which will play a key role, is described in the following lemma.

**Lemma 2.2.** Suppose that (2.9) holds. Let $t^*$ be the unique zero of $\varphi_\beta$ in $[0, r_*)$ and let $\{t_n\}$ be the sequence generated by (2.11). Then

$$t_n < t_{n+1} < t^* \quad \text{for each } n \in \mathbb{N} \tag{2.12}$$

Consequently, $\{t_n\}$ converges increasingly to $t^*$. Moreover, in the special case when $\beta > 0$,

$$t_{n+1} - t_n \leq \frac{\lambda}{1+\theta} \quad \text{for each } n \in \mathbb{N} \tag{2.13}$$

**Proof.** We will verify (2.12) by mathematical induction. To prove (2.12) holds for $n = 0$, we note that

$$t_1 = \begin{cases} \lambda, & \beta = 0; \\ \frac{\lambda}{1+\theta}, & \beta > 0. \end{cases} \tag{2.14}$$

Then $t_1 \leq r_*$. In fact, if $\beta > 0$, then by (2.7) and (2.6)

$$b = \sigma r_*^2 + \beta^2 2^{1-\beta} \theta r_*^{1+\beta} \leq (\sigma r_* + 2^{1-\beta} (1 + \beta) \theta r_*^3) r_* = (1 + \theta) r_*; \tag{2.15}$$

hence $\frac{b}{1+\theta} \leq r_*$ and so $t_1 = \frac{\lambda}{1+\theta} \leq \frac{b}{1+\theta} \leq r_*$ by (2.9). If $\beta = 0$, then by (2.7) and (2.6)

$$b = \sigma r_*^2 \leq \sigma r_*^2 = (1-\theta) r_* \leq r_*; \tag{2.16}$$

hence $t_1 = \lambda \leq b \leq r_*$ by (2.9).
Furthermore, since
\[
\varphi_\beta(t_1) = \begin{cases} 
\frac{\sigma}{2} \lambda^2 + \theta \lambda > 0, \\
\frac{\sigma}{2} \left( \frac{\lambda}{1+\theta} \right)^2 + 2^{1-\beta} \theta \left( \frac{\lambda}{1+\theta} \right)^{1+\beta} > 0, 
\end{cases} \beta = 0; \\
\beta > 0,
\]
(2.12) holds for \( n = 0 \) because \( \varphi_\beta \) is strictly decreasing on \([0, r_*]\) and \( \varphi_\beta(t^*) = 0 \).

Now assume that \( t_1 < t_2 < \cdots < t_n < t^* \). Then,
\[
\varphi_\beta(t_n) > 0 \tag{2.17}
\]
and \( t_n < r_* \). Thus, thanks to (2.5), (2.6) and the fact that \( \varphi'_\beta \) is strictly increasing, one has
\[
-\psi'_\beta(t_n) \geq -\varphi'_\beta(t_n) > -\varphi'_\beta(r_*) = 0. \tag{2.18}
\]

This and (2.17) imply that
\[
t_{n+1} = t_n - \frac{\varphi_\beta(t_n)}{\psi'_\beta(t_n)} \geq t_n.
\]

On the other hand, by (2.5) and (2.17), \( t_{n+1} \leq t_n - \frac{\varphi_\beta(t_n)}{\psi'_\beta(t_n)} \). Since the function \( t \mapsto t - \frac{\varphi_\beta(t)}{\psi'_\beta(t)} \) is monotonically increasing on \([0, t^*]\), it follows that
\[
t_{n+1} \leq t_n - \frac{\varphi_\beta(t_n)}{\psi'_\beta(t_n)} < t^* - \frac{\varphi_\beta(t^*)}{\psi'_\beta(t^*)} = t^*.
\]

Hence, (2.12) holds for each \( n \geq 0 \). This means that \( \{t_n\} \) converges increasingly to a point, say \( \varsigma \). Clearly \( \varsigma \in [0, t^*] \) and \( \varsigma \) is a zero of \( \varphi_\beta \) (cf. (2.11)). Noting that \( t^* \) is the unique zero of \( \varphi_\beta \) in \([0, r_*]\), one has that \( \varsigma = t^* \) and so \( \{t_n\} \) converges increasingly to \( t^* \).

Finally, suppose that \( \beta > 0 \). By (2.5) and (2.17),
\[
t_{n+1} - t_n = \frac{\varphi_\beta(t_n)}{\psi'_\beta(t_n)} \leq \frac{\varphi_\beta(t_n)}{\varphi'_\beta(t_n)} \leq \frac{\varphi_\beta(t_0)}{\varphi'_\beta(t_0)} \tag{2.19}
\]
for each \( n = 0, 1, \cdots \),

where the last inequality holds because the function \( M \) defined by \( M(t) := -\frac{\varphi_\beta(t)}{\varphi'_\beta(t)} \) for each \( 0 \leq t \leq t^* \) is monotonically decreasing. Noting that \( -\frac{\varphi_\beta(t_0)}{\varphi'_\beta(t_0)} = \frac{\lambda}{1+\theta} \), we see that (2.13) holds by (2.19). This completes the proof. \( \square \)

3 Kantorovich-type convergence criterion

Recall that \( F: D \to Y \) is an operator with the continuous Fréchet derivative denoted by \( F' \). Let \( x_0 \in D \) be such that the inverse \( F'(x_0)^{-1} \) exists. We say that \( F'(x_0)^{-1}F' \) satisfies Lipschitz condition on \( B(x_0, r) \) with the Lipschitz constant \( \gamma \) if
\[
\|F'(x_0)^{-1}(F'(x) - F'(y))\| \leq \gamma \|x - y\| \quad \text{for all} \ x, y \in B(x_0, r). \tag{3.1}
\]
Assume that \( B(x_0, \frac{1}{\gamma}) \subseteq D \) throughout the whole paper. The following lemma can be proved by Banach’s Lemma with a standard argument, see for example [21].

**Lemma 3.1.** Let \( r \leq \frac{1}{\gamma} \) and let \( x \in B(x_0, r) \). Then \( F'(x) \) is invertible and satisfies that

\[
\| F'(x)^{-1}F'(x_0) \| \leq \frac{1}{1 - \gamma \|x - x_0\|}. \tag{3.2}
\]

For the remainder of the present paper, we assume that the residuals \( \{r_n\} \) satisfy (1.9) and that \( \eta := \sup_{n \geq 0} \eta_n < 1 \). Let

\[
\alpha := \| F'(x_0)^{-1}F(x_0) \|. \tag{3.3}
\]

Without loss of generality, we may assume throughout that \( x_0 \) is not a zero of \( F \). This means that \( \alpha > 0 \). Write

\[
\sigma := \frac{\gamma(1 + \eta^{-1})}{(1 + \alpha \gamma \eta(1 + \eta))(1 + \eta(\alpha^2 + 1))}, \quad \theta := \frac{\eta}{1 + \eta(\alpha^2 - 1)} \tag{3.4}
\]

and

\[
\lambda := \begin{cases} 
\alpha(1 + \eta), & \beta = 0; \\
\alpha \eta^{-1}(1 + \theta), & \beta > 0.
\end{cases} \tag{3.5}
\]

Recall that \( r_* \) is the unique positive zero of \( \varphi'_\beta \) (that is, \( r_* \) satisfies (2.6)) and that \( \{t_n\} \) denotes the sequence generated by (2.11) with \( \sigma, \theta \) and \( \lambda \) given by (3.4) and (3.5). The main theorem now can be stated as follows.

**Theorem 3.1.** Let \( \sigma \) and \( \theta \) be given by (3.4). Suppose that \( F'(x_0)^{-1}F' \) satisfies the Lipschitz condition on \( B(x_0, t^*) \) with the Lipschitz constant \( \gamma \). Suppose that

\[
\alpha \leq \begin{cases} 
\frac{(1 - \eta)^2}{\gamma(1 + \eta)(2 + \eta) - \eta(1 - \eta)^2}, & \beta = 0; \\
\min \left\{ \frac{\sigma t^2_* + 2(1 - \beta) \theta r_1 \eta^{2\beta}}{(1 + \beta)(1 + \eta^{2\beta})}, \frac{\eta^{-\beta}}{1 + \theta} \right\}, & \beta > 0.
\end{cases} \tag{3.6}
\]

Let \( \{x_n\} \) be the sequence generated by Algorithm A[\( r_n; x_0 \)]. Then \( \{x_n\} \) converges to a solution \( x^* \) of (1.1) and the following assertion holds:

\[
\|x_n - x^*\| \leq t^* - t_n \quad \text{for each } n \in \mathbb{N}. \tag{3.7}
\]

**Proof.** Let \( b \) be defined by (2.7). We assert that

condition (3.6) holds \( \iff [\lambda \leq b \text{ and } \alpha^\beta \eta^{\frac{\beta}{2 + \beta^2}} \leq 1] \). \tag{3.8}
In the case when $\beta > 0$, (3.8) is a direct consequence of the definitions of $b$ and $\lambda$ respectively given in (2.7) and (3.5). In the case when $\beta = 0$, (2.6) and (2.7) reduce to

$$\sigma r_s = 1 - \theta \quad \text{and} \quad b = \frac{\sigma}{2} r_s^2.$$  

Furthermore,

$$\sigma = \frac{\gamma (1 + \eta)}{1 + \alpha \gamma \eta (1 + \eta)}, \quad \theta = \eta \quad \text{and} \quad \lambda = \alpha (1 + \eta).$$

It follows that

$$b = \frac{\alpha}{2} r_s^2 = \frac{(1 - \theta)^2}{2\sigma} = \frac{(1 - \eta)^2 (1 + \alpha \gamma \eta (1 + \eta))}{2\gamma (1 + \eta)}.$$  

Consequently,

$$\alpha (1 + \eta) \leq \frac{(1 - \eta)^2 (1 + \alpha \gamma \eta (1 + \eta))}{2\gamma (1 + \eta)} \iff \lambda \leq b. \quad (3.9)$$

This is equivalent to the assertion (3.8) because $\alpha \beta^\eta \eta^{1-\eta} \leq 1$ holds automatically for $\beta = 0$. Thus (3.8) is proved and Lemma 2.2 is applicable to concluding that (2.12) and (2.13) hold. To complete the proof, it is sufficient to verify that

$$\| F'(x_0)^{-1} F(x_n) \| \leq t_{n+1} - t_n \quad (3.10)$$

and

$$\| x_{n+1} - x_n \| \leq t_{n+1} - t_n \quad (3.11)$$

hold for each $n \in \mathbb{N}$. In fact, granting this, one has from (3.11) that

$$\| x_{n+m} - x_n \| \leq \sum_{i=1}^{m} \| x_{n+i} - x_{n+i-1} \| \leq \sum_{i=1}^{m} (t_{n+i} - t_{n+i-1}) = t_{n+m} - t_n. \quad (3.12)$$

Therefore, $\{ x_n \}$ is a Cauchy sequence and (3.7) holds by letting $m \to \infty$ in (3.12). Moreover, letting $n \to \infty$ in (3.10) shows that the limit $x^*$ is a solution of (1.1).

Below, we will prove (3.10) and (3.11) by mathematical induction. Note first that, by (1.9), if $n \in \mathbb{N}$ is such that $x_n$ is well-defined, then

$$\| F'(x_0)^{-1} r_n \| \leq \eta \| F'(x_0)^{-1} F(x_n) \|^{1+\beta} \leq \eta \| F'(x_0)^{-1} F(x_n) \|^{1+\beta}. \quad (3.13)$$

Since $t_1 = \alpha (1 + \eta^{1+\beta})$, (3.10) is clear for $n = 0$ by the definition of $\alpha$. By (3.13),

$$\| x_1 - x_0 \| = \| F'(x_0)^{-1} F(x_0) + F'(x_0)^{-1} r_0 \| \leq \alpha + \eta \alpha^{1+\beta}. \quad (3.14)$$

Using (3.6), one has that $\eta \alpha^{\beta} \leq \eta^{1+\beta}$. This together with (3.14) yields that

$$\| x_1 - x_0 \| \leq \alpha + \eta \alpha^{1+\beta} \leq \alpha (1 + \eta^{1+\beta}) = t_1 - t_0.$$
This shows that (3.11) holds for $n = 0$. Assume now that (3.10) and (3.11) hold for all $n \leq m - 1$. We have to prove (3.10) and (3.11) hold for $n = m$. For this end, we apply Algorithm A[$r_n; x_0$] to get that

$$F(x_m) = F(x_m) - F(x_{m-1}) - F'(x_{m-1})(x_m - x_{m-1}) + r_{m-1}$$

$$= \int_0^1 F'(x_{m-1}) = F'(x_{m-1})d\tau(x_m - x_{m-1}) - F'(x_{m-1})(x_m - x_{m-1}) + r_{m-1},$$

where $x_m = x_n + \tau(x_{n+1} - x_n)$ for each $0 \leq \tau \leq 1$. Hence,

$$\|F'(x_0)^{-1}F(x_m)\| \leq \left\| F'(x_0)^{-1}\int_0^1 \left( F'(x_{m-1}) - F'(x_m) \right) (x_m - x_{m-1})d\tau(x_{m-1}) \right\| + \|F'(x_0)^{-1}r_{m-1}\|. \quad (3.15)$$

Below we will show that

$$\left\| F'(x_0)^{-1}\int_0^1 \left( F'(x_{m-1}) - F'(x_m) \right) (x_m - x_{m-1})d\tau(x_{m-1}) \right\| \leq \frac{\gamma}{2}(t_m - t_{m-1})^2 \quad (3.16)$$

and

$$\|F'(x_0)^{-1}r_{m-1}\| \leq \eta(1 + \eta^{1+\beta})^{-1-\beta}(t_m - t_{m-1})^{1+\beta}. \quad (3.17)$$

Since (3.11) holds for all $n \leq m - 1$, one has that

$$\|x_{m-1}^\tau - x_0\| \leq \sum_{i=0}^{m-2} \|x_{i+1} - x_i\| + \tau\|x_m - x_{m-1}\| \leq \sum_{i=0}^{m-2} (t_{i+1} - t_i) + \tau(t_m - t_{m-1}). \quad (3.18)$$

Hence, by (2.12),

$$\|x_{m-1}^\tau - x_0\| \leq \tau t_m + (1 - \tau)t_{m-1} < t^*. \quad (3.19)$$

In particular,

$$\|x_m - x_0\| \leq t_m < t^* \quad \text{and} \quad \|x_{m-1} - x_0\| \leq t_{m-1} < t^*. \quad (3.20)$$

Thus, applying the Lipschitz condition, we get that

$$\left\| F'(x_0)^{-1}\int_0^1 \left( F'(x_{m-1}) - F'(x_m) \right) (x_m - x_{m-1})d\tau \right\|$$

$$\leq \int_0^1 \gamma \|x_{m-1}^\tau - x_{m-1}\|\|x_m - x_{m-1}\|d\tau$$

$$= \frac{\gamma}{2}\|x_m - x_{m-1}\|^2$$

$$\leq \frac{\gamma}{2}(t_m - t_{m-1})^2, \quad (3.21)$$

where the last inequality holds because of (3.11) (with $n = m - 1$). This completes the proof of (3.16). To show (3.17), we note by (3.10) (with $n = m - 1$) that

$$\|F'(x_0)^{-1}F(x_m)\| \leq (1 + \eta^{1+\beta})^{-1}(t_m - t_{m-1}).$$
Thus, we have by (3.13) that
\[ \|F'(x_0)^{-1}r_{m-1}\| \leq \eta \|F'(x_0)^{-1}F(x_{m-1})\|^{1+\beta} \leq \eta(1 + \eta^{1+\beta})^{-1-\beta}(t_m - t_{m-1})^{1+\beta}. \] (3.22)

Therefore, (3.17) is proved. Combining (3.15)-(3.17), one has that
\[ \|F'(x_0)^{-1}F(x_m)\| \leq \frac{\gamma}{2}(t_m - t_{m-1})^2 + \eta(1 + \eta^{1+\beta})^{-1-\beta}(t_m - t_{m-1})^{1+\beta}. \] (3.23)

Hence,
\[ \frac{1 + \eta^{1+\beta}}{1 - \gamma t_m} \|F'(x_0)^{-1}F(x_m)\| \leq \frac{\gamma}{2}(t_m - t_{m-1})^2 + \frac{\eta(1 + \eta^{1+\beta})^{-\beta}}{1 - \gamma t_m}(t_m - t_{m-1})^{1+\beta}. \] (3.24)

We claim that
\[ \sigma(1 - \gamma t) \geq -\psi'_\beta(t)\gamma(1 + \eta^{1+\beta}) \geq 0 \quad \text{for each } t \in [t_1, r_*]. \] (3.25)

In fact, let \( t \in [t_1, r_*] \). Then, by the definition of \( \lambda \),
\[ \alpha(1 + \eta^{1+\beta}) = t_1 \leq t. \] (3.26)

Since \( \eta \leq \eta^{1+\beta} \) (noting \( \eta < 1 \)), it follows that
\[ 1 + \alpha \gamma \eta(1 + \eta) \leq 1 + \alpha \gamma \eta^{1+\beta}(1 + \eta^{1+\beta}) \leq 1 + \gamma \eta^{1+\beta} t. \] (3.27)

Noting that \( \theta = \frac{\eta}{1 + \eta(\alpha^{\beta}-1)} \) and thanks to (3.26), one has that
\[ \frac{1}{1 + \eta(\alpha^{\beta}-1)} = 1 + \theta(1 - \alpha^{\beta}) \geq 1 + \theta - (1 + \beta)\theta\alpha^{\beta}(1 + \eta^{1+\beta})^{\beta} \geq 1 + \theta - (1 + \beta)\theta^{\beta}. \] (3.28)

In view of (2.4), \( \psi'_\beta(t) = \sigma t + (1 + \beta)\theta^{\beta} - (1 + \theta) \). This and (3.28) imply that
\[ \frac{1}{1 + \eta(\alpha^{\beta}-1)} \geq -\psi'_\beta(t) + \sigma t. \] (3.29)

Noting that \( \sigma = \frac{\gamma(1 + \eta^{1+\beta})}{(1 + \alpha \gamma \eta(1 + \eta))(1 + \eta(\alpha^{\beta}-1))} \), we use (3.28) and (3.27) to get that
\[ \sigma \geq \frac{\gamma(1 + \eta^{1+\beta})(-\psi'_\beta(t) + \sigma t)}{1 + \gamma \eta^{1+\beta} t} = \frac{\gamma(1 + \eta^{1+\beta})(-\psi'_\beta(t) + \sigma t)}{1 - \gamma t + \gamma(1 + \eta^{1+\beta}) t}, \] (3.30)

which is equivalent that
\[ \sigma(1 - \gamma t) \geq -\psi'_\beta(t)\gamma(1 + \eta^{1+\beta}). \]
Since \(-\psi'_\beta(t) \geq -\varphi'_\beta(t) \geq -\varphi'_\beta(r_*) = 0\), one sees that (3.25) is proved and the claim stands. In particular, (3.25) implies that \(\sigma(1 - \gamma t_m) \geq -\psi'_\beta(t_m)\gamma(1 + \eta^{1+\beta})\) and so

\[
\frac{\gamma(1 + \eta^{1+\beta})}{1 - \gamma t_m} \leq \frac{\sigma}{-\psi'_\beta(t_m)}, \tag{3.31}
\]

because \(t_m \in (t_1, r_*)\) and \(-\psi'_\beta(t_m) > 0\). Consequently, in view of the definitions of \(\sigma\) and \(\theta\),

\[
\frac{\eta(1 + \eta^{1+\beta})^{-\beta}}{1 - \gamma t_m} \leq \frac{\eta(1 + \eta^{1+\beta})^{-\beta\sigma}}{\psi'_\beta(t_m)\gamma(1 + \eta^{1+\beta})} = \frac{\eta(1 + \eta^{1+\beta})^{-\beta}}{\psi'_\beta(t_m)(1 + \alpha \gamma (1 + \eta))(1 + \eta(\alpha^2 - 1))} \leq \frac{\theta(t_m - t_{m-1})^{1+\beta}}{\psi'_\beta(t_m)}. \tag{3.32}
\]

This together with (3.31) and (3.24) implies

\[
\frac{1 + \eta^{1+\beta}}{1 - \gamma t_m} \| F'(x_0)^{-1} F(x_m) \| \leq \frac{\sigma(t_m - t_{m-1})^2}{-2\psi'_\beta(t_m)} + \frac{\theta(t_m - t_{m-1})^{1+\beta}}{-\psi'_\beta(t_m)}. \tag{3.33}
\]

Noting the following well-known inequality

\[
t^{1+\beta} + (1 + \beta)t \leq 2^{1-\beta}((1 + t)^{1+\beta} - 1) \quad \text{for each} \quad t \geq 0,
\]

we have

\[
\left(\frac{t_m - t_{m-1}}{t_{m-1}}\right)^{1+\beta} + (1 + \beta)\frac{t_m - t_{m-1}}{t_{m-1}} \leq 2^{1-\beta}\left(\left(1 + \frac{t_m - t_{m-1}}{t_{m-1}}\right)^{1+\beta} - 1\right). \tag{3.34}
\]

It follows that

\[
(t_m - t_{m-1})^{1+\beta} \leq 2^{1-\beta}t_m^{1+\beta} - 2^{1-\beta}t_{m-1}^{1+\beta} - (1 + \beta)t_{m-1}^{1+\beta}(t_m - t_{m-1}). \tag{3.35}
\]

Combining this with (3.33), one has that

\[
\frac{1 + \eta^{1+\beta}}{1 - \gamma t_m} \| F'(x_0)^{-1} F(x_m) \| \leq \frac{1}{\psi'_\beta(t_m)} \left(\frac{\sigma}{2}(t_m - t_{m-1})^2 + 2^{1-\beta}\theta t_m^{1+\beta} - 2^{1-\beta}\theta t_{m-1}^{1+\beta} - (1 + \beta)\theta t_{m-1}^{1+\beta}(t_m - t_{m-1})\right) = \frac{\varphi_\beta(t_m) - \varphi_\beta(t_{m-1}) - \psi'_\beta(t_m)(t_m - t_{m-1})}{\psi'_\beta(t_m)} = t_{m+1} - t_m; \tag{3.36}
\]
and hence (3.10) holds for \( n = m \). To verify (3.11) for \( n = m \), we take \( t = t^* \) in (3.25) to get
\[
\sigma(1 - \gamma t^*) \geq -\psi'_\beta(t^*)\gamma(1 + \eta^{1+\beta}) \geq -\varphi'_\beta(t^*)\gamma(1 + \eta^{1+\beta}) \geq -\varphi'_\beta(r_\ast)\gamma(1 + \eta^{1+\beta}) = 0, \tag{3.37}
\]
which implies \( t^* \leq \frac{1}{\gamma} \). Thus, by (3.20), Lemma 3.1 is applicable to concluding that \( F'(x_m)^{-1} \) exists and
\[
\|F'(x_m)^{-1}F'(x_m)\| \leq \frac{1}{1 - \gamma\|x_m - x_0\|} \leq \frac{1}{1 - \gamma t_m}. \tag{3.38}
\]
Then by (3.13), we conclude that
\[
\|x_{m+1} - x_m\| \leq \frac{1}{1 - \gamma t_m}\|F'(x_0)\| + \eta\|F'(x_0)^{-1}F'(x_m)\|^{1+\beta}. \tag{3.39}
\]
We assert that
\[
\eta\|F'(x_0)^{-1}F'(x_m)\|^{1+\beta} \leq \eta^{1+\beta}\|F'(x_0)^{-1}F'(x_m)\|. \tag{3.40}
\]
Indeed, (3.40) is trivial in the case when \( \beta = 0 \). For the case when \( \beta > 0 \), by (3.10) (with \( n = m \)) and (2.13),
\[
\eta^{1+\beta}\|F'(x_0)^{-1}F'(x_m)\|^{1+\beta} \leq \frac{\eta^{1+\beta}\|F'(x_0)^{-1}F'(x_m)\|_{1+\beta}^{1+\beta}(t_{m+1} - t_m)^{\beta}}{1 + \eta^{1+\beta}\|F'(x_0)^{-1}F'(x_m)\|^{1+\beta}} \leq \frac{\eta^{1+\beta}\lambda^{\beta}}{1 + \eta^{1+\beta}\|F'(x_0)^{-1}F'(x_m)\|^{1+\beta}}.
\]
Recollecting that \( \lambda = \alpha(1 + \eta^{1+\beta})(1 + \theta) \), it follows from (3.6) that
\[
\eta^{1+\beta}\|F'(x_0)^{-1}F'(x_m)\|^{1+\beta} \leq \frac{\eta^{1+\beta}\lambda^{\beta}}{(1 + \eta^{1+\beta})(1 + \theta)^{\beta}} = \alpha^{\beta}\eta^{1+\beta} \leq 1,
\]
which implies (3.40) holds. Hence
\[
\|x_{m+1} - x_m\| \leq \frac{1 + \eta^{1+\beta}}{1 - \gamma t_m}\|F'(x_0)^{-1}F'(x_m)\| \leq t_{m+1} - t_m, \tag{3.41}
\]
thanks to (3.10) proved just for \( n = m \). Therefore, (3.11) holds for \( n = m \) and the proof is complete. \( \square \)

In particular, in the case when \( \eta_\ast \equiv 0 \), Algorithm A\([r_n; x_0]\) reduces to Newton’s method, and furthermore one has \( \sigma = \gamma', \theta = 0, \lambda = \alpha \). Thus \( r_\ast = \frac{1}{\gamma} \). In this case the sequence \( \{t_n\} \) defined by (2.11) reduce to Newton’s sequence with initial point \( t_0 = 0 \). Hence, by a well-known result, see for example [9, 13, 20] and [21], we have that
\[
\hat{t}^* - t_n = \frac{\xi^{2^{n-1}}}{\sum_{j=0}^{2^n-1} \xi^j}\hat{t}^* \quad \text{for each } n \in \mathbb{N}, \tag{3.42}
\]
where
\[
\hat{t}^* = 1 - \sqrt{1 - 2\alpha\gamma} \quad \text{and} \quad \xi = 1 - \sqrt{1 - 2\alpha\gamma}. \tag{3.43}
\]
Thus, Theorem 3.1 reduces to Kantorovich’s theorem.
Corollary 3.1. Let $\eta_n \equiv 0$ for each $n \geq 0$. Suppose that $F'(x_0)^{-1}F'$ satisfies the Lipschitz condition (3.1) with $r = \frac{1-\sqrt{1-2\alpha\gamma}}{\gamma}$ and that

$$0 < \alpha \gamma \leq \frac{1}{2}.$$ 

Then Newton’s sequence $\{x_n\}$ converges to a solution $x^*$ of (1.1) satisfying

$$\|x_n - x^*\| \leq \frac{\xi^{2^n-1}}{\sum_{j=0}^{2^n-1} \xi^j} \hat{t}^*,$$

where $\xi$ and $\hat{t}^*$ are defined by (3.43).

Note that the convergence criterion given in Theorem 3.1 is implicit in the case when $\beta > 0$. An explicit convergence criterion is given in the following corollary.

Corollary 3.2. Suppose that $\beta > 0$ and that $F'(x_0)^{-1}F'$ satisfies the Lipschitz condition (3.1) on $B(x_0, t^*)$. Suppose also that

$$\alpha \leq \min \left\{ \frac{2^{1-\frac{1}{\beta}}}{\mu^{\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{1}{\beta}} \eta(1+\eta\alpha\beta)}{\mu^{1+\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{2}{\beta}} \gamma(1+\eta\alpha\beta)}{\mu^{\frac{2}{\beta}} v^{2-\frac{1}{\beta}}}, \frac{1}{\nu^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{1}{\beta}} \eta(1+\eta\alpha\beta)}{\mu^{1+\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{2}{\beta}} \gamma(1+\eta\alpha\beta)}{\mu^{\frac{2}{\beta}} v^{2-\frac{1}{\beta}}} \right\},$$

(3.44)

where $\mu = \gamma + (1+\beta)\eta$ and $\nu = 1 + \eta^{1+\beta}$. Let $\{x_n\}$ be the sequence generated by Algorithm A$[r_n; x_0]$. Then $\{x_n\}$ converges to a solution $x^*$ of (1.1) and the following assertion holds:

$$\|x_n - x^*\| \leq t^* - t_n \quad \text{for each } n \in \mathbb{N}.$$  

Proof. By Theorem 3.1, it suffices to prove that (3.44) implies (3.6) in the case when $\beta > 0$. To do this, we suppose that (3.44) holds. Then

$$(1+\eta\alpha\beta)\nu \alpha \leq \frac{2^{1-\frac{1}{\beta}} (1+\eta\alpha\beta)}{\mu^{\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{1}{\beta}} \eta(1+\eta\alpha\beta)}{\mu^{1+\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{2}{\beta}} \gamma(1+\eta\alpha\beta)}{\mu^{\frac{2}{\beta}} v^{2-\frac{1}{\beta}}} \leq \frac{2^{1-\frac{1}{\beta}} (1+\eta\alpha\beta)}{\mu^{\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{1}{\beta}} \eta(1+\eta\alpha\beta)}{\mu^{1+\frac{1}{\beta}} v^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{2}{\beta}} \gamma(1+\eta\alpha\beta)}{\mu^{\frac{2}{\beta}} v^{2-\frac{1}{\beta}}} (1+\alpha\gamma\eta(1+\eta)).$$

(3.45)

Noting that $\beta > 0$ and $\nu = 1 + \eta^{1+\beta}$, one has by the definitions of $\sigma$, $\theta$ and $\lambda$ (cf. (3.4) and (3.5)) that

$$(1+\eta)(\alpha^\beta - 1))\sigma = \frac{\gamma\nu}{(1+\alpha\gamma\eta(1+\eta))},$$

$$(1+\eta(\alpha^\beta - 1))\theta = \eta,$$

$$(1+\eta)(\alpha^\beta - 1))(1+\theta) = 1 + \eta\alpha\beta$$

and

$$(1+\eta)(\alpha^\beta - 1))\lambda = (1+\eta\alpha\beta)\nu \alpha.$$
Substituting these into $\varphi_\beta$ defined by (2.2), we get that, for each $t > 0$,

$$(1 + \eta(\alpha^\beta - 1))\varphi_\beta(t) = \frac{\gamma \nu}{(1 + \alpha \gamma \eta(1 + \eta))} \frac{t^2}{2} + \eta 2^{1-\beta} t^{1+\beta} - (1 + \eta \alpha^\beta) t + (1 + \eta \alpha^\beta) \nu \alpha.$$ 

Consequently,

$$(1 + \eta(\alpha^\beta - 1))\varphi_\beta(2^{1-\beta} \mu^{-\frac{1}{\beta}} \nu^{-\frac{1}{\beta}})$$

$$= \frac{\gamma \nu}{(1 + \alpha \gamma \eta(1 + \eta))} \frac{2}{(2^{1-\beta} \mu^{-\frac{1}{\beta}} \nu^{-\frac{1}{\beta}})^2} + \eta 2^{1-\beta} (2^{1-\beta} \mu^{-\frac{1}{\beta}} \nu^{-\frac{1}{\beta}})^{1+\beta}$$

$$- (1 + \eta \alpha^\beta)(2^{1-\beta} \mu^{-\frac{1}{\beta}} \nu^{-\frac{1}{\beta}}) + (1 + \eta \alpha^\beta) \nu \alpha,$$

$$= \frac{2^{1-\beta} \gamma}{\mu^{\frac{1}{\beta}} \nu^{\frac{1}{\beta} - 1} (1 + \alpha \gamma \eta(1 + \eta))} + \frac{2^{1-\beta} \eta}{\mu^{\frac{1}{\beta}} \nu^{\frac{1}{\beta} - 1}} - \frac{2^{1-\beta} (1 + \eta \alpha^\beta)}{\mu^{\frac{1}{\beta}} \nu^{\frac{1}{\beta}}} + (1 + \eta \alpha^\beta) \nu \alpha.$$ 

This together with (3.45) implies that $\varphi_\beta(2^{1-\beta} \mu^{-\frac{1}{\beta}} \nu^{-\frac{1}{\beta}}) \leq 0$. By the definition of $r_*$, $r_*$ is the minimizer of $\varphi_\beta$ on $(0, +\infty)$. It follows that $\varphi_\beta(r_*) \leq \varphi_\beta(2^{1-\beta} \mu^{-\frac{1}{\beta}} \nu^{-\frac{1}{\beta}}) \leq 0$ and so $\lambda \leq b$ by (2.8), that is,

$$\alpha \leq \frac{\sigma r_*^2 + 2 b 2^{1-\beta} \theta r_*^{1+\beta}}{2(1 + \theta)(1 + \eta \alpha - 1)}.$$ 

The proof is complete. 

Below we consider the special case when $\beta = 1$. Then,

$$\sigma = \frac{\gamma (1 + \sqrt{\eta})}{(1 + \alpha \gamma \eta(1 + \eta))(1 + \eta(\alpha - 1))}, \quad \theta = \frac{\eta}{1 + \eta(\alpha - 1)}$$

(3.46)

and

$$\lambda = \frac{\alpha(1 + \sqrt{\eta})(1 + \alpha \eta)}{1 + \eta(\alpha - 1)}.$$ 

(3.47)

Note that in this case the sequence $\{t_n\}$ defined by (2.11) also reduces to Newton’s sequence with initial point $t_0 = 0$. Hence, as have noted in Corollary 3.1, we have that

$$i^* - t_n = \frac{x^{2n-1}}{\sum_{j=0}^{2n-1} \xi^j} i^* \quad \text{for each } n \in \mathbb{N},$$

(3.48)

where

$$i^* = \frac{1 + \theta - \sqrt{(1 + \theta)^2 - 2(\sigma + 2\theta) \lambda}}{\sigma + 2\theta} \quad \text{and} \quad \xi = \frac{1 + \theta - \sqrt{(1 + \theta)^2 - 2(\sigma + 2\theta) \lambda}}{1 + \theta + \sqrt{(1 + \theta)^2 - 2(\sigma + 2\theta) \lambda}}.$$ 

(3.49)

We have the following corollary, which gives an explicit criterion sharper than the corresponding one in Corollary 3.2, and an explicit estimate for the error $\|x_n - x^*\|$. 

Corollary 3.3. Suppose that $\beta = 1$ and that $F'(x_0)^{-1}F'$ satisfies the Lipschitz condition (3.1) with $r = \hat{t}^*$. Suppose also that

$$\alpha \leq \min \left\{ \frac{1}{2\gamma(1 + \sqrt{\eta})^2 + 4\eta(1 + \sqrt{\eta}) - \eta} , \frac{1}{\sqrt{\eta}} \right\}. \quad (3.50)$$

Then the sequence $\{x_n\}$ generated by Algorithm A[$r_n; x_0$] converges to a solution $x^*$ of (1.1) satisfying

$$\|x_n - x^*\| \leq \frac{\xi^{2^n - 1}}{\sum_{j=0}^{2^n-1} \xi^j} \hat{t}^* \quad \text{for each } n \in \mathbb{N}.$$

where $\hat{t}^*$ and $\xi$ are defined by (3.49).

Proof. Using the expressions of $\sigma$ and $\theta$ in (3.46), we conclude from (2.6) that

$$r_s = \frac{1 + \theta}{\sigma + 2\theta} = \frac{1 + \alpha\eta}{\gamma(1 + \sqrt{\eta})^2 + 4\eta(1 + \sqrt{\eta})}, \quad (3.51)$$

Therefore, noting that $\beta = 1$ and thanks to (3.46) and (3.51), the criterion (3.6) is equivalent that

$$\alpha \leq \min \left\{ \frac{1 + \alpha\eta}{2\gamma(1 + \sqrt{\eta})^2 + 4\eta(1 + \sqrt{\eta})}, \frac{1}{\sqrt{\eta}} \right\}. \quad (3.52)$$

On the other hand, (3.50) is equivalent that

$$\alpha \leq \min \left\{ \frac{1 + \alpha\eta}{2\gamma(1 + \sqrt{\eta})^2 + 4\eta(1 + \sqrt{\eta})}, \frac{1}{\sqrt{\eta}} \right\}. \quad (3.53)$$

Since $\alpha > 0$, (3.53) implies criterion (3.52); hence (3.6) holds. This completes the proof by Theorem 3.1 and (3.48).

4 Concluding remarks

By using a majorizing function technique, we have established in the previous section a general convergence criterion for the inexact Newton method with the controls given by (1.9). Our approach is different from that employed by Argyros in [1], where the Lipschitz continuity assumption on the second Fréchet derivative is assumed. The remainder of this section is devoted to some comparison of our results with some known results reported in [10] and [15].

For the special case when $\beta = 0$, Guo established in [10, Theorem 2.6] the following convergence criterion:

$$\alpha_\gamma \leq \frac{\sqrt{(4\eta + 5)^2 - 2\eta^2 - 14\eta - 11}}{(1 + \eta)(1 - \eta)^2}. \quad (4.1)$$
By Theorem 3.1 of the present paper, we have the following convergence criterion for \( \beta = 0 \):
\[
\alpha \gamma \leq \frac{(1 - \eta)^2}{(1 + \eta)(2(1 + \eta) - \eta(1 - \eta)^2)}.
\] (4.2)

Define the functions \( g_1 \) and \( g_2 \) respectively by
\[
g_1(\eta) = \frac{\sqrt{(4\eta + 5)^3} - 2\eta^2 - 14\eta - 11}{(1 + \eta)(1 - \eta)^2} \quad \text{for each } \eta \in [0, 1)
\]
and
\[
g_2(\eta) = \frac{(1 - \eta)^2}{(1 + \eta)(2(1 + \eta) - \eta(1 - \eta)^2)} \quad \text{for each } \eta \in [0, 1).
\]

Then (4.1) and (4.2) are respectively equivalent that
\[
\alpha \gamma \leq g_1(\eta) \quad \text{and} \quad \alpha \gamma \leq g_2(\eta).
\]

The graphs of the functions \( g_1 \) and \( g_2 \) are illustrated in Figure 4-1, from which one can see that the criterion (4.2) is sharper than (4.1) for \( \eta \in [0, \eta_0) \) while for \( \eta \in (\eta_0, 1) \) the criterion (4.1) is a little better than (4.2) but the difference is much smaller, where \( \eta_0 = 0.5267 \cdots \) is the solution of the equation \( g_1(\eta) = g_2(\eta) \) in \([0, 1]\).

![Figure 4-1: graphs of \( g_1(\eta) \) and \( g_2(\eta) \)]

For the case when \( \beta > 0 \), the authors of the present paper have established in [15] the following explicit convergence criterion for the inexact Newton method under the \((L, p)\)-Hölder condition:
\[
\alpha \leq \begin{cases}
\frac{\beta}{(1 + \beta)(1 + \eta^{1+\beta})} & \min \left\{ 2^{\frac{1}{1+\beta}}(1 + \eta^{\frac{\beta}{1+\beta}})^{-\frac{1}{\beta}}(\gamma + (1 + \beta)\eta)^{-\frac{1}{\beta}}, 1 \right\} \quad 0 < \beta < 1, \\
\min \left\{ \frac{1}{2(\gamma + 2\eta)(1 + \sqrt{\eta})^2}, \frac{1}{\sqrt{\eta}} \right\} & \beta = 1.
\end{cases}
\] (4.3)
Consider the special case when $\beta = 1$. Since
\[
\frac{1}{2\gamma(1 + \sqrt{\eta})^2 + 4\eta(1 + \sqrt{\eta}) - \eta} \geq \frac{1}{2\gamma(1 + \sqrt{\eta})^2 + 4\eta(1 + \sqrt{\eta})^2} = \frac{1}{2(\gamma + 2\eta)(1 + \sqrt{\eta})^2},
\] it follows that the convergence criterion (3.50) in Corollary 3.3 is sharper than that given in (4.3).

Now consider the case when $0 < \beta < 1$. We assert that our convergence criterion (3.44) obtained in Corollary 3.2 is sharper than the one in (4.3) in the case when $\gamma \geq 1$. To show this, suppose that (4.3) holds and that $\gamma \geq 1$. Recall that $\mu = \gamma + (1 + \beta)\eta$ and $\nu = 1 + \eta^{\frac{1}{1+\beta}}$. Then $\mu, \nu \geq 1$. Thus, thanks to the fact that $1 + \beta \leq 2\beta$, one has
\[
1 - \frac{(1 + \beta)\eta}{\mu} - \frac{(1 + \beta)\gamma}{2\beta \mu^\frac{1}{\beta}} = \frac{\gamma}{\mu} - \frac{(1 + \beta)\gamma}{2\beta \mu^\frac{1}{\beta}} \geq 0,
\] which is equivalent to
\[
1 - \frac{\eta}{\mu} - \frac{\gamma}{2\beta \mu^\frac{1}{\beta}} \geq \frac{\beta}{1 + \beta}.
\] Thus (4.3) gives
\[
\alpha \leq \frac{2\beta 1 - \frac{1}{\beta}}{(1 + \beta)\mu^\beta \nu^{1+\frac{1}{\beta}}} \leq \frac{2^{1-\frac{1}{\beta}}}{\mu^\beta \nu^{1+\frac{1}{\beta}}} \left(1 - \frac{\eta}{\mu} - \frac{\gamma}{2\beta \mu^\frac{1}{\beta}}\right) = \frac{2^{1-\frac{1}{\beta}}}{\mu^\beta \nu^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{1}{\beta}} \eta}{\mu^\beta \nu^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{2}{\beta}} \gamma}{\mu^\beta \nu^{1+\frac{1}{\beta}}}.
\] Hence
\[
\alpha \leq \frac{2^{1-\frac{1}{\beta}}}{\mu^\beta \nu^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{1}{\beta}} \eta}{\mu^\beta \nu^{1+\frac{1}{\beta}}} - \frac{2^{1-\frac{2}{\beta}} \gamma}{\mu^\beta \nu^{1+\frac{1}{\beta}}}.
\] This shows that (3.44) is sharper than the one in (4.3) when $\gamma \geq 1$.

We end this paper with an example for which Theorem 3.1 (with $\beta = 0$) is applicable but not [10, Theorem 2.6].

**Example 4.1.** Let $\mathbb{R}^2$ be endowed with the $l_2$-norm. Let $x_0 = (\xi_1^0, \xi_2^0)^T = (0.95, 0.96)^T$. Define $F : \mathbb{R}^2 \to \mathbb{R}^2$ by
\[
F(x) := \left(\frac{1}{2} \xi_1^2 - \frac{1}{2} \xi_2^2, \frac{1}{2} \xi_1 - \frac{23}{50}\right)^T \quad \text{for each } x = (\xi_1, \xi_2)^T \in \mathbb{R}^2.
\] Then
\[
F'(x) = \begin{pmatrix} \xi_1 & -\xi_2 \\ \frac{1}{2} & 0 \end{pmatrix} \quad \text{for each } x = (\xi_1, \xi_2)^T \in \mathbb{R}^2
\] and
\[
F'(x) - F'(y) = \begin{pmatrix} \xi_1 - \xi_1 & -\xi_2 \\ 0 & \xi_2 - \xi_2 \end{pmatrix} \quad \text{for each } x = (\xi_1, \xi_2)^T, y = (\zeta_1, \zeta_2)^T \in \mathbb{R}^2.
\]
Thus, one has that
\[
\|F'(x_0)^{-1}(F'(x) - F'(y))\|_2 \leq \|F'(x_0)^{-1}\|_F \|F'(x) - F'(y)\|_F \\
\leq \|F'(x_0)^{-1}\|_F \sqrt{(\xi_1 - \zeta_1)^2 + (\xi_2 - \zeta_2)^2} \\
= \|F'(x_0)^{-1}\|_F \|x - y\|_2.
\]

This means that $F'(x_0)^{-1}F'$ satisfies Lipschitz condition (3.1) with the Lipschitz constant
\[
\gamma = \|F'(x_0)^{-1}\|_F. \text{ Noting that } \alpha = \|F'(x_0)^{-1}F(x_0)\|_2, \text{ we obtain}
\]
\[
\alpha \gamma = \|F'(x_0)^{-1}\|_F \|F'(x_0)^{-1}F(x_0)\|_2 = 0.149144421 \cdots.
\]

On the other hand, taking $\eta = 0.15$, one has
\[
\frac{(1 - 2\eta)^2}{(1 + \eta)(2(1 + \eta) - \eta(1 - 2\eta)^2)} = 0.191370741 \cdots > \alpha \gamma
\]
and
\[
\frac{\sqrt{4\eta + 5)^3 - 2\eta^2 - 14\eta - 11}}{(1 + \eta)(1 - \eta)^2} = 0.128802424 \cdots < \alpha \gamma.
\]

Thus, Theorem 3.1 (with $\beta = 0$) is applicable but not [10, Theorem 2.6].

References


