On Newton’s method under Hölder continuous derivative ✤

Huang Zhengda

Department of Mathematics, Xixi Campus, Zhejiang University, Tianmusan Road 34, Hangzhou 310028, PR China

Received 11 October 2001
Submitted by L. Debnath

Abstract

A Mysovskii-type theorem for Newton’s method under \((k, p)\)-Hölder continuous derivative is considered in this paper. For the application studied by others, the new condition is weaker than ones in the literature. Also we prove that the optimal convergent order is \(p + 1\) for \(0 < p < 1\). © 2002 Elsevier Science (USA). All rights reserved.

Keywords: Newton’s method; Semilocal convergence; Convergent order

1. Introduction

In this study, we want to discuss the semilocal convergence of Newton’s method defined by

\[
x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \quad x_0 \in D,
\]

for solving the equation

\[
f(x) = 0,
\]

where \(f\) is a Fréchet differentiable nonlinear operator defined on an open convex subset \(D\) of a Banach space \(E_1\) with value in a Banach space \(E_2\).
The main hypotheses concerning semilocal convergence of (1) in most papers summarize to Lipschitz-type inequalities
\[ \|f'(x) - f'(y)\| \leq \frac{\int L(u) \, du}{\|x-z\|} \quad \forall x, y \in D, \tag{3a} \]

or
\[ \|f'(z)^{-1}[f'(x) - f'(y)]\| \leq \frac{\int L(u) \, du}{\|x-z\|} \quad \forall x, y \in D, \tag{3b} \]

for a proper choice of \( z \in D \) and \( L(u) \), the nondecreasing, bounded and integrable function in an interval (see [1–3]), or in other forms (see [4,5]). For example, if we choose \( L(u) \equiv K \), (3) becomes the well-known Lipschitz continuity
\[ \|f'(x) - f'(y)\| \leq K \|x - y\| \quad \forall x, y \in D, \]

or
\[ \|f'(z)^{-1}[f'(x) - f'(y)]\| \leq K \|x - y\| \quad \forall x, y \in D, \]

which is used by the classical Kantorovich theorem [6] or Mysovskii theorem [7]. If we choose \( z = x_0 \), \( L(u) = \gamma + Ku \), where \( \gamma \geq 0 \), \( K > 0 \), (3b) is
\[ \|f'(z)^{-1}[f'(x) - f'(y)]\| \leq \frac{\int L(u) \, du}{\|x-z\|}, \]

which can be deduced from the hypotheses in [8,9]. One can also see [2,3] for details about the different choices of \( L(u) \). However, there are many interesting problems in the literature; cf., for example, [10–13], where conditions (3) are not satisfied, but the Hölder continuity of derivative
\[ \|f'(x) - f'(y)\| \leq K \|x - y\|^p \quad \forall x, y \in D, \tag{4a} \]

or
\[ \|f'(z)^{-1}[f'(x) - f'(y)]\| \leq K \|x - y\|^p \quad \forall x, y \in D, \tag{4b} \]

is satisfied for some \( z \in D \), where \( K > 0 \) and \( 0 < p < 1 \).

There are papers consider the Kantorovich-type conditions for Newton’s method (1) under the hypotheses (4) (see [10,13,14]). Such conditions always need to judge the existence of positive zeros of a transcendental equation, which is complex sometimes. On the other hand, using such conditions with \( 0 < p < 1 \), results in the literature get the convergent order equal to \( p + 1 \) at least, but it seems there are no paper which proves that \( p + 1 \) is optimal, or it can be achieved, which is very different from the case if we use the hypotheses (3).
In the following, under the hypotheses (4), we will get a Mysovskii-type condition to guarantee the semilocal convergence of (1). When it is applied to a two points boundary value problem which has been studied by other authors [10, 13,14], it will be seen that the new condition is more relaxed. Meanwhile, we will prove that the optimal convergent order of (1) is $p + 1$ for $0 < p < 1$.

2. Semilocal convergence and its application

For some $z \in D$, $r > 0$, let $O(z, r)$, $\overline{O(z, r)}$ be the open and closed ball with center $z$ and radius $r$, respectively. We have

**Theorem 1.** Let $f(x)$ be defined in Section 1, (4a) holds on $D$ for $0 < p \leq 1$, $x_0 \in D$ and

\[
\left\| f'(x)^{-1} \right\| \leq B \quad \forall x \in D,
\]

\[
\left\| f'(x_0)^{-1} f(x_0) \right\| \leq \eta, \quad H = \frac{BK\eta^p}{1 + p} < 1.
\]

If $\overline{O(x_1, \sigma \eta)} \subset D$, where

\[
\sigma = \sum_{i=1}^{\infty} H((1+p)^{i-1})/p,
\]

then the sequence $\{x_n\}$ defined by (1) starting from $x_0$ converges to the simple root $x^*$ of Eq. (2) with optimal convergent order equal to $1 + p$ and

\[
\|x^* - x_n\| \leq \eta \sum_{i=n}^{\infty} H((1+p)^{i-1})/(p).
\]

**Proof.** Since $H < 1$ and $((1+p)^i - 1)/p \geq i$ for $i \geq 1$, we have $\sigma < H/(1 - H) < \infty$. Assume that for some $k \geq 1$, $x_k \in O(x_1, \sigma \eta)$; then

\[
\|x_{k+1} - x_k\| = \left\| f'(x_k)^{-1} f(x_k) \right\| \leq B \int_{0}^{1} K \theta^p d\theta \|x_k - x_{k-1}\|^{1+p}
\]

\[
\leq \frac{BK}{1 + p} \|x_k - x_{k-1}\|^{1+p}
\]

follows from (4a) and

\[
f(x_k) = f(x_k) - f(x_{k-1}) - f'(x_{k-1})(x_k - x_{k-1})
\]

\[
= \int_{0}^{1} [f'(x_{k-1} + \theta(x_k - x_{k-1})) - f'(x_k)] d\theta(x_k - x_{k-1}).
\]
Hence, by induction, we obtain
\[ \|x_{n+1} - x_n\| \leq \eta H^{((1+p)n-1)/p} \quad \forall n \geq 0, \]
and
\[ \|x_{n+j} - x_n\| \leq \sum_{i=n}^{n+j} \|x_{i+1} - x_i\| \leq \eta \sum_{i=n}^{n+j} H^{((1+p)i-1)/p} \quad \forall j \geq 0, n \geq 0, \]
which implies that for all \( n \geq 0, x_n \in O(x_1, \sigma \eta) \) and \( \{x_n\} \) is a Cauchy sequence.
So \( \exists x^* \in O(x_1, \sigma \eta) \) such that \( x_n \to x^* \) as \( n \to \infty \). By letting \( j \to \infty \) in (6), (5) holds and consequently the convergent order of (1) is at least \( 1 + p \). Now, to complete the proof of the theorem, it is sufficient to prove that \( 1 + p \) can be achieved. As usual, we construct a one-dimensional function and discuss the convergent order of Newton’s method applying to it. Let \( E_1 = E_2 = \mathbb{R} \) and
\[
\phi(t) = \begin{cases} 
-t + \frac{K}{1+p} t^{1+p}, & t \geq 0, \\
-t - \frac{K}{1+p} (-t)^{1+p}, & t \leq 0,
\end{cases}
\]
with a zero \( t^* = 0 \); then (4a) holds since
\[ |\phi'(u) - \phi'(v)| = K |u|^p - |v|^p \leq K |u - v|^p \quad \forall u, v \in (-\infty, +\infty) \]
follows from the equality \((1+x)^p \leq 1 + x^p \ \forall x > 0 \) and \( 0 \leq p \leq 1 \). Newton’s method applying to \( \phi(t) \) starting from \( 0 < t_0 < \sqrt[1+p]{1/K} \) generates a sequence defined by
\[
t_{n+1} = \begin{cases} 
\frac{pKt_0^{1+p}}{(p+1)(1+Kt_n^{p})} t_n^{1+p}, & t_n \geq 0, \\
\frac{pK}{(p+1)(1-K(-t_n)^{p})} (-t_n)^{p} t_n, & t_n < 0,
\end{cases} \quad n \geq 0.
\]
It can be checked that
\[
B = \frac{1}{1 - K t_0^p}, \quad \eta = \frac{pKt_0^{1+p}}{(1+p)(1-Kt_0^p)},
\]
and
\[
H = \frac{p^p [K t_0^p]^{1+p}}{(1+p)^2 (1-K t_0^p)^2}.
\]
Since inequality
\[
\frac{p^p}{(1+p)^2} x^{1+p} < (1-x)^2
\]
is true near \( x = 0 \), so long as we choose \( t_0 > 0 \) small enough such that \( x = K t_0^p < (1+p)/(1+2p) \) satisfies the equality, then \( H < 1 \) and (7) converges to \( t^* \). From (7) we have
\[
\lim_{n \to \infty} \frac{t_{2n+1} - t^*}{(t_{2n} - t^*)^{1+p}} = \lim_{n \to \infty} \frac{t^* - t_{2n+2}}{(t^* - t_{2n+1})^{1+p}} = -\frac{pK}{1+p}.
\]
which tells us that $1 + p$ can be achieved and is optimal. The proof is completed. □

**Corollary 1.** Suppose $p = 1$, $x_0 \in D$, $\|x_1 - x_0\| \leq \eta$, and $\|f'(x)^{-1}\| \leq B \ \forall x \in D$. If $O(x_1, \sigma \eta) \subset D$, $H = KB\eta < 2$, then the sequence $\{x_n\}$ defined by (1) starting from $x_0$ converges to a zero $x^*$ of Eq. (2) and

$$
\|x^* - x_n\| \leq \left[ 1 + \sigma \left( \left( \frac{H}{2} \right)^{2^n} \right) \right] \left( \frac{H}{2} \right)^{2^n - 1} \eta, \ n \geq 0.
$$

The condition of Corollary 1, which is more relaxed than that of Mysovskii theorem [7], was proved in [15].

Doing similarly as in Theorem 1, we can get the following convergent results for $p = 0$.

**Theorem 2.** Let $f(x)$ be defined in Section 1, (4a) holds on $D$ for $p = 0$, $x_0 \in D$ and

$$
\left\| f'(x)^{-1} \right\| \leq B \ \forall x \in D,
$$

$$
\left\| f'(x_0)^{-1} f(x_0) \right\| \leq \eta, \quad H = BK < 1.
$$

If $O(x_1, H\eta/(1 - H)) \subset D$, then the sequence $\{x_n\}$ defined by (1) starting from $x_0$ converges to the simple root $x^*$ of Eq. (2) with convergent order equal to 1 at least and

$$
\|x^* - x_n\| \leq \frac{H^n \eta}{1 - H}.
$$

In the following, we apply Theorem 1 to the two points boundary value problem

\begin{equation}
\begin{aligned}
x'' + x^{1+p} &= 0, \quad p \in [0.1], \\
x(0) &= x(1) = 0,
\end{aligned}
\end{equation}

which was studied by [10,13,14]. As did before, we divide the interval $[0, 1]$ into $m$ subintervals and we get $h = 1/m$. Let $v_0, v_1, \ldots, v_m$ be points of subdivision with $0 = v_0 < v_1 < \cdots < v_m = 1$. An approximation for the second derivative may be chosen as

$$
x_i'' = \frac{x_{i-1} - 2x_i + x_{i+1}}{h^2}, \quad x_i = x(v_i), \ i = 1, 2, \ldots, m - 1,
$$

$$
x_0 = x_1 = 0.
$$

To get an approximation to the solution of (8), we need to solve the nonlinear equation

\begin{equation}
\begin{aligned}
f(x) = Mx + h^2 \phi(x) &= 0,
\end{aligned}
\end{equation}

where the operator $f: \mathbb{R}^{m-1} \to \mathbb{R}^{m-1}$,

$$M = \begin{pmatrix} -2 & 1 & & & & \frac{1}{p} \\ 1 & -2 & 1 & & & \frac{1}{p} \\ & \ddots & \ddots & \ddots & \vdots & \vdots \\ & & 1 & -2 & 1 & \frac{1}{p} \\ & & & & & \frac{1}{p} \end{pmatrix}, \quad \phi(x) = \begin{pmatrix} x_1^{1+p} \\ \vdots \\ x_{m-1}^{1+p} \end{pmatrix},$$

and

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m-1} \end{pmatrix}.$$

Let $x \in \mathbb{R}^{m-1}$, $G \in \mathbb{R}^{m-1} \times \mathbb{R}^{m-1}$ and define the norm of $x$ and $G$ by

$$\|x\| = \max_{1 \leq j \leq m-1} |x_j|, \quad \|G\| = \max_{1 \leq i \leq m-1} \sum_{j=1}^{m-1} |g_{ij}|.$$

Since

$$f'(x) = M + h^2 (p+1) \begin{pmatrix} x_1^p \\ x_2^p \\ \vdots \\ x_{m-1}^p \end{pmatrix}$$

for all $x, z \in \mathbb{R}^{m-1}$, for which $x_i > 0, z_i > 0, i = 1, 2, \ldots, m-1$, we get

$$\|f'(x) - f'(y)\| = \frac{3}{2} h^2 \max_{1 \leq j \leq m-1} |x_j^p - z_j^p| \leq \frac{3}{2} h^2 \left\{ \max_{1 \leq j \leq m-1} |x_j - z_j| \right\}^p \leq \frac{3}{2} h^2 \|x - z\|^p.$$

(10)

Let $m = 10$, $p = 1/2$, and

$$D = \left\{ \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{array} \right| \begin{array}{c} 29.958 \leq x_1 \leq 29.958 + 4 \\ 61.964 \leq x_2 \leq 61.964 + 4 \\ 88.66 \leq x_3 \leq 88.66 + 4 \\ 106.492 \leq x_4 \leq 106.492 + 4 \\ 112.770 \leq x_5 \leq 112.770 + 4 \\ 106.488 \leq x_6 \leq 106.485 + 4 \\ 88.653 \leq x_7 \leq 88.653 + 4 \\ 61.958 \leq x_8 \leq 61.958 + 4 \\ 29.956 \leq x_9 \leq 29.956 + 4 \end{array} \right\}.$$

We can verify that

$$\|f'(x)^{-1}\| \leq 46.52 \quad \forall x \in D.$$

(11)
In fact, if we divide $D$ into $2^7$ hypercubes with length of edge equal to 2, then by Matlab5.3 it can be computed that at center $x_c$ of each hypercube holds $\|f'(x_c)^{-1}\| \leq 27.40$, and consequently, (11) follows from Banach Lemma and (10). Let

$$z_0 = \begin{pmatrix} 40.1524 \\ 76.3785 \\ 105.135 \\ 123.611 \\ 129.999 \\ 123.675 \\ 105.257 \\ 76.5462 \\ 40.3495 \end{pmatrix}$$

as in [10,13,14]. After 2 iterates of Newton’s method $z_{n+1} = z_n - f'(z_n)^{-1}f(z_n)$, $n \geq 0$, we obtain

$$z_1 = \begin{pmatrix} 33.9597 \\ 65.9638 \\ 92.6580 \\ 110.4912 \\ 116.7693 \\ 110.4878 \\ 92.6524 \\ 65.9579 \\ 33.9559 \end{pmatrix} \in D, \quad z_2 = \begin{pmatrix} 33.5775 \\ 65.2094 \\ 91.5757 \\ 109.1792 \\ 115.3752 \\ 109.1792 \\ 91.5757 \\ 65.2094 \\ 33.5775 \end{pmatrix} \in D.$$

If we choose $z_1, z_2$ as $x_0, x_1$ for our theorem, respectively, our computations by Matlab5.3 lead to the following results:

$$\eta = \|f'(x_0)^{-1}f(x_0)\| \leq 1.39408, \quad K = 0.015, \quad H \leq 0.55 < 1,$$

$$\sigma = 0.84093829161\ldots, \quad \sigma \eta \leq 1.173.$$

Hence the hypercube $O(x_1, \sigma \eta) \subset D$ combined with (10) and (11) deduces that the hypotheses of Theorem 1 are all satisfied. So, $\{z_n\}$ defined by Newton’s method (1) starting from $z_0$ converges to a simple zero $x^*$ of (9), and an approximation to the solution of (8) with $p = 1/2$ is obtained.

Our condition is more relaxed than those in [10,13,14] for the problem discussed, because it is needed two more iterations and to choose $z_4$ as $x_0$ to verified the convergence there.

In [12], (4) is replaced by a weaker hypotheses

$$\|f(x) - f(x_0)\| \leq \omega(\|x - x_0\|) \quad \forall x \in D,$$

where $x_0 \in D$ and $\omega(t)$ is a nondecreasing positive function. If we can verify that $R$, the small positive root of a equation, satisfies $\omega(R) < 1/3, \overline{O(x_0, R)} \subset D$
and other conditions, then the sequence defined by (1) starting from $x_0$ converges to the unique zero in $\overline{O(x_0, R)}$ by Theorem 3 in [12]. For the problem (8) with $p = 1/2$, we have $\omega(t) = t^{0.5}$ and $R < 0.5774$. Since $\|z_1 - z_0\| > \|z_1 - z_2\| > 1 > 0.5774$, therefore we can not choose $z_1$ or $z_2$ as $x_0$ of the Theorem 3 in [12] to judge the convergence of the sequence. Hence our condition is also more relaxed than that in [12] for the problem discussed.

References