A COMPARATIVE STUDY BETWEEN CONVERGENCE RESULTS FOR NEWTON'S METHOD

IOANNIS K. ARGYROS\textsuperscript{a} AND SAÏD HILOUT\textsuperscript{b}

ABSTRACT. We present a new theorem for the semilocal convergence of Newton's method to a locally unique solution of an equation in a Banach space setting. Under a gamma-type condition we show that we can extend the applicability of Newton's method given in [12]. We also provide a comparative study between results using the classical Newton-Kantorovich conditions ([6], [7], [10]), and the ones using the gamma-type conditions ([12], [13]). Numerical examples are also provided.

1. INTRODUCTION

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

\begin{equation}
F(x) = 0,
\end{equation}

where $F$ is a twice-Frèchet-differentiable operator defined on a convex subset $\mathcal{D}$ of a Banach space $\mathcal{X}$ with values in a Banach space $\mathcal{Y}$.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is driven by the equation $\dot{x} = Q(x)$, for some suitable operator $Q$, where $x$ is the state. Then the equilibrium states are determined by solving equation (1.1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations),

\textsuperscript{a} Received by the editors December 5, 2007. Accepted October 28, 2008.
\textsuperscript{b} 2000 Mathematics Subject Classification. 65G99, 65K10, 47H17, 49M15.
\textit{Key words and phrases.} Newton's method, Banach space, Newton-Kantorovich-type hypotheses, gamma-type condition, majorizing sequence, local/semilocal convergence, radius of convergence, Lipschitz conditions.

or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative—when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

Newton’s method

\[
    x_{n+1} = x_n - F'(x_n)^{-1} F(x_n) \quad (n \geq 0), \quad (x_0 \in D)
\]

is undoubtedly the most popular method for generating a sequence approximating \(x^*\). Here \(F'(x) \in \mathcal{L}(\mathcal{X}, \mathcal{Y})\), the space of bounded linear operators from \(\mathcal{X}\) into \(\mathcal{Y}\), denotes the Fréchet–derivative of operator \(F\) ([6], [10]).

A survey on local as well as semilocal convergence theorems for Newton’s method (1.2) under Newton-Kantorovich-type or \(\gamma\)-type conditions can be found in [6], [10], and the references there (see [1]-[5], [7]-[9], [11]-[13]).

In section 3, we provide a new semilocal convergence theorem under a \(\gamma\)-type condition (see (3.7)), which extends the applicability of the corresponding elegant theorem of Wang [12].

In sections 3 and 4 we compare semilocal and local results on Newton’s method in order for us to answer to the question: (which is the motivation for writing this paper)

Can you find conditions under which Newton-Kantorovich-type results can be used in cases \(\gamma\)-type results on Newton’s method (1.2) cannot be used and vice versa ?

Numerical examples are also provided.

2. Preliminaries

Let \(b \geq 0\), and \(\gamma > 0\) be fixed. It is convenient for us to define function \(f\) on \([0, \frac{1}{\gamma}]\) by

\[
    f(t) = b - t + \frac{\gamma \cdot t^2}{1 - \gamma \cdot t};
\]

constants \(\alpha, \xi^*\) and \(\xi^{**}\)

\[
    \alpha = b \cdot \gamma,
\]
\[
\begin{align*}
  t_1^* &= \frac{1 + \alpha - \sqrt{(1 + \alpha)^2 - 8 \alpha}}{4 \gamma}, \\
  t_1^{**} &= \frac{1 + \alpha + \sqrt{(1 + \alpha)^2 - 8 \alpha}}{4 \gamma}.
\end{align*}
\]

If

\[(2.2) \quad \alpha \leq 3 - 2 \sqrt{2}, \]

then \( f \) has \( t_1^* \) and \( t_1^{**} \) as real roots, and by \([12]\):

\[(2.3) \quad b \leq t_1^* \leq \left(1 + \frac{1}{\sqrt{2}}\right) b \leq \left(1 - \frac{1}{\sqrt{2}}\right) \gamma = r_0 \leq t_1^{**} \leq \frac{1}{2 \gamma}.\]

3. **Semilocal Convergence Analysis of Newton’s Method (1.2)**

We state first our semilocal convergence theorem, followed by the corresponding one by Wang:

**Theorem 3.1** ([3]). Let \( F : \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y} \) be a Fréchet-differentiable operator. If there exists \( x_0 \in \mathcal{D} \) with \( F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X}) \), and non-negative constants \( l_0 \) and \( l \) such that

\[(3.1) \quad \| F'(x_0)^{-1} F(x_0) \| \leq b,\]

\[(3.2) \quad \| F'(x_0)^{-1} (F'(x) - F'(y)) \| \leq l \| x - y \|,\]

\[(3.3) \quad \| F'(x_0)^{-1} (F'(x) - F'(x_0)) \| \leq l_0 \| x - x_0 \|,\]

for all \( x, y \in \mathcal{D} \),

\[(3.4) \quad h_0 = 2 L b \leq 1, \quad L = \frac{l + l_0}{2},\]

and

\[(3.5) \quad \overline{U}(x_0, r_1 = 2b) \subseteq \mathcal{D},\]

then, sequence \( \{x_n\} \) \((n \geq 0)\) generated by the Newton’s method (1.2), is well defined, remains in \( \overline{U}(x_0, r_1) \) for all \( n \geq 0 \), and converges to the unique solution \( x^* \) of equation \( F(x) = 0 \) in \( \overline{U}(x_0, r_1) \).

**Theorem 3.2** ([13]). Let \( F : \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y} \) be a twice-Fréchet-differentiable operator. If there exists \( x_0 \in \mathcal{D} \) with \( F'(x_0)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X}) \), and a positive constant \( \gamma \) such that

\[(3.6) \quad \| F'(x_0)^{-1} F(x_0) \| \leq b,\]
\[
\left\| F'(x_0)^{-1} F''(x) \right\| \leq \frac{2 \gamma}{(1 - \gamma \left\| x - x_0 \right\|)^3} \quad \text{for all } x \in D,
\]
\[
\alpha = b \gamma \leq 3 - 2\sqrt{2},
\]
and
\[
\overline{U}(x_0, r_0) \subseteq D,
\]

then, sequence \( \{x_n\} \) (\( n \geq 0 \)) generated by the Newton's method (1.2), is well defined, remains in \( \overline{U}(x_0, r_0) \) for all \( n \geq 0 \), and converges to the unique solution \( x^* \) of equation \( F(x) = 0 \) in \( \overline{U}(x_0, r_0) \).

### 3.1. A Comparison Study between Theorems 3.1 and 3.2

It can then easily be seen by simply comparing the hypotheses of the above theorems that the following hold true:

1. Hypotheses of Theorem 3.1 imply hypotheses of Theorem 3.2 provided that

\[
L \left( 1 - \frac{1}{\sqrt{2}} \right) \leq \gamma \leq 2L (3 - 2\sqrt{2}).
\]

In this case we must use the theorem providing the smaller ratio of convergence.

2. Case favorable to Theorem 3.1: (that is the hypotheses of Theorem 3.1 hold true, but hypotheses of Theorem 3.2 are violated).

Under the hypotheses of Theorem 3.1, further assume

\[
h_0 < \frac{L}{\gamma} \left( 1 - \frac{1}{\sqrt{2}} \right),
\]

and

\[
\gamma > L \left( 1 - \frac{1}{\sqrt{2}} \right),
\]

then Theorem 3.2 cannot apply since \( r_0 > r_1 \).

3. Hypotheses of Theorem 3.2 imply hypotheses of Theorem 3.1 provided that

\[
\gamma \geq 2(3 - 2\sqrt{2}) L,
\]

and

\[
\alpha \leq \frac{1}{2} \left( 1 - \frac{1}{\sqrt{2}} \right) \quad (< 3 - 2\sqrt{2}).
\]

In this case we again use the result providing the smaller ratio of convergence.
(4) Case favorable to Theorem 3.2: Under the hypotheses of Theorem 3.2, further assume

\[
\frac{1}{2} \left(1 - \frac{1}{\sqrt{2}}\right) < \alpha \leq 3 - 2\sqrt{2},
\]

then Theorem 3.1 cannot apply since \(r_1 > r_0\).

(5) Let [10]:

\[
h = 2lb \leq 1,
\]

\[
\bar{U}(x_0, r_2) \subseteq D,
\]

where

\[
r_2 = \frac{1 - \sqrt{1 - 2h}}{l}, \quad l \neq 0.
\]

Condition (3.16) is the famous Newton–Kantorovich hypothesis, which is the sufficient condition for the convergence of Newton’s method (1.2) to a unique solution \(x^*\) of equation \(F(x) = 0\), provided that \(\bar{U}(x_0, r_2) \subseteq D\).

Note that in general

\[
l_0 \leq l
\]

holds and \(\frac{l}{l_0}\) can be arbitrarily large [3]–[6]. We also have that

\[
h \leq 1 \implies h_0 \leq 1,
\]

but not vice versa unless if \(l_0 = l\).

Then under hypotheses of Theorem 3.2, (3.2), further assume

\[
2\gamma < l \leq 3.216162028 \gamma,
\]

and

\[
\alpha = b\gamma \leq \frac{b(2 - b)}{2(4 - b)}, \quad b = 4 \left(1 - \frac{\sqrt{2l}}{\gamma}\right),
\]

then (3.16) holds. Moreover if \(r_2 \leq r_0\), then the hypotheses of Theorem 3.1 imply the Newton–Kantorovich hypotheses (3.16) and (3.17). In this case we use the results providing the smaller ratio of convergence.

(6) A more interesting case is given below, where we have convergence under hypotheses of Theorem 3.2, but not under the Newton–Kantorovich hypotheses
Indeed it can easily be seen that if together with (3.21), condition
\begin{equation}
\frac{b(2-b)}{2(4-b)} < \alpha \leq 3 - 2\sqrt{2}
\end{equation}
holds, then \( h > 1 \).

3.2. Examples

**Example 3.3.** Let \( X = Y = \mathbb{R} \), \( D = [a, 2-a] \), \( a \in [0, \frac{1}{2}] \), \( x_0 = 1 \), and define function \( F \) on \( D \) by
\begin{equation}
f(x) = x^3 - a.
\end{equation}
Using (3.24), we get \( b = \frac{1}{3} (1 - a) \), \( l = 2(2-a) \), \( l_0 = 3 - a \), and \( \gamma = 1 \). The Newton–Kantorovich hypothesis is violated, since
\begin{equation}
h = \frac{4}{3} (1 - a) (2 - a) > 1 \quad \text{for all } a \in [0, \frac{1}{2}),
\end{equation}
whereas hypothesis (3.4):
\begin{equation}
h_0 = \frac{1}{3} (1 - a) (3 - a + 2(2 - a)) \leq 1
\end{equation}
provided that \( a \in \left[ \frac{5 - \sqrt{13}}{3}, \frac{1}{2} \right) \),
where \( \frac{5 - \sqrt{13}}{3} = .464816242 \). Moreover hypothesis (2.2):
\begin{equation}
\alpha = b \leq 3 - 2\sqrt{2}
\end{equation}
for
\begin{equation*}
a \in \left[ .485281374, \frac{1}{2} \right).
\end{equation*}
Note that in [3] using a variation of hypothesis (3.4) we extended the interval to at least \( \left[ .4505, \frac{1}{2} \right) \).

The next example involves a nonlinear integral equation appearing in radiative transfer ([4], [8]).

**Example 3.4.** Let \( X = C[0,1] \) equipped with the sup-norm, \( \lambda \in \mathbb{R} \), \( x_0(s) = 1 \), and define operator \( G \) on \( X \) by
\begin{equation}
G(x(s)) = \lambda x(s) \int_0^1 \frac{s}{s+t} x(t) \, dt - x(s) + 1.
\end{equation}
Let $\gamma = 1.0606 \lambda$. We have [6], [8], [10]:

$$\| G'(x_0(s)) \| \leq 1.53039421 = b_0,$$

and

$$\| G'(x_0(s))^{-1} G(x_0(s)) \| \leq b_0 \| G(x_0(s)) \| \leq b_0 |\lambda| \ln 2,$$

Set $b = b_0 |\lambda| \ln 2$. It can then easily be seen that all hypotheses of Theorem 3.2 are satisfied provided that

$$\alpha = b \gamma = 1.125072211 |\lambda|^2 \leq 3 - 2 \sqrt{2},$$

which is true for $|\lambda| \leq .390511759$. However, our Theorem 3.1 extends the range of $\lambda$ to

$$|\lambda| \leq .394464158.$$

4. LOCAL CONVERGENCE ANALYSIS OF NEWTON’S METHOD (1.2)

We now state four local convergence theorem for Newton’s method in order to compare them to each other:

**Theorem 4.1 ([11]).** Let $F : D \subseteq X \rightarrow Y$ be a Fréchet–differentiable operator. If there exists $x^* \in D$ with $F'(x^*)^{-1} \in \mathcal{L}(Y, X)$, with $F(x^*) = 0$, and constant $K \geq 0$ such that

$$\| F'(x^*)^{-1} (F'(x) - F'(y)) \| \leq K \| x - y \|,$$

for all $x, y \in D$.

Then, sequence $\{x_n\}$ $(n \geq 0)$ generated by the Newton’s method (1.2), is well defined, remains in $U(x^*, R_1)$ for all $n \geq 0$, provided that $x_0 \in \overline{U}(x^*, R_1)$ and

$$\overline{U}(x^*, R_1) \subseteq D,$$

where,

$$R_1 = \frac{2}{3K}.$$

Moreover sequence $\{x_n\}$ converges quadratically to $x^*$, so that

$$\| x_{n+1} - x^* \| \leq \frac{K}{2(1 - K \| x_n - x^* \|)} \| x_n - x^* \|^2 (n \geq 0).$$

Note that it follows from (4.1) that there exists $K_0 \geq 0$ such that

$$\| F'(x^*)^{-1} (F'(x) - F'(x^*)) \| \leq K_0 \| x - x^* \|,$$
for all \( x \in \mathcal{D} \).

Then we have

\[ R_2 = \frac{2}{2K_0 + K}, \]

and

\[ \| x_{n+1} - x^* \| \leq \frac{K \| x_n - x^* \|^2}{2(1 - K_0 \| x_n - x^* \|)} \quad (n \geq 0). \]

Note that in general

\[ K_0 \leq K \]

holds and \( \frac{K}{K_0} \) can be arbitrarily large [3]–[6].

Moreover by (4.2), (4.4) and (4.5) we have

\[ R_1 \leq R_2. \]

If strict inequality holds in (4.5), so does in (4.6).

**Theorem 4.3** ([13]). Let \( F : \mathcal{D} \subseteq \mathcal{X} \rightarrow \mathcal{Y} \) be an analytic operator. Assume there exist \( x^* \in \mathcal{D} \) \( F'(x^*)^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X}) \), and \( \gamma^* \geq 0 \) given by

\[ \gamma^* = \sup_{k \geq 2} \left\| \frac{F'(x^*)^{-1}F^{(k)}(x^*)}{k!} \right\|^{\frac{1}{k-1}} \]

such that

\[ \| F'(x^*)^{-1}F''(x) \| \leq \frac{2 \gamma^*}{(1 - \gamma^* \| x - x^* \|)^3} \]

for all \( x \in \mathcal{D} \), then the conclusions of Theorem 4.1 hold in the ball \( \overline{U}(x^*, R_3) \), where

\[ R_3 = \frac{1}{2 \gamma^*} (3 - 2\sqrt{2}). \]

**Theorem 4.4.** Under the hypotheses of Theorem 4.3, condition (4.8) is replaced by

\[ \| F'(x^*)^{-1}F''(x) \| \leq \frac{2 \gamma}{(1 - \gamma \| x - x^* \|)^3} \]

for some \( \gamma \geq 0 \), and all \( x \in \mathcal{D} \). Then the conclusions of Theorem 4.1 hold true in the ball \( \overline{U}(x^*, R_4) \), where

\[ R_4 = \frac{5 - \sqrt{13}}{6 \gamma}, \]
and
\[ \| x_{n+1} - x^* \| \leq \frac{\gamma (1 - \| x_n - x^* \|)}{2 (1 - \gamma \| x_n - x^* \|)^2 - 1} \| x_n - x^* \|^2 \quad (n \geq 0). \]

Note that for \( \gamma = \gamma^* \)
\[ R_3 < R_4. \]

**Proof.** We use mathematical induction to arrive at the estimates:

\[ (4.12) \quad \| F'(x_n)^{-1} F'(x^*) \| \leq \frac{(1 - \gamma \| x_n - x^* \|)^2}{2 \left( 1 - \gamma \| x_n - x^* \| \right)^2 - 1}, \]

\[ (4.13) \quad \| F'(x^*)^{-1} \int_0^1 F''(x^* + t(x_n - x^*)) (1 - t) \, dt \| \| x_n - x^* \| \]
\[ \leq \int_0^1 \frac{2 \gamma \| x_n - x^* \|^2 (1 - t)}{(1 - \gamma \| x_n - x^* \| t)^3} \, dt \leq \frac{\gamma \| x_n - x^* \|^2}{1 - \gamma \| x_n - x^* \|}, \]

which together with the identity
\[ x_{n+1} - x^* = -(F'(x_n)^{-1} F'(x^*)) F'(x^*)^{-1} \]
\[ \times \int_0^1 F''(x^* + (1 - t)(x_n - x^*)) (1 - t) (x_n - x^*)^2 \, dt, \]

lead to
\[ \| x_{n+1} - x^* \| \leq \| x_n - x^* \| < R_4, \]

which imply that for \( x_0 \in U(x^*, R_4), x_n \in U(x^*, R_4), \) and \( \lim_{t \to \infty} x_n = x^*. \)

That completes the proof of the theorem. \( \square \)

In order for us to compare the smaller ratios, and larger radii provided in our Theorems 4.2 and 4.4, let us denote by \( p, q \) the ratios in Theorems 4.2 and 4.4 respectively.

Define scalar function \( g \) by

\[ (4.15) \quad g(t) = (2 K - K_0) t^2 + (K_0 + \gamma - 4 K) t + 2 K - \gamma. \]

Then by (4.4) and (4.11) we have

\[ (4.16) \quad \frac{R_4}{R_2} = \frac{.11620406 \cdot 2 K_0 + K}{\gamma} = \begin{cases} < 1 & \text{if } 2 K_0 + K < 8.605551303 \gamma, \\ 1 & \text{if } K_0 + K = 8.605551303 \gamma, \\ > 1 & \text{if } 2 K_0 + K > 8.605551303 \gamma. \end{cases} \]

Concerning the ratios we have the following cases for \( r = \min\{R_2, R_4\} \) and \( U(t^*, r) \subseteq D: \)

1. If \((2 - \sqrt{3}) \gamma < K_0 < (2 + \sqrt{3}) \gamma, \) then \( p > q. \)
(2) If \((2 - \sqrt{3}) \gamma \geq K_0\) or \(K_0 \geq (2 + \sqrt{3}) \gamma\), and \(2K - \gamma > 0\), \(K_0 + \gamma - 4K > 0\), then \(p > q\).

(3) If \((2 - \sqrt{3}) \gamma \geq K_0\) or \(K_0 \geq (2 + \sqrt{3}) \gamma\), and \(2K - \gamma > 0\), \(K_0 + \gamma - 4K < 0\), \(r < R_S < \frac{1}{\gamma} \min\{R_2, R_4\}\), then \(p < q\), where \(\gamma R_S\) is the smaller zero of function \(g\).

**Remark 4.5.** As noted in [1], [5], [6], [7], [10], [14] the local results obtained here can be used for projection method such as Arnold’s, the generalized minimum residual method (GMRES), the generalized conjugate residual method (GCR), for combined Newton/finite projection methods, and in connection with the mesh independence principle to develop the cheapest and most efficient mesh refinement strategies.

**Remark 4.6.** The local results obtained can also be used to solve equation of the form \(F(x) = 0\), where \(F'\) satisfies the autonomous differential equation [4]:

\[(4.17) \quad F'(x) = P(F(x)),\]

where \(P : \mathcal{Y} \rightarrow \mathcal{X}\) is a known continuous operator. Since \(F'(x^*) = P(F(x^*)) = P(0)\), we can apply our results without actually knowing the solution of \(x^*\) of equation (1.1).

We compare the results of this section using a numerical example:

**Example 4.7.** Let \(\mathcal{X} = \mathcal{Y} = \mathbb{R}, \mathcal{D} = U(0, 1)\), and define function \(F\) on \(\mathcal{D}\) by

\[(4.18) \quad F(x) = e^x - 1.\]

Note that we can set \(P(x) = x + 1\) in (4.17).

We can easily seen that \(K_0 = e - 1\), \(K = e\) and we can set \(\gamma = \frac{1}{2}\), with \(\gamma^* \leq \frac{1}{2}\).

Therefore we obtain

\[R_1 = .245252961, \quad R_2 = .324947231,\]

\[R_3 = .171573 \quad \text{and} \quad R_4 = .46481624.\]

Hence, we have

\[R_1 < R_2 \quad \text{and} \quad R_3 < R_4.\]

That is our results provide the largest radii under the same computational cost both under Newton-Kantorovich and gamma-type hypotheses.
REFERENCES


*Cameroon University, Department of Mathematics Sciences, Lawton, OK 73505, USA Email address: iargyros@cameron.edu

*Poitiers University, Laboratoire de Mathématiques et Applications, Bd. Pierre et Marie Curie, Téléport 2, B.P. 30179, 86962 Futuroscope Chasseneuil Cedex, France Email address: said.hilout@math.univ-poitiers.fr*