NEWTON’S METHODS FOR VARIATIONAL INCLUSIONS UNDER CONDITIONED FRÉCHET DERIVATIVE

Abstract. Estimates of the radius of convergence of Newton’s methods for variational inclusions in Banach spaces are investigated under a weak Lipschitz condition on the first Fréchet derivative. We establish the linear convergence of Newton’s and of a variant of Newton methods using the concepts of pseudo-Lipschitz set-valued map and ω-conditioned Fréchet derivative or the center-Lipschitz condition introduced by the first author.

1. Introduction. This paper is concerned with the problem of approximating a solution of the variational inclusion

\begin{equation}
0 \in f(x) + G(x)
\end{equation}

where \( f : X \rightarrow Y \) is a continuous function and \( G : X \rightarrow 2^Y \) is a set-valued map with closed graph; \( X \) and \( Y \) are Banach spaces. Some problems related to mathematical programming, complementarity problems, optimal control and to other fields can be represented in the form (1.1) (see for example [21, 22]).

For approximating locally the unique solution \( x^* \) of (1.1), we consider the Newton method based on the following partial linearization:

\begin{equation}
\begin{aligned}
x_0 & \text{ is given as starting point,} \\
0 & \in f(x_k) + \nabla f(x_k) (x_{k+1} - x_k) + G(x_{k+1}),
\end{aligned}
\end{equation}

where \( \nabla f(x) \) denotes the Fréchet derivative of \( f \) at \( x \). Dontchev [8] proved that Newton’s method (1.2) is quadratically convergent to \( x^* \) under the pseudo-Lipschitzianity of \( (f + G)^{-1} \) and the Lipschitz continuity of \( \nabla f \) in a
neighborhood $V$ of $x^*$ with constant $L$:

$$
||\nabla f(x) - \nabla f(y)|| \leq L||x - y||, \quad x, y \in V.
$$

Piétrus [20] obtained superlinear convergence whenever the Fréchet derivative of $f$ satisfies a Hölder condition

$$
||\nabla f(x) - \nabla f(y)|| \leq L||x - y||^p, \quad x, y \in V, \quad p \in [0, 1].
$$

A similar condition to (1.4) on the first order divided difference is used in [13]–[16] to study the local convergence for the secant and Steffensen-type methods. Some convergence analysis of (1.2) is presented in [2] using a condition on the $m$th ($m \geq 2$) Fréchet derivative $\nabla^{(m)} f$:

$$
||\nabla^{(m)} f(x) - \nabla^{(m)} f(x_0)|| \leq L||x - x_0||
$$

for $x$ in some neighborhood of $x_0$.

For the case of nonlinear equations ($G \equiv 0$ in (1.1)), algorithm (1.2) is reduced to Newton’s method for solving $f(x) = 0$ and has been widely studied these last years (see for example [1, 3, 4, 11] and the references given there).

In [17], the authors present the following variant of Newton’s method:

$$
\begin{cases}
y_0 \text{ is given as starting point,} \\
0 \in f(y_k) + h \nabla f(y_k)(y_{k+1} - y_k) + G(y_{k+1}),
\end{cases}
$$

where $h$ is a constant ($h \neq 1$). Local linear convergence is investigated in [17] under Lipschitz condition (1.3).

In this paper, we use different conditions from the previous ones to study the convergence of Newton’s method (1.2) and of the variant (1.6). We relax the usual Lipschitz and Hölder conditions. So, the main conditions required are

$$
\begin{align*}
||\nabla f(x) - \nabla f(y)|| &\leq \omega(||x - y||) \quad \text{for } x, y \text{ in } V, \\
||\nabla f(x) - \nabla f(x^*)|| &\leq \mu(||x - x^*||) \quad \text{for } x \text{ in } V,
\end{align*}
$$

where $\omega, \mu : \mathbb{R}_+ \to \mathbb{R}_+$ are continuous nondecreasing functions. When (1.7) is satisfied, we say that $\nabla f$ is $\omega$-conditioned. Condition (1.8) is called the center-Lipschitz condition on the operator $\nabla f$. Conditions (1.7) and (1.8) are used in [1, 4, 5, 12] to solve nonlinear equations.

Our main tool used for obtaining linear convergence is the Aubin continuity of $(f(x^*) + \nabla f(x^*)(\cdot - x^*) + G(\cdot))^{-1}$ at $(0, x^*)$. Before defining this concept, let us give some standard notations. We denote by $B_r(x)$ the closed ball centered at $x$ with radius $r$, and by $|| \cdot ||$ the norms of $X$ and $Y$; the distance from a point $x \in X$ to a subset $A \subset X$ is defined as $\text{dist}(x, A) = \inf_{a \in A} ||x - a||$. Let $A : X \to 2^Y$ be a set-valued map. We write $\text{gph} A = \{(x, y) \in X \times Y : y \in A(x)\}$ and $A^{-1}(y) = \{x \in X : y \in A(x)\}$. A set-valued map $\Gamma : X \to 2^Y$ is said to be $M$-pseudo-Lipschitz around
(x₀, y₀) ∈ gph Γ (M > 0) (or Aubin continuous at (x₀, y₀)) if there exist constants a and b such that
\[ e(Γ(x₁) ∩ Bₐ(y₀), Γ(x₂)) ≤ M∥x₁ - x₂∥, \quad ∀x₁, x₂ ∈ Bₐ(x₀), \]
where the excess from the set A to the set C is defined by \( e(C, A) = \sup_{x ∈ C} \text{dist}(x, A) \).

The pseudo-Lipschitz property has been introduced by Aubin (see [6]). A basic characterization of this property of the inverse of a set-valued map is given by the Graves theorem (see [9]). Other characterizations of Aubin continuity have been obtained by Rockafellar [23] using the Lipschitz continuity of the distance function \( (x, y) ↦ \text{dist}(y, Γ(x)) \) around \((x₀, y₀)\), and by Mordukhovich [18, 24] via the concept of coderivative of multifunctions \( D^*Γ(x/y) \), where
\[ v ∈ D^*Γ(x/y)(u) ⇔ (v, -u) ∈ N_{gph} r(x, y). \]

The Mordukhovich criterion says that Γ with a closed graph is pseudo-Lipschitz around \((x₀, y₀)\) if and only if
\[ \|D^*Γ(x₀/y₀)\|⁺ = \sup_{u ∈ B₁(0)} \sup_{v ∈ D^*Γ(x₀/y₀)(u)} \|v\| < ∞. \]

In fact, the Mordukhovich criterion plays a fundamental role in variational analysis and its applications. We refer the reader to [6, 7, 9, 10, 18, 19, 23, 24] and the references given there for more details and applications of Aubin continuity.

2. Preliminaries and assumptions. As the main tool of our analysis we will use the following lemmas. The first is the fixed point theorem for set-valued map proved by Dontchev and Hager [9]. This theorem is a generalization of the Picard fixed point theorem restricted to single-valued mappings. The second lemma gives an estimate using condition (1.7).

**Lemma 2.1** (see [9]). Let φ be a set-valued map from X into the closed subsets of X, let η₀ ∈ X, and let r and λ be such that 0 ≤ λ < 1 and

(a) \( \text{dist}(η₀, φ(η₀)) ≤ r(1 - λ) \).
(b) \( e(φ(x₁) ∩ Bₗ(η₀), φ(x₂)) ≤ λ∥x₁ - x₂∥ \) for all \( x₁, x₂ ∈ Bₗ(η₀) \).

Then φ has a fixed point in \( Bₗ(η₀) \), that is, there exists \( x ∈ Bₗ(η₀) \) such that \( x ∈ φ(x) \). If φ is single-valued, then x is the unique fixed point of φ in \( Bₗ(η₀) \).

**Lemma 2.2.** Suppose that the assumption (1.7) is satisfied on a convex neighborhood V. Then for all \( x \) and \( y \) in V,
\[ \|f(x) - f(y) - \nabla f(y)(x - y)\| ≤ \mu(\|x - y\|)∥x - y∥. \]
In particular, if the assumption (1.8) is satisfied then for all \( x \) in \( V \),
\[
\| f(x) - f(x^*) - \nabla f(x^*)(x - x^*) \| \leq \mu(\|x - x^*\|)\|x - x^*\|.
\]

**Proof.** For all \( x \) and \( y \) in \( V \) we can write
\[
f(x) - f(y) - \nabla f(y)(x - y) = \left( \int_0^1 \nabla f(x + t(y - x)) \, dt - \int_0^1 \nabla f(y) \, dt \right)(x - y).
\]
Hence
\[
\| f(x) - f(y) - \nabla f(y)(x - y) \| \leq \| x - y \| \int_0^1 \| \nabla f(x + t(y - x)) - \nabla f(y) \| \, dt. \]

Before stating the main result of this study, we need to introduce some notations. First, for \( k \in \mathbb{N} \) and \((x_k)\) defined in (1.2), define the set-valued mappings \( Q : X \to 2^Y \) and \( \psi_k : X \to 2^X \) by
\[
(2.1) \quad Q(\cdot) := f(x^*) + \nabla f(x^*)(\cdot - x^*) + G(\cdot), \quad \psi_k(\cdot) := Q^{-1}(Z_k(\cdot)),
\]
where \( Z_k : X \to Y \) is defined by
\[
(2.2) \quad Z_k(x) := f(x^*) + \nabla f(x^*)(x - x^*) - f(x_k) - \nabla f(x_k)(x - x_k).
\]
Note that \( x_1 \) is a fixed point of \( \psi_0 \) if and only if \( 0 \in f(x_0) + \nabla f(x_0)(x_1 - x_0) + G(x_1) \).

We will make the following assumptions in an open convex neighborhood \( V \) of \( x^* \):

- (\( H_1 \)) The condition (1.7) is satisfied on \( V \).
- (\( H_2 \)) The set-valued map \( (f(x^*) + \nabla f(x^*)(\cdot - x^*) + G(\cdot))^{-1} \) is \( M \)-pseudo-Lipschitz around \((0, x^*)\) with constants \( a \) and \( b \) (see the definition of Aubin continuity) and \( M\omega(a) < 1/2 \).

### 3. Convergence analysis

In this section we are concerned with the existence of sequences \((x_n)\) satisfying (1.2) and the \( q \)-linear convergence of \((x_n)\) to the solution \( x^* \) of (1.1) under the previous assumptions. The main result of this study is as follows.

**Theorem 3.1.** Let \( x^* \) be a solution of (1.1). Suppose that assumptions (\( H_1 \))–(\( H_2 \)) are satisfied. For every constant \( C \) such that
\[
\frac{M\omega(a)}{1 - M\omega(a)} < C < 1,
\]
one can find \( \delta > 0 \) such that for every starting point \( x_0 \neq x^* \) in \( B_\delta(x^*) \), there exists a sequence \((x_k)\) satisfying (1.2) which is \( q \)-linearly convergent to \( x^* \), i.e.,
\[
(3.1) \quad \|x_{k+1} - x^*\| \leq C\|x_k - x^*\|.
\]
The proof of Theorem 3.1 is by induction on \( k \). We prove the existence of a starting point \( x_1 \) for \( x_0 \) in \( V \).

**Proposition 3.2.** Under the assumptions of Theorem 3.1, one can find \( \delta > 0 \) such that for every starting point \( x_0 \neq x^* \) in \( B_\delta(x^*) \), the set-valued map \( \psi_0 \) has a fixed point \( x_1 \) in \( B_\delta(x^*) \) satisfying

\[ \|x_1 - x^*\| \leq C\|x_0 - x^*\|, \]

where \( C \) is as in Theorem 3.1.

**Proof.** By hypothesis \((H2)\) we have

\[ e(Q^{-1}(y') \cap B_a(x^*)) \leq M\|y' - y''\|, \quad \forall y', y'' \in B_b(0). \]  

Fix \( \delta > 0 \) such that

\[ \delta < \min \left\{ a, \frac{b}{2\omega(a)} \right\}. \]

The main idea of the proof is to show that both assertions (a) and (b) of Lemma 2.1 hold, where \( \eta_0 = x^* \), \( \phi = \psi_0 \) defined in (2.1), and where \( r \) and \( \lambda \) are numbers to be set. According to the definition of the excess \( e \), we have

\[ \text{dist}(x^*, \psi_0(x^*)) \leq e(Q^{-1}(0) \cap B_\delta(x^*), \psi_0(x^*)). \]

For all \( x_0 \) in \( B_\delta(x^*) \), by assumptions \((H1), (3.4)\) and Lemma 2.2 we have

\[ \|Z_0(x^*)\| = \|f(x^*) - f(x_0) - \nabla f(x_0)(x^* - x_0)\| \leq \omega(\|x_0 - x^*\|)\|x_0 - x^*\| \leq \omega(a)\|x_0 - x^*\|. \]

Then (3.4) yields \( Z_0(x^*) \in B_b(0) \). Consequently,

\[ e(Q^{-1}(0) \cap B_\delta(x^*), \psi_0(x^*)) = e(Q^{-1}(0) \cap B_\delta(x^*), Q^{-1}[Z_0(x^*)]) \leq M\omega(a)\|x_0 - x^*\|. \]

By inequality (3.5), we get

\[ \text{dist}(x^*, \psi_0(x^*)) \leq M\omega(a)\|x_0 - x^*\|. \]

Since \( C(1 - M\omega(a)) > M\omega(a) \), there exists \( \lambda \in [M\omega(a), 1/2[ \) such that \( C(1 - \lambda) \geq M\omega(a) \) and

\[ \text{dist}(x^*, \psi_0(x^*)) \leq C(1 - \lambda)\|x_0 - x^*\|. \]

By choosing \( r := r_0 = C\|x_0 - x^*\| \) we deduce from (3.9) that assertion (a) in Lemma 2.1 is satisfied.

By (3.4) we have \( r_0 \leq \delta \leq a \). Moreover, for \( x \in B_\delta(x^*) \), by \((H1), (3.4)\) and Lemma 2.2 we have
The inequality \( \lambda \geq M \omega(a) \) implies

\[
(3.12) \quad e(\psi_0(x') \cap \mathbb{B}_{r_0}(x^*), \psi_0(x'')) \leq \lambda \|x'' - x'\|,
\]

so condition (b) of Lemma 2.1 is satisfied. Hence \( \psi_0 \) has a fixed point \( x_1 \in \mathbb{B}_{r_0}(x^*) \).

Proof of Theorem 3.1. Keep \( \eta_0 = x^* \) and set \( r := r_k = C \|x^* - x_k\| \).

Proposition 3.2 applied to the map \( \psi_k \) gives the existence of a fixed point \( x_{k+1} \) for \( \psi_k \), which is an element of \( \mathbb{B}_{r_k}(x^*) \). This finishes the proof of Theorem 3.1.

Now, we are interested in the sequence \( (y_n) \) given by the variant (1.6) of Newton’s method. We present a linear convergence result for algorithm (1.6) under the condition (1.7). Before stating the main result, we give some notations. For \( (y_k) \) defined in (1.6), define the set-valued mappings \( P : X \to 2^Y \) and \( \Gamma_k : X \to 2^X \) by

\[
(3.13) \quad P(\cdot) := f(x^*) + h\nabla f(x^*)(\cdot - x^*) + G(\cdot), \quad \Gamma_k(\cdot) := P^{-1}(W_k(\cdot)),
\]

where \( W_k \) is defined by

\[
(3.14) \quad W_k(x) := f(x^*) + h\nabla f(x^*)(x - x^*) - f(y_k) - h\nabla f(y_k)(x - y_k).
\]

We will make the following assumptions in an open convex neighborhood \( V \) of \( x^* \):

\( (H0)^* \) There exists \( K > 0 \) such that \( \|\nabla f(y_0)\| < K \).

\( (H2)^* \) The set-valued map \( (f(x^*) + h\nabla f(x^*)(\cdot - x^*) + G(\cdot))^{-1} \) is \( M' \)-pseudo-Lipschitz around \((0, x^*)\) with constants \( a' \) and \( b' \) (given by the definition of Aubin continuity) and \( M'((1 + |h|)\omega(a') + |1 - h|K) < 1 \).
The main result is as follows.

**Theorem 3.3.** Let $x^*$ be a solution of (1.1). Suppose that assumptions $(\mathcal{H}0)^*$, $(\mathcal{H}1)$ and $(\mathcal{H}2)^*$ are satisfied. For every constant $C'$ such that

$$\frac{M'(\omega(a')) + |1 - h|K}{1 - M'|h|\omega(a')} < C' < 1,$$

one can find $\gamma > 0$ such that for $y_0 \neq x^*$ in $\mathbb{B}_\gamma(x^*)$, there exists a sequence $(y_k)$ defined by (1.6) satisfying

$$\|y_{k+1} - x^*\| \leq C'\|y_k - x^*\|.$$  \hspace{1cm} (3.15)

The proof of Theorem 3.3 is based on the following proposition.

**Proposition 3.4.** Under the assumptions of Theorem 3.3, one can find $\gamma > 0$ such that for every starting point $y_0 \neq x^*$ in $\mathbb{B}_\gamma(x^*)$, the set-valued map $\Gamma_0$ has a fixed point $y_1$ in $\mathbb{B}_\gamma(x^*)$ satisfying

$$\|y_1 - x^*\| \leq C'\|y_0 - x^*\|,$$  \hspace{1cm} (3.16)

where $C'$ is as in Theorem 3.3.

**Proof.** The proposition can be proved in the same way as Proposition 3.2. The constant $\gamma$ is selected such that

$$\gamma < \min\left\{a', \frac{b'}{(1 + |h|)\omega(a') + |1 - h|K}\right\}.$$  \hspace{1cm} (3.17)

Using assumptions $(\mathcal{H}0)^*$, $(\mathcal{H}1)$, $(\mathcal{H}2)^*$ and Lemma 2.2, obtain

$$\|W_0(x^*)\| = \|f(x^*) - f(y_0) - h\nabla f(y_0)(x^* - y_0)\|
\leq \|f(x^*) - f(y_0) - \nabla f(y_0)(x^* - y_0)\|
\quad + |1 - h| \|\nabla f(y_0)\| \|y_0 - x^*\|
\leq (\omega(\|y_0 - x^*\|) + |1 - h| \|\nabla f(y_0)\|)\|y_0 - x^*\|
\leq (\omega(a') + |1 - h|K)\|y_0 - x^*\|.$$  \hspace{1cm} (3.18)

Thus $W_0(x^*) \in \mathbb{B}_{b'}(0)$. Consequently,

$$\text{dist}(x^*, \Gamma_0(x^*)) \leq M(\omega(a') + |1 - h|K)\|y_0 - x^*\|.$$  \hspace{1cm} (3.19)

As $C'(1 - M'|h|\omega(a')) > M'(\omega(a') + |1 - h|K)$, there is $\lambda' \in [M'|h|\omega(a'), 1]$ such that $C'(1 - \lambda') \geq M'(\omega(a') + |1 - h|K)$ and

$$\text{dist}(x^*, \Gamma_0(x^*)) \leq C'(1 - \lambda')\|y_0 - x^*\|.$$  \hspace{1cm} (3.20)

Choosing $r'_0 = C'\|y_0 - x^*\|$, we have $r'_0 \leq \gamma \leq a'$. Moreover, for $x \in \mathbb{B}_\gamma(x^*)$
we obtain
\begin{equation}
\|W_0(x)\| = \|f(x^*) + h \nabla f(x^*)(x - x^*) - f(y_0) - h \nabla f(y_0)(x - y_0)\|
\leq \|f(x^*) - f(y_0) - \nabla f(y_0)(x - y_0)\|
+ |h| \|\nabla f(x^*) - \nabla f(y_0)\| \|x - x^*\|
+ |1 - h| \|\nabla f(y_0)\| \|y_0 - x^*\|
\leq ((1 + |h|) \omega(a') + |1 - h| K) \gamma. \tag{3.21}
\end{equation}
Thus by (3.17) we deduce that $W_0(x) \in B_{\gamma}(0)$ for all $x \in B_{\gamma}(x^*)$. It follows that for all $x', x'' \in B_{r_0}(x^*)$,
\begin{equation}
e(\Gamma_0(x')) \cap B_{r_0}(x^*), \Gamma_0(x'')) \leq M' \|W_0(x') - W_0(x'')\|
= M'|h| \|(\nabla f(x^*) - \nabla f(y_0))(x'' - x')\|
\leq M'|h| \|\nabla f(x^*) - \nabla f(x_0)\| \|x'' - x'\|
\leq M'|h| \omega(a') \|x'' - x'\| \leq \lambda' \|x'' - x'\|. \tag{3.22}
\end{equation}
The existence of a fixed point $y_1 \in B_{r_0}(x^*)$ for the map $\Gamma_0$ is ensured. This finishes the proof of Proposition 3.4. Theorem 3.3 is deduced by induction. }

Remark 3.5. Theorems 3.1 and 3.3 remain true under the center-Lipschitz condition (1.8).

References


Department of Mathematical Sciences
Cameron University
Lawton, OK 73505, U.S.A.
E-mail: ioannisa@cameron.edu

Department of Applied Mathematics
and Computation
Faculty of Science and Technics of Béni-Mellal
B.P. 523, Béni-Mellal 23000, Morocco
E-mail: said_hilout@yahoo.fr

Received on 19.4.2007 (1869)