LOCAL CONVERGENCE OF NEWTON’S METHOD FOR PERTURBED GENERALIZED EQUATIONS

IOANNIS K. ARGYROS

ABSTRACT. A local convergence analysis of Newton’s method for perturbed generalized equations is provided in a Banach space setting. Using center Lipschitzian conditions which are actually needed instead of Lipschitzian hypotheses on the Fréchet-derivative of the operator involved and more precise estimates under less computational cost we provide a finer convergence analysis of Newton’s method than before [5]–[7].

1. INTRODUCTION

In this study we are concerned with the problem of approximating a solution of the equation

(1) \[ o \in f(x) + g(x) + F(x), \]

where \( X, Y \) are Banach spaces, \( f: X \to Y \) is a Fréchet-differentiable operator, \( g: X \to Y \) is a continuous operator, and \( F: Y \to Y \) is a closed set-valued mapping.

Equation (1) is the perturbed problem for

(2) \[ o \in f(x) + F(x), \]

where \( g \) in (1) is the perturbed operator.

Many problems, e.g. in engineering and economics can be viewed as special cases of equation (1) [2]–[11].

The most popular method for generating a sequence approximating a solution of equation (1) is undoubtedly Newton’s method in the form

(3) \[ o \in f(x_n) + g(x_n) + F'(x_n)(x_{n+1} - x_n) + F(x_{n+1}), \quad (n \geq 0) \]

Received by the editors March 30, 2006 and, in revised form, August 12, 2006.

2000 Mathematics Subject Classification. 65K10, 65G99, 47H04, 49M15.

Key words and phrases. Newton’s method, Banach space, generalized equation with perturbation, Fréchet derivative, center-Lipschitz, Lipschitz condition, local convergence, variational inequalities.

where $F'(x)$ denotes the Fréchet-derivative of operator $F$ [9], and $x_0$ is an initial guess in some neighborhood of the solution denoted by $x^*$. A local as well as semilocal convergence analysis for method (3) involving nonlinear equations has been given in [2], [3] and the references there.

In the case of generalized equations of the form (1) Geoffory and Pietrus provided a local convergence analysis for method (3) in [7]. Here we noticed that some of their hypotheses are not really needed in the proof. Therefore, we managed under weaker hypotheses and less computational cost to provide a finer convergence analysis including more precise estimates on the distances involved.

A survey on results involving generalized equations can be found in [1]–[11] and the references there.

2. Local Convergence Analysis of Method (3)

In order for us to introduce our results we also first need to introduce some terminology and a fixed point theorem already used in [6].

As in [2], [7] we denote by $A(x,y)$ the approximation of $f(x) + g(x) + F(x)$. That is we set

$$A(x, y) = f(y) + f'(y)(x - y) + g(y) + F(x) \text{ for all } x, y \in X.$$  \hspace{1cm} (4)

It is convenient for us to define operator $Q_n: X \to Y$ by

$$Q_n(x) = f(x^*) + f'(x^*)(x - x^*) + g(x^*) - f(x_n)$$

$$- f'(x_n)(x - x_n) - g(x_n) \quad (n \geq 0),$$

and set-valued map $T_n: X \rightrightarrows Y$ by

$$T_n(x) = A(\cdot, x^*)^{-1}[Q_n(x)].$$  \hspace{1cm} (5)

Note that $x_1 \in X$ is a fixed point of $T_0$ if and only if the following implication holds true:

$$x_1 \in T_0(x_1) \iff Q_0(x_1) \in A(x_1, x^*)$$

$$\iff o \in f(x_0) + g(x_0) + f'(x_0)(x_1 - x_0) + F(x_1).$$

That is $x_1$ satisfies (3). In general if $x_n$ plays the role of $x_0$, method (3) is used to show $x_{n+1}$ is a fixed point of $T_n$ etc. This way we generate a sequence $\{x_n\}$ satisfying (3).

We will make the assumptions:
(A₁) Operator $f : X \to Y$ is Fréchet-differentiable and its derivative is $L$-Lipschitz continuous and $L_0$-center-Lipschitz continuous in a neighborhood $U$ of $x^*$. That is

$$
\|F'(x) - F'(y)\| \leq L\|x - y\| \quad \text{for all } x, y \in U,
$$

and

$$
\|F'(x) - F'(x^*)\| \leq L_0\|x - x^*\| \quad \text{for all } x \in U.
$$

(A₂) Operator $g : X \to Y$ is $K_0$-center-Lipschitz in a neighborhood $U$ of $x^*$.

(A₃) The set-valued mapping $A(\cdot, x^*)^{-1} : Y \rightrightarrows X$ is $M$-pseudo-Lipschitz at 0 for $x^*$, i.e. there exist neighborhoods $U$ of $x^*$ and $V$ of 0 such that

$$
e(A(\cdot, x^*)^{-1}(y) \cap U, A(\cdot, x^*)(z)) \leq M\|y - z\|
$$

for all $M$ such that

$$
\alpha_0 = M\left(\frac{L}{2} + K_0\right) < 1,
$$

where,

$$
e(A, B) = \sup_{x \in A} \text{dist}(x, B)
$$

denotes the excess $e$ from a set $B$ to the set $A$. The importance of introducing such a type of continuity due to Aubin has been explained in detail in [1], [5], [6], [11].

From now on we denote for $x \in X$, $r > 0$

$$
U(x, r) = \{v \in X \mid \|x - v\| \leq r\}.
$$

We need the following generalization of a fixed point theorem by Ioffe–Tikhomirov [6], [8]:

**Lemma 1.** Let $(X, \rho)$ be a Banach space. Let $T$ be a map from $X$ into the closed subsets of $X$, let $q_0 \in X$ and let $r > 0$ and $\lambda \in [0, 1)$ be such that:

$$
\text{dist}(q_0, T(q_0)) \leq r(1 - \lambda),
$$

and

$$
e(T(x_1) \cap U(q_0, r), T(x_2)) \leq \lambda \rho(x_1, x_2) \quad \text{for all } x_1, x_2 \in U(q_0, r).
$$

Then, $T$ has a fixed point in $U(q_0, r)$. Moreover if $T$ is single-valued, then $x$ is the unique fixed point of $T$ in $U(q_0, r)$.

We can show the main local convergence result of Newton's method (3):
Theorem 2. Under assumptions (A₁)–(A₃) and for any \( c \in (0,1) \) there exists \( \delta > 0 \) such that for any initial guess \( x₀ \in U(x^*, \delta) \) there exists a sequence \( \{x_n\} \) generated by Newton’s method (3) such that

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

To prove Theorem 2 we need the auxiliary result:

Proposition 3. Under the hypotheses of Theorem 2 there exist \( \delta > 0 \) such that for all \( x₀ \in U(x^*, \delta) \) \( x₀ \neq x^* \), the map \( T₀ \) has a fixed point \( x₁ \) in \( U(x^*, \delta) \).

Proof. By (A₃) there exist positive constants \( a \) and \( b \) such that

\[
e(A(\cdot, x^*)^{-1}(y) \cap U(x^*, a), A(\cdot, x^*)^{-1}(z)) \leq M\|y - z\| \quad \text{for all } y, z \in U(0, b).
\]

Choose \( \delta > 0 \) to be fixed and

\[
\delta \in (0, \delta₀),
\]

where,

\[
\delta₀ = \min \left\{ \frac{a}{c}, \frac{b}{2(L + 2K₀)} \right\}.
\]

Let \( q₀ = x^* \). We will show conditions (14) and (15) of Lemma 1 hold true.

Let \( x₀ \neq x^*, x₀ \in U(x^*, \delta) \). Using (5), (A₂), (8) and (9) we get

\[
\|Q₀(x^*)\| = \|f(x^*) - f(x₀) - f'(x₀)(x^* - x₀) + g(x^*) - g(x₀)\|
\]

\[
\leq \frac{L}{2}\|x^* - x₀\|^2 + K₀\|x^* - x₀\|.
\]

For \( \delta \) sufficiently small and (18)

\[
\|Q₀(x^*)\| \leq \left( \frac{L}{2} + K₀ \right)\|x^* - x₀\| \leq b.
\]

In view of (17) we have:

\[
e(\lambda A(\cdot, x^*)^{-1}(0) \cap U(x^*, a), A(\cdot, x^*)^{-1}(Q₀(x^*))) \leq M\|Q₀(x^*)\|,
\]

and

\[
\text{dist}(x^*, T₀(x^*)) \leq M \left( \frac{L}{2} + K₀ \right)\|x^* - x₀\|.
\]

By the choice of \( c \) there exists \( \lambda \in (0,1) \) such that \( c(1 - \lambda) \geq M \left( \frac{L}{2} + K₀ \right) \), and hence

\[
\text{dist}(x^*, T₀(x^*)) \leq c(1 - \lambda)\|x^* - x₀\|.
\]
Let $q_0 = x^*$, $r = r_0 = c\|x^* - x_0\|$. It follows (14) holds.

We shall show (15) also holds true.

In view of $\delta \leq \frac{a}{c}$, we get $r_0 \leq a$. Let $x \in U(x^*, \delta)$. We can obtain using (8), (9), (A2), and the choice of $\delta$:

$$
\|Q_0(x)\| \leq \|f(x^*) - f(x) - f'(x^*)(x - x^*)\|
+ \|f(x) - f(x_0) - f'(x_0)(x - x_0)\| + \|g(x^*) - g(x_0)\|
\leq \frac{L_0}{2}\|x^* - x_0\|^2 + \frac{L}{2}\|x - x_0\|^2 + K_0\|x^* - x_0\|
\leq 4\delta \left(\frac{L}{2} + K_0\right) \leq b,
$$

where,

$$
L = \frac{L + L_0}{2}.
$$

Moreover, for $x^1, x^2 \in U(x^*, r_0)$, we get

$$
e(T_0(x^1) \cap U(x^*, r_0), T_0(x^2)) \leq e(T_0(x^1) \cap U(x^*, \delta), T_0(x^2))
\leq M\|Q_0(x^1) - Q_0(x^2)\|
\leq M\|F'(x^*)(x^1 - x^2) - F'(x_0)(x^1 - x^2)\|
\leq ML_0\|x^* - x_0\| \|x^1 - x^2\|
\leq ML_0\delta \|x^1 - x^2\|.
$$

We can assume that without loss of generality

$$
\delta < \frac{\lambda}{ML_0} = \delta_1,
$$

which implies (15). Therefore all conditions of Lemma 1 hold true. Hence, we deduce the existence of a fixed point $x_1 \in U(x^*, r_0)$ for the map $T_0$.

That completes the proof of Proposition 3. \qed

Proof of Theorem 2. In view of $x_1 \in U(x^*, r_0)$ we get

$$
\|x_1 - x^*\| \leq r_0 = c\|x_0 - x^*\|.
$$

Using induction for $q_0 = x^*, r_k = c\|x_k - x^*\|^2$, following the proof of Proposition 3 for the map $T_k$ we conclude the existence of a fixed point $x_{k+1}$ for $T_k$ in $U(x^*, r_k)$. That is

$$
\|x_{k+1} - x^*\| \leq c\|x_k - x^*\|^2.
$$

That completes the induction and the proof of the theorem. \qed
Remark 4. In general

\[ L_0 \leq L \]

and

\[ K_0 \leq K \]

holds and \( \frac{L}{L_0}, \frac{K}{K_0} \) can be arbitrarily large \([2], [3]\), where \( K \) is the Lipschitz constant of operator \( g \) in some neighborhood \( V \) of \( x^* \), a hypothesis used in \([7]\) corresponding to our Assumption \((A_2)\). If equality holds in both \((31)\) and \((32)\) then our results reduce to the corresponding ones in \([7]\). Otherwise our results constitute an improvement since they allow: a larger \( \delta \), which implies a wider choice of initial guesses \( x_0 \); a smaller choice of \( c \) which improves the ratio of the quadratic convergence of Newton’s method \((3)\) given by \((16)\).

These observations/improvements are important in computational mathematics \([2], [3], [6], [7], [8], [11]\).

REFERENCES


DEPARTMENT OF MATHEMATICAL SCIENCES, CAMERON UNIVERSITY, LAWTON, OK 73505, USA

Email address: iargyros@cameron.edu