Korean J. Math. **20** (2012), No. 1, pp. 117–123

ON THE CONVERGENCE OF NEWTON'S METHOD FOR SET VALUED MAPS UNDER WEAK CONDITIONS

Ioannis K. Argyros

Abstract. We provide a convergence analysis of Newton's method for set valued maps under center Hölder continuity conditions on the Fréchet derivative of the operator involved. This approach extends the applicability of earlier works [4,5,7].

1. Introduction

Let X, Y be Banach spaces, $f\colon X \to Y$ be a Fréchet differentiable operator and $F\colon X \to 2^Y$ be a multi-valued operator with a closed graph.

In this study we are concerned with the problem of approximating a solution $x \in X$ of the generalized equation:

$$y \in f(x) + F(x), \tag{1.1}$$

where y is a given parameter.

Note that: if $F = 0$, then (1.1) is a nonlinear equation [1];

If F is the positive orthant in $\mathbf{R}^i$, then (1.1) is a system of inequalities;

If F is a normal cone to a convex and closed set in X, then (1.1) may represent variational inequalities.

For other examples and a survey on results concerning the solution of equation (1.1) we refer the reader to [5] and the references there.

The most popular method for generating a sequence approximating a solution of (1.1) is undoubtedly Newton's method in the form

$$y \in f(x_n) + \nabla f(x_n)(x_{n+1} - x_n) + F(x_{n+1}) \quad (n \geq 0), \tag{1.2}$$

where ∇f denotes the Fréchet-derivative of the operator f.

Received January 20, 2012. Revised February 26, 2012. Accepted March 5, 2012.

2010 Mathematics Subject Classification: 65H10, 65B05, 65G99, 47H17, 49M15.

Key words and phrases: Set valued maps, Newton's method, Aubin continuity, generalized equations, center-Hölder continuity.

A usual condition is given by the Hölder continuity assumption

$$\|\nabla f(x) - \nabla f(\overline{x})\| \le L\|x - \overline{x}\|^{\lambda} \quad \text{for all } x, \overline{x} \in D \subseteq X \text{ and some } \lambda \in (0,1],\ L > 0. \tag{1.3}$$

The case when $\lambda = 1$ was studied in [5] by Dontchev, whereas the general case was investigated by Pietrus in [7].

Here we further weaken (1.3). Indeed let x^* be a solution of (1.1).

We assume the center-Hölder continuity assumption

$$\|\nabla f(x) - \nabla f(x^*)\| \le L_0\|x - x^*\|^{\lambda_0} \quad \text{for all } x \in D \text{ and some } \lambda_0 \in (0,1],\ L_0 > 0. \tag{1.4}$$

Note that in general

$$L_0 \le L \tag{1.5}$$

and

$$\lambda_0 \ge \lambda \tag{1.6}$$

hold in general and $\frac{L}{L_0}$, $\frac{\lambda_0}{\lambda}$ can be arbitrarily large [2,3]. Clearly there are cases when (1.4) holds but not (1.3). Therefore our results can be used in cases not covered before.

Using the concept of Aubin continuity [4,6], we provide a convergence analysis of method (1.2).

2. Preliminaries

In order for us to make the paper as self-contained as possible we briefly restate some concepts already in the literature [3]–[8].

Let $r > 0$, $x \in X$. Then we set

$$U(x, r) := \{\overline{x} \in X : \|x - \overline{x}\| \le r\}. \tag{2.1}$$

Recall that a set-valued map Γ from Y to the subsets of a Banach space Z is said to be M-pseudo-Lipschitz around

$$(y_0, z_0) \in \operatorname{Graph} \Gamma := \{(y, z) \in Y \times Z : z \in \Gamma(y)\}$$

if there exist neighborhoods V of y_0 and U of z_0 such that

$$\sup \operatorname{dist}(z, \Gamma(y_2)) \le M\|y_1 - y_2\| \quad \text{for all } y_1,\ y_2 \text{ in } V, \text{ and } z \in \Gamma(y_1). \tag{2.2}$$

Equivalently, Γ is M-pseudo-Lipschitz around $(y_0, z_0) \in \text{Graph}\,\Gamma$ with constants ℓ and m if for every $y_1, y_2 \in U(y_0, m)$ and for every $z_1 \in \Gamma(y_1) \cap U(0, \ell)$ there exists $z_2 \in \Gamma(y_2)$ such that

$$\|z_1 - z_2\| \le M\|y_1 - y_2\|. \tag{2.3}$$

Let A and C be two subsets of X. We denote by $e(C, A)$ the excess from A to C given by

$$e(C, A) = \sup\{\text{dist}(x, A)\colon x \in C\}. \tag{2.4}$$

Then, we can equivalently replace (2.2) by

$$e\big(\Gamma(y_1) \cap U, \Gamma(y_2)\big) \le M\|y_1 - y_2\|. \tag{2.5}$$

The above condition is usually called the Aubin continuity and the maps satisfying this property are called Aubin continuous maps [4,6].

We will also need the Lemma [5]:

Lemma 2.1. *Let $(\overline{x}, \overline{y}) \in Graph(f + F)$ and f be a Fréchet differentiable operator in an open neighborhood D of $\overline{x}$, whose derivative ∇f is continuous at $\overline{x}$.*

If F has a closed graph and the map $(f + F)^{-1}$ is Aubin continuous at $(\overline{y}, \overline{x})$, then there exist positive constants r, s and M such that for every $x \in U(\overline{x}, r)$ if

$$P_x = \big[f(x) + \nabla f(x)(\cdot - x) + F(\cdot)\big]^{-1}, \tag{2.6}$$

then

$$e\big(P_x(y') \cap U(\overline{x}, r), P_x(y'')\big) \le M\|y' - y''\| \quad \text{for all } y', y'' \in U(\overline{x}, r). \tag{2.7}$$

3. Convergence analysis of method (1.2)

We show the main result of the study:

Theorem 3.1. *Let x^* be a solution of (1.1) for $y = 0$, f a Fréchet-differentiable operator in an open neighborhood D of x^*, and F a set-valued map with a closed graph. We suppose that the Fréchet-derivative ∇f of f is (ε, x^*) center-continuous and satisfies condition (1.4) on D. Further suppose that the map $(f + F)^{-1}$ is Aubin continuous at $(0, x^*)$.*

Then, there exist positive constants σ, and b such that for every $y \in U(0, b)$ and $x_0 \in U(x^*, \sigma)$ there exists a Newton sequence $\{x_n\}$ for (1.1) *generated by (1.2), starting from* x_0 *which converges to a solution* x *of (1.1) for* y.

Moreover, there exists a constant α *such that*

$$\|x_{n+1} - x\| \leq \alpha \|x_n - x\| \quad (n \geq 0). \tag{3.1}$$

Proof. In view of the Aubin continuity of $(f+F)^{-1}$ at $(0, x^*)$ with constants ℓ, m and modulus c we deduce that for all y_1 and $y_2 \in U(0, m)$ and for all $x \in (f+F)^{-1}(y_1) \cap U(x^*, \ell)$ there exists $\overline{x} \in (f+F)^{-1}(y_2)$ satisfying

$$\|x - \overline{x}\| \leq c\|y_1 - y_2\|. \tag{3.2}$$

Letting $\delta = m$, $y_1 = 0$, $y_2 = y$, $x = x^*$ and $\overline{x} = x$ in the above assertion we obtain the existence of $\delta > 0$ such that for every $y \in U(0, \delta)$ there exists $x \in (f+F)^{-1}(y) \cap U(x^*, c\|y\|)$.

Let us assume that σ and b satisfy:

$$\sigma \leq \frac{r}{2}, \tag{3.3}$$

$$b \leq \min\left\{\frac{s}{2}, \delta, \frac{r}{2c}\right\}, \tag{3.4}$$

$$\sigma + cb \leq \min\left\{\left(\frac{s}{4L_0}\right)^{\frac{1}{1+\lambda_0}}, \left(\frac{r}{4ML_0}\right)^{\frac{1}{1+\lambda_0}}, \left(\frac{1}{2ML_0}\right)^{\frac{1}{\lambda_0}}\right\}, \tag{3.5}$$

where r, s and M are given by Lemma 2.1 with $\overline{x} = x^*$ and $\overline{y} = 0$.

We shall show that for a suitable initial point x_0, we can obtain x_1 satisfying (1.2) and (3.1) with $n = 0$.

Let $x_0 \in U(x^*, \sigma)$, $y \in U(0, b)$ and $x \in (f+F)^{-1}(y) \cap U(x^*, c\|y\|)$. Note that we have $\|x - x^*\| \leq cb \leq r$. In view of $y \in f(x) + F(x)$ it follows

$$x \in P_{x_0}\big(y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)\big) \cap U(x^*, r). \tag{3.6}$$

Using (1.4), (3.4)–(3.6) we obtain in turn

$$\begin{aligned}
&\|y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)\| \\
&\le \|y\| + \left\{ \int_0^1 \|[\nabla f(x + t(x_0 - x)) - \nabla f(x^*)]dt\| \right. \\
&\qquad \left. + \|\nabla f(x^*) - \nabla f(x_0)\| \right\} \|x - x_0\| \\
&\le b + L_0 \left[\int_0^1 (t\|x_0 - x^*\| + (1-t)\|x - x^*\|)^{\lambda_0} dt \right. \\
&\qquad \left. + \|x_0 - x^*\|^{\lambda_0} \right] \|x - x_0\| \\
&\le b + L_0 \big[(\sigma + cb)^{\lambda_0} + \sigma^{\lambda_0}\big] \|x - x_0\| \\
&\le b + 2L_0(\sigma + cb)^{\lambda_0} \|x - x_0\| \qquad (3.7)\\
&\le b + 2L_0(\sigma + cb)^{1+\lambda_0} \le \frac{s}{2} + \frac{s}{2} = s, \qquad (3.8)
\end{aligned}$$

which implies that $z = y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0) \in U(0, s)$. Since $x_0 \in U(x^*, \sigma) \subset U(x^*, r)$ and $(f + F)^{-1}$ is Aubin continuous at $(0, x^*)$ it follows from Lemma 2.1 that

$$\begin{aligned}
&e\big(P_{x_0}(z) \cap U(x^*, r), P_{x_0}(y)\big) \\
&\qquad \le M\| - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)\|. \qquad (3.9)
\end{aligned}$$

Therefore, there exists $x_1 \in P_{x_0}(y)$ such that

$$\begin{aligned}
\|x - x_1\| &\le M\| - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)\| \\
&\le 2ML_0(\sigma + cb)^{\lambda_0}\|x - x_0\|. \qquad (3.10)
\end{aligned}$$

In view of $x \in U(x^*, cb)$ and $\|x_1 - x^*\| \le \|x - x_1\| + \|x - x^*\|$, we get in turn:

$$\|x^* - x_1\| \le 2ML_0(\sigma + cb)^{\lambda_0}\|x - x_0\| + cb \le \frac{r}{2} + \frac{r}{2} = r, \qquad (3.11)$$

which implies $x_1 \in U(x^*, r)$.

Assuming the existence of $x_1, x_2, \ldots, x_k$ elements of $U(x^*, r)$ satisfying (1.2) and (3.1) we shall show that x_{k+1} does. In view of (3.4) we get

$$\|x - x_j\| \le 2ML_0(\sigma + cb)^{1+\lambda_0} \quad \text{for all } 0 \le j \le k. \qquad (3.12)$$

We shall show

$$x \in P_{x_k}\big(y - f(x) + f(x_k) + \nabla f(x_k)(x - x_k)\big) \cap U(x^*, r). \qquad (3.13)$$

Using Lemma 2.1 and (3.5) we can obtain in turn

$$\begin{aligned}(3.14)\qquad &\|y - f(x) + f(x_k) + \nabla f(x_k)(x - x_k)\| \\ &\leq 2L_0(\sigma + cb)^{\lambda_0}\|x - x_k\| + b \\ &\leq 2L_0(\sigma + cb)^{\lambda_0} 2ML_0(\sigma + cb)^{1+\lambda_0} + b \\ &\leq \frac{s}{2} + \frac{s}{2} = s,\end{aligned}$$

which implies the existence of $x_{k+1} \in P_{x_k}(y)$ such that

$$\begin{aligned}(3.15)\qquad \|x - x_{k+1}\| &\leq M\| - f(x) + f(x_k) + \nabla f(x_k)(x - x_k)\| \\ &\leq \alpha\|x - x_k\|,\end{aligned}$$

where,

$$\alpha = 2ML_0(\sigma + cb)^{\lambda_0},$$

which completes the induction for (3.1).

Finally we shall show $\{x_n\}$ $(n \geq 0)$ is a convergent sequence. Let $\varepsilon > 0$ be such that $2M\varepsilon < 1$. By the center-continuity of ∇f at x^*, and (1.4) we have:

$$(3.16)\qquad \|\nabla f(u) - \nabla f(x^*)\| \leq \varepsilon \quad \text{for all } u \in U(x^*, r),$$

by restricting $r \in \big(0, \big(\frac{\varepsilon}{L_0}\big)\frac{1}{\lambda_0}\big]$. Moreover we also have that for $x_k \in U(x^*, r)$:

$$\begin{aligned}(3.17)\qquad \|x_{k+1} - x_k\| &\leq M\|f(x_k) - f(x_{k-1}) - \nabla f(x^*)(x_k - x_{k-1})\| \\ &\quad + M\|(\nabla f(x^*) - \nabla f(x_{k-1}))(x_k - x_{k-1})\| \\ &\leq 2M\varepsilon\|x_k - x_{k-1}\| \leq \cdots \leq (2M\varepsilon)^k\|x_1 - x_0\|.\end{aligned}$$

It follows by (3.17) that sequence $\{x_k\}$ is Cauchy in a Banach space X and as such it converges to x.

That completes the proof of the theorem. □

References

[1] I.K. Argyros, *The Newton–Kantorovich method under mild differentiability conditions and the Ptâk error estimates*, Monatsh. Math. **101** (1990), 175–193.

[2] I.K. Argyros, *A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space*, J. Math. Anal. Appl. **298** (2004), 374–397.

[3] I.K. Argyros, *Newton Methods*, Nova Science Publ. Inc., New York, 2005.

[4] A.L. Dontchev, *Local convergence of the Newton method for generalized equations*, C. R. Acad. Sci. Paris, Sér. I Math. **332** (1996), 327–331.

[5] A.L. Dontchev *Uniform convergence of the Newton method for Aubin continuous maps*, Serdica Math. J. **22** (1996), 385–398.
[6] A.D. Ioffe and V.M. Tikhomirov, *Theory of Extremal Problems*, North Holland, Amsterdam, 1979.
[7] A. Pietrus, *Does Newton's method for set-valued maps converge uniformly in mild differentiability context?*, Rev. Colombiana Mat. **30** (2000), 49–56.
[8] R.T. Rockafellar, *Lipschitz properties of multifunctions*, Nonlinear Anal. **9** (1985), 867–885.

Cameron University
Department of Mathematical Sciences
Lawton, OK 73505, USA
E-mail: iargyros@cameron.edu