Revista Colombiana de Matemáticas
Volumen 34 (2000), páginas 49–56

Does Newton's method for set-valued maps converges uniformly in mild differentiability context?

Alain Pietrus

Université de Poitiers, Futuroscope Chasseneuil, FRANCE

Abstract. In this article, we study the existence of Newton-type sequence for solving the equation $y \in f(x) + F(x)$ where y is a small parameter, f is a function whose Fréchet derivative satisfies a Hölder condition of the form $||\nabla f(x_1) - \nabla f(x_2)|| \leq K||x_1 - x_2||^d$ and F is a set-valued map between two Banach spaces X and Y. We prove that the Newton-type method $y \in f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1})$, is locally convergent to a solution of $y \in f(x) + F(x)$ if the set valued map $\Big(f(x^*)+\nabla f(x^*)(\cdot - x^*)+F(\cdot)\Big)^{-1}$ is Aubin continuous at $(0, x^*)$ where x^* is a solution of $0 \in f(x) + F(x)$. Moreover, we show that this convergence is superlinear uniformly in the parameter y and quadratic when $d = 1$.

Key words and phrases. Set-valued maps, Aubin continuity, generalized equations, Newton's method, superlinear uniform convergence.

1991 Mathematics Subject Classification. Primary 47HO4. Secondary 65K10.

1. Introduction

In a previous paper [P1], we have studied a Newton-type method for solving generalized equation of the form:

$$\text{Find} \quad x \in X \quad \text{such that} \quad 0 \in f(x) + F(x) \tag{1}$$

where f is a function and F is a set-valued map.

In this study X and Y are two Banach spaces, $f : X \longrightarrow Y$ is Fréchet differentiable function and $F : X \longrightarrow 2^Y$ a multi-valued function with closed graph.

Let us recall that equation (1) is an abstract model for various problems.

- When $F = 0$, (1) is an equation.
- When F is the positive orthant in $\mathbb{R}^m$, (1) is a system of inequalities.
- When F is the normal cone to a convex and closed set in X, (1) may represent variational inequalities.

For other examples, the reader could refer to [D2].

Generally, for solving equations of kind (1), the authors (see [D1]) in order to obtain convergence have to suppose that ∇f (the Fréchet derivative of f) is Lipschitz (as in the classical Newton method) in a neighborhood Ω of a solution of (1). In [P1], we show that equation (1) can be solved by a Newton type method in the case where ∇f satisfies a Hölder condition on Ω:

$\exists K > 0,\ d \in (0, 1]$ such that

$$||\nabla f(x_1) - \nabla f(x_2)|| \leq K||x_1 - x_2||^d \ \forall x_1, x_2 \in \Omega. \quad (2)$$

The question of the existence of a Newton type method to solve perturbed equation associated to (1) is quite natural. For this question Dontchev in [D2] give an affirmative answer keeping the same assumption on ∇f as in [D1].

Our aim in this paper is to show that the answer remain true if ∇f satisfies condition (2). The results obtained in this study include some results of Dontchev obtained in [D2].

The perturbed equation mentioned above is in fact equation (1) in which 0 is replaced by a parameter y. The resulting problem is

$$\text{Find} \quad x \in X \quad \text{such that} \quad y \in f(x) + F(x) \quad (3)$$

The Newton-type method for solving (1) is: If x_k is the current iterate, the next iterate is found from the relation:

$$0 \in f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad k = 0, 1, \ldots \quad (4)$$

where $\nabla f(x)$ is the Fréchet derivative of f at x and x_0 is a given starting point.

It is obvious that the Newton-type method for solving (3) is the Newton-type method for solving (1) in which 0 is replaced by y. This method reads:

If x_k is the current iterate, the next iterate is found from the relation:

$$y \in f(x_k) + \nabla f(x_k)(x_{k+1} - x_k) + F(x_{k+1}), \quad k = 0, 1, \ldots \quad (5)$$

Note that when $F = \{0\}$, (4) is the classical Newton method for solving $f(x) = 0$ whereas (5) is the classical Newton method for solving $y = f(x)$ and if (4) represents the variational system associated with an optimization problem, then (5) is the corresponding version of the sequential quadratic programming method.

In the literature, the class of functions described by (2) is often called $\mathcal{C}^{1,d}$. Note that when $d = 1$, we have the Lipschitz condition for ∇f.

Condition (2) has been considered by several authors for solving operator equations. In [A], when f satisfies condition (2), one can solve $f(x) + g(x) = 0$ in some closed ball with the help of the iteration

$$x_{n+1} = x_n - f'(x_n)^{-1}\Big(f(x_n) + g(x_n)\Big). \tag{6}$$

Let us precise that when $d = 1$, the above iteration has been considered by Nguen and Zabrejko [NZ].

Condition (2) is often called the Vertgejm condition, in fact the first basic results are due to Vertgejm [V1], [V2]. In [AELZ], condition (2) gives new results on the approximation of Newton-Kantorovich and furnishes applications to nonlinear integral equations.

Throughout this paper all the norms are denoted by $||\ .\ ||$, the distances by dist, the inverse of a map G by G^{-1} and we will denote by $B_r(x)$ the closed ball centered at x with radius r.

Recall that a set-valued map Γ from Y to the subsets of 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

$$\sup_{z\in\Gamma(y_1)\cap U} \operatorname{dist}(z, \Gamma(y_2)) \le M||y_1 - y_2|| \quad \forall y_1 \text{ and } y_2 \text{ in } \quad V. \tag{7}$$

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

$$||z_1 - z_2|| \le M||y_1 - y_2||.$$

Let A and C be two subsets of X, if we denote by $e(C, A)$ the excess from the set A to the set C (semi-distance of Haussdorff)

$$e(C, A) = \sup\{\operatorname{dist}(x, A) : x \in C\},$$

then we have an equivalent definition of a M-pseudo-Lipschitz set-valued map, replacing (7) by:

$$e(\Gamma(y_1) \cap U, \Gamma(y_2)) \le M||y_1 - y_2||. \tag{8}$$

In [DR], the above property is called the Aubin continuity and the maps satisfying this property are called Aubin continuous maps. In [DH] the above property has been used in order to study the problem of inverse for set-valued maps. Let us precise also that the constant M is often called the modulus of Aubin continuity.

For more information about the Aubin continuity, the reader could refer to [AF], [M] and [R].

In the present paper, we follow an idea of the author in [D2], to show that the Aubin continuity is a sufficient property to allows equation (1) to be stable under small perturbation. what is the minimum one can hope for a numerical process.

2. Uniform convergence in mild differentiability context

In the case when $y = 0$ in equation (4), we have showed in a previous paper [P1] that if x^* is a solution of (1), the Aubin continuity of $(f+F)^{-1}$ at $(0, x^*)$ gives the existence of a superlinear Newton sequence which converges to x^*. We are going to show that this result remains true if we replace 0 by some small y.

The main result of this study is the following:

Theorem. *Let x^* be a solution of (1) , f a function which is Fréchet differentiable in an open neighborhood Ω of x^* and F a set-valued map with closed graph. We suppose that the Fréchet derivative ∇f of f is continuous and satisfies condition (2) in Ω with constant K.*

If we suppose that the map $(f+F)^{-1}$ is Aubin continuous at $(0, x^*)$. Then there exists positive constants σ, and b such that for every $y \in B_b(0)$ and $x_0 \in B_\sigma(x^*)$ there exists a Newton sequence (x_n) for (3) defined by (5), starting from x_0 and which converge to a solution x of (3) for y.

Furthermore, there exists a constant γ such that

$$||x_{k+1} - x|| \leq \gamma ||x_k - x||^{1+d} \tag{9}$$

that is (x_k) is superlinearly convergent to x.

As it has been announced in the introduction, this theorem differs from this given by Dontchev because of the properties of the Fréchet derivative ∇f of f. In [D2], the author suppose that ∇f is Lipschitz continuous in Ω, whereas here, we only suppose that ∇f is Hölder continuous in Ω that is more general and in our opinion very interesting in applications. Obviously, when the Hölder exponent is equal to one, we find again the Dontchev result.

3. Proof of the theorem

Before proving the theorem, some results will be useful for the sequel.

Lemma 1. *Let U be a convex set. If $||\nabla f(x_1) - \nabla f(x_2)|| \le K||x_1 - x_2||^d$ ($d \in (0,1]$), $\forall x_1, x_2 \in U$, then*

$$||f(x_2) - f(x_1) - \nabla f(x_1)(x_2 - x_1)|| \le \frac{1}{d+1} K ||x_2 - x_1||^{1+d}.$$

Proof. We have

$$f(x_2) - f(x_1) - \nabla f(x_1)(x_2 - x_1) = \int_0^1 \nabla f(tx_2 + (1-t)x_1) dt (x_2 - x_1) - \int_0^1 \nabla f(x_1) dt (x_2 - x_1).$$

From this, we deduce

$$||f(x_2) - f(x_1) - \nabla f(x_1)(x_2 - x_1)|| \le ||x_2 - x_1|| \int_0^1 ||\nabla f(tx_2 + (1-t)x_1) - \nabla f(x_1)|| dt,$$

and thus

$$||f(x_2) - f(x_1) - \nabla f(x_1)(x_2 - x_1)|| \le K||x_2 - x_1||^{1+d} \int_0^1 t^d dt = \frac{K}{1+d} ||x_2 - x_1||^{1+d}. \quad ☑$$

Lemma 2 [D2]. *Let $(\tilde{x}, \tilde{y}) \in$ Graph$(f + F)$ and f be a function which is Fréchet differentiable in an open neighborhood Ω of $\tilde{x}$, whose derivative ∇f is continuous at $\tilde{x}$.*

If we suppose that F has closed graph and the map $(f + F)^{-1}$ is Aubin continuous at $(\tilde{x}, \tilde{y})$, then there exist positive constants r, s and M such that for every $x \in B_r(\tilde{x})$ if $P_x = [f(x) + \nabla f(x)(. - x) + F(.)]^{-1}$, then

$$e\Big(P_x(y') \cap B_r(\tilde{x}), P_x(y'')\Big) \le M||y' - y''||$$

for every $y', y'' \in B_s(\tilde{x})$.

Proof of the theorem. Firstly, let us remark that the Aubin continuity of $(f + F)^{-1}$ at $(0, x^*)$ with constants l, m and modulus c implies that for all

y_1 and $y_2 \in B_m(0)$ and for all $x_1 \in (f+F)^{-1}(y_1) \cap B_l(x^*)$ the existence of $x_2 \in (f+F)^{-1}(y_2)$ which satisfies $||x_1 - x_2|| \le c||y_1 - y_2||$.

Taking $\delta = m$, $y_1 = 0$, $y_2 = y$, $x_1 = x^*$ and $x_2 = x$ in the above assertion, we obtain the existence of $\delta > 0$ such that for every $y \in B_\delta(0)$ there exists $x \in (f+F)^{-1}(y) \cap B_{c||y||}(x^*)$.

Now, let us assume that σ and b satisfy the following:

(i) $\sigma \le \frac{r}{2}$,

(ii) $b \le \min\{\frac{s}{2}, \delta, \frac{r}{2c}\}$,

(iii) $cb + \sigma \le \min\Big\{\left(\dfrac{s(1+d)}{2K}\right)^{\frac{1}{1+d}}, \left(\dfrac{r(1+d)}{2MK}\right)^{\frac{1}{1+d}}, \left(\dfrac{(1+d)}{MK}\right)^{\frac{1}{d}}\Big\}$

where r, s and M are given by lemma 2 with $\tilde{x} = x^*$ and $\tilde{y} = 0$.

For every $y \in B_b(0)$, we have to prove the existence of a Newton sequence (x_n) which converge to x which is solution of (3).

We proceed by induction for the rest of the proof. More precisely, we are going to show that starting from a suitable x_0, we obtain x_1 which verify (4) and (9) with $k = 0$ and so and so for all $k \in \mathbb{N}$.

Let $x_0 \in B_\sigma(x^*)$, $y \in B_b(0)$ and $x \in (f+F)^{-1}(y) \cap B_{c||y||}(x^*)$ which allows $||x - x^*|| \le cb \le r$. Let us also remark that $y \in f(x) + F(x)$ is equivalent to

$$x \in P_{x_0}\big(y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)\big) \cap B_r(x^*). \tag{10}$$

and with lemma 1, we obtain

$$\begin{aligned}||y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)|| &\le \frac{K}{1+d}||x - x_0||^{\frac{1}{1+d}} + b \\ &\le \frac{K}{1+d}(cb + \sigma) + b.\end{aligned}$$

The use of the hypotheses (ii) and (iii) gives

$$||y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)|| \le \frac{s}{2} + \frac{s}{2} = s. \tag{11}$$

The inequality (11) means that $z = y - f(x) + f(x_0) + \nabla f(x_0)(x - x_0) \in B_s(x^*)$. Since $x_0 \in B_\sigma(x^*) \subset B_r(x^*)$ and $(f+F)^{-1}$ is Aubin continuous at $(0, x^*)$ thus an application of lemma 2 gives

$$e\bigg(P_{x_0}(z) \cap B_r(x^*), P_{x_0}(y)\bigg) \le M|| - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)||. \tag{12}$$

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

$$||x - x_1|| \le M|| - f(x) + f(x_0) + \nabla f(x_0)(x - x_0)||. \tag{13}$$

Thanks to lemma 1, (13) becomes

$$||x - x_1|| \leq \frac{MK}{1+d}||x - x_0||^{1+d} \tag{14}$$

and since $x \in B_{cb}(x^*)$ and $||x_1 - x^*|| \leq ||x - x_1|| + ||x - x^*||$, we obtain

$$||x^* - x_1|| \leq \frac{MK}{1+d}||x - x_0||^{1+d} + cb \leq \frac{r}{2} + \frac{r}{2} = r$$

thus $x_1 \in B_r(x^*)$.

Let us suppose that we have proved the existence of $x_1,\ x_2, \ldots x_k$ element of $B_r(x^*)$ and which verify the relation (4) and (9). We are going to show that we can find x_{k+1} with the same property. Firstly, it is easy to see by induction using (ii) that

$$||x - x_l|| \leq \frac{MK}{1+d}(cb + \sigma)^{1+d},\ \forall\, 2 \leq l \leq k. \tag{15}$$

Starting with x_k, as before, we show that

$$x \in P_{x_k}\big(y - f(x) + f(x_k) + \nabla f(x_k)(x - x_k)\big) \cap B_r(x^*). \tag{16}$$

Furthermore, using lemma 1,(iii) and the recursion relation, we obtain

$$\begin{aligned}
||y - f(x) + f(x_k) + \nabla f(x_k)(x - x_k)|| &\leq \frac{K}{1+d}||x - x_k||^{\frac{1}{1+d}} + b \\
&\leq \frac{K}{1+d}\Big(\frac{MK}{d+1}(cb+\sigma)^{d+1}\Big)^{1+d} + b \\
&\leq \frac{K}{1+d}\big(cb+\sigma\big)^{1+d} + b \\
&\leq \frac{s}{2} + \frac{s}{2} = s.
\end{aligned}$$

By lemma 2 there exists a Newton iterate $x_{k+1} \in P_{x_k}$ such that

$$||x - x_{k+1}|| \leq M|| - f(x) + f(x_k) + \nabla f(x_k)(x - x_k)|| \leq \frac{MK}{1+d}||x - x_k||^{1+d} \tag{17}$$

That gives the inequality (3) of the theorem at the step $k+1$.

Now, let us show that the previous sequence is convergent.

Let $\varepsilon > 0$ be such that $M\varepsilon < 1$. Since ∇f is continuous, we can suppose that

$$||\nabla f(z) - \nabla f(u)|| \leq \varepsilon\ \forall\, z, u\ \in B_r(x^*).$$

We also have all the x_k in $B_r(x^*)$ and they satisfy

$$||x_{k+1} - x_k|| \leq M||f(x_k) - f(x_{k-1}) - \nabla f(x_{k-1})(x_k - x_{k-1})||$$

This inequality and the continuity of ∇f give

$$||x_{k+1} - x_k|| \leq M\varepsilon||x_{k-1} - x_k|| \leq \cdots \leq (M\varepsilon)^k||x_1 - x_0||$$

This last inequality show that (x_k) is a Cauchy sequence thus (x_k) converges to x and the proof of the theorem is complete. ☑

References

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

[AELZ] J. APPELL, E. DE PASCALE, J. V. LYSENKO & P. P. ZABREJKO, *New results on Newton-Kantorovich approximations with applications to nonlinear integral equations*, Numer. Funct. Anal. and Optimiz. **18** no. 1 and 2 (1997), 1–17.

[D1] A.L. DONTCHEV, *Local convergence of the Newton method for generalized equations*, C. R. Acad. Sc. Paris série I **322** (1996), 327–329.

[D2] A. L. DONTCHEV, *Uniform convergence of the Newton method for Aubin continuous maps*, Serdica Math. J. **22** (1996), 385–398.

[DH] A. L. DONTCHEV & W. W. HAGER, *An inverse function theorem for set-valued maps*, Proc. Amer. Math. Soc. **121** (1994), 481–489.

[DR] A. L. DONTCHEV, R. T. ROCKAFELLAR, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim. **6** (1996), 1087–1105.

[IT] A. D. IOFFE & V. M. TIKHOMIROV, *Theory of extremal problems*, North Holland, 1979.

[M] B. S. MORDUKHOVICH, *Complete characterization of openness, metric regularity and Lipschitzian properties of multifunctions*, Trans. Amer. Math. Soc. **340** (1993), 1–36.

[NZ] D. F. NGUEN & P. P. ZABREJKO, *The majorant method in the theory of Newton-Kantorovich approximations and the Ptâk error estimates*, Numer. Funct. Anal. Optimiz. **9** (1987), 671–684.

[P1] A. PIETRUS, *Generalized equations under mild differentiability conditions*, To appear in the Revista Real Acad. de Ciencias, Madrid.

[R] R. T. ROCKAFELLAR, *Lipschitz properties of multifunctions*, Nonlinear Anal. **9** (1985), 867–885.

[V1] B. A. VERTGEJM, *On conditions for the applicability of Newton's method*, in Russian, Dokl. Akad. Nauk SSSR **110** (1956), 719–722.

[V2] B. A. VERTGEJM, *On some methods of solution of nonlinear operator equations in Banach spaces*, in Russian, Uspekhi Mat. Nauk **12** (1957), 166–169, Translated in AMS Transl. II. **16**, 378–382.

(Recibido en mayo de 2000; revisado por el autor en septiembre de 2000)

UFR SCIENCES-SP2MI
DÉPARTEMENT DE MATHÉMATIQUES
TÉLÉPORT 2, BOULEVARD MARIE ET PIERRE CURIE
B.P. 30179, 86962 FUTUROSCOPE CHASSENEUIL CEDEX
FRANCE

e-mail: pietrus@mathrs.sp2mi.univ-poitiers.fr