RACSAM
Rev. R. Acad. Cien. Serie A. Mat.
VOL. **102** (2), 2008, pp. 355–361
Matemática Aplicada / Applied Mathematics

On the convergence of some methods for variational inclusions

C. Jean–Alexis and A. Piétrus

Abstract. In this paper, we study variational inclusions of the following form $0 \in f(x) + g(x) + F(x) \quad (*)$ where f is differentiable in a neighborhood of a solution x^* of $(*)$ and g is differentiable at x^* and F is a set-valued mapping with closed graph acting in Banach spaces. The method introduced to solve $(*)$ is superlinear and quadratic when ∇f is Lipschitz continuous.

Sobre la convergencia de algunos métodos para inclusiones variacionales

Resumen. En este artículo se estudian inclusiones variacionales de la forma $0 \in f(x) + g(x) + F(x) \quad (*)$ donde f es diferenciable en un entorno de la solución x^* de $(*)$, g es diferenciable en x^* y F es una aplicación con gráfica cerrada entre espacios de Banach. El método introducido para resolver $(*)$ es superlineal y cuadrático cuando ∇f es continuo y verifica la condición de Lipschtz.

1 Introduction

In this study, we are concerned with the problem of approximating a solution x^* of the following variational inclusion

$$0 \in f(x) + g(x) + F(x) \tag{1}$$

where $f\colon X \to Y$ is differentiable in a neighborhood of a solution x^* of (1), $g\colon X \to Y$ is differentiable at x^* but may be not differentiable in a neighborhood of x^* while $F\colon X \to 2^Y$ denotes a set-valued mapping with closed graph and X, Y are two Banach spaces.

Inclusion (1) can be viewed as a perturbed problem of the following one

$$0 \in f(x) + F(x), \tag{2}$$

the function g being the perturbation function.

For solving (2), several iterative methods have been presented. When ∇f is locally Lispchitz on a neighborhood of a solution x^* of (2), Dontchev [5] established a quadratically convergent Newton-type method under a pseudo-Lipschitz property for set-valued mappings and in [6], he proved the stability of the method. Following Dontchev's method, Piétrus [17] obtained superlinear convergence when ∇f is Hölder on a neighborhood of x^* and he also showed the stability of the method in this mild differentiability context (see [18]).

Presentado por / Submitted by Ildefonso Díaz.
Recibido / Received: 1 de agosto de 2008. *Aceptado / Accepted*: 15 de octubre de 2008.
Palabras clave / Keywords: Set-valued mappings, variational inclusions, superlinear convergence, pseudo-Lipschitz mappings, divided difference.
Mathematics Subject Classifications: 49J53, 47H04, 65K10.

When f is only continuous and differentiable at x^*, for solving (2), Hilout and Piétrus [12] considered the sequence

$$\begin{cases} x_0 \text{ and } x_1 \text{ are given starting points} \\ y_k = \alpha x_k + (1-\alpha)x_{k-1}; & \alpha \text{ is fixed in } [0,1[\\ 0 \in f(x_k) + [y_k, x_k; f](x_{k+1} - x_k) + F(x_{k+1}) \end{cases} \tag{3}$$

where $[y_k, x_k; f]$ is a first order divided difference of f on the points y_k and x_k. This operator will be defined in section 2. They prove the superlinear convergence of this method.

For solving (1), when $F = \{0\}$, f is differentiable and g is continuous function admitting first and second order divided differences, Cătinas [4] proposed a combination of Newton's method with the secant's method . An extension of this method to variational inclusions is studied in [10] where Geoffroy et al. proved the superlinear convergence under an assumption on the second order divided difference. So, these two methods are valid if g posseses a second order divided difference.

According to this idea of combination, we propose, in this paper, a method of the form

$$0 \in f(x_k) + g(x_k) + \Big(\nabla f(x_k) + [2x_{k+1} - x_k, x_k; g]\Big)(x_{k+1} - x_k) + F(x_{k+1}). \tag{4}$$

For proving the convergence of (4), we don't use the concept of second order divided difference which is a paramount notion in the work evocated previously. We obtain an order of convergence better than those obtained by the authors in [10, 11]; moreover, in the Lipschitz case, we have a quadratic convergence. Because of the presence of x_{k+1} in the divided difference instead of x_k, from a numerical viewpoint, (4) seems to be better than the method presented in [11].

This work is organized as follows: in section 2, we recall a few preliminary results about regularity of set-valued mappings, divided differences and we state a fixed point theorem which is very important for obtaining the algorithm. In section 3, we show the existence and the convergence of the sequence defined by (4).

2 Preliminary results

Let us give some notation. We denote by $\mathbb{B}_r(x)$ the closed ball centered at x with radius r, by $\|\cdot\|$ all the norms, by $\mathrm{L}(X,Y)$ the space of linear operators from X to Y. The distance from a point $x \in X$ and a subset $A \subset X$ is defined as $\operatorname{dist}(x,A) = \inf\limits_{a\in A} \|x-a\|$, the excess from the set A to the set C is defined by $e(C,A) = \sup_{x\in C} \operatorname{dist}(x,A)$ and the graph of a set-valued mapping $F\colon X \to 2^Y$ is denoted by $\operatorname{Gph} F = \{\,(x,y) \in X\times Y \mid y \in F(x)\,\}$.

Recall the definition of a pseudo-Lipschitz (or Lipschitz-like) set-valued mapping:

Definition 1 *A set-valued mapping $\Gamma\colon X \to 2^Y$ is said to be M-pseudo-Lipschitz around $(x_0,y_0) \in \operatorname{Gph}\Gamma$ if there exist constants a and b such that*

$$e(\Gamma(x_1) \cap \mathbb{B}_a(y_0), \Gamma(x_2)) \le M\|x_1 - x_2\|.$$

The pseudo-Lipschitz property has been introduced by Aubin, it is the reason for which this property is sometimes called "Aubin continuity", see [2, 3].

Characterizations of this property are obtained by Rockafellar [19] using the Lipschitz continuity of the distance function $\operatorname{dist}(y,\Gamma(x))$ and by Mordukhovich [15] via the concept of coderivative of set-valued mappings.

The pseudo-Lipschitz property of a set-valued mapping Γ is also equivalent to the metric regularity of Γ^{-1} and to the openness with linear rate of Γ^{-1}, see [8, 9, 14, 16, 20].

Definition 2 *An operator $[x, y; f] \in \mathrm{L}(X, Y)$ is called a divided difference of first order of the function f at the points x and y in X ($x \neq y$) if the following property holds:*

$$[x, y; f](y - x) = f(y) - f(x). \tag{5}$$

Let us remark that if f is Fréchet-differentiable at x then $[x, x; f] = \nabla f(x)$ where ∇f is the Fréchet derivative of f.

For a best understanding of the theory of divided differences of nonlinear operators, the reader could refer to [1].

Next comes an extension to the set-valued setting of a local version of the Banach fixed point theorem which has been proved in [7].

Lemma 1 *Let ϕ be a set-valued mapping from X into the closed subsets of X, let $\eta_0 \in X$ and let r and λ be such that $0 < \lambda < 1$ and*

(a) $\mathrm{dist}(\eta_0, \phi(\eta_0)) < r(1 - \lambda)$,

(b) $e(\phi(x_1) \cap \mathbb{B}_r(\eta_0), \phi(x_2)) \leq \lambda \, \|x_1 - x_2\|, \ \ \forall x_1, x_2 \in \mathbb{B}_r(\eta_0)$,

then ϕ has a fixed point in $\mathbb{B}_r(\eta_0)$. That is, there exists $x \in \mathbb{B}_r(\eta_0)$ such that $x \in \phi(x)$. If ϕ is single–valued, then x is the unique fixed point of ϕ in $\mathbb{B}_r(\eta_0)$.

This lemma is a generalization of fixed point theorem in [13], where in assertion (b) the excess e is replaced by the Hausdorff distance.

3 Convergence analysis

Throughout this section, we suppose that:

(H1) The function $f\colon X \to Y$ is Fréchet-differentiable and its derivative is (L, p)-Hölder on a neighborhood Ω of a solution x^* of (1) that means :

$$\exists L > 0, \quad \|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|^p, \qquad p \in (0, 1], \quad \forall x, y \in \Omega.$$

(H2) The function $g\colon X \to Y$ is Fréchet-differentiable at x^* and admits a first order divided difference satisfying the following condition :

there exists $\nu > 0$ such that for all x, y, u and $v \in \Omega$ ($x \neq y$, $u \neq v$),

$$\|[x, y; g] - [u, v; g]\| \leq \nu(\|x - u\|^p + \|y - v\|^p), \qquad p \in [0, 1].$$

That means that the first order divided difference of g satisfies a (ν, p)-Hölder condition.

(H3) The set-valued mapping $F\colon X \to 2^Y$ with closed graph is such that $[f + g + F]^{-1}$ is pseudo-Lipschitz around $(0, x^*)$.

Remark 1 *From a result in* [7], *the assumption* (H3) *implies that the map $[f(x^*) + \nabla f(x^*)(. - x^*) + g(.) + F(.)]^{-1}$ is pseudo-Lipschitz around $(0, x^*)$.*

In the sequel, we denote by M its modulus.

The main theorem of this study reads as follow.

Theorem 1 *Under the assumptions* (H1)–(H3) *and for every $c > M\left(\dfrac{L}{p+1} + \nu\right)$, one can find $\delta > 0$ such that for every distinct starting point $x_0 \in \mathbb{B}_\delta(x^*)$, there exists a sequence (x_k) defined by* (4) *which satisfies*

$$\|x_{k+1} - x^*\| \leq c\|x_k - x^*\|^{p+1}. \tag{6}$$

Before proving this theorem, let us introduce some notation. First, define the set-valued mapping Q from X into the subsets of Y by

$$Q(x) = f(x^*) + g(x) + \nabla f(x^*)(x - x^*) + F(x).$$

Then for $k \in \mathbb{N}$ and $x_k \in X$, set

$$\begin{aligned} Z_k(x) = f(x^*) + g(x) + \nabla f(x^*)(x - x^*) - f(x_k) - g(x_k) \\ - \big(\nabla f(x_k) + [2x - x_k, x_k; g]\big)(x - x_k). \end{aligned}$$

Finally, define $\phi_k \colon X \to 2^X$ by $\phi_k(x) = Q^{-1}[Z_k(x)]$.

PROOF OF THEOREM 1. By assumption (H3), $Q^{-1}(\cdot)$ is M-pseudo-Lipschitz around $(0, x^*)$ then there exist positive constants a and b such that

$$e(Q^{-1}(y') \cap \mathbb{B}_a(x^*), Q^{-1}(y'')) \leq M\|y' - y''\|, \qquad \forall y', y'' \in \mathbb{B}_b(0). \tag{7}$$

Let us choose $\delta > 0$ such that

$$\delta < \min\left\{a, \left(\frac{b(p+1)}{L(2^{p+1}+1) + 2^{p+1}\nu(p+1)}\right)^{\frac{1}{p+1}}, \left(\frac{1}{c}\right)^{\frac{1}{p}}\right\}. \tag{8}$$

We apply Lemma 1 to the map ϕ_0 with $\eta_0 = x^*$ and r and λ are numbers to be set. Let us check that both assertions (a) and (b) of Lemma 1 hold. According to the definition of the excess e, we have

$$\begin{aligned} \mathrm{dist}(x^*, \phi_0(x^*)) &\leq e(Q^{-1}(0) \cap \mathbb{B}_\delta(x^*), \phi_0(x^*)) \\ &\leq e(Q^{-1}(0) \cap \mathbb{B}_\delta(x^*), Q^{-1}[Z_0(x^*)]). \end{aligned}$$

Moreover, for all $x_0 \in \mathbb{B}_\delta(x^*)$ ($x_0 \neq x^*$) we have

$$\begin{aligned} \|Z_0(x^*)\| &= \|f(x^*) + g(x^*) - f(x_0) - g(x_0) - \nabla f(x_0)(x^* - x_0) \\ &\quad - [2x^* - x_0, x_0; g](x^* - x_0)\| \\ &\leq \|f(x^*) - f(x_0) - \nabla f(x_0)(x^* - x_0)\| \\ &\quad + \|g(x^*) - g(x_0) - [2x^* - x_0, x_0; g](x^* - x_0)\| \\ &\leq \frac{L}{p+1}\|x_0 - x^*\|^{p+1} + \|[x^*, x_0; g] - [2x^* - x_0, x_0; g]\|\|x^* - x_0\| \\ &\leq \frac{L}{p+1}\|x_0 - x^*\|^{p+1} + \nu\|x_0 - x^*\|^p\|x_0 - x^*\| \\ &\leq \left(\frac{L}{p+1} + \nu\right)\delta^{p+1}. \end{aligned}$$

Then (8) yields $Z_0(x^*) \in \mathbb{B}_b(0)$. Hence, by (H3), we have

$$\mathrm{dist}(x^*, \phi_0(x^*)) \leq M\|Z_0(x^*)\| \leq M\left(\frac{L}{p+1} + \nu\right)\|x_0 - x^*\|^{p+1}.$$

Since $c > M\left(\frac{L}{p+1} + \nu\right)$, one can find $\lambda \in]0, 1[$ such that $c(1 - \lambda) > M\left(\frac{L}{p+1} + \nu\right)$. Hence,

$$\mathrm{dist}(x^*, \phi_0(x^*)) < c(1 - \lambda)\|x_0 - x^*\|^{p+1}. \tag{9}$$

By setting $r = r_0 = c\|x_0 - x^*\|^{p+1}$, we can deduce from the previous inequality that the assertion (a) of Lemma 1 is satisfied.

Now, we show that condition (b) of Lemma 1 is satisfied. It is clear that $r_0 < \delta < a$ and moreover for $x \in \mathbb{B}_\delta(x^*)$, we have

$$\begin{aligned}
\|Z_0(x)\| &= \|f(x^*) + g(x) + \nabla f(x^*)(x - x^*) - f(x_0) - g(x_0) \\
&\qquad - \nabla f(x_0)(x - x_0) - [2x - x_0, x_0; g](x - x_0)\| \\
&\le \|f(x^*) - f(x) - \nabla f(x^*)(x^* - x)\| + \|f(x) - f(x_0) - \nabla f(x_0)(x - x_0)\| \\
&\qquad + \|g(x) - g(x_0) - [2x - x_0, x_0; g](x - x_0)\| \\
&\le \frac{L}{p+1}\|x^* - x\|^{p+1} + \frac{L}{p+1}\|x - x_0\|^{p+1} + \nu\|x - x_0\|^p\|x - x_0\| \\
&\le \left(\frac{L(2^{p+1}+1)}{p+1} + 2^{p+1}\nu\right)\delta^{p+1}.
\end{aligned}$$

Then by (8), we deduce that for all $x \in \mathbb{B}_\delta(x^*)$, $Z_0(x) \in \mathbb{B}_b(0)$; it follows that for all $x', x'' \in \mathbb{B}_{r_0}(x^*)$ we have

$$\begin{aligned}
&e(\phi_0(x') \cap \mathbb{B}_{r_0}(x^*), \phi_0(x'')) \\
&\qquad \le e(\phi_0(x') \cap \mathbb{B}_\delta(x^*), \phi_0(x'')) \\
&\qquad \le M\|Z_0(x') - Z_0(x'')\| \\
&\qquad \le M\Big[\|\nabla f(x^*) - \nabla f(x_0)\|\|x' - x''\| + \|g(x') - g(x'') - [2x' - x_0, x_0; g](x' - x_0) \\
&\qquad\qquad + [2x'' - x_0, x_0; g](x'' - x_0)\|\Big] \\
&\qquad \le ML\|x^* - x_0\|^p\|x' - x''\| + M\|[x', x''; g] - [2x' - x_0, x_0, g]\|\,\|x' - x''\| \\
&\qquad\qquad + M\|[2x'' - x_0, x_0; g] - [2x' - x_0, x_0; g]\|\,\|x'' - x_0\| \\
&\qquad \le ML\delta^p\|x' - x''\| + M\nu(\|x_0 - x'\|^p + \|x'' - x_0\|^p)\|\,\|x' - x''\| \\
&\qquad\qquad + M\nu\|2x'' - 2x'\|^p\,\|x'' - x_0\| \\
&\qquad \le M\big(L\delta^p + 2^{p+1}\nu\delta^p + 2^{2p}\nu\delta^p\big)\|x' - x''\| \\
&\qquad \le M(L + 2^{p+1}\nu + 4^p\nu)\delta^p\|x' - x''\|.
\end{aligned}$$

Without loss of generality, we can assume that $\delta < \left(\dfrac{\lambda}{M(L + 2^{p+1}\nu + 4^p\nu)}\right)^{\frac{1}{p}}$ thus condition (b) of Lemma 1. Since both conditions of Lemma 1 are fulfilled, we can deduce the existence of a fixed point $x_1 \in \mathbb{B}_{r_0}(x^*)$ for the map ϕ_0 which implies that the inequality (6) is checked for $k = 0$.

Proceeding by induction, keeping $\eta_0 = x^*$ and setting $r = r_k = c\|x_k - x^*\|^{p+1}$; we prove, in the same way, the existence of a fixed point $x_{k+1} \in \mathbb{B}_{r_k}(x^*)$ for ϕ_k. This fact implies that

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

Hence the proof of Theorem 1 is complete. ■

Concluding remarks

- If $p = 1$ then the Fréchet derivative of f and the fisrt order divided difference of g satisfy a Lipschitz condition, we obtain the quadratic convergence of the method (4).
- If in the assumptions (H1) and (H2), we take different exponents (p_1 and p_2) for inequalities satisfied by ∇f and the divided difference of g, we find similar results of order $(\beta + 1)$ by setting $\beta = \min\{p_1, p_2\}$.

- If we use a weaker condition (called Lipschitz center-Hölder) on the first order divided difference of g by replacing the inequality in (H2) by

$$\|[x,y;g]-\nabla g(x^*)\| \le \nu(\|x-x^*\|^p+\|y-x^*\|^p), \quad p\in[0,1],$$

we find similar results but assuming that g admits a bounded second order divided difference.

References

[1] ARGYROS, I. K., (2007). *Computational theory of iterative methods*, Stud. Comp. Math., Elsevier, Amsterdam.

[2] AUBIN, J-P., (1984). Lipschitz behavior of solutions to convex minimization problems, *Math. Oper. Res.*, **9** , 87–111.

[3] AUBIN, J-P. AND FRANKOWSKA, H., (1990). *Set-valued analysis*, Birkhäuser, Boston.

[4] CĂTINAS, E., (1994). On some iterative methods for solving nonlinear equations, *Rev. Anal. Nume. Theo. Approx.*, **23**, 17–53.

[5] DONTCHEV, A. L., (1996). Local convergence of the Newton method for generalized equation, *C.R.A.S Paris*, **322**, Serie I , 327.-331.

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

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

[8] DONTCHEV, A. L. AND ROCKAFELLAR, R. T., (2004). Regularity and conditioning of solution mappings in variational analysis, *Set-Valued Anal.*, **1–2**, 79–109.

[9] DONTCHEV, A. L., QUINCAMPOIX, M. AND ZLATEVA, N., (2006). Aubin criterion for metric regularity, *J. Convex Anal.*, **13** , 2, 281–297.

[10] GEOFFROY, M. H. AND PIÉTRUS, A., (2004). Local convergence of some iterative methods for solving generalized equations, *J. Math. Anal. Appl.*, **290**, 497–505.

[11] HILOUT, S., JEAN-ALEXIS, C. AND PIÉTRUS, A. A secant-type method for variational inclusions under center-Hölder conditions, Preprint.

[12] HILOUT, S. AND PIÉTRUS, A., (2006). A semilocal convergence of the secant-type method for solving a generalized equations, *Positivity*, **10**, 673–700.

[13] IOFFE, A. D. AND TIKHOMIROV, V. M., (1979). *Theory of extremal problems*, Studies in Mathematics and its Applications, Amsterdam, New-York.

[14] MORDUKHOVICH, B. S., (1993). Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, *Trans. Amer. Math. Soc.*, **340**, 1, 1–35.

[15] MORDUKHOVICH, B. S., (1994). Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis, *Trans. Amer. Math. Soc.*, **343**, 609–657.

[16] MORDUKHOVICH, B. S., (2006). *Variational Analysis and Generalized Differentiation I: Basic Theory*, Springer-Verlag, Berlin.

[17] PIÉTRUS, A., (2000). Does Newton's method for set-valued maps converges uniformly in mild differentiability context?, *Rev. Colombiana Mat.*, **32**, 49–56.

[18] PIÉTRUS, A., (2000). Generalized equations under mild differentiability conditions, *Rev. R. Acad. Cienc. Exact. Fis. Nat.*, **94**, 1, 15–18.

[19] ROCKAFELLAR, R. T., (1984). Lipschitzian properties of multifunctions, *Nonlinear Anal.*, **9**, 867–885.

[20] ROCKAFELLAR, R. T. AND WETS, R., (1998). *Variational analysis*, Ser. Com. Stu. Math. **317**, Springer-Verlag, Berlin.

C. Jean–Alexis
Laboratoire Analyse, Optimisation, Contrôle,
Université des Antilles et de la Guyane,
Département de Mathématiques et Informatique,
Campus de Fouillole,
F–97159 Pointe–à–Pitre, France,
celia.jean-alexis@univ-ag.fr

A. Piétrus
Laboratoire Analyse, Optimisation, Contrôle,
Université des Antilles et de la Guyane,
Département de Mathématiques et Informatique,
Campus de Fouillole,
F–97159 Pointe–à–Pitre, France,
apietrus@univ-ag.fr