ON A SECANT-LIKE METHOD FOR SOLVING GENERALIZED EQUATIONS

Ioannis K. Argyros, Lawton, Saïd Hilout, Poitiers

(Received April 26, 2007)

Abstract. In the paper by Hilout and Piétrus (2006) a semilocal convergence analysis was given for the secant-like method to solve generalized equations using Hölder-type conditions introduced by the first author (for nonlinear equations). Here, we show that this convergence analysis can be refined under weaker hypothesis, and less computational cost. Moreover finer error estimates on the distances involved and a larger radius of convergence are obtained.

Keywords: secant-like method, generalized equations, Aubin continuity, radius of convergence, divided difference

MSC 2000: 65G99, 65K10, 49M15

1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution x^* of the generalized equation

$$0 \in f(x) + G(x), \tag{1.1}$$

where f is a continuous function from a Banach space X into a Banach space Y and G is a set-valued map from X into the subsets of X with closed graph. Many problems from applied sciences can be formulated like equation (1.1), see [4], [11]–[21]. A survey on results concerning solution of the generalized equation (1.1) can be found in [1], [4], [5]–[21], and the references there.

As in the work [12] we use the secant-like method

$$0 \in f(x_k) + [y_k, x_k; f](x_{k+1} - x_k) + G(x_{k+1}), \tag{1.2}$$

where for x_0 and x_1 being given initial guesses, $y_k = \alpha x_k + (1-\alpha)x_{k-1}$ and $\alpha \in [0,1)$ is fixed.

A local convergence analysis was provided for method (1.2) under the following Hölder-type conditions introduced by us in [1], [4], [5] to solve nonlinear equations:

$$\|[x, y; f] - [u, v; f]\| \leqslant \nu(\|x-u\|^p + \|y-v\|^p) \quad \text{for all } x, y, u, v \in X,\ x \neq y,\ u \neq v,\ \text{some } \nu > 0 \text{ and } p \in [0,1], \tag{1.3}$$

where $[x, y; f] \in \mathcal{L}(X, Y)$, the space of bounded linear operators from X into Y, is called a divided difference of order one at the points x and y, satisfying

$$[x, y; f](y-x) = f(y) - f(x) \quad \text{for all } x, y \text{ in } X \text{ with } x \neq y. \tag{1.4}$$

Note that if f is Fréchet-differentiable, then $[x, x; f] = \nabla f$ (see [2]–[4]).

In general, ν and p in (1.3) are not easy to compute. This is our motivation for introducing weaker hypotheses

$$\|[x, y; f] - [z, y; f]\| \leqslant \nu_1 \|x - z\|^p \tag{1.5}$$

and

$$\|[x^*, x; f] - [y, x; f]\| \leqslant \nu_0 \|y - x^*\|^p \tag{1.6}$$

for all x, y, z in X, $x \neq y$, $z \neq y$, $x \neq x^*$ and some $\nu_1 > 0$, $\nu_0 > 0$.

Note that in general

$$\nu_0 \leqslant \nu_1 \leqslant \nu \tag{1.7}$$

holds, and ν/ν_0, ν/ν_1, ν_1/ν_0 can be arbitrarily large [2]–[4]. Note that parameters ν_0 and ν_1 are easier to determine than ν. Moreover, as it turns out, conditions (1.5) and (1.6) are actually needed in the proof of semilocal convergence of the secant-like method (1.2). Using the above observations we provide under weaker hypotheses and at less computational cost a local convergence analysis with the following advantages: finer error estimates on the distances $\|x_n - x^*\|$ $(n \geqslant 0)$, and a larger radius of convergence which allows a larger choice of initial guesses x_0 and x_1.

These observations are very important in computational mathematics [2], [4].

2. Preliminaries and assumptions

In order to make the paper as self-contained as possible, we recall some definitions that can also be found in [5], [6], [10]–[12], [15], [21].

Definition 2.1. The distance from a point x to a set A in a metric space (Z, ϱ) is defined by $\operatorname{dist}(x, A) = \inf\{\varrho(x, y),\ y \in A\}$ and the excess e from the set A to a set C is given by $e(C, A) = \sup\{\operatorname{dist}(x, A),\ x \in C\}$. Let $\Lambda\colon X \rightrightarrows Y$ be a set-valued map; we denote $\operatorname{gph} \Lambda = \{(x, y) \in X \times Y,\ y \in \Lambda(x)\}$ and $\Lambda^{-1}(y) = \{x \in X,\ y \in \Lambda(x)\}$. We denote by $B_r(x)$ the closed ball centered at x with radius r.

Definition 2.2. A set-valued Λ is pseudo-Lipschitz around $(x_0, y_0) \in \operatorname{gph} \Lambda$ with modulus M if there exist constants a and b such that

$$\sup_{z \in \Lambda(y') \cap B_a(y_0)} \operatorname{dist}(z, \Lambda(y'')) \leqslant M\|y' - y''\| \quad \text{for all } y' \text{ and } y'' \text{ in } B_b(x_0). \tag{2.1}$$

Using the excess, we have an equivalent definition replacing the inequality (2.1) by

$$e(\Lambda(y') \cap B_a(y_0), \Lambda(y'')) \leqslant M\|y' - y''\| \quad \text{for all } y' \text{ and } y'' \text{ in } B_b(x_0). \tag{2.2}$$

This property is also called "Aubin continuity" [6]. Characterizations of the pseudo-Lipschitz property were also obtained by Rockafellar using the Lipschitz continuity of the distance function $\operatorname{dist}(y, \Lambda(x))$ around (x_0, y_0) in [20] and by Mordukhovich in [16] via the concept of coderivative of multifunctions.

We need a lemma on fixed points whose proof can be found in [9], [15].

Lemma 2.3. *Let (Z, ϱ) be a complete metric space, φ a set-valued map from Z into the closed subsets of Z, let $\eta_0 \in Z$ and let r and λ be such that $0 \leqslant \lambda < 1$ and*
(a) $\operatorname{dist}(\eta_0, \varphi(\eta_0)) \leqslant r(1 - \lambda)$,
(b) $e(\varphi(x_1) \cap B_r(\eta_0), \varphi(x_2)) \leqslant \lambda\varrho(x_1, x_2),\ \forall x_1, x_2 \in B_r(\eta_0)$.

Then φ has a fixed point in $B_r(\eta_0)$. That is, there exists $x \in B_r(\eta_0)$ such that $x \in \varphi(x)$. If φ is single-valued, then x is the unique fixed point of φ in $B_r(\eta_0)$.

Throughout this work, the distance ϱ in Lemma 2.3 is replaced by the norm.

We make the following assumptions:

(H1) The set-valued map $(f(x^*) + G)^{-1}$ is pseudo-Lipschitz with modulus M around $(0, x^*)$.

(H2) For all $x, y \in V$ we have $\|[x, y; f]\| \leqslant d$, $\|f(x) - f(x^*)\| \leqslant d_0\|x - x^*\|$ and $Md < 1$.

3. Local convergence analysis for the secant-like method (1.2)

We need to introduce some notation. First, define a set-valued map $Q\colon\ X \rightrightarrows Y$ by

$$Q(x) = f(x^*) + G(x). \tag{3.1}$$

For $k \in \mathbb{N}^*$ and x_k, y_k defined in (1.2), we consider the quantity

$$Z_k(x) := f(x^*) - f(x_k) - [y_k, x_k; f](x - x_k). \tag{3.2}$$

Finally, define a set-valued map $\psi_k\colon\ X \rightrightarrows X$ by

$$\psi_k(x) := Q^{-1}(Z_k(x)). \tag{3.3}$$

We will show the following main local result for the method (1.2):

Theorem 3.1. *Let x^* be a solution of* (1.1). *Suppose that assumptions* (1.5), (1.6), (H1) *and* (H2) *are satisfied. For every $C > (1-Md)^{-1}M\nu_0[(1-\alpha)^p+\alpha^p] = C_0$, one can find $\delta > 0$ such that for every distinct starting points x_0 and x_1 in $B_\delta(x^*)$ there exists a sequence (x_k) defined by* (1.2) *which satisfies*

$$\|x_{k+1} - x^*\| \leqslant C\|x_k - x^*\| \max\{\|x_k - x^*\|^p, \|x_{k-1} - x^*\|^p\}. \tag{3.4}$$

To prove Theorem 3.1, we first prove the following proposition:

Proposition 3.2. *Under the assumptions of Theorem* 3.1, *one can find $\delta > 0$ such that for every distinct starting points x_0 and x_1 in $B_\delta(x^*)$ (x_0, x_1 and x^* distinct), the set-valued map ψ_1 has a fixed point x_2 in $B_\delta(x^*)$ satisfying*

$$\|x_2 - x^*\| \leqslant C\|x_1 - x^*\| \max\{\|x_1 - x^*\|^p, \|x_0 - x^*\|^p\}. \tag{3.5}$$

Note that the point x_2 is a fixed point of ψ_1 if and only if

$$0 \in f(x_1) + [y_1, x_1; f](x_2 - x_1) + G(x_2). \tag{3.6}$$

Once x_k is computed, we will show that the function ψ_k has a fixed point x_{k+1} in X. This process is useful for proving existence of a sequence (x_k) satisfying (1.2).

Proof of Proposition 3.2. Since the iterate y_1 in (1.2) is defined by $y_1 = \alpha x_1 + (1-\alpha)x_0$, it is clear that $y_1 \in B_\delta(x^*)$.

By hypothesis (H1) there exist positive numbers M, a and b such that

$$e(Q^{-1}(y') \cap B_a(x^*), Q^{-1}(y'')) \leqslant M\|y' - y''\|, \ \forall y', y'' \in B_b(0). \tag{3.7}$$

Fix $\delta > 0$ such that

$$\delta < \delta_0 = \min\left\{a;\ \sqrt[p+1]{\frac{b}{\nu_0((1-\alpha)^p + \alpha^p)}};\ \frac{1}{\sqrt[p]{C}};\ \frac{b}{2d_0};\ \sqrt[p+1]{\frac{b}{2^{p+2}\nu_1}}\right\}. \tag{3.8}$$

To prove Proposition 3.2 we intend to show that both the assertions (a) and (b) of Lemma 2.3 hold, where $\eta_0 := x^*$, φ is the function ψ_1 defined by (3.3) and r and λ are numbers to be set. According to the definition of the excess e, we have

$$\operatorname{dist}(x^*, \psi_1(x^*)) \leqslant e(Q^{-1}(0) \cap B_\delta(x^*), \psi_1(x^*)). \tag{3.9}$$

Moreover, for all points x_0 and x_1 in $B_\delta(x^*)$ (x_0, x_1 and x^* distinct) we have

$$\|Z_1(x^*)\| = \|f(x^*) - f(x_1) - [y_1, x_1; f](x^* - x_1)\|.$$

By assumption (1.6) we deduce

$$\begin{aligned}\|Z_1(x^*)\| &= \|([x^*, x_1; f] - [y_1, x_1; f])(x^* - x_1)\| \\ &\leqslant \|[x^*, x_1; f] - [y_1, x_1; f]\|\|x^* - x_1\| \\ &\leqslant \nu_0\|x^* - y_1\|^p\|x^* - x_1\| \\ &\leqslant \nu_0((1-\alpha)\|x^* - x_0\| + \alpha\|x^* - x_1\|)^p\|x^* - x_1\|.\end{aligned} \tag{3.10}$$

Thus

$$\|Z_1(x^*)\| \leqslant \nu_0[(1-\alpha)^p\|x^* - x_0\|^p + \alpha^p\|x^* - x_1\|^p]\|x^* - x_1\|. \tag{3.11}$$

Then (3.8) yields $Z_1(x^*) \in B_b(0)$. Hence from (3.7) one has

$$\begin{aligned}e(Q^{-1}(0) \cap B_\delta(x^*), \psi_1(x^*)) &= e(Q^{-1}(0) \cap B_\delta(x^*), Q^{-1}[Z_1(x^*)]) \\ &\leqslant M\nu_0[(1-\alpha)^p\|x^* - x_0\|^p + \alpha^p\|x^* - x_1\|^p]\|x^* - x_1\|.\end{aligned} \tag{3.12}$$

By (3.9) we get

$$\begin{aligned}\operatorname{dist}(x^*, \psi_1(x^*)) &\leqslant M\nu_0[(1-\alpha)^p\|x^* - x_0\|^p + \alpha^p\|x^* - x_1\|^p]\|x^* - x_1\| \\ &\leqslant M\nu_0[(1-\alpha)^p + \alpha^p]\|x^* - x_1\| \max\{\|x_1 - x^*\|^p, \|x_0 - x^*\|^p\}.\end{aligned} \tag{3.13}$$

Since $C(1-Md) > M\nu_0[(1-\alpha)^p+\alpha^p]$, there exists $\lambda \in [Md, 1[$ such that $C(1-\lambda) \geqslant M\nu_0[(1-\alpha)^p+\alpha^p]$ and

$$\operatorname{dist}(x^*, \psi_1(x^*)) \leqslant C(1-\lambda)\|x^*-x_1\| \max\{\|x_1-x^*\|^p, \|x_0-x^*\|^p\}. \tag{3.14}$$

By setting $r := r_1 = C\|x^*-x_1\| \max\{\|x_1-x^*\|^p, \|x_0-x^*\|^p\}$ we can deduce from the inequality (3.14) that the assertion (a) in Lemma 2.3 is satisfied.

Now, we show that condition (b) of Lemma 2.3 is satisfied.

By (3.8) we have $r_1 \leqslant \delta \leqslant a$ and, moreover, for $x \in B_\delta(x^*)$ we have

$$\begin{aligned} \|Z_1(x)\| &= \|f(x^*) - f(x_1) - [y_1, x_1; f](x-x_1)\| \\ &\leqslant \|f(x^*) - f(x)\| + \|[x, x_1; f] - [y_1, x_1; f]\|\|x-x_1\|. \end{aligned} \tag{3.15}$$

Using the assumptions (1.5) and (H2) we obtain

$$\begin{aligned} \|Z_1(x)\| &\leqslant d_0\|x^*-x\| + \nu_1\|x-y_1\|^p\|x-x_1\| \\ &\leqslant d_0\|x^*-x\| + \nu_1(\|x-x^*\| + \|x^*-y_1\|)^p\|x-x_1\| \\ &\leqslant d_0\delta + \nu_1(2\delta)^p 2\delta = d_0\delta + \nu_1 2^{p+1}\delta^{p+1}. \end{aligned} \tag{3.16}$$

Then by (3.8) we deduce that for all $x \in B_\delta(x^*)$ we have $Z_1(x) \in B_b(0)$. Then it follows that for all $x', x'' \in B_{r_0}(x^*)$ we have

$$e(\psi_1(x') \cap B_{r_1}(x^*), \psi_1(x'')) \leqslant e(\psi_1(x') \cap B_\delta(x^*), \psi_1(x'')),$$

which yields by (3.7)

$$\begin{aligned} e(\psi_1(x') \cap B_{r_1}(x^*), \psi_1(x'')) &\leqslant M\|Z_1(x') - Z_1(x'')\| \\ &\leqslant M\|[y_1, x_1; f](x''-x')\| \\ &\leqslant M\|[y_1, x_1; f]\|\|x''-x'\|. \end{aligned} \tag{3.17}$$

Using (H2) and the fact that $\lambda \geqslant Md$, we obtain

$$e(\varphi_0(x') \cap B_{r_1}(x^*), \psi_1(x'')) \leqslant Md\|x''-x'\| \leqslant \lambda\|x''-x'\| \tag{3.18}$$

and thus condition (b) of Lemma 2.3 is satisfied. Since both conditions of Lemma 2.3 are fulfilled, we can deduce existence of a fixed point $x_2 \in B_{r_1}(x^*)$ for the map ψ_1. Thus the proof of Proposition 3.2 is complete. □

P r o o f T h e o r e m 3.1. Proceeding by induction, keeping $\eta_0 = x^*$ and setting

$$r := r_k = C\|x^*-x_k\| \max\{\|x_k-x^*\|^p, \|x_{k-1}-x^*\|^p\},$$

the application of Proposition 3.2 to the map ψ_k gives the existence of a fixed point x_{k+1} for ψ_k which is an element of $B_{r_k}(x^*)$. This last fact gives the inequality (3.4) and the proof of Theorem 3.1 is complete. □

4. Concluding remarks

When $\alpha = 1$, our method is no longer applicable, but if we suppose that f is Fréchet-differentiable then (1.2) is equivalent to a Newton-type method for solving (1.1). In this case conditions on ∇f give quadratic convergence (see [8]) and superlinear convergence (see [18]) and in both cases the convergence is uniform (see [8] and [19]). These results were further improved in [5].

When $\alpha = 0$ the sequence (1.2) is reduced to the method introduced by M. Geoffroy and A. Piétrus in [11]. Let us note that the problem studied in [11] can be seen as a perturbation of (1.1) by a Fréchet differentiable function. In both cases, we obtain a superlinear convergence using different assumptions, but in the present paper the existence of second order divided differences is not required.

Finally, in order to compare our results with the corresponding ones in [12], under hypotheses (1.3), (H1) and (H2)$'$ given by

(H2)$'$ for all $x, y \in V$ we have $\|[x,y;f]\| \leqslant d$ and $Md < 1$,

let us define

$$\delta_0' = \min\left\{a;\ \sqrt[p+1]{\frac{b}{\nu((1-\alpha)^p + \alpha^p)}};\ \frac{1}{\sqrt[p]{C'}};\ \frac{b}{2d};\ \sqrt[p+1]{\frac{b}{2^{p+2}\nu}}\right\} \tag{3.19}$$

and

$$C_0' = \frac{M\nu[(1-\alpha)^p + \alpha^p]}{1 - Md}. \tag{3.20}$$

In view of (1.7), (3.8), (3.19), (3.20) and the definitions of C_0 and C' ($C' > C_0'$) we have

$$C_0 \leqslant C_0' \tag{3.21}$$

and

$$\delta_0' \leqslant \delta_0. \tag{3.22}$$

Note also that if strict inequality holds on the right hand side of inequality (1.7), then so it does in (3.21) and (3.22). Hence, the claims made in the introduction about the advantages of our approach over the corresponding ones in [12] have been justified.

References

[1] *I. K. Argyros*: A new convergence theorem for Steffensen's method on Banach spaces and applications. Southwest J. Pure Appl. Math. *1* (1997), 23–29. zbl

[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. zbl

[3] *I. K. Argyros*: New sufficient convergence conditions for the secant method. Czech. Math. J. *55* (2005), 175–187. zbl

[4] *I. K. Argyros*: Approximate Solution of Operator Equations with Applications. World Scientific Publ. Comp., New Jersey, USA, 2005. zbl

[5] *I. K. Argyros*: An improved convergence analysis of a superquadratic method for solving generalized equations. Rev. Colombiana Math. *40* (2006), 65–73.

[6] *J. P. Aubin, H. Frankowska*: Set-Valued Analysis. Birkhäuser, Boston, 1990. zbl

[7] *E. Cătinas*: On some iterative methods for solving nonlinear equations. Rev. Anal. Numér. Théor. Approx. *23* (1994), 17–53. zbl

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

[9] *A. L. Dontchev, W. W. Hager*: An inverse mapping theorem for set-valued maps. Proc. Amer. Math. Soc. *121* (1994), 481–489. zbl

[10] *M. H. Geoffroy, S. Hilout, A. Piétrus*: Acceleration of convergence in Dontchev's iterative method for solving variational inclusions. Serdica Math. J. *29* (2003), 45–54. zbl

[11] *M. H. Geoffroy, A. Piétrus*: Local convergence of some iterative methods for solving generalized equations. J. Math. Anal. Appl. *290* (2004), 497–505. zbl

[12] *S. Hilout, A. Piétrus*: A semilocal convergence analysis of a secant-type method for solving generalized equations. Positivity *10* (2006), 693–700. zbl

[13] *M. A. Hernández, M. J. Rubio*: Semilocal convergence of the secant method under mild convergence conditions of differentiability. Comput. Math. Appl. *44* (2002), 277–285. zbl

[14] *M. A. Hernández, M. J. Rubio*: ω-conditioned divided differences to solve nonlinear equations. Monografías del Semin. Matem. García de Galdeano *27* (2003), 323–330. zbl

[15] *A. D. Ioffe, V. M. Tihomirov*: Theory of Extremal Problems. North Holland, Amsterdam, 1979. zbl

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

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

[18] *A. Piétrus*: Generalized equations under mild differentiability conditions. Rev. Real. Acad. Ciencias de Madrid *94* (2000), 15–18. zbl

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

[20] *R. T. Rockafellar*: Lipschitzian properties of multifunctions. Nonlinear Analysis *9* (1984), 867–885. zbl

[21] *R. T. Rockafellar, R. J.-B. Wets*: Variational Analysis. A Series of Comprehensives Studies in Mathematics, Springer, 317, 1998. zbl

Authors' addresses: *Ioannis K. Argyros*, Cameron university, Department of Mathematics Sciences, Lawton, OK 73505, USA, e-mail: `ioannisa@cameron.edu`; *Saïd Hilout*, Département d'Applications des Mathématiques, Université de Poitiers, bd. Marie & Pierre Curie, Téléport 2, BP 30179, 86962 Futuroscope Chasseneuil Cedex, France, e-mail: `said.hilout@ac-poitiers.fr`.