East Asian Mathematical Journal
Vol. 25 (2009), No. 4, pp. 425–431

LOCAL CONVERGENCE OF NEWTON-LIKE METHODS FOR GENERALIZED EQUATIONS

Ioannis K. Argyros

Abstract. We provide a local convergence analysis for Newton-like methods for the solution of generalized equations in a Banach space setting. Using some ideas of ours introduced in [2] for nonlinear equations we show that under weaker hypotheses and computational cost than in [7] a larger convergence radius and finer error bounds on the distances involved can be obtained.

1. Introduction

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

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

where X, Y are Banach spaces, $f\colon X \to Y$ is a Fréchet-differentiable operator in a neighborhood U of x^*, $g\colon X \to Y$ is continuous at x^* and F denotes a set-valued map from X into the subsets of Y.

If $F = \{0\}$ and $g = 0$ equation (1) reduces to a regular nonlinear equation. If $F = \{0\}$ and $g \neq 0$ equation is again a regular nonlinear equation studied in [2] and the references there. Here we are interested in generating a sequence $\{x_n\}$ $(n \geq 0)$ approximating x^* in cases when $F = \{0\}$ and $g = 0$ or not.

The most popular method for approximating x^* is undoubtedly Newton-like method of the form

$$o \in f(x_n) + g(x_n) + \big(f'(x_n) + [x_{n-1}, x_n, g]\big)(x_{n+1} - x_n) + F(x_{n+1}) \tag{2}$$

where $f'(x)$ denotes the Fréchet-derivative of operator f and $[x, y; g]$ simply denoted by $[x, y]$ is the first order divided difference of g at the points x, y satisfying $[x, y] \in L(X, Y)$, and

$$[x, y](y - x) = g(y) - g(x) \quad \text{for } x \neq y. \tag{3}$$

If g is Fréchet-differentiable at $x \in X$ then $[x, x] = g'(x)$.

Received October 9, 2008; Accepted April 8, 2009.
2000 *Mathematics Subject Classification.* 65H10, 65G99, 47H17, 49M15.
Key words and phrases. Newton-like methods, Banach space, local convergence, radius of convergence, generalized equations, Fréchet derivative, Lipschitz/center-Lipschitz condition.

Geoffroy and Pietrus provided a local convergence analysis for method (2) in [7]. Here we are motivated by this paper, our work in [2] and optimization considerations. Using more precise error estimates and a combination of Lipschitz as well as center Lipschitz conditions on f' and g we provide a finer convergence analysis than before [5]–[7] with the advantages already stated in the abstract of this paper.

2. Local Convergence Analysis of Method (2)

We need the definition of a divided difference of order 2 [9], the definition Aubin continuity of a set-valued map [1] and a generalization of the Ioffe–Tikhomirov theorem on fixed points of operators [6], [8].

Definition 1. We say that an operator in $L(X, L(X, Y))$ denoted by $[x, y, z; g]$ or simply $[x, y, z]$ is called a divided difference of order two of the operator $y\colon X \to Y$ at the points $x, y, z \in X$ if

$$[x, y, z](z - x) = [y, z] - [x, y] \quad \text{for all distinct points } x,\ y \text{ and } z \text{ from } X. \tag{4}$$

If g is twice Fréchet-differentiable at $x \in X$ then

$$[x, x, x] = \frac{g''(x)}{2}\,.$$

Definition 2. A set-valued map $\Gamma X \leftrightarrows Y$ is said to be M-pseudo-Lipschitz about

$$(x_0, y_0) \in \operatorname{Graph}\Gamma = \{(x, y) \in X \times Y \mid y \in \Gamma(x)\}$$

if there exist neighborhoods V of y_0 and U of x_0 such that

$$e\big(\Gamma(v) \cap U, \Gamma(w)\big) \le M\|v - w\| \quad \text{for all } v, w \in V. \tag{5}$$

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

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

Lemma 3. *Let (X, ρ) be a Banach space, let T map X to the closed subsets of X, let $q_0 \in X$, and let $r > 0$, and $\lambda \in [0, 1)$ be such that the following hold true*:

$$\operatorname{dist}(q_0, T(q_0)) < r(1 - \lambda), \tag{6}$$

$$e\big(T(v) \cap U(q_0, r), T(w)\big) \le \lambda\rho(v, w) \quad \textit{for all } v, w \in U(q_0, r). \tag{7}$$

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

We will make the following assumptions:

(A_1) F has a closed graph;
(A_2) f is Fréchet differentiable in some neighborhood V of x^*;
(A_3) g is differentiable at x^*;

(A_4) f' is L-Lipschitz on V and L_0-center Lipschitz on V. That is there exist positive constants L and L_0 such that

$$\|F'(y_1) - F'(y_2)\| \leq L\|y_1 - y\| \tag{8}$$

and

$$\|F'(y) - F'(x^*)\| \leq L_0\|y - x^*\| \quad \text{for all } y, y_1, y_2 \in V; \tag{9}$$

(A_5) there exists a positive constant K such that for all $x, y, z \in V$,

$$\|[x, y, z]\| \leq K; \tag{10}$$

(A_6) the set-valued map

$$G(x)^{-1} = [f(x^*) + f'(x^*)(x - x^*) + g(x) + F(x)]^{-1} \tag{11}$$

is M-pseudo-Lipschitz around $(0, x^*)$.

We can state the main local convergence result for method (2):

Theorem 4. *Under assumptions* (A_1)–(A_6) *the following hold true*:

for every $c > M\big(\frac{L}{2} + K\big) = c_0$ *there exists* $\delta > 0$ *such that for any distinct initial guesses* $x_0, x_1 \in U(x^*, \delta)$, *there exists a sequence* $\{x_n\}$ $(n \geq 0)$ *generated by Newton-like method (2) such that*

$$\|x_{n+1} - x^*\| \leq c\|x_n - x^*\| \max\{\|x_n - x^*\|, \|x_{n-1} - x^*\|\} \ (n \geq 1). \tag{12}$$

Before starting the proof it is convenient to define operators R_n and T_n by

$$\begin{aligned} R_n(x) &= f(x^*) + g(x) + f'(x^*)(x - x^*) \\ &\quad - f(x_n) - g(x_n) - (f'(x_n) + [x_{n-1}, x_n])(x - x_n) \quad (n \geq 1) \end{aligned} \tag{13}$$

and

$$T_n(x) = G^{-1}[R_n(x)] \quad (n \geq 1). \tag{14}$$

Note that x_{k+1} is a fixed point of T_k if and only if $R_k(x_{k+1}) \in G(x_{k+1})$, i.e., if and only if

$$o \in f(x_k) + g(x_k) + \big(f'(x_k) + [x_{k-1}, x_k]\big)(x_{k+1} - x_k) + F(x_{k+1}). \tag{15}$$

We also need the auxiliary result:

Proposition 5. *Under the hypotheses of Theorem 4, there exists* $\delta > 0$ *such that for all* $x_0, x_1 \in U(x^*, \delta)$ $(x_0, x_1, x^*$ *distinct), the map* T_1 *has a fixed point* x_2 *in* $U(x^*, \delta)$ *satisfying*

$$\|x_2 - x^*\| \leq c\|x_1 - x^*\| \max\{\|x_1 - x^k\|, \|x_0 - x^*\|\}. \tag{16}$$

Proof. In view of (A_6) there exist positive constants a and b such that

$$\begin{aligned} &e\big(G^{-1}(y_1) \cap U(x^*, a), G^{-1}(y_2)\big) \\ &\qquad \leq \ M\|y_1 - y_2\| \quad \text{for all } y_1, y_2 \in U(0, b). \end{aligned} \tag{17}$$

Choose a fixed $\delta \in (0, \delta_0)$ where

$$\delta_0 = \min\left\{a, \frac{1}{c}, \left(\frac{2b}{4L + L_0 + 8K}\right)^{1/2}\right\}. \tag{18}$$

We shall show conditions (6) and (7) of Lemma 3 hold true where $q_0 = x^*$ and $T = T_1$, for some constants r and λ to be determined.

We first note that

$$\operatorname{dist}(x^*, T_1(x^*)) \leq e\big(G^{-1}(0) \cap U(x^*, \delta), T_1(x^*)\big). \tag{19}$$

Let $x_0, x_1 \in U(x^*, \delta)$ such that x_0, x_1 and x^* are distinct, then we obtain in turn by (3), (4), (8)–(10) and (18)

$$\begin{aligned}
\|R_1(x^*)\| &\leq f(x^*) + g(x^*) - f(x_1) - g(x_1) - (f'(x_1) + [x_0, x_1])(x^* - x_1))\| \\
&\leq \|f(x^*) - f(x_1) - f'(x_1)(x^* - x_1)\| \\
&\quad + \|g(x^*) - g(x_1) - [x_0, x_1](x^* - x_1)\| \\
&= \|f(x^*) - f(x_1) - f'(x_1)(x^* - x_1)\| \\
&\quad + \|[x_0, x_1, x^*](x^* - x_0)(x^* - x_1)\| \\
&\leq \frac{L}{2}\|x^* - x_1\|^2 + K\|x^* - x_0\| \cdot \|x^* - x_1\| \\
&\leq \left(\frac{L}{2}\|x^* - x_1\| + K\|x^* - x_0\|\right)\|x^* - x_1\| \\
&\leq \left(\frac{L}{2} + K\right)\delta\|x^* - x_1\| \leq \left(\frac{L}{2} + K\right)\delta^2 \leq b,
\end{aligned} \tag{20}$$

by the choice of δ.

In view of (17) we get

$$\begin{aligned}
&e\big(G^{-1}(0) \cap U(x^*, \delta), T_1(x^*)\big) \\
&\quad = e\big(G^{-1}(0) \cap U(x^*, \delta), G^{-1}[R_1(x^*)]\big) \\
&\quad \leq M\left(\frac{L}{2}\|x^* - x_1\| + K\|x^* - x_0\|\right)\|x^* - x_1\|.
\end{aligned} \tag{21}$$

Using (19) we obtain in turn

$$\begin{aligned}
\operatorname{dist}(x^*, T_1(x^*)) &\leq M\left[\frac{L}{2}\|x^* - x_1\| + K\|x^* - x_0\|\right]\|x^* - x_1\| \\
&\leq M\left(\frac{L}{2} + K\right)\|x^* - x_1\| \max\{\|x^* - x_0\|, \|x^* - x_1\|\}.
\end{aligned} \tag{22}$$

Choose c fixed and $c > M\big(\frac{L}{2} + K\big)$. Then there exist $\lambda \in (0,1)$ such that $M\big(\frac{L}{2} + K\big) \leq c(1-\lambda)$. That is

$$\operatorname{dist}(x^*, T_0(x^*)) \leq c(1-\lambda)\|x^* - x_1\| \max\{\|x^* - x_0\|, \|x^* - x_1\|\}. \tag{23}$$

Letting $q_0 = x^*$, $r = r_1 = c\|x^* - x_1\| \max\{\|x^* - x_0\|, \|x^* - x_1\|\}$ condition (6) holds true.

We shall show condition (7) also holds true. By $\delta c < 1$ and $x_0, x_1 \in U(x^*, \delta)$ we have $r_1 \le \delta \le a$. Let $x \in U(x^*, \delta)$, then we get in turn

$$\begin{aligned} \|R_1(x)\| &\le \|f(x^*) - f(x) - f'(x^*)(x^* - x)\| \\ &\quad + \|f(x) - f(x_1) - f'(x_1)(x - x_1)\| \\ &\quad + \|g(x) - g(x_1) - [x_0, x_1](x - x_1)\| \\ &\le \left(\frac{L_0 + 4L}{2} + 4K\right)\delta^2, \end{aligned} \tag{24}$$

which implies $z_1(x) \in U(0, b)$ for $x \in U(x^*, \delta)$ by the choice of δ.

Let $w, z \in U(x^*, r_1)$ then by (17)

$$\begin{aligned} & e\big(T_1(w) \cap U(x^*, r_1), T_1(z)\big) \\ &\quad \le e\big(T_1(w) \cap U(x^*, \delta), T_1(z)\big) \le M\|R_1(w) - R_1(z)\| \\ &\quad \le M\|(F'(x^*) - F'(x_1))(w - z)\| + M\|g(w) - g(z) - [x_0, x_1](w - z)\| \\ &\quad \le M\|(F'(x^*) - F'(x_1))(w - z)\| + M\|\big([x_1, z, w](w - x_1) \\ &\qquad + [x_0, x_1, w](w - x_0)\big)(z - w)\| \le M\delta(L_0 + 4K)\|z - w\|. \end{aligned} \tag{25}$$

Without loss of generality we may assume

$$\delta < \frac{\lambda}{M(L_0 + 4K)} = \delta_1, \tag{26}$$

which implies condition (7). By Lemma 3 there exists a fixed point $x_2 \in U(x^*, r_1)$ for the map T_1.

That completes the proof of Proposition 5. □

Proof of Theorem 4. Using induction on $k \ge 1$ and setting

$$q_0 = x^*, \quad r_k = c\|x_k - x^*\| \max\{\|x_{k-1} - x^*\|, \|x_k - x^*\|\}$$

we conclude by Proposition 5 that the map T_k has a fixed point x_{k+1} in $U(x^*, r_k)$. It follows that

$$\|x_{k+1} - x^*\| \le c\|x_k - x^*\| \max\{\|x_k - x^*\|, \|x_{k-1} - x^*\|\} \ \ (k \ge 1).$$

That completes the proof of Theorem 4. □

As in [7] we consider two modifications of method (2):

Remark 6. (a) If (2) is replaced by

$$o \in f(x_n) + g(x_n) + \big(f'(x_n) + [x_0, x_n]\big)(x_{n+1} - x_n) + F(x_{n+1}) \tag{27}$$

then under hypotheses (A_1)–(A_6) the conclusions of Theorem 4 hold with (12) replaced by

$$\|x_{n+1} - x^*\| \le c\|x_n - x^*\| \max\{\|x_n - x^*\|, \|x_0 - x^*\|\}. \tag{28}$$

Note that regular-false method (27) [3] is slower than method (2).

(b) If (2) is replaced by

$$o \in f(x_n) + y(x_n) + \big(f'(x_n) + [x_{n+1}, x_n]\big)(x_{n+1} - x_n) \tag{29}$$

or

$$o \in f(x_n) + f'(x_n)(x_{n+1} - x_n) + g(x_{n+1}) + F(x_{n+1}) \tag{30}$$

then if $c > c_0$ is replaced by $c > c_1 = \frac{ML}{2}$ and (H_5) is dropped under hypotheses (A_1)–(A_4) and (A_6) the conclusions of Theorem 4 hold true with (12) replaced by the faster (quadratic convergence):

$$\|x_{n+1} - x^*\| \le c\|x_n - x^*\|^2. \tag{31}$$

Remark 7. In general

$$L_0 \le L \tag{32}$$

holds and $\frac{L}{L_0}$ can be arbitrarily large [2]–[4]. If equality holds in (32), then our results reduce to the corresponding ones in [7]. Otherwise they constitute an improvement. Indeed denote by δ_0^0 and δ_1^1 used in [7] and given by

$$\delta_0^0 = \min\left\{a, \frac{1}{c}, \left(\frac{2b}{4L + L + 8K}\right)^{1/2}\right\} \tag{33}$$

and

$$\delta_1^1 = \frac{\lambda}{M(L + 4K)}\,. \tag{34}$$

It follows from (18), (26), (33) and (34) that

$$\delta_0^0 \le \delta_0 \tag{35}$$

and

$$\delta_1^1 \le \delta_1. \tag{36}$$

Note also that the choice of δ influences the choice of c. In view of (35) and (36) we conclude that under the same computational cost (since in practice the computation of constant L requires the computation of L_0) and hypotheses a larger convergence radius δ and a smaller ratio c can be obtained. These observations are very important in computational mathematics [2]–[11].

References

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

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

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

[4] Argyros, I. K., *On the secant method for solving nonsmooth equations*, J. Math. Anal. Appl. (to appear, 2006).

[5] Dontchev, A. L. and Hager, W. W., *An inverse function theorem for set-valued maps*, Proc. Amer. Math. Soc. **121** (1994), 481–484.

[6] Dontchev, A. L., *Local convergence of the Newton method for generalized equations*, C. R. Acad. Sci. Paris, t. 332, Série I (1996), 327–331, Mathematical Analysis.

[7] Geoffory, M. H. and Pietrus, A., *A local convergence of some iterative methods for generalized equations*, J. Math. Anal. Applic. **290** (2004), 497–505.

[8] Ioffe, A. D. and Tikhomirov, V. M., *Theory of Extremal Problems*, North Holland, Amsterdam, 1979.

[9] Kantorovich, L. V. and Akilov, G. P., *Functional Analysis in Normed Spaces*, Pergamon Press, Oxford, 1982.

[10] Rockafellar, R. T., *Lipschitz properties of multifunctions*, Nonlinear Analysis, **9** (1985), 867–885.

[11] Xiao, B. and Harker, P. T., *A nonsmooth Newton method for variational inequalities I: Theory*, Math. Programming, **65** (1994), 151–194.

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