**Revista Colombiana de Matemáticas**
Volumen 40 (2006), páginas 65–73

# An improved convergence analysis of a superquadratic method for solving generalized equations

IOANNIS K. ARGYROS
Cameron University, USA

ABSTRACT. We provide a finer local convergence analysis than before [6]–[9] of a certain superquadratic method for solving generalized equations under Hölder continuity conditions.

*Keywords and phrases.* Superquadratic convergence, generalized equations, radius of convergence, Aubin continuity, pseudo-Lipschitz map.

*2000 Mathematics Subject Classification.* Primary: 65K10, 65G99. Secondary: 47H04, 49M15.

RESUMEN. Nosotros hacemos un análisis de convergencia local más fino que el proporcionado antes de [6]–[9] de cierto método supercuadrático para resolver ecuaciones generalizadas bajo ciertas condiciones de continuidad de Hölder.

## 1. Introduction

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

$$o \in F(x) + G(x), \tag{1.1}$$

where $F$ is a twice Fréchet differentiable operator defined on a Banach space $X$ with values in a Banach space $Y$, and $G$ is a set-valued map from $X$ to the subsets of $Y$.

Local results providing sufficient conditions for the existence of $x^*$ have been provided by several authors using various iterative methods and hypotheses [2]–[9], [11]. Here in particular, we use the method

$$o \in F(x_n) + \nabla F(x_n)(x_{n+1} - x_n) + \frac{1}{2}\nabla^2 F(x_n)(x_{n+1} - x_n)^2 + G(x_{n+1}) \tag{1.2}$$

to generate a sequence approximating $x^*$.

In the paper by Geoffroy and Pietrus [9] local convergence results were provided for method (1.2) using Hölder continuity conditions on $\nabla^2 F$. Here we are motivated by this paper, our paper [1], and optimization considerations.

In particular using the same hypotheses but more precise error bounds we provide a larger convergence radius and finer error bounds on the distances $\|x_n - x^*\|$ $(n \geq 0)$.

Some numerical examples are provided to justify our theoretical results. The same examples are used to compare favorably our results with the corresponding ones in [9].

The paper is organized as follows: In Section 2 we have collected a number of necessary results [6], [9], [10] needed in our local convergence analysis appearing in Section 3.

## 2. Preliminaries

We need a definition about the Aubin continuity [5]–[7]:

**Definition 2.1.** *A set-valued map $\Gamma\colon X \to Y$ is said to be $M$-pseudo-Lipschitz around $(x_0, y_0) \in$ graph $F = \{(x,y) \in X \times Y \mid y \in \Gamma(x)\}$ if there exist neighborhoods $V$ of $x_0$ and $U$ of $y_0$ such that*

$$\sup_{y \in \Gamma(u) \cap U} \operatorname{dist}(y, \Gamma(v)) \leq M\|u - v\| \quad \textit{for all } x, y \in V. \tag{2.1}$$

The Aubin continuity of $\Gamma$ is equivalent to the openess with linear rate of $\Gamma^{-1}$ and the metric regularity of $\Gamma^{-1}$.

Let $A$ and $C$ be two subsets of $X$. Then the excess $e$ from the set $A$ to the set $C$ is given by

$$e(C, A) = \sup_{x \in C} \operatorname{dist}(x, A). \tag{2.2}$$

Estimate (2.1) using (2.2) can be written

$$c\,(\Gamma(u) \cap U, \Gamma(v)) \leq M\|u - v\| \quad \text{for all } u, v \in V. \tag{2.3}$$

We also need a lemma about fixed points [10]:

**Lemma 2.2.** *Let $(X, \rho)$ be a Banach space, let $T$ be a map from $X$ into the closed subsets of $X$, let $p \in X$ and let $r$ and $\lambda$ be such that $0 \leq \lambda < 1$, and*

$$\operatorname{dist}(p, T(p)) \leq r(1 - \lambda), \tag{2.4}$$

$$e(T(u) \cap U(p,r), T(v)) \leq \lambda \rho(u, v), \quad \textit{for all } u, v \in U(p, r) \tag{2.5}$$

*where*

$$U(p, r) = \{x \in X \ \|x - p\| \leq r\}. \tag{2.6}$$

*Then $T$ has a fixed point in $U(p,r)$. Moreover if $T$ is single-valued, then $x$ is the unique fixed point of $T$ in $U(p,r)$.*

Let $x^*$ be a solution of (1.1). We assume:

(A1) $F$ is Fréchet-differentiable on some neighborhood $V$ of $x^*$;
(A2) $\nabla^2 F$ is bounded by $L$ on $V$ and $\|\nabla^2 F(x^*)\| \leq L_0$;

(A3) $\nabla^2 F$ is $\alpha$-Hölder on $V$ with constant $K$, i.e.

$$\|\nabla^2 F(x) - \nabla^2 F(y)\| \leq K\|x-y\|^\alpha \quad \text{for all } x, y \in V, \tag{2.7}$$

where $K$ satisfies

$$K \geq 5(\alpha+1)(\alpha+2)\overline{L}, \quad \overline{L} = \frac{L_0 + L}{2}; \tag{2.8}$$

(A4) $\nabla^2 F$ is $\alpha$-center-Hölder on $V$ at $x^*$ with constant $K_0$, i.e.

$$\|\nabla^2 F(x) - \nabla^2 F(x^*)\| \leq K_0\|x-x^*\|^\alpha \quad \text{for all } x \in V; \tag{2.9}$$

(A5) The application

$$\left[F(x^*) + \nabla F(x^*)(\cdot - x^*) + \frac{1}{2}\nabla^2 F(x^*)(\cdot - x^*)^2 + G(\cdot)\right]^{-1} \tag{2.10}$$

is $M$-pseudo-Lipschitz around $(0, x^*)$ and $G$ has closed graph.

We can now compare our hypotheses with the corresponding ones in [9]:

**Remark 2.3.** In general

$$K_0 \leq K, \quad L_0 \leq L, \tag{2.11}$$

hold in general and $\frac{K}{K_0}$ can be arbitrarily large [1], [2]. If $K_0 = K$ our hypotheses reduce to the ones in [9]. Otherwise our hypotheses can be used to improve the results in [9] as stated in the Introduction. Note that in practice the computation of $K$ requires that of $K_0$. That is the computational cost of our hypotheses (A1)–(A5) is the same as the corresponding one in [9] using (A1)–(A3) and (A5).

## 3. Local convergence analysis of method (1.2)

We will follow the proof routine in [9] but we will also stretch the differences where the really needed condition (2.9) is used instead of the stronger (2.7) used in [9].

We state the main local convergence result for method (1.2):

**Theorem 3.1.** *Let $x^*$ be a solution of (1.1). Under hypotheses* (A1)–(A5) *and for*

$$c > \frac{MK}{(\alpha+1)(\alpha+1)} \tag{3.1}$$

*there exists $\delta > 0$ such that for every starting guess $x_0 \in U(x^*, \delta)$ there exists a sequence $\{x_n\}$ $(n \geq 0)$ generated by method (1.2) satisfying*

$$\|x_{n+1} - x^*\| \leq c\|x_n - x^*\|^{2+\alpha} \quad (n \geq 0). \tag{3.2}$$

In order for us to prove this theorem we first need some notations. Let us define the set-valued map $Q$ from $X$ to the subsets of $Y$ by

$$Q(x) = F(x^*) + \nabla F(x^*)(x-x^*) + \frac{1}{2}\nabla^2 F(x^*)(x-x^*) + G(x). \tag{3.3}$$

Let

$$Z_n(x) = F(x^*) + \nabla F(x^*)(x - x^*) + \frac{1}{2}\nabla^2 F(x^*)(x - x^*)^2 \\ - F(x_n) - \nabla F(x_n)(x - x_n) - \frac{1}{2}\nabla^2 F(x_n)(x - x_n)^2, \tag{3.4}$$

and define $T_n\colon X \to Y$ by

$$T_n(x) = Q^{-1}[Z_n(x)]. \tag{3.5}$$

Clearly $x_1$ is a fixed point of $T_0$ if and only if:

$$F(x^*) + \nabla F(x^*)(x_1 - x^*) + \frac{1}{2}\nabla^2 F(x^*)(x_1 - x^*) - F(x_0) \\ - \nabla F(x_0)(x_1 - x_0) - \frac{1}{2}\nabla^2 F(x_0)(x_1 - x_0)^2 \in Q(x_1), \tag{3.6}$$

or equivalently

$$o \in F(x_0) + \nabla F(x_0)(x_1 - x_0) + \frac{1}{2}\nabla^2 F(x_0)(x_1 - x_0)^2 + G(x_1). \tag{3.7}$$

We need the proposition:

**Proposition 3.2.** *Under the hypotheses of Theorem 3.1, there exists $\delta > 0$ such that for all $x_0 \in U(x^*, \delta)$ $(x_0 \neq x^*)$, the map $T_0$ has a fixed point $x_1$ in $U(x^*, \delta)$.*

*Proof.* By (A5) there exist positive numbers $a$ and $b$ such that

$$e\left(Q^{-1}(y_1) \cap U(x^*, a), Q^{-1}(y_2)\right) \leq M\|y_1 - y_2\|, \text{ for all } y_1, y_2 \in U(0, b). \tag{3.8}$$

Choose $\delta > 0$ such that

$$\delta < \delta_0, \tag{3.9}$$

where

$$\delta_0 = \min\left\{a, \left[\frac{b(\alpha+1)(\alpha+2)}{K_0 + K2^{2+\alpha}}\right]^{\frac{1}{2+\alpha}}, \frac{(\alpha+1)(\alpha+2)}{MK} - \frac{1}{c}, \frac{1}{\sqrt[1+\alpha]{c}}\right\}. \tag{3.10}$$

We shall show condition (2.4) and (2.5) of Lemma 2.2 hold true for $p = x^*$, $T$ being $T_0$ and $r$ and $\lambda$ parameters to be determined.

We first have

$$\operatorname{dist}(x^*, T_0(x^*)) \leq e\left(Q^{-1}(0) \cap U(x^*, \delta), T_0(x^*)\right). \tag{3.11}$$

Using (2.7), (3.4), and (3.9) we obtain in turn:

$$\begin{aligned}
\|Z_0(x^*)\| &= \left\| F(x^*) - F(x_0) - \nabla F(x_0)(x^* - x_0) - \frac{1}{2}\nabla^2 F(x_0)(x^* - x_0)^2 \right\| \\
&= \left\| \int_0^1 (1-t)\nabla^2 F(x_0 + t(x^* - x_0))(x^* - x_0)^2 dt \right. \\
&\quad \left. - \frac{1}{2}\nabla^2 F(x_0)(x^* - x_0)^2 \right\| \\
&\leq K \left| \int_0^1 (1-t)t^\alpha dt \right| \|x^* - x_0\|^{2+\alpha} \\
&= \frac{K}{(\alpha+1)(\alpha+2)} \|x^* - x_0\|^{2+\alpha} < b. \qquad (3.12)
\end{aligned}$$

It follows from (3.8):

$$\begin{aligned}
e\big(Q^{-1}(0) \cap U(x^*, \delta), T_0(x^*)\big) &= e\big(Q^{-1}(0) \cap U(x^*, \delta), Q^{-1}[T_0(x^*)]\big) \\
&\leq \frac{MK}{(\alpha+1)(\alpha+2)} \|x_0 - x^*\|^{2+\alpha} \qquad (3.13)
\end{aligned}$$

and by (3.11)

$$\operatorname{dist}(x^*, T_0(x^*)) \leq \frac{MK}{(\alpha+1)(\alpha+2)} \|x^* - x_0\|^{2+\alpha}. \qquad (3.14)$$

Moreover by (3.9)

$$\operatorname{dist}(x^*, T_0(x^*)) < c\left[1 - \frac{MK\delta}{(\alpha+1)(\alpha+2)}\right] \|x^* - x_0\|^{2+\alpha}, \qquad (3.15)$$

since,

$$c\left[1 - \frac{MK\delta}{(\alpha+1)(\alpha+2)}\right] > \frac{MK}{(\alpha+1)(\alpha+2)}. \qquad (3.16)$$

Note that by the choice of $c$

$$\frac{MK\delta}{(\alpha+1)(\alpha+2)} < 1. \qquad (3.17)$$

By setting $p = x^*$, $\lambda = \frac{MK\delta}{(\alpha+1)(\alpha+2)}$ and $r = r_0 = c\|x_0 - x^*\|^{2+\alpha}$ we deduce (2.4). We shall show (2.5). We have $r_0 \leq \delta < a$, since $\delta \leq \frac{1}{1+\sqrt[3]{c}}$ for $\|x_0 - x^*\| \leq \delta$.

In view of (2.7), (2.9) and (3.4) we can obtain in turn

$$\begin{aligned}\|Z_0(x)\| \leq & \left\|F(x^*) - F(x) + \nabla F(x^*)(x - x^*) + \frac{1}{2}\nabla^2 F(x^*)(x - x^*)^2\right\| \\ & + \left\|F(x) - F(x_0) - \nabla F(x_0)(x - x_0) - \frac{1}{2}\nabla^2 F(x_0)(x - x_0)^2\right\| \\ & \leq \frac{K_0}{(\alpha+1)(\alpha+2)}\|x - x^*\|^{2+\alpha} + \frac{K}{(\alpha+1)(\alpha+2)}\|x - x_0\|^{2+\alpha} \\ & \leq \frac{K_0}{(\alpha+1)(\alpha+2)}\|x - x^*\|^{2+\alpha} \\ & + \frac{K}{(\alpha+1)(\alpha+2)}(\|x - x^*\| + \|x_0 - x^*\|)^{2+\alpha} \\ & \leq \frac{(K_0 + K \cdot 2^{2+\alpha})\delta^{2+\alpha}}{(\alpha+1)(\alpha+2)} \leq b, \end{aligned} \tag{3.18}$$

and $Z_0(x) \in U(0, b)$. That is for all $u, v \in U(x^*, r_0)$ we have

$$\begin{aligned} & e(T_0(u) \cap U(x^*, r_0), T_0(v)) \\ & \leq e(T_0(u) \cap U(x^*, \delta), T_0(v)) \leq M\|Z_0(u) - Z_0(v)\| \\ & \leq M\Big\|\nabla F(x^*)(u - v) - \nabla F(x_0)(u - v) + \frac{1}{2}\nabla^2 F(x^*)(u - v + v - u)^2 \\ & - \frac{1}{2}\nabla^2 F(x^*)(v - x^*)^2 + \frac{1}{2}\nabla^2 F(x_0)(v - x_0)^2 \\ & - \frac{1}{2}\nabla^2 F(x_0)(u - v + v - x_0)^2\Big\| \leq 5M\overline{L}\delta\|u - v\|, \end{aligned} \tag{3.19}$$

which shows (2.5). It follows by Lemma 2.2 $x_1 \in U(x^*, r_0)$ is a fixed point of $T_0$.

That completes the proof of Proposition 3.2. ☑

*Proof of Theorem 3.1.* We have $x_1 \in U(x^*, r_0)$. That is

$$\|x_1 - x^*\| \leq r_0 = c\|x_0 - x^*\|^{2+\alpha}. \tag{3.20}$$

We continue using induction on $n \geq 0$. Set $p = x^*$, $\lambda = \frac{MK\delta}{(\alpha+1)(\alpha+2)}$ and $r_n = c\|x_n - x^*\|^{2+\alpha}$ to obtain again from the application of Proposition 3.2 to $T_n$ the existence of a fixed point $x_{n+1}$ of $T_n$ in $U(x^*, r_n)$, which implies (3.2).

That completes the proof of Theorem 3.1. ☑

**Corollary 3.3.** *Let $x^*$ be a simple solution of (1.1). Under assumptions* (A1)–(A5) *for*

$$c > \frac{MK}{(\alpha+1)(\alpha+2)} = c_0 \tag{3.21}$$

*there exists $\delta > 0$ such that any sequence $\{x_n\}$ generated by (1.2) with $x_n \in U(x^*, \delta)$ satisfies (3.2).*

*Proof.* Let $\delta > 0$ be a number satisfying (3.9) and

$$\delta < \delta_1, \tag{3.22}$$

where,

$$\delta_1 = \min\left\{\frac{1}{3M\overline{L}}, \frac{(\alpha+1)(\alpha+2)c - MK}{3(\alpha+1)(\alpha+2)cM\overline{L}}\right\}. \tag{3.23}$$

We assume without loss of generality that $x^*$ is a unique solution of (1.1) in a certain neighborhood of $x^*$, since $x^*$ is a simple zero of (1.1). Let us choose it to be $U(x^*, \delta)$. Set $x^* = Q^{-1}(0) \cap U(x^*, \delta)$. By Theorem 3.1

$$x_{n+1} = Q^{-1}[Z_n(x_{n+1})].$$

In view of (2.2), (2.3), (2.7) and (2.8) we obtain in turn:

$$\begin{aligned}
\operatorname{dist}(x_{n+1}, Q^{-1}(0)) &= \|x_{n+1} - x^*\| \\
&\le e\big(Q^{-1}[Z_n(x_{n+1})] \cap U(x^*, \delta), Q^{-1}(0)\big) \le M\|Z_n(x_{n+1})\| \\
&\le M\Big\|F(x^*) + \nabla F(x^*)(x_{n+1} - x^*) + \frac{1}{2}\nabla^2 F(x^*)(x_{n+1} - x^*)^2 \\
&\quad - F(x_n) - \nabla F(x_n)(x_{n+1} - x_n) - \frac{1}{2}\nabla^2 F(x_n)(x_{n+1} - x_n)^2\Big\| \\
&\le M\Big\|F(x^*) + \nabla F(x^*)(x_{n+1} - x^*) + \frac{1}{2}\nabla^2 F(x^*)(x_{n+1} - x^*)^2 \\
&\quad - F(x_n) - \nabla F(x_n)(x_{n+1} - x^* + x^* - x_n) \\
&\quad - \frac{1}{2}\nabla^2 F(x_n)(x_{n+1} - x^* + x^* - x_n)^2\Big\| \\
&\le M\left[\frac{K_0}{(\alpha+1)(\alpha+2)}\|x^* - x_n\|^{2+\alpha} + 3\overline{L}\delta\|x_{n+1} - x^*\|\right], \qquad (3.24)
\end{aligned}$$

or

$$\begin{aligned}
\|x_{n+1} - x^*\| &\le \frac{MK}{(\alpha+1)(\alpha+2)(1 - 3M\overline{L}\delta)}\|x_n - x^*\|^{2+\alpha} \\
&\le c\|x_n - x^*\|^{2+\alpha}.
\end{aligned}$$

That completes the proof of Corollary 3.3. ☑

**Remark 3.4.** If $L_0 = L$ and $K_0 = K$, then our results are reduced to the corresponding ones in [9]. Otherwise they constitute an improvement. Indeed, let us denote by $\overline{\delta}_0$, $\overline{\delta}_1$ parameters obtained from $\delta_0$ and $\delta_1$ respectively by replacing $K_0$ and $L_0$ by $K$ and $L$ respectively. Then, we get

$$\overline{\delta}_0 \le \delta_0 \tag{3.25}$$

and

$$\overline{\delta}_1 \le \delta_1. \tag{3.26}$$

That is we can obtain a larger convergence radius for method (1.2), which implies that a wider choice of initial choices $x_0$ becomes available, and finer error bounds on the distances $\|x_n - x^*\|$ $(n \geq 0)$. These observations are important in computational mathematics [1], [2], [6].

**Remark 3.5.** The local results obtained here can be used to solve equations where $F''$ satisfies the autonomous differential equation [1], [2]

$$F''(x) = P(F(x)), \tag{3.27}$$

where $P\colon Y \to X$ is a known continuous operator. Since $F''(x^*) = P(F(x^*)) = P(0)$, we can apply our results without actually knowing the solution $x^*$ of equation (1.1).

We complete this study with two numerical examples where we show that strict inequality can hold in (2.11).

**Example 3.6.** Let $X = Y = \mathbf{R}$, $x^* = 0$, and define $F$ on $U(0,1)$ by

$$F(x) = e^x - x. \tag{3.28}$$

It can easily be seen that

$$\alpha = 1, \quad L_0 = 1, \quad L = K = e \quad \text{and} \quad K_0 = e - 1. \tag{3.29}$$

**Example 3.7.** Let $X = Y = \mathbf{R}$, $x^* = \frac{9}{4}$, $U(x^*, r) \subset D = [.81, 6.25]$, and define function $F$ on $D$ by

$$F(x) = \frac{4}{15}x^{5/2} - \frac{1}{2}x^2. \tag{3.30}$$

We obtain

$$\alpha = \frac{1}{2}, \; L_0 = \frac{1}{2}, \; L = \sqrt{6.25} - 1, \; K_0 = \frac{1}{2} \quad \text{and} \quad K = 1. \tag{3.31}$$

## References


[1] I. K. Argyros, 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.

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

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

[4] I. K. Argyros, D. Chen, & M. Tabatabai, The Halley–Werner method in Banach spaces, *Revue d'Analyse Numerique et de theorie de l'Approximation*, **1** (1994), 1–14.

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

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

[7] A. L. Dontchev, Local convergence of the Newton method for generalized equations, *C.R.A.S. Paris* **332** Ser. I (1996), 327–331.

[8] A. L. DONTCHEV & W. W. HAGER, An inverse function theorem for set-valued maps, *Proc. Amer. Math. Soc.* **121** (1994), 481–489.
[9] M. H. GEOFFROY & A. A. PIETRUS, Superquadratic method for solving generalized equations in the Hölder case, *Ricerche di Matematica* **LII** fasc. 2 (2003), 231–240.
[10] A. D. IOFFE & V. M. TIKHOMIROV, *Theory of Extremal Problems*, North Holland, Amsterdam, 1979.
[11] S. M. ROBINSON, Strong regular generalized equations, *Math. Oper. Res.* **5** (1980), 43–62.

(Recibido en marzo de 2006. Aceptado en mayo de 2006)

DEPARTMENT OF MATHEMATICAL SCIENCES
CAMERON UNIVERSITY
OK 73505
LAWTON, USA
*e-mail:* iargyros@cameron.edu