Nonlinear Functional Analysis and Applications
Vol. 13, No. 1 (2008), pp. 147-158

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm

ON THE SECANT METHOD FOR SOLVING NONSMOOTH EQUATIONS AND NONDISCRETE INDUCTION

Ioannis K. Argyros

Cameron University, Department of Mathematical Sciences,
Lawton, OK 73505, USA

e-mail: `ioannisa@cameron.edu`

Abstract. In this study we are concerned with the problem of approximating a locally unique solution of an operator equation in Banach space using the Secant method, and nondiscrete mathematical induction. The differentiability of the operator involved is not assumed. Using a flexible point-based approximation we provide a local as well as a semilocal convergence analysis for the Secant method. Our results are justified by numerical examples that cannot be handled with earlier works.

1. Introduction

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

$$F(x) = 0, \tag{1.1}$$

where operator F is a continuous operator defined on a closed subset D of a Banach space X with values in a Banach space Y.

A large number of problems in applied mathematics and also in engineering are solved by finding the solutions of certain equations [2], [12], [14], [16]. For example, dynamic systems are mathematically modeled by difference or differential equations, and their solutions usually represent the states of the systems. For the sake of simplicity, assume that a time-invariant system is

[0]Received December 20, 2006. Revised September 29, 2007.
[0]2000 Mathematics Subject Classification: 65H10, 65G99, 47H10, 49M15.
[0]Keywords: Banach space, secant method, Newton-like methods, point-based approximation, Banach lemma on invertible operators, nondiscrete mathematical induction.

driven by the equation $\dot{x} = T(x)$ (for some suitable operator T), where x is the state. Then the equilibrium states are determined by solving equation (1). Similar equations are used in the case of discrete systems. The unknowns of engineering equations can be functions (difference, differential, and integral equations), vectors (systems of linear or nonlinear algebraic equations), or real or complex numbers (single algebraic equations with single unknowns). Except in special cases, the most commonly used solution methods are iterative, when starting from one or several initial approximations a sequence is constructed that converges to a solution of the equation. Iteration methods are also applied for solving optimization problems. In such cases, the iteration sequences converge to an optimal solution of the problem at hand. Since all of these methods have the same recursive structure, they can be introduced and discussed in a general framework.

The most popular iterative procedures for approximations x^* are the so-called Newton-like methods. The essence of these methods is to replace F by an approximate operator (linearization) that can be solved more easily. Recent developments on local and semilocal results concerning such methods can be found in [2]–[4], [12], [14], and the references there.

When operator F is nonsmooth, the linearization is no longer available. In [16] a replacement was introduced through the notion of a point-based-approximation (to be precised later). The properties of this approximation are similar to those of linearization and were successfully used for Newton's method. However we noticed (see the numerical example at the end of the study) that such an approximation may not exist. Therefore in order to solve a wider range of problems we introduce a more flexible and precise point-based-approximation which is more suitable for Newton-like methods and in particular for Secant-type iterative procedures [1], [2], [5]–[11]. We will use the method of nondiscrete mathematical induction due to Porta and Ptak [14] for our convergence analysis (see also [13]).

A semilocal convergence analysis for the Secant method is provided. Our approach is justified through numerical examples.

2. Preliminary results

We need a definition of a point-based-approximation (PBA) for operator F which is suitable for the Secant method.

Definition 1. Let F be an operator from a closed subset D of a metric space (X, d) into a normed linear space Y. Operator F has a (PBA) on D at the point $x_0 \in D$ if there exists an operator $A\colon D \times D \times D \to Y$ and scalars ℓ_0, ℓ such that u, v, w, x, y and z in D,

$$\|F(w) - A(u, v, w)\| \leq \ell d(u, w) d(v, w), \tag{2.1}$$

$$\|[A(x,y,z)-A(x_0,x_0,z)]-[A(x,y,w)-A(x_0,x_0,w)]\| \\ \leq \ \ell_0[d(x,x_0)+d(y,x_0)]d(z,w), \tag{2.2}$$

and

$$\|[A(x,y,z)-A(u,v,z)]-[A(x,y,w)-A(u,v,w)]\| \leq \ell[d(x,u)+d(u,v)]d(z,w). \tag{2.3}$$

We then say A is a (PBA) for F.

This definition is suitable for the application of the Secant method. Indeed let X be also a normed linear space, D is convex and F has a divided difference of order one on $D\times D$ denoted by $[x,y;F]$ and satisfying the standard condition [2], [12], [13]:

$$\|[u,v;F]-[w,x;F]\| \leq \ell(\|u-w\|+\|v-x\|) \tag{2.4}$$

for all u, v, w and x in D. If we set

$$A(u,v,w)=F(v)+[u,v;F](w-v) \tag{2.5}$$

then (2) becomes

$$\|F(w)-F(v)-[u,v;F](w-v)\| \leq \ell\|u-w\|\,\|v-w\|, \tag{2.6}$$

whereas (3) and (4) are equivalent to property (5) of linear operator $[\cdot,\cdot;F]$. Note that a (PBA) does not imply differentiability.

It follows by (2) that one way of finding a solution x^* of equation (1) is to solve for w the equation

$$A(x,y,w)=0 \tag{2.7}$$

provided that x and y are given.

We now need a definition also used in [15], [16] which amounts to the reciprocal of a Lipschitz constant for the inverse operator.

Definition 2. Let X, D, and Y be as in Definition 1, and let $G\colon D\to Y$. Then

$$\delta(G,D)=\inf\left\{\frac{\|G(u)-G(v)\|}{d(u,v)},\ u\neq v,\ u,v\in D\right\}. \tag{2.8}$$

Note that G is one-to-one if and only if $\delta(G,D)\neq 0$.

We state and prove the following generalization of the classical Banach Lemma on invertible operators [12, Th. 4 (2.V)]:

Lemma 1. *Let X, D and Y be as in Definition 1. Assume further X is a Banach space. Let F and G be operators from D into Y with G being Lipschitzian with modulus ℓ and center-Lipschitzian with modulus ℓ_0. Let $x_0\in D$ with $F(x_0)=y_0$. Assume that*:

$$\overline{U}(y_0,\alpha)=\{y\in Y \mid \|y-y_0\|\leq\alpha\}\subseteq F(D); \tag{2.9}$$

$$0 \le \ell < d = \delta(F, D); \tag{2.10}$$
$$\overline{U}(x_0, d^{-1}\alpha) \subseteq D, \tag{2.11}$$

and

$$\theta_0 = (1 - \ell_0 d^{-1})\alpha - \|G(x_0)\| \ge 0. \tag{2.12}$$

Then the following hold:

$$\overline{U}(y_0, \theta_0) \subseteq (F + G)(\overline{U}(x_0, d^{-1}\alpha)) \tag{2.13}$$

and

$$\delta(F + G, D) \ge d - \ell. \tag{2.14}$$

Proof. Define operator $Ty(x) = F^{-1}(y - G(x))$, for each fixed $y \in \overline{U}(y_0, \theta_0)$, and $x \in \overline{U}(x_0, \delta^{-1}\alpha)$. We can get:

$$\begin{aligned} \|y - G(x) - y_0\| &\le \|y - y_0\| + \|G(x) - G(x_0)\| + \|G(x_0)\| \\ &\le \theta_0 + \ell_0 \delta^{-1}\alpha + \|G(x_0)\| = \alpha. \end{aligned}$$

Therefore $Ty(x)$ is a singleton set since $\delta > 0$. That is Ty is an operator on $\overline{U}(x_0, \delta^{-1}\alpha)$. This operator maps $\overline{U}(x_0, \delta^{-1}\alpha)$ into itself. Indeed for $x \in \overline{U}(x_0, \delta^{-1}\alpha)$:

$$d(Ty(x), x_0) = d(F^{-1}(y - G(x)), F^{-1}(y_0)) \le \delta^{-1}\alpha.$$

Moreover let u, v be in $\overline{U}(x_0, \delta^{-1}\alpha)$, then

$$\begin{aligned} d(Ty(u), Ty(v)) &\le d(F^{-1}(y - G(u)), F^{-1}(y - G(v))) \\ &\le \delta^{-1}\ell d(u, v). \end{aligned} \tag{2.15}$$

It follows by the contraction mapping principle [12, Th. 1 (1.XVI)] and (11) that operator Ty is a strong contraction, and as such it has a fixed point $x(y)$ in $\overline{U}(x_0, \delta^{-1}\alpha)$ with $(F + G)(x(y)) = y$. Such a point $x(y)$ in D is unique in D since

$$\begin{aligned} \delta(F + G, D) &= \inf\left\{ \frac{\|[F(u) - F(v)] + [G(u) - G(v)]\|}{d(u, v)},\ u \ne v,\ u, v \in D \right\} \\ &\ge \delta(F, D) - \sup\left\{ \frac{\|G(u) - G(v)\|}{d(u, v)},\ u \ne v,\ u, v \in D \right\} \\ &\ge \delta - \ell \\ &> 0. \end{aligned}$$

That is $F + G$ is one-to-one on D. □

Remark 1. In general

$$\ell_0 \le \ell \tag{2.16}$$

holds and $\frac{\ell}{\ell_0}$ can be arbitrarily large [3], [4]. If equality holds (17) our Lemma 1 reduces to the corresponding Lemma 3.1 in [15, p. 298]. Otherwise our

Lemma 1 improves (enlarges) the range for θ given in [15, p. 298], and under the same computational cost since in practice the computation of ℓ requires that of ℓ_0. This observation is important in computational mathematics, since it enlarges the ball $\overline{U}(y_0, \theta_0)$.

The following lemma is used to show uniqueness of the solution in the semilocal case.

Lemma 2. *Let X and Y be normed linear spaces, and let D be a closed subset of X. Let $F\colon D \to Y$, and let A be a (PBA) for operator F on D at the point $x_0 \in D$. Denote by d the quantity $\delta(A(x_0, x_0, \cdot), D)$. If $\overline{U}(x_0, \rho) \subseteq D$, then*

$$\delta(F, \overline{U}(x_0, \rho)) \geq d - (2\ell_0 + \ell)\rho. \tag{2.17}$$

In particular, if $d - (2\ell_0 + \ell)\rho > 0$, then F is one-to-one on $\overline{U}(x_0, \rho)$.

Proof. Let w, z be points in $\overline{U}(x_0, \rho)$. We can write

$$\begin{aligned} F(w) - F(z) \quad = \quad & [F(w) - A(x, y, w)] + [A(x, y, w) - A(x, y, z)] \\ & + [A(x, y, z) - F(z)] \end{aligned} \tag{2.18}$$

By (2) we can have

$$\|F(w) - A(x, y, w)\| \leq \ell \|x - w\| \, \|y - w\|$$

and

$$\|F(z) - A(x, y, z)\| \leq \ell \|x - z\| \, \|y - z\|.$$

Moreover we can find

$$\begin{aligned} \|A(x, y, u) - A(x, y, v)\| &\geq \|A(x_0, x_0, u) - A(x_0, x_0, v)\| \\ &\quad - \|[A(x, y, u) - A(x_0, x_0, u)] - [A(x, y, v) - A(x_0, x_0, v)]\| \end{aligned}$$

and therefore

$$\begin{aligned} &\delta(A(x, y, \cdot), D) \\ &\geq \delta(A(x_0, x_0, \cdot), D) \\ &\quad - \sup\left\{ \frac{\|[A(x, y, u) - A(x_0, x_0, u)] - [A(x, y, v) - A(x_0, x_0, v)]\|}{\|u - v\|}, \right. \\ &\qquad\qquad \left. u \neq v, \ u, v \in D \right\} \\ &\geq d_0 - \ell_0(\|x - x_0\| + \|y - x_0\|) \geq d - 2\ell_0 \rho. \end{aligned}$$

Furthermore, we can now have

$$\begin{aligned} \|F(w) - F(z)\| &\geq (d - 2\ell_0 \rho)\|w - z\| - \ell[\|x - w\| \, \|y - w\| + \|x - z\| \, \|y - z\|] \\ &\geq (d - 2\ell_0 \rho)\|w - z\| - \frac{\ell}{2}\|w - z\|^2 \end{aligned}$$

and for $w \neq z$,

$$\frac{\|F(w) - F(z)\|}{\|w - z\|} \geq d - (2\ell_0 + \ell)\rho.$$

That completes the proof of Lemma 2. $\square$

Remark 2. In order for us to compare our result with the corresponding Lemma 2.4 in [16, p. 294], first note that if:

(a) equality holds in (17), $u = v$ and $x = y$ in (2)–(4), then our result reduces to Lemma 2.4 by setting $\frac{k}{2} = \ell = \ell_0$.

(b) Strict inequality holds in (17), $u = v$ and $x = y$ then our Lemma 2 improves (enlarges) the range for ρ, and under the same computational cost. The implication of that is: (see Remark 3 that follows): in the semilocal case the uniqueness ball is more precise.

3. Convergence analysis

We use the notation $\omega^{(n)}$ to denote the nth iterate of a function ω from a set S into itself. That is $\omega^n(s)$ is the member of the sequence given by $s_0 = s$, $s_{i+1} = \omega(s_i)$ $(i \geq 0)$. A rate of convergence on a real interval T of the form $(0, t_0]$ or $(0, +\infty)$ is a function $\omega\colon T \to T$ such that the series $s(t) = \sum_{k=0}^{\infty} \omega^{(k)}(t)$ converges for each $t \in T$. The sum $s(t)$ is called the estimate function corresponding to ω. If $Z(t)$ is a family of sets defined on T of the above form, then by $Z(0^+)$ we denote the limit superior of $Z(t)$ as $t \to 0$. That is the set of all points z that are limits of sequences $\{z_k\}$ such that $z_k \in Z(t_k)$ for each k, $t_k \in T$, and $t_k \to 0$ as $k \to \infty$.

We will use the following theorem whose proof can be found in [13, Th. 1.9].

Theorem 1. *Let (E, d) be a Banach space and let G be a function from a subset D of E into E. Let $x_0 \in D$. Suppose that there exists a rate of convergence ω on T and a family of subsets $Z(t) \subseteq D$ $(t \in T)$ such that:*

(a) *for some $r_0 \in T$, $x_0 \in Z(r_0)$, and*

(b) *for each $t \in T$ and each $x \in Z(t)$, $d(G(x), x) \leq t$ and $G(x) \in Z(\omega(t))$.*

Then the following hold:

(1) *the sequence $\{x_n\}$ generated from x_0 by $x_{n+1} = G(x_n)$ is well defined, and it converges to a point $x^* \in Z(0^+) \subseteq \mathrm{c}\ell\, D$.*

(2) *For each $n, x_n \in Z[\omega^{(n)}(r_0)]$, $d(x_n, x_{n+1}) \leq \omega^{(n)}(r_0)$, and $d(x_n, x^*) \leq s[\omega^{(n)}(r_0)]$.*

(3) *If for some n and some $d_n \in T$ one has $x_{n-1} \in Z(d_n)$, then $d(x_n, x^*) \leq s(d_n) - d_n = s(\omega(d_n))$.*

We state and prove the main semilocal convergence theorem for the Secant method. The proof uses Theorem 1.

Theorem 2. *Let X and Y be Banach spaces, D a closed convex subset of X, x_{-1}, and $x_0 \in D$ with $\|x_0 - x_{-1}\| \leq q_0$, $\|x_0 - x_1\| \leq c_0$, and F a continuous operator from D into Y. Suppose operator F has a (PBA) on D. Moreover assume:*

(a) $\delta(A(x_{-1}, x_0, \cdot), D) \geq d_0 > 0$;

(b) $0 \leq h = (1 - d_0^{-1} h_0 q_0)^2 - 4 d_0^{-1} h_0 r_0$;

(c) *for each $y \in \overline{U}(0, d_0(c - r_0))$ the equation $A(x_{-1}, x_0, y) = y$ has a solution x, where*

$$c = \frac{d_0}{2h_0}\left(1 - d_0^{-1} h_0 q_0 - \sqrt{h}\right); \tag{3.1}$$

(d) *the solution $G(x_{-1}, x_0) = 0$ of $A(x_{-1}, x_0, G(x_{-1}, x_0)) = 0$ satisfies*

$$\|G(x_0) - x_0\| \leq r_0;$$

(e) $\overline{U}(x_0, c) \subseteq D$;
and

(f) $c_0 \geq c$.

Then the Secant iteration defining x_{n+1} by

$$A(x_{n-1}, x_n, x_{n+1}) = 0 \tag{3.2}$$

remains in $\overline{U}(x_0, c)$ and converges to a solution $x^* \in \overline{U}(x_0, c)$ of equation $F(x) = 0$.

Moreover the following estimates hold:

$$\|x_{n+1} - x_n\| \leq \omega^{(n)}(r_0), \tag{3.3}$$

and

$$\|x_n - x^*\| \leq s(\omega^{(n)}(r_0)) \quad \text{for all } n \geq 0, \tag{3.4}$$

where functions ω, s are the rate of convergence and its associated estimate function [13, p. 433], and are given by

$$\omega(r) = r\left(d_0^{-1} h_0 r + 1 - 2\sqrt{d_0^{-2} h_0^2 a^2 + d_0^{-1} r}\right) \tag{3.5}$$

$$s(r) = \sqrt{a^2 + d_0 h_0^{-1} r} - a + r, \tag{3.6}$$

respectively, for $r > 0$ and

$$a = \frac{d_0}{2h_0}\sqrt{h}. \tag{3.7}$$

Proof. We need to define functions $g(r) = c - g(r)$, $\beta(r) = h_0(a + s(r))$ and $\alpha(r) = \beta(r)s(w(r))$. We denote $\delta(A(x, x, \cdot), D)$ by $d(x)$ from now on for simplicity.

For each positive r define $Z(r)$ to be the set of all $x \in D$ such that

(i) $\|x - x_0\| \le g(r)$,
(ii) $d(x) \ge \beta(r)$,
(iii) for each $y \in \overline{U}(0, \alpha(r))$, the equation $A(\cdot, x, z) = y$ has a unique solution z in D, and
(iv) the solution $G(\cdot, x)$ of $A(\cdot, x, G(\cdot, x)) = 0$ satisfies $\|x - G(\cdot, x)\| \le r$.

The first step in the application of the method of nondiscrete induction using the set-function Z is to show that $G(\cdot, x) \in Z(\omega(r))$ if $r > 0$ and $x \in Z(r)$ for all $\omega \in D$. By (iii) $G(\cdot, x)$ exists. Set $\overline{x} = G(\cdot, x)$. Using (i) and (iv), we get

$$\|\overline{x} - x_0\| \le \|x - x_0\| + \|\overline{x} - x\| \le g(r) + r = g(\omega(r)). \tag{3.8}$$

This shows (i) for $\overline{x}$. We also have $c > g(\omega(r))$ and $\overline{U}(x_0, c) \subseteq D$ (by hypothesis (e)) which in fact imply $\overline{x} \in D$. Concerning (ii) using (3) we get for $v, \overline{v} \in Z$

$$\begin{aligned} \delta(A(\overline{x}, v, \cdot), D) &\ge \delta(A(\overline{v}, \overline{x}, \cdot), D) - h_0(\|\overline{x} - \overline{v}\| + \|v - \overline{v}\|) \\ &\ge \beta(\omega(r)). \end{aligned} \tag{3.9}$$

For (iii) we must show that if $y \in \overline{U}(0, \alpha(\omega(r))$ then $A(\cdot, \overline{x}, z) = y$ has a unique solution z in D. To achieve this we use Lemma 1 with F, T, x_0, and y_0 replaced by $A(\cdot, x, \cdot)$, $A(\cdot, \overline{x}, \cdot) - A(\cdot, x, \cdot)$, $\overline{x}$ and 0 respectively.

Hypothesis (a) of the lemma follows from property (iii) of $Z(r)$. The function $\beta(\omega(r))$ is positive and therefore by (28) hypothesis (b) of Lemma 1 holds for $d = \beta(r)$.

Hypothesis (c) of Lemma 1 will be verified if $\overline{U}(\overline{x}, \beta(r)^{-1}\alpha(r)) \subseteq D$. We have

$$\overline{U}(\overline{x}, s(\omega(r)) \subseteq \overline{U}(x_0, g(\omega(r)) + s(\omega(r))) = \overline{U}(x_0, c) \subseteq D$$

where we used hypothesis (e) of the theorems together with the estimates $\beta(r)^{-1}\alpha(r) = s(\omega(r))$ and $\|\overline{x} - x_0\| \le g(\omega(r))$.

Hypothesis (d) of Lemma 1 will be verified if we show that θ_0 i.e.

$$[1 - h_0 r\beta(r)^{-1}]\beta(r)s(\omega(r)) - \|A(\cdot, x, \overline{x}) - A(\cdot, \overline{x}, \overline{x})\| \tag{3.10}$$

is nonnegative. Instead we will show that a certain upper bound of (29) is nonnegative.

Using (2) we get

$$\begin{aligned} \|A(v, \overline{v}, \overline{x}) - A(\overline{x}, \overline{x}, \overline{x})\| &\le h_0(\|v - \overline{x}\| + \|\overline{v} - \overline{x}\|)\|\overline{v} - \overline{x}\| \\ &\le \omega(r). \end{aligned} \tag{3.11}$$

By replacing the norm in (2) by $\omega(r)$ and using the properties of functions β, ω and s [13] it is simple algebra to show that the resulting expression is nonnegative, which imply the same for θ_0.

For each $y \in \overline{U}(0, \alpha(\omega(r))$, the equation $A(\omega, \overline{x}, z) = y$ has a solution z which is unique in D because $d(\overline{x}) > 0$. Therefore $\overline{x} \in Z(\omega(r))$. To show point (iv), we must show that if $\overline{\overline{x}}$ is the solution of $A(\omega, \overline{x}, z) = 0$, then $\|\overline{\overline{x}} - \overline{x}\| \leq \omega(r)$. Using (30) to bound $\|F(\overline{x})\|$ and (28) to bound $d(\overline{x})$ we get by (9)

$$\|\overline{\overline{x}} - \overline{x}\| \leq d(\overline{x})\|A(\overline{x}, \overline{x}, \overline{\overline{x}}) - A(\overline{x}, \overline{x}, \overline{x})\| \leq \omega(r),$$

which implies point (iv) also holds.

We now have shown that if $r > 0$ and $x \in Z(r)$, then $G(x) \in Z(\omega(r))$. We still need to show $x_0 \in Z(r_0)$. Note that $x_0 \in D$. Clearly $g(r_0) = c - s(r_0) = 0$, so $g(r_0) \geq 0$ holds which implies (i) of the definition of $Z(r_0)$. For point (ii) we must show $d_0 \geq \beta(r_0)$ which follows immediately from the definition of β and r_0. Point (iii) follows from $\alpha(r_0) = d(c - r_0)$, and hypothesis (c) of the theorem. Finally, hypothesis (d) of the theorem ensures that point (iv) holds, which imply $x_0 \in Z(r_0)$.

By Theorem 1 $\{x_n\}$ is a well defined sequence that converges to some $x^* \in Z(0^+)$ with $x_n \in Z[\omega^{(n)}(r_0)]$ for each n. We must have

$$\|x^* - x_0\| = \lim_{n\to\infty} \|x_n - x_0\| \leq c.$$

Moreover, since $x^* \in Z(0^+)$, there exist two sequences $\{t_j\}$ converging to zero in $\mathbf{R}$ and $\{z_j\}$ converging to x^* with $z_j \in Z(t_j)$ for each j. Writing z'_j for the unique solution of $A(\omega, z_j, z) = 0$, we have $\|z'_j - z_j\| \leq t_j$ for each j according to point (iv) of the definition of the set $Z(t_j)$. Therefore $\{z'_j\}$ converges to x^*. Furthermore by (2)

$$\begin{aligned} \|F(z'_j)\| &= \|F(z'_j) - A(z, z_j, z'_j)\| \\ &\leq h_0(\|z'_j - z_j\| + \|z - z_j\|)\|z'_j - z_j\| \to 0 \end{aligned}$$

as $j \to \infty$. By the continuity of F we deduce $F(x^*) = 0$.

Estimates (22) and (23) follow from Theorem 1 and the properties of function β, ω and s (see also Theorem 1 in [13, p. 434]).

That completes the proof of Theorem 1. □

Remark 3. The uniqueness of the solution x^* was not considered in Theorem 2. Indeed, we do not know if under the conditions stated above the solution x^* is unique, say in $\overline{U}(x_0, c)$. However using Lemma 2 we can obtain a uniqueness result, so that if ρ satisfies

$$c < \rho < \frac{d}{2\ell_0 + \ell},$$

then operator F is one-to-one in a neighborhood of x^*, since $x^* \in \overline{U}(x_0, c)$. That is x^* is an isolated zero of F in this case.

4. Numerical examples

In this section we show how to choose operator A in cases not covered in [1], [4]–[11], [13], [16]. Let $X = Y = (\mathbf{R}^2, \|\cdot\|_\infty)$. Consider the system

$$\begin{aligned} 3x^2y + y^2 - 1 + |x-1| &= 0 \\ x^4 + xy^3 - 1 + |y| &= 0. \end{aligned} \tag{4.1}$$

Set for $v = (v_1, v_2)$, $\|v\|_\infty = \|(v_1, v_2)\|_\infty = \max\{|v_1|, |v_2|\}$, $F(v) = P(v)+Q(v)$, $P = (P_1, P_2)$, $Q = (Q_1, Q_2)$. Define

$$P_1(v) = 3v_1^2v_2 + v_2^2 - 1, \quad P_2(v) = v_1^4 + v_1v_2^3 - 1,$$

$$Q_1(v) = |v_1 - 1|, \quad Q_2(v) = |v_2|.$$

We shall take divided differences of order one $[x, y; P]$, $[x, y; Q] \in M_{2\times 2}(\mathbf{R})$ to be for $w = (w_1, w_2)$:

$$[v, w, P]_{i,1} = \frac{P_i(w_1, w_2) - P_i(v_1, w_2)}{w_1 - v_1},$$

$$[v, w, P]_{i,2} = \frac{P_i(v_1, w_2) - P_i(v_1, v_2)}{w_2 - v_2}$$

provided that $w_1 \neq v_1$ and $w_2 \neq v_2$. If $w_1 = v_1$ or $w_2 = v_2$ replace $[x, y; P]$ by P'. Similarly we define

$$[v, w; Q]_{i,1} = \frac{Q_i(w_1, w_2) - Q_i(v_1, w_2)}{w_1 - v_1},$$

$$[v, w; Q]_{i,2} = \frac{Q_i(v_1, w_2) - Q_i(v_1, v_2)}{w_2 - v_2}$$

for $w_1 \neq v_1$ and $w_2 \neq v_2$. If $w_1 = v_1$ or $w_2 = v_2$ replace $[x, y; Q]$ by the zero 2×2 matrix in $M_{2\times 2}(\mathbf{R})$.

We consider three interesting choices for operator A:

$$A(v, v, w) = P(v) + Q(v) + P'(v)(w - v), \tag{4.2}$$

$$A(u, v, w) = P(v) + Q(v) + ([u, v; P] + [u, v; Q])(w - v) \tag{4.3}$$

and

$$A(u, v, w) = P(v) + Q(v) + (P'(v) + [u, v; Q])(w - v). \tag{4.4}$$

Using method (32) for $x_0 = (1, 0)$, and both methods (33) and (34) for $x_{-1} = (5, 5)$, $x_0 = (1, 0)$ we obtain the following three tables respectively:

n	$x_n^{(1)}$	$x_n^{(2)}$	$\|x_n - x_{n-1}\|$
0	1	0	
1	1	0.333333333333333	3.333E–1
2	0.906550218340611	0.354002911208151	9.344E–2
3	0.885328400663412	0.338027276361322	2.122E–2
4	0.891329556832800	0.326613976593566	1.141E–2
5	0.895238815463844	0.326406852843625	3.909E–3
6	0.895154671372635	0.327730334045043	1.323E–3
7	0.894673743471137	0.327979154372032	4.809E–4
8	0.894598908977448	0.327865059348755	1.140E–4
9	0.894643228355865	0.327815039208286	5.002E–5
10	0.894659993615645	0.327819889264891	1.676E–5
11	0.894657640195329	0.327826728208560	6.838E–6
12	0.894655219565091	0.327827351826856	2.420E–6
13	0.894655074977661	0.327826643198819	7.086E–7
...			
39	0.894655373334687	0.327826521746298	5.149E–19

n	$x_n^{(1)}$	$x_n^{(2)}$	$\|x_n - x_{n-1}\|$
−1	5	5	
0	1	0	5.000E+00
1	0.9898008742107820	0.012627489072365	1.262E–02
2	0.921814765493287	0.307939916152262	2.953E–01
3	0.900073765669214	0.325927010697792	2.174E–02
4	0.894939851625105	0.327725437396226	5.133E–03
5	0.894658420586013	0.327825363500783	2.814E–04
6	0.894655375077418	0.327826521051833	3.045E–04
7	0.894655373334698	0.327826521746293	1.742E–09
8	0.894655373334687	0.327826521746298	1.076E–14
9	0.894655373334687	0.327826521746298	5.421E–20

n	$x_n^{(1)}$	$x_n^{(2)}$	$\|x_n - x_{n-1}\|$
−1	5	5	
0	1	0	5
1	0.909090909090909	0.363636363636364	3.636E–01
2	0.894886945874111	0.329098638203090	3.453E–02
3	0.894655531991499	0.327827544745569	1.271E–03
4	0.894655373334793	0.327826521746906	1.022E–06
5	0.894655373334687	0.327826521746298	6.089E–13
6	0.894655373334687	0.327826521746298	2.710E–20

We did not verify the hypotheses of Theorem 2 for the above starting points. However, it is clear that the hypotheses of Theorem 2 are satisfied for all three methods for starting points closer to the solution

$$x^* = (.894655373334687, .327826521746298)$$

chosen from the lists of the tables displayed above.

Note that the results in [16] cannot apply here because operator A no matter how it is chosen cannot satisfy the Lipschitz conditions (a) or (b) in Definition 2.1 in [16, p. 293] needed for the application of Theorem 3.2 in the same paper.

Other possible applications of operators equations with a (PBA) are already noted in [2], [16] and the references there.

References

[1] S. Amat, S. Busquier and V. Candela, *A class of quasi-Newton generalized Steffensen methods on Banach spaces*, J. Comput. Appl. Math., **149** (2002), 397–408.

[2] I. K. Argyros, *Advances in the Efficiency of Computational Methods and Applications*, World Scientific Publ. Co., River Edge, NJ, 2000.

[3] I. K. Argyros, *An improved error analysis for Newton-like methods under generalized conditions*, J. Comput. Appl. Math., **157**, 1 (2003), 169–185.

[4] I. K. Argyros, *New sufficient convergence conditions for the Secant method*, Checkoslovak Mathematical Journal (to appear).

[5] J. M. Gutierrez, *A new semilocal convergence theorem for Newton's method*, J. Comput. Appl. Math., **79** (1997), 131–145.

[6] J. M. Gutierrez, M. A. Hernandez and M. A. Salanova, *Accessibility of solutions by Newton's method*, Internat. J. Comput. Math., **57** (1995), 239–247.

[7] J. M. Gutierrez, M. A. Hernandez and M. A. Salanova, *Resolution of quadratic equations in Banach space*, Numer. Funct. Anal. and Optim., **17**, (1 and 2) (1996), 113–121.

[8] M. A. Hernandez and M. A. Salanova, *A family of Chebyshev–Halley-type methods*, Internat. J. Comp. Math., **47** (1993), 56–63.

[9] M. A. Hernandez, *Chebyshev's approximation algorithms and applications*, Comput. Math. Appl., **41**, 3–4 (2001), 433–445.

[10] M. A. Hernandez, M. J. Rubio and J. A. Ezquerro, *Secant-like methods for solving nonlinear integral equations of the Hammerstein type*, J. Comput. Appl. Math., **115** (2000), 245–254.

[11] M. A. Hernandez and M. J. Rubio, *Semilocal convergence of the Secant method under mild convergence conditions of differentiability*, Comput. Math. Appl., **44**, 3–4 (2002), 277–285.

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

[13] F. A. Potra, *An error analysis for the Secant method*, Numer. Math. **38** (1982), 427–445.

[14] F. A. Potra and V. Pták, *Nondiscrete Induction and Iterative Processes*, Pitman, Boston, 1984.

[15] S. M. Robinson, *An implicit function theorem for a class of nonsmooth functions*, Mathematics of Operations Research, **16**, 2 (1991), 292–309.

[16] S. M. Robinson, *Newton's method for a class of nonsmooth functions*, Set-Valued Analysis, **2** (1994), 291–305.