Mathematical Modelling and Analysis
Volume 29, Issue 2, 347–366, 2024
https://doi.org/10.3846/mma.2024.18680

A Modified Newton-Secant Method for Solving Nonsmooth Generalized Equations

Vitaliano de Sousa Amaral[a], Paulo Sérgio Marques dos Santos[b], Gilson N. Silva[a] and Sissy Souza[b]

[a] *Universidade Federal do Piauí, Departamento de Matemática*
Teresina, PI, Brazil
[b] *Universidade Federal do Delta do Parnaíba, Departamento de Matemática*
Parnaíba, PI, Brazil
E-mail(*corresp.*): vitalianoamaral@ufpi.edu.br
E-mail: psergio@ufpi.edu.br
E-mail: gilson.silva@ufpi.edu.br
E-mail: sissy@ufpi.edu.br

Received February 23, 2023; accepted March 8, 2024

Abstract. In this paper, we study the solvability of nonsmooth generalized equations in Banach spaces using a modified Newton-secant method, by assuming a Hölder condition. Also, we generalize a Dennis-Moré theorem to characterize the superlinear convergence of the proposed method applied to nonsmooth generalized equations under strong metric subregularity. Numerical examples are provided to illustrate the effectiveness of our approach.

Keywords: Newton-Kantorovich theorem, divided differences, Newton-secant method, generalized equations.

AMS Subject Classification: 65K05; 65K15; 49M15; 49M37.

1 Introduction

In this study, we aim to investigate the Newton-Kantorovich method for solving a nonsmooth generalized equation

$$f(x) + g(x) + F(x) \ni 0, \tag{1.1}$$

where $g : \mathbb{X} \to \mathbb{Y}$ is a continuous function that admits first and second order divided differences, $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ is a multifunction, $f : \mathbb{X} \to \mathbb{Y}$ is a Fréchet differentiable function, and $\mathbb{X}$ and $\mathbb{Y}$ are Banach spaces.

One of the main reasons for this is these equations arise from the reformulation of some problems in mathematical programming, for instance, nonlinear equations, systems of equations and inequalities, nonlinear complementarity, variational inequality, and equilibrium problems, see for instance [1,15,24] and references therein. Owing to the large number of applications wherein this problem appears, there are many numerical techniques that deal with them.

The standard Newton-secant method for solving (1.1), and proposed by Geffroy and Piétrus [16], uses the iteration

$$f(x_k) + g(x_k) + (f'(x_k) + [x_{k-1}, x_k; g])(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad (1.2)$$

$k = 1, 2, \ldots$, with a given guess point $x_0 \in \mathbb{X}$, and $[x_{k-1}, x_k; g]$ represents the divided difference operator related to g. The above iterative scheme was studied by Catinas in [3] for the particular case $F \equiv 0$. Hernández and Rubio [18] also presented a semilocal convergence analysis for the above method (1.2) to the particular case as in [3]. Other important variants of the method (1.2) has been studied, for instance, in [19] Jean-Alexis and Piétrus presented the following iterative scheme for solving (1.1)

$$f(x_k) + g(x_k) + (f'(x_k) + [2x_{k-1} - x_k, x_k; g])(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad (1.3)$$

$k = 1, 2, \ldots$. Under suitable conditions, they proved in [19] a local convergence analysis for the previous method with superlinear rate. Some years later, Rashid, Wang and Li [22] also presented a local convergence analysis for the above method, but with different assumptions compared with [19].

It is well-known that Newton's method has some disadvantages in practice. For instance, on one hand, one has to compute the Jacobian at every iteration and, on the other hand, a linear system must be solved exactly. This sometimes makes the Newton's method inefficient especially when the problem size is large, see for instance [20].

To overcome these disadvantages we propose the following method for solving (1.1) where the Fréchet derivative $f'(x_k)$ is replaced by a perturbed Fréchet derivative $A(x_k)$ which is much easier or computationally less expensive to calculate:

$$f(x_k) + g(x_k) + \big(A(x_k) + [x_{k-1}, x_k; g]\big)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \quad (1.4)$$

$k = 1, 2, \ldots$. We assume that $f : \mathbb{X} \to \mathbb{Y}$ is Fréchet differentiable, f has a Hölder derivative, $g : \mathbb{X} \to \mathbb{Y}$ is a continuous function that admits first and second order divided differences, $(f(x_1) + (A(x_1) + [x_0, x_1; g])(\cdot - x_1) + g(x_1) + F(\cdot))^{-1}$ is Aubin continuous at 0 for x_2, and $A : \mathbb{X} \to \mathbb{Y}$ stands for an approximation of $f' : \mathbb{X} \to \mathcal{L}(\mathbb{X}, \mathbb{Y})$, which satisfies a kind of Hölder-relaxed condition. That is, we assume there exist positive constants K, M, $\overline{M}$ and real numbers $\upsilon \in [0, 1]$ and $\xi \in [0, 1]$ satisfying

$$\|f'(x) - f'(y)\| \leq K\|x - y\|^{\xi}, \quad \|[x, y, z; g](z - x)\| \leq \overline{M}$$
$$\|f'(x_1) - A(x)\| \leq M\|x_1 - x\|^{\upsilon} + m, \quad m \geq 0.$$

A similar result to the case $g = 0$ and $F \equiv 0$ was first obtained by J. Rokne in [23]. Thus, our convergence result constitutes an extension to the one obtained by Rokne.

The remainder of this study is arranged as follows. In Section 2, notations and important results for supporting our main results are introduced. In Section 3, the central result is proposed and proved as well a numerical example satisfying all the conditions of the main theorem. In Section 4, we study a version of the Dennis-Moré theorem for solving (1.1). In Section 5, numerical examples are presented to demonstrate the effectiveness of the proposed method. Finally, concluding remarks are provided in Section 6.

2 Preliminaries

In this section, we present some notations and important results required during the development of this study. Let $\mathbb{X}$ and $\mathbb{Y}$ be Banach spaces. Let the open and closed balls of radius $\delta > 0$, centered at x, be denoted by

$$\mathbb{B}_\delta(x) := \{y \in \mathbb{X} : \|x - y\| < \delta\}, \qquad \mathbb{B}_\delta[x] := \{y \in \mathbb{X} : \|x - y\| \le \delta\}.$$

We denote by $\mathcal{L}(\mathbb{X}, \mathbb{Y})$ the vector space consisting of all continuous linear mappings $T : \mathbb{X} \to \mathbb{Y}$, and the norm of T is defined by $\|T\| := \sup\{\|Tx\| : \|x\| \le 1\}$. Let $f : \Omega \to \mathbb{Y}$ be differentiable in an open set $\Omega \subseteq \mathbb{X}$. The linear mapping $f'(x) : \mathbb{X} \to \mathbb{Y}$, which is continuous, denotes the derivative of f at x. Let $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a multifunction. Then, the graph, domain, and range of the multifunction F are the sets gph $F := \{(x, u) \in \mathbb{X} \times \mathbb{Y} \ : \ u \in F(x)\}$, dom $F := \{x \in \mathbb{X} \ : \ F(x) \neq \varnothing\}$, and rge $F := \{u \in \mathbb{Y} \ : \ u \in F(x)$ for some $x \in \mathbb{X}\}$. The multifunction $F^{-1} : \mathbb{Y} \rightrightarrows \mathbb{X}$ defined by $F^{-1}(u) := \{x \in \mathbb{X} \ : \ u \in F(x)\}$ denotes the inverse of F. Let C and D be subsets of $\mathbb{X}$; then,

$$d(x, D) = \inf_{y \in D} \|x - y\|, \qquad e(C, D) := \sup_{x \in C} d(x, D)$$

respectively define the distance from x to D and excess of C beyond D. The following conventions are adopted: $d(x, D) = +\infty$, when $D = \varnothing$, $e(\varnothing, D) = 0$, when $D \neq \varnothing$, and $e(\varnothing, \varnothing) = +\infty$. The following definition is an important consideration in the subsequent analysis.

Definition 1. Let $\mathbb{X}$ and $\mathbb{Y}$ be two Banach spaces. The first order divided difference of the operator $g : \mathbb{X} \to \mathbb{Y}$ on the points $x, y \in \mathbb{X}$ is denoted by $[x, y; g]$, if the following conditions hold:

$$[x, y; g](y - x) = g(y) - g(x), \quad x, y \in \mathbb{X}, \ x \neq y.$$

The second order divided difference of g on the points $x, y, z \in \mathbb{X}$ is denoted by $[x, y, z; g]$, if the following conditions hold:

$$[x, y, z; g](z - x) = [y, z; g] - [x, y; g], \ \ x, y, z \in \mathbb{X}, \ x \neq y, \ x \neq z, \ y \neq z.$$

If g is Fréchet differentiable at x, we define $[x, x; g]$ as $g'(x)$ and, if g is twice differentiable at x, then $[x, x, x; g]$ is defined as $\frac{1}{2}g''(x)$.

DEFINITION 2. A multifunction $\Gamma : \mathbb{Y} \rightrightarrows \mathbb{X}$ is Aubin continuous at $(y_0, x_0) \in gph(\Gamma)$, with modulus $\kappa > 0$ and radii $a > 0$ and $b > 0$, if the following inequality holds

$$e(\Gamma(y_1) \cap \mathbb{B}_a(x_0), \Gamma(y_2)) \leq \kappa \, \|y_1 - y_2\|, \quad \text{for all} \quad y_1, y_2 \in \mathbb{B}_b(y_0).$$

For more details about the Aubin property the reader can see [9,10,14]. In the following, we present the notion of metric regularity, which plays an important role in mathematical analysis.

DEFINITION 3. Let $\Omega \subset \mathbb{X}$ be open and nonempty. A set-valued mapping $H : \Omega \rightrightarrows \mathbb{Y}$ is said to be metrically regular at $\bar{x} \in \Omega$ for $\bar{u} \in \mathbb{Y}$, when $\bar{u} \in H(\bar{x})$, the graph of H is locally closed at $(\bar{x}, \bar{u})$ and there exist constants $\kappa > 0$, $a > 0$ and $b > 0$ such that $B_a[\bar{x}] \subset \Omega$ and

$$d(x, H^{-1}(u)) \leq \kappa d(u, H(x)), \quad \forall (x, u) \in \mathbb{B}_a[\bar{x}] \times \mathbb{B}_b[\bar{u}].$$

In addition, if the mapping $B_b[\bar{u}] \ni u \mapsto H^{-1}(u) \cap \mathbb{B}_a[\bar{x}]$ is single-valued, then H is called strongly metrically regular at $\bar{x} \in \Omega$ for $\bar{u} \in \mathbb{Y}$ with associated constants $\kappa > 0$, $a > 0$ and $b > 0$.

Remark 1. It is known that a multifunction $\Gamma : \mathbb{Y} \rightrightarrows \mathbb{X}$ has the Aubin property at y_0 for x_0 with constant $\kappa > 0$ if and only if $\Gamma^{-1} : \mathbb{Y} \rightrightarrows \mathbb{X}$ is metrically regular at x_0 for y_0 with the same constant κ, see [14, Theorem 5A.3, p. 255]. If f is a function which is strictly differentiable at some point x_0, then the Aubin continuity of f^{-1} at $(f(x_0), x_0)$ is equivalent to the surjectivity of $f'(x_0)$, by the Graves theorem, see [8,17]. Moreover, it is known that if $f : \mathbb{X} \to \mathbb{Y}$ is a function which is strictly differentiable at x_0, $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ is a multifunction with closed graph and $y_0 \in f(x_0) + F(x_0)$, then the Aubin continuity of the multifunction $(f + F)^{-1}$ at (y_0, x_0) is equivalent to the Aubin continuity of the multifunction $(f(x_0) + f'(x_0)(\cdot - x_0) + F(\cdot))^{-1}$ at (y_0, x_0). Also, it is shown in [13] that Aubin property of $f(x_0) + f'(x_0)(\cdot - x_0) + N_C(\cdot)$ at (y_0, x_0) is equivalent to the strong metric regularity of the same multifuntion at the same point. See a detailed discussion on this topic in [12].

Next, we present a weak regularity assumption, namely the strong metric subregularity concept. This property will be used in the sequel to show the superlinear convergence rate of our proposed method.

DEFINITION 4. A multifunction $\Gamma : \mathbb{X} \rightrightarrows \mathbb{Y}$ is said to be strongly metrically subregular at $\bar{x}$ for $\bar{y}$ when $(\bar{x}, \bar{y}) \in gph(F)$ and there is a constant $\kappa > 0$ together with a neighborhood U of $\bar{x}$ such that

$$\|x - \bar{x}\| \leq \kappa d(\bar{y}, F(x)), \tag{2.1}$$

for all $x \in U$.

Next, we present a class of mappings f and F for which the multifunction

$$\Gamma_{f+F}(x, y) := f(x) + f'(x)(y - x) + F(y).$$

has the Aubin property. Firstly, we will define the following concept:

Definition 5. A mapping $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be monotone if

$$\langle y' - y, x' - x\rangle \geq 0, \quad \forall (x, y),\, (x', y') \in \operatorname{gph} G.$$

If G is monotone and its graph is maximal with respect to this property, i.e., it is not properly contained in the graph of any other monotone operator, then we say that G is maximal monotone.

Proposition 1. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be maximal monotone. Assume that $x_* \in \mathbb{R}^n$ and $\beta > 0$ satisfy the following condition:*

$$\langle f'(x_*)\, p, p\rangle \geq \beta \|p\|^2, \qquad \forall\, p \in \mathbb{R}^n.$$

Then, rge $\Gamma_{f+F}(x_*, \cdot) = \mathbb{R}^n$, and for each $\bar{x} \in \mathbb{R}^n$ and $\bar{u} \in \Gamma_{f+F}(x_*, \bar{x})$, the multifunction $\Gamma_{f+F}(x_*, \cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ has the Aubin property at $\bar{x} \in \mathbb{R}^n$ for $\bar{u} \in \mathbb{R}^n$, with constants $\kappa = 1/(\alpha + \beta)$, $a = +\infty$, and $b = +\infty$.

Proof. See [5, Proposition 1]. □

Remark 2. The previous result says that if f is a continuously differentiable function and F is a maximal monotone operator then $\Gamma_{f+F}(x, y)$ has the Aubin property. In particular, $\Gamma_{f+F}(x, y)$ is strongly metrically regular, see Remark 1.

We end this section by presenting a generalization of the contraction mapping principle for multifunctions, whose proof is found in [14, Theorem 5E.2, p. 313].

Theorem 1. *Let $\Gamma : \mathbb{X} \rightrightarrows \mathbb{Y}$ be a multifunction and let $x_0 \in \mathbb{X}$. Suppose that there exist scalars $\rho > 0$ and $\lambda \in (0, 1)$ such that the set* $\operatorname{gph} \Gamma \cap (\mathbb{B}_\rho[x_0] \times \mathbb{B}_\rho[x_0])$ *is closed and the following conditions hold:*

(i) $d(x_0, \Gamma(x_0)) \leq \rho(1 - \lambda)$;

(ii) $e\,(\Gamma(p) \cap \mathbb{B}_\rho[x_0], \Gamma(q)) \leq \lambda \|p - q\|$ *for all* $p, q \in \mathbb{B}_\rho[x_0]$.

Then, Γ has a fixed point in $\mathbb{B}_\rho[x_0]$. That is, there exists $y \in \mathbb{B}_\rho[x_0]$ such that $y \in \Gamma(y)$.

3 Newton-Kantorovich theorem

This section establishes a Newton-Kantorovich theorem by assuming the Aubin property (see Definition 2), the Hölder continuity of the derivative f', and the continuity of the function g. Besides, we suppose that the function g admits first and second order divided differences. It is important to note that although we assume the differentiability of f, which means we need some assumptions about f', we also suppose that there is some approximation for f' and, in each iteration of our proposed method (1.4) we do not need to compute $f'(x_k)$.

Theorem 2. *Let $x_0, x_1 \in \mathbb{X}$ and $\eta > 0$ be given. Suppose $f : \mathbb{X} \to \mathbb{Y}$ is Fréchet differentiable in a neighborhood $\mathcal{O}$ of x_1, $g : \mathbb{X} \to \mathbb{Y}$ is a continuous function that admits first and second order divided differences in $\mathcal{O}$, and $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ is a multifunction with closed graph. Suppose the following conditions are hold for all $x, y, z \in \mathcal{O}$:*

(i) there exists $K > 0$, $\overline{M} > 0$, and $\xi \in [0,1]$ such that

$$\|f'(x) - f'(y)\| \leq K\|x - y\|^{\xi}, \qquad \|[x, y, z; g](z - x)\| \leq \overline{M}; \tag{3.1}$$

(ii) there exists a linear mapping $A : \mathbb{X} \to \mathbb{Y}$ stands for an approximation of $f' : \mathbb{X} \to \mathcal{L}(\mathbb{X}, \mathbb{Y})$, with the positive constant M, m and exponent $\upsilon \in [0,1]$ such that

$$\|f'(x_1) - A(x)\| \leq M\|x_1 - x\|^{\upsilon} + m; \tag{3.2}$$

(iii) assume that $\|x_1 - x_0\| \leq \frac{\eta}{2}$, and let x_2 be obtained by the Newton-secant method (1.4) from x_0 and x_1 such that $\|x_2 - x_1\| \leq \frac{\eta}{2}$ and the multifunction

$$(f(x_1) + g(x_1) + (A(x_1) + [x_0, x_1; g])(\cdot - x_1) + F(\cdot))^{-1}$$

is Aubin continuous at 0 for x_2 with the associated radii a and b and modulus $\kappa > 0$;

(iv) there exists a number $s > \eta$ satisfying

$$\begin{cases} 2s - \eta/2 & < \quad a, \\ \frac{K}{\xi+1}(s)^{\xi+1} + ms + (M(s)^{\upsilon} + 2m)s + 3\overline{M}s & \leq \quad b, \\ \kappa\big(Ms^{\upsilon} + 2m + 2\overline{M}\big) & < \quad 1, \\ s(h-1) + \eta & \leq \quad 0, \end{cases}$$

where

$$h := \frac{\kappa}{1-\kappa\big(Ms^{\upsilon}+2m+2\overline{M}\big)}\left(\frac{K}{\xi+1}\eta^{\xi}+Ks^{\xi} + Ms^{\upsilon} + m + \overline{M}\right); \tag{3.3}$$

(v) Suppose that $t_0 = 0$, $t_1 = \eta/2$, $t_2 = \eta$, and for $k \geq 2$

$$t_{k+1} - t_k = \lambda(t_k)\left(\frac{K}{\xi+1}(t_k - t_{k-1})^{\xi} + Kt_{k-1}^{\xi} + Mt_{k-1}^{\upsilon} + m + \overline{M}\right), \tag{3.4}$$

where $\lambda(t_k) = \kappa(t_k - t_{k-1})/\big(1 - \kappa(Mt_k^{\upsilon} + 2m + 2\overline{M})\big)$.

Then, the sequence $\{x_k\}$ generated by (1.4) is well defined for every $k \geq 0$, $x_k \in \mathbb{B}_{t^*}(x_0)$, and converges to a point $x_* \in \mathbb{B}_{t^*}(x_0)$ such that $f(x_*) + g(x_*) + F(x_*) \ni 0$, where t^* is the limit point of the sequence $\{t_k\}$ defined in (3.4).

Remark 3. Assumption (iii) above is extremely important in the convergence analysis of the quasi-Newton method (1.4). For example, if $g = 0$ and f satisfies (2.3) then it follows from Proposition 1 that it is always verified, because in the definition of Aubin property $a = +\infty$ and $b = +\infty$. That is, from x_0 and x_1 we can always derive a point x_2 such that

$$(f(x_1) + A(x_1)(\cdot - x_1) + F(\cdot))^{-1}$$

is Aubin continuous at 0 for x_2. In Section 5, we will present another important example where Assumption (iii) is fulfilled.

The proof of Theorem 2 is based on the following two results:

Lemma 1. *Let $\{t_k\}$ be the sequence defined by (3.4), and let condition (iv) of Theorem 2 hold. Then, there exists $t^* \le s$ such that $\{t_k\}$ converges to t^*.*

Proof. Taking into account that s and η are positive constants and $s(h-1)+\eta \le 0$, by (iv), we conclude that $h < 1$. As a consequence we have $\frac{\eta}{1-h} \le s$. In the subsequent steps, we use induction to prove that $\{t_k\}$ is a non-decreasing and bounded sequence. By (3.4), with $k = 2$, we have

$$\begin{aligned} t_3 - t_2 =& \frac{\kappa}{1-\kappa\big(M\eta^{\upsilon}+2m+2\overline{M}\big)}\left[\frac{K}{\xi+1}\left(\frac{\eta}{2}\right)^{\xi} + K\left(\frac{\eta}{2}\right)^{\xi}\right]\frac{\eta}{2} \\ &+ \frac{\kappa}{1-\kappa\big(M\eta^{\upsilon}+2m+2\overline{M}\big)}\left[M\left(\frac{\eta}{2}\right)^{\upsilon} + m + \overline{M}\right]\frac{\eta}{2} \ge 0. \end{aligned}$$

Using the condition $\eta < s$ and the definition of h in (3.3), we obtain

$$\begin{aligned} t_3 - t_2 \le& \frac{\kappa}{1-\kappa\big(Ms^{\upsilon}+2m+2\overline{M}\big)}\left(\frac{K}{\xi+1}\eta^{\xi} + Ks^{\xi} + Ms^{\upsilon} + m + \overline{M}\right)\eta \\ =& \, h\eta < \eta. \end{aligned}$$

Thus, we obtain that

$$t_3 \le t_2 + h\eta = \eta(1+h) = \frac{\eta}{1-h}(1-h^2) < \frac{\eta}{1-h} \le s.$$

Now, we assume, for all $n \le k$, that

$$t_n \le s, \;\; t_n - t_{n-1} \le \eta, \;\; t_n \ge t_{n-1}.$$

From (3.4), using that $(t_k - t_{k-1}) \ge 0$ and $t_k \le s$, we have

$$t_{k+1} - t_k = \lambda(t_k)\left(\frac{K}{\xi+1}(t_k - t_{k-1})^{\xi} + Kt_{k-1}^{\xi} + Mt_{k-1}^{\upsilon} + m + \overline{M}\right) \ge 0. \quad (3.5)$$

Moreover, using $t_k \le s$, $0 \le t_k - t_{k-1} \le \eta$ and $t_k \ge t_{k-1}$, from the above equality, we find that

$$t_{k+1} - t_k \le \frac{\kappa}{1-\kappa\big(Ms^{\upsilon}+2m+2\overline{M}\big)}\left(\frac{K}{\xi+1}\eta^{\xi} + Ks^{\xi} + Ms^{\upsilon} + m + \overline{M}\right)(t_k - t_{k-1}).$$

Using the condition $\eta < s$ and the definition of h in (3.3), we obtain

$$t_{k+1} - t_k \le h(t_k - t_{k-1}). \quad (3.6)$$

Given that $h < 1$ and $t_k - t_{k-1} \leq \eta$, we have

$$t_{k+1} - t_k \leq h\eta < \eta.$$

Repeating (3.6) successively, we obtain

$$t_{k+1} \leq t_k + h(t_k - t_{k-1}) \leq \ldots \leq t_1 + (h + \ldots + h^{k-1} + h^k)(t_1 - t_0). \quad (3.7)$$

Because $t_0 = 0$, $t_1 = \frac{\eta}{2}$ and $h < 1$, we conclude from (3.7) that

$$t_{k+1} \leq (t_1 - t_0)(1 + h + \cdots + h^k) = \frac{\eta}{2}\frac{1 - h^{k+1}}{1 - h}.$$

Again, using $h < 1$, we have $t_{k+1} \leq \frac{\eta}{1-h} \leq s$. Hence, using (3.5) and the above inequality, we conclude that $\{t_k\}$ is a monotone and bounded sequence, that is, there exists $t^* > 0$ such that $\lim t_k = t^* \leq s$. □

Lemma 2. *Suppose all of the assumptions of Theorem 2 hold. Then the sequence* $\{x_k\}$ *generated by the Newton iteration* (1.4) *satisfies*

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \quad k \geq 0, \quad (3.8)$$

where $\{t_k\}$ *is defined by* (3.4).

Proof. We prove (3.8) by induction on k. As $t_0 = 0$, $t_1 = \frac{\eta}{2}$, and $t_2 = \eta$, from assumption (iii), we have

$$\|x_1 - x_0\| \leq \eta/2 = t_1 - t_0 \text{ and } \|x_2 - x_1\| \leq \eta/2 = t_2 - t_1,$$

that is, (3.8) is true for $k = 0$ and $k = 1$. Now, we assume that $\|x_n - x_{n-1}\| \leq t_n - t_{n-1}$ for all $1 < n \leq k$. Thus, if $n \leq k$ then,

$$\|x_n - x_1\| \leq \sum_{j=2}^{n} \|x_j - x_{j-1}\| \leq \sum_{j=2}^{n} t_j - t_{j-1} \leq t_n - t_1 \leq t^* - t_1, \quad (3.9)$$

$$\|x_n - x_0\| \leq \sum_{j=1}^{n} \|x_j - x_{j-1}\| \leq \sum_{j=1}^{n} t_j - t_{j-1} = t_n \leq t^*. \quad (3.10)$$

Combining (3.9), (3.10), for every $x \in \mathbb{B}_{\|x_k - x_0\|}(x_k)$, we have

$$\|x - x_1\| \leq \|x - x_k\| + \|x_k - x_1\| \leq t^* + t^* - t_1 \leq 2t^* - \frac{\eta}{2} \leq 2s - \frac{\eta}{2} \leq a. \quad (3.11)$$

Now, we assume, for all $2 \leq n \leq k$, that x_n satisfies Theorem 2. For every $x \in \mathbb{B}_{\|x_k - x_0\|}(x_k)$, we define

$$P(x) := f(x_1) + g(x_1) + (A(x_1) + [x_0, x_1; g])(x - x_1) + F(x)$$

and the multifunction

$$\Phi_k(x) := P^{-1}[R(x)],$$

where

$$\begin{aligned}R(x) :=& f(x_1) + g(x_1) + (A(x_1) + [x_0, x_1; g])(x - x_1) - \big(f(x_k) + g(x_k)\big) \\ & - \big([x_{k-1}, x_k; g] + A(x_k)\big)(x - x_k).\end{aligned}$$

Next, we check all the conditions in Theorem 1. First, as (3.1) holds,

$$\|f(x) - f(x_{k-1}) - f'(x_{k-1})(x - x_{k-1})\| \leq \frac{K}{\xi + 1}\|x - x_{k-1}\|^{\xi+1} \tag{3.12}$$

is true for all $k \geq 1$. Meanwhile, taking into account Definition 1, we have

$$\begin{aligned}&\|g(x_1) - g(x_{k-1}) + [x_0, x_1; g](x - x_1) - [x_{k-2}, x_{k-1}; g](x - x_{k-1})\| \\ &\quad = \|[x_{k-2}, x_{k-1}, x_1; g](x_1 - x_{k-2})(x_1 - x_{k-1}) \\ &\qquad + [x_{k-2}, x_{k-1}; g](x_1 - x) + [x_0, x_1; g](x - x_1)\|.\end{aligned} \tag{3.13}$$

On the other hand, by using again Definition 1 we obtain

$$\begin{aligned}&\|[x_{k-2}, x_{k-1}; g](x_1 - x) - [x_0, x_1; g](x_1 - x)\| \\ &\leq \|[x_1, x_{k-2}, x_{k-1}; g](x_{k-1} - x_1) + [x_0, x_1, x_{k-2}; g](x_{k-2} - x_0)\|\|x_1 - x\|.\end{aligned} \tag{3.14}$$

Now, we combine (3.13) and (3.14) and use the second inequality in (3.1) to conclude that

$$\begin{aligned}&\|g(x_1) - g(x_{k-1}) + [x_0, x_1; g](x - x_1) - [x_{k-2}, x_{k-1}; g](x - x_{k-1})\| \\ &\qquad \leq \overline{M}\|x_1 - x_{k-1}\| + 2\overline{M}\|x_1 - x\|.\end{aligned} \tag{3.15}$$

Hence, combining (3.12), (3.15) and using (3.2) we have

$$\begin{aligned}&\|f(x_1) + g(x_1) + (A(x_1) + [x_0, x_1; g])(x - x_1) - \big(f(x_{k-1}) + g(x_{k-1})\big) \\ &\quad - \big([x_{k-2}, x_{k-1}; g] + A(x_{k-1})\big)(x - x_{k-1})\| \\ &\leq \|f(x_1) - f(x_{k-1}) - f'(x_1)(x_1 - x_{k-1})\| + \|(f'(x_1) - A(x_1))(x_1 - x_{k-1})\| \\ &\quad + \|(A(x_1) - A(x_{k-1}))(x - x_{k-1})\| + \|g(x_1) - g(x_{k-1}) \\ &\quad + [x_0, x_1; g](x - x_1) - [x_{k-2}, x_{k-1}; g](x - x_{k-1})\| \\ &\leq \frac{K}{\xi+1}\|x_1 - x_{k-1}\|^{\xi+1} + m\|x_1 - x_{k-1}\| + \big(M\|x_{k-1} - x_0\|^{\upsilon} + 2m\big)\|x - x_{k-1}\| \\ &\quad + \overline{M}\|x_1 - x_{k-1}\| + 2\overline{M}\|x_1 - x\|.\end{aligned} \tag{3.16}$$

Furthermore, taking into account that $x \in \mathbb{B}_{\|x_k - x_1\|}(x_k)$, (3.9), (3.10), (3.16) and $t^* \leq s$ we can conclude from (3.16) that

$$\begin{aligned}&\|f(x_1) + g(x_1) + (A(x_1) + [x_0, x_1; g])(x - x_1) - \big(f(x_{k-1}) + g(x_{k-1})\big) \\ &\quad - \big([x_{k-2}, x_{k-1}; g] + A(x_{k-1})\big)(x - x_{k-1})\| \\ &\leq \frac{K}{\xi + 1}(t^*)^{\xi+1} + mt^* + (M(t^*)^{\upsilon} + 2m)t^* + 3\overline{M}t^* \leq b.\end{aligned}$$

Secondly, we note that $x_k \in P^{-1}[S(x_k)]$, where

$$S(x_k) := f(x_1)+g(x_1)+(A(x_1)+[x_0,x_1;g])(x_k-x_1) - \big(f(x_{k-1})+g(x_{k-1})\big) \\ - \big([x_{k-2},x_{k-1};g]+A(x_{k-1})\big)(x_k-x_{k-1}).$$

Since $P^{-1}(\cdot)$ is Aubin continuous at 0 for x_2 with modulus κ and constants a and b, we obtain that

$$d(x_k,\Phi_k(x_k)) \le e\Big\{P^{-1}\big[S(x_k)\big]\cap \mathbb{B}_a(x_2),\Phi_k(x_k)\Big\} \\ \le \kappa(\|f(x_k)-f(x_{k-1})-f'(x_{k-1})(x_k-x_{k-1})\| \\ +\|f'(x_{k-1})-f'(x_1)\|\|x_k-x_{k-1}\|) + +\kappa(\|f'(x_1)-A(x_{k-1})\|\|x_k-x_{k-1}\| \\ +\|g(x_k)-g(x_{k-1})-[x_{k-2},x_{k-1};g](x_k-x_{k-1})\|).$$

Again, we use (3.12), (3.13), (3.2), and the second inequality in (3.1) to obtain the following estimate:

$$d(x_k,\Phi_k(x_k)) \le \frac{\kappa K}{\xi+1}\|x_k-x_{k-1}\|^{\xi+1} + \kappa K\|x_{k-1}-x_1\|^{\xi}\|x_k-x_{k-1}\| \\ +\kappa(M\|x_{k-1}-x_1\|^{\upsilon}+m+\overline{M})\|x_k-x_{k-1}\| \\ = \rho\big(1-\kappa(M\|x_k-x_1\|^{\upsilon}+2m+2\overline{M})\big), \\ \rho = C_k\|x_k-x_{k-1}\|, \tag{3.17}$$

where

$$C_k := \frac{\frac{\kappa K}{\xi+1}\|x_k-x_{k-1}\|^{\xi}+\kappa K\|x_{k-1}-x_1\|^{\xi}+\kappa(M\|x_{k-1}-x_1\|^{\upsilon}+m+\overline{M})}{1-\kappa\big(M\|x_k-x_1\|^{\upsilon}+2m+2\overline{M}\big)}.$$

However, using (3.11), if $p,q \in \mathbb{B}_{\|x_k-x_1\|}(x_k)$, after some manipulations, we have

$$e\big\{\Phi_k(p)\cap\mathbb{B}_{\|x_k-x_1\|}(x_k),\Phi_k(q)\big\} \le e\big\{\Phi_k(p)\cap\mathbb{B}_a(x_2),\Phi_k(q)\big\} \\ \le \kappa\|(A(x_1)-A(x_k))\|\|p-q\| + \kappa\|([x_0,x_1;g]-[x_{k-1},x_k;g])(p-q)\|.$$

Using the Definition 1 and the second inequality in (3.1), we have

$$\|[x_0,x_1;g]-[x_{k-1},x_k;g]\| = \|[x_0,x_1;g]-[x_k,x_0;g]+[x_k,x_0;g]-[x_{k-1},x_k;g]\| \\ = \|[x_k,x_0,x_1;g](x_1-x_k)-[x_{k-1},x_k,x_0;g](x_0-x_{k-1})\| \le 2\overline{M}.$$

Since $\|x_k-x_1\| \le t^* \le s$, $\|x_k-x_0\| \le t^* \le s$ and (3.2) holds, we conclude that

$$e\{\Phi_k(p)\cap\mathbb{B}_{\|x_k-x_1\|}(x_k),\Phi_k(q)\} \\ \le \kappa(\|A(x_1)-f'(x_k)\|+\|f'(x_k)-A(x_k)\|)\|p-q\| + 2\kappa\overline{M} \\ \le \kappa\big(M\|x_k-x_1\|^{\upsilon}+2m+2\overline{M}\big)\|p-q\| \le \kappa\big(Ms^{\upsilon}+2m+2\overline{M}\big)\|p-q\|.$$

Because $\kappa\big(Ms^{\upsilon}+2m+2\overline{M}\big) < 1$, we can apply Theorem 1 with $\Phi = \Phi_k$, $\overline{x} = x_k$, and $\lambda = \kappa\big(Ms^{\upsilon}+2m+2\overline{M}\big)$ to conclude that there exists $x_{k+1} \in \mathbb{B}_\rho[x_k]$ such that $x_{k+1} \in \Phi_k(x_{k+1})$. That is, x_{k+1} is a Newton iteration to (1.4) obtained from x_k. Moreover, we have

$$\|x_{k+1} - x_k\| \leq \rho, \tag{3.18}$$

where ρ is defined in (3.17). Using (3.18), (3.9), (3.10), $\|x_{k-2} - x_1\| \leq t_{k-2} \leq t_{k-1}$ and $\|x_n - x_{n-1}\| \leq t_n - t_{n-1}$ for $0 < n \leq k$, we have

$$\|x_{k+1}-x_k\| \leq \lambda(t_k)\left(\frac{K}{\xi+1}(t_k-t_{k-1})^{\xi}+Kt_{k-1}^{\xi}+Mt_{k-1}^{v}+m+\overline{M}\right) = t_{k+1}-t_k.$$

□

Proof. **[Proof of Theorem 2]** Using Lemma 1, $\{t_k\}$ converges to t^*, and using Lemma 2 we have

$$\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\| \leq \sum_{k=k_0}^{\infty} t_{k+1} - t_k = t_* - t_{k_0} < +\infty,$$

for any $k_0 \in \mathbb{N}$. Hence, we conclude that $\{x_k\}$ is a Cauchy sequence in $\mathbb{B}_{t^*}(x_0)$, and it converges to some x_*. Therefore, we can also conclude that:

$$\|x_* - x_k\| \leq t^* - t_k. \tag{3.19}$$

By (3.10) and (3.19) we have

$$\|x_* - x_0\| \leq \|x_* - x_k\| + \|x_k - x_0\| \leq t^* - t_k + t_k = t^*,$$

that is, $x_* \in \mathbb{B}_{t^*}(x_0)$. Because the definition of $\{x_k\}$ implies that

$$f(x_k) + g(x_k) + \big(A(x_k) + [x_{k-1}, x_k; g]\big)(x_{k+1} - x_k) + F(x_{k+1}) \ni 0, \tag{3.20}$$

$k = 1, 2, \ldots$, we use the continuity of f, g, and $A(x_k)$, and that F has closed graph, to conclude after passing the limit on k in (3.20) that

$$f(x_*) + g(x_*) + F(x_*) \ni 0, \quad x_* \in \mathbb{B}_{t^*}(x_0).$$

□

Corollary 1. Under the assumptions of Theorem 2 we have the following estimate

$$\|x_n - x_*\| \leq \frac{h}{1-h}(t_n - t_{n-1}), \quad \text{for all} \quad n \geq 1.$$

In particular we have linear convergence of $\{x_n\}$, i.e.,

$$\|x_n - x_*\| \leq \frac{\eta}{1-h}h^n, \quad \text{for all} \quad n \geq 1.$$

Proof. Since $\|x_{k+1} - x_k\| \leq t_{k+1} - t_k \leq h(t_k - t_{k-1}) \leq h^k\eta$, we show that

$$\|x_{m+n} - x_n\| \leq h(1 + h + \ldots + h^{m-1})(t_n - t_{n-1}).$$

Hence, taking into account $h < 1$ we obtain

$$\|x_{m+n} - x_n\| \leq \frac{h}{1-h}(t_n - t_{n-1}).$$

Letting $m \to +\infty$ we prove the corollary. □

Remark 4. We end this section by presenting a sufficient condition under which assumption (3.2) is satisfied by a suitable finite difference approximation A. Define

$$A = \left[\frac{f(x+he_1)-f(x)}{h}, \dots, \frac{f(x+he_n)-f(x)}{h}\right], \tag{3.21}$$

for some $h > 0$ and $e_j \in \mathbb{R}^n$ the $j-$th orthonormal vector of the canonical basis for $\mathbb{R}^n$. Since f' is $K-$Hölder continuous, then

$$\|f(x+he_i) - f(x) - hf'(x)e_i\| \leq \frac{K}{\xi+1} h^{\xi+1}$$

holds for all $x, y \in \mathbb{R}^n$. The previous inequality implies in

$$\left\| \left(\frac{f(x+he_i)-f(x)}{h}\right) - f'(x)e_i \right\| \leq \frac{K}{\xi+1} h^{\xi}.$$

Then, by using the definition of A in (3.21) we obtain that

$$\|(A - f'(x))e_i\| \leq \frac{K}{\xi+1} h^{\xi}.$$

Hence,

$$\|A - f'(x)\|^2 \leq \|A - f'(x)\|_F^2 = \sum_{i=1}^{n} \|(A - f'(x))e_i\|_2^2 \leq n\left(\frac{K}{\xi+1}\right)^2 h^{2\xi},$$

where $\|\cdot\|_F$ is the Frobenius norm. From the previous inequality we can conclude the following one:

$$\|A - f'(x)\| \leq \frac{\sqrt{n}K}{\xi+1} h^{\xi}. \tag{3.22}$$

Then, combining (3.1) and (3.22), we get

$$\|A - f'(x_1)\| \leq \|A - f'(x)\| + \|f'(x) - f'(x_1)\| \leq \frac{\sqrt{n}K}{\xi+1} h^{\xi} + K\|x - x_1\|^{\xi}.$$

Thus, if we take h satisfying

$$h^{\xi} < \frac{(\xi+1)(M\|x-x_1\|^{\upsilon} + m - K\|x - x_1\|^{\xi})}{\sqrt{n}K},$$

then (3.2) holds. In the particular case $\upsilon = \xi = 1$ and $M = K$, then we choose $A = f'$ in (3.2).

4 Dennis-Moré theorem

In this section, we consider the Newton-secant method (1.4) for solving the generalized equation (1.1), with $A(x_k) \in \mathcal{L}(\mathbb{X}, \mathbb{Y})$ meaning some kind of approximation for $f'(x_k)$. Firstly, we observe that if $f\colon \mathbb{X} \to \mathbb{Y}$ is Fréchet differentiable around $\bar{x}$, f' is continuous at $\bar{x}$ and $\{x_k\}$ is a sequence converging to $\bar{x}$, $x_{k+1} \neq x_k$ for all k, then

$$\lim_{k\to\infty} \frac{\|f(x_{k+1}) - f(x_k) - f'(\bar{x})s_k\|}{\|s_k\|} = 0, \tag{4.1}$$

where $s_k := x_{k+1} - x_k$.

In the remainder of this section, we link the analysis presented so far with a central result in the theory of quasi-Newton methods, namely, the Dennis-Moré theorem. This theorem, first published in [6], gives a characterization for the q-superlinear convergence of a quasi-Newton method applied to a smooth equation $f(x) = 0$ with a zero at $\bar{x}$ at which the derivative mapping $f'(\bar{x})$ is invertible. That is, if a quasi-Newton method generates a sequence x_k which stays near $\bar{x}$ and $x_{k+1} \neq x_k$ for all k, then x_k is convergent q-superlinearly if and only if it is convergent and, in addition,

$$\lim_{k\to\infty} \|\mathcal{E}_k s_k\|/\|s_k\| = 0, \tag{4.2}$$

where $\mathcal{E}_k := A(x_k) - f'(\bar{x})$.

The next result is a version of the classical Dennis-Moré theorem applied to nonsmooth generalized equations. We need to assume a strong assumption, namely, (4.3) below. In short, it shows that if (4.2) holds for some $A(x_k)$, $f + g + F$ is strongly metrically subregular at $\bar{x}$ for 0 and (4.3) holds, then every convergent sequence, in particular those whose existence is claimed in Theorem 2, is actually convergent superlinearly. To the best of our knowledge, by assuming the strong metric subregularity is the only way to prove the theorem below, see for instance [1, Theorem 4.9] and [4, Theorem 3.2].

Theorem 3. *Suppose that $f\colon \mathbb{X} \to \mathbb{Y}$ is a continuously differentiable function at $\bar{x}$, $g\colon \mathbb{X} \to \mathbb{Y}$ is a nonsmooth function that admits first order divided differences, $F : \mathbb{X} \rightrightarrows \mathbb{Y}$ is a multifunction with closed graph. Let $\{x_k\}$ be a sequence generated by the method* (1.4) *and assume that $\{x_k\}$ converges to $\bar{x}$ with $x_k \neq \bar{x}$ for all $k \geq 0$. Then, we can show that:*

a) If (4.2) *holds, then $\bar{x}$ is a solution of the generalized equation* (1.1).

b) If (4.2) *holds, the multifunction $f+g+F$ is strongly metrically subregular at $\bar{x}$ for* 0 *and*

$$\|[x_{k-1}, x_k, x_{k+1}; g]s_k\| \leq \eta\|s_k\|, \tag{4.3}$$

for each $\eta > 0$, then $x_k \to \bar{x}$ superlinearly.

Proof. We firstly assume that (4.2) holds. Taking into account (1.4), there exists a sequence $w_k \in F(x_{k+1})$ such that

$$0 = f(x_k) + g(x_k) + (A(x_k) + [x_{k-1}, x_k; g])s_k + w_k. \tag{4.4}$$

Then, since $x_k \to \bar{x}$, as $k \to \infty$, g is a continuous operator we obtain that $(A(x_k) + [x_{k-1}, x_k; g])s_k \to 0$, as $k \to \infty$. Indeed,

$$\|(A(x_k) + [x_{k-1}, x_k; g])s_k\| \leq \|\mathcal{E}_k s_k\| + \|[x_{k-1}, x_k; g]\|\|s_k\| + \|f'(\bar{x})\|\|s_k\|.$$

On the other hand, we know that $x_k \to \bar{x}$. Hence, passing limit on k in the Equation (4.4) we conclude that $w_k \to -f(\bar{x}) - g(\bar{x})$. Therefore, after passing the limit in the inclusion $w_k \in F(x_{k+1})$ we get $-f(\bar{x}) - g(\bar{x}) \in F(\bar{x})$, i.e., $\bar{x}$ is a solution of (1.1).

Finally, we assume that (4.2) holds and the multifunction $f + g + F$ is strongly metrically subregular at $\bar{x}$ for 0. Then, by using (2.1), we firstly conclude that there exists some $\kappa > 0$ such that, for k big enough, we have

$$\|x_{k+1} - \bar{x}\| \le \kappa d(0, f(x_{k+1}) + g(x_{k+1}) + F(x_{k+1})). \tag{4.5}$$

The inclusion (1.4) implies that

$$\begin{aligned} -f(x_k) - g(x_k) - f'(\bar{x})s_k - \mathcal{E}_k s_k - [x_{k-1}, x_k; g]s_k + f(x_{k+1}) + g(x_{k+1}) \\ \in f(x_{k+1}) + g(x_{k+1}) + F(x_{k+1}), \end{aligned} \tag{4.6}$$

for all $k \ge 0$. Thus, we combine (4.5) and (4.6) to conclude that

$$\begin{aligned} \|x_{k+1} - \bar{x}\| &\le \kappa\|f(x_{k+1}) - f(x_k) - f'(\bar{x})s_k\| + \kappa\|\mathcal{E}_k s_k\| \\ &\quad + \kappa\|g(x_{k+1}) - g(x_k) - [x_{k-1}, x_k; g]s_k\| \\ &= \kappa\|f(x_{k+1}) - f(x_k) - f'(\bar{x})s_k\| + \kappa\|\mathcal{E}_k s_k\| + \kappa\|[x_{k-1}, x_k, x_{k+1}; g]s_k\|. \end{aligned} \tag{4.7}$$

We shrink η, if necessary, to obtain $\eta < \epsilon$ for some $\epsilon > 0$. Thus, it follows from (4.3) that

$$\|[x_{k-1}, x_k, x_{k+1}; g]s_k\| \le \epsilon\|s_k\|.$$

Using this inequality, (4.1) and (4.2), we obtain from (4.7) that

$$\|x_{k+1} - \bar{x}\| \le 3\kappa\epsilon\|x_{k+1} - \bar{x}\| + 3\kappa\epsilon\|x_k - \bar{x}\|, \tag{4.8}$$

for k large enough and $\epsilon > 0$. Therefore, we use (4.8), and since we can take $\epsilon < 1/(3\kappa)$, we conclude that

$$\frac{\|x_{k+1} - \bar{x}\|}{\|x_k - \bar{x}\|} \le \frac{3\kappa\epsilon}{1 - 3\kappa\epsilon},$$

that is, $x_k \to \bar{x}$ superlinearly, since ϵ is positive and arbitrary. □

It is important to emphasize that this is the first time that a Dennis-Moré theorem is presented for such kind of problem (1.1). To an extensive study about the Dennis-Moré theorem for solving generalized equations we recommend the references [4, 11].

Remark 5. In next section, we will present an example that satisfies the hypothesis (4.3), see Example 4.

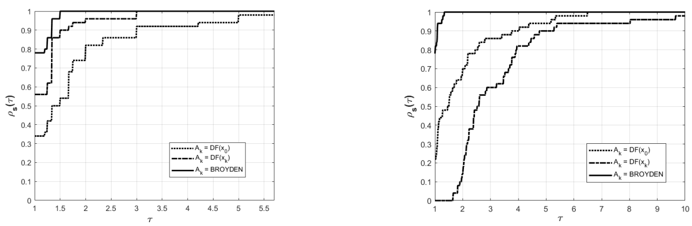
5 Numerical examples

In this section, we apply our proposed method to solve four finite-dimensional problems with non-isolated solutions. The algorithms were conducted in MATLAB R2019b on an 8 GB RAM Intel Core i7 notebook. Therein, d^k was computed using the subroutine FMINCON in the coded tests. The stop condition is $(f + g)(x^k) < 10^{-8}$. In the comparative study of the first two examples, following [7], the performance ratios are defined by $r_{p,s} = t_{p,s}/\min\{t_{p,j} : j \in S\}$ for $p \in \mathcal{P}$, $s \in S$, and the overall assessment of the performance of a particular solver s is given by $\rho_s(\tau) = \frac{1}{n_p}\text{card}\{p \in \mathcal{P} : r_{s,p} \le \tau\}$, where n_p is the number of problems in the set $\mathcal{P}$.

Example 1. Based on [21], let us consider the problem defined by:

$f, g\colon \mathbb{R}^4 \to \mathbb{R}^8,\ \ f(x) = (-\cos(0.2(x_1 - x_4)) - 3.2, x_3^2 - 4, -x_2 - x_3 - 2,$
$x_1 - x_4 - 5, x_4 - x_1 - 5, x_1^2 + x_2^2 + x_3^2 + x_4^2 - 8, x_1 + 2x_2 - 2, -2x_3 + x_4 + 2),$
$g(x) = (2|x_1 - x_2 + 2| + |x_2 - x_3 + 2x_4| + |x_3 + x_4 - 4|, |2x_1 - x_2|, -|x_4 + 2|,$
$0, 0, 0, 0, 0)^T$ and $F(x) = \mathbb{R}^6_+ \times \{0\} \times \{0\}$.

The point $x_* = (-0.64, 1.32, t, 2t - 2)$ is a solution of the problem for all $t \in [0.77, 1.11]$.



a) Number of iterations b) CPU time

Figure 1. Performance profiles for Example 1.

We have considered fifty randomly generated starting points x_0 in $[-2, 2]$, by setting the stop condition as $(u_1, u_2, \ldots, u_6, |u_7|, |u_8|) < 10^{-8}$, where $u = (f + g)(x_k)$. The results corresponding to the solved instances are represented in the performance profiles in Figure 1 according to the number of iterations and required CPU time. In this test, we observe that $A_k = A(x_k)$ defined by Quasi-Newton method with Broyden update gives better results in terms of the number of iterations and CPU time.

Example 2. Let us consider the nonsmooth system:

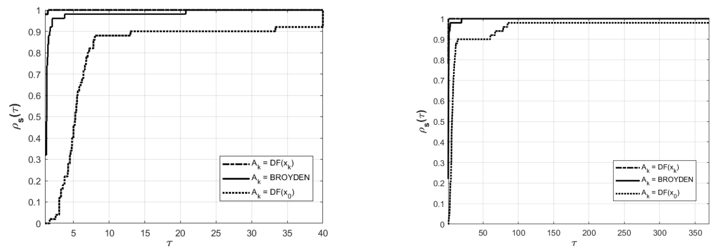
$$\begin{cases} x_1^2 + x_2^2 - |x_1 - 0.5| - 1 \le 0, \\ x_1^2 + (x_2 - 1)^2 - |x_1 - 0.5| - 1 \le 0, \\ (x_1 - 1)^2 + (x_2 - 1)^2 - 1 = 0. \end{cases}$$

Following [2], the point $x_* = \left(\frac{1}{2}, 1 - \frac{1}{2}\sqrt{3}\right)^T = (0.5, 0.1339745962)^T$ is one of the solutions of the system. In fact, the set of solutions is given by the arc of the circle with center $(1, 1)^T$ and radius $r = 1$ located between x_* and the point

$$x^{**} = \left(\frac{11}{26} - \frac{3}{13}\sqrt{3}, \frac{8}{13} + \frac{9}{26}\sqrt{3}\right)^T = (0.0233728904, 1.214940664)^T.$$

In this problem, we consider $f(x_1, x_2) = (x_1^2 + x_2^2 - 1, x_1^2 + (x_2 - 1)^2 - 1, (x_1 - 1)^2 + (x_2 - 1)^2 - 1)^T$, $g(x_1, x_2) = (-|x_1 - 0.5|, -|x_1 - 0.5|, 0)^T$, and $F(x) = \mathbb{R}^2_+ \times \{0\}$.

In Example 2, we have considered fifty randomly generated starting points x_0 in $[-2, 2]$, by setting the stop condition as $(u_1, u_2, |u_3|) < 10^{-8}$, where



a) Number of iterations b) CPU time

Figure 2. Performance profiles for Example 2.

$u = (f+g)(x_k)$. In Figure 2, we have presented performance profiles according to the number of iterations and required CPU time. Also, in Example 2, $A_k = A(x_k)$ defined by Newton method gives better results in terms of the number of iterations and CPU time.

In the last test, we reinforce the importance of our proposed method (1.4) as a good alternative scheme to solve (1.1).

Example 3. Based on the Structured Jacobian Problem 3 in [20], we consider the following system of inequalities:

$$\begin{cases} -2x_1^2 + 3x_1 - 2x_2 + 3x_{n-4} - x_{n-3} - x_{n-2} + 0.5x_{n-1} - x_n + 1 & \leq & 0, \\ -2x_j^2 + 3x_j - 2x_{j+1} + 3x_{n-4} - x_{n-3} - x_{n-2} + 0.5x_{n-1} - x_n + 1 & \leq & 0, \\ -2x_n^2 + 3x_n - x_{n-1} + 3x_{n-4} - x_{n-3} - x_{n-2} + 0.5x_{n-1} - x_n + 1 & \leq & 0, \end{cases}$$

$j = 2, \dots, n-1$ where, we take $x = (x_1, \dots, x_n)$ and consider:

$$f(x) = \left(-2x_1^2 + 3x_1 - 2x_2, \dots, -2x_j^2 + 3x_j - 2x_{j+1}, \dots, -2x_n^2 + 3x_n - x_{n-1}\right)^T,$$
$$g(x) = (3x_{n-4} - x_{n-3} - x_{n-2} + 0.5x_{n-1} - x_n + 1)(1, \dots, 1)^T, \quad F(x) = \mathbb{R}^n_+.$$

In Table 1, we have presented results by comparing (1.3) and our proposed method (1.4), in terms of optimality and CPU time. For both schemes, the optimality condition was reached in the first iteration, by setting the stop condition as $(f+g)(x_k) < 10^{-8}$.

Table 1. Numerical results with $x^0 = (0.9, \dots, 0.9)^T$.

Dimension (n)	Newton method (1.3), [19]		(1.4) with Broyden method	
	$\max(f+g)(x^k)$	CPU-time (s)	$\max(f+g)(x^k)$	CPU-time (s)
5	-0.071328	0.2702	-1.0741	0.1259
10	-0.071328	0.7775	-0.1490	0.4666
25	-0.071328	15.6580	-0.1955	12.6728
50	-0.071328	219.4207	-0.1738	171.3098

The last example illustrated how the proposed scheme (1.4) with Broyden's update can be a good solver for ill-conditioned problems. Moreover, we have that $x^* = -(t, t, \dots, t)$ with $t \in [1, 2]$ solves the problem.

We end this section by presenting an impontant example in nonlinear physical phenomena. This application can be seen in chemical physics, fluid mechanics, plasma physics and biology.

Example 4. We will consider the problem of finding a solution of the following nonlinear equation:

$$x(s) = h(s) + \frac{1}{2}\int_a^b K(s,t)\Phi(x(t))dt,\ \alpha \le s \le \beta, \tag{5.1}$$

and $K : [\alpha, \beta] \times [\alpha, \beta] \to \mathbb{R}$ is the Green's function. We first note that (5.1) can become a particular case of (1.1). In fact, we take $f \equiv 0$, $F \equiv 0$ and then we have the problem

$$g(x)(s) = x(s) - h(s) - \frac{1}{2}\int_\alpha^\beta K(s,t)\Phi(x(t))dt = 0.$$

We approximate the integral by the Gauss-Legendre quadrature formula

$$\int_\alpha^\beta K(s,t)\Phi(x(t))dt \approx \sum_{j=1}^{p} \omega_j K(t_i, t_j)\Phi(x(t_j)),$$

where the nodes t_i and the weights ω_i are known. Let $x_i := x(t_i)$ and $h_i = h(t_i)$. Thus,

$$x(s) - h(s) - \frac{1}{2}\int_\alpha^\beta K(s,t)\Phi(x(t))dt \approx x_i - h_i - \frac{1}{2}\sum_{j=1}^{p} a_{ij}\Phi(x_j),\ i = 1, 2, \ldots, p,$$

where

$$a_{ij} = \omega_j K(t_i, t_j) = \begin{cases} \omega_j \dfrac{(\beta - t_i)(t_j - \alpha)}{\beta - \alpha}, & j \le i, \\ \omega_j \dfrac{(\beta - t_j)(t_i - \alpha)}{\beta - \alpha}, & j \ge i. \end{cases}$$

Hence, (5.1) can be rewritten as

$$g(x) \equiv x - h - \frac{1}{2}Az = 0,\ g : \mathbb{R}^p \to \mathbb{R}^p, \tag{5.2}$$

where $x{=}(x_1, \ldots, x^p)^T$,$h{=}(h_1, \ldots, h_p)^T$, $A = (a_{ij})_{i,j=1}^p$,$z{=}(\Phi(x_1), \ldots, \Phi(x_p))^T$. We choose $K : [0, 1] \times [0, 1] \to \mathbb{R}$ as the Green's function, $\Phi(x(t)) = (x(t))^3 + |x(t)|$ and $h \equiv 0$. Then, (5.2) is the same as

$$g(x) \equiv x - \frac{1}{2}A(x_1^3 + |x_1|, \ldots, x_p^3 + |x_p|) = 0.$$

We can consider the first order divided differences that do not require g to be differentiable. Let $i, j = 1, 2, \ldots, p$. Then,

$$\begin{aligned}
[x, y; g]_{ij} &= \frac{1}{x_j - y_j}[g_i(x_1, \ldots, x_j, y_{j+1}, \ldots, y_p) - g_i(x_1, \ldots, x_{j-1}, y_j, \ldots, y_p)] \\
&= -\frac{1}{2(x_j{-}y_j)}(a_{i1}, \ldots, a_{in})(x_1^3{+}|x_1|, \ldots, x_j^3 + |x_j|, y_{j+1}^3{+}|y_{j+1}|, \ldots, y_p^3{+}|y_p|) \\
&+ \frac{1}{2(x_j{-}y_j)}(a_{i1}, \ldots, a_{in})(x_1^3{+}|x_1|, \ldots, x_{j-1}^3{+}|x_{j-1}|, y_j^3{+}|y_j|, \ldots, y_p^3 + |y_p|),
\end{aligned}$$

which implies in

$$[x,y;g]_{ij} = -\frac{a_{ij}}{2(x_j-y_j)}[x_j^3+|x_j|-(y_j^3+|y_j|)]$$
$$= -\frac{a_{ij}}{2}\left[\frac{x_j^3-y_j^3}{x_j-y_j}+\frac{|x_j|-|y_j|}{x_j-y_j}\right] = -\frac{a_{ij}}{2}\left[x_j^2+x_jy_j+y_j^2+\frac{|x_j|-|y_j|}{x_j-y_j}\right].$$

Let us note that

$$|[x,y;g]_{ij}| \le \left|\frac{a_{ij}}{2}\right| \left|\left[x_j^2+x_jy_j+y_j^2+\frac{|x_j|-|y_j|}{x_j-y_j}\right]\right|$$
$$\le \left|\frac{a_{ij}}{2}\right| \left[\left|x_j^2+x_jy_j+y_j^2\right|+|(|x_j|-|y_j|)/(x_j-y_j)|\right]$$
$$= \left|\frac{a_{ij}}{2}\right| \left[\left|x_j^2+x_jy_j+y_j^2\right|+\frac{||x_j|-|y_j||}{|x_j-y_j|}\right] = \frac{|a_{ij}|}{2}\left[\left|x_j^2+x_jy_j+y_j^2\right|+1\right].$$

Using Triangular's inequality, we get

$$|[x,y;g]_{ij}-[y,z;g]_{ij}| \le |[x,y;g]_{ij}|+|[y,z;g]_{ij}|$$
$$\le \frac{1}{2}|a_{ij}|\left[\left|x_j^2+x_jy_j+y_j^2\right|+\left|y_j^2+y_jz_j+z_j^2\right|+2\right].$$

As a consequence, there exists $\overline{m}_{ij}>0$ such that $|[x,y;g]_{ij}-[y,z;g]_{ij}| \le \overline{m}_{ij}$ for all $x,y,z\in\mathcal{O}$, a neighborhood of x_1. Hence, we can conclude that there exists $\overline{M}$ such that

$$\|[x,y,z;g](x-z)\| \le \overline{M}, \quad \forall\ x,y,z\in\mathcal{O}.$$

By taking all this into account we can conclude that

$$[x,y;g] = I-\frac{1}{2}A\,\mathrm{diag}\left(\begin{pmatrix} x_1^2+x_1y_1+y_1^2 \\ \vdots \\ x_p^2+x_py_p+y_p^2\end{pmatrix}+\begin{pmatrix}\frac{|x_1|-|y_1|}{x_1-y_1} \\ \vdots \\ \frac{|x_p|-|y_p|}{x_p-y_p}\end{pmatrix}\right). \quad (5.3)$$

Consequently, (5.3) guarantees the existence of the first order divided difference for the function g, as long as $x_j \neq y_j$. Hence, since $f=0$ and $F\equiv 0$, assumption (iii) in Theorem 2 is verified, since it reduces to find a solution of

$$g(x_1)+[x_0,x_1;g](\cdot - x_1) = 0,$$

which exists because the existence of first order divided difference for g.

Therefore, since all the hypotheses of Theorem 2 are fulfilled, we can apply this theorem to guarantee that the sequence defined in (1.4) converges to a solution of the problem (5.1).

Remark 6. Since x,y,z are in the same neighborhood of x_1, then we can take, in particular, $\|z-x\|$ small enough, that is, $\|z-x\|\le\eta$, for all $\eta>0$. Hence, if $\|A\|\le\frac{2\epsilon}{2+\eta}$, for all $\epsilon>0$, then $\|[x,y,z;g](x-z)\|\le\epsilon$, and the assumption in (4.3) holds.

6 Conclusions

We studied the solvability of generalized equations using divided differences in Banach spaces by considering a Kantorovich-like technique. As a byproduct, some results of Dontchev, Geoffroy, and Rokne, have been extended in part, by assuming a Hölder condition. A generalization of a Dennis-Moré theorem was also proved in order to obtain the superlinear convergence of the proposed method. Additionally, numerical examples are reported to illustrate the effectiveness of our approach.

Acknowledgements

The authors would like to thank the editor and referees for his valuable remarks and comments. The second author was supported in part by CNPq grant 309552/2022-2. The third author was supported in part by CNPq grants 401864/2022-7 and 306593/2022-0.

References

[1] F. J. Aragón Artacho, A. Belyakov, A.L. Dontchev and M. López. Local convergence of quasi-Newton methods under metric regularity. *Comput. Optim. Appl.*, **58**(1):225–247, 2014. https://doi.org/10.1007/s10589-013-9615-y.

[2] S. Bernard, C. Cabuzel, S.P. Nuiro and A. Piétrus. Extended semismooth Newton method for functions with values in a cone. *Acta Appl. Math.*, **155**(1):85–98, 2018. https://doi.org/10.1007/s10440-017-0146-x.

[3] E. Catinas. On some iterative methods for solving nonlinear equations. *Rev. Anal. Numér. Théor. Approx.*, **23**(1):47–53, 1994.

[4] R. Cibulka, A.L. Dontchev and M.H. Geoffroy. Inexact Newton methods and Dennis–Moré theorems for nonsmooth generalized equations. *SIAM J. Control Optim.*, **53**(2):1003–1019, 2015. https://doi.org/10.1137/140969476.

[5] F.R. de Oliveira, O.P. Ferreira and G.N. Silva. Newton's method with feasible inexact projections for solving constrained generalized equations. *Comput. Optim. Appl.*, **72**(1):159–177, 2019. https://doi.org/10.1007/s10589-018-0040-0.

[6] J.E. Dennis and J.J. Moré. A characterization of superlinear convergence and its application to quasi-Newton methods. *Math. Comput.*, **28**(126):549–560, 1974. https://doi.org/10.1090/S0025-5718-1974-0343581-1.

[7] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Math. Program.*, **91**(16):201–213, 2002. https://doi.org/10.1007/s101070100263.

[8] A.L. Dontchev. The Graves theorem revisited. *J. Convex Anal.*, **3**:45–53, 1996.

[9] A.L. Dontchev. Local analysis of a Newton-type method based on partial linearization. *Lect. Appl. Math.*, **32**(1):295–306, 1996.

[10] A.L. Dontchev. Uniform convergence of the Newton method for Aubin continuous maps. *Serdica Math. J.*, **22**(3):385–398, 1996. Available from Internet: `http://eudml.org/doc/11643`.

[11] A.L. Dontchev. Generalizations of the Dennis–Moré theorem. *SIAM J. Optim.*, **22**(3):821–830, 2012. https://doi.org/10.1137/110833567.

[12] A.L. Dontchev and W.W. Hager. An inverse mapping theorem for set-valued maps. *Proc. Amer. Math. Soc.*, **121**(2):481–489, 1994. https://doi.org/10.1090/S0002-9939-1994-1215027-7.

[13] A.L. Dontchev and R.T. Rockafellar. Characterizations of strong regularity for variational inequalities over polyhedral convex sets. *SIAM J. Optim.*, **6**(4):1087–1105, 1996. https://doi.org/10.1137/S1052623495284029.

[14] A.L. Dontchev and R.T. Rockafellar. *Implicit functions and solution mappings: A view from variational analysis*, volume 616. Springer, 2009. https://doi.org/10.1007/978-1-4939-1037-3.

[15] O.P. Ferreira and G.N. Silva. Local convergence analysis of Newton's method for solving strongly regular generalized equations. *J. Math. Anal. Appl.*, **458**(1):481–496, 2018. https://doi.org/10.1016/j.jmaa.2017.09.023.

[16] M.H. Geoffroy and A. Piétrus. Local convergence of some iterative methods for generalized equations. *J. Math. Anal. Appl.*, **290**(2):497–505, 2004. https://doi.org/10.1016/j.jmaa.2003.10.008.

[17] L.M. Graves. Some mapping theorems. *Duke Math. J.*, **17**:111–114, 1950. https://doi.org/10.1215/S0012-7094-50-01713-3.

[18] M.A. Hernández and M.J. Rubio. Semilocal convergence of the secant method under mild convergence conditions of differentiability. *Comp. Math. Appl.*, **44**(3-4):277–285, 2002. https://doi.org/10.1016/S0898-1221(02)00147-5.

[19] C. Jean-Alexis and A. Piétrus. On the convergence of some methods for variational inclusions. *Rev. R. Acad. Cien. serie A. Mat.*, **102**(2):355–361, 2008. https://doi.org/10.1007/BF03191828.

[20] W.J. Leong, M.A. Hassan and M.W. Yusuf. A matrix-free quasi-Newton method for solving large-scale nonlinear systems. *Comput. Math. Appl.*, **62**(5):2354–2363, 2011. https://doi.org/10.1016/j.camwa.2011.07.023.

[21] Q. Liu and J. Wang. A one-layer recurrent neural network for constrained nonsmooth optimization. *IEEE Trans. Syst. Man Cybern. Syst. Cybernetics*, **41**(5):1323–1333, 2011. https://doi.org/10.1109/TSMCB.2011.2140395.

[22] M.H. Rashid, J.H. Wang and C. Li. Convergence analysis of a method for variational inclusions. *Appl. Anal.*, **91**(10):1943–1956, 2012. https://doi.org/10.1080/00036811.2011.618127.

[23] J. Rokne. Newton's method under mild differentiability conditions with error analysis. *Numer. Math.*, **18**(5):401–412, 1971. https://doi.org/10.1007/BF01406677.

[24] P.S.M. Santos, G.N. Silva and R.C.M. Silva. Newton-type method for solving generalized inclusion. *Numer. Algor.*, **88**(1):1811–1829, 2021. https://doi.org/10.1007/s11075-021-01096-8.