TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Sự hội tụ địa phương của một kiểu phương pháp Newton gần đúng sử dụng mô hình tối ưu trong bài toán con

Trần Ngọc Nguyên, Nguyễn Văn Vũ*

Khoa Toán và Thống kê, Trường Đại học Quy Nhơn, Việt Nam
Ngày nhận bài: 23/02/2021; Ngày sửa bài: 14/04/2021;
Ngày nhận đăng: 26/04/2021; Ngày xuất bản: 28/06/2021

TÓM TẮT

Bài báo này tập trung vào nghiên cứu lược đồ lặp kiểu Newton gần đúng giải các phương trình suy rộng bao hàm ánh xạ đa trị trong trường hợp hữu hạn chiều. Chúng tôi đề xuất một chiến lược cập nhật động các bước lặp mới bằng cách đưa vào tại mỗi bước một mô hình quy hoạch toán học dựa trên dạng tuyến tính hóa của phần đơn trị xuất hiện trong phát biểu bài toán gốc. Chúng tôi cũng đưa ra kết quả phân tích hội tụ địa phương của lược đồ được đề xuất và áp dụng để xây dựng một thuật toán cấu trúc cho một lớp quan trọng là các bài toán bù. Một vài thực nghiệm số cũng được xem xét nhằm đánh giá bước đầu tính khả thi thực tiễn của phương pháp.

Từ khóa: *Ánh xạ đa trị, phương trình suy rộng, phương pháp Newton, tính nửa ổn định.*

*Tác giả liên hệ chính.
Email: nguyenvanvu@qnu.edu.vn

https://doi.org/10.52111/qnjs.2021.15306

QUY NHON UNIVERSITY
JOURNAL OF SCIENCE

Local convergence of an inexact Newton-type method involving optimization model on subproblems

Tran Ngoc Nguyen, Nguyen Van Vu*

Faculty of Mathematics and Statistics, Quy Nhon University, Vietnam
Received: 23/02/2021; Accepted: 14/04/2021;
Accepted: 26/04/2021; Published: 28/06/2021

ABSTRACT

This paper deals with inexact Newton-type scheme for solving generalized equation governed by set-valued mappings defined on finitely dimensional spaces. We proposed a new dynamical updating strategy by adapting in a mathematical program modeling based on the linearization of the single-valued part at each step. We investigated the local convergence behavior of the proposed framework and applied it to design a structural algorithm for solving complementarity problems. Implementation of several numerical tests was also considered to illustrate the feasibility of such framework.
Keywords: *Set-valued mapping, generalized equation, Newton-type method, semistability.*

1. Introduction

The Newton (or Newton-Raphson) method together with its extensions have been well-known in the literature as among of popular and efficient strategies for finding the zeros to a system of nonlinear functions. This is due to the good behavior of concrete algorithms designed from such manner, especially, the high growth of the convergence under mild assumptions on the input data. Particularly, when the functions defining the system are sufficiently smooth (of $C^{1,1}$ class for example), the corresponding Newton-based algorithms might be locally quadratically convergent (see[1]).

As motivated from certain problems in applications, many authors have extended the classical Newton framework to deal with the general model called by *generalized equation* (GE). Mathematically, an abstract GE defined on finite dimensional spaces can be formulated as inclusions involved set-valued map

$$0 \in \Phi(x) + N(x), \qquad (1)$$

where, the single-valued term $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is assumed to be smooth up to the necessary order, and the set-valued part $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ has closed graph. Problem of type (1) covers many concrete situations under the suitable choice of N. One of the most important case is the variational inequalities[2] which is closely related to mathematical program induced by selecting N to be the normal cone mapping associated with feasible region. For the sack of further reading, we refer to.[1,3] and the references therein.

One of the earliest framework dealing with abstract problem (1) is the famous Josephy-Newton method.[1,4] The core idea behind is to perturb (1) by replacing Φ with its linearization at each step and searching the next iterate as a solution to the auxiliary problem

$$0 \in \Phi(x^k) + \Phi'(x^k)(x - x^k) + N(x). \qquad (2)$$

Here and in what follows, x^k is meant to be the k^{th} iterate of the principal loop, and Φ' stands for the first-order derivative of Φ. More general, (2) can be subsumed as a particular case of the following scheme

$$0 \in A(x^k, x) + N(x), \qquad (3)$$

in which a perturbed term $A(x^k, \cdot)$ is in the position of Φ. (The typical Josephy-Newton framework

*Corresponding author.
Email: nguyenvanvu@qnu.edu.vn

https://doi.org/10.52111/qnjs.2021.15306

is induced from (3) if we let $A(u,v) = \Phi(u) + \Phi'(u)(v-u)$.) Under some mild assumptions, the reference methods will produce an iterative sequence (x^k) that converges to a solution of (1) with linear/superlinear/quadratic rate. The readers should consult[1,3,5,6] and the references therein for more about those topics.
As mentioned in,[1] one of the sharpest trend for the study of local convergence of (2) was appeared in the work,[7] and then, was considered in the later paper.[8] Another line for dealing with (2), sometimes is said to be semilocal convergence result, was begun at least from the work,[9] and then, by some later ones, e.g.[6,10,11] Perhaps the major difference between the two aforementioned strategies is that, the local convergence involves the information concerning an existing solution, while the other one almost concentrates on initial guess point.

Toward the numerical implementation aspect, the Newton-type frameworks mentioned above operate as follows: after getting an iterate, one constructs a partial linearization system corresponding to the original problem, and then solve exactly the immediate system to produces the next step. (Particularly, (2) and (3) are typical in such a manner.) Being quite different, some authors proposed several inexact schemes in order to solve variational inclusion (1). For instance, the paper[12] introduced an abstract iterative scheme based on the subproblem

$$\big(\Phi(x^k)+\Phi'(x^k)(x-x^k)+N(x)\big)\cap R_k(x^k,x) \neq \emptyset \quad (4)$$

for a family of set-valued maps $R_k : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^m$. While the authors[1] developed the algorithms for (1) using the updating process obtained by considering the perturbed inclusion

$$0 \in \Phi(x^k) + \Phi'(x^k)(x-x^k) + \Omega(x^k, x-x^k) + N(x) \quad (5)$$

whenever x^k is present. The inexact models allow us to apply suitably dynamical strategy of selecting the solvers to handle auxiliary problems that still ensure the convergence of overall process.

The current paper follows the idea of inexact Newton-type methods proposed in the literature with a few relaxations. Suppose now an iterate $x^k \in \mathbb{R}^n$ is known, then we would compute next step x^{k+1} via the perturbed GE

$$0 \in \Phi(x^k)+\Phi'(x^k)(x^{k+1}-x^k)+w^k+N(x^{k+1}), \quad (6)$$

for a perturbation term $w^k \in \mathbb{R}^n$. The performance of the overall procedure is strongly concerned with how efficiently one selects w^k and solve the inclusion (6). For instance, with aim of obtaining locally superlinear convergence of the resulting sequence (x^k), we can require that w^k satisfies the condition

$$\|w^k\| = o\big(\|x^{k+1}-x^k\| + \|x^k - x^*\|\big) \quad (7)$$

as mentioned,[1] where x^* is assumed to be an existing solution. Unfortunately, this dynamical choice for w^k seems to be slightly difficult to verify in practice, since it involves a posterior estimation. To avoid that drawback, we proposed a modified version of (7) by solving simultaneously x^{k+1} and w^k with some additional constraints. Precisely, let us introduce several auxiliary variables $d := x - x^k$, $z \in \mathbb{R}^n$ and consider the optimization problem

$$\begin{array}{ll} \min_{t\geqslant 0} & t \\ \text{subject to} & \Phi(x^k)+\Phi'(x^k)(d)+w+z=0, \\ & \|w\|^2 - t\|d\|^2 \leqslant 0, \\ & z \in N(x^k+d), \end{array} \quad (8)$$

for unknowns t, d, w and z. Once an exact/inexact solution to (8) is found, we update the next step $x^{k+1} = x^k + d^k$ (d^k is extracted from the previous procedure) and continue. Note that if (7) does hold, problem (8) will admit a feasible solution (t_k, d^k, w^k, z^k) such that $t_k \downarrow 0$. This demonstrates the possibility of implementing (8) in practice with the help of available optimization solvers.

The paper is organized as follows. In the next section we recall some basic notions used throughout the paper. Section 3 is devoted to introduce our generally inexact Newton-type framework and investigate its local convergence. In Section 4, we apply our method to a concrete important class of variational problem. The last section presents some numerical experiments to consider the practical performance of algorithms based on our approach.

2. Preliminaries

For the convenience of reading, we start by recalling some notions that will be used through the paper. The convention of notations used in the monograph will be applied throughout the paper. We frequently work with the a set-valued map $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ assigning to each $x \in \mathbb{R}^n$ a subset $N(x) \subset \mathbb{R}^m$ (may be empty). Such an object could be identified with its graph, defined by $\mathrm{Gr}(N) := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in N(x)\}$. For any usual map $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (equivalent terminology single-valued map), its (Frechét) derivative will be denoted by Φ'. While, notation Φ'' is meant to be the second-order corresponding derivative. When Φ is real-valued, we write $\nabla\Phi(x)$ and $\nabla^2\Phi(x)$ respectively to indicate its gradient vector and Hessian matrix at a given point x. Conventionally, all single-valued maps appeared in the paper are assumed to be differentiable up to the necessary order.

The problem of our interest is an abstract GE of the form

$$0 \in \Phi(x) + N(x), \tag{9}$$

where $\Phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a smooth map, and $N : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is set-valued whose graph is a closed set. The closedness assumption to $\mathrm{Gr}(N)$ allows for preserving the inclusion involving N after passing to the limit. We are interested in the iterative scheme governed by solving the subproblem of the form

$$0 \in \Phi(x^k) + \Phi'(x^k)(x - x^k) + \Omega(x^k, x - x^k) + N(x), \tag{10}$$

for some set-valued term $\Omega : \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$. To study the convergence behavior of such scheme, the next definition is useful.

Definition 1 (Semistable solution,[1])**.** Suppose that $x^* \in \mathbb{R}^n$ is a solution to GE (9). x^* is said to be semistable if for every $r \in \mathbb{R}^n$ any solution $u(r)$ to the perturbed inclusion

$$r \in \Phi(x) + N(x) \tag{11}$$

being close enough to x^* satisfies the estimate

$$\|u(r)\| = O(\|r\|) \quad \text{as } \|r\| \to 0. \tag{12}$$

Concern Definition 1, the next result is fruitful when dealing with the local convergence analysis of scheme updated through (10).

Theorem 2 ([1] Theorem 3.6)**.** *Let x^* be a semistable solution to the GE (9) for which the derivative Φ' is continuous at x^*. Let $\Omega : \mathbb{R}^n \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-value map satisfying the following assumption: for x close enough to x^*, the GE*

$$0 \in \Phi(x) + \Phi'(x)(u) + \Omega(x, u) + N(x + u) \tag{13}$$

has a solution $u(x)$ such that $u(x) \to 0$ as $x \to x^*$ and the estimate

$$\|\omega\| = o(\|u\| + \|x - x^*\|) \tag{14}$$

holds as $x \to x^*$, $u \to 0$ uniformly for $\omega \in \Omega(x, u)$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^n$ obeying the inclusion

$$0 \in \Phi(x) + \Phi'(x)(u) + \omega + N(x + u). \tag{15}$$

Then there exists $\delta > 0$ such that for any starting point $x^0 \in$ close enough to x^*, there exists a sequence $(x^k) \subset \mathbb{R}^n$ such that x^{k+1} is a solution to the GE (10) for each $k = 0, 1, 2, \dots,$ satisfying

$$\|x^{k+1} - x^k\| \leqslant \delta. \tag{16}$$

For any such sequence, x^k converges to x^* superlinearly. Moreover, the rate of convergence is quadratic provided the derivative Φ' is locally Lipschitz-continuous with respect to x^*, and provided (14) can be replaced with the estimate

$$\|\omega\| = O\big(\|u\|^2 + \|x - x^*\|^2\big). \tag{17}$$

Here, superlinear rate of convergence is meant to be $x^k \to x^*$ with

$$\limsup_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|} = 0. \tag{18}$$

If the relation

$$\limsup_{k\to\infty} \frac{\|x^{k+1} - x^*\|}{\|x^k - x^*\|^2} < +\infty \tag{19}$$

is in position of (18), one says that $x^k \to x^*$ quadratically. Further, local Lipschitz continuity of Φ' w.r.t x^* is equivalent to the assertion that

$$\limsup_{x \neq x' \to x^*} \frac{\|\Phi'(x) - \Phi'(x')\|}{\|x - x'\|} < +\infty. \tag{20}$$

3. Local analysis for a class of inexact Josephy-Newton method via optimization model

The current section is devoted to study an iterative scheme of inexact type for solving GE (9) by modifying the well-known Josephy-Newton method in the literature (see more, e.g.[4,6]). Such a framework based on the optimization model (8) in order to update $x^{k+1} = x^k + d^k$, where d^k is computed from an exact/approximating optimum to the nonlinear programming

$$\begin{array}{ll} \min_{t \geqslant 0} & t \\ \text{s.t.} & \Phi(x^k) + \Phi'(x^k)(d) + w + z = 0 \\ & d^T d \leqslant \rho_k^2 \\ & w^T w - t d^T d \leqslant 0 \\ & z \in N(x^k + d), \end{array} \tag{21}$$

where (ρ_k) is a positive sequence that determines the upper bound for actual step d^k. Comparing with (8), such a constraint is included to avoid the quite long step and ensure the local convergence property. This model can be viewed as a relaxation of the original subproblem involved in the Josephy-Newton scheme

$$0 \in \Phi(x^k) + \Phi'(x^k)(x - x^k) + N(x) \tag{22}$$

after adding the slack variable w. Also, (21) is subsumed as a particular case to (18), where $\Omega(x^k, x - x^k)$ coincides with the w component of feasible region to problem (21). Algorithm 1 outlines the overall process of our proposed method.

https://doi.org/10.52111/qnjs.2021.15306

Remark 3. The core phase in Algorithm 1 is subproblem defined in (21) which simultaneously generates both the search direction d and the term of inexactness w. This is slightly different from the strategy given by (10), where the suitable candidate for w should be chosen via an existing set-valued map Ω. To the practical implementation aspect, our proposed algorithm using optimization model (21) is quite flexible rather than the one by means of (10). Nevertheless, the performance of overall process depends on how efficient we solve (21). Further, for concrete applications, it is also necessary to invoke some reasonable globalization framework in order to avoid the sensitivity when dealing with difficult problems. Those issues should not be in the scope of this paper, and for the analysis here, we assume in general that any of subproblem (21) is solved successfully throughout the principal loop of Algorithm 1.

Concern the behavior of Algorithm 1, we consider some assumptions below.

(A1). The sequence of parameters (ρ_k) appear in model (21) is chosen in some manner such that $\rho_k \downarrow 0$.

(A2). The mathematical program (21) admits at least one (optimal or almost optimal) solution that can be successfully computed at each step.

(A3). The optimal value t_k^{opt} returned by solving (21) satisfies

$$\limsup_{k\to\infty} \{t_k^{opt}\} = 0. \tag{23}$$

```
input  : x^0, Φ, N
output: sequence x^0, x^1, x^2, ...
k ← 0 ;
repeat                    /*principal loop*/
    Set up the model (21) ;
    Solve the model (exact/inexact);
    if solve (21) successfully then
        Extract d^k from solution to (21);
        Update x^{k+1} ← x^k + d^k;
        k ← k + 1
    else               /*subproblem fail*/
        Terminate the loop ;
    endif
until stopping criterion reached;/*end
 loop*/
```

Algorithm 1. Inexact Newton-type method involving optimization model of subproblems

The next theorem summarizes the local convergence result for the proposed method based on Algorithm 1.

Theorem 4 (Local analysis for Algorithm 1)**.** *Consider GE (9) whose single-valued part Φ is differentiable. Let x^* be semistable in the sense of Definition 1 and Φ' be continuous around x^*. If all assumptions (A1), (A2) and (A3) are fulfilled, then by starting at x^0 being close enough to x^*, the sequence (x^k) generated via Algorithm 1 converges to x^* superlinearly.*

Proof. We notice first that, if we set

$$r^k := \Phi(x^{k+1}) - \Phi(x^k) - \Phi'(x^k)(d^k) - w^k, \tag{24}$$

then it is obvious to see that x^{k+1} solves the GE

$$r^k \in \Phi(x) + N(x). \tag{25}$$

For the main proof, let us mimic the line of Theorem 2. Choosing some parameters $\kappa > 0$ and $0 < \delta, \epsilon < 1$ for which the following does hold: when $\|r\| \leqslant 2\epsilon$, any solution $u(r)$ to the perturbed GE $r \in \Phi(u) + N(u)$ with $\|u(r) - x^*\| \leqslant 2\delta$ will satisfy the estimate

$$\|u(r) - x^*\| \leqslant \kappa\|r\|. \tag{26}$$

Scaling $\epsilon > 0$ if necessary, we can suppose that

$$\|\Phi'(x) - \Phi'(x^*)\| \leqslant \delta, \ \forall \|x - x^*\| \leqslant 2\epsilon. \tag{27}$$

Finally, since $\rho_k \downarrow 0$, after skipping a few first indexes, it is possible to require $\rho_k \leqslant \epsilon$.

We now start by x^0 such that $\|x^0 - x^*\| \leqslant \epsilon$. By assumption (A2), (21) produces some triplet (t_0, d^0, w^0) which solves the GE

$$0 \in \Phi(x^0) + \Phi'(x^0)(d^0) + w^0 + N(x^0 + d^0) \tag{28}$$

such that $\|w^0\| \leqslant t_0\|d^0\|$ and $\|d^0\| \leqslant \rho_0$. As mentioned at the beginning, $x^1 = x^0 + d^0$ is a solution to the inclusion

$$r^0 \in \Phi(x) + N(x) \tag{29}$$

for $r^0 = \Phi(x^1) - \Phi(x^0) - \Phi'(x^0)(d^0) - w^0$. The Taylor's expansion applied to Φ yields

$$r^0 = \int_0^1 \left\{[\Phi'(x^0 + sd^0) - \Phi'(x^0)](d^0)\right\} ds - w^0. \tag{30}$$

Because of $\|x^0 - x^*\| \leqslant \epsilon$ and $\|d^0\| \leqslant \rho_0$, it does hold $\|x^0 + sd^0 - x^*\| \leqslant 2\epsilon$ when s varies in the interval $[0, 1]$. Thus, the estimate

$$\|\Phi'(x^0 + sd^0) - \Phi'(x^*)\| \leqslant \delta \tag{31}$$

is straightforward for $0 \leqslant s \leqslant 1$. The combination between (27), (30) and (31) gives us

$$\|r^0\| \leqslant 2\delta\|d^0\| + \|w^0\| \leqslant (2\delta + t_0)\|d^0\|. \tag{32}$$

Hence, assumption of Theorem 4 implies $\|r^0\| \leqslant 2\epsilon$. Consequently, we obtain

$$\|x^1 - x^*\| \leqslant \kappa\|r^0\| \leqslant \kappa(2\delta + t_0)\|d^0\|. \tag{33}$$

Taking into account the inequality $\|d^0\| \leqslant \|x^1 - x^*\| + \|x^0 - x^*\|$, we deduce from (33)

$$\|x^1 - x^*\| \leqslant \frac{\kappa(2\delta + t_0)}{1 - \kappa(2\delta + t_0)}\|x^0 - x^*\| < \|x^0 - x^*\|, \quad (34)$$

if the following is active

$$2\kappa(2\delta + t_0) < 1. \quad (35)$$

Because $\delta > 0$ can be made arbitrarily small and $t_k^{opt} \downarrow 0$, (35) could be supposed to satisfied.

In summary, under the condition (35) and $\|x^0 - x^*\| \leqslant \epsilon$, we obtain x^1 via Algorithm 1 and moreover

$$\|x^1 - x^*\| \leqslant \sigma\|x^0 - x^*\| \quad (36)$$

for some $0 \leqslant \sigma < 1$. This shows that the process above still applicable if x^1 is in position of x^0. As a result, the sequence (x^k) is well-defined with

$$\|x^{k+1} - x^*\| \leqslant \sigma\|x^k - x^*\|. \quad (37)$$

It can be derived from (37) that $\|x^k - x^*\| \to 0$. The superlinear rate of convergence is attained by applying the result presented in[1] Proposition 3.4. □

4. Application to nonlinear complementarity problems (NCPs)

The NCP corresponding to a smooth map $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ could be formulated as follows (see, e.g.[2])

$$x \geqslant 0, F(x) \geqslant 0, x^T F(x) = 0, \quad (38)$$

in which we denote as usual x^T the transpose of x being written as a column matrix. If F coincides with the gradient ∇f to a smooth real-valued function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, the NCP (38) reduces to the KKT system associated with minimization program[1]

$$\min_x \quad f(x) \quad \text{s.t.} \quad x \geqslant 0. \quad (39)$$

By setting $\Phi(x) = F(x)$ and

$$N(x) := \begin{cases} \{z \in \mathbb{R}^n \mid z \leqslant 0, z^T x = 0\}, & \text{if } x \geqslant 0 \\ \emptyset, & \text{otherwise}, \end{cases} \quad (40)$$

we recover GE (9) from (38). The optimization model (21) now reads

$$\begin{array}{ll} \min_{t \geqslant 0} & t \\ \text{s.t.} & F(x^k) + F'(x^k)(d) + w + z = 0 \\ & z^T(x^k + d) = 0 \\ & -d^T d + \rho_k^2 \geqslant 0 \\ & -w^T w + t d^T d \geqslant 0 \\ & x^k + d \geqslant 0 \\ & -z \geqslant 0. \end{array} \quad (41)$$

Under such configuration, the iterative process to find a solution of (38) based on our proposed scheme is described in Algorithm 2.

```
input  : x^0, F
output: sequence x^0, x^1, x^2, ...
k ← 0 ;
repeat                    /*principal loop*/
    Compute ρ_k ;
    Set up the model (41) ;
    Solve the model (exact/inexact);
    if successful solving then
        Extract d^k from solution to (41);
        Update x^{k+1} ← x^k + d^k;
        k ← k + 1;
    else                  /*subproblem fail*/
        Terminate the loop ;
    endif
until stopping criterion reached;
```

Algorithm 2. Inexact Newton-type method for NCP (38)

5. Numerical experiments

We implement Algorithm 2 with open-source software `GNU Octave`.[1] The testing model involving mathematical program (39) whose objective function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ belongs to the well-known family of Rosenbrock type[13]

$$f(x) := \sum_{i=1}^{n-1} [r(x_{i+1} - x_i^2)^2 + (x_i - 1)^2], \quad (42)$$

where $r > 0$ is some parameter. Throughout the tests, each input sample is randomly generated via the built-in program `randn`. To handle subproblems in the form of (41), we invoke some certain hooking solvers `ipopt`,[14] `minos`,[15] `snopt`.[16] At the moment, we benchmark the performance by comparing several features:

- the number of samples that stops at successful criteria;
- the number of samples that assumptions (A2), (A3) are fulfilled throughout the principal loop.

To decide whether termination is optimal or not, we adopt the merit function $\omega(x) = \|\text{diag}(x)F(x)\|$, where $\text{diag}(x)$ is the diagonal matrix whose diagonal entries are components of vector x. The threshold tolerance for stopping criteria maintained during the test is chosen as 10^{-6}, while the maximum number of iterations in principal loops is set

[1] Homepage: `https://www.gnu.org/software/octave/`

https://doi.org/10.52111/qnjs.2021.15306

to be 200. Here, we update dynamically $\rho_{k+1} := \min(\rho_k, 1/\log(k+1))$ and check assumption (A3) by comparing with the slowly convergent sequence $e_k = 1/k$. Tables 1 and 2 show the benchmarking results obtained after specific test which consists in $N = 20$ samples of dimension $n = 5$.

Features	ipopt	minos	snopt
nb. optimal	4	2	3
nb. limit	16	12	6
nb. failure	0	6	11

Table 1. Numerical implementation of Algorithm 2: `nb. optimal` (resp.`nb. limit`,`nb. failure`) indicates the number of founding optimal iterate (reaching limit of threshold/updating fail)

Features	ipopt	minos	snopt
nb. valid	10	13	8
nb. failure	10	6	12
nb. other	0	1	0

Table 2. Information concern Assumption (A3): `nb. valid` (resp. `nb. failure`, `nb. other`) shows the number of samples that satisfy (resp. not satisfy, be unsure)

From Table 1, we can see that our proposed model works pretty well. More specifically, the percentage of problems which are successfully solved by our model combined with `ipopt`, (resp. `minos` and `snopt`) is 100%, (resp. 70% and 50%). This means that the successful rate of our model is more than 70%. It is worth noting that `ipopt`, `minos` and `snopt` are general-purposed solvers which are not properly designed for our new model. It is really interesting to propose a particular algorithm for solving (41) which is more efficient than these above solvers.

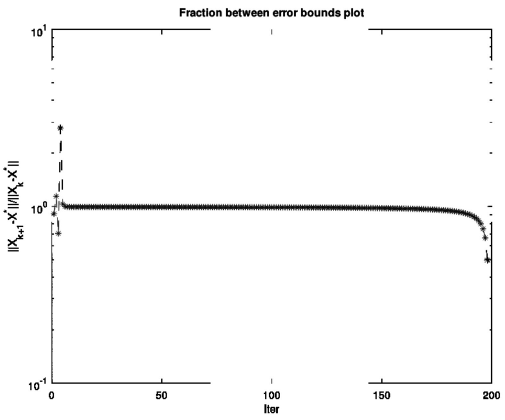

Figure 1. Superlinear convergence of Algorithm 2 with `ipopt` solver

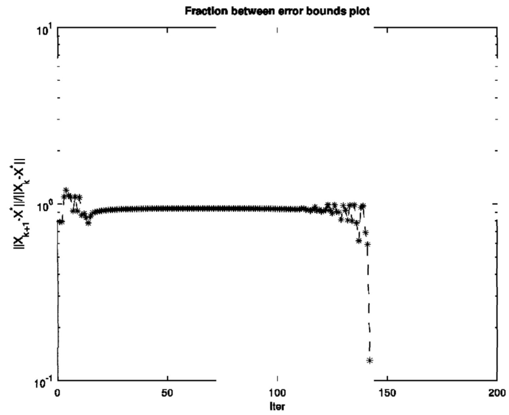

Figure 2. Superlinear convergence of Algorithm 2 with `minos` solver

Table 2 shows us the validity of assumption (A3). It can be seen that over 50% (31/60) of problems satisfy this assumption. According to Theorem 4, Algorithm 2 will attain a superlinear rate of convergence in these problems. To verify this assertion, we plot the behavior of Algorithm 2 when solving a problem satisfying assumption (A3) in Figures 1, 2 and 3. These figures give us the error distance of iterates in one certain successful sample that is extracted randomly. We can see that the error distances decrease faster and faster on last iterations. This implies the superlinear convergence in the corresponding cases.

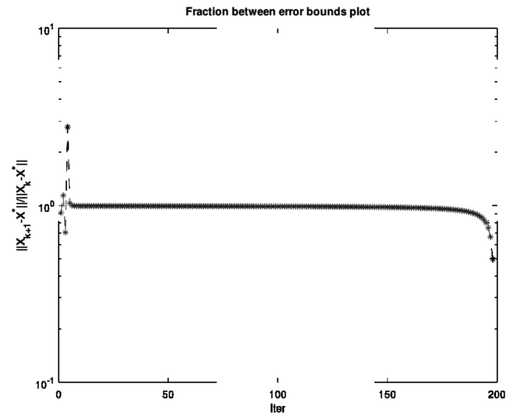

Figure 3. Superlinear convergence of Algorithm 2 with `snopt` solver

Remark 5. In Table 2 the only of interest is actually Assumption (A3), because the lack of validity for Assumption (A2) is almost the same with failure of overall process given in Table 1.

Acknowledgments: *This research is conducted within the framework of science and technology projects at institution level of Quy Nhon University under the project code T2020.653.01.*

References

1. A. F. Izmailov and M. V. Solodov, *Newton-type methods for optimization and variational problems*, Springer International Publishing, 2014.

2. F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems, Vol I*, 1st edition, Springer-Verlag New York, 2003.

3. A. L. Dontchev and R. T. Rockafellar, *Implicit functions and solution mappings: A view from variational analysis*, 2nd edition, Springer-Verlag New York, 2014.

4. N. H. Josephy, Newton's method for generalized equations, Technical report, University of Wisconsin, Madison, 1979.

5. S. Adly, R. Cibulka, and H. V. Ngai, Newton's method for solving inclusions using set-valued approximations, *SIAM J. Optim.*, **2015**, *25*, 159–184.

6. R. Cibulka, A. L. Dontchev, J. Preininger, V. Veliov, and T. Roubal, Kantorovich-type theorems for generalized equations, *J. Convex Anal.*, **2018**, *25*, 459–486.

7. J. F. Bonnans, Local analysis of Newton-type methods for variational inequalities and nonlinear programming, *Appl. Math. Optim.*, **1994**, *29*, 161–186.

8. A. F. Izmailov, A. S. Kurennoy, and M. V. Solodov, The Josephy–Newton method for semismooth generalized equations and semismooth SQP for optimization, *Set-Valued Var. Anal*, **2013**, *21*, 17–45.

9. A. L. Dontchev, *Local analysis of a Newton-type method based on partial linearization*, The mathematics of numerical analysis. 1995 AMS-SIAM summer seminar in applied mathematics, July 17–August 11, 1995, Park City, UT, USA. Providence, RI: American Mathematical Society, 1996, 295–306.

10. S. Adly, H. V. Ngai, and V. V. Nguyen, Newton's method for solving generalized equations: Kantorovich's and Smale's approaches, *J. Math. Anal. Appl.*, **2016**, *439*, 396–418.

11. M. H. Rashid, S. H. Yu, C. Li, and S. Y. Wu, Convergence analysis of the Gauss–Newton-type method for Lipschitz-like mappings, *J. Optim. Theory Appl.*, **2013**, *158*, 216–233.

12. A. L. Dontchev and R. T. Rockafellar, Convergence of inexact Newton methods for generalized equations, *Math. Program.*, **2013**, *139*, 115–137.

13. H. H. Rosenbrock, An automatic method for finding the greatest or least value of a function, *The Computer Journal*, **1960**, *3*, 175–184.

14. A. Wächter and L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Program.*, **2006**, *106*, 25–57.

15. B. A. Murtagh and M. A. Saunders, *A projected Lagrangian algorithm and its implementation for sparse nonlinear constraints*, Berlin, Heidelberg: Springer Berlin Heidelberg, 1982, 84–117.

16. P. E. Gill, W. Murray, and M. A. Saunders, Snopt: An sqp algorithm for large-scale constrained optimization, *SIAM Review*, **2005**, *47*, 99–131.

https://doi.org/10.52111/qnjs.2021.15306