Cent. Eur. J. Math. • 6(3) • 2008 • 469-481
DOI: 10.2478/s11533-008-0039-3

VERSITA

Central European Journal of **Mathematics**

A nonsmooth version of the univariate optimization algorithm for locating the nearest extremum (locating extremum in nonsmooth univariate optimization)

Research Article

Marek J. Śmietański*

Faculty of Mathematics and Computer Science, University of Łódź, 90-238 Łódź, ul. Banacha 22, Poland

Received 24 October 2007; accepted 14 May 2008

Abstract: An algorithm for univariate optimization using a linear lower bounding function is extended to a nonsmooth case by using the generalized gradient instead of the derivative. A convergence theorem is proved under the condition of semismoothness. This approach gives a globally superlinear convergence of algorithm, which is a generalized Newton-type method.

MSC: 65K05, 49M15

Keywords: univariate optimization • unconstrained optimization • linear bounding function • semismooth function

© *Versita Warsaw and Springer-Verlag Berlin Heidelberg.*

1. Introduction

This paper presents an algorithm for solving nonsmooth univariate minimization problem. A similar algorithm for smooth problems was introduced by Tseng in [18]. The iterative methods for higher dimensional problems involve steps which search for extrema along certain directions in R^n. These steps are equivalent to one-dimensional problems. Finding the step size along the direction vector involves solving the minimization subproblem which is an unidimensional search problem. Hence, the unidimensional search methods are most indispensable and the efficiency of any algorithm partly depends on them.
An unconstrained univariate optimization can be applied to solve many practical problems from the line search in multidimensional optimization via global optimization to minimax problems. The most popular and widely used derivative-free methods of minimization for univariate function are the uniform and dichotomous search, the golden section and the Fibonacci search methods (see e.g. Bazaraa, Sherali and Shetty [1]). All of these methods are only linearly convergent, since they successively reduce the width of the interval in every step. It is important to study these problems because mathematical approaches developed to solve them can be generalized to the multidimensional case

* *E-mail: smietan@math.uni.lodz.pl*

by numerous schemes (e.g. one-point based, diagonal, simplicial, space-filling curves and many more other popular approaches). On the other hand we can find some applications in global optimization (see e.g. Hansen, Jaumard and Lu [8]). Electrotechnics and electronics are among the fields where one-dimensional optimization methods can be used successfully (see Sergeyev et al. [15]). Because in many problems functions may not possess a sufficient degree of smoothness, algorithms for solving nondifferentiable unconstrained problems can be important tools. Moreover, constrained minimization problems with inequality constraints can be reduced to the unconstrained ones by using the penalty scheme $f_P(x) = f_0(x) + P\max\{g_1(x), g_2(x), \ldots, g_{N_v}(x), 0\}$ with the penalty coefficient P (see Famularo, Sergeyev and Pugliese [7]).
In this paper we apply Tseng's algorithm to the semismooth problem. We show that under some assumptions on a linear lower bounding function the algorithm is globally convergent. We prove that the algorithm for semismooth functions is superlinearly convergent, which characterizes Newton-like methods. Finally, this paper also presents examples and results of numerical experiments.

2. Preliminaries

Suppose that $f : R^n \to R$ and $F : R^n \to R^n$ are locally Lipschitz functions.

Definition 2.1.

The *generalized gradient* of f at x introduced by Clarke [4] is

$$\partial f(x) := \mathrm{co}\left\{\lim \nabla f\left(x^{(k)}\right) : x^{(k)} \to x,\ x^{(k)} \notin S,\ x^{(k)} \in D_f\right\},$$

where S is some set with Lebesgue measure 0 in R^n and D_f denotes the set of points at which f is differentiable.

Clarke showed [4] that if f is C^1 or differentiable at x and regular, then

$$\partial f(x) = \{\nabla f(x)\}.$$

Moreover, many other important properties of the generalized gradient were presented by Clarke; the most important of which are that $\partial f(x)$ is a compact subset of R^n and that ∂f is upper semicontinuous.
According to Rademacher's Theorem, the local Lipschitz continuity of F implies that F is differentiable almost everywhere.

Definition 2.2.

The *generalized Jacobian* of F at x in the sense of Clarke [4] is

$$\partial F(x) := \mathrm{co}\left\{\lim JF\left(x^{(k)}\right) : x^{(k)} \to x, x^{(k)} \in D_F\right\},$$

where D_F is the set of points at which F is differentiable and $JF\left(x^{(k)}\right)$ denotes the usual Jacobian matrix of partial derivatives of F at $x^{(k)}$.

Remark 2.1.

If $n = 1$, then $\partial F(x) = \partial F_1(x)$, i.e. the generalized gradient and the generalized Jacobian coincide (see Clarke [4]).

Qi and Sun in [14] proved the following:

Proposition 2.1.
Let x be a point in R^n. Assume that for any $h \in R^n$

$$\lim_{V \in \partial F(x+th), t \downarrow 0} Vh$$

exists. Then the classic directional derivative exists and is equal to this limit, i.e.

$$F'(x;h) = \lim_{V \in \partial F(x+th), t \downarrow 0} Vh.$$

The notion of semismoothness was originally introduced for functionals by Mifflin in [11]. Obviously, convex functions and smooth functions are semismooth. Scalar products and sums of semismooth functions are still semismooth functions (see Mifflin [11]). Moreover, piecewise smooth functions and maximum of a finite number of smooth functions are semismooth, too. The following definition is taken from Qi and Sun [14]:

Definition 2.3.
F is *semismooth* at x if F is locally Lipschitz at x and

$$\lim_{V \in \partial F(x+th'), h' \to h, t \downarrow 0} Vh'.$$

exists for every $h \in R^n$.

Semismooth functions have several properties which are very important in convergence analysis of methods in nonsmooth optimization. We need some of these properties for our later discussion:

Theorem 2.1 (Qi and Sun [14]).
Suppose that $F'(x;h)$ exists for every h at x. Then

(i) $F'(x;\cdot)$ is Lipschitz;

(ii) for any h there exists a $V \in \partial F(x)$ such that $F'(x;h) = Vh$.

Theorem 2.2 (Qi and Sun [14]).
The following statements are equivalent:

(i) F is semismooth at x;

(ii) for every $V \in \partial F(x+h)$ as $h \to 0$,

$$Vh - F'(x;h) = o(\|h\|);$$

(iii) $\lim_{h \to 0} \frac{F'(x+h;h) - F'(x;h)}{\|h\|} = 0$.

Remark 2.2.
In the original version of statement (iii) in the above theorem the assumption $x + h \in D_F$ is added. Without loss of generality, this assumption may be removed – the proof is analogous. From statement (iii) of the above theorem, it follows that if F has a strong Fréchet derivative at x, then F is semismooth at x.

Potra, Qi and Sun [13] presented the following:

Lemma 2.1.
Suppose that $F : R \to R$ is semismooth at x. Then the lateral derivatives $F'(x+) = F'(x,1)$ and $F'(x-) = F'(x,-1)$ exist and we have

$$\partial F(x) = \mathrm{co}\left\{F'(x+), F'(x-)\right\}.$$

Definition 2.4.

If for any $V \in \partial F(x+h)$, as $h \to 0$,

$$Vh - F'(x,h) = O\left(\|h\|^{1+p}\right),$$

where $0 < p \leq 1$, then we call F *p-order semismooth* at x.

Clearly, p-order semismoothness implies semismoothness. Following Qi and Sun [14] we remark that if F is semismooth at x, then for any $h \to 0$,

$$F(x+h) - F(x) - F'(x;h) = o\left(\|h\|\right),$$

and if F is p-order semismooth at x, then for any $h \to 0$,

$$F(x+h) - F(x) - F'(x;h) = O\left(\|h\|^{1+p}\right).$$

Potra, Qi and Sun [13] called the function F *strongly semismooth* if $p = 1$. Piecewise C^2 functions are examples of strongly semismooth functions.

Definition 2.5 (Pang [12]).

A function F is said to be *B-differentiable* at x if there exists a function $BF : R^n \to R^n$ called the *B-derivative* of F at x, which is positively homogeneous of degree 1 (i.e. $BF(x)(th) = tBF(x)h$ for all $h \in R^n$ and all $t \geq 0$), such that

$$\lim_{h \to 0} \frac{F(x+h) - F(x) - BF(x)h}{\|h\|} = 0.$$

Shapiro showed in [16] that a locally Lipschitz function F is B-differentiable at x if and only if it is directionally differentiable at x. In this case, the B-derivative and directional derivative are identical (see Harker and Xiao [9]).

Proposition 2.2 (Qi and Sun [14]).

If the B-derivative $BF(\cdot)$ is Lipschitz at x, then F is strongly semismooth at x.

We denote by $B(x,r)$ the closed ball in R^n with center x and radius r.

3. The linear bounding functions

First, we present the concept of linear lower bounding function (LLBF), which was introduced to develop optimization algorithms by Bromberg and Chang [2]. An LLBF for continuously differentiable functions was used in an algorithm for finding the nearest root of a function in a search direction by Tseng [18].

Definition 3.1.

Let $f : R \to R$ be a semismooth function and let I be a closed interval in R. f is said to have a *linear lower bounding function* (LLBF) l at the end point $p \in I$ *over* I, if l is a linear function which satisfies the following condition

$$f(x) \geq l(x) = \gamma(x-p) + f(p),$$

for all $x \in I$ and $\gamma \in R$ which is bounded by some G, i.e. $|\gamma| < G$.

Remark 3.1.

A linear upper bounding function (LUBF) l at a point p can be defined in the similar way, replacing $\geq$ by $\leq$.

Suppose that an LLBF (or LUBF) l of a function f is obtained over some interval $[a, b] \subset R$. We will call l a *left LLBF* (or a *left LUBF*) if l is obtained at a and a *right LLBF* (or a *right LUBF*) if l is obtained at b. When we refer to both LLBF and LUBF we shall use a shorter name, i.e. a *linear bounding function* (LBF).
Now we formulate some assumptions on LBF. Let f be a locally Lipschitz function, let y be any point and

$$l(x) = \gamma(\delta, y)(x - y) + f(y)$$

be a left LBF for f on $[y, y + \delta]$. We assume that $\gamma(\delta, y)$ have the following properties:

(A1) $\gamma : \mathbb{R}^+ \setminus \{0\} \times R \to R$ is an upper semicontinuous function;

(A2) for any $y \in R$

$$\gamma(\delta_1, y) \le \gamma(\delta_2, y) \text{ for } \delta_1 > \delta_2;$$

(A3) for any $y \in R$

$$\lim_{\delta \downarrow 0} \gamma(\delta, y) \in \partial f(y),$$

where $\partial f(y)$ denotes the generalized gradient of f at y.

All of the above assumptions (A1)–(A3) remain the same for a right LBF defined on $[y - \delta, y]$ except (A2) in which we assume that

$$\gamma(\delta_1, y) \le \gamma(\delta_2, y) \text{ for } \delta_1 < \delta_2.$$

Example 3.1.

We establish a semismooth function for which we give an LBF. Mifflin [11] proposed $f(x) = \log(1 + |x|)$ as an example of a semismooth function on R, which is neither convex nor differentiable. Note that in a neighborhood of $x = 0$,

$$f(x) = \max\{\log(1 + x), \log(1 - x)\},$$

that is, f is the pointwise maximum of two smooth functions. Moreover,

$$\partial f(x) = \begin{cases} \frac{1}{1+x} & \text{for } x > 0, \\ \operatorname{conv}(-1, 1) & \text{for } x = 0, \\ \frac{-1}{1-x} & \text{for } x < 0. \end{cases}$$

A right LLBF for $\log(1 + |x|)$ on $[a, b]$, $b > 0$, is represented by

$$\gamma(b - a, b)(x - b) + \log(1 + b),$$

where

$$\gamma(b - a, b) = \max\left\{\frac{1}{1 + b}, \frac{\log(1 + |a|) - \log(1 + |b|)}{\max(0, a) - b}\right\}.$$

Now we present a relationship between f' and γ. The following theorem is the nonsmooth version of this one given in [18].

Theorem 3.1.

Suppose $x^* \in \mathbb{R}$ is such that $f(x^*) = 0$ and $f'(x^*+) \neq 0$. Then there exists $r > 0$ such that $f'(x+) \ge \frac{1}{2}\gamma(\delta, y)$ for all $x, y \in B(x^*, r)$ and for all $\delta \in (0, 2r]$.

Proof. Assume $f'(x^*+) < 0$. Since f is semismooth, by statement (iii) of Theorem 2.2 we have that for a given $\varepsilon > 0$ there exists $r_1 > 0$ such that for any $h \in R$ satisfying $|h| < r_1$,

$$|f'(x^*; h) - f'(x; h)| \leq \varepsilon |h| .$$

Let $\varepsilon = -\frac{1}{8}\frac{f'(x^*+)}{|h|}$. For $h = 1$ we obtain from the above inequality

$$|f'(x+) - f'(x^*+)| \leq -\frac{1}{8} f'(x^*+), \text{ for all } x \in B(x^*, r_1). \tag{1}$$

From assumptions (A1) and (A3), Theorem 2.1 and Lemma 2.1 we have

$$\lim_{\delta\downarrow 0, x\to x^*} \gamma(\delta, x) \in \partial f(x).$$

So, by Lemma 2.1 there exist $r_2, \bar{\delta} > 0$ such that

$$|f'(x^*+) - \gamma(\delta, x)| \leq -\frac{1}{8} f'(x^*+), \text{ for all } x \in B(x^*, r_2) \text{ and for all } \delta \in \left(0, \bar{\delta}\right]. \tag{2}$$

The above inequality also implies that

$$\frac{f'(x^*+)}{\gamma(\delta, x)} \leq 2, \text{ for all } x \in B(x^*, r_2) \text{ and for all } \delta \in \left(0, \bar{\delta}\right]. \tag{3}$$

Let $r = \min\left\{r_1, r_2, \frac{1}{2}\bar{\delta}\right\}$. From (1) and (2) we obtain

$$\begin{aligned} |f'(x+) - \gamma(\delta, x)| \leq |f'(x+) - f'(x^*+)| + |f'(x^*+) - \gamma(\delta, x)| \leq -\frac{1}{4} f'(x^*+), \\ \text{for all } x, y \in B(x^*, r) \text{ and for all } \delta \in (0, 2r]. \end{aligned} \tag{4}$$

Finally, we have from (3) and (4)

$$\frac{f'(x^*+)}{\gamma(\delta, y)} \geq 1 - \frac{f'(x^*+)}{4\gamma(\delta, y)} \geq \frac{1}{2}, \text{ for all } x, y \in B(x^*, r) \text{ and for all } \delta \in (0, 2r].$$

The proof is similar, if $f'(x^*+) > 0$. □

4. Algorithm and its convergence

First, we present the problem of searching for the nearest root of f on the right-hand side of some point x_0. Assume f is a semismooth function and $f(x_0) > 0$. The problem is the following:

$$\left\{ \begin{array}{l} \max x \\ f(y) \geq 0 \text{ for all } y \in [x_0, x] \end{array} \right. \tag{P}$$

Similar as in Tseng [18], we suppose that $l_0 = \gamma_0(x - x_0) + f(x_0)$ is a right LLBF over $[x_0, x_0 + \Delta x_0]$. By the definition of a right LLBF, $f(y) \geq l_0(y)$ for all $y \in [x_0, x_0 + \Delta x_0]$. Hence we can formulate a trust region type method for solving (P):

$$\left\{ \begin{array}{l} \max\ x \\ l_0(x) = \gamma_0(x - x_0) + f(x_0) \geq 0 \\ x_0 \leq x \leq x_0 + \Delta x_0, \end{array} \right. \tag{$\bar{\text{P}}$}$$

where $\Delta x_0 > 0$ is the size of the trust region. $(\bar{P})$ is a linear programming problem (see e.g. Conn, Gould and Toint [5]). The LLBF's are used to eliminate these regions which do not contain zeros so that the nearest zero along the search direction can be reached.

The algorithm below allows us to find the nearest root on the right of the starting point x_0 for which $f(x_0) > 0$. Obviously, if $f(x_0) < 0$ we have to use a different LBF and suitable modifications in some steps of algorithms are required (interval in Step 1 and iteration in Step 3).

Algorithm UOA:

Assume that we have some constant $k > 1$ and a small number $\varepsilon > 0$. Let $x_0 \in R$ be a starting point such that $f(x_0) > 0$ and let $\Delta x_0 > 0$ be a starting step which is the size of the trust region. Let the search direction be rightward. Given a point x_i, the steps for obtaining x_{i+1} are: Step 0: $i = 0$;
Step 1: Construct a left LLBF l_i for f over $I_i = [x_i, x_i + \Delta x_i]$ of the form

$$l_i(x) = \gamma_i (x - x_i) + f(x_i);$$

If $|f(x_i)| = |l_i(x_i)| \leq \varepsilon$ then stop;
Step 2: If $\gamma_i \neq 0$ then compute an auxiliary step size

$$h_i = -\frac{f(x_i)}{\gamma_i};$$

Step 3: Compute the next iteration and a new trust region size:
if $0 < h_i < \Delta x_i$ then

$$\begin{aligned} x_{i+1} &= x_i + h_i \\ \Delta x_{i+1} &= h_i, \end{aligned}$$

otherwise

$$\begin{aligned} x_{i+1} &= x_i + \Delta x_i \\ \Delta x_{i+1} &= k\Delta x_i \end{aligned}$$

Step 4: $i = i + 1$. Return to Step 1.
Tseng [18] described in detail how we can apply the algorithm of locating the nearest root to solve an unconstrained univariate optimization problem. If we replace f by some g' then we obtain the nearest stationary point in the search direction. Moreover, in the nonsmooth case we can take some element from the generalized gradient of g instead of g'.

The convergence properties of a similar algorithm for continuously differentiable functions was proved by Tseng [18]. We may formulate analogous lemma for a semismooth function without loss of generality. Let $\{x_i\}$, $\{\Delta x_i\}$ and $\{\gamma_i\}$ be the sequences generated by the algorithm UOA. Then the following holds:

Lemma 4.1 (Tseng [18]).

If at some iteration i, $\Delta x_i > 0$ is finite and $f(x_i) > 0$, the following statements are true:

(i) $f(x) > 0$ *for all* $x \in [x_i, x_{i+1})$;

(ii) If $\Delta x_i \leq \frac{f(x_i)}{G}$ *then* $x_{i+1} \notin (x_i, x_i + \Delta x_i)$;

(iii) $x_{i+1} > x_i$ *and* $\Delta x_{i+1} > 0$.

Now, we establish a convergence theorem for the semismooth function f.

Theorem 4.1.
Let $\bar{x}$ be an accumulation point of $\{x_i\}$ and let x^* be the nearest root on the right-hand side of x_0. Then

(i) $x_i \uparrow \bar{x}$ *and* $\Delta x_i \downarrow 0$ *as* $i \to \infty$*;*

(ii) $\bar{x} = x^*$*;*

(iii) $\lim_{i\to\infty} \gamma_i \in \partial f(x^*)$*;*

(iv) If $f'(x^*+) \neq 0$*, there exists* $i_0 > 0$ *such that*

$$\Delta x_{i+1} = -\frac{f(x_i)}{\gamma_i}, \text{ for all } i > i_0,$$

i.e. the conditions $\gamma_i \neq 0$ *and* $0 < h_i < \Delta x_i$ *always hold after the* i_0*-th iteration.*

Proof. Proofs of the statements (i) and (ii) are similar those of [18], however we will write them out in detail since it might be useful on account of other notation and approaches in Steps 2 and 3.
(i) An accumulation point is also a limit point, since $\{x_i\}$ is a strictly increasing sequence by statement (iii) of Lemma 4.1. At the i-th iteration, if $0 < h_i < \Delta x_i$ then

$$\Delta x_{i+1} = |x_{i+1} - x_i|. \tag{5}$$

Otherwise $\Delta x_i = x_{i+1} - x_i$ and

$$\Delta x_{i+1} = k\Delta x_i = k|x_{i+1} - x_i|, \text{ for } k > 1. \tag{6}$$

So, from (5) and (6) we have

$$\Delta x_{i+1} \leq k|x_{i+1} - x_i|, \text{ for all } i = 1, 2, \ldots$$

Hence, $\Delta x_i \downarrow 0$ as $i \to \infty$, because $x_i \uparrow \bar{x}$.
(ii) The sequences $\{x_i\}$ and $\{\Delta x_i\}$ are bounded, since $x_i \uparrow \bar{x}$ and $\Delta x_i \downarrow 0$ as $i \to \infty$. From the definition of LBF, there exists a $G > 0$ such that $|\gamma_i| = |\gamma(\Delta x_i, x_i)| < G$ for all $i = 1, 2, \ldots$. Since the linear lower bounding functions of f always bound f below, it follows that $\bar{x} \leq x^*$. Suppose on the contrary that $\bar{x} \neq x^*$, so $f(\bar{x}) > 0$ and $f(x_i) > 0$ for all $i = 1, 2, \ldots$. Consider an arbitrary iteration, e.g. the k-th iteration. If we let

$$\varepsilon = \frac{1}{2}\min\left\{kG\Delta x_k, \frac{f(\bar{x})}{G}, \min_i f(x_i)\right\} > 0,$$

then we have two possible cases.
(a) $\Delta x_k > \varepsilon/G$: If $0 < h_k < \Delta x_k$, then $\gamma_k < 0$ and

$$\Delta x_k \geq \Delta x_{k+1} = -\frac{f(x_k)}{\gamma_k} > \frac{\varepsilon}{G}.$$

Otherwise

$$\Delta x_{k+1} = k\Delta x_k > \frac{\varepsilon}{G}.$$

Hence, the $k+1$-th iteration reduces to the same case again and we conclude that

$$\begin{aligned}\Delta x_i &> \frac{\varepsilon}{G} \text{ for all } i \geq k,\\ x_{i+1} &> x_i + \frac{\varepsilon}{G} \text{ for all } i \geq k.\end{aligned} \tag{7}$$

(b) $\Delta x_k \le \varepsilon/G$: From statement (ii) of Lemma 4.1 we conclude that the condition $0 < h_k < \Delta x_k$ does not hold and from the definition of ε we have $\Delta x_{k+1} = k\Delta x_k > \varepsilon/G$. It reduces to case (a) at the $k+1$-th iteration. From (7) we have also

$$x_{i+1} > x_i + \frac{\varepsilon}{G}, \text{ for all } i \ge k+1. \tag{8}$$

From (7) and (8), $x_i \to \infty$ as $i \to \infty$, which contradicts the fact that $x_i \to \bar{x}$. Thus $\bar{x} = x^*$.
(iii) Since $x_i \uparrow x^*$, $\Delta x_i \downarrow 0$ as $i \to \infty$. Also $\gamma_i = \gamma(\Delta x_i, x_i)$ and $\gamma(\cdot,\cdot)$ is an upper semicontinuous function (by assumption (A1)), and so by assumption (A3) and Lemma 2.1,

$$\lim_{i\to\infty}\left(\gamma_i - f'(x_i+)\right) = \lim_{i\to\infty}\left(\gamma(\Delta x_i, x_i) - f'(x_i+)\right) = 0.$$

Therefore $\lim_{i\to\infty}\gamma_i \in \partial f(x^*)$.
(iv) Suppose on the contrary that for any $i_0 > 0$, there always exists an $\hat{\imath} > i_0$ such that $\gamma_{\hat{\imath}} \ne 0$ and $0 < h_{\hat{\imath}} < \Delta x_{\hat{\imath}}$ does not hold. From Theorem 3.1 it can be shown that there exists $\hat{r} > 0$ and $\bar{\imath} > 0$ such that $[x_i, x_i + \Delta x_i] \subset B(x^*, \hat{r})$ for all $i > \bar{\imath}$ and

$$\frac{f'(x+)}{\gamma_i} = \frac{f'(x+)}{\gamma(\Delta x_i, x_i)} \ge \frac{1}{2} \text{ for all } x \in B(x^*, \hat{r}) \text{ and for all } i > \bar{\imath}. \tag{9}$$

Let $j > \bar{\imath}$. (9) implies that if $\gamma_j \ne 0$ and $0 < h_j < \Delta x_j$ holds, then by the local Lipschitz continuity of f we have

$$\Delta x_j \ge \Delta x_{j+1} = -\frac{f(x_j)}{\gamma_j} = \frac{f'(\xi+)\left(x^* - x_j\right)}{\gamma_j} \ge \frac{1}{2}\left(x^* - x_j\right) \tag{10}$$

where $\xi \in \left(x_j, x^*\right) \subset B(x^*, \hat{r})$. Now (10) implies that $x^* \in \left[x_{j+1}, x_{j+1} + \Delta x_{j+1}\right]$ and $0 < h_{j+1} < \Delta x_{j+1}$ holds. Recursively using (9) and (10), we have that $x^* \in [x_i, x_i + \Delta x_i]$ and $0 < h_i < \Delta x_i$ holds for all $i \ge j+1$. So the only possibility is that there exists $\bar{j} > \bar{\imath}$ such that $0 < h_i < \Delta x_i$ does not hold for all $i \ge \bar{j}$. But this will imply that $\Delta x_i \to \infty$ as $i \to \infty$, which contradicts the fact that $\Delta x_i \downarrow 0$. $\square$

As Tseng remarked in [18], his algorithm belongs to the modified Newton methods and has superlinear convergence. It follows from conditions (i)–(iii) of Theorem 4.1 and the characterization theorem for superlinear convergence proved by Dennis and Moré in [6]. The superlinear convergence in the nonsmooth case follows from the convergence of the nonsmooth version of Newton's method presented by Qi and Sun [14] in the form

$$x^{(k+1)} = x^{(k)} - V_k^{-1}F(x^{(k)}),$$

where $F: R^n \to R^n$ is semismooth, and the matrix $V_k \in \partial F(x^{(k)})$, $\partial F(x^{(k)})$ is the generalized Jacobian of F at $x^{(k)}$. Obviously, as we remarked earlier, the generalized gradient and the Jacobian coincide for $n = 1$. So in the algorithm UOA, γ_i is some approximation of generalized gradient of f.

Chen [3] introduced the notion of a smooth approximation function for locally Lipschitz functions. It is easy to verify that LBF satisfies all requested conditions (from the definition of l and statement (iii) of Theorem 4.1). Moreover, from assumption (A3) it follows that LBF satisfies the Jacobian consistence property. Since f is semismooth, then LBF also satisfies the directional derivative consistence property (see Lemma 2.3 in [3]).
Now we show that it is possible to reach faster convergence, when the function f is higher order semismooth.

Corollary 4.1.

Suppose that $x_i \uparrow x^*$ as $i \to \infty$, $f'(x^*+) \ne 0$ and f is p-order semismooth at x^*. If

$$O\left(|\gamma_i - f'(x_i+)|\right) = O\left(|x_i - x^*|\right)$$

then

$$|x_{i+1} - x^*| = O\left(|x_i - x^*|^{1+p}\right).$$

Proof. Statement (iv) of Theorem 4.1 implies that

$$f(x_i) + \gamma_i (x_{i+1} - x_i) = 0$$

holds for a suitable large $i > 0$. So

$$\gamma_i (x_{i+1} - x^*) = \gamma_i (x_i - x^*) - f(x_i).$$

Hence

$$\begin{aligned} |x_{i+1} - x^*| &= \left| x_i - x^* - \frac{f(x_i)}{\gamma_i} \right| \le \\ &\le \left| \frac{1}{\gamma_i} \left(\gamma_i (x_i - x^*) - f'(x^*; x_i - x^*) \right) \right| \\ &+ \left| \frac{1}{\gamma_i} \left(f(x_i) - f(x^*) - f'(x^*; x_i - x^*) \right) \right|. \end{aligned}$$

From the definition of p-order semismoothness of f at x^* we have

$$Vh - f'(x^*, h) = O\left(|h|^{1+p}\right) \text{ for any } V \in \partial f(x^* + h), h \to 0.$$

Moreover, for any $h \to 0$,

$$f(x^* + h) - f(x^*) - f'(x^*; h) = O\left(|h|^{1+p}\right).$$

Let $h = x_i - x^*$. By both of the above equalities and statement (iii) of Theorem 4.1, we obtain

$$|x_{i+1} - x^*| = O\left(|x_i - x^*|^{1+p}\right).$$

□

So, if the convergence in (A3) is at least as fast as the convergence of sequence $\{x_i\}$, then the sequence $\{x_i\}$ is convergent with rate $1 + p$. For example, quadratic convergence of the algorithm is implied by the strong semismoothness of f.

5. Numerical experiments and conclusions

In this section we show some practical performance of the proposed method. Four test minimization problems were solved: two with smooth functions and two with nonsmooth ones. The algorithm was numerically compared with some methods with and without derivatives proposed by Bazaraa, Sherali and Shetty in [1] and with the generalized Newton method (see [14]) used some approximation of generalized Jacobian (see e.g. [17]). We used the numbers of iterations (n) and the value of $f'(x_n)$ (for smooth cases) to test the efficiency. The algorithms were implemented in C++ and the results are carried out under double precision on a Pentium computer. The LBF for the considered test problems can be constructed based on the methods given in [18] and in Example 3.1 of this paper.
The practical termination criterion was used with the condition $|\Delta x_i| \le \varepsilon$, where $\varepsilon = 1.0E-10$. Moreover, we let $k = 2$ and $\Delta x_0 = 1.0$.

Example 5.1 ([18]).
Consider the unconstrained minimization problem with the function

$$f(x) = \sin x + \sin(2x/3).$$

This function has an infinite number of local minimum points in R. The initial interval for the search methods was $[-2, 6]$ and the initial point for the classical Newton method and algorithm UOA was $x_0 = 2$.

Table 1. The results from Example 5.1.

method	n	x_n	$f'(x_n)$	remarks
dichotomous search	36	5.3622462787	10^{-6}	
golden section	53	5.3622475449	10^{-9}	
regula falsi	10	1.8103569739	10^{-11}	local maximum
classical Newton	4	1.8103569739	10^{-11}	local maximum
UOA	6	−1.8103569739	10^{-11}	

Example 5.2 ([10]).

The function

$$f(x) = x^4 - 8.5x^3 - 31.0625x^2 - 7.5x + 45$$

has only one minimum point $x^* = 8.27846234384511$. Here the initial interval for the search methods was $[8, 9]$ and the initial point for the classical Newton method and algorithm UOA was $x_0 = 8.5$.

Table 2. The results from Example 5.2.

method	n	x_n	$f'(x_n)$
dichotomous search	33	8.2784638583	10^{-4}
golden section	48	8.2784623010	10^{-5}
regula falsi	6	8.2784623438	10^{-8}
classical Newton	5	8.2784623438	10^{-8}
UOA	8	8.2784623438	10^{-8}

Example 5.3.

Consider the nonsmooth function

$$g(x) = x\log(|x| + 0.1) - x + 0.1\,\mathrm{sgn}\,x\log(x + 0.1\,\mathrm{sgn}\,x).$$

which has minimum at $x^* = 0.9$. Note that $g'(x) = \log(|x| + 0.1)$ is semismooth function. We initialized the computations with the interval $[-1, 4]$ or the point $x_0 = 2$.

Table 3. The results from Example 5.3.

method	n	x_n	remarks
dichotomous search	36	0.899999904146962	
golden section	52	0.900000000019380	
regula falsi	×	×	looped
generalized Newton	7	0.900000000000021	
UOA	11	0.900000000000815	

Example 5.4.

Consider another nonsmooth function

$$g(x) = 0.25(|x - 1.5| - |\sin(10x)| + 3.0)$$

which has an infinite number of local minimum points in R. A global minimum point is $x^* = 1.71785921735823$. Here we initialized the computations with the interval $[0.2, 3]$ or the point $x_0 = 1.6$.

Table 4. The results from Example 5.4.

method	n	x_n	remarks
dichotomous search	36	2.66033697635430	local minimum
golden section	51	0.2	close to local minimum
regula falsi	41	11.30973355296672	local maximum
generalized Newton	6	2.66033701343517	local minimum
UOA	8	1.71785921737004	

Additionally, in Table 5 we establish initial intervals which allow us to find the global minimum point x^*. The intervals for search methods have bigger length, more iterations are needed and we can easyily miss a good approximation of the solution. In turn, the generalized Newton method has the shortest interval of convergence, therefore it is the fastest one.

Table 5. Initial intervals for Example 5.4.

method	initialization
dichotomous search	$[0.84, 2.43]$
golden section	$[1.07, 2.21]$
regula falsi	$[1.58, 1.88]$
generalized Newton	$[1.61, 1.84]$
UOA	$[1.58, 1.88]$

As stated above, we have developed a generalization of the univariate optimization algorithm for solving nonsmooth unconstrained problems. The performance of the method is evaluated in terms of the number of iterations required and compared with some popular methods using test problems. The important conclusion is that the algorithm of searching for the nearest extremum allows us to find the solution localized on the required side of an initial point. The other practical methods may lead to very different points (even maximum points or point outside given interval) if we do not start computations from a sufficient good initial point (or interval). Moreover, the relationship between the initial point and the obtained solution is unpredictable. The experimental results indicate that the new method has at least superlinear convergence, which confirms that algorithm UOA has better convergence properties than search methods without derivatives. Moreover, the results pointed out that Δx_i converges to 0 as $i \to \infty$, which was proved in statement (i) of Theorem 4.1.

References

[1] Bazaraa M.S., Sherali H.D., Shetty C.M., Nonlinear programming, theory and algorithms, John Wiley & Sons Inc., New York, 1993

[2] Bromberg M., Chang T.S., Global optimization using linear lower bounds: one dimensional case, In: Proceedings of the 29th Conference on Decision and Control, Honolulu, Hawaii, 1990

[3] Chen X., Superlinear convergence of smoothing quasi-Newton methods for nonsmooth equations, J. Comput. Appl. Math., 1997, 80, 105–126

[4] Clarke F.H., Optimization and nonsmooth analysis, John Wiley & Sons Inc., New York, 1983

[5] Conn A.R., Gould N.I.M., Toint P.L., Trust-region methods, SIAM, Philadelphia, 2000

[6] Dennis Jr. J.E., Moré J.J., A characterization of superlinear convergence and its application to quasi-Newton methods, Math. Comp., 1974, 28, 549–560

[7] Famularo D., Sergeyev Ya.D., Pugliese P., Test problems for Lipschitz univariate global optimization with multi-extremal constraints, In: Dzemyda G., Saltenis V., Žilinskas A. (Eds.), Stochastic and global optimization, Kluwer Academic Publishers, Dordrecht, 2002
[8] Hansen P., Jaumard B., Lu S.H., Global optimization of univariate Lipschitz functions I: survey and properties, Math. Program., 1992, 55, 251–273
[9] Harker P.T., Xiao B., Newton's method for the nonlinear complementarity problem: a B-differentiable equation approach, Math. Program., 1990, 48, 339–357
[10] Kahya E., A class of exponential quadratically convergent iterative formulae for unconstrained optimization., Appl. Math. Comput., 2007, 186, 1010–1017
[11] Mifflin R., Semismooth and semiconvex functions in constrained optimization, SIAM J. Control Optim., 1977, 15, 959–972
[12] Pang J.S., Newton's method for B-differentiable equations, Math. Oper. Res., 1990, 15, 311–341
[13] Potra F.A., Qi L., Sun D., Secant methods for semismooth equations, Numer. Math., 1998, 80, 305–324
[14] Qi L., Sun J., A nonsmooth version of Newton's method, Math. Program., 1993, 58, 353–367
[15] Sergeyev Ya.D., Daponte P., Grimaldi D., Molinaro A., Two methods for solving optimization problems arising in electronic measurements and electrical engineering, SIAM J. Optim., 1999, 10, 1–21
[16] Shapiro A., On concepts of directional differentiability, J. Optim. Theory Appl., 1990, 66, 477–487
[17] Śmietański M.J., A new versions of approximate Newton method for solving nonsmooth equations, Ph.D. thesis, University of Łódź, Poland, 1999 (in Polish)
[18] Tseng C.L., A Newton-type univariate optimization algorithm for locating the nearest extremum, Eur. J. Oper. Res., 1998, 105, 236–246