Yugoslav Journal of Operations Research
23 (2013), Number 1, 59-71
DOI: 10.2298/YJOR110501024D

ON AN ALGORITHM IN NONDIFFERENTIAL CONVEX OPTIMIZATION

Nada I. ĐURANOVIĆ-MILIČIĆ

Department of Mathematics,
Faculty of Technology and Metallurgy
University of Belgrade, Belgrade, Serbia
nmilicic@tmf.bg.ac.rs

Milanka GARDAŠEVIĆ-FILIPOVIĆ

Vocational College of Technology, Arandjelovac, Serbia
milankafilipovic@yahoo.com

Received: May 2011 / Accepted: October 2012

Abstract: In this paper an algorithm for minimization of a nondifferentiable function is presented. The algorithm uses the Moreau-Yosida regularization of the objective function and its second order Dini upper directional derivative. The purpose of the paper is to establish general hypotheses for this algorithm, under which convergence occurs to optimal points. A convergence proof is given, as well as an estimate of the rate of the convergence.

Keywords: Moreau-Yosida regularization, non-smooth convex optimization, directional derivative, second order Dini upper directional derivative, uniformly convex functions.

MSC: 90C30; 90C25; 65K05.

1. INTRODUCTION

The following minimization problem is considered:

$$\min_{x \in R^n} f(x) \tag{1.1}$$

where $f: R^n \to R \cup \{+\infty\}$ is a convex and not necessarily differentiable function with a nonempty set X^* of minima.

Many approaches have been presented for non-smooth programs, but they are often restricted to the convex unconstrained case. The reason for the restriction is the fact that a constrained problem can be easily transformed to an unconstrained problem using a distance function. In general, the various approaches are based on combinations of the following methods: subgradient methods, bundle techniques and the Moreau-Yosida regularization.

For a convex function f it is very important that its Moreau-Yosida regularization is a new function with the same set of minima as f and is differentiable with Lipschitz continuous gradient, even when f is not differentiable. In [13], [14] and [23] the second order properties of the Moreau-Yosida regularization of a given function f are considered.

Having in mind that the Moreau-Yosida regularization of a proper closed convex function is an LC^1 function, we present an optimization algorithm (using the second order Dini upper directional derivative (described in [1])) based on the results from [3], which is the main idea of this paper.

We shall present an iterative algorithm for finding an optimal solution of problem (1.1) by generating the sequence of points $\{x_k\}$ of the following form:

$$x_{k+1} = x_k + \alpha_k s_k + \alpha_k^2 d_k \quad k = 0,1,..., s_k \neq 0, d_k \neq 0 \tag{1.2}$$

where the step-size α_k and the directional vectors s_k and d_k are defined by the particular algorithms.

The paper is organized as follows: in the second section, some basic theoretical preliminaries are given; in the third section, the Moreau-Yosida regularization and its properties are described; in the fourth section, the definition of the second order Dini upper directional derivative and its basic properties are given; in the fifth section, the semi-smooth functions and conditions for their minimization are described. Finally, in the sixth section, a model algorithm is suggested, its convergence is proved and an estimate rate of its convergence is given, too.

2. THEORETICAL PRELIMINARIES

Throughout the paper we will use the following notation. A vector s refers to a column vector, and ∇ denotes the gradient operator $\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, ..., \frac{\partial}{\partial x_n}\right)^T$. The Euclidean product is denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ is the associated norm; $B(x, \rho)$ is the ball centred at x with radius ρ. For a given symmetric positive definite linear operator M, we set $\langle \cdot, \cdot \rangle_M := \langle M \cdot, \cdot \rangle$; hence, it is shortly denoted by $\|x\|_M^2 := \langle x, x \rangle_M$. The smallest and the largest eigenvalue of M we denote by λ and Λ, respectively.

The *domain* of a given function $f: R^n \to R \cup \{+\infty\}$ is the set $dom(f) = \{x \in R^n \,|\, f(x) < +\infty\}$. We say that f is proper if its domain is nonempty.

The point $x^* = arg\min_{x \in R^n} f(x)$ refers to the minimum point of a given function $f: R^n \to R \cup \{+\infty\}$.

The *epigraph* of a given function $f: R^n \to R \cup \{+\infty\}$ is the set $epi\, f = \{(\alpha, x) \in R \times R^n \,|\, \alpha \geq f(x)\}$. The concept of the epigraph enables us to define convexity and closure of a function in a new way. We say that f is convex if its epigraph is a convex set, and f is closed if its epigraph is a closed set.

In this section, we will give definitions and statements that are necessary for this work.

Definition 1. *A vector $g \in R^n$ is said to be a subgradient of a given proper convex function $f: R^n \to R \cup \{+\infty\}$ at a point $x \in R^n$ if the next inequality*

$$f(z) \geq f(x) + g^T \cdot (z - x) \tag{2.1}$$

holds for all $z \in R^n$. The set of all subgradients of $f(x)$ at the point x, called the *subdifferential* at the point x, is denoted by $\partial f(x)$. The subdifferential $\partial f(x)$ is a nonempty set if and only if $x \in dom(f)$.

For a convex function f it follows that $f(x) = \max_{z \in R^n} \{f(z) + g^T (x - z)\}$ holds, where $g \in \partial f(z)$ (see [7]).

The concept of the subgradient is a simple generalization of the gradient for non-differentiable convex functions.

Lemma 1. *Let $f: S \to R \cup \{+\infty\}$ be a convex function defined on a convex set $S \subseteq R^n$, and $x' \in \operatorname{int} S$. Let $\{x_k\}$ be a sequence such that $x_k \to x'$, where $x_{k+1} = x_k + \varepsilon_k s_k + \varepsilon_k^2 d_k$, $k = 0,1,..., s_k \neq 0, d_k \neq 0$, $\varepsilon_k > 0$, $\varepsilon_k \to 0$ and $s_k \to s$, $d_k \to d$ and $g_k \in \partial f(x_k)$. Then all accumulation points of the sequence $\{g_k\}$ lie in the set $\partial f(x')$.*

Proof. *Since $g_k \in \partial f(x_k)$, then the inequality $f(y) \geq f(x_k) + g_k^T \cdot (y - x_k)$ holds for any $y \in S$. Hence, taking any subsequence for which $g_k \to g'$, it follows that $f(y) \geq f(x') + g'^T \cdot (y - x')$, which means that $g' \in \partial f(x')$.* ■

Definition 2. *The directional derivative of a real function f defined on R^n at the point $x' \in R^n$ in the direction $s \in R^n$, denoted by $f'(x', s)$, is*

$$f'(x', s) = \lim_{t \downarrow 0} \frac{f(x' + t \cdot s) - f(x')}{t} \tag{2.2}$$

when this limit exists.

Hence, it follows that if the function f is convex and $x' \in dom\, f$, then

$$f(x'+t\cdot s)=f(x')+t\cdot f'(x',s)+o(t) \tag{2.3}$$

holds, which can be considered as one linearization of the function f (see in [8]).

Lemma 2. *Let $f:S\to R\cup\{+\infty\}$ be a convex function defined on a convex set $S\subseteq R^n$, and $x'\in \operatorname{int} S$. If the sequence $x_k\to x'$, where $x_k=x'+\varepsilon_k s_k$, $\varepsilon_k>0$, $\varepsilon_k\to 0$ and $s_k\to s$ then the next formula:*

$$f'(x',s)=\lim_{k\to\infty}\frac{f(x_k)-f(x')}{\varepsilon_k}=\max_{g\in\partial f(x')} s^T g \tag{2.4}$$

holds.

Proof. See in [9] or [17].

Lemma 3. *Let $f:S\to R\cup\{+\infty\}$ be a convex function defined on a convex set $S\subseteq R^n$. Then $\partial f(x)$ is bounded for $\forall x\in B\subset int\,S$, where B is a compact.*

Proof. See in [10] or [12].

Proposition 1 *Let $f:R^n\to R\cup\{+\infty\}$ be a proper convex function. The condition:*

$$0\in\partial f(x) \tag{2.5}$$

is a first order necessary and sufficient condition for a global minimizer at $x\in R^n$. This can be stated alternatively as:

$$\forall s\in R^n,\ \|s\|=1\ \max_{g\in\partial f(x)} s^T g\ge 0\,. \tag{2.6}$$

Proof. See [16].

Lemma 4. *If a proper convex function $f:R^n\to R\cup\{+\infty\}$ is a differentiable function at a point $x\in dom(f)$, then:*

$$\partial f(x)=\{\nabla f(x)\}\,. \tag{2.7}$$

Proof. The statement follows directly from Definition 2.

Lemma 5. *Let $f_i:R^n\to R\cup\{+\infty\}$ for $i=\{1,2,...,n\}, n\in N$ be convex functions, and $f(x)=\max_{i\in\{1,2,...,n\}} f_i(x)$. Then the function f is a convex function, and its subgradient g at the point $x\in R^n$ is given as follows:*

$$g=\sum_{i\in\hat{I}}\lambda_i g_i \tag{2.8}$$

where $\sum_{i \in \hat{I}} \lambda_i = 1$ *and* $\lambda_i \geq 0$, $g_i \in \partial f_i(x)$ *for* $i \in \hat{I}$, *and* $\hat{I}$ *is the set*

$\hat{I} = \{i \in I | f(x) = f_i(x)\}$.

Proof. See in [7].

Definition 3. *The real function* f *defined on* R^n *is* LC^1 *function on the open set* $D \subseteq R^n$ *if it is continuously differentiable and its gradient* ∇f *is locally Lipschitz, i.e.*

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| \text{ for } x, y \in D \tag{2.9}$$

for some $L > 0$.

3. THE MOREAU-YOSIDA REGULARIZATION

Definition 4. *Let* $f : R^n \to R \cup \{+\infty\}$ *be a proper closed convex function. The Moreau-Yosida regularization of a given function* f, *associated to the metric defined by M, denoted by F, is defined as follows:*

$$F(x) := \min_{y \in R^n} \left\{ f(y) + \frac{1}{2}\|y - x\|_M^2 \right\} \tag{3.1}$$

The above function is an *infimal convolution*. In [18] it is proved that the infimal convolution of a convex function is also a convex function. Hence, the function defined by (3.1) is a convex function and has the same set of minima as the function f (see in [8]), so the motivation of the study of Moreau-Yosida regularization is due to the fact that $\min_{x \in R^n} f(x)$ is equal to $\min_{x \in R^n} F(x)$.

Definition 5. The minimum point $p(x)$ of the function (3.1):

$$p(x) := \operatorname{argmin}_{y \in R^n} \left\{ f(y) + \frac{1}{2}\|y - x\|_M^2 \right\} \tag{3.2}$$

is called the *proximal point* of x.

Proposition 2. *The function* F *defined by (3.1) is always differentiable.*

Proof. See in [8].

The first order regularity of F is well known (see in [8] and [13]): without any further assumptions, F has a Lipschitzian gradient on the whole space R^n. More precisely, for all $x_1, x_2 \in R^n$ the next formula:

$$\|\nabla F(x_1) - \nabla F(x_2)\|^2 \leq \Lambda \langle \nabla F(x_1) - \nabla F(x_2), x_1 - x_2 \rangle \tag{3.3}$$

holds (see in [13]), where $\nabla F(x)$ has the following form:

$$G := \nabla F(x) = M(x - p(x)) \in \partial f(p(x)) \tag{3.4}$$

and $p(x)$ is the unique minimum in (3.1). So, according to the above consideration and Definition 3, we conclude that F is an LC^1 function (see in [14]).

Note that the function f has nonempty subdifferential at any point p of the form $p(x)$. Since $p(x)$ is the minimum point of the function (3.1), then (see in [8] and [13]):

$$p(x) = x - M^{-1}g \text{ where } g \in \partial f(p(x)) \tag{3.5}$$

In [13] it is also proved that for all $x_1, x_2 \in R^n$ the next formula:

$$\|p(x_1) - p(x_2)\|_M^2 \le \langle M(x_1 - x_2), p(x_1) - p(x_2) \rangle \tag{3.6}$$

is valid, namely the mapping $x \to p(x)$, where $p(x)$ is defined by (3.2), is Lipschitzian with constant $\frac{\Lambda}{\lambda}$ (see Proposition 2.3. in [13]).

Lemma 6: The following statements are equivalent:

(i) x minimizes f;

(ii) $p(x) = x$

(iii) $\nabla F(x) = 0$

(iv) x minimizes F;

(v) $f(p(x)) = f(x)$;

(vi) $F(x) = f(x)$

Proof. See in [8] or [23].

4. DINI SECOND UPPER DIRECTIONAL DERIVATIVE

We shall give some preliminaries that will be used in the remainder of the paper.

Definition 6. [22] The *second order Dini upper directional derivative* of the function $f \in LC^1$ at the point $x \in R^n$ in the direction $d \in R^n$ is defined to be $f_D''(x,d) = \limsup_{\alpha \downarrow 0} \frac{[\nabla f(x + \alpha d) - \nabla f(x)]^T \cdot d}{\alpha}$. If ∇f is directionally differentiable at x_k, we have that $f''_D(x_k,d) = f''(x_k,d) = \lim_{\alpha\downarrow 0} \frac{[\nabla f(x+\alpha d) - \nabla f(x)]^T \cdot d}{\alpha}$ *holds* for all $d \in R^n$.

Since the Moreau-Yosida regularization of a proper closed convex function f is an LC^1 function, we can consider its second order Dini upper directional derivative at the point $x \in R^n$ in the direction $d \in R^n$. Using (3.4) we can state that:

$$F''_D(x,d) = \limsup_{\alpha\downarrow 0} \frac{g_1 - g_2}{\alpha} d,$$

where $F(x)$ is defined by (3.1) and $g_1 \in \partial f(p(x+\alpha d)), g_2 \in \partial f(p(x))$.

***Lemma 7**: Let* $f : R^n \to R$ be a closed convex proper function and F is its Moreau – Yosida regularization. Then the next statements are valid.

(i) $F''_D(x_k, kd) = k^2 F''_D(x_k,d)$

(ii) $F''_D(x_k, d_1+d_2) \le 2(F''_D(x_k,d_1) + F''_D(x_k,d_2))$

(iii) $|F''_D(x_k,d)| \le K \cdot \|d\|^2$, where K is some constant.

Proof. See in [1] and [22].

***Lemma 8.** Let $f : R^n \to R$ be a closed convex proper function and let F be its Moreau – Yosida regularization. Then the next statements are valid.*

(i) $F''_D(x,d)$ is upper semicontinous with respect to (x,d), i.e. $\limsup_{i\to\infty} F''_D(x_i,d_i) \le F''_D(x,d)$ when $(x_i,d_i) \to (x,d)$

(ii) $F''_D(x,d) = \max\{d^T V d \mid V \in \partial^2 F(x)\}$

Proof. See in [1] and [22].

5. SEMI-SMOOTH FUNCTIONS AND OPTIMALITY CONDITIONS

***Definition 7:** A function $\nabla F : R^n \to R^n$ is said to be semi-smooth at the point $x \in R^n$ if ∇F is locally Lipschitzian at $x \in R^n$ and the limit $\lim_{\substack{h\to d\\ \lambda\downarrow 0}} \{Vh\}$, $V \in \partial^2 F(x+\lambda h)$ exists for any $d \in R^n$.*

Note that for a closed convex proper function, the gradient of its Moreau-Yosida regularization is a semi-smooth function.

***Lemma 9.** [22]: If $\nabla F : R^n \to R^n$ is semi-smooth at the point $x \in R^n$, then ∇F is directionally differentiable at $x \in R^n$ and for any $V \in \partial^2 F(x+h), h \to 0$ we have that: $Vh - (\nabla F)'(x,h) = o(\|h\|)$. Similarly we have that $h^T V h - F''(x,h) = o(\|h\|^2)$.*

Lemma 10: *Let $f : R^n \to R$ be a proper closed convex function and let F be its Moreau-Yosida regularization. So, if $x \in R^n$ is a solution of the problem (1.1), then $F'(x,d) = 0$ and $F''_D(x,d) \geq 0$ for all $d \in R^n$.*
Proof. See in [6].
Lemma 11. *Let $f : R^n \to R$ be a proper closed convex function, F its Moreau-Yosida regularization, and x a point from R^n. If $F'(x,d) = 0$ and $F''_D(x,d) > 0$ for all $d \in R^n$, then $x \in R^n$ is a strict local minimizer of the problem (1.1).*
Proof. See in [6].

6. A MODEL ALGORITHM

In this section an algorithm for solving the problem (1.1) is introduced. We suppose that at each $x \in R^n$ it is possible to compute $f(x), F(x)$, $\nabla F(x)$ and $F''_D(x,d)$ for a given $d \in R^n$.

At the *k*-th iteration the direction vector $s_k \neq 0$ in (1.2) is any vector satisfying the nonascent property, i.e. $\nabla F^T(x_k)s_k \leq 0$ holds, and the direction vector d_k presents a solution of the problem

$$\min_{d \in R^n} \Phi_k(d),\ \Phi_k(d) = \nabla F(x_k)^T d + \frac{1}{2} F''_D(x_k, d) \tag{6.1}$$

where $F''_D(x_k, d)$ stands for the second order Dini upper directional derivative at x_k in the direction d. Note that if Λ is a Lipschitzian constant for F, it is also a Lipschitzian constant for ∇F. The function $\Phi_k(d)$ is called an iteration function. It is easy to see that $\Phi_k(0) = 0$, and $\Phi_k(d)$ is Lipschitzian on R^n.

For the given q, where $0 < q < 1$, the step-size $\alpha_k > 0$ is a number satisfying $\alpha_k = q^{i(k)}$,where $i(k)$ is the smallest integer from $i = 0,1,...$ such that the following two inequalities are satisfied:

$$F(x_k + \alpha_k s_k + \alpha_k^2 d_k) - F(x_k) \leq \sigma \left[\alpha_k \nabla F^T(x_k) s_k - \frac{1}{4} \alpha_k^4 F''_D(x_k, d_k) \right] \tag{6.2}$$

and

$$F(x_k + 2\alpha_k s_k + (2\alpha_k)^2 d_k) - F(x_k) > \sigma \begin{bmatrix} 2\alpha_k \nabla F^T(x_k) s_k \\ -\frac{1}{2}(2\alpha_k)^4 F''_D(x_k, d_k) \end{bmatrix} \tag{6.3}$$

where $0 < \sigma < 1$ is a reassigned constant, and $x_0 \in R^n$ is a given point.

We make the following assumptions.

A1. We suppose that there exist constants $c_2 \geq c_1 > 0$ such that

$$c_1 \|d\|^2 \leq F_D''(x_k, d) \leq c_2 \|d\|^2 \text{ for every } d \in R^n \tag{6.4}$$

A2. $\|d_k\| = 1$ and $\|s_k\| = 1$, $k = 0,1,...$

A3. There exists a value $\gamma > 0$ such that

$$-\nabla F^T(x_k)s_k \geq \gamma \|s_k\|^2, k = 0,1,2,... \tag{6.5}$$

A4. $\nabla F^T(x_k)s_k \to 0 \Rightarrow \|\nabla F(x_k)\| \to 0, k \to \infty$

It follows from Lemma 3.1 in [22] that under the assumption A1, the optimal solution of the problem (6.1) exists.

In order to have a finite value $i(k)$ satisfying (6.2), it is sufficient that s_k and d_k have descent properties, i.e. $\nabla F^T(x_k)s_k < 0$ and $\nabla F^T(x_k)d_k < 0$ whenever $\nabla F(x_k) \neq 0$. The first relation follows from (6.5). Relating to the second condition, if $d_k \neq 0$ is a solution of (6.1), it follows that $\Phi_k(d_k) \leq 0 = \Phi_k(0)$. Consequently, using (6.4), we get

$$\nabla F^T(x_k)d_k \leq -\frac{1}{2}F_D''(x_k, d) \leq -\frac{1}{2}c_1 \|d\|^2 < 0 \tag{6.6}$$

i.e. d_k is a descent direction at x_k.

Now, suppose that for some $\alpha_k = q^{i(k)}$ the inequality (6.2) holds, but the inequality (6.3) does not hold. In other words, we have the following:

$$F(x_k + \alpha_k s_k + \alpha_k^2 d_k) - F(x_k) \leq \sigma\left[\alpha_k \nabla F^T(x_k)s_k - \frac{1}{4}\alpha_k^4 F_D''(x_k, d_k)\right]$$

and

$$F(x_k + 2\alpha_k s_k + (2\alpha_k)^2 d_k) - F(x_k) \leq \sigma\left[2\alpha_k \nabla F^T(x_k)s_k - \frac{1}{2}(2\alpha_k)^4 F_D''(x_k, d_k)\right]$$

that is, if there is no j for which $2^j \alpha_k$ satisfies (6.3), we shall obtain

$$F(x_k + 2^j \alpha_k s_k + (2^j \alpha_k)^2 d_k) - F(x_k) \leq \sigma\begin{bmatrix} 2^j \alpha_k \nabla F^T(x_k)s_k \\ -\frac{1}{2}(2^j \alpha_k)^4 F_D''(x_k, d_k) \end{bmatrix}, j = 0,1,2,...$$

The right side of the above inequality tends to $-\infty$ as $j \to \infty$, that is, $F(x_k + 2^j \alpha_k s_k + (2^j \alpha_k)^2 d_k) - F(x_k) \to -\infty$ as $j \to \infty$, which is the contradiction since F is, because of the assumptions, bounded below on the compact set $L(x_0)$. Therefore, for some j both inequalities (6.2) and (6.3) will be satisfied. Consequently, our algorithm is well-defined.

The inequality (6.3) guarantees a suitable reduction in F, which means that $F(x)$ is decreased by at least a multiple of the modulus of the directional derivative $\alpha_k \nabla F^T(x_k)s_k$ and $-\frac{1}{2}\alpha_k^4 F_D''(x_k, d_k)$. As $\alpha_k = 0$ satisfies this inequality, it is necessary to introduce another condition that prevents too small α_k to be chosen. This is the purpose of the inequality (6.2).

Proposition 3. *If the Moreau-Yosida regularization $F(\cdot)$ of the proper closed convex function $f(\cdot)$ satisfies the condition (6.4), then:*

(i) the function $F(\cdot)$ is uniformly convex, and hence, strictly convex;

(ii) the level set $L(x_0) = \{x \in R^n : F(x) \leq F(x_0)\}$ is a compact convex set, and

(iii) there exists a unique point x^* such that $F(x^*) = \min_{x \in L(x_0)} F(x)$.

Proof. See in [6].

Convergence theorem . *If the Moreau-Yosida regularization $F(\cdot)$ of the proper closed convex function $f(\cdot)$ satisfies the assumptions A1, A2 ,A3 and A4, then for any initial point $x_0 \in R^n, x_k \to \bar{x}$, as $k \to \infty$, where $\bar{x}$ is a unique minimal point of the function f.*

Proof. From (6.2), (6.5) and (6.6) it follows that

$$F(x_{k+1}) - F(x_k) \leq \sigma\left[q^{i(k)} \nabla F^T(x_k)s_k - \frac{1}{4} q^{4i(k)} F_D''(x_k, d_k) \right] < 0 \tag{6.7}$$

Hence $\{F(x_k)\}$ is a decreasing sequence, and consequently $\{x_k\} \subset L(x_0)$. Since $L(x_0)$ is by Proposition 3 a compact convex set, it follows that the sequence $\{x_k\}$ is bounded. Therefore there exist accumulation points of $\{x_k\}$. Since ∇F is by assumption continuous, then, if $\nabla F(x_k) \to 0$ as $k \to \infty$, it follows that every accumulation point $\bar{x}$ of the sequence $\{x_k\}$ satisfies $\nabla F(\bar{x}) = 0$. Since F is by Proposition 3 strictly convex, it follows that there exists a unique point $\bar{x} \in L(x_0)$ such that $\nabla F(\bar{x}) = 0$. Hence, $\{x_k\}$ has a unique limit point $\bar{x}$ – and it is a global minimizer. Therefore, we have to prove that $\nabla F(x_k) \to 0$ as $k \to \infty$.

We first show that the set of indices $\{i(k)\}$ is uniformly bounded above by a number I, i.e. $i(k) \leq I < \infty$. Suppose the contrary. By the definition of $i(k)$ from (6.2) it follows that

$$F(x_k + q^{i(k)-1}s_k + q^{2i(k)-2}d_k) - F(x_k) > \sigma\left[q^{i(k)-1} \nabla F^T(x_k)s_k - \frac{1}{4} q^{4i(k)-4} F_D''(x_k, d_k) \right]$$

By the definition of the Dini derivative and by (6.4), we have

$$
\begin{aligned}
&F(x_k + q^{i(k)-1} s_k + q^{2i(k)-2} d_k) - F(x_k) \\
&\quad = q^{i(k)-1} \nabla F(x_k)^T s_k + q^{2i(k)-2} d_k \nabla F(x_k)^T d_k + \frac{1}{2} F_D''(x_k, q^{i(k)-1} s_k + q^{2i(k)-2} d_k) + o(q^{2i(k)-2}) \\
&\quad \geq q^{i(k)-1} \nabla F(x_k)^T s_k + q^{2i(k)-2} d_k \nabla F(x_k)^T d_k + \frac{1}{2} c_1 \left\| q^{i(k)-1} s_k + q^{2i(k)-2} d_k \right\|^2 + o(q^{2i(k)-2}) \\
&\quad = q^{i(k)-1} \nabla F(x_k)^T s_k + q^{2i(k)-2} d_k \nabla F(x_k)^T d_k + \frac{1}{2} c_1 q^{2i(k)-2} \left\| s_k + q^{i(k)-1} d_k \right\|^2 + o(q^{2i(k)-2}) \\
&\quad > \sigma \left[q^{i(k)-1} \nabla F(x_k)^T s_k - \frac{1}{2} q^{4i(k)-4} F_D''(x_k, d_k) \right]
\end{aligned}
$$

Accumulating all terms of order higher than $O(q^{2i(k)-2})$ into the term $o(q^{2i(k)-2})$ (because $\|s_k\| = \|d_k\| = 1$) and using the fact that $\nabla F(x_k)^T d_k \leq 0$ (by (6.6)) yields $\frac{1}{2} c_1 q^{2i(k)-2} \|s_k\|^2 + o(q^{2i(k)-2}) > (\sigma - 1) q^{i(k)-1} \nabla F(x_k)^T s_k \geq 0$ since $0 < \sigma < 1$ and $\nabla F(x_k)^T s_k \leq 0$. Dividing by $q^{i(k)-1}$ yields $\frac{1}{2} c_1 q^{i(k)-1} \|s_k\|^2 + o(q^{i(k)-1}) > (\sigma - 1) \nabla F(x_k)^T s_k \geq 0$. Dividing by $\frac{1}{2q} c_1 \|s_k\|^2 = \frac{1}{2q} c_1$ yields $q^{i(k)-1} > \frac{2(\sigma-1)q}{c_1} \nabla F(x_k)^T s_k + \frac{o(q^{i(k)-1})q}{c_1}$. Taking the limit as $k \to \infty$, and having in view (6.5), we get $q^{i(k)} > \frac{2(1-\sigma)q}{c_1} \gamma + o(q^{i(k)-1}) > 0.$

Hence, $\alpha_k = q^{i(k)}$ is bounded away from zero, which contradicts the assumption that the sequence $\{i(k)\}$ is unbounded. Hence $i(k) \leq I < \infty$ for all k. From (6.2), it follows that

$$
\begin{aligned}
&F(x_{k+1}) - F(x_k) \leq \sigma \left[q^{i(k)} \nabla F(x_k)^T s_k - \frac{1}{2} q^{4i(k)} \nabla F_D''(x_k; d_k) \right] \\
&\leq \sigma \left[q^I \nabla F(x_k)^T s_k - \frac{1}{2} q^{4I} \nabla F_D''(x_k; d_k) \right]
\end{aligned} \tag{6.8}
$$

Hence, multiplying this inequality by (–1), we get $F(x_k) - F(x_{k+1}) \geq \sigma \left[q^I \nabla F(x_k)^T s_k - \frac{1}{2} q^{4I} \nabla F_D''(x_k; d_k) \right].$

Since $\{F(x_k)\}$ is bounded below (on the compact set $L(x_0)$) and monotone (by (6.8)), it follows that $F(x_{k+1}) - F(x_k) \to 0$ as $k \to \infty$. Hence from (6.8) it follows that $\nabla F(x_k)^T s_k \to 0$ and $\nabla F_D''(x_k; d_k) \to 0$ as $k \to \infty$. Finally, from A4, it follows that

$\|\nabla F(x_k)\| \to 0, k \to \infty$, and from Lemma 6, it follows that $\bar{x}$ is a unique minimal point of the function f.

<u>Convergence rate theorem</u> *Under the assumptions of the previous theorem we have that the following estimate holds for the sequence* $\{x_k\}$ *generated by the algorithm.*

$$F(x_n) - F(\bar{x}) \le \mu_0 \left[1 + \frac{\mu_0}{\eta^2} \sum_{k=0}^{n-1} \frac{F(x_k) - F(x_{k+1})}{\|\nabla F(x_k)\|^2} \right]^{-1}, \textit{for } n = 1,2,3,...$$

where $\mu_0 = F(x_0) - F(\bar{x})$, *and* $\operatorname{diam} L(x_0) = \eta < \infty$ *(since by Proposition 3 it follows that* $L(x_0)$ *is bounded).*

<u>*Proof*</u>. The proof directly follows from Theorem 9.2, page 167, in [11].

CONCLUSION

The algorithm presented in this paper is based on the algorithms from [2], [3], [22] and [6]. The convergence is proved under mild conditions. This method uses minimization along a plane, defined by vectors s_k and d_k, to generate a new iterative point at each iteration. Relating to the algorithm in [3], the presented algorithm is defined and it converges for noonsmooth convex function.

ACKNOWLEDGMENT

This research was supported by Science Fund of Serbia, grant number 44007.

REFERENCES

[1] Clarke F. H.: *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.
[2] Djuranovic Milicic Nada: "A generalization of the Cea-Goldstein step-size algorithm", *European Journal of Operational Research* 63 (1992) 1-9.
[3] Djuranovic Milicic Nada: "On an algorithm for C1,1 optimization", *Filomat*, 21 (2007) 17-24.
[4] Djuranovic-Milicic, N., "An algorithm for LC1 optimization", *YUJOR* ,15 (2) (2005) 301-300.
[5] Djuranovic-Milicic, N., " A multi-step curve search algorithm in nonlinear optimization", *YUJOR* ,18 (1) (2008) 47-52.
[6] Djuranovic-Milicic, N., and Gardasevic-Filipovic M.: "An algorithm for minimization of a nondifferentiable convex function", *The 2009 International Conference of Applied and Engineering Mathematics, World Congress on Engineering 2009, Proceedings,* 2004, 1241-1246.
[7] Демьянов В. Ф., Васильев: *Недифференцируемая оптимизация*, Москва „Наука", 1981.

[8] Fuduli A.: “Metodi numerici per la minimizzazione di funzioni convesse nondifferenziabili”, PhD Thesis, Departimento di Electronica Informatica e Sistemistica, Univerzita della Calabria, 1998.

[9] Fletcher R.: *Practical Methods of Optimization, Volume 2, Constrained Optimization*, Department of Mathematics, University of Dundee, Scotland, U.K,Wiley-Interscience, Publication, 1981.

[10] Иоффе А. Д., Тихомиров В. М.: *Теория экстремальныц задач*, Москва „Наука“, 1974.

[11] Karmanov V.G: “Mathematical Programming”. Nauka, Moscow, 1975 (in Russian).

[12] Кусарев А. Г., Кутателадзе С. С.:*Субдифференциальное исчисление*, Новосибирск, Издатеьство „Наука“, 1987.

[13] Lemaréchal C. , and Sagastizabal C.: “Practical aspects of the Moreau-Yosida regularization I: theoretical preliminaries”, *SIAM Journal of Optimization*, 7 (1997) 367-385.

[14] Meng F., and Zhao G.:”On second-order properties of the Moreau-Yosida regularization for constrained nonsmooth convex programs”, Numer. Funct. Anal. Optim , VOL 25; NUMB 5/6, pages 515-530, 2004

[15] Пшеничнй Б. Н.: *Выпуклый анализ и экстремальные задачи*, Москва „Наука“, 1980.

[16] Пшеничнй Б. Н.: *Неопходимые условя экстремума*, Москва „Наука“, 1982.

[17] Пшеничнй Б. Н.: *Метод линеаризации*, Москва „Наука“, 1983.

[18] Rockafellar R.: *Convex Analysis*, Princeton University Press, New Jersey, 1970.

[19] Polak E.: *Computational Methods in Optimization. A Unified Approach*, Mathematics in Science and Engineering, Vol. 77, Academic Press, New York - London, 1971.

[20] Pytlak R.:”Conjugate gradient algorithms in nonconvex optimization”, Library of Congress Control Number: 2008937497, *Springer-Verlag*, Berlin Heidelberg, 2009.

[21] Shi,Z.J., “Convergence of multi-step curve search method for unconstrained optimization”, *J. Numer. Math*., 12 (2004) 297-309.

[22] Sun W., Sampaio R.J.B., Yuan J.: “Two algorithms for LC^1 unconstrained optimization”, *Journal of Computational Mathematics*, 18 (2000) 621-632.

[23] Qi L.: “Second order analysis of the Moreau-Yosida regularization”, School of Mathematics, The University of New Wales Sydney, New South Wales 2052, Australia, 1998.