U.P.B. Sci. Bull., Series A, Vol. 83, Iss. 3, 2021 ISSN 1223-7027

EXACTNESS OF THE ABSOLUTE VALUE EXACT PENALTY FUNCTION METHOD FOR OPTIMIZATION PROBLEMS VIA UPPER CONVEXIFICATORS

Tadeusz Antczak, [1]

In this paper, the exactness property of the classical absolute value exact penalty function method is analyzed for a new class of nonsmooth optimization problems. Thus, the conditions guaranteeing the equivalence between the sets of optimal solutions in the considered nonsmooth constrained optimization problem and its associated penalized optimization problem with the l_1 exact penalty function are derived in terms of upper convexificators under appropriate convexity hypotheses.

Keywords: absolute value exact penalty function method; penalized optimization problem; exactness of the penalization; upper convexificator; Karush-Kuhn-Tucker optimality conditions.

MSC2010: 90C30, 90C25, 49J52, 49M37

1. Introduction

Exact penalty function methods used for the solution of a constrained optimization problem are based on the construction of a (penalty) function whose a unconstrained minimizer is also a solution of the constrained optimization problem. One of the most frequently used type of an exact penalty function for solving constrained optimization problem is the absolute value exact penalty function, also known as the l_1 exact penalty function (see, for example, [1]-[7], [11], [12], [15], [17], [20], [27], [29], [30]-[31], and others).

The notion of convexificators was first introduced by Demyanov [9]. Recently, the idea of convexificators has been employed to extend and strengthen various results in nonsmooth analysis and optimization (see, for example, [10], [13], [14], [19], [22]-[26], [28], and others).

In this paper, we use the concept of an upper convexificator to derive sharpen results for the l_1 exact penalty function method which is used for solving a nondifferentiable optimization problem. We analyze the main property of the l_1 exact penalty function method, that is, exactness of the penalization. We prove the equivalence between the set of optimal solutions of the considered convex optimization problem and the set of minimizers of its associated exact penalized optimization problem with the absolute value penalty function which is the objective function in such an unconstrained mathematical programming problem. The result established here shows that there does exist a lower bound for the penalty parameter which is equal to the largest Lagrange multiplier, associated to a Karush-Kuhn-Tucker point in the original nonlinear optimization problem such that, for any penalty parameter exceeding than the mentioned above threshold, this equivalence holds. The main result is established for the considered nondifferentiable constrained optimization problem in which every involved function admits an upper convexificator ∂^* and, moreover, is a ∂^*-convex function. The results established in the paper are illustrated by the example of a nonsmooth optimization problem in which the involved functions are ∂^*-convex.

2. **Preliminaries and Problem Formulation**

In this section, we give some basic definitions and results, which will be used in the sequel.

By $\langle \cdot, \cdot \rangle$, we denote the inner product of the vectors. Let S be a nonempty subset of R^n. The convex hull of S is denoted by $conv\, S$.

Lemma 2.1. *Let S_1 and S_2 be two nonempty subsets of R^n. Then,*

$$conv\,(S_1 + S_2) = conv\, S_1 + conv\, S_2.$$

Now, for a general reader, we recall the following definitions of lower and upper Dini derivatives and also lower and upper convexificators in the sense of Demyanov (see [9]).

Definition 2.1. *Let a mapping $f : R^n \to \overline{R} := R \cup \{\infty\}$ be an extended real-valued function, $u \in R^n$ and let $f(u)$ be finite. The lower and upper Dini derivatives of the function f at u in the direction $d \in R^n$ are defined, respectively, by*

$$f^-(u; d) := \liminf_{\alpha \downarrow 0} \frac{f(u + \alpha d) - f(u)}{\alpha},$$

$$f^+(u; d) := \limsup_{\alpha \downarrow 0} \frac{f(u + \alpha d) - f(u)}{\alpha}.$$

It is worthwhile to mention that, in the case where f is locally Lipschitz at u, $f^-(x; d)$ and $f^+(x; d)$ are continuous in d.

Along the lines of Jeyakumar and Luc [22] (see also Dutta and Chandra [14], Golestani and Nobakhtian [19]), we give now the definitions of lower and upper convexificators that will be useful in the sequel.

Definition 2.2. *The function $f : R^n \to \overline{R}$ is said to have an upper convexificator $\partial^* f(u) \subset R^n$ at $u \in R^n$ if $\partial^* f(u)$ is closed and, for each $d \in R^n$,*

$$f^-(u; d) \leq \sup_{\xi \in \partial^* f(u)} \langle \xi, d \rangle .$$

Definition 2.3. *The function $f : R^n \to \overline{R}$ is said to have a lower convexificator $\partial_* f(u) \subset R^n$ at $u \in R^n$ if $\partial_* f(u)$ is closed and, for each $d \in R^n$,*

$$f^+ (u; d) \geq \inf_{\xi \in \partial_* f(u)} \langle \xi, d \rangle .$$

Definition 2.4. *The function $f : R^n \to \overline{R}$ is said to have a convexificator $\partial f(u) \subset R^n$ at $u \in R^n$ if $\partial f(u)$ is both lower and upper convexificator of f at u.*

Definition 2.5. *The function $f : R^n \to \overline{R}$ is said to have a semiregular lower convexificator $\partial_* f(u) \subset R^n$ at $u \in R^n$ if $\partial_* f(u)$ is closed and, for each $d \in R^n$,*

$$f^- (u; d) \geq \inf_{\xi \in \partial_* f(u)} \langle \xi, d \rangle . \tag{1}$$

If equality holds in (1), then $\partial_* f(u)$ is called a lower regular convexificator of f at u.

Definition 2.6. *The function $f : R^n \to \overline{R}$ is said to have a semiregular upper convexificator $\partial^* f(u) \subset R^n$ at $u \in R^n$ if $\partial^* f(u)$ is closed and, for each $d \in R^n$,*

$$f^+ (u; d) \leq \sup_{\xi \in \partial^* f(u)} \langle \xi, d \rangle . \tag{2}$$

If equality holds in (2), then $\partial^* f(u)$ is called an upper regular convexificator of f at u.

Definition 2.7. *The function f is said to be directionally differentiable at $u \in R^n$ if, for every direction $d \in R^n$, the usual one-sided directional derivative of f at u defined by*

$$f'(u; d) := \lim_{\alpha \downarrow 0} \frac{f(u + \alpha d) - f(u)}{\alpha}$$

exists.

Remark 2.1. *Obviously, if f is directionally differentiable at $u \in R^n$, then, for every $d \in R^n$,*

$$f'(u; d) = f^- (u; d) = f^+ (u; d) . \tag{3}$$

A real-valued function $f : R^n \to R$ is said to be locally Lipschitz on R^n if, for any $x \in R^n$, there exist a neighborhood U of x and a positive constant $K_x > 0$ such that, for every $y, z \in U$, it holds $|f(y) - f(z)| \leqq K_x \|y - z\|$.

Definition 2.8. [8] *The Clarke generalized directional derivative of a locally Lipschitz function $f : X \to R$ at $u \in X$ in the direction $d \in R^n$, denoted $f^0 (u; d)$, is given by*

$$f^0(u; d) = \limsup_{\substack{y \to u \\ \alpha \downarrow 0}} \frac{f(y + \alpha d) - f(y)}{\alpha}. \tag{4}$$

Definition 2.9. [8] *The Clarke generalized subgradient of f at $u \in R^n$, denoted $\partial_C f(u)$, is defined by*

$$\partial_C f(u) = \left\{ \xi \in R^n : f^0(u; d) \geq \langle \xi, d \rangle \text{ for all } d \in R^n \right\}. \tag{5}$$

It follows by the above definition that, for any $d \in R^n$, we have

$$f^0(u; d) = \max \left\{ \langle \xi, d \rangle : \xi \in \partial_C f(u) \right\}. \tag{6}$$

Remark 2.2. *Note that, for every fixed $u \in R^n$, the Clarke generalized directional derivative $f^0(u; d)$ is sublinear in d on R^n and $\partial_C f(u)$ is a nonempty compact subset of R^n. Further, the following inequalities*

$$f^0(u; d) \geq f^+(u; d) \geq f^-(u; d)$$

hold for any direction $d \in R^n$. Therefore, by Definition 2.2 and (6), the Clarke subdifferential $\partial_C f(u)$ is a compact and convex upper convexificator of f at u. Further, it has been shown (see, e.g., [22]*) that, for a locally Lipschitz function, many important subdifferentials are convexificators and they may contain the convex hull of a convexificator.*

Remark 2.3. *It is well-known that if f is locally Lipschitz at $u \in R^n$ and regular at u in the sense of Clarke (see* [8]*), then it is directionally differentiable at u and $f'(u; d) = f^0(u; d)$. Hence, by (3), it follows that the following relations*

$$f^-(u; d) = f'(u; d) = f^0(u; d) = f^+(u; d)$$

hold for every direction $d \in R^n$ and, moreover, the functions $f'(u; \cdot)$, $f^-(u; \cdot)$, $f^+(u; \cdot)$ are sublinear.

Remark 2.4. *Note that since $\operatorname{conv} \partial^* f(u) \subset \partial_C f(u)$, certain results that are expressed in terms of upper convexificators may provide sharp conditions even for locally Lipschitz functions.*

Lemma 2.2. *Let $f : R^n \to \overline{R}$ and let $\partial_* f(u)$ and $\partial^* f(u)$ be the lower and upper convexificators of f at $u \in R^n$, respectively. If $\beta > 0$, then $\beta \partial^* f(u)$ is an upper convexificator of βf at $u \in R^n$. If $\beta < 0$, then $\beta \partial_* f(u)$ is an upper convexificator of βf at $u \in R^n$.*

Lemma 2.3. *Assume that the functions f_1, $f_2 : R^n \to R$ admit upper convexificators $\partial^* f_1(u)$ and $\partial^* f_2(u)$ at u, respectively, and that one of the convexificators is upper regular at u. Then, $\partial^* f_1(u) + \partial^* f_2(u)$ is an upper convexificator of $f_1 + f_2$ at u.*

Definition 2.10. *Let $f : R^n \to \overline{R}$ be an extended real-valued function that has an upper convexificator ∂^* at the given point $u \in R^n$. f is said to be ∂^*-convex at u on R^n if the following inequality*

$$f(x) - f(u) \geq \langle \xi, x - u \rangle \tag{7}$$

holds for any $\xi \in \partial^* f(u)$ *and for all* $x \in R^n$. *If the inequality (7) is satisfied at each point* u, *then* f *is said to be a* ∂^**-convex function on* R^n. *If* X *is a nonempty convex subset of* R^n *and (7) is satisfied for all* $x, u \in X$, *then* f *is said to be* ∂^**-convex on* X.

Consider the following constrained optimization problem:

$$\begin{array}{c} \text{minimize } f(x) \\ \text{subject to} \quad g_i(x) \leq 0, \ i \in I = \{1, ...m\}, \quad x \in R^n, \end{array} \tag{P}$$

where $f : R^n \to \overline{R}$ and $g_i : R^n \to \overline{R}$, $i \in I$, are extended real-valued functions defined on R^n. Let $D := \{x \in R^n : g_i(x) \leq 0, \ i \in I\}$ be the set of all feasible solutions of (P). We assume that f and g_i, $i \in I$, admit upper convexificators at every feasible solution. Further, we denote the set of active inequality constraints at point $\overline{x} \in R^n$ by $I(\overline{x}) = \{i \in I : g_i(\overline{x}) = 0\}$.

The Karush-Kuhn-Tucker necessary optimality conditions for nondifferentiable optimization problems by using upper convexificators were proved in the optimization literature (see, for example, by Dutta and Chandra [13], [14], Golestani and Nobakhtian [19], Li and Zhang [25], Luu [26], and others). Now, we give them in the case of the considered nonsmooth scalar optimization problem (P).

Theorem 2.1. *(Karush-Kuhn-Tucker necessary optimality conditions). Let* $\overline{x} \in D$ *be an optimal solution of the considered nondifferentiable optimization problem (P) and some suitable constraint qualification be satisfied at* $\overline{x}$. *Further, assume that the function* $f : R^n \to R$ *admits an upper semiregular convexificator* $\partial^* f(\overline{x})$ *and each constraint function* $g_i : R^n \to R$, $i \in I(\overline{x})$, *admits an upper convexificator* $\partial^* g_i(\overline{x})$. *Then, there exists a Lagrange multiplier* $\overline{\lambda} \in R^m$ *such that*

$$0 \in cl \left(conv\, \partial^* f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i conv\, \partial^* g_i(\overline{x}) \right), \tag{8}$$

$$\overline{\lambda}_i g_i(\overline{x}) = 0, \ i \in I, \tag{9}$$

$$\overline{\lambda}_i \geq 0, \quad i \in I. \tag{10}$$

Remark 2.5. *In addition, if convexificator* $\partial^* f(\overline{x})$ *is bounded, then the Karush-Kuhn-Tucker necessary optimality condition (8) reduces to the relation*

$$0 \in conv\, \partial^* f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i conv\, \partial^* g_i(\overline{x})$$

(see, for example, [25]*).*

Definition 2.11. *The point* $\overline{x} \in D$ *is said to be a Karush-Kuhn-Tucker point (a KKT point, for short) if there exists a Lagrange multiplier* $\overline{\lambda} \in R^m$ *such that the conditions (8)-(10) are satisfied at* $\overline{x}$.

We shall assume that the suitable constraint qualification is fulfilled at each optimal solution of the considered optimization problem (P).

3. The exactness of the l_1 exact penalty function method

For the considered optimization problem (P), we consider the penalty function defined by $P(x,c) = f(x) + c\varphi(x)$, where $\varphi : R^n \to R_+$ is given. If a threshold value $\overline{c}$ for the penalty parameter c exists such that, for every $c \geq \overline{c}$,

$$\arg\min\{f(x) : x \in D\} = \arg\min\{P(x,c) : x \in R^n\},$$

then the function $P(x,c)$ is called an exact penalty function.

The most popular nondifferentiable exact penalty function is the absolute value penalty function also called the l_1 exact penalty function. For the considered optimization problem (P), the associated penalized optimization problem (P(c)) constructed in the the l_1 exact penalty function method is given by

$$\text{minimize } P(x,c) = f(x) + c\sum_{i=1}^{m} g_i^+(x), \qquad (\mathrm{P}(c)) \tag{11}$$

where, for a given constraint $g_i(x) \leq 0$, the function g_i^+ is defined by

$$g_i^+(x) = \begin{cases} 0 & \text{if } g_i(x) \leq 0, \\ g_i(x) & \text{if } g_i(x) > 0. \end{cases} \tag{12}$$

We call the unconstrained optimization problem defined by (11) the penalized optimization problem with the absolute value exact penalty function.

Now, we prove the equivalence between the sets of optimal solutions in the original constrained optimization problem (P) and its associated penalized optimization problem (P(c)) with the l_1 exact penalty function for any penalty parameter c exceeding a given threshold. This result is derived in terms of upper convexificators under convexity hypotheses.

First, we prove that a Karush-Kuhn-Tucker point of the considered extremum problem (P) is a minimizer of its associated penalized optimization problem (P(c)).

Theorem 3.1. *Let $\overline{x} \in D$ be a Karush-Kuhn-Tucker point of the constrained optimization problem (P) with Lagrange multiplier $\overline{\lambda} \in R^m$. Furthermore, assume that the function f admits an upper semiregular convexificator at $\overline{x}$, each constraint function g_i, $i \in I$, admits an upper convexificator at $\overline{x}$ and, moreover, they are ∂^*-convex at $\overline{x}$ on R^n. If c is assumed to be sufficiently large (it is sufficient to set $c \geq \max\{\overline{\lambda}_i,\, i \in I\}$), then $\overline{x}$ is also a minimizer of its penalized optimization problem (P(c)).*

Proof. Assume that $\overline{x} \in D$ is a Karush-Kuhn-Tucker point of the constrained optimization problem (P). By assumption, functions f, g_i, $i \in I$, are ∂^*-convex at $\overline{x}$ on D. Hence, by Definition 2.10, the following inequalities

$$f(x) - f(\overline{x}) \geq \langle \xi, x - \overline{x} \rangle, \tag{13}$$

$$g_i(x) - g_i(\overline{x}) \geq \langle \zeta_i, x - \overline{x} \rangle, \quad i \in I \tag{14}$$

hold for all $x \in R^n$ and for any $\xi \in \partial^* f(\overline{x})$, $\zeta_i \in \partial^* g_i(\overline{x})$, $i \in I$, respectively. Multiplying each inequality (14) by $\overline{\lambda}_i \geq 0$, $i \in I$, and then adding both sides of the resulting inequalities and (13), we get that the inequality

$$f(x) + \sum_{i=1}^{m} \overline{\lambda}_i g_i(x) - \left(f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i g_i(\overline{x}) \right) \geq \langle \xi, x - \overline{x} \rangle + \sum_{i=1}^{m} \overline{\lambda}_i \langle \zeta_i, x - \overline{x} \rangle \quad (15)$$

holds for any $\xi \in \partial^* f(\overline{x})$, $\zeta_i \in \partial^* g_i(\overline{x})$, $i \in I$, and for all $x \in R^n$. Since inequality (15) is satisfied for any $\xi \in \partial^* f(\overline{x})$, $\zeta_i \in \partial^* g_i(\overline{x})$, $i \in I$, and for all $x \in R^n$, we have that the inequality

$$f(x) + \sum_{i=1}^{m} \overline{\lambda}_i g_i(x) - \left(f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i g_i(\overline{x}) \right) \geq \quad (16)$$

$$\left\langle \sup_{\xi \in \partial^* f(\overline{x})} \xi, x - \overline{x} \right\rangle + \sum_{i=1}^{m} \overline{\lambda}_i \sup_{\zeta_i \in \partial^* g_i(\overline{x})} \langle \zeta_i, x - \overline{x} \rangle$$

holds. By assumption, $\overline{x} \in D$ is a Karush-Kuhn-Tucker point in the constrained optimization problem (P) and, moreover, the Karush-Kuhn-Tucker necessary optimality conditions (8)-(10) are satisfied at $\overline{x}$ with a Lagrange multiplier $\overline{\lambda} \in R^m$. Hence, by Lemma 2.1, the Karush-Kuhn-Tucker necessary optimality condition (8) gives

$$0 \in cl\, conv \left(\partial^* f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i \partial^* g_i(\overline{x}) \right). \quad (17)$$

Let us denote $\Omega(\overline{x}) = \partial^* f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i \partial^* g_i(\overline{x})$. Since $\overline{\lambda}_i \geq 0$, $i = 1, ..., m$, by Lemma 2.2 and Lemma 2.3, it follows that $\Omega(\overline{x})$ is an upper convexificator of the function $\partial^* f(\cdot) + \sum_{i=1}^{m} \overline{\lambda}_i \partial^* g_i(\cdot)$ at $\overline{x}$. Hence, we have that the following relation

$$\sup_{\vartheta \in \Omega(\overline{x})} \langle \vartheta, x - \overline{x} \rangle = \sup_{\xi \in \partial^* f(\overline{x})} \langle \xi, x - \overline{x} \rangle + \sum_{i=1}^{m} \overline{\lambda}_i \sup_{\zeta_i \in \partial^* g_i(\overline{x})} \langle \zeta_i, x - \overline{x} \rangle \quad (18)$$

holds for all $x \in R^n$. By (16), it follows that there exists $\overline{\vartheta} \in cl\, conv\, \Omega(\overline{x})$ such that $\left\langle \overline{\vartheta}, x - \overline{x} \right\rangle = 0$. This implies that the inequality

$$\sup_{\vartheta \in cl\, conv\, \Omega(\overline{x})} \langle \vartheta, x - \overline{x} \rangle \geq 0. \quad (19)$$

holds for all $x \in R^n$. It is well-known that the support function of a nonempty set $S \subset R^n$ is also the support function of the closure of S, and even of the closed convex hull of S (see [16]). In other words, the support functional of any set and its closed convex hull are identical. Hence, by the upper convexificator $\Omega(\overline{x})$ and the usual calculus of support functions, we observe that, for all $x \in R^n$,

$$\sup_{\vartheta \in cl\, conv\, \Omega(\overline{x})} \langle \vartheta, x - \overline{x} \rangle = \sup_{\vartheta \in \Omega(\overline{x})} \langle \vartheta, x - \overline{x} \rangle. \quad (20)$$

Thus, combining (16)-(20), we get that the following inequality

$$f(x) + \sum_{i=1}^{m} \overline{\lambda}_i g_i(x) \geq f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i g_i(\overline{x}) \tag{21}$$

holds for all $x \in R^n$. By (12), it follows that

$$f(x) + \sum_{i=1}^{m} \overline{\lambda}_i g_i^+(x) \geq f(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i g_i^+(\overline{x}). \tag{22}$$

For $c \geq \max\left\{\overline{\lambda}_i,\ i \in I\right\}$, we have that (18) gives, for all $x \in R^n$,

$$f(x) + c\sum_{i=1}^{m} g_i^+(x) \geq f(\overline{x}) + c\sum_{i=1}^{m} g_i^+(\overline{x}). \tag{23}$$

Therefore, by (11), we conclude that the following inequality $P(x,c) \geq P(\overline{x},c)$ holds for all $x \in R^n$. This means that $\overline{x}$ is a minimizer of the penalized optimization problem (P(c)). Thus, the proof of this theorem is completed. □

Corollary 3.1. *Let $\overline{x}$ be an optimal point of the considered optimization problem (P). Furthermore, assume that all hypotheses of Theorem 3.1 are fulfilled. Then $\overline{x}$ is also a minimizer of the penalized optimization problem (P(c)) with the absolute value exact penalty function.*

Now, by an example of a nonsmooth optimization problem, we illustrate the above result in terms of upper convexificators under convexity hypotheses.

Example 3.1. *Consider the following nonsmooth optimization problem*

$$\min f(x) = \begin{cases} x^2\left|\cos\frac{\pi}{x}\right| + x & \text{if } \ x \neq 0 \\ 0 & \text{if } \ x = 0 \end{cases} \qquad \textit{(P1)}$$

$$g(x) = |x| - x \leq 0.$$

Note that $D = \{x \in R : x \geq 0\}$ and $\overline{x} = 0$ is an optimal solution of the considered nonsmooth optimization problem (P1). Since we use the l_1 exact penalty method for solving the extremum problem (P1), we construct the following unconstrained optimization problem

$$P(x,c) = f\,(x) + c\max\{0, |x| - x\} \to \min. \qquad \textit{(P1(c))}$$

It can be seen that $f^-(0;d) = f^+(0;d) = d$ for every $d \in R$. Note that $\overline{x} = 0$ is a Karush-Kuhn-Tucker point of the problem (P1) and the Karush-Kuhn-Tucker necessary optimality conditions (8)-(10) are fulfilled at $\overline{x}$. Indeed, the Lagrange multiplier $\overline{\lambda} = \frac{1}{2}$ satisfies the condition: $0 \in conv\,\partial^* f(\overline{x}) + \overline{\lambda} conv\,\partial^* g(\overline{x})$, where $\partial^* f(\overline{x}) = \{1\}$ and $\partial^* g(\overline{x}) = \{-2; 0\}$ and, moreover, the condition $\overline{\lambda} g(\overline{x}) = 0$. Note, by Definition 2.5, $\partial^* f(\overline{x})$ and $\partial^* g(\overline{x})$ are regular convexificators f and g at $\overline{x}$, respectively. Further, it can be shown, by Definition 2.10, that the objective function f and the constraint function g are ∂^*-convex at $\overline{x}$ on R. Then, by

Theorem 3.1, it follows that $\overline{x} = 0$ *is also a minimizer in the penalized optimization problem (P1(c)) with the absolute value penalty function for any parameter* $c \geq \frac{1}{2}$*. Note that* $conv\, \partial^* f(\overline{x}) \subset \partial_C f(\overline{x}) = [-\pi + 1, \pi + 1]$*. Hence, this result is sharper than the similar result established in the literature (see, for example,* [2], [6], [17]*).*

Proposition 3.1. *Let* $\overline{x}$ *be a minimizer of the penalized optimization problem (P(c)) with the* l_1 *exact penalty function. Further, assume that the objective function* f *is lower semicontinuous at* $\overline{x}$ *and the set* D *of all feasible solutions in the problem (P) is compact. Then there exists a Karush-Kuhn-Tucker point of the considered optimization problem (P).*

Proof. Assume that $\overline{x}$ is a minimizer of the penalized optimization problem (P(c)). Then, the following inequality $P(x,c) \geq P(\overline{x},c)$ holds for all $x \in R^n$. Hence, by the definition of the penalized optimization problem (P(c)), it follows that the following inequality

$$f(x) + c\sum_{i=1}^{m} g_i^+(x) \geq f(\overline{x}) + c\sum_{i=1}^{m} g_i^+(\overline{x})$$

holds for all $x \in R^n$. Thus, the following inequality $f(x) \geq f(\overline{x})$ holds for all $x \in D$. This means that f is lower bounded on the compact set D. Since f is a lower semicontinuous function, there exists $\widetilde{x} \in D$ at which f achieves its minimum on D. Hence, there exists the Lagrange multiplier $\widetilde{\lambda} \in R^m$ such that the Karush-Kuhn-Tucker necessary optimality conditions (8)-(10) are satisfied at $\widetilde{x}$ with this Lagrange multiplier. Then, by Definition 2.11, $\widetilde{x} \in D$ is a Karush-Kuhn-Tucker point of the considered nonsmooth optimization problem (P). □

Theorem 3.2. *Let* f *be a lower semicontinuous function and the point* $\overline{x}$ *be a minimizer of the penalized optimization problem (P(c)) with the* l_1 *exact penalty function. Furthermore, assume that the functions* f*,* g_i*,* $i \in I$*, are* ∂^**-convex at any Karush-Kuhn-Tucker point of the considered constrained optimization problem (P) and the set* D *of all feasible solutions in (P) is compact. If the penalty parameter* c *is sufficiently large (it is sufficient that* c *satisfies the following condition* $c > \max\left\{\widetilde{\lambda}_i,\ i \in I\right\}$*), where* $\widetilde{\lambda}$ *is the Lagrange multiplier associated to any Karush-Kuhn-Tucker point* $\widetilde{x}$ *in (P) with Lagrange multiplier* $\widetilde{\lambda} \in R^m$*, then* $\overline{x}$ *is also optimal of the nonsmooth optimization problem (P).*

Proof. Assume that $\overline{x}$ is a minimizer of the penalized optimization problem (P(c)). Then, by Proposition 3.1, there exists a KKT point $\widetilde{x}$ of the problem (P). Let the Karush-Kuhn-Tucker necessary optimality conditions (8)-(10) be satisfied with a Lagrange multiplier $\widetilde{\lambda} \in R^m$.

Now, we prove that $\overline{x}$ is also optimal in the considered optimization problem (P). First, we show that $\overline{x}$ is feasible for the problem (P). By contradiction, suppose that $\overline{x}$ is not feasible in (P). As we have established above, the given constrained optimization problem (P) has an optimal solution $\widetilde{x}$. Hence, since a constraint qualification is satisfied at $\widetilde{x}$, there exist Lagrange multiplier vector $\widetilde{\lambda} \in R^m$ such that the Karush-Kuhn-Tucker necessary optimality conditions (8)-(10) are satisfied at $\widetilde{x}$. By assumption, the functions f, g_i, $i \in I$, are ∂^*-convex at any Karush-Kuhn-Tucker point of (P). Therefore, by Definition 2.10, it follows that the inequalities

$$f(\overline{x}) - f(\widetilde{x}) \geq \langle \xi, \overline{x} - \widetilde{x} \rangle, \quad \forall \xi \in \partial^* f(\widetilde{x}), \tag{24}$$

$$g_i(\overline{x}) - g_i(\widetilde{x}) \geq \langle \zeta_i, \overline{x} - \widetilde{x} \rangle, \quad \forall \zeta_i \in \partial^* g_i(\overline{x}), \quad \forall i \in I \tag{25}$$

hold. Multiplying (25) by the associated Lagrange multiplier and then adding them to (24), we get

$$f(\overline{x}) - f(\widetilde{x}) + \sum_{i=1}^{m} \widetilde{\lambda}_i g_i(\overline{x}) - \sum_{i=1}^{m} \widetilde{\lambda}_i g_i(\widetilde{x}) \geq \left\langle \xi + \sum_{i=1}^{m} \widetilde{\lambda}_i \zeta_i, \overline{x} - \widetilde{x} \right\rangle.$$

Using (12) with the Karush-Kuhn-Tucker necessary optimality conditions (8), (9) and the feasibility of $\widetilde{x}$ in the original extremum problem (P), in the similar way as in the proof of Theorem 3.1, we get

$$f(\overline{x}) + \sum_{i=1}^{m} \widetilde{\lambda}_i g_i^+(\overline{x}) \geq f(\widetilde{x}). \tag{26}$$

By assumption, the penalty parameter c is sufficiently large (it is sufficient that $c > \max\left\{\widetilde{\lambda}_i,\, i \in I\right\}$). Since $\overline{x}$ is assumed to be not feasible in the given optimization problem (P), at least one of $g_i^+(\overline{x})$ must be nonzero. Therefore, (26) yields

$$f(\overline{x}) + c\sum_{i=1}^{m} g_i^+(\overline{x}) > f(\widetilde{x}). \tag{27}$$

Hence, by $\widetilde{x} \in D$ and (27), we have

$$f(\overline{x}) + c\sum_{i=1}^{m} g_i^+(\overline{x}) > f(\widetilde{x}) + c\sum_{i=1}^{m} g_i^+(\widetilde{x}).$$

Then, by the definition of the l_1 exact penalty function (see (11)), it follows that the following inequality $P(\overline{x}, c) > P(\widetilde{x}, c)$ holds, which is a contradiction to the assumption that $\overline{x}$ is a minimizer of the penalized optimization problem (P(c)) with the absolute value penalty function. Thus, we have proved that $\overline{x}$ is feasible in the given constrained optimization problem (P). Hence, the optimality of $\overline{x}$ in the considered constrained optimization problem (P) follows directly from (**??**). This completes the proof of this theorem. □

Hence, the main result in the paper is formulated in the theorem below:

Theorem 3.3. *Let all hypotheses of Corollary 3.1 and Theorem 3.2 be fulfilled. Then, $\overline{x}$ is an optimal solution in the considered extremum problem (P) if and only if it is a minimizer in its associated exact penalized optimization problem (P(c))* *with the absolute value penalty function for all penalty parameters exceeding the threshold which is equal to the maximal value of the Lagrange multiplier corresponding to a KKT point of the problem (P). In other words, the set of optimal solutions in the considered nonsmooth extremum problem (P) and the set of minimizers in its associated exact penalized optimization problem (P(c)) with the absolute value penalty function coincide for all penalty parameters exceeding the threshold which is equal to the maximal value of the Lagrange multiplier corresponding to a KKT point of the problem (P).*

4. **Conclusions**

In this paper, exactness of the penalization of the classical exact l_1 penalty function method has been investigated. Hence, new conditions guarantying the equivalence of optimal solutions in the original nonsmooth minimization problem and its penalized optimization problem constructed in this method have been derived in the terms of upper convexificators of the involved functions. It is known that convexificators can be viewed as a weaker version of the notion of subdifferentials so that it will lead in sharper results in nonsmooth analysis. Therefore, the results established in the paper for the classical l_1 exact penalty function method are sharpen in comparison to those ones existing in the literature.

REFERENCES

[1] *T. Antczak*, Exact penalty functions method for mathematical programming problems involving invex functions, European J. Oper. Res. **198** (2009), 29-36.

[2] *T. Antczak*, The exact l_1 penalty function method for constrained nonsmooth invex optimization problems, in: D. Hömberg, F, Tröltzsch (Eds.) System Modeling and Optimization. Vol. **391** of the series IFIP Advances in Information and Communication Technology, pp. 461-470, 2013.

[3] *M. S. Bazaraa, H. D. Sherali, C. M. Shetty*, Nonlinear Programming: Theory and Algorithms, John Wiley and Sons; New York, 1991

[4] *D. P. Bertsekas*, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, Inc., 1982.

[5] *D. P. Bertsekas, A. E. Koksal*, Enhanced optimality conditions and exact penalty functions, Proceedings of Allerton Conference; September 2000, 2000

[6] *J. F. Bonnans, J. Ch. Gilbert, C. Lemaréchal, C. A. Sagastizábal*, Numerical Optimization. Theoretical and Practical Aspects, Springer-Verlag Berlin Heidelberg New York, 2003.

[7] *C. Charalambous*, A lower bound for the controlling parameters of the exact penalty functions, Math. Program. **15** (1978), 278-290.

[8] *F. H. Clarke*, Optimization and Nonsmooth Analysis, A Wiley-Interscience Publication, John Wiley&Sons, Inc., 1983.

[9] *V. F. Demyanov*, Convexification and concavification of a positively homogenous function by the same family of linear functions, Universita di Pisa, Report 3, 208, 802, 1994.

[10] *V. Demyanov, A. Rubinov*, Quasidifferentiability and Related Topics, Nonconvex Optimization and its Applications Vol. **43** Springer-Science+Business Media, B.V., 2000.

[11] *G. Di Pillo, L. Grippo*, Exact penalty functions in constrained optimization, SIAM J. Control Optim. **27** (1989), 1333-1360.

[12] *S. Dolecki, S. Rolewicz*, Exact penalty for local minima, SIAM J. Control Optim. **17** (1979), 111-124

[13] *J. Dutta, S. Chandra*, Convexificators, generalized convexity, and optimality conditions, J. Optim. Theory Appl. **113** (2002), 41-64.

[14] *J. Dutta, S. Chandra*, Convexificators, generalized convexity and vector optimization, Optimization **53** (2004), 77-94.

[15] *S. P. Han, O. L. Mangasarian*, Exact penalty functions in nonlinear programming, Math. Program. **17** (1979), 251-269.

[16] *J. -B. Hiriart-Urruty, C. Lemarechal*, Convex Analysis and Minimization, Algorithm I: Fundamentals, Springer-Verlag, Berlin Heilderberg, 1993.

[17] *F. Facchinei*, Exact penalty functions and Lagrange multipliers, Optimization **22** (1991), 579-606.

[18] *N. Gadhi*, Necessary and sufficient optimality conditions for fractional multiobjective problems, Optimization **57** (2008), 527-537.

[19] *M. Golestani, S. Nobakhtian*, Nonsmooth multiobjective programming and constraint qualifications, Optimization **62** (2013), 783-795.

[20] *S. M. H. Janesch, L. T. Santos*, Exact penalty methods with constrained subproblems, Investigatión Operativa **7** (1997), 55-65.

[21] *A. Jayswal, I. Stancu-Minasian, J. Banerjee*, Optimality conditions and duality for interval-valued optimization problems using convexifactors, Rend. Circ. Mat. Palermo **65** (2016), 17-32.

[22] *V. Jeyakumar, D. T. Luc*, Nonsmooth calculus, minimality, and monotonicity of convexiflcators, J. Optim. Theory Appl. **101** (1999), 599-621.

[23] *A. Kabgani, M. Soleimani-damaneh, M. Zamani*, Optimality conditions in optimization problems with convex feasible set using convexificators, Math. Methods Oper. Res. **86** (2017), 103-121.

[24] *X. F. Li, J. Z. Zhang*, Stronger Kuhn-Tucker type conditions in nonsmooth multiobjective optimization: locally Lipschitz case, J. Optim. Theory Appl. **127** (2005), 367-388.

[25] *X. F. Li, J. Z. Zhang*, Necessary optimality conditions in terms of convexificators in Lipschitz optimization, J. Optim. Theory Appl. **131** (2006), 429-452.

[26] *D. V. Luu*, Necessary and sufficient conditions for efficiency via convexificators, J. Optim. Theory Appl. **160** (2014), 510-526.

[27] *O. L. Mangasarian*, Sufficiency of exact penalty minimization, SIAM J. Control Optim. **23** (1985), 30-37.

[28] *Y. Pandey, S. K. Mishra*, Duality for nonsmooth optimization problems with equilibrium constraints using convexificators, J. Optim. Theory Appl. **171** (2016), 694-707.

[29] *A. L. Peressini, F. E. Sullivan, J. J. Uhl Jr.*, The Mathematics of Nonlinear Programming, Springer-Verlag New York Inc., 1998.

[30] *E. Rosenberg*, Exact penalty functions and stability in locally Lipschitz programming, Math. Program. **30** (1984), 340-356.

[31] *W. Sun, Y.-X. Yuan*, Optimization, Theory and Methods: Nonlinear Programming, Optimization and its Applications Vol. **1**, Springer, 2006.

[1]Associate Professor, Faculty of Mathematics and Computer Science,
University of Łódź, Banacha 22, 90-238 Łódź,
e-mail: tadeusz.antczak@wmii.uni.lodz.pl