CrossMark

ORIGINAL PAPER

The exact absolute value penalty function method for identifying strict global minima of order m in nonconvex nonsmooth programming

Tadeusz Antczak[1]

Abstract In this paper, it is demonstrated that the exact absolute value penalty function method is useful for identifying the special sort of minimizers in nonconvex nonsmooth optimization problems with both inequality and equality constraints. The equivalence between the sets of strict global minima of order m in nonsmooth minimization problem and of its associated penalized optimization problem with the exact l_1 penalty function is established under nondifferentiable (F, ρ)-convexity assumptions imposed on the involved functions. The threshold of the penalty parameter, above which this result holds, is also given.

1 Introduction

The notion of a strict local minimizer of order m plays an important role in the convergence analysis of iterative numerical methods (see, for example, [1]) and in stability analysis (see, for example, [2,3]). Some results and optimality conditions concerning characterizations of such minimizers for nonlinear constrained mathematical programming problems have been derived by Auslender [4], Studniarski [5], Ward [6]. These results, in general, suggest that these minimizers are often exactly those satisfy an

✉ Tadeusz Antczak
    antczak@math.uni.lodz.pl

[1] Faculty of Mathematics and Computer Science, University of Łódź,
     Banacha 22, 90-238 Łódź, Poland

"m-th derivative test". In this paper, we present a different approach for identifying of such minimizers.

In the past few years, considerable attention has been given to devising methods for solving nonlinear programming problems via unconstrained minimization techniques. One of classes of such methods which has emerged as very promising is the class of exact penalty function methods. The most popular nondifferentiable exact penalty function is the absolute value penalty function method (also called the exact l_1 penalty function method) and it has been investigated in [7–26], and others. Most of the results established on the nondifferentiable exact l_1 penalty function is devoted to the study of conditions ensuring that an optimal solution in the given convex optimization problem is also an unconstrained solution of the penalty function.

In this paper, we present a new characterization of the exact penalty method with the absolute value penalty function used to solve a class of nonconvex nondifferentiable optimization problems involving both inequality and equality constraints in which the functions constituting them are locally Lipschitz (F, ρ)-convex functions of order m, not necessarily with respect to the same ρ. Namely, we use the exact l_1 penalty function method to find a strict global minimizer of order m in the considered nonconvex nondifferentiable optimization problem involving both inequality and equality constraints. Indeed, we associate a strict global minimizer of order m in the nonsmooth constrained extremum problem with a strict global minimizer of order m in a unconstrained optimization problem (called penalized optimization problem) constructed in this approach in which the absolute value penalty function is minimized. Further, we also establish the converse result, that is, that a strict global minimizer of order m in the penalized optimization problem with the exact l_1 penalty function is also a strict global minimizer of order m in the nonlinear constrained optimization problem with both inequality and equality constraints. In this way, we prove that the sets of strict global minimizers of order m in both optimization problems coincide for a larger class of optimization problems than convex ones. The treshold of the penalty parameter above which this result holds is equal to the largest Lagrange multiplier in the absolute value. The results established in the paper are illustrated by suitable examples of nonconvex nondifferentiable optimization problems solving by using the exact l_1 penalty function method.

2 Preliminaries and problem formulation

Throughout this section, X is a nonempty open subset of R^n.

Definition 1 [27] The Clarke generalized subgradient of f at $x \in X$, denoted $\partial f(x)$, is defined by $\partial f(x) = \left\{ \xi \in R^n : f^0(x; v) \geq \xi^T v \text{ for all } v \in R^n \right\}$, where $f^0(x; v)$ is the Clarke generalized directional derivative of a locally Lipschitz function $f : X \to R$ at $x \in X$ in the direction $v \in R^n$ given by $f^0(x; v) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{f(y + \lambda v) - f(y)}{\lambda}$.

Definition 2 A functional $F : X \times X \times R^n \to R$ is sublinear (with respect to the third component) if, for all $x, u \in X$,

(i) $F(x, u; q_1 + q_2) \leq F(x, u; q_1) + F(x, u; q_2), \forall q_1, q_2 \in R^n$,

(ii) $F(x, u; \alpha q) = \alpha F(x, u; q), \forall \alpha \in R_+, \forall q \in R^n$.

By (ii), it is clear that

$$F(x, u; 0) = 0. \tag{1}$$

Now, we give the definition of a nondifferentiable (F, ρ)-convex function of order m (see [28]).

Definition 3 A locally Lipschitz function $f : X \to R$ is said to be (F, ρ)-convex of order m at $u \in X$ on X if, there exist a sublinear (with respect to the third component) functional $F : X \times X \times R^n \to R$, an integer $m \geq 1$ and a real number ρ such that, the following inequality

$$f(x) - f(u) \geq F(x, u; \xi) + \rho \|x - u\|^m$$

holds for every $\xi \in \partial f(u)$ and all $x \in X$. If the above inequality is satisfied for all $u \in X$, then f is said to be (F, ρ)-convex on X.

In the paper, we consider the following constrained optimization problem:

$$\begin{aligned}
&\text{minimize } f(x) \\
\text{subject to } \quad &g_i(x) \leq 0, i \in I = \{1, \ldots p\} \qquad\qquad \text{(P)} \\
&h_j(x) = 0, j \in J = \{1, \ldots, s\},
\end{aligned}$$

where $f : X \to R$ and $g_i : X \to R, i \in I, h_j : X \to R, j \in J$, are locally Lipschitz functions on a nonempty open set $X \subset R^n$.

For the purpose of simplifying our presentation, we will next introduce some notations which will be used frequently throughout this paper. Let $D := \{x \in X : g_i(x) \leq 0, i \in I, h_j(x) = 0, j \in J\}$ be the set of all feasible solutions of problem (P). Further, by $I(\overline{x}) = \{i \in I : g_i(\overline{x}) = 0\}$, we denote the set of active inequality constraints at point $\overline{x} \in D$.

The concept of a strict local minimizer of order m was defined by Cromme [1], under the name "strongly unique" minimizer, in a study of iterative numerical methods.

Definition 4 Let $m \geq 1$ be an integer. We say that $\overline{x}$ is a strict global minimizer of order m in the considered optimization problem (P) if there exists $\beta > 0$ such that

$$f(x) \geq f(\overline{x}) + \beta \|x - \overline{x}\|^m$$

for all $x \in D$.

It is well known (see, for example, [27,29,30]) that, if $\overline{x} \in D$ is an optimal solution for problem (P), then the following conditions, known as the generalized form of the Karush–Kuhn–Tucker conditions, are satisfied:

Theorem 5 (Generalized Karush–Kuhn–Tucker necessary optimality conditions). *Let $\overline{x} \in D$ be an optimal solution of problem (P) and a suitable constraint qualification be satisfied at $\overline{x}$. Then, there exist $\overline{\lambda} \in R^p$, $\overline{\mu} \in R^s$ such that*

$$0 \in \partial f(\overline{x}) + \sum_{i=1}^{p} \overline{\lambda}_i \partial g_i(\overline{x}) + \sum_{j=1}^{s} \overline{\mu}_j \partial h_j(\overline{x}), \tag{2}$$

$$\overline{\lambda}_i g_i(\overline{x}) = 0, \quad i \in I, \tag{3}$$

$$\overline{\lambda}_i \in R_+, \quad i \in I. \tag{4}$$

Remark 6 Note that, following Hiriart-Urruty [30], a constraint qualification assuring the conclusion of the above theorem is the following one: there exists $v \in R^n$, $g_i^0(\overline{x}; v) < 0$, $i \in I(\overline{x})$. In the presence of inequality constraints, from Fritz John type optimality conditions, Clarke [27] established generalized Karush–Kuhn–Tucker necessary optimality conditions under the assumptions of "calmness" of the optimization problem. This constraint qualification has the advantage to be present in most problems, even if it seems difficult to verify it in general. Further, it is possible to use the following Cottle constraint qualification: either $g_i(\overline{x}) < 0$, $i \in I$ or $0 \notin conv\{\partial g_i(\overline{x}) : i \in I(\overline{x})\}$. Since the results in the paper have been established for nondifferentiable optimization problems with generalized convex functions (that is, locally Lipschitz (F, ρ)-convex of order m), it is possible to use also generalized Slater constraint qualification.

In the paper, we will assume that a suitable constraint qualification is satisfied at any optimal solution in the considered nonlinear constrained optimization problem (P).

Definition 7 The point $\overline{x} \in D$ is said to be a Karush–Kuhn–Tucker point (*a KKT point, for short*) if there exist the Lagrange multipliers $\overline{\lambda} \in R^p$, $\overline{\mu} \in R^s$ such that the conditions (2)–(4) are satisfied at $\overline{x}$ with these Lagrange multipliers.

3 The exact l_1 penalty method for optimization problems with locally Lipschitz (F, ρ)-convex functions of order m

The unconstrained optimization problem with the exact l_1 penalty function constructed in the exact l_1 penalty function method for the considered constrained minimization problem (P) can be written in the following form

$$\text{minimize } P(x, c) = f(x) + c\left[\sum_{i \in I} g_i^+(x) + \sum_{j \in J} |h_j(x)|\right], \quad \text{(P(c))} \tag{5}$$

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

$$g_i^+(x) = \begin{cases} 0 & \text{if } g_i(x) \leq 0, \\ g_i(x) & \text{if } g_i(x) > 0. \end{cases} \tag{6}$$

We will call the unconstrained optimization problem (P(c)) the penalized optimization problem (with the absolute value penalty function).

Theorem 8 *Let $\overline{x} \in D$ be a Karush–Kuhn–Tucker point of the constrained optimization problem (P), at which the Generalized Karush–Kuhn–Tucker conditions (2)–(4) are satisfied with the Lagrange multipliers $\overline{\lambda} \in R^p$ and $\overline{\mu} \in R^s$. Let $J^+(\overline{x}) = \{j \in J : \overline{\mu}_j > 0\}$ and $J^-(\overline{x}) = \{j \in J : \overline{\mu}_j < 0\}$. Furthermore, assume the following hypotheses are satisfied:*

(a) *the objective function f is locally Lipschitz $\left(F, \rho_f\right)$-convex of order m at $\overline{x}$ on X,*
(b) *the inequality constraints g_i, $i \in I(\overline{x})$, are locally Lipschitz $\left(F, \rho_{g_i}\right)$-convex of order m at $\overline{x}$ on X,*
(c) *the equality constraints h_j, $j \in J^+(\overline{x})$, are locally Lipschitz $\left(F, \rho_{h_j}^+\right)$-convex of order m at $\overline{x}$ on X,*
(d) *the functions $-h_j$, $j \in J^-(\overline{x})$, are locally Lipschitz $\left(F, \rho_{h_j}^-\right)$-convex of order m at $\overline{x}$ on X,*
(e) *$\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- > 0.$*

If the penalty parameter c is assumed to be sufficiently large (it is sufficient to set $c \geq \max \left\{ \overline{\lambda}_i, i \in I, \left|\overline{\mu}_j\right|, j \in J \right\}$, where $\overline{\lambda}_i$, $i = 1, \ldots, p$, $\overline{\mu}_j$, $j = 1, \ldots, s$, are the Lagrange multipliers associated with the constraint g_i and h_j, respectively), then $\overline{x}$ is also a strict global minimizer of order m in its associated penalized optimization problem (P(c)) with the exact l_1 penalty function.

Proof By assumption, $\overline{x}$ is a Karush–Kuhn–Tucker point of the constrained optimization problem (P). Then, there exist the Lagrange multipliers $\overline{\lambda} \in R^p$ and $\overline{\mu} \in R^s$ such that the Generalized Karush–Kuhn–Tucker conditions (2)–(4) are satisfied at $\overline{x}$. Since assumptions (a)–(d) are fulfilled, by Definition 3, the following inequalities

$$f(x) - f(\overline{x}) \geq F(x, \overline{x}; \xi) + \rho_f \|x - \overline{x}\|^m, \quad \forall \xi \in \partial f(\overline{x}), \tag{7}$$

$$g_i(x) - g_i(\overline{x}) \geq F(x, \overline{x}; \zeta_i) + \rho_{g_i} \|x - \overline{x}\|^m, \quad \forall \zeta_i \in \partial g_i(\overline{x}), \quad i \in I(\overline{x}), \tag{8}$$

$$h_j(x) - h_j(\overline{x}) \geq F(x, \overline{x}; \varsigma_j) + \rho_{h_j}^+ \|x - \overline{x}\|^m, \quad \forall \varsigma_j \in \partial h_j(\overline{x}), \quad j \in J^+(\overline{x}), \tag{9}$$

$$-h_j(x) + h_j(\overline{x}) \geq F(x, \overline{x}; -\varsigma_j) + \rho_{h_j}^- \|x - \overline{x}\|^m, \quad \forall \varsigma_j \in \partial h_j(\overline{x}), \quad j \in J^-(\overline{x}) \tag{10}$$

hold for all $x \in X$. Multiplying the inequalities (8) and (9) by the corresponding Lagrange multipliers and the inequalities (10) by $-\overline{\mu}_j$, $j \in J^-(\overline{x})$, then, adding both sides of the obtained inequalities, we get, for all $x \in X$,

$$\sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(x) - \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(\overline{x})$$
$$\geq \sum_{i \in I(\overline{x})} \overline{\lambda}_i F(x, \overline{x}; \zeta_i) + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} \|x - \overline{x}\|^m, \quad \forall \zeta_i \in \partial g_i(\overline{x}), \tag{11}$$

$$\sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(\overline{x})$$
$$\geq \sum_{j \in J^+(\overline{x})} \overline{\mu}_j F\left(x, \overline{x}; \varsigma_j\right) + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ \|x - \overline{x}\|^m, \quad \forall \varsigma_j \in \partial h_j(\overline{x}), \quad (12)$$

$$\sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \geq \sum_{j \in J^-(\overline{x})} \left(-\overline{\mu}_j\right) F\left(x, \overline{x}; -\varsigma_j\right)$$
$$- \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \|x - \overline{x}\|^m, \quad \forall \varsigma_j \in \partial h_j(\overline{x}). \quad (13)$$

Since the functional $F : X \times X \times R^n \to R$ is sublinear (with respect to the third component), by (11)–(13), we have, for all $x \in X$, respectively,

$$\sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(x) - \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(\overline{x}) \geq F\left(x, \overline{x}; \sum_{i \in I(\overline{x})} \overline{\lambda}_i \zeta_i\right)$$
$$+ \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} \|x - \overline{x}\|^m, \quad \forall \zeta_i \in \partial g_i(\overline{x}), \quad (14)$$

$$\sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \geq F\left(x, \overline{x}; \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \varsigma_j\right)$$
$$+ \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ \|x - \overline{x}\|^m, \quad \forall \varsigma_j \in \partial h_j(\overline{x}), \quad (15)$$

$$\sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \geq F\left(x, \overline{x}; \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \varsigma_j\right)$$
$$- \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \|x - \overline{x}\|^m, \quad \forall \varsigma_j \in \partial h_j(\overline{x}). \quad (16)$$

Adding both sides of (7) and (14)–(16), we get

$$f(x) - f(\overline{x}) + \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(x) - \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(\overline{x}) + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(x)$$
$$- \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(\overline{x}) + \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \geq F(x, \overline{x}; \xi)$$
$$+ F\left(x, \overline{x}; \sum_{i \in I(\overline{x})} \overline{\lambda}_i \zeta_i\right) + F\left(x, \overline{x}; \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \varsigma_j\right) + F\left(x, \overline{x}; \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \varsigma_j\right)$$
$$+ \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^-\right) \|x - \overline{x}\|^m.$$

Since the functional $F : X \times X \times R^n \to R$ is sublinear (with respect to the third component), we have, for all $x \in X$,

$$f(x) - f(\overline{x}) + \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(x) - \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(\overline{x}) + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(x)$$
$$- \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(\overline{x}) + \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x})$$
$$\geq F\left(x, \overline{x}; \xi + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \zeta_i + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \varsigma_j + \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \varsigma_j\right)$$
$$+ \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^-\right) \|x - \overline{x}\|^m.$$

Hence, by the Generalized Karush–Kuhn–Tucker condition (2), it follows that the inequality

$$f(x) - f(\overline{x}) + \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(x) - \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(\overline{x}) + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(x)$$
$$- \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(\overline{x}) + \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \geq F(x, \overline{x}; 0)$$
$$+ \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \right) \|x - \overline{x}\|^m$$

holds for all $x \in X$. Thus, (1) gives

$$f(x) - f(\overline{x}) + \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(x) - \sum_{i \in I(\overline{x})} \overline{\lambda}_i g_i(\overline{x}) + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(x)$$
$$- \sum_{j \in J^+(\overline{x})} \overline{\mu}_j h_j(\overline{x}) + \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x})$$
$$\geq \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \right) \|x - \overline{x}\|^m.$$

By the Generalized Karush–Kuhn–Tucker condition (3) and taking the Lagrange multipliers equal to 0, we have, for all $x \in X$,

$$f(x) + \sum_{i=1}^{p} \overline{\lambda}_i g_i(x) + \sum_{j=1}^{s} \overline{\mu}_j h_j(x) \geq f(\overline{x}) + \sum_{j=1}^{s} \overline{\mu}_j h_j(\overline{x})$$
$$+ \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \right) \|x - \overline{x}\|^m.$$

Using $\overline{x} \in D$ together with (6), we obtain that the inequality

$$f(x) + \sum_{i=1}^{p} \overline{\lambda}_i g_i^+(x) + \sum_{j=1}^{s} \left| \overline{\mu}_j h_j(x) \right| \geq f(\overline{x}) + \sum_{i=1}^{p} \overline{\lambda}_i g_i^+(\overline{x})$$
$$+ \sum_{j=1}^{s} \left| \overline{\mu}_j h_j(\overline{x}) \right| + \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ \right.$$
$$\left. - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \right) \|x - \overline{x}\|^m$$

holds for all $x \in X$. By assumption, the penalty parameter c is assumed to satisfy the condition $c \geq \max \{\overline{\lambda}_i, i \in I, |\overline{\mu}_j|, j \in J\}$). Since $\overline{x} \in D$, the inequality above gives for all $x \in X$,

$$f(x) + c \left[\sum_{i=1}^{p} g_i^+(x) + \sum_{j=1}^{s} |h_j(x)| \right]$$
$$\geq f(\overline{x}) + c \left[\sum_{i=1}^{p} g_i^+(\overline{x}) + \sum_{j=1}^{s} |h_j(\overline{x})| \right]$$
$$+ \left(\rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^- \right) \|x - \overline{x}\|^m.$$

By definition of the objective function in the penalized optimization problem (P(c)) with the exact l_1 penalty function, it follows that the inequality

$$P(x, c) \geq P(\overline{x}, c) + \beta \|x - \overline{x}\|^m,$$ (17)

holds for all $x \in X$, where

$$\beta = \rho_f + \sum_{i \in I(\overline{x})} \overline{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\overline{x})} \overline{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\overline{x})} \overline{\mu}_j \rho_{h_j}^-.$$ (18)

Since $\beta > 0$, by (17) and Definition 4, we conclude that $\overline{x}$ is a strict global minimizer of order m in the penalized optimization problem (P(c)) with the absolute value penalty function. Thus, the conclusion of theorem is established.   □

The result below follows directly from Theorem 8.

Corollary 9 *Let $\overline{x}$ be a strict global minimizer of order m for the considered constrained optimization problem (P) and the suitable constraint qualification be satisfied at $\overline{x}$. Furthermore, assume that all assumptions of Theorem 8 are fulfilled. Then $\overline{x}$ is also a strict global minimizer of order m in the penalized optimization problem (P(c)) with the absolute value penalty function.*

In the example below, we consider a nonconvex nonsmooth optimization problem with (F, ρ)-convex functions of order 3, for which assumption e) of Theorem 8 is not satisfied. We show that a strict global minimizer of order 3 in the considered nonconvex nonsmooth optimization problem is not a strict global minimizer of order 3 in its penalized optimization problem (P(c)) with the absolute value penalty function.

Example 10 Consider the following nonsmooth constrained optimization problem

$$f(x) = x^3 + \frac{1}{2}|x| \rightarrow \min$$
$$g(x) = -x \leq 0.$$ (P1)

Note that $D = \{x \in R : x \geq 0\}$ and $\overline{x} = 0$ is a strict global minimizer of order 3 in the considered nonsmooth optimization problem. It can be shown, by Definition 3, that the objective function f is locally Lipschitz (F, ρ_f)-convex of order 3 at $\overline{x}$ on R and the constraint function g is locally Lipschitz (F, ρ_g)-convex of order 3 at $\overline{x}$ on R, where a functional $F : R \times R \times R \rightarrow R$ is defined by $F(x, \overline{x}; \alpha) = \alpha |x|$ and $\rho_f = -1$, $\rho_g = 0$. Note that assumption e) in Theorem 8 is not satisfied in this case. However, we use the exact l_1 penalty function method to solve the considered nonsmooth optimization problem (P1). Therefore, we construct the following unconstrained optimization problem

$$P(x, c) = x^3 + \frac{1}{2}|x| + c \max\{0, -x\} \rightarrow \min. \quad (P1(c))$$

Further, it is not difficult to see that, for any $c > 0$, $\overline{x} = 0$ is not a strict global minimizer of order 3 in the above unconstrained optimization problem generated in the exact l_1 penalty method (more exactly, it is not a strict global minimizer of any

order). This follows from the fact that the downward order of growth of f exceeds the upward of growth of g at $\overline{x}$ when moving from $\overline{x}$ towards smaller values. Indeed, note that $\inf_{x \in R} P(x, c) \to -\infty$ when $x \to -\infty$ for any $c > 0$. As it follows from this example, although the functions constituting the original optimization problem are locally Lipschitz (F, ρ)-convex of order 3 at $\overline{x}$ on R, then a feasible point $\overline{x} = 0$, being a strict global minimizer of order 3 in the given constrained optimization problem (P1), is not a strict global minimizer of order 3 in its associated penalized optimization problem generated in the exact l_1 penalty function method. This is a consequence of the fact that the assumption (e) in Theorem 8 is not satisfied in the considered case. Hence, as it follows even from this example, assumption (e) in Theorem 8 is essential to prove the result in this theorem and it can not be omitted.

Now, under some stronger assumptions, we prove the converse result.

Theorem 11 *Let the point $\overline{x}$ be a strict global minimizer of order m for the penalized optimization problem (P(c)) with the absolute value penalty function. Also, let $\widetilde{x}$ be any Karush–Kuhn–Tucker point of the original mathematical programming problem (P) and the Generalized Karush–Kuhn–Tucker necessary optimality conditions be satisfied at $\widetilde{x}$ with the Lagrange multipliers $\widetilde{\lambda}_i, i = 1, \ldots, p, \widetilde{\mu}_j, j = 1, \ldots, s,$ associated with the constraints g_i and h_j, respectively. Furthermore, assume that:*

(a) *the objective function f is locally Lipschitz $\left(F, \rho_f\right)$-convex of order m at $\widetilde{x}$ on X,*
(b) *the inequality constraints $g_i, i \in I(\widetilde{x})$, are locally Lipschitz $\left(F, \rho_{g_i}\right)$-convex of order m at $\widetilde{x}$ on X,*
(c) *the equality constraints $h_j, j \in J^+(\widetilde{x})$, are locally Lipschitz $\left(F, \rho_{h_j}^+\right)$-convex of order m at $\widetilde{x}$ on X,*
(d) *the functions $-h_j, j \in J^-(\widetilde{x})$, are locally Lipschitz $\left(F, \rho_{h_j}^-\right)$-convex of order m at $\widetilde{x}$ on X,*
(e) $\rho_f + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^- > 0,$
(f) *the set D of all feasible solutions in the constrained optimization problem (P) is compact.*

If the penalty parameter c is sufficiently large (namely, it is sufficient that c satisfies the condition $c > \max\left\{\widetilde{\lambda}_i, i \in I, \left|\widetilde{\mu}_j\right|, j \in J\right\}$), then $\overline{x}$ is also a strict global minimizer of order m for problem (P).

Proof Assume that $\overline{x}$ is a global minimizer of order m for the penalized optimization problem (P(c)) with the absolute value penalty function. Then, by definition of the penalized problem (P(c)) and Definition 4, the following inequality

$$f(x) + c \left(\sum_{i=1}^{p} g_i^+(x) + \sum_{j=1}^{s} \left|h_j(x)\right|\right)$$
$$\geq f(\overline{x}) + c \left(\sum_{i=1}^{p} g_i^+(\overline{x}) + \sum_{j=1}^{s} \left|h_j(\overline{x})\right|\right) + \beta \|x - \overline{x}\|^m$$

holds for all $x \in X$. Using (6) together with the definition of the absolute value, we get that the following inequality

$$f(x) + c \left(\sum_{i=1}^{p} g_i^+(x) + \sum_{j=1}^{s} \left| h_j(x) \right| \right) \geq f(\overline{x}) + \beta \left\| x - \overline{x} \right\|^m$$

holds for all $x \in X$. Therefore, it is also satisfied for all $x \in D$. Hence, by (6), the inequality

$$f(x) \geq f(\overline{x}) + \beta \left\| x - \overline{x} \right\|^m \tag{19}$$

holds for all $x \in D$. The inequality above means that values of the objective function f are bounded below on the set D. Since f is a continuous function bounded below on the compact set D, therefore, by Weierstrass' theorem, f admits its minimum $\widetilde{x}$ on D.

Now, we prove that $\overline{x}$ is a strict global minimizer of order m in the constrained optimization problem (P). First, we show that $\overline{x}$ is feasible in the given extremum problem (P). By means of contradiction, we suppose that $\overline{x}$ is not feasible in problem (P). As we have established above, the considered constrained optimization problem (P) has an optimal solution $\widetilde{x}$. Since the suitable constraint qualification is satisfied at $\widetilde{x}$, there exist the Lagrange multipliers $\widetilde{\lambda} \in R^p$ and $\widetilde{\mu} \in R^s$ such that the Generalized Karush–Kuhn–Tucker necessary optimality conditions (2)–(4) are satisfied at $\widetilde{x}$. By hypotheses (a)–(d) and Definition 3, the inequalities

$$f(\overline{x}) - f(\widetilde{x}) \geq F(\overline{x}, \widetilde{x}; \xi) + \rho_f \left\| \overline{x} - \widetilde{x} \right\|^m, \quad \forall \xi \in \partial f(\widetilde{x}), \tag{20}$$

$$g_i(\overline{x}) - g_i(\widetilde{x}) \geq F(\overline{x}, \widetilde{x}; \zeta_i) + \rho_{g_i} \left\| \overline{x} - \widetilde{x} \right\|^m, \quad \forall \zeta_i \in \partial g_i(\widetilde{x}), \quad i \in I(\widetilde{x}), \tag{21}$$

$$h_j(\overline{x}) - h_j(\widetilde{x}) \geq F(\overline{x}, \widetilde{x}; \varsigma_j) + \rho_{h_j}^+ \left\| \overline{x} - \widetilde{x} \right\|^m, \quad \forall \varsigma_j \in \partial h_j(\widetilde{x}), \quad j \in J^+(\widetilde{x}), \tag{22}$$

$$- h_j(\overline{x}) + h_j(\widetilde{x}) \geq F(\overline{x}, \widetilde{x}; -\varsigma_j) + \rho_{h_j}^- \left\| \overline{x} - \widetilde{x} \right\|^m, \quad \forall \varsigma_j \in \partial h_j(\widetilde{x}), \quad j \in J^-(\widetilde{x}) \tag{23}$$

hold. Multiplying inequalities (21)–(23) by the corresponding Lagrange multipliers, respectively, by the sublinearity of the functional F (with respect to the third component), it follows that

$$f(\overline{x}) - f(\widetilde{x}) + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i g_i(\overline{x}) - \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i g_i(\widetilde{x}) + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j h_j(\overline{x})$$

$$- \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j h_j(\widetilde{x}) + \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j h_j(\overline{x}) - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j h_j(\widetilde{x})$$

$$\geq F\left(\overline{x}, \widetilde{x}; \xi + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \zeta_i + \sum_{j \in J(\widetilde{x})} \overline{\mu}_j \varsigma_j \right)$$

$$+ \left(\rho_f + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^- \right) \left\| \overline{x} - \widetilde{x} \right\|^m.$$

Since the Generalized Karush–Kuhn–Tucker necessary optimality condition (2) is satisfied at $\widetilde{x}$, by (1), the above inequality implies

$$
f(\overline{x}) - f(\widetilde{x}) + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i g_i(\overline{x}) - \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i g_i(\widetilde{x}) + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j h_j(\overline{x})
$$
$$
- \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j h_j(\widetilde{x}) + \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j h_j(\overline{x}) - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j h_j(\widetilde{x})
$$
$$
\geq \left(\rho_f + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^- \right) \|\overline{x} - \widetilde{x}\|^m .
$$
(24)

Since the Generalized Karush–Kuhn–Tucker necessary optimality condition (3) is satisfied at $\widetilde{x}$, we have

$$
\sum_{i=1}^{p} \widetilde{\lambda}_i g_i(\widetilde{x}) = 0.
$$
(25)

Using the feasibility of $\widetilde{x}$ in the original optimization problem (P) together with (6) and (25), we obtain

$$
\sum_{i=1}^{p} \widetilde{\lambda}_i g_i(\widetilde{x}) = \sum_{i=1}^{p} \widetilde{\lambda}_i g_i^+(\widetilde{x}), \quad \sum_{j=1}^{s} \widetilde{\mu}_j h_j(\widetilde{x}) = \sum_{j=1}^{s} \left| \widetilde{\mu}_j h_j(\widetilde{x}) \right| .
$$
(26)

Since $\overline{x}$ is not feasible in problem (P), (6) implies

$$
\sum_{i=1}^{p} \widetilde{\lambda}_i g_i^+(\overline{x}) \geq \sum_{i=1}^{p} \widetilde{\lambda}_i g_i(\overline{x}), \quad \sum_{j=1}^{s} \widetilde{\mu}_j \left| h_j(\widetilde{x}) \right| \geq \sum_{j=1}^{s} \widetilde{\mu}_j h_j(\overline{x}).
$$
(27)

Combining (24)–(27), we get

$$
f(\overline{x}) + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i g_i^+(\overline{x}) + \sum_{j \in J(\widetilde{x})} \left| \widetilde{\mu}_j \right| \left| h_j(\overline{x}) \right|
$$
$$
\geq f(\widetilde{x}) + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i g_i^+(\widetilde{x}) + \sum_{j \in J(\widetilde{x})} \left| \widetilde{\mu}_j \right| \left| h_j(\widetilde{x}) \right|
$$
$$
+ \left(\rho_f + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^- \right) \|\overline{x} - \widetilde{x}\|^m .
$$
(28)

Since $c > \max \left\{ \widetilde{\lambda}_i, i \in I, \left| \widetilde{\mu}_j \right|, j \in J \right\}, \widetilde{x} \in D$ gives

$$
f(\overline{x}) + c \left(\sum_{i \in I(\widetilde{x})} g_i^+(\overline{x}) + \sum_{j \in J(\widetilde{x})} \left| h_j(\overline{x}) \right| \right)
$$
$$
> f(\widetilde{x}) + c \left(\sum_{i \in I(\widetilde{x})} g_i^+(\widetilde{x}) + \sum_{j \in J(\widetilde{x})} \left| h_j(\widetilde{x}) \right| \right)
$$
$$
+ \left(\rho_f + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^- \right) \|\overline{x} - \widetilde{x}\|^m .
$$
(29)

We denote $\widetilde{\beta} = \rho_f + \sum_{i \in I(\widetilde{x})} \widetilde{\lambda}_i \rho_{g_i} + \sum_{j \in J^+(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^+ - \sum_{j \in J^-(\widetilde{x})} \widetilde{\mu}_j \rho_{h_j}^-$. By assumption (e), it follows that $\widetilde{\beta} > 0$. Thus, by the definition of the objective function in problem (P(c)), the following inequality

$$P(\overline{x}, c) > P(\widetilde{x}, c) + \widetilde{\beta} \|\overline{x} - \widetilde{x}\|^m \tag{30}$$

holds. Hence,

$$P(\overline{x}, c) - \widetilde{\beta} \|\overline{x} - \widetilde{x}\|^m > P(\widetilde{x}, c).$$

Since $\widetilde{\beta} > 0$, it follows that the following inequality

$$P(\overline{x}, c) + \beta \|\overline{x} - \widetilde{x}\|^m > P(\widetilde{x}, c)$$

holds for every $\beta > 0$. But the inequality above contradicts the assumption that $\overline{x}$ is a global minimizer of order m in the penalized optimization problem (P(c)) with the exact absolute penalty function. Thus, we have established that $\overline{x}$ is feasible in the constrained optimization problem (P). Then, the conclusion of theorem, that is, the result that $\overline{x}$ is a global minimizer of order m in the given constrained optimization problem (P) follows directly from (19).   □

The following result follows directly from Corollary 9 and Theorem 11:

Corollary 12 *Let all hypotheses of Corollary 9 and of Theorem 11 be fulfilled. Then, the set of strict global minimizers of order m for the given constrained optimization problem (P) and the set of strict global minimizers of order m for its associated penalized optimization problem (P(c)) with the exact absolute value penalty function coincide.*

Now, we illustrate the results established in the paper by the help of nonconvex nonsmooth optimization problems with locally Lipschitz (F, ρ)-convex of order m functions.

Example 13 Consider the following nonsmooth constrained optimization problem

$$\begin{aligned} f(x) &= \arctan(x^2) + |x| + 1 \to \min \\ g(x) &= \ln(x^2 - |x| + 1) \le 0. \end{aligned} \tag{P2}$$

Note that $D = \{x \in R : -1 \le x \le 1\}$ and it is not difficult to show by Definition 4 that $\overline{x} = 0$ is a strict global minimizer of order 1 for the considered nonsmooth optimization problem (P2). Therefore, the Generalized Karush–Kuhn–Tucker necessary optimality conditions (2)–(4) are fulfilled at this point with the Lagrange multiplier $\overline{\lambda}$ satisfying the following condition: $0 \in \partial f(\overline{x}) + \overline{\lambda} \partial g(\overline{x})$, where $\partial f(\overline{x}) = [-1, 1]$ and $\partial g(\overline{x}) = [-1, 1]$. If we set $F(x, \overline{x}; \vartheta) = \frac{1}{8} |x - \overline{x}| \vartheta$, $\rho_f = \frac{7}{8}$, $\rho_g = -\frac{9}{8}$, then, by Definition 3, the objective function f is (F, ρ_f)-convex of order 1 at $\overline{x}$ on R and the constraint function g is locally Lipschitz (F, ρ_g)-convex of order 1 at $\overline{x}$ on R, and, moreover, the condition $\rho_f + \overline{\lambda} \rho_g \ge 0$ is satisfied. Since we solve problem (P2) by using the exact

l_1 penalty function method, we construct the following unconstrained optimization problem

$$P(x, c) = \arctan\left(x^2\right) + |x| + 1 + c \max\left\{0, \ln\left(x^2 - |x| + 1\right)\right\}. \quad \text{(P2(c))}$$

Then, by Theorem 8, it is follows that, for any penalty parameter c satisfying $c > \bar{\lambda}$, the point $\bar{x} = 0$ is also a strict global minimizer of order 1 for the penalized optimization problem (P2(c)) with the absolute value penalty function given above. Furthermore, since both the objective function f is (F, ρ_f)-convex of order 1 and the constraint function g is locally Lipschitz (F, ρ_g)-convex of order 1 at any Karush–Kuhn–Tucker point of problem (P2) on R, by Theorem 11, $\bar{x} = 0$, being a strict global minimizer of order 1 for the penalized optimization problem (P2(c)) for all penalty parameters no less than 0, is also a strict global minimizer of order 1 for the considered optimization problem (P2) Thus, in fact, there is the equivalence between the sets of strict global minimizers of order 1 for optimization problems (P2) and (P2(c)).

In the next example, we consider a nonconvex nonsmooth optimization problem in which not all involved functions are (F, ρ)-convex functions of the same order with respect to any sublinear functional $F : R \times R \times R \to R$. We show in this case that there is no equivalence between the considered nonsmooth optimization problem and its associated penalized optimization problem with the absolute value penalty function in the sense discussed in the paper. In other words, the sets of strict global minimizers of the same order are not the same in these two optimization problems.

Example 14 Consider the following nonsmooth constrained optimization problem

$$\begin{aligned} f(x) &= x^3 + |x| \to \min \\ g(x) &= x^2 + x \le 0. \end{aligned} \quad \text{(P3)}$$

Note that $D = \{x \in R : -1 \le x \le 0\}$ and $\bar{x} = 0$ is a strict global minimizer of order 3 for the considered optimization problem. It can be proved that the objective function f is not (F, ρ_f)-convex of order m on R with respect to any sublinear (with respect to the third component) functional $F : R \times R \times R \to R$ (see [28]). In other words, there do not exist real numbers ρ_f, ρ_g (with $\rho_f > 0$ or $\rho_g > 0$) and the same sublinear functional $F : R \times R \times R \to R$ with respect to which the objective function f is (F, ρ_f)-convex on R and the constraint function g is (F, ρ_g)-convex on R. Since we use the exact l_1 penalty method to solve the considered optimization problem (P3), we construct the following unconstrained optimization problem:

$$P(x, c) = x^3 + |x| + c \max\left\{0, x^2 + x\right\} \to \min. \quad \text{(P3(c))}$$

Then note that $\bar{x} = 0$ is not a strict global minimizer of order 3 for its associated penalized optimization problem (P3(c)) with the exact l_1 penalty function for any penalty parameter c, and the more so, for any penalty parameter c satisfying the condition $c > \bar{\lambda}_1$. Indeed, it is not difficult to see that the downward order of growth of f exceeds the upward of growth of g at $\bar{x}$ when moving from $\bar{x}$ towards smaller

values. In fact, note that $\inf_{x \in R} P_c(x) \to -\infty$ when $x \to -\infty$ for any $c > 0$. This follows from the fact that the functions constituting the constrained optimization problem (P3) are not locally Lipschitz (F, ρ)-convex of order m with respect to the same (sublinear with respect to the third component) functional $F : R \times R \times R \to R$.

Now, we give an example of such a nonsmooth optimization problem in which the objective function is coercive [23], but it is not (F, ρ_f)-convex of order m on R with respect to any sublinear (with respect to the third component) functional $F : R \times R \times R \to R$ (see [28]). We show also in such a case that there is no the equivalence between the considered nonsmooth optimization problem and its associated penalized optimization problem with the absolute value penalty function in the sense discussed in the paper. In other words, the corecivity property of the objective function is not sufficient to ensure that the sets of strict global minimizers of the same order are the same in these two optimization problems.

Example 15 Consider the following nonsmooth constrained optimization problem

$$f(x) = 2|x+1| - 2|x| + |x-1| \to \min \\ g(x) = -x \le 0. \tag{P4}$$

Note that $D = \{x \in R : x \ge 0\}$ and it can be shown, by Definition 4, that $\overline{x} = 1$ is a strict global minimizer of order 1 in the considered nonsmooth optimization problem (P4). Further, note that the objective function f is coercive (see [23]), but it is not (F, ρ_f)-convex of order 1 on R with respect to any sublinear (with respect to the third component) functional $F : R \times R \times R \to R$ (see [28]). Since we use the exact l_1 penalty method to solve the considered optimization problem (P4), we construct the following unconstrained optimization problem:

$$P(x, c) = 2|x+1| - 2|x| + |x-1| + c \max\{0, -x\} \to \min. \tag{P4(c)}$$

Then note that $\overline{x} = 1$ is not a strict global minimizer of order 1 in its associated penalized optimization problem (P4(c)) with the exact l_1 penalty function for any penalty parameter c satisfying the condition $c > \overline{\lambda}_1 = 1$. Further, it can be shown, by Definition 4, that $\overline{x} = -1$ is a strict global minimizer of order 1 for penalty parameters $c \in [1, 2)$. In other words, the treshold of penalty parameters above which there is the equivalence between the sets of strict global minimizers of order 1 in problems (P4) and (P4(c)) is not equal to the largest Lagrange multiplier in the absolute value. This follows from the fact that the the objective function in the constrained optimization problem (P4) is not locally Lipschitz (F, ρ)-convex of order 1 with respect to any sublinear with respect to the third component functional $F : R \times R \times R \to R$.

References

1. Cromme, L.: Strong uniqueness: a far-reaching criterion for the convergence of iterative procedures. Numerische Mathematik **29**, 179–193 (1978)
2. Klatte, D.: Stable local minimizers in semi-infinite optimization: regularity and second-order conditions. J. Comput. Appl. Math. **56**, 137–157 (1994)
3. Studniarski, M.: Characterizations of strict local minima for some nonlinear programming problems. Nonlinear Anal. **30**, 5363–5367 (1997). **(Proc. 2nd World Congress of Nonlinear Analysis)**
4. Auslender, A.: Stability in mathematical programming with nondifferentiable data. SIAM J. Control Optim. **22**, 239–254 (1984)
5. Studniarski, M.: Sufficient conditions for the stability of local minimum points in nonsmooth optimization. Optimization **20**, 27–35 (1989)
6. Ward, D.E.: Characterizations of strict local minima and necessary conditions for weak sharp minima. J. Optim. Theory Appl. **80**, 551–571 (1994)
7. Antczak, T.: The l_1 penalty function method for nonconvex differentiable optimization problems with inequality constraints. Asia Pacific J. Op. Res. **27**, 1–18 (2010)
8. Antczak, T.: The l_1 exact G-penalty function method and G-invex mathematical programming problems. Math. Comput. Model. **54**, 1966–1978 (2011)
9. Antczak, T.: The exact l_1 penalty function metod for nonsmooth invex optimization problems. Hömberg, D., Tröltzsch, F. (eds.) System Modelling and Optimization, 25th IFIP TC Conference, CSMO 2011, AITC 391, Berlin, Germany, September 2011, Springer (2013), pp. 461–471
10. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms. Wiley, New York (1991)
11. Bertsekas, D.P.: Constrained Optimization and Lagrange Multiplier Methods. Academic Press, Inc., USA (1982)
12. Bertsekas, D.P., Koksal, A.E.: Enhanced optimality conditions and exact penalty functions. In: Proceedings of Allerton Conference, September 2000
13. Charalambous, C.: A lower bound for the controlling parameters of the exact penalty functions. Math. Program. **15**, 278–290 (1978)
14. Di Pillo, G., Grippo, L.: Exact penalty functions in constrained optimization. SIAM J. Control Optim. **27**, 1333–1360 (1989)
15. Di Pillo, G., Lucidi, S., Rinaldi, F.: An approach to constrained global optimization based on exact penalty functions. J. Glob. Optim. **54**, 251–260 (2012)
16. Dolecki, S., Rolewicz, S.: Exact penalty for local minima. SIAM J. Control Optim. **17**, 111–124 (1979)
17. Fletcher, R.: An exact penalty function for nonlinear programming with inequalities. Math. Program. **5**, 129–150 (1973)
18. Han, S.P., Mangasarian, O.L.: Exact penalty functions in nonlinear programming. Math. Program. **17**, 251–269 (1979)
19. Janesch, S.M.H.: Exact penalty functions a lower bound to the penalty parameter. Int. Math. Forum **2**, 75–86 (2007)
20. Luenberger, D.: Control problem with kinds. IEEE Trans. Autom. Control **15**, 570–574 (1970)
21. Mandal, P., Nahak, C.: The l_1 exact exponential penalty function method with (p, r)-ρ-(η, θ)-invexity. J. Adv. Math. Stud. **5**, 127–148 (2012)
22. Mangasarian, O.L.: Sufficiency of exact penalty minimization. SIAM J. Control Optim. **23**, 30–37 (1985)
23. Peressini, A.L., Sullivan, F.E., Uhl Jr, J.J.: The Mathematics of Nonlinear Programming. Springer, New York (1988)
24. Rosenberg, E.: Exact penalty functions and stability in locally Lipschitz programming. Math. Program. **30**, 340–356 (1984)
25. Zangwill, W.I.: Nonlinear programming via penalty functions. Manag. Sci. **13**, 344–358 (1967)
26. Zaslavski, A.J.: A sufficient condition for exact penalty in constrained optimization. SIAM J. Optim. **16**, 250–262 (2005)
27. Clarke, F.H.: Optimization and Nonsmooth Analysis. A Wiley-Interscience Publication, John Wiley & Sons Inc., New York (1983)
28. Caprari, E.: ρ-Invex functions and (F, ρ)-convex functions: properties and equivalences. Optimization **52**, 65–74 (2003)

29. Hiriart-Urruty, J.B.: On optimality conditions in nondifferentiable programming. Math. Program. **14**, 73–86 (1978)
30. Hiriart-Urruty, J.B.: Refinements of necessary optimality conditions in nondifferentiable programming. Appl. Math. Optim. **5**, 63–82 (1979)