The Exact l_1 Penalty Function Method for Constrained Nonsmooth Invex Optimization Problems

Tadeusz Antczak

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

Abstract. The exactness of the penalization for the exact l_1 penalty function method used for solving nonsmooth constrained optimization problems with both inequality and equality constraints is considered. Thus, the equivalence between the sets of optimal solutions in the nonsmooth constrained optimization problem and its associated penalized optimization problem with the exact l_1 penalty function is established under locally Lipschitz invexity assumptions imposed on the involved functions.

Keywords: exact l_1 penalty function method, absolute value penalty function, penalized optimization problem, locally Lipschitz invex function, Generalized Karush-Kuhn-Tucker optimality conditions.

1 Introduction

Considerable attention has been given in recent years to devising methods for solving nonlinear programming problems via unconstrained minimization techniques. One class of methods which has emerged as very promising is the class of exact penalty function methods. Methods using exact penalty function transform a constrained extremum problem into a single unconstrained optimization problem. The constraints are placed into the objective function via a penalty parameter c in a way that penalizes any violation of the constraints.

One important property that distinguishes exact penalty functions is the exactness of the penalization. The concept of exact penalization is sometimes ambiguous, or at least varies from author to author. One of the definitions of the exactness of the penalization is the following: there is an appropriate penalty parameter choice such that a single unconstrained minimization of the penalty function yields a solution of the constrained optimization problem.

Nondifferentiable exact penalty functions were introduced for the first time by Eremin [6] and Zangwill [17]. In almost all of the introduced penalized approaches the notion of convexity plays a dominant role. In 1970, Luenberger [13] showed that, under convex assumptions, there is a lower bound for a penalty parameter c, equal to the largest Lagrange multiplier in absolute value, associated to one of the constraints of the nonlinear constrained optimization problem.

D. Hömberg and F. Tröltzsch (Eds.): CSMO 2011, IFIP AICT 391, pp. 461–470, 2013.

Later, Charalambous [3] generalized the result of Luenberger for the absolute value penalty function, assuming the second-order sufficient conditions. Under the assumptions that the minimization problem is solvable and that it satisfies the relaxed Slater constraint qualification, Mangasarian [14] characterized solutions of the convex optimization problem in terms of minimizers of the exact penalty function for a single value of the penalty parameter exceeding some treshold. Bazaraa et al. [2] also used the exact l_1 penalty function method to solve nonlinear convex optimization problems with both inequality and equality constraints. They assumed that the objective function and the inequality constraints are convex and the equality constraints are affine functions to prove that a Karush-Kuhn-Tucker point in the original optimization problem is a minimizer of the exact l_1 penalty function in the associated penalized optimization problem with sufficiently large value of a penalty parameter. In the mentioned above works, the lower bound of the penalty parameter above which, for all penalty parameters, any optimal solution of the original nonlinear optimization problem is also a minimizer of the penalized problem has been given for differentiable optimization problems involving convex functions. However, from the practical point of view, the converse result is also important.

In recent years, some numerous generalizations of convex functions have been derived which proved to be useful for extending optimality conditions and some classical duality results, previously restricted to convex programs, to larger classes of nonconvex optimization problems. One of them is the invexity notion introduced by Hanson [10] for differentiable scalar functions and later generalized from different points of view, also in the case of nondifferentiable functions (see [1], [5], [8], [11], [12], [16], [18], and others).

Now, we show that there is the equivalence between the set of optimal solutions in a nondifferentiable nonconvex optimization problem and the set of minimizers in its associated exact penalized problem with the absolute value penalty function. It turns out that this property is not true only for (differentiable) convex optimization problems, but it still holds for nonlinear optimization problems involving locally Lipschitz invex functions with respect to the same function η (with the exception of those equality constraint functions for which the associated Lagrange multipliers are negative – these functions should be assumed to be incave with respect to the same function η). The result established here shows that there does exist a lower bound for a penalty parameter c, equal to the largest Lagrange multiplier in absolute value, associated to a Karush-Kuhn-Tucker point in the original nonlinear optimization problem, above which this equivalence holds. Further, in the case when at least one of the functions constituting the nondifferentiable constrained optimization problem is not locally Lipschitz invex and in the case when the objective function is coercive but not invex, then the equivalence in the sense discussed here might not hold between these optimization problems.

2 Preliminaries and Problem Formulation

Throughout this section, X is a nonempty subset of R^n. A real-valued function $f : X \to R$ is said to be locally Lipschitz on X if, for any $x \in X$, there exist a neighborhood U of x and a positive constant $K_x > 0$ such that, for every $y, z \in U$, it holds $|f(y) - f(z)| \leqq K_x \|y - z\|$. The Clarke generalized directional derivative [4] of a locally Lipschitz function $f : X \to R$ at $x \in X$ in the direction $v \in R^n$, denoted $f^0(x; v)$, is given by $f^0(x; v) = \limsup_{\substack{y \to x \\ \lambda \downarrow 0}} \frac{f(y+\lambda v)-f(y)}{\lambda}$.

Definition 1. *The Clarke generalized subgradient [4] of f at $x \in X$, denoted $\partial f(x)$, is defined by $\partial f(x) = \left\{\xi \in R^n : f^0(x; v) \geq \xi^T v \text{ for all } v \in R^n\right\}$.*

The following definition is a generalization of the definition of a class of differentiable convex functions to the case of a class of locally Lipschitz invex functions (see [10]).

Definition 2. *[10] Let a function $f : X \to R$ be a locally Lipschitz function on X and $u \in X$. If there exists a vector-valued function $\eta : X \times X \to R^n$ such that, for each $x \in X$, the inequality $f(x) - f(u) \geq \xi^T \eta(x, u)$ holds for any $\xi \in \partial f(u)$, then f is said to be a locally Lipschitz invex function at u on X with respect to η. If the inequality above is satisfied at any point u, then f is said to be a locally Lipschitz invex function on X with respect to η.*

In order to define an analogous class of Lipschitz incave functions with respect to η, the direction of the inequality in the definition of invex functions should be changed to the opposite one.

Definition 3. *[15] A continuous function $f : R^n \to R$ is said to be coercive if $\lim_{\|x\| \to \infty} f(x) = \infty$.*

Consider the following constrained optimization problem:

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

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 set $X \subset R^n$.

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, we denote a set of active inequality constraints at point $\overline{x} \in X$ by $I(\overline{x}) = \{i \in I : g_i(\overline{x}) = 0\}$.

Theorem 4. *[4], [18] (Generalized Karush-Kuhn-Tucker necessary optimality conditions). Let $\overline{x} \in D$ be an optimal solution in problem (P) and some suitable constraint qualification be satisfied at $\overline{x}$. Then, there exist $\overline{\lambda} \in R^m$, $\overline{\mu} \in R^s$ such that*

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

$$\overline{\lambda}_i g_i(\overline{x}) = 0, \; i \in J, \tag{2}$$

$$\overline{\lambda}_i \in R_+, \quad i \in J. \tag{3}$$

We will assume that a suitable constraint qualification is satisfied at any optimal point in problem (P).

Definition 5. *The point $\overline{x} \in D$ is said to be Karush-Kuhn-Tucker point (a KKT point, for short) if there exist the Lagrange multipliers $\overline{\lambda} \in R^m$, $\overline{\mu} \in R^s$ such that the conditions (1)-(3) are satisfied at $\overline{x}$.*

3 The Exactness of the Exact l_1 Penalty Function Method

The most popular nondifferentiable exact penalty function is the absolute value penalty function also called the exact l_1 penalty function. Its definition, for the considered optimization problem (P), is the following

$$\text{minimize } P(x,c) = f(x) + c\left[\sum_{i\in I} g_i^+(x) + \sum_{j\in J} |h_j(x)|\right], \qquad (\mathrm{P}(c)) \tag{4}$$

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

$$g_i^+(x) = \begin{cases} 0 & \text{if } g_i(x) \leq 0, \\ g_i(x) & \text{if } g_i(x) > 0. \end{cases} \tag{5}$$

The unconstrained optimization problem defined above, we call the penalized optimization problem with the absolute value penalty function.

It is known (see, for example, [2]) that under suitable convexity assumptions and a constraint qualification, there exists a finite value c that will recover an optimal solution in the constrained optimization problem (P) via the minimization of the exact penalty function being the objective function in the exact penalized optimization problem (P(c)). Now, we generalize this result by weakening the convexity assumption imposed on the functions constituting the considered nonsmooth optimization problem (P).

Theorem 6. *Let $\overline{x} \in D$ be a Karush-Kuhn-Tucker point in the constrained optimization problem (P), at which the Generalized Karush-Kuhn-Tucker conditions (1)-(3) are satisfied with the Lagrange multipliers $\overline{\lambda} \in R^m$ 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 that the functions f, g_i, $i \in I$, h_j, $j \in J^+(\overline{x})$, are locally Lipschitz invex at $\overline{x}$ on X with respect to the same function η and the functions h_j, $j \in J^-(\overline{x})$, are locally Lipschitz incave at $\overline{x}$ on X with respect to the same function η. If c is assumed to be sufficiently large (it is sufficient to set $c \geq \max\{\overline{\lambda}_i,\ i \in I,\ |\overline{\mu}_j|,\ j \in J\}$, where $\overline{\lambda}_i$, $i = 1,...,m$, $\overline{\mu}_j$, $j = 1,...,s$, are the Lagrange multipliers associated to the constraints g_i and h_j, respectively), then $\overline{x}$ is also a minimizer of its penalized optimization problem (P(c)) with the absolute value penalty function.*

Proof. By assumption, $\overline{x}$ is a Karush-Kuhn-Tucker point in the constrained optimization problem (P), at which the Generalized Karush-Kuhn-Tucker conditions (1)-(3) are satisfied with the Lagrange multipliers $\overline{\lambda} \in R^m$ and $\overline{\mu} \in R^s$. Since $c \geq \max\left\{\overline{\lambda}_i,\ i \in I,\ \left|\overline{\mu}_j\right|,\ j \in J\right\}$, then, by definition of the objective function in the penalized optimization problem (P(c)), it follows that

$$P\left(x,c\right) = f(x) + c\sum_{i=1}^{m} g_i^+(x) + c\sum_{j=1}^{s}\left|h_j(x)\right| \geq f(x) + \sum_{i=1}^{m}\overline{\lambda}_i g_i^+(x) + \sum_{j=1}^{s}\left|\overline{\mu}_j h_j(x)\right|. \tag{6}$$

Thus, (5) gives

$$f(x) + \sum_{i=1}^{m}\overline{\lambda}_i g_i^+(x) + \sum_{j=1}^{s}\left|\overline{\mu}_j h_j(x)\right| \geq f(x) + \sum_{i=1}^{m}\overline{\lambda}_i g_i(x) + \sum_{j=1}^{s}\overline{\mu}_j h_j(x). \tag{7}$$

By assumption, the inequality constraints g_i, $i \in I$, and the equality constraints h_j, $j \in J^+\left(\overline{x}\right)$, are locally Lipschitz invex at $\overline{x}$ on X and the equality constraints h_j, $j \in J^-\left(\overline{x}\right)$, are locally Lipschitz incave at $\overline{x}$ on X. Hence, by the Generalized Karush-Kuhn-Tucker conditions (2) and (3) together with the feasibility of $\overline{x}$ in problem (P), it follows that the inequality

$$f(x) + \sum_{i=1}^{m}\overline{\lambda}_i g_i(x) + \sum_{j=1}^{s}\overline{\mu}_j h_j(x) \geq f(x) + \sum_{i=1}^{m}\overline{\lambda}_i \zeta_i^T \eta\left(x,\overline{x}\right) + \sum_{j=1}^{s}\overline{\mu}_j \gamma_j^T \eta\left(x,\overline{x}\right) \tag{8}$$

holds for any $\zeta_i \in \partial g_i\left(\overline{x}\right)$, $i = 1,...,m$, and for any $\gamma_j \in \partial h_j\left(\overline{x}\right)$, $j = 1,...,s$. Then, using the Generalized Karush-Kuhn-Tucker condition (2) , we get

$$\begin{aligned} f(x) + \textstyle\sum_{i=1}^{m}\overline{\lambda}_i\left[g_i(\overline{x}) + \zeta_i^T\eta\left(x,\overline{x}\right)\right] + \textstyle\sum_{j=1}^{s}\overline{\mu}_j\left[h_j(\overline{x}) + \gamma_j^T\eta\left(x,\overline{x}\right)\right] \\ = f(x) + \textstyle\sum_{i=1}^{m}\overline{\lambda}_i\zeta_i^T\eta\left(x,\overline{x}\right) + \textstyle\sum_{j=1}^{s}\overline{\mu}_j\gamma_j^T\eta\left(x,\overline{x}\right). \end{aligned} \tag{9}$$

Thus, by the Generalized Karush-Kuhn-Tucker necessary optimality condition (1), it follows that

$$f(x) + \sum_{i=1}^{m}\overline{\lambda}_i\zeta_i^T\eta\left(x,\overline{x}\right) + \sum_{j=1}^{s}\overline{\mu}_j\gamma_j^T\eta\left(x,\overline{x}\right) = f(x) - \xi^T\eta\left(x,\overline{x}\right), \tag{10}$$

where $\xi \in \partial f(\overline{x})$. By assumption, f is locally Lipschitz invex at $\overline{x}$ on X also with respect to the function η. Using Definition 2 together with the feasibility of $\overline{x}$ in problem (P), we get

$$f(x) - \xi^T\eta\left(x,\overline{x}\right) \geq f(\overline{x}) = f(\overline{x}) + c\sum_{i=1}^{m} g_i^+(\overline{x}) + c\sum_{j=1}^{s}\left|h_j(\overline{x})\right| = P\left(\overline{x},c\right). \tag{11}$$

Then, by (6)-(11), we conclude that the inequality $P\left(x,c\right) \geq P\left(\overline{x},c\right)$ holds for all $x \in X$. This means that $\overline{x}$ is a minimizer of the penalized optimization problem (P(c)) with the absolute value penalty function and the proof of theorem is complete. ■

Corollary 7. *Let $\overline{x}$ be an optimal point in the considered optimization problem (P). Furthermore, assume that all hypotheses of Theorem 6 are fulfilled. Then $\overline{x}$ is also a minimizer in the penalized optimization problem (P(c)) with the absolute value penalty function.*

Theorem 8. *Let the point $\overline{x}$ be a minimizer of the penalized optimization problem (P(c)) with the absolute value penalty function. Furthermore, assume that the functions f, g_i, $i \in I$, h_j, $j \in J^+(\widetilde{x})$, are locally Lipschitz invex at $\widetilde{x}$ on X with respect to the same function η, and the functions h_j, $j \in J^-(\widetilde{x})$, are locally Lipschitz incave at $\widetilde{x}$ on X with respect to the same function η, where $\widetilde{x}$ is any Karush-Kuhn-Tucker point in problem (P), at which the Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are satisfied with the Lagrange multipliers $\widetilde{\lambda} \in R^m$ and $\widetilde{\mu} \in R^s$. If the set of all feasible solutions in the constrained optimization problem (P) is compact and the penalty parameter c is sufficiently large (it is sufficient if c satisfies the following condition $c > \max\left\{\widetilde{\lambda}_i,\ i \in I,\ |\widetilde{\mu}_j|,\ j \in J\right\}$), then $\overline{x}$ is also optimal in problem (P).*

Proof. We assume that $\overline{x}$ is a minimizer in the penalized optimization problem (P(c)) with the absolute value penalty function. Then, by the definition of the penalized optimization problem (P(c)) and (5), the following inequalities $f(x) + c\left(\sum_{i=1}^m g_i^+(x) + \sum_{j=1}^s |h_j(x)|\right) \geq f(\overline{x}) + c\left(\sum_{i=1}^m g_i^+(\overline{x}) + \sum_{j=1}^s |h_j(\overline{x})|\right) \geq f(\overline{x})$ hold for all $x \in X$. Thus, for all $x \in D$, the following inequality

$$f(x) \geq f(\overline{x}) \tag{12}$$

holds. The inequality above means that values of the function f are bounded below on the set D of all feasible solutions in the constrained optimization problem (P). 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 also optimal in the considered optimization problem (P). First, we show that $\overline{x}$ is feasible in problem (P). By means of contradiction, suppose that $\overline{x}$ is not feasible in problem (P). As we have established above, the given constrained optimization problem (P) has an optimal solution $\widetilde{x}$. Since a constraint qualification is satisfied at $\widetilde{x}$, then there exist the Lagrange multipliers $\widetilde{\lambda} \in R^m$ and $\widetilde{\mu} \in R^s$ such that the Generalized Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are satisfied at $\widetilde{x}$. By assumption, the functions f, g_i, $i \in I$, h_j, $j \in J^+(\widetilde{x})$, are invex at $\widetilde{x}$ on X with respect to the same function η and the functions h_j, $j \in J^-(\widetilde{x})$, are incave at $\widetilde{x}$ on X with respect to the same function η. Therefore, by Definition 2, respectively, it follows that the inequalities

$$f(\overline{x}) - f(\widetilde{x}) \geq \xi^T \eta(\overline{x}, \widetilde{x}), \tag{13}$$

$$g_i(\overline{x}) - g_i(\widetilde{x}) \geq \zeta_i^T \eta(\overline{x}, \widetilde{x}), \quad i \in I, \tag{14}$$

$$h_j(\overline{x}) - h_j(\widetilde{x}) \geq \gamma_j^T \eta(\overline{x}, \widetilde{x}), \quad j \in J^+(\widetilde{x}), \tag{15}$$

$$h_j(\overline{x}) - h_j(\widetilde{x}) \leq \gamma_j^T \eta(\overline{x}, \widetilde{x}), \quad j \in J^-(\widetilde{x}) \tag{16}$$

hold for each $\xi \in \partial f(\widetilde{x})$, $\zeta_i \in \partial g_i(\widetilde{x})$, $i = 1, ..., m$, and $\gamma_j \in \partial h_j(\widetilde{x})$, $j = 1, ..., s$. Multiplying (14), (15) and (16) by the associated Lagrange multiplier and then adding both sides of the obtained inequalities and both sides of (13), we get

$$\begin{aligned} f(\overline{x}) - f(\widetilde{x}) + \textstyle\sum_{i=1}^m \widetilde{\lambda}_i g_i(\overline{x}) - \sum_{i=1}^m \widetilde{\lambda}_i g_i(\widetilde{x}) + \sum_{j=1}^s \widetilde{\mu}_j h_j(\overline{x}) - \sum_{j=1}^s \widetilde{\mu}_j h_j(\widetilde{x}) \\ \geq \left[\xi^T + \textstyle\sum_{i=1}^m \widetilde{\lambda}_i \zeta_i^T + \sum_{j=1}^s \widetilde{\mu}_j \gamma_j^T\right] \eta(\overline{x}, \widetilde{x}). \end{aligned}$$

Using (5) with the Karush-Kuhn-Tucker necessary optimality conditions (1), (2) and the feasibility of $\widetilde{x}$ in problem (P), we get

$$f(\overline{x}) + \sum_{i=1}^m \widetilde{\lambda}_i g_i^+(\overline{x}) + \sum_{j=1}^s \widetilde{\mu}_j |h_j(\overline{x})| \geq f(\widetilde{x}). \tag{17}$$

By assumption, the penalty parameter c is sufficiently large (it is sufficient that $c > \max\left\{\widetilde{\lambda}_i,\ i \in I,\ |\widetilde{\mu}_j|,\ j \in J\right\}$. Since $\overline{x}$ is assumed to be not feasible in the given optimization problem (P), therefore, at least one of $g_i^+(\overline{x})$ and $|h_j(\overline{x})|$ must be nonzero. Therefore, (17) yields

$$f(\overline{x}) + c\left[\sum_{i=1}^m g_i^+(\overline{x}) + \sum_{j=1}^s |h_j(\overline{x})|\right] > f(\widetilde{x}). \tag{18}$$

Then, by $\widetilde{x} \in D$ and (2), we get

$$f(\overline{x}) + c\left[\sum_{i=1}^m g_i^+(\overline{x}) + \sum_{j=1}^s |h_j(\overline{x})|\right] > f(\widetilde{x}) + c\left[\sum_{i=1}^m g_i^+(\widetilde{x}) + \sum_{j=1}^s |h_j(\widetilde{x})|\right].$$

Then, by the definition of the exact l_1 penalty function (see (4)), 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 in 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 problem (P) follows directly from (12). ■

Corollary 9. *Let the hypotheses of Corollary 7 and Theorem 8 are fulfilled. Then, the set of optimal solutions in the considered extremum problem (P) and the set of minimizers in its associated exact penalized optimization problem (P(c)) with the absolute value penalty function coincide.*

Example 10. Consider the following nonsmooth optimization problem

$$\begin{aligned} f(x) &= \arctan(|x|) \to \min \\ g(x) &= \tfrac{1}{2}\left(e^{|x|-x} - 1\right) \leq 0. \end{aligned} \quad \text{(P1)}$$

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

$$P(x,c) = \arctan(|x|) + c \max\left\{0, \frac{1}{2}\left(e^{|x|-x} - 1\right)\right\} \to \min. \quad \text{(P1}(c))$$

Note that $\overline{x} = 0$ is feasible in problem (P1) and the Generalized Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are fulfilled at $\overline{x}$ 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 $\partial g(\overline{x}) = [-1, 0]$. Further, it can be established by Definition 2 that the objective function f and the constraint function g are locally Lipschitz invex at $\overline{x}$ on R with respect to the same function η defined by $\eta(x, \overline{x}) = \frac{1}{2}(\arctan(|x|) - \arctan(|\overline{x}|))$. Then, by Theorems 6 and 8, it follows that, for any penalty parameter c satisfying $c > \overline{\lambda}$, there is the equivalence between the sets of optimal solutions in optimization problems (P1) and (P1(c)). Further, note that not all functions involved in problem (P1) are differentiable and convex. Therefore, in order to show that the point $\overline{x} = 0$, being optimal in (P1), is also a minimizer in the unconstrained optimization problem (P1(c)), we can not use the conditions for convex smooth optimization problems (see, for instance, Theorem 9.3.1 [2]).

Example 11. *Consider the following nonsmooth constrained optimization problem*

$$\min \ f(x) = \begin{cases} -x + 4 & \text{if} \quad x < -4, \\ \frac{1}{2}x + 10 & \text{if} -4 \leq x < 0, \\ -5x + 10 & \text{if} \ \ 0 \leq x < 2, \\ x - 2 & \text{if} \quad x \geq 2, \end{cases} \qquad \textit{(P2)}$$
$$g(x) = x - \tfrac{1}{4} \leq 0,$$

in which not all functions are locally Lipschitz invex. Note that $D = \left\{x \in R : x \leq \frac{1}{4}\right\}$ *and* $\overline{x} = -4$ *is an optimal solution in the considered optimization problem (P2). Since* $0 \in \partial f(0) = \left[-5, \frac{1}{2}\right]$, *then* $\widetilde{x} = 0$ *is a stationary point of* f. *It is not difficult to show that* $\widetilde{x}$ *is not a global minimizer of* f. *Then the objective function* f *is not locally Lipschitz invex on* R *with respect to any function* η *defined by* $\eta : R \times R \to R$ *(see, for example, [16]). However, we use the exact* l_1 *penalty method to solve the considered optimization problem (P2). Therefore, we construct the following unconstrained optimization problem*

$$P(x,c) = f(x) + c \max\left\{0, x - \frac{1}{4}\right\} \to \min \qquad \textit{(P2(c))}$$

Note that $\overline{x} = -4$, *being an optimal solution in problem (P2), is not a global minimizer in the associated penalized optimization problem (P2(c)) for all values of the penalty parameter c satisfying the condition* $c > \overline{\lambda} = 0$, *(where* $\overline{\lambda}$ *is the Lagrange multiplier associated to the inequality constraint g satisfying the*

Karush-Kuhn-Tucker necessary optimality conditions (1)-(3)). However, for every penalty parameter $c \in (0, \frac{32}{7})$*, the point* $\widehat{x} = 2$ *is a global minimizer in the above penalized optimization problem (P2(c)). Therefore, there is no the equivalence between the sets of optimal solutions in problems (P2) and (P2(c)) for any penalty parameter* c *satisfying the condition* $c > \overline{\lambda}$*. This follows from the fact that not all functions constituting the considered optimization problem (P2) are locally Lipschitz invex on R.*

Remark 12. *Peressini et al. [15] considered differentiable convex optimization problems and solved them by using the exact* l_1 *penalty function method. Under assumption that the objective function in the constrained optimization problem is coercive (see Definition 3), they proved that, for sufficiently large values of the penalty parameter* c*, the constrained optimal solution in (P) is also a minimizer in its associated penalized optimization problem (P(c)) with the exact* l_1 *penalty function. But the finite value of the penalty parameter* c*, above which this result holds, was not given in [15]. Note that the objective function in the optimization problem (P2) considered in Example 11 is coercive. However, for not all values of the penalty parameter* c *satisfying the condition* $c > \overline{\lambda}$*, an optimal solution in the considered optimization problem (P2) yields a minimizer in its associated penalized optimization problem (P(c)) with the exact* l_1 *penalty function. But the result proved in the paper shows that, under invexity assumptions imposed on the functions constituting the constrained nonsmooth optimization problem (P), for every value of the penalty parameter* c *satisfying the condition* $c > \max\left\{\overline{\lambda}_i,\ i \in I,\ \left|\overline{\mu}_j\right|,\ j \in J\right\}$*, the sets of optimal solutions in problems (P) and (P(c)) coincide. Hence, this example shows that in the case when the objective function is coercive but not invex the result established in the paper might not be true for such optimization problems.*

References

[1] Antczak, T.: Lipschitz r-invex functions and nonsmooth programming. Numerical Functional Analysis and Optimization 23, 265–283 (2002)

[2] Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear programming: theory and algorithms. John Wiley and Sons, New York (1991)

[3] Charalambous, C.: A lower bound for the controlling parameters of the exact penalty functions. Mathematical Programming 15, 278–290 (1978)

[4] Clarke, F.H.: Optimization and nonsmooth analysis. A Wiley-Interscience Publication, John Wiley&Sons, Inc. (1983)

[5] Craven, B.D.: Invex functions and constrained local minima. Bulletin of the Australian Mathematical Society 24, 357–366 (1981)

[6] Eremin, I.I.: The penalty method in convex programming. Doklady Akad. Nauk SSSR 143, 748–751 (1967)

[7] Di Pillo, G., Grippo, L.: Exact penalty functions in constrained optimization. SIAM Journal of Control and Optimization 27, 1333–1360 (1989)

[8] Egudo, R.R., Hanson, M.A.: On sufficiency of Kuhn-Tucker conditions in nonsmooth multiobjective programming, FSU Report No. M-888 (1993)

[9] Han, S.P., Mangasarian, O.L.: Exact penalty functions in nonlinear programming. Mathematical Programming 17, 251–269 (1979)
[10] Hanson, M.A.: On sufficiency of the Kuhn-Tucker conditions. Journal of Mathematical Analysis and Applications 80, 545–550 (1981)
[11] Kaul, R.N., Suneja, S.K., Lalitha, C.S.: Generalized nonsmooth invexity. Journal of Information and Optimization Sciences 15, 1–17 (1994)
[12] Kim, M.H., Lee, G.M.: On duality theorems for nonsmooth Lipschitz optimization problems. Journal of Optimization Theory and Applications 110, 669–675 (2001)
[13] Luenberger, D.: Control problem with kinds. IEEE Transaction on Automatic Control 15, 570–574 (1970)
[14] Mangasarian, O.L.: Sufficiency of exact penalty minimization. SIAM Journal of Control and Optimization 23, 30–37 (1985)
[15] Peressini, A.L., Sullivan, F.E., Uhl, Jr., J.J.: The mathematics of nonlinear programming. Springer-Verlag New York Inc. (1988)
[16] Reiland, T.W.: Nonsmooth invexity. Bulletin of the Australian Mathematical Society 42, 437–446 (1990)
[17] Zangwill, W.I.: Nonlinear programming via penalty functions. Management Science 13, 344–358 (1967)
[18] Zhao, F.: On sufficiency of the Kuhn-Tucker conditions in nondifferentiable programming. Bulletin of the Australian Mathematical Society 46, 385–389 (1992)