TAIWANESE JOURNAL OF MATHEMATICS
Vol. 17, No. 2, pp. 559-581, April 2013
DOI: 10.11650/tjm.17.2013.1823
This paper is available online at http://journal.taiwanmathsoc.org.tw

SADDLE POINT CRITERIA AND THE EXACT MINIMAX PENALTY FUNCTION METHOD IN NONCONVEX PROGRAMMING

Tadeusz Antczak

Abstract. A new characterization of the exact minimax penalty function method is presented. The exactness of the penalization for the exact minimax penalty function method is analyzed in the context of saddle point criteria of the Lagrange function in the nonconvex differentiable optimization problem with both inequality and equality constraints. Thus, new conditions for the exactness of the exact minimax penalty function method are established under assumption that the functions constituting considered constrained optimization problem are invex with respect to the same function η (exception with those equality constraints for which the associated Lagrange multipliers are negative - these functions should be assumed to be incave with respect to the same function η). The threshold of the penalty parameter is given such that, for all penalty parameters exceeding this treshold, the equivalence holds between a saddle point of the Lagrange function in the considered constrained extremum problem and a minimizer in its associated penalized optimization problem with the exact minimax penalty function.

1. Introduction

In the paper, we consider the following differentiable optimization problem with both inequality and equality constraints:

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

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

Received March 14, 2012, accepted September 4, 2012.
Communicated by Biagio Ricceri.
2010 *Mathematics Subject Classification*: 49M30, 90C26, 90C30.
Key words and phrases: Exact minimax penalty function method, Minimax penalized optimization problem, Exactness of the exact minimax penalty function, Saddle point, Invex function.

We will write $g := (g_1, ..., g_m) : X \to R^m$ and $h := (h_1, ..., h_s) : X \to R^s$ for convenience.

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 (P).

Further, we denote the set of active constraints at point $\overline{x} \in D$ by

$$I(\overline{x}) = \{i \in I : g_i(\overline{x}) = 0\}.$$

It is well-known (see, for example, [8]) that the following Karush-Kuhn-Tucker conditions are necessary for optimality of a feasible solution $\overline{x}$ in the considered differentiable constrained optimization problem.

Theorem 1. *Let $\overline{x}$ be an optimal solution in problem (P) and a suitable constraint qualification [8] be satisfied at $\overline{x}$. Then there exist the Lagrange multipliers $\overline{\xi} \in R^m$ and $\overline{\mu} \in R^s$ such that*

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

$$\overline{\xi}_i g_i(\overline{x}) = 0,\ \ i \in I, \tag{2}$$

$$\overline{\xi} \geq 0. \tag{3}$$

Definition 2. The point $\overline{x} \in D$ is said to be a Karush-Kuhn-Tucker point in problem (P) if there exist the Lagrange multipliers $\overline{\xi} \in R^m$, $\overline{\mu} \in R^s$ such that the Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are satisfied at this point with these Lagrange multipliers.

In the field of mathematical programming many efforts have been devoted to solve nonlinear constrained optimization problems. It is well known that penalty function methods are popular methods in solving nonlinear constrained optimization problems. A particular subclass of penalty functions are the so-called exact penalty functions. They can be subdivided, in turn, into two main classes: nondifferentiable exact penalty functions and continuously differentiable exact penalty functions. Nondifferentiable exact penalty functions were introduced for the first time by Eremin [17] and Zangwill [32].

In the exact penalty functions methods, the original constrained optimization problem is replaced by an unconstrained problem, in which the objective function is the sum of a certain "merit" function (which reflects the objective function of the original extremum problem) and a penalty term which reflects the constraint set. The merit function is chosen, in general, as the original objective function, while the penalty term is obtained by multiplying a suitable function, which represents the constraints, by a positive parameter c, called the penalty parameter. A given penalty parameter c is called an exact penalty parameter when every solution of the given extremum problem can be found by solving the unconstrained optimization problem with the penalty function associated with c. Further, from theoretical point of view, exact penalty functions are important because of the relationship to the necessary optimality conditions for a minimum in constrained optimization.

In [13], Charalambous introduced a class of nondifferentiable exact penalty functions defined as follows

$$P_p(x,\alpha,\beta,c) = f(x) + c\left(\sum_{i=1}^{m}\left[\alpha_i g_i^+(x)\right]^p + \sum_{j=1}^{s}\left[\beta_j \left|h_j(x)\right|\right]^p\right)^{1/p},$$

where $p \geq 1$, $\alpha_i > 0$, $i = 1, ..., m$, $\beta_j > 0$, $j = 1, ..., s$. For a given constraint $g_i(x) \leq 0$, the function $g_i^+(x)$ is defined by

$$g_i^+(x) = \begin{cases} 0 & \text{if } g_i(x) \leq 0, \\ g_i(x) & \text{if } g_i(x) > 0. \end{cases} \tag{4}$$

It follows by (4) that the function $g_i^+(x)$ is equal to zero for all x that satisfy the constraint $g_i(x) \leq 0$ and that it has a positive value whenever this constraint is violated. Moreover, large violations in the constraint $g_i(x) \leq 0$ result in large values for $g_i^+(x)$. Thus, the function $g_i^+(x)$ has the penalty features relative to the single constraint $g_i(x) \leq 0$.

Later, Han and Magasarian [22] introduced the class of penalty functions defined as follows

$$P_p(x,c) = f(x) + c\left\|g^+(x), h(x)\right\|_p,$$

where $c > 0$, $\|\cdot\|_p$ denotes the l_p norm over R^{m+s} for $1 \leq p \leq \infty$.

For $p = 1$, we get the most known nondifferentiable exact penalty function, called the exact l_1 penalty function (also the absolute value penalty function). The exact l_1 penalty function method has been introduced by Pietrzykowski [29]. Most of the literature on nondifferentiable exact penalty functions methods for optimization problems is devoted to the study of conditions ensuring that a (local) optimum of the original constrained optimization problem is also an unconstrained (local) minimizer of the exact penalty function. In the literature, it can be found a lot of researches which has been developed the exactness of the exact l_1 penalty function method (see, for example, [8, 12, 13, 18, 19, 21, 24, 26, 28, 29, 33]). Further, Antczak [2] established this result for nonlinear smooth optimization problems (with both inequality and equality constraints) under suitable invexity assumptions. He showed that there is a lower bound $\overline{c}$

for the penalty parameter c, equal to the largest Lagrange multiplier such that, for any $c > \overline{c}$, optimal solutions sets of the original mathematical programming problem (P) and its penalized optimization problem (P(c)) with the absolute value penalty function are the same.

For $p = \infty$, we obtain the so-called exact minimax penalty function. It is given by

$$P_{\infty}(x,c) = f(x) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^{+}(x), |h_j(x)| \right\}.$$

The exact minimax penalty function method has been used also by Bandler and Charalambous [9].

The Lagrange multipliers of a nonlinear optimization problem and the saddle points of its Lagrangian function have been studied by many authors (see, for instance, [7, 8, 25, 30], and others). But in most of the studies, an assumption of convexity on the problems was made. Recently, several new concepts concerning a generalized convex function have been proposed. Among these, the concept of an invex function has received more attention (see Hanson [23]).

Problems of finding saddle point form a large class of problems encountered in various types of game situations (see, for instance, [11]) and also in intrinsically mathematical problems, for example, in the problem of nonlinear programming under Lagrangian formulations. Antczak [1] used the so-called η-approximation method to obtain new saddle point criteria for differentiable optimization problems. He defined an η-saddle point and an η-Lagrange function in the η-approximated optimization problem associated to the considered constrained extremum problem. Further, under invexity assumptions, he established the equivalence between an optimal solution in the given optimization problem and an η-saddle point in its associated η-approximated optimization problem. Later, using the η-approximation method, Antczak [6] established the equivalence between a second order η-saddle point and a second order η-Lagrange function and an optimal solution in twice differentiable optimization problems. In [4], Antczak defined the so-called G-saddle points of the so-called G-Lagrange function in constrained optimization problems and he proved the new saddle point criteria for nonconvex constrained optimization problems involved G-invex functions with respect to the same function η and with respect to, not necessarily, the same function G. In [16], Demyanov and Pevnyi used penalty function method for finding saddle points. The strategy used by them is to transform the constrained problem into a sequence of unconstrained problems which are considerably easier to solve than the original optimization problem. Zhao et al. [34] defined a new class of augmented Lagrangian functions for nonlinear programming problem with both equality and inequality constraints. They proved relationship between local saddle points of this new augmented Lagrangian and local optimal solutions in the considered optimization problems. In [5], a quadratic penalization technique was applied to establish strong Lagrangian duality property for an invex program under the assumption that the objective function is coercive.

In the paper, we are motivated to give a new characterization of the exact minimax penalty function method used for solving the nonconvex differentiable optimization problem with both inequality and equality constraints. The main goal of this paper is, therefore, to relate a saddle point in the considered constrained minimization problem and a minimizer of its associated penalized optimization problem with the exact minimax penalty function. A lower bound $\overline{c}$ of the penalty parameter c is given in the function of the Lagrange multipliers such that, for all penalty parameters exceeding this treshold, the equivalence mentioned above holds. In order to prove this result, we assume that the functions constituting the considered constrained optimization problem are invex with respect to the same function η (exception with those equality constraints for which the associated Lagrange multipliers are negative - these functions are assumed incave also with respect to the same function η). The results established in the paper are illustrated by a suitable example of a nonconvex smooth optimization problem solved by using the exact minimax penalty function method.

2. Basic Notations and Preliminary Definitions

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 optimization problems. One of them is invexity notion introduced by Hanson [23]. He extended the concept of convex functions and applied them to prove optimality conditions and duality results for nonlinear constrained optimization problems. Later, Craven [15] named on those functions, as invex functions.

Now, we recall the definition of an invex function introduced by Hanson [23].

Definition 3. Let X be a nonempty subset of R^n and $f : X \to R$ be a differentiable function defined on X. If there exists a vector-valued function $\eta : X \times X \to R^n$ such that, for all $x \in X$ $(x \neq u)$,

$$f(x) - f(u) \geq [\eta(x,u)]^T \nabla f(u), \quad (>) \tag{5}$$

then f is said to be (strictly) invex at $u \in X$ on X with respect to η. If the inequality (5) is satisfied for each $u \in X$, then f is an invex function on X with respect to η.

Remark 4. In the case when $\eta(x,u) = x - u$, we obtain the definition of a differentiable convex function.

Definition 5. Let X be a nonempty subset of R^n and $f : X \to R$ be a differentiable function defined on X. If there exists a vector-valued function $\eta : X \times X \to R^n$ such that, for all $x \in X$ $(x \neq u)$,

$$f(x) - f(u) \leq [\eta(x,u)]^T \nabla f(u), \quad (<) \tag{6}$$

then f is said to be (strictly) incave at $u \in X$ on X with respect to η. If (6) is satisfied for each $u \in X$, then f is an incave function on X with respect to η.

Remark 6. In the case when $\eta(x,u) = x - u$, we obtain the definition of a differentiable concave function.

Before we prove the main results for the considered optimization problem (P), we need the following useful lemma. The simple proof of this lemma is omitted in this work.

Lemma 7. *Let φ_k, $k = 1, ..., p$, be real-valued functions defined on $X \subset R^n$. For each $x \in X$, one has*

$$\max_{1 \leq k \leq p} \varphi_k(x) = \max_{\lambda \in \Lambda} \sum_{k=1}^{p} \lambda_k \varphi_k(x),$$

where $\Lambda = \{\lambda = (\lambda_1, ..., \lambda_p) \in R_+^p : \sum_{k=1}^p \lambda_k = 1\}$.

Now, for the reader's convenience, we also recall the definition of a coercive function.

Definition 8. A continuous function $f : R^n \to R$ is said to be coercive if

$$\lim_{\|x\| \to \infty} f(x) = \infty.$$

This means that, for any constant M, there must be a positive number β_M such that $f(x) \geq M$ whenever $\|x\| \geq \beta_M$. In particular, the values of f cannot remain bounded on a set X in R^n that is not bounded.

Remark 9. For f to be coercive, it is not sufficient that $f(x) \to \infty$ as each coordinate tends to ∞. Rather f must become infinite along any path for which $\|x\|$ becomes infinite.

3. Exact Minimax Penalty Function Method and Saddle Point Criteria

In this section, we use the exact minimax penalty function method for solving the nonconvex differentiable extremum problem (P) with both inequality and equality constraints. In this method, for the considered optimization problem (P) with both equality and inequality constraints, we construct the following unconstrained optimization problem as follows

$$\text{(7)} \qquad \text{minimize } P_\infty(x,c) = f(x) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(x), |h_j(x)| \right\}. \qquad (\mathrm{P}_\infty(c))$$

We call the unconstrained optimization problem defined above the penalized optimization problem with the exact minimax penalty function.

The idea of the exact minimax penalty function method is to solve the original nonlinear constrained optimization problem (P) by means of a single unconstrained minimization problem (P$_\infty(c)$). Roughly speaking, a minimax penalty function for problem (P) is a function $P_\infty(\cdot, c)$ given by (7), where $c > 0$ is the penalty parameter, with the property that there exists a lower bound $\overline{c} \geq 0$ such that for $c > \overline{c}$ any optimal solution in problem (P) is also a minimizer in the penalized optimization problem (P$_\infty(c)$) with the exact minimax penalty function.

The saddle point theory is of great importance not only in optimization theory but also in finding optimal solutions of the problems associated with exact penalty functions. Therefore, we prove the equivalence between a saddle point $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \in D \times R^m_+ \times R^p$ in the given constrained extremum problem (P) and a minimizer $\overline{x}$ in the penalized optimization problem (P$_\infty(c)$) with the exact minimax penalty function. In order to prove this result, we assume that the functions constituting the considered constrained optimization problem are invex with respect to the same function η at $\overline{x}$ on X (exception with those equality constraints for which the associated Lagrange multipliers are negative - they should be assumed incave also with respect to the same function η).

Before we do it, we recall the definition of the Lagrange function in problem (P) and the definition of a saddle point of the Lagrange function for the considered optimization problem (P) given by Bazaraa et al. [8].

Definition 10. The Lagrange function or the Lagrangian L in the considered optimization problem (P) with both inequality and equality constraints is defined by

$$L\left(x, \xi, \mu\right) := f(x) + \xi^T g(x) + \mu^T h(x). \tag{8}$$

Definition 11. A point $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \in D \times R^m_+ \times R^s$ is said to be a saddle point in the optimization problem (P) if

(i) $L\left(\overline{x}, \xi, \mu\right) \leq L\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \quad \forall \xi \in R^m_+, \forall \mu \in R^s$,

(ii) $L\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \leq L\left(x, \overline{\xi}, \overline{\mu}\right) \quad \forall x \in X$.

Theorem 12. *Let $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \in D \times R^m_+ \times R^s$ be a saddle point in the considered optimization problem (P). If c is assumed to satisfy $c \geq \sum_{i=1}^m \overline{\xi}_i + \sum_{j=1}^s \left|\overline{\mu}_j\right|$, where $\overline{\xi}_i$, $i = 1, ..., m$, $\overline{\mu}_j$, $j = 1, ..., s$, are the Lagrange multipliers associated to the inequality constraints g_i and the equality constraints h_j, respectively, then $\overline{x}$ is a minimizer in the associated penalized optimization problem (P$_\infty$(c)) with the exact minimax penalty function.*

Proof. By assumption, $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right)$ is a saddle point in the considered constrained extremum problem (P). Then, by Definition 11 i) and Definition 10, we have, for any

$\xi \in R_{+}^{m}$ and any $\mu \in R^{s}$,

$$f(\overline{x}) + \xi^{T} g(\overline{x}) + \mu^{T} h(\overline{x}) \leq f(\overline{x}) + \overline{\xi}^{T} g(\overline{x}) + \overline{\mu}^{T} h(\overline{x}). \tag{9}$$

Then, (9) implies

$$\overline{\xi}^{T} g(\overline{x}) = 0. \tag{10}$$

By Definition 11 ii) and Definition 10, we have, for all $x \in X$,

$$f(\overline{x}) + \overline{\xi}^{T} g(\overline{x}) + \overline{\mu}^{T} h(\overline{x}) \leq f(x) + \overline{\xi}^{T} g(x) + \overline{\mu}^{T} h(x). \tag{11}$$

Using (4) together with the definition of the absolute value, we get

$$f(x) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}(x) + \sum_{j=1}^{s} \overline{\mu}_{j} h_{j}(x) \leq f(x) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}^{+}(x) + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| \left|h_{j}(x)\right|. \tag{12}$$

Combining (11) and (12), we obtain

$$f(\overline{x}) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}(\overline{x}) + \sum_{j=1}^{s} \overline{\mu}_{j} h_{j}(\overline{x}) \leq f(x) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}^{+}(x) + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| \left|h_{j}(x)\right|. \tag{13}$$

Then, by $\overline{x} \in D$ and (10), it follows that

$$f(\overline{x}) \leq f(x) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}^{+}(x) + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| \left|h_{j}(x)\right|. \tag{14}$$

Using the feasibility of $\overline{x}$ together with (4), we get

$$f(\overline{x}) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}^{+}(\overline{x}) + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| \left|h_{j}(\overline{x})\right| = f(\overline{x}). \tag{15}$$

Combining (14) and (15), we obtain that the inequality

$$f(\overline{x}) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}^{+}(\overline{x}) + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| \left|h_{j}(\overline{x})\right| \leq f(x) + \sum_{i=1}^{m} \overline{\xi}_{i} g_{i}^{+}(x) + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| \left|h_{j}(x)\right| \tag{16}$$

holds for all $x \in X$.

Now, we consider two cases:

(a) $\sum_{i=1}^{m} \overline{\xi}_{i} + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| > 0$.

We divide the both sides of the inequality (16) by $\sum_{i=1}^{m} \overline{\xi}_{i} + \sum_{j=1}^{s} \left|\overline{\mu}_{j}\right| > 0$. Thus,

$$(17)\qquad \begin{aligned}&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(\overline{x})+\sum_{i=1}^{m}\frac{\overline{\xi}_i}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}g_i^{+}(\overline{x})\\ &+\sum_{j=1}^{s}\frac{\left|\overline{\mu}_j\right|}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}\left|h_j(\overline{x})\right|\leq\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(x)\\ &+\sum_{i=1}^{m}\frac{\overline{\xi}_i}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}g_i^{+}(x)+\sum_{j=1}^{s}\frac{\left|\overline{\mu}_j\right|}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}\left|h_j(x)\right|.\end{aligned}$$

We denote

$$(18)\qquad \overline{\lambda}_k=\frac{\overline{\xi}_k}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|},\ \ k=1,...,m,$$

$$(19)\qquad \overline{\lambda}_{m+k}=\frac{\overline{\mu}_k}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|},\ \ k=1,...,s,$$

$$(20)\qquad \varphi_k(x)=g_k^{+}(x),\ \ k=1,...,m,$$

$$(21)\qquad \varphi_{m+k}(x)=\left|h_k(x)\right|,\ \ k=1,...,s.$$

Note that

$$(22)\qquad \overline{\lambda}_k\geq 0,\quad k=1,...,m+s,\qquad \sum_{k=1}^{m+s}\overline{\lambda}_k=1.$$

Combining (17)-(22), we obtain, for all $x\in X$,

$$(23)\qquad \begin{aligned}&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(\overline{x})+\sum_{k=1}^{m+s}\overline{\lambda}_k\varphi_k(\overline{x})\\ \leq&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(x)+\sum_{k=1}^{m+s}\overline{\lambda}_k\varphi_k(x).\end{aligned}$$

Using the feasibility of $\overline{x}$ in problem (P) together with (18)-(22), we have

$$(24)\qquad \max_{\lambda\in\Lambda}\sum_{k=1}^{m+s}\lambda_k\varphi_k(\overline{x})=\sum_{k=1}^{m+s}\overline{\lambda}_k\varphi_k(\overline{x}),$$

where $\Lambda=\{\lambda=(\lambda_1,...,\lambda_{m+s})\in R_{+}^{m+s}:\sum_{k=1}^{m+s}\lambda_k=1\}$. Thus, (23) and (24) imply

$$(25)\qquad \begin{aligned}&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(\overline{x})+\max_{\lambda\in\Lambda}\sum_{k=1}^{m+s}\lambda_k\varphi_k(\overline{x})\\ \leq&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(x)+\max_{\lambda\in\Lambda}\sum_{k=1}^{m+s}\lambda_k\varphi_k(x).\end{aligned}$$

By Lemma 7, it follows that

$$\begin{aligned}&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(\overline{x})+\max_{1\le k\le m+s}\varphi_k(\overline{x})\\ \le\ &\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(x)+\max_{1\le k\le m+s}\varphi_k(x).\end{aligned}$$

Therefore, (20) and (21) yield

$$\begin{aligned}&\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(\overline{x})+\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(\overline{x}),\left|h_j(\overline{x})\right|\right\}\\ \le\ &\frac{1}{\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|}f(x)+\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(x),\left|h_j(x)\right|\right\}.\end{aligned}\tag{26}$$

We multiply the inequality above by $\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|>0$. Thus,

$$\begin{aligned}&f(\overline{x})+\left(\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|\right)\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(\overline{x}),\left|h_j(\overline{x})\right|\right\}\\ \le\ &f(x)+\left(\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|\right)\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(x),\left|h_j(x)\right|\right\}.\end{aligned}\tag{27}$$

By assumption, $c\ge\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|$. Since $\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(\overline{x}),\left|h_j(\overline{x})\right|\right\}=0$, then the following inequality

$$f(\overline{x})+c\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(\overline{x}),\left|h_j(\overline{x})\right|\right\}\le f(x)+c\max_{\substack{1\le i\le m\\1\le j\le s}}\left\{g_i^{+}(x),\left|h_j(x)\right|\right\}$$

holds for all $x\in X$. By definition of the minimax penalized optimization problem ($\mathrm{P}_\infty(c)$), it follows that the inequality

$$P_\infty\left(\overline{x},c\right)\le P_\infty\left(x,c\right)\tag{28}$$

holds for all $x\in X$. This means that $\overline{x}$ is a minimizer in the associated penalized optimization problem ($\mathrm{P}_\infty(c)$) with the exact minimax penalty function.

(b) $\sum_{i=1}^{m}\overline{\xi}_i+\sum_{j=1}^{s}\left|\overline{\mu}_j\right|=0$.

By Definition 11 ii) and Definition 10, it follows that the inequality

$$f\left(\overline{x}\right)\le f\left(x\right)$$

holds for all $x \in X$. Thus, the assumption $\sum_{i=1}^{m} \overline{\xi}_i + \sum_{j=1}^{s} \left|\overline{\mu}_j\right| = 0$ implies that the inequality

$$f(\overline{x}) + \left(\sum_{i=1}^{m} \overline{\xi}_i + \sum_{j=1}^{s} \left|\overline{\mu}_j\right| \right) \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\overline{x}), \left|h_j(\overline{x})\right| \right\}$$
$$\leq f(x) + \left(\sum_{i=1}^{m} \overline{\xi}_i + \sum_{j=1}^{s} \left|\overline{\mu}_j\right| \right) \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(x), \left|h_j(x)\right| \right\}$$

holds for all $x \in X$. From the feasibility of $\overline{x}$ in problem (P), it follows that

$$\max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\overline{x}), \left|h_j(\overline{x})\right| \right\} = 0.$$

Thus, the assumption $c \geq \sum_{i=1}^{m} \overline{\xi}_i + \sum_{j=1}^{s} \left|\overline{\mu}_j\right| = 0$ implies that the inequality

$$f(\overline{x}) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\overline{x}), \left|h_j(\overline{x})\right| \right\} \leq f(x) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(x), \left|h_j(x)\right| \right\}$$

is satisfied for all $x \in X$. By definition of the minimax penalized optimization problem (P$_\infty(c)$), it follows that the inequality

$$P_\infty\left(\overline{x}, c\right) \leq P_\infty\left(x, c\right) \tag{29}$$

holds for all $x \in X$. This means that $\overline{x}$ is a minimizer in the associated penalized optimization problem (P$_\infty(c)$) with the exact minimax penalty function.

In both considered cases a) and b), we prove that, for all penalty parameters c exceeding the treshold value equal to $\sum_{i=1}^{m} \overline{\xi}_i + \sum_{j=1}^{s} \left|\overline{\mu}_j\right|$, the point $\overline{x}$ is a minimizer in the penalized optimization problem (P$_\infty(c)$) with the exact minimax penalty function. Thus, the proof of theorem is completed. ■

Remark 13. Note that Theorem 12 was established without any invexity assumption imposed on the functions constituting the considered optimization problem (P).

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

Theorem 14. *Let $\overline{x}$ be a minimizer in the penalized optimization problem (P$_\infty$(c)) with the exact minimax penalty function and the penalty parameter c be assumed to satisfy the condition $c > \sum_{i=1}^{m} \widetilde{\xi}_i + \sum_{j=1}^{s} |\widetilde{\mu}_j|$, where $\widetilde{\xi}_i$, $i = 1, ..., m$, $\widetilde{\mu}_j$, $j = 1, ..., s$, are the Lagrange multipliers associated with the inequality constraints g_i and the equality constraints h_j, respectively, satisfying the Karush-Kuhn-Tucker necessary optimality conditions at $\widetilde{x} \in D$. Further, assume that the objective function f, the* *inequality constraints g_i, $i \in I(\widetilde{x})$, and the equality constraints h_j, $j \in J^+(\widetilde{x}) := \{j \in J : \widetilde{\mu}_j > 0\}$, are invex on X with respect to the same function η, and the equality constraints h_j, $j \in J^-(\widetilde{x}) := \{j \in J : \widetilde{\mu}_j < 0\}$, are incave on X with respect to the same function η. If the set D of all feasible solutions in problem (P) is compact, then there exist $\overline{\xi} \in R^m_+$ and $\overline{\mu} \in R^s$ such that $(\overline{x}, \overline{\xi}, \overline{\mu})$ is a saddle point in the considered constrained optimization problem (P).*

Proof. By assumption, $\overline{x}$ is a minimizer in the penalized optimization problem ($\mathrm{P}_\infty(c)$) with the exact minimax penalty function. Thus, for all $x \in X$, the following inequality

$$P_\infty(x, c) \geq P_\infty(\overline{x}, c)$$

holds. By the definition of the exact minimax penalty function $P_\infty(\cdot, c)$, we have that the inequality

$$f(x) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{g_i^+(x), |h_j(x)|\right\} \geq f(\overline{x}) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{g_i^+(\overline{x}), |h_j(\overline{x})|\right\} \tag{30}$$

holds for all $x \in X$. Thus, by (4), it follows that the inequality

$$f(x) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{g_i^+(x), |h_j(x)|\right\} \geq f(\overline{x}) \tag{31}$$

holds for all $x \in X$. Therefore, it is also satisfied for all $x \in D$. By (4), it follows that $g_i^+(x) = 0$, $i \in I$, for all $x \in D$. Moreover, by the feasibility of x in problem (P), we have that $h_j(x) = 0$, $j \in J$. Thus, (31) yields that the inequality

$$f(x) \geq f(\overline{x}) \tag{32}$$

holds for all $x \in D$.

In order to prove that there exist the Lagrange multipliers $\overline{\xi}_i$, $i = 1, ..., m$, $\overline{\mu}_j$, $j = 1, ..., s$, associated to the inequality constraints g_i and the equality constraints h_j, respectively, such that $(\overline{x}, \overline{\xi}, \overline{\mu})$ is a saddle point of the Lagrange function in the given constrained 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). 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. Therefore, the considered optimization problem (P) has an optimal solution $\widetilde{x}$. Thus, the Karush-Kuhn-Tucker necessary optimality conditions are satisfied at $\widetilde{x}$ with the Lagrange multipliers $\widetilde{\xi} \in R^m_+$ and $\widetilde{\mu} \in R^s$. By assumption, the objective function f and the constraint functions g_i, $i \in I(\widetilde{x})$, h_j, $j \in J^+(\widetilde{x})$, are invex at $\widetilde{x}$ on X with respect to the same function η, and, moreover, the constraint functions h_j, $j \in J^-(\widetilde{x})$, are incave at $\widetilde{x}$ on X with respect to the same function η. Hence, by Definitions 3 and 5, respectively, the following inequalities

$$\begin{aligned} f(x) - f(\widetilde{x}) &\geq [\eta(x,\widetilde{x})]^T \nabla f(\widetilde{x}), \\ g_i(x) - g_i(\widetilde{x}) &\geq [\eta(x,\widetilde{x})]^T \nabla g_i(\widetilde{x}), \quad i \in I(\widetilde{x}), \\ h_j(x) - h_j(\widetilde{x}) &\geq [\eta(x,\widetilde{x})]^T \nabla h_j(\widetilde{x}), \quad j \in J^+(\widetilde{x}), \\ h_j(x) - h_j(\widetilde{x}) &\leq [\eta(x,\widetilde{x})]^T \nabla h_j(\widetilde{x}), \quad j \in J^-(\widetilde{x}) \end{aligned}$$

hold for all $x \in X$. Therefore, they are also satisfied for $x = \overline{x}$. Thus,

$$f(\overline{x}) - f(\widetilde{x}) \geq [\eta(\overline{x},\widetilde{x})]^T \nabla f(\widetilde{x}), \tag{33}$$

$$g_i(\overline{x}) - g_i(\widetilde{x}) \geq [\eta(\overline{x},\widetilde{x})]^T \nabla g_i(\widetilde{x}), \quad i \in I(\widetilde{x}), \tag{34}$$

$$h_j(\overline{x}) - h_j(\widetilde{x}) \geq [\eta(\overline{x},\widetilde{x})]^T \nabla h_j(\widetilde{x}), \quad j \in J^+(\widetilde{x}), \tag{35}$$

$$h_j(\overline{x}) - h_j(\widetilde{x}) \leq [\eta(\overline{x},\widetilde{x})]^T \nabla h_j(\widetilde{x}), \quad j \in J^-(\widetilde{x}). \tag{36}$$

Since the Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are satisfied at $\widetilde{x}$ with Lagrange multipliers $\widetilde{\xi} \in R^m_+$ and $\widetilde{\mu} \in R^s$, then

$$\widetilde{\xi}_i g_i(\overline{x}) - \widetilde{\xi}_i g_i(\widetilde{x}) \geq [\eta(\overline{x},\widetilde{x})]^T \widetilde{\xi}_i \nabla g_i(\widetilde{x}), \quad i \in I(\widetilde{x}), \tag{37}$$

$$\widetilde{\mu}_j h_j(\overline{x}) - \widetilde{\mu}_j h_j(\widetilde{x}) \geq [\eta(\overline{x},\widetilde{x})]^T \widetilde{\mu}_j \nabla h_j(\widetilde{x}), \quad j \in J^+ \cup J^-(\widetilde{x}). \tag{38}$$

Adding both sides of (33), (37) and (38) and taking into account the Lagrange multipliers equal to 0, we obtain

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

By the Karush-Kuhn-Tucker necessary optimality conditions (1), we get

$$f(\overline{x}) - f(\widetilde{x}) + \sum_{i=1}^{m} \widetilde{\xi}_i \left(g_i(\overline{x}) - g_i(\widetilde{x}) \right) + \sum_{j=1}^{s} \widetilde{\mu}_j \left(h_j(\overline{x}) - h_j(\widetilde{x}) \right) \geq 0.$$

Thus,

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

From the feasibility of $\widetilde{x}$ in problem (P) and the Karush-Kuhn-Tucker necessary optimality condition (2), it follows that

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

Hence, by (4) and the definition of the absolute value, we get

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

Now, assume that $\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|>0$. We divide the inequality above by $\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|>0$. Thus,

$$\begin{aligned}\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\overline{x})+\sum_{i=1}^{m}\frac{\widetilde{\xi}_i}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}g_i^+(\overline{x})\\+\sum_{j=1}^{s}\frac{|\widetilde{\mu}_j|}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}\,|h_j(\overline{x})|\geq\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\widetilde{x}).\end{aligned}\tag{39}$$

We denote

$$\widetilde{\lambda}_k=\frac{\widetilde{\xi}_k}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|},\ \ k=1,...,m,\tag{40}$$

$$\widetilde{\lambda}_{m+k}=\frac{\widetilde{\mu}_k}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|},\ \ k=1,...,s,\tag{41}$$

$$\varphi_k(\overline{x})=g_k^+(\overline{x}),\ \ k=1,...,m,\tag{42}$$

$$\varphi_{m+k}(\overline{x})=|h_k(\overline{x})|\,,\ \ k=1,...,s.\tag{43}$$

Note that

$$\widetilde{\lambda}_k\geq 0,\ \ k=1,...,m+s,\quad \sum_{k=1}^{m+s}\widetilde{\lambda}_k=1.\tag{44}$$

Combining (40)-(44), we obtain

$$\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\overline{x})+\sum_{k=1}^{m+s}\widetilde{\lambda}_k\varphi_k(\overline{x})\geq\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\widetilde{x}).$$

By (40)-(44), it follows that

$$\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\overline{x})+\max_{\lambda\in\Lambda}\sum_{k=1}^{m+s}\lambda_k\varphi_k(\overline{x})\geq\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\widetilde{x}),$$

where $\Lambda=\{\lambda=(\lambda_1,...,\lambda_{m+s})\in R_+^{m+s}:\sum_{k=1}^{m+s}\lambda_k=1\}$. Thus, by Lemma 7, it follows that

$$\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\overline{x})+\max_{1\leq k\leq m+s}\varphi_k(\overline{x})\geq\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\widetilde{x}).$$

Hence, by (42) and (43), we have

$$\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\overline{x})+\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\overline{x}),|h_j(\overline{x})|\right\}\geq\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\widetilde{x}).$$

Using (4) together with the feasibility of $\widetilde{x}$ in the constrained optimization problem (P), we get

$$\begin{aligned}&\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\overline{x})+\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\overline{x}),|h_j(\overline{x})|\right\}\\&\geq\frac{1}{\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|}f(\widetilde{x})+\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\widetilde{x}),|h_j(\widetilde{x})|\right\}.\end{aligned}\tag{45}$$

We multiply the inequality above by $\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|>0$. Thus,

$$\begin{aligned}&f(\overline{x})+\left(\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|\right)\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\overline{x}),|h_j(\overline{x})|\right\}\\&\geq f(\widetilde{x})+\left(\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|\right)\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\widetilde{x}),|h_j(\widetilde{x})|\right\}.\end{aligned}\tag{46}$$

By assumption, $c>\sum_{i=1}^{m}\widetilde{\xi}_i+\sum_{j=1}^{s}|\widetilde{\mu}_j|$. Therefore, (46) gives

$$f(\overline{x})+c\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\overline{x}),|h_j(\overline{x})|\right\}>f(\widetilde{x})+c\max_{\substack{1\leq i\leq m\\1\leq j\leq s}}\left\{g_i^+(\widetilde{x}),|h_j(\widetilde{x})|\right\}.$$

Hence, by the definition of the penalized optimization problem ($\mathrm{P}_\infty(c)$) (see (7)), it follows that the inequality

$$P_\infty(\overline{x},c)>P_\infty(\widetilde{x},c)$$

holds, contradicting the optimality of $\overline{x}$ in the penalized optimization problem ($\mathrm{P}_\infty(c)$) with the exact minimax penalty function. Thus, we have established that $\overline{x}$ is feasible in the given minimization problem (P).

Now, we consider the case when $\sum_{i=1}^{m} \widetilde{\xi}_i + \sum_{j=1}^{s} |\widetilde{\mu}_j| = 0$. Hence, using the Karush-Kuhn-Tucker necessary optimality conditions (1) together with (33), we get

$$f(\overline{x}) \geq f(\widetilde{x}). \tag{47}$$

Since $\sum_{i=1}^{m} \widetilde{\xi}_i + \sum_{j=1}^{s} |\widetilde{\mu}_j| = 0$, then

$$\begin{aligned} & f(\overline{x}) + \left(\sum_{i=1}^{m} \widetilde{\xi}_i + \sum_{j=1}^{s} |\widetilde{\mu}_j| \right) \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\overline{x}), |h_j(\overline{x})| \right\} \\ & \geq f(\widetilde{x}) + \left(\sum_{i=1}^{m} \widetilde{\xi}_i + \sum_{j=1}^{s} |\widetilde{\mu}_j| \right) \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\widetilde{x}), |h_j(\widetilde{x})| \right\}. \end{aligned}$$

By assumption, $c > \sum_{i=1}^{m} \widetilde{\xi}_i + \sum_{j=1}^{s} |\widetilde{\mu}_j| = 0$. Thus,

$$f(\overline{x}) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\overline{x}), |h_j(\overline{x})| \right\} > f(\widetilde{x}) + c \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq s}} \left\{ g_i^+(\widetilde{x}), |h_j(\widetilde{x})| \right\}.$$

Hence, by the definition of the penalized optimization problem (P$_\infty(c)$) (see (7)), it follows that the inequality

$$P_\infty(\overline{x}, c) > P_\infty(\widetilde{x}, c)$$

holds, contradicting the optimality of $\overline{x}$ in the penalized optimization problem (P$_\infty(c)$) with the exact minimax penalty function. Thus, we have established that $\overline{x}$ is feasible in the considered constrained optimization problem (P).

Hence, by (32), it follows that $\overline{x}$ is optimal in the given optimization problem (P). Then, there exist the Lagrange multipliers $\overline{\xi} \in R_+^m$ and $\overline{\mu} \in R^s$ such that the Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are satisfied at $\overline{x}$.

Now, by Definition 11, we show that $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \in D \times R_+^m \times R^s$ is a saddle point of the Lagrange function in the considered constrained optimization problem (P).

First, we prove the relation i) in Definition 11. By the Karush-Kuhn-Tucker optimality condition (2), the inequality

$$\xi^T g(\overline{x}) \leq \overline{\xi}^T g(\overline{x})$$

holds for all $\xi \in R_+^m$. Hence,

$$f\left(\overline{x}\right) + \xi^T g(\overline{x}) + \mu^T h(\overline{x}) \leq f\left(\overline{x}\right) + \overline{\xi}^T g(\overline{x}) + \overline{\mu}^T h(\overline{x}).$$

By definition of the Lagrange function (8), the following inequality

$$L\left(\overline{x}, \xi, \mu\right) \leq L\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \tag{48}$$

holds for all $\xi \in R_+^m$ and $\mu \in R^s$.

Now, we prove the inequality ii) in Definition 11. By assumption, f, g_i, $i \in I(\overline{x})$, h_j, $j \in J^+(\overline{x})$, are invex functions at $\overline{x}$ with respect to the same function η on X and, moreover, h_j, $j \in J^-(\overline{x})$, are incave at $\overline{x}$ with respect to the same function η on X. Then, by Definitions 3 and 5, respectively, the inequalities

(49) $$f(x) - f(\overline{x}) \geq [\eta(x,\overline{x})]^T \nabla f(\overline{x}),$$

(50) $$g_i(x) - g_i(\overline{x}) \geq [\eta(x,\overline{x})]^T \nabla g_i(\overline{x}), \ \ i \in I(\overline{x}),$$

(51) $$h_j(x) - h_j(\overline{x}) \geq [\eta(x,\overline{x})]^T \nabla h_j(\overline{x}), \ \ j \in J^+(\overline{x}),$$

(52) $$h_j(x) - h_j(\overline{x}) \leq [\eta(x,\overline{x})]^T \nabla h_j(\overline{x}), \ \ j \in J^-(\overline{x})$$

hold for all $x \in X$. Multiplying (50)-(53) by the Lagrange multipliers $\overline{\xi}_i \geq 0$, $i \in I$, $\overline{\mu} > 0$, $j \in J^+(\overline{x})$, $\overline{\mu} < 0$, $j \in J^-(\overline{x})$, respectively, we get

(53) $$\overline{\xi}_i g_i(x) - \overline{\xi}_i g_i(\overline{x}) \geq [\eta(x,\overline{x})]^T \overline{\xi}_i \nabla g_i(\overline{x}), \ i \in I(\overline{x}),$$

(54) $$\overline{\mu}_j h_j(x) - \overline{\mu}_j h_j(\overline{x}) \geq [\eta(x,\overline{x})]^T \overline{\mu}_j \nabla h_j(\overline{x}), \ \ j \in J^+(\overline{x}) \cup J^-(\overline{x}).$$

Adding both sides of (53) and (54), we obtain

(55) $$\sum_{i\in I(\overline{x})} \overline{\xi}_i g_i(x) - \sum_{i\in I(\overline{x})} \overline{\xi}_i g_i(\overline{x}) \geq [\eta(x,\overline{x})]^T \sum_{i\in I(\overline{x})} \overline{\xi}_i \nabla g_i(\overline{x}),$$

(56)
$$\sum_{j\in J^+(\overline{x})\cup J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j\in J^+(\overline{x})\cup J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \geq [\eta(x,\overline{x})]^T \sum_{j\in J^+(\overline{x})\cup J^-(\overline{x})} \overline{\mu}_j \nabla h_j(\overline{x}).$$

Now, adding both sides of (49), (55) and (56), we get

$$\begin{aligned}
&f(x) - f(\overline{x}) + \sum_{i\in I(\overline{x})} \overline{\xi}_i g_i(x) - \sum_{i\in I(\overline{x})} \overline{\xi}_i g_i(\overline{x}) \\
&+ \sum_{j\in J^+(\overline{x})\cup J^-(\overline{x})} \overline{\mu}_j h_j(x) - \sum_{j\in J^+(\overline{x})\cup J^-(\overline{x})} \overline{\mu}_j h_j(\overline{x}) \\
\geq\ & [\eta(x,\overline{x})]^T \left[\nabla f(\overline{x}) + \textstyle\sum_{i=1}^m \overline{\xi}_i \nabla g_i(\overline{x}) + \sum_{j\in J^+(\overline{x})\cup J^-(\overline{x})} \overline{\mu}_j \nabla h_j(\overline{x})\right].
\end{aligned}$$

Taking into account the Lagrange multipliers equal to 0, then the Karush-Kuhn-Tucker necessary optimality condition (1) yields that the inequality

$$f(x) + \sum_{i\in I} \overline{\xi}_i g_i(x) + \sum_{j\in J} \overline{\mu}_j h_j(x) \geq f(\overline{x}) + \sum_{i\in I} \overline{\xi}_i g_i(\overline{x}) + \sum_{j\in J} \overline{\mu}_j h_j(\overline{x})$$

holds for all $x \in X$. Thus, by the definition of the Lagrange function (8), the following inequality

$$L\left(x, \overline{\xi}, \overline{\mu}\right) \geq L\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \tag{57}$$

holds for all $x \in X$. Inequalities (48) and (57) mean, by Definition 11, that $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right)$ is a saddle point of the Lagrange function in the considered optimization problem (P). ∎

In Theorems 12 and 14, the equivalence between a saddle point in the considered optimization problem (P) and a minimizer in its associated penalized optimization problem (P$_\infty(c)$) with the exact minimax penalty function has been established for all penalty parameters exceeding the given treshold. The main tool to prove this equivalence turns out a suitable invexity assumption imposed on the functions constituting the considered nonlinear optimization problem (P).

Now, we illustrate the results established in the paper by the help of an example of a nonconvex constrained optimization problem with invex functions. In order to solve it, we use the exact minimax penalty function method.

Example 15. Consider the following nonlinear constrained optimization problem with both inequality and equality constraints

$$\begin{aligned} f(x) &= \arctan^2(x_1) + \arctan(x_1) + \arctan^2(x_2) + \arctan(x_2) \to \min \\ g_1(x) &= \arctan^2(x_1) - \arctan(x_1) \leq 0, \\ g_2(x) &= -\arctan(x_2) \leq 0, \\ h(x) &= \arctan(x_1) - \arctan(x_2) = 0, \\ X &= R^2. \end{aligned} \tag{P1}$$

Note that the set of all feasible solutions in problem (P1) is the set $D = \left\{(x_1, x_2) \in R^2 : 0 \leq x_1 \leq \frac{\pi}{4} \ \wedge \ x_2 \geq 0 \ \wedge \ x_1 = x_2\right\}$. We use the exact minimax penalty function method for solving the considered optimization problem (P1). Therefore, we construct the following unconstrained optimization problem

$$\begin{aligned} P1_\infty(x, c) &= \arctan^2(x_1) + \arctan(x_1) + \arctan^2(x_2) + \arctan(x_2) + \\ & c\left\{\max\left\{0, \arctan^2(x_1) - \arctan(x_1)\right\}, \ \max\left\{0, -\arctan(x_2)\right\}, \right. \\ & \left. |\arctan(x_1) - \arctan(x_2)|\right\} \to \min . \end{aligned} \tag{P1$_\infty(c)$}$$

Note that $\overline{x} = (0, 0)$ is such a feasible solution in problem (P1) at which the Karush-Kuhn-Tucker necessary optimality conditions (1)-(3) are satisfied with the Lagrange multipliers $\overline{\xi} = \left(\overline{\xi}_1, \overline{\xi}_2\right)$, where $\overline{\xi}_1 = \overline{\mu} + 1$, $\overline{\xi}_2 = -\overline{\mu} + 1$ and $-1 \leq \overline{\mu} \leq 1$. Further, it is can be showed, by Definition 11, that $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \in D \times R^2_+ \times R$ is a saddle point of the Lagrange function in the considered constrained optimization problem (P1). Therefore, by Theorem 12, for every penalty parameter $c \geq \sum_{i=1}^{2} \overline{\xi}_i + |\overline{\mu}| = 2 + |\overline{\mu}| = 3$, it follows that $\overline{x} = (0,0)$ is a minimizer in any associated penalized optimization problem (P1$_\infty(c)$) with the exact minimax penalty function. Now, we show the converse result. In fact, it can be established, by Definition 3, that both the objective function f and the constraint function $g = (g_1, g_2)$ are invex on R^2 with respect to the same function η defined by

$$\eta(x,u) = \begin{bmatrix} \eta_1(x,u) \\ \\ \eta_2(x,u) \end{bmatrix} = \begin{bmatrix} \left(1+u_1^2\right)(\arctan(x_1) - \arctan(u_1)) \\ \\ \left(1+u_2^2\right)(\arctan(x_2) - \arctan(u_2)) \end{bmatrix}.$$

Also it can be showed that, if $\overline{\mu}$ is positive, then the equality constraint h is invex on R^2 with respect to the function η defined above and, in the case when $\overline{\mu}$ is negative, then the equality constraint h is incave on R^2 with respect to the function η defined above. Further, the set D is a compact subset of R^2. Thus, all hypotheses of Theorem 14 are fulfilled. Since the point $\overline{x} = (0,0)$ is a minimizer in the penalized optimization problem (P1$_\infty(c)$) with the exact minimax penalty function for any penalty parameter c satisfying $c > \overline{\xi}_1 + \overline{\xi}_2 + |\overline{\mu}| = 3$, therefore $\left(\overline{x}, \overline{\xi}, \overline{\mu}\right) \in D \times R_+^2 \times R$ is a saddle point of the Lagrange function in the considered optimization problem (P1). Thus, there exists the lower bound $\overline{c}$ equal to $\sum_{i=1}^{2} \overline{\xi}_i + |\overline{\mu}|$ such that, for every penalty parameter c exceeding this threshold value, the considered minimization problem (P1) and any its penalized optimization problem (P1$_\infty(c)$) with the exact minimax penalty function are equivalent in the sense discussed in the paper.

In the next example, we consider a constrained optimization problem in which not all functions are invex. It turns out that, for such constrained optimization problems, the equivalence might not hold between the set of saddle points in the given constrained optimization problem and the set of minimizers in its associated penalized optimization problem with exact minimax penalty function.

Example 16. Consider the following constrained optimization problem

$$\begin{aligned} f(x) = \tfrac{1}{4}x^4 - \tfrac{1}{3}x^3 - x^2 + 1 &\to \min \\ g_1(x) = 2 - x &\leq 0, \qquad\qquad \text{(P2)} \\ g_2(x) = -x^2 + 3x - 2 &\leq 0. \end{aligned}$$

Note that $D = \{x \in R : x \geq 2\}$ and a point $\left(\overline{x}, \overline{\xi}_1, \overline{\xi}_2\right) = (2,0,0)$ is a saddle point in the considered optimization problem (P2). Further, by Theorem 1 [10], it follows that the objective function f and the constraint function g_2 are not invex on R with respect to any function η defined by $\eta : R \times R \to R$. However, we use the exact minimax penalty function method to solve problem (P2). Then, in order to solve (P2) by this method, we construct the following unconstrained optimization problem

$$P2_{\infty}(x,c) = \frac{1}{4}x^4 - \frac{1}{3}x^3 - x^2 + 1 + \\ c\max\left\{\max\left\{0, 2-x\right\}, \max\left\{0, -x^2+3x-2\right\}\right\} \to \min. \qquad (\mathrm{P2}_{\infty}(c))$$

It is not difficult to show that $P2_{\infty}(x,c)$ does not have a minimizer at $\overline{x} = 2$ for any $c > 0$, that is, for every penalty parameter exceeding the treshold given in the paper. This is a consequence of the fact that not all functions constituting problem (P2) are not invex on R (with respect to any function $\eta : R \times R \to R$). Therefore, since all functions constituting the given constrained optimization problem are not invex with respect to any function $\eta : R \times R \to R$, then the treshold value of the penalty parameter is not equal to the sum of the absolute value of the Lagrange multipliers satisfying the Karush-Kuhn-Tucker necessary optimality conditions. As it follows even from this example, if this assumption is violated, then the equivalence between the given optimization problem (P) and its penalized optimization problem with the exact minimax penalty function might not hold for every penalty parameter exceeding this threshold.

Remark 17. Note that the objective function f in the optimization problem (P2) in Example 16 is coercive (see [28]). Hence, as it follows from Example 16, the coercive assumption of the objective function is not sufficient condition to prove the equivalence between the set of saddle points of the Lagrange function in the given constrained optimization problem and the set of minimizers in its penalized optimization problem with the exact minimax penalty function. This is also a consequence of the fact that not all functions constituting problem (P2) are invex on R (with respect to any function $\eta : R \times R \to R$).

4. Conclusion

In the paper, the exact minimax penalty function method has been used for solving nonconvex differentiable optimization problems involving both inequality and equality constraints. In this method, for the considered constrained minimization problem (P), an associated penalized optimization problem ($\mathrm{P}_{\infty}(c)$) with the exact minimax penalty function is constructed. A lower bound on the penalty parameter c has been given in the paper such that, for every penalty parameter c exceeding this treshold, the equivalence holds between a saddle point $(\overline{x}, \overline{\xi}, \overline{\mu}) \in D \times R^m_+ \times R^s$ in the given nonconvex constrained optimization problem (P) and a minimizer $\overline{x}$ in its associated penalized optimization problem ($\mathrm{P}_{\infty}(c)$) with the exact minimax penalty function. This result has been established under assumption that the functions constituting the given constrained optimization problem (P) are invex with respect to the same function η (exception with these equality constraints for which the associated Lagrange multipliers are negative - those functions should be assumed incave with respect to η). Further, it has been shown that, in the absence of invexity assumption, the equivalence between the set of saddle points of the Lagrange function in the given constrained optimization problem and the set of minimizers in its penalized optimization problem with the exact minimax penalty function might not hold for every penalty parameter exceeding the threshold given in the paper. Also the coercive assumption of the objective function is not sufficient to prove this equivalence for all penalty parameters above the given treshold. As it follows from the paper, the concept of invexity is a useful tool to prove the the equivalence between the set of saddle points of the Lagrange function in the given constrained optimization problem and the set of minimizers in its penalized optimization problem with the exact minimax penalty function. The principal motivation for the exact minimax penalty function method presented in this paper is that, in the comparison to the classical exact l_1 function method (see [2]), there is no the lower bound on the penalty parameter c in the optimization literature such that, for all penalty parameters exceeding this treshold, the equivalence mentioned above holds for nonconvex constrained extremum problem and its associated penalized optimization problem with the exact minimax penalty function.

References

1. T. Antczak, Saddle-point criteria in an η-approximation method for nonlinear mathematical programming problems involving invex functions, *Journal of Optimization, Theory and Applications*, **132** (2007), 71-87.
2. T. Antczak, Exact penalty functions method for mathematical programming problems involving invex functions, *European Journal of Operational Research*, **198** (2009), 29-36.
3. T. Antczak, The l_1 penalty function method for nonconvex differentiable optimization problems with inequality constraints, *Asia-Pacific Journal of Operational Research*, **27** (2010), 1-18.
4. T. Antczak, G-saddle point criteria and G-Wolfe duality in differentiate mathematical programming, *Journal of Information & Optimization Sciences*, **31** (2010), 63-85.
5. T. Antczak, Penalty function methods and a duality gap for invex optimization problems, *Nonlinear Analysis*, **71** (2009), 3322-3332.
6. T. Antczak, Saddle points criteria via a second order η-approximation approach for nonlinear mathematical programming involving second order invex functions, *Kybernetika*, **47** (2011), 222-240.
7. K. J. Arrow, F. J. Gould and S. M. Howe, A general saddle point result for constrained optimization, *Mathematical Programming*, **5** (1973), 225-234.
8. M. S. Bazaraa, H. D. Sherali and C. M. Shetty, *Nonlinear programming: theory and algorithms*, John Wiley and Sons, New York, 1991.
9. J. W. Bandler and C. Charalambous, Nonlinear programming using minimax techniques, *Journal of Optimization, Theory and Applications*, **13** (1974), 607-619.

10. A. Ben-Israel and B. Mond, What is invexity? *Journal of Australian Mathematical Society Ser. B*, **28** (1986), 1-9.

11. D. P. Bertsekas, *Constrained optimization and Lagrange multiplier methods*, Academic Press, Inc. 1982.

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

13. Ch. Charalambous, A lower bound for the controlling parameters of the exact penalty functions, *Mathematical Programming*, **15** (1978), 278-290.

14. A. R. Conn, Constrained optimization using a nondifferentiable penalty function, *SIAM Journal of Numerical Analysis*, **10** (1973), 760-784.

15. B. D. Craven, Invex functions and constrained local minima, *Bulletin of the Australian Mathematical Society*, **24** (1981), 357-366.

16. V. Demyanov and A. Pevnyi, Numerical methods for finding saddle points, *USSR Computational Mathematics and Mathematical Physics*, **12** (1972), 11-52.

17. I. I. Eremin, The penalty method in convex programming, *Doklady Akad. Nauk SSSR*, **143** (1967), 748-751.

18. J. P. Evans, F. J. Gould and J. W. Tolle, Exact penalty functions in nonlinear programming, *Mathematical programming*, **4** (1973), 72-97.

19. G. Di Pillo and L. Grippo, Exact penalty functions in constrained optimization, *SIAM Journal of Control and Optimization*, **27** (1989), 1333-1360.

20. R. Fletcher, A class of methods for nonlinear programming with termination and convergence properties, in: *Integer and nonlinear programming*, J. Abadie, ed. Amsterdam, 1970, pp. 157-173.

21. R. Fletcher, An exact penalty function for nonlinear programming with inequalities, *Mathematical Programming*, **5** (1973), 129-150.

22. S. P. Han and O. L. Mangasarian, Exact penalty functions in nonlinear programming, *Mathematical Programming*, **17** (1979), 251-269.

23. M. A. Hanson, On sufficiency of the Kuhn-Tucker conditions, *Journal of Mathematical Analysis and Applications*, **80** (1981), 545-550.

24. D. Luenberger, Control problem with kinks, *IEEE Transaction on Automatic Control*, **15** (1970), 570-574.

25. O. L. Mangasarian, *Nonlinear programming*, McGraw-Hill, New York, 1969.

26. O. L. Mangasarian, Sufficiency of exact penalty minimization, *SIAM Journal of Control and Optimization*, **23** (1985), 30-37.

27. D. H. Martin, The essence of invexity, *Journal of Optimization Theory and Applications*, **42** (1985), 65-76.

28. A. L. Peressini, F. E. Sullivan and J. J. Uhl, Jr., *The mathematics of nonlinear programming*, Springer-Verlag New York Inc. 1988.

29. T. Pietrzykowski, An exact potential method for constrained maxima, *SIAM Journal of Numerical Analysis*, **6** (1969), 294-304.

30. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

31. T. Weir and V. Jeyakumar, A class of nonconvex functions and mathematical programming, *Bulletin of the Australian Mathematical Society*, **38** (1988), 177-189.

32. W. I. Zangwill, Nonlinear programming via penalty functions, *Management Science*, **13** (1967), 344-358.

33. A. J. Zaslavski, *Optimization on metric and normed spaces*, Sprnger optimization and its Applications, Springer, New York, 2010.

34. W. Zhao, J. Zhang and J. Zhou, *Existence of local saddle points for a new augmented Lagrangian function*, Mathematical Problems in Engineering, Vol. 2010, Article ID 324812.

Tadeusz Antczak
Faculty of Mathematics and Computer Science
University of Lódź
Banacha 22, 90-238 Lódź
Poland
E-mail: antczak@math.uni.lodz.pl