J Optim Theory Appl (2017) 173:770–792
DOI 10.1007/s10957-016-0998-7

Global Optimality Conditions in Nonconvex Optimization

Alexander S. Strekalovsky[1]

Received: 14 July 2015 / Accepted: 16 August 2016 / Published online: 7 April 2017

Abstract In this paper, we address the nonconvex optimization problem, with the goal function and the inequality constraints given by the functions represented by the difference of convex functions. The effectiveness of the classical Lagrange function and the max-merit function is being investigated as the merit functions of the original problem. In addition to the classical apparatus of optimization theory, we apply the new global optimality conditions for the auxiliary problems related to the Lagrange and max-merit functions.

Keywords D.c. functions · Inequality constraints · Global optimality conditions · Lagrange function · Saddle point · Linearized problem

Mathematics Subject Classification 90C26

1 Introduction

It is well known that contemporary (classical) optimization methods [1–4] certainly remain a much more popular choice for a "generic" user of the optimization software. Meanwhile, some existing implementations of the widely used global optimization methods that offer theoretical guarantees of finding a global solution [5–7] also attract a permanent attention from researchers. At the same time, according to the conver-

Communicated by Asen L. Dontchev.

This paper is dedicated to the memory of Vladimir F. Demyanov.

✉ Alexander S. Strekalovsky
strekal@icc.ru

1 Matrosov Institute for System Dynamics and Control Theory SB RAS, Lermontov St., 134, Irkutsk, Russia 664033

gence theorems [1–4] for classical methods, the latter do not provide, in general, an approximate global solution, but only a stationary one, for example, a KKT-point. Nonetheless, in real-life problems, employment of the classical optimization methods [1–4] often yields a rather satisfactory (for a user) feasible point. On the other hand, the global optimization methods [5–7] also have numerous examples of successful applications, regardless of the fact that these global search schemes suffer the curse of dimension, when the volume of computations grows exponentially along with the growth of the problem's dimension [5–7].

In addition, in the seventies of twentieth century, important discoveries and achievements were made in the area of nondifferential optimization and variational principles [8–14], which enabled the use of nonsmooth auxiliary functions. It is worth noting that Demyanov [9,10] was not only one of the pioneers of the nonsmooth optimization in the USSR, later in Russia, but he was also an outstanding researcher, the founder and supervisor of the widely known Saint Petersburg school of the nonsmooth optimization.

Nevertheless, the discoveries mentioned above do not influence considerably the development of new methods for nonconvex problems. The situation can be partially explained by the fact that the KKT-theorem and its modern generalizations [8–12] turn out to be ineffective when it comes to a characterization of a global solution to nonconvex problems. Hence, it is clear that users, as well as researchers, need new theoretical tools (optimality conditions, numerical methods, etc.) allowing not only to escape stationary or local solutions, but also to design new numerical procedures (computationally implementable) that would be able to jump out of local pits and improve the goal function at the same time.

One of the first attempts to develop advanced mathematical apparatus was undertaken in [15–19] for special (elementary) nonconvex optimization problems such as convex maximization, d.c. minimization and reverse-convex problems [5–7]. It is important to highlight that our approach (the global search theory) successfully employs all classical optimization methods (whenever suitable) and corresponding software packages (IBM CPLEX, Xpress, Gurobi, etc.) within the schemes of special local and global search methods [15–19].

In this paper, we address more general problems where the goal function and inequality constraints are given by the (d.c.) function represented by the difference of convex functions. After the statement of the problem in Sects. 2 and 5, we reduce the original problem to the two auxiliary problems with the minimization of corresponding merit functions: the so-called max-merit function and the classical Lagrangian. In Sects. 4 and 5, for each merit function, we establish necessary and sufficient global optimality conditions (GOC) and study its properties. Furthermore, we test the global optimality conditions (GOC) on nonconvex examples, simultaneously comparing the effectiveness of the merit functions.

The analysis of the research results will show that, although the GOC developed below proven effective, each merit function has its own objective different from this one of the original problem. In particular, the Lagrangian aims to find a saddle point. This becomes possible to reveal only with the help of the new global optimality conditions for each of the merit functions.

The conclusions are presented in Sect. 6.

2 Statement of the Problem

Let us consider the following problem:

$$(\mathcal{P}): \qquad \left.\begin{array}{c} \min\limits_x f_0(x) := g_0(x) - h_0(x), \quad x \in S, \\ f_i(x) := g_i(x) - h_i(x) \le 0, \quad i \in I = \{1, \dots, m\}, \end{array}\right\} \tag{1}$$

where the functions $g_i(\cdot)$, $h_i(\cdot), i = 0, 1, \dots, m$, are convex on $I\!R^n$, so that the functions $f_i(\cdot)$, $i = 0, 1, \dots, m$, are the (d.c.) functions of A.D. Alexandrov, represented as a difference of two convex functions [4,8,11].

In order to avoid some singularities [4,11], we assume that $S \subset \operatorname{int}\operatorname{dom} g_i \cap \operatorname{int}\operatorname{dom} h_i$, $\forall i \in I \cup \{0\}$, where int and dom φ stand for the interior and the domain $\operatorname{dom}\varphi = \{x \ : \ \varphi(x) < \infty\}$, respectively, and the set $S \subset I\!R^n$ is convex and may be given by the relations

$$S := \{x \in S_0 \ : \ c_k(x) \le 0, \ k = 1, \dots, K, \\ \alpha_j(x) := \langle a^j, x\rangle + \beta_j = 0, \ j = 1, \dots, K_1\},$$

where $c_k(\cdot), k = 1, \dots, K$, are convex, and $\alpha_j(\cdot)$, $j = 1, \dots, K_1$, are affine functions on $I\!R^n$, and the set $S_0 \subset I\!R^n$ is convex as well. Besides, suppose that the set $Sol(\mathcal{P})$ of global solutions to Problem $(\mathcal{P})$, $Sol(\mathcal{P}) = \{z \in D : f_0(z) = \mathcal{V}(\mathcal{P})\}$, is nonempty, where $D := \{z \in S : f_i(x) \le 0, \quad i \in I\} \ne \varnothing$ and

$$\mathcal{V}(\mathcal{P}) := \inf(f_0, D) = \inf_x\{f_0(x) : x \in S, \quad f_i(x) \le 0, \ \ i \in I\} > -\infty,$$

are the feasible set and the optimal value of Problem $(\mathcal{P})$, respectively.

3 The Auxiliary Problem

Let us consider the following function [4,12–14]

$$F(x, \eta) := \max\{f_0(x) - \eta; \ f_1(x), \dots, f_m(x)\}, \tag{2}$$

where $\eta \in I\!R$. Below, for a feasible (in $(\mathcal{P})$) point $z \in D$, we denote $\zeta := f_0(z)$.

Proposition 3.1 [4,12–14] *Suppose z is a solution to Problem $(\mathcal{P})$, $z \in Sol(\mathcal{P})$. Then, the point z is a solution to the following auxiliary problem*

$$(\mathcal{P}_\zeta): \qquad \min_x F(x, \zeta), \quad x \in S. \tag{3}$$

Proposition 3.2 [4] *The point $z \in D$ is a solution to Problem $(\mathcal{P})$ if and only if z is a solution to the auxiliary problem $(\mathcal{P}_\zeta)$ with $\zeta = \mathcal{V}(\mathcal{P})$. Under latter conditions, the equality $Sol(\mathcal{P}) = Sol(\mathcal{P}_\zeta)$ holds.*

Below we will use Proposition 3.1 in the contrapositive form as follows.

Lemma 3.1 *Suppose that the feasible for Problem $(\mathcal{P})$ point $z \in D$ is not a solution to problem $(\mathcal{P}_\zeta)$, so that there exists a point $u \in S$, such that $F(u, \zeta) < 0 = F(z, \zeta)$. Then, the point $z \in D$ cannot be a solution to Problem $(\mathcal{P})$: $z \notin Sol(\mathcal{P})$.*

Proof From the inequality $F(u, \zeta) < 0$, it follows that $u \in S$, $f_i(u) < 0$, and $f_0(u) < f_0(z) = \zeta$, so that u is feasible for Problem $(\mathcal{P})$. Hence, $z \notin Sol(\mathcal{P})$. $\square$

4 Optimality Conditions

First of all, let us show that the objective function $F(x, \eta)$ of Problem $(\mathcal{P}_\eta)$–(3), given in (2), is a d.c. function. For that purpose, we employ the equality $F(x, \eta) = F(x, \eta) + H(x) - H(x)$ where

$$H(x) = \sum_{i=0}^{m} h_i(x). \tag{4}$$

From (4), due to (1) and (2), it immediately follows that

$$\begin{aligned} F(x, \eta) &= \max\left\{ f_0(x) + H(x) - \eta; \quad \max_{1 \le i \le m} [f_i(x) + H(x)] \right\} - H(x) \\ &= \max\left\{ g_0(x) + \sum_{j=1}^{m} h_j(x) - \eta; \quad \max_{1 \le i \le m} \left[g_i(x) + \sum_{j \ne i} h_j(x) \right] \right\} \\ &\quad - \sum_{i=0}^{m} h_i(x) = G(x, \eta) - H(x), \end{aligned} \tag{5}$$

where

$$\begin{aligned} G(x, \eta) &= \max\left\{ f_0(x) + H(x) - \eta; \quad \max_{1 \le i \le m} [f_i(x) + H(x)] \right\} \\ &= \max\left\{ g_0(x) + \sum_{j=1}^{m} h_j(x) - \eta; \quad \max_{1 \le i \le m} \left[g_i(x) + \sum_{j \ne i} h_j(x) \right] \right\}. \end{aligned} \tag{6}$$

It can be readily seen that the functions $G(\cdot, \eta)$ and $H(\cdot)$ are convex, because the set of convex functions is a convex cone [4,8,11–14]. Hence, the objective function $F(x, \eta)$ turns out to be a d.c. function. Now we present the major result of this section.

Theorem 4.1 *Suppose that the point z is a solution to Problem $(\mathcal{P})$ and $\zeta := f_0(z)$. Then, for every pair $(y, \beta) \in I\!R^n \times I\!R$, such that*

$$H(y) = \beta, \tag{7}$$

the following inequality holds

$$G(x,\zeta)-\beta \geq \sum_{i=0}^{m}\langle h_i'(y), x-y\rangle \quad \forall x \in S, \tag{8}$$

with any subgradients $h_i'(y) \in \partial h_i(y)$ *of the functions* $h_i(\cdot)$ *at the point* y, $i = 0, 1, \ldots, m$.

Proof Since $z \in Sol(\mathcal{P})$, according to Proposition 3.1 $z \in Sol(\mathcal{P}_\zeta)$, so that

$$0 = F(z,\zeta) \leq F(x,\zeta) = G(x,\zeta) - H(x) \quad \forall x \in S,$$

whence, on account of (7), it follows $0 \leq G(x,\zeta) - \beta + H(y) - H(x) \quad \forall x \in S$.

Further, due to convexity of the functions $h_i(\cdot)$, $i = 0, \ldots, m$, we derive from the latter inequality that

$$G(x,\zeta)-\beta \geq \sum_{i=0}^{m}[h_i(x)-h_i(y)] \geq \sum_{i=0}^{m}\langle h_i'(y), x-y\rangle \quad \forall x \in S,$$

which coincides with the inequality (8). □

Remark 4.1 Note that Theorem 4.1 proposes to consider (for every pair $(y,\beta) \in I\!R^n \times I\!R$ satisfying (7)) the convex linearized problem $\left(\sum_{i=0}^{m} h_i'(y) =: H'(y) \in \partial H(y)\right)$

$$(\mathcal{P}L(y)): \qquad \min_x \Phi_\zeta(x) := G(x,\zeta) - \langle H'(y), x\rangle, \quad x \in S. \tag{9}$$

Besides, the linearization is performed with respect to the "unified" nonconvexity of Problem $(\mathcal{P})$ (or, which is the same, with respect to the basic nonconvexity of Problem $(\mathcal{P}_\zeta)$) given by the function $H(x) = \sum\limits_{i=0}^{m} h_i(x)$.

Hence, one can say that, in a sense, the message of Theorem 4.1 consists in reducing the nonconvex problem $(\mathcal{P})$ (with the help of Proposition 3.1) to the family of convex Problems $(\mathcal{PL}(y))$–(9) depending also on (y,β), as parameters. Besides, the inequality (8) can be rewritten, for example, as follows:

$$\mathcal{V}(\mathcal{PL}(y)) \geq \beta - \langle H'(y), y\rangle =: \mathcal{N}(y,\beta), \tag{10}$$

where $\mathcal{V}(\mathcal{PL}(y))$ is the optimal value of the linearized problem $(\mathcal{PL}(y))$ [4,8,11–14].

Remark 4.2 Suppose that there have been found a point $u \in S$ and a pair $(y,\beta) \in I\!R^n \times I\!R$: $H(y) = \beta$, such that (8) is violated, i.e., $0 > G(u,\zeta) - \beta - \langle H'(y), u-y\rangle$. Then, due to convexity of $H(\cdot) = \sum\limits_{i=0}^{m} h_i(\cdot)$, we obtain with the help of (7) $0 > G(u,\zeta) - \beta - H(u) + H(y) = G(u,\zeta) - H(u) = F(u,\zeta)$.

It yields $F(u,\zeta) < 0 = F(z,\zeta)$, and, in virtue of Lemma 3.1, we conclude that z is not a global solution to Problem $(\mathcal{P})$. Moreover, the point $u \in S$ is feasible for

Problem $(\mathcal{P})$ and, besides, it is better than the point z in question, since $f_0(z) = \zeta > f_0(u)$. Hence, the conditions (7)–(8) of Theorem 4.1 possess the classical constructive (algorithmic) property (once the conditions are violated, one can find a feasible point which is better than the point under consideration). Let us demonstrate the effectiveness of the constructive property by an example.

Example 4.1 Consider the problem

$$\min_x f_0(x) := -\frac{1}{2}x^2, \quad x \in I\!R, \quad f_1(x) := x^2 - x - 2 \le 0. \tag{11}$$

(a) Introduce the Lagrange function: $\mathcal{L}(x, \lambda) = -\frac{1}{2}x^2 + \lambda(x^2 - x - 2)$ and consider the KKT-system

$$\frac{\partial \mathcal{L}}{\partial x}(x, \lambda) = -x + 2\lambda x - \lambda = 0, \quad f_1(x) \le 0, \quad \lambda \ge 0, \quad \lambda f_1(x) = 0. \tag{12}$$

It is easy to see that the point $z_1 = -1$, $f_1(z_1) = 0$ satisfies (12) with $\lambda = \frac{1}{3}$.

(b) In order to verify whether the point $z_1 = -1$ is a global solution to (11) or not, let us apply the global optimality conditions (GOC) (7)–(8) of Theorem 4.1. First, it can be readily seen that

$$\left.\begin{array}{c} g_0(x) \equiv 0,\ h_0(x) := \dfrac{1}{2}x^2 = -f_0(x),\ g_1(x) = x^2 - x - 2 = f_1(x),\ h_1(x) \equiv 0, \\ \zeta := f_0(z_1) = -\dfrac{1}{2}, \quad H(y) = h_0(x) = \dfrac{1}{2}x^2. \end{array}\right\} \tag{13}$$

Therefore, from (13) and (6), we obtain $G(x, \zeta) := \max\{g_0(x) + h_1(x) - \zeta; g_1(x) + h_0(x)\} = \max\{\frac{1}{2}; \frac{3}{2}x^2 - x - 2\}$.

(c) Let us set $y := 1$. Then we have, according to (7), $\beta = H(y) = \dfrac{1}{2}y^2 = \dfrac{1}{2}$.

Furthermore, for the point $u = \dfrac{4}{3}$, we have $\dfrac{3}{2}u^2 - u - 2 = -\dfrac{2}{3} < \dfrac{1}{2}$, and $G(u, \zeta) = \dfrac{1}{2}$.

Besides, $\nabla H(y) = y = 1$, $\langle \nabla H(y), u - y \rangle = \frac{1}{3}$, $G(u, \zeta) = \frac{1}{2} < \frac{1}{2} + \frac{1}{3} = \beta + \langle \nabla H(y), u - y \rangle$, so that the inequality (8) is violated. Moreover, $F(z, \zeta) = 0 > -\frac{4}{9} = F(u, \zeta)$.

Hence, according to Theorem 4.1 and Lemma 3.1, the point $z_1 = -1$ is not a global solution to (11). Actually, $f_0(u) = -\dfrac{1}{2}u^2 = -\dfrac{8}{9} < -\dfrac{1}{2} = f_0(z)$. So, the point $u = \dfrac{4}{3}$ violating (8) turns out to be better in the sense of Problem $(\mathcal{P})$ than the critical point $z_1 = -1$ provided by the KKT-system (12). □

Remark 4.3 Let us set in (7) $y = z$. Then, we have $\beta = H(z) = \sum_{i=0}^{m} h_i(z)$. Besides, from (8) it follows

$$G(x,\zeta) - \langle H'(z), x\rangle \geq \beta - \langle H'(z), z\rangle = H(z) - \langle H'(z), z\rangle \ \ \forall x \in S.$$

It means that for the linearized problem $(\mathcal{P}L(z))$–(9), the number $\varkappa := H(z) - \langle H'(z), z\rangle$ is a lower bound, so that $\Phi_\zeta(x) \geq \varkappa \ \forall x \in S$. On the other hand, it can be readily seen that

$$\begin{aligned} G(z,\zeta) := \max &\left\{ g_0(z) + \sum_{j=1}^{m} h_j(z) - \zeta;\ \ g_1(z) + \sum_{j\neq 1}^{m} h_j(z); \ldots, g_m(z) + \sum_{j}^{m-1} h_j(z) \right\} \\ = \max &\left\{ H(z);\ \max_{1\leq i\leq m} [f_i(z) + H(z)] \right\} = H(z), \end{aligned} \tag{14}$$

because $\max_{1\leq i\leq m} [f_i(z)] \leq 0$. Therefore, $\Phi_\zeta(z) := G(z,\zeta) - \langle H'(z), z\rangle = H(z) - \langle H'(z), z\rangle = \varkappa$.

So, z is a solution to the convex Problem $(\mathcal{PL}(z))$, and since the set S is convex, the corresponding necessary and sufficient conditions [4,8–14] hold

$$0 \in \partial G(z,\zeta) - \sum_{i=0}^{m} h_i'(z) + N(z\ ;\ S), \tag{15}$$

where $\partial G(z,\zeta)$ is the subdifferential of $G(\cdot,\zeta)$ at z, and $N(z\ ;\ S)$ is the normal cone at z to S [4,8–14]. It means that the global optimality conditions (GOC) (7)–(8) of Theorem 4.1 are somehow connected with the classical optimization theory [1–3]. □

In order to open the possibility to construct numerical methods of global search, such results as Theorem 4.1 can be used in the contrariwise form as follows.

Theorem 4.2 *Let us be given a feasible in Problem* $(\mathcal{P})$ *point* z, $f_0(z) =: \zeta$, *and, in addition, the following assumption holds:*

$$(\mathcal{H}): \qquad \exists v \in I\!R^n : F(v,\zeta) > F(z,\zeta) = 0. \tag{16}$$

If z *is not a solution to Problem* $(\mathcal{P}_\zeta)$, *then one can find a pair* $(y,\beta) \in I\!R^n \times I\!R$, *a point* $u \in \mathcal{D}$ *and a collection of subgradients* $\{h'_{00}(y), h'_{10}(y), \ldots, h'_{m0}(y)\}$, *where* $h'_{i0}(y) \in \partial h_i(y)$, $i = 0, 1, \ldots, m$, *such that*

$$H(y) := \sum_{i=0}^{m} h_i(y) = \beta, \tag{7}$$

$$G(y,\zeta) \leqslant \beta, \tag{17}$$

$$G(u,\zeta) - \beta < \sum_{i=0}^{m} \langle h'_{i0}(y), u - y \rangle. \tag{18}$$

Proof 1) Due to the assumption, there exists a point $u \in I\!R^n$, such that

$$u \in S,\ F(u,\zeta) < F(z,\zeta) = 0, \quad G(u,\zeta) < H(u). \tag{19}$$

As above, we see that $f_i(u) < 0, i \in I, f_0(u) < \zeta$, so that $u \in \mathcal{D}$ and the point u is also better than $z \in \mathcal{D}$ in the sense of Problem $(\mathcal{P})$. Let us introduce the following convex set C

$$C := epi\, H(\cdot) = \{(x,\gamma) \ :\ H(x) \leqslant \gamma\} \subset I\!R^n \times I\!R. \tag{20}$$

It is easy to see now that the latter inequality in (19) is equivalent to the relation

$$(u, G(u,\zeta)) \notin C = epi\, H. \tag{21}$$

Due to the assumption $(\mathcal{H})$–(16), we have $G(v,\zeta) > H(v)$, which is equivalent to the inclusion

$$(v, G(v,\zeta)) \in int\, C = int[epi\, H]. \tag{22}$$

2) Furthermore, due to the properties of the convex set C, and on account of (21) and (22), there exists a number α, $0 < \alpha < 1$, such that $(y,\alpha) = \alpha(u, G(u,\zeta)) + (1-\alpha)(v, G(v,\zeta)) \in bdC$, where $bd\, C = \{(x,\gamma) : H(x) = \gamma\}$ stands for the boundary of the convex set C. Hence, we have

$$y = \alpha u + (1-\alpha)v, \quad \beta = \alpha G(u,\zeta) + (1-\alpha)G(v,\zeta) = H(y). \tag{23}$$

Thus, the pair (y,β) satisfies (7). In addition, with the help of convexity of the function $x \mapsto G(x,\zeta)$, we obtain that $G(y,\zeta) \leqslant \alpha G(u,\zeta) + (1-\alpha)G(v,\zeta) = \beta$, so that (17) is also proven. Besides, from (23), it follows

$$u = \alpha^{-1}[y - (1-\alpha)v], \quad G(u,\zeta) = \alpha^{-1}[\beta - (1-\alpha)G(v,\zeta)]. \tag{23$'$}$$

3) Suppose now that, despite the assertion of Theorem 4.2, there does not exist a tuple $h'_{i0}(y) \in \partial h_i(y), i = 0, 1, \ldots, m$, such that (18) takes place with the pair (y,β) constructed above, satisfying (7) and (17), and with the feasible point u. It means that $G(u,\zeta) - \beta \geqslant \sum_{i=0}^{m} \langle h'_i(y), u - y \rangle$ for any subgradients $h'_i(y) \in \partial h_i(y), i = 0, 1, \ldots, m$. Then, we obtain with the help of the presentation (23$'$)

$$\begin{aligned} 0 &\geqslant \beta - \alpha^{-1}[\beta - (1-\alpha)G(v,\zeta)] + \langle H'(y), \alpha^{-1}[y - (1-\alpha)v] - y \rangle \\ &= \frac{1-\alpha}{\alpha}[G(v,\zeta) - \beta] + \frac{1-\alpha}{\alpha}\langle H'(y), y - v \rangle, \end{aligned}$$

where $H'(y) := \sum_{i=0}^{m} h_i'(y)$. Furthermore, due to convexity of the function $H(\cdot)$ and the assumption $(\mathcal{H})$–(16), and with the help of (7), we derive that

$$0 \geqslant \frac{1-\alpha}{\alpha}[G(v,\zeta)-\beta+H(y)-H(v)] = \frac{1-\alpha}{\alpha}[G(v,\zeta)-H(v)] = \frac{1-\alpha}{\alpha}F(v,\zeta) > 0,$$

which is impossible. Hence, the hypothesis in the very beginning of the part 3) of the proof turns out to be wrong. So, (18), together with Theorem 4.2, are proved. □

Remark 4.4 As it was shown in Remark 4.1, Theorem 4.1 proposes to reduce Problem $(\mathcal{P}_\zeta)$ to the family of linearized problems $(\mathcal{P}L(y))$–(9) depending on the parameters $(y,\beta) \in I\!R^n \times I\!R$, satisfying (7), with the consecutive verification of (8). On the other hand, it was pointed out that Problem $(\mathcal{P}L(y))$–(9) is not smooth, while the data of Problem $(\mathcal{P})$ can be differentiable. It turns out that this shortcoming of the approach can be easily overcome with the help of the following result (see, for instance, [1,3] and [14], exercise 3.32).

Consider the following problem

$$(Q): \qquad \left.\begin{array}{c} \min\limits_x \Theta(x) := \sigma(x) + f(x), \quad x \in S, \\ \sigma(x) := \max\limits_j \big\{\varphi_j(x) \ : \ j \in J = \{1,\ldots,N\}\big\}. \end{array}\right\} \tag{24}$$

Together with Problem (Q)–(24), let us consider the auxiliary problem

$$(QA): \qquad \min_{(x,t)} \Phi(x,t) := t + f(x), \quad x \in S,\ t \in I\!R, \quad \varphi_j(x) \leqslant t,\ j \in J. \tag{25}$$

Lemma 4.1 *[1,3,14] A point z is a solution to Problem (Q)–(24) if and only if the pair (z,t_*) is a solution to Problem (QA)–(25), where*

$$t_* = \sigma(z) = \max_j\{\varphi_j(z) \ : \ j \in J\}. \tag{26}$$

Lemma 4.2 *Let the quadratic function $q(x) := \frac{1}{2}\langle x, Ax\rangle - \langle b, x\rangle$ with positive definite matrix $A = A^\top$ and $b \in I\!R^n$ be given. Consider the optimization problem (with a parameter $u \in I\!R^n$)*

$$\max_{(y,\beta)} \Theta(y,\beta) := \beta + \langle \nabla q(y), u - y\rangle, \quad (y,\beta) \in I\!R^{n+1}: \quad q(y) = \beta - \gamma. \tag{27}$$

Then, the solution (y_*,β_*) to the problem (27) is provided by the equalities

$$y_* = u, \quad \beta_* = q(y_*) + \gamma. \tag{28}$$

Proof Due to the equality constraint in (27), it can be readily seen that $\Theta(y,\beta) = q(y) + \gamma + \langle \nabla q(y), u - y\rangle = \frac{1}{2}\langle y, Ay\rangle - \langle b, y\rangle + \gamma + \langle Ay - b, u - y\rangle =: \psi(y) =$

$\gamma - \frac{1}{2}\langle y, Ay\rangle + \langle Ay, u\rangle - \langle b, u\rangle$. Hence, the function $y \mapsto \psi(y)$ is strongly concave, and the problem (27) amounts to the following unconstrained one: $\max_y \psi(y), \quad y \in I\!R^n$, which can be solved by the equation $\nabla\psi(y) = A(u-y) = 0$ that proves the relations (28). □

Let us now demonstrate the effectiveness of Theorem 4.2, Lemmas 4.1 and 4.2 by examples.

Example 4.2 (See example 12.9 in [1]) Consider the following problem ($x \in I\!R^2$)

$$\min_x f_0(x) = x_2^2 - 0,1(x_1-4)^2, \quad f_1(x) = 1 - x_1^2 - x_2^2 \leqslant 0, \; x \in S := [-2,4] \times I\!R. \tag{29}$$

It can be readily seen that $g_0(x) = x_2^2$, $h_0(x) = 0.1(x_1-4)^2$, $g_1(x) \equiv 1$, $h_1(x) = x_1^2 + x_2^2$. Besides, according to (5)–(6), we have

$$H(x) = 0,1(x_1-4)^2 + x_1^2 + x_2^2, \tag{30}$$

$$\begin{aligned} G(x,\eta) &= \max\{f_0(x) + H(x) - \eta, \, f_1(x) + H(x)\} \\ &= \max\{2x_2^2 + x_1^2 - \eta; \, 1 + 0.1(x_1-4)^2\}. \end{aligned} \tag{31}$$

Note that (see [1]) the point $z = (1,0)^\top$ satisfies the KKT-conditions with $\lambda_1 = 0.3$. Let us verify whether the point z is a global solution to (29). First, according to Theorems 4.1 and 4.2, we have to choose parameters $(y,\beta) \in I\!R^2 \times I\!R$, fulfilling the quadratic equality

$$\beta = H(y) = y_1^2 + y_2^2 + 0.1(y_1-4)^2, \tag{30$'$}$$

and a point $u \in S$ in order to violate (8) so that $G(u,\zeta) < \beta + \langle\nabla H(y), u-y\rangle$. Here $\zeta := f_0(z) = -0.1(z_1-4)^2 = -0.9$ and, besides, $G(u,\zeta) = \max\left\{u_1^2 + 2u_2^2 + 0.9; \, 1 + 0.1(u_1-4)^2\right\}$.

Let us look now for points $y = (y_1, y_2)$ and $u = (u_1, u_2)$ such that $y_2 = 0 = u_2$.

Then, denoting $y := y_1$ and $u := u_1$, we have $\beta = H((y,0)) = y^2 + 0.1(y-4)^2 = 1.1y^2 - 0.8y + 1.6$. Due to Lemma 4.2, we set $y_* = u$, and in order to find a point $u \in S$, we try to solve the problem

$$\min_u G(u,\zeta) = \max\left\{u^2 + 0.9; \, 1 + 0.1(u-4)^2\right\}, \quad u \in [-2,4]. \tag{32}$$

Furthermore, due to Lemma 4.1, the problem (32) amounts to the following one

$$\begin{aligned} &\min_{(u,t)} t, \quad u \in [-2,4], \; t \in I\!R, \quad \varphi_0(u,t) := u^2 + 0.9 - t \leqslant 0, \\ &\varphi_1(u,t) := 0.1(u-4)^2 + 1 - t \leqslant 0. \end{aligned} \tag{32$'$}$$

This problem is obviously convex, and therefore, let us seek for a solution of the KKT-system satisfying $-2 < u < 4$, which yields us $\mathcal{L}(u,t,\mu) = t + \mu_0\left(u^2 + \frac{9}{10} - t\right) + \mu_1[1 + 0.1(u-4)^2 - t]$,

$$\left.\begin{array}{ll}(a)\ \frac{\partial\mathcal{L}}{\partial t}=1-\mu_0-\mu_1=0;\ \mu_0\geqslant 0,\ \mu_1\geqslant 0; & (b)\ \frac{\partial\mathcal{L}}{\partial u}=2\mu_0 u+0.2\mu_1(u-4)=0;\\ (c)\ \mu_0(u^2+0.9-t)=0; & (d)\ \mu_1[1+0.1(u-4)^2-t]=0.\end{array}\right\} \tag{33}$$

Due to $\mu_0+\mu_1=1$, the condition (b) yields the equality $0.9\mu_0 u+0.1u+0.4\mu_0-0.4=0$. Then, the cases i) $\mu_0=1$, $\mu_1=0$ and ii) $\mu_0=0$, $\mu_1=1$ lead us to points i) $(u,t)=(0;0.9)^\top$ and ii) $(u,t)=(4,1)^\top$ which are unfeasible for ($32'$).

iii) $0<\mu_0,\mu_1<1$, $\mu_0+\mu_1=1$. In this case, according to (33)(c)(d), we obtain the system of quadratic equations: $u^2+0.9=t,\quad 1+0.1(u^2-8u+16)=t$, which can be reduced to the single equation $0.9u^2+0.8u-1.7=0$. The latter equation has two solutions $u=1$ and $u=-\frac{17}{9}=-1\frac{8}{9}>-2$. The first one produces the point $z=(1,0)^\top$ under consideration, so it is not interesting for us.

Now, let us verify (18) only for $u=-\frac{17}{9}=-1\frac{8}{9}$, for which $G(u,\zeta)=t_*=u^2+0.9=1+0.1(u-4)^2\approx 4.4679$. On the other hand ($y_1^*=u_1=-\frac{17}{9}$, $y_2^*=u_2=0$), we have

$$\begin{aligned}\theta(y_*,u_*)&:=\beta+\langle\nabla H(y),u-y\rangle=1.1u^2-0.8u+1.6>5.4+1.6\\&=7.0>4.4679\approx G(u,\zeta).\end{aligned}$$

So, we see that (8) has been violated. Hence, according to Theorems 4.1 and 4.2, the point $z=(1,0)^\top$ with $f_0(z)=-0.9$ is not a global solution to Problem $(\mathcal{P}_\zeta)$. At the same time, due to Lemma 3.1, it is not a solution to (29), which is confirmed by the inequality $f_0(z)=-0.9>f_0((-\frac{17}{9},0))\approx -3.4679$.

To sum up, one can say that the global optimality conditions (GOC) of Theorems 4.1 and 4.2 allow us not only to show that the stationary (critical) point $z=(1,0)^\top$ is not a global solution to Problem (29), but, in addition, to construct a point $u=(-\frac{17}{9},0)^\top$, which is better than $z=(1,0)^\top$ and closer to the global solution $(-2,0)^\top\in Sol(29)$. □

Remark 4.5 (Remark 4.3 Revisited) Recall that $z\in Sol(\mathcal{P})$ is a solution to the problem

$$(\mathcal{P}L(z)):\qquad \min_x \Phi_\zeta(x):=G(x,\zeta)-\langle H'(z),x\rangle,\quad x\in S, \tag{9'}$$

and the following equalities hold $G(z,\zeta)=H(z)$, $\mathcal{V}(\mathcal{P}L(z)):=\Phi_\zeta(z)=H(z)-\langle H'(z),z\rangle=:\varkappa$.

All this yielded the inclusion (15). Now we will use another approach.

For that purpose, suppose that $S=I\!R^n$, and all the functions $g_i(x)$, $h_i(\cdot)$, $i=0,1,\ldots,m$, are differentiable on $I\!R^n$. Then, due to Lemma 4.1, Problem $(\mathcal{PL}(z))$–($9'$) is equivalent to the following problem

$$\left.\begin{array}{c}\min\limits_{(x,t)}\Big(t-\sum\limits_{i=0}^{m}\langle\nabla h_i(z),x\rangle\Big),\quad (x,t)\in I\!R^{n+1},\\ f_0(x)+H(x)-\zeta-t\leqslant 0,\quad f_i(x)+H(x)-t\leqslant 0,\ i=1,\ldots,m.\end{array}\right\} \tag{34}$$

The latter problem is obviously convex. Let us introduce the Lagrange function $\mathcal{L}_Q(x,t,\lambda)$ for Problem (34) (where $\lambda=(\lambda_0,\lambda_1,\ldots,\lambda_m)^T\in I\!R_+^{m+1}$) as follows

$$\mathcal{L}_Q(x,t,\lambda) := t - \langle \nabla H(z), x\rangle + \lambda_0[f_0(x) + H(x) - \zeta - t] + \sum_{i=1}^{m} \lambda_i [f_i(x) + H(x) - t].$$

Note that here the multiplier, relative to the cost function in (34), has been taken equal to 1, because in Problem (34), the Slater assumption is always satisfied with any $x \in S$ and rather large $\hat{t}$ such that

$$f_0(x) + H(x) - \zeta - \hat{t} < 0, \quad f_i(x) + H(x) - \hat{t} < 0, i = 1, \dots, m.$$

Thus, for the pair (z, t_*), $t_* = G(z,\zeta) = H(z)$ to be a solution to Problem (34), the KKT-theorem provides existence of Lagrange multipliers $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_m)^T \in I\!R^{m+1}_+$, such that the following conditions hold

$$(a)\ \frac{\partial \mathcal{L}_Q(z, t_*, \lambda)}{\partial t} = 0 \in I\!R, \qquad (b)\ \frac{\partial \mathcal{L}_Q(z, t_*, \lambda)}{\partial x} = 0_n \in I\!R^n; \tag{35}$$

$$\begin{aligned} &\lambda_0[f_0(z) + H(z) - \zeta - t_*] = \lambda_0[H(z) - t_*] = 0\\ &\lambda_i[f_i(z) + H(z) - t_*] = \lambda_i f_i(z) = 0, i = 1, \dots, m. \end{aligned} \tag{36}$$

From (35), it simply follows

$$\left.\begin{aligned} (a)\ \sum_{i=0}^{m} \lambda_i = 1,\ \ (b)\ 0_n &= -\nabla H(z) + \sum_{i=0}^{m} \lambda_i [\nabla H(z) + \nabla f_i(z)]\\ &= \sum_{i=0}^{m} \lambda_i \nabla f_i(z) = \sum_{i=1}^{m} \lambda_i [\nabla g_i(z) - \nabla h_i(z)]. \end{aligned}\right\} \tag{35$'$}$$

Hence, from the conditions (7)–(8) of Theorem 4.1, we easily derived, in particular $(y = z)$, the classical KKT-results for Problem $(\mathcal{P})$. So, we have ascertained the real connection between the optimality conditions (7)–(8) of Theorem 4.1 and the classical optimization theory. □

Now let turn to the case when the necessary conditions (7)–(8) of Theorem 4.1 become sufficient.

Theorem 4.3 *Suppose that for a feasible (in Problem $(\mathcal{P})$) point $z \in D$, $\zeta := f_0(z)$, the assumption $(\mathcal{H})$–(16) is fulfilled. Suppose, in addition, that for every pair $(y, \beta) \in I\!R^n \times I\!R$, such that*

$$(a)\ H(y) = \sum_{i=0}^{m} h_i(y) = \beta, \qquad (b)\ G(y,\zeta) \le \beta, \tag{37}$$

there exists a fixed collection of subgradients $\{h'_{0*}(y), h'_{1*}(y), \dots, h'_{m*}(y)\}$, $h'_{i*}(y) \in \partial h_i(y)$, $i = 0, 1, \dots, m$, for which the following inequality holds

$$G(x,\zeta) - \beta \geq \sum_{i=0}^{m} \langle h'_{i*}(y), x - y \rangle \quad \forall x \in S. \tag{38}$$

Then, the point $z \in D$ is a global solution to Problem $(\mathcal{P}_\zeta)$: $z \in Sol(\mathcal{P}_\zeta)$.

Proof 1) Suppose that despite the assertion of Theorem 4.3, there exists a point $u \in I\!R^n$ such that,

$$u \in S, \quad F(u,\zeta) < F(z,\zeta) = 0, \tag{19}$$

meanwhile the condition (37)–(38) takes place. Furthermore, we perform the similar operations as it was done in the proof of Theorem 4.2 until the part 3).

3) Since the pair (y,β) constructed in parts 1) and 2) of the proof of Theorem 4.2 satisfies (23) and (37) (see (7) and (17)), then, according to the assumption of Theorem 4.3, (38) must hold with $x = u \in S$ and some ensemble of subgradients $h'_{i*}(y) \in \partial h_i(y)$, $i = 0, 1, \ldots, m$, so that

$$0 \geq \beta - G(u,\zeta) + \sum_{i=0}^{m} \langle h'_{i*}(y), u - y \rangle. \tag{39}$$

The latter inequality leads us to the absurdity $0 \geq \alpha^{-1}(1-\alpha)F(v,\zeta) > 0$, as it was shown in the proof of Theorem 4.2. Thus, Theorem 4.3 is proved. □

Remark 4.6 Let us now pay attention to the fact that Theorem 4.3 provides the sufficient conditions for $z \in S$ to be a solution to Problem $(\mathcal{P}_\zeta)$, $\zeta = f_0(z)$, but not for the original Problem $(\mathcal{P})$. Only if we added the equality $\zeta = \mathcal{V}(\mathcal{P})$ (which is, in particular, rather difficult to verify in the majority of the applied problems, see Proposition 3.2), then we would be able to make a conclusion about the global solution property in Problem $(\mathcal{P})$ of the feasible point $z \in D$ in question. On the other hand, the equality $f_0(z) = \zeta = \mathcal{V}(\mathcal{P})$ for a feasible point $z \in D$ immediately provides that $z \in Sol(\mathcal{P})$ without any supplementary conditions. □

Example 4.3 Consider the problem

$$\left.\begin{array}{c} \min\limits_{x} f_0(x) = \frac{1}{2}(x_1-4)^2 + (x_2+2)^2, \\ f_1(x) = (x_1-1)^2 - (x_2+1)^2 \leq 0, \quad f_2(x) = (x_2-2)^2 - (x_1+2)^2 \leq 0. \end{array}\right\} \tag{40}$$

It is easy to see that the point $z_* = (4,-2)^\top$ is the global minimum of the strongly convex function $f_0(\cdot)$ on $I\!R^2$, and $f_0(z_*) = 0$ provides a lower bound for $\mathcal{V}(40) = \inf(f_0, D) \geq 0$. Note that z_* is unfeasible in (40), since $f_1(z_*) = 8 > 0$. Let us consider another point $z = \left(\frac{4}{3}, -\frac{2}{3}\right)^\top$ which is feasible for (40), since $f_1(z) = 0$ and $f_2(z) = -4 < 0$. Besides, it can be readily seen that z satisfies the KKT-conditions with $\lambda_0 = 1$, $\lambda_1 = 4 > 0$, $\lambda_2 = 0$, $\zeta := f_0(z) = 5\frac{1}{3}$.

In addition, because $f_1(x) = (x_1 - x_2 - 2)(x_1 + x_2) \le 0$, $f_2(x) = (x_2 - x_1 - 4)(x_2 + x_1) \le 0$, it can be readily shown that the feasible set $D := \{x \in I\!R^2 \ : \ f_i(x) \le 0, \quad i = 1, 2\}$ is represented by the union of two convex parts: $D = D_1 \cup D_2$, $D_1 = \{x \ : \ x_1 + x_2 = 0\}$, $D_2 = \{x \ : \ x_1 + x_2 \ge 0, \quad x_2 - x_1 - 4 \le 0, \quad x_1 - x_2 - 2 \le 0\}$.

Hence, from the geometrical view point, it is easy to see that the point $z_0 = (\frac{8}{3}, -\frac{8}{3})^\top$ is the global solution to the problem (40) with the optimal value $\mathcal{V}(40) = f_0(z_0) =: \zeta_0 = \frac{4}{3} = 1\frac{1}{3}$. So, the point $z = (\frac{4}{3}, -\frac{2}{3})^\top$ turns out not to be a global solution to (40), since $f_0(z) = 5\frac{1}{3} = \zeta$. However, the goal function $F(x, \zeta)$ of Problem $(\mathcal{P}_\zeta)$ does not distinguish between these two points. Actually,

$$\begin{aligned} F(z_0, \zeta) &= \max\{f_0(z_0) - \zeta;\ f_1(z_0),\ f_2(z_0)\} = 0 \\ &= \max\{f_0(z) - \zeta;\ f_1(z);\ f_2(z)\} = F(z, \zeta), \end{aligned}$$

because $z \in D$, $z_0 \in D = \{x \in I\!R^2 \ : \ f_i(x) \le 0, \quad i = 1, 2\}$.

Moreover, for all feasible (in Problem (40)) points which are better (in the sense of the problem (40)) than the point z, i.e., $u \in \{x \in I\!R^2 \ : \ x_1 + x_2 = 0, \quad f_0(x) < \zeta = 5\frac{1}{3}\}$, we have the same results: $F(u, \zeta) = 0$, because $f_1(u) = 0 = f_2(u)$. For instance, for any point $x(\alpha)$ of the form

$$x(\alpha) = (v(\alpha), -v(\alpha))^\top, \quad v(\alpha) = 1.1\alpha + 4.2(1 - \alpha), \quad \alpha \in [0, 1],$$

we have $F(x(\alpha), \zeta) = 0$, meanwhile $f_0(x(\alpha)) < f_0(z) = 5\frac{1}{3}\ \forall \alpha \in [0, 1]$.

On the other hand, one can show that the GOC of Theorems 4.1 and 4.2 fail to improve the point $z = \left(\frac{4}{3}, -\frac{2}{3}\right)^\top$, albeit there exist a lot of points which are better than z in the sense of Problem (40). □

So, Example 4.3 demonstrates that Problem $(\mathcal{P}_\eta)$ is not sufficiently adequate to Problem $(\mathcal{P})$. More precisely, taking into account Propositions 3.1 and 3.2, it is easy to see that the set $Sol(\mathcal{P}_\zeta)$ may contain a lot of points which do not belong to $Sol(\mathcal{P})$, so that the inclusion $Sol(\mathcal{P}) \subset Sol(\mathcal{P}_\zeta)$ may be really proper. Moreover, the inequality $\zeta > \zeta_0 = \mathcal{V}(\mathcal{P})$ holds among with the inclusion $Sol(\mathcal{P}) \subset Sol(\mathcal{P}_\zeta)$, which is inadmissible. Therefore, we move to another type of merit (or penalty, in a rough sense) function.

5 The Lagrange Function

Consider the standard (normal) Lagrange function for Problem $(\mathcal{P})$: $\mathcal{L}(x, \lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$.

It is common to call a pair (z, λ) a saddle point of the Lagrange function $\mathcal{L}(x, \lambda)$: $(z, \lambda) \in Sdl(\mathcal{L})$, if the following inequalities are satisfied [4,8–14]:

$$\forall \mu \in I\!R^m_+ \quad \mathcal{L}(z, \mu) \le \mathcal{L}(z, \lambda) \le \mathcal{L}(x, \lambda) \quad \forall x \in S. \tag{41}$$

Lemma 5.1 *[4,8,11–14] For a pair $(z, \lambda) \in S \times I\!R^m_+$, the following two assertions are equivalent:*

$$i)\quad \max_{\mu}\{\mathcal{L}(z, \mu)\ :\ \mu \in I\!R^m_+\} = \mathcal{L}(z, \lambda); \tag{42}$$

$$ii)\quad z \in D, \quad \lambda \in I\!R^m_+, \quad \lambda_i f_i(z) = 0, \quad i = 1, \ldots, m. \tag{43}$$

Recall that a vector $\lambda \in I\!R^m_+$ satisfying the KKT-conditions, including (43), is called a Lagrange multiplier [4,8,11–14] at a point $z \in D$. The set of all Lagrange multipliers at z will be denoted by $\mathcal{M}(z)$. Remember, in addition, that for a convex optimization problem of type $(\mathcal{P})$–(1), when $h_i(x) \equiv 0\ \forall i \in \{0\} \cup I$, we have $\mathcal{M}(z_1) = \mathcal{M}(z_2) = \mathcal{M}$, if $z_i \in Sol(\mathcal{P})$, $i = 1, 2$ [4].

Proposition 5.1 *[4,8,11,13,14] If the pair $(z, \lambda) \in S \times I\!R^m_+$ is a saddle point of the Lagrange function $\mathcal{L}(x, \mu)$ on the set $S \times I\!R^m_+$, then the point z is a global solution to Problem $(\mathcal{P})$.*

We will employ this assertion in what follows in another form.

Proposition 5.2 *Suppose $z \in D$, z is a KKT-point to $(\mathcal{P})$, but it is not a global solution to Problem $(\mathcal{P})$.*

Then, there does not exist a Lagrange multiplier $\lambda \in \mathcal{M}(z)$ such that $(z, \lambda) \in Sdl(\mathcal{L})$.

Furthermore, since $f_i(x) = g_i(x) - h_i(x)$, $i = 0, 1, \ldots, m$, $\mathcal{L}(x, \lambda)$ has a very clear d.c. representation

$$\left.\begin{array}{ll} (a) & \mathcal{L}(x, \lambda) = G_\lambda(x) - H_\lambda(x), \\ (b) & G_\lambda(x) = g_0(x) + \sum\limits_{i=1}^{m} \lambda_i g_i(x), \quad H_\lambda(x) = h_0(x) + \sum\limits_{i=1}^{m} \lambda_i h_i(x). \end{array}\right\} \tag{44}$$

Taking into account (44), let us look at the normal Lagrange function from the view point of the global optimality conditions (GOC) [15–19].

Theorem 5.1 *Suppose $(z, \lambda) \in Sdl(\mathcal{L})$, $\lambda_0 = 1$, $\zeta := f_0(z)$. Then, $\forall (y, \beta) \in I\!R^n \times I\!R$ such that*

$$H_\lambda(y) := \sum_{i=0}^{m} \lambda_i h_i(y) = \beta - \zeta, \tag{45}$$

the following inequality holds

$$G_\lambda(x) - \beta \geq \sum_{i=0}^{m} \lambda_i \langle h_i'(y), x - y\rangle \quad \forall x \in S, \tag{46}$$

for any subgradients $h_i'(y) \in \partial h_i(y)$ of the functions $h_i(\cdot)$ at the point y, $i \in I \cup \{0\}$.

Proof According to the assumption, we have the chain

$$\zeta := f_0(z) = \sum_{i=0}^{m} \lambda_i f_i(z) = \mathcal{L}(z, \lambda) \leq \sum_{i=0}^{m} \lambda_i f_i(x) = \mathcal{L}(x, \lambda) \quad \forall x \in S,$$

from which, due to (44) and (45), it follows $\beta - H_\lambda(y) = \zeta \leq \mathcal{L}(x, \lambda) = G_\lambda(x) - H_\lambda(x) \quad \forall x \in S$.

Then, by the convexity of $H_\lambda(\cdot) = \sum_{i=0}^{m} \lambda_i h_i(\cdot), \lambda_i \geq 0, i \in I$, we obtain

$$G_\lambda(x) - \beta \geq H_\lambda(x) - H_\lambda(y) \geq \sum_{i=0}^{m} \lambda_i \langle h_i'(y), x - y \rangle \quad \forall x \in S,$$

which coincides with (46). □

Remark 5.1 Due to Proposition 5.1, it can be readily seen that for a global solution $z \in Sol(\mathcal{P})$, for which one can find a multiplier $\lambda \in \mathcal{M}(z)$ such that $(z, \lambda) \in Sdl(\mathcal{L})$, the global optimality conditions (GOC) (45)–(46) turn out to be necessary optimality conditions.

Remark 5.2 As it was in Theorem 4.1, we see that Theorem 5.1 reduces the nonconvex problem

$$(\mathcal{L}): \qquad \min_x \mathcal{L}(x, \lambda), \quad x \in S,$$

to the verification of (46) for the family of parameters (y, β): $H_\lambda(y) = \beta - \zeta$, or, more precisely, to solving the family of the convex linearized problems

$$(\mathcal{L}L(y)): \qquad \min_x \Phi_\lambda(x) = G_\lambda(x) - \langle H_\lambda'(y), x \rangle, \quad x \in S, \tag{47}$$

with the subsequent verification of (46) with $x = u = u(y, \beta) \in Sol(\mathcal{L}L(y))$.

Example 4.1 (Revisited) Taking into account that $z_1 = 1$ is a KKT-point with $\lambda_1 = \frac{1}{3}$, $f_0(z_1) = -\frac{1}{2} =: \zeta$, we get $\mathcal{L}(x, \lambda) = -\frac{1}{2}x^2 + \frac{1}{3}(x^2 - x - 2) = G_\lambda(x) - H_\lambda(x)$, $G_\lambda(x) = \frac{1}{3}(x^2 - x - 2)$, $H_\lambda(x) = \frac{1}{2}x^2$, $H_\lambda'(x) = x$. By setting $y = 0$, we obtain $H_\lambda(y) = 0 = \beta - \zeta$, whence we derive $\beta = \zeta = -\frac{1}{2}$, $H_\lambda'(y) = y = 0$. Then, the convex linearized (at y) problem

$$(\mathcal{L}L(y)): \qquad \min_x \Phi_\lambda(x) := G_\lambda(x) - \langle H_\lambda'(y), x \rangle = \frac{1}{3}(x^2 - x - 2) - \langle y, x \rangle, \quad x \in I\!R,$$

yields the point $u = \frac{1}{2}$, which is feasible: $f_1(u) = 0.25 - 0.5 - 2 = -2.25 < 0$. Then, $\theta(y, \beta) := \beta + \langle H_\lambda'(y), u - y \rangle = \beta = -\frac{1}{2} > -\frac{3}{4} = \frac{1}{3}(u^2 - u - 2) = G_\lambda(u)$. The latter inequality means that (46) is violated, so that $(z_1, \lambda_1) = (-1, \frac{1}{3})$ is not a saddle point. In addition, the GOC of Theorem 5.1 provided the point $u = \frac{1}{2}$, with $f_0(u) = -\frac{1}{8}$, which is worse in the sense of Problem (29) in comparison with $z_1 = -1$, since

$f_0(z_1) = -\frac{1}{2} < f_0(u) = -\frac{1}{8}$. On the other hand, matching the values of the Lagrange function, we see that $\mathcal{L}(u, \lambda_1) := -\frac{1}{2}u^2 + \frac{1}{3}(u^2 - u - 2) = -\frac{7}{8} < \mathcal{L}(z_1, \lambda_1) = -\frac{1}{2}$. So, the improvement happened, but at the value of the Lagrange function. □

Remark 5.3 Let us now set in (45) $y = z$. Then, we have $\beta = f_0(z) + \sum_{i=0}^{m} \lambda_i h_i(z) = \mathcal{L}(z, \lambda) + H_\lambda(z) = G_\lambda(z)$. Furthermore, from the inequality (46), it follows $G_\lambda(x) - G_\lambda(z) \geq \langle H'_\lambda(z), x - z\rangle \quad \forall x \in S$.

It means that the point z is a solution to the convex problem $(\mathcal{L}L(y))$–(47) with $y = z$, $z \in Sol(\mathcal{L}L(z))$.

Hence, the following optimality condition takes place [4,8,11–14]

$$0_* \in \partial G_\lambda(z) - H'_\lambda(z) + N(z; S). \tag{48}$$

The inclusion (48) in the smooth case is very standard for experts in the optimization theory [4,8–11,13,14]. Actually, in the smooth case where $z \in \text{int}\, S$ or $S = I\!R^n$, the inclusion (48) transmutes itself into the equality

$$\sum_{i=0}^{m} \lambda_i[\nabla g_i(z) - \nabla h_i(z)] = \sum_{i=0}^{m} \lambda_i \nabla f_i(z) = \nabla \mathcal{L}(z, \lambda) = 0 \in I\!R^n. \tag{48$'$}$$

It can be readily seen that those are the KKT-conditions, because the complementarity conditions can be produced from (48) [1,2,4,12–14] as well. Hence, we conclude that the conditions (45)–(46) of Theorem 5.1 are connected, in a natural way, with the classical optimality conditions [1–4,8,11–14].

Remark 5.4 Furthermore, suppose that there exists a tuple (y, β, u), such that (y, β) satisfies the equality (45) and violates (46), i.e., $G_\lambda(u) - \beta < \langle H'_\lambda(y), u - y\rangle$. Then, due to convexity of $H_\lambda(\cdot) = \sum_{i=0}^{m} \lambda_i h_i(\cdot)$, it follows $G_\lambda(u) - \beta < H_\lambda(u) - H_\lambda(y)$. Moreover, on account of (44) and (45), we have

$$\mathcal{L}(u, \lambda) = G_\lambda(u) - H_\lambda(u) < \beta - H_\lambda(y) = \zeta = f_0(z) = \mathcal{L}(z, \lambda),$$

where $\lambda \in \mathcal{M}(z)$. Hence, the right-hand side inequality in (41) is violated with $u \in S$. It means that the pair (z, λ) is not a saddle point: $(z, \lambda) \notin Sdl(\mathcal{L})$. So, the GOC (45),(46) possess the constructive (algorithmic) property, which is demonstrated by the following example. □

Example 5.1 (Example 4.2 Revisited) Let us return to the problem (29), where $z = (1, 0)^\top$ is a KKT-point with $\lambda_1 = 0.3$, so that $\mathcal{L}(x, \lambda) = f_0(x) + 0.3 f_1(x) = x_2^2 - 0.1(x_1 - 4)^2 + 0.3(1 - x_1^2 - x^2)$, $S = [-2, 4] \times I\!R$.

Then, we see that $G_\lambda(x) = x_2^2 + 0.3$, $H_\lambda(x) = 0.1(x_1 - 4)^2 + 0.3(x_1^2 + x_2^2)$. Let us show that $z = (1, 0)^\top$ with $\zeta := f_0(z) = -0.9$ is not a global solution to (29).

First, by setting $y = (0, 0)^\top$, $u = (-0.5; 0)^\top$, we see that $H_\lambda(y) = 1.6 = \beta - \zeta$, $\beta = 0.7$,

$$\begin{aligned} \nabla H_\lambda(y) &= (0.2(y_1 - 4) + 0.6y_1, 0.6y_2)^\top = (-0.8, 0)^\top, \\ \langle \nabla H_\lambda(y), u - y\rangle &= \langle (-0.8, 0)^\top, (-0.5, 0)^\top\rangle = 0.4, \\ \theta(y, \beta) &:= \beta + \langle \nabla H_\lambda(y), u - y\rangle = 1.1, \end{aligned}$$

while $G_\lambda(u) = 0.3$. Hence, we obtain $G_\lambda(u) = 0.3 < 1.1 = \beta + \langle \nabla H(y), u - y\rangle$, so that (z, λ_1) is not a saddle point of $\mathcal{L}(x, \lambda_1)$ in virtue of Theorems 5.1. Besides, $z = (1, 0)^\top$ is not a global solution to (29), since $f_0(u) = -0.1(-\frac{1}{2}-4)^2 = -2.025 < -0.9 = f_0(z)$ that confirms the assertion. □

Furthermore, regardless of the effectiveness of GOC demonstrated in Examples 4.1 and 4.2 (Revisited), the question arises when there exists such a tuple (y, β, u) which violates (46). An answer is given by the following result.

Theorem 5.2 *Let a KKT-point $z \in D$ with the corresponding multipliers $\lambda \in \mathcal{M}(z)$, $\lambda_0 = 1$, be given. Besides, the following assumption takes place*

$$(\mathcal{H}): \qquad \exists v \in I\!R^n : \quad \mathcal{L}(v, \lambda) > \mathcal{L}(z, \lambda) = f_0(z) =: \zeta. \tag{49}$$

Suppose, in addition, that the pair (z, λ) is not a saddle point of $\mathcal{L}(z, \lambda)$ on $S \times I\!R^m_+$.

Then, one can find a tuple (y, β, u), where $(y, \beta) \in I\!R^{n+1}$, $u \in S$, and a fixed collection of subgradients $\{h'_{00}(y), h'_{10}(y), \ldots, h'_{m0}(y)\}$, $h'_{i0}(y) \in \partial h_i(y)$, $i \in \{0\} \cup I$, such that

$$H_\lambda(y) := \sum_{i=0}^{m} \lambda_i h_i(y) = \beta - \zeta, \tag{50}$$

$$G_\lambda(y) \le \beta, \tag{51}$$

$$G_\lambda(u) - \beta < \sum_{i=0}^{m} \lambda_i \langle h'_{i0}(y), u - y\rangle. \tag{52}$$

Proof 1) According to the assumption, there exists a point $u \in S$ such that $\mathcal{L}(u, \lambda) < \mathcal{L}(z, \lambda) = f_0(z) = \zeta$.

In particular, it may be a feasible in Problem $(\mathcal{P})$ point, which is better than z: $f_i(u) \le 0$, $i \in I$, $f_0(u) < f_0(z)$. In fact, since $\lambda_i \ge 0$, $i \in I$, we have $\mathcal{L}(u, \lambda) := f_0(u) + \sum_{i=1}^{m} \lambda_i f_i(u) \le f_0(u) < f_0(z) = \mathcal{L}(z, \lambda)$.

The proof is carried out further in the same fashion as the proof of Theorem 4.2.

Actually, introducing the convex set $C := \mathrm{epi}\big[H_\lambda(\cdot) + \zeta\big] = \{(x, \gamma) \in I\!R^{n+1} : H_\lambda(x) + \zeta \le \gamma\}$, we establish the relations as follows $(u, G_\lambda(u)) \notin C$, $\quad (v, G_\lambda(v)) \in \mathrm{int}\, C$.

Then, there exists a number α, $0 < \alpha < 1$, such that $(y, \beta) = \alpha(u, G_\lambda(u)) + (1 - \alpha)(v, G_\lambda(v)) \in \mathrm{bd}\, C$, or, which is the same, $y = \alpha u + (1 - \alpha)v$, $\beta = \alpha G_\lambda(u) + (1 - \alpha)G_\lambda(v) = H_\lambda(y) + \zeta$.

So, (50) is proved, and (51) can be proved by the convexity of $x \mapsto G_\lambda(x)$.

Furthermore, the negation of the assertion of Theorem 5.2 leads us to the absurdity

$$0 \geq \beta - G_\lambda(u) + \sum_{i=0}^{m} \lambda_i \langle h_i'(y), u - y\rangle = \frac{1-\alpha}{\alpha}\big[G_\lambda(v) - \beta\big] + \frac{1-\alpha}{\alpha}\langle H_\lambda'(y), y - v\rangle$$
$$\geq \frac{1-\alpha}{\alpha}\big[\mathcal{L}(v,\lambda) - \zeta\big] > 0.$$

The latter inequality follows from $(\mathcal{H})$–(49). Hence, Theorem 5.2 is proven. □

Now let us compare the effectiveness of the GOC of Theorems 5.1, 5.2 and of Theorems 4.1, 5.1.

Example 5.2 (Example 4.3 Revisited) Recall that Theorems 4.1 and 4.2 fail to help escape the KKT-point $z = \left(\frac{4}{3}, -\frac{2}{3}\right)^\top$ with $\lambda_1 = 4$, $\lambda_2 = 0$ and improve the value of the function $F(z,\zeta) = 0$, $\zeta := f_0(z) = 5\frac{1}{3}$. Meanwhile, there exist points feasible in the problem (40), which are better than z in the sense of the problem (40).

Now, let us employ the Lagrange function $\mathcal{L}(x,\lambda)$ with $\lambda = (4,0) \in \mathcal{M}(z)$:

$$\mathcal{L}(x,\lambda) = f_0(x) + \lambda_1 f_1(x) = \frac{1}{2}(x_1-4)^2 + (x_2+2)^2 + 4[(x_1-1)^2 - (x_2+1)^2].$$

Then, we have $\mathcal{L}(x,\lambda) = G_\lambda(x) - H_\lambda(x)$, where $G_\lambda(x) = \frac{1}{2}(x_1-4)^2 + (x_2+2)^2 + 4(x_1-1)^2$, $H_\lambda(x) = 4(x_2+1)^2$. Let us choose $y = (0, -\frac{1}{2})^\top$, $u = \left(\frac{4}{3}, 0\right)^\top$, $f_1(u) = -\frac{8}{9} < 0$, $f_2(u) = -7\frac{1}{9} < 0$. Besides, it yields $\nabla H_\lambda(y) = (0, 8(y_2+1))^\top = (0,4)^\top$. Then, we obtain that $\beta = H_\lambda(y) + \zeta = 4(-\frac{1}{2}+1)^2 + 5\frac{1}{3} = 6\frac{1}{3}$, $\langle \nabla H_\lambda(y), u - y\rangle = \big\langle (0,4)^\top, (\frac{4}{3}, 0.5)^\top \big\rangle = 2$. Now it can be readily seen that

$$\psi(y,\beta) := \beta + \langle \nabla H_\lambda(y), u - y\rangle = 8\frac{1}{3}, \quad G_\lambda(u) = \frac{1}{2}\left(\frac{4}{3} - 4\right)^2 + 2^2 + 4\left(\frac{4}{3} - 1\right)^2 = 8.$$

Hence, we see that $G_\lambda(u) = 8 < 8\frac{1}{3} = \beta + \langle \nabla H_\lambda(y), u - y\rangle$. So, (46) is violated. Due to Theorems 5.1 and 5.2, it means that (z,λ) is not a saddle point of the Lagrange function. Actually $\mathcal{L}(u,\lambda) := f_0(u) + 4 f_1(u) = 7\frac{5}{9} + 4\cdot\left(-\frac{8}{9}\right) = 4 < 5\frac{1}{3} = f_0(z) = \mathcal{L}(z,\lambda)$. On the other hand, it is easy to compute that

$$f_0(u) = \frac{1}{2}\left(\frac{4}{3} - 4\right)^2 + 2^2 = \frac{32}{9} + 4 = 7\frac{5}{9} > f_0(z) = 5\frac{1}{3},$$

so that there is no improvement at all in the problem (40). Hence, we see that, in contrast to the GOC of Theorems 4.1 and 4.2, those ones of Theorems 5.1 and 5.2 allow us to improve the point z in the sense of the Lagrange function, since they aim at minimizing the Lagrange function with respect to the variable x.

However, they do not aim at the minimization of $f_0(x)$ over D, i.e., to solve Problem $(\mathcal{P})$. □

Remark 5.5 The volume of the paper allows only a few words about the practical use of GOC of Theorems 4.1–4.3 and 5.1–5.2. First, a local search method, which is "a brick" or "a workhorse" of a general global search, can be produced as follows. Let us give an approximate $x^s \in S$ for problem $(\mathcal{L})$, then next iterate can be chosen as an approximate solution to the linearized (convex) problem

$$(\mathcal{L}L_s): \qquad \min_x \Phi^s_\lambda(x) := G_\lambda(x) - \langle H'_\lambda(x^s), x\rangle, \quad x \in S.$$

The convergence of the method can be investigated in the same manner as in [15,17,18]. Furthermore, we consider Theorems 4.2 and 5.2 as a hint of a procedure enabling to escape a local pitfall with the help of parameters (y, β): $H_\lambda(y) = \beta - \zeta_k$, where $\zeta_k = f_0(z^k)$, z^k is a current iterate, $z^k \in D$, which is not a solution to $(\mathcal{P})$.

For finding (numerically) such a pair (y, β), we fixe some $\beta \in [\beta_-, \beta_+]$ (see [15,18]) and approximate the level surface $U_H(\beta) = \{y \in I\!R^n : H_\lambda(y) = \beta - \zeta_k\}$ by a finite approximation $\mathcal{A}_k(\beta) = \{y^i \in I\!R^n : H_\lambda(y^i) = \beta - \zeta_k,\ i = 1, \dots, N\}$. For instance, in the case $H_\lambda(x) = \|x\|^2$ or $H_\lambda(x) = \frac{1}{2}\langle x, A_\lambda x\rangle$ with a positive definite matrix $A = A^\top$, one can use $y^i = \pm\mu_i e^i$, with e^i from a basis in $I\!R^n$ and μ_i as a solution to the quadratic equations with respect to μ_i, $i = 1, \dots, n$. Furthermore, we solve the linearized problem

$$(\mathcal{P}L_i): \qquad \min_x (G_\lambda(x) - \langle H'_\lambda(y^i), x\rangle), \quad x \in S,$$

providing $u^i \in S$. Finally, we have to verify the inequality $G_\lambda(u^i) - \beta \geq \langle H'_\lambda(y^i), u^i - y^i\rangle$ $\forall i = 1, \dots, n$ and decide by means of constructive property of GOC about the effectiveness of the points $u^i \in S$, $i = 1, \dots, n$. It is a simplest and rough scheme of global search. □

More practical and effective GS methods were developed in [15–17]. The sufficient GOC for $(\mathcal{L})$ are presented below.

Theorem 5.3 *Let us be given a feasible in Problem* $(\mathcal{P})$ *point* $z \in D$, $\zeta := f_0(z)$, *and, besides, the following assumption takes place*

$$(\mathcal{H}(\mathcal{L})): \qquad \exists v \in I\!R^n: \quad \exists \lambda = (\lambda_1, \dots, \lambda_m)^\top \in \mathcal{M}(z), \quad \lambda_0 = 1, \textit{ for which}$$
$$\mathcal{L}(v, \lambda) := G_\lambda(v) - H_\lambda(v) = \sum_{i=0}^m \lambda_i f_i(v) > \zeta = f_0(z) = \mathcal{L}(z, \lambda). \tag{53}$$

Suppose, in addition, that for every pair $(y, \beta) \in I\!R^n \times I\!R$, *such that*

$$(a)\ H_\lambda(y) = \sum_{i=0}^m \lambda_i h_i(y) = \beta - \zeta, \qquad (b)\ G_\lambda(y) \leq \beta, \tag{45$'$}$$

there exists a fixed collection of subgradients $\{h'_{00}(y), h'_{10}(y), \dots h'_{m0}(y)\}$, *where* $h'_{i0}(y) \in \partial h_i(y)$, $i = 0, 1, \dots, m$, *for which the following inequality holds*

$$G_\lambda(x) - \beta \geq \sum_{i=0}^{m} \lambda_i \langle h'_{i0}(y), x - y\rangle \quad \forall x \in S. \tag{46'}$$

Then, the pair $(z, \lambda) \in S \times I\!R^m_+$ *turns out to be a saddle point of the function* $\mathcal{L}(x, \mu)$. □

The proof of Theorem 5.3 is similar to those ones of Theorems 4.2 and 4.3 presented in Sect. 4, and it is therefore omitted. Besides, it can be readily seen that in Theorem 5.3, for every pair (y, β), satisfying GOC (45′), the inequality (46′) must be fulfilled for only one fixed tuple of subgradients $h'_{i0}(y) \in \partial h_i(y)$ (and not for any $h'_i(y) \in \partial h_i(y)$, $i \in \{0\} \cup I$, as it was in Theorem 5.1). On the other hand, with the help of Propositions 5.1 and 5.2, we obtain from Theorem 5.3 the following results.

Corollary 5.1 *Suppose all the assumptions of Theorem 5.3 are fulfilled. Then, the point* $z \in D$ *is a global solution to Problem* $(\mathcal{P})$.

However, let us stress the fact that, in virtue of Theorem 5.1, the conditions (45)–(46) are not regretfully necessary for a point $z \in D$ to be a global solution to $(\mathcal{P})$. On the other hand, Theorem 5.3 provides for the inclusion $(z, \lambda) \in Sdl(\mathcal{L})$, $\lambda \in \mathcal{M}(z)$, which, in turn, yields that $z \in Sol(\mathcal{P})$.

At the same time, note that, if (z, λ) is not a saddle point of $\mathcal{L}(x, \lambda)$, meanwhile z is a KKT-point, then we are not able to say that $z \in Sol(\mathcal{P})$, which is our major goal. Naturally, the question arises whether for every global solution z to Problem $(\mathcal{P})$, there exists a corresponding Lagrange multiplier $\lambda \in \mathcal{M}(z)$, such that $(z, \lambda) \in Sdl(\mathcal{L})$, which can help us to find a solution to Problem $(\mathcal{P})$. We propose to see a supplementary example.

Example 5.3 (of G.R.Walsh, [20], p. 67) Consider the problem

$$\left.\begin{array}{c} \min\limits_x f_0(x) = x_1x_2 - 2x_1^2 - 3x_2^2, \\ f_1(x) = 3x_1 + 4x_2 - 12 \leq 0,\; f_2(x) = x_2^2 - x_1^2 + 1 \leq 0, \\ f_3(x) = -x_1 \leq 0, \quad f_4(x) = x_1 - 4 \leq 0, \quad f_5(x) = -x_2 \leq 0, \quad f_6(x) = x_2 - 3 \leq 0. \end{array}\right\} \tag{54}$$

Let us study the unique solution $z = (4, 0)^\top$ to the problem (54), $f_0(z) = \zeta = -32$. Clearly, the Lagrange function at z takes the form $\mathcal{L}(x, \lambda) = x_1x_2 - 2x_1^2 - 3x_2^2 + \lambda_1(3x_1 + 4x_2 - 12) + \lambda_4(x_1 - 4) - \lambda_5 x_2$, $S = I\!R^2$, because $\lambda_2 = \lambda_3 = \lambda_6 = 0$ due to the complementarity conditions: $\lambda_i f_i(x) = 0$, $i = \overline{1, 6}$. In addition, with the help of the KKT-conditions at $z = (4, 0)^\top$, we derive that $\lambda_1 = 2$, $\lambda_4 = 10$, $\lambda_5 = 12$, so that

$$\mathcal{L}(x, \lambda) = x_1x_2 - 2x_1^2 - 3x_2^2 + 16x_1 - 4x_2 - 64. \tag{55}$$

Besides, $\mathcal{L}(z, \lambda) = -32 = \zeta = f_0(z)$, as it should be. Furthermore, it can be readily seen that the function $x \mapsto \mathcal{L}(x, \lambda)$ is a d.c. one. We use the d.c. representation as follows: $\mathcal{L}(x, \lambda) = G_\lambda(x) - H_\lambda(x)$, where

$$G_\lambda(x) = 2(x_1^2 + x_2^2) + 16x_1 - 4x_2 - 64, \quad H_\lambda(x) = 4x_1^2 + 5x_2^2 - x_1x_2. \tag{56}$$

Therefore, one can see that $\beta = H_\lambda(y)+\zeta = H_\lambda(y)-32$, $\nabla H_\lambda(x) = (8x_1-x_2, 10x_2-x_1)^\top$, $\langle \nabla H(y), u-y\rangle = 8y_1u_1 - y_2u_1 + 2y_1y_2 - 8y_1^2 - 10y_2^2 + 10y_2u_2 - y_1u_2$, from what it follows that

$$\begin{aligned}\theta(y,\beta) &:= \beta + \langle \nabla H(y), u-y\rangle = \psi(y)\\ &:= (8u_1 - u_2)y_1 + (10u_2 - u_1)y_2 + y_1y_2 - 4y_1^2 - 5y_2^2 - 32.\end{aligned} \tag{57}$$

Moreover, Lemma 4.2 leads us to the equality: $y_* = u$. Now, let us choose the vector u as $u = (-\frac{1}{5}, -\frac{4}{5})^\top$, since $S = I\!R^2$. Then, one can compute that $G_\lambda(u) = -62\frac{16}{25} < -42 = \psi(y_*) = \theta(y_*, \beta_*)$, so that (46) is violated. Therefore, due to Theorems 5.1 and 5.2, it implies that the pair (z, λ) is not a saddle point of $\mathcal{L}(u, \lambda)$. The latter assertion can be easily verified by the direct calculations $\mathcal{L}(u, \lambda) = -65\frac{21}{25} < -32 = \mathcal{L}(z, \lambda)$. So, we see that even for unique (global) solution $z \in Sol(\mathcal{P})$, there does not exist a Lagrange multiplier λ such that $(z, \lambda) \in Sdl(\mathcal{L})$. □

6 Conclusions

In this paper, we addressed the general optimization problem with the d.c. goal function and the d.c. inequality constraints. Besides, the original problem has been reduced to optimization problems with two types of the cost function given by the classical Lagrangian and the max-merit function.

The objective was to evaluate these merit functions from the view point of a global solution to the original problem via the global optimality conditions (GOC) and the classical OC, mainly the KKT-theorem.

The results of investigations demonstrated that both merit functions do not adequately reflect the features of the original problem $(\mathcal{P})$, regardless of the effectiveness of the new GOG developed above. Besides, one merit function has some advantages in comparison with another one in a number of examples and vice versa in other examples. Moreover, both merit functions have their own objective different to that one of the original problem. For instance, the GOC and examples showed that it may happen that there is no a saddle point for a global solution to the original problem. Hence, it is necessary to consider other types of the merit or penalty functions [1–4,8–12] more relevant to the problem in question.

Acknowledgements This work has been supported by the Russian Science Foundation, project No. 15-11-20015. The author expresses his particular gratitude to a reviewer whose remarks allowed to improve considerably the presentation of the paper.

References

1. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, New York (2006)
2. Bonnans, J.-F., Gilbert, J.C., Lemaréchal, C., Sagastizábal, C.A.: Numerical Optimization: Theoretical and Practical Aspects, 2nd edn. Springer, Berlin (2006)

3. Izmailov, A.F., Solodov, M.V.: Newton-Type Methods for Optimization and Variational Problems. Springer, New York (2014)
4. Hiriart-Urruty, J.-B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms. Springer, Berlin (1993)
5. Floudas, C.A., Pardalos, P.M. (eds.): Frontiers in Global Optimization. Kluwer Academic Publishers, Dordrecht (2004)
6. Horst, R., Tuy, H.: Global Optimization. Deterministic Approaches. Springer, Berlin (1993)
7. Tuy, H.: D.C. optimization: theory, methods and algorithms. In: Horst, R., Pardalos, P.M. (eds.) Handbook of Global Optimization, pp. 149–216. Kluwer Academic Publisher, Dordrecht (1995)
8. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, New York (1998)
9. Demyanov, V.F., Vasiliev, L.V.: Nondifferentiable Optimization. Translation Series in Mathematics and Engineering. Springer-Optimization Software, New York (1985)
10. Demyanov, V.F., Rubinov, A.M.: Constructive Nonsmooth Analysis. Peter Lang, Frankfurt a/M (1995)
11. Hiriart-Urruty, J.-B.: Generalized differentiability, duality and optimization for problems dealing with difference of convex functions. In: Ponstein, J. (ed.) Convexity and Duality in Optimization. Lecture Notes in Economics and Mathem. Systems, vol. 256, pp. 37–69. Springer, Berlin (1985)
12. Clarke, F.H.: Optimization and Nonsmooth Analysis. Wiley-Interscience, New York (1983)
13. Ioffe, A.D., Tihomirov, V.M.: Theory of Extremal Problems. Elsevier Science Ltd., Amsterdam (1979)
14. Hiriart-Urruty, J.-B.: Optimisation et Analyse Convex. Presses Universitaires de France, Paris (1998)
15. Strekalovsky, A.S.: On solving optimization problems with hidden nonconvex structures. In: Rassias, T.M., Floudas, C.A., Butenko, S. (eds.) Optimization in Science and Engineering, pp. 465–502. Springer, New York (2014)
16. Strekalovsky, A.S.: Global optimality conditions for optimal control problems with functions of A.D. Alexandrov. J. Optim. Theory Appl. **159**, 297–321 (2013)
17. Strekalovsky, A.S.: On local search in d.c. optimization problems. Appl. Math. Comput. **255**, 73–83 (2015)
18. Strekalovsky, A.S.: Elements of Nonconvex Optimization. Nauka, Novosibirsk (2003). **(in Russian)**
19. Strekalovsky, A.S.: On the minimization of the difference of convex functions on a feasible set. Comput. Math. Math. Phys. **43**(3), 399–409 (2003)
20. Walsh, G.R.: Methods of Optimization. Wiley, New York (1975)