Global Optimality Conditions for Optimization Problem with D.C. Inequality and Equality Constraints

Alexander S. Strekalovsky
Matrosov Institute for System Dynamics and Control Theory SB RAS
Lermontov st., 134,
664033 Irkutsk, Russia
strekal@icc.ru

Abstract

This paper addresses the nonconvex optimization problem with the cost function and constraints given by d.c. functions. The original problem is reduced to a problem without inequality and equality constraints by means of the exact penalization techniques. Furthermore, the penalized problem is presented as a d.c. minimization problem. For the latter problem we develop the global optimality conditions (GOCs) which reduce the nonconvex optimization problem to a family of convex problems. In the paper the properties of the GOCs are investigated. The effectiveness of the GOCs is demonstrated by examples.

1 Statement of the Problem

Consider the following problem:

$$(\mathcal{P}): \qquad \left.\begin{array}{c} f_0(x) := g_0(x) - h_0(x) \downarrow \min\limits_x, \quad x \in S, \\ f_i(x) := g_i(x) - h_i(x) \leq 0, \quad i \in I = \{1, \ldots, m\}, \\ f_i(x) := g_i(x) - h_i(x) = 0, \quad i \in \mathcal{E} = \{m+1, \ldots, l\}; \end{array}\right\}$$

where the functions $g_i(\cdot)$, $h_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, are convex on $I\!R^n$, so that the functions $f_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, are the d.c. functions[Floudas et al., 2004, Horst et al., 1993, Tuy, 1995, Hiriart-Urruty et al., 1993, Hiriart-Urruty, 1985]. Recall that any continuous function can be approximated by d.c. function with any desirable accuracy. Let all functions in $(\mathcal{P})$ be smooth.

Besides, assume that the set $S \subset I\!R^n$ is convex and compact.

Furthermore, suppose that the set $Sol(\mathcal{P})$ of global solutions to Problem $(\mathcal{P})$, $Sol(\mathcal{P}) := \{x \in \mathcal{F} \mid f_0(z) = \mathcal{V}(\mathcal{P})\}$ and the feasible set $\mathcal{F}$ of Problem $(\mathcal{P})$, $\mathcal{F} := \{x \in S \mid f_i(x) \leq 0,\ i \in I,\ f_i(x) = 0,\ i \in \mathcal{E}\}$, are non-empty. Besides, in what follows the optimal value $\mathcal{V}(\mathcal{P})$ of Problem $(\mathcal{P})$ is supposed to be finite:

$$\mathcal{V}(\mathcal{P}) := \inf(f_0, \mathcal{F}) := \inf_x \{f_0(x) \mid x \in \mathcal{F})\} > -\infty.$$

In: Yu. G. Evtushenko, M. Yu. Khachay, O. V. Khamisov, Yu. A. Kochetov, V.U. Malkova, M.A. Posypkin (eds.): Proceedings of the OPTIMA-2017 Conference, Petrovac, Montenegro, 02-Oct-2017, published at http://ceur-ws.org

2 Exact Penalty

Introduce the penalty function $W(\cdot)$ for Problem $(\mathcal{P})$ as follows

$$W(x) := \max\{0, f_1(x), \ldots, f_m(x)\} + \sum_{j\in\mathcal{E}} |f_j(x)|\,. \tag{1}$$

Further, along with Problem $(\mathcal{P})$, consider the penalized problem without the inequality and equality constraints:

$$(\mathcal{P}_\sigma): \qquad \theta_\sigma(x) \triangleq f_0(x) + \sigma W(x) \downarrow \min_x, \quad x \in S, \tag{2}$$

where $\sigma \geq 0$ is a penalty parameter.

As well-known, if $z \in Sol(\mathcal{P}_\sigma)$, and z is feasible in $(\mathcal{P})$, i.e. $z \in \mathcal{F}$, then z turns out to be a global solution to $(\mathcal{P})$: $z \in Sol(\mathcal{P})$ [Nocedal et al., 2006, Bonnans et al., 2006, Izmailov et al., 2014, Hiriart-Urruty et al., 1993, Clarke, 1983, Burke, 1991]. On the other hand, the inverse implementation does not, in general, hold.

Hence, the crucial moment of the exact penalization (EP) theory is the existence of a threshold value $\sigma_* \geq 0$ of the penalty parameter $\sigma \geq 0$ for which $Sol(\mathcal{P}_\sigma) \subset Sol(\mathcal{P})$ $\forall \sigma \geq \sigma_*$. In other words, for $\sigma \geq \sigma_*$ Problems $(\mathcal{P})$ and $(\mathcal{P}_\sigma)$ turn out to be equivalent in the sense that $Sol(\mathcal{P}) = Sol(\mathcal{P}_\sigma)$ (see Chapt. VII, Lemma 1.2.1 in [Hiriart-Urruty et al., 1993]).

On the other hand, the existence of the threshold exact penalty parameter $\sigma_* \geq 0$ allows us to solve a single unconstrained problem instead of a sequence of unconstrained problems with $\sigma_k \to \infty$ [Byrd et al., 2012, Di Pillo et al., 2012, Di Pillo et al., 2015].

Recall that under various constraint qualification (CQ) conditions (MFCQ, etc. [Robinson, 1976, Burke, 1991, Zaslavski, 2013, Kruger, 2015, Kruger et al., 2014]), the error bound properties [Nocedal et al., 2006, Bonnans et al., 2006, Izmailov et al., 2014, Robinson, 1976, Burke, 1991, Han et al., 1979, Kruger, 2015, Kruger et al., 2014], the metric sub-regularity conditions, calmness of constraints systems can help to prove the existence of the exact penalty threshold $\sigma_* \geq 0$ even for a global solution [Clarke, 1983, Burke, 1991, Cococcioni et al., 2017, Zaslavski, 2013, Di Pillo et al., 2012, Di Pillo et al., 2015].

Assume that some regularity condition is fulfilled that ensures the existence of such threshold value $\sigma_* \geq 0$ of penalty parameter.

3 Global Optimality Conditions (GOC)

Before all, we will prove that the cost function $\theta_\sigma(\cdot)$ of Problem $(\mathcal{P}_\sigma)$ is a d.c. function, i.e. it can be represented as a difference of convex functions. Indeed, since

$$|f_i(x)| = \max\{g_i(x) - h_i(x), h_i(x) - g_i(x)\} \pm [g_i(x) + h_i(x)] = 2\max\{g_i(x), h_i(x)\} - [g_i(x) + h_i(x)],$$

it can be readily seen that

$$\theta_\sigma(x) \overset{\triangle}{=} f_0(x) + \sigma \max\{0, f_i(x),\ i \in I\} + \sigma \sum_{i\in\mathcal{E}} |f_i(x)| = G_\sigma(x) - H_\sigma(x), \tag{3}$$

where

$$H_\sigma(x) := h_0(x) + \sigma \Big[\sum_{i\in I} h_i(x) + \sum_{j\in\mathcal{E}} (g_j(x) + h_j(x))\Big], \tag{4}$$

$$G_\sigma(x) := \theta_\sigma(x) + H_\sigma(x) = g_0(x) + \sigma \max\Big\{\sum_{j\in I} h_j(x);\ \Big[g_j(x) + \sum_{j\in I}^{j\neq i} h_j(x)\Big],\ i \in I\Big\} + \\ + 2\sigma \sum_{i\in\mathcal{E}} \max\{g_i(x); h_i(x)\}. \tag{5}$$

Obviously, $G_\sigma(\cdot)$ and $H_\sigma(\cdot)$ are both convex functions [Hiriart-Urruty et al., 1993, Rockafellar et al., 1998, Rockafellar, 1970], so that the function $\theta_\sigma(\cdot)$ is a d.c. function, as claimed. Besides, it is clear, that for a feasible (in $(\mathcal{P})$) point $z \in S$ we have

$$W(z) \overset{\triangle}{=} \max\{0, f_1(z), \ldots, f_m(z)\} + \sum_{i\in\mathcal{E}} |f_i(z)| = 0,$$

and therefore, for $\zeta := f_0(z)$, we obtain

$$\theta_\sigma(z) = f_0(z) + \sigma W(z) = f_0(z) = \zeta. \tag{6}$$

Theorem 3.1. *Let a point $z \in \mathcal{F}$ be a solution to Problem $(\mathcal{P})$ and $\sigma \geq \sigma_* > 0$, where $\sigma_* \geq 0$ is a threshold value of penalty parameter.*

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

$$H_\sigma(y) = \beta - \zeta, \tag{7}$$

the following inequality holds

$$G_\sigma(x) - \beta \geq \langle \nabla H_\sigma(y), x - y \rangle \quad \forall x \in S. \tag{8}$$

Remark 3.1. It is not difficult to note that Theorem 3.1 reduces the solution of the nonconvex Problem $(\mathcal{P}_\sigma)$ to an investigation of the family of the convex (linearized) problems

$$(\mathcal{P}_\sigma L(y)): \quad \Phi_{\sigma y}(x) := G_\sigma(x) - \langle \nabla H_\sigma(y), x \rangle \downarrow \min_x, \quad x \in S, \tag{9}$$

depending on the pairs $(y, \beta) \in I\!R^{n+1}$ which fulfill the equation (7) (or, what is the same),

$$(\mathcal{P}_\sigma L(y)): \quad \Phi_{\sigma y}(x) := G_\sigma(x) - \langle \nabla h_0(y) + \sigma \Big[\sum_{i \in I} \nabla h_i(y) + \sum_{j \in \mathcal{E}} (\nabla g_j(y) + \nabla h_j(y)) \Big], x \rangle \downarrow \min_x, \quad x \in S. \tag{9$'$}$$

It is worth noting that the linearization is carried out with respect to the "unified" nonconvexity of Problem $(\mathcal{P})$ accumulated by the function $H_\sigma(\cdot)$ (see $(\mathcal{P})$–(1) and (4)) that includes all the functions $h_i(\cdot)$, $i \in \{0\} \cup I \cup \mathcal{E}$, $g_j(\cdot)$, $j \in \mathcal{E}$, which generate all nonconvexity in Problems $(\mathcal{P})$ and $(\mathcal{P}_\sigma)$ (according to the representations (3)–(5)).

Hence, the verification of the principal inequality (8) can be performed by solving the linearized problems $(\mathcal{P}_\sigma L(y))$ and varying the parameters (y, β) satisfying (7). Besides, we have to verify (8), which can be rewritten as follows

$$\mathcal{V}(\mathcal{P}_\sigma L(y)) \geq \beta - \langle \nabla H_\sigma(y), y \rangle =: N(y, \beta), \tag{8$'$}$$

where $\mathcal{V}(\mathcal{P}_\sigma L(y))$ is the optimal value of the linearized problem $(\mathcal{P}_\sigma L(y))$

Remark 3.2. Suppose, we found a triple (y, β, u), $(y, \beta) \in I\!R^n \times I\!R$, $H_\sigma(y) = \beta - \zeta$, $u \in S$, such that the principal inequality (8) is violated, i.e.

$$0 > G_\sigma(u) - \beta - \langle \nabla H_\sigma(y), u - y \rangle,$$

Then, using the equation (7) and the convexity of the function $H_\sigma(\cdot)$, we derive

$$0 > G_\sigma(u) - \beta - H_\sigma(u) + H_\sigma(y) = \theta_\sigma(u) - \zeta = \theta_\sigma(u) - \theta_\sigma(z),$$

or, $\theta_\sigma(z) > \theta_\sigma(u)$, $z \in \mathcal{F}$, $u \in S$. Hence, the point z can not be a solution to $(\mathcal{P}_\sigma)$.

Moreover, if z and u are feasible in $(\mathcal{P})$, $z, u \in \mathcal{F}$, and since $W(u) = 0$, we obtain $f_0(z) = \theta_\sigma(z) > \theta_\sigma(u) = f_0(u)$. It means that $z \notin Sol(\mathcal{P})$ and $u \in \mathcal{F}$ is a vector better than $z \in \mathcal{F}$.

Hence, the conditions (7)–(8) of Theorem 3.1 possess the classical constructive (algorithmic) property (once the conditions are violated, one can find a feasible vector which is better than the point under investigation).

Let us demonstrate the effectiveness of this property by an example.

Example 3.1. Consider the problem ([Nocedal et al., 2006, Example 12.20])

$$\left. \begin{array}{c} f_0(x) = 4x_1x_2 \downarrow \min\limits_x, \quad x \in I\!R^2, \\ f_1(x) = x_1^2 + x_2^2 - 1 = 0. \end{array} \right\} \tag{10}$$

It is easy to see that the point $z = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)^\top$, $\zeta := f_0(z) = 2$, is feasible: $f_1(z) = 0$ and satisfies the KKT-equation $\nabla f_0(z) + \lambda_1 \nabla f_1(z) = 0 \in I\!R^2$ with $\lambda_1 = -2$. However, it is not clear whether the point z is a global solution. In order to decide on it, let us apply Theorem 3.1.

Since $f_0(x) = 4x_1x_2 = (x_1 + x_2)^2 - (x_1 - x_2)^2$, it can be readily seen that

$$g_0(x) = (x_1 + x_2)^2, \quad h_0(x) = (x_1 - x_2)^2, \qquad g_1(x) = x_2^2 + x_2^2, \quad h_1(x) \equiv 1. \tag{11}$$

Besides, let set $\sigma := 3 > |\lambda_1| = 2$. Then, according to (4) and (5), we have

$$\left.\begin{array}{r} H_\sigma(x) = h_0(x) + \sigma\big[g_1(x) + h_1(x)\big] = (x_1 - x_2)^2 + 3(x_1^2 + x_2^2 + 1), \\ G_\sigma(x) = g_0(x) + 2\sigma \max\{g_1(x); h_1(x)\} = (x_1 + x_2)^2 + 6\max\{x_1^2 + x_2^2; 1\}. \end{array}\right\} \tag{12}$$

Let choose, now, $y = (-1, 0.5)^\top$ which is unfeasible in the problem (10). Then we have

$$H_\sigma(y) = (y_1 - y_2)^2 + 3(y_1^2 + y_2^2 + 1) = 9$$

and, as a consequence, we derive $\beta = H_\sigma(y) + \zeta = 9 + 2 = 11$. Furthermore, let choose a feasible point $u = (-0.6; 0.8)^\top$, $u_1^2 + u_2^2 = 1$, and compute $G_\sigma(u)$ (see (12))

$$G_\sigma(u) = (u_1 + u_2)^2 + 6\max\{u_1^2 + u_2^2; 1\} = (0.2)^2 + 6 = 6.04.$$

Besides, it is not difficult to compute that $u - y = (-0.6; 0.8)^\top - (-1; 0.5)^\top = (0.4; 0.3)^\top$,

$$\nabla H_\sigma(y) = 2(y_1 - y_2; y_2 - y_1)^\top + 6(y_1, y_2)^\top = 2(4y_1 - y_2; 4y_2 - y_1)^\top = (-9; 6)^\top.$$

Whence we immediately derive that

$$\begin{array}{c} \langle \nabla H_\sigma(y), u - y\rangle = \langle (-9; 6)^\top, (0.4; 0.3)^\top\rangle = -3.6 + 1.8 = -1.8, \\ \beta + \langle \nabla H_\sigma(y), u - y\rangle = 11 - 1.8 = 9.2 > 6.04 = G_\sigma(u). \end{array}$$

The latter inequality means that in Problem (10) the principal inequality (8) of Theorem 3.1 is violated.

Hence, the point $z = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)^\top$ is not a global solution to the problem (10) in virtue of Theorem 3.1.

Indeed, it is confirmed by the inequality $f_0(u) = -0.48 < \zeta = f_0(z) = 2$. □

Let consider now possible relations between the conditions (7)–(8) of Theorem 3.1 and the classical optimality conditions, in particular, the KKT theorem for Problem $(\mathcal{P})$. For this purpose, suppose that a feasible (in Problem $(\mathcal{P})$) point z satisfies the conditions (7)–(8) of Theorem 3.1.

First, let set in (7)–(8) $y = z$. Then we immediately derive that $\beta := H_\sigma(z) + \zeta = G_\sigma(z)$.

Therefore, from (8) it follows the validity of the inequality

$$G_\sigma(x) - G_\sigma(z) \geq \langle \nabla H_\sigma(z), x - z\rangle \quad \forall x \in S.$$

It implies that the point z (satisfying (7)–(8)) is a solution to the linearized convex problem as follows

$$(\mathcal{P}_\sigma L(z)) \colon \quad G_\sigma(x) - \langle \nabla H_\sigma(z), x\rangle \downarrow \min_x, \quad x \in S.$$

Since $(\mathcal{P}_\sigma L(z))$ is a convex problem, then the following inclusion is, as well-known, the necessary and sufficient optimality condition for z being a solution to $(\mathcal{P}_\sigma L(z))$:

$$0_n \in \partial G_\sigma(z) - \nabla H_\sigma(z) + N(z \mid S). \tag{13}$$

When $S = I\!R^n$, the inclusion (13) implies

$$\nabla H_\sigma(z) \subset \partial G_\sigma(z), \tag{13'}$$

which is the necessary optimality condition for Problem $(\mathcal{P}_\sigma)$ with $S = I\!R^n$ [Hiriart-Urruty, 1985, Strekalovsky, 2003]. Thus, the conditions (7)–(8) of Theorem 3.1 entail the well-known optimality conditions (13) and (13′) [Nocedal et al., 2006, Bonnans et al., 2006, Izmailov et al., 2014, Floudas et al., 2004, Strekalovsky, 2013, Strekalovsky, 2014, Strekalovsky, 2017, Strekalovskiy, 2003] for Problem $(\mathcal{P}_\sigma)$.

Nevertheless, the natural question arises on whether it is possible to find a triple $(y, \beta, u) \in I\!R^{2n+1}$, satisfying (7) and which violates the inequality (8).

Theorem 3.2. *Assume, that a feasible in Problem* $(\mathcal{P})$ *point* z *is not an* ε*-solution to* $(\mathcal{P})$*, i.e.*

$$\inf(f_0, \mathcal{F}) + \varepsilon = \mathcal{V}(\mathcal{P}) + \varepsilon < \zeta := f_0(z). \tag{14}$$

In addition, let a vector $v \in I\!R^n$ *satisfy the following inequality*

$$(\mathcal{H}): \qquad f_0(v) > \zeta - \varepsilon. \tag{15}$$

Then, for any penalty parameter $\sigma > 0$ *one can find a tuple* (y, β, u)*,* $(y, \beta) \in I\!R^{n+1}$*,* $u \in \mathcal{F}$*, the following conditions take place*

$$\left.\begin{array}{ll} (a) & H_\sigma(y) = \beta - \zeta + \varepsilon; \\ (b) & G_\sigma(y) \le \beta, \\ (c) & G_\sigma(u) - \beta < \langle \nabla H_\sigma(y), u - y \rangle. \end{array}\right\} \tag{16}$$

□

Now let us demonstrate the effectiveness of the GOCs of Theorems 3.1 and 3.2 on another example.

Example 3.2. Consider the problem

$$\left.\begin{array}{c} f_0(x) = x_1^2 - 2x_2^2 + x_3^2 \downarrow \min\limits_x, x \in I\!R^3, \\ f_1(x) = x_3^2 - x_1^2 - x_2^2 = 0, \quad f_2(x) = 4x_1x_3 = 0, \quad -2 \le x_2 \le 1. \end{array}\right\} \tag{17}$$

It can be readily seen that the point $z = (0,0,0)^\top$, $\zeta := f_0(z) = 0$ is a degenerate KKT point in the problem (17), since $f_1(z) = f_2(z) = 0$, $\nabla f_0(z) = \nabla f_1(z) = \nabla f_2(z) = (0,0,0)^\top$. However, it is not clear whether the KKT vector z is a global solution to (17) or not. Therefore, let us apply Theorems 3.1 and 3.2 to clarify the situation.

It is easy to see that in the problem (17) we have $g_0(x) = x_1^2 + x_3^2$, $h_0(x) = 2x_2^2$, $g_1(x) = x_3^2$, $h_1(x) = x_1^2 + x_2^2$. In addition, using the d.c. representation $f_2(x) = 4x_1x_3 = (x_1 + x_3)^2 - (x_1 - x_3)^2$, we obtain $g_2(x) = (x_1 + x_3)^2$, $h_2(x) = (x_1 - x_3)^2$.

For simplicity of presentation, we will apply the denotation $S = [-2, 1]$ for bounding the variable $x_2 \in I\!R$, but in the investigation of the linearized problems we use two inequality constraints $x_2 \le 1$, $x_2 + 2 \ge 0$.

Hence, according to (3)–(5) we have

$$\begin{array}{c} H_\sigma(x) = h_0(x) + \sigma \sum\limits_{j \in \mathcal{E}} [g_j(x) + h_j(x)] = \\ = 2x_2^2 + \sigma[(x_3^2 + x_1^2 + x_2^2) + (x_1 + x_3)^2 + (x_1 - x_3)^2] = 2x_2^2 + \sigma[3x_1^2 + 3x_3^2 + x_2^2]; \end{array} \tag{18}$$

$$\begin{array}{c} G_\sigma(x) = g_0(x) + 2\sigma \sum\limits_{j \in \mathcal{E}} \max\{g_j(x); h_j(x)\} = \\ = x_1^2 + x_3^2 + 2\sigma[\max\{x_3^2; x_1^2 + x_2^2\} + \max\{(x_1 + x_3)^2; (x_1 - x_3)^2\}]. \end{array} \tag{19}$$

Let us set $\sigma := 1$, $y = (\frac{1}{6}, 1, \frac{7}{6})^\top \notin \mathcal{F}$. Then we obtain

$$\nabla H_\sigma(x) = (0, 4x_2, 0)^\top + \sigma(6x_1, 2x_2, 6x_3)^\top = 6(x_1, x_2, x_3)^\top,$$

besides, $\nabla H_\sigma(y) = (1, 6, 7)^\top$.

In order to find a suitable point in $u \in \mathcal{F}$, consider the linearized problem as follows

$$\begin{array}{rl} (\mathcal{P}_\sigma L(y)): \quad & G_\sigma(x) - \langle \nabla H_\sigma(y), x \rangle = x_1^2 + x_3^2 + 2\max\{x_3^2; x_1^2 + x_2^2\} + \\ & + 2\max\{(x_1 + x_3)^2; (x_1 - x_3)^2\} - \langle (1,6,7)^\top, x \rangle \downarrow \min\limits_x, \quad x \in I\!R^3, \quad -2 \le x_2 \le 1. \end{array} \tag{20}$$

It is not difficult to see that the problem (20) amounts to the following one [Hiriart-Urruty, 1998]

$$\left.\begin{array}{c} x_1^2 + x_3^2 + 2\gamma_1 + 2\gamma_2 - x_1 - 6x_2 - 7x_3 \downarrow \min\limits_{(x,\gamma)}, \\ x_3^2 \le \gamma_1, \quad x_1^2 + x_2^2 \le \gamma_1, \quad \gamma = (\gamma_1, \gamma_2) \in I\!R^2, \quad (x_1 + x_3)^2 \le \gamma_2, \quad (x_1 - x_3)^2 \le \gamma_2, \\ x_2 \le 1, \quad x_2 + 2 \ge 0, \quad x \in I\!R^3. \end{array}\right\} \tag{20'}$$

Besides, as above, it can be readily seen that the Slater condition holds in (20′). Furthermore, the solution vector $(u, \gamma_*) \in I\!R^5$ satisfies the complementarity conditions as follows

$$\left.\begin{array}{c}\eta_1(x_3^2-\gamma_1)=0=\eta_2(x_1^2+x_2^2-\gamma_1), \qquad \eta_3\big[(x_1+x_3)^2-\gamma_2\big]=0=\eta_4\big[(x_1-x_3)^2-\gamma_2\big],\\ \mu_1(x_2-1)=0=\mu_2(x_2+2).\end{array}\right\} \tag{21}$$

and, besides, we have for $\gamma_* = (\gamma_{*1}, \gamma_{*2})^\top$

$$\gamma_{*1} = \max\{\, u_3^2; u_1^2 + u_2^2 \,\}; \qquad \gamma_{*2} = \max\{\, (u_1+u_3)^2; (u_1-u_3)^2 \,\}. \tag{22}$$

In addition, since the Lagrange function for the problem (20′) has thee following form

$$\begin{aligned}&\mathcal{L}(x,\gamma;\eta_1,\eta_2,\eta_3,\eta_4,\mu_1,\mu_2) = x_1^2 + x_3^2 + 2\gamma_1 + 2\gamma_2 - x_1 - 6x_2 - 7x_3 +\\ &+\eta_1(x_3^2-\gamma_1)+\eta_2(x_1^2+x_2^2-\gamma_1)+\eta_3\big[(x_1+x_3)^2-\gamma_2\big]+\eta_4\big[(x_1-x_3)^2-\gamma_2\big]+\mu_1(x_2-1)-\mu_2(x_2+2),\end{aligned} \tag{23}$$

(an, besides, $(\eta_1,\eta_2,\eta_3,\eta_4,\mu_1,\mu_2) \in I\!R^6_+$), then the KKT system contains the following equations [Rockafellar, 1993]

$$\left.\begin{array}{l}\dfrac{\partial\mathcal{L}(u,\gamma_*)}{\partial\gamma_1} = 2-\eta_1-\eta_2 = 0, \quad \text{i.e. } \eta_1+\eta_2 = 2,\\ \dfrac{\partial\mathcal{L}(u,\gamma_*)}{\partial\gamma_2} = 2-\eta_3-\eta_4 = 0, \quad \text{i.e. } \eta_3+\eta_4 = 2;\end{array}\right\} \tag{24}$$

$$\left.\begin{array}{ll}(a) & \dfrac{\partial\mathcal{L}(u,\gamma_*)}{\partial x_1} = 2u_1 - 1 + 2\eta_2 u_1 + 2\eta_3(u_1+u_3) + 2\eta_4(u_1-u_3) = 0,\\ (b) & \dfrac{\partial\mathcal{L}(u,\gamma_*)}{\partial x_2} = -6 + 2\eta_2 u_2 + \mu_1 - \mu_2 = 0,\\ (c) & \dfrac{\partial\mathcal{L}(u,\gamma_*)}{\partial x_3} = 2u_3 - 7 + 2\eta_1 u_3 + 2\eta_3(u_1+u_3) + 2\eta_4(u_3-u_1) = 0.\end{array}\right\} \tag{25}$$

It can be readily seen that the point $u = (0,1,1)^\top$ satisfies the KKT conditions (21),(24),(25) with $\gamma_* = (\gamma_{*1},\gamma_{*2})^\top = (1,1)^\top$ (see (22)). Indeed, the equation (25) with $u = (0,1,1)^\top$ take the form

$$\left.\begin{array}{ll}(a) & 2\eta_3 - 1 - 2\eta_4 = 0, \qquad \text{or } \eta_3 - \eta_4 = \frac{1}{2},\\ (b) & 2\eta_2 - 6 + \mu_1 - \mu_2 = 0,\\ (c) & 2 - 7 + 2\eta_1 + 2\eta_3 + 2\eta_4 = 0.\end{array}\right\} \tag{25′}$$

Then, from (25′) (a) we derive $\eta_3 = \frac{5}{4}$, $\eta_4 = \frac{3}{4}$. Further, from (25′) (c) with the help of (24) it follows that $2\eta_1 = 5 - 2(\eta_3+\eta_4) = 1$, i.e. $\eta_1 = \frac{1}{2}$, $\eta_2 = \frac{3}{2}$.

On the other hand, thanks to (21) we see that $\mu_2 = \mu_2(u) = 0$.Then (25′) (b) provides that $\mu_1 = 3$. Hence, the point $u = (0,0,1)^T$ really is a KKT point in (20′), and, due to convexity of problem (20′), u is also a solution to (20) ($(u,\gamma_*)^T$ is a solution to (20′)).

Now let us verify whether the principal inequality (8) of Theorem 3.1 holds with (y,β,u) where $\beta = H_\sigma(y)+\zeta$, $\zeta = f_0(z) = 0$. First compute $H_\sigma(y)$ with $y = \left(\dfrac{1}{6}, 1, \dfrac{7}{6}\right)$: $H_\sigma(y) = 3(y_1^2+y_2^2+y_3^2) = 7\dfrac{1}{6}$.

Thus, $\beta = H_\sigma(y) = 7\dfrac{1}{6}$. Furthermore, $\beta + \langle\nabla H(y),(u-y)\rangle = 5\frac{5}{6}$. On the other hand, it can be readily computed, that $G_\sigma(u) = u_1^2 + u_3^2 + 2\gamma_{*1} + 2\gamma_{*2} = 5$.

Therefore, we have $G_\sigma(u) = 5 < 5\dfrac{5}{6} = \beta + \langle\nabla H(y), u-y\rangle$.

Hence, the principal inequality (8) of Theorem 3.1 is violated, and, as a consequence, the degenerate KKT point $z = (0,0,0)^T$ is not a global solution to the problem (17).

Moreover, by solving the linearized problem $(\mathcal{P}L(y))$ with $y = \left(\dfrac{1}{6}, 1, \dfrac{7}{6}\right)^T$, we constructed the feasible in (17) point $u = (0,1,1)^T$, which is better than z, since $f_0(u) = -1 < \zeta_0 = f_0(z) = 0$.

Furthermore, it can be readily seen, as above, that the point $u = (0,1,1)^T$ is also a KKT point in the original problem (17), but not a global solution to (17). Moreover we can show this fact, by repeating the same procedure of finding another pair (y_1, β_1), such that $H_\sigma(y_1) = \beta_1 - \zeta_1$, where $\zeta_1 := f_0(u) = f_0(z_1)$, $z_1 := u$, and by solving the linearized problem $(\mathcal{P}L_1) := (\mathcal{P}L(y_1))$, which provides the point u_1 such that

$$f_0(u_1) =: \zeta_2 < \zeta_1 = f_0(u) = f_0(z_1).$$

So, by the procedure described above we give a hint how may be constructed one of the simplest global search procedures which is able to escape stationary points and local solutions in non-convex Problem $(\mathcal{P})$.

4 Sufficient Optimality Conditions

Now we turn to the question on when the conditions (7)–(8) of Theorem 3.1 become sufficient for a feasible point being a global solution to nonconvex Problem $(\mathcal{P})$.

Theorem 4.1. *Suppose that for a feasible in Problem $(\mathcal{P})$ point z, $\zeta := f_0(z)$, the condition $(\mathcal{H})$–(15) is fulfilled. In addition, let some penalty parameter $\sigma > 0$ be given. Finally, assume that for every pair $(y, \beta) \in I\!R^n \times I\!R$, satisfying the relation*

$$(a) \quad H_\sigma(y) = \beta - \zeta + \varepsilon, \qquad (b) \quad G_\sigma(y) \leq \beta, \tag{26}$$

the following inequality holds

$$G_\sigma(x) - \beta \geq \langle \nabla H_\sigma(y), x - y \rangle \quad \forall x \in S. \tag{27}$$

Then, the point $z \in \mathcal{F}$ turns out to be an ε-global solution to Problem $(\mathcal{P}_\sigma)$ as well as to Problem $(\mathcal{P})$.

Acknowledgements

This work was supported by the Russian Science Foundation (Project No. 15-11-20015).

References

[Bonnans et al., 2006] Bonnans, J.-F., Gilbert, J.C., Lemaréchal, C., & Sagastizábal, C.A. (2006) *Numerical Optimization: Theoretical and Practical Aspects, 2nd edn.* Berlin, Heidelberg: Springer-Verlag.

[Burke, 1991] Burke, J. (1991) An exact penalization viewpoint of constrained optimization. *SIAM Journal on Control and Optimization, 29*, 968-998.

[Byrd et al., 2012] Byrd, R., Lopez-Calva, G., & Nocedal, J. (2012) A line search exact penalty method using steering rules. *Mathematical Programming, Ser. A, 133*, 39-73. doi: 10.1007/s10107-010-0408-0

[Clarke, 1983] Clarke, F.H. (1983) *Optimization and Nonsmooth Analysis.* New York: Springer-Verlag

[Cococcioni et al., 2017] Cococcioni, M., Pappalardo, M., & Sergeyev, Ya.D. (2017) Lexicographic multi-objective linear programming using grossone methodology: Theory and algorithm. *Applied Mathematics and Computation.* doi:10.1016/j.amc.2011.07.042

[Di Pillo et al., 2012] Di Pillo, G., Lucidi, S., & Rinaldi, F. (2012) An approach to constrained global optimization based on exact penalty functions. *Journal of Global Optimization, 54*, 251-260. doi: 10.1007/s10898-010-9582-0

[Di Pillo et al., 2015] Di Pillo, G., Lucidi, S., & Rinaldi, F. (2015) A derivative-free algorithm for constrained global optimization based on exact penalty functions. *Journal of Optimization Theory and Applications, 164*, 862-882. doi: 10.1007/s10957-013-0487-1

[Floudas et al., 2004] Floudas, C.A. & Pardalos, P.M. (eds.) (2004) *Frontiers in Global Optimization.* Dordrecht: Kluwer Academic Publishers.

[Han et al., 1979] Han, S. & Mangasarian, O. (1979) Exact penalty functions in nonlinear programming. *Mathematical Programming, 17*, 251–269.

[Hiriart-Urruty, 1985] Hiriart-Urruty, J.-B. (1985) 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, 256.* (pp. 37–69). Berlin: Springer-Verlag.

[Hiriart-Urruty, 1998] Hiriart-Urruty, J.-B. (1998) *Optimisation et Analyse Convex.* Paris: Presses Universitaires de France (in French).

[Hiriart-Urruty et al., 1993] Hiriart-Urruty, J.-B. & Lemaréchal, C. (1993) *Convex Analysis and Minimization Algorithms.* Berlin: Springer-Verlag.

[Horst et al., 1993] Horst R. & Tuy H. (1993) *Global Optimization. Deterministic Approaches.* Berlin: Springer-Verlag.

[Izmailov et al., 2014] Izmailov, A.F. & Solodov, M.V. (2014) *Newton-Type Methods for Optimization and Variational Problems.* New York: Springer.

[Kruger et al., 2014] Kruger, A., Minchenko, L., & Outrata, J. (2014) On relaxing the Mangasarian-Fromovitz constraint qualiffcation. *Positivity, 18*, 171-189. doi: 10.1007/s11117-013-0238-4

[Kruger, 2015] Kruger, A. (2015) Error bounds and metric subregularity. *Optimization, 64*, 49-79. doi: 10.1080/02331934.2014.938074

[Nocedal et al., 2006] Nocedal, J. & Wright, S.J. (2006) *Numerical Optimization.* New York: Springer.

[Robinson, 1976] Robinson, S. (1976) Stability theory for systems of inequalities, part II: differentiable nonlinear systems. *SIAM Journal on Numerical Analysis, 13*, 497–513.

[Rockafellar, 1970] Rockafellar, R.T. (1970) *Convex Analysis.* Princeton: Princeton University Press.

[Rockafellar, 1993] Rockafellar, R.T. (1993) Lagrange multipliers and optimality. *SIAM Review, 35(2)*, 183-238

[Rockafellar et al., 1998] Rockafellar, R.T. & Wets, R.J.-.B. (1998) *Variational Analysis.* New York: Springer.

[Strekalovsky, 2003] Strekalovsky, A.S. (2013) *Elements of Nonconvex Optimization.* Novosibirsk: Nauka (in Russian)

[Strekalovskiy, 2003] Strekalovskiy, A.S. (2003) On the minimization of the difference of convex functions on a feasible set. *Computational Mathematics and Mathematical Physics, 43*, 399-409.

[Strekalovsky, 2013] Strekalovsky, A.S. (2013) Global optimality conditions for optimal control problems with functions of A.D. Alexandrov. *Journal of Optimization Theory and Applications, 159*, 297-321. doi: 10.1007/s10957-013-0355-z

[Strekalovsky, 2014] Strekalovsky, A.S. (2014) 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). New York: Springer.

[Strekalovsky, 2015] Strekalovsky, A.S. (2015) On local search in d.c. optimization problems. *Applied Mathematics and Computation, 255*, 73-83. doi: http://dx.doi.org/10.1016/j.amc.2014.08.092

[Strekalovsky, 2017] Strekalovsky, A.S. (2017) Global Optimality Conditions in Nonconvex Optimization. *Journal of Optimization Theory and Applications, 173(3)*, 770-792. doi:10.1007/s10957-016-0998-7

[Tuy, 1995] Tuy, H. (1995) D.c. Optimization: Theory, Methods and Algorithms. In Horst, R., Pardalos, P.M. (Eds.) *Handbook of Global optimization.* (pp. 149-216). Dordrecht: Kluwer Academic Publisher.

[Zaslavski, 2013] Zaslavski, A.J (2013) Exact Penalty Property in Optimization with Mixed Constraints via Variational Analysis. *SIAM Journal on Optimization, 23(1)*, 170-187. doi: 10.1137/120870840