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:

\[
(P): \begin{aligned}
 f_0(x) &:= g_0(x) - h_0(x) \downarrow \min_{x} \quad x \in S, \\
 f_i(x) &:= g_i(x) - h_i(x) \leq 0, \quad i \in I = \{1, \ldots, m\}, \\
 f_i(x) &:= g_i(x) - h_i(x) = 0, \quad i \in E = \{m+1, \ldots, l\}; 
\end{aligned}
\]

where the functions \(g_i(\cdot), h_i(\cdot), i \in \{0\} \cup I \cup E,\) are convex on \(\mathbb{R}^n,\) so that the functions \(f_i(\cdot), i \in \{0\} \cup I \cup E,\) 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 \((P)\) be smooth.

Besides, assume that the set \(S \subset \mathbb{R}^n\) is convex and compact.

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

\[V(P) := \inf(f_0, F) := \inf_{x} \{f_0(x) \mid x \in F\} > -\infty.\]

Copyright © by the paper’s authors. Copying permitted for private and academic purposes.

2 Exact Penalty

Introduce the penalty function \( W(\cdot) \) for Problem (\( \mathcal{P} \)) as follows
\[
W(x) := \max\{0, f_1(x), \ldots, f_m(x)\} + \sum_{j \in \mathcal{E}} |f_j(x)|.
\] (1)

Further, along with Problem (\( \mathcal{P} \)), consider the penalized problem without the inequality and equality constraints:
\[
(\mathcal{P}_\sigma): \quad \theta_\sigma(x) = f_0(x) + \sigma W(x) \downarrow \min_x \quad x \in S,
\] (2)

where \( \sigma \geq 0 \) is a penalty parameter.

As well-known, if \( z \in \text{Sol}(\mathcal{P}_\sigma) \), and \( z \) is feasible in (\( \mathcal{P} \)), i.e. \( z \in \mathcal{F} \), then \( z \) turns out to be a global solution to (\( \mathcal{P} \)): \( z \in \text{Sol}(\mathcal{P}) \) [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 \( \text{Sol}(\mathcal{P}_\sigma) \subset \text{Sol}(\mathcal{P}) \) \( \forall \sigma \geq \sigma_* \). In other words, for \( \sigma \geq \sigma_* \), Problems (\( \mathcal{P} \)) and (\( \mathcal{P}_\sigma \)) turn out to be equivalent in the sense that \( \text{Sol}(\mathcal{P}) = \text{Sol}(\mathcal{P}_\sigma) \) (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
\[
|f_i(x)| = \max\{g_i(x) - h_i(x), h_i(x) - g_i(x)\} = g_i(x) + h_i(x) - 2 \max\{g_i(x), h_i(x)\} - [g_i(x) + h_i(x)],
\]
it can be readily seen that
\[
\theta_\sigma(x) \overset{\Delta}{=} f_0(x) + \sigma \max\{0, f_i(x), i \in I\} + \sigma \sum_{i \in \mathcal{E}} |f_i(x)| = G_\sigma(x) - H_\sigma(x),
\] (3)

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

\[
G_\sigma(x) := \theta_\sigma(x) + H_\sigma(x) = g_0(x) + \sigma \max\left\{\sum_{j \in \mathcal{E}} h_j(x); \left\{g_j(x) + \sum_{j \notin i} h_j(x)\right\}, i \in I\right\} + 2\sigma \sum_{i \in \mathcal{E}} \max\{g_i(x); h_i(x)\}.
\] (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
\[
W(z) \overset{\Delta}{=} \max\{0, f_1(z), \ldots, f_m(z)\} + \sum_{i \in \mathcal{E}} |f_i(z)| = 0,
\]

540
Theorem 3.1. Let a point \( z \in F \) be a solution to Problem (\( P \)) and \( \sigma \geq \sigma_0 > 0 \), where \( \sigma_0 \geq 0 \) is a threshold value of penalty parameter.

Then, for every pair \((y, \beta) \in \mathbb{R}^n \times \mathbb{R}\) such that

\[
H_\sigma(y) = \beta - \zeta,
\]

the following inequality holds

\[
G_\sigma(x) - \beta \geq \langle \nabla H_\sigma(y), x - y \rangle \quad \forall x \in S.
\]

Remark 3.1. It is not difficult to note that Theorem 3.1 reduces the solution of the nonconvex Problem (\( P_\sigma \)) to an investigation of the family of the convex (linearized) problems

\[
(P_\sigma L(y)):
\]

\[
\Phi_{\sigma y}(x) := G_\sigma(x) - \langle \nabla H_\sigma(y), x \rangle \downarrow \min_x, \quad x \in S,
\]

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

\[
(P_\sigma L(y)):
\]

\[
\Phi_{\sigma y}(x) := G_\sigma(x) - \langle \nabla h_0(y) + \sigma \left[ \sum_{i \in I} \nabla h_i(y) + \sum_{j \in \mathcal{E}} \nabla g_j(y) \right], x \rangle \downarrow \min_x, \quad x \in S.
\]

It is worth noting that the linearization is carried out with respect to the “unified” nonconvexity of Problem (\( P \)) accumulated by the function \( H_\sigma(\cdot) \) (see (\( P \))–(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 (\( P \)) and (\( P_\sigma \)) (according to the representations (3)–(5)).

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

\[
V(P_\sigma L(y)) \geq \beta - \langle \nabla H_\sigma(y), y \rangle =: N(y, \beta),
\]

where \( V(P_\sigma L(y)) \) is the optimal value of the linearized problem (\( P_\sigma L(y) \)).

Remark 3.2. Suppose, we found a triple \((y, \beta, u) \in \mathbb{R}^n \times \mathbb{R}, \) \( H_\sigma(y) = \beta - \zeta, \) \( u \in S, \) 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, \( \theta_\sigma(z) > \theta_\sigma(u), \) \( z \in F, \ u \in S. \) Hence, the point \( z \) can not be a solution to (\( P_\sigma \)).

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

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])

\[
\begin{align*}
 f_0(x) &= 4x_1x_2 \downarrow \min_x, \quad x \in \mathbb{R}^2, \\
 f_1(x) &= x_1^2 + x_2^2 - 1 = 0.
\end{align*}
\]

It is easy to see that the point \( z = \left( \sqrt{2}, \sqrt{2} \right)^T, \) \( \zeta := f_0(z) = 2, \) is feasible: \( f_1(z) = 0 \) and satisfies the KKT-equation \( \nabla f_0(z) + \lambda_1 \nabla f_1(z) = 0 \in \mathbb{R}^2 \) 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 \( g_0(x) = (x_1 + x_2)^2, \quad h_0(x) = (x_1 - x_2)^2, \quad g_1(x) = x_1^2 + x_2^2, \quad h_1(x) = 1. \) (11)

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

\[
H_\sigma(x) = h_0(x) + \sigma [g_1(x) + h_1(x)] = (x_1 - x_2)^2 + 3(x_1^2 + x_2^2 + 1),
\]

\[
G_\sigma(x) = g_0(x) + 2\sigma \max\{g_1(x); h_1(x)\} = (x_1 + x_2)^2 + 6\max\{x_1^2 + x_2^2; 1\}. \]

(12)

Let choose, now, \( y = (-1, 0.5)^T \) 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 \( u = (-0.6; 0.8)^T, u_1^2 + u_2^2 = 1 \), 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 \( u - y = (-0.6; 0.8)^T - (-1; 0.5)^T = (0.4; 0.3)^T \),

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

Whence we immediately derive that

\[
\langle \nabla H_\sigma(y), u - y \rangle = ((-9; 6)^T, (0.4; 0.3)^T) = -3.6 + 1.8 = -1.8,
\]

\[
\beta + \langle \nabla H_\sigma(y), u - y \rangle = 11 - 1.8 = 9.2 > 6.04 = G_\sigma(u).
\]

The latter inequality means that in Problem (10) the principal inequality (8) of Theorem 3.1 is violated. 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 (P). For this purpose, suppose that a feasible (in Problem (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

\[
(P_\sigma L(z)): \quad G_\sigma(x) - \langle \nabla H_\sigma(z), x \rangle \downarrow \min_x, \quad x \in S.
\]

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

\[
0_n \in \partial G_\sigma(z) - \nabla H_\sigma(z) + N(z \mid S).
\]

(13)

When \( S = \mathbb{R}^n \), the inclusion (13) implies

\[
\nabla H_\sigma(z) \subset \partial G_\sigma(z), \quad (13')
\]

which is the necessary optimality condition for Problem \((P_\sigma)\) with \( S = \mathbb{R}^n \) [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 \((P_\sigma)\).

Nevertheless, the natural question arises on whether it is possible to find a triple \((y, \beta, u) \in \mathbb{R}^{2n+1},\) satisfying (7) and which violates the inequality (8).
**Theorem 3.2.** Assume, that a feasible in Problem (P) point $z$ is not an $\varepsilon$-solution to (P), i.e.

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

In addition, let a vector $v \in \mathbb{R}^n$ satisfy the following inequality

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

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

$$\begin{align*}
(a) \quad & H_\sigma(y) = \beta - \zeta + \varepsilon;
(b) \quad & G_\sigma(y) \leq \beta,
(c) \quad & G_\sigma(u) - \beta < \langle \nabla H_\sigma(y), u - y \rangle.
\end{align*}$$ \hfill (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

$$\begin{align*}
& f_0(x) = x_1^2 - 2x_2^2 + x_3^2 \downarrow \min, x \in \mathbb{R}^3, \\
& f_1(x) = x_2^3 - x_1^2 - x_2^2 = 0, \quad f_2(x) = 4x_1x_3 = 0, \quad -2 \leq x_2 \leq 1.
\end{align*}$$ \hfill (17)

It can be readily seen that the point $z = (0, 0, 0)^T, \ \zeta := f_0(z) = 0$ is a degenerate KKT point in the problem (17), since $f_1(z) = f_2(z) = 0, \ \nabla f_0(z) = \nabla f_1(z) = \nabla f_2(z) = (0, 0, 0)^T$. 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 $g_0(x) = x_1^2 + x_3^2, \ h_0(x) = 2x_2^2, \ g_1(x) = x_2^2, \ h_1(x) = x_1^2 + x_2^2$. 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 $x_2 \in \mathbb{R}$, but in the investigation of the linearized problems we use two inequality constraints $x_2 \leq 1, \ x_2 + 2 \geq 0$.

Hence, according to (3)-(5) we have

$$H_\sigma(x) = h_0(x) + \sigma \sum_{j \in E} [g_j(x) + h_j(x)] =
= 2x_2^2 + \sigma[(x_2^3 + x_1^2 + x_2^2) + (x_1 + x_3)^2 + (x_1 - x_3)^2] = 2x_2^2 + \sigma[3x_1^2 + 3x_3^2 + x_2^2];$$

$$G_\sigma(x) = g_0(x) + 2\sum_{j \in E} \max\{g_j(x); h_j(x)\} =
= x_2^3 + 2\sigma\max\{x_2^3, x_2^2 + x_2^2\} + \max\{(x_1 + x_3)^2; (x_1 - x_3)^2\}].$$ \hfill (19)

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

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

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

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

$$(\mathcal{P}_\sigma L(y)): \quad G_\sigma(x) - \langle \nabla H_\sigma(y), x \rangle = x_2^2 + x_3^2 + 2\max\{x_2^3; x_2^3 + x_2^2\} +
+ 2\max\{(x_1 + x_3)^2; (x_1 - x_3)^2\} - (1, 6, 7)^T, \ x \downarrow \min, \ x \in \mathbb{R}^3, \ -2 \leq x_2 \leq 1.$$ \hfill (20)

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

$$\begin{align*}
x_1^2 + x_3^2 + 2\gamma_1 + 2\gamma_2 + x_1 - 6x_2 - 7x_3 \downarrow \min, \ x \in \mathbb{R}^3, \ x_2 \leq 1, \ x_2 + 2 \geq 0, \end{align*}$$ \hfill (20')
Besides, as above, it can be readily seen that the Slater condition holds in (20'). Furthermore, the solution vector \( (u_1, u_2) \in \mathbb{R}_+^3 \) satisfies the complementarity conditions as follows

\[
\eta_1(x_1^2 - \gamma_1) = 0 = \eta_2(x_1^2 + x_2^2 - \gamma_1), \quad \eta_3[(x_1 + x_3)^2 - \gamma_2] = 0 = \eta_4[(x_1 - x_3)^2 - \gamma_2],
\]

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

\[
\gamma_{\ast 1} = \max\{ u_2^2; u_1^2 + u_3^2 \}; \quad \gamma_{\ast 2} = \max\{ (u_1 + u_3)^2; (u_1 - u_3)^2 \}.
\]

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

\[
L(x, \gamma; \eta_1, \eta_2, \eta_3, \eta_4, \mu_1, \mu_2) = x_1^2 + x_2^3 + 2\gamma_1 + 2\gamma_2 - x_1 - 6x_2 - 7x_3 + \\
+ \eta_1(x_3^2 - \gamma_1) + \eta_2(x_1^2 + x_2^2 - \gamma_1) + \eta_3[(x_1 + x_3)^2 - \gamma_2] + \eta_4[(x_1 - x_3)^2 - \gamma_2] + \mu_1(x_2 - 1) - \mu_2(x_2 + 2),
\]

(23)

Thus, it can be readily seen that the point \( (y_0, \beta_0, u_0) \) is a KKT point in (20), and, due to convexity of problem (20'), it is also a solution to (20) \( (u_0, \gamma_{\ast 0})^\top \) is a solution to (20').

Now let us verify whether the principal inequality (8) of Theorem 3.1 holds with \( (y, \beta, u) \) where

\[
\beta = H_\sigma(y) + \zeta, \quad \zeta = f_0(z).
\]

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

\[
2\eta_1 = 5 - 2(\eta_3 + \eta_4) = 1, \quad i.e. \quad \eta_1 = \frac{1}{2}, \eta_2 = \frac{3}{2}.
\]

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 \( u = (0,0,1,1)^\top \) really is a KKT point in (20), and, due to convexity of problem (20'), \( u \) is also a solution to (20) \( (u_0, \gamma_{\ast 0})^\top \) is a solution to (20').

Now let us verify whether the principal inequality (8) of Theorem 3.1 holds with \( (y, \beta, u) \) where

\[
\beta = H_\sigma(y) + \zeta, \quad \zeta = f_0(z).
\]

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

\[
2\eta_1 = 5 - 2(\eta_3 + \eta_4) = 1, \quad i.e. \quad \eta_1 = \frac{1}{2}, \eta_2 = \frac{3}{2}.
\]

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 \( u = (0,0,1,1)^\top \) really is a KKT point in (20), and, due to convexity of problem (20'), \( u \) is also a solution to (20) \( (u_0, \gamma_{\ast 0})^\top \) is a solution to (20').

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

\[
2\eta_1 = 5 - 2(\eta_3 + \eta_4) = 1, \quad i.e. \quad \eta_1 = \frac{1}{2}, \eta_2 = \frac{3}{2}.
\]

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

\[
2\eta_1 = 5 - 2(\eta_3 + \eta_4) = 1, \quad i.e. \quad \eta_1 = \frac{1}{2}, \eta_2 = \frac{3}{2}.
\]

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 \( u = (0,0,1,1)^\top \) really is a KKT point in (20), and, due to convexity of problem (20'), \( u \) is also a solution to (20) \( (u_0, \gamma_{\ast 0})^\top \) is a solution to (20').
of finding another pair \((y_1, \beta_1)\), 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 \((PL_1) := (PL(y_1))\), 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 \((P)\).

\section{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 \((P)\).

\textbf{Theorem 4.1.} Suppose that for a feasible in Problem \((P)\) 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 \((y, \beta) \in \mathbb{R}^n \times \mathbb{R}\), satisfying the relation
\[
(\alpha) \quad H_\sigma(y) = \beta - \zeta + \varepsilon, \\
(\beta) \quad G_\sigma(y) \leq \beta,
\]
the following inequality holds
\[
G_\sigma(x) - \beta \geq \langle \nabla H_\sigma(y), x - y \rangle \quad \forall x \in S.
\]
Then, the point \(z \in F\) turns out to be an \(\varepsilon\)-global solution to Problem \((P_\sigma)\) as well as to Problem \((P)\).

\textbf{Acknowledgements}

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

\textbf{References}


545