On a Special Nonlinear Problem Arising in the Study of Convex SIP Problems

Kostyukova O.I.,* Tchemisova T.V. †, Yermalinskaya S.A. ‡

Abstract

We continue a study of convex problems of Semi-Infinite Programming (SIP) started in [6, 7]. In the Implicit Optimality Criterion from [6], we formulated the optimality conditions for convex SIP problem in terms of such the conditions for a special Nonlinear Programming (NLP) problem. In the present paper, we study some specific properties of this nonlinear problem and obtain efficient optimality conditions for it. We show that in the case when the constraint function of the initial SIP problem is analytical, the optimality conditions take the form of criterion.

Key words: Semi-Infinite Programming (SIP), Nonlinear Programming (NLP), Convex Programming (CP), the Slater condition, optimality conditions, irregularity

1 Introduction

In our papers [6, 7], we studied convex problems of Semi-Infinite Programming (SIP). A new approach to solution of such the problems was proposed that is based on new concepts of immobile points and the immobility orders. These concepts were used to formulate and prove the Implicit Optimality Criterion that has permitted us to replace the optimality conditions for the convex SIP problem by such the conditions for a Nonlinear Programming (NLP) problem of a special type, denoted by $NLP(I_*(x^0))$, where x^0 is some feasible solution and $I_*(x^0)$ is some index set. The structure of this nonlinear problem is determined by the properties of the initial SIP problem. We show that in spite of the fact that, in general, the problem $NLP(I_*(x^0))$ is strongly irregular (the mapping corresponding to the equality constraints is degenerated, so the extremum is abnormal in terms of [1]), it is possible to obtain efficient necessary and sufficient optimality conditions for it. Moreover, in the case of the analytical constraint function, these conditions take the form of criterion.

*Institute of Mathematics, Belorussian Academy of Sciences, Belarus

†Department of Mathematics, University of Aveiro, Portugal. The work has been supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal - in the scope of the Program of Pluriannual Funding of CEOC - Centro de Estudos em Optimização e Controlo - of the University of Aveiro.

‡Belorussian State University of Informatics and Radio-electronics, Belarus

The paper is organized as follows. In section 2, we state the convex SIP problem, recall the notions of immobility orders and immobile points, describe shortly the algorithm that defines the immobile points (DIO algorithm). Given a feasible solution x^0 of the convex SIP problem, we use the information about the immobile points of this problem to construct a nonlinear problem $NLP(I_*(x^0))$ and formulate the Implicit Optimality Criterion. The properties of the problem $NLP(I_*(x^0))$ are studied in section 3. A special attention is paid to the case when the constraint function of problem $NLP(I_*(x^0))$ is analytical. In section 4, we formulate necessary and sufficient optimality conditions for the problem $NLP(I_*(x^0))$ and show that in the case of analytical constraint function these conditions take a form of criterion. Section 5 contains a short discussion and some conclusions.

2 Implicit Optimality Criterion for SIP

2.1 Convex SIP problem: definition and basic notions

Consider a SIP problem in the form

$$\begin{aligned} & c(x) \longrightarrow \min, \\ \text{s.t.} \quad & f(x,t) \leq 0, \quad t \in T = [t_*, t^*], \ t_*, t^* \in \mathbb{R}, \end{aligned} \tag{2.1}$$

where $x \in \mathbb{R}^n$; functions $c : \mathbb{R}^n \to \mathbb{R}$ and $f : \mathbb{R}^n \times T \to \mathbb{R}$ are sufficiently smooth in $\mathbb{R}^n$ and $\mathbb{R}^n \times T$, respectively. Suppose that $c(x)$ and $f(x,t)$ are convex w.r.t. x.

Denote by X the feasible set of problem (2.1):

$$X = \{x \in \mathbb{R}^n : \ f(x,t) \leq 0, \ t \in T\}.$$

Problem (2.1) is said to satisfy the Slater condition if there exists $\bar{x} \in X$ such that $f(\bar{x},t) < 0, \ \forall t \in T$.

For any $x \in X$, let

$$T_a(x) = \{t \in T : \ f(x,t) = 0\}$$

be the corresponding set of active points from T.

The following notations will be also used in this paper:

$$\begin{gathered} \nabla c(x) = \partial c(x)/\partial x; \ \nabla_x f(x,t) = \partial f(x,t)/\partial x, \ \nabla_{xx} f(x,t) = \partial^2 f(x,t)/\partial x^2; \\ f^{(0)}(x,t) = f(x,t), \ f^{(s)}(x,t) = \partial^s f(x,t)/\partial t^s, \ s \in \{1,2,\dots\}; \\ \mathcal{N}(q) = \emptyset, \text{ if } q < 0, \ \mathcal{N}(q) = \{0, \ 1, \ \dots, \ q\} \text{ if } q \geq 0, \ q \in Z. \end{gathered}$$

Assumption 2.1 *Suppose* $X \neq \emptyset$ *and there exists* $\bar{x} \in X$ *such that* $|T_a(\bar{x})| < \infty$ *and* $f^{\rho(t)}(\bar{x},t) \neq 0$ *with some* $\rho(t) < \infty$ *for all* $t \in T_a(\bar{x})$.

Definition 2.1 *Given* $t \in T$, *a number* $q(t) \in \{-1,0,1,\dots\}$ *is called the order of immobility (or the immobility order) of* t *in SIP problem (2.1), if*

1. *for each* $x \in X$ *the equalities* $f^{(s)}(x,t) = 0, \ s \in \mathcal{N}(q(t))$, *hold true;*
2. *there exists* $x(t) \in X$ *such that* $f^{(q(t)+1)}(x(t),t) \neq 0$.

It follows from the definition above that problem (2.1) satisfies the Slater condition if and only if $q(t) = -1, \forall t \in T$.

To simplify the further laying out, we can make the following assumption, without loss of generality.

Assumption 2.2 *Suppose that* $q(t_*) = q(t^*) = -1$.

Note that all the results of the paper can be easily extended for the convex SIP problems, where $q(t_*) \geq -1,\ q(t^*) \geq -1$, as it was made in [5].

From Definition 2.1 and Assumption 2.2 it follows that $q(t) + 1$ is even and $f^{(q(t)+1)}(x(t), t) < 0$ for all $t \in T$.

Definition 2.2 *A point* $t \in T$ *is called an immobile point of problem (2.1), if* $q(t) > -1$.

In [6], the finite algorithm (DIO algorithm) that determines immobility orders of all points of the set T in the convex SIP problem (2.1) was justified. In the next subsection, we give a brief description of the algorithm.

2.2 Brief description of DIO algorithm

Initialization. Given any $\bar{x} \in X$, let $T_a(\bar{x}) = \{t_i,\ i \in I(\bar{x})\},\ I(\bar{x}) = \{1,\ 2,\ \ldots,\ p(\bar{x})\}$, $p(\bar{x}) < \infty$. Set: $k = 0,\ I = I(\bar{x}),\ q_i^{(0)} = -1,\ i \in I$.

General iteration ($k \geq 0$). Iteration starts with the odd numbers $q_i^{(k)},\ i \in I$. Denote

$$X_i^{(k)} = \{z \in \mathbb{R}^n :\ f^{(s)}(z,\ t_i) = 0,\ s \in \mathcal{N}(q_i^{(k)});\ f^{(q_i^{(k)}+1)}(z,\ t_i) \leq 0\}; \tag{2.2}$$

$$X^{(k)} = \bigcap_{i \in I} X_i^{(k)}. \tag{2.3}$$

For every $i \in I$ consider a nonlinear problem

$$f^{(q_i^{(k)}+1)}(z,\ t_i) \longrightarrow \min, \text{ s.t. } z \in X^{(k)}. \tag{2.4}$$

Since $\bar{x} \in X^{(k)}$, then either problem (2.4) has an optimal solution, or the objective function $f^{(q_i^{(k)}+1)}(z,\ t_i)$ of this problem is not bounded from below on the feasible set $X^{(k)}$.

Denote by $x^{(i)}$ an optimal solution of problem (2.4), case it exists; otherwise let $x^{(i)}$ be a feasible solution of (2.4) satisfying the inequality $f^{(q_i^{(k)}+1)}(x^{(i)},\ t_i) < 0$.

Consider the set $I^{(k)} = \{i \in I :\ f^{(q_i^{(k)}+1)}(x^{(i)},\ t_i) = 0\}$.

In the case $I^{(k)} \neq \emptyset$, set $q_i^{(k+1)} = q_i^{(k)} + 2,\ i \in I^{(k)};\quad q_i^{(k+1)} = q_i^{(k)},\ i \in I \backslash I^{(k)}$, and start the next iteration with $k := k + 1$.

In the case $I^{(k)} = \emptyset$, algorithm stops with $q(t_i) = q_i^{(k)},\ i \in I;\quad q(t) = -1,\ t \in T \backslash T_a(\bar{x})$.

Note that in [6] it was showed that a number of iterations of the DIO algorithm is finite.

2.3 Implicit Optimality Criterion for convex SIP

It was shown in [6] that the concepts of immobility orders and immobile points are the important characteristics of the constraints of the convex SIP problem (2.1), making it possible to formulate optimality conditions for this problem (with an infinite number of constraints) in terms of optimality conditions for a certain NLP problem (with the finite number of constraints).

In what follows, we suppose that both, Assumption 2.1 and Assumption 2.2, are satisfied.

Denote by T^* the subset of the set T defined as follows:

$$T^* = \{t \in T: \ f(x,t) = 0, \ \forall x \in X\} = \{t \in T: \ q(t) > -1\}.$$

Evidently, $T^* \subseteq T_a(x), \ \forall x \in X$. From Assumption 2.1 we have $|T^*| \leq |T_a(\bar{x})| < \infty$ for some $\bar{x} \in X$. Let us use the following presentation of the set T^*:

$$T^* = \{t_j^0, \ j = 1, \dots, p^0\}, \ p^0 := |T^*| < \infty. \tag{2.5}$$

Consider any $x^0 \in X$ and the corresponding set of active points $T_a(x^0)$. It follows from (2.5) that

$$T^* = \{t_j^0, j = 1, \dots, p^0\} \subset T_a(x^0)$$

and

$$q_j := q(t_j^0) > -1, \ \forall j \in \{1, \dots, p^0\}, \ q(t) = -1, \ \forall t \in T_a(x^0) \setminus T^*.$$

Let us construct a nonlinear problem in the form

$$\begin{array}{lll} & & c(x) \longrightarrow \min, \\ \mathbf{NLP}(I_*(x^0)): & \text{s.t.} & f^{(s)}(x, t_j^0) = 0, \ s \in \mathcal{N}(q_j), \ f^{(q_j+1)}(x, t_j^0) \leq 0, \ j = 1, \dots, p^0, \\ & & f(x, t_j^0) \leq 0, \ \ j \in I_*(x^0), \end{array} \tag{2.6}$$

where

$$\{t_j^0, j \in I_*(x^0)\} \subset T_a(x^0) \setminus T^*, \ I_*(x^0) \subset \{p^0+1, p^0+2, \dots\}, \ |I_*(x^0)| < \infty. \tag{2.7}$$

The following theorem can be proved based on the results from [6].

Theorem 2.1 *[Implicit Optimality Criterion] A feasible solution $x^0 \in X$ is optimal to the convex SIP problem (2.1) if and only if there exists a set of points*

$$\{t_j^0, j \in I_*(x^0)\} \subset T_a(x^0) \setminus T^*, \ \ |I_*(x^0)| < \infty \tag{2.8}$$

such that x^0 is optimal to the problem $NLP(I_*(x^0))$.

Note that there is no assumption about finiteness of the set $T_a(x^0)$.

From Theorem 2.1 it follows that to study optimality of a feasible solution to the convex SIP problem (2.1) it is enough to study optimality of this solution to some nonlinear problem in the form (2.6). The latest problem has a special form as it is constructed taking into account the immobile points and the correspondent immobility orders of problem (2.1).

In the next section, we consider some specific properties of the problem $NLP(I_*(x^0))$ that will be used later to obtain optimality conditions for this problem.

3 Properties of the nonlinear problem $NLP(I_*(x^0))$

Here we consider a nonlinear problem in the form (2.6) with a set $I_*(x^0)$ given by (2.7). Our aim is to study the properties of this problem following from the way it was constructed.

3.1 Convexity of the problem $NLP(I_*(x^0))$

To show that the feasible set of problem (2.6) is convex, let us analyze the algorithm of determination of immobility orders (DIO algorithm) suggested in [6] and shortly described in section 2.

Let k_* be the number of iteration where algorithm DIO has stopped. Then $q(t_i) = q_i^{(k_*)}, i \in I = \{1, \dots, p(\bar{x})\}$.

Theorem 3.1 *For any* $k \in \{0, \dots, k_*\}$, *the set* $X^{(k)}$ *constructed on the correspondent iteration of DIO algorithm, is convex and functions* $f^{(q_i^{(k)}+1)}(x, t_i)$, $i \in I$, *are convex w.r.t.* x *in* $X^{(k)}$.

Proof. From DIO algorithm we have $q_i^{(k+1)} \geq q_i^{(k)}$, hence the following inclusions are valid:

$$X_i^{(k+1)} \subseteq X_i^{(k)},\ i \in I,\ k \in \{0, \dots, k_* - 1\}. \tag{3.1}$$

Taking into account (2.3), we obtain

$$X^{(k+1)} \subseteq X^{(k)},\ k \in \{0, \dots, k_* - 1\}. \tag{3.2}$$

We will prove the statement of the theorem by induction.

First suppose that $k = 0$. All the sets $X_i^{(0)} = \{x \in \mathbb{R}^n : \ f(x, t_i) \leq 0\},\ i \in I$, are convex as $f(x, t)$ is convex w.r.t. x for all $t \in T$. Then from (2.3) we conclude that $X^{(0)}$ is convex too, being the intersection of the convex sets.

By construction, $q_i^{(0)} = -1$ for any $i \in I$. The functions $f^{(0)}(x, t_i) = f(x, t_i)$, $i \in I$, are convex w.r.t. x in $\mathbb{R}^n$ and, therefore they are convex in the set $X^{(0)}$. Thus the theorem is valid for $k = 0$.

Suppose now that the statement of the theorem is true for $k < k_*$. Let us prove the theorem for $k + 1$.

By the algorithm, $q_i^{(k+1)} = q_i^{(k)}$ for all $i \in I \backslash I^{(k)}$. Consequently,

$$X_i^{(k+1)} = X_i^{(k)},\ \forall i \in I \backslash I^{(k)}. \tag{3.3}$$

Consider now any $i \in I^{(k)}$. Then by DIO algorithm we have

$$\min_{x \in X^{(k)}} f^{(q_i^{(k)}+1)}(x, t_i) = 0,\ \forall i \in I^{(k)}. \tag{3.4}$$

For $i \in I^{(k)}$, let us show that the following set is convex:

$$S_i := \{x \in X^{(k)} : f^{(q_i^{(k)}+1)}(x, t_i) = 0, f^{(q_i^{(k)}+2)}(x, t_i) = 0, f^{(q_i^{(k)}+3)}(x, t_i) \leq 0\}. \tag{3.5}$$

If $x, y \in S_i$, then

$$x, y \in X^{(k)}, \tag{3.6}$$

$$f^{(q_i^{(k)}+1)}(x, t_i) = 0, f^{(q_i^{(k)}+1)}(y, t_i) = 0, f^{(q_i^{(k)}+2)}(x, t_i) = 0, f^{(q_i^{(k)}+2)}(y, t_i) = 0, \tag{3.7}$$

$$f^{(q_i^{(k)}+3)}(x, t_i) \leq 0,\ f^{(q_i^{(k)}+3)}(y, t_i) \leq 0. \tag{3.8}$$

Denote

$$x(\alpha) = \alpha x + (1 - \alpha)y.$$

By assumption of induction, the set $X^{(k)}$ is convex and the functions $f^{(q_i^{(k)}+1)}(x, t_i)$ are convex in $X^{(k)}$ for any $i \in I$. Consequently,

$$x(\alpha) \in X^{(k)},\ \alpha \in [0, 1], \tag{3.9}$$

and

$$f^{(q_i^{(k)}+1)}(x(\alpha), t_i) \leq \alpha f^{(q_i^{(k)}+1)}(x, t_i) + (1 - \alpha) f^{(q_i^{(k)}+1)}(y, t_i),\ i \in I^{(k)},\ \alpha \in [0, 1]. \tag{3.10}$$

Taking into account (3.7) and (3.10), we obtain

$$f^{(q_i^{(k)}+1)}(x(\alpha), t_i) \leq 0, \quad i \in I^{(k)},\ \alpha \in [0, 1]. \tag{3.11}$$

Relations (3.4) and (3.9) together with (3.11) imply

$$f^{(q_i^{(k)}+1)}(x(\alpha), t_i) = 0, \quad i \in I^{(k)},\ \alpha \in [0, 1]. \tag{3.12}$$

Furthermore, since $x,\ x(\alpha)$, and y belong to $X^{(k)}$, then the following equalities hold true:

$$f^{(s)}(x, t_i) = f^{(s)}(x(\alpha), t_i) = f^{(s)}(y, t_i) = 0,\ s \in \mathcal{N}(q_i^{(k)}),\ i \in I,\ \alpha \in [0, 1]. \tag{3.13}$$

From convexity of $f(x, t)$ w.r.t. x and from its sufficient smoothness, we conclude that

$$f(x(\alpha), t_i + \Delta t) \leq \alpha f(x, t_i + \Delta t) + (1 - \alpha) f(y, t_i + \Delta t),\ i \in I,\ \alpha \in [0, 1], \tag{3.14}$$

and that for any $m > 0$, $z \in \mathbb{R}^n$ the following Taylor expansion is valid:

$$f(z, t_i + \Delta t) = \sum_{l=0}^{m} \frac{1}{l!} f^{(l)}(z, t_i) \Delta t^l + o(\Delta t^m),\ i \in I^{(k)}. \tag{3.15}$$

Substitute (3.15) with $z = x(\alpha),\ z = x,\ z = y$, and $m = q_i^{(k)} + 2$ in (3.14). Taking into consideration (3.7), (3.12), (3.13), we obtain

$$f^{(q_i^{(k)}+2)}(x(\alpha), t_i) \Delta t^{(q_i^{(k)}+2)} + o(\Delta t^{(q_i^{(k)}+2)}) \leq o(\Delta t^{(q_i^{(k)}+2)}),\ i \in I^{(k)},\ \alpha \in [0, 1]. \tag{3.16}$$

Divide (3.16) by $\Delta t^{(q_i^{(k)}+2)}$ and pass to limit, taking into account that all $q_i^{(k)}$ are odd:

$$\lim_{\Delta t \to +0} \frac{f^{(q_i^{(k)}+2)}(x(\alpha), t_i) \Delta t^{(q_i^{(k)}+2)} + o(\Delta t^{(q_i^{(k)}+2)})}{\Delta t^{(q_i^{(k)}+2)}} \leq \lim_{\Delta t \to +0} \frac{o(\Delta t^{(q_i^{(k)}+2)})}{\Delta t^{(q_i^{(k)}+2)}} = 0,$$

$$\lim_{\Delta t \to -0} \frac{f^{(q_i^{(k)}+2)}(x(\alpha), t_i)\Delta t^{(q_i^{(k)}+2)} + o(\Delta t^{(q_i^{(k)}+2)})}{\Delta t^{(q_i^{(k)}+2)}} \geq \lim_{\Delta t \to -0} \frac{o(\Delta t^{(q_i^{(k)}+2)})}{\Delta t^{(q_i^{(k)}+2)}} = 0,$$

Hence,

$$f^{(q_i^{(k)}+2)}(x(\alpha), t_i) = 0, \; i \in I^{(k)}, \; \alpha \in [0, 1]. \tag{3.17}$$

Now, let us substitute in (3.14) the expansion (3.15) with $m = q_i^{(k)} + 3$ for $z = x(\alpha), \; z = x, \; z = y$. Then, considering (3.7), (3.12), and (3.17), we have

$$f^{(q_i^{(k)}+3)}(x(\alpha), t_i)\Delta t^{(q_i^{(k)}+3)} + o(\Delta t^{(q_i^{(k)}+3)}) \leq \alpha f^{(q_i^{(k)}+3)}(x, t_i)\Delta t^{(q_i^{(k)}+3)} +$$

$$+(1-\alpha) f^{(q_i^{(k)}+3)}(y, t_i)\Delta t^{(q_i^{(k)}+3)} + o(\Delta t^{(q_i^{(k)}+3)}), i \in I^{(k)}, \alpha \in [0, 1].$$

Divide the inequality above by $\Delta t^{(q_i^{(k)}+3)}$ and let $\Delta t \longrightarrow 0$. Since $q_i^{(k)}, i \in I^{(k)}$, are odd, then

$$f^{(q_i^{(k)}+3)}(x(\alpha), t_i) \leq \alpha f^{(q_i^{(k)}+3)}(x, t_i) + (1-\alpha) f^{(q_i^{(k)}+3)}(y, t_i), i \in I^{(k)}, \alpha \in [0, 1]. \tag{3.18}$$

From (3.8) and (3.18) we obtain

$$f^{(q_i^{(k)}+3)}(x(\alpha), t_i) \leq 0, \; i \in I^{(k)}, \; \alpha \in [0, 1]. \tag{3.19}$$

Relations (3.9), (3.12), (3.17), and (3.19) yield $x(\alpha) \in S_i$ for all $\alpha \in [0, 1]$. Therefore, the set S_i is convex for any $i \in I^{(k)}$.

From (2.3) and (3.1)-(3.3), we get

$$X^{(k+1)} = X^{(k)} \bigcap \left(\bigcap_{i \in I^{(k)}} X_i^{(k+1)} \right) = \bigcap_{i \in I^{(k)}} \left(X^{(k)} \bigcap X_i^{(k+1)} \right) = \bigcap_{i \in I^{(k)}} S_i. \tag{3.20}$$

Taking into account convexity of the sets $S_i, \; i \in I^{(k)}$, we conclude that the set $X^{(k+1)}$ is convex. By DIO algorithm, $q_i^{(k+1)} = q_i^{(k)}, \; i \in I \backslash I^{(k)}$. Then functions

$$f^{(q_i^{(k+1)}+1)}(x, t_i) = f^{(q_i^{(k)}+1)}(x, t_i), \; i \in I \backslash I^{(k)},$$

are convex w.r.t. x in $X^{(k+1)}$ since they are convex in $X^{(k)}$ and (3.2) is true.

For $i \in I^{(k)}$ we have $q_i^{(k+1)} = q_i^{(k)} + 2$. Hence

$$f^{(q_i^{(k+1)}+1)}(x, t_i) = f^{(q_i^{(k)}+3)}(x, t_i), \; i \in I^{(k)},$$

all these functions being convex w.r.t. x in S_i that follows from inequalities (3.18). Taking into account (3.20), we can state also that these functions are convex w.r.t. x in $X^{(k+1)}$.

Therefore, we have proved that the statement of the theorem is valid for $k+1$, that concludes the proof. ■

Let $\bar{Y}$ be the set defined as follows:

$$\bar{Y} = \{x \in \mathbb{R}^n : \ f^{(s)}(x, t_j^0) = 0, \ s \in \mathcal{N}(q_j), \ f^{(q_j+1)}(x, t_j^0) \leq 0, \ j = 1, \dots, p^0\}. \quad (3.21)$$

According to Remark 3.1 in [6], there exists a vector $\tilde{x} \in X$ such that

$$f^{(s)}(\tilde{x}, t) = 0, \ s \in \mathcal{N}(q(t)), \ f^{(q(t)+1)}(\tilde{x}, t) < 0, \quad \forall t \in T. \quad (3.22)$$

Having assumed in DIO algorithm that $\bar{x} = \tilde{x}$ with $\tilde{x}$ satisfying (3.22), we obtain $T_a(\tilde{x}) = T^*$, $I = I(\tilde{x}) = \{1, \dots, p^0\}$ and $\bar{Y} = X^{(k_*)}$. Therefore, by Theorem 3.1, the set $\bar{Y}$ is convex.

Let us rewrite problem (2.6) in the equivalent form

$$\begin{aligned} c(x) &\longrightarrow \min, \\ \text{s.t.} \quad x \in \bar{Y}, \quad f(x, t_j^0) &\leq 0, \ j \in I_*(x^0). \end{aligned} \quad (3.23)$$

Since functions $c(x)$, $f(x,t)$ are convex w.r.t. x and $\bar{Y}$ is the convex set, then for any $I_*(x^0) \subset \{p^0+1, \dots, p(x^0)\}$ problem (3.23), and, therefore, problem (2.6) is a Convex Programming (CP) problem (see [3]). The following theorem is proved.

Theorem 3.2 *For any $x^0 \in X$ and any $I_*(x^0)$ defined in (2.7) the problem NLP($I_*(x^0)$) is a Convex Programming problem.*

3.2 Reducing the number of constraints in the problem $NLP(I_*(x^0))$

As a rule, the number of constraints in the CP problem (3.23) is too large. Therefore, the following result is of special interest.

Theorem 3.3 *Let x^0 be an optimal solution of the problem*

$$\begin{aligned} c(x) &\longrightarrow \min, \\ \textit{s.t.} \quad f_i(x) \leq 0, \quad i &\in I = \{1, \dots, m\}, \ x \in H, \end{aligned} \quad (3.24)$$

where $c(x)$, $f_i(x)$, $i \in I$, $x \in \mathbb{R}^n$, are sufficiently smooth convex functions and $H \subset \mathbb{R}^n$ is a convex set. Suppose that problem (3.24) satisfies the Slater condition, i.e. for some $\bar{x} \in H$ the inequalities $f_i(\bar{x}) < 0$, $i \in I$, hold true. If $|I| = m > n$, then there exists a subset $I_* \subset I$, $|I_*| \leq n$ such that vector x^0 is the optimal solution of the problem

$$\begin{aligned} c(x) &\longrightarrow \min, \\ \textit{s.t.} \quad f_i(x) \leq 0, \quad i &\in I_*, \ x \in H. \end{aligned} \quad (3.25)$$

Proof. First, let us show that for any $m > n$ there exists a subset $I_* \subset I$ such that $|I_*| = m - 1$ and x^0 is the solution of problem (3.25).

Taking into account the assumptions of the theorem, x^0 is optimal in (3.24) if and only if there exists a set of numbers (see [9], p.237–239) $\lambda_i \geq 0$, $i \in I$, such that $\lambda_i f_i(x^0) = 0$, $i \in I$, and

$$\left(\nabla c(x^0) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x^0) \right)' h \geq 0, \quad \forall h \in \mathcal{A}(x^0 | H), \quad (3.26)$$

with

$$\mathcal{A}(x^0|H) = \{h \in \mathbb{R}^n : \ h = x - x^0, \quad x \in H\}.$$

If $\lambda_{i_0} = 0$ for some $i_0 \in I$, then set $I_* := I \backslash \{i_0\}$, $\bar{\lambda}_i := \lambda_i, \ i \in I_*$. Evidently, $|I_*| = m - 1$ and the following inequalities take place

$$\left(\nabla c(x^0) + \sum_{i \in I_*} \bar{\lambda}_i \nabla f_i(x^0)\right)' h \geq 0, \forall h \in \mathcal{A}(x^0|H), \ \bar{\lambda}_i \geq 0, \ \bar{\lambda}_i f_i(x^0) = 0, i \in I_*. \quad (3.27)$$

Then x^0 optimal to problem (3.25) too.

Let us suppose now that $\lambda_i > 0, \ i \in I$. Since $m > n$, then there exists a non-vanishing vector $\alpha = (\alpha_1, \ \dots, \ \alpha_m)'$ such that

$$\sum_{i=1}^{m} \alpha_i \nabla f_i(x^0) = 0. \quad (3.28)$$

Without loss of generality we can suppose that $I^- = \{i \in I : \ \alpha_i < 0\} \neq \emptyset$ (indeed, otherwise instead of α we can take vector $-\alpha$). Let

$$\Theta_0 = \min_{i \in I^-} \left(-\frac{\lambda_i}{\alpha_i}\right) = -\frac{\lambda_{i_*}}{\alpha_{i_*}} > 0, \ i_* \in I^-.$$

Multiply (3.28) by Θ_0 and add to the expression in brackets in the left hand side of (3.26). It results in

$$\left(\nabla c(x^0) + \sum_{i=1}^{m} \tilde{\lambda}_i \nabla f_i(x^0)\right)' h \geq 0, \quad \forall h \in \mathcal{A}(x^0|H),$$

where $\tilde{\lambda}_i = \lambda_i + \Theta_0 \alpha_i \geq 0, \ i \in I$. Note that $\tilde{\lambda}_{i_*} = 0$. Set $I_* = I \backslash \{i_*\}, \quad \bar{\lambda}_i = \tilde{\lambda}_i, \ i \in I_*$. Then relations (3.27) are valid and x^0 is the optimal solution of problem (3.25), with $|I_*| = m - 1$. If $|I_*| \leq n$, the theorem is proved. Otherwise, set $I := I_*$ and repeat the same reasoning. ■

From the theorem we can deduce that if in the convex problem (3.23) the index set $I_*(x^0)$ satisfies the conditions $n < |I_*(x^0)| < \infty$, then there exists a subset $\hat{I}_* \subset I_*(x^0)$ such that $|\hat{I}_*| \leq n$ and any optimal solution to (3.23) is also optimal to the problem

$$\begin{aligned} & c(x) \longrightarrow \min, \\ \text{s.t.} \quad & x \in \bar{Y}, f(x, t_j^0) \leq 0, \ j \in \hat{I}_*. \end{aligned} \quad (3.29)$$

Therefore Theorem 2.1 can be formulated with the inequality in (2.8) replaced by $|I_*(x^0)| \leq n$.

3.3 The case of analytical constraint function

Here we will show that in the case when the constraint function of SIP problem is analytical, the equality constraints of the correspondent nonlinear problem $NLP(I_*(x^0))$ can be equivalently replaced by linear ones.

First consider NLP problem in the form

$$c(x) \to \min, \; x \in \mathcal{Y} = H \cap G, \tag{3.30}$$

where

$$H = \{x \in \mathbb{R}^n : \; h(x) = 0\}, \; G = \{x \in \mathbb{R}^n : \; g_i(x) \le 0, \; i = 1, \ldots, m\} \tag{3.31}$$

with $h(x) = (h_j(x), \; j = 1, \ldots, l)$. The following proposition is evident.

Proposition 3.1 *If* $\mathcal{Y} \neq \emptyset$, *then there exists a finite set of vectors* $y_i \in \mathcal{Y}, \; i = 0, 1, \ldots, r$, *such that*

$$rank\,(y_i - y_0, \; i = 1, \ldots, r) = r, \tag{3.32}$$

$$rank\,(x - y_0, \; y_i - y_0, \; i = 1, \ldots, r) = r, \; \forall x \in \mathcal{Y}. \tag{3.33}$$

Evidently, here $0 \le r \le n$. From relation (3.33) we conclude that for any $x \in \mathcal{Y}$ there exist numbers $\alpha_i, \; i = 1, \ldots, r$, such that $x - y_0 = \sum\limits_{i=1}^{r} \alpha_i (y_i - y_0)$. Then, having denoted $\alpha_0 = 1 - \sum\limits_{i=1}^{r} \alpha_i$, we obtain

$$\mathcal{Y} \subset \mathcal{X}, \tag{3.34}$$

where

$$\mathcal{X} := \{x \in \mathbb{R}^n : \; x = \sum_{i=0}^{r} \alpha_i y_i, \; \sum_{i=0}^{r} \alpha_i = 1\} \tag{3.35}$$

with some $y_i \in \mathcal{Y}, \; i = 0, \ldots, r$, satisfying conditions (3.32), (3.33).

Let us suppose now that in (3.31) functions $h_j(x), \; j = 1, \ldots, l$, are analytical.

Theorem 3.4 *Suppose that the feasible set* $\mathcal{Y}$ *of problem (3.30) is nonempty and convex, function* $h(x)$ *in (3.31) being analytical. Then the following representation holds true:*

$$\mathcal{Y} = \mathcal{X} \cap G. \tag{3.36}$$

Proof. Suppose first that $r = 0$, i.e. the feasible set $\mathcal{Y}$ contains a unique element y_0. Then $\mathcal{Y} = \mathcal{X} = \{y_0\} \subset G$, and (3.36) is satisfied.

Now let us consider a situation, where $r \ge 1$. Relations (3.34) and $\mathcal{Y} \subset G$ imply

$$\mathcal{Y} \subset \mathcal{X} \cap G. \tag{3.37}$$

If we manage to prove

$$\mathcal{X} \subset H \tag{3.38}$$

then, evidently, we obtain $\mathcal{Y} \subset \mathcal{X} \cap G \subset H \cap G = \mathcal{Y}$ and (3.36) will be proved.

Thus, let us prove (3.38). Since $\mathcal{Y} \subset H$, then it is sufficient to prove the following implication:

$$\mathcal{X} \backslash \mathcal{Y} \subset H. \tag{3.39}$$

Consider the set $\mathcal{X}\backslash\mathcal{Y}$. Suppose that there exists some vector z such that $z \in \mathcal{X}\backslash\mathcal{Y}$ and $z \notin H$. Since $z \in \mathcal{X}$, then for some numbers $\alpha_i \in \mathbb{R}$, $i = 0, \dots, r$, we have

$$z = \sum_{i=0}^{r} \alpha_i y_i, \quad \sum_{i=0}^{r} \alpha_i = 1. \tag{3.40}$$

Vectors y_i, $i = 0, \dots, r$, belong to the convex set $\mathcal{Y}$, therefore

$$y_* := \sum_{i=0}^{r} \frac{1}{r+1} y_i \in \mathcal{Y}. \tag{3.41}$$

Consider vector $z(\mu) = \mu z + (1-\mu) y_*$, $\mu \in \mathbb{R}$. Taking into account (3.40) and (3.41), we have

$$z(\mu) = \mu \sum_{i=0}^{r} \alpha_i y_i + (1-\mu) \sum_{i=0}^{r} \frac{1}{r+1} y_i = \sum_{i=0}^{r} \beta_i(\mu) y_i, \tag{3.42}$$

where

$$\beta_i(\mu) = \mu \alpha_i + (1-\mu) \frac{1}{r+1}, \; i = 0, \dots, r. \tag{3.43}$$

From (3.40) and (3.43) we obtain

$$\sum_{i=0}^{r} \beta_i(\mu) = \mu \sum_{i=0}^{r} \alpha_i + (1-\mu) \sum_{i=0}^{r} \frac{1}{r+1} = 1, \; \forall \mu \in \mathbb{R}, \tag{3.44}$$

hence $z(\mu) \in \mathcal{X}$. Let $\mu_* = \min\limits_{i=0,\dots,r} \mu_i$, where

$$\mu_i = \begin{cases} \dfrac{1}{1-\alpha_i(r+1)} & \text{, if } \alpha_i < 0, \\ \dfrac{1}{r} & \text{, if } 0 \leq \alpha_i \leq 1 \\ \dfrac{r}{\alpha_i(r+1)-1} & \text{, if } \alpha_i > 1. \end{cases}$$

It is easy to verify that $\mu_* > 0$ and

$$\beta_i(\mu) \geq 0, \; i = 0, \dots, r, \quad \forall \mu \in [0, \mu_*]. \tag{3.45}$$

Since $y_i \in \mathcal{Y}$, $i = 0, \dots, r$, where $\mathcal{Y}$ is convex, then, taking into account (3.42), (3.44), and (3.45), we obtain $z(\mu) \in \mathcal{Y}$, $\forall \mu \in [0, \mu_*]$. Hence

$$h_j(z(\mu)) \equiv 0, \; \forall \mu \in [0, \mu_*], \quad j = 1, \dots, l. \tag{3.46}$$

Since it was assumed that the functions $h_j(x)$, $x \in \mathbb{R}^n$, $j = 1, \dots, l$, are analytical, then the functions $\bar{h}_j(\mu) = h_j(z(\mu))$, $\mu \in \mathbb{R}$, $j = 1, \dots, l$, are analytical too and by (3.46) we have

$$\bar{h}_j(\mu) = h_j(z(\mu)) \equiv 0, \; \forall \mu \in \mathbb{R}, \quad j = 1, \dots, l.$$

Therefore, $z(\mu) \in H$ for all $\mu \in \mathbb{R}$ and, in particular, $z(1) = z \in H$. The contradiction obtained proves that there is no any $z \in \mathcal{X}\backslash\mathcal{Y}$ such that $z \notin H$ and, consequently, inclusion in (3.39) holds true and the theorem is proved. ■

Let $\psi_i \in \mathbb{R}^n,\ i = 1, \dots, r$, be a set of vectors satisfying the conditions

$$\begin{aligned} &\operatorname{rank}(\psi_i,\ i = 1, \dots, r) = r, \\ &M^k(y_0, \psi_i) = 0,\ k = 1, 2, \dots;\ \ i = 1, \dots, r, \end{aligned} \tag{3.47}$$

where $y_0 \in \mathcal{Y}$,

$$M^k(x, \psi) = D^k(x, \psi)\psi,\ D^k(x, \psi) = \frac{\partial M^{k-1}(x, \psi)}{\partial x},\ D^1(x, \psi) = D^1(x) = \frac{\partial h(x)}{\partial x}.$$

Denote

$$\mathcal{X}(\psi_1, \dots, \psi_r) = \{x \in \mathbb{R}^n :\ x = y_0 + \sum_{i=1}^{r} \alpha_i \psi_i\}.$$

Here $\alpha_i \in \mathbb{R},\ i = 1, \dots, r$.

Proposition 3.2 *Assume that there exists $y_0 \in \mathcal{Y}$ such that $g_j(y_0) < 0$, $i = 1, \dots, m$, and suppose that conditions of Theorem 3.4 are fulfilled. Then the set $\mathcal{X}$ defined in (3.35) can be represented in the form*

$$\mathcal{X} = \mathcal{X}(\psi_1, \dots, \psi_r) \tag{3.48}$$

with an arbitrary set of vectors $\psi_1, \dots, \psi_r$ satisfying (3.47).

Proof. First, we will show that there exists a set of vectors satisfying (3.47). Really, let $y_i, i = 0, 1, \dots, r$, be vectors satisfying conditions of Proposition 3.1. It is easy to check that the vectors

$$\tilde{\psi}_i := y_i - y_0,\ \ i = 1, \dots, r, \tag{3.49}$$

satisfy (3.47).

Now let us show that there are no $r + 1$ linearly independent vectors $\psi_i,\ \ i = 1, \dots, r + 1$, such that

$$M^k(y_0, \psi_i) = 0,\ k = 1, 2, \dots;\ \ i = 1, \dots, r + 1. \tag{3.50}$$

Arguing by contradiction, suppose that there exist $r+1$ of linearly independent vectors $\psi_i \in \mathbb{R}^n,\ i = 1, \dots, r + 1$, that satisfy (3.50). Since $g_j(y_0) < 0$ for all $i = 1, \dots, m$, there exists $\varepsilon_0 > 0$ such that

$$B(y_0, \varepsilon_0) = \{x \in \mathbb{R}^n :\ \|x - y_0\| \leq \varepsilon_0\} \subset G. \tag{3.51}$$

The equalities (3.50) are satisfied for all the vectors $\psi_i,\ i = 1, \dots, r + 1$. Therefore, (3.50) are satisfied also for vectors

$$\psi_i(\alpha) = \alpha\psi_i,\ i = 1, \dots, r + 1,\ \forall \alpha \in \mathbb{R}. \tag{3.52}$$

Denote

$$\bar{\alpha} = \frac{\varepsilon_0}{\|\psi_i\|},\ x_i = y_0 + \psi_i(\bar{\alpha}),\ i = 1, \dots, r + 1. \tag{3.53}$$

By construction, $x_i \in B(y_0, \varepsilon_0),\ i = 1, \dots, r+1$, and, hence, $x_i \in G,\ i = 1, \dots, r+1$. Furthermore, since vectors $\psi_i(\bar{\alpha}),\ i = 1, \dots, r+1$, satisfy equalities (3.47), then $x_i \in H,\ i = 1, \dots, r+1$, and, consequently,

$$x_i \in \mathcal{Y},\ i = 1, \dots, r+1. \tag{3.54}$$

On the other hand, by assumption, all the vectors $\psi_i,\ i = 1, \dots, r+1$ are linearly independent. Therefore,

$$rank(x_i - y_0,\ i = 1, \dots, r+1) = rank(\psi_i(\bar{\alpha}),\ i = 1, \dots, r+1) = r+1.$$

Since $y_0 \in \mathcal{Y}$, $x_i \in \mathcal{Y},\ i = 1, \dots, r+1$, then the latest formula contradicts to (3.33). The contradiction obtained proves that there exist exactly r vectors $\psi_i \in \mathbb{R}^n,\ i = 1, \dots, r$, satisfying (3.47).

Now let us show that

$$\mathcal{X}(\psi_1, \dots, \psi_r) = \mathcal{X}(\bar{\psi}_1, \dots, \bar{\psi}_r), \tag{3.55}$$

where $\psi_1, \dots, \psi_r$ and $\bar{\psi}_1, \dots, \bar{\psi}_r$ are two set of vectors satisfying (3.47). Due to the fact that there are no $r+1$ linearly independent vectors satisfying (3.50), we conclude that for each vector $\bar{\psi}_k$, $k = 1, \dots, r$, there exist numbers β_{ik}, $i = 1, \dots, r$, such that

$$\bar{\psi}_k = \sum_{i=1}^{r} \beta_{ik} \psi_i, \quad k = 1, \dots, r.$$

Then

$$\mathcal{X}(\bar{\psi}_1, \dots, \bar{\psi}_r) := \{x \in \mathbb{R}^n : x = y_0 + \sum_{k=1}^{r} \bar{\alpha}_k \bar{\psi}_k\}$$

$$= \{x \in \mathbb{R}^n : x = y_0 + \sum_{i=1}^{r} \alpha_i \psi_i\} =: \mathcal{X}(\psi_1, \dots, \psi_r),$$

where $\alpha_i := \sum_{k=1}^{r} \bar{\alpha}_k \beta_{ik}$. Relation (3.55) is proved.

The statement of Proposition 3.2 (relation (3.48)) follows from (3.55) and the condition $\mathcal{X} = \mathcal{X}(\tilde{\psi}_1, \dots, \tilde{\psi}_r)$, where $\tilde{\psi}_1, \dots, \tilde{\psi}_r$ are defined in (3.49). ■

Let us consider now the problem $NLP(I_*(x^0))$ in the form (2.6). The set Y of its feasible solutions can be represented in the form $Y = H \cap G$, where

$$H = \{x \in \mathbb{R}^n :\ h(x) = 0\}, \quad G = \{x \in \mathbb{R}^n :\ g_j(x) \le 0,\ j \in J(x^0)\},$$
$$h(x) = (h_{js}(x), s \in \mathcal{N}(q_j),\ j = 1, \dots, p^0), \tag{3.56}$$

$$h_{js}(x) = f^{(s)}(x, t_j^0),\ s \in \mathcal{N}(q_j); \ g_j(x) = f^{(q_j+1)}(x, t_j^0),\ j = 1, \dots, p^0; \tag{3.57}$$
$$g_j(x) = f(x, t_j^0), j \in I_*(x^0),$$

$$J(x^0) := \{1, \dots, p^0\} \cup I_*(x^0). \tag{3.58}$$

Suppose that the function $f(x,t),\ x\in\mathbb{R}^n,\ t\in T\subset\mathbb{R}$, is analytical. Then the derivatives of all orders of $f(x,t)$ are analytical functions too.

Consequently, the feasible set Y of problem (2.6) satisfies all conditions of Theorem 3.4 and Proposition 3.2 with $y_0=\tilde{x}$ where $\tilde{x}$ is defined in (3.22). Hence it can be presented in the form $Y=\mathcal{X}\cap G$. Then, taking into account representation (3.48) obtained before, the problem $NLP(I_*(x^0))$ can be written in the equivalent form

$$\begin{gathered} c(x)\to\min_{x,\alpha},\ x\in\mathbb{R}^n,\ \alpha\in\mathbb{R}^r,\\ x-\tilde{x}-\Psi\alpha=0,\\ f^{(q_j+1)}(x,t_i^0)\le 0,\ i=1,\dots,p^0,\ f(x,t_j^0)\le 0,\ j\in I_*(x^0), \end{gathered}\tag{3.59}$$

where $\Psi=(\psi_i,\ i=1,\dots,r)$, and the vectors $\psi_i\in\mathbb{R}^n,\ i=1,\dots,r$, satisfy (3.47) with $h(x)$ defined in (3.56).

Observation. For the linearly independent vectors $\psi_i\in\mathbb{R}^n,\ i=1,\dots,r$, from (3.47), there always exist $n-r$ linearly independent vectors $\psi_i\in\mathbb{R}^n,\ i=r+1,\dots,n$, orthogonal to them. Therefore we obtain a system of n linearly independent vectors $\psi_j,\ j=1,\dots,n$, satisfying

$$\psi_i'\psi_j=0,\ i=r+1,\dots,n,\ j=1,\dots,r,$$

that form in $\mathbb{R}^n$ a full basis. Then the set $\mathcal{X}$ can be presented as follows:

$$\mathcal{X}=\{x\in\mathbb{R}^n:\ Ax=A\tilde{x}\},$$

where $A'=(\psi_i,\ i=r+1,\dots,n)$, and problem (3.59) takes the form

$$\begin{gathered} c(x)\to\min,\ x\in\mathbb{R}^n,\\ Ax-A\tilde{x}=0,\\ f^{(q_j+1)}(x,t_i^0)\le 0,\ i=1,\dots,p^0,\ f(x,t_j^0)\le 0,\ j\in I_*(x^0). \end{gathered}\tag{3.60}$$

Therefore, we have proved here that in the case when the equality constraints of the problem $NLP(I_*(x^0))$ (represented, generally, by nonlinear functions) are analytical, a matrix Ψ exists such that the problem $NLP(I_*(x^0))$ is equivalent to problem (3.59) with linear equality constraints. By construction, problem (3.59) satisfies the Slater condition. We will use late this property to obtain optimality conditions for the problem $NLP(I_*(x^0))$ in the form of criterion.

4 Optimality conditions for the problem $NLP(I_*(x^0))$

In this section, we study the optimality conditions for the problem $NLP(I_*(x^0))$ in the form (2.6).

Using notations (3.57), (3.58), we can rewrite problem (2.6) in the form

$$\begin{aligned} & c(x)\longrightarrow\min,\\ \text{s.t.}\quad & h_{js}(x)=0,\ \ s\in\mathcal{N}(q_j),\ j=1,\dots,p^0,\\ & g_j(x)\le 0,\ j\in J(x^0). \end{aligned}\tag{4.1}$$

Then its feasible set Y is given by

$$Y=\{x\in\mathbb{R}^n:\ h_{js}(x)=0,\ s\in\mathcal{N}(q_j),\ j=1,\dots,p^0;\ g_j(x)\le 0,\ j\in J(x^0)\}.$$

From Theorem 3.2 it follows that the set Y is convex. The Lagrange function for problem (4.1) has the form

$$\mathcal{L}(x,\lambda) = \lambda_0 c(x) + \sum_{j=1}^{p^0} \sum_{s\in\mathcal{N}(q_j)} \lambda_{js} h_{js}(x) + \sum_{j\in J(x^0)} \mu_j g_j(x) \tag{4.2}$$

with the Lagrange multipliers vector

$$\lambda = (\lambda_0, \lambda_{js}, s \in \mathcal{N}(q_j),\ j = 1,\ldots,p^0;\ \mu_j, j \in J(x^0)). \tag{4.3}$$

Given $x^0 \in Y$, let $J_A(x^0) = \{j \in J(x^0) : g_j(x^0) = 0\}$ be the correspondent active index set.

4.1 Necessary Optimality Conditions

In [7], it is proved that when SIP problem (2.1) does not satisfy the Slater condition, application of the classical necessary optimality conditions of NLP (see [3], for example) to the problem $NLP(I_*(x^0))$ will not give any informative optimality conditions since these conditions are trivially fulfilled for any feasible $x \in X$. It happens because the equality constraints of the problem $NLP(I_*(x^0))$ determine a mapping that is irregular. Nontrivial necessary optimality conditions for such cases (generalized necessary conditions) are suggested in [1, 2, 8], et al., where in the absence of regularity a so called $2-$ or $p-$ regularity is studied. In the specific case of the problem $NLP(I_*(x^0))$, the order of irregularity is greater (as a rule) than two and to obtain necessary optimality conditions using the techniques from [1, 2, 8] one needs to use constructions that are rather bulky (see [1]).

4.2 Sufficient Optimality Conditions

4.2.1 The first order sufficient optimality conditions

Consider the problem $NLP(I_*(x^0))$ in the form (4.1). Let us prove the following theorem.

Theorem 4.1 *[The first order sufficient optimality condition] Suppose that for a feasible $x^0 \in Y$ there exist numbers*

$$\lambda_{js}, s \in \mathcal{N}(q_j),\ j = 1,\ldots,p^0;\ \mu_j \geq 0,\ j \in J(x^0), \tag{4.4}$$

such that

$$\mu_j g_j(x^0) = 0, \quad j \in J(x^0),$$

$$\nabla c(x^0) + \sum_{i=1}^{p^0} \sum_{s\in\mathcal{N}(q_j)} \lambda_{js} \nabla h_{js}(x^0) + \sum_{j\in J(x^0)} \mu_j \nabla g_j(x^0) = 0. \tag{4.5}$$

Then x^0 is an optimal solution of the convex problem (4.1).

Proof. Arguing by contradiction, suppose that there exists $y \in Y$ such that

$$c(y) < c(x^0). \tag{4.6}$$

Consider the vector

$$x(\alpha) = \alpha y + (1-\alpha)x^0 = x^0 + \alpha(y - x^0) = x^0 + \alpha l, \text{ with } l = y - x^0, \ \alpha \in [0,1]. \tag{4.7}$$

Since Y is convex, then $x(\alpha) \in Y$ for any $\alpha \in [0,1]$. Taking into account the latest inclusion and convexity of function $c(x)$, we obtain

$$\begin{aligned} \bar{c}(\alpha) &:= c(x(\alpha)) = c(x^0 + \alpha l) \leq c(x^0) + \alpha(c(y) - c(x^0)), \\ \bar{h}_{js}(\alpha) &:= h_{js}(x(\alpha)) \equiv 0, \ s \in \mathcal{N}(q_j), \ j = 1, \ldots, p^0, \\ \bar{g}_j(\alpha) &:= g_j(x(\alpha)) \leq 0, \ j \in J_A(x^0), \end{aligned} \tag{4.8}$$

where $\alpha \in [0,1]$.

Consider the following Taylor-series expansions of the functions from (4.8) for $\alpha > 0$ sufficiently small:

$$\bar{c}(\alpha) = \bar{c}(x^0) + \alpha \frac{d\bar{c}(0)}{d\alpha} + o(\alpha) = c(x^0) + \alpha l' \nabla c(x^0) + o(\alpha), \tag{4.9}$$

$$\bar{h}_{js}(\alpha) = \bar{h}_{js}(0) + \alpha \frac{d\bar{h}_{js}(0)}{d\alpha} + o(\alpha) = h_{js}(x^0) + \alpha l' \nabla h_{js}(x^0) + o(\alpha), \ s \in \mathcal{N}(q_j), \ j = 1, \ldots, p^0,$$

$$\bar{g}_j(\alpha) = \bar{g}_j(0) + \alpha \frac{d\bar{g}_j(0)}{d\alpha} + o(\alpha) = g_j(x^0) + \alpha l' \nabla g_j(x^0) + o(\alpha), \ j \in J_A(x^0).$$

Then taking into account (4.6), (4.8), and (4.9), we have

$$l' \nabla c(x^0) \leq c(y) - c(x^0) < 0, \tag{4.10}$$

$$l' \nabla h_{js}(x^0) = 0, \ s \in \mathcal{N}(q_j), \ j = 1, \ldots, p^0; \quad l' \nabla g_j(x^0) \leq 0, \ j \in J_A(x^0). \tag{4.11}$$

Multiplying equation (4.5) from the left by l' and considering (4.11), we obtain

$$l' \nabla c(x^0) = - \sum_{j \in J_A(x^0)} \mu_j \, l' \nabla g_j(x^0) \geq 0. \tag{4.12}$$

The last inequality contradicts (4.10), that concludes the proof. ■

4.2.2 The second order sufficient optimality conditions

Given a feasible solution x^0 of problem (4.1) and the corresponding active index set $J_A(x^0)$, consider the *cone of strictly critical directions* in x^0.

$$\begin{aligned} \bar{\mathcal{K}}(x^0) := \{ \xi \in \mathbb{R}^n : \ & \xi' \nabla c(x^0) < 0, \, \xi' \nabla h_{js}(x^0) = 0, \, s \in \mathcal{N}(q_j), j = 1, \ldots, p^0, \\ & \xi' \nabla g_j(x^0) \leq 0, \, j \in J_A(x^0) \}, \end{aligned}$$

and the set

$$\Lambda(x^0) = \{ \lambda : \ \lambda \neq 0, \ \nabla_x \mathcal{L}(x^0, \lambda) = 0, \lambda_0 \geq 0, \mu_j \geq 0, \ \mu_j g_j(x^0) = 0, j \in J(x^0) \}.$$

Theorem 4.2 *Let x^0 be a feasible solution of (4.1). If $\Lambda(x^0) \neq \emptyset$ and*

$$\max_{\lambda \in \Lambda(x^0), ||\lambda||=1} \xi' \nabla_{xx} \mathcal{L}(x^0, \lambda) \xi > 0, \ \forall \xi \in \bar{\mathcal{K}}(x^0), \tag{4.13}$$

then x^0 is a minimizer to the problem $NLP(I_*(x^0))$ in the form (4.1).

Proof (by contradiction). Suppose that the assumptions of the theorem are satisfied for some $x^0 \in Y$ that is not optimal to (4.1). Then there exists $y \in Y$ such that

$$c(y) < c(x^0). \tag{4.14}$$

Since $\lambda_0 \geq 0$ for any $\lambda \in \Lambda(x^0)$, then we can separately consider the following two cases:

I) there exists $\lambda^0 \in \Lambda(x^0)$ such that $\lambda_0^0 > 0$; II) $\lambda_0 = 0$ for all $\lambda \in \Lambda(x^0)$.

Consider, first, case I). Evidently, here the first order sufficient optimality conditions are satisfied for x^0 with $\lambda = \frac{1}{\lambda_0^0} \lambda^0$. Then by Theorem 4.1, x^0 is optimal in problem (4.1) and the contradiction is obtained .

Suppose now that case II) has occurred. Let $x(\alpha)$ be defined by (4.7) and let functions $\bar{h}_{js}(\alpha)$, $\bar{g}_j(\alpha)$ be defined by (4.8). The same reasoning as that used in the proof of Theorem 4.1 leads to relations (4.10) and (4.11), that, obviously, give

$$l \in \bar{\mathcal{K}}(x^0). \tag{4.15}$$

Having expanded the functions $\bar{h}_{js}(\alpha)$ in the Taylor series for a sufficiently small $\alpha > 0$, we obtain

$$\bar{h}_{js}(\alpha) = h_{js}(x^0) + \alpha l' \nabla h_{js}(x^0) + \alpha^2 l' \nabla_{xx} h_{js}(x^0) l + o(\alpha^2) = 0, \ s \in \mathcal{N}(q_j), \ j = 1, \ldots, p^0,$$

wherefrom, taking into account (4.11), we get

$$l' \nabla_{xx} h_{js}(x^0) l = 0, \ s \in \mathcal{N}(q_j), \ j = 1, \ldots, p^0. \tag{4.16}$$

By construction, any $\lambda = (\lambda_0 \geq 0, \ \lambda_{js}, s \in \mathcal{N}(q_j), \ j = 1, \ldots, p^0; \ \mu_j \geq 0, \ j \in J(x^0)) \in \Lambda(x^0)$ satisfies the relations

$$\mu_j g_j(x^0) = 0, \ j \in J(x^0), \tag{4.17}$$

$$\nabla_x \mathcal{L}(x^0, \lambda) = \lambda_0 \nabla c(x^0) + \sum_{j=1}^{p^0} \sum_{s \in \mathcal{N}(q_j)} \lambda_{js} \nabla h_{js}(x^0) + \sum_{j \in J(x^0)} \mu_j \nabla g_j(x^0) = 0. \tag{4.18}$$

Multiplying (4.18) from the left by the vector l', setting $\lambda_0 = 0$, and taking into account (4.11), we obtain

$$\sum_{j \in J(x^0)} \mu_j \, l' \nabla g_j(x^0) = 0. \tag{4.19}$$

Now, let us consider the Taylor-series expansions

$$\bar{g_j}(\alpha) = g_j(x^0) + \alpha l' \nabla g_j(x^0) + \alpha^2 l' \nabla_{xx} g_j(x^0) l + o(\alpha^2),\ j \in J(x^0)$$

for a sufficiently small $\alpha > 0$. Since $\mu_j \geq 0,\ j \in J(x^0)$, then, considering (4.8), (4.17), and (4.19), we have

$$\sum_{j \in J(x^0)} \mu_j \bar{g_j}(\alpha) = \alpha^2 \sum_{j \in J(x^0)} \mu_j\, l' \nabla_{xx} g_j(x^0) l + o(\alpha^2) \leq 0,$$

and, therefore

$$\sum_{j \in J(x^0)} \mu_j\, l' \nabla_{xx} g_j(x^0) l \leq 0, \quad \forall \lambda \in \Lambda(x^0). \tag{4.20}$$

From (4.15), (4.16), and (4.20), taking into account $\lambda_0 = 0,\ \forall \lambda \in \Lambda(x^0)$ we obtain

$$l' \nabla_{xx} \mathcal{L}(x^0, \lambda) l = \sum_{j \in J(x^0)} \mu_j\, l' \nabla_{xx} g_j(x^0) l \leq 0, \quad \forall \lambda \in \Lambda(x^0),$$

that contradicts (4.13), making case II) impossible too.

Thus we have proved that the assumption made in the beginning of the proof is false and x^0 is an optimal solution of (4.1). ■

4.3 Optimality conditions for the problem $NLP(I_*(x^0))$ with analytical constraint function

It was shown in subsection 3.3 that if the function $f(x,t),\ x \in \mathbb{R}^n,\ t \in T \subset \mathbb{R}$, is analytical, then the problem $NLP(I_*(x^0))$ can be written in the equivalent form (3.59) or, in terms of (3.57) and (3.58), as follows:

$$\begin{gathered} c(x) \to \min_{x,\alpha},\ x \in \mathbb{R}^n,\ \alpha \in \mathbb{R}^r, \\ x - \tilde{x} - \Psi\alpha = 0, \\ g_j(x) \leq 0,\ j \in J(x^0), \end{gathered} \tag{4.21}$$

where $\Psi = (\psi_i, i = 1, \ldots, r)$ and $\psi_i \in \mathbb{R}^n,\ i = 1, \ldots, r$, are some vectors satisfying (3.47). The feasible set of problem (4.21) takes the form

$$Y = \{(x, \alpha) :\ x \in \mathbb{R}^n, \alpha \in \mathbb{R}^r,\ x - \tilde{x} - \Psi\alpha = 0,\ g_j(x) \leq 0,\ j \in J(x^0)\}.$$

Evidently, $(\tilde{x}, 0) \in Y$ where $\tilde{x}$ satisfies (3.22), $0 \in \mathbb{R}^r$. From (3.22), taking into account (3.57), we obtain

$$g_j(\tilde{x}) < 0,\ j \in J(x^0),$$

that means that problem (4.21) satisfies the Slater condition. The following theorem is true then.

Theorem 4.3 *[The first order necessary and sufficient optimality conditions] A feasible solution $x^0 \in Y$ is optimal in $NLP(I_*(x^0))$ if and only if there exist multipliers $\mu_j \geq 0,\ j \in J(x^0)$, such that*

$$\mu_j g_j(x^0) = 0, \quad j \in J(x^0),$$
$$\Psi'(\nabla c(x^0) + \sum_{j \in J(x^0)} \mu_j \nabla g_j(x^0)) = 0, \tag{4.22}$$

where functions $g_j(x), j \in J(x^0)$, are defined in (3.57); $\Psi = (\psi_i, i = 1, \ldots, r)$, and $\psi_i, i = 1, \ldots, r$, are some vectors satisfying (3.47) with $y_0 = \tilde{x}$, and vector $\tilde{x}$ satisfies (3.22) .

Note that Theorem 4.3 gives a *criterion* of optimality for the problem $NLP(I_*(x^0))$. To apply Theorem 4.3 we have to construct a matrix Ψ defined by linearly independent vectors $\psi_i \in \mathbb{R}^n,\ i = 1, \ldots, r$, that satisfy relations (3.47). The algorithm of construction of Ψ is not connected with the optimization problem $NLP(I_*(x^0))$, but does with the properties of the set $\bar{Y}$ (see (3.21)) that does not depend on vector x^0.

Example 3.1. Consider the convex SIP problem

$$\begin{array}{c} 3x_1 + x_2 + 3x_3 + x_4 + \frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} + \frac{x_4^2}{2} \longrightarrow \min, \\ \text{s.t.} \ -t^2 x_1 + t x_2 + \sin(t) x_3 + x_4^4 - t^2 - t^3/6 - 5t^4 \leq 0, \quad t \in T = [-1,\ 2], \end{array} \tag{4.23}$$

where $x \in \mathbb{R}^4$.

It is easy to show that this problem has a unique immobile point $t_1^0 = 0$ with $q_1 = q(t_1^0) = 1$, $p^0 = 1$. Since $q_1 > -1$, we can conclude that the Slater condition is not satisfied here.

Consider the feasible solution $x^0 = (-1, 1, -1, 0)' \in X$ to the convex SIP problem (4.23). Evidently for problem (4.23) we have $I_*(x^0) = \emptyset$.

Then the corresponding problem $NLP(I_*(x^0))$ takes the form

$$\begin{array}{c} 3x_1 + x_2 + 3x_3 + x_4 + \frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} + \frac{x_4^2}{2} \longrightarrow \min, \\ \text{s.t.} \quad x_4^4 = 0, \quad x_2 + x_3 = 0, \quad -2x_1 - 2 \leq 0, \end{array} \tag{4.24}$$

and the Lagrange function for it is given by

$$\mathcal{L}(x, \lambda) = \lambda_0 \Big(3x_1 + x_2 + 3x_3 + x_4 + \frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} + \frac{x_4^2}{2} \Big) + \lambda_{10} x_4^4 + \lambda_{11}(x_2 + x_3) + \mu_1(-2x_1 - 2).$$

Let us apply Theorems 4.1, 4.2 and 4.3 to the feasible solution $x^0 = (-1, 1, -1, 0)' \in X \subset Y$ to check if it is optimal in problem (4.24).

The condition $\nabla_x \mathcal{L}(x^0, \lambda) = 0$ can be written as

$$\lambda_0 \begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix} + \lambda_{10} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_{11} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} -2 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0.$$

It is evident that the latest system cannot be satisfied with $\lambda_0 > 0$ and the condition of Theorem 4.1 is not fulfilled.

Next let us verify the conditions of Theorem 4.2. In our example $\nabla_{xx}\mathcal{L}(x^0,\lambda) = \mathbb{O} \in \mathbb{R}^{4\times 4}$ for all $\lambda \in \Lambda(x^0)$. Therefore, even without constructing the cone of strictly critical directions $\bar{\mathcal{K}}(x^0)$ we see that

$$\max_{\lambda\in\Lambda(x^0),\ ||\lambda||=1} \xi'\nabla_{xx}\mathcal{L}(x^0,\lambda)\xi = 0,\ \forall \xi \in \bar{\mathcal{K}}(x^0),$$

and the second order sufficient optimality conditions formulated in Theorem 4.2 are not fulfilled.

Finally, let us apply to x^0 Theorem 4.3. It is easy to check that the vectors $y_0 = (0,1,-1,0)'$, $\psi_1 = (0,1,-1,0)$, $\psi_2 = (1,0,0,0)$ satisfy conditions (3.47). Put

$$\Psi' := (\psi_1,\psi_2)' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix}.$$

Then, evidently,

$$\Psi'(\nabla c(x^0) + \mu_1 \nabla g_1(x^0)) = (0, 2-2\mu_1)',$$

where $c(x) = 3x_1 + x_2 + 3x_3 + x_4 + \frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} + \frac{x_4^2}{2}$, $g_1(x) = -2x_1 - 2$, and conditions (4.22) of Theorem 4.3 are satisfied with $\mu_1 = 1$. Therefore, vector $x^0 = (-1,1,-1,0)' \in X \subset Y$ is optimal to NLP problem (4.24) and hence (see Theorem 2.1) it is optimal to the convex SIP problem (4.23) too.

Remark 4.1 *It was mention above that the problem $NLP(I_*(x^0))$ is equivalent to problem (3.60). For our example, this problem takes the form*

$$3x_1 + x_2 + 3x_3 + x_4 + \frac{x_1^2}{2} + \frac{x_2^2}{2} + \frac{x_3^2}{2} + \frac{x_4^2}{2} \longrightarrow \min,$$
$$s.t. \quad x_4 = 0, \quad x_2 + x_3 = 0, \quad -2x_1 - 2 \le 0.$$

5 Conclusion

In the paper, we study the optimality conditions for the special nonlinear problem $NLP(I_*(x^0))$ constructed for the convex SIP problem (2.1) and connected with it by means of the Implicit Optimality Criterion. Since this Criterion permits to replace the optimality conditions for the convex SIP problem by such the conditions for the problem $NLP(I_*(x^0))$, the optimality conditions for the latest problem are of a special interest. It is shown in the paper that in spite of the fact that the problem $NLP(I_*(x^0))$ is strongly degenerated, it possesses the properties that permit to obtain nontrivial necessary optimality conditions. The sufficient optimality conditions formulated in the paper use the specificity of the problem $NLP(I_*(x^0))$ and allow us to obtain new efficient optimality conditions for convex SIP. It is showed in the paper that in the case when the equality constraints of the problem $NLP(I_*(x^0))$ are given by analytical functions, the optimality conditions take form of criterion.

The results of the paper can be applied for study of the optimality conditions as well as for constructing of numerical methods for the convex SIP. We believe also that some of results can be extended for non convex cases.

References

[1] Arutiunov A.V. *Optimality Conditions: abnormal and degenerate problems*, Kluwer Academic Publishers, 2000.

[2] Avakov E.R. *Necessary conditions for the minimum for smooth abnormal problems with equality and inequality-type constraints.* Mathematicheskie Zametki. V. 45., 1989, pp. 3-11.

[3] Bazaraa M. S., Sherali H.D., and Shetty C.M. *Nonlinear Programming: theory and algorithms*, John Wiley and Sons Inc., 1993.

[4] Brezhneva O.A., Tret'yakov A. *Optimality conditions for degenerate extremum problems with equality constraints*, SIAM J. Control Optim., Vol.42, N. 2, 2003, pp. 729-745.

[5] Kostyukova O.I., *Investigation of the linear extremal problems with continuum of constraints*, Tech.Rep., Preprint 26 (336), Institute of Mathematics, Academy of Sciences of BSSR, Minsk, Belarus, 1988.

[6] Kostyukova O.I., Tchemisova T.V., Yermalinskaya S.A. *Convex Semi-Infinite Programming: Implicit Optimality Criterion based on the concept of immobile points.* CM 05, I-38. Preprint of University of Aveiro, Department of Mathematics, Portugal, 2005.

[7] Kostyukova O.I., Tchemisova T.V. *Convex Semi-Infinite Programming: explicit optimality conditions.* CM 06/I-07, Preprint of University of Aveiro, Department of Mathematics, Portugal, 2006.

[8] Ledzewicz U., Schattler H. *High-order approximations and generalized necessary conditions for optimality.* SIAM J. Control Optim. V. 37, N 1., 1998, pp. 33-53.

[9] Vasil'ev F.P. *Numeric methods of solution of extremal problems* (in Russian), Publishing House "Nauka", Moskow, 1988, pp. 237-239.