Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Nondegenerate Necessary Conditions of Optimality for Problems Without Normality Assumptions

Aram Arutyunov and Fernando Lobo Pereira

***Abstract*— Nondegenerate second-order necessary conditions of optimality for general nonlinear optimization problems are presented and discussed in this article. Besides functional equality and inequality constraints, we also consider constraints in the form of an inclusion into a given closed set. Without assuming a priori normality, our conditions remain informative for abnormal points, and, under very general assumptions also take into account the second order effect of the curvature of the set in the inclusion constraints.**

I. Introduction

The main goal of this article is to present and discuss first- and second-order necessary conditions of optimality, firstly proved in [5], for the following general nonlinear constrained optimization problem:

$$(P_1) \qquad \text{Minimize} \quad f(x) \qquad \text{subject to} \quad \begin{cases} F_1(x) \le 0 \\ F_2(x) = 0 \\ x \in C \end{cases}$$

where X is a vector space, $C \subseteq X$ is a given closed set, $f : X \to \mathbb{R}$, $F_1 : X \to \mathbb{R}^{k_1}$ and $F_2 : X \to \mathbb{R}^{k_2}$ are given smooth mappings, and k_1 and k_2 are also given positive integers. Remark that the inequality above is naturally understood in a componentwise sense.

An important example of an instance of the above class of problems is the following optimal control problem discussed in detail in [3]

$$\begin{aligned} (OPC) \quad & \text{Minimize} && J(x_0, u, \mu) \\ & \text{subject to} && dx(t) = f(x(t), u(t), t)dt \\ & && \qquad +G(t)d\mu(t),\ t \in [t_0, t_1], \\ & && W_1(p) \le 0,\ W_2(p) = 0, \\ & && \mu \in \mathcal{K}, \end{aligned}$$

where $p = (x(t_0), x(t_1))$, $x(t_0) = x_0$, $x(t_1) = x_1$ for given t_0 and t_1 with $t_0 < t_1$ and

$$J(x_0, u, \mu) = W_0(p) + \int_{t_0}^{t_1} f^0(x(t), u(t), t)dt + \int_{[t_0,t_1]} g^0(t)d\mu(t).$$

This work was not supported by any organization

Aram Arutyunov is with Differential Equations and Functional Analysis Dept., Peoples Friendship University of Russia 6, Mikluka-Maklai St., Moscow, 117198, Russia arutun@orc.ru

Fernando Lobo Pereira is with the Instituto de Sistemas e Robótica, Faculdade de Engenharia da Universidade do Porto, R. Dr. Roberto Frias, 4200-465 Porto, Portugal flp@fe.up.pt

The mappings f^0, g^0, f, G, W_i, $i = 0, 1, 2$, are endowed with the required smoothness assumptions.

The cone $\mathcal{K}$ is defined by

$$\mathcal{K} = \Big\{ \mu \in C^*([t_0, t_1]; \mathbb{R}^k) \colon \forall \text{ continuous } \phi \text{ s.t. } \phi(t) \in K^0\ \forall t, \text{ and } \int_B \phi(t)d\mu \ge 0,\ \forall \text{ Borel } B \subset [t_0, t_1] \Big\},$$

K being a given convex, closed, pointed cone from $\mathbb{R}^k$, and K^0 being is its dual.

An admissible control is any pair (u, μ), where $\mu \in \mathcal{K}$ and $u \in L^m_\infty[t_0, t_1]$. A trajectory $x : [t_0, t_1] \to \mathbb{R}^n$, associated with a control policy (u, μ) is a function of bounded variation such that

$$x(t) = x(t_0) + \int_{t_0}^{t} f(x(\tau), u(\tau), \tau)d\tau + \int_{[t_0,t]} G(\tau)d\mu(\tau)$$

for all $t \in (t_0, t_1]$. An admissible control process is a triple (x_0, u, μ), where (u, μ) is an admissible control and the corresponding trajectory satisfies the given endpoint constraints. The process (x_0^*, u^*, μ^*) is a local minimizer of the problem (OPC) if it possesses an appropriate neighborhood such that process (x_0^*, u^*, μ^*) brings the minimum to the problem (OPC).

Additional comments on this problem, the statement of first- and second-order necessary conditions of optimality are, and their proof are presented in [3].

It is not difficult to see that (P_1) can be regarded as a particular instance of the following closely related problem that has been widely investigated in the literature (see, for example, [9], [8], [7] and references therein):

$$(P_2) \qquad \text{Minimize} \quad \bar{f}(\bar{x}) \qquad \text{subject to} \quad \bar{F}(\bar{x}) \in \bar{C}.$$

On the other hand, note that if we define $x := \text{col}(\bar{x}, \tilde{x})$, $\tilde{x} := \bar{F}(\bar{x})$, $F_2(x) := \bar{F}(\bar{x}) - \tilde{x}$, $f(x) := \bar{f}(\bar{x})$, and $C := \bar{C} \times \{0\}$, (P_2) becomes a particular case of (P_1). We will show how the optimality conditions for this problem follow from the ones derived for (P_1).

Given the extremely wide range of classes of optimization problems such as, for example, nonlinear programming, optimal control, calculus of variations, semi-definite and semi-infinite programming, composite parameterized optimization problems, and variational problems in mechanics, covered by these problem paradigms, it is not surprising that a vast amount of literature addressing a large number of research issues such as optimality conditions (necessary, sufficient, first-order, second-order), existence conditions, sensitivity analysis, optimization algorithms, etc., (see, among the references more pertinent to this article, for example, [1], [6], [7], [9], [10], [14], [16], [21], [22], [23]).

The result discussed here follows naturally from previous work of the authors on optimization, [1], [2], and applications optimal control, [3], [4], [5], where inclusion constraints appear naturally.

We improve on the best existing results on necessary conditions of optimality (see [16], [17], [12], [8]) in that our conditions reflect the second order effect of the curvature of the set inclusion constraint under assumptions weaker than those usually considered in the literature. In fact, we require the set C (or $\bar{C}$) to be merely closed, and, hence, we dispense with the usually assumed property of convexity.

Another important feature of our results concern the fact that, in contrast with many results on necessary conditions of optimality (see, for example [6], [16], [11]), our conditions remain informative for abnormal problems and hold without a priori normality assumptions on the data of the problem. This feature is due to the fact that additional information from second-order conditions is used in order to select an appropriate subset of the set of multipliers satisfying the local necessary conditions of optimality.

The proof of the necessary conditions of optimality discussed in this article is presented in detail in [5]. It is based on the removal of the equality and inequality constraints by using a penalty method. Our conditions are initially proved for the case in which X is finite dimensional and, then, they are extended to the ones stated in our result.

This article is organized as follows:

In section II, we present some preliminary definitions required to state our results. The concepts and some pertinent properties of normal cone, first- and second-order tangent cones, as well as that of linear invariant subspace of a set are discussed. In section III, some additional definitions are given and the necessary optimality conditions for (P_1) are presented and discussed. Here, we also include a brief outline of the proof. In the following section, section, problem (P_2) is addressed. Besides presenting the optimality conditions, we also include a discussion on their relation with those of (P_1). Finally, we include in section V, an example illustrating these conditions.

II. Preliminary Definitions

In this section, we introduce some basic general concepts and objects that will be used throughout the next sections.

Given a vector space X, let us consider the functions $f : X \to R$ and $F : X \to Y$. Let $\mathcal{M}$ be the collection of all finite dimensional subspaces of X and denote its finite topology by τ. We say that the set $A \subset X$ is open in the finite topology τ if, for any $M \in \mathcal{M}$, $A \cap M$ is open in the unique Hausdorff vector topology of M. Let f and F be twice continuously differentiable in a neighborhood of a given point $x_0 \in X$ with respect to the finite topology τ. This implies that

$$f(x) = f(x_0) + \langle a, x - x_0\rangle + \frac{1}{2}q[x - x_0]^2 + \alpha_0(x - x_0)$$

$$F(x) = F(x_0) + A(x - x_0) + \frac{1}{2}Q[x - x_0]^2 + \alpha(x - x_0)$$

for some linear functional $a \in X^*$, linear operator $A : X \to Y$, bilinear form $q : X \times X \to \mathbb{R}$, bilinear mapping $Q : X \times X \to Y$, and mappings $\alpha_0 : X \to \mathbb{R}^1$, and $\alpha : X \to Y$, such that, $\forall x \in X$ and, for any $M \in \mathcal{M}$, such that $x \in M$,

$$\frac{\alpha_0(x - x_0)}{\|x - x_0\|_M^2} \to 0, \text{ and } \frac{|\alpha(x - x_0)|}{\|x - x_0\|_M^2} \to 0,$$

as x $\to x_0$.

Here and in what follows, $\|\cdot\|_M$ is a finite-dimensional norm in M and $B[x, x]$ or $B[x]^2$ denote a bilinear mapping B. The mappings a and q are called, respectively, the first- and second-order derivatives of f and, from now on, denoted by $\frac{\partial f}{\partial x}(x_0)$ and $\frac{\partial^2 f}{\partial x^2}(x_0)$. A similar notation is used for the mapping F and other functions.

Next, we introduce the type of normal cone used in our definition of generalized Lagrange multiplier. Assume that the set $C \subset X$ is closed in the finite topology τ of X. The normal cone to C at x in the sense of Mordukhovich, $N_C(x)$, firstly introduced in [18]) for infinite dimensional spaces may also be defined as follows

$$N_C(x) = \bigcup_{M \in \mathcal{M}} N_C^M(x)$$

where $\mathcal{M}$ is, as above, the set of all finite dimensional subspaces of X, and

$$N_C^M(x) = \limsup_{\substack{\bar{x} \in M \\ \bar{x} \to x}} \bigcup_{r>0} \{r[x - W_{M \cap C}(x)]\}$$

with $W_{M\cap C}(x) = \inf_{\xi \in C \cap M} \{\|\xi - x\|\}$.

It is worth to remind some important properties of the Mordukhovich normal cone, (see, for example, [21], [19]). The normal cone $N_C(x)$ is closed, possibly, nonconvex, and $N_C^M(\cdot)$ is upper semicontinuous in $C \cap M$ (here, $M \in \mathcal{M}$). It is well known (see, for example, [19]) that $\frac{\partial f}{\partial x}(x_0) \in -N_C(x_0)$ is a necessary condition for x_0 to be a minimizer of f over C. These properties and the fact that it is the smallest normal cone, makes it the most natural one to derive necessary conditions of optimality.

We will also need the tangent cone

$$T_C(x) = \bigcup_{M \in \mathcal{M}} T_{C\cap M}(x)$$

where $T_{C\cap M}(x)$ is the contingent (Bouligand) cone to the set $C \cap M$ at the point x, given by

$$\{d \in M : \exists\, \varepsilon_n \downarrow 0,\, dist_M(x + \varepsilon_n d, C \cap M) = o(\varepsilon_n)\}.$$

The inner and the outer second order tangent cones to C at x in a direction d are, respectively, given by

$$T^2_C(x,d)=\bigcup_{M\in\mathcal{M}} T^2_{C\cap M}(x,d),\ O^2_C(x,d)=\bigcup_{M\in\mathcal{M}} O^2_{C\cap M}(x,d),$$

where $M\in\mathcal{M}$ is an arbitrary finite-dimensional linear subspace containing x, and d, and $T^2_{C\cap M}(x,d)$ and $O^2_{C\cap M}(x,d)$ are, respectively, given by

$$\Big\{w\in X:\ dist_M(x+\varepsilon d+\frac{1}{2}\varepsilon^2 w, C)=o(\varepsilon^2),\ \varepsilon\geq 0\Big\},$$

and

$$\Big\{w\in X\colon \exists\varepsilon_n\downarrow 0 \text{ s.t. } dist_M(x+\varepsilon_n d+\frac{1}{2}\varepsilon_n^2 w, C)=o(\varepsilon_n^2)\Big\}.$$

Obviously, $T^2_C(x,d)\subset O^2_C(x,d)$. Furthermore, from [7], it is asserted that both $T^2_{C\cap M}(x,d)$ and $O^2_{C\cap M}(x,d)$ and that $O^2_C(x,d)\neq\emptyset$ only if $d\in T_C(x)$. Furthermore, while $O^2_C(x,d)$ may fail to be convex, $T^2_C(x,d)$ is always convex whenever C is a convex set.

Now, following [5], we introduce the concept of *invariant linear subspace* (*ILS*) relatively to a closed set C at x. A linear subspace is *ILS* relatively to C at x, denoted by $\mathcal{I}_C(x)$, if $x+\mathcal{I}_C(x)\subseteq C$, $\forall x\in C$. If this property holds for all $x\in C$, then it is said to be *ILS* with respect to C.

Since, for a given closed set C, any linear subspace of an *ILS* also is an invariant linear subspace an *ILS* is not, in general, unique. For any $x\in C$, put $\mathcal{I}_C(x)=\cap_{r\neq 0} r[C-x]$, being the intersection taken over all $r\in\mathbb{R}$, $r\neq 0$.

It is proved in [5] that if the set C is convex, then, $\forall x\in C$, $\mathcal{I}_C(x)$ is the maximal *ILS* relatively C and, therefore, it does not depend on x.

Note also that, if the C is a convex cone, then $\mathcal{I}_C=C\cap(-C)$ is the maximal *ILS*.

III. Optimality Conditions for (P_1)

Let us consider the problem (P_1).

Let $\lambda=(\lambda_0,\lambda_1,\lambda_2)$ with $\lambda_0\in\mathbb{R}$, $\lambda_1\in\mathbb{R}^{k_1}$, and $\lambda_2\in\mathbb{R}^{k_2}$, and define the generalized Lagrangian by

$$\mathcal{L}(x,\lambda)=\lambda_0 f(x)+\langle\lambda_1,F_1(x)\rangle+\langle\lambda_2,F_2(x)\rangle.$$

Let $\Lambda=\Lambda(x_0)$ denote the set of the generalized Lagrange multipliers λ that correspond to the point x_0 according to the Lagrange multiplier rule (see [19], [20], [7]), i.e., satisfying

$$\begin{cases} \dfrac{\partial\mathcal{L}}{\partial x}(x_0,\lambda)\in -N_C(x_0) \\ \lambda_0\geq 0, \lambda_1\geq 0,\ \langle\lambda_1,F_1(x_0)\rangle=0,\ |\lambda|=1. \end{cases} \tag{1}$$

In what follows, let $F_{1,j}$ denote the $j^t h$ coordinate of the function F_1.

Let the *critical cone* of the problem (P) at the point x_0, $\mathcal{K}(x_0)$, to be defined by

$$\Big\{h\in T_C(x_0):\langle\frac{\partial f}{\partial x}(x_0),h\rangle\leq 0,\ \frac{\partial F_2}{\partial x}(x_0)h=0, \text{and}\ \frac{\partial F_{1,j}}{\partial x}(x_0)h\leq 0\ \forall j \text{ s.t. } F_{1,j}(x_0)=0\Big\}.$$

Notice that $\mathcal{K}(x_0)$ is convex if the set C is convex and that, since it contains zero, is always nonempty.

From now on, we assume without loss of generality that $F_1(x_0)=0$. This can always be achieved by omitting the nonactive components of the inequality constraints.

Take any linear subspace $M\subseteq X$ and denote by $\Lambda(x_0,M)$ the set of all Lagrange multipliers $\lambda\in\Lambda$ for which there exists a linear subspace $\Pi\subseteq M$ (possibly depending on λ) such that

$$\begin{aligned} \mathrm{codim}_M\Pi &\leq k, \\ \Pi &\subseteq \mathrm{Ker}\,\frac{\partial F}{\partial x}(x_0), \\ \frac{\partial^2\mathcal{L}}{\partial x^2}(x_0,\lambda)[x,x] &\geq 0,\ \forall x\in\Pi, \end{aligned}$$

where, and from now on, codim_M and $\mathrm{Ker}\,A$ denote, respectively, the codimension relative to the subspace M and the kernel of the linear operator $A: X\to Y$. Each set $\Lambda(x_0,M)$ is obviously compact (but it may be empty).

For a given set $C\subseteq X$, we denote by $\sigma(\cdot,C)$ its support function, i.e., for $x^*\in X^*$,

$$\sigma(x^*,C)=\sup_{x\in C}\langle x^*,x\rangle.$$

Let us now state our necessary conditions of optimality for problem (P_1).

Theorem 1: Consider the problem (P_1) with the function f and mappings F_1 and F_2 are twice continuously differentiable and the set $C\subset X$ is closed.

Let x_0 be a point of local minimum with respect to the finite topology τ of X.

Then, for each ILS $\mathcal{I}$ with respect to C, the set $\Lambda(x_0,\mathcal{I})$ is nonempty, and, moreover, for each $h\in\mathcal{K}(x_0)$ and any convex set $\mathcal{T}(h)\subseteq O^2_C(x_0,h)$,

$$\max_{\lambda\in\Lambda_a}\Bigg\{\frac{\partial^2\mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h] -\sigma\Big(-\frac{\partial\mathcal{L}}{\partial x}(x_0,\lambda),\mathcal{T}(h)\Big)\Bigg\}\geq 0. \tag{2}$$

Here, $\Lambda_a=\mathrm{conv}\Lambda(x_0,\mathcal{I})$ and conv denotes the convex hull of a set.

Note that this theorem was derived in [1] for the case $C=X$, and in [2], [3] for the case in which C is a convex cone and $h\in C+span\{x_0\}$. Also, if C is a convex cone, then the tangent cone $T_C(x)$ coincides with closure of the set $C+span\{x\}$ for each $x\in C$.

Additionally, if $h\in\mathcal{I}_C(x_0)$, an ILS relatively to C, then

$$\max_{\lambda\in\Lambda_a}\Big\{\frac{\partial^2\mathcal{L}}{\partial x^2}(x_0,\lambda)[h,h]\Big\}\geq 0. \tag{3}$$

To see this, just note that $0\in O^2_C(x_0,h)$ and the conclusion is obtained by taking $\mathcal{T}(h)=\{0\}$ in (2).

Now, let us see that if C is a convex set, then

$$\sigma\Big(-\frac{\partial\mathcal{L}}{\partial x}(x_0,\lambda),\mathcal{T}(h)\Big)\leq 0\quad\forall\,\lambda\in\Lambda. \tag{4}$$

First, notice that it follows from $h \in \mathcal{K}(x_0)$ that $\langle \frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda), h \rangle \leq 0$ holds.

Also, from the convexity of the set C, we have that $-N(x, C) = (T_C(x))^*$, $\forall x \in C$ (K^* denotes the positive dual cone to a cone $K \subseteq X$), and hence, according to (1), we obtain

$$\frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda) \in (T_C(x_0))^*.$$

Since, for each $w \in O_C^2(x_0, h)$, we have

$$\varepsilon_n h + \frac{1}{2}\varepsilon_n^2 w + o(\varepsilon_n^2) \in C - x_0 \subseteq T_C(x_0)$$

(being the last inclusion due to the convexity of C) and, by using the first inequality obtained in this proof, we obtain, $\varepsilon_n^2 \langle \frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda), w \rangle + o(\varepsilon_n^2) \geq 0$. This and the arbitrariness of w proves (4).

We also proved that, if the set C is convex, then condition (3) becomes stronger then (2).

In the case C is a convex cone, if either $h \in C + span\{x_0\}$, or the cone C has a finite number of faces, then $\sigma(x^*, T_C^2(x_0, h)) = 0$, $\forall x^* \in N_C(x_0)$.

In [8], [11], [17] many examples can be found for which the additional term σ does not disappear. The following simple example shows that, if the set C is not convex, then the term $\sigma\left(-\frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda), \mathcal{T}(h)\right)$ may become strictly positive.

Let $C = \{x = (x_1, x_2) \in \mathbb{R}^2 : x_1^l \geq x_2^m\}$ where l and m are given positive integers such that $l \leq 2m$. If $h = (1, 0)$, then it is obvious that the outer second-order tangent set $O_C^2(0, h)$ contains a ball δB, $B = \{w \in \mathbb{R}^2 : \|w\| \leq 1\}$, for some $\delta > 0$. Obviously, $\sigma(\zeta, \delta B) > 0$, $\forall \zeta \neq 0$. Obviously, for the set C and the vector h constructed above, condition (2) is stronger than condition (3) if $\frac{\partial \mathcal{L}}{\partial x}(x_0, \lambda) \neq 0$.

Let us outline the proof of theorem 1. We start by considering X to be finite dimensional and show, in a first stage, that the set of multipliers Λ_a is nonempty. This involves the formulation of sequence of optimization problems originated from (P_1) by removing the equality and the inequality constraints by penalization, that is, only the set inclusion constraint is present. Then, after the first- and second-order necessary conditions of optimality are proved and expressed in terms of the data of the original problem, the limit is taken to reveal the desired conclusion.

The ensuing stage consists in proving, still in the finite dimensional context, a second-order variational inequality. This is achieved by constructing a sequence of auxiliary optimization problems with reference to a sequence of points converging to the solution of the original problem and depending explicitly on two directions, one in the critical cone and another in the second-order contingent cone. After showing that the corresponding sequence of solutions converges to the one of the original problem, the result proved in stage one is applied, being the desired conclusion obtained in the limit.

The conclusion of stage two is derived for infinite dimensional X in the third stage. Here, a key role is played by the finite topology τ introduced earlier, together with the fact that the family of sets of appropriate multipliers obtained by considering $X \cap M$, indexed by the finite dimensional subspace M, is a centered system of sets.

The proof is concluded by showing that the order of the *inf* and *max* operations can be changed.

IV. Optimality Conditions for (P_2)

Given the fact that many important and interesting results are stated for problems formulated as (P_2), we consider here the optimality conditions for this class of problems. Let $x_0 \in X$ be the solution to the problem:

$$\begin{array}{lll} (P_2) & \text{Minimize} & f(x) \\ & \text{subject to} & F(x) \in C, \end{array}$$

where C is a given closed set in $Y = \mathbb{R}^k$, and f and F are twice continuously differentiable with respect to the finite topology τ in a neighborhood of x_0.

As in the statement of conditions for problem (P_1), we require the following definitions.

Let $\mathcal{L}^{P_2}(x, \lambda) = \lambda_0 f(x) + \langle \bar{\lambda}, F(x) \rangle$ with $\lambda = (\lambda_0, \bar{\lambda})$, $\lambda_0 \in \mathbb{R}^1$, and $\bar{\lambda} \in Y^*$.

Denote by $\Lambda^{P_2} = \Lambda^{P_2}(x_0)$ the set of the Lagrange multipliers λ associated with problem (P_2) corresponding to the point x_0 according to the Lagrange multiplier rule:

$$\frac{\partial \mathcal{L}^{P_2}}{\partial x}(x_0, \lambda) = 0,\ \lambda_0 \geq 0,\ \bar{\lambda} \in N_C(F(x_0)),\ |\lambda| = 1.$$

Take any linear subspace $M \subseteq Y$ and consider the set of all Lagrange multipliers $\lambda \in \Lambda^{P_2}$ for which there exists a linear subspace $\Pi \subseteq X$ (depending on λ) such that

$$\begin{aligned} \text{codim}\Pi &\leq k, \\ \Pi &\subseteq \left(\frac{\partial F}{\partial x}(x_0)\right)^{-1}(M), \\ \frac{\partial^2 \mathcal{L}^{P_2}}{\partial x^2}(x_0, \lambda)[h, h] &\geq 0, \quad \forall h \in \Pi \end{aligned}$$

(Here, $\text{codim} = \text{codim}_X$.) We denote this set of Lagrange multipliers by $\Lambda^{P_2}(x_0, M)$.

Theorem 2: Let x_0 be a point of local minimum with respect to the finite topology τ of the problem (P_2).

Then, for each ILS $\mathcal{I}$ with respect to C, the set $\Lambda^{P_2}(x_0, \mathcal{I})$ is nonempty, and, moreover, for each $h \in \mathcal{K}^{P_2}$, defined by

$$\left\{h :\ \frac{\partial F}{\partial x}(x_0)h \in T_C(F(x_0)),\ \langle \frac{\partial f}{\partial x}(x_0), h \rangle \leq 0\right\},$$

and, for each convex set $\mathcal{T}(h) \subseteq O_C^2(F(x_0), \frac{\partial F}{\partial x}(x_0)h)$, the following condition holds

$$\max_{\lambda \in \Lambda_a^{P_2}} \left(\frac{\partial^2 \mathcal{L}^{P_2}}{\partial x^2}(x_0, \lambda)[h, h] - \sigma\left(\bar{\lambda}, \mathcal{T}(h)\right)\right) \geq 0. \quad (5)$$

Here, $\Lambda_a^{P_2} = \mathrm{conv}\Lambda^{P_2}(x_0, \mathcal{I})$.

Proof: As it was observed in the introduction, this problem can be regarded as particular instance of the problem (P_1).

By applying theorem 1 to this particular instance of problem (P_1), we obtain the Lagrangian

$$\mathcal{L}(x, y, \lambda) = \lambda_0 f(x) + \langle \bar{\lambda}, F(x) - y \rangle$$

and conditions (1) become

$$\frac{\partial \mathcal{L}}{\partial y}(x_0, F(x_0), \lambda) = -\bar{\lambda} \in -N_C(F(x_0)).$$

Now, by using the later inclusion, formulas (1) and (2) and definition of the set $\Lambda_a^{P_2}$ we obtain the desired conclusion. The theorem is proved. ■

There is a significant set of publications, notably [7], [8], addressing (P_2). By assuming the convexity of the set C and Robinson's constraint qualification for arbitrary Banach space Y, the following optimality condition was derived:

$$\max_{\lambda \in \Lambda^{P_2}} \left(\frac{\partial^2 \mathcal{L}^{P_2}}{\partial x^2}(x_0, \lambda)[h, h] - \sigma\left(\bar{\lambda}, \mathcal{T}(h)\right) \right) \geq 0 \; \forall h \in \mathcal{K}^{P_2}. \tag{6}$$

Clearly, theorem 2 shows that, for finite dimensional Y, the result of [8] holds for generalized Lagrangian without Robinson's constraint qualification and without the convexity assumption on C.

Moreover, since, in general, $\Lambda_a^{P_2} \subseteq conv\Lambda^{P_2}$, (5) is stronger than (6).

At the same time, under Robinson's constraint qualification, we can guarantee in theorem 2 that $\lambda_0 > 0$, $\forall \lambda \in \Lambda^{P_2}$.

Finally, note that if C is a closed and pointed convex cone, that is $C \cap [-C] = \{0\}$, then the maximal ILS $\mathcal{I}_C$ is equal to $\{0\}$ and the inclusion above becomes an equality.

The necessary conditions of optimality (6) for the problem without Robinson's constraint qualification and for which the convex set C has nonempty interior was obtained in [8]. Another type of necessary conditions of optimality was obtained in [14], [15] for the problem without Robinson's constraint qualification and for which C has nonempty interior and has a finite number of faces.

V. Example

Let us consider the following semi-definite programming problem

$$\begin{array}{lll} (P) & \text{Minimize} & f(x) \\ & \text{subject to} & F_1(x) \in \mathcal{S}_-^p \\ & & F_2(x) = 0, \end{array}$$

being

- $X = \mathbb{R}^n$, $Y = \mathcal{S}^p \times \mathbb{R}^{k_2}$, where $\mathcal{S}^p$ is the space of $p \times p$ symmetric matrices,
- n, p, k_2 given integers,
- $C = \mathcal{S}_-^p \times \{0\}$, where $\mathcal{S}_-^p$ the cone of negative semi-definite matrices,
- $f : X \to \mathbb{R}$ a given smooth function which is assumed (without loss of generality) to satisfy $\dfrac{\partial f}{\partial x}(0) \neq 0$,
- $F_1 : X \to \mathcal{S}^p$ defined by $F_1(x) := \sum_{i,j} x_i x_j S_{i,j} + \psi_1(x)$, where x_i are coordinates of vector x, and $S_{i,j}$ are given symmetric matrices,
- $F_2 : X \to \mathbb{R}^{k_2}$ defined by $F_2(x) = \dfrac{1}{2}Q(x) + \psi_2(x)$ where the bilinear mapping $Q : X \times X \to \mathbb{R}$ is defined by $Q(x) = (\langle Q_1 x, x \rangle, ..., \langle Q_{k_2} x, x \rangle)$ where Q_i are given symmetric matrices, being
  ψ_l, $l = 1, 2$, given smooth mappings such that $\psi_l(0)=0$, $\dfrac{\partial \psi_l}{\partial x}(0) = 0$, $\dfrac{\partial^2 \psi_l}{\partial x^2}(0) = 0$, $l = 1, 2$.

Let $k := p(p+1)/2 + k_2$, $F = (F_1, F_2)$ and equip the space $\mathcal{S}^p$ with scalar product $A \bullet B = trace(AB)$. Then, the dual cone to $\mathcal{S}_-^p$ is the cone $\mathcal{S}_+^p$ of all positive semi-definite matrices. (See details in [8].)

Consider the point $x = 0$. It can be readily seen that this is an abnormal point and, hence, because $\dfrac{\partial F}{\partial x}(0) = 0$, Robinson's constraint qualification is not satisfied. Therefore, known necessary conditions, [8], cannot be applied.

On the other hand, by applying theorem 2, we obtain the following necessary conditions of optimality:

If $x = 0$ is a local minimum for (P), then $\lambda_0 = 0$ (again, from the fact that $\dfrac{\partial F}{\partial x}(0) = 0$) and, hence, $\forall\, h \in \mathbb{R}^n$, $\exists \lambda_1 \in \mathcal{S}_+^p$, $\lambda_2 \in \mathbb{R}^{k_2}$, such that

$$\left\{ \begin{array}{c} \mathrm{ind}\left(\sum_{i,j} h_i h_j S_{i,j} \bullet \lambda_1 + \langle Q(h), \lambda_2 \rangle \right) \leq k, \\ (\lambda_1, \lambda_2) \neq 0, \\ \sum_{i,j} h_i h_j S_{i,j} \bullet \lambda_1 + \langle Q(h), \lambda_2 \rangle \geq 0. \end{array} \right. \tag{7}$$

where, $\mathrm{ind}\, q$ denotes the index of the quadratic form q on a given space V which can be defined as the dimension of a subspace of V of maximum dimension where the quadratic form q is negative definite.

Notice that we used the fact that if $h \notin \mathcal{K}^P$, then $-h \in \mathcal{K}^P$ and that $\sigma(\lambda_1, T^2_{\mathcal{S}_-^p}(0, 0)) = 0$ for each $\lambda_1 \in \mathcal{S}_+^p$.

Therefore, we conclude that if there exists $h \in \mathbb{R}^n$ such that

$$\sum_{i,j} h_i h_j S_{i,j} \bullet \lambda_1 + \langle Q(h), \lambda_2 \rangle < 0$$

for any (λ_1, λ_2) which satisfies (7) (in particular, if $\mathrm{ind}\left(\sum_{i,j} h_i h_j S_{i,j} \bullet \lambda_1 + \langle Q(h), \lambda_2 \rangle \right) > k$, $\forall (\lambda_1, \lambda_2) \neq 0$ such that $\lambda_1 \in \mathcal{S}_+^p$), then $x = 0$ cannot be a local minimum of the problem.

VI. Acknowledgments

The first author acknowledges the support of Russian Foundation for Basic Researches, grant N 05-01-00193; the second author acknowledges the support of Fundação para a Ciência e Tecnologia and Fundação Calouste Gulbenkian.

References

[1] A. V. Arutyunov, *Optimality Conditions: Abnormal and Degenerate Problems*, Kluwer Academic Publishers, 2000.
[2] A. V. Arutyunov, Necessary Extremum Conditions and a Inverse Function Theorem without a priori Normality Assumptions, *Proceedings of the Steklov Institute of Mathematics*, vol. 236, 2002, pp. 25–36.
[3] A. V. Arutyunov, V. Jacimovic, F. Pereira, Second Order Necessary Conditions for Optimal Impulsive Control Problems, *Journal of Dynamical and Control Systems*, vol. 9, 2003, pp. 131–153.
[4] A. V. Arutyunov, V. Dykhta, F. Pereira, Necessary Conditions for Impulsive Nonlinear Optimal Control Problems without a priori Normality Assumptions, *Journal of Optimization Theory and Applications*, vol. 124, 2005, pp. 55–77.
[5] A. V. Arutyunov, F. Pereira, Second-order necessary optimality conditions for problems without a priori normality assumptions, *Mathematics Operations Research*, to appear, 2005.
[6] A. Ben-Tal, J. Zowe, A unified theory of first and second order conditions for extremum problems in topological vector spaces, *Math. Programming Study*, vol 19, 1982, pp. 39–76.
[7] J. F. Bonnans, R. Cominetti, A. Shapiro, Second order optimality conditions based on parabolic second order tangent sets, *SIAM J. Optimization*, vol. 9, 1999, pp. 466-492.
[8] J. F. Bonnans, A. Shapiro, *Perturbation Analysis of Optimization Problems*, Series in Operations Research, Springer-Verlag New-York Inc, 2000.
[9] J. F. Bonnans, A. Shapiro, Optimization problems with perturbations: a guided tour, *SIAM Review*, vol. 40, 1998, pp. 228–264.
[10] F. H. Clarke, *Optimization and nonsmooth analysis*, A Wiley–Interscience Publication, John Wiley Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1983.
[11] R. Cominetti, Metric regularity, tangent sets, and second-order optimality conditions, *Applied Mathematics and Optimization*, vol 21, 1990, pp. 265–287.
[12] R. Cominetti, J. Penot, Tangent sets of order one and two to positive cones of some functional spaces, *Applied Mathematics and Optimization*, vol. 36, 1997, pp. 291–312.
[13] I. Ekeland, R. Temam, *Convex analysis and variational problems*, North-Holland Publishing company, Amsterdam, 1976.
[14] A. F. Izmailov, Optimality conditions in extremal problems with nonregular inequality constraints, *Mathematical Notes*, vol. 66, 1999, pp. 72–81.
[15] A. F. Izmailov, M. V. Solodov, Optimality conditions for irregular inequality-constrained problems, *SIAM J. on Control and Optimization*, vol. 42, 2002, pp. 1080–1295.
[16] H. Kawasaki, An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems, *Mathematical Programming*, vol. 41, 1988, pp. 73–96.
[17] H. Kawasaki, Second-order necessary optimality conditions for minimizing a sup-type function, *Mathematical Programming*, vol. 49, 1991, pp. 213–229.
[18] A. Y. Kruger, B. S. Mordukhovich, Extremal points and the Euler-Lgrange equation in nonsmooth optimization, *Dokl. Akad. Nauk BSSR*, vol. 24, 1980, pp. 684–687.
[19] B. S. Mordukhovich, Complete characterezation of openness, metric regularity, and Lipshitzian properties of multifunctions, *Transactions of the American Mathematical Society*, vol. 340, 1993, pp. 1–36.
[20] B. S. Mordukhovich, Maximum principle in problems of time optimal control with nonsmooth constraints, *Appl. Math. Mech.*, vol. 40, 1976, pp. 960–969.
[21] R. Rockafellar, R. Wets, *Variatitional Analysis*, Springer-Verlag, 1998.
[22] A. Shapiro, First- and second-order analysis of nonlinear semidefinite programs, *Mathematical Programming*, vol. 77, 1997, pp. 301–320.
[23] R. B. Vinter, *Optimal Control*, Birkhäuser, Boston, 2000.