Jaroslav Doležal

A note on necessary conditions in mathematical programming

*Kybernetika*, Vol. 17 (1981), No. 6, 555--560

Persistent URL: `http://dml.cz/dmlcz/125636`

KYBERNETIKA — VOLUME 17 (1981), NUMBER 6

# A NOTE ON NECESSARY CONDITIONS IN MATHEMATICAL PROGRAMMING

JAROSLAV DOLEŽAL

A problem of necessary optimality conditions in general mathematical programming problems is investigated. It is shown that a unifying approach of Boltjanskij to the solution of this class of optimization problems remains valid when some other kind of the so-called first-order conical approximation to a set is used.

## 1. INTRODUCTION

Necessary optimality conditions for a general class of mathematical programming problems were studied by a number of authors in the past. In [1] the abstract theory developed by Neustadt [2] was applied to obtain necessary conditions also for problems with implicit set constraints. To do this, a concept of the so-called conical approximation to a given set was introduced also originating from the general ideas of [2]. Later, similar concept was used by Boltjanskij [3, 4] to deal with mathematical programming problem having variety of possible constraints.

In [1] the equality type constraints were alternatively treated separately as suggested by the original scheme given in [2]. Such approach enables to overcome certain difficulties resulting from equality type constraints and offers an additonal possibility to refine the obtained conditions as the "constraints qualification" is concerned. On the other hand, the unifying approach developed in [3] can be applied practically to any type of a mathematical programming problem not a priori distinguishing equality type and other constraints.

In this note it is briefly demonstrated that also somewhat more general concept of a first-order conical approximation can be alternatively used within the context of [3, 4]. The corresponding proof of the respective basic theorem of [3, 4] is indicated in brief. Its idea is due to Miljutin — see [4, footnote on p. 16]. In this way the desired basic result is obtained in a straight-forward manner on applying general separation theorem of [2]. Some applications of this result to mathematical programming problems are also reviewed.

## 2. PRELIMINARY RESULTS

Let us mention some basic definitions and results used in the sequel. For a set $\Omega \subset R^n$ let us denote $\bar{\Omega}$ and co $\Omega$ its closure in $R^n$ and its conical hull, respectively. If $C \subset R^n$ is a convex cone with vertex in $\hat{x}$, then $C' = \{a \in R^n \mid \langle x - \hat{x}, a\rangle \leqq 0, x \in C\}$ represents the polar (dual) cone to $C$ in $R^n$. As usual, $\langle \cdot, \cdot \rangle$ denotes the scalar product in $R^n$.

**Definition 1.** Let $C_1, \ldots, C_s$ be a family of convex cones in $R^n$ having common vertex $\hat{x}$. We say that this family exhibits a separation property in $R^n$ if there is one cone which can be separated (in the classical sense) from the intersection of the remaining ones.

In this form the defined concept is due to Boltjanskij [3, 4], where also the following alternative characterization is given.

**Proposition 1.** The family of convex cones $C_1, \ldots, C_s$ exhibits a separation property iff there exist vectors $a_i \in C_i'$, $i = 1, \ldots, s$, with at least one of them being nonzero, and such that $a_1 + \ldots + a_s = 0$.

**Definition 2.** Let $\Omega \subset R^n$. A convex cone $C(\hat{x}, \Omega)$ with vertex at $\hat{x}$ will be called a first-order conical approximation to the set $\Omega$ at $\hat{x} \in \Omega$ if for any finite collection $x_1, \ldots, x_k$ of vectors in $C(\hat{x}, \Omega)$ and arbitrary neighbourhood $U$ of the origin there exists a positive number $\varepsilon_0$ such that for every $\varepsilon$, $0 < \varepsilon < \varepsilon_0$, there is a continuous map $\zeta_\varepsilon$ from co $\{x_1, \ldots, x_k\}$ into $\Omega$ such that $\zeta_\varepsilon(x) \in x + \varepsilon(x - \hat{x}) + \varepsilon U$ for all $x \in$ co $\{x_1, \ldots, x_k\}$.

Slightly different definition of the so-called conical approximation of the second kind was used in [1] or [5]. In a more general setting the above concept was introduced by Neustadt [2] assuming only "convex set" approximation in order to deal with abstract optimization problems. Also the well-known results of Dubovickij and Miljutin [6] are closely related to this subject. For a sake of comparison let us finally recall the alternative concept of a "mantle" used in similar meaning by Boltjanskij [3, 4].

**Definition 3.** Let $x \in \Omega \subset R^n$ and let $K$ be a convex cone with vertex in $\hat{x}$. The cone $K$ is called a mantle to the set $\Omega$ at $\hat{x}$ if there is a neighbourhood $U(\hat{x})$ of $\hat{x}$ and a continuous map $\psi$ from $K \cap U(\hat{x})$ into $\Omega$, such that $\psi(x) = x + o(x - \hat{x})$, where $o(x) = 0$ for $x = 0$ and $\lim_{\|x\| \to 0} \|o(x)\| / \|x\| = 0$.

One can show that $K$ being a mantle to $\Omega$ at $\hat{x}$ is in the same time also a first-order conical approximation $C(\hat{x}, \Omega)$. Especially in our finite dimensional setting the both concepts are practically equivalent[1]). Alternatively, in [4] the additional requirement of $\psi$ being continuously differentiable is imposed. Also then the validity of most basic results is still maintained.

## 3. BASIC THEOREM

Now let us formulate the fundamental result which will show that the mentioned general scheme of Boltjanskij [3, 4] is applicable also in the case of convex approximations given according to Definition 2. Further details are to be found in the indicated references.

**Theorem 1.** Let $\Omega_0, \Omega_1, \ldots, \Omega_s$ be a family of sets in $R^n$ the intersection of which consists of a single point $\hat{x}$, and let $C(\hat{x}, \Omega_i)$, $i = 0, 1, \ldots, s$, be the corresponding first-order conical approximations to these sets at $\hat{x}$. Assume that at least one of these cones is not a hyperplane (of any dimension). Then there exist vectors $a_i \in \in C'(\hat{x}, \Omega_i)$, $i = 0, 1, \ldots, s$, not all zero, and such that $a_0 + a_1 + \ldots + a_s = 0$.

Alternative conclusion due to Definition 1 is that the family of cones $C(\hat{x}, \Omega_i)$, $i = 0, 1, \ldots, s$, exhibits a separation property. For the case of a "mantle" the somewhat lengthy proof of this result is given in [3]. On the other hand, if the mentioned continuous differentiability of $\psi$ in Definition 3 holds, the corresponding proof is considerably simpler [4]. Here the idea of Miljutin will be used, as discussed earlier, together with the general result of [2].

Proof of Theorem 1. Consider a product-space $\boldsymbol{R} = R^{(s+1)n} = R^n \times R^n \times \ldots \ldots \times R^n$, i.e. $(s + 1)$-times the original space $R^n$. In this space define the sets

$$K = C(\hat{x}, \Omega_0) \times C(\hat{x}, \Omega_1) \times \ldots \times C(\hat{x}, \Omega_s), \tag{1}$$

and

$$\boldsymbol{\Omega} = \Omega_0 \times \Omega_1 \times \ldots \times \Omega_s. \tag{2}$$

It is not very difficult to see that $\boldsymbol{K}$ is the first-order conical approximation to the set $\boldsymbol{\Omega}$ at the point $\hat{\boldsymbol{x}} = (\hat{x}, \hat{x}, \ldots, \hat{x})$, $\hat{\boldsymbol{x}} \in \boldsymbol{R}$, i.e. $\boldsymbol{K} = C(\hat{\boldsymbol{x}}, \boldsymbol{\Omega})$, which is not a plane. Further let

$$\Delta = \{\boldsymbol{x} \in \boldsymbol{R} \mid \boldsymbol{x} = (x, x, \ldots, x),\ x \in R^n\}. \tag{3}$$

The set $\Delta$ is a proper linear subspace of $\boldsymbol{R}$, as can be directly verified. This in turn implies that $C(\boldsymbol{x}, \Delta) = \Delta$ for any $\boldsymbol{x} \in \Delta$.

Moreover, the sets $\boldsymbol{\Omega}$ and $\Delta$ have only the point $\hat{\boldsymbol{x}}$ in common. As an easy con-

[1]) Note added in proof. An exhaustive classification of various conical approximations was given recently by D. H. Martin, R. J. Gardner, G. G. Watkins: Indicating cones and the intersection principle for tangential approximants in abstract multiplier rules, J. Optim. Theory Appl. *33* (1981), 4, 515—537, where also a similar construction was presented.

sequence of [2, Theorem 2.2] one obtains that then there is a nonzero vector $\boldsymbol{a} \in \boldsymbol{R}$, $\boldsymbol{a} = (a_0, a_1, \ldots, a_s)$, $a_i \in R^n$, $i = 0, 1, \ldots, s$, such that

$$(4) \qquad \langle \boldsymbol{x} - \hat{\boldsymbol{x}}, \boldsymbol{a} \rangle \leqq 0 \leqq \langle \boldsymbol{y} - \hat{\boldsymbol{x}}, \boldsymbol{a} \rangle \quad \text{for all} \quad \boldsymbol{x} \in \boldsymbol{K}, \quad \boldsymbol{y} \in \boldsymbol{\Delta} .$$

As $\boldsymbol{x} = (x_0, x_1, \ldots, x_s) \in \boldsymbol{K}$ implies $x_i \in C(\hat{x}, \Omega_i)$, $x_i \in R^n$, $i = 0, 1, \ldots, s$, for the left-hand inequality in (4) can be written

$$(5) \qquad \sum_{i=0}^{s} \langle x_i - \hat{x}, a_i \rangle \leqq 0 \quad \text{for all} \quad x_i \in C(\hat{x}, \Omega_i), \quad i = 0, 1, \ldots, s .$$

Then necessarily $a_i \in C'(\hat{x}, \Omega_i)$, $i = 0, 1, \ldots, s$.

Owing to (3) the right-hand inequality in (4) simply says that

$$(6) \qquad 0 \leqq \langle x - \hat{x}, a_0 + a_1 + \ldots + a_s \rangle \quad \text{for all} \quad x \in R^n ,$$

and thus $a_0 + a_1 + \ldots + a_s = 0$. This completes the proof of the theorem. □

## 4. MATHEMATICAL PROGRAMMING PROBLEMS

Now it is easy to use the methodology of [3, 4] to formulate similar theorems for a variety of mathematical programming problems. In general, it will be assumed that the aim is to minimize a function $f(x)$, where

$$(7) \qquad f : R^n \to R^1 ,$$

subject to the constraints

$$(8) \qquad x \in \Omega = \Omega_1 \cap \ldots \cap \Omega_s ,$$

with $\Omega_i \subset R^n$, $i = 1, \ldots, s$.

It will be shown that the respective necessary optimality conditions for this problem are obtained in a straightforward way on applying Theorem 1. However, let us first recall some important special cases of first-order conical approximations. These results can be easily established realizing the related theory described in [1]. The obvious analogy with [3, 4] is also helpful in this respect. By a subscript let us denote the respective gradient.

**Proposition 1.** Let $h : R^n \to R^1$ be continuously differentiable and let $h_x(\hat{x}) \neq 0$. Then the set $\{x \in R^n \mid \langle x - \hat{x}, h_x(\hat{x}) \rangle = 0\}$ is a first-order conical approximation to $\{x \in R^n \mid h(x) = 0\}$ at $\hat{x}$.

**Proposition 2.** Let $g : R^n \to R^1$ be continuously differentiable and let $g(\hat{x}) = 0$, $g_x(\hat{x}) \neq 0$. Then the set $\{x \in R^n \mid \langle x - \hat{x}, g_x(\hat{x}) \rangle \leqq 0\}$ is a first-order conical approximation to both $\{x \in R^n \mid g(x) \leqq 0\}$, and $\{x \in R^n \mid g(x) < 0\} \cup \{\hat{x}\}$ at $\hat{x}$.

**Proposition 3.** Let $\Omega \subset R^n$ be convex. Then the radial (support) cone $RC(\hat{x}, \Omega)$

to $\Omega$ at any $\hat{x} \in \Omega$, given as $RC(\hat{x}, \Omega) = \{x \in R^n \mid x = \gamma(x' - \hat{x}),\ \gamma \geqq 0,\ x' \in \Omega\}$, is a conical approximation to $\Omega$ at $\hat{x}$.

It can be further realized that in all these cases also other mentioned types of approximations have the same form. However, this is not true in general. Knowledge of these particular cases, which appear very often in practical problems, enables to formulate corresponding optimality conditions in a more familiar form as usual in classical mathematical programming theory.

**Theorem 2.** Let $\hat{x}$ be a solution to a mathematical programming problem (7)–(8), with $f$ being continuously differentiable. Further let $C(\hat{x}, \Omega_i)$ be a first-order conical approximations to the respective sets $\Omega_i$, $i = 1, \ldots, s$. Then there is a number $\mu \leqq 0$ and vectors $a_i \in C'(\hat{x}, \Omega_i)$, $i = 1, \ldots, s$, such that the following conditions are satisfied.

(a) If $\mu = 0$, then at least one of the vectors $a_i$, $i = 1, \ldots, s$ is nonzero.

(b) $\mu f_x(\hat{x}) = a_1 + \ldots + a_s$.

Proof. To verify this result one can follow analogical scheme of [3, 4]. If $f_x(\hat{x}) = 0$, the conclusion of the theorem becomes trivial. Assuming therefore $f_x(\hat{x}) \neq 0$ denote as

$$\Omega_0 = \{x \in R^n \mid f(x) - f(\hat{x}) < 0\} \cup \{\hat{x}\}\,. \tag{9}$$

According to Proposition 2, the set

$$C(\hat{x}, \Omega_0) = \{x \in R^n \mid \langle x - \hat{x}, f_x(\hat{x})^0 \leqq 0\} \tag{10}$$

is a first-order conical approximation to $\Omega_0$ at $\hat{x}$. As in $\hat{x}$ the minimum of $f$ over $\Omega$ is attained, then clearly the intersection $\Omega_0 \cap \Omega_1 \cap \ldots \cap \Omega_s = \{\hat{x}\}$. Thus the all requirements for the application of Theorem 1 to this particular case are met. This, in turn, implies the existence of $a_i \in C'(\hat{x}, \Omega_i)$, $i = 0, 1, \ldots, s$, not all of them being zero, and such that $a_0 + a_1 + \ldots + a_s = 0$. Realizing that $a_0 = -\mu f_x(\hat{x})$, $\mu \leqq 0$, yields the desired result. □

In this way all results described in [3, 4] are valid using the concept of a first-order conical approximation instead of a "mantle". The interested reader can consult these references for a number of various particular cases including equality and/or inequality type constraints, cases with some of the conical approximations in (8) not exhibiting a separation property, etc. For example, if only the overall constraining set $\Omega$, with a first-order conical approximation $C(\hat{x}, \Omega)$ at the minimizing point $\hat{x}$ is considered, the condition (b) of Theorem 2 can be alternatively expressed as

$$\langle x - \hat{x},\ \mu f_x(\hat{x})\rangle \leqq 0 \quad \text{for all} \quad x \in \bar{C}(\hat{x}, \Omega)\,. \tag{11}$$

Similar form of necessary conditions was used through [1] and [5].

## 5. CONCLUSIONS

The aim of this note was to make more explicit the idea briefly sketched in [4]. In this way fairly general and unifying scheme of Boltajnskij [3, 4] has shown to be applicable also if somewhat more general concept of a first-order conical approximation to a set is assumed. This generalization evidently implies to some other problems, e.g. of the minimax type studied in [7]. Moreover, it is felt that on defining a "local" analogy of Definition 1 in the same way as in [4] the concept of a local mantle was introduced, a further generalization of Theorem 1 would be possible.

On the other hand, it seems that in the studied finite-dimensional case one can hardly expect some basic generalizations in the future without removing a differentiability assumption concerning various functions in a problem description. Especially in this direction a number of interesting results was achieved quite recently, e.g. see [8, 9].

(Received March 24, 1981.)

## REFERENCES

[1] M. D. Canon, C. D. Cullum, E. Polak: Theory of Optimal Control and Mathematical Programming. McGraw-Hill, New York 1970.

[2] L. W. Neustadt: An abstract variational theory with applications to a broad class of optimization problems, I., General theory. SIAM J. Control *4* (1966), 3, 505—527.

[3] V. G. Boltjanskij: Optimal Control of Discrete Systems. Nauka, Moscow 1973. In Russian.

[4] V. G. Boltjanskij: The method of tents in the theory of extremal problems. Uspechi matematičeskich nauk *30* (1975), 3, 3—55. In Russian.

[5] M. Canon, C. Cullum, E. Polak: Constrained minimization problems in finite dimensional spaces. SIAM J. Control *4* (1966), 3, 528—547.

[6] A. J. Dubovickij, A. A. Miljutin: Extremum problems involving constraints. Žurnal vyč. matematiki i mat. fiziki *5* (1965), 3, 395—453. In Russian.

[7] V. G. Boltjanskij, I. S. Čebotaru: Necessary conditions in a minimax problem. Dokl. Akad. Nauk SSSR *213* (1973), 2, 257—260. In Russian.

[8] F. H. Clarke: A new approach to Lagrange multipliers. Mathem. Oper. Res. *1* (1976), 2, 165 to 174.

[9] J. B. Hiriart-Urruty: Refinements of necessary optimality conditions in non-differentiable programming I. Appl. Math. Optim. *5* (1979), 1, 63—82.

*Ing. Jaroslav Doležal, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.*