# An Optimality Condition for a Nondifferentiable Convex Program*

**Henry Wolkowicz**
*Department of Mathematics, The University of Alberta, Edmonton, Alberta, Canada T6G 2G1*

**An optimality condition for the ordinary convex programming problem which utilizes the directional derivatives of the constraints is studied.**

## 1. INTRODUCTION

Consider the ordinary convex programming problem:

$$\text{(P)} \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g^k(x) \le 0,\ k \in P = \{1, \ldots, m\}, \end{array}$$

where $f, g^k \colon R^n \to R$ are convex, not necessarily differentiable, functions. We study an optimality condition for (P) that utilizes the directional derivatives of the binding constraints. We show that this optimality condition holds under *any* constraint qualification and so is implied by the Karush–Kuhn–Tucker conditions and is, in fact, equivalent to those conditions in the case of differentiable constraints.

## 2. PRELIMINARIES

Following [1], we introduce the *cone of directions of decrease* of the convex function $h$ at $x$:

$$D_h^{<}(x) = \{d\colon \text{there exists } \bar{\alpha} > 0 \text{ with } h(x + \alpha d) < h(x), \text{ for all } 0 < \alpha \le \bar{\alpha}\}.$$

The *directional derivative* of $h$ at $x$ in the direction $d$ is defined as

$$\nabla h(x;d) = \lim_{t \downarrow 0} \frac{h(x + td) - h(x)}{t}.$$

Finite-valued convex functions have the nice property that the directional derivatives exist universally. A vector $\phi \in R^n$ is said to be a *subgradient* of $h$ at $x$ if

$$h(z) \ge h(x) + \phi(z - x), \qquad \text{for all } z \in R^n.$$

The set of all subgradients of $h$ at $x$ is then called the subdifferential of $h$ at $x$ and is

*This research was partly supported by the National Science and Engineering Research Council of Canada, and was done while the author was situated at Dalhousie University, Halifax, Canada.

*Naval Research Logistics Quarterly*, Vol. 30, pp. 415–418 (1983)
 CCC 0028-1441/83/030415-04$01.40

denoted $\partial h(x)$. In general (when $h$ is finite-valued), $\partial h(x)$ is a nonempty, compact, convex set and

$$\nabla h(x;d) = \max\{\phi d:\ \phi \in \partial h(x)\}, \tag{1}$$

where for two vectors $u$ and $v$ in $R^n$ we let $uv$ denote the dot product. Furthermore,

$$D_h^{<}(x) = \{d:\ \nabla h(x;d) < 0\}; \tag{2}$$

$$-D_h^{<}(x)^+ = \bigcup_{\lambda \geq 0} \lambda \partial h(x), \qquad \text{when } 0 \notin \partial h(x), \tag{3}$$

where, if $M$ is a set in $R^n$, then the *nonnegative polar* of $M$ is

$$M^+ = \{\phi:\ \phi m \geq 0, \qquad \text{for all } m \in M\}.$$

(For the above and other related facts, see, e.g., [1] and [4].)

The *linearizing cone* at $x$ is

$$C(x) = \{d:\ \phi d \leq 0, \qquad \text{for all } \phi \in \partial g^k(x) \text{ and all } k \in P(x)\},$$

where $P(x)$ is the set of *binding* (*active*) *constraints* at $x$, i.e.,

$$P(x) = \{k \in P:\ g^k(x) = 0\}.$$

By Equation (1), we see that

$$C(x) = \{d:\ \nabla g^k(x;d) \leq 0, \qquad \text{for all } k \in P(x)\}. \tag{4}$$

The *cone of subgradients* at $x$ is

$$B(x) = \{\phi:\ \phi = \Sigma \lambda_k \phi^k, \qquad \text{for some } \lambda_k \geq 0,\ \phi^k \in \partial g^k(x) \text{ and } k \in P(x)\}. \tag{5}$$

The linearizing cone and the cone of subgradients have the following dual property (see, e.g., [5]):

$$\text{closure } -B(x) = C^+(x). \tag{6}$$

The Karush–Kuhn–Tucker conditions of optimality (see, e.g., [4]) can now be expressed as

$$\partial f(x) - B(x) \neq \emptyset \tag{7}$$

or equivalently as

$$0 \in \partial f(x) + \Sigma \lambda_k \partial g^k(x), \qquad \text{for some } \lambda_k \geq 0 \text{ and } k \in P(x). \tag{8}$$

If $x$ is feasible, then these conditions are always sufficient for optimality. Necessity may fail unless some constraint qualification holds at $x$ (see, e.g., [1]).

## 3. THE OPTIMALITY CONDITION

Consider the optimality condition

$$C(x) \cap D_f^{<}(x) = \emptyset. \tag{9}$$

We will now show that the condition (9) is equivalent to the Karush–Kuhn–Tucker conditions (7) [or (8)] whenever $B(x)$ is closed. [Note that $B(x)$ is a finitely generated polyhedral cone and thus closed, if the constraints $g^k$ are differentiable.]

THEOREM 3.1: Suppose that $B(x)$ is closed. Then (9) is equivalent to (7) [or (8)].

PROOF: If $0 \in \partial f(x)$, then $D_f^{<}(x) = \emptyset$ and the result follows. Now suppose that $0 \notin \partial f(x)$. Then $D_f^{<}(x)$ is a nonempty, open, convex cone, $C(x)$ is a convex cone, and the Hahn–Banach theorem (see, e.g., [2]) implies that (9) holds if and only if there exists

$$\begin{aligned} 0 \neq \phi &\in -D_f^{<}(x)^+ \cap C(x)^+ \\ &= \bigcup_{\lambda \geq 0} \lambda \partial f(x) \cap -B(x), \end{aligned}$$

by (3) and (6). Q.E.D.

Thus the condition (9) is equivalent to the Karush–Kuhn–Tucker conditions (7) if the constraints $g^k$ are differentiable. In fact, it is easy to see that (9) is always a sufficient condition for optimality. Moreover, necessity will hold under any constraint qualification whatsoever, since the closure of $B(x)$ is implied by every constraint qualification (see, e.g., [5]).

We were motivated in studying (9) by a result of Mond and Schechter [3]. They studied program (P) with the added assumptions that (i) the constraints $g^k$ are differentiable, (ii) the objective function $f$ is of a special type, and (iii) the generalized Slater condition holds, i.e., there exists a feasible point $x_*$ such that $g^k(x_*) < 0$ whenever $g^k$ is nonaffine. Our results show that these assumptions are not required for (9) to hold.

If $x$ solves (P) and $f$ is some particular fixed objective function, then both (7) and (9) may hold even though $B(x)$ fails to be closed. A trivial example of this occurs when $0 \in \partial f(x)$, i.e., when $x$ is a global minimum of $f$. However, it may happen that (9) holds but (7) fails—(9) is the weaker necessary condition. [Note that if $B(x)$ is not closed, then one can always find a (linear) objective function $f$ for which (7) will fail at the optimum point $x$ (see, e.g., [5]).]

*EXAMPLE 3.1**

Let

$$K = \{x \doteq (x_i) \in R^2 : x_1^2 + (x_2 - 1)^2 \leq 1\};$$

$$g^1(x) = \max\{\phi x : \phi \in K\}; \text{ and } f(x) = x_1.$$

Then $x = 0$ solves (P), (9) holds but (7) fails. [Note that $B(x)$ is the origin union the upper, open half-plane while $C(x)$ is the negative $x_2$ axis.]

## REFERENCES

[1] Ben-Israel, A., Ben-Tal, A., and Zlobec, S., *Optimality in Nonlinear Programming: A Feasible Directions Approach*, Wiley, New York, 1981.

*This example is due to Professor J. M. Borwein.

[2] Girsanov, I. V., *Lectures on Mathematical Theory of Extremum Problems*, Lecture Notes in Economics and Mathematical Systems, No. 67, Springer-Verlag, New York, 1972.
[3] Mond, B., and Schechter, M., "On a constraint qualification in a nondifferentiable programming problem," *Naval Research Logistics Quarterly*, **23**, 611–613 (1976).
[4] Rockafellar, R. T., *Convex Analysis*, Princeton U. P., Princeton, 1970.
[5] Wolkowicz, H., "Geometry of optimality conditions and constraint qualifications: the convex case," *Mathematical Programming*, **19**, 32–60 (1980).