A Strengthened Test for Optimality

H. Wolkowicz

Communicated by A. V. Fiacco

Abstract. In this paper, we strengthen recent characterizations of optimality for the convex program

\[ \mu = \inf \{ f(x) : g_k(x) \leq 0, \ k = 1, \ldots, m \}, \]

where the functions \( f \) and \( g_k, k = 1, \ldots, m \), are convex functions on \( \mathbb{R}^n \). The characterizations presented here are stronger, in the sense that the Lagrange multiplier relation holds over a larger set. This strengthens information about the stability of the solution with respect to perturbations in the right-hand side of the constraints. In particular, we show that, in the characterizations of optimality in Refs. 1–2, the set \( D_{\partial^-(x^*)} \), the intersection of the cones of directions of constancy of the equality constraints, can be replaced by the larger and simpler set \( D_h(x^*) \), the cone of directions of constancy of a single function \( h \). We also discuss how to choose \( h \) to get the strongest characterization.

Key Words. Characterization of optimality, Lagrange multipliers, faithfully convex functions, gradients, cones of directions of constancy, strongest optimality conditions.

1. Introduction

Consider the convex program

\[ \mu = \inf \{ f(x) : g_k(x) \leq 0, \ k \in \mathcal{P} = \{1, \ldots, m\} \}, \]

where \( f, g^k : \mathbb{R}^n \to \mathbb{R} \) are differentiable convex functions. Without loss of generality, we assume that none of the functions is constant. We will require that some of the constraints \( g^k \) be faithfully convex functions, i.e., convex

---

1 This research was partially supported by NSERC-A-3388.
2 Assistant Professor, Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada.
functions which are not affine along any line segment, unless they are affine along the entire line extending the segment (e.g., Ref. 3). The class of faithfully convex functions is large and it includes all analytic convex functions as well as all strictly convex functions. Characterizations of optimality for Program (P) have been given in Refs. 1–2. These characterizations hold without any constraint qualification, and they implicitly include a Lagrange multiplier relation which holds on the convex set

$$x^* + D_{\mathcal{P}^0}^{-}(x^*),$$

where $D_{\mathcal{P}^0}^{-}(x^*)$ is the intersection of the cones of directions of constancy of the equality constraints at the optimum $x^*$. It is of interest to have strong optimality conditions, in the sense that the Lagrange multiplier relation holds on as large a convex set as possible (e.g., Ref. 4). For example, if we have $\lambda_k \geq 0$ satisfying

$$f(x) + \sum_{k \in \mathcal{K}} \lambda_k g^k(x) \geq \mu, \quad \text{for all } x \in \Omega,$$  \hspace{1cm} (1)

then, if $\Omega$ is all of $\mathbb{R}^n$, our solution is stable with respect to (feasible) perturbations of the right-hand sides of the constraints (e.g., Ref. 5). Similarly, the larger the set $\Omega$ is, the more we can guarantee that the solution is stable with respect to certain perturbations (e.g., Refs. 5–6).

In this paper, we show that we can replace $D_{\mathcal{P}^0}^{-}(x^*)$ by various larger cones. In particular, if

$$\alpha_k \geq 0, \quad \text{for all } k \in \mathcal{P}^0,$$

with $\alpha_k > 0$ if $g^k$ is not affine, then we can use the cone $D_h^{-}(x^*)$ when

$$h = \sum_{k \in \mathcal{P}^0} \alpha_k g^k$$

is faithfully convex; i.e., we can replace the intersection of the cones of constancy $D_{\mathcal{P}^0}^{-}(x^*)$ by the larger and simpler cone of constancy of the single function $h$. Note that the $\alpha_k$ are arbitrary (nonnegative) constants. Thus we can change the values of the constants $\alpha_k$ and obtain a variety of different cones $D_h^{-}(x^*)$. This theory simplifies the algorithm for finding the set $\mathcal{P}^0$, given in Ref. 1 (see Remark 3.2.).

In addition, we show how to weaken the differentiability and faithfully convex assumptions; we discuss how to find the scalars $\alpha_k$ which give the strongest optimality conditions, and compare our results with those in Ref. 7, which use the badly behaved set of constraints.

Note that our results in Sections 2 and 3 hold when $\mathbb{R}^n$ is replaced by a locally convex (Hausdorff) topological vector space.