

OPTIMALITY CONDITIONS FOR A CONE-CONVEX PROGRAMMING PROBLEM

HÉLÈNE M. MASSAM

(Received 14 November 1977; revised 2 May 1978)

Communicated by A. P. Robertson

Abstract

Optimality conditions without constraint qualifications are given for the convex programming problem:

Maximize $f(x)$ such that $g(x) \in B$,

where f maps X into R and is concave, g maps X into R^m and is B -concave, X is a locally convex topological vector space and B is a closed convex cone containing no line. In the case when B is the nonnegative orthant, the results reduce to some of those obtained recently by Ben-Israel, Ben-Tal and Zlobec.

Subject classification (Amer. Math. Soc. (MOS) 1970): 90 C 25.

1. Introduction

This paper gives both necessary and sufficient conditions of optimality without any constraint qualification for the following convex programming problem:

(P) Maximize $f(x)$ such that $g(x) \in B$,

where f maps X into R , g maps X into R^m , X is a locally convex topological vector space, B is a closed convex cone in R^m containing no line. The function f is concave and g is B -concave.

Karush–Kuhn–Tucker-type optimality conditions have been given for similar convex programming problems over cones (see Craven (1974) and (1977)). For $B = R_+^m$, the nonnegative orthant, optimality conditions without constraint qualification can be found in Ben-Tal et al. (1976); see also Ben-Israel et al. (1976).

Here, the problem is considered for a general closed convex cone B containing no line, and no regularity conditions such as Slater's constraint qualification is needed for the given characterizations of optimality.

In order to state optimality conditions without constraint qualifications, Ben-Tal et al. (1976) introduced a particular cone, called the cone of directions of constancy of the function g at the feasible point \bar{x} . Similarly here, we pay special attention to certain directions forming a cone called the generalized cone of constancy of g at \bar{x} with respect to B .

In the next section, all the definitions and lemmata necessary to the study of this cone and necessary for obtaining characterizations of optimality are presented. In Section 3, the 'generalized cone of constancy' of g at \bar{x} with respect to B is introduced. In Section 4, two primal and one dual characterization of optimality are given. The primal theorems state that a feasible point \bar{x} is optimal if and only if a system involving \bar{x} and a direction d in the primal space X is inconsistent. The dual theorem says that \bar{x} is optimal if and only if a system expressed in terms of multipliers in the dual space is consistent. These results are illustrated by an example. Finally, in Section 5, the connection with the optimality conditions given by Ben-Israel et al. (1976), when $B = R_+^m$, is established.

2. Preliminary results

In this section we recall some of the definitions and lemmata that we shall need in the proof of the characterizations of optimality in Section 4.

A subset C of X is a cone if, for any x in C and $\lambda \geq 0$, λx is in C . A function $f: X \rightarrow R^m$ defined on a convex subset D of X is said to be convex with respect to the closed convex cone B or B -convex, (see for example Craven (1974)), if for all x_1, x_2 in D and all λ in $[0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) - \lambda f(x_1) - (1 - \lambda)f(x_2) \in -B.$$

We shall use the classical definitions of the polar cone (see Guignard (1969)), the tangent cone and the pseudotangent cone. The *polar cone* C^+ to the cone C is defined by

$$C^+ = \{x' \in X' \text{ such that } x'(x) \geq 0, \text{ for all } x \text{ in } C\}.$$

Given a subset A of X and \bar{x} in A , the *tangent cone* to A at \bar{x} is

$$T(A, \bar{x}) = \{z \in X: \text{there exists } \{z_n\}_{n=1}^\infty \subset A \text{ and}$$

$$\{\lambda_n\}_{n=1}^\infty \subset R_+ \text{ such that } z_n \rightarrow \bar{x} \text{ and } \lambda_n(z_n - \bar{x}) \rightarrow z\}.$$

The *pseudotangent cone* to A at \bar{x} , denoted $P(A, \bar{x})$, is the closure of the convex hull of $T(A, \bar{x})$. The *support cone to a convex set C at \bar{x} in C* is defined (see Holmes (1972)) as the closure of the union of rays issued from \bar{x} and going through C . It is denoted by $S(C, \bar{x})$ and defined by

$$S(C, \bar{x}) = \overline{\{\bar{x} + \lambda(x - \bar{x}) \text{ such that } x \in C\}}.$$

Supporting hyperplane, relative interior and face are also concepts that we are going to use extensively.

DEFINITION 2.1. A nonzero continuous linear functional f on X is said to be a *supporting functional* for a set A at $x_0 \in A$ if $f(x) \geq f(x_0)$ for all $x \in A$. The closed hyperplane $H = \{x : f(x) = f(x_0)\}$ is called a *supporting hyperplane* to A at the point x_0 . The closed halfspace determined by the supporting hyperplane containing A is called a *supporting halfspace* to A at x_0 . The set A is said to 'lie' on one side of the hyperplane. Note that a supporting hyperplane to a cone C at a boundary point of C separates C from $-C$.

DEFINITION 2.2. A subset A of R^n is called an *affine set* if $(1 - \lambda)x + \lambda y \in A$ for every x, y in A and λ in R . The *affine hull* of a set A is the smallest affine set containing A and is denoted $\text{aff } A$.

DEFINITION 2.3. The relative interior of a convex set C in R^n , denoted $\text{ri } C$, is the interior of the set C regarded as a subset of its affine hull. The *relative boundary* of C , denoted ∂C , is defined by $\partial C = \text{cl } C \setminus \text{ri } C$.

Every convex set in R^n has a relative interior. This statement is not true in abstract spaces.

DEFINITION 2.4. A *face* of a convex set C is a convex subset C' of C such that every closed line segment in C with a relative interior point in C' has both end points in C' . The empty set and C itself are faces of C . If C' is the set of points where a certain linear function h achieves its maximum over C , then C' is a face of C . A face is called *exposed* if it is equal to the intersection of C with a nontrivial supporting hyperplane H to C , that is, a supporting hyperplane not containing all of C . It is clear that an exposed face of a closed convex cone is a closed convex cone.

We shall also need to know some properties of the cones we consider. The most important ones are stated in the following lemmata. All of them are known results. The proof of the first two results can be found in Massam (1977). The other four are proved in Rockafellar (1970). The set B is a closed convex cone in

\mathbb{R}^m containing no line. Unless otherwise specified, \bar{y} in B is assumed different from the vertex 0 of B . Without loss of generality, we can also assume that B has dimension m .

LEMMA 2.1. *The support cone to B at \bar{y} , $S(B, \bar{y})$, is equal to the intersection of the supporting halfspaces to B at \bar{y} .*

LEMMA 2.2. *Let $\bar{y} \in B$, then $S(B, \bar{y}) = T(B, \bar{y}) = T(B, \bar{y}) \pm \bar{y}$ and $\partial T(B, \bar{y})$ is included in the union of supporting hyperplanes to B at \bar{y} .*

Note that if $\bar{y} = 0$ is the vertex of B , then $T(B, \bar{y}) = S(B, \bar{y}) = B$ and all the above lemmata remain true.

LEMMA 2.3. *If F' and F'' are faces of a convex set C such that $\text{ri } F'$ and $\text{ri } F''$ have a point in common, then actually $F' = F''$.*

LEMMA 2.4. *If F is a face of a convex set C and D is a convex set such that $F \subset D \subset C$, then F is a face of D .*

LEMMA 2.5. *If F is a face of F' and F' a face of C , then F is a face of C .*

LEMMA 2.6. *Let C be a nonempty convex set and let U be the collection of all relative interiors of nonempty faces of C . Then U is a partition of C , that is, the sets in U are disjoint and their union is C . Every relatively open convex subset of C is contained in one of the sets in U and these are the maximal relatively open convex subsets of C .*

Finally let us give a property of a B -concave function.

LEMMA 2.7. *Let g be defined on a convex set D , B -concave and Fréchet-differentiable, then $g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -B$ for any $\alpha \geq 0$ such that $\bar{x} + \alpha d \in D$.*

PROOF. By definition of B -concavity,

$$g(\lambda(\bar{x} + \alpha d) + (1 - \lambda)\bar{x}) - \lambda g(\bar{x} + \alpha d) - (1 - \lambda)g(\bar{x}) \in B,$$

that is,

$$g(\bar{x} + \lambda \alpha d) - \lambda g(\bar{x} + \alpha d) - (1 - \lambda)g(\bar{x}) \in B.$$

But, B being a cone,

$$\frac{g(\bar{x} + \lambda \alpha d) - g(\bar{x})}{\lambda} - (g(\bar{x} + \alpha d) - g(\bar{x})) \in B.$$

Let $\lambda \rightarrow 0$, then

$$\nabla g(\bar{x})(\alpha d) - g(\bar{x} + \alpha d) + g(\bar{x}) \in B.$$

Let us note that the previous lemma and all subsequent results are still valid if X is no longer a normed space but simply a locally convex topological vector space and if functions defined on X are not Fréchet but only compact-differentiable (see definition of compact-differentiability in Averbukh and Smolyanov (1968)). The notation $\nabla g(\bar{x})(\cdot)$ will then stand for the compact derivative of g at \bar{x} .

3. The generalized cone of constancy

In this section, we give the definition of the generalized cone of constancy. We also try to give the reader a geometric insight for the definition of this cone, by examining a particular example.

Let \mathcal{P} be the indexing set of all exposed faces F_k of B . We denote by $\mathcal{P}(x)$ the indexing set of all exposed faces of B containing x , that is, $\mathcal{P}(x) = \{k \in \mathcal{P} \text{ such that } g(x) \in F_k\}$. The set $A = \{x \in X : g(x) \in B\}$ is the feasible set of the problem (P). Let us also define \mathcal{P}^- as $\mathcal{P}^- = \{k \in \mathcal{P} \text{ such that } x \in A \text{ implies } g(x) \in F_k\}$ and B^- as $B^- = \bigcap_{k \in \mathcal{P}^-} F_k$.

DEFINITION 3.1. Let \bar{x} be a feasible solution of (P). The *generalized cone of constancy* of g with respect to B is defined by

$$C(g, B, \bar{x}) = \{d \in X : \text{there exists } \bar{\alpha} > 0 \text{ such that for all } \alpha \in (0, \bar{\alpha}] \\ g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri } B^-, \quad g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -B^-\}.$$

In Section 4, two primal characterizations of optimality will be given. One involves only one face of B , namely B^- and is expressed in function of $C(g, B, \bar{x})$. The other involves several faces of B and is expressed in term of another cone that we shall call the *multi-face generalized cone of constancy* of g at \bar{x} with respect to B . Let us define this new cone.

DEFINITION 3.2. Let \bar{x} be a feasible solution of (P). The *multi-face generalized cone of constancy* of g at \bar{x} with respect to B is defined as the union, over all subsets I of $\mathcal{P}(\bar{x})$, of sets \tilde{C}_I defined as follows:

$$\tilde{C}_I(g, B, \bar{x}) = \{d \in X : \text{there exists } \bar{\alpha} > 0 \text{ such that for all } \alpha \in (0, \bar{\alpha}] \\ \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} H_i, \quad \nabla g(\bar{x})(\alpha d) \notin \bigcap_{i \in I'} H_i \text{ if } I' \supset I, \\ g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri } \bigcap_{i \in I} F_i, \quad g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in - \bigcap_{i \in I} F_i\},$$

where H_i 's are the supporting hyperplanes to B at $g(\bar{x})$. So

$$\tilde{C}(g, B, \bar{x}) = \bigcup_{I \subset \emptyset(\bar{x})} \tilde{C}_I(g, B, \bar{x}).$$

An example of the cone $C(g, B, \bar{x})$ is now given.

EXAMPLE 3.1. Let

$$g(x_1, x_2) = \begin{bmatrix} g_1(x_1, x_2) \\ g_2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} 1 - x_2 \\ -(1 - x_1)^2 \end{bmatrix}.$$

Let $B = \{(y_1, y_2)^t : y_2 \geq 0, y_1 + y_2 \geq 0\}$ and let $\bar{x} = (1, 1)^t$ so that $g(\bar{x}) = (0, 0)^t \in \partial B$.

The feasible set is $A = \{(x_1, x_2)^t : x_1 = 1, x_2 \leq 1\}$. The supporting hyperplanes to B at $g(\bar{x})$ are $H_1 = \{(y_1, y_2)^t : y_2 = 0\}$ and $H_2 = \{(y_1, y_2)^t : y_1 + y_2 = 0\}$. The corresponding faces are

$$F_1 = \{(y_1, y_2)^t : y_2 = 0, y_1 \geq 0\} \quad \text{and} \quad F_2 = \{(y_1, y_2)^t : y_1 + y_2 = 0, y_2 \geq 0\}.$$

If $x \in A$, $g(x) \in F_1$ so that

$$B^\circ = \{(y_1, y_2)^t : y_1 \geq 0, y_2 = 0\} \quad \text{and} \quad \text{ri } B^\circ = \{(y_1, y_2)^t : y_1 > 0, y_2 = 0\}$$

$$\nabla g(\bar{x}) = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \quad \text{so that} \quad \nabla g(\bar{x})(d) = \begin{bmatrix} -d_2 \\ 0 \end{bmatrix} \in H_1,$$

g is R_+^2 -concave and therefore B -concave because $R_+^2 \subset B$,

$$C(g, B, \bar{x}) = \{(d_1, d_2)^t : \text{there exists } \bar{\alpha} > 0 \text{ such that for all } \alpha \in (0, \bar{\alpha}]\}$$

$$g_1(\bar{x}) + \nabla g_1(\bar{x})(\alpha d) > 0,$$

$$g_1(\bar{x} + \alpha d) - g_1(\bar{x}) - \nabla g_1(\bar{x})(\alpha d) \leq 0,$$

$$g_2(\bar{x}) + \nabla g_2(\bar{x})(\alpha d) = 0,$$

$$g_2(\bar{x} + \alpha d) - g_2(\bar{x}) - \nabla g_2(\bar{x})(\alpha d) = 0\},$$

$$C(g, B, \bar{x}) = \{(d_1, d_2)^t : \text{there exists } \bar{\alpha} > 0 \text{ such that for all } \alpha \in (0, \bar{\alpha}]\}$$

$$-\alpha d_2 > 0, \quad -\alpha d_2 + \alpha d_2 \leq 0,$$

$$-\alpha^2 d_1^2 = 0, \quad -\alpha d_2 > 0\},$$

$$C(g, B, \bar{x}) = \{(d_1, d_2)^t : d_1 = 0, d_2 < 0\}.$$

Let us now consider an example where the usual Kuhn–Tucker conditions do not apply. This example will motivate the later theory and show why the cones C and \tilde{C} have been introduced. In order to do so, we shall consider the ice-cream cone B given in Example 2.5 of Ben-Israel (1969) and a point $g(\bar{x})$ on the boundary of B . The different ways in which $g(\bar{x} + \alpha d)$ behaves when α varies from 0 to a small $\bar{\alpha} > 0$ will be outlined. Finally, for the function $g(\cdot)$ as given in Example 2.5 of Ben-Israel (1969), we shall compute the cone $C(g, B, \bar{x})$.

Consider the following cone

$$B = \{(y_1, y_2, y_3)^t : 2y_1 y_3 \geq y_2^2, y_1 \geq 0, y_3 \geq 0\}.$$

Suppose that a function $g: X \rightarrow R$ is such that, for some \bar{x} in X ,

$$g(\bar{x}) = (g_1(\bar{x}), g_2(\bar{x}), g_3(\bar{x}))^t = (0, 0, 1)^t.$$

The tangent cone $T(B, g(\bar{x}))$ to B at $g(\bar{x})$ is the closed halfspace determined by the supporting hyperplane $H = \{(y_1, y_2, y_3)^t : y_1 = 0\}$ to B at $g(\bar{x})$, and containing B . If for some direction d , $\nabla g(\bar{x})(d)$ belongs to the interior of $T(B, g(\bar{x}))$, as it will be shown in the proof of Theorem 4.1, for α sufficiently small, $g(\bar{x} + \alpha d) \in B$. If $\nabla g(\bar{x})(d)$ does not belong to $T(B, g(\bar{x}))$, it will also be shown that there is no $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$. Suppose now that $\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x}))$, then one does not know whether $g(\bar{x} + \alpha d)$ will be inside or outside B . However, $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$ belongs to H for any positive α and by B -concavity of g ,

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -B \text{ for sufficiently small } \alpha.$$

This means that for sufficiently small α , $g(\bar{x} + \alpha d)$ belongs to the half-space determined by H but not containing B . It is then clear that if $g(\bar{x} + \alpha d)$ is to belong to B for sufficiently small α , it has to belong to H also, that is, to $B \cap H$. The direction d will, then, have to satisfy the following condition

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -(B \cap H) = -F,$$

which is the second property of any direction of the multi-face generalized cone of constancy. We noticed that $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$ belongs to H . It is also geometrically clear that if it does not belong to B , then $g(\bar{x} + \alpha d)$ cannot be in B for sufficiently small $\alpha > 0$. So,

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \text{ belongs to } B \cap H = F.$$

Thus, it has been roughly indicated that, if for some point \bar{x} in X and some direction d in X , $g(\bar{x}) \in \partial B$ and $\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x}))$, then $g(\bar{x} + \alpha d)$ belongs to B for

sufficiently small α only if d is such that,

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in F, \quad g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -F.$$

for α sufficiently small.

Suppose now that $g(x) = g(x_1, x_2) = (0, x_1, x_2)^t$, with $g(\bar{x}) = (0, 0, 1)^t$. The directional derivative of g at \bar{x} is

$$\nabla g(\bar{x})(d) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} (d_1 \ d_2)^t = (0, d_1, d_2)^t.$$

The feasible set A is equal to $A = \{(x_1, x_2)^t : x_1 = 0, x_2 \geq 0\}$. So if $x \in A$,

$$g(x) = (0, 0, y_3)^t$$

where y_3 is nonnegative. Therefore $\mathcal{P}^* \neq \emptyset$ and the face B^* is

$$B^* = \{(y_1, y_2, y_3)^t : y_1 = y_2 = 0, y_3 \geq 0\}.$$

Moreover, if H is, as defined before, the supporting hyperplane to B at $g(\bar{x})$,

$$B \cap H = \{(y_1, y_2, y_3)^t : y_1 = y_2 = 0, y_3 \geq 0\}$$

so that, $F = B^*$ and F is the only face of B containing $g(\bar{x})$. Consequently $C(g, B, \bar{x}) = \check{C}(g, B, \bar{x})$. We can now compute the generalized cone of constancy of g with respect to B at $g(\bar{x})$.

$$C(g, B, \bar{x}) = \{d \in \mathbb{R}^2 : \text{there exists } \bar{\alpha} > 0 \text{ such that for all } \alpha \in (0, \bar{\alpha})$$

$$0 + \alpha d_1 = 0, \quad 1 + \alpha d_2 > 0,$$

$$\bar{x}_1 + \alpha d_1 - \bar{x}_1 - \alpha d_1 = 0, \quad \bar{x}_2 + \alpha d_2 - \bar{x}_2 - \alpha d_2 \leq 0\}$$

$$= \{d \in \mathbb{R}^2 : \text{there exists } \bar{\alpha} > 0 \text{ such that for all } \alpha \in (0, \bar{\alpha})$$

$$d_1 = 0, \quad \alpha d_2 > -1\}$$

$$= \{(d_1, d_2)^t : d_1 = 0, d_2 \text{ arbitrary}\}.$$

Hence, in this particular example, if $\nabla g(\bar{x})(d)$ belongs to H , $g(\bar{x} + \alpha d)$ belongs to B for $\alpha > 0$ sufficiently small only if d is such that $d_1 = 0$, that is, only if $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$ belongs to the positive y_3 -axis for $\alpha > 0$ sufficiently small.

We shall consider the general case in the next section.

4. Characterization of optimality

Three necessary and sufficient conditions of optimality for arbitrary but fixed feasible point \bar{x} are given in this section, Theorems 4.1, 4.2, and 4.3. The first one is a primal characterization using the cone $\check{C}(g, B, \bar{x})$. From this, a primal characterization using the cone $C(g, B, \bar{x})$ can be easily deduced. It is this result that we shall consider as the main result. Finally, using the Dubovitskii–Milyutin Theorem of the alternative, a dual characterization of optimality is also given.

Before we give the main results, let us recall that $\mathcal{P}^=$ is the indexing set of all faces F_k such that if, $g(x) \in B$, then $g(x) \in F_k$. If $\mathcal{P}^= = \emptyset$, then there is a feasible \hat{x} such that $g(\hat{x}) \in \text{int } B$. We shall say that Slater’s constraint qualification is satisfied. Then, when $B = R_+^m$, our results are the same as those provided by the traditional Karush–Kuhn–Tucker theory (see Kuhn and Tucker (1951), Mangasarian (1969)). In the proof of the main theorem, Theorem 4.2, the following Lemma 4.1 is needed. In the proof of the dual theorem, Theorem 4.3, the following Lemmata 4.2 and 4.3 are needed. Let us state and prove them.

LEMMA 4.1. Assume that $\mathcal{P}^= = \emptyset$. Then the consistency of the system

$$(S_1) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$$

is equivalent to the consistency of the system

$$(S_2) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in T(B, g(\bar{x})).$$

PROOF. If (S_1) is consistent, it is obvious that (S_2) is also consistent. Assume now that (S_2) is consistent. Then there exists $d \in X$ such that

$$\nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in T(B, g(\bar{x})).$$

As $\mathcal{P}^= = \emptyset$, there exists \hat{x} feasible such that $g(\hat{x}) \in \text{int } B$. Let $\hat{d} = \hat{x} - \bar{x}$, then by B -concavity of g ,

$$g(\hat{x}) - g(\bar{x}) - \nabla g(\bar{x})(\hat{d}) \in -B$$

that is,

$$\nabla g(\bar{x})(\hat{d}) \in g(\hat{x}) - g(\bar{x}) + B \subset g(\hat{x}) - g(\bar{x}) + T(B, g(\bar{x})) \subset \text{int } T(B, g(\bar{x})).$$

Consider now $d = \hat{d} + \alpha \hat{d}$,

$$\nabla f(\bar{x})(d) = \nabla f(\bar{x})(\hat{d}) + \alpha \nabla f(\bar{x})(\hat{d}) > 0 \quad \text{for sufficiently small } \alpha,$$

and

$$\nabla g(\bar{x})(d) = \nabla g(\bar{x})(\hat{d}) + \alpha \nabla g(\bar{x})(\hat{d}) \in \text{int } T(B, g(\bar{x})),$$

so that the system (S_1) is consistent and the lemma is proved.

LEMMA 4.2. *The set $M = \{x: g(x) \in B^-\}$ is convex and when $\mathcal{P}^- \neq \emptyset$, the restriction of g on M is B^- -convex.*

PROOF. Let x_1, x_2 be in M , let λ be in $[0, 1]$. By B -concavity of g ,

$$g(\lambda x_1 + (1 - \lambda)x_2) - \lambda g(x_1) - (1 - \lambda)g(x_2) \in B$$

and therefore $g(\lambda x_1 + (1 - \lambda)x_2) \in B^- + B^- + B = B$. But, by definition of B^- , if $g(\lambda x_1 + (1 - \lambda)x_2)$ belongs to B , it belongs to B^- . So $g(\lambda x_1 + (1 - \lambda)x_2) \in B^-$ and M is convex. We also just proved that if x_1, x_2 belong to M and λ to $[0, 1]$, then $g(x_1), g(x_2), g(\lambda x_1 + (1 - \lambda)x_2)$ belong to the supporting hyperplane H to B determined by $H \cap B = B^-$. Thus, for any such x_1, x_2, λ ,

$$g(\lambda x_1 + (1 - \lambda)x_2) - \lambda g(x_1) - (1 - \lambda)g(x_2) \in B \cap H = B^-,$$

that is, g defined on M is B^- -convex.

LEMMA 4.3. *Let $\mathcal{P}^- \neq \emptyset$. Then there exists a feasible point x such that $g(x)$ belongs to the relative interior of B^- .*

PROOF. Suppose that for any x in M , $g(x)$ belongs to the relative boundary of B^- . Then, by B^- -concavity of g on M , $g(x)$ belongs to the intersection of B^- with a supporting hyperplane H' , not containing B^- but containing part of its boundary. Without loss of generality, we can assume that H' is also supporting B at a boundary point of B^- . Then $B^- \cap H'$ is a face of B and for any x in M , $g(x)$ belongs to $B^- \cap H'$, which is strictly included in B^- . This contradicts the definition of B^- . So, there exists x feasible, that is, x in M such that $g(x) \in \text{ri } B^-$.

Let us now state and prove the primal characterization of optimality expressed in terms of the multiface generalized cone of constancy $\tilde{C}(g, B, \bar{x})$.

THEOREM 4.1. *Let \bar{x} be a feasible solution of (P) . Then \bar{x} is optimal if, and only if, the following system is inconsistent.*

$$(\tilde{S}) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in T(B, g(\bar{x}))$$

with $\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x}))$ only if $d \in \tilde{C}(g, B, \bar{x})$.

PROOF. Any $x \neq \bar{x}$ can be written $x = \bar{x} + \alpha d$ where $d \in X, \alpha > 0$. The feasible point \bar{x} is not a maximum if, and only if, there exist $\bar{\alpha} > 0, d \in X$ such that

$$f(\bar{x} + \alpha d) > f(\bar{x}) \quad \text{for } \alpha \in (0, \bar{\alpha}]$$

and

$$g(\bar{x} + \alpha d) \in B \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

The function f being concave, the first condition is equivalent to $\nabla f(\bar{x})(d) > 0$. If $g(\bar{x}) \in \text{int } B$, by continuity of g , for any d , there exists $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in \text{int } B$ for $\alpha \in (0, \bar{\alpha}]$ and then, \bar{x} is maximal, if and only if, $\nabla f(\bar{x})(d) > 0$ for all $d \neq 0$, that is, $\nabla f(\bar{x}) = 0$.

Let us now consider the case when $g(\bar{x}) \in \partial B$. In the next several pages of proof, we show that the direction d is feasible, that is, there exists a scalar $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$, either when $\nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$ or when

$$\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x})) \quad \text{and} \quad d \in \check{C}(g, B, \bar{x}),$$

but in no other cases. That will allow us to conclude that \bar{x} is not a maximum if and only if one of the following two systems is consistent:

$$(S_1) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$$

or

$$(S_2) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x})), \quad d \in \check{C}(g, B, \bar{x}),$$

which is the desired result.

In order to obtain this result, we must consider three cases.

Case I : $\nabla g(\bar{x})(d) \notin T(B, g(\bar{x}))$,

Case II : $\nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$,

Case III : $\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x}))$.

Case I. $\nabla g(\bar{x})(d) \notin T(B, g(\bar{x}))$

The set $T(B, g(\bar{x}))$ is a closed convex cone. Therefore its complement $T^c(B, g(\bar{x}))$ is an open cone. Then if $\nabla g(\bar{x})(d)$ belongs to $T^c(B, g(\bar{x}))$, there exists an open set V such that $\nabla g(\bar{x})(d) \in V \subset T^c(B, g(\bar{x}))$, that is, there exists an open V such that

$$\lim_{\alpha \rightarrow 0^+} \frac{g(\bar{x} + \alpha d) - g(\bar{x})}{\alpha} \in V \subset T^c(B, g(\bar{x})).$$

Hence there exists $\delta > 0$ small enough, such that for any $\alpha \in (0, \delta)$,

$$\frac{g(\bar{x} + \alpha d) - g(\bar{x})}{\alpha} \in T^c(B, g(\bar{x})).$$

But $T^c(B, g(\bar{x}))$ being a cone, this implies that $g(\bar{x} + \alpha d) - g(\bar{x}) \in T^c(B, g(\bar{x}))$ for $\alpha \in (0, \delta)$. This can also be written $g(\bar{x} + \alpha d) \in g(\bar{x}) + T^c(B, g(\bar{x}))$. By Lemma 2.2,

$$g(\bar{x}) + T^c(B, g(\bar{x})) = (g(\bar{x}) + T(B, g(\bar{x})))^c = T^c(B, g(\bar{x})).$$

Therefore, for $\alpha \in (0, \delta)$, $g(\bar{x} + \alpha d) \in T^c(B, g(\bar{x})) \subset B^c$, that is, $g(\bar{x} + \alpha d)$ does not belong to B . So we can conclude that if $\nabla g(\bar{x})(d) \notin T(B, g(\bar{x}))$, there is no $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$.

Case II. $\nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$

In this case, there exists a neighbourhood N of 0 such that

$$\begin{aligned} \nabla g(\bar{x})(d) + N &\subset \text{int } T(B, g(\bar{x})) = \text{int } S(B, g(\bar{x})) \\ &= \text{int } (S(B, g(\bar{x})) - g(\bar{x})), \quad \text{by Lemma 2.2.} \end{aligned}$$

By definition of the Fréchet-derivative, for any neighbourhood N of 0, there exists $\bar{\alpha} > 0$ small enough such that for $\alpha \in (0, \bar{\alpha}]$,

$$g(\bar{x} + \alpha d) \in g(\bar{x}) + \nabla g(\bar{x})(\alpha d) + \alpha N \in g(\bar{x}) + \text{int } \alpha [S(B, g(\bar{x})) - g(\bar{x})]$$

because $S(B, g(\bar{x}))$ is a cone.

Therefore $g(\bar{x} + \alpha d) = g(\bar{x}) + \alpha \mu (b - g(\bar{x}))$ for some $b \in B$ and $\mu \geq 0$. For α small enough, we can consider that $\alpha \mu$ belongs to $[0, 1]$ and then, by convexity of B ,

$$g(\bar{x} + \alpha d) = (1 - \alpha \mu)g(\bar{x}) + \alpha \mu b \in B \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

Thus, if $\nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$, there exists $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$.

Case III. $\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x}))$

We want to prove that in this case there exists $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$ if, and only if, $d \in \tilde{C}(g, B, \bar{x})$. We shall prove first that if $d \in \tilde{C}(g, B, \bar{x})$ $g(\bar{x} + \alpha d)$ is feasible for $\alpha \in (0, \bar{\alpha}]$ where $\bar{\alpha}$ is some positive scalar. Let I be the subset of $\mathcal{P}(\bar{x})$ such that $\nabla g(\bar{x})(d) \in \bigcap_{i \in I} H_i$ but $\nabla g(\bar{x})(\alpha d) \notin \bigcap_{i \in I'} H_i$ if $I' \supset I$. Let $H = \text{Aff } \bigcap_{i \in I} F_i$. Because $d \in \tilde{C}(g, B, \bar{x})$,

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in - \bigcap_{i \in I} F_i$$

for $\alpha \in (0, \bar{\alpha}]$. Moreover, g is Fréchet-differentiable at \bar{x} , and therefore, for any neighbourhood N of 0, there exists $\alpha = \alpha(N)$ small enough such that for $\alpha \in (0, \alpha(N))$,

$$(1) \quad g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in N \cap H.$$

Also, because $d \in \tilde{C}(g, B, \bar{x})$, $g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri } \bigcap_{i \in I} F_i$. Moreover, $\text{ri } \bigcap_{i \in I} F_i$ is open with respect to H and, therefore, there exists N' such that

$$(2) \quad g(\bar{x}) + \nabla g(\bar{x})(\alpha d) + N' \cap H \subset \text{ri } \bigcap_{i \in I} F_i.$$

By (1) and (2), there exists $\tilde{\alpha}$ small enough such that, for $\alpha \in (0, \tilde{\alpha})$,

$$g(\bar{x} + \alpha d) \in g(\bar{x}) + \nabla g(\bar{x})(\alpha d) + N' \cap H \subset \text{ri } \bigcap_{i \in I} F_i.$$

Hence there exists $\tilde{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$, for $\alpha \in (0, \tilde{\alpha}]$.

Let us now prove the converse. Assume $d \notin \tilde{C}(g, B, \bar{x})$. Now either condition

$$(3) \quad g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri} \bigcap_{i \in I} F_i, \quad \text{for all } \alpha \in (0, \bar{\alpha}]$$

is not satisfied or condition

$$(4) \quad g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in - \bigcap_{i \in I} F_i, \quad \text{for all } \alpha \in (0, \bar{\alpha}]$$

is not satisfied.

Suppose then, that (4) is not satisfied, that is, that there exists no $\bar{\alpha} > 0$ such that

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -F_i = -(B \cap H_i), \quad \text{for some } i \in I, \text{ for } \alpha \in (0, \bar{\alpha}].$$

By B -concavity of g ,

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -B.$$

Therefore, if (4) is not satisfied, for some i , there exists no $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in H_i$ for $\alpha \in (0, \bar{\alpha}]$. However, by definition of I , $\nabla g(\bar{x})(d) \in H_i$ and $g(\bar{x})$ always belongs to H_i , so that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in H_i.$$

Therefore there is no $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in H_i$ for $\alpha \in (0, \bar{\alpha}]$. As H_i separates B from $-B$, $g(\bar{x} + \alpha d)$ belongs to the open half-space determined by H_i and containing $-(B \setminus F_i)$ and not intersecting B . So if (4) is not satisfied, there is no $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$.

Suppose now that (4) is satisfied but (3) is not satisfied, that is,

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in - \bigcap_{i \in I} F_i, \quad \text{for } \alpha \in (0, \bar{\alpha}],$$

but there is no $\bar{\alpha} > 0$ for which

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri} \bigcap_{i \in I} F_i, \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

We shall first note that $\bigcap_{i \in I} F_i$ is the smallest face containing $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$. Indeed, assume that there exists an F_0 such that $0 \notin I$ and $g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in F_0$. Let H_0 be the supporting hyperplane to B at $g(\bar{x})$ such that $F_0 = B \cap H_0$, then $\nabla g(\bar{x})(\alpha d) \in H_0$ but this contradicts the fact that $\nabla g(\bar{x})(\alpha d) \notin \bigcap_{i \in I'} H_i$ if $I' \supset I$. So $\bigcap_{i \in I} F_i$ is the smallest face containing $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$. Thus, if there is no $\bar{\alpha} > 0$ for which (3) is satisfied, there is no $\bar{\alpha}' > 0$ for which

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}'].$$

We shall accept this assertion for the moment and prove it as Lemma 4.4 after the proof of this theorem is completed. Then, in this case, as $g(\bar{x}) \in F_i$ for any $i \in I$, $g(\bar{x} + \alpha d) \notin \bigcap_{i \in I} F_i$. Otherwise, by relation (4),

$$-(g(\bar{x}) + \nabla g(\bar{x})(\alpha d)) \in -\bigcap_{i \in I} F_i - \bigcap_{i \in I} F_i = -\bigcap_{i \in I} F_i,$$

and this would contradict the fact that (3) is not satisfied. However,

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} H_i$$

and, as (4) is satisfied and $F_i = B \cap H_i$,

$$g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} H_i$$

so that $g(\bar{x} + \alpha d) \in \bigcap_{i \in I} H_i$. Therefore, there exists no $\bar{\alpha}$ such that $g(\bar{x} + \alpha d) \in \bigcap_{i \in I} F_i$ while $g(\bar{x} + \alpha d) \in \bigcap_{i \in I} H_i$ for $\alpha \in (0, \bar{\alpha}]$. This implies that there is no $\bar{\alpha} > 0$ such that for $\alpha \in (0, \bar{\alpha}]$, $g(\bar{x} + \alpha d) \in B$. Thus if (3) is not satisfied, there is no $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$. We can now conclude that $d \in \tilde{C}_I(g, B, \bar{x})$ for some I , that is, $d \in \tilde{C}(g, B, \bar{x})$ if, and only if, there exists $\bar{\alpha} > 0$ such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$.

We have thus proved that if \bar{x} is optimal, there is no d such that (\tilde{S}) is satisfied. Cases I, II and III exhaust all the possible alternatives for $\nabla g(\bar{x})(d)$, so that the reciprocal statement is true and Theorem 4.1 is proved.

Let us now prove Lemma 4.4 that was needed in the previous theorem to show that if (3) is not satisfied, there is no $\bar{\alpha}' > 0$ such that $g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} F_i$ for $\alpha \in (0, \bar{\alpha}']$.

LEMMA 4.4. *Let the notation and assumptions be the same as in Case III of Theorem 4.1. Then, if there is no $\bar{\alpha} > 0$ for which*

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri} \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}]$$

then, there is no $\bar{\alpha}' > 0$ for which

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}'].$$

PROOF. Indeed, let us assume that there exists an $\bar{\alpha} > 0$ such that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}]$$

but that there is no $\bar{\alpha} > 0$ such that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri} \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

As $g(\bar{x})$ belongs to $\partial \bigcap_{i \in I} F_i$, that implies that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \partial \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

Let y_1, y_2 be the following points in $\bigcap_{i \in I} F_i$

$$y_1 = g(\bar{x}) + \nabla g(\bar{x})(\alpha_1 d), \quad y_2 = g(\bar{x}) + \nabla g(\bar{x})(\alpha_2 d),$$

where α_1 and α_2 are fixed scalars in $(0, \bar{\alpha}]$.

There exists a relatively open convex subset D of $\bigcap_{i \in I} F_i$ containing both y_1 and y_2 if and only if there is a line segment in $\bigcap_{i \in I} F_i$ having both y_1, y_2 in its relative interior. The line segment L joining $g(\bar{x})$ and $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$ lies in $\bigcap_{i \in I} F_i$ by assumption. The points y_1 and y_2 are in $\text{ri } L$. So, there exists a relatively open convex subset D of $\bigcap_{i \in I} F_i$ containing both these points and therefore containing all points $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$ for $0 < \alpha < \bar{\alpha}$. By Lemma 2.6, this relatively open subset must be contained in the relative interior of a non-empty face F of $\bigcap_{i \in I} F_i$. By Lemma 2.3, F is either equal to $\text{ri } \bigcap_{i \in I} F_i$ or included in $\partial \bigcap_{i \in I} F_i$. But we assumed that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \partial \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

So $F \subset \partial \bigcap_{i \in I} F_i$. It is therefore a proper face of $\bigcap_{i \in I} F_i$ and also a face of B by Lemma 2.5. This contradicts the fact that $\bigcap_{i \in I} F_i$ is the smallest face of B containing $g(\bar{x}) + \nabla g(\bar{x})(\alpha d)$ for $\alpha \in (0, \bar{\alpha}]$, as was shown in the previous theorem. Therefore, we proved that if there is no $\bar{\alpha} > 0$ such that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \text{ri } \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}],$$

there exists an $i \in I$ such that there is no $\bar{\alpha} > 0$ satisfying

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in F_i \quad \text{for } \alpha \in (0, \bar{\alpha}].$$

So, there is certainly no $\bar{\alpha}' > 0$ such that

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in \bigcap_{i \in I} F_i \quad \text{for } \alpha \in (0, \bar{\alpha}'],$$

and Lemma 4.4 is proved.

This completes the proof of the optimality conditions using all the possible intersections of faces of B . It will be shown now that we need consider only one face, namely B^- . And so, the following characterization of optimality is expressed in terms of the generalized cone of constancy $C(g, B, \bar{x})$.

THEOREM 4.2. *Let \bar{x} be a feasible solution of (P). Then \bar{x} is optimal if, and only if, the following system is inconsistent.*

$$(S) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in T(B, g(\bar{x}))$$

with $\nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x}))$ only if $d \in C(g, B, \bar{x})$.

PROOF. By Theorem 4.1, if \bar{x} is optimal, there is no d belonging to any of the $\tilde{C}_I(g, B, \bar{x})$, where I is any subset of $\mathcal{P}(\bar{x})$, such that, also $\nabla f(\bar{x})(d) > 0$ and $\nabla g(\bar{x})(d) \in T(B, g(\bar{x}))$. In particular, there is no such d belonging to $\tilde{C}_{\mathcal{P}^-}(g, B, \bar{x})$. But, as can be proved easily, $\tilde{C}_{\mathcal{P}^-}(g, B, \bar{x}) = C(g, B, \bar{x})$. So, by Theorem 4.1, if x is optimal, there is no d satisfying (S) and the necessary condition is proved.

Suppose now that \bar{x} is not optimal. We shall show that (S) is consistent. By Theorem 4.1, either $\mathcal{P}^- = \emptyset$ and

$$(S_1) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))$$

is consistent, or $\mathcal{P}^- \neq \emptyset$ and

$$(S_2) \quad \nabla f(\bar{x})(d) > 0, \quad \nabla g(\bar{x})(d) \in \partial T(B, g(\bar{x})), \quad d \in \tilde{C}(g, B, \bar{x})$$

is consistent.

If (S₁) is consistent, there is nothing to prove. If (S₂) is consistent, there exists a feasible direction d such that $g(\bar{x} + \alpha d) \in B$ for $\alpha \in (0, \bar{\alpha}]$. By definition of \mathcal{P}^- , $g(\bar{x} + \alpha d) \in B^-$ for $\alpha \in (0, \bar{\alpha}]$. Also, for any feasible direction d ,

$$\nabla g(\bar{x})(\alpha d) = \lim_{\alpha \rightarrow 0} \frac{g(\bar{x} + \alpha d) - g(\bar{x})}{\alpha}$$

belongs to the hyperplane H^- determining B^- , that is, the supporting hyperplane to B at $g(\bar{x})$ such that $B^- = B \cap H^-$. So there exist $\bar{\alpha} > 0$ and d such that, for $\alpha \in (0, \bar{\alpha}]$

$$\nabla f(\bar{x})(d) > 0, \quad g(\bar{x}) + \nabla g(\bar{x})(\alpha d) \in B^-, \quad g(\bar{x} + \alpha d) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d) \in -B^-.$$

By Lemma 4.3, there exists \hat{x} such that $g(\hat{x}) \in \text{ri } B^-$, so that if we denote $d^\lambda = \hat{x} - \bar{x}$, for α sufficiently small,

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d^\lambda) \in \text{ri } B^- \quad \text{and} \quad g(\bar{x} + \alpha d^\lambda) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d^\lambda) \in -B^-.$$

Let $d_\lambda = \lambda d^\lambda + (1 - \lambda) d$ with $\lambda \in (0, 1)$. Then

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d_\lambda) \in \text{ri } B^- \quad \text{for } \alpha \in (0, \bar{\alpha}'],$$

for some positive scalar $\bar{\alpha}'$, and by B -concavity of g

$$g(\bar{x} + \alpha d_\lambda) - g(\bar{x}) - \nabla g(\bar{x})(\alpha d_\lambda) \in -B^-.$$

Hence, $d_\lambda \in C(g, B, \bar{x})$ for $\alpha \in (0, \alpha']$. This implies that $\nabla g(\bar{x})(d_\lambda) \in \partial T(B, g(\bar{x}))$. Indeed, by Lemma 2.2, $T(B, g(\bar{x})) = T(B, g(\bar{x})) + g(\bar{x})$. So to prove that $\nabla g(\bar{x})(d_\lambda)$ or $\nabla g(\bar{x})(\alpha d_\lambda)$, for $\alpha > 0$, belongs to $\partial T(B, g(\bar{x}))$ is equivalent to prove that $g(\bar{x}) + \nabla g(\bar{x})(\alpha d_\lambda)$ belongs to $\partial T(B, g(\bar{x}))$. But, as $d_\lambda \in C(g, B, \bar{x})$,

$$g(\bar{x}) + \nabla g(\bar{x})(\alpha d_\lambda) \in B^-$$

which certainly belongs to the boundary of the smallest cone with vertex $g(\bar{x})$ containing B , that is, $\partial T(B, g(\bar{x}))$. So $\nabla g(\bar{x})(d_\lambda) \in \partial T(B, g(\bar{x}))$.

Moreover,

$$f(\bar{x} + \alpha d_\lambda) \geq (1 - \lambda)f(\bar{x} + \alpha d) + \lambda f(\bar{x} + \alpha d') > f(\bar{x}), \text{ for sufficiently small } \lambda.$$

Therefore d_λ is such that

$$\nabla f(\bar{x})(d_\lambda) > 0, \quad \nabla g(\bar{x})(d_\lambda) \in \partial T(B, g(\bar{x})), \quad d_\lambda \in C(g, B, \bar{x}).$$

The system (S) is consistent and the theorem is proved.

By the Dubovitskii–Milyutin Theorem and the Minkowski–Farkas Lemma (see Girsanov (1970)), the dual characterization of optimality, Theorem 4.3, follows quite directly from Theorem 4.2. The cones to be considered in the proof of the dual theorem are:

$$\begin{aligned} K_1 &= \{d \in X: \nabla f(\bar{x})(d) > 0\}, \\ K_2 &= \{d: \nabla g(\bar{x})(d) \in \text{int } T(B, g(\bar{x}))\}, \\ K_3 &= C(g, B, \bar{x}). \end{aligned}$$

THEOREM 4.3. *Let \bar{x} be a feasible solution of (P). Then \bar{x} is optimal if, and only if, there exists $\mu \in T^+(B, g(\bar{x}))$ such that*

$$\nabla^t f(\bar{x}) + \mu \nabla^t g(\bar{x}) = 0, \quad \text{if } \mathcal{P}^- = \emptyset$$

or

$$\nabla^t f(\bar{x}) \in -C^+(g, B, \bar{x}), \quad \text{if } \mathcal{P}^- \neq \emptyset.$$

Let us illustrate Theorems 4.2 and 4.3 by an example.

EXAMPLE 4.1. Consider the following problem:

$$\begin{aligned} &\text{Maximize } f(x_1, x_2) = -x_1 + x_2 \\ &\text{such that } g(x_1, x_2) = \begin{bmatrix} 1 - x_2 \\ -(1 - x_1)^2 \end{bmatrix} \in B \end{aligned}$$

where $B = \{(y_1, y_2)^t : y_2 \geq 0, y_1 + y_2 \geq 0\}$. We already considered the constraint $g(x_1, x_2) \in B$ in Section 3. Let us recall that the feasible set is

$$A = \{(x_1, x_2)^t : x_1 = 1, x_2 \leq 1\}.$$

The faces of B are

$$F_1 = \{(y_1, y_2)^t : y_2 = 0, y_1 \geq 0\},$$

$$F_2 = \{(y_1, y_2)^t : y_2 \geq 0, y_1 + y_2 = 0\}.$$

Let $\bar{x} = (1, 1)^t$. Then $g(\bar{x}) = (0, 0)^t$, $T(B, g(\bar{x})) = B$ and if $x \in A$, $g(x) \in F_1$, so that $B^* = F_1$. We saw in Example 2.1, that $C(g, B, \bar{x}) = \{(d_1, d_2)^t : d_1 = 0, d_2 < 0\}$. Moreover $\nabla f(\bar{x})(d) = -2d_1 + d_2$, and

$$\nabla g(\bar{x})(d) = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} -d_2 \\ 0 \end{bmatrix} \in \partial T(B, g(\bar{x})) \text{ for all } d,$$

such that $d_2 \leq 0$. The point \bar{x} is optimal and Theorems 4.2 and 4.3 are verified.

These results, Theorems 4.2 and 4.3, also find an immediate application in the computing algorithm called MELP developed in Ben-Tal and Zlobec (1975). The MELP is a feasible direction method used for solving differentiable convex programs with nonlinear concave constraints of the type:

$$(C) \quad \text{Max } f^0(x) \text{ such that } f^k(x) \geq 0, \quad k \in \mathcal{P} = \{1, 2, \dots, p\}.$$

This method is based on the following theorem (Ben-Tal and Zlobec (1975)), requiring the solution of a program with cone-constraints.

THEOREM. *Let $f^k : k \in \{0\} \cup \mathcal{P}$, be differentiable concave functions: $R^n \rightarrow R$ and \bar{x} be a feasible solution of problem (C). For a given subset Ω of $\mathcal{P}(\bar{x})$, let the $(n+1)$ -tuple $(\lambda(\Omega), d(\Omega))$ be an optimal solution of the linear program over a cone*

(C, Ω) Maximize λ such that

$$d^t \nabla f^0(\bar{x}) - \lambda \geq 0; \quad d^t \nabla f^k(\bar{x}) - \lambda \geq 0, \quad k \in \Omega;$$

$$d \in D_{\bar{x}}^-(\bar{x}), \quad k \in \mathcal{P} \setminus \Omega; \quad |d_i| \leq 1, \quad i = 1, \dots, n.$$

Then

- (a) \bar{x} is an optimal solution of (C) if, and only if, for every $\Omega \subset \mathcal{P}(\bar{x})$, $\lambda(\Omega) = 0$;
- (b) a vector $\bar{d} \in R^n$ is a feasible direction of descent at \bar{x} if, and only if, there exist a subset $\bar{\Omega}$ of $\mathcal{P}(\bar{x})$ and $\bar{\lambda} > 0$ such that $(\bar{\lambda}, \bar{d})$ is a feasible solution of (C, $\bar{\Omega}$).

The sets $D_{\bar{x}}^-(\bar{x})$ are the cone of directions of constancy of the function f_k at \bar{x} introduced in Ben-Tal et al. (1976) (the definition of $D_{\bar{x}}^-(\bar{x})$ is recalled in the next section). Hence in order to use the previous theorem, one must solve convex programs (C, Ω) over a cone and our theory could then be applied.

Let us mention, to finish this section, that a saddle point characterization of optimality can also be given, without constraint qualification, for problem (P). Of course, such characterization does not make use of the generalized cone of constancy. Such results can be found in Massam (1977).

5. Reduction to the classical results

Consider the classical problem

(C) Maximize $f(x)$ such that

$$g_i(x) \geq 0, \quad i = 1, 2, \dots, m,$$

where f and $g_i, i = 1, \dots, m$, are concave and Fréchet-differentiable. The functions f and $g_i, i = 1, \dots, m$, map R^n into R . So in this case $B = R^n$. Let A be the feasible set. Let \bar{x} be feasible, that is,

$$\begin{cases} g_j(\bar{x}) = 0, & j = 1, \dots, p, \\ g_j(\bar{x}) > 0, & j = p + 1, \dots, m, \end{cases} \quad \text{where } 1 \leq p \leq m,$$

or $g_j(\bar{x}) > 0, \text{ for } j = 1, \dots, p, \dots, m.$

This last case represents an unconstrained case and therefore is of no interest. Let us consider the first case. The point $g(\bar{x})$ belongs to the supporting hyperplanes $H_j = \{y \in R^m: y_j = 0, j \in \mathcal{P}(\bar{x})\}$, with the following notation

$$\mathcal{P} = \{1, 2, \dots, p, \dots, m\}, \quad \mathcal{P}(\bar{x}) = \{1, \dots, p\},$$

$$\mathcal{P}^- = \{j \in \mathcal{P}: x \in A \text{ implies } g_j(x) = 0\}.$$

The corresponding faces are $F_j = \{y \in R^m: y_j = 0, y_i \geq 0 \text{ for } i \in \mathcal{P}, i \neq j\}$. Further, $T(B, g(\bar{x})) = \{y \in R^m: y_1 \geq 0, \dots, y_p \geq 0; y_{p+1}, \dots, y_m \text{ arbitrary}\}$,

$$\partial T(B, g(\bar{x})) = \bigcup_{i \in \mathcal{P}(\bar{x})} \{y \in R^m: y_i = 0, y_j \geq 0 \text{ for } j \neq i \text{ and } j \in \mathcal{P}(\bar{x}), y_l \text{ arbitrary for } l \notin \mathcal{P}(\bar{x})\}.$$

$$B^- = \bigcap_{j \in \mathcal{P}^-} F_j = \{y \in R^m: y_j = 0 \text{ for } j \in \mathcal{P}^-, y_j \geq 0 \text{ for } j \notin \mathcal{P}^-\}$$

and

$$\text{ri } B^- = \{y \in R^m: y_j = 0 \text{ for } j \in \mathcal{P}^-, y_j > 0 \text{ for } j \notin \mathcal{P}^-\}.$$

Then

$$C(g, B, \bar{x}) = \{d \in R^n : \text{there exists } \bar{\alpha} > 0 \text{ such that for } \alpha \in (0, \bar{\alpha}]$$

$$g_j(\bar{x}) + \nabla g_j(\bar{x})(\alpha d) = 0, j \in \mathcal{P}^-,$$

$$g_j(\bar{x}) + \nabla g_j(\bar{x})(\alpha d) > 0, j \in \mathcal{P}/\mathcal{P}^-,$$

$$g_j(\bar{x} + \alpha d) - g_j(\bar{x}) - \nabla g_j(\bar{x})(\alpha d) = 0, j \in \mathcal{P}^-,$$

$$g_j(\bar{x} + \alpha d) - g_j(\bar{x}) - \nabla g_j(\bar{x})(\alpha d) \leq 0, j \in \mathcal{P}/\mathcal{P}^- \}.$$

When $j \notin \mathcal{P}(\bar{x})$, that is, $g_j(\bar{x}) > 0$, the previous conditions impose no constraint on $\nabla g_j(\bar{x})(d)$ or $g_j(\bar{x} + \alpha d)$. So,

$$C(g, B, \bar{x}) = \{d \in R^n : \text{there exists } \bar{\alpha} > 0 \text{ such that for } \alpha \in (0, \bar{\alpha}]$$

$$\nabla g_j(\bar{x})(d) = 0, j \in \mathcal{P}^-,$$

$$\nabla g_j(\bar{x})(d) > 0, j \in \mathcal{P}(\bar{x})/\mathcal{P}^-,$$

$$g_j(\bar{x} + \alpha d) = g_j(\bar{x}), j \in \mathcal{P}^-,$$

$$0 < g_j(\bar{x} + \alpha d) \leq \nabla g_j(\bar{x})(d), j \in \mathcal{P}(\bar{x})/\mathcal{P}^- \}.$$

Therefore

$$C(g, B, \bar{x}) = \{d \in R^n : \text{there exists } \bar{\alpha} > 0 \text{ such that for } \alpha \in (0, \bar{\alpha}]$$

$$\nabla g_j(\bar{x})(d) > 0, j \in \mathcal{P}(\bar{x})/\mathcal{P}^-, g_j(\bar{x} + \alpha d) = g_j(\bar{x}), j \in \mathcal{P}^- \}$$

$$= \bigcap_{j \in \mathcal{P}^-} D_{g_j}^-(\bar{x}) \cap \{d \in R^n : \nabla g_j(\bar{x})(d) > 0, j \in \mathcal{P}(\bar{x})/\mathcal{P}^-\},$$

where $D_{g_j}^-(\bar{x})$ is the cone of directions of constancy for the function g_j at \bar{x} as introduced in Ben-Tal et al. (1976). Theorems 4.3 and 4.4 reduce to the following characterizations of optimality.

THEOREM 5.1. *Let \bar{x} be a feasible solution of (C), \bar{x} is optimal if, and only if, either (I) or (II) is verified:*

(I) *The system*

$$\nabla f(\bar{x})(d) > 0, \nabla g_j(\bar{x})(d) > 0, j \in \mathcal{P}(\bar{x})/\mathcal{P}^-,$$

$$\nabla g_j(\bar{x})(d) = 0 \text{ only if } d \in \bigcap_{j \in \mathcal{P}^-} D_{g_j}^-(\bar{x})$$

is inconsistent, where

$$D_{g_j}^-(\bar{x}) = \{d \in R^n : \text{there exists } \bar{\alpha} > 0 \text{ such that } g_j(\bar{x} + \alpha d) = g_j(\bar{x}) \text{ for } \alpha \in (0, \bar{\alpha}]\}.$$

(II) There exists $\mu_i \geq 0$, $i \in \mathcal{P}(\bar{x})$, not all zero, such that

$$\nabla f(\bar{x}) + \sum_{i \in \mathcal{P}(\bar{x})/\mathcal{P}^-} \mu_i \nabla g_i(\bar{x}) \in - \left(\bigcap_{j \in \mathcal{P}^-} D_{g_j}^-(\bar{x}) \right)^+, \quad \text{if } \mathcal{P}^- \neq \emptyset,$$

$$\nabla f(\bar{x}) + \sum_{i \in \mathcal{P}(\bar{x})} \mu_i \nabla g_i(\bar{x}) = 0, \quad \text{if } \mathcal{P}^- = \emptyset.$$

These conditions are exactly the primal and dual 'single I characterization' of optimality given by Ben-Tal et al. (1976). A connection with their results is clearly established. When $\mathcal{P}^- \neq \emptyset$, Slater's condition is satisfied and the results obtained are the classical Karush–Kuhn–Tucker results.

REMARKS. 1. In problem (C), the condition $g_j(x) = 0$ only means that $g(x)$ belongs to the face F_j of the cone $B = R_+^m$ and therefore the notion of face is just as important in the study of (C) as in the study of (P).

2. Before giving the characterizations involving only one subset of \mathcal{P} , namely the subset \mathcal{P}^- , Ben-Tal et al. (1976) gave characterizations where every subset I of $\mathcal{P}(\bar{x})$ was involved. This primal and therefore also dual characterization could be directly deduced from Theorem 4.2 given in terms of the cone $\tilde{C}(g, B, \bar{x})$.

References

- V. I. Averbukh and O. A. Smolyanov (1968), 'The various definitions of the derivative in linear topological spaces', *Russian Math. Surveys* **23**, 67–114.
- A. Ben-Israel (1969), 'Linear equations and inequalities on finite dimensional, real or complex, vector spaces: A unified theory', *J. Math. Anal. Appl.* **27**, 367–389.
- A. Ben-Israel, A. Ben-Tal and S. Zlobec (1976), 'Optimality conditions in convex programming' presented at the IX International Symposium on Mathematical Programming, Budapest, Hungary.
- A. Ben-Tal, A. Ben-Israel, and S. Zlobec (1976), 'Characterization of optimality in convex programming without a constraint qualification', *J. Optimization Theory Appl.* **20**, 417–437.
- A. Ben-Tal and S. Zlobec (1975), 'A new class of feasible direction methods', Research Report CCS 216, Center for Cybernetic Studies, The University of Texas, Austin, Texas.
- B. D. Craven (1974), 'A transposition theorem, applied to convex programming in Banach spaces', Mathematics Department, University of Melbourne, Australia, 1974.
- B. D. Craven (1977), 'Lagrangian condition and quasiduality', *Bull. Austral. Math. Soc.* **16**, 325–339.
- I. V. Girsanov (1970), *Lectures on mathematical theory of extremum problems*, No. 67 (Springer-Verlag, Berlin).
- M. Guignard (1969), 'Generalized Kuhn–Tucker conditions for mathematical programming problems in Banach spaces', *SIAM J. Control Optimization* **7**, 232–241.

- R. B. Holmes (1972), *A course on optimization and best approximation*, Lecture notes in mathematics, No. 257 (Springer-Verlag, Berlin).
- H. W. Kuhn and A. W. Tucker (1951), 'Nonlinear programming', *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman, University of California, Berkeley, California.
- O. L. Mangasarian (1969), *Nonlinear programming* (McGraw-Hill, New York).
- H. M. Massam (1977), *Convex programming with cones* (Ph.D. Thesis, McGill University, Montreal, Quebec, Canada).
- R. T. Rockafellar (1970), *Convex analysis* (Princeton University Press, Princeton, New Jersey).

Department of Mathematics
University of Toronto, Toronto
Canada