CONE-CONVEX PROGRAMMING: STABILITY AND AFFINE CONSTRAINT FUNCTIONS

J.M. Borwein* and H. Wolkowicz†

We present strengthened optimality conditions for the abstract convex program. In particular, we consider the special case when the range space of the constraint is finite dimensional and when the constraint function is affine. Applications include a sensitivity theorem and a generalization of Farkas' Lemma.

1. INTRODUCTION

In this paper we consider both: the general abstract convex program

$$(P_g) \quad \mu = \inf\{p(x) : g(x) \in -S,\ x \in \Omega\}, \tag{1}$$

where $p: X \to R$ is a convex functional, $g: X \to Y$ is an S-convex function, S is a convex cone in Y and Ω is a convex set in X; and the special abstract convex program

$$(P_A) \quad \mu = \inf\{p(x) : Ax \in -S,\ x \in \Omega\}, \tag{2}$$

where now $Ax = Lx - b$ is an affine operator and Ω is a polyhedral set, i.e., L is a linear operator, b is a vector and Ω is a finite intersection of closed half spaces.

Such programs arise in several situations. For example, the semi-infinite program

$$\mu = \inf\{p(x) : h(x,t) \leq 0, \text{ for all t in T and x in } \Omega\}$$

can be written in the form (P_g) if we set $g(x) = h(x,\cdot)$, Y some subspace of R^T and S the cone of nonnegative functions in Y. If T is compact, $h(\cdot,t)$ convex for each t and $h(x,\cdot)$ continuous for each x, then we can choose $Y = C[T]$, the continuous functions on T. The Lagrange multipliers are then taken from the dual space of Borel measures on T, see e.g. Borwein and Wolkowicz (1979b).

---

*Research partially supported by NSERC A4493.

†Research partially supported by NSERC A3388.

ISBN 0-12-621120-5

Optimal control problems with linear dynamics, an initial condition, a final target set and a control set can be rewritten in the form (P_A). The constraint $Ax \in -S$ represents the state hitting the target for a given control x, see e.g. Luenberger (1969).

Many problems in stability of differential equations can be phrased in terms of the feasibility of an operator equation in the space of linear operators endowed with the symmetric ordering, i.e. the ordering induced by the cone of positive semidefinite operators, see e.g. Berman (1973).

Characterizations of optimality for (P_g) with constraint qualification have been given by many authors, see e.g. Luenberger (1969). The usual constraint qualification used is Slater's condition, i.e. there exists a feasible point $\hat{x}$ whose image under g is in the topological interior of -S. This condition has been weakened by Craven and Zlobec (1980) to require only nonempty relative algebraic interior of the feasible set and nonempty interior of the cone S, while a characterization of optimality for (P_g) without any constraint qualification has been given by Borwein and Wolkowicz (1979b).

Massam (1979) considered (P_g) in the special case that Y is finite dimensional. She used the "minimal exposed face" of -S which contained the image of the feasible set to get optimality conditions which held when the minimal exposed face actually coincided with the minimal face. This was pointed out and strengthened in Borwein and Wolkowicz (1981a). In Borwein and Wolkowicz (1981b) we presented a Lagrange multiplier theorem for (P_g) which holds without any constraint qualification. This result differs from the standard Lagrange multiplier theorem in two ways. First the Lagrange multiplier is chosen from the dual cone of the minimal face (or "minimal cone") of (P_g), denoted S^f, rather than the (smaller) dual cone of S; second, the Lagrangian is restricted to $x \in \Omega \cap g^{-1}(S^f - S^f)$.

In Borwein and Wolkowicz (1979a) we presented an algorithm which "regularizes" (P_g) by finding S^f and $g^{-1}(S^f-S^f)$.

In this paper we restrict ourselves to the special case when Y is finite dimensional. We first recall and strengthen the Lagrange multiplier characterization of optimality for (P_g) given in Borwein and Wolkowicz (1981b). This is then applied to the program (P_A) to yield a Lagrange multiplier result which differs from the standard case only in that the multiplier is chosen from a larger dual cone. This differs from previous results where the multiplier is chosen from a larger dual cone *and* the variable x is restricted to a smaller set than Ω.

These results are then applied to extend the sensitivity theorem for (P_g) given in Luenberger (1969) and derive a generalization of the Farkas' lemma given in Ben-Israel (1969). These results hold without any constraint qualification.

2. PRELIMINARIES

Let us first consider the *general abstract convex program*

$$(P_g) \quad \mu = \inf\{p(x): g(x) \in -S,\ x \in \Omega\}$$

where $p: X \to R$ is a continuous convex functional (on Ω); $g: X \to R^m$ is a continuous S-*convex function* (on Ω), i.e.

$$tg(x) + (1-t)g(y) - g(tx+(1-t)y) \in S$$

for all $0 \leq t \leq 1$ and x, y in Ω; S is a convex cone, i.e. $S + S \subset S$ and $tS \subset S$ for all $t \geq 0$; Ω is a convex set in X (not necessarily polyhedral); and X is a locally convex (Hausdorff) space. The *feasible set* of (P_g) is

$$F = \Omega \cap g^{-1}(-S). \tag{3}$$

We assume throughout that

$$F \neq \emptyset. \tag{4}$$

The *polar cone* of a set K in R^m is

$$K^+ = \{\phi \in R^m: \phi x \geq 0 \text{ for all } x \text{ in } K\}, \tag{5}$$

where ϕx denotes the inner product in R^m. The *annihilator* of a set K is

$$K^{\perp} = K^{+} \cap -K^{+}.$$

Recall that for two closed convex cones K and L,

$$(K \cap L)^{+} = \overline{K^{+} + L^{+}}\ , \tag{6}$$

where $\bar{\cdot}$ denotes closure, e.g. Luenberger (1969).

We will need to use the smallest face of S which contains $-g(F)$.

Definition 1: (i) K is a *face* of a convex cone S if K is a convex cone and

$$x,\ y \in S,\ x + y \in K \ \text{ implies } \ x,\ y \in K. \tag{7}$$

(ii) The *minimal cone* for (P_g), denoted S^f, is defined to be the smallest face of S which contains $-g(F)$. (Note that S^f is the intersection of all faces of S which contain $-g(F)$).

The following characterization of optimality for (P_g) was given in Borwein and Wolkowicz (1981b) Theorem 4.1.

THEOREM 1: For the program (P_g),

$$\mu = \inf\{p(x) + \lambda g(x) : x \in F^f\} \tag{8}$$

for some λ *in* $(S^f)^+$ *and* $F^f = \Omega \cap g^{-1}(S^f - S^f)$. *In addition, if* $\mu = p(a)$ *with* a *in* F, *then*

$$\lambda g(a) = 0 \qquad \text{(complementary slackness)} \tag{9}$$

and (8) and (9) characterize optimality of a *in* F. □

To prove the above result one shows that

$$F^f = \Omega \cap g^{-1}(S^f - S^f) = \Omega \cap g^{-1}(S^f - S), \tag{10}$$

$$g(F) \cap -\mathrm{ri}S^f \neq \emptyset \tag{11}$$

where ri denotes *relative interior*, and that

$$g \ \text{ is } \ S^f\text{ - convex on } F^f.$$

Then, the standard Lagrange multiplier theorem is applied to the equivalent program

$$\mu = \inf\{p(x) : g(x) \in -S^f,\ x \in F^f\}.$$

The details are given in Borwein and Wolkowicz (1981b). Note that by the standard Lagrange multiplier theorem we mean Theorem 1 with $F^f = \Omega$ and $(S^f)^+$ replaced by S^+, see e.g. Luenberger (1969). This theorem requires that Slater's condition hold, i.e. that

$$\Omega \cap g^{-1}(-\text{int } S) \neq \emptyset , \tag{12}$$

where int denotes interior.

We will now present a version of Theorem 1 which is stronger in the sense that the Lagrange multiplier relation holds on a larger set. A similar result was given in Borwein and Wolkowicz (1979a), Theorem 6.2 and Remark 6.3. This result will be applied in Section 3. First we need the following lemma. Note that a *polyhedral function* is the maximum of a finite number of affine functions.

LEMMA 1: Consider the ordinary convex program

$$(P_o) \quad \mu = \inf\{p(x): g^k(x) \leq 0, \ k = 1,\ldots,m, \ x \in V\}$$

where $g^k: X \to R$, $k = 1,\ldots,m$, *are continuous convex functionals* (on V) *and* V *is a polyhedral set in* X. *Suppose that there exists an* $\hat{x} \in V$ *such that* $g^k(\hat{x}) \leq 0$, $k = 1,\ldots,m$, *with strict inequality if* g^k *is not polyhedral on* V. *Then*

$$\mu = \inf\{p(x) + \lambda g(x): x \in V\} \tag{13}$$

for some $\lambda = (\lambda_k) \geq 0$. *Moreover, if* $\mu = p(a)$ *for some* $a \in F$, *then*

$$\lambda g(a) = 0 \tag{14}$$

and (13) and (14) characterize optimality of a *in* F.

PROOF. Let $I = \{k: g^k \text{ is polyhedral}\}$, $J = \{1,\ldots,m\}\setminus I$ and

$$W = \{x \in X: g^k(x) \leq 0, \text{ for all k in I}\}.$$

Then (P_o) is equivalent to the program

$$\mu = \inf\{p(x): g^k(x) \leq 0, \text{ for all k in J}, \ x \in V \cap W\}. \tag{15}$$

The point $\hat{x}$ now satisfies Slater's condition for this program. Therefore there exist nonnegative scalars α_k, k in J, such that

$$\mu = \inf\{p(x) + \sum_{k\in J} \alpha_k g^k(x): x \in V \cap W\}.$$

Now since V is polyhedral,

$$V = \{x \in X: \psi^k(x) = \phi^k x - b_k \leq 0, \ k = 1,\ldots,t$$

for some ϕ^k in X*, the dual of X, and b_k in R. Moreover, for each k in I, $g^k(x) = \max_{i\in I^k} \eta_i^k(x)$ for some affine functions η_i^k and *finite* index sets I^k.

Therefore

$$\mu = \inf\{p(x) + \sum_{k\in J} \alpha_k g^k(x): \ \eta_i^k(x) \leq 0, \ i \text{ in } I^k, \ k \text{ in } I,$$
$$\psi^k(x) \leq 0, \ k = 1,\ldots,t\}.$$

This program is linearly constrained which implies that there exist further nonnegative scalars α_i^k, $i \in I^k$, k in I, and a nonnegative vector λ in R^t such that

$$\mu = \inf\{p(x) + \Sigma\alpha_i^k\eta_i^k(x) + \lambda\psi(x)\} \leq \inf\{p(x) + \alpha g(x) + \lambda\psi(x)\}$$

where $\alpha_k = \sum_{i\in I^k} \alpha_i^k$, $\alpha = (\alpha_k)$ and $\psi = (\psi^k)$. Since $\lambda\psi(x) \leq 0$ for all x in V,

$$\mu \leq \inf\{p(x) + \alpha g(x): \ x \in V\}. \tag{16}$$

The reverse inequality follows from the definition of (P_o), since $\alpha g(x) \leq 0$ for all feasible x. This proves (13). The rest of the proof, i.e. complementary slackness and sufficiency, is standard. □

If $\hat{x} \in \text{ri } V$, then the above lemma holds with V convex but not necessarily polyhedral, see Rockafellar (1970a, Theorem 28.1).

We now present the strengthened version of Theorem 1.

THEOREM 2: Suppose that

$$g \ \text{ is } \ S^f\text{-convex on } F^H, \tag{17}$$

where

$$F \subset F^H \subset \Omega\ . \tag{18}$$

Moreover, suppose that F^H *is polyhedral. Then Theorem 1 holds with* F^f *replaced by* F^H.

PROOF. We can rewrite (P_g) as the equivalent program

$$(P_H) \quad \mu = \inf\{p(x)\colon g(x) \in -S^f,\ x \in F^H\}.$$

By (11) and (18), a generalized Slater's condition holds for this program and thus the standard Lagrange multiplier theorem holds. Let us show this by applying Theorem 1 and Lemma 1. Now, by Theorem 1 applied to (P_H), we get

$$\mu = \inf\{p(x) + \bar{\lambda}g(x)\colon x \in F^H \cap g^{-1}(S^f - S^f)\} \tag{19}$$

for some $\bar{\lambda}$ in $(S^f)^+$. Since $Y = R^m$ is finite dimensional, we can find ϕ_i, $i = 1,\ldots,t$, in $(S^f)^\perp$ such that

$$(S^f - S^f) = \bigcap_{i=1}^{t} \{\phi_i\}^+.$$

Thus $x \in g^{-1}(S^f - S^f)$ if and only if $\phi_i g(x) \le 0$, $i = 1,\ldots,t$, and (19) is equivalent to

$$\mu = \inf\{p(x) + \bar{\lambda}g(x)\colon \phi_i g(x) \le 0,\ i = 1,\ldots,t,\ x \in F^H\}. \tag{20}$$

Since g is S^f-convex on F^H and $\{\phi_i\} \subset (S^f)^\perp \subset (S^f)^+$, we conclude that both $\phi_i g$ and $-\phi_i g$ are convex (on F^H), which in turn implies that

$$\phi_i g \text{ is affine (on } F^H),\ i = 1,\ldots,t$$

and, without loss of generality, we can assume them affine on X. Now since F^H is polyhedral, we can apply Lemma 1 to (20). Thus, there exist nonnegative scalars α_i such that

$$\mu = \inf\{p(x) + \bar{\lambda}g(x) + \sum_{i=1}^{t} \alpha_i\phi_i g(x)\colon x \in F^H\}. \tag{21}$$

We can now let

$$\lambda = \bar{\lambda} + \sum_{i=1}^{t} \alpha_i\phi_i .$$

Then clearly λ is in $(S^f)^+$ and this shows that (8) holds with F^f replaced by F^H. That (9) holds follows, for if $a \in F$, then

$$\lambda g(a) = \bar{\lambda}g(a) \tag{22}$$

since $\sum_{i=1}^{t} \alpha_i \phi_i$ is in $(S^f)^\perp$. The sufficiency of (8) and (9) is clear. □

The above two theorems differ from the standard case in two ways. First, x is restricted to a set smaller than Ω while the multiplier λ is chosen from the larger cone $(S^f)^+$ rather than S^+. Theorem 2 shows how to weaken the restriction on x. This strengthens the optimality conditions. We now show that we can further strengthen the conditions by replacing $(S^f)^+$ with a smaller cone.

COROLLARY 1: Suppose that in Theorem 2 we have

$$S^f \subset L \cap -H; \; F^H = \Omega \cap g^{-1}(H). \tag{23}$$

Then Theorem 2 holds with L^+ *replacing* $(S^f)^+$ *if*

$$L^+ - H^+ = (S^f)^+ \tag{24}$$

or equivalently, when equality holds in (23) and H *and* L *are closed convex cones, if*

$$L^+ - H^+ \text{ is closed.} \tag{25}$$

Moreover, if equality holds in (23) and H *is a subspace, then the above condition (24) or (26) is also necessary (independent of the particular functions* p *and* g*). In particular, we can choose* $H = S^f - S^f$ *and* $L = S$.

PROOF. Note that when equality holds in (23) then (24) and (25) are equivalent since, by (6)

$$(S^f)^+ = (L \cap -H)^+ = \overline{L^+ - H^+},$$

when H and L are closed convex cones. Now, if (24) holds and λ satisfies (8) and (9) (with F^H instead of F^f), then one can solve $\lambda = \phi - h$ with ϕ in L^+ and h in H^+. Hence for any x in F^H we get

$$\begin{aligned} \lambda g(x) &= \phi g(x) - h g(x) \\ &\le \phi g(x), \end{aligned} \tag{26}$$

since $g(x) \in H$ when $x \in F^H$. Thus

$$\begin{aligned} \mu &= \inf\{p(x) + \lambda g(x) : x \in F^H\} \\ &\le \inf\{p(x) + \phi g(x) : x \in F^H\}. \end{aligned} \tag{27}$$

The reverse inequality holds also. For, $\phi g(x) \leq 0$, for all $x \in F \subset F^H$, since $g(F) \subset -S^f \subset -L$. The complementary slackness condition is proved in the usual way, i.e. if $\mu = p(a)$ with $a \in F$, then

$$\begin{aligned} p(a) &= \mu \\ &= \inf\{p(x) + \phi g(x) : x \in F^H\} \\ &\leq p(a) + \phi g(a), \text{ since } a \in F \subset F^H \\ &\leq p(a), \qquad \text{since } g(F) \subset -S^f \subset -L. \end{aligned}$$

Conversely, suppose ϕ lies in $(S^f)^+$, equality holds in (23) and H is a subspace. Assuming that Theorem 2 holds with L^+ replacing $(S^f)^+$, we need to show that (24) holds. Let P be the orthogonal projection on H. Consider the program

$$\mu = \inf\{\phi P(y) : -Py \in - \text{ cone } L,\ y \in R^m\}$$

where cone L denotes the convex cone generated by L. Then pp^{-1} (cone L) $\subset pp^{-1}(L \cap -H) = S^f$ so $\mu = 0$. Also $-P^{-1}(H) = R^m$ so that Theorem 2 yields

$$0 = \mu = \inf\{\phi Py + \lambda(-Py) : y \in R^m\}.$$

Since we now assume that $\lambda \in L^+$, we get $\phi P - \lambda P = 0$ and

$$\begin{aligned} \phi &= \phi - (\phi P - \lambda P) \\ &= (\phi - \lambda)(I-P) + \lambda \in -H^+ + L^+. \end{aligned}$$ □

Remark 1. As an example of the above theorem, let us consider the case $S = R^m_+$, i.e. (P_g) is now equivalent to the ordinary finite dimensional convex program with convex (on Ω) constraints $g^k(x) \leq 0$, $k \in P = \{1,\ldots,m\}$, $x \in \Omega$. Let $P^= = \{k \in P: g^k(x) = 0,$ for all $x \in F\}$ be the minimal indexing set of binding constraints, e.g. Abrams and Kerzner (1978). Then it is easy to see that

$$S^f = \{y = (y_k) \in R^m_+ : y_k = 0, \text{ for all } k \in P^=\}$$

is the minimal cone of (P_g) and Theorem 1 yields

$$\mu = \inf\{p(x) + \lambda g(x) : x \in \Omega \text{ and } g^k(x) = 0, \text{ for all } k \in P^=\},$$

for some $\lambda = (\lambda_k)$ in $(S^f)^+$, i.e. $\lambda_k \geq 0$ for $k \in P \setminus P^=$, λ_k arbitrary for $k \in P^=$. This result extends the characterization of optimality for the ordinary convex program given in e.g. Abrams and Kerzner (1978), Ben-Israel et al. (1976) and Ben-Tal and Ben-Israel (1976). The extension is in the sense that the infimum may be unattained and the constraint $x \in \Omega$ is included. (Note that one might also be able to use the approach of Abrams (1975) for problems with unattained infima.) Since S is polyhedral, Corollary 1 implies we can assume that $\lambda \in S^+$, i.e. $\lambda_k \geq 0$ for all k in P.

Now suppose that Ω is polyhedral and the constraints g^k, $k \in P^=$, are analytic convex (or only piecewise faithfully convex, see Ben-Israel et al. (1976) and Rockafellar (1970b)) on Ω. In Theorem 2 set $H = \{\alpha\}^\perp$, where $\alpha = (\alpha_k)$ with $\alpha_k > 0$ if $k \in P^=$ and g^k is not polyhedral (on Ω), $\alpha_k \geq 0$ if $k \in P^=$ and g^k is polyhedral (on Ω) and $\alpha_k = 0$ otherwise. Then

$$F^H = \Omega \cap g^{-1}(H)$$

$$= \Omega \cap \{x \in X: \sum_{k \in P^=} \alpha_k g^k(x) = 0\}.$$

Let x be any feasible point for (P) and

$$D_h = \{d \in X: \text{ there exists } \bar{\alpha} > 0 \text{ with } h(x+\alpha d) = h(x), \text{ for all } 0 < \alpha \leq \bar{\alpha}\}$$

be the cone of directions of constancy at x of h, e.g. Ben-Israel et al. (1976), where $h = \Sigma \alpha_k g^k$. Then D_h is a subspace (or polyhedral cone) independent of the point x (see e.g. Ben-Tal and Ben-Israel (1979) for X finite dimensional and Wolkowicz (1980a) for the general case). Note that Theorem 2 holds with F^H replaced by any polyhedral set G such that

$$F \subset G \subset F^H$$

since the program (P_H) remains equivalent to the original program (P_g). We now set

$$G = \Omega \cap (\hat{x}+D_h) \ .$$

To apply Theorem 2 with F^H replaced by G, we need only show that g is

S^f-convex on G. Now if $x,y \in G$ and $0 \le \lambda \le 1$, then $\lambda x + (1-\lambda)y \in G$, since D_h is a subspace and so convex, and

$$g^k(\lambda x+(1-\lambda)y) - \lambda g^k(x) - (1-\lambda)g^k(y) \le 0 \qquad (28)$$

for all $k \in P$, since the constraints g^k are convex on Ω. To show that g is S^f-convex on G we need to show that equality holds in (28) for all $k \in P^=$. In fact, it is sufficient to show that g^k is affine on G for all $k \in P^=$. Now suppose that $k_o \in P^=$ and g^{k_o} is not affine (on Ω), then $\alpha_{k_o} > 0$, $h(x) = h(\hat{x}) = 0$ for all x in G and so

$$-\alpha_{k_o} g^{k_o}(x) = \sum_{k \in P \setminus \{k_o\}} \alpha_k g^k(x)$$

for all $x \in G$. Thus $-g^{k_o}$ as well as g^{k_o} are convex functions on G which implies that g^{k_o} is affine on G. Theorem 2 now yields

$$\mu = \inf\{p(x) + \lambda g(x): \ x \in G\}$$

for some λ in $(S^f)^+$. Moreover, by Corollary 1 we can assume that $\lambda \in S^+$, i.e. $\lambda \ge 0$. This result yields the optimality conditions given in Wolkowicz (1980b). Note that, if $\{x \in X: \ h(x) = 0\}$ is convex, then $F^H = G$.

The above set G can be found computationally in the analytic convex (or faithfully convex) case by calculating the cones of directions of constancy of the appropriate functions, see Wolkowicz (1978).

In the case that Slater's condition (12) holds, the above theorems reduce to the standard Lagrange multiplier theorem (see e.g. Luenberger (1969)) i.e. in this case $g^{-1}(S^f - S^f)$ becomes all of X, while $(S^f)^+$ becomes S^+.

Note that A is S-convex (on Ω) since it is affine (on Ω). Thus the above preliminaries all hold for the program (P_A), see (2).

Remark 2. We can extend the functions $p: \ X \to R \cup \{\infty\}$ and $g: \ X \to R^m \cup \{\infty\}$ and remove the continuity assumptions on p and g.

Theorem 1 (see Wolkowicz and Borwein (1981b)) would then require that dom $p \supset F$, where dom p denotes the essential domain of p. Lemma 1 needs $x \in$ dom p and p continuous at some feasible point and g continuous on V. Theorem 2 needs dom $p \supset F$ and $p(x) + \bar{\lambda}g(x)$ continuous at some point in $F^H \cap g^{-1}(S^f - S^f)$.

3. A LAGRANGE MULTIPLIER THEOREM FOR (P_A)

If the cone S is polyhedral, then the program (P_A) is linearly constrained and the standard Lagrange multiplier theorem always holds. (Recall that Ω is a polyhedral set in (P_A).) However, we now see that even when S is an arbitrary convex cone, then program (P_A) allows a Lagrange multiplier theorem which differs from the standard case only in that the multiplier λ is chosen from a larger cone containing S^+, namely from $(S^f)^+$. Moreover, when A is a linear operator and Ω is a subspace, we get a geometric weakest constraint qualification.

THEOREM 3: For the program (P_A),

$$\mu = \inf\{p(x) + \lambda Ax : x \in \Omega\}, \tag{29}$$

for some λ *in* $(S^f)^+$. *In addition, if* $\mu = p(a)$ *with* a *in* F, *then*

$$\lambda g(a) = 0 \tag{30}$$

and (29) and (30) characterize optimality of a *in* F.

PROOF: By Theorem 1, we get that

$$\mu = \inf\{p(x) + \bar{\lambda}Ax : x \in \Omega \cap A^{-1}(S^f - S^f)\} \tag{31}$$

for some $\bar{\lambda}$ in $(S^f)^+$ and if $\mu = p(a)$ with a in F, then

$$\bar{\lambda}g(a) = 0 \tag{32}$$

and (31) and (32) characterize optimality of a in F. Now since $Y = R^m$ is finite dimensional, we see that

$$A^{-1}(S^f - S^f) = \{x \in X : \phi_i Ax = 0,\ i = 1,\dots,k\}, \tag{33}$$

for some $\phi_i \in (S^f - S^f)^\perp = (S^f)^\perp$ with

$$S^f - S^f = \bigcap_{i=1}^{k} \{\phi_i\}^{\perp}. \tag{34}$$

Lemma 1 now implies that

$$\mu = \inf\{p(x) + \bar{\lambda}Ax + \sum_{i=1}^{k} \alpha_i\phi_i Ax\colon x \in \Omega\}, \tag{35}$$

for some α_i in R_+. We now let

$$\lambda = \bar{\lambda} + \sum_{i=1}^{k} \alpha_i\phi_i . \tag{36}$$

Since the functionals ϕ_i are in (S^f), we get that

$$\lambda \in (S^f)^+ \tag{37}$$

and moreover, since $g(a) \in g(F) \subset -S^f$,

$$\lambda g(a) = \bar{\lambda}g(a). \tag{38}$$

The conclusion now follows by substituting λ into (35) and noting that the program (35) is equivalent to (31) and that, by (38), $\lambda g(a) = 0$ if and only if $\bar{\lambda}g(a) = 0$. □

Remark 3. In (31) we can assume $\bar{\lambda}$ is in S^+ rather than $(S^f)^+$ if

$$S^+ + (S^f)^{\perp} = (S^f)^+ \tag{39}$$

(see Corollary 1). Thus if

$$A(\Omega) \subset (S^f - S^f) \tag{40}$$

then $A^{-1}(S^f - S^f)$ is redundant in (31) and (29) holds with λ in S^+. Thus (39) and (40) is a constraint qualification for (P_A).

We can further strengthen the above theorem by using Theorem 2. In fact, in this case we see that regularity of the problem, in the sense that the standard Lagrange multiplier theorem holds, independent of p, depends solely on the condition (39), i.e. solely on the geometry of the cone S.

THEOREM 4: First, for the program (P_A), there exists a polyhedral cone H in R^m such that

$$S^f = S \cap -H \text{ and } A(\Omega) \subset H. \tag{41}$$

Now if K *and* H *in* R^m *satisfy*

$$\text{(i)}\quad S^f \subset K \cap -H \text{ and } \quad \text{(ii)}\quad A(\Omega) \subset H \tag{42}$$

with H *polyhedral, then Theorem 3 holds with* $(S^f)^+$ *replaced by* K^+, *if*

$$K^+ - H^+ = (S^f)^+, \tag{43}$$

or equivalently, when equality holds in (42)(i) and H *and* K *are closed convex cones, if*

$$K^+ - H^+ \text{ is closed.} \tag{44}$$

Moreover, if H *is a subspace and equality holds in (42)(i), then the above conditions are also necessary for* K^+ *to replace* $(S^f)^+$.

PROOF: If $S^f = S$, we can choose $H = R^m$. If S^f is a proper face of S, then $A(\Omega) \cap -\text{ri } S = \emptyset$ and the Hahn-Banach Theorem implies that there exists $\phi_1 \in S^+$ such that

$$\phi_1(A(\Omega)) \geq 0, \quad \phi_1(\text{ri } S) > 0. \tag{45}$$

This implies that $-S^f \subset -S \cap \{\phi_1\}^+$. If equality holds, then we can set $H = \{\phi_1\}^+$. If not, then we repeat the same process but with the cone $-S \cap \{\phi_1\}^+$ replacing the cone $-S$. Since R^m is finite dimensional this process must stop in a finite number of steps. We then set

$$H = \bigcap_{i=1}^{t} \{\phi_i\}^+ .$$

This H then satisfies (41).

Now suppose that (42) holds. Since A is affine, we see that A is S^f-convex on the polyhedral set

$$F^H = \Omega \cap A^{-1}(H) \tag{46}$$

and Theorem 2 implies that we can replace F^f by F^H in (8), i.e.

$$\mu = \inf\{p(x) + \lambda Ax: \; x \in \Omega \cap A^{-1}(H)\}, \tag{47}$$

for some λ in $(S^f)^+$. By Corollary 1, we see exactly when $(S^f)^+$ can be replaced by K^+, but with $x \in \Omega \cap A^{-1}(H)$ rather than $x \in \Omega$ as desired. But since $A(\Omega) \subset H$, we see that $x \in \Omega$ if and only if $x \in \Omega \cap A^{-1}(H)$. □

Slater's condition (12) is a *constraint qualification* for (P_g), i.e. it is a sufficient condition which guarantees that the standard Lagrange multiplier theorem holds for (P_g), independent of the objective function p. A *weakest constraint qualification* is a necessary and sufficient constraint qualification. Weakest constraint qualifications for the ordinary convex program, i.e. for the program (P_g) with S polyhedral and $\Omega = X$, have been given in Wolkowicz (1980a). The nonconvex case is treated in Gould and Tolle (1972). See also Bazaraa et al. (1976). We now see that the above theorem yields a very elegant weakest constraint qualification for (P).

COROLLARY 2: Consider the program (P_A) *when* A *is linear and* Ω *is a subspace. Then there exists a subspace* H *satisfying (41) and moreover, the condition*

$$S^+ - H^+ = (S^f)^+ \tag{48}$$

or equivalently, if S *is closed,*

$$S^+ - H^+ \text{ is closed} \tag{49}$$

is a weakest constraint qualification.

PROOF: Since $A(\Omega)$ is now a subspace, let us choose H to be any subspace satisfying (41), e.g. $H = A(\Omega) + S^f - S^f$. The result now follows from Theorem 4. □

4. STABILITY AND A GENERALIZATION OF FARKAS' LEMMA

We now apply the results in the previous section. First, we have the following sensitivity theorem for a perturbed program $(P_{g,\varepsilon})$, when the perturbation ε is restricted to $S^f - S$.

THEOREM 5: Suppose that μ *and* $\mu(\varepsilon)$ *are the optimal values of the program* (P_g) *and the perturbed program*

$$(P_{g,\varepsilon}) \quad \begin{array}{l} \textit{minimize} \ \ p(x) \\ \textit{subject to} \ \ g(x) - \varepsilon \in -S \ \textit{and} \ x \in \Omega \end{array}$$

respectively, where $\varepsilon \in S^f - S$ *(resp.* $\in S^f_\varepsilon - S$*). Suppose that* λ *and* λ_ε *are the (restricted) Lagrange multipliers for* (P_g) *and* $(P_{g,\varepsilon})$ *respectively, found using Theorem 1. Then*

$$\mu - \mu(\varepsilon) \leq \lambda\varepsilon \ (\text{resp} \ \geq -\lambda_\varepsilon \varepsilon). \tag{50}$$

PROOF: Suppose that S^f_ε is the minimal cone for the perturbed program $(P_{g,\varepsilon})$ with $\varepsilon \in S^f - S$. Let us first show that $S^f_\varepsilon \subset S^f$. Suppose not, i.e. suppose that there exists $x \in \Omega$ such that $g(x) - \varepsilon = -s \in -S$ but $s \notin S^f$. Then $g(x) = -s + \varepsilon \in S^f - S$ but $-s + \varepsilon \notin S^f - S^f$, since $(S-S^f) \cap (S^f-S) = S^f - S^f$. Now by (11), there exists $\hat{x} \in \Omega$ such that $g(\hat{x}) = -\hat{s} \in -\text{ri } S^f$. Thus for $0 < \lambda < 1$ and λ sufficiently small, we see that $(1-\lambda)\hat{x} + \lambda x \in \Omega$, $-\hat{s} + \lambda(\hat{s}+\varepsilon) \in -\text{ri } S^f - S$ and

$$\begin{aligned} g((1-\lambda)\hat{x}+\lambda x) &= (1-\lambda)g(\hat{x}) + \lambda g(x) - s_1, \quad \text{for some } s_1 \in S \\ &= -\hat{s} + \lambda(\hat{s}+\varepsilon) - \lambda s - s_1 \in -S \backslash -S^f, \\ &\qquad\qquad \text{since } s \in S^f. \end{aligned} \tag{51}$$

This contradicts the definition of the minimal cone S^f. Thus

$$S^f_\varepsilon \subset S^f. \tag{52}$$

Now by Theorem 1

$$\begin{aligned} \mu - p(x) &\leq \lambda g(x), \quad \text{for all } x \text{ in } F^f = \Omega \cap g^{-1}(S^f-S) \\ &\leq \lambda\varepsilon, \quad \text{for all } x \text{ in } F_\varepsilon, \text{ the feasible set of} \\ &\qquad (P_{g,\varepsilon}), \text{ by (52) and since } \lambda \in (S^f)^+. \end{aligned}$$

The first inequality of (50) now follows by taking the infimum over x in F_ε. The second inequality follows symmetrically. □

This result reduces to the standard stability result if Slater's Condition holds, i.e. in this case the perturbation ε is no longer restricted since $S^f - S = S - S = R^m$, e.g. Luenberger (1969; 222) or Geoffrion (1971). Thus $\mu(\varepsilon)$ is a continuous function of ε. (In Geoffrion (1971) it was shown that μ is a continuous function of ε when restricted to the subspace $S^f - S^f$. This follows from the above since $S^f = S^f_\varepsilon$ if $\varepsilon \in S^f - S^f$.)

We now derive a generalization of Farkas' lemma. Recall that the usual Farkas' lemma (e.g. Mangasarian (1969)) holds with the assumption that S is polyhedral. In the following A^t: $R^m \to R^n$ is the transpose of the m×n matrix A and S^f is the minimal cone for the constraint $Ax \in S$.

THEOREM 6: Suppose that A *is an* m×n *matrix,* ϕ *is a vector in* R^n, H *is a subspace (whose existence is promised by Corollary 2) and* K *is a closed convex cone which satisfies*

$$S^f = K \cap -H,\ A(R^n) \subset H \text{ and } (44). \tag{53}$$

Then the following are equivalent:

i) The system

$$A^t\lambda = \phi\ ,\quad \lambda \in (S^f)^+ \tag{54}$$

is consistent.

ii) $Ax \in S \Longrightarrow \phi x \geq 0$.

iii) The system

$$A^t\lambda = \phi\ ,\quad \lambda \in K^+ \tag{55}$$

is consistent

iv) $Ax \in K \Longrightarrow \phi x \geq 0$.

PROOF: Consider the abstract convex program

$$(P) \quad \mu = \inf\{\phi x:\ -Ax \in -S,\ x \in \Omega = X\}.$$

Then (ii) is equivalent to the fact that $x^* = 0$ solves (P). By Theorem 3, this is equivalent to

$$0 = \mu = \inf\{\phi x - \lambda Ax:\ x \in R^n\}$$

for some λ in $(S^f)^+$. This in turn is equivalent to

$$0 = \phi - A^t\lambda \ , \quad \text{for some } \lambda \in (S^f)^+,$$

since the gradient

$$\nabla(\phi x - \lambda Ax) = \phi - A^t\lambda.$$

Thus (i) is equivalent to (ii). The remaining equivalences follow from Theorem 4 and Corollary 2. □

Note that if (39) holds, then (iii) and (iv) yield the extension of Farkas' Lemma, with $K^+ = S^+$ (see e.g. Ben-Israel (1969))

$$A^t\lambda = \phi \ , \quad \lambda \in S^+ \text{ is consistent}$$

if and only if

$$Ax \in S \Longrightarrow \phi x \geq 0.$$

This result holds if and only if $A(S^+)$ is closed, or equivalently, $N(A) + S$ is closed which is equivalent to (39).

REFERENCES

ABRAMS, R.A. (1975). "Projections of Convex Programs With Unattained Infima." *SIAM J. Control 13,* 706-718.

ABRAMS, R.A. and KERZNER, L. (1978). "A Simplified Test for Optimality." *J. of Optimization Theory and Applications 25,* 161-170.

BAZARAA, M.S., SHETTY, C.M., GOODE, J.J. and NASHED, M.Z. (1976). "Nonlinear Programming Without Differentiability in Branch Spaces: Necessary and Sufficient Constraint Qualification." *Applicable Analysis 5,* 165-173.

BEN-ISRAEL, A. (1969). "Linear Equations and Inequalities on Finite Dimensional, Real or Complex, Vector Spaces: A Unified Theory." *J. of Mathematical Analysis and Its Applications 27,* 367-389.

BEN-ISRAEL, A., BEN-TAL, A. and ZLOBEC, S. (1976). "Optimality Conditions in Convex Programming." The IX International Symposium on Mathematical Programming, Budapest.

BEN-TAL, A. and BEN-ISRAEL, A. (1979). "Characterizations of Optimality in Convex Programming: The Nondifferential Case." *Applicable Analysis 9,* 137-156.

BERMAN, A. (1973). "Cones, Matrices and Mathematical Programming." Lecture Notes in Economics and Mathematical Systems 79, Springer-Verlag, New York.

BORWEIN, J. and WOLKOWICZ, H. (1981a). "Facial Reduction for a Cone-Convex Programming Problem." *J. of the Australian Mathematical Society.30,* 369-380.

BORWEIN, J. and WOLKOWICZ, H. (1981b). "Characterizations of Optimality for the Abstract Convex Program With Finite Dimensional Range." *J. of the Australian Mathematical Society 30,* 390-411.

BORWEIN, J. and WOLKOWICZ, H. (1979a). "Regularizing the Abstract Convex Program." *J. of Mathematical Analysis and Its Applications,* (forthcoming).

BORWEIN, J. and WOLKOWICZ, H. (1979b). "Characterizations of Optimality Without Constraint Qualification for the Abstract Convex Program." *Mathematical Programming Study*, (forthcoming).

CRAVEN, B.D. and ZLOBEC, S. (1980). "Complete Characterization of Optimality for Convex Programming in Banach Spaces." *Applicable Analysis 11*, 61-78.

GEOFFRION, A.M. (1971). "Duality in Nonlinear Programming: A Simplified Applications-Oriented Development." *SIAM Review 13*, 1-37.

GOULD, F.J. and TOLLE, J.W. (1972). "Geometry of Optimality Conditions and Constraint Qualifications." *Mathematical Programming 2*, 1-18.

LUENBERGER, D.G. (1969). *Optimization by Vector Spaces Methods*. John Wiley & Sons Inc., New York.

MANGASARIAN, O. (1969). *Nonlinear Programming*. McGraw-Hill, New York.

MASSAM, H. (1979). "Optimality Conditions for a Cone-Convex Programming Problem." *J. of the Australian Mathematical Society (Series A), 27*, 141-162.

MOREAU, J.J. (1966). "Convexity and Duality." In *Functional Analysis and Optimization* (E.R. Caranells, ed.), Academic Press, New York, 145-169.

PERESSINI, A.L. (1967). *Ordered Topological Vector Spaces*. Harper and Row, New York.

ROCKAFELLAR, R.T. (1970a). *Convex Analysis*. Princeton University Press, Princeton, N.J.

ROCKAFELLAR, R.T. (1970b). "Some Convex Programs Whose Duals are Linearly Constrained." In *Nonlinear Programming* (J.B. Rosen, O.L. Mangasarian and K. Ritter, eds.), Academic Press, New York, 293-322.

WOLKOWICZ, H. (1978). "Calculating the Cone of Directions of Constancy." *J. of Optimization Theory and Applications 25*, 451-457.

WOLKOWICZ, H. (1980a). "Geometry of Optimality Conditions and Constraint Qualifications: The Convex Case." *Mathematical Programming 19*, 32-60.

WOLKOWICZ, H. (1980b). "A Strengthened Test for Optimality." *J. of Optimization Theory and Applications*, (forthcoming).