JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 83, 495–530 (1981)

Regularizing the Abstract Convex Program

JON BORWEIN*

Department of Mathematics, Carnegie–Mellon University, Pittsburgh, Pennsylvania 15213

AND

HENRY WOLKOWICZ†

Department of Mathematics, University of Alberta, Edmonton, Canada

Submitted by G. Leitmann

Characterizations of optimality for the abstract convex program

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

where S is an arbitrary convex cone in a finite dimensional space, Ω is a convex set, and p and g are respectively convex and S-convex (on Ω), were given in [10]. These characterizations hold without any constraint qualification. They use the "minimal cone" S^f of (P) and the cone of directions of constancy $D_g^=(S^f)$. In the faithfully convex case these cones can be used to regularize (P), i.e., transform (P) into an equivalent program (P_r) for which Slater's condition holds. We present an algorithm that finds both S^f and $D_g^=(S^f)$. The main step of the algorithm consists in solving a particular complementarity problem. We also present a characterization of optimality for (P) in terms of the cone of directions of constancy of a convex functional $D_{\phi g}^=$ rather than $D_g^=(S^f)$.

1. INTRODUCTION

We consider the (abstract) convex program

$$\textit{minimize } p(x) \textit{ subject to } g(x) \in -S, \qquad x \in \Omega, \tag{P}$$

where p is an extended convex functional on X, g is an extended S-convex function on X into Y, X and Y are locally convex spaces with Y being finite dimensional, $\Omega \subset X$ is convex, and S is a convex cone. Primal and dual characterizations of optimality for (P) have been given in [10]. These characterizations use (i) the smallest face of S containing the image of the

* Research partially supported by NSERC A4493.
† Research partially supported by NSERC A3388.

0022-247X/81/100495–36$02.00/0

feasible set (denoted S^f) and (ii) the cone of directions of constancy of g at the optimum point a (denoted $D_g^=(S^f, a)$). These characterizations generalize the so-called "BBZ conditions" [5] and do not require any constraint qualification. Applications include, for example, (i) the optimal control problem where the initial state and final target are given in Y [14, 18] and (ii) the linear estimation problem where we seek the best unbiased etimator in the cone S of positive semi-definite matrices [16].

Our main purpose is to present an algorithm that finds the cones S^f and $D_g^=(S^f, a)$ and also to show how these cones can be used to regularize (P) so that Slater's condition holds.

Section 2 presents several preliminary notions and definitions. In Section 3 we develop the necessary theory dealing with the faces of a finite dimensional convex cone. In particular we discuss exposed faces and introduce the notion of a projectionally exposed face. We then show (see Proposition 3.9 and Remark 3.1) that every face E of S is the intersection of projectional images of S with the intersection being finite if E is polyhedral and facially exposed (in S). This notion allows a simplification of the optimality criteria.

Section 4 extends the notion of the cones of directions given in [5]. We also extend several results for these cones which have proven useful. Section 5 introduces faithfully S-convex functions. We then extend the properties that (i) a convex function bounded on a line is constant on that line [22] and (ii) the cone of directions of constancy of a faithfully convex function is a subspace of X independent of the point $x \in X$ [4], [21].

In Section 6 we recall and strengthen a characterization of optimality for (P) given in [10]. The new characterization is in terms of the cones of constancy of convex functionals rather than $D_g^=(S^f, a)$ and is strengthened in the sense that the Lagrange multiplier relation holds over a larger set. We then present our regularization technique. This technique essentially restricts the program (P) to subspaces of X and Y so that Slater's condition holds.

The algorithm to find S^f and $D_g^=(S^f, a)$ is presented in Section 7. This algorithm is given for weakly faithfully S-convex functions g, i.e., S-convex functions g for which $\phi g(\cdot)$ is faithfully convex for all ϕ in the dual space for which ϕg is convex. Note that all analytic convex and all strictly convex functions are faithfully convex and so this algorithm can be applied to a wide class of functions. The non-faithfully convex case is outlined.

We conclude with several examples in Section 8.

2. Preliminaries

We consider the convex programming problem

$$\textit{minimize } p(x) \textit{ subject to } g(x) \in -S, \qquad x \in \Omega, \tag{P}$$

where p: $X \to R \cup \{+\infty\}$; $g: X \to Y \cup \{+\infty\}$; X and Y are real locally convex (separated topological vector) spaces; Y is finite dimensional with an abstract maximal element $+\infty$; $\Omega \subset X$ is convex; $S \subset Y$ is a convex cone; i.e., $S + S \subset S$ and $\lambda S \subset S$ for all $\lambda \geqslant 0$; p is an extended real convex functional (on Ω) and g is S-convex (on Ω); i.e.,

$$tg(x_1) + (1-t)\, g(x_2) - g(tx_1 + (1-t)\, x_2) \in S \tag{2.1}$$

for any x_1, x_2 (in Ω) and t in $[0, 1]$. The cone S induces in Y a transitive and reflexive *ordering* $\geqslant_S$:

$$x_1 \geqslant x_2 \qquad \text{iff} \quad x_1 - x_2 \in S. \tag{2.2}$$

Unless otherwise specified, it will be assumed that the order is induced by the cone S; e.g., $x = y$ denotes $x =_s y$, etc.... Moreover

$$x_1 > x_2 \qquad \text{iff} \quad x_1 - x_2 \in \operatorname{ri} S \qquad \text{(relative interior of } S\text{).}$$

Further notations are as in [10, 13]. We briefly summarize several essential notations and known results:

$$F = g^{-1}(-S) \cap \Omega; \tag{2.3}$$

$$\operatorname{dom} g \text{ is the } \textit{domain} \text{ of } g;$$

$$\operatorname{dom} p \supset F;$$

X^*, Y^* are the continuous duals of X and Y, respectively, equipped with the w^*-topology; K^+ is the (nonnegative) dual cone of the set K;

$$K^{++} = (K^+)^+ = \overline{\operatorname{cone}}\, K, \tag{2.4}$$

the closure of the convex cone generated by K;

$$(\bar{S}_1 \cap \bar{S}_2)^+ = \overline{S_1^+ + S_2^+}, \tag{2.5}$$

for any two convex cones S_1 and S_2 in Y; $K^\perp = K^+ \cap (-K^+)$ is the *annihilator* of K and $\phi^\perp = \{\phi\}^\perp$ for any vector ϕ; $\nabla g(a; d)$ is the *directional derivative* of g at a and it exists for each direction d if g is convex on X, continuous at a and S is closed and pointed, i.e., $S \cap -S = \{0\}$; $\partial g(a)$ is the *subdifferential* of g at a; $\partial g(a)$ is non-empty when X is a weakly compactly generated Banach space and g is S-convex on X and continuous at a with S closed and pointed. In this case, for any ϕ in S^+ and d in X [28],

$$\phi\, \nabla g(a; d) = \max_{T \in \partial g(a)} \phi T(d). \tag{2.6}$$

When $Y = R$ and $S = R_+$, then (2.6) holds in any locally convex space X. Unless otherwise specified we will assume that (2.6) holds. $\mathcal{N}(\cdot)$ and $\mathcal{R}(\cdot)$ denote null space and range, respectively. The symbol 0 is used for both zero element and subspace of a vector space. I denotes the identity matrix, A^+ denotes the generalized inverse of the matrix A [6], A^t denotes the transpose of A and P_N denotes the (orthogonal) projection on the subspace N.

3. Faces of a Finite Dimensional Cone

In this section we summarize several useful results on the faces of a cone. For more details and missing proofs see, e.g., [2, 3, 10, 12, 15].

DEFINITION 3.1. A subcone K of S is called a *face* of S denoted $K \lhd S$, if

$$x \in K,\ 0 \leqslant y \leqslant x, \quad \text{implies} \quad y \in K. \tag{3.1}$$

PROPOSITION 3.1. *$K \lhd S$ if and only if*

$$0 \leqslant y, x, x+y \in K \quad \textit{implies} \quad x, y \in K. \tag{3.2}$$

PROPOSITION 3.2. (a) *If $K \lhd L \lhd S$, then $K \lhd S$.*

(b) *If $K \lhd S$ and $K \subset L \subset S$, then $K \lhd L$.*

DEFINITION 3.2. (a) A face of S is *exposed* if there exists ϕ in S^+ such that

$$K = \{s \in S : \phi s = 0\}. \tag{3.3}$$

(b) The convex cone S is called *facially exposed* if every face of S is exposed.

PROPOSITION 3.3. *S is facially exposed if and only if exposed faces of exposed faces of S are themselves exposed faces of S.*

Note that the faces of a convex cone are convex cones and are closed when S is closed. Moreover, every polyhedral cone is facially exposed. An example of a cone which is not facially exposed is given in [10].

PROPOSITION 3.4. *If $K \lhd S$, then*

(a) $(K-S) \cap S = (K-K) \cap S = (K-S) \cap K = K$; (3.4)

(b) $(K-S) \cap (S-K) = K-K$. (3.5)

PROPOSITION 3.5. *Let C be an arbitrary subset of S. Then*

(a) *there is a* unique minimal face C^f *containing C;*

(b) *there is a* unique minimal exposed face C^{ef} *containing C.*

PROPOSITION 3.6. *Suppose that E is a proper face of S. Then*

$$E^{ef} \neq S.$$

Proof. By hypothesis

$$E \cap \operatorname{ri} S = \phi.$$

Since $E-E \subset S-S$, the Hahn–Banach theorem says that there exists $0 \neq \phi \in Y^*$ such that

$$\phi s > 0, \qquad \text{for all} \quad s \in \operatorname{ri} S, \qquad \phi s \leqslant 0, \qquad \text{for all} \quad e \in E.$$

Since $E \subset S$, we conclude that E is contained in the face exposed by the positive functional ϕ; i.e.,

$$E \subset \{s \in S: \phi s = 0\} \not\subseteq S. \quad \blacksquare$$

Though not all faces are exposed, the above lemma shows that every face is an exposed face of some larger face of S. For, if E is a proper face of E^{ef}, we then repeat the process in E^{ef}. We eventually must stop since Y is finite dimensional. This allows a reduction process by means of exposed faces, see Section 7.

DEFINITION 3.3. The *minimal cone for* (P), denoted S^f, is defined by

$$S^f = (-g(F))^f, \tag{3.6}$$

where F is the feasible set for (P). Similarly, the *minimal exposed cone for* (P), denoted S^{ef}, is defined by

$$S^{ef} = (-g(F))^{ef}. \tag{3.7}$$

The minimal cones S^f and S^{ef} have the following properties.

PROPOSITION 3.7. [10]. (a) *g is S^f-convex on F*; (3.8)

(b) *$g(F) + S^f$ is convex*; (3.9)

(c) *$g(F) \cap -\operatorname{ri} S^f \neq \phi$, when $F \neq \phi$*; (3.10)

(d) $S^{ef} = (g(F)+S)^{+\perp} \cap S$

$$= \overline{\bigcup_{\lambda \geqslant 0} \lambda(g(F)+S)} \cap S. \tag{3.11}$$

COROLLARY 3.1. *$y \in \operatorname{ri}\{y\}^f$, for all $y \in S$.*

Proof. Choose g constant and equal to $-y$ in (c) above. ∎

PROPOSITION 3.8. *Suppose that E is an exposed face of S. Then E is an exposed face of any subcone of S which contains E.*

Proof. By hypothesis, there exists a supporting hyperplane H to S such that

$$E = S \cap H.$$

Therefore, if K is a subcone of S which contains E,

$$E = K \cap H. \quad ∎$$

Projections onto faces will play a role in our optimality conditions. The following lemma shows that each face can be expressed using projections.

PROPOSITION 3.9. *Suppose that $E \lhd S$ and that A_i, $i \in I$, are all the points of S^+ for which E is not a subset of the hyperplane $A_i^\perp$. Let P_i be a projection onto E–E satisfying*

$$\mathcal{N}(P_i) \subset A_i^\perp, \qquad \text{for all} \quad i \in I, \tag{3.12}$$

and let P be the orthogonal projection onto E–E. Then

$$E = \bigcap_I (P_i S) \cap PS. \tag{3.13}$$

Proof. First note that since E is not a subset of $A_i^\perp$, we can find a subspace L_i of $A_i^\perp$ such that

$$L_i \cap (E\text{–}E) = 0; \qquad L_i \oplus (E\text{–}E) = R^m, \tag{3.14}$$

where $\oplus$ denotes direct sum. Thus we can choose P_i to be the projection onto E–E along L_i; i.e.,

$$\mathcal{R}(P_i) = E\text{–}E; \qquad \mathcal{N}(P_i) = L_i. \tag{3.15}$$

This satisfies (3.12). Now since $P_i E = E$, for all $i \in I$, we get that

$$E \subset \bigcap_I P_i S \cap PS. \tag{3.16}$$

This then implies that

$$\operatorname{span} \bigcap_I P_i S \cap PS = E\text{–}E. \tag{3.17}$$

Thus, if $e \in E\text{–}E$ but $e \notin E$, then the result (3.13) follows if we can show that

$$e \notin P_i S, \qquad \text{for some} \quad i \in I. \tag{3.18}$$

Now, since $E \subset \Lambda_i^{\perp}$ implies $(E\text{–}E) \subset \Lambda_i^{\perp}$, we get

$$\begin{aligned} E &= S \cap (E\text{–}E) \qquad \text{by (3.4)} \\ &= \bigcap_I \Lambda_i^+ \cap (E\text{–}E). \end{aligned}$$

Therefore, there exists $i \in I$ such that

$$e \notin \Lambda_i^+. \tag{3.19}$$

Since $\mathcal{N}(P_i) \subset \Lambda_i^{\perp}$, we get that

$$e + y \notin \Lambda_i^+, \qquad \text{for all} \quad y \in \mathcal{N}(P_i). \tag{3.20}$$

This yields (3.18), since $S \subset \Lambda_i^+$. ∎

Remark 3.1. If E is polyhedral, then E is uniquely determined by its maximal proper faces [12] and these are finite in number. Therefore, if the maximal proper faces of E are exposed (in S), we may take the index set I in the above Proposition to be finite; i.e.,

$$E = \bigcap_{i=1}^{k} P_i S \cap PS. \tag{3.21}$$

In the case that the dimension of E is 0 or 1, this gives

$$E = PS. \tag{3.22}$$

Note that we have not assumed that E itself need be exposed in (3.21).

Being able to express the minimal cone S^f as a finite intersection of projectional images of S will simplify the optimality conditions in Section 6. We now introduce the following projectional notion.

DEFINITION 3.4. The convex cone S is called *projectionally exposed* if every face of S is the image of S under some projection; i.e.,

$$E \lhd S \text{ implies } E = PS, \qquad \text{for some projection } P. \tag{3.23}$$

EXAMPLE 3.1. Consider the cone S of all positive semi-definite matrices in the space Y of all $m \times m$ real symmetric matrices. The matrices are represented by their distinct upper triangular parts and thus $Y = R^{(m^2+m)/2}$. In

[10] it was shown that S is both facially and projectionally exposed. In fact, $E \lhd S$ if and only if

$$E = \{a \in S : \mathcal{N}(a) \supset \mathcal{R}(q)\}, \tag{3.24}$$

for some projection matrix $q \in S$, and then

$$E = (I - q)\, S(I - q), \tag{3.25}$$

where I denotes the identity matrix in Y. That is $E = PS$, where the projection

$$P\cdot = (I - q) \cdot (I - q). \tag{3.26}$$

4. Cones of Directions for Convex Functions

We now extend the notion of the cones of directions, given in [8] for convex functionals, to S-convex functions.

Definition 4.1. For $E \lhd S$, $a \in \operatorname{dom} g$ and the relation $\mathcal{R}$, we denote

$$D_g^{\mathcal{R}}(E, a) = \{d : \exists \bar{\alpha} > 0 \text{ and } g(a + \alpha d)\, \mathcal{R}_{S-E}\, g(a), \text{ for all } 0 < \alpha \leqslant \bar{\alpha}\}. \tag{4.1}$$

When $\mathcal{R}$ is $=$, $\leqslant$, $<$, $>$, and $\geqslant$ then these are the *cones of directions of constancy*, *nonincrease*, *decrease*, *increase*, and *nondecrease*, respectively. (If $Y = R \cup +\{\infty\}$, i.e., g is an extended real convex functional, and if $E = \{0\}$ and $S = R_+$, then (4.1) reduces to the cones of directions given in [8].) For simplicity of notation, we will delete E in the case $E = 0$; e.g., $D_p^{<}(0, a) = D_p^{<}(a)$.

Definition 4.2. Suppose that $C \subset X$. For $a \in C$, the set of *feasible directions* at a is

$$C(a) = \{d : \text{there exists } \bar{\alpha} > 0 \text{ with } a + \alpha d \in C, \text{ for all } 0 < \alpha \leqslant \bar{\alpha}\}.$$

If we choose the face E properly, then the directions of nonincrease $D_g^{\leqslant}(E, a)$ are exactly the feasible directions at a for (P).

Proposition 4.1. *Suppose that g is continuous at $a \in \Omega$, $g(a) \leqslant 0$ and $E = \{-g(a)\}^f$. Then*

$$F(a) = D_g^{\leqslant}(E, a). \tag{4.2}$$

Proof. Suppose that $d \in F(a)$. Then there exists $\bar{\alpha} > 0$ such that

$$g(a + \alpha d) \leqslant 0, \qquad \text{for all} \quad 0 < \alpha \leqslant \bar{\alpha}. \tag{4.3}$$

Thus

$$g(a + \alpha d) \leqslant_{S-E} g(a), \qquad \text{for all} \quad 0 < \alpha \leqslant \bar{\alpha}. \tag{4.4}$$

This shows that

$$F(a) \subset D_g^{\leqslant}(E, a).$$

Conversely, suppose that $d \in D_g^{\leqslant}(E, a)$; i.e., there exists $\bar{\alpha} > 0$ such that (4.4) holds. Since g is continuous at a and $g(a) \in -\operatorname{ri} E$ by Corollary 3.1, we get that (4.3) holds, though possibly for a smaller $\bar{\alpha} > 0$. That $a + \alpha d \in \Omega$ for small $\alpha > 0$ follows similarly. ∎

The above proposition shows that $D_g^{\leqslant}(\{-g(a)\}^f, a)$ is convex. We now see that this holds in a more general case.

PROPOSITION 4.2. *Suppose that $g(a) \leqslant 0$ and E is a face of S. Then*

$$D_g^{\leqslant}(E, a) \textit{ is convex.} \tag{4.5}$$

Proof. Let $d_1, d_2 \in D_g^{\leqslant}(E, a)$ and $d = \lambda d_1 + (1 - \lambda) d_2$, $0 \leqslant \lambda \leqslant 1$. Then there exists $\bar{\alpha} > 0$ such that

$$g(a + \alpha d_i) \leqslant_{S-E} g(a), \qquad \text{for all} \quad 0 < \alpha \leqslant \bar{\alpha}.$$

Thus, for all $0 < a \leqslant \bar{\alpha}$,

$$\begin{aligned} g(a + \alpha d) &\leqslant_S \lambda g(a + \alpha d_1) + (1 - \lambda)\, g(a + \alpha d_2), \qquad \text{since } g \text{ is } S\text{-convex} \\ &\leqslant_{S-E} \lambda g(a) + (1 - \lambda)\, g(a) = g(a). \end{aligned}$$

Thus $d \in D_g^{\leqslant}(E, a)$. ∎

The feasible directions can be used to characterize optimality.

LEMMA 4.1. [8] *A feasible solution $a \in F$ of* (P) *is optimal if and only if*

$$D_p^{<}(a) \cap F(a) = \phi. \tag{4.5}$$

Note that we can replace $F(a)$ in (4.5) by its closure if p is continuous at a [10].

5. Cone-Faithfully Convex Functions

Definition 5.1. The S-convex function g is *faithfully convex* (with respect to the face E) if: g maps a line segment into E–E only if g maps the whole line containing that line segment into E–E. We then say that g is S-faithfully convex if it is faithfully convex with respect to every face of S.

The following proposition extends the result by Rockafellar [22] that: a convex functional bounded on a line is constant on that line.

Proposition 5.1. *Suppose that S is closed. Let $y \in Y$ and $a, d \in X$. If*

$$g(a + \alpha d) \leqslant y, \qquad \textit{for all} \quad \alpha \in R, \tag{5.1}$$

then

$$g(a + \alpha d) = g(a), \qquad \textit{for all} \quad \alpha \in R. \tag{5.2}$$

Proof. Since the function $g(\cdot) - y$ is also S-convex, we can assume without loss of generality that $y = 0$. Now (5.1) implies

$$\phi g(a + \alpha d) \leqslant 0, \qquad \text{for all} \quad \alpha \in R, \quad \text{for all} \quad \phi \in S^+. \tag{5.3}$$

Rockafellar's result for convex functionals yields

$$\phi(g(a + d) - g(a)) = 0, \qquad \text{for all} \quad \alpha \in R, \quad \text{for all} \quad \phi \in S^+, \tag{5.4}$$

which in turn implies (5.2) since S is closed. ∎

Corollary 5.1. *Under the assumptions of the proposition,*

$$d \in D_g^=(S, a). \tag{5.5}$$

Faithfully convex functionals introduced by Rockafellar [21] have proven very useful in optimization theory [5, 24, 26]. Since all analytic as well as all strictly convex functionals are faithfully convex, applications are widespread. One property which has proven extremely useful in applications to algorithms [24, 26], is that: the cone of directions of constancy of a continuous faithfully convex functional is a closed subspace of X independent of the point x under consideration. We now extend this property to S-convex functions.

Theorem 5.1. *Suppose that E is a face of S, g is continuous and S-faithfully convex and both S and S–E are closed. Then $D_g^=(E, a) = D_g^=(E)$ is a closed subspace of X independent of a.*

Proof. First, let us show that

$$D_g^=(E, a) \text{ is a subspace (closed).} \tag{5.6}$$

Let $d = d_1 + d_2$ and $d_i \in D_g^=(E, a)$. If $\alpha \in R$, then

$$\begin{aligned} g(a + \alpha d) &= g(\tfrac{1}{2}(a + 2\alpha d_1) + \tfrac{1}{2}(a + 2\alpha d_2)) \\ &\leqslant_S \tfrac{1}{2} g(a + 2\alpha d_1) + \tfrac{1}{2} g(a + 2\alpha d_2) \\ &\leqslant_{S-E} g(a). \end{aligned}$$

Thus (5.6) follows by Corollary 5.1 and Proposition 3.4. (The closure follows by continuity.)

Let us now show independence of the point. Let $x_1, x_2 \in X$ and $d \in D_g^=(E, x_1)$. We need to show that $d \in D_g^=(E, x_2)$. By Corollary 5.1, it is sufficient to show that

$$g(x_2 + \alpha d) \leqslant_{S-E} z, \tag{5.7}$$

for all $\alpha \in R$ and some $z \in Y$. Now choose $y \in Y$ so that

$$g(x_2) \leqslant_{S-E} g(x_1) + y. \tag{5.8}$$

For example, $y = g(x_2) - g(x_1) - s + e$, for some $s \in S$ and $e \in E$. Fix $\alpha \in R$ and let

$$\begin{gathered} 0 < t_k < 1; \quad t_k \to 0 \quad \text{as} \quad k \to \infty; \quad \gamma_k = 1/t_k; \\ z^k = \alpha d + t_k(x_1 - x_2). \end{gathered}$$

Since $d \in D_g^=(E, x_1)$ and g is S-faithfully convex, we get that

$$g(x_1) =_{S-E} g(x_1 + \gamma_k \alpha d) \tag{5.9}$$

which implies that

$$g(x_1) =_{S-E} g(x_2 + \gamma_k z^k). \tag{5.10}$$

But

$$\begin{aligned} g(x_2 + z^k) &\leqslant_S (1 - t_k)\, g(x_2) + t_k\, g(x_2 + \gamma_k z^k), \qquad \text{by } S\text{-convexity of } g, \\ &\leqslant_{S-E} (1 - t_k)\, g(x_1) + (1 - t_k)\, y + t_k\, g(x_2 + \gamma_k z^k), \qquad \text{by (5.8)}, \\ &=_{S-E} g(x_1) + (1 - t_k)\, y, \qquad \text{by (5.10)}. \end{aligned} \tag{5.11}$$

Now

$$\begin{aligned} g(x_2 + \alpha d) &= \lim_{k \to \infty} g(x_2 + z^k), \qquad \text{by continuity of } g, \\ &\leqslant_{S-E} \lim_{k \to \infty} g(x_1) + (1 - t_k) y, \qquad \text{by (5.11)}, \\ &= g(x_1) + y. \end{aligned}$$

Since both $g(x_1)$ and y are independent of α, we have proven (5.7). ∎

The condition that E–S be closed may be restrictive in applications. E–S is always closed in the polyhedral case. It is still an open question whether one can relax the closure condition. Note that, as in the case of a real convex functional, analytic S-convex functions are faithfully convex as are strictly S-convex functions.

In the algorithm presented in Section 7, we will assume that g is weakly faithfully convex; i.e., g is faithfully convex for each $\phi \in Y^*$ for which ϕg is convex (on Ω). This removes the requirement that E–S be closed and also shows that $D_g^=(S^f)$ is a subspace independent of the point $x \in X$.

6. Characterizations of Optimality and Regularization

In [10], we presented the following characterization of optimality for (P).

THEOREM 6.1. (a) *Suppose that μ is the finite optimal value of* (P). *Then*

$$p(x) + \lambda g(x) \geqslant \mu, \qquad \textit{for all} \quad x \in F^f, \tag{6.1}$$

for some λ in $(S^f)^+$ and $F^f = g^{-1}(S^f – S) \cap \Omega$.

(b) *If μ is actually attained by $p(a)$, $a \in F$, then in addition*

$$\lambda g(a) = 0. \tag{6.2}$$

Remark 6.1. In certain cases the multiplier in (6.1) may be supposed to be in S^+ rather than just in $(S^f)^+$ (independent of p and g). This situation is characterized by

$$S^+ + (S^f)^\perp = (S^f)^+, \tag{6.3}$$

or equivalently, when S is closed, by

$$S^+ + (S^f)^\perp \text{ is closed.} \tag{6.4}$$

Thus multipliers in S^+ exist whenever S is polyhedral. In fact, this is true whenever S^f is polyhedral and facially exposed in S, for then S^f can be written as the intersection of a finite number of projectional images of S (this follows by Remark 3.1 and a proof similar to the proof of Theorem 6.3 below). For the same reason, multipliers in S^+ exist whenever S is projectionally exposed. In particular (see Example 3.1) multipliers in S^+ exist whenever S is the cone of $m \times m$ psd matrices.

The above theorem and remark characterize optimality for (P) without any constraint qualification. The multiplier relationship in (6.1) is restricted to the set S^f. For stability and related results it is of interest to get the "strongest" optimality conditions, i.e., to have the set F^f as large as possible [11]. In fact to ensure stability for all feasible perturbations, one needs $F^f = \Omega$. We now show that a larger F^f is possible. First we will need the following lemma, which will also prove useful in our algorithm in Section 7.

LEMMA 6.1. *Let H be a subspace of Y. Then*

$$g \text{ is } (S \cap H)\text{-convex on } F^H = g^{-1}(H) \cap \Omega. \tag{6.5}$$

Proof. Let $0 < t < 1$, $x_t = tx_1 + (1-t)x_2$ and $x_1, x_2, x_t \in F^H$, i.e., $g(x_1)$, $g(x_2)$, $g(x_t) \in H$. Therefore

$$tg(x_1) + (1-t)\, g(x_2) - g(x_t) \in (S \cap H),$$

since g is S-convex and H is a subspace. ∎

Note that F^H need not be convex in the above. Note also that

$$F^f = g^{-1}(S^f - S^f) \cap \Omega \tag{6.6}$$

[10, Proposition 4.1(c)] and so possibly $F^H \supset F^f$; i.e., F^H is larger than F^f with equality if $H = S^f - S^f$. Before presenting the strengthened optimality characterization, we first present the following optimality conditions which hold under a "generalized Slater's condition."

LEMMA 6.2. *Suppose that g is continuous and weakly faithfully S-convex (on Ω), Ω is the intersection of a polyhedral set and a closed linear manifold, and* (P) *satisfies the generalized Slater's condition: there exists*

$$\hat{x} \in \Omega \qquad \text{with} \quad g(\hat{x}) \in -\operatorname{ri} S. \tag{6.7}$$

Then the standard Lagrange multiplier theorem holds; i.e., *Theorem* 6.1 *holds with F^f replaced by Ω and $(S^f)^+$ replaced by S^+.*

Proof. By (6.7), we get that $S^f = S$. Since Y is finite dimensional, we can find ϕ_i, $i = 1,..., t$, in $S^\perp$ such that

$$S\text{–}S = \bigcap_{i=1}^{t} \phi_i^+. \tag{6.8}$$

Now by Theorem 6.1 and (6.6), there exists $\lambda \in S^+$ such that (P) is equivalent to the program

$$\inf\{p(x) + \lambda g(x): x \in F^f = g^{-1}(S\text{–}S) \cap \Omega\}$$

which, by (6.8), is itself equivalent to

$$\inf\{p(x) + \lambda g(x): (\phi_i g)(x) \leqslant 0,\ i = 1,..., t,\ x \in \Omega\}. \tag{6.9}$$

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

$$\phi_i g \text{ is affine}, \qquad i = 1,..., t \text{ (on } \Omega).$$

Since x is restricted to Ω in (6.9), we can assume that the functions $\phi_i g$ are affine on all of X. Suppose that

$$\Omega = P \cap V,$$

where V is a closed subspace and

$$P = \{x: \psi_i x - a_i \leqslant 0,\ i = 1,..., k\}, \qquad \psi_i \in X^*, \tag{6.10}$$

is polyhedral. Then (6.9) is equivalent to the linearly constrained convex program

$$\inf\{p(x) + \lambda g(x): \phi_i g(x) \leqslant 0,\ i = 1,..., t;$$
$$\psi_i g(x) \leqslant a_i,\ i = 1,..., k,\ x \in V\}. \tag{6.11}$$

Since the $t + k$ constraints for this program are all linear and finite in number, and any feasible point for this program is in ri V (when V is a closed subspace) we have satisfied the generalized Slater's condition for the ordinary convex program; i.e., there exists a feasible point in the relative interior of the constraint set (V in our case) which satisfies with strict inequality all the inequality constraints which are not affine (none in our case). We can now obtain Kuhn–Tucker multipliers (see, e.g., [22, Theorem 28.2]) $\lambda_i \geqslant 0$, $i = 1,..., t$, corresponding to the constraints $\phi_i g$. The result now follows since $\lambda + \sum_{i=1}^{t} \lambda_i \phi_i \in S^+$.

Rather than apply the result in [22], which is phrased in finite dimensions, we can apply Pshenichnyii's condition [13, p. 87] to the program (6.11) and get that

$$\partial(p+\lambda g)(a) \cap (\{x\colon \phi_i g(x) \leqslant 0,\ i=1,\dots,t\} \cap \Omega - a)^+ \neq \varnothing,$$

which reduces to

$$\partial(p+\lambda g)(a) \cap (\{x\colon \phi_i g(x) \leqslant 0,\ i=1,\dots,t\} - a)^+ + (\Omega - a)^+) \neq \varnothing,$$

by (2.5) (closure holds since the polar of a polyhedral set is finitely generated), or equivalently

$$(\partial p(a) + \partial \lambda g(a)) \cap (\operatorname{cone}\{\phi_i g\}_{i=1}^{t} + (\Omega - a)^+) \neq \varnothing$$

since g is continuous and $\phi_i g(a) = 0$, $i = 1,\dots, t$, or equivalently,

$$\left(\partial p(a) + \partial \lambda g(a) + \sum_{i=1}^{t} \alpha_i \phi_i g\right) \cap (\Omega - a)^+ \neq \varnothing,$$

where $\alpha_i \geqslant 0$, $i = 1,\dots, t$, or equivalently, a solves the program

$$\inf\{(p+\bar{\lambda} g)(x)\colon x \in \Omega\},$$

where $\bar{\lambda} = \lambda + \sum_{i=1}^{t} \alpha_i \phi_i$ is in S^+. ∎

We now show that we can strengthen Theorem 6.1 when g is weakly faithfully S-convex, Ω is polyhedral, and S^f is exposed. (See Remark 6.3 below for S^f not necessarily exposed.)

THEOREM 6.2. *Suppose that g is continuous and weakly faithfully S-convex, Ω is polyhedral, and S^f is exposed; i.e.,*

$$\phi \in S^+; \qquad H = \phi^{\perp}; \qquad S^f = S \cap H. \tag{6.12}$$

Let K^H be any convex set which satisfies

$$F^f \subset K^H \subset F^H, \tag{6.13}$$

where F^f and F^H are as in (6.1) *and* (6.5), *respectively. Then Theorem* 6.1 *holds with F^f replaced by the* (*larger*) *convex set K^H.*

Proof. Let

$$K^L = (\hat{x} + D^{=}_{\phi g}(\hat{x})) \cap \Omega, \tag{6.14}$$

where $\hat{x} \in F$. Let us show that K^L is the largest (closed) convex set which satisfies (6.13). That K^L is closed and convex follows from the continuity and faithful convexity of g and the polyhedrality of Ω. Now let $x \in K^L$. Then $x \in \Omega$ and $x = \hat{x} + \alpha d$, for some $\alpha \geqslant 0$ and $d \in D^{=}_{\phi g}(\hat{x})$. Thus $\phi g(x) = 0$, which yields $x \in F^H$. Conversely suppose K is convex and satisfies (6.13), and $x \in K$. Then, since $K \subset F^H$, we get that $x \in \Omega$ and $\phi g(x) = 0$. Moreover, since K is convex and contains F^f, $\hat{x} + \alpha(x - \hat{x})$ is in $K \subset F^H$, for all $0 \leqslant \alpha \leqslant 1$; i.e., $\phi g(\hat{x} + \alpha d) = 0$, for all $0 \leqslant \alpha \leqslant 1$ and $d = x - \hat{x}$. Thus

$$x \in (\hat{x} + D^{=}_{\phi g}(\hat{x})) \cap \Omega = K^L; \tag{6.15}$$

i.e., this shows that $K \subset K^H$.

Since

$$F \subset F^f \subset K^L, \tag{6.16}$$

we can rewrite (P) as

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

The result (for K^L in (6.14)) now follows from Lemma 6.2, (3.10), the polyhedrality of Ω and the faithful convexity of ϕg. Note that g is S^f-convex on K^L by Lemma 6.1. Now if λ is the Lagrange multiplier found in (6.1) (for the largest K^H as given in (6.14)), then (6.1) and (6.2) clearly hold for any convex subset of this K^H which contains the feasible set F. Thus, since $F \subset F^f$, we can choose any K^H which satisfies (6.13). ∎

The above theorem gives us a variety of optimality conditions. First, we can choose the subspace H which satisfies (6.12). Then we choose the desired K^H in (6.13). Note that if we choose $H = S^f - S^f$, then we recover Theorem 6.1. In this case we no longer require the assumptions of faithful convexity or of polyhedrality. These assumptions can be weakened but cannot be eliminated entirely. See [27] for examples in the case $S = R^m_+$.

Remark 6.2. Corresponding to Remark 6.1, we get that the multiplier in (6.1), with F^f replaced by K^H, may be supposed to be in S^+, rather than just $(S^f)^+$, exactly when

$$S^+ + H^+ = (S^f)^+, \tag{6.17}$$

or equivalently, when S is closed,

$$S^+ + H^+ \text{ is closed.} \tag{6.18}$$

Proof. The proof is similar to the proof of Corollary 4.2 in [10]. We include it here for completeness. Note that (6.17) and (6.18) are equivalent by (2.5). Now if (6.17) holds and λ satisfies (6.1) and (6.2) (with K^H instead of F^f) then one can solve $\lambda = \phi + h$ with ϕ in S^+ and h in H^+. Hence for any x in $K^H \subset F^H$

$$\lambda g(x) = \phi g(x) + hg(x) = \phi g(x) \tag{6.19}$$

since $g(F^H) \subset H$ and $H^+ = H^\perp$. Thus $\phi g(x)$ may be substituted for $\lambda g(x)$ in (6.1) and (6.2).

Conversely, suppose ϕ lies in $(S^f)^+$. Let P be the orthogonal projection on H. Consider

$$\mu = \inf\{\phi P(x)\colon -Px \in -S,\ x \in X\}. \tag{P}$$

Then $PP^{-1}(S) = PP^{-1}(H \cap S) \subset H$ and so $\mu = 0$. Also $-P^{-1}(H) = X$ so that (6.1) (with K^H the largest convex set satisfying (6.12) replacing F^f) yields

$$\phi Px + \lambda(-Px) \geqslant 0, \qquad \text{for all} \quad x \in X. \tag{6.20}$$

Since we now assume that $\lambda \in S^+$ we derive that

$$\begin{aligned} \phi &= \phi - (\phi P - \lambda P) \\ &= (\phi - \lambda)(I - P) + \lambda \in H^\perp + S^+. \quad \blacksquare \end{aligned}$$

As above, we note that multipliers in S^+ exist whenever S is polyhedral. Furthermore, since $F^f \subset K^H$, we get that (6.17) $\Leftrightarrow$ (6.3), or equivalently, when S is closed, that (6.18) $\Leftrightarrow$ (6.4). Now if S is closed and S^f is exposed, i.e., $S^f = S \cap \phi^\perp$, then (6.18) becomes

$$S^+ + \operatorname{span}\{\phi\} \text{ is closed.} \tag{6.21}$$

Remark 6.3. Primal and dual characterizations of optimality, using directional derivatives and subdifferentials, were given in [10]. These follow directly from Theorem 6.1, by applying Pshenichnyii's condition [13, p. 87], and employ the cones of directions $D_g^=(S^f, a)$ and $D_g^{\leqslant}(S^f, a)$. We now see that Theorem 6.2 allows us to replace the above mentioned cones by cones of directions of constancy of convex functionals. In particular, by repeated application of Proposition 3.6, we get the following equivalent programs to (P) (with $S_1 = S$, S_i^{ef} denotes the smallest exposed face of S_i containing S^f and $\hat{x} \in F$):

$$\inf\{p(x)\colon g(x) \in -S_2,\ x \in \Omega_2 = (\hat{x} + D_{\phi_1 g}) \cap \Omega\}, \tag{P$_2$}$$

$$\phi_1 \in S_1^+, \qquad S_2 = S_1^{ef} = (S_1 \cap \phi_1^\perp),$$

$$\inf\left\{p(x)\colon g(x) \in -S_3,\ x \in \Omega_3 = \left(x + \bigcap_{i=1}^{2} D^=_{\phi_i g}\right) \cap \Omega\right\}, \tag{P$_3$}$$

$$\phi_2 \in S_2^+, \qquad S_3 = S_2^{ef} = (S_2 \cap \phi_2^\perp),$$

$$\inf \left\{ p(x) \colon g(x) \in -S_t, x \in \Omega_t = \left(x + \bigcap_{i=1}^{t} D^{=}_{\phi_i g} \right) \cap \Omega \right\}, \tag{P$_t$}$$

$$S_1 = S^f, \qquad \phi_i \in S_i^t, \qquad S_{i+1} = (S_i \cap \phi_i^{\perp}).$$

Thus we can choose

$$H = \bigcap_{i=1}^{t} \phi_i^{\perp} \tag{6.22}$$

in (6.12) and

$$K^H = \hat{x} + \bigcap_{i=1}^{t} D^{=}_{\phi_i g}$$

in (6.13). (For more details see the proof of Lemma 7.2 below.) Therefore, by Pshenichnyii's condition, we can replace $D_g^{=}(S^f, a)$ by $\bigcap_{i=1}^{t} D^{=}_{\phi_i g}$. Moreover, if S^f is exposed, i.e., $S^f = \phi^{\perp} \cap S$, $\phi \in S^f$, then the characterizations and duality results are in terms of the cone of directions of constancy of the single convex functional ϕg; e.g., if a is feasible and we define the *restricted Lagrangian* by

$$L^H(\lambda) = \inf\{p(x) + \lambda g(x) \colon x \in \hat{x} + D^{=}_{\phi g}\}, \tag{6.24}$$

where $\hat{x}$ is any feasible point, then

$$\mu = p(a) \qquad \text{(i.e., } a \text{ is optimal for (P))}$$

if and only if

$$0 \in \partial p(a) + \partial s^+ g(a) - (D^{=}_{\phi g} \cap \operatorname{cone}(\Omega - a))^+$$
$$\text{for some } s^+ \in (S^f)^+ \text{ with } s^+ g(a) = 0$$

if and only if

$$\mu = \sup\{L^H(\lambda) \colon \lambda \in (S^f)^+)\}.$$

Again $(S^f)^+$ may be replaced by S^+ if (6.3) holds. Note that if S is a polyhedral cone, then (6.3) always holds and S^f is always exposed. For example, suppose that S is R^m_+, the nonnegative orthant in R^m, $g = (g^k)$ and $\phi = (\phi_k)$ is in R^m_+ with $\phi_k = 0$ if $k \notin \mathscr{P}^{=}$, $\phi_k > 0$ if $k \in \mathscr{P}^{=}$, where $\mathscr{P}^{=}$ is the set of "equality constraints," i.e., the set of constraints g^k which are identically zero on the feasible set [1]. Then S^f is exposed by ϕ and the above characterization of optimality simplifies and strengthens the result in [1, 4]. One can also allow $\phi_k = 0$ if $k \in \mathscr{P}^{=}$ and g^k is affine (see [27]).

Now suppose that the cone of constancy $D_g^=(S^f)$ is a subspace of X independent of the point $a \in X$. We then get the following regularization technique. Recall that Slater's condition for (P) is

$$\text{there exists } \hat{x} \in \Omega \text{ such that } g(\hat{x}) \in -\operatorname{int} S.$$

THEOREM 6.3. *Let $a \in F$ and $D_g^=(S^f)$ be a subspace independent of $x \in F$. Suppose that $B: Z \to X$ is a linear operator satisfying*

$$D_g^=(S^f) = \mathscr{R}(B),$$

where Z is a locally convex space, and that Q is a full row rank matrix satisfying

$$S^f\text{–}S^f = \mathscr{R}(Q^t).$$

Consider the program, in the variable $z \in Z$,

$$\text{minimize } p(a + Bz) \text{ subject to } Qg(a + Bz) \in -QS^f, \text{ and}$$
$$z \in \bar{\Omega} = \{z: Bz \in \Omega - a\}. \qquad (\mathrm{P_r})$$

Then Slater's condition is satisfied for $(\mathrm{P_r})$ *and $z = 0$ is a feasible point of* $(\mathrm{P_r})$. *Moreover, if z^* solves* $(\mathrm{P_r})$, *then $a + Bz^*$ solves* (P).

In addition, if S^f is projectionally exposed, i.e., if $S^f = PS$ for some projection P, then we can replace QS^f in the definition of $(\mathrm{P_r})$ *by QS as long as we choose Q so that $Q^\dagger Q = P$.*

Proof. We write the following equivalent programs to (P):

$$\text{minimize } p(x) \text{ subject to } g(x) \in -S^f, \quad x \in \Omega,\ g(x) \in S^f - S^f; \qquad (\mathrm{P_1})$$

$$\text{minimize } p(x) \text{ subject to } Qg(x) \in -QS^f, \quad x \in \Omega,\ g(x) \in S^f\text{–}S^f; \qquad (\mathrm{P_2})$$

$$\text{minimize } p(x) \text{ subject to } Qg(x) \in -QS^f, \quad x \in \Omega,\ x - a \in \mathscr{R}(B); \qquad (\mathrm{P_3})$$

$$\text{minimize } p(a + Bz) \text{ subject to } Qg(a + Bz) \in -QS^f, \quad Bz \in \Omega - a. \qquad (\mathrm{P_r})$$

Thus $(\mathrm{P_r})$ is equivalent to the original program (P). That Slater's condition holds for $(\mathrm{P_r})$ follows from (3.10) and from the fact that Q is onto.

That we can replace QS^f by QS if S^f is projectionally exposed follows from the relation [6]

$$QQ^\dagger Q = Q. \quad \blacksquare$$

In the polyhedral case, the above regularization reduces to the one in [25]. It is now of interest to be able to calculate the cones S^f and $D_g^=(S^f)$. Note that even if g is *not* faithfully S-convex and if $\bar{\Omega} = \Omega \cap (D_g^{=}(S^f, a) + a)$, then the program

$$\textit{minimize } p(x) \textit{ subject to } Qg(x) \in -QS^f, \qquad x \in \bar{\Omega}, \qquad (\mathrm{P}_{\mathrm{r,a}})$$

is equivalent to (P) and regular *at a*. In fact it satisfies Slater's condition at a.

7. Facial Reduction to Find S^f, $D_g^{=}(S^f)$

We now present an algorithm to find the cones S^f and $D_g^{=}(S^f)$. The algorithm is a finite iterative method. The basic step involves reducing the problem to an equivalent problem on an exposed face of S containing S^f. It is interesting to note that S^f is found even though we might have $S^f \neq S^{ef}$, i.e., even though S^f may not be exposed. In the case that S is polyhedral, the algorithm is equivalent to the one in [26] which itself was a modification of the one originally given in [1].

The reduction step is based on the following lemma. (We assume that g is continuous in the sequel.)

Lemma 7.1. *Suppose that $a \in F$. Then the system*

$$\begin{aligned} &(\Omega - a)^+ \cap \partial\phi g(a) \neq \varnothing, \\ &\phi \in S^+, \qquad \phi g(a) = 0, \end{aligned} \tag{7.1}$$

is consistent only if

$$S^f \subset \phi^\perp \cap S. \tag{7.2}$$

Proof. Suppose that (7.1) holds. Then a is a global minimum for the convex function $\phi g(\cdot)$ on the convex set Ω, which implies that $\phi g(x) = 0$, for all $x \in F$. Thus the exposed face $\phi^\perp \cap S$ contains $-g(F)$ and therefore also contains S^f. ∎

We will also need the following theorem of the alternative.

Theorem 7.1. *Suppose that $a \in F$. Then* exactly one *of the following two systems is consistent.*

$$(\Omega - a)^+ \cap \partial\phi g(a) \neq \varnothing, \qquad 0 \neq \phi \in S^+, \qquad \phi g(a) = 0. \tag{7.3}$$

$$g(\hat{x}) \in -\operatorname{int} S, \qquad \hat{x} \in \Omega \quad (\textit{Slater's condition}). \tag{7.4}$$

Proof. Suppose that (7.3) holds. Then $0 = \phi g(a)$ is a global minimum of the convex function $\phi g(\cdot)$ on the convex set Ω. Thus the system

$$\phi g(x) < 0, \qquad x \in \Omega, \tag{7.5}$$

is inconsistent, which implies that (7.4) fails.

Conversely, suppose that (7.4) fails. We can assume $\operatorname{int} S \neq \varnothing$. Otherwise choose $\phi \in (S-S)^{\perp}$ to satisfy (7.3). Thus

$$0 \notin g(\Omega) + \operatorname{int} S \text{ (open, convex set)}, \tag{7.6}$$

which implies that there exists $0 \neq \phi \in Y^*$ such that

$$\phi(g(\Omega) + \operatorname{int} S) \geqslant 0.$$

But then

$$\phi \in S^+; \qquad \phi(g(\Omega)) \geqslant 0. \tag{7.7}$$

Since $g(\Omega) \subset -S$, we get $\phi g(a) = 0$. Thus $0 = \phi g(a)$ is again a global minimum of $\phi g(\cdot)$ on Ω and (7.3) follows by Pshenichnyii's condition [13, p. 87]. ∎

COROLLARY 7.1. *Suppose that $a \in F$ and the system* (7.3) *is inconsistent. Then*

$$S^f = S; \qquad D_g^=(S^f) = X.$$

Remark 7.1. If $f: R^n \to R$ is a faithfully convex functional, then the cone of directions of constancy of f, $D_f^=$, can be found using the algorithm in [24] (f differentiable) or in [26] (f nondifferentiable). Let us refer to this algorithm as algorithm A.

We now present the algorithm that finds S^f and $D_g^=(S^f)$.

Algorithm B

Initialization. Let $a \in F$; $\Omega_0 = \Omega - a$; $m_0 = \dim Y$; $Q_0 = I_{m_0 \times m_0}$; $S_0 = S$; $n_0 = \dim X$; $P_0 = I_{n_0 \times n_0}$; $i = 0$.

i-th *step* $(0 \leqslant i \leqslant t)$. If $m_i > 0$ *and* the system

$$\begin{gathered} \Omega_i^+ \cap [\partial \phi_i Q_i g(a)] P_i \neq \varnothing, \\ \phi_i Q_i g(a) = 0, \qquad 0 \neq \phi_i = S_i^+ \end{gathered} \tag{7.8}$$

is consistent, then use algorithm A to find the $n_i \times n_{i+1}$ matrix A_{i+1} satisfying

$$\mathscr{R}(A_{i+1}) = D^{=}_{\phi_i Q_i g \circ P_i}. \tag{7.9}$$

Then set

$$m_{i+1} = m_i - 1 = \dim \phi_i^{\perp};$$

$$B_{i+1} : \phi_i \xrightarrow{\text{onto}} R^{m_{i+1}};$$

$$\mathscr{N}(B_{i+1}) = \operatorname{span}\{\phi_i\} \qquad (\text{with } B_{i+1} = [1] \text{ if } m_{i+1} = 0);$$

$$P_{i+1} = P_i A_{i+1}$$

$$Q_{i+1} = B_{i+1} Q_i;$$

$$E_i = S_i \cap \phi_i^{\perp};$$

$$S_{i+1} = B_{i+1} E_i;$$

$$\Omega_{i+1} = A_{i+1}^{\dagger}\{\Omega_i \cap \mathscr{R}(A_{i+1})\}.$$

Now proceed to step $i + 1$.

If the system (7.8) is inconsistent or $m_i = 0$, then STOP.

Conclusion.

$$S^f = B_1^{\dagger} B_2^{\dagger} \cdots B_i^{\dagger} S_i,$$
$$D_g^{=}(S^f) = \mathscr{R}(P_i). \tag{7.10}$$

LEMMA 7.2. *Let*

$$p^k(y) = p(a + P_k y)$$
$$g^k(y) = Q_k g(a + P_k y)$$
$$F^k = \{y \in R^{n_k}: g^{(k)}(y) \in -S_k, \ y \in \Omega_k\}$$

and consider the programs

$$\inf\{p^k(y): y \in F^k\}. \tag{P_k}$$

Then, for $k = 0, 1, \ldots, t$ (*or* $t - 1$ *depending on the context*):

(a) $g^{k+1}(y) = B_{k+1} g^k(A_{k+1} y)$;

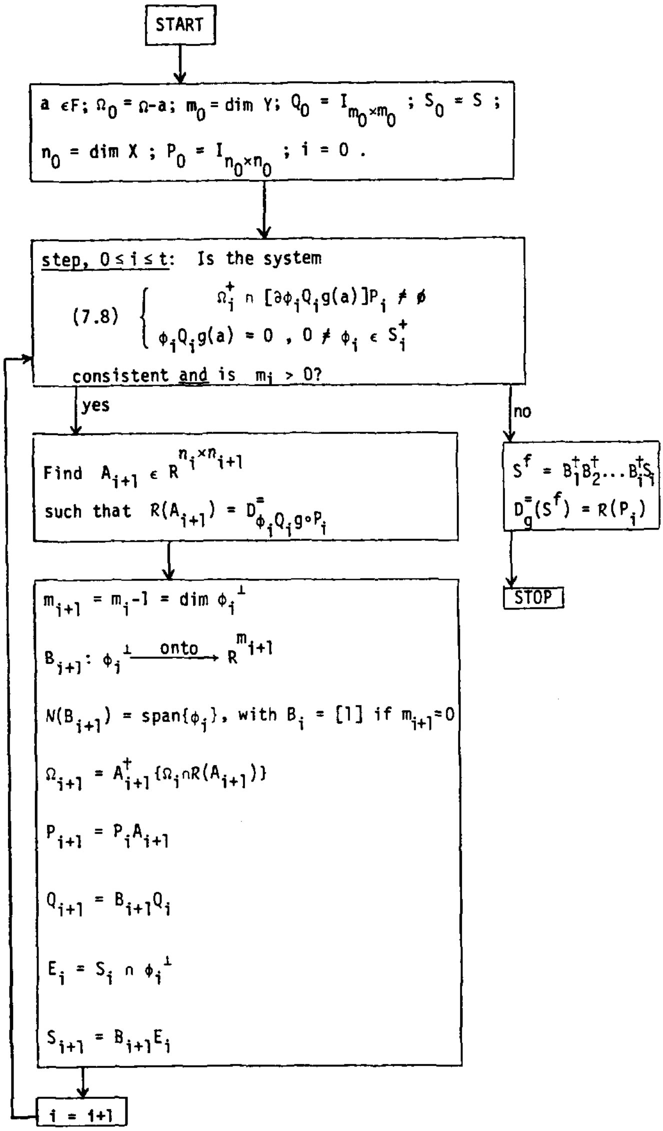
(b) $0 \in F^k$;

(c) g^k *is* S_k*-convex and* $D^{=}_{\phi_k Q_k g \circ p_k} = D^{=}_{\phi_k g^k}$;

(d) $F^k \subset A_{k+1} F^{k+1}$;

(e) $S_k^f = B_{k+1}^{\dagger} S_{k+1}^f$ *and* $D^{=}_{g^k}(S_k^f) = A_k D^{=}_{g^{k+1}}(S_{k+1}^f)$, *where* S_k^f *is the minimal cone of* (P_k).

ALGORITHM B



Proof.

(a)
$$\begin{aligned} g^{k+1}(y) &= Q_{k+1}\, g(a + P_{k+1} y) \\ &= B_{k+1} Q_k\, g(a + P_k A_{k+1} y) \qquad (7.11) \\ &= B_{k+1}\, g^k(A_{k+1} y). \end{aligned}$$

(b) The result holds for $k=0$ by hypothesis. So let us assume that the result holds for g^k and prove it for g^{k+1}. By Lemma 7.1, we get that

$$S_k^f \subset \phi_k^\perp \cap S_k. \tag{7.12}$$

Note that

$$\begin{aligned} \partial\phi_k g^k(0) &= \partial\phi_k Q_k g(a+P_k y) \qquad \text{at} \quad y=0 \\ &= [\partial\phi_k Q_k g(a)]\, P_k, \qquad \text{by the chain rule,} \end{aligned} \tag{7.13}$$

while

$$\phi_k g^k(0) = \phi_k Q_k g(a). \tag{7.14}$$

Thus $g^k(0) \in -S_k^f \subset E_k$ and

$$\begin{aligned} g^{k+1}(0) &= B_{k+1} g^k(A_{k+1}0) \qquad \text{by (a)} \\ &= B_{k+1} g^k(0) \\ &\in -B_{k+1}E_k = -S_{k+1}. \end{aligned}$$

That $0 \in \Omega_{k+1}$ is clear.

(c) We prove the result by (finite) induction. By hypothesis, the result holds for $k=0$. So let us assume that the result holds for g^k and prove it for g^{k+1}. Let $x, y \in R^{n_{k+1}}$ and $0<t<1$. To show that g^{k+1} is S_{k+1}-convex we need to show that

$$tg^{k+1}(x) + (1-t)\, g^{k+1}(y) - g^{k+1}(tx+(1-t)\,y) \in S_{k+1},$$

or equivalently, by (a), that

$$B_{k+1}[tg^k(A_{k+1}x) + (1-t)\, g^k(A_{k+1}y) - g^k(A_{k+1}(tx+(1-t)\,y)) \in B_{k+1}E_k.$$

But this follows since

$$\begin{aligned} g^k(0) &\in -S_k^f \subset \phi_k^\perp, && \text{by (b) and (7.12);} \\ \mathscr{R}(A_{k+1}) &= D^=_{\phi_k Q_k g \circ P_k} = D^=_{\phi_k g_k}, && \text{by the induction step;} \\ & g^k \text{ is } S^k\text{-convex}, && \text{by the induction step;} \end{aligned}$$

and

$$E_k = S_k \cap \phi_k^\perp.$$

Moreover

$$d \in D^=_{\phi_{k+1} g^{k+1}}$$

iff

$$\phi_{k+1} g^{k+1}(d) = \phi_{k+1} g^{k+1}(0), \qquad \text{by faithful convexity,}$$

iff

$$\phi_{k+1} Q_{k+1} g(a + P_{k+1} d) = \phi_{k+1} Q_{k+1} g(a)$$

iff

$$\phi_{k+1} Q_{k+1} g(P_{k+1} d) = \phi_{k+1} Q_{k+1} g(0), \qquad \text{by faithful convexity,}$$

iff

$$d \in D^{=}_{\phi_{k+1} Q_{k+1} g \circ P_{k+1}}.$$

(d) Now $y \in F^{k+1}$ *implies*

$$y \in \Omega_{k+1}, \qquad g^{k+1}(y) \in -S^f_{k+1}$$

implies, since $A_{k+1} A^\dagger_{k+1} = P_{\mathscr{R}(A_{k+1})}$,

$$A_{k+1} y \in \Omega_k, \qquad B_{k+1} g^k(A_{k+1} y) \in -S^f_{k+1}$$

implies, since $B^\dagger_{k+1} B_{k+1} = P_{\phi_k^\perp}$,

$$A_{k+1} y \in \Omega_k, \qquad g^k(A_{k+1} y) \in -B^\dagger_{k+1} S^f_{k+1} = -E_k.$$

Thus $A_{k+1} y \in F^k$.

(e) Consider program (P_{k+1}). By Proposition 3.7(c), we know that there exists

$$\hat{y} \in \Omega_{k+1} \qquad \text{with} \quad g^{k+1}(\hat{y}) \in -\operatorname{ri} S^f_{k+1}.$$

This implies that

$$\hat{y} = A^\dagger_{k+1} \hat{z}, \qquad \text{for some} \quad \hat{z} \in \Omega_k \cap \mathscr{R}(A_{k+1}),$$

with

$$B_{k+1} g^k(A_{k+1} \hat{y}) \in -\operatorname{ri} S^f_{k+1}, \qquad \text{by (a),}$$

equivalently

$$\hat{z} \in \Omega_k \qquad \text{with} \quad B_{k+1} g^k(\hat{z}) \in -\operatorname{ri} S^f_{k+1}, \quad \text{since} \quad A_{k+1} A^\dagger_{k+1} = P_{\mathscr{R}(A_{k+1})},$$

equivalently

$$\hat{z} \in \Omega_k, \qquad B^\dagger_{k+1} B_{k+1} g^k(\hat{z}) \in -\operatorname{ri} B^\dagger_{k+1} S^f_{k+1}, \qquad \text{since } B_{k+1} \text{ is onto,}$$

equivalently

$$\hat{z} \in \Omega_k, \qquad g^k(\hat{z}) \in -\operatorname{ri} B^\dagger_{k+1} S^f_{k+1},$$

since $B^\dagger_{k+1} B_{k+1} = P_{\phi_k^\perp}$ and $\hat{z} = A_{k+1}\hat{y} \in D^=_{\phi_k g^k}$ by (c), yields $\phi_k g^k(\hat{z}) = 0$. Now $B^\dagger_{k+1} S^f_{k+1}$ is a face of E_k and thus a face of S_k. Therefore

$$S^f_k \supset B^\dagger_{k+1} S^f_{k+1}.$$

The converse inclusion follows since

$$\begin{aligned} g^k(F^k) &\subset g^k(A_{k+1}F^{k+1}) \\ &= B^\dagger_{k+1} B_{k+1} g^k(A_{k+1}F^{k+1}), \qquad \text{since} \quad B^\dagger_{k+1} B_{k+1} = P_{\phi_k^\perp} \\ &= B^\dagger_{k+1} g^{k+1}(F^{k+1}), \qquad \text{by (a)}, \\ &\subset -B^\dagger_{k+1} S^f_{k+1}. \end{aligned}$$

Moreover

$$d \in D^=_{g^{k+1}}(S^f_{k+1}, 0)$$

iff

$$g^{k+1}(\alpha d) \in S^f_{k+1} - S^f_{k+1}, \qquad 0 < \alpha \leqslant \bar{\alpha},$$

iff

$$B_{k+1} g^k(A_{k+1} \alpha d) \in S^f_{k+1} - S^f_{k+1}, \qquad 0 < \alpha \leqslant \bar{\alpha},$$

iff

$$B^\dagger_{k+1} B_{k+1} g^k(A_{k+1} \alpha d) \in B^\dagger_{k+1}(S^f_{k+1} - S^f_{k+1}), \qquad 0 < \alpha \leqslant \bar{\alpha},$$

iff

$$g^k(A_{k+1} \alpha d) \in B^\dagger_{k+1}(S^f_{k+1} - S^f_{k+1}), \qquad 0 < \alpha \leqslant \bar{\alpha},$$

iff

$$g^k(A_{k+1} \alpha d) \in S^f_k - S^f_k, \qquad 0 < \alpha \leqslant \bar{\alpha},$$

i.e.,

$$A_{k+1} d \in D^=_{g^k}(S^f_k, 0). \quad \blacksquare$$

THEOREM 7.2. *Suppose that $a \in F$ and g is weakly faithfully S-convex* (on Ω); i.e., *ϕg is faithfully convex for all $\phi \in Y^*$ for which ϕg is convex (on Ω). Then the above algorithm finds S^f* and *$D^=_g(S^f)$ in at most $t = \dim Y - \dim S^f + 1$ steps. Moreover, the program* (P_t) (see Lemma 7.2) *yields the regularized program of Theorem* 6.3.

Proof. There are two cases to consider.

(i) (7.8) is inconsistent at step t and $m_t > 0$. Now by (7.13), (7.14), Theorem 7.1, and Corollary 7.1, we get that Slater's condition is satisfied for (P_t) and

$$S_t^f = S_t; \qquad D_{g^t}^{=}(S_t^f) = R^{n_t}.$$

Therefore Lemma 7.2(e) yields the conclusion (7.10). Furthermore, Theorem 6.3 shows that (P_t) is the regularized program for (P).

(ii) $$m_t = 0.$$

By Lemma 7.1 and step $t-1$, we get that

$$S_{t-1}^f = 0.$$

As above the result still follows from Lemma 7.2(d). ∎

Remark 7.1. The algorithm will still work if g is not weakly faithfully S-convex. In this case we no longer can substitute the matrices P_i to get the equivalent programs and must modify the system (7.8) to read

$$\left(\bigcap_i D_{\phi_i Q_i g}^{=} \cap (\Omega - a)\right)^+ \cap \partial\phi_i Q_i g(a) \neq \varnothing,$$

$$\phi_i Q_i g(a) = 0, \qquad 0 \neq \phi_i \in S_i^+.$$

We restrict ourselves to the faithfully convex case as it seems preferable for applications. Recall that all analytic convex and strictly convex functions are faithfully convex. The algorithm may also be modified to use the notion of faithfully S-convex introduced in Section 5. In this case we find A_{i+1} so that $\mathscr{R}(A_{i+1}) = D_{Q_i g \circ P_i}^{=}(E_i)$. One may also choose B_{i+1} so that $\mathscr{N}(B_{i+1}) = (E_i - E_i)^\perp$. Both these changes speed up the algorithm.

Remark 7.2. Once S^f and $D_g^{=}(S^f)$ are found, we can apply Theorem 6.3 to get an equivalent program for which Slater's condition is satisfied. In fact, as seen above, (P_t) is the regularized program. Known methods for this case can now be applied (see, e.g., [14]). However, if the original optimal point was not a Kuhn–Tucker point, then stability problems may arise. Note that solving the complementarity problem (7.8) may also pose a problem. Robinson [19, 20] discusses an extension of Newton's method for cone constraints, while Tuy [23] presents an algorithm for the complementarity problem with nonpolyhedral constraints.

Remark 7.3. The above algorithm regularizes program (P) once a feasible starting point $a \in F$ is found. Finding a feasible starting point is itself a problem when Slater's condition fails. The case when $S = R_+^m$ was treated in [26]. The method there involves starting with all the constraints in the objective function and iterating while simultaneously removing any constraints which are satisfied from the objective function, using them as constraints again and regularizing. This is a modification of the standard (phase I) process of finding a feasible point. One can also modify the technique in [17] which takes the objective function into account while finding the feasible point. The algorithm in [26] seems intuitively clear though the proof was long and technical. In our case S is not polyhedral. Finding a feasible starting point then appears to be equivalent to discretizing the dual cone while applying the above mentioned technique in [26].

8. EXAMPLES

EXAMPLE 8.1. Let us consider the polyhedral case with $S = R^m_+$ and $\Omega = R^n$. In this case the algorithm is a modification of the one given in [26] which was a modification of the one in [1] for the faithfully convex case. The algorithm is also a modification of the one given in [4]. The following set of constraints are taken from [4].

Let $S = R^7_+$ and $g = (g_i)\colon R^5 \to R^7$, where

$$\begin{aligned}
g_1(x) &= e^{x_1} + x_2^2 - 1\\
g_2(x) &= x_1^2 + x_2^2 + e^{-x_3} - 1\\
g_3(x) &= x_1 + x_4^2 + x_5^2 - 1\\
g_4(x) &= e^{-x_2} - 1\\
g_5(x) &= (x_1 - 1)^2 + x_2^2 - 1\\
g_6(x) &= x_1 + e^{-x_4} - 1\\
g_7(x) &= x_2 + e^{-x_5} - 1.
\end{aligned}$$

Initialization. Let

$$a = (0, 0, 1, \tfrac{1}{2}\sqrt{2}, \tfrac{1}{2}\sqrt{2})^t.$$

By the complementary slackness condition in (7.8), we will only have to consider the binding constraints g_1, g_3, g_4, g_5 whose gradients are

$$\begin{aligned}
\nabla g_1(a) &= (1, 0, 0, 0, 0)\\
\nabla g_3(a) &= (1, 0, 0, \sqrt{2}, \sqrt{2})\\
\nabla g_4(a) &= (0, -1, 0, 0, 0)\\
\nabla g_5(a) &= (-2, 0, 0, 0, 0).
\end{aligned}$$

We also have $\Omega_0 = R^5$, $m_0 = 7$, $n_0 = 5$, and $S_0 = R^7_+$.

Step 0. The vector

$$\phi_0 = (\tfrac{2}{3}, 0, 0, 0, \tfrac{1}{3}, 0, 0)$$

solves (7.8). Now

$$\phi_0 Q_0 g \circ P_0 = \tfrac{2}{3} g_1 + \tfrac{1}{3} g_5.$$

Then

$$P_1 = A_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Q_1 = B_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$E_0 = \{s = (s_i) \in R^7_+ : s_1 = s_5 = 0\}$$

$$S_1 = \{s = (s_1) \in R^6_+ : s_4 = 0\}$$

$$\Omega_1 = R^3.$$

Step 1. The vector

$$\phi_1 = (0, 0, 1, 0, 0, 0)$$

solves (7.8). Now

$$\phi_1 Q_1 g \circ P_1 = g_4 \circ P_1.$$

Then

$$A_2 = I_{3\times 3}$$

$$P_2 = P_1$$

$$B_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$E_1 = \{s = (s_i) \in R^6_+ : s_3 = s_4 = 0\}$$

$$S_2 = \{s = (s_i) \in R^5_+ : s_3 = 0\}$$

$$\Omega_2 = R^3.$$

Step 2. The vector

$$\phi_2 = (0, 0, 1, 0, 0)$$

solves (7.8). Now

$$\phi_2 Q_2 g \circ P_3 = 0.$$

Then

$$A_3 = I_{3\times 3}$$

$$P_3 = P_2$$

$$B_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$Q_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$E_2 = S_2$$

$$S_3 = R^4_+$$

$$\Omega_3 = R^3.$$

Step 3. Let

$$\phi_3 = (\lambda_i) \in S_3^+.$$

Then (7.8) becomes

$$\lambda_i \geqslant 0, \qquad \text{not all zero,}$$

$$0 = \phi_3 Q_3 g(a) = (\lambda_1 g_2 + \lambda_2 g_3 + \lambda_3 g_6 + \lambda_4 g_7)(a)$$

$$0 \in [\partial\phi_3 Q_3 g(a)] P_3 = [(\lambda_1 \nabla g_2 + \lambda_2 \nabla g_3 + \lambda_3 \nabla g_6 + \lambda_4 \nabla g_7)(a)] P_3.$$

since g_2, g_6, and g_7 are not binding, the above is equivalent to

$$\lambda_1 = \lambda_3 = \lambda_4 = 0, \qquad \lambda_2 > 0$$
$$0 = [\lambda_2 \nabla g_3(a)] P_3$$

which is inconsistent. We now conclude by (7.10) that the minimal cone

$$\begin{aligned} S^f &= B_0^{\dagger} B_1^{\dagger} B_2^{\dagger} B_3^{\dagger} S_3 \\ &= B_1^t B_2^t B_3^t R_+^4 \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \{s = (s_i)\ R_+^6 : s_3 = s_4 = 0\} \\ &= \{s = (s_i) \in R_+^7 : s_1 = s_4 = s_5 = 0\}, \end{aligned}$$

while the cone of directions of constancy

$$\begin{aligned} D_g^{=}(D^f) &= \mathscr{R}(P_3) \\ &= \{d = (d_i) \in R^5 : d_1 = d_2 = 0\}. \end{aligned}$$

This coincides with the results found in [4]. Note that though $Q_3^{\dagger} \neq B_1^{\dagger} B_2^{\dagger} B_3^{\dagger}$ here, we still get that $Q_3^{\dagger} S_3 = S^f$. Moreover the constraint (mapping R^3 to R^4) in the regularized program ($\mathrm{P_r}$) is

$$\begin{aligned} Q_3\, g(a_0 + P_3 z) &= \begin{pmatrix} g_2(a + (0, 0, z_1, z_2, z_3)^t) \\ g_3(\quad \cdots \quad) \\ g_6(\quad \cdots \quad) \\ g_7(\quad \cdots \quad) \end{pmatrix} \\ &= \begin{pmatrix} e^{-z_1} & -1 \\ z_2^2 + z_3^2 & -1 \\ e^{-z_2} & -1 \\ e^{-z_3} & -1 \end{pmatrix}. \end{aligned}$$

Remark 8.1. Even though S is polyhedral, the above application of our algorithm differs from that in [1, 4, 26]. It is interesting to note that after solving (7.8) we find the cone of directions of constancy of the single convex functional $\phi_i Q_i g \circ P_i$, which is a linear combination of the convex functionals $(Q_i g \circ P_i)_j$. This differs from [1, 4, 26] where the intersection of the cones of directions of constancy are needed. In particular, if $\phi_0 = (\lambda_j)$ solves (7.8) at step 0, the above argument suggests that

$$D^{=}_{\phi \circ g} = \bigcap_j D^{=}_{\lambda_j g_j}. \tag{8.1}$$

However, this is not true in general as can be seen by considering the two linear functionals $g_1(x) = x_1 + x_2$ nd $g_2 = -g_1$ with $\lambda_1 = \lambda_2 = 1$. Let us see what happens for this example. (For more details see [27].)

EXAMPLE 8.2. Let $S = R^2_+$ and $g(x) = (g_i(x))$, where $g_1(x) = x_1 + x_2$ and $g_2 = -g_1$.

Initialization. Let

$$a = (0, 0)^t;$$

then

$$g_1(a) = -\nabla g_2(a) = (1, 1).$$

Step 0. The vector

$$\phi_0 = (1, 1)$$

solves (7.8). Now

$$\phi_0 Q_0 g \circ P_0 = 0.$$

Then

$$\begin{aligned} P_1 &= A_1 = I_{2\times 2} \\ Q_1 &= B_1 = [1 - 1] \\ E_0 &= 0 \qquad (\text{in } R^2) \\ S_1 &= 0 \qquad (\text{in } R^1). \end{aligned}$$

Step 1. The vector

$$\phi_1 = (1)$$

solves (7.8). Now

$$\phi_1 Q_1 g \circ P_1(x) = 2x_1 + 2x_2.$$

Then

$$P_2 = A_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
$$B_2 = [1]$$
$$Q_2 = [1 \ \ -1]$$
$$E_1 = 0 \qquad (\text{in } R^1).$$

Step 2. Since $m_2 = 0$, we conclude that the minimal cone

$$S^f = 0 \qquad (\text{in } R^2),$$

while the cone of directions of constancy

$$D_g^{=}(S^f) = \{d = (d_i) \in R^2 : d_1 = -d_2\}.$$

The constraint $(R^1 \to R^1)$ in the regularized program $(\mathrm{P_r})$ is

$$\begin{aligned} Q_2 g(a_0 + P_2 z) &= g_1((z_1, -z_1)^t) - g_2((z_1, -z_1)^t) \\ &= 0. \end{aligned}$$

Thus the program $(\mathrm{P_r})$ is the unconstrained program

$$\textit{minimize } p((z_1, -z_1)^t), \qquad z_1 \in R.$$

9. Conclusion

In this paper we have studied the abstract convex program (P) with the cone constraint $g(x) \in -S$ and set constraint $x \in \Omega$, where $S \subset Y$ is finite dimensional. We have presented several results on the faces of the convex cone S and have generalized known results on cones of directions and faithful convexity to the case of S-convex functions. We have then shown how to use these results to characterize optimality for (P). This follows similar results in [10]. In particular, in the case that g is weakly faithfully S-convex and Ω is polyhedral, we have seen that we can strengthen the characterization presented in [10] in the sense that the multiplier relationship holds on the larger sets K^H (see Theorem 6.2 and Remark 6.3).

In the faithfully S-convex case, the (generalized) cone of directions of constancy $D_g^{=}(S^f)$ is a subspace independent of x in X. In this case, we can regularize program (P) so that Slater's condition holds (see Theorem 6.3). The algorithm presented in Section 7 reduces program (P) to obtain this regularized program $(\mathrm{P_r})$. This algorithm is presented in the case that g is weakly faithfully S-convex; i.e., ϕg is a faithfully convex functional for each $\phi \in Y^*$ such that ϕg is convex (on Ω). This hypothesis holds for example whenever g is weakly analytic.

Many open questions still remain to be studied. Several known results for the cones of directions and for faithfully convex functions remain to be extended. It is hoped that these extensions will lead to new optimality criteria as well as stability results. For the strengthened characterization of optimality presented in Theorem 6.2, the question arises of finding H so that G^H is as large as possible. If S^f is exposed and g is weakly faithfully S-convex, this reduces to the question of finding $\phi \in (S^f)^+$ with $S \cap \Phi^{\perp} = S^f$ such that the subspace $D^{=}_{\phi g}$ has the largest dimension possible. In the case of the ordinary convex program with $S = R^m_+$, we want to find *positive* scalars α_k such that the subspace

$$D^{=}_{\sum_{k \in \mathscr{P}^=} \alpha_k g^k}$$

has the largest dimension possible [27] ($\mathscr{P}^=$ is the set of equality constraints [1].). Another question which arises is that: if the multiplier relationship in Theorem 6.1 holds with F^f replaced by Ω, can one always find H in Theorem 6.2 with $K^H = \Omega$; i.e., is the optimality criteria the strongest possible?

The examples given in Section 8 treat only the polyhedral case. In this case our algorithm simplifies the algorithm presented in [1] for finding $\mathscr{P}^=$, the equality set of constraints. This simplification is due to the fact that at each step we find only $D^{=}_{\phi g}$, which is the cone of directions of constancy of a single convex functional, rather than the intersection of cones of directions of constancy of several convex functionals. Substituting the matrices P_i, Q_i, thus reducing the dimensions of the image and domain spaces, also speeds up the algorithm. It still remains to study the algorithm in the nonpolyhedral case. The main question which arises is how to treat the complementarity problem (7.8) in the nonpolyhedral case. Tuy [23] discusses a fixed point algorithm that can be applied to the complementarity problem in this case. Stability of the algorithm with respect to round off and truncation errors may also pose a serious problem, especially when calculating $A_i^\dagger$ and $B_i^\dagger$ when A_i is not chosen to be of full column rank or A_i and B_i are ill-conditioned.

As mentioned in the introduction, applications for this theory include finding the unbiased nonnegative estimator in the "ice-cream" cone of nonnegative matrices and also the optimal control problem where the target is a finite dimensional convex set with empty interior. Further applications include the semi-infinite programming problem or polynomial approximation problem where the constraint g is a linear operator on a finite dimensional domain. In this case, though the range space and the cone S may not be finite dimensional, the image of the linear operator g and so also the minimal cone S^f are finite dimensional. This allows one to formulate such problems with Y chosen to be the (self-dual) Hilbert space of square integrable functions on the interval $[a, b]$ (denoted $L_2[a, b]$), where the nonnegative cone S has empty interior, rather than being restricted to choose Y as the space of continuous functions $C[a, b]$ so that S has nonempty interior. The Lagrange multipliers are then functions in $L_2[a, b]$ rather than measures in $C[a, b]^*$.

REFERENCES

1. R. A. ABRAMS AND L. KERZNER, A simplified test for optimality, *J. Optim. Theory Appl.* **25** (1978), 161–170.
2. G. P. BARKER, The lattice of faces of a finite dimensional cone, *Linear Algebra Appl.* **7** (1973), 71–82.
3. G. P. BARKER AND D. CARLSON, Cones of diagonally dominant matrices, *Pacific J. Math.* **57** (1975), 15–32.
4. A. BEN-ISRAEL, A. BEN-TAL, AND S. ZLOBEC, "Optimality in Nonlinear Programming: A Feasible Directions Approach," Wiley, New York, 1981.
5. A. BEN-ISRAEL, A. BEN-TAL, AND S. ZLOBEC, Optimality conditions in convex programming, *in* "The IX International Symposium on Mathematical Programming, Budapest, Hungary, 1976."
6. A. BEN-ISRAEL AND T. N. E. GREVILLE, "Generalized Inverses Theory and Applications," Wiley–Interscience, New York, 1973.
7. A. BEN-TAL AND A. BEN-ISRAEL, Characterizations of optimality in convex programming: The nondifferentiable case, *Appl. Anal.* **9** (1979), 137–156.
8. A. BEN-TAL, A. BEN-ISRAEL, AND S. ZLOBEC, Characterizations of optimality in convex programming without a constraint qualification, *J. Optim. Theory Appl.* **20** (1976), 417–437.
9. J. BORWEIN AND H. WOLKOWICZ, Characterizations of optimality without constraint qualification for the abstract convex program, Research Report No. 14, Dalhousie University, Canada, 1979, *Math. Programming*, in press.
10. J. BORWEIN AND H. WOLKOWICZ, Characterizations of optimality for the abstract convex program with finite dimensional range, *J. Austral. Math. Soc. Ser. A* **30** (1981), 390–411.
11. F. J. GOULD AND J. W. TOLLE, Geometry of optimality conditions and constraint qualifications, *Math. Programming* **2** (1972), 1–18.
12. B. GRUNBAUM, "Convex Polytopes," Wiley, London/New York/Sydney, 1967.
13. R. B. HOLMES, "Geometric Functional Analysis and Its Applications," Springer–Verlag, New York/Berlin, 1975.
14. D. G. LUENBERGER, "Optimization by Vector Space Methods," Wiley, New York, 1969.
15. H. MASSAM, Optimality conditions for a cone-convex programming problem, *J. Austral. Math. Soc. Ser. A* **27** (1979), 141–162.
16. F. PUKELSHEIM, "Linear Models and Convex Programs: Unbiased Nonnegative Estimation in Variance Component Models," Report 104, Stanford, 1977.
17. E. POLAK, R. TRAHAN, AND D. Q. MAYNE, Combined phase I–phase II methods of feasible directions," *Math. Programming* **17** (1979), 61–73.
18. M. PACHTER AND R. J. STERN, On state restrained linear quadratic control problems, *J. Math. Anal. Appl.* **69** (1979), 315–328.

19. S. M. ROBINSON, Extension of Newton's method to nonlinear functions with value in a cone, *Numer. Math.* **19** (1972), 341–347.
20. S. M. ROBINSON, "Strongly Regular Generalized Equations," Technical Summary No. 1877, University of Wisconson, 1978.
21. R. T. ROCKAFELLAR, Some convex programs whole duals are linearly constrained, *in* "Nonlinear Programming" (J. B. Rosen, O. L. Mangasarian, and K. Ritter, Eds.), pp. 293–322, Academic Press, New York, 1970.
22. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, N. J., 1972.
23. H. TUY, "Some Results on Fixed Point Algorithmic Theory," presented at the Tenth International Symposium on Mathematical Programming.
24. H. WOLKOWICZ, Calculating the cone of directions of constancy, *J. Optim. Theory Appl.* **25** (1978), 451–457.
25. H. WOLKOWICZ, Geometry of optimality conditions and constraint qualifications: The convex case, *Math. Programming* **19** (1980), 32–60.
26. H. WOLKOWICZ, The method of reduction in convex programming, Research Report No. 21, Dalhousie University, Canada, 1979.
27. H. WOLKOWICZ, A strengthened test for optimality, Research Report, The University of Alberta, Canada (1979), *J. Optim. Theory Appl.*, in press.
28. J. ZOWE, Subdifferentiability of convex functions with values in an ordered vector space, *Math. Scand.* **34** (1974), 69–83.