J. Austral. Math. Soc. (Series A) 30 (1981), 390–411

CHARACTERIZATION OF OPTIMALITY FOR THE ABSTRACT CONVEX PROGRAM WITH FINITE DIMENSIONAL RANGE

JON M. BORWEIN and HENRY WOLKOWICZ

(Received 21 March 1980; revised 28 July 1980)

Communicated by J. B. Miller

Abstract

This paper presents characterizations of optimality for the abstract convex program

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

when 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). These characterizations, which include a Lagrange multiplier theorem and do not presume any *a priori* constraint qualification, subsume those presently in the literature.

1980 *Mathematics subject classification* (*Amer. Math. Soc.*): primary 90 C 25.

Keywords: convexity, abstract programming, Lagrange multipliers, Slater's condition, subgradients, duality principles, facial reduction, symmetric matrices, cone of positive semi-definite matrices.

1. Introduction

In this paper we continue the study of the abstract convex program (P), for which the constraint has finite dimensional range. Massam (1979) used the faces of the cone S to present a characterization of optimality for (P). In Borwein and Wolkowicz (1980), we presented a corrected and strengthened version of this characterization and showed that this characterization does not yield a meaningful Lagrange multiplier relation. The main result of this paper, see Theorem 4.1, presents a Lagrange multiplier result which characterizes optimality for (P) without any constraint qualification. In particular, we show that the *BBZ* conditions, presented for ordinary convex programs in Abrams and Kerzner

©Copyright Australian Mathematical Society 1981

This research was partially supported by NSERC A4493 and NSERC A3388.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

(1978) and Ben–Israel and others (1976) are directly contained in our results, see Example 6.1. In section two we define our optimization problem (P) and present the necessary preliminaries. In section three we introduce the necessary facial notions and the concept of a minimal cone for (P). In section four we derive our central results. These results produce primal and dual characterizations of optimality for (P) with no constraint qualifications assumed, and include a strong duality theorem and a first order subgradient characterization. In section five we sketch extensions of our development to various concepts of vector optimization. Section six contains a decomposition principle for product orders and concludes with several examples. In particular we show that the *BBZ* conditions, presented for ordinary convex programs in Ben–Israel and others (1976) and Ben–Tal and Ben–Israel (1976) are directly contained in our results. Section seven applies our theory to an approximation problem involving symmetric matrices.

Our approach in this paper is to reduce (P) to an equivalent problem to which the "standard Lagrange multiplier theorem" can be applied. This reduction can be used whenever constraint regularization is desirable. This approach was used in Borwein and Wolkowicz (1979) for the case when S is not finite dimensional. Another approach to this problem is given in Craven and Zlobec (1980).

2. Preliminaries

In this paper we consider the convex programming problem

$$(2.1) \qquad (P) \qquad \mu = \inf\{p(x): g(x) \in -S, x \in \Omega\},$$

where $p\colon X \to \mathbf{R} \cup \{+\infty\}$, $g\colon X \to Y \cup \{+\infty\}$; X and Y are real locally convex (separated topological vector) spaces; Y is finite dimensional with an abstract maximal element $+\infty$ (Peressini (1967)); $\Omega \subset X$ and $S \subset Y$ are convex and moreover S is a *cone* (not necessarily closed), that is λs lies in S whenever s is in S and λ is non-negative; p is a convex (extended) functional and g is S-convex (on Ω), that is

$$(2.2) \qquad tg(x_1) + (1-t)g(x_2) - g(tx_1 + (1-t)x_2) \in S,$$

for any x_1, x_2 (in Ω) and any t in $[0, 1]$.

The set of x for which $g(x)$ is finite is the *domain* of g, $\operatorname{dom} g$. From (2.2) it follows that $\operatorname{dom} g$ is convex. As is well known a convex cone S induces an *ordering* $\geqslant_s$ on Y given by

$$(2.3) \qquad x_1 \geqslant_s x_2 \Leftrightarrow x_1 - x_2 \in S$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

which is transitive and reflexive. It is antisymmetric exactly when S is pointed, that is $S \cap -S = \{0\}$. It will be convenient to introduce the following notation

$$F = g^{-1}(-S), \qquad A = g^{-1}(-S) \cap \Omega. \tag{2.4}$$

Thus A is the feasible set for (P) and we make the additional routine assumption that

$$\operatorname{dom} p \supset A \neq \emptyset. \tag{2.5}$$

Let X^* and Y^* denote the continuous dual spaces of X and Y respectively. We suppose throughout that X^* is endowed with the weak star topology $\sigma(X^*, X)$ (see Robertson and Robertson (1964) for details).

Given any set K in X the *dual cone* of K is the set in X^*

$$K^+ = \{x' \in X^*: x'x \geqslant 0, \text{ if } x \in K\}. \tag{2.6}$$

Correspondingly if K' is in X^*

$$K'^+ = \{x \in X: x'x \geqslant 0, \text{ if } x' \in K'\}. \tag{2.7}$$

Then $K^+(K'^+)$ is always a closed convex cone and

$$K^{++} = (K^+)^+ \; ((K^+)^+) = \overline{\text{cone}\, K}, \tag{2.8}$$

where $\overline{\text{cone}\, K}$ denotes the closure of the convex cone generated by K. In particular, for a convex cone S, $S^{++} = \bar{S}$. The functionals in S^+ will be said to be positive (for S). Given any two convex cones S_1 and S_2 in X

$$(\bar{S}_1 \cap \bar{S}_2)^+ = \overline{S_1^+ + S_2^+}. \tag{2.9}$$

For proofs of these and other related results the reader is referred to Borwein (1978) and Holmes (1975). It will also be convenient to denote the annihilator of a set K in X

$$K^\perp = K^+ \cap (-K^+). \tag{2.10}$$

The *directional derivative* of g at a is defined by

$$\nabla g(a; d) = \lim_{t \downarrow 0} \frac{g(a + td) - g(a)}{t}. \tag{2.11}$$

Then $\nabla g(a; d)$ will exist for each direction d if g is convex on X, continuous at a and S is closed and pointed, Zowe (1974).

A continuous linear operator $T: X \to Y$ is a *subgradient* for g at a if

$$Td \leqslant_s g(a + d) - g(a) \quad \text{for all } d \text{ in } X. \tag{2.12}$$

The set of all such subgradients is denoted $\partial g(a)$. In case $Y = \mathbf{R}$, $S = \mathbf{R}_+$ these definitions reduce to the standard ones and so apply to p. It follows from a result of Zowe's (1974) and (1975a) that when X is a weakly compactly generated Banach space (for example X is reflexive or separable), and g is convex on X and continuous at a, with S closed pointed and convex, then $\partial g(a)$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

is non-empty and for any s^+ in S^+

$$s^+\nabla g(a; d) = \max_{T \in \partial g(a)} s^+T(d), \tag{2.13}$$

for each d in X. When $Y = \mathbf{R}$ and $S = \mathbf{R}_+$, then (2.13) holds true in any locally convex space X. Note that when $\nabla g(a; td)$ exists, then for $t \geqslant 0$

$$g(a + td) - g(a) - \nabla g(a; td) \in \bar{S}. \tag{2.14}$$

The reader is refered to Zowe (1974) and Borwein (1980) for more details.

Any other terms are, whenever possible, consistent with usage in Holmes (1975). We will use the symbol 0 for both the zero element and subspace of a vector space.

3. Facial reduction

Our results are based on identifying the smallest face of S which contains $-g(F)$. In this section we introduce the necessary facial notions and prove several preliminary results.

DEFINITION 3.1. (a) K is a *face* of a convex cone S if K is a convex cone, and

$$s_1, s_2 \in S, \quad s_1 + s_2 \in K \Rightarrow s_1, s_2 \in K. \tag{3.1}$$

(b) S^f denotes the (unique) smallest face of S which contains $-g(A)$.

Note that S^f is just the intersection of all the faces of S which contain $-g(A)$.

PROPOSITION 3.1. *If K is a face of S then*

$$(K - K) \cap S = K. \tag{3.2}$$

PROOF. Suppose $s = k_1 - k_2$ with s in S and k_1, k_2 in K. Then $s + k_2 = k_1$ is in K and by (3.1) s lies in K. The reverse inclusion is clear.

Recall that every convex set C in finite dimensions has non-empty *relative interior*, denoted ri C, which is the interior of C viewed as a subset of its affine span, Rockafellar (1970). The next lemma is fundamental to our subsequent results.

LEMMA 3.1. *If A is non-empty, then*

$$\text{(a)} \quad g(A) \cap -\operatorname{ri} S^f \neq \varnothing. \tag{3.3}$$

$$\text{(b)} \quad g(A) + S^f \quad \textit{is convex}. \tag{3.4}$$

PROOF. See Borwein and Wolkowicz (1980), Lemmas 3.1 and 3.2.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

The above lemma shows that when S has interior, then $S^f = S$ exactly when Slater's condition holds for (P), that is there exists

$$\hat{x} \in \Omega \quad \text{with } g(\hat{x}) \in -\operatorname{int} S. \tag{3.5}$$

where int denotes interior.

4. A general multiplier theorem

In this section we present the Lagrange multiplier characterization of optimality for (P), see Theorem 4.1. This characterization holds without any constraint qualification and is further used to derive primal and dual characterizations of optimality, see Corollaries 4.1 and 4.3. First, we need the following two definitions of cones.

DEFINITION 4.1.

(4.1) (a) $$D_g^{=}(a) = \{d\colon \exists \alpha(d) > 0 \text{ with } g(a + td) \in S^f - S^f \text{ if } 0 \leqslant t \leqslant \alpha(d)\}.$$

(4.2) (b) $$D_g^{\leqslant}(a) = \{d\colon \exists \alpha(d) > 0 \text{ with } g(a + td) \in S^f - S \text{ if } 0 \leqslant t \leqslant \alpha(d)\}.$$

If $Y = \mathbf{R}^m$ and $S = \mathbf{R}^m_+$, that is (P) is equivalent to the ordinary convex program with m convex inequality constraints, then (4.1) (resp. (4.2)) corresponds to the intersection of the cones of directions of constancy denoted $D^{=}_{\mathcal{P}^=}(a)$ (resp. non-increase denoted $D^{\leqslant}_{\mathcal{P}^=}(a)$) of the equality constraints, that is the constraints which are identically 0 on the feasible set. These cones are the cones employed in the BBZ conditions. These relationships are made explicit in section 6. We now have the following two propositions. Note that part (d) of Proposition 4.1 shows that $D^{=}_{\mathcal{P}^=}(a)$ in the BBZ conditions is always convex.

PROPOSITION 4.1. *Suppose that $a \in A$. Then*

(4.3) (a) $$g^{-1}(S^f - S) \cap \Omega = g^{-1}(S^f - S^f) \cap \Omega.$$

(b) *Hence*

$$D_g^{\leqslant}(a) \cap \operatorname{cone}(\Omega - a) = D_g^{=}(a) \cap \operatorname{cone}(\Omega - a). \tag{4.4}$$

(c) *Both $g^{-1}(S^f - S^f) \cap \Omega$ and $D_g^{=}(a) \cap \operatorname{cone}(\Omega - a)$ are convex.*

(d) *If $\Omega = X$, then $D_g^{\leqslant}(a) = D_g^{=}(a)$ and both are convex.*

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

PROOF. (a) Let x lie in $\Omega \cap g^{-1}(S^f - S)$. By Lemma 3.1, choose $\hat{x}$ in $\Omega \cap g^{-1}(-\text{ri } S^f)$. Then

$$g(x) = s_1^f - s_1 \quad \text{with } s_1^f \in S^f, s_1 \in S. \tag{4.5}$$

For $0 < \lambda < 1$, $x_\lambda = \lambda x + (1 - \lambda)\hat{x} \in \Omega$ and

$$\begin{aligned} g(\lambda x + (1 - \lambda)\hat{x}) &\in \lambda g(x) + (1 - \lambda)g(\hat{x}) - S, \quad \text{since } g \text{ is } S\text{-convex,} \\ &\subset \lambda s_1^f + (1 - \lambda)g(\hat{x}) - (S + \lambda s_1). \end{aligned}$$

For small λ, $\lambda s_1^f + (1 - \lambda)g(\hat{x})$ lies in $-\text{ri } S^f$ and hence in $-S$. Thus x_λ is feasible and $g(x_\lambda)$ lies in $-S^f$. Now, again by S-convexity of g, for small λ

$$g(x) \in \lambda^{-1}g(x_\lambda) - \lambda^{-1}(1 - \lambda)g(\hat{x}) + S \subset S - S^f. \tag{4.6}$$

Thus by (4.5) and (4.6)

$$g(x) = s_1^f - s_1 = s_2 - s_2^f, \quad s_1, s_2 \in S;\ s_1^f, s_2^f \in S^f.$$

Hence

$$s_1 + s_2 = s_1^f + s_2^f \in S^f$$

and since S^f is a face of S, s_1 lies in S^f. Thus by (4.5), $g(x)$ lies in $S^f - S^f$ which yields the desired containment. The reverse containment is immediate.

(b) It suffices to show that

$$D_g^{\leqslant}(a) \cap \text{cone}(\Omega - a) \subset D_g^{=}(a) \cap \text{cone}(\Omega - a).$$

Now if d lies in $D_g^{\leqslant}(a) \cap \text{cone}(\Omega - a)$, there is some $\delta > 0$ with

$$g(a + td) \in S^f - S, \qquad a + td \in \Omega, \text{ for all } 0 \leqslant t < \delta.$$

By part (a)

$$g(a + td) \in S^f - S^f, \qquad a + td \in \Omega \text{ for all } 0 \leqslant t < \delta,$$

and d lies in $D_g^{=}(a) \cap \text{cone}(\Omega - a)$.

(c) $g^{-1}(S^f - S)$ is convex since $S^f - S$ is a convex cone and g being S convex is $S - S^f$ convex. Thus $g^{-1}(S^f - S) \cap \Omega$ is convex as is $D_g^{\leqslant}(a) \cap \text{cone}(\Omega - a)$ which is just the smallest cone containing $(g^{-1}(S^f - S) \cap \Omega) - a$.

(d) is now immediate.

PROPOSITION 4.2. *The constraint g is S^f-convex on*

$$A^f = \Omega \cap g^{-1}(S^f - S).$$

PROOF. Let x_1, x_2 lie in A^f. Let $0 \leqslant t \leqslant 1$. Then $x_t = tx_1 + (1 - t)x_2$ lies in A^f and so by (4.3)

$$g(x_1), g(x_2), g(x_t) \text{ lie in } S^f - S^f.$$

Thus

$$tg(x_1) + (1 - t)g(x_2) - g(x_t) \in (S^f - S^f).$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

Again as g is S-convex on Ω

$$tg(x_1) + (1 - t)g(x_2) - g(x_t) \in (S^f - S^f) \cap S = S^f$$

as claimed.

We can now present our central result.

THEOREM 4.1. (a) *Suppose that* μ *is the finite optimal value of* (P) *given by* (2.1). *Then*

$$p(x) + \lambda g(x) \geqslant \mu, \quad \text{for all } x \in \Omega \cap g^{-1}(S^f - S) \tag{4.7}$$

for some λ *in* $(S^f)^+$.

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

$$\lambda g(a) = 0 \tag{4.8}$$

and (4.7) *and* (4.8) *characterize optimality of* a *in* A.

PROOF. We rewrite (P) as (P^f)

$$\mu = \inf\{p(x): g(x) \in -S^f, x \in \Omega \cap g^{-1}(S^f - S)\}, \tag{P^f}$$

which clearly has the same optimum as (P) since

$$A \subset g^{-1}(-S^f) \cap A^f \subset g^{-1}(-S) \cap \Omega \subset A. \tag{4.9}$$

Now g considered as a mapping g^f defined only on A^f is S^f-convex and maps A^f into $S^f - S^f$, by Propositions 4.1 and 4.2. Then if $\hat{x}$ lies in A and $g(\hat{x})$ lies in $-\text{ri } S^f$, as promised by Lemma 3.1, $\hat{x}$ is a *Slater point* for

$$\mu = \inf\{p(x): g^f(x) \in -S^f \text{ and } x \in \Omega \cap g^{-1}(S^f - S)\}. \tag{4.10}$$

Both (a) and (b) now follow from the standard Lagrange multiplier theorem (see for example Holmes (1975) and Luenberger (1969)) on extending λ arbitrarily from $S^f - S^f$ to Y.

We now list two further characterizations of optimality of a feasible point a in A. Recall that $D_g^{\leqslant}(a) \cap \text{cone}(\Omega - a) = D_g^{=}(a) \cap \text{cone}(\Omega - a)$, by (4.4).

COROLLARY 4.1. *Suppose that both* g *and* f *have finite directional derivatives at the feasible point* a *in* A. *Then*

(i) *a is optimal for (P) if and only if*

(ii) *there is some λ in $(S^f)^+$ with*

$$\nabla p(a; d) + \lambda(\nabla g(a; d) + g(a)) \geqslant 0 \tag{4.11}$$

for all d in $D_g^{\leqslant}(a) \cap \text{cone}(\Omega - a)$, if and only if

$$\text{(iii)} \qquad \nabla p(a; d) < 0 \tag{4.12}$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

$$\nabla g(a; d) + g(a) \in -S^f \tag{4.13}$$

$$d \in D_g^{\leqslant}(a) \cap \operatorname{cone}(\Omega - a) \tag{4.14}$$

has no solution d in X.

PROOF. (i) $\Rightarrow$ (ii) follows from (4.7) and (4.8) on taking directional derivatives and observing once more that $\operatorname{cone}(A^f - a) = D_g^{\leqslant}(a) \cap \operatorname{cone}(\Omega - a)$.

(ii) $\Rightarrow$ (iii) is immediate.

(iii) $\Rightarrow$ (i). If $a \in A$ is not optimal one can find d with $p(a + d) < p(a)$ and $a + d$ in A. Then as g is S^f-convex on A

$$\nabla g(a; d) + g(a) \in -\overline{S^f} + g(a + d) \subset -S^f \tag{4.15}$$

and

$$\nabla p(a; d) \leqslant p(a + d) - p(a) < 0. \tag{4.16}$$

Finally that d satisfies (4.14) is clear. A little more work allows one to replace $\overline{S^f}$ by $\operatorname{ri} S^f$.

In certain cases the multiplier λ in (4.7) may be supposed to be in S^+, as in the standard Lagrange multiplier theorem, and not just in the larger cone $(S^f)^+$. We now characterize this situation.

COROLLARY 4.2. *Theorem* 4.1 *and Corollary* 4.1 *hold, with S^+ replacing $(S^f)^+$, for all linear (S-convex) g exactly when*

$$S^+ + (S^f)^{\perp} = (S^f)^+ \tag{4.17}$$

or equivalently, when S is closed, if

$$S^+ + (S^f)^{\perp} \text{ is closed.} \tag{4.18}$$

PROOF. (4.17) and (4.18) are equivalent by virtue of (2.9) and Proposition 3.1. Now if (4.17) holds and λ satisfies (4.7) and (4.8) one can solve $\lambda = s^+ + s^{\perp}$ with s^+ in S^+, $s^{\perp}$ in $(S^f)^{\perp}$. Hence for any x in A^f

$$\lambda g(x) = s^+ g(x) + s^{\perp} g(x) = s^+ g(x) \tag{4.19}$$

since $g(A^f) \subset S^f - S^f$. Thus $s^+ g(x)$ may be substituted for $\lambda g(x)$ in (4.6) and (4.7).

Conversely, suppose ϕ lies in $(S^f)^+$ and let P be the orthogonal projection of Y on $S^f - S^f$. Consider

$$(\hat{P}) \qquad \mu = \inf\{\phi P(y) : -Py \in -S \text{ and } y \in Y\}. \tag{4.20}$$

Then $PP^{-1}(S) = PP^{-1}(S^f) = S^f$ and so $\mu = 0$. Also $-P^{-1}(S^f - S^f) = Y$ so that (4.7) yields

$$\phi Py + \lambda(-Py) \geqslant 0, \quad \text{for all } y \in Y. \tag{4.21}$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

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

$$\phi = (\phi - \lambda)(I - P) + \lambda \in (S^f)^\perp + S^+. \tag{4.22}$$

Since ϕ is an arbitrary member of $(S^f)^+$ we see that (4.17) holds.

Corollary 4.2 shows that multipliers in S^+ exist whenever S is polyhedral since this ensures that (4.17) holds, Rockafellar (1970).

EXAMPLE 4.1. Let $S_1 = \{(x, y, z)\colon 2xy \geqslant z^2,\ x \geqslant 0,\ y \geqslant 0\}$ and $S_2 = \operatorname{cone}\{S_1 \cup (1, 0, 1)\}$.

(a) When $S = S_1$, there are three possibilities for S_1^f in (P).

(i) Slater's condition holds and $S_1 = S_1^f$, $S_1^f - S_1^f$ is $\mathbf{R}^3$. This is standard;

(ii) S_1^f is a boundary ray and $S_1^f - S_1^f$ is a single line;

(iii) $S_1^f = \{0\}$ whence $g(A)$ is empty or $\{0\}$ while A^f is also either empty or $\{0\}$. Thus only case (ii) is interesting. Although S_1 is not polyhedral $S_1^+ + (S_1^f)^\perp$ is always closed and Corollary 4.2 is applicable. The closure of $S_1^+ + (S_1^f)^\perp$ can be checked directly. It also follows from the observation that the projection of $\mathbf{R}^3$ on $S_1^f - S_1^f$ sends S_1 to S_1^f.

(b) When $S = S_2$ it is possible for (4.17) to fail. Let us illustrate this with

$$\mu = \min\{z\colon (-x, 0, -z) \in -S_2\}. \tag{P_2}$$

This satisfies (2.1) with

$$p = (0, 0, 1), \quad g = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad \Omega = \mathbf{R}^3.$$

An examination of S_2, shows that $g(g^{-1}(-S_2)) = -S_2^f$ is the convex cone generated by (1, 0, 1) and (1, 0, 0). $S_2^f - S_2^f$ is the plane $y = 0$ and in this plane z is non-negative. Thus $\mu = 0$ and (4.7) yields λ in $(S_2^f)^+$ with

$$z + \lambda(-x, 0, z) \geqslant 0, \quad \text{for all } (x, y, z) \in g^{-1}(S_2^f - S_2^f) = \mathbf{R}^3. \tag{4.23}$$

Now solving for $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ in (4.23) yields $\lambda_1 = 0$, $\lambda_3 = 1$ and so $\lambda = (0, \lambda_2, 1)$ for some λ_2 in $\mathbf{R}$. A possible λ in $(S_2^f)^+$ is (0, 0, 1). There are no possibilities in $(S_2)^+$. Indeed as $S_1 \subset S_2$

$$(S_2)^+ \subset S_1^+ = S_1$$

since S_1^+ is self dual. But $(0, \lambda_2, 1) \in S_1$ implies $0 = 0\lambda_2 \geqslant (1)^2$ which is impossible.

REMARK 4.1. (i) As in the previous example, if there is any projection of Y on $S^f - S^f$ which sends S to S^f, positive multipliers (in S^+) will exist and (4.17) will hold.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

(ii) Although (4.17) will not hold in general there will generally be an asymptotic version of (4.7) and (4.8). Indeed one has, for x in A^f and λ in $(S^f)^+$,

$$\lambda g(x) = \lim_{n\to\infty} s_n^+ g(x) \tag{4.24}$$

for appropriate s_n^+ in S^+. Indeed when S is closed

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

always holds.

Corollary 4.3. *Suppose that* (4.17) *holds. Suppose also that S is closed and pointed with p and g continuous at a and g S-convex on X with X a weakly compactly generated Banach space. Then $a \in A$ is optimal if and only if*

$$0 \in \partial p(a) + s^+ \partial g(a) - \left(D_g^<(a) \cap \operatorname{cone}(\Omega - a)\right)^+ \tag{4.25}$$

for some s^+ in S^+ with $s^+ g(a) = 0$.

Proof. Since (4.17) holds it follows from (4.7) and (4.8) that for some s^+ in S^+

$$0 \in \partial(p + s^+ g + i_{A^f})(a) \tag{4.26}$$

where $i_{A^f}(x)$ is the indicator function of A^f: that is $i_{A^f}(x)$ is 0 in A^f and $+\infty$ elsewhere. Since p and $s^+ g$ are continuous at a and i_{A^f} is finite at a

$$0 \in \partial p(a) + \partial(s^+ g)(a) - \left(D_g^<(a) \cap \operatorname{cone}(\Omega - a)\right)^+ \tag{4.27}$$

since $(D_g^<(a) \cap \operatorname{cone}(\Omega - a))^+$ is $-\partial i_{A^f}(a)$ (see Rockafellar (1970a), Corollary 23.5.4). It remains to observe that

$$\partial(s^+ g)(a) = s^+ \partial g(a).$$

This follows from the central result in Zowe (1974) or (2.13).

It is possible to formulate a version of Corollary 4.3 without assuming (4.17) to hold. One may do that by restricting attention entirely to g^f rather than g. We can now easily establish the following duality theorem. Define L^f, the *restricted Lagrangian* by

$$L^f(\lambda) = \inf\{p(x) + \lambda g(x) : x \in A^f\} \tag{4.28}$$

and the restricted dual problem (D^f) by

$$(D^f) \qquad d = \sup\{L^f(\lambda) : \lambda \in (S^f)^+\}. \tag{4.29}$$

Then (D^f) is a concave optimization problem and one obtains

Theorem 4.2. (Restricted Strong Duality). *When (P) has a finite optimum,*

$$\inf\{p(x) : g(x) \in -S,\ x \in \Omega\} = \max\{L^f(\lambda) : \lambda \in (S^f)^+\}. \tag{4.30}$$

Moreover, if (4.17) *holds, $(S^f)^+$ may be replaced by S^+ in* (4.29) *and* (4.30).

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

PROOF. It suffices to observe that for any λ in $(S^f)^+$ and any x in A

$$L^f(\lambda) \leqslant p(x). \tag{4.31}$$

Thus $\mu \geqslant d$. Now Theorem 4.1 guarantees the existence of λ in $(S^f)^+$ with $L^f(\lambda) \geqslant \mu$. This yields (4.30). The final conclusion follows from Corollary 4.2.

It seems worth noting that while weaker constraint qualifications can be used to derive Lagrange multiplier theorems, Slater's condition is the weakest consistent with a bounded set of Lagrange multipliers, Gauvin and Tolle (1977) and Zowe and Kurcyusz (1978). We only prove the following:

COROLLARY 4.4. *Let Λ^f denote the (non-empty) solution set for (D^f). Then Λ^f considered as a subset of $(S^f - S^f)^*$ is compact.*

PROOF. It is easily verified that Λ^f is closed. Now let $\hat{x} \in A$ with $g(\hat{x}) \in -\mathrm{ri}\, S^f$. Let $N_0 = \{x \in S^f - S^f\colon \|x\| < \delta\}$, where $\delta > 0$ and $\|\cdot\|$ is Euclidean norm, be a bounded symmetric neighbourhood of 0 in $S^f - S^f$ with $g(\hat{x}) + N_0 \subset -S^f$. Then $g(\hat{x}) + S^f$ contains N_0 and it follows that

$$\rho(\hat{x}) + \lambda(n_0) \geqslant \mu$$

for any λ in Λ^f and n_0 in N_0. Since N_0 is symmetric

$$\|\lambda\|_{N_0} = \sup_{n_0 \in N_0} |\lambda(n_0)| \leqslant \rho(\hat{x}) - \mu. \tag{4.32}$$

Now $\|\cdot\|_{N_0}$ is a norm on $(S^f - S^f)^*$ and as $S^f - S^f$ is finite dimensional, induces the original topology on $S^f - S^f$. Thus Λ^f is bounded.

Finally let us observe that the considerations of this section apply immediately to programs in which g and Ω are convex but p is not. One such case, in which p is quasiconvex, has been studied in Luenberger (1968) and Borwein (1977).

5. Vector optimization

The substance of section four involves only the constraint structure and not the objective function. It is relatively straight forward to extend the results of the last section to various types of vector optimization. We suppose $p\colon X \to Z$ is K-convex where Z is a partially ordered topological vector space ordered by a convex cone K. We consider three types of vector optimum and sketch the extensions.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

DEFINITION 5.1. p we say has a *weak minimum* over A at a in A if

$$A \cap \{x: p(x) - p(a) \in -\operatorname{int} K\} = \varnothing. \tag{5.1}$$

In this case one may separate the convex sets $p(A) + K$ and $p(a) - \operatorname{int} K$ by some non-zero functional k^+ in K^+ (assuming of course that int K is non-empty). It follows that

$$(P_k^+) \qquad \min\{k^+p(x): g(x) \in -S, x \in \Omega\} \tag{5.2}$$

is an equivalent scalar problem to which the results of section four may be applied. Note also that any optimum for (P_{k^+}) is automatically a weak-optimum for p over A. The reader is referred to Borwein (1980a) or Penot (1978) for further details.

DEFINITION 5.2. We say p has a *Pareto minimum* or *efficient solution* over A at a in A if

$$A \cap \{x: p(x) - p(a) \in -K \setminus 0\} = \varnothing. \tag{5.3}$$

If int K is non-empty any Pareto optimum is a weak optimum. If K has a weakly-compact base and a is a *proper efficient point*, Geoffrion (1968), Penot (1978), Borwein (1977a), Borwein (1980a), then there is a *strictly positive* multiplier k^+ in K^+ (that is $k^+(k) > 0$ if $k \neq 0 \in K$) such that a minimizes (P_{k^+}), given by (5.2). Again, conversely, any minimum of (P_{k^+}) with k^+ strictly positive yields a proper efficient point for p over A.

DEFINITION 5.3. We say p has a *strong minimum* over A at a if

$$A \subset \{x: p(x) \in p(a) + K\}. \tag{5.4}$$

Clearly any strong minimum is a Pareto minimum. If K is an order complete convex cone one may apply the duality results in Zowe (1975b) or Borwein (1980b) to

$$(VP) \qquad p(a) = \min_K\{p(x): g(x) \in -S^f, x \in A^f\} \tag{5.5}$$

and derive the existence of a continuous linear operator T mapping Y into Z such that

$$p(x) + Tg(x) \geqslant_K p(a) \quad \text{for all } x \in A^f \tag{5.6}$$

where $T(S^f) \subset K$ and $Tg(z) \in K \subset -K$.

Note that in this case the continuity of T follows trivially since Y is finite dimensional and hence K need not be assumed normal, Peressini (1967). In this case, as distinct from the previous two, it is critical that (VP) actually has a Slater-point and does not just satisfy a weaker constraint qualification.

If p and g are continuous at a, (5.6) may be replaced by

$$0 \in \partial p(a) + \partial(Tg)(a) - \left(D_g^{\leqslant}(a) \cap \operatorname{cone}(\Omega - a)\right)^K \tag{5.7}$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

when K is normal and C^K denotes the continuous linear operators mapping C into K.

6. Decomposition of product orderings

A generalization of the BBZ conditions, Ben–Israel and others (1976), will be exhibited below. We begin with a preliminary series of results about product orderings.

Consider g, S, Y, A as before but now specialized so that g, Y and S are finite products:

$$g = \prod_{\mathcal{P}} g_i; \qquad g_i \colon X \to Y_i \cup \{+\infty\}; \tag{6.1}$$

$$Y = \prod_{\mathcal{P}} Y_i; \qquad S = \prod_{\mathcal{P}} S_i.$$

Thus g is S-convex exactly when each g_i is S_i-convex and so on. To make the identification complete

$$Y \cup \{+\infty\} = \prod_{\mathcal{P}} (Y_i \cup \{+\infty\}) \tag{6.2}$$

as long as any product elements with an infinite entry are identified. Let $\mathcal{P}$ have finite cardinality m.

Definition 6.1.

$$S_i^f = (-g_i(A))^f. \tag{6.3}$$

That is: S_i^f is the minimal face of S_i containing $-g_i(A)$, which necessarily lies in S_i.

Proposition 6.1. *When A is non-empty*

$$\prod_{\mathcal{P}} S_i^f = S^f. \tag{6.4}$$

Proof. Since the product of faces is a face, it is clear that $\prod_{\mathcal{P}} S_i^f$ is a face of S containing $g(A)$ and hence S^f. To show the converse containment it suffices to show that

$$\operatorname{ri}\Big(\prod_{\mathcal{P}} S_i^f\Big) \cap S^f \neq \varnothing \tag{6.5}$$

as containment will follow from the extremality of S^f. Analogous reasoning to Lemma 3.1, applied to g_i and S_i on A, shows that since g_i is S_i^f-convex on A one

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

can solve

$$\{g_i(a_i)\} \cap -\text{ri } S_i^f \neq \varnothing, \qquad a_i \in A. \tag{6.6}$$

Now set $\hat{a} = 1/m\Sigma_{k\in\mathcal{P}}a_k$. Then for each i in $\mathcal{P}$

$$\begin{aligned} g_i(\hat{a}) &\in \frac{1}{m}\sum_{k\in\mathcal{P}} g_i(a_k) - S_i^f \quad (\text{by } S_i^f\text{-convexity}),\\ &\subset \frac{1}{m}g_i(a_i) - S_i^f \subset -\text{ri } S_i^f - S_i^f, \quad \text{by (6.6)},\\ &= -\text{ri } S_i^f. \end{aligned}$$

Thus

$$g(\hat{a}) \in -\prod_{\mathcal{P}} \text{ri } S_i^f = -\text{ri}\Big(\prod_{\mathcal{P}} S_i^f\Big). \tag{6.7}$$

But since $g(\hat{a})$ lies in $-S^f$, (6.7) establishes (6.5).

Defining $D_{g_i}^{\leqslant}(a)$, $D_{g_i}^{=}(a)$ analogously with $D_g^{\leqslant}(a)$, $D_g^{=}(a)$ but with S_i replacing S throughout (4.1) and (4.2) we gain a decomposition result.

THEOREM 6.1. *Let a lie in A.*

$$\text{(a)} \quad D_g^{=}(a) = \bigcap_{\mathcal{P}} D_{g_i}^{=}(a); \tag{6.8}$$

$$\text{(b)} \quad D_g^{\leqslant}(a) = \bigcap_{\mathcal{P}} D_{g_i}^{\leqslant}(a). \tag{6.9}$$

PROOF. We prove only (a). Now d lies in $D_g^{=}(a)$ exactly when for some $\bar{\alpha} > 0$ and all $0 \leqslant \alpha < \bar{\alpha}$

$$\begin{aligned} g(a+\alpha d) &\in S^f - S^f\\ &= \prod_{\mathcal{P}} S_i^f - \prod_{\mathcal{P}} S_i^f, \quad \text{by (6.3)},\\ &= \prod_{\mathcal{P}} (S_i^f - S_i^f), \end{aligned}$$

whence $g_i(a+\alpha d)$ lies in $S_i^f - S_i^f$ as desired. The opposite containment is identical since $\mathcal{P}$ is finite.

One may combine (6.3), (6.8), (6.9) and

$$(S^f)^+ = \prod_{\mathcal{P}} (S_i^f)^+ \tag{6.10}$$

to decompose the results in the previous sections. We will examine this for a special case which includes the *BBZ* conditions and Slater's condition.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

EXAMPLE 6.1. In (6.1) let $\mathcal{P} = \{0, 1, \ldots, m-1\}$. Let Y_0 be a finite dimensional space and let Y_i $(1 \leqslant i \leqslant m-1)$ be $\mathbf{R}$ and S_i $(1 \leqslant i \leqslant m-1)$ be $\mathbf{R}_+$. Suppose A is non-empty. Then for $1 \leqslant i \leqslant m-1$, either

$$g_i(A) = 0 \quad \text{and} \quad S_i^f = 0 \tag{6.11}$$

or

$$g_i(A) \neq 0 \quad \text{and} \quad S_i^f = \mathbf{R}_+. \tag{6.12}$$

The indices for which (6.11) holds coincide with the set of equality constraints $\mathcal{P}^=$ used in the *BBZ* conditions. Indeed if i is in $\mathcal{P}^=$, then

$$D_{g_i}^=(a) = \{d\colon \exists\alpha(d) > 0, g_i(a + \alpha d) = g_i(a), 0 \leqslant \alpha \leqslant \alpha(d)\}; \tag{6.13}$$

$$D_{g_i}^{\leqslant}(a) = \{d\colon \exists\alpha(d) > 0, g_i(a + \alpha d) \leqslant g_i(a), 0 \leqslant \alpha \leqslant \alpha(d)\}; \tag{6.14}$$

which coincide respectively with the *cones of constancy* and *nonincrease* defined in Ben-Israel and others (1976). If i is not in $\mathcal{P}^=$, it is immediate that

$$D_{g_i}^=(a) = D_{g_i}^{\leqslant}(a) = \operatorname{cone}(\operatorname{dom} g_i - a) \tag{6.15}$$

as $S_i^f - S_i^f = Y_i = \mathbf{R}$ in this case. Notice also that if

$$g_0(A) \cap \operatorname{int} S_0 \neq \varnothing \quad \text{(Slater's condition)} \tag{6.16}$$

then similarly

$$D_{g_0}^=(a) = D_{g_0}^{\leqslant}(a) = \operatorname{cone}(\operatorname{dom} g_0 - a). \tag{6.17}$$

Indeed in the case $Y_i = \mathbf{R}$, $S_i = \mathbf{R}_+$ either (6.11) holds or Slater's condition does. For general S this dichotomy fails. Suppose that a is interior to dom g_i for each i in $\mathcal{P}$. Then supposing (6.16) holds we derive from (4.17), (4.25) and our discussion:

$$a \text{ in } A \text{ is optimal for } (P) \tag{6.18}$$

is equivalent to

$$0 \in \partial f(a) + \partial\lambda_0 g_0(a) + \sum_{i=1}^{m-1} \lambda_i \partial g_i(a) + \Big(\bigcap_{\mathcal{P}^=} D_{g_i}^{\leqslant}(a) \cap \operatorname{cone}(\Omega - a)\Big)^+ \tag{6.19}$$

for some λ_0 in S_0^+ and $\lambda_1, \ldots, \lambda_{m-1}$ non-negative. Letting Ω be X and g_0 some constant in $-\operatorname{int} S_0$ we have rederived the dual *BBZ* conditions. Even here we have an improvement in that our development allows descent cones rather than just equality cones in (6.19) (and avoids the redundant use of conv $\bigcap_{\mathcal{P}^=} D_{g_i}^=(a)$). Letting $m = 1$, (6.19) reduces to the classical Lagrange multiplier theorem. Note that it is an immediate consequence of (6.4) that when all the cones involved are either polyhedral or satisfy Slater's condition on A, then (4.17) holds and multipliers exist in S^+. Note also that S^f, $S^f - S$ and $S^f - S^f$ have overt descriptions as follows from (6.4), (6.11), (6.12) and (6.16).

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

The primal characterization for the ordinary convex program can also be specialized as above. Moreover, unlike Ben–Israel et al. (1976) our results still apply if in (P) the optimum is not attained. We may still use (6.4) to decouple the constraint set A^f.

The descent algorithms suggested in Ben–Israel and others (1976) for the ordinary convex program have immediate application to (6.19). Indeed, in this case it is possible to compute $\mathcal{P}^{=}$, Abrams and Kerzner (1978), effectively and, with suitable restrictions on g, also the relevant cones of decrease or constancy.

EXAMPLE 6.2. Proposition 6.1 and Theorem 6.1 allow one to derive the previous results for arbitrary convex Ω from the case $\Omega = X$. Indeed the constraint set A given by

$$g_0(x) \in -S_0, \qquad x \in \Omega \tag{6.20}$$

can be rewritten as

$$g(x) = (g_0(x), g_1(x)) \in -(S_0, S_1) = S \tag{6.21}$$

where $g_1\colon X \to \mathbf{R} \cup \{+\infty\}$ is the indicator function of Ω and $S_1 = \mathbf{R}_+$. Then from the definitions

$$S_0^f = \left(-g_0\left(g^{-1}(-S)\right)\right)^f = \left(-g_0(A)\right)^f \quad (\text{in } S_0), \tag{6.22}$$

$$S_1^f = \left(-g_1(A)\right)^f = (0)^f = 0 \quad (\text{in } S_1). \tag{6.23}$$

Now $D_{g_0}^{=}(a)$, $D_{g_0}^{\leqslant}(a)$ are exactly as before (with g_0, S_0 replacing g, S) while

$$\begin{aligned} D_{g_1}^{=}(a) = D_{g_1}^{\leqslant}(a) &= \{d\colon \exists \alpha(d) > 0, g_i(a + \alpha d) \leqslant 0, 0 < \alpha < \alpha(d)\} \\ &= \operatorname{cone}(\Omega - a), \end{aligned} \tag{6.24}$$

since g_1 is zero on Ω and $+\infty$ elsewhere. Hence

$$D_g^{=}(a) = D_{g_0}^{=}(a) \cap \operatorname{cone}(\Omega - a) \tag{6.25}$$

and this intersection is convex. Thus we could have developed the results of the previous sections only for $\Omega = X$ and then substituted (6.25) and other similar formulae to reach the general case.

Equally if Ω is closed in a normed space, one can set

$$g_1(x) = \inf\{\|x - \omega\|\colon \omega \in \Omega\} \tag{6.26}$$

and still derive (6.24), (6.25) while keeping g_1 everywhere finite.

EXAMPLE 6.3. Consider (P) with $X = Y = \mathbf{R}^2$, $S = \mathbf{R}_+^2$ and $\Omega = X$. Set

$$g(x_1, x_2) = \left(x_2^2 + x_1 - 1,\ 1 - x_1\right). \tag{6.27}$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

Then $A = g^{-1}(-S) = \{(1, 0)\}$, $S^f = 0$ and

(6.28) $$D_g^{\leqslant}(1, 0) = \{(d_1, d_2): \exists\alpha(d) > 0, \alpha^2 d_2^2 + \alpha d_1 \leqslant 0, -\alpha x_1 \leqslant 0, \text{ for } 0 \leqslant \alpha \leqslant \alpha(d)\} = 0 = D_g^{=}(1, 0).$$

Also

(6.29) $$D_{g_1}^{\leqslant}(1, 0) = 0 \cup \{(d_1, d_2): d_1 > 0\};$$

while

(6.30) $$D_{g_1}^{=}(1, 0) = 0$$

and

(6.31) $$D_{g_2}(1, 0) = \{(d_1, d_2): d_1 \geqslant 0\},$$

while

(6.32) $$D_{g_2}^{=}(1, 0) = \{(d_1, d_2): d_1 = 0\}.$$

Thus, while (6.28), confirms the promised equality of Propositions 4.1 and 6.1, the constancy and non-increase cones differ for both g_1 and g_2. Note, in addition, that (6.29) shows that $D_{g_1}^{\leqslant}(1, 0)$ is not closed even though g_1 is faithfully convex, Rockafellar (1970b). It is also reasonably easy to give examples in which $D_{g_i}^{=}(a)$ is not convex, Ben–Tal and others (1976). If, however, g is differentiable we have the following general result.

PROPOSITION 6.2. *Let $g: X \to Y$ be S-convex and differentiable. Let E be a face of S containing $-g(a)$. Then when S is closed*

(6.33) $$C(E) = \{d: \exists\alpha(d) > 0 \text{ with } g(a + \alpha d) \in E - E \text{ for } 0 \leqslant \alpha \leqslant \alpha(d)\}$$

is convex. As a consequence when g is differentiable both $D_g^{=}(a)$ and $D_{g_i}^{=}(a)$ are convex (independent of Ω).

PROOF. Let d_0, d_1 lie in $C(E)$. Let $d_t = td_1 + (1 - t)d_0$ for $0 \leqslant t \leqslant 1$. Then for small positive α

$$g(a + \alpha d_t) - g(a) \leqslant_S tg(a + \alpha d_1) + (1 - t)g(a + \alpha d_0) - g(a)$$

and, as $g(a)$ lies in E,

(6.34) $$g(a + \alpha d_t) - g(a) \in E - E - S = E - S.$$

In addition, for $i = 0, 1$,

(6.35) $$\nabla g(a)(d_i) = \lim_{\alpha \downarrow 0} \frac{g(a + \alpha d_i) - g(a)}{\alpha} \in E - E$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

as $E - E$ is closed. Thus, for $\alpha > 0$,

$$g(a + \alpha d_t) - g(a) \geqslant_s \alpha \nabla g(a) d_t = \alpha t \nabla g(a)(d_1) + \alpha(1 - t)\nabla g(a)(d_0), \tag{6.36}$$

because g is S-convex and differentiable.

Thus by (6.34), (6.35), (6.36)

$$g(a + \alpha d_t) - g(a) \in (S - E) \cap (E - S) \tag{6.37}$$

for small positive α. It follows from the extremality of E that $(S - E) \cap (E - S) \subset E - E$ and the first conclusion is obtained. The second conclusion follows on specializing E to be S^f or S_i^f.

7. Applications in matrix theory

Let S be the cone of $m \times m$ psd matrices in the space $Y = \mathbf{R}^{(m^2+m)/2}$ of $m \times m$ real symmetric matrices and suppose that we are given the matrix b in Y and the subspace L in $\mathbf{R}^m$. (By abuse of notation we also use lower case letters, such as a, b, x, s^+, to denote matrices.) Consider the problem:

(7.1) Find the closest (in Euclidean norm) matrix to b which is negative semi-definite (nsd) on L.

We let the inner-product in Y be, where tr denotes trace,

$$\langle x, y \rangle = \operatorname{tr} xy.$$

We now solve (7.1) by considering the problem as a programming problem and using the fact that every symmetric matrix can be diagonalized by a unitary transformation (of eigenvectors). Recall that S is self dual, Berman and Ben–Israel (1969).

THEOREM 7.1. *Suppose that the matrix b in Y and the subspace L in $\mathbf{R}^m$ are given. Then the unique closest matrix (in Euclidean norm) to b which is nsd on L is*

$$a = U\Lambda_- U^t + \mathcal{P} b, \tag{7.2}$$

where: $P_L b P_L = U \Lambda U^t = U\Lambda_+ U^t + U\Lambda_- U^t$; P_L is the (orthogonal) projection in Y on the subspace L in $\mathbf{R}^m$; U is a unitary matrix of eigenvectors; Λ is the diagonal matrix of eigenvalues; Λ_+ and Λ_- are the diagonal matrices of positive and negative eigenvalues; and the projection

$$\mathcal{P} \cdot = I - P_L \cdot P_L$$

PROOF. Choose an $m \times m$ matrix D, not necessarily symmetric, such that

$$L = \mathcal{R}(D), \tag{7.3}$$

where $\mathcal{R}(D)$ is the range of D. Now x in Y is nsd on L if and only if $y = D^t x D$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

is nsd, since

$$(D^t x D\alpha, \alpha) = (x(D\alpha), (D\alpha)), \qquad \alpha \in \mathbf{R}^m. \tag{7.4}$$

Thus our problem is equivalent to the abstract convex program

$$(P) \qquad \begin{cases} \text{minimize } p(x) = \frac{1}{2}\|x - b\|^2 \\ \text{subject to } g(x) = D^t x D \in -S, x \in X, \end{cases} \tag{7.5}$$

where $X = Y$. Let us solve (P) using Corollary 4.3. Since L need not be all of $\mathbf{R}^m$, the matrix D may be singular in which case Slater's condition fails for (P). Let

$$P = \text{projection on } \mathfrak{N}(D), \tag{7.6}$$

where $\mathfrak{N}(D)$ is the null space of D. First, we show that

$$S^f = S \cap \{P\}^{\perp}. \tag{7.7}$$

We need only show that, see Barker and Carlson (1975)

$$S^f = \{y \in S \colon \mathfrak{N}(y) \supset \mathfrak{N}(D)\}. \tag{7.8}$$

Suppose $-y \in g(F)$, i.e. $-y = g(x) = D^t x D \in -S$. Then $\mathfrak{N}(y) \supset \mathfrak{N}(D)$ is clear. Conversely, suppose $y \in -S$ and $\mathfrak{N}(y) \supset \mathfrak{N}(D)$. We need to find x such that $y = D^t x D$. Choose

$$x = D^{t\dagger} y D^{\dagger}, \tag{7.9}$$

where $\cdot^{\dagger}$ denotes the generalized inverse, Ben–Israel and Greville (1973). Then

$$\begin{aligned} D^t x D &= D^t D^{t\dagger} y D^{\dagger} D \\ &= P_{\mathfrak{R}(D^t)} y P_{\mathfrak{R}(D^t)} \\ &= y, \quad \text{since } \mathfrak{N}(y) \supset \mathfrak{N}(D). \end{aligned}$$

Thus we have shown that

$$-g(F) \subset S \cap \{P\}^{\perp} \subset -g(F) \tag{7.10}$$

which proves (7.7). In fact, we get that

$$-g(F) = S^f. \tag{7.11}$$

Moreover,

$$(I - P)S(I - P) = S^f. \tag{7.12}$$

For if $y = (I - P)s(I - P)$ and $s \in S$, then $\mathfrak{N}(y) \supset \mathfrak{N}(I - P) = \mathfrak{R}(P)$ and $y \in S$ which implies that $y \in S^f$. Conversely, if $y \in S^f$, then

$$\begin{aligned} y &= -D^t x D, \quad \text{for some } x \in X \text{ by (7.11)} \\ &= -(I - P) D^t x D (I - P), \quad \text{since } I - P = P_{\mathfrak{R}(D^t)}. \end{aligned}$$

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

Thus (7.12) exhibits a projection of S onto S^f. This implies that Corollary 4.3 is *applicable and we conclude that a solves (P) if and only if*

$$0 \in \nabla p(a) + \nabla s^+ g(a) - \left(D_g^=(a)\right)^+, \tag{7.13}$$

for some $s^+ \in S^+$ and $g(a) \in -S$ with $s^+ g(a) = 0$. Now, by (4.1) and linearity of g, we get

$$D_g^=(a) = \{d: D'dD \in S^f - S^f\} = X,$$

since

$$\begin{aligned} S^f - S^f &= (I - P)(S - S)(I - P), \quad \text{by (7.12)} \\ &= (I - P)Y(I - P) \end{aligned}$$

and $I - P = P_{\mathcal{R}(D')}$. Therefore, (7.13) becomes

$$b = a + Ds^+D', g(a) \in -S, s^+ \in S^+, \tag{7.14}$$

and

$$\operatorname{tr} s^+ D'aD = 0. \tag{7.15}$$

Set

$$a = U\Lambda_- U' + \mathcal{P}b \tag{7.16}$$

and

$$s^+ = D^\dagger U \Lambda_+ U' D'^\dagger. \tag{7.17}$$

Then

$$\begin{aligned} g(a) &= D'aD \\ &= D'(U\Lambda_- U' + \mathcal{P}b)D \\ &= D'(P_L U \Lambda_- U' P_L + \mathcal{P}b)D, \\ &\qquad \text{by the definitions in the statement of the theorem,} \\ &= D'U\Lambda_- U'D, \quad \text{since } L = \mathcal{R}(D) \\ &\in -S, \quad \text{since } \Lambda_- \text{ is diagonal and nsd;} \\ s^+ \in S^+ &= S, \quad \text{since } \Lambda_+ \text{ is diagonal and psd;} \\ a + Ds^+D' &= \mathcal{P}b + U\Lambda U', \quad \text{since } U\Lambda_+U' = P_L U \Lambda_+ U' P_L, \\ &= b; \end{aligned}$$

and

$$\begin{aligned} \operatorname{tr} s^+ D'aD &= \operatorname{tr} Ds^+D'a \\ &= \operatorname{tr} U\Lambda_+ U'a \\ &= \operatorname{tr} 0 = 0, \end{aligned}$$

since $P_L \cdot P_L$ and $\mathcal{P}$ are orthogonal projections, $\Lambda_+\Lambda_- = 0$ and $\mathcal{R}(U\Lambda_+U') \subset L$.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

Thus (7.14) and (7.15) are satisfied and a, given in (7.16), solves (P). The uniqueness follows by strict convexity of p.

The theorem gives the following *unique* decomposition of a symmetric matrix.

COROLLARY 7.1. *Suppose that b is a real $m \times m$ symmetric matrix and L is a subspace of $\mathbf{R}^m$. Then b can be decomposed uniquely as*

$$b = c + d \quad \textit{with } \operatorname{tr} cd = 0,$$

where c is nsd on L, d is psd and $\mathcal{R}(d) \subset L$.

PROOF. Set $c = a$ and $d = Ds^+D^t$ in (7.14). Then c is nsd on L since $D^taD \in -S$. Moreover, $\mathcal{R}(d) \subset L$, since $\mathcal{N}(d) \supset \mathcal{N}(D^t) = L^\perp$, and d is psd since s^+ is psd.

Note that if $L = \mathbf{R}^m$, then we get the unique decomposition of b into orthogonal psd and nsd matrices.

Acknowledgement

The authors would like to thank the referees for their many helpful comments.

References

R. A. Abrams and L. Kerzner (1978), "A simplified test for optimality", *J. Optimization Theory Appl.* **25**, 161–170.

G. P. Barker and D. Carlson (1975), "Cones of diagonally dominant matrices", *Pacific J. Math.* **57**, 15–32.

A. Ben-Israel, A. Ben-Tal and S. Zlobec (1976), "Optimality conditions in convex programming", *The IX International Symposium on Mathematical Programming* (Budapest, Hungary, August).

A. Ben-Israel and T. N. E. Greville (1973), *Generalized inverse theory and applications* (Wiley-Interscience, New York).

A. Ben-Tal and A. Ben-Israel (1979), "Characterizations of optimality in convex programming: the nondifferentiable case", *Applicable Anal.* **9**, 137–156.

A. Ben-Tal, A. Ben-Israel and S. Zlobec (1976), "Characterixations of optimality in convex programming without a constraint qualification", *J. Optimization Theory Appl.* **20**, 417–437.

A. Berman and A. Ben-Israel (1969), "Linear equations over cones with interior: a solvability theorem with applications to matrix theory", (Report No. 69-1 Series in Applied Math., Northwestern University).

J. Borwein (1977a), "Proper efficient points for maximizations with respect to cones", *SIAM J. Control Optimization* **15**, 57–63.

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press

J. Borwein (1977b), "Multivalued convexity and optimization: a unified approach to inequality and equality constraints", *Math. Programming* **13**, 183–199.
J. Borwein (1978), "Weak tangent cones and optimization in a Banach space", *SIAM J. Control Optimization* **16**, 512–522.
J. Borwein (1980a), "The geometry of Pareto efficiency over cones", *Math. Operationsforsch. Statist.*, **11**, 235–248.
J. Borwein (1980b), "A Lagrange multiplier theorem and a sandwich theorem for convex relations" (Report No. 80-1, Dalhousie University, Canada) *Math. Scand.*, to appear.
J. Borwein and H. Wolkowicz (1979), "Characterizations of optimality without constraint qualification for the abstract convex program", (Research Report No. 14, Dalhousie University, Halifax, N. S., Canada).
J. Borwein and H. Wolkowicz (1981), "Facial reduction for a cone-convex programming problem", *J. Austral. Math. Soc. Ser. A.* **30**, 000–000.
B. D. Craven and S. Zlobec, (1981) "Complete characterization of optimality for convex programming in Banach spaces", *Applicable Anal.* **11**, 61–78.
J. Gauvin (1977), "A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming", *Math. Programming* **12**, 136–138.
J. Gauvin and J. Tolle (1977), "Differential stability in nonlinear programming, *SIAM J. Control Optimization* **15**, 294–311.
A. M. Geoffrion (1968), "Proper efficiency and the theory of vector maximization", *J. Math. Anal. Appl.* **22**, 618–630.
R. B. Holmes (1975), *Geometric functional analysis and its applications*, (Springer-Verlag).
D. G. Luenberger (1968), "Quasi-convex programming", *SIAM J. Appl. Math.* **16**, 1090–1095.
D. G. Luenberger (1969), *Optimization by vector space methods* (John Wiley & Sons, Inc.).
H. Massam (1979), "Optimality conditions for a cone-convex programming problem", *J. Austral. Math. Soc. Ser. A.* **27**, 141–162.
J. P. Penot (1978), "L'optimisation à la Pareto: deux ou trois choses que je sais d'elle" Communication au Colloque "Structure Economiques et Econometric".
A. L. Peressini (1967), *Ordered topological vector spaces* (Harper and Row).
A. P. & J. W. Robertson (1964), *Topological vector spaces* (Cambridge University Press).
R. T. Rockafellar (1970a), *Convex analysis* (Princeton University Press).
R. T. Rockafellar (1970b), 'Some convex programs whose duals are linearly constrained", *Nonlinear Programming*, edited by J. B. Rosen, O. L. Mansasarian and K. Ritter, pp. 293–322 (Academic Press, N. Y.).
H. Wolkowicz (1978), "Calculating the cone of directions of constancy", *J. Optimization Theory Appl.* **25**, 451–457.
H. Wolkowicz (1980), "Geometry of optimality conditions and constraint qualifications: the convex case", *Math. Programming* **19**, 32–60.
J. Zowe (1974), "Subdifferentiability of convex functions with values in an ordered vector space", *Math. Scand.* **34**, 69–83.
J. Zowe (1975a), "Linear maps majorized by a sublinear map", *Arch. Math.* (Basel) **26**, 637–645.
J. Zowe (1975b), "A duality theorem for a convex programming problem in order complete vector lattices", *J. Math. Anal. Appl.* **50**, 273–287.
J. Zowe (1978), "Regularity and stability for the mathematical programming problem in Banach spaces", Preprint No. 37.

Department of Mathematics
Dalhousie University
Halifax
Canada

Department of Mathematics
University of Alberta
Edmonton
Canada

https://doi.org/10.1017/S1446788700017882 Published online by Cambridge University Press