Some Applications of Optimization in Matrix Theory*

Henry Wolkowicz
Department of Mathematics
The University of Alberta
Edmonton, Alberta, Canada T6G 2G1

Submitted by Richard A. Brualdi

ABSTRACT

We apply a recent characterization of optimality for the abstract convex program with a cone constraint to three matrix theory problems: (1) a generalization of Farkas's lemma; (2) paired duality in linear programming over cones; (3) a constrained matrix best approximation problem. In particular, these results are not restricted to polyhedral or closed cones.

1. INTRODUCTION

In this paper we apply a recent characterization of optimality for the abstract convex program with a cone constraint to three matrix theory problems: (1) a generalization of Farkas's lemma; (2) paired duality in linear programming over cones; (3) a constrained matrix best approximation problem. In particular, these results are not restricted to polyhedral or closed cones.

2. PRELIMINARIES

First, let us consider the abstract convex program

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

where $p: X\to R\cup\{+\infty\}$, $g: X\to R^m\cup\{+\infty\}$; X is a topological vector space; $\Omega\subset X$ is convex and S is a convex cone in R^m, i.e. $S+S\subset S$ and $tS\subset S$ for all

*Presented at the Matrix Theory Conference held at Auburn University, 19–22 March 1980. Research partially supported by NSERC A3388.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/81/060101+18$02.50

$t \geq 0$; and p is a convex (extended) functional (on Ω) and g is S-convex (on Ω), i.e.

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

for all $x, y \in \Omega$ and $0 \leq t \leq 1$. For greater generality, we have attached an abstract maximal element to R^m (see e.g. [13]). Note that R^m is linearly ordered by S, i.e. $x \geq_S y$ if and only if $x-y \in S$. This ordering is transitive and reflexive. It is antisymmetric exactly when S is pointed, i.e. $S \cap -S = \{0\}$.

We will also need the following notation: the cone $K \subset S$ is a *face* of S if

$$x, y \in S \text{ and } x+y \in K \quad \Rightarrow \quad x, y \in K.$$

The *feasible set* of (P) is

$$F = \{x \in \Omega : g(x) \in -S\}.$$

The *minimal cone* of (P) is

$$S^f = \bigcap \{\text{faces of } S \text{ containing } -g(F)\}.$$

The minimal cone has the following property (see [7]):

$$-g(F) \cap \operatorname{ri} S^f \neq \varnothing, \tag{2.2}$$

where ri denotes relative interior. For a set T in R^m, the *polar cone* of T is

$$T^+ = \{\phi \in R^m : \phi y \geq 0 \text{ for all } y \text{ in } T\},$$

where ϕy denotes the inner product of ϕ and y in R^m. The orthogonal complement of T is

$$T^\perp = T^+ \cap -T^+. \tag{2.3}$$

We now have the following characterization of optimality for (P) which holds without any constraint qualification.

THEOREM 2.1 [7]. *Suppose that μ, the optimal value of* (P), *is finite.*

Then

$$\mu=\inf\{p(x)+\lambda g(x): x\in F^f\} \tag{2.4}$$

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

$$\lambda g(a)=0 \tag{2.5}$$

and (2.4) *and* (2.5) *characterize optimality of* a *in* F.

The above theorem differs from the standard Lagrange multiplier theorem (e.g. [12]) in three ways. First, the Lagrange multiplier λ is found in the (possibly) larger cone $(S^f)^+$ rather than S^+; second, the variable x is restricted to the (possibly) smaller set F^f rather than Ω; and third, the above theorem holds irrespective of any constraint qualification. In the presence of Slater's qualification,

$$\exists x\in\Omega \quad \text{such that} \quad g(x)\in -\operatorname{int} S,$$

where int denotes interior, the above theorem reduced to the standard theorem.

There are situations where we can strengthen the above theorem in the sense that we can replace $(S^f)^+$ by a smaller cone and replace F^f by a larger set. Thus we get closer to the standard result. We now include several results of this type.

COROLLARY 2.1 [7]. *Theorem* 2.1 *holds with* S^+ *replacing* $(S^f)^+$ *exactly when*

$$S^+ +(S^f)^\perp=(S^f)^+, \tag{2.6}$$

or equivalently, when S *is closed, exactly when*

$$S^+ +(S^f)^\perp \text{ is closed.} \tag{2.7}$$

COROLLARY 2.2 [9]. *Suppose that* g *is affine and* Ω *is a finite dimensional subspace. Then Theorem* 2.1 *holds with* F^f *replaced by* Ω.

COROLLARY 2.3. *Suppose that* (P) *satisfies the generalized Slater's condition*

$$\exists \hat{x} \in \operatorname{ri} \Omega \quad \text{with} \quad g(\hat{x}) \in -\operatorname{ri} S. \tag{2.8}$$

Then the standard Lagrange multiplier theorem holds, i.e., Theorem 2.1 *holds with* F^f *replaced by* Ω *and* $(S^f)^+$ *replaced by* S^+.

Proof. By (2.2), we get that $S^f = S$. Now choose ϕ_i, $i=1,\dots,t$, in $S^\perp$ such that

$$S - S = \bigcap_{i=1}^{t} \phi_i^+. \tag{2.9}$$

By Theorem 2.1, we have

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

for some λ in S^+. This, by (2.9), is equivalent to

$$\mu = \inf\{p(x) + \lambda g(x) : (\phi_i g)(x) \leqslant 0,\ i=1,\dots,t,\ x \in \Omega\}. \tag{2.10}$$

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 implies that

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

Thus the program (2.10) is linearly constrained, and moreover, by (2.8),

$$\hat{x} \in \operatorname{ri} \Omega, \qquad \phi_i g(\hat{x}) \leqslant 0, \quad i=1,\dots,t. \tag{2.11}$$

This is the generalized Slater's condition for an ordinary convex program, since the constraints $\phi_i g$ are affine. Therefore (see [14, Theorem 28.2]), there exist Kuhn-Tucker multipliers $\alpha_i \geq 0$, $i=1,\ldots,t$, such that

$$\mu=\inf\left\{p(x)+\lambda g(x)+\sum \alpha_i\phi_i g(x): x\in\Omega\right\}. \tag{2.12}$$

The results now follows, since

$$\lambda+\sum \alpha_i\phi_i \in S^+. \qquad \blacksquare$$

We now include the following duality result. We define L^f, the *restricted Lagrangian*, by

$$L^f(\lambda)=\inf\left\{p(x)+\lambda g(x): x\in F^f\right\},$$

and the *restricted dual problem*

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

Then (D^f) is a concave optimization problem, and we get:

THEOREM 2.2 [7]. *If $\mu<\infty$, then*

$$\mu=\sup\left\{L^f(\lambda): \lambda\in(S^f)^+\right\}. \tag{2.13}$$

Moreover, if (2.6) holds, then $(S^f)^+$ can be replaced by S^+.

COROLLARY 2.4. *If $\mu<\infty$, Ω is a finite dimensional subspace, and g is affine, then (2.13) holds with F^f replaced by Ω (in the definition of $L^f(\lambda)$).*

Proof. Note that for every λ in $(S^f)^+$ and $x\in\Omega$,

$$L^f(\lambda)\leq p(x).$$

Thus $\mu\geq d$. Now Corollary 2.2 guarantees the existence of λ in $(S^f)^+$ with $\mu=L^f(\lambda)$. $\blacksquare$

We now present the applications of this theory. We restrict ourselves to X also finite dimensional.

3. A GENERALIZATION OF FARKAS'S LEMMA

The original Farkas's lemma (e.g. [10]) gives the equivalence of the consistency of the system

$$Ax=b, \qquad x\in K, \tag{3.1}$$

with the statement

$$A^t y\in K^+ \Rightarrow by \geq 0, \tag{3.2}$$

where A is an $m\times n$ matrix, A^t is its transpose, and $K=K^+=R^n_+$. This has been extended to K a closed convex cone in [3] under the assumption that $A(K)$ [or equivalently $K+\mathfrak{N}(A)$] is closed, where $\mathfrak{N}(\cdot)$ denotes null space. We now present a generalization of Farkas's lemma without this extra assumption. We let S^f be the minimal cone for the constraint $-A^t y\in -S$.

THEOREM 3.1 [9]. *The system*

(i) $$Ax=b, \qquad x\in(S^f)^+,$$

is consistent if and only if

(ii) $$A^t y\in S \Rightarrow by\geq 0.$$

Proof. Since S is a convex cone and A is linear, statement (ii) is equivalent to the fact that

$$0=\mu=\inf\{by: -A^t y\in -S\}. \tag{3.3}$$

Corollary 2.2 implies that this is equivalent to

$$0=\mu=\inf\{by+\lambda(-A^t y)\} \tag{3.4}$$

for some λ in $(S^f)^+$. Since the inf is achieved at $y=0$, we can differentiate to get statement (i), i.e.

$$0=\frac{d}{dy}\left[by-\lambda(A^ty)\right]$$
$$=b-A\lambda. \qquad \blacksquare$$

4. PAIRED DUALITY IN LINEAR PROGRAMMING OVER CONES

We first consider the following pair of linear programs with cone constraints. *Both* S and Ω are now convex cones (not necessarily polyhedral or closed) while $g(x)=b-Ax$, where b is an m-vector and A is an $m\times n$ matrix. We again denote the feasible set of (P)

$$F=\{x\in\Omega: Ax-b\in S\},$$

and set the minimal cones

$$S^f=\bigcap\ \{\text{faces of } S \text{ containing } A(F)-b\},$$

$$\Omega^f=\bigcap\ \{\text{faces of } \Omega \text{ containing } F\}.$$

Note that if we consider the constraint

$$\bar{g}(x)=\begin{bmatrix} I & -A \\ 0 & -I \end{bmatrix}\begin{pmatrix} b \\ x \end{pmatrix}\in-\begin{pmatrix} S \\ \Omega \end{pmatrix}=-(S\times\Omega),$$

then

$$(S\times\Omega)^f=S^f\times\Omega^f.$$

The dual pair is

(P) $$\begin{cases} \mu=\inf cx \\ \text{subject to}\quad Ax-b\in S,\quad x\in\Omega, \end{cases}$$

(D) $$\begin{cases} \sup by \\ \text{subject to}\quad c-A^ty\in(\Omega^f)^+,\quad y\in(S^f)^+. \end{cases}$$

The *Lagrangian* of (P) and (D) is

$$L(x,y)=cx+y(b-Ax)$$
$$=by+(c-A^t y)x. \tag{4.1}$$

The point $(x^0,y^0)\in\Omega^f\times(S^f)^+$ is a *saddle point* of $L(x,y)$ with respect to $\Omega^f\times(S^f)^+$ if

$$L(x^0,y)<L(x^0,y^0)<L(x,y^0) \qquad \text{for all} \quad x\in\Omega^f, \quad y\in(S^f)^+. \tag{4.2}$$

THEOREM 4.1. *Consider the paired programs* (P) *and* (D). *Then:*

(a) *If one of the problems is inconsistent, then the other is inconsistent or unbounded.*

(b) *Let the two problems be consistent, and let* x^0 *be a feasible solution of* (P) *and* y^0 *be a feasible solution of* (D). *Then*

$$cx^0>by^0. \tag{4.3}$$

(c) *If both* (P) *and* (D) *are consistent, then they have optimal solutions and their optimal values are equal.*

(d) *Let* x^0 *and* y^0 *be feasible solutions of* (P) *and* (D) *respectively. Then* x^0 *and* y^0 *are optimal if and only if*

$$y^0(Ax^0-b)+(c-A^t y^0)x^0=0,$$

or equivalently, if and only if

$$y^0(Ax^0-b)=(c-A^t y^0)x^0=0.$$

(e) *If* S^f *is closed (which holds if* S *is closed), then the vectors* $x^0\in R^n$ *and* $y^0\in R^m$ *are optimal solutions of* (P) *and* (D) *respectively if and only if the point* (x^0,y^0) *is a saddle point of* $L(x,y)$ *for all* (x,y) *in* $\Omega^f\times(S^f)^+$, *and then*

$$L(x^0,y^0)=cx^0=by^0. \tag{4.4}$$

Proof. (a): If (P) is inconsistent, then by [5, Theorem 3.3], there exists a vector ϕ such that

$$A^t\phi\in\Omega^+, \qquad -\phi\in S^+, \qquad b\phi>0. \tag{4.5}$$

Since $\Omega^f\subset\Omega$ and $S^f\subset S$, we get that $\Omega^+\subset(\Omega^f)^+$ and $S^+\subset(S^f)^+$. Thus (D) is unbounded.

Conversely, suppose that (P) is consistent and bounded. Now (P) is equivalent to

$$\mu=\inf\{cx: Ax-b\in S^f, x\in\Omega^f\}. \tag{4.6}$$

Moreover, if we rewrite the constraints as

$$g(x)=\begin{bmatrix} I & -A\\ 0 & -I\end{bmatrix}\begin{pmatrix} b\\ x\end{pmatrix}\in-\begin{pmatrix} S\\ \Omega\end{pmatrix}=-(S\times\Omega),$$

we see that $(S\times\Omega)^f=S^f\times\Omega^f$, and thus by (2.2), there exists

$$\hat{x}\in\operatorname{ri}\Omega^f \quad\text{with}\quad A\hat{x}-b\in\operatorname{ri}S^f.$$

Therefore, from (4.6) and Corollary 2.3, we get

$$\mu=\inf\{cx+y(b-Ax): x\in\Omega^f\}$$

for some y in $(S^f)^+$. Since Ω^f is a cone, this implies that

$$cx-yAx\geq 0 \qquad\text{for all}\quad x\in\Omega^f,$$

i.e., we have $c-A^ty\in(\Omega^f)^+$ and $y\in(S^f)^+$. Thus y is a feasible solution of (D).

(b):

$$cx^0\geq cx^0+y^0(b-Ax^0),\text{ since } b-Ax^0\in-S^f \text{ while } y^0\in(S^f)^+$$

$$=(c-A^ty^0)x^0+y^0b\geq y^0b,\text{ since } c-A^ty^0\in(\Omega^f)^+ \text{ while } x^0\in F\subset\Omega^f.$$

(c): Since both programs are consistent, part (b) implies that μ, the optimal value of (P), is not $-\infty$. Thus Theorem 2.2 applied to (4.6) implies

$$\mu = \sup_{y \in (S^f)^+} \inf_{x \in \Omega^f} \{cx + y(b - Ax)\}. \tag{4.7}$$

Since Ω^f is a cone, the inf for a fixed y is either yb or $-\infty$. Since we are then taking the sup, we can assume that the inf is yb, i.e. we can add the constraint

$$cx - yAx = (c - A^t y)x \geq 0 \qquad \text{for all} \quad x \in \Omega^f,$$

i.e.

$$c - A^t y \in (\Omega^f)^+.$$

Moreover, this inf is achieved with $x=0$. Thus (4.7) becomes

$$\mu = \sup_{y \in (S^f)^+} \left\{ yb : c - A^t y \in (\Omega^f)^+ \right\}.$$

(d): First suppose that x^0 and y^0 are both optimal. Thus

$$y^0 b = \mu = cx^0 \geq cx^0 + y^0(b - Ax^0), \text{ since } b - Ax^0 \in -S^f \text{ when } x^0 \in F$$

$$= y^0 b + (c - A^t y^0)x^0 \geq \mu,$$

i.e. $(c - A^t y^0)x^0 = y^0(b - Ax^0) = 0$. Conversely, suppose x^0 and y^0 are feasible and $(c - A^t y^0)x^0 = y^0(b - Ax^0) = 0$. Then by (b)

$$\begin{aligned} y^0 b &= y^0 b + (c - A^t y^0)x^0 \\ &= cx^0 + y^0(b - Ax^0) \\ &= cx^0 \geq by^0. \end{aligned} \tag{4.8}$$

Then $cx^0 = by^0$, and by (b) and (c), both x^0 and y^0 are optimal.

(e): Let (x^0, y^0) be a saddle point of $L(x, y)$, with respect to $\Omega^f \times (S^f)^+$. Then, for $x \in \Omega^f$,

$$\begin{aligned} L(x, y^0) &= by^0 + (c - A^t y^0)x \\ &\geq L(x^0, y^0) \\ &= by^0 + (c - A^t y^0)x^0. \end{aligned}$$

Thus $(c - A^t y^0)(x - x^0) \geq 0$ for all $x \in \Omega^f$. Setting $x = 0$ and $x = 2x^0$ implies that $(c - A^t y^0)x \geq 0$. Thus $c - A^t y^0 \in (\Omega^f)^+$, i.e., y^0 is a feasible point of (D). Similarly, for any $y \in (S^f)^+$,

$$\begin{aligned} L(x^0, y) &= cx^0 + y(b - Ax^0) \\ &\leq L(x^0, y^0). \end{aligned}$$

Thus $(y - y^0)(b - Ax^0) \leq 0$ for all $y \in (S^f)^+$. Again, setting $y = 0$ and $y = 2y_0$ shows that $y_0(b - Ax^0) = 0$ and thus $Ax^0 - b \in (S^f)^{++} = S^f$, since S^f is closed. This implies that x^0 is also feasible. Substituting $x = y = 0$ in the definition of the saddle point implies that $cx^0 \leq L(x^0, y^0) \leq by^0$, which, by (b) and (c), proves the optimality of x^0 and y^0.

Conversely, let x^0 and y^0 be optimal solutions of (P) and (D) respectively. Then $cx^0 = by^0$ by (c), and (4.4) follows from (d). Moreover, for any $x \in \Omega^f$,

$$\begin{aligned} L(x, y^0) &= by^0 + (c - A^t y^0)x \\ &\geq by^0, \qquad \text{since} \quad c - A^t y^0 \in (\Omega^f)^+ \\ &= L(x^0, y^0), \qquad \text{by (d)}. \end{aligned}$$

Similarly, for any $y \in (S^f)^+$, we have $L(x^0, y) \leq L(x^0, y^0)$, and thus (x^0, y^0) is a saddle point of $L(x, y)$ with respect to $\Omega^f \times (S^f)^+$. ■

The above theorem is an extension of the results given for polyhedral cones in [1] (see also [5]).

5. A BEST APPROXIMATION PROBLEM

Consider the following problem:

PROBLEM. Given the real symmetric $n\times n$ matrix B and the three subspaces L_1, L_2, and L_3 of R^n, find the (unique) real symmetric $n\times n$ matrix A which is closest to B in Frobenius norm (Hilbert-Schmidt norm) and which is negative semidefinite (nsd) on L_1, positive semidefinite (psd) on L_2, and 0 on L_3.

Solution. First, it is clear that A must be 0 on $L_1\cap L_2$. Thus, we can rewrite the problem so that A must be: 0 on $L_3+L_1\cap L_2$; nsd on L_1', any complementary subspace of $L_1\cap L_2+L_1\cap L_3$ in L_1; and psd on L_2', any complementary subspace of $L_1\cap L_2+L_2\cap L_3$ in L_2. Now set $L_3'=L_3+L_1\cap L_2$, and let: P_1 be the projection on L_1' along any complementary subspace of R^n which contains $L_2'+L_3'$; P_2 be the projection on L_2' along any complementary subspace which contains $L_1'+L_3'$; and P_3 be the projection on L_3' along any complementary subspace which contains $L_1'+L_2'$. We now define the unitary diagonalizations

$$\begin{aligned} P_iBP_i^t &= U_iD_iU_i^t \\ &= U_iD_i^+U_i^t+U_iD_i^-U_i^t, \quad i=1,2,3, \end{aligned}$$

where the U_i are the unitary matrices of eigenvectors, D_i are the diagonal matrices of eigenvalues, and D_i^+ and D_i^- are the diagonal matrices of positive and negative eigenvalues. We let S be the cone of $n\times n$ psd matrices, in the space $Y=R^{(n^2+n)/2}$ of $n\times n$ real symmetric matrices represented by their distinct triangular parts. The norm in Y is given by the Euclidean inner product

$$\langle A,B\rangle=\operatorname{tr}AB,$$

the trace of the matrix product AB. We define the projection in Y,

$$\mathcal{P}\cdot=I\cdot-P_4\cdot P_4,$$

where P_4 is the (orthogonal) projection on $L_1'+L_2'+L_3'$. Let us now show that the solution is

$$A=U_1D_1^-U_1^t+U_2D_2^+U_2^t+\mathcal{P}B. \tag{5.1}$$

Choose the matrices E_1, E_2, and E_3 so that

$$L_i' = \mathcal{R}(E_i),$$

where $\mathcal{R}(\ )$ denotes range space. The matrix X in Y is nsd (psd) on L_1' (L_2') if and only if $E_1^t X E_1$ ($E_2^t X E_2$) is nsd (psd) on all of R^n, since

$$(E_i^t X E_i y, y) = (X(E_i y), (E_i y)) \qquad \text{for} \quad y \text{ in } R^n.$$

Now we can rewrite the problem as the abstract convex program

$$(\mathrm{P}) \qquad \begin{cases} \text{minimize} & p(X) = \frac{1}{2}\|X-B\|^2 = \frac{1}{2}\operatorname{tr}(X-B)^2 \\ \text{subject to} & g_1(X) = E_1^t X E_1 \in -S, \\ & g_2(X) = -E_2^t X E_2 \in -S, \\ & g_3(X) = E_3^t X E_3 \in \{0\}. \end{cases}$$

The generalized Lagrange multiplier theorem states that if A satisfies the constraints and

$$0 = \nabla p(A) + \nabla\langle \lambda_1, g_1(A)\rangle + \nabla\langle \lambda_2, g_2(A)\rangle + \nabla\langle \lambda_3, g_3(A)\rangle, \tag{5.2}$$

$$0 = \langle \lambda_i, g_i(A)\rangle, \qquad i = 1, 2, 3,$$

for some matrices λ_1, λ_2 in S^+, λ_3 in $\{0\}^+ = Y$, where ∇ denotes gradient, then A solves (P).

Let A be as in (5.1) and set

$$\lambda_1 = E_1^{\dagger} U_1 D_1^+ U_1^t E_1^{\dagger t},$$

$$\lambda_2 = E_2^{\dagger} U_2 D_2^- U_2^t E_2^{\dagger t},$$

$$\lambda_3 = E_3^{\dagger} U_3 D_3 U_3^t E_3^{\dagger t},$$

where $E_i^{\dagger}$ denotes the generalized inverse of E_i (see e.g. [4]) which satisfies

$$E_i E_i^{\dagger} = P_i.$$

Then

$$\begin{aligned}\varepsilon_1(A)&=E_1^tAE_1\\&=E_1^t\left(U_1D_1^-U_1^t+U_2D_2^+U_2^t+\mathcal{P}B\right)E_1\\&=E_1^tP_1^t\left(U_1D_1^-U_1^t+P_2^tU_2D_2^+U_2^tP_2\right)P_1E_1\\&=E_1^tU_1D_1^-U_1^tE_1, \qquad \text{since} \quad P_2P_1=0\\&\in -S, \qquad \text{since} \quad D_1^- \text{ is nsd.}\end{aligned}$$

Similarly

$$\varepsilon_2(A)=-E_2^tU_2D_2^+U_2^tE_2\in -S$$

and

$$\varepsilon_3(A)=E_3^tAE_3=0.$$

Moreover, both λ_1 and λ_2 are in $S^+=S$ (see e.g. [6]), since both D_1^+ and $-D_2^-$ are psd, while $\lambda_3\in\{0\}^+=Y$. We have left to show that (5.2) holds, or equivalently, after differentiating, that

$$B=A+E_1\lambda_1E_1^t-E_2\lambda_2E_2^t+E_3\lambda_3E_3^t,$$

$$\operatorname{tr}\lambda_iE_i^tAE_i=0, \qquad i=1,2,3.$$

Now

$$\begin{aligned}A+E_1\lambda_1E_1^t-E_2\lambda_2E_2^t+E_3\lambda_3E_3^t&=A+U_1D_1^+U_1^t+U_2D_2^-U_2^tP_3BP_3\\&=U_1D_1U_1^t+U_2D_2U_2^t+P_3BP_3+\mathcal{P}B\\&=P_1BP_1+P_2BP_2+P_3BP_3+(I-P_4)B(I-P_4)\\&=B;\end{aligned}$$

while

$$\begin{aligned}\operatorname{tr}\lambda_1 E_1^t A E_1 &= \operatorname{tr} E_1\lambda_1 E_1^t A\\ &= \operatorname{tr} U_1 D_1^+ U_1^t A\\ &= 0,\end{aligned}$$

since the projections are mutually complementary and $D_1^+ D_1^- = 0$. Similarly

$$\operatorname{tr}\lambda_i E_i^t A E_i = 0, \qquad i = 2,3.$$

Uniqueness follows from the strict convexity of the objective function $p(X)$. ∎

REMARK 5.1. We were able to use the standard Lagrange multiplier theorem in the above, even though Slater's condition fails for (P). The cone S of psd matrices is very well behaved in general. In fact, if K is a face of S, then

$$S^+ + K^\perp \text{ is closed.} \tag{5.3}$$

For (see [2]) there exists a projection P in S such that $PS = K$, and moreover

$$S^+ + K^\perp = S + \mathfrak{N}(P),$$

(where $\mathfrak{N}$ denotes null space) is closed if and only if PS is closed (see [11]). Recall that the condition (5.3) is the one in Corollary 2.1 which allows one to replace $(S^f)^+$ by S^+ in Theorem 2.1.

REMARK 5.2. It is well known (see e.g. [12, p. 222]) that the Lagrange multipliers are sensitivity coefficients, i.e., if a solves the original program (P) in Section 2 with Lagrange multipler λ, while a_z solves the perturbed program (P_z) with the perturbed constraint $g(x) \in z - S$, then

$$p(a) - p(a_z) < \lambda z. \tag{5.4}$$

For the above matrix problem, suppose that we allow the following perturbation:

$$A \text{ must be "almost" nsd on } L_1,$$

i.e., we are given the scalar ε and we require

$$(Ax,x)\leq\varepsilon(x,x)\qquad\text{for all}\quad x\text{ in }L_1.$$

Then this is equivalent to the perturbation

$$(E_1^tAE_1y,y)\leq\varepsilon(E_1^tE_1y,y)\qquad\text{for all } y,$$

or equivalently

$$E_1^tAE_1-\varepsilon E_1^tE_1\in-S.$$

If A_ε is the solution of the perturbed problem, then, using the value of λ_1 found in our solution, we get

$$p(A)-p(A_\varepsilon)\leq\lambda_1(\varepsilon E_1^tE_1)=\varepsilon\operatorname{tr}\lambda_1E_1^tE_1;$$

expanding both sides yields

$$\begin{aligned}\tfrac{1}{2}\operatorname{tr}\left((B-A)^2-(B-A_\varepsilon)^2\right)&=\tfrac{1}{2}\operatorname{tr}\left(A^2-A^2-2B(A-A_\varepsilon)\right)\\&\leq\varepsilon\operatorname{tr}E_1^tU_1D_1^+U_1^tE_1^{t\,t}E_1^tE_1\\&=\varepsilon\operatorname{tr}U_1D_1^+U_1^t.\end{aligned}$$

As may have been expected, the sensitivity depends on the nonnegative eigenvalues of P_1BP_1. Note also that if $\varepsilon<0$, then the feasible set $F=\varnothing$.

Similarly, if A_ε solves the problem with the three perturbed constraints

$$(Ax,x)\leq\varepsilon_1(x,x)\qquad\text{for all}\quad x\text{ in }L_1,$$

$$(-Ax,x)\leq\varepsilon_2(x,x)\qquad\text{for all}\quad x\text{ in }L_2,$$

$$(Ax,x)=\varepsilon_3(x,x)\qquad\text{for all}\quad x\text{ in }L_3,$$

where ε_1, ε_2, and ε_3 are three scalars, then

$$p(A)-p(A_\varepsilon)\leq\varepsilon_1\operatorname{tr}U_1D_1^+U_1^t+\varepsilon_2\operatorname{tr}U_2D_2^-U_2^t+\varepsilon_3\operatorname{tr}U_3D_3U_3^t.$$

In addition, if we had Lagrange multipliers corresponding to A_ϵ, we would also get a lower bound for $p(A)-p(A_\epsilon)$.

6. CONCLUDING REMARKS

In [9] we have presented various strengthened versions of Theorem 2.1 which, under certain hypotheses, use smaller cones than $(S^f)^+$, though not necessarily S^+. This has led to various cones replacing $(S^f)^+$ in the generalization of Farkas's lemma, i.e. various equivalent statements.

These strengthenings can also be applied to the paired duals in Section 4. For example, under certain closure conditions, we would get a family of paired duals

$$(\mathrm{P}_{K,L}) \qquad \begin{cases} \mu = \inf cx \\ \text{subject to} \quad Ax-b\in K, \quad x\in L, \end{cases}$$

$$(\mathrm{D}_{M,N}) \qquad \begin{cases} \sup by \\ \text{subject to} \quad c-A^t y\in M^+, \quad y\in N^+, \end{cases}$$

where K, L, M, and N are convex cones which satisfy certain closure hypotheses as well as the inclusions $S^f\subset K\subset S$, $\Omega^f\subset L\subset\Omega$, $S^+\subset M^+\subset(\Omega^f)^+$, and $\Omega^+\subset N^+\subset(S^f)^+$. These closure conditions hold in the event that both Ω and S are polyhedral cones, in which case we can choose $M=\Omega$ and $N=S$. Theorem 4.1 then reduces to a result given in [1] (see also [5]).

These strengthenings of Theorem 4.1 will be presented in a future study.

REFERENCES

1 R. A. Abrams and A. Ben-Israel, Complex mathematical programming, in *Proceedings of the Third Annual Israel Conference on Operations Research*, Tel Aviv, July 1969.

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

3 A. Ben-Israel, Linear equations and inequalities on finite dimensional, real or complex, vector spaces: a unified theory, *J. Math. Anal. Appl.* 27:367–389 (1969).

4 A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley-Interscience, New York, 1973.

5 A. Berman, *Cones, Matrices and Mathematical Programming*, Lecture Notes in Economics and Mathematical Systems, Springer, 1973.

6 A. Berman and A. Ben-Israel, Linear equations over cones with interior: a solvability theorem with applications to matrix theory, Report No. 69-1, Series in Applied Math., Northwestern Univ., 1969.
7 J. M. Borwein and H. Wolkowicz, Characterizations of optimality for the abstract convex program with finite dimensional range, *J. Austral. Math. Soc.*, 30:390–411 (1981).
8 J. M. Borwein and H. Wolkowicz, Regularizing the abstract convex program, *J. Math. Anal. Appl.*, to appear.
9 J. M. Borwein and H. Wolkowicz, Cone-convex programming: stability and affine constraint functions, in *Proceedings of the NATO Conference on Generalized Concavity*, Vancouver, 1980, to appear.
10 J. Farkas, Uber die Theorie des einfachen Ungleichungen, *J. Reine Angew Math.* 124:1–24 (1902).
11 R. B. Holmes, *Geometric Functional Analysis and Its Applications*, Springer, 1975.
12 D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, 1969.
13 A. L. Peressini, *Ordered Topological Vector Spaces*, Harper and Row, 1967.
14 R. T. Rockafellar, *Convex Analysis*, Princeton U.P., Princeton, N.J., 1970.

Received 25 April 1980; revised 24 November 1980