NORTH-HOLLAND

Explicit Solutions for Interval Semidefinite Linear Programs*

Henry Wolkowicz†

Department of Combinatorics and Optimization
Faculty of Mathematics
University of Waterloo
Waterloo, Ontario N2L 3G1, Canada

Submitted by Hans Schneider

ABSTRACT

We consider the special class of semidefinite linear programs

(IVP) maximize $\operatorname{trace} CX$ subject to $L \preceq A(X) \preceq U$,

where C, X, L, U are symmetric matrices, A is an (onto) linear operator, and $\preceq$ denotes the Löwner (positive semidefinite) partial order. We present explicit representations for the general primal and dual optimal solutions. This extends the results for standard linear programming that appeared in Ben-Israel and Charnes [3]. This work is further motivated by the explicit solutions for a different class of semidefinite problems presented recently in Yang and Vanderbei [15].

1. INTRODUCTION

We study the semidefinite linear programming problem with interval constraints

$$\text{(IVP)} \qquad \begin{array}{ll} \text{maximize} & \operatorname{trace} CX \\ \text{subject to} & L \preceq A(X) \preceq U, \end{array}$$

where C, X are symmetric $n \times n$ matrices; L, U are symmetric $m \times m$ matrices; A is a linear operator from the space of $n \times n$ symmetric matrices

*This report is available by anonymous ftp at `orion.uwaterloo.ca` in the directory `pub/henry/reports`. E-mail:`henry@orion.waterloo.ca`. URL: `http://orion.uwaterloo.ca/~hwolkowi`.

†The author thanks The National Science and Engineering Research Council Canada for their support.

LINEAR ALGEBRA AND ITS APPLICATIONS 236:95–104 (1996)

655 Avenue of the Americas, New York, NY 10010

0024-3795/96/$15.00
SSDI 0024-3795(94)00130-6

onto the space of $m \times m$ symmetric matrices; and $\preceq$ denotes the Löwner partial order, i.e. $X \preceq Y$ if and only if $Y - X$ is positive semidefinite. We consider the space of symmetric matrices as a vector space with the *trace inner product* $\langle C, X\rangle = \operatorname{trace} CX$. The corresponding norm is the *Frobenius matrix norm* $\|X\| = \sqrt{\operatorname{trace} X^2}$.

There has recently been a resurgence of interest in problems with semidefinite constraints. This is partly due to new applications to integer programming and min-max eigenvalue problems as well as to successful new solution techniques using interior point methods; see e.g., [1, 6, 7, 10, 11, 13]. It is interesting and surprising that many of the results from linear programming follow through to these nonlinear problems.

In the case that the partial order is the coordinatewise ordering, (IVP) reduces to the ordinary interval linear programs studied in [3, 12]. An explicit solution and an algorithm for general linear programming problems based on these solutions is provided therein. In this note, we show that the results from [3] can be extended to the class (IVP) of semidefinite programming problems. This paper is further motivated by the explicit solutions, provided recently in [15], of the following class of semidefinite programs:

$$\begin{array}{ll} \text{maximize} & \operatorname{trace} CX \\ \text{subject to} & MXM^t = B, \quad X \succeq 0, \end{array} \tag{1.1}$$

where C is a given symmetric matrix, M is a full row rank matrix, and the dual of the program (1.1) is strictly feasible.

Note that the results in this paper still hold in the more general setting when X is a vector in a given vector space $\mathcal{V}$ and A is a linear operator from $\mathcal{V}$ onto the space of $m \times m$ matrices. Therefore these problems fall into the class of programs with finite dimensional range studied in [5].

Our main result is the explicit representation of the general solution of (IVP) presented as Theorem 2.2. Section 3 discusses two Matlab programs that find these explicit solutions.

2. THEORETICAL RESULTS

We let A^- denote a *generalized inverse* of A, i.e., a linear operator that satisfies

$$AA^-A = A; \tag{2.1}$$

see e.g., [4]. Then the general solution of the linear equation $A(X) = Y$ is $X \in A^-(Y) + \mathcal{N}(A)$, where $\mathcal{N}(A)$ is the *null space* of A.

LEMMA 2.1. *Suppose that* (IVP) *is feasible. Then* (IVP) *has a bounded solution if and only if*

$$C \perp \mathcal{N}(A), \tag{2.2}$$

where $\perp$ *denotes orthogonal.*

Proof. If (2.2) fails, then it is clear that the objective function of (IVP) can be made arbitrarily large. Conversely, if (2.2) holds, then the objective value $\operatorname{trace} CX = \operatorname{trace} CP_{\mathcal{R}(A^-)}(X)$, where $P_{\mathcal{R}(A^-)} = A^-A$ denotes the *projection on the range of* A^-. Moreover, feasibility of X implies that $\|A(X)\| \leq \max\{\|L\|\,\|U\|\}$. Therefore,

$$\begin{aligned} \operatorname{trace} CX &\leq \|C\|\,\|X\| \leq \|C\|\,\|A^-\|\,\|A(X)\| \\ &\leq \|C\|\,\|A^-\| \max\{\|L\|, \|U\|\}, \end{aligned}$$

where the operator norm $\|A^-\|$ is induced by the vector (Frobenius) norm, i.e., $\|A^-\| = \max_{\|Y\|=1} \|A^-(Y)\|$. ∎

Thus, under the boundedness assumption, (IVP) has the following simplification

$$\text{(SIVP)} \qquad \begin{array}{ll} \text{maximize} & \operatorname{trace} CA^-(Y) = \operatorname{trace} \overline{C}Y \\ \text{subject to} & L \preceq Y \preceq U, \end{array}$$

where

$$\overline{C} = (A^-)^* C \tag{2.3}$$

and $(A^-)^*$ is the *adjoint* of A^-, i.e., for the appropriate inner products, $\langle Z, A^-(Y)\rangle = \langle (A^-)^*(Z), Y\rangle \ \forall Z, Y$. Note that the definition of $\overline{C}$ means that it is symmetric and it is a least squares solution of $A^*(Z) = C$. The condition for boundedness in Lemma 2.1 implies that this least squares solution $\overline{C}$ is, in fact, a solution.

We now present the optimality conditions for this simplified program (SIVP). We first need some definitions. K is a (convex) cone if $K + K \subset K$, and $tK \subset K \ \forall t \geq 0$. The cone $T \subset K$ is a *face* of the cone K, denoted $T \lhd K$, if

$$x, y \in K \quad x + y \in T \quad \Rightarrow \quad x, y \in T.$$

The *feasible set* of (SIVP) is

$$\mathcal{F} = \{Y : L \preceq Y \preceq U\},$$

and we call (SIVP) *feasible* if $\mathcal{F} \neq \emptyset$. The *minimal cones* of (SIVP) are

$$U^f = \bigcap\{\text{faces of } \mathcal{P} \text{ containing } U - \mathcal{F}\},$$
$$L^f = \bigcap\{\text{faces of } \mathcal{P} \text{ containing } \mathcal{F} - L\},$$

where $\mathcal{P}$ denotes the *cone of positive semidefinite matrices* in the appropriate space. The *polar cone* of a set C is

$$C^+ = \{\phi : \langle \phi, c\rangle \geq 0 \ \forall c \in C\}.$$

From the commutativity of the trace, it can be shown that $\mathcal{P}$ is *self-polar*, i.e.

$$\mathcal{P} = \mathcal{P}^+. \tag{2.4}$$

The faces of $\mathcal{P}$ can be completely characterized in terms of the null spaces of the matrices, i.e., the matrix P is in the minimal face containing the matrix Q if and only if $\mathcal{N}(Q) \supset \mathcal{N}(P)$; see e.g., [2]. Moreover, the faces are *exposed*, i.e., they are equal to the intersection of a hyperplane with $\mathcal{P}$. This can be used to show that the minimal cones are equal, i.e. if $\Phi \succeq 0$ and $\operatorname{trace}\Phi(U - Y) = 0 \ \forall Y \in \mathcal{F}$, then $L \in \mathcal{F}$ and $0 = \operatorname{trace}\Phi(U - L) = \operatorname{trace}\Phi(U - Y + Y - L) = \operatorname{trace}\Phi(Y - L) \ \forall Y \in \mathcal{F}$. Thus every hyperplane containing the minimal cone U^f also contains L^f. The converse follows similarly. We denote the *minimal cone* by

$$\mathcal{P}^f = L^f = U^f.$$

We now state the optimality conditions (see [14]).

THEOREM 2.1. *Suppose that*

$$L \preceq U.$$

Then an optimal solution exists for (SIVP). *Moreover, Y solves* (SIVP) *if and only if the following system is consistent:*

$$\begin{array}{rll} \overline{C} = S_2 - S_1 & (\textit{dual feasibility}), & \\ U - L = Z_1 + Z_2 & (\textit{primal feasibility}), & (2.5) \\ \operatorname{trace} S_1 Z_1 = \operatorname{trace} S_2 Z_2 = 0 & (\textit{complementary slackness}), & \end{array}$$

with

$$Y = L + Z_1 = U - Z_2 \ \textit{and} \ S_1, S_2 \in (\mathcal{P}^f)^+, \qquad Z_1, Z_2 \in \mathcal{P}.$$

Proof. That (SIVP) is feasible and bounded is clear. Now, we can rewrite the two sided constraint of (SIVP) as two constraints. Then the Lagrangian becomes

$$\operatorname{trace}[\overline{C}Y - S_1(L - Y) - S_2(Y - U)],$$

with dual variables (Lagrange multipliers) $S_1 \in (U^f)^+$, $S_2 \in (L^f)^+$; see [14]. Differentiating yields the dual feasibility equation. The primal feasibility equation comes from adding the slack variables Z_1, Z_2 and eliminating Y. The characterization of optimality using complementary slackness follows from Theorem 4.1 in [14]. ∎

REMARK 2.1. The dual of (SIVP) is

$$\text{(DSIVP)} \qquad \begin{array}{ll} \text{minimize} & \operatorname{trace}(-LS_1 + US_2) \\ \text{subject to} & -S_1 + S_2 = \overline{C}, \\ & S_1, S_2 \in (\mathcal{P}^f)^+. \end{array}$$

The dual variables do not change for (IVP), since A is onto. Under the boundedness assumption, the dual problem for (IVP) is the same as for (SIVP) after multiplying the equality constraint by the adjoint A^*. Note that $\overline{C}$ is then replaced by C.

REMARK 2.2. In the case that Slater's condition

$$\text{there exists } \widehat{Y} \text{ such that } L \prec \widehat{Y} \prec U \tag{2.6}$$

holds, we have $\mathcal{P}^f = \mathcal{P}$. Since $\mathcal{P}$ is self-polar, the above optimality conditions simplify in the sense that both

$$(U^f)^+ = (L^f)^+ = \mathcal{P}. \tag{2.7}$$

However, even though (SIVP) has the trivial identity constraint, the standard duality results given by (2.7) can fail if Slater's condition does not hold. For example, suppose that

$$\overline{C} = \begin{bmatrix} -8 & 6 \\ 6 & 8 \end{bmatrix} \quad \text{and} \quad U - L = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}. \tag{2.8}$$

Then $U - L$ is a point on a ray which is a minimal face of $\mathcal{P}$. Therefore, complementary slackness implies that one of S_1 and S_2 (say S_2) must be on the orthogonal ray, i.e., the ray through the matrix

$$T = \begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix}.$$

But then dual feasibility implies that S_1 must have a negative element on the diagonal, a contradiction to finding multipliers in $\mathcal{P}$. (We continue this example in Section 3.)

REMARK 2.3. The decomposition of $\overline{C}$ into positive and negative parts is unique if the two parts are orthogonal and positive semidefinite, i.e.,

$$S_1S_2 = 0, \quad S_1, S_2 \in \mathcal{P}.$$

This is called the Moreau decomposition [9]. Therefore, uniqueness of the optimal dual solution implies that we have obtained a Moreau decomposition. Note that if $\mathcal{P}^f = \mathcal{P}$, then complementary slackness implies the stronger condition $S_iZ_i = 0, \ i = 1, 2$.

We now find an explicit solution to the optimality conditions (2.5). Let

$$E = U - L, \tag{2.9}$$

and let Q be the nonsingular matrix formed from the scaled eigenvectors of E so that

$$\overline{E} = QEQ^t = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}, \qquad Q \text{ nonsingular}, \tag{2.10}$$

where I_k denotes the $k \times k$ identity matrix. Let P be the $k \times k$ orthogonal matrix that diagonalizes the upper left $k \times k$ block of $\overline{C}$ after congruence by the inverse of Q, i.e.

$$[P \quad 0]Q^{-t}\,\overline{C}Q^{-1} \begin{bmatrix} P^t \\ 0 \end{bmatrix} = D, \tag{2.11}$$

with

$$PP^t = I, \qquad D = \begin{bmatrix} 0 & 0 \\ 0 & \underline{D} \end{bmatrix}_{k\times k} \text{ diagonal}, \quad \underline{D}_{(k-h)\times(k-h)} \text{ nonsingular}.$$

We let

$$\overline{P} = \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix}, \qquad \overline{D} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$

$$G = \overline{P}Q^{-t}\overline{C}Q^{-1}\overline{P}^t, \qquad \overline{G} = G - \overline{D}. \tag{2.12}$$

Note that trace $\overline{D}\,\overline{G} = 0$. We now construct two $n \times n$ symmetric positive semidefinite matrices with arbitrary upper left $h \times h$ blocks. The remaining elements of these two matrices are zero, except for the diagonal elements, which depend on the signs of the diagonal elements of D. Choose T so that

$$T \text{ is arbitrary, symmetric with } 0 \preceq T_{h\times h} \preceq I. \tag{2.13}$$

Define

$$D_1(i,j) = \begin{cases} T(i,j) & \text{if } \; i,j \leq h, \\ 1 & \text{if } \; i=j>h \text{ and } \overline{D}(i,i)>0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.14}$$

and

$$D_2 = \overline{E}' - D_1. \tag{2.15}$$

We now present the explicit representations of the primal and dual optimal solutions. We use the various matrices defined above.

THEOREM 2.2. *Suppose that* (IVP) *is feasible and has a finite solution value, i.e., that* (2.2) *holds. Let R be an arbitrary symmetric $n \times n$ matrix with the top left $k \times k$ block identically* 0. *Set*

$$Z_1 = Q^{-1}\overline{P}^t D_1 \overline{P} Q^{-t}, \qquad Z_2 = Q^{-1}\overline{P}^t D_2 \overline{P} Q^{-t}, \tag{2.16}$$

and

$$S_1 = -Q^t\overline{P}^t(D_2\overline{D} + R)\overline{P}Q, \qquad S_2 = Q^t\overline{P}^t(D_1\overline{D} + (\overline{G} - R))\overline{P}Q. \tag{2.17}$$

Then, for all such arbitrary R and arbitrary T defined as part of D_1, the matrices S_1, S_2 are optimal for the dual program of (SIVP) *as well as for the dual program of* (IVP). *The general solution for the simplified problem* (SIVP) *is*

$$Y = L + Z_1 \; (= U - Z_2).$$

The general solution for (IVP) *is*

$$X \in A^-(L + Z_1) + \mathcal{N}(A) \; [= A^-(U - Z_2) + \mathcal{N}(A)]. \tag{2.18}$$

Proof. We use the matrices defined above to reduce (IVP) to an ordinary linear programming problem. The key step is the reduction to a problem for which the principal parts (corresponding to the minimal faces) of the Lagrange multipliers and slack variables all commute and so are mutually diagonalizable.

From the definition of E, the constraint for (SIVP) can be further simplified to $0 \preceq T \preceq E$, with $Y = L + T$. We then make the substitution of variables $V = QTQ^t$ and replace the objective function matrix with

$Q^{-t}\overline{C}Q^{-1}$. We get the equivalent problem

$$\begin{array}{ll} \text{maximize} & \operatorname{trace}(Q^{-t}\overline{C}Q^{-1})V \\ \text{subject to} & 0 \preceq V \preceq \overline{E}. \end{array}$$

The simple structure of $\overline{E}$ identifies the minimal faces, i.e., the feasible variables V have to be positive semidefinite with nonzeros corresponding only to the nonzero block I_k of $\overline{E}$. Moreover (assume for simplicity that $k = n$; otherwise, consider the principal leading blocks), if we use the primal feasibility optimality conditions in (2.5) to substitute for Z_1 in the complementary slackness condition, then we see that S_1Z_2 is symmetric and so S_1, Z_2 commute and are mutually, orthogonally diagonalizable. That S_1, Z_1 commute follows from $S_1Z_1 = 0$. This shows that the Lagrange multipliers and slack variables can be mutually diagonalized and we can reduce the problem to an ordinary linear program. This we now do.

We diagonalize the upper left block of the new objective function using orthogonal P. Simultaneously we substitute the variables $W = \overline{P}V\overline{P}^t$ so as not to change the value of the objective function or $\overline{E}$. This last equivalent problem is

$$\begin{array}{ll} \text{maximize} & \operatorname{trace} GW \\ \text{subject to} & 0 \preceq W \preceq \overline{E}. \end{array} \tag{2.19}$$

Since W is zero except possibly in the top left $k \times k$ block, the objective function is equivalent to the diagonal objective trace $\overline{D}W$. Therefore this is a simple ordinary interval linear program on the diagonal elements of W. The zero elements of D in $\overline{D}$ allow an arbitrary positive semidefinite block in the optimal solution. Otherwise, the remainder of the solution must be diagonal, where the sign of $D(i,i)$ determines whether the diagonal element is 0 or 1. (This follows from the fact that the eigenvalues of a symmetric matrix majorize the diagonal elements; see e.g., [8]) Thus, for arbitrary T defined in (2.13), the matrices D_1 provide the general solution of (2.19). We can now reverse the reduction steps (using the congruences by Q and $\overline{P}$) to get the general solutions for our general semidefinite programs (IVP).

We have obtained the general solution without explicitly using the optimality conditions in Theorem 2.1. However, direct substitution shows that the Lagrange multipliers and slacks defined in the theorem satisfy the optimality conditions (2.5). ■

3. CONCLUSION

We have provided explicit expressions for the general solutions of (IVP). This is accomplished by reduction to the simpler problem (SIVP), where the objective function matrix $\overline{C}$ can be found by solving a system of linear equations. The generalized inverse A^- need never be evaluated explicitly. This result can be combined with the simple constraint in (1.1) to provide explicit solutions to interval constraints of the type $L \preceq MA(X)M^t \preceq U$, where A is still an onto operator while M is a general matrix of appropriate dimensions.

It is interesting and surprising that so many results from ordinary linear programming, such as the above explicit solutions, follow through to the nonlinear semidefinite partial order. The standard linear programming duality results follow through if a constraint qualification holds; in the absence of a CQ, they still hold if the minimal cones are taken into account: see [14]. The efficiency of interior point methods follows. (See the various references mentioned above.) It seems possible that the finite pivot algorithm for general ordinary linear programs, based on the explicit solutions for interval linear programs (see [12]), can be extended as well.

Rather than present actual examples to illustrate our theory, we have written two Matlab programs which find the explicit solutions of (SIVP). The first program initializes a random problem of type (SIVP). The user can guarantee nonuniqueness of the primal and/or dual optimal solutions. The second program finds the optimal primal and dual solutions and verifies the optimality conditions. These programs are available using anonymous ftp at the site `orion.uwaterloo.ca`, in the directory `pub/henry/matlab/semidef`. They are called `initsemi.m` and `algorsemi.m`, respectively. The programs were tested on thousands of random problems. We have not included the operator A, since the reduction to (SIVP) can be done by solving the linear system of equations for $\overline{C}$. Examples of linear operators A can be found in the various references on semidefinite programming.

For the example in Remark 2.2, we set $L = 0$. The Matlab program finds that $Y = U$ is primal optimal. A pair of dual optimal matrices is

$$S_1 = \begin{bmatrix} 0.1508 & -0.1929 \\ -0.1929 & 0.1552 \end{bmatrix}, \qquad S_2 = \begin{bmatrix} 7.8492 & 5.8071 \\ 5.8071 & 8.1552 \end{bmatrix}.$$

Then $Z_1 = Y$, $Z_2 = 0$. Note that complementary slackness ($\operatorname{trace} S_1 Z_1 = 0$) holds but does not imply $S_1 Z_1 = 0$, since S_1 is not positive semidefinite.

REFERENCES

1 F. Alizadeh, Combinatorial optimization with semidefinite matrices, in *Proceedings of the Second Annual Integer Programming and Combinatorial Optimization Conference*, Carnegie-Mellon Univ., 1992.

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

3 A. Ben-Israel and A. Charnes, An explicit solution of a special class of linear programming problems, *J. Oper. Res. Soc.* 16(6):1166–1175 (1968).
4 A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley-Interscience, 1974.
5 J. M. Borwein and Henry Wolkowicz, Characterizations of optimality for the abstract convex program with finite dimensional range, *J. Austral. Math. Soc. Ser. A* 30:390–411 (1981).
6 R. W. Freund and F. Jarre, An Interior-Point Method for Convex Fractional Programming, Technical Report Manuscript 93-03, AT&T Bell Lab., Murray Hill, N.J., 1993.
7 F. Jarre, An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices, *SIAM J. Control Optim.* 31:1360–1377 (1993).
8 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic, New York, 1979.
9 J. J. Moreau, Décomposition orthogonale d'un espace hilbertien selon deux cônes mutuellement polaires, *C. R. Acad. Sci. Paris* 255:238–240 (1962).
10 Y. E. Nesterov and A. S. Nemirovsky, *Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms*, SIAM, Philadelphia, 1993.
11 F. Rendl, R. J. Vanderbei, and H. Wolkowicz, Max-Min Eigenvalue Problems, Primal-Dual Interior Point Algorithms, and Trust Region Subproblems, Technical Report CORR 93-30, Dept. of Combinatorics and Optimization, Univ. of Waterloo; in *Proceedings of NATO Conference on Nonlinear Programming*, Il-Ciocco, Italy, 1993, to appear.
12 P. D. Robers and A. Ben-Israel, Interval programming, *I&EC Process Design and Devel.* 8:496–501 (1969).
13 L. Vandenberghe and S. Boyd, Primal-Dual Potential Reduction Method for Problems Involving Matrix Inequalities, Technical Report, Electrical Engineering Dept., Stanford Univ., Stanford, CA 94305, 1993.
14 Henry Wolkowicz, Some applications of optimization in matrix theory, *Linear Algebra Appl.* 40:101–118 (1981).
15 B. Yang and R. J. Vanderbei, The simplest semidefinite programs are trivial, *Math. Oper. Res.* to appear.

Received 12 November 1993; final manuscript accepted 21 May 1994