The Simplest Semidefinite Programs are Trivial

Robert J. Vanderbei [*] Bing Yang

Program in Statistics & Operations Research
Princeton University
Princeton, NJ 08544

January 10, 1994

Technical Report SOR-93-12

Abstract

We consider optimization problems of the following type:

$$\min\{\mathrm{tr}(CX) : A(X) = B, X \text{ positive semidefinite}\}.$$

Here, $\mathrm{tr}(\cdot)$ denotes the trace operator, C and X are symmetric $n \times n$ matrices, B is a symmetric $m \times m$ matrix and $A(\cdot)$ denotes a linear operator. Such problems are called semidefinite programs and have recently become the object of considerable interest due to important connections with max-min eigenvalue problems and with new bounds for integer programming. In the context of symmetric matrices, the simplest linear operators have the following form:

$$A(X) = MXM^T,$$

where M is an arbitrary $m \times n$ matrix. In this paper, we show that for such linear operators the optimization problem is trivial in the sense that an explicit solution can be given.

Key words: semidefinite programming.

AMS 1991 Subject Classification: Primary 65K10, Secondary 90C25

---

[*]Research partially support by AFOSR through grant AFOSR-91-0359.

1 Introduction.

For each integer n, let $\mathcal{S}^n$ denote the space of symmetric $n \times n$ matrices and let $\succeq$ denote the *Löwner* partial order on $\mathcal{S}^n$ induced by the cone of positive semidefinite matrices. The trace operator provides an inner product on $\mathcal{S}^n$ by the following formula

$$\langle C, D \rangle = \mathrm{tr}(CD).$$

With this inner product, $\mathcal{S}^n$ is a Hilbert space.

The *semidefinite programming problem* is defined as

$$\begin{array}{rc} \text{minimize} & \mathrm{tr}(CX) \\ \text{subject to} & A(X) = B \\ & X \succeq 0. \end{array} \tag{1.1}$$

Here, X and C belong to $\mathcal{S}^n$, B is an element of $\mathcal{S}^m$ $(m \leq n)$, and A is a linear operator from $\mathcal{S}^n$ to $\mathcal{S}^m$. This class of optimization problems is a subclass of the general problems considered in [4]. They have important applications in min-max eigenvalue problems (see e.g. [5, 6, 7]) and also in obtaining excellent bounds for branch-and-bound algorithms [2, 3, 1].

The dual of (1.1) is

$$\begin{array}{rc} \text{maximize} & \mathrm{tr}(BY) \\ \text{subject to} & A^T(Y) + Z = C \\ & Z \succeq 0. \end{array} \tag{1.2}$$

Here, A^T denotes the adjoint (or transpose) of A determined by the inner products on $\mathcal{S}^n$ and $\mathcal{S}^m$.

2 Linear Operators from $\mathcal{S}^n$ to $\mathcal{S}^m$.

It is interesting to consider what constitutes a linear operator from $\mathcal{S}^n$ to $\mathcal{S}^m$. The simplest such operators have the following form:

$$A(X) = MXM^T, \tag{2.1}$$

where M is an arbitrary $m \times n$ matrix. In this case, it is easy to see that

$$A^T(Y) = M^TYM.$$

Indeed, this fact follows from the following trace identity:

$$\mathrm{tr}(YMXM^T) = \mathrm{tr}(M^TYMX).$$

The second simplest such operators are of the following type:

$$A(X) = MXN^T + NXM^T, \tag{2.2}$$

where M and N are arbitrary $m \times n$ matrices. In this case,

$$A^T(Y) = M^TYN + N^TYM.$$

These operators play an important role in certain Lyaponov inequalities [8], which arise, for example, in the stability analysis of dynamical systems.

Theorem 1 *Every linear operator from $\mathcal{S}^n$ to $\mathcal{S}^m$ can be decomposed as a sum of operators of type (2.2).*

Proof. First consider linear operators from the space $\mathcal{R}^{n^2}$ of $n \times n$ matrices to the space $\mathcal{R}^{m^2}$ and for each i, j, k, l let L_{ijkl} be such an operator defined by

$$L_{ijkl}X = E_{ij}XE_{kl}^T,$$

where E_{ij} is the $m \times n$ matrix with all zeros except for a one in the (i, j)th position. It is easy to see that the L_{ijkl} are linear operators, that they are independent and since there are n^2m^2 of them they form a basis for the space of all such linear transformations. Hence, any such linear transformation A can be decomposed as

$$A(X) = \sum_{ijkl} a_{kl}^{ij} E_{ij}XE_{kl}^T.$$

Now to get the desired result for linear operators from $\mathcal{S}^n$ to $\mathcal{S}^m$, simply "lift" the operator to act in the larger space, employ the above decomposition and then "project" back down to the smaller space. $\square$

3 Duality

The following weak duality theorem connects the primal (1.1) to the dual (1.2):

Theorem 2 *Given X, Y, and Z, suppose that:*

1. *X is feasible for the primal;*
2. *the pair (Y, Z) is feasible for the dual;*
3. *the pair (X, Z) is complementary in the sense that $XZ = 0$.*

Then, X is optimal for the primal and the pair (Y, Z) is optimal for the dual.

Proof. The usual proof works. Indeed, first we use primal and dual feasibility to compute the duality gap:

$$\begin{aligned} \operatorname{tr}(CX) - \operatorname{tr}(ZX) &= \operatorname{tr}((C - Z)X) \\ &= \operatorname{tr}(A^T(Y)X) \\ &= \operatorname{tr}(YA(X)) \\ &= \operatorname{tr}(YB) \\ &= \operatorname{tr}(BY). \end{aligned}$$

Hence,

$$\operatorname{tr}(CX) - \operatorname{tr}(BY) = \operatorname{tr}(ZX) \geq 0.$$

Then, we remark that $\operatorname{tr}(ZX) = 0$ if and only if $ZX = 0$. (Actually, we only need the trivial direction, but for a proof of both directions, see [1].) □

4 The Semidefinite Program for Simple Operators.

In this section, we study the semidefinite programming problem in the case where the linear operator A is "simple". Indeed, throughout this section, we make the following assumptions:

Assumptions.

1. The linear operator A has the following simple form:

$$A(X) = MXM^T.$$

2. The primal problem is feasible.
3. The dual problem is strictly feasible.
4. The matrix M has full rank.

Hence, the semidefinite programming problem we wish to study in this section is

$$\begin{array}{rc} \text{minimize} & \operatorname{tr}(CX) \\ \text{subject to} & MXM^T = B \\ & X \succeq 0 \end{array} \tag{4.1}$$

and its dual is

$$\begin{array}{rc} \text{maximize} & \operatorname{tr}(BY) \\ \text{subject to} & M^TYM + Z = C \\ & Z \succeq 0. \end{array} \tag{4.2}$$

Note that the feasibility assumptions imply that B is positive semidefinite.

We shall show that the primal-dual pair (4.1),(4.2) can be solved explicitly. To this end, we first find an explicit solution to the following system:

$$MXM^T = B \tag{4.3}$$
$$M^TYM + Z = C \tag{4.4}$$
$$XZ = 0. \tag{4.5}$$

Then by magic we will see that X and Z are positive semidefinite. Therefore, Theorem 2 of the previous section will imply that the explicit solution indeed solves the semidefinite programming problem.

Let Q and R denote the matrices obtained from a QR-factorization of the $n \times n$ matrix $\left[\begin{array}{c|c} M^T & 0 \end{array}\right]$:

$$QR = \left[\begin{array}{c|c} M^T & 0 \end{array}\right].$$

Then, Q and R are both $n \times n$ matrices, $QQ^T = I$, and

$$R = \left[\begin{array}{c|c} P^T & 0 \\ \hline 0 & 0 \end{array} \right],$$

where P^T is an $m \times m$ upper triangular matrix. Let

$$J = \left[\begin{array}{c|c} I & 0 \end{array} \right]. \tag{4.6}$$

Then, M can be written as

$$M = PJQ^T.$$

Substituting this expression for M into (4.3) and (4.4), we get

$$\begin{aligned} PJQ^T XQJ^T P^T &= B \qquad (4.7) \\ QJ^T P^T YPJQ^T + Z &= C. \qquad (4.8) \end{aligned}$$

Now, we multiply (4.7) on the left and right by P^{-1} and P^{-T}, respectively. And, we multiply (4.8) on the left and right by Q^T and Q, respectively. The result is

$$\begin{aligned} JRJ^T &= G \qquad (4.9) \\ J^T SJ + T &= H, \qquad (4.10) \end{aligned}$$

where

$$\begin{aligned} R &= Q^T XQ, \qquad (4.11) \\ S &= P^T YP, \qquad (4.12) \\ T &= Q^T ZQ, \qquad (4.13) \\ G &= P^{-1} BP^{-T}, \qquad (4.14) \\ H &= Q^T CQ. \qquad (4.15) \end{aligned}$$

Now, partition R, T, and H to match the partition of J given in (4.6) and rewrite (4.9) and (4.10) accordingly to get

$$\left[\begin{array}{c|c} I & 0 \end{array} \right] \left[\begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{12}^T & R_{22} \end{array} \right] \left[\begin{array}{c} I \\ \hline 0 \end{array} \right] = G \tag{4.16}$$

and

$$\left[\begin{array}{c} I \\ \hline 0 \end{array}\right] S \left[\begin{array}{c|c} I & 0 \end{array}\right] + \left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline T_{12}^T & T_{22} \end{array}\right] = \left[\begin{array}{c|c} H_{11} & H_{12} \\ \hline H_{12}^T & H_{22} \end{array}\right]. \tag{4.17}$$

From (4.16), we see that

$$R_{11} = G \tag{4.18}$$

and from (4.17), we find that

$$S + T_{11} = H_{11} \tag{4.19}$$

$$T_{12} = H_{12} \tag{4.20}$$

$$T_{22} = H_{22} \tag{4.21}$$

Soon, we will need to know that H_{22} is invertible. To see this, first note that it depends only on the data for the problem:

$$H_{22} = (Q^T C Q)_{22}.$$

But also, (4.21) together with (4.13) implies that for any dual feasible pair (Y, Z), we have

$$H_{22} = T_{22} = (Q^T Z Q)_{22}.$$

Now our assumption of strict dual feasibility allows us to pick a positive definite Z from which we see that H_{22} is positive definite.

If we view (4.19) as defining S, then the arrays R_{12}, R_{22} and T_{11} still need to be determined. We also have not yet used the requirement that $XZ = 0$. This will give us more conditions which we will then use to determine these three remaining arrays. From (4.11) and (4.13), we see that

$$RT = Q^T X Q Q^T Z Q = Q^T X Z Q = 0. \tag{4.22}$$

Writing this equation in terms of partitioned matrices, we have

$$\left[\begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{12}^T & R_{22} \end{array}\right] \left[\begin{array}{c|c} T_{11} & T_{12} \\ \hline T_{12}^T & T_{22} \end{array}\right] = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \end{array}\right].$$

From this we get four equations:

$$R_{11}T_{11} + R_{12}T_{12}^T = 0 \tag{4.23}$$

$$R_{11}T_{12} + R_{12}T_{22} = 0 \tag{4.24}$$

$$R_{12}^T T_{11} + R_{22}T_{12}^T = 0 \tag{4.25}$$

$$R_{12}^T T_{12} + R_{22}T_{22} = 0. \tag{4.26}$$

Now, we use (4.24) to solve for R_{12},

$$\begin{aligned} R_{12} &= -R_{11}T_{12}T_{22}^{-1} \\ &= -GH_{12}H_{22}^{-1}, \end{aligned} \tag{4.27}$$

and then use (4.26) and (4.27) to solve for R_{22},

$$\begin{aligned} R_{22} &= -R_{12}^T T_{12}T_{22}^{-1} \\ &= H_{22}^{-T}H_{12}^T G^T H_{12}H_{22}^{-1}, \end{aligned} \tag{4.28}$$

and lastly use (4.26) and (4.23) to solve for T_{11},

$$\begin{aligned} T_{11} &= -R_{11}^{-1}R_{12}T_{12}^T \\ &= G^{-1}GH_{12}H_{22}^{-1}H_{12}^T \\ &= H_{12}H_{22}^{-1}H_{12}^T. \end{aligned} \tag{4.29}$$

With the formulas given by (4.27), (4.29), (4.28), and (4.20) for R_{12}, T_{11}, R_{22}, and T_{12}, respectively, it is easy to check that (4.25) holds automatically.

At this point, we reassemble matrices R and T from their respective submatrices to get

$$\begin{aligned} R &= \left[\begin{array}{c|c} G & -GH_{12}H_{22}^{-1} \\ \hline -H_{22}^{-T}H_{12}^T G & H_{22}^{-T}H_{12}^T GH_{12}H_{22}^{-1} \end{array}\right] \\ &= \left[\begin{array}{c} I \\ \hline -H_{22}^{-T}H_{12}^T \end{array}\right] G \left[\begin{array}{c|c} I & -H_{12}H_{22}^{-1} \end{array}\right] \end{aligned} \tag{4.30}$$

and

$$\begin{aligned} T &= \left[\begin{array}{c|c} H_{12}H_{22}^{-1}H_{12}^T & H_{12} \\ \hline H_{12}^T & H_{22} \end{array}\right] \\ &= \left[\begin{array}{c} H_{12}H_{22}^{-1/2} \\ \hline H_{22}^{1/2} \end{array}\right] \left[\begin{array}{c|c} H_{22}^{-1/2}H_{12}^T & H_{22}^{1/2} \end{array}\right] \end{aligned} \tag{4.31}$$

(where we have used the symmetry of G and H_{22} to obtain the given expression for R). Our feasibility assumption that B is positive semidefinite together with (4.14) implies that G is positive semidefinite. Hence, we see from (4.30) and (4.31) that R and T are positive semidefinite. Therefore, (4.11) and (4.13) show that X and Z are positive semidefinite too.

We have proved the following theorem:

Theorem 3 *The matrices*

$$\begin{aligned} X &= QRQ^T \\ Y &= P^{-T}SP^{-1} \\ Z &= QTQ^T, \end{aligned}$$

with R given by (4.30), S given by (4.19), and T given by (4.31) solves problems (4.1) and (4.2).

Perhaps a particularly interesting application of this theorem arises when M is simply a row vector a^T. In this case, the dual is especially attractive. Indeed, it is the problem of finding the largest scalar multiple of a symmetric rank one matrix that is not larger (in the Löwner sense) than a given matrix C:

$$\begin{array}{rc} \text{maximize} & y \\ \text{subject to} & yaa^T \preceq C. \end{array}$$

As far as we know, the explicit solution described in Theorem 3 is the first such solution given for this problem.

Acknowledgement: The authors would like to thank Henry Wolkowicz for introducing us to this problem area and for carefully critiquing early drafts of this paper.

References

[1] F. Alizadeh. Combinatorial optimization with semidefinite matrices. In *Proceedings of the Second Annual Integer Programming and Combinatorial Optimization Conference*, pages 373–395, Carnegie-Mellon University, 1992.

[2] N.K. Karmarkar and S.A. Thakur. An interior point approach to a tensor optimization problem with application to upper bounds in integer quadratic optimization problems. Technical report, AT&T Bell Laboratories, Murray Hill, NJ, 1992.

[3] L. Lovasz and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. Optimization*, 1(2):166–190, 1991.

[4] Y.E. Nesterov and A.S. Nemirovsky. *Interior Point Polynomial Methods in Convex Programming : Theory and Algorithms*. SIAM Publications. SIAM, Philadelphia, USA, 1993.

[5] M.L. Overton. On minimizing the maximum eigenvalue of a symmetric matrix. *SIAM J. Matrix Analysis and Applications*, 9:256–268, 1988.

[6] M.L. Overton. Large-scale optimization of eigenvalues. *SIAM J. Optimization*, 2:88–120, 1992.

[7] F. Rendl, R.J. Vanderbei, and H. Wolkowicz. Interior-point methods for max-min eigenvalue problems. Technical Report SOR 92-15, Princeton, University, 1992.

[8] L. Vandenberghe and S. Boyd. Primal-dual potential reduction method for problems involving matrix inequalities. Technical report, Electrical Engineering Department, Stanford University, Stanford, CA 94305, 1993.