The Simplest Semidefinite Programs are Trivial

Robert J. Vanderbei \(\ast\)  \hspace{1cm}  \text{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:

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

Here, \(\text{tr}(\cdot)\) 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

\(\ast\)Research partially support by AFOSR through grant AFOSR-91-0359.
1 Introduction.

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

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

With this inner product, $S^n$ is a Hilbert space.

The semidefinite programming problem is defined as

$$\begin{align*}
\text{minimize} & \quad \text{tr}(CX) \\
\text{subject to} & \quad A(X) = B \\
& \quad X \succeq 0.
\end{align*}$$

(1.1)

Here, $X$ and $C$ belong to $S^n$, $B$ is an element of $S^m$ ($m \leq n$), and $A$ is a linear operator from $S^n$ to $S^m$. 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

$$\begin{align*}
\text{maximize} & \quad \text{tr}(BY) \\
\text{subject to} & \quad A^T(Y) + Z = C \\
& \quad Z \succeq 0.
\end{align*}$$

(1.2)

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

2 Linear Operators from $S^n$ to $S^m$.

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

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

(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:
\[ \text{tr}(YMXM^T) = \text{tr}(M^TYMX). \]

The second simplest such operators are of the following type:
\[ A(X) = MXX^T + NXM^T, \quad (2.2) \]
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 Lyapunov inequalities [8], which arise, for example, in the stability analysis of dynamical systems.

**Theorem 1**  Every linear operator from \( S^n \) to \( S^m \) 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
\[ A(X) = \sum_{ijkl} a_{ijkl}^i X E_{kl}^T. \]

Now to get the desired result for linear operators from \( S^n \) to \( S^m \), 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:

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

Hence,

$$\text{tr}(C X) - \text{tr}(B Y) = \text{tr}(Z X) \geq 0.$$ 

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

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

$$\begin{align*}
\text{minimize} & \quad \text{tr}(CX) \\
\text{subject to} & \quad MXM^T = B \\
& \quad X \succeq 0
\end{align*}$$

and its dual is

$$\begin{align*}
\text{maximize} & \quad \text{tr}(BY) \\
\text{subject to} & \quad MTYM + Z = C \\
& \quad Z \succeq 0.
\end{align*}$$

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:

$$\begin{align*}
MXM^T &= B \quad (4.3) \\
MTYM + Z &= C \quad (4.4) \\
XZ &= 0. \quad (4.5)
\end{align*}$$

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 $\left[ \begin{array}{c} MT \\ 0 \end{array} \right]$

$$QR = \left[ \begin{array}{c} MT \\ 0 \end{array} \right].$$
Then, $Q$ and $R$ are both $n \times n$ matrices, $QQ^T = I$, and

$$R = \begin{bmatrix} P^T & 0 \\ 0 & 0 \end{bmatrix},$$

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

$$J = \begin{bmatrix} I & 0 \end{bmatrix}.$$  \hspace{1cm} (4.6)

Then, $M$ can be written as

$$M = PJQ^T.$$  

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

$$PJQ^T XQJ^T P^T = B \hspace{1cm} (4.7)$$

$$QJ^TP^TYPJQ^T + Z = C. \hspace{1cm} (4.8)$$

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

$$JRJ^T = G \hspace{1cm} (4.9)$$

$$J^T SJ + T = H \hspace{1cm} (4.10)$$

where

$$R = Q^T XQ, \hspace{1cm} (4.11)$$

$$S = P^TYP, \hspace{1cm} (4.12)$$

$$T = Q^TZQ, \hspace{1cm} (4.13)$$

$$G = P^{-1}BP^{-T}, \hspace{1cm} (4.14)$$

$$H = Q^TCQ. \hspace{1cm} (4.15)$$

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

$$\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} R_{11} & 0 \\ 0 & R_{22} \end{bmatrix} \begin{bmatrix} I \\ 0 \end{bmatrix} = G \hspace{1cm} (4.16)$$
and
\[
\begin{bmatrix}
I \\
0
\end{bmatrix}
S
\begin{bmatrix}
I & 0
\end{bmatrix}
+ \begin{bmatrix}
T_{11} & T_{12} \\
T_{12}^T & T_{22}
\end{bmatrix}
= \begin{bmatrix}
H_{11} & H_{12} \\
H_{12}^T & H_{22}
\end{bmatrix}.
\] (4.17)

From (4.16), we see that
\[R_{11} = G\] (4.18)
and from (4.17), we find that
\[
\begin{align*}
S + T_{11} &= H_{11} \\
T_{12} &= H_{12} \\
T_{22} &= H_{22}
\end{align*}
\] (4.19) (4.20) (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:
\[H_{22} = (Q^T CQ)_{22}.\]

But also, (4.21) together with (4.13) implies that for any dual feasible pair \((Y, Z)\), we have
\[H_{22} = T_{22} = (Q^T ZQ)_{22}.\]

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
\[RT = Q^T XQQ^T ZQ = Q^T XZQ = 0.\] (4.22)

Writing this equation in terms of partitioned matrices, we have
\[
\begin{bmatrix}
R_{11} & R_{12} \\
R_{12}^T & R_{22}
\end{bmatrix}
\begin{bmatrix}
T_{11} & T_{12} \\
T_{12}^T & T_{22}
\end{bmatrix}
= \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix}.
\]
From this we get four equations:
\[
\begin{align*}
R_{11}T_{11} + R_{12}T_{12}^T &= 0 \\
R_{11}T_{12} + R_{12}T_{22} &= 0 \\
R_{12}^T T_{11} + R_{22}T_{12}^T &= 0 \\
R_{12}^T T_{12} + R_{22} T_{22} &= 0.
\end{align*}
\] (4.23) (4.24) (4.25) (4.26)
Now, we use (4.24) to solve for $R_{12}$,
\[
R_{12} = -R_{11} T_{12} T_{22}^{-1}
= -G H_{12} H_{22}^{-1}, \tag{4.27}
\]
and then use (4.26) and (4.27) to solve for $R_{22}$,
\[
R_{22} = -R_{12}^T T_{12} T_{22}^{-1}
= H_{22}^{-T} H_{12} G^T H_{12} H_{22}^{-1}, \tag{4.28}
\]
and lastly use (4.26) and (4.23) to solve for $T_{11}$,
\[
T_{11} = -R_{11}^{-1} R_{12} T_{12}^T
= G^{-1} G H_{12} H_{22}^{-1} H_{12}^T
= H_{12} H_{22}^{-1} H_{12}^T. \tag{4.29}
\]

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
\[
R = \begin{bmatrix}
G & -G H_{12} H_{22}^{-1} \\
-H_{22}^{-1} H_{12}^T G & H_{22}^{-1} H_{12}^T G H_{12} H_{22}^{-1}
\end{bmatrix}
= \begin{bmatrix}
I & -H_{12} H_{22}^{-1}
\end{bmatrix}
G \begin{bmatrix}
I & -H_{12} H_{22}^{-1}
\end{bmatrix}
\tag{4.30}
\]
and
\[
T = \begin{bmatrix}
H_{12} H_{22}^{-1} H_{12}^T & H_{12} \\
H_{12}^T & H_{22}
\end{bmatrix}
= \begin{bmatrix}
H_{12} H_{22}^{-1/2} \\
H_{22}^{1/2}
\end{bmatrix}
\begin{bmatrix}
H_{22}^{-1/2} H_{12}^T & H_{12}^{1/2}
\end{bmatrix}
\tag{4.31}
\]
(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

\[
X = QRQ^T \\
Y = P^{-T}SP^{-1} \\
Z = QTQ^T,
\]

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 \( C \):

\[
\text{maximize } y \\
\text{subject to } yaa^T \preceq C.
\]

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


