Proceedings of ALGORITMY 2009
pp. 412–418

SEMIDEFINITE REPRESENTABILITY OF THE TRACE OF TOTALLY POSITIVE LAURENT POLYNOMIAL MATRIX FUNCTIONS*

MICHEL BAES†

Abstract. The function that maps a positive semidefinite matrix to the trace of one of its nonnegative integer power is semidefinite representable. In this note, we reduce the size of this semidefinite representation from $\mathcal{O}(kn)$ linear matrix inequalities of dimension n, where k is the desired power and $n \times n$ the size of the matrix to $\mathcal{O}(\log_2(k))$ linear matrix inequalities of dimension $2n$. We also propose a variant of our strategy that can deal with traces of negative powers.

Key words. Semidefinite optimization, power function, duality.

AMS subject classifications. 90C22, 15A48.

1. Introduction. A central preoccupation in optimization is to ascertain that a certain problem can be reliably solved within a predictably reasonably short time-span. Among the classes of problems that can be solved efficiently figures the class of *semidefinite optimization problems*, that is, as optimization problems with a linear objective, some linear equality constraints, and a semidefinite constraint. This class is now well-studied and many new applications have emerged, where semidefinite optimization plays a decisive role (see [2] and the references therein).

Nesterov and Nemirovski [5], and later Ben-Tal and Nemirovski [1], have defined formally the set of objects that can be used as building blocks for a semidefinite optimization problem. These objects should possess the property of *semidefinite representability*. Below, we denote by $\mathbf{S}^n$ the set of symmetric $n \times n$ matrices, and by $\mathbf{S}^n_+ \subset \mathbf{S}^n$ the cone of positive semidefinite matrices. We write $A \in \mathbf{S}^n_+$ and $A \succeq 0$ indifferently, and $A \succ 0$ when $A \in \operatorname{int} \mathbf{S}^n_+$. Also $A \succeq$ [resp. $\succ$]B iff $A - B \succeq$ [resp. $\succ$]0.

DEFINITION 1.1. *Let $Q \subseteq \mathbf{R}^m$ be a closed convex set. We say that Q is* semidefinite representable (SDr) *if and only if there exists two positive integers n, p, a linear operator $\mathcal{A} : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{S}^n$, and a matrix $B \in \mathbf{S}^n$ such that:*

$$x \in Q \iff \mathcal{A}(x, u) \succeq B.$$

We say that a convex function $f : S \subseteq \mathbf{R}^m \to \mathbf{R}$, where S is a convex set, is *semidefinite representable* if and only if its epigraph is SDr. Optimization problems involving a SDr objective function and linear equality constraints can be written as semidefinite optimization problems and solved efficiently using standard semidefinite optimization software such as SeDuMi [6] or SDPT3 [7], provided that the size of the resulting problem remains moderate.

Note that a function $f : Q \subseteq \mathbf{R}^m \to \mathbf{R}$ is SDr if it can be represented in the

*This work was supported by: ⋄ Research Council KUL: GOA AMBioRICS, CoE EF/05/006 Optimization in Engineering(OPTEC), IOF-SCORES4CHEM; ⋄ Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011).

†Institute For Operations Research, ETH Zürich, 101 Rämisstrasse, 8092 Zürich, Switzerland. Part of this work has been done while the author was at SCD/ESAT, 10 Kasteelpark Arenberg, 3001 Heverlee, Belgium.

following form:

$$\begin{aligned} f(x) := \min & \langle c, (x; u) \rangle \\ & \text{s.t.} \mathcal{A}(x, u) \succeq B, \end{aligned}$$

where $\mathcal{A}$ and B are defined as in Definition 1.1. Indeed, the epigraph of f can be described as:

$$\operatorname{epi} f = \{(t, x) : \exists u \in \mathbf{R}^p \text{ such that } \mathcal{A}(x, u) \succeq B \text{ and } \langle c, (x; u) \rangle \leq t\}.$$

We have borrowed the Matlab notation for $(x; u) := [x^T, u^T]^T$.

2. Semidefinite representability of power trace. We are investigating in this section the semidefinite representability of functions of the type:

$$\begin{aligned} f_k : \mathbf{S}^n_+ & \to \mathbf{R} \\ X & \mapsto f_k(X) := \operatorname{tr}(X^k), \end{aligned}$$

where $k \in \mathbf{Z}$. The semidefinite representability of totally positive Laurent polynomials, that is, functions of the type:

$$f(X) = \sum_{k \in \mathbf{Z}} a_k f_k(X),$$

where $\{a_k : k \in \mathbf{Z}\}$ is a sequence of nonnegative reals only a finite number of which are nonzero, would follow immediately.

It is well-known (e.g. as an application of the convexity result of Davis [3]) that f_k is a convex function for those k's for which the function $\mathbf{R}^+ \to \mathbf{R}^+, t \mapsto t^k$ is convex itself, that is, for $k \notin]0, 1[$. For $k \in]0, 1[$, the corresponding function f_k is concave.

Actually, a self-scaled representation of this function exists for $k \geq 0$, as a consequence of Proposition 4.2.2 in [1]. However, this representation is rather expensive. Their construction starts from a second-order representation of the function $x \mapsto g(x) := \sum_{i=1}^n x_i^k$ (See Section 3.3 in [1]) — this construction takes as much as $\mathcal{O}(k)$ second-order cones of dimension 3. We write this representation of $\{(x, t) : g(x) \leq t\}$ as $\exists u \in \mathbf{R}^l : \mathcal{A}(x, u, t) \in \mathcal{K}$, where $\mathcal{A}$ is affine and $\mathcal{K}$ an appropriate self-scaled cone. Now,

$$\begin{aligned} f_k(X) \leq t \iff & \exists x \in \mathbf{R}^n, u \in \mathbf{R}^l \text{ such that:} \\ & \mathcal{A}(x, u, t) \in K, \\ & x_1 \geq x_2 \geq \cdots \geq x_n, \\ & \sum_{i=1}^k \lambda_i(X) \leq \sum_{i=1}^k x_i \ \text{ for every } 1 \leq k \leq n. \end{aligned}$$

The last set of constraints are semidefinite representable. When $k = n$, the corresponding constraint is just linear. But for other values of k, the constraint can be represented by no less than two linear matrix inequalities of dimension n.

We describe below an alternative representation of the function f_k, which proves to be much cheaper. The fundamental principle on which our representation is based is the well-known Schur complement Lemma (see Theorem 7.7.6 in [4]), which we recall below.

LEMMA 2.1. *Let*

$$X := \begin{pmatrix} A & B^T \\ B & C \end{pmatrix}$$

be a real symmetric matrix. Then $X \succ 0$ *if and only if* $C \succ 0$ *and* $A \succ B^T C^{-1} B$. *Also if* $C \succ 0$ *and* $A \succeq B^T C^{-1} B$*, then* $X \succeq 0$.

For positive integers k, the matrix inequalities in the SDr of f_k can be constructed according to the following algorithm.

ALGORITHM 2.1.
Input: $k \in \mathbf{N}_0$.
Let $i := 0$, $m := k$.
while $m \geq 2$ ***do***
  if m *is odd*
    Add $\begin{pmatrix} X_i & X_{i+1} \\ X_{i+1} & X \end{pmatrix} \succeq 0$ *to the list of LMI's.*
    Let $i := i + 1$ *and* $m := (m + 1)/2$.
  end
  if m *is even*
    Add $\begin{pmatrix} X_i & X_{i+1} \\ X_{i+1} & I_n \end{pmatrix} \succeq 0$ *to the list of LMI's.*
    Let $i := i + 1$ *and* $m := m/2$.
  end
end
Add $X_i = X$ *to the set of constraints.* Observe that every matrix X_i is necessarily symmetric. The size of our representation is of the order of $\mathcal{O}(\log_2(k))$ linear matrix inequalities of dimension $2n \times 2n$, which is much smaller in terms of k and of n than the representation of Ben-Tal and Nemirovski.

The following proposition shows that the inequalities constructed in the previous algorithm correspond indeed to a SDr of f_k.

PROPOSITION 2.1. *For every positive* k*, the function* f_k *can be written as* $f_k(X) = \min\{\mathrm{tr}(X_0)$*, subject to the set of constraints generated by Algorithm 2.1 with input* $k\}$.

Proof. We denote by C_k the set of constraints generated by Algorithm 2.1 with input k.

First, we proceed to prove the inequality $f_k(X) \geq \min\{\mathrm{tr}(X_0) : C_k\}$. Let us fix $k \in \mathbf{N}_0$ and define the matrices X_i^* according to the following procedure:
  let $i := 0$, $m := k$.
  while $m \geq 2$ **do**
    $X_i^* := X^m$. Let $i := i + 1$ and $m := m/2 + (1 - (-1)^m)/4$.
  end
  $X_i^* := X$.
A simple verification shows that these matrices $X_0^*, X_1^*, \ldots, X_i^*$ satisfy the inequalities C_k. As $\mathrm{tr}(X_0^*) = \mathrm{tr}(X^k) = f_k(X)$, we have proved the desired inequality.

We use a duality argument to prove the reverse inequality. We fix a power k, and, for every iteration number $0 \leq i < N$, we denote by m_i the value of the variable m at the beginning of loop i. We write A_i for the $(2, 2)$-block of the i-th constraint matrix constructed in the algorithm (2.1). That is, $A_i = X$ if the variable m_i is odd, and $A_i = I_n$ otherwise. We also denote by N the value of the variable i at the end of the algorithm — actually, $N = \lceil \log_2(k) \rceil$.

The original problem can be written in the primal form

$$\min\left\{\langle C, M\rangle_F : \mathcal{A}M = B, M \in \left(\mathbf{S}_+^{2n}\right)^N\right\},$$

with $C = [I_n; 0_{n\times n}; \ldots; 0_{n\times n}]$, $B = [0_{2n\times n}; A_1; \cdots; 0_{2n\times n}; X; X; A_N]$, and, denoting

$$M = \operatorname{diag}\left(\begin{pmatrix} F_1 & G_1^T \\ G_1 & H_1 \end{pmatrix}, \ldots, \begin{pmatrix} F_N & G_N^T \\ G_N & H_N \end{pmatrix}\right),$$

the linear operator $\mathcal{A}$ takes the form:

$$\begin{array}{c} \left| \; F_1\, G_1^T G_1 H_1 \right| \; F_2\, G_2^T G_2 H_2 \left| \quad \cdots \quad \right| \; F_N\, G_N^T G_N H_N \left.\right| \\ \mathcal{A} = \left[\begin{array}{ccc|ccc|c|ccc} I_n & & & -I_n & & & & & & \\ & I_n & & -I_n & & & & & & \\ & & I_n & & & & & & & \\ \hline & & & I_n & & & & & & \\ & & & & I_n & & \ddots & & & \\ & & & & & I_n & & & & \\ \hline & & & & & & \ddots & -I_n & & \\ & & & & & & & -I_n & & \\ \hline & & & & & & & I_n & & \\ & & & & & & & & I_n & \\ & & & & & & & & & I_n \end{array}\right]. \end{array}$$

The dual of the above problem writes:

$$\max\left\{\langle B, Y\rangle_F : \mathcal{A}^T Y + S = C, C \in \left(\mathbf{S}_+^{2n}\right)^N\right\}.$$

After a few elementary manipulation, this dual takes the following form:

$$\begin{array}{rl} \max & \sum_{i=0}^{N-1}\langle Y_i, A_i\rangle + 2\langle Z_N, X\rangle \\ \text{s.t.} & \begin{pmatrix} I_n & -Z_1^T \\ -Z_1 & -Y_0 \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} Z_1 + Z_1^T & -Z_2^T \\ -Z_2 & -Y_1 \end{pmatrix} \succeq 0, \ldots, \\ & \begin{pmatrix} Z_{N-1} + Z_{N-1}^T & -Z_N^T \\ -Z_N & -Y_{N-1} \end{pmatrix} \succeq 0. \end{array}$$

Now, we proceed to construct a feasible dual point for which the objective value equals $\operatorname{tr}(X^k)$, thereby proving the theorem and actually providing all the information on the sensitivity of each primal constraint.

We set $Z_i^* := 2^i X^{k-m_i}$, which is therefore symmetric, and $Y_i^* := -2^i X^k A_i^{-1}$, so that $\langle Y_i^*, A_i\rangle = -2^i \operatorname{tr}(X^k)$. Observe that $Z_0^* = I_n$. Also, we set

$$Z_N^* := 2^{N-1} X^{k-m_N} = 2^{N-1} X^{k-1}.$$

The objective's value is therefore

$$\sum_{i=0}^{N-1} -2^i \operatorname{tr}(X^k) + 2^N \operatorname{tr}(X^k) = \operatorname{tr}(X^k).$$

If m_i is odd, then $A_i = X$, and the i-th matrix takes the form:

$$\begin{pmatrix} 2Z_i^* & -Z_{i+1}^* \\ -Z_{i+1}^* & -Y_i^* \end{pmatrix} = \begin{pmatrix} 2\cdot 2^{i-1}X^{k-m_i} & -2^i X^{k-(m_i+1)/2} \\ -2^i X^{k-(m_i+1)/2} & 2^i X^{k-1} \end{pmatrix}$$
$$= 2^i D^{1/2} \begin{pmatrix} X^{(1-m_i)/2} \\ -I_n \end{pmatrix} \begin{pmatrix} X^{(1-m_i)/2} & -I_n \end{pmatrix} D^{1/2}$$
$$\succeq 0,$$

where

$$D := \begin{pmatrix} X^{k-1} & 0 \\ 0 & X^{k-1} \end{pmatrix}.$$

If m_i is even, we have $A_i = I$, and the i-th constraint matrix is:

$$\begin{pmatrix} 2Z_i^* & -Z_{i+1}^* \\ -Z_{i+1}^* & -Y_i^* \end{pmatrix} = \begin{pmatrix} 2\cdot 2^{i-1}X^{k-m_i} & -2^i X^{k-m_i/2} \\ -2^i X^{k-m_i/2} & 2^i X^{k} \end{pmatrix}$$
$$= 2^i D^{1/2} \begin{pmatrix} X^{-m_i/2} \\ -I_n \end{pmatrix} \begin{pmatrix} X^{-m_i/2} & -I_n \end{pmatrix} D^{1/2} \succeq 0.$$

As the point we have constructed is feasible for the dual problem and attains an objective's value of $\mathrm{tr}(X^k)$, we have proved the inequality $f_k(X) \le \min\{\mathrm{tr}(X_0) : C_k\}$. ∎

For negative indices k, we use the following consequence of Schur's Lemma:

$$\begin{pmatrix} X & I \\ I & Y \end{pmatrix} \succ 0 \iff X \succ 0 \text{ and } Y \succ X^{-1}. \tag{1}$$

ALGORITHM 2.2.
Input: *k,* negative *integer.*
Let $i := 0$, $m := -k$.
while *$m \ge 2$* ***do***
  if *m is odd*
    Add $\begin{pmatrix} X_i & X_{i+1} \\ X_{i+1} & X \end{pmatrix} \succeq 0$ *to the list of LMI's.*
    Let $i := i+1$ and $m := (m-1)/2$.
  end
  if *m is even*
    Add $\begin{pmatrix} X_i & X_{i+1} \\ X_{i+1} & I_n \end{pmatrix} \succeq 0$ *to the list of LMI's.*
    Let $i := i+1$ and $m := m/2$.
  end
end
Add $\begin{pmatrix} X_i & I_n \\ I_n & X \end{pmatrix} \succeq 0$ *to the set of constraints.*

The proof of this algorithm follows the same lines as the one for positive k's.

PROPOSITION 2.2. *For every negative k, the function f_k can be written as $f_k(X) = \min\{\mathrm{tr}(X_0)$, subject to the set of constraints generated by Algorithm 2.2 with input $k\}$.*

Proof. We denote by D_k the set of constraints generated by Algorithm 2.2 with input k.

Let us fix $k < 0$, and decompose it as $k = -\sum_{i=0}^{N} a_i 2^i$, with $a_i \in \{0, 1\}$. We also define $m_i := \sum_{j=i}^{N} a_j 2^{j-i}$, which is exactly the value of the variable m at the beginning of loop i in Algorithm 2.2. We finally write $A_i := X^{a_i}$, so that $A_N = X$.

It can be easily checked that the matrices $X_i^* := X^{-m_i}$ for $0 \leq i \leq N$ satisfy D_k. Therefore $\mathrm{tr}(X^k) = \mathrm{tr}(X_0^*) \geq \min\{\mathrm{tr}(X_0) : D_k\}$.

Now, the problem $\min\{\mathrm{tr}(X_0) : D_k\}$ can be written in the primal form

$$\min\left\{\langle C, M\rangle_F : \mathcal{A}M = B, M \in \left(\mathbf{S}_+^{2n}\right)^N\right\},$$

where

$$C = [I_n; 0_{n\times n}; \ldots; 0_{n\times n}]$$

and

$$B = [0_{2n\times n}; A_1; \cdots; A_{N-1}; I_n; I_n; A_N].$$

The matrix $\mathcal{A}$ is as in the proof of Proposition 2.1. The dual of this problem, after a few trivial simplifications reads as:

$$\begin{array}{ll} \max & \sum_{i=0}^{N-1} \langle Y_i, A_i\rangle + 2\langle Z_N, I_n\rangle \\ \text{s.t.} & \begin{pmatrix} I_n & -Z_1^T \\ -Z_1 & -Y_0 \end{pmatrix} \succeq 0, \\ & \begin{pmatrix} 2Z_1 & -Z_2^T \\ -Z_2 & -Y_1 \end{pmatrix} \succeq 0, \ldots, \\ & \begin{pmatrix} 2Z_{N-1} & -Z_N^T \\ -Z_N & -Y_{N-1} \end{pmatrix} \succeq 0. \end{array}$$

A feasible point is given by $Y_i^* := -2^i X^{k-a_i}$ for $0 \leq i < N$, and $Z_i^* := 2^{i-1} X^{k+m_i}$ for $0 \leq i \leq N$, which is symmetric. Note that $Z_0^* = I_n/2$. We have:

$$\begin{aligned} \begin{pmatrix} 2Z_i^* & -Z_i^{*T} \\ -Z_i^* & -Y_i^* \end{pmatrix} &= \begin{pmatrix} 2\cdot 2^{i-1} X^{k+m_i} & -2^i X^{k+(m_i-a_i)/2} \\ -2^i X^{k+(m_i-a_i)/2} & 2^i X^{k-a_i} \end{pmatrix} \\ = 2^i \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}^{\frac{k+m_i}{2}} &\begin{pmatrix} -I_n \\ X^{-\frac{m_i-a_i}{2}} \end{pmatrix} \begin{pmatrix} -I_n & X^{-\frac{m_i-a_i}{2}} \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & X \end{pmatrix}^{\frac{k+m_i}{2}} \\ \succeq 0. \end{aligned}$$

Finally, this feasible point brings the dual objective to a value of:

$$-\sum_{i=0}^{N-1} 2^i \langle X^{k-a_i}, X^{a_i}\rangle + 2\cdot 2^{N-1}\langle X^{k+m_N}, I_n\rangle = \mathrm{tr}(X^k),$$

since $m_N = 0$. The desired inequality is thereby proved. □

Acknowledgements. I warmly thank Prof. Yurii Nesterov for having introducing me to the problem during a stimulating discussion.

REFERENCES

[1] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS/SIAM Series on Optimization, vol. 2, SIAM, 2001.

[2] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
[3] C. Davis, *All convex invariant functions of Hermitian matrices*, Archiv der Mathematik, 8 (1957), pp. 276278.
[4] R. Horn and Ch. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.
[5] Y. Nesterov and A. Nemirovski, *Interior Point Polynomial Algorithms in Convex Programming*, Studies in Applied Mathematics, no. 13, SIAM, 1993.
[6] J. Sturm, *Using sedumi 1.02, a Matlab toolbox for optimization over symmetric cones*, Optimization Methods and Software, 11 (1999), pp. 625653.
[7] R. Tutuncu, K. Toh, and M. Todd, *Solving semidenite-quadratic-linear programs using SDPT3*, Mathematical Programming, Series B, 95 (2003), 189217.