ISE Industrial and Systems Engineering

On the computational complexity of gap-free duals for semidefinite optimization

Imre Pólik
Lehigh University

Tamás Terlaky
Lehigh University

Report: 09T-016

On the computational complexity of gap-free duals for semidefinite optimization

Imre Pólik*, Tamás Terlaky

Department of Industrial & Systems Engineering, Lehigh University
200 West Packer Ave, Bethlehem, PA, USA, 18015

Abstract

We consider the complexity of gap-free duals in semidefinite programming. Using the theory of homogeneous cones we provide a new representation of Ramana's gap-free dual and show that the new formulation has a much better complexity than originally proved by Ramana.

Key words: gap-free duals, Ramana-dual, semidefinite optimization, Schur complement, Siegel cone, homogeneous cones

1. Introduction

In his seminal paper [6], Ramana presented an exact duality theory for semidefinite optimization (SDP) without any constraint qualification. Later, in a joint paper [8] with Tunçel and Wolkowicz, they showed how the result can be derived from a more general theorem for convex problems [1]. The title of Ramana's original paper proves that he was aware of the complexity implications of his dual problem. Here we show that the complexity of a gap-free dual can be better than what Ramana had proved, as we improve it by roughly a factor of 2. For this we use a construction from the theory of homogeneous cones: the Siegel cone.

1.1. Notations

Let us briefly review the notation we are using in the paper. The set of nonnegative (positive) numbers is denoted by $\mathbb{R}_+$ ($\mathbb{R}_{++}$). Vectors throughout the paper are denoted by lowercase Latin letters and are assumed to be column vectors. Matrices are denoted by uppercase letters. If v is a vector then v_i is its i^{th} component. Let $\mathbb{S}^n$ denote the set of real symmetric matrices, $\mathbb{S}^n_+$ the set of $n \times n$ symmetric positive semidefinite matrices and $\mathbb{S}^n_{++}$ the set of $n \times n$ symmetric positive definite matrices. We will write $X \succeq Y$ ($X \succ Y$) to denote that $X - Y$ is positive semidefinite (definite). For matrices X and Y the scalar product will be denoted by $X \bullet Y = \operatorname{Tr}(X^T Y)$. In general, $\mathcal{K}$ will be a closed, convex, pointed cone with non-empty interior.

2. The Lagrange dual for SDP

For the purpose of this paper the primal semidefinite optimization problem is

$$\begin{aligned} \max\ & b^T y \\ \sum_{i=1}^m y_i A_i + S &= C \\ S &\succeq 0, \end{aligned} \tag{SDP$_\mathrm{P}$}$$

where $C, S, A_i, i = 1, \ldots, m$ are real, symmetric, $n \times n$ matrices, $b, y \in \mathbb{R}^m$. The standard Lagrange-Slater dual of (SDP$_\mathrm{P}$) is:

$$\begin{aligned} \min\ & C \bullet X \\ A_i \bullet X &= b_i,\ i = 1, \ldots, m \\ X &\succeq 0, \end{aligned} \tag{SDP$_\mathrm{D}$}$$

where $X \in \mathbb{S}^n$. The basic duality relations of the primal and dual problems are (see, e.g., [10]):

Weak duality: For any feasible solution y and S of (SDP$_\mathrm{P}$) and feasible solution X of (SDP$_\mathrm{D}$)

*Corresponding author.
Email addresses: imre@polik.net (Imre Pólik), terlaky@lehigh.edu (Tamás Terlaky)

we have $b^T y \leq C \bullet X$. In particular the optimal value of ($\mathrm{SDP_P}$) is always less than or equal to the optimal value of ($\mathrm{SDP_D}$).

Strong duality: If problem ($\mathrm{SDP_P}$) (or ($\mathrm{SDP_D}$)) has a strictly feasible solution, then the optimal values of the two problems are equal, and the optimum is attained in problem ($\mathrm{SDP_D}$) (or ($\mathrm{SDP_P}$)). Moreover, if both problems are strictly feasible, then the optimal values are equal and they are attained on both the primal and dual sides.

In general, the existence of a strictly feasible solution (or some other form of constraint qualification) is necessary to establish strong duality for the Lagrange-dual. The following example with positive duality gap is taken from [5]:

$$\begin{gathered} \max y_2 \\ \begin{pmatrix} y_2 & 0 & 0 \\ 0 & y_1 & y_2 \\ 0 & y_2 & 0 \end{pmatrix} \preceq \begin{pmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{gathered} \tag{1}$$

For any feasible solution of this problem we have $y_2 = 0$ so the optimal objective value is 0. However, the dual problem

$$\begin{aligned} \min \alpha X_{11}& \\ X_{22} &= 0 \\ X_{11} + 2X_{23} &= 1 \\ \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} &\succeq 0 \end{aligned} \tag{2}$$

has $X_{22} = X_{23} = 0$ and $X_{11} = 1$, thus the optimal value is α. This also shows that the duality gap can be arbitrarily large. The set of feasible solutions for the primal problem is $y_1 \leq 0, y_2 = 0$, so there is no strictly feasible primal solution. Moreover, the dual problem has $X_{22} = 0$, so it does not have a positive definite feasible solution either.

2.1. Strong duality by regularization

Another approach to establish strong duality is to employ an equivalent reformulation of the primal problem, and then dualize that. This process is called regularization, and was first developed for general convex optimization by Borwein and Wolkowicz [1]. For the case of semidefinite optimization, the regularized dual was also independently obtained by Ramana [6], see also [7, 8]. He established the following extended dual problem:

$$\begin{aligned} \min C \bullet (X^{k+1} + Z^{k+1}) & \qquad (\mathrm{SDP_{Ramana}}) \\ A_i \bullet (X^{k+1} + Z^{k+1}) &= b_i, \; i = 1, \ldots, m \\ C \bullet (X^j + Z^j) &= 0, \; j = 1, \ldots, k \\ A_i \bullet (X^j + Z^j) &= 0, 1 \leq j \leq k, 1 \leq i \leq m \\ Z^1 &= 0 \\ X^{k+1} &\in \mathbb{S}^n_+ \\ \begin{pmatrix} X^j & {Z^{j+1}}^T \\ Z^{j+1} & I \end{pmatrix} &\in \mathbb{S}^{2n}_+, \; j = 1, \ldots, k, \end{aligned}$$

where $k \leq \min\{n, m\}$ is an integer constant depending on the geometry of the problem. Using the Schur complement (see, e.g., [13]) the last constraint is equivalent to

$$X^j - {Z^{j+1}}^T Z^{j+1} \in \mathbb{S}^n_+, \; j = 1, \ldots, k.$$

The special form of this constraint motivates our generalization.

2.2. The complexity of the original Ramana dual

Optimization problems over convex cones are usually solved with interior point methods (IPMs). From the viewpoint of the present paper the internal workings of these algorithms are not so important, the interested reader is directed to the literature, most importantly [5, 9]. These methods solve problems ($\mathrm{SDP_P}$) and ($\mathrm{SDP_D}$) to precision ε in at most $\mathcal{O}(\sqrt{n} \log(1/\varepsilon))$ iterations.

In general, to solve an optimization problem over a closed, convex, pointed, solid cone $\mathcal{K}$ to precision ε, IPMs take $\mathcal{O}\left(\sqrt{\vartheta} \log(1/\varepsilon)\right)$ iterations, where ϑ is a complexity parameter depending only on the cone $\mathcal{K}$. It is important to note that the iteration complexity does not depend on the dimension of $\mathcal{K}$ or the number of linear equalities, however, the cost of one iteration is determined by these quantities. There are several estimates for ϑ depending on the geometric and algebraic properties of the cone. For the cone of $n \times n$ positive semidefinite matrices we have $\vartheta(\mathbb{S}^n_+) = n$. This implies that the complexity parameter of Ramana's gap-free dual ($\mathrm{SDP_{Ramana}}$) is $2nk + n$, since the last constraint uses $2n \times 2n$ matrices.

3. A different look at the Schur complement

In what follows we will provide a new analysis of the complexity parameter of the last constraint in ($\mathrm{SDP_{Ramana}}$).

3.1. Homogeneous cones

First, we need some background material about homogeneous cones. We only provide the material that is needed for this paper, for more details and proofs see [11, 12].

Definition 3.1. *A closed, convex cone $\mathcal{K} \subseteq \mathbb{R}^n$ with nonempty interior is* homogeneous *if the group of automorphisms of $\mathcal{K}$ acts transitively on the interior of $\mathcal{K}$.*

In some sense the definition requires that all the interior points of the cone behave the same way. Typical examples of homogeneous cones are n-dimensional polyhedral cones in $\mathbb{R}^n$ with exactly n extreme rays (e.g., the nonnegative orthant), the set of symmetric positive semidefinite matrices and the second-order or Lorentz cone.

3.2. The Siegel cone

The standard construction of homogeneous cones is via T-algebras, but they can also be constructed in a recursive way, due to Vinberg [11]. Here we follow [4].

Definition 3.2. *Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a homogeneous cone, and consider a symmetric bilinear mapping $B : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^n$ such that for every $u, v \in \mathbb{R}^p$ and $\lambda_1, \lambda_2 \in \mathbb{R}$*

1. $B(\lambda_1 u_1 + \lambda_2 u_2, v) = \lambda_1 B(u_1, v) + \lambda_2 B(u_2, v)$,
2. $B(u, v) = B(v, u)$,
3. $B(u, u) \in \mathcal{K}$,
4. $B(u, u) = 0$ *implies* $u = 0$.

A symmetric bilinear form B is called homogeneous *if $\mathcal{K}$ is a homogeneous cone and there is a transitive subgroup $G \subsetneq \mathrm{Aut}(\mathcal{K})$ such that for every $g \in G$ there is a linear transformation $\tilde{g}$ on $\mathbb{R}^p$ such that $g(B(u,v)) = B(\tilde{g}(u), \tilde{g}(v))$. In other words, the diagram*

$$\begin{tikzcd} \mathbb{R}^p \times \mathbb{R}^p \arrow[r, "\tilde{g}\times\tilde{g}"] \arrow[d, "B"'] & \mathbb{R}^p \times \mathbb{R}^p \arrow[d, "B"] \\ \mathbb{R}^n \arrow[r, "g"] & \mathbb{R}^n \end{tikzcd} \tag{3}$$

commutes.

Having such a bilinear function we can define a new set:

Definition 3.3. *The Siegel cone $\mathrm{SC}(\mathcal{K}, B)$ of $\mathcal{K}$ and B is defined as the closure*

$$\overline{\{(x, u, t) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}_{++} : tx - B(u, u) \in \mathcal{K}\}}.$$

This concept is analogous to the Schur complement for semidefinite matrices. Using the Siegel cone we can express a quadratic relation with a linear one of higher rank and dimension.

Theorem 3.4. *The Siegel cone has the following properties (see [4] for proofs):*

1. *If $\mathcal{K}$ is a homogeneous cone and B is a homogeneous symmetric bilinear form then $\mathrm{SC}(\mathcal{K}, B)$ is a homogeneous cone.*
2. *Every homogeneous cone can be obtained as the Siegel cone of another homogeneous cone using an appropriate bilinear function B.*
3. *The complexity parameter of a Siegel cone is*
$$\vartheta(\mathrm{SC}(\mathcal{K}, B)) = \vartheta(\mathcal{K}) + 1.$$
4. *A self-concordant barrier function (see [5]) with barrier parameter $\vartheta(\mathcal{K}) + 1$ for the Siegel cone is*
$$\phi_{\mathrm{SC}(\mathcal{K},B)}(x, u, t) = \phi_{\mathcal{K}}\left(x - \frac{B(u,u)}{t}\right) - \ln t,$$
where $\phi_{\mathcal{K}}(.)$ is a self-concordant barrier function for cone $\mathcal{K}$ with barrier parameter $\theta(\mathcal{K})$.

The last result in crucial in applying an interior-point method to solve an optimization problem over homogeneous cones, or the Siegel cone in particular.

3.3. The Schur complement as a Siegel cone

Now we can see how the Schur complement constraint in ($\mathrm{SDP}_{\mathrm{Ramana}}$) could be expressed differently. The cone of symmetric, positive semidefinite matrices is homogeneous. The bilinear function $B(U,V) = \frac{1}{2}(U^T V + V^T U)$ is a symmetric bilinear function. We need to show that it is a homogeneous form with respect to $\mathbb{S}^n_+$. For this, note that the automorphisms of $\mathbb{S}^n_+$ are of the form $g : X \mapsto P^T X P$, where P is an orthogonal matrix. Using this we have

$$\begin{aligned} g(B(U,V)) &= P^T \frac{U^T V + V^T U}{2} P \\ &= \frac{(P^T U^T P)(P^T V P) + (P^T V^T P)(P^T U P)}{2} \\ &= B(P^T U P, P^T V P) = B(g(U), g(V)), \end{aligned}$$

which proves that B is homogeneous with respect to $\mathbb{S}^n_+$. By Theorem 3.4 this implies that the set

$$\mathrm{SC}(\mathbb{S}^n_+, B) = \overline{\{(X, U, t) \in \mathbb{S}^n_+ \times \mathbb{R}^{n\times n} \times \mathbb{R}_{++} : tX - U^T U \in \mathbb{S}^n_+\}}$$

is a Siegel cone, thus it is homogeneous and its complexity parameter is $\vartheta(\mathrm{SC}(\mathbb{S}^n_+, B)) = n + 1$.

3.4. The complexity of the improved dual

On the other hand, the relation $tX - U^T U \in \mathbb{S}^n_+$ is routinely expressed using the Schur complement as

$$\begin{pmatrix} X & U^T \\ U & tI \end{pmatrix} \in \mathbb{S}^{2n}_+.$$

Using this construction we can replace the last constraint in problem ($\mathrm{SDP}_{\mathrm{Ramana}}$) with a Siegel cone constraint. This yields the following homogeneous cone optimization problem

$$\begin{aligned} \min\ & C \bullet (X^{k+1} + Z^{k+1}) \qquad\qquad (\mathrm{SDP'}_{\mathrm{Ramana}}) \\ A_i \bullet & \left(X^{k+1} + Z^{k+1}\right) = b_i,\ i = 1, \ldots, m \\ & C \bullet \left(X^j + Z^j\right) = 0,\ j = 1, \ldots, k \\ & A_i \bullet \left(X^j + Z^j\right) = 0,\ 1 \le j \le k, 1 \le i \le m \\ & Z^1 = 0 \\ & X^{k+1} \in \mathbb{S}^n_+ \\ & \left(X^j, Z^{j+1}, 1\right) \in \mathrm{SC}\left(\mathbb{S}^n_+, B\right), j = 1, \ldots, k \end{aligned}$$

The complexity parameter of this problem is only $(n+1)k+n$, which is roughly half of the complexity parameter of ($\mathrm{SDP}_{\mathrm{Ramana}}$).

3.5. Another look at the Schur complement

There is another interpretation of this result. Homogeneous cones can also be represented as slices of the positive semidefinite cone, more specifically we have:

Lemma 3.5 (Chua, [2]). *If $\mathcal{K} \subset \mathbb{R}^n$ is a homogeneous cone then there exist an $n' \le n$ and an injective linear map $M : \mathbb{R}^n \to \mathbb{S}^{n'}$ such that $M(\mathcal{K}) = \mathbb{S}^{n'}_+ \cap M(\mathbb{R}^n)$. In other words, any n-dimensional homogeneous cone can be represented as the intersection of an affine space and the cone of $n \times n$ positive semidefinite matrices.*

A similar result was obtained independently by Faybusovich, [3]. We have to note here, that not all slices of the positive semidefinite cone are homogeneous, see [2] for a counterexample. In light of this result, the Schur complement is a very efficient way to represent the homogeneous cone $\mathrm{SC}\left(\mathbb{S}^n_+, B\right)$ as a slice of $\mathbb{S}^{2n}_+$, as Lemma 3.5 only guarantees the existence of a representation with matrices of size $n(n+1)/2 + n^2 + 1$.

We also have an explicit self-concordant barrier function for the Siegel cones (see Thm (3.4)), so we can solve the extended dual problem ($\mathrm{SDP'}_{\mathrm{Ramana}}$) with interior-point methods.

4. Conclusions

We showed that the complexity parameter of a new representation of Ramana's dual for semidefinite programming is at most $(n+1)L+n$, improving the original bound of $2nL + n$. On the other hand, the new formulation ($\mathrm{SDP'}_{\mathrm{Ramana}}$) is a general homogeneous cone optimization problem. In particular, it is not self-dual, although its dual can be generated easily. Also, it cannot be solved directly with current optimization solvers.

The problem of finding an *explicitly formed* gap-free dual problem to ($\mathrm{SDP}_{\mathrm{P}}$) with the lowest possible complexity parameter is still open. It may even be possible to further improve the complexity of the Ramana dual by exploiting the overall structure of the problem.

Acknowledgements

The authors were supported by the NSERC Discovery Grant #5-48923, the Canada Research Chair Program, MITACS and the start-up grants of Lehigh University.

The idea to consider homogeneous cones in the context of strong conic duality was inspired by a seminar talk by Chek Beng Chua at McMaster University in May 2006.

References

[1] Borwein, J., Wolkowicz, H., 1981. Regularizing the abstract convex program. Journal of Mathematical Analysis and Applications 83, 495–530.

[2] Chua, C. B., 2003. Relating homogeneous cones and positive definite cones via T-algebras. SIAM Journal on Optimization 14 (2), 500–506.

[3] Faybusovich, L., 2002. On Nesterov's approach to semi-infinite programming. Acta Applicandae Mathematicae 74, 195–215.

[4] Güler, O., Tunçel, L., 1998. Characterization of the barrier parameter of homogeneous convex cones. Math. Programming 81, 55–76.

[5] Nesterov, Y., Nemirovski, A., 1994. Interior-Point Polynomial Algorithms in Convex Programming. SIAM Publications, Philadelphia, PA.

[6] Ramana, M. V., 1997. An exact duality theory for semidefinite programming and its complexity implications. Math. Programming 77, 129–162.

[7] Ramana, M. V., Freund, R. M., 1996. On the ELSD duality theory for SDP, unpublished manuscript.

[8] Ramana, M. V., Tunçel, L., Wolkowicz, H., 1997. Strong duality for semidefinite programming. SIAM Journal on Optimization 7 (3), 641–662.

[9] Renegar, J., 2001. A Mathematical View of Interior-Point Methods in Convex Optimization. MPS/SIAM Series on optimization. SIAM.

[10] Rockafellar, R. T., 1970. Convex Analysis. Princeton University Press, Princeton, N. J.
[11] Vinberg, E. B., 1963. The theory of convex homogeneous cones. Transactions of the Moscow Mathematical Society 12, 340–403.
[12] Vinberg, E. B., 1965. The structure of the group of automorphisms of a homogeneous convex cone. Transactions of the Moscow Mathematical Society 13, 63–93.
[13] Zhang, F. (Ed.), 2005. The Schur complement and its applications. Vol. 4 of Numerical Methods and Algorithms. Springer.