Necessary and sufficient global optimality conditions for NLP reformulations of linear SDP problems

L. Grippo*, L. Palagi*, V Piccialli*

*Sapienza Università di Roma
Dipartimento di Informatica e Sistemistica "A. Ruberti"
Via Ariosto 25 - 00185 Roma - Italy

e-mail (Grippo): grippo@dis.uniroma1.it
e-mail (Palagi): palagi@dis.uniroma1.it
e-mail (Piccialli): piccialli@dis.uniroma1.it

Abstract

In this paper we consider the standard linear SDP problem, and its low rank nonlinear programming reformulation, based on a Gramian representation of a positive semidefinite matrix. For this nonconvex quadratic problem with quadratic equality constraints, we give necessary and sufficient conditions of global optimality expressed in terms of the Lagrangian function.

Keywords: semidefinite programming - low-rank factorization - optimality conditions.

1 Introduction

The standard linear SDP problem we consider is of the form:

$$\begin{array}{ll} \min & \operatorname{trace}(QX) \\ & \operatorname{trace}(A_iX) = b_i, \quad i = 1,\dots,m \\ & X \succeq 0, \quad X \in \mathcal{S}^n, \end{array} \tag{SDP}$$

where the data matrices Q and A_i for $i = 1,\dots,m$ are $n \times n$ real symmetric matrices, trace (QX) denotes the trace-inner product of matrices, and the $n \times n$ matrix variable X is required to be symmetric and positive semidefinite, as indicated with the notation $X \succeq 0,\ X \in \mathcal{S}^n$, where $\mathcal{S}^n$ is the space of real $n \times n$ symmetric matrices.
This class of problems contains important problems as special cases, such as linear or quadratic programming, and arises in a wide variety of applications in system and control theory, combinatorial optimization, approximation theory, robust optimization, and mechanical and electrical engineering.
Among the main approaches for solving linear SDP problems are interior point methods (see for example the survey [10] and references therein), and first order non linear programming methods ([2]). Interior point methods are in general able to solve SDP problems of small or medium size with high accuracy, and are proved to converge in polynomial time to an ϵ optimal solution. However, in practice, the dimension of the problem that can be solved is still limited to a maximum of a few thousand variables for the most efficient codes. First order nonlinear programming methods work efficiently in practice and can solve much larger problems, but convergence to a global solution is not guaranteed. We are interested in this class of methods, and the aim of this paper is to give a theoretical result that further justifies this approach.
Burer and Monteiro in [2, 3] recast a general linear SDP problem as a low rank semidefinite programming problem (LRSDP) by applying the change of variables $X = RR^T$, where R is a $n \times r$, $r < n$, rectangular matrix. They get the following problem

$$\begin{array}{ll} \min & \operatorname{trace}(QRR^T) \\ & \operatorname{trace}(A_iRR^T) = b_i, \quad i = 1,\dots,m \\ & R \in \mathbb{R}^{n\times r}, r \le n. \end{array} \tag{1}$$

The value of r is chosen by exploiting the result proved in Barvinok [1] and Pataki [12], that states that, under suitable assumptions, there exists an optimal solution of a linearly constrained SDP problem with rank r satisfying $r(r+1) \le m$, where m is the number of linear constraints. Problem (1) is a nonlinear programming problem; although it has been

proved [1, 12, 2] that a value of r exists such that there is a one-to one correspondence between global solutions of Problem (1) and global solutions of Problem (SDP), Problem (1) is a non convex problem, so that recognizing a global solution is a difficult task.
In this paper, under suitable assumptions on Problem (SDP), we state necessary and sufficient global optimality conditions in terms of the Lagrangian function. Similar conditions have already been proved to be sufficient in [3]. These conditions extend the necessary and sufficient ones proved in [5] for the special case of the semidefinite relaxation of the max cut problem and can be related to the necessary and sufficient global optimality conditions established for some classes of nonconvex quadratic problems (see e.g. [6, 11, 9]).

2 Low rank SDP formulations

It has been proved in [2, 5] that if Problem (SDP) admits a solution X^* or rank r, this can be found by solving Problem (1). Actually Problem (1) can be rewritten as a standard nonlinear programming problem, by setting $R = (v_1 \dots v_n)^T$, $v_k \in \mathbb{R}^r$ for $k = 1, \dots, n$ so that $(RR^T)_{ij} = v_i^T v_j$ and we get:

$$\min \quad q_r(v) := \sum_{i=1}^{n} \sum_{j=1}^{n} q_{ij} v_i^T v_j \tag{2}$$
$$\sum_{k=1}^{n} \sum_{j=1}^{n} (A_i)_{kj} v_k^T v_j = b_i, \; i = 1, \dots, m, \; v_k \in \mathbb{R}^r, k = 1, \dots, n.$$

The following result proved in [1] and [12] provides a useful upper bound on the value of r, that can be easily computed.

Proposition 1 *Suppose that the feasible set of Problem* (SDP) *has an extreme point. Then there exists an $X \in \mathcal{S}^n$ optimal solution of* (SDP) *with rank r satisfying the inequality*

$$r(r+1)/2 \leq m. \qquad \square$$

This result implies that $r \leq \widehat{r}$ where

$$\widehat{r} = \max\{k \in N : k(k+1)/2 \leq m\} = \left\lfloor \frac{\sqrt{1+8m}-1}{2} \right\rfloor. \tag{3}$$

Therefore, for sufficiently large values of r, a global solution of Problem (2) gives a global solution of Problem (SDP). In particular, Problem (2) gives a global solution of Problem (SDP) for all $r \geq \widehat{r}$.

We can recast Problem (2) in compact standard vector notation of nonlinear programming form by using Kronecker products $\otimes$ (see, for instance, [8]). We recall that given two matrices A $m \times n$ and B $p \times q$, the Kronecker product $A \otimes B$ is the $mp \times nq$ matrix given by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}B & a_{n2}B & \dots & a_{mn}B \end{bmatrix}.$$

Given a matrix $A \in \mathcal{S}^n$, with spectrum $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ and a matrix $B \in \mathcal{S}^m$, with spectrum $\sigma(B) = \{\mu_1, \dots, \mu_m\}$, it is known (see [8]) that the spectrum of $A \otimes B$ is given by:

$$\sigma(A \otimes B) = \{\lambda_i \mu_j \ : i = 1, \dots, n;\ j = 1, \dots, m\}.$$

Hence, letting $e_i \in \mathbb{R}^n$, we can write the vector $v \in \mathbb{R}^{nr}$ as

$$v = \sum_{i=1}^{n} (e_i \otimes v_i) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

so that we have

$$v_i = (e_i \otimes I_r)^T v.$$

Therefore we can write the objective function of Problem (2) as:

$$\begin{aligned} \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij} v_i^T v_j = & \sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij} \left((e_i \otimes I_r)^T v\right)^T (e_j \otimes I_r)^T v \\ = & \sum_{i=1}^{n}\sum_{j=1}^{n} v^T (q_{ij} e_i e_j^T \otimes I_r) v = v^T (Q \otimes I_r) v. \end{aligned}$$

With similar reasoning the constraints can be written as

$$\sum_{k=1}^{n}\sum_{j=1}^{n} (A_i)_{kj} v_k^T v_j = v^T (A_i \otimes I_r) v.$$

Thus, we obtain the nonlinear programming problem

$$\begin{aligned} \min \quad & v^T (Q \otimes I_r) v \ = \ q_r(v) \\ & v^T (A_i \otimes I_r) v = b_i \qquad i = 1, \dots, m, \end{aligned} \tag{NLP$_\mathrm{r}$}$$

which is the problem we will focus on in the rest of the paper.