1 SDP duality

Consider the following SDP in inequality form:

$$\begin{aligned} \min_{x \in \mathbb{R}} \ & x \\ \text{s.t.} \ & \begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix} \succeq 0. \end{aligned} \tag{1}$$

(a) Draw the feasible set. Is it convex?

Solution:
The symmetric matrix $\begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix}$ is positive semidefinite iff its diagonal entries and its determinant are nonnegative, i.e. $x \geq 0$, $y \geq 0$ and $xy \geq 1$. The region of (x, y) defined by these conditions is the epigraph of the convex function $y = \frac{1}{x}$ with domain $\mathbb{R}_{++}$. Being the epigraph of a convex function, this set is a convex subset of $\mathbb{R}^2$. See Figure 1.

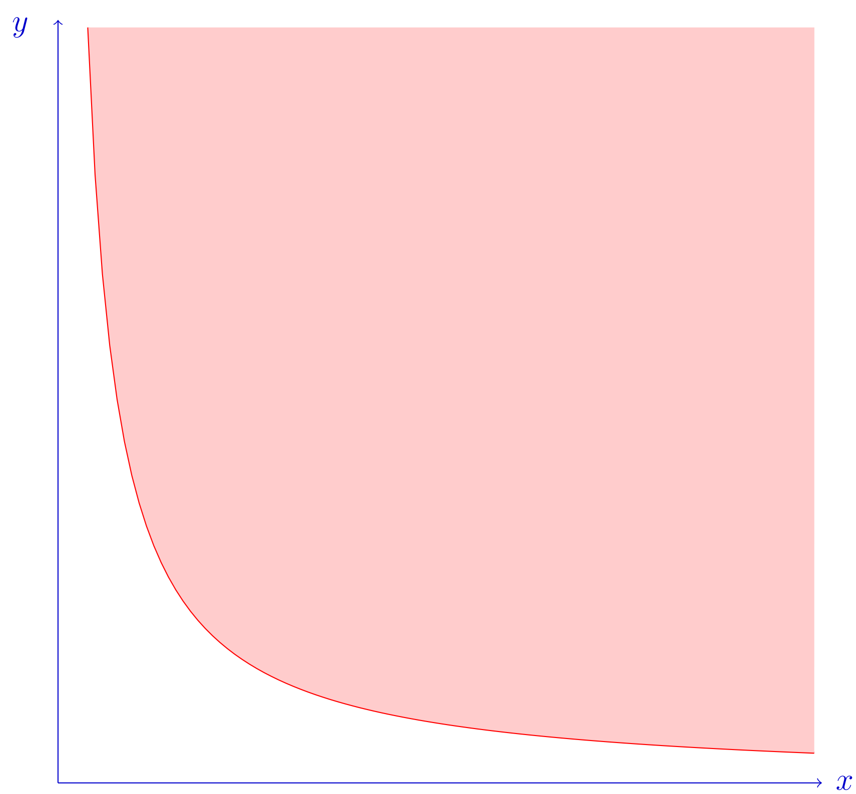

Figure 1: The domain of the primal SDP is the closed region above the curve $xy = 1$ in the nonnegative quadrant. This is a convex set.

(b) Write the conic dual SDP. This will be in standard form.

Remark: The textbook of Calafiore and El Ghaoui only describes how write a conic dual for a primal SDP in standard form (in which case the dual will be in inequality form) and the formula there is wrong because of a sign error. The duality formulas in the textbook of Boyd and Vandenberghe can be used (but note that their sign convention for linear matrix inequalities is different from ours). You could also use the formulas in the lecture slides in Lecture 21 and Lecture 22.

Solution:
If the primal SDP in inequality form is given by

$$\begin{aligned} \min_{x\in\mathbb{R}^m} \ & c^T x \\ \text{s.t. } & F_0 + \sum_{i=1}^m x_i F_i \succeq 0, \end{aligned}$$

where $F_0, F_1, \ldots, F_m \in \mathbb{S}^n$ and $c \in \mathbb{R}^n$, its dual is the SDP in standard form given by

$$\begin{aligned} \max_{X\in\mathbb{S}^n} \ & \text{trace}(-F_0 X) \\ \text{s.t. } & \text{trace}(F_i X) = c_i, \ i = 1, \ldots, n, \\ & X \succeq 0. \end{aligned}$$

Here we have $n = 2$, with the vectors in $\mathbb{R}^2$ written as $\begin{bmatrix} x & y \end{bmatrix}^T$, and

$$c = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ F_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ F_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ F_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The dual of the primal SDP in (1) is given by

$$\begin{aligned} \max_{x_{11}, x_{12}, x_{22}} \ & -2x_{12} \\ \text{s.t. } & x_{11} = 1, \\ & x_{22} = 0, \\ & \begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0. \end{aligned} \tag{2}$$

(c) Is the primal SDP feasible? Is it strictly feasible?

Remark: The SDP in inequality form

$$\begin{aligned} \min_{x\in\mathbb{R}^m} \ & c^T x \\ \text{s.t. } & F_0 + \sum_{i=1}^m x_i F_i \succeq 0, \end{aligned}$$

where $F_0, F_1, \ldots, F_m \in \mathbb{S}^n$, $c \in \mathbb{R}^n$, is said be strictly feasible if there is some $x \in \mathbb{R}^n$ such that $F(x) \in \mathbb{S}^n_{++}$, i.e. $F(x)$ is positive definite. Here $F(x)$ denotes $F_0 + \sum_{i=1}^m x_i F_i$.

Solution:
Since we can choose x and y such that $\begin{bmatrix} x & 1 \\ 1 & y \end{bmatrix}$ is positive definite, the primal SDP in (1) is strictly feasible (and hence also feasible).

(d) Is the dual SDP feasible? Is it strictly feasible?

Remark: The SDP in standard form

$$\begin{aligned} \min_{X \in \mathbb{S}^n} \ & \text{trace}(CX) \\ \text{s.t.} \ & \text{trace}(A_i X) = b_i, \ i = 1, \dots, m, \\ & X \succeq 0, \end{aligned}$$

where $C, A_1, \dots, A_m \in \mathbb{S}^m$, $b_1, \dots, b_m \in \mathbb{R}$, is said to be strictly feasible if there is some $X \in \mathbb{S}^n_{++}$ satisfying the equality constraints trace$(A_i X) = b_i$ for $i = 1, \dots, m$.

Solution:

From the dual SDP given in (2), we see that since we must have $x_{22} = 0$, the condition $\begin{bmatrix} x_{11} & x_{12} \\ x_{12} & x_{22} \end{bmatrix} \succeq 0$ forces $x_{12} = 0$. Since we also require that $x_{11} = 1$, there is only one matrix X in the feasible set, namely $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. This matrix is not positive definite. We conclude that the dual SDP is feasible but not strictly feasible.

(e) Find the optimal primal value p^* and the optimal dual value d^*. Does strong duality hold?

Solution:

From Figure 1 it is clear that the optimal primal value is $p^* = 0$ (which is not attained, i.e. there is no optimal point). To see this analytically, first note that we must have $x \geq 0$ at every feasible point. Further, for any $\epsilon > 0$, however small, the matrix $\begin{bmatrix} \epsilon & 1 \\ 1 & \epsilon^{-1} \end{bmatrix}$ is positive semidefinite, so the objective function of the primal problem can be made to equal ϵ.

As for the dual problem, it has only one feasible point and the value of the dual objective at that feasible point is 0, so we have $d^* = 0$.

Since $p^* = d^*$, strong duality holds.

2 Ellipsoid intersection

Consider the following QCQP with a single quadratic constraint

$$\begin{aligned} p^* := \min_{x \in \mathbb{R}^n} \ & x^T A_0 x + 2b_0^T x + c_0 \\ \text{s.t.} \ & x^T A_1 x + 2b_1^T x + c_1 \leq 0, \end{aligned} \tag{3}$$

where $A_0, A_1 \in \mathbb{S}^n_+, b_0, b_1 \in \mathbb{R}^n, c_0, c_1 \in \mathbb{R}$.

Note that this problem formulation allows us to determine whether a given ellipsoid in $\mathbb{R}^n$ intersects another ellipsoid. This is because, checking whether the ellipsoid

$$\{x \in \mathbb{R}^n : (x - x_1)^T A_1 (x - x_1) \leq d_1\}$$

intersects the ellipsoid

$$\{x \in \mathbb{R}^n : (x - x_0)^T A_0 (x - x_0) \leq d_0\}$$

is the same as checking whether the set of $x \in \mathbb{R}^n$ satisfying

$$x^T A_1 x - 2x_1^T A_1 x + x_1^T A_1 x_1 - d_1 \leq 0$$

intersects the set of $x \in \mathbb{R}^n$ satisfying

$$x^T A_0 x - 2x_0^T A_0 x + x_0^T A_0 x_0 - d_0 \leq 0.$$

If we could solve the optimization problem in (3) with

$$\begin{aligned} b_0 &:= -A_0^T x_0, \quad c_0 := x_0^T A_0 x_0 - d_0, \\ b_1 &:= -A_1^T x_1, \quad c_1 := x_1^T A_1 x_1 - d_1 \end{aligned}$$

then this intersection would be nonempty iff the value of this optimization problem is nonpositive.

(a) Write the Lagrangian of the primal problem in (3) and calculate the dual objective function.

Solution:

The Lagrangian of the primal problem can be written as

$$\mathcal{L}(x, \lambda) = x^T (A_0 + \lambda A_1) x + 2(b_0 + \lambda b_1)^T x + (c_0 + \lambda c_1).$$

The domain of this function is $\mathbb{R}^n \times \mathbb{R}$, but we are only interested in it for $\lambda \geq 0$.

The dual objective function can be written as

$$g(\lambda) := \inf_{x \in \mathbb{R}^n} \mathcal{L}(x, \lambda) = \begin{cases} -\infty, & \text{if} A_0 + \lambda A_1 \not\succeq 0 \\ -\infty, & \text{if } A_0 + \lambda A_1 \succeq 0 \text{ and } (b_0 + \lambda b_1) \notin \mathcal{R}(A_0 + \lambda A_1) \\ S, & \text{if } A_0 + \lambda A_1 \succeq 0 \text{ and } (b_0 + \lambda b_1) \in \mathcal{R}(A_0 + \lambda A_1) \end{cases}$$

where $S := (c_0 + \lambda c_1) - (b_0 + \lambda b_1)^T (A_0 + \lambda A_1)^\dagger (b_0 + \lambda b_1)$, with $(A_0 + \lambda A_1)^\dagger$ denoting the Moore-Penrose inverse of $A_0 + \lambda A_1$.

(b) Using the Schur complement rule and a slack variable to represent the objective function, show that the dual problem can be written as the following SDP in inequality form:

$$
\begin{aligned}
d^* := \max_{\lambda\in\mathbb{R},\gamma\in\mathbb{R}} \quad & \gamma \\
\text{s.t.} \quad & \begin{bmatrix} A_0+\lambda A_1 & b_0+\lambda b_1 \\ b_0^T+\lambda b_1^T & c_0+\lambda c_1-\gamma \end{bmatrix} \succeq 0, \\
& \lambda \geq 0.
\end{aligned} \tag{4}
$$

Remark: For this problem you will need to use to a generalized version of the Schur complement rule. A proof is provided in the the addendum of the discussion set, but you can take this for granted if you wish.

The traditional Schur complement rule, which we have proved in class, says that if we have a symmetric block matrix

$$M := \begin{bmatrix} A & X \\ X^T & B \end{bmatrix},$$

and if we assume that A is positive definite, then:

(i) M is positive semidefinite iff $B - X^T A^{-1} X$ is positive semidefinite;

(ii) M is positive definite iff $B - X^T A^{-1} X$ is positive definite.

Here we use the fact that A is nonsingular to be able to write A^{-1}.

The more general version of the Schur complement rule works even if A is only positive semidefinite but not necessarily positive definite. It says

(iii) M is positive semidefinite iff $B - X^T A^{\dagger} X$ is positive semidefinite and $\mathcal{R}(X) \subseteq \mathcal{R}(A)$.

Here $A^{\dagger}$ denotes the Moore-Penrose inverse of A.

Solution:

Introducing a slack variable to represent the dual objective function, the dual can be written as

$$
\begin{aligned}
d^* := \max_{\lambda\in\mathbb{R},\gamma\in\mathbb{R}} \quad & \gamma \\
\text{s.t.} \quad & (c_0+\lambda c_1) - (b_0+\lambda b_1)^\top (A_0+\lambda A_1)^{\dagger}(b_0+\lambda b_1) \geq \gamma, \\
& A_0+\lambda A_1 \succeq 0, \\
& \lambda \geq 0, \\
& (b_0+\lambda b_1) \in \mathcal{R}(A_0+\lambda A_1),
\end{aligned}
$$

which by the use of the Schur complement rule is the same as

$$
\begin{aligned}
d^* := \max_{\lambda\in\mathbb{R},\gamma\in\mathbb{R}} \quad & \gamma \\
\text{s.t.} \quad & \begin{bmatrix} A_0+\lambda A_1 & b_0+\lambda b_1 \\ b_0^T+\lambda b_1^T & c_0+\lambda c_1-\gamma \end{bmatrix} \succeq 0, \\
& \lambda \geq 0.
\end{aligned}
$$

(c) Since the dual is an SDP in inequality form, we can now take the dual of the dual (called the bidual), which will be an SDP in standard form. Show that, after some simplification, the bidual problem can be written as

$$\begin{aligned} \min_{x\in\mathbb{R}^n, X\in\mathbb{S}^n} \quad & \text{trace}(A_0X) + 2b_0^Tx + c_0 \\ \text{s.t.} \quad & \text{trace}(A_1X) + 2b_1^Tx + c_1 \le 0, \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0. \end{aligned}$$

Solution:

Since we are used to forming duals for optimization problems presented in minimization form we first rewrite the dual in minimization form as

$$\begin{aligned} -\min_{\lambda\in\mathbb{R},\gamma\in\mathbb{R}} \quad & -\gamma \\ \text{s.t.} \quad & \begin{bmatrix} A_0+\lambda A_1 & b_0+\lambda b_1 & 0_{n,1} \\ b_0^T+\lambda b_1^T & c_0+\lambda c_1-\gamma & 0 \\ 0_{1,n} & 0 & \lambda \end{bmatrix} \succeq 0, \end{aligned}$$

where we have combined the two LMIs into a single one.

This can be dualized using the pattern for dualizing an SDP in inequality form by setting

$$c = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \; F_0 = \begin{bmatrix} A_0 & b_0 & 0_{n,1} \\ b_0^T & c_0 & 0 \\ 0_{1,n} & 0 & 0 \end{bmatrix}, \; F_1 = \begin{bmatrix} A_1 & b_1 & 0_{n,1} \\ b_1^T & c_1 & 0 \\ 0_{1,n} & 0 & 1 \end{bmatrix}, \; F_2 = \begin{bmatrix} 0_{n,n} & 0_{n,1} & 0 \\ 0_{1,n} & -1 & 0 \\ 0_{1,n} & 0 & 0 \end{bmatrix},$$

where we have temporarily indicated the sizes of the zero matrices for clarity (but will now stop doing so).

We therefore get the bidual as the SDP in standard form (where we retain the overall negative sign) given by

$$\begin{aligned} -\max_{W\in\mathbb{S}^{n+2}} \quad & \text{trace}(-F_0W) \\ \text{s.t.} \quad & \text{trace}(F_1W) = 0, \\ & \text{trace}(F_1W) = -1, \\ & W \succeq 0. \end{aligned}$$

Let us write

$$W = \begin{bmatrix} X & x & v \\ x^T & y & s \\ v^T & s & t \end{bmatrix},$$

where $X \in \mathbb{S}^n$, $x, v \in \mathbb{R}^n$, $y, s, t \in \mathbb{R}$. The bidual can then be written as

$$\begin{aligned} -\max_{X,x,v,y,s,t} \quad & -\text{trace}(A_0X) - 2b_0^Tx - c_0y \\ \text{s.t.} \quad & \text{trace}(A_1X) + 2b_1^Tx + c_1y + t = 0, \\ & -y = -1, \\ & \begin{bmatrix} X & x & v \\ x^T & y & s \\ v^T & s & t \end{bmatrix} \succeq 0. \end{aligned}$$

The second equality constraint tell us that $y = 1$, so we can eliminate y and substitute 1 for it. The positive semidefiniteness condition tells us that we must have $t \geq 0$, so we can eliminate t and replace the first equality constraint by an inequality constraint. From the positive semidefiniteness condition, with $y = 1$, we notice that if $t \geq 0$ and $\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0$ then we can simply set $v \in \mathbb{R}^n$ and $s \in \mathbb{R}$ to be zero, because they do not show up in either the objective function or the other constraints, and we will still be satisfying the positive semidefiniteness condition when we do this. Putting these observations together and bringing the overall minus sign inside to write the bidual as a minimization problem, it becomes of the desired form

$$\begin{aligned} \min_{x \in \mathbb{R}^n, X \in \mathbb{S}^n} \quad & \text{trace}(A_0 X) + 2b_0^T x + c_0 \\ \text{s.t.} \quad & \text{trace}(A_1 X) + 2b_1^T x + c_1 \leq 0, \\ & \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0. \end{aligned}$$

(d) What is the relation of the initial primal problem with the bidual derived in the previous part of the question?

Solution:

Observe that the primal problem can be written as

$$\begin{aligned} \min_{x \in \mathbb{R}^n, X \in \mathbb{S}^n} \quad & \text{trace}(A_0 X) + 2b_0^T x + c_0 \\ \text{s.t.} \quad & \text{trace}(A_1 X) + 2b_1^T x + c_1 \leq 0, \\ & X = xx^T, \end{aligned}$$

while the bidual can be written as

$$\begin{aligned} \min_{x \in \mathbb{R}^n, X \in \mathbb{S}^n} \quad & \text{trace}(A_0 X) + 2b_0^T x + c_0 \\ \text{s.t.} \quad & \text{trace}(A_1 X) + 2b_1^T x + c_1 \leq 0, \\ & X \succeq xx^T, \end{aligned}$$

as can be seen from an application of the Schur complement rule.

Since A_0 and A_1 are positive semidefinite, these problems are equivalent. Every (x, X) in the feasibility set of the primal is also feasible for the bidual and the objective function of the two problems is identical, so the value of the bidual is no bigger than that of the primal. On the other hand, if (x, X) is feasible for the bidual, then (x, xx^T) will also be feasible for the bidual (because $A_1 \in \mathbb{S}_+^n$) and the value of the objective at (x, xx^T) will be no bigger than that at (x, X) (because $A_0 \in \mathbb{S}_+^n$). This means that the value of the bidual is the same as the value of the primal.

Addendum: (**Optional**)

We will prove the more general version of the Schur complement rule. Consider the symmetric block matrix

$$M := \begin{bmatrix} A & X \\ X^T & B \end{bmatrix},$$

and assume that $A \in \mathbb{S}^n$ is positive semidefinite. Then we can write

$$A = U_r \Sigma_r U_r^T,$$

where $U_r \in \mathbb{R}^{n \times r}$ has columns forming an orthonormal basis for $\mathcal{R}(A)$ and Σ_r is an invertible $r \times r$ diagonal matrix. Here $r = \text{rank}(A)$. For concreteness, let $B \in \mathbb{S}^m$.

Suppose first that $B - X^T A^\dagger X$ is not positive semidefinite. Then there is $v \in \mathbb{R}^m$ such that

$$v^T(B - X^T A^\dagger X)v < 0.$$

Given $w \in \mathbb{R}^n$, consider

$$\begin{aligned}\begin{bmatrix} w^T & v^T \end{bmatrix} M \begin{bmatrix} w \\ v \end{bmatrix} &= \begin{bmatrix} w^T & v^T \end{bmatrix} \begin{bmatrix} U_r \Sigma_r U_r^T & X \\ X^T & B \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} \\ &= w^T U_r \Sigma_r U_r^T w + 2w^T X v + v^T B v\end{aligned}$$

Now choose $w = -A^\dagger X v = -U_r \Sigma_r^{-1} U_r^T X v$. This gives

$$\begin{bmatrix} w^T & v^T \end{bmatrix} M \begin{bmatrix} w \\ v \end{bmatrix} = v^T(B - X^T A^\dagger X)v < 0,$$

which shows that M is not positive semidefinite.

Next suppose $\mathcal{R}(X) \subsetneq \mathcal{R}(A)$. Then there is $w \in \mathcal{N}(A)$ and $v \in \mathbb{R}^m$ such that $w^T X v < 0$. For $\alpha \in \mathbb{R}$ we write

$$\begin{aligned}\begin{bmatrix} w^T & \alpha v^T \end{bmatrix} M \begin{bmatrix} w \\ \alpha v \end{bmatrix} &= \begin{bmatrix} w^T & \alpha v^T \end{bmatrix} \begin{bmatrix} A & X \\ X^T & B \end{bmatrix} \begin{bmatrix} w \\ \alpha v \end{bmatrix} \\ &= w^T A w + 2\alpha w^T X v + \alpha^2 v^T B v \\ &= 2\alpha w^T X v + \alpha^2 v^T B v,\end{aligned}$$

where the last step is because $w^T A w = 0$, which comes from $w \in \mathcal{N}(A)$. For $\alpha > 0$ sufficiently small, this is negative, because $w^T X v < 0$. Hence M is not positive semidefinite.

We have shown that if M is positive semidefinite then we must have both $B - X^T A^\dagger X$ is positive semidefinite and $\mathcal{R}(X) \subseteq \mathcal{R}(A)$. We will now assume that we have both $B - X^T A^\dagger X$ is positive semidefinite and $\mathcal{R}(X) \subseteq \mathcal{R}(A)$ and show that this implies that M is positive semidefinite, which will complete the proof.

Note that $\mathcal{R}(X) \subseteq \mathcal{R}(A)$ implies that $X = U_r U_r^T X$. Hence we have

$$M = \begin{bmatrix} U_r \Sigma_r U_r^T & X \\ X^T & B \end{bmatrix} = \begin{bmatrix} U_r \Sigma_r U_r^T & U_r U_r^T X \\ X^T U_r U_r^T & B \end{bmatrix} = \begin{bmatrix} U_r & 0_{n,m} \\ 0_{m,r} & I_m \end{bmatrix} \begin{bmatrix} \Sigma_r & U_r^T X \\ X^T U_r & B \end{bmatrix} \begin{bmatrix} U_r^T & 0_{r,m} \\ 0_{m,n} & I_m \end{bmatrix}.$$

The traditional Schur complement rule tells us that since $B - X^T A^\dagger X$ is positive semidefinite we have that $\begin{bmatrix} \Sigma_r & U_r^T X \\ X^T U_r & B \end{bmatrix}$ is positive semidefinite (to see this, recall that $A^\dagger = U_r \Sigma_r^{-1} U_r^T$). But M is written above as a congruence transformation of this matrix. Hence M is positive semidefinite. This completes the proof.