Semidefinite Programming	**Lecture 12**
OR 6327 Spring 2012	**March 1, 2012**
Scribe: Dmitriy Drusvyatskiy	

Consider the primal-dual pair of semidefinite programs,

$$(P)\quad \begin{array}{rrcl} \min_X & C \bullet X & & \\ & \mathcal{A}X & = & b, \\ & X & \succeq & 0, \end{array} \qquad (D)\quad \begin{array}{rrcrcl} \max_{y,S} & b^T y & & & & \\ & \mathcal{A}^* y & + & S & = & C, \\ & & & S & \succeq & 0. \end{array}$$

We have seen that weak duality holds, meaning

$$C \bullet X \geq b^T y \text{ for all } X \text{ feasible for } (P) \text{ and all } (y, S) \text{ feasible for } (D).$$

Recall that the Lagrangian dual of

$$\min\{f(x) : g(x) = 0, x \in D\},$$

is the problem

$$\max_{y \in \mathbb{R}^m} (\min_{x \in D} \{f(x) + y^T g(x)\}).$$

Furthermore the objective value of the dual problem is always no greater than the objective value of the original problem.

Evidently (D) is the Lagrangian dual of (P) obtained by dualizing the linear constraints $\mathcal{A}X = b$. Indeed the Lagrangian dual of (P) is

$$\begin{aligned} \max_{y \in \mathbb{R}^m} (\min_{X \succeq 0} \{C \bullet X + y^T (b - \mathcal{A}X)\}) &= \max_{y \in \mathbb{R}^m} \{b^T y + \min_{X \succeq 0} (C - \mathcal{A}^* y) \bullet X\} \\ &= \max_{y \in \mathbb{R}^m} \{b^T y : C - \mathcal{A}^* y \succeq 0\}, \end{aligned}$$

where the last equality follows from the fact that the cone of positive semidefinite matrices is self-dual.

Exercise 1 Check that (P) is the Lagrangian dual of (D) if we dualize $\mathcal{A}^* y + S = C$.

The main question we consider now is: when does strong duality hold? We first investigate a few examples showing some complications that may arise in the semidefinite duality theory. (Note that there is a more complicated duality theory avoiding these exceptions. See the paper by Ramana, Tuncel, and Wolkowicz in the references part of the home page.)

Example 1 Consider the problem

$$(D)\qquad \begin{array}{l} \max -y_1, \\ \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} y_1 + \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} y_2 \preceq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \end{array}$$

Hence we would like to minimize y_1 subject to the constraint $\begin{bmatrix} y_1 & 1 \\ 1 & y_2 \end{bmatrix} \succeq 0$. This constraint in turn amounts to the requirements $y_1 \geq 0, y_2 \geq 0, y_1 y_2 \geq 1$. Here the optimal value is 0, but it is not attained.

Now consider the dual problem

$$(P) \qquad \begin{aligned} \min\ & x_{21} + x_{12}, \\ & x_{11} = 1, \\ & x_{22} = 0, \\ & \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \succeq 0. \end{aligned}$$

The only feasible (and hence optimal) solution of this problem is $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Consequently the optimal value is 0. Summarizing, we see that there is no duality gap for the pair (P) and (D), but the optimal value for (D) is not attained.

Example 2 Consider the problem

$$(P) \qquad \begin{aligned} \min\ & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} X, \\ & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} X = 1, \\ & \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} X = 0, \\ & X \succeq 0. \end{aligned}$$

The constraint $x_{22} = 0$ forces $x_{12} = x_{21} = 0$. Consequently $x_{33} = 1$ and the matrix $X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ is optimal with value 1.

Consider now the dual problem

$$(D) \qquad \begin{aligned} & \max\ y_1, \\ & \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} y_1 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} y_2 \preceq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Rewriting the constraint in (D), we have

$$\begin{bmatrix} 0 & -y_1 & 0 \\ -y_1 & -y_2 & 0 \\ 0 & 0 & 1 - y_1 \end{bmatrix} \succeq 0.$$

The top left entry being zero forces y_1 to be zero, and consequently $y = (0; 0)$ is optimal with value 0.

In summary, the optimal values of (P) and (D) are attained but there is a positive duality gap. Note that this property is fragile. If we change $b_2 = 0$ to any $\epsilon > 0$, the optimal value of (P) jumps to 0. If we change $c_{11} = 0$ to any $\epsilon > 0$, the optimal value of (D) jumps to 1.

Exercise 2 What happens if we make both of these changes simultaneously?

To avoid the problems illustrated by the two examples above, we add regularity conditions to the primal-dual pairs.

Definition 1 Given the primal problem (P), we define the sets of feasible and strictly feasible solutions as

$$\mathcal{F}(P) := \{X \in \mathbb{M}^n : \mathcal{A}X = b, X \succeq 0\}, \text{ and}$$
$$\mathcal{F}^o(P) := \{X \in \mathbb{M}^n : \mathcal{A}X = b, X \succ 0\},$$

respectively. Employing analogous notation for the dual problem (D), we have

$$\mathcal{F}(D) := \{(y, S) \in \mathbb{R}^n \times \mathbb{M}^n : \mathcal{A}^* y + S = C, S \succeq 0\}, \text{ and}$$

$$\mathcal{F}^o(D) := \{(y, S) \in \mathbb{R}^n \times \mathbb{M}^n : \mathcal{A}^* y + S = C, S \succ 0\}.$$

Theorem 1 *If $\mathcal{F}(P)$ and $\mathcal{F}^o(D)$ are nonempty, then*

1. *(P) has a nonempty compact set of optimal solutions; and*
2. *there is no duality gap.*

Proof: We first prove 1. Let $\hat{X} \in \mathcal{F}(P)$ and $(\hat{y}, \hat{S}) \in \mathcal{F}^o(D)$. Our goal is to apply the Weierstrass theorem, which states that a continuous function on a compact set attains its minimum. Here $\mathcal{F}(P)$ is nonempty and closed, but not necessarily bounded. Observe that (P) has exactly the same set of optimal solutions as

$$\min\left\{C \bullet X : \mathcal{A}X = b, C \bullet X \leq C \bullet \hat{X}, X \succeq 0\right\}.$$

Notice

$$C \bullet X - C \bullet \hat{X} = (C \bullet X - b^T \hat{y}) - (C \bullet \hat{X} - b^T \hat{y}) = \hat{S} \bullet X - \hat{S} \bullet X.$$

So the set of optimal solutions of (P) is contained in

$$\left\{X \in \mathbb{M}^n : \mathcal{A}X = b, \hat{S} \bullet X \leq \hat{S} \bullet \hat{X}, X \succeq 0\right\}.$$

Since we have $\hat{S} \succ 0$, the set above is bounded. We conclude that the set of optimal solutions to (P) is nonempty by the Weierstrass theorem, and is compact since it is a closed subset of the compact set above.

Before proving 2, we need to recall the separating hyperplane theorem, which we do now.

Theorem 2 *If $C \subset \mathbb{R}^n$ is a closed, convex set and $\bar{x} \notin C$, then there exist a nonzero vector $s \in \mathbb{R}^n$ and a real $\eta \in \mathbb{R}$ with*

$$s^T \bar{x} < \eta < \inf\{s^T x : x \in C\}.$$

Definition 2 A nonzero vector $z \in \mathbb{R}^n$ is a *direction of recession* for a closed, convex set $C \subset \mathbb{R}^n$ if for all $x \in C$ the inclusion $\{x + \lambda z : \lambda \geq 0\} \subset C$ holds.

Theorem 3 *If $C, D \subset \mathbb{R}^n$ are closed, convex sets with no common direction of recession and $C \cap D = \emptyset$, then there exist a nonzero vector $s \in \mathbb{R}^n$ and a real $\eta \in \mathbb{R}$ with*

$$\sup\{s^T x : x \in D\} < \eta < \inf\{s^T x : x \in C\}.$$

Proof: Consider the set $C - D$. Clearly this set is convex. Since C and D have no common direction of recession, one can show that $C - D$ is also closed. Finally, $0 \notin C - D$. An application of Theorem 2 yields the result. $\square$

Proof of part 2 of Theorem 1:

Let ζ be the optimal value of (P) and let $\epsilon > 0$. We want to prove that there exists a feasible dual solution with objective value no smaller than $\zeta - \epsilon$. To this end define the two sets

$$\begin{aligned} \mathcal{C} &= \mathbb{M}^n_+, \\ \mathcal{D} &= \{X \in \mathbb{M}^n : \mathcal{A}X = b, C \bullet X \leq \zeta - \epsilon\}. \end{aligned}$$

Clearly $\mathcal{C}$ and $\mathcal{D}$ are disjoint, closed, convex sets in $\mathbb{M}^n$. We claim that $\mathcal{C}$ and $\mathcal{D}$ have no common direction of recession. Indeed, suppose Z were a common direction of recession. Then we have a nonzero Z with $Z \succeq 0$, $\mathcal{A}Z = 0$, and $C \bullet Z \leq 0$, thus contradicting 1. So applying Theorem 3, we deduce that there exists a nonzero matrix $S \in \mathbb{M}^n$ and a real $\eta \in \mathbb{R}$ with

$$\sup\{S \bullet X : \mathcal{A}X = b, C \bullet X \leq \zeta - \epsilon\} < \eta < \inf\{S \bullet X : X \succeq 0\}.$$

From the right-hand side of the inequality, we deduce $S \succeq 0$ and $\eta < 0$. Hence the implication

$$-\mathcal{A}X = -b,\ C \bullet X \leq \zeta - \epsilon \Rightarrow S \bullet X < \eta < 0,$$

holds.

We deduce that the implied inequality is an appropriate linear combination of the left-hand constraints. So there exists a vector $y \in \mathbb{R}^m$ and a real $\lambda \geq 0$ with

$$-\mathcal{A}^* y + \lambda C = S,\ -b^T y + \lambda(\zeta - \epsilon) \leq \eta.$$

Suppose for the sake of contradiction $\lambda = 0$. Then we have

$$\mathcal{A}^* y + S = 0, \qquad b^T y \geq -\eta > 0.$$

So $0 = \hat{X} \bullet (\mathcal{A}^* y + S) = y^T \mathcal{A}\hat{X} + \hat{X} \bullet S \geq b^T y > 0$, which is a contradiction. We conclude that $\lambda > 0$, and by scaling y, S, λ, and η, we can assume that $\lambda = 1$. Consequently we obtain

$$\begin{aligned} \mathcal{A}^* y + S &= C, \\ S &\succeq 0, \end{aligned}$$

and $b^T y \geq \zeta - \epsilon - \eta > \zeta - \epsilon$. Hence the value of the dual is no smaller than $\zeta - \epsilon$. $\square$