Lecture 9: Duality in Semidefinite Programs

Rajat Mittal

IIT Kanpur

We have already seen the strong connections of dual and primal program in case of linear programming. Similar connections exist in the world of semidefinite programming.

In this lecture note, we will introduce what are dual cones, learn how to take dual of a general cone program and apply this knowledge to semidefinite programming. Please note that *strong duality* does not need to always hold in the case of a semidefinite program as was the case in linear programs.

To extend the theory of duality for semidefinite program, we need the concept of *dual cone*.

1 Dual cone

Exercise 1. What is the definition of a cone, convex cone, proper cone and a finitely generated cone?

Given a cone C in $\mathbb{R}^n$, we can define the dual of cone C by

$$C^* = \{y \in \mathbb{R}^n : y^T x \geq 0, \ \forall x \in C\}$$

Here $y^T x$ is the inner product on $\mathbb{R}^n$. In general, the same definition can be used in other inner-product spaces using the inner product in that space.

Suppose $y \in C^*$, then the entire cone C lies on one side of the hyperplane which is normal to y. The cone C contains origin (0). In other words, the hyperplane with normal y (passing through origin) is a supporting hyperplane at 0.

Given two hyperplanes which are supporting at 0, their convex combination (convex combination of normals) will also give a supporting hyperplane. Geometrically, the convex combination will lie *in between* the original two hyperplanes. Hence,

$$y_1, y_2 \in C^* \rightarrow \theta y_1 + (1-\theta) y_2 \in C^*, \ \forall \theta \in \{0, 1\}$$

It is again easy to see that

$$y \in C^*, \ \theta \geq 0 \Rightarrow \theta y \in C^*.$$

So the dual cone is a convex cone irrespective of whether the original cone was convex or not. Notice from the definition, the dual cone is always closed too (all limits exist in the dual cone). Actually, the starting set C need not be a cone. For any set C, we can define the dual and it will be a convex cone.

Exercise 2. Can you describe, what will be the dual cone of a set C?

Note 1. Polar cone is just the negative of a dual cone,

$$C' = \{y : y^T x \leq 0, \ \forall x \in C\}.$$

We will mostly be dealing with cones which are closed, convex, pointed (contain no line) and have non-empty interior. Such cones are called *proper cones*. Using a proper cone S, we can define corresponding *generalized inequality* as a partial order.

$$x \geq y \Leftrightarrow x - y \in S$$

Exercise 3. Show that the positive orthant is a proper cone. What is the generalized inequality with respect to that cone?

1.1 Examples

Let's take a look at few examples and find the dual of some cones.

- Subspace: First, we can show that a linear subspace L (all linear combinations are present in the set) is a cone. We leave it as an easy exercise.
  A vector y is in L^* iff $y^T x \geq 0, \ \forall x \in L$. But in a subspace, if x is present then so is $-x$.
  So,
  $$y^T x = 0, \ \forall x \in L.$$
  Hence, the dual cone to the subspace is all vectors normal to it.
- Positive orthant ($\mathbb{R}^n_+$): We discussed that this set is a cone before and the usual $\geq$ inequality is the generalized inequality with respect to this cone.

  Exercise 4. Show that the positive orthant is the dual cone of itself

  Such cones are called *self dual cones*.
- Set of all positive semidefinite matrices are called the *positive semidefinite cone*. We will show later that it is also a self dual cone.

  Exercise 5. Show that if $C \subseteq D$, then $D^* \subseteq C^*$.

- Another interesting case of a dual cone is the case of a finitely generated cone. Suppose we are given a cone $C = cone(x_1, \cdots, x_k)$. How will its dual cone look?
  $$C^* = \{y : y^T x \geq 0, \ \forall x \in C\}$$
  Clearly, for all the generators of C $(x_1, \cdots, x_k)$, $y^T x_i \geq 0$. Since any element of the cone is a conic combination of these vectors, this condition is actually sufficient too. So,
  $$C^* = \{y : y^T x_i \geq 0, \ i = 1, \cdots, k\}.$$
  Notice that this gives the definition of cone in terms of inequalities. Affine Weyl theorem tells us that these two are equivalent formulations. So, C^* is also a finitely generated cone.
  It is important to note that the generators of the original cone are the inequalities for the dual cone and vice versa. So in the case of finitely generated cones, it is easy to specify the dual cone.

  Exercise 6. What will be the dual cone in $\mathbb{R}^2$ generated by two vectors.

1.2 Dual of the dual cone

The natural question is, what is the dual cone of C^* for a closed convex cone C?

Suppose $x \in C$ and $y \in C^*$, then we know $y^T x \geq 0$. Since this condition is symmetric, we get that every element of C will be in C^{**}, i.e., $C \subseteq C^{**}$. We will show that these two sets are indeed equal.

For the other direction, we need to show that any element x not in C has negative inner product with at least one $y \in C^*$.

Suppose $x \notin C$, then by Farkas lemma, there exists a hyperplane which separates the point x and the cone C. So there exists y, s.t., $y^T x < 0 \leq y^T z, \ \forall z \in C$.

Then $y \in C^*$ by definition. We know, $y^T x < 0$, so $x \notin C^{**}$. Hence, for a closed convex cone C,

$$C = C^{**}.$$

Exercise 7. Dual cone can be defined for arbitrary sets also. What will be the dual of the dual in this case?

As discussed last time, we will mostly be interested in proper cones. Remember that they are convex, closed cones which do not contain any line and have nonempty interior. We will give an informal argument that the dual of a proper cone is a proper cone.

Since C has nonempty interior, there exists a ball of non-zero radius inside the cone. Suppose the dual cone contains a line $\theta y,\ \theta \in \mathbb{R}$. Then, for every point $x \in C$, $y^T x = 0$ (why).

It implies that $y^T x = 0$ for the entire ball contained in the cone too. Since the inner product of y with center of the ball is zero and in every direction we move the inner product is zero. This implies the inner product with entire space is zero, and hence $y = 0$. So C^* does not contain a line.

Similarly, suppose C has empty interior. Since it is closed and convex, it should be contained in an affine space of lower dimension. Then the normal to that affine space will be contained in C^*. Actually, any scalar multiple of normal will be contained in C^*. Hence, C^* contains a line.

From the previous two paragraphs the cone C has non-empty interior iff C^* does not contain a line. Now, suppose C^* is empty. It implies $C^{**} = C$ contains a line. The contrapositive is that if C does not contain a line then C^* is non-empty.

Hence, it is clear that the dual of a proper cone is proper.

1.3 $\mathcal{S}_n$ is a self dual cone

The inner product in the space of matrices is the $\bullet$ operation between matrices.

$$A \bullet B = \sum_{i,j} A_{ij} B_{ij} = tr(A^T B)$$

The dual cone of $\mathcal{S}_n$ is the cone $\mathcal{S}'_n$, s.t.,

$$\mathcal{S}'_n = \{M : M \bullet N \geq 0,\ \forall N \in \mathcal{S}_n\}$$

Consider the Hadamard product of two positive semidefinite matrices. It was proved before that $M \circ N \succeq 0$ (using the gram matrix definition of positive semidefiniteness).

Exercise 8. If $M \circ N \succeq 0$, prove that $M \bullet N \geq 0$.

Hence every positive semidefinite matrix is part of the dual cone $\mathcal{S}'_n$. It implies

$$\mathcal{S}_n \subseteq \mathcal{S}'_n. \tag{1}$$

Now consider a symmetric matrix $M \notin \mathcal{S}_n$. There exist at least one negative eigenvalue λ and an eigenvector v corresponding to it. So,

$$0 > \lambda = v^T M v = M \bullet (vv^T).$$

This implies that $M \notin \mathcal{S}'_n$. Hence,

$$\mathcal{S}'_n \subseteq \mathcal{S}_n. \tag{2}$$

From Eqn. 1 and Eqn. 2, we get $\mathcal{S}'_n = \mathcal{S}_n$. The cone of positive semidefinite cone is a self dual cone.

2 Dual for a cone program

In this section we will see how the dual of a general cone program can be obtained. Suppose there is an optimization program,

$$\begin{aligned} & \max c^T x \\ \text{s.t.} \qquad & Ax \leq b \\ x \in S. \quad & (S \text{ is some cone}) \end{aligned}$$

For this case, say we take the linear combination y of the rows of matrix A. Then, we want $y^T b$ to be an upper bound on the value of $c^T x$ for any feasible x. One possible way is to satisfy,

$$c^T x \leq y^T Ax \leq y^T b.$$

For the second inequality, y should be positive and for the first one $y^T A - c \in S^*$. Here, S^* is the dual cone of S. So, the dual program looks like,

$$\begin{aligned} &\min y^T b \\ \text{s.t. } \ & y^T A - c \in S^* \\ & y \geq 0, \end{aligned}$$

Let's look at the two programs slightly differently.

$$\begin{aligned} &\max c^T x \\ \text{s.t. } \ & Ax - b \leq 0 \Leftrightarrow Ax - b \in C \\ & x \in S. \quad (S, C \text{ are cones}) \end{aligned}$$

$$\begin{aligned} &\min y^T b \\ \text{s.t. } \ & y^T A - c \in S^* \\ & -y \in C^*, \end{aligned}$$

Cone C is the cone of all negative vectors, $\mathbb{R}^n_-$.

If the given program is a maximization problem, then if the primal variable is in cone S then the dual constraint is the membership in S^*. If the primal constraint is membership in cone C then the negative of dual variable is in cone C^*. For the minimization problem the relationship is opposite.

	Primal variable in S	Primal constraint in C
Max	Dual constraint in S^*	Negative of dual variable in C^*
Min	Negative of Dual constraint in S^*	Dual variable in C^*

Exercise 9. Check and verify the table.

3 Dual of a semidefinite program

Let's look at the primal-dual pair for semidefinite programming.

Primal

$$\begin{aligned} &\max C \bullet X \\ \text{s.t. } \ & A_i \bullet X = b_i \quad \forall i \\ & X \in \mathcal{S}_n \end{aligned}$$

Dual

$$\begin{aligned} &\min y^T b \\ \text{s.t. } \ & \sum_i y_i A_i - C \in \mathcal{S}_n \end{aligned}$$

The dual variable y is unconstrained because there is equality in the primal constraint. The dual program is in the second standard form previously discussed in the class. You will verify in the assignment that dual of the dual is the primal SDP.

Exercise 10. Remember the SDP for finding the maximum eigenvalue of a matrix M. Take its dual and show that it computes the maximum eigenvalue (without using duality).

From the formulation of a dual program, it turns out that the dual program gave an upper/lower bound depending on whether the problem was maximization/minimization respectively.

Are these bounds tight? It turns out that these bounds are tight in most cases of interest like linear programming. The bounds are not tight in a few contrived cases.

When the primal and dual values agree, *strong duality* is said to hold. We will later look at the conditions under which strong duality holds.

First, it is instructive to see the direct proof of weak duality. Given any feasible solution X for primal and y for dual,

$$C \bullet X \leq (\sum_i y_i A_i) \bullet X = \sum_i y_i b_i = y^T b. \tag{3}$$

Notice that the above equation implies, all feasible solutions of dual give an upper bound on the optimal value of primal. Similarly any feasible solution for primal gives a lower bound on the optimal value of dual.

Suppose the optimal value of primal is p^*, attained at X^*. Similarly the optimal value of dual is d^* and obtained for y^*. Weak duality implies that $p^* \leq d^*$. Assume that this two values are equal, i.e., $p^* = d^*$ (strong duality). Then from Eqn. 3,

$$\begin{aligned} C \bullet X &= (\textstyle\sum_i y_i A_i) \bullet X \\ \Rightarrow (\sum_i y_i A_i - C) \bullet X &= 0 \end{aligned}$$

The condition above is called the *complementary slackness* condition. So, for optimal X^*, y^* with strong duality, the complementary slackness condition holds. Conversely, if X, y are feasible solutions of primal and dual respectively and satisfy the complementary slackness condition then strong duality holds and $p^* = d^*$.

Exercise 11. What does the complementary slackness condition tell us in case of linear programming?

The discussion above is done for semidefinite program, but with little more effort can be generalized to cone programs.

4 Extra reading: Strong duality, Slater's condition

It turns out that in most of the applications of semidefinite programming to real world, strong duality holds. Hence the optimal value of primal is same as optimal value of dual. Strong duality can be obtained by verifying *Slater condition*. Specifically, if the semidefinite program satisfies *Slater conditions* then it has strong duality.

Theorem 1. *Given a semidefinite program in standard form with parameters C, A_i, b, suppose the feasible set of primal is $\mathcal{P}$ and feasible set of dual is $\mathcal{D}$. Then strong duality holds if either*

- *If $\mathcal{D} \neq \emptyset$ and there exists a strictly feasible $X \in \mathcal{P}$, i.e., $X \succ 0, A_i \bullet X = b_i \ \ \forall i$.*
  or
- *If $\mathcal{P} \neq \emptyset$ and there exists a strictly feasible $y \in \mathcal{D}$, i.e., $\sum_i y_i A_i - C \succ 0$.*

In other words, if the primal is feasible and dual is strictly feasible or vice versa then strong duality holds. In the usual cases, the feasible region would be expected to be non-empty and even have non-empty interior. That would imply strong duality from Slater's condition.

Proof. We will prove the first part. Since the dual of the dual is primal, the second part will follow.

Consider the set,

$$M = \{Z, u, v : \exists X, \text{ s.t.}, X \succeq Z, \forall i \ \ u_i = A_i \bullet X - b_i, v \leq C \bullet X\}$$

This can be thought of as the hypograph (the point below the graph) for the function

$$X, (\forall i \ \ A_i \bullet X - b_i), C \bullet X,$$

on the first and last parts. Here, Z is the symmetric matrix, u is a vector and v is a scalar. Since the function is concave (linear), so the hypograph is a convex set. Consider another convex set,

$$N = \{(0, 0, t) : t > p^*\}$$

Here p^* is the optimal value of primal (it is finite because $\mathcal{D} \neq \varnothing$). Hence, set M and N do not intersect (why?). By separating hyperplane theorem, there exist $\lambda_1, \lambda_2, \lambda_3, \alpha$, s.t.,

$$\lambda_1 \bullet Z + \lambda_2^T u + \lambda_3 v \le \alpha \quad \forall (Z, u, v) \in M \tag{4}$$

$$\lambda_1 \bullet 0 + \lambda_2^T 0 + \lambda_3 t \ \ge \alpha \quad \forall (0, 0, t) \in N. \tag{5}$$

The left hand side of Eqn. 4 can only be upper bounded if $\lambda_1 \succeq 0$ and $\lambda_3 \ge 0$. Eqn. 5 implies $\lambda_3 p^* \ge \alpha$. Assume $\lambda_3 > 0$, we will show this later by using the strict feasibility. Then,

$$\lambda_1' \bullet Z + \lambda_2'^T u + v \le p^* \quad \forall (Z, u, v) \in M. \tag{6}$$

Where $\lambda_1' = \frac{\lambda_1}{\lambda_3}$ and $\lambda_2' = \frac{\lambda_2}{\lambda_3}$. Using the definition of M, we get,

$$\lambda_1' \bullet X + \sum_i \lambda_{2,i}' (A_i \bullet X - b_i) + C \bullet X \qquad \le p^* \quad \forall X. \tag{7}$$

$$(\lambda_1' + \sum_i \lambda_{2,i}' A_i + C) \bullet X \le p^* + b^T \lambda_2 \quad \forall X \tag{8}$$

Here $\lambda_{2,i}'$ is the i^{th} entry of λ_2'. Eqn. 8 implies $p^* + b^T \lambda_2 \ge 0$ and $\lambda_1' + \sum_i \lambda_{2,i}' A_i + C = 0$.

Exercise 12. Prove the last statement.

Substituting $\lambda_2' = -\lambda$, we get (remember that $\lambda_1' \succeq 0$),

$$p^* \ge b^T \lambda \text{ and } \sum_i \lambda_i A_i - C \in \mathcal{S}_n$$

So, we get a feasible dual solution λ. Since dual is a minimization problem, from weak duality $p^* \le d^*$. This implies,

$$p^* \ge b^T \lambda \ge d^* \Rightarrow p^* = d^*$$

Hence, strong duality holds. The only part left to prove is: $\lambda_3 \neq 0$ using strict feasibility (notice that we have not used strict feasibility is till now).

Suppose $\lambda_3 = 0$, then Eqn. 4 and Eqn. 5 implies

$$\lambda_1 \bullet X + \sum_i \lambda_{2,i} (A_i \bullet X - b_i) \le \alpha \le \lambda_3 p^* = 0 \quad \forall X.$$

Take a strictly feasible X, then $A_i \bullet X - b_i = 0$. So $\lambda_1 \bullet X \le 0 \Rightarrow \lambda_1 = 0$, since $\lambda_1 \succeq 0$. Then,

$$\sum_i \lambda_{2,i} (A_i \bullet X - b_i) \le \alpha \le \lambda_3 p^* = 0 \quad \forall X.$$

For simplicity, Let's assume that A_i's span the whole space. For a strictly feasible X, left hand side is zero. Then, if we move in some direction $X + \epsilon Y$, the value will be non-zero (otherwise $\lambda_2 = 0$ and the separating hyperplane is trivial).

So, in either that direction or the negative one, the left hand side will be positive and will violate inequality. Hence λ_3 cannot be zero.

□

4.1 Counterexample for strong duality

From the last section we infer that strong duality holds for most of the cases in semidefinite programming. But still there exist cases when strong duality does not hold. Below, we give an example.

Consider the semidefinite programs for 3×3 matrices, Primal

$$\begin{aligned} \max \quad & -x_{11} - x_{22} \\ \text{s.t.} \quad & x_{11} = 0, 2x_{13} + x_{22} = 1 \\ & X \succeq 0 \end{aligned}$$

Dual

$$\begin{aligned} & \min y_2 \\ \text{s.t.} \quad & \begin{pmatrix} y_1 + 1 & 0 & y_2 \\ 0 & y_2 + 1 & 0 \\ y_2 & 0 & 0 \end{pmatrix} \succeq 0 \end{aligned}$$

For the primal problem, the first digonal entry forces x_{13} to be zero. Hence the optimal value is -1. For the dual problem, the last diagonal entry forces $y_2 = 0$. So the optimal value is 0. Hence, strong duality does not hold.

Exercise 13. Show that Slater's condition does not hold in this case.

Notice that the first constraint in primal is a weird way to say that first row and column are zero. Show that if we remove first row and column from the primal problem, then strong duality holds.

5 Assignment

Exercise 14. Given two matrices $M, N \succeq 0$, s.t., $M \bullet N = 0$. Then, show that the eigenvectors of M, N corresponding to non-zero eigenvalues are orthogonal to each other.

Exercise 15. Show that the programs given in the *counterexample to strong duality* section are dual of each other.

Exercise 16. Convert the dual of the standard SDP into the standard form, take dual and show that you get the original SDP.

Exercise 17. What is the dual of the following program?

$$\begin{aligned} & \max \textstyle\sum_{(i,j) \in E} \frac{1 - y_i^T y_j}{2} \\ \text{s.t.} \quad & \|y_i\| = 1 \quad \forall i \end{aligned} \tag{9}$$