Semidefinite Programming
Lecture 3

scribe: ÜNSAL ÖZDOĞRU

February 13 1995

1 Rules for taking duals

Like in Linear Programming, we have the following primal-constraint dual-variable relationships in cones.

	MIN		MAX	
				C
V	$\geq_{\mathcal{K}} 0$	$\longleftrightarrow$	$\geq_{\mathcal{K}^*}$	O
A	$\leq_{\mathcal{K}} 0$	$\longleftrightarrow$	$\leq_{\mathcal{K}^*}$	N
R	unrestricted	$\longleftrightarrow$	$=$	S
				T
C				
O	$\geq_{\mathcal{K}}$	$\longleftrightarrow$	$\geq_{\mathcal{K}^*} 0$	V
N	$\leq_{\mathcal{K}}$	$\longleftrightarrow$	$\geq_{\mathcal{K}^*} 0$	A
S	$=$	$\longleftrightarrow$	unrestricted	R
T				

Figure 1: Rules for taking dual of a mixed linear program over cone $\mathcal{K}$. Variables in one program gives rise to constraints in another and vice-versa.

Let's try to understand this for the Lovàsz θ function

$$\theta(G) = \min\{\lambda_1(X + J) : X_{ij} = 0, \quad \forall\{i,j\} \in E \text{ or } i = j\}$$

λ_1 corresponds to largest eigenvalue. This problem can be written as a semidefinite program

$$\begin{array}{rll} \textbf{Primal :} & \min & z \\ & \text{st} & zI - X \succeq J \\ & & E_{ij} \bullet X = 0 \ \ \forall i,j \in E \text{ or } i = j \end{array}$$

Dual of this problem

$$\begin{array}{rl} \max & J \bullet Y \\ \text{st} & I \bullet Y = 1 \\ -Y + \sum_{i=j}(i,j) \in E u_{ij} E_{ij} = 0 & \\ & Y \succeq 0 \end{array}$$

This maximization problem is equivalent to

$$\begin{array}{lll} \max & \sum Y_{ij} & \\ \text{st} & \text{trace } Y = 1 & \\ & Y_{ij} = 0 & \forall i,j \notin E \\ & Y \succeq 0 & \end{array}$$

2 Complementary Slackness Theorem for Cones

In an SDP problem, if strong duality holds then duality gap is zero, that is $X \bullet S = 0$. However, in the presence of $X \succeq 0$ and $S \succeq 0$, one gets a stronger result:

THEOREM 2.1. *Let $X \succeq 0$ and $S \succeq 0$. Then $X \bullet S = 0$ iff $XS = 0$.*

Proof. Obviously if $XS = 0$ then trace $XS = X \bullet S = 0$. To see the converse note that

$$0 = X \bullet S = \text{trace } XS = \text{trace } X^{1/2} S X^{1/2}$$

But $X^{1/2} S X^{1/2} \succeq 0$ and its trace is zero, thus all of its eigenvalues are zero and so

$$(X^{1/2} S^{1/2})^T (X^{1/2} S^{1/2}) = X^{1/2} S X^{1/2} = 0$$

and so $X^{1/2} S^{1/2} = 0$ and therefore $XS = 0$. ■

Let us just mention a couple of other complementary slackness theorems for some familiar cones. In general at the optimum the pair of primal and dual vectors $\mathbf{x}$ and $\mathbf{z}$ the three conditions

$$\begin{aligned} \mathbf{x} &\geq_{\mathcal{K}} 0 \\ \mathbf{z} &\geq_{\mathcal{K}^*} \\ \mathbf{x}^T \mathbf{z} &= 0 \end{aligned}$$

induces an n-dimensional variety; in other words, the condition $\mathbf{x}^T\mathbf{z} = 0$, in the presence of inequalities $\mathbf{x} \geq_K 0$ and $\mathbf{z} \geq_{K^*}$ is equivalent to n equations $f_i(\mathbf{x}, \mathbf{z}) = 0, i = 1 \ldots n$. For example in case of semidefinite programming we have $X \succeq 0$ and $Z \succeq 0$, then $X \bullet Z = 0$ if and only if $XZ = 0$. Thus, $X \bullet Z = 0$ is equivalent to $\frac{n(n+1)}{2}$ equations. let us now study a couple of other cones, and their complementarity relations.

2.1 Ice Cream Cone:

$$Q := \{\mathbf{x} : x_0 \geq \sqrt{x_1^2 + \ldots + x_n^2} = ||\mathbf{x}||_2\}$$

Exercise 1. Prove that Q is self-polar.

Notice that the vector $\mathbf{x}$ is indexed from zero, and from definition we must have: $x_0 \geq 0$. Write $\mathbf{x} \geq_Q 0$ whenever $\mathbf{x} \in Q$. Now in order to find out the complementary slackness theorem for the linear programming for the cone Q, we need to first find out what do the three conditions

$$\mathbf{x} \geq_Q 0$$
$$\mathbf{z} \geq_Q 0$$
$$\mathbf{x}^T\mathbf{z} = 0$$

imply. Both $\mathbf{x}$ and $\mathbf{z}$ are in cone Q and they are orthogonal. Upon a moment's reflection we see that the projection of $\mathbf{x}$ and the projection of $\mathbf{z}$ on the plane $x_0 = 0$ must be in opposite directions, see figure ??. In other words:

LEMMA 2.1. *If* $\mathbf{x} \geq_Q 0$ *and* $\mathbf{z} \geq_Q 0$, *then* $\mathbf{x}^T\mathbf{z}$ *iff*

$$\frac{-x_0}{z_0} = \frac{x_1}{z_1} = \ldots = \frac{x_n}{z_n} = -r$$

Proof.

$$(*) \qquad \begin{array}{l} x_0^2 \geq (x_1^2 + \ldots + x_n^2) \\ z_0^2 \geq (z_1^2 + \ldots + z_n^2) \end{array} \qquad \longleftrightarrow \qquad x_0^2 \geq (\tfrac{x_0 z_1}{z_0})^2 + \ldots + (\tfrac{x_0 z_n}{z_0})^2$$

And we also know that $x_0 z_0 = -x_1 z_1 - \cdots - x_n z_n$. If we multiply both sides with $2x_0^2/z_0^2$.

$$(**) \qquad 2x_0^2 = -2\frac{x_0 z_1}{z_0}x_1 - \ldots - 2\frac{2x_0 z_1}{z_0}x_1$$

after adding both sides of (*) subtracting (***) we get

$$0 \geq (x_1 + x_0 z_1/z_0)^2 + \ldots + (x_n + x_0 z_n/z_0)^2$$

which implies the result. ■

Exercise 2. Optimization over the ice cream cone can be done by solving a system of equation. In other words the optimization problem

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{st} & A\mathbf{x} = \mathbf{b} \\ & x \geq_Q 0 \end{array}$$

can be solved analytically. Show how this can be accomplished.
Hint: The intersection of an affine space and the ice cream cone is an ellipsoid in the dimension of the affine space. Optimization over an ellipsoid is easy.

A less trivial problem is:

$$\begin{array}{lll} \min & \mathbf{c}_1\mathbf{x}_1 + \cdots + \mathbf{c}_k\mathbf{x}_k & \\ \text{st} & A_1\mathbf{x}_1 + \cdots + A_k\mathbf{x}_k = \mathbf{b} & \\ & \mathbf{x}_k \geq_Q \mathbf{0} & \text{for } i = 1, \cdots, k \end{array} \qquad \begin{array}{lll} \max & \mathbf{b}^T\mathbf{y} & \\ \text{st} & A_k^T\mathbf{y} + \mathbf{z}_k = \mathbf{c} & \text{for } i = 1, \cdots, k \\ & \mathbf{z}_k \geq_Q \mathbf{0} & \text{for } i = 1, \cdots, k \end{array} \tag{2.1}$$

the primal feasibility, dual feasibility and complementary slackness together are equivalent to the following system:

$$\begin{array}{ll} A_1\mathbf{x}_1+\cdots+A_k\mathbf{x}_k=\mathbf{b} & \\ A_i^T\mathbf{y}+\mathbf{z}_i=\mathbf{c}_i & \text{for } i=1,\cdots,k \\ (\mathbf{x}_i)^T\mathbf{z}_i=0 & \text{for } i=1,\cdots,k \\ (\mathbf{x}_i)_0(\mathbf{z}_i)_j+(\mathbf{x}_i)_j(\mathbf{z}_i)_0=0 & \text{for } j=1,\cdots,n_j, \ \text{for } i=1,\cdots,k \end{array}$$

Exercise 3. Prove that optimization over the ice cream cone is equivalent to convex quadratically constrained quadratic programming; that is the following optimization problem:

$$\begin{array}{lll} \min & \mathbf{x}^TC_0\mathbf{x}+\mathbf{b}_0^T\mathbf{x} & \\ \text{st} & \mathbf{x}^TC_i\mathbf{x}+\mathbf{b}_i^T\mathbf{x}\le 0 & \text{for } i=1,\cdots,k \end{array}$$

for positive semidefinite matrices C_i, $i=0,\cdots,k$ can be transformed into 2.1.

Optimization over the ice cream cone can also be formulated as a semidefinite program Notice that

$$\mathbf{x}\ge_Q \longleftrightarrow \begin{pmatrix} x_0 & x_1 & \cdots & x_n \\ x_1 & x_0 & & \\ \vdots & & \ddots & \\ x_n & & & x_0 \end{pmatrix} = \begin{pmatrix} x_0 & \bar{\mathbf{x}}^T \\ \bar{\mathbf{x}} & x_0 I \end{pmatrix} \succeq 0$$

where $\bar{\mathbf{x}}=(x_1,\cdots,x_n)^T$, (recall that $\mathbf{x}$ is indexed from 0, $\bar{\mathbf{x}}$ is $\mathbf{x}$ without x_0.) The matrix above is sum of a multiple of identity and a rank two matrix. Therefore, it is easy to compute its eigenvalues.

2.2 Complementary Slackness for p-cones:

Define p-cones analogously as:

$$\mathbf{x}\ge_p 0 \Leftrightarrow \mathbf{x}\in\mathcal{K}_p{:=}\{\mathbf{x}:\ x_0\ge (x_i^p+\ldots+x_n^p)^{1/p}\}$$

Polar of $\mathcal{K}_p$ is the cone $\mathcal{K}_q$ where $p\ge 1$ and $1/p+1/q=1$. Note that for $p=\infty$, $q=1$ and the optimization problem becomes a linear program. For $p=q=2$ we have the ice cream cones. We saw that in this case the optimization problem can be solved analytically. For p other two this is not the case. Consider the primal and dual problems:

$$\begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{st} & A\mathbf{x}=\mathbf{b} \\ & \mathbf{x}\ge_p 0 \end{array} \qquad\qquad \begin{array}{ll} \max & \mathbf{b}^T\mathbf{y} \\ \text{st} & A^T\mathbf{y}+\mathbf{z}=\mathbf{c} \\ & \mathbf{z}\ge_q 0 \end{array}$$

At the optimum we have:

$$\begin{array}{l} \mathbf{x}\ge_p 0 \\ \mathbf{z}\ge_q 0 \\ \mathbf{x}^T\mathbf{z}=0 \end{array}$$

These three conditions are equivalent to:

$$\frac{x_0^p}{z_o^q} = \frac{|x_1|^p}{|z_1|^q} = \ldots = \frac{|x_n|^p}{|z_n|^q}$$

equations. This result is a consequence of Hölder's inequality, more precisely the conditions under which Hölder's inequality becomes an equality.

LEMMA 2.2. *Let $1/p + 1/q = 1$ and $p, q \leq 1$. If $\mathbf{x} \geq_p 0$, $\mathbf{z} \geq_q 0$, and $\mathbf{x}^T\mathbf{z} = 0$, then*

$$\frac{x_0^p}{z_o^q} = \frac{|x_1|^p}{|z_1|^q} = \ldots = \frac{|x_n|^p}{|z_n|^q} \tag{2.2}$$

Proof. We have

$$\begin{aligned} x_0 z_0 &= -x_1 z_1 - \ldots - x_n z_n \\ &\leq |x_1 z_1| + \ldots + |x_n z_n| \\ &\leq \sqrt[p]{|x_1|^p + \ldots + |x_n|^p}\sqrt[q]{|z_1|^q + \ldots + |z_n|^q} && \text{(by Hölder's inequality)} \\ &\leq x_0 z_0 && \text{since } \mathbf{x} \geq_p 0 \text{ and } \mathbf{z} \geq_q 0 \end{aligned}$$

■

Degeneracy in SDP We now focus on the primal and dual standard SDP

$$\begin{array}{llll} min & CX & max & \mathbf{b}^T\mathbf{y} \\ \text{st} & \mathcal{A}\mathbf{vec}X = \mathbf{b} & \text{Mat}(\mathcal{A}^T\mathbf{y}) + Z = C & \\ & X \succeq 0 & & Z \succeq 0 \end{array} \tag{2.3}$$

We have seen that at the optimal points $XZ = 0$. Product of two symmetric matrices is another symmetric matrix and therefore X and Z must commute. Therefore they share a system of eigenvectors. We can write :

$$X = Q\text{Diag}(x_1, \ldots, x_n)Q^T \text{and} \qquad Z = Q\text{Diag}(z_1, \ldots, z_n)Q^T$$

Therefore $XZ = 0$ is the same as saying X and Z share an orthogonal system of eigenvectors and eigenvalues satisfy

$$x_i S_i = 0, \text{ for } i = 1 \ldots n$$

The following lemma follows almost immediately.

LEMMA 2.1. *At the optimum*

$$\text{rank }(X) + \text{rank }(Z) \leq n$$

Note: This lemma implies that when the shared system of eigenvectors, that is the orthogonal matrix Q imposes an order on the eigenvalues of X and Z, then at least one of the corresponding eigenvalues must be zero. Let $x_1, \cdots, x_n$ be the eigenvalues of X, and similarly $z_1, \cdots, z_n$ eigenvalues of Z, where the order is determined by the shared eigenvectors matrix Q.

DEFINITION 2.1. *For the optimal solutions X and Z strict complementarity holds if exactly one of x_i and z_i is zero for each $i = 1, \cdots, 0$.*

Think of Primal and Dual relations in LP and assume that neither primal nor dual degeneracy holds. In LP primal nondegeneracy means that dual has unique solution; dual nondegeneracy means that the primal has unique solutions. In that case one also has strict complementarity. In fact in LP, when we have primal and dual nondegeneracy, precisely m of x_i's are nonzero and precisely $n - m$ of z_i's are nonzero.

In SDP the situation is different. First observe the following immediate lemma:

COROLLARY 2.0.1. *Strict Complementarity iff* $\operatorname{rank}(X) + \operatorname{rank}(Z) = n$.

Unlike LP, however, number of nonzero eigenvalues of optimal solutions are not some fixed numbers, but they may fall in a range. Even for given m, number of equality constraints, and n, dimension of, different SDP'

$$X = (x_1, \ldots, x_n, 0, \ldots, 0) S = (z_1, \ldots, z_n, 0, \ldots, 0)$$

These are valid values for r and s. rank $(X)+$ rank $(S) \leq n$. $r + s = n$ is valid only in the case of strict complementarity.

Warning

A key statement in the lecture claimed that for each pair of primal and dual SDP's there exist one set of optimal solutions which satisfy strict complementarity. This claim is false. In fact, there are SDP's with unique primal and dual solutions for which strict complementarity does not hold. For the correct version of proofs refer to the paper "Complementarity and nondegeneracy in semidefinite programming" by F. Alizadeh, J.P. Haeberly and M. Overton. This paper is posted in semidefinite programming home page under "duality theory, facial structure and degeneracy."