

# LINMA2471 : Optimization models and methods: Lecture 8

Antoine Noé, Ad'Oul Ismail, Dehem Boris

S10 - November, 18th 2015

From week 1 to week 3, we saw how to model convex problems. From week 4 to 6, we studied methods that deal with that type of problems. The disadvantage of those methods is that they can be quite slow, especially when compared with linear solvers (for linear systems). We now look at a certain class of convex problems, conic problems, that conserve some of the properties that make linear systems so effective. In particular, we want our problems to have duality, because a dual problem will be a much better problem to certify optimality. We start by looking at quadratic functions  $x^T Q x$

$$\begin{array}{l} x^T x \\ \downarrow \\ x^T Q \\ \updownarrow \\ (L^T x)^T (L^T x) \end{array} \quad Q \geq 0 \iff Q = LL^T$$

This last expression  $(L^T x)^T (L^T x)$  is a linear transformation of  $x^T x$  so it preserves convexity.

## 1 Exercise session on conic optimization

In this section, we will do a quick reminder on some exercises of Exercise session 5 on conic modelling: second order cone and duality.

### Exercise 2

In this exercise, we have a rotated second order cone  $\mathbb{L}_R^n$

$$\mathbb{L}_R^n = \{(x, y, z) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}_+ \mid x_1^2 + x_2^2 + \dots + x_n^2 \leq yz\}$$

We can prove that this set is a cone by replacing it by a second order cone. Why do we then use the rotated second order cone? The rotated second order cone is usefull because it makes some models easier (for example, the models in exercise 3).

### Exercise 4

The question here is "how do we deal with conic modelling of **functions**"? Lets take example (d).

Function  $(x, z) \rightarrow \frac{\|x\|^2}{z}$  (on the domain  $z > 0$ ). We have

$$\min_{\substack{x \in \mathbb{R}^n \\ z \in \mathbb{R}_{>0}}} \frac{\|x\|^2}{z}$$

This is not a conic problem because there is no linear objective. Our goal is to transform it in order to have a linear objective:

$$\begin{aligned} \min_{\substack{x \in \mathbb{R}^n \\ z \in \mathbb{R} > 0}} \frac{\|x\|^2}{z} \\ \Updownarrow \\ \min_{\substack{t \in \mathbb{R} \\ x \in \mathbb{R}^n \\ z \in \mathbb{R} > 0}} t \quad s.t. \\ \|x\|^2 \leq tz \end{aligned}$$

Now, we have a linear objective, and constraints of the rotated second order cone type.

## Exercise 6 : Conic duality

In this exercise, we need to find the dual for some optimization problems, using an equivalent conic formulation. To compute a dual problem, we need to reformulate it to a standard conic formulation.

Lets consider the dual and the primal problem:

$$\begin{aligned} (D) \quad \max b^T y \quad & \longleftrightarrow \quad \min c^T x \quad (P) \\ A^T y \preceq_K c \quad & Ax = b \\ & x \succeq_{K^*} 0 \end{aligned}$$

where (D) stands for the dual problem and (P) the primal problem. Note that for the positive real axis, second order cones and the semi definite cone:  $K = K^*$ .

## 2 Example: Duality

### 2.1 Primal problem

Lets consider the following problem:

$$\begin{aligned} \min \sqrt{z^T Q z} \quad (Q \geq 0) \\ Cz \leq d \end{aligned}$$

The function we minimize is quadratic, and Q is positive semi-definite. The constraint  $Cz \leq d$  is a linear constraint. We need to transform this formulation into one of the two formulations above.

Let's say we first want to take care of the objective (since it is quadratic).

$$\begin{aligned} \min t \\ \sqrt{z^T Q z} \leq t \quad Q = LL^T \\ \|L^T z\| \leq t \end{aligned}$$

This is equivalent

$$\begin{pmatrix} t \\ L^T z \end{pmatrix} \in \mathbb{L}^n$$

Now, we want to reformulate the linear constraint:

$$Cz \leq d \iff d - Cz \geq 0$$

This can be written as

$$d - Cz \in \mathbb{R}_+^m$$

Now, we are going to write everything using the cones:

Let  $y = \begin{pmatrix} t \\ z \end{pmatrix} \in \mathbb{R}^{n+1}$  ( $t$  is composed of 1 element and  $z$  of  $n$  elements).

Objective  $-\max b^T y$  with  $b = \begin{pmatrix} -1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$

Constraint:  $c - A^T u \geq_K 0$  with  $K = \mathbb{R}_+^m \times \mathbb{L}^n$  (big vector with the top part being  $\mathbb{R}_+^m$  and the bottom part  $\mathbb{L}^n$ )

$$\begin{bmatrix} d - Cz \\ t \\ L^T z \end{bmatrix} \in K, \text{ where } d - Cz \in \mathbb{R}_+^m \text{ and } \begin{bmatrix} t \\ L^T z \end{bmatrix} \in \mathbb{L}^n$$

Note that  $d - Cz$  is composed of  $m$  elements,  $t$  of one element and  $L^T z$  of  $n$  elements.

We are not completely done yet because we still need  $c$  and  $A^T$ :

$$c = \begin{pmatrix} d \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \quad A^T = \begin{pmatrix} 0 & C \\ -1 & 0^{1 \times n} \\ 0^{n \times 1} & -L^T \end{pmatrix}$$

The first column of  $A^T$  corresponds to the coefficients of  $t$ .

By telling our solver the different matrices and the cone, it will solve the problem. Why is this useful? Because now we can write the dual without a problem.

## 2.2 Dual Problem

$$K^* = (\mathbb{R}_+^m \times \mathbb{L}^n)^* = (\mathbb{R}_+^m)^* \times (\mathbb{L}^n)^* = \mathbb{R}_+^m \times \mathbb{L}^n = K$$

Lets consider  $x = \begin{pmatrix} u \\ v \end{pmatrix}$  with  $u \in \mathbb{R}_+^m$  and  $v \in \mathbb{L}^n$

The dual is

$$\begin{aligned} & -\min \begin{bmatrix} d \\ 0 \end{bmatrix}^T \begin{pmatrix} u \\ v \end{pmatrix} \\ & \begin{bmatrix} 0^{1 \times n} & -1 & 0^{1 \times m} \\ C^T & 0^{k \times 1} & -L \end{bmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -1 \\ 0^{k \times 1} \end{pmatrix} \end{aligned} \quad (1)$$

Where  $n$ ,  $m$  and  $k$  are the dimensions of  $C$  and  $L$ .

After simplifying, we have

$$\begin{aligned}
& -\min d^T u \\
& u \geq 0 \\
& v_0 \geq \| (v_1 \dots v_n) \| \\
& C^T u - L \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = 0 \\
& -v_0 = -1
\end{aligned}$$

By putting  $w = (v_1 \dots v_n)$ , we get the following simplified model

$$\begin{aligned}
& -\min d^T u \\
& u \geq 0 \\
& \| w \| \leq 1 \\
& C^T u = Lw
\end{aligned}$$

This model cannot be simplified any further.

### 3 Theorems and properties of Dual

After this example of conic dual computation, we want to study the properties of the conic dual optimization and some of its theorems. One interesting design allows to prove some theorems. Let us first explain this design to eventually find the dual of the primal conic problem (P):

$$\begin{aligned}
& \min \mathbf{b}^T \mathbf{y} \\
& \mathbf{A}^T \mathbf{y} \preceq_K \mathbf{c}
\end{aligned}$$

First, to build the standard dual problem, We write the constraints as a system of inequations where coefficients  $\lambda_i \geq 0$  multiplies  $c_i$  (positivity of  $\lambda$  assures the preservation of inequalities):

$$\frac{\begin{bmatrix} a_1^T \mathbf{y} \leq c_1 \\ a_2^T \mathbf{y} \leq c_2 \\ \vdots \\ a_n^T \mathbf{y} \leq c_n \end{bmatrix} \times \begin{matrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{matrix}}{(\sum_{i=1}^n \lambda_i a_i)^T \leq \sum_{i=1}^n \lambda_i c_i}$$

We can write the new system system as:

$$(\mathbf{A}\lambda)^T \mathbf{y} \leq \mathbf{c}^T \lambda$$

The dual goal is to find the best upper bound for this system. This is, that best upper bound is found by minimizing the dual objective.

When  $A\lambda = b$ , we get the system  $b^T y \leq c^T \lambda$ . The upper bound depends on  $\lambda$  and the dual problem (D) becomes:

$$\begin{aligned} & \min c^T \lambda \\ & \text{such that } A\lambda = b \\ & \lambda \geq 0 \end{aligned}$$

In a conic problem, we have:

$$\begin{bmatrix} a_1^T y \\ a_2^T y \\ \vdots \\ a_n^T y \end{bmatrix} \preceq_K \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

A simple example on the Lorentz cone show that the standard dual does not always work in conic optimization: let us have:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \preceq_{\mathbb{L}^2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

This is valid since on the Lorentz cone we must have in this case the condition  $1 \geq \|(-1, 0)\|$  satisfied, which is the case. However, if we now multiply  $c^T$  by a positive vector  $\lambda = (2, 10, 0)$ , we get then:

$$\begin{array}{ccc} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} & \preceq_{\mathbb{L}^2} & \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{array}{l} \times 2 \\ \times 10 \\ \times 0 \end{array} \\ \hline & \overset{?}{0} \leq & 2 - 10 + 0 \end{array}$$

In this example, the positivity of  $\lambda$  does not preserve the inequalities as we need to in order to use the standard dual formulation.

### Necessary and sufficient condition

The idea is that when starting with something positive and multiplying by  $\lambda$ , we must stay positive. For instance a good multiplier on the Lorentz cone would be such that:

$$\begin{array}{ccc} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} & \preceq_{\mathbb{L}^2} & \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{array}{l} \times \lambda_1 \\ \times \lambda_2 \\ \vdots \\ \times \lambda_n \end{array} \\ \hline & 0 \leq & \sum_{i=1}^n \lambda_i x_i \end{array}$$

The condition to have a correct  $\lambda$  is the following: whatever the  $x$  we start with,  $\lambda \in K^*$  if  $\lambda^T x \geq 0$ :

$$\lambda^T x \geq 0 \quad \forall x \in K \Leftrightarrow \lambda \in K^*$$

Where  $K^*$  is the dual cone.

### Example 1 : dual cone for $\mathbb{R}_+$ and $\mathbb{R}_+^n$

$(\mathbb{R}_+)^* = ?$

$$(\mathbb{R}_+)^* = \{\lambda | x\lambda \geq 0, \forall x \in \mathbb{R}_+\}$$

- $\lambda \geq 0$  : this works because  $x\lambda \geq 0$  when  $x, \lambda \geq 0$
- $\lambda < 0$  : one can pick  $x = 1$  and observe that  $x\lambda < 0$ , that is  $\lambda \notin (\mathbb{R}_+)^*$

$(\mathbb{R}_+^n)^* = ?$

$$(\mathbb{R}_+^n)^* = \{\lambda | x^T \lambda \geq 0, \forall x \in \mathbb{R}_+^n\}$$

- $\lambda \geq 0$  : this works because  $x^T \lambda \geq 0$  when  $x, \lambda \geq 0$
- $\lambda_i < 0$  : one can pick a  $x$  such that one of its elements is negative, but not negative everywhere:

$$\lambda = \begin{bmatrix} ? \\ \vdots \\ < 0 \\ \vdots \\ ? \end{bmatrix}, x = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

This way,  $\lambda_i^T x_i < 0$ , and that implies dual cone is  $(\mathbb{R}_+^n)$ .

### Example 2 : dual cone for $\mathbb{L}^n$

$(\mathbb{L}^n)^* = ?$

$$(\mathbb{L}^n)^* = \{\lambda | x^T \lambda \geq 0, \forall x \in \mathbb{L}^n\}$$

- $\lambda \in \mathbb{L}^n$  :  $x^T \lambda = x_0 \lambda_0 + [x_1 \dots x_n][\lambda_1 \dots \lambda_n]^T$ :  
Because we know  $x_0 \geq \|x_1 \dots x_n\|$  and  $\lambda_0 \geq \|\lambda_1 \dots \lambda_n\|$ , we can say:

$$\begin{aligned} x^T \lambda &\leq x_0 \lambda_0 + \|x_1 \dots x_n\| \|\lambda_1 \dots \lambda_n\| \\ &\geq x_0 \lambda_0 - \|x_1 \dots x_n\| \|\lambda_1 \dots \lambda_n\| \geq 0 \end{aligned}$$

- $\lambda \notin \mathbb{L}^n$  :  $\lambda_0 < \|\lambda_1 \dots \lambda_n\|$   
One can pick  $x$  the following way:  $x_0 = \|\lambda_1 \dots \lambda_n\|$ ;  $[x_1 \dots x_n] = -[\lambda_1 \dots \lambda_n]$   
In that case,

$$\begin{aligned} x^T \lambda &= \lambda_0 \|\lambda_1 \dots \lambda_n\| - \|\lambda_1 \dots \lambda_n\|^2 \\ &= (\lambda_0 - \|\lambda_1 \dots \lambda_n\|) \|\lambda_1 \dots \lambda_n\| < 0 \end{aligned}$$

This implies the second order cone is self-dual.

### Theorems

We can conclude what has been done with two theorems on conic duality:

- **Weak duality** : if  $x$  is feasible for the primal (P) and  $y$  is feasible for the dual (D), then

$$c^T x \geq b^T y$$

- Under some assumptions, optimal value of (P) is equal to the optimal value of (D). This is detailed in the next lecture.