CM7 : LINMA 2471 : Optimization methods and models : Lecture 7

Ciamarra Geoffrey   Losseau Bruno   Lucas Joachim

November, 4th 2015

1 Linear and sublinear convergence

For methods with sublinear convergence, the value of the error firstly downs very quickly but after some value, the decreases becomes really slow. For methods with linear convergence, the decrease of the error is, at the beginning, slowler then sublinear methods. However, after the value where the decrease for sublinear methods becomes slow, the decrease for linear method remains constant on a logarithmic scale.

Thus, the choice of the method depends on the application. If one needs a lot of accuracy, it is preferable to use linear methods but if high accuracy is not that much important (i.e. we only look for one digit), one then should use sublinear methods. Some graphs to illustrated these results are available on the *Linear convergence is faster than sublinear convergence (eventually) [missing figure from week 7]* pdf file.

2 Conic optimization

In order to introduce the conic optimization concept, let's begin with an linear example.

Example 2.1 (Linear optimisation problem)**.**

$$\begin{aligned} \max_{y_i} \quad & 2y_1 + 3y_2 + 2y_3 \\ \text{subject to} \quad & y_1 + y_2 \leq 1 \quad\quad (1) \\ & y_2 + y_3 \leq 2 \quad\quad (2) \\ & y_3 \leq 3 \quad\quad (3) \end{aligned}$$

A feasible point is given by $(y_1, y_2, y_3) = (1, 0, 2)$ which provides a cost function value 6.

How to know whether 6 *is strictly lower than or strictly higher than or equal to the optimal objective value ?*

Since a feasible point gives 6 as cost function value, the optimal objective value is not strictly lower than 6. The value 6 can be increased to 7 by taking the feasible point $(y_1, y_2, y_3) = (2, -1, 3)$. So the optimal objective value is not equal to 6 but higher than 6.
Concerning the value 7, we could ask our question again, *how to prove that the optimal objective value is higher than or equal to* 7 *?*

We need an upper bound on the value for the objective function. This can be achieved by taking a linear combination of the constraints.

Indeed,

$$2 \cdot \text{ constraint } 1 + 1 \cdot \text{ constraint } 2 + 1 \cdot \text{ constraint } 3 = 2y_1 + 3y_2 + 2y_3$$
$$\leq 7$$

7 is an upper bound for the objective value function. Since we have found a feasible point for which the objective value is equal to 7, we can claim that the optimal objective value is equal to 7.

It is possible to find a bound for the objective value for any linear problem. We want to extend the notion of linear problems while keeping their 'nice' properties of linear problems (i.e. duality and efficient algorithms).

In order to extend this notion, we will generalize the inequalities $\leq$ and $\geq$ in the linear problems.

With $K \subseteq \mathbb{R}^n$, we define an generalized type of inequalities

$$a \succeq_K 0 \Leftrightarrow a \in K$$

On the basis of this definition, we also have the following properties

$$a \succeq_K b \Leftrightarrow a - b \succeq_K 0 \Leftrightarrow a - b \in K$$
$$a \preceq_K b \Leftrightarrow b \succeq_K a \Leftrightarrow b - a \succeq_K 0 \Leftrightarrow b - a \in K$$

Let us also impose two sensible properties for an order

1. $a \succeq_K 0 \Rightarrow \lambda a \succeq_K 0$ for any $\lambda \geq 0$
2. $a \succeq_K 0$ and $b \succeq_K 0 \Rightarrow a + b \succeq_K 0$

The properties 1. means that K is a *cone* and the properties 2. means that K is *closed under addition*. The K must be convex, K is said to be a convex cone. In fact, for any cone K (closed under nonnegative scalar multiplication), we have the equivalence

$$K \text{ is closed under addition } \Leftrightarrow K \text{ is convex}$$

2.1 Generalization

From the usual formulation of the problem, one can generalize :

$$\max b^T y \text{ such that } A^T y \leqslant c$$
$$\Downarrow$$
$$\max b^T y \text{ such that } A^T y \preceq_K c$$

This problem is convex.
Considering the standard linear case, $K = \mathbb{R}^n_+$.

Example 2.2. *Going back to example 2.1, how would one define the set K ?*
As :
$$A^T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Defining $K = \mathbb{R}^3_+$ implies :

$$C - A^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} y \subset K = \mathbb{R}^3_+$$

$$\begin{pmatrix} 1 - y1 - y2 \\ 2 - y2 - y3 \\ 3 - y3 \end{pmatrix} \subset K = \mathbb{R}^3_+$$

2.2 Solving a problem on different cones

As one consider polyhedron cones, the boundary is linear. Also, applying a linear operation to a linear problem keeps the problem linear. As a result, applying a rotation or a scaling operation to a cone and solving the problem on the newly obtained cone is equivalent to solving it on the first cone.

The conic optimization is useful since it is more general than linear optimization. The sets considered are not polyhedrons anymore but rather curved sets. Thus, it can cover more problems.
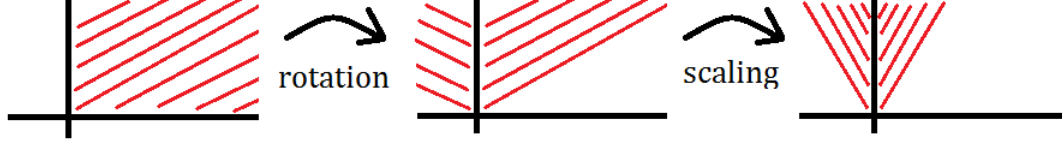
FIGURE 1 – Solving a problem on one cone is equivalent to solve it on a newly obtained cone via rotation and scaling.

2.3 Lorentz cone

A common example of a cone is the Lorentz cone or Ice cream cone. It is defined such as :

$$\mathbb{L}^n = \{(x_0, ... x_n) \in \mathbb{R}^{n+1} \mid \sqrt{x_1^2 + ... + x_n^2} \leq x_0\}$$

One can verify it is indeed a cone by checking that when multiplying a value in the cone by a scalar it stays in the cone ; and that addition is preserved (intuition : rule of parallelogram). Also, checking if the hull is a cone is the same as checking whether it is convex or not.

Example 2.3. *Let the problem :*

$$\max(y_1 + y_2)$$
$$y_1^2 + y_2^2 \leq 7$$

One notes that the constraint is quadratic and semi-definite positive. One would like to formulate the problem as linear, except for the cone. All that is non-linear shall be put in the cone, hence :

$$y_1^2 + y_2^2 \leq 7 \Leftrightarrow \sqrt{y_1^2 + y_2^2} \leq \sqrt{7}$$

$$\Updownarrow$$

$$\begin{pmatrix} \sqrt{7} \\ y1 \\ y2 \end{pmatrix} \in \mathbb{L}^2 \Leftrightarrow \begin{pmatrix} \sqrt{7} \\ y1 \\ y2 \end{pmatrix} \succeq_{\mathbb{L}^2} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Updownarrow$$

$$\begin{pmatrix} \sqrt{7} \\ 0 \\ 0 \end{pmatrix} \succeq_{\mathbb{L}^2} \begin{pmatrix} 0 \\ -y_1 \\ -y_2 \end{pmatrix}$$

Finally :

$$C = \sqrt{7}, \qquad A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

And the problem can be written as :

$$\max(y_1 + y_2)$$
$$-y_1 - y_2 \preceq_{\mathbb{L}^2} \sqrt{7}$$

2.4 Requirements for K

1. $x \succeq 0$ and $x \preceq 0 \Rightarrow x = 0$
   which means $K \cap (-K) = \{0\}$ (the cone is *pointed*)
2. We define the strict inequality by $a \succ 0 \Leftrightarrow a \in int\ K$ (and $a \succ b$ iff $a - b \in int\ K$)
   Hence we require $int\ K \neq \emptyset$ (the cone is *solid*)
3. Finally, we would like to be able to take limits :

$$\text{If } \{x_i\}_{i \to \infty} \text{ with } x_i \succeq_K 0\ \forall i, \text{ then } \lim_{i \to \infty} x_i = \bar{x} \Rightarrow \bar{x} \succeq_K 0$$

which is equivalent to saying that K is *closed*.

A convex cone K that is solid, pointed and closed will be called a *proper cone.*

As seen in subsection 2.2, in the conic optimization theory, all the cones can always be merged as an only one according to the following definition.

Definition 2.1. *Considering several conic constraints*

$$A_1^T y \preceq_{K_1} c_1 \text{ and } A_2^T y \preceq_{K_2} c_2$$

which are equivalent to

$$c_1 - A_1^T y \in K_1 \text{ and } c_2 - A_2^T y \in K_2$$

one introduces the product cone $K = K_1 \times K_2$ to write

$$(c_1 - A_1^T y, c_2 - A_2^T y) \in K_1 \times K_2$$

$$\Leftrightarrow \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} A_1^T \\ A_2^T \end{pmatrix} \in K_1 \times K_2 \Leftrightarrow \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} - \begin{pmatrix} A_1^T \\ A_2^T \end{pmatrix} \succeq_{K_1 \times K_2} 0$$

If K_1 and K_2 are proper, $K_1 \times K_2$ is also proper.

However, in this course, and for the exercises, we will not merge the cones.

2.5 Equivalence with convex otpimization

Since a cone is convex, conique optimization can be related to convex optimization : conic optimization is clearly a special case of convex optimization. What about the reverse statement ?

$$\min_{x \in \mathbb{R}^n} f(x) \text{ such that } x \in X \subseteq \mathbb{R}^n$$

– The objective of a convex problem can be assumed *w.l.o.g.* to be *linear* w.l.o.g. : $f(x) = c^T x$
– The feasible region of a convex problem can be assumed *w.l.o.g.* to be in the *conic* standard format :

$$X = \{x \in K \text{ and } Ax = b\}$$

$\Rightarrow$ conic optimization *equivalent* to convex optimization.

One can prove that everything in the use of conique optimization, must be convex. Considering the conic constraint :

$$c - A^T y \in K$$

choosing K to be convex is a good way to ensure the constraint to be convex too since linear transformations preserve convexity.

2.6 Conclusion with duality

What about the conic dual problems ? Since we generalized

$$\max b^T y \text{ s.t. } A^T y \leq c$$

to

$$\max b^T y \text{ s.t. } A^T y \preceq_K c$$

it could be tempting to generalize

$$\min c^T x \text{ s.t. } Ax = b \text{ and } x \geq 0$$

to

$$\min c^T x \text{ s.t. } Ax = b \text{ and } x \succeq_K 0.$$

However, this is not the right primal-dual pair ! Actually, we must also dualize the cone ; and the corresponding correct primal-dual pair is given by

$$\max b^T y \text{ such that } A^T y \preceq_K c$$

$$\min c^T x \text{ such that } Ax = b \text{ and } x \succeq_{K^*} 0$$

where we introduce the dual cone K^* defined as

$$K^* = \{z \in \mathbb{R}^n \text{ such that } x^T z \geq 0 \ \forall x \in K\}$$

– K^* is a convex cone, called the *dual* cone of K ;
– K^* is always *closed*, and if K is closed, $(K^*)^* = K$;
– K is *pointed* (resp. solid) $\Rightarrow K^*$ is *solid* (resp. pointed)
– Cartesian products : $(K_1 \times K_2)^* = K_1^* \times K_2^*$;
– $(\mathbb{R}^n+)^* = \mathbb{R}^n+, (\mathbb{L}^n)^* = \mathbb{L}^n, (\mathbb{S}_+^n)^* = \mathbb{S}_+^n$:
  these cones are self-dual, but there are (many) cones which are not.