

LINMA2471 : Optimization II - Models and Methods

Course Notes - 30th September 2015

CHIEM Benjamin - 2174 1200
GERARD Céline - 2590 1200
HAVELANGE Andine - 2583 1200

October 2015

0.1 Modelling Tricks

After studying the standard form of some optimization problems, we will see some modelling tricks which permit to transform a non-linear or non-convex problem into a linear or convex optimization problem.

0.1.1 Monotonicity

Definition 1 *A monotonic function over an interval is a function that is either increasing or decreasing over this interval.*

The transformation that is described in the following lines uses monotonic functions to turn an optimization problem into a simpler one. These operations do not change the problem but can change the solution and the optimal value of the objective of the problem. Let's see some examples.

Example 1 $\min \|x\|_2$ with $x \in X$ is equivalent to $\min \|x\|_2^2$ with $x \in X$. In fact, $z \rightarrow z^2$ is a monotonic function (increasing in this case) over \mathbb{R} . By solving this problem, we will find the same optimal solution than the first model but the value of the objective function will be different.

Example 2 These functions are monotonic and can be used like in the previous example to simplify the problem :

. $z \rightarrow e^z$

. $z \rightarrow \log(z), (z > 0)$

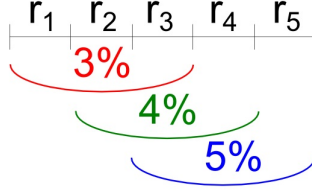


Figure 1: Rates

. $z \rightarrow -\frac{1}{z}$, ($z > 0$) for example, we can change $\min \frac{1}{\|x\|_2}$ to $\min -\|x\|_2$

We not only use this trick to modify the objective function, but also some constraints. For example :

$$f(x) \leq b \Leftrightarrow e^{f(x)} \leq e^b$$

Example : Advertisement for bank account : This example allows us to look at the effects of monotonicity in real life optimization problems. We put money into an account where we can't take our money back sooner than 5 years after. The bank guarantees that we have a high percentage when we take back our money after 5 years and assures the three rate over 3 years given on Figure 1.

What is the worst global rate compatible with this ? Therefore, we are looking for the cumulative effect of the 5 rates (given over a year).

To solve this problem, we introduce the variables r_i corresponding to the rate per year where $i = 1, \dots, 5$. We don't want a rate which is negative or null so we have the following optimization problem :

$$\min_{r_i > 0} r_1 r_2 r_3 r_4 r_5$$

under the constraints :

$$r_1 r_2 r_3 \geq 1.03$$

$$r_2 r_3 r_4 \geq 1.04$$

$$r_3 r_4 r_5 \geq 1.05$$

This problem is not linear. We can transform it into a linear problem if we use the logarithm function. Indeed, let $y_i = \log(r_i) \forall i = 1, \dots, 5$. We can do this because the logarithm is a monotonic increasing function. The problem can be written as :

$$\min y_1 + y_2 + y_3 + y_4 + y_5$$

under the constraints :

$$y_1 + y_2 + y_3 \geq \log(1.03)$$

$$y_2 + y_3 + y_4 \geq \log(1.04)$$

$$y_3 + y_4 + y_5 \geq \log(1.05)$$

This problem is now linear (and thus convex).

Remark

If we don't make the change of variable $y_i = \log(r_i) \forall i = 1, \dots, 5$, the variables r_i appear in \log so the problem isn't linear.

0.1.2 Change of variables

We use change of variables in order to transform the problem into a linear or convex problem. For example, if every variable appears in a logarithm, then we can use the change of variable to remove the logarithm.

NB : every appearance of the variables needs to match the change of variables!

For example, signomials¹ can be converted into convex functions thanks of this trick. Let's consider the signomial $\frac{x_1 x_2^2}{x_3^{\frac{1}{2}}}$. Let $x_i = e^{y_i}$, we obtain the following expression by change of variable $e^{y_1 + 2y_2 - \frac{1}{2}y_3}$. Moreover, the exponential of a linear function is convex and the change of variables conserves the convexity (See later).

0.1.3 Misleading/Deceptive appearances

In this part, we are looking for a polynomial $p(x)$, $x \in \mathbb{R}$ of degree D which fulfils some characteristic (See Figure 2):

- . For $x \in [0; f_1]$, $p(x) \geq 3$
- . For $x \in [f_2; f_3]$, $p(x) \leq 0.5$

Let $p(x) = \sum_{i=0}^D a_i x^i$. On the Figure 2, we observe that the polynomial will not be linear or convex. We then have the impression that the optimization problem is not linear. But it is not the case, given that the variables are not the x_i but the a_i . Furthermore, the constraints (2 examples of constraints are given below) are linear.

$$\sum_{i=0}^D a_i 50^i \geq 3 \Leftrightarrow p(50) \geq 3$$

$$\sum_{i=0}^D a_i 100^i = 1 \Leftrightarrow p(100) = 1$$

¹A signomial is an algebraic function of one or more variables of the form:
 $f(x_1, x_2, \dots, x_n) = \sum_i (c_i \prod_j x_j^{a_{ij}})$

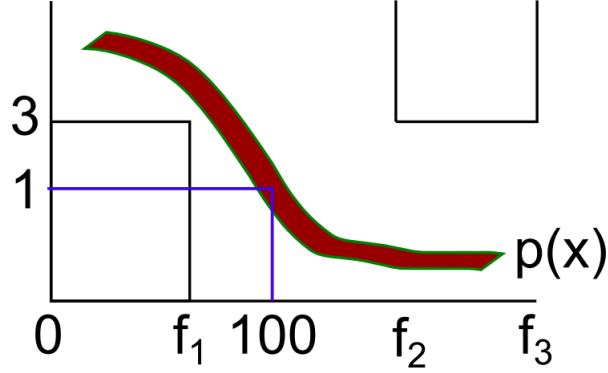


Figure 2: Polynomial $p(x)$

0.1.4 Flexibility

The flexibility of an optimization problem is its capacity to be solved in different ways depending on which criterion we really want to optimize. Let us consider an approximation problem. We have a set of points (x_i, y_i) in \mathbb{R}^2 , and we want to find the function (which is often a polynomial) which approximates *at best* these points. In other words, we want to find a function f that minimizes the errors ϵ_i at the approximation points:

$$\epsilon_i = |y_i - f(x_i)|$$

The question now is to choose the way to minimize errors ϵ_i . Indeed, we could choose to minimize overall errors with a least square criterion:

$$\min ||\epsilon||_2 = \sqrt{\sum_i \epsilon_i^2}$$

Another choice would be to minimize the sum of the errors:

$$\min ||\epsilon||_1 = \sum_i |\epsilon_i|$$

Finally, a last (often used) way to minimize errors is to minimize the maximum error:

$$\min \max_i \epsilon_i$$

All these considerations prove that, depending on the criterion we choose to optimize (i.e. depending on the context), the result could be different. For

example, the problem

$$\begin{array}{ll} \max & \text{ safety} \\ \text{cost} & \leq m \end{array}$$

will not lead to the same solution than

$$\begin{array}{ll} \min & \text{ cost} \\ \text{ safety} & \geq b \end{array}$$

0.1.5 Charnes and Cooper

Consider the non-linear problem which is a division of two linear expression :

$$\min \frac{c^T x + d}{f^T x + g}$$

under the constraints :

$$Ax \leq b$$

In this case, taking the logarithm of the objective function will not change anything: we will have the same problem.

We make the hypothesis that $f^T x + g > 0 \forall x$ such that $Ax \leq b$. If not, we have the solution $-\infty$ which is not an interesting solution.

To solve this problem and make it linear, we are going to homogenize the objective function. Let $x = \frac{y}{t}$ with $y \in \mathbb{R}^n$ and $t > 0 \in \mathbb{R}$. If we take $t = 1$ then we get back to the original problem. So, one solution in x correspond to several solutions in (y, t) (for example : $(x, 1)$, $(2x, 2)$, ... $(\lambda x, \lambda)$).

We can write the problem as :

$$\min \frac{\frac{c^T y}{t} + d}{\frac{f^T y}{t} + g}$$

with :

$$A \frac{y}{t} \leq b$$

By simplification, we obtain the following problem :

$$\min \frac{c^T y + dt}{f^T y + gt}$$

with :

$$Ay \leq bt$$

Now, we notice that the objective function's numerator and denominator are linear as the constraint. This problem has a property called homogeneity. If you take any solution, you can multiply any component by the same constant

and nothing changes. Mathematically :
 (y, t) solution $\Rightarrow (\lambda y, \lambda t)$ solution $\forall \lambda > 0$

We are going to choose solutions satisfying : $f^T y + gt = 1$. We can do this using the property of homogeneity. This step results in selecting one solution among the collection of solutions multiple of each other. The objective function gets simpler and linear and we add one constraint. So, we have the following linear optimization problem :

$$\min c^T y + dt$$

with :

$$\begin{aligned} Ay - bt &\leq 0 \\ f^T y + gt &= 1 \\ t &\geq 0 \end{aligned}$$

We compute y^* and t^* from this problem and take $x^* = \frac{y^*}{t^*}$ the solution of the original problem.

NOTE : If we have $t = 0$ at the optimum then the problem is unbounded and $x^* \rightarrow \infty$. Example :

$$\min \frac{1}{x}$$

with :

$$x \geq 1$$

This problem have an optimal value of 0 so the solution $x^* \rightarrow +\infty$.

0.2 Convex Optimization : Theorems and properties

0.2.1 Convex sets

Definition and examples

Definition 2 A set X is convex if and only if

$$x, y \in X \Rightarrow \lambda x + (1 - \lambda)y \in X \quad \forall 0 \leq \lambda \leq 1$$

Thus, a set is convex if and only if it contains all the segments joining any pair of its points.

Examples:

- $\mathbb{R}^n, \mathbb{R}_+^n, \emptyset$
- hyperplans ($\{x \mid b^T x = \beta\}$)
- open or closed half-spaces ($\{x \mid b^T x < \beta\}$ and $\{x \mid b^T x \leq \beta\}$)
- open and closed balls ($\{x \mid \|x - a\| < r\}$ and $\{x \mid \|x - a\| \leq r\}$)

Properties

Property 1 Given a collection of convex sets $\{C_i\}_{i \in I} \subseteq \mathbb{R}^n$ (I can be arbitrary), then $\bigcap_{i \in I} C_i$ is convex too.

It follows that polyhedrons are convex because they are intersection of half-spaces.

Property 2 Given a collection of convex sets $C_1, C_2, C_3, \dots, C_n$, their cartesian product $C_1 \times C_2 \times C_3 \times \dots \times C_n$ is convex too.

Property 3 If $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^n$ are convex then the Minkowski sum of X and Y , $X + Y = \{x + y \mid x \in X \text{ and } y \in Y\}$ is convex too.

Remark

The union of convex sets is not always convex!

0.2.2 Convex functions

Definition and examples

Definition 3 A function f with domain D is a convex function of and only if

D is convex and

$$x, y \in D \Rightarrow f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad \forall 0 \leq \lambda \leq 1$$

Examples:

- linear and affine functions are convex ($x \rightarrow \alpha x$, $x \rightarrow b^T x$ and $x \rightarrow b^T x + \alpha$)
- the norm function and the square of the norm function are convex functions ($x \rightarrow \|x\|$ and $x \rightarrow \|x\|^2$)
- quadratic forms ($x \rightarrow x^T Q x$) are convex functions when the matrix $Q \in \mathbb{R}^{n \times n}$ is semi positive definite
- the functions $x \rightarrow e^x$, $x \rightarrow \log(x)$ and $x \rightarrow |x|^p$ ($1 \leq p$) are convex

Definition 4 A function f is concave $\Leftrightarrow -f$ is convex.

Remark

Linear and affine functions are convex and concave.

Properties

To know whether a function is convex or not, we have to transform it into its epigraph and check if it is convex or not. But there are some useful properties of convex functions that we can use to spare time.

Property 4 If f is a convex function and $c \in \mathbb{R}_0^+$, then cf is convex.

Property 5 If f and g are convex functions, then $f + g$ is convex.

Property 6 Given a collection of convex functions $\{f_i\}_{i \in I} : \mathbb{R}^n \rightarrow \mathbb{R}$, $\sup_{i \in I} f_i$ is convex too.

$$\text{with } \left[\sup_{i \in I} f_i \right] (x) = \sup_{i \in I} f_i(x)$$

Property 7 Given $f(x, s)$ (with $x \in \mathbb{R}^n$ and $s \in \mathbb{R}$, a parameter) such that $x \rightarrow f(x, s)$ is convex for any s ,

$$\int_{s \in S} f(x, s) ds \text{ is convex}$$

Convexity and differential calculus

Property 8 Let f be a differentiable function of which the domain D is open. f is convex if and only if D is convex and

$$\forall x, y \in D, f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

This property means that, if f is convex, it will be above all its Taylor approximations of first order. This signifies that at any point, the tangent of the function is under the function. This is useful to make a piecewise approximation of f by linear functions (an example of such an approximation is shown on Figure 3). In order to obtain such an approximation, we have to choose n points at which we calculate the tangent of the function f and then take the maximum of the n tangents (in the example, $n = 3$).

Besides, the value of the approximation is lower than the real value on any point. It allows us to obtain lower bounds.

Property 9 Let f be a twice differentiable function of which the domain D is open. f is convex if and only if D is convex and

$$\forall x \in D, \nabla^2 f(x) \geq 0$$

0.2.3 Convexity and linear transformations

Linear transformations preserve convexity. Indeed,

Property 10 If $S \subseteq \mathbb{R}^n$ is convex and $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^m : x \rightarrow Ax + b$ is a linear transformation, then the image of S by Φ ,

$$\Phi(S) = \{\Phi(x) \mid x \in S\}$$

is convex too.

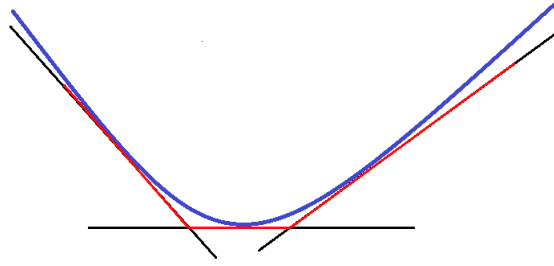


Figure 3: Illustration of the piecewise linear approximation of a convex function

Property 11 *If $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^m : x \rightarrow Ax + b$ is a linear transformation and $f : x \rightarrow f(x)$ is a convex function, then the composition*

$$f \circ \Phi = f(\Phi(x)) = f(Ax + b)$$

is convex too.

Property 12 *If $S \subseteq \mathbb{R}^n$ is convex and $\Theta : \mathbb{R}^m \rightarrow \mathbb{R}^n : x \rightarrow ax + b$ is a linear transformation, then the image of S by the inverse of Θ ,*

$$\Theta^{-1}(S) = \{x \mid \Theta(x) \in S\}$$

is convex too.