Tutorial 8 : Penalties and Barrier methods

Operations research, 3rd semester.

2024

Exercice 1 — *Beltrami penalties*

We want to minimize $f(x) = x_1 + x_2$ such that $x_1^2 - x_2 \leq 2$.

1. Solve the problem using the penalties method, with the penalty of Courant-Beltrami.
2. Same question with the equivalent following problem : minimize $f(x) = x_1 + x_2$ such that $x_1^2 - x_2 + x_3^2 = 2$.

▶ **Correction**

1. We set $g(x) = x_1^2 - x_2 - 2$ and $P(x) = g^+(x)^2 = \max(g(x), 0)^2$; we want to minimize $q(x, \mu) = f(x) + \mu \cdot P(x)$ for all μ.
   An optimal solution x_μ to this problem is a critical point and satisfies

$$\begin{aligned} \nabla(q(x_\mu, \mu)) &= 0 \\ \nabla f(x_\mu) + \mu \nabla P(x_\mu) &= 0 \\ \nabla f(x_\mu) + 2\mu \max(g(x_\mu), 0) \nabla g(x_\mu) &= 0 \end{aligned}$$

   We consider two cases : $g(x_\mu) \leq 0$ et $g(x_\mu) > 0$

   If $g(x_\mu) \leq 0$, :

$$\begin{aligned} \nabla f(x_\mu) &= 0 \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= 0 \end{aligned}$$

   This is a contradiction, considering the other case : $g(x_\mu) > 0$

$$\begin{aligned} \nabla f(x_\mu) + 2\mu g(x_\mu) \nabla g(x_\mu) &= 0 \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\mu(x_{\mu 1}^2 - x_{\mu 2} - 2) \begin{pmatrix} 2x_{\mu 1} \\ -1 \end{pmatrix} &= 0 \\ 1 + 2\mu(x_{\mu 1}^2 - x_{\mu 2} - 2)(2x_{\mu 1}) = 0 \text{ and } (x_{\mu 1}^2 - x_{\mu 2} - 2) &= \frac{1}{2\mu} \\ x_1 = -\frac{1}{2} \text{ and } x_2 &= -\frac{7}{4} - \frac{1}{2\mu} \end{aligned}$$

   We have $g(x) \geq 0$ as $x_1^2 - x_2 = 2 + \frac{1}{2k} \geq 2$. Finally :

$$\lim_{\mu \to +\infty} x_\mu = \begin{pmatrix} -1/2 \\ -7/4 \end{pmatrix}$$

   This point is the only critical point of the function ; moreover, the function q does not tend to $-\infty$ at infinity, so it is an optimal solution.
   Therefore, the sequence of x_μ converges, and it converges to an optimal solution of f.

2. We set $h(x) = x_1^2 - x_2 + x_3^2 - 2$ and $P(x) = h(x)^2$; we seek, for all μ, to minimize $q(x, \mu) = f(x) + \mu \cdot P(x)$.
   An optimal solution x_μ to this problem is a critical point and satisfies

$$\begin{aligned} \nabla(q(x_\mu, \mu)) &= 0 \\ \nabla f(x_\mu) + \mu \nabla P(x_\mu) &= 0 \\ \nabla f(x_\mu) + 2\mu h(x_\mu) \nabla h(x_\mu) &= 0 \\ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 2\mu(x_{\mu1}^2 - x_{\mu2} + x_{\mu3}^2 - 2) \begin{pmatrix} 2x_{\mu1} \\ -1 \\ 2x_{\mu3} \end{pmatrix} &= 0 \end{aligned}$$

   The third equality implies that $x_{\mu3} = 0$ or $(x_{\mu1}^2 - x_{\mu2} + x_{\mu3}^2 - 2) = 0$.

   If $(x_{\mu1}^2 - x_{\mu2} + x_{\mu3}^2 - 2) = 0$ then the two first equalities give a contradiction

   If $x_{\mu3} = 0$ then we get again

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\mu(x_{\mu1}^2 - x_{\mu2} - 2) \begin{pmatrix} 2x_{\mu1} \\ -1 \end{pmatrix} = 0$$

   We then have the same solution as in the first question.

Exercice 2 — *Barrier method*

Solve the following problem with the barrier method : minimize $f(x) = x_1 + x_2$ such that $x_1^2 - x_2 \leq 2$.

▶ **Correction**

We will try the two barriers $B(x) = \frac{-1}{g(x)}$ and $B(x) = -\log(-g(x))$.

Let $g(x) = x_1^2 - x_2 - 2$ and $B(x) = \frac{-1}{g(x)}$. We minimize $q(x, \mu) = f(x) + \mu B(x)$ such that $g(x) < 0$.

An optimal solution x_μ to this problem is a critical point and satisfies

$$\begin{aligned} \nabla(q(x_\mu, \mu)) &= 0 \\ \nabla f(x_\mu) + \mu \nabla B(x_\mu) &= 0 \\ \nabla f(x_\mu) + \mu \frac{1}{g(x_\mu)^2} \nabla g(x_\mu) &= 0 \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{\mu}{(x_{\mu1}^2 - x_{\mu2} - 2)^2} \begin{pmatrix} 2x_{\mu1} \\ -1 \end{pmatrix} &= 0 \end{aligned}$$

The second equation implies $\frac{\mu}{(x_{\mu1}^2 - x_{\mu2} - 2)^2} = 1$, so

$$\begin{aligned} x_{\mu1} &= \frac{-1}{2} \text{ and } \mu = (x_{\mu1}^2 - x_{\mu2} - 2)^2 \\ x_{\mu1} &= \frac{-1}{2} \text{ and } \mu = (x_{\mu2} - \frac{7}{4})^2 \\ x_{\mu1} &= \frac{-1}{2} \text{ and } x_{\mu2} = -\frac{7}{4} \pm \sqrt{\mu} \end{aligned}$$

As $g(x_\mu) < 0$, we can only have $x_{\mu2} = -\frac{7}{4} - \sqrt{\mu}$

$$x_{\mu1} = \frac{-1}{2} \text{ and } x_{\mu2} = -\frac{7}{4} + \sqrt{\mu}$$

This point is the only critical point of the function ; furthermore, the function q is convex, thus it is an optimal solution.

$$\lim_{\mu \to 0} x_\mu = \begin{pmatrix} -1/2 \\ -7/4 \end{pmatrix}$$

The sequence of x_μ converges, so it converges to an optimal solution of f.

Let us use the other barrier :

Let's set $g(x) = x_1^2 - x_2 - 2$ and $B(x) = -\log(-g(x))$. We aim to minimize $q(x, \mu) = f(x) + \mu B(x)$ subject to the constraint that $g(x) < 0$.

An optimal solution x_μ to this problem is a critical point and satisfies

$$\begin{aligned}
\nabla(q(x_\mu, \mu)) &= 0 \\
\nabla f(x_\mu) + \mu \nabla B(x_\mu) &= 0 \\
\nabla f(x_\mu) - \mu \frac{-1}{-g(x_\mu)} \nabla g(x_\mu) &= 0 \\
\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{\mu}{(x_{\mu 1}^2 - x_{\mu 2} - 2)} \begin{pmatrix} 2x_{\mu 1} \\ -1 \end{pmatrix} &= 0
\end{aligned}$$

the second equality implies that $\frac{-\mu}{(x_{\mu 1}^2 - x_{\mu 2} - 2)} = 1$, so

$$\begin{aligned}
x_{\mu 1} &= \frac{-1}{2} \text{ and } \mu = (-x_{\mu 1}^2 + x_{\mu 2} + 2) \\
x_{\mu 1} &= \frac{-1}{2} \text{ and } x_{\mu 2} = -\frac{7}{4} + \mu
\end{aligned}$$

We then have $g(x) = -\mu < 0$

This point is the only critical point of the function ; furthermore, the function q is convex, so it is the optimal solution.

$$\lim_{\mu \to 0} x_\mu = \begin{pmatrix} -1/2 \\ -7/4 \end{pmatrix}$$

The sequence of x_μ converges, so it converges to an optimal solution of f.

Exercice 3 — *Penalties and Lagrange multipliers*

We consider a generic problem in which we want to minimize $f(x)$ on $\mathbb{R}^n$ such that $g_i(x) \leq 0$ for every $i \in [\![1; m]\!]$ and such that $h_j(x) = 0$ for every $j \in [\![1; p]\!]$. We apply the penalties method with the Penalty of Beltrami P. We write $q(x, \mu) = f(x) + \mu P(x)$ and $x_k = \arg\min q(x, k)$.

We assume that f, g_i and h_j are C^1. Let x^* be an optimal solution of f, we assume x^* satisfy the linear independant constraint of qualification. Finally, we assume that the sequence x_k converges toward x^*.

1. Recall the Kuhn-Tucker conditions for that problem at x^*, we write λ_i and μ_j the Langrange multipliers respectively associated with the functions g_i and h_j.
2. May there exist multiple values for λ_i, μ_j satisfying the conditions ?
3. Write the gradient of $q(x, k)$ at x_k, what is the numerical value of that gradient ?
4. Deduce that $\lim_{k \to +\infty} 2k g_i^+(x_k) = \lambda_i$ and $\lim_{k \to +\infty} 2k h_j(x_k) = \mu_j$ if those limits exist.

▶ **Correction**

1. We write $I(x) = \{i | g_i(x) = 0\}$
   The conditions are
   There exists $\lambda_i \geq 0$ and $\mu_j \in \mathbb{R}$ such that $\nabla f(x^*) + \sum_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_j \nabla h_j(x^*) = 0$.

2. No, because the qualification for linear independence is satisfied. The vectors $\nabla g_i(x)$, for i such that $g_i(x) = 0$, and $\nabla h_j(x)$ are independent ; thus the values of the λ_i and μ_j are unique.
3. We know that x_k is a solution, therefore it is a critical point, so $\nabla q(x,k)(x_k) = 0$
   The gradient of $q(x,k)$ can be written as :

$$\nabla q(x,k)(x_k) = \nabla f(x_k) + k\sum_{i=1}^{m} 2g_i^+(x_k)\nabla g_i(x_k) + k\sum_{j=1}^{p} 2h_j(x_k)\nabla h_j(x_k) = 0$$

4. If we take the limit, since f, g_i, and h_j are C^1, and $\lim_{k\to+\infty} x_k = x^*$

$$\lim_{k\to+\infty} \nabla f(x_k) = \nabla f(x^*)$$
$$\lim_{k\to+\infty} \nabla g_i(x_k) = \nabla g_i(x^*)$$
$$\lim_{k\to+\infty} \nabla h_j(x_k) = \nabla h_j(x^*)$$
$$\lim_{k\to+\infty} \nabla q(x,k)(x_k) = 0$$

   Donc

$$\nabla f(x^*) + \sum_{i=1}^{m}(\lim_{k\to+\infty} 2kg_i^+(x_k))\nabla g_i(x^*) + \sum_{j=1}^{p}(\lim_{k\to+\infty} 2kh_j(x_k))\nabla h_j(x^*) = 0$$

   By continuity of the g_i, we can notice that, for any i such that $g_i(x^*) < 0$, from a certain rank onward, $g_i(x_k) < 0$ and therefore $g_i^+(x_k) = 0$.

   Thus as $\lim_{k\to+\infty} 2kg_i^+(x_k)$ and $\lim_{k\to+\infty} 2kh_j(x_k)$ exist,

$$\nabla f(x^*) + \sum_{i\in I(x^*)}(\lim_{k\to+\infty} 2kg_i^+(x_k))\nabla g_i(x^*) + \sum_{j=1}^{p}(\lim_{k\to+\infty} 2kh_j(x_k))\nabla h_j(x^*) = 0$$

   We deduce the value of these limits by the uniqueness of the coefficients of the Kuhn-Tucker equality.

Exercice 4 — *Internal barrier*

We consider the following problem

$$\min_{x\in\mathbb{R}} f(x) = x^2 \quad \text{s.c.} \begin{cases} x - 1 \leq 0 & (1) \\ -x - 1 \leq 0 & (2) \\ -x^2 \leq 0 & (3) \end{cases}$$

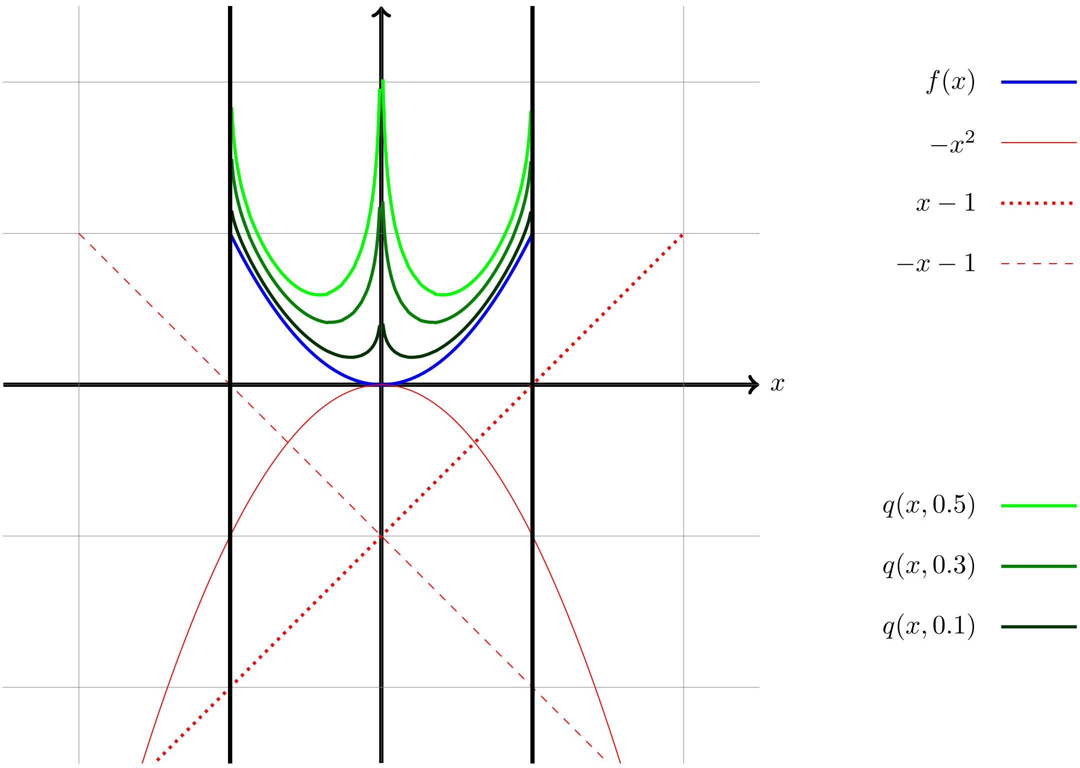
1. Draw the graphical representation of the problem and an associated logarithmical barrier. Does the problem satisfy the hypothesis that are needed to use the logarithmical and inverse barriers ?
2. Whatever the answer to the previous quesiton is, try to solve the problem using a barrier method, using a logarithmical barrier.
3. We replace the constraint $-x^2 \leq 0$ by $g_3(x) \leq 0$ with

$$g_3(x) = \text{ s.c.} \begin{cases} -(x+\frac{1}{2})^2 \text{ if } x \leq \frac{-1}{2} \\ -(x-\frac{1}{2})^2 \text{ if } x \geq \frac{1}{2} \\ 0 \text{ otherwise} \end{cases}$$

   Draw the graphical representation of the problem and an associated barrier.
4. Can we solve the problem with the barrier method ? How could we get around that problem ?

▶ **Correction**

1. Detail, the constraint (3) is trivial. This can be easily seen from the diagram. q is, as in the lecture, the function $f(x) + \mu B(x)$ where B is a barrier (here logarithmic).



The problem does not satisfy the hypothesis $g_3(x) < 0$ if and only if $x \in \mathring{S}$, indeed, $S = [-1, 1]$ thus $\mathring{S} =]-1, 1[$ and $g_3(0) = 0$. Therefore, we should not use logarithmic and inverse barriers (according to the course).
However, we can see from the drawing that the minimum of the function $q(x, \mu)$ seems to approach 0 as μ approaches 0.

2. Let us apply the logarithmic barrier nonetheless.
Let $B(x) = -\log(-x+1) - \log(x+1) - \log(-x^2)$. We aim to minimize $q(x, \mu) = f(x) + \mu B(x)$ given that $g(x) < 0$.
An optimal solution x_μ to this problem is a critical point and satisfies

$$\nabla(q(x_\mu, \mu)) = 0$$
$$f'(x_\mu) + \mu \nabla B'(x_\mu) = 0$$
$$2x_\mu + \mu \frac{1}{-x_\mu + 1} + \mu \frac{-1}{x_\mu + 1} + \mu \frac{2}{-x_\mu} = 0$$

We multiply by the denomiators and simplify :

$$x_\mu^4 + (-1 - 2\mu)x_\mu^2 + \mu = 0$$

we set $X = x_\mu^2$

$$X^2 + (-1 - 2\mu)X + \mu = 0$$

We solve the system

$$X = \frac{1}{2}((2\mu+1) \pm \sqrt{4\mu^2+1}) > 0$$
$$x_\mu = \pm\sqrt{\frac{1}{2}((2\mu+1) \pm \sqrt{4\mu^2+1})}$$

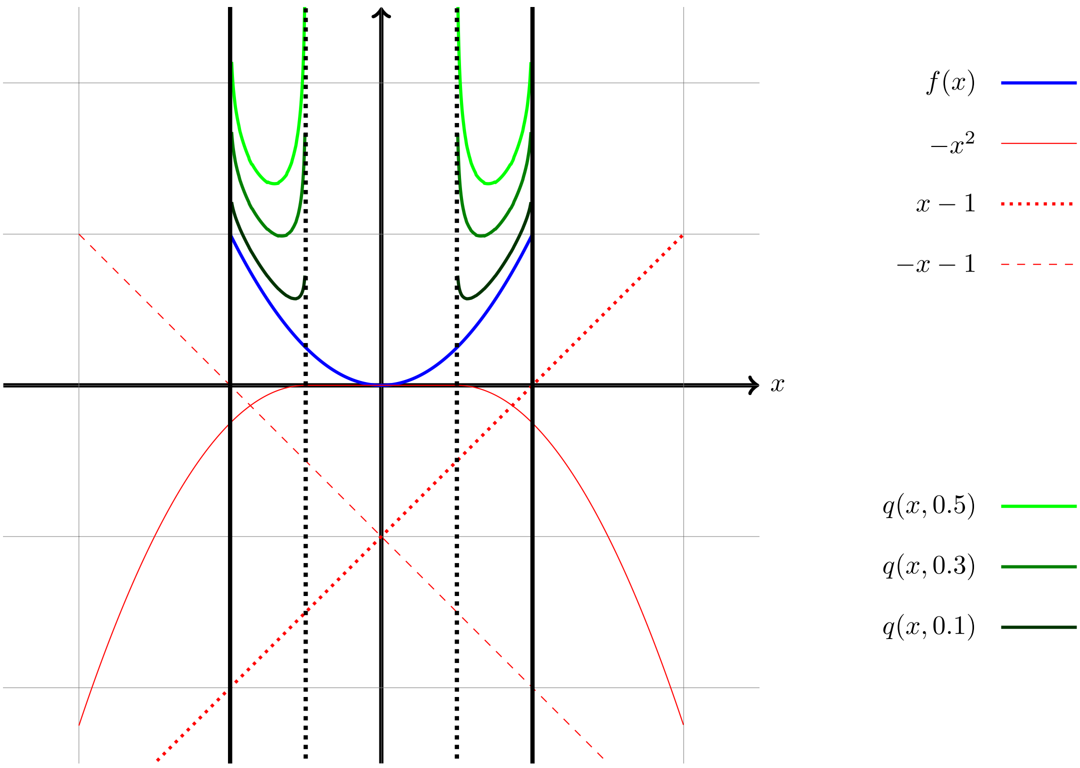
We check that $g_1(x_\mu) < 0$, $g_2(x_\mu) < 0$, and $g_3(x_\mu) < 0$ for all $\mu > 0$.

These points are the only critical points of the function ; moreover, the function q is convex, so one of them is an optimal solution. In this case, it is $\sqrt{\frac{1}{2}((2\mu+1) - \sqrt{4\mu^2+1})}$

$$\lim_{\mu \to 0} x_\mu = 0.$$

Each of the sequences of x_μ converges, so it converges to an optimal solution of $f : x = 0$.

3. We get the following drawing



4. If we apply the method, we see that we will tend towards $x = \pm\frac{1}{2}$ instead of 0. We could solve this problem by relaxing the constraints a bit :

   We solve $\min\limits_{x\in\mathbb{R}} f(x) = x^2$ s.c. $\begin{cases} g_1(x) - \varepsilon \leq 0 & (1) \\ g_2(x) - \varepsilon \leq 0 & (2) \\ g_3(x) - \varepsilon \leq 0 & (3) \end{cases}$

   We then let ε tend towards 0. The method has the drawback of possibly using the set of infeasible solutions, unlike the classical barrier method.