TD 8 : Méthodes des pénalités et des barrières

Recherche opérationnelle S3.

2024

Exercice 1 — *Pénalité de Beltrami*

On souhaite minimiser $f(x) = x_1 + x_2$ sous contrainte $x_1^2 - x_2 \leq 2$.

1. Résoudre le problème avec la méthode des pénalités de Beltrami.
2. Même question en se ramenant à un problème d'égalité : minimiser $f(x) = x_1 + x_2$ sous contrainte $x_1^2 - x_2 + x_3^2 = 2$.

▶ **Correction**

1. On pose $g(x) = x_1^2 - x_2 - 2$ et $P(x) = g^+(x)^2 = \max(g(x), 0)^2$; on cherche, pour tout μ à minimiser $q(x, \mu) = f(x) + \mu \cdot P(x)$.
   Une solution optimale x_μ de ce problème est un point critique et vérifie

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

On considère alors 2 cas : $g(x_\mu) \leq 0$ et $g(x_\mu) > 0$

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

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

Ce qui est exclus. Considérons l'autre cas, $g(x_\mu) > 0$

$$\begin{aligned}\nabla f(x_\mu) + 2\mu g(x_\mu)\nabla g(x_\mu) &= 0\\ \begin{pmatrix}1\\1\end{pmatrix} + 2\mu(x_{\mu1}^2 - x_{\mu2} - 2)\begin{pmatrix}2x_{\mu1}\\-1\end{pmatrix} &= 0\\ 1 + 2\mu(x_{\mu1}^2 - x_{\mu2} - 2)(2x_{\mu1}) &= 0 \text{ et } (x_{\mu1}^2 - x_{\mu2} - 2) = \frac{1}{2\mu}\\ x_1 &= -\frac{1}{2} \text{ et } x_2 = -\frac{7}{4} - \frac{1}{2\mu}\end{aligned}$$

On a bien $g(x) \geq 0$ car $x_1^2 - x_2 = 2 + \frac{1}{2k} \geq 2$. Enfin :

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

Ce point est le seul point critique de la fonction ; de plus la fonction q ne tend pas vers $-\infty$ en l'infini , il est donc solution optimale.
La suite des x_μ converge donc elle converge vers une solution optimale de f.

2. On pose $h(x) = x_1^2 - x_2 + x_3^2 - 2$ et $P(x) = h(x)^2$; on cherche, pour tout μ à minimiser $q(x,\mu) = f(x) + \mu \cdot P(x)$.
Une solution optimale x_μ de ce problème est un point critique et vérifie

$$\begin{aligned}
\nabla(q(x_\mu,\mu)) &= 0\\
\nabla f(x_\mu) + \mu \nabla P(x_\mu) &= 0\\
\nabla f(x_\mu) + 2\mu h(x_\mu)\nabla h(x_\mu) &= 0\\
\begin{pmatrix}1\\1\\0\end{pmatrix} + 2\mu(x_{\mu 1}^2 - x_{\mu 2} + x_{\mu 3}^2 - 2)\begin{pmatrix}2x_{\mu 1}\\-1\\2x_{\mu 3}\end{pmatrix} &= 0
\end{aligned}$$

La 3e égalité implique que $x_{\mu 3} = 0$ ou $(x_{\mu 1}^2 - x_{\mu 2} + x_{\mu 3}^2 - 2) = 0$.

Si $(x_{\mu 1}^2 - x_{\mu 2} + x_{\mu 3}^2 - 2) = 0$ alors les 2 premières égalités donnent $1 = 0$, ce qui est exclu.

Si $x_{\mu 3} = 0$ alors on retombe sur

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

Et on a donc bien la même solution qu'à la première question.

Exercice 2 — *Pénalité des barrière*

On souhaite minimiser $f(x) = x_1 + x_2$ sous contrainte $x_1^2 - x_2 \leq 2$. Résoudre le problème avec la méthode des barrières.

▶ Correction

On va utiliser les 2 barrières : $B(x) = \frac{-1}{g(x)}$ et $B(x) = -\log(-g(x))$.

Posons $g(x) = x_1^2 - x_2 - 2$ et $B(x) = \frac{-1}{g(x)}$. On cherche à minimiser $q(x,\mu) = f(x) + \mu B(x)$ sachant que $g(x) < 0$.

Une solution optimale x_μ de ce problème est un point critique et vérifie

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

La 2e équation implique que $\frac{\mu}{(x_{\mu 1}^2 - x_{\mu 2} - 2)^2} = 1$, donc

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

Et puisque $g(x_\mu) < 0$, on ne peut avoir $x_{\mu 2} = -\frac{7}{4} - \sqrt{\mu}$

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

Ce point est le seul point critique de la fonction ; de plus la fonction q est convexe, il est donc solution optimale.

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

La suite des x_μ converge donc elle converge vers une solution optimale de f.

Utilisons maintenant l'autre barrière :

Posons $g(x) = x_1^2 - x_2 - 2$ et $B(x) = -\log(-g(x))$. On cherche à minimiser $q(x,\mu) = f(x) + \mu B(x)$ sachant que $g(x) < 0$.

Une solution optimale x_μ de ce problème est un point critique et vérifie

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

La 2e équation implique que $\frac{-\mu}{(x_{\mu 1}^2 - x_{\mu 2} - 2)} = 1$, donc

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

On a bien $g(x) = -\mu < 0$

Ce point est le seul point critique de la fonction ; de plus la fonction q est convexe, il est donc solution optimale.

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

La suite des x_μ converge donc elle converge vers une solution optimale de f.

Exercice 3 — *Pénalité et multiplicateurs de Lagrange*

Soit un problème générique, on souhaite minimiser $f(x)$ sur $\mathbb{R}^n$ tel que $g_i(x) \leq 0$ pour tout $i \in [\![1;m]\!]$ et tel que $h_j(x) = 0$ pour tout $j \in [\![1;p]\!]$. On applique la méthode des pénalités avec la Pénalité P de Beltrami. On note $q(x,\mu) = f(x) + \mu P(x)$ et $x_k = \arg\min q(x,k)$.

On suppose que f, les g_i et h_j sont C^1. Soit x^* une solution optimale de f, on suppose que x^* vérifie le critère de qualification de l'indépendance linéaire. Enfin, on suppose que la suite x_k converge vers x^*.

1. Rappelez les conditions de Kuhn-Tucker pour ce problème en x^*, on notera λ_i et μ_j les coefficients multiplicateurs de Lagrange associés aux fonctions g_i et h_j.
2. Est-il possible qu'il existe plusieurs λ_i, μ_j vérifiant ces conditions ?
3. Écrire le gradient de $q(x,k)$ en x_k. À quelle valeur est égal ce gradient ?
4. En déduire que $\lim\limits_{k\to+\infty} 2kg_i^+(x_k) = \lambda_i$ et $\lim\limits_{k\to+\infty} 2kh_j(x_k) = \mu_i$, si ces limites existent.

▶ Correction

1. On note $I(x) = \{i | g_i(x) = 0\}$
   Les conditions sont :
   Il existe $\lambda_i \geq 0$ et $\mu_j \in \mathbb{R}$ tels que $\nabla f(x^*) + \sum\limits_{i\in I(x^*)} \lambda_i \nabla g_i(x^*) + \sum\limits_{i=1}^{p} \mu_j \nabla h_j(x^*) = 0$.

2. Non car la qualification de l'indépendance linéaire est vérifiée. Les vecteurs $\nabla g_i(x)$, pour i tel que $g_i(x) = 0$ et $\nabla h_j(x)$ sont indépendants ; donc les valeurs des λ_i et des μ_j sont uniques.
3. On sait que x_k est solution, il est donc point critique, donc $\nabla q(x,k)(x_k) = 0$
   Le gradient de $q(x,k)$ s'écrit :
   $$\nabla q(x,k)(x_k) = \nabla f(x_k) + k\sum_{i=1}^{m} 2g_i^+(x_k)\nabla g_i(x_k) + k\sum_{j=1}^{p} 2h_j(x_k)\nabla h_j(x_k) = 0$$
4. Si on passe à la limite, puisque f, g_i et h_j sont C^1, et que $\lim_{k\to+\infty} x_k = x^*$
   $$\lim_{k\to+\infty} \nabla f(x_k) = \nabla f(x^*)$$
   $$\lim_{k\to+\infty} \nabla g_i(x_k) = \nabla g_i(x^*)$$
   $$\lim_{k\to+\infty} \nabla h_j(x_k) = \nabla h_j(x^*)$$
   $$\lim_{k\to+\infty} \nabla q(x,k)(x_k) = 0$$
   Donc
   $$\nabla f(x^*) + \sum_{i=1}^{m}(\lim_{k\to+\infty} 2kg_i^+(x_k))\nabla g_i(x^*) + \sum_{j=1}^{p}(\lim_{k\to+\infty} 2kh_j(x_k))\nabla h_j(x^*) = 0$$
   Par continuité des g_i, on peut remarquer que, pour tout i tel que $g_i(x^*) < 0$, à partir d'un certain rang, $g_i(x_k) < 0$ et donc $g_i^+(x_k) = 0$.

   Donc, puisque $\lim_{k\to+\infty} 2kg_i^+(x_k)$ et $\lim_{k\to+\infty} 2kh_j(x_k)$ existent,
   $$\nabla f(x^*) + \sum_{i\in I(x^*)}(\lim_{k\to+\infty} 2kg_i^+(x_k))\nabla g_i(x^*) + \sum_{j=1}^{p}(\lim_{k\to+\infty} 2kh_j(x_k))\nabla h_j(x^*) = 0$$
   On déduit la valeur de ces limites par unicité des coefficients de l'égalité de Kuhn Tucker.

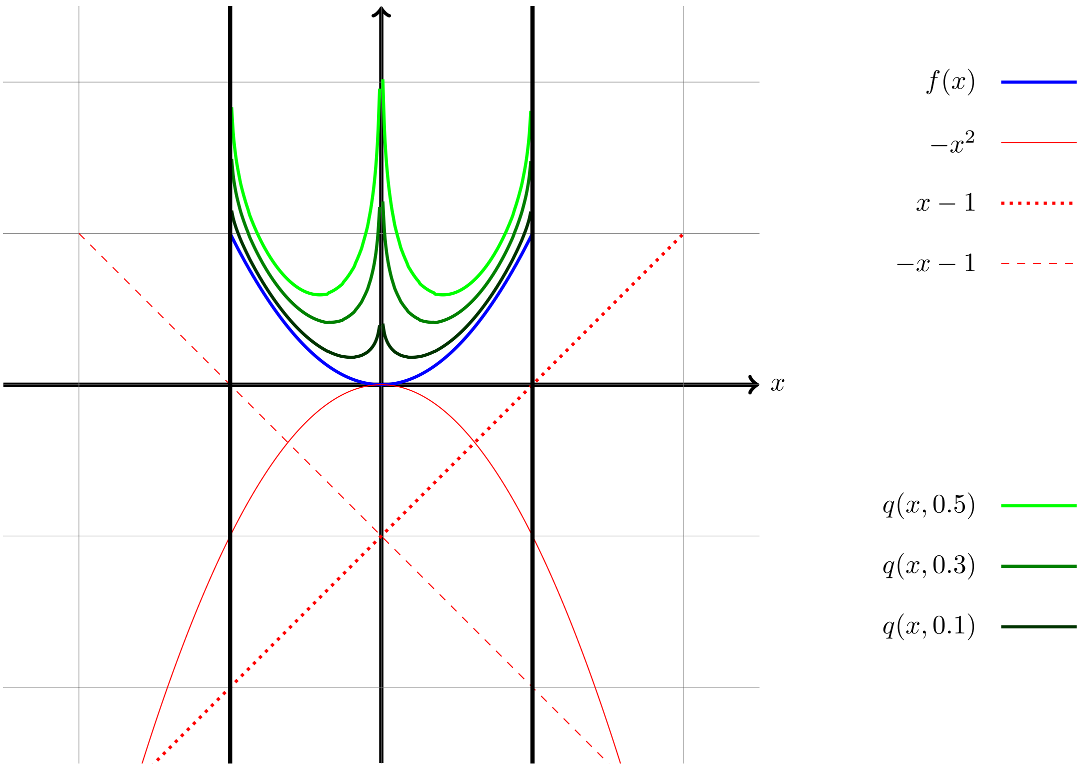
Exercice 4 — *Barrière interne*

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

1. Représenter ce problème et une barrière logarithmique associée graphiquement. Est-ce que ce problème satisfait les hypothèses nécessaires à l'utilisation des barrières logarithmique et inverse ?
2. Quelle que soit la réponse à la question précédente, tenter de résoudre ce problème en utilisant la méthode de la barrière avec une barrière logarithmique.
3. On remplace la contrainte $-x^2 \leq 0$ par $g_1(x) \leq 0$ avec
   $$g_3(x) = \text{ s.c. } \begin{cases} -(x+\frac{1}{2})^2 \text{ si } x \leq \frac{-1}{2} \\ -(x-\frac{1}{2})^2 \text{ si } x \geq \frac{1}{2} \\ 0 \text{ sinon} \end{cases}$$
   Représenter ce problème et un barrière associée graphiquement.
4. Peut-on résoudre ce problème avec la méthode des barrières ? Comment pourrait-on contourner ce problème ?

▶ **Correction**

1. Détail, la contrainte (3) est triviale. On peut le voir aisément sur le dessin. q est, comme dans le cours, la fonction $f(x) + \mu B(x)$ où B est une barrière (ici logarithmique).



Le problème ne satisfait pas l'hypothèse $g_3(x) < 0$ ssi $x \in \mathring{S}$, en effet, $S = [-1, 1]$ donc $\mathring{S} =]-1, 1[$ et $g_3(0) = 0$. Donc, on ne devrait pas utiliser les barrières logarithmiques et inverses (d'après le cours).

Cependant, on voit sur le dessin que le minimum des fonction $q(x, \mu)$ semble tendre vers 0 quand μ tend vers 0.

2. Appliquons malgré tout la barrière logarithmique.

Posons $B(x) = -\log(-x + 1) - \log(x + 1) - \log(-x^2)$. On cherche à minimiser $q(x, \mu) = f(x) + \mu B(x)$ sachant que $g(x) < 0$.

Une solution optimale x_μ de ce problème est un point critique et vérifie

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

On multiplie par les dénominateurs, on simplifie et on trouve

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

on pose $X = x_\mu^2$

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

On résout le système :

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

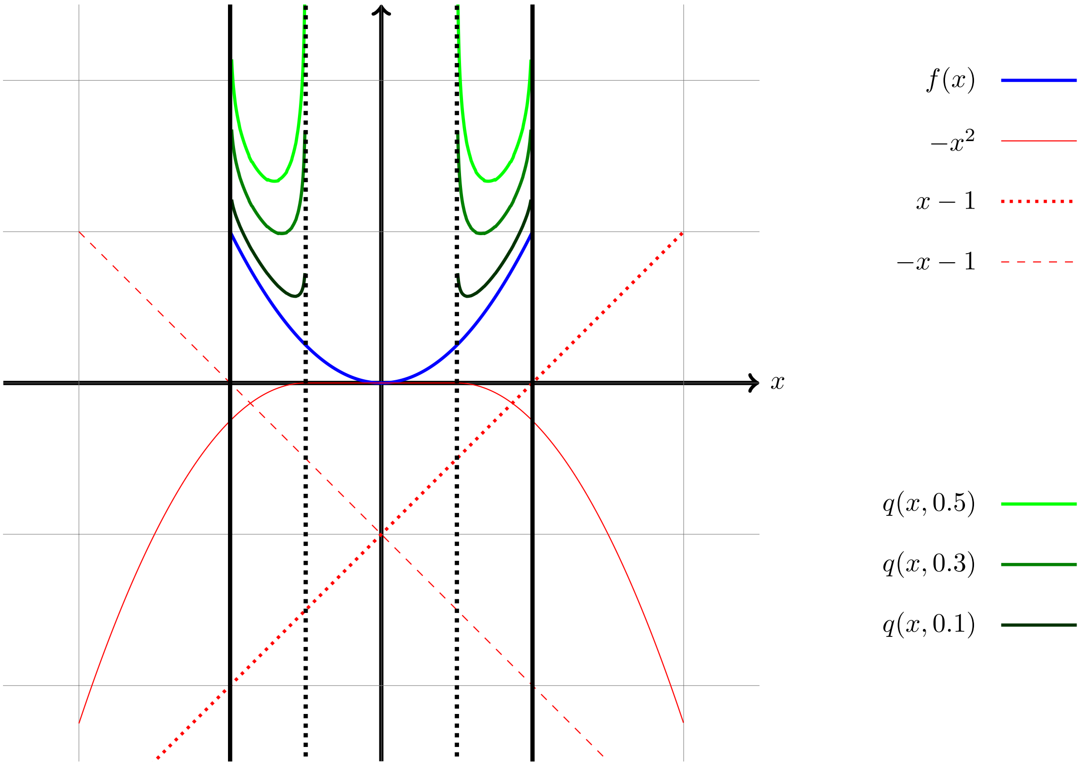
On vérifie bien $g_1(x_\mu) < 0$, $g_2(x_\mu) < 0$ et $g_3(x_\mu) < 0$ pour tout $\mu > 0$.

Ces points sont les seuls points critiques de la fonction ; de plus la fonction q est convexe, l'un d'eux est donc solution optimale. En l'occurence il s'agit de $\sqrt{\frac{1}{2}((2\mu+1) - \sqrt{4\mu^2+1})}$

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

Chacune des suites des x_μ converge donc elle converge vers une solution optimale de f : $x = 0$.

3. On obtient le dessin suivant



4. Si on applique la méthode, on voit qu'on va tendre vers $x = \pm\frac{1}{2}$ au lieu de 0. On pourrait régler ce problème en relâchant un peu les contraintes :

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

   On fait ensuite tendre ε vers 0. La méthode a le défaut de possiblement se balader dans l'ensemble des solutions non réalisables ; contrairement à la méthode des barrières classique.