TD 10 : Gradient réduit et méthode de pénalité

lionel.rieg@ensiie.fr

1 Gradient réduit

Exercice 1

Résoudre le problème suivant avec la méthode du gradient réduit :

$$\min x_1^2 + 4x_2^2 \qquad \text{s.c.} \qquad \begin{cases} x_1 + 2x_2 \geq 1 \\ -x_1 + x_2 \leq 0 \\ x_1 \geq 0, x_2 \geq 0 \end{cases}$$

On partira du point $P_0 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}$ et on prendra la décomposition $B = \{1, 3\}$ et $N = \{2, 4\}$.

2 Méthode de pénalité

Exercice 2

On considère le problème (P) suivant :

$$\min f(x) = 3x_1^2 + 2x_2^2 \qquad \text{s.c.} \qquad x_1 + x_2 \geq 1 \tag{P}$$

1. Résoudre (P) par la méthode de pénalité de Beltrami en prenant $\mu_k = k$.
2. Retrouver le multiplicateur de Lagrange des conditions de Kuhn-Tucker.

Solutions

▶ Exercice 1

On commence par réécrire les contraintes avec des égalités en introduisant des variables supplémentaires :

$$\begin{cases} x_1 + 2x_2 - x_3 = 1 \\ -x_1 + x_2 + x_4 = 0 \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0 \end{cases} \qquad \text{ce qui donne} \quad A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} \quad \text{et} \quad \nabla f(x) = \begin{pmatrix} 2x_1 \\ 8x_2 \\ 0 \\ 0 \end{pmatrix}$$

Première itération : L'énoncé nous impose $B = \{1, 3\}$ et $N = \{2, 4\}$, d'où l'on tire $A_B = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}$ et $A_N = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$. On vérifie que A_B est inversible car son déterminant est -1. Calculons A_B^{-1} et $-A_B^{-1}A_N$ pour en déduire le gradient réduit de f en P_0 : $r = \frac{\partial \bar{f}}{\partial x_N} = -\frac{\partial f}{\partial x_b}(P_0)A_b^{-1}A_N + \frac{\partial f}{\partial x_N}(P_0)$:

$$A_B^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix} \qquad -A_B^{-1}A_N = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} \qquad r = \begin{pmatrix} 2 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 8 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 0 \end{pmatrix}$$

On en déduit alors $d_N = \begin{pmatrix} -10 \\ 0 \end{pmatrix}$ et puisqu'il est non nul, $d_B = -A_B^{-1}A_N d_N = \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} -10 \\ 0 \end{pmatrix} = \begin{pmatrix} -10 \\ -30 \end{pmatrix}$. La direction de descente en P_0 est donc $d_0 = \begin{pmatrix} -10 \\ -10 \\ -30 \\ 0 \end{pmatrix}$ et le nouveau point P_1 est $P_1 = P_0 + \alpha d_0$.

On commence par calculer la borne supérieure de α en imposant que P_1 doit être réalisable, c'est-à-dire que ses coordonnées sont positives ou nulles.

$$\begin{cases} 1 - 10\alpha \geq 0 \\ 1 - 10\alpha \geq 0 \\ 2 - 30\alpha \geq 0 \\ 0 \geq 0 \end{cases} \iff \begin{cases} \alpha \leq \frac{1}{10} \\ \alpha \leq \frac{1}{15} \end{cases} \iff \alpha \leq \frac{1}{15}$$

Pour déterminer P_2, on minimise $f(P_2)$ qui est une fonction de α. On a $f(P_2)(\alpha) = (1 - 10\alpha)^2 + 4(1 - 10\alpha)^2 = 5(1 - 10\alpha)^2$ qui atteint son minimum lorsque $1 - 10\alpha = 0$, donc en $\frac{1}{10}$. Il est hors de l'intervalle autorisé donc on prend la borne de l'intervalle : $\alpha = \frac{1}{15}$. On a donc $P_1 = P_0 + \frac{1}{15}d_0 = \begin{pmatrix} 1 - 10\frac{1}{15} \\ 1 - 10\frac{1}{15} \\ 2 - 30\frac{1}{15} \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 0 \\ 0 \end{pmatrix}$.

On remarque que la troisième composante de P_1 est nulle alors que $3 \in B$, on doit donc effectuer un changement de base. on échange 3 avec l'indice de la plus grande coordonnée dans N, ici la seconde qui vaut $1/3$.

Deuxième itération : On a $B = \{1, 2\}$ et $N = \{3, 4\}$ donc $A_B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$ et $A_N = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Refaisons les calculs comme à l'étape précédente :

$$A_B^{-1} = \frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} \qquad -A_B^{-1}A_N = \frac{1}{3}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \qquad r = \begin{pmatrix} 2/3 & 8/3 \end{pmatrix}\frac{1}{3}\begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} = \begin{pmatrix} 10/9 & -4/9 \end{pmatrix}$$

On en déduit $d_N = \begin{pmatrix} 0 \\ 4/9 \end{pmatrix}$ puis $d_B = -A_B^{-1} A_N d_N = \frac{1}{3} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 4/9 \end{pmatrix} = \begin{pmatrix} 8/27 \\ -4/27 \end{pmatrix}$, d'où $d_1 = \begin{pmatrix} 8/27 \\ -4/27 \\ 0 \\ 4/9 \end{pmatrix}$.

Les contraintes du problèmes en $P_2 = P_1 + \alpha d_1$ sont :

$$\begin{cases} 1/3 + 8/27\ \alpha \geq 0 \\ 1/3 - 4/27\ \alpha \geq 0 \\ 0 \geq 0 \\ 4/9\ \alpha \geq 0 \end{cases} \iff \begin{cases} \alpha \geq -9/8 \\ \alpha \leq 9/4 \\ \alpha \geq 0 \end{cases} \iff 0 \leq \alpha \leq 9/4$$

Pour calculer le minimum de $f(P_2)$ comme fonction de α, on va utiliser la dérivée directionnelle (on pourrait tout aussi bien faire comme précédemment et étudier $f(P_2)$ comme une fonction réelle) : le point P_2 est le minimum de f dans la direction d_1 donc sa dérivée en P_2 selon la direction d_1 doit être nulle : $\nabla f(P_2).d_1 = 0$.

$$\nabla f(P_2).d_1 = 0 \iff \begin{pmatrix} 2(1/3 + \alpha\, 8/27) \\ 8(1/3 - \alpha\, 4/27) \\ 0 \\ 0 \end{pmatrix} . \frac{4}{27} \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1/3 \end{pmatrix} = \frac{\cancel{4}}{\cancel{27}} \left(4\Big(\frac{1}{3} + \alpha\,\frac{8}{27}\Big) - 8\Big(\frac{1}{3} - \alpha\,\frac{4}{27}\Big) \right) = 0 \iff \frac{4}{3} + \frac{64}{27}\alpha = 0$$

On obtient finalement $\alpha = 9/16$ qui est bien dans l'intervalle autorisé et on a donc $P_2 = \begin{pmatrix} 1/2 \\ 1/4 \\ 0 \\ 1/4 \end{pmatrix}$.

Aucune des composantes d'indice dans B n'est nulle, il n'y a donc pas besoin de faire de changement de base.

Troisième itération : On a toujours $B = \{1, 2\}$ et $N = \{3, 4\}$ donc on peut reprendre le calcul de $-A_B^{-1} A_N$. Calculons r : $r = \begin{pmatrix} 1 & 2 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ d'où on déduit $d_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Puisqu'il est nul, on a atteint l'optimum.

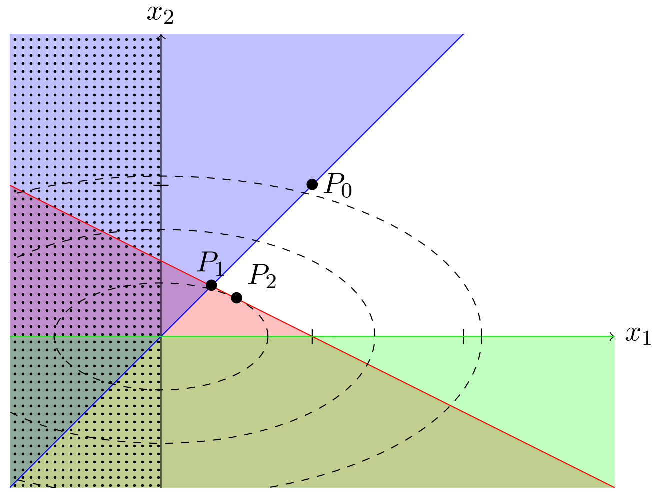
Vérification : On peut vérifier que le point trouvé vérifie les conditions de Kuhn-Tucker, à savoir :

$$\begin{cases} \nabla f(x) + \mu A - \lambda = 0 & \\ \lambda \geq 0 & \text{(positivité)} \\ \lambda_i x_i = 0 & \text{(complémentarité)} \end{cases}$$

Comme dans le cours, on pose $\mu = -\frac{\partial f}{\partial x_B}(x) A_B^{-1}$, $\lambda_B = 0$ et $\lambda_N = r$.
Ici, cela donne $\mu = -\begin{pmatrix} 1 & 2 \end{pmatrix} \frac{1}{3} \begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \end{pmatrix}$, $\lambda_1 = 0$, $\lambda_2 = 0$, $\lambda_3 = 1$ et $\lambda_4 = 0$. La positivité est évidente. Seul λ_3 est non nul mais on a $x_3 = 0$ donc les conditions de complémentarité sont vérifiées.

$$\begin{aligned} \nabla f(x) + \mu A - \lambda &= \begin{pmatrix} 1 & 2 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 0 & 0 \end{pmatrix} + \begin{pmatrix} -1 & -2 & 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Représentation graphique dans le plan (x_1, x_2) **:**



▶ Exercice 2

1. La pénalité de Beltrami est donnée par $P(k) = \sum (g_i^+(x))^2$ où $g^+(x) = \max(0, g(x))$. Ici, cela donne $P(x) = (g^+(x))^2 = (\max(0, -x_1 - x_2 - 1))^2$ et la suite de problèmes avec pénalités (donc sans contraintes) est donc $(P_k) \quad \min f(x) + \mu_k P(x)$, c'est-à-dire $\min f_k(x) = 3x_1^2 + 2x_2^2 + k(g^+(x))^2$ puisqu'on prend $\mu_k = k$.
   Le gradient de (P_k) est :

$$\nabla f(x) + 2kg^+(x)\nabla g(x) = \begin{pmatrix} 6x_1 \\ 4x_2 \end{pmatrix} + 2kg^+(x) \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

   On considère deux cas, suivant le signe de $g(x)$:

   Cas $g(x) < 0$ **:** On a donc $g^+(x) = 0$ et $\nabla f_k = \nabla f$. La condition $\nabla f_k(x) = 0$ donne $x_1 = x_2 = 0$. En revanche, on a $g(\begin{pmatrix} 0 \\ 0 \end{pmatrix}) = 1$ ce qui est incompatible avec l'hypothèse $g(x) < 0$.
   Il n'y a donc pas de solution dans ce cas.

   Cas $g(x) \geq 0$ **:** On a donc $g^+(x) = g(x)$ et $\nabla f_k(x) = \nabla f(x) + 2kg(x)\nabla g(x)$. La condition $\nabla f_k(x) = 0$ donne

$$\begin{pmatrix} 6x_1 \\ 4x_2 \end{pmatrix} + 2k(-x_1 - x_2 - 1) \begin{pmatrix} -1 \\ -1 \end{pmatrix} = 0 \iff \begin{cases} 6x_1 + 2k(x_1 + x_2 + 1) = 0 \\ 4x_2 + 2k(x_1 + x_2 + 1) = 0 \end{cases} \iff \begin{cases} 6x_1 + 2k(x_1 + \frac{3}{2}x_1 + 1) = 0 \\ 6x_1 = 4x_2 \iff x_2 = \frac{3}{2}x_1 \end{cases}$$

   On obtient finalement $x_1 = \frac{2k}{5k+6}$ et $x_2 = \frac{3k}{5k+6}$. La solution x_k de (P_k) est donc $\begin{pmatrix} \frac{2k}{5k+6} \\ \frac{3k}{5k+6} \end{pmatrix}$. La limite de cette suite est $x = \begin{pmatrix} 2/5 \\ 3/5 \end{pmatrix}$ qui est donc solution du problème (P) initial.

2. Si (x_k) est une suite qui tend vers une solution de (P), les multiplicateurs de Lagrange sont donnés par $2\mu_k g_i^+(x_k)$ qui tend vers λ_i.
   Dans notre cas, cela donne :

$$2kg^+(x_k) = 2kg(x_k) = 2k\left(1 - \frac{2k}{5k+6} - \frac{3k}{5k+6}\right) = 2k\left(\frac{5k + 6 - 2k - 3k}{5k+6}\right) = \frac{12k}{5k+6} \to \frac{12}{5}$$