# TD 7 : Algorithme du gradient réduit

Recherche opérationnelle S3.

2024

**Exercice 1 — *Exemple simple.***

Soit $(P)$ le problème suivant :

$$\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + 4x_2^2 \quad \text{s.c.} \quad \begin{cases} x_1 + 2x_2 \geq 1 \\ -x_1 + x_2 \leq 0 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

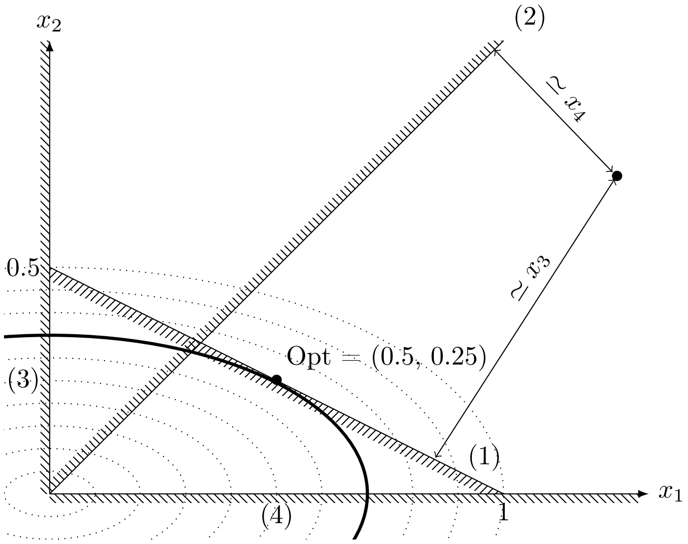
1. Écrire $(P)$ sous forme standard avec deux variables $x_3$ et $x_4$.
2. Représenter graphiquement le problème, include $x_3$ et $x_4$ sur le dessin.
3. Appliquer l'algorithme à partir de $P_0 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}$ et la base $\mathcal{B} = \{1, 3\}$ (3 itérations).
4. Vérifier que le point trouvé satisfait les conditions de Karush-Khun-Tucker.

▶ **Correction**

1. $\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + 4x_2^2 \quad \text{s.c.} \quad \begin{cases} x_1 + 2x_2 - x_3 = 1 \\ -x_1 + x_2 + x_4 = 0 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$
2. L'équation d'une ellipse de demi-grand axes $a$ et $b$ centrée en l'origine est $x^2/a^2 + y^2/b^2 = 1$. Donc $x_1^2 + 4x_2^2 = R$ est l'équation d'une ellipse telle que :

$$\begin{cases} \frac{\sqrt{R}}{a^2} & = 1 \\ \frac{\sqrt{R}}{b^2} & = 4 \end{cases}$$

Donc $a = 2b = \sqrt{R}$.
On trace donc des ellipses autours de l'origine de plus en plus petite avec $a = 2b$ jusqu'à sortir de l'ensemble des solutions réalisables ; le point où on sort indique la solution optimale.




3. On part du point $P_0 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 0 \end{pmatrix}$ et de la base $B = \{1, 3\}$, donc $N = \{2, 4\}$.

   La matrice $A$ des coefficients est $\begin{pmatrix} 1 & 2 & -1 & 0 \\ -1 & 1 & 0 & 1 \end{pmatrix}$ ; elle est bien de rang 2.

   Le gradient est $\nabla f(P_0) = \begin{pmatrix} 2 \\ 8 \\ 0 \\ 0 \end{pmatrix}$

   On calcule $A_B$, $A_N$, $x_B$, $x_N$, $\nabla f_B$ et $\nabla f_N$ :
   $A_B = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}, \quad A_N = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$
   $x_B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad x_N = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
   $\nabla f_B = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \nabla f_N = \begin{pmatrix} 8 \\ 0 \end{pmatrix}$
   La matrice $A_B$ est inversible et $A_B^{-1} = \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}$.
   On peut donc calculer le gradient réduit :

$$\begin{aligned} {}^t\nabla \bar{f}(x_N) &= -{}^t\nabla f_B \cdot A_B^-1 \cdot A_N + {}^t\nabla f_N \\ &= \begin{pmatrix} 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 3 & 1 \end{pmatrix} + \begin{pmatrix} 8 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 10 & 2 \end{pmatrix} \end{aligned}$$

   Autre méthode : calculer explicitement $\bar{f}(x_N)$.

$$\begin{aligned} f(x) &= x_1^2 + 4x_2^2 \\ \bar{f}(x_N) &= (x_2 + x_4)^2 + 4x_2^2 \\ &= 5x_2^2 + x_4^2 + 2x_2x_4 \end{aligned}$$

   Donc $\nabla \bar{f}(x_N) = \begin{pmatrix} 10x_2 + 2x_4 \\ 2x_4 + 2x_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 2 \end{pmatrix}$

On calcule maintenant la direction $d_N = \begin{pmatrix} -10 \\ 0 \end{pmatrix} \begin{matrix} \leftarrow \nabla \bar{f}(x_N)_{x_2} > 0 \text{ et } x_2 \neq 0 \\ \leftarrow \nabla \bar{f}(x_N)_{x_4} > 0 \text{ et } x_4 = 0 \end{matrix}$

On en déduits $d_B = -A_B^{-1} A_N d_N = \begin{pmatrix} -10 \\ -30 \end{pmatrix} \begin{matrix} (x_1) \\ (x_3) \end{matrix}$

La direction est non nulle, on cherche donc $\alpha_1 =$ le $\alpha \geq 0$ maximum tel que $P_0 + \alpha \cdot d \geq 0$.

On a donc $\begin{cases} 1 - 10\alpha & \geq 0 \\ 1 - 10\alpha & \geq 0 \\ 2 - 30\alpha & \geq 0 \\ 0 - 0\alpha & \geq 0 \end{cases}$

Donc $\alpha_1 = \frac{1}{15}$.

On cherche maintenant le $\alpha$ entre 0 et $\alpha_1$ minimisant $f(P_0 + \alpha d)$.

$$\begin{aligned} g(\alpha) = f(P_0 + \alpha d) &= (1 - 10\alpha)^2 + 4(1 - 10\alpha)^2 \\ &= 5(1 - 10\alpha)^2 \\ g'(\alpha) &= -100(1 - 10\alpha) \\ g'(\alpha) &\leq 0 \\ \Leftrightarrow \alpha &\leq \frac{1}{10} \end{aligned}$$

La fonction $g$ est donc décroissante entre 0 et $\alpha_1 \leq \frac{1}{10}$, on pose $\alpha_2 = \alpha_1 = \frac{1}{15}$.

On se déplace donc vers $P_1 = P_0 + \frac{1}{15} d = \begin{pmatrix} 1/3 \\ 1/3 \\ 0 \\ 0 \end{pmatrix}$

On a $x_3$ qui s'annule. Il faut donc changer de base. On essaie de remplacer $x_3$ par $x_2$ dans la base (puisque $x_2 \geq x_4$).

Soit $B' = \{1, 2\}$, on a $A_{B'} = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$ de déterminant 3, donc inversible. On peut donc effectuer ce changement de base.

On a donc maintenant $B = \{1, 2\}$ et $N = 3, 4$.

Le gradient est $\nabla f(P_0) = \begin{pmatrix} 2/3 \\ 8/3 \\ 0 \\ 0 \end{pmatrix}$

On calcule $A_B$, $A_N$, $x_B$, $x_N$, $\nabla f_B$ et $\nabla f_N$ :

$A_B = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}, \quad A_N = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$

$x_B = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}, \quad x_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\nabla f_B = \begin{pmatrix} 2/3 \\ 8/3 \end{pmatrix}, \quad \nabla f_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

La matrice $A_B$ est inversible et $A_B^{-1} = \begin{pmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{pmatrix}$.

On peut donc calculer le gradient réduit :

$$\begin{aligned} {}^t\nabla \bar{f}(x_N) &= -\, {}^t\nabla f_B \cdot A_B^-1 \cdot A_N + {}^t\nabla f_N \\ &= \begin{pmatrix} 2/3 & 8/3 \end{pmatrix} \cdot \begin{pmatrix} -1/3 & -1/3 \\ -2/3 & 1/3 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 10/9 & -4/9 \end{pmatrix} \end{aligned}$$

On calcule maintenant la direction $d_N = \begin{pmatrix} 0 \\ 4/9 \end{pmatrix} \begin{matrix} \leftarrow \nabla \bar{f}(x_N)_{x_3} > 0 \text{ et } x_3 = 0 \\ \leftarrow \nabla \bar{f}(x_N)_{x_4} < 0 \text{ et } x_4 \neq 0 \end{matrix}$

On en déduits $d_B = -A_B^{-1} A_N d_N = \begin{pmatrix} 8/27 \\ -4/27 \end{pmatrix} \begin{matrix} (x_1) \\ (x_2) \end{matrix}$

La direction est non nulle, on cherche donc $\alpha_1 =$ le $\alpha \geq 0$ maximum tel que $P_1 + \alpha \cdot d \geq 0$.

On a donc $\begin{cases} 1/3 + 8/27\alpha & \geq 0 \\ 1/3 - 4/27\alpha & \geq 0 \\ 0 + 0\alpha & \geq 0 \\ 0 + 4/9\alpha & \geq 0 \end{cases}$

Donc $\alpha_1 = \frac{9}{4}$.

On cherche maintenant le $\alpha$ entre 0 et $\alpha_1$ minimisant $f(P_0 + \alpha d)$.

$$\begin{aligned} g(\alpha) = f(P_0 + \alpha d) &= (1/3 + 8/27\alpha)^2 + 4(1/3 - 4/27\alpha)^2 \\ g'(\alpha) &= 2 * 8/27 * (1/3 + 8/27\alpha) - 8 * 4/27 * (1/3 - 4/27\alpha) \\ g'(\alpha) &= -1/3 + 16/27 \cdot \alpha \\ g'(\alpha) &\leq 0 \\ \Leftrightarrow \alpha &\leq \frac{9}{16} \end{aligned}$$

La fonction $g$ est donc décroissante entre 0 et $\frac{9}{16} \leq \alpha_1$, on pose $\alpha_2 = \frac{9}{16}$.

On se déplace donc vers $P_2 = P_1 + \frac{9}{16} d = \begin{pmatrix} 1/2 \\ 1/4 \\ 0 \\ 1/4 \end{pmatrix}$

On a pas d'annulation de $x_1$ ou $x_2$, on conserve la même base.

On calcule $x_B$, $x_N$, $\nabla f_B$ et $\nabla f_N$ :

$x_B = \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}, \quad x_N = \begin{pmatrix} 0 \\ 1/4 \end{pmatrix}$

$\nabla f_B = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \nabla f_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

On peut donc calculer le gradient réduit :

$$\begin{aligned} {}^t\nabla \bar{f}(x_N) &= -{}^t\nabla f_B \cdot A_B^-1 \cdot A_N + {}^t\nabla f_N \\ &= \begin{pmatrix} 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} -1/3 & -1/3 \\ -2/3 & 1/3 \end{pmatrix} + \begin{pmatrix} 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \end{aligned}$$

On calcule maintenant la direction $d_N = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{matrix} \leftarrow \nabla \bar{f}(x_N)_{x_3} > 0 \text{ et } x_3 = 0 \\ \leftarrow \nabla \bar{f}(x_N)_{x_4} = 0 \text{ et } x_4 \neq 0 \end{matrix}$

La direction $d_N$ est nulle, donc on s'arrête et les conditions (KKT) sont vérifiées.

4. On cherche $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0, \mu_1, \mu_2 \in \mathbb{R}$ tels que

$$\begin{cases} \nabla f(P_2) + \sum\limits_{i=1}^{4} \lambda_i \nabla g_i(P_2) + \sum\limits_{i=1}^{2} \mu_j \nabla h_j(P_2) & = 0 \\ \lambda_i g_i(P_2) & = 0 \ \forall i \leq 4 \end{cases}$$

$$\begin{cases} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & = 0 \\ \lambda_1 1/2 & = 0 \\ \lambda_2 1/4 & = 0 \\ \lambda_3 0 & = 0 \\ \lambda_4 11/4 & = 0 \end{cases}$$

$$\begin{cases} \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \mu_1 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 0 \end{pmatrix} + \mu_2 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix} & = 0 \\ \lambda_1 & = 0 \\ \lambda_2 & = 0 \\ \lambda_4 & = 0 \end{cases}$$

On pose donc $\mu_2 = 0$, $\mu_1 = -1$ et $\lambda_3 = 1$ et le système est vérifié.

**Exercice 2 — *Gradient réduit et programme standard***

Soit $(P)$ le programme suivant :

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.c.} \quad \begin{cases} \sum\limits_{j=1}^{n} a_{ij} x_j \leq b_i \ \forall i \in [\![1;m]\!] \\ x \geq 0 \end{cases}$$

On suppose que les fonctions $g_i$ sont linéaires et qu'il existe $z > 0$ réalisable.

1. Ecrire la forme standard de $(P)$ en rajoutant $m$ variable. On notera $y_1, y_2, \dots, y_{n+m}$ les variables de ce programme et $f'$ sa fonction objectif.
2. Montrer que si on fixe les valeurs de $\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$ à $z$ alors, pour toute valeur de $i \in [\![1;m]\!]$, $y_{n+i} \geq 0$. On note $y_0$ cette solution.
3. Montrer qu'il est possible de démarrer l'algorithme du gradient réduit depuis $y_0$ avec la base $B = \{n+i, i \in [\![1;m]\!]\}$.
4. Calculez la direction $d$ obtenue à la première itération et montrer que, pour tout . $i \leq n$, $d_i = -(\nabla f(z))_i$.

### ▶ Correction

1. La forme standard est
$$\min_{y \in \mathbb{R}^{n+m}} f'(y) \quad \text{s.c.} \quad \begin{cases} \sum\limits_{j=1}^{n} a_{ij} y_j + y_{n+i} = b_i \ \forall i \in [\![1;m]\!] \\ y \geq 0 \end{cases} .$$
2. Soit $i \leq m$, alors $\sum\limits_{j=1}^{n} a_{ij} z_j \leq b_i$ ou $\sum\limits_{j=1}^{n} a_{ij} z_j + y_{n+i} = b_i$ alors $y_{n+i} = b_i - \sum\limits_{j=1}^{n} a_{ij} z_j \geq 0$.

3. $y_0$ est une solution réalisable et, la matrice $A_B$ associée à cette base est l'identité, donc elle est inversible. On peut donc démarrer l'algorithme depuis ce point avec cette base.
4. $A_B = I_m$, $A_N$ est la matrice des coefficients $a_{ij}$.
   $x_N = x_0$
   $\nabla f'_N : (\nabla f'(y_0))_i = (\nabla f(x_0))_i$ si $i \in N$
   $\nabla f'_B : (\nabla f'(y_0))_i = 0$ si $i \in B$
   donc ${}^t\nabla \bar{f}'(x_N) = -A_N \nabla f'_B + \nabla f'_N = \nabla f'_N$.
   Pour tout $i \in N$, autrement dit $i \leq n$, on a $y_i > 0$. Donc $d_i = -\nabla \bar{f}'(x_N)_i = -(\nabla f_N)_i = -(\nabla f(x_0))_i$.

**Exercice 3 — *Objectif linéaire***

Soit $(P)$ le programme suivant :

$$\min_{x\in\mathbb{R}^n} c\cdot x \quad \text{s.c.} \begin{cases} Ax = b \\ x \geq 0 \end{cases}$$

où $A$ est une matrice de taille $m \times n$ avec $m \leq n$ et $b$ un vecteur de taille $m$.

Ce type de programme peut être résolu avec l'algorithme du simplexe, similaire à l'algorithme du gradient réduit. On rappelle que, dans le simplexe, toute variable de la non-base $N$ est nulle.

1. On se place en une solution de base réalisable : une solution réalisable $x$ et une base $B$ telle que pour tout $i \in N$, $x_N = 0$. Calculer le gradient réduit et retrouvez la formule des coûts réduits de l'algorithme du simplexe.
2. On suppose que l'espace $Ax = b; x \geq 0$ est borné. Montrer que si, lors d'une itération de l'algorithme du gradient réduit, $d \neq \vec{0}$, il y a nécessairement une variable de $x$ qui est nulle à l'issue de l'itération.
3. Y a-t-il nécessairement changement de base dans ce cas ?
4. On rappelle que l'algorithme du simplexe se déplace de point en point en inversant une variable de base et une variable de non base à chaque itération. Montrer que, même si on commence au même point avec la même base, il existe des cas où l'algorithme du gradient réduit choisit une direction différente de celle de l'algorithme du simplexe.

**► Correction**

1. On a
   $x_N = 0$
   $\nabla f_N = \big(c_i\big)_{i\in N}$ et $\nabla f_B = \big(c_i\big)_{i\in B}$
   donc ${}^t\nabla \bar{f}'(x_N) = -{}^t\nabla f_B \cdot A_B^{-1} \cdot A_N + {}^t\nabla f_N$.
   On note $\bar{A} = A_B^{-1} \cdot A_N$ matrice de taille $|B| \times |N|$
   donc $\nabla \bar{f}'(x_N)_i = -({}^t\nabla f_B \cdot \bar{A})_i + (\nabla f_N)_i$
   donc $\nabla \bar{f}'(x_N)_i = -(\sum_{j\in B} \bar{a}_{ji} c_j) + c_i$
   On retrouve bien la formule du coût réduit.
2. Si $d \neq \vec{0}$, alors on suit la direction $x + \alpha d$ avec :
   — $\alpha_1 = \max\{\alpha | x + \alpha d \geq 0\}$
   — $\alpha_2 = \arg\min\{c \cdot (x + \alpha d) | 0 \leq \alpha \leq \alpha_1\}$
   Or $c \cdot (x + \alpha d) = c \cdot x + \alpha c \cdot d$. Montrons que $c \cdot d < 0$.
   On a alors la fonction qui est décroissante avec $\alpha$, donc on a nécessairement $\alpha_1 = \alpha_2$.
   On peut montrer qu'il existe $d_i < 0$ pour au moins un $i$. Dans le cas contraire, pour tout $\alpha \geq 0$, $x + \alpha_1 d > 0$ ; or on sait que $A(x + \alpha d) = b$ (résultat du cours). On aurait donc $\alpha_1 = +\infty$ ce qui contredirait l'hypothèse que $Ax = b; x \geq 0$ est borné.
   Posons alors $i = \arg\min\{-x_j/d_j | j \leq n, d_j < 0\}$ ; alors $\alpha_1 = -x_i/d_i$ et $(x + \alpha_1 d)_j = 0$. Donc la $i^{\text{e}}$ variable de $i$ s'annule à l'issue de l'itération.
3. Il n'y a pas nécessairement changement de base.
   Par exemple avec le programme suivant :
   $$\min_{x\in\mathbb{R}^n} x_1 \quad \text{s.c.} \begin{cases} x_1 + x_2 + x_3 = 1 \\ x \geq 0 \end{cases}$$
   On part du point $(0.25, 0.25, 0.5)$. On prend la base $B = \{3\}$. On calcule la direction, on trouve $d = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ (cf exo 2 pour pouvoir faire le calcul de tête).

   On se déplace dans la direction donnée, et on tombe sur le point $\begin{pmatrix} 0 \\ 0.25 \\ 0.75 \end{pmatrix}$. Il n'y a pas changement de base car $x_1$ n'est pas dans la base, son annulation ne change donc rien.

4. Par exemple avec le programme suivant :

$$\min_{x\in\mathbb{R}^n} -x_1 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 + x_3 = 4 \\ -x_1 + x_2 + x_4 = 1 \\ x_1 + x_2 - x_5 = 2 \\ x \geq 0 \end{cases}$$

Dans ce programme tout point visité par le simplexe a au moins 2 coordonnées nulles.
On part du point $(0.5, 1.5, 2, 0, 0)$. On prend la base $B = \{1, 2, 3\}$. On a bien $x_N = 0$. On calcule la direction, on trouve comme coûts réduits $-0.5$ pour $x_4$ et $x_5$. $d = \begin{pmatrix} 0.5 \\ 0 \\ -0.5 \\ 0.5 \\ 0.5 \end{pmatrix}$.

On se déplace dans la direction donnée, on a $\alpha_1 = \alpha_2 = 4$, et on tombe sur le point $\begin{pmatrix} 2.5 \\ 1.5 \\ 0 \\ 2 \\ 2 \end{pmatrix}$.

On n'est donc plus sur un point avec 2 coordonnées nulles, ce ne peut donc pas être la direction choisie par l'algorithme du simplexe.