TD 6 : Algorithme du gradient Projeté

Recherche opérationnelle S3.

2024

Exercice 1 — ***Un exemple simple.***

Soit (P) le problème suivant :

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

1. Représenter graphiquement le problème.
2. Vérifier que tout x réalisable de (P) (i.e. satisfaisant les contraintes) vérifie la qualification de l'indépendance linéaire. En déduire que seules la ou les solutions optimales vérifient les conditions de Karush Kuhn-Tucker.
3. Appliquer l'algorithme à partir de $P_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (3 itérations).
4. Vérifier que le point trouvé satisfait les conditions de Karush Khun-Tucker.

▶ **Correction**

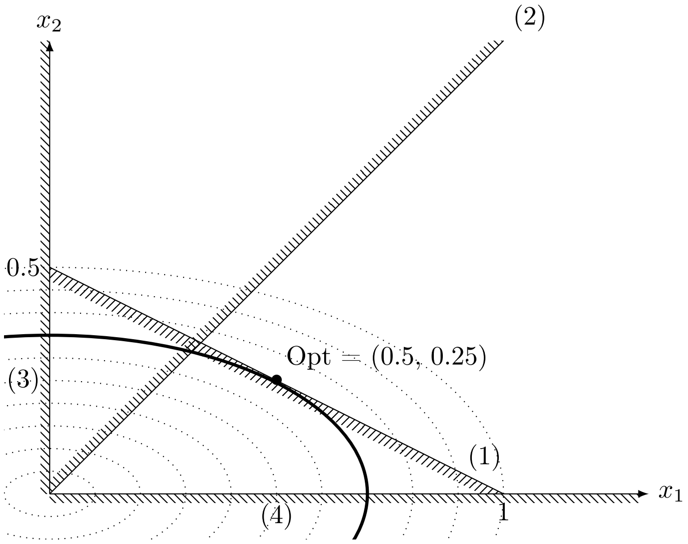
On réécrit le problème sous la bonne forme :

$$\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + 4x_2^2 \;\text{ s.c. } \begin{cases} -x_1 - 2x_2 \leq -1 & (1) \\ -x_1 + x_2 \leq 0 & (2) \\ -x_1 \leq 0 & (3) \\ -x_2 \leq 0 & (4) \end{cases}$$

1. 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.



2. La matrice des coefficients est $A_s = \begin{pmatrix} -1 & -2 \\ -1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$.

   On peut voir sur le dessin que, sur l'ensemble des solutions réalisables, seules les cas suivants sont possibles pour $I(x)$, l'ensemble des inégalités saturées par un point x :
   - $I(x) = \emptyset$
   - $I(x) = \{1\}$
   - $I(x) = \{2\}$
   - $I(x) = \{4\}$
   - $I(x) = \{1, 2\}$
   - $I(x) = \{1, 4\}$

   Dans les 4 premiers cas, $(\nabla g_i(x) | i \in I(x))$ contient 0 ou 1 vecteur, donc il s'agit d'une famille libre.

   Les deux autres cas sont
   - $(\nabla g_i(x) | i \in I(x)) = (\begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix})$ qui est libre (la matrice associée est carrée et de déterminant -3, donc inversible).
   - $(\nabla g_i(x) | i \in I(x)) = (\begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix})$ qui est libre (la matrice associée est carrée et de déterminant 1, donc inversible)..

     On en déduit donc que tout minimum local (ou global) vérifie (KKT). Or, f est convexe et les g_i sont convexes (car linéaires), donc tout point qui vérifie (KKT) est un minimum global. Il y a donc équivalence.

3. Partons de $P_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. On a $I(x) = \{2\}$ et $\nabla f(x) = \begin{pmatrix} 2 \\ 8 \end{pmatrix}$.

   On calcule la projection sur $L = \{y | -y_1 + y_2 = 0\}$:
   - via $d = (I - {}^tA_s \cdot (A_s \cdot {}^tA_s)^{-1} \cdot A_s) \cdot (-\nabla f(x)) = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ -8 \end{pmatrix} = \begin{pmatrix} -5 \\ -5 \end{pmatrix}$
   - ou via $y \in L \Leftrightarrow y_1 = y_2$; donc $d = \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et

   $$(-\nabla f(x)) \cdot d = d^2 \Rightarrow -2\beta - 8\beta = \beta^2 \cdot 2 \Rightarrow \beta = -5$$

   .

On peut vérifier que $d \in L$ et que $(-\nabla f(x)) - d = \begin{pmatrix} 3 \\ -3 \end{pmatrix} \perp d$,

$d \neq \vec{0}$, donc on se déplace le plus loin possible dans la direction indiquée. On note S l'ensemble réalisable.

$$\begin{aligned}
\alpha_1 &= \max_{0 \leq \alpha} \{\alpha | P_0 + \alpha \cdot d \in S\} \\
&= \max_{0 \leq \alpha} \left\{ \alpha | \begin{pmatrix} 1 - 5\alpha \\ 1 - 5\alpha \end{pmatrix} \in S \right\} \\
&= \max_{0 \leq \alpha} \left\{ \alpha | \begin{array}{c} -3 + 15\alpha \leq -1 \\ 0 \leq 0 \\ -1 + 5\alpha \leq 0 \\ -1 + 5\alpha \leq 0 \end{array} \right\} \\
&= \frac{2}{15}
\end{aligned}$$

$$\begin{aligned}
\alpha_2 &= \arg \min_{0 \leq \alpha \leq \alpha_1} \{g(\alpha) = f(P_0 + \alpha \cdot d)\} \\
&= \arg \min_{0 \leq \alpha \leq \alpha_1} \{(1 - 5\alpha)^2 + 4(1 - 5\alpha)^2\} \\
&= \arg \min_{0 \leq \alpha \leq \alpha_1} \{5(1 - 5\alpha)^2\} \\
g'(\alpha) &= 5 \cdot (-5) \cdot 2(1 - 5\alpha) \leq 0 \\
\Leftrightarrow \alpha &\leq \frac{1}{5}
\end{aligned}$$

Donc $g'(\alpha)$ est négatif sur $[0, \alpha_1]$, donc

$$\alpha_2 = \alpha_1 = \frac{2}{15}$$

On se déplace donc vers $x = P_0 + \frac{2}{15} d = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}$.

On a $I(x) = \{1, 2\}$ et $\nabla f(x) = \begin{pmatrix} 2/3 \\ 8/3 \end{pmatrix}$.

On calcule la projection sur $L = \{y| - y_1 + y_2 = 0; -y_1 - 2y_2 = 0\} = \{0\}$, donc $d = \vec{0}$. (on peut aussi recalculer l'opérateur de projection, qui est la matrice nulle).

Donc on exprime $-\nabla f(x)$ sous la forme $\lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) = \lambda_1 \begin{pmatrix} -1 \\ -2 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

- soit en résolvant le système
- soit avec $(A_s \cdot {}^t A_s)^{-1} \cdot A_s \cdot (-\nabla f(x))$ qui donne le vecteur λ.

$(A_s \cdot {}^t A_s)^{-1} \cdot A_s = \begin{pmatrix} -1/3 & -1/3 \\ -2/3 & 1/3 \end{pmatrix}$ donc $\lambda = \begin{pmatrix} 10/9 \\ -4/9 \end{pmatrix}$

On retire la contrainte (2) de $I(x)$. On calcule la projection sur $L = \{y| - y_1 - 2y_2 = 0\}$.

$d = (I - {}^t A_s \cdot (A_s \cdot {}^t A_s)^{-1} \cdot A_s) \cdot (-\nabla f(x)) = \begin{pmatrix} 0.8 & -0.4 \\ -0.4 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} -2/3 \\ -8/3 \end{pmatrix} = \begin{pmatrix} 8/15 \\ -4/15 \end{pmatrix}$

On peut encore vérifier que $d \in L$ et que $(-\nabla f(x)) - d = \begin{pmatrix} -6/5 \\ -12/5 \end{pmatrix} \perp d$

$d \neq \vec{0}$, donc on se déplace le plus loin possible dans la direction indiquée.

$$\begin{aligned}
\alpha_1 &= \max_{0 \leq \alpha} \{\alpha | P_1 + \alpha \cdot d \in S\} \\
&= \max_{0 \leq \alpha} \left\{ \alpha | \begin{pmatrix} 1/3 + 8/15\alpha \\ 1/3 - 4/15\alpha \end{pmatrix} \in S \right\} \\
&= \max_{0 \leq \alpha} \left\{ \alpha | \begin{matrix} -1 \leq -1 \\ -12/15\alpha \leq 0 \\ -1/3 - 8/15\alpha \leq 0 \\ -1/3 + 4/15\alpha \leq 0 \end{matrix} \right\} \\
&= \frac{5}{4}
\end{aligned}$$

$$\begin{aligned}
\alpha_2 &= \arg \min_{0 \leq \alpha \leq \alpha_1} \{g(\alpha) = f(P_1 + \alpha \cdot d)\} \\
&= \arg \min_{0 \leq \alpha \leq \alpha_1} \{(1/3 + 8/15\alpha)^2 + 4(1/3 - 4/15\alpha)^2\} \\
&= \frac{5}{16}
\end{aligned}$$

On se déplace donc vers $x = P_1 + \frac{5}{16}d = \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}$

On a toujours $I(x) = \{1\}$.

On calcule la projection sur $L = \{y| - y_1 - 2y_2 = 0\}$.

$d = (I - {}^tA_s \cdot (A_s \cdot {}^tA_s)^{-1} \cdot A_s) \cdot (-\nabla f(x)) = \begin{pmatrix} 0.8 & -0.4 \\ -0.4 & 0.2 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$d = \vec{0}$ (puisqu'on est en 2D et que l'opérateur de projection est non nul, ça signifie tout simplement que le gradient est orthogonal au bord où on se trouve ; donc son projeté sur le bord est nul).

Donc on exprime $-\nabla f(x) = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$ sous la forme $\lambda_1 \nabla g_1(x) = \lambda_1 \begin{pmatrix} -1 \\ -2 \end{pmatrix}$.

On a trivialement $\lambda_1 = 1$.

Tous les λ_i sont positifs, donc on s'arrête et on renvoie la solution $\begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix}$

4. On a $\nabla f \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix} + \nabla g_1 \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix} = 0$ et $g_1 \begin{pmatrix} 1/2 \\ 1/4 \end{pmatrix} = 0$

   donc, en posant $\lambda_1 = 1$ et $\lambda_2, \lambda_3, \lambda_4 = 0$ $\begin{cases} \nabla f(x) + \sum_{i=1}^{4} \lambda_i \nabla g_i(x) & = 0 \\ \lambda_1 \cdot g_1(x) & = 0 \\ \lambda_2 \cdot g_2(x) & = 0 \\ \lambda_3 \cdot g_3(x) & = 0 \\ \lambda_4 \cdot g_4(x) & = 0 \end{cases}$

   Les conditions (KKT) sont bien vérifiées.

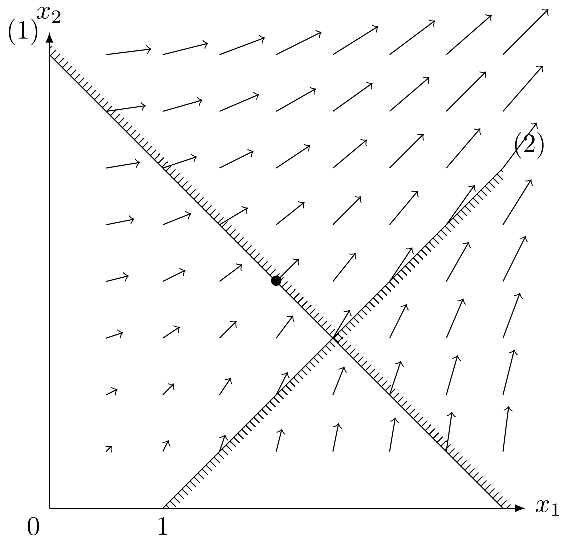
Exercice 2 — *Avec des égalités*

Soit (P) le problème suivant :

$$\max_{x \in \mathbb{R}^2} f(x) = x_1 \cdot x_2 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 + x_3 = 4 & (1) \\ x_1 - x_2 + x_4 = 1 & (2) \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

Mêmes questions que pour l'exercice 1. Pour l'excution de l'algorithme, on partira du point $\begin{pmatrix} 1.5 \\ 0.5 \\ 2 \\ 0 \end{pmatrix}$).

▶ **Correction**

1. 

   On a tracé le gradient sur le dessin. L'optimal est en (2, 2). La fonction n'est pas convexe mais il n'existe qu'un seul point vérifiant les conditions (KKT).

2. On peut réécrire le problème ainsi :

   $$\min_{x\in\mathbb{R}^2} f(x) = -x_1 \cdot x_2 \text{ s.c. } \begin{cases} x_1 + x_2 + x_3 = 4 & (1) \\ x_1 - x_2 + x_4 = 1 & (2) \\ -x_1, -x_2, -x_3, -x_4 \leq 0 & (3)...(6) \end{cases}$$

   La matrice des coefficients est $A = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$

   On peut être dans les cas suivants :
   - $I(x) = \emptyset$
   - $I(x) = \{3\}$
   - $I(x) = \{4\}$
   - $I(x) = \{5\}$
   - $I(x) = \{6\}$
   - $I(x) = \{3, 4\}$
   - $I(x) = \{3, 5\}$
   - $I(x) = \{4, 6\}$
   - $I(x) = \{5, 6\}$

   Par exemple, on peut vérifier l'indépendance linéaire dans le dernier cas directement. On a alors la famille de vecteurs $\left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}\right)$

   La matrice carrée associée est de déterminant -2, donc inversible, donc la famille est libre. Et donc les sous-familles sont libres aussi. La qualification sous l'indépendance linéaire est vérifiée si $I(x) \subset \{5, 6\}$.

3. On a ici les grandes lignes de l'algo, la méthode étant la même que dans l'exo précédent :

   On part du point $\begin{pmatrix} 1.5 \\ 0.5 \\ 2 \\ 0 \end{pmatrix}$

On a $\nabla f(x) = \begin{pmatrix} -0.5 \\ -1.5 \\ 0 \\ 0 \end{pmatrix}$

On a $I(x) = \{6\}$. On projette sur l'espace $L = \{y|y_4 = 0; y_1 + y_2 + y_3 = 0; y_1 - y_2 + y_4 = 0\} = \{y|y_4 = 0; y_3 = -2y_1; y_1 = y_2\}$.

On trouve $d = \begin{pmatrix} 1/3 \\ 1/3 \\ -2/3 \\ 0 \end{pmatrix}$.

$d \neq \vec{0}$. On se déplace donc dans le sens $x + \alpha d$. On trouve $\alpha_1 = \alpha_2 = 3$

On arrive sur $x = \begin{pmatrix} 2.5 \\ 1.5 \\ 0 \\ 0 \end{pmatrix}$.

On a $\nabla f(x) = \begin{pmatrix} -1.5 \\ -2.5 \\ 0 \\ 0 \end{pmatrix}$

On a $I(x) = \{5, 6\}$. Il y a 2 inégalités saturées et 2 égalités, donc l'espace sur lequel on projette est nécessairement $\{0\}$ (car 1. on projette sur un espace orthogonal à un espace à 4 dimension. Puisqu'on est dans $\mathbb{R}^4$, cet espace est de dimension 0, donc il s'agit de $\{0\}$. 2. On a A_S qui est une matrice carrée et inversible. Si on calcule l'opérateur de projection, on aura la matrice nulle.)

Donc on exprime $-\nabla f(x)$ sous la forme $\lambda_5 \nabla g_5(x) + \lambda_6 \nabla g_6(x) + \mu_1 \nabla h_1(x) + \mu_2 \nabla h_2(x)$. On trouve $\lambda_5 = 2$, $\lambda_6 = -0.5$, $\mu_1 = 2$, $\mu_2 = -0.5$. On supprime donc 6 de $I(x)$.

On a $I(x) = \{5\}$. On projette maintenant sur $L = \{y|y_3 = 0; y_1 + y_2 + y_3 = 0; y_1 - y_2 + y_4 = 0\} = \{y|y_3 = 0; y_2 = -y_3; y_4 = y_2/2\}$. On trouve $d = \begin{pmatrix} -1/6 \\ 1/6 \\ 0 \\ 1/3 \end{pmatrix}$.

$d \neq \vec{0}$. On se déplace donc dans le sens $x + \alpha d$. On trouve $\alpha_1 = 15, \alpha_2 = 3$

On arrive sur $x = \begin{pmatrix} 2 \\ 2 \\ 0 \\ 1 \end{pmatrix}$.

On a $\nabla f(x) = \begin{pmatrix} -2 \\ -2 \\ 0 \\ 0 \end{pmatrix}$

On projette sur le même espace que précédemment, $I(x)$ n'a pas changé et vaut toujours $\{5\}$. On trouve $d = \vec{0}$.

Donc on exprime $-\nabla f(x)$ sous la forme $\lambda_5 \nabla g_5(x) + \mu_1 \nabla h_1(x) + \mu_2 \nabla h_2(x)$. On trouve $\lambda_5 = 2$, $\mu_1 = 2$, $\mu_2 = 0$.

Tous les λ_i (ici seulemement λ_5) sont positifs. On s'arrête.

4. Le point vérifie bien (KKT) avec $\lambda_1 = \lambda_2 = \lambda_6 = 0$ et $\lambda_5 = 2$, $\mu_1 = 2$, $\mu_2 = 0$. On ne peut pas en déduire l'optimalité globale car la fonction n'est pas convexe. Mais on pourrait remarquer que c'est le seul point réalisable qui vérifie ces conditions. Donc il n'y a pas d'autre maximum local dans l'espace réalisable. Il est donc optimal.

Exercice 3 — *Opérateur de projection.*

Connaissant un vecteur g de $\mathbb{R}^n$, on souhaite minimiser $\frac{1}{2}\|g - p\|^2$ pour tout vecteur p de $\mathbb{R}^n$ vérifiant $Ap = 0$, sachant que A est une matrice $m \times n$ de rang $m < n$.

Autrement dit, on souhaite trouver le projeté de g sur l'espace $\{p \backslash Ap = 0\}$.

1. Écrire et résoudre les conditions de Kuhn-Tucker pour ce problème.
2. Retrouver, par ce biais, la formule de l'opérateur de projection sur l'espace $L = \{p \backslash Ap = 0\}$.

▶ **Correction**

Tous les points vérifient la qualification de l'indépendance linéaire car le rang de la matrice est m ; les minimums locaux vérifient donc tous KKT et $A\,{}^tA$ est inversible.

1. $$(KKT) : \begin{cases} -(g-p) + {}^tA\mu & = 0 \quad (1) \\ Ap & = 0 \quad (2) \end{cases}$$

$$\begin{aligned} A \cdot (-(g-p) + {}^tA\mu) &= 0 \\ -Ag + Ap + A\,{}^tA\mu &= 0 \\ A\,{}^tA\mu &= Ag \\ \mu &= (A\,{}^tA)^{-1}Ag \end{aligned}$$

2. On reporte dans (1)

$$\begin{aligned} -(g-p) + {}^tA(A\,{}^tA)^{-1}Ag &= 0 \\ p &= (I - {}^tA(A\,{}^tA)^{-1}A) \cdot g \end{aligned}$$

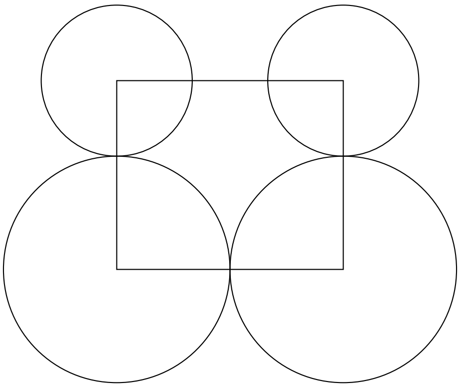
La fonction est convexe, les égalités sont linéaires, donc (KKT) est suffisant pour prouver l'optimalité globale ; p est bien le projeté de g sur L.

Exercice 4 — *Problème non convexe*

On considère le problème suivant :

Soient un rectangle de côté 10×12 ; on souhaite placer 4 disques, centrés en les 4 sommets du rectangle dont l'aire est **maximum** tels que les intérieurs des disques ne s'intersectent pas ; autrement dit, les disques peuvent uniquement se toucher sur leur frontière.

Par exemple, le dessin suivant indique une solution réalisable maximale :



1. Modéliser ce problème sous forme d'un problème de programmation mathématique. On posera $\gamma = \sqrt{244} \simeq 15.62$.
2. Montrer que, si on applique le gradient projeté en partant du point où tous les rayons sont nuls, l'algorithme s'arrête à la première itération.
3. Montrer que, si on applique le gradient projeté en partant du point donné sur la figure (en supposant que les rayons des 2 cercles du bas sont égaux), l'algorithme s'arrête à la première itération.
4. Montrer que ces solutions ne sont pas optimales.

► **Correction**

1. On peut écrire le problème sous la forme suivante, avec x_1, x_2, x_3, x_4 les rayons des disques avec le centre 1 en bas à gauche, le 2 en bas à droite, le 3 en haut à droite et le 4 en haut à gauche.

$$\max_{x \in \mathbb{R}^2} f(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 \leq 12 \\ x_1 + x_3 \leq \gamma \\ x_1 + x_4 \leq 10 \\ x_2 + x_3 \leq 10 \\ x_2 + x_4 \leq \gamma \\ x_3 + x_4 \leq 12 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

On peut le réécrire ainsi :

$$\min_{x \in \mathbb{R}^2} f(x) = -x_1^2 - x_2^2 - x_3^2 - x_4^2 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 \leq 12 \quad (1) \\ x_1 + x_3 \leq \gamma \quad (2) \\ x_1 + x_4 \leq 10 \quad (3) \\ x_2 + x_3 \leq 10 \quad (4) \\ x_2 + x_4 \leq \gamma \quad (5) \\ x_3 + x_4 \leq 12 \quad (6) \\ -x_1, -x_2, -x_3, -x_4 \leq 0 \quad (7), (8), (9), (10) \end{cases}$$

2. Remarque : Ce point est un maximum local, le gradient est donc nul, donc l'algorithme va nécessairement s'arrêter.

Si $x = 0$, alors on a $I(x) = \{7, 8, 9, 10\}$. On calcule la projection sur $L = \{y | -y_1 = -y_2 = -y_3 = -y_4 = 0\} = \vec{0}$.
Autre méthode, la matrice A_s est $-I$, donc ${}^tA_s \cdot (A_s \cdot {}^tA_s)^{-1} \cdot A_s = I$, donc l'opérateur de projection est la matrice nulle.
Autre méthode, le gradient est nul, sa projection est donc nulle aussi.
Donc on exprime $-\nabla f(x) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ sous la forme $\sum_{i=7}^{10} \lambda_i \nabla g_i(x)$. On a trivialement $\lambda_i = 0$ pour tout i, donc on s'arrête.

3. Sur la figure, on peut déduire les rayons du système suivant
$\begin{cases} x_1 + x_2 & = 12 \quad (1) \\ x_1 + x_4 & = 10 \quad (3) \\ x_2 + x_3 & = 10 \quad (4) \end{cases}$ avec $x_1 = x_2$ et $x_3 = x_4$.
On a donc $\begin{cases} x_1 & = 6 \\ x_2 & = 6 \\ x_3 & = 4 \\ x_4 & = 4 \end{cases}$
On a donc $\nabla f(x) = \begin{pmatrix} -12 \\ -12 \\ -8 \\ -8 \end{pmatrix}$
On a $I(x) = \{1, 3, 4\}$ et $A_s = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$.
$I - {}^tA_s \cdot (A_s \cdot {}^tA_s)^{-1} \cdot A_s = \begin{pmatrix} 0.25 & -0.25 & 0.25 & -0.25 \\ -0.25 & 0.25 & -0.25 & 0.25 \\ 0.25 & -0.25 & -0.25 & 0.25 \\ -0.25 & 0.25 & 0.25 & -0.25 \end{pmatrix}$
Donc $d = (I - {}^tA_s \cdot (A_s \cdot {}^tA_s)^{-1} \cdot A_s) \cdot (-\nabla f(x)) = \vec{0}$.
Donc on exprime $-\nabla f(x)$ sous la forme $\sum_{i=1,3,4} \lambda_i \nabla g_i(x) = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \lambda_4 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$.
On a donc $\begin{cases} \lambda_1 + \lambda_3 & = 12 \\ \lambda_1 + \lambda_4 & = 12 \\ \lambda_4 & = 4 \\ \lambda_3 & = 4 \end{cases}$
On en déduit que $\lambda_1 = 8$
Tous les λ_i sont positifs, donc on s'arrête.

4. Ces deux solutions ne sont pas optimales.
La première a pour valeur 0 et la seconde -104.
Par exemple, la solution $x = \begin{pmatrix} 10 \\ 2 \\ \gamma - 10 \\ 0 \end{pmatrix}$ a pour valeur -135.59

Exercice 5 — *Projection et forme standard*

Soit les programmes suivants :

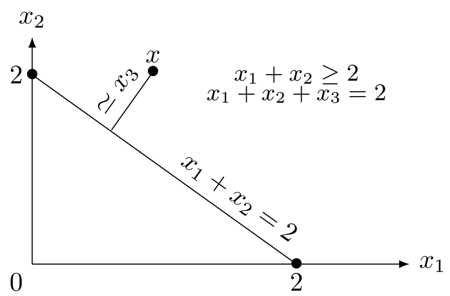
$$\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + x_2^2 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 \geq 2 \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

$$\min_{x\in\mathbb{R}^2} f(x) = x_1^2 + x_2^2 \quad \text{s.c.} \quad \begin{cases} x_1 + x_2 - x_3 = 2 \\ x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0 \end{cases}$$

1. Montrer qu'ils sont équivalents, dans le sens où ils ont même valeur optimale et où toute solution optimale de l'un peut se déduire d'une solution optimale de l'autre.
2. Représenter graphiquement, sur le même dessin, ces deux programmes.
3. Montrer que si on applique une itération du gradient projeté sur le premier programme, depuis le point $x = (3, 2)$, la direction suivie est $(-6, -4)$.
4. Montrer que si on applique une itération du gradient projeté sur le second programme, depuis le point $x = (3, 2, 3)$, la direction suivie est $(-8/3, -2/3, -10/3)$ et le point atteint $(0.6, 1.4, 0)$.
5. Comment expliquer que les programmes, bien qu'équivalents, ne suivent pas la même direction sur le dessin ?

► Correction

1. Ils sont équivalents puisque, s'il existe (x_1, x_2) solution réalisable du premier problème alors (x_1, x_2, x_1+x_2-2) est solution du second avec la même valeur objectif. De même si (x_1, x_2, x_3) est solution du premier alors $x_1 + x_2 = 2 + x_3 \geq 2$ donc (x_1, x_2) est solution du premier programme avec la même valeur objectif.



2.

3. Dans le premier programme, au point $(3, 2)$, $I(x) = \emptyset$, on projet donc sur $\mathbb{R}^2$, donc $d = -\nabla f(x) = \begin{pmatrix} -2x_1 \\ -2x_2 \end{pmatrix} = \begin{pmatrix} -6 \\ -4 \end{pmatrix}$
4. Dans le second programme, au points $(3, 2, -3)$ on a $I(x) = \emptyset$ mais on a une égalité. Donc on projette sur l'espace $L = \{y | y_1 + y_2 - y_3 = 0\}$.
   En utilisant l'opérateur de projection on a : $A_S = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$; donc $PL = I - {}^tA_S \cdot (A_S \cdot {}^tA_S)^{-1} \cdot A_S = \begin{pmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{pmatrix}$ donc $d = PL \cdot \begin{pmatrix} -6 \\ -4 \\ 0 \end{pmatrix} = \begin{pmatrix} -8/3 \\ -2/3 \\ -10/3 \end{pmatrix}$
5. Ceci s'explique par le fait que l'opérateur de projection n'est pas équivalent dans les deux programmes. Dans le second programme, on projette sur un espace légèrement différent du premier. En effet, dans le premier on projette sur $L_1 = \{y \in \mathbb{R}^2\} = \{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\}$ et dans le second, sur $L_2 = \{y | y_1 + y_2 - y_3 = 0\} = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_1 + y_2 - 2 \end{pmatrix} \right\}$
   On peut voir que le premier est, en queque sorte, inclu dans le second ; puisqu'on est libre de choisir n'importe quelle coordonnée y_1 et y_2 dans L_2. Mais l'existence non libre de cette troisième coordonnée fait que le point le plus proche du gradient de f dans L_2 n'a pas exactement les mêmes 2 premières coordonnées, il est un peu dévié.

Il existe toutefois des points où le projeté contient bien $-\nabla f(x)$.

Il s'agit des points où $x_1 = x_2$. Le gradient est alors de la forme $\begin{pmatrix} \alpha \\ \alpha \\ 0 \end{pmatrix}$ et $PL \cdot \begin{pmatrix} -\alpha \\ -\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} -\alpha \\ -\alpha \\ -2\alpha \end{pmatrix}$

On peut expliquer cette différence autrement. Au lieu de faire le dessin en 2D, il faut le faire en 3D. On est dans le plan $P = \{x_3 = x_1 + x_2 - 2\}$. Ce plan est un plan penché. Si le gradient a les deux premières coordonnées égales, alors, comme expliqué précédemment, le projeté a aussi ses deux premières coordonnées égales, on suit donc bien la même direction dans les deux programmes.

Si, à l'inverse ces coordonnées sont différentes, alors le gradient n'est pas dans le sens de la descente, et lorsqu'on le projette sur le plan, alors le projeté dévie un peu. Plus on s'éloigne de la droite $x_1 = x_2$, plus le projeté dévie.