Université Paris Cité
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

Master 1$^{\text{ère}}$ année, MMA, 2023-2024
Optimisation sous contraintes

Examen du 21/05/2024

Durée 2h. Une feuille recto-verso de notes est autorisée.

Exercice 1

Soit $n \in \mathbb{N}^*$, on note $\mathcal{B} = (e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$ et on rappelle la notation suivante

$$\forall x \in \mathbb{R}^n, \qquad x \succeq 0 \quad \Leftrightarrow \quad \forall i \in \{1, \dots, n\},\, x_i \geqslant 0.$$

Dans la suite on considère l'ensemble

$$C = \{x \in \mathbb{R}^n \,/\, x \succeq 0 \text{ et } \|x\|_1 \leqslant 1\}\,.$$

Partie I : questions préliminaires

1. Montrer que C est un convexe fermé non vide.
2. On rappelle que $x \in C$ est un point extrême de C si pour tous $y, \tilde{y} \in C$ tels que $x = \frac{y+\tilde{y}}{2}$ alors $y = \tilde{y} = x$.
   a) Soit C_1, C_2 deux ensembles convexes de $\mathbb{R}^n$, tels que $C_1 \subset C_2$ et soit $x \in C_1$. Montrer que si x est un point extrême de C_2, alors x est un point extrême de C_1. La réciproque est-elle vraie ?
   b) On rappelle que l'ensemble des points extrêmes de $\bar{B}_1(0,1) = \{x \in \mathbb{R}^n \,/\, \|x\|_1 \leqslant 1\}$ est
   $$\{e_1, -e_1, e_2, -e_2, \dots, e_n, -e_n\}.$$
   En déduire n points extrêmes de C.
   c) Pensez-vous que C ne contient que n points extrêmes ? Justifier.
3. Proposer une écriture de C, la plus simple possible, sous la forme
   $$C = \{x \in \mathbb{R}^n \,/\, \forall i \in \{1, \dots, m\},\, f_i(x) \leqslant 0\},$$
   pour un certain $m \in \mathbb{N}^*$, et des fonctions $f_i : \mathbb{R}^n \to \mathbb{R}$, pour $1 \leqslant i \leqslant m$, $\mathcal{C}^\infty$ sur $\mathbb{R}^n$ à déterminer.
4. Soit $z \in \mathbb{R}^n$ fixé. On considère le problème d'optimisation suivant
   $$\inf_{x \in C} \frac{1}{2} \|x - z\|_2^2\,.$$
   a) Justifier que ce problème admet une unique solution et fournir une interprétation géométrique. On notera $p_C(z)$ cette solution.
   b) Si $z \in C$, que vaut $p_C(z)$?
   c) Si $z \notin C$, montrer que l'ensemble des contraintes actives en $p_C(z)$, noté $I(p_C(z))$, est non vide.

Partie II : détermination de p_C lorsque $n = 2$.

Dans cette partie uniquement on suppose que $n = 2$. On admet que l'on peut écrire

$$C = \{x \in \mathbb{R}^n \,/\, f_1(x) \leqslant 0,\ f_2(x) \leqslant 0 \text{ et } f_3(x) \leqslant 0\},$$

avec $f_1 : x \in \mathbb{R}^2 \mapsto -x_1$, $f_2 : x \in \mathbb{R}^2 \mapsto -x_2$ et $f_3 : x \in \mathbb{R}^2 \mapsto x_1 + x_2 - 1$.

On fixe $z = (z_1, z_2) \notin C$. Pour simplifier l'écriture, on notera au besoin $z^\star = p_C(z)$ et $z^\star = (z_1^\star, z_2^\star)$.

5. Montrer que tous les points de C sont réguliers par rapport aux contraintes.

6. Soit $x \in C$. Montrer à l'aide des résultats utilisant les conditions KKT qu'il y a équivalence entre

- $x = z^\star$ (x est le minimum global de $\cancel{f}\ g : x \in \mathbb{R}^n \mapsto \frac{1}{2}\,\|x - z\|_2^2$ sur C),
- il existe $\lambda_1, \lambda_2, \lambda_3$ des réels positifs tels que

$$\begin{aligned}
&\lambda_1 x_1 = 0,\\
&\lambda_2 x_2 = 0,\\
&\lambda_3(x_1 + x_2 - 1) = 0,\\
&x = \Big(z_1 + \lambda_1 - \lambda_3, z_2 + \lambda_2 - \lambda_3\Big).
\end{aligned} \tag{1}$$

7. L'objectif de cette question est de déterminer l'expression de $p_C(z)$ en fonction de z à l'aide de l'équivalence démontrée à la question précédente. Pour cela on va distinguer six situations en fonction de l'appartenance de z aux zones A_1, A_2, B_1, C_1, D_1 et D_2 représentées sur la Figure 1.

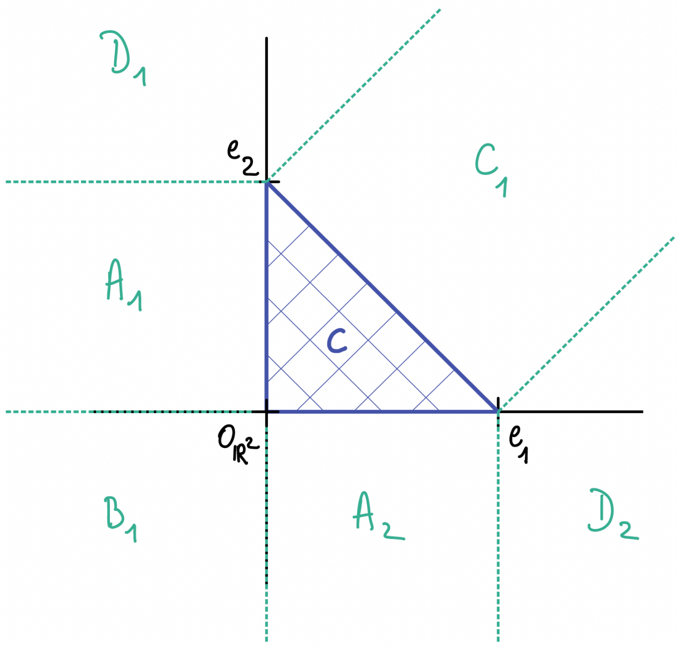

Figure 1. Représentation de C et des zones $A_1, A_2, B_1, C_1, D_1, D_2$ considérées dans la question 7. pour la différenciation des cas lors de la détermination de $p_C(z)$ pour $z \notin C$.

Vous donnerez pour chacune des six zones (sauf C_1) l'expression de $p_C(z)$ en fonction de z. Vous démontrerez rigoureusement ce résultat pour

— A_1 ou A_2 (cas similaires),

— B_1,

— D_1 ou D_2 (cas similaires).

Vous préciserez bien à chaque fois les conditions d'appartenance de z à ces zones.

Indication : vous pourrez dans un premier temps intuiter la projection puis utiliser la CNS de la précédente question.

Partie III : étude d'un problème d'optimisation sur C

On s'intéresse au problème d'optimisation suivant

$$(\mathcal{P}) \qquad \inf_{x\in C} f(x),$$

avec $f : x \in \mathbb{R}^n \mapsto \frac{1}{2}\|Ax-y\|_2^2$ avec $A \in \mathcal{M}_{p,n}(\mathbb{R})$ telle que $\ker(A) = \{0_{\mathbb{R}^n}\}$ et $y \in \mathbb{R}^p$.

8. Montrer que le problème $(\mathcal{P})$ est bien posé et que la solution est unique.

9. a) Que vaut $T_C(0_{\mathbb{R}^n})$? Justifier.

   b) En déduire, en justifiant, $T_C(0_{\mathbb{R}^n})^*$.

   c) On suppose que $A^T y \preceq 0$. Montrer alors que $0_{\mathbb{R}^n}$ est l'unique solution de $(\mathcal{P})$.

   d) Retrouver ce dernier résultat à l'aide des conditions KKT.

10. Rappeler la définition de l'algorithme du gradient projeté et étudier sa convergence lorsqu'il est appliqué au problème $(\mathcal{P})$.

11. On admet que l'on peut utiliser l'algorithme Frank-Wolfe pour $(\mathcal{P})$. On rappelle qu'il produit une suite $(x_k)_{k\in\mathbb{N}}$ d'éléments de C telle que pour $x_0 \in C$ choisi par l'utilisateur, on a

$$\forall k \in \mathbb{N}, \qquad \begin{cases} s_k \in \operatorname{argmin}_{s\in C} \langle \nabla f(x_k),\, s\rangle, \\ \lambda_k = \frac{2}{k+2}, \\ x_{k+1} = (1-\lambda_k)x_k + \lambda_k s_k. \end{cases}$$

   a) Déterminer l'ensemble des points extrêmes de C, que l'on notera $\mathrm{Ext}(C)$.

   b) Justifier que pour tout $\alpha \in \mathbb{R}^n$, le problème

$$(\mathcal{L}) \qquad \inf_{x\in C} \langle \alpha,\, x\rangle,$$

   admet au moins une solution.

   c) On admet qu'au moins une solution de $(\mathcal{L})$ est dans $\mathrm{Ext}(C)$. Proposez, en pseudo-code, une fonction permettant de fournir une solution s_k de

$$\min_{s\in C} \langle \nabla f(x_k),\, s\rangle.$$

   d) <u>Bonus</u> : démontrer la propriété admise à la précédente question. *Indication : toute réponse sera appréciée (notamment dessins etc...). Pour obtenir une preuve rigoureuse, une méthode est d'exploiter les conditions KKT.*