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.

Total : 43.5 = 8 + 12.5 + 23 dont 9 bonus.

Exercice 1 (43.5pt)

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 (8pt)

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 $n = 2$. (12.5pt)

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

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

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.

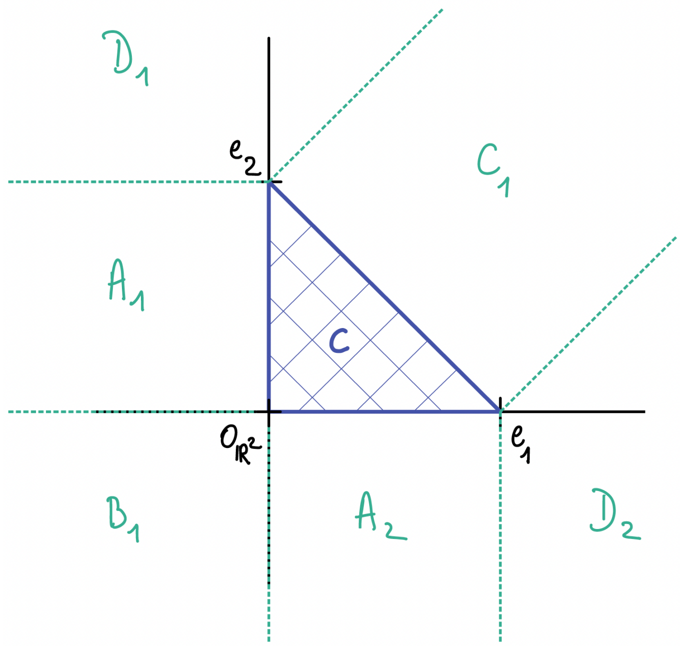

Figure 1. Représentation de C et des zones $A_1, A_2, B_1, C_1, D_1, D_2$ considérées dans la question 7. pour la différenciation des cas lors de la détermination de $p_C(z)$ pour $z \notin C$.

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 (23pt)

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

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

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

Correction.

Partie I

8 = 1.25 + (1.25 + 0.75 + 1) + 1 + (1 + 0.25 + 1.5)

1. L'ensemble C est non vide car contient $0_{\mathbb{R}^n}$ (0.25pt).
   On a $C = \bar{B}_1(0,1) \cap (\mathbb{R}_+)^n$ où $\bar{B}_1(0,1)$ est la boule unité fermée pour la norme ℓ_1 qui est donc est convexe. De plus $(\mathbb{R}_+)^n$ est fermé (comme produit cartésien de fermés ou alternativement directement par la caractérisation séquentielle des fermés) et convexe (en retournant à la définition par exemple). Ainsi C est un convexe fermé comme intersection de convexes fermés (1pt).

2. a) Notons que puisque $x \in C_1$, on a $x_2 \in C_2$ comme $C_1 \subset C_2$, donc l'hypothèse x point extrême de C_2 fait sens. Montrons que x est un point extrême de C_1. Soit donc $y, \tilde{y} \in C_1$ tels que $x = \frac{y+\tilde{y}}{2}$. Comme $C_1 \subset C_2$, on a $y, \tilde{y} \in C_2$. Or comme x est un point extrême de C_2, on a donc $y = \tilde{y} (= x)$. D'où la conclusion (0.5pt).

   La réciproque à cette proposition est fausse. Contre-exemple : soit les deux convexes $C_1 = \{0\}$ et $C_2 = \mathbb{R}$, $C_1 \subset C_2$. Alors $0 \in C_1$ est trivialement un point extrême de C_1 comme C_1 est réduit au singleton $\{0\}$. Mais 0 n'est pas un point extrême de $C_2 = \mathbb{R}$ car $\mathbb{R}$ ne contient aucun point extrême : pour tout $x \in \mathbb{R}$, on peut toujours écrire $x = \frac{(x-1)+(x+1)}{2}$ avec $\frac{x-1}{2}, \frac{x+1}{2} \in C_2$ et $\frac{x+1}{2} \neq \frac{x-1}{2}$ (0.75pt).

   b) On a que le convexe $C = \bar{B}_1(0,1) \cap (\mathbb{R}_+)^n$ est un sous-ensemble du convexe $\bar{B}_1(0,1)$. Les vecteurs $e_1, e_2, \ldots, e_n$ sont des éléments de C et des points extrêmes de $\bar{B}_1(0,1)$, donc par le résultat de la question précédentes, ce sont des points extrêmes de C (0.75pt).

c) En faisant par exemple un dessin de C en dimension 2, on intuite que $0_{\mathbb{R}^n} \notin \{e_1, e_2, \ldots, e_n\}$ est également un point extrême de C. Montrons cette affirmation.

Soit $y, \tilde{y} \in C$ tels que $0_{\mathbb{R}^n} = \frac{y+\tilde{y}}{2}$. On a donc pour tout $i \in \{1, \ldots, n\}$, $0 = \frac{y_i+\tilde{y}_i}{2}$. Or $y_i, \tilde{y}_i \geqslant 0$. D'où $y_i = \tilde{y}_i = 0$, i.e. $y = \tilde{y} = 0_{\mathbb{R}^n}$.

Ainsi C contient au moins $n+1$ points extrêmes : $\{0_{\mathbb{R}^n}, e_1, \ldots, e_n\}$ (1pt).

> On peut montrer qu'il n'y en a pas d'autre ! Nous ferons la preuve dans le cas $n = 2$ ci-dessous.

3. Soit $x \in \mathbb{R}^n$, alors on a $x \in C$ si et seulement si pour tout $i \in \{1, \ldots, n\}$, $-x_i \leqslant 0$ et $\sum_{i=1}^n |x_i| \leqslant 1$. Soit $x \in C$, on a donc $x_i \geqslant 0$ d'où $\sum_{i=1}^n |x_i| = \sum_{i=1}^n x_i$. Ainsi en posant pour tout $i \in \{1, \ldots, n\}$, $f_i : x \in \mathbb{R}^n \mapsto -x_i$ et $f_{n+1} : x \in \mathbb{R}^n \mapsto x_1 + \ldots x_n - 1$, qui sont des fonctions $\mathcal{C}^\infty$ sur $\mathbb{R}^n$, on a

$$C = \{x \in \mathbb{R}^n \,/\, f_1(x) \leqslant 0, \ldots, f_n(x) \leqslant 0 \text{ et } f_{n+1}(x) \leqslant 0\}\,. \,(1pt)$$

Pour la suite il pourra être utile de calculer les gradients de ces fonctions. Soit $x \in \mathbb{R}^n$, on a pour tout $i \in \{1, \ldots, n\}$, $f_i(x) = \langle -e_i, x\rangle$ et $f_{n+1}(x) = \langle e_1 + \ldots + e_n, x\rangle - 1$, donc $\nabla f_i(x) = -e_i$ et $\nabla f_{n+1}(x) = e_1 + \ldots + e_n$.

4. a) On peut faire l'étude de ce problème en remarquant qu'il est bien posée car la fonction objectif est continue, coercive et C fermé non vide, puis qu'il admet une unique solution car la fonction objectif est strictement convexe sur le convexe C.

Alternativement on peut remarquer que ce problème d'optimisation est bien connu puisque correspond à trouver le point du convexe fermé non vide C le plus proche de z. On sait d'après le cours que ce problème admet une unique solution qui est la projection de z sur C (1pt).

b) Si $z \in C$, alors pour $x = z$, on a $\frac{1}{2}\|x - z\|_2^2 = 0$ et donc z est l'unique solution du problème d'optimisation, i.e. $p_C(z) = z$ (0.25pt).

c) Si $z \notin C$ alors on sait par le cours que $p_C(z) \in \partial C$. Or si pour tout $i \in \{1, \ldots, n+1\}$, $f_i(p_C(z)) < 0$, par continuité des f_i, on a $p_C(z) \in \overset{\circ}{C}$. Contradiction avec $p_C(z) \in \partial C = \bar{C} \setminus \overset{\circ}{C} = C \setminus \overset{\circ}{C}$. D'où il existe $i_0 \in \{1, \ldots, n+1\}$, tel que $f_{i_0}(p_C(z)) = 0$, i.e. $i_0 \in I(p_C(z)) \neq \varnothing$ (1.5pt).

Partie II

12.5 = 2 + 4.5 + 6

5. Soit $x \in C$. Si $x \in \overset{\circ}{C}$, alors $I(x) = \varnothing$ et donc x est un point régulier pour les contraintes définissant C.

Sinon $x \in \partial C$ et on sait d'après ce qui précède qu'il existe $i_0 \in \{1, 2, 3\}$ tel que $f_{i_0}(x) = 0$. Donc $I(x) = \varnothing$. De plus on ne peut avoir $I(x) = \{1, 2, 3\}$ car alors $x_1 = x_2 = 0$ et donc $x_1 + x_2 = 0 \neq 1$. D'où $I(x)$ contient un ou deux éléments. Or $\nabla f_1(x) = -e_1$, $\nabla f_2 = -e_2$ et $\nabla f_3(x) = e_1 + e_2$, donc n'importe quel sous famille de cardinal deux (ou un) parmi ces trois vecteurs est libre. Ainsi x est également régulier (2pt).

6. Notons $g : x \in \mathbb{R}^2 \mapsto \frac{1}{2}\|x - z\|_2^2$ la fonction objectif du problème d'optimisation, comme suggéré dans l'énoncé, qui est $\mathcal{C}^\infty$ sur $\mathbb{R}^n$.

($\Rightarrow$) Pour le sens direct, si $x = z^\star$, comme tous les points de C sont réguliers, c'est en particulier le cas en $x \in C$. Ainsi les contraintes sont qualifiées en x. Or c'est le minimum global de g sur C, donc les conditions KKT sont satisfaites en $x = z^\star = p_C(z)$ (0.75pt) i.e. il existe $\lambda_1, \lambda_2, \lambda_3$ des réels positifs (conditions d'admissibilité duale) tels que

$$\begin{aligned}
&\lambda_1 f_1(x) = 0,\\
&\lambda_2 f_2(x) = 0,\\
&\lambda_3 f_3(x) = 0,\\
&\nabla g(x) + \lambda_1 \nabla f_1(x) + \lambda_2 \nabla f_2(x) + \lambda_3 \nabla f_3(x) = 0_{\mathbb{R}^2}. \,(1pt)
\end{aligned}$$

Les trois premières égalités sont les conditions de complémentarité et la dernière la condition de stationnarité. Comme $\nabla g(x) = x - z$, on obtient après simplifications les relations attendues (0.75pt).

($\Leftarrow$) Si la réciproque est vraie, alors comme montré au-dessus les conditions (1) correspondent aux conditions de complémentarité et stationnarité pour le problème $\inf_{x \in C} g$ et la condition d'admissibilité primale $x \in C$ et la condition d'admissibilité duale $\lambda_1, \lambda_2, \lambda_3 \geqslant 0$ sont aussi satisfaites, par conséquent les conditions KKT pour le problème sont satisfaites en $x \in C$. Comme la fonction g est convexe sur $\mathbb{R}^n$ donc sur C, les conditions KKT pour le problème, satisfaites en x, sont suffisantes pour que $x \in C$ soit un minimum global de g sur C, donc le minimum global de g sur C, i.e. $x = z^\star$ (2pt).

7. — Supposons $z \in A_1$. C'est le cas si et seulement si $z_1 < 0$ et $0 \leqslant z_2 \leqslant 1$ (0.25pt).
Conjecture. On intuite que $x = (0, z_2) \in C$ est la projection de z sur C et dans ce cas $I(x) \subset \{1\}$, sauf les cas extrêmes $z_2 = 0$ ou $z_2 = 1$, car $f_1(x) = 0$, $f_2(x) < 0$ (sauf si $z_2 = 0$) et $f_3(x) = z_2 - 1 < 0$ (sauf si $z_2 = 1$).
Posons donc $\lambda_2 = 0$, $\lambda_3 = 0$. On a $x = (0, z_2) = (z_1 + \lambda_1, z_2)$ si et seulement si $\lambda_1 = -z_1 > 0$. On a donc bien trouvé $\lambda_1, \lambda_2, \lambda_3$ positifs tels que pour $x = (0, z_2) \in C$ les conditions (1) sont satisfaites. Par la condition nécessaire et suffisante démontrée à la question précédente, on déduit que $p_C(z) = x = (0, z_2)$ (1.5pt).

On s'attend à ce que $p_C(z) = (z_1, 0)$ lorsque $z \in A_2$ (0.5pt).

— Supposons $z \in B_1$. C'est le cas si et seulement si $z_1 < 0$ et $z_2 \leqslant 0$ ou $z_1 \leqslant 0$ ou $z_2 < 0$ (0.25pt).
Conjecture. On intuite que $x = (0, 0) \in C$ est la projection de z sur C. Dans ce cas $I(x) = \{1, 2\}$ (car $f_1(x) = 0$, $f_2(x) = 0$ et $f_3(x) = -1 < 0$).
Posons donc $\lambda_3 = 0$. On a $x = (0, 0) = (z_1 + \lambda_1, z_2 + \lambda_2)$ si et seulement si $\lambda_1 = -z_1 \geqslant 0$ et $\lambda_2 = -z_2 \geqslant 0$. On a donc bien trouvé $\lambda_1, \lambda_2, \lambda_3$ positifs tels que pour $x = (0, 0) \in C$ les conditions (1) sont satisfaites. Par la condition nécessaire et suffisante démontrée à la question précédente, on déduit que $p_C(z) = (0, 0)$ (1.5pt).

— Supposons que $z \in D_1$. C'est le cas si et seulement si $z_2 \geqslant 1$ et $z_2 - z_1 > 1$ ou $z_2 > 1$ et $z_2 - z_1 \geqslant 1$ (0.25pt).
Conjecture. On intuite que $x = e_2$ est la projection de z sur C et dans ce cas on a $I(x) = \{1, 3\}$.
Posons donc $\lambda_2 = 0$, et on cherche λ_1, λ_3 tels que $x = e_2 = (0, 1) = (z_1 + \lambda_1 - \lambda_3, z_2 - \lambda_3)$. On a donc $\lambda_3 = z_2 - 1 \geqslant 0$ (car $z_2 \geqslant 1$) puis $\lambda_1 = \lambda_3 - z_1 = z_2 - z_1 - 1 \geqslant 0$ (car $z_2 - z_1 \geqslant 1$). Ainsi pour $x = e_2 \in C$, on a trouvé $\lambda_1, \lambda_2, \lambda_3$ des réels positifs tels que les conditions (1) sont satisfaites. Par la CNS, on donc $p_C(z) = e_2$ (1.5pt).

On s'attend à ce que $p_C(z) = e_1$ lorsque $z \in D_2$ (0.25pt).

Partie III

$23 = 2.75 + (1 + 1.5 + 1.5 + 3.75) + 3 + (3 + 0.5 + 3 + 3)$

8. La fonction f est continue sur $\mathbb{R}^n$ et C est fermé borné donc compact, d'où f admet un minimum global sur C. Ainsi le problème $(\mathcal{P})$ est bien posé (0.5pt).

La fonction f est $\mathcal{C}^2$ sur $\mathbb{R}^n$, et pour tout $x \in \mathbb{R}^n$, $\nabla f(x) = A^T(Ax - b)$, puis $\nabla^2 f(x) = A^T A$ (0.75pt). On a pour tout $h \in \mathbb{R}^n$, $\langle \nabla^2 f(x) h, h \rangle = (Ah)^T(Ah) = \|Ah\|_2^2 \geqslant 0$. Et $\langle \nabla^2 f(x) h, h \rangle = 0$ si et seulement si $Ah = 0$ i.e. $h = 0$ comme $\ker(A) = \{0\}$. Ainsi pour tout $x \in \mathbb{R}^n$, $\nabla^2 f(x) \succ 0$ (0.75pt), ce qui implique que f est strictement convexe sur $\mathbb{R}^n$, donc également sur le convexe C. Ainsi f admet un unique minimum global sur C (0.75pt).

> Remarque post-correction. Attention la matrice A est a priori rectangulaire, donc on ne peut avoir des arguments à base de A est définie positive car $\ker(A) = \{0\}$ (même si elle était carré, il faudrait aussi qu'elle soit symétrique positive, ce que l'énoncé ne suppose pas). Ne pas aller trop vite !

9. a) Comme C est convexe, on a $T_C(0) = \overline{\mathbb{R}_+^*(C - 0)}$. Or $\mathbb{R}_+^*(C - 0) = \mathbb{R}_+^* C = (\mathbb{R}_+)^n$, qui est un ensemble fermé, donc $T_C(0) = (\mathbb{R}_+)^n$ (1pt).

> Remarque post-correction. Ne pas hésiter à s'aider d'un dessin dans un premier temps pour intuiter la réponse et se guider vers la preuve. Nous avons traité de nombreux exemples similaires en classe. Une réponse alternative originale était de dire que $0_{\mathbb{R}^n}$ est un point régulier de C par rapport aux contraintes (a été montré précédemment), donc les contraintes sont qualifiées en 0, d'où $T_C(0) = T_C^\ell(0)$. Et $T_C^\ell(0)$ se détermine facilement. Mais c'était plus long !

b) On a donc $T_C(0)^* = (\mathbb{R}_+)^n$. En effet :
($\supset$) Soit $v \in (\mathbb{R}_+)^n$, alors pour $h \in (\mathbb{R}_+)^n$, on a bien $\langle v, h \rangle \geqslant 0$, d'où $v \in T_C(0)^*$ (0.5pt).
($\subset$) Soit $v \notin (\mathbb{R}_+)^n$, alors il existe $k_0 \in \{1, \ldots, n\}$ tel que $v_{k_0} < 0$. Posons $h = (0, \ldots, 0, -v_{k_0}, 0, \ldots, 0) \in (\mathbb{R}_+)^n$ où le coefficient non nul de h est en position k_0. Alors $\langle v, h \rangle = -v_{k_0}^2 < 0$, d'où $v \notin T_C(0)^*$ (1pt).

> Remarque post-correction. Idem commencer par faire un dessin en cas de besoin.

c) On a $\nabla f(0) = -A^T y$, donc par hypothèse $\nabla f(0) \succeq 0_{\mathbb{R}^n}$ i.e.

$$\nabla f(0) \in T_C(0)^*. \tag{2}$$

Or f est convexe sur C, donc la condition (2) d'optimalité d'ordre 1 est suffisante pour que $0_{\mathbb{R}^n}$ soit un (donc le, vue la question 8.) minimum global de f sur C (1.5pt).

d) Comme le problème d'optimisation est convexe (fonction objectif convexe sur le convexe C), on sait que si $0_{\mathbb{R}^n} \in C$ satisfait les conditions KKT associées au problème, alors c'est un (donc le par la question 8.) minimum global de f sur C (0.5pt).

Il s'agit donc de montrer qu'il existe $\lambda_1, \ldots, \lambda_{n+1}$ des réels positifs tels que pour tout $i \in \{1, \ldots, n+1\}$, $\lambda_i f_i(0_{\mathbb{R}^n}) = 0$ et $\nabla f(0_{\mathbb{R}^n}) + \sum_{i=1}^{n+1} \lambda_i \nabla f_i(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^n}$ (0.5pt). Raisonnons par analyse synthèse.

Analyse. Supposons un instant que les conditions sont satisfaites et regardons ce que l'on peut en déduire. La contrainte associée à f_{n+1} n'est pas active en $0_{\mathbb{R}^n}$ puisque $f_{n+1}(0_{\mathbb{R}^n}) = -1 < 0$, donc on doit nécessairement avoir $\lambda_{n+1} = 0$. Puis pour n'importe quels $\lambda_1, \ldots, \lambda_n$ positifs, on a bien $\lambda_i \underbrace{f_i(0_{\mathbb{R}^n})}_{=0} = 0$ pour tout $i \in \{1, \ldots, n\}$, donc ces conditions de complémentarité n'induisent aucunes contraintes sur ces λ_i. Regardons ce qu'implique la condition de stationnarité en utilisant $\lambda_{n+1} = 0$, on doit avoir

$$-A^T y = \sum_{i=1}^{n} \lambda_i e_i.$$

Comme $-A^T y \succeq 0$, en posant pour tout $i \in \{1, \ldots, n\}$, $\lambda_i = (-A^T y)_i \geqslant 0$, l'égalité précédente est donc bien satisfaite.

Synthèse. Posons pour tout $i \in \{1, \ldots, n\}$, $\lambda_i = (-A^T y)_i \geqslant 0$ et $\lambda_{n+1} = 0$. Alors les conditions de complémentarité sont satisfaites, ainsi que la condition de stationnarité par les calculs faits dans la phase d'analyse. Ainsi les conditions KKT pour le problème $(\mathcal{P})$ sont satisfaites en $0_{\mathbb{R}^n}$, d'où la conclusion (2.75pt).

10. L'algorithme du gradient projeté s'applique à la minimisation d'une fonction $\mathcal{C}^1$ sur un ouvert contenant l'ensemble des contraintes qui doit être convexe fermé non vide. C'est bien le cas ici puisque f est $\mathcal{C}^1$ (et même $\mathcal{C}^\infty$) sur $\mathbb{R}^n$ et C est un convexe fermé non vide. Après le choix d'un point initial $x_0 \in \mathbb{R}^n$ et d'un pas $\tau > 0$, l'algorithme définie une suite $(x_k)_{k \in \mathbb{N}}$ telle que

$$\forall k \in \mathbb{N}, \quad x_{k+1} = p_C(x_k - \tau \nabla f(x_k)).$$

Si f est m-fortement convexe sur C, pour $m > 0$, et ∇f est L-Lipschitzienne sur C, pour $L > 0$, alors la suite $(x_k)_{k \in \mathbb{N}}$ converge vers l'unique minimum global de f sur C si $\tau \in]0, \frac{2m}{L^2}[$ (1.5pt).

Montrons ces deux points. On sait déjà que pour tout $x \in \mathbb{R}^n$, $\nabla^2 f(x) = A^T A$ avec $A^T A$ une matrice symétrique définie positive. Elle admet donc une plus petite valeur propre $m > 0$ et alors $\nabla^2 f(x) \succeq m I_n$ pour tout $x \in \mathbb{R}^n$. D'où f est bien m-fortement convexe. Puis pour $x, x' \in \mathbb{R}^n$, on a

$$\left\|\nabla f(x) - \nabla f(x')\right\|_2 = \left\|A^T A(x - x')\right\|_2 \leqslant \left\|A^T A\right\|_2 \left\|x - x'\right\|_2,$$

où $\left\|A^T A\right\|_2$ est la norme d'algèbre pour les matrices carrés de taille $n \times n$ associée à la norme euclidienne sur $\mathbb{R}^n$. En posant $L = \left\|A^T A\right\|_2 > 0$, on a bien que ∇f est L-Lipschitzienne (1.5pt).

11. a) On a déjà vu dans la première partie que $\mathrm{Ext}(C) \subset \{0_{\mathbb{R}^n}, e_1, \ldots, e_n\}$. Montrons qu'il n'y a pas d'autres points extrêmes. En fait ce n'est pas si facile... Sera complété plus tard. J'ai mis la question en bonus.

Bonus : 3pt

b) La fonction $h : x \in \mathbb{R}^n \mapsto \langle \alpha, x \rangle$ est continue sur $\mathbb{R}^n$ car linéaire, on a déjà vu que C est compact non vide, d'où h admet un minimum global sur C, i.e. le problème est bien posé (0.5pt).

> Remarque post-correction. J'ai beaucoup vu que cette fonction h est coercive sur $\mathbb{R}^n$! D'ailleurs l'erreur a aussi été commise dans l'interro no3 (et ou le DM). Ce n'est pas correct ! En effet, prenons un $x_0 \neq 0$ orthogonal à α. Alors pour tout $t \in \mathbb{R}$, $h(tx_0) = t \langle \alpha, x_0 \rangle = 0$. Donc $\lim_{|t| \to +\infty} h(tx_0) = 0 \neq +\infty$. On ne peut donc avoir $\lim_{\|x\|_2 \to +\infty} h(tx) = +\infty$.
> Ici on est dans un cas tout simple : fonction continue sur un compact non vide C, d'où l'existence d'un min global sur C. Pas besoin de coercivité pour obtenir l'existence.
> Enfin, bien que la fonction h soit convexe (car linéaire), cette propriété ne nous intéresse pas ici pour montrer l'existence d'un minimum global. Pour rappel, la convexité permet typiquement le passage de min local à min global. La stricte convexité fournit l'unicité (mais pas l'existence !

c) Sera complété plus tard Bonus : 3pt.

d) Sera complété plus tard Bonus : 3pt.