OPT202 – Optimisation Différentiable II

Contrôle des connaissances

- Durée : 3 h (on ne peut probablement pas tout faire...).
- Dans chaque sujet (ou sous-sujet), on peut admettre des résultats énoncés dans des questions ou sous-questions *précédentes* pour traiter la question ou sous-question courante.
- On peut consulter tous les documents *distribués au cours*, ainsi que ses propres notes.
- Ne pas écrire en rouge, couleur réservée aux observations des correcteurs.
- *Éteindre tout appareil connecté à un réseau.*

Partie 1 : Quatre sujets indépendants

1. Convexité. Soient $\|\cdot\|$ la norme euclidienne de $\mathbb{R}^n$, a un point de $\mathbb{R}^n$ et P une partie non vide quelconque de $\mathbb{R}^n$. Montrez que l'ensemble

$$C := \{x \in \mathbb{R}^n : \|x-a\| \leqslant \|x-y\| \text{ pour tout } y \in P\}$$

est convexe et fermé.

2. Projeté d'un polyèdre convexe.

On considère le polyèdre convexe

$$P := \{(x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} : A_1 x_1 + A_2 x_2 \leqslant b\},$$

où les matrices A_i sont de type $m \times n_i$ et $b \in \mathbb{R}^m$. Le projeté de P sur $\mathbb{R}^{n_1}$ est, par définition, l'ensemble

$$P_1 := \{x_1 \in \mathbb{R}^{n_1} : \text{il existe un } x_2 \in \mathbb{R}^{n_2} \text{ tel que } (x_1, x_2) \in P\},$$

Montrez que

$$P_1 = \{x_1 \in \mathbb{R}^{n_1} : y^\mathsf{T}(A_1 x_1 - b) \leqslant 0, \text{ pour tout } y \in \mathbb{R}^m_+ \cap \mathcal{N}(A_2^\mathsf{T})\}, \tag{1}$$

où $\mathcal{N}(A_2^\mathsf{T})$ est le noyau de A_2^T.

3. Meilleure résolution d'un système linéaire non réalisable. On considère le problème d'optimisation en $x \in \mathbb{R}^n$ suivant

$$(P) \qquad \inf_{x \in \mathbb{R}^n} \|Ax - b\|,$$

où $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ et $\|\cdot\|$ est une norme *arbitraire* sur $\mathbb{R}^m$. On note $f : \mathbb{R}^n \to \mathbb{R}$ la fonction *convexe* définie en $x \in \mathbb{R}^n$ par $f(x) = \|Ax - b\|$.

1. Montrez que le sous-différentiel de f en $x \in \mathbb{R}^n$, pour le produit scalaire euclidien, s'écrit

$$\partial f(x) = \{A^\mathsf{T} y : \|y\|_D \leqslant 1, \; y^\mathsf{T}(Ax - b) = \|Ax - b\|\}, \tag{2}$$

où $v \in \mathbb{R}^m \mapsto \|v\|_D := \sup\{v^\mathsf{T} u : \|u\| \leqslant 1\}$ est la norme duale de $\|\cdot\|$ pour le produit scalaire euclidien.

2. On considère à présent le problème d'optimisation suivant, qui est équivalent au problème (P) :

$$(\tilde{P}) \qquad \left\{ \begin{array}{l} \inf_{(x,v)\in\mathbb{R}^n\times\mathbb{R}^m} \ \|v\| \\ Ax - b + v = 0. \end{array} \right.$$

Montrez dans quel sens on peut considérer le problème (D) ci-dessous comme le problème dual lagrangien de $(\tilde{P})$ (et donc de (P)) :

$$(D) \qquad \left\{ \begin{array}{l} \sup_{y\in\mathbb{R}^m} \ b^\mathsf{T} y \\ A^\mathsf{T} y = 0 \\ \|y\|_{_D} \leqslant 1. \end{array} \right.$$

3. Montrez que (P) et (D) ont une solution et qu'il n'y a pas de saut de dualité.
4. (a) Montrez que l'ensemble

$$C := \{x \in \mathcal{R}(A^\mathsf{T}) : \|Ax - b\| = \mathrm{val}(P)\}$$

est convexe (on a noté $\mathcal{R}(A^\mathsf{T})$ l'image de A^T et $\mathrm{val}(P)$ la valeur optimale de (P)).

(b) Montrez que l'ensemble des solutions de (P) s'écrit

$$\mathrm{Sol}(P) = C + \mathcal{N}(A).$$

(c) Montrez que C est réduit à un point lorsque $\|\cdot\|$ est la norme ℓ_2 (ou euclidienne) de $\mathbb{R}^m$.

5. Montrez que, lorsque la norme utilisée dans (P) est la norme $\|\cdot\|_1 : y \in \mathbb{R}^m \mapsto \|y\|_1 := \sum_{i=1}^m |y_i|$, le problème (P) peut s'écrire comme un problème d'optimisation linéaire.

4. Biconjuguée du compteur de composantes non nulles. On note $\|x\|_p$ la norme ℓ_p d'un vecteur $x \in \mathbb{R}^n$ ($p = 1$ ou ∞), $B_1 := \{x \in \mathbb{R}^n : \|x\|_1 \leqslant 1\}$ la boule unité fermée de $\mathbb{R}^n$ pour la norme ℓ_1 et

$$\|x\|_0 := |\{i \in [1:n] : x_i \neq 0\}|$$

le *compteur de composantes non nulles* d'un vecteur $x \in \mathbb{R}^n$ (dans cette définition, $|\cdot|$ est utilisé pour désigner le cardinal d'un ensemble). On considère la fonction $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ définie en $x \in \mathbb{R}^n$ par

$$f(x) = \|x\|_0 + \mathcal{I}_{B_1}(x),$$

où $\mathcal{I}_{B_1}$ est l'indicatrice de B_1 ($\mathcal{I}_{B_1}(x) = 0$ si $x \in B_1$ et $\mathcal{I}_{B_1}(x) = +\infty$ si $x \notin B_1$).

On note $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ et $f^{**} : \mathbb{R}^n \to \overline{\mathbb{R}}$ les conjuguée et biconjuguée de f pour le produit scalaire euclidien de $\mathbb{R}^n$.

1. Montrez que $f^*(x^*) \geqslant 0$ pour tout $x^* \in \mathbb{R}^n$.
2. Montrez que $f^*(x^*) \leqslant (\|x^*\|_\infty - 1)^+$ pour tout $x^* \in \mathbb{R}^n$, où, pour un scalaire t, on a noté $t^+ := \max(0, t)$.
3. Montrez que, pour tout $x^* \in \mathbb{R}^n$, on a

$$f^*(x^*) = (\|x^*\|_\infty - 1)^+. \tag{3}$$

[*Indication* : On pourra examiner successivement les cas où $\|x^*\|_\infty \leqslant 1$ et $\|x^*\|_\infty > 1$.]

4. Montrez que, pour tout $x \in \mathbb{R}^n$, on a

$$f^{**}(x) = \|x\|_1 + \mathcal{I}_{B_1}(x). \tag{4}$$

[*Indication* : On pourra examiner successivement les cas où $\|x\|_1 \leqslant 1$ et $\|x\|_1 > 1$.]

Partie 2 : Ensemble saillant de minimiseurs

Soient $\mathbb{E}$ un espace euclidien de produit scalaire noté $\langle\cdot,\cdot\rangle$ (norme associée $\|\cdot\|$), $B := \{x \in \mathbb{E} : \|x\| \leqslant 1\}$ la boule unité fermée de $\mathbb{E}$, $f \in \overline{\operatorname{Conv}}(\mathbb{E})$ (ce qui signifie que f est une fonction convexe, propre et fermée) et $f_{\min} := \inf\{f(x) : x \in \mathbb{E}\}$. On suppose que $f_{\min}$ est fini et que f a un ensemble de minimiseurs $S := \{x \in \mathbb{E} : f(x) = f_{\min}\}$ non vide.

1. Montrez que, quel que soit $x \in \mathbb{E}$, le problème $\inf\{\|x-y\| : y \in S\}$ a une solution et une seule. On la note $\mathrm{P}_S(x)$.

On note $\mathrm{d}_S(x) := \|x - \mathrm{P}_S(x)\|$ la distance de x à S, pour la norme $\|\cdot\|$. On dit que l'ensemble des minimiseurs S est *saillant* s'il existe $\alpha > 0$ tel que

$$\forall x \in \mathbb{E} : \quad f(x) \geqslant f_{\min} + \alpha\, \mathrm{d}_S(x). \tag{5}$$

Dans une première partie, on se propose de montrer que cette propriété est *équivalente* à l'inclusion suivante (pour le même $\alpha > 0$)

$$S + \alpha B \subset \bigcup_{x \in S} \Big(x + \partial f(x)\Big), \tag{6}$$

où le sous-différentiel $\partial f(x)$ de f en $x \in \mathbb{E}$ est calculé pour le produit scalaire $\langle\cdot,\cdot\rangle$.

Supposons dans un premier temps que (5) ait lieu et l'on se donne pour objectif de démontrer (6). Soient $\bar{x}_0 \in S$, $u \in B$ et $\bar{x} := \mathrm{P}_S(\bar{x}_0 + \alpha u)$. On introduit $g := \bar{x}_0 + \alpha u - \bar{x}$.

2. Montrez que $\|g\| \leqslant \alpha$.

3. Montrez que $\forall x \in \mathbb{E}$, $f(x) \geqslant f(\bar{x}) + \langle g, x - \bar{x}\rangle$.

4. Montrez que (6) est vérifiée.

Supposons à présent que (6) ait lieu et l'on se donne pour objectif de démontrer (5). Soient $x \in \mathbb{E} \backslash S$ et $\bar{x} := \mathrm{P}_S(x)$.

5. Montrez qu'il existe un $\bar{x}_1 \in S$ et $g_1 \in \partial f(\bar{x}_1)$ tels que

$$\bar{x} + \alpha \frac{x - \bar{x}}{\|x - \bar{x}\|} = \bar{x}_1 + g_1.$$

6. Montrez que $f(x) \geqslant f_{\min} + \langle g_1, x - \bar{x}_1\rangle$ et en déduire (5).

Dans une seconde partie, on considère l'algorithme de minimisation de f, qui calcule en $x \in \mathbb{E}$ le nouvel itéré x_+ comme solution de

$$\inf_{y \in \mathbb{E}} \left(f(y) + \frac{1}{2}\|y - x\|^2\right). \tag{7}$$

7. Sachant que les fonctions de $\overline{\operatorname{Conv}}(\mathbb{E})$ ont une minorante affine, montrez que le problème (7) a une solution et une seule.

8. Supposons que f ait un ensemble saillant de minimiseurs. Montrez que si x est suffisamment proche de S, alors $x_+ \in S$.

OPT202 – Optimisation Différentiable II
Contrôle des connaissances (solutions succinctes)

Partie 1 : Quatre sujets indépendants

1. Comme

$$C = \bigcap_{y \in P} C_y, \qquad \text{avec } C_y := \{x \in \mathbb{R}^n : \|x-a\| \leqslant \|x-y\|\},$$

il suffit de montrer que, pour y fixé dans P, C_y est convexe et fermé pour qu'il en soit de même pour C. Or l'inégalité $\|x-a\| \leqslant \|x-y\|$ s'écrit aussi (on élève au carré et on développe) $x^{\mathsf{T}}(y-a) \leqslant (\|y\|^2 - \|a\|^2)/2$; elle est donc affine en x. Donc C_y est un demi-espace fermé de $\mathbb{R}^n$ (c'est la *description de Voronoï* d'un demi-espace), donc un convexe fermé.

Remarque. On notera que la fonction $\varphi : x \mapsto \|x-a\| - \|x-y\|$ (avec a et y fixés) n'est pas convexe. Par exemple le long de $t \mapsto x_t := (1-t)a + ty$, on a

$$\varphi(x_t) = \begin{cases} -\|y-a\| & \text{si } t < 0 & \text{(constante)} \\ (2t-1)\|y-a\| & \text{si } \quad 0 \leqslant t \leqslant 1 & \text{(affine)} \\ \|y-a\| & \text{si } \qquad 1 < t & \text{(constante).} \end{cases}$$

2. On note P_1' l'ensemble à droite dans (1).

$[P_1 \subset P_1']$ Cette inclusion s'obtient facilement en multipliant chaque membre de l'inégalité $A_1x_1 + A_2x_2 \leqslant b$ par un $y \geqslant 0$ dans le noyau de A_2^{T}.

$[P_1 \supset P_1']$ Voici deux approches possibles.

- Si $x_1 \in P_1'$, on a

$$\begin{aligned} b - A_1x_1 &\in (\mathbb{R}^m_+ \cap \mathcal{N}(A_2^{\mathsf{T}}))^+ \\ &= \{y : (A_2 \;\; -A_2 \;\; I)^{\mathsf{T}} y \geqslant 0\}^+ \\ &= \{A_2x_2 + s : s \geqslant 0\}, \end{aligned}$$

  par le lemme de Farkas. Dès lors, il existe un x_2 tel que $A_1x_1 + A_2x_2 \leqslant b$.

  On peut aussi procéder comme suit :

$$\begin{aligned} b - A_1x_1 &\in (\mathbb{R}^m_+ \cap \mathcal{N}(A_2^{\mathsf{T}}))^+ \\ &= \overline{(\mathbb{R}^m_+)^+ + (\mathcal{N}(A_2^{\mathsf{T}}))^+} \\ &= \overline{\mathbb{R}^m_+ + \mathcal{R}(A_2)} \qquad (\mathbb{R}^m_+)^+ = \mathbb{R}^m_+ \text{ et } \mathcal{N}(A_2^{\mathsf{T}})^+ = \mathcal{R}(A_2) \\ &= \mathbb{R}^m_+ + \mathcal{R}(A_2), \end{aligned}$$

  où la première égalité vient de ce que $(K_1 \cap K_2)^+ = \overline{K_1^+ + K_2^+}$ si K_1 et K_2 sont des cônes convexes fermés et la seconde vient du fait que $\mathbb{R}^m_+ + \mathcal{R}(A_2)$ est fermé (c'est un polyèdre convexe comme somme de 2 polyèdres convexes). Comme ci-dessus, il en résulte qu'il existe un x_2 tel que $A_1x_1 + A_2x_2 \leqslant b$.

- On observe que le problème d'optimisation linéaire

$$\begin{cases} \inf\, (b - A_1x_1)^{\mathsf{T}} y \\ A_2^{\mathsf{T}} y = 0 \\ y \geqslant 0 \end{cases}$$

a une solution (par exemple $y = 0$, car son objectif est $\geqslant 0$ sur son ensemble admissible, par hypothèse). On prend comme lagrangien la fonction

$$(y,(x_2,s)) \mapsto (b - A_1x_1)^\mathsf{T} y - x_2^\mathsf{T} A_2^\mathsf{T} y - s^\mathsf{T} y.$$

Les contraintes d'un problème d'optimisation linéaire étant qualifiées, les conditions d'optimalité affirment alors l'existence de multiplicateurs (x_2, s) tels que

$$A_1x_1 + A_2x_2 + s = b \qquad \text{et} \qquad s \geqslant 0.$$

Donc $A_1x_1 + A_2x_2 \leqslant b$, c'est-à-dire $(x_1, x_2) \in P$, ce qui signifie que $x_1 \in P_1$.

3. 1. On a en effet $\partial f(x) = A^\mathsf{T} \partial(\|\cdot\|)(Ax - b)$ (sous-différentiel de la composition d'une fonction convexe et d'une application affine). Comme $\partial(\|\cdot\|)(v) = \{y : \|y\|_D \leqslant 1,\ y^\mathsf{T} v = \|v\|\}$ (séance de TD7), le résultat s'ensuit.

2. On peut écrire $(\tilde{P})$ comme le problème

$$\inf_{(x,v)\in\mathbb{R}^n\times\mathbb{R}^m} \ \sup_{y\in\mathbb{R}^m} \ \|v\| - y^\mathsf{T}(Ax - b + v).$$

De ce point de vue, on obtient comme problème dual

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^m} \ \inf_{(x,v)\in\mathbb{R}^n\times\mathbb{R}^m} \ \|v\| - y^\mathsf{T}(Ax - b + v) \qquad (8)\\
&\quad = \sup_{y\in\mathbb{R}^m} \left(b^\mathsf{T} y - \sup_x (A^\mathsf{T} y)^\mathsf{T} x - \sup_v \left[y^\mathsf{T} v - \|v\| \right] \right)\\
&\quad = \sup_{y\in\mathbb{R}^m} \begin{cases} b^\mathsf{T} y & \text{si } A^\mathsf{T} y = 0 \text{ et } \|y\|_D \leqslant 1\\ -\infty & \text{sinon,} \end{cases}
\end{aligned}$$

où l'on a utilisé la conjuguée $\mathcal{I}_{B_D}$ de la norme. Ce dernier problème est équivalent au problème (D).

3. - Le problème (P) peut aussi s'écrire $\inf\{\|y - b\| : y \in \mathcal{R}(A)\}$ qui a une solution $y \in \mathcal{R}(A)$ (car le critère est continu coercif et l'ensemble admissible est fermé non vide), donc il existe un $x \in \mathbb{R}^n$ tel que $y = Ax$ et ce x est solution de (P).
   - Le problème (D) consiste à maximiser une fonction continue sur un compact non vide ; il a donc une solution.
   - *Typiquement, lorsque (P) et (D) ont une solution, l'absence de saut de dualité se montre au moyen des conditions d'optimalité.*

   Comme une contrainte de (D) est non différentiable et que l'on n'a pas vu de conditions d'optimalité pour ce cas, on utilise la condition d'optimalité de (P) en une solution x, qui s'écrit $0 \in \partial f(x)$. Alors, par (2), il existe un $y' \in \mathbb{R}^m$ tel que $A^\mathsf{T} y' = 0$, $\|y'\|_D \leqslant 1$ et $(y')^\mathsf{T}(Ax - b) = \|Ax - b\|$ ou $-(y')^\mathsf{T} b = \|Ax - b\| = \mathrm{val}(P)$. On voit alors que $y = -y'$ vérifie les contraintes de (D) avec un coût égal à $\mathrm{val}(P)$. Comme $\mathrm{val}(D) \leqslant \mathrm{val}(P)$, on en déduit que y est solution de (D) et qu'il n'y a pas de saut de dualité.

4. (a) L'ensemble C est l'intersection du convexe $\mathcal{R}(A^\mathsf{T})$ et de l'ensemble des solutions de (P), qui s'écrit $\mathrm{Sol}(P) := \{x \in \mathbb{R}^n : \|Ax - b\| = \mathrm{val}(P)\} = \{x \in \mathbb{R}^n : \|Ax - b\| \leqslant \mathrm{val}(P)\}$ (il est convexe comme ensemble de sous-niveau de fonction convexe $x \mapsto \|Ax - b\|$). C'est donc un convexe.

   (b) Si $x \in \mathrm{Sol}(P)$, on peut écrire $x = x_0 + x_1$ avec $x_0 \in \mathcal{N}(A)$ et $x_1 \in \mathcal{N}(A)^\perp = \mathcal{R}(A^\mathsf{T})$. Comme $Ax - b = Ax_1 - b$, il s'ensuit que $x_1 \in \mathcal{R}(A^\mathsf{T}) \cap \mathrm{Sol}(P) = C$. Donc $x \in C + \mathcal{N}(A)$.

   Inversement, si $x = x_1 + x_0$ avec $x_1 \in C$ et $x_0 \in \mathcal{N}(A)$, on a $Ax - b = Ax_1 - b$, si bien que $\|Ax - b\| = \mathrm{val}(P)$, ce qui signifie que $x \in \mathrm{Sol}(P)$.

   (c) L'ensemble C contient les solutions de (P) qui sont dans l'image de A^T. Lorsque $\|\cdot\| = \|\cdot\|_2$, cet ensemble contient l'unique solution de norme minimale de (P).

5. En effet, (P) se récrit alors comme suit

$$\begin{cases} \min_{(x,t)\in\mathbb{R}^n\times\mathbb{R}^m} \sum_{i=1}^m t_i \\ -t_i \leqslant (Ax-b)_i \leqslant t_i, \quad \forall i \in [1:m], \end{cases}$$

qui est bien un problème d'optimisation linéaire (minimisation d'une fonction linéaire sur un polyèdre convexe).

Remarque. Si $\|\cdot\|$ est la norme ℓ_1, en résolvant (P), on cherche à satisfaire exactement le plus grand nombre possible d'équations scalaires du système $Ax = b$ (qui est éventuellement non réalisable), tout en minimisant l'erreur $\|Ax-b\|_1$, sans savoir a priori quelles sont ces équations scalaires qui peuvent être résolues exactement. Ceci est illustré en dimension $m = 2$ à la figure 1. On voit que la solution $\bar{x}_1$ obtenue avec la norme ℓ_1 rend la seconde composante de $A\bar{x}_1$ égale

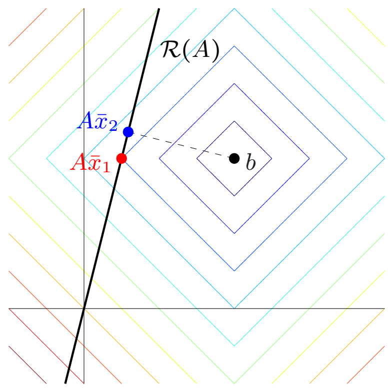

Figure 1: Meilleures solutions $\bar{x}_1$ en norme ℓ_1 et $\bar{x}_2$ en norme ℓ_2, ainsi que les courbes de niveau de la fonction $y \in \mathbb{R}^2 \mapsto \|y-b\|_1$.

à celle de b, donc $(A\bar{x}_1-b)_2 = 0$; alors que la solution $\bar{x}_2$, obtenue avec la norme ℓ_2, rend toutes les composantes de $A\bar{x}_2$ différentes de celles de b. Par ailleurs, la résolution en norme ℓ_2, qui ne demande que la résolution d'un système linéaire, est moins coûteuse que la résolution en norme ℓ_1, qui requiert la résolution d'un problème d'optimisation linéaire.

4. 1. La conjuguée s'écrit en x^*:

$$f^*(x^*) = \sup_{x\in B_1} (x^*)^\mathsf{T} x - \|x\|_0.$$

En prenant $x = 0 \in B_1$, on obtient

$$f^*(x^*) \geqslant 0. \tag{9}$$

2. En utilisant l'inégalité de Hölder, on obtient

$$\begin{aligned} f^*(x^*) &\leqslant \sup_{x\in B_1} \underbrace{\|x^*\|_\infty \|x\|_1 - \|x\|_0}_{\begin{cases} = 0 & \text{si } x = 0 \\ \leqslant \|x^*\|_\infty - 1 & \text{si } x \in B_1 \setminus \{0\} \end{cases}} \\ &\leqslant (\|x^*\|_\infty - 1)^+. \end{aligned} \tag{10}$$

3. • Supposons d'abord que $\|x^*\|_\infty \leqslant 1$. En utilisant les inégalités (9) et (10), on en déduit que $f^*(x^*) = 0$.
   • Supposons maintenant que $\|x^*\|_\infty > 1$. Si $\|x^*\|_\infty = |x_i^*|$, on obtient en prenant $x = \operatorname{sgn}(x_i^*)e^i \in B_1$ que $f^*(x^*) \geqslant \|x^*\|_\infty - 1$. Alors (10) implique que $f^*(x^*) = \|x^*\|_\infty - 1$.

4. Par définition, on a

$$f^{**}(x) = \sup_{x^*\in\mathbb{R}^n} (x^*)^\mathsf{T} x - f^*(x^*) = \sup_{x^*\in\mathbb{R}^n} (x^*)^\mathsf{T} x - (\|x^*\|_\infty - 1)^+.$$

- Supposons dans un premier temps que $\|x\|_1 \leqslant 1$. Considérons successivement les supremums en $x^* \in \mathbb{R}^n$ avec les contraintes additionnelles $\|x^*\|_\infty \leqslant 1$ et $\|x^*\|_\infty \geqslant 1$:

$$\sup_{\|x^*\|_\infty \leqslant 1} (x^*)^\mathsf{T} x - f^*(x^*) = \sup_{\|x^*\|_\infty \leqslant 1} (x^*)^\mathsf{T} x = \|x\|_1$$

et

$$\begin{aligned}
&\sup_{\|x^*\|_\infty \geqslant 1} (x^*)^\mathsf{T} x - f^*(x^*) \\
&= \sup_{\|x^*\|_\infty \geqslant 1} (x^*)^\mathsf{T} x - (\|x^*\|_\infty - 1) \\
&\leqslant \sup_{\|x^*\|_\infty \geqslant 1} \|x^*\|_\infty \underbrace{(\|x\|_1 - 1)}_{\leqslant 0} + 1 \qquad [\text{Hölder}] \\
&= \|x\|_1 \qquad [\text{en prenant } \|x^*\|_\infty = 1, \text{ le plus petit possible}].
\end{aligned}$$

Dès lors

$$\|x\|_1 \leqslant 1 \quad \Longrightarrow \quad f^{**}(x) = \|x\|_1.$$

- Supposons maintenant que $\|x\|_1 > 1$. Alors, il existe un x_1^* tel que $\|x_1^*\|_\infty \leqslant 1$ et $(x_1^*)^\mathsf{T} x > 1$. En prenant x^* de la forme $t x_1^*$ avec $t > 0$, on voit que

$$f^{**}(x) \geqslant \sup_{t \geqslant 1/\|x_1^*\|_\infty} t(x_1^*)^\mathsf{T} x - t\|x_1^*\|_\infty + 1 \geqslant \sup_{t \geqslant 1/\|x_1^*\|_\infty} t\underbrace{[(x_1^*)^\mathsf{T} x - \|x_1^*\|_\infty]}_{>0} + 1 = +\infty.$$

Conclusion : on a montré que, sur la boule unité de la norme ℓ_1, $\|\cdot\|_1$ est la plus grande fonction convexe qui minore la fonction non convexe $\|\cdot\|_0$ (alors que sur $\mathbb{R}^n$ cette plus grande fonction convexe est nulle).

Partie 2 : Ensemble saillant de minimiseurs

1. L'ensemble S s'écrit aussi $S \coloneqq \{x \in \mathbb{E} : f(x) \leqslant f_{\min}\}$. C'est un convexe fermé (car f est convexe fermée et S est un ensemble de sous-niveau de f), non vide (par hypothèse). Donc le projeté d'un point $x \in \mathbb{E}$ sur S existe et est unique.

2. Voici deux démonstrations possibles.

- Comme $\bar{x}$ est le projeté de $\bar{x}_0 + \alpha u$ sur S et $\bar{x}_0 \in S$, on a $\|g\| = \|(\bar{x}_0 + \alpha u) - \bar{x}\| \leqslant \|(\bar{x}_0 + \alpha u) - \bar{x}_0\| = \alpha\|u\| = \alpha$.
- On prend le produit scalaire de $g \coloneqq \bar{x}_0 + \alpha u - \bar{x}$ avec g et on utilise le fait que $\langle g, \bar{x}_0 - \bar{x}\rangle \leqslant 0$. On trouve $\|g\|^2 \leqslant \alpha\langle g, u\rangle \leqslant \alpha\|g\|$ (Cauchy-Schwarz). Donc $\|g\| \leqslant \alpha$.

3. Soit $x \in \mathbb{E}$. On a

$$\begin{aligned}
\langle g, x - \bar{x}\rangle &= \langle g, x - \mathrm{P}_S(x)\rangle + \langle g, \mathrm{P}_S(x) - \bar{x}\rangle \\
&\leqslant \langle g, x - \mathrm{P}_S(x)\rangle \qquad [\bar{x} = \mathrm{P}_S(\bar{x}_0 + \alpha u) \text{ et } \mathrm{P}_S(x) \in S] \\
&\leqslant \alpha\|x - \mathrm{P}_S(x)\| \qquad [\text{Cauchy-Schwarz et } \|g\| \leqslant \alpha] \\
&\leqslant \alpha\, \mathrm{d}_S(x) \qquad [\text{définition de } \mathrm{d}_S(x)] \\
&\leqslant f(x) - f(\bar{x}) \qquad [(5)],
\end{aligned}$$

qui est l'inégalité recherchée.

4. Par le point 3, $g \in \partial f(\bar{x})$. Alors, on a montré que l'élément générique de l'ensemble à gauche dans (6), $\bar{x}_0 + \alpha u$, qui est égal à $\bar{x} + g$, est donc dans l'ensemble à droite dans (6). Ceci démontre (6).

5. On utilise directement (6) avec $\bar{x} \in S$ et $(x - \bar{x})/\|x - \bar{x}\| \in B$.

6. En utilisant le fait que $g_1 \in \partial f(\bar{x}_1)$, $\bar{x}_1 \in S$ et en substituant la valeur de

$$g_1 = \bar{x} - \bar{x}_1 + \alpha(x - \bar{x})/\|x - \bar{x}\|,$$

on trouve

$$f(x) \geqslant f_{\min} + \langle g_1, x - \bar{x}_1 \rangle \geqslant f_{\min} + \frac{\alpha}{\|x - \bar{x}\|}\langle x - \bar{x}, x - \bar{x}_1 \rangle + \underbrace{\langle \bar{x}_1 - \bar{x}, \bar{x}_1 - x \rangle}_{\geqslant 0}.$$

On se libère du dernier terme, qui est positif par une caractérisation de la projection. En partageant $x - \bar{x}_1 = (x - \bar{x}) + (\bar{x} - \bar{x}_1)$ dans le second terme, on obtient

$$f(x) \geqslant f_{\min} + \alpha\, \mathrm{d}_S(x) + \frac{\alpha}{\|x - \bar{x}\|} \underbrace{\langle \bar{x} - x, \bar{x}_1 - \bar{x} \rangle}_{\geqslant 0}.$$

On obtient (5) en observant que le dernier terme est positif par une caractérisation de la projection.

7. Comme $\varphi_x := f + \frac{1}{2}\| \cdot - x\|^2$ est fermée (car f l'est), le problème (7) aura une solution si l'on montre que $\varphi_x(y) \to +\infty$ lorsque $\|y\| \to \infty$. Soit $y \mapsto \langle \xi, y \rangle + \alpha$ une minorante affine de f. En utilisant Cauchy-Schwarz, on trouve

$$\begin{aligned}
\varphi_x(y) \quad &\geqslant \quad \langle \xi, y \rangle + \alpha + \frac{1}{2}\|y - x\|^2 \\
&\geqslant \quad -\|\xi\|\,\|y\| + \alpha + \frac{1}{2}\|y\|^2 - \|y\|\,\|x\| + \frac{1}{2}\|x\|^2 \\
&\geqslant \quad \left(\frac{1}{2}\|y\| - \|\xi\| - \|x\|\right)\|y\| + \alpha + \frac{1}{2}\|x\|^2.
\end{aligned}$$

On en déduit que $\varphi_x(y) \to +\infty$ lorsque $\|y\| \to \infty$ et donc que (7) a au moins une solution.

L'unicité vient de la stricte convexité de φ_x, dérivant elle-même de la stricte convexité de $\|\cdot - x\|^2$.

8. Si $x \in S + \alpha B$, d'après (6), il existe un $\bar{x} \in S$ tel que $x = \bar{x} + g$, avec $g \in \partial f(\bar{x})$, ce qui s'écrit aussi $0 \in \partial f(\bar{x}) + \bar{x} - x$. Ceci montre que $\bar{x}$ est l'unique solution de (7). Dès lors $x_+ = \bar{x}$ et donc $x_+ \in S$.