SOLUTIONS

1 Cône dual d'une intersection de deux cônes

Il y a beaucoup de démonstrations possibles. Dans un premier temps, essayons de voir s'il n'y a pas une inclusion évidente.

- C'est probablement le cas de l'inclusion « $\supset$ », car il suffit de vérifier qu'une propriété a lieu. Voyons cela. Comme $(K_1 \cap K_2)^+$ est un fermé, il suffit de montrer que

$$(K_1 \cap K_2)^+ \supset K_1^+ + K_2^+.$$

Soit donc $d_1 \in K_1^+$ et $d_2 \in K_2^+$. Il s'agit de montrer que $d_1 + d_2 \in (K_1 \cap K_2)^+$. Pour cela, on prend $x \in K_1 \cap K_2$ et on vérifie la positivité de

$$\langle d_1 + d_2, x\rangle = \underbrace{\langle \underbrace{d_1}_{\in K_1^+}, \underbrace{x}_{\in K_1} \rangle}_{\geqslant 0} + \underbrace{\langle \underbrace{d_2}_{\in K_2^+}, \underbrace{x}_{\in K_2} \rangle}_{\geqslant 0} \geqslant 0.$$

- Partant de rien, l'inclusion réciproque $(K_1 \cap K_2)^+ \subset \overline{K_1^+ + K_2^+}$ est plus compliquée à démontrer, car il faut représenter un élément de $(K_1 \cap K_2)^+$ comme une (suite de) somme(s) d'éléments de K_1^+ et K_2^+. On retrouve la même problématique que dans la démonstration du lemme de Farkas, laquelle s'y prenait en raisonnant par l'absurde, ce qui permettait d'utiliser un théorème de séparation de convexes (Hahn-Banach). Au lieu d'utiliser la même stratégie, on regarde si l'on ne peut pas appliquer directement le lemme de Farkas, à savoir

$$\{y \in \mathbb{F} : A^* y \in K^+\}^+ = \overline{A(K)}. \tag{13.11}$$

On pourrait d'ailleurs obtenir directement (13.1), de ce raisonnement (mais la première étape, maintenant inutile, nous a permis de nous mettre sur la bonne piste). En comparant les membres de droite de (13.1) et (13.11), on voit qu'il est judicieux de prendre pour application linéaire A et pour cône convexe K :

$$A : (x_1, x_2) \in \mathbb{E} \times \mathbb{E} \mapsto x_1 + x_2 \in \mathbb{E} \qquad \text{et} \qquad K = K_1^+ \times K_2^+.$$

En effet, le membre de droite de (13.1) est alors bien l'adhérence de $A(K)$. Il reste à voir ce que devient le membre de gauche de (13.11) pour ces choix de A et K.

  - L'adjoint $A^* : \mathbb{E} \to \mathbb{E} \times \mathbb{E}$ de A s'obtient comme suit

$$\begin{aligned}
\langle A^* y, (x_1, x_2)\rangle_{\mathbb{E}\times\mathbb{E}} &= \langle y, A(x_1, x_2)\rangle_{\mathbb{E}} \quad \text{[définition de l'adjoint]} \\
&= \langle y, x_1 + x_2\rangle_{\mathbb{E}} \quad \text{[définition de } A\text{]} \\
&= \langle y, x_1\rangle_{\mathbb{E}} + \langle y, x_2\rangle_{\mathbb{E}} \\
&= \langle (y, y), (x_1, x_2)\rangle_{\mathbb{E}\times\mathbb{E}} \quad \text{[définition du produit scalaire de } \mathbb{E}\times\mathbb{E}\text{]}.
\end{aligned}$$

Donc $A^*(y) = (y, y)$.

- Par ailleurs
$$K^+ = (K_1^+ \times K_2^+)^+ = K_1^{++} \times K_2^{++} = K_1 \times K_2,$$
car $K_i^{++} = K_i$ pour des cônes convexes fermés K_i[2].

Le membre de gauche de (13.11) s'érit donc
$$\{y \in \mathbb{F} : (y, y) \in K_1 \times K_2\}^+ = \{y \in \mathbb{F} : y \in K_1 \cap K_2\}^+ = (K_1 \cap K_2)^+,$$
qui est bien le membre de gauche de (13.1), qui est ainsi démontré.

2 Dualisation lagrangienne de la projection sur un polyèdre convexe

1) On obtient le même résultat en dualisant la contrainte d'égalité ou les deux contraintes.

   - Dualiser la contrainte d'égalité de (13.2) revient à écrire ce problème sous la forme équivalente suivante
   $$\inf_{x \geqslant 0} \sup_{y \in \mathbb{R}^m} \frac{1}{2}\|x - z\|^2 - y^\mathsf{T}(Ax - b). \tag{13.12}$$
   Le dual s'écrit alors
   $$\begin{aligned} &\sup_{y \in \mathbb{R}^m} \inf_{x \geqslant 0} \frac{1}{2}\|x - z\|^2 - y^\mathsf{T}(Ax - b) \\ &\quad = \sup_{y \in \mathbb{R}^m} \inf_{x \geqslant 0} \frac{1}{2}\|x - (A^\mathsf{T}y + z)\|^2 - \frac{1}{2}\|A^\mathsf{T}y + z\|^2 + \frac{1}{2}\|z\|^2 + b^\mathsf{T}y \\ &\quad = \sup_{y \in \mathbb{R}^m} \frac{1}{2}\|(A^\mathsf{T}y + z)^-\|^2 - \frac{1}{2}\|A^\mathsf{T}y + z\|^2 + \frac{1}{2}\|z\|^2 + b^\mathsf{T}y \\ &\quad = \sup_{y \in \mathbb{R}^m} -\frac{1}{2}\|(A^\mathsf{T}y + z)^+\|^2 + \frac{1}{2}\|z\|^2 + b^\mathsf{T}y, \end{aligned}$$
   où l'on a noté $v^- := \max(0, v)$ et utilisé $\|v\|^2 = \|v^+\|^2 + \|v^-\|^2$.
   - Dualiser les deux contraintes de (13.2) revient à écrire ce problème sous la forme équivalente suivante
   $$\inf_{x \in \mathbb{R}^n} \sup_{\substack{y \in \mathbb{R}^m \\ s \geqslant 0}} \frac{1}{2}\|x - z\|^2 - y^\mathsf{T}(Ax - b) - s^\mathsf{T}x. \tag{13.13}$$
   Le dual s'écrit alors
   $$\sup_{\substack{y \in \mathbb{R}^m \\ s \geqslant 0}} \inf_{x \in \mathbb{R}^n} \frac{1}{2}\|x - z\|^2 - y^\mathsf{T}(Ax - b) - s^\mathsf{T}x \tag{13.14}$$
   $$= \sup_{\substack{y \in \mathbb{R}^m \\ s \geqslant 0}} -\frac{1}{2}\|A^\mathsf{T}y + s\|^2 - (A^\mathsf{T}y + s)^\mathsf{T}z + b^\mathsf{T}y$$
   $$= \sup_{y \in \mathbb{R}^m} \left[-\inf_{s \geqslant 0} \left(\frac{1}{2}\|A^\mathsf{T}y + s + z\|^2\right) + \frac{1}{2}\|z\|^2 + b^\mathsf{T}y\right] \tag{13.15}$$

[2] Voir l'exercice VI.3.2.a.

$$
\begin{aligned}
&= \sup_{y\in\mathbb{R}^m} \; -\frac{1}{2}\| \underbrace{[-(A^\mathsf{T} y+z)] - [-(A^\mathsf{T} y+z)]^+}_{=[-(A^\mathsf{T} y+z)]^- = (A^\mathsf{T} y+z)^+} \|^2 + \frac{1}{2}\|z\|^2 + b^\mathsf{T} y,\\
&= \sup_{y\in\mathbb{R}^m} \; -\frac{1}{2}\|(A^\mathsf{T} y+z)^+\|^2 + \frac{1}{2}\|z\|^2 + b^\mathsf{T} y,
\end{aligned}
$$

où le minimum en (13.14) est obtenu en $x = z + A^\mathsf{T} y + s$ et le minimum en (13.15) est obtenu en $s = [-(A^\mathsf{T} y + z)]^+$.

2) Voici deux méthodes.

- *On utilise les conditions d'optimalité du problème primal* (13.2). Comme ce problème s'écrit comme un problème de projection sur un convexe fermé non vide (par hypothèse), il a une solution unique, disons $\bar{x}$. Comme les contraintes de (13.2) sont qualifiées (car affines), il existe $\bar{y} \in \mathbb{R}^m$ et $\bar{s} \in \mathbb{R}^n$ tels que

$$
\left\{
\begin{array}{l}
\bar{x} - z - A^\mathsf{T}\bar{y} - \bar{s} = 0\\
A\bar{x} = b\\
0 \leqslant \bar{x} \perp \bar{s} \geqslant 0.
\end{array}
\right. \tag{13.16}
$$

Par complémentarité, lorsque $\bar{x}_i > 0$, on a $\bar{s}_i = 0$ et donc $\bar{x}_i = (A^\mathsf{T}\bar{y}+z)_i$; et lorsque $\bar{x}_i = 0$, $(A^\mathsf{T}\bar{y}+z)_i = -\bar{s}_i \leqslant 0$; par conséquent,

$$
\bar{x} = (A^\mathsf{T}\bar{y}+z)^+. \tag{13.17}
$$

Par ailleurs,

$$
b^\mathsf{T}\bar{y} = \bar{x}^\mathsf{T} A^\mathsf{T}\bar{y} = \bar{x}^\mathsf{T}(\bar{x} - z - \bar{s}) = \bar{x}^\mathsf{T}(\bar{x} - z). \tag{13.18}
$$

On a alors

$$
\begin{aligned}
\mathrm{val}(D) &\geqslant -\frac{1}{2}\|(A^\mathsf{T}\bar{y}+z)^+\|^2 + b^\mathsf{T}\bar{y} + \frac{1}{2}\|z\|^2\\
&= -\frac{1}{2}\|\bar{x}\|^2 + \bar{x}^\mathsf{T}(\bar{x}-z) + \frac{1}{2}\|z\|^2 \qquad [(13.17) \text{ et } (13.18)]\\
&= \frac{1}{2}\|\bar{x}-z\|^2\\
&= \mathrm{val}(P).
\end{aligned}
$$

Comme $\mathrm{val}(D) \leqslant \mathrm{val}(P)$ par dualité faible, les valeurs optimales primale et duale sont égales.

- *On utilise la condition d'optimalité du problème dual* (13.3). Celle-ci s'obtient en dérivant son critère, qui est différentiable. On obtient

$$
A(A^\mathsf{T}\bar{y}+z)^+ = b. \tag{13.19}
$$

La valeur optimale duale est alors

$$
\begin{aligned}
\mathrm{val}(D) &= -\frac{1}{2}\|(A^\mathsf{T}\bar{y}+z)^+\|^2 + b^\mathsf{T}\bar{y} + \frac{1}{2}\|z\|^2\\
&= -\frac{1}{2}\|(A^\mathsf{T}\bar{y}+z)^+\|^2 + [(A^\mathsf{T}\bar{y}+z)^+]^\mathsf{T} A^\mathsf{T}\bar{y} + \frac{1}{2}\|z\|^2 \qquad [(13.19)]\\
&= -\frac{1}{2}\|(A^\mathsf{T}\bar{y}+z)^+\|^2 + ([A^\mathsf{T}\bar{y}+z]^+)^\mathsf{T}(A^\mathsf{T}\bar{y}+z)
\end{aligned}
$$

$$\begin{aligned}
& \quad - ([A^\mathsf{T}\bar{y} + z]^+)^\mathsf{T} z + \frac{1}{2}\|z\|^2 \\
= & \;\frac{1}{2}\|(A^\mathsf{T}\bar{y} + z)^+\|^2 - ([A^\mathsf{T}\bar{y} + z]^+)^\mathsf{T} z + \frac{1}{2}\|z\|^2 \\
= & \;\frac{1}{2}\|(A^\mathsf{T}\bar{y} + z)^+ - z\|^2 \\
\leqslant & \;\operatorname{val}(P) \qquad [\text{dualité faible}].
\end{aligned}$$

Mais $\bar{x} := (A^\mathsf{T}\bar{y} + z)^+$ est positif et vérifie $A\bar{x} = b$ par (13.19), si bien que $\bar{x}$ est un point admissible primal. Alors la dernière inégalité est en réalité une égalité (et $\bar{x}$ est solution primale). Il n'y a donc pas de saut de dualité.

3) Le problème (13.3) peut aussi s'écrire comme le problème de la minimisation de la fonction convexe $\varphi : \mathbb{R}^m \to \mathbb{R}$ définie en $y \in \mathbb{R}^m$ par

$$\varphi(y) := \frac{1}{2}\|(A^\mathsf{T}y + z)^+\|^2 - b^\mathsf{T}y - \frac{1}{2}\|z\|^2.$$

- $[(a) \Rightarrow (b)]$ Si (b) n'a pas lieu, on peut trouver une direction $p \in \mathbb{R}^m$ telle que $p \neq 0$, $A^\mathsf{T}p \leqslant 0$ et $b^\mathsf{T}p \geqslant 0$. Si $\bar{y}$ minimise φ (donc est solution de (13.3)) et $t \geqslant 0$, on a $A^\mathsf{T}(\bar{y} + tp) + z \leqslant A^\mathsf{T}\bar{y} + z$ composante par composante, si bien que $0 \leqslant [A^\mathsf{T}(\bar{y} + tp) + z]^+ \leqslant [A^\mathsf{T}\bar{y} + z]^+$ et donc $\varphi(\bar{y} + tp) \leqslant \varphi(\bar{y})$. Donc $\bar{y} + tp$ est encore solution de (13.3). Comme $p \neq 0$, l'ensemble des solutions de (13.3) n'est pas borné.
- $[(b) \Rightarrow (c)]$ Il est clair que A est surjective, sinon on pourrait trouver $p \neq 0$ tel que $A^\mathsf{T}p = 0$; comme ce $p \in \mathcal{R}(A)^\perp$ et $b \in \mathcal{R}(A)$, on aurait $b^\mathsf{T}p = 0$, ce qui contredirait (b). Par ailleurs, (b) s'écrit aussi $\{p : A^\mathsf{T}p \leqslant 0,\ b^\mathsf{T}p \geqslant 0\} = \{0\}$ ou $\{p : (-A^\mathsf{T}p, b^\mathsf{T}p) \in \mathbb{R}^{n+1}_+\}^+ = \mathbb{R}^m$. Par le lemme de Farkas, cela revient à dire que $\{-Ax + \alpha b : (x, \alpha) \in \mathbb{R}^{n+1}_+\} = \mathbb{R}^m$. On peut donc trouver $x_1 \geqslant 0$ et $\alpha_1 \geqslant 0$ tels que $A(x_1 + e) = \alpha_1 b$. Alors $x := (x_0 + x_1 + e)/(\alpha_1 + 1)$ vérifie $x > 0$ et $Ax = b$.
- $[(c) \Rightarrow (a)]$ Voici deux méthodes possibles.
  - *On utilise la fonction asymptotique de* φ (concept non vu dans le cours). La fonction asymptotique de φ prend en $p \in \mathbb{R}^m$ la valeur

    $$\varphi^\infty(p) = \begin{cases} -b^\mathsf{T}p & \text{si } A^\mathsf{T}p \leqslant 0 \\ +\infty & \text{sinon.} \end{cases}$$

    Supposons que $\varphi^\infty(p) \leqslant 0$. Alors $A^\mathsf{T}p \leqslant 0$ et $b^\mathsf{T}p \geqslant 0$. Par (c), on peut trouver un $x > 0$ tel que $Ax = b$. En multipliant par p, on trouve $0 \leqslant b^\mathsf{T}p = p^\mathsf{T}Ax \leqslant 0$, si bien que $A^\mathsf{T}p = 0$; donc $p = 0$ par l'injectivité de A^T. On conclut en utilisant l'équivalence $(iii) \Leftrightarrow (iv)$ de la proposition 3.28 du syllabus.
  - *On raisonne par l'absurde.* Supposons que $\operatorname{Sol}(D)$ ne soit pas borné. Il existe alors une suite de solutions duales $\{y_k\}$ qui tendent en norme vers l'infini. En extrayant une sous-suite au besoin, on peut supposer que

    $$\frac{y_k}{\|y_k\|} \to p \neq 0.$$

    En passant à la limite dans $\varphi(y_k)/\|y_k\|^2 = \operatorname{val}(D)/\|y_k\|^2 \to 0$, on trouve que $\|(A^\mathsf{T}p)^+\|^2 = 0$ ou que $A^\mathsf{T}p \leqslant 0$. En passant à la limite dans l'inégalité ci-dessous

    $$-\frac{b^\mathsf{T}y_k}{\|y_k\|} - \frac{1}{2}\frac{\|z\|^2}{\|y_k\|} \leqslant \frac{\varphi(y_k)}{\|y_k\|} \to 0,$$

on trouve que $b^\mathsf{T} p \geqslant 0$. Puis on s'y prend comme dans la première méthode pour trouver une contradiction : par (c), on peut trouver un $x > 0$ tel que $Ax = b$; en multipliant par p, on trouve $0 \leqslant b^\mathsf{T} p = p^\mathsf{T} Ax \leqslant 0$, si bien que $A^\mathsf{T} p = 0$; donc $p = 0$, ce qui contredit le fait que $p \neq 0$.

4) 4.1) Si $b \in \mathcal{R}(A)$, $Ax - b \in \mathcal{R}(A)$ et, au lieu de (13.12), on peut écrire (13.2) comme suit

$$\inf_{x \geqslant 0} \sup_{y \in \mathcal{R}(A)} \frac{1}{2}\|x - z\|^2 - y^\mathsf{T}(Ax - b). \tag{13.20}$$

En effet, si $Ax \neq b$, le suppremum en $y \in \mathcal{R}(A)$ sera $+\infty$ en prenant $y = t(Ax - b) \in \mathcal{R}(A)$ et $t \to \infty$. Comme au point 1, la dualisation de ce problème conduit au dual (13.4) de (13.2).

4.2) Il suffit de montrer que l'on peut trouver une solution $(\bar{x}, \bar{y}, \bar{s})$ de (13.16) avec $\bar{y} \in \mathcal{R}(A)$. Soit $(\bar{x}, \bar{y}_0, \bar{s})$ une solution de (13.16) et $\bar{y}$ la projection de $\bar{y}_0$ sur $\mathcal{R}(A)$. Alors $\bar{y}_0 - \bar{y} \in \mathcal{R}(A)^\perp = \mathcal{N}(A^\mathsf{T})$. Dès lors, $(\bar{x}, \bar{y}, \bar{s})$ est encore solution de (13.16).

4.3) Le problème (13.4) peut aussi s'écrire comme le problème de la minimisation de la fonction convexe $\varphi : \mathbb{R}^m \to \mathbb{R}$ définie en $y \in \mathbb{R}^m$ par

$$\varphi(y) := \frac{1}{2}\|(A^\mathsf{T} y + z)^+\|^2 - b^\mathsf{T} y - \frac{1}{2}\|z\|^2 + \mathcal{I}_{\mathcal{R}(A)}(y).$$

- $[(a) \Rightarrow (b)]$ Si (b) n'a pas lieu, on peut trouver une direction $p \in \mathcal{R}(A)$ telle que $p \neq 0$, $A^\mathsf{T} p \leqslant 0$ et $b^\mathsf{T} p \geqslant 0$. Si $\bar{y}$ minimise φ (donc est solution de (13.4)) et $t \geqslant 0$, on a comme précédemment $\varphi(\bar{y} + tp) \leqslant \varphi(\bar{y})$. Donc $\bar{y} + tp$ est encore solution de (13.4). Comme $p \neq 0$, l'ensemble des solutions de (13.4) n'est pas borné.
- $[(b) \Rightarrow (c)]$ [3] (b) s'écrit aussi $\{p \in \mathcal{R}(A) : A^\mathsf{T} p \leqslant 0,\ b^\mathsf{T} p \geqslant 0\} = \{0\}$ ou

$$\begin{aligned}
\mathbb{R}^m &= \{p \in \mathcal{R}(A) : (-A^\mathsf{T} p, b^\mathsf{T} p) \in \mathbb{R}^{n+1}_+\}^+ \\
&= [\{p : (-A^\mathsf{T} p, b^\mathsf{T} p) \in \mathbb{R}^{n+1}_+\} \cap \mathcal{R}(A)]^+ \\
&= \overline{\{p : (-A^\mathsf{T} p, b^\mathsf{T} p) \in \mathbb{R}^{n+1}_+\}^+ + \mathcal{R}(A)^+} && [(13.1)] \\
&= \{p : (-A^\mathsf{T} p, b^\mathsf{T} p) \in \mathbb{R}^{n+1}_+\}^+ + \mathcal{N}(A^\mathsf{T}) && [\text{convexe dense}] \\
&= \{-Ax + \alpha b : (x, \alpha) \in \mathbb{R}^{n+1}_+\} + \mathcal{N}(A^\mathsf{T}) && [\text{lemme de Farkas}].
\end{aligned}$$

  En prenant Ae comme élément de $\mathbb{R}^m$ à gauche de l'identité obtenue, on peut trouver $x_1 \geqslant 0$, $\alpha_0 \geqslant 0$ et $h \in \mathcal{N}(A^\mathsf{T})$ tels que $A(x_1 + e) = \alpha b + h$. Comme $b \in \mathcal{R}(A)$, il vient que $h \in \mathcal{N}(A^\mathsf{T}) \cap \mathcal{R}(A) = \{0\}$; donc $h = 0$ pour ce choix de $Ae \in \mathbb{R}^m$. Alors $x := (x_0 + x_1 + e)/(\alpha + 1)$ vérifie $x > 0$ et $Ax = b$.
- $[(c) \Rightarrow (a)]$ La fonction asymptotique de φ prend en $p \in \mathbb{R}^m$ la valeur $-b^\mathsf{T} p$ si $p \in \mathcal{R}(A)$ et $A^\mathsf{T} p \leqslant 0$ et prend la valeur $+\infty$ autrement. Supposons que $\varphi^\infty(p) \leqslant 0$. Alors $p \in \mathcal{R}(A)$, $A^\mathsf{T} p \leqslant 0$ et $b^\mathsf{T} p \geqslant 0$. Par (b), on peut trouver un $x > 0$ tel que $Ax = b$. En multipliant par p, on trouve $0 \leqslant b^\mathsf{T} p = p^\mathsf{T} Ax \leqslant 0$,

[3] La démonstration proosée de cette implication suppose que l'on a démontré (13.1) (exercice VIII.1) et que l'on sait qu'un convexe dense dans un espave vectoriel ne peut être que cet espace vectoriel (C convexe dans un espace vectoriel $\mathbb{E}$ et $\overline{C} = \mathbb{E} \Rightarrow C = \mathbb{E}$, c'est ce que l'on a entendu par la phrase « convexe dense »), une propriété probablement difficile à démontrer sans la notion d'intérieur relatif, non vue dans ce cours.

si bien que $A^\mathsf{T} p = 0$. Donc $p \in \mathcal{N}(A^\mathsf{T}) \cap \mathcal{R}(A)$, ce qui implique que $p = 0$. On conclut en utilisant l'équivalence $(iii) \Leftrightarrow (iv)$ de la proposition 3.28 du syllabus.

3 Poursuite de base ou recouvrement ℓ_1

1. • Si $b \notin \mathcal{R}(A)$, (P_1) n'a pas de solution (problème non réalisable).
   • Si $b \in \mathcal{R}(A)$, (P_1) a un ensemble admissible fermé et non vide et a un critère continu et coercif, donc le problème a une solution.

2. On peut en effet récrire successivement le problème (P_1) comme suit

$$\left\{\begin{array}{l} \inf_{x\in\mathbb{R}^n} \ \sum_{i=1}^n |x_i| \\ Ax = b. \end{array}\right.$$

Puis

$$\left\{\begin{array}{l} \inf_{(x,t)\in\mathbb{R}^n\times\mathbb{R}^n} \ \sum_{i=1}^n t_i \\ |x_i| \leqslant t_i \\ Ax = b. \end{array}\right.$$

Et enfin

$$\left\{\begin{array}{l} \inf_{(x,t)\in\mathbb{R}^n\times\mathbb{R}^n} \ \sum_{i=1}^n t_i \\ -t_1 \leqslant x_i \leqslant t_i \\ Ax = b, \end{array}\right.$$

qui est un problème d'optimisation linéaire (critère linéaire et contraintes affines).

3. • En prenant $x = 0$, on voit que l'on a toujours

$$\sup_{x\in\mathbb{R}^m} \Big(y^\mathsf{T} x - \|x\|_1 \Big) \geqslant 0.$$

   • Si $\|y\|_\infty \leqslant 1$, on a

$$\sup_{x\in\mathbb{R}^m} \Big(y^\mathsf{T} x - \|x\|_1 \Big) \leqslant \sup_{x\in\mathbb{R}^m} \left(\sum_i |y_i|\,|x_i| - \|x\|_1 \right) \leqslant \sup_{x\in\mathbb{R}^m} \left(\sum_i |x_i| - \|x\|_1 \right) \leqslant 0.$$

   On peut aussi utiliser directement Hölder.

   • Si $|y_j| > 1$, on a en prenant $x_i = t\,\mathrm{sgn}(y_j)\delta_{ij}$ avec $t > 0$:

$$\sup_{x\in\mathbb{R}^m} \Big(y^\mathsf{T} x - \|x\|_1 \Big) \geqslant t|y_j| - t = t(|y_j| - 1),$$

   qui tend vers l'infini lorsque $t \to \infty$.

4. Comme le problème (P_1) s'écrit aussi $\inf_{x\in\mathbb{R}^n} \sup_{y\in\mathbb{R}^m} (\|x\|_1 - y^\mathsf{T}(Ax - b))$, son dual lagrangien s'écrit

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^m} \inf_{x\in\mathbb{R}^n} \Big(\|x\|_1 - y^\mathsf{T}(Ax - b) \Big) \\
&= \sup_{y\in\mathbb{R}^m} \left(b^\mathsf{T} y - \sup_{x\in\mathbb{R}^n} \, (A^\mathsf{T} y)^\mathsf{T} x - \|x\|_1 \right) \\
&= \sup_{y\in\mathbb{R}^m} \Big(b^\mathsf{T} y - \mathcal{I}_{B_\infty}(A^\mathsf{T} y) \Big) \qquad [(13.5)] \\
&= \sup_{\|A^\mathsf{T} y\|_\infty \leqslant 1} \ b^\mathsf{T} y.
\end{aligned}$$

5. - $[\Rightarrow]$ Si $b \notin \mathcal{R}(A) = \mathcal{N}(A^\mathsf{T})^\perp$, la projection orthogonale b_0 de b sur $\mathcal{N}(A^\mathsf{T})$ est non nulle. Avec $y_t := tb_0$ où $t \in \mathbb{R}$, on a $A^\mathsf{T} y_t = 0$ (donc y_y est admissible pour (D_1)) et $b^\mathsf{T} y_t = t\|b_0\|_2^2$ tend vers $+\infty$ lorsque $t \to +\infty$. Donc (D_1) n'a pas de solution.
   - $[\Leftarrow]$ Remarquons d'abod que, si A n'est pas surjective (et on ne fait pas l'hypothèse de surjectivité de A), le critère de (D_1) n'est pas coercif sur son ensemble admissible. En effet, alors il existe $y_1 \in \mathcal{N}(A^\mathsf{T}) \setminus \{0\}$, ty_1 est admissible pour tout $t \in \mathbb{R}$ et $b^\mathsf{T}(ty_1) = 0$ (car $b \in \mathcal{R}(A) = \mathcal{N}(A^\mathsf{T})^\perp$ et $y_1 \in \mathcal{N}(A^\mathsf{T})$).

     Voici deux démonstrations de la réciproque.

     - (D_1) peut s'écrire comme un problème d'optimisation linéaire (objectif linéaire et contraintes $-e \leqslant A^\mathsf{T} y \leqslant e$), réalisable (par $y = 0$). Donc, s'il n'a pas de solution sa valeur optimale est $+\infty$. Par dualité faible, la valeur optimale de (P_1) est aussi $+\infty$, si bien que (P_1) n'est pas réalisable, c'est-à-dire $b \notin \mathcal{R}(A)$.
     - Si $b \in \mathcal{R}(A)$, il existe un x_0 tel que $b = Ax_0$ et la valeur optimale du problème (D_1) est celle du problème en $z \in \mathbb{R}^n$ obtenu en prenant $z = A^\mathsf{T} y$:

       $$\sup_{\|A^\mathsf{T} y\|_\infty \leqslant 1} x_0^\mathsf{T} A^\mathsf{T} y = \sup_{z \in \mathcal{R}(A^\mathsf{T}) \cap B_\infty} x_0^\mathsf{T} z.$$

       Comme $\mathcal{R}(A^\mathsf{T}) \cap B_\infty$ est compact non vide et $z \mapsto x_0^\mathsf{T} z$ est continue le problème en z a une solution, disons $\bar{z}$. Comme $\bar{z} \in \mathcal{R}(A^\mathsf{T})$, il existe un $\bar{y} \in \mathbb{R}^n$ tel que $\bar{z} = A^\mathsf{T}\bar{y}$ et $\bar{y}$ est alors solution de (D_1).

4 Convexification du compteur de composantes non nulles

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

   $$f^*(x^*) \geqslant 0. \tag{13.21}$$

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

   $$\begin{aligned} f^*(x^*) &\leqslant \sup_{x \in B_1} \underbrace{\|x^*\|_\infty \|x\|_1 - \|x\|_0}_{\begin{cases} = 0 & \text{si } x = 0 \\ \leqslant \|x^*\|_\infty - 1 & \text{si } x \in B_1 \setminus \{0\} \end{cases}} \\ &\leqslant (\|x^*\|_\infty - 1)^+. \end{aligned} \tag{13.22}$$

3. - Supposons d'abord que $\|x^*\|_\infty \leqslant 1$. En utilisant les inégalités (13.21) et (13.22), 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 (13.22) 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 tx_1^* avec $t > 0$, on voit que

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

5 Ensemble saillant de minimiseurs

1. L'ensemble S s'écrit aussi $S := \{x \in \mathbb{E} : f(x) \leqslant f_{\min}\}$. 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 $\|\bar{g}\| = \|(\bar{x}_0 + \alpha u) - \bar{x}\| \leqslant \|(\bar{x}_0 + \alpha u) - \bar{x}_0\| = \alpha\|u\| = \alpha$.
   - On prend le produit scalaire de $\bar{g} := \bar{x}_0 + \alpha u - \bar{x}$ avec $\bar{g}$ et on utilise le fait que $\langle \bar{g}, \bar{x}_0 - \bar{x}\rangle \leqslant 0$. On trouve $\|\bar{g}\|^2 \leqslant \alpha\langle \bar{g}, u\rangle \leqslant \alpha\|\bar{g}\|$ (Cauchy-Schwarz). Donc $\|\bar{g}\| \leqslant \alpha$.

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

$$\begin{aligned}
\langle \bar{g}, x - \bar{x}\rangle &= \langle \bar{g}, x - \mathrm{P}_S(x)\rangle + \langle \bar{g}, \mathrm{P}_S(x) - \bar{x}\rangle \\
&\leqslant \langle \bar{g}, x - \mathrm{P}_S(x)\rangle \qquad [\bar{x} = \mathrm{P}_S(\bar{x}_0 + \alpha u) \text{ et } \mathrm{P}_S(x) \in S] \\
&\leqslant \alpha\|x - \mathrm{P}_S(x)\| \qquad [\text{Cauchy-Schwarz et } \|\bar{g}\| \leqslant \alpha] \\
&\leqslant \alpha \operatorname{dist}(x, S) \qquad [\text{définition de } \operatorname{dist}(x, S)] \\
&\leqslant f(x) - f(\bar{x}) \qquad [(13.8)],
\end{aligned}$$

   qui est l'inégalité recherchée.

4. Par le point 3, $\bar{g} \in \partial f(\bar{x})$. Alors, on a montré que l'élément générique de l'ensemble à gauche dans (13.9), $\bar{x}_0 + \alpha u$, qui est égal à $\bar{x} + \bar{g}$, est donc dans l'ensemble à droite dans (13.9). Ceci démontre (13.9).

5. On utilise directement (13.9) avec $\bar{x} \in S$ et $(x-\bar{x})/\|x-\bar{x}\| \in B$.
6. On a
$$g_1 = \bar{x} - \bar{x}_1 + \alpha \frac{x-\bar{x}}{\|x-\bar{x}\|}. \tag{13.23}$$
Ensuite
$$\begin{array}{rcll} f(x) & \geqslant & f(\bar{x}_1) + \langle g_1, x-\bar{x}_1\rangle & [g_1 \in \partial f(\bar{x}_1)] \\ & = & f_{\min} + \dfrac{\alpha}{\|x-\bar{x}\|} \underbrace{\langle \bar{x}-x, \bar{x}_1 - x\rangle}_{\geqslant \|\bar{x}-x\|^2} + \underbrace{\langle \bar{x}_1-\bar{x}, \bar{x}_1 - x\rangle}_{\geqslant 0} & [f(\bar{x}_1) = f_{\min} \text{ et } (13.23)] \\ & \geqslant & f_{\min} + \alpha \|x-\bar{x}\| & [\text{caractérisations de la projection}] \\ & = & f_{\min} + \alpha \operatorname{dist}(x,S). & \end{array}$$
On obtient donc (13.8).
7. Comme $\varphi_x := f + \frac{1}{2}\|\cdot - x\|^2$ est fermée (car f l'est), le problème (13.10) 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
$$\begin{array}{rcl} \varphi_x(y) & \geqslant & \langle \xi, y\rangle + \alpha + \dfrac{1}{2}\|y-x\|^2 \\ & \geqslant & -\|\xi\|\,\|y\| + \alpha + \dfrac{1}{2}\|y\|^2 - \|y\|\,\|x\| + \dfrac{1}{2}\|x\|^2 \\ & \geqslant & \left(\dfrac{1}{2}\|y\| - \|\xi\| - \|x\|\right)\|y\| + \alpha + \dfrac{1}{2}\|x\|^2. \end{array}$$
On en déduit que $\varphi_x(y) \to +\infty$ lorsque $\|y\| \to \infty$ et donc que (13.10) 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 (13.9), il existe un $\bar{x} \in S$ tel que $x = \bar{x} + \bar{g}$, avec $\bar{g} \in \partial f(\bar{x})$, ce qui s'écrit aussi $0 \in \partial f(\bar{x}) + \bar{x} - x$. Ceci montre que $\bar{x}$ est l'unique solution de (13.10). Dès lors $x_+ = \bar{x}$ et donc $x_+ \in S$.

6 Sous-différentiel de la fonction duale

Proposition 13.25 ou notes manuscrites du cours.

7 Sous-différentiel de la fonction valeur

1. Sous-différentiel de la fonction marginale : proposition 3.68.
2. Sous-différentiel de la fonction valeur : proposition 4.48.

Bibliographie

[1] J.D. Blanchard, C. Cartis, J. Tanner (2011). Compressed sensing: how sharp is the restricted isometry property? *SIAM Review*, 53(1), 105–125. [doi].

[2] E.J. Candès, T. Tao (2005). Decoding by linear programming. *IEEE Transactions on Information Theory*, 51(11), 4203–4215. [doi].

[3] S.S. Chen, D.L. Donoho, M.A. Saunders (1998). Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20, 33–61. [doi].

[4] D.L. Donoho (2006). Compressed sensing. *IEEE Transactions on Information Theory*, 52(4), 1289–1306. [doi].

[5] Z. Harchaoui, A. Juditsky, A. Nemirovski (2014). Conditional gradient algorithms for norm-regularized smooth convex optimization. *Mathematical Programming*. [doi].