DM 1 — Optimisation Convexe

Titouan Le Breton

Exercice 1

1. On peut écrire
$$S = \{x \mid a^\top x \geq \alpha\} \cap \{x \mid a^\top x \leq \beta\},$$
c'est-à-dire l'intersection de deux demi-espaces convexes. Par conséquent
$$\boxed{S \text{ est convexe.}}$$

2. Chaque contrainte $x_i \leq \beta_i$ et $x_i \geq \alpha_i$ définit un demi-espace convexe. Leur intersection (un hyper-rectangle) est donc convexe :
$$\boxed{S \text{ est convexe.}}$$

3. C'est l'intersection de deux demi-espaces convexes. Par conséquent :
$$\boxed{S \text{ est convexe.}}$$

4. Pour un y fixé, la condition $\|x - x_0\|^2 \leq \|x - y\|^2$ équivaut à
$$2x^\top(y - x_0) \leq \|y\|^2 - \|x_0\|^2,$$
soit une inégalité linéaire en x. Pour chaque y, cela définit un demi-espace convexe, et S est leur intersection. Donc :
$$\boxed{S \text{ est convexe.}}$$

5. $$\boxed{\text{L'espace n'est pas convexe.}}$$
Contre-exemple :
Soit $S_1 = \{-1, 1\}$ et $S_2 = \{0\}$ dans $\mathbb{R}$. Alors pour $x = -2$ et $x = 2$, on a $\text{dist}(x, S_1) = 1 \leq 2 = \text{dist}(x, S_2)$, donc ces deux points appartiennent à S. Mais pour $x = 0$, $\text{dist}(0, S_1) = 1 > 0 = \text{dist}(0, S_2)$, donc $0 \notin S$. L'ensemble S n'est donc pas convexe.

6. Si $x, y \in C$ et $\lambda \in [0, 1]$, alors pour tout $s \in S_2$:
$$x + s \in S_1, \quad y + s \in S_1.$$
Par convexité de S_1,
$$\lambda(x + s) + (1 - \lambda)(y + s) = (\lambda x + (1 - \lambda)y) + s \in S_1.$$
Donc $(\lambda x + (1 - \lambda)y) + S_2 \subseteq S_1$, ce qui prouve que $\lambda x + (1 - \lambda)y \in C$. Ainsi :
$$\boxed{C \text{ est convexe..}}$$

7. **Cas $\boldsymbol{\theta = 1}$:** l'inégalité devient

$$\|x-a\| \leq \|x-b\|,$$

ce qui définit un hyperplan, donc l'espace est convexe.

Cas $\boldsymbol{0 \leq \theta < 1}$: En partant de

$$(1-\theta^2)x^\top x - 2(a-\theta^2 b)^\top x + (a^\top a - \theta^2 b^\top b) \leq 0,$$

on divise par $(1-\theta^2) > 0$ puis on complète le carré :

$$\begin{aligned}
& x^\top x - 2\left(\tfrac{a-\theta^2 b}{1-\theta^2}\right)^\top x + \tfrac{a^\top a - \theta^2 b^\top b}{1-\theta^2} \leq 0 \\
\iff & \left\|x - \tfrac{a-\theta^2 b}{1-\theta^2}\right\|^2 - \left\|\tfrac{a-\theta^2 b}{1-\theta^2}\right\|^2 + \tfrac{a^\top a - \theta^2 b^\top b}{1-\theta^2} \leq 0 \\
\iff & \left\|x - \tfrac{a-\theta^2 b}{1-\theta^2}\right\|^2 \;\leq\; \left\|\tfrac{a-\theta^2 b}{1-\theta^2}\right\|^2 - \tfrac{a^\top a - \theta^2 b^\top b}{1-\theta^2}.
\end{aligned}$$

Le membre de droite ne dépend pas de x et est positif. On obtient donc l'équation d'une boule fermée, donc l'espace est convexe.
On conclut :

$$\boxed{\text{L'espace est convexe.}}$$

Exercice 2

1. La norme $\|\cdot\|$ est convexe et la composition avec une application linéaire conserve la convexité, donc $x \mapsto \|A^{(i)}x - b^{(i)}\|$ est convexe pour chaque i. Le maximum de fonctions convexes étant convexe, on en déduit que

$$\boxed{g(x) = \max_i \|A^{(i)}x - b^{(i)}\| \quad \text{est convexe.}}$$

2. $$\forall\, i_1 < \cdots < i_r \in \{1, \ldots, n\}, \quad x \mapsto |x_{i_1}| + \cdots + |x_{i_r}| \text{ est convexe.}$$

Ainsi, on peut écrire

$$f(x) = \max_{1 \leq i_1 < \cdots < i_r \leq n} \left(|x_{i_1}| + \cdots + |x_{i_r}|\right).$$

Or, le maximum d'une famille de fonctions convexes est convexe.

$$\boxed{f \text{ est donc une fonction convexe.}}$$

Exercice 3

Soient $x, y \in I$ et $\lambda \in [0,1]$. Posons $z = \lambda x + (1-\lambda)y$. Par convexité de f et g, on a :

$$f(z) \leq \lambda f(x) + (1-\lambda)f(y), \quad g(z) \leq \lambda g(x) + (1-\lambda)g(y).$$

Comme $f, g \geq 0$, on peut multiplier ces inégalités :

$$f(z)g(z) \leq [\lambda f(x) + (1-\lambda)f(y)][\lambda g(x) + (1-\lambda)g(y)].$$

En développant :

$$f(z)g(z) \leq \lambda^2 f(x)g(x) + (1-\lambda)^2 f(y)g(y) + \lambda(1-\lambda)[f(x)g(y) + f(y)g(x)].$$

Or, puisque f et g sont de même monotonie,

$$(f(x)-f(y))(g(x)-g(y)) \geq 0 \quad \iff \quad f(x)g(y)+f(y)g(x) \leq f(x)g(x)+f(y)g(y).$$

En remplaçant dans l'inégalité précédente, on obtient :

$$f(z)g(z) \leq \lambda f(x)g(x) + (1-\lambda)f(y)g(y).$$

Ainsi :

$$\boxed{h(x) = f(x)g(x) \text{ est convexe}}$$

Exercice 4

4.1 Supposons d'abord qu'un des y_i soit négatif. Si $y_k < 0$ pour un certain k, on peut choisir $x = -te_k$ avec $t > 0$. On a alors $\max_i x_i = 0$ et $y^\top x = -ty_k \to +\infty$ lorsque $t \to +\infty$. Par conséquent, $f^*(y) = +\infty$. Ainsi, pour que $f^*(y)$ soit fini, il faut nécessairement $y_i \geq 0$ pour tout i.

Considérons ensuite la somme des composantes de y. Prenons $x = t\mathbf{1}$, avec $t > 0$. On obtient

$$\max_i x_i = t, \qquad y^\top x = t\sum_i y_i,$$

et donc

$$y^\top x - \max_i x_i = t\big(\sum_i y_i - 1\big).$$

Pour que la valeur du suprémum ne tende pas vers $+\infty$ quand $t \to +\infty$, il faut que $\sum_i y_i \leq 1$.

En revanche, si l'on prend $x = -t\mathbf{1}$ avec $t > 0$, on a

$$\max_i x_i = -t, \qquad y^\top x = -t\sum_i y_i,$$

et donc

$$y^\top x - \max_i x_i = -t\big(\sum_i y_i - 1\big).$$

Pour que ce terme ne tende pas vers $+\infty$ lorsque $t \to +\infty$, il faut cette fois que $\sum_i y_i \geq 1$. Les deux conditions combinées imposent donc $\sum_i y_i = 1$.

Enfin, supposons que $y_i \geq 0$ pour tout i et que $\sum_i y_i = 1$. Montrons que, dans ce cas, on a toujours $y^\top x - \max_i x_i \leq 0$.

Comme chaque y_i est positif ou nul, on peut majorer chaque produit $y_i x_i$ par y_i multiplié par la plus grande des composantes de x, c'est-à-dire par $\max_j x_j$. En effet, pour tout i, on a

$$x_i \leq \max_j x_j \quad \Rightarrow \quad y_i x_i \leq y_i \max_j x_j.$$

En additionnant ces inégalités pour $i = 1, \ldots, n$, on obtient

$$\sum_i y_i x_i \leq \sum_i y_i \max_j x_j = (\max_j x_j) \sum_i y_i.$$

Or, par hypothèse, $\sum_i y_i = 1$, donc

$$y^\top x = \sum_i y_i x_i \leq \max_j x_j.$$

Cela montre que

$$y^\top x - \max_i x_i \leq 0 \quad \text{pour tout } x \in \mathbb{R}^n.$$

En conclusion,

$$\boxed{f^*(y) = \begin{cases} 0, & \text{si } y_i \geq 0 \text{ et } \sum_i y_i = 1, \\ +\infty, & \text{sinon.} \end{cases}}$$

4.2 Prenons $x = t\mathbf{1}$ avec $t > 0$. Alors $\max_i x_i = t$ et, plus généralement, les r plus grandes composantes valent aussi t, donc

$$f(x) = \sum_{i=1}^{r} x_{[i]} = rt, \qquad y^\top x = t \sum_{i=1}^{n} y_i.$$

Ainsi

$$y^\top x - f(x) = t\Big(\sum_{i=1}^{n} y_i - r\Big).$$

Pour que le suprémum ne tende pas vers $+\infty$ quand $t \to +\infty$, il faut $\sum_{i=1}^{n} y_i \le r$.

En revanche, si l'on prend $x = -t\mathbf{1}$ avec $t > 0$, on a encore

$$f(x) = \sum_{i=1}^{r} x_{[i]} = r(-t) = -rt, \qquad y^\top x = -t \sum_{i=1}^{n} y_i,$$

et donc

$$y^\top x - f(x) = -t\Big(\sum_{i=1}^{n} y_i - r\Big).$$

Pour que cette quantité ne tende pas vers $+\infty$ lorsque $t \to +\infty$, il faut $\sum_{i=1}^{n} y_i \ge r$.

Les deux conditions combinées imposent donc $\sum_{i=1}^{n} y_i = r$.

Pour $x = te_i$ avec $t > 0$, on a $f(x) = t$ et $y^\top x = (y_i)t$, d'où

$$y^\top x - f(x) = (y_i - 1)t.$$

Pour éviter $+\infty$ quand $t \to +\infty$, il faut $y_i \le 1$.

Pour $x = -te_i$ avec $t > 0$, on a $f(x) = 0$ et $y^\top x = -y_i t$, d'où

$$y^\top x - f(x) = -y_i t.$$

Pour éviter $+\infty$ quand $t \to +\infty$, il faut $y_i \ge 0$.

Donc pour avoir $f^*(y) < +\infty$, on doit avoir $\quad 0 \le y_i \le 1$ pour tout i, $\quad$ et $\sum_{i=1}^{n} y_i = r$. Supposons désormais que $y \in \mathbb{R}^n$ vérifie $0 \le y_i \le 1$ pour tout i et $\sum_{i=1}^{n} y_i = r$. Nous montrons que pour tout $x \in \mathbb{R}^n$,

$$y^\top x \;\le\; \sum_{i=1}^{r} x_{[i]} \;=\; f(x),$$

ce qui impliquera $f^*(y) = \sup_x\{y^\top x - f(x)\} \le 0$, et comme $x = 0$ donne l'égalité, on aura $f^*(y) = 0$.

Par symétrie (réarrangement), on peut supposer $x_{[1]} \ge \cdots \ge x_{[n]}$. Considérons alors la quantité $\sum_{i=1}^{n} y_i x_{[i]}$ avec $0 \le y_i \le 1$ et $\sum_i y_i = r$. Si pour certains $i < j$ on a $y_i < 1$ et $y_j > 0$, transférer une masse $\delta = \min\{1 - y_i,\, y_j\}$ de l'indice j vers i augmente la somme de $\delta(x_{[i]} - x_{[j]}) \ge 0$. En itérant, on obtient que la somme est maximisée quand $y_1 = \cdots = y_r = 1$ et $y_{r+1} = \cdots = y_n = 0$. Ainsi, pour tout y admissible,

$$y^\top x \;=\; \sum_{i=1}^{n} y_i x_{[i]} \;\le\; \sum_{i=1}^{r} x_{[i]} \;=\; f(x).$$

Donc $y^\top x - f(x) \le 0$ pour tout x, d'où $\sup_x(y^\top x - f(x)) \le 0$. Comme $x = 0$ donne $y^\top 0 - f(0) = 0$, on conclut $f^*(y) = 0 \quad$ si $\quad 0 \le y_i \le 1\ \forall i,\ \ \sum_i y_i = r$.

En combinant avec la première partie, on a bien

$$\boxed{f^*(y) = \begin{cases} 0, & \text{si } y \in [0,1]^n \text{ et } \mathbf{1}^\top y = r, \\ +\infty, & \text{sinon.} \end{cases}}$$

4.3 Cas particulier : $m = 1$

Alors $f(x) = a_1 x + b_1$, donc

$$f^*(y) = \begin{cases} 0 & \text{si } y = a_1, \\ +\infty & \text{sinon.} \end{cases}$$

1er cas : $y < a_1$

Alors pour tout i, on a $y - a_i < 0$, donc

$$(y - a_i)x - b_i \to +\infty \quad \text{lorsque } x \to -\infty.$$

Ainsi $f^*(y) = +\infty$.

2ème cas : $y > a_m$

Cas symétrique : $(y - a_i)x - b_i \to +\infty$ lorsque $x \to +\infty$, donc $f^*(y) = +\infty$.

3ème cas : $a_i \leq y \leq a_{i+1}$

On remarque que si $a_i = a_{i+1}$ et $b_i > b_{i+1}$, alors pour tout x,

$$(y - a_i)x - b_i < (y - a_{i+1})x - b_{i+1},$$

on peut donc éliminer le terme i de la minimisation. On suppose donc désormais $a_1 < a_2 < \cdots < a_m$. .

Soient i, j tels que

$$y - a_i \geq 0, \qquad y - a_j \leq 0,$$

et que l'intersection entre les droites

$$(y - a_i)x - b_i \quad \text{et} \quad (y - a_j)x - b_j$$

soit la *plus petite* des intersections entre une droite de pente positive et une droite de pente négative.

En ce point,

$$x_{ij} = \frac{b_j - b_i}{a_i - a_j}.$$

Pour $x < x_{ij}$, toutes les droites actives ont pente positive, donc $\min_i((y - a_i)x - b_i)$ est croissante. Pour $x > x_{ij}$, les droites actives ont pente négative, donc la fonction est décroissante. Ainsi, le maximum du $\sup_x$ est atteint en $x = x_{ij}$.

On en déduit :

$$f^*(y) = (y - a_i)x_{ij} - b_i = (y - a_i)\frac{b_j - b_i}{a_i - a_j} - b_i,$$

où i, j sont tels que la condition ci-dessus est satisfaite.

$$\boxed{f^*(y) = \begin{cases} +\infty, & y < a_1 \text{ ou } y > a_m, \\ (y - a_i)\dfrac{b_j - b_i}{a_i - a_j} - b_i, & \text{sinon, pour } (i, j) \text{ réalisant la plus petite} \\ & \text{intersection entre droites de pentes opposées.} \end{cases}}$$