OPT202 – Optimisation Différentiable II
Contrôle des connaissances

- Durée : 3 h (on ne peut probablement pas tout faire... ; mieux vaut démontrer moins de choses très bien que beaucoup de choses approximativement).
- 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. *Éteindre tout appareil connecté à un réseau.*
- Ne pas écrire en rouge, couleur réservée aux observations des correcteurs.

Notations générales

- L'ensemble $[1:n]$ est celui des n premiers entiers non nuls $[1,n]\cap\mathbb{N}$ et $\overline{\mathbb{R}} := \mathbb{R}\cup\{-\infty,+\infty\}$.
- Les inégalités vectorielles doivent se comprendre composante par composante. Par exemple, si u et $v \in \mathbb{R}^n$, l'inégalité $u \leqslant v$ (resp. $u < v$) signifie que pour tout $i \in [1:n]$, $u_i \leqslant v_i$ (resp. $u_i < v_i$).
- Le produit scalaire d'un espace euclidien est toujours noté $\langle\cdot,\cdot\rangle$ et la norme associée $\|\cdot\|$.
- On note $\operatorname{Conv}(\mathbb{E})$ l'ensemble des fonctions convexes propres définies sur un espace vectoriel $\mathbb{E}$ à valeurs dans $\overline{\mathbb{R}}$.

Partie 1 : Questions diverses indépendantes

1. Un premier problème. On considère le problème d'optimisation en $x \in \mathbb{R}^n$ suivant

$$\min_{x\geqslant 0} \ \|x+a\|,$$

où $a \in \mathbb{R}^n$ et $\|\cdot\|$ est la norme euclidienne de $\mathbb{R}^n$. Calculez sa solution et montrez que sa valeur optimale est $\|a^+\|$, où $a^+ := \max(0,a)$ composante par composante. Justifiez bien vos réponses.

2. Calcul de solution. On considère le problème d'optimisation à deux variables $x = (x_1,x_2) \in \mathbb{R}^2$ suivant :

$$\begin{cases} \min_x \ (x_1-2)^2 + 3(x_1-2)(x_2-1) + (x_2-1)^2 \\ x_1 \geqslant 0 \\ x_2 \geqslant 0. \end{cases} \tag{1}$$

1. Montrez théoriquement que le problème (1) a au moins une solution (*théoriquement* veut dire ici *sans calculer une solution*, ce qui est le sujet du point suivant).
2. Calculer une solution *globale* du problème (1) en utilisant les conditions d'optimalité du premier ordre (dont on justifiera l'utilisation).

3. Minimisation d'une fonction linéaire sur une enveloppe convexe fermée. Soient $\mathbb{E}$ un espace euclidien, $c \in \mathbb{E}$ et X une partie non vide (*non nécessairement convexe*) de $\mathbb{E}$. On note $\overline{\operatorname{co}}\, P$ l'enveloppe convexe fermée d'une partie P de $\mathbb{E}$ et $\arg\min\{\langle c,x\rangle : x \in P\}$ l'ensemble des solutions du problème d'optimisation $\inf\{\langle c,x\rangle : x \in P\}$.

1. Montrez que
$$\inf_{x \in X} \langle c, x \rangle = \inf_{x \in \overline{\mathrm{co}}\, X} \langle c, x \rangle, \tag{2}$$
même si ces problèmes n'ont pas de solution.
2. Montrez que
$$\overline{\mathrm{co}} \left(\arg\min_{x \in X} \langle c, x \rangle \right) \subset \arg\min_{x \in \overline{\mathrm{co}}\, X} \langle c, x \rangle. \tag{3}$$

4. Sous-différentiabilité et sous-lipschitzianité. Soient $\mathbb{E}$ un espace euclidien, $f \in \mathrm{Conv}(\mathbb{E})$ et $x \in \operatorname{dom} f$. On cherche à montrer que f est sous-différentiable en x (c'est-à-dire que son sous-différentel $\partial f(x)$ est non vide) si, et seulement si,

$$\boxed{\forall\, y \in \mathbb{E} : \quad f(y) \geqslant f(x) - L\|y - x\|.} \tag{4}$$

1. Montrez que si $\partial f(x) \neq \varnothing$, alors il existe une constante $L \geqslant 0$ telle que (4) ait lieu.

On souhaite à présent démontrer la réciproque de l'affirmation précédente. On suppose donc que (4) a lieu pour une certaine constante $L \geqslant 0$. On introduit les ensembles suivants :

$$C_1 = \operatorname{epi} f, \qquad \text{et} \qquad C_2 = \{(y, \alpha) \in \mathbb{E} \times \mathbb{R} : \alpha < f(x) - L\|y - x\|\},$$

où $\operatorname{epi} f := \{(y, \alpha) \in \mathbb{E} \times \mathbb{R} : f(y) \leqslant \alpha\}$ désigne l'épigraphe de f.

2. Montrez que C_1 et C_2 sont des ensembles convexes.
3. Montrez que l'on peut trouver $(\xi, t) \in \mathbb{E} \times \mathbb{R}$ et $\gamma \in \mathbb{R}$ tels que $(\xi, t) \neq 0$ et
$$\forall (y_1, \alpha_1) \in C_1, \quad \forall (y_2, \alpha_2) \in C_2 : \quad \langle \xi, y_1 \rangle - t\alpha_1 \leqslant \gamma \leqslant \langle \xi, y_2 \rangle - t\alpha_2.$$
4. Montrez que pour les mêmes $(\xi, t) \in \mathbb{E} \times \mathbb{R}$ et $\gamma \in \mathbb{R}$, on a
$$\forall (y_1, \alpha_1) \in C_1, \quad \forall (y_2, \alpha_2) \in C_2 : \quad \langle \xi, y_1 \rangle - t\alpha_1 \leqslant \gamma < \langle \xi, y_2 \rangle - t\alpha_2.$$
où la deuxième inégalité est devenue stricte.
5. Montrez que $\gamma = \langle \xi, x \rangle - tf(x)$ et que $t > 0$.
6. Montrez que $\partial f(x) \neq 0$.

Partie 2 : Une représentation variationnelle du projecté cartésien d'un ensemble convexe

1. Des conditions pour qu'un problème dual lagrangien soit non borné

Soit $\mathbb{E}$ un espace euclidien. On considère le problème d'optimisation

$$(P) \quad \begin{cases} \min\ f(x) \\ c(x) \leqslant 0 \\ x \in X, \end{cases}$$

où $f : \mathbb{E} \to \mathbb{R}$ est convexe, les conposantes de $c : \mathbb{E} \to \mathbb{R}^m$ sont convexes et X est une partie non vide convexe de $\mathbb{E}$. On note

$$X^p := \{x \in X : c(x) + p \leqslant 0\}$$

l'ensemble admissible de (P), perturbé par le vecteur $p \in \mathbb{R}^m$. La *fonction valeur* associée à cette perturbation est la fonction $v : \mathbb{R}^m \to \overline{\mathbb{R}}$ qui prend en $p \in \mathbb{R}^m$ la valeur

$$v(p) := \inf_{x \in X^p} f(x).$$

Le *dual lagrangien* du problème (P) s'écrit

$$(D) \quad \sup_{\lambda \geqslant 0} \inf_{x \in X} f(x) + \lambda^{\mathsf{T}} c(x).$$

On note $\operatorname{val}(D)$ la valeur optimale du problème dual et $\delta : \mathbb{R}^m \to \overline{\mathbb{R}}$ la fonction duale, qui en $\lambda \in \mathbb{R}^m$ prend la valeur

$$\delta(\lambda) = -\left(\inf_{x \in X} f(x) + \lambda^{\mathsf{T}} c(x)\right).$$

On cherche à démontrer le lien suivant entre la fonction valeur de (P) et la valeur optimle du problème dual :

$$\boxed{0 \notin \overline{\operatorname{dom} v} \quad \text{et} \quad \operatorname{val}(D) > -\infty \qquad \Longrightarrow \qquad \operatorname{val}(D) = +\infty,} \tag{5}$$

où $\operatorname{dom} v := \{p \in \mathbb{R}^m : v(p) < +\infty\}$ est le domaine de v et $\overline{\operatorname{dom} v}$ est son adhérence.

Pour démontrer (5), on suppose donc que $0 \notin \overline{\mathscr{D}}$, où $\mathscr{D} := \operatorname{dom} v$, et que $\operatorname{val}(D) > -\infty$. On admettra que, sous les hypothèses énoncées, v est convexe (proposition 4.46 du syllabus).

1. Montrez qu'il existe $\mu \in \mathbb{R}^m$ tel que
$$\sup_{p \in \mathscr{D}} \mu^{\mathsf{T}} p < 0. \tag{6}$$

2. Montrez que $\mu \geqslant 0$.

3. Montrez que
$$\inf_{x \in X} \mu^{\mathsf{T}} c(x) > 0. \tag{7}$$

4. Montrez que l'on peut trouver un $\lambda \geqslant 0$ tel que $\delta(\lambda) \in \mathbb{R}$.

5. En utilisant les λ et $\mu \in \mathbb{R}^m_+$ trouvés aux points 1 et 4, montrez que $\mathrm{val}(D) = +\infty$.

2. Représentation variationnelle du projecté cartésien d'un ensemble convexe

On considère un ensemble défini dans un espace produit

$$\mathscr{E} := \{(x,y) \in X \times Y : c(x,y) \leqslant 0\},$$

où

- X est un *convexe* d'un espace vectoriel $\mathbb{E}$,
- Y est un ensemble non vide arbitraire,
- $c : X \times Y \to \mathbb{R}^m$ est une fonction *à valeurs finies*, telle que pour tout $y \in Y$:
  - $c(\cdot, y)$ est convexe,
  - $P_c := \{p \in \mathbb{R}^m : \text{il existe un } x \in X \text{ tel que } c(x,y) + p \leqslant 0\}$ est *fermé*.

Le projeté cartésien de $\mathscr{E}$ sur Y est noté et défini par

$$\mathscr{E}_Y := \{y \in Y : \text{il existe un } x \in X \text{ tel que } c(x,y) \leqslant 0\}.$$

On cherche à décrire « de manière variationnelle » l'ensemble $\mathscr{E}_Y$, c'est-à-dire par l'équivalence suivante : pour tout $y \in Y$,

$$\boxed{y \in \mathscr{E}_Y \quad\Longleftrightarrow\quad \sup_{\gamma \in \Delta_m} \inf_{x \in X} \gamma^{\mathsf{T}} c(x,y) \leqslant 0,} \tag{8}$$

où $\Delta_m := \{\gamma \in \mathbb{R}^m_+ : e^{\mathsf{T}}\gamma = 1\}$ est le simplexe unité de $\mathbb{R}^m$ (on a noté e le vecteur dont les composantes valent toutes 1).

6. Monrez l'implication "$\Rightarrow$" dans (8).

On cherche maintenant à démontrer l'implication réciproque "$\Leftarrow$" dans (8). On suppose donc que l'inégalité de droite dans (8) a lieu.

7. Montrez que
$$\sup_{\gamma \geqslant 0} \inf_{x \in X} \gamma^{\mathsf{T}} c(x,y) \leqslant 0. \tag{9}$$

8. Montrez que $y \in \mathscr{E}_Y$.

   [*Indication* : On pourra utiliser l'implication (5), qui requiert l'introduction d'un certain problème d'optimisation.]

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

Partie 1 : Questions diverses indépendantes

1. Voici deux solutions. La première suppose que l'on sait ce qu'est la projection sur des contraintes de bornes (exercice IX.2). La seconde, passant par les conditions d'optimalité, est un peu plus laborieuse.

- Comme le problème s'écrit
$$\min_{x \geqslant 0} \|x - (-a)\|,$$
il consiste à projeter $(-a)$ sur l'orthant positif. Sa solution est donc $(-a)^+ = \max(0, -a) =: a^-$ et sa valeur optimale est $\|a^- - (-a)\| = \|a^- + a\| = \|a^- + (a^+ - a^-)\| = \|a^+\|$.
- On écrit les conditions d'optmalité du problème (il a une solution et les contraintes sont qualifiées), en utilisant le lagrangien $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \frac{1}{2}\|x + a\|^2 - \lambda^\mathsf{T} x$, on obtient
$$x + a = \lambda \qquad \text{et} \qquad 0 \leqslant \lambda \perp x \geqslant 0.$$
Ensuite on peut raisonner de plusieurs manières ; par exemple, en fonction du signe des conposantes de a :
  - $a_i \geqslant 0 \Rightarrow x_i = 0$ (sinon $x_i > 0$ [pas d'autre possibilité] et $\lambda_i > 0$ [par la première condition], et la complémentarité n'aurait pas lieu),
  - $a_i < 0 \Rightarrow x_i \geqslant -a_i > 0 \Rightarrow \lambda_i = 0 \Rightarrow x_i = -a_i$.

  On a montré que la solution est $\max(0, -a) =: a^-$. Comme précédemment, la valeur optimale est alors $\|a^- - (-a)\| = \|a^- + a\| = \|a^- + (a^+ - a^-)\| = \|a^+\|$.

2.

1. L'ensemble admissible étant un fermé non vide (mais non borné) et le critère, disons f, étant semi-continu inférieurement (il est continu comme polynôme), il suffit de montrer que le critère est coercif sur $\mathbb{R}^2_+$.
Supposons donc que $\|x\| \to \infty$ dans l'ensemble admissible. Alors soit $x_1 \to \infty$, soit $x_2 \to \infty$. Suppsons que $x_1 \to \infty$ (étant donné la forme du critère, le raisonnement est le même lorsque $x_2 \to \infty$). Alors $x_1 - 2 \geqslant 0$ dès que x_1 est assez grand, et donc $3(x_1 - 2)(x_2 - 1) \geqslant -3(x_1 - 2)$, si bien que
$$f(x) \geqslant (x_1 - 2)^2 - 3(x_1 - 2).$$
Donc $f(x) \to \infty$.
2. On utilise le lagrangien $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ défini en $(x, \lambda) \in \mathbb{R}^2 \times \mathbb{R}^2$ par
$$\ell(x, \lambda) = (x_1 - 2)^2 + 3(x_1 - 2)(x_2 - 1) + (x_2 - 1)^2 - \lambda_1 x_1 - \lambda_2 x_2.$$
Comme les contraintes sont qualifiées (car affines), il existe un multiplicateur $\lambda \in \mathbb{R}^2$ associé à la solution $x \in \mathbb{R}^2$ tel que
$$\left\{\begin{array}{l} 2x_1 + 3x_2 - \lambda_1 = 7 \\ 3x_1 + 2x_2 - \lambda_2 = 8 \\ 0 \leqslant \lambda \perp x \geqslant 0. \end{array}\right.$$
Calculons des points stationnaires (ils sont représentés à la figure 1). On choisit de considérer les quatre cas possibles en fonction de la valeur prise par les multiplicateurs.

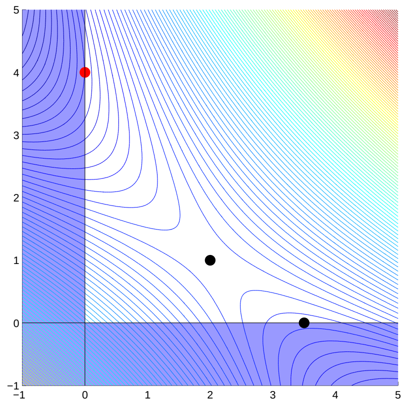

Figure 1: La solution (point rouge) du problème (1), l'autre point stationnaire (point noir) et les courbes de niveau du critère (qui est coercif sur l'ensemble admissible, mais pas sur $\mathbb{R}^2$)

- $\boldsymbol{\lambda_1 = \lambda_2 = 0}$. Alors $x = (2, 1)$, qui est strictement admissible. C'est donc un point stationnaire, mais ce n'est pas un minimum local car
$$\nabla^2 f(x) = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix} \not\succcurlyeq 0.$$
- $\boldsymbol{\lambda_1 > 0}$ **et** $\boldsymbol{\lambda_2 = 0}$. Alors $x = (0, 4)$, qui est admissible, et $\lambda = (5, 0) \geqslant 0$. C'est donc un point stationnaire. La valeur du critère en ce point vaut -5.
- $\boldsymbol{\lambda_1 = 0}$ **et** $\boldsymbol{\lambda_2 > 0}$. Alors $x = (7/2, 0)$, qui est admissible, et $\lambda = (0, 5/2) \geqslant 0$. C'est donc un point stationnaire. La valeur du critère en ce point vaut $-5/4$.
- $\boldsymbol{\lambda_1 > 0}$ **et** $\boldsymbol{\lambda_2 > 0}$. Alors $x = 0$, mais $\lambda = (-7, -8)$ n'est pas $\geqslant 0$. Ce n'est donc pas un point stationnaire.

Comme on a calculé tous les points stationnaires et que la solution (globale) en fait partie, il suffit de comparer les valeurs du critère en ces points pour obtenir la solution (un algorithme ne calculerait qu'un seul point stationnaire, d'où l'utilité des conditions du second ordre pour détecter la minimalité locale). La solution est donc $\bar{x} = (0, 4)$ et la valeur optimale est -5.

3. 1. Comme $X \subset \overline{\mathrm{co}}\, X$, on a certainement
$$\inf\{\langle c, x\rangle : x \in X\} \geqslant \inf\{\langle c, x\rangle : x \in \overline{\mathrm{co}}\, X\}. \tag{10}$$
Voici trois manières de raisonner, de la plus courte à la plus longue. On note
$$\alpha := \inf\{\langle c, x\rangle : x \in X\} \qquad \text{et} \qquad H^+(c, \alpha) := \{x \in \mathbb{E} : \langle c, x\rangle \geqslant \alpha\}.$$

- Pour tout $x \in X$, on a $\langle c, x\rangle \geqslant \alpha$, ce qui s'écrit $X \subset H^+(c, \alpha)$. Comme $H^+(c, \alpha)$ est un demi-espace fermé, on a aussi $\overline{\mathrm{co}}\, X \subset H^+(c, \alpha)$. Cette dernière inclusion implique que $\inf\{\langle c, x\rangle : x \in \overline{\mathrm{co}}\, X\} \geqslant \alpha = \inf\{\langle c, x\rangle : x \in X\}$.
- On raisonne par l'absurde en supposant que l'inégalité précédente est stricte. Alors, on peut trouver $x_0 \in \overline{\mathrm{co}}\, X$ tel que
$$\langle c, x_0\rangle < \alpha. \tag{11}$$
Ceci signifie que le demi-espace fermé $H^+(c, \alpha)$ contient X mais pas x_0 :
$$X \subset H^+(c, \alpha) \qquad \text{et} \qquad x_0 \notin H^+(c, \alpha). \tag{12}$$
De la première affirmation de (12) on déduit que $\overline{\mathrm{co}}\, X \subset H^+(c, \alpha)$ et donc $x_0 \in \overline{\mathrm{co}}\, X \subset H^+(c, \alpha)$, ce qui est en contradiction avec la seconde affirmation de (12).

- On montre d'abord que
$$\inf\{\langle c,x\rangle : x \in \operatorname{co} X\} \geqslant \alpha. \tag{13}$$
Tout élément de $x \in \operatorname{co} X$ s'écrit comme une somme finie $x = \sum_i t_i x_i$ avec les $t_i \geqslant 0$ et $\sum_i t_i = 1$ et des $x_i \in X$ (non vu au cours mais apparemment connu de beaucoup ; attention à ne pas limiter le nombre de x_i à deux). Comme on a $\langle c,x_i\rangle \geqslant \alpha$, on en déduit que $\langle c,x\rangle = \sum_i t_i\langle c,x_i\rangle \geqslant \sum_i t_i\alpha = \alpha$. On a donc démontré (13).

  On utilise ensuite $\overline{\operatorname{co}}\, X = \overline{\operatorname{co} X}$ (mentionné au cours sans démonstration) et le fait que minimiser sur un ensemble ou sur son adhérence donne la même valeur optimale (utilisation d'une suite).

2. Soient $\alpha := \inf\{\langle c,x\rangle : x \in \overline{\operatorname{co}}\, X\}$ et $H^=(c,\alpha) := \{x \in \mathbb{E} : \langle c,x\rangle = \alpha\}$. Alors
$$\operatorname*{arg\,min}_{x\in\overline{\operatorname{co}}\, X} \langle c,x\rangle = (\overline{\operatorname{co}}\, X) \cap H^=(c,\alpha), \tag{14}$$
si bien que l'ensemble à droite dans (14) (et donc $\operatorname{arg\,min}\{\langle c,x\rangle : x \in \overline{\operatorname{co}}\, X\}$) est un convexe fermé (éventuellement vide). Il suffit donc de montrer qu'il contient $\operatorname{arg\,min}\{\langle c,x\rangle : x \in X\}$. Voici deux manières de raisonner.

   - Soit $\bar{x} \in \operatorname{arg\,min}\{\langle c,x\rangle : x \in X\}$. Alors $\langle c,\bar{x}\rangle = \min\{\langle c,x\rangle : x \in X\} = \inf\{\langle c,x\rangle : x \in \overline{\operatorname{co}}\, X\}$ (par le point 1). Comme $\bar{x} \in X \subset \overline{\operatorname{co}}\, X$, il advient que $\bar{x} \in \operatorname{arg\,min}\{\langle c,x\rangle : x \in \overline{\operatorname{co}}\, X\}$.
   - Soit $\bar{x} \in \operatorname{arg\,min}\{\langle c,x\rangle : x \in X\}$, si bien que $\langle c,\bar{x}\rangle = \alpha$. Alors $H^+(c,\alpha)$ est un demi-espace fermé qui contient X, donc qui contient $\overline{\operatorname{co}}\, X$, ce qui implique que $\langle c,\bar{x}\rangle \leqslant \langle c,x\rangle$ pour tout $x \in \overline{\operatorname{co}}\, X$. Comme $\bar{x} \in X \subset \overline{\operatorname{co}}\, X$, il advient que $\bar{x} \in \operatorname{arg\,min}\{\langle c,x\rangle : x \in \overline{\operatorname{co}}\, X\}$.

4.
1. Soit en effet $x^* \in \partial f(x)$, supposé non vide. Alors, en utilisant Cauchy-Schwarz, on a que, pour tout $y \in \mathbb{E}$,
$$f(y) \geqslant f(x) + \langle x^*, y-x\rangle \geqslant f(x) - L\|y-x\|,$$
où $L := \|x^*\|$.
2. L'ensemble C_1 est convexe comme épigraphe d'une fonction convexe (c'est une définition). De même, l'application $(y,\alpha) \in \mathbb{E}\times\mathbb{R} \mapsto \alpha + L\|y-x\|$ étant convexe (raisons : somme de deux fonctions convexes, une norme est convexe, la composition d'une application affine et d'une fonction convexe est convexe), C_2 est convexe comme ensemble de sous-niveau stricte de cette fonction.
3. Lorsque (4) a lieu, les ensembles C_1 et C_2 sont convexes disjoints, donc on peut les séparer. C'est ce qui est fait.
4. L'inégalité stricte est justifiée par le fait que C_2 est ouvert. En effet, supposons que l'on ait l'égalité pour un certain $(y_2,\alpha_2) \in C_2$: $\gamma = \langle \xi, y_2\rangle - t\alpha_2$. Comme C_2 est ouvert, le couple $(y_2 - \varepsilon\xi, \alpha_2 + \varepsilon t)$ serait encore dans C_2 pour $\varepsilon > 0$ petit. Alors, on aurait
$$\gamma \leqslant \langle \xi, y_2 - \varepsilon\xi\rangle - t(\alpha_2 + \varepsilon t) = \gamma - \varepsilon(\|\xi\|^2 + t^2).$$
Ceci contredirait le fait que $(\xi,t) \neq 0$.
5. On exploite les inégalités du point 4.
   - En prenant $(x, f(x)) \in C_1$, on obtient $\langle \xi, x\rangle - tf(x) \leqslant \gamma$. En prenant $(x, f(x)-\varepsilon) \in C_2$, avec $\varepsilon > 0$ arbitraire, on obtient $\gamma \leqslant \langle \xi, x\rangle - tf(x)$. Donc $\gamma = \langle \xi, x\rangle - tf(x)$.
   - Alors, en prenant $(x, f(x)-1) \in C_2$, on obtient maintenant $t > 0$.
6. En introduisant $x^* := \xi/t$, la première inégalité devient
$$\forall (y_1,\alpha_1) \in C_1 : \quad f(x) + \langle x^*, y_1 - x\rangle \leqslant \alpha_1.$$
Soit $y \in \mathbb{E}$ quelconque. En prenant $(y, f(y)) \in C_1$, on obtient $f(x) + \langle x^*, y-x\rangle \leqslant f(y)$, qui implique que x^* est un sous-gradient de f en x.

Partie 2 : Une représentation variationnelle du projecté cartésien d'un ensemble convexe

1. Comme v est convexe, son domaine $\mathscr{D}$ est aussi convexe. Maintenant, le convexe compact $\{0\}$ et le convexe fermé $\overline{\mathscr{D}}$ sont disjoints et peuvent donc être strictement séparés : $\exists\, \mu \in \mathbb{R}^m$ tel que

$$\sup_{p\in\overline{\mathscr{D}}} \mu^\mathsf{T} p < \inf_{p\in\{0\}} \mu^\mathsf{T} p = 0.$$

   Comme $\sup\{\mu^\mathsf{T} p : p \in \mathscr{D}\} \leqslant \sup\{\mu^\mathsf{T} p : p \in \overline{\mathscr{D}}\}$ (en réalité on a l'égalité), l'inégalité (6) a bien lieu.

2. En prenant $x \in X$ (qui est non vide), $i \in [1:m]$, $t \geqslant 0$ et $p := -c(x) - te^i$, on voit que $x \in X^p$. Alors $v(p) < +\infty$, ce qui s'écrit aussi $p \in \mathscr{D}$. Dès lors $\mu^\mathsf{T} p \leqslant 0$ ou $-\mu^\mathsf{T} c(x) - t\mu_i \leqslant 0$ pour tout $t \geqslant 0$, dont on déduit que $\mu_i \geqslant 0$.

3. Pour $x \in X$, $p := -c(x) \in \mathscr{D}$. On déduit alors de (6)

$$0 > \sup_{p\in\mathscr{D}} \mu^\mathsf{T} p \geqslant \sup_{x\in X} -\mu^\mathsf{T} c(x) = -\inf_{x\in X} \mu^\mathsf{T} c(x),$$

   qui n'est autre que (7).

4. Par la définition de δ et $X \neq \varnothing$, on a $\delta > -\infty$.

   Par ailleurs, $\operatorname{val}(D) \neq -\infty$ implique qu'il existe un $\lambda \geqslant 0$ tel que $\delta(\lambda) < +\infty$. Pour ce $\lambda \geqslant 0$, on a donc $\delta(\lambda) \in \mathbb{R}$.

5. Soit $\lambda \in \operatorname{dom}\delta$, donné par le point précédent, et μ, donné par le point 1. Alors pour tout $t \geqslant 0$, il vient

$$\begin{aligned} \delta(\lambda + t\mu) &= -\inf_{x\in X} f(x) + (\lambda + t\mu)^\mathsf{T} c(x) \\ &\leqslant \underbrace{-\inf_{x\in X} \ell(x,\lambda)}_{\delta(\lambda)\in\mathbb{R}} - t \underbrace{\inf_{x\in X} \mu^\mathsf{T} c(x)}_{>0 \text{ par } (7)}, \end{aligned}$$

   qui tend donc vers $-\infty$ lorsque $t \to +\infty$. Comme $\lambda + t\mu \geqslant 0$, on obtient $\operatorname{val}(D) = +\infty$.

6. Soit $y \in \mathscr{E}_Y$. Par définition de $\mathscr{E}_Y$, il existe $x_y \in X$ tel que $c(x_y, y) \leqslant 0$. Comme les $\gamma \in \Delta_n$ ont leurs composantes positives, on en déduit que $\gamma^\mathsf{T} c(x_y, y) \leqslant 0$. On a donc montré que

$$\exists\, x \in X, \quad \forall\, \gamma \in \Delta_m, \quad \gamma^\mathsf{T} c(x,y) \leqslant 0. \tag{15}$$

   On peut ensuite raisonner de deux manières différentes.

   - L'affirmation (15) implique que

$$\inf_{x\in X} \sup_{\gamma\in\Delta_m} \gamma^\mathsf{T} c(x,y) \leqslant 0.$$

     Par dualité faible, on obtient le résultat.

- L'affirmation (15) implique aussi certainement que

$$\forall\, \gamma \in \Delta_m, \quad \exists\, x \in X, \quad \gamma^\mathsf{T} c(x,y) \leqslant 0$$

et donc

$$\sup_{\gamma \in \Delta_m} \inf_{x \in X} \gamma^\mathsf{T} c(x,y) \leqslant 0.$$

7. Si ce n'était pas le cas, il existerait un $\gamma_0 \geqslant 0$ tel que $\inf_{x\in X} \gamma_0^\mathsf{T} c(x,y) > 0$. Alors ce γ_0 est positif non nul, si bien que $e^\mathsf{T}\gamma_0 > 0$ et en introduisant $\gamma_1 := \gamma_0/(e^\mathsf{T}\gamma_0)$, qui est dans Δ_m, on aurait

$$\inf_{x\in X} \gamma_1^\mathsf{T} c(x,y) = \frac{1}{e^\mathsf{T}\gamma_0} \left(\inf_{x\in X} \gamma_0^\mathsf{T} c(x,y) \right) > 0.$$

Ceci contredirait l'hypothèse du cas.

8. Par (9), le problème dual de

$$\begin{cases} \inf_{x\in X} \ 0 \\ c(x,y) \leqslant 0, \end{cases} \tag{16}$$

a une valeur optimale nulle (obtenue en prenant $\gamma = 0$), donc finie. Dès lors, 0 adhère au domaine de la fonction valeur du problème (16) (implication (5), X convexe et $c(\cdot,y)$ convexe quel que soit $y \in Y$), qui est P_c. Comme P_c est fermé par hypothèse (elle n'intervient qu'ici), il vient $0 \in P_c$, ce qui s'écrit aussi $y \in \mathscr{E}_Y$.