OPT201 – Optimisation Différentiable I

Contrôle des connaissances

- Durée : 1 h 30 (on ne peut probablement pas tout faire...).
- Dans chaque sujet, on peut admettre des résultats énoncés dans des questions ou sous-questions *précédentes* pour traiter la question ou sous-question courante.
- Ne pas écrire en rouge.
- On peut consulter tous les documents *distribués au cours*, ainsi que ses propres notes.
- *Éteindre tout appareil connecté à un réseau.*

Trois sujets indépendants

1. Fonction d'appui. Soient $\mathbb{E}$ un espace euclidien (produit scalaire noté $\langle\cdot,\cdot\rangle$). On note $\|\cdot\|$ une norme sur $\mathbb{E}$ (pas nécessairement celle associée au produit scalaire) et $B = \{x \in \mathbb{E} : \|x\| \leqslant 1\}$ la boule unité associée. On note aussi $\|\cdot\|_D$ la norme duale de $\|\cdot\|$ pour le produit scalaire $\langle\cdot,\cdot\rangle$ et $B_{\text{D}} := \{d \in \mathbb{E} : \|d\|_D \leqslant 1\}$ la boule unité associée. La *fonction d'appui* de B est la fonction $\sigma_B : \mathbb{E} \to \mathbb{R}$ définie en $d \in \mathbb{E}$ par

$$\sigma_B(d) = \sup_{x \in B} \langle d, x\rangle.$$

Montrez que l'ensemble de sous-niveau 1 de σ_B est la boule-unité duale, ce qui s'écrit

$$\{d \in \mathbb{E} : \sigma_B(d) \leqslant 1\} = B_{\text{D}}. \tag{1}$$

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

$$\begin{cases} \min_x \; (x_1 - 1)^2 + 2(x_2 - 1)^2 + 5x_1x_2 \\ 2x_2 \geqslant x_1 \geqslant 0. \end{cases} \tag{2}$$

1. Montrez théoriquement que ce problème a au moins une solution (*théoriquement* veut dire ici *sans calculer une solution*, ce qui est le sujet des points suivants, mais par des arguments de nature topologique).
2. Montrez que les contraintes du problème sont qualifiées en tout point admissible.
3. Calculez tous les points stationnaires du problème, composantes primales et duales (c'est-à-dire, les solutions des conditions d'optimalité du premier ordre).
4. Déterminez la ou les solutions (globales) du problème.

3. Dualité forte pour fonction séparable. Pour un entier $p \in \mathbb{N}_*$, on suppose donnés un ensemble X et p fonctions $\varphi_i : X \to \overline{\mathbb{R}}$.

1. Justifiez l'inégalité suivante (en particulier l'utilisation du "max" plutôt que du "sup")

$$\inf_{x \in X} \max_{i \in [1:p]} \varphi_i(x) \geqslant \max_{i \in [1:p]} \inf_{x \in X} \varphi_i(x) \tag{3}$$

et donnez un exemple où l'on n'a pas l'égalité.

Le problème à gauche (resp. à droite) dans (3) est appelé *problème primal* (resp. *problème dual*). Une *solution primale* (resp. *solution duale*) est une solution du problème primal (resp. dual).

Hypothèse de séparabilité. On suppose dorénavant que X est un produit Cartésien $X = X_1 \times \cdots \times X_p$ de p ensembles X_i et que chaque fonction φ_i ne dépend que de la i-ième composante $x_i \in X_i$ de $x = (x_1, \ldots, x_p)$.

2. On cherche à montrer que

$$\inf_{x \in X} \max_{i \in [1:p]} \varphi_i(x_i) = \max_{i \in [1:p]} \inf_{x_i \in X_i} \varphi_i(x_i). \tag{4}$$

   (a) Soit $\varepsilon > 0$. Montrez que, pour tout $i \in [1:p]$, on peut trouver $x_i^\varepsilon \in X_i$ tel que

$$\max_{i \in [1:p]} \varphi_i(x_i^\varepsilon) \leqslant \max_{i \in [1:p]} \inf_{x_i \in X_i} \varphi_i(x_i) + \varepsilon. \tag{5}$$

   (b) Montrez que l'on a (4) (bien justifier votre réponse).

On note $\bar{v}$ la valeur commune des deux membres de l'identité (4).

3. Montrez que l'ensemble des solutions primales s'écrit $\bar{X} := \bar{X}_1 \times \cdots \times \bar{X}_p$, où

$$\bar{X}_i := \{x_i \in X_i : \varphi_i(x_i) \leqslant \bar{v}\}. \tag{6}$$

4. Montrez que l'ensemble des solutions duales s'écrit

$$\bar{I} := \{i \in [1:p] : \varphi_i(x_i) \geqslant \bar{v} \text{ pour tout } x_i \in X_i\}. \tag{7}$$

5. Supposons que, pour toute solution duale $\bar{\imath} \in \bar{I}$, le problème $\inf\{\varphi_{\bar{\imath}}(x_{\bar{\imath}}) : x_{\bar{\imath}} \in X_{\bar{\imath}}\}$ a une solution $\hat{x}_{\bar{\imath}}$. Montrez qu'alors on peut trouver un point $\bar{x} \in X$ tel

$$\bar{x} \in \bar{X}. \tag{8a}$$
$$\bar{x}_{\bar{\imath}} = \hat{x}_{\bar{\imath}} \text{ pour tout } \bar{\imath} \in \bar{I}, \tag{8b}$$
$$\operatorname*{Arg\,max}_{i \in [1:p]} \varphi_i(\bar{x}) = \bar{I}. \tag{8c}$$

6. Supposons que $\bar{X} \neq \varnothing$. Alors,
   (a) $\bar{\imath} \in \bar{I}$ si, et seulement si, pour tout $\bar{x} \in \bar{X}$, $\bar{\imath}$ maximise $i \in [1:p] \mapsto \varphi_i(\bar{x}_i)$,
   (b) $\forall\, \bar{\imath} \in \bar{I}$, $\forall\, \bar{x} \in \bar{X}$, $\bar{x}_{\bar{\imath}}$ minimise $\varphi_{\bar{\imath}}$ sur $X_{\bar{\imath}}$ et $\varphi_{\bar{\imath}}(\bar{x}_{\bar{\imath}}) = \bar{v}$.

OPT201 – Optimisation Différentiable I

Contrôle des connaissances (solutions succinctes)

1. On a les équivalences suivantes:

$$\sigma_B(d) \leqslant 1 \iff \sup_{x \in B} \langle d, x \rangle \leqslant 1 \iff \|d\|_D \leqslant 1 \iff d \in B_{\mathrm{D}}.$$

On en déduit l'identité (1).

2.

1.
   - L'ensemble admissible X est non vide ($0 \in X$) et fermé (défini par des inégalités au sens large et continues), mais non borné ($(0, t) \in X$ pour $t \to \infty$).
   - Le critère f est continu (c'est un polynôme), mais n'est pas coercif sur $\mathbb{R}^2$ ($f(t, -t) \to -\infty$ lorsque $t \to +\infty$)
   - f est coercif sur $\mathbb{R}^2_+$ ($(x_1 - 1)^2 + 2(x_2 - 1)^2 + 5x_1x_2 \geqslant \|x - e\|^2$ pour la norme $x \mapsto \|x\| = \sqrt{x_1^2 + 2x_2^2}$) donc sur l'ensemble admissible (qui est contenu dans $\mathbb{R}^2_+$).

   Donc le problème a au moins une solution (corollaire du théorème de Weierstrass).
2. L'ensemble admissible est décrit par des contraintes linéaires ((QC-A) a lieu).
3. Le problème et sa solution sont représentés à la figure 1.

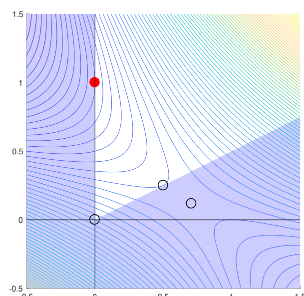

Figure 1: L'ensemble admissible du problème (en clair), sa solution (point rouge) et les courbes de niveau du critère. Les points cerclés sont des candidats aux conditions d'optimalité, mais qui ne vérifient pas les contraintes (pour $x = (12, 2)/17$) ou la positivité des multiplicateurs (pour $x = 0$ et $x = (1/2, 1/4)$).

Le lagrangien est la fonction $\ell : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie en $(x, \lambda) \in \mathbb{R}^2 \times \mathbb{R}^2$ par

$$\ell(x, \lambda) = (x_1 - 1)^2 + 2(x_2 - 1)^2 + 5x_1x_2 - \lambda_1 x_1 + \lambda_2(x_1 - 2x_2).$$

Les conditions de KKT s'écrivent

$$\begin{cases} 2(x_1 - 1) + 5x_2 - \lambda_1 + \lambda_2 = 0 \\ 4(x_2 - 1) + 5x_1 - 2\lambda_2 = 0 \\ 0 \leqslant \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \perp \begin{pmatrix} x_1 \\ 2x_2 - x_1 \end{pmatrix} \geqslant 0. \end{cases}$$

- **Cas 1 :** $\lambda_1 = 0$ et $\lambda_2 = 0$.

Il faut alors résoudre

$$\begin{cases} 2(x_1-1)+5x_2=0 \\ 4(x_2-1)+5x_1=0 \\ 0 \leqslant \begin{pmatrix} 0 \\ 0 \end{pmatrix} \perp \begin{pmatrix} x_1 \\ 2x_2-x_1 \end{pmatrix} \geqslant 0. \end{cases}$$

Les deux premières équations donnent $x=(12,2)/17$, ce qui n'est pas compatible avec la contrainte $2x_2 \geqslant x_1$.

- **Cas 2 :** $\lambda_1>0$ et $\lambda_2=0$.

  Il faut alors résoudre

$$\begin{cases} 2(x_1-1)+5x_2-\lambda_1=0 \\ 4(x_2-1)+5x_1=0 \\ x_1=0 \\ 0 \leqslant \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 0 \\ 2x_2-x_1 \end{pmatrix} \geqslant 0. \end{cases}$$

  Les trois premières équations donnent $x=(0,1)$ et $\lambda=(3,0)$, ce qui est compatible avec la complémentarité.

- **Cas 3 :** $\lambda_1=0$ et $\lambda_2>0$.

  Il faut alors résoudre

$$\begin{cases} 2(x_1-1)+5x_2+\lambda_2=0 \\ 4(x_2-1)+5x_1-2\lambda_2=0 \\ 2x_2=x_1 \\ 0 \leqslant \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix} \perp \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \geqslant 0. \end{cases}$$

  Les trois premières équations donnent $x=(1/2,1/4)$ et $\lambda=(0,-1/4)$, ce qui n'est pas compatible avec la positivité requise des multiplicateurs.

- **Cas 4 :** $\lambda_1>0$ et $\lambda_2>0$.

  Il faut alors résoudre

$$\begin{cases} 2(x_1-1)+5x_2-\lambda_1+\lambda_2=0 \\ 4(x_2-1)+5x_1-2\lambda_2=0 \\ x_1=x_2=0 \\ 0 \leqslant \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \perp \begin{pmatrix} 0 \\ 0 \end{pmatrix} \geqslant 0. \end{cases}$$

  Les 4 premières équations donnent $x=(0,0)$ et $\lambda=(-4,-2)$, ce qui n'est pas compatible avec la positivité requise des multiplicateurs.

4. Il reste donc un seul point stationnaire $x=(0,1)$, de multiplicateur $\lambda=(3,0)$, qui est donc l'unique solution du problème (car toute solution doit vérfier les conditions d'optimalité de KKT).

3. 1. • C'est l'inégalité de dualité faible. Le "max" peut être utilisé à la place du "sup" parce que le supremum est pris sur un nombre fini de quantités ; il est donc atteint.

- Comme exemple avec saut de dualité, il suffit de prendre celui donné à la figure 13.1 du syllabus, qui est du genre $X=\mathbb{R}$, $p=2$, $\varphi_1(x)=(x+1)^2$ et $\varphi_2(x)=(x-1)^2$, auquel cas, le membre de gauche vaut 1 et le membre de droite vaut 0.

2. (a) Soit $\varepsilon>0$. Pour tout $i \in [1:p]$, on peut trouver un $x_i^\varepsilon \in X_i$ tel que

$$\varphi_i(x_i^\varepsilon) \leqslant \inf_{x_i \in X_i} \varphi_i(x_i) + \varepsilon.$$

Alors, (5) s'ensuit.

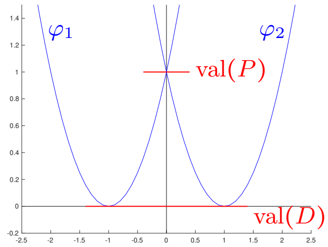

Figure 2: Exemple avec saut de dualité.

(b) Comme le membre de gauche de (5) est la valeur en $x^\varepsilon := (x_1^\varepsilon, \dots, x_p^\varepsilon)$ de la fonction $x = (x_1, \dots, x_p) \in X \mapsto \max_{i\in[1:p]} \varphi_i(x_i)$, l'inégalité suivante a certainement lieu

$$\inf_{x\in X} \max_{i\in[1:p]} \varphi_i(x_i) \leqslant \max_{i\in[1:p]} \varphi_i(x_i^\varepsilon).$$

En combinant avec (5), on trouve

$$\inf_{x\in X} \max_{i\in[1:p]} \varphi_i(x_i) \leqslant \max_{i\in[1:p]} \inf_{x_i\in X_i} \varphi_i(x_i) + \varepsilon.$$

Puisque $\varepsilon > 0$ est arbitraire, l'inégalité "$\leqslant$" a bien lieu dans (4).

Par l'inégalité de dualité faible (3) et le fait que φ_i ne dépend que de la i-ième composante de x, l'inégalité "$\geqslant$" dans (4) se déduit de (3), si bien que l'on a effectivement l'égalité dans (4).

3. On a les équivalences suivantes :

$$\begin{aligned} \bar{x} \in \mathrm{Sol}(P) \quad &\iff \quad \max_{i\in[1:p]} \varphi_i(\bar{x}_i) \leqslant \inf_{x\in X} \max_{i\in[1:p]} \varphi_i(x_i) =: \bar{v}, \\ &\iff \quad \varphi_i(\bar{x}_i) \leqslant \bar{v}, \qquad \forall\, i \in [1:p], \\ &\iff \quad \bar{x} \in \bar{X}. \end{aligned}$$

Donc $\mathrm{Sol}(P) = \bar{X}$.

4. On a les équivalences suivantes :

$$\begin{aligned} \bar{\imath} \in \mathrm{Sol}(D) \quad &\iff \quad \inf_{x_{\bar{\imath}}\in X_{\bar{\imath}}} \varphi_i(x_{\bar{\imath}}) \geqslant \max_{i\in[1:p]} \inf_{x_i\in X_i} \varphi_i(x_i) =: \bar{v}, \\ &\iff \quad \varphi_{\bar{\imath}}(x_{\bar{\imath}}) \geqslant \bar{v}, \qquad \forall\, x_{\bar{\imath}} \in X_{\bar{\imath}}, \\ &\iff \quad \bar{\imath} \in \bar{I}. \end{aligned}$$

Donc $\mathrm{Sol}(D) = \bar{I}$.

5. • Par (7), si $i \notin \bar{I}$, il existe un $\tilde{x}_i \in X_i$ tel que $\varphi_i(\tilde{x}_i) < \bar{v}$. On définit alors $\bar{x} \in X$ par

$$\bar{x}_i = \begin{cases} \hat{x}_i & i \in \bar{I} \\ \tilde{x}_i & \text{sinon.} \end{cases}$$

Ce $\tilde{x}$ vérifie (8b).

• Si $i \in \bar{I}$, qui est nécessairement non vide selon (4), on a

$$\begin{aligned} \varphi_i(\bar{x}_i) &= \varphi_i(\hat{x}_i) \qquad [\text{définition de } \bar{x}] \\ &= \min\{\varphi_i(x_i) : x_i \in X_i\} \qquad [\text{hypothèse faite sur } \hat{x}_i] \\ &= \bar{v} \qquad [i \in \bar{I} \text{ et définition de } \bar{v}]. \end{aligned}$$

Par ailleurs, si $i \notin \bar{I}$, on a

$$\begin{aligned} \varphi_i(\bar{x}_i) &= \varphi_i(\tilde{x}_i) \qquad [\text{définition de } \bar{x}] \\ &< \bar{v} \qquad [i \notin \bar{I} \text{ et définition de } \tilde{x}]. \end{aligned}$$

Dès lors, $\operatorname{Arg\,max}\{\varphi_i(\bar{x}_i) : i \in [1:n]\} = \bar{I}$ et l'on a (8c).

- Enfin, $\max\{\varphi_i(\bar{x}_i) : i \in [1:n]\} = \bar{v}$, ce qui montre que $\bar{x}$ est solution primale ou (8a).

6. (a) $\boxed{\Rightarrow}$

Soient $\bar{\imath} \in \bar{I}$ et $\bar{x} \in \bar{X}$. Comme il n'y a pas de saut de dualité, $(\bar{x}, \bar{\imath})$ est un point-selle de $(x, i) \in X \times [1:p] \mapsto \varphi_i(x)$, si bien que

$$\forall\, x \in X, \quad \forall\, i \in [1:p]: \qquad \varphi_i(\bar{x}_i) \leqslant \varphi_{\bar{\imath}}(\bar{x}_{\bar{\imath}}) \leqslant \varphi_{\bar{\imath}}(x_{\bar{\imath}}). \tag{9}$$

De l'inégalité de gauche on déduit que $\bar{\imath} \in \operatorname{Arg\,max}\{\varphi_i(\bar{x}_i) : i \in [1:p]\}$.

$\boxed{\Leftarrow}$

Attention : si on prend une solution primale quelconque $\bar{x} \in \bar{X}$ et $\bar{\imath} \in \operatorname{Arg\,max}\{\varphi_i(\bar{x}_i) : i \in [1:p]\}$, on a l'inégalité de gauche dans (9), mais pas nécessairement l'inégalité de droite ; donc $(\bar{x}, \bar{\imath})$ n'est pas nécessairement un point-selle et $\bar{\imath}$ n'est pas nécessairement une solution duale.

On peut s'y prendre par contraposition. Supposons que $\bar{\imath} \notin \bar{I}$. Alors, par (7), on peut trouver $\tilde{x}_{\bar{\imath}} \in X_{\bar{\imath}}$ tel que $\varphi_{\bar{\imath}}(\tilde{x}_i) < \bar{v}$. Soit $\hat{x}$ une solution primale (qui est supposée exister). On construit alors le point $\bar{x} \in X$ dont la composante i est donnée par

$$\bar{x}_i = \begin{cases} \hat{x}_i & \text{si } i \neq \bar{\imath} \\ \tilde{x}_{\bar{\imath}} & \text{si } i = \bar{\imath}. \end{cases}$$

Comme $\varphi_i(\bar{x}_i) \leqslant \bar{v}$ pour tout $i \in [1:n]$ (car $\varphi_i(\hat{x}_i) \leqslant \bar{v}$ pour tout $i \in [1:n]$ par (6) et $\varphi_{\bar{\imath}}(\tilde{x}_i) < \bar{v}$), le vecteur $\bar{x}$ est aussi une solution primale (par (6) également). Cependant, $\varphi_{\bar{\imath}}(\bar{x}_{\bar{\imath}}) = \varphi_{\bar{\imath}}(\tilde{x}_{\bar{\imath}}) < \bar{v} = \max_i \varphi_i(\bar{x}_i)$ montre que $\bar{\imath}$ ne maximise pas $i \in [1:p] \mapsto \varphi_i(\bar{x}_i)$.

(b) Soient $\bar{\imath} \in \bar{I}$ et $\bar{x} \in \bar{X}$. Comme il n'y a pas de saut de dualité, $(\bar{x}, \bar{\imath})$ est un point-selle de $(x, i) \in X \times [1:p] \mapsto \varphi_i(x)$, si bien que l'on a (9) et $\varphi_{\bar{\imath}}(\bar{x}_{\bar{\imath}}) = \bar{v}$ (second affirmation). La première affirmation se déduit de l'inégalité de droite dans (9).