Examen du cours Optimisation et Contrôle
Partie I

12 Juin 2014

Résumé

Une série d'exercices indépendants ou (♠) désigne un exercice bonus !

1 Transformée de Fenchel

Soit I un intervalle de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction supposée strictement décroissante et continue. On admet que $J = f(I)$ est un intervalle de $\mathbb{R}$ de même nature que I (par exemple si $I = (a, b]$ alors $J = (c, d])$, et que $f^{-1} : J \to \mathbb{R}$ est aussi strictement décroissante et continue sur J. On supposera de plus que f est convexe sur I. On peut étendre la fonction f en dehors de I en lui donnant la valeur $+\infty$ et calculer alors sa transformée de Fenchel $f^\star$.

Question 1 *Montrer que la transformée de Fenchel de f ne dépends que des valeurs de f sur l'intérieur de I.*

Correction : c'est clair, if faut juste faire attention aux points du bord de l'intervalle, mais par continuité de f les valeurs ne dépendent que des valeurs de f sur l'intérieur de I.

Question 2 *Montrer que f^{-1} est (strictement décroissante et) convexe sur J.*

Correction : On sait que f^{-1} est définie sur $J = f(I)$ et on a admis que f^{-1} est continue et strictement décroissante. Il faut juste montrer la convexité (qui permet en fait aussi de retrouver la continuité sur $\mathrm{int}(J)$.
Pour montrer la convexité, soient $(u, v) \in J^2$, alors il existe $(x, y) \in I^2$ tels que $u = f(x)$ et $v = f(y)$. On écrit l'inégalité de convexité pour (x, y) et on compose avec f^{-1} qui est décroissante.

On note

$$a := \inf I$$

(a est l'extrémité gauche de I ou $-\infty$ si I n'est pas borné à gauche).

Question 3 *Soit $p < 0$, montrer que $(f^{-1})^*(p) = -pf^*(1/p)$.*

Correction : Soit $p \in \mathbb{R}$ on a

$$(f^{-1})^*(p) = \sup_{y \in J} py - f^{-1}(y)$$

où $J = f(I)$

$$\begin{aligned} &= \sup_{x \in I} pf(x) - x \\ &= \sup_{x \in I} -p(-f(x) - x/p) \\ &= (-p) \sup_{x \in I}(-f(x) - x/p) = -pf^*(1/p) \end{aligned}$$

Question 4 *Montrer que* $(f^{-1})^*(0) = -a$.

Correction : Soit $p = 0$. On a $f^{-1})^*(p) = \sup_{x \in I} pf(x) - x$

$$(f^{-1})^*(0) = \sup_{x \in I} -x = -\inf I$$

Question 5 *Soit* $p > 0$*, montrer que* $(f^{-1})^*(p) = f(a^+)p - a$*, si* a *et* $f(a^+)$ *sont finis ; et* $+\infty$ *dans les autres cas (i.e, si* $a = -\infty$ *ou* a *fini et* $f(a^+) = +\infty$*).*

Correction : Soit $p > 0$, la fonction $pf(x) - x$ est décroissante sur I, il faut regarder sa limite en a^+ pour conclure.

2 Conditions d'optimalité

Soit $f : \mathbb{X} \to \mathbb{R}$ une fonction réelle différentiable (au sens de Gateaux) dont la dérivée est continue, C un convexe fermé et $\mathbb{X}$ un espace de Hilbert. On rappelle qu'une condition nécessaire d'optimalité pour que $x^\star$ soit solution du problème $\min_{x \in C} f(x)$ est

$$\langle \nabla f(x^\star), x^\star - x \rangle \leq 0, \quad \forall x \in C\,. \tag{1}$$

On considère la fonction $\theta : \mathbb{X} \to \mathbb{R}$ définie par

$$\theta(x) = \min_{y \in C} g(x, y) \quad \text{avec} \quad g(x, y) := \langle \nabla f(x), y - x \rangle + \frac{1}{2} \|y - x\|^2\,. \tag{2}$$

Question 6 *Montrer que la fonction* θ *est bien définie (problème de minimisation bien posé). Montrer que* θ *est une fonction négative.*

Correction : pour x donné, on minimise une fonction convexe coercive sur un convexe fermé. Le problème de minimisation est bien posé et a une solution unique. Comme $g(x, x) = 0$ la fonction θ est négative

Question 7 *Pour x fixé, écrire les conditions d'optimalité du problème (2) de minimisation de $g(x,y)$ en $y \in C$. Est-ce que les conditions d'optimalité sont nécessaires et suffisantes ?*

Correction : pour x donné, on a un problème convexe, les conditions nécessaire d'optimalité sont aussi suffisantes. Elles s'écrivent

$$\left\langle \nabla f(x) + y - x, y - y' \right\rangle \leq 0 \quad \forall y' \in C \tag{3}$$

Question 8 *Montrer que si $x^\star$ est un optimum local pour f sur C, alors $\theta(x^\star) = 0$.*

Correction : Si $x^\star$ est un optimum local pour f, il vérifie les conditions nécessaires d'optimalité (1). On voit alors que $y = x^\star$ vérifie les conditions d'optimalité (3) pour $x = x^\star$. Pour le problème (2) les conditions d'optimalité sont nécessaires et suffisantes, on obtient donc que le minimum en y de $g(x^\star, y)$ est obtenu pour $y = x^\star$, ce qui donne $\theta(x^\star) = 0$.

3 Fonction support

Soit A un ensemble non vide d'un espace de Hilbert $\mathbb{X}$ et σ_A la fonction support de A. On pose

$$\overline{h} := \operatorname*{argmin}_{h \in \overline{\text{co}}(A)} \|h\|^2$$

Question 9 *Montrer que $\overline{h}$ est bien défini.*

Correction : C'est la projection de 0 sur un convexe fermé non vide $\overline{\text{co}}(A)$. Il existe et il est unique.

Question 10 *Montrer que $\sigma_A(-\overline{h}) = -\left\|\overline{h}\right\|^2$. En déduire que si $0 \notin \overline{\text{co}}(A)$, alors $\sigma_A(-\overline{h}) < 0$.*

Correction : On peut se simplifier la vie en posant $B = \overline{\text{co}}(A)$. On sait alors que $\sigma_A = \sigma_B$. On veut donc montrer que $\sigma_B(-\overline{h}) = -\left\|\overline{h}\right\|^2$, où $\overline{h}$ est la projection de 0 sur le convexe fermé B. La condition d'optimalité pour $\overline{h}$ s'écrit :

$$\left\langle 0 - \overline{h}, x - \overline{h} \right\rangle \leq 0 \quad \forall x \in B \tag{4}$$

Soit encore

$$\left\langle x, -\overline{h} \right\rangle \leq -\left\|\overline{h}\right\|^2 \quad \forall x \in B \tag{5}$$

Par ailleurs on a $\overline{h} \in B$ par définition de la projection. On obtient finalement

$$\sup_{x \in B} \left\langle x, -\overline{h} \right\rangle = -\left\|\overline{h}\right\|^2 . \tag{6}$$

La fin est facile. En effet, si $0 \notin \overline{\text{co}}(A)$ alors $\overline{h} \neq 0$ et donc $\sigma_A(-\overline{h}) < 0$.

Question 11 *Montrer qu'on a l'équivalence*

$$\sigma_A \geq 0 \iff 0 \in \overline{\text{co}}(A)\,. \tag{7}$$

Correction : $(\Leftarrow)$ Si $0 \in \overline{\text{co}}(A)$, alors pour tout $h \in \mathbb{X}$, $\sigma_{\overline{\text{co}}(A)}(h) \geq \langle 0, h\rangle = 0$, et comme $\sigma_A = \sigma_{\overline{\text{co}}(A)}$ on obtient $\sigma_A \geq 0$.

Correction : $(\Rightarrow)$ On montre la réciproque par l'absurde. Si $0 \notin \overline{\text{co}}(A)$, alors on sait par la question 10 que $\sigma_A(-\overline{h}) = -\left\|\overline{h}\right\|^2$ et $\overline{h}$ est non nul. On ne peut donc avoir $\sigma_A \geq 0$

Question 12 *Montrer que le problème de minimisation suivant est bien posé (i.e montrer que le min existe).*

$$\min_{y \in \mathbb{X}} \sup_{x \in \overline{\text{co}}(A)} \langle x, y\rangle + \frac{1}{2}\|y\|^2\,. \tag{8}$$

Quand $\overline{\text{co}}(A)$ est borné, montrer qu'on peut écrire une égalité de point selle et en déduire la valeur du point selle.

Correction : On cherche à minimiser en y la fonction $\sigma_{\overline{\text{co}}(A)}(y) + \frac{1}{2}\|y\|^2$ qui est convexe s.c.i et coercive. Quand $\overline{\text{co}}(A)$ est borné, on vérifie qu'on est bien dans le cadre du théorème du cours pour avoir un point selle. La valeur du point selle s'obtient alors en calculant le problème dual :

$$\max_{x \in \overline{\text{co}}(A)} \inf_{y \in \mathbb{X}} \langle x, y\rangle + \frac{1}{2}\|y\|^2 \tag{9}$$

$$= \max_{x \in \overline{\text{co}}(A)} -\frac{1}{2}\|x\|^2 = -\min_{x \in \overline{\text{co}}(A)} \|x\|^2 = -\left\|\overline{h}\right\|^2. \tag{10}$$

Une application de ce qui précède. Soit J un ensemble fini, on se donne $\psi(x) = \max_{j \in J} f_j(x)$ et on suppose que les fonctions $f_j : \mathbb{X} \to \mathbb{R}$ sont continuement différentiables au sens de Gateaux. On note par ∇f_j leur dérivée au sens de Gateaux.

Question 13 *Montrer que la dérivée directionnelle en x dans la direction h de ψ existe et vaut*

$$d\psi(x;h) = \max_{j \in J_x} \langle \nabla f_j(x), h\rangle \quad \textit{où} \quad J_x = \underset{j \in J}{\text{argmin}}\, f_j(x)\,. \tag{11}$$

Correction : On doit calculer

$$\lim t \downarrow 0 \frac{\psi(x+th) - \psi(x)}{h}\,. \tag{12}$$

On a :

$$\frac{\psi(x+th) - \psi(x)}{h} = \max_{j \in J} \frac{f_j(x+th) - \psi(x)}{h}\,. \tag{13}$$

et

$$\lim_{t \to \infty} \frac{f_j(x+th) - \psi(x)}{h} = \begin{cases} df_j(x;h) \text{ si } j \in J_x \\ -\infty \text{ sinon.} \end{cases} \tag{14}$$

On obtient donc

$$\lim t \downarrow 0 \frac{\psi(x+th)-\psi(x)}{h} = \max_{j\in J_x} df_j(x;h) = \max_{j\in J_x} \langle \nabla f_j(x), h\rangle \ . \tag{15}$$

Question 14 *Montrer que la condition nécessaire d'optimalité pour ψ en $x^\star$ s'écrit aussi*

$$0 \in \overline{\mathrm{co}}_{j\in J_x} \{\nabla f_j(x^\star)\} \ . \tag{16}$$

Correction : Une condition nécessaire d'optimalité est $d\psi(x^\star;h) \geq 0$ pour tout $h \in \mathbb{X}$. Posons $A := \{\nabla f_j(x^\star) | j \in J_{x^\star}\}$, la condition d'optimalité équivaut à $\sigma_A \geq 0$ d'après la question précédente. On a vu que cette condition équivaut à $0 \in \overline{\mathrm{co}}(A)$ qui est précisément la condition à montrer.

On considère maintenant

$$\theta(x) := \min_{h\in\mathbb{X}} \gamma(h,x) \quad \text{avec} \quad \gamma(h,x) := \max_{j\in J} f_j(x) - \psi(x) + \langle \nabla f_j(x), h\rangle + \frac{1}{2}\|h\|^2 \ . \tag{17}$$

Question 15 *Montrer que le problème de minimisation* (17) *en h est bien posé et que $\theta(x) \leq 0$.*

Correction : On minimise en h une fonction convexe sci comme max de fonction convexe sci et la fonction est coercive. Pour $h = 0$ cette fonction vaut $\max_{j\in J} f_j(x) - \psi(x) = 0$, on a donc $\theta(x) \leq 0$

Question 16 *Montrer que*

$$d\psi(x;h(x)) \leq \theta(x) - \frac{1}{2}\|h(x)\|^2$$

où $h(x)$ est la valeur qui réalise le min de $\gamma(h,x)$ pour x fixé.

Correction : On a

$$d\psi(x;h(x)) = \max_{j\in J_x} \langle \nabla f_j(x), h(x)\rangle \tag{18}$$

sachant que $f_j(x) = \psi(x)$ pour $j \in J_x$ on a

$$= \max_{j\in J_x} f_j(x) - \psi(x) + \langle \nabla f_j(x), h(x)\rangle \tag{19}$$

$$\leq \max_{j\in J} f_j(x) - \psi(x) + \langle \nabla f_j(x), h(x)\rangle \tag{20}$$

$$\leq \theta(x) - \frac{1}{2}\|h(x)\|^2 \tag{21}$$

Question 17 *Montrer que si $x^\sharp$ est un minimum local pour ψ, alors $\theta(x^\sharp) = 0$.*

Correction : *D'après la question précédente on a $d\psi(x;h(x)) \leq \theta(x)$. Si $x^\sharp$ est un minimum local pour ψ il vérifie alors les conditions nécessaires d'optimalité $d\psi(x;h) \geq 0$ pour tout h. On a donc $0 \leq d\psi(x^\sharp;h(x^\sharp)) \leq \theta x^\sharp$ et donc $\theta x^\sharp = 0$ puisqu'on sait par ailleurs que $\theta x^\sharp \leq 0$.*

Question 18 *(♠) Montrer que si $d\psi(x^\star;h) > 0$, $\forall h \in \mathbb{X}$, $h \neq 0$ alors $x^\star$ est un minimiseur strict de Ψ. On utilisera un théorème des valeurs intermédiaires (pour $(x,y) \in \mathbb{X}^2$, il existe $s \in [0,1]$ tel que $f_j(y) = f_j(x) + \langle \nabla f_j(x + s(y-x)), y-x \rangle$).*

Correction : *On raisonne par l'absurde. Si $x^\star$ n'est pas un minimum local strict, alors il existe une suite $(x_i)_{i\in\mathbb{N}}$ de points disticts de $x^\star$ et qui converge vers $x^\star$ et telle que $\psi x_j \leq \psi x^\star$. Pour i et j fixé on trouve s_j^i tel que*

$$f_j(x_i) = f_j(x^\star) + \left\langle \nabla f_j(x^\star + s_i^j(x_i - x^\star)), x_i - x^\star \right\rangle \tag{22}$$

On obtient alors

$$\Psi x_i = \max_{j\in J} f_j(x^\star) + \left\langle \nabla f_j(x^\star + s_i^j(x_i - x^\star)), x_i - x^\star \right\rangle \tag{23}$$

et donc

$$\max_{j\in J} f_j(x^\star) + \left\langle \nabla f_j(x^\star + s_i^j(x_i - x^\star)), x_i - x^\star \right\rangle \leq \psi x^\star \tag{24}$$

Comme $J_{x^\star} \subset J$ on obtient

$$\max_{j\in J_{x^\star}} f_j(x^\star) + \left\langle \nabla f_j(x^\star + s_i^j(x_i - x^\star)), x_i - x^\star \right\rangle \leq \psi x^\star\,. \tag{25}$$

Pour $i \in J_{x^\star}$ on a $f_j(x^\star) = \psi x^\star$, l'inégalité se simplifie sous la forme

$$\max_{j\in J_{x^\star}} \left\langle \nabla f_j(x^\star + s_i^j(x_i - x^\star)), x_i - x^\star \right\rangle \leq \psi x^\star\,. \tag{26}$$

Il suffit maintenant de choisir une sous suite en i telle que $(x_{\sigma(i)} - x^\star)/\left\|x_{\sigma(i)} - x^\star\right\|_{i\in\mathbb{N}}$ et $(s^j_{\sigma(i)})_{i\in\mathbb{N}}$ convergent pour tout $j \in J_{x^\star}$ pour trouver une direction h telle que $d\psi(x^\star;h) \leq 0$.