Université Paris Cité
UFR de Mathématiques et Informatique
45, rue des Saints-Pères, 75006, Paris.

M1 MMA 2023-2024
OPTIMISATION SOUS-CONTRAINTES

Feuille de TD n°1

Généralités.

Pour $f_i : \mathbb{R}^n \to \mathbb{R}$ $(1 \leqslant i \leqslant m)$ et $h_j : \mathbb{R}^n \to \mathbb{R}$ $(1 \leqslant j \leqslant p)$, on considère

$$S = \Big\{ x \in \mathbb{R}^n / f_i(x) \leqslant 0,\ i \in \{1, \ldots, m\} \text{ et } h_j(x) = 0,\ j \in \{1, \ldots, p\} \Big\} \tag{1}$$

Exercice 7

Soit f et g deux fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. On suppose que f est différentiable sur $\mathbb{R}^n$ et que g est convexe sur $\mathbb{R}^n$. Soit C un sous-ensemble convexe non vide de $\mathbb{R}^n$, et $x \in C$. On considère les deux propriétés suivantes

(i) $f + g$ admet un minimum en x sur C.

(ii) Pour tout $y \in C$,

$$\langle \nabla f(x),\, y - x \rangle + g(y) - g(x) \geqslant 0.$$

1. Montrer que (i) implique (ii).
2. On suppose de plus que f est convexe. Montrer que (i) et (ii) sont équivalents.

Correction.

1. Supposons (i) vrai. Soit $y \in C$. On a pour tout $t \in [0, 1]$, $x + t(y - x) = (1 - t)x + ty \in C$ puisque C est convexe.

Comme $f + g$ admet un minimum en x sur C, on a donc

$$\forall t \in [0, 1], \quad f(x + t(y - x)) + g(x + t(y - x)) \geqslant f(x) + g(x),$$

et donc par convexité de g

$$\forall t \in [0, 1], \quad f(x + t(y - x)) + g(x) + t(g(y) - g(x)) \geqslant f(x + t(y - x)) + g(x + t(y - x)),$$

ainsi finalement

$$\forall t \in [0, 1], \quad f(x + t(y - x)) + t(g(y) - g(x)) \geqslant f(x). \tag{2}$$

Or f étant différentiable sur $\mathbb{R}^n$, il existe $\varepsilon : V \to \mathbb{R}$ une fonction définie sur V voisinage de $0_{\mathbb{R}^n}$ telle que ε admet 0 comme limite en $0_{\mathbb{R}^n}$ et pour tout $h \in V$, $f(x + h) = f(x) + \langle \nabla f(x),\, h \rangle + \|h\|\, \varepsilon(h)$. Il existe $t_0 \in]0, 1]$ tel que pour tout $t \in]0, t_0]$, on a $t(y - x) \in V$ et donc $f(x + t(y - x)) = f(x) + t \langle \nabla f(x),\, y - x \rangle + t\, \|y - x\|\, \varepsilon(t(y - x))$. Par conséquent en utilisant (2) et en simplifiant par $t \neq 0$, on obtient

$$\forall t \in]0, t_0], \quad \langle \nabla f(x),\, y - x \rangle + \|y - x\|\, \varepsilon(t(y - x)) + g(y) - g(x) \geqslant 0.$$

En passant à la limite lorsque $t \to 0$, on obtient l'inégalité souhaitée.

> La méthode utilisée pour obtenir la réponse est classique : x est un minimum de $f + g$ donc pour obtenir une condition d'optimalité, on perturbe $f + g$ en x dans des directions appropriés. Ici on peut effectuer la perturbation dans la direction $y - x$ pour tout $y \in C$ car C est convexe (c'est une direction admissible à C en x et donc également une direction tangente à C en x, cf chapitre 2).
> Notez que dans l'énoncé x est supposé être un minimum global pour $f + g$ sur C, mais la conclusion restait vraie si x était seulement un minimum local (il suffit de se restreindre au début du raisonnement à un voisinage relatif de x dans C).

2. Montrons que (ii) implique (i). Soit $y \in C$, comme f est convexe et différentiable en x, on a

$$f(y) \geqslant f(x) + \langle \nabla f(x),\, y - x \rangle\, .$$

Ainsi comme $\langle \nabla f(x),\, y - x \rangle + g(y) \geqslant g(x)$, on obtient que $f(y) + g(y) \geqslant f(x) + \langle \nabla f(x),\, y - x \rangle + g(y) \geqslant f(x) + g(x)$. Par conséquent $x \in C$ est un minimum global de f sur C.

Exercice 8 (Méthode de pénalisation extérieure)
Soit $f:\mathbb{R}^n\to\mathbb{R}$ une fonction continue, coercive et strictement convexe. On considère $C=\{x\in\mathbb{R}^n \,/\, g(x)\leqslant 0\}$, où $g:\mathbb{R}^n\to\mathbb{R}$ est une fonction convexe continue, et on suppose que C est non vide. On note $x^\star$ le minimum de f sur C. Pour tout $k\in\mathbb{N}$, on pose

$$\forall x\in\mathbb{R}^n,\quad f_k(x)=f(x)+k\alpha(x),\qquad \text{avec } \alpha(x)=(\max(g(x),0))^2.$$

1. Montrer que pour tout $k\in\mathbb{N}$, f_k admet un unique minimum sur $\mathbb{R}^n$, que l'on note x_k.
2. Montrer que pour tout $k\in\mathbb{N}$, $f(x_k)\leqslant f(x^\star)$, puis que la suite $(x_k)_{k\in\mathbb{N}}$ est bornée. On note $\tilde{x}$ une valeur d'adhérence de cette suite.
3. Montrer que que pour tout $k\in\mathbb{N}^*$, $\alpha(x_k)\leqslant \frac{1}{k}(f(x^\star)-f(x_k))$ et en déduire que $\tilde{x}\in C$.
4. Montrer que $\tilde{x}=x^\star$, puis que $(x_k)_{k\in\mathbb{N}}$ converge vers $x^\star$.

Correction.

> Notez que f admet bien un unique minimum global sur C puisque C est fermé (car g continue) et non vide, et f continue coercive et strictement convexe sur le convexe C.

1. Soit $k\in\mathbb{N}$. Comme g est continue sur $\mathbb{R}^n$, c'est aussi le cas de $\alpha=\max(0,g)=\frac{g+|g|}{2}$ et donc f_k est continue sur $\mathbb{R}^n$. Comme $f_k\geqslant f$ et f coercive sur $\mathbb{R}^n$, on déduit donc que f_k est coercive sur $\mathbb{R}^n$. L'ensemble C étant fermé et non vide, on obtient que f_k admet un minimum global sur C.

   Montrons que la fonction α est convexe sur $\mathbb{R}^n$. Soit $x,x'\in\mathbb{R}^n$ et $t\in[0,1]$. Si $g((1-t)x+tx')\leqslant 0$, alors $\alpha((1-t)x+tx')=0$ et puisque $\alpha\geqslant 0$, on a donc $0=\alpha((1-t)x+tx')\leqslant(1-t)\alpha(x)+t\alpha(x')$. Sinon $g((1-t)x+tx')>0$ et comme par convexité de g, on a $g((1-t)x+tx')\leqslant(1-t)g(x)+tg(x')$, on a donc $\alpha((1-t)x+tx')\leqslant(1-t)g(x)+tg(x')\leqslant(1-t)\alpha(x)+t\alpha(x')$ (car $\alpha\geqslant g$).

   Ainsi f étant strictement convexe sur $\mathbb{R}^n$, on déduit que $f_k=f+k\alpha$ est également strictement convexe sur $\mathbb{R}^n$. L'ensemble C étant convexe, f est donc strictement convexe sur C et par conséquent f_k admet au plus un minimum global sur C.

   Finalement, f_k admet bien un unique minimum global sur C.
2. Soit $k\in\mathbb{N}$, comme x_k est le minimum global de f_k sur $\mathbb{R}^n$, on a $f_k(x_k)\leqslant f_k(x^\star)$. Or $f_k(x^\star)=f(x^\star)$ car $x^\star\in C$ d'où $g(x^\star)\leqslant 0$ et donc $\alpha(x^\star)=0$. De plus $f_k(x_k)\geqslant f(x_k)$ puisque $\alpha\geqslant 0$, d'où finalement $f(x_k)\leqslant f(x^\star)$.

   Comme la fonction f est coercive, il existe $r>0$ tel que pour tout $\|x\|>r$, on a $f(x)>f(x^\star)$. Par conséquent $\|x_k\|\leqslant r$ et donc $(x_k)_{k\in\mathbb{N}}$ est bornée.
3. Soit $k\in\mathbb{N}^*$, on a vu que $f_k(x_k)\leqslant f_k(x_k)=f(x^\star)$ d'où $\alpha(x_k)\leqslant\frac{1}{k}(f(x^\star)-f(x_k))$.

   Soit φ une extractrice telle que x_φ converge vers $\tilde{x}\in\mathbb{R}^n$. La fonction f étant continue, on déduit par le critère séquentielle de la continuité que $f(x_{\varphi(k)})\underset{k\to+\infty}{\to}f(\tilde{x})$. En particulier $(f(x_{\varphi(k)}))_{k\in\mathbb{N}}$ est bornée, donc il existe $M>0$ tel que pour tout $k\in\mathbb{N}^*$

$$0\leqslant\alpha(x_{\varphi(k)})\leqslant\frac{M}{k}. \tag{3}$$

   Par continuité de α et le critère séquentielle de la continuité, on a $\alpha(x_{\varphi(k)})\underset{k\to+\infty}{\to}\alpha(\tilde{x})$ et par passage à la limite dans (3) on obtient $\alpha(\tilde{x})=0$, d'où $g(\tilde{x})\leqslant 0$ i.e. $\tilde{x}\in C$.
4. On sait d'après 2. que pour tout $k\in\mathbb{N}$, $f(x_{\varphi(k)})\leqslant f(x^\star)$, puis d'après 3. que $(f(x_{\varphi(k)}))_{k\in\mathbb{N}}$ tend vers $f(\tilde{x})$, ainsi par passage à la limite dans l'inégalité on obtient $f(\tilde{x})\leqslant f(x^\star)$. Or $\tilde{x}\in C$ et $x^\star$ est l'unique minimum global de f sur C, d'où $\tilde{x}=x^\star$.

   Comme $\tilde{x}$ était *une* valeur d'adhérence quelconque de $(x_k)_{k\in\mathbb{N}}$, on déduit donc que $(x_k)_{k\in\mathbb{N}}$ admet $x^\star$ comme unique valeur d'adhérence. Or $(x_k)_{k\in\mathbb{N}}$ est bornée, donc $(x_k)_{k\in\mathbb{N}}$ converge vers $x^\star$.