INSA de Toulouse - 4GMM **Mardi 05/01/2016**

ANALYSE CONVEXE & OPTIMISATION
Durée approximative : 2h.

Seuls le polycopié et les notes de cours sont autorisés.

Il est vivement recommandé de lire tout l'énoncé au départ, car la majorité des questions sont dépendantes.

Les questions précédées d'une (*) sont plus difficiles (mais abordables). La question (***) est bien plus ouverte et complexe.

Régularisation de Moreau-Yosida

La régularisation de Moreau-Yosida est une technique générique pour obtenir des approximations régulières de fonctions non lisses. Le but de ce problème est d'établir quelques propriétés de cette régularisation dans le cas convexe.

Soit
$$\Gamma_0(\mathbb{R}^n) = \{f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\},\ f \text{ est convexe fermée}\}.$$
Soit $f \in \Gamma_0(\mathbb{R}^n)$. On note pour $\epsilon \geq 0$:
$$f_\epsilon(x) = \min_{u\in\mathbb{R}^n} f(u) + \frac{1}{2\epsilon}\|x-u\|_2^2, \tag{1}$$
où $\|\cdot\|_2^2$ désigne la norme ℓ^2 usuelle sur $\mathbb{R}^n$. L'application $f \mapsto f_\epsilon$ est appelée régularisation de Moreau-Yosida.

Partie I (Questions de cours $\sim$ 11pts)

1. Soit $g_\epsilon(u) = f(u) + \frac{\epsilon}{2}\|x-u\|_2^2$. La fonction g_ϵ est-elle différentiable pour tout $f \in \Gamma_0(\mathbb{R}^n)$? Est-elle fortement convexe pour tout $f \in \Gamma_0(\mathbb{R}^n)$?
2. Le problème (1) est-il convexe ? Admet-il une solution ? Si oui, est-elle unique ?
3. Calculer la régularisée de Moreau-Yosida de la fonction $h(u) = |u|$, où $u \in \mathbb{R}$.
4. Montrer que pour $h(u) = |u|$, la fonction h_ϵ est C^1.
5. (*) Calculer la constante de Lipschitz de la dérivée de h_ϵ.
6. Calculer la régularisée de Moreau-Yosida de la fonction $f(u) = \|u\|_1$, où $u \in \mathbb{R}^n$ (vous pourrez utiliser les résultats sur h_ϵ directement).
7. Dessiner les lignes de niveau de la fonction $f(u) = \|u\|_1$ dans $\mathbb{R}^2$. Sur le même graphe, dessiner les lignes de niveau de sa régularisée f_ϵ pour $\epsilon = 1$.
8. Déterminer le minimiseur (l'argmin) du problème (1) pour la fonction $f(u) = \frac{1}{2}\|Au-b\|_2^2$, où $A \in \mathbb{R}^{m\times n}$ et $b \in \mathbb{R}^m$.

Partie II (Problème $\sim$ 8pts + X)

Pour analyser la fonction f_ϵ, on va se servir de relations de dualité. On rappelle que la transformée de Fenchel de f est notée f^* et que pour $f \in \Gamma_0(\mathbb{R}^n)$, elle satisfait $f^{**} = f$, soit encore, pour tout $u \in \mathbb{R}^n$:

$$f(u) = \sup_{y \in \mathbb{R}^n} \langle u, y \rangle - f^*(y).$$

On peut donc écrire que :

$$\begin{aligned} f_\epsilon(x) &= \min_{u \in \mathbb{R}^n} \sup_{y \in \mathbb{R}^n} \langle u, y \rangle - f^*(y) + \frac{1}{2\epsilon}\|x-u\|_2^2 \\ &= \sup_{y \in \mathbb{R}^n} \min_{u \in \mathbb{R}^n} \langle u, y \rangle - f^*(y) + \frac{1}{2\epsilon}\|x-u\|_2^2. \end{aligned}$$

La possibilité d'intervertir le min et le sup est un résultat de dualité de Fenchel-Rockefeller (non vue en cours). Ce résultat fondamental est équivalent à la dualité Lagrangienne vue dans la première partie du cours. Dans tout le problème, on suppose que $\mathrm{dom}(f^*)$ est borné, c'est-à-dire qu'il existe $R \geq 0$ tel que $\forall x \in \mathbb{R}^n$, $\|x\|_2 > R$, $f^*(x) = +\infty$.

1. Montrer que
$$f_\epsilon(x) = \sup_{y \in \mathbb{R}^n} \langle x, y \rangle - f^*(y) - \frac{\epsilon}{2}\|y\|_2^2. \tag{2}$$
2. Montrer que f_ϵ est une fonction convexe.
3. Montrer l'encadrement suivant :
$$f(x) - \frac{\epsilon}{2}R^2 \leq f_\epsilon(x) \leq f(x).$$
Ce résultat établit que la suite $(f_\epsilon)_{\epsilon \in \mathbb{R}_+}$ converge uniformément vers f lorsque $\epsilon \to 0$.
4. Soit $y(x) \in \mathbb{R}^n$ le point qui réalise le suprémum dans (2). Justifier que ce point est unique.
5. (*) Soit $x_0 \in \mathbb{R}^n$. On rappelle qu'un élément $\eta \in \partial f_\epsilon(x_0)$ satisfait pour tout $x \in \mathbb{R}^n$:
$$f_\epsilon(x) \geq f_\epsilon(x_0) + \langle \eta, x - x_0 \rangle.$$
Montrer que $y(x_0) \in \partial f_\epsilon(x_0)$.
6. Le but de cette question est de montrer que f_ϵ est différentiable et de déterminer son gradient. On donne pour ce faire le théorème de Danskin (1967).

   Théorème 1. *Soit Y un sous-ensemble compact de $\mathbb{R}^m$. Soit $\phi : \mathbb{R}^n \times Y \to \mathbb{R}$ une fonction convexe en la variable x. Soit $g(x) = \sup_{y \in Y} \phi(x, y)$. On suppose que le supremum est atteint en un seul point $y(x)$.*
   Alors g est convexe et différentiable. De plus $\nabla g(x) = \frac{\partial \phi(x, y(x))}{\partial x}$.

   En utilisant ce résultat, montrer que f_ϵ est différentiable et déterminer son gradient.
7. (*** Hors-barême) Montrer que ∇f_ϵ satisfait :
$$\|\nabla f_\epsilon(x_1) - \nabla f_\epsilon(x_2)\|_2 \leq \frac{1}{\epsilon}\|x_1 - x_2\|_2.$$
Indication : on peut utiliser les conditions d'optimalité $x_1 - \partial f(y(x_1)) - \epsilon y(x_1) \ni 0$.

Partie III (Application algorithmique $\sim$ 7pts)

Dans cette partie, on s'intéresse au problème :

$$\min_{x \in \mathbb{R}^n} f(x)$$

où f est une fonction convexe fermée non-différentiable qui admet au moins un minimum x^*. On suppose que $\text{dom}(f^*)$ est borné de rayon $R > 0$. Pour résoudre ce problème, on se propose de le régulariser en résolvant à place :

$$\min_{x \in \mathbb{R}^n} f_\epsilon(x), \tag{3}$$

et d'utiliser une descente de gradient accélérée sur (3). On note $(x_k)_{k \in \mathbb{N}}$ la suite générée par cette descente.

1. Rappeler le taux de convergence de la suite $(x_k)_{k \in \mathbb{N}}$.
2. (*) En utilisant le résultat de la question II.3, montrer que :

$$f(x_k) - f(x^*) \leq \frac{\|x_0 - x_\epsilon^*\|_2^2}{\epsilon k^2} + \frac{\epsilon R^2}{2}.$$

3. Soit $c > 0$. On souhaite obtenir un point x_k satisfaisant :

$$f(x_k) - f(x^*) \leq c,$$

avec l'algorithme décrit ci-dessus. Comment choisir ϵ et le nombre d'itérations de façon optimale (i.e. quel est le nombre d'itération minimum nécessaire pour obtenir cette précision) ?