Optimisation avancée
2018-2019. Examen d'entraînement

Exercice 1 Quizz

1. Répondre par *Vrai* ou *Faux* aux assertions suivantes. Donnez une preuve pour chaque *Vrai* et un contre-exemple pour chaque *Faux*.
   a) Soit $\beta > 0$. Si f est convexe et β-régulière et g est convexe, alors $f + g$ est β-régulière.
   b) Si f est une fonction convexe définie sur un ensemble convexe $E \subseteq \mathbb{R}^d$, alors f est différentiable sur l'intérieur de E.
   c) Soit $f : \mathbb{R}^d \to \mathbb{R}$ et $\alpha > 0$. Alors, f est α-fortement convexe si et seulement si pour tout $u \in \mathbb{R}^d$ et pour tout $x \in \mathbb{R}^d$, $\phi_{u,x}$ est $\alpha\|u\|^2$-fortement convexe, où $\phi_{u,x}(t) = f(x + tu), \forall t \in \mathbb{R}$.
2. Rappeler l'algorithme de la descente de gradient pour les fonctions convexes β-régulières sur $\mathbb{R}^d$, en rappelant l'interprétation géométrique du choix du pas.

Exercice 2 Convexité

1. Soit E un sous-ensemble convexe et fermé de $\mathbb{R}^d$ et soit π_E la projection sur E. Montrer que π_E est 1-Lipschitzienne.
2. Soit f une fonction convexe sur un ensemble convexe ouvert E et $L > 0$. Montrer que f est L-Lipschitzienne sur E si et seulement si pour tout $x \in E$, l'ensemble des sous-gradients de f en x est inclus dans la boule Euclidienne centrée en l'origine et de rayon L.

Exercice 3 Descente de gradient conditionnel: l'algorithme de Frank-Wolfe

Soit $E \subseteq \mathbb{R}^d$ un ensemble convexe et compact et $f : E \to \mathbb{R}$ une fonction convexe et β-régulière, où $\beta > 0$ est connu. On considère l'algorithme ci-dessous, où $\gamma_s = \frac{2}{s}$, pour tout $s = 1, \ldots, T$. Soit D le diamètre de E.

1. Montrer que pour tout $t = 1, \ldots, T$, $f(x_t) - f(x_{t-1}) \leq \gamma_t \nabla f(x_{t-1})^\top (y_{t-1} - x_{t-1}) + \dfrac{\beta}{2}\gamma_t^2 D^2$.
2. Pour $t = 1, \ldots, T$, on pose $\delta_t = f(x_t) - f(x^*)$. En utilisant la définition de y_{t-1}, en déduire que $\delta_t \leq (1 - \gamma_t)\delta_{t-1} + \frac{\beta}{2}\gamma_t^2 D^2$.
3. Conclure que $f(\hat{x}) - f(x^*) \leq \frac{\beta D^2}{T}$

Algorithm 1 Algorithme de Frank-Wolfe

$x_0 \in \mathbb{R}^d$.

for $t = 1$ TO T **do**

$\quad y_{t-1} \in \underset{y \in E}{\arg\min} \nabla f(x_{t-1})^\top y$

$\quad x_t = (1 - \gamma_t) x_{t-1} + \gamma_t y_{t-1}$

end for

$\hat{x} = x_T$