Module optimisation non lisse pour le machine learning
Pas de document autorisé

Sujet d'examen méthode sous-gradient et prox : 21 février 2020

Nom :

Prénom :

Exercice : convergence de la méthode de sous-gradient

On suppose que la fonction $f : \mathbb{R}^n \to \mathbb{R},\ x \mapsto f(x)$ est convexe. On suppose aussi que $\exists G > 0, \forall x \in \mathbb{R}^n,\ \forall g \in \partial f(x), \|g\| \ \leq\ G$ On utilise la méthode du sous-gradient pour résoudre le problème $\min_{x \in \mathbb{R}^n} f(x)$. On suppose que ce problème admet une solution notée x_*, on note $R = \|x_0 - x_*\|$, et on pose $f_{\text{best}}^K = \min_{0 \leq k \leq K} f(x_k)$.

1. Montrer (3 lignes) que pour tout k, $\|x_{k+1} - x_*\|^2 \ \leq\ \|x_k - x_*\|^2 - 2\alpha_k g_k^\top (x_k - x_*) + \alpha_k^2 \|g_k\|^2$ puis que (2 lignes) $\|x_{k+1} - x_*\|^2 \leq\ \|x_k - x_*\|^2 - 2\alpha_k (f(x_k) - f_*) + \alpha_k^2 \|g_k\|^2$.
   Réponse :

2. En déduire que $f_{\text{best}}^K - f(x^\star) \ \leq\ \frac{R^2 + G^2 \sum_{k=0}^K \alpha_k^2}{2 \sum_{k=0}^K \alpha_k}$ et que la méthode converge pour $\alpha_k = 1/(k+1)$. Réponse :

3. Utiliser la question 2. pour le cas où $\alpha_k = 1/\sqrt{k+1}$. Réponse :

On rappelle que $\operatorname{prox}_h(x) = \operatorname{argmin}_u h(u) + \frac{1}{2}\|u - x\|^2$. On suppose les fonctions citées ci-dessous convexes. Montrer que

4. si $f(x) = ag(x) + b$ avec $a > 0$, alors $\operatorname{prox}_f(x) = \operatorname{prox}_{ag}(x)$. Réponse :

5. si $f(x) = g(x) + a^T x - b$, alors $\operatorname{prox}_f(x) = \operatorname{prox}_g(x - a)$. Réponse :

6. si $f(x) = g(ax + b)$, avec $a \neq 0$, alors $\operatorname{prox}_f(x) = \frac{1}{a}\left(\operatorname{prox}_{a^2 g}(ax + b) - b\right)$. Réponse :