Université Claude Bernard - Lyon 1
M2 Statistique, Modélisation et Science des données
11 décembre 2023

Optimisation Convexe : Algorithmes et Applications en Apprentissage

Contrôle terminal. Durée : 2h. Tous les documents sont autorisés.
Dans tout le sujet ci-dessous, $||x||_p$ désigne la norme ℓ_p : $||x||_p := (\sum_i |x_i|^p)^{1/p}$ si $p < \infty$, $||x||_\infty := \max_i |x_i|$.

Exercice 1 (6 points)**.** Soit $f : \mathbb{R}^n \to \mathbb{R}$ la fonction définie par $f(x) := \sqrt{1 + ||x||_2^2}$. Calculer f^*. En quels points le sous-différentiel de f^* est-il non-vide ?

Exercice 2 (6 points)**.** Étant donné une matrice $A \in \mathbb{R}^{m\times n}$ on considère les deux problèmes suivants :

$$\min \left\{ \sum_{i=1}^n x_i \log(x_i) \ : \ x \in \mathbb{R}^n,\, x_i \geq 0 \text{ for all } i, ||Ax||_\infty \leq 1 \right\} \quad \text{et} \quad \max \left\{ - \sum_{i=1}^n e^{-(A^t\xi)_i - 1} - ||\xi||_1 \ : \xi \in \mathbb{R}^m \right\}.$$

1. Expliquer pourquoi les deux problèmes sont en dualité.
2. Suggérer quels algorithmes on pourrait utiliser pour résoudre l'un et l'autre et discuter leur faisabilité.

Exercice 3 (8 points)**.** Étant donnés une matrice $A \in \mathbb{R}^{m\times n}$ et deux vecteurs $y \in \mathbb{R}^n$ et $c \in \mathbb{R}^m$ on veut projeter y sur l'ensemble $C = \{x \in \mathbb{R}^n \ : \ Ax \leq c\}$, où l'inégalité entre vecteurs est à prendre composante par composante. On regarde donc

$$\min \left\{ \frac{||x-y||_2^2}{2} \ : \ Ax \leq c \right\}.$$

1. Un minimiseur x existe-t-il ? est-il unique ?
2. Suggérer un algorithme pour résoudre ce problème, en décrivant explicitement ses différentes étapes.
3. Peut-on utiliser l'algorithme de gradient projeté pour résoudre ce problème ?

Exercice 4 (10 points)**.** On s'intéresse à l'approximation d'un vecteur $y \in \mathbb{R}^n$ donné par un vecteur à composantes positives.

1. Trouver les cordonnées du point $\bar{x}(y)$ solution de

$$\min \left\{ \frac{||x-y||_2^2}{2} \ : \ x \geq 0 \right\}.$$

2. Si on accepte que certaines composantes soient négatives, on peut se donner $\lambda > 0$ et résoudre

$$\min \left\{ \frac{||x-y||_2^2}{2} + \lambda \sum_{i=1}^n [x_i]_- \right\},$$

   où $[s]_- := \max\{-s, 0\}$ indique la partie négative du nombre $s \in \mathbb{R}$. Prouver que ce problème admet une unique solution et donner la formule pour la trouver, en termes de y et λ. On appelle cette solution $\bar{x}(\lambda, y)$.
3. Prouver que l'on a $\lim_{\lambda\to+\infty} \bar{x}(\lambda, y) = \bar{x}(y)$.
4. Étant donnés une matrice $A \in \mathbb{R}^{m\times n}$ on veut maintenant approcher un vecteur $y \in \mathbb{R}^m$ par un vecteur de la forme Ax où x est un vecteur à composantes positives, ou dont peu des composantes sont négatives. On considère alors

$$\min \left\{ \frac{||Ax-y||_2^2}{2} + \lambda \sum_{i=1}^n [x_i]_- \right\}.$$

   Suggérer un algorithme pour résoudre ce problème, en décrivant explicitement ses différentes étapes.
5. Pour la convergence de l'algorithme choisi il est sans doute nécessaire de supposer que le problème d'optimisation ci-dessus admet une solution. Prouver qu'il en admet une lorsque A est une matrice dont tous les coefficients sont strictement positifs.