UFR
DE MATHÉMATIQUES
ET INFORMATIQUE

UNIVERSITÉ
PARIS
DESCARTES

Optimisation (MML1E31) (M1 MM, 2017-2018)

Examen partiel du lundi 20 novembre 2017

Nombre de pages du sujet : 3. Durée 2h.

Les documents suivants sont autorisés :
— Polycopiés distribués en cours et notes de cours manuscrites correspondantes,
— Sujets de TP imprimés et notes manuscrites correspondantes.
La consultation de tout autre document (livres, etc.) et l'utilisation de tout appareil électronique sont interdites.

Exercice 1. *(sur environ 6 points)*

Soient m et n deux entiers non nuls. Soient $f : \mathbb{R}^m \to \mathbb{R}$ une fonction, $A \in M_{m,n}(\mathbb{R})$ une matrice rectangulaire et $b \in \mathbb{R}^m$ un vecteur colonne. On pose $g : \mathbb{R}^n \to \mathbb{R}$ la fonction définie par

$$\forall x \in \mathbb{R}^n, \quad g(x) = f(Ax+b).$$

1. Rappeler la définition d'une fonction strictement convexe sur $\mathbb{R}^n$.
2. On suppose que $\mathrm{Ker}(A) \neq \{0\}$. Montrer que g n'est pas strictement convexe sur $\mathbb{R}^n$.
3. Montrer que si f est strictement convexe sur $\mathbb{R}^m$ et $\mathrm{Ker}(A) = \{0\}$ alors la fonction g est également strictement convexe sur $\mathbb{R}^n$.
4. On considère la fonction $h : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$h(x) = e^{2x_1 - x_2 - 1} + e^{-x_1+2} + e^{x_2}.$$

   Démontrer que h est strictement convexe sur $\mathbb{R}^2$ à l'aide de la question précédente.
5. Donner un script Octave qui :
   — Définit une fonction `h = @(x)` qui évalue $h(x)$ en s'assurant que cette fonction puisse être évaluée sur des listes de p points $(x^{(1)}, x^{(2)}, \dots, x^{(p)})$ données sous la forme d'une matrice de taille $n \times p$.
   — Affiche le graphe de la fonction h sur le domaine $[0,2] \times [0,2]$.

Exercice 2. *(sur environ 17 points)*

Soit $n \geqslant 2$ un entier. On note $\|x\| = \sqrt{\langle x, x\rangle}$ la norme euclidienne usuelle de $\mathbb{R}^n$,

$$\|x\|_\infty = \max_{i \in \{1,\dots,n\}} |x_i| \quad \text{et} \quad \|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

Pour $x \in \mathbb{R}^n$ on note $\mathrm{sign}(x)$ le vecteur de $\{-1,1\}^n$ tel que pour tout $i \in \{1,\dots,n\}$

$$(\mathrm{sign}(x))_i = \begin{cases} 1 & \text{si } x_i \geqslant 0, \\ -1 & \text{si } x_i < 0. \end{cases}$$

Dans la suite de l'exercice, $f : \mathbb{R}^n \to \mathbb{R}$ désigne une fonction deux fois différentiable sur $\mathbb{R}^n$, fortement convexe et telle que

$$\forall x, h \in \mathbb{R}^n, \quad m\|h\|^2 \leqslant \langle \nabla^2 f(x)h, h\rangle \leqslant M\|h\|^2$$

avec $0 < m \leqslant M$. On note $x^\star$ le point où f atteint son minimum global et $p^\star = f(x^\star)$ la valeur minimale de f.

Le but de l'exercice est d'étudier la convergence d'un algorithme de descente qui prend pour direction de descente à l'étape k le vecteur

$$d^{(k)} = -\|\nabla f(x^{(k)})\|_1 \operatorname{sign}(\nabla f(x^{(k)})).$$

Plus précisément, on considère l'Algorithme 1 ci-dessous.

Algorithme 1 : Algorithme de descente selon la direction de plus grande pente pour $\|\cdot\|_1$

Données : Un point initial $x^{(0)} \in \mathbb{R}^n$, un seuil de tolérance $\varepsilon > 0$
Résultat : Un point $x \in \mathbb{R}^n$ proche de $x^\star$
Initialiser x :
  $x \leftarrow x^{(0)}$;
  $k \leftarrow 0$;
tant que $\|\nabla f(x)\| > \varepsilon$ **faire**

1. Calculer la direction de descente $d^{(k)} = -\|\nabla f(x^{(k)})\|_1 \operatorname{sign}(\nabla f(x^{(k)}))$.
2. Déterminer un pas de descente $t^{(k)} > 0$ par la méthode exacte pour la direction de descente $d^{(k)} = -\|\nabla f(x^{(k)})\|_1 \operatorname{sign}(\nabla f(x^{(k)}))$
   (ou par la méthode de rebroussement de paramètres α et β).
3. Mettre à jour x :
   $x \leftarrow x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)}$;
   $k \leftarrow k + 1$;

fin

Partie I : Implémentation de l'Algorithme 1 avec la méthode de rebroussement *(sur environ 4 points)*

On suppose que les fonctions `f = @(x)` et `gradf = @(x)` qui évaluent respectivement $f(x)$ et $\nabla f(x)$ sont définies dans Octave. On rappelle qu'en Octave si `u` est un vecteur $u = (u_1, \ldots, u_n)$, alors `abs(u)` renvoie le vecteur $(|u_1|, \ldots, |u_n|)$.

1. Ecrire une fonction Octave `sign` qui prend en entrée un vecteur colonne u et renvoie le vecteur $\operatorname{sign}(u) \in \{-1, 1\}^n$.
2. Ecrire une fonction Octave

```
function x = descente_norme_un(x0, eps, alph, bet, f, gradf)
```

   qui prend en entrées un point initial $x^{(0)} \in \mathbb{R}^n$, un seuil de tolérance $\varepsilon > 0$, et deux paramètres α et β, et les fonctions `f` et `gradf` et renvoie le point x calculé par l'Algorithme 1 utilisant la méthode de rebroussement pour le calcul du pas de descente (avec les paramètres $\alpha \in [0, \frac{1}{2}[$ et $\beta \in]0, 1[$).

Partie II : Quelques propriétés générales *(sur environ 6 points)*

3. Montrer que pour tout $x \in \mathbb{R}^n$, $\|x\|_1 \geqslant \|x\|$.
4. Soit $x \in \mathbb{R}^n$. Montrer que
$$\langle x, -\operatorname{sign}(x)\rangle = \min_{u \in \mathbb{R}^n,\ \|u\|_\infty \leqslant 1} \langle x, u\rangle.$$
5. Montrer que pour tout $y \in \mathbb{R}^n$,
$$f(y) \geqslant f(x) + \langle \nabla f(x), y - x\rangle + \frac{m}{2}\|y - x\|^2.$$
6. Montrer que $g : y \mapsto f(x) + \langle \nabla f(x), y - x\rangle + \frac{m}{2}\|y - x\|^2$ est une fonctionnelle quadratique pour laquelle on précisera la matrice A, le vecteur b et la constante c.
7. Déterminer le minimum de g et en déduire que
$$\|\nabla f(x)\|^2 \geqslant 2m(f(x) - p^\star).$$

Partie III : Etude de la convergence de l'Algorithme 1 avec la méthode exacte *(sur environ 7 points)*

Le but de la fin de l'exercice est de démontrer la convergence de l'Algorithme 1 **avec la méthode exacte** pour le calcul du pas de descente. On suppose qu'à l'étape k de l'Algorithme 1 le point $x^{(k)}$ est différent de $x^\star$ et on pose $d^{(k)} = -\|\nabla f(x^{(k)})\|_1 \operatorname{sign}(\nabla f(x^{(k)}))$.

8. Montrer que $\langle \nabla f(x^{(k)}), d^{(k)}\rangle = -\|\nabla f(x^{(k)})\|_1^2$ et justifier que $d^{(k)}$ est bien une direction de descente pour le point x.
9. Montrer que pour tout $t > 0$,
$$f(x^{(k+1)}) \leqslant f(x^{(k)} + td^{(k)}) \leqslant f(x^{(k)}) - t\|\nabla f(x^{(k)})\|_1^2 + t^2\frac{M}{2}n\|\nabla f(x^{(k)})\|_1^2.$$
10. En déduire que
$$f(x^{(k+1)}) \leqslant f(x^{(k)}) - \frac{1}{2nM}\|\nabla f(x^{(k)})\|_1^2.$$
11. En déduire à l'aide des inégalités de la Partie II que
$$f(x^{(k+1)}) - p^\star \leqslant \left(1 - \frac{m}{nM}\right)\left(f(x^{(k)}) - p^\star\right).$$
12. Conclure sur la convergence de l'Algorithme 1.