UFR
DE MATHÉMATIQUES
ET INFORMATIQUE

UNIVERSITÉ
PARIS
DESCARTES

Optimisation, algorithmique (MML1E31) (M1 Maths, 2016-2017)

Examen partiel du mercredi 16 novembre 2016

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 4 points)*

Soient $n, p \in \mathbb{N}^*$ deux entiers non nuls. Soient $A \in \mathcal{M}_{p,n}(\mathbb{R})$ une matrice rectangulaire, $b \in \mathbb{R}^p$ un vecteur et $\lambda > 0$. On définit la fonction $f : \mathbb{R}^n \to \mathbb{R}$ par

$$f(x) = \|Ax - b\|^2 + \lambda \|x\|^2$$

(attention on utilise la même notation pour la norme dans $\mathbb{R}^n$ et dans $\mathbb{R}^p$, mais les vecteurs ne sont a priori pas de la même taille). On admet que f est deux fois différentiable sur $\mathbb{R}^n$.

1. Donner l'expression du gradient $\nabla f(x)$ et de la matrice hessienne $\nabla^2 f(x)$ de f en tout point $x \in \mathbb{R}^n$.
2. Montrer que la fonction f est fortement convexe.
3. Est-ce que l'algorithme de descente de gradient converge pour cette fonction f ?

Exercice 2. *(sur environ 6 points)*

1. Rappeler l'algorithme de calcul de pas de descente par la méthode de rebroussement en spécifiant le domaine de validité des paramètres α et β.
2. On suppose que les fonctions

   `function fx = f(x)` et `function gfx = gradf(x)`

   sont définies dans scilab. Donner le code d'une fonction scilab

   ```
   function t = rebroussement(x, d, alph, bet)
   ```

   qui prend en entrées les variables représentant le point $x \in \mathbb{R}^n$, la direction de descente $d \in \mathbb{R}^n$ et les paramètres α et β et renvoie le pas de descente t calculé par la méthode de rebroussement pour la fonction f.

Dans la suite de l'exercice on suppose que f est une fonction deux fois différentiable sur $\mathbb{R}$ et fortement convexe 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.

Soient $x \in \mathbb{R}^n$, $x \neq x^\star$, et $d \in \mathbb{R}^n$ une direction de descente au point x pour la fonction f.

3. Montrer que
$$f(x+td) \leqslant f(x) + t\langle \nabla f(x), d\rangle + \frac{t^2}{2} M\|d\|^2.$$
4. Montrer que le critère d'arrêt de la méthode de rebroussement est vérifié pour tout t tel que
$$0 < t \leqslant -2(1-\alpha)\frac{\langle \nabla f(x), d\rangle}{M\|d\|^2}.$$
5. En déduire une majoration du nombre d'itérations de la méthode de rebroussement au point x pour la direction de descente d.

Exercice 3. *(sur environ 10 points)*

1. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction convexe différentiable. Montrer que pour tous $x, y \in \mathbb{R}^n$,
$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \geqslant 0.$$

Dans la suite de l'exercice on suppose que f est une fonction deux fois différentiable sur $\mathbb{R}$ et fortement convexe 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.

2. Montrer que la fonction $g(x) = f(x) - \frac{m}{2}\|x\|^2$ est convexe.
3. En déduire que pour tous $x, y \in \mathbb{R}^n$,
$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \geqslant m\|y - x\|^2.$$

Dans la suite de l'exercice on cherche à étudier la convergence de l'algorithme de descente de gradient à pas fixe, à savoir

Algorithme 1 : Algorithme de descente de gradient à pas fixe

Données : Un point initial $x^{(0)} \in \mathbb{R}^n$, un seuil de tolérance $\varepsilon > 0$, un pas fixe $t > 0$
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. Mettre à jour x avec le pas fixe t dans la direction de descente $-\nabla f(x^{(k)})$:
    $x \leftarrow x^{(k+1)} = x^{(k)} - t\nabla f(x^{(k)})$;
    $k \leftarrow k + 1$;
fin

On suppose tout d'abord que l'algorithme converge.

4. Justifier en quelques mots l'intérêt de cet algorithme par rapport à l'algorithme de descente de gradient utilisant la méthode de rebroussement pour le calcul du pas de descente ?

5. On suppose que les fonctions

   `function fx = f(x)` et `function gfx = gradf(x)`

   sont définies dans scilab. Donner le code d'une fonction scilab

   ```
   function x = desc_gradient_pas_fixe(x0, eps, t)
   ```

   qui prend en entrées un point initial $x^{(0)} \in \mathbb{R}^n$, un seuil de tolérance $\varepsilon > 0$, et un pas fixe $t > 0$ et renvoie le point x calculé par l'algorithme de descente de gradient à pas fixe.

On passe maintenant à l'étude de la convergence. Soient $t > 0$ et $x^{(0)} \in \mathbb{R}^n$. On considère la suite de points définie par $x^{(k+1)} = x^{(k)} - t\nabla f(x^{(k)})$. On rappelle que comme pour tous $x, h \in \mathbb{R}^n$, $\langle \nabla^2 f(x)h, h\rangle \leqslant M\|h\|^2$, la fonction $x \mapsto \nabla f(x)$ est M-lipschitzienne, c'est-à-dire,

$$\forall x, y \in \mathbb{R}^n, \quad \|\nabla f(y) - \nabla f(x)\| \leqslant M\|y - x\|.$$

6. Justifier que l'on a

$$x^{(k+1)} - x^\star = x^{(k)} - x^\star - t(\nabla f(x^{(k)}) - \nabla f(x^\star)).$$

7. En déduire que

$$\|x^{(k+1)} - x^\star\|^2 \leqslant (1 - 2mt + M^2t^2)\|x^{(k)} - x^\star\|^2.$$

8. En déduire que pour tout $t \in \left]0, \frac{2m}{M^2}\right[$, l'algorithme de descente de gradient à pas fixe t converge linéairement.