Optimisation

Master 1 Statistique & Data Science, Ingénierie Mathématique, 2020-2021

Examen partiel du mercredi 31 mars 2021

Nombre de pages du sujet : 4. Durée 2h. **Aucun document n'est autorisé**

Exercice 1. *(sur environ 4 points)*

Soient $\alpha \in \mathbb{R}$ et $f_\alpha : \mathbb{R}^2 \to \mathbb{R}$ la fonctionnelle définie par

$$f_\alpha(x) = \alpha^2 x_1^2 + \frac{1}{2} x_2^2 + \alpha x_1 x_2 + \alpha^2 x_1 + x_2 - 5.$$

1. Montrer que f est une fonctionnelle quadratique et préciser la matrice symétrique A, le vecteur b et le réel c pour lesquels

$$f_\alpha(x) = \frac{1}{2}\langle Ax, x\rangle - \langle b, x\rangle + c.$$

2. Pour quelles valeurs de α la fonction f_α est-elle
   2.a) convexe ?
   2.b) strictement convexe ?
3. Pour quelles valeurs de α le problème d'optimisation

$$\inf_{x \in \mathbb{R}^2} f_\alpha(x)$$

admet des solutions ? En cas d'existence préciser s'il y a unicité de la solution. *On ne demande pas de calculer les solutions.*

Exercice 2. *(sur environ 10 points)*

1. Rappeler la définition d'une fonction convexe sur $\mathbb{R}^n$.
2. Soit $f : \mathbb{R}^n \to \mathbb{R}$. Montrer que f est convexe sur $\mathbb{R}^n$ si et seulement si

$$\forall p \geqslant 2, \forall x_1, \dots, x_p \in \mathbb{R}^n, \forall \theta \in [0,1]^p \text{ tel que } \sum_{k=1}^{p} \theta_k = 1, \quad f\left(\sum_{k=1}^{p} \theta_k x_k\right) \leqslant \sum_{k=1}^{p} \theta_k f(x_k).$$

3. Soit $p \geqslant 2$. Montrer que

$$\forall x \in \mathbb{R}^p, \forall \theta \in [0,1]^p \text{ tel que } \sum_{k=1}^{p} \theta_k = 1, \quad \langle \theta, x\rangle^2 \leqslant \sum_{k=1}^{p} \theta_k x_k^2.$$

On fixe maintenant un entier $p \geqslant 2$. On note

$$\mathcal{P} = \left\{ q \in]0,1[^p \text{ tel que } \sum_{k=1}^{p} q_k = 1 \right\}.$$

et on fixe un point $q \in \mathcal{P}$. On pose $g : \mathbb{R}^p \to \mathbb{R}$ la fonction définie par

$$g(x) = \ln\left(\sum_{k=1}^{p} e^{x_k}\right) - \langle q, x\rangle$$

et $S : \mathbb{R}^p \to \mathbb{R}^p$ la fonction définie par

$$(S(x))_k = \frac{e^{x_k}}{\sum_{j=1}^{p} e^{x_j}}, \quad k = 1, \dots, p.$$

4. Montrer que pour tout $x \in \mathbb{R}^p$, $S(x) \in \mathcal{P}$.
5. Justifier que g est différentiable et donner l'expression de son gradient en fonction de q et S.
6. Justifier que g est deux fois différentiable et montrer que pour tout $x \in \mathbb{R}^p$,

$$\nabla^2 g(x) = \operatorname{diag}(S(x)) - S(x)S(x)^T.$$

7. Montrer que g est convexe à l'aide de l'inégalité de la question 3.
8. En déduire que l'ensemble des minima globaux de g est la droite affine

$$\{x \in \mathbb{R}^p,\ \exists \lambda \in \mathbb{R},\ \forall k \in \{1, \dots, p\},\ x_k = \lambda + \ln q_k\}.$$

   et préciser la valeur minimale atteinte.
9. La fonction g est-elle strictement convexe ? Justifier.

Exercice 3. *(sur environ 11 points)*

Soit $n \geqslant 2$ un entier. Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice symétrique définie positive. On note $\|x\| = \sqrt{\langle x, x\rangle}$ la norme euclidienne usuelle de $\mathbb{R}^n$, $\|x\|_P = \sqrt{\langle Px, x\rangle}$ la norme associée à P et $\|x\|_{P^{-1}} = \sqrt{\langle P^{-1}x, x\rangle}$ la norme associée à P^{-1} la matrice inverse de P. On note par la suite $\gamma = \lambda_{\min}(P) > 0$ la valeur propre minimale de P et $\delta = \lambda_{\max}(P)$ la valeur propre maximale de P.

1. Démontrer que pour tout $x \in \mathbb{R}^2$ on a les inégalités

$$\gamma\|x\|^2 \leqslant \|x\|_P^2 \leqslant \delta\|x\|^2 \quad \text{et} \qquad \frac{1}{\delta}\|x\|^2 \leqslant \|x\|_{P^{-1}}^2 \leqslant \frac{1}{\gamma}\|x\|^2.$$

Dans la suite de l'exercice, $f : \mathbb{R}^n \to \mathbb{R}$ désigne une fonction deux fois différentiable sur $\mathbb{R}^n$ 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 et $p^\star = f(x^\star)$ la valeur minimale de f. On rappelle que pour tout $x \in \mathbb{R}^n$, on a

$$p^\star \geqslant f(x) - \frac{1}{2m}\|\nabla f(x)\|^2.$$

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

$$d^{(k)} = -P\nabla f(x^{(k)}).$$

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

Algorithme 1 : Algorithme de descente à étudier

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 $d^{(k)} = -P\nabla f(x^{(k)})$.
2. Déterminer un pas de descente $t^{(k)} > 0$ pour la direction de descente $d^{(k)}$ (par la méthode exacte ou 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

Implémentation de l'Algorithme 1 avec la méthode de rebroussement On suppose que les fonctions

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

sont définies dans scilab.

2. ***(sur 4 points)*** Ecrire une fonction scilab

```
function [Xk,FXk] = algo_descente(x0, eps, alph, bet, f, gradf, P)
```

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, les fonctions pour évaluer f et ∇f et la matrice P. La fonction renvoie la matrice `Xk` (de taille n par nombre d'itérations) contenant la liste des vecteurs $x^{(k)}$ donnée par l'Algorithme 1 **avec le calcul du pas par méthode de rebroussement** ainsi que la liste des valeurs `Fxk` de la fonctionnelle f aux points $x^{(k)}$. On veillera à ajouter une condition pour limiter le nombre d'itérations à 10 000.

Etude de la convergence de l'Algorithme 1 avec la méthode exacte 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 $x^\star$.

3. Montrer que
$$\|d^{(k)}\|^2 \leqslant \delta \|\nabla f(x^{(k)})\|_P^2.$$

4. Montrer que pour tout $t > 0$,
$$f(x^{(k+1)}) \leqslant f(x^{(k)} + td^{(k)}) \leqslant f(x^{(k)}) + \left(-t + \frac{M\delta t^2}{2}\right) \|\nabla f(x^{(k)})\|_P^2.$$

5. En déduire que
$$f(x^{(k+1)}) \leqslant f(x^{(k)}) - \frac{\gamma}{2M\delta}\|\nabla f(x^{(k)})\|^2.$$
6. Conclure sur la convergence de l'Algorithme 1.