Optimisation

Janvier 2019 **Tous documents autorisés**

Toute affirmation intuitive non argumentée sera à éviter

Exercice 1. Le problème de Képler (10pt).

Soient a, b, c trois réels positifs. On appelle $\mathcal{E}$ l'ensemble des points M de coordonnées (x, y, z) tels que $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$ (ellipsoide). On pose $\phi(x, y, z) = x^2/a^2 + y^2/b^2 + z^2/c^2 - 1$ et $f(x, y, z) = -xyz$.

(1) Donner une interprétation géométrique du problème $(\mathcal{P})$ suivant, dit de Képler :

$$\begin{array}{ll} \quad\max & xyz \\ (x,y,z) \in \mathcal{E} & \\ x \geq 0,\, y \geq 0,\, z \geq 0 & \end{array}$$

et montrer que ce problème admet un ensemble de solutions non vide.

La suite de l'exercice est centrée sur les conditions nécessaires d'optimalité au premier ordre (KKT).

(2) Supposer dans un premier temps qu'aucune des contraintes d'inégalité n'est active.
(2.1) Montrez que les contraintes sont qualifiées en tout point de l'ellipsoide $\mathcal{E}$.
(2.2) Exprimer la condition nécessaire d'optimalité au premier ordre et montrer qu'elle conduit à la résolution du système (S) :

$$\begin{cases} (x, y, z) \in \mathcal{E} \\ -yz + 2\lambda x/a^2 = 0 \\ -zx + 2\lambda y/b^2 = 0 \\ -xy + 2\lambda z/c^2 = 0 \end{cases}$$

(2.3) Montrer que ces équations impliquent $-3xyz + 2\lambda = 0$. En déduire en remplaçant dans (S) que $3x^2 = a^2$ et achever la résolution de S.
(2.4) Conclure les questions **(2)** précisément en revenant au problème d'optimisation de départ.

(3) On considère que la seule contrainte d'inégalité active est $x = 0$. Exprimer à nouveau la condition nécessaire d'optimalité du premier ordre. Montrez que cette condition n'admet pas de solution.

(4) En utilisant la condition nécessaire d'optimalité du premier ordre, montrer qu'il n'est pas possible d'avoir deux contraintes d'inégalité actives en un point solution. Est-il possible d'avoir les 3 contraintes d'inégalité actives en une solution ?

(5) En guise de résumé de l'exercice, décrire l'ensemble solution du problème $(\mathcal{P})$.

Problème. Plus profonde descente (10pt).

On note $\| \bullet \|_2$ la norme Euclidienne. Soit f une fonction continûment différentiable de $\mathbb{R}^n$ dans $\mathbb{R}$, minorée. On s'intéresse au problème

$$\min_{x\in\mathbb{R}^n} f(x);$$

On appelle *direction de plus profonde descente normalisée* (DDN) en x, tel que $\nabla f(x) \neq 0$, toute solution du problème en lavariable d suivant

$$\mathcal{P} \ : \ \min\{\nabla f(x)^T d, d \in \mathbb{R}^n, \|d\|_2 = 1\}.$$

(1) Préliminaire (peut-être admis en première lecture). Soit pour cette question x tel que $\nabla f(x) \neq 0$. Montrer que $d = -\frac{1}{\|\nabla f(x)\|}\nabla f(x)$ est une direction DNN. Est-ce l'unique DNN en x ?

On suppose de plus que la Hessienne de f est bornée au sens suivant

$$\exists M > 0, \forall x \in \mathbb{R}^n, \ \max_{d\neq 0} \frac{|d^T \nabla^2 f(x) d|}{\|d\|^2} \leq M.$$

On considère l'algorithme de descente suivant :

Plus profonde descente	
Initialisation	Choisir $\alpha \in]0, 1/2[$, $\beta \in]0, 1[$, un point de départ x_0
Itération	For $j = 0, 1, 2, \ldots$ 1. Si $\nabla f(x) = 0$, arrêt de l'algorithme 2. On pose $d_j = -\frac{1}{\|\nabla f(x_j)\|}\nabla f(x_j)$ 3. Recherche linéaire : Poser $t_j = \beta^i$ où i est le plus petit $i \in \mathbb{N}$ tel que $f(x + \beta^i \|\nabla f(x_j)\|_2\, d_j) \leq f(x) + \alpha\beta^i \|\nabla f(x_j)\|_2 \nabla f(x_j)^T d_j$ 4. Mettre à jour : $x_{j+1} = x_j + t_j \|\nabla f(x_j)\|_2 d_j$ finFor

(2) Etude de la convergence de la recherche linéaire.

(2.1) Montrer par un développement de Taylor-Lagrange sans reste que pour tout $x \in \mathbb{R}^n$, pour tout $t > 0$, et pour tout $d \in \mathbb{R}^n$, $f(x + t\,d) \leq f(x) + t\nabla f(x)^T d + \frac{1}{2}Mt^2\|d\|^2$.

(2.2) On pose $d = -\frac{1}{\|\nabla f(x)\|}\nabla f(x)$. Vérifier que que pour tout $t > 0$, $f(x + t\|\nabla f(x)\|_2 d) \leq f(x) - t\|\nabla f(x)\|_2^2 + \frac{1}{2}Mt^2\|\nabla f(x)\|_2^2$.

(2.3) En déduire que pour $\hat{t} = 1/M$, on a $f(x + \hat{t}\|\nabla f(x)\|_2 d) \leq f(x) + \alpha \frac{\|\nabla f(x)\|_2}{M} \nabla f(x)^T d$.

(2.4) Montrer que si $0 \leq t \leq 1/M$, on a $-t + Mt^2/2 \leq -t/2$. Déduire alors de 2.2 que la recherche linéaire est bien définie (l'étape 3. de l'algorithme s'arrête en un nombre fini de tests de valeurs pour i croissant).

(2.5) Question difficile. Montrer que t_j de l'étape 3. de l'algorithme vérifie

$$f(x_j) - f(x_{j+1}) \geq \alpha \min(1, \beta/M)\|\nabla f(x_j)\|_2^2 \geq 0.$$

On pourra, pour $d = -\frac{1}{\|\nabla f(x)\|}\nabla f(x)$ représenter pour $t > 0$ les fonctions $t \mapsto f(x - t\,d)$, $t \mapsto f(x) - \alpha\, t\|\nabla f(x)\|^2$ et $t \mapsto f(x) - t\|\nabla f(x)\|_2^2 + \frac{1}{2}Mt^2\|\nabla f(x)\|_2^2$.

(3) Quelle propriété de f permet-elle d'obtenir un résultat de convergence de l'algorithme à partir de **(2.5)** ? Enoncer précisément ce résultat de convergence.