Algorithmes d'optimisation – L3 MINT/DL 2017/2018 – Université Paris-Sud

Partiel du 13/03/2018

Durée : 3h, à répartir entre questions écrites et numériques.

Les documents (notes de cours, ancien TP, etc) sont interdits. La qualité de la rédaction sera prise en compte dans la correction. Vous pouvez bien sûr admettre les résultats d'une question afin de traiter les questions suivantes.

1 Problème : point de Fermat d'un triangle

Soit $a, b, c \in \mathbb{R}^2$ trois points distincts et $p > 0$. On considère le problème d'optimisation

$$\min_{\mathbb{R}^2} f(x) \text{ où } f(x) = \|x-a\|^p + \|x-b\|^p + \|x-c\|^p\,, \tag{1}$$

où $\|\cdot\|$ est la norme euclidienne. On appelle *point de Fermat* du triangle a, b, c un minimiseur de (1) quand $p = 1$.

Q1.[Existence] Montrer que le problème d'optimisation admet un minimiseur pour tout $p > 0$.

Q2.[Convexité] Montrer que les fonctions $x \mapsto \|x\|$ et $x \mapsto \|x\|^p$ ($p \geq 1$) sont convexes. En déduire que la fonction f est convexe pour $p \geq 1$.

Q3.[Cas $p = 2$] Dans le cas $p = 2$, calculer $\nabla f(x)$, et montrer que l'unique minimiseur de f est $x^* = \frac{1}{3}(a + b + c)$.

Q4.[Non-différentiabilité] Démontrer que $x \mapsto \|x\|$ n'est pas différentiable en $(0, 0)$, et en déduire que pour $p = 1$, f n'est pas différentiable aux points a, b, c.

1.1 Régularisation du problème avec $p = 1$

Comme f n'est pas différentiable sur $\mathbb{R}^2$, on ne peut pas (a priori) utiliser la méthode de descente de gradient pour résoudre (1). On introduit donc une approximtion régulière de la norme euclidienne,

$$n_\varepsilon(x) = \sqrt{\varepsilon^2 + x_1^2 + x_2^2} = \sqrt{\varepsilon^2 + \|x\|^2},$$

et on va approcher le problème de Fermat (1) par

$$\min_{\mathbb{R}^2} f_\varepsilon(x) \text{ où } f_\varepsilon(x) = n_\varepsilon(x-a) + n_\varepsilon(x-b) + n_\varepsilon(x-c), \tag{2}$$

Q5.[Gradient et hessienne de n_ε] Démontrer que $n_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^2)$, et que

$$\frac{\partial n_\varepsilon}{\partial x_i}(x) = \frac{x_i}{n_\varepsilon(x)}, \qquad \frac{\partial^2 n_\varepsilon}{\partial x_i \partial x_j}(x) = \frac{n_\varepsilon(x)^2 \delta_{ij} - x_i x_j}{n_\varepsilon(x)^3},$$

où $\delta_{ij} = 1$ si $i = j$ et 0 sinon. En déduire l'expression du gradient et de la hessienne de n_ε.

Q6.[Gradient et hessienne de f_ε.] Montrer que

$$\nabla f_\varepsilon(x) = \nabla n_\varepsilon(x-a) + \nabla n_\varepsilon(x-b) + \nabla n_\varepsilon(x-c)$$

$$\mathrm{D}^2 f_\varepsilon(x) = \mathrm{D}^2 n_\varepsilon(x-a) + \mathrm{D}^2 n_\varepsilon(x-b) + \mathrm{D}^2 n_\varepsilon(x-c)$$

Q7.[Estimations sur $\mathrm{D}^2 n_\varepsilon$] Démontrer que $\langle \mathrm{D}^2 n_\varepsilon(x)v|v\rangle = \frac{n_\varepsilon(x)^2\|v\|^2-\langle x|v\rangle^2}{n_\varepsilon(x)^3}$, et en déduire que pour tout point de x tel que $n_\varepsilon(x) \leq R$, on a, en posant $M_R = 1/\varepsilon$ et $m_R = \varepsilon^2/R^3$,

$$M_R \|v\|^2 \geq \langle \mathrm{D}^2 n_\varepsilon(x)v|v\rangle \geq m_R \|v\|^2 .$$

Q8.[Convergence de la descente de gradient avec backtracking] Soit $x^{(0)} \in \mathbb{R}^2$, $R = f_\varepsilon(x^{(0)})$ et soit $S = \{x \in \mathbb{R}^2 \mid f_\varepsilon(x) \leq R\}$.

1. Montrer que S est compact.
2. Montrer que pour tout $x \in S$, $n_\varepsilon(x - z) \leq R$ où $z \in \{a, b, c\}$. En déduire que

$$3M_R \|v\|^2 \geq \langle \mathrm{D}^2 f_\varepsilon(x)v|v\rangle \geq 3m_R \|v\|^2$$

3. Conclure à l'aide d'un théorème du cours.

Q9.[Convergence du problème régularisé (*)] Démontrer que (2) admet une unique solution x^*_ε. Montrer que si (1) admet une unique solution x^*, alors $\lim_{\varepsilon\to 0} x^*_\varepsilon = x^*$.

1.2 Résolution numérique du problème régularisé

Télécharger `http://quentin.mrgt.fr/files/cours/m315/partiel-2018-sujet.ipynb` *et le renommer nom-prenom.ipynb / Écrire les réponses dans les cellules prévues.*

2 Exercice : Méthode de gradient à pas fixe

Soit $f \in \mathcal{C}^2(\mathbb{R}^d)$ une fonction convexe. On considère l'algorithme du gradient à pas fixe : pour $x^{(0)} \in \mathbb{R}^d$ et $\tau > 0$, on définit

$$x^{(k+1)} = x^{(k)} - \tau \nabla f(x^{(k)}) \tag{3}$$

On suppose qu'il existe des constantes $0 < m < M$ telles que

$$\forall x \in \mathbb{R}^d, \forall v \in \mathbb{R}^d, m \|v\|^2 \leq \langle \mathrm{D}^2 f(x)v|v\rangle \leq M \|v\|^2 . \tag{4}$$

Q10.[Estimations]

1. Montrer que $\nabla f(y) = \nabla f(x) + \int_0^1 \mathrm{D}^2 f((1-t)x + ty)(y - x)\mathrm{d}t$ *(Indication : appliquer le théorème fondamental de l'analyse à $g : t \in [0, 1] \mapsto \frac{\partial f}{\partial x_i}(x_t)$ où $x_t = (1-t)x + ty$).*
2. Déduire de 1. que $\langle \nabla f(x) - \nabla f(y)|x - y\rangle \geq m \|x - y\|^2$.
3. Soit A symétrique positive telle que $\forall v$, $\langle Av|v\rangle \leq M \|v\|^2$. Montrer que $\|Av\| \leq M \|v\|$. (*Indication : diagonaliser A en base orthonormée, montrer que les valeurs propres de A sont comprises entre 0 et M.*).
4. Déduire de 1. et 3. que $\|\nabla f(x) - \nabla f(y)\| \leq M \|x - y\|$.

Q11.[Contraction] On considère la fonction $F_\tau : \mathbb{R}^d \to \mathbb{R}^d$ définie par $F_\tau(x) = x - \tau\nabla f(x)$. Nous allons donner une condition sur τ garantissant que F_τ est contractante.

1. Démontrer que pour tout $x, y \in \mathbb{R}^d$,

$$\|F_\tau(x) - F_\tau(y)\|^2 = \|x - y\|^2 + \tau^2 \|\nabla f(x) - \nabla f(y)\|^2 - 2\tau\langle \nabla f(x) - \nabla f(y)|x - y\rangle$$

2. Déduire de la question précédente que $\|F_\tau(x) - F_\tau(y)\|^2 \leq (1 + M^2\tau^2 - 2m\tau) \|x - y\|^2$.

Q12.[Convergence] Déduire de la question précédente et du théorème du point fixe contractant que si $0 < \tau < 2\frac{m}{M^2}$, la suite $x^{(k)}$ définie par (3) converge vers l'unique minimiseur de f sur $\mathbb{R}^d$.

Q13.[Taux optimal] Pour quel τ la constante de Lipschitz de F_τ donnée par Q11.2 est elle minimale ?