**Algorithmes d'optimisation – L3 MINT/DL 2018/2019 – Université Paris-Sud**

# Partiel du 26/02/2019

Durée : 3h, à répartir entre questions écrites et numériques (partie 1.2).

*Les documents (notes de cours, ancien TP, etc) sont interdits. La qualité de la rédaction sera prise en compte dans la correction ; en particulier, il faut citer les hypothèses des théorèmes utilisés.* ***Vous pouvez bien sûr admettre les résultats d'une question afin de traiter les questions suivantes.***

## 1 Problème : projection sur le simplexe

On s'intéressera au problème d'optimisation suivant :

$$\min_{x=(x_1,x_2)\in\Omega} f(x) \text{ où } f(x) = \frac{1}{2}\|x-z\|^2 + h_\varepsilon(x) \tag{1}$$

Dans ce problème, $z = (z_1, z_2) \in \mathbb{R}^2$ est fixé, $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0, x_1 + x_2 < 1\}$ et $h_\varepsilon(x) := -\varepsilon\left[\log(x_1) + \log(x_2) + \log(1 - (x_1 + x_2))\right].$

### 1.1 Étude théorique

**Q1.** [Existence] Soit $x^{(0)} \in \Omega$ et $S = \{x \in \Omega \mid f(x) \le f(x^{(0)})\}$.

1. Montrer que $f \in \mathcal{C}^0(\Omega)$.
2. Montrer que tous les termes apparaissant dans la définition de $f$ sont positifs. En déduire que $S \subseteq \Delta_\eta$, où $\eta = \exp(-f(x^{(0)})/\varepsilon)$ et où
$$\Delta_\eta = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \ge \eta, x_2 \ge \eta, x_1 + x_2 \le 1 - \eta\} \subseteq \Omega.$$
3. En déduire que l'ensemble $S$ est fermé.
4. Montrer que le problème (1) admet un minimiseur.

**Q2.** [Convexité et unicité]

1. Montrer que l'ensemble $\Omega$ est convexe.
2. Montrer que si $g : \Omega \to \mathbb{R}$ est strictement convexe et $h : \Omega \to \mathbb{R}$ est convexe, alors la fonction $g + h$ est strictement convexe.
3. Sans calculer de dérivée, en déduire que la fonction $f$ est strictement convexe, et que le problème (1) admet un unique minimiseur.

**Q3.** [Gradient et hessienne] Montrer que

$$\nabla h_\varepsilon(x) = -\varepsilon\left[\left(\frac{1}{x_1}, \frac{1}{x_2}\right) + \frac{1}{x_1 + x_2 - 1}(1, 1)\right] \qquad \nabla f(x) = x - z + \nabla h_\varepsilon(x)$$

$$\mathrm{D}^2 h_\varepsilon(x) = \varepsilon\begin{pmatrix} \frac{1}{x_1^2} & 0 \\ 0 & \frac{1}{x_2^2}\end{pmatrix} + \frac{\varepsilon}{(1 - x_1 - x_2)^2}\begin{pmatrix} 1 & 1 \\ 1 & 1\end{pmatrix} \qquad \mathrm{D}^2 f(x) = \mathrm{id} + \mathrm{D}^2 h_\varepsilon(x)$$

**Q4.** [Conditionnement du problème]

1. Montrer que $\forall v \in \mathbb{R}^2, \langle \mathrm{D}^2 f(x)v|v\rangle \ge \|v\|^2$.
2. Montrer que $\forall v \in \mathbb{R}^2, \langle \mathrm{D}^2 h_\varepsilon(x)v|v\rangle \le \varepsilon\left[\max\left(\frac{1}{x_1^2}, \frac{1}{x_2^2}\right) + \frac{2}{(1-x_1-x_2)^2}\right]\|v\|^2.$

3. Déduire que $\forall x \in \Delta_\eta$, $\forall v \in \mathbb{R}^2$, $\|v\|^2 \leq \langle \mathrm{D}^2 f(x) v | v\rangle \leq M \|v\|^2$ où $M = \left(1 + \frac{3\varepsilon}{\eta^2}\right)$.

On considère les deux méthodes d'optimisation suivantes, partant d'un point $x^{(0)} \in \Omega$ :

$$\text{(Gradient)} \begin{cases} d^{(k)} = -\nabla f(x^{(k)}) \\ t^{(k)} = \text{pas\_armijo}(x^{(k)}, d^{(k)}) \\ x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)} \end{cases} \qquad \text{(Newton)} \begin{cases} d^{(k)} = -[\mathrm{D}^2 f(x^{(k)})]^{-1} \nabla f(x^{(k)}) \\ t^{(k)} = \text{pas\_armijo}(x^{(k)}, d^{(k)}) \\ x^{(k+1)} = x^{(k)} + t^{(k)} d^{(k)} \end{cases}$$

où $\text{pas\_armijo}(x, v) = \max\{t \mid \ell \in \mathbb{N}, t = \beta^\ell, x + tv \in \Omega, f(x + tv) \leq f(x) + t\alpha\langle \nabla f(x) | v\rangle\}$. Dans la suite, on prendra $\beta = \frac{1}{2}$, $\alpha = \frac{3}{10}$.

**Q5.** [Convergence] Justifier la convergence de la méthode (Gradient), en citant précisément les hypothèses du théorème utilisé. Donner une borne sur le taux de convergence.

### 1.2 Résolution numérique du problème

*Copier le fichier `/commun/doc/merigot/M315-partiel.ipynb` dans le répertoire `reMise_copie` situé sur le bureau.* ***Renommer le M315-nom-prenom.ipynb****, et répondre aux questions dans les cellules prévues.*

## 2 Exercice : Optimisation coordonnée par coordonnée

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^2$ et vérifiant pour tout $x \in \mathbb{R}^2$, et $i \in \{1, 2\}$, $\left|\frac{\partial^2 f}{\partial e_i^2}(x)\right| \leq M$.

**Q6.** Soit $x = (x_1, x_2) \in \mathbb{R}^2$.

1. Montrer que $\forall t \in \mathbb{R}, f(x_1 + t, x_2) \leq f(x_1, x_2) + t\frac{\partial f}{\partial e_1}(x_1, x_2) + Mt^2/2$.
2. En déduire que $\min_{t\in\mathbb{R}} f(x_1 + t, x_2) \leq f(x_1, x_2) - \frac{1}{2M}\left[\frac{\partial f}{\partial e_1}(x_1, x_2)\right]^2$

On se donne un point $x^{(0)} \in \mathbb{R}^2$ tel que l'ensemble $S = \{x \in \mathbb{R}^2 \mid f(x) \leq f(x^{(0)})\}$ est compact, et on considère l'algorithme suivant :

$$\forall k \in \mathbb{N}, \quad \begin{cases} t^{(k)} = \arg\min_{t\in\mathbb{R}} f\left(x_1^{(k)} + t, x_2^{(k)}\right), \\ x^{(k+\frac{1}{2})} = x^{(k)} + (t^{(k)}, 0) \\ s^{(k+\frac{1}{2})} = \arg\min_{s\in\mathbb{R}} f\left(x_1^{(k+\frac{1}{2})}, x_2^{(k+\frac{1}{2})} + s\right) \\ x^{(k+1)} = x^{(k+\frac{1}{2})} + (0, s^{(k+\frac{1}{2})}) \end{cases} \tag{2}$$

**Q7.** Déduire de la question précédente que pour tout $k \geq 1$,

$$f\left(x^{(k+\frac{1}{2})}\right) \leq f(x^{(k)}) - \frac{1}{2M}\left(\frac{\partial f}{\partial e_1}(x^{(k)})\right)^2 = f(x^{(k)}) - \frac{1}{2M}\left\|\nabla f(x^{(k)}\right\|^2. \tag{3}$$

*(Indication : montrer que $\forall k \geq 0$, $\frac{\partial f}{\partial e_2}(x^{(k+1)}) = 0$ en utilisant la troisième ligne de (2).)*

On démontrerait de la même manière que

$$f\left(x^{(k+1)}\right) \leq f\left(x^{(k+\frac{1}{2})}\right) - \frac{1}{2M}\left\|\nabla f(x^{(k+\frac{1}{2})}\right\|^2. \tag{4}$$

**Q8.** [Convergence] Déduire de (3)–(4) que $\sum_{k\geq 1}\left\|\nabla f(x^{(k)})\right\|^2 < +\infty$, puis que $\lim_{k\to+\infty} \nabla f(x^{(k)}) = 0$. Que peut-on dire de plus si $f$ est supposée strictement convexe ?