MASTER 1 INFO 2016–2017

Optimisation Algorithmique

EXAMEN, durée 2h

Tout document est interdit.
Tout appareil électronique, même à titre d'horloge, est également interdit.
Nombre de pages de l'énoncé : 2
Il faut justifier les réponses. Il sera tenu compte de la rédaction de la copie.

Questions de cours

Soit f une fonction définie sur $\mathbb{R}^2$, on note $\nabla f(x)$ le gradient de f au point x et $H_f(x)$ la matrice hessienne de f en $x \in \mathbb{R}^2$.
On suppose qu'il existe un minimum unique $x^* = \arg\min_x f(x)$ et l'on note $(x^{(\mathrm{k})})_{k\in\mathbb{N}} = \left({x_1}^{(\mathrm{k})}, {x_2}^{(\mathrm{k})}\right)_{k\in\mathbb{N}}$ une suite de $\mathbb{R}^2$ qui converge vers $x^* \in \mathbb{R}^2$ quand k tend vers l'infini.

1. Sous quelles condition(s) un vecteur $d \in \mathbb{R}^2$ est une direction de descente de f en $x \in \mathbb{R}^2$?

2. Pour déterminer le pas de descente σ de f en x dans la direction d on utilise souvent le “backtracking”. Expliquez de quoi il s'agit. Donner l'algorithme (pseudo-code).

3. On rappelle que la méthode itérative qui permet de calculer la suite $(x^{(\mathrm{k})})$ est d'ordre $r > 1$ s'il existe $C \in \mathbb{R}_+^*$ tel que pour k suffisamment grand
$$\frac{\left\|x^{(\mathrm{k+1})} - x^*\right\|}{\left\|x^{(\mathrm{k})} - x^*\right\|^r} \leq C \; .$$
Comment définit on la précision $prec(k)$?
Donner une majoration de $prec(k+1)$ par $prec(k)$ pour une méthode d'ordre 2.
Expliquez le résultat grâce à un exemple où $prec(0) = 1$.

Exercice 1

Soit $f(x_1, x_2) = (x_1 + x_2 - 1)^2 + 10(x_1 - x_2)^2$.

1. Déterminer $v^* = \min_x f(x)$ et $x^* = \arg\min_x f(x)$.

2. Calculer $\nabla f(x)$ et $H_f(x)$ au point $x = (x_1, x_2) \in \mathbb{R}^2$.
En déduire que f peut s'écrire comme une fonction quadratique que l'on précisera.

3. Au point $x^{(0)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, calculer $d^{(0)}$, la direction de descente selon la méthode du gradient.

4. Déterminer le pas σ_0 qui minimise f le long de cette direction.
   Pour cela, poser $\varphi(\sigma) = f(x^{(0)} + \sigma d^{(0)})$ et minimiser φ sur $\mathbb{R}_+^*$.
   Calculer ensuite $f(x^{(1)})$. Conclusion ?

5. Au point $\tilde{x}^{(0)} = \begin{pmatrix} \frac{10}{11} \\ 1 \end{pmatrix}$, calculer $\tilde{d}^{(0)}$, la direction de descente selon la méthode du gradient.
   Sans calculs, prévoir la forme de $\tilde{d}^{(1)}$. Que peut-on prévoir comme comportement de la suite $(\tilde{x}^{(\mathrm{n})})$?

6. Comparez les suites obtenues pour la méthode du gradient à partir des valeurs initiales $x^{(0)}$ et $\tilde{x}^{(0)}$.

7. Justifier (sans calculs) que, indépendamment de la valeur de $x^{(0)}$, la méthode de Newton converge en une itération.
   Expliquez.

Exercice 2

On pose $g(x_1, x_2) = -\Big(\ln(x_1) + \ln(x_2) + \ln(1 - x_1) + \ln(1 - x_2)\Big)$ pour $(x_1, x_2) \in]0, 1[^2$.

1. Déterminer $\nabla g(x_1, x_2)$ et $H_g(x_1, x_2)$.

2. Déterminer $x^* = (x_1^*, x_2^*) \in \mathbb{R}^2$ tel que $v_m = g(x_1^*, x_2^*) = \min\limits_{(x_1,x_2)\in \mathrm{dom}(g)} g(x_1, x_2)$.

3. Comment faut-il adapter l'algorithme de descente de gradient (avec backtracking) pour pouvoir l'appliquer à g ?

4. Est-ce qu'il faut prendre les mêmes précautions pour la méthode de Newton ?