EPFL
ENAC INTER TRANSP-OR
Prof. M. Bierlaire

RECHERCHE OPÉRATIONNELLE
GC/SIE
Printemps 2009/2010

Corrigé 8

Problème 1

On calcule le gradient :

$$\nabla f(x) = \begin{pmatrix} 2x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{pmatrix},$$

le Hessien :

$$H(x) = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$$

et son inverse :

$$H^{-1}(x) = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}.$$

On s'aperçoit que l'on a :

$$H^{-1}(x)\nabla f(x) = \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 2x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x.$$

Ainsi, quel que soit $x^{(0)} \in \mathbb{R}^2$, on a

$$x^{(1)} = x^{(0)} - H^{-1}(x^{(0)})\nabla f(x^{(0)}) = x^{(0)} - x^{(0)} = 0.$$

Ainsi, en une itération et indépendamment du point de départ $x^{(0)}$, la méthode de Newton converge vers le minimum de la fonction. On peut montrer que cette propriété est vérifiée pour toute fonction quadratique définie positive.

Problème 2

a) Le gradient et la matrice hessienne de f sont :

$$\begin{aligned} \nabla f(\boldsymbol{x}) &= \begin{pmatrix} x_1 \\ 9x_2 \end{pmatrix} \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^2 \\ \nabla^2 f(\boldsymbol{x}) &= \begin{pmatrix} 1 & 0 \\ 0 & 9 \end{pmatrix} \qquad \forall\, \boldsymbol{x} \in \mathbb{R}^2 \end{aligned}$$

Le pas de l'algorithme de la plus forte pente est donné par :

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k - \left(\frac{\nabla f(\boldsymbol{x}_k)^T \nabla f(\boldsymbol{x}_k)}{\nabla f(\boldsymbol{x}_k)^T \nabla^2 f(\boldsymbol{x}_k) \nabla f(\boldsymbol{x}_k)} \right) \nabla f(\boldsymbol{x}_k) \\ &= \begin{pmatrix} (x_1)_k \\ (x_2)_k \end{pmatrix} - \left(\frac{(x_1^2)_k + 81(x_2^2)_k}{(x_1^2)_k + 729(x_2^2)_k} \right) \begin{pmatrix} (x_1)_k \\ 9(x_2)_k \end{pmatrix}. \end{aligned}$$

On obtient les valeurs suivantes données par l'algorithme :

k	$\boldsymbol{x}_k$	$f(\boldsymbol{x}_k)$
0	$\begin{pmatrix} 9 & 1 \end{pmatrix}^T$	45
1	$\begin{pmatrix} \frac{36}{5} & \frac{-4}{5} \end{pmatrix}^T$	28.8
2	$\begin{pmatrix} \frac{144}{25} & \frac{16}{25} \end{pmatrix}^T$	18.432
3	$\begin{pmatrix} \frac{576}{125} & \frac{-64}{125} \end{pmatrix}^T$	11.796

b) Le gradient et la matrice hessienne de g sont :

$$\begin{aligned} \nabla g(\boldsymbol{x}) &= \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} & \forall\, \boldsymbol{x} \in \mathbb{R}^2 \\ \nabla^2 g(\boldsymbol{x}) &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} & \forall\, \boldsymbol{x} \in \mathbb{R}^2 \end{aligned}$$

Le pas de l'algorithme de la plus forte pente est donné par :

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k - \left(\frac{\nabla g(\boldsymbol{x}_k)^T \nabla g(\boldsymbol{x}_k)}{\nabla g(\boldsymbol{x}_k)^T \nabla^2 g(\boldsymbol{x}_k) \nabla g(\boldsymbol{x}_k)} \right) \nabla g(\boldsymbol{x}_k) \\ &= \begin{pmatrix} (x_1)_k \\ (x_2)_k \end{pmatrix} - \begin{pmatrix} (x_1)_k \\ (x_2)_k \end{pmatrix} \end{aligned}$$

On obtient les valeurs suivantes données par l'algorithme :

k	$\boldsymbol{x}_k$	$g(\boldsymbol{x}_k)$
0	$\begin{pmatrix} 9 & 1 \end{pmatrix}^T$	82
1	$\begin{pmatrix} 0 & 0 \end{pmatrix}^T$	0

On voit que $\boldsymbol{x}_1 = \boldsymbol{x}_*$.

On constate que si la fonction à optimiser est bien conditionnée l'algorithme de la plus forte pente converge en une itération. Notons que ceci arrive très rarement en pratique.

c) On passe d'une itération à l'autre avec l'algorithme de Newton de la manière suivante :

$$\begin{aligned} \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \boldsymbol{s}_k \\ \text{avec} \qquad \nabla^2 f(\boldsymbol{x}_k) \boldsymbol{s}_k &= -\nabla f(\boldsymbol{x}_k) \end{aligned}$$

On obtient les valeurs suivantes données par l'algorithme :

k	$\boldsymbol{x}_k$	$\boldsymbol{s}_k$	$f(\boldsymbol{x}_k)$
0	$\begin{pmatrix} 9 & 1 \end{pmatrix}^T$	$\begin{pmatrix} -9 & -1 \end{pmatrix}^T$	45
1	$\begin{pmatrix} 0 & 0 \end{pmatrix}^T$	$-$	0

On constate que $\boldsymbol{x}_1 = \boldsymbol{x}_*$.

On voit donc que pour une fonction quadratique définie positive, l'algorithme de Newton converge en une seule itération.

Problème 3

a) $\nabla f(\boldsymbol{x}) = \begin{pmatrix} 4x_1^3 + 2x_1 \\ 2x_2 \end{pmatrix} \qquad \forall \boldsymbol{x} \in \mathbb{R}^2$

b) $\boldsymbol{d}_c$ est une direction de descente, car :

$$\nabla f(\boldsymbol{x}_c)\boldsymbol{d}_c = \begin{pmatrix} 6 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix} = -20 < 0$$

c) Pour chacun des α, vérifions les conditions (1) et (2).

- α_1 vérifie la condition (2) mais pas la condition (1)
  (1) $f(-2,0) = 20 \not\leq 1 = 3 + 0.1 \cdot 1 \cdot \begin{pmatrix} 6 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix}$
  (2) $\begin{pmatrix} -26 & 0 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix} = 108 \geq -10 = 0.5 \cdot \begin{pmatrix} 6 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix}$
- α_2 vérifie la condition (1) mais pas la condition (2)
  (1) $f(0.7, 0.9) = 1.54 \leq 2.8 = 3 + 0.1 \cdot 0.1 \cdot (-20)$
  (2) $\begin{pmatrix} 2.772 & 1.8 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix} = -10.116 \not\geq -10 = 0.5 \cdot \begin{pmatrix} 6 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix}$
- α_3 vérifie les deux conditions
  (1) $f(-0.5, 0.5) = 0.5625 \leq 2 = 3 + 0.1 \cdot 0.5 \cdot (-20)$
  (2) $\begin{pmatrix} -1.5 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix} = 3.5 \geq -10 = 0.5 \cdot \begin{pmatrix} 6 & 2 \end{pmatrix} \begin{pmatrix} -3 \\ -1 \end{pmatrix}$

Remarque :

- α_1 se trouve à droite de la région admissible des α possibles.
- α_2 se trouve à gauche de la région admissible des α possibles.
- α_3 se trouve dans la région admissible des α possibles.

Problème 4

(1) Aucune recherche linéaire n'est effectuée. En se déplaçant d'un pas complet le long de la direction de la plus forte pente en x_k, rien ne garantit que l'on a $f(x_{k+1}) < f(x_k)$!

(2) La direction de recherche d_k est constante et ne fait pas intervenir d'informations sur f ! On n'a donc pas la garantie que d_k constitue une direction de descente en x_k pour une fonction f quelconque.

(3) La direction de recherche d_k est la direction de la plus forte préconditionnée par une matrice diagonale définie positive et constitue donc bien une direction de descente. La recherche linéaire inexacte d'Armijo est bien définie et il existe des pas α qui remplissent cette condition.

(4) La direction de recherche est la direction de Newton modifiée de façon à obtenir une direction de descente. En effet, la définition du poids que l'on met sur la diagonale de la matrice hessienne garantit la définie positivité de la matrice perturbée. Une recherche linéaire exacte est utilisée pour déterminer le déplacement le long de la direction de recherche.

(5) D'une part, la suite des directions de recherche $\{d_k\}_{k \in \mathbb{N}}$ tend vers 0 ! D'autre part, le paramètre β étant choisi à 1 dans la condition d'Armijo, il n'est pas possible de trouver un pas α remplissant cette condition !

(6) La récurrence de cette méthode est telle que la direction de recherche prise est finalement $-d_k$, ce qui représente l'opposé de la direction de la plus forte pente et constitue donc une direction de montée !

April 22, 2010 – mbi/mfe