Corrigé 4

Problème 1

a) $\nabla f(x_1, x_2) = \begin{pmatrix} 6x_1 \\ 4x_2^3 \end{pmatrix}$.

b) $f(x_k) = 19$, $\nabla f(x_k) = \begin{pmatrix} 6 \\ -32 \end{pmatrix}$, $d = -\nabla f(x_k) = \begin{pmatrix} -6 \\ 32 \end{pmatrix}$.

It	α	$x_k + \alpha \cdot d$	$f(x_k + \alpha \cdot d)$	$f(x_k) + \alpha\beta\nabla f(x_k)^T d$
0	1	$\begin{pmatrix} -5 \\ 30 \end{pmatrix}$	810075	19-106
2	0.25	$\begin{pmatrix} -0.5 \\ 6 \end{pmatrix}$	1296.76	19-26.5
4	0.0625	$\begin{pmatrix} 0.625 \\ 0 \end{pmatrix}$	1.171875	19-6.625

$x_{k+1} = \begin{pmatrix} 0.625 \\ 0 \end{pmatrix}$.

c)

It	α	$x_k + \alpha \cdot d$	$f(x_k + \alpha \cdot d)$	$f(x_k) + \alpha\beta\nabla f(x_k)^T d$
0	1	$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$	0	19-7

$x_{k+1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

d)

$x = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $f(x) = 0$.

Problème 2

$$f(\boldsymbol{x}) = f(x, y) = x^2 + 2y^2$$

L'itération pour chacun des algorithmes mentionnés dans la donnée est défini comme :

$$\boldsymbol{x_{k+1}} = \boldsymbol{x_k} - \alpha_k D_k \nabla f(\boldsymbol{x_k}),$$

où

1) $D_k = I$ pour l'algorithme de la plus forte pente,

2) $D_k = (\nabla^2 f(\boldsymbol{x_k}))^{-1}$ pour l'algorithme de Newton.

1) Pour l'algorithme de la plus forte pente, la condition de Wolfe revient à :

$$f\left(\boldsymbol{x_k} - \alpha_k \nabla f(\boldsymbol{x_k})\right) \leq f(\boldsymbol{x_k}) - \alpha\beta \nabla f(\boldsymbol{x_k})^T \nabla f(\boldsymbol{x_k})$$

$\nabla f(\boldsymbol{x}) = \begin{pmatrix} 2x \\ 4y \end{pmatrix}$ et on a donc (afin de simplifier la notation, on omet l'indice k) :

$$f(\boldsymbol{x} - \alpha \nabla f(\boldsymbol{x})) = (x - \alpha 2x)^2 + 2(y - \alpha 4y)^2 = x^2(1-2\alpha)^2 + 2y^2(1-4\alpha)^2$$

et

$$f(\boldsymbol{x}) - \alpha\beta \nabla f(\boldsymbol{x})^T \nabla f(\boldsymbol{x}) = x^2 + 2y^2 - \alpha\beta(4x^2 + 16y^2)$$

Quelques manipulations algébriques permettent de reformuler la condition de Wolfe :

$$0 \leq \alpha_k \leq (1-\beta)\frac{x_k^2 + 4y_k^2}{x_k^2 + 8y_k^2}$$

Pour $\boldsymbol{x_0} = (9\ 1)^T$ et $\beta = 1/100$, on obtient :

(S_0) $0 \leq \alpha_0 \leq 1,044$. En choisissant $\alpha_0 = 0.9$,
$\boldsymbol{x_1} = \boldsymbol{x_0} - \frac{9}{10}\nabla f(\boldsymbol{x_0}) = \begin{pmatrix} 1 - 81/5 \\ 9 - 18/5 \end{pmatrix}$.

Les itérations suivantes suivent le même schéma de calcul.

2) Pour l'algorithme de Newton, la condition de Wolfe s'écrit :

$$f\left(\boldsymbol{x_k} - \alpha_k(\nabla^2 f(\boldsymbol{x_k}))^{-1}\nabla f(\boldsymbol{x_k})\right) \leq f(\boldsymbol{x_k}) - \alpha\beta \nabla f(\boldsymbol{x_k})^T (\nabla^2 f(\boldsymbol{x_k}))^{-1} \nabla f(\boldsymbol{x_k})$$

La matrice hessienne de f est donnée par $\begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$. Ainsi $D_K = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \end{pmatrix}$.

Après quelques calculs, on obtient :

$$0 \leq \alpha_k \leq 2(1-\beta)$$

En choisissant $\alpha_k = 1/2$, on obtient la suite de points définie par $x_{k+1} = \frac{1}{2}x_k$.

Problème 3

(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 !

March 21, 2011 – mbi/mfe