Université de Versailles Saint-Quentin-En-Yvelines
L3, Optimisation et Applications (LSMA651)
Année 2011-2012
Enseignants: L. Dumas, J.P. Bartier, M.Z. Dauhoo
`http://dumas.perso.math.cnrs.fr/LSMA651.html`

TD 5 : METHODES DE DESCENTE

Exercice 1.
On munit $\mathbb{R}^n$ du produit scalaire euclidien usuel, noté $< \cdot, \cdot >$, la norme euclidienne associée étant notée $\| \, . \, \|$. Pour tout $v \in \mathbb{R}^n$, $\| v \| = \sqrt{v^T v}$ sur $\mathbb{R}^n$. On définit la norme d'opérateur de M comme suit:

$$\| M \| = \max_v \{ \| Mv \|, | \| v \| = 1 \} .$$

Montrer que les propriétés suivantes sont équivalentes pour une matrice M symétrique ($h > 0$ est fixé):

(a) M est inversible et $\| M^{-1} \| \leq \dfrac{1}{h}$

(b) $\| Mv \| \geq h. \| v \| \; \forall v \in \mathbb{R}^n$

(c) $| \lambda_i(M) | \geq h$ pour tout valeur propre $\lambda_i(M)$ de M, $i = 1, \cdots, n$

Exercice 2.
On suppose que f est C^2 sur $\mathbb{R}^n$. Montrer que

$$\forall (x, z) \in \mathbb{R}^n \times \mathbb{R}^n, \quad \nabla f(z) - \nabla f(x) = \int_0^1 [H(x + t(z - x))] \, (z - x) dt,$$

où $H(y)$ représente la matrice hessienne de f au point $y \in \mathbb{R}^n$.

Exercice 3.
Soit f C^2 sur $\mathbb{R}^n$ et $x^* \in \mathbb{R}^n$ tel que $\nabla f(x^*) = 0$. On suppose que $H(x)$ vérifie les conditions suivantes:

(1) il existe $h > 0$ tel que $\| H^{-1}(x^*) \| \leq \dfrac{1}{h}$.

(2) Il existe $\beta > 0$ et $L > 0$ tel que

$$\| H(x) - H(x^*) \| \leq L \| x - x^* \| \text{ pour tout } x \text{ vérifiant } \| x - x^* \| \leq \beta.$$

Soient $\| x - x^* \| < \gamma := \min(\beta, \frac{2h}{3L})$ et $x_N := x - H^{-1}(x)\nabla f(x)$. Montrer les inégalites suivantes:

(a) $\| x_N - x^* \| \leq \| x - x^* \|^2 \dfrac{L}{2(h - L \| x - x^* \|)}$.

(b) $\| x_N - x^* \| \leq \| x - x^* \| < \gamma$.

Exercice 4.
La méthode de Descente de Gradient
Soit la fonction $f(\mathbf{x})$ définie et C^1 sur $\mathbb{R}^n$. On sait que dans un voisinage d'un point $\mathbf{a} \in \mathbb{R}^n$, f diminue le plus rapidement si l'on passe dans la direction de la pente négative de f à $\mathbf{a}$, c'est à dire la direction $-\nabla \mathbf{f}(\mathbf{a})$.

On commence avec une estimation, x_0, pour un minimum local de f et considère la séquence $x_0, x_1, x_2, \cdots$, où

$$\mathbf{x}_{i+1} \;=\; \mathbf{x}_i \;-\; \alpha \nabla \mathbf{f}(\mathbf{x_i}), i = 0, 1, 2 \cdots$$

tel que

$$\mathbf{f}(\mathbf{x_0}) \geq \mathbf{f}(\mathbf{x_1}) \geq \mathbf{f}(\mathbf{x_2}) \cdots$$

$f(x)$ est définie comme suit:

$$\mathbf{f}(\mathbf{x}) \;=\; \frac{1}{2}\mathbf{x}^T \begin{pmatrix} 6 & -2 \\ -2 & 6 \end{pmatrix} \mathbf{x} \;+\; [-1 \;\; -1]\mathbf{x}$$

.

(i) Quel est l'unique minimum (global) x^* de f?

(ii) En commençant avec $\mathbf{x_0} = [0 \; 0]^T$ et $\alpha = 0.1$, calculer les deux itérés suivants $\mathbf{x_1}$ et $\mathbf{x_2}$.

(iii) Trouver la taille de pas maximum (α) pour que la méthode converge vers x^* quel que soit le point x_0.