Université de Versailles Saint-Quentin-En-Yvelines
L3, Optimisation et Applications (LSMA651)
Année 2010-2011
Enseignants : L. Dumas, T. Horsin
`http://www.math.uvsq.fr/~dumas/LSMA651`

TD5 : OPTIMISATION SANS CONTRAINTE, METHODE DU GRADIENT ET DE NEWTON (suite)

L'objectif de cet exercice est de définir une nouvelle règle de recherche linéaire et de démontrer qu'elle génère une méthode de descente convergente.

On considère f une fonction C^1 de $\mathbb{R}^n$ dans $\mathbb{R}$, telle que $\lim_{||x|| \to +\infty} f(x) = +\infty$. On note g la fonction gradient de f définie de $\mathbb{R}^n$ dans $\mathbb{R}^n$. On suppose que g est Lipschitzienne sur tout ensemble $S_{x_0} = \{x \in \mathbb{R}^n, f(x) \leq f(x_0)\}$.

1. Montrer que f possède un minimum global x^* pour lequel $g(x^*) = 0$.

2. On cherche à construire une méthode de descente $(x_k)_{k\in\mathbb{N}}$ définie à partir de $x_0 \in \mathbb{R}^n$ par la formule :

$$x_{k+1} = x_k + t_k d_k$$

où d_k désigne une direction de descente (c'est à dire telle que $(d_k, g(x_k)) < 0$) et t_k le pas dans cette direction déterminé par les conditions suivantes : on note $q(t) = f(x_k + td_k)$ et on choisit deux réels m_1 et m_2 tels que $0 < m_1 < m2 < 1$. Alors :

a) t_k est satisafaisant si $q(t_k) \leq q(0) + m_1 t_k q'(0)$ et $q'(t_k) \geq m_2 q'(0)$.
b) t_k est trop grand si $q(t_k) > q(0) + m_1 t_k q'(0)$.
c) t_k est trop petit si $q(t_k) \leq q(0) + m_1 t_k q'(0)$ et $q'(t_k) < m_2 q'(0)$

Donner un exemple graphique des valeurs de t_k admissibles dans le cas d'une fonction tracée 'à la main'. Comment interpréter la condition a) ?

3. Pour déterminer une telle valeur de t_k, on utilise l'algorithme suivant (de paramètres $t_{init} > 0$ et $\lambda > 1$) :

- Etape 0: on part de $t = t_{init} > 0$. On note $t_g = 0$ et $t_d = 0$.
- Etape 1: On teste t suivant les critères a) b) et c):
  - si a), c'est fini
  - si b) (t trop grand), on note $t_d = t$ et on passe à l'étape 2.
  - si c) (t trop petit), on note $t_g = t$ et on passe à l'étape 2.
- Etape 2
  - si $t_d = 0$, on calcule $t_{new} = \lambda t$
  - Si $t_d > 0$, on calcule $t_{new} = \frac{t_g + t_d}{2}$.
- Etape 3: on boucle avec l'étape 1 avec le nouveau t.

On cherche à montrer que cet algorithme converge en un nombre fini d'étapes pour la fonction f choisie en préambule.

3.1 Montrer qu'il est impossible de répéter une infinité de fois la première condition de l'étape 2.

3.2 On suppose qu'on répète une infinité de fois la deuxième condition de l'étape 2. Montrer dans ce cas que t_g et t_d sont deux suites adjacentes de limite t^*.

3.3 Montrer alors qu'en ce point, $q(t^*) = q(0) + m_1 t^* q'(0)$. En utilisant un taux d'accroissement bien choisi, en déduire que $q'(t^*) \geq m_1 q'(0)$ ainsi que $q'(t^*) \leq m_2 q'(0)$. Conclure.

4. On cherche à montrer que la méthode de descente ainsi construite est convergente vers un point critique de f lorsque $d_k = -g(x_k) = -g_k$ (direction du gradient).

4.1 Montrer que $q'(0) = -||g_k||^2$ dans ce cas.

4.2 Montrer que $m_1 ||g_k|| \cdot ||x_{k+1} - x_k|| \leq f(x_k) - f(x_{k+1})$.

4.3 Montrer que $(1 - m_2)||g_k||^2 \leq (g_k - g_{k+1}, g_k)$ puis $(1 - m_2)||g_k|| \leq L||x_{k+1} - x_k||$ où L désigne la constante de Lipschitz de g dans S_{x_0}.

4.4 En composant et en sommant ces relations, montrer que $\lim_{k\to+\infty} g_k = 0$.