TD 2 Optimisation : méthodes de gradient

Exercice 1 (*CC 2015*)

On s'intéresse ici à un nouveau principe de recherche linéaire pour une méthode à direction de descente.

Soit f une fonction C^1 de $\mathbb{R}^n$ dans $\mathbb{R}$, telle que $\lim_{||x||\to+\infty} f(x) = +\infty$. On note g le gradient de la fonction f défini de $\mathbb{R}^n$ dans $\mathbb{R}^n$. On suppose que g est Lipschitz continu sur $\mathbb{R}^n$ avec une constante de Lipschitz égale à L.

On rappelle qu'une méthode de descente consiste à définir une suite $(x_k)_{k\in\mathbb{N}}$ partant d'un point $x_0 \in \mathbb{R}^n$ avec la relation :

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

où d_k est une direction de descente (telle que $< d_k, g(x_k) >< 0$) et t_k est le pas correspondant.

Ici, le pas est donné par la règle suivante : $t_k = 0$ si $g(x_k) = 0$ et sinon :

$$t_k = \inf\{t \geq 0, \quad h_k'(t) = 0,\ h_k(t) < h_k(0)\}$$

où pour tout $t \geq 0$, $h_k(t) = f(x_k + td_k)$.

1. Prouver que la condition choisie est valide, c'est à dire que t_k existe pour tout $k \geq 0$. Représenter graphiquement un exemple de tel pas pour une fonction arbitraire (non convexe).

2. On dit que la condition Z est vérifié pour un pas de descente si on a :

$$f(x_{k+1}) \leq f(x_k) - C||g(x_k)||^2 \cos^2(\theta_k)$$

pour une constante $C > 0$ donnée indépendante de k et où

$$\cos(\theta_k) = \frac{- < d_k, g(x_k) >}{||g(x_k)||.||d_k||}$$

On cherche à prouver que la condition Z est bien vérifiée pour le pas choisi ici.

2 a) En utilisant une égalité de Taylor, prouver que pour tout $t \in [0, t_k]$

$$h(t_k) \leq h(t) \leq h(0) + t(d_k, g(x_k)) + t^2 L\frac{||d_k||^2}{2} \tag{1}$$

2b) Trouver la valeur t, nommée t_{min} où le terme à droite dans l'inégalité (1) est minimal.

2c) Prouver que

$$0 \leq (d_k, g(x_k)) + t_k L ||d_k||^2$$

et en déduire que

$$f(x_{k+1}) \leq f(x_k) - \frac{(d_k, g(x_k))^2}{2L||d_k||^2}$$

2c) Conclure.

3. Prouver que la condition Z implique :

$$\sum_{k=1}^{+\infty} ||g(x_k)||^2 \cos^2(\theta_k) < +\infty$$

4. On suppose que $d_k = -g(x_k)$ (direction de descente du gradient). Montrer que la méthode de descente ainsi construite converge en un sens à préciser.