

TD Optimisation: méthodes de descente d'ordre 1

Exercice 1 (examen 2016)

On cherche à minimiser sur $\mathbb{R}^3$ la fonction :

$$J(x, y, z) = x^2 + 2y^2 + z^2 - x + y - z$$

1. Donner, en le justifiant, la solution exacte au problème considéré.
2. On veut utiliser la méthode de descente suivant la direction du gradient avec une stratégie de type backtracking avec condition d'Armijo pour approcher la solution à partir du point $X_0 = (0, 1, 1)$. Que vaut la direction de descente à la première itération ?

On rappelle que la condition d'Armijo pour un point de départ X_0 et une direction de descente d s'écrit :

$$J(X_0 + \alpha d) \leq J(X_0) + \beta \alpha < d, \nabla J(X_0) >$$

Calculer explicitement puis représenter graphiquement dans le cas présent la fonction $\alpha \mapsto J(X_0 + \alpha d)$ ainsi que les valeurs de α vérifiant la condition d'Armijo pour $\beta = 0.1$.

3. On prend $\alpha_{init} = 1$ et $\tau = 0.1$ (constante de backtracking). Que vaut X_1 ?

Exercice 2

On considère f une fonction C^1 de $\mathbb{R}^n$ dans $\mathbb{R}$, telle que $\lim_{\|x\| \rightarrow +\infty} f(x) = +\infty$. 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 suite $(x_k)_{k \in \mathbb{N}}$ convergeant vers un minimum (local ou global) de f . On suppose qu'il est possible de définir correctement celle-ci à partir de la donnée de $x_0 \in \mathbb{R}^n$ quelconque et de la relation :

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

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 supposé satisfaire la conditions suivante :

$$q(t_k) \leq q(0) + m_1 t_k q'(0) \quad \text{et} \quad q'(t_k) \geq m_2 q'(0)$$

où on a noté $q(t) = f(x_k + t d_k)$ et où m_1 et m_2 sont deux réels tels que $0 < m_1 < m_2 < 1$.

Donner un exemple graphique des valeurs de t_k admissibles dans le cas d'une fonction à une variable tracée 'à la main'.

3. On peut 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). Proposer un algorithme permettant de construire la suite x_k , c'est à dire en particulier de déterminer un réel t_k convenable.