Université de Versailles Année 2023-2024
http://dumas.perso.math.cnrs.fr/M1-optim.html semestre 1

TD 3 Optimisation: optimisation locale sans contraintes

Exercice 1 (*une itération "à la main", 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 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 vérfiant 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 (*point de Cauchy, 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.

Exercice 3 (*règle de Wolfe, CC 2020*)

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 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 $q(t) = f(x_k + td_k)$ et où m_1 et m_2 sont deux réels tels que $0 < m_1 < m2 < 1$.

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 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).

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

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

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

3.4 En déduire que $\lim_{k \to +\infty} g_k = 0$.

Question subsidiaire : proposer un algorithme permettant de construire la suite x_k, c'est à dire en particulier de déterminer un réel t_k convenable.