Université Paris-Dauphine
Département MIDO
Chargé de cours : F.-X. Vialard.
Chargés de TD/TP : G. Jankowiak, A. Frouvelle.
DUMI2E troisième année

Optimisation numérique
année 2013/2014

TD n^{o}3 : Première partie

Une méthode de quasi-Newton très simple

1. Introduction

On va étudier une méthode proposée par J. Barzilai et J. Borwein en 1988. C'est une méthode proche de la méthode de descente de gradient habituelle mais qui utilise un propriété de type quasi-Newton pour définir le pas. C'est une méthode qui en pratique accélère la méthode de descente de gradient tout en conservant un coût comparable.

2. Présentation de l'algorithme

Initialisation de $x_0 \in \mathbb{R}^n$, $g_0 = \nabla f(x_0)$ et une tolérance $tol \in \mathbb{R}$
Choix du pas β_0 par section dorée.
$x_1 = x_0 - \beta_0 g_0$.
Tant que ($\|g_k\| > tol$) **faire**
- $s_k = x_k - x_{k-1}$
- $y_k = g_k - g_{k-1}$
- $\beta_k = \frac{\langle s,s \rangle}{\langle s,y \rangle}$
- $g_k = \nabla f(x_k)$
- $x_{k+1} = x_k - \beta_k g_k$
- $k = k + 1$

Fin
Retourner x_k.

Algorithme 1: Algorithme de Barzilai et Borwein

Dans cette partie, on considère $f : \mathbb{R}^n \to \mathbb{R}$ une fonction C^2. On cherche à approcher la Hessienne de f par αId avec α un réel.

Question 1. Écrire y_k en fonction de la Hessienne de f par une formule d'intégration.

On dit qu'une matrice H satisfait la condition de quasi-Newton pour s et y si $Hs_k = y_k$. On s'intéresse à des matrices $H = \alpha Id$ pour $\alpha \in \mathbb{R}$.

Question 2. Trouver α tel que $\|y_k - \alpha s_k\|^2$ soit minimal.

Question 3. Rappeler le choix de la direction de descente dans la méthode de Newton et retrouver la formulation de l'algorithme.

Dans la question 2, l'approximation se fait sur la Hessienne de f. On peut aussi faire cette approximation directement sur l'inverse de la Hessienne de f.

Question 4. Écrire la condition de quasi-Newton correspondante (en fonction de H^{-1}). En déduire une autre valeur possible pour β_k en suivant la même approche que dans la question 2.

3. Analyse de la convergence

On suppose maitenant que f est une fonction quadratique. On considère dans la suite une matrice symétrique définie positive $A \in M_n(\mathbb{R})$ et $b \in \mathbb{R}^n$. On s'intéresse au problème de minimisation de la fonction $f(x) = \frac{1}{2}\langle x, Ax\rangle + \langle x, b\rangle$ avec $\langle \cdot, \cdot \rangle$ le produit scalaire usuel. On pose $\alpha_k = 1/\beta_k$.

Question 5. Montrer que $g_{k+1} = \frac{1}{\alpha_k}(\alpha_k Id - A)g_k$.

On note $\lambda_1 \leq \ldots \leq \lambda_n$ les valeurs propres (strictement positive) de la matrice A et $B = (v_1, \ldots, v_n)$ une base de vecteurs propres orthonormée associée.

Question 6. Soit $k \geq 1$. Calculer les coordonnées (z_i^k) de g_k en fonction de la suite α_k et des coordonnées de g_1 dans la base B.

Question 7. Montrer $\alpha_k = \frac{\langle g_{k-1}, Ag_{k-1}\rangle}{\langle g_{k-1}, g_{k-1}\rangle}$.

Question 8. Soit $\delta = |\frac{\lambda_n}{\lambda_1} - 1|$. En déduire que

$$|z_i^{k+1}| \leq \delta |z_i^k|\,,$$

et en déduire la convergence de la méthode dans le cas $\delta < 1$.

Question 9. Démontrer la convergence dans le cas où $n = 2$ et δ quelconque.

Question 10. Proposer une approche pour prouver la convergence en dimension $n \geq 2$ et δ est quelconque.