ENS Cachan, DPT Maths

Optimisation numérique M1 – TD3 – Newton, Quasi-Newton

Florian De Vuyst, Adrien Le Coënt, Lara Raad
CMLA UMR 8536, ENS Cachan

29 Septembre 2016

1 Méthode de la sécante

Pour $f : \Omega \subset \mathbb{R} \to \mathbb{R}$, cette méthode consiste à remplacer, dans la méthode de Newton, $f'(x_n)$ par son approximation par différences finies

$$\frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

On obtient ainsi la *méthode de la sécante*

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n), \quad x = 1, 2, \ldots .$$

Cette méthode est souvent donnée sous la forme

$$x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}.$$

Montrer que si f est deux fois continûment dérivable au voisinage de x et si x est racine simple de f, alors il existe un voisinage de x tel que, pour tout couple $(x_0, x_1 \neq 0)$ appartenant à ce voisinage, la suite (x_n) obtenue par la méthode de la sécante converge vers x avec une convergence d'ordre $(1 + \sqrt{5})/2 \simeq 1.618$ au moins.

Indication : en notant $e_n = x_n - x$, on pourra montrer en utilisant un développement de Taylor, que

$$e_{n+1} \sim e_n e_{n-1} \frac{f''(x)}{2f'(x)}$$

2 Méthode de Davidon-Fletcher-Powell (DFP)

Soit $J : \mathbb{R}^n \to \mathbb{R}$ une fonctionnelle de classe $\mathscr{C}^1$. On rappelle qu'en optimisation, une matrice S^{k+1} satisfait la condition de quasi-Newton (CQN) sur J si

$$S^{k+1}\left(\nabla J(u^{k+1}) - \nabla J(u^k)\right) = u^{k+1} - u^k.$$

Les méthodes de correction de rang 2 recherchent S^{k+1} sous la forme

$$S^{k+1} = S^k + a\, v^k (v^k)^T + b\, w^k (w^k)^T,$$

où a et b sont deux réels et v^k et w^k sont deux vecteurs tels que

$$\begin{aligned} v^k &= u^{k+1} - u^k, \\ w^k &= S^k\left(\nabla J(u^{k+1}) - \nabla J(u^k)\right), \end{aligned}$$

avec a et b tels que

$$\begin{aligned} a\,\langle v^k, \nabla J(u^{k+1}) - \nabla J(u^k)\rangle &= 1, \\ b\,\langle w^k, \nabla J(u^{k+1}) - \nabla J(u^k)\rangle &= -1 \end{aligned}$$

répondent à la contrainte CQN.

La méthode DFP consiste donc à choisir un vecteur u^0 quelconque, une matrice S^0 symétrique définie positive, une direction

$$d^k = -S^k \nabla J(u^k),$$

à effectuer une recherche linéaire

$$\alpha^k = \arg\min_{\alpha \in \mathbb{R}^+} J(u^k + \alpha\, d^k)$$

et une mise à jour

$$\begin{aligned} u^{k+1} &= u^k + \alpha^k\, d^k, \\ S^{k+1} &= S^k + a\, v^k (v^k)^T + b\, w^k (w^k)^T, \end{aligned}$$

où S^{k+1} est construite comme indiquée ci-dessus.

On utilise les notations suivantes :

$$\begin{aligned} \delta^k &= u^{k+1} - u^k, \\ \gamma^k &= \nabla J(u^{k+1}) - \nabla J(u^k). \end{aligned}$$

Montrer que

1. A l'étape k, si S^k est symétrique définie positive et que $\langle \delta^k, \gamma^k\rangle > 0$, alors la matrice S^{k+1} préserve cette propriété.

2. Si J est quadratique de Hessien A, matrice symétrique définie positive, alors la méthode D.F.P. est telle que

$$\langle \delta^i, \delta^j \rangle_A = 0 \quad 0 \leq i < j \leq k,$$
$$\left(S^{k+1}A\right)\delta^i = \delta^i, \quad 0 \leq i \leq k$$

3. En déduire que la méthode converge en n itérations et que

$$S^n = A^{-1}.$$