Ensimag 2A – automne 2015
TDs optimisation
Alejandro Blumentals, Franck Iutzeler, Jérôme Malick

# TP 3 – une méthode de quasi-Newton : BFGS

**Objectif :** Cette séance termine l'écriture d'un code d'optimisation (pour des problèmes différentiables, sans contraintes), reprenant les TPs précédents et s'attachant maintenant au calcul d'une bonne direction de descente.

**Rappel :** La méthode de BFGS consiste à réaliser l'itération

$$x_{k+1} = x_k - t_k W_k \nabla f(x_k)$$

où $t_k$ est donné par la recherche linéaire de Wolfe et la matrice symétrique définie positive $W_k$ est calculée par la formule de récurrence

$$W_{k+1} = W_k - \frac{s_k y_k^T W_k + W_k y_k s_k^T}{y_k^T s_k} + \left[1 + \frac{y_k^T W_k y_k}{y_k^T s_k}\right] \frac{s_k s_k^T}{y_k^T s_k},$$

avec $s_k = x_{k+1} - x_k$ et $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$. Le schéma général d'une méthode de quasi-Newton est alors :

1. avec l'itéré initial $x_0$, se donner une matrice initiale $W_0$ symétrique définie positive ;
2. connaissant le gradient $\nabla f(x_k)$, calculer la direction $d_k = -W_k \nabla f(x_k)$ ;
3. calculer le pas $t_k$ par recherche linéaire de Wolfe ;
4. connaissant le nouvel itéré $x_{k+1}$, appeler le simulateur et calculer la nouvelle matrice $W_{k+1}$.

**Exercice 1 – Écriture de la méthode BFGS.** Programmer la méthode BFGS : pour cela, récupérer la structure générale du TP1, insérer le calcul du pas du TP2 puis modifier la direction en ajoutant la matrice de Quasi-Newton. On prendra à la matrice identité (commande `eye` en scilab) pour $W_0$ et un pas initial égal à 1 en entrée de la recherche linéaire de Wolfe.

**Exercice 2 – Tests.**

**a)** Évaluer ce code avec la fonction de Rosenbrock

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - {x_1}^2)^2$$

partant du point initial $(-1, 1.2)$. Comparer avec les autres méthodes (gradient à pas constant, gradient avec Wolfe, quasi-Newton à pas constant...). Même chose avec d'autres points initiaux (par exemple $(0, 1.2)$).

**b)** Visualiser les déplacements des itérés sur une "carte" de la fonction.

**c)** Estimer les vitesses de convergence asymptotique - en tracant par exemple la suite des valeurs

$$\log f(x_k) \qquad \text{et éventuellement} \qquad \log \|\nabla f(x_k)\|.$$

Comparer la vitesse de convergence avec celle de la methode de gradient avec Wolfe (que l'on récupère en enlevant la mise à jour de $W_k$ dans quasi-Newton).

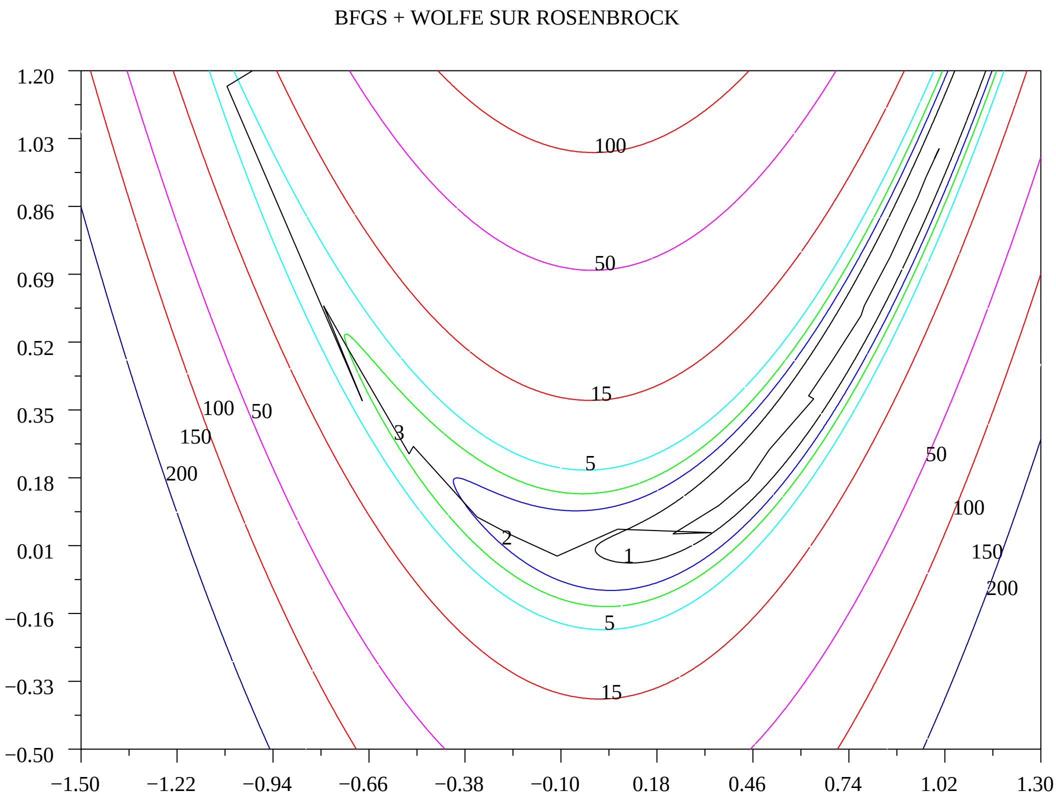


**Figure 1** Minimisation de la fonction Rosenbrock par BFGS, en partant de $x_0 = (-0.9, 1.2)$