École des Ponts ParisTech
Mars 2018 - Avril 2018

Cours Optimisation et Contrôle

DistribEauPti - Projet sur les réseaux de distribution d'eau

Marc-Antoine Augé et Matthieu Roux

Table des matières

1 Séance 1 - Définition de l'oracle

1.1 Calculs différentiels

On pose

$$F(q) = \frac{1}{3} < q, r \circ q \circ |q| > + < p_r, A_r q >$$

et

$$q(q_c) = q^{(0)} + Bq_c$$

On cherche à calculer $\nabla F(q(q_c))$ et $HF(q(q_c))$ le Hessien.
Remarquons tout d'abord que les matrices sont à coefficients réels donc transposition et adjonction sont deux opérations identiques.
Commencons par $\nabla F(q)$ en écrivant le produit de Hadamard terme à terme et le produit scalaire sous forme de somme :

$$F(q) = \frac{1}{3} \sum_{i=1}^{n} q_i^2 . r_i . |q_i| + < p_r, A_r q >$$

On note alors ϵ_i le signe de q_i :

$$F(q) = \frac{1}{3} \sum_{i=1}^{n} \epsilon_i . r_i . q_i^3 + < p_r, A_r q >$$

D'où immédiatement, étant donné que le gradient du second terme est $A_r^T . p_r$:

$$\nabla F(q) = (\epsilon_i . r_i . q_i^2)_{1 \leq i \leq n} + A_r^T . p_r$$

On peut le réécrire sans la notation ϵ_i et en utilisant un produit de Hadamard :

$$\boxed{\nabla F(q) = (r_i . q_i . |q_i|)_{1 \leq i \leq n} + A_r^T . p_r = r \circ q \circ |q| + A_r^T . p_r}$$

Puis par composition, comme $q(q_c) = q^{(0)} + q_c$, on a, en notant par abus de notation : $F(q_c) = F(q(q_c))$:

$$\boxed{\nabla F(q_c) = B^T \nabla F(q) = B^T (r \circ q \circ |q| + A_r^T . p_r)}$$

Pour calculer le Hessien, on a tout d'abord, en notant $\mathrm{diag}(X)$ la matrice diagonale possédant sur sa diagonale les coefficients de X :

$$HF(q) = 2.\mathrm{diag}(r \circ |q|)$$

Par composition, étant donné que le $H(q(q_c)) = 0$:

$$\boxed{HF(q_c) = HF(q(q_c)) = 2.B^T.\mathrm{diag}(r \circ |q|).B}$$

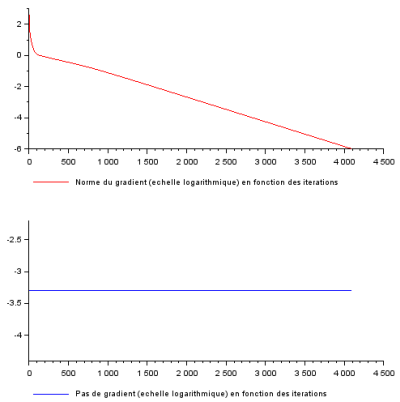
2 Séances 2 et 3 - Implémentations d'algorithmes

Nous avons cherché à résoudre le problème d'optimisation avec différents algorithmes vus en cours :

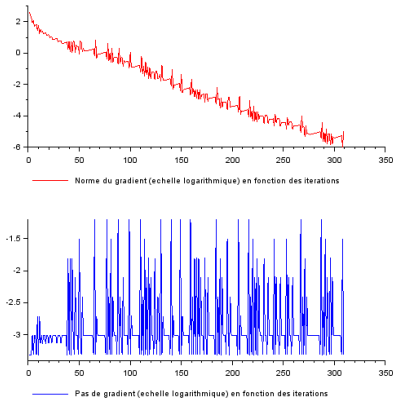
- Mise en place d'une recherche linéaire vérifiant les conditions de Wolfe avec l'algorithme de Fletcher-Lemaréchal. Cette étape permet de trouver un pas optimal et conduit notamment à l'**algorithme de gradient à pas variable**.
- Mise en place de l'algorithme de **Polak-Ribière**, algorithme de gradient conjugué non-linéaire où la direction étudiée dépend de la direction de l'étape précédente.
- Mise en place de **l'algorithme de BFGS** (Broyden-Fletcher-Goldfarb-Shanno) qui est une méthode de quasi-Newton où on utilise une approximation de l'inverse du Hessien de la fonction à minimiser.
- Mise en place de **l'algorithme de Newton** qui n'approxime pas l'inverse du Hessien

Nous avons obtenu les résultats présentés sur la Figure **??** et synthétisés sur le Tableau 2. Nous remarquons notamment les points suivants :

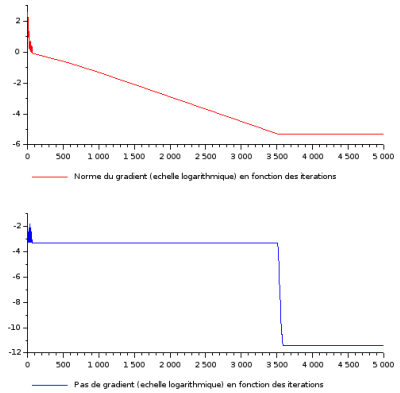
- L'ajout de la recherche linéaire via l'algorithme de Fletcher-Lemaréchal est une très bonne amélioration qui demande un peu plus de calculs par itération mais en en faisant près de $10\times$ moins demeure deux fois plus rapide qu'un algorithme de gradient à pas fixe.
- L'algorithme de Newton est extrêmement rapide (6 itérations) mais demande une bonne connaissance du problème (le Hessien) : toutefois une approximation de ce Hessien, comme à travers l'algorithme BFGS donne également d'excellents résultats.
- L'initialisation dans l'algorithme de Fletcher-Lemaréchal est importante. Si on initialise avec le α précédent, comme sur la Figure **??**, alors la convergence est très mauvaise.
- On remarque que l'algorithme de Fletcher-Lemaréchal fait osciller le pas de gradient.



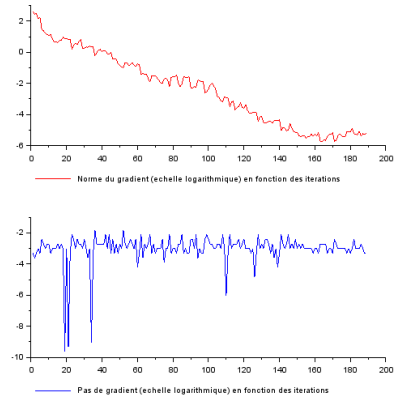
(a) Gradient à pas fixe



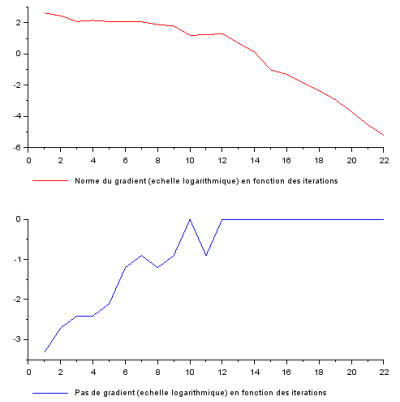
(b) Gradient à pas variable (conditions de Wolfe)



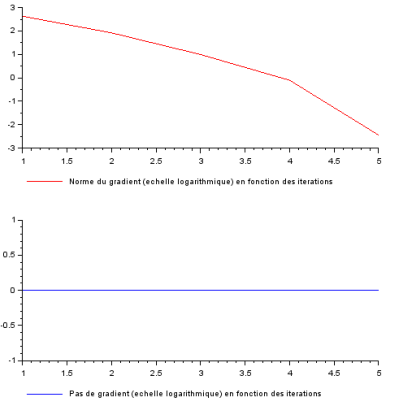
(c) Gradient à pas variable où Fletcher-Lemaréchal initialité avec α précédent



(d) Polak-Ribière



(e) BFGS



(f) Newton

Figure 1 – Comparaison des différents algorithmes implémentés

Méthode	Scilab	Pas fixe	Pas variable	Polak-Ribiere	BFGS	Newton
Itérations	-	4094	310	190	23	6
Temps CPU (s)	0	0.34375	0.15625	0.109375	0.03125	0
Critère optimal	-3.734007	-3.734007	-3.734007	-3.734007	-3.734007	-3.734007
Norme du gradient	$10^{-}7$	$10^{-}6$	$9.10^{-}7$	$7.10^{-}7$	$4.10^{-}7$	$7.85.10^{-}8$

Figure 2 – Comparaison des différents algorithmes