WIKIPÉDIA

BFGS

En mathématiques, la méthode de **Broyden-Fletcher-Goldfarb-Shanno** (**BFGS**) est une méthode permettant de résoudre un problème d'optimisation non linéaire sans contraintes.

La méthode BFGS est une solution souvent utilisée lorsque l'on veut un algorithme à directions de descente.

L'idée principale de cette méthode est d'éviter de construire explicitement la matrice hessienne et de construire à la place une approximation de l'inverse de la dérivée seconde de la fonction à minimiser, en analysant les différents gradients successifs. Cette approximation des dérivées de la fonction permet l'application de la méthode de Quasi-Newton (une variante de la méthode de Newton) de manière à trouver le minimum dans l'espace des paramètres.

La matrice hessienne n'a pas besoin d'être recalculée à chaque itération de l'algorithme. Cependant, la méthode suppose que la fonction peut être approchée localement par un développement limité quadratique autour de l'optimum.

Sommaire

Base

La recherche de la direction de descente $\mathbf{p}_k$ à l'étape k est donnée par la solution de l'équation suivante, équivalente à l'équation de Newton :

$$\mathrm{B}_k \mathbf{p}_k = -\nabla f(\mathbf{x}_k).$$

Une recherche linéaire dans la direction $\mathbf{p}_k$ est alors utilisée pour trouver le prochain point $\mathbf{x}_{k+1}$.

Plutôt que d'imposer de calculer B_{k+1} comme la matrice Hessienne au point $\mathbf{x}_{k+1}$, la hessienne approchée à l'itération k est mise à jour en ajoutant deux matrices.

$$\mathrm{B}_{k+1} = \mathrm{B}_k + \mathrm{U}_k + \mathrm{V}_k$$

U_k et V_k sont des matrices symétriques de rang 1 mais ont des bases différentes. Une matrice est symétrique de rang 1 si et seulement si elle peut s'écrire sous la forme $c \cdot a \cdot a^{\mathrm{T}}$, où a est une matrice colonne et c un scalaire.

De manière équivalente, U_k et V_k produisent une matrice de mise à jour de rang 2 qui est robuste vis-à-vis des problèmes d'échelle, qui pénalisent souvent les méthodes de gradient (comme la méthode de Broyden (en), l'analogue multidimensionel de la méthode de la sécante). Les conditions imposées pour la mise à jour sont:

$$\mathrm{B}_{k+1}(\mathbf{x}_{k+1}-\mathbf{x}_k)=\nabla f(\mathbf{x}_{k+1})-\nabla f(\mathbf{x}_k).$$

Algorithme

À partir d'une valeur initiale $\mathbf{x}_0$ et une matrice Hessienne approchée B_0 les itérations suivantes sont répétées jusqu'à ce que $\mathbf{x}$ converge vers la solution.

1. Trouver $\mathbf{p}_k$ en résolvant : $\mathrm{B}_k\mathbf{p}_k=-\nabla f(\mathbf{x}_k)$.
2. Effectuer une recherche linéaire pour trouver le pas optimal α_k dans la direction trouvée dans la première partie, et ensuite mettre à jour $\mathbf{x}_{k+1}=\mathbf{x}_k+\alpha_k\mathbf{p}_k=\mathbf{x}_k+\mathbf{s}_k$.
3. $\mathbf{y}_k=\nabla f(\mathbf{x}_{k+1})-\nabla f(\mathbf{x}_k)$.
4. $\mathrm{B}_{k+1}=\mathrm{B}_k+(\mathbf{y}_k\mathbf{y}_k^\top)/(\mathbf{y}_k^\top\mathbf{s}_k)-(\mathrm{B}_k\mathbf{s}_k\mathbf{s}_k^\top\mathrm{B}_k)/(\mathbf{s}_k^\top\mathrm{B}_k\mathbf{s}_k)$.

La fonction $f(\mathbf{x})$ est la fonction à minimiser. La convergence peut être testée en calculant la norme du gradient, $|\nabla f(\mathbf{x}_k)|$. En pratique, B_0 peut être initialisé avec $\mathrm{B}_0 = \mathrm{I}$, et la première itération sera alors équivalente à celle de l'algorithme du gradient, mais les autres itérations le raffineront de plus en plus grâce à B, l'approximation de la Hessienne.

On peut calculer l'intervalle de confiance de la solution à partir de l'inverse de la matrice Hessienne finale.

Bibliographie

- C. G. Broyden, « The Convergence of a Class of Double-rank Minimization Algorithms », *Journal of the Institute of Mathematics and Its Applications*, vol. 6, 1970, p. 76-90.
- R. Fletcher, « A New Approach to Variable Metric Algorithms », *Computer Journal*, vol. 13, 1970, p. 317-322.
- D. Goldfarb, « A Family of Variable Metric Updates Derived by Variational Means », *Mathematics of Computation*, vol. 24, 1970, p. 23-26.
- D. F. Shanno, « Conditioning of Quasi-Newton Methods for Function Minimization », *Mathematics of Computation*, vol. 24, 1970, p. 647-656.
- Mordecai Avriel, *Nonlinear Programming: Analysis and Methods*, Dover Publishing, 2003 (ISBN 0-486-43227-0).

Voir aussi

- Méthode de Newton

Références

- **(en)** Cet article est partiellement ou en totalité issu de l'article de Wikipédia en anglais intitulé « BFGS method (https://en.wikipedia.org/wiki/BFGS_method?oldid=116346046) » (voir la liste des auteurs (https://en.wikipedia.org/wiki/BFGS_method?action=history)).

Ce document provient de « https://fr.wikipedia.org/w/index.php?title=BFGS&oldid=135232842 ».

La dernière modification de cette page a été faite le 8 mars 2017 à 17:28.

Droit d'auteur : les textes sont disponibles sous licence Creative Commons attribution, partage dans les mêmes conditions ; d'autres conditions peuvent s'appliquer. Voyez les conditions d'utilisation pour plus de détails, ainsi que les crédits graphiques. En cas de réutilisation des textes de cette page, voyez comment citer les auteurs et mentionner la licence.
Wikipedia® est une marque déposée de la Wikimedia Foundation, Inc., organisation de bienfaisance régie par le paragraphe 501(c)(3) du code fiscal des États-Unis.