⊖ Besoin de calculer la matrice hessienne $H_f(x_k)$ en chaque itéré (long/coûteux + pas nécessairement de garanties qu'elle existe) puis son inverse $H_f(x_k)^{-1}$ (long/coûteux + instable numériquement si mal conditionnée)

⊖ La méthode est très sensible au choix du point de départ $x_0 \in \mathbb{R}^n$ → elle peut converger vers un point selle ou diverger. Ses garanties de convergence sont seulement locales

Améliorations potentielles

* Commencer par une descente de gradient (warm start) pour se rapprocher de la solution, puis finir avec la méthode de Newton
* Calculer la nouvelle direction d_k comme solution du système linéaire $H_f(x_k)d_k = -\nabla f(x_k)$ (via une factorisation de type Choleski) plutôt que d'inverser la hessienne pour avoir $d_k = -H_f(x_k)^{-1}\nabla f(x_k)$
* Remplacer le calcul de la hessienne $H_f(x_k)$ par une approximation B_k → utilisation d'une méthode de quasi Newton

Méthode de quasi Newton

À supposer que x_{k+1} et x_k sont assez proches, on a $\nabla f(x_{k+1}) \simeq \nabla f(x_k) + H_f(x_k)(x_{k+1}-x_k)$ (DL$_1$ de ∇f en x_k)

→ la méthode de quasi Newton cherche à construire une matrice B_{k+1} telle que $\nabla f(x_{k+1}) - \nabla f(x_k) \simeq B_{k+1}(x_{k+1}-x_k)$ de manière à approximer au mieux $H_f(x_k)$ par B_{k+1}, et que son inverse $H_{k+1} = B_{k+1}^{-1}$ approxime au mieux l'inverse de la hessienne $H_f(x_k)^{-1}$

Schéma d'une méthode de quasi-Newton

Initialisation : $x_0 \in \mathbb{R}^n$, $B_0 = I_d$ et $H_0 = B_0^{-1}$

Boucle : Tant que critère d'arrêt non satisfait

→ $d_k = -H_k \nabla f(x_k)$ (pour $k=0$ → descente de gradient)

→ η_k le pas défini par recherche linéaire (méthode d'Armijo)

→ $x_{k+1} = x_k + \eta_k d_k$

→ mise à jour de B_k / H_k

→ $\delta x_k = x_{k+1} - x_k$ et $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$

→ méthode BGFS (Broyden-Fletcher-Goldfarb-Shanno)

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T \delta x_k} - \frac{B_k \delta x_k (B_k \delta x_k)^T}{\delta x_k^T B_k \delta x_k}$$

$$H_{k+1} = \left(I_d - \frac{\delta x_k y_k^T}{y_k^T \delta x_k}\right) H_k \left(I_d - \frac{y_k \delta x_k^T}{y_k^T \delta x_k}\right) + \frac{\delta x_k \delta x_k^T}{y_k^T \delta x_k}$$