Université de M'sila
Faculté des Mathématiques de l'Informatique
Département des mathématiques
Licence 3ième année/ Optimisation
Année universitaire: 2020/2021

Méthodes de descentes

Une méthode de descente pour résoudre un problème d'optimisation $\min_{x\in\mathbb{R}^n} f(x)$ consiste à engendrer une suite de la forme

$$x_{n+1} = x_n + \rho_n d_n$$

telle que en assurant que la valeur de f décroit à chaque itération, c'est à dire:

$$f(x_{n+1}) < f(x_n), n = 0, 1, ...$$

Noter que:
1) d_n est une direction de descente de f en point x_n, (il suffit de vérifier $\nabla f(x_n)^\intercal d_n < 0$).
2) ρ_n est le pas de descente, qui prend trois formes:

$$\begin{cases} \textbf{Pas fixe: } \text{(toujours fixé)} \\ \textbf{Pas optimal}\text{: On résout à chaque étape le problème } \min_{\rho>0} f(x_n + \rho d_n) \end{cases}$$

3) **Critère d'arrêt**: Quand on commence les itérations de l'algorithme ci-dessus, on a besoin d'une condition pour arrêter les itérations de cet algorithme. Il existe plusieurs critères, par exemple:

$$\begin{cases} 1) \ \|\nabla f(x_{n+1})\| < \varepsilon \\ 2) \ \|x_{n+1} - x_n\| < \varepsilon \\ 3) \ |f(x_{n+1}) - f(x_n)| < \varepsilon \end{cases}$$

Chaque critère a des avantages et des inconvénients.

Algorithme de base:
- Soit $x_0 \in \mathbb{R}^n$. tel que $\nabla f(x_0) \neq 0$. Posons $n = 0$.
- Choisir d_k tel que $\nabla f(x_n)^\intercal d_n < 0$.
- Choisir (Recherche linéaire) $\rho_n > 0$
- Poser $x_{n+1} = x_n + \rho_n d_n$
- Critère d'arrêt.

Méthode du gradient: est la méthode la plus simple, juste on remarque que

$$d_n = -\nabla f(x_n) \neq 0$$

est bien une direction de descente de f en point x_n. L'algorithme du gradient est donné par

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

<u>Méthode de Newton:</u> Dans cette méthode, on prend la direction suivante:

$$d_n = -\left(\nabla^2 f(x_n)\right)^{-1} \nabla f(x_n),$$

avec des condistions sur f, d_n peut être une direction de descente (par exemple si $\left(\nabla^2 f(x_n)\right)^{-1}$ est définie positive). L'algorithme du Newton est donné par

$$x_{n+1} = x_n - \rho_n \left(\nabla^2 f(x_n)\right)^{-1} \nabla f(x_n).$$

<u>Méthode du gradient conjugué:</u>
Cett méthode a été introduite principalement pour résoudre le système linéaire $Ax = b$ où la matrice A est symétrique définie positive. On peut l'utiliser également pour les problèmes d'optimisation quadratiques. Elle converge en un nombre au plus égal à la dimension de la matrice A. On pose tout d'abord

$$d_0 = -g_0 = -\nabla f(x_0),$$

puis on essaie de trouver d_1 qui est conjuguée à d_0 c'est à dire

$$d_0^{\intercal} A d_1 = 0.$$

On suppose que

$$d_1 = -g_1 + \beta_0 d_0.$$

Après un simple calcul, on trouve:

$$\beta_0 = \frac{g_1^{\intercal} A d_0}{d_0^{\intercal} A d_0}.$$

Puis on obtient

$$x_2 = x_1 + \rho_1 d_1.$$

On fait la même chose pour calculer les autres directions de descentes

$$d_{k+1} = -g_{k+1} + \beta_k d_k$$

en calculant β_k sachant que $d_k^{\intercal} A d_{k+1} = 0$.