

Projet 15 : Descentes de gradient et applications à l'apprentissage profond

Ce projet est constitué de deux parties.

1. On s'intéressera d'abord aux algorithmes de descentes de gradient, dont le but est d'approcher un minimum d'une fonction différentiable de plusieurs variables $E : \mathbb{R}^N \rightarrow \mathbb{R}$, typiquement une fonction d'erreur. Le principe de ces algorithmes est le suivant :
 - (a) On part d'un point $x_0 \in \mathbb{R}^N$.
 - (b) On calcule $-\text{grad}_{x_0} E$, qui est le vecteur indiquant la *plus forte descente* au point x_0 .
 - (c) Pour un certain pas $\alpha_0 \in]0, +\infty[$, on pose $x_1 = x_0 - \alpha_0 \cdot \text{grad}_{x_0} E$.
 - (d) On réitère les étapes (b) et (c) jusqu'à l'obtention d'un x_n vérifiant un critère d'arrêt, par exemple $\|\text{grad}_{x_n} E\| < \varepsilon$, où $\varepsilon > 0$ est un seuil préalablement fixé.
2. On s'intéressera ensuite à un problème d'apprentissage profond, à savoir trouver une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R} \dots$
 - (a) ... dont on connaît les valeurs $y_i \in \mathbb{R}$ en certain points $X_i \in \mathbb{R}^n$,
 - (b) ... qui est de la forme $f(X) = F(w_1, w_2, \dots, w_N; X)$ où la fonction $F : \mathbb{R}^N \times \mathbb{R}^n \rightarrow \mathbb{R}$ est un *réseau de neurones* et où on appelle les paramètres $(w_1, w_2, \dots, w_N) \in \mathbb{R}^N$ les *poids synaptiques* de f ,
 - (c) ... qui minimise l'erreur globale quadratique

$$E(f) = E(w_1, w_2, \dots, w_N) = \sum_{i \in I} (f(X_i) - y_i)^2.$$

L'objectif du projet est la compréhension des objets et problèmes décrits ci-dessus, ainsi que l'implémentation d'algorithmes résolvant les problèmes.

Les ouvertures possibles sont nombreuses (améliorations des algorithmes de descente de gradient, architectures possibles pour les réseaux de neurones, création d'IA pour un problème précis, etc.).

Mots-clefs :

calcul différentiel, descente de gradient, apprentissage profond, perceptron.

Références possibles :

- [KW19] Mykel J KOCHENDERFER et Tim A WHEELER : *Algorithms for optimization*. Mit Press, 2019.
- [Wik25a] WIKIPÉDIA : Algorithme du gradient — wikipédia, l'encyclopédie libre. http://fr.wikipedia.org/w/index.php?title=Algorithme_du_gradient&oldid=230925644, 2025.
- [Wik25b] WIKIPÉDIA : Réseau de neurones artificiels — wikipédia, l'encyclopédie libre. http://fr.wikipedia.org/w/index.php?title=R%C3%A9seau_de_neurones_artificiels&oldid=230882445, 2025.