

TP - Descente de Gradient

December 23, 2022

L'objectif du TP est de se familiariser avec les éléments clés de la descente de gradient. Rédiger un rapport ainsi que le code couvrant les points suivants.

Partie 1 - Descente de gradient sur des fonctions en dimension 2

- Étudier, comprendre et illustrer les rôles des paramètres dans la descente de gradient (point initial, learning rate, momentum,..)
- Illustrer les pièges d'une descente de gradients: ravines, plateaux, minimum locaux,..
- Implémenter la descente simple, Momentum, Nesterov et Adam. Exemple de fonctions simples:

$$f(x, y) = x^2 + \gamma y^2 \text{ avec } \gamma \text{ un entier, } g(x, y) = 1 - \exp(-10x^2 - y^2)$$

$$h(x, y) = x^2y - 2xy^3 + 3xy + 4, \quad \text{etc}$$

- Essayer les fonctions classiques de test: Beale, Himmelblau, Rosenbrock,.. Quels sont les difficultés rencontrées?
https://en.wikipedia.org/wiki/Test_functions_for_optimization

$$\text{Rosenbrock}(x, y) = (1 - x)^2 + 100(y - x^2)^2$$

$$\text{Beale}(x, y) = (1.5 - x + xy)^2 + (2.25 - x + xy^2)^2 + (2.625 - x + xy^3)^2$$

$$\text{Booth}(x, y) = (x + 2y - 7)^2 + (2x + y - 5)^2$$

etc.

Partie 2 - Descente de gradient et régression linéaire

- Se choisir (ou générer) un échantillon d'apprentissage $\mathcal{A} = (x_i, y_i)$.
- Estimer les paramètres d'un modèle linéaire par Brute force (expliquer/illustrer les inconvénients), analytiquement, et par descente de gradient.