Partie 3: Methode du gradient stochastique

Florentin Goyens

12 octobre 2022

Intro

- Learning is an optimization problem over the training set, in the hope that it generalizes to unseen data.

The finite sum minimization problem

$$\text{minimize}_{x \in \mathbb{R}^d} \sum_{i=1}^{n} (h(x_i) - y_i)^2 = \sum_{i=1}^{n} f_i(x)$$

$$F(x) = \sum_{i=1}^{n} f_i(x)$$

$$\nabla F(x) = \sum_{i=1}^{n} \nabla f_i(x)$$

Title

$$\text{minimize}_{x \in \mathbb{R}^d} \ F(x) = \sum_{i=1}^{n} f_i(x)$$

Gradient descent method:

$$x_{k+1} = x_k - \gamma \nabla F(x_k)$$

Stochastic gradient method:

$$x_{k+1} = x_k - \gamma \nabla f_i(x_k)$$