

STOCHASTIC GRADIENT DESCENT

flatten() : pour aplatis une liste de listes. C'est-à-dire fusionner toutes les listes de la liste en une seule. Ex: [[1,2,3] , [4,5,6]] devient [1,2,3,4,5,6]

Gradient descent (cas de la régression linéaire) :

$$f_{\theta}(x_i) = \theta_1 + \theta_2 x_i$$

Define the loss function for the simple linear least squares regression as follows:

$$J(\theta) = \frac{1}{2} \sum_{i=1}^m (f_{\theta}(x_i) - y_i)^2$$

$$J(\theta) = \sum_{i=1}^m J_i(\theta)$$

$$\theta^{(t+1)} = \theta^{(t)} - \eta \nabla_{\theta} J(\theta^{(t)}) = \theta^{(t)} - \eta \sum_{i=1}^m \nabla_{\theta} J_i(\theta^{(t)})$$

Batch gradient descent : faire évoluer $\theta^{(t)}$ en utilisant toutes les données en même temps dans le calcul de la fonction d'erreur. Le batch gradient descent converge vers le inimum local si la fonction d'erreur est convexe, sinon vers le minimum local.