

Machine Learning

Gradient de l'erreur

Gradient et mise à jour

$$y = a.x + b$$

$$E_{\Omega} = \frac{1}{2n} \sum_{i=[1..n]} (\hat{y}_i - y_i)^2$$

$$E_{\Omega} = \frac{1}{2n} \sum_{i=[1..n]} (\hat{y}_i - (a.x_i + b))^2$$

...

$$\frac{\partial E_{\Omega}}{\partial a} = \frac{1}{n} \sum_{i=[1..n]} (a.x_i + b - \hat{y}_i).x_i$$

$$\frac{\partial E_{\Omega}}{\partial b} = \frac{1}{n} \sum_{i=[1..n]} (a.x_i + b - \hat{y}_i)$$

$$U^2' = 2U' * U$$

M.A.J :

$$a \leftarrow a - \gamma \cdot \frac{\partial E_{\Omega}}{\partial a}$$

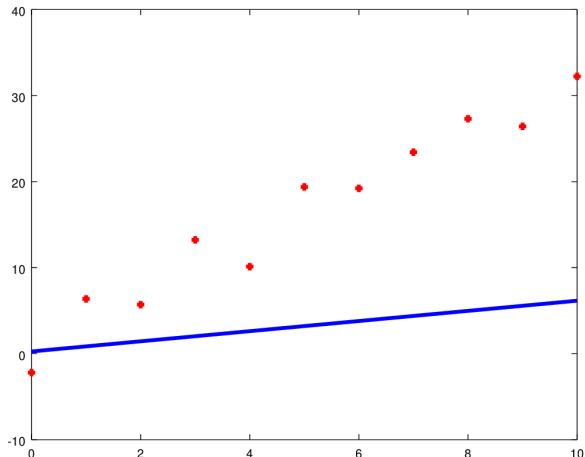
$$b \leftarrow b - \gamma \cdot \frac{\partial E_{\Omega}}{\partial b}$$

où $1 > \gamma > 0$ (learning rate)

Gradient et mise à jour

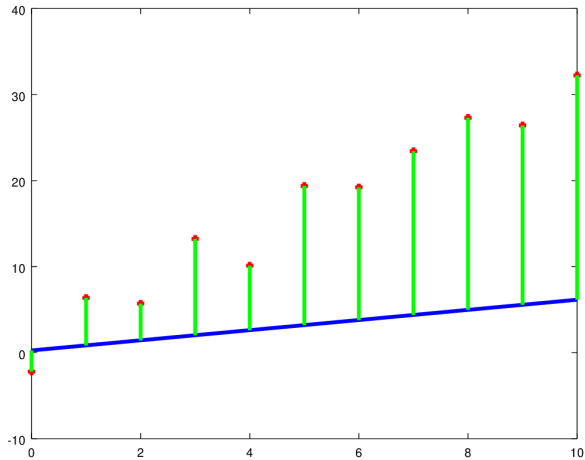
initialisation au hasard ($\gamma = 0.01$)

- $a = 0.58$ ($\hat{a} = 3.0$)
- $b = 0.25$ ($\hat{b} = 0.5$)



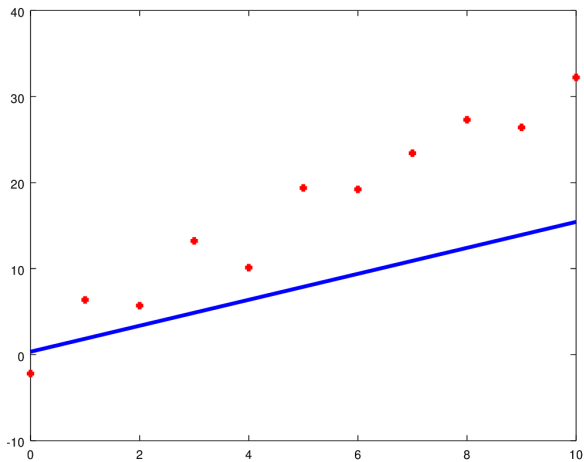
Gradient et mise à jour

- $a = 0.58$ ($\hat{a} = 3.0$)
- $b = 0.25$ ($\hat{b} = 0.5$)



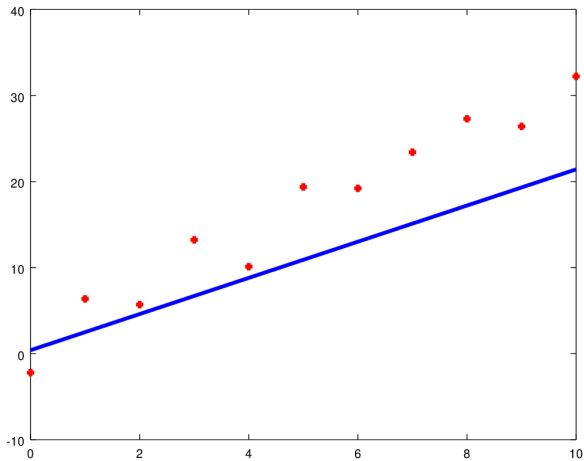
Gradient et mise à jour

- $a = 1.50$ ($\hat{a} = 3.0$)
- $b = 0.35$ ($\hat{b} = 0.5$)



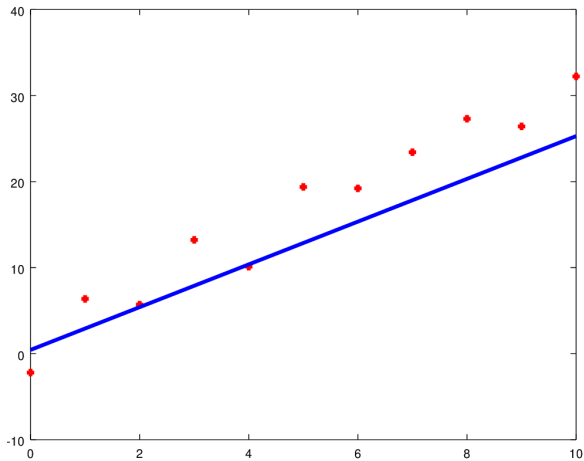
Gradient et mise à jour

- $a = 2.10$ ($\hat{a} = 3.0$)
- $b = 0.40$ ($\hat{b} = 0.5$)



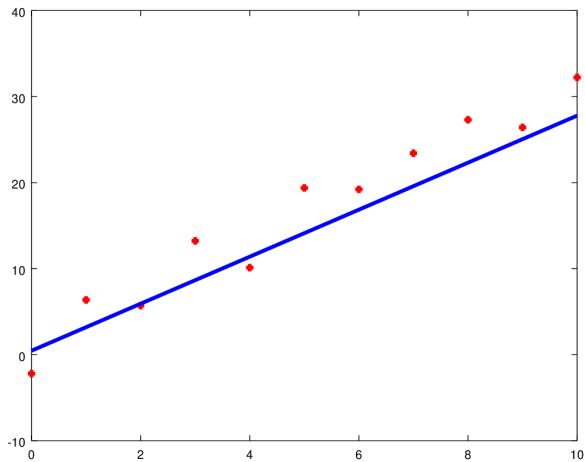
Gradient et mise à jour

- $a = 2.48$ ($\hat{a} = 3.0$)
- $b = 0.43$ ($\hat{b} = 0.5$)



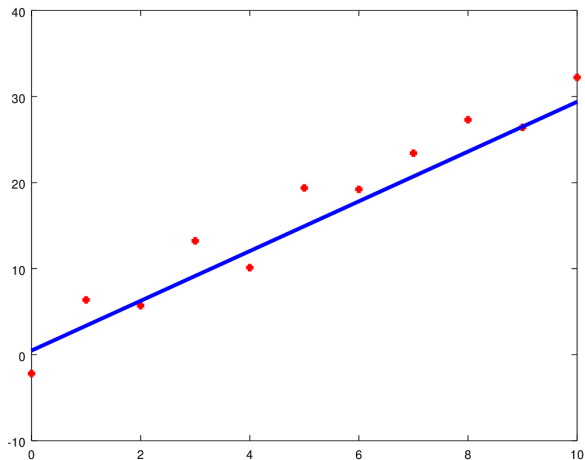
Gradient et mise à jour

- $a = 2.73$ ($\hat{a} = 3.0$)
- $b = 0.46$ ($\hat{b} = 0.5$)



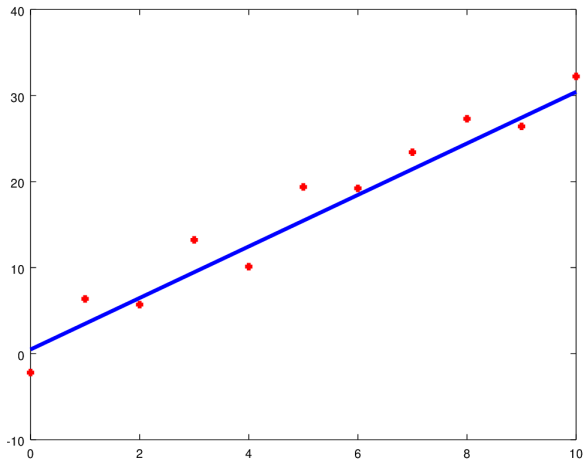
Gradient et mise à jour

- $a = 2.89$ ($\hat{a} = 3.0$)
- $b = 0.47$ ($\hat{b} = 0.5$)



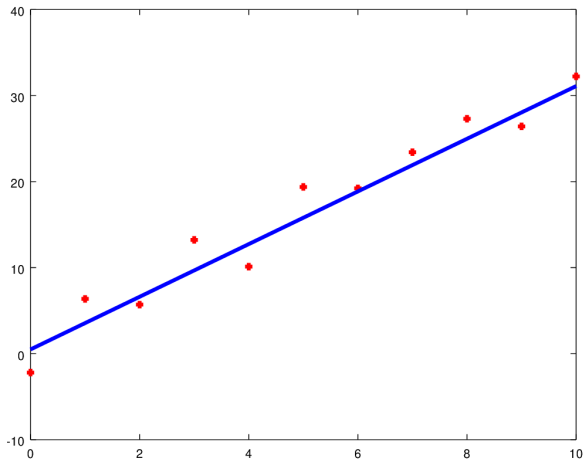
Gradient et mise à jour

- $a = 2.99$ ($\hat{a} = 3.0$)
- $b = 0.48$ ($\hat{b} = 0.5$)



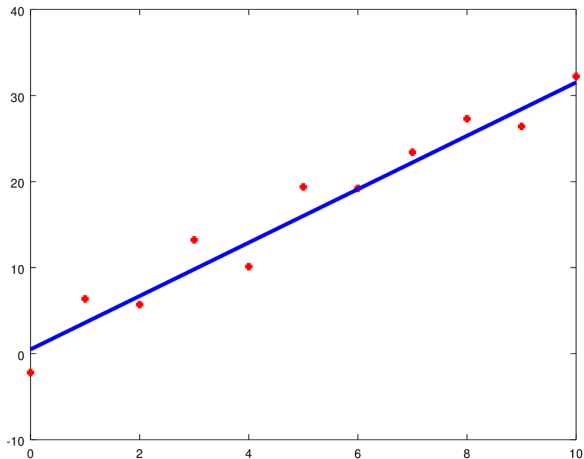
Gradient et mise à jour

- $a = 3.06$ ($\hat{a} = 3.0$)
- $b = 0.49$ ($\hat{b} = 0.5$)



Gradient et mise à jour

- $a = 3.10$ ($\hat{a} = 3.0$)
- $b = 0.49$ ($\hat{b} = 0.5$)



Gradient et mise à jour

- $a = 3.13$ ($\hat{a} = 3.0$)
- $b = 0.50$ ($\hat{b} = 0.5$)

