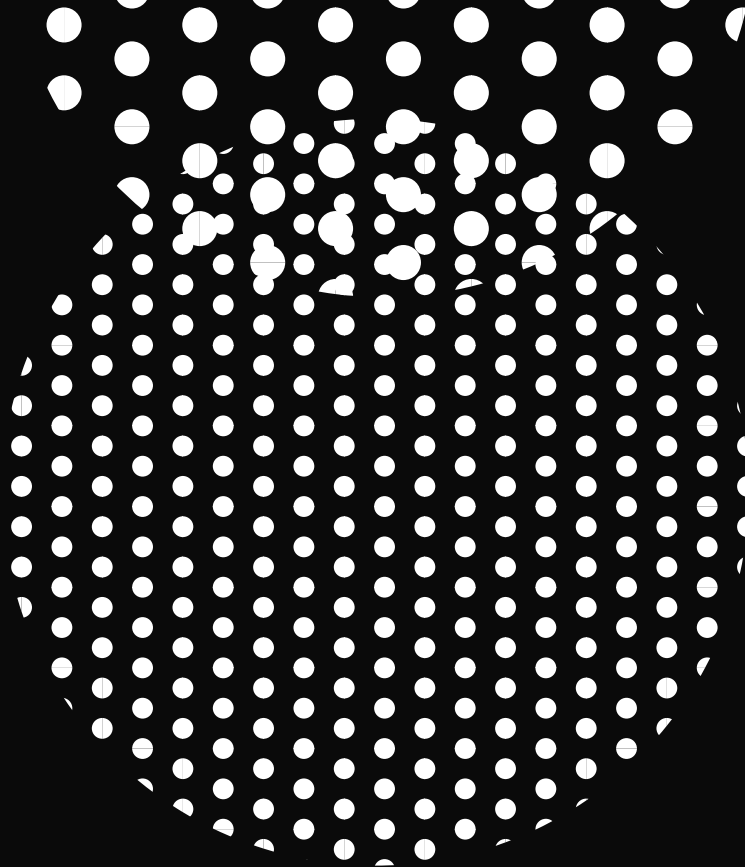




Exercice 4.1

$$X \in \mathbb{R}^{N \times D}$$

N : nbre d'exemples et D : nbre de données

$$X = \begin{bmatrix} x_1^{(1)} \\ \vdots \\ x_i^{(1)} \\ \vdots \\ x_i^{(N)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_D^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(N)} & x_2^{(N)} & \dots & x_D^{(N)} \end{bmatrix}$$

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad \vec{w} = [b, w_1, \dots, w_D]$$

$y_i \in \{-1, 1\}$

$$\vec{w} \in \mathbb{R}^D$$

$$y_{\text{real}}^{(n)} = y_n$$

$$f_{\vec{w}}(x^{(n)}) = \underbrace{\vec{w} \cdot x^{(n)} + b}_{= \vec{w} \cdot \vec{x}^{(n)}}$$

$$J(\vec{w}, X, Y) = \frac{1}{N} \sum_{n=1}^N \max(0, -(\vec{w} \cdot \vec{x}^{(n)}) y_n) = -\frac{1}{N} \sum_{n \in Y} (\vec{w} \cdot \vec{x}^{(n)}) y_n$$

Min à jouer des params, Descente de gradient

$\eta_{>0}$ taux d'apprentissage $\vec{w} \leftarrow \vec{w} - \eta \vec{\nabla}_{\vec{w}} J(\vec{w}, X, Y)$