

CORRECTION DES TRAVAUX DIRIGÉS DE
MACHINE LEARNING
CYCLE PLURIDISCIPLINAIRE D'ÉTUDES SUPÉRIEURES
UNIVERSITÉ PARIS SCIENCES ET LETTRES

Joon Kwon

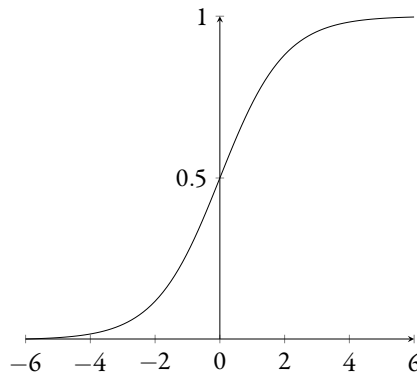
vendredi 27 mars 2020



1) Le cours donne :

$$(\hat{w}, \hat{b}) = \arg \min_{(w, b) \in \mathbb{R}^d \times \mathbb{R}} \left\{ \frac{1}{n} \sum_{i=1}^n \log(1 + e^{-y_i(\langle w, x_i \rangle + b)}) \right\}.$$

2) Tracé de la fonction :



3) Soit $z \in \mathbb{R}$ et $y \in]0, 1[$.

$$\begin{aligned} y = \frac{1}{1 + e^{-z}} &\iff e^{-z} = \frac{1}{y} - 1 \\ &\iff e^z = \frac{y}{1 - y} \\ &\iff z = \log \frac{y}{1 - y}. \end{aligned}$$

ϕ_{sig} est donc bijective de $\mathbb{R} \rightarrow]0, 1[$, et

$$\forall y \in]0, 1[, \quad \phi_{\text{sig}}^{-1}(y) = \log \frac{y}{1 - y}.$$

4) De la définition de (\hat{w}, \hat{b}) il découle que

$$g_{\hat{w}, \hat{b}} = \arg \min_{g \in \mathcal{L}_d} \left\{ \frac{1}{n} \sum_{i=1}^n \log(1 + e^{-y_i g_{w,b}(x_i)}) \right\}.$$

Dans l'expression ci-dessus, on peut tout à fait remplacer $g_{w,b}$ par $\phi_{\text{sig}}^{-1} \circ \phi_{\text{sig}} \circ g_{w,b}$. Cela donne :

$$g_{\hat{w}, \hat{b}} = \arg \min_{g \in \mathcal{L}_d} \left\{ \frac{1}{n} \sum_{i=1}^n \log(1 + e^{-y_i \phi_{\text{sig}}^{-1}(\phi_{\text{sig}} \circ g_{w,b})(x_i)}) \right\}.$$

En utilisant la définition de \mathcal{F}_0 , on peut écrire :

$$\hat{f}_0 = \phi_{\text{sig}} \circ g_{\hat{w}, \hat{b}} = \arg \min_{f \in \mathcal{F}_0} \left\{ \frac{1}{n} \sum_{i=1}^n \log(1 + e^{-y_i (\phi_{\text{sig}}^{-1} \circ f)(x_i)}) \right\}.$$

On note $y_{0,i} = \frac{y_i + 1}{2}$ l'étiquette de l'exemple $i \in [n]$ dans le problème auxiliaire. En utilisant l'expression de ϕ_{sig}^{-1} , on peut écrire :

$$\hat{f}_0 = \arg \min_{f \in \mathcal{F}_0} \left\{ \frac{1}{n} \sum_{i=1}^n \log \left(1 + \exp \left(-(2y_{0,i} - 1) \log \left(\frac{f(x_i)}{1 - f(x_i)} \right) \right) \right) \right\}$$

Ce qu'on peut réécrire comme une ERM :

$$\hat{f}_0 = \arg \min_{f \in \mathcal{F}_0} \left\{ \frac{1}{n} \sum_{i=1}^n \ell_0(y_{0,i}, f(x_i)) \right\}$$

où on a posé pour fonction de perte :

$$\ell_0(y, y') = \log \left(1 + \exp \left(-(2y - 1) \log \left(\frac{y'}{1 - y'} \right) \right) \right)$$

qu'on peut simplifier en :

$$\ell_0(y, y') = \log \left(1 + \left(\frac{1 - y'}{y'} \right)^{2y-1} \right).$$

5) On vérifie facilement que $\text{sign} \circ \phi_{\text{sig}} = \text{sign}$. On peut donc écrire :

$$\text{sign} \circ \hat{f}_0 = \text{sign} \circ \phi_{\text{sig}} \circ g_{\hat{w}, \hat{b}} = \text{sign} \circ g_{\hat{w}, \hat{b}} = \hat{f}.$$

