

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

Joon Kwon

mercredi 13 avril 2022



**EXERCICE 1.** —

1)

$$\begin{aligned}\langle w, v \rangle + b &= \langle w, x - (\langle w, x \rangle + b)w \rangle + b \\ &= \langle w, x \rangle - (\langle w, x \rangle + b) \|w\|_2^2 + b \\ &= 0,\end{aligned}$$

car  $\|w\|_2 = 1$  par hypothèse.

2) On a

$$\begin{aligned}\min_{\substack{x' \in \mathbb{R}^d \\ \langle w, x' \rangle + b}} \|x - x'\|_2 &\leq \|x - v\|_2 = \sqrt{\langle x - v, x - v \rangle} \\ &= \sqrt{\langle (\langle w, x \rangle + b)w, (\langle w, x \rangle + b)w \rangle} \\ &= \sqrt{(\langle w, x \rangle + b)^2 \|w\|_2^2} \\ &= |\langle w, x \rangle + b|.\end{aligned}$$

3)

$$\begin{aligned}\|x - u\|_2^2 &= \|x - v + v - u\|_2^2 \\ &= \|x - v\|_2^2 + 2 \langle x - v, v - u \rangle + \|v - u\|_2^2 \\ &\geq \|x - v\|_2^2 + 2 \langle x - v, v - u \rangle.\end{aligned}$$

Or, on a

$$\begin{aligned}\langle x - v, v - u \rangle &= (\langle w, x \rangle + b) \langle w, v - u \rangle = (\langle w, x \rangle + b)(\langle w, v \rangle - \langle w, u \rangle) \\ &= (\langle w, x \rangle + b)(-b + b) = 0,\end{aligned}$$

d'où le résultat.

4) On a donc

$$\min_{\substack{x' \in \mathbb{R}^d \\ \langle w, x' \rangle + b}} \|x - x'\|_2 = |\langle w, x \rangle + b|.$$

## EXERCICE 2. —

1) Du fait que  $(\hat{w}, \hat{b})$  est une solution de  $(*)$ , il découle que  $\mu > 0$ . On a donc :

$$\begin{aligned}h_{\tilde{w}, \tilde{b}}(x) &= \text{sign}(\langle \tilde{w}, x \rangle + \tilde{b}) = \text{sign}(\langle \mu \hat{w}, x \rangle + \mu \hat{b}) \\ &= \text{sign}(\mu(\langle \hat{w}, x \rangle + \hat{b})) = \text{sign}(\langle \hat{w}, x \rangle + \hat{b}) = h_{\hat{w}, \hat{b}}.\end{aligned}$$

2)  $(\tilde{w}, \tilde{b})$  satisfait les contraintes de  $(**)$  : en effet, pour tout  $i \in [n]$ , on a

$$\begin{aligned}y_i(\langle \tilde{w}, x_i \rangle + \tilde{b}) &= y_i(\langle \mu \hat{w}, x_i \rangle + \mu \hat{b}) = \mu y_i(\langle \hat{w}, x_i \rangle + \hat{b}) \\ &= \left( \min_{i' \in [n]} y_{i'}(\langle \hat{w}, x_{i'} \rangle + \hat{b}) \right)^{-1} \left( y_i(\langle \hat{w}, x_i \rangle + \hat{b}) \right) \\ &\geq 1.\end{aligned}$$

Montrons que  $\tilde{w}$  minimise la fonction objectif  $w \mapsto \|w\|_2$ . Soit  $(w, b) \in \mathbb{R}^d \times \mathbb{R}$  qui satisfait également les contraintes de  $(**)$  :

$$\forall i \in [n], \quad y_i(\langle w, x_i \rangle + b) \geq 1,$$

et montrons  $\|w\|_2 \geq \|\tilde{w}\|_2$ . On considère  $w_0 = w / \|w\|_2$  et  $b_0 = b / \|w\|_2$ . On a alors  $\|w_0\| = 1$  et pour tout  $i \in [n]$ ,

$$y_i(\langle w_0, x_i \rangle + b_0) = \frac{1}{\|w\|_2} y_i(\langle w, x_i \rangle + b) > 0.$$

Autrement dit,  $(w_0, b_0)$  satisfait les contraintes de (\*). Donc,  $(\hat{w}, \hat{b})$  étant par définition solution de (\*) on a :

$$\min_{i \in [n]} |\langle \hat{w}, x_i \rangle + \hat{b}| \geq \min_{i \in [n]} |\langle w_0, x_i \rangle + b_0|.$$

En remplaçant ci-dessus  $w_0$  et  $b_0$  par leurs définitions, on a :

$$\begin{aligned} \min_{i \in [n]} |\langle \hat{w}, x_i \rangle + \hat{b}| &\geq \min_{i \in [n]} \left| \left\langle \frac{w}{\|w\|_2}, x_i \right\rangle + \frac{b}{\|w\|_2} \right| = \frac{1}{\|w\|_2} \left( \min_{i \in [n]} |\langle w, x_i \rangle + b| \right) \\ &\geq \frac{1}{\|w\|_2}, \end{aligned}$$

où la dernière inégalité découle du fait que  $(w, b)$  satisfait par définition les contraintes de (\*\*) et que  $y_i \in \{-1, 1\}$ . On a donc

$$\|w\|_2 \geq \left( \min_{i \in [n]} |\langle \hat{w}, x_i \rangle + \hat{b}| \right)^{-1} = \mu = \mu \|\hat{w}\| = \|\tilde{w}\|,$$

D'où le résultat.

