Correction des travaux dirigés de Machine learning

Cycle pluridisciplinaire d'études supérieures
Université Paris sciences et lettres

Joon Kwon

mercredi 18 mai 2022

Exercice 1. — Soit w_* la projection orthogonale de $\hat{w}$ sur $\mathrm{Vect}_{i\in[n]}\,\psi(x_i)$. Il existe donc $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tels que

$$w_* = \sum_{i=1}^n \alpha_i \psi(x_i).$$

Et puisque $\hat{w} - w_* \in (\mathrm{Vect}_{i\in[n]}\,\psi(x_i))^\perp$, on a pour tout $i \in [n]$:

$$\langle \hat{w} - w_*, \psi(x_i)\rangle = 0,$$

autrement dit :

$$\langle \hat{w}, \psi(x_i)\rangle = \langle w_*, \psi(x_i)\rangle .$$

Ainsi, pour tout $i \in [n]$,

$$f_{w_*,\hat{b}}(\psi(x_i)) = \phi(\langle w_*, \psi(x_i)\rangle + \hat{b}) = \phi(\langle \hat{w}, \psi(x_i)\rangle + \hat{b}) = f_{\hat{w},\hat{b}}(\psi(x_i)).$$

$(\hat{w}, \hat{b})$ étant une solution par définition, on a :

$$\frac{1}{n}\sum_{i=1}^n \ell(y_i, f_{\hat{w},\hat{b}}(\psi(x_i))) + \frac{\lambda}{2}\left(\|\hat{w}\|^2 + \hat{b}^2\right) \leqslant \frac{1}{n}\sum_{i=1}^n \ell(y_i, f_{w_*,\hat{b}}(\psi(x_i))) + \frac{\lambda}{2}\left(\|w_*\|^2 + \hat{b}^2\right),$$

ce qui, en utilisant ce qui précède et le fait que $\lambda > 0$, se simplifie en :

$$\|\hat{w}\|^2 = \|w_*\|^2 .$$

Or w_* étant la projection orthogonale de $\hat{w}$, on a :

$$\|w_*\|^2 + \|\hat{w} - w_*\|^2 = \|\hat{w}\|^2 = \|w_*\|^2 .$$

Cela entraîne $\|\hat{w} - w_*\|^2 = 0$, autrement dit $\hat{w} = w_*$.

Exercice 2. — Pour $1 \leqslant i, k \leqslant m$,

$$\mathrm{G}_{ik} = \mathrm{K}(x_i, x_k) = \langle \psi(x_i), \psi(x_k) \rangle = \langle \psi(x_k), \psi(x_i) \rangle = \mathrm{K}(x_k, x_i) = \mathrm{G}_{ki}.$$

G est symmétrique. Soit $u \in \mathbb{R}^m$.

$$\begin{aligned} u^\top \mathrm{G} u &= \sum_{1 \leqslant i,k \leqslant m} u_i \mathrm{K}(x_i, x_k) u_k = \sum_{1 \leqslant i,k \leqslant m} u_i \langle \psi(x_i), \psi(x_k) \rangle u_k = \sum_{1 \leqslant i,k \leqslant m} \langle u_i \psi(x_i), \psi(x_k) u_k \rangle \\ &= \left\langle \sum_{i=1}^m u_i \psi(x_i), \sum_{k=1}^m u_k \psi(x_k) \right\rangle = \left\| \sum_{i=1}^m u_i \psi(x_i) \right\|^2 \geqslant 0. \end{aligned}$$

G est bien semi-définie positive.