Travaux dirigés de
Machine learning
Cycle pluridisciplinaire d'études supérieures
Université Paris sciences et lettres

mercredi 18 mai 2022

On rappelle le théorème qui donne l'existence et l'unicité de la projection orthogonale sur un sous-espace de dimension finie d'un espace préhibertien. Il sera utile dans le premier exercice.

Théorème. — *Projection orthogonale.* — *Soit* $(\mathrm{E}, \langle \cdot\,, \cdot \rangle)$ *un espace préhilbertien et* F *un sous-espace de dimension finie. Pour* $x \in \mathrm{E}$*, il existe un unique* $x_{\mathrm{F}} \in \mathrm{F}$ *tel que :*

$$x - x_{\mathrm{F}} \in \mathrm{F}^{\perp}.$$

De plus,

$$\|x\|^2 = \|x_{\mathrm{F}}\|^2 + \|x - x_{\mathrm{F}}\|^2 .$$

Exercice 1. — *Théorème de représentation.* — Soit $\mathscr{X}$, $\mathscr{Y}$ deux ensembles quelconques, et $\ell : \mathscr{Y} \times \mathscr{Y} \to \mathbb{R}$ une fonction. Soit également $(\tilde{\mathscr{X}}, \langle \cdot\,, \cdot \rangle)$ un espace préhilbertien, $\| \cdot \|$ la norme associée, et $\psi : \mathscr{X} \to \tilde{\mathscr{X}}$ une application. Soit $\phi : \mathbb{R} \to \mathscr{Y}$ et on note pour $(w, b) \in \tilde{\mathscr{X}} \times \mathbb{R}$:

$$\forall \tilde{x} \in \tilde{\mathscr{X}}, \quad f_{w,b}(\tilde{x}) = \phi(\langle w, \tilde{x} \rangle + b).$$

Soit $\lambda > 0$, $n \geqslant 1$ un entier et $(x_i, y_i)_{i\in[n]} \in \mathscr{S}(\mathscr{X}, \mathscr{Y})$ un échantillon. On considère le problème d'optimisation suivant.

$$\begin{aligned} \text{minimiser} \quad & \frac{1}{n}\sum_{i=1}^{n} \ell\left(y_i, f_{w,b}(\psi(x_i))\right) + \frac{\lambda}{2}\left(\|w\|^2 + b^2\right) \\ \text{sachant} \quad & (w, b) \in \tilde{\mathscr{X}} \times \mathbb{R}. \end{aligned} \tag{$*$}$$

Soit $(\hat{w}, \hat{b})$ une solution du problème d'optimisation ($*$).

1) Montrer qu'il existe $w_* \in \tilde{\mathscr{X}}$ et $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tels que :

$$w_* = \sum_{i=1}^{n} \alpha_i \psi(x_i) \qquad \text{et} \qquad \forall i \in [n],\ (f_{w_*,\hat{b}} \circ \psi)(x_i) = (f_{\hat{w},\hat{b}} \circ \psi)(x_i).$$

2) En déduire que $\hat{w} = w_*$.

Exercice 2 (*Caractérisation des noyaux*). — Soit $\mathscr{X}$ un ensemble quelconque.

1) On suppose dans cette question qu'on a $(\tilde{\mathscr{X}}, \langle\cdot\,,\,\cdot\rangle)$ espace préhilbertien, $\psi : \mathscr{X} \to \tilde{\mathscr{X}}$, et $\mathrm{K} : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ le noyau associé. Soit $x_1, \dots, x_m \in \mathscr{X}$. Montrer que la matrice :

$$\mathrm{G} = (\mathrm{K}(x_i, x_k))_{1\leqslant i,k\leqslant m}$$

est symétrique semi-définie positive.

On souhaite à présent établir la réciproque. Soit $\mathrm{K} : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ une application telle que pour tout entier $m \geqslant 1$ et tous $x_1, \dots, x_m \in \mathscr{X}$, la matrice $(\mathrm{K}(x_i, x_k))_{1\leqslant i,k\leqslant m}$ est symétrique semi-définie positive.

Le but de l'exercice est de montrer que K est un noyau, autrement dit qu'il existe une application de redescription à valeurs dans un espace préhilbertien dont le noyau associé est K.

Pour $x \in \mathscr{X}$, on note K_x la fonction $\mathrm{K}_x : x' \mapsto \mathrm{K}(x, x')$. Soit $\tilde{\mathscr{X}}$ le sous-espace vectoriel de $\mathbb{R}^{\mathscr{X}}$ engendré par les fonctions $(\mathrm{K}_x)_{x\in\mathscr{X}}$. Soit A l'ensemble des vecteurs $\alpha = (\alpha_x)_{x\in\mathscr{X}} \in \mathbb{R}^{\mathscr{X}}$ ayant un nombre fini de coefficients non-nuls. $\tilde{\mathscr{X}}$ peut donc s'écrire :

$$\tilde{\mathscr{X}} = \left\{ \sum_{x\in\mathscr{X}} \alpha_x \mathrm{K}_x \right\}_{\alpha\in\mathrm{A}},$$

où la somme est bien définie car seul un nombre fini de termes est non-nul.

Soit $\tilde{x}, \tilde{x}' \in \tilde{\mathscr{X}}$. Il existe $\alpha, \beta \in A$ tels que

$$\tilde{x} = \sum_{x \in \mathscr{X}} \alpha_x K_x \qquad \text{et} \qquad \tilde{x}' = \sum_{x \in \mathscr{X}} \beta_x K_x.$$

On pose alors :

$$\langle \tilde{x}, \tilde{x}' \rangle = \sum_{x, x' \in \mathscr{X}} \alpha_x \beta_{x'} K(x, x'),$$

où la somme a bien un sens car seul un nombre fini de termes est non-nul. L'objectif des questions ci-dessous est de montrer que $\langle \cdot , \cdot \rangle$ est bien un produit scalaire sur $\tilde{\mathscr{X}}$ et qu'il existe une application $\psi : \mathscr{X} \to \tilde{\mathscr{X}}$ telle que K est le noyau associé.

2) Montrer que la valeur de $\langle \tilde{x}, \tilde{x}' \rangle$ ne dépend pas du choix des coefficients $(\alpha_x)_{x \in \mathscr{X}}$ et $(\beta_x)_{x \in \mathscr{X}}$ (qui peuvent ne pas être uniques).
3) Montrer que $\langle \cdot , \cdot \rangle$ est bilinéaire et symétrique.
4) Soit $\tilde{x} \in \tilde{\mathscr{X}}$. Montrer que $\langle \tilde{x}, \tilde{x} \rangle \geqslant 0$.
5) Soit $\tilde{x} \in \tilde{\mathscr{X}}$ et $x \in \mathscr{X}$. Pour $\lambda \in \mathbb{R}$, en considérant la quantité

$$\langle \lambda \tilde{x} + \tilde{x}(x) K_x, \lambda \tilde{x} + \tilde{x}(x) K_x \rangle,$$

montrer que :

$$\tilde{x}(x)^4 \leqslant \tilde{x}(x)^2 \langle \tilde{x}, \tilde{x} \rangle K(x, x).$$

6) En déduire que si $\tilde{x} \in \tilde{\mathscr{X}}$ est tel que $\langle \tilde{x}, \tilde{x} \rangle = 0$, alors $\tilde{x} = 0$.
7) Montrer que K est le noyau associé à une application de redescription $\psi : \mathscr{X} \to \tilde{\mathscr{X}}$ à préciser.