Correction des travaux dirigés de
Machine learning
Cycle pluridisciplinaire d'études supérieures
Université Paris sciences et lettres

vendredi 28 avril 2023

Exercice 1. — Pour $x, x' \in \mathbb{R}^2$,

$$\mathrm{K}(x, x') = \langle x, x' \rangle = (x_1 x_1' + x_2 x_2')^2 = x_1^2 (x_2')^2 + 2 x_1 x_1' x_2 x_2' + x_2^2 (x_2')^2.$$

On pose $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ définie par

$$\psi : (x_1, x_2) \mapsto (x_1^2, \sqrt{2} x_1 x_2, x_2^2).$$

On a bien alors :

$$\forall x, x' \in \mathbb{R}^2, \quad \langle \psi(x), \psi(x') \rangle_{\mathbb{R}^3} = x_1^2 (x_2')^2 + 2 x_1 x_1' x_2 x_2' + x_2^2 (x_2')^2 = \mathrm{K}(x, x'),$$

K est le noyau associé à ψ.

Exercice 2. — *Noyau polynomial.* — On note $[d] = \{1, \dots, d\}$. Pour $x, x' \in \mathbb{R}^d$,

$$\begin{aligned}
\mathrm{K}(x, x') &= \langle x, x' \rangle^m \\
&= (x_1 x_1' + \cdots + x_d x_d') \times (x_1 x_1' + \cdots + x_d x_d') \times \cdots \times (x_1 x_1' + \cdots + x_d x_d') \\
&= \sum_{(j_1, \dots, j_m) \in [d]^m} \prod_{k=1}^{m} x_{j_k} x_{j_k}' = \sum_{(j_1, \dots, j_m) \in [d]^m} \left(\prod_{k=1}^{m} x_{j_k} \right) \left(\prod_{k=1}^{m} x_{j_k}' \right).
\end{aligned}$$

Soit $\tilde{\mathscr{X}} = \mathbb{R}^{([d]^m)}$ On considère $\psi : \mathbb{R}^d \to \tilde{\mathscr{X}}$ définie par

$$\psi : \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \longmapsto \left(\prod_{k=1}^{m} x_{j_k} \right)_{(j_1,\ldots,j_m)\in[d]^m}$$

On considère sur $\tilde{\mathscr{X}}$ le produit scalaire canonique : pour $u, v \in \tilde{\mathscr{X}}$,

$$\langle u, v \rangle_{\tilde{\mathscr{X}}} = \sum_{(j_1,\ldots,j_m)\in[d]^m} u_{(j_1,\ldots,j_m)} v_{(j_1,\ldots,j_m)}.$$

Alors, on a pour $x, x \in \mathbb{R}^d$:

$$\langle \psi(x), \psi(x') \rangle_{\tilde{\mathscr{X}}} = \sum_{(j_1,\ldots,j_m)\in[d]^m} \left(\prod_{k=1}^{m} x_{j_k} \right) \left(\prod_{k=1}^{m} x'_{j_k} \right) = \mathrm{K}(x, x').$$

K est donc bien un noyau.

Exercice 3. — *Noyau gaussien*

1) Soit $\mathrm{N} \geqslant 0$. En utilisant l'inégalité de Cauchy-Schwarz :

$$\sum_{m=0}^{\mathrm{N}} |u_m v_m| \leqslant \sqrt{\sum_{m=0}^{\mathrm{N}} u_m^2} \sqrt{\sum_{m=0}^{\mathrm{N}} v_m^2} \leqslant \sqrt{\sum_{m=0}^{+\infty} u_m^2} \sqrt{\sum_{m=0}^{+\infty} v_m^2}.$$

La série $\sum u_m v_m$ converge absolument, donc converge.

2) $\tilde{\mathscr{X}}$ est inclus par définition dans l'espace vectoriel $\mathbb{R}^{\mathbb{N}}$. Montrons que $\tilde{\mathscr{X}}$ est un sous-espace. Pour $u, v \in \tilde{\mathscr{X}}$ et $\lambda \in \mathbb{R}$, on a que la série $\sum(\lambda u_m)^2$ converge, donc $\lambda u \in \tilde{\mathscr{X}}$. Pour $m \geqslant 0$, on a :

$$(u_m + v_m)^2 = u_m^2 + 2u_m v_m + v_m^2.$$

On sait que les séries $\sum u_m^2$, $\sum v_m^2$ et $\sum u_m v_m$ convergent, donc $\sum(u_m + v_m)^2$ aussi. Donc : $u + v \in \tilde{\mathscr{X}}$. $\tilde{\mathscr{X}}$ est bien un sous-espace.

3) Soit $u, v, w \in \tilde{\mathscr{X}}$ et $\lambda \in \mathbb{R}$.
— *Bilinéaire*. — On a pour $\mathrm{N} \geqslant 0$,

$$\sum_{m=0}^{\mathrm{N}} (u_m + v_m) w_m = \sum_{m=0}^{\mathrm{N}} u_m w_m + \sum_{m=0}^{\mathrm{N}} v_m w_m.$$

et

$$\sum_{m=0}^{N} \lambda u_m v_m = \lambda \sum_{m=0}^{N} u_m v_m.$$

En passant à la limite quand $N \to +\infty$, on obtient :

$$\langle u+v, w\rangle_{\tilde{\mathscr{X}}} = \langle u, w\rangle_{\tilde{\mathscr{X}}} + \langle v, w\rangle_{\tilde{\mathscr{X}}} \quad \text{et} \quad \langle \lambda u, v\rangle_{\tilde{\mathscr{X}}} = \lambda \langle u, w\rangle_{\tilde{\mathscr{X}}}.$$

— *Symmétrique*. — Pour $N \geqslant 0$,

$$\sum_{m=0}^{N} u_m v_m = \sum_{m=0}^{N} v_m u_m,$$

et en passant à la limite quand $N \to +\infty$, on obtient $\langle u, v\rangle_{\tilde{\mathscr{X}}} \langle v, u\rangle_{\tilde{\mathscr{X}}}$.

— *Positive*. — Pour $N \geqslant 0$, on a

$$\sum_{m=0}^{N} u_m u_m \geqslant 0$$

et $\langle u, u\rangle_{\tilde{\mathscr{X}}} \geqslant 0$ en passant à la limite quand $N \to +\infty$.

— *Définie*. — Si $\langle u, u\rangle = 0$, on a pour tout $m \geqslant 0$:

$$0 \leqslant u_m^2 \leqslant \sum_{m'=0}^{N} u_{m'}^2,$$

ce qui entraîne $0 \leqslant u_m^2 \leqslant 0$ en passant à la limite quand $N \to +\infty$, autrement dit $u_m = 0$. Donc $u = 0$.

4) Pour $x, x' \in \mathbb{R}$,

$$\begin{aligned}
K(x, x') &= \exp\left(-\frac{(x-x')}{2}\right) = \exp\left(-\frac{x^2}{2} + xx' - \frac{(x')^2}{2}\right) \\
&= \exp\left(-\frac{x^2}{2}\right)\exp\left(-\frac{(x')^2}{2}\right)\sum_{m=0}^{+\infty}\frac{(xx')^m}{m!} \\
&= \sum_{m=0}^{+\infty}\left(\frac{e^{-x^2/2}x^m}{\sqrt{m!}}\right)\left(\frac{e^{-(x')^2/2}(x')^m}{\sqrt{m!}}\right).
\end{aligned}$$

On définit :

$$\forall x \in \mathbb{R}, \quad \psi(x) = \left(\frac{e^{-x^2/2}x^m}{\sqrt{m!}}\right)_{m\geqslant 0}.$$

Pour tout $x \in \mathbb{R}$, on a bien $\psi(x) \in \tilde{\mathscr{X}}$ car la série $\sum(\frac{x^m}{\sqrt{m!}})^2$ converge (de limite e^{x^2}). Et on a bien :

$$\forall x, x' \in \mathbb{R}, \quad \langle \psi(x), \psi(x') \rangle_{\tilde{\mathscr{X}}} = \mathrm{K}(x, x').$$