Correction de l'examen de Machine learning

Cycle pluridisciplinaire d'études supérieures
Université Paris sciences et lettres

Exercice 3 (*Détection de sous-mots*). —

1) On considère $\widetilde{\mathscr{X}} = \mathbb{R}^{\mathscr{X}_d}$ muni de son produit scalaire canonique et on définit ψ par :
$$\forall x \in \mathscr{X}_d, \quad \psi(x) = \left(\mathbb{1}_{\{v \prec x\}}\right)_{v \in \mathscr{X}_d},$$
où $\mathbb{1}_{\{v \prec x\}} = 1$ si $v \prec x$, et $\mathbb{1}_{\{v \prec x\}} = 0$ sinon. Alors, pour $x, x \in \mathscr{X}_d$, on a :
$$\begin{aligned}\langle \psi(x), \psi(x') \rangle &= \sum_{v \in \mathscr{X}_d} \mathbb{1}_{\{v \prec x\}} \mathbb{1}_{\{v \prec x'\}} = \sum_{v \in \mathscr{X}_d} \mathbb{1}_{\{v \prec x \text{ et } v \prec x'\}} \\ &= \operatorname{Card}\{v \in \mathscr{X}_d \mid v \prec x \text{ et } v \prec x'\} = \mathrm{K}(x, x').\end{aligned}$$

2) Soit $v \in \mathscr{X}_d$. On pose $\phi(u) = 2u - 1$ et on note e_v le vecteur de la base canonique de $\mathbb{R}^{\mathscr{X}_d}$ associé à v, autrement dit $e_v = (\mathbb{1}_{\{v=v'\}})_{v' \in \mathscr{X}_d}$. On note g_{e_v} l'application définie par $g_{e_v}(\tilde{x}) = \langle e_v, \tilde{x} \rangle$, pour $\tilde{x} \in \widetilde{\mathscr{X}}$. Alors, $\tilde{f}_v = \phi \circ g_{e_v}$ est bien un prédicteur linéaire sur $\mathscr{X}$. Et pour $x \in \widetilde{\mathscr{X}}$,
$$(\tilde{f}_v \circ \psi)(x) = \phi\left(\langle e_v, \psi(x) \rangle\right) = \phi\left(\mathbb{1}_{\{v \prec x\}}\right) = f_v(x).$$

Exercice 4 (*Caractérisation des noyaux*). —

1) Voir TD.

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

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

Supposons qu'il existe $\alpha' \in A$ tel que $\tilde{x}$ s'écrit aussi :

$$\tilde{x} = \sum_{x \in \mathscr{X}} \alpha'_x K_x.$$

Alors, on a :

$$\begin{aligned}\langle \tilde{x}, \tilde{x}' \rangle &= \sum_{x, x' \in \mathscr{X}} \alpha_x \beta_{x'} K(x, x') = \sum_{x' \in \mathscr{X}} \beta_{x'} \sum_{x \in \mathscr{X}} \alpha_x K(x, x') = \sum_{x' \in \mathscr{X}} \beta_{x'} \sum_{x \in \mathscr{X}} \alpha_x K_x(x') \\ &= \sum_{x' \in \mathscr{X}} \beta_{x'} \tilde{x}(x') = \sum_{x' \in \mathscr{X}} \beta_{x'} \sum_{x \in \mathscr{X}} \alpha'_x K_x(x') = \sum_{x, x' \in \mathscr{X}} \alpha'_x \beta_{x'} K(x, x').\end{aligned}$$

$\langle \tilde{x}, \tilde{x}' \rangle$ ne dépend donc pas du choix des coefficients $(\alpha_x)_{x \in \mathscr{X}}$. De même pour les coefficients $(\beta_x)_{x \in \mathscr{X}}$.

3) Soit $\tilde{x}, \tilde{x}', \tilde{x}'' \in \tilde{\mathscr{X}}$ s'écrivant :

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

Soit $\{x_1, \dots, x_n\}$ l'ensemble des $x \in \mathscr{X}$ tels que α_x ou β_x est non-nul. Par hypothèse, la matrice $(K(x_i, x_k))_{1 \leqslant i,k \leqslant n}$ est symétrique. On a alors

$$\begin{aligned}\langle \tilde{x}, \tilde{x}' \rangle &= \sum_{x, x' \in \mathscr{X}} \alpha_x \beta_{x'} K(x, x') = \sum_{1 \leqslant i,k \leqslant n} \alpha_{x_i} \beta_{x_k} K(x_i, x_k) \\ &= \sum_{1 \leqslant i,k \leqslant n} \beta_{x_k} \alpha_{x_i} K(x_k, x_i) = \langle \tilde{x}', \tilde{x} \rangle .\end{aligned}$$

$\langle \cdot , \cdot \rangle$ est donc symétrique.

Soit $\lambda \in \mathbb{R}$. On a :

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

On a également :

$$\begin{aligned}\langle \tilde{x}, \tilde{x}' + x'' \rangle &= \sum_{x,x' \in \mathscr{X}} \alpha_x(\beta_{x'} + \gamma_{x'})\mathrm{K}(x, x') \\ &= \sum_{x,x' \in \mathscr{X}} \alpha_x\beta_{x'}\mathrm{K}(x, x') + \sum_{x,x' \in \mathscr{X}} \alpha_x\gamma_{x'}\mathrm{K}(x, x') \\ &= \langle \tilde{x}, \tilde{x}' \rangle + \langle \tilde{x}, \tilde{x}'' \rangle .\end{aligned}$$

$\langle \cdot , \cdot \rangle$ est bien bilinéaire.

4) Soit $\alpha \in \mathrm{A}$ tel que $\tilde{x} = \sum_{x \in \mathscr{X}} \alpha_x \mathrm{K}_x$, et $\{x_1, \dots, x_n\}$ l'ensemble des $x \in \mathscr{X}$ tels que α_x est non-nul. En notant $\mathrm{G} = (\mathrm{K}(x_i, x_k))_{1 \leqslant i,k \leqslant n}$, et $u = (\alpha_{x_i})_{1 \leqslant i \leqslant n}$, on a :

$$\langle \tilde{x}, \tilde{x} \rangle = \sum_{x,x' \in \mathscr{X}} \alpha_x\alpha_{x'}\mathrm{K}(x, x') = \sum_{1 \leqslant n,k \leqslant n} \alpha_{x_n}\alpha_{x_k}\mathrm{K}(x_i, x_k) = u^\top \mathrm{G} u \geqslant 0.$$

car G est semi-définie positive par hypothèse. $\langle \cdot , \cdot \rangle$ est donc positive.

5) La quantité considérée est positive d'après la question précédente. En développant grâce à la bilinéarité, on a :

$$\begin{aligned}0 &\leqslant \langle \lambda\tilde{x} + \tilde{x}(x)\mathrm{K}_x, \lambda\tilde{x} + \tilde{x}(x)\mathrm{K}_x \rangle \\ &= \lambda^2 \langle \tilde{x}, \tilde{x} \rangle + 2\lambda\tilde{x}(x) \langle \tilde{x}, \mathrm{K}_x \rangle + \tilde{x}(x)^2 \langle \mathrm{K}_x, \mathrm{K}_x \rangle \\ &= \lambda^2 \langle \tilde{x}, \tilde{x} \rangle + 2\lambda\tilde{x}(x)^2 + \tilde{x}(x)^2\mathrm{K}(x, x).\end{aligned}$$

Cette quantité est un polynôme de degré 2 en λ, et est toujours positive. Le discriminant est donc négatif :

$$4\tilde{x}(x)^4 - 4 \langle \tilde{x}, \tilde{x} \rangle \tilde{x}(x)^2\mathrm{K}(x, x) \leqslant 0,$$

d'où le résultat.

6) Pour $\tilde{x} \in \tilde{\mathscr{X}}$, si $\langle \tilde{x}, \tilde{x} \rangle = 0$ l'inégalité de la question précédente entraîne que $\tilde{x}(x) = 0$ pour tout $x \in \mathscr{X}$, autrement dit $\tilde{x} = 0$.

7) D'après les questions précédentes, $\langle \cdot , \cdot \rangle$ est bien un produit scalaire. On considère l'application $\psi : x \mapsto \mathrm{K}_x$. K est alors le noyau associé à ψ car pour tous $x, x' \in \mathscr{X}$:

$$\langle \psi(x), \psi(x') \rangle = \langle \mathrm{K}_x, \mathrm{K}_{x'} \rangle = \mathrm{K}(x, x'),$$

où la dernière égalité découle de la définition de $\langle \cdot , \cdot \rangle$.