MVA: Kernel methods in machine learning (2022/2023)

Assignment 3

Solution by Matias Etcheverry March 8, 2023

1 Exercice 1: B_n-splines

For $x \in \mathbb{R}$, we define $I(x) = \mathbb{1}_{|x|\leq 1}$ and $B_n = I^{\star n}$. Let $k_n(x, y) = B_n(x - y)$ defined over $\mathbb{R} \times \mathbb{R}$.

- We show that k_n is a pd kernel.
  - We first show that k_n is symmetric, with a recurrence on n. k_1 is symmetric, as I is even. We then suppose that k_n is symmetric, i.e. B_n is even, and we have:

$$\begin{aligned}
B_{n+1}(x) &= B_n \star I(x) \\
&= \int_{\mathbb{R}} B_n(u) I(x-u) \mathrm{d}u \\
&= \int_{\mathbb{R}} B_n(u) I(-x+u) \mathrm{d}u \quad (I \text{ even}) \\
&= \int_{\mathbb{R}} B_n(-u) I(-x-u) \mathrm{d}u \\
&= \int_{\mathbb{R}} B_n(u) I(-x-u) \mathrm{d}u \quad (\text{by hypothesis, } B_n \text{ supposed even}) \\
&= B_{n+1}(-x)
\end{aligned}$$

    This shows that B_{n+1} is even. Thus k_{n+1} is symmetric for $n \in \mathbb{N}^*$. The recurrence is proved.
  - Next, we show that $a^T k_n a \geq 0$, for $a \in \mathbb{R}^n$. First, let's compute B_2:

$$\begin{aligned}
B_2(x) &= I \star I(x) \\
&= \int_{\mathbb{R}} \mathbb{1}_{|u|\leq 1}(u) \mathbb{1}_{|x-u|\leq 1}(u) \mathrm{d}u \\
&= \int_{-1}^{1} \mathbb{1}_{-1+x\leq u\leq 1+x}(u) \mathrm{d}u \\
&= \begin{cases} 0 & \text{if } x \geq 2 \text{ or } x \leq -2 \\ 2+x & \text{if } 0 \leq x \leq 2 \\ 2-x & \text{if } -2 \leq x \leq 0 \end{cases}
\end{aligned}$$

    B_2 is a triangular pulse, as shown in 1. The Fourier transform of B_2 is thus given by:

$$\hat{B}_2(w) = 4 \left(\frac{\sin w}{w} \right)^2$$

    This Fourier transform is positive, real-value and symmetric. Thus, by Böchner theorem, the shift-invariant kernel k_2 associated to the continuous function B_2 is positive-definite. In particular, we have, for $(a_1, \cdots, a_n) \in \mathbb{R}^n$, and $(x_1, \cdots, x_n) \in \mathbb{R}^n$:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j k_2(x_i, x_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j B_2(x_i - x_j) \geq 0$$

We then apply a recurrence on n to show that $\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j k_n(x_i, x_j) \geq 0$ for $n \geq 2$. Let's suppose that $\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j k_n(x_i, x_j) \geq 0$,

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j k_{n+1}(x_i, x_j) &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j B_{n+1}(x_i - x_j) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j (B_n \star I)(x_i - x_j) \\
&= \underbrace{\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j B_n\right)}_{\geq 0 \text{ by hypothesis}} \star I(x_i - x_j) \quad \text{(by linearity of convolution)} \\
&\geq 0
\end{aligned}$$

Thus, the recurrence is done.

We showed that $k_n(x, y) = B_n(x - y)$ is a kernel on $\mathcal{X} \times \mathcal{X}$ for $n \geq 2$. I couldn't prove it for $n = 1$.

- We are dealing with a translation invariant pd kernel.
  - First, let's show that B_n is integrable. As before, we do a recurrence on n. B_1 is integrable as I is bounded by 1 and has $[-1, 1]$ as support. Let's suppose B_n is integrable:

$$\begin{aligned}
\int_{\mathbb{R}} |B_{n+1}(x)| \mathrm{d}x &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} B_n(u) I(x - u) \mathrm{d}u \right| \mathrm{d}x \\
&\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |B_n(u)||I(x - u)| \mathrm{d}u \mathrm{d}x \\
&\leq \int_{\mathbb{R}} |B_n(u)| \int_{\mathbb{R}} |I(x - u)| \mathrm{d}x \mathrm{d}u \quad \text{(by Fubini-Tonelli Theorem)} \\
&\leq \int_{\mathbb{R}} |B_n(u)|) \mathrm{d}u \int_{\mathbb{R}} |I(x)| dx \\
&< \infty \quad \text{(by hypothesis)}
\end{aligned}$$

  This shows that B_{n+1} is integrable, i.e. the recurrence is shown.
  - Next, we show that the Fourier transform of B_n is integrable. Let's first recall the Fourier transform of B_1, which is a rectangle pulse:

$$\hat{B}_1(\omega) = 2\frac{\sin \omega}{\omega}$$

  Then, we have:

$$\begin{aligned}
\hat{B}_{n+1}(\omega) &= \int_{\mathbb{R}} B_n \star I(x) e^{-i\omega x} \mathrm{d}x \\
&= \int_{\mathbb{R}} B_n(x) e^{-i\omega x} \mathrm{d}x \times \int_{\mathbb{R}} I(x) e^{-i\omega x} \mathrm{d}x \quad \text{(convolution in time is multiplication in frequency)} \\
&= \hat{B}_n(\omega) \hat{B}_1(\omega)
\end{aligned}$$

  Thus, by recurrence, we end up with, for all $n \geq 2$:

$$\hat{B}_n(\omega) = \left(2\frac{\sin \omega}{\omega}\right)^n$$

  Finally, $\hat{B}_n$ is integrable for $n \geq 2$:

  - ∗ it is integrable when $\omega \to 0$, because $\left|\left(2\frac{\sin\omega}{\omega}\right)^n\right| \sim 2^n$, which is a constant.
  - ∗ it is integrable when $\omega \to +\infty$, because $\left|\left(2\frac{\sin\omega}{\omega}\right)^n\right| \leq \frac{2^{n+1}}{x^n}$ which is integrable for $n \geq 2$.

Finally, as B_n and $\hat{B}_n$ are integrable for $n \geq 2$, we can construct $\mathcal{H}_n$ the subset of $L_2(\mathbb{R})$ that consists of integrable and continuous functions f such that:

$$\begin{aligned} \|f\|^2_{\mathcal{H}_n} &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{|\hat{f}(\omega)|^2}{\hat{B}_n(\omega)}\mathrm{d}\omega < \infty \\ \text{i.e. } \|f\|^2_{\mathcal{H}_n} &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{|\hat{f}(\omega)|^2}{\left(2\frac{\sin\omega}{\omega}\right)^n}\mathrm{d}\omega < \infty \end{aligned}$$

endowed with the inner product:

$$\langle f \mid g\rangle_{\mathcal{H}_n} = \frac{1}{2\pi}\int_{\mathbb{R}} \frac{\hat{f}(\omega)\overline{\hat{g}(\omega)}}{\left(2\frac{\sin\omega}{\omega}\right)^n}\mathrm{d}\omega$$

is the RKHS of k_n as rk.

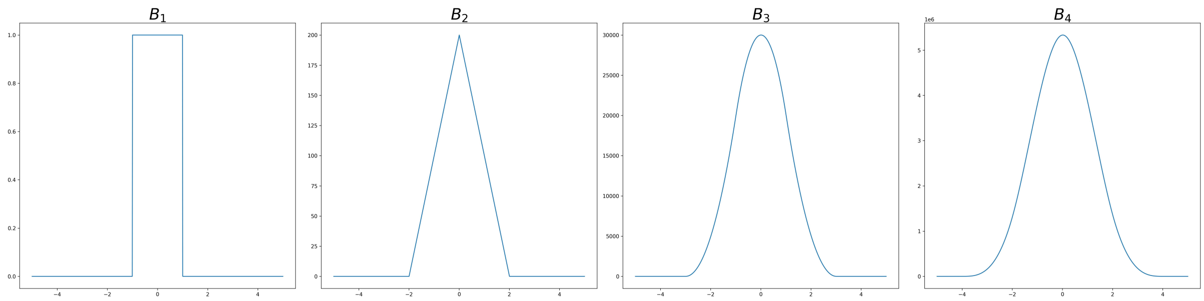

Figure 1: B_n-splines for $n = 1, 2, 3, 4$

2 Exercice 2: Sobolev spaces

1. Let $\mathcal{H} = \{f : [0;1] \mapsto \mathbb{R}, \text{absolutely continuous}, f' \in L^2([0;1]), f(0) = 0\}$ with $\langle f \mid g\rangle_{\mathcal{H}} = \int_0^1 f'(u)g'(u)\mathrm{d}u$.
   - (a) We show that $\mathcal{H}$ is a pre-Hilbert space:
     - i. it is a sub-vector space of continuous function $L^2([0,1])$
     - ii. $\langle f \mid g\rangle_{\mathcal{H}}$ is a bilinear form and $\langle f \mid f\rangle_{\mathcal{H}} \geq 0$
     - iii. We have, for all $x \in [0,1]$:

       $$f(x) = \int_0^x f'(u)\mathrm{d}u$$

       Thus, $\langle f \mid f\rangle_{\mathcal{H}} = 0 \Rightarrow f'(x) = 0$ for all $x \in [0,1]$, i.e. $f(x) = 0$
   - (b) We show that $\mathcal{H}$ is complete for the induced norm.
     - i. Let $(f_n)_{n\geq 0}$ be a Cauchy sequence in $\mathcal{H}$. It converges to f.
     - ii. Thus, $(f'_n)_{n\geq 0}$ is a Cauchy sequence in $L^2([0,1])$ (complete space) which converges to $g \in L^2([0,1])$.

iii. We have

$$\begin{aligned} f(x) &= \lim_{n\to\infty} f_n(x) \\ &= \lim_{n\to\infty} \int_0^x f_n'(u)\mathrm{d}u \\ &= \lim_{n\to\infty} \int_0^x g(u)\mathrm{d}u \end{aligned}$$

Thus, we have f absolutely continuous, $f' = g$ i.e. $f' \in L^2([0,1])$ and $f(0) = \lim_{n\to\infty} f_n(0) = 0$. Finally, $f \in \mathcal{H}$.

iv. Finally,

$$\begin{aligned} \|f_n - f\|_{\mathcal{H}} &= \|f_n' - f'\|_{L^2([0,1])} \\ &\xrightarrow[n\to\infty]{} 0 \end{aligned}$$

Finally, $\underline{\mathcal{H} \text{ is a Hilbert space.}}$

(c) We show that the reproducing kernel of $\mathcal{H}$ is K such that:

$$\boxed{\forall x \in [0,1], \forall t \in [0,1] \quad K_x(t) = K(x,t) = \min(x,t)}$$

i. K_x is 1-Lipshitz. Thus, K_x is absolutely continuous. $K_x' \in L^2([0,1])$. Finally $K_x(0) = 0$. Thus, $K_x \in \mathcal{H}$.

ii. For any $x \in [0,1]$ and $f \in \mathcal{H}$, we have:

$$\begin{aligned} \langle f \mid K_x \rangle_{\mathcal{H}} &= \int_0^1 f'(u) K_x'(u)\mathrm{d}u \\ &= \int_0^x f'(u)\mathrm{d}u \\ &= f(x) \end{aligned}$$

$\underline{\mathcal{H} \text{ is a RKHS with kernel } K.}$

2. Let $\mathcal{H} = \{f : [0;1] \mapsto \mathbb{R}, \text{absolutely continuous}, f' \in L^2([0;1]), f(0) = f(1) = 0\}$ with $\langle f \mid g \rangle_{\mathcal{H}} = \int_0^1 f'(u)g'(u)\mathrm{d}u$.

(a) We show that $\mathcal{H}$ is a pre-Hilbert space. We use the exact same proof as above.

(b) We show that $\mathcal{H}$ is complete for the induced norm. We use the same proof as above. We should not forget to mention that $f(1) = \lim_{n\to\infty} f_n(1) = 0$

(c) We are now looking for a reproducing kernel K of $\mathcal{H}$. To do so, we look for K_x as piecewise linear:

$$\forall x \in [0,1], \forall t \in [0,1] \quad K_x(t) = \begin{cases} \alpha t & \text{if } t \leq x \\ -\beta t + \gamma & \text{if } t \geq x \end{cases}$$

i. the continuity condition at $t = x$ and $K_x(1) = 0$ implies $\gamma = \beta$ and $(\alpha + \beta)x = \beta$

ii. K_x is $\max(\alpha, \beta)$-Lipshitz. Thus, K_x is absolutely continuous. $K_x' \in L^2([0,1])$. Finally $K_x(0) = 0$ and $K_x(1) = 0$ (with the above conditions). Thus, $K_x \in \mathcal{H}$.

iii. For any $x \in [0,1]$ and $f \in \mathcal{H}$, we have:

$$\begin{aligned} \langle f \mid K_x \rangle_{\mathcal{H}} &= \int_0^1 f'(u) K_x'(u)\mathrm{d}u \\ &= \alpha(f(x) - f(0)) - \beta(f(1) - f(x)) \\ &= f(x)(\alpha + \beta) \end{aligned}$$

The last equation implies $\alpha + \beta = 1$. Finally, all the combined conditions give us:

$$\forall x \in [0,1], \forall t \in [0,1] \quad K_x(t) = K(x,t) = \begin{cases} (1-x)t & \text{if } t \leq x \\ -xt + x & \text{if } t \geq x \end{cases}$$

i.e. $$\boxed{\forall x \in [0,1], \forall t \in [0,1] \quad K(x,t) = \min(x,t) - xt}$$

<u>$\mathcal{H}$ is a RKHS with kernel K.</u>

3. Let $\mathcal{H} = \{f : [0;1] \mapsto \mathbb{R}, \text{absolutely continuous}, f \in L^2([0;1]), f' \in L^2([0;1]), f(0) = f(1) = 0\}$ with $\langle f \mid g \rangle_{\mathcal{H}} = \int_0^1 f(u)g(u) + f'(u)g'(u)\mathrm{d}u$.
   (a) We show that $\mathcal{H}$ is a pre-Hilbert space. We use the exact same proof as above.
   (b) We show that $\mathcal{H}$ is complete for the induced norm. We use the same proof as above. We notice that

   $$\|f_n - f\|_{\mathcal{H}}^2 = \underbrace{\|f_n - f\|}_{\xrightarrow[n\to\infty]{}0}{}_{L^2([0,1])} + \underbrace{\|f_n' - f'\|}_{\xrightarrow[n\to\infty]{}0}{}_{L^2([0,1])}$$

   due to the completeness of $L^2([0,1])$.
   (c) We are now looking for a reproducing kernel K of $\mathcal{H}$. To do so, we integrate by parts the reproducing property. For any $x \in [0,1]$ and $f \in \mathcal{H}$, let P_x denote a primitive of K_x:

   $$\begin{aligned} \langle f \mid K_x \rangle_{\mathcal{H}} &= \int_0^1 f(u)K_x(u) + f'(u)K_x'(u)\mathrm{d}u \\ &= K_x(1)f(1) - K_x(0)f(0) + \int_0^1 f'(u)\left[K_x'(u) - P_x(u)\right]\mathrm{d}u \\ &= \int_0^1 f'(u)\left[K_x'(u) - P_x(u)\right]\mathrm{d}u \end{aligned}$$

   In order to have $\langle f \mid K_x \rangle_{\mathcal{H}} = f(x)$, we can choose K_x such that

   $$K_x''(u) - K_x(u) = \mathbb{1}_{u \leq x}$$

   This is a second order linear differential equation:
   - the general solution is: $K_x^0 : u \mapsto \lambda e^{-u} + \mu e^u$
   - I didn't find a particular solution for this equation. However, let's call a particular solution K_x^1

   Thus, we have the kernel K:

   $$\boxed{\forall x \in [0,1], \forall t \in [0,1] \quad K_x(t) = K(x,t) = \lambda e^{-u} + \mu e^u + K_x^1(u)}$$

   The absolute continuity and the border conditions on K should provide enough information to solve λ and μ.

<u>$\mathcal{H}$ is a RKHS with kernel K.</u>

4. Let $\mathcal{H} = \{f : [0;1] \mapsto \mathbb{R}, \text{absolutely continuous}, f' \in L^2([0;1]), f(0) = f'(0) = 0\}$ with $\langle f \mid g \rangle_{\mathcal{H}} = \int_0^1 f''(u)g''(u)\mathrm{d}u$.
   (a) We show that $\mathcal{H}$ is a pre-Hilbert space. We use the same proof as above. However, it is still important to mention, for $f \in \mathcal{H}$:

   $$\begin{aligned} \|f\|_{\mathcal{H}} = \int_0^1 f''(u)^2\mathrm{d}u = 0 &\Rightarrow f''(x) = 0 \quad \forall x \in [0,1] \\ &\Rightarrow f'(x) = \text{constant} = 0 \quad \forall x \in [0,1] \quad (\text{because } f'(0) = 0) \\ &\Rightarrow f(x) = \text{constant} = 0 \quad \forall x \in [0,1] \quad (\text{because } f(0) = 0) \end{aligned}$$

   i.e. $\|f\|_{\mathcal{H}} = 0$ implies $f = 0$.

(b) We show that $\mathcal{H}$ is complete for the induced norm. We use the same proof as above. We notice that

$$\|f_n - f\|_{\mathcal{H}}^2 = \underbrace{\|f_n'' - f''\|}_{\xrightarrow[n\to\infty]{} 0}{}_{L^2([0,1])}$$

due to the completeness of $L^2([0,1])$, as f_n'' is a Cauchy sequence in $L^2([0,1])$.

(c) We are now looking for a reproducing kernel K of $\mathcal{H}$. To do so, we integrate by parts the reproducing property. For any $x \in [0,1]$ and $f \in \mathcal{H}$, let P_x denote a primitive of K_x:

$$\langle f \mid K_x \rangle_{\mathcal{H}} = \int_0^1 f''(u) K_x''(u) \mathrm{d}u = f'(1) K_x''(1) - \int_0^1 f'(u) K_x'''(u) \mathrm{d}u \tag{1}$$

We look for K_x as piecewise polynomial:

$$\forall x \in [0,1], \forall t \in [0,1] \quad K_x(t) = \begin{cases} -t^3 + \beta t^2 + \gamma t + \delta & \text{if } t \leq x \\ bt^2 + ct + d & \text{if } t \geq x \end{cases}$$

Then, we have several conditions which solve the unknowns:

- Respecting the reproducing property in Eq. (1) leads to $K_x''(1) = 0$ i.e. $b = 0$.
- $K_x(0) = K_x'(0) = 0$ leads to $\gamma = \delta = 0$
- The continuity condition on K_x at $t = x$ leads to $-x^3 + \beta x^2 = cx + d$
- While it is not explicitly written, we need K_x to be twice derivable, i.e. we need the continuity condition on K_x' at $t = x$. This leads to $-3x^2 + 2\beta x = c$

Thus, we have the kernel K:

$$\boxed{\forall x \in [0,1], \forall t \in [0,1] \quad K(x,t) = \begin{cases} -t^3 + 2xt^2 & \text{if } t \leq x \\ x^2 t & \text{if } t \geq x \end{cases}}$$

$\mathcal{H}$ is a RKHS with kernel K.