MVA "Kernel methods in machine learning" Homework 1

Julien Mairal and Jean-Philippe Vert

Upload your answers (in PDF) to:
`https://goo.gl/XtQqdo`
before March 13, 2019, 1pm (Paris time).

Exercice 1. Kernels
Show that the following kernels are positive definite:

1. On $\mathcal{X} = \mathbb{R}$:
$$\forall x, y \in \mathbb{R}\,, \quad K(x, y) = \cos(x - y)\,.$$

2. On $\mathcal{X} = \{x \in \mathbb{R}^p : \|x\|_2 < 1\}$:
$$\forall x, y \in \mathcal{X}\,, \quad K(x, y) = 1/(1 - x^\top y)\,.$$

3. Given a probability space $(\Omega, \mathcal{A}, P)$, on $\mathcal{X} = \mathbb{R}$:
$$\forall A, B \in \mathcal{A}\,, \quad K(A, B) = P(A \cap B) - P(A)P(B)\,.$$

4. Let $\mathcal{X}$ be a set and $f, g : \mathcal{X} \to \mathbb{R}_+$ two non-negative functions:
$$\forall x, y \in \mathcal{X} \quad K_4(x, y) = \min(f(x)g(y), f(y)g(x))$$

5. Given a non-empty finite set E, on $\mathcal{X} = \mathcal{P}(E) = \{A : A \subset E\}$:
$$\forall A, B \subset E\,, \quad K(A, B) = \frac{|\,A \cap B\,|}{|\,A \cup B\,|}\,,$$
where $|\,F\,|$ denotes the cardinality of F, and with the convention $\frac{0}{0} = 0$.

Exercice 2. RKHS

1. Let K_1 and K_2 be two positive definite kernels on a set $\mathcal{X}$, and α, β two positive scalars. Show that $\alpha K_1 + \beta K_2$ is positive definite, and describe its RKHS.

2. Let $\mathcal{X}$ be a set and $\mathcal{F}$ be a Hilbert space. Let $\Psi : \mathcal{X} \to \mathcal{F}$, and $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be:

$$\forall x, x' \in \mathcal{X}, \quad K(x, x') = \langle \Psi(x), \Psi(x') \rangle_{\mathcal{H}} \, .$$

   Show that K is a positive definite kernel on $\mathcal{X}$, and describe its RKHS.

Exercice 3. Sobolev spaces

1. Let

$$\mathcal{H} = \left\{ f : [0,1] \to \mathbb{R}\,,\ \text{absolutely continuous}, f' \in L^2([0,1]), f(0) = 0 \right\} \, ,$$

   endowed with the bilinear form

$$\forall f, g \in \mathcal{H}\,, \quad \langle f, g \rangle_{\mathcal{H}} = \int_0^1 f'(u) g'(u) du \, .$$

   Show that $\mathcal{H}$ is an RKHS, and compute its reproducing kernel.

2. Same question when

$$\mathcal{H} = \left\{ f : [0,1] \to \mathbb{R}\,,\ \text{absolutely continuous}, f' \in L^2([0,1]), f(0) = f(1) = 0 \right\} \, ,$$

3. Same question, when $\mathcal{H}$ is endowed with the bilinear form:

$$\forall f, g \in \mathcal{H}\,, \quad \langle f, g \rangle_{\mathcal{H}} = \int_0^1 \left(f(u) g(u) + f'(u) g'(u) \right) du \, .$$

Exercice 4. Duality

Let $(x_1, y_1), \ldots, (x_n, y_n)$ a training set of examples where $x_i \in \mathcal{X}$, a space endowed with a positive definite kernel K, and $y_i \in \{-1, 1\}$, for $i = 1, \ldots, n$. $\mathcal{H}_K$ denotes the RKHS of the kernel K. We want to learn a function $f : \mathcal{X} \mapsto \mathbb{R}$ by solving the following optimization problem:

$$\min_{f \in \mathcal{H}_K} \frac{1}{n} \sum_{i=1}^{n} \ell_{y_i}(f(x_i)) \quad \text{such that} \quad \| f \|_{\mathcal{H}_K} \leq B\,, \tag{1}$$

where ℓ_y is a convex loss functions (for $y \in \{-1, 1\}$) and $B > 0$ is a parameter.

a. Show that there exists $\lambda \geq 0$ such that the solution to problem (1) can be found be solving the following problem:

$$\min_{\alpha \in \mathbb{R}^n} R(K\alpha) + \lambda \alpha^\top K \alpha\,, \tag{2}$$

where K is the $n \times n$ Gram matrix and $R : \mathbb{R}^n \mapsto \mathbb{R}$ should be explicited.

b. Compute the Fenchel-Legendre transform[1] R^* of R in terms of the Fenchel-Legendre transform ℓ_y^* of ℓ_y.

c. Adding the slack variable $u = K\alpha$, the problem (1) can be rewritten as a constrained optimization problem:

$$\min_{\alpha \in \mathbb{R}^n, u \in \mathbb{R}^n} R(u) + \lambda \alpha^\top K \alpha \quad \text{such that} \quad u = K\alpha\,. \tag{3}$$

Express the dual problem of (3) in terms of R^*, and explain how a solution to (3) can be found from a solution to the dual problem.

d. Explicit the dual problem for the logistic and squared hinge losses:

$$\ell_y(u) = \log(1 + e^{-yu})\,.$$

$$\ell_y(u) = \max(0, 1 - yu)^2\,.$$

[1]For any function $f : \mathbb{R}^N \mapsto \mathbb{R}$, the *Fenchel-Legendre transform* (or *convex conjugate*) of f is the function $f^* : \mathbb{R}^N \mapsto \mathbb{R}$ defined by

$$f^*(u) = \sup_{x \in \mathbb{R}^N} x^\top u - f(x)\,.$$