Kernel Methods

HW 1

① Consider 2 kernels K^1 and K^2.

$$K_p(x,y) = K^1(x,y) \cdot K^2(x,y) \rightarrow \text{Inner products}$$

$$K^1 = \langle \phi_x^{(1)}, \phi_y^{(1)} \rangle = \sum_{i=1}^{M} \phi_i^{(1)}(x)\, \phi_i^{(1)}(y)$$

$$K^2 = \langle \phi_x^{(2)}, \phi_y^{(2)} \rangle = \sum_{j=1}^{N} \phi_j^{2}(x)\, \phi_j^{2}(y)$$

$$\Rightarrow \quad K_p(x,y) = \sum_{i=1}^{M} \sum_{j=1}^{N} \phi_i^{(1)}(x)\, \phi_i^{(1)}(y)\, \phi_j^{(2)}(x)\, \phi_j^{(2)}(y)$$

Let $\phi_{ij}(z) = \phi_i(z)\phi_j(z) =$ scalar

$$\Rightarrow \quad K_p(x,y) = \sum_{i=1}^{M} \sum_{j=1}^{N} \phi_{ij}(x)\, \phi_{ij}(y)$$

$$= \langle \phi(x), \phi(y) \rangle_H.$$

$\therefore$ product of kernels is a kernel

(7) $x, x' \in \mathbb{R}^d$

$d \geq 1$

$m \geq 1 \quad m \in \mathbb{Z}$

$c \geq 0 \quad c \in \mathbb{R}$

$$k(x, x') = (\langle x, x' \rangle + c)^m$$

now consider

$$\langle x, x' \rangle + c = k(x, x')$$

Consider the following featuremap

$$\phi : x \to [x, \sqrt{c}]^T$$

$$\langle \phi(x), \phi(x_1) \rangle = \langle [x, \sqrt{c}], [x_1, \sqrt{c}] \rangle$$

$$= \langle x, x_1 \rangle + c$$

$\therefore$ the given function is a valid kernel.

Quoting the literature

If k is a kernel

$k + c$ is a kernel $\forall\, c \geq 0$.

Now consider

K_1 and K_2 be 2 kernel functions

then let $K = K_1 \circ K_2$

$$\Rightarrow K = \sum_{j=1}^{n}\sum_{i=1}^{n} \lambda_i (u_i u_i^T) \circ \mu_j (v_j v_j^T)$$

$$K = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \mu_j (u_i \odot v_j)(u_i \odot v_j)^T$$

$\odot$ being the element wise multiplication.

$$K = \sum_{k=1}^{n^2} \gamma_k w_k w_k^T \qquad w_k = u_i \odot v_j$$

$$w_k = u_{\lfloor k/n \rfloor} \odot v_{k \bmod n}$$

$$\gamma_k = \lambda_{\lfloor k/n \rfloor} \odot \mu_{k \bmod n}.$$

$\therefore \forall a \in \mathbb{R}^n$

$a^T K a \geq 0$ for such a function

$\Rightarrow$ product of kernels is a kernel.

$\therefore \left(\langle x, x_1 \rangle + c\right)^m$ is a valid kernel.

$\forall m \in N$

$c \geq 0$

③ Moore Aronszajn Theorem

Let $k: X \times X \to \mathbb{R}$ be a PD function then $\exists$ a unique RKHS whose reproducing kernel is k.

Let $\phi: X \to \mathbb{R}^X$

$$\phi = k(\cdot, x) \text{ for } x \in X$$

$$f(\cdot) = \sum_{i=1}^{n} \alpha_i k(\cdot, x_i) \quad \phi(x_i)$$

$$f(x) = \sum_{i=1}^{n} \alpha_i k(x, x_i)$$

Consider a vector space

$$f(x) = \sum_{i=1}^{n} \alpha_i k(x, x_i) \quad \begin{array}{l} n \in \mathbb{N} \\ \alpha_i \in \mathbb{R} \\ x_i \in X \end{array}$$

III^{ly} $$g(x) = \sum_{j=1}^{n'} \beta_j k(x, x_j) \quad \begin{array}{l} n' \in \mathbb{N} \\ \beta_j \in \mathbb{R} \\ x_j \in X \end{array}$$

the dot product results in

$$\langle f,g\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n'} \alpha_i \beta_j k(x_i, x_j)$$

This is bilinear.

$$\langle f,g\rangle = \sum_{j=1}^{n'} \beta_j f(x_i) = \sum_{i=1}^{n} \alpha_i g(x_j)$$

also

$$\langle f,g\rangle = \langle g,f\rangle$$

and $\langle f,f\rangle \geq 0$ from defn.

let $\alpha_i \alpha_j = P_K$

$$\therefore \langle f,g\rangle = \sum_{i,j=1}^{n,n'} P_K \langle a_i, a_j\rangle \qquad \underline{k = n \times n'}$$

By doing these constructions we end up at a pre-RKHS. Call it H'

Now define H to be the set of functions

$f \in \mathbb{R}^{\mathcal{X}}$ for which there exists an

H_0 Cauchy seq.n $\{f_n\}$ converging to f

pointwise. Then H is an RKHS

$$H' \subseteq H$$

① Inner product is well defined b/w,

f and g.

② $\langle f, f \rangle_H = 0$ iff $f = 0$

③ Evaluating functionals are continuous on H

④ H is complete

RK $\Leftrightarrow$ Kernel $\Leftrightarrow$ PD

$\therefore$ There are infinitely many feature space representations. Diff. feature maps give coeff's of canonical feature map in terms of simpler functions