Note of STAT 671

Statistical Learning I

Octorber 16 2019

Contents

Two Definitions of RKHS (why equvalent)

$X \neq \phi$, $\mathcal{H}$: Hilbert Space of function $\mathcal{X} \mapsto \mathbb{R}$

Example: $\mathcal{X} = \{x_1, ..x_n\}$; $\begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \subset \{\text{vector of } \mathbb{R}^n\}$

Definition 1:

$\mathcal{H}$ is a RKHS when there is a function $\mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$, $K(\cdot,\cdot)$ such that

- A: $t \mapsto k(t,x) \in \mathcal{H}$ for each x
- B: $\langle f, k(\cdot,x)\rangle_{\mathcal{H}} = f(x)$ for each $f \in \mathcal{H}, x \in \mathcal{X}$
  – Reproducing Property

Definition 2:

$\mathcal{H}$ is a RKHS when the evaluation functions

$$\begin{aligned} F_x : \ \mathcal{H} &\mapsto \mathbb{R} \\ f &\mapsto f(x) \quad \text{are continuous.} \end{aligned}$$

Definition 1 $\implies$ Definition 2

F_x is continuous. if

$$\begin{aligned} \|f-g\|_{\mathcal{H}} &< \delta \quad \text{(might depend on x)} \\ \implies \quad |f(x)-g(x)| &< \varepsilon \end{aligned}$$

F_x is *C-Lipschitz* continuous when

$$|f(x)-g(x)| \leq c\|f-g\|_{\mathcal{H}}, \quad c>0, \quad \text{for any } f,\ g \in \mathcal{H}$$

C-Lipschitz $\implies$ continuity.

$$|f(x)-g(x)| = |(f-g)(x)| = |\langle f-g, k(\cdot,x)\rangle_{\mathcal{H}}| \leq \|f-g\|_{\mathcal{H}} \ \underbrace{\langle k(\cdot,x), k(\cdot,x)\rangle^{\frac{1}{2}}}_{k^{\frac{1}{2}}(x,x)}$$

Definition 2 $\implies$ Definition 1

> *Riesz Representation Theorem*: In any Hilber Space of function $\mathcal{X} \mapsto \mathbb{R}$ for which F_x is continuous for each $x \in \mathcal{X}$, then there is an unique element of $\mathcal{H}$, notated g_x, for which $f(x) = \langle f, g_x\rangle_{\mathcal{H}}$ for each $f \in \mathcal{H}$, $\quad g_x(\cdot) = k(\cdot, x)$.

Examples

Example 0: $\mathcal{X} \in \mathbb{R}^d, \quad k(x,y) = x^T y$

Example 1: $\mathcal{X} = \{x_1, ..x_n\}$,

notate $\underset{(n,n)}{k}$; $[k]_{ij} = k(x_i, x_j)$. k is symmetric and positive semi-definite.

Assume that k is positive definite,

$$f : \ \mathcal{X} \mapsto \mathbb{R}, \quad \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \subset \ \mathbb{R}^n$$

$$k(\cdot, x_i) = \begin{bmatrix} k_{1i} \\ \vdots \\ k_{ni} \end{bmatrix} = k_i \,; \quad k = (k_1, ..k_n)$$

$$\begin{aligned} \mathcal{H} &= \{\alpha_1 k_1 + \cdots + \alpha_n k_n;\ \alpha_1, \cdots, \alpha_n \in \mathbb{R}\} \\ &= \text{Span}\{k_1, ..k_n\} = \mathbb{R}^n \quad \text{is a vector space.} \end{aligned}$$

$$\begin{aligned} \langle f, g \rangle_{\mathcal{H}} &= f^T k^{-1} g \\ \langle f, k(\cdot, x_i) \rangle &= \langle f, k e_i \rangle, \quad e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i \\ &= f^T \underbrace{k^{-1} k}_{I} e_i \\ &= f^T e_i \\ &= \begin{bmatrix} f(x_1) & \ldots & f(x_n) \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i \\ &= f(x_i) \end{aligned}$$

Example 2: $\mathcal{X} \in \mathbb{R}^n, \quad k(x, y) = (x^T y)^2$

$$\mathcal{H} = \{f :\ f(x) = x^T S_x; \quad \underset{(n,n)}{S} \text{ is a symmetric Matrix}\}$$

verify this is a Hilbert Space.

$$\langle f_{S_1}, f_{S_2} \rangle_{\mathcal{H}} = \langle S_1, S_2 \rangle_{\mathcal{F}} = \sum_{i,j=1}^{n} [S_1]_{ij} [S_2]_{ij}$$

$$\langle f_{S_1}, k(\cdot, x_i) \rangle = f_{S_1}(x) \quad \text{check it}$$

$$k(y, x) = (y^T x)(y^T x) = y^T \cdot \underbrace{x x^T}_{\substack{(n,n) \\ \text{symmetric} \\ \text{matrix}}} \cdot y$$