- Vector space

A vector space is a set V with addition $+ : V \times V \to V$, multiplication $\times : \mathbb{R} \times V \to V$

with property $\begin{cases} \text{commutative} & a+b=b+a \\ \text{associative} & (a+b)+c = a+(b+c) \\ \text{distributive} & k\cdot(a+b) = ka+kb \\ \text{negativity} & v \in V \quad -v \in V \\ \vdots & \end{cases}$

- Norm & Normed space

A norm on a vector space is a function $\|\cdot\| : V \to [0,+\infty)$

with property $\begin{cases} \|v\|=0 \Leftrightarrow v=0 \\ \|\alpha v\| = \alpha \|v\| \\ \|x+y\| \leq \|x\|+\|y\| \end{cases}$

A vector space with a norm on it is a normed space

- Banach Space

A normed space V is a Banach space

if V is complete wrt the metric induced by the norm $d(x,y) = \|x-y\|$

- Pre-Hilbert Space

A pre-Hilbert space $\mathcal{H}$ is a vector space over $\mathbb{C}$ (or $\mathbb{R}$)

with a Hermitian inner product

$$\langle \cdot,\cdot \rangle_{\mathcal{H}} : \mathcal{H} \times \mathcal{H} \to \mathbb{C} \quad \text{s.t.} \begin{cases} \forall \lambda_1, \lambda_2 \in \mathbb{C} & \langle \lambda_1 v_1 + \lambda_2 v_2, w \rangle_{\mathcal{H}} = \lambda_1 \langle v_1, w_1 \rangle_{\mathcal{H}} + \lambda_2 \langle v_2, w_2 \rangle_{\mathcal{H}} \\ \forall v, w \in \mathcal{H} & \langle v, w \rangle_{\mathcal{H}} = \overline{\langle w, v \rangle_{\mathcal{H}}} \\ \forall v \in \mathcal{H} & \langle v, v \rangle_{\mathcal{H}} \geq 0 \text{ real}, \quad \langle v,v \rangle_{\mathcal{H}} \Leftrightarrow v=0 \end{cases}$$

can prove Cauchy-Schwartz from these 3 condition

then $\|v\|_{\mathcal{H}} = \langle v,v \rangle_{\mathcal{H}}^{1/2}$ is a norm on $\mathcal{H}$

- Hilbert space

A Hilbert space is a pre-Hilbert space

that is complete wrt norm $\|x\|_{\mathcal{H}} = \langle x,x \rangle_{\mathcal{H}}^{1/2}$

Eg. $\mathbb{C}^n$: $\langle u,v \rangle = \sum_i u_i \overline{v_i}$

l^2: $\{ \{x_n\} : \sum_{i=1}^{\infty} |x_i|^2 < +\infty \}$ $\quad \langle x,y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}$

L^2: $\{ f : \int |f|^2 < +\infty \}$ $\quad \langle f,g \rangle = \int f \cdot \overline{g}$

- Construct Reproducing Kernel Hilbert Space

⓪: Start with a PSD function $k: \mathcal{X}\times\mathcal{X} \mapsto \mathbb{R}$ s.t. $\begin{cases} \text{if } K \in \mathbb{R}^{m\times m} \text{ and } K_{ij} = k(x_i, x_j) \\ \text{then } K \in S_+^n \end{cases}$

①: define the feature map $\phi: \mathcal{X} \to (\mathcal{X} \to \mathbb{R})$

$\phi(x) = k(x, \cdot)$

②: Construct vectors in vector space

$f = \sum_{i=1}^{m} \alpha_i \phi(x_i) = \sum_{i=1}^{m} \alpha_i k(x_i, \cdot)$ any linear combination of $k(x_i, \cdot)$

the vector space is $\text{span}(\{\phi(x) : x \in \mathcal{X}\}) = \{ f = \sum_{i=1}^{m} \alpha_i k(x_i, \cdot) : m \in \mathbb{N}\ \alpha_i \in \mathbb{R}\ x \in \mathcal{X} \}$

③: To make it a pre-Hilbert space, define the inner product

$$\left.\begin{aligned} f &= \sum_i \alpha_i k(x_i, \cdot) \\ g &= \sum_j \beta_j k(x_j, \cdot) \end{aligned}\right] \langle f, g\rangle_{\mathcal{H}} = \sum_i \sum_j \alpha_i \beta_j k(x_i, x_j)$$

Reproducing property: $f = \sum_i \alpha_i k(x_i, \cdot)$

$\langle f, k(x, \cdot)\rangle_{\mathcal{H}} = \sum_i \alpha_i k(x_i, x) = f(x)$

$\langle k(x, \cdot), k(\cdot, y)\rangle_{\mathcal{H}} = k(x, y)$

can show that $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ satisfies the properties required in pre-Hilbert space

(1): Symmetry: By symmetry of $k(\cdot, \cdot)$

(2): linearity: $\langle f_1 + f_2, g\rangle_{\mathcal{H}} = \langle \sum_{i=1}^{m} \alpha_i k(x_i, \cdot) + \sum_{i=m+1}^{n} \alpha_i k(x_i, \cdot), \sum_j \beta_j k(x_j, \cdot)\rangle_{\mathcal{H}}$

$$\begin{aligned} &= \langle \sum_{i=1}^{n} \alpha_i k(x_i, \cdot), \sum_j \beta_j k(x_j, \cdot)\rangle_{\mathcal{H}} \\ &= \sum_{i=1}^{n} \sum_j \alpha_i \beta_j k(x_i, x_j) \\ &= \sum_{i=1}^{m} \sum_j \alpha_i \beta_j k(x_i, x_j) + \sum_{i=m+1}^{n} \alpha_i \beta_j k(x_i, x_j) \\ &= \langle f_1, g\rangle_{\mathcal{H}} + \langle f_2, g\rangle_{\mathcal{H}} \end{aligned}$$

(3). PSD.

$f = \sum_i \alpha_i k(x_i, \cdot)$ $\quad \langle f, f\rangle_{\mathcal{H}} = \sum_i \sum_j \alpha_i \alpha_j k(x_i, x_j) = \alpha^T K \alpha \geq 0$

$$\begin{aligned} f(x)^2 &= \langle k(x, \cdot), f\rangle_{\mathcal{H}}^2 \\ &\leq \langle k(x, \cdot), k(\cdot, x)\rangle_{\mathcal{H}}^2 \cdot \langle f, f\rangle_{\mathcal{H}}^2 \\ &= k(x, x)^2 \cdot \langle f, f\rangle_{\mathcal{H}}^2 \end{aligned}$$

$$\left.\begin{aligned} &\text{if } \langle f, f\rangle_{\mathcal{H}} = 0, \text{ then } f(x) = 0\ \forall x, \text{ thus } f = 0 \\ &\text{if } f = 0, \text{ then } \langle f, f\rangle_{\mathcal{H}} = 0 \text{ is obvious} \end{aligned}\right] \langle f, f\rangle_{\mathcal{H}} = 0 \Leftrightarrow f = 0$$

④ For $\mathcal{H}$ to be a Hilbert space, we need it to be complete

$$\mathcal{H} = \overline{\left\{ f = \sum_{i=1}^{m} \alpha_i k(x_i, \cdot) : \alpha_i \in \mathbb{R},\ x_i \in \mathcal{X},\ m \in \mathbb{N} \right\}}$$

Definition : RKHS

$\mathcal{H}$ is a RKHS if

$$\exists k : \mathcal{X} \times \mathcal{X} \to \mathbb{R} \quad \text{s.t.} \quad \begin{cases} k \text{ has the reproducing property } f(x) = \langle f, k(\cdot, x) \rangle_{\mathcal{H}} \\ k \text{ spans } \mathcal{H} : \quad \mathcal{H} = \overline{\text{span}\{ k(x, \cdot) : x \in \mathcal{X} \}} \end{cases}$$

Theorem :

$\forall$ psd function $k(\cdot,\cdot)$, $\exists$ a unique RKHS